# Cover sheet

BAA number: **DARPA-BAA-11-13**

Proposal title :**Human Computing and Machine Learning Assisted Decentralized UAV Control**

Lead Organization submitting proposal : Polytechnic University of Puerto Rico

Type of business, selected among the following categories: “LARGE BUSINESS”, “SMALL DISADVANTAGED BUSINESS”, “OTHER SMALL BUSINESS”, “HBCU”, “MI”, “OTHER EDUCATIONAL”, OR “OTHER NONPROFIT”

Contractor’s reference number

Team members (if applicable) and type of business for each

BAA Technical Thrust Area Addressed: **2.a Unmanned Systems**

Technical point of contact:
Arturo Geigel
Polytechnic University of Puerto Rico
377 Ponce de León Ave. Hato Rey, PR 00918
Phone: (787) 622-8000 ext 472     Cel:(787)298-3672
ageigel@pupr.edu

Administrative point of contact to include: salutation, last name, first name, street address, city, state, zip code, telephone, fax (if available), electronic mail (if available), total funds requested from DARPA, and the amount of cost share (if any) AND

Date proposal was submitted.

Affirmation of existing SETA support contracts : none

**Innovation**.

The proposed research aims at providing a proof of concept of decentralized control of unmanned vehicle. The concept aims to solve two major problems in remotely operated vehicle decentralization. The first problem that will be tackled is the problem of deciding the best alternative course of action based on a competitive voting scheme of multiple inputs from various operators. The second problem to address is that of clearance for the multiple operators that will be competing for the control of the unmanned vehicle. The proposed solution to both problems is to provide a “real time” morphed video sequence that changes the recognizable features of persons and objects to operators allowing non-cleared personnel to carry out the mission and allows multiple operators without clearance to view the unmanned vehicle video stream. The proposed solution rests in current technical advances in massive parallel architectures such as GPU computing and current state of the art image transformations, image retrieval and machine learning techniques. GPU technology allows for greater efficiency in computing real time images and provides an alternative platform on which to carry out identification and transformation of images that will allow smooth interpolation between the image stream of the unmanned vehicle and stored images.

Current state of the art research is focused on providing a single operator for multiple unmanned vehicles. The proposed research on the other hand aims to allow multiple operators to operate a single vehicle providing redundancy of the decision process of the operator by optimizing the course of action to be followed. The same stream of research can be applied to allow multiple operators to operate multiple vehicles. In this case, the process will eliminate single operator point of failure to multiple vehicle operation and allow for plane formation division to engage individual operators in case of unexpected events where a single operator cannot provide granular responses to a formation of planes.

**Expected Results**

The expected deliverable will be a proof of concept that will explore the feasibility of such real time operations on a scaled model of currently deployed military unmanned vehicles. The prototype will include an infrastructure that can be adapted to unmanned vehicles in use by the military that, in principle, will allow for the use of non-cleared personnel to carry out missions without the need for briefing with details of the missions.

It is expected that most of the basic research on computer architecture, computer configuration, and real time morphing will be implemented and documented in the three year span of the project. It is also expected that details of latency issues in image transformation as well as communications delay will be documented.

A basic prototype of an aerial reconnaissance remotely operated vehicle with its controlling infrastructure will be delivered at the end of the project. The expected deliverable will also allow future research to integrate multiple vehicles in formation.

**Technical Rationale**.

The impact on mission deployments will be substantial since the system could, in principle, allow a person without security clearance to pilot the vehicle without knowing the true objective of the mission. This will have an advantage of minimizing the need to know in sensitive operations. Through image morphing of the principal mission elements and the scenery, the operator will not be able to identify the scenery nor the persons involved in the mission and will not be able to distinguish a mission from a training exercise or a game.

The other benefit will come through redundancy in operational command of the vehicle which will be decentralized, this means that the equipment as well as the operators of the vehicle can be geographically distributed and avoid a single point of failure.

**Technical Approach**.

The technical approach involves four main components that will be used in creating the necessary infrastructure to support the feasibility of the project as shown on **Figure 1**.

