Towards Human-Friendly Efficient Control of Multi-Robot Teams

Adrian Stoica[1], Theodoros Theodoridis[2], David F. Barrero[3], Huosheng Hu[2], and Klaus McDonald-Maier[2]
[1]Jet Propulsion Laboratory, Pasadena, California, USA
[2]University of Essex, School of CSEE, Colchester, UK
[3]Universidad de Alcalá, Alcalá de Henares, Spain
[1]*adrian.stoica@jpl.nasa.gov*
[2]*{ttheod,hhu,kdm}@essex.ac.uk*
[3]*david@aut.uah.es*

***Abstract*— This paper explores means to increase efficiency in performing tasks with multi-robot teams, in the context of natural Human-Multi-Robot Interfaces (HMRI) for command and control. The motivating scenario is an emergency evacuation by a transport convoy of unmanned ground vehicles (UGVs) that have to traverse, in shortest time, an unknown terrain. In the experiments the operator commands, in minimal time, a group of rovers through a maze. The efficiency of performing such tasks depends on both, the levels of robots' autonomy, and the ability of the operator to command and control the team. The paper extends the classic framework of levels of autonomy (LOA), to *levels/hierarchy of autonomy characteristic of Groups* (G-LOA), and uses it to determine new strategies for control. An *UGV-oriented command language* (UGVL) is defined, and a mapping is performed from the human-friendly gesture-based HMRI into the UGVL. The UGVL is used to control a team of 3 robots, exploring the efficiency of different G-LOA; specifically, by (a) controlling each robot individually through the maze, (b) controlling a leader and *cloning its controls to followers*, and (c) controlling the entire group. Not surprisingly, commands at increased G-LOA lead to a faster traverse, yet a number of aspects are worth discussing in this context.**

***Index Terms*—Multi-robot control, human-robot interfaces, robot language, sliding autonomy, adaptive autonomy, autonomy of robot teams, group levels of autonomy**

I. Introduction

As the cost of robotic platforms continues to reduce, an increasing number of applications involve multiple robots. The efficiency of performing tasks with robotic teams (as well as for mixed teams of robots and humans) depends on both, the levels of autonomy, and the ability of humans to command and control the team; in particular through efficient interfaces [20]. The transition from the current state of the art that requires several human operators for the control a single robot, to having a single human control multiple robots, has been identified as one of the main challenges in robotics.

For a rich communication with robots, a human-friendly robot-oriented language is needed to adequately specify a wide range of control commands from high-level objectives, to direct commands (task goals). For high bandwidth, reduced attention burden and fatigue, as well as increased mobility and capability to handle various objects, an operator should have a freedom of movements leaving the hands unencumbered. Future Multi-robot Operator Control Units (MOCU) will likely be using natural human-multi-robot interfaces (HMRI) such as gesture, speech, etc. Ideally the level of effort in coordinating robots should not be higher than coordinating fellow humans.

This paper explores new modalities of increasing the efficiency of controlling groups of robots. The application scenario is an emergency evacuation in which, under the instruction of a coordinator, a transport convoy of unmanned ground vehicles (UGV) would traverse, in shortest time, an unknown terrain. To increase control efficiency, we explore means to control multiple robots at once, we propose a UGV-oriented language (UGVL), and a mapping between a natural hand gesture-based HMRI into the UGVL. The hand gestures are recognized with a 16-channel EMG sensor array, the JPL's BioSleeve, donned on the forearm.

A. Group Levels of Autonomy

The current categorizations of Levels of Autonomy (LOA), introduced by Sheridan [17], refined in consequent formulations, such as NASA SMART, and nuanced in NIST ALFUS [27], along dimensions of Human Independence,

