StereoCNC: A Stereovision-guided Robotic Laser System

Guangshen Ma [1], Weston Ross [2] and Patrick J. Codd [1,2]

Abstract— This paper proposes a stereovision-guided robotic laser system that can conduct laser ablation on targets selected by human operators in the color image, referred as *StereoCNC*. Two digital cameras are integrated into a previously developed robotic laser system to add a color sensing modality and formulate the stereovision. A calibration method is implemented to register the coordinate frames between stereo cameras and the laser system, modelled as a 3D-to-3D Least-squares problem. This problem is solved by a RANSAC-based 3D rigid transformation method and the calibration reprojection errors are used to characterize a 3D error field by Gaussian process regression. This regression error model is used to predict an error value for each data point of a stereo-reconstructed point cloud and an optimization problem is formulated to adjust the surgical site to a new position with minimum reprojection errors. Based on the calibrated system and the error model, a stereovision-guided laser-tissue removal pipeline is proposed to precisely locate, target, and ablate a surface region. The pipeline is validated by the experiments on phantoms with color texture and various geometric shapes. The overall targeting accuracy of the system achieves an average RMSE of $0.13 \pm 0.02 \ mm$ and maximum error of $0.34 \pm 0.06 \ mm$, as measured by pre- and post-laser ablation images. The results show potential applications of using the developed stereovision-guided robotic system for superficial laser surgery, including dermatologic applications or removal of exposed tumorous tissue in neurosurgery.

I. INTRODUCTION

Robotic laser system can assist surgeons in the use of laser scalpels by providing precise position control with a two-axis laser galvanometric system (LGS) and high-fidelity vision with an external sensor [1], [2], [3], [4]. Commonly employed sensors include Optical Coherence Tomography (OCT) [5], fluorescence microscopy [6], monocular cameras and stereovision [2], [3]. Of these, monocular and stereovision are the most user-friendly as they capture visible light and present it in a clinically tractable format such as a color digital image. Stereovision is a well-studied technique to extract 3D information from two digital cameras and it can be easily integrated to stereo endoscope systems [7] and robotic laser systems [2]. Such systems enable image-guided surgery and can provide higher level analysis such as object tracking and 3D reconstruction of the surgical site [2], [3]. Stereovision is advantageous due to the fast 3D reconstruction with colorized texture information without the need for advanced optics and structured light, especially working with biological tissue [8].

The benefit of stereovision motivates this work to integrate it as an additional sensing modality into our prior

[1] Brain Tool Lab, Department of Mechanical Engineering and Materials Science, Duke University

[2] Department of Neurosurgery, Duke University Medical Center
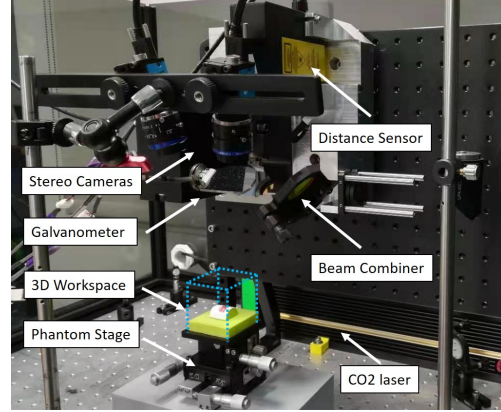
* Accepted by IROS 2021 with final submission

Fig. 1: StereoCNC system components: a free space CO_2 laser is guided into a beam combiner before entering the two-axis galvanometer system to be directed onto the surgical surface on the 3D stage. A visible light distance sensor also enters the galvanometer at the beam combiner. The stereovision system is aligned such that its field of view (FOV) overlaps with the galvanometer.

robotic laser platform [9] (referred as *TumorCNC*). The new system is referred to as the *StereoCNC* and it combines a stereovision and a CNC-like laser system for superficial tissue ablation. *StereoCNC* is intended to provide interactive image guided surgery with the color vision feedback from the stereovision system, which includes texture and topological information from triangulation surface reconstruction. This new system enables users to choose a pixel region in a digital color image and generate a 3D surface point cloud for superficial laser-tissue removal.

In this work, we incorporate stereovision to the existing *TumorCNC* system [9]. This system consists of a CO_2 cutting laser (invisible), a laser (visible) distance sensor and a two-axis galvanometer to guide the laser towards the surgical site. *TumorCNC* can scan a 3D point and generate a surface profile. The stereovision can also track the visible laser and calculate the 3D coordinate of the laser spot simultaneously, with respect to its frame of reference. With 3D corresponding points measured in the two frames, the system calibration is to find the rigid transformation between stereovision and the *TumorCNC* system, which can be modelled as a 3D-to-3D Least-squares problem.

However, the direct solution for the 3D Least-squares problem can be affected by two error sources: stereo 3D reconstruction error and the *TumorCNC* system error. The stereovision reconstruction error can increase quadratically

with depth [10], [11] while the *TumorCNC* system error was less than 0.1 mm as reported in [9], [12]. Stereovision is expected to contribute greatly to the calibration error. This indicates the distribution of the calibration error is not uniform in the 3D workspace. In addition, the effective calibrated region is related to the field of view and the depth of field from the stereo cameras. An effective calibrated region should exist and we hypothesis that a region with lower calibration error could be determined.

In summary, this study proposes an efficient system calibration method to precisely achieve stereovision-guided superficial laser-tissue removal, including three main objectives: estimate the correct transformation between the stereovision and the laser galvanometric scanner (LGS) by solving the 3D Least-squares problem, characterize the 3D calibration error field by using Gaussian process regression [13] and develop an effective algorithm to adjust the surgical site to a low calibration error region. This calibration method uses a regression method to model the End-to-End system calibration error, and does not require a complex modelling between the stereo cameras and the LGS. Two phantom studies validate the feasibility of the proposed calibration pipeline in stereovision-guided laser surgery.