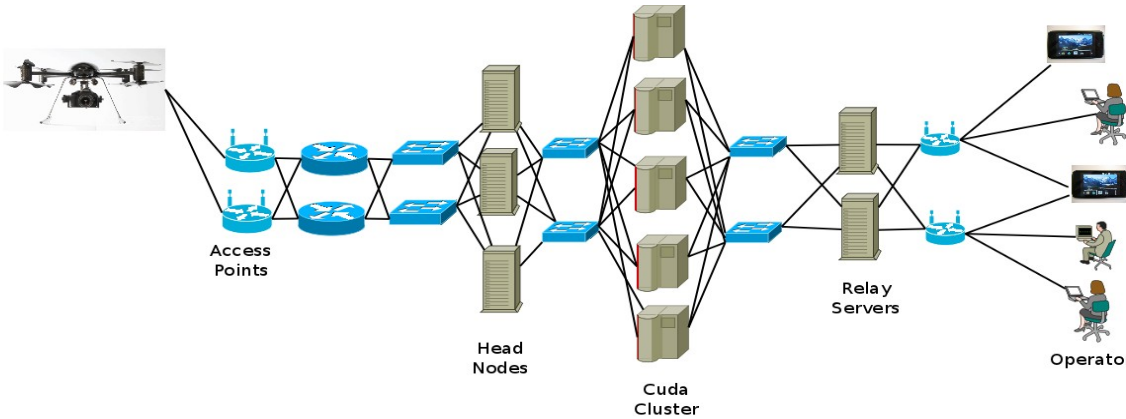



**Figure 1.**

The infrastructure will consist of the unmanned vehicle which will communicate through wireless access points that will, in turn, transmit the information to the head nodes. The head nodes will distribute the load to the CUDA clusters which will do the image processing part of the project and will also hold access to the image databases. After the images have been processed, the CUDA cluster nodes will forward the information to the relay servers which will organize the information and send it to the individual operators.

**Components of the proposed Research**

The proposed infrastructure will include the following components that make the core of the project:

1. Human computing games that will allow tagging of the images
2. CUDA processing for image identification and morphing
3. The communications infrastructure to handle the load and provide operator redundancy

The first component will aim at creating *human computing* games [1] which will allow tagging the images from video stills. These tagging will allow validation of current state of the art clustering techniques that will be used for the project. This phase will allow the creation of a tagged database that will be used for the project.

The second component will be the main thrust of the research. This stageinvolves the setting of a computing infrastructure that will provide the required computational bandwidth to process both incoming stream data as well as stored tagged data. This infrastructure will be controlled through custom software that has been optimized for parallel execution on CUDA processors. In addition, efficient distributed databases must be deployed to reduce query time. The use of carefully indexed data based on the human computing and clustering tags will allow for distributed storage of the images. The retrieval phase will be developed using methods such as inverted files [2] as well as more current techniques such as map reduce. The identification for retrieval will be based on two different processes. The first, is the inverted file query based on neural networks and, the second, will be keyword-based for redundancy. Both will depend on efficient relevant features extracted from the incoming stream of images. The features will be extracted using traditional features algorithms (PCA, Sobel, etc.) as well as statistical methods [3] as shown on figure 2.

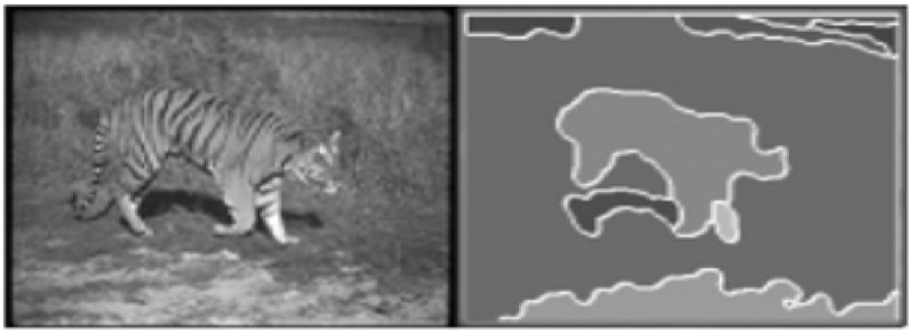


**Figure 2 Feature Selection**

Once the features have been extracted and the candidate images have been extracted, the selection of mesh control points will be carried out. These control points will again be determined with statistical methods as well as traditional backpropagation and other neural network architecture [4][5]methods trained from the stored data. From there, a mesh warping technique [6] will be used to transform the image into a composite of the incoming stream image overlapped with the stored image (see figure 3). The main objective is to transform objects within the same category such as person to person building to building, etc.

