

Intraoperative Fiducial Tracking in TORS

EN.600.646 CIS2 Project #15

Spring 2014

Project Report

Xiao Hu

Project Mentors: Wen.P Liu

Anton Deguet

Part I. Technical Summary

1. Project Background

(1). Overall background

TransOral robotic surgery (TORS) is a minimally invasive surgical intervention for resection of base of tongue tumors, which has become a significant health care concern. Because most base of tongue tumors are buried deep in the musculature of the tongue, when doing the TORS, expert surgeons always rely on experience to remain correctly oriented with respect to critical anatomy. Such practice leaves considerable room for improvement.

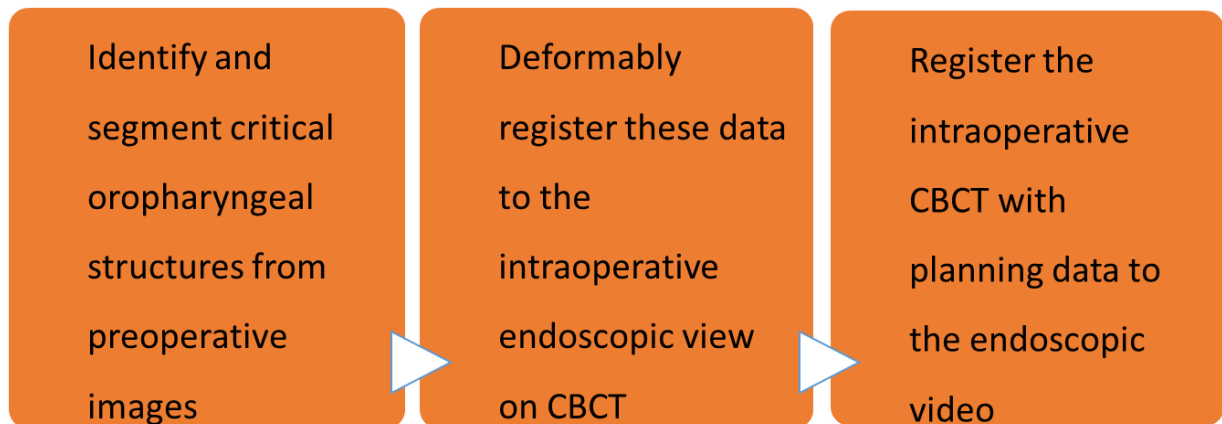
So, an image guidance system with stereoscopic video augmentation has been proposed. It integrates a preoperative surgical plan with the intraoperative endoscopic scope to do video augmentation during the surgery in the operating room. The rough steps of this method are as below.

- First, a detailed surgical plan is made by the surgeon from the preoperative CT or MR of the patient. Basically, the surgical plan is the position of the tongue tumor in the CT or MR coordinate.
- Second, when the patient is positioned for the surgery, a CBCT image is acquired immediately to capture the intraoperative deformation.
- Third, using the deformable registration between the preoperative CT (MR) and the intraoperative CBCT to get the intraoperative surgical plan on CBCT.
- Then, a rigid transformation, resulting from the tracking of the fiducial which

is attached directly above the resection area at the beginning of the surgery, between the CBCT and the endoscopic video is calculated.

- Next, the planned data can be registered to the robotic endoscopic video using this transformation, to form the augmented video during the surgery.

Those steps could be simplified as below.



The basic idea is two coordinate system transformations:

$${}^{\text{Video}}T_{\text{CT}} = ({}^{\text{Video}}T_{\text{CBCT}})({}^{\text{CBCT}}T_{\text{CT}}),$$

$$Data_{\text{Video}} = ({}^{\text{Video}}T_{\text{CT}})(Data_{\text{CT}}).$$

One method to find ${}^{\text{Video}}T_{\text{CBCT}}$ is to use rigid fiducials.

(2). Fiducials tracking background

Fiducials are attached directly above the resection area at the beginning of the surgery, before getting CBCT image of the tongue. The position of the fiducial in the CBCT coordinate system could be got manually or by some image processing. Assuming the relative position between the fiducial and the tumor is unchangeable,

which makes sense, knowing the 3D positions of each of the 3 fiducials in the two coordinate systems, the transformation between the two coordinate systems, here ${}^{\text{Video}}T_{\text{CBCT}}$, can be calculated.

Current existing method to find ${}^{\text{Video}}T_{\text{CBCT}}$ is a manual process which is based on user input, either by the operator using 3DUI virtual cursors or MTMs, or by placing tool tips of PSMs directly onto the sphere, which would require robot-to-tooltip calibration. A non-user input method should be developed.

2. Project goal

The goal of this project is to develop and implement an intraoperative fiducial tracking method that does not require user input to do the tracking, based on 2D images.

