Omnidirectional Vision applied to Unmanned Aerial Vehicles UAVs attitude and heading estimation

Iván F. Mondragón[*,a], Pascual Campoy[a], Carol Martinez[a], Miguel Olivares[a]

[a]*Computer Vision Group, Universidad Politécnica de Madrid, C. José Gutiérrez Abascal 2. 28006 Madrid, Spain*

Abstract

This paper presents an aircraft attitude and heading estimator using catadioptric images as a principal sensor for UAV or as a redundant system for IMU (Inertial Measure Unit) and gyros sensors. First, we explain how the unified theory for central catadioptric cameras is used for attitude and heading estimation, explaining how the skyline is projected on the catadioptric image and how it is segmented and used to calculate the UAV's attitude. Then, we use appearance images to obtain a visual compass, and we calculate the relative rotation and heading of the aerial vehicle. Finally the tests and results using the UAV COLIBRI platform and the validation of them in real flights are presented, comparing the estimated data with the inertial values measured onboard.

Key words: Omnidirectional Images, Catadioptric systems, Unmanned Aerial Vehicles (UAV), Sky segmentation.

1. Introduction

Omnidirectional vision can be a useful sensor in estimating attitude in UAV. It can be used as the only sensor or as complementary sensor for inertial and GPS information. Its main advantage is that pitch and roll are estimated accordingly to the horizon line and therefore it regards the land level variations, as well as yaw is estimated accordingly to visual objects that

---

[*]Corresponding author

   Email address: `imondragon@etsii.upm.es` (Iván F. Mondragón)

   URL: `www.disam.upm.es/colibri` (Iván F. Mondragón)

can be used for trajectory planning and see & avoid strategies.

In general, there are twelve variables that define the UAV state [1], which are inertial latitude, longitude and altitude of the UAV, body frame velocities (rates); and roll, pitch and yaw angles and rates. In order to make a low level controller for UAVs, it is necessary to have a direct measurement of the attitude angles (Roll, Pitch and Yaw) or their rates. Frequently, these variables are measured using rate gyros as a part of a more complex sensor (the Inertial Measurement Unit (IMU)) that involves gyroscopes and accelerometers to estimate the relative position, velocity and acceleration of a vehicle in motion. There has been an active improvement in precision, cost and size of this sensor in the last years from mechanical devices through optical systems to MEMs sensors. Gyroscopes and IMU units are extremely sensitive to measurements errors caused by drift, generating wrong estimation in orientation after a long operation period, making it necessary to reference them to an external measurement system like GPS framework [2]. Also, they still can be affected by structural and mechanical fatigue, vibrations, temperature changes, electric interferences and others [3], [4] causing a erroneous data or a sensor failure. In addition, small and micro UAV sometimes have restrictions in cost and sensors payload capacity restricting the use of other sensors.

The idea of using visual information for UAV attitude estimation is not new. The first experiments attempted to obtain the skyline from images taken by a single perspective camera looking forward on the aircraft, using this to estimate the roll angle with a horizontal reference [5], [6],[7],[8]. These works differ in the way that they segment the sky and ground and in how they estimate the horizon line.

Omnidirectional vision has also been used for UAV control and attitude estimation. Hrabar [9] use an omnidirectional system for sideways-looking sensing on an autonomous helicopter by applying image unwrapping. Demonceaux *et al.*[10], use a similar approach to the one presented in this paper, showing the advantages of using omnidirectional rather than perspective images for attitude estimation. They detect the horizon line on the catadioptric image using a Markov Random Fields segmentation and then project it on the equivalent sphere projection model for a Catadioptric system, showing a good performance (without a ground true validation), on an off-line processed video sequence.

Catadioptric systems also have been used for robot odometry and relative orientation on outdoor vehicles [11], [12] showing that it is possible to estimate relative orientation and position of a mobile robot using appearance panoramic images.

In this work, we propose the use of an omnidirectional vision system as an attitude sensor for simultaneous Roll, Pitch and relative heading or Yaw angle measurement. Section 2 shows the UAV platform employed for tests. Section 3 shows the catadioptric system employed and the equivalent projection model. Then, in section 4 an image processing algorithm to isolate the skyline from the catadioptric image is implemented with the aim to estimate the attitude of the bodyframe. Relative yaw or heading is also obtained using the so called visual compass on appearance images. Finally, in section 5 we present the feasibility and reliability of this approach when the obtained results are compared with the Inertial Measurement Unit (IMU) of the UAV testbed employed.

2. UAV System Description

The Colibri project has three totally operational UAV platforms: One electric helicopter with a 1400 W motor (Figure 1), and two gasoline powered industrial twin helicopter with an 52 cc engine and 8 hp capable to carry up to 12 kg payload. The COLIBRI testbeds [13], are equipped with a Xscale-based flight computer augmented with sensors (GPS, IMU, Magnetometer, etc fused with a Kalman filter for state estimation). Additionally they include a Pan and Tilt servo controlled platform for many different cameras and sensors. In order to enable it to perform vision processing, it also has a VIA mini-ITX 1.5 GHz onboard computer with 2 GB RAM, a wireless interface and support for many type of Firewire cameras including Mono (BW), RAW Bayer, color, and stereo heads for images acquisition. It is possible to use IP cameras and analog cameras as well.

The system runs in a client-server architecture using TCP/UDP messages. The computers run Linux OS working in a multi-client wireless 802.11g ad-hoc network, allowing the integration of vision system and visual tasks with flight control. This architecture allows embedded applications to run onboard the autonomous helicopter while it interacts with external processes

3

Figure 1: UPM-COLIBRI III Electric Helicopter platform during an Omnidirectional attitude and heading estimation tests using a catadioptric camera

through a high level switching layer. The visual control system and additional external processes are integrated with the flight control through this layer using TCP/UDP messages. The layer is based on a communications API where all the messages and data types are defined. The helicopter's low-level controller is based on simple PID control loops to ensure its stability. The higher level controller uses various sensing mechanisms such as GPS and/or vision to perform tasks such as navigation, landing, visual tracking, etc.

3. Central Catadioptric Cameras

Catadioptric cameras are devices that combine reflective elements (catoptric) and refractive systems (dioptric) to form a projection onto the image plane of the camera. They can be classified as central and non-central catadioptric cameras according to the single effective viewpoint criteria. Baker and Nayar [14], [15], define the configurations that satisfy the constraints of a single viewpoint, finding that a central catadioptric system can be built combining a perspective camera with a hyperbolic, elliptical or planar mirror, or using an orthographic camera with a parabolic mirror.

Geyer and Daniilidis [16], [17] proposed an unified model for the projective geometry induced by central catadioptric systems, showing that these projections are isomorphic to a projective mapping from a sphere (centered

on the effective viewpoint) to a plane with the projection center on the perpendicular axis to the plane.