II. RELATE WORK

The proposed work is mostly related to the calibration of the stereovision and the laser galvanometric scanner (referred as Stereo-to-LGS calibration), with the major difference that our study uses a laser distance sensor with the LGS. This is a challenging calibration problem because of the complexity in geometry modelling and the sensitivity of the systematic errors. The solutions to the calibration problem can be summarized as Model-based and Model-free methods. This section also reviews the studies of Automating surgical instrument positioning and Calibration error modelling.

Model-based Calibration: Model-based calibration generally aims to formulate an optimization problem based on the geometric relation between the stereovision camera model and the configuration of LGS. Qi *et al.* proposed a calibration strategy to estimate the 2D-to-3D transformation between the LGS image plane and the 3D stereo image frame [14]. However, this study does not include a laser distance sensor and mainly was designed for a long baseline stereovision system in order to reduce the 3D reconstructed error, which makes it difficult to be integrated into the existing robotic system. Andreff *et. al* employed a trifocal configuration between a two-cameras system and a LGS for laser steering [3]. This work presents an effective visual servoing control algorithm for positioning a servo-guided laser onto the surface as indicated by a reference trajectory. In addition, Schoob *et. al* proposed a system that uses stereo cameras for surface tracking during tissue deformation [2]. This system implements a simplified laser-stereovision system for image-guided laser surgery, while not focusing on the 3D geometric analysis, since their stereovision system was not used for 3D reconstruction. Additionally, the study in [15], though not with stereovision, discussed a Random

sample consensus (RANSAC) based 3D-to-3D registration method to estimate the transformation between a 3D range sensor (e.g. Kinect RGB-D camera) and the LGS, which is mostly related to our proposed work.

Model-free Calibration: Model-free calibration aims to solve the Stereo-to-LGS calibration problem by using the statistical regression methods or artificial neural networks (ANN) [16]. As the Model-based methods can cause ill-conditioned matrix problems due to the sensor noise or the system errors, Model-free solutions show a major benefit of modelling the complicated calibration problem by using an End-to-End configuration (e.g. ANN) and data-driven techniques. Tu *et al.* proposed an End-to-End Stereo-to-LGS calibration strategy to find the mapping between the mirror input digital voltage and the output laser beam, by using a extreme learning machine method [4]. Similarly, Wissel *et al.* discussed a comparison study of various data-driven techniques in calibrating the galvanometric laser scanners, including the look-up table, Gaussian process regression and ANN [17]. These studies show feasibility of using purely data-driven methods for calibration, however, these methods tend to require a large amount of sampled data for model training and the interpretation problems of ANN remain unsolved for system calibration.

Calibration Error Modelling and Automated Positioning: Several studies have been proposed to model the End-to-End system calibration error by ANN or the regression methods. For example, the study in [18] proposed a method of modelling the stereovision volumetric error by using ANN. Similarly, our proposed study aims to use the calibration reprojection errors to formulate a 3D error field. A common non-parametric regression method, such as Gaussian process regression [13], can be used to fit the reprojection errors and predict an error value for each data point of a given point cloud. Therefore, a following problem is the determination of an optimal object position to minimize the total calibration errors as it will greatly affect the precision of laser ablation. Several studies on automating surgical instrument positioning have been proposed for laparoscopic and eye surgery [19], [20]. Differently, our work aims to optimize the positioning of the surgical site to minimize the calibration errors.

III. METHODS

A. Hardware Architecture

The *StereoCNC* includes a stereovision system, a 3D laser distance sensor (Mechanical Technology Inc., Albany, NY), a 10W CO_2 cutting laser (Synrad Inc., Mukilteo, WA), a two-axis galvanometer system (Cambridge Technology, Bedford, MA) and a mirror system for laser beam steering, as shown in Fig. 1. A manual control stage (Thorlabs, Newton, NJ) is used to control the position of surgical site with a resolution of 0.02 mm. The stereovision system consists of two 1280×720 USB 3.0 industrial cameras (Model: DFK 33UP1300, The Imaging Source, NC USA) and Two 8 mm lenses (Model: TCL 0814, The Imaging Source, NC USA) mounted in a fixed stage. The target object is localized within the camera focus range. In this prototype, the galvanometer and

the distance sensor are controlled by a multi-function data acquisition (DAQ) device (National Instrument, model USB-6211) and programmed in Python.

B. Review of the StereoCNC Research Platform

There are two lasers in the system: the red, visible triangulation sensor laser and the invisible, CO_2 cutting laser. Both lasers enter the galvanometer from the beam combiner, but their trajectories are not perfectly aligned. For a given galvanometer position, the lasers are incident at different points in a 3D workspace, as shown in Fig. 2. In previous work, a system calibration procedure was developed to estimate the positions of the distance sensor and the CO2 laser [9], [12]. The galvanometer mirrors angles can thus be calculated and the system can steer the CO_2 laser to the same 3D point targeted by the distance sensor (with visible laser spot). Apart from the stereovision system, this robotic surgical platform performs two main functions:

- Locate a 3D point by the visible laser and describe it at the world frame $\{W\}$, referred as $p^W \in \mathbb{R}^3$. $\{W\}$ is defined at the specified position of the Y-mirror [9].
- Given p^W, the mirror angles (θ_x, θ_y) can be calculated such that the CO_2 cutting laser can be guided to target at the point on the surface.

The full calibration procedure and details of the whole system are discussed in [9]. In this study, we assume that the angular conversion function f_θ is already known. f_θ is an implementation of a binary search to find the optimal mirror angles to reach at a 3D coordinate with a desired laser output vector, until an accepted precision is satisfied [9]. Given a 3D point defined as $p^W = (x_w, y_w, z_w)$, the mirror angles can be computed as $\theta_x, \theta_y = f_\theta(p^W)$.
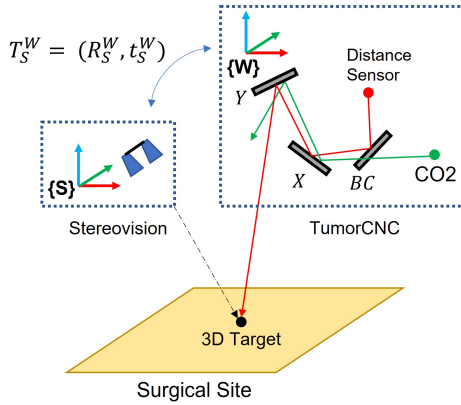

Fig. 2: StereoCNC system model. **X** and **Y** denote the two-axis galvo-mirrors. **BC** is the beam combiner. $\{W\}$ and $\{S\}$ are the world and stereovision coordinate frames.

