

VISION-BASED AUTONOMOUS APPROACH AND LANDING FOR AN AIRCRAFT USING A DIRECT VISUAL TRACKING METHOD

Tiago F. Gonçalves, José R. Azinheira
IDMEC, IST/TULisbon, Av. Rovisco Pais N.1, 1049-001 Lisboa, Portugal

Patrick Rives
INRIA-Sophia Antipolis, 2004 Route des Lucioles, BP93, 06902 Sophia-Antipolis, France

Keywords: Aircraft autonomous approach and landing, Vision-based control, Linear optimal control, Dense visual tracking.

Abstract: This paper presents a feasibility study of a vision-based autonomous approach and landing for an aircraft using a direct visual tracking method. Auto-landing systems based on the *Instrument Landing System* (ILS) have already proven their importance through decades but general aviation stills without cost-effective solutions for such conditions. However, vision-based systems have shown to have the adequate characteristics for the positioning relatively to the landing runway. In the present paper, rather than points, lines or other features susceptible of extraction and matching errors, dense information is tracked in the sequence of captured images using an *Efficient Second-Order Minimization* (ESM) method. Robust under arbitrary illumination changes and with real-time capability, the proposed visual tracker suits all conditions to use images from standard CCD/CMOS to *Infrared* (IR) and radar imagery sensors. An optimal control design is then proposed using the homography matrix as visual information in two distinct approaches: reconstructing the position and attitude (pose) of the aircraft from the visual signals and applying the visual signals directly into the control loop. To demonstrate the proposed concept, simulation results under realistic atmospheric disturbances are presented.

1 INTRODUCTION

Approach and landing are known to be the most demanding flight phase in fixed-wing flight operations. Due to the altitudes involved in flight and the consequent nonexisting depth perception, pilots must interpret position, attitude and distance to the runway using only two-dimensional cues like perspective, angular size and movement of the runway. At the same time, all six degrees of freedom of the aircraft must be controlled and coordinated in order to meet and track the correct glidepath till the touchdown.

In poor visibility conditions and degraded visual references, landing aids must be considered. The *Instrument Landing System* (ILS) is widely used in most of the international airports around the world allowing pilots to establish on the approach and follow the ILS, in autopilot or not, until the decision height is reached. At this point, the pilot must have visual contact with the runway to continue the approach and proceed to the flare manoeuvre or, if it is not the

case, to abort. This procedure has proven its reliability through decades but landing aids systems that require onboard equipment are still not cost-effective for most of the general airports. However, in the last years, the *Enhanced Visual Systems* (EVS) based on *Infrared* (IR) allowed the capability to proceed to non-precision approaches and obstacle detection for all weather conditions. The vision-based control system proposed in the present paper intends then to take advantage of these emergent vision sensors in order to allow precision approaches and autonomous landing.

The intention of using vision systems for autonomous landings or simply estimate the aircraft position and attitude (pose) is not new. Flight tests of a vision-based autonomous landing relying on feature points on the runway were already referred by (Dickmanns and Schell, 1992) whilst (Chatterji et al., 1998) present a feasibility study on pose determination for an aircraft night landing based on a model of the *Approach Lighting System* (ALS). Many others have followed in using vision-based control on

fixed/rotary wings aircrafts, and even airships, in several goals since autonomous aerial refueling ((Kimmet et al., 2002), (Mati et al., 2006)), stabilization with respect to a target ((Hamel and Mahony, 2002), (Azinheira et al., 2002)), linear structures following ((Silveira et al., 2003), (Rives and Azinheira, 2004), (Mahony and Hamel, 2005)) and, obviously, automatic landing ((Sharp et al., 2002), (Rives and Azinheira, 2002), (Proctor and Johnson, 2004), (Bourquardez and Chaumette, 2007a), (Bourquardez and Chaumette, 2007b)). In these problems, different types of visual features were considered including geometric model of the target, points, corners of the runway, binormalized Plucker coordinates, the three parallel lines of the runway (at left, center and right sides) and the two parallel lines of the runway along with the horizon line and the vanishing point. Due to the standard geometry of the landing runway and the decoupling capabilities, the last two approaches have been preferred in problems of autonomous landing.

In contrast with feature extraction methods, direct or dense methods are known by their accuracy because all the information in the image is used without intermediate processes, reducing then the sources of errors. The usual disadvantage of such method is the computational consuming of the sum-of-squared-differences minimization due to the computation of the Hessian matrix. The *Efficient Second-order Minimization* (ESM) (Malis, 2004) (Behimane and Malis, 2004) (Malis, 2007) method does not require the computation of the Hessian matrix maintaining however the high convergence rate characteristic of the second-order minimizations as the Newton method. Robust under arbitrary illumination changes (Silveira and Malis, 2007) and with real-time capability, the ESM suits all the requirements to use images from the common CCD/CMOS to IR sensors.

In vision-based or visual servoing problems, a planar scene plays an important role since it simplifies the computation of the projective transformation between two images of the scene: the planar homography. The Euclidean homography, computed with the knowledge of the calibration matrix of the imagery sensor, is here considered as the visual signal to use into the control loop in two distinct schemes. The *position-based visual servoing* (PBVS) uses the recovered pose of the aircraft from the estimated projective transformation whilst the *image-based visual servoing* (IBVS) uses the visual signal directly into the control loop by means of the interaction matrix. The controller gains, from standard output error LQR design with a PI structure, are common for both schemes whose results will be then compared with a sensor-based scheme where precise measures are considered.