A modified version of this unified model is presented by Barreto and Araujo in [18], [19], where the mapping between points in the 3D world and points in the catadioptric image plane is split into three steps. First, a linear function maps the world into an oriented projective plane. Then a non-linear function transforms points between two oriented projective planes. Finally there is a collineation function depending on the mirror parameters and the camera calibration matrix (intrinsic parameters). Figure 2 shows the general unit sphere projection for modeling catadioptric systems.
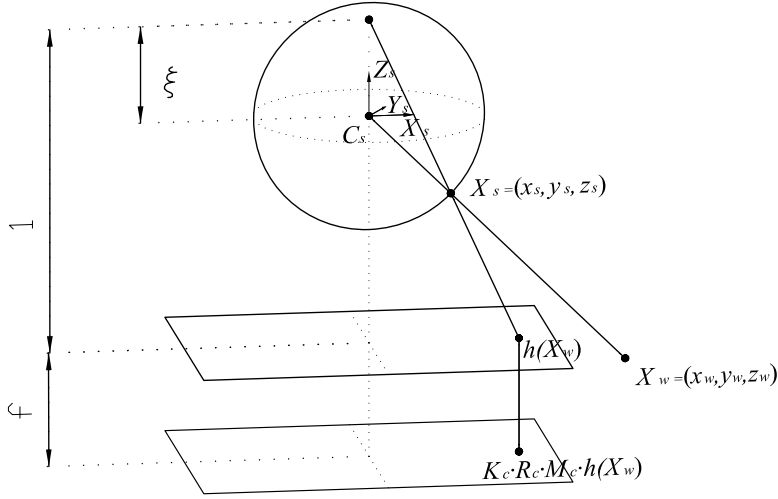

Figure 2: Catadioptric projection modelled by the unit sphere.

Consider a point in space (visible to the catadioptric system), with Cartesian coordinates $X_w = (x_w, y_w, z_w)^T$ in the catadioptric reference (focus). This point is mapped onto point $X_s = (x_s, y_s, z_s)^T$ on the unitary sphere centered on the effective view point by equation 1.

$$X_s = \frac{X_w}{\sqrt{x_w^2 + y_w^2 + z_w^2}} = \frac{X_w}{\|X_w\|} \tag{1}$$

5

Table 1: Parameters ξ and ψ for central catadioptric systems (d is distance between focus and $4p$ is the Lactus Rectum

	Parabolic	Hyperbolic	Elliptical	Planar
ξ	1	$\frac{d}{\sqrt{d^2+4p^2}}$	$\frac{d}{\sqrt{d^2+4p^2}}$	0
ψ	$1+2p$	$\frac{d+2p}{\sqrt{d^2+4p^2}}$	$\frac{d-2p}{\sqrt{d^2+4p^2}}$	1

To each projective point X_s, corresponds a projective point $X_c = (x_c, y_c, z_c)^T$ in a coordinate system with origin at the camera projection center. This projection is a non-linear mapping between two projective planes and is defined by equation 2.

$$X_c = (x_c, y_c, z_c)^T = M_c \cdot \hbar(X_w)$$

where

$$M_c = \begin{bmatrix} \psi - \xi & 0 & 0 \\ 0 & \xi - \psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad (2)$$

$$\hbar(X_w) = \left(x_w, y_w, z_w + \xi\sqrt{x_w^2 + y_w^2 + z_w^2} \right)^t$$

where the Matrix M_c depends on the mirror parameters ξ and ψ, defined for each one of the central catadioptric projections, as is shown in Table 1.

Finally, the image in the catadioptric plane is obtained after a collineation between the image and the Projective plane depending of the camera's intrinsic parameters K_c (where m_x and m_y are the pixels per unit distance in image coordinates, f is the focal distance and (x_0, y_0) are the coordinates of the principal point), and the rotation of the camera R_c. The projection of a world point on the catadioptric image is defined by equation 3.

$$H_c = K_c \cdot R_c \cdot M_c$$
$$X_i = H_c \cdot \hbar(X_w)$$
$$K_c = \begin{bmatrix} fm_x & s & x_0 \\ 0 & fm_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \qquad (3)$$

6

$\hbar(X_w)$ is a homogenous positive injective fucytion, with an inverse defined by $\hbar^{-1}(X_w)$. This function maps points in a projective plane onto the unitary sphere. The non-linear inverse function is defined by

$$(x_s, y_s, z_s)^t = \hbar^{-1}(H_c^{-1}X_i) = (\lambda_c x_c, \lambda_c y_c, \lambda_c z_c - \xi)^t$$

where

$$\lambda_c = \frac{z_c \xi + \sqrt{z_c^2 + (1 - \xi^2)(x_c^2 + y_c^2)}}{x_c^2 + y_c^2 + z_c^2}$$

$$(4)$$

3.1. Catadioptric system

Svoboda, Pajdla and Hlaváč [20], [21], developed a projection model for a perspective camera with a hyperbolic mirror and gave a general approach to design and construct a useful catadioptric system for mobile robots. They also explain the appropriated assembly and alignment of camera and mirror in order to obtain the adequate images with central projection center. This approach assumes a hyperbolic shaped mirror centered in one of its focal points F' which general function is defined by equation 5:

$$\left(\frac{z + \sqrt{a^2 + b^2}}{a} \right)^2 - \left(\frac{x}{b} \right)^2 - \left(\frac{y}{b} \right)^2 = 1$$

$$(5)$$

Where a,b are the mirror parameters. The mirror eccentricity is $e = \sqrt{a^2 + b^2}$ and the distance between focus is defined as $d = 2e$.
The maximum view angle taking into account the dimension of the mirror and the distance to the camera is $\alpha = \frac{\pi}{2} + \arctan\left(\frac{h - 2e}{r_{top}} \right)$, where r_{top} is the radius of any point on the mirror top rim ($r_{top}^2 = x_{top}^2 + y_{top}^2$), and h is the distance between the the top of the mirror and the camera center.

The employed catadioptric system is based on the hyperbolic shape proposed by Okamoto *et. al* [22], designed using the Svoboda method [20]. It has a height from the top mirror to camera center of $h = 100mm$ and the radius of the mirror is set to $r_{rim} = 20mm$. A standard firewire 30 fps, CCD camera with resolution of 640x480 pixels is selected and looking for the maximum occupancy on the image, the projected radius is defined to $r_{pix_{rim}} = 240$ pixels. Assuming a pixel size of $0.01mm$, a lens with focal distance of $f = 12mm$ has to be used. The mirror shape parameters are $b = 100\sqrt{4 + 1} - 2\sqrt{100^2 + 20^2} = 19.6461$ and $a = 2b = 39.2922$, with a

maximum vertical angle of view of $\alpha = 121.3$ degrees. Figure 3 shows the camera mirror distribution. A transparent extruded thermoplastic acrylic tube is used to make the final assembly. Finally, the catadioptric camera system is calibrated using the omnidirectional camera calibration toolbox developed by Mei and Rives [23].

Figure 3: Catadioptric system design and assembly using a transparent extruded thermoplastic acrylic tube.

4. Omnidirectional Image Processing

To be able to measure the body frame attitude based on the catadioptric image, it is necessary to know, how the skyline is projected onto the unitary sphere and onto the catadioptric plane. Geyer and Daniilidis [17] demonstrate that a line on the space projects as a great circle on the unitary sphere and as a conic on the image plane. Later, Ying and Hu [24] demonstrate that the occluding contour of a sphere in space is projected onto a circle on the unit sphere or onto a conic in the catadioptric image plane. Considering the skyline the occluding contour on the earth sphere surface, finding it requires to look for a small circle on the unitary sphere model or for a conic or ellipse on the image plane as proposed by Demonceaux *et al.* [10], (see figure 4 for an illustration).

