Visual Sense-and-Avoid for Small Size Unmanned Aerial Vehicles

Theses of the Ph.D. dissertation

Tamás Zsedrovits

Scientific Advisor:
Ákos Zarándy, D.Sc.

Supervisor:
✝ Tamás Roska, D.Sc.

Pázmány Péter Catholic University
Faculty of Information Technology and Bionics
Multidisciplinary Technical Sciences Doctoral School

Budapest, 2014

1 Introduction and aims

In the last decade Unmanned Aircraft Systems (UAS) – beforehand Unmanned Aerial Vehicles (UAVs) -technology has evolved considerably. Besides the military applications now there is a great opportunity to use UAS in commercial applications as well. More and more companies start to develop applications and services based on the UAS platform. According to many aviation experts pilotless aircrafts are going to revolutionize air transport in the near future. As written in the cover story of December 2011 issue of IEEE Spectrum Magazine: “A pilotless airliner is going to come; it's just a question of when,” said James Albaugh, the president and CEO of Boeing Commercial Airlines.

Nevertheless, in order to use UAS in these fields their reliability needs to be increased as well as their capabilities need to be extended further, their ease of use needs to be improved and their cost have to be decreased. At the same time the regulatory challenge of integrating UAS into national and international air space has to be solved. One of the most important problems which has to be solved is the collision avoidance or sense-and-avoid capability. These functions have to be run on-board even if the connection between the aircraft and the control station is lost or some of the on-board sensors fail.

Provided that the size and the energy consumption of the UAV are limited, a camera based avoidance system would provide cost and weight advantages against the systems currently in use on bigger aircrafts, like cooperative systems for example TCAS. Furthermore near airfields, because of a great density of aircrafts and the limited frequency resources of air traffic controllers the camera-based approach seems to be more feasible then others.

Today’s kilo-processor chips allow us to implement complex algorithms in real time with low power consumption, which allows us to run the image processing. My work was done in a research group funded by the Office of Naval Research (ONR) and ONR Global within the framework of the grant N62909-10-1-7081. The main goal of this research is to develop an autonomous mid-size fixed-wing safety critical UAV for civil applications.

The development of the actual hardware elements went parallel with the development of the algorithmic framework which included the research on vision based SAA for UAVs. At the time of this research there were no complete, visual SAA system for mid-size and small UAS and the properties of this kind of systems had not been described yet.

The aim of this work is to introduce and analyse visual methods for the UAS SAA problem. In particular, what kind of information can be extracted from the image flow if the intruder airplane is close enough and what can we expect from various visual space reconstruction algorithms in the case of own aircraft’s attitude estimation.

2 Methods

In my work I used methods and theory from the field of multiple view geometry and image processing. The new algorithm for the relative direction angle estimation uses the geometry of the actual scene and the projective camera model. In the experiments for the analysis of this algorithm I used a multi-fovea image processing with cellular operators for the image segmentation.

The main experimental framework is a hardware in the loop simulator for tests on the ground and the developed UAV for the airborne tests. The flight control is running on hardware in the loop system, and the aircrafts are simulated in MATLAB/Simulink. For the own aircraft a high fidelity mathematical model has been used. The intruder is modelled as a simple double integrator. For the own aircraft a trajectory tracking controller has been designed, which runs on an MPC555 embedded microprocessor. The flight simulator PC communicates with the image rendering and processing computer via Ethernet and the rendering is done by the FlightGear simulator program. The image processing is included into a modified FlightGear, which sends the results of the image processing (the subtented angle and the size) to an FPGA via USB. The FPGA runs an EKF in order to predict the relative 3D position of the intruder. In the current system a Spartan 3 FPGA runs the motion prediction task. The results are sent back to the control part where the risk estimation and the trajectory generation take place.

The relative direction angle estimation algorithm is tested on images generated by the HIL simulator as well as on real videos with a fixed model airplane. The aircraft’s wingtip points were selected both by a human expert and by an algorithm for the tests. The error is calculated with a simple subtraction in order to show the direction of the error as well.

For the pose estimation four algorithms were selected from the literature: A (i) homography based solution as a basic algorithm with small computational need but with less accuracy. The (ii) eight point algorithm, as standard algorithm in epipolar geometry. The (iii) five point algorithm, as one of the state of the art algorithms with higher computational need, but with promising stability over the various scenes. Finally, (iv) MLESAC, as an iterative, stochastic solution.

