IMAGE-BASED VISUAL SERVOING FOR VANISHING FEATURES AND GROUND LINES TRACKING: APPLICATION TO A UAV AUTOMATIC LANDING

José Raul Azinheira† and Patrick Rives*‡

† IDMEC / Instituto Superior Técnico, av Rovisco Pais, n.1, 1049-001 Lisbon, Portugal

email: jraz@dem.ist.utl.pt

‡ INRIA/Sophia-Antipolis, 2004 rte des Lucioles, BP93, 06902 Sophia-Antipolis, France

email: Patrick.Rives@sophia.inria.fr

Abstract

A unified modeling framework is defined and an image-based visual servoing scheme is introduced to evaluate the feasibility of a vision-based guidance for a UAV model during survey missions and for automatic landing. Relevant image features are selected to allow for the proposed objective, namely decoupling rotation and translation in the image, and allowing to respect the natural separation between the lateral and longitudinal motions of the aircraft. The controller design is adapted to include the aircraft dynamics and the vision output in an linear optimal control approach. Simulation results are presented in order to validate the proposed solution and allow for a critical evaluation.

Index Terms

Image-based visual servoing, Unmanned aerial vehicle (UAV), auto-landing, dynamic modeling, robust control

Corresponding author. Email: Patrick.Rives@sophia.inria.fr

$\mathbf{\Delta}_l, \mathbf{\Delta}_r, \mathbf{\Delta}_H$ left, right and horizon lines

$(\delta_a, \delta_e, \delta_r, \delta_T)$ aileron, elevator, rudder and throttle deflections

$\mathbf{\Omega} = [p, q, r]^T$ angular velocity in body frame

(ϕ, θ, ψ) Roll, pitch and yaw Euler angles

(ρ, θ) polar representation of straight line in image plane

$\mathbf{A}, \mathbf{B}, \mathbf{C}$ matrices of linear model

$\mathbf{K}$ state feedback gain for controller

$\mathbf{L}^T$ interaction matrix

$\mathbf{P} = [N, E, D]^T$ aircraft position in earth frame (North, East, Down)

$\mathbf{Q}, \mathbf{R}$ matrices of LQR cost function

$\bar{\mathbf{r}}$ 6D pose of camera w.r.t. earth frame

$\mathbf{S}$ transformation matrix from earth to body frame

$\mathbf{s}$ vector of visual signals

$\mathbf{T}_{CT}$ velocity screw

V_t aircraft airspeed

$\mathbf{V} = [u, v, w]^T$ inertial velocity in body frame

(x_H, y_H) coordinates of vanishing point in image

$\mathbf{x}, \mathbf{u}, \mathbf{y}, \mathbf{z}$ state, input and output vectors of linear model

$\mathbf{X} \, \mathbf{U}$ state and input vector in dynamic model

$\mathbf{x} = (x, y)$ projected point coordinates in image plane

$\mathbf{X} = (X, Y, Z)$ point coordinates in image frame

1. INTRODUCTION

The main objective of the present work is the search for a feedback control strategy applied to an unmanned air vehicle (UAV) using the image provided by an airborne camera. The application to the automatic landing of an aircraft is the more demanding flight phase, both in terms of required accuracy and necessary robustness. A specific characteristic of the landing (partly justifying the previous statement) is the fact that the relevant position variables are relative to a local ground frame and not to an absolute frame. Such a characteristic is also true for a large class of survey missions where the control objective is defined in terms of a relative positioning w.r.t. structures lying on the ground, like roads, rivers or a landing strip.

The landing problem may thus be considered as the extreme case of a need for perfect path tracking in a ground relative frame, for which a vision sensor (an airborne camera) may be regarded as a sensor with adequate characteristics, useful for the precise positioning relatively to the ground.

In terms of control strategy, two options are usually available:

- a first solution, referred as the position-based visual servoing (PBVS) or reconstruction approach, consists in considering the image output as another sensor specifically used to help for the landing phase in order to get a better estimate of the vehicle position, as compared to a desired path. The aircraft position w.r.t the ground is estimated from the image using vanishing points or lines or by computing an homography under the assumption that the ground plane is flat (Rathinam et al. 2005), (Saripalli et al. 2003), (Templeton et al. 2007).The problem of control is then stated as a trajectory tracking in the cartesian space.

- a second solution, so called image-based visual servoing (IBVS) approach, consists in switching from the flight control scheme to a landing scheme where the vehicle motion is only referenced to the image as compared to a reference image to be tracked. Until now, only a few IBVS approches have been applied to the control of aerial unmanned vehicles (see, for example, the USC Autonomous Flying Vehicle Project (Mejias et al. 2005), (Silveira et al 2003).

The position-based visual servoing approach is a more conservative solution, probably easier to implement: the flight control scheme is almost unchanged, only switching to an image enhanced estimation of the position.

On the other hand the reconstruction relies on a good calibration of the camera and the effect of the image noise or inaccuracy must be secured.

Some papers already touched the problem of vision based flight control, like (Dickmanns 1994), (Schell and Dickmanns 1994), (Chatterji et al. 1998) or (Kimmett et al. 2002). The first and the second papers are representative of an impressive amount of works on vision-based flight control developed at UniBwM (Universitaet der Bundeswehr Munich) since 1985. The authors use full state reconstruction by recursive estimation and state feedback techniques for the design of longitudinal and lateral flight controllers. In complement to image features, inertial data are also used to deal with gusts and wind inputs. The third paper presents a pose reconstruction technique to be used as a pilot aid during landing, with no automatic control. The last paper presents the simulation results of a reconstruction control scheme for in-flight refueling.

The image-based visual servoing approach is more ambitious and a cautious analysis is required before it may be tested in real flight conditions. It must be secured that the control in the image does not excite the unstable or marginally stable dynamics of the vehicle. It is however a solution that is not so dependent on the camera calibration, as long as an acceptable reference image is provided for the visual tracking (Espiau et al. 1992),(Rives et al. 1996). In contrast with the usual image-based control in robotics, in our particular case the target system may not be assumed as a pure velocity integrator, the aircraft is under-actuated and the dynamics and couplings between the axes may not be neglected and are to be included into the controller design. Another approach to overpass this problem is presented by (Hamel and Mahony 2002), with a dynamical visual servo to stabilize a scale helicopter over a landing pad and which treats image measurements as unit vectors defining a mapping from the image plane to the unit sphere which allows a decoupling between translations and rotations.

The approach presented here was originally introduced in (Rives and Azinheira 2004) for the tracking of ground features with an airship. The scheme is adapted in the present paper in order to allow the automatic landing of an unmanned aircraft. A similar approach was recently presented in (Bourquardez and Chaumette 2007) but dealing with the alignment phase and not considering the wind disturbance.