The present paper is organized as follows: In the Section 2, some useful notations in computer vision are presented, using as example the rigid-body motion equation, along with an introduction of the considered frames and a description of the aircraft dynamics and the pinhole camera models. In the same section, the two-view geometry is introduced as the basis for the IBVS control law. The PBVS and IBVS control laws are then presented in the Section 3 as well as the optimal controller design. The results are shown in Section 4 while the final conclusions are presented in Section 5.

2 THEORETICAL BACKGROUND

2.1 Notations

The rigid-body motion of the frame b with respect to frame a by ${}^a\mathbf{R}_b \in \mathbb{SO}(3)$ and ${}^a\mathbf{t}_b \in \mathbb{R}^3$, the rotation matrix and the translation vector respectively, can be expressed as

$${}^a\mathbf{X} = {}^a\mathbf{R}_b {}^b\mathbf{X} + {}^a\mathbf{t}_b \quad (1)$$

where, ${}^a\mathbf{X} \in \mathbb{R}^3$ denotes the coordinates of a 3D point in the frame a or, in a similar manner, in homogeneous coordinates as

$${}^a\mathcal{X} = {}^a\mathbf{T}_b {}^b\mathcal{X} = \begin{bmatrix} {}^a\mathbf{R}_b & {}^a\mathbf{t}_b \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} {}^b\mathbf{X} \\ 1 \end{bmatrix} \quad (2)$$

where, ${}^a\mathbf{T}_b \in \mathbb{SE}(3)$, $\mathbf{0}$ denotes a matrix of zeros with the appropriate dimensions and ${}^a\mathcal{X} \in \mathbb{R}^4$ are the homogeneous coordinates of the point ${}^a\mathbf{X}$. In the same way, the Coriolis theorem applied to 3D points can also be expressed in homogenous coordinates, as a result of the derivative of the rigid-body motion relation in Eq. (2),

$$\begin{aligned} {}^a\dot{\mathcal{X}} &= {}^a\dot{\mathbf{T}}_b {}^b\mathcal{X} = {}^a\dot{\mathbf{T}}_b {}^a\mathbf{T}_b^{-1} {}^a\mathcal{X} = \\ &= \begin{bmatrix} \widehat{\boldsymbol{\omega}} & \mathbf{v} \\ \mathbf{0} & 0 \end{bmatrix} {}^a\mathcal{X} = {}^a\widehat{\mathbf{V}}_{ab} {}^a\mathcal{X}, \quad {}^a\widehat{\mathbf{V}}_{ab} \in \mathbb{R}^{4 \times 4} \end{aligned} \quad (3)$$

where, the angular velocity tensor $\widehat{\boldsymbol{\omega}} \in \mathbb{R}^{3 \times 3}$ is the skew-symmetric matrix of the angular velocity vector $\boldsymbol{\omega}$ such that $\boldsymbol{\omega} \times \mathbf{X} = \widehat{\boldsymbol{\omega}}\mathbf{X}$ and the vector ${}^a\widehat{\mathbf{V}}_{ab} = [\mathbf{v}, \boldsymbol{\omega}]^\top \in \mathbb{R}^6$ denotes the velocity screw and indicates the velocity of the frame b moving relative to a and viewed from the frame a . Also important in the present paper is the definition of stacked matrix, denoted by the superscript "s", where each column is rearranged into a single column vector.

2.2 Aircraft Dynamic Model

Let \mathcal{F}_0 be the earth frame, also called NED for North-East-Down, whose origin coincides with the desirable

touchdown point in the runway. The latter, unless explicitly indicated and without loss of generality, will be considered aligned with North. The aircraft linear velocity $\mathbf{v} = [u, v, w] \in \mathbb{R}^3$, as well as its angular velocity $\boldsymbol{\omega} = [p, q, r] \in \mathbb{R}^3$, is expressed in the aircraft body frame \mathcal{F}_c whose origin is at the center of gravity where u is defined towards the aircraft nose, v towards the right wing and w downwards. The attitude, or orientation, of the aircraft with respect to the earth frame \mathcal{F}_0 is stated in terms of Euler angles $\Phi = [\phi, \theta, \psi] \in \mathbb{R}^2$, the roll, pitch and yaw angles respectively.

The aircraft motion in atmospheric flight is usually deduced using Newton's second law and considering the motion of the aircraft in the earth frame \mathcal{F}_0 , assumed as an inertial frame, under the influence of forces and torques due to gravity, propulsion and aerodynamics. As mentioned above, both linear and angular velocities of the aircraft are expressed in the body frame \mathcal{F}_b as well as for the considered forces and moments. As a consequence, the Coriolis theorem must be invoked and the kinematic equations appear naturally relating the angular velocity rate $\boldsymbol{\omega}$ with the time derivative of the Euler angles $\dot{\Phi} = \mathbf{R}^{-1}\boldsymbol{\omega}$ and the instantaneous linear velocity \mathbf{v} with the time derivative of the NED position $[\dot{N}, \dot{E}, \dot{D}]^T = \mathbf{S}^T \mathbf{v}$.

