NOTICE:

EVERYTHING HAS BEEN FINALIZED TO PDF AND SUBMITTED TO DARPA, EMAIL OR CALL JUN HWANG TO MAKE ANY CHANGES
jun@artemek.com

Due: Apr 9, 2015

All full proposals must be in the formats given below. Nonconforming proposals may be rejected without review. Proposals shall consist of two volumes. All pages shall be electronically formatted for a page of 8.5” by 11.0” with type not smaller than 12 points. Smaller font may be used for figures, tables and charts. The page limitation for full proposals includes all figures, tables, and charts. Volume I, Technical and Management Proposal, described in Proposal Content below, may include an attached bibliography of relevant technical papers or research notes (published and unpublished) which document the technical ideas and approach upon which the proposal is based. Copies of not more than three (3) relevant papers can be included with the submission. The bibliography and attached papers are not included in the page counts given below. The submission of other supporting materials along with the proposals is strongly discouraged and will not be considered for review. Except for the attached bibliography and Section I, Volume I shall not exceed 30 pages (40 pages if the proposal dollar value is > $1 million). Maximum page lengths for each section are shown in braces { } below. All full proposals must be written in English.

Proposals will be evaluated using the following criteria: (1) Overall Scientific and Technical Merit; (2) Potential Contribution and Relevance to the DARPA/TTO Mission; (3) Cost Realism; (4) Realism of Proposed Schedule; and (5) Proposer’s Capabilities and/or Related Experience.

Volume I

Section I. Administrative

a. BAA Number: DARPA BAA-13-22

b. Proposal Title: Automated Shared Awareness Through Sensor Data Classification

c. Technical Areas: TTO: Ground Systems: Soldier/Squad Technologies and Tactical Operations in Urban Environments

d. Lead Organization Submitting Proposal: Artemek Technologies INC

e. Type of business: Small Business

f. All Other Team Members and Type of Business Each:

- University of Southern California, other educational

g. BAA Technical Focus Area: Ground Systems

h. Technical POC 1:
Mr. Li, Chris
University of Southern California, Hedco Neuroscience Building, 3641 Watt Way Los Angeles, CA 90089-2520
chris@artemek.com

Technical POC 2:
Mr. Hwang, Jun
University of Southern California, Hedco Neuroscience Building, 3641 Watt Way Los Angeles, CA 90089-2520
jun@artemek.com

i. Administrative POC:
Mr. Hwang, Jun
University of Southern California, Hedco Neuroscience Building, 3641 Watt Way Los Angeles, CA 90089-2520
jun@artemek.com

j. Award instrument requested: Grant

k. Place(s) / Period(s) of Performance:
Research done at University of Southern California iLab and Artemek
Phase I: Aug 15th 2014 – Apr 15th 2015
Phase II: 4/15/15 – 4/15/16
Phase III: 4/16/16 – 12/16/17

l. Summary of the Costs of proposed research including total base cost, estimates of base cost in each year of effort, estimates of itemized options in each year of effort, and cost sharing if relevant.
Phase I: $168,200 (grant)

Phase II (option): $256,800(grant)
Phase III (option): $171,200 (grant)

m. Defense Contract Manufacturing Agency (DMCA) Administration Office: n/a

n. Defense Contract Audit Agency (DCAA) Audit Office: n/a

o. Date Proposal Was Prepared: Jan 9, 2014

p. DUNS number: 079252316

q. TIN number: 46-4474438

r. CAGE code: 72YZ2

s. Proposal validity period (minimum 180 days): 180 days

t. Affirmation of existing SETA support contacts: none

u. Affirmation of Human Subject Research: none

v. Affirmation of Animal Research: none

w. Statement of Unique Capability Provided by Government or Government-Funded Team Member: Not Applicable

2. Table of Contents {no page limit}

Section II. Summary of Proposal

## 1. Innovation

Currently, almost all information on the battlefield is relayed to command via radio and then processed by analysts throughout the chain of command. The drawn-out process of radioing information and having humans process it results in latency that can extend into hours–a process is too slow for the modern battlefield. Furthermore, plenty of critical information goes unreported, since it relies on the soldiers to judge situations where the information is useful, e.g. dehydration and early signs of injury. To solve these challenges, we will use lightweight, low-energy sensors to detect events of interest and transmit this data to command, where software can filter and present the information to help extract actionable insights.

The sensors themselves are attached to individual soldiers. They may include motion sensors, atmospheric sensors, and biometric sensors. From the raw data, the sensors use event detection to infer more practical information such as stress levels, rifle orientation, dehydration, etc. This information is delivered in real-time to commanders, who will be able to act on this information with minimal delay. In addition, combining data from multiple soldiers will yield further insights. We can triangulate enemies based on multiple rifle orientations and estimate ammunition supplies based on fired rounds. Finally, the data are stored so that missions can be replayed for training or learning, which allows us to learn things like where enemies are likely to be based out of (based on the location of prior engagements), and accurate details of what happened in a firefight for forensics or training.

## 2. Results

We intend to produce a suite of sensors for this purpose, starting with the rifle sensor, which can detect orientation and the exact time of each shot fired. Following the agile approach, we will produce working products at the end of each Phase.

For the hardware: In Phase I, we will have a works-like prototype with commodity hardware ready for a pilot test. In Phase II, we will have a ruggedized version working with existing communication infrastructure, ready for field-testing. In Phase III, we will apply learnings from Phase I and II in producing biometric sensors.

For the software: In Phase I, we will focus on the orientation and shot data streams, which alone are enough to provide low-latency information to commanders. In Phase II, we will automate the analysis of this information using custom classification algorithms (e.g. to infer duress, filtering for false positives). In Phase III, we will leverage the availability of multiple devices to do even higher-level learning, for instance combining the orientations of two rifles pointed at the same object and triangulating the location of their mutual target.

## 3. Technical Rationale

Having information pushed and filtered by software rather than pulled and analyzed by humans alleviates two critical factors that impede commanders: transparency and latency.

Much information on the battlefield goes unreported or gets lost. For instance, cases of dehydration and cases of concussion go unreported until the symptoms manifest, which by then is too late to mitigate without affecting the mission. As a result, commanders do not have a complete picture of what is going on and may have their missions impeded by these surprises. By bringing this information to the commander's attention, they will be able to make better decisions, increasing mission effectiveness.

Other information that sensors can report easily is difficult or impossible today. For instance, triangulating enemy locations based on two or more rifle orientations improves situational awareness, and should be more accurate than visually estimating location. The ability to replay missions by reviewing the stored data is a powerful way to reconstruct what happened for forensics or training purposes. Every round fired, rifle raised, elevated heartbeat and more are available to piece together past missions.

Finally, latency is much improved with this system. When a soldier engages an enemy, his rifle will rise, and he may fire a round minutes before he finds time to radio for help. These precious seconds matter and can be used to call for backup or medical attention. In addition, commanders can know the signs of dehydration or fatigue before they impact the soldier and jeopardize the mission. With these sensors coupled with event detection, we can greatly reduce the decision loop time, which leads to a more agile, responsive military force.

## 4. Technical Approach

We will start with the rifle orientation sensor, which transmits orientation and shot times over the network, in Phase I. We will then ruggedize the sensor to make it field-testable in Phase II. Finally, in Phase III we will use our learnings from Phase I and II to produce networked biometric sensors.

- Highly accurate attitude and Heading Reference System (AHRS) with commodity hardware, accounting for magnetic disturbances from the rifle and motion from gunshots. (Phase I)

- Create first-generation hardware. (Phase I)
- Event detection from rifle orientation or raw motion to infer the soldier's activity (idle, engaging an enemy, in duress). (Phase II)
- Use existing military communication infrastructure. (Phase II)
- Long 8hr+ battery life and ruggedized against impact, water, and heat. (Phase II)
- Deriving insights from multiple sensors: enemy triangulation, supply estimation, high-level combat overview (Phase III)
- Developing hardware for atmospheric and biometric sensors (Phase III)

**5. Experience**

Chris Li completed his undergraduate research at the University of Southern California (Los Angeles), where he built hardware and software for autonomous robots. His primary emphasis is in distributed systems, which include work done in industry for Apple and eBay. His open-source contributions include patches in Apache Hadoop as well as other widely downloaded projects (20k downloads). At Artemek, he is responsible for leading product development and engineering.

Jun Hwang is an undergraduate student (on leave) in computer science at Boise State University (Boise), where he previously researched innovative methods to detect protein markers for innovative methods for cellular regeneration. Although studies are within life sciences, his primary emphasis is product management which he has done work at Apple and eBay and as an strategy/M&A/corporate development analyst (FT) at The Walt Disney Company. At Artemek, he is responsible for leading business, finance/operations.

Laurent Itti received his M.S. degree in Image Processing from the Ecole Nationale Superieure des Telecommunications (Paris, France) in 1994, and his Ph.D. in Computation and Neural Systems from Caltech (Pasadena, California) in 2000. He is a Professor of Computer Science, Psychology, and Neuroscience at the University of Southern California. Dr. Itti has co-authored over 200 publications in peer-reviewed journals, books and conferences, three patents, and several open-source neuromorphic vision software toolkits.

