AUTONOMOUS BEHAVIOR MODELS FOR AR DRONE

Stefan Spasovski, Nevena Ackovska
Faculty of Computer Science and Engineering
"Ss. Cyril and Methodius" University, Skopje, Macedonia

ABSTRACT

The aim of this paper is to provide a model for enabling the Parrot AR Drone to infer intelligent behaviour. Following the bioinspred approaches implemented in robotics, the paper presents a comparison between the thinking centres in the living organisms and proposes a reasoning line for the drone. Sensory inputs have been explored, thus analysis for feasibility of adding them to the robot is done.

I. INTRODUCTION

There are many existing drones that are able to maintain stable flight. One of the most popular drone used for student research in airborne robotics is the SA. Parrots' AR Drone [1], depicted in figure 1.

These drones are able to execute a wide range of tasks. An example task is guarding a predefined yard, which is not an easy task, yet it is achievable, since one could give the robotpredefined navigation points and route it through those points. However, a drone companion that walks a child to the bus stop gives much more purpose to thetask.

Figure 1: Image of the Parrot AR. Drone.

A drone that can follow objects has many uses. For example, it can be used to inspect the behavior of birds that have nests in hardly reachable places, follow kids to the school ensuring their safety or engage a flight chase with other flying objectsto measure their maximal speed. On the other hand, new flight safe aircraft maneuvers can be inspired by observing the birds behavior. The research elaborated in this paper addresses the bird chase flight and proposes one very encouraging solution.

The main goal of the research presented in this paper is to make the AR Drone an autonomous robot. The goal is not only to make it act on its own, but also to be able to act intelligently given a specific purpose.

The paper is organized as follows. The next section elaborates on how the autonomous thinking is done in the living world. Afterwards, the most common sensors are explained and how they can contribute to the autonomy of the drone. Next, a discussion on what can be done in the robotics to bridge the limitations of the living world is presented. Finally, a discussion about the selection of the best model given the purpose for future practical usage is presented.

II. REASONING IN THE LIVING ORGANISMS

This section gives insight in the reasoning of the living world on planet Earth. All familiar living organisms have the reasoning center built in themselves, as is elaboratedin the following subsections. Even the simplest beings show intelligence that leads best to their survival. According to the classification recognized by the International Society of Protistologists living organisms are divided in three domains [2, 3]: Bacteria, Archaea and Eukaryote. They are explained in detail in the following subsections.

A. *Bacteria*

Bacteria are organisms that lack a central core, nucleus and other compartments. It was long believed that these creatures are not able to think, but rather only execute a predefined reaction when certain circumstances occur. However, in 2009, Pilpel et. al.[4] showed that bacteria can prepare themselves and learn from experiences. Their decision making is done by chemical chain reactions [5] within the whole body of the bacterium.

Most of the representatives of this kingdom are unicellular, but not all of them. Quite unique are the colonial formations that bacteria can form. For instance the Myxobacteriaareable to form multi cellular bodies in which each cell communicates with the others via acids and hormones. This is characteristic example for social behavior [6].

B. *Archaea*

The Archaea have many similarities with the representatives of the Bacteria Domain. One of them is that they both lack the central core and their thinking centers similar. Just like the bacteria their decision making is also done by chemical chain reactions within the whole body.

C. *Eukaryota*

1) *Unicellular*

Unicelular eukaryotes are organisms that have their decision making unit in the central core of the organism, the nucleus [8]. The kingdoms of Protozoa, Chromista and Fungi are representatives of the unicellular eukaryotes, besides them, multy-cellular eucariotes can be classified in the kingdoms of animals and plants.

2) *Animals*

The kingdom of animals is the kingdom in which humans belong. In eukaryote organisms, although the cells can act on their own, yet the decision making process is done in the neural system - specialized reasoning subsystem.

3) *Plants*

The last but not least are the vegetative organisms, the plants. They lack actuators, but they show active anticipation of events in the environment. Stanislaw et. al. [7] showed that they have memory and inter-cellular communication similar to humans' neurons.

III. Reasoning for AR Drone

As in the living organisms, in robotics an information needs to be propagated to a decision making center and the decision needs to be propagated to the correspondingobject.The sensory input needs to be sent to the reasoning unit from which commandsare sent to the actuators.