In order to simplify the controller design, it is common to linearize the non-linear model around an given equilibrium flight condition, usually a function of airspeed \mathbf{V}_0 and altitude h_0 . This equilibrium or trim flight is frequently chosen to be a steady wings-level flight, without presence of wind disturbances, also justified here since non-straight landing approaches are not considered in the present paper. The resultant linear model is then function of the perturbation in the state vector \mathbf{x} and in the input vector \mathbf{u} as

$$\begin{bmatrix} \dot{\mathbf{x}}_v \\ \dot{\mathbf{x}}_h \end{bmatrix} = \begin{bmatrix} \mathbf{A}_v & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_h \end{bmatrix} \begin{bmatrix} \mathbf{x}_v \\ \mathbf{x}_h \end{bmatrix} + \begin{bmatrix} \mathbf{B}_v \\ \mathbf{B}_h \end{bmatrix} \begin{bmatrix} \mathbf{u}_v \\ \mathbf{u}_h \end{bmatrix} \quad (4)$$

describing the dynamics of the two resultant decoupled, lateral and longitudinal, motions. The longitudinal, or vertical, state vector is $\mathbf{x}_v = [u, w, q, \theta]^T \in \mathbb{R}^4$ and the respective input vector $\mathbf{u}_v = [\delta_E, \delta_T]^T \in \mathbb{R}^2$ (elevator and throttle) while, in the lateral case, the state vector is $\mathbf{x}_h = [v, p, r, \phi]^T \in \mathbb{R}^4$ and the respective input vector $\mathbf{u}_h = [\delta_A, \delta_R]^T \in \mathbb{R}^2$ (ailerons and rudder). Because the equilibrium flight condition is slowly varying for manoeuvres as the landing phase, the linearized model in Eq. (4) can then be considered constant along all the glidepath.

2.3 Pinhole Camera Model

The onboard camera frame \mathcal{F}_c , rigidly attached to the aircraft, has its origin at the center of projection of the camera, also called pinhole. The corresponding z -axis, perpendicular to the image plane, lies on the optical axis while the x - and y -axis are defined towards right and down, respectively. Note that the camera frame \mathcal{F}_c is not in agreement with those usually defined in flight mechanics.

Let us consider a 3D point \mathcal{P} whose coordinates in the camera frame \mathcal{F}_c are ${}^c\mathbf{X} = [X, Y, Z]^T$. This point is perspective projected onto the normalized image plane $I_m \in \mathbb{R}^2$, distant one-meter from the center of projection, at the point $\mathbf{m} = [x, y, 1]^T \in \mathbb{R}^2$ such that

$$\mathbf{m} = \frac{1}{Z} {}^c\mathbf{X}. \quad (5)$$

Note that, computing the projected point \mathbf{m} knowing coordinates \mathbf{X} of the 3D point is a straightforward problems but the inverse is not true because Z is one of the unknowns. As a consequence, the coordinates of the point \mathbf{X} could only be computed up to a scale factor, resulting on the so-called lost of depth perception.

When a digital camera is considered, the same point \mathcal{P} is projected onto the image plane I , whose distance to the center of projection is defined by the focal length $f \in \mathbb{R}^+$, at the pixel $\mathbf{p} = [p_x, p_y, 1] \in \mathbb{R}^3$ as

$$\mathbf{p} = \mathbf{K}\mathbf{m} \quad (6)$$

where, $\mathbf{K} \in \mathbb{R}^{3 \times 3}$ is the camera intrinsical parameters, or calibration matrix, defined as follows

$$\mathbf{K} = \begin{bmatrix} f_x & fs & p_{x0} \\ 0 & f_y & p_{y0} \\ 0 & 0 & 1 \end{bmatrix} \quad (7)$$

The coordinates $\mathbf{p}_0 = [p_{x0}, p_{y0}, 1]^T \in \mathbb{R}^3$ define the principal point, corresponding to the intersection between the image plane and the optical axis. The parameter s , zero for most of the cameras, is the skew factor which characterizes the affine pixel distortion and, finally, f_x and f_y are the focal lengths in the both directions such that when $f_x = f_y$ the camera sensor presents square pixels.

2.4 Two-views Geometry

Let us consider a 3D point \mathcal{P} whose coordinates ${}^c\mathbf{X}$ in the current camera frame are related with those ${}^0\mathbf{X}$ in the earth frame by the rigid-body motion in Eq. (1) of \mathcal{F}_c with respect to \mathcal{F}_0 as

$${}^c\mathbf{X} = {}^c\mathbf{R}_0 {}^0\mathbf{X} + {}^c\mathbf{t}_0. \quad (8)$$

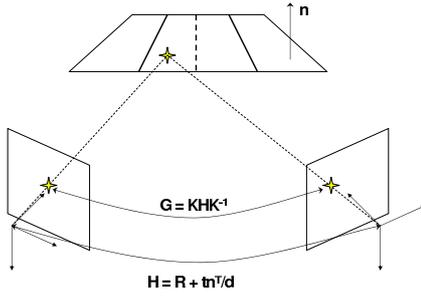


Figure 1: Perspective projection induced by a plane.

Let us now consider a second camera frame denoted reference camera frame \mathcal{F}_* in which the coordinates of the same point \mathcal{P} , in a similar manner as before, are

$${}^*\mathbf{X} = {}^*\mathbf{R}_0 {}^0\mathbf{X} + {}^*\mathbf{t}_0. \quad (9)$$

By using Eq. (8) and Eq. (9), it is possible to relate the coordinates of the same point \mathcal{P} between reference \mathcal{F}_* and current \mathcal{F}_c camera frames as

$$\begin{aligned} {}^c\mathbf{X} &= {}^c\mathbf{R}_0 {}^*\mathbf{R}_0^T {}^*\mathbf{X} + {}^c\mathbf{t}_0 - {}^c\mathbf{R}_0 {}^*\mathbf{R}_0^T {}^*\mathbf{t}_0 = (10) \\ &= {}^c\mathbf{R}_* {}^*\mathbf{X} + {}^c\mathbf{t}_* \end{aligned}$$

