Monitoring electromagnetic tracking error using redundant sensors

Vinyas Harish, Eden Bibic, Andras Lasso, Matthew S. Holden, Thomas Vaughan, Zachary Baum, Tamas Ungi, Gabor Fichtinger

Laboratory for Percutaneous Surgery, Queen's University, Kingston, ON, Canada

ABSTRACT

PURPOSE: The intraoperative measurement of tracking error is crucial to ensure the reliability of electromagnetically navigated procedures. For intraoperative use, methods need to be quick to set up, easy to interpret, and not interfere with the ongoing procedure. Our goal was to evaluate the feasibility of using redundant electromagnetic sensors to alert users to tracking error in a navigated intervention setup.

METHODS: Electromagnetic sensors were fixed to a rigid frame around a region of interest and on surgical tools. A software module was designed to detect tracking error by comparing real-time measurements of the differences between inter-sensor distances and angles to baseline measurements. Once these measurements were collected, a linear support vector machine-based classifier was used to predict tracking errors from redundant sensor readings.

RESULTS: Measuring the deviation in the reported inter-sensor distance and angle between the needle and cautery served as a valid indicator for electromagnetic tracking error. The highest classification accuracy, 86%, was achieved based on readings from the cautery when the two sensors on the cautery were close together. The specificity of this classifier was 93% and the sensitivity was 82%.

CONCLUSION: Placing redundant electromagnetic sensors in a workspace seems to be feasible for the intraoperative detection of electromagnetic tracking error in controlled environments. Further testing should be performed to optimize the measurement error threshold used for classification in the support vector machine, and improve the sensitivity of our method before application in real procedures.

Keywords: Quality Assurance System**,** Electromagnetic Tracking Error, Image-Guided Intervention, Image-Guided Therapy, 3D Slicer, SlicerIGT, PLUS Toolkit

1. INTRODUCTION

Electromagnetic (EM) tracking offers continuous localization of medical instruments and patient anatomy without line-of-sight requirements. However, magnetic field distortions and EM interference can lead to decreased tracking accuracy. Therefore, it is important to monitor the reliability of EM tracking systems in clinical environments. In their 2014 literature review, Franz *et al.* have outlined a recommended protocol for validating EM tracked computer-navigated intervention systems [1]. A good protocol is expected to provide evaluation and analysis of precision, accuracy, dynamic effects, and magnetic field distortions. Solutions for monitoring intraoperative EM tracking error must be quick to deploy, easy to use, and cause minimal interference to the ongoing procedure. Given these requirements, using large phantoms or robots to evaluate tracking error is not suitable for intraoperative use [2, 3].

Currently, work is being done in the Laboratory for Percutaneous Surgery to develop a multistage solution for monitoring EM tracking error for computer-navigated procedures. This solution, termed the Tracking Error Inspector has proven to be a useful tool for conducting workspace analysis in both laboratory and operating room environments as part of a preoperative quality assurance workflow [4]. Optical tracking is unaffected by field distortions, and has been used to collect ground truth measurements in previous work [5]. Thus, in our application, tracking error readings are computed by concurrently tracking a tool both optically and electromagnetically and comparing the measurements between the tracking modalities. However, setting up the optical tracker requires additional space, setup time, introduces line-of-sight constraints, and in general may be cumbersome in a busy operating room. Moreover, optical trackers are generally more expensive than adding an additional EM tracker.

In order to overcome these challenges, our goal was to develop a proof-of-concept for a system that could detect EM tracking error using redundant EM sensors. Jain *et al*. have described an EM tracking error characterization method using redundant EM sensors [6]. In their system, multiple sensors are affixed to different locations on a tool. The ground truth tool geometry, or distances and angles between sensors without tracking error, can be compared to the measured tool geometry to provide a gauge of tracking error in real-time.

The addition of unobtrusive, intraoperative EM tracking error monitoring represents a natural extension of the Tracking Error Inspector system. This extension would be achieved using redundant EM sensors. We propose using redundant EM position sensors in the EM field to monitor tracking accuracy throughout the procedure and indicate when errors are suspected to have exceeded a safe limit.

