

A five-camera vision system for UAV visual attitude calculation and collision warning

Akos Zarandy^{**†}, Zoltan Nagy[†], Balint Vanek^{*}, Tamas Zsedrovits^{*}, Andras Kiss^{†*},
Mate Nemeth^{*}

email: zarandy.akos@sztaki.mta.hu

^{*}Institute for Computer Science and Control, 13-17 Kende Street, Budapest, H-1111, Hungary

[†]Pazmany Peter Catholic University, 50 Prater Street,

Abstract. A five-camera vision system was developed for UAV visual attitude calculation and collision warning. The vision system acquires images by using five miniature cameras, stores, and evaluates the visual data real-time with a multi-core processor system implemented in FPGA. The system was designed to be able to operate on a medium sized UAV platform, which raised numerous strict physical constraints.

Keywords: vision system, UAV, low power, multi-camera, FPGA, multi-core processing, image processing, visual navigation, collision warning.

1 Introduction

Unmanned Aerial Vehicle (UAV) technology reached an advanced level, which enables them technically to fly autonomously a predefined paths and complete different missions. However, legally they are not allowed to fly fully autonomously, since flight authorities identified various safety shortcomings [1]. One of the problems is that they are not robust enough due to the lack of on-board sensor and actuator redundancies. Another missing capability is the collision avoidance [2,3,4,5], because the GPS based control and navigation system makes the UAV flying practically blindly, hence it can collide with any other aircraft, or with any stationary object, which is not correctly on the map (new building, antenna tower, pillar of a bridge, crane, ski lift, etc.). The introduced vision system was designed to help in these problems, by making the Inertial Navigation System (INS) more robust by adding an extra angular velocity sensor source, and by identifying collision threats in time.

Naturally the vision system should fulfill numerous tough specification criteria. Its resolution and field of view (FOV) should be high enough to identify intruder aircraft from large distance; it should be able to perform real-time processing; its size, weight, and power consumption parameters should satisfy on-board UAV operation requirements; and finally, it should be affordable. From functionality point of view, it is expected to calculate the attitude of the aircraft by calculating the differential orientation changes between consequent frames (yaw, pitch, roll angles) and detect intruder air-

crafts, which are on a collision course; store all the acquired images in full resolution for archiving and for off-line testing purposes.

The paper is organized on a way that first we describe the state-of-the-art and the related work in this field (Section 2). Then, system specification is given in Section 3. After that the system is described in Section 4. In Section 5 the multi-core processor array implementation is briefly shown. Finally, measurement results are given in Section 6.

2 Related work

Naturally, avoiding mid-air collisions is not a new problem. Traditionally, there are two different approaches to address the airborne collision avoidance. The first assumes cooperation among the aircrafts. In this case each aircraft transmits its position, velocity, and planned route, and based on a predefined protocol the aircrafts avoids approaching each-other. The nowadays used version of the system is called TCAS (traffic collision avoidance system) [6]. A new version of it, called ADS-B (automatic dependent surveillance-broadcast) is currently introduced, and will be mandatory in most larger aircrafts from 2020 [7]. Though cooperative approaches are relatively simple, and does not require sensing of remote aircrafts, however US and European agencies require having non-cooperating solution on board as well.

Modern big airliners utilize sophisticated radar and computer systems, which identify the position and the velocity of intruder aircrafts, warns the pilot if they are on a collision course, and even make an avoidance maneuver automatically if pilot does not react. However, this solution cannot be applied to small aircrafts due economic and weight considerations.

For large UAVs sensor fusion is a commonly used approach, to make the collision avoidance system operational in all flight conditions. The system, described in [8], is based on a pulsed Ka-band radar, two kinds of visible cameras, two IR cameras, and two PCs. For small UAVs vision only systems are currently developed in different places. One is described in [9], in which one piece of 1024x768 resolution camera and a PC with GPU is used to identify the intruder. Compared to this system, our system has significantly higher resolution, smaller weight, size and power consumption, thanks to its compact design and FPG image processor engine.

3 System specification

The vision system has two important roles, namely the attitude data calculation and the collision warning. From these, the more challenging task from image acquisition point of view is the collision warning, because detection of potentially dangerous intruder aircrafts in time requires the permanent monitoring of the field of view of $220^{\circ}\times 70^{\circ}$ in front of our UAV [10] with high resolution as we will see below. Fig. 1 illustrates the requirements of the safe avoidance. According to the flight safety requirements [10], there should be a certain separation volume around each aircraft, in

which nothing else can be. The size of the separation volume (separation minima) differs from airplane to airplane and situation to situation.

