Seminar Presentation

3D Tool Tracking in the Presence of Microscope Motion

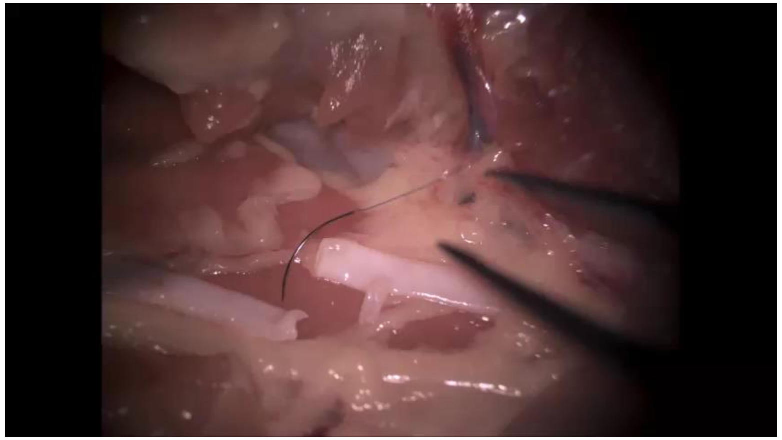
Group 4:
Molly O'Brien
PhD Student
mobrie38@jhu.edu

Mentors:
Dr. Austin Reiter,
Dr. Russell Taylor

3D Tool Tracking in the Presence of Microscope Motion

- Problem:

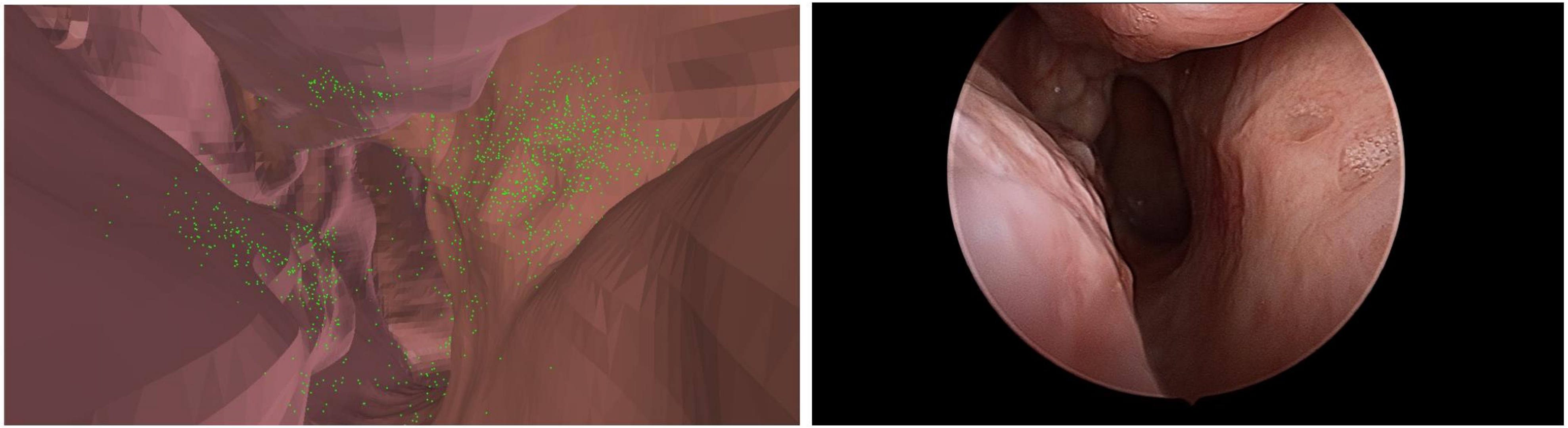
https://www.youtube.com/watch?v=HXTEFoFJ9iA&t=617s



- Goal:

Paper Selection and why

1. S. Leonard, A. Reiter, A. Sinha, M. Ishii, R. Taylor, and G. Hager, “Image-Based Navigation for Functional Endoscopic Sinus Surgery Using Structure From Motion,” in *SPIE*, San Diego, 2016.
2. B. Allen, F. Kasper, G. Nataneli, E. Dutson, and P. Faloutos, “Visual Tracking of Laparoscopic Instruments in Standard Training Environments,” in *MMVR,* Newport Beach 2011.



[1]

(b) Point cloud registration with non-erectile tissues (nasopharynx).

Figure 6: Registration of two point clouds to CT scans.

