

Representing Autonomous Systems' Self-Confidence through Competency Boundaries

Andrew R. Hutchins, M. L. Cummings
Humans and Autonomy Laboratory, Duke University, Durham, NC, USA

Mark Draper
Supervisory Control and Cognition Branch, Air Force Research Laboratory, Dayton, OH, USA

Thomas Hughes
Infoscitex Corporation, Dayton, OH, USA

A method for determining the self-confidence of autonomous systems is proposed to assist operators in understanding the state of unmanned vehicles under control. A sensing-optimization/verification-action (SOVA) model, similar to the perception-cognition-action human informational processing model, has been developed to illustrate how autonomous systems interact with their environment and how areas of uncertainty affect system performance. LIDAR and GPS were examined for scenarios where sensed surroundings could be inaccurate, while discrete and probabilistic algorithms were surveyed for situations that could result in path planning uncertainty. Likert scales were developed to represent sensor and algorithm uncertainties, and these scales laid the foundation for the proposed Trust Annunciator Panel (TAP) consisting of a series of uncertainty level indicators (ULIs). The TAP emphasizes the critical role of human judgment and oversight, especially when autonomous systems operate in clustered or dynamic environments.

INTRODUCTION

In human-autonomy interactions, a human operator's insight into the ability of an autonomous agent to achieve a successful level of performance in complex task environments is often obscured by the limited communication channel characteristics of such interactions. This is particularly problematic when task demands approach, or exceed, the competency boundaries of assigned agents. It would seem that the ability of autonomous agents to self-report a level of confidence under such circumstances would be extremely helpful in supporting effective human-agent coordination.

While computers are capable of performing complex computations much faster than humans, there still exists the need for operators to oversee and make critical decisions that computers are not yet capable of making. This inherent brittleness is exacerbated in autonomous systems that must deal with significant environmental uncertainty as well as uncertainty inherent in probabilistic reasoning (Cummings, 2014). Because humans rely on the information relayed by these systems to make critical decision, it would be useful for these autonomous systems to relay confidence of derived information and the potential courses of action, whether through sensors or decision algorithms, to operators who need to calibrate their trust in decisions and actions.

Issues related to human trust in autonomous systems have been well established as a fundamental issue in human-robot collaboration (Lee & See, 2004; Hoffmann, Johnson, et al., 2013); however it is possible that by bridging the path from robot-to-operator, the operator's trust in the autonomous system will be calibrated to the robot's potential (Clare, 2013). Figure 1 illustrates the human-machine trust loop in which an operator assigns a task to an autonomous system, in this case an Unmanned Vehicle (UV), with some trust that the task can be completed, and the feedback loop indicates the "awareness"

that the UV has of its surroundings and the vehicle's confidence that it can complete the assigned mission.

To evaluate the ability of a UV to generate some measure of self-confidence to an operator, the methods by which the vehicle perceives the environment will be investigated. In addition, representative algorithms that use this collection of sensor data will be evaluated to determine the scenarios where (1) false perception of the UV's environment is present, or (2) an inability to properly use the sensor data due to mission constraints is possible.

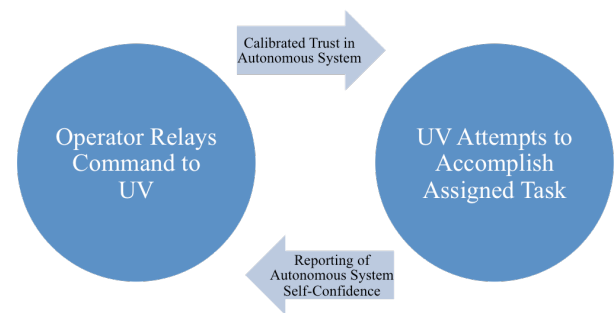


Figure 1: Human-Machine Trust Loop

This paper will focus on the navigation aspect of unmanned aerial vehicle (UAV) control as it relates to path planning, which is a common UV function. Two of the most common methods of UAV perception (LIDAR and GPS), two discrete path planning algorithms (Dijkstra's Algorithm and A*), and two probabilistic algorithms (probabilistic road mapping and rapidly exploring random trees (RRT)) are investigated for scenarios that could result in mission failure due to incomplete or inefficient path planning. By scaling the uncertainty in the discussed scenarios, a configural display for growing uncertainties is discussed. Finally, a Trust Annunciator Panel

(TAP) is described that could increase the operator’s situational awareness as it relates to the UAV’s world model across a set of functions, beyond just that of path planning.

