State of the Art of Image-aided Navigation Techniques for Aircraft Approach and Landing

Jérémy VEZINET, *ENAC Telecom Lab, France*
Anne-Christine ESCHER, *ENAC Telecom Lab, France*
Alain GUILLET, *AIRBUS, France*
Christophe MACABIAU, *ENAC Telecom Lab, France*

BIOGRAPHIES

Jérémy VEZINET graduated as an electronics engineer in 2010 from the ENAC (French National School for Civil Aviation) in Toulouse, France. Since 2011, he is a PhD student in the Signal Processing and Navigation Research Group (SIGNAV) of ENAC in Toulouse, France. The PhD project is in collaboration with the inertial navigation systems department of Airbus. He is currently working on multi-sensor hybridization.

Dr. Anne-Christine ESCHER graduated as an electronics engineer in 1999 from the ENAC in Toulouse, France. Since 2002, she has been working as a researcher and as a lecturer in the signal processing and navigation research group (SIGNAV) of the TELECOM Lab in the ENAC. She received her Ph.D. in 2003.

Alain GUILLET graduated as aeronautic engineer in 1984 from the ENSICA in Toulouse, France. Since 1985, he has been working in AIRBUS where he held different positions, mainly in the Engineering. Since 1995 he has been in charge of various projects on Air Data & Inertial systems, on the whole development cycle: Specification, Development, Certification, In-Service Support as well as Research.

Dr. Christophe MACABIAU graduated as an electronics engineer in 1992 from the ENAC in Toulouse, France. Since 1994, he has been working on the application of satellite navigation techniques to civil aviation. He received his PhD in 1997 and has been in charge of the signal processing lab of ENAC since 2000

ABSTRACT

Global Navigation Satellite Systems (GNSS) and Inertial Navigation Systems (INS) are currently the main navigation systems for aircraft. INS navigation is based on dead-reckoning principle so small errors in the measurements of vehicle accelerations and rotation rates can cause non-negligible integration drift. As an alternative navigation mean, more and more considered in navigation applications, image-aided inertial navigation is a complete autonomous navigation opportunity because it is only based on sensors onboard that provide information from the dynamic of the vehicle and the observation of the scenery. This system of navigation might be used in the particular case of loss of other systems of navigation that need additional exterior equipment (Global Navigation Satellite System (GNSS) or Instrument Landing System (ILS) for example). Besides in the particular context of civil aviation, during precision approaches (currently done with ILS or GPS augmented with GBAS or SBAS), it could be interesting to use an additional autonomous mean of navigation like video measurement that could overtake the need of additional ground or space infrastructure. But the performance has to be analyzed so as to determine if such a system can satisfy the precision approaches requirements which are very stringent.

Development of techniques of navigation with imaging sensors does not mean a complete replacement of the currently widely used hybridization technique that consists in coupling Global Positioning System (GPS) with Inertial Navigation System (INS). Imaging sensors would be more an additional navigation means, to be combined with existing navigation means. Lots of studies concerning the use of visual measurements for navigation have been achieved and cover a large range of applications, from precise guidance of a UAV during a landing [5], to the estimation of translation or velocity between two successive image of the landscape [7]. From the review of the applications, one of the most robust solutions to use visual measurements is to extract some geometrical measurements from the detection of targets in the image and to use them in a Kalman filter with GPS and INS measurements [1]. The optimal solution would be to couple these geometrical visual measurements, obtained through the detection of particular features, with other means currently used on board (mainly INS measurements and GPS pseudo-range measurements).

The information extracted by visual measurements can characterize precisely the relative position of the camera with respect to the feature detected in the image. However, the knowledge of a distance (range between

camera and feature, height of camera above ground or dimension of an object on the ground) is needed to determine a metric scale which is absolutely needed to interpret the image measurements.

The current paper proposes a state of the art of image-aided navigation methods. A description of the key elements or characteristics used in these methods is done. Based on this, a proposition is done for a video-based navigation system for approach and landing operations. This is a first step of feasibility study for an aircraft landing system based on video.

INTRODUCTION

Visual navigation is one of the oldest known navigation method based on the observation of the heavens (it was called celestial navigation). Some of the navigators used equipment to determine angles between stars and horizon or vertical, and then to estimate their position. Basic principle of visual navigation is defined by a simple fact: the observation of the world and objects around us is the most reliable information for deducing our relative position with respect to our environment.

High-end transport aircraft currently uses Global Navigation Satellite System (GNSS) and Inertial navigation System (INS) as the main navigation means. They are even commonly coupled in a hybridized architecture to provide a level of performance that can reach requirements from en route operations down to Non-Precision Approaches (NPA) and Required Navigation Performances (RNP) operations. However it still exists some operations with very stringent requirements, in terms of accuracy, continuity, availability and integrity, like precision approaches of category I, II and III, where GNSS and INS are not sufficient.

As a way to improve performance of the navigation systems and to reach the most stringent operations, a solution is to add new sources of information (already available on board or not) and to use their measurements in a more global hybridized architecture. The main advantage of fusion of different sources is the compensation of drawbacks of each source and the improvement of the integrity and accuracy of the estimated navigation parameters. A lot of studies have been done about the coupling of GPS and/or INS with various others sensors (Radio-Altimeter, Wheel Speed Sensors, Air Data Sensors, Magnetometers,…) but one type of sensor that seems to offer the largest set of applications is the video sensors, providing optical measurements.

Improvement of low-cost, light and high resolution video sensors has led to an interest in extracting accurate navigation information such as position, velocity or attitude from an optical measurement. Cameras are currently available onboard of some aircraft and they are mainly used to assist the pilot for ground navigation or to entertain the passengers during flights. However, observation of the surrounding scenery can be considered as a good source of information for navigation purpose. For instance, an image flow measurement can be representative of the position, the velocity and the orientation of the aircraft. But a correct transcription of the details in the landscape can only be done taking into consideration physical limitations and characteristics of a video sensor: resolution, field of view, dimension and position of the sensor.

Visual measurements can provide a lot of information from a simple image. A basic digital optical sensor measures the intensity of the light entering an aperture with a Charge-Coupled Device (CCD) or a Complementary Metal-Oxyde-Semiconductor (CMOS). This measurement, as a snapshot of the surrounding scenery, provides information of light intensity at each pixel that constitutes the sensor. Therefore, the information associated to one pixel has to be associated to their coordinates so as to extract the location of the pixel in the image frame. An optical sensor is most of the time associated to an image processing algorithm to be able to detect particular pixels in the image that are associated to area of interest in the scenery. Finally the location of those pixels, images of features (or point of interest), can contain geometric information that can be used for navigation purpose.