4

The present paper is structured as follows. In a first section, we briefly present the dynamic model of the aircraft, to be used for control design. The second section is devoted to the image modeling, introduces the visual servoing approach and the control design. The third section presents simulation results, which allow to evaluate the behaviour of the controled UAV and validate the approach. Some brief conclusions are finally drawn in the last section.

2. DYNAMIC MODELING OF THE AIRCRAFT

The basic equations describing the aircraft motion in atmospheric flight may be found in the bibliography on Flight Mechanics, by instance in (McLean 1990), (Stevens and Lewis 1992), (Phillips 2004).

The present section only gathers the basic notations and equations further used in the simulation setup or for the control design phase.

2.1. Frames and Notations

The aircraft trajectory and the landing road coordinates are given in the earth frame (or NED, for North-East-Down), with the center by instance on the road axis (Figure 1) -for simplicity and without loss of generality,the road is chosen as aligned with the North axis. The local frame, linked to the aircraft (ABC, aircraft body centered), where the aircraft velocity $\mathbf{V} = [u, v, w]^T$ is given, is centered at its center of gravity, u directed towards the aircraft nose, v towards the right wing and w downwards. The angular velocity is also expressed in the local frame: $\mathbf{\Omega} = [p, q, r]^T$. The airborne camera is rigidly attached to the aircraft, with its optical axis aligned with the aircraft longitudinal axis. The change from earth frame to local frame is defined by the transformation matrix $\mathbf{S}$, which may as usual be stated in terms of the Euler angles (ϕ, θ, ψ), respectively roll, pitch and yaw angles.

2.2. Flight Mechanics and Dynamic Modeling

The mechanics of atmospheric flight is generally deduced from the application of Newton's second law, considering the motion of the aircraft in the earth frame, assumed as an inertial frame, under the influence of forces and torques due to gravity, aerodynamics and propulsion.

If the aircraft motion is described as usual in the body frame by its inertial and angular velocity, and by the Euler angles, the dynamic system may be put in a synthetic form as the following state space equation:

$$\dot{\mathbf{X}} = f(\mathbf{X}, \mathbf{U}, \mathbf{D}) \tag{1}$$

where:

- $\dot{\mathbf{X}} = \frac{d\mathbf{X}}{dt}$ is the usual notation for time derivation

- $\mathbf{X} = [u, v, w, p, q, r, \phi, \theta, \psi]^T$ is the state vector,

- $\mathbf{U} = [\delta_a, \delta_e, \delta_r, \delta_T]^T$ is the input vector, with respectively the aileron, elevator and rudder deflections and the throttle input,

- the disturbance vector is usually representing the wind velocity, with its six components

  $\mathbf{D} = [u_w, v_w, w_w, p_w, q_w, r_w]^T$

The state equation (Eq.1) is established for a nearly constant air density and the aircraft motion only depends on its velocity and attitude (angular position) and the aircraft cartesian position $\mathbf{P} = [N, E, D]^T$ in the earth frame does not appear. Since this position is to be controlled, three position states are added, as integrators of the aircraft inertial velocity $\mathbf{V}$, considered in the earth frame:

$$\dot{\mathbf{P}} = \mathbf{S}^T \mathbf{V} \tag{2}$$

2.3. Linearized Models

The equations non-linearity, their complexity and a certain level of approximation in the aircraft models have justified the search for simplified versions and, as a first step, it is common to linearize the equations for small perturbations around an equilibrium flight. This equilibrium or trim flight is frequently taken as a horizontal straight leveled flight, with no wind.

Under these conditions, the equations are written as functions of the perturbations in the state vector $\mathbf{x}$, in the input vector $\mathbf{u}$ or in the disturbance vector $\mathbf{d}$, resulting in two differential matrix equations describing the dynamics of two independent (decoupled) motions:

$$\dot{\mathbf{x}}_v = \mathbf{A}_v\mathbf{x}_v + \mathbf{B}_v\mathbf{u}_v + \mathbf{E}_v\mathbf{d}_v$$

$$\dot{\mathbf{x}}_h = \mathbf{A}_h\mathbf{x}_h + \mathbf{B}_h\mathbf{u}_h + \mathbf{E}_h\mathbf{d}_h$$

(3)

where:

- the index v is associated with the longitudinal motion, in the vertical plane, and

- the index h is associated with the lateral/directional motion, mostly in the horizontal plane.

The longitudinal state vector is $\mathbf{x}_v = [u, w, q, \theta]^T$ and the input vector is $\mathbf{u}_v = [\delta_e, \delta_T]^T$ (where all the variables are changes from the trim value). In the lateral case, the state vector is $\mathbf{x}_h = [v, p, r, \phi]^T$ and the input vector $\mathbf{u}_h = [\delta_a, \delta_r]^T$.

The linear models described by Eq. 3 depend on the trim point chosen for the linearization: namely they are function of the airspeed V_o and the altitude h_o. The validity of these linear models is obviously limited to a surrounding near the trim point, namely by the linearization of the angles, and by the validity of the aerodynamic laws, within the flight envelope, which state that the airspeed must remain greater than the stall speed and lower than a maximum allowed airspeed. The inputs are also limited, in value or rate, both for the surface deflections and for the thrust input.

2.4. Experimental Setup: the ARMOR X7 UAV

The modelling process described above has been applied to the experimental UAV ARMOR X7 currently under development at IST in Lisbon (Figure 2). The half-scale flying model in use in our early experiments has 3 m wing span and 18 kg nominal weight. Its cruise airspeed is about 18 m/s.

The fully non-linear dynamic model (Eq. 1) was defined in order to enable a realistic simulation, including the effect of wind and atmospheric turbulence, as well as the ground effect when the aircraft is near touchdown. A *Matlab/Simulink* simulation platform was developed to test control solutions and evaluate strategies for the desired autonomous or semi-autonomous operation (Costa 1999). The linearized model (Eq. 3) is used in the control design phase, as by instance in (Azinheira et al. 1998).

The simulation platform allows to handle simple models of the 3D scene. A ground road was defined, with a width of $5\ m$, and a simulated camera was introduced, which outputs the visual signals to be used in the image-based visual servoing.

3. IMAGE-BASED VISUAL SERVOING

In opposite to a 3D visual servoing method (Dickmanns 1994), (Furst and Dickmanns 1998), an image-based visual servoing does not require an explicit 3D reconstruction of the scene. The basic idea is to assume that a task can be fully specified in terms of the desired configuration of a set of geometric features in the image. The task will be perfectly achieved when such a configuration is reached (Samson et al. 1990). In terms of control, this can be formulated as a problem of regulation to zero of a certain output function directly defined in the image frame.