Since the aim is to create an object following authonomus robot, the content of the reasoning unit is of substantial importance.Object reconition algorithms must be present in the reasonong center of the drone in ordred to give an aim of the tasks that it has to fulfill. No matter how the view is provided, the algorithms used for interpretation of where the object being followed is, can generally be divided into two groups.If one needs to recognize a member from agroup of objects, there are lots of algorithms that can do classification based on an image input. Examples are Neural-Based target detectors, Edge detection etc. [21].On the other hand, if one wants to track certain object, there are relatively new algorithms that are doing excellent job in the on-line tracking task. For instance, Kalal's tracking learning detection, inspired by the natural way of thinking[26] or the Visual Decomposition by Lee et. al[25]. Benchmark tests for most of these algorithms and comparison for their results are already made [22, 23, 24].

IV. Sensors and Their characteristics

Reasoning in living organisms is done within the organisms, but are their sensors embeded in them as well? One could inconsiderately say that they are, but it depends on the interpretation of the sensor. Through two trivial examples these different points of view areshown.

Example 1: *A flame is seen, than almost every entity will run away*. Here, the visionary sensor is on the entity; it sees a fire and reacts repulsively on it.

Example 2: *Someone tells the firefighter that a flame was seen and the firefightergoes towards it*. He doesn't need to actually see the fire, somebody else can be the messanger.

There is an analogy in the action of the firefighter and the action of a remotely controlled AR Drone. Indeed, the sound sensor corresponds to WiFi sensor of the drone and the sound waves the firefighter heard correspond to the WiFi waves the drone received. As a consequence, both the Firefighter and the Drone acted based on the signals received from outer sources. Truly, the drone was steered using the controller's vision and in the same manner the firefighter undertook action for fire he did not see.

Because the aim of this research is enhancement of the AR Drone, analysis of what the benefits of having certain sensors are and what they are capable of, is presentedin the following sections. Not only sensors that are AR Drone embeded are covered, but also those that it can be equipped with.

A. *Vision Sensor*

Humans' most utilized sensor is the vision sensor. A plethora of hard work is invested for computers to understand and build cognitive image of the things that lay in front of the camera. Tasks that seem effortless for any human, like recognizing the difference between an apple and a pear, actually is one of the hardest thing to do for a computer. With immense computing power, object recognition and tracking can be made.

1) *Camera*

When it comes to input interpretation the camera sensors are one of the most dificultto interpret. They also are closest to the humans most reliable sensory input, the visual system. Unlike most of the other sensors, immense quantities of data needs to be processed. This requires not only fast analysis, because of the high input rate (the frames per second), but also lots of memory, because each image takes a lot of space. As a result, it is impossible for the majority of the micro-controllers to handle this task. Consequently, thick clients or other higher computational units have to be used for this task.

Because the camera is mounted on the drone, which is a moving object, other independent moving objects can be detected [9] and even 3D object can be mapped [12].

2) Stereo Vision Camera

Since stereo vision cameras are two regular cameras with aligned optical axes, thereby having focus at infinity, larger part of section 3.1 is also applicable here.

Because of the cameras'characteristic positioning, they can provide relative depth to the images being compared. This is a big perk in the process of perceiving the image input.This gives the ability of approximating the position of the objects that are seen. A system that maps a 3D out of stereo vision has appeared in the distant 1976. It provides three-dimensional measurements with resolution of ±l mm at distances of about 2 meters [10].Later, Little and Don [11], produced a robot that can navigate itself, without colliding usingnothing but a stereo vision camera.

B. Sound Sensor

Microphones are long existent sound sensors, yet analyzing their input is not a trivial task. After the input is analyzed, it is used for interpretation of commands or engaging conversations. In order to make a localization of where the sound is coming from, at least two sensors are needed. When using two sensors placed symmetricallyit can be determinedwhich sensor is closer to the source. The assumption made by determining which of the sensors senses louder voice is an rough, approximate assumption. Moreover, if the source is right before the robot and it only has two microphones on the sides, it cannot determine whether the source is in front, below, behind or above it.

These problems can be solved by sampling from different positions, and comparing the measurements in order to get a single cognitive image of the source. Even determination from more microphones, rather than fewer from different positions, is possible. By using an array of sensors, the plausibility of determining the source accurately is increasing[13, 14].

