Wisdom of the swarm for cooperative decision-making in human-swarm interaction

Article *in* Proceedings - IEEE International Conference on Robotics and Automation · June 2015

DOI: 10.1109/ICRA.2015.7139432

Wisdom of the Swarm for Cooperative Decision-Making in Human-Swarm Interaction

J. Nagi*, H. Ngo*, L. M. Gambardella, Gianni A. Di Caro

***Abstract*— Human-swarm interaction (HSI) is a developing field of research in which the problem of gesture-based control has been attracting an increasing attention, being at the same time a natural form of interaction and an effective way to point and select individual or groups of robots in the swarm. Gesture-based interaction usually requires vision-based recognition and classification of the gesture from the swarm. At this aim, existing methods for cooperative sensing and recognition make use of distributed consensus algorithms, which include for instance averaging and frequency counting. In this work we present a distributed consensus protocol that allows robot swarms to learn efficiently gestures from online interactions with a human teacher. The protocol also facilitates the integration of different consensus algorithms. Experiments have been performed in emulation using on real data acquired by a swarm of robots. The results indicate that effectively exploiting the collective decision-making of the swarm is a viable way to rapidly achieve good learning performance.**

I. Introduction

The general context of this work refers to the interaction between humans and multi-robot systems, namely, *human-swarm interaction* (HSI). In particular, the focus is on the use of *uninstrumented methods*, (i.e., methods that do not use sophisticated hardware devices from the human side), to perform the interaction, and in particular on the use of *hand gestures*. In our previous work [2], [3], [4], we have investigated the use of hand gestures for letting a human selecting robots and communicating mission commands to robot swarms. Since a requisite for such a form of interaction is that both individual robots and the swarm as a whole are able to recognize and classify the selected gesture vocabulary, initially we developed a batch supervised learning approach for vision-based cooperative sensing and recognition of hand gestures [2]. A *distributed consensus protocol* for fusing opinions between robot members, based on position-dependent estimates of reliability was introduced at the aim. A tunable parameter has been used in the protocol to balance the trade-off between the *time taken to reach consensus decisions* (convergence speed) and the *amount of evidence collected* from multiple observations by the swarm, for a single gesture that is *shown for a duration of time* to the swarm. In other words, multiple observations are gathered by every robot in the swarm for the same gesture over the course of some time, which relates to the amount of evidence collected for that gesture. Shifting attention towards *online* methods, in which the human plays the role of a teacher for letting the swarm learning the desired gestures, in [3] an online incremental framework for learning hand gestures from multiple viewpoints was presented, with consensus being built *offline* using a simple *averaging* technique. Following this, a *weighted-average method* was introduced in [4] using a bagging method, where swarm-level consensus was built offline by taking into account the confidence weight (online learning accuracy) of each robot.

Several challenges can be faced during consensus-building in swarm robotic systems, including difficulties arising in *deployment* and *distributed sensing* (e.g., due to illumination conditions, or bad sensing viewpoints). Distributed sensing systems, such as robot swarms, can concurrently gather perceptual information in parallel. This results in different observations (i.e., different inputs) being sensed and learned by different robots from multiple viewpoints (locations), hence different learned classifiers can be evolved with varying performance.

The motivation of this work comes from the idea of *combing expert advice* among different multiple learners, for building consensus decisions in ensemble-based learning systems (i.e., multi-classifier systems for distributed learning). The goal is to make more robust and effective the overall swarm consensus process by proposing a distributed consensus protocol that facilitates the integration of different consensus algorithms for interactive learning tasks in an *online* (incremental) setting [5], [6]. In relation to previous work [2], this work builds consensus decisions by fusing multiple observations from each single gesture. In other words, every robot in the swarm gathers a *single* observation from each gesture, which is then used by every robot to build a distributed consensus at the swarm-level. To assess the efficacy of our contribution in the context of distributed consensus protocol for interactive learning tasks, we provide a comparison with existing consensus methods, such as: *averaging* [3], [4], *frequency counting* (weighted-average) [2], aggregation methods, and learning methods based on expert advice [7].

The online setting that we consider is that of an interactive multi-class learning scenario where a human operator acts as an instructor (i.e., a *teacher*) for supervising the learning of hand gestures to a robot swarm. At the beginning of an *interaction round*, each robot acquires an observation (i.e., the instructor shows a gesture to the swarm) based on its field of view from the human. Next, robots in the swarm make a unified decision (i.e., build a *distributed consensus*) based on their individual predictions, by using their *local* (individual) observations and learning parame-

(*) These authors have equal contributions. All authors are with the Dalle Molle Institute for Artificial Intelligence (IDSIA), Lugano, Switzerland. Email: {jawad,hung,luca,gianni}@idsia.ch

ters. This consensus process outputs a class label, from a set of predefined classes. Depending upon the accuracy of the consensus decision (i.e., correct or incorrect) the human can choose to provide feedback to the swarm. If the swarm predicts the gesture correctly, the human does not provide feedback (e.g., hiding hands implicitly indicates a correct prediction), and the second interaction round starts. Alternatively, if the gesture is predicted incorrectly, the human needs to communicate the true label to the swarm as a *one-sided feedback*. This one-sided feedback can be in the form of a predefined number manually/electronically sent to the swarm. Upon receiving feedback, robots in the swarm update their individual learning models, as well as the aggregation weights for making consensus. Our robot swarm aims to obtain a swarm-level prediction (using the proposed distributed consensus protocol) as accurately as possible in each interactive learning round.