Because the original datum obtained is the image projection, the skyline detection focused on isolating the sky from the ground in this image and then estimating the best adjusted ellipse to the skyline. To isolate the sky from the ground we use an approach based on the method employed by Cornall *et al.* [7] in which the RGB components of each pixel are weighted using the function $f(RGB) = 3B^2/(R + G + B)$. This function has shown very good

8

Figure 4: Skyline is the occluding contour of the earth sphere surface, whose projection on the equivalent unitary sphere model through a plane that intersect it forms the depicted red circle.

results for sky-ground segmentation under different light and cloud conditions. For each resulting grayscale image from function $f(RGB)$ a Pyramid Segmentation [25] followed by a Gaussian Adaptive Threshold function [26] is used to obtain a sky-ground binary image. The pyramid segmentation allows us to reduce the effects of sunlight or brightness and clouds shadows in the image under variable lighting conditions. This threshold method is very fast, and produces good results in real-time. Figure 5 shows the segmentation on a catadioptric image weighted using the method described above.

Once we have a sky-ground thresholded image, the ground contour on the image can be easily defined. This contour represents the skyline and is used by a fitting function to obtain the ellipse with the best approximation to the contour. As can be seen on figure 5(c), if there are some high objects like mountains, buildings and trees that not totally occlude the horizon, the segmentation method continues working, because the approximate ellipse is obtained using all the segmented contour and not only parts of the skyline. Figure 6 shows some examples of the best fitted ellipse of the skyline on original catadioptrics images obtained during a UAV flights under different
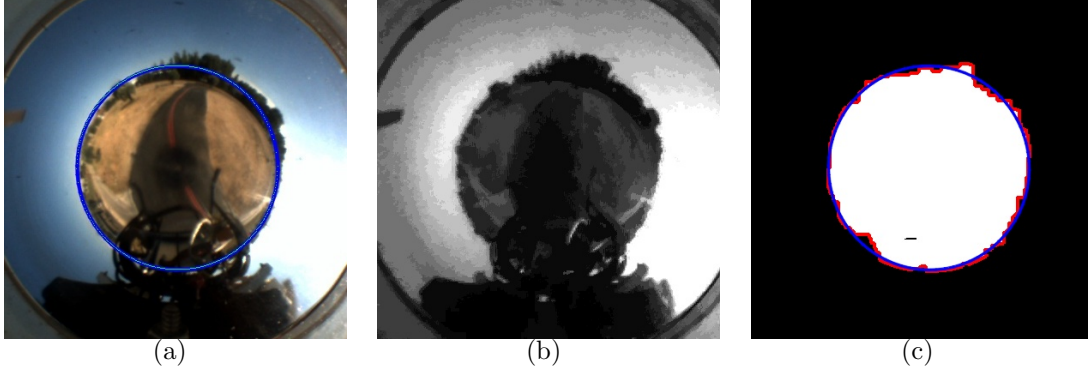
9

(a)　　　　　　　　(b)　　　　　　　　(c)

Figure 5: Segmentation on a catadioptric image (5(a)), weighted using the function $f(RGB) = 3B^2/(R + G + B)$ (5(b)), and the obtained sky ground binary image (5(c)).

sunlight and weather conditions.

4.1. Skyline backprojection on sphere.

The segmented skyline is defined by the points of the contour that represent the ground border or by the adjusted ellipse points $SKY_{img} = (x_{SKY_{img}}, y_{SKY_{img}}, 1)$. These points are backprojected onto the unitary sphere using equation 4 obtaining $SKY_s = (x_{SKY_s}, y_{SKY_s}, z_{SKY_s})$ as shown in figure 7. The circle formed by the skyline points on the sphere forms a plane that intersects with the unitary sphere. To obtain the parameters of the skyline in the unitary sphere, it is sufficient to find the plane with normal equation $N_x x_{SKY_s} + N_y y_{SKY_s} + N_z z_{SKY_s} + D = 0$ that best adjusts the backprojected points of the skyline contour or the adjusted ellipse on the image plane.

For each point of the backprojected skyline, the normal equation of the plane is obtained by $z_{SKY_s}^i = N_x x_{SKY_s}^i + N_y y_{SKY_s}^i + D$ with $i = 1, ..., n$ and an overdetermined linear system of the form $(Ax = b)$ is solved using the *pseudo-inverse method* to obtain the plane $\pi_{sky} = (N_x, N_y, 1, D)$ (Equation 6).
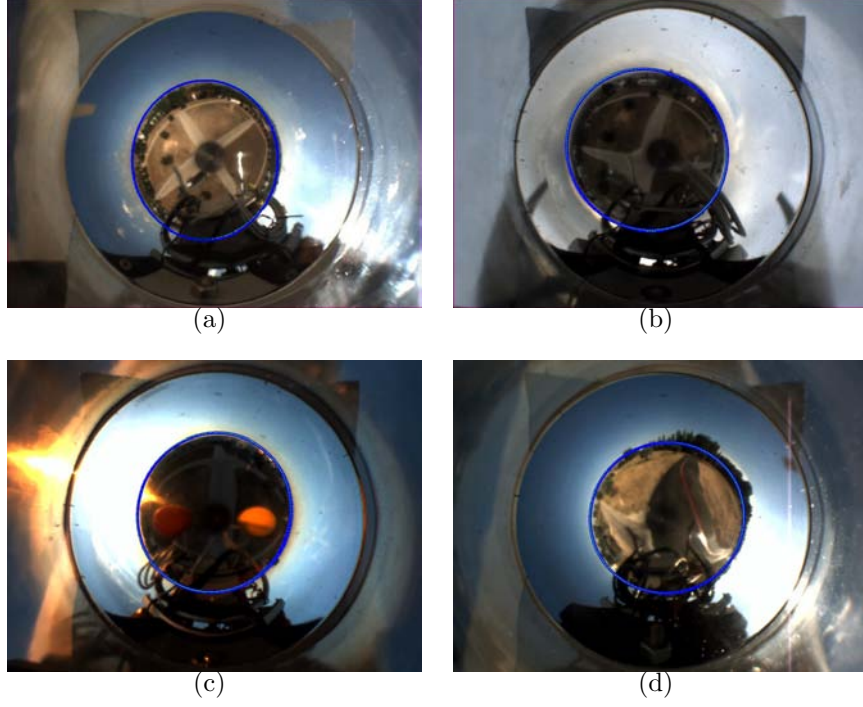
10

(a)

(b)

(c)

(d)

Figure 6: The best fitted ellipse (blue) to the skyline on original catadioptric images obtained during a UAV flight under different sunlight and weather conditions . Figure 6(a) is a flight under a fall season clear sky, figure 6(b) is a winter cloudy day, figure 6(c) is on a winter partly cloudy sunset, and 6(d) is a takeoff and flight in a summer sunny day.