These algorithms were tested on simulated feature point coordinates generated by EGT toolbox for MATLAB. Two real flight paths were selected: (i) a sinusoidal and (ii) a zigzag path, both with at least 7000 points, which corresponds to a path of 70s flight sampled with 100 Hz. The feature points were placed randomly near to the ground and were projected to the image plane using scaled version of the camera calibration matrix of the micro camera used on board. To characterize the performance of each algorithm the absolute error of the three Euler angles are used. Also the mean, the median and the standard deviation of the error is used.

3 New scientific results

1 Thesis: *Development of a relative direction angle estimation algorithm for visual sense and avoid system for autonomous unmanned aerial systems:*

I have introduced a new algorithm for relative direction angle estimation and shown the reachable accuracy in various situations. The algorithm is based on the assumption that the two approaching aircrafts are on a straight path and we have calibrated camera. I have also shown a simple algorithm for the extraction of the aircraft's wingtip points on the images. The accuracy of the relative direction angle is measured in pure simulation, on rendered frames and on recorded videos as well. Furthermore, the accuracy of the algorithm on wingtip coordinates marked by a human expert and extracted by the algorithm is shown.

1.1 I have introduced a new algorithm for relative direction angle estimation for autonomous UAV visual SAA systems in the case when the two approaching aircrafts follow a straight path. I have shown that the accuracy of the algorithm in pure simulations when there is no noise or rounding to coordinates added is comparable with the numerical precision.

The relative direction angle, α can be calculated from the following formula:

$$\cos\alpha = \frac{\langle \bar{\boldsymbol{p}}' - \bar{\boldsymbol{q}}\,; \bar{\boldsymbol{r}} - \bar{\boldsymbol{s}} \rangle}{\|\bar{\boldsymbol{p}}' - \bar{\boldsymbol{q}}\| \|\bar{\boldsymbol{r}} - \bar{\boldsymbol{s}}\|} \tag{1}$$

where p_{p3} and p_4 are measured on the image plane and P_1 and P_2 are estimated based on the camera matrix, and the assumptions made on the two aircraft’s path. In this model the instances rotated by 180° are equal and the $\alpha = \cos^{-1} X$ function gives good solution in $\alpha = [0°; 180°]$ range.

The relative angle α should be in the $[-90°; 90°]$ range, so it is transformed according to the following rules. If $\alpha > 90°$, then $\alpha = 180° - \alpha$, if $\alpha < -90°$, then $\alpha = -180° - \alpha$. With these calculations the expected results are obtained consistently.

If the intruder is on the xy horizontal plane, p_{p3} equals p_4 and the α angle cannot be estimated with this algorithm. The altitude of our UAV can be easily changed with acceleration or deceleration.

With pinhole camera model, the given centroid point of the intruder is projected back from image plane to space to several distances. The wingspan of the intruder is 11m (36 ft. 1 in), which is the wingspan of Cessna 172, a typical light aircraft that shares the airspace with our UAV. Thus the wing is represented by an 11m line segment and is rotated in the previously calculated point. The field of view and resolution of the camera and the distance along x axis is required for the calculation. The fuselage of the aircraft is neglected. With these calculations the lower bound of the error is approximated.

1.2 I have investigated through simulations how the relative position of the intruder changes the accuracy. I have shown experimentally that the closer the intruder is to the horizontal (y) axis the bigger the error of the α. And similarly the bigger the distance along the x axis the smaller the intruder is in the image, therefore the spatial discretization gives higher error value. Furthermore, the proximity to y has a greater effect on the error than in the smaller distance case.

The measurements was made with the same pinhole camera and airplane model that I used in the first case, except that the calculated points are rounded, like in the case of a real camera.

The relative distance along the x axis is 1 km (0.62 miles), the resolution is 1920x1080 pixels, the horizontal field of view is 50° and the pixels are squares. The wingspan of the intruder is 11m (36 ft. 1 in), which is the wingspan of Cessna 172. The size of intruder in the image plane is between 15 and 20 pixels, depending on the rotation angle and the position.

I have shown that the azimuth angle has technically no effect on the accuracy, but the change in the distance of the intruder to the camera and in the elevation change the accuracy. The reason is that the larger the distance the smaller the intruder in the image and the bigger the altitude difference the more you observe the wing of the intruder.

