Human-Robot Cooperation: Fast, Interactive Learning from Binary Feedback

Jawad Nagi*, Hung Ngo* (*Joint First Authors)
Jürgen Schmidhuber, Luca M. Gambardella, Gianni A. Di Caro
Dalle Molle Institute for Artificial Intelligence (IDSIA), CH-6928, Lugano, Switzerland
{jawad,hung,juergen,luca,gianni}@idsia.ch

Categories and Subject Descriptors

I.2.9 [**Robotics**]; I.4 [**Image Processing and Computer Vision**]; I.5 [**Pattern Recognition**]: Interactive Systems

General Terms

Algorithms, Methods, Design

1. VIDEO DESCRIPTION

This video presents results of a Swiss-funded project, focusing on *symbiotic human-robot cooperation.* We consider an *interactive learning* scenario and select the use of hand gestures to allow a human instructor to interact with a mobile robot. In our scenario, we use hand gestures to encode commands the robot will execute. At this aim, we define a vocabulary of $N = 6$ hand gestures (classes) that represent *finger counts* from 1 to 6. Using 6 hand gestures, the robot can perform 6 actions. In our case, each action represents that the robot must move to the *coloured landmark* corresponding to the recognized gesture class.

In this work, we exploit *feedback information* given by the human instructor to the robot during the learning and classification process. We use the notion of feedback as a means for improving the robot's learning and recognition accuracy over time, i.e., after a few interaction rounds with the human instructor. Our online learning and classification algorithm works in interaction rounds, and a single interaction round is described as follows. When the robot blinks its beacon, this signals the human that it is ready for interaction. Next, the robot acquires an observation (i.e., the human shows a gesture) using its onboard camera, and processes it (i.e., color-based segmentation and feature extraction) to compute the features of the hand silhouette using a pretrained Convolutional Neural Network (CNN). Using the resulting feature vector, the robot predicts the observation and produces an output gesture class Y, where $1 \leq Y \leq 6$. After prediction, the robot blinks its LEDs Y times with a RGB color corresponding to the predicted class. Then, the human provides a feedback to the robot to inform it, if the robot predicted the observation correctly or not. After feedback is given the interaction round terminates. If the gesture is recognized correctly, the robot performs the task associated with it, however if prediction is incorrect, then the gesture from the previous round is shown again to the robot and the next interaction round starts.

As feedback must be communicated to the robot, which amounts to another classification problem, we consider the use of *binary feedback* as it is more flexible than full feedback (i.e., the true class labels) and comes at a lower cost. Since binary feedback provides less information during communication than full feedback, this communication modality poses an implicit exploration-exploitation tradeoff—the robot may not necessarily want to select its current best prediction, in favor of the potentially most informative prediction for accelerating the learning progress. Our Upper Confidence Weighted Learning strategy (UCWL; [1]) provides an efficient solution to this setting. UCWL supports *online learning*, which neither assumes the data are independent and identically distributed, nor considers separate training and testing phases. It is formulated using an adaptive, large-margin update rule [2] for fast learning. To provide informed exploration-exploration trade-off, UCWL combines both score and uncertainty in making predictions.

The real-time demo implemented in this video makes use of our UWCL strategy, where binary feedback from the human is encoded using two separate *human actions* (i.e., right or wrong). When the human hides his hand (fist) behind his back, this action indicates that the robot recognized the gesture correctly. If the prediction was incorrect, then the human waves his hand vigorously to indicate a wrong prediction. Using these two predefined actions that can be reliably sensed in natural interactions (such as hiding or waving the hand) is quite feasible, while communicating a large (potentially very large) number of different classes requires more attention from, and is a larger burden upon, the human.

2. REFERENCES

[1] H. Ngo, M. Luciw, N. A. Vien, and J. Schmidhuber. Upper confidence weighted learning for efficient exploration in multiclass prediction with binary feedback. In *Proc. of the 23rd Intl. Joint Conf. on Artificial Intelligence (IJCAI)*, pages 2488–2494, 2013.

[2] J. Wang, P. Zhao, and S. Hoi. Exact soft confidence-weighted learning. In *Proc. of the 29th International Conference on Machine Learning (ICML)*, pages 121–128, 2012.

HRI'14, March 3–6, 2014, Bielefeld, Germany.
ACM 978-1-4503-2658-2/14/03.
http://dx.doi.org/10.1145/2559636.2559644.