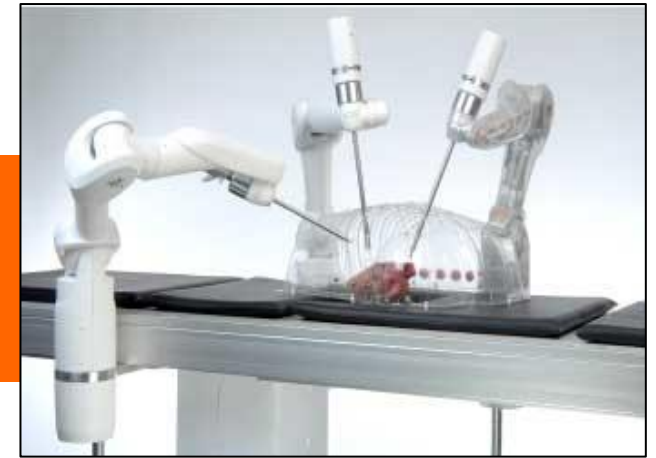


“Pose Estimation of Medical Instruments with a Robotic Vision System”

Case Study Project
Summer semester 2024

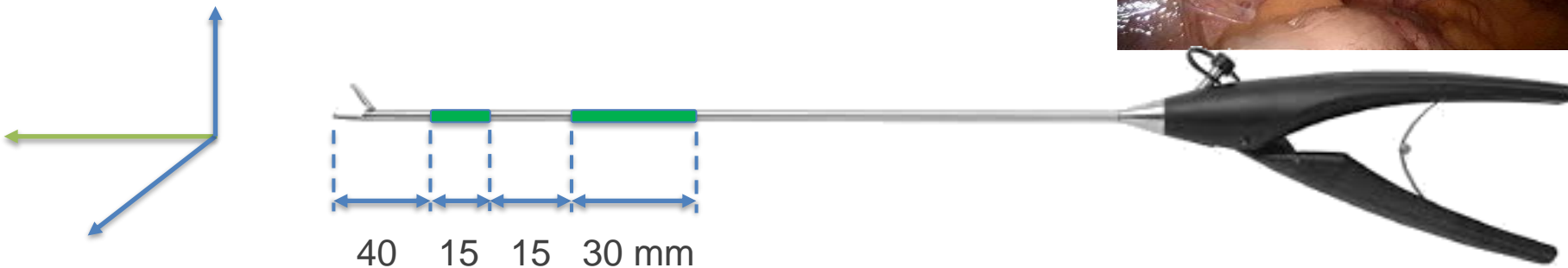
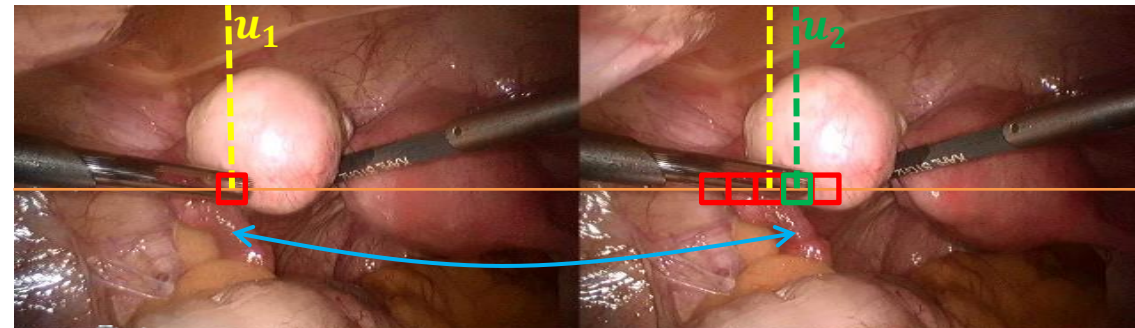
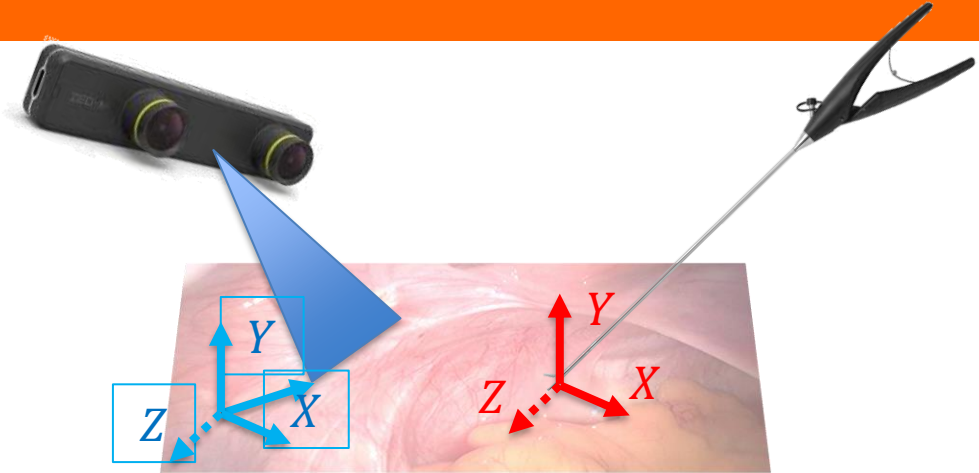
Dortmund, 17.06.2024