To be able to avoid the separation minima, the intruder should be detected from a distance, which is not smaller than the traffic avoidance threshold. If the intruder is not detected before crossing the traffic avoidance threshold, but detected before the collision avoidance threshold, the collision can be still avoided. For human pilots 12.5 second before collision is the last time instant, when collision can be avoided with high probability [Hiba! A hivatkozási forrás nem található.]. Naturally, to avoid scaring the pilots and the passengers of the other aircraft, and to increase the safety level, earlier initialization of the avoidance maneuver is required, which certainly assumes earlier detection. Since the tracks of the small and medium size UAVs do not interfere with streamliners, or high speed jets, we have to be prepared for other UAVs and Cessna 172 type manned crafts. This means that the maximal joint approaching speed is 100 m/s, therefore we need to detect them from 2000 meters (20 seconds before collision), to be able to safely avoid them. In these cases the separation minima is 2000ft (~660m) collision volume is 500ft (~160m).

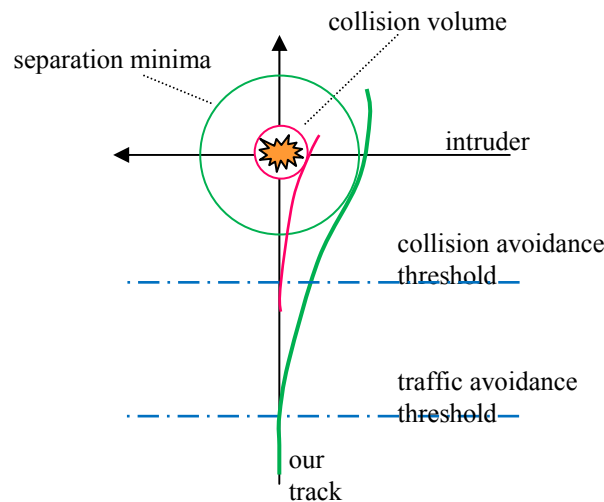


Fig. 1. Traffic (green) and collision (magenta) avoidance courses

To be able to perform robust visual detection of an aircraft, it should be at least 3 pixels large in the captured image. For a Cessna 172 class aircraft, with 10 meters wingspan, 0.1 degree/pixel resolution is minimum required. This means an overall minimum 2200x700 pixel resolution.

Other important system requirement is the speed. The control system is expecting 20 navigation parameter update in a second, therefore, the frame rate should be minimum 20FPS. Naturally, the image processing part of the vision system should be able to perform the complete processing on this speed also.

For real-time attitude calculation the same resolution and speed, but smaller FOV is satisfactory. Therefore the system with the above specification can calculate the angular changes of the aircraft orientation.

The system should be able to fit and operate on a UAV platform, which introduces strong limitations to its size, weight and power parameters. Our target was to fit the device to medium sized UAVs with 3m wingspan, which limits the weight to maximum 0,5kg including batteries. Another important requirement is that the vision and storage system should be resonance tolerant.

4 System description

In this section first, the selection of the main components is described, then the system architecture, the interconnections of the components, the power distribution, and the system integration are shown.

4.1 Camera selection

The key component of a special purpose vision system is the camera. During the design phase, one has to consider different types of camera. One would think that the most straightforward solution would be to use one piece of high resolution (like 2500x1000) camera with a low distortion ultra-wide angle optics. However, the problem with this setup is that the size and the weight of the camera and especially the ultra-wide view angle optics is way beyond the acceptable limits.

Therefore we have decided to apply multiple small cameras. We have studied three different classes of cameras:

1. Micro cameras with integrated lenses (mobile phone class);
2. Miniature cameras with S-mount (M12) lenses;
3. Small industrial cameras with C-mount or CS-mount lenses.

Micro cameras. The advantage of the micro cameras is that they are cheap, have sufficient pixel resolution up to 8 megapixels (e.g. Framos ules: <http://www.framos-imaging.com/sensormodules.html?&L=1>), and they are ultra-compact and low power. However, the price of the miniaturization is poor optical quality, and rolling shutter sensors, which makes them unusable for UAV navigation application, where it is critical to capture the entire image at the same time.

