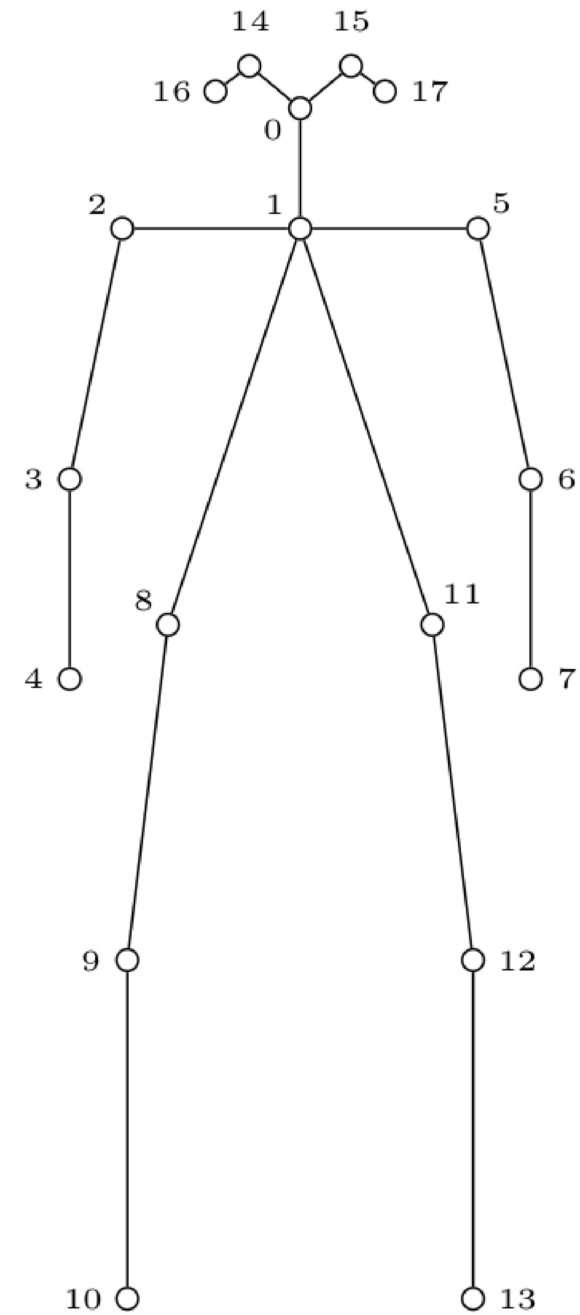
Humaintude Teaching Assistant

- Portable
  - *On site data gathering in real-world conditions*
  - *Support for using multiple RGBD-cameras*
  - *Sensor agnostic, open source and cross platform*
- 3D skeleton
  - *Invariant to capture-configuration*
  - *Full skeleton*
  - *Keypoint refinement*
- Tracking
  - *Separate data for patient and practitioner*
  - *Time-dependent information*
  - *Extract Region of Interest*

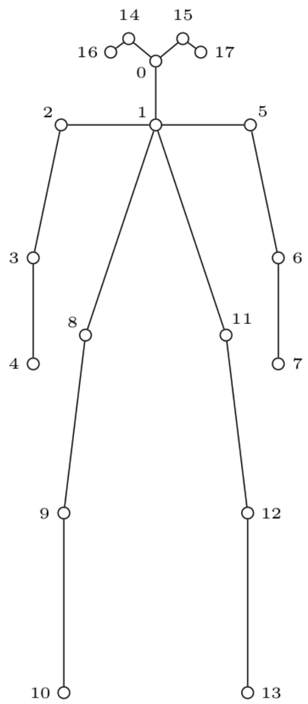


Future work

- Data gathering
  - *Record for each technique*
  - *Multiple subjects*
  - *Machine learning models*
- Teaching Program
  - *Step-by-step instructions*
  - *Validation using Machine Learning*

Humaintude Teaching Assistant

- Portable
  - *On site data gathering in real-world conditions*
  - *Support for using multiple RGBD-cameras*
  - *Sensor agnostic, open source and cross platform*
- 3D skeleton
  - *Invariant to capture-configuration*
  - *Full skeleton*
  - *Keypoint refinement*
- Tracking
  - *Separate data for patient and practitioner*
  - *Time-dependent information*
  - *Extract Region of Interest*



The proposed system has been implemented using OpenPose’s keypoint extractor on an IR image provided by an RGBD camera, and extracts a scaled 3D skeleton using the additional depth channel. The system features support for simultaneous data capture from multiple angles, and subject tracking.

