Using arm and hand gestures to command robots during stealth operations

Adrian Stoica*, Chris Assad, Michael Wolf, Ki Sung You, Marco Pavone,
Terry Huntsberger, Yumi Iwashita
Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Drive, Pasadena, CA, 91109

ABSTRACT

Command of support robots by the warfighter requires intuitive interfaces to quickly communicate high degree-of-freedom (DOF) information while leaving the hands unencumbered. Stealth operations rule out voice commands and vision-based gesture interpretation techniques, as they often entail silent operations at night or in other low visibility conditions. Targeted at using bio-signal inputs to set navigation and manipulation goals for the robot (say, simply by pointing), we developed a system based on an electromyography (EMG) "BioSleeve", a high density sensor array for robust, practical signal collection from forearm muscles. The EMG sensor array data is fused with inertial measurement unit (IMU) data. This paper describes the BioSleeve system and presents initial results of decoding robot commands from the EMG and IMU data using a BioSleeve prototype with up to sixteen bipolar surface EMG sensors. The BioSleeve is demonstrated on the recognition of static hand positions (e.g. palm facing front, fingers upwards) and on dynamic gestures (e.g. hand wave). In preliminary experiments, over 90% correct recognition was achieved on five static and nine dynamic gestures. We use the BioSleeve to control a team of five LANdroid robots in individual and group/squad behaviors. We define a gesture composition mechanism that allows the specification of complex robot behaviors with only a small vocabulary of gestures/commands, and we illustrate it with a set of complex orders.

Keywords: Human-robot interfaces, gesture recognition, electromyography, EMG sensor arrays, stealth operations

1. INTRODUCTION

1.1 The need for efficient means to communicate commands and exercise control over robots

Robots and various forms of unmanned platforms are gradually becoming a common tool in support of soldiers in the field. The current means of controlling them, however, are not soldier-centric or responsive to the needs of the field personnel. Soldier command of supporting robots and unmanned platforms requires intuitive interfaces to communicate fast, high DOF information, yet leaving the hands unencumbered. Clearly these platforms should be enhancers and not deterrents to the mission due to inefficient means of control. The level of effort in coordinating with robots should not be higher than coordinating with a fellow soldier, and ideally would use similar gestures and signals. Stealth requirements rule out voice commands and vision-based gesture interpretation techniques for soldier's intent during silent operations at night or in other low visibility conditions.

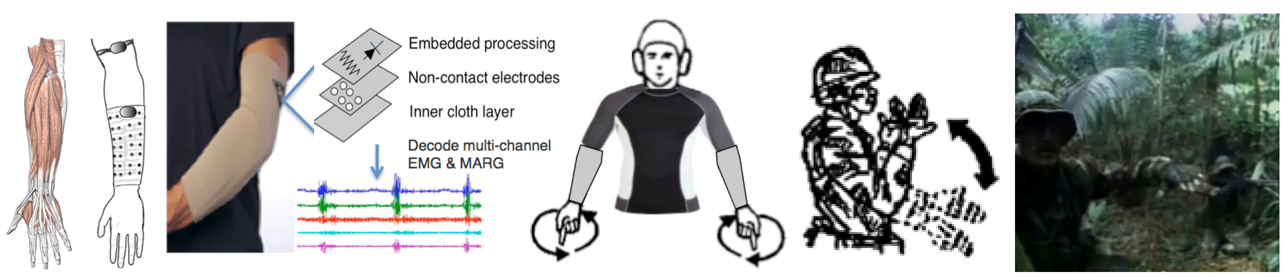

Figure 1. BioSleeve/BioSuit system concept. BioSleeve could monitor over 20 muscles and DOFs in the arm and hand.

Email: adrian.stoica@jpl.nasa.gov; Phone 1 818 354-2190.

Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2012,
Proc. of SPIE Vol. 8407, 84070G · © 2012 SPIE · CCC code: 0277-786X/12/$18 · doi: 10.1117/12.923690

1.2 A vision for interfaces controlled by bio-signals

The focus of our work has been interfaces and control systems that use biological signals to control robots. Our initial motivation was the need to provide astronauts with better ways of controlling manipulators in Extra-Vehicular Activity (EVA) activities, while having to deal with difficulty of using the EVA suit/gloves. Electromyogram (EMG) signals provide a direct, higher bandwidth and reliable modality for command interfaces. These extend to control of prosthetic limbs and further, to controlling not only one robot with multiple degrees of freedom, but also teams of robots. The interfaces have wide use, from setting navigation and manipulation goals for the robot (say, simply by pointing) to precise control of movement when needed. A primary goal has been the design of a wearable sleeve interface ("BioSleeve") for practical signal collection from forearm muscles, incorporating an integrated high density array of surface EMG sensors, several strategically placed inertial sensors, and in-sleeve sensor processing to fuse and decode all signals. As a first step, an initial prototype BioSleeve was developed with a sensor array of 8-16 surface EMG sensors and a 6-axis inertial measurement unit (IMU) mounted on the back of the hand.

1.3 Technical challenges of EMG systems

The main challenges of surface EMG systems come from: (1) sensor-to-skin interface issues that cause non-stationarity and signal degradation; (2) noise and other artifacts from motion of electrodes relative to the skin/muscle; (3) reliability of the array packaging, (4) separating signals that distinguish deeper muscles and individual fingers, and (5) the time-varying stochastic nature of the surface EMG signal itself, particularly for dynamic gestures. Issues 1 and 3 are primarily hardware related, issue 2 requires a combination of hardware and software improvements, and issues 4 and 5 require improved decoding algorithms. EMG data analysis is challenging in general, because the signals are stochastic and noisy, active muscles overlap for various movements, and forearm movements such as twists tend to shift the electrodes with the skin over the underlying muscles. Initial studies at JPL indicate that individual finger motions and twisting motions of the forearm are distinguishable with enough channels on the forearm.

Conventional EMG electrodes in use today are predominately passive "wet" electrode types with Ag/AgCl adhesive gel interfaces; these electrodes can be bothersome to mount and lose signal quality as they dry over time. Dry contact electrodes have also been used, particularly in clinical studies, but they also have interface issues and artifacts from relative motion and require constant mechanical pressure to maintain good skin contact. Practical non-contact sensors are now available, potentially resolving many of the above issues.[1,2]. However, further study is required for practical usage issues, including sensitivity to skin-electrode separation distance and saturation from motion artifacts and/or friction-induced electrostatic charge. The current version of the BioSleeve uses conventional electrodes, though the methods we present can be adapted for dry or non-contact electrodes.

A variety of papers have addressed recognition of EMG signals, but most of the work focused on a small number of sensors, typically wet contact sensors. Hand and individual finger tracking has been previously demonstrated from small forearm EMG arrays [3-6], with the focus on classification of discrete static gestures and not dynamic, temporally varying gestures. The BioSleeve classifies both static and dynamic gestures.

1.4 Paper outline

Section 2 describes the BioSleeve system, focusing on sensors, data acquisition and the software platform. The corresponding learning and classification algorithms and their results in the recognition of static and dynamic gestures are presented in section 3. Section 4 focuses on the use of the gestures to commands and control a group of five Landroid robots. Section 5 focuses on future work, presenting a path towards technology maturation and deployment.

2. BIOSLEEVE SYSTEM

The BioSleeve system and application concepts are shown in Figure 1. The system integrates several technologies to enable detailed and accurate arm and hand tracking: (1) EMG sensors, which can be embedded into clothing and be unobtrusive to the operator, (2) IMU sensors, which can be used to estimate limb orientation with respect to the body, and (3) advanced decoding algorithms for EMG gesture recognition. The BioSleeve system can be expanded to two arms with all degrees of freedom, to estimate position and orientation of carried equipment, and to add wearable sensors to monitor leg, head, and body posture (we refer to this complete system as BioSuit).

