JOHNS HOPKINS UNIVERSITY

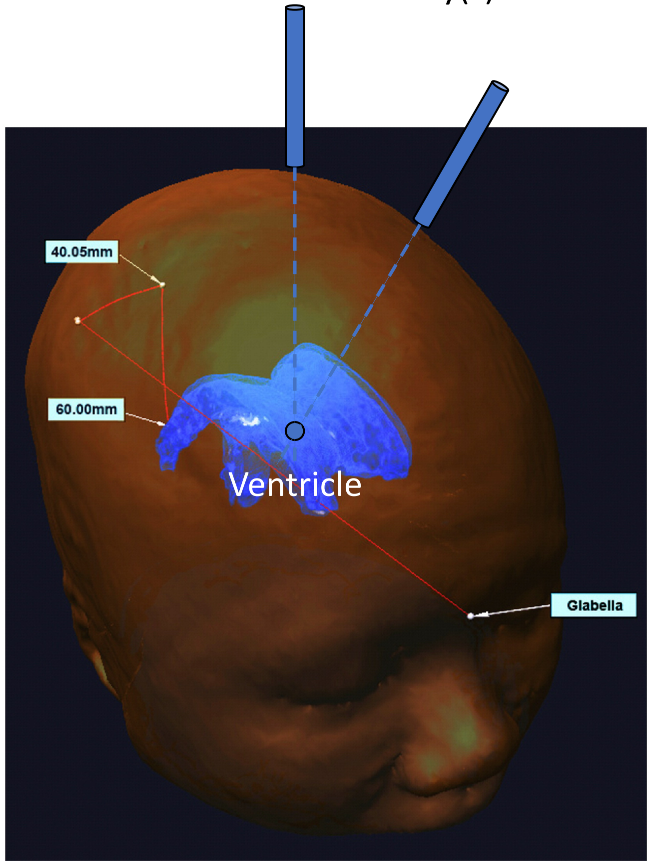
Seminar Presentation: A Fast, Low-cost, Computer Vision Approach for Tracking Surgical Tools

Group 15:
Mingyi Zheng

Project Goal

- The goal is to introduce image guidance via augmented reality on HoloLens

- The image guidance is AR overlay of ventricle model from CT image and catheter guide overlay.



Paper Selection

R. Dockter, R. Sweet and T. Kowalewski,

“A fast, low-cost, computer vision approach for tracking surgical tools,”

2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, Chicago, IL, 2014, pp. 1984-1989. doi: 10.1109/IROS.2014.6942826

- Fast surgical tool tracking algorithm with accurate result
- Applicable to catheter tracking for my project

Paper Objective

Summary of Problem:

Recent developments in training curricula require motion metrics like path length, economy of motion, time, motion smoothness, and response orientation to provide information for discriminating expert from novice surgeon

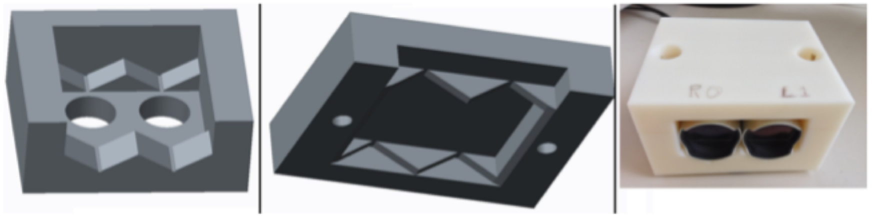
Goal:

- Design a computer vision solution which can fast and accurately track tools in 3D space in real-time with low-cost
- Low-cost and platform independent hardware to increase the chance of adoption in cost-constrained medical training and simulation applications.

