Image-Based Navigation for Functional Endoscopic Sinus Surgery Using Structure From Motion

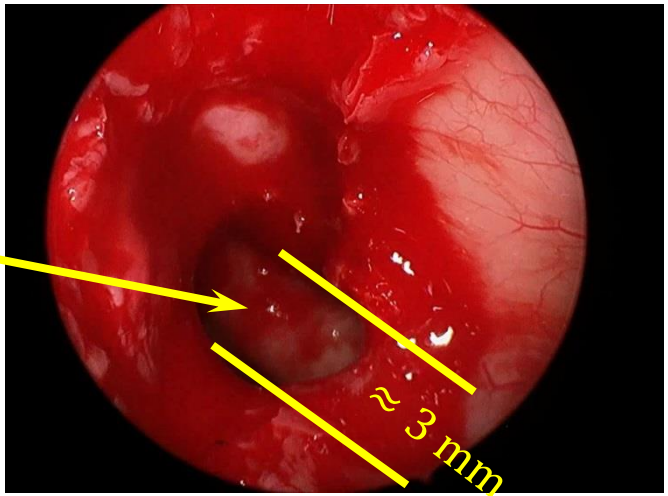
Simon Leonard, Austin Reiter, **Ayushi Sinha**, Masaru Ishii, Russell H. Taylor and Gregory D. Hager

The Johns Hopkins University

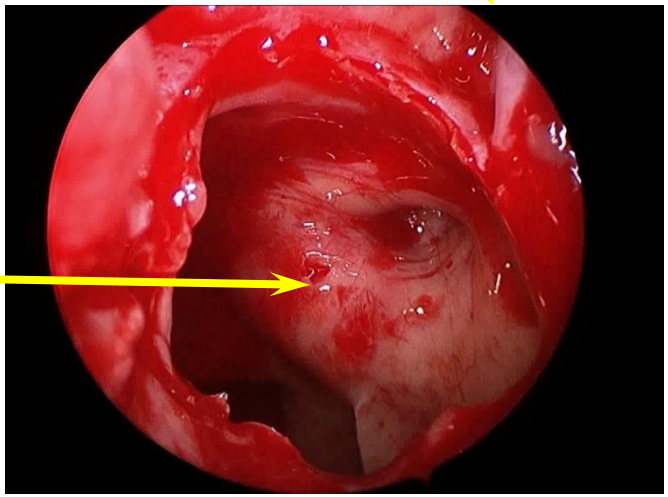
Introduction

- Functional Endoscopic Sinus Surgery (FESS) is a challenging procedure for otolaryngologists
- Over 250,000 FESS are performed annually in the USA
- Use to treat common conditions such as chronic sinusitis
- Surgeons remove several layers of cartilage and tissues that are within millimeters of critical anatomical structures (nerves, arteries, ducts)