The first prototype BioSleeve (shown in Figure 2) consists of an array of bipolar surface EMG sensors on the forearm, with an IMU worn on the hand. A small, low power, differential EMG amplifier circuit was designed, built, and integrated in the sleeve. The circuit is based around a surface mount instrumentation amplifier (INA321 from Texas Instruments), analog bandpass filtering and output buffer, and snap buttons for electrode attachment. An array of these circuits fit in an elastic sleeve material for mounting on the user's forearm. The EMG signals are amplified, bandpass filtered, and buffered, and then transmitted through a wire tether to an off-board computer for digitization. Circuit characteristics: Power input = 228 µW (76 µA at 3V) during operation, <5 µA in sleep mode. Gain = 100 V/V. Frequency pass band = 16 to 600 Hz. Typical raw EMG signals are shown in Figure 3.

We experimented with both wet clinical electrodes and dry (no contact gel) electrodes in elastic skin-tight sleeves, and identified necessary design improvements. Using the wet adhesive electrodes make the sleeve/sensor array difficult to mount. Dry electrodes have a potential advantage in ease of use for mounting the BioSleeve on the user's arm, but can have lower signal to noise ratio if not in good skin contact. They require constant mechanical pressure to maintain good skin contact. A system with sixteen channels of bipolar sensor circuits for the in-sleeve array was constructed. Packaging the array in elastic sleeve materials proved to be the major challenge for reliability, because breaks in the array wiring from motion caused most experiments to be run with twelve or fewer working channels. The current prototype uses eight bipolar sensors with two wet adhesive electrodes per sensor.

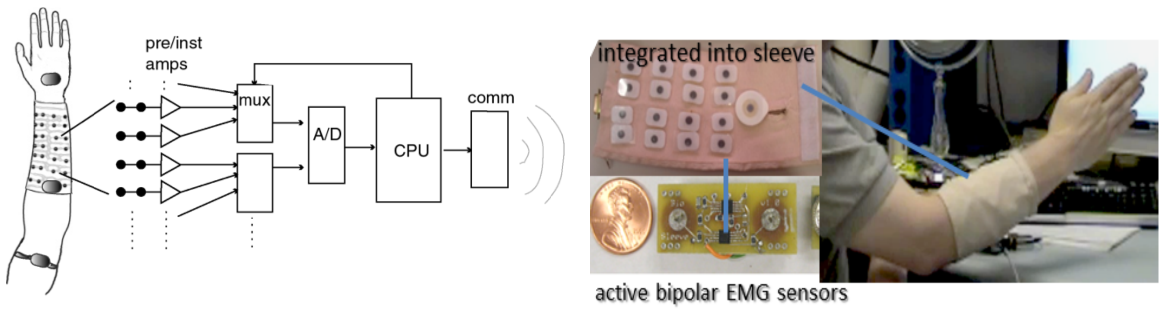

Figure 2. BioSleeve system architecture (left), and initial implementation (wired) with wet electrodes integrated into sleeve array in the current prototype (right).

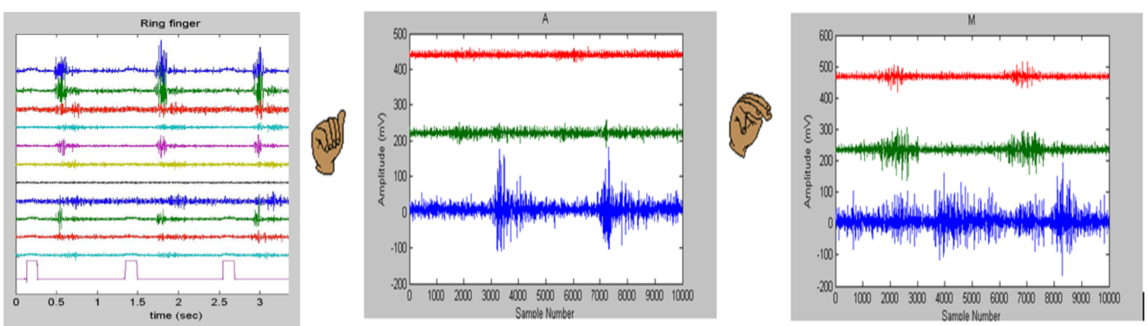

Figure 3. Left: example data from individual finger motion. Right: Sample raw EMG data from 3 channels during two similar letters of the American Sign Language alphabet.