Hardware Design

1. Experimental Webcam Unit

- stereo camera mount comprised of two USB Lifecam Studio cameras from Microsoft
- Frame rates of 30 FPS at standard resolutions costs $50 each
- A rapid prototype stereo camera mount
- Interocular distance: 29.1mm



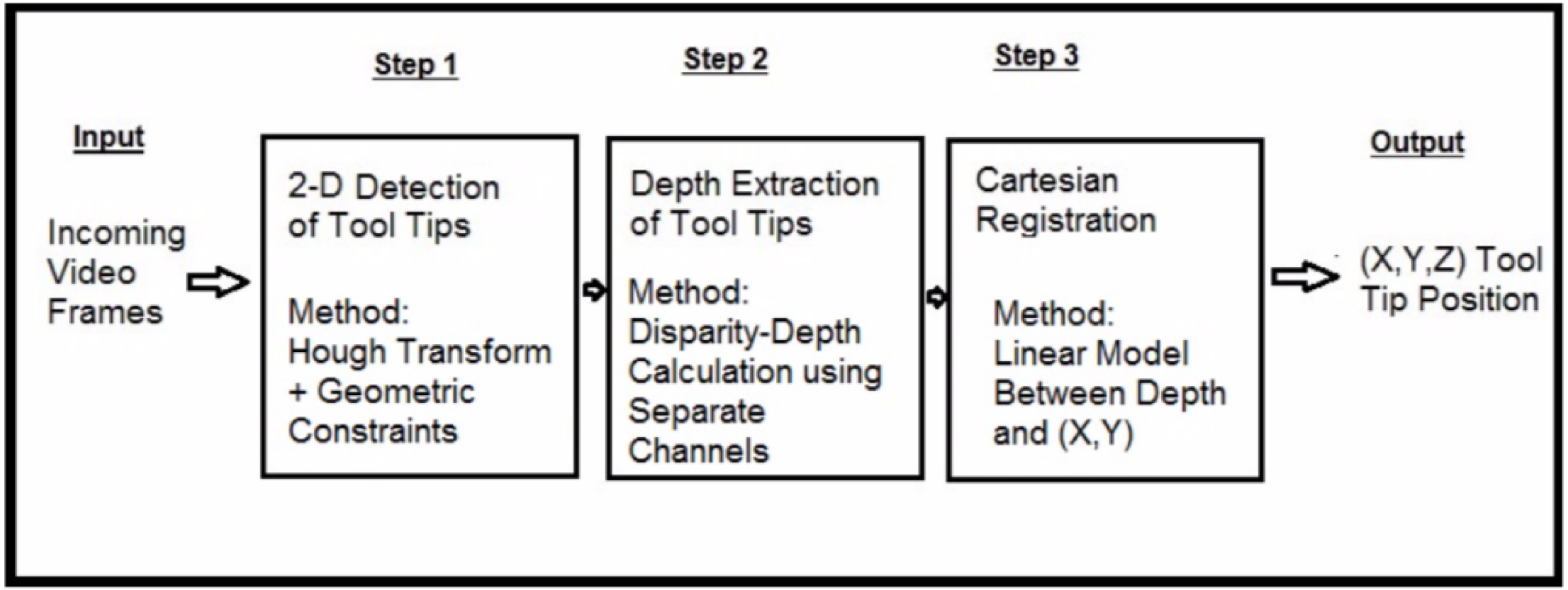
2. Da Vinci Camera Unit(Evaluation)

- The endoscope is capable of resolutions up to 1920x1080 and provides a framerate up to 100 FPS
- Interocular distance: 5.1mm

R. Dockter, R. Sweet and T. Kowalewski, "**A fast, low-cost, computer vision approach for tracking surgical tools**," *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Chicago, IL, 2014, pp. 1984-1989. doi: 10.1109/IROS.2014.6942826

Software and Algorithm Design

3-Step approach for Object Detection



R. Dockter, R. Sweet and T. Kowalewski, "**A fast, low-cost, computer vision approach for tracking surgical tools**," *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Chicago, IL, 2014, pp. 1984-1989. doi: 10.1109/IROS.2014.6942826

Step 1 - 2D Detection

- Dynamic edge gradient threshold
  - Threshold is set by computing the number of edge pixels in the image. If the number of edge pixels is too high (or low), dependent on resolution, the gradient threshold value is then increased (or decreased).