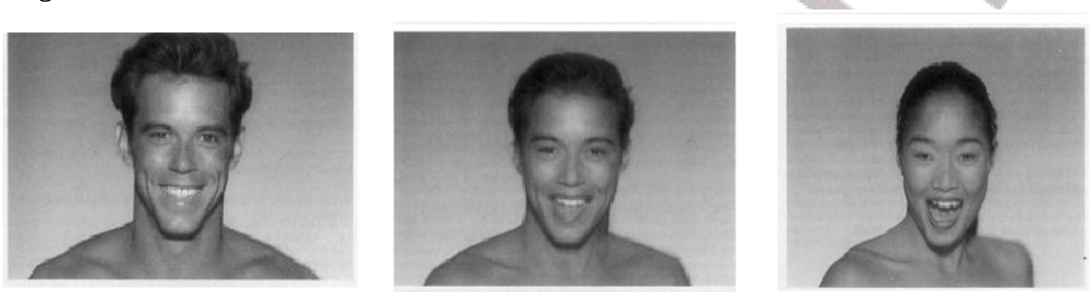



**Figure 3 Mesh warping**

The algorithms described above will all be parallelized for CUDA parallel processing execution. Careful segmentation of the processing stages will be determined during the experimental part of the proposal to determine optimal execution.

The third component will consist on providing the communications infrastructure to connect the unmanned vehicle to the computing infrastructure that carries out the required computation. This infrastructure will consist of incoming wireless signals that are transmitted from the unmanned vehicle to the receiving stations. These receiving stations will be redundant to avoid single point of failure. A "capture the flag" configuration will be implemented where both stations transmit the same information to the processing head nodes. The fastest receiving station to deliver the message to the processing head nodes will be the winner and the rest is discarded. The head nodes will determine the processing pipeline based on the current load of the system and priorities will be assigned based on feature salience and predetermined importance (person's head vs. trees, etc.). The processing servers will carry out the above described process in a pipeline and produce the final morphed image which will then be sent to relay stations which will transmit the signal to multiple operators. Multiple operators will then compete for the optimal decision of the execution path of the vehicle. A voting scheme algorithm [7][8] will be implemented where the best solution based on consensus between the information stored on a neural network and that provided by the operators will take place. The action with the most votes will be the one carried out by the vehicle. Due to expected delays that will be quantified during the scope of the project, a collision avoidance algorithm also has to be implemented in the unmanned vehicle.

Initial testing of the prototype will be limited to daylight conditions to test the concept and reduce the number of experimental variables to tackle in the project. Latency issues associated with the image processing overhead and network hops are the main identified risks of the project and one of the most important issues to tackle to guarantee the success of the prototype.

**Experience**.

Arturo Geigel (principal)- Professor Geigel has over 13 years in the computer security field and 5 years in the machine learning field. Professor Geigel has designed the neural Trojan which is a novel compromise on these architectures. His current research in this area is focusing on the theoretical as well as practical embedding of long sequences of payloads. He is also researching the compromise on other architectures such as K-means, Kohonen networks, adaptive resonance and support vector machines. He is also working on theoretical research of backpropagation and simulated annealing in deep architectures and applied research on deep learning for speech recognition and image processing. Geigel has also worked as a computer forensic consultant on 50 cases including speech enhancement, speech authentication, image verification, image artifact analysis as well as data carving. Professor Geigel is working with students on morphing algorithms, and cell phone control of remote servo devices. He is also working with students on a Homeland Security grant on research of data mining for terrorism detection.