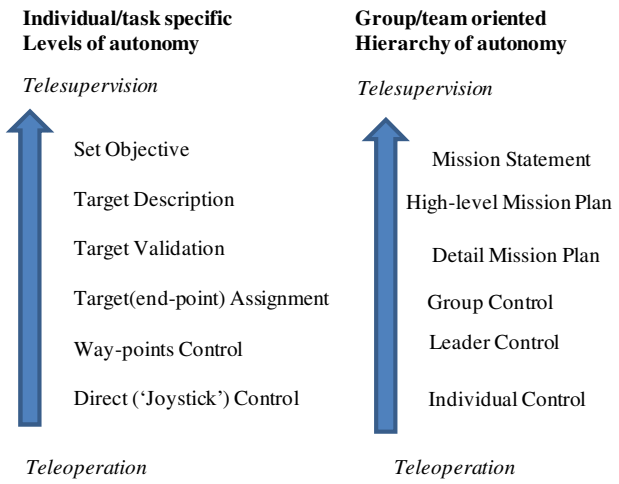


Fig 1. Extension of individual levels of autonomy to a group/team autonomy hierarchy – G-LOA.

Mission Complexity and Environmental Complexity, do not have a dimension that allows a constructive use of the model for developing strategies in multi-robot control. Aiming to fill this gap we propose a dimension of classification that specifically deals with the control of groups of robots, denoted as group-levels of autonomy (G-LOA). An example for vehicles that have to traverse a region, towards target end points, is illustrated in Fig. 1.

G-LOA has at its lowest level of autonomy (i.e. highest degree of teleoperation) the individual control of each member of the group. The next level is the leader control, a teleoperation of the leader, with an assumption of means of control for followers. Higher in the hierarchy comes the group control with different granularity (subgroup and group). Plans can be provided at various levels of detail, for the entire group, with specificities for special group members. The highest autonomy following this description is the mission statement. This extension is used in this paper to develop strategies of traversing the terrain with groups of robots of various sizes.

B. Related Work

In [13] algorithms and display concepts allow soldiers to efficiently interact with a robotic swarm, that is participating in a representative convoy mission. The focus there is on keeping soldiers cognizant of swarm operations through an interface that allows them to monitor status and/or institute corrective actions. [14] focuses on the required flexibility of group formations when traversing from one point to the next, in ground-based military maneuvers. The work is done in simulations. For a human-led team of semi-autonomous agents, a certain level of awareness demonstrated by the agents regarding the quality of the formation. Through the use of a Multi-Robot System (MRS), this work combines leader-follower principles augmented by an assistive formation maintenance (AFM) method, used to improve formation by keeping and demonstrating a *formation-in-motion* concept. The goal is to provide a military application that allows a soldier to efficiently teleoperate a semi-autonomous MRS capable of keeping formation in a cluttered environment.

In the context of gesture-based HMRI the Swarmanoid project [18], and its successor NCCR Robotics projects, address the gesture interaction for swarm commands [10]. The focus there was on gesture recognition from vision, distributed on the robots. The work employs robots that recognize through vision a number of finger gestures, observed from different viewing angles, where the interpretation of classes is fused by a single robot. By associating gestures with commands to different robots (6 to 13), the control performance is simplified by splitting a group of robots in 2.

Significant work in human-robot interaction and control of robot teams has been done by Goodrich and collaborators (for a survey see [24]). In [25] they refer to a team-level autonomy. One of their experiments tested terrain reversibility of three robots (they employ a 'playbook'-style management [26]).

The rest of the paper is structured as follows. Section II defines the UGV-oriented language (UGVL), and the hand gesture-based HMRI. Section III presents experiments in which a human uses hand gestures and the UGVL to control a group of rovers through a maze, exploring the efficiency of strategies based on different G-LOA levels. Section IV discusses the results and plans of future work.

TABLE I

UGV-L PRIMITIVES: A SET OF COMMANDS DEFINING THE VOCABULARY FOR CONTROLLING A TEAM OF ROBOTS/UGVS

Command Class	**Command (example, incomplete for numbers, compass, etc)**					
Team Selection	Entire team	Sub-team	Individual	-	-	-
Role	Leader	Deputy	Follower	Target	Friend	Enemy
Actions	Move	Transmit Video	Record Video	Launch	Clone	-
Action Step	Go/Start	Stop	Wait	Execute	Cancel	-
Degree /increment	minimum	A bit	Quite a bit	More	Much more	maximum
Direction (relative to heading)	Forward	Back	Left	Right	Half- Right	Half-left
Distance	Close	Far	Precise (unit)	-	-	-
Direction (absolute, pointing)	There (point to space)	To object (point to object)	In that direction (point direction)	-	-	-
Turn (relative, absolute, style)	To the right	To the left	O'clock	Compass	Sharp	Smooth
Scale	Tenth of	Unitary	Tens of	Fraction of	Times	-
Units	m	s	m/s	-	-	-
Formation	Encircle	Y	R-Edge	L-Edge	V	Zig-zag
Speed	Slower	Faster	% (may be increment)	m/s	Min	Max
Number	0	1	2	3	4	5
Compass	N	NE	E	SE	S	SW
Behaviors	Approach	Patrol	Explore	Circle	Attack	Retreat