C. Proximity

1) Range Sensor, Sonar

For determining distance between object, usually infrared or ultrasound sonars are used. Range sensor is a device that measures the distance between the sensor and the first object on its axis of alignment. If put on a static body, the change in distance created by an object that passes in the space between can be detected. If put on a moving airborne object, it can measure altitude at which the object is and if precisely navigated, it can even make topological maps. Furthermore, if a sensor is put on a rotating body, it comprises a radar.

Most of the airborne drones have range sensors for altitude input.

2) 3D depth sensor

3D depth sensor is usually used in combination with the vision sensor, as in Kinect[19]. This way it provides easier way of interpreting the data given from the camera. Itgives the ability to create a cognitive map of the surrounding in real time. This is very useful, since onecan use it to roughly locate oneselfin the already passed places.

D. Positioning Sensors

There is a set of sensors that can tell where the robot is and in what position (state) it is. Autonomy only with these sensors cannot be provided, thus combining with other inputs is obligatory.

1) Gyroscope

The Gyroscope is a mandatory sensor for airborne robots that gives information of their current yaw, pitch and roll. It is of uttermost importance for airborne crafts because the hovering is done by implying a force in the opposite direction of the gravitational force, with the same amount. If the craft is rolled a bit, and the force is not increased, then the craft would lean in that direction thereby loosing altitude which leads to grounding of the craft.

This sensor can help significantly in adjusting the readings from the other sensors. For instance, if a front faced camera is placed on a drone, and it receives an image with an object on the top center of the frustum, it would imply that the object is horizontally centered, vertically above the drone. However, if a sensor sends information for the pitch of the drone, that is facing downward, even though the object is in the top center of the frustum, it actually is on the same altitude as the drone, as shown in Figure 2.

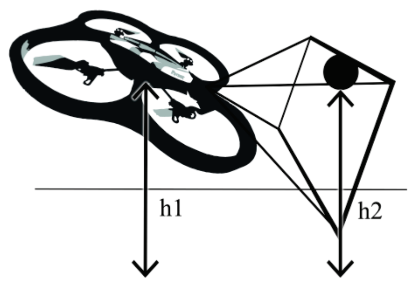

Figure 2: Object at the same height as the drone, at the top of its frustum, where h1=h2 are the hights of the drone and the object seen corenspodently

2) Global Positioning System

The Global Positioning System (GPS) sensor, gives the robot the ability of knowing its position in longitude and latitude. This sensory input does not give information about the surrounding, thus a detailed path direction and collision proof path without additional sensors is impossible. Its usefulness is significantwhen using it in wide open spaces or longer paths.

V. HARDWARE MODELS

Unlike the living organisms, a robot has the flexibility of having more physically separable parts, if the means of communication between these parts permit and provide a way for them to work as intended. This implies that we can dividethe robot onmore interconnected parts. In this case one or more of those partscan be included inthe decision making, whereas others can be used for information gathering or laboring. Having this concept in mind, it challenges the definition of the robot autonomy.

A. The Standard Model

In the standard model, the decision making center is part of the flying robot. It can be achieved in at least two ways. The processing unit that does the scheduling of the actuators will be responsible for the decision making. That unit in AR Drones is ARM9 468 MHz processor [18]. The alternative is equipping the drone with additional arbiter, which has to be lightweight and energy efficient, so it would not impact the actuators' performance, depicted on Figure 3.

More than few personal micro computers have the computational power and are lightweight as needed, including Raspberry Pi [15] with 700MHz processor, Cubieboard [16] 1.0Ghz, MK 802 [17] with 1.0GHz etc. all from the cortex A8 Arduino family.

A down side is that Raspberry Pi, the only unit that has GPIO (General Purpose Input Output) pins for custom needs, is the one with the slowest computing power.

Figure 3: An image depicting the AR Drone hardwired to a Raspberry Pi GPIO pins

B. Extended Model

The second concept presented in this paperwill be referred to as the Extended model. Extending the robot with a separate static part, called an arbiter and the existing mobile part, called drone, comprises the Extended Model. The arbiter is the part that does the calculations needed for determining drones actions. The drone is the movable part of the robot, the actuator. Advantage of the extended model is that it provides biger computational force, which allows more complex algorithms to be executed.