3. LEARNING/CLASSIFICATION ALGORITHMS FOR GESTURE RECOGNITION

The signals acquired and filtered by the BioSleeve are sent off-board for gesture recognition processing. In the current implementation, static gestures are classified using the EMG signals in a Support Vector Machine (SVM) algorithm and dynamic gestures use IMU data in a custom technique founded on a spatial pattern recognition / dynamic programming technique. Future work will integrate EMG and IMU signals for both types of gestures. After donning the BioSleeve, the user completes a 2–5 minute calibration exercise, which collects data in each gesture to train the classifiers.

3.1 Static gesture recognition

Static gestures were implemented to enable a "virtual joystick" with five commands: left, right, forward, backward, and stop. These were accomplished with hand position only (allowing any arm position/orientation) and thus required only EMG signals. The classification approach is founded on multiclass SVM [7,8], which reduces the single multiclass problem into multiple binary classification problems. A five-class SVM was applied to an eight-channel EMG feature vector, where a feature is defined as the standard deviation of the EMG amplitude over the last 0.5-second window. One challenge of using amplitude-based signals, however, is that they can vary as the battery voltage supplied to the BioSleeve decreases. This effect was compensated for in software by offsetting the feature vector by the minimum value of the vector.

The BioSleeve classification algorithms and command interfaces were implemented on an off-board computer, which then sent the interpreted commands to up to five LANdroid mobile robots (iRobot Corp., USA) in real time. Figure 4 shows the labeled EMG data for the five static gestures. The separability of the classes indicates the BioSleeve provides high signal quality and data that is valuable for detecting the user's hand position. The classes were mapped to commands sent to the robots, so the user could drive the robot with these hand positions. Classification accuracy was consistently over 90%, with some tests indicating 100% accuracy (over 600 consecutive time steps), although current results are limited to a single user operating within about 30 minutes of calibration.

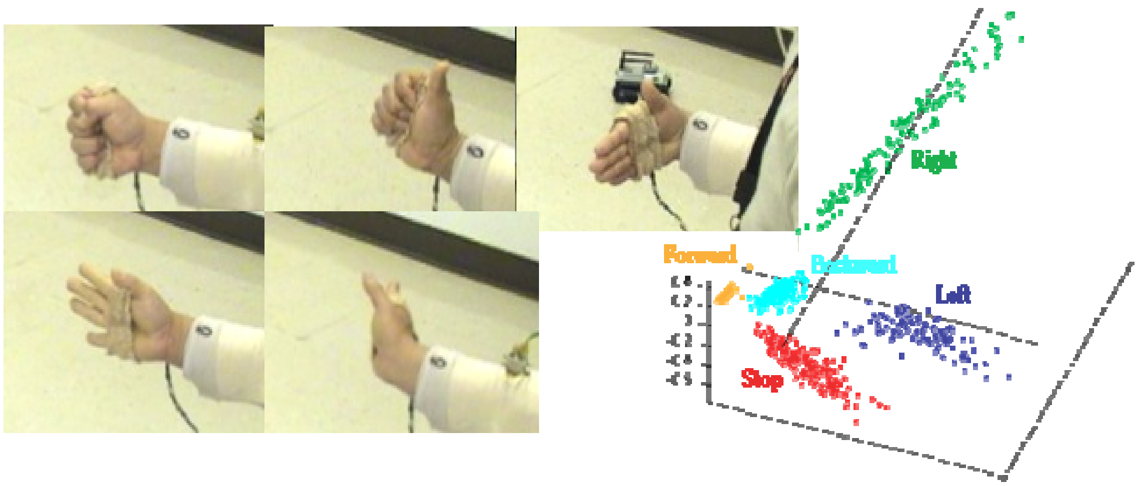

Figure 4: View of five static gestures (S1 to S5) and corresponding EMG signals displayed in their 3-dimensional principal components. The signals are well separated, indicating high classification accuracy is feasible.