However, considering that \mathcal{P} lies on a plane Π , the plane equation applied to the coordinates of the same point in the reference camera frame gives us

$${}^*\mathbf{n}^T {}^*\mathbf{X} = d^* \Leftrightarrow \frac{1}{d^*} {}^*\mathbf{n}^T {}^*\mathbf{X} = 1 \quad (11)$$

where, ${}^*\mathbf{n}^T = [n_1, n_2, n_3]^T \in \mathbb{R}^3$ is the unit normal vector of the plane Π with respect to \mathcal{F}_* and $d^* \in \mathbb{R}^+$ the distance from the plane Π to the optical center of same frame. Thus, substituting Eq. (11) into Eq. (10) results on

$${}^c\mathbf{X} = \left({}^c\mathbf{R}_* + \frac{1}{d^*} {}^c\mathbf{t}_* {}^*\mathbf{n}^T \right) {}^*\mathbf{X} = {}^c\mathbf{H}_* {}^*\mathbf{X} \quad (12)$$

where, ${}^c\mathbf{H}_* \in \mathbb{R}^{3 \times 3}$ is the so-called Euclidean homography matrix. Applying the perspective projection from Eq. (5) along with the Eq. (6) into the planar homography mapping defined in Eq. (12), the relation between pixels coordinates \mathbf{p} and ${}^*\mathbf{p}$ illustrated in Figure 1 is obtained as follows

$${}^c\mathbf{p} \propto \mathbf{K} {}^c\mathbf{H}_* \mathbf{K}^{-1} {}^*\mathbf{p} \propto {}^c\mathbf{G}_* {}^*\mathbf{p} \quad (13)$$

where, $\mathbf{G} \in \mathbb{R}^{3 \times 3}$ is the projective homography matrix and " \propto " denotes proportionality.

3 VISION-BASED AUTONOMOUS APPROACH AND LANDING

3.1 Visual Tracking

The visual tracking is achieved by directly estimating the projective transformation between the image

taken from the airborne camera and a given reference image. The reference images are then the key to relate the motion of the aircraft \mathcal{F}_b , through its airborne camera \mathcal{F}_c , with respect to the earth frame \mathcal{F}_0 . For the PBVS scheme, it is the known pose of the reference camera with respect to the earth frame that will allow us to reconstruct the aircraft position with respect to the same frame. What concerns the IBVS, where the aim is to reach a certain configuration expressed in terms of the considered feature, the path planning is then an implicit need of such scheme. For example, if lines are considered as features, the path planning is defined as a function of the parameters which define those lines. In the present case, the path planning shall be defined by images because it is the dense information that is used in order to estimate the projective homography ${}^c\mathbf{G}_*$.

3.2 Visual Servoing

3.2.1 Linear Controller

The standard LQR optimal control technique was chosen for the controller design, based on the linearized models of both longitudinal and lateral motions in Eq. (4). Since not all the states are expected to be driven to zero but to a given reference, the control law is more conveniently expressed as an optimal output error feedback. The objective of the following vision-based control approaches is then to express the respective control laws as a function of the visual information, which is directly or indirectly related with the pose of the aircraft. As a consequence, the pose state vector $\mathbf{P} = [n, e, d, \phi, \theta, \psi]^T \in \mathbb{R}^6$, in agreement to the type of vision-based control approach, is given differently from the velocity screw $\mathbf{V} = [u, v, w, p, q, r]^T \in \mathbb{R}^6$, which could be provided from an existent *Inertial Navigation System* (INS) or from some filtering method based on the estimated pose. Thus, the following vision-based control laws are more correctly expressed as

$$\mathbf{u} = -\mathbf{k}_p(\mathbf{P} - \mathbf{P}^*) - \mathbf{k}_v(\mathbf{V} - \mathbf{V}^*) \quad (14)$$

where, \mathbf{k}_p and \mathbf{k}_v are the controller gains relative to the pose and velocity states, respectively.

3.2.2 Position-based Visual Servoing

In the position-based, or 3D, visual servoing (PBVS) the control law is expressed in the Cartesian space and, as a consequence, the visual information computed into the form of planar homography is used to reconstruct explicitly the pose (position and attitude). The airborne camera will be then considered as only

another sensor that provides a measure of the aircraft pose.

In the same way that, knowing the relative pose between the two cameras, \mathbf{R} and \mathbf{t} , and the planar scene parameters, \mathbf{n} and d , it is possible to compute the planar homography matrix \mathbf{H} it is also possible to recover the pose from the decomposition of the estimated projective homography \mathbf{G} , with the additional knowledge of the calibration matrix \mathbf{K} . The decomposition of \mathbf{H} can be performed by singular value decomposition (Faugeras, 1993) or, more recently, by an analytical method (Vargas and Malis, 2007). These methods result into four different solutions but only two are physically admissible. The knowledge of the normal vector \mathbf{n} , which defines the planar scene Π , allows us then to choose the correct solution.

Therefore, from the decomposition of the estimated Euclidean homography

$${}^c\tilde{\mathbf{H}}_* = \mathbf{K}^{-1} {}^c\tilde{\mathbf{G}}_* \mathbf{K}, \quad (15)$$

both ${}^c\tilde{\mathbf{R}}_*$ and ${}^c\tilde{\mathbf{t}}_*/d^*$ are recovered being respectively, the rotation matrix and normalized translation vector. With the knowledge of the distance d^* , it is then possible to compute the estimated rigid-body relation of the aircraft frame \mathcal{F}_b with respect to the inertial one \mathcal{F}_0 as

$${}^0\tilde{\mathbf{T}}_b = {}^0\mathbf{T}_* \left({}^b\mathbf{T}_c {}^c\tilde{\mathbf{T}}_* \right)^{-1} = \begin{bmatrix} {}^0\tilde{\mathbf{R}}_b & {}^0\tilde{\mathbf{t}}_b \\ \mathbf{0} & 1 \end{bmatrix} \quad (16)$$

