**University of Kent: School of Engineering and Digital Arts**

**EL600 Project**

Literature Review

AVINSoR Project

**Artificial Visual Intelligence and Navigation System for application in Robotics**

[www.](http://www.Birring.net/Tej)birring.net/tej

**By**

Gurtej Singh Birring ([gsb4@kent.ac.uk](mailto:gsb4@kent.ac.uk))

**Dr Peter Lee**

# Introduction

The AVINSoR project’s initial aims are to investigate *biologically-inspired* and   
*“alternative-architecture”* embedded systems that attempt to recreate intelligence – i.e. systems which make *smart, adaptable decisions* and havethe ability to learn from the environment in which the operate - this should ideally comprise: *the ability to learn from feedback of their own behaviour, “extrasensory” inputs* and *training data* from the larger system of which they are a part*.* This project is particularly interested in the application and implementation of artificial intelligence in self-navigating automobiles, robots, quadcopters and drones – but only when the central control system is constrained to the scale of an embedded system.

## Aims and Objectives

The expected outcome of this endeavour, other than an *extensive* study of the *state-of-the-art*, is the design and development of an *intelligent, trainable control system* with the ability to relate data from the *visual input field* and other “*input” fields* to an array of digital outputs – which will be harnessed by external circuitry to create kinetic movement in a miniature (non-human size) autonomous land vehicle (demonstration configuration). The actual demonstration of the system will be limited to land – largely due to time constraint and budget.

### *Naturally* Learning Control System

The emphasis on *biological-inspiration* and *machine learning* is to reduce, to the maximum extent possible, the need to program and re-program the system for different applications, environments and situations (*adaptability)*, and to reduce the number of external instructions to the greatest extent possible - other than for *“training”*) – making the *learning process* of the machine seem *as natural (and effortless) as possible* (i.e. ***ideally*** the system should be *trainable* but not *programmable –* as ultimately, there would be no need to rectify the output response of the system to the visual stimuli through anything other than *training*)*.*

### Self-Navigation

It may also be worth noting, at this point, that the word “navigation” itself may imply different types of behaviour (or “navigationaloutcome”) for different purposes/applications. For example, if the purpose of the *robotic vehicle* is solely *discovery*, then there is no *final destination* as such - and the spatial perception of the environment is generated from scratch and does not, in any way, rely on previously available external data such as GPS street or terrain maps, but rather new “sensory information” (and knowledge of how sensory information is collated to “perceive” objects). For other applications, such as navigating to a particular *road*, existing knowledge is required for identifying the relative direction and distance of the selected *destination* from the position and direction of the current. If such data is readily available at a suitable engineering cost, it *should* be used.

# Current Work

There have been major advances in the field of self-navigating air and land vehicles in the last few decades, in fact, the deployment of unmanned aerial vehicles (UAVs) has risen vastly – and often utilize automated navigation systems. Miniature self-navigating vehicles such as quadcopters and other forms have also been realized over the last few years.

## Self-Navigation Systems

Before we investigate advancements in autonomous (self-navigating) vehicles, it is important that our particular niche is defined – so that knowledge from *a wide range of* studies of (or related to) this field can be appropriately identified.

### The Niche

Firstly, the central aim of this project is to design and develop a ***core***i.e. a *visually-intelligent control system with the ability to learn output behaviour* by being *trained* to correlate sensory (visual and non-visual) data. The visual perception aspect (i.e. all the necessary image processing, object detection, tracking, and recognition operations) should be integrated within the system’s architecture.

Secondly, *miniature autonomous land vehicle* demonstration aside, the core *should* be *versatile* i.e. able to be used for other *miniature* robotic applications such as that of a quadcopter or ornithopter. This does *not* mean research and innovation focussed on more complex (large/human-sized) vehicles should not be considered.

### State of the Art

#### Miniature Autonomous Vehicles

This is the area of research and innovation most directly relevant to the AVINSoR project, as navigational automation and central control of these miniature vehicles would already be implemented in the format of a suitable-sized embedded system.

An autonomous navigation system was implemented on a small quadcopter using the Odroid-U2 development platform[[1]](#footnote-1). The method utilizes *visual* and inertial sensors for GPS-independent *3D pose estimation* which would be especially useful in a GPS-denied environment, and *3D reconstruction of overflown terrain* to find potential landing sites and generate landing trajectories for the purpose of *autonomous landing* [1].