1) Personal Computer as Arbiter

The first suggested model is to put a personal computer to receive all input from the drone, do all the computational and decision making processes and finally send commands to the drone, depicted on Figure 4.

Figure 4: AR Drone in wireless communication with a personal computer

2) Server Farm as Arbiter

Since we can use a personal computer for the calculations, no limitations on using a single computer are implied. If the demand of calculations needed to be performed is high, and if paralleling them is feasible, then the server farm as arbiter is reasonable choice. It is shown on Figure 5.

At least one computer stillneeds to communicate with the flying part of the robot. That computer will schedule the work for the other units of the server farm and a whole farm will compute the drones' next move. This provides a better computational force than using a single computer.

3) Personal Computer as Gateway to Grid or Cloud as Arbiter

If the criteria for the server farm are met and if a reasonable scheduler can be made, then grids are also viable option for this kind of computing.

Figure 5: AR Drone in wireless communication with the gateway PC and it with a server farm

VI. AR Drone Possibilities

The Parrot AR Drone has front facing VGA camera at 15 frames per second (FPS), vertical QCIF camera at 60 FPS, vertical ultrasound sensor, both accelerometer and gyroscope, WiFi 802.11b/g, additional USB socket, and all these run on ARM 9 RISC 32 bit processor at 468 MHz, and 128 MB DDR RAM[18].

Usinggyroscopes, AR Drone has the inputs needed in order to maintain a stable flight. The floor facing range sensor provides enough input for the AR Drone to stay at some particular altitude. Sufficient information for whether the drone is in flight or on ground is also provided.

However, none of these inputs give enough information for the environment around the robot. That is why utilizing cameras is mandatory for more intelligent behaviour.

With the front facing camera, the robot is able to recognize and/or track objects. However,since that is immensely demanding task, the 468 MHz processor would not be able to do the tracking efficiently.

Combiningthe floor facing camera and the range sensor, if the floor facing camera can recognize patterns, it is possible to calculate the distance the drone is passing. If the velocity is known, and the rate at which the object seen with the front facing camera is getting bigger, than a calculation of the distance between the object and the drone is also possible. This is a great perk.

From the sensors mentioned in section 3, the only input that is not fully provided with the stock sensors of the AR Drone is 3D depth. Currently there are two leading products that provide this, the Microsoft's Kinect [19] and OpenNI's RD1[20]. The problem with bothis that they are too big and weight more than 3 times than the drone itself. Thus, the drone is unable to carry them.

The on-line learning algorithms seem more suitable for the predatory behavior we are up to achieve, then the classification algorithms. The computing power required by the vast majority of them is suitable for running on a PC.Those algorithms are designed to be run on a PC.For on-line object tracking needed, the processor embedded in the AR Drone is not nearly enough. Because of the inability to execute the on-line tracking algorithms concurrently, the model in section 4.2 is suggested.

Adding a microcomputer on the craft itself for now is not an option. As mentioned before, Raspberry Pi, the only unit that has GPIO pins, is the one with the slowest computing power.

The communication between the drone and the personal computer can be done using the already equipped WiFi 802.11b/g component which gives the maximum outdoor range of 140 meters and 54Mbit/s.

VII. Conclusion and Future Work

After we covered the reasoning in the living organisms, and showed that the robots not necessarily have to mimic the nature, we suggested several models that challenge the definition of autonomy and recommended a use of personal computers as arbiters for calculations. In this paper the types of sensors that can contribute to the autonomus flight are covered. The research showed that the AR Drone was well designed and equipped with all the required sensors.

In future, benchmarks on the on-line tracking algorithms are to be created which can give better insight which one is most suitable for object detection from a moving camera. That will enable implementation of the model and benchmark it in its final product environment.

VIII. Acknowledgements

This work was partially financed by the Faculty of Computer Science and Engineering at the "Ss.Cyril and Methodius University.

References

[1] Parrot, S. A. "AR. Drone." *website, http://ardrone. parrot. com* (2011).