Miniature cameras with S-mount (M12) lenses. Miniature cameras are good candidates for low volume, low weight applications. In this camera class, the optics is already replaceable, and one can find high resolution (megapixel) lightweight optics (<http://www.sedeco.nl/sedeco/index.php/lenses/smout>) for them with different view angles. The resolution of the rolling shutter ones are going up or beyond 5 megapixels (http://www.mobisensesystems.com/pages_en/aptina_modules.html), while the global shutter ones have lower resolutions, like WVGA or the soon available 1.2 megapixels one (http://www.mobisensesystems.com/pages_en/camera_modules.html). Here the typical power consumption is less than 200mW, and the weight is around 10g including optics.

The output of these cameras is either parallel raw data or USB.

Small industrial cameras with C-mount or CS-mount lenses. There are a very large number of cameras in this class. One can find them in different resolution (from VGA up to 8 megapixels), size (from 3x3x3cm), weight (from 40g), and both rolling and global shutter types (e.g. <http://www.ptgrey.com/products/index.asp>). However, here the weight of a precision lens is significant as well (60-200g) (<http://www.edmundoptics.com/imaging/imaging-lenses/>), hence the overall weight is above 100g. This weight is much larger than the cameras in the second category, but as an exchange, the precision of the lens and the optical alignment is much better. The power consumption of these cameras is watts rather than hundred milliwatts, mostly because they use power hungry high-speed serial output data channels.

The outputs of these cameras are typically Gige, USB 2, USB 3, Camera Link, or Fire-wire.

Selection. Since we need global shutter sensor, we can select cameras from the second or the third category only. The second category makes possible to build vision system for small and medium sized UAVs, where weight of the vision system should not exceed 500g.

Our other selection criterion was the data interface. For us the parallel digital raw data IO was the optimal, since for short distance it consumes much less power than high speed serial interfaces (Gige, USB, Fire-wire, Camera Link) which were designed for long distance communications.

We have selected 5 pieces WVGA (752x480) cameras (MBSV034M-FFC from Mobisens) to cover the required resolution with necessary overlap. For this 1/3 inch camera module, we have selected 3.66 mm focal length High Resolution Infinite Conjugate μ -Video™ Imaging Lenses from Edmund.

4.2 Data storage unit

In an airborne application the data storage can be implemented in some kinds of flash memory device. The options are memory card, USB stick, or solid state disk. The data-rate to save in this device is $5 \times 752 \times 480 \times 20 = 36 \text{ Mbyte/sec}$ (2.1 Gbyte/min) raw data assuming 5 cameras, WVGA image size, and 20fps. Though data compression is a widely used option for image storage, in our application, where very small remote objects are needed to be identified, the artifacts introduced by the compression is intolerable.

Therefore we needed a device which can cope with 36Mbyte/sec data flow. This is way beyond the write speed of an SD card (2-10Mbyte/sec) or a USB stick (4-25Mbyte/sec). Moreover, we need to store up to 20 minutes flight data during a test data acquisition flight, hence 45Gbyte data storage space is needed. This fits already to a small sized SSD (64 Gbyte). The system enables easy up scaling, since SSDs go up to 600Gbyte.

4.3 Processor selection

Nowadays the high-performance image processing platforms are based either on GPUs, or on DSPs, or on FPGAs. In case of strict power, weight, and size budget, the power hungry GPU platforms with their heavy cooling radiators cannot be an option,

even though there have been some platforms already developed for military UAVs (<http://defense.ge-ip.com/products/gpgpu/c497>).

By comparing DSPs and FPGAs, the DSPs are more flexible, and their programming time is much shorter, however, the processing performance of the FPGAs is much higher. Since a five-camera data acquisition, processing and storing system requires high computational speed and flexible data communication channels, the FPGA solution was the better choice. We have selected a small form factor FPGA board with a Spartan 6 XC6SLX45T FPGA (EXPARTAN-6T, <http://www.tokudenkairo.co.jp/exp6t/>), which had enough user IO ports to collect the data from the five cameras, and had SATA interface to save the acquired image flows.

4.4 System architecture and interconnections

The block diagram of the system is shown in **Fig. 2**. It contains off-the-shelf components (the cameras, the FPGA board, and the SSD), and a custom designed interface card. The cameras and the interface card are connected with 30 wire Flexible Flat Cable (FFC). The interface card is connected to the FPGA card with a board to board 80 pin connector. The SSD is connected to the FPGA board with SATA cable.