The work presented here provides two main contributions. Firstly, an integrated framework to allow robot swarms to learn efficiently from online interactions is presented, which combines online learning technologies with distributed consensus-building mechanisms. Secondly, a distributed consensus protocol is proposed that facilitates the integration of different consensus algorithms.

The rest of the paper is organized as follows. Section 2 discusses the related work in different domains. Section 3 presents the distributed consensus protocol and related technologies for each robot to learn. Section 4 reports experimental findings, and Section 5 presents concluding remarks.

II. Related Work

A. Gesture-based Interaction with Multi-robot Systems

The field of human-swarm interaction (HSI) is a relatively new area of research that aims at investigating techniques and methods suitable for interaction and cooperation between humans and robot swarms (i.e., multi robot systems). Due to negligence of research, HSI has received little attention [8], and not much is known about interaction issues related to *cooperative learning and decision-making mechanisms*.

Uninstrumented methods for interacting with robot swarms [1] (i.e., methods that require robot swarms to be within physical proximity of humans for sensing signals) are increasingly gaining attention, as illustrated in Figure 1, since they can overcome the limitations of teleoperated platforms and ease robot autonomy. Among several means of communication, hand gestures have been often selected [9], [10] as an efficient interaction modality to effectively communicate mission instructions to multi-robot systems in an uninstrumented way. Examples of use of hand gestures include instructing robots [11] and human-robot interaction (HRI) tasks [12].

With the goal of high-level human control of robot swarms, the work presented here uses hand gestures for letting humans express core spatial concepts. In simpler words, allowing humans to effectively select and command robots (see Section IV), as well as to indicate directions. However, in order to *use* hand gestures, the swarm needs to be able to recognize the gestures. Therefore, here the human is used a teacher supervising the swarm-level process of learning hand gestures.

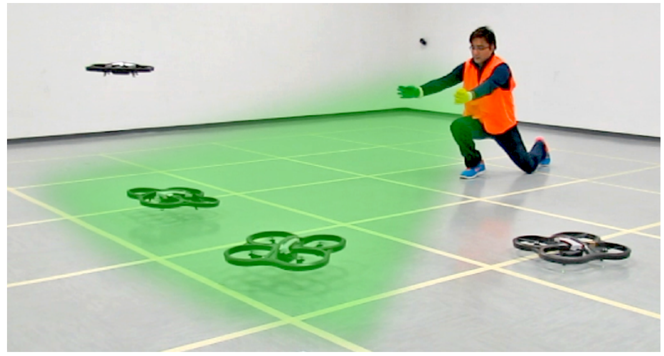

Fig. 1: A human operator selecting a group of 2 spatially-located robots from a swarm of 4 UAVs.

B. Consensus in Sensor Networks & Multi-robot Systems

Consensus algorithms serve as fundamental tools in wireless sensor networks. A swarm of robots connected through a (multi-hop) mobile ad-hoc network (MANET) can effectively sense information in a distributed wayand/or in parallel from multiple viewpoints, acting as an enhanced sensor. In order to exploit this capability, however the robots in the swarm need to reach a global agreement regarding the object of interest (e.g., a gesture), which define the need to develop *distributed consensus protocols* [13] that would ensure guaranteed convergence towards a common outcome. For decentralized data fusion tasks, in recent times a variety of distributed consensus strategies for applications with sensor networks [14] and multi-robot systems [15] have been investigated. A detailed overview of common consensus algorithms and applications is reported in [16]. Distributed consensus algorithms have been adopted in many robotic applications, including dynamic task allocation [17], collective map building [18] and obstacle avoidance [19].

Similarly, distributed camera networks have also used distributed consensus mechanisms for vision-based classification tasks [20], such as multi-camera surveillance and monitoring (i.e., intruder detection). Fusion of observations from multiple wide-baseline static cameras has been adopted in many perception applications including object classification [21] and pose estimation [22], with multiple viewpoints providing valuable inputs for reconstructing 3D information and overcoming the limits (i.e., occlusions, range) of each sensor. In [23] a face recognition system was presented, where multiple agents participated to build a fully-distributed classifier, that took advantage from the joint information contained in observations from multiple viewpoints. The most standard approach for fusing vision-based data from an array of distributed sensors requires computing features from sensed images, which are then aggregated and centrally classified, as adopted in [24] for human action detection.

C. Distributed Online Algorithms for Learning Consensus

Online learning in distributed sensing networks is emerging as a research topic for developing adaptive and intelligent routing protocols [25]. Having attracted considerable attention within the machine learning community, ensemble-based learning [26], namely *ensemble of classifiers* (i.e., multi-classifier systems) have been adopted in existing works, for distributed learning tasks that include: bagging, boosting and mixture of experts.