2. METHODS

2.1 Hardware Design

We designed an adjustable, rectangular frame using the SolidWorks software (2015 SP03, Dassault Systèmes, Vélizy, France) to hold four Ascension Model 800 EM sensors (Northern Digital Inc., ON, Canada) (Figure 1). The frame is rigid, and has the sensors in a fixed geometric arrangement. Thus, any measured deviations in the distances and angles between sensors in the frame can be attributed to EM tracking error. The frame is designed to surround a specific region of interest within the workspace. The design files are available online, free for use and modification without restrictions in the PLUS toolkit 3D model catalog (www.plustoolkit.org).

Figure 1. Mechanical design of the frame to hold Ascension EM sensors.

2.2 Software Design and Computational Details

We developed a module for 3D Slicer (www.slicer.org) that monitors the deviations in the pairwise distances and angles between an arbitrary set of EM position and orientation sensors (Figure 2). This platform was also used for developing the other aspects of the Tracking Error Inspector. Building upon an open-source platform has the additional advantages of allowing for reproducibility and enables other researchers to adapt this system for their needs. The PLUS toolkit's PlusServer application[7] was used to connect hardware devices and stream live measurement data to 3D Slicer using the OpenIGTLink communication protocol [8].

Several sensor configurations were specified: mounting sensors on a rigid frame around the region of interest, and on tracked tools. The baseline measurements of the distances and angles between sensors were computed with samples collected over 30 seconds. This step was completed in a clean field, far from potential error sources.

From the PLUS toolkit, we can obtain transformations from a sensor's coordinate frame to the coordinate frame of the EM tracker. For two arbitrary sensors A and B, these transformations are $T_{A\to Tracker}$ from the coordinate frame of sensor A *{A}* to the EM tracker *{Tracker}* and $T_{B\to Tracker}$ from the coordinate frame of sensor B *{Sensor B}* to the EM tracker *{Tracker}*.

The pairwise distance between sensors A and B was calculated as a Euclidean distance shown in equation (1), where a and b represent the position vectors of the sensors obtained from their respective transformations.

$$d\,(A,B) = \sqrt{(b_x - a_x)^2 + \left(b_y - a_y\right)^2 + (b_z - a_z)^2} \quad (1)$$

After the sampling period, the average of the Euclidean distance between pairs of redundant sensors, d_{avg}, was calculated.

Computing the average angle between two sensors required several steps, as shown in equations (2) through (4). First, the transformation between two sensors, A and B, was calculated as in equation (2).

$$T_{A\to B} = (T_{B\to Tracker})^{-1} T_{A\to Tracker} \quad (2)$$

Next, the rotation matrix from the $T_{A\to B}$ transformation ($R_{A\to B}$) was isolated and converted into a quaternion ($q_{A\to B}$). To determine the average rotation difference, w, between sensors, the average rotation between the two sensors was first computed. This was done as a cumulative moving average taken over all timestamps of recording, using quaternion spherical-linear interpolation (SLERP) to determine the average between the previous average rotation, q_m, and the new rotation sample, q_n:

$$q_i = [w_i, x_i, y_i, z_i] = \frac{q_m sin((1-t)\theta) + q_n sin(t\theta)}{sin\theta} \quad (3)$$

Where q_i represents the interpolated quaternion, q_m and q_n are the input quaternions, and t is a scalar value between 0.0 (at q_m) and 1.0 (at q_n). At each iteration, we take t to be the inverse of one greater than the number of samples used to calculate q_m. The result of SLERP is an interpolated quaternion. Then, we computed the average angle, a_{avg}, using the w_i component of the interpolated quaternion as shown in equation (5).

$$a_{avg} = 2(\arccos(w_i)) \quad (4)$$

As a quality assurance measure for the baseline measurements, we computed the standard deviation of the pairwise distances and angle differences between the sensors. If the standard deviation was high, it suggested that the baseline measurement should be retaken.