Data: Frame
Result: Cartesian Tool Location
Convert frame to grayscale;
Blur grayscale image;
Sobel edge detection;
Dynamic edge gradient threshold;
Probabilistic hough transform;
for *Each line in PHT line array* **do**
  Extract unique line pair;
  Compute line parameters;
  if *Parameters best match constraints* **then**
    Save line;
  else
    Extract 2 new lines;
  end
end
Calculate endpoints closest to prior location;
Save tool tip location;

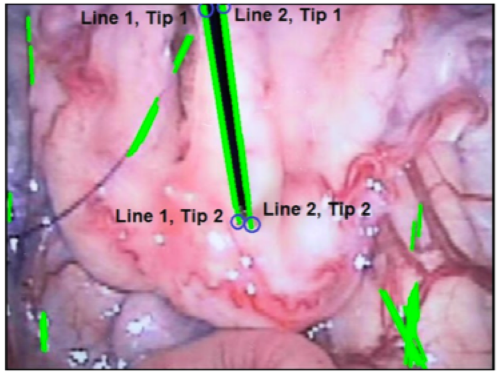
R. Dockter, R. Sweet and T. Kowalewski, "**A fast, low-cost, computer vision approach for tracking surgical tools**," *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*,

Step 1 – PHT + Constraint

- Probabilistic Hough Transform(PHT)
  - Improve frame rate
  - Endpoints of line is determined
  - Limits the loss of lines due to occlusions along the shaft
- Geometric Constraint - Endpoint
  - For initialization, Tool endpoint is closest to the center of image
  - After initialization, endpoints is closest to last known location
  - Midpoint of the two endpoints is wrist
  - If multiple tools, the line endpoint is sorted according to previous constraints and separated according to prior spatial information and tool shaft angle

Constraint	Condition
Lines along the tool shaft are parallel	$\Delta\theta < \theta_{threshold}$
Endpoints of the lines should be near to each other.	$d(p(i)_{1,1}, p(i)_{2,1}) < d_{threshold}$
Length of two lines should be longer than any other set of parallel lines	$d(p(i)_{1,1}, p(i)_{1,2}) > length_{max}$
The resultant endpoints should be 'near' the last known location.	$d(p(i)_{1,1}, p(i-1)_{1,1}) < d_{near}$

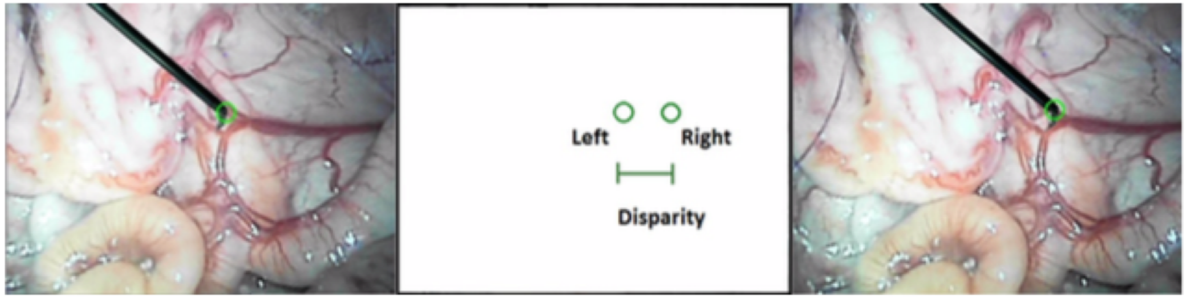
TABLE I. Geometric constraints and their implementation in code. d() represents the euclidean distance formula and $p(k)_{m,n}$ represents a point (x,y) within frame k, line m, index n