II. A HUMAN-FRIENDLY HMRI INTERFACE: LANGUAGE AND GESTURE-BASED COMMANDS

In this section we describe a robot-oriented, yet human-friendly language, UGVL presented in Table I, and a gesture interface that enables an operator to command efficiently a team of robots performing a task. In order to increase the expressiveness of the interface, the language allows the composition of simple symbols, ie. gestures, to build composite constructs named sentences that describe complex behaviors. In this way, one reaches a high representation power, which allows very fine control of the team, sub-teams, or individuals in teams, while keeping a limited number of symbols.

A. A command language for unmanned ground vehicles

We propose a UGVL command language which includes the following classes of commands, summarized in Table I. Below, definitions of the major classes are listed analytically, yet leaving others which are more straight-forward.

- *Team (Group) Selection*: Selects the team/group constituency; can be the whole team/group, a sub-team/group, or individual robots. Indexing is needed to identify the sub-team/group or individual robot.
- *Role*: Defines roles, can be the leader, the deputy which is next in line if leader is canceled, or a follower; can be targets, friend or enemy; etc.
- *Formation*: The formation that the robots are instructed to move into.
- *Speed*: It selects the velocity of the selected robot or robots. The speed is specified in relative terms (*slower* or *faster)*, with *incremental* increase/decrease as percentage of the maximum speed, in absolute terms as percentage of the maximum speed, in absolute value such as miles per hour, or the min or max accepted.
- *Numbers*: 0 to 10, are parameters which depend on a category selected; for example, time in seconds.
- *Compass*: Indicates direction e.g. North East (NE).
- *Behaviors*: Include a predetermined or learned sequences of things UGV-s can do part of a mission scenario.

B. Gestures: primitives for the language

The main input device for this research was the JPL's BioSleeve [22, 23], which is a hand gesture recognition system with a 16-channel EMG sensor array donned on the forearm; in one of its versions it also includes IMUs (Fig. 2).

The BioSleeve recognizes 28 simple gestures (Fig. 3) with correct classification rate (CCR) more than 97% (this particular classification did not use IMU information, which is mostly used in complex/dynamic gestures). The signals acquired and filtered by the BioSleeve offline for the later gesture recognition process. In the current implementation, static gestures are classified using the EMG signals in a Support Vector Machine (SVM).

After donning the BioSleeve, the user completes a 2–5 minute calibration exercise, which collects data in each gesture to train the classifiers. Details on the use of BioSleeve are given in [9, 21, 22].

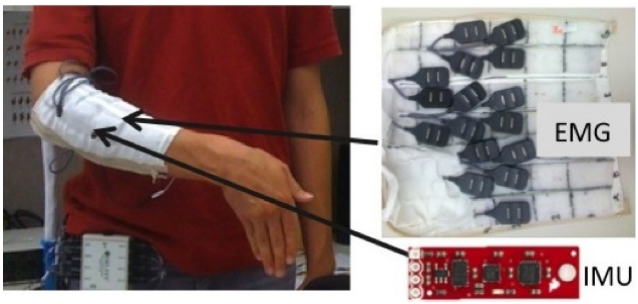


Fig. 2. JPL BioSleeve mounted on the forearm, with EMG and IMU sensors.

C. Mapping gestures to a UGVL

For most languages that cover rich forms of expression it is common that the number of primitives used is comparatively small. Letters of an alphabet can be only a few, yet a large vocabulary of words can be created, and words can be further concatenated (based on grammars) to create meaningful sentences. Similarly, with a relatively small number of gestures one can form richer composite structures for the commands required to control the team of robots. A combination of gestures in a sequence allows obtaining an arbitrarily large number of commands. This idea, proposed in [9] is extended here to define a UGVL, and a mapping between the human-friendly gesture-based HMRI into the UGVL.