Once the baseline distances and angle differences were calculated, real-time tracking error was monitored by comparing measurements with the baseline measurements. First, the current position distance between the two sensors is computed as in equation (1). To compute the position error, E_P, this current distance is then subtracted from the baseline distance measurement between the two sensors. Finally, the absolute value of this difference is taken, as shown in equation (5).

$$E_P = \left| d_{avg} - d\ (A, B)_{Current} \right| \quad (5)$$

Computing the rotation error between sensors involved several steps. First, the transformation between current positions of the two sensors, A and B, was calculated as in equation (2). Then, we computed rotation between the baseline transformation, $TB_{A\to B}$, and current transformation, $TC_{A\to B}$, between the two sensors. Next, the rotation matrix was isolated from the prior calculated matrix product, M, and orthogonalized to avoid rounding errors as shown in equation (7). Finally, the rotational error was computed as an angle using the axis-angle representation shown in equation (8).

$$M = TB_{A\to B}(TC_{A\to B})^{-1} \quad (6)$$

$$R_M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \to \hat{R}_M = \begin{bmatrix} \hat{m}_{11} & \hat{m}_{12} & \hat{m}_{13} \\ \hat{m}_{21} & \hat{m}_{22} & \hat{m}_{23} \\ \hat{m}_{31} & \hat{m}_{32} & \hat{m}_{33} \end{bmatrix} \quad (7)$$

$$E_R = arccos\left((\hat{m}_{11} + \hat{m}_{22} + \hat{m}_{33} - 1)/2\right) \quad (8)$$

After the sampling period, we computed the average and standard deviation of the distance error, E_P, and the rotation error, E_R.

Finally, as each Ascension sensor contains multiple sensor coils, the tracker provides a 'quality' value for each sensor by evaluating the consistency between measurements provided by these coils. Quality values were recorded along with position and orientation data.

