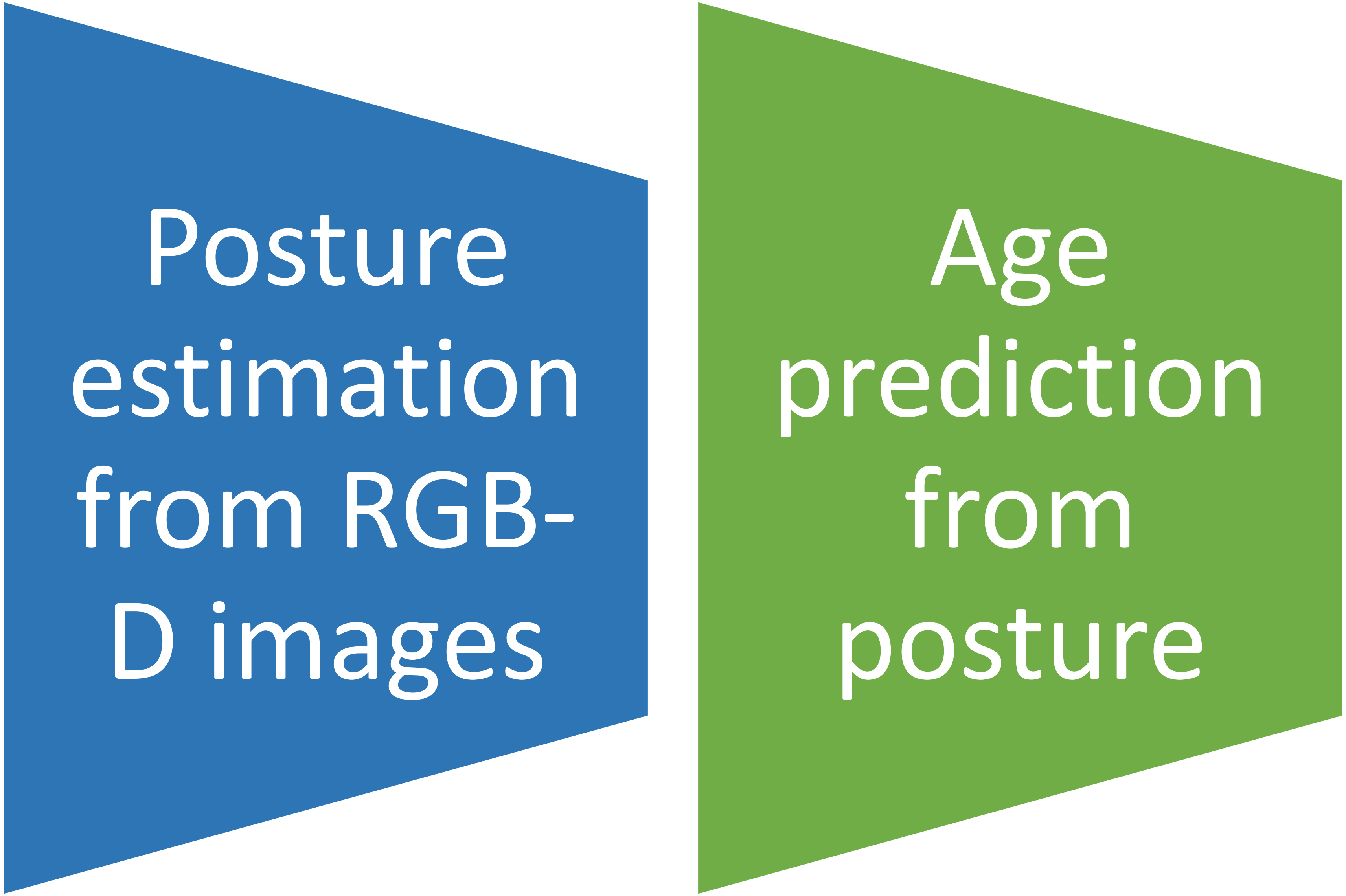
Age Prediction from Posture Estimation

Lotem Nadir, October 2021



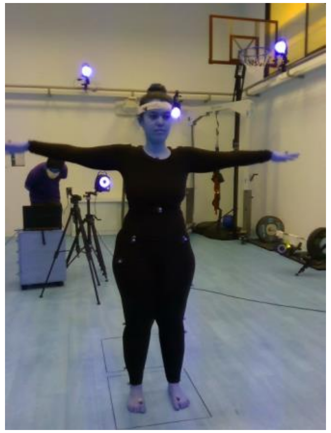
Posture estimation from RGB-D images

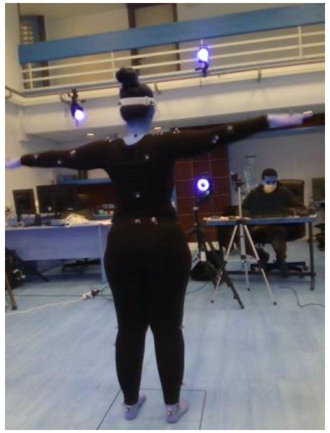
Age prediction from posture

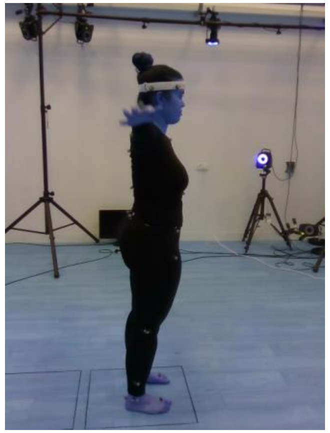
Posture estimation from RGB-D images

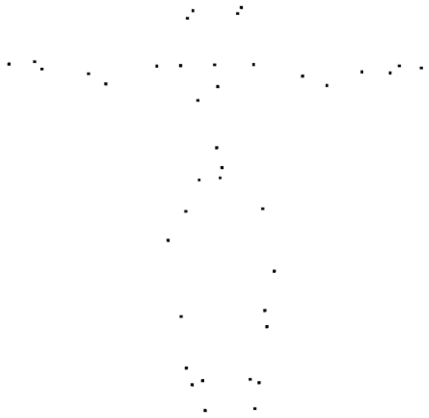
Dataset - Overview

- 15 recording sessions.
- In each recording session, a person was recorded using 3 RealSense cameras (RGB-D) and Vicon sensors (3D coordinates).