Ensemble methods such as bagging [27] and boosting [28], work by combining relatively *weak* learners, and have been extended to online versions [29]. *Mixture of experts* algorithms, on the other hand, aim to combine learners that are experts in *specific* input regions [30]. The distributed consensus protocol presented in this paper is motivated from the fact that different learning algorithms have different working mechanisms derived under different assumptions. Therefore, some learning algorithms may perform well in some domains, or some input regions, while others may perform poorly in other domains or regions. As a result, combining different classifiers can improve the overall classification performance of distributed consensus-building mechanisms. Data fusion strategies that have commonly been adopted in ensemble learning, include: algebraic operations (e.g., mean, median), weighting-based methods (e.g., weighted average, weighted majority voting) and entropy-based methods. Online ensemble learning methods [31], [32] provide an advantage compared to offline methods. This is because, as every sensor suffers a loss after learning the truth of its prediction, on every iteration (i.e., incremental update), the weights of the fused decision are updated taking into account the incremental performance of each sensor [33].

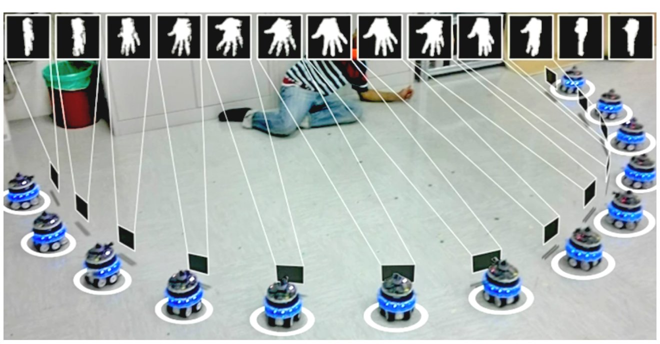

Fig. 2: A swarm of $N = 13$ robots sense a hand gesture from multiple points of view. Each robot sees an image from a different angle, and segments the image (based on color information) to obtain a hand silhouette. From such silhouettes, numerical features are computed that are used for learning and classification.

The research area most closely related to the setting considered in this paper is that dealing with issuing predictions with *expert advice* [7]. In this paradigm, a set of predictors, called experts, is given and the objective is to develop an adaptive aggregation rule to combine their predictions. Specifically, at each round, each predictor issue a prediction based on its *local* observation and expertise, followed by an overall consensus prediction made through an aggregation rule. After this, the true outcome is revealed (i.e., in our case the outcomes are class labels). Subsequently the aggregation rule is updated based on the performance of the experts, before advancing to the next round. The update rule aims to find the best aggregation, in terms of the cumulative prediction mistakes, as quickly as possible.

III. Consensus-driven Swarm Learning

In this section, we first describe our distributed protocol for swarm consensus, then we go into the technical details of vision-based learning employed on individual robots.

A. Distributed Protocol for Swarm-level Consensus

For letting a robot swarm to cooperatively sense, learn and recognize gestures from multiple viewpoints (locations), robots first need to be deployed in formations that optimize the spatial distribution of the swarm. In this context, robot swarms can make use of *distributed and parallel sensing* capabilities while *exploiting mobility* as shown in Figure 2. Our recent works in [2] present approaches for optimizing sensing coverage, maintaining wireless connectivity, and human localization [34], [1]. After a robot swarm is deployed (see Figure 4), the information sensed from an array of distributed and parallel sensors (i.e., the robots) is processed on-board of each robot. In order to efficiently fuse information from multiple robots into a single mutual swarm-level agreement, decentralized information fusion strategies such as distributed consensus-building mechanisms are required.

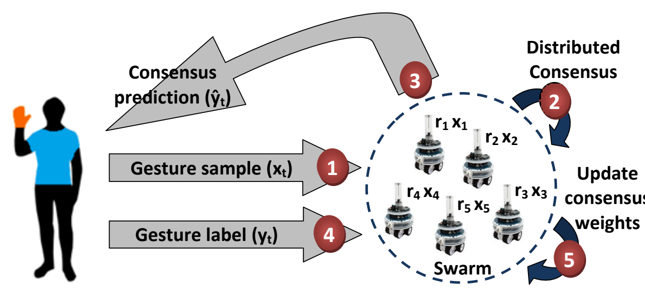

Fig. 3: The online interactive learning process supervised by a human, illustrating a *single interaction round.*

The online interactive learning setting introduced in this work is illustrated in Figure 3, where a human supervises the learning process of a robot swarm. This entire process is referred to as a single *interaction round*, where firstly the swarm predicts a given gesture observation using consensus-building strategies, and then incrementally learns this observation using the true label provided by the human. Our distributed consensus protocol, as described in Algorithm 1, is designed so that we can easily plug in a wide range of algorithms for multi-class prediction with expert advice [7]. Without loss of generality, we assume that the algorithm is running on robot #1 in the swarm; the other robots, numbered

$r \in \{1,2,\ldots,N\}$, are also running the same copy of the algorithm.