In the simplest mapping between gestures and commands, we composed two or more gestures to codify a command. In order to make this interface easy to use, we grouped similar commands and identified them with a certain gesture. Thus, the first gesture identifies the commands class, and the second one provides the specific command to execute. Some commands require additional information, such as duration time (in seconds), or the robot index. To deal with the risk of a command misinterpretation (e.g. from gesture CCR of 97%) a common practice is to have the recognition system confirm the recognized command, which if wrong can be canceled and re-expressed, at the price of a delay associated with the acknowledgement/validation procedure. An optimal allocation of gestures to commands has to consider the existence of a domain-specific gesture command language.

Fig. 3 Hand gestures recognized by JPL BioSleeve, an embedded EMG sensor array.

TABLE II

A SUBSET OF COMMANDS USED IN THE EXPERIMENTS FOR ROBOT CONTROL

Gesture	Description	Command					
		G2	G7	G8	G9	G10	G11
G13	Robot	all	group	single	-	-	-
G6	Index	0	1	2	3	4	5
G16	Control	-	direct	auto	-	-	-
G25	Action	stop	forward	backward	left	right	-
G18	Speed (%)	0	20	40	60	80	100
G5	Time (sec)	0	1	2	3	4	5
G20	Cloning	false	true	-	-	-	-

D. Command composition

The use of function selectors provides a richer class of control. The software recognizes the first gesture, and depending on the meaning associated to it, it interprets the second gesture differently. With respect to Table II, used as an experimental baseline in UGVL, G13+G2, means 'select the entire group of robots', while by contrast G13+G7, G6+G7 means 'select sub-group 1'. Furthermore, even thought when the proposed language is able to represent a wide set of commands, in order to accomplish the task in the scenario under consideration, it is necessary to represent not only a command, but a composition of commands that we name it a *sentence*. A sentence is synthesized by a BNF gesture-grammar in an expression-like form.

Example BNF gesture sentences:

```
<exp>::=<group><index><control><action><speed>
<time><clone>
<exp1>::=com<G13 G7> // Select robot sub-group
    <G6  G7>  // Choose sub-group 1
    <G16 G8>  // Perform automatic control
    <G25 G7>  // Move forward
    <G18 G8>  // Speed at 40%
    <G5 G11>  // Travel for 5 sec
    <G20 G2>  // Do not use cloning
<exp2>::=com<G13 G8> // Select single robot
    <G6  G8>  // Choose robot 3
    <G16 G8>  // Perform automatic control
    <G25 G10> // Turn right
    <G18 G7>  // Speed at 20%
    <G5  G8>  // Travel for 2 sec
    <G20 G2>  // Do not use cloning
```

III. EXPERIMENTS AND DISCUSSION

The scenario requires driving a team of robots through a simple maze, with the goal of minimizing the traverse time for the entire team. We tested the efficiency of the language (UGVL), and investigated the most effective G-LOA strategy to accomplish the goal. The robots used (Brookstone AC13) only supported an adjustable speed and a 2-DOF heading and had no odometry. The hardware limitations impacted the tasks' setup and the design of the interface.

Three sets of experiments were run, which correspond to different levels in G-LOA hierarchy (as defined in Fig. 1):

- *Individual control* (Teleoperation of individuals robots)
- *Leader control* (Teleoperation of a leader and *cloning* of its behavior to its followers)
- *Group control* (Teleoperation of all the robots).

In each case, corrections were made at individual robot level. Snapshots of the robots in the three experiments are illustrated in Fig. 4. The first set of images (Fig. 4(a)) shows the individual control, the operator's sequence of gestures driving one robot at a time through the maze. Fig. 4(b) illustrates leader control, in which case the leader is driven by the operator, and its commands cloned. *By cloning we mean that the sequence of commands applied to the leader robot gets ported to other robot (in some respect the leader is in fact a teleoperated scout on which commands are tried and then duplicated on others).* Fig. 4(c) illustrates the group control, where all the team members receive the same commands from the operator (however, corrections were applied individually).