S. Leonard, A. Reiter, A. Sinha, M. Ishii, R. Taylor, and G. Hager

“IMAGE-BASED NAVIGATION FOR FUNCTIONAL ENDOSCOPIC SINUS SURGERY USING STRUCTURE FROM MOTION”

Paper 1

- **Problem:**
  - Functional Endoscopic Sinus Surgery (FESS)
    - Used to treat chronic sinusitis, remove polyps, open up passageways
    - Near critical structures like nerves and arteries
  - Navigation in FESS is difficult
    - Register endoscope video to CT
    - Current systems can only get 1mm accuracy
- **Goal:**

1. Register endoscope video to preoperative CT
2. Overlay areas of interest on the endoscope video

https://www.youtube.com/watch?v=lUgrdIOVMUg

S. Leonard, A. Reiter, A. Sinha, M. Ishii, R. Taylor, and G. Hager, "Image-Based Navigation for Functional Endoscopic Sinus Surgery Using Structure From Motion," in *SPIE*, San Diego, 2016.

Key Result and Significance

- Key Result:

Registration Error	
Erectile Tissue	Non-Erectile Tissue
1.21 mm	0.91 mm

- Significance of Key result
  - Sub-millimeter registration results

S. Leonard, A. Reiter, A. Sinha, M. Ishii, R. Taylor, and G. Hager, "Image-Based Navigation for Functional Endoscopic Sinus Surgery Using Structure From Motion," in *SPIE*, San Diego, 2016.

Background

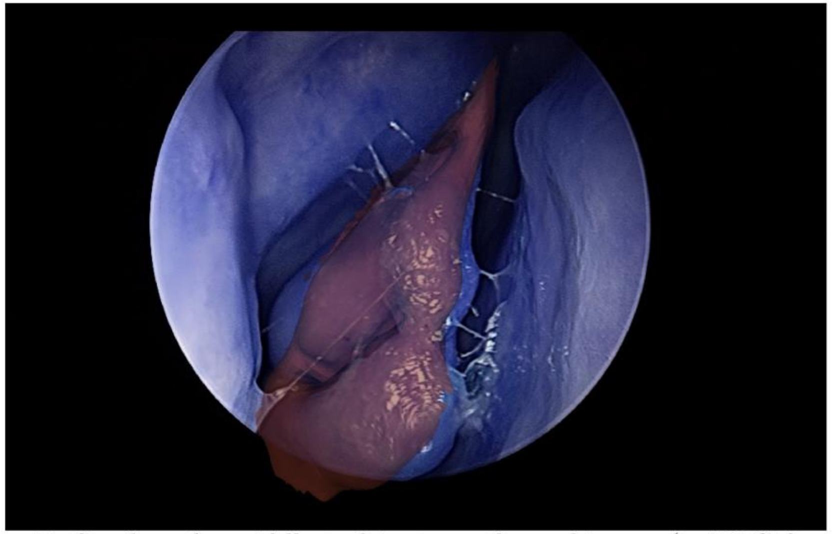
- Functional Endoscopic Sinus Surgery (FESS)
  - Removes cartilage and tissue from nasal passageway
  - Near critical structures like optic nerves, arteries
  - Complications include: cerebrospinal fluid leaks, blindness, difficulty controlling eye movement, excessive bleeding

https://www.youtube.com/watch?v=lUgrdIOVMUg

S. Leonard, A. Reiter, A. Sinha, M. Ishii, R. Taylor, and G. Hager, "Image-Based Navigation for Functional Endoscopic Sinus Surgery Using Structure From Motion," in *SPIE*, San Diego, 2016.

What did they do?

1. Generate 3D point cloud from endoscope video using Structure from Motion and Bundle Adjustment
2. Register to CT scan using ICP
3. Overlay areas of interest segmented in the CT on the endoscope image



[1]

Figure 7: Overlay of a middle turbinate on the real image (in BRG format).

S. Leonard, A. Reiter, A. Sinha, M. Ishii, R. Taylor, and G. Hager, "Image-Based Navigation for Functional Endoscopic Sinus Surgery Using Structure From Motion," in *SPIE*, San Diego, 2016.

Structure from Motion and Bundle Adjustment

- Generate a Point Cloud
  - HMA (Hierarchical Multi-Affine) algorithm with SURF features to get accurate feature matching
  - Match points between each pair of endoscope images
  - Use multiple endoscope frames (as few as 15) to get a registration
- Use point cloud to estimate 3D structure
- Bundle Adjustment to estimate camera motion
- 6 DOF magnetic tracker used to find ambiguous scale factor in camera motion and 3D structure