**6. Cost and Schedule**

Phase I: Pilot testable product (8 months / $168,200)

- Works-like prototypes built and mountable to rifle
- Orientation sensor has highly-accurate AHRS

Phase II: Field testable product (8 months / $256,800)

- Orientation sensor detecting events such as raising of a rifle vs. resting behavior
- Acts like prototype: custom board, 8+ hour battery life
- Reliable communication over existing military infrastructure
- Small manufacturing run (~100 devices)

Phase III: Atmospheric / biometric sensors (8 months / $171,200)
- Sensors capable of detecting stress levels of soldiers
- Atmospheric sensors for heat, pressure, humidity
- Multi-sensor insights (enemy triangulation, supply estimation)

Section III. Detailed Proposal Information

**1. Statement of Work (SOW)**

**a. Objectives**

We plan to enhance shared awareness in the military by combining networked data collection hardware with event classification software. Currently, most information is transmitted via radio to human analysts, where it is passed up the chain of command and collated along with other intelligence, surveillance and reconnaissance (ISR), until a decision can be made. Thereafter, the decision is passed down through the various levels until a soldier can take action (respond).

**i. Current State of Technology**

Currently, the military transmits GPS location through **Blue Force Tracking** (BFT), which provides military commanders and forces a unified view of friendly forces' locations. However, much of the information on the battlefield is contextual, such as the presence of enemies and the health of individual soldiers, and this information is still radioed in. Current users of the BFT systems include the United States Army, the United States Marines Corps, the United States Air Force, the United States Navy ground-based expeditionary forces (i.e., United States Naval Special Warfare Command (NSWC) and Navy Expeditionary Combat Command (NECC) units), and the United Kingdom. Furthermore, BFT does not have the capability to track individual soldiers real-time.

The **Integrated Blast Effects Sensor Suite** (I-BESS) program is a sensor system which collects acceleration and pressure data. It is designed to detect concussions and other trauma that a soldier might not notice when returning to base. However, the I-BESS program is limited in scope to blast detection, and does not transmit data in real-time, which does not solve the issue of battlefield latency.

**Individual Gunshot Detectors** (IGD) are soldier-issued versions of the **Boomerang** (a gunfire locator). It detects the location of enemy snipers or soldiers and provides the information to soldiers. Our gunshot detector will detect shots only from the rifle on which it is mounted, which is an entirely different set of information. As for triangulation of enemies, boomerang is best suited for quick instances where enemies are firing gunshots, but will not help with locating enemies who are not firing rounds within range of the sensor. Our proposed system will allow soldiers to cooperatively tag targets or general locations with their rifles.

## ii. General Description of Tasks

### Highly accurate Attitude and Heading Reference System (AHRS)

These systems, while typically known for their usage on aircraft, will allow our rifle sensor to track its orientation over time. We will be using commodity hardware, since we plan on producing the device at a cost-effective price. Where our AHRS departs from the standard aircraft software is that it will need to track accurately through gunshots, which produce large impulses in acceleration. In addition, it will need to track correctly when mounted on a rifle, which will affect the compass readings. This will be done by the USC iLab in Phase I.

### Infer soldier activity from motion data

While rifle orientation is useful, ultimately context is more important. For instance, if the soldier is pointing the rifle towards the ground, they are probably idle, but if raising the rifle, they may be in duress. In addition to these tasks, we would also like to know things like if the soldier is running, or if the soldier is aiming. Finally, these insights can be combined with gunshot detection to further improve its efficacy. This will be done by the iLab in Phase II.

### Physical design of product

In Phase I, we will have works-like prototypes with 3D printed enclosures designed for rapid iteration. In Phase II we will begin integrating hardware from Phase I into ruggedized cases ready for field-testing. We expect the battery to last beyond what typical missions require, and to be either field-swappable or rechargeable back at base. This will be done by Artemek in collaboration with contracted industrial designers, design firms and contract manufacturers.

For communication: In Phase I, we will use off-the-shelf consumer Bluetooth radios to transmit data, which will be sufficient for a pilot test. However, in order to do a field test, we will need to interoperate with existing military infrastructure. This will be done by Artemek in Phase II.

### Multi-sensor insights

By combining sensor data from multiple rifle sensor devices, we can learn even more about the battlefield. One critical capability is enemy triangulation. If we have two or more soldiers aim at a single target, and we know their location and the direction of their rifles, we can then triangulate the location of their target. Another capability is supply estimation. By counting the number of rounds fired, we can determine which areas are combat-heavy and will likely need ammunition to be resupplied. This will be done by the iLab in Phase III.

**Biometric/atmospheric sensors**
From our learnings during Phase I and II, we will develop biometric sensors and atmospheric sensors with our existing framework. The atmospheric sensors can be used to determine things like elevated risk for dehydration, elevation (for improving GPS accuracy), and humidity (for calibrating weapons). The biometric sensors can determine stress, injury, or dehydration. We choose to do this in Phase III to mitigate risk–by now we will have learned from our experiences in making the rifle sensor. This will be done by Artemek.

## b. Technical Rationale and Approach:

### Highly Accurate Altitude and Heading Reference System (AHRS)

**Rationale and Background**. We propose to use solid-state microelectronics sensors (MEMS) mounted on a rifle, which will provide 9 independent measurements: linear acceleration in X, Y, Z, rotational velocity about X, Y, Z, and magnetic field direction in X, Y, Z. The first six (acceleration and gyroscopic rotation) can be integrated over time to track position and 3D pose (aka attitude) of the sensor. However, because the sensors are noisy, integration over extended time periods give rise to accumulated errors. Thus, the estimate of pose obtained through successive integrations of acceleration and rotational velocity drifts over time. To correct drift, absolute sensing is necessary as it allows the sensor to recalibrate its absolute zero. One could use GPS (typically, for high-motion applications such as aircraft, as GPS is not very accurate), or another common approach is to use a 3D compass (magnetometer). Integration of all 9 measurements is carried through a relatively sophisticated algorithm (e.g., extended Kalman filter) that also filters out noise to give rise to a real-time estimate of the 3D pose of the sensor. Thus, some computational power is required in addition to the sensing elements themselves. Several manufacturers already propose high-end AHRS systems, sometimes with additional features such as temperature compensation to maximize accuracy. These sensors, however, tend to be bulky, power-hungry, and, most importantly here, expensive (e.g., around $2,000 for Microstrain AHRS units). While this may be acceptable on large, expensive equipment (e.g., an aircraft), we here propose to explore a low-cost, low-power, yet accurate solution for rifle-mounted AHRS.

**Preliminary results:** We have built a prototype using the Invensense MPU-9150 as the IMU and a Teensy 3.0 to transmit data serially over USB. It streams data at 1kHz. No specific effort was mode to optimize for power consumption at this stage (see details in later section on physical design)

**Approach**: Specifically, we will save on hardware costs by using ARM microprocessors, which are both cost-effective and powerful. Commercial MEMS IMUs such as the Invensense 9600 MPU are being increasingly integrated into consumer smartphones such as the iPhone and many Android phones. Though less accurate than their high-end counterparts, we can tolerate small positional errors in our application (that autonomous aircraft that the expensive sensors are designed for, cannot).

By using these sensors and limiting their update rates, we can lower power consumption. As these devices will be mission critical, they will need to last the duration of the mission and longer. Soldiers on the field often carry 3V Lithium-Ion batteries for other devices, and one option would be making the battery swappable in-field. Another would be to have inductive or contact charging back at base, as radios currently have. Though there are concerns with the volatility of Lithium-Ion batteries near soldier's bodies, we believe that the distance due to being mounted on a rifle alleviates these concerns. We will also explore safer Lithium Ferrous Oxide (LiFe) batteries.

**Ferromagnetic interference and calibration.** One complication comes from the fact that the magnetometer measurements are affected by the nearby presence of any ferromagnetic material or large electrical currents. Typically, then, it is recommended that the compass sensor be mounted as far away as possible from those (e.g., on a non-ferromagnetic mast that sticks out of a vehicle). While electrical currents are likely not problematic here, we need to mount the sensor very close to the metallic rifle, and sticking out significantly is not an option as it could severely impact rifle usability. It is unlikely that we can fully account for magnetic distortions either through closed-form mathematical analysis or through simulation (given, e.g., a 3D model of the small arms being used). Hence, we will instead develop machine vision and machine learning tools to learn the distortions induced by a specific rifle, during a factory calibration phase. The most realistic situation is when a person is holding the rifle and also wearing/carrying the typical gear the eventual end-user would. Thus, we will replicate this situation in the lab using actors, and we will film the actors as they manipulate the rifle in various ways. Using machine vision software, we will extract the true pose of the rifle at any time. This will then be compared to the raw, uncalibrated measurements obtained from our rifle-mounted AHRS system, and the difference will be used to calibrate the sensors:

- we will use a Kinect (RGB + depth) sensor to film the actor and rifle at 30 frames/s. This sensor yields not only full-color video, but also a depth map, which makes it easy to detect the rifle and to accurately identify its pose. We have extensive experience with these RGB-D sensors and the associated data processing.
- before a calibration session, the Kinect will be placed in a fixed position, and true (reference) magnetic field will be measured in the arena using a highly accurate compass.