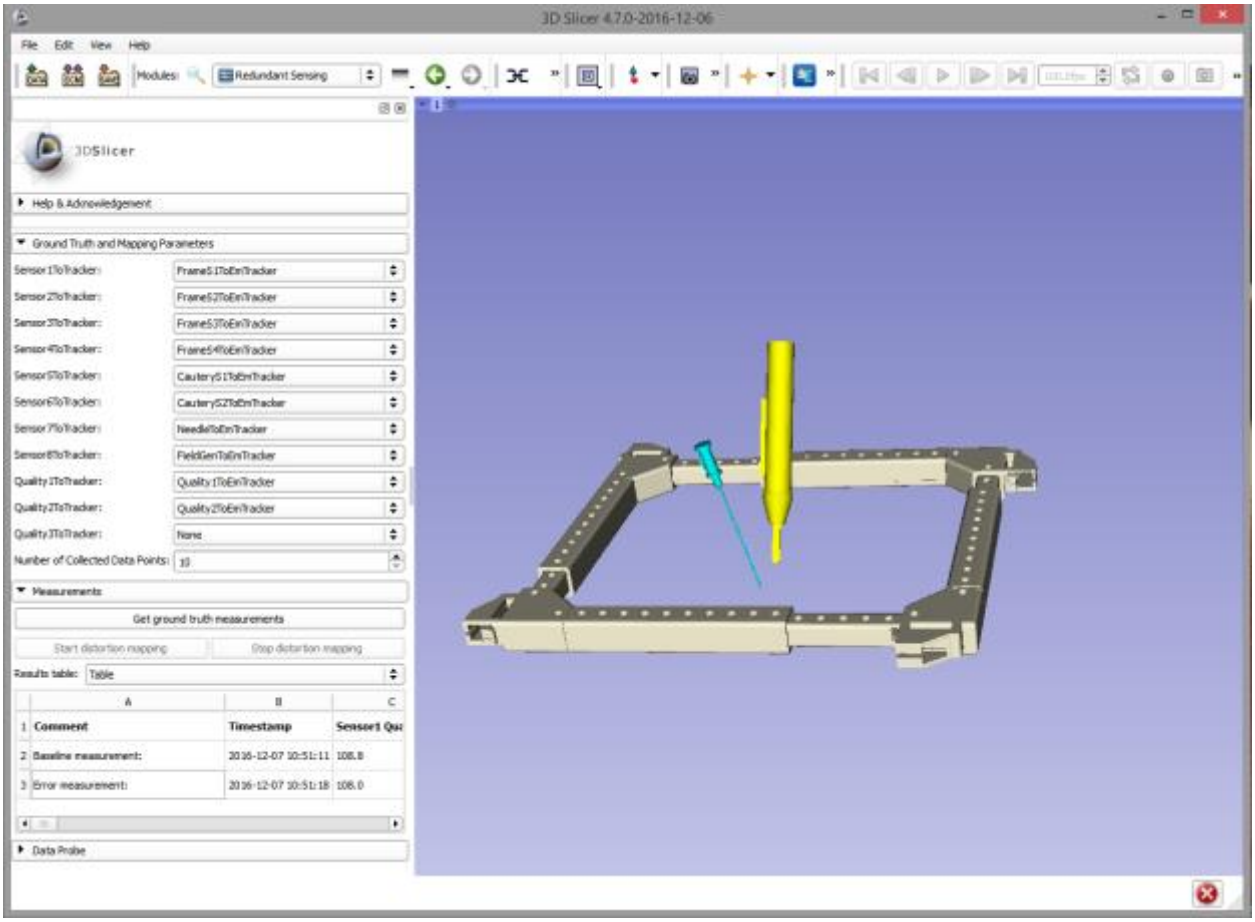

Figure 2. 3D Slicer user interface displaying our designed software module in a setup mimicking EM-navigated breast-conserving surgery, further explained in Section 2.3.

2.3 Experimental Design

We tested our system in a laboratory setup mimicking EM-navigated breast-conserving surgery, where EM sensors are attached to a tissue-locking needle and electrosurgical cautery [9]. The cautery and tumor geometry are tracked relative to the tissue-locking needle. It is important to assess the tracking accuracy of the needle relative to the cautery, as tracking errors between them can lead to an incomplete resection. The tracking error between these two tools is defined as the measurement error.

To examine how various aspects of the navigated breast-conserving surgery workflow affect tracking accuracy, five additional sensors were introduced to the setup: four sensors were affixed to a rigid frame around the region of interest, and an extra sensor was attached to the cautery (Figure 3). The two sensors on the cautery were tested in two separate configurations, one where the sensors were close together (1 cm apart) and one where they were further apart (6 cm apart). Although it was predicted that the 6 cm configuration would be more sensitive, we wanted to determine if the difference was negligible. The 1 cm configuration was more compact and further away from the cautery's tip, which may be more convenient to use in practice.

After baseline measurements were obtained with the field generator at different distances and positions, several tests were done to determine if increases in the measurement error were induced by various objects in the workspace: a Microsoft

Surface tablet was brought to different distances from the workspace, and a surgical retractor was placed in eight different positions around the region of interest (Figure 2 Right). A Microsoft Surface tablet was used to induce error in this experiment because it is used in Ungi *et al.*'s system as a part of the clinical workflow. The retractor was either perpendicular to a side of the frame or on one of the corners of the frame at a 45-degree angle. Real-time error monitoring took place over periods of 30 seconds, and the averages and standard deviations of the positional tracking error, rotational tracking error, and quality values were recorded. It was also explored if the measurement error corresponded to the errors reported from the redundant sensors.