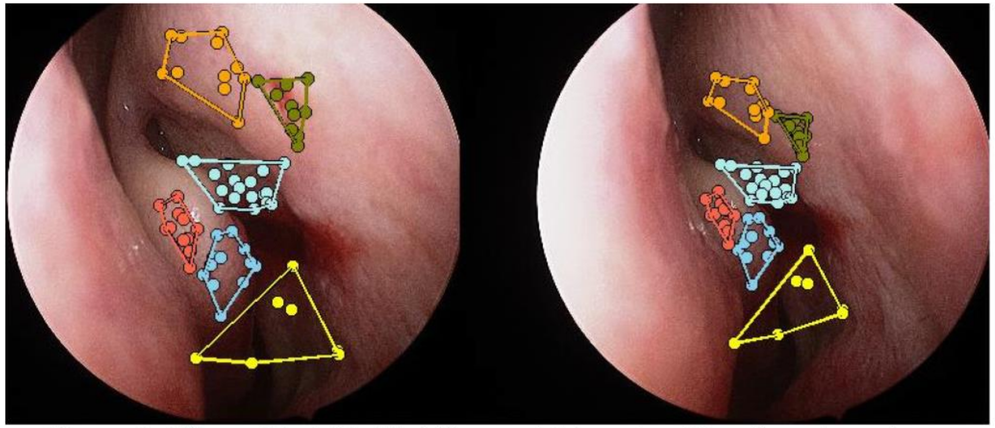


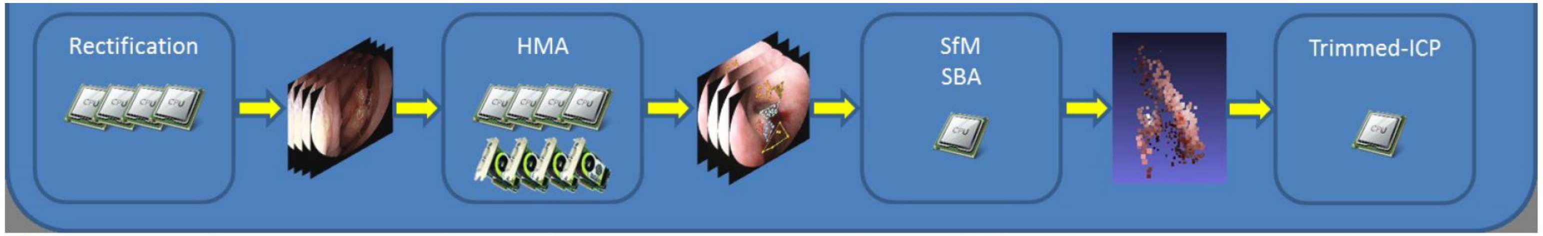
Figure 2: Hierarchical Multi-Affine matching algorithm from two views.

[1]

S. Leonard, A. Reiter, A. Sinha, M. Ishii, R. Taylor, and G. Hager, "Image-Based Navigation for Functional Endoscopic Sinus Surgery Using Structure From Motion," in *SPIE*, San Diego, 2016.

Register Point Cloud to CT Scan

- Use trimmed ICP (Iterative Closest Point) to register 3D point cloud to CT scan
  - Trimmed so it is robust to outliers
- Can solve for scale factor



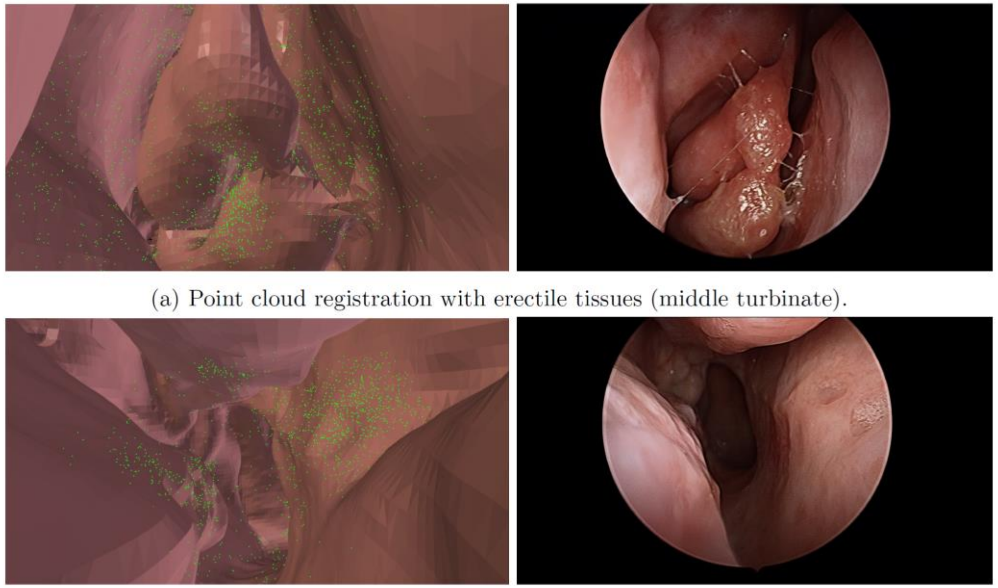
[1]

Figure 1: Block diagram of our image-based navigation system. The images are sent to a computing server to compute the structure from motion and to register the result to the patients CT scan.

