Assignment for Technical Evaluation

Assignment: 3D Pointer and Simple Stroke Recognition

For this technical evaluation, we ask you to create a prototype-level application written in C++, using data from DepthSense cameras.

Given a sequence of depth and confidence image pairs recorded from two models of DepthSense cameras[1], create an application that isolates the hand in the scene, computes a pointer located in that hand (in depth-image space), and uses the sequence of positions to classify the stroke. We suggest to use the *$1 Unistroke Recognizer*[2], which provides at least two C++ implementations, but you are free to use another solution that you find suitable. Pay attention to the different scenarios in the given image sequences: some gestures are performed with two hands.

Deliverables and Evaluation Modalities

Document and explain the rationale of all the assumptions you are making, such as the known working conditions and limitations of your solution. We will discuss them during the next interview.

For the next interview, the project sources should be in a state where it is possible to modify, rebuild and run the application. During that interview, the following aspects will be discussed:

- The technical challenges encountered, and how you solved them
- The various paths explored, including those that were discarded, and why
- A quick review of the code structure and how the different pieces interact together

We will use a videoconferencing software with screen sharing capabilities so you may show a live demo, support material such as a slideshow or drawings, and so we can browse through the code. The modalities for that call will be arranged by email.

[1] See details in the *Resources* section of this document

[2] https://depts.washington.edu/madlab/proj/dollar/impact.html

Resources

Datasets

This document is accompanied by a link to an archive containing several image sequences, saved as numbered tiff files. You may use the *imread()* function from OpenCV [3]to load these files, and simulate a video stream in the application. The sequences were recorded at 30 fps.

These image sequences present different scenarios, to develop and test your tracking and detection methods.

The dataset includes a file name named ***camera_parameters.txt*** for each camera, which lists the intrinsic parameters of the depth camera used to capture the images. The pinhole parameters are given in pixel units.

Development Tools

We recommend using Visual Studio 2019[4] for these assignments. We suggest the use OpenCV for image processing and basic visualization. These recommendations are not requirements. You may use another C++ IDE, operating system and support libraries for image processing and visualization if you are more comfortable with them.

Additional Information

Feel free to send questions to clarify any point in this assignment.

[3] http://opencv.org/

[4] https://www.visualstudio.com/en-us/products/visual-studio-community-vs.aspx