- the actor will then enter the calibration arena and take various poses, making sure to cover the full range of possible rotations of the rifle. To ensure highly accurate synchronization between the pose seen by the Kinect and the measurements of the sensors on the rifle, we will blink a light on the sensors using a particular binary pattern of blinking. As this pattern is observed in the Kinect video, it will be used to recover accurate clock synchronization between the Kinect and the rifle sensors. The rifle will be detected in the RGB-D pointcloud data, which should be achievable with high accuracy in most, but the most severely occluded poses given that a complete 3D model of the exact rifle will be available (for example, scanned beforehand using the Kinect and under no-occlusion, ideal-lighting conditions). In case we encounter difficulties in localizing the rifle in the pointcloud data, we will add visible markers to the rifle to help with initial location and pose estimation, then using the 3D data to refine this estimate and accurately match the 3D model of the rifle to the pointcloud data. If occlusions make detecting the rifle impossible in some configurations, we may add one or more extra Kinect devices at different locations around the actor.
- we will start with simple regression between reference magnetic field, plus the pose measured by Kinect, and pose measured by our rifle-mounted sensors. This will yield the sensor calibration. If regression is not sufficiently accurate, we will explore more sophisticated methods (e.g., K nearest neighbors, Locally Linear Embedding (Roweis & Saul, 2000)).
- because each MEMS sensor potentially needs to be calibrated differently, some amount of calibration likely will be necessary for every unit produced. We propose here to look for a decomposition of the calibration function into two sub-parts: 1) correction for distortions due to the metallic rifle, 2) compass calibration to true North. Likely 1) will be similar from one rifle to another, which could then, after we learn it once through the full calibration procedure just outlined, apply to other rifles. Then, 2) might become a much simpler calibration that can be done by the end-user shortly after purchase of the equipment (e.g., similar to how new cars equipped with a compass ask you to drive on a few tight circles to calibrate the compass). Thus, our last research task here will be to find the best way to split the full calibration into these two sub-parts.

**Special considerations for gunshot detection and other strong transient accelerations.**

Unlike typical AHRS design, one complication for the proposed system is that it will be subjected to deliberate strong accelerations, for example, each time a shot is fired, or the rifle is dropped or otherwise colliding with other hard objects or surfaces. Since the core AHRS algorithm to convert raw sensor measurements into pose is a filter, one needs to pay special attention to the operation of this filter across sudden transient acceleration. For example, Pourtakdoust & Ghanbarpour (2007) reported a modified, adaptive unscented Kalman filter to explicitly account for high transient accelerations, although in their AUV application those accelerations were still likely smaller than what will be experienced in our application. The research tasks here are:

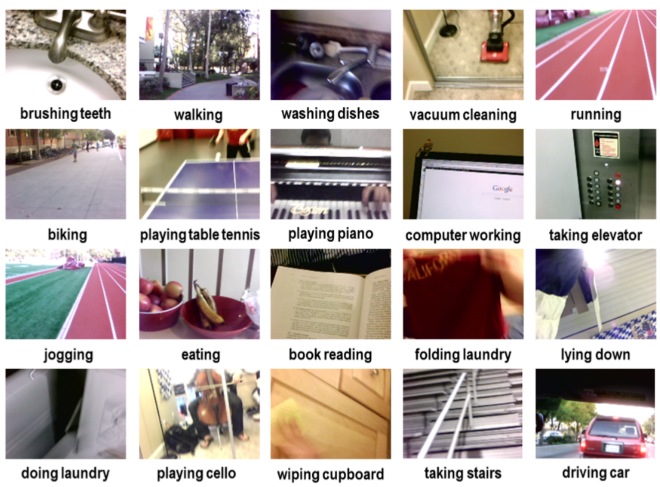
- measure raw rifle-mounted sensor data during shots, drops, and collisions. This will guide the choice of MEMS sensor (range of acceptable accelerations) and will allow us to build typical profiles of acceleration during these events.
- develop a continuous-discrete unscented Kalman filter to account for two possible regimens of operation (carrying vs. shot/drop/collision) and switching between the two. Possibly, we will need to create different regimens or models for shots vs. drops vs. collisions.
- test using out Kinect-based test arena how well the new AHRS algorithm is able to maintain accurate pose information across shots, drops and other collisions.

**Inferring Soldier Activity From Motion Data**

**Rationale.** We hypothesize that machine learning applied to streaming accelerometer, as well as possibly visual, audio or other streaming sensor data, can successfully decode user activity, situation and intentions. This will allow the system to constantly report and possibly adapt to user state and intent, making it more responsive and "aware" of user situation. This is part of our effort to avoid bombarding commanders with large amounts of irrelevant sensor data, delivering instead context-aware, actionable information about each soldier in theater.

**Background: Decoding activity, cognitive state, and user intentions.** In one system, Peters and Itti (2008) monitored eye movements while individuals played video games and were able to predict from the eye movement patterns, up to 2 seconds in advance, when the player was about to pull the trigger in a flight combat game or to shift gears in a car racing game. In further research, we have been able to predict, highly significantly above chance, which object a user will look at next in driving and puzzle-type games, through joint monitoring of eye movements, video frames, and joystick actions (Borji, Sihite & Itti 2012a; 2012b). When eye-tracking is not available, studies on activity classification have used inertial measurement units (IMUs; *i.e.*, accelerometers, gyroscopes), possibly combined with pressure sensors (Zhang & Tang 2012) or a camera (Spriggs et al 2009), placed on various body parts: thigh, waist, forearm, chest, knee, ankle, neck, foot (Zhang & Tang 2012), or head (Bao & Intille 2004; Yang & Hsu 2010). Machine learning algorithms identify patterns in the sensor data streams, giving rise to classification into different activities. Most studies thus far have used small databases, from five to nine activities. Five activities were successfully recognized by a wrist-worn accelerometer (94.13% correct; Chernbumroong et al 2011) and a waist-worn accelerometer (99.5%, Lee et al 2009). Six activities also yielded good results with a pocket-worn (91.7%, Kwapisz et al 2010) or a belt-worn (82.8%, Zhang et al 2010) accelerometer. An IMU on the front hip could differentiate between nine activities (90%, Zhang & Sawchuk 2011). Distributed inertial sensor networks used up to six sensor modules on different body parts (Yang & Hsu 2010), scoring 84% accuracy on 20 different activities (Bao & Intille 2004). An inertial sensor network combined with a camera yielded 61% on 29 kitchen activities (Spriggs et al 2009). *We believe that similar machine learning and pattern recognition techniques can be leveraged here to achieve a new level of activity decoding and automated situation awareness.*

**Preliminary results.** In a pilot project with Google's "Project Glass" team, we have been able to successfully decode 20 daily activities by monitoring, over time, data streamed by a head-mounted 9-degrees-of-freedom inertial measurement unit (IMU; 3 accelerometers, 3 gyroscopes, and 3 magnetometers; Fig. 1). We achieved over 80% correct classification, highly significantly above chance level of 5% correct classification.




| Confusion Matrix [%] — Predicted class / Actual class | biking | brushing teeth | computer working | doing laundry | driving car | eating | folding laundry | Jogging | lying down | playing cello | playing piano | playing table tennis | reading book | running | taking elevator | taking stairs | vacuum cleaning | Walking | washing dishes | wiping cupboard |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| biking | 90 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 2 | 0 | 0 | 0 | 0 |
| brushing teeth | 0 | 98 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| computer working | 0 | 0 | 79 | 0 | 4 | 2 | 0 | 0 | 12 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 |
| doing laundry | 0 | 0 | 0 | 98 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| driving car | 6 | 0 | 14 | 0 | 34 | 0 | 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 26 | 0 | 0 | 0 | 0 | 0 |
| eating | 0 | 2 | 0 | 0 | 0 | 96 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 |
| folding laundry | 0 | 0 | 0 | 0 | 4 | 0 | 86 | 0 | 1 | 7 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| jogging | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 97 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 0 | 0 |
| lying down | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 0 | 91 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| playing cello | 0 | 3 | 0 | 7 | 0 | 0 | 8 | 0 | 0 | 82 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| playing piano | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 95 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| playing table tennis | 1 | 0 | 0 | 16 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 79 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 |
| reading book | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 91 | 0 | 7 | 0 | 0 | 0 | 0 | 0 |
| running | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 96 | 0 | 1 | 0 | 0 | 0 | 0 |
| taking elevator | 0 | 4 | 0 | 0 | 0 | 5 | 9 | 0 | 5 | 11 | 0 | 0 | 0 | 0 | 66 | 0 | 0 | 0 | 0 | 0 |
| taking stairs | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 92 | 0 | 0 | 0 | 0 |
| vacuum cleaning | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 91 | 0 | 0 | 0 |
| walking | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 15 | 0 | 85 | 0 | 0 |
| washing dishes | 0 | 9 | 0 | 0 | 0 | 1 | 8 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 81 | 0 |
| wiping cupboard | 52 | 0 | 0 | 1 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 37 | 0 | 0 | 0 | 4 |