DelFly Explorer, a dragonfly imitating ornithopter, with its initial origins as an undergraduate BSc project at Delft University of Technology, attracted media attention for being the world’s smallest and lightest self-navigating drone [2]. The ornithopter is currently able to perform collision-free flight on its own, utilizing stereo-vision imaging and algorithms capable of navigating through indoor environments that lack visual texture [3] [4].

#### Autonomous Land Vehicles

Kuang-Hsiung Chen and Wen-Hsiang Tsai describe a mathematical model for calculating trajectories of movement around (multiple) detected objects to avoid collision in an outdoor road environment [5].

Google’s Self-Driving Car project, and the many automated vehicles contesting in the DARPA Grand Challenges (a series of competitions for American autonomous vehicles held and funded by DARPA - Defence Advanced Research Projects Agency) over the last decade, are examples of innovation in the field of autonomous land vehicles.

Non-visual sensory data is often used to map the surrounding environment, for example, Google’s Self-Driving Car utilizes the Velodyne 64-beam laser to generate a 3D map, as well as four radars, a GPS and an inertial measurement unit – it may also be worth noting that a camera is only used to detect traffic lights [6]. Stanley (DARPA Grand Challenge 2005 winner) also used a similar approach, generating a 3D map of the terrain in front using a LIDAR (light detection and ranging) unit, because the team soon realized that GPS units where never enough to handle all off-road hazards [7].

Tartan Racing (DARPA Urban Challenge contender) decomposed the problems of urban driving into five components: *mission planning layer* to determine the most effective route amongst a network of roads, a *behavioural layer* to execute the route by adapting to local traffic and exceptional situations, a *motion planning layer* to safeguard the vehicle by identifying and selecting the most appropriate trajectories, the *perception layer* that combines data from LIDAR, radar and vision systems to map location of static objects, shape of road, and other vehicles, and the *mechatronic system layer* designed to provide power, sensing and mobility [8].

A section of IEEE Spectrum magazine’s website is dedicated exclusively to articles on the latest research and innovation with regards to self-driving cars [9]. Just over a decade ago, Sandia National Laboratories (USA) published a report on the state-of-the-art in Autonomous Land Vehicle Systems and Technology [10]. Though much progress has been made over the years, the report does explain concepts such as Mobility, Localisation, Mission and Task Planning with regards to ALVs.

#### Autonomous Air Vehicles

There are many methods of nature-inspired target tracking mechanisms which are utilized in UAVs for automation of navigation as well as other applications [11].

A purely visual method for 3D UAV motion estimation was also proposed at Imperial College London [12].

## Specific Areas of Research

As we have discovered, navigation systems for autonomous applications are based around research and technical innovations of various fields, some of which are considered outside the bounds of computing and electronic engineering (in the conventional sense at least). Hereon, we will discuss the latest work with regards to human visual perception, object recognition, machine learning and anything else of interest to autonomous navigation systems.

### Human Visual Perception

Human vision and human visual perception are highly complex fields of research, and yet, still not fully understood. Many (but not all) features of human vision can be implemented in computer vision positively. One such example is stereoscopy, which is the remarkable ability to convert two somewhat separate images into a single perception of objects in a three dimensional space.

The understanding of human visual perception may allow us to better the design of existing approach. However, not all features and processes of human vision can be directly implemented as positives in computer vision.

Human visual perception involves the study of three biological *devices* [13]:

1. **The eye** – our understanding of this is largely as a physical model. The function of this sensory organ is determined by its physiology (biological functions of its organic chemical components).
2. **The neural system** - our understanding of this is only as an experimental mathematical/logical model, as it’s considered difficult to determine the exact neural connections of the brain (and connected organs, tissues, and cells) precisely.
3. **The brain** - a psychological model – as we cannot access or model the brains processing directly, hence we can only determine behaviour by experimentation and inference.