Let us consider the airborne camera C, which can be viewed as a mechanical system with several actuated degrees of freedom. The pose (position and orientation) of C is an element $\bar{\mathbf{r}}$ of $\mathbb{R}^3 \times \mathbb{SO}_3$, which is a six dimensional differential manifold. C interacts with its environment. We assume that the image given by C (see Figure 3) fully characterizes the relative position of C with respect to the NED frame attached to the scene. Moreover, let us consider that the information in the image may be modeled as a set of *visual signals* characterizing the geometric features which result from the projection onto the image of the 3D objects belonging to the scene. Each elementary signal $\mathbf{s}_i(\bar{\mathbf{r}})$ defines a differentiable mapping from $\mathbb{R}^3 \times \mathbb{SO}_3$ to $\mathbb{R}$. As shown in (Espiau et al. 1992), the differential of this mapping (so-called *interaction matrix*) $\mathbf{L}_i^T$ relates the variation of the signal $\mathbf{s}_i(\bar{\mathbf{r}})$ observed in the image to the motion between the camera and the 3D target expressed by the camera velocity screw $\mathbf{T}_{CT}$.

$$\dot{\mathbf{s}}_i = {\mathbf{L}_i}^T \mathbf{T}_{CT} \tag{4}$$

An analytical expression for the interaction matrix when the image features are general algebraic curves can be derived (for more details, see (Chaumette et al. 1993), (Rives et al. 1996)). In our peculiar case, we consider the set of geometric primitives in the 3D scene will be constituted by points and straight lines.

3.1. Modeling the Interaction Matrix

Let us assume hereafter that we use a pinhole camera model with a focal length equal to 1 and that both the points and lines in the 3D scene and their projection in the 2D normalized image plane are expressed in the camera frame $\mathcal{F}_s$.

Case of points:

Any point M with coordinates $\mathbf{X} = (X, Y, Z)$ projects onto the image plane as a point m with coordinates $\mathbf{x} = (x, y)$ such that:

$$x = X/Z \ , \ y = Y/Z \tag{5}$$

By differentiating Eq.(5), it is obvious to compute the interaction matrix linking the 2D motion observed in the image to the camera motion in the 3D scene.

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -\frac{1}{Z} & 0 & \frac{x}{Z} & xy & -1-x^2 & y \\ 0 & -\frac{1}{Z} & \frac{y}{Z} & 1+y^2 & -xy & -x \end{bmatrix} \mathbf{T}_{CT} \tag{6}$$

Case of straight lines:

A straight line in the 3D scene is here represented as the intersection of two planes described in the implicit form $h(\underline{\mathbf{X}}, \underline{\mathbf{Q}}) = 0$ such that :

$$h(\underline{\mathbf{X}}, \underline{\mathbf{Q}}) = \begin{cases} a_1 X + b_1 Y + c_1 Z = 0 \\ a_2 X + b_2 Y + c_2 Z + d_2 = 0 \end{cases} \tag{7}$$

with $d_2 \neq 0$ in order to exclude degenerated cases. In these equations, $\underline{\mathbf{X}} = (X, Y, Z, 1)$ denotes the homogeneous coordinates, expressed in the camera frame, of the 3D points lying on the 3D line, and $\underline{\mathbf{Q}}$ denotes a parameterization of the 3D lines manifold.

The equation of the 2D projected line in the image plane (see Figure 3) can also be written in an implicit form $g(\underline{\mathbf{x}}, \underline{\mathbf{q}}) = 0$ such that :

$$g(\underline{\mathbf{x}}, \underline{\mathbf{q}}) = x \cos \theta + y \sin \theta - \rho = 0 \tag{8}$$

9

with

$$
\begin{cases}
\cos\theta = a_1/\sqrt{a_1^2 + b_1^2} \\
\sin\theta = b_1/\sqrt{a_1^2 + b_1^2} \\
\rho = -c_1/\sqrt{a_1^2 + b_1^2}
\end{cases}
$$

where $\underline{\mathbf{x}} = (x, y, 1)$ denotes the homogeneous coordinates in the image of the 2D points lying on the 2D line, and $\underline{\mathbf{q}}$ denotes a parameterization of the 2D lines manifold, hereafter the polar representation (ρ, θ).

A general form of the interaction matrix may then be obtained for each line:

$$
\begin{aligned}
\mathbf{L}_\theta^T = [\quad & \lambda_\theta \cos\theta & \lambda_\theta \sin\theta & -\lambda_\theta\rho & \ldots \\
& -\rho\cos\theta & -\rho\sin\theta & -1 &] \\
\mathbf{L}_\rho^T = [\quad & \lambda_\rho \cos\theta & \lambda_\rho \sin\theta & -\lambda_\rho\rho & \ldots \\
& (1+\rho^2)\sin\theta & -(1+\rho^2)\cos\theta & 0 &]
\end{aligned}
\tag{9}
$$

with $\lambda_\theta = (a_2\sin\theta - b_2\cos\theta)/d_2$ and $\lambda_\rho = (a_2\rho\cos\theta + b_2\rho\sin\theta + c_2)/d_2$.

3.2. Modeling the Visual Signals

A first step in order to include vision in the control loop is to define a reference scene, the image of which, as viewed from an airborne camera, would allow for a good vertical and lateral or attitude positioning, but leaving freedom enough to cope with the limitations of the vehicle dynamics.

As presented above (see Figure 1), let us consider a scene composed by a strip lying on the ground. We assume a smooth and limited curvature (piecewise linear), and the strip is parametrized by two parallel curves (i.e. the two sides of the road or river). These curves project onto the image as shown in Figure 3.

Let us define in the image the two tangents $\boldsymbol{\Delta}_r$ and $\boldsymbol{\Delta}_l$ to the right and left border lines of the road at a given image coordinate y_T. $\boldsymbol{\Delta}_r$ and $\boldsymbol{\Delta}_l$ converge to a vanishing point $\mathbf{x}_H$ which belongs to the horizon line $\boldsymbol{\Delta}_H$. The parameters of the lines $\boldsymbol{\Delta}_r$, $\boldsymbol{\Delta}_l$ and $\boldsymbol{\Delta}_H$ and the coordinates of the vanishing point $\mathbf{x}_H$ in the image depend on the relative position and attitude of the camera frame w.r.t. the road.

In terms of a survey task for an aerial vehicle, following a road on the ground consists in keeping the longitudinal axis of the vehicle collinear with the road median axis and, simultaneously, keeping the height over the road constant. The longitudinal speed may be controlled separately and tunes the survey task.

These constraints between the reference ground frame and the camera frame have to be translated in terms of reference visual signals to be observed in the image. In order to cope with the natural decoupling of the aircraft dynamics (see section 2), it is interesting to choose these visual signals such that rotations and translations are, as much as possible, decoupled.