$$[N_x, N_y, D]^t = \arg\min_x \|Ax - b\|$$

where

$$A = \begin{bmatrix} x^i_{SKY_s} & y^i_{SKY_s} & 1 \\ \vdots & \vdots & \vdots \\ x^n_{SKY_s} & y^n_{SKY_s} & 1 \end{bmatrix}$$

$$x = \begin{bmatrix} N_x \\ N_y \\ D \end{bmatrix}$$

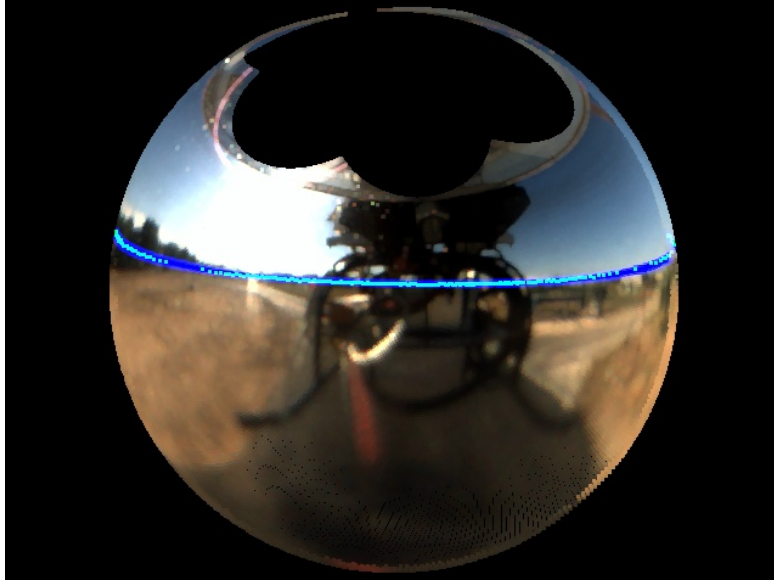$$b = \begin{bmatrix} z^i_{SKY_s} \\ \vdots \\ z^n_{SKY_s} \end{bmatrix}$$

(6)

11

Figure 7: The best fitted ellipse (blue) to the skyline is backprojected on a unitary sphere model, forming a plane that intersects the sphere (which forms a small circle). The normal vector to this plane defines the attitude of the camera and the UAV

4.2. Pitch And Roll Estimation

The normal vector to the plane formed by the skyline and the unitary sphere is defined as $N = [N_x, N_y, 1]^t$. Assuming that the camera frame is aligned with the UAV frame so that the x axis is the heading of the UAV and the y axis is aligned with the UAV wing, it is possible to obtain the desired *roll* (ϕ) and *pitch* (θ) angles, using equation 7.

$$\theta = \arccos\left(\frac{N_x}{\sqrt{N_x^2 + N_y^2 + 1}}\right)$$
$$\phi = \arccos\left(\frac{N_y}{\sqrt{N_x^2 + N_y^2 + 1}}\right) \tag{7}$$

4.3. Yaw estimation using a visual compass

The relative heading of the UAV is calculated using the so called visual compass on appearance images. This method was used by Labrosse [11] and later by Scaramuza [12]. It consists of a part of a panoramic image obtained

12

from a catadioptric image using a polar to Cartesian coordinates change or unwrapping process employing equation 8

$$I(\alpha, R) = I_c(R\cos(\alpha) + u_0, R\sin(\alpha) + v_0) \qquad (8)$$

where (u_0, v_0) are the coordinates of the catadioptric image center, α is a linear function with maximum range $[0, 2\pi]$ and R is a linear function that scans along the image Radius. The steps and range for α and R are defined according to the desired panoramic image resolution and size. The unwrapping process excludes the parts of the catadioptric images that involves the camera reflex and the outside of the mirror.

If the catadioptric image corresponds to a scene captured with an almost perfect vertical camera to the ground plane, then pure rotation will appear on the appearance image as a pure pixel column-wise shift. The relative rotation between two consecutive images is obtained, by finding the best match based on the images' column shift using the Euclidean distance. Equation 9 shows the Euclidean distance between two panoramic images I_m and I_n with the same size and space color, as a function of the column-wise shift on the Image I_n by α pixels (horizontal rotation). Figure 8 shows two consecutive appearance images obtained by a unwrapping process with a small rotation.

$$d(I_m, I_n, \alpha) = \sqrt{\sum_{i=1}^{width} \sum_{j=1}^{height} \sum_{k=1}^{Nchan} (I_m(i,j,k) - I_n(i+\alpha, j, k))^2} \qquad (9)$$

The best shift α_{min} that minimize the distance function $d(I_m, I_n, \alpha_{min}) \leq d(I_m, I_n, \alpha) \forall \alpha \in \mathbb{R}$ is the best pixel rotation between these two images. Figure 9 shows the Euclidean distance between the two appearance images on figure 8 as a function of a column-wise shift on the second image.

The rotation angle or *yaw* ψ between images is directly related to the obtained column shift between images, considering only the angular resolution of the appearance images defined by the images field of view FOV and the images width as shown in equation 10.

$$\psi_{(I_m, I_n)} = \alpha_{min} \frac{FOV}{imgWidth} \qquad (10)$$

To obtain the final rotation relative to the first image, it is necessary to add the obtained value to a counter.
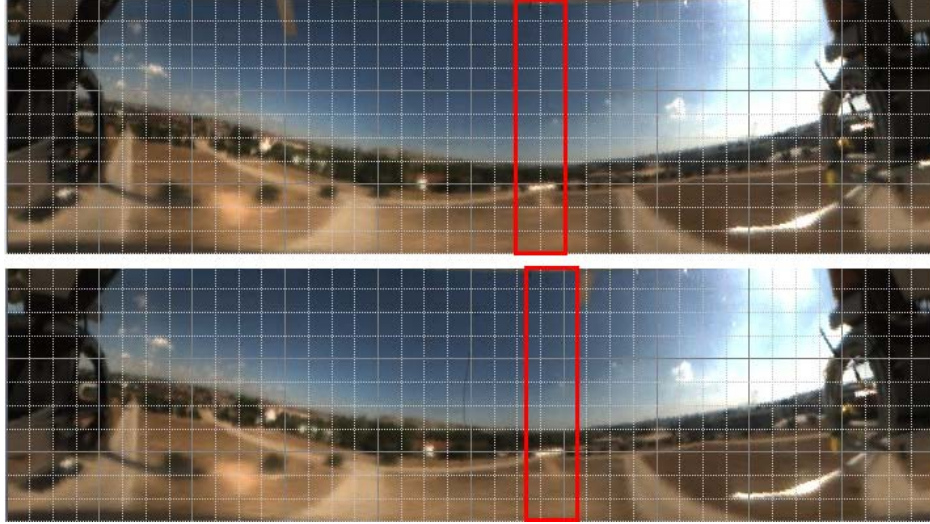
Figure 8: Two consecutive appearance images with a small rotation between them. A white grid is superimposed to reference and the red box shows clearly the column shift between images