[2] Alastair G.B. Simpson and , Andrew J. Roger *The real kingdoms of eukaryotes*, Current Biology 14 (17): R6936, doi:10.1016/j.cub.2004.08.038, (2004)

[3] Adl, Sina M.; et al. *The New Higher Level Classification of Eukaryotes with Emphasis on the Taxonomy of Protists*, Journal of Eukaryotic Microbiology 52 (5): 399, doi:10.1111/j.1550-7408.2005.00053.x (2005)

[4] Amir Mitchell, Gal H. Romano, Bella Groisman, Avihu Yona, Erez Dekel, Martin Kupiec, Orna Dahan and Yitzhak Pilpel *Adaptive prediction of environmental changes by microorganisms*, Nature, June 2009

[5] T. Kalmar, C. Lim, P. Hayward, S. Munoz-Descalzo, J. Nichols, J. Garcia-Ojalvo, and A. Martinez Arias *Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells.* PLoS Biol.7, 2009

[6] Martin Dworkin Recent Advances in the Social and Developmental Biology of the Myxobacteria Microbiological Reviews, Mar. 1996, p. 70–102

[7] Magdalena Szechyska-Hebdaa, Jerzy Krukc, Magdalena Greckaa, Barbara Karpiskaa and Stanisaw Karpiskia *Evidence for Light Wavelength-Specific Photoelectrophysiological Signaling and Memory of Excess Light Episodes in Arabidopsis* The Plant Cell July vol. 22 no. 7, 2010 2201–2218

[8] Gabor Balazsi, Alexander van Oudenaarden, and James J. Collins *Cellular Decision Making and Biological Noise: From Microbes to Mammals* Cell144, March 18, 2011 910–925

[9] P.H.S. Torr and D. W. Murray *Statistical Detection of Independent Movement from a Moving Camera* Image and Vision Computing 11, 1993 180–187

[10] Y. Vakimovsky and R. Cunningham *A System for Extracting 3-Dimensional Measurements From a Stereo Pair of TV Cameras* JPL Technical Memorandum, May 15, 1976 33–769

[11] Murray, Don and James J. Little. *Using real-time stereo vision for mobile robot navigation.* Autonomous Robots 8.2, 2000 161–171

[12] Wolfgang Niem, Jochen Wingbermhle *Automatic Reconstruction of 3D Objects Using a Mobile Monoscopic Camera* Image and Vision Computing Vol. 17, Issue 2, February 1999 125–134

[13] Jean-Marc Valin, Francois Michaud, Jean Rouat, Dominic Letourneau *Robust Sound Source Localization Using a Microphone Array on a Mobile Robot* Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on Oct. 2003 Vol 2. 1228–1233

[14] Michael S. Brandstein and Harvey F. Silverman A Practical Methodology for Speech Source Localization With Microphone Arrays November 13,1996

[15] Raspberry Pi Specification website, http://www. raspberrypi. org/ faqs

[16] Cubieboard Schematic website, http://dl. cubieboard. org/ hardware/ cubieboard_schematic_2012-08-08.pdf

[17] MK802 Specification website, http://www. miniand. com/ products/ MK802%20Android%20Mini%20PC #specifications

[18] AR Drone Specification website, http://parrot-drone. co. uk/ parrot-ar-drone-technical. php (2011)

[19] Kinect Specification website, http://msdn. microsoft. com/ en-us/ library/ jj131033.aspx (2013)

[20] RD1 Specification website, http://www. openni. org/ rd1-08-specifications/ (2012)

[21] Javidi, Bahram. Image recognition and classification: algorithms, systems, and applications. Vol. 78. CRC Press, 2002

[22] Wang, Qing, et al. *An experimental comparison of online object-tracking algorithms.* SPIE: Image and Signal Processing (2011): 81381A-81.

[23] Spampinato, C., S. Palazzo, and D. Giordano. *Evaluation of tracking algorithm performance without ground-truth data.* Image Processing (ICIP), 2012 19th IEEE International Conference on. IEEE, 2012.

[24] Chau, Duc Phu, Franois Bremond, and Monique Thonnat. *Online evaluation of tracking algorithm performance.* Crime Detection and Prevention (ICDP 2009), 3rd International Conference on. IET, 2009.

[25] Kwon, Junseok, and Kyoung Mu Lee. *Visual tracking decomposition* Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. IEEE, 2010.

[26] Kalal, Zdenek, Krystian Mikolajczyk, and Jiri Matas. *Tracking-learning-detection* Pattern Analysis and Machine Intelligence, IEEE Transactions on 34.7 (2012): 1409-1422.