The algorithm works with any multiclass learner, whose output is a probability vector over the M possible classes, as denoted by a vector $\bar{\boldsymbol{m}}$ in the pseudocode of Algorithm 1. Note that the issued prediction is usually different for different robots in the swarm, as they observe different inputs for a given gesture. At each time step, to make distributed consensus over multiclass predictions, each robot r in the swarm exchanges its prediction probability vector $\bar{\boldsymbol{m}}_{t,r}$ with other robots.

Algorithm 1: HSI Distributed Consensus Protocol.

```
//Initialization
{c_{0,r} = 1}_{r=1}^{N} //Consensus weights
//Main interactive learning loop
for t = 1,2,... do
    Receive new observation x_t ∈ ℝ^d
    Output prediction probability vector m̄.
    // BEGIN swarm consensus phase
    Exchange m̄ among N robots in the swarm.
    // On receiving all m̄'s
    Compute consensus probability vector m̄_c.
    // END swarm consensus phase
    Output consensus label ŷ_t = argmax_{i=1,...,M} (m̄_c^i).
    Observe feedback y_t ∈ {1,...,M}.
    Update consensus weights {c_{t,r}}_{r=1}^{N}.
    Update M-class learning parameters using y_t.
end
```

Initially, to each robot r is assigned a unit consensus weight, $c_{0,r} = 1$. As the interactive learning process unfolds, the robots with more accurate predictions will have higher weights, while those with more mistakes will have diminishing weights. The consensus prediction is made based on the weighted prediction vector, $\bar{\boldsymbol{m}}_c$, where $\bar{m}_c^i = \sum c_{t,r} m_r^i$ for each class $i = 1 \ldots M$. After a true label is revealed, each consensus weight is updated based on the current loss of its learner. For instance, the Weighted Average Algorithm (WdAA; [35]) uses a multiplicative update $c_t = c_{t-1} e^{-\lambda_t}$, with λ_t the prediction loss on current example. In this paper, we use a square-loss function for multiclass prediction. The Weak Aggregating Algorithm (WkAA; [36]) employs a time- and loss-dependent update rule, $c_t = c_{t-1} e^{-\lambda_t/\sqrt{t}}$. The multiclass extension of Aggregating Algorithm (AA; [37]) uses a more involved update rule.

Besides these state-of-the-art aggregation rules, we also evaluated two other consensus algorithms. The first one employs a simple averaging strategy (over the predicted probabilities for each class) and does not need to maintain a consensus weight vector. The second one calculates the frequency of correct prediction for each class, and weights the prediction probability vector $\bar{\boldsymbol{m}}_{t,r}$ of each robot with this frequency, before averaging.

B. Single-robot Learning and Recognition

At the level of the individual robots, for the interactive task of visual learning and recognition, we follow a basic computer vision approach based on *color-based segmentation*, *feature extraction*, and *supervised learning*.

To simplify the recognition task without losing generality, we assume that human operators wear *tangible input devices* with "known characteristic colors" (e.g., colored gloves), as shown in Figure 1). We segment these color-based gestures by exploiting the characteristic colors of the gloves, as discussed in previous works [2], [3], [4]. From each of the segmented hand contours images, a set of $F = 40$ numerical shape features are computed, that correspond to a *feature vectors*, as reported in [2], [3], [38].

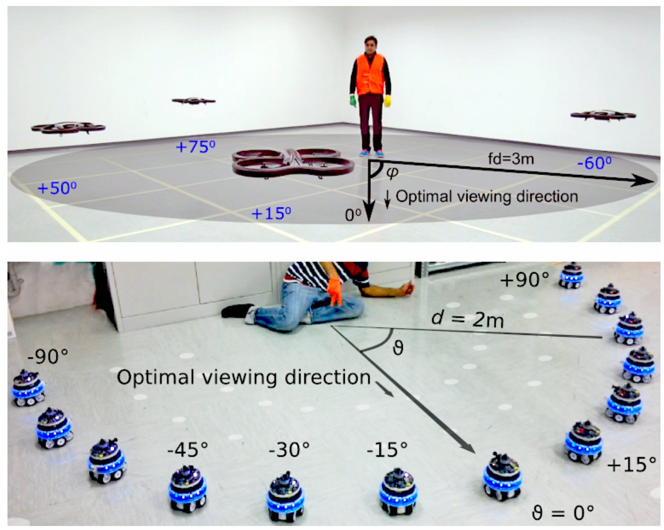

Fig. 4: Experimental setup for dataset acquisition. Top: Using $N = 4$ Parrot drones. Bottom: Using $N = 13$ foot-bots.

The feature vectors computed from the segmented hand contours are used as training instances for *multi-class online learning* and recognition tasks, by a Confidence-Weighted (CW; [39]) linear classifier. This is an *aggressive*, large-margin learning algorithm, which updates its learning parameters not only in rounds with prediction mistakes, but also in rounds with correct prediction but the margins are smaller than some threshold. Such effective update rules make CW the state of the art second-order online learning algorithm. The normalized prediction vector for classifying a sample from a trained CW multiclass-classifier, produces a prediction probability vector.