- The RealSense cameras record 3 different angles: front, back and side.









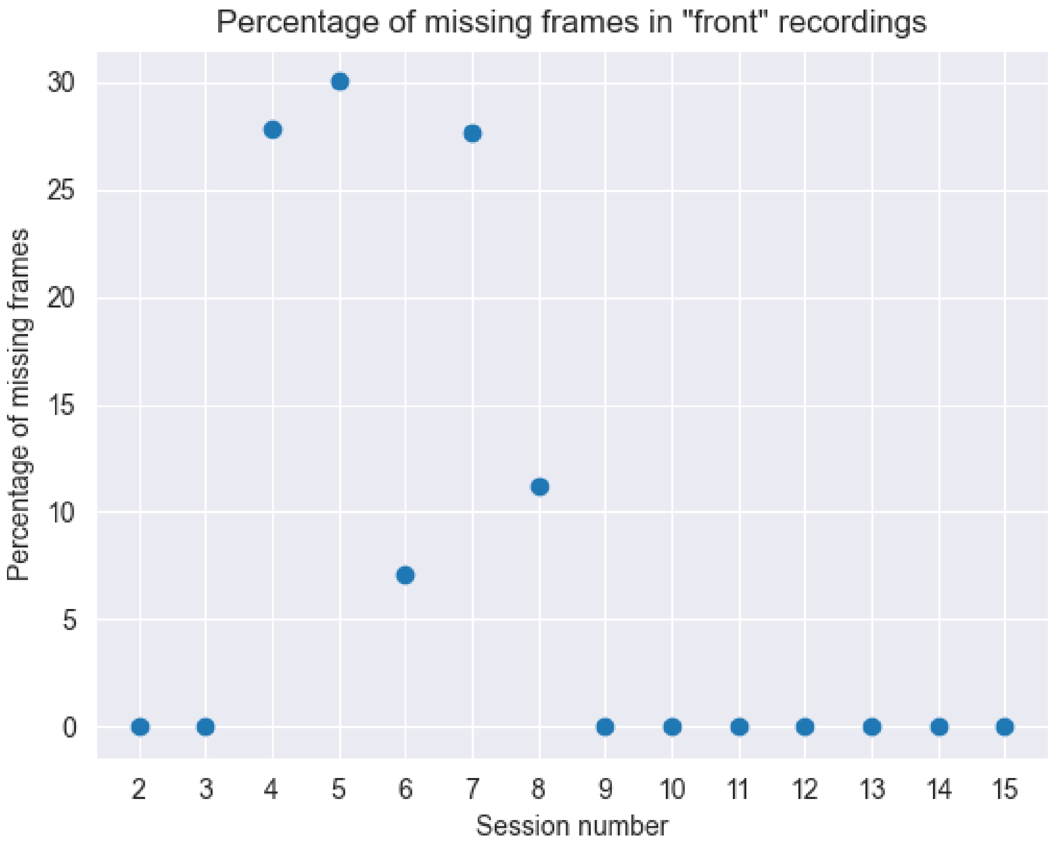
Posture estimation from RGB-D images

Dataset - Challenges

- The recordings are not synchronized by default, i.e. in the first frame in each recording the object is in a slightly different position.
- The FPS of the RealSense is 30, and the FPS of the Vicon is 120. In some recordings, the RealSense’s FPS is 15.
- Since 2 of the RealSense cameras in each session are connected to the same laptop, there is a frame drop in the output of these cameras.
- There are no timestamps in the Vicon data.