Partially exposed artery

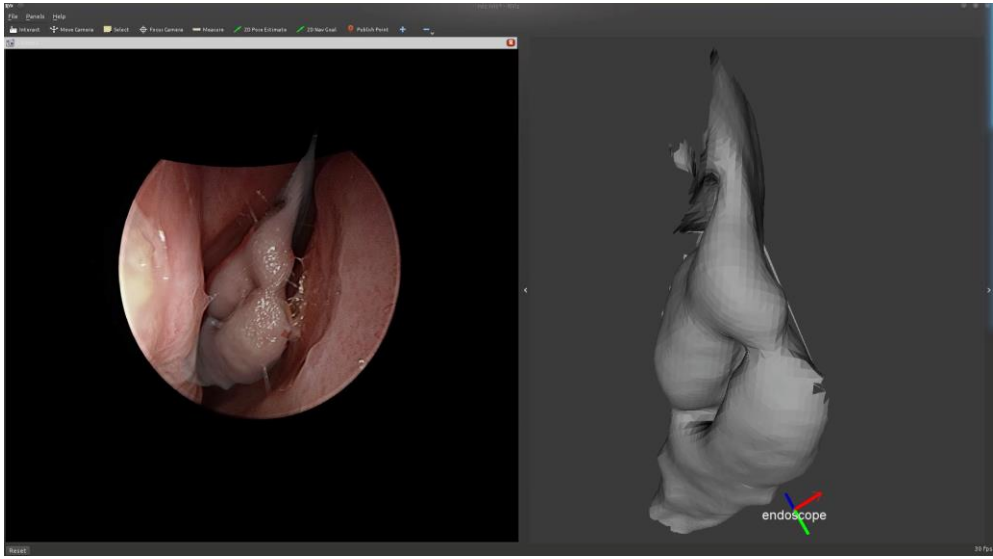


Same exposed artery



Motivation

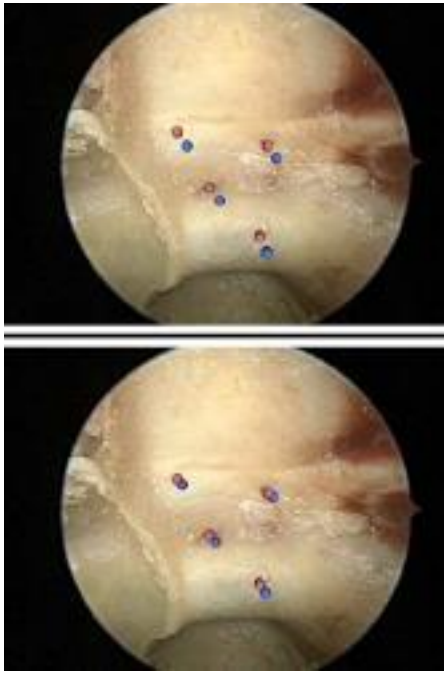
- Complications of FESS
  - Major: 0.31-0.47%
  - Minor: 1.37-5.6%
- Safety and efficiency are improved by using navigation systems
- State of the art navigation systems have reported accuracy greater than 1 mm which is large considering the scale of the sinus cavities



Overlay of middle turbinate (CT) on video images

Previous Work

- Previous Video-CT registration used pairs of cadaver images
  - Using needles as fiducials
  - Using image pairs
  - Reprojection distance error of 1.28 mm (1.82 mm for tracker)



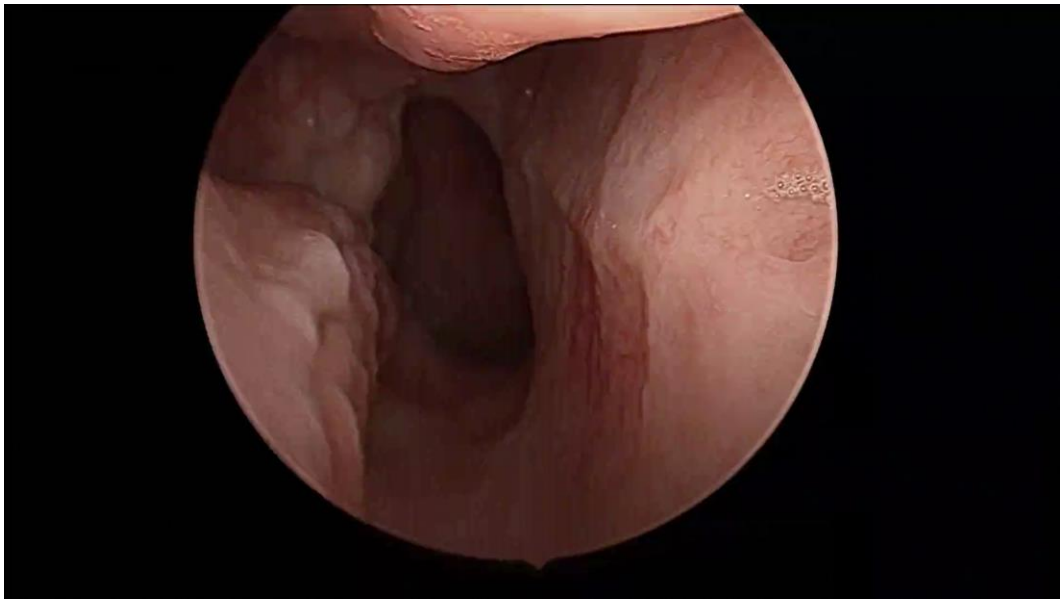
Tracker-based

Image-based

Reprojection error between tracker based and image-based Methods
Mirota IEEE TMI 2013