Juan F. Torres

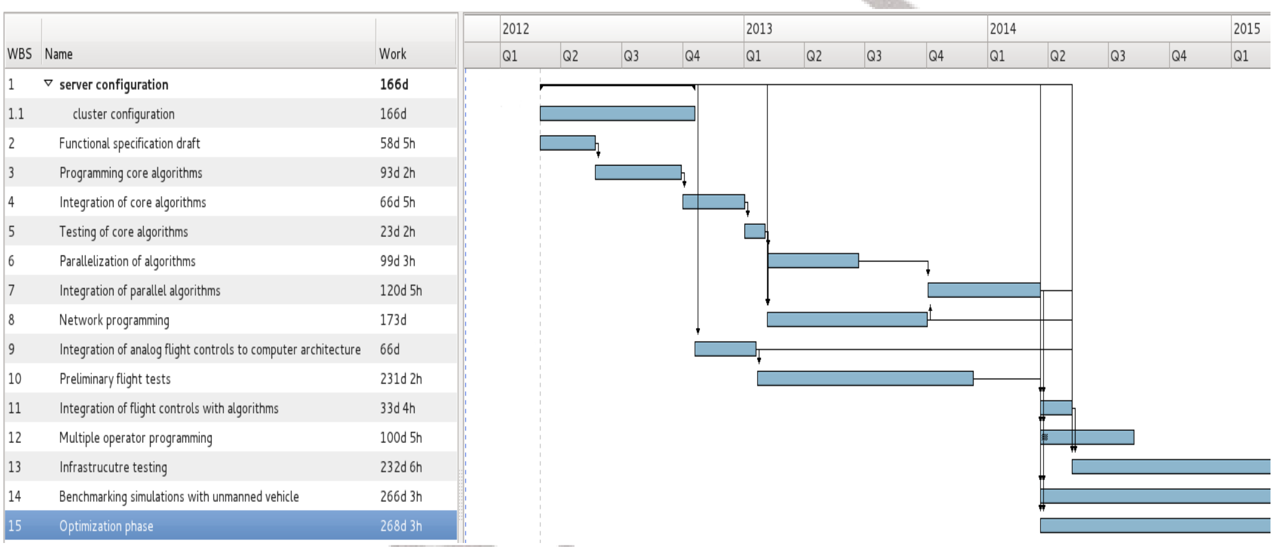
Dr. Juan F. Torres earned a PhD in Electrical and Computer Engineering from Georgia Institute of Technology (GT) in May 2010. He is an Associate Professor at Polytechnic University (PUPR), Dept. of Electrical & Computer Engineering and Computer Science (ECECS), where he teaches graduate courses in Pattern Recognition, Computer Networking, and Algorithms. For two consecutive years, he mentored undergraduate students within an NSF-sponsored summer research program (REU) at Georgia Tech, where he also developed and administered workshops on Digital Signal Processing topics as part of an outreach program for three consecutive years. Dr. Torres has conducted and published research in the area of speech processing, particularly in the parameterization of speaker's voices, the recognition of speaker emotion and affective state, and the possibility of detecting deception in speech. Most of his previous research has involved extensive application of pattern recognition techniques.

Felix Nevarez- Professor Felix Nevárez has been working in the computational engineering area for the past 10 years. From 1997 to 2000 was part of the Thermal Radiation Group at Virginia Tech where he designed and developed a computer program for Monte Carlo ray trace Used graphical environment at Virginia Tech for the analysis and design of space borne based radiometers. From 2003 to 2006 he worked as a consultant for Pratt & Whitney where he modify the code developed at Virginia Tech to predict the IR signature behind turbine engines due to hot surfaces and gases. For the last year he has been working with other professors in developing a computational model to predict the spatial distribution of hot electron rings and their properties such as their energy distribution function, trajectories and spatial confinement inside an Electron Cyclotron Resonance (ECR) machine in the Plasma Lab at the Polytechnic University of Puerto Rico. Recently Dr. Marvi Teixeira and Prof. Nevárez won a DOD grant to implement and study a parallel cyclic convolution algorithm developed by Dr. Teixeira. This algorithm will to be implemented in a 64 cores Beowulf cluster located in the High Performance Lab at PUPR.