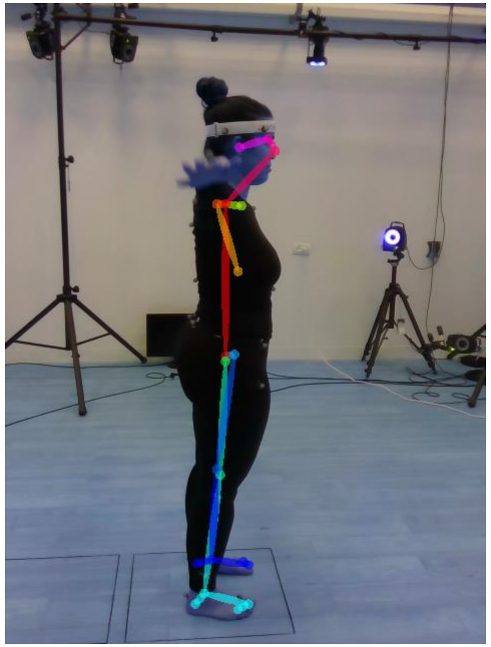
Posture estimation from RGB-D images

Dataset – Challenges



Dataset – Main steps were done

- Synchronizing the recordings:
  - Tried to use OpenPose for detecting the T-pose at the beginning of each recording. Bad results on the side angle, required manual fixes on the front and back angles.
  - Manually detected the T-pose in each recording.



OpenPose output on “side” images

Posture estimation from RGB-D images

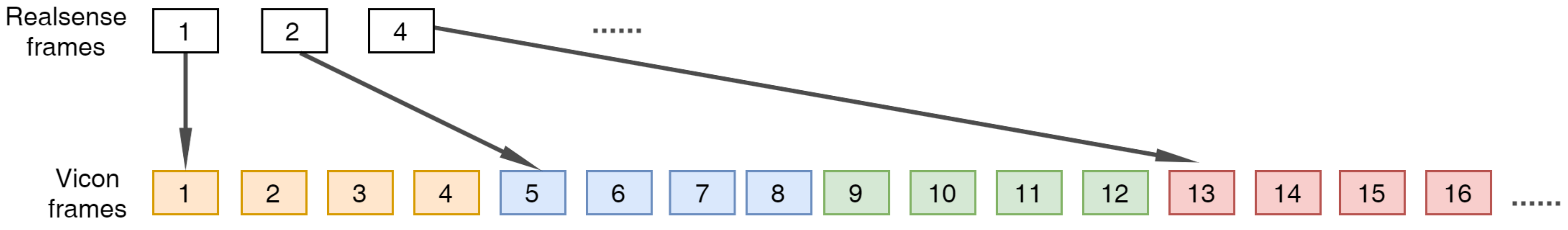
Dataset – Main steps were done

- Different FPS:
  - I have decided to take every 4th (or 8th) frame from the Vicon data. I've also tried to average every 4 (or 8) frames. The measurement to check which method is better was to calculate an angle in the neck. There was an average difference of ~0.05 degrees per second between the 2 methods.

Posture estimation from RGB-D images

Dataset – Main steps were done

- Frame drop in RealSense videos:
    - I have extracted for each RealSense recording the frames numbers from the bag file, and used the differences in the frames in order to "keep" only the corresponding frames from the Vicon data.



Dataset – Main steps were done

- Implemented trimming scripts.
- Validation of the synchronizing was done manually as well. Sometimes, due to the frame-drop, the T-pose was not detected correctly, and I had to pick another frame as first frame.