3.2 Dynamic gesture recognition

To classify dynamic gestures, patterns of feature vector changes over time need to be detected and extracted. Dynamic programming (DP) was used in the analysis, because it was previously successfully demonstrated for gesture recognition due to its ability to optimally compensate for nonlinear time fluctuations of gestures [9]. During training and recognition the movements were repeated a number of times. Fig 5 shows five significant frames over a complete period, or hand/arm movement return to the starting position (which may not be needed for all gestures). Given a vector sequence $\mathrm{R_c = R_{c,1}, \cdots, R_{c,t}, \cdots, R_{c,T}}$ for a registered reference gesture pattern of category *c* from the calibration data and a vector sequence $\mathrm{I = I_1, I_2, \cdots, I_\tau}$, for an input gesture sequence comprised of several gestures (each feature vector I represents the value of IMU signal at a frame τ as shown in Fig. 6), the input gesture I is recognized with the DP algorithm by calculating the matching cost between I and $\mathrm{R_c}$. Gesture separability is over 99%, indicated by the accuracy of classifying the training data of 1000 time samples for each of nine gestures. True cross-validated accuracy has not yet been formally tested.

Figure 5. Image sequences illustrating nine dynamic gestures (D1 to D9)

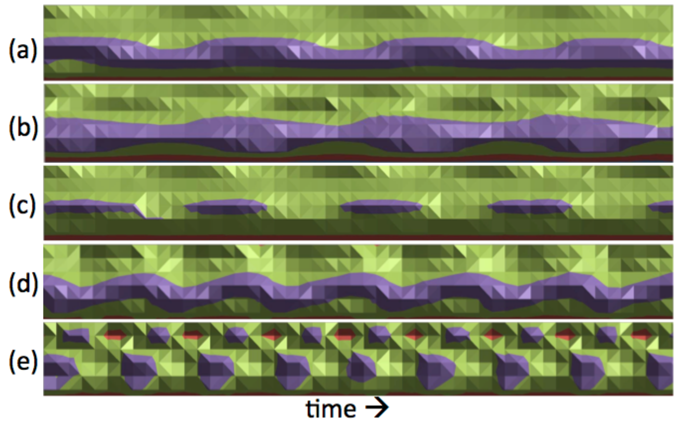

Figure 6: Heat map of IMU signals over time. The y-axis of each subplot is each IMU signal (3 rotational and 3 translational accelerations), and the x-axis is time. Five different repeated temporal gestures (a–e) so that the pattern of each can be recognized.

3.3 Discussion

The results reported here are promising but are preliminary – all results are from the same subject (N=1), and without removal/replacement of the sensor sleeve between calibration and testing. However, we expect that the system will generalize well to others because of the similarities in anatomy of the human arm and muscles. EMG data analysis for detailed gestures and hand/finger poses is more challenging, because the signals are stochastic and noisy, active muscles overlap for various movements (many to many mapping between muscles and fingers), and forearm movements such as twist tend to shift the electrodes with the skin over the underlying muscles. Our experiments in collecting simultaneous EMG data from the forearm indicate that, with a sufficient density of sensors in the array and active channel selection, one can distinguish patterns of muscle activities underlying different hand and finger motions, including individual finger motions and twisting motions of the forearm. This discrimination capability will be particularly important for correct classification between two similar hand/finger configurations, such as those shown in Figure 3.