This paper presents some applications that are considered as a snapshot of the vision-based navigation method. Our objective is to identify some important elements when dealing with video. It is thus organized as follows: the first section presents an overview of some applications using video. It is a succinct description of some image-based navigation methods with an identification of the main assumptions and the principle of the algorithm. The second section is a first classification of these methods. This classification aims at identifying advantages, drawbacks and context of the methods in order to select the parts that can be considered in the context of approach and landing of aircraft. The third section is a proposition of a method that could be used during an approach or a landing operation and the description of the process. This is a first step of feasibility study for an aircraft landing system based on video.

SOME EXAMPLES OF APPLICATION

This part is a review of some papers found in the literature that present a video-based navigation method. It aims at covering the largest range of methods so as to identify most of the key elements and characteristics involved with the use of a video-based navigation system.

Topological localization

Topological cartography deals with a discrete representation of the environment without scale metric necessity. The localization then aims at recovering a location in the topological space. This topological representation appears in video-based navigation methods described in [6], [7] and [8]. Ulrich and Nourbakhsh, in [5] used a topological approach for the localization of a mobile robot in a relatively closed environment (indoor or along a road). Figure 1 presents an example of a topological map of an environment.

Topological localization is directly opposed to geometric localization because it avoids maintaining a metric map of the environment and allows operating directly in image space. Indeed, geometric localization usually uses a grid as a map representation (in two or three dimensions). They attempt to keep track of the mobile's exact position with respect to the map's coordinate system.

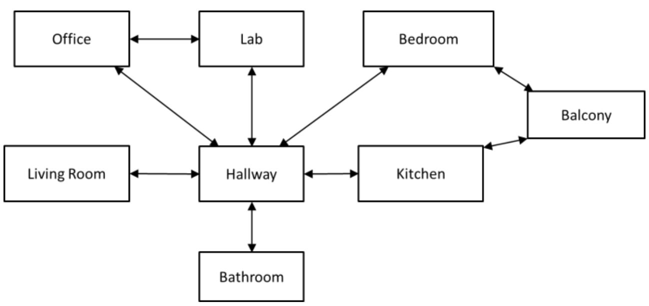


Figure 1 – Environment graph of an apartment [5]

In the same way, Segvic, Remazeilles, Diosi and Chaumette in [4] present a method of localization for an autonomous mobile robot in two steps: a learning step and a localization step. The first step is called mapping components and it aims at acquiring images through a learning stage and then extracting interest points (or features) in these images. Construction of the "map", called environment graph (see Figure 2), is done during a previous navigation procedure with other means of navigation (or human interaction). During this procedure, some images provided by the camera are recorded as key images, also called nodes, in order to constitute the graph. The selection is done based on a criterion of difference between two successive nodes: they have to be sufficiently separated to minimize the number of nodes in the final graph but they have to have enough similarities to find common sets of features. Once the map has been constructed, the graph is completed with the set of features X_i and the scale metric s_i in each image I_i ; and the two-view geometry W_i (including rotation, translation and metric between the two surrounding images) and match arrays M_i between common features in I_{i-1} and I_i in each arc *i*.

The second step in [4] is the localization component and aims at locating in the map previously created the current image registered by the camera. The principle of localization among the nodes of the graph is based on comparison between the current image and the reference image. It allows localizing the vehicle among the topological graph. The process of tracking features during the passing from a node to another is done by computing the two view-geometries and three view geometries between the surrounding nodes and current image (see Figure 3). The process aims at keeping the tracking of consistent features and updating the topological localization of the vehicle.

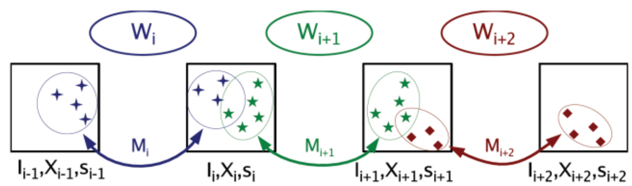


Figure 2 – Linear environment graph [4]

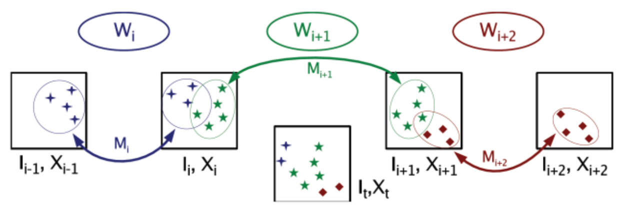


Figure 3 – Localisation of the current image in the environment graph [4]

Finally, most of the applications presented below involve repetitive trajectories like a tour or indoor areas. Even if the topological localization approach allows freeing from a metric constraint, it seems not really adapted for an application in an open and unknown place.

Visual servoing

Visual servoing can be defined as the use of vision sensors to provide closed-loop feedback control of some moving component. In visual servoing, the system aims at minimizing an error function. Visual servoing methods are usually used for controlling the pose of industrial robots arms but it is also used in aircraft landing applications. [9], [10], [11], [12], [13] and [14] present image-based visual servoing for an aircraft during approach and landing. Miller, Shah and Harper in [15] use, in addition, image registration for landing a UAV on a runway. This method is based on the comparison between a test frame and a reference frame previously registered. The approach presented in [15] only uses visual measurements and a stack of reference frames. The navigation process is in 3 steps: the localization of the runway in each image, the estimation of the attitude of the UAV and the steering of the UAV towards the runway maintaining the correct glideslope. A forward pointing camera is mounted on the UAV and its intrinsic calibration matrix C is known (C is one of the matrix that links the homogeneous coordinates of a 3D-point with coordinates of the point projected in the image plan, it depends on the focal length, the size of the image and the coordinates of the optical center).

Navigation method used in this paper implies that during a previous flight, images have been registered to create a stack of reference frames. During this previous flight the flown trajectory is considered as the reference trajectory, the ideal glide path (see Figure 4). The set of reference images contains frames taken as the UAV gets closer to the runway and that are sampled at an increasing frequency as the altitude decreases. A preprocessing step is necessary to annotate two particular points: the vanishing point (intersection between the horizon line and runway axis) and the beginning of the runway (more precisely the spot where the UAV should touch the ground). They are circled in yellow in Figure 4.

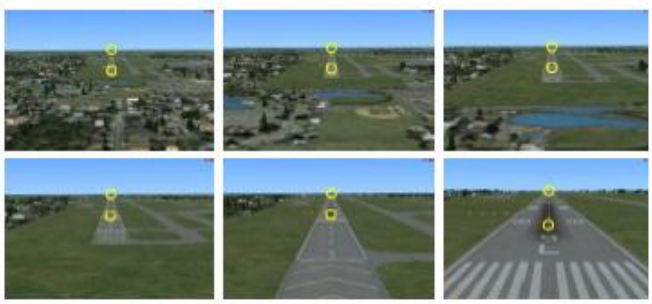


Figure 4 – Set of reference frames taken from a video as the UAV gets closer to the runway. The vanishing point and a point at the beginning of the runway (the spot where the UAV should touch down) are annotated in each frame. [15]

From the measured image (the current test frame) provided by the video during landing, a Scale-Invariant Feature Transformation (SIFT) algorithm is used to compute the planar homography H (the planar homography relates any point on the ground in a particular view (the reference view) to the corresponding point in a different view (the current view)) between the test frame and the reference frame so that they have the most correlations. Once the best reference frame identified and H matrix computed, it is possible to project the two points annotated in the reference frame to the corresponding points in the test frame (see Figure 5).