There are many somewhat complex organic image processing operations involved in the human visual brain which allow us to adapt to low light conditions and changes in (or lack of) luminosity [13]. Deducing depth from motion, finding edges and boundaries, forming an image from the individual pulses from alpha, beta and gamma cells, perceiving objects and creating a hierarchical neural structure to determine how an object/feature should be efficiently and effectively perceived is something which the human visual system naturally does [14].

### Computer Vision and Object Recognition

Object detection and object recognition techniques are mathematical models, usually based on the organic image processing functions of the eyes that are used to detect edges, boundaries, etc. in order to determine objects of

A survey on real time object detection, tracking and recognition published earlier this year in the International Journal of Computer Applications [15] outlines many recent studies on object *detection* and *tracking* that could potentially be applied to automated navigation:

* An approach to real-time object detection based on Modified Directional Lifting-Based Discrete Wavelet Transform (MDLDWT) performed particularly well detecting *multiple* moving objects, even when configured with low resolution input [16].
* A method of detecting salient objects combining two types of object saliency – *centre surround* saliency and *integrated* saliency [17].
* A real-time object detection method for video surveillance applications utilizing enhanced edge localization mechanism and gradient directional masking performed particularly well compared to other edge-based methods of object detection, and also in varying light conditions [18].
* An approach to real-time object detection and tracking in low-light conditions utilizing contrast analysis algorithms [19].
* A method for detecting and tracking vehicles in traffic surveillance using background subtraction actually allows for the background to be *reconstructed* on demand in any traffic condition [20], a multiple object detection and tracking method designed to overcome illumination variations, shadow interference and object occlusion problems using stereo-vision [21].
* A method of object detection in images can be trained to detect specific objects using a *multiple* example-based *trainable* detectors. In a demonstration, four example-based detectors where used to recognize human bodies – trained using head, legs, left arm and right arm, respectively [22].
* Use of contour cues by means of CENTRIST visual descriptor allows for an object detection framework capable of operating at an impressive 20 frames per second using only a single processing thread [23].

# Equipment

Currently, the following hardware equipment is available for the purposes of this project:

* **Nexys™3 Spartan-6 FPGA Board.** Implementation of central control logic, memory management, peripheral and IO logic (VGA graphics), and neural network / machine learning architecture.
* **VmodCAM Stereo Camera Module.** Stereo camera module to allow depth perception.
* 24x **Atmel ATmega32u2** based miniature development boards *for* *auxiliary demonstration circuitry.*