Objectives

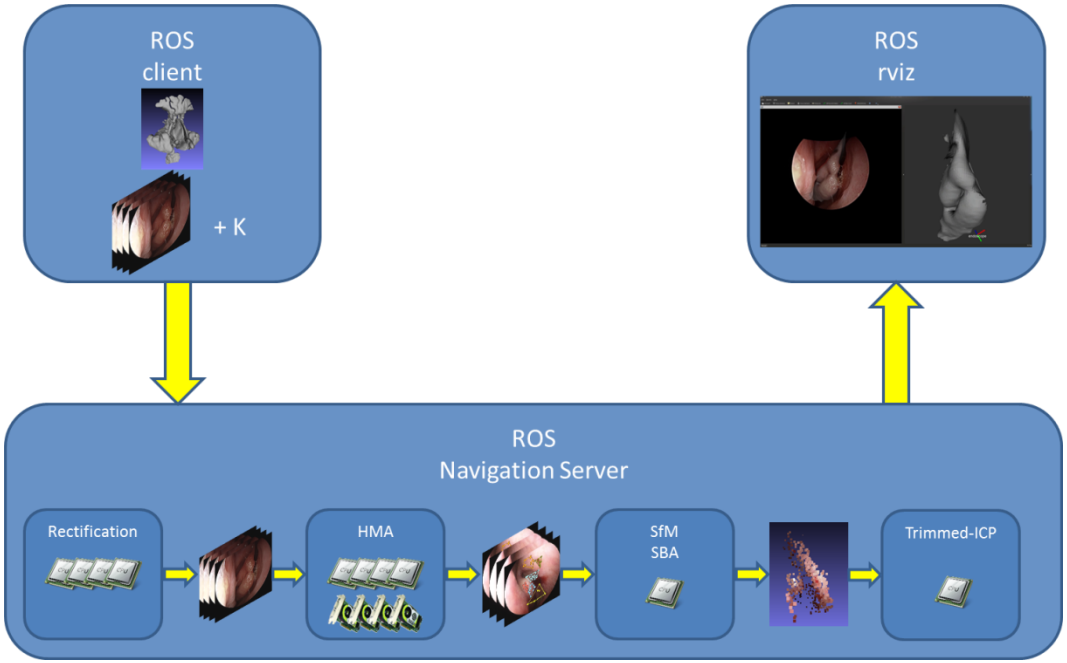
1. Test video-CT registration with in-vivo data
   - Test registration for erectile tissues
   - Less “feature rich” images
2. Use a greater set of video images for registration
3. Manageable computation time



Sample video sequence (1 second)