IV. Experimental Results

A. Experimental Setup

The system presented in this work addresses two fundamental problems in the domain of HSI, namely, *selecting robots from a swarm* (P-SEL) and *commanding selected robots* (P-COM). To be able to perform experiments, a swarm of robots was used to acquire large amounts of gesture images from different points of view. For each of the two problems (P-SEL and P-COM), one dataset was acquired:

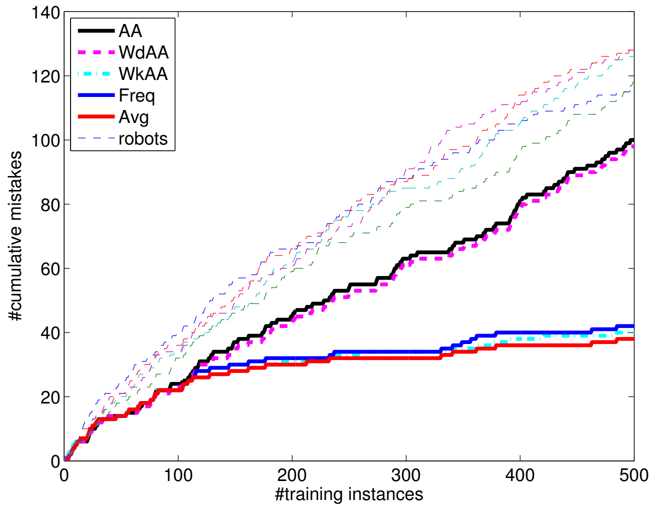

Fig. 5: A swarm size of $N = 5$ robots, where all are robots are deployed only at *good sensing positions*.

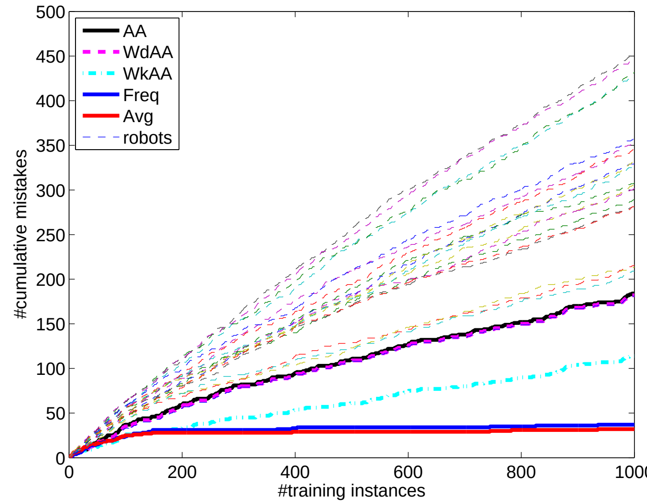

Fig. 7: A swarm size of $N = 20$ robots, where robots are deployed at different (good, bad and mixed) sensing positions.

- P-SEL Dataset [1]: Use of $N = 4$ UAVs positioned in formations as shown in Figure 4 top. The dataset comprises of $M = 4$ classes of gestures for selecting: individual, groups, individuals and groups, and all robots from a swarm respectively (see Figure 6 top).
- P-COM Dataset [2]: Use of $N = 13$ small UGVs as shown in Figure 4 bottom. The datasets comprises of $M = 6$ gestures, where gestures encoded as *finger-counts* from 0 to 5 represent numerical quantities (see Figure 6 bottom).

The datasets (acquired by a physical swarm) are tagged with labelled ground truth information (i.e., distance d, angle ϑ, gesture class y), and have been used for running quantitative *emulation experiments*: robot observations are sampled from these dataset of real images, and realistic simulations are built. Each observation x_t in the dataset is associated to a sensing position $(\vartheta, d)_x$ and a ground truth label (class) $y_t \in \{1, \ldots, M\}$, where M represents the number of classes. Interactions in which human shows a gesture are simulated as follows: each robot r is assigned to a randomly defined position $(\vartheta, d)_r$. Then, the observation x_t that robot r receives is randomly selected from the subset of dataset images, whose ground truth class is y_t, and whose associated position $(\vartheta, d)_x$ is as close as possible to $(\vartheta, d)_r$.

Fig. 6: Gesture categories (classes) in the acquired datasets. Top: P-SEL dataset. Bottom: P-COM dataset.

The results reported in the subsections below are emulated using different realizations of random variables (i.e., robot positions, observations sampled from the dataset, gesture sequences) from both datasets. Although we evaluated our protocol on other multiclass forecasting datasets [37] used in previous works, we do not present them here, as there is no significant difference, and the conclusions are still valid.

B. Swarm-level consensus accuracy

In this subsection, we report the performance of the swarm consensus in terms of online prediction accuracy. Except for the results plotted in Figures 5 and 7, which are the outcome of a single typical run in our experiments, the other results are calculated by averaging the results of 10 trials per experiment scenario. In Figures 5 and 7 the noticeable repeating pattern is that performance (in terms of number of mistakes) of consensus using the swarm is better than that of individual robots, and averaging outperforms other consensus algorithms.