AREAS OF UNCERTAINTY

Human supervisory control paradigms typically involve an operator overseeing the performance of a robot or autonomous system and making decisions that the system is not capable of making on its own. In such supervisory control settings, both humans and autonomous systems have a parallel information processing structure, which follows an input-processing-output model (Parasuraman, Sheridan, & Wickens, 2000). We propose that for autonomous systems, this model can be represented by the sensing-optimization/verification-action (SOVA) model in Figure 2.

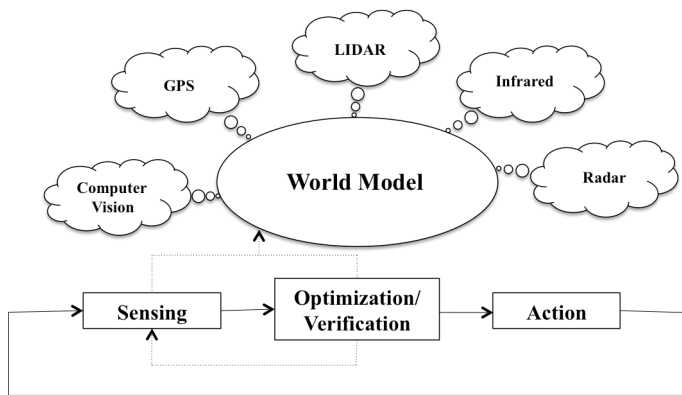


Figure 2: SOVA Model for Autonomous Systems

The method by which an autonomous UV perceives its surroundings is through series of sensors that work in concert to accomplish assigned tasks. The methods of perceiving the world model for the robot are quite different from those of the human model. While humans rely on their five senses, memory, and/or expectations to intervene with the surrounding world, the robotic system must rely on various types of sensors to view its much different world model. For example, LIDAR, GPS, infrared, and radar are all common methods of robot sensors used for environment perception (Hargis, Lang, et al., 1998).

Progressing past sensing the UV’s environment is optimization/verification of the perceived surroundings. During this section of the SOVA model the underlying algorithms determine what course of action the autonomous vehicle should take, such that the assigned task can be completed in a safe and efficient manner. Navigation via path planning and image recognition are examples of two such tasks that require complex processing.

The final block, the action block, is where the autonomous system relies upon preprogrammed scripts to determine how to execute the plan based primarily on the world model formed by the sensors, which will undoubtedly be incomplete.

The SOVA model is of particular interest when supervising non-deterministic (stochastic) controlled UVs. The intermediate feedback loop between sensing and optimization/verification allows for replanning when

unexpected events occur, such as interlopers, rapid weather change, and/or waypoint changes.

In order to determine whether an autonomous system can estimate its own competency limits, we propose that the uncertainty surrounding the SOVA elements must be understood and communicated to the operator. Towards this end, we will discuss how this could possibly be done, given current state-of-the-art technologies and algorithms.

Sensing

Two commonly used sensing methods were investigated that are often used on autonomous systems in an effort to generate a world model. The first of these methods was LIDAR. LIDAR is commonly used in air and ground UVs for mapping and object detection. However, during rainfall LIDAR provides inaccurate data due to the scattering of light from the sweeping laser. In addition, LIDAR only partially detects some transparent objects, such as glass, due to light being transmitted and detecting objects through these surfaces (Hernandez-Marin, Wallace, & Gibson, 2008).