The method described above was developed under the assumption of a pure camera rotation on its vertical axis which is perpendicular to the horizontal plane. In the general case, the UAV has translational components and roll and pitch variations, causing the camera vertical axis to not always be perpendicular to the horizon. However, as shown by Labrosse [11] and by Scaramuza [12] the visual compass method based on appearance images is still valid under translation and attitude variations if the camera has small displacements or the distance to the objects is large compared with the displacement. Because images are captured at high frequency, small variations of pitch and roll are present between consecutive images; therefore, the pure rotation assumption is still valid. Finally, because the translational movements contribution to the optical flow is not homogeneous on the omnidirectional image, but the rotation movements contributes equally on all image, is enough to use the regions of the images on which the translation contribution to optical flow is reduced. Considering that a forward/backward translation has more contribution to the optical flow on the image regions corresponding to the sides of the vehicles and little in from and back regions of the vehicle, working only with the from and/or back portion of the image is enough to reduce the effects of a translation on the image rotation estimation. In our
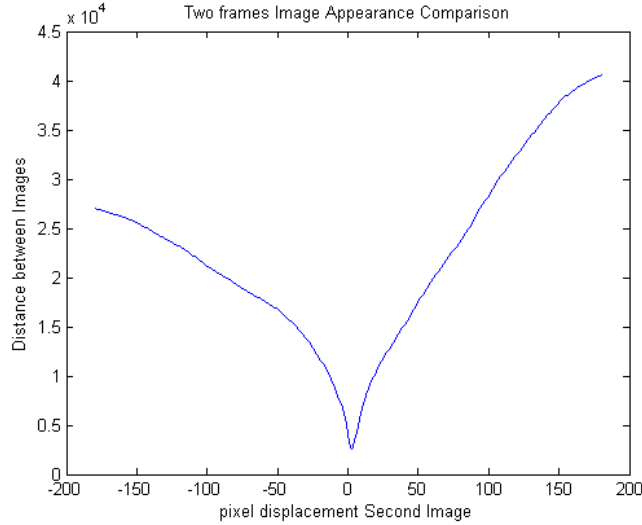
Figure 9: Euclidean Distance between the two appearance images on figure 8 as function of the column-wise shift on the second image

implementation we only use a FOV of 120 degrees corresponding to the front central part of the panoramic image.

5. Tests and Results

Several tests have been made using the Colibri testbeds [27] in different seasons, weather and illumination conditions. Table 2 describes the flight and weather conditions in which these tests have been done. In these tests, a series of flights were performed in both autonomous and manual modes. In autonomous mode, the helicopter can take both, a previously defined trajectory, or a hover (stationary) flight. In manual mode, a safety pilot takes a free flight with strong movements of the helicopter. The algorithm is tested during these flights (including a takeoff process) and an image sequence is stored, associating to each of the processed images the UAV attitude information estimated by the omnidirectional system. Also, a flightlog is created with the GPS position, IMU data (heading, body frame angles and displacement velocities), and the helicopter position estimated by the Kalman Filter of the controller on the local plane with reference to the takeoff point. These values are used for later comparisons with the estimated data using the catadioptric system.

15

Table 2: Flight test description and weather conditions

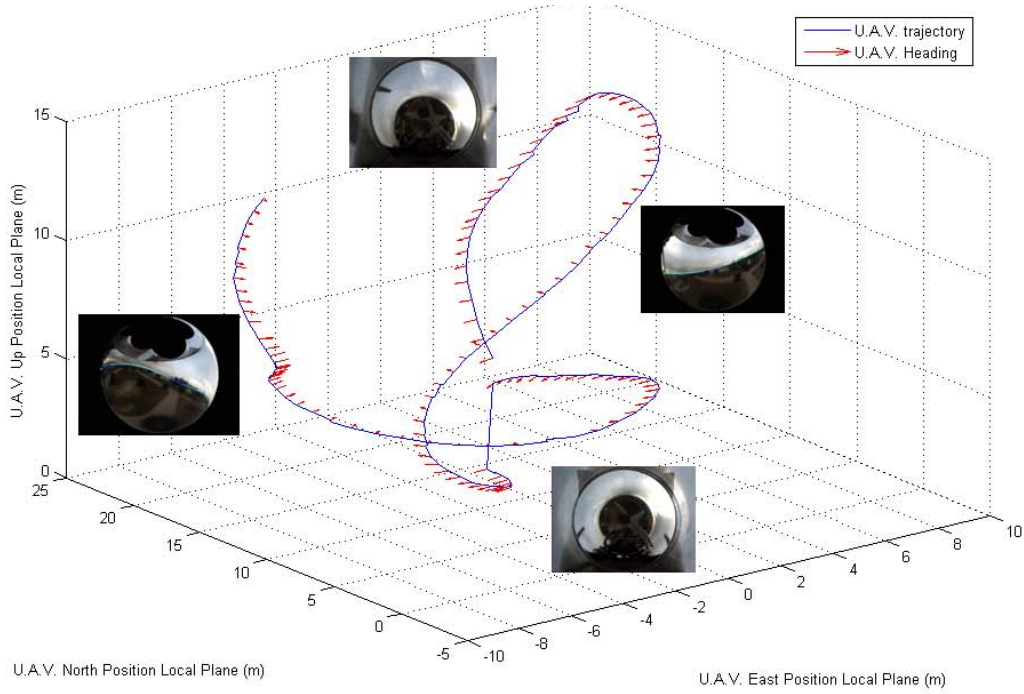| Flight	Season	Weather & Lighting	Flight Mode
Flt. 1	autumn	Clear and windy day, 19°C	Auto Hover, high level
Flt. 2	winter	Cloudy Sky 2°C	Manual mode, high level
Flt. 3	winter	Partly Cloudy, Sunset 8°C	Manual mode, high level
Flt. 4	summer	Sunny Sky 37°C	Manual mode, takeoff-low level

Figure 10: Flight number 2, 3D trajectory and UAV heading reconstruction, obtained using the flightlog data. The blue line depicts the translational movement and the red arrows represent the heading direction of the UAV (pitch and yaw angles). Superimposed Images shows some of the different catadioptric and sphere result obtained during the flight sequence. In this flight, every five hundred frames (approximately 15 seg), the algorithm makes a small break to store the acquired and processed images and IMU data in the on board hard disk, causing this strong change in the stored trajectory values