S. Leonard, A. Reiter, A. Sinha, M. Ishii, R. Taylor, and G. Hager, "Image-Based Navigation for Functional Endoscopic Sinus Surgery Using Structure From Motion," in *SPIE*, San Diego, 2016.

Experimental Setup

- 90s data collection
- Endoscope inserted in both airways
- Some cases had significant congestion differences between the CT and endoscope images
- Initial pose estimate for trimmed ICP given manually



(a) Point cloud registration with erectile tissues (middle turbinate).

(b) Point cloud registration with non-erectile tissues (nasopharynx).

Figure 6: Registration of two point clouds to CT scans.

[1]

S. Leonard, A. Reiter, A. Sinha, M. Ishii, R. Taylor, and G. Hager, "Image-Based Navigation for Functional Endoscopic Sinus Surgery Using Structure From Motion," in *SPIE*, San Diego, 2016.

Results

- Registration Accuracy
  - No ground truth
  - TRE (error between CT surface and 3D point cloud)

Table 1: Average 70th percentile registration error

	Non-Erectile Tissues	Erectile Tissues
$\bar{e}_{ICP}$	0.91 mm (0.2 mm)	1.21 mm (0.3 mm)

[1]

- Camera Pose error:
  - No ground truth
  - Segmented structure (middle turbinate) overlap of 86%

S. Leonard, A. Reiter, A. Sinha, M. Ishii, R. Taylor, and G. Hager, "Image-Based Navigation for Functional Endoscopic Sinus Surgery Using Structure From Motion," in *SPIE*, San Diego, 2016.

Conclusion

- Achieve state of the art registration accuracy

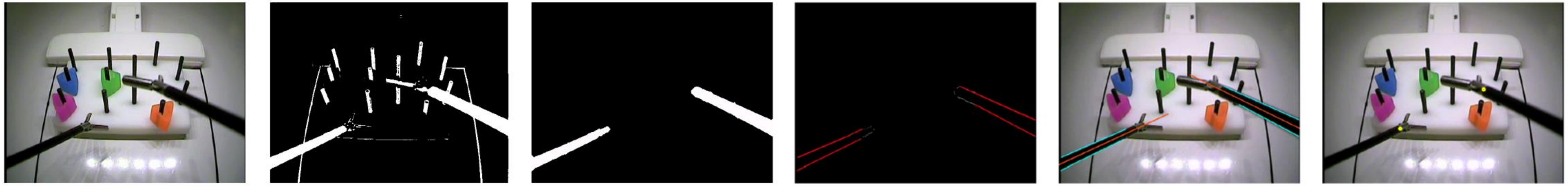
- Future work
  - Clinical validation
  - Improve robustness to initial guess, erectile tissue

S. Leonard, A. Reiter, A. Sinha, M. Ishii, R. Taylor, and G. Hager, "Image-Based Navigation for Functional Endoscopic Sinus Surgery Using Structure From Motion," in *SPIE*, San Diego, 2016.

My Assessment

- Punchline:
  - Endoscope: triangulate points like stereo by using multiple views
  - My problem is easier bc have stereo, I can try an easier approach like ICP first
- Weakness
  - True registration error unclear:
    - How good are your triangulated points?
- Relevance
  - Triangulating 3D point cloud
  - May do bundle adjustment to find camera movement between frames

S. Leonard, A. Reiter, A. Sinha, M. Ishii, R. Taylor, and G. Hager, “Image-Based Navigation for Functional Endoscopic Sinus Surgery Using Structure From Motion,” in *SPIE*, San Diego, 2016.



(a) Unmodified frame from the FLS camera during a training task. (b) Binary probability map of black regions. (c) Binary mask with isolated instruments. (d) Extracted lateral contours of instruments. (e) Instrument direction estimated using line-fitting. (f) Tracked position in 2D.