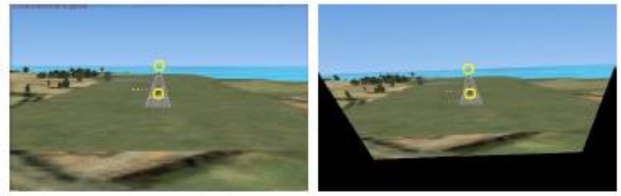


Figure 5 – Projection of two points in the test frame (on the left the reference frame, on the right the tranformed reference frame in the same view as the test frame)

The relative position of these two points in the test frame allows estimating the UAV attitude and steering. This technique is similar to the runway analysis conducted by a pilot during landing. Finally geometrical information red in the test image is directly converted into a command for the actuators of the UAV. The underlying property is that the relative position of the two points is related with the UAV attitude because usually this is based on the same interpretation as that done by the pilots when landing. The reading of the coordinates of the two annotated points (vanishing and threshold points) permits to estimates three geometrical parameters: the runway offset, the runway angle and the runway distance (see Figure 6).

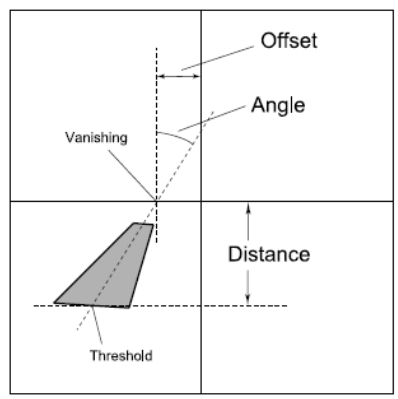


Figure 6 – interpretation of the measured parameters in the test frame [15]

Extraction of geometric features from image measurement allows using visual servoing techniques to minimize error of position or orientation during landing. Furthermore, it is possible to interpret these geometric features in a different way by using a metric scale. Image measurements can then provide real position parameters (height, range, and attitude). Such a method is detailed in [16].

Visual servoing described in [15] appears as a way to extract geometric parameters that characterize deviation with respect to a previous trajectory, considered as the reference one. In the next part of the method, these parameters are used as steering commands for piloting the UAV along the ideal glide path. The major interest of this study is the principle of estimation of deviation parameter by computing the homography between two images.

Path planning

Sinopoli, Micheli, Donato and John Koo present in [17] a video-based navigation method for a UAV (a helicopter) without a complete knowledge of the environment. It is based on optimal path planning trough a hierarchical approach. The method is detailed through a four-part algorithm implemented on a Flight Management System (FMS). A vision system is coupled with the FMS. The first two steps of the algorithm, called the Strategic Planner (SP) and the Tactical Planner (TP), are respectively an offline path planning and an online local trajectory computation.

The SP step consists in creating a set of waypoints. This first part of the algorithm is not performed with the video system. However a wavelet transformation from a Digital Elevation Model (DEM) of the environment then a Dijkstra optimization algorithm are employed to find the shortest path between two waypoints on the transformed grid (see Figure 7).

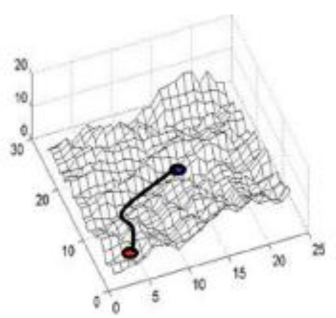


Figure 7 – Path between two waypoints on a DEM [18]

The TP step is then a vision-based technique and is based on local obstacles avoidance. TP consist in connecting the waypoints provided by the SP. The TP builds a sub-grid (see Figure 8) of known dimensions, between two successive waypoints and compute a risk map from the video measurements. The measurements associated with the position and the linear and angular velocity of the UAV (provided by GPS and inertial sensors) generate a depth map (i.e. the distance between the camera and the object filmed). Finally, the depth measured between the camera and the supposed point P (corresponding to the center of a cell in the sub-grid) is compared to the real distance between the point P (of known coordinates) and the camera. The difference between the supposed range between the UAV and the targeted cell P and the measured depth measured characterizes the risk of presence of an obstacle. These values (i.e. for each points of the sub-grid) create a risk map all along the cells of the sub-grid between the two waypoints. The optimal trajectory is considered as the one with the least risk.

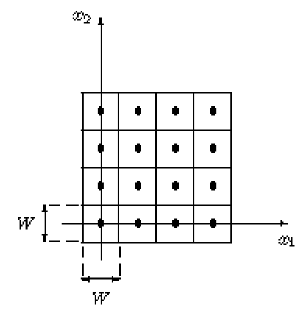


Figure 8 – Sub-grid decomposition [18]

The paper thus presents a particular method to find an optimal path between two selected waypoints. The navigation is done by coupling GPS, inertial measurements and a video system for obstacles detection. Such a method can be employed for autonomous navigation in a constrained area (with high relief variations or urban area). The area has to be partially known to establish an initial guess about the optimal path. The video contribution is to update the path in real time in case of erroneous a-priori DEM, the apparition of an obstacle or error in the path navigation. The method is adapted to the particular context presented here and may be considered as an additional means for collision avoidance. However precise navigation issues cannot be ensured by such an algorithm.

Simultaneous Localization and Mapping and visual odometry

Simultaneous Localization and Mapping (SLAM) techniques deal with the problem of building a map of an environment unknown by the mobile while navigating this environment using the created map. SLAM algorithms generally consist of multiple parts: landmark extraction, data association, state estimation, state update and landmark update. Each of these parts can be done in many ways [18]. SLAM algorithms are usually implemented from a basic odometry system coupled with a range measurement sensor to locate the landmarks (laser, sonar or stereovision).