Control of rotations : In the peculiar case of vision-based control applied to an aerial vehicle, we can take advantage of the projective properties of the sensor. O.D Faugeras et al. (Faugeras and Luong 2001) and R. Hartley et al. (Hartley and Zisserman 2000) have shown that multiple views provided by a moving camera are related by projective constraints such that any point in the first image has a correspondence in the second image lying on a so-called epipolar line. The fundamental relationship between the i-th point in two different images is:

$$\frac{Z_{2i}}{Z_{1i}}\mathbf{x}_{2i} = {}^2\overset{\infty}{\mathbf{H}}_1 \left(\mathbf{x}_{1i} + \frac{1}{Z_{1i}}\mathbf{c}_1 \right) \tag{10}$$

where ${}^2\overset{\infty}{\mathbf{H}}_1 = \mathbf{K}\,{}^2\mathbf{R}_1\,\mathbf{K}^{-1}$ is the homography of the plane at infinity and $\mathbf{c}_1 \propto \mathbf{Kt}$ is the epipole in the first image ($\mathbf{K}$ is the calibration matrix of the camera).

Let us consider now peculiar points in the scene which belong to the plane at infinity (in our case, the vanishing point and the horizon line), then $Z_{1i} = Z_{2i} = \infty$ and Eq. 10 yields:

$$\overset{\infty}{\mathbf{x}}_{2i} = \mathbf{K}\,{}^2\mathbf{R}_1\,\mathbf{K}^{-1}\,\overset{\infty}{\mathbf{x}}_{1i} \tag{11}$$

From this equation, it appears that the observed motion in the image of the points which belong to the plane at infinity only depends on the rotation part of the camera displacement and are independent on the translations. Thus, controling such points in the image-based visual servoing allows a perfect decoupling of rotations and translations of the camera. To control the three rotations, we have chosen to use (see Figure 3)

as visual signals $\mathbf{s}_1 = [\mathbf{x}_H, \theta_H]^T$, where $\mathbf{x}_H = [x_H, y_H]^T$ are the coordinates of the vanishing point in the image, and θ_H is the angle of the horizon line $\boldsymbol{\Delta}_H = (\rho_H, \theta_H)$ expressed in polar coordinates.

Using visual signals $\mathbf{s}_1$ in Eq. 6 yields the analytical form of the interaction matrix :

$$\dot{\mathbf{s}}_1 = \begin{bmatrix} 0 & \mathbf{L}_{rot1}^T \end{bmatrix} \mathbf{T}_{CT} \tag{12}$$

$$\begin{bmatrix} \dot{x}_H \\ \dot{y}_H \\ \dot{\theta}_H \end{bmatrix} = \begin{bmatrix} x_H y_H & -(1 + x_H^2) & y_H \\ (1 + y_H^2) & -x_H y_H & x_H \\ -\rho_H c\theta_H & -\rho_H s\theta_H & -1 \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix}$$

where $c\theta = \cos\theta$, $s\theta = \sin\theta$ and $(\omega_x, \omega_y, \omega_z)$ is the camera angular velocity.

Control of translations : As was said at the beginning of this section, in a path tracking task, we consider that the longitudinal speed V_x of the aerial vehicle does not require to be controlled from the vision task (the longitudinal speed is indeed to be regulated according to the vehicle requirements expressed in terms of airspeed or angle of attack). So, we want to select visual signals adequate to control the lateral and vertical translations. That can be done by relating the projection of the border lines observed in the image to the lateral and vertical position of the camera computed at the desired reference attitude (i.e. aligned with the road) (Figure 4).

Assuming the width L of the road is known and the optical axis of the camera is aligned with the road, we can compute the angles of the lines $\boldsymbol{\Delta}_r$, $\boldsymbol{\Delta}_l$, depending on the altitude h and the lateral position error e of the camera:

$$\tan(\theta_r) = \frac{L + 2e}{2h}$$

$$\tag{13}$$

$$\tan(\theta_l) = \frac{-L + 2e}{2h}$$

From this equation, a good choice of visual signals for controling the altitude and the lateral position of the camera will be:

$$\mathbf{s}_2 = \begin{bmatrix} t_m \\ \\ t_d \end{bmatrix} = \begin{bmatrix} \frac{(\tan\theta_r + \tan\theta_l)}{2} \\ \\ \frac{(\tan\theta_r - \tan\theta_l)}{2} \end{bmatrix} = \begin{bmatrix} \frac{e}{h} \\ \\ \frac{L}{2h} \end{bmatrix} \tag{14}$$

The two lines are lying on the ground plane at a finite distance, thus the corresponding interaction matrix will depend both on rotations and translations. Derivating the tangent :

$$s = \tan\theta \; \Rightarrow \; \dot{s} = \frac{1}{\cos^2\theta} \, \dot{\theta} \tag{15}$$

and using the general form (Eq. 9), yields an analytical form of the interaction matrix :

$$\dot{\mathbf{s}}_2 = \begin{bmatrix} \mathbf{L}_{tran2}^T & \mathbf{L}_{rot2}^T \end{bmatrix} \; \mathbf{T}_{CT} \tag{16}$$

Finally, combining Eqs. 12 and 16, we obtain the global interaction matrix $\mathbf{L}^T$, which is a lower triangular matrix with good decoupling properties :

$$\dot{\mathbf{s}} = \mathbf{L}^T \, \mathbf{T}_{CT}$$

$$\dot{\mathbf{s}} = \begin{bmatrix} \dot{\mathbf{s}}_1 \\ \\ \dot{\mathbf{s}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{L}_{rot1}^T \\ \\ \mathbf{L}_{tran2}^T & \mathbf{L}_{rot2}^T \end{bmatrix} \begin{bmatrix} \mathbf{V}_{CT} \\ \\ \mathbf{\Omega}_{CT} \end{bmatrix} \tag{17}$$