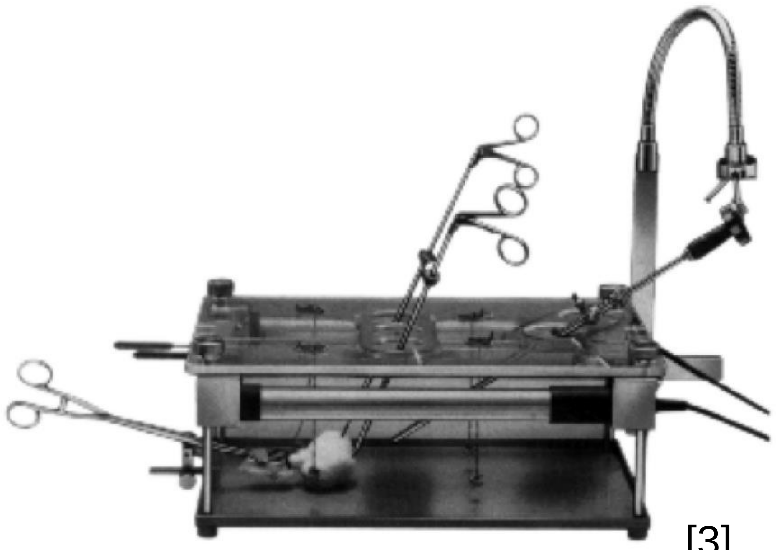
[3]

B. Allen, F. Kasper, G. Nataneli, E. Dutson, and P. Faloutos

"VISUAL TRACKING OF LAPAROSCOPIC INSTRUMENTS IN STANDARD TRAINING ENVIRONMENTS"

"Visual Tracking of Laparoscopic Instruments in Standard Training Environments"

- Problem:
  - Laparoscopic surgery is hard because of physical constraints, want a way to accurately gauge surgeon skill
  - Tracking tool movement better than just metrics like number tools dropped, can measure path length etc.
- Goal:
  - Track motion of laparoscopic tools from video



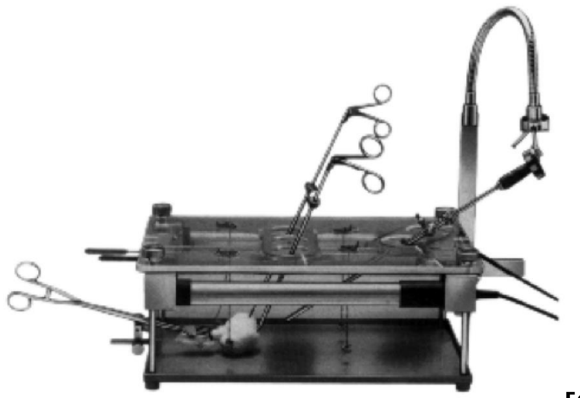
[3]

(b) A standard FLS box trainer.

B. Allen, F. Kasper, G. Nataneli, E. Dutson, and P. Faloutos, "Visual Tracking of Laparoscopic Instruments in Standard Training Environments," in *MMVR,* Newport Beach 2011.

Paper 2

- Key Result:
  - Tracking system in FLS (Fundamentals of Laparoscopic Surgery) tool box
    - Tool shaft direction
    - Tool tip
    - Trocar position
    - Camera position



(b) A standard FLS box trainer.

[3]

- Significance of Key result
  - Quantitative measure of laparoscopic surgical skill
  - Tool tracking alg

B. Allen, F. Kasper, G. Nataneli, E. Dutson, and P. Faloutos, “Visual Tracking of Laparoscopic Instruments in Standard Training Environments,” in *MMVR,* Newport Beach 2011.

Background: Fundamentals of Laparoscopic Surgery Toolbox

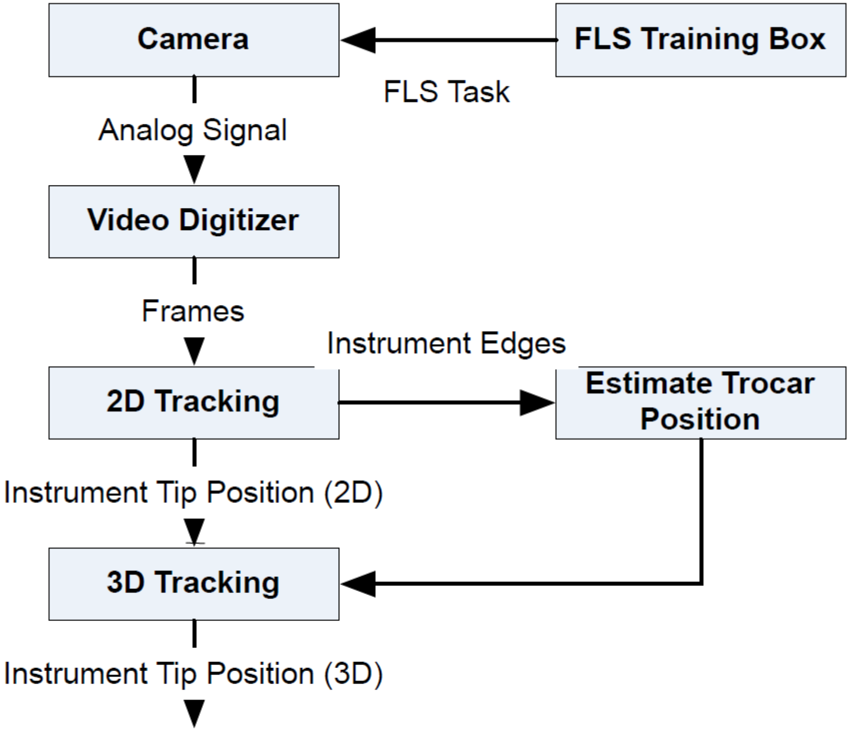
- Laparoscopic surgery requires specialized movements for tight spaces
- "training surgeons have little ability to self-assess" [2]
- Current tool tracking systems:
  - Magnetic
  - Mechanical links
  - Virtual Reality environment (use a joystick)
- Drawbacks:
  - Physical attachments
  - Unrealistic movements