Mirisola and Dias present in [19] a vision-based navigation method using an Attitude and Heading Reference Sensor (AHRS). The inertial sensor is mounted on the camera and provides orientation measurements. The navigation method deals with the reconstruction of a trajectory from an images sequence. The employed process is denoted as visual odometry because the video system aims at estimating translation vector between two successive images. The principle detailed in [19] is based on the estimation of the translation vector by pure homography, that is to say the determination of the relation between two sets of homogeneous pixel coordinates that represent the same points imaged from two different positions. In that case the homography (1.1) provides the rotation and the translation between the two views (see Figure 9). In that equation, λ represents the scale factor, R the rotation matrix, t/d the translation vector and n the 3D plan normal.

$$\lambda H = \lambda\left(R - \frac{t}{d}n^T\right) \tag{1.1}$$

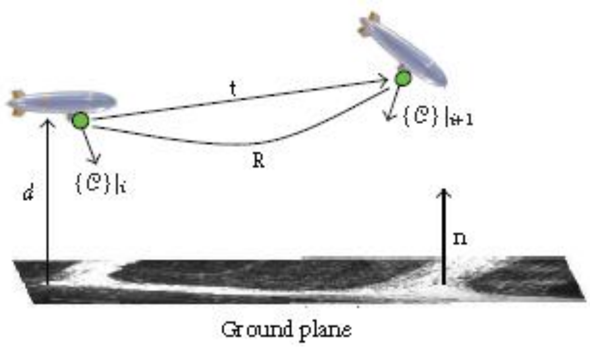


Figure 9 – Image of a 3D plane by a moving camera [19]

Besides the method developed presents a variation with a problem that comes down to the estimation of a pure translation vector. Indeed, this other process starts with the projection of the recorded image on a virtual horizontal plane (i.e. plane with normal n parallel to gravity). This transformation is done with the computation of the infinite homography. The infinite homography literally represents the transformation generated when the plane is moved at the infinity but it is also the homography between two images taken from the same point but only rotated (the rotation corresponds to the orientation of the camera with respect to the normal view of the horizontal plane). In that case, the computation of the infinite homography is only done from the AHRS attitude measurements (see [19] for more details).

When the entire set of images is transformed through this infinite homography, the sequence corresponds to image measurements taken from a camera constantly oriented along the vertical. Then after the selection and matching of a set of features and their correspondence in the next image, a procruste procedure is done to find the transformation and generate the 2D translation vector, the rotation matrix and the scale factor between the two images (see Figure 10).

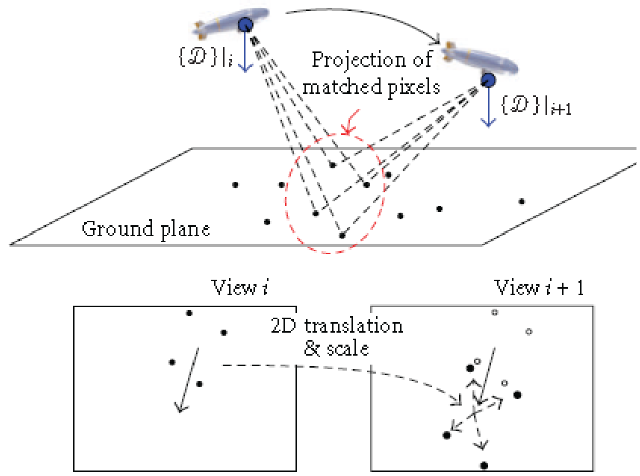


Figure 10 – Principle of trajectory recovery [19]

The method detailed in [19] is at the end coupled with a SLAM algorithm that aims at reducing the visual odometry. Some methods involving SLAM are also described in [20] and [7]. However a straight movement seems to be a major constraint in visual odometry and SLAM techniques. Then navigation ensured by SLAM algorithms for high dynamic mobile could not verify such a constraint and guarantee high performance.

Fusion of imaging and inertial sensors

The last method in literature deals with considering video sensors as a source of measurements that can be coupled to INS and/or GPS measurements through a hybridization filter. Veth, Giebner, Raquet, Ebcin and Fletcher detailed in [21], [22], [23] and [24] a tight-coupled image-aided inertial navigation through a Kalman Filter (KF) or an extended Kalman Filter (EKF). In [21], the measurement model from video sensors is described with its integration in a KF. The particularity of the state vector of the KF is that it contains positions of targets that are considered as features for the video. Indeed, these targets are at unknown locations. The paper also presents relations between the measurements and the target locations so as to generate the measurement function (see (1.2) to (1.4)).

Image measurements for a target in the environment are defined as angles depicted in Figure 11.

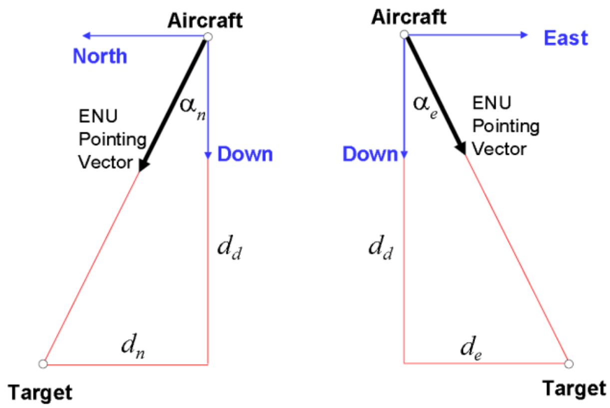


Figure 11 – Definition of image measurements [21]

$$d_n = R_N\left(lat_i - lat_{ins} - \delta lat\right) \qquad (1.2)$$
$$d_e = R_E\left(lon_i - lon_{ins} - \delta long\right) \qquad (1.3)$$
$$d_d = alt_{ins} + \delta lat - alt_i \qquad (1.4)$$

In these equations d_n, d_e and d_d correspond to North, East and Down distance as depicted in Figure 11, lat_{ins}, $long_{ins}$ and alt_{ins} are position estimated by the INS, δlat, $\delta long$ and δalt are the INS position estimation error and lat_i, $long_i$ and alt_i are the target positions. The proposed method is based on navigation using video when GPS outages and the paper shows that position error is reduced in that case.

However in [22], [23] and [24], the algorithm is depicted as a pure image-aided inertial navigation. GPS is not used in the method and the video system is based on stochastic feature projection aided by inertial measurements. Finally this tight fusion of optical and inertial sensors can provide an autonomous navigation and good performance. The model of video measurements and their integration in a KF presented in [21] represents a very good way to provide a navigation solution that can reach high level of performance.

VIDEO-BASED NAVIGATION KEY ELEMENTS

As depicted in the previous section through few examples, there are various methods for visual-based navigation. Even if they are applied in different contexts and for different applications, they have common key elements or characteristics, and in this part we identify some of them.