### Cost

| Description | | Cost year 1 | Cost year 2 | Cost year 3 | Total |
|---|---|---|---|---|---|
| Unmanned vehicle (Draganflyer X6 Helicopter<br>**Model:** DF-X6-STILL-PKG ) | | - | $60,000 | - | $60,000 |
| Servers and Telecommunications equipment | | $80,000 | $30,000 | - | $110,000 |
| | | | | | |
| Additional parts, equipment and supplies | | - | $15,000 | $15,000 | $30,000 |
| | | | | | |
| Personnel | | $50,000 | $50,000 | $50,000 | $150,000 |
| Fringe benefits | | $16,500 | $16,500 | $16,500 | $49500 |
| Programmers | | $110,000 | $110,000 | $110,000 | $330,000 |
| Consultants | | - | $3,000 | $3,000 | $6,000 |
| | | | | | |
| Total Direct costs | | $256,500 | $284500 | $194500 | $735,500 |
| Total indirect costs | | | | | |

### Schedule




The proposed schedule is broken down into 15 sub tasks as shown in the projects Gantt chart.

### Milestones

Each of the 15 major tasks has a defined deliverable as the milestone for that particular task. The milestones are divided into four broad categories that contain multiple tasks each with its own deliverable.

<u>Algorithm Milestone Deliverables</u>

1. Connected network infrastructure to carry out the described work
2. Functional and design specification for the project
3. Programmed non parallel sub-components that will make all the infrastructure processing
4. Integration of the core algorithm into the projects functional components
5. Refined code that will be obtained by the testing phase

<u>Parallel Programming Milestone Deliverables</u>

6. The initial code in parallel form that will use CUDA programming and possible OpenMP routines
7. Programming that has the integration of both CUDA and OpenMP programming, as well as integration of MPI programming from task 8
8. MPI routines needed for the project's networked infrastructure communication that will be used in deliverable 7

<u>Unmanned Flight Vehicle Components Milestone Deliverables</u>

9. The code and circuits that will connect the analog flight controllers to the digital components of the infrastructure
10. Documented flights testing the infrastructure performance
11. Integration of flight controls with the algorithms of deliverable 7, merged into a single functioning unit

<u>Benchmarks and Testing of the Infrastructure Milestones</u>

12. Test flight data and performance of the algorithms that control multiple operators
13. Consolidated test flight data analysis
14. Benchmarking of the infrastructure and identification of poor performance elements in the infrastructure
15. Optimization phase of any code that shows poor performance in deliverable 14

**References**

[1] Luis von Ahn and Laura Dabbish. General Techniques for Designing Games with a Purpose. Communications of the ACM,. pp 58-67,2008 .

[2] H. Muller, D. M. Squire, W. Muller, and T. Pun. Efficient access methods for content-based image retrieval with inverted files. In S. Panchanathan, S.-F. Chang, and C.-C. J. Kuo, editors, Multimedia Storage and Archiving Systems IV (VV02), volume 3846 of SPIE Proceedings, Boston, Massachusetts, USA, September 20–22 1999.

[3] B. C. Ko, J. Peng, and H. Byun, “Region-based image retrieval using probabilistic feature relevance feedback,” Pattern Anal. Applic. (PAA), vol. 4, pp. 174–184, 2001.

[4] ByoungChul Ko, Hyeran Byun, "Probabilistic Neural Networks Supporting Multi-Class Relevance Feedback in Region-Based Image Retrieval," Pattern Recognition, International Conference on, p. 40138, 16th International Conference on Pattern Recognition (ICPR'02) - Volume 4, 2002

[5] G.E. Hinton and R.R. Salakhutdinov, Reducing the Dimensionality of Data with Neural Networks, Science, 28 July 2006, Vol. 313. no. 5786, pp. 504 - 507.

[6] G. Wolberg, “Image Morphing: A Survey,” The Visual Computer 14, pp. 360 –372 ,1998.

[7] L. I. Kuncheva. A theoretical study on six classifier fusion strategies. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(2):281–286, 2002.

[8] M. van Erp, L. Vuurpijl, and L. Schomaker, “An overview and comparison of voting methods for pattern recognition,” in Proc. 8th Int.Workshop Frontiers Handwriting Recogn., 2002, pp. 195–200.