*Fig. 1. (left) Snapshots from 20 daily activities of an individual. (right) Confusion matrix of activity decoding from head-mounted inertial data. Overall accuracy was 80.3% (Windau & Itti, 2013).*

**Technical approach.** We will develop new algorithms to decode what soldiers are doing or intending to do, first based on IMU data, then possibly adding other sensors including visual or audio. Activities of interest here include:

*Stress Roles:*

- Relaxed with weapon lowered
- Relaxed with weapon raised (soldier using rifle to point, or accidentally raising rifle)
- In duress with weapon raised
- In duress with weapon lowered (soldier firing from top of building for instance)

*Physical Activities:*

- Running
- Walking
- Crawling
- Completely immobile
- Taking cover/being fired upon

*Urban Combat:*

- Climbing Stairs
- Climbing Walls
- Breaking a door
- Throwing a grenade
- Sweeping a room

*Marksmanship:*

- Long-distance Aiming
- Short-distance Aiming (CQB)
- Suppressing fire (non aiming)
- Reloading

The proposed general architecture can be structured in three major processing steps (Fig. 2). Step 1 prefilters sensor data and transforms it from a local dynamic coordinate system into a stable normalized coordinate system. Step 2 handles the feature extraction for IMU data (and possibly additional camera or other sensor data). IMU data is segmented into windows, followed by the extraction of statistical and physical features from each window. For each image in the camera (or audio, or other sensor) data, one statistical summary or so-called "GIST" feature vector is calculated (Siagian & Itti 2007). Step 3 is performing classification of activities/intentions by using a network of multiple classifiers. The result is a list of activities with assigned probabilities.

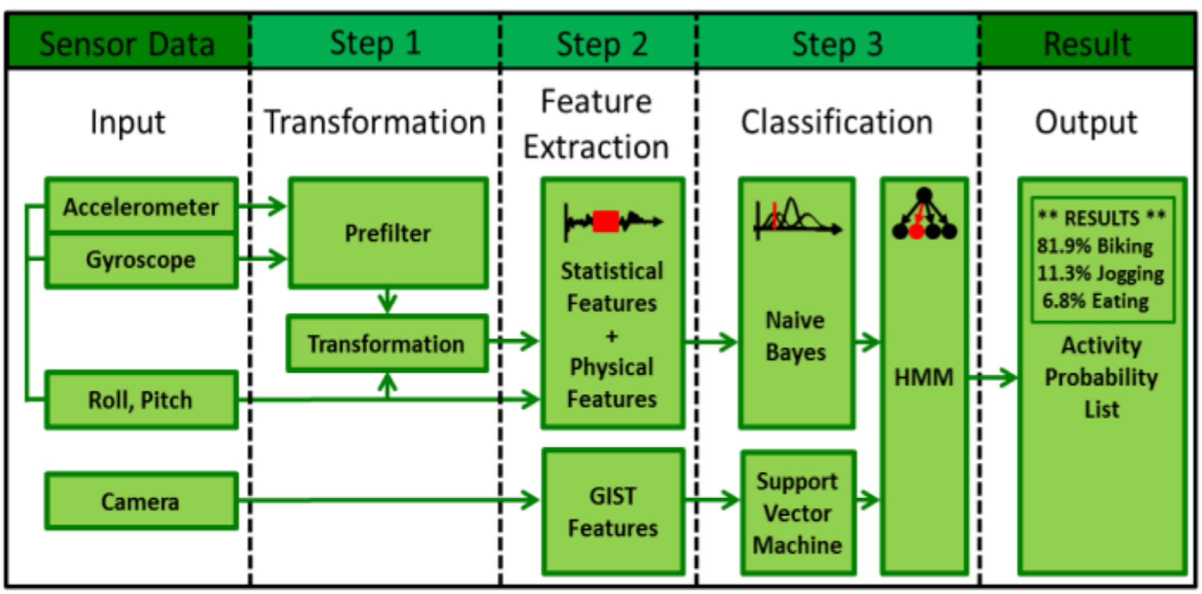



*Fig. 2. General proposed system architecture. In Phase I, we will focus on IMU data, possibly adding camera, auditory and other sensors in later phases.*

**Step 1: transformation.** It is important for classifiers to perform robustly when classifying activities such as walking or running to extract features that are not affected by the exact orientation of the weapon or the exact way a person is holding it. When a person performs a movement, the local sensor coordinate system will change accordingly. The key idea of our approach is to keep the sensor data in a static coordinate system that will stay stable even when the weapon moves (for decoding of activities related to soldier motion), in addition to measuring absolute weapon attitude (for decoding that involve raising or lowering the weapon). Thus, the sensor data needs to be transformed from its dynamic local sensor coordinate system into a normalized coordinate system. This normalized coordinate system is defined as the x-axis pointing out of the weapon, the y-axis points perpendicularly to the right and the z-axis points vertically down. Here we will use a simple matrix transformation as described previously in Windau & Itti (2013). Fig. 3 shows examples of accelerometer data before / after transformation.

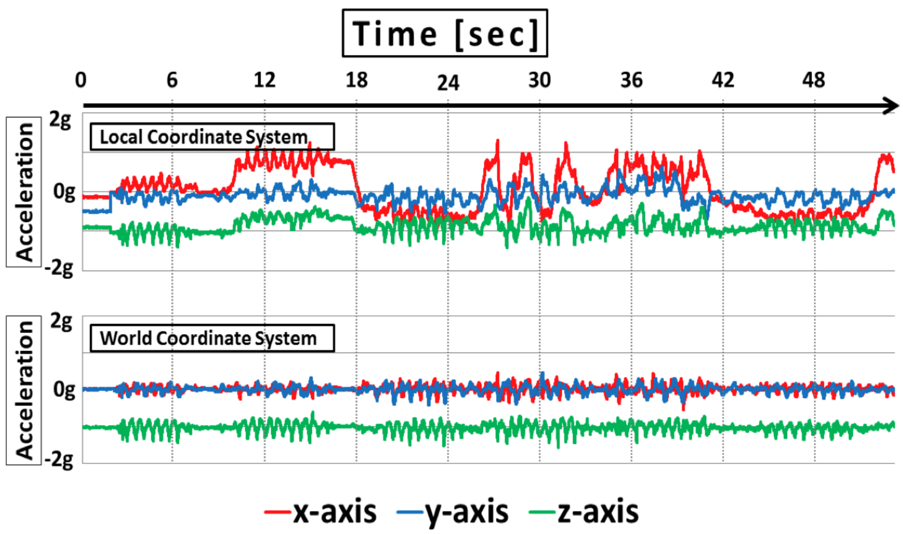



*Fig. 3. Accelerometer data before and after transformation into a stable coordinate system. While the raw data in local coordinate system shows large variations here linked to adjusting the pose of the sensor (e.g., adjusting the manner in which the soldier is carrying a weapon), those are irrelevant to decoding activities such as walking vs. running. The transformed data in world coordinate system eliminate these, making the decoding of walking vs. running more robust.*

**Step 2: Features.** Initially, for simplicity we propose to use 22 simple features extracted from inertial sensor data (Fig. 4). Seven features (energy, periodicity) are created by transforming sensor data into the frequency spectrum via FFT (Fast Fourier Transformation).

| Statistical Features | | |
|---|---|---|
| Type | Sensor Data | Dimensions |
| Mean | Accelerometer | 3 |
| | Gyroscope | 3 |
| Variance | Accelerometer | 3 |
| | Gyroscope | 3 |
| Total Number of Statistical Features | | 12 |

| Physical Features | | |
|---|---|---|
| Type | Sensor Data | Dimensions |
| Movement Intensity | Accelerometer | 2 |
| Energy | Accelerometer | 4 |
| Energy Consumption | Accelerometer | 1 |
| Periodicity | Gyroscope | 3 |
| Total Number of Physical Features | | 10 |

| Total Number of Features | 22 |
|---|---|

*Fig 4: Initial set of features computed from IMU data streams. This will be replaced in later years by an optimal set of features learned using a dictionary learning techniques.*

**Mean** is used to measure the average acceleration and angular velocity for each sensor axis over one window length. Mean is larger for activities with strong body motions. **Variance** describes how far acceleration and angular velocity are spread out along an axis. Fast and wide motions are larger. **Movement Intensity** (MI) specifies the intensity of motions. Mean and variance of MI are calculated over one window (Zhang & Sawchuk 2011).

$$MI[t] = \sqrt{a_x[t]^2 + a_y[t]^2 + a_z[t]^2}$$

**Energy** describes the motion quantity (Zhang & Sawchuk 2011). $E_i$ measures the energy for each axis; E is the energy over the entire system.

$$E_i = \frac{1}{N}\sum_{f=0}^{F} M_{a_i}[f]^2 \qquad E = \frac{1}{N}\sum_{f=0}^{F}(M_{a_X}[f]^2 + M_{a_Y}[f]^2 + M_{a_Z}[f]^2)$$