Image registration

Visual-based navigation methods are often used in a closed space or a well-known environment (for example indoor, or along a defined trajectory). The purpose of the navigation is to locate with respect to a known reference that can be a fixed point or a previous trajectory. In the second case, to be able to locate itself in the environment with respect to a previous path, the visual system needs a set of images considered as references (images recorded during a previous trajectory navigated with other means). The definition of reference frames in [15] recorded during a previous landing, or of a graph of reference images in [4] in order to locate the test frame among the references and then extract and/or extrapolate data in the surrounding reference images, illustrates this principle. In that way it is possible to compare each current test frame recorded by the camera with the images in the database, in order to establish the best correspondence. Once the best reference image has been identified, it is possible to extract data from the reference image (features or position when the image was taken for example) to estimate the real parameters corresponding to the current frame (feature prediction or position extrapolation). However, this method has few drawbacks. The first one is that the coverage area must be restrained: references are available only in a known area and it is not possible to use references everywhere. But the major constraints is that the reference images have to be recorded in a database and this database needs to be sufficiently large to contain images and exact position and orientation of the camera in order to precisely describe the area. Yet the method presented in [15] shows that it is possible to reduce the number of images in the database (or to adapt the number with the trajectory) when velocity is low, while when velocity increases the number of image has to be higher. But the advantage is just a reduction of the size of the database. Such a method cannot be generalized to any area because it needs reference images for every area where we want to navigate.

The interest in having a reference corresponding to the ideal path is that the system will be able to localize itself with respect to the database in real-time. This solution allows freeing from a scale metric needed to estimate an absolute position from images. However in case of absence of any reference measurement this scale metric is needed.

Knowledge of a metric

As it is explained in [19], estimation of navigation parameters (orientation, position, altitude) using visual aids requires the knowledge of geometrical information (range, height or dimension) in the surrounding scenery. This metric allows traducing a distance in pixels on the image to a distance in meters in the real environment. Thus in most of the methods described in the previous parts, video sensor is coupled with a sensor that is able to measure the ranging distance or the altitude (LIDAR, stereopsis, altimeter or even GPS or INS when video appears as an aiding measurement).

However, if an image processing algorithm can detect at least two known specific features in an image, then we can estimate the metric scale. For example, Figure 12 shows the side view of a forward pointing camera horizontally mounted on an aircraft during landing. During an approach operation, if we consider that an image processing algorithm is able to detect (automatically or with pilot assistance) the beginning and the end of the runway (depicted on the figure as threshold and horizon points), we know that the real distance between those two points is the exact length of the runway. Then with the knowledge of the characteristics of the camera (focal length and sensor dimensions especially), it is clearly possible to estimate an external quantity such as an height H, a range (R_H and R_T for the horizon and threshold range) or a horizontal distance (D_H and D_T) to one of these points (see (2.1) and (2.2) for the equations). Based on the usual pinhole camera model, in Figure 12, P_1 and P_2 are respectively the image of the threshold and horizon points of the runway with x_1 and x_2 their vertical coordinate in the image (the center of the image is considered as the origin). The camera is considered as calibrated and the distortion effects have been canceled.

$$\mathrm{H} = \mathrm{D_T} \times \tan\left(\frac{x_2}{f}\right) \qquad (2.1)$$

$$\mathrm{H} = (\mathrm{D} + \mathrm{D_T}) \times \tan\left(\frac{x_1}{f}\right) \qquad (2.2)$$

Then from the measurement of the coordinates of the two images points (x_1 and x_2) it is possible to estimate D_T and H by resolving the system above.

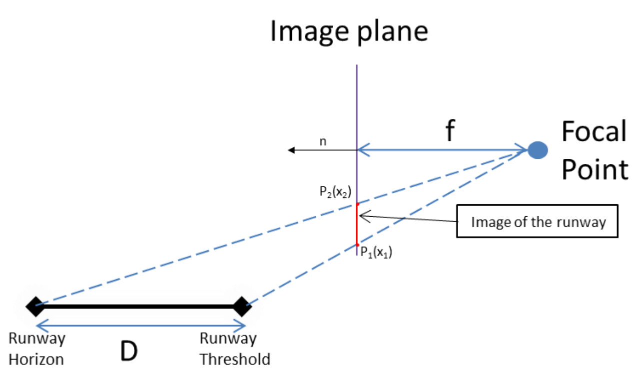


Figure 12 – Determination of the metric from a known dimension

Image measurement model

There are different ways to use measurements obtained by a video sensor. The simplest way is through video-servoing methods in which the steering of the mobile is commanded by the difference between the measured value and the desired value. This solution follows the principle that a camera mounted on a vehicle can follow a line (printed on the ground for example) and correct the trajectory so as to be always aligned with this line. We can extend the principle saying that the camera can detect features with a particular pattern in the image and send data of navigation that aim at keeping always these points with the same pattern. For example, if we consider the same two points as in the previous part (runway threshold and horizon points), a camera mounted on a UAV in the forward direction and horizontal can detect if the two features are aligned with the vertical axis of the image or not (see Figure 13). A signal proportional to the misalignment may thus be provided to the actuators to align the UAV with the runway axis. This is the principle used in [15].

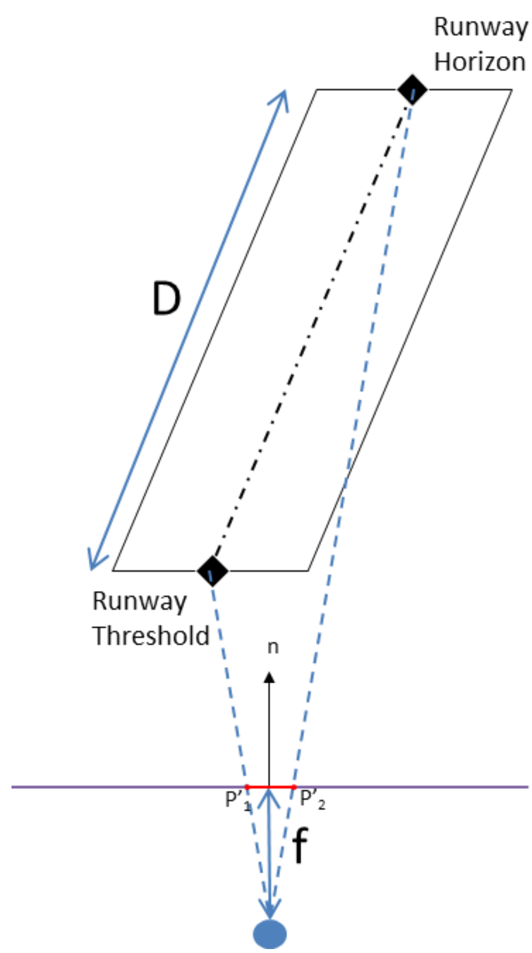


Figure 13 – upper view during a landing operation

In fact, the method depicted in [15] can be seen as a video-servoing technique in which the navigation system of the UAV provides steering information from visual measurements but cannot estimate a precise measurement of the position of the UAV. It only acts on the actuators of the UAV trying to compensate the deviation between the real trajectory and the ideal one but does not provide the values of the deviation.

