Human-Swarm Interaction through Distributed Cooperative Gesture Recognition

Alessandro Giusti
Dalle Molle Institute for Artificial Intelligence (IDSIA)
Lugano, Switzerland
alessandrog@idsia.ch

Jawad Nagi
Dalle Molle Institute for Artificial Intelligence (IDSIA)
Lugano, Switzerland
jawad@idsia.ch

Luca M. Gambardella
Dalle Molle Institute for Artificial Intelligence (IDSIA)
Lugano, Switzerland
luca@idsia.ch

Stéphane Bonardi
Biorobotics Laboratory (BioRob)
École Polytechnique Fédérale de Lausanne (EPFL)
Lausanne, Switzerland
stephane.bonardi@epfl.ch

Gianni A. Di Caro
Dalle Molle Institute for Artificial Intelligence (IDSIA)
Lugano, Switzerland
gianni@idsia.ch

Categories and Subject Descriptors

I.2.9 [**Robotics**]; I.2 [**Distributed Artificial Intelligence**]: Coherence and coordination; C.2 [**Computer Communication Networks**]: Distributed applications; I.4 [**Image Processing and Computer Vision**]:

General Terms

Algorithms

1. VIDEO OVERVIEW

The video presents the first results of a Swiss-funded project focusing on *symbiotic peer-to-peer interaction and cooperation between humans and robot swarms*. As a first step, we considered human-swarm interaction, and selected the use of hand gestures to let a human communicate with a swarm of relatively simple mobile robots. In our scenario, a hand gesture encodes a command, that the swarm will execute. The robots that we used are the *foot-bots*, developed in the *Swarmanoid* project [1].

Hand gestures are a powerful and intuitive way to communicate, and do not require the use of additional devices. However, real-time *vision-based* recognition of hand gestures is a challenging task for the single foot-bot, due to its limited processing power and field of view. We investigated how to exploit *robot mobility*, *swarm spatial distribution*, and *multi-hop wireless communications*, to let the robots in the swarm: (i) implement a *distributed and cooperative sensing* of hand gestures, and (ii) robustly reach a *consensus* about a gesture.

The first step was to use 13 foot-bots to collect 70,000 hand gestures images representing finger counts (from 0 to 5) and five furniture-like shapes. With this data set we trained a *Support Vector Machine*, which is used by the robots for individual gesture classification and generation of an *opinion vector*, assigning a probability to each known gesture [3].

Following glove detection, the robots *move* to adapt their viewpoint for better sensing and maximize gesture information mutually gathered by the swarm as a whole. Using *multi-hop message relaying*, robots spread their opinions throughout the swarm. A *distributed consensus protocol* let the robots integrate and weight the different opinions.

When a robot has gathered enough evidence, it takes a *decision* for the hand gesture, and sends it into the swarm. Different decisions compete with each other, with the one assessed with the *highest confidence* that eventually wins.

The video shows two examples of command *execution*. For the furniture-like shapes, robots send their decisions to the *Roombot* system, where modular robots 'build' the furniture [2]. For the finger count, after receiving a 'two', the swarm splits in two groups moving in opposite ways.

2. ACKNOWLEDGMENTS

This research was supported by the Swiss National Science Foundation (SNSF) through the National Centre of Competence in Research (NCCR) Robotics (`www.nccr-robotics.ch`).

3. REFERENCES

[1] *Swarmanoid: Towards humanoid robotic swarms.* FET-OPEN project funded by the European Commission, `http://www.swarmanoid.org`.

[2] *Roombots: Modular robotics for adaptive and self-organizing furniture.* `http://biorob.epfl.ch/roombots`.

[3] J. Nagi, F. Ducatelle, G. A. Di Caro, D. Cireşan, U. Meier, A. Giusti, F. Nagi, J. Schmidhuber, and L. M. Gambardella. Max-pooling convolutional neural networks for vision-based hand gesture recognition. In *Proc. of the 3rd IEEE International Conference on Signal & Image Processing and Applications (ICSIPA)*, Kuala Lumpur, Malaysia, November 16–18, 2011.

.