1.3 I have investigated the reachable accuracy of the algorithm on wingtip coordinates extracted from rendered images and from real videos. I have shown experimentally that the accuracy can be close to the theoretical value with wingtip points selected by a human expert and extracted with a simple algorithm on rendered images. I have also shown experimentally that on real videos with a simple time average the noise introduced by the wind can be filtered out.

In our simulation environment pictures are taken and the wingtip pixel coordinates are selected by a human expert. With pinhole camera model, the given centroid point of the intruder is projected back from image plane to space to several distances and in every position it is rotated by specific angles in the xy plane. The resolution is 1920x1080 pixels and the horizontal field of view is 50° and the pixels are squares. The measurements have shown that with good wingtip coordinates in realistic situation the error can be close to the theoretical minimum.

The wingtip points were also extracted with a simple algorithm, which determines the wingtip coordinates from the segmented images. The extreme of y and z coordinates are used in appropriate order to get the coordinates. In this case when the intruder had been rotated with 80° and with -80° angles, the error of the estimation is bigger, because the simple algorithm could not distinguish between the pixels of the wing and the pixels of the tail. In contrast, in the mid-range the performance of this really simple algorithm is almost the same as the performance of the human expert (close to the theoretical limit).

2 Thesis: *Error analysis of the 4 chosen camera pose estimation algorithms in the case of UAV SAA application for the rotation calculation:*

I have chosen four feature point based relative pose estimation algorithm. A homography based solution as a basic algorithm with small computational need but with less accuracy. The eight point algorithm, as standard algorithm in epipolar geometry. The five point algorithm, as one of the state of the art algorithms with higher computational need, but with promising stability over the various scenes. Finally, MLESAC, as an iterative, stochastic solution. The aim of the investigation is to show the strengths and weaknesses of these algorithms in the aircraft attitude estimation task.

2.1 I have investigated the performance of the four chosen algorithms in simulations using two different real flight paths and synthetized images with randomly placed feature points and taken into account the model of the camera used on board with different resolution. I have experimentally shown that without any feature point coordinate error the five point algorithm is the best. The error of the five point algorithm is close to the numerical precision of the calculations. The errors of other two epipolar geometry based solutions are also at least one order of magnitude smaller than the 1 pixel angular resolution. And the homography has got an error that remains below 1 pixel.

For the tests 350 feature points are placed randomly with uniform distribution in a right prism which is 2000m wide, 3000m long and 30m tall. The point coordinates are between -1000 and 1000 in the Y direction and from 0 to 3000 in the X direction. The maximum altitude of the points is 23 m and the Z coordinate starts from 3 m beyond the ground level to simulate small holes.

For the camera projection the calibration matrix of one of our miniature camera is used. The internal calibration matrix is scaled in order to simulate cameras with different resolutions.

First, tests with absolute feature point precision are run. In this case the best achievable results are obtained because there is practically no spatial discretization, the effect of the temporal resolution change can be investigated independently.

2.2 I have investigated the effect of the translation on the performance of the four chosen algorithm. I have experimentally shown that the error is bigger as the time step is bigger in between the frames except for the five point algorithm in some situations.

I have shown the results of the pitch angle, which is most affected. Theoretically due to the bigger baseline separation bigger translation between the two frames could be advantageous for the three algorithms which are based on the epipolar constraint (five point, eight point and MLESAC). It can be seen in the figure practically this is not true, the error is bigger as the step is bigger in between the frames except for the five point algorithm in some situations. One possible explanation is that the number of the feature points which can be seen in both frames is reduced and the feature points are more drifted to the side of the image. On the other hand, the integral error altogether is smaller for the whole path.

2.3 I have investigated the possibility of the use of feature extraction algorithms with subpixel capability with the four algorithms. I have experimentally shown that except the five point algorithm, the pose estimation can benefit from the subpixel feature point calculation.

The sub pixel feature point extraction is simulated by random, normal distribution noise (0 mean and 0.5 pixel standard deviation) on absolute precise feature point coordinates. Surprisingly, the five point algorithm cannot benefit from the subpixel resolution. The eight point algorithm and the MLESAC have lower mean error values.