C. Stereovision System Calibration

1) System Model: Two fixed cameras can formulate a stereovision and be used for 3D reconstruction. Given a rectified stereo configuration, the corresponding pixels determined by disparity matching in both cameras are used

to estimate a depth and a 3D point can be reconstructed [21]. We define the left camera as a reference frame $\{S\}$ for stereovision and the 3D reconstructed coordinate as p^S.

To find the galvanometer mirrors angles for laser cutting, p^S is first transformed to the world frame by a transformation $T_S^W = (\boldsymbol{R}_S^W, \boldsymbol{t}_S^W)$:

$$p^W = \boldsymbol{R}_S^W p^S + \boldsymbol{t}_S^W \tag{1}$$

Where $\boldsymbol{R}_S^W$ is a standard 3×3 rotation matrix, $\boldsymbol{t}_S^W$ is a 3D translation vector. Substituting p^W into $f_\theta(\cdot)$, the mirror angles for the cutting laser are calculated as:

$$\theta_x, \theta_y = f_\theta(\boldsymbol{R}_S^W p^S + \boldsymbol{t}_S^W) \tag{2}$$

θ_x and θ_y are mirror angles that control the laser beam to be incident at the targets. Equation (2) describes the fundamental equation for the 2D-to-3D stereovision guided laser system, which can estimate the galvanometer mirror angles to control the CO_2 laser to reach at a 3D point identified by a 2D pixel position in the (left) color image.

2) 3D-to-3D Rigid Transformation: To determine T_S^W, we formulate the calibration as a 3D-to-3D Least-squares problem. Points within the field of view of both systems are simultaneously recorded in each system frame of reference. This requires the system to collect sample data points from various positions (X-Y-Z) in the 3D workspace. The calibration board is moved to N depth positions to collect M sampled data within a raster scanning region. The calibration is formulated as a Least-squares optimization problem that aims to minimize the 3D distance square errors between the measured points defined in $\{W\}$ and $\{S\}$:

$$\min_{\boldsymbol{R}_S^W, \, \boldsymbol{t}_S^W} \sum_i^M \sum_j^N ||p_{ij}^W - (\boldsymbol{R}_S^W p_{i,j}^S + \boldsymbol{t}_S^W)||_2^2 \tag{3}$$

Where $|| \cdot ||_2$ denotes the L2-norm. The i-th is the index for each measurements and j-th is the index of the depth of object. This is a well-defined 3D Least-squares problem which can be solved by Singular Value Decomposition [22].

3) RANSAC-based Calibration: In a large workspace, the 3D Least-squares solution can be affected by the outliers caused by the stereo reconstruction errors or the points locating outside of the camera focus distance [10]. This problem can be solved by Random sample consensus (RANSAC), which is an iterative method to estimate the correct model parameters by detecting outliers from the observed data points [23]. Sampled data points were obtained by controlling the visible laser to scan a $25.4 \times 25.4 \times 10.5 \ mm^3$ 3D region in the effective workspace. For each measurement, the stereovision cameras can precisely detect the 2D center laser spots by setting the camera to a low exposure time. This will block all the image pixels except for the brighter laser spots and enables a precise 3D stereo reconstruction. For generality of the calibration, the scanning region is moved to different depths to collect enough sample data (Fig. 3).

In this study, the proposed RANSAC-calibration method is slightly different from the generic RANSAC method [23] and
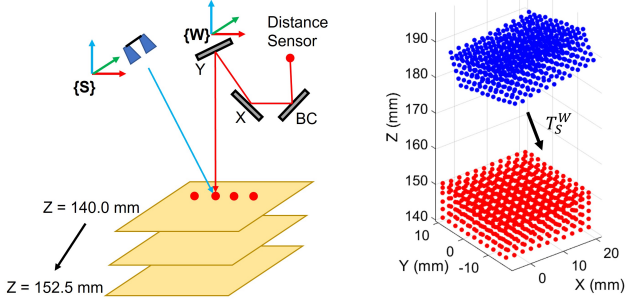
Fig. 3: Stereo-to-World transformation for system calibration. The raster scanning is performed at various depths to cover the entire region of interest. T_S^W is the transformation between $\{W\}$ and $\{S\}$.

the selection of the parameters are based on the experimental tests. In theory, three non-collinear points are required for 6-DOF rigid transformation. The minimum number of points for the model is set as a higher value for a more robust model estimation. The inlier ratio is set as 75% and the error threshold is defined as 0.50 mm. The best fit from RANSAC was obtained based on the fitting error threshold and the ratio of inliers.

D. Calibration Error Field Modelled by Gaussian Process

Based on RANSAC calibration, the reprojection error $E_{i,j}^{proj}$ for each sampled data point is obtained by applying the optimal transformation $\overline{T_S^W} = (\overline{R_S^W}, \overline{t_S^W})$ for the sample data $p_{i,j}^W$:

$$E_{i,j}^{proj} = ||p_{i,j}^W - (\overline{R_S^W} p_{i,j}^S + \overline{t_S^W})||_2 \qquad (4)$$

The calibration error distribution is characterized by $E_{i,j}^{proj}$ where each 3D coordinate in the sample data is assigned with an error value. It is important to use the posterior error distribution to make predictions for the unobserved data points by Gaussian process regression (GPR) [13]. The calibration reprojection errors and the corresponding positions are used to estimate the GPR model parameters.