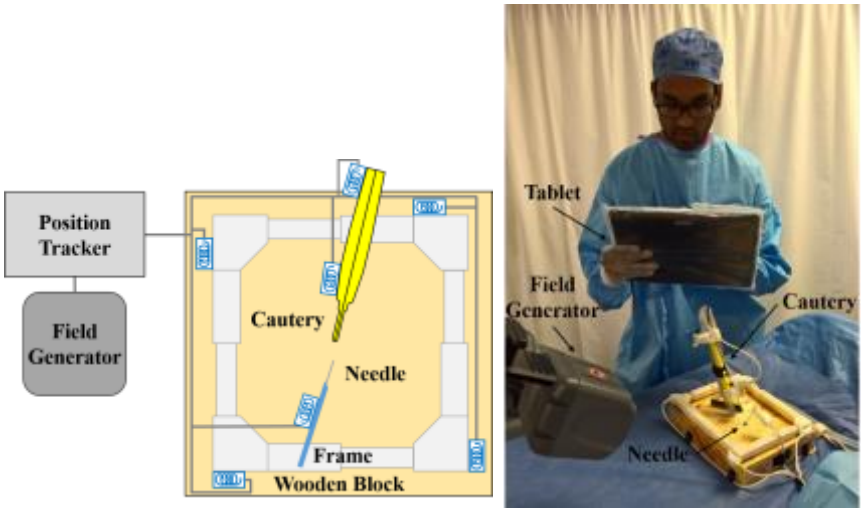

Figure 3. Our experimental setup using redundant EM sensors to characterize tracking accuracy of in a mock navigated lumpectomy (left). An experiment involving the use of a tablet computer to induce tracking error (right).

2.4 Designing a Predictor

In our experiment, we could directly ascertain the measurement error of the navigated lumpectomy system. Any movements of the needle and cautery could be attributed to tracking error because they were both firmly fixed in the wooden block. However, in a real setup, this would not be possible as the needle and cautery are constantly in motion relative to each other. Thus, we can only use measurements from the redundant sensors that have been added as an indicator for potential EM tracking error in the surgical workspace. In other words, we must use the readings from the redundant sensors to predict when there is likely to be measurement error in the system. Due to the sheer volume of data at hand, we explored if machine learning could be used to determine a threshold for the errors between a subset of the redundant sensors that is also associated with measurement error.

This problem of designing a predictor is a supervised machine learning problem, as the goal is to infer a function representing an error threshold using labelled training data. Due to the complexities of EM tracking and field distortion, predicting the actual amount of measurement error expected between the needle and cautery would be exceedingly difficult. Thus, a simplification that can be made is that we are only interested in predicting if there is likely to be measurement error in the system or not. This simplification reduces our problem to a binary classification problem, where the goal of the chosen machine learning algorithm will be to determine the best possible means of differentiating between two possible class labels using input attributes. In his "Review of Classification Techniques", Kotsiantis explores the strengths and weaknesses to several supervised machine learning algorithms, including support vector machines (SVMs)

[10]. SVMs are found to perform well when working with multidimensional data and continuous attributes. All our attributes are measurements from the redundant sensors, and are thus continuous values. Because of the multitude of possible combinations of redundant sensor measurements, our data is multidimensional. Moreover, SVMs perform well when non-linear relationships exist between the input and output data, which is also expected due to the complexities of EM tracking. Based on these reasons, it was believed that SVMs would be a good option for developing a predictor for measurement errors in our experimental setup.

We developed two linear support vector machine-based classifiers to predict measurement error using RapidMiner Studio Community Edition (RapidMiner Inc., Boston, MA, USA). One classifier was trained using information gathered from the sensors on the frame. Attributes used to train the frame-based classifier included all averages and standard deviations of the pairwise distances and angles between sensors, as well as the averages and standard deviations of the quality values from the frame. The other classifier was trained with information gathered from the sensors on the cautery. Attributes used to train the cautery-based classifier included the average and standard deviation of the difference in distance between the two sensors on the cautery, and the averages and standard deviations of the quality values of these sensors. A true measurement error was defined when either the position or the orientation error between the needle and cautery exceeded 1.4 mm or 0.5 degrees, which is the manufacturer's reported RMS accuracy of the tracker.

3. RESULTS AND DISCUSSION

Our experiments showed that the tablet computer had minimal influence on the tracking accuracy, as the largest measurement errors observed were under 0.43 mm and 0.46 degrees (Table 1).