4. GESTURE BASED COMMANDS FOR ROBOT CONTROL

The commands coming from static and dynamic gestures can be used to control a rich set of behaviors for a robot or a squad.

4.1 Gesture composition

To enrich the space of possible behaviors one can define categories of gestures that can be concatenated to produce a more complex behavior. One assignment is illustrated in Table 1.

Table 1: Combination of gestures leads to complex behaviors by combining selectors and sequencing gestures

Selector		D1	D2	D3	D4	D5	D6	D7	D8	D9
S1	Mode /Behavior	DOG mode	UAV-control	UGV-control	DOC pack	Column	Front	V-form	Semi-circle	Follow-lead
S2	Member	ALL	R1	R2	R3	R4	R5	LEAD	Non-L	SouthS
S3	Direction/ Speed	Advan ce	Retreat	Stop	Right	Left	U-turn	Faster	Same	Slower
S4	Attention	Freeze	Safe mode	Self-destroy	Avoid	Go	Come	There	Far	Close
S5	Validation									
S2- D1	Dog mode	Sit	Come	Retrieve	Search	Good	No!	Bite	Track	Crawl
S2-D2	UAV-con	Tk-off	Land	Hover	Follow	Back	Ahead	Photo	Video	Goal

According to Table 1, using function selectors provides a richer class of control.. The software recognizes the first gesture and depending on the meaning associated to it interprets the second gesture differently.Thus, S1-D1 means that there has been a mode select to the "DOG mode". By contrast, S3-D1 means 'Advance'

4.2 Describing complete orders as sequences

Gestures allows real-time tele-operation as well as mission specification after which the robots start execution of a sequence and autonomously interrupt or being corrected when needed, freeing the operator for his own part in the mission.

In order to simplify, the sequence will be expressed in a natural language, followed by the corresponding sequence of gestures.
"Platform 1, Dog mode, Search, (Validation). Platform 2, UAV control (Validation) Take-off , Ahead, Send Video" which in terms of gestures will be reflected in the sequences:
S2—D2, S1-D1, S3-D4, S5. S2-D3, S1-D2, S5, D1, D6, D8. S5.

The platforms would perform autonomously, would interrupt when needed, while the operator in turn can be providing guidance/corrections at any time (i.e., "Platform 2, Move faster" (S2-D3, S3-D7)).

4.3 Examples

Different gestures were programmed to trigger stored sequences on the LANdroid robot (for example, U-turn). Some gestures were used to indicate responses in a dialogue with the computer interface (e.g., pointing forward means "confirm command" - Validation). Figure 7 shows three example single robot actions, triggered by the corresponding gesture, and Figure 8 shows individual robot behavior control within a group.

