2016 NDIA GROUND VEHICLE SYSTEMS ENGINEERING AND TECHNOLOGY SYMPOSIUM
AUTONOMOUS GROUND SYSTEMS (AGS) TECHNICAL SESSION
AUGUST 2-4, 2016 - NOVI, MICHIGAN

INTUITIVE ROBOTIC OPERATOR CONTROL (IROC): INTEGRATION OF GESTURE RECOGNITION WITH AN UNMANNED GROUND VEHICLE AND HEADS UP DISPLAY

Lisa Baraniecki
Jack Vice
Jonathan Brown
Josh Nichols
AnthroTronix, Inc.
Silver Spring, MD

Dave Stone
Dawn Dahn
Marine Corps Warfighting Laboratory
Futures Directorate
Quantico, VA

ABSTRACT

Currently, fielded ground robotic platforms are controlled by a human operator via constant, direct input from a controller. This approach requires constant attention on the part of the operator, decreasing situational awareness (SA). In scenarios where the robotic asset is non-line-of-sight (non-LOS), the operator must monitor visual feedback, which is typically in the form of a video feed and/or visualization. With the increasing use of personal radios, smart devices/wearable computers, and network connectivity by individual warfighters, the need for an unobtrusive means of robotic control and feedback is becoming more necessary. A proposed intuitive robotic operator control (IROC) involving a heads up display (HUD), instrumented gesture recognition glove, and ground robotic asset is described in this paper. Under the direction of the Marine Corps Warfighting Laboratory (MCWL) Futures Directorate, AnthroTronix, Inc. (ATinc) is implementing the described integration for completion and demonstration by 30 September 2016.

Background

Currently fielded ground robotic platforms are directly controlled by an operator using direct human input via a controller, often with a gamepad-like joystick controller operated by hand. Robots are operated using these direct operator control interfaces either where the operator has view of the robot and hence direct visual feedback of its performance or where the operator cannot see the robot, and visual feedback is provided by either a video feed and/or a visualization of the robot within its environment. At the same time, individual warfighters are increasing their use of (both in possessing and being tied into) personal radios, smart devices/wearable computers, and network connectivity at the squad level. Direct operator control of robots is tactically undesirable, and limits the robot's usefulness because it requires (at least) one warfighter's attention to operate, and usually more to provide the operator with security since the operator is 'heads-down," rendering him vulnerable in tactical situations. These current operational conditions for using ground robots therefore diminish a unit's (e.g. squad)'s warfighting capability.

The multitude of demands caused by operating a ground robot on a warfighter's attention can negatively impact performance (Mitchell, Samms, Glumm, Krausman, Brelsford, & Garrett, 2004). These demands can lead to poor decision-making, reduced response time, and generally poor overall performance as the individual must divert their attentional resources toward processing information as opposed to performing tasks (Wickens, 2002, 2008). Given the increase in cognitive demands on solders in the form of complex technological systems, a need exists for more intuitive operator control of ground robots.

Military operations are inherently dynamic environments and thus present physical and cognitive challenges to the human operators. In order for human teams to work together effectively, they must have accurate shared mental models, which are defined to be knowledge structures held by team members that enable them to understand task conditions which are used to coordinate their actions and adapt their behavior to task demands and the actions of the other team members (Cannon-Bowers, 1993). In the case of human-robot teams, these shared mental models are also necessary for successful team performance. However, humans and robots do not always perceive and process information in the same manner, which creates a barrier to information and task sharing. Moreover, the current state of artificial intelligence (AI) is such that humans are still necessary for direct or supervisory control to perform most tasks.

Effort Goals and Scope

AnthroTronix, Inc. (ATinc), a research and development engineering firm specializing in advanced human-machine interface devices, has extensive experience developing multi-modal interfaces for communication and command/control of computer-based systems such as wearable computers and robotic platforms (Vice et al, 2001, Vice et al 2005). ATinc has expertise in basic and applied research and development related to military training, and has conducted extensive research and development involving multimodal interfaces, including sensor-based motion tracking and gestural interface technologies, multimodal feedback devices, and mobile computing systems.