R. Dockter, R. Sweet and T. Kowalewski, "**A fast, low-cost, computer vision approach for tracking surgical tools**," *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*,

Step-2 Depth Extraction

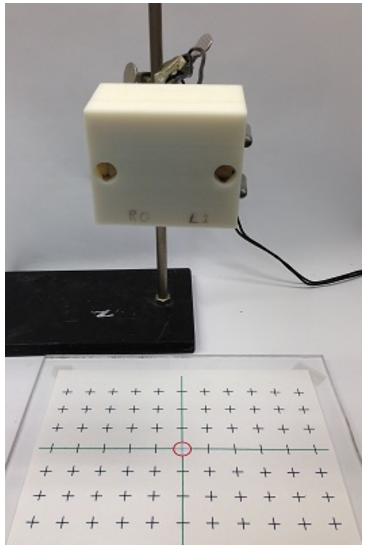
- A full stereo correspondence is computational expensive
- Single tool tip disparity calculation is simpler and used
- Disparity is calculated by Euclidan distance formula
- Disparity is then used to determine the depth to the tool tip and Cartesian coordinates

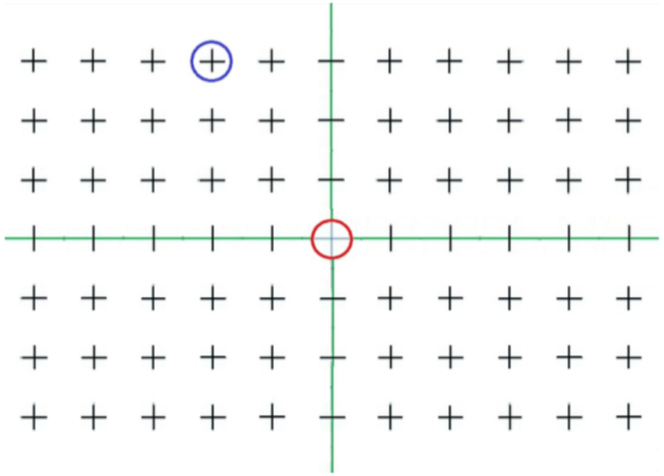


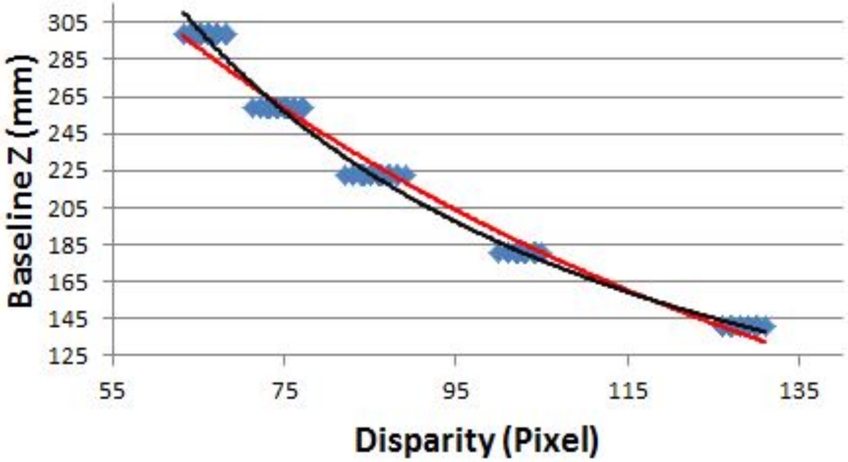
R. Dockter, R. Sweet and T. Kowalewski, "**A fast, low-cost, computer vision approach for tracking surgical tools**," *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Chicago, IL, 2014, pp. 1984-1989. doi: 10.1109/IROS.2014.6942826

Step-2 Depth Extraction

- Rather than standard linear transformation model to calculate depth from disparity, the depth model is determined by non-linear curve fitting with calibration data to have minimal drop in accuracy