The second sensing method analyzed was GPS. GPS has become the standard for automobile, air traffic, and maritime navigation over the past several years and can work in conjunction with LIDAR for UV location and path planning (and potentially other inertial navigation systems). However, for present-day UAV systems, GPS is typically a stand-alone navigation system, but it has limitations with coverage and location accuracy. UAVs that are tasked with operating in either an urban environment, densely vegetated area, mountainous locations, etc. are at risk of GPS signal loss, especially if the network is in low-altitude, uncontrolled airspace, typically under 1,200 ft.

The loss of GPS is especially problematic for control of UAVs if GPS is the only way that the operator can locate the vehicle or team of vehicles. In addition to reception loss, GPS struggles to relay accurate altitude measurements. In some cases altitude inaccuracies can be up to 2-3 times the uncertainty associated with planar GPS measurement error (U.S. Geological Survey, 2014).

We propose that there are two core dimensions of competency boundary estimation for autonomous sensors, that of output quality (i.e., can the sensor generate accurate data?), and mission impact (i.e., if the sensor output is inaccurate, how would the mission be affected?). While the first dimension provides clear quantitative metrics, the second is more contextual, i.e., a poorly performing sensor may not affect a mission if high degrees of accuracy are not needed.

Figure 3 represents these two uncertainty scales that rank problematic scenarios for LIDAR and GPS sensors. The initial scale in Figure 3(a) is an output quality scale with Likert rankings from 1-5, with 1 meaning a very accurate reading from the sensing device and a 5 meaning that there is a greater than 50% chance that the collected reading is inaccurate. These thresholds are notional and could be modified for the sensor under consideration.

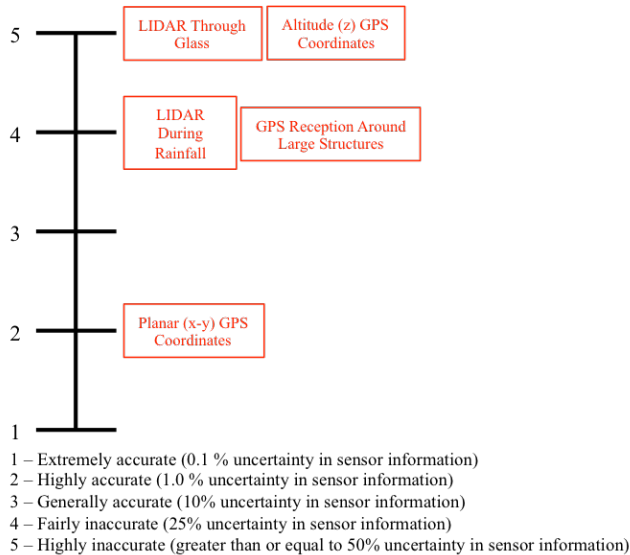
The increase in uncertainty for Figure 3(b) is based on whether the acquired data from the sensing device has the potential to impact mission performance. Therefore, if anomalies in the collected data have a high probability of

jeopardizing the mission, that sensor will be ranked high on the mission impact uncertainty scale.

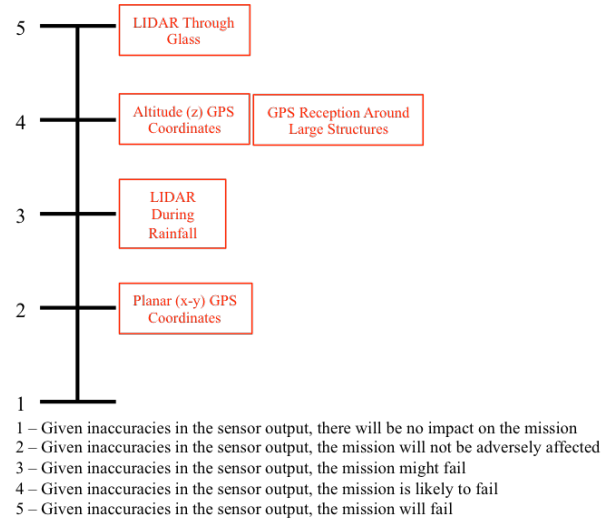
For example, it is known that readings for planar (x-y) GPS coordinates are reasonably accurate (within a few inches of the true value) (Du, Du, & Wang, 2009). Therefore, for the output quality scale, planar GPS was ranked as a 2, indicating that the acquired measurements are highly accurate. Given that planar GPS measurements might be inaccurate, as ranked in Figure 3(a), this inaccuracy would not be high enough to unfavorably affect the mission. Thus, it was ranked as a 2 on the mission impact scale, indicating that any inaccuracies in the sensor output will not adversely affect the mission.