[4] https://www.youtube.com/watch?v=ROUGZ79Paxk

B. Allen, F. Kasper, G. Nataneli, E. Dutson, and P. Faloutos, "Visual Tracking of Laparoscopic Instruments in Standard Training Environments," in *MMVR,* Newport Beach 2011.

What did they do?

1. Extract tool contours from color image
2. Estimate direction of tool shaft
3. Search to find most probable tool tip position



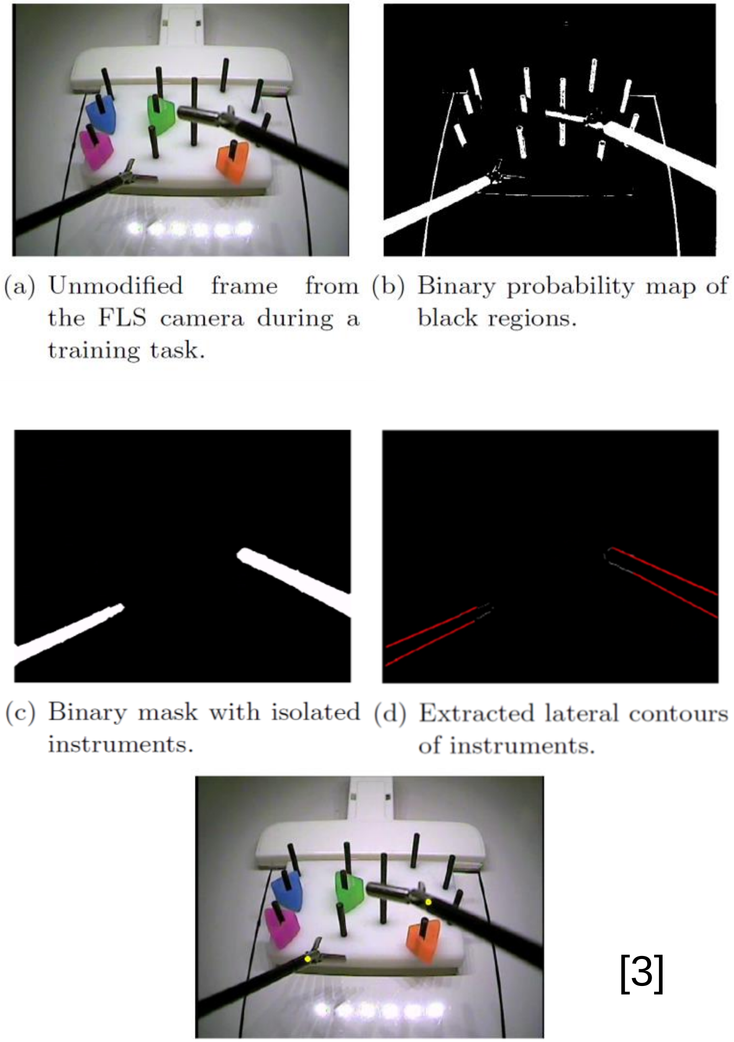
(a) Overview of the process showing data flow. [3]

B. Allen, F. Kasper, G. Nataneli, E. Dutson, and P. Faloutos, "Visual Tracking of Laparoscopic Instruments in Standard Training Environments," in *MMVR,* Newport Beach 2011.

Extract tool contours, tool tip from color image

- Color threshold to get black parts of image (b)
- Erosion and dilation to get only tool shaft (c)
- Search along shaft for abrupt change at tip
  - Color space
  - Gradient space
  - Confidence measure

$$P(T_G) = 1 - \frac{\|T_G - T_C\|}{\beta}$$



(a) Unmodified frame from the FLS camera during a training task. (b) Binary probability map of black regions.

(c) Binary mask with isolated instruments. (d) Extracted lateral contours of instruments.