Dr.-Ing. Tai Fei



Project “Tracking of the Needle holder”

- 3D Tracking of Needle Holder Trajectory During Surgery.
- Primary Goal: Object Detection and Trajectory Tracking of the Needle Holder Tip
 - Data: Stereo images from an embedded vision system
 - Detection of colored objects within the surgical scene
 - 3D-Reconstruction, orientation of needle holder rod
 - Use marker on the plan of the medical scene picture as coordinate origin.
 - Track the trajectory of the needle holder tip and annotate the orientation vector of the rod at each position.
 - **Optional:** Is the pose of needle holder completely defined? What is missing? Any suggestion for solution?



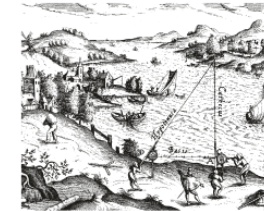
Documentation: MATLAB, Stereo Vision

- **ZED with Matlab**
 - <https://www.stereolabs.com/docs/matlab>
- **ZED Camera calibration parameter**
 - <https://support.stereolabs.com/hc/en-us/articles/360007497173-What-is-the-calibration-file>
- **Assistance on the Topic of Stereo Vision**
 - Chapter 14, „Robotics, Vision and Control“, Peter Corke, 2nd Editor
- **Documentation Matlab Toolbox of Machine Vision (Peter Corke)**
 - <https://www.petercorke.com/MVTB/vision.pdf>
- **Matlab Tutorial for Stereo Vision**
 - <https://de.mathworks.com/discovery/stereo-vision.html>
- **40 rectified images reference to validate the rectification in folder “reference_rectified”**