A similar method was used to scale the altitude GPS coordinates, GPS reception, LIDAR during rainfall, and LIDAR through glass, recalling that just the function of navigation is the context. It was determined that the most uncertain scenarios for these sensors is in GPS altitude measurements (z) and LIDAR detection through transparent objects, such as glass (Figure 3(b)). Moreover, since GPS altitude measurements can be so inaccurate, low altitude operations for UAVs could be a significant driver of system trust (or the lack thereof).

The scales in Figure 3(a) and 3(b) are ordinal representations of the limits of the sensors' capabilities, and thus, represent the competency boundaries of these systems. By using such a scale, humans can better understand how sensor output can change under different conditions and when they should adjust their perception and confidence about a system's ability to meet mission goals. Once data has been sensed, it must be processed to lead to an actionable decision, as demonstrated by the optimization/verification block in the SOVA model (Figure 2). However, this processing can introduce additional layers of uncertainty. For the example of UV navigation, significant computation occurs in the task of path planning. Thus, two well-known discrete path planning algorithms and two common probabilistic path planning algorithms were investigated to determine what scenarios could result in uncertainty when applying sensor information to construct global paths.



(a)



(b)

Figure 3: Uncertainty Scales for Autonomous System Sensing (a) Output Quality Scale, (b) Mission Impact Scale

Path Planning Algorithms

The first algorithm investigated was Dijkstra's Algorithm, which is a greedy algorithm that will always find the least costly path without using a heuristic. However, since Dijkstra's Algorithm searches all possible local paths for the least costly, a large nodal network could require a vast amount of computational time and capabilities to determine a solution, which might be outside of the mission constraints in terms of time available.

A second discrete path planning algorithm that was examined is A* Search. A* applies a heuristic, which generates the sum of the cost to get to a state plus the estimated cost from that state to the goal. While A* performs better than Dijkstra's Algorithm in larger nodal networks, it might find a suboptimal solution in a multi-goal scenario due to the choice of the heuristic.

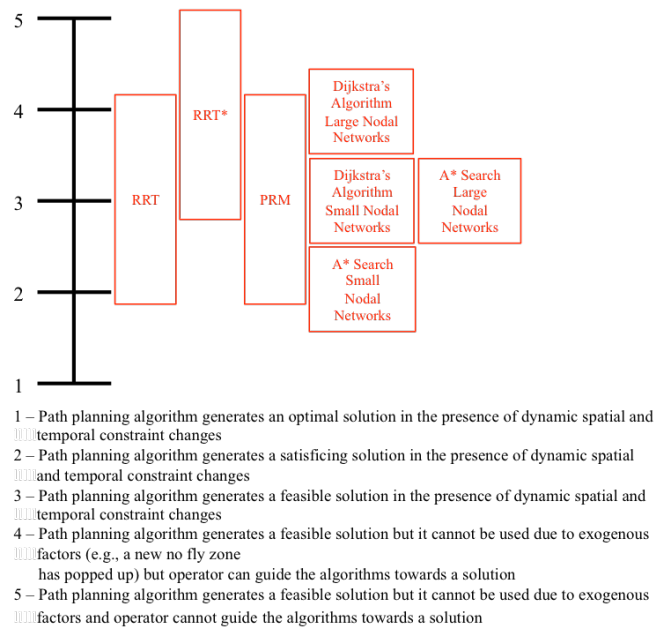
In addition to discrete path planning algorithms, two popular probabilistic path planning algorithms were considered, which generally treat the problem space as continuous instead of discrete. In Probabilistic Road Mapping (PRM), a sample of the UAV's location, velocity, etc. is used to predict collision-free paths. After possible paths have been projected, graph algorithms, such as Dijkstra's or A* can be used to develop a global solution. One of the central issues with probabilistic path planning algorithms is the time associated with developing a global path. While constructing local paths is relatively straightforward, the construction of a global path from these local paths can be time consuming.

