Human-Animal Affective Robot Touch (HAART) Gesture Data Set

General Information about HAART

This is the HAART dataset, collected from 10 participants performing 7 touch gestures over 12 different conditions on the same pressure sensor. The training set features all the labels and data collected from 7 of the participants. The remaining 3 participants' data forms the test set. Predictions should be made on the gesture label.

The pressure sensor is a 10x10 pixel grid where each pixel is approximately 1 square inch. Each pixel reacts to contact, providing a pressure reading and a location (ie the activated pixel).

Participants were instructed to perform 7 touch gestures:

- No Touch (no contact with the sensor)
- Constant (constant contact; much like Press)
- Pat (rhythmic patting by palm and fingers)
- Rub (back and forth consistent contact with fingers and/or palm)
- Scratch (back and forth motion mostly contact with fingernails)
- Stroke (unidirectional touch with flat of hand)
- Tickle (light fingertip motion)

All gestures were performed on different combinations of substrate + cover conditions with the sensor lying between the substrate and cover. The possible substrates were (1) None - firm flat back (2) Foam - flat foam (3) Curve - curved foam. The covers were (1) Fur - dense synthetic fur (2) None - no cover (3) Long - sparse long synthetic fibres (4) Short - dense short synthetic fibres. This forms 12 possible different conditions ie No Substrate - No Cover; Foam Substrate - Short Fur Cover; to name a few.

Since participants mainly performed gestures towards the centre of the sensor, in order to mitigate any noise from anchoring the sensor at the edges, we have trimmed the data and provided the centre 64 pixels with the cell labels as displayed below. Data capture was collected at a rate of 54 frames every second or 54 Hz. Each gesture was performed continuously for 10 seconds, though we have again trimmed the first and last second to provide the 8 seconds of continuous, uninterrupted capture. Each frame is the snapshot over the entire grid at a particular instant.

Data Structure

The data is formatted in a csv wherein each row represents one frame. Each row is labelled with the participant number, substrate condition, cover condition, gesture, and the 64 cells. The 432 frames for a set with the same participant number, substrate, cover and gesture (including the 64 cells per frame) forms the continuous sequential capture of that gesture over 8 seconds (54Hz * 8 sec = 432 frames). There may be value in acknowledging the order of each set of 432 rows as a time series.

The 64 cells are numbered by row-major order as displayed in Figure 1 below.

	1	2	3	4	5	6	7	8	
	9	10	11	12	13	14	15	16	
	17	18	19	20	21	22	23	24	
	25	26	27	28	29	30	31	32	
	33	34	35	36	37	38	39	40	
	41	42	43	44	45	46	47	48	
	49	50	51	52	53	54	55	56	
	57	58	59	60	61	62	63	64	

Figure 1. Cell labels as they correspond to sensor pixels

Missing Data

Please note that the following data points are not available.

From the Training Set~
Participant 3, Substrate-Foam, Cover-None, Gesture-Rub is unavailable.
Participant 5, Substrate-Foam, Cover-None, Gesture-ALL are unavailable.
Participant 6, Substrate-Foam, Cover-None, Gesture-Rub is unavailable.
Participant 10, Substrate-Curve, Cover-Fur, Gesture-Constant is unavailable.

From the Test Set~
Participant 2, Substrate-None, Cover-Fur, Gesture-Constant is unavailable.

Goal

Create a model from the Training Set that can predict the missing Gesture labels from the Test Set.

Any questions can be directed to Laura Cang cang@cs.ubc.ca at the University of British Columbia