The pictures point out cases where corrections were needed for the direction of movement. Due to various influencing factors, such as different level of battery, the movement of the clone ends up slightly different, and may require compensation. While the effect is an artifact of the hardware limited platforms, it is useful for simulating a real-world effect, which may appear due to the terrain non-uniformity and specific path the rovers take, with different friction or slippage. The time for traverse for the three cases is shown in Table III.

TABLE III
TIME REQUIRED TO ACCOMPLISH THE SCENARIO ON 3 LEVELS OF G-LOA

G-LOA Level	Traverse modality	Total Time
Individual control	Total time for all 3	96 sec
Leader control	Followers clone leader's control	79 sec
Group control	All robots moved	58 sec

The results indicate, perhaps not surprisingly, that commands at increased level of hierarchy in G-LOA lead to a faster traverse. This is in agreement with the results in [21], which addressed a similar problem of motion of a team of 2 robots through a maze, including serial (one by one), parallel (all at once, similar to group), and manual control of one and several degrees of LOA in the other. Parallel control (our group control) turned out to be the fastest method, and the one that generated the lower perceived workload to the operator.

IV. CONCLUSIONS AND FUTURE DIRECTIONS

The paper demonstrated an efficient methodology of controlling a group of ground robots using a UGV command gesture-based language, and the construction of gesture-based grammatical expressions designated for robot commanding and control. We tested a level of hierarchy/autonomy framework for determining new control strategies, which has successfully been applied in scenarios for guiding a single and groups of robots, as well as cloning a route performed by single robot to multiples. Findings from this work revealed that a higher efficiency in terms of shorter time to execute the mission is obtained by controlling at a higher level of G-LOA.

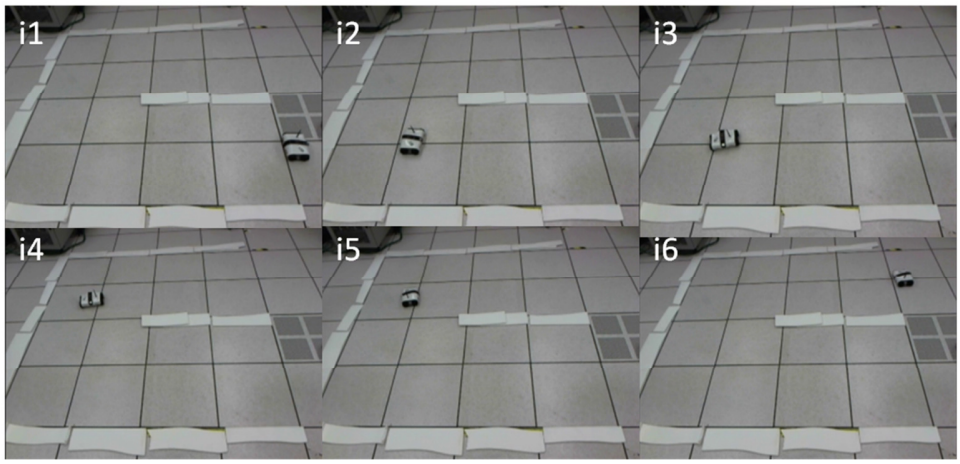
Our future work will continue in several directions. Primarily, we plan to implement a complete set of a UGVL language, and test the efficiency of the vocabulary and grammar, as well as modify and expand it as needed. In addition, an implementation of a multi-modal interface (speech and gesture) is reckoned to increase performance time in general, robustness from errors, and ease of use in particular. The use of higher levels of individual autonomy in each platform, we believe to offer flexibility for the operator, by engaging less time for rectifying the robots' pose. Eventually, we are planning to continue exploring various control methods, and scenarios with more levels in the hierarchy of G-LOA, by deploying a larger number of robots.