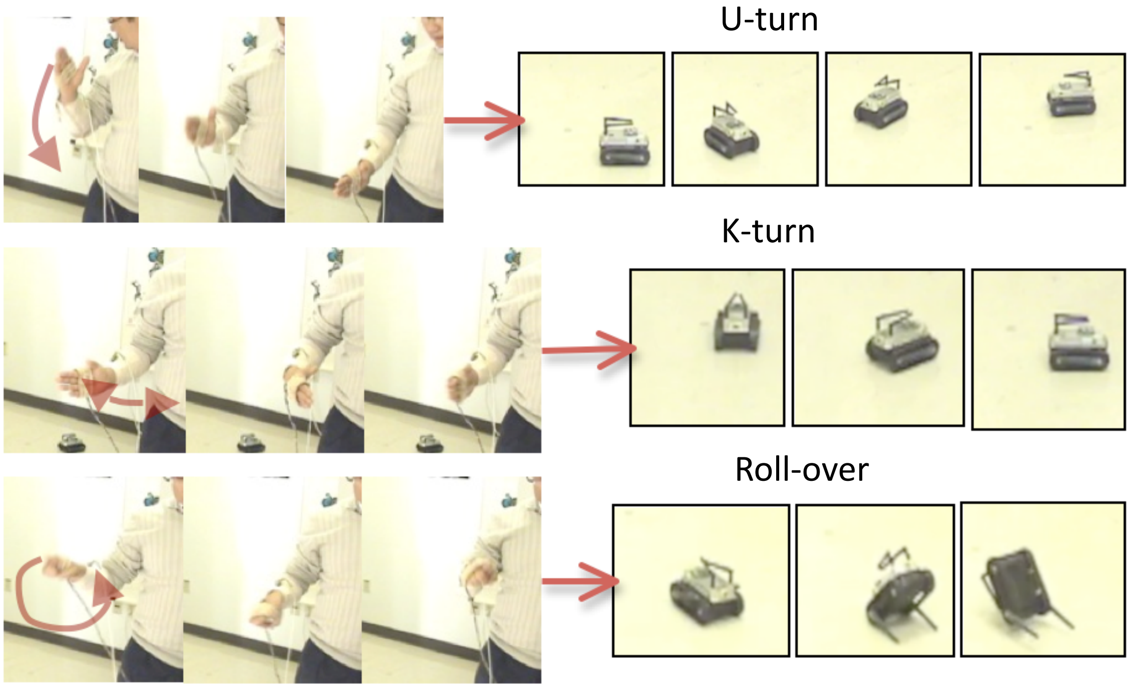

Figure 7. Left: Example dynamic gestures and corresponding robot action. Three dynamic gesture sequences are displayed with a sequence of frames each. To the right, the robot executes the stored procedure indicated by the gesture. From top to bottom, the actions are: U-turn, K-turn, and roll over.

Figure 8 illustrates group behaviors – entire squad following same order – to advance (top) and to group to center (middle), or a specific individual in the squad is ordered to execute a different action

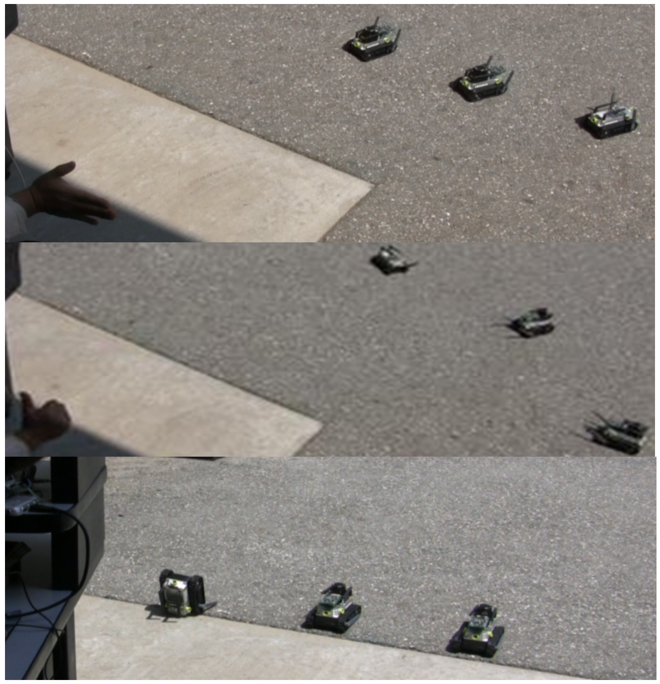

Figure 8. BioSleeve prototype in use to control a squad of LANdroid robots, as a group or with individual behaviors. UP: All robots advance as a front in the same direction. MIDDLE: Robots have surrounded a target and advanced in concentric circles towards center. BOTTOM: a member of the team has been singled and is executing an individual behavior. Each individual team member can be given a different behavior when needed.