The metal retractor induced considerable measurement error, up to 17.34 mm and 7.71 degrees in certain cases when placed around the needle. As expected, the error was largest when the retractor was placed in between the field generator and sensors. The error also increased when the retractor was positioned around the needle, as opposed to the cautery or other areas. The sensor on the needle was much closer to the retractor in certain positions compared to the sensor on the cautery, leading us to believe this caused a greater impact on tracking accuracy of the needle. Because the retractor had a large influence on the tracking accuracy depending on its positioning, eliminating the use of ferromagnetic retractors altogether during EM navigated procedures could be considered.

Table 1. A summary of the true positional and rotational measurement errors (reported as the average plus/minus the standard deviation). The 'retractor' setup is the average of the positional and rotational errors caused by the various positions of the retractor in the workspace. The 'tablet' setup is the average of the positional and rotational errors caused by the tablet when held by an assistant at distances of 40, 45, and 50 cm away from the needle and cautery.

Setup	Field generator 20 cm away	Field generator 25 cm away	Field generator 30 cm away
Clean field	0.05 ± 0.03 mm 0.07 ± 0.01 deg	0.05 ± 0.03 mm 0.02 ± 0.01 deg	0.05 ±0.02 mm 0.01 ± 0.00 deg
Retractor	2.90 ± 0.03 mm 2.11 ± 0.00 deg	2.98 ± 0.04 mm 2.02 ± 0.01 deg	3.67 ± 0.10 mm 2.40 ± 0.01 deg
Tablet	0.08 ± 0.05 mm 0.45 ± 0.01 deg	0.07 ± 0.06 mm 0.29 ± 0.02 deg	0.40 ± 0.06 mm 0.34 ± 0.01 deg

The positional and rotational error between the two cautery sensors seemed to be a promising indicator of true measurement error. There appeared to be a positive correlation between instances of error from the two sensors on the cautery (the predicted error) and the true measurement error (Figure 4). We found no such patterns in the readings from the four sensors on the frame. The measurement error was correlated to readings from different sensors on the frame dependent on the

location of the error source. For example, placing a retractor near sensors one and two on the frame would cause an increased error in measurements related to sensor one or two, while not affecting the other sensors. Thus, determining a metric that encompassed multiple readings from the frame in a useful manner proved to be challenging. For example, Figure 5 illustrates using the maximum positional or rotational error from all pairwise combinations of sensors as a predictor for error. Although the predictor for error could be quite high, for example when the retractors were placed directly on a sensor, this did not necessarily mean there would be a measurement error in the system or vice-versa. Due to this complication, all inter-sensor positional and rotational errors and the quality values from each sensor on the frame were used to train our classifier. There were no discernable patterns in the reported quality values from the sensors that warranted developing a classifier solely based on those values.

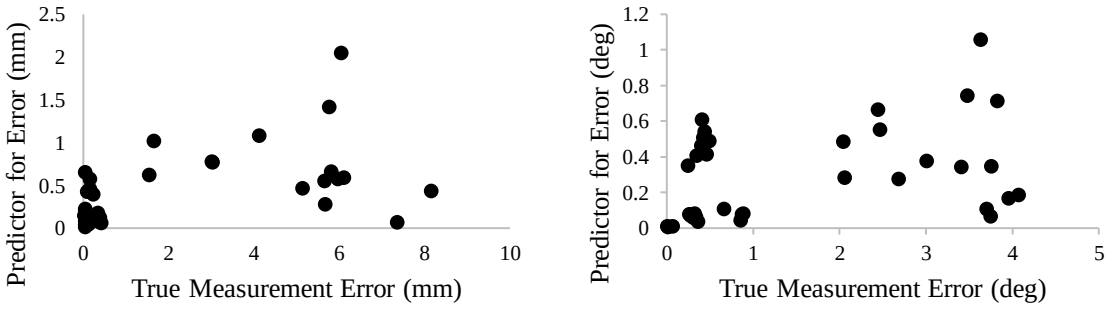

Figure 4. Scatterplots displaying positional error (left) and rotational error (right) between the predictors for error and true measurement errors. The predictors for error are those reported between the two redundant sensors on the cautery.