# References

|  |  |
| --- | --- |
| [1] | R. Brockers, M. Humenberger and S. Weiss, “Towards Autonomous Navigation of Miniature UAV,” in *Computer Vision and Pattern Recognition Workshops (CVPRW), CVPR2014*, Columbus, OH, USA, IEEE, 2014, pp. 631-637. |
| [2] | J. Vincent, “'Robot dragonfly' is the smallest and lightest self-navigating drone,” 2013. [Online]. Available: http://www.independent.co.uk/life-style/gadgets-and-tech/robot-dragonfly-is-the-smallest-and-lightest-selfnavigating-drone-9008723.html. [Accessed 15 October 2014]. |
| [3] | Delft University of Technology, “DelFly Explorer,” 2014. [Online]. Available: http://www.delfly.nl/explorer.html. [Accessed 15 October 2014]. |
| [4] | C. D. Wagter, S. Tijmons, B. Remes and G. d. Croon, “Autonomous Flight of a 20-gram Flapping Wing MAV with a 4-gram Onboard Stereo Vision System,” Delft University of Technology, Delft Netherlands, 2013. |
| [5] | K.-H. Chen and W.-H. Tsai, “Vision-based obstacle detection and avoidance for autonomous land vehicle navigation in outdoor roads,” *Automation in Construction,* vol. 10, no. 1, pp. 1-25, 2000. |
| [6] | IEEE Spectrum, “How Google's Self-Driving Car Works,” 2011. [Online]. Available: http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/how-google-self-driving-car-works. [Accessed 14 10 2014]. |
| [7] | Popular Mechanics, “How Stanley Works,” 9 January 2006. [Online]. Available: http://www.popularmechanics.com/technology/engineering/robots/2169012, http://www.popularmechanics.com/cm/popularmechanics/pdf/PMX0106DARPA2.pdf. [Accessed 2014 10 15]. |
| [8] | C. Urmson, “Tartan Racing: A Multi-Modal Approach to the DARPA Urban,” 13 April 2007. [Online]. Available: http://archive.darpa.mil/grandchallenge/TechPapers/Tartan\_Racing.pdf. [Accessed 15 October 2014]. |
| [9] | IEEE Spectrum, “Special Report: The Self-Driving Car,” 2014. [Online]. Available: http://spectrum.ieee.org/static/the-self-driving-car. [Accessed 15 October 2014]. |
| [10] | H. Durrant-Whyte, “A Critical Review of the State-of-the-Art in Autonomous Land Vehicle Systems and Technology,” Sandia National Laboratories, New Mexico USA, 2001. |
| [11] | Y. Watanabe, P. Fabiani and G. L. Besnerais, “Visual Air-to-Ground Target Tracking and self-localization of a UAV in a GPS denied Environment,” 26 March 2009. [Online]. Available: https://gt-uav.hds.utc.fr/lib/exe/fetch.php?id=fr%3Areunions&cache=cache&media=fr:reunion26mars2009:gdruav\_260309.pdf. [Accessed 15 October 2014]. |
| [12] | M. Angelopoulou and C.-S. Bouganis, “Feature Selection with Geometric Constraints for Vision-Based Unmanned Aerial Vehicle Navigation,” 1-13 August 2009. [Online]. Available: http://cas.ee.ic.ac.uk/people/ccb98/papers/MariaICIP11.pdf. [Accessed 15 October 2014]. |
| [13] | R. L. Gregory, Eyes and Brain, The Psychology of Seeing, Fifth Edition, Oxford UK: Oxford University Press, 1998. |
| [14] | V. Bruce, P. R. Green and M. A. Georgeson, Visual Perception; physiology and ecology., 4th ed., New York USA: Psychology Ptess, 2003. |
| [15] | C. Hemalatha, S. Muruganand and R. Maheswaran, “Survey on Real Time Object Detection, Tracking and Recognition in Image Processing,” *International Journal of Computer Applications,* vol. 91, no. 16, pp. 38-42, 2014. |
| [16] | C.-H. Hsia and J.-M. Guo, “Efficient modified directional lifting-based discrete wavelet transform for moving object detection,” *Signal Processing,* vol. 96, pp. 138-152, 2014. |
| [17] | V. Yanulevskaya, J. Uijlings and J.-M. Geusebroek, “Salient object detection: From pixels to segments,” *Image and Vision Computing,* vol. 31, no. 1, pp. 31-42, 2013. |
| [18] | P. K. Dhar, M. I. Khan, A. K. S. Gupta, D. Hasan and J.-M. Kim, “An Efficient Real Time Moving Object Detection Method for Video Surveillance System,” *International Journal of Signal Processing, Image Processing and Pattern Recognition,* vol. 5, no. 3, 2012. |
| [19] | K. Huang, L. Wang, T. Tan and S. Maybank, “A real-time object detecting and tracking system for outdoor night surveillance,” *Pattern Recognition,* vol. 41, no. 1, pp. 432-444, 2008. |
| [20] | N. A. Mandellos, I. Keramitsoglou and C. T. Kiranoudis, “A background subtraction algorithm for detecting and tracking vehicles,” *Expert Systems with Applications: An International Journal,* vol. 38, no. 3, pp. 1619-1631, 2011. |
| [21] | L. Cai, L. He, Y. Xu, Y. Zhao and X. Yang, “Multi-object detection and tracking by stereo vision,” *Pattern Recognition,* vol. 43, no. 12, pp. 4028-4041, 2010. |
| [22] | A. Mohan, C. Papageorgiou and T. Poggio, “Example-Based Object Detection in Images by Components,” *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 23, no. 4, pp. 349-361, 2001. |
| [23] | J. Wu, N. Liu, C. Geyer and J. M. Rehg, “C4 : A Real-Time Object Detection Framework,” *Image Processing, IEEE Transactions on,* vol. 22, no. 10, pp. 4096 - 4107, 2013. |

1. A 1.7GHz ARM Cortex-A9 Quad-Core processor based micro-sized development platform designed to execute Android and Ubuntu-Linux operating systems. [↑](#footnote-ref-1)