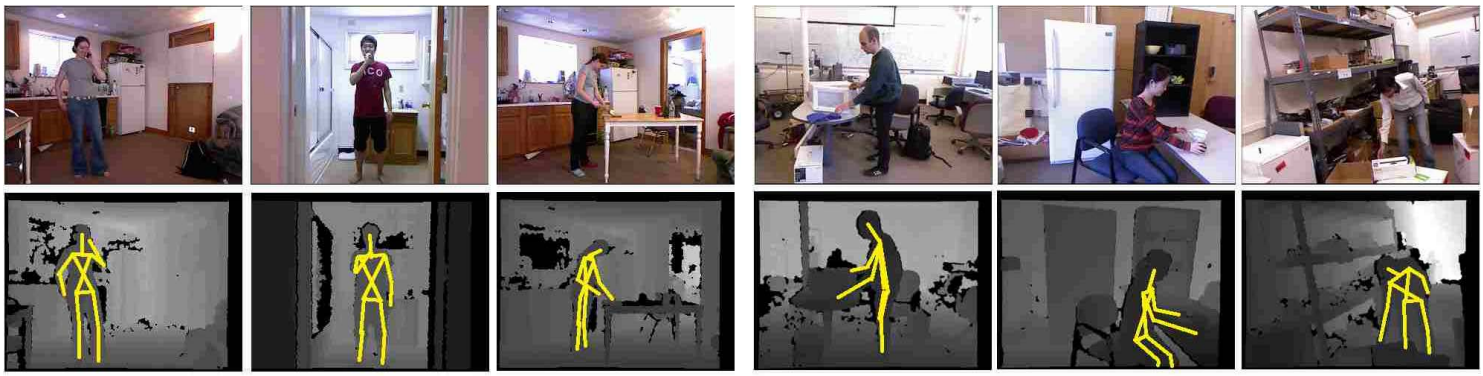
Robot Learning Lab

Personal Robotics, Co-Robots, Robotic Perception. Computer Science Department, Cornell University.

Cornell Activity Datasets: CAD-60 & CAD-120



Data

The CAD-60 and CAD-120 data sets comprise of RGB-D video sequences of humans performing activities which are recording using the Microsoft Kinect sensor.

CAD-60

CAD-60 dataset features:

- 60 RGB-D videos
- 4 subjects: *two male, two female, one left-handed*
- 5 different environments: *office, kitchen, bedroom, bathroom, and living room*
- 12 activities: *rinsing mouth, brushing teeth, wearing contact lens, talking on the phone, drinking water, opening pill container, cooking (chopping), cooking (stirring), talking on couch, relaxing on couch, writing on whiteboard, working on computer*
- tracked skeletons

Each video come with RGB images, Depth images, and the tracked skeletons.

Information:	README	Sample Images	State of the art results

Download RGB-D + Skeleton:	Person 1	Person 2	Person 3	Person 4

CAD-120

CAD-120 dataset features:

- 120 RGB-D videos of long daily activities
- 4 subjects: *two male, two female, one left-handed*
- 10 high-level activities: *making cereal, taking medicine, stacking objects, unstacking objects, microwaving food, picking objects, cleaning objects, taking food, arranging objects, having a meal*
- 10 sub-activity labels: *reaching, moving, pouring, eating, drinking, opening, placing, closing, scrubbing, null*
- 12 object affordance labels: *reachable, movable, pourable, pourto, containable, drinkable, openable, placeable, closable, scrubbable, scrubber, stationary*
- tracked skeletons

Click here for samle images.

RGB-D images:	Person 1	Person 2	Person 3	Person 4	README
RGB-D text data:	Person 1	Person 2	Person 3	Person 4	README
Annotations:	Person 1	Person 2	Person 3	Person 4	README
Features:	Features	README			

Data Format

Skeleton Data Format

Skeleton data consists of 15 joints. There are 11 joints that have both joint orientation and joint position. And, 4 joints that only have joint position. Each row follows the following format.

```
Frame#,ORI(1),P(1),ORI(2),P(2),...,P(11),J(11),P(12),...,P(15)

  Frame# => integer starting from 1

  ORI(i) => orientation of ith joint
              0 1 2
              3 4 5
              6 7 8
            3x3 matrix is stored as followed by CONF
              0,1,2,3,4,5,6,7,8,CONF
            Read NITE PDF (see below) to get more detail about the matrix

  P(i)   => position of ith joint followed by CONF
              x,y,z,CONF
            values are in milimeters

  CONF   => boolean confidence value (0 or 1)
            Read NITE PDF (see below) to get more detail about the confidence value


Joint number -> Joint name
   1 -> HEAD
   2 -> NECK
   3 -> TORSO
   4 -> LEFT_SHOULDER
   5 -> LEFT_ELBOW
   6 -> RIGHT_SHOULDER
   7 -> RIGHT_ELBOW
   8 -> LEFT_HIP
   9 -> LEFT_KNEE
  10 -> RIGHT_HIP
  11 -> RIGHT_KNEE
```

```
12 -> LEFT_HAND
13 -> RIGHT_HAND
14 -> LEFT_FOOT
15 -> RIGHT_FOOT
```

Read page 10~13 of NITE 1.3 PDF for more detail on skeleton orientation, position, and confidence values.

RGBD Data Format

RGBD data has resolution of 240 by 320. RGB is saved as three-channel 8-bit PNG file. And, Depth is saved as single-channel 16-bit PNG file. Due to alignment of Depth and RGB data, some pixels on the edges will have value of 0. Refer to "Feature Extraction" in Sung et al. code below to look at how to parse these PNG files.

Code

All of the code described in our papers can be retrieved by running following command.

1. Human Activity Feature Extraction for Detection: Sung et al. AAAI PAIR 2011, ICRA 2012

   ```
   git clone git://github.com/jysung/activity_detection.git
   ```

   The human_activity_detection repository has code for the following:

   - Feature generation code
   - Skeleton visualization code

2. Activity Labeling Code: Koppula et al. IJRR 2013

   ```
   git clone git://github.com/hemakoppula/human_activity_labeling.git
   ```

   The human_activity_labeling repository has code for the following:

   - Feature generation code: ROS and PCL packages.
   - Learning and inference code (using python interface to svm_struct).

3. Activity Anticipation Code, Koppula and Saxena. RSS 2013

   ```
   git clone git://github.com/hemakoppula/human_activity_anticipation.git
   ```

   The human_activity_anticipation repository has code for the following:

   - Anticipation code: PCL package.
   - Learning code (using python interface to svm_struct).

Skeleton Visualization Tool

Available in Sung et al. code repository above.

Turns specific frame of skeleton data into Matlab 3D plot. It shows each joints location as well as orientation of joints that have such information. Refer to README for more details.
Requires: Matlab (tested on R2010a)