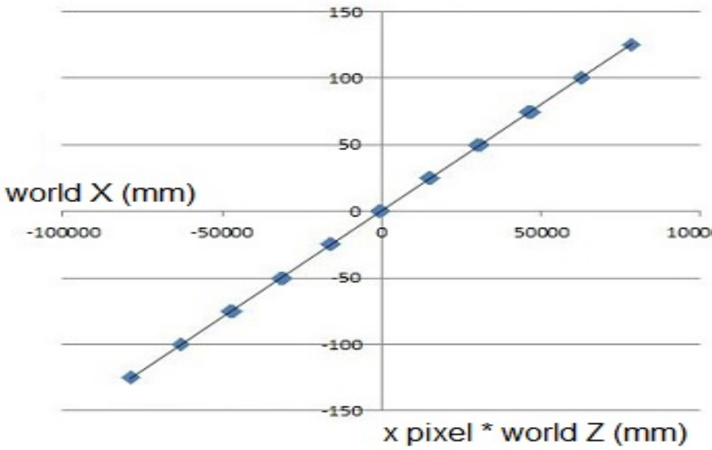


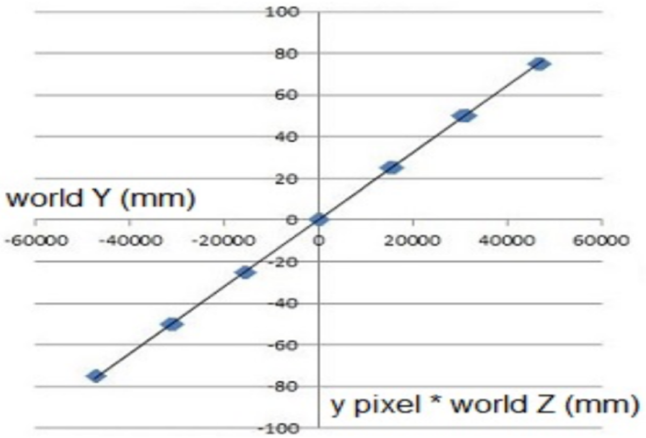


R. Dockter, R. Sweet and T. Kowalewski, "**A fast, low-cost, computer vision approach for tracking surgical tools**," *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Chicago, IL, 2014, pp. 1984-1989. doi: 10.1109/IROS.2014.6942826

Step-3 Cartesian Registration

- Mapping equation from (x,y,disparity) values into (Xw,Yw,Zw) global coordinates.
- A linear fit is applied to the plot of xp ∗ Zw vs Xw/Yw.
- Separate calibrations are required for experimental camera setup and the da Vinci Endoscope.





Webcam setup:

$$Z_w = a_1(disp^{a_2}) + a_3 x_p + a_4 y_p + a_5 \tag{1}$$

Endoscope:

$$disp_o = \begin{bmatrix} b_0 & \dots & b_n \end{bmatrix} * \begin{bmatrix} 1 & x & y & x^2 & xy & y^2 & x^3 & x^2y & xy^2 & y^3 \end{bmatrix}^T \tag{2}$$

$$Z_w = c_0 e^{c_1(disp - disp_o)} \tag{3}$$

(a_i,b_i,c_i) : coefficients determined using linear regression fit

(x_p,y_p): X Y Pixel location

$disp_0$ = disparity offset for endoscope

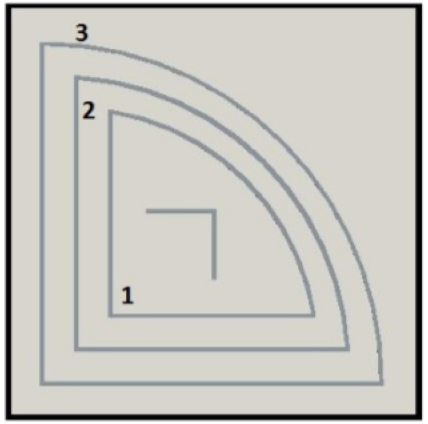
R. Dockter, R. Sweet and T. Kowalewski, "**A fast, low-cost, computer vision approach for tracking surgical tools**," *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Chicago, IL, 2014, pp. 1984-1989. doi: 10.1109/IROS.2014.6942826