	Five point	Eight point	Homo-graphy	MLESAC
Absolute precision	$3.2 \cdot 10^{-11}$	$2.1 \cdot 10^{-3}$	$5.1 \cdot 10^{-2}$	$1.3\ 10^{-3}$
Subpixel	$1.2 \cdot 10^{-1}$	$1.1 \cdot 10^{-2}$	$7.2 \cdot 10^{-2}$	$2.0\ 10^{-3}$
Pixelized	$9.4 \cdot 10^{-2}$	$5.5 \cdot 10^{-1}$	$1.2 \cdot 10^{-1}$	$3.2\ 10^{-1}$

Table 1 Roll error of the four algorithms changing with different feature point precision for the CPAR=0.093°/px camera

2.4 I have investigated the performance of the algorithms in more general case, when the feature point coordinates are rounded, or are rounded and contain noise as well. I have experimentally shown that the five point algorithm performs the best with mean error value around 1 pixel. I have experimentally shown that the homography algorithm can perform almost as good as the five point, with mean error around 1.5 pixels. The computational need of the homography algorithm is 2 orders of magnitude smaller than the computational demand of five point algorithm in the number of the multiplications. I have experimentally shown that the pixelization has got a smaller effect on the homography algorithm than on the others. It can be stated that the homography algorithm can be used in those situations where the computational power is restricted.

4 Application of the results

As it is stated at the introduction, the aim of the research project is to develop an autonomous UAS with SAA capability. The UAS, which is developed, is the main application area of the results of this dissertation.

The EKF based motion estimation can benefit from the results of the relative direction angle measurements. When the wingtips of the intruder can be extracted, the direction angle can be used as a measurement for the EKF. Another possibility is that the calculated α angle is used at the initialization of the EKF.

The results from the error analysis of the four algorithms are used to develop an attitude estimator which can fuse the conventional GPS/INS data with the estimates from the camera. As a first step the results from the five point algorithm and the results from the homography are used to test the fused estimation. The tests confirmed that the homography is indeed less affected by the noise.

Furthermore the results and the used test environment can help to choose the appropriate camera and algorithms for a given situation. If the desired accuracy is known, based on the results and other requirements, like processing speed a suitable camera and algorithm can be chosen. The camera is determined by its CPAR and the algorithm is determined by the accuracy and the spatial-temporal constraints. Additionally more pose estimation algorithms can be tested easily in the simulation environment.

5 Acknowledgement

First of all I would like to thank and acknowledge the support of my scientific advisor and supervisor Ákos Zarándy. He helped me a lot to find my way at the beginning of the PhD, guided me throughout my work and motivated me with his enthusiasm. I would like to thank also Tamás Roska for his great support and fatherly guidance during my studies.

I am grateful to Péter Szolgay, Judit Nyékyné Gaizler, Ferenc Kovács and Ágnes Bércesné Novák for their continuous support and encouragement during the years.

I am also thankful to my closest collaborators for their help, and for all the discussions we had: Bálint Vanek, Zoltán Nagy, András Kiss, Péter Bauer, Tamás Péni, Borbála Pencz, Máté Németh and Arnold Leitli.

I am very grateful to my colleagues especially to those with whom we shared an office: Miklós Koller, Gábor Tornai, Mihály Radványi, Tamás Fülöp, Ádám Rák, András Horváth, Attila Stubendek, Domonkos Gergelyi, Csaba Józsa, László Füredi, Kálmán Tornai, Bence Borbély, István Reguly, Csaba Nemes, László Laki, Ádám Balogh, Norbert Sárkány, Zoltán Tuza, János Rudan, Balázs Varga, Dávid Tisza, Tamás Pilissy, Zoltán Kárász, László Kozák, Andrea Kovács, Dániel Szolgay, Ádám Fekete, András Gelencsér, Dóra Bihary, Zsolt Gelencsér, Antal Hiba, Endre László, Borbála Siklósi, András Rácz, Balázs Oláh, Anna Horváth, András Laki, Imre Juhász and Mátyás Jani for their help in any ways.

I am also thankful to our postdoctoral fellows András Oláh, György Cserey, Kristóf Karacs, Kristóf Iván, Zoltán Gáspári and Miklós Gyöngy for their inspiring work and for their encouraging worlds.

I acknowledge also the kind help of the personnel of Students' Office, in particular Katinka Tivadarné Vida, the Dean's Office, the Financial Department and the IT Department. I am also very thankful to Viktória Sifter from the Library.

I am especially grateful to my wife Kriszti who has encouraged me and believed in me all the time. I am very grateful to my mother, my father and my sister and to my whole family who always believed in me and supported me in all possible ways. I am very grateful to my daughter Julcsi, who let me work close to the deadlines.