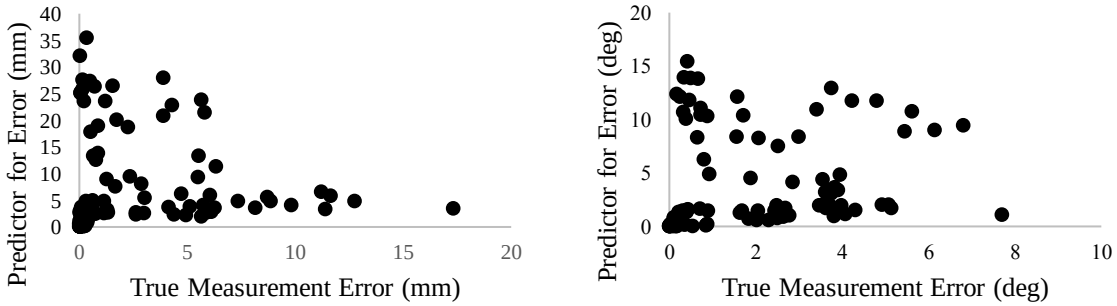

Figure 5. Scatterplots displaying positional error (left) and rotational error (right) between the predictors for error and true measurement errors. The predictors for error are the maximum of the pairwise errors between the four sensors of the frame.

The "leave-one-out" cross-validation technique was used to verify the performance of the classifiers developed in RapidMiner. The accuracy of the frame-based classifier (63%) was lower than that of the cautery-based classifier (86%) (Table 2). The specificities of both classifiers were over 84%; however, the main issue with both the classifiers was their sensitivities, or their ability to avoid false negatives. False negatives are especially important to minimize in this application, otherwise the classifier will fail to warn surgeons of true measurement errors that compromise surgical outcomes. The sensitivity of the cautery-based classifier increased by 20% when the two sensors were closer together,

indicating that this configuration of sensors is better. Thus, the instances where the sensors were close together were treated as a separate case in developing the cautery-based classifier.

Table 2. A summary of the performance of the readings from the frame-based classifier and cautery-based classifier in predicting true measurement error. The left-hand column in the confusion matrix refers to the number of predicted cases, while the top row refers to the number of actual cases in each category. The sensors on the cautery were close together.

Frame-Based Classifier Confusion Matrix			**Cautery-Based Classifier Confusion Matrix**		
	Act. False	*Act. True*		*Act. False*	*Act. True*
Pred. False	29	22	*Pred. False*	13	4
Pred. True	5	16	*Pred. True*	1	18
Frame-Based Classifier Performance Measures			**Cautery-Based Classifier Performance Measures**		
Specificity	85%		*Specificity*	93%	
Sensitivity	42%		*Sensitivity*	82%	
Accuracy	63%		*Accuracy*	86%	

A greater volume of data is needed to refine our classification methods and further reduce the false negative error rates. However, it does seem feasible to use redundant sensors as predictors of measurement error. Once a suitable error threshold is determined, a 3D Slicer module could use it to provide real-time error monitoring feedback during navigated procedures. Ultimately, redundant sensors would provide an error monitoring solution that is easy to set up and causes no interference to the ongoing surgical procedure—circumventing the limitations of previously mentioned error monitoring methods.

4. CONCLUSION

Redundant electromagnetic sensors were added to a navigated intervention setup. Readings from different subsets of these redundant sensors were used to train two linear support vector machine-based classifiers to predict tracking error. Although the classifiers had sufficiently high specificities to demonstrate a proof-of-concept, their sensitivities were too low for safe use within a real procedure. Further experiments will be done with different electromagnetic sensors and in different surgical setups to improve our classifier's performance. Our system is a useful research tool, as it is configurable and open-source. Upon further refinement, we believe that the presented method can be used for intraoperative quality assurance feedback of electromagnetically navigated interventions.