The input data for GPR is a set of training data D_t, collected from the RANSAC-calibration with registration error Y taken from 3D sampled points at X_t, denoted as $\{D_t = (X_t, Y_t)|(X_t, Y_t) = (\mathbf{x_1}, y_1), (\mathbf{x_2}, y_2), \cdots, (\mathbf{x_n}, y_n), X_t \in \mathbb{R}^3, Y_t \in \mathbb{R}\}$ for a set of n sample data. Given D_t, the posterior distribution of the learned GPR function $f(\cdot)$ for an unobserved data point x^* with mean μ_{x^*} and $\sigma_{x^*}^2$ are denoted as:

$$p(f(x^*)|x^*, X_t, Y_t) = \mathcal{N}(f(x^*); \mu_{x^*}, \sigma_{x^*}^2) \qquad (5)$$

Where $\mu_{x^*} = k_*^T(K + \sigma_n^2 I)^{-1} Y_t$ and $\sigma_{x^*}^2 = k(x^*, x^*) - k(x^*, X_t)^T(K + \sigma_n^2 I)^{-1} k(x^*, X_t)$. k_*^T is the covariance vector between the test points and the training inputs. $k(x^*, X_t)$ is a covariance vector of the predicted position x^* and the training data X_t, by using a popular squared exponential kernel [13]. K is the covariance matrix of X_t and σ_n^2 is the

noise of the data (Gaussian i.i.d. noise). Given a new 3D position x^*, the predicted error E^{GPR} can be estimated by:

$$E^{GPR}(x^*) = \mu_{x^*} = k_*^T(K + \sigma_n^2 I)^{-1} Y_t \qquad (6)$$

Based on Equation (6), all the registration error values at a given point cloud can be estimated by the GPR model.

E. Optimal Object Position by Genetic Algorithm

Given an object in 3D, it is very likely that the initial position is not located at a region with the minimal calibration errors. We seek to find a new position with globally minimized calibration error where the laser can precisely target at each point. Given a 3D point cloud measured from the object surface with K data points, we define a cost function based on the GPR error field:

$$\min_{\Delta x_{opt} \in \mathbb{R}^3} \sum_{i=1}^{K} E_i^{GPR}(x_i + \Delta x_{opt}) \qquad (7)$$

Where $x_i \in \mathbb{R}^3$ is the i-th 3D coordinate of a point cloud when the object is at the original position. Δx_{opt} is the optimal offset between the initial position and the final position. The new position $x_i + \Delta x_{opt}$ for each iteration is defined in a searching region of $12.7 \times 17.8 \times 10.1\ mm^3$. This region is smaller than the raster scanning region since only the effective sampling data points (with reprojection error less than 0.50 mm) are selected for GPR fitting. A Genetic Algorithm (GA) was applied to solve this global optimization problem by using the MATLAB Global Optimization Toolbox. GA first generates an initial, randomized population in the total error field and looks for the seed candidates with a lower fitness value. The optimization procedure aims to find lower fitness function values and terminates when the difference between each generation is smaller than the function tolerance defined as 10^{-6}. The output of the GA is an optimal offset Δx_{opt} between the old and new positions that can be adjusted by the motion stage.

IV. EXPERIMENTS AND RESULTS

A. RANSAC-based Calibration and GPR Error Field

The calibration errors are calculated by Equation (4) and used for fitting the GPR model. For visualization, GPR was applied in a $12.7 \times 17.8 \times 10.1\ mm^3$ mesh grid to show the calibration error field. We define a "low-error" region based on a reference threshold of 0.35 mm and this is used for validating the proposed calibration method. Fig. 5 (a) illustrates the full representation of a 3D error field predicted by GPR and (b) depicts a low-error field with a threshold of 0.35 mm. The 0.35 mm error field is within the range of $10.4 \times 17.8 \times 10.1\ mm^3$ and this region provides adequate space for regular pathological tissue ablation such as skin lesion (tumor size < 5.0 mm) in dermatology. Furthermore, Fig. 5 (c) and (d) illustrate two low-error regions with thresholds of 0.25 mm and 0.20 mm, respectively. The error distribution is consistent since the data points inside the volumetric region show lower error attributes than the outward data. This demonstrates that the
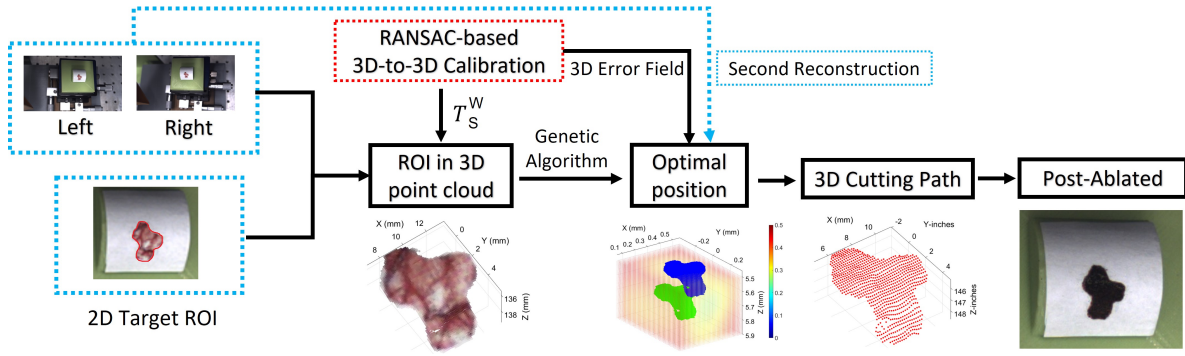
Fig. 4: Stereovision-guided pipeline for laser surgery: The stereo reconstruction point cloud is first transformed to the world frame by the calculated transformation T_S^W. The Genetic Algorithm (GA) is applied to find an optimal phantom position (blue and green point cloud only for visualization). The object is moved to a new position and the stereo reconstruction is applied again to create a 3D cutting path. The post-ablated error is measured in pixel distance.

RANSAC-calibration method can localize a low-error region with a greater standard of error tolerance.