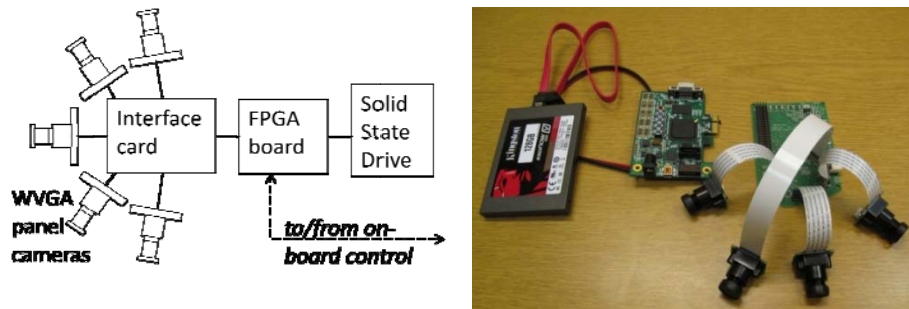


Fig. 2. The block diagram of the vision system and the photo of the connected components

4.5 Operation

The cameras are initialized through separate I²C bus. They are running synchronized. Their integration times are the same, and they receive the same system clock, and exposure trigger signal. Therefore the individual frames are captured at the same time.

The vision system is connected to the on-board control computer of the UAV through two I²C buses. Through these connections, the vision system receives the attitude estimation calculated by the navigation computer, and based on it, the vision system also calculates its yaw, pitch, and roll figures, what is sent back to the navigation system. In case of intruder aircraft detection, its position and size is sent to the navigation and control computer to initialize an avoidance maneuver.

4.6 Power supply

The total power consumption of vision system is about 7.5 W. Most of it is consumed by the SSD, which is 4.8 W alone (<http://www.legitreviews.com/article/1980/1/>). The energy source of the entire system is a 1200mAh 7.4V Lithium Polymer battery (2S1P). It can provide continuously 30 amps (25C), which ensures that the battery will not be overloaded. It enables close to 1 hour continuous operation.

4.7 System integration

Physical system integration is always a key point of a complex embedded system. It is especially true for an airborne vision system with multiple cameras, where the relative camera orientations are critical. Therefore a horseshoe like solid aluminum frame was constructed for holding the cameras and cancelling any cross vibrations (**Fig. 3**). The interface and the FPGA cards were put in and behind the horseshoe between two aluminum planes. The vision system is mounted to the nose of a two engine aircraft on a way that the axis of the front camera is aligned with the horizontal axis of the aircraft (**Fig. 4**).

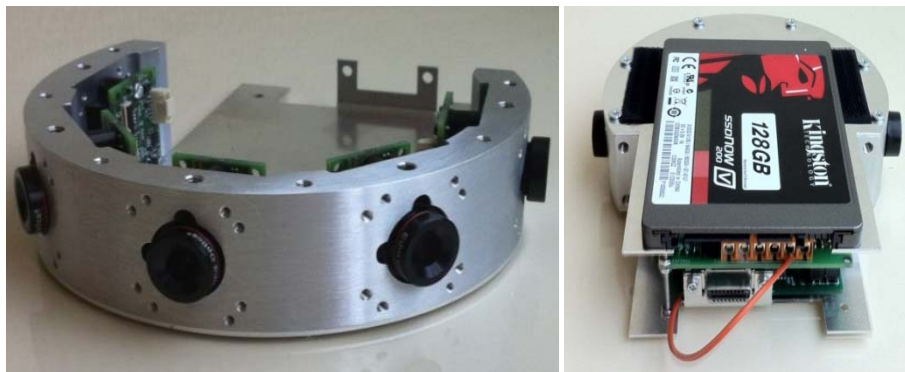


Fig. 3. Camera holder aluminum frame with the cameras (left), and the entire vision system without the power units (right)

5 Multi-core processor architecture in the FPGA

The image processing system should execute the following parallel tasks:

- calculating the attitude changes of the aircraft;
- identifying intrude aircrafts;
- communicating with the control and navigation processor of the UAV;
- and transferring the raw image data towards the SSD.