The other probabilistic path planning algorithm considered was rapidly exploring random trees (RRT). For RRTs, random nodes are generated and a branch is formed from the nearest node already on the tree and extends out towards the new random node. The algorithm builds local paths from the new branches until a global solution is found. RRT* is an extension of RRT that continuously optimizes its paths to determine if a more optimal solution exists.

Figure 4 depicts similar uncertainty scales to the sensing uncertainty scales in Figure 3. Output quality (i.e., can the

algorithm generate feasible solutions?), and mission impact (i.e., if the solution is viable, how could variation in output affect the mission?) are representations of the competency boundaries of the optimization/verification block in Figure 2. Both discrete and probabilistic algorithms are listed in the scales together. The probabilistic algorithm blocks span multiple categories due to the random nature of the algorithms. Therefore, the environment that the UAV is operating in strongly influences the performance of these probabilistic algorithms. The uncertainty rankings in the output quality scale in Figure 4(a) are dependent on whether an operator has any ability to rectify a potentially infeasible solution. The mission impact scale in Figure 4(b) addresses the spatial and temporal constraints of the mission and how useful the solutions are.

For Figure 4, it is important to discuss the notion of feasibility. While the algorithms in Figure 4 will generate feasible solutions given known variables and state spaces, one problem common to military command and control systems is the rapidly and ambiguously changing battlefield. For example, if state space changes, such as a rapidly emerging no-fly zones, cannot be communicated in time to mission planners, then despite producing technically ‘feasible’ solutions, the solutions will be practically infeasible since any path planned through the no-fly zone (unknown to the algorithm) is an unacceptable path to the operator.

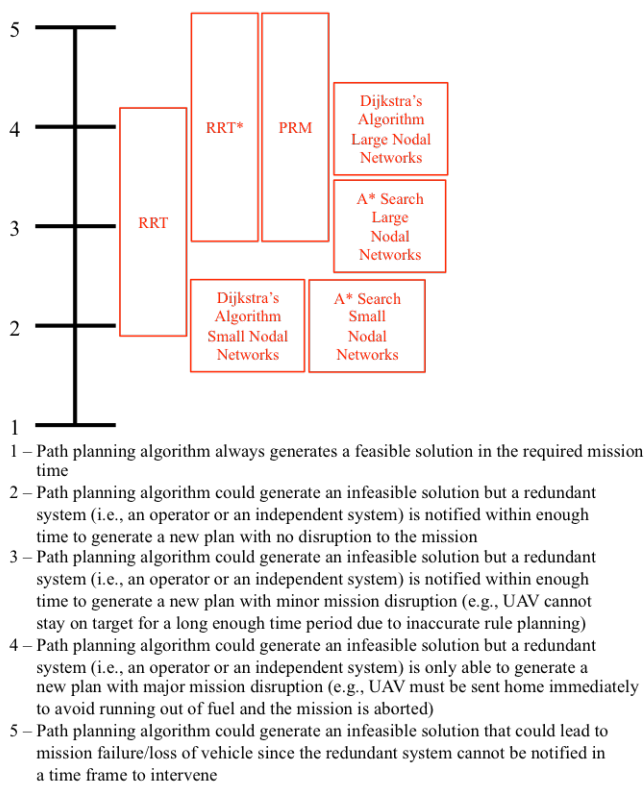


(b)
Figure 4: Uncertainty Scales for Autonomous System Path Planning Algorithms (a) Output Quality Scale, (b) Mission Impact Scale