Using the flightlog it is possible to reconstruct the 3D trajectory of the vehicle and the camera and/or helicopter attitude. Figure 10 shows a 3D reconstruction of one flight made in manual mode in which is possible to see
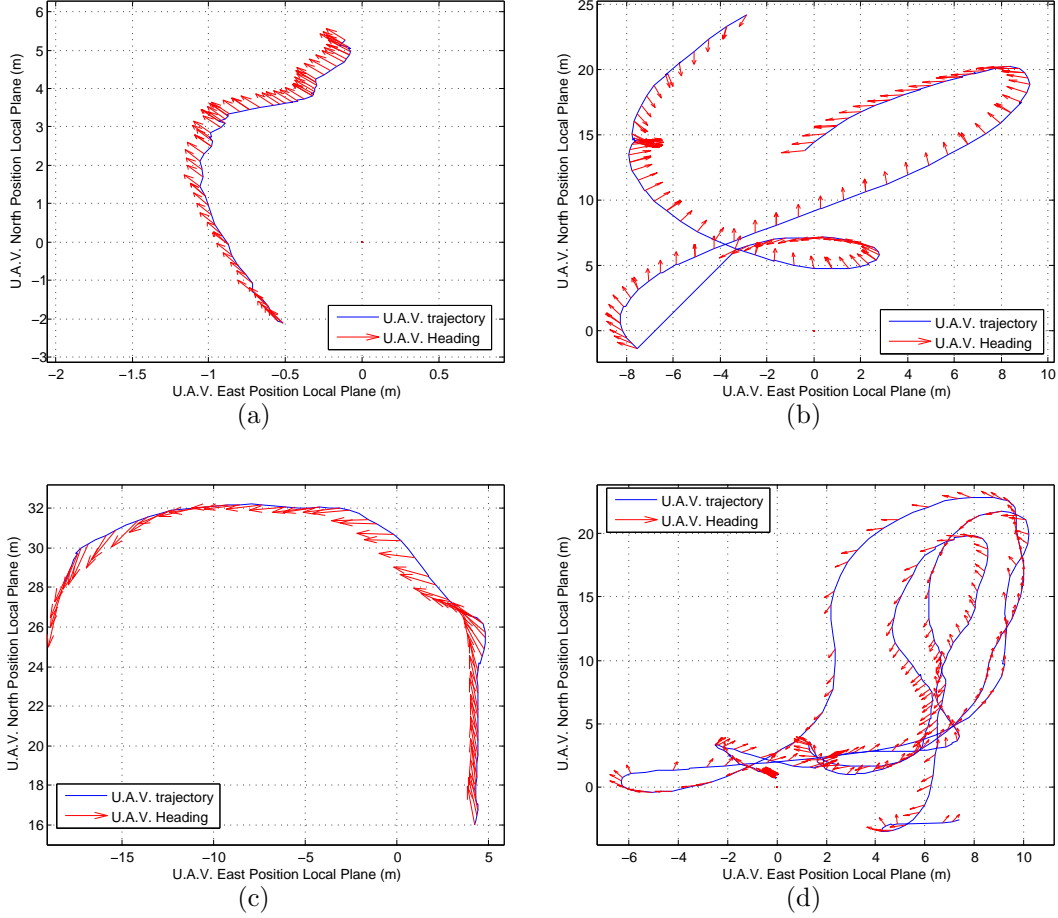
Figure 11: 2D trajectories and UAV heading reconstruction. Figure 11(a) is a high level hover flight in autonomous mode, figure 11(b) is a manual high level flight, figure 11(c) is manual high level during a sunset and 11(d) is a manual takeoff and low level flight.

big changes in the attitude and orientation of the UAV. Figure 11 shows the corresponding 2D reconstructions of flight tests described on table 2, showing for each one, the trajectory and helicopter heading.

The algorithm developed, estimates the absolute Roll and Pitch angles of the Camera bodyframe on the UAV and the relative Yaw rotation to the first image. For these tests, the camera is located on the Pan and Tilt platform of the helicopter in such a way that the vertical axes of the camera

Table 3: RMSE for estimated *roll* (ϕ), *pitch* (θ) and *Relative Yaw* (ψ)

	roll (ϕ)	pitch (θ)	R. Yaw (ψ)
Flight 1	0.1807	4.3710	1.2533
Flight 2	2.7528	2.8267	10.7844
Flight 3	0.9363	4.0918	6.0413
Flight 4	2.8647	1.9183	5.0086

and helicopter are parallel (by adjusting the camera platform tilt). In this way, the hyperbolic mirror faces downward, and the camera looks up. This positioning ensures that all axes of the helicopter and camera are coincident, so that the obtained roll and pitch angles for the camera, are the same for the helicopter frame, as shown in figure 4.

The estimated values of roll, pitch and yaw from test flights (Table 2, Fig. 11) are compared with the corresponding stored IMU values. Figures 12, 13 and 14 shows these results and table 3 has the Root Mean Squared Error (RMSE) of the estimated values compared with the IMU values as ground truth. In flight 2 (Fig. 10), every five hundred frames (15 seg), the algorithm made a small break to store the acquired and processed images and IMU data in the on board hard disk (other tests, the process is continuous), causing this strong change in the measured values, because there are not processed images during this saving time.

The estimated roll values have a very similar behavior to the one performed by the helicopter during the tests. The estimated values are very close to the IMU values and have a small RMSE against absolute values measured in the IMU.

Pitch values are also estimated accurately compared with the ground truth IMU. However when the helicopter has a high nose-up angle, a portion of the ground is occluded on the catadioptric image by the platform's structure and the UAV's reflections, causing a small error in the adjustment of the skyline on the equivalent sphere projection, and the pitch estimation. Additionally, this causes that in general, the pitch RMSE has higher values than the Roll, although these values are still a high-quality measurement. The error caused by the ground occlusion can be solved changing the camera position to be totally below the UAV bodyframe reducing the portion of the
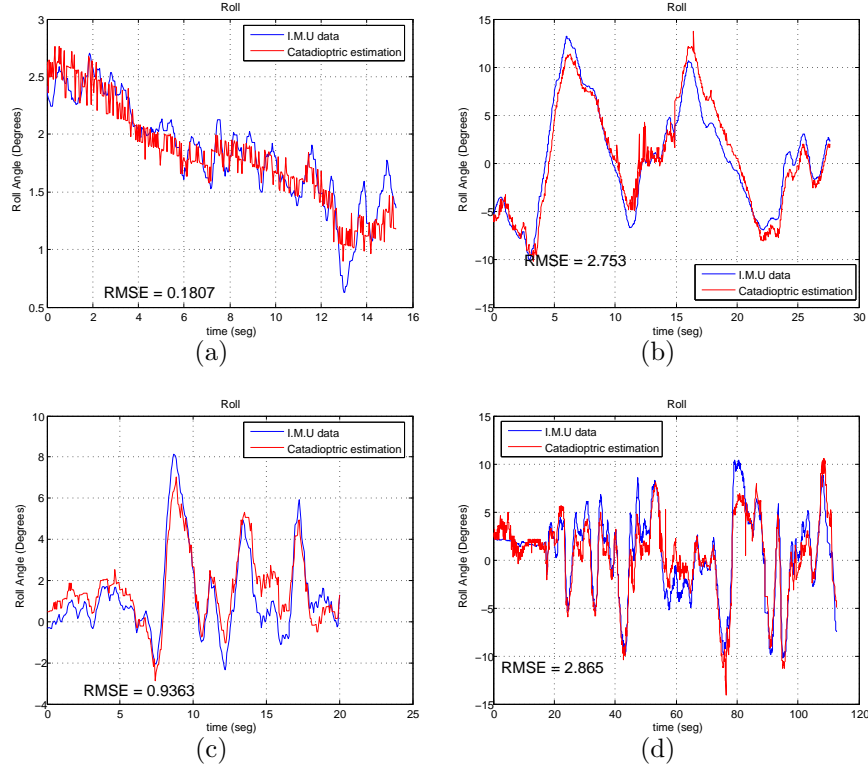
Figure 12: Estimated Roll compared with the Roll Angle measured by IMU (Roll angle in degrees $-180 < \phi < 180$. Zero is level with the horizon and increasing is right wing down.). Figures 12(a) to 12(a) corresponds to the flights shown in figure 11.

camera platform reflected on the mirror.