3. Approaches

The approaches to achieve the goal consist mainly of three steps:

- Use color and edge characteristics to detect the frame which encompasses the three fiducials
- Use color and position characteristics to detect the fiducials
- Apply Kalman filter to track the fiducials in videos

(1). Fiducial frame detection

- a. Use color information to detect the green pixels



Figure 1

- b. Use edge detector to detect the edges in the image



Figure 2

- c. Combine the two images to get the possible fiducial frame contour

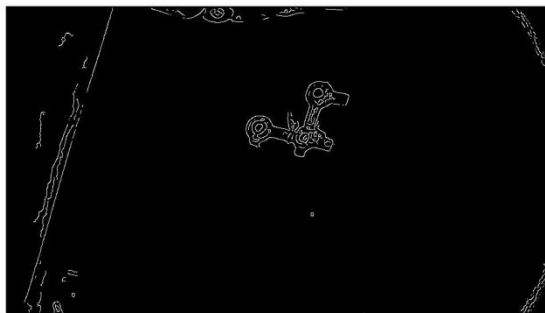


Figure 3

Since most of the points that are wrongly detected as the green frame are on some device surface with similar color, and especially in the intraoperative scene it can be inferred that those points are rare in the regions with large gradient, this step can effectively get rid of most of the wrongly detected points.

The RGB value of the green frame is the essential factor for this detection. When the value gets closer to Green and gets more unique, the better the detection result is. Figure 4 shows a much clearer frame detection than Figure 3, because the RGB value of its frame is more distinguished.

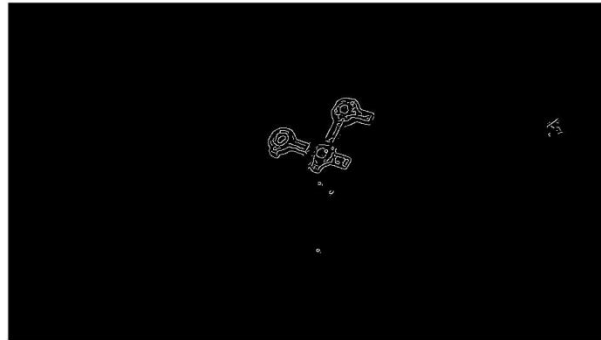


Figure 4

However, using figure 3 as the input for next steps can still get reasonable result.

(2). Fiducial Detection

a. Use the frame contour and the color information to detect the three colored fiducials, and delete some noise points

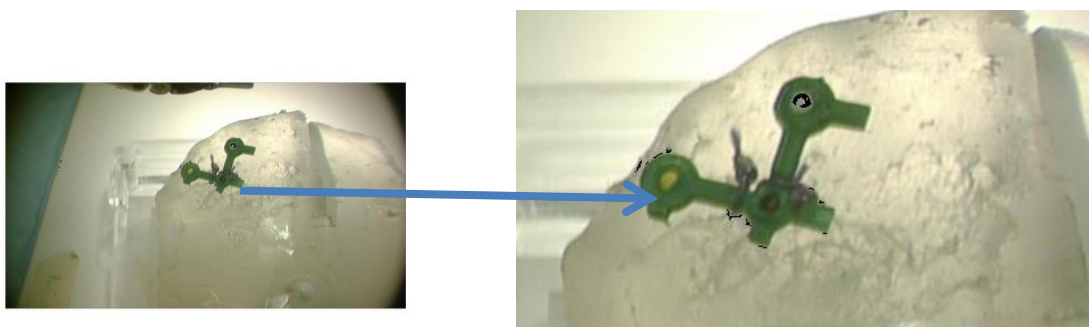


Figure 5 take the white fiducial for example

b. Use connectivity to group pixels that might belong to a fiducial

For example for Figure 5, to group the pixels whose $RGB=[0,0,0]$.

- c. Use a weighting function to select the fiducial group from the candidate groups



Figure 6 Fiducial detection. The center of each square is the detected fiducial center, and the half length of the square is the detected fiducial radius

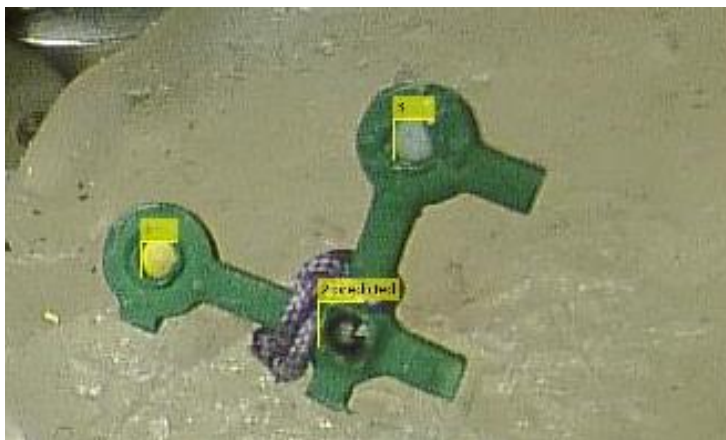
(3). Fiducial tracking in videos

We've applied the Matlab "multiObjectTracking" and associated functions for implementation. It applies Kalman filter for tracking, which is used to find fiducials of which the detections are missing. The tracking is based on motion with the assumption that all objects move in a straight line with constant speed.