In order to address the complications of tactical situations, ATinc, under contract to and with direction from the Marine Corps Warfighting Lab (MCWL), is implementing an intuitive, integrated, interactive approach to robotic control. This approach covers both the method as well as implementation of the robotic control. Method refers to command input via hand gestures while implementation refers to the use of NuGlove, an instrumented glove that recognizes hand gestures. The user will NuGlove and make gestures that correspond to commands to be sent to the robot. Video feedback will be displayed on the heads up display. This comprehensive system allows the warfighter both control and information access without introducing an interruption into the task flow.

The complete system that ATinc is implementing is comprised of a Heads Up Display (HUD), a NuGlove Instrumented Glove, an Android Device, and an Endeavor Robotics PackBot 510 with FasTac. Figure 1 shows how the components of the system will interact.

Figure 1. Diagram of IROC System

Heads Up Display (HUD)

Heads up Displays (HUDs) allow users to view data and information without requiring that they move their heads or look away from their normal viewpoints. Users do not have to switch between heads down and heads up in order to obtain crucial mission information. This is especially relevant to tactical environments, where situation awareness maintenance is key. The HUD is able to provide necessary information on command, however the user is also able to return focus to the current task almost instantly.

Figure 2. Heads Up Display

NuGlove Instrumented Gesture Recognition System

During combat maneuvers, dismount warfighters will typically use hand-and-arm signals for communication. The warfighters will use an established set of hand-and-arm signals, which aids in the maintenance of shared mental models within teams. It also allows the warfighter to maintain noise discipline. The use of NuGlove to capture and relay this information allows the commands to be sent to multiple team members simultaneously and without the need for line-of-sight. NuGlove uses sensors that are small, lightweight, and unobtrusively incorporated into the warfighters' current field gloves. The concept of gesture recognition is a key component of what developers refer to as a perceptual user interface (PUI). The goal of such a design is to enhance the efficiency and ease of use for the underlying application design in order to maximize usability. The use of inertial measurement unit (IMU) sensor technologies for gesture recognition allows for a technically-feasible, near-term approach within uncontrolled environments.

Figure 3. NuGlove Instrumented Glove

NuGlove for Robotic Control

NuGlove provides an efficient means of control over a robotic asset. It provides a solution that allows for single-hand control, whereas standard gamepad controllers require two-handed control. The control is directly scalable to the range of motion of the hand. Additionally, hand gestures are an intuitive motion known to humans. The range of hand postures provides a wide range of potential commands for a robotic asset. Dynamic gesture recognition also allows for an intuitive means of direct control without the interaction with an additional controller.

Methods of Gesture Recognition

There are many ways to accomplish gesture recognition as a whole. The capabilities of the NuGlove IMUs, specifically, provide multiple methods of gesture recognition implementation. Gesture recognition is broken down into two main categories: static and dynamic gestures.

Static Gestures

Recognition of static gestures can be separated into two methods, based on the way the software recognizes the user input.

- **Discrete Hand Postures** – this means the hand itself is oriented in a unique manner, recognized by the software. A basic example of this would be the difference between making a “point” gesture and a halt (closed fist) gesture, with the hand location in space staying the same.
- **Unique Overall Hand Positions** – this means the hand posture can be the same, however the position of the hand changes over time. An example would be making a two-fingers (“peace sign”) gesture while moving the wrist to change hand location but keeping the hand posture the same.

Dynamic Gestures

Dynamic gesture recognition can be accomplished through multiple methods. In this categorization, it is important to highlight that dynamic gestures are being characterized by both their user implementation as well as the implementation of the commands by the software. By using this categorization, we are provided with an overall depiction of the dynamic gesture recognition process.