Furthermore, as observed from Figure 5, the consensus performance of the swarm deployed in good positions is better than that of the case where individual robots are deployed in bad positions, as it was expected. Good sensing positions refer to locations that provide better quality of sensed information (e.g., facing directly in front of the human, central field of view, at a shorter distance from the human), whereas bad positions refer to locations with worse sensing conditions (e.g., rear field of view, partial occlusions, excessive distance from human).

The impact of distinctive deployment positions is further analysed in Figure 8. From the results, it is obvious that consensus performance decreases when robots are deployed at bad sensing positions, and vice versa for good positions. However, one interesting phenomenon is that, in bad deployment conditions, the improvement of averaging methods compared to all other consensus methods, is the highest.

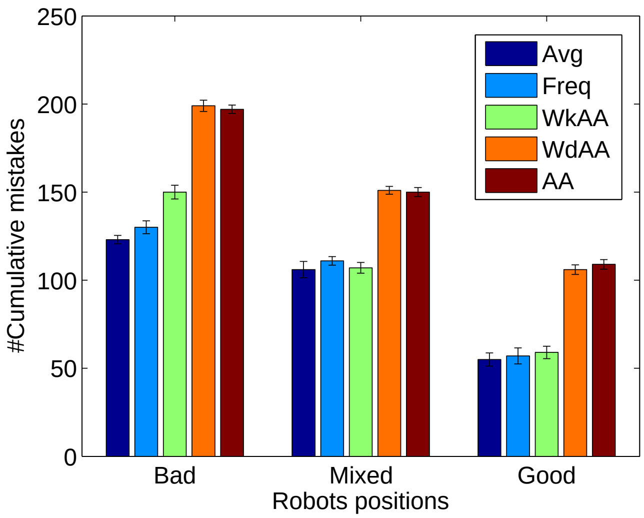

Fig. 8: Impact of different deployment positions (i.e., good, bad, and mixed) on the cumulative mistakes made by the swarm of $N = 13$ robots, after 500 interaction rounds.

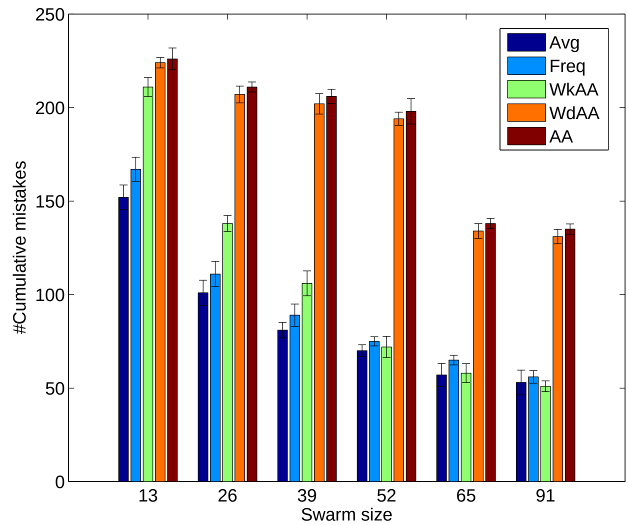

Fig. 9: Effect of swarm sizes $N = [13, 26, 39, 52, 65, 91]$ robots, on the cumulative mistakes made by the swarm, after 500 interaction rounds.

C. *Impact of swarm size on consensus*

The impact of different swarm sizes on consensus-building strategies is reported in Figure 9. Tests have been performed for swarm sizes of $N = [13, 26, 39, 52, 65, 91]$ robots. The main trend observable is that, larger robot swarms yield less mistakes in the interactive learning process. As not much significant difference is present when using $N = [65, 91]$ robots, this indicates that swarms with approximately $N = [60, 70]$ robots can provide similar performance as compared to swarms of 100 robots. In addition, for large swarm sizes (i.e., $N > 50$ robots) the WkAA approach, slightly outperforms simple averaging.

D. *Effect of swarm mobility on consensus*

The effect of using swarm mobility on the performance of consensus-building strategies is reported in Figure 10, after 500 interaction rounds. Swarm mobility is emulated by switching (changing) robot deployment positions after every $[1, 10, 50, 100]$ interaction rounds. The general conclusion is that, when robot positions are switched more frequently during interactive learning (i.e., smaller mobility step sizes), the performance of the swarm as a whole degrades for all consensus algorithms. However, even if robot positions are changed after every interaction round (i.e., after a step size of 1), the consensus performance of the swarm is still good, and robust to different deployment positions.