In the same way, the method in [7] does not use a particular optical measurement. The video is used to identify a displacement between two successive images and to estimate the translation vector between them. In fact the measurement is associated to the image and the set of features but there is no measurement that corresponds to a particular feature.

In the contrary, [21] uses video in a tight coupled architecture and uses angular visual measurements that correspond to angles between the normal vector of the image and the vector pointing to the target (the line of sight). Two angles can thus entirely describe the position of the object seen by the camera and can easily be related to the position of the image of the target. This angular representation is similar to azimuth and elevation coordinates for a satellite for example. The description of that representation is illustrated in Figure 11. This representation permits to relate the pixel coordinates of a detected landmark to position in the NED navigation frame (North, East, Down). The interest of that is that a simple relation can be found between the state vector when using a hybridization algorithm and the measurements by computing the partial derivative of equations in (1.2), (1.3) and (1.4) as presented in [21].

APPLICATION FOR CIVIL AVIATION AIRCRAFT APPROACH AND LANDING OPERATIONS

Landing and approach procedures require high level performance in terms of accuracy, integrity, continuity and availability. Some categories of approaches require special equipment on board and/or on the ground (ILS, GBAS,…) but they are not available on every plane and every airport. Some video-based navigation methods represent an autonomous means of navigation because they only need a video sensor (camera) and an image processing algorithm to provide data that can be used for position or velocity estimation. However, we saw from the previous part that image measurements can be related to geometrical data only if we have an external metric (height, range, dimensions of a visible object,…). This part details a proposition of a video-based navigation method for approach and landing operations from elements and ideas identified in the previous parts. The entire method has to be considered as a scenario of approach for an aircraft using a video landing system.

The image processing part that can detect the runway or a particular point in the runway will not be studied because it is not in the frame of this PhD work. So it represents a quite strong hypothesis but some studies have shown that this detection is possible and allow justifying our hypothesis. For example [2] presents a robust method to detect the runway axis in all weather, time and visibility conditions. This method is based on detection of symmetry. Some other methods showed the capability of detecting horizon line by working on the contrast [15].

Finally for the scenario proposed, a basic pinhole model will be used for the video sensor.

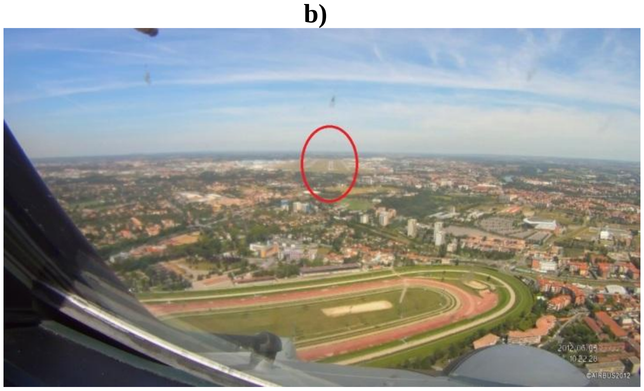
Context definition

The work done in the state of the art on vision-based navigation allowed imagining a video system that could be used on civil aviation aircraft for precise approach and landing operations. The starting point is to take inspiration from the current landing systems (ILS, GBAS, …) at the availability and coverage level and to define a service volume. This service volume area would be defined through three bounds defined later and illustrated in Figure 15. Thus, we consider that, with ideal visibility conditions, clear sky and the runway located in an isolated area, the maximum distance at which the runway can be seen is around 30 Nm (i.e. 55km). This distance is proposed based on what human can do in these conditions. However it seems non-realistic to consider that image processing algorithm can detect an object such as a runway farther than 30Nm.

Figure 14 presents different views from a camera mounted on an Airbus aircraft during landing at Toulouse-Blagnac Airport. The visibility conditions are good but the airport is in a dense urban environment. We can identify the runway from few kilometers (around 10km) and aided by human interaction (selection by the pilot) in order to reduce the area of searching, the image processing algorithm can detect and track the points corresponding to the beginning and the end of the runway. The range of 30Nm is taken to represent the largest limit from which the video system starts analyzing the images and detecting features. The second range that has to be considered represents the limit for stopping tracking the runway. If after this limit the video system is not able to find and lock the runway, this means that the video system cannot be used for landing. The last range that has to be considered is close to the runway when the video system is able to detect and track two features on both sides of the runway (so not only the threshold and vanishing/horizon points of the runway). This last range can be estimated around 1Nm before the threshold of the runway. The Figure 15 depicted the various areas defined above.

a)

b)



c)

Figure 14 – Image from a camera mounted on a A400M – Toulouse – a) at 16km – b) at 10km – c) at 5km

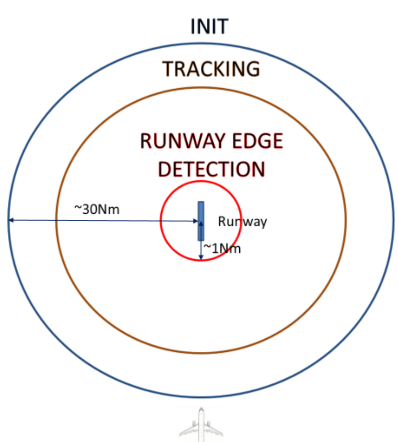


Figure 15 – Diagram of a standard approach using video system

The definition of these different limits is done considering that a video sensor is at least as good as the human vision. These limits determine the coverage area of the video system but also determine the performance characteristics of the video system.. This is described in the next part.

Video Sensor characteristics

This part deals with establishing the relation between the coverage areas presented in the previous part and the characteristics of the video sensor (focal length, size, field of view and resolution). As it is depicted in Figure 12 and equations (2.1) and (2.2), the relation between the distance between two features and the coordinates of the points in the image is simple. For the study we will consider a video sensor with a constant focal length of 50mm. This corresponds to a usual video sensor and also more or less to the human vision.

Once the focal length is defined, the field of view and the size of the video sensor can be defined as they determine the aperture of the camera (i.e. the size of the image that can be recorded). They can be defined by deciding that when the aircraft reach the runway threshold, the runway width has to be seen entirely by the video sensor, to be able to see the runway side lines when landing. In such a case, features can be detected on each side of the runway, and as previously explained, knowing the width of the runway we can compute the distance between the two corresponding pixel. This will define the minimal width of the video sensor. Besides to determine the minimal height of the video sensor, the same process can be done considering that the runway length has to be seen entirely by the video sensor, to be able to see the runway threshold and vanishing/horizon points when landing.