- **Proportional Control** — the movement of a static gesture is tied to the output response. This is typically used for direct control over the movement of the responding system. An example of this would be direct drive control via hand movement.
- **Dynamic Gesture Recognition via a Series of Static Gestures** – this is accomplished by recognizing a series of discrete static gestures in succession, during a distinct time period. This is the best way to implement dynamic gesture recognition using IMUs. Additionally, the way in which this is implemented on the side of the user is the classic example of gesture recognition.
- **Static Gesture Moving Through Space** – the user would implement this in the same manner as proportional control, but the software interpretation would be different. This would take the recognition of the static gesture moving through space and assume a discrete dynamic gesture. This would then be tied to a single software command, as opposed to direct control of some aspect of the system.

- **Dynamic Gesture Moving Through Space** – due to the constraints of IMUs, this method is vastly under researched and therefore rarely implemented. There is also limited applicability and need, in terms of use-cases.

Gestures
In an effort to leverage current military communications procedures of using hand-and-arm signals, the library of gestures selected for commands are based on current Marine Corps signals. Due to hardware and/or software restrictions, some signals needed to be modified for this project.

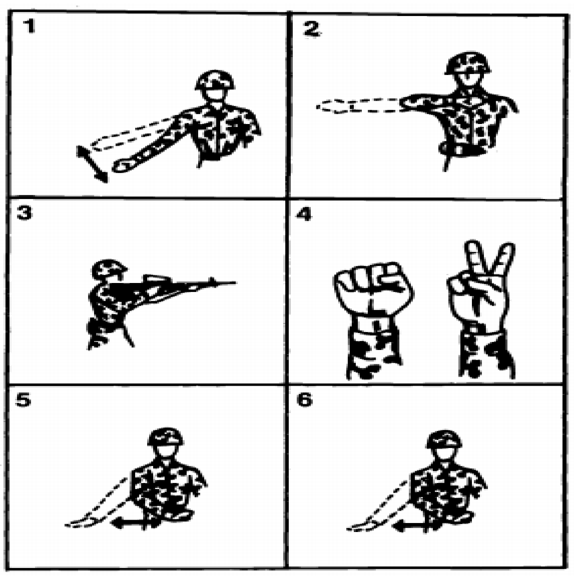

Figure 4. Standard Marine Hand-and-Arm Signals (U.S. Marine Corps, 2004)

NuGlove Gesture Recognition Algorithm
NuGlove contains one mainboard CPU, with an IMU sensor, as well as 9 other satellite sensors, also known as fingerboards. The sensor values are converted to quaternions for gesture recognition. The quaternion at each sensor site is taken in relation to the sensor located on the back of the hand. This is saved and compared to the previously recorded saved gestures. Depending on the difference between the saved and created gesture, a gesture is recognized.

Android Device
The Android operating system was selected for its compatibility with the Nett Warrior End User Device. The Samsung Galaxy S4 device was selected for the system integration. The device is currently running Android version 5.0 (commonly referred to as "Lollipop"). The Android device is currently connected to the glove via wired USB. The device runs the gesture recognition software, which operates as a part of the larger system architecture (see "System Architecture" for further detail).

Figure 5. Samsung Galaxy S4

PackBot 510 with FasTac
Endeavor Robotics' PackBot 510 with FasTac is an explosive ordinance disposal (EOD) unmanned ground vehicle. Its multi-mission flexibility makes it an ideal choice for various military applications. The capabilities of the asset enable the warfighter to expand their effectiveness in the field.

Figure 6. PackBot 510 with FasTac

There are many ways to implement control over the aspects provided by the PackBot Robotic asset. For example, the Camera Arm and Manipulator Arm are the two main features of the vehicle. Depending on various use-case scenario constraints combined with hardware/software capabilities, there are multiple ways to control the arms.

- **Single-joint control** – the user is able to control the individual degrees of freedom of the arm as a whole via the individual joints
- **End-effector control** – implementing inverse kinematics, the user controls the position of the end effector of the robotic arm
- **Master-slave control** – typically used in scenarios where the mechanical arm mirrors or closely mimics that of the human arm, or another anatomical

structure (i.e. finger), this method directly ties the human movements to that of the robotic arm

It should be noted, however, that this is not an exhaustive list, especially given that a particular control method can incorporate multiple means of control.