where, ${}^b\mathbf{T}_c$ corresponds to the pose of the airborne camera frame \mathcal{F}_c with respect to the aircraft body frame \mathcal{F}_b and ${}^0\mathbf{T}_*$ to the pose of the reference camera frame \mathcal{F}_* with respect to the earth frame \mathcal{F}_0 . Finally, without further considerations, the estimated pose $\tilde{\mathbf{P}}$ obtained from ${}^0\tilde{\mathbf{R}}_b$ and ${}^0\tilde{\mathbf{t}}_b$ could then be applied to the control law in Eq. (17) as

$$\mathbf{u} = -\mathbf{k}_p(\tilde{\mathbf{P}} - \mathbf{P}^*) - \mathbf{k}_v(\mathbf{V} - \mathbf{V}^*) \quad (17)$$

3.2.3 Image-based Visual Servoing

In the image-based, or 2D, visual servoing (IBVS) the control law is expressed directly in the image space. Then, in contrast with the previous approach, the IBVS does not need the explicit aircraft pose relative to the earth frame. Instead, the estimated planar homography $\tilde{\mathbf{H}}$ is used directly into the control law as some kind of pose information such that reaching a certain reference configuration \mathbf{H}^* the aircraft presents the intended pose. This is the reason why an IBVS scheme needs implicitly for path planning expressed in terms of the considered features.

In IBVS schemes, an important definition is that of interaction matrix which is the responsible to relate

the time derivative of the visual signal vector $\mathbf{s} \in \mathbb{R}^k$ with the camera velocity screw ${}^c\mathbf{V}_{c^*} \in \mathbb{R}^6$ as

$$\dot{\mathbf{s}} = \mathbf{L}_s {}^c\mathbf{V}_{c^*} \quad (18)$$

where, $\mathbf{L}_s \in \mathbb{R}^{k \times 6}$ is the interaction matrix, or the feature jacobian. Let us consider, for a moment, that the visual signal vector \mathbf{s} is a matrix and equal to the Euclidean homography matrix ${}^c\mathbf{H}_*$, the visual feature considered in the present paper. Thus, the time derivative of \mathbf{s} , admitting the vector ${}^*\mathbf{n}/d^*$ as slowly varying, is

$$\dot{\mathbf{s}} = {}^c\dot{\mathbf{H}}_* = {}^c\dot{\mathbf{R}}_* + \frac{1}{d^*} {}^c\dot{\mathbf{t}}_* {}^*\mathbf{n}^\top \quad (19)$$

Now, it is known that both ${}^c\dot{\mathbf{R}}_*$ and ${}^c\dot{\mathbf{t}}_*$ are related with the velocity screw ${}^c\mathbf{V}_{c^*}$, which could be determined using Eq. (3), as follows

$$\begin{aligned} {}^c\hat{\mathbf{V}}_{c^*} &= {}^c\hat{\mathbf{T}}_* {}^c\mathbf{T}_*^{-1} = \\ &= \begin{bmatrix} {}^c\dot{\mathbf{R}}_* {}^c\mathbf{R}_*^\top & {}^c\dot{\mathbf{t}}_* - {}^c\dot{\mathbf{R}}_* {}^c\mathbf{R}_*^\top {}^c\mathbf{t}_* \\ \mathbf{0} & 1 \end{bmatrix} \end{aligned} \quad (20)$$

from where, ${}^c\dot{\mathbf{R}}_* = {}^c\hat{\omega} {}^c\mathbf{R}_*$ and ${}^c\dot{\mathbf{t}}_* = {}^c\mathbf{v} + {}^c\hat{\omega} {}^c\mathbf{t}_*$. By using such results back in Eq. (19) results on

$$\begin{aligned} {}^c\dot{\mathbf{H}}_* &= {}^c\hat{\omega} \left({}^c\mathbf{R}_* + \frac{1}{d^*} {}^c\mathbf{t}_* {}^*\mathbf{n}^\top \right) + \frac{1}{d^*} {}^c\mathbf{v} {}^*\mathbf{n}^\top = \\ &= {}^c\hat{\omega} {}^c\mathbf{H}_* + \frac{1}{d^*} {}^c\mathbf{v} {}^*\mathbf{n}^\top \end{aligned} \quad (21)$$

Hereafter, in order to obtain the visual signal vector, the stacked version of the homography matrix ${}^c\tilde{\mathbf{H}}_*^s$ must be considered and, as a result, the interaction matrix is given by

$$\dot{\mathbf{s}} = {}^c\dot{\mathbf{H}}_*^s = \begin{bmatrix} \mathbf{I}(3) {}^*\mathbf{n}_1/d^* & -{}^c\hat{\mathbf{H}}_{*1} \\ \mathbf{I}(3) {}^*\mathbf{n}_2/d^* & -{}^c\hat{\mathbf{H}}_{*2} \\ \mathbf{I}(3) {}^*\mathbf{n}_3/d^* & -{}^c\hat{\mathbf{H}}_{*3} \end{bmatrix} \begin{bmatrix} {}^c\mathbf{v} \\ {}^c\omega \end{bmatrix} \quad (22)$$

where, $\mathbf{I}(3)$ is the 3×3 identity matrix and \mathbf{H}_i is the i th column of the matrix as well as \mathbf{n}_i is the i th element of the vector. Note that, $\hat{\omega}\mathbf{H}$ is the external product of ω with all the columns of \mathbf{H} and $\omega \times \mathbf{H}_1 = -\mathbf{H}_1 \times \omega = -\hat{\mathbf{H}}_1\omega$.