Experimental Design

- Time:

$$T_{tracking}(ms) = Max(T_{detect,right}, T_{detect,left}) + T_{3DLocate}$$

- Accuracy:
  - moving the tool tip around a known fixed trajectory
  - compute the error between world coordinates and world trajectory
- Noise:
  - Calculate the deviation of a fixed surgical tool



R. Dockter, R. Sweet and T. Kowalewski, "**A fast, low-cost, computer vision approach for tracking surgical tools**," *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Chicago, IL, 2014, pp. 1984-1989. doi: 10.1109/IROS.2014.6942826

Result

- Endoscope has larger tracking error overall than Webcam, but it is less than tool shaft diameter
- Webcam setup achieves 99.4% for tool localization
- Background and lighting invariant

Performance Metric	Webcam	Endoscope
Computation Time (ms)	39.9	33.99
Frame Rate (FPS)	25.86	26.98
Depth Reprojection Error (mm)	4.09	7.89
Localization Noise (Total) (mm)	1.29	7.92
Average 3D Error (mm)	3.05	8.68
Average 2D Error (mm)	1.59	1.88
95^{th} Percentile Error (mm)	5.47	15.06
Percent Within 8 mm	99.4 %	55.22 %

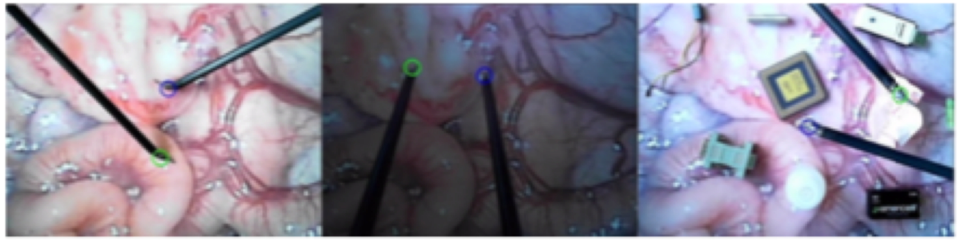


Fig. 8. Tool tracking configurations using the webcam setup (Left Channel).

R. Dockter, R. Sweet and T. Kowalewski, "**A fast, low-cost, computer vision approach for tracking surgical tools**," *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Chicago, IL, 2014, pp. 1984-1989. doi: 10.1109/IROS.2014.6942826

Conclusion

- Tool tracking algorithm is invariant to end-effector type and lighting since it only tracks tool shaft and is independent on color.
- For the endoscopic camera setup, the depth extraction model is responsible for most of the error due to narrow interocular separation as well as the uncertainties in camera optics.
- The range of disparities found in this camera setup was found to be about 10 pixels for a depth range of 200 mm. This results in a high signal to noise ratio.
- Future work
  - Geometric computer vision methods for 3D coordinate extraction,
  - Epipolar geometry to constrain search space for 2nd stereo channel,
  - Incorporation to real surgical application

R. Dockter, R. Sweet and T. Kowalewski, "**A fast, low-cost, computer vision approach for tracking surgical tools**," *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Chicago, IL, 2014, pp. 1984-1989. doi: 10.1109/IROS.2014.6942826