Parameters are N (number of samples per window length), $M_{ai}$ (discrete FFT component magnitude of acceleration along the axis i), f (frequency), and F (maximal frequency of window). **Energy Expenditure** (EE) is also known as the normalized signal magnitude area [8] and describes the amount of energy used for an activity. T is specified as the time of one window. **Periodicity** ($f_{peak}$) detects recurring motions. $F_{peak}$ determines the highest dominant frequency for an axis i. Parameters are $M_{wi}$ (discrete FFT component magnitude of angular velocity along the axis i), c (minimum required magnitude threshold) to avoid noise peaks.

$$f_{peak} = \mathrm{MAX}(M_{\omega_i} > c)$$

In later phases, we proposed to use dictionary learning to create an optimal set of features from training data. In this approach, we will collect data from all the desired activities. We will segment the data into small windows as done previously (e.g., 9 IMU measurements x 50 samples). We will then compute a sparse set of basic functions that can best represent these windows, for example using the Lasso algorithm (or any other sparse dictionary learning approach; Olshausen & Field 1996; Friedman et al 2008). A window of data collected during testing will then be approximated as a linear combination of the basic functions. The coefficients of this linear combination will be the new feature vector used for decoding. With this approach we expect that some of the basic functions will capture some prototypical "elementary action primitives", such as putting the left foot down while walking. Thus, the decoder will be able to work with these elementary action primitives as opposed to the semantically poorer statistical and physical features described above. Such an approach has proven quite effective in computer vision, where elementary pieces of objects can be learned in a similar manner from many training image patches, then allowing an object recognition algorithm to reason about assemblages of these pieces as opposed to raw pixel data (Zhu et al 2010a; 2010b).

**Step 3: Activity and intention classification / decoding.** In a first simple approach, we propose to use a Naive Bayes classifier followed by a Hidden Markov model (HMM) to decode and filter activities over time. We will then investigate more complex Dynamic Bayesian Networks, shown in our previous eye movement research to provide superior performance, as well as supporting integration of multiple, multimodal data streams (Borji et al 2012a; 2012b).

The Naive Bayes + HMM approach has been described in Windau & Itti (2013) and has yielded robust performance on decoding activities using a head-mounted IMU. Here we will extend this approach to the different types of activities considered in this proposal.

We will then develop a real-time mathematical framework to combine information from all available sensors (first, just IMU, then possibly adding new sensors). The sensor data streams will be fed to a dynamic Bayesian network to infer the (hidden-variable) cognitive state of the user along several simple dimensions (e.g., walking, crouching, shooting). These inferred variables will provide awareness of user state and intentions.

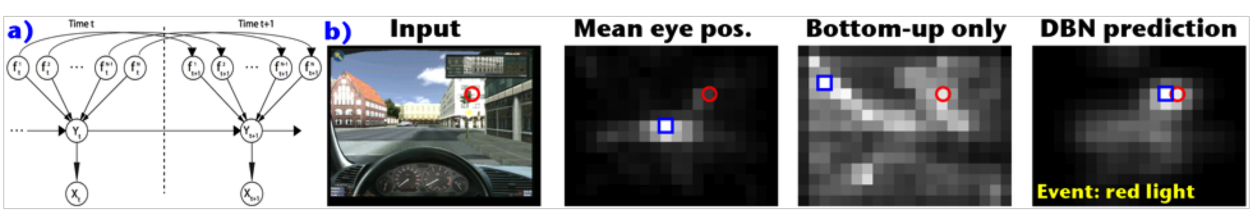



*Fig. 5: (a) Dynamic Bayesian network (DBN) framework (see text for notations). (b) Illustrative prototype from previous work, using camera and eye-tracking features in a driving scenario (eye-tracking will here be replaced by inertial data). From left to right: Input video and current eye position of a driver (red circle, on a traffic light); predicted eye position (blue square) by a trivial model that just averages all previous eye positions; and by a model*

*that only uses bottom-up saliency cues (both here fail to capture that the driver is monitoring the traffic light); predicted eye position by the DBN of (a), on average significantly closer to the subsequent eye movement of the driver (Borji et al., 2012a; 2012b). The DBN also predicts probabilities for 8 possible "situations" (driving straight, in a left turn, in a right turn, waiting at a red light, etc), over 90% correct in a leave-one-out test with 11 subjects (train on 10, test on the 11$^{th}$).*

We will use the Dynamic Bayesian Network (DBN) framework to provide a principled way to combine the feature inputs into inferred state variables. We pose the problem as follows. Low-level feature detection algorithms will extract features $f_t^i$ in real-time from inertial, and possibly other data streams. From these features, a DBN will iteratively update an internal estimate of the (hidden variable) user state, $Y_t$, and then predict in real-time the probability $X_t$ for the next user action (this prediction will thus be a probability vector over all known actions). The main difficulty in this work is in determining the state space for $Y_t$. At first, we will use clustering methods to determine it from training data. For example, using a simple k-means clustering, we will be able to investigate how increasing the dimensionality of this state space may improve subsequent decoding performance. As the relationship between $f_t^i$ and $Y_t$ may not be as simple as shown in Fig. 5, we will then investigate structure-learning methods to derive a possibly more complex graphical model than the one shown in Fig. 5. This DBN approach will hence require that we study from a theoretical standpoint how to best structure the graphical model which expresses the conditional dependencies (and lack thereof) between features, hidden variables, and output variables (Fig. 5). The simple DBN of Fig. 5 is likely to become too limited and a more elaborate network with additional intermediary nodes will likely be required. The next challenge is inference in the presence of noisy sensors and hidden (non-measurable) state variables, which we will solve by designing a Bayes filter, which we will implement as a Particle Filter for computational reasons, similar to our previous robotics work (Siagian & Itti, 2009). Note that while this inference may seem intractable, we have demonstrated ways in which it can efficiently be achieved (Borji *et al.,* 2012a; 2012b).

## Physical Design and Electronics

Our hardware design utilizes both off-the-shelf components and custom-made ones to speed development time (Fig. 6). Our existing and planned hardware revisions follow:

**Phase 0 (pre-DARPA) Off The Shelf**: used to develop proof of concepts and initial shot detection. Electronics were handled by the Texas Instruments SensorTag and the enclosure was 3D printed using a MakerBot Replicator Z18. The sensor attaches to the Picatinny rail of any small arms (rifle, SMG or pistol).

*Fig 6: (left) the custom enclosure for the TI SensorTag (right) the enclosure mounted on the Picatinny rail of a rifle (standard M4 "style" Carbine used by most infantry type soldier).*

**Phase 0 (pre-DARPA) Custom**: used to develop on custom hardware, though larger, the custom hardware is more powerful. The device samples at kilohertz rates and transmits a raw byte stream over Bluetooth. Electronics are off-the-shelf, and use wire-wrapped assembly for flexibility. They include the Teensy 3.0 ARM 96MHz mpu, Invensense MPU-9150 9DOF IMU, Analog Pressure Sensor, BlueSMiRF Bluetooth 3.0 wireless board, and a 3.7V LiPo battery pack.

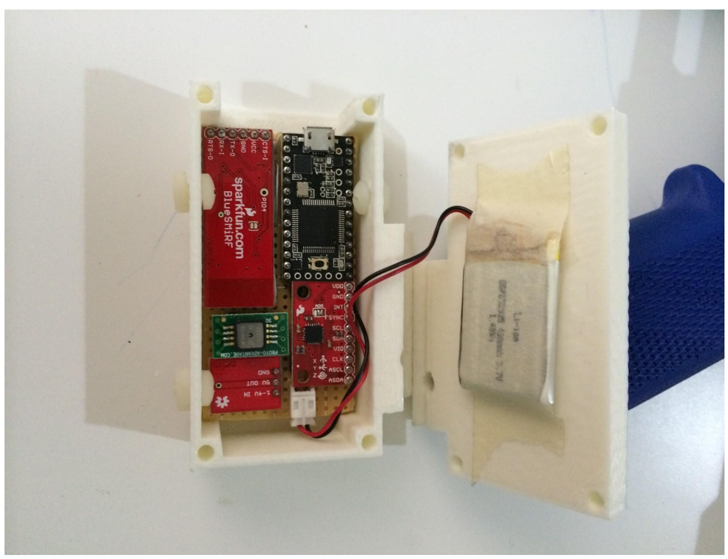



*Fig 7: One revision of wire-wrapped custom prototypes in its 3D printed case.*

**Phase I Off The Shelf**: We plan to use high-end IMUs (such as those made by Microstrain or Xens) as a reference to begin development of data visualization and user-facing software while the iLab works on using custom hardware. This will let us get started immediately without being blocked on progress of hardware.

**Phase I Custom**: We will produce other wire-wrapped boards so we can begin development of the AHRS software immediately. While that continues, we will design and manufacture PCBs of that design so we can make more of them. We typically order PCBs from contract manufacturers and assemble them in the iLab, which is equipped with soldering/reflow stations and a P&P machine.

**Phase II Custom**: In phase II, we will start back at wire-wrapped boards if we need to swap our communication chipset out, depending on what existing technology we choose for wireless communication in-field. After we pick and prototype, we will produce another PCB, coordinating with industrial designers on the final form of the device. Special consideration will be given to the antenna, so as to allow maximal range and efficiency when placed near a large ferrous body (the rifle).