All of these functionalities are handled by a custom designed multi-core processor architecture implemented in a Spartan 6 LX45T FPGA. The basic concept of the processor design was to mimic the human foveal vision on a way that a pre-processor

examines the entire frame and identifies those locations, which needs more attention. Then, the focus of the processing is shifted to these locations one after the other, similarly as our fovea focuses to different important details of a scene.



Fig. 4. The vision system mounted on the nose of the aircraft (left) and the enlarged aircraft nose (right)

The architecture of the multi-core foveal processor is shown in **Fig. 5**. As it is shown on the figure, the five parallel 8 bit data flows arriving synchronously from the cameras are combined to one, time multiplexed 8 bit data flow. The combined data flow goes to the SATA core and to the full-frame streaming pre-processor as well.

The pre-processor has a streaming architecture, means that it cannot randomly access the entire frame, but it receives it row-wise sequentially as the image is read out from the sensor. To be able to calculate neighborhood operators, it collects a few lines of the frame and processes those lines together. As the data stream flows through the processor, it finds those high contrast corner-like locations where displacement vector will be calculated when the next frame arrives. It also identifies those objects which might turn out during the post processing phase to be an intruder aircraft. The pre-processor sends the coordinates of the identified locations to the internal micro-processor (MicroBlaze), and saves the raw frame and the some processed data to the external memory.

The MicroBlaze is a general purpose 32 bit soft-core processor implemented in the Xilinx FPGAs. It has relatively low computational power (~200MHz clock speed), which means that it cannot perform image processing tasks. It can be used to be the control processor of the system, and also to perform some decision making and communication.

The MicroBlaze then goes through the identified suspicious locations and performs foveal (region of interest, ROI) processing one after the other, by instructing the binary and the grayscale ROI processors to cut out the required windows, copy them into the internal block memories of the FPGA, and execute the program sequences. The detailed description of the pre-processor and the foveal processor can be found in [11], while the algorithm description in [12].

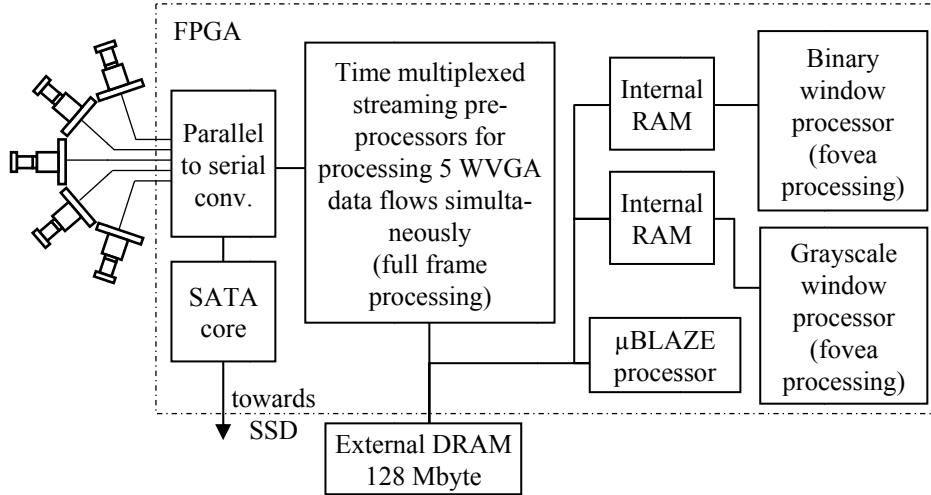


Fig. 5. The block diagram of the image processing architecture

6 Measurement results

We have executed multiple successful flights for areal image acquisition. The captured image sequences are synchronized with the data recordings of the inertial measurement unit (IMU), hence the aircraft position and attitude at each frame capturing time intend is known.

On these image sequences, the displacement vectors were calculated in 8 different characteristic locations, and based on them, the attitude was calculated using the five-point method [13]. As it can be seen in **Fig. 6** the calculated yaw data of the UAV and the measured results are closely correlated. (Other two angles and the detailed description of the algorithm are shown in [13]. **Fig. 6** also shows an identified intruder aircraft. The algorithm performs well. It can identify all the intruder aircrafts against clear-sky or cloudy background in our image sequences captured from UAV platform or from the ground. On the other hand, we are at the beginning of the algorithm evaluation both in the attitude calculation and the intruder aircraft identification.