Computing the desired visual signals $\mathbf{s}^* = [\mathbf{s}_1^* \; \mathbf{s}_2^*]^T$ *:* Since a control in the image is used, the path reference to be tracked by the controller is converted into an image reference to be compared with the visual output and then the error is used by the controller. So, we need to express the Cartesian path tracking task in terms of image features trajectory. Let us consider a road following task at a constant altitude h^*, centered w.r.t the road ($e = 0$) and a constant longitudinal speed V_x^*. We want to be aligned along the median axis of the road, which means in the image: $\theta^* = -\theta_r^* = \theta_l^*$, $x_H^* = 0$ and $\theta_H^* = 0$. If, for simplicity, we assume also that the camera optical axis z_c is horizontal and aligned with the road axis, then $y_H^* = 0$, $\rho_H = \rho_l = \rho_r = 0$, and the equation of the ground plane is such that $(a = 0, b = 1, c = 0, d = h^*)$. Using these values, the desired visual signal is $\mathbf{s}^* = (0, 0, 0, 0, \theta^*)^T$ and the interaction matrix $\mathbf{L}_{\mathbf{s}=\mathbf{s}^*}^T$ computed at the desired position is :

$$\mathbf{L}^T_{\mathbf{s}=\mathbf{s}^*} = \begin{bmatrix} 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{1}{h^*} & 0 & 0 & 0 & 0 & -1 \\ 0 & \frac{L}{2(h^*)^2} & 0 & 0 & 0 & 0 \end{bmatrix} \tag{18}$$

Let us note that only one coupling between the translation and the rotation remains, in the control of the translation along the x-axis of the camera (y-axis of the aircraft). We can also verify that the translation along the optical z-axis of the camera (x-axis of the aircraft) is not controlled by the visual output $\mathbf{s}$.

3.3. Controller Design

The image-based auto-pilot is implemented according to the Block Diagram of Figure 5, with the use of an airspeed sensor and a camera as only feedback sensors, and with a controller regulating the airspeed and tracking the image reference constructed as above.

The idea of a visual control for an unstable platform of 12th order is challenging but implies a great concern with the robustness of the solution. As a first tentative, a solution is searched using optimal control, based on the linearized model of the vehicle dynamics and image output (Eqs. 3 and 17), looking for a pure gain applied on the measured output error:

$$\mathbf{u} = \mathbf{D}\left(\mathbf{y}^* - \mathbf{y}\right) \tag{19}$$

where the output $\mathbf{y} = [u, x_H, y_H, \theta_H, t_d, t_m]^T$ includes the longitudinal speed u and the visual output $\mathbf{s}$.

The non-linear dynamics of the air vehicle (Eq. 1) is first linearized as was described above, around a trim equilibrium state corresponding to a stabilized leveled flight at a constant airspeed (V_o) and at a constant altitude (h_o) above the road and aligned with the road axis, the longitudinal and lateral variables of Eq. 3 are joined in a single state, and the deterministic case is considered:

14

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{20}$$

The state $\mathbf{x} = \left[\mathbf{v}^T, \mathbf{p}^T\right]^T$ of this model includes the 6D change in velocity $\mathbf{v} = [u, v, w, p, q, r]^T$ and 6D change in position $\mathbf{p} = [n, e, d, \phi, \theta, \psi]^T$ of the aircraft frame with respect to the ground frame. The input ($\mathbf{u}$) includes the 3 control surface deflections $(\delta_a, \delta_e, \delta_r)$ and the engine thrust change (δ_T).

The visual output is also linearized, for the same trim condition, using the Jacobian of the image function (or interaction matrix Eq. 18) and including the change from aircraft frame to image frame ($\mathbf{S}_a$):

$$\mathbf{s} = \mathbf{L}^T(h_o)\mathbf{S}_a\mathbf{p} \tag{21}$$

which, together with the longitudinal speed change (u), gives the output equation:

$$\mathbf{y} = \mathbf{C}\left(h_o\right)\mathbf{x} = \begin{bmatrix} \mathbf{C}^u & 0 \\ 0 & \mathbf{L}^T\mathbf{S}_a \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{p} \end{bmatrix} \tag{22}$$

where $\mathbf{C}^u = [1, 0, 0, 0, 0, 0]$ extracts the longitudinal velocity component. This output corresponds to the use of only two sensors: the camera, and an airspeed sensor as an approximate measure of the longitudinal speed error $u \simeq V_t - V_o$.

For a fixed airspeed and altitude, the optimal state feedback gain of the LTI system is obtained with the standard *Matlab* LQR function, corresponding to the minimization of a cost function weighting the output error and control needs:

$$J = \int_0^\infty \left(\mathbf{y}^T\mathbf{Q}\mathbf{y} + \mathbf{u}^T\mathbf{R}\mathbf{u}\right) dt \tag{23}$$

through the definition of the appropriate weighting matrices $\mathbf{Q}$ and $\mathbf{R}$.

The state feedback gain $\mathbf{K}$ is finally transformed into an output error feedback gain using the pseudo-inverse of the output matrix:

$$\mathbf{D} = \mathbf{KC}^{\dagger} \tag{24}$$

PI controller :

In order to reduce the static error appearing in the previous solution with a constant wind disturbance[1], an error integrator $\mathbf{y}_i$ was added in the previous procedure and an augmented output $\mathbf{z}$ is considered:

$$
\begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{y}}_i \end{bmatrix} = \begin{bmatrix} \mathbf{A} & 0 \\ \mathbf{C} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y}_i \end{bmatrix} + \begin{bmatrix} \mathbf{B} \\ 0 \end{bmatrix} \mathbf{u} \\
\mathbf{z} = \begin{bmatrix} \mathbf{C} & 0 \\ 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y}_i \end{bmatrix} \tag{25}
$$

The resulting system was discretized and an optimal output error feedback gain was obtained, yielding the discrete controller:

$$
\begin{aligned}
\mathbf{x}_{k+1}^c &= \mathbf{A}^c \mathbf{x}_k^c + \mathbf{B}^c \left(\mathbf{z}_k^* - \mathbf{z}_k \right) \\
\mathbf{u}_k &= \mathbf{C}^c \mathbf{x}_k^c + \mathbf{D}^c \left(\mathbf{z}_k^* - \mathbf{z}_k \right)
\end{aligned} \tag{26}
$$

Sliding Gain :

Since there is a change in the linearized system as altitude is changing, and namely because the Jacobian of the image and then $\mathbf{C}(h_o)$ is dramatically increasing when the vehicle is near touchdown, the interaction matrix is computed at each sample time according to the current desired altitude and the applied feedback gain is updated:

$$\mathbf{u}_k = \mathbf{C}_k^c \mathbf{x}_k^c + \mathbf{D}_k^c \left(\mathbf{z}_k^* - \mathbf{z}_k \right) \tag{27}$$

The weighting matrices are also corrected as altitude reduces in order to integrate the strict constraints near touchdown.

---

[1]In the case of a lateral wind component, the error cannot be completely cancelled because the reference output $\mathbf{z}^*$ is not a stable trim solution.

Obviously, the visual control is only valid in flight and is switched off once the aircraft has landed and is rolling on the ground.

4. SIMULATION RESULTS

In order to allow for a fair comparison, a similar setup was used for all simulations, in agreement with the Block Diagram presented above. The control was implemented with a $100\ ms$ $(10\ Hz)$ sampling rate.