Yaw estimation uses the first image taken by the algorithm as a reference, calculating rotation with respect to this reference image. Absolute Yaw data measured by the IMU is rotated according to the first image angle and changed to a range between $0 < \psi < 360$ for easy comparison with omnidirectional data. Results show that the rotation between frames, as well as the total rotation, are both a good approximation to real values, however there are some cases in which the rotation direction is correctly estimated, but the magnitude is not correct, as can be seen on Fig. 14(b). This happens, because Yaw estimation is obtained using a counter of rotations between consecutive images, causing an accumulative error along all the estimation process. In
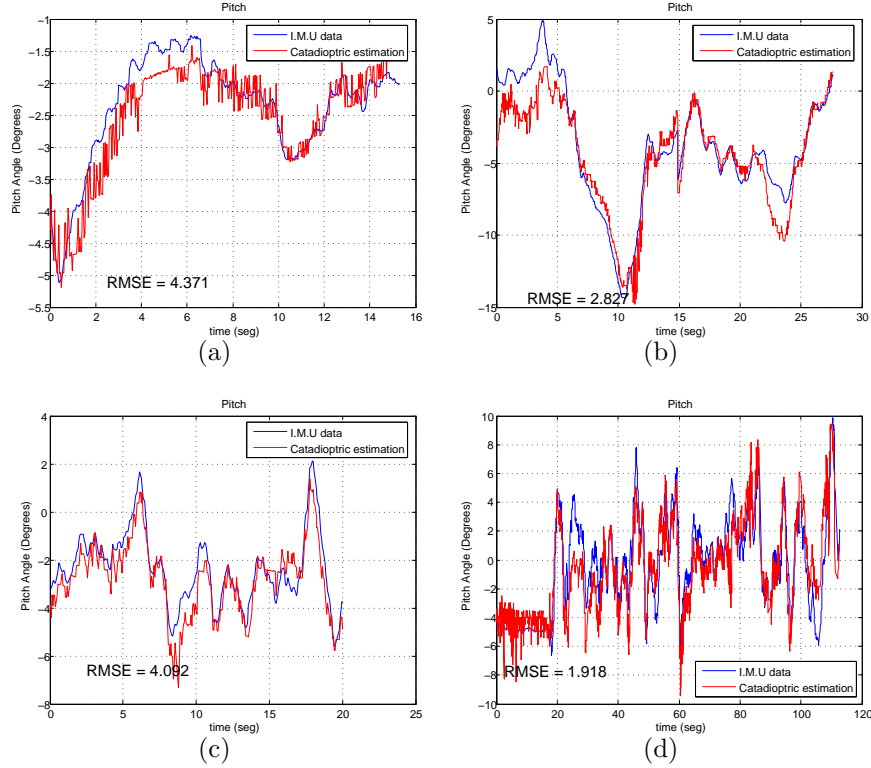
19

Figure 13: Estimated Pitch compared with the Pitch Angle measured by IMU (Pitch angle in degrees $-90 < \theta < 90$. Zero is level with the horizon and increasing is nose up.). Figures 13(a) to 13(d) corresponds to the flights shown in figure 11.

addition, the unwrapped panoramic image only has a resolution of 0.5 pixels per degree, causing that small rotations between consecutive images cannot be detected and not added to the counter. Roll and Pitch estimation do not have this accumulative error problem, because they are calculated using only the information available in the last catadioptric image. Currently, it is under test a new method to reduce the Yaw estimation error, by increasing the resolution of the unwrapped image and using a Kalman filter in order to improve the estimation robustness, however there are not available results yet.

In Flight 4 (Figure 11(d)) the algorithm was running from the helicopter's takeoff phase in order to test the robustness of the skyline detection at low al-
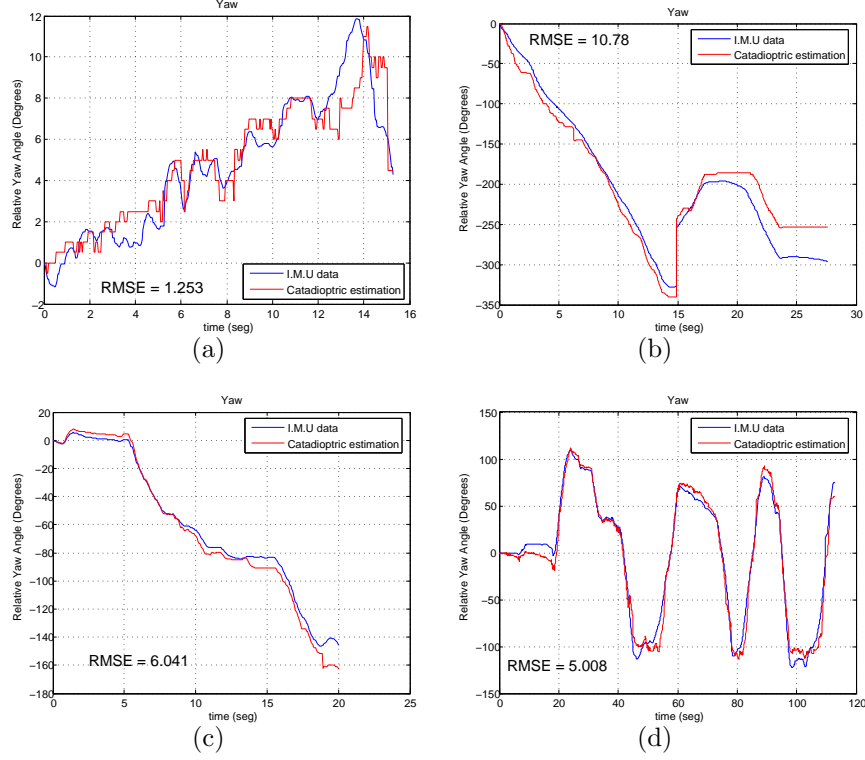
Figure 14: Estimated Relative Yaw compared with the Relative Yaw Angle measured by IMU (Heading (yaw) angle in Degrees. $-180 < \psi < 180$, zero is due north and increasing is rotating clockwise when viewed from above in absolute measuring and $-360 < \psi < 360$ in relative mode.). Figures 14(a) to 14(d) corresponds to the flights shown in figure 11.

titudes. During the first 20 seconds, the helicopter is on the ground, making a rotor speeding up in order to reach the necessary revolutions (R.P.M) to lift the helicopter on air, causing vibration on the camera platform. Skyline segmentation works fine on the captured images, but because the platform has a vibration component, the attitude angles have a big noise component on the estimated values during the beginning of this test as figures 12(d), 13(d) and 14(d) show. This noise component during the takeoff produces a small increment of the RMSE values for this flight.

Tests have been done under different weather and sky conditions, flying at different heights and including images sequences taken with the helicopter on the ground and during a takeoff phase. In all cases, results are good,

showing the feasibility of using a catadioptric system as a UAV attitude and heading estimator or as a redundant visual system. The total video sequences for these flights and additional tests in manual and autonomous mode are available on the Colibri Project Web Page [27].

6. Conclusions

This paper deals with the research and results of the use of omnidirectional computer vision techniques onboard a UAV. These techniques are not merely used to acquire environmental visual information that can be used afterward by off-line processing. Rather, this paper has shown that computer vision can play an important role on-line during the flight itself in order to acquire the essential UAV state information in order to be used as flight controller's sensor.