Chapter

14

Using Multiple Images



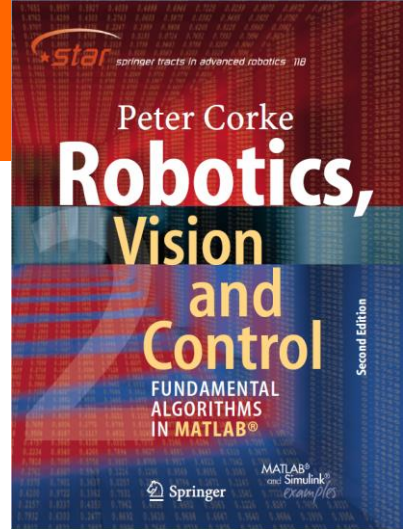
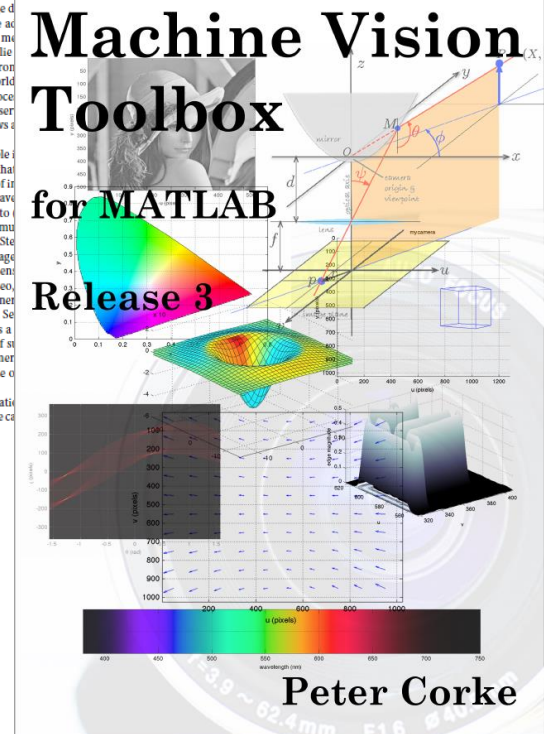
In the previous chapter we learned about corner detectors which find particularly distinctive *points* in a scene. These points can be reliably detected in different views of the same scene irrespective of viewpoint or lighting conditions. Such points are characterized by high image gradients in orthogonal directions and typically occur on the corners of objects. However the 3-dimensional coordinate of the corresponding world point was lost in the perspective projection process which we discussed in Chap. 11 – we mapped a 3-dimensional world point to a 2-dimensional image coordinate. All we know is that the world point lies along some ray in space corresponding to the pixel coordinate, as shown in Fig. 11.6. To recover the missing third dimension we need additional information. In Sect. 11.2.3 the additional information was camera calibration

Almost if we can determine the translation of the camera only up to an unknown scale factor, that is, the translation is $\lambda \mathbf{t} \in \mathbb{R}^3$ where the direction of \mathbf{t} is known but λ is not.

to estimate the object's 3-dimensional pose from 2-dimensional image data. In this chapter we consider an alternative approach in which the information comes from *multiple* views of the same scene. As already mentioned, pixel coordinates from a single view constrain the world point to lie on a ray. If we can locate the same world point in another image, taken from a different pose, we can determine another ray along which that world point lies. The world point lies at the intersection of these two rays – a process known as triangulation or 3D reconstruction. Even more powerfully, if we observe many points, we can estimate the 3D motion of the camera between the views.

The underlying challenge is to find the same world point in multiple images. This is the *correspondence problem*, an important but nontrivial problem that we discuss in Sect. 14.1. In Sect. 14.2 we revisit the fundamental geometry of stereo vision developed in Chap. 11 for the case of a single camera. If you have that chapter, or it's been a while since you read it, it would be helpful to yourself with that material. We extend the geometry to encompass multiple cameras and show the geometric relationship between pairs of images. Stereo vision is an important technique for robotics where information from two images taken from different viewpoints, is combined to determine the 3-dimensional structure of the world. We discuss sparse and dense approaches to stereo reconstruction, in some detail in Sect. 14.3. Bundle adjustment is a very general technique for combining information from many cameras and is introduced in Sect. 14.4. 3-dimensional information that is created is typically represented as a set of 3D points, and techniques for plane fitting and alignment of surfaces are introduced in Sect. 14.5. For some applications we can use RGBD camera turn depth as well as color information and the underlying principle of light is introduced in Sect. 14.6.

We finish this chapter, and this part of the book, with four applications based on the concepts we have learned. Section 14.7.1 describes how we can



Robotic Vision Project Topic Definition SoSe 2024

- Case Study Project (ESE, IT, BMIT, RMS)
 - Calendar week 25 ~ 28 → @Home/Remote or @A604 on Wednesday, 8.30am-13.30am
 - Oral exam → 15.07. - 26.07.2024 (tbd)
 - Submission (slides, files) → Saturday, 13.07.2024, 12:00