In agreement with the decoupling properties, the lateral and longitudinal behaviours are initially checked separately. The lateral tracking of the ground road is first presented, with a simulation at constant altitude; a pure longitudinal landing case is then analysed, and a realistic landing is finally considered, with a lateral initial error and wind disturbances.

For all the figures, the relevant parameters during visual control are in solid lines and in dotted lines before visual control or after touchdown. When they exist, the references are presented in dashed lines.

4.1. Lateral Tracking

The lateral behaviour of the visual tracking is analysed in a simulation with a $20\ m$ constant altitude reference, a constant $16\ m/s$ reference airspeed, and the simulation includes the correction of an initial alignment error of $18\ m$ to the right of the road, and then the tracking of a road S turn (see Figure 6). Two sub-cases were considered: a nominal case without wind disturbance, and a wind case, with an intensity of $5\ m/s$, blowing with $15°$ from the left of the path.

The horizontal path presented in Figure 6 clearly shows a smooth alignment and then a very good tracking of the road axis; the motion is fast and well damped. In the nominal case (left), the aircraft is perfectly aligned on the road axis and only the corners of the S shape are a little smoothed. On the other hand, the right figure shows the influence of the wind disturbance, introducing a static error on the horizontal tracking (in this case, an arrow was drawn to indicate the wind direction).

17

In Figure 7 are shown the time evolutions of horizontal and vertical tracking errors, as well as the corresponding visual signals t_m and t_d. The two sub-cases are very similar, the difference only revealing the static error, with $2.8\ m$ in the lateral and $0.18\ m$ in the vertical tracking.

The static error on the visual signal x_H (see the windy case in Figure 8) is due to the characteristics of the aircraft dynamic model when it is flying with no side-slip (natural coordinated flight), expressing that an equilibrium flight above the road with a side wind component necessarily has a heading offset. The other visual signals are well regulated to zero but, as it is clear in Figure 7, the signal t_m is also offset (with a value of 0.14), which corresponds to the lateral tracking error. A little coupling in the altitude is hardly visible in signal t_d.

4.2. Longitudinal Landing

The following landing conditions were assumed:

- start at $700\ m$ away from the touchdown point on the road; after a stabilization period, the landing control is switched on at $500\ m$ from the desired touchdown, with alignment first and then descent;
- initial altitude at $20\ m$ reference;
- initial airspeed at reference speed, equal to $16\ m/s$, which is an airspeed adequate for the approach phase, with a pitch attitude acceptable until touchdown (the model stall speed is slightly below $13\ m/s$).

Two altitude profiles were first considered:

- an usual linear descent at constant sinking speed, with a glide slope of $3°$, and flare for touchdown, with a final reduction of airspeed before ground contact;
- a cosine descent, varying continuously from the initial altitude to touchdown, with also an airspeed reduction for ground contact.

The simulation results comparing the nominal landings without wind and for the two altitude profiles are shown in Figure 9, with, from top to bottom, the altitude and airspeed curves, along with their reference profiles, the pitch angle and the two longitudinal inputs, thrust demand and elevator deflection.

The altitude profiles demonstrate a fair tracking of the reference, with a little lag at the start of descent, smoothly corrected before touchdown. The airspeed profiles exhibit the influence of the descent on the airspeed regulation, but the reference airspeed is still well responding for the deceleration before touchdown. Both the flare in the altitude profiles and the airspeed curves seem to be more precise in the cosine case.

The pitch curves and the input curves clearly show the difference between the two profiles, with the cosine solution giving smoother curves but implying a steeper descent at mid altitude.

The relevant *longitudinal* visual signals, y_H and t_d, respectively associated with the pitch angle and altitude, are presented in Figure 10, along with their references. The characteristics are similar for the two descent profiles, maybe with a closer tracking in the linear case. The vanishing point coordinate exhibits a little overshoot at its maximum value, whereas the t_d signal is fairly well tracked.

In terms of airplane automatic landing, the performance of this nominal case may be analysed through the impact vertical velocity (sinking speed) as presented in Figure 11, which is to be compared to the proposed UAVs regulation limit of $2\ m/s$ (SBAC 1991). Both solutions are well inside the regulation limits, with an impact velocity near $0.2\ m/s$. The flare phase appears however very sudden for the linear case, whereas the cosine profile yields a very continuous curve till touchdown. According to these curves, and looking for a safer touchdown, the cosine profile was then selected for the more realistic landing simulations.

4.3. Realistic landing with wind

In order to have a first evaluation of the validity and robustness of the visual control scheme, a realistic windy landing simulation was run, again from a height of $20\ m$ to touchdown, with a cosine descent and with an initial lateral error of $16\ m$ to the right of the road. The wind conditions were defined with:

- a mean nose component with $5\ m/s$, with a $15°$ angle to the left of the road (the wind intensity corresponds to 31% of the aircraft airspeed, regulated to $16\ m/s$, and is quite significant);

- plus an atmospheric turbulence component, simulated by a Dryden model, with an intensity of $3\ m/s$, which, in a scale from 0 to $7\ m/s$, corresponds to an intermediate gust case.

The simulation results are presented in Figure 12. To some extent, the characteristics of this simulation are in agreement with the pure longitudinal and lateral cases, with the coupling from lateral to longitudinal during the initial alignment (say between $500\ m$ and $400\ m$ before the touchdown point), and then the influence from altitude to the lateral tracking, visible in the diminishing static error in the lateral track error (top-right curve). Globally the landing is well behaved and smooth, with a lateral error of $0.9\ m$ at touchdown and an impact velocity near $0.2\ m/s$ (the roll and yaw angles are respectively $-0.2°$ and $-1.2°$).

The visual signals are presented in Figure 13, showing again characteristics similar to the lateral and longitudinal cases. The influence of the altitude on the lateral tracking is here more visible, mostly on the t_m signal which seems to go out of control: remember however that a trade-off has to be made for touchdown, allowing some lateral error in order to ensure the aircraft attitude is acceptable and permits that the undercarriage touches the ground safely, and then the aircraft starts to roll along the road (using a specific ground controller). The influence of the atmospheric turbulence is also more visible in the visual signals, namely in the angular signals x_H, y_H and θ_H, but this influence remains very little.

5. CONCLUSIONS

In the work described in this paper, a linear structures following task by an aircraft is proposed using image-based visual servoing. This control approach has been analysed in order to set a first exploratory evaluation of a visual servoing technique applied to the automatic landing of an unmanned aircraft (UAV).

The classical image-based approach was adapted to the specific case:

- an adequate scene and image features were selected to allow for the proposed objective, namely decoupling rotation and translation in the image, and allowing to respect the natural separation between the lateral and longitudinal motions of the aircraft;

- the controller design was defined to include the aircraft dynamic characteristics and a sliding gain optimal control was chosen as a first robust solution.