However, the velocity screw in Eq. (18), as well as in Eq. (22), denotes the velocity of the reference frame \mathcal{F}_* with respect to the airborne camera frame \mathcal{F}_c and viewed from \mathcal{F}_c which is not in agreement with the aircraft velocity screw that must be applied into the control law in Eq. (17). Instead, the velocity screw shall be expressed with respect to the reference camera frame \mathcal{F}_* and viewed from aircraft body frame \mathcal{F}_b , where the control law is effectively applied. In this manner, and knowing that the velocity tensor ${}^c\hat{\mathbf{V}}_{c^*}$ is a skew-symmetric matrix, then

$${}^c\hat{\mathbf{V}}_{c^*} = -{}^c\hat{\mathbf{V}}_{c^*}^\top = -{}^c\hat{\mathbf{V}}_{*c} \quad (23)$$

Now, assuming the airborne camera frame \mathcal{F}_c rigidly attached to the aircraft body frame \mathcal{F}_b , to change the velocity screw from the aircraft body to the airborne camera frame, the adjoint map must be applied as

$$\begin{aligned} {}^c\hat{\mathbf{V}} &= {}^b\mathbf{T}_c^{-1}b\hat{\mathbf{V}}b\mathbf{T}_c = \\ &= \begin{bmatrix} {}^b\mathbf{R}_c^\top b\hat{\boldsymbol{\omega}}b\mathbf{R}_c & {}^b\mathbf{R}_c^\top b\mathbf{v} + {}^b\mathbf{R}_c^\top b\hat{\boldsymbol{\omega}}b\mathbf{t}_c \\ \mathbf{0} & 0 \end{bmatrix} \end{aligned} \quad (24)$$

from where, ${}^b\mathbf{R}_c^\top b\hat{\boldsymbol{\omega}}b\mathbf{R}_c = \widehat{{}^b\mathbf{R}_c^\top b\boldsymbol{\omega}}$ and ${}^b\mathbf{R}_c^\top b\hat{\boldsymbol{\omega}}b\mathbf{t}_c = -{}^b\mathbf{R}_c^\top b\hat{\mathbf{t}}_c {}^c\boldsymbol{\omega}$ and, as a result, the following velocity transformation ${}^c\mathbf{W}_b \in \mathbb{R}^{6 \times 6}$ is obtained

$${}^c\mathbf{V} = {}^c\mathbf{W}_b {}^b\mathbf{V} = \begin{bmatrix} {}^b\mathbf{R}_c^\top & -{}^b\mathbf{R}_c^\top b\hat{\mathbf{t}}_c \\ \mathbf{0} & {}^b\mathbf{R}_c^\top \end{bmatrix} \begin{bmatrix} b\mathbf{v} \\ b\boldsymbol{\omega} \end{bmatrix} \quad (25)$$

Using the Eq. (25) into the Eq. (22) along with the result from Eq. (23) results as follows

$$\dot{\mathbf{s}} = {}^c\mathbf{H}_*^s = -\mathbf{L}_s {}^c\mathbf{W}_b {}^b\mathbf{V}_{*c} \quad (26)$$

Finally, let us consider the linearized version of the previous result as

$$\mathbf{s} - \mathbf{s}^* = {}^c\mathbf{H}_*^s - \mathbf{H}^{*s} = -\mathbf{L}_s {}^c\mathbf{W}_b {}^b\mathbf{W}_0 (\mathbf{P} - \mathbf{P}^*) \quad (27)$$

where,

$${}^b\mathbf{W}_0 = \begin{bmatrix} \mathbf{S}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_0 \end{bmatrix} \quad (28)$$

are the kinematic and navigation equations, respectively, linearized for the same trim point as for the aircraft linear model $[\phi, \theta, \psi]_0^\top = [0, \theta_0, 0]^\top$. It is then possible to relate the pose error $\mathbf{P} - \mathbf{P}^*$ of the aircraft with the Euclidean homography error ${}^c\mathbf{H}_*^s - \mathbf{H}^{*s}$. For the present purpose, the reference configuration is $\mathbf{H}^* = \mathbf{I}(3)$ which corresponds to match exactly both current I and reference I^* images. The proposed homography-based IBVS visual control law is then expressed as

$$\mathbf{u} = -\mathbf{k}_p (\mathbf{L}_s {}^c\mathbf{W}_0)^\dagger ({}^c\tilde{\mathbf{H}}_*^s - \mathbf{H}^{*s}) - \mathbf{k}_v (\mathbf{V} - \mathbf{V}^*) \quad (29)$$

where, $\mathbf{A}^\dagger = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top$ is the Moore-Penrose pseudo-inverse matrix.

4 RESULTS

The vision-based control schemes proposed above have been developed and tested in an simulation framework where the non-linear aircraft model is implemented in Matlab/Simulink along with the control aspects, the image processing algorithms in C/C++ and the simulated image is generated by the Flight-Gear flight simulator. The aircraft model considered

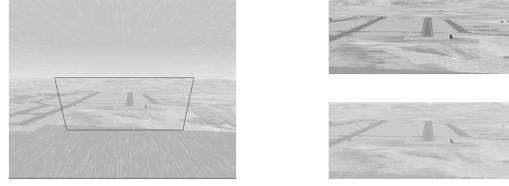


Figure 2: Screenshot from the dense visual tracking software. The delimited zone (left) corresponds to the bottom-right image, warped to match with the top-right image. The warp transformation corresponds to the estimated homography matrix.

corresponds to a generic category B business jet aircraft with $50m/s$ of stall speed, $265m/s$ of maximum speed and $20m$ wing span. This simulation framework has also the capability to generate atmospheric condition effects like fog and rain as well as vision sensors effects like pixels spread function, noise, colorimetry, distortion and vibration of different types and levels.