**Phase III Off The Shelf**: We will our next round of biometric and atmospheric sensors with off the shelf hardware once again to allow our software efforts to advance unimpeded. We have already identified the TI SensorTag as a good candidate, for its low energy and environmental sensing's tolerance of low sampling rates. In addition, we will make use of existing commercial heart rate monitors (such as portable ones made for measuring heart rates of runners, like the Polar Bluetooth heart rate monitor).

**Phase III Custom**: While these sensors will get us started, they will likely be unfit for use on the battlefield due to the unique circumstances soldiers face. It will be our task to` create new sensors or contract new sensors for future deployment.

## Multi-Sensor Insights

The proposed Dynamic Bayesian Network approach naturally extends to scenarios where multiple sensors are analyzed jointly. In such case, we will modify the structure of the DBN to add nodes that represent possible joint intentions of a team, and that influence sub-networks, one for each team member, similar to those used to infer activity and intentions for a single individual.

Learning for such a more complex network would be significantly more involved, but likely can be split into two parts: first learn the DBN associated with each individual, and then fix it; second, only learn the top-level DBN that links the different individuals together.

Note that this extension to teams is simple when the team has a known, fixed size (or maximum size, as removing members from a team should not affect the functioning of the DBN). Here, we assume that the maximum size of a team will be known in advance and will remain modest (e.g., less than 100), as larger teams would typically split into sub-teams in a hierarchical manner, thereby prompting a hierarchical DBN. This approach fits well in the current command structure of the Army, for instance, which is organized into squads of 4-10 soldiers, platoons of 16-40 soldiers, companies of 100-200 soldiers, and so on, with an associated commander at each rank.

Activities to decode here include engagement of a squad, platoon or company, and estimation of their objective or objectives. In a manner similar to using the DBN approach to predict where in space one might look next while driving (Fig. 5), one of the outputs of the team DBN will be a spatial map that highlights the joint interest of the team over particular locations in space (an extension to simply triangulating a joint target from multiple rifles orientations). This approach has proven particularly effective in decoding locations of high interest or high value for future eye movements based on very noisy eye movement input data (Borji et al., 2012b; Fig. 5). In addition to this probabilistic decoding of team-based intentions, simpler metrics will be accumulated at the team level, such as depletion of ammunition based simply on counting the number of shots fired.

**Biometric/Atmospheric sensors**

Solid state atmospheric sensors such as humidity detectors, temperature detectors, and pressure sensors are cheap and low energy. The low update rates required (once every few seconds) means low power and longer battery life. In fact, a few of the atmospheric sensors (temperature, pressure) will already be present in the Phase I sensor. If atmospheric sensing is important to DARPA, we can integrate them into the rifle sensor and enable their use in Phase III with over-the-air firmware upgrades.

Our strategy for biometric sensors is to focus on off-the-shelf technology for the hardware and utilize their sensor outputs directly in our software for multi-sensor insights. In recent years, commercial biometric sensors have taken off, with sensors such as Bluetooth heart rate monitoring bands and stress detectors that detect the electrical conductance of skin. These sensors exist alone, but by networking them into our system, we can learn even more through correlations.

### c. Exit Criteria

Phase I: Produce a works-like device that can be demonstrated in a pilot test by military commanders at the different levels. The system will:

- Detect and transmit gunshot and orientation
- Provide graphical UI for commanders (brigade and below)

Phase II: Produce a looks-like device that is ready for field-testing in realistic conditions. The device will be heat, shock, and water tolerant. Device should communicate over a channel that is in use in the field today. The device will:

- Infer soldier activity from motion data
- Communicate in a bandwidth constrained environment over military infrastructure

Phase III: Combine information from multiple sensors (option, does not depend on Phase II)

- Multi-sensor insights for firearm sensor
- Develop new biometric/atmospheric sensors as works-like prototypes

**d. Deliverables**

We expect the following deliverables:

- Rifle sensor hardware in Phase I and II
- Biometric and atmospheric sensors in Phase III
- New techniques to track firearms in 3D space
- New classification algorithms to classify activities from biometric sensors
- New software for aggregating and visualizing relevant information for commanders
- New algorithms for aggregating sensory data at multiple levels: e.g. squad, platoon, company, etc.
- A system for storing raw sensor streams into data warehouses for later use

**2. Risk and Risk Reduction**

**Risk Area:** Existing communications may not be ready for our technology. The systems might not use battery friendly protocols or wavelengths, or the data rates may be too slow, or too intermittent.

**Risk Reduction:** To get started quickly, we will focus on Bluetooth initially. We chose Bluetooth since there is industry movement towards standardizing Bluetooth as a local wireless protocol (such as the Motorola XTS or APX series radio used in law enforcement). From here, we will evaluate whether adoption of Bluetooth enabled radios is imminent, or if we should adopt one of the many other communication technologies available on the battlefield, such as Wi-Fi.

Slow data rates are another concern. The amount of bandwidth available is measured in sub-kilobit levels today. To address this, we take the following approaches:

1. We do much of the processing device-side, so that we transfer only filtered data
2. We further filter on the device side so that we aren't always transmitting (if we detect the soldier is idle, we will transmit at a slower rate
3. We will save data to the device for offline access in the event the network is too busy. Some activities are more important than others (soldier in duress vs. soldier idle), and will take precedence during network congestion. This will ensure that data is still available for forensics or training purposes but doesn't necessarily congest the network.

**Risk Area:** Battery technology will not allow the device to last a reasonable amount of time.

**Risk Reduction:** One solution is to have the battery be swappable, as some devices today are. Soldiers already carry spare batteries on the battlefield. Another solution is to slow the data update rates or batch non-critical data updates so the radio chipset can sleep.

## 3. Expected Results

### a. Transferable Technology and Transfer Paths

In Phase I, we will deliver a working rifle orientation and shot sensor and associated user interfaces for pilot testing for the military. In Phase II, we will make our prototype into a field-testable device for the military. In Phase III we will produce pilot testable biometric and atmospheric sensors.

Artemek will produce and sell its shared awareness system to the military, including device manufacturing and software licensing/integration. To do this, we will be partnering with domestic contract manufacturers and hiring in-house engineers.

The iLab will be publishing papers and contributing their work to the research community.

### b. Proprietary Claims

Artemek claims the following intellectual property:

- Existing devices made by Artemek
- Existing shot detection systems previously done by Artemek
- Existing AHRS systems previously done by Artemek
- Any prior art (diagrams, etc.) done by Artemek as demonstrated in its provisional patents or documentation

Patent application information:

| Application Number | Inventor | Assignee | Title | Filing Date |
| --- | --- | --- | --- | --- |
| 61/846,487 | Chris Li, Jun Hwang | Artemek | FIREARM TRAINING SYSTEM FOR SHOOTING ACCURACY | July 15th, 2013 |
| 61/932,040 | Chris Li, Jun Hwang | Artemek | SENSOR ENABLED MOBILE WEAPON COORDINATION | Jan 27th, 2014 |

**4. Experience**

Chris Li completed his undergraduate research at the University of Southern California (Los Angeles), where he built hardware and software for autonomous robots. His primary emphasis is in distributed systems, which include work done in industry for Apple and eBay. His open source contributions include patches in Apache Hadoop as well as other widely downloaded projects (20k downloads). At Artemek, he is responsible for leading product development and engineering.

Jun Hwang is an undergraduate student in computer science at Boise State University (Boise), where he previously researched innovative methods to detect protein markers for innovative methods for cellular regeneration. Although studies are within life sciences, his primary emphasis is product management which he has done work at Apple and eBay and as an strategy/M&A/corporate development analyst (FT) at The Walt Disney Company. At Artemek, he is responsible for leading business, finance/operations.

Laurent Itti received his M.S. degree in Image Processing from the Ecole Nationale Superieure des Telecommunications (Paris, France) in 1994, and his Ph.D. in Computation and Neural Systems from Caltech (Pasadena, California) in 2000. He is a Professor of Computer Science, Psychology, and Neuroscience at the University of Southern California. Dr. Itti has co-authored over 200 publications in peer-reviewed journals, books and conferences, three patents, and several open-source neuromorphic vision software toolkits.

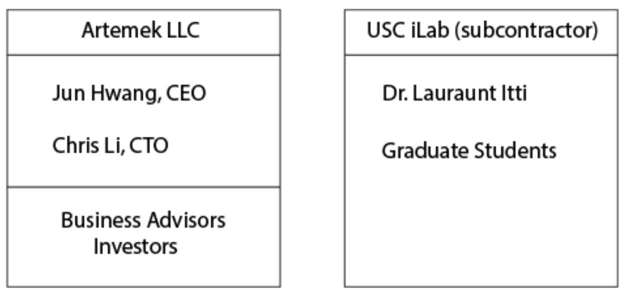
**5. Facilities**

The iLab is equipped with a CNC machine (Tormach 770), several 3D printers (including professional grade printers), hot-air soldering stations, conventional soldering stations, precision pick-and-place machine for assembly of circuit boards with tiny components (NeoDen TM-240A), solder reflow oven with precision-controlled temperature profile to finalize circuit board assembly, and a variety of robotics components. We use Altium Designer to design circuits, then outsource printing of circuit boards, and then use our lab equipment to assemble boards. We use SolidWorks for 3D mechanical design, then transferring to HSMWorks for CNC planning, and finally either carving a part from a block of aluminum using our CNC or building it from plastic using our 3D printer.