The simulations used to analyse the close loop characteristics and the behavior of the control solution show quite a good performance, well in agreement with the specifications, and the visual-servoing scheme seems

clearly able to land the aircraft in nominal and intermediate wind conditions. The conclusion is then that the idea looks feasible, and clearly justifies further studies to complete the validation and eventually implement such a visual servoing scheme on the real aircraft.

REFERENCES

Azinheira, J., J. Rente, and M. Kellett 1998. Longitudinal auto-landing for a low cost uav. Proceedings of *3rd IFAC Symposium on Intelligent Autonomous Vehicles*, 495–500.

Bourquardez, O. and F. Chaumette 2007. Visual servoing of an airplane for alignment with respect to a runaway. Proceedings of *IEEE Int. Conf. on Robotics and Automation*, 1330–1335.

Chatterji, G., P. Menon, and B. Sridhar 1998. Vision based position and attitude determination for aircraft night landing. *Journal of Guidance, Control and Dynamics 21*(1): 84–91.

Chaumette, F., P. Rives, and B. Espiau 1993. *Visual Servoing*, Chapter Classification and realization of the different vision-based tasks,*World Scientific Series in Robotics and Automated Systems - Vol.7*, Ed. Koichi Hashimoto. Singapore, World Scientific Press.

Costa, N. 1999. Modelação da aeronave robotizada armor x7. Technical report, Instituto Superior Técnico.

Dickmanns, E.1994. The 4D-approach to dynamic machine vision. Proceedings of *33rd Conference on Decision and Control*, 3770–3775.

Espiau, B., F. Chaumette, and P. Rives 1992. A new approach to visual servoing in robotics. *IEEE Trans. on Robotics and Automation 8*(3): 313–326.

Faugeras, O. and Q. Luong 2001. *The geometry of multiple images*. Cambridge, Massachussets, USA: The MIT Press.

Furst, S. and E. Dickmanns 1998. A vision based navigation system for autonomous aircraft. Proceedings of *5th Int. Conf. on Intelligent Autonomous Vehicles -IAS-5*, 765–774.

Hamel, T. and R. Mahony 2002. Visual servoing of an under-actuated dynamic rigid-body system: An image-based approach. *IEEE Trans. on Robotics and Automation 18*(2): 187–198.

Hartley, R. and A. Zisserman 2000. *Multiple View Geometry*. Cambridge, United Kingdom, Cambride University Press.

Kimmett, J., J. Valasek, and J. Junkins 2002. Vision based navigation system for autonomous aerial refueling. Proceedings of *CCA 2002*, 1138–1143.

McLean, D. 1990. *Automatic Flight Control Systems*. New Jersey, USA, Prentice-Hall.

Mejias, L., S. Saripalli, G. S. Sukhatme, and P. Cervera 2005. A visual servoing approach for tracking features in urban areas using an autonomous helicopter. Proceedings of *IEEE Int. Conf. on Robotics and Automation*, 2503–2508 .

Phillips, W. 2004. *Mechanics of Flight*. New Jersey, USA, John Wiley and Sons.

Rathinam, S., Z. Kim, A. Soghikian, and R. Sengupta 2005. Vision based following of locally linear structures using an unmanned aerial vehicle. Proceedings of *44th IEEE Conference on Decision and Control*, 6085–6090.

Rives, P. and J. Azinheira 2004. Linear structures following by an airship using vanishing point and horizon line in a visual servoing scheme. Proceedings of *IEEE Int. Conf. on Robotics and Automation*, Vol. 1, 255–260.

Rives, P., R. Pissard-Gibollet, and L. Pelletier 1996. Sensor-based tasks: From the specification to the control aspects. Proceedings of *6th Int. Symposium on Robotics and Manufacturing (WAC)*, 827–833.

Samson, C., B. Espiau, and M. L. Borgne 1990. *Robot control: the task function approach*. Oxford, United Kingdom, Oxford University Press.

Saripalli, S., J. Montgomery, and G.S.Sukhatme 2003. Visually-guided of an unmanned aerial vehicle. *IEEE Trans. on Robotics and Automation 19*(3): 371–381.

SBAC 1991. *Unmanned Air Vehicles: Guide to the Procurement, Design and Operation of UAV Systems*. SBAC Guided Weapons Ranges Commitee.

Schell, F. and E. Dickmanns 1994. Autonomous landing of airplanes by dynamic machine vision. *Machine Vision and Application 7*(3): 127–134.

Silveira, G., J. R. Azinheira, P. Rives, and S. S. Bueno. 2003. Line following visual servoing for aerial robots combined with complementary sensors. Proceedings of *IEEE International Conference on Advanced Robotics*, 1160-1165.

Stevens, B. and F. Lewis 1992. *Aircraft Control and Simulation*. New Jersey, USA, John Wiley & Sons.

Templeton, T., D. Shim, C. Geyer, and S. Sastry 2007. Autonomous vision-based landing and terrain mapping using an mpc-controlled unmanned rotorcraft. Proceedings of *IEEE International Conference on Robotics and Automation*, 1349–1356.