The last characteristic of the video sensor is the resolution (the resolution corresponds to the density of pixels in the image and not the number of pixels, it is expressed in dot per inch: dpi). The resolution of the sensor can be determined by computing the size of the smallest object detectable. We can consider that we want to be able to track the threshold and the vanishing/horizon points of the runway, from a distance of several nautical miles, and considering we know the real distance between these two points (so the length of the runway). Thus we can compute the corresponding distance in the image, between the two image points (always using the technique depicted in Figure 12). This distance in the image gives us the distance between two pixels that can be differentiated by an image processing algorithm. In ideal condition we could consider than two juxtaposed pixels can detect two different features. But in a realistic way, taking into account the sensor noise, a distance of several pixels is needed to detect two different features. Figure 16 illustrates the repartition of pixels in a video sensor. In that configuration, if the upper pixel corresponds to the image of the horizon point of the runway and the lower pixel to the image of the threshold point, there is a separation of five pixels between the two runway points, which corresponds to the length of the runway.

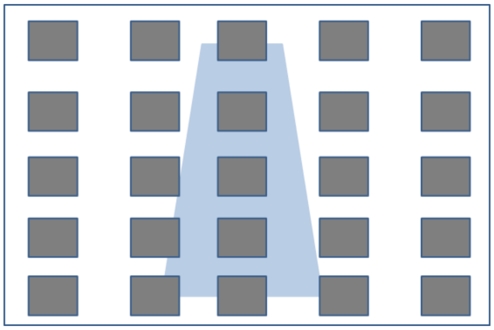


Figure 16 – Resolution determination

Video measurement and hybridization

One main idea is this study for using a video system in navigation is to consider that the video sensor is an additional navigation source that can be used or not by the navigation system. In the same manner as the current sources of navigation measurements (GNSS receiver, air data sensors, inertial navigation system,…), the availability of an additional source of measurements allow to increase the number of inputs in the global architecture of an hybridization algorithm. Current hybridization algorithms are usually done between GPS and baro-inertial measurements.

The angular measurement model proposed in [21] can be proposed and extended. If we consider a forward pointing camera mounted on an aircraft, Figure 17 illustrates the image obtained during a flight when a particular landmark is detected. X and Y are considered as the coordinates of the pixel image.

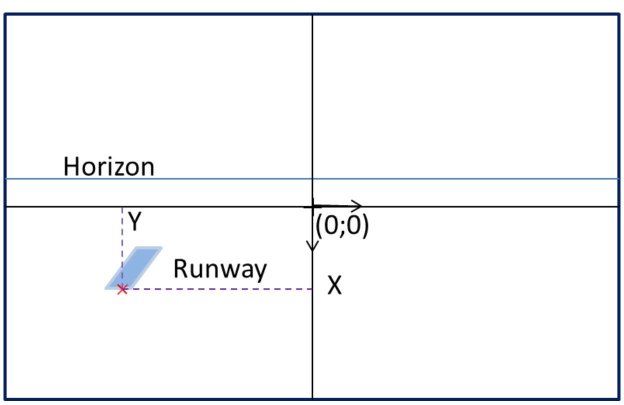


Figure 17 – Image measurement

In this method we consider that the camera is horizontally mounted and in the forward direction (i.e. the normal to the image plan is aligned with the longitudinal axis of the aircraft. The angular measurement corresponding to the point depicted in Figure 17 is computed as follows:

$$\begin{aligned} \alpha_x &= \arctan\left(\frac{x}{f}\right) \\ \alpha_y &= \arctan\left(\frac{y}{f}\right) \end{aligned} \tag{3.1}$$

These angles characterize the angular deviation between the vector normal to the camera plan and the vector pointing toward the detected feature. The two angles are easily computable from the pixel coordinates of the detected point in the image.

Then the angles shall be expressed with respect to the navigation frame, so with respect to the North and East axes. The computation is done from the knowledge of the attitudes of the camera (computed with an inertial system that provides aircraft so camera attitudes). These angles are the parameters that optimally link the position of a point in pixel coordinates in the image with the same point in latitude& longitude coordinates in the outside scenery.

The use of optical measurements in an hybridization filter (for example a Kalman filter) is detailed in papers and thesis of Veth, Giebner, Raquet and Ebcin ([1], [3], [22], [23] and [24]). The interest of having a video system that generates optical measurements is that they can be processed with other filters (KF, SLAM) and combined with other measurements: GPS (as in [19]), INS (as in [1] and [3]).

The three areas identified in the diagram of Figure 15 are defined to be flexible in term of number of features or measurements provided by the video sensor. The choice of optical angular measurements allows generating angles for any number of features. It has also to be noticed that when the aircraft approaches the runway, the image quality increases as well as the ability to more precisely detect the features.

In the first area, the video system can receive the assistance of the pilot or any other means, to select a sub area where the runway is, or to select some features among a too important number of features.

In the second area, the detection is done on the runway axis (the horizon point and the threshold point) and the angle measurements characterize two points that implicitly define glide and localizer angles (like an ILS).

The third area is defined from the moment when the two previous points are not sufficient or no more visible. In this case it is possible to add two additional tracked points, laterally on the left and right side of the runway. In the same manner than the runway threshold and horizon points, these two additional points and their relative position can provide information of alignment with the runway. The optimal solution is to define a precise point of contact (chosen automatically for instance from a data base, or selected by the pilot on the runway). This point will be considered as the ideal landing point. Once the ideal landing point is defined and locked, the detection algorithm selects two lateral points on both sides of the runway (Figure 18 illustrates the configuration with the additional points). When the aircraft is too close to the runway, the aperture of the sensor will not be sufficient to keep tracking of the runway threshold point and the previous metric information is not anymore available (the length of the runway is not available). The new available metric information is now the lateral distance between the points on both sides of the runway (which corresponds to the runway width).

Finally, when the aircraft is on the runway, the horizon point of the runway is no more helpful (because estimation of pitch is not needed) or is no more visible. The configuration of the tracked points can change again and the two tracked points can be the ones on both sides of the runway but with the particularity that their vertical coordinate in the image plan is fixed, so representing 2 points "sliding" together with the aircraft, as long as the aircraft move on along the runway (or representing the fact that the runway side lines virtually contains an infinite number of points). So the goal is now to maintain the same distance between the two side lines. In that way, the configuration is similar to a system which purpose is to follow a line printed on the ground..

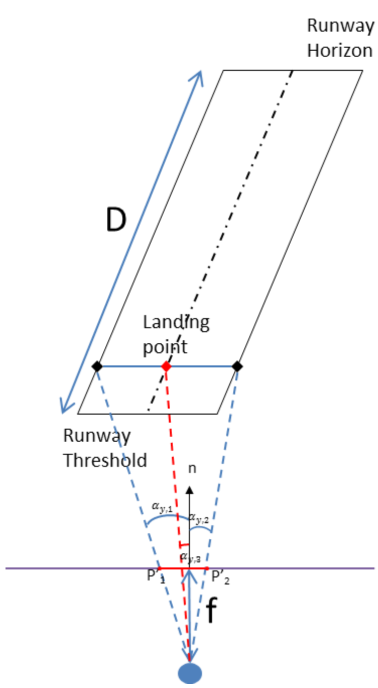


Figure 18 – Upper view during the last part of the approach

This figure presents the configuration of the tracked features when the aircraft is in the third area of coverage (see Figure 15). It illustrates the y-component of the angular measurements.