TRUST ANNUNCIATOR PANEL (TAP)

Using the scales presented in Figures 3 and 4, a configural display (Gibson, 1979) was designed to better illustrate the potential for growing uncertainty based on the navigation environment of one or more UAVs. The display shown in Figure 5 is called the uncertainty level indicator (ULI), which is composed of four quadrants, each of which represent one of the four scales presented in Figures 3 and 4. It is intended to represent uncertainty along a single dimension, which in this case is the navigation function.

The empty ULI in Figure 5(a) indicates that the UAV is operating in an area that has few obstacles and detrimental conditions, therefore allowing the sensing hardware and algorithms to function properly. A ULI begins to fill due to some environmental constraint that causes the level of uncertainty to increase relative to the inaccuracies provided by sensing systems (Figure 5(b)). After the second tick mark has been reached, the sensor inaccuracy rises to a level such that the quality and the mission impact of recommended solutions from the algorithm are brought into question (i.e., the Sensor Mission Impact quadrant is red and the Algorithm Output Quality and Mission Impact quadrants are yellow) (Figure 5(c)). The final color change occurs after the third tick mark where the quadrant(s) turn red indicating high levels of uncertainty. As the situation approaches these conditions, reliance on the solutions provided by the algorithms should be questioned in terms of both their quality and mission impact. At this point, the operator should contemplate planning the path for the UAV to ensure a safe path is constructed in order for the UAV to complete its assigned mission (Figure 5(d)).



(a)

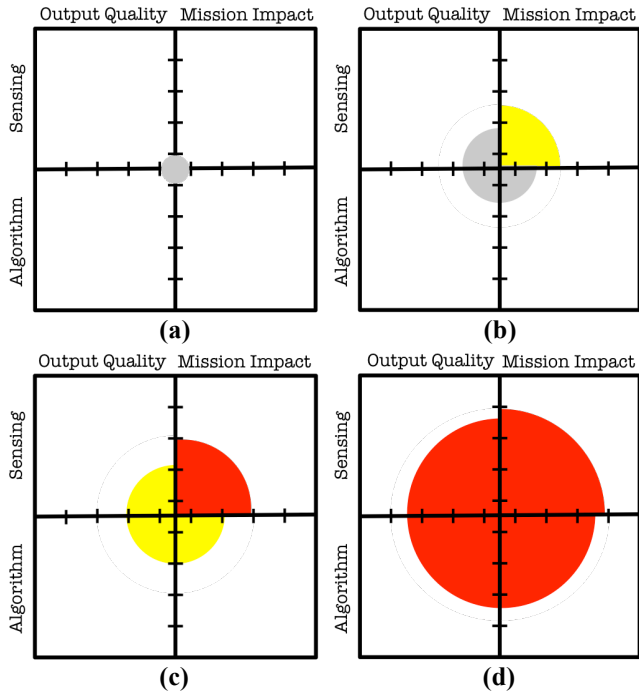


Figure 5: Uncertainty Level Indicator (ULI) with Increasing Uncertainties Proceeding From (a) to (d)

The navigation ULI only expresses the uncertainty of the ability of the sensors and algorithms to formulate a feasible and robust path. Given that this ULI only addresses one dimension of UAV control, other aspects of the mission might need to be included, such as payload management and communications. So to accurately represent to the user a multi-dimensional representation of uncertainty across various UV functions, ULIs could be grouped.

Figure 6 illustrates the Trust Annunciator Panel (TAP), which is a combination of ULIs that account for various mission aspects. The operator would have the TAP available for a high level status review of one or more UAV's self-confidence for its assigned tasks and mission. This communication method allows for an operator to quickly grasp when a problem has emerged or has worsened. For example, when the lower algorithm quadrants are red, the operator knows that a manual path needs to be constructed due to some intervening object or mission time constraint.