The chosen airport scenario was the Marseille-Marignane Airport with an nominal initial position defined by an altitude of $450m$ and a longitudinal distance to the runway of $9500m$, resulting into a 3 degrees descent for an airspeed of $60m/s$. In order to have illustrative results and to verify the robustness of the proposed control schemes, two sets of simulations results are presented. The first with an initial lateral error of $50m$, an altitude error of $30m$ and a steady wind composed by $10m/s$ headwind and $1m/s$ of crosswind. The latter, with a different initial lateral error of $75m$, considers in addition the presence of turbulence. What concerns the visual tracking aspects, a database of $200m$ equidistant images along the runway axis till the $100m$ height, and $50m$ after that, was considered and the following atmospheric conditions imposed: fog and rain densities of 0.4 and 0.8 ($[0,1]$). The airborne camera is considered rigidly attached to the aircraft and presents the following pose ${}^b\mathbf{P}_c = [4m, 0m, 0.1m, 0, -8 \text{ degrees}, 0]^\top$. The simulation framework operates with a $50ms$, or $20Hz$, sampling rate.

For all the following figures, the results of the two simulations are presented simultaneously and identified in agreement with the legend in Figure 3(a). When available, the corresponding references are presented in black dashed lines. Unless specified, the detailed discussion of results is referred to the case without turbulence.

Let us start with the longitudinal trajectory illustrated in Figure 3(a) where it is possible to verify immediately that the PBVS results are almost coincident with the ones where the sensor measurements were considered ideal (Sensors). Indeed, because the same con-

trol law is used for these two approaches, the results differ only due to the pose estimation errors from the visual tracking software. For the IBVS approach, the first observation goes to the convergence of the aircraft trajectory with respect to the reference descent that occurs later than for the other approaches. This fact is a consequence not only of the limited validity of the interaction matrix in Eq (27), computed for a stabilized descent flight, but also of the importance of the camera orientation over the position, for high altitudes, when the objective is to match two images. In more detail, the altitude error correction in Figure 3(b) shows then the IBVS with the slowest response and, in addition, a static error not greater than $2m$ as a cause of the wind disturbance. In fact, the path planning does not contemplate the presence of the wind, from which the aircraft attitude is dependent, leading to the presence of static errors. These same aspects are verified in the presence of turbulence but now with a global altitude error not greater than $8m$, after stabilization. The increasing altitude error at the distance of $650m$ before the touchdown corresponds to the natural loss of altitude when proceeding to the pitch-up, or flare, manoeuvre (see Figure 3(c)) in order to reduce the vertical speed and correctly land the aircraft. What concerns the touchdown distances, both Sensors and PBVS results are very close and at a distance around $330m$ after the threshold line while, for the IBVS, this distance is of approximately $100m$. In the presence of turbulence, these distances became shorter mostly due to the oscillations in the altitude control during the flare manoeuvre. Again, Sensors and PBVS touchdown points are very close and about $180m$ from the runway threshold line while, for the IBVS, this distance is about $70m$.

The lateral trajectory illustrated in Figure 3(e) shows a smooth lateral error correction for all the three control schemes, where both visual control laws maintain an error below the $2m$ after convergence. Once more, the oscillations around the reference are a consequence of pose estimation errors from visual tracking software, which become more important near the Earth surface due to the high displacement of the pixels in the image and the violation of the planar assumption of the region around the runway. The consequence of these effects are perceptible in the final part not only in the lateral error correction but also in the yaw angle of the aircraft in Figure 3(f). For the latter, the static error is also an influence of the wind disturbance which imposes an error of 1 degree with respect to the runway orientation of exactly 134.8 degrees North.

In the presence of turbulence, Sensors and PBVS control schemes present a different behavior during the

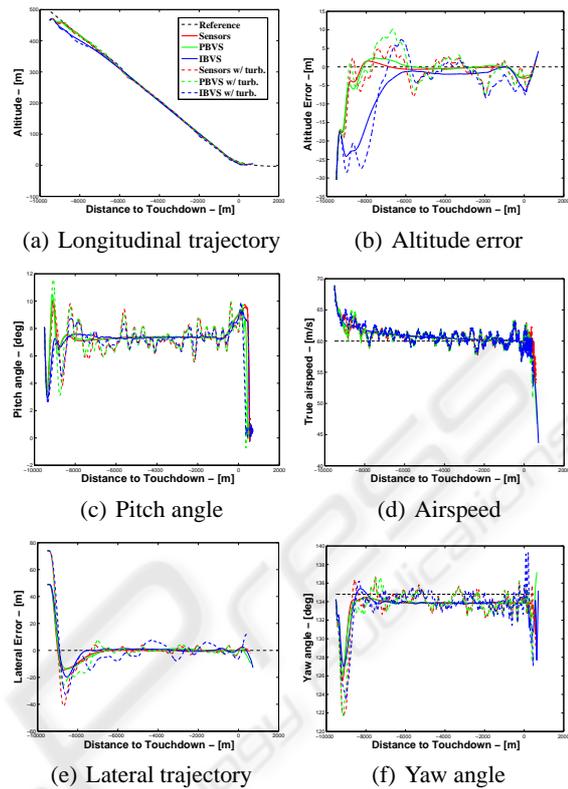


Figure 3: Results from the vision-based control schemes (PBVS and IBVS) in comparison with the ideal situation of precise measurements (Sensors).

lateral error correction manoeuvre. Indeed, due to the important bank angle and the high pitch induced by the simultaneous altitude error correction manoeuvre, the visual tracking algorithm lost information on the near-field of the camera essential for the precision of the estimated translation. The resultant lateral error estimative, greater than it really is, forces the lateral controller to react earlier in order to minimize such error. As for the longitudinal case, the IBVS presents a slower response on position error corrections resulting into a lateral error not greater than $8m$ which contrasts with the $4m$ from the other two approaches.