(a) GP Error Field (Effective area measured from a 12.71 x 17.81 x 10.08 mm³ range)

(b) GP Error Field (0.35 mm threshold 10.43 x 17.81 x 10.08 mm³ range)

(c) GP Error Field (0.25 mm threshold and 5.59 x 15.68 x 6.45 mm³ range)

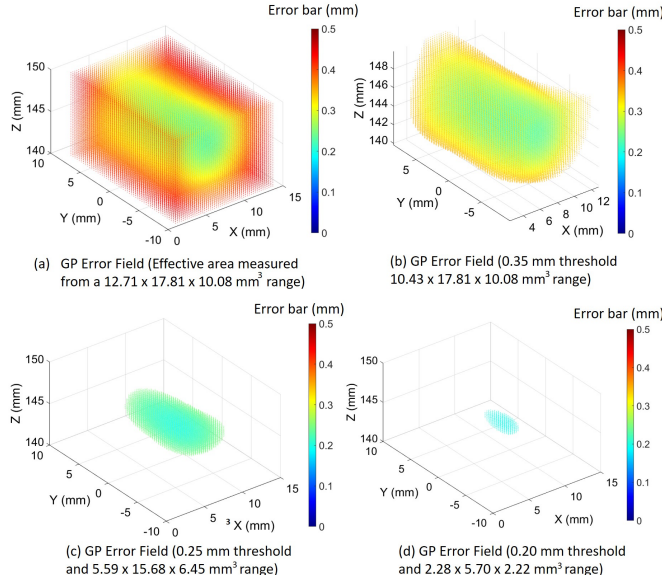(d) GP Error Field (0.20 mm threshold and 2.28 x 5.70 x 2.22 mm³ range)

Fig. 5: 3D calibration error field modelled by Gaussian process regression.

B. Study 1: Ablated Point Measurement

Study 1 aims to measure the accuracy of the 3D error fields predicted by GPR. As the stereo reconstruction error contributes to the system error with a greater ratio, the error between pre- and post- laser ablations (guided by stereovision) should be consistent with the GPR error fields. To test this hypothesis, we designed an experiment to measure the point ablation error in various positions (Fig. 6). We used a planar phantom with discrete colorized fiducials that can show clear features for precise 3D centroid detection with stereovision. The detected 3D positions were transformed to the world coordinate frame and the angles of galvanometer were calculated for laser cutting. The laser ablation at each fiducial point was conducted and the post-ablated image was

captured for error measurement, as shown in Fig. 7. As the effective GPR error field in Z-axis ranges from 142.5 mm to 150.0 mm, the planar phantom was moved to 6 equally spaced depths and the study was repeated for each position.

The result of *Study 1* is shown in Fig. 7 from Z_1 to Z_6. The post-ablation center (black dot) was compared with the pre-ablation center labelled by the operator. The distance error was measured in pixel distance and converted to millimeter unit based on the ratio of 0.109 (1 pixel $\approx$ 0.109 mm in a working distance between 140.0 mm to 152.5 mm in $\{W\}$). The distributions of low-error fiducials (0.35 mm and 0.25 mm, labelled in blue and red circles) are more consistent from $Z_1 = 140.0$ mm to $Z_4 = 147.5$ mm and become more sparse at Z_5 and Z_6. This is because Z_5 and Z_6 are close to the boundary of the low-error region and the distribution of the calibration error becomes inconsistent. Nevertheless, the existence of a consistent low-error region from Z_1 to Z_4 demonstrates that the low-error region for laser ablation follows the trend of the GPR calibration error field. This suggests that a consistent 3D low-error region can be localized with the RANSAC-calibration method.

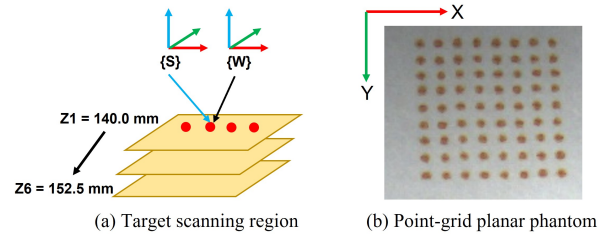(a) Target scanning region   (b) Point-grid planar phantom

Fig. 6: Setup of *Study 1*. (a) shows the testing studies at different depths. (b) depicts the planar phantom with colorized fiducials whose 3D centers can be easily detected by stereovision.

C. Study 2: Stereovision-guided Targeted Surface Ablation

The goal of this experiment was to validate that the proposed stereovision pipeline can precisely locate, target,
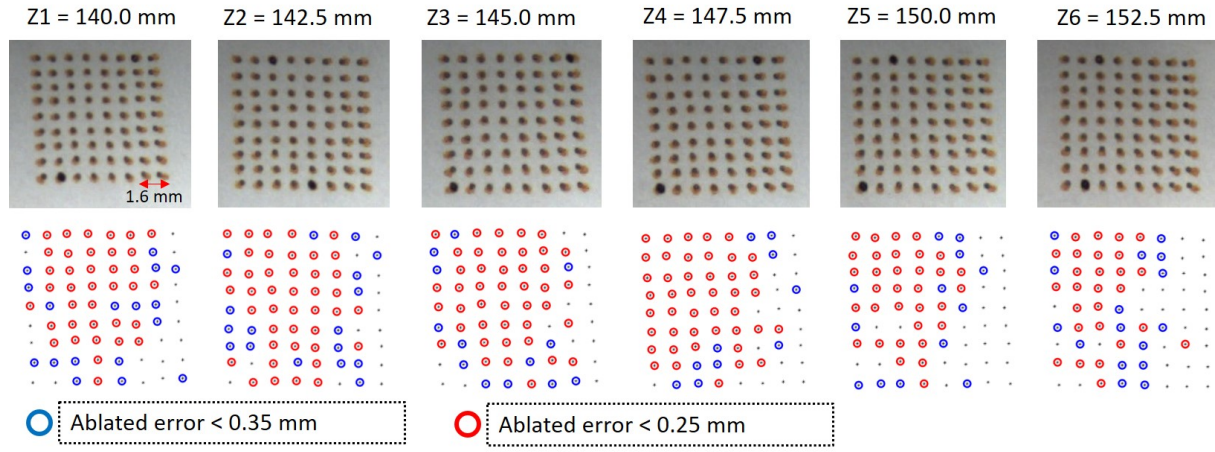
Fig. 7: Result of *Study 1*. Point ablations based on a planar calibration board placed at various depths. The first row from Z_1 to Z_6 shows the images with ablated centers. The second row shows the ablated errors of each fiducial position.