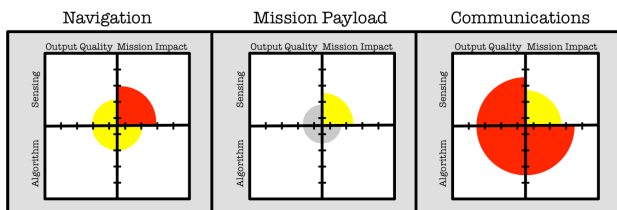


Figure 6: Proposed Trust Annunciator Panel (TAP)

CONCLUSIONS AND FUTURE WORK

In this exploratory effort a configural display has been proposed that allows for operators to monitor the self-confidence of a single, or team of UVs. The Trust Annunciator Panel (TAP) is comprised of multiple uncertainty level indicators (ULIs) that inform the operator of the potential

inaccuracies in the sensing hardware, as well as the underlying algorithms.

The ULIs and TAP represent both a bottoms up and top down approach to communicating a system's confidence level to an operator so that appropriate trust can be developed. The scales, as represented in Figures 3 and 4, represent both a data driven approach (i.e., sensor output quality) as well as a goal-based cognitive approach (i.e., can the operator intervene to correct underlying poor assumptions or constraints?) The configural implementation of the TAP display reduces the level of reasoning to perception-based and reinforces the human-value added when completing missions in dynamic or unknown environments. Thus the ULI design enables direct perception interaction, which is a critical design consideration for operators under high workload and time constraints.

The preliminary work presented here only addresses the navigation ULI, which is strictly the path planning process for the UAV. Future work is needed to conduct a similar analysis for mission payload and stakeholder communication, both of which are fundamental in various missions such as vehicle tracking, UAV team operations, supply delivery, etc. Moreover, human-in-the-loop testing is needed to further calibrate the TAP/ULI displays, as well as determine if it is feasible to combine more than a single functionality into one ULI, and determine/communicate the utility of the manifestation of an autonomous agent's self-confidence for improving human-agent teaming.

ACKNOWLEDGEMENTS

The work presented was supported by the Air Force Research Lab (AFRL) and Infoscitex Corporation.

REFERENCES

- Clare, A. S. (2013). Modeling real-time human-automation collaborative scheduling of unmanned vehicles. *Ph.D. Thesis, MIT, Cambridge, MA.*
- Cummings, M. L. (2014). Man versus machine or man + machine? *IEEE Intelligent Systems, 29(5)*, 62-69.
- Du, X., Du, D., & Wang, K. (2009). Position accuracy evaluation of GPS receiver under small UAV flight environment. *Proceedings of the 9th International Conference on Electronic Measurements & Instruments, 2-29.*
- Gibson, J. J. (1979). *The ecological approach to visual perception.* Psychology Press.
- Hargis Jr, P. J., Lang, A. R., Schmitt, R. L., Henson, T. D., Daniels, J. W., Jordan, J. D., Shroder, K. L., & Shokair, I. R. (1998). Sandia multispectral airborne lidar for UAV deployment. *Proceedings of the 4th Joint Workshop on Standoff Detection for Chemical and Biological Defense, 26-30.*
- Hernandez-Marin, S., Wallace, A. M., & Gibson, G. J. (2008). Multilayered 3D LiDAR image construction using spatial models in a Bayesian framework. *IEEE Transactions on Pattern Analysis and Machine Integration, 30(6)*, 1028-1040.
- Hoffmann, R. R., Johnson, M., Bradshaw, J. M., & Underbrink, R. (2013). Trust in automation. *Intelligent Systems, IEEE, 28(1)*, 84-88.
- Lee, J. D., & See, K. A. (2004). Trust in automation: designing for appropriate reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 46(1)*, 50-80.
- Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, 30(3)*, 286-297.
- U.S. Geological Survey. (2014, September 3). USGS global positioning application and practice. Retrieved February 17, 2015, from <http://water.usgs.gov/osw/gps/>.