Playful exploration of a robot's gesture production and recognition abilities

Jan de Wit[1], Bram Willemsen[1], Mirjam de Haas[1], Pieter Wolfert[1,2], Paul Vogt[1], Emiel Krahmer[1]

[1] Tilburg University [2] Radboud University Nijmegen

A robot's ability to communicate non-verbally can increase understanding between human and robot, and can help to maintain an engaging interaction. However, gestures in most studies with robots tend to rely on the designers' frame of reference, and their perspective on the robot's physical limitations. We propose a system, based on a gesture guessing game, where the robot learns many different examples of gestures from human players, which it can then replicate during subsequent interactions. Because players attempt to guess the meaning of the robot's gestures, the robot is able to identify those examples that best represent the target objects, given its limited expressive abilities. A first iteration of this exploratory study is set to take place with a SoftBank NAO robot, at the NEMO science museum in Amsterdam.

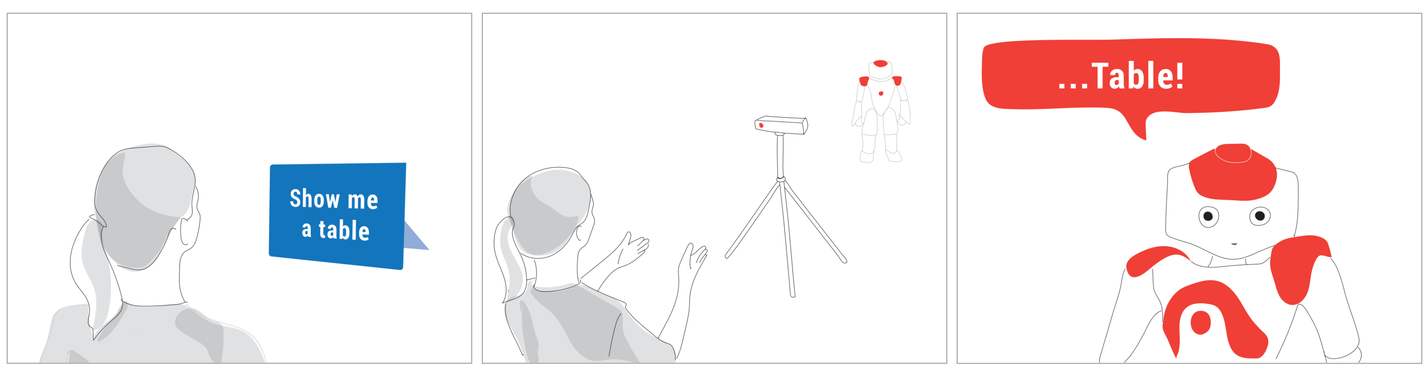

Figure 1: Example of the robot's turn to guess.

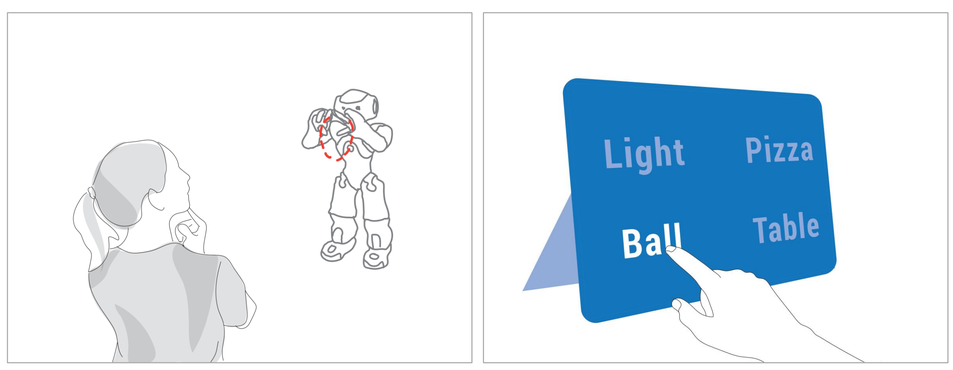

Figure 2: Example of the player's turn to guess.

Feature extraction

To detect similarity between gestures regardless of variations in the speed or size of the motion, we extract the *gist* of the gesture, based on the inflection points in the motion's trajectory [1]. We then find the nearest extreme point of the hand's position (Figure 3) and map this to a quadrant that is relative to the person's shoulder (Figure 4). This results in a description of a gesture that consists of a sequence of salient points from the motion's trajectory.

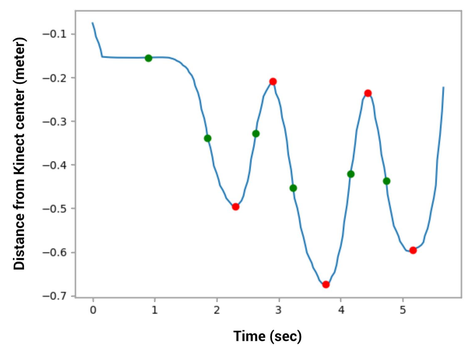

Figure 3: Inflection points (green) and peaks (red) of a motion.

Gesture recognition

Gesture recognition is currently implemented with a k-nearest neighbors approach, where the similarity between gestures is measured by aligning them using the Needleman-Wunsch algorithm with a custom scoring matrix.

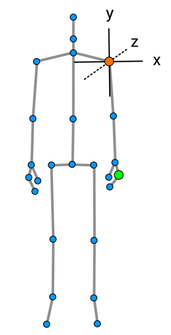

Figure 4: Quadrants of the hand position relative to the shoulder.

Gesture generation

Recorded examples of gestures for each object are clustered based on their similarity. Clusters may thus represent different strategies or variations within the gestures. The robot picks the next gesture to perform based on the weights assigned to each cluster and to individual examples within the cluster (through exploration and exploitation). Depending on whether the player guesses the object correctly or not, the weights are either increased or decreased.

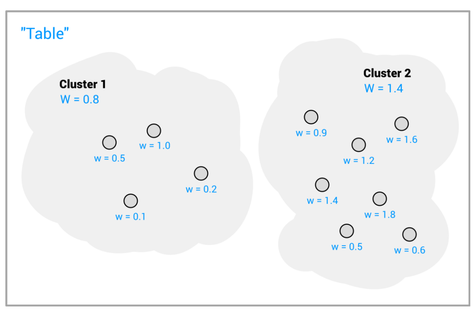

Figure 5: Example of clustered gestures with weights.

[1] Maria Eugenia Cabrera and Juan Wachs. A Human-Centered approach to One-Shot Gesture Learning. Frontiers in Robotics and AI 4 (2017).