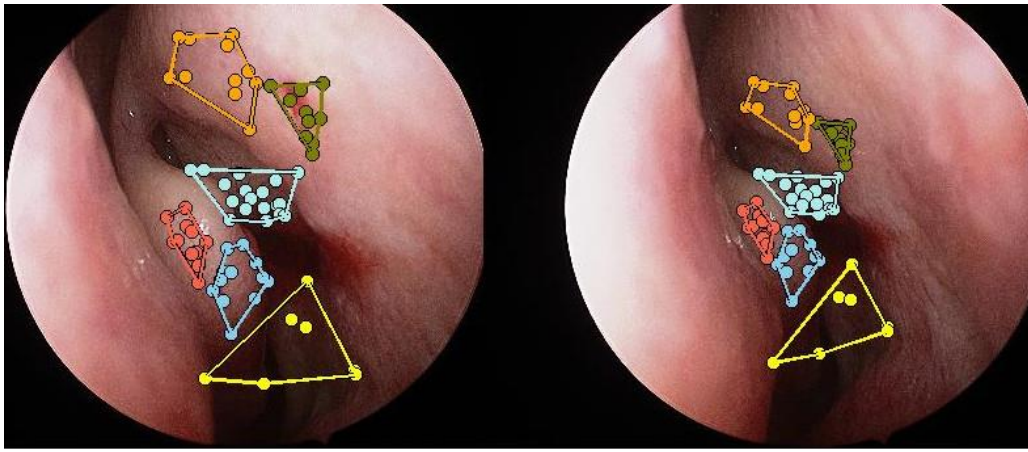
Method

1. Use a set of images to compute structure from motion (SfM)
2. Scale structure and motion to CT using the magnitude of tracking trajectory
   - Magnetic tracker is used to estimate the scale of the motion
3. Register the 3D structure to CT using trimmed-ICP (TriICP) with scale
   - Initial guess must be provided



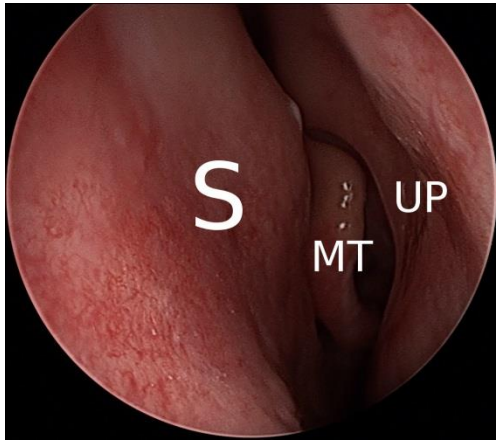
SfM
Hierarchical Multi-Affine Matching

- SfM is computed from a set of matched image-features (SIFT or SURF)
- Endoscope images and motion are challenging for conventional matching algorithms
- HMA matches increases quantity and quality of matches
  - SURF are extracted from a pool of GPU
  - Initial SURF matches using brute force algorithm
  - HMA matches computed using a pool of CPUs



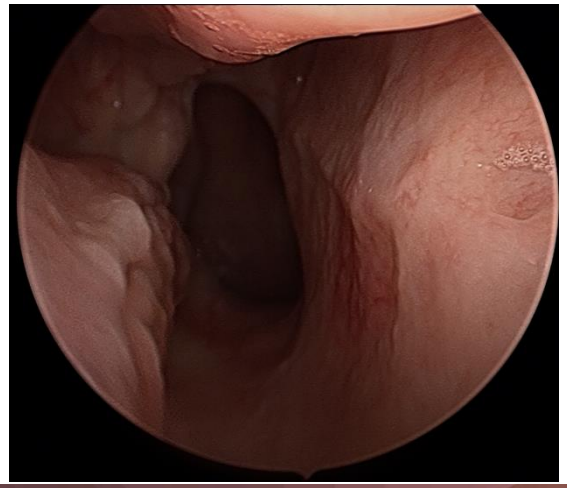
SfM
Sparse Bundle Adjustment

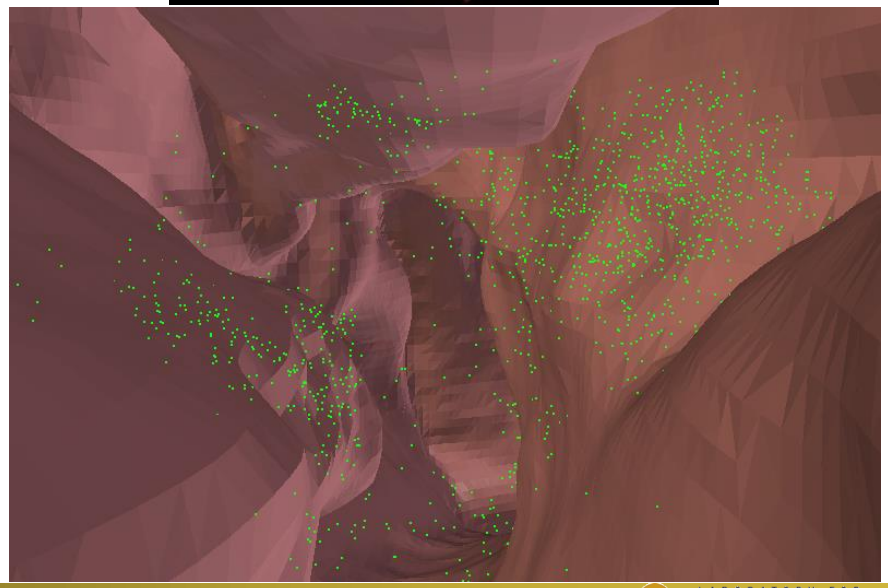
- SfM with SBA is computed from HMA matches
- The structure and motion are scaled according to the magnitude of the motion as measured from the magnetic tracker
- One second of video (~30 frames) yields between 800 and 1000 points