ACKNOWLEDGMENT

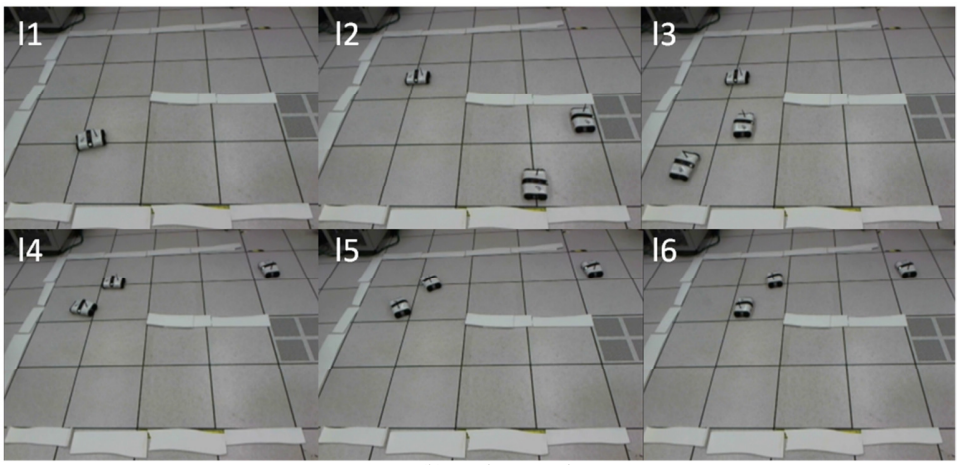
The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. The work of authors Dr T. Theodoridis, Dr H. Hu, and Dr K. D. McDolnald-Maier was supported by the UK EPSRC Global Engagements grant EP/K004638. Dr D. Barerro was supported by the University of Alcalá Mobility Grant. We thank Dr Y. Iwashita and Mr L. B. Clark for their support in the implementation of the experiments. We thank the anonymous reviewers for their suggestions in improving this paper.