6 Publications

6.1 The author's journal publications

[1] T. Zsedrovits, A. Zarandy, B. Vanek, T. Peni, J. Bokor, and T. Roska, “Estimation of Relative Direction Angle of Distant, Approaching Airplane in Sense-and-Avoid,” *J. Intell. Robot. Syst.*, vol. 69, no. 1–4, pp. 407–415, Jan. 2013.

[2] A. Zarandy, M. Nemeth, Z. Nagy, A. Kiss, L. Santha, and T. Zsedrovits, “A real-time multi-camera vision system for UAV collision warning and navigation,” *J. Real-Time Image Process.*, Sep. 2014.

6.2 The author's international conference publications

[3] T. Zsedrovits, A. Zarandy, B. Vanek, T. Peni, J. Bokor, and T. Roska, “Collision avoidance for UAV using visual detection,” in *Proc. of 2011 IEEE Int. Sym. of Circuits and Systems (ISCAS)*, 2011, pp. 2173–2176.

[4] T. Zsedrovits, A. Zarandy, B. Vanek, T. Peni, J. Bokor, and T. Roska, “Visual Detection and Implementation Aspects of a UAV See and Avoid System,” in *Proc. of 2011 20th European Conference on Circuit Theory and Design (ECCTD),* 2011, pp. 472–475.

[5] T. Zsedrovits, A. Zarandy, B. Vanek, T. Peni, J. Bokor, and T. Roska, “Estimation of Relative Direction Angle of Distant, Colliding Airplane in Sense-and-avoid and Tracking,” Presented at the *International Conference on Unmanned Aircraft Systems*, Philadelphia, Pennsylvania, 2012.

[6] T. Zsedrovits, A. Zarandy, B. Vanek, T. Peni, J. Bokor, and T. Roska, “Azimuth estimation of distant, approaching airplane in See-and-avoid Systems,” in *Proc. of 2012 13th International Workshop on Cellular Nanoscale Networks and their Applications*, Turin, Italy, 2012, pp. 1–6.

[7] T. Zsedrovits, P. Bauer, A. Zarandy, B. Vanek, J. Bokor, and T. Roska, “Towards Real-Time Visual and IMU Data Fusion,” Presented at the *AIAA Guidance, Navigation, and Control Conference*, Reston, Virginia, 2014.

[8] T. Zsedrovits, P. Bauer, A. Zarandy, B. Vanek, J. Bokor, and T. Roska, “Error Analysis of Algorithms for Camera Rotation Calculation in GPS/IMU/Camera Fusion for UAV Sense and Avoid Systems,” Presented at the *International Conference on Unmanned Aircraft Systems*, Orlando, Florida, 2014.

6.3 The author’s other journal and conference publicaitons

[9] B. Vanek, T. Peni, J. Bokor, T. Zsedrovits, A. Zarandy, and T. Roska, “Performance analysis of a vision only Sense and Avoid system for small UAVs,” Presented at the *AIAA Guidance, Navigation, and Control Conference*, Reston, Virigina, 2011.

[10] B. Vanek, T. Péni, Á. Zarándy, J. Bokor, T. Zsedrovits, and T. Roska, “Performance Characteristics of a Complete Vision Only Sense and Avoid System,” Presented at the *AIAA Guidance,* Navigation, *and Control Conference*, 2012.

[11] Z. Nagy, A. Kiss, A. Zarandy, T. Zsedrovits, B. Vanek, T. Peni, J. Bokor, and T. Roska, “Volume and power optimized high-performance system for UAV collision avoidance,” in *Proc. of the 2012 IEEE Int. Symp. on Circuits and Systems*, 2012, pp. 189–192.

[12] A. Zarandy, T. Zsedrovits, Z. Nagy, A. Kiss, and T. Roska, “Visual sense-and-avoid system for UAVs,” in *Proc. of 2012 13th International Workshop on Cellular Nanoscale Networks and their Applications*, 2012, pp. 1–5.

[13] A. Zarandy, T. Zsedrovits, Z. Nagy, A. Kiss, P. Szolgay, and T. Roska, “Cellular processor array based UAV safety system,” in *Proc. of 2012 13th International Workshop on Cellular Nanoscale Networks and their Applications*, 2012, pp. 1–2.

[14] A. Zarandy, Z. Nagy, B. Vanek, and T. Zsedrovits, “A five-camera vision system for UAV visual attitude calculation and collision warning,” *Comput. Vis. Syst. Lect. Notes Comput. Sci.*, vol. 7963, pp. 11–20, 2013.