This variety of options for control, in conjunction with the plethora of gesture recognition capabilities, allows for the exploration of many combinations of recognition methods and approaches to arm control. The NuGlove can be used to achieve each of these methods of control. For example, master-slave control has been demonstrated using the NuGlove system with a 3-link arm. Using sensors placed on the index finger, the movement of the Operator's index finger was tied to the movement of the robotic arm. NuGlove has also been implemented for single-joint control and end-effector control.

System Architecture

The main processing component of the IROC system runs on the Android platform. The NuGlove Gesture Recognition Software runs on the Android device. The software recognizes a given gesture input and outputs the necessary commands to be sent to the PackBot. The PackBot accepts communication wirelessly via an external controller connected to the robot via the payload port Ethernet connection. This implementation is at the direct suggestion of Endeavor Robotics, the new name for the former Defense & Security Division of iRobot. The intent for the HUD integration is to act as an interface display for the Android Application.

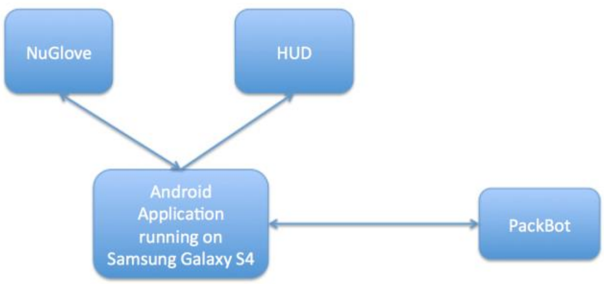

Figure 7. Diagram of System Architecture

Current Progress

At the Sea, Air, & Space Conference at National Harbor, MD, ATinc provided an interim demonstration of the current system. The demo included static and dynamic gesture recognition via glove running on an Android system, which wirelessly communicated commands to the PackBot asset. The commands implemented included chassis drive and camera arm deployment. This demo was presented alongside other demonstrations of current MCWL technologies.

Future Efforts

Potential follow-on efforts might focus on the integration of the current IROC system with an existing Autonomous Control software system. The IROC system provides an unobtrusive, intuitive means of accessing the system architecture, thus allowing for simple deployment of autonomous assets.

REFERENCES

[1] Cannon-Bowers, J. A., Salas, E., & Converse, S. A. (1993). Shared mental models in complex systems. Current issues in individual and group decision makinig, 221-246.

[2] Lathan, C., Vice, J. M., Tracey, M., Plaisant, C., Druin, A., Edward, K., & Montemayor, J. (2001, March). Therapeutic play with a storytelling robot. In CHI'01 Extended Abstracts on Human Factors in Computing Systems (pp. 27-28). ACM.

[3] Lathan, C. E., Tracey, M. R., Vice, J. M., Druin, A., & Plaisant, C. (2005). U.S. Patent No. 6,895,305. Washington, DC: U.S. Patent and Trademark Office.

[4] Mitchell, D. K., Samms, C., Glumm, M., Krausman, A., Brelsford, M., & Garrett, L. (2004). Improved performance research integration tool (IMPRINT) model analyses in support of the situational understanding as an enabler for unit of action maneuver team soldiers science and technology objective (STO) in support of future combat systems (FCS). Aberdeen Proving Ground, MD: US Army Research Laboratory

[5] U.S. Marine Corps, Marine Rifle Squad, MCWP 3-11.2, (Washington, DC: Headquarters, U.S. Marine Corps, 2002), 8-1 – 8-44.

[6] Wickens, C. D. (2002). Multiple resources and performance prediction. Theoretical issues in ergonomics science, 3(2), 159-177

[7] Wickens, C. D. (2008). Multiple resources and mental workload. Human Factors: The Journal of the Human Factors and Ergonomics Society, 50(3), 449-455.