Posture estimation from RGB-D images

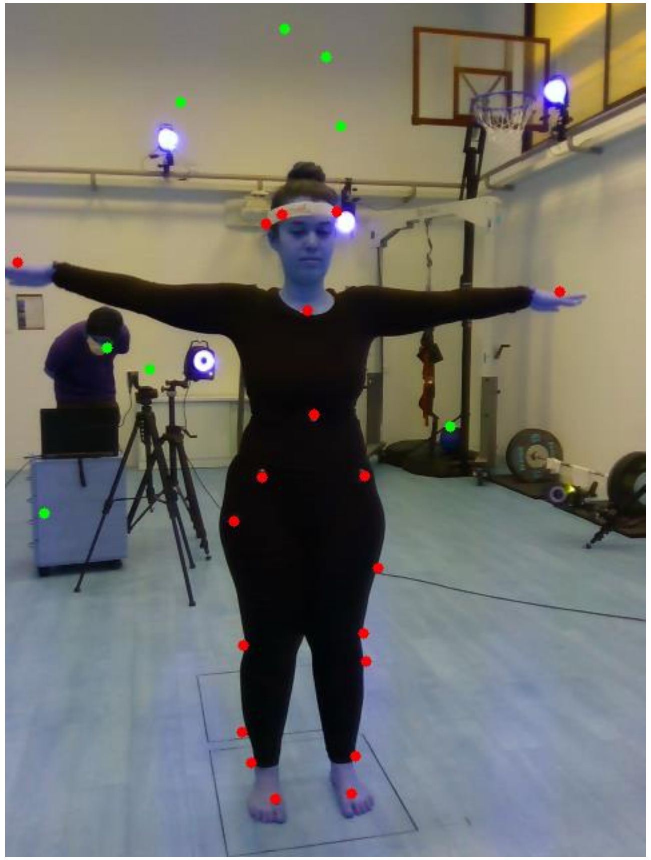
Projecting Vicon points into RealSense pixels

- OpenPose is a model for 2d pose estimation from RGB images.
- We had to project the Vicon points into the RealSense pixels.
- The RealSense cameras and the Vicon sensors are not calibrated.
- Kabsch algorithm is used to find the transformation between the Vicon system and the RealSense 3D system.
- Improvements to Kabsch:
    - Averaging the depth value (z) for each point with neighbor pixels.
    - Removing noisy depth values – For each point, if its depth (z) value is greater than some clipping distance, the point is removed.
    - Running Kabsch and finding error for each point, removing points with high error rates and re-calculating Kabsch.

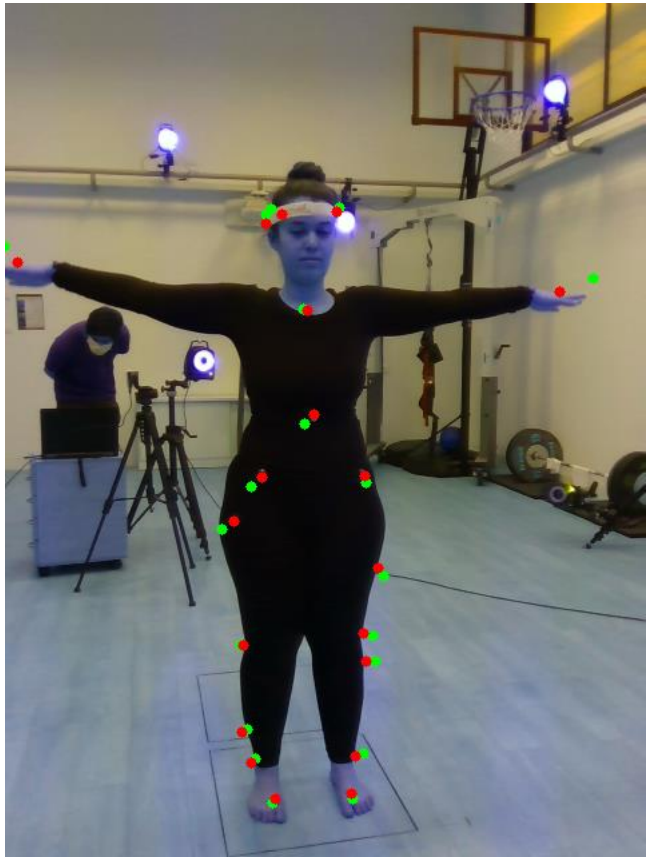
Current error after improvements is ~60mm.

Posture estimation from RGB-D images

Projecting Vicon points into RealSense pixels



without improvements



with improvements

Projecting Vicon points into RealSense pixels

- Currently working on improving the transformation error:
  - For future recordings by placing a static object for calibration.
  - For previous recordings by picking frames where the object is as static as possible.

Age prediction from posture

- Generated dataset: ~250,000 samples, each sample has 39 3d points, and age as label. Many of the samples are “duplicates” due to the Vicon high FPS.
- Currently training PointNet on the dataset, for classifying the objects as ‘old’ or ‘young’.