4. Implementation results

We have tested our algorithm in three videos: the one that the fiducial is pitching, the one it is yawing, and the one it is occluded.

(1). Pitching:

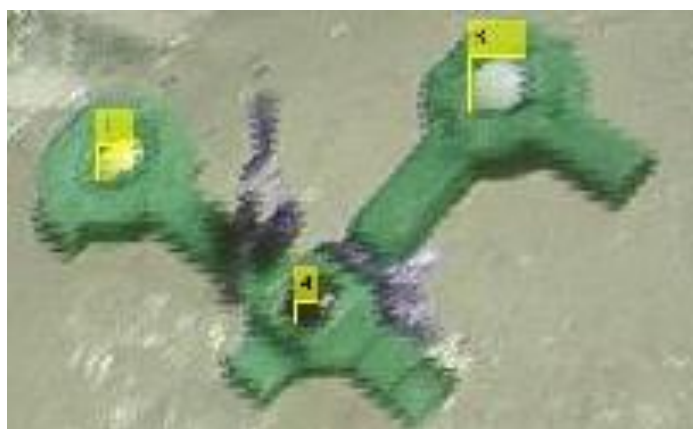
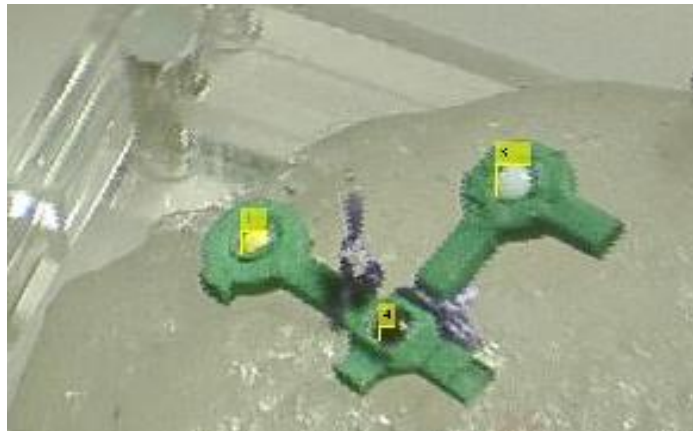


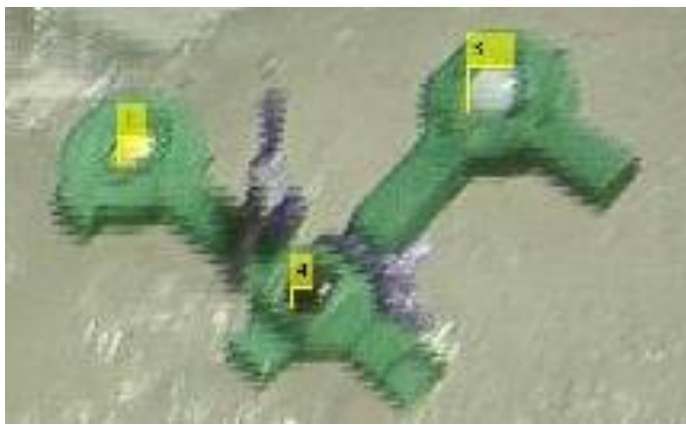
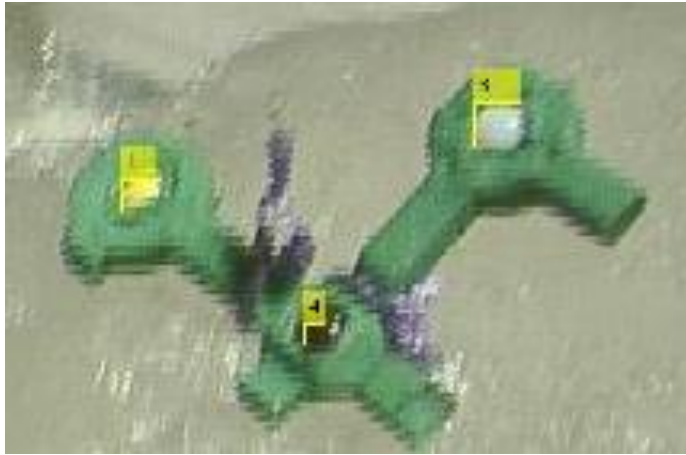