5. FUTURE WORK

5.1 Advancing the BioSleeve/BioSuit

In order to enhance capabilities of the BioSleeve/BioSuit, the hardware system is being improved for increased usability—most notably by using non-contact electrodes for "slip on" use without special preparation of the forearm, with accompanying adaptive calibration algorithms to both minimize initial calibration time and limit the errors due to slippage over time. A large array of surface EMG sensors will be positioned over most of the forearm and upper-arm muscles, with up to twenty separate muscles monitored for high DOF motions. The EMG signals will be particularly useful to monitor force and position information from the hand and fingers. To complement the EMG array, several IMU sensors will monitor gross limb DOFs, with respect to the body. On-board processing will fuse data from EMG and IMU sensors, and gesture recognition software will minimize the bandwidth to communicate warfighter's actions. Data can be efficiently stored onboard for offline analysis, or wirelessly transmitted for real-time observation. All sensors will be non-contact and embedded into wearable sensor arrays, and thus can be conveniently donned and worn as part of clothing. Sleeve layer materials will be chosen with criteria based on elastic characteristics to hold the sensors close to the skin, durability, and user comfort. This packaging will allow for free mobility, having low mass, low power, and no external wires to constrain movement, and should improve system reliability. Importantly, the user's hands remain unencumbered, an important advantage when carrying equipment, manipulation tasks, or wearing complementary gloves/glove-based devices.

5.2 Applications in control of robots and unmanned platforms

We will continue to mature the robot control system, by expanding the vocabulary of gesture based control, integrating with other modalities such as voice commands, further merging the human commands with robot autonomy, and developing an appropriate human–robot dialogue for user feedback and exception control. We also explore modalities that are appropriate for control of large groups of platforms, in swarm or other formation, in particular with an interest to the control of swarms of expendable platforms.

5.3 Applications in tracking warfighter body pose and team communications

A wearable "BioSleeve" sensor system and accompanying software algorithms could continuously measure the configuration of a warfighter's arm, hand, and fingers, including recognition of communication gestures, and perhaps also determine the position and orientation of hand-carried equipment relative to the body. By fusing biological, inertial, and magnetic sensors, we believe the BioSleeve system could capture body and arm joint angles to within 1° and classify detailed hand positions with >90% accuracy, without placing any test equipment on the warfighter's hands.

6. CONCLUSIONS

We developed an EMG sensor array based BioSleeve able to recognize five static and nine dynamic gestures with over 90% accuracy. We used the BioSleeve to control a team of five LANdroid robots in individual and group/squad behaviors. We defined a gesture composition mechanism that allows the specification of complex robot behaviors with only a small vocabulary of gestures/commands, and we illustrated the giving of such complex orders with a set of examples.

ACKNOWLEDGEMENT

The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES

[1] Mathews R., McDonald N.J., Hervieux P., Turner P., Steindorf J. "A wearable physiological sensor suite for unobtrusive monitoring of physiological and cognitive state" Proc. 29th Annual Int. Conf. IEEE EMBS,(2007)
[2] Cognionics, Inc.: http://cognionics.com/
[3] Wheeler K.R., Jorgensen C.C. "Gestures as input: Neuroelectric joysticks and keyboards", IEEE Pervasive Computing, 2(2), pp.56-61 (2003)
[4] Wheeler K.R., "Device Control Using Gestures Sensed from EMG", Proc. IEEE International Workshop on Soft Computing in Industrial Applications, (2003)
[5] Shenoy P., Miller K.J., Crawford B., Rao R.P.N. "Online Electromyographic Control of a Robotic Prosthesis", IEEE Transactions on Biomedical Engineering, 55(3), pp. 1128-1135 (2008)
[6] Tenore F.V.G., Ramos A., Fahmy A., Acharya S., Etienne-Cummings R., Thakor N.V., "Decoding of individuated finger movements using surface electromyography", IEEE Transactions on Biomedical Engineering, (2009)
[7] Cortes C., Vapnik V. "Support-Vector Networks" Machine Learning, 20, (1995)
[8] Hsu C.W., Lin C.J. "A Comparison of Methods for Multiclass Support Vector Machines" IEEE Transactions on Neural Networks, (2002)
[9] Darrell T., Pentland A. "Space-time gestures" Proc. CVPR, pp.335-340 (1993)