and ablate the surface region identified by the user (Fig. 11). A study of 9 tests were conducted on 3D-printed phantom with planar, concave and convex shapes to test the generality of the proposed method. The phantom were attached with colorized texture tapes with clear features for successful stereo reconstruction. The 9 tests were conducted with a stereovision-guided pipeline to target at a 3D region of interest (ROI) identified by the contour of the 2D pixel region, as shown in Fig. 4. The point cloud at the optimal position was first projected to a 2D perpendicular plane, re-sampled and a cutting path was created by uniformly aligning the laser points at the 2D projected points.

1) Stereo Reconstruction for Surface Point Cloud: Stereo Reconstruction was mainly implemented by using Python-OpenCV toolbox. The stereo cameras were calibrated to find the intrinsic and extrinsic parameters for stereo rectification. Given two pixel coordinates in left and right image frames, the depth of a 3D point was estimated by disparity matching via a generic SGBM algorithm [24]. A 3D surface was created based on the pixel region selected in the color image and the reconstructed point cloud was transformed by T_S^W for laser cutting (Fig 4).

2) Genetic Algorithm Optimization: A Genetic Algorithm was employed to optimize the location of a 3D object surface in the GPR low-error region based on a 0.35 mm threshold. It is expected that GA can minimize the total calibration errors of a given point cloud. For example, Fig. 8 depicts the error predictions on a given point cloud before and after GA optimization. The Root-mean-square of the calibration error (RMSE) reduces from 0.36 mm to 0.22 mm via GA optimization. This decrease of RMSE validates the feasibility of using GA for total error minimization.

The performance of GA is evaluated by tracking the fitness function values over the number of generation. Fig. 9 illustrates the fitness values of the GA performances with all the 9 tests with three various shapes (planar, convex, and concave). The mean and best fitness values can successfully converge to the global minimums at the final generation,
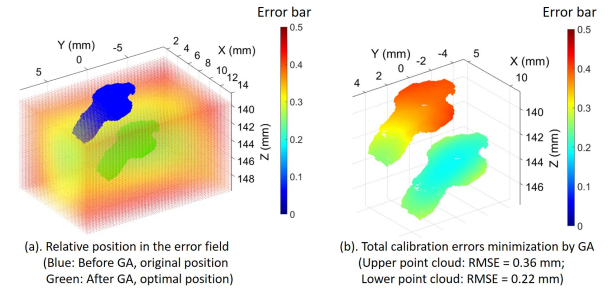

Fig. 8: Error fields comparison before and after GA optimization. Specifically in (a), the color (blue and green) does not represent the error distribution (only for visualization).
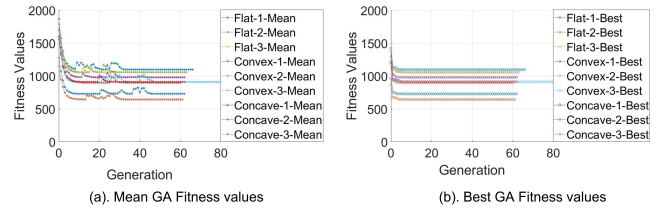
which takes around 10.0 seconds for computation.

Fig. 9: Best and mean Fitness values in GA optimization. The fitness value is expected to decrease for the objective minimization.

3) Error Measurement: The performance of the study was evaluated primarily with two metrics: contour error and surface error. For surface ablation, E_{dis} measures the distance error between the pre-ablation object contour and the post-ablation boundary that were labelled by the operator, referred C_{pre} and C_{post}. Similar to a target path tracking for precision evaluation in [25], the distance error E_{dis} is assumed to be the minimum distance (Infimum) from a query point in C_{pre} to C_{post}:

$$E_{dis} = ||p_{pre} - \inf_{p_{post} \in C_{post}} (||p_{pre} - p_{post}||_2)||_2 \quad (8)$$

Where $p_{pre} \in C_{pre}$. The root-mean-square error (RMSE) and maximum error (MAE) were calculated based on E_{dis}, which could be used to measure the deviation of pre- and post-ablation contours (Fig. 10 (c)).
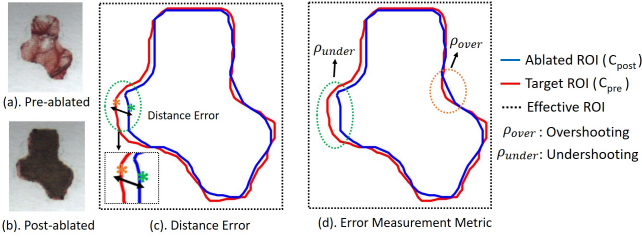

Fig. 10: An example of error measurement: (a) and (b) shows pre- and post-ablation images. (c) shows the diagram of distance error measurement. (d) depicts the region of overshooting and undershooting.

For surface errors, the DICE similarity coefficient has been commonly used to evaluate the similarity of overlapping regions and thus utilized to compare the ablated and target area [26]. The DICE coefficient is defined as $DICE = \frac{2 \times Area(overlapping)}{Area(Target) + Area(Ablated)}$. We also introduce the ratios of overshooting (ρ_{over}) and undershooting (ρ_{under}) as the important factors of evaluating the precision and robustness of a laser cutting procedure. The overshooting and undershooting ratios are defined as $\rho_{over} = \frac{Area(Overshoot)}{Area(Target)}$ and $\rho_{under} = \frac{Area(Undershoot)}{Area(Target)}$ respectively, as shown in Fig. 10 (d). The definition of ρ_{under} and ρ_{over} share the same denominator so as to compare the undershooting and overshooting effects on the system performance. ρ_{under} denotes the percentage of pixel area not covered by the laser irradiation that should have been targeted, while ρ_{over} describes the overshooting area as an indication of shooting the laser to an incorrect area outside of the targeted region. The results of ρ_{over} and ρ_{under} are shown in Table. I.

TABLE I: Statistical Measurements of 3D Objects in Gaussian Error Fields (rounded to 2 decimals places.)

	RMSE (mm)	MAE (mm)	ρ_{under} (%)	ρ_{over} (%)	DICE (%)
Planar-1	0.12	0.33	0.63	3.14	97.97
Planar-2	0.13	0.33	3.80	2.02	96.81
Planar-3	0.18	0.43	5.58	2.47	95.66
Convex-1	0.15	0.43	5.33	0.65	96.76
Convex-2	0.14	0.31	4.55	1.25	96.77
Convex-3	0.10	0.24	1.83	2.24	97.75
Concave-1	0.12	0.31	0.95	3.97	97.31
Concave-2	0.11	0.33	1.76	3.71	97.19
Concave-3	0.15	0.39	3.25	4.03	96.00
Mean	0.13	0.34	3.08	2.61	96.91
Standard Deviation	0.02	0.06	1.86	1.20	0.75

V. DISCUSSION

The proposed Stereo-to-LGS calibration pipeline overcomes the common reconstruction errors of standard stereovision systems and models the calibration as a 3D Least-squares problem. The results indicate that a RANSAC-based 3D transformation method can be used to solve the calibration problem. The 3D reprojection error field, as characterized by Gaussian process regression, can be used for controlling the object position to a lower calibration error region and performing the precise laser-tissue removal. The error field has a reference threshold of 0.35 mm (all errors < 0.35 mm) and an estimated dimension of $10.4 \times 17.8 \times 10.1 \ mm^3$, and clinically this would be adequate for superficial tissue ablation for many dermatological applications such as typical skin lesion with sizes around $5 \times 5 \ mm^2$.

One of our hypotheses was that the reconstruction error contributes the greatest to the total calibration error, and this would affect the accuracy of the 3D Least-squares calibration. Results from *Study 1* demonstrate that the dimension of the ablated error region matches the GPR calibration error field based on a 0.35 mm error threshold, supporting our hypothesis. This suggests the proposed RANSAC-calibration method can successfully find the low-error region for precise laser surgery. For each surgical case, either the surgical site or the laser scalpel can be controlled to a new position to achieve a precise stereovision-guided laser-tissue removal.

The second experiment evaluates the performance of the proposed stereovision-guided laser ablation pipeline with phantoms of various shapes (convex, concave, and planar). The average of RMSE and maximum laser cutting error among the 9 testing cases are reported as $0.13 \pm 0.02 \ mm$ and $0.34 \pm 0.06 \ mm$. These results validate that the proposed laser surgical platform can robustly perform laser ablation on surface and the precision is not sensitive to the change of object shapes. The mean laser overshooting ratio (2.61%) is smaller than the mean undershooting ratio (3.08%) for surface ablation as measured by the target region and post-ablated area. This suggests that the system has a slightly higher chance of missing a labelled region than cutting outside of the target region into healthy tissue that is not labelled by the user. Depending on the surgical procedure, over-resection vs under-resection carry different risks. Furthermore, the 96.91% DICE similarity score demonstrates that the ablated region can successfully match the region of target.

System integration with different sub-components will inevitably cause various error sources. This study proposes a data-driven technique to model the End-to-End calibration errors and does not require a complex description of the systematic error. Future studies include the improvements of stereo reconstruction algorithms and the *ex-vivo* animal tissue study.
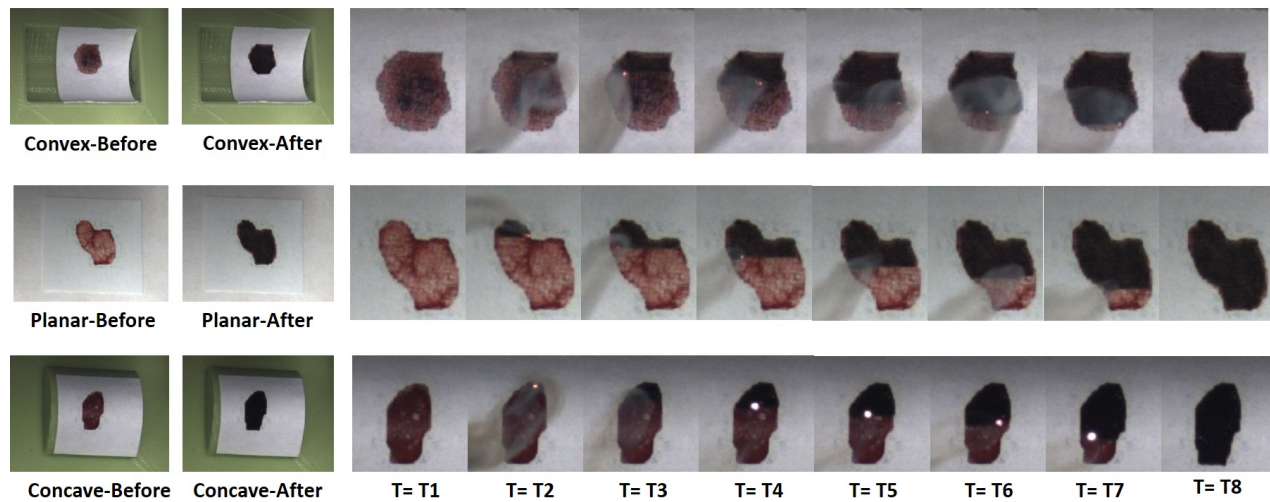
ACKNOWLEDGMENT

Fig. 11: Laser ablation at each time step. The ablation acts like a CNC-process to remove the tissue line-by-line.