Total video frames: 52

Correctly tracked: 33 (including the first 7 training frames)

(2). Yawing:

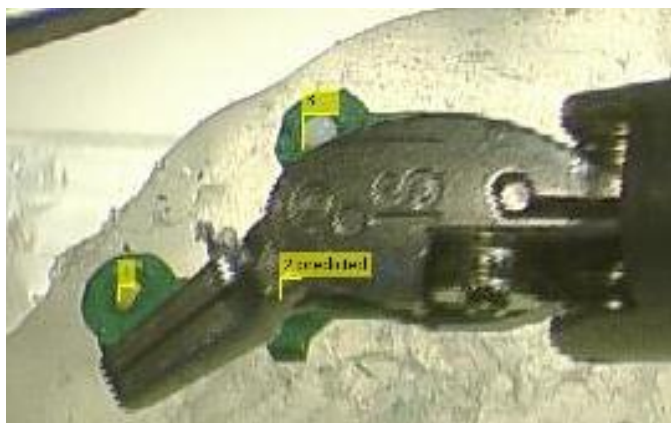
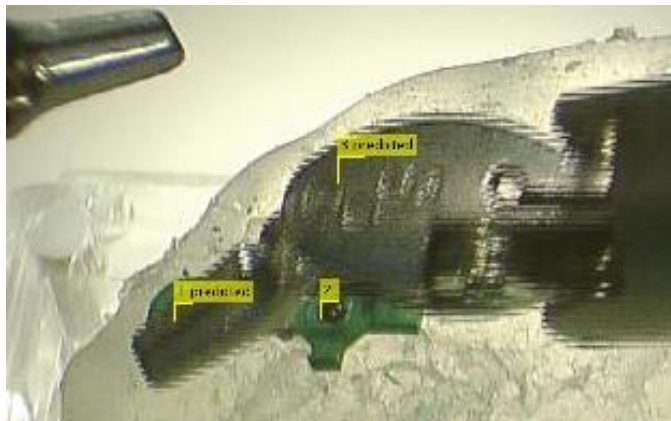
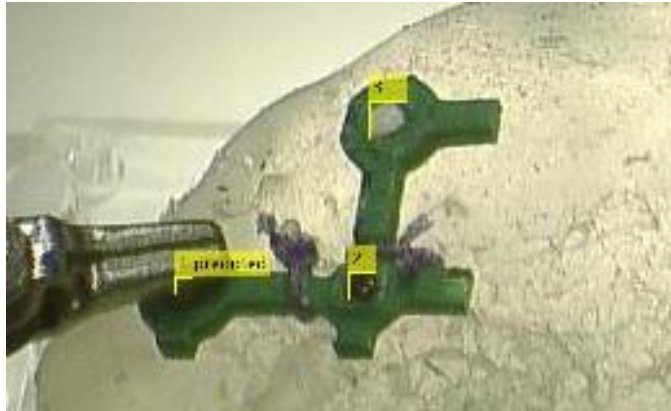




Total video frames:70

Correctly tracked: 65 (including the first 7 training frames)

(3). Occlusion, the fiducials not moving



Total video frames: 30

Correctly tracked: 30 (including the first 7 training frames)

5. Discussions

The algorithm is robust for yawing and occlusion, but for pitching, when the pitch angle is over approximately 70 degrees, the black fiducial can't be tracked. It is because the RGB value of the black fiducial changes significantly according to shading and reflection during the process, and at the same time, shading can make the RGB value of other pixels quite close to that of the black fiducial, which results in the wrong detection.

I think the method that might solve this problem is to redesign the black fiducial, changing its color into blue or some other color more distinguishable.

Also, the tracking algorithm can't run in real-time, having approximate 1 second delay. To solve this, I think an initial user input to get the region of the green frame would help.

Part II. Management Summary

1. What was accomplished vs. planned

- Implementation of fiducial detection of the recorded endoscopic images
Done 4/14
- Implementation of fiducial detection of the recorded stereo video
Done 4/22
- Real-time fiducial tracking
Done tracking 4/25, but not quite real-time
- Fiducial tracking of the intraoperative stereo video
No
- Video recording for the intraoperative fiducial tracking process
No
- A new fiducial for better and more accurate tracking
No

2. What might be next

- Redesign the black fiducial
- Check the tracking result of the intraoperative videos
- Rewrite the code using cisst

3. What I've learnt through this project

- Color image segmentation
- Active Contours without edges algorithm
- Some of the Cisst libraries
- And don't underestimate the difficulty of a project until having achieved the desired outcome