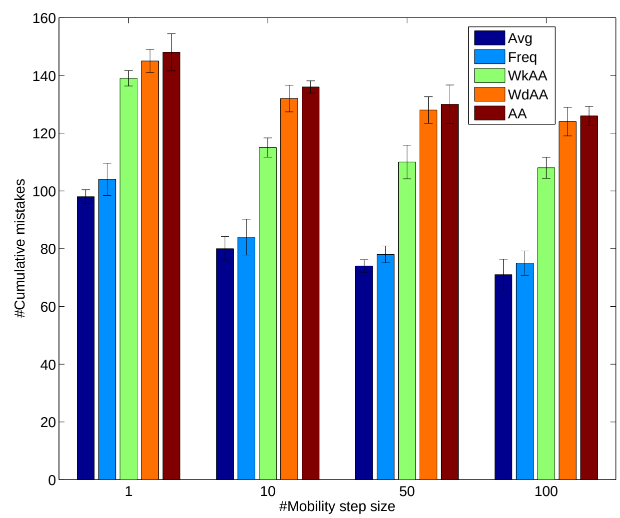

Fig. 10: Effect of swarm mobility after $[1, 10, 50, 100]$ interaction round(s) vs. the cumulative mistakes, after 500 interaction rounds, made by the swarm of $N = 13$ robots, deployed as shown in Figure 4 top.

V. Conclusions

This paper has presented a distributed consensus protocol that allows robot swarms to learn efficiently from online interactions, and facilitates the integration of different consensus algorithms. The techniques and strategies reported in this work have been experimentally evaluated on two different image datasets of gestures, acquired by ground robots and flying drones.

Multiple findings have resulted as a consequence of this work. Firstly, the performance of consensus decisions made by a swarm of robots outperforms the decisions of an individual robot (even robots at the best sensing positions). Secondly, small and medium size swarms give the best performance with consensus-based averaging. Thirdly, for large swarm sizes learning methods based on expert advice (i.e., WkAA) have shown slightly better performance than averaging methods. Lastly, the performance of the swarm based on distributed consensus is robust to the mobility of the swarm (i.e., when robots switch positions after interaction rounds), swarm size, and deployment positions of the robots.

ACKNOWLEDGMENTS

This research was partially supported by the Swiss National Science Foundation (SNSF) through the National Centre of Competence in Research (NCCR) Robotics (www.nccr-robotics.ch). The authors participate to Project 4: *Distributed Robotics*, sub-project 4.5, *Symbiotic human-swarm cooperation*.

REFERENCES

[1] J. Nagi, A. Giusti, L. M. Gambardella, and G. A. Di Caro, "Human-swarm interaction using spatial gestures," in *Proc. of IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2014, pp. 3834–3841.

[2] A. Giusti, J. Nagi, L. M. Gambardella, and G. A. Di Caro, "Cooperative sensing and recognition by a swarm of mobile robots," in *Proc. of 25th IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2012, pp. 551–558.

[3] J. Nagi, H. Ngo, A. Giusti, L. M. Gambardella, J. Schmidhuber, and G. A. Di Caro, "Incremental learning using partial feedback for gesture-based human-swarm interaction," in *Proc. of 21st IEEE Intl. Symposium on RO-MAN*, 2012, pp. 898–905.

[4] J. Nagi, A. Giusti, F. Nagi, L. M. Gambardella, and G. A. Di Caro, "Online feature extraction for the incremental learning of gestures in human-swarm interaction," in *Proc. of IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2014, pp. 3331–3338.

[5] H. Ngo, M. Luciw, N. Vien, and J. Scmidhuber, "Upper confidence weighted learning for efficient exploration in multiclass prediction with binary feedback," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*. AAAI Press, 2013, pp. 2488–2494.

[6] H. Ngo, M. Luciw, N. Vien, J. Nagi, A. Förster, and J. Schmidhuber, "Efficient interactive multiclass learning from binary feedback," *ACM Transaction on Interactive Intelligent Systems*, 2014.

[7] N. Cesa-Bianchi and G. Lugosi, *Prediction, learning, and games*. Cambridge University Press, 2006.

[8] S. T. Hayes and J. A. Adams, "Human-swarm interaction: Sources of uncertainty," in *Proc. of Intl. Conf. on HRI*, 2014, pp. 170–171.

[9] G. Podevijn, R. O'Grady, Y. S. G. Nashed, and M. Dorigo, "Gesturing at subswarms: Towards direct human control of robot swarms," in *Proc. the 14th Conference TAROS*, 2013.

[10] S. Pourmehr, V. M. Monajjemi, R. T. Vaughan *et al.*, "You two! take off!: Creating, modifying and commanding groups of robots using face engagement and indirect speech in voice commands," in *Proc. of 26th IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2013, pp. 137–142.

[11] S. Mitra and T. Acharya, "Gesture recognition: A survey," *IEEE Transactions on Systerms, Man, and Cybernetics–Part C*, vol. 43, no. 3, pp. 311–324, 2007.

[12] M. A. Goodrich and A. Schultz, "Human-robot interaction: a survey," *Foundations & Trends in HCI*, vol. 1, no. 3, pp. 203–275, 2007.

[13] T. Li, M. Fu, L. Xie, and J.-F. Zhang, "Distributed consensus with limited communication data rate," *IEEE Transactions on Automatic Control*, vol. 56, no. 2, pp. 279–292, 2011.

[14] R. Olfati-Saber and J. Shamma, "Consensus filters for sensor networks and distributed sensor fusion," in *Proc. of the 44th IEEE Conference on Decision and Control (CDC)*, 2005, pp. 6698–6703.