It should be noted the precision of the dense visual tracking software. Indeed, the attitude estimation errors are often below 1 degree for transient responses and below 0.1 degrees in steady state. Depending on the quantity of information available in the near field of the camera, the translation error could vary between the $1m$ and $4m$ for both lateral and altitude errors and between $10m$ and $70m$ for the longitudinal distance. The latter is usually less precise due to its alignment with the optical axis of the camera.

5 CONCLUSIONS

In the present paper, two vision-based control schemes for an autonomous approach and landing of an aircraft using a direct visual tracking method are proposed. For the PBVS solution, where the vision system is nothing more than a sensor providing position and attitude measures, the results are naturally very similar with the ideal case. The IBVS approach based on a path planning defined by a sequence of images shown clearly to be able to correct an initial pose error and land the aircraft under windy conditions. Despite the inherent sensitivity of the vision tracking algorithm to the non-planarity of the scene and the high pixels displacement in the image for low altitudes, a shorter distance between the images of reference was enough to deal with potential problems. The inexistence of a filtering method, as the Kalman filter, is the proof of the robustness of the proposed control schemes and the reliability of the dense visual tracking. This clearly justifies further studies to complete the validation and the eventual implementation of this system on a real aircraft.

ACKNOWLEDGEMENTS

This work is funded by the FP6 3rd Call European Commission Research Program under grant Project N.30839 - PEGASE.

REFERENCES

- Azinheira, J., Rives, P., Carvalho, J., Silveira, G., de Paiva, E.C., and Bueno, S. (2002). Visual servo control for the hovering of an outdoor robotic airship. In *IEEE International Conference on Robotics and Automation*, volume 3, pages 2787–2792.
- Behimane, S. and Malis, E. (2004). Real-time image-based tracking of planes using efficient second-order minimization. In *IEEE International Conference on Intelligent Robot and Systems*, volume 1, pages 943–948.
- Bourquardez, O. and Chaumette, F. (2007a). Visual servoing of an airplane for alignment with respect to a runway. In *IEEE International Conference on Robotics and Automation*, pages 1330–1355.
- Bourquardez, O. and Chaumette, F. (2007b). Visual servoing of an airplane for auto-landing. In *IEEE International Conference on Intelligent Robots and Systems*, pages 1314–1319.
- Chatterji, G., Menon, P., K., and Sridhar, B. (1998). Vision-based position and attitude determination for aircraft night landing. *AIAA Journal of Guidance, Control and Dynamics*, 21(1).
- Dickmanns, E. and Schell, F. (1992). Autonomous landing of airplanes by dynamic machine vision. *IEEE Workshop on Application on Computer Vision*, pages 172–179.
- Faugeras, O. (1993). *Three-dimensional computer vision: a geometric view point*. MIT Press.
- Hamel, T. and Mahony, R. (2002). Visual servoing of an under-actuated dynamics rigid-body system: an image-based approach. In *IEEE Transactions on Robotics and Automation*, volume 18, pages 187–198.
- Kimmitt, J., Valasek, J., and Junkins, J. L. (2002). Vision based controller for autonomous aerial refueling. In *Conference on Control Applications*, pages 1138–1143.
- Mahony, R. and Hamel, T. (2005). Image-based visual servo control of aerial robotic systems using linear images features. In *IEEE Transaction on Robotics*, volume 21, pages 227–239.
- Malis, E. (2004). Improving vision-based control using efficient second-order minimization technique. In *IEEE International Conference on Robotics and Automation*, pages 1843–1848.
- Malis, E. (2007). An efficient unified approach to direct visual tracking of rigid and deformable surfaces. In *IEEE International Conference on Robotics and Automation*, pages 2729–2734.
- Mati, R., Pollini, L., Lunghi, A., and Innocenti, M. (2006). Vision-based autonomous probe and drogue aerial refueling. In *Conference on Control and Automation*, pages 1–6.
- Proctor, A. and Johnson, E. (2004). Vision-only aircraft flight control methods and test results. In *AIAA Guidance, Navigation, and Control Conference and Exhibit*.
- Rives, P. and Azinheira, J. (2002). Visual auto-landing of an autonomous aircraft. Research Report 4606, INRIA Sophia-Antipolis.
- Rives, P. and Azinheira, J. (2004). Linear structure following by an airship using vanishing point and horizon line in visual servoing schemes. In *IEEE International Conference on Robotics and Automation*, volume 1, pages 255–260.
- Sharp, C., Shakernia, O., and Sastry, S. (2002). A vision system for landing an unmanned aerial vehicle. In *IEEE International Conference on Robotics and Automation*, volume 2, pages 1720–1727.
- Silveira, G., Azinheira, J., Rives, P., and Bueno, S. (2003). Line following visual servoing for aerial robots combined with complementary sensors. In *IEEE International Conference on Robotics and Automation*, pages 1160–1165.
- Silveira, G. and Malis, E. (2007). Real-time visual tracking under arbitrary illumination changes. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–6.
- Vargas, M. and Malis, E. (2007). Deeper understanding of the homography decomposition for vision-base control. Research Report 6303, INRIA Sophia-Antipolis.