Trimmed Iterative Closest Point with Scale

- The scaled structured is registered to the mesh of the CT scan
- Use 70% of inliers to register structure points to CT
- Allow to scale structure to compensate for tracker inaccuracy
- Initial guess must be provided

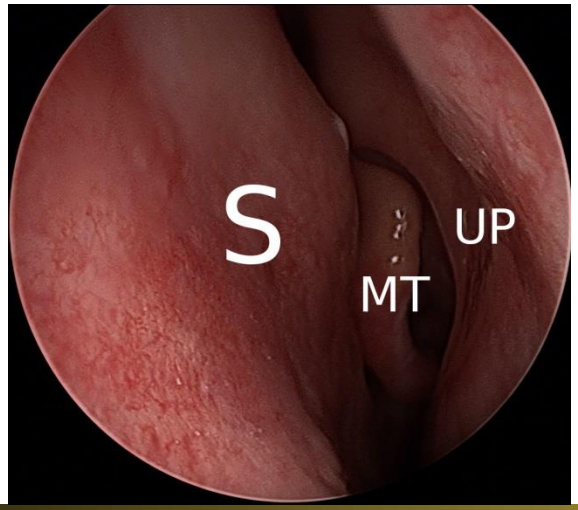




Erectile/Non-Erectile tissues

- Account for erectile/non-erectile tissues
  - Congestion cause some tissues to swell
  - Discrepancy if CT was obtained weeks before video

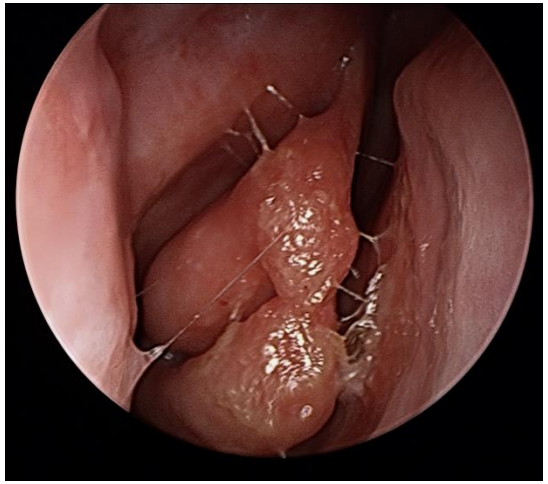
Decongested view (video)
of the middle turbinate