LIST OF FIGURES

Fig. 1.   Frames and notations

Fig. 2.   Experimental UAV ARMOR X7
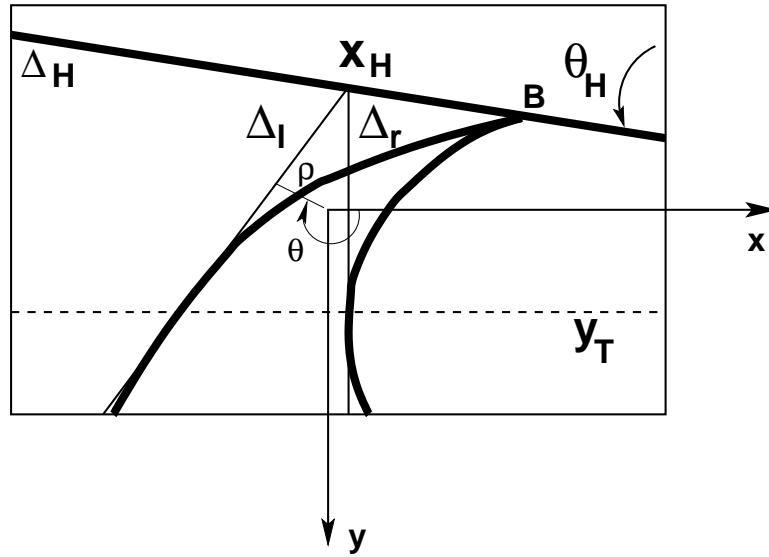
Fig. 3. Image of the road as viewed by the airborne camera
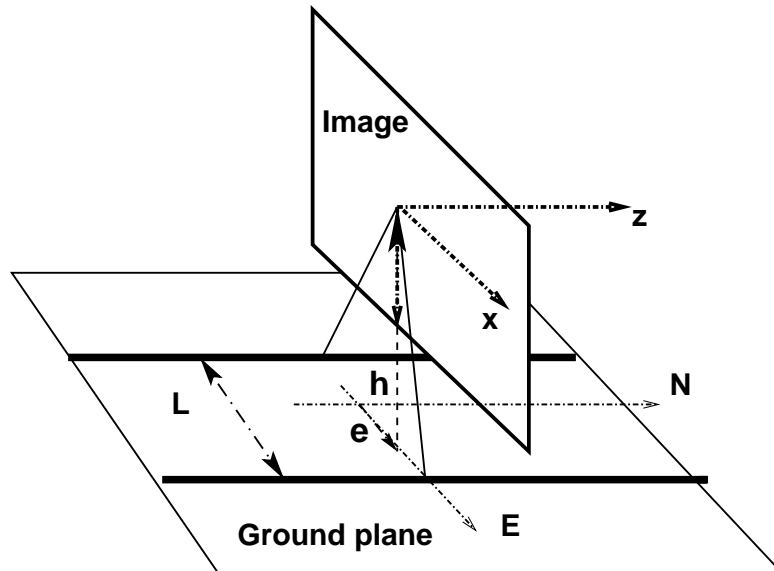
Fig. 4. Computing the reference visual signals: L is the width of the road, h is the altitude and e is the lateral position of the camera w.r.t. the road axis
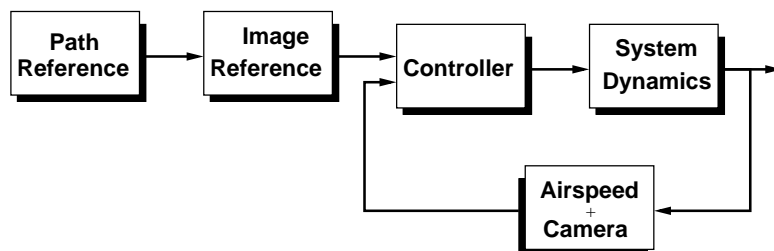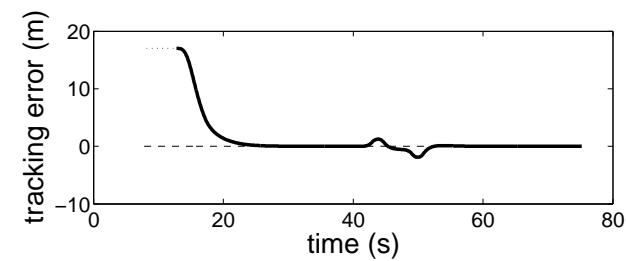

Fig. 5. Image-based Control Block Diagram
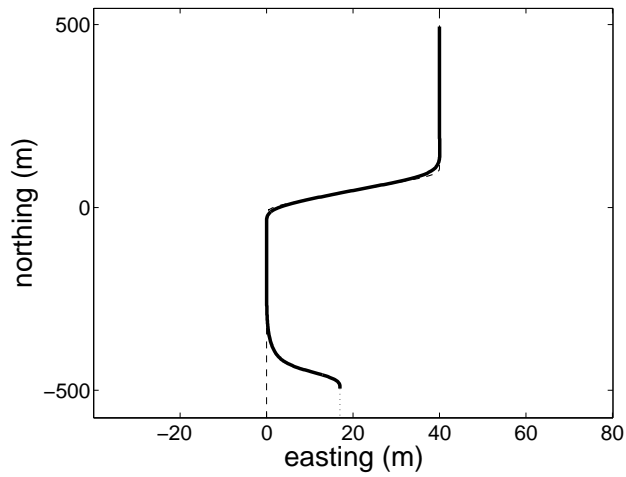
(a) without wind
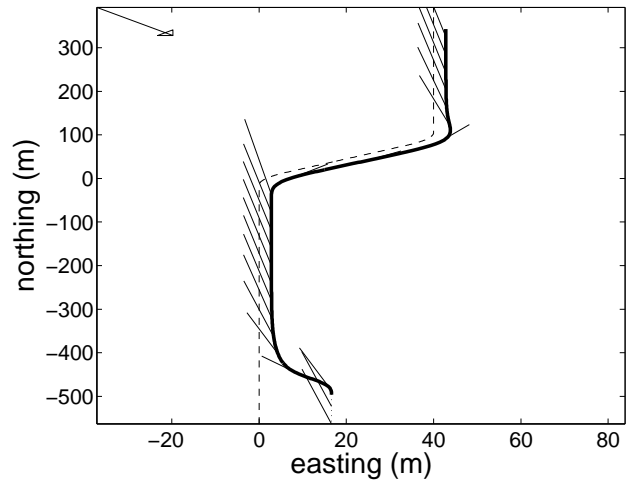
(b) constant wind

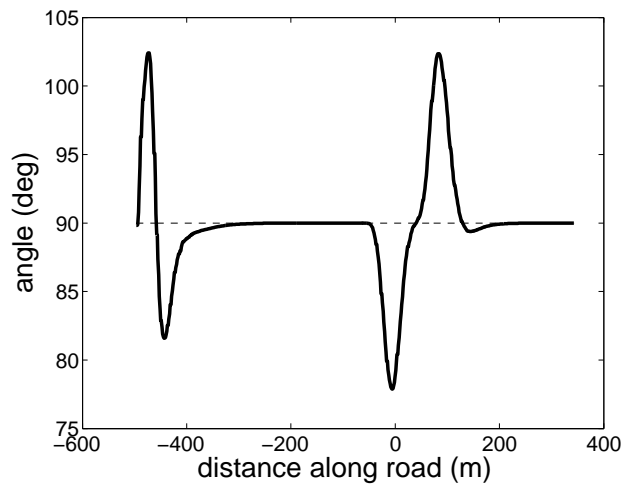Fig. 6. Lateral tracking horizontal path: the arrow indicates the wind direction
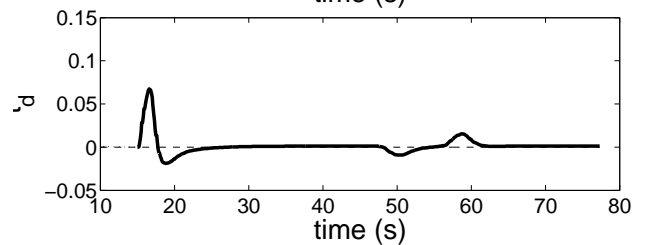