We have developed and tested a method for UAV attitude (roll and pitch) and heading estimation based totally on visual information taken by a catadioptric camera. This approach has been validated against inertial measures using a UAV testbed, showing that the estimated values are a very good approximation of the real state of the aircraft, demonstrating the feasibility of using this kind of system as a main sensor on UAVs with reduced sensor payloads or as a redundant system for IMU and gyroscopes in cases of failure or malfunction of these systems. The pitch and roll are calculated in relation to the horizon and they are thereafter referenced to the actual land inclination, which can varies during the flight. Similarly the heading is calculated in reference to external objects, so it can used with adventage in trajectory planning.

We have tested the algorithm under different weather conditions, showing the adaptability of the proposed method to many sunlight and illumination conditions and the robustness in extreme cases like a totally cloudy sky or a sunset in which the skyline isolation is difficult to obtain.

Tests have shown a good performance at different fly levels, with a high robustness at low altitudes in which high objects (trees, mountains and buildings) often occlude parts of the contour of the earth sphere surface, making difficult the skyline detection and segmentation. At high altitudes, this objects are under the horizon line level, allowing a excellent sky ground seg-

mentation and algorithm efficiency.

Estimated values using the omnidirectional system, have shown a good response to strong changes on the aircraft flying state with a near real time computational cost. These visual based measurements have been proved to vary accordingly to the measurements of the classical sensors, showing discrepancies of only a few degrees. These discrepancies can be used for sensor fusion, as well as for obstacle avoidance and trajectory updating. Considering these facts, the inertial data measured with the catadioptric system are appropriate for a flight controller based on visual sensor with additional features like object tracking and servoing.

Based on the results of our work, we conclude that the UAV field has reached an important stage of maturity, in which the possibility of using vision as a main sensor for UAVs control can now be imagined and in some cases attained. We have experimentally demonstrated capabilities to measure the attitude and the heading of an autonomous helicopter by using only visual information.

Acknowledgement

References

[1] R. W. Beard, State estimation for micro air vehicles, in: Innovations in Intelligent Machines (1), 2007, pp. 173–199.

[2] B. Siciliano, O. Khatib (Eds.), Springer Handbook of Robotics, Springer, Berlin, Heidelberg, 2008. doi:http://dx.doi.org/10.1007/978-3-540-30301-5.

[3] J. A. Walraven, Failure mechanisms in M.E.M.S., Internantional Test Conference (2003) 828.

[4] A. Dumai, A. Winkler, Reliability prediction model for gyroscopes, Annual Reliability and Maintainability Symposium, 1990. Proceedings. (1990) 5–9doi:10.1109/ARMS.1990.67921.

[5] S. M. Ettinger, M. C. Nechyba, P. G. Ifju, M. Waszak, Vision-guided flight stability and control for micro air vehicles, in: IEEE International Conference on Intelligent Robots and Systems, IEEE, 2002.

[6] S. Todorovic, M. Nechyba, P. Ifju, Sky/ground modeling for autonomous MAV flight, Robotics and Automation, 2003. Proceedings. ICRA '03. IEEE International Conference on 1 (2003) 1422–1427 vol.1. doi:10.1109/ROBOT.2003.1241791.

[7] T. Cornall, G. Egan, A. Price, Aircraft attitude estimation from horizon video, Electronics Letters 42 (13) (2006) 744–745. doi:10.1049/el:20060547.

[8] D. Dusha, W. Boles, R. Walker, Fixed-wing attitude estimation using computer vision based horizon detection, in: In Proceedings 12th Australian International Aerospace Congress, Melbourne, Australia, 2007, pp. 1–19.

[9] S. Hrabar, G. Sukhatme, Omnidirectional vision for an autonomous helicopter, in: In IEEE Internation Conference on Robotics and Automation, 2003, pp. 558–563.

[10] C. Demonceaux, P. Vasseur, C. Pgard, Omnidirectional vision on UAV for attitude computation, in: IEEE International Conference on Robotics and Automation 2006 (ICRA'06), IEEE, Orlando, FL, US, 2006, pp. 2842–2847.

[11] F. Labrosse, The visual compass: Performance and limitations of an appearance-based method, Journal Of Field Robotics 23 (10) (2006) 913–941.

[12] D. Scaramuzza, R. Siegwart, Appearance guided monocular omnidirectional visual odometry for outdoor ground vehicles, IEEE Transactions on Robotics, Special Issue on Visual SLAM. In press. Guest editors:

Jose' Neira, Andrew Davison, John J. Leonard, Publication date: October 2008 (2008).

[13] P. Campoy, J. F. Correa, I. Mondragón, C. Martínez, M. Olivares, L. Mejías, J. Artieda, Computer vision onboard UAVs for civilian tasks, Journal of Intelligent and Robotic Systems. 54 (1-3) (2009) 105–135. doi:http://dx.doi.org/10.1007/s10846-008-9256-z.

[14] S. Nayar, S. Baker, A theory of catadioptric image formation, tech report CUCS-015-97, Department of Computer Science, Columbia University, 1997.

[15] S. Baker, S. K. Nayar, A theory of single-viewpoint catadioptric image formation, International Journal of Computer Vision 35 (2) (1999) 1 – 22.

[16] C. Geyer, K. Daniilidis, A unifying theory for central panoramic systems and practical applications, in: ECCV (2), 2000, pp. 445–461.

[17] C. Geyer, K. Daniilidis, Catadioptric projective geometry, Journal of Computer Vision 43 (2001) 223–243.

[18] J. a. Barreto, H. Araujo, Issues on the geometry of central catadioptric image formation, Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on 2 (2001) II–422–II–427 vol.2. doi:10.1109/CVPR.2001.990992.

[19] J. a. Barreto, H. Araujo, Geometric properties of central catadioptric line images, in: ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part IV, Springer-Verlag, London, UK, 2002, pp. 237–251.

[20] T. Svoboda, T. Pajdla, V. Hlaváč, Central panoramic cameras: Geometry and design, Research report K335/97/147, Czech Technical University, Faculty of Electrical Engineering, Center for Machine Perception, FEL ČVUT, Karlovo náměstí 13, Praha, Czech Republic, available at ftp://cmp.felk.cvut.cz/pub/cmp/articles/svoboda/TR-K335-97-147.ps.gz (December 1997).

[21] T. Svoboda, T. Pajdla, Epipolar geometry for central catadioptric cameras, Int. J. Computer Vision. 49 (1) (2002) 23–37.

[22] J. Okamoto Jr, V. Grassi Jr., Visual servo control of a mobile robot using omnidirectional vision, in: Proceedings of Mechatronics 2002, University of Twente, Netherlands, 2002, pp. 413–422.

[23] C. Mei, P. Rives, Single view point omnidirectional camera calibration from planar grids, in: IEEE International Conference on Robotics and Automation, 2007.

[24] X. Ying, Z. Hu, Catadioptric camera calibration using geometric invariants, Pattern Analysis and Machine Intelligence, IEEE Transactions on 26 (10) (2004) 1260–1271. doi:10.1109/TPAMI.2004.79.

[25] H. J. Antonisse, Image segmentation in pyramids, j. Comput Graphics and Image Process 19 (4) (1982) 367–383.

[26] G. Bradski, A. Kaehler, Learning OpenCV: Computer Vision with the OpenCV Library, O'Reilly, Cambridge, MA, 2008.

[27] COLIBRI, Universidad Politécnica de Madrid. Computer Vision Group. COLIBRI Project, http://www.disam.upm.es/colibri (2009).