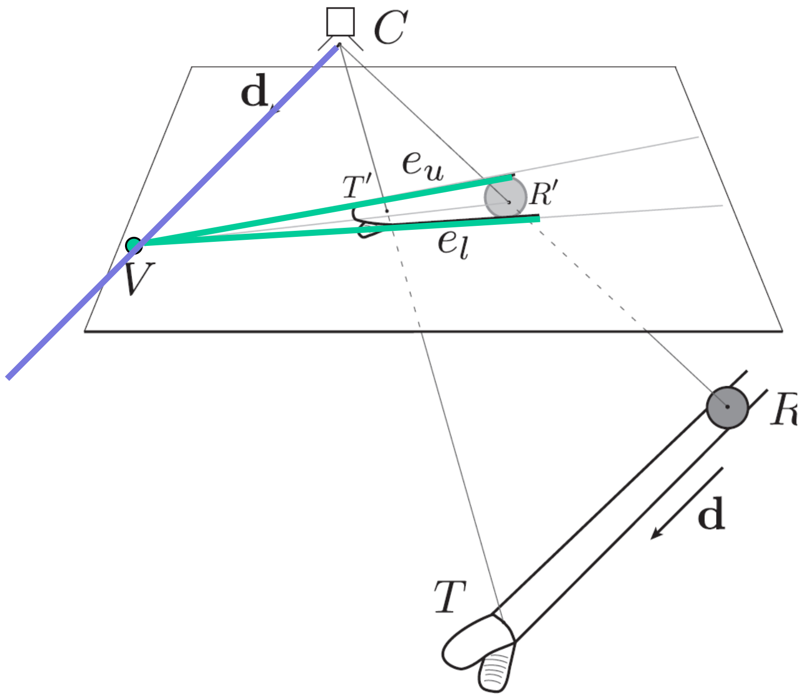
(f) Tracked position in 2D.

[3]

B. Allen, F. Kasper, G. Nataneli, E. Dutson, and P. Faloutos, “Visual Tracking of Laparoscopic Instruments in Standard Training Environments,” in *MMVR,* Newport Beach 2011.

Estimate direction of tool shaft

- Find edges of tool, vanishing point in the image
- Lines parallel to the tool shaft in the world will pass through the vanishing point in the image
- d: direction of the tool in the world
- $d = \frac{V}{\|V\|}$

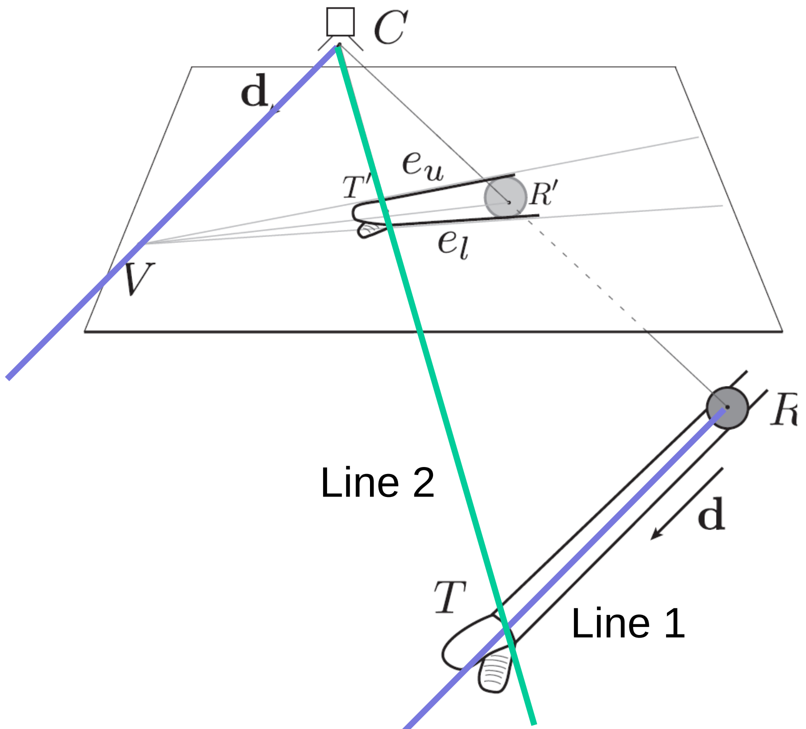


(g) The geometry of the image formation of the instrument. [3]

B. Allen, F. Kasper, G. Nataneli, E. Dutson, and P. Faloutos, “Visual Tracking of Laparoscopic Instruments in Standard Training Environments,” in *MMVR,* Newport Beach 2011.