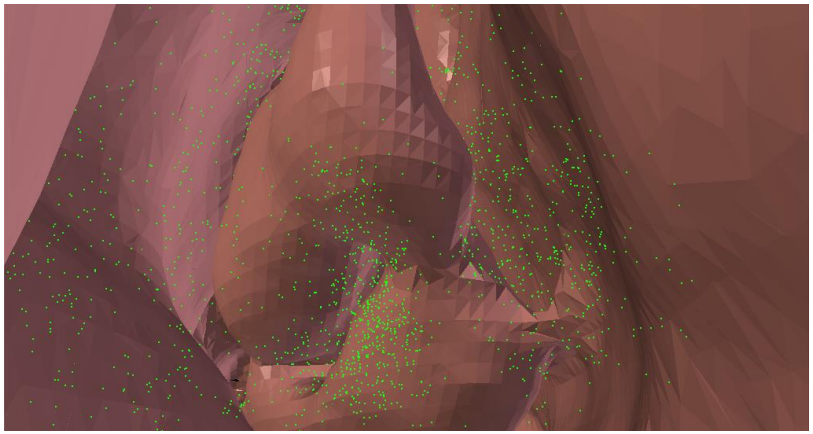


Congested view (CT)
of the middle turbinate

Results

- Video data was captured from JHOC
  - ~90 seconds per patient
  - Divided in one second video sequence (~30 frames)
- Five sequences with erectile tissues and five sequence with non-erectile tissues
  - Non-erectile tissues TriICP residual: 0.91 mm (0.2 mm)
  - Erectile tissues TriICP residual: 1.21 mm (0.3 mm)
- Average computation time
  - 10.2 seconds (1.3) for 30 frames



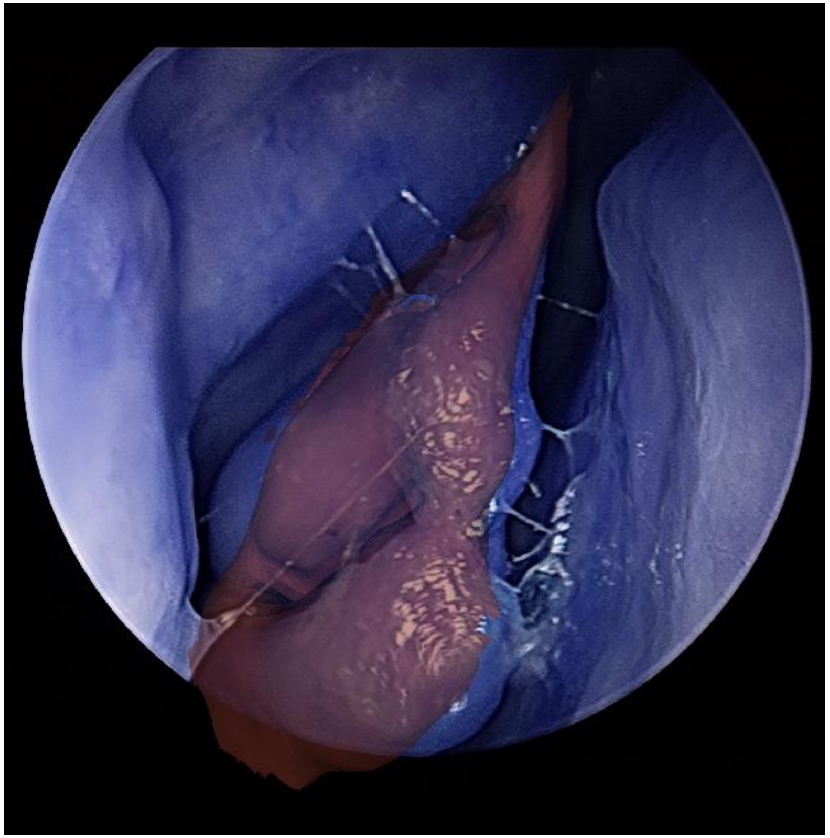


Results

- No ground truth available
- Measure reprojection error using overlaying of anatomical structure (middle turbinate)

$$\frac{\sum(I_A \oplus I_B)}{\sum I_B}$$

- 86% mean reprojection accuracy (std 3%)
  - Middle turbinate is surrounded by erectile tissues



Conclusion and Future Work

- FESS is commonly used for treatments of chronic sinus diseases
- Current navigation systems have limited accuracy
- Our research introduces a video-CT registration with sub-millimeter accuracy for non-erectile tissues with in-vivo data
- Proposed video-CT registration enables overlay of CT structures (visible or occluded) on video data
- Computation time is comparable to state of the art navigation systems (inserting and removing markers)
- Future work includes analysis of robustness to initial registration guess

Thank you!

Questions?

Acknowledgement: This work is funded by NIH R01-EB015530: Enhanced Navigation for Endoscopic Sinus Surgery through Video Analysis.