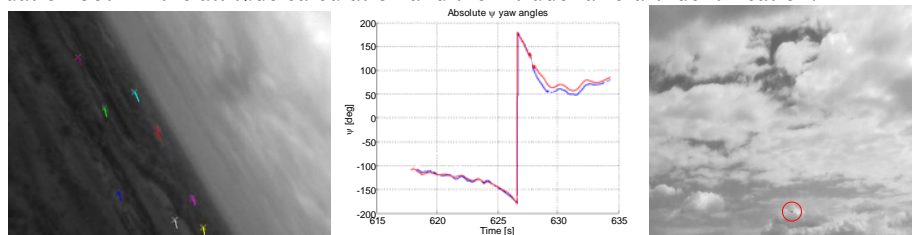


Fig. 6. The local displacement vectors (left), the yaw angle calculation with different methods (middle) (blue: five-point method [13], red: IMU data), and a detected intruder (right)

7 Conclusion

A five-camera vision system was introduced. The system was designed to be able to operate on UAV platforms. Its role is real-time attitude (orientation angle) calculation, vision based collision warning, and visual flight data acquisition. The system has been built and partially verified on a UAV platform.

8 Acknowledgement

The ONR Grants (N62909-11-1-7039, N62909-10-1-7081) is greatly acknowledged., The authors express their thanks to grants TÁMOP- 4.2.1.B-11/2/KRM-2011-0002 and TÁMOP-4.2.2/B-10/1-2010-0014.

References

1. W. Felder, Unmanned System Integration into the National Airspace System. Philadelphia, PA, USA: Keynote Presented at ICUAS 2012, June 2012.
2. Debadepta Dey, Christopher Geyer, Sanjiv Singh and Matt Digioia "Passive, Long-Range Detection of Aircraft: Towards a Field Deployable Sense and Avoid System" Field and Service Robotics, Springer Tracts in Advanced Robotics, 2010, Volume 62/2010, 113-123, DOI: 10.1007/978-3-642-13408-1_11
3. Federal Aviation Administration, Fact Sheet - Unmanned Aircraft Systems (UAS). 2010
4. Department of Defense, "Unmanned Aircraft System Airspace Integration Plan," Tech. Rep. March, Department of Defense, 2011
5. Federal Aviation Administration, Integration of Unmanned Aircraft Systems into the National Airspace System Concept of Operations. 2012.
6. C. Livadas, J. Lygeros, A. Lynch, Nancy, "High-level modeling and analysis of the traffic alert and collision avoidance system (TCAS)", Proceedings of the IEEE, Volume: 88 , Issue: 7 Page(s): 926- 948, 2000
7. Federal Aviation Administration, Fact Sheet - Automatic Dependent Surveillance-Broadcast (ADB-S). 2010.
8. Giancarmine Fasano, Domenico Accardo, Antonio Moccia, Ciro Carbone, Umberto Ciniglio, Federico Corrado, and Salvatore Luongo. "Multi-Sensor-Based Fully Autonomous Non-Cooperative Collision Avoidance System for Unmanned Air Vehicles", Journal of Aerospace Computing, Information, and Communication, Vol. 5, No. 10 (2008), pp. 338-360.
9. L.Mejias, S. McNamara, J. Lai, J. Ford, "Vision-based detection and tracking of aerial targets for UAV collision avoidance", International Conference on Intelligent Robots and Systems (IROS), pp.: 87 - 92, 2010
10. International Civil Aviation Organization, "Air Traffic Management", ICAO Doc 4444, fifteenth edition — 2007
11. Z. Nagy, A. Kiss, Á. Zarándy, B. Vanek, T. Péni, J. Bokor, T. Roska, „Volume and power optimized high-performance system for UAV collision avoidance” ISCAS-2012, Seoul, Korea
12. Ákos Zarándy, Tamás Zsedrovits, Zoltán Nagy, András Kiss, Tamás Roska, „On-board see-and-avoid system”, Conference of the Hungarian Association for Image Processing and Pattern Recognition (Képf 2013) pp 604-617. Bakonybel, 2013
13. Tamas Zsedrovits, Akos Zarandy, Balint Vanek, Tamas Peni, Jozsef Bokor, Tamas Roska, „Estimation of Relative Direction Angle of Distant, Approaching Airplane in Sense-and-avoid” Journal of Intelligent and Robotic Systems, Volume 69, Issue 1-4 , pp 407-415, 2013