Search to find most probable tool tip position

- Tool tip in the world is the intersection of Line 1 and Line 2
  - If they don't intersect, find midpoint of shortest line between them

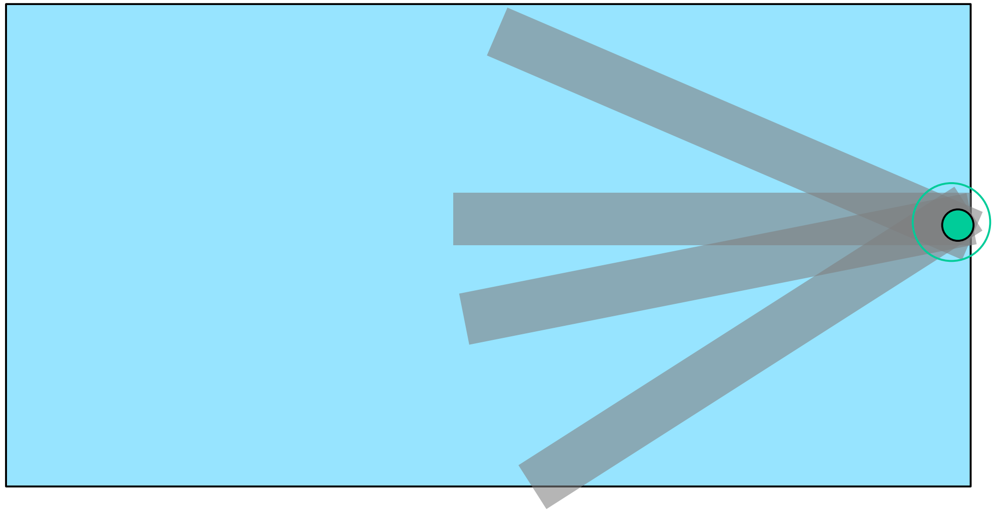


(g) The geometry of the image formation of the instrument. [3]

B. Allen, F. Kasper, G. Nataneli, E. Dutson, and P. Faloutos, "Visual Tracking of Laparoscopic Instruments in Standard Training Environments," in *MMVR,* Newport Beach 2011.

Locating the trocar

- an image point R0 is found as the point closest to all of the tool edge points



http://www.medtronic.com/content/dam/covidien/library/global/en/product/trocars-and-access/versaport-v2-trocar-f1.jpg

B. Allen, F. Kasper, G. Nataneli, E. Dutson, and P. Faloutos, "Visual Tracking of Laparoscopic Instruments in Standard Training Environments," in *MMVR,* Newport Beach 2011.

Conclusion

- Get accurate tool tip tracking, 3D tool tip position from one camera
- Confidence measure lets them stop tracking if tool is occluded or tool close to camera (blurry)

B. Allen, F. Kasper, G. Nataneli, E. Dutson, and P. Faloutos, "Visual Tracking of Laparoscopic Instruments in Standard Training Environments," in *MMVR,* Newport Beach 2011.

Conclusion

- Get accurate tool tip tracking, 3D tool tip position from one camera
- Confidence measure lets them stop tracking if tool is occluded or tool close to camera (blurry)

B. Allen, F. Kasper, G. Nataneli, E. Dutson, and P. Faloutos, "Visual Tracking of Laparoscopic Instruments in Standard Training Environments," in *MMVR,* Newport Beach 2011.

My Assessment

- Interesting: get 3D geometry from 1 image, but we don't know how accurate it is
  - "validated visually"
- Weakness
  - Only specific training setting (they say this upfront), doesn't address problems in real surgical data like blood, water, specularity on tissue
  - Fixed camera (different from me)
  - No quantitative evaluation of results "validated visually"
- Possible next steps for this work
  - Apply tracking algorithm to real surgical data

B. Allen, F. Kasper, G. Nataneli, E. Dutson, and P. Faloutos, "Visual Tracking of Laparoscopic Instruments in Standard Training Environments," in *MMVR,* Newport Beach 2011.

References

[1] S. Leonard, A. Reiter, A. Sinha, M. Ishii, R. Taylor, and G. Hager, “Image-Based Navigation for Functional Endoscopic Sinus Surgery Using Structure From Motion,” in *SPIE*, San Diego, 2016.

[2] https://www.youtube.com/watch?v=lUgrdIOVMUg

[3] B. Allen, F. Kasper, G. Nataneli, E. Dutson, and P. Faloutos, “Visual Tracking of Laparoscopic Instruments in Standard Training Environments,” in *MMVR,* Newport Beach 2011.

[4] https://www.youtube.com/watch?v=ROUGZ79Paxk

[5]http://www.medtronic.com/content/dam/covidien/library/global/en/product/trocars-and-access/versaport-v2-trocar-f1.jpg