REFERENCES

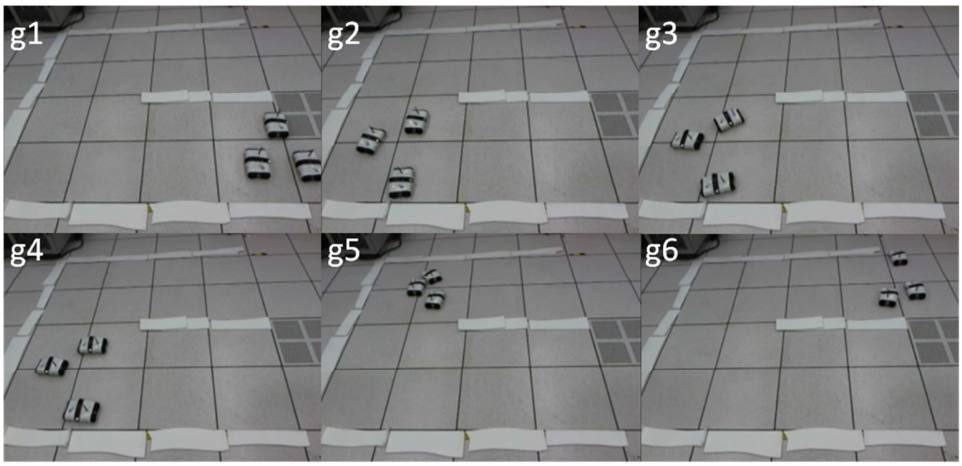
[1] Mathews R., et al. "A wearable physiological sensor suite for unobtrusive monitoring of physiological and cognitive state" Proc. 29th Annual Int. Conf. IEEE EMBS(2007)
[2] Wheeler K.R., Jorgensen C.C. "Gestures as input: Neuroelectric joysticks and keyboards", Pervasive Computing 2(2), 56-61 (2003)
[3] Wheeler K.R., "Device Control Using Gestures Sensed from EMG", Proc. IEEE Int. Workshop on Soft Computing in Ind. App., (2003)
[4] Shenoy P., Miller K.J., Crawford B., Rao R.P.N. "Online Electromyographic Control of a Robotic Prosthesis", IEEE Trans. on Biomedical Engineering, 55(3), pp. 1128-1135 (2008)
[5] Tenore F.V.G., *et al*, "Decoding of individuated finger movements using surface electromyography", IEEE Trans Biomed Eng (2009)
[6] Cortes C., Vapnik V. "Support-Vector Networks" Machine Learning, 20, (1995)
[7] Hsu C.W., Lin C.J. "A Comparison of Methods for Multiclass Support Vector Machines" IEEE Trans on Neural Networks, (2002)
[8] Darrell T., Pentland A. "Space-time gestures" CVPR, pp.335-340 (1993)
[9] Stoica A., et al.: Using arm and hand gestures to command robots during stealth operations SPIE DSS 2012 Baltimore, 2012
[10] A. Giusti, J. Nagi, L. Gambardella, G. A. Di Caro, Cooperative sensing and recognition by a swarm of mobile robots, Proceedings of the 25th IEEE/RSJ Int Conf on Intelligent Robots and Systems (IROS), Vilamoura, Portugal, October 7-12, 2012.
[11] Dorigo M. et al Swarmanoid: a novel concept for the study of heterogeneous robotic swarms. Robotics & Automation Mag, 2012.
[12] Haas, E.C. et al, "Multimodal controls for soldier/swarm interaction," RO-MAN, vol., no., pp.223-228, July 31 2011-Aug. 3 2011
[13] Fields, et al; , "Effective robot team control methodologies for battlefield applications," Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ Int Conf , vol., no., pp.5862-5867, 10-15 Oct. 2009
[14] Parker, L.T.; Howard, A.M.; , "Assistive formation maintenance for human-led multi-robot systems," Systems, Man and Cybernetics, SMC, IEEE Int Conf, vol., no., pp.2350-2355, 11-14 Oct. 2009
[15] Curran, K.; Condell, J.; Knox, J.; , "Guiding robots through wireless location positioning," Tech. for Practical Robot App. (TePRA), 2012 IEEE Int Conf. on , vol., no., pp.25-30, 23-24 April 2012
[16] Nathan P. Lucasa, Abhilash K. Pandyaa, R. Darin Ellis Review of multi-robot taxonomy, trends and applications for defense and space Proc. of SPIE Vol. 8387 83871N-2
[17] Sheridan, T. B. "Teleoperation, telepresence, and telerobotics: Research needs for space". Human Factors in Automated and Robotic Space Systems: Proc of a Symposium. 1987.
[18] http://www.swarmanoid.org/
[19] (http://www.swarms.org/)
[20] Chen, J.Y.C.; Barnes, M.J.; Harper-Sciarini, M.; , "Supervisory Control of Multiple Robots: Human-Performance Issues and User-Interface Design," SMC, Part C: Applications and Reviews, IEEE Trans on , vol.41, no.4, pp.435-454, July 2011
[21] J. M. Riley and L. D. Strater: Effects of Robot Control Mode on Situation Awareness and Performance in a Navigation Task. Proc. of the Human Factors and Ergonomics Society
[22] Wolf et al, Decoding Static and Dynamic Arm and Hand Gestures from the JPL Biosleeve IEEE Aerospace Conference, 2013
[23] Assad, C. et al BioSleeve: a Natural EMG-Based Interface for HRI, HRI 2013
[24] Goodrich, M. A. and Schultz, A.C Human-Robot Interaction: A Survey Foundations and Trends in Human-Computer Interaction, 1(3), 2007, pp 203-275.
[25] M. A. Goodrich, *et al.* Managing Autonomy in Robot Teams: Observations from Four Experiments. In Proceedings of the ACM SIGCHI/SIGART IEEE RAS Conf on Human Robot Interaction. Arlington, Virgina, USA. March 8-11, 2007
[26] C. A. Miller, et al A playbook interface for mixed initiative control of multiple unmanned vehicle teams. In Proc. Digital Avionics Systems Conference, volume 2, pages E4–1 – 7E4–13, 2002
[27] Autonomy Levels For Unmanned Systems, http://www.nist.gov/el/isd/ks/autonomy_levels.cfm



(a) Individual control.



(b) Leader control.



(c) Group control.

Fig. 4 Experiments with different levels/ hierarchy of group autonomy. (a) Controlling each robot separately. (b) The leader is controlled, while other robots are 'cloned' (execute the same commands as sent to leader) - individual corrections are needed l4→l5 maneuver as the leader requires further corrections of orientation as in l6. (c) All robots obey the same command, yet individual difference and those induced by different terrain in their path lead to differences that receive compensation on individual level. The turning maneuver g2→g3, requires further corrections as in g4.