Artemek is equipped with an electronics lab, a 3D printer (MakerBot Replicator), and CAD software (SolidWorks, AutoCAD) for designing and manufacturing enclosures and boards (EAGLE). In addition, Artemek has relationships with outside design firms that specialize in industrial design, engineering testing, contract manufacturing, and RF testing.

Shot detection has been tested thus far on public shooting ranges. We expect to continue live-fire testing on public shooting ranges, but any pilot testing or field testing will be done at the military's preferred location.

### 6. Organization




Artemek is a registered Delaware LLC (in the process of changing to Delaware C Corp). Jun Hwang is responsible for business development/partnership outreach and all finance/business functions. He will ensure Artemek's operations and finances. Chris Li is responsible for engineering. He will ensure the product does what is needed by customers and is on time.

The University of Southern California's iLab also has a subcontractor relationship with Artemek to carry out tasks that relate to their specialty. Dr. Laurent Itti is responsible for coordinating iLab work, which involves meeting with graduate students and the general approach to solving technical challenges.

During the BAA, we plan on working closely together. The day-to-day engineering tasks will be carried out by Chris Li, Jun Hwang and graduate students in the iLab. These will involve sprint meetings (SCRUM) and planning sessions as well with various stake holders. On a monthly scale, we will meet to discuss timelines and the pace of the overall project tasks. As a former iLab researcher, Chris Li has done projects for the lab, and thus we anticipate a strong working relationship.

## 7. Project Management

### a. Management Plan:

We plan to utilize the agile methodology for development in weeklong sprints. Each week we will meet with a stakeholder, which will be one of Artemek's contacts in the military (active duty). They will help us prioritize our backlog of items to work on. Our team will then move items from the backlog into the sprint to work on. Artemek has daily standup meetings, which are open to iLab members as well. In our standup meetings, we quickly go over what we've done, what we're going to do, and any blockers or impediments. This lasts until the end of the sprint, where we host a sprint demo and retrospective, where we can go over what we accomplished and talk about how to improve.

The core team, which is composed of iLab personnel working on the project and Artemek employees, will meet in person at least once a month for the duration of the project. In the early phases, Artemek will do much of its development in the iLab while we finalize our hardware. We will synchronize files and documents through Google Drive (or another enterprise cloud storage), and financials will be shared quarterly during Artemek all-hands meetings.

Artemek's management team also consists of its advisory board, which includes Bob Gourley, former CTO of the DIA; Bob Flores, former CTO of the CIA; Courtlandt Gates, former CEO of Clearwater Analytics; and Andrew Rogers, Co-founder, and CTO of Space Curve.

Brief Resumes

Laurent Itti (USC)

Education
Ph.D, California Institute of Technology, Pasadena, California
M.S., Ecole Nationale Supérieure des Télécommunications, Paris, France
Mathématiques Supérieures et Spéciales M', Tours, France

Experience
*2013 – present*
Professor of Computer Science, University of Southern California, Los Angeles

*2006 – 2013*
Associate Professor of Computer Science, University of Southern California, Los Angeles

*2000 – 2006*
Assistant Professor of Computer Science, University of Southern California, Los Angeles

*2002 – present*
Voting faculty member, USC Neuroscience Graduate Program, Los Angeles

*2001 – 2002*
Faculty member, USC Neuroscience Graduate Program, Los Angeles

*2000 – present*
Adjunct assistant professor of Psychology, USC, Los Angeles

*2000 – 2000*
Postdoc in Neuroimaging Research, Harbor-UCLA Medical Center, Torrance, California

*1993 - 1998*
Neuroimaging Research Associate, Harbor-UCLA Medical Center, Torrance, California

Chris Li (Artemek)

Education
B.S., Electrical Engineering University of Southern California, Los Angeles

Experience
*2013 – 2014*
Hadoop Platform Engineer, eBay Inc., Bellevue, Washington

*2012 – 2012*
OTA/Antenna Design Engineer (Intern), Apple Inc., Cupertino, California

*2011 – 2013*
Undergraduate Researcher, University of Southern California iLab, Los Angeles, California

Jun Hwang (Artemek)

Education
B.A. Computer Science (incomplete - on hold), Boise State University, Boise, Idaho

Experience
*2013 – 2013*
Product Manager (Intern), eBay Inc., San Jose, California

*2012 – 2012*

Product Manager (Intern), Apple Inc., Cupertino, California

*2011 – 2012*

Analyst -Strategy & Business Development, The Walt Disney Company | Disney Consumer Product, Glendale, California

*2011 – 2011*

Quantitative Analyst (Intern), Standard & Poor's | S&P Capital IQ, New York City, New York

**b. Schedule**

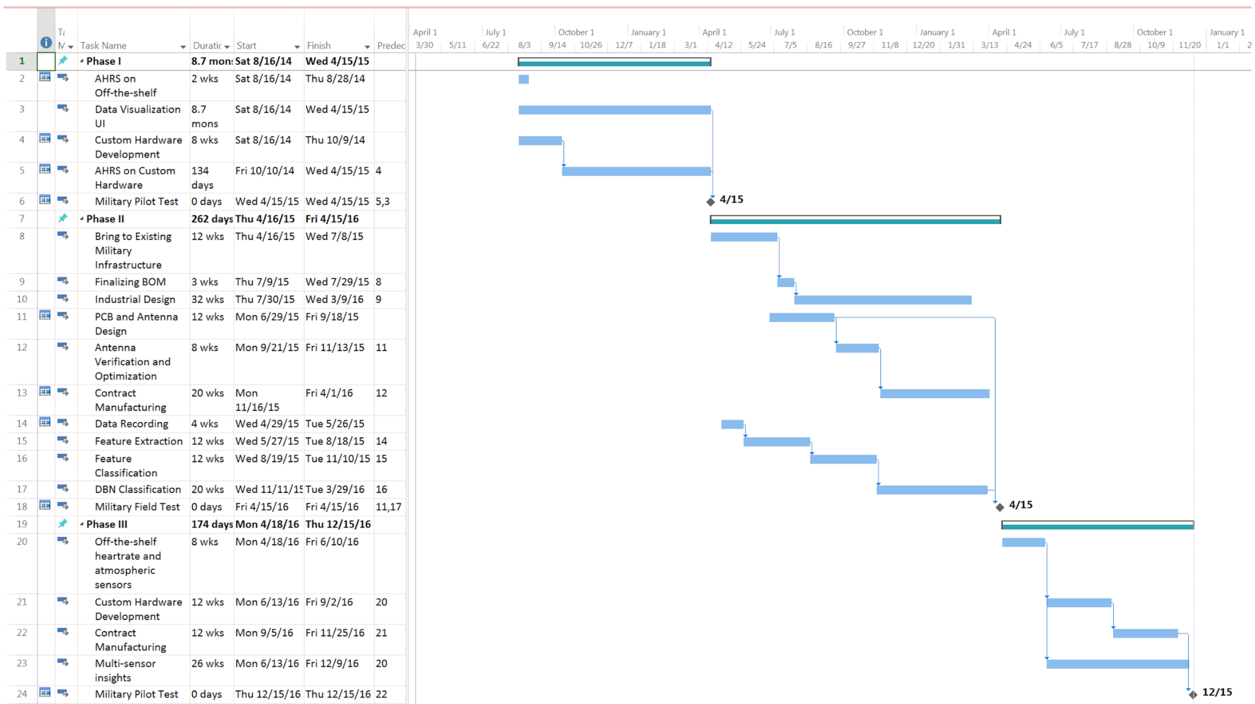



*Fig 8: Gantt chart showing expected times for tasks in each phase. The MS Project mpp file is included in the submission.*

<u>Milestones</u>

Phase I + 2 months: AHRS running on off-the-shelf hardware and transmitting data to remote computer, demo-able product

Phase I + 5 months: 1. Custom hardware completed and running shot detection, transmitting data remotely to 2. user-facing software which helps commanders visualize data

Phase I end (+8 months): Culminates in a pilot test for the military, demonstrating 3D orientation running on commodity hardware transmitted across the network in a controlled environment.

Phase II + 3 months: Custom hardware adapted to use existing military infrastructure. Feature data recorded.

Phase II + 5 months: BOM finalized and first hardware product design completed and ready for manufacture. Feature extraction completed.

Phase II + 8 months: Antenna testing completed and validated. Feature classification via Naive Bayes completed.

Phase II + 10 months: Small scale production complete. Feature classification via Dynamic Bayesian Networks completed.

Phase II end (+12 months): Culminates in a field test for the military, demonstrating the product in real operating environments using real communication infrastructure.

Phase III + 2 months: User facing software accepts data from off-the-shelf devices. Multi-sensor insights can do basic triangulation.