CONCLUSION

The current paper goes through some vision-based methods of navigation with a presentation of practical and distinct applications.

The first section reviews the various navigation methods that use video sensors. They represent a sampling of techniques using video with different context, constraints, applications and objectives. It shows multiple ways for exploiting image measurements autonomously or coupled with other sensors (GPS, INS, radio-altimeters,…).

The second section presents a personal study and classification of the key elements or characteristics of the video-based techniques, as the determination of a metric scale, the advantages and drawbacks of image registration and a possible image measurement model.

The last section proposes a method of video-based navigation, taking into account the properties and characteristics of a video-based navigation system. This parts aims at providing a first step of feasibility study of an aircraft landing system based on video. It describes a scenario of an approach and landing operation using a forward pointing camera to detect features on the runway and able to estimate the position of the aircraft or deviation with respect to a reference axis. The proposed scenario only deals with providing specific features with a particular configuration that contains indications of the relative position of the aircraft (threshold and vanishing/horizon points of the runway and points located on both sides of the runway). These landmarks can even be associated to the absolute position of the runway through the image measurement model of the features (optical angular measurements) that links position of the aircraft and position of the landmark. The knowledge of a metric scale obtained by tracking the threshold and vanishing/horizon points of the runway (separated by a known length) can also provide a height measurement. The values presented in the third part that define the three areas of the scenario can be refined by linking them with the requirements of approach and landing operations.

From the methods presented in the state of the art, it has been shown that video measurements can help the navigation of an aircraft or improve the estimation of some navigation parameters. Objective of the study was to determine how video system can be used and hybridized with other navigation measurements such as GPS or INS. This paper represents the exploitation of the state of the art and is a first step of feasibility study for an aircraft landing system based on video.

The next step of this work will be to develop a model of video measurements and a model of hybridization filter using several sources of navigation measurements and to perform simulations.

BIBLIOGRAPHY

[1] Capt. M. G. Giebner, "Tightly-Coupled Image-Aided inertial Navigation System via a Kalman Filter", PhD Thesis, 2010, Wright-Patterson Air Force Base, Ohio.

[2] Dr. K. Schertler, "Towards a Robust and Efficient Runway Detection for Image Based Landing", EADS Innovation Works, IPDAYS 2012.

[3] Maj. M. J. Veth, "Fusion of Imaging and Inertial Sensors for Navigation", PhD Dissertation, 2009, Wright-Patterson Air Force Base, Ohio.

[4] S. Segvic, A. Remazeilles, A. Diosi, F. Chaumette, "Large scale vision-based navigation without an accurate global reconstruction", 2007, IEEE Int. Conf. on Computer Vision and Pattern Recognition.

[5] I. Ulrich, I. Nourbakhsh, "Appearance-Based Place Recognition for Topological Localization", IEEE International Conference on Robotics and Automation, San Francisco, CA, April 2000.

[6] J. Gaspar, J. Santos-Victor, "Vision-based Navigation and Environmental Representations with an Omni-directional Camera", IEEE Transaction on Robotics and Automation, vol. 16, N°6, December 2000.

[7] A. Angeli, S. Doncieux, J-A. Meyer, "Visual topological SLAM and global localization", ICRA'09 Proceedings of the 2009 IEEE international conference on Robotics and Automation

[8] F. Fraundorfer, C. Engels, D. Nister, "Topological mapping, localization and navigation using image collections", International Conference on Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ.

[9] O. Bourquardez, F. Chaumette, "Visual servoing of an airplane for auto-landing", Conference on Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International.

[10] T. Gonccalves, J. Azinheira, P. Rives, "Homography-based visual servoing of an aircraft for automatic approach and landing", Conference on Robotics and Automation (ICRA), 2010 IEEE International.

[11] J. R. Azinheira, P. Rives, "Image-Based Visual Servoing for Vanishing Features and Ground Lines Tracking: Application to a UAV Automatic Landing", International Journal of Optomechatronics, 2008.

[12] P. Rives, J. R. Azinheira, "Linear structures following by an airship using vanishing point and horizon line in a visual servoing scheme", International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE.

[13] F. Le Bras, T. Hamel, C. Barat, R. Mahony, "Nonlinear Image-Based Visual Servo controller for automatic landing guidance of a fixed-wing Aircraft", European Control Conference, 2009.

[14] L. Coutard, F. Chaumette, J-M. Pflimlin, "Automatic landing on aircraft carrier by visual servoing", IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'11.

[15] A. Miller, M. Shah, D. Harper, "Landing a UAV on a Runway Using Image Registration", 2008, Proceedings – IEEE International Conference on Robotics and Automation.

[16] H-U. Doehler, B. Korn, "Robust Position Estimation Using Images From an Uncalibrated Camera", The 22nd Digital Avionics Systems Conference, 2003. DASC '03.

[17] B. Sinopoli, M. Micheli, G. Donato, T. J. Koo, "Vision-based Navigation for Unmanned Aerial Vehicle", 2001, procedings of the IEEE international conference on robotics and automation.

[18] S. Riisgaard, M. R. Blas, "SLAM for Dummies - A Tutorial Approach to Simultaneous Localization and Mapping", 2005.

[19] L. G. B. Mirisola, J. Dias, "Exploiting Attitude Sensing in Vision-Based Navigation for an Airship", 2009, Journal of Robotics.

[20] H. Huang, "Bearing-only SLAM - A Vision-Based Navigation System for Autonomous Robots", 2008, PhD Thesis.

[21] J. Raquet, Capt M. G. Giebner, "Navigation Using Optical Measurements of Objects at Unknown Locations", ION 59th Annual Meeting/CIGTF 22nd Guidance, June 2003.

[22] Maj. M. J. Veth, J. Raquet, "Two-dimensional Stochastic Projections for Tight Integration of Optical and Inertial Sensors for Navigation", ION NTM 2006, 18-20 January 2006, Monterey CA.

[23] Maj. M. J. Veth, J. Raquet , "Fusion of Imaging and Inertial Sensors for Navigation", ION GNSS 19th International Technical Meeting of the Satellite Division, 26-29 September 2006, Fort Woth, TX.

[24] S. Ebcin, Maj. M. J. Veth, "Tightly-Coupled Image-Aided Inertial Navigation Using the Unscented Kalman Filter", ION GNSS 20th International Technical Meeting of the Satellite Division, 25-28, September 2007, Fort Worth, TX.