ACKNOWLEDGEMENTS

This work was supported in part Discovery Grants Program of the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Applied Cancer Research Unit program of Cancer Care Ontario with funds provided by the Ontario Ministry of Health and Long-Term Care. Vinyas Harish and Zachary Baum were supported by the NSERC Undergraduate Student Research Awards (USRA) program. Eden Bibic was supported by the Queen's University Internships In Computing (QUIC). Matthew S. Holden and Thomas Vaughan are supported by the NSERC Canada Graduate Scholarship (CGS). Gabor Fichtinger is supported as a Cancer Care Ontario Research Chair in Cancer Imaging.

REFERENCES

[1] Franz, A.M., Haidegger, T., Birkfellner, W., Cleary, K., Peters, T.M., and Maier-Hein, L., "Design and application of an assessment protocol for electromagnetic tracking systems," *Medical Physics* 32(7), 2371-2379 (2005).

[2] Hummel, J.B., Bax, M.R., Figl, M.L., Kang, Y., Maurer, C. Jr., Birkfellner, W.W., Bergmann, H., and Shahidi, R., "Design and application of an assessment protocol for electromagnetic tracking systems," *Medical Physics* 32(7), 2371-2379 (2005).

[3] Shen, E., Shechter, G., Kruecker, J., and Stanton, D., "Quantification of AC electromagnetic tracking system accuracy in a CT scanner environment," *Proc. SPIE Med. Imag.: Visualization and Image-Guided Procedures 6509*, 1-10 (2007).

[4] Harish, V., Baksh, A., Ungi, T., Lasso, A., Baum, Z., Gauvin, G., Engel, J., Rudan, J., and Fichtinger, G., "Measurement of electromagnetic tracking error in a navigated breast surgery set-up," *Proc. SPIE Med. Imag.: Image-Guided Procedures, Robotic Interventions, and Modeling 9786*, 1-8 (2016).

[5] Vaccarella, A., De Momi, E., Enquobahrie, A., and Ferrigno, G., "Unscented Kalman Filter Based Sensor Fusion for Robust Optical and Electromagnetic Tracking in Surgical Navigation," *Instrumentation and Measurement, IEEE Transactions on* 62(7), 2067-2081 (2013).

[6] Jain A.K., Matinfar, M.B., Chan, R., Parthasarthy, V., and Stanton D.A., "Method and System for Characterizing and Visualizing Electromagnetic Tracking Errors," Patent WO 2011110966 A2 (2011).

[7] Lasso A., Heffter, T., Rankin A., Pinter, C., Ungi, T., and Fichtinger, G., "PLUS: open-source toolkit for ultrasound-guided intervention systems," *Biomedical engineering, IEEE Transactions on* 61(10), 2527-2537 (2014).

[8] Tokuda,J., Fischer, G.S., Papademetris, X., Yaniv, Z., Ibanez, L., Cheng, P., Liu, H., Blevins, J., Arata, J., Golby, A.J., Kapur, T., Pieper, S., Burdette, E.C., Fichtinger, G., Tempany, C.M., and Hata, N., "OpenIGTLink: an open network protocol for image-guided therapy environment," *International Journal of Medical Robotics and Computer Assisted Surgery* 5(4), 423-434 (2009).

[9] Ungi T., Gauvin G., Lasso, A., Yeo, C., Pezeshki, P., Vaughan, T., Carter, K., Rudan, J., Engel, J., and Fichtinger, G., "Navigated Breast Tumor Excision Using Electromagnetically Tracked Ultrasound and Surgical Instruments," *Biomedical engineering, IEEE Transactions on* 63(3), 600-606 (2016).

[10] Kotsiantis, S.B., "Supervised Machine Learning: A Review of Classification Techniques," *Proc. Emerging Artificial Intelligence Applications in Computer Engineering*, 3-24 (2007).