[15] A. Fagiolini, M. Pellinacci, G. Valenti, G. Dini, and A. Bicchi, "Consensus-based distributed intrusion detection for multi-robot systems," in *Proc. of IEEE ICRA*, 2008, pp. 120–127.

[16] R. Olfati-Saber, J. Fax, and R. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[17] G. Das, T. McGinnity, S. Coleman, and L. Behera, "A fast distributed auction and consensus process using parallel task allocation and execution," in *Proc. of Intl. Conf. on IROS*, 2011, pp. 4716–4721.

[18] R. Aragues, J. Cortes, and C. Sagues, "Distributed consensus on robot networks for dynamically merging feature-based maps," *IEEE Transactions on Robotics*, vol. 28, no. 4, pp. 840–854, 2012.

[19] I. Navarro and F. Matía, "A framework for the collective movement of mobile robots based on distributed decisions," *Robotics and Autonomous Systems*, vol. 59, no. 10, pp. 685–697, 2011.

[20] H. Aghajan, C. Wu, and R. Kleihorst, "Distributed vision networks for human pose analysis," in *Signal Processing Techniques for Knowledge Extraction & Information Fusion*, D. Mandic, M. Golz, A. Kuh *et al.*, Eds. Springer, 2008, pp. 181–200.

[21] W. Schriebl, T. Winkler, A. Starzacher, and B. Rinner, "A pervasive smart camera network architecture applied for multi-camera object classification," in *Proc. of Intl. Conf. on Dist. Smart Cameras*, 2009.

[22] A. Jorstad, D. DeMenthon, I. Wang, and P. Burlina, "Distributed consensus on camera pose," *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2396–2407, 2010.

[23] P. F. E. Kokiopoulou, "Distributed classification of multiple observation sets by consensus," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 104–114, Jan. 2011.

[24] C.-H. Yu and R. Nagpal, "Self-adapting modular robotics: A generalized distributed consensus framework," in *Proc. of IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2009, pp. 1881–1888.

[25] G. A. Di Caro, M. Kudelski, E. Flushing, J. Nagi, I. Ahmed, and L. M. Gambardella, "On-line supervised learning of link quality estimates in wireless networks," in *Proc. of 12th IEEE/IFIP Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net)*, 2013, pp. 69–76.

[26] R. Polikar, "Ensemble based systems in decision making," *IEEE. Circuits and Systems Magazine*, vol. 6, no. 3, pp. 21–45, 2006.

[27] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.

[28] J. Quinlan, "Bagging, boosting, and C4.5," in *Proc. of 13th Intl. Conf. on Artificial Intelligence*, 1996, pp. 725–730.

[29] N. Oza, "Online bagging and boosting," in *Proc. of IEEE Intl. Conf. on Systems, Man, and Cybernetics*, 2005, pp. 2340–2345.

[30] R. Jacobs, M. Jordan, S. Nowlan, and G. Hinton, "Adaptive mixtures of local experts," *Neural Computation*, vol. 3, no. 1, pp. 79–87, 1991.

[31] R. Polikar, L. Upda, S. S. Upda, and V. Honavar, "Learn++: an incremental learning algorithm for supervised neural networks," *IEEE Trans. on SMC—Part C*, vol. 31, no. 4, pp. 497–508, 2001.

[32] D. Parikh and R. Polikar, "An ensemble-based incremental learning approach to data fusion," *IEEE Transactions on Systems, Man, and Cybernetics–Part B*, vol. 37, no. 2, pp. 437–450, 2007.

[33] R. Elwell and R. Polikar, "Incremental learning of concept drift in nonstationary environments," *IEEE Transactions on Neural Networks*, vol. 22, no. 10, pp. 1517–1531, 2011.

[34] J. Nagi, G. A. Di Caro, A. Giusti, and L. M. Gambardella, "Learning symmetric face pose models online using locally weighted projectron regression," in *Proc. of 21st IEEE Intl. Conf. on Image Processing (ICIP)*, 2014, pp. 1400–1404.

[35] J. Kivinen and M. Warmuth, "Averaging expert predictions," in *Computational Learning Theory*. Springer, 1999, pp. 153–167.

[36] Y. Kalnishkan and M. Vyugin, "The weak aggregating algorithm and weak mixability," in *Annual Conference on Learning Theory*. Springer, 2005, pp. 188–203.

[37] V. Vovk and F. Zhdanov, "Prediction with expert advice for the Brier game," *Jour. of Mach. Learn. Res.*, vol. 10, pp. 2445–2471, 2009.

[38] G. A. Di Caro, A. Giusti, J. Nagi, and L. M. Gambardella, "A simple and efficient approach for cooperative incremental learning in robot swarms," in *Proc. of Intl. Conf. on Advanced Robotics (ICAR)*, 2013, pp. 1–8.

[39] M. Dredze, K. Crammer, and F. Pereira, "Confidence-weighted linear classification," in *Proc. of the 25th ICML*, 2008, pp. 264–271.