Phase III + 5 months: Custom devices developed for pilot testing, multi-sensor insights working at the squad level with Dynamic Bayesian Networks.

Phase III end (+8 months): Pilot test for the military, includes multi-sensor insights at the squad and company level. The small arms sensor from Phase II ready for field deployment.

## Section IV. Additional Information
## References

L. Bao and S. Intille, “Activity Recognition from User-Annotated Acceleration Data”, Proceedings of the Second International Conference on Pervasive Computing (PERVASIVE 2004), Linz/Vienna, Austria, LNCS 3001, A. Ferscha, F. Matterns, Eds., Springer-Verlag Berlin, Heidelberg, (2004), pp. 1-17.

A. Borji, D. N. Sihite, and L. Itti, “Probabilistic Learning of Task-Specific Visual Attention”, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2012), Providence, Rhode Island, 1-8, (2012a).

A. Borji, D. N. Sihite, and L. Itti, “What/Where to Look Next? Modeling Top-down Visual Attention in Complex Interactive Environments”, IEEE Transactions on Systems, Man, and Cybernetics, Part A - Systems and Humans, 1-16, (in press), (2012b).

S. Chernbumroong, A. S. Atkins, and H. Yu: “Activity Classification Using a Single Wrist-Worn Accelerometer”, Proceedings of the 2011 5th IEEE International Conference on Software, Knowledge, Information Management and Applications, Benevento, Italy, (2011).

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. Biostatistics, 9(3), 432-441.

M. Lee, J. Kim, K. Kim, I. Lee, S. H. Jee, and S. K. Yoo, “Physical Activity Recognition Using a Single Tri-Axis Accelerometer”, Proceedings of the World Congress on Engineering and Computer Science 2009 (WCECS 2009), Vol I, San Francisco, California, (2009).

J. R. Kwapisz, G. M. Weiss, and S. A. Moore, “Activity Recognition Using Cell Phone Accelerometers”, Proceedings of the 4th International Workshop on Knowledge Discovery from Sensor Data (SensorKDD 2010), Washington, DC, (2010).

Olshausen, B. A. and Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature, 381(6583), 607-609.

Pourtakdoust, S. H., & Asl, H. G. (2007). An adaptive unscented Kalman filter for quaternion-based orientation estimation in low-cost AHRS. *Aircraft Engineering and Aerospace Technology*, *79*(5), 485-493.

R. J. Peters, L. Itti, Congruence between model and human attention reveals unique signatures of critical visual events, In: Advances in Neural Information Processing Systems, Vol. 20 (NIPS*2007), pp. 1145-1152, Cambridge, MA:MIT Press, Jun 2008.

Roweis, S. T., & Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. Science, 290(5500), 2323-2326.

C. Siagian, L. Itti, Rapid Biologically-Inspired Scene Classification Using Features Shared with Visual Attention, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 29, No. 2, pp. 300-312, Feb 2007.

C. Siagian, L. Itti, Biologically Inspired Mobile Robot Vision Localization, IEEE Transactions on Robotics, Vol. 25, No. 4, pp. 861-873, July 2009.

E. H. Spriggs, F. Torre, and M. Hebert, “Temporal Segmentation and Activity Classification from First-Person Sensing”, Proceedings of the IEEE Workshop on Egocentric Vision, Miami, Florida, 17-24, (2009).

J. Windau, L. Itti, Situation awareness via sensor-equipped eyeglasses, In: Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 5674-5679, Nov 2013

C. Yang and Y. Hsu, “A Review of Accelerometry-Based Wearable Motion Detectors for Physical Activity Monitoring”, Sensors, 10, 7772-7788, (2010).

S. Zhang, P. McCullagh, C. Nugent, and H. Zheng, “Activity Monitoring Using a Smart Phone’s Accelerometer with Hierarchical Classification“, Proceedings of the 6th International Conference on Intelligent Environments, Kuala Lumpur, Malaysia, (2010).

T. Zhang and W. Tang, “Classification of Posture and Activities by Using Decision Trees”, Proceedings of the 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS 2012), San Diego, California, (2012).

M. Zhang and A. A. Sawchuk, “A Feature Selection-Based Framework for Human Activity Recognition Using Wearable Multimodal Sensors,” Proceedings of the 6th International ICST Conference on Body Area Networks (BodyNets), Beijing, China, (2011).

Zhu, L., Chen, Y., Torralba, A., Freeman, W., & Yuille, A. (2010a). Part and appearance sharing: Recursive compositional models for multi-view. In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on (pp. 1919-1926). IEEE.

Zhu, L., Chen, Y., & Yuille, A. (2010b). Learning a hierarchical deformable template for rapid deformable object parsing. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 32(6), 1029-1043.

Volume II. Cost Proposal

Section I. Administrative

**1. Cover sheet:**

a. BAA number: DARPA-BAA-13-22

b. Proposal title: Automated Shared Awareness Through Sensor Data Classification

c. Lead Organization submitting proposal: Artemek Technologies INC

d. Type of business, selected among the following categories: Other Small Business

e. All team members (if applicable) and type of business for each: University of Southern California (Other Educational)

f. BAA Technical Focus Area Addressed: Ground Systems

h. Technical POC 1:
Mr. Li, Chris
University of Southern California, Hedco Neuroscience Building, 3641 Watt Way Los Angeles, CA 90089-2520
chris@artemek.com

Technical POC 2:
Mr. Hwang, Jun
University of Southern California, Hedco Neuroscience Building, 3641 Watt Way Los Angeles, CA 90089-2520
jun@artemek.com

i. Administrative POC:
Mr. Hwang, Jun
University of Southern California, Hedco Neuroscience Building, 3641 Watt Way Los Angeles, CA 90089-2520
jun@artemek.com

i. Award instrument requested: Grant

j. Place(s) and period(s) of performance:
Research done at University of Southern California iLab and Artemek
Phase I: Aug 16th 2014 – Apr 15th 2015

Phase II: 4/15/15 – 4/15/16
Phase III: 4/16/16 – 12/16/17

k. Total proposed cost separated by basic award and option(s) (if any):
Phase I: $168,200 (grant)
Phase II (option): $256,800(grant)
Phase III (option): $171,200 (grant)

l. Name, address, and telephone number of the proposer's cognizant Defense Contract Management Agency (DCMA) administration office (if known): n/a

n. (14) Name, address, and telephone number of the proposer's cognizant Defense Contract Audit: n/a

o. Agency (DCAA) audit office (if known): n/a
(15) Date proposal was prepared: Jan 9 2014

p. (16) DUNS number: 079252316
(17) TIN number: 46-4474438
(18) Cage Code: 72YZ2
q. (19) Subcontractor Information; and

r. (20) Proposal validity period: 180 days

Section II. Detailed Cost Proposal

**1. Cost Proposal**

| | Phase I (8 mo) | Phase II (12 mo) | Phase III (8 mo) |
|---|---|---|---|
| **Personnel** | | | |
| Mr. Chris Li | $24,000 | $36,000 | $24,000 |
| Mr. Jun Hwang | $24,000 | $36,000 | $24,000 |
| **Fringe Benefits** | | | |
| Mr. Chris Li | $1,600 | $2,400 | $1,600 |
| Mr. Jun Hwang | $1,600 | $2,400 | $1,600 |
| **Equipment** | | | |
| 2x Precalibrated IMU | $9,000 | | |
| **Travel** | | | |
| Domestic Travel | $2,000 | $2,000 | $2,000 |
| **Participant Support Costs** | | | |
| Stipends | $2,000 | $2,000 | $2,000 |
| Travel | $2,000 | $2,000 | $2,000 |
| **Other Costs** | | | |
| Materials and Supplies | $2,000 | $6,000 | $8,000 |
| Consultant Services | $4,000 | $24,000 | $10,000 |
| Subcontract Costs (iLab) | $96,000 | $144,000 | $96,000 |

| | | | |
|---|---|---|---|
| **Total Cost to Agency** | $168,200 | $256,800 | $171,200 |

**2. Cost Justification**

**Personnel**

Personnel costs were calculated for a base salary of $3000/month each for Mr. Jun Hwang and Mr. Chris Li.

**Fringe Benefits**

Fringe benefits cover health insurance costs at $200 per person, per month.

**Participant/trainee Costs**

In order to cover domestic travel for military members whom we test our devices on, we request $4000 per phase for their flights, lodging, meals, and misc fees. Bringing in service members allows us to better focus our product designs and features for their needs.

**Travel Costs**

Artemek will fly to various locations for conferences and meetings from Los Angeles frequently, including: the greater Washington D.C. metro area, Fayetteville, North Carolina (Fort Bragg where JSOC, USAOC, FORSCOM resides), Tampa, Florida (MacDill Air Force Base where USSOCOM and USCENTCOM resides), Naval Amphibious Base Coronado and Marine Corps Base Camp Lejeune (where MARSOC resides).

**Consultant Services**

Artemek outsources work (such as industrial design, hardware testing, and manufacturing) to other design agencies and CMs.

**3. Subcontractor Cost Proposal and Justification**
Attached on next page.