Portable

The system is portable, and can operate on 1 or more RGBD cameras. This allows us to bring the system to real locations to gather accurate data for later reference and training.

As mentioned, the system allows for using multiple RGBD cameras, to increase accuracy and valid recording area.

It is also written in ROS, and as long as depth, ir, and RGBD images can be provided, the system can work on any cameras. That being said, it has thus far only been tested on Microsoft’s Kinect v2 camera.

Since it is written in C++ for ROS, we can deploy the system on a wide variety of setups, as long as ROS support is provided

3D skeleton

Since the 3D data is recorded, we can relate the data to any world-frame. (For example a selected keypoint in one of the observed persons.) This allows the system to be setup agnostic, and training data can be recorded on different camera configurations than for example validation data. We can also train the system on different camera configurations for more robust models. Accuracy can also be improved by the number of cameras used to track the subject.

A full skeleton is provided to the system by combining data from multiple RGBD cameras, and by guessing the position of unobserved keypoints. The guesses are either based on the parent keypoints, or by the previous position if it was observed at an earlier moment.

As mentioned earlier, OpenPose is used to detect the humans in the IR scene, and the 3D positions for each keypoint is then calculated using the depth information and the cameras intrinsic parameters. A good calibration file is therefore recommended for best possible results.

Tracking

Tracking is needed to separate the patient’s data from the practitioner’s data. It is also needed to provide better estimation for unobserved limbs, and storing time-dependent information.

Currently the tracking system is realized by assgning IDs based on the last observed each observed persons center of mass. It is however desired to create a filter and a more advanced tracking method to ensure that the IDs are not confused as time progresses.

Our system also provides an easy extraction method for Regions of interest. A window size in meters, as well as the pixel positions of the center is provided (for example the nose-keypoint of a person), and an image containing the region is calculated. This allows us to reduce the area extracted if a person is farther away from the camera.

Figure The figure shows the keypoints provided by the OpenPose system, and is therefore the keypoints our system tracks

Figure keypoints	
0	Nose
1	Neck
2	R shoulder
3	R elbow
4	R hand
5	L shoulder
6	L elbow
7	L hand
8	R hip
9	R knee
10	R foot
11	L hip
12	L knee
13	L foot
14	R eye
15	L eye
16	R ear
17	L ear

Future work

- Data gathering
  - *Record for each technique*
  - *Multiple subjects*
  - *Machine learning models*
- Teaching Program
  - *Step-by-step instructions*
  - *Validation using Machine Learning*

Data gathering

Records should be made for each technique on location in the real world, and in a controlled environment in the lab with multiple cameras, if possible.

Need to gather data for each technique we want the students to learn. Multiple subjects of different sizes should be recorded for each technique.

Then a machine learning model should be trained to recognize each motion. Playback of the motion might be difficult, but this would be an interesting problem to solve. We propose one way of doing this as averaging all the training data. *Here some new ideas are needed.*

Teaching program

We propose a system where we will use Machine learning to validate the students execution of each technique.

Step by step instructions are given to the student, where a 3D avatar shows how the technique should be preformed. Then, if we define the patient as a common frame between the current data and the data modeled from training the network. a 3D representation could be shown with the students current position in relation to the correct motion. Then, it would be easy for the student to align themselves in the correct position.

When the program validates the motion as correct, the next instruction are given. At the end, the student should receive a short summary on the performance, mentioning what can be done different. (Such as “Position yourself closer to the patient” or “Your center of mass should be further back when preforming the lifting maneuver”.)