REFERENCES

[1] Q. Peng, A. Juzeniene, J. Chen, L. Svaasand, T. Warloe, K.-E. Giercksky, and J. Moan, "Lasers in medicine," *Reports on Progress in Physics*, vol. 71, p. 056701, 04 2008.

[2] A. Schoob, D. Kundrat, L. A. Kahrs, and T. Ortmaier, "Stereo vision-based tracking of soft tissue motion with application to online ablation control in laser microsurgery," *Medical image analysis*, vol. 40, pp. 80–95, 2017.

[3] N. Andreff and B. Tamadazte, "Laser steering using virtual trifocal visual servoing," *The International Journal of Robotics Research*, vol. 35, no. 6, pp. 672–694, 2016.

[4] J. Tu and L. Zhang, "Effective data-driven calibration for a galvanometric laser scanning system using binocular stereo vision," *Sensors*, vol. 18, no. 1, p. 197, 2018.

[5] W. Huang, C. Gao, Y. Lan, S. Zeng, J. L. Pathak, M. Zhou, L. Ge, and J. Zhang, "Optical coherence tomography characterizes the roughness and thickness of the heterogeneous layer on cortical bone surface induced by er: Yag laser ablation at different moisture contents," *Quantitative imaging in medicine and surgery*, vol. 10, no. 3, p. 713, 2020.

[6] M. B. Tucker, S. Joseph, W. Ross, G. Ma, P. Chongsathidkiet, P. Fecci, and P. Codd, "Creation of a non-contact, automated brain tumor detection device for use in brain tumor resection," in *Clinical and Translational Neurophotonics 2020*, vol. 11225. International Society for Optics and Photonics, 2020, p. 112250C.

[7] Y. Li, F. Richter, J. Lu, E. K. Funk, R. K. Orosco, J. Zhu, and M. C. Yip, "Super: A surgical perception framework for endoscopic tissue manipulation with surgical robotics," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2294–2301, 2020.

[8] J. Wilm, S. H. N. Jensen, and H. Aanæs, "Structured light scanning of skin, muscle and fat," 2015.

[9] W. Hill, "The tumorcnc: Development and evaluation of a first-prototype automated tumor resection device," Ph.D. dissertation, 2016.

[10] D. Gallup, J.-M. Frahm, P. Mordohai, and M. Pollefeys, "Variable baseline/resolution stereo," in *2008 IEEE conference on computer vision and pattern recognition*. IEEE, 2008, pp. 1–8.

[11] S. D. Blostein and T. S. Huang, "Error analysis in stereo determination of 3-d point positions," *IEEE transactions on pattern analysis and machine intelligence*, no. 6, pp. 752–765, 1987.

[12] G. Ma, W. A. Ross, I. Hill, N. Narasimhan, and P. J. Codd, "A novel laser scalpel system for computer-assisted laser surgery," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 386–392.

[13] M. Seeger, "Gaussian processes for machine learning," *International journal of neural systems*, vol. 14, no. 02, pp. 69–106, 2004.

[14] L. Qi, Y. Zhang, S. Wang, Z. Tang, H. Yang, and X. Zhang, "Laser cutting of irregular shape object based on stereo vision laser galvano-metric scanning system," *Optics and Lasers in Engineering*, vol. 68, pp. 180–187, 2015.

[15] S. Sels, B. Bogaerts, S. Vanlanduit, and R. Penne, "Extrinsic calibration of a laser galvanometric setup and a range camera," *Sensors*, vol. 18, no. 5, p. 1478, 2018.

[16] L. N. Smith and M. L. Smith, "Automatic machine vision calibration using statistical and neural network methods," *Image and Vision Computing*, vol. 23, no. 10, pp. 887–899, 2005.

[17] T. Wissel, B. Wagner, P. Stüber, A. Schweikard, and F. Ernst, "Data-driven learning for calibrating galvanometric laser scanners," *IEEE Sensors Journal*, vol. 15, no. 10, pp. 5709–5717, 2015.

[18] M. A. Isa, D. Sims-Waterhouse, S. Piano, and R. Leach, "Volumetric error modelling of a stereo vision system for error correction in photogrammetric three-dimensional coordinate metrology," *Precision Engineering*, vol. 64, pp. 188–199, 2020.

[19] A. Krupa, J. Gangloff, C. Doignon, M. F. De Mathelin, G. Morel, J. Leroy, L. Soler, and J. Marescaux, "Autonomous 3-d positioning of surgical instruments in robotized laparoscopic surgery using visual servoing," *IEEE transactions on robotics and automation*, vol. 19, no. 5, pp. 842–853, 2003.

[20] T. Tayama, Y. Kurose, T. Nitta, K. Harada, Y. Someya, S. Omata, F. Arai, F. Araki, K. Totsuka, T. Ueta, *et al.*, "Image processing for autonomous positioning of eye surgery robot in micro-cannulation," *Procedia CIRP*, vol. 65, pp. 105–109, 2017.

[21] C. Loop and Z. Zhang, "Computing rectifying homographies for stereo vision," in *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, vol. 1. IEEE, 1999, pp. 125–131.

[22] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Sensor fusion IV: control paradigms and data structures*, vol. 1611. International Society for Optics and Photonics, 1992, pp. 586–606.

[23] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[24] H. Hirschmuller, "Stereo processing by semiglobal matching and mutual information," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 30, no. 2, pp. 328–341, 2007.

[25] L. S. Mattos, G. Dagnino, G. Becattini, M. Dellepiane, and D. G. Caldwell, "A virtual scalpel system for computer-assisted laser micro-surgery," in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*. IEEE, 2011, pp. 1359–1365.

[26] K. H. Zou, S. K. Warfield, A. Bharatha, C. M. Tempany, M. R. Kaus, S. J. Haker, W. M. Wells III, F. A. Jolesz, and R. Kikinis, "Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports," *Academic radiology*, vol. 11, no. 2, pp. 178–189, 2004.