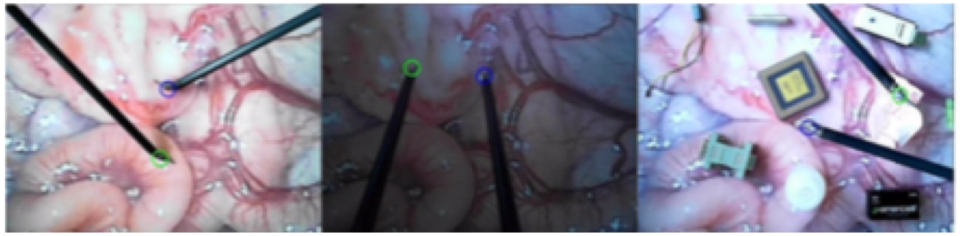
Assessment

Positive:

- Provide a fast and accurate tool tracking algorithm
- Clear description for 2D object detection
- Explain the reason for choices (PHT)

Negative:

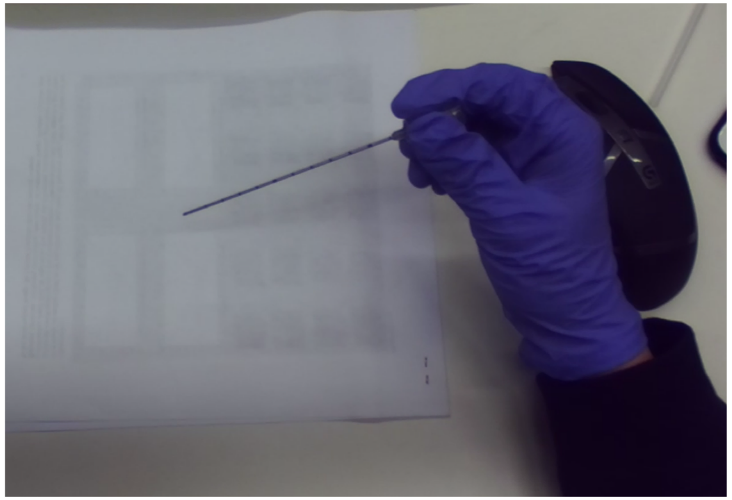
- Lack of explanation for depth extraction model and cartesian registration
- Unclear experimental setup
  - Distance of Webcam to 3d print board
  - Setup for test with real surgical background
  - Setup for Endoscope test

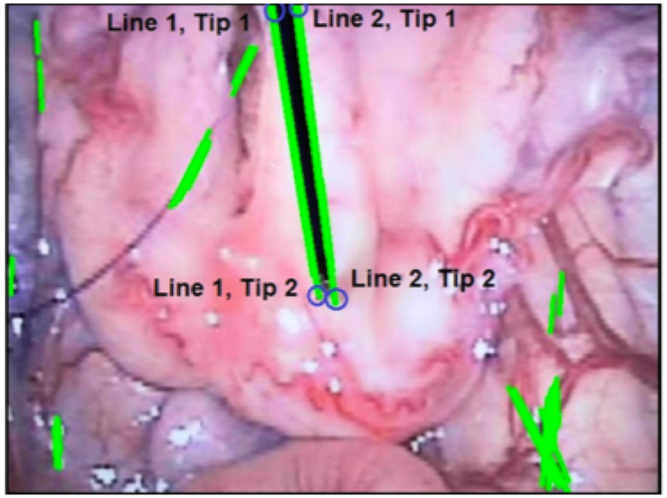


R. Dockter, R. Sweet and T. Kowalewski, "**A fast, low-cost, computer vision approach for tracking surgical tools**," *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Chicago, IL, 2014, pp. 1984-1989. doi: 10.1109/IROS.2014.6942826

Relevance

- Tools look similar
- The tracking algorithm (PHT + Geometric Constraint) is applicable
- Experiment setup is useful for catheter tracking evaluation





References

R. Dockter, R. Sweet and T. Kowalewski, "**A fast, low-cost, computer vision approach for tracking surgical tools**," *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Chicago, IL, 2014, pp. 1984-1989. doi: 10.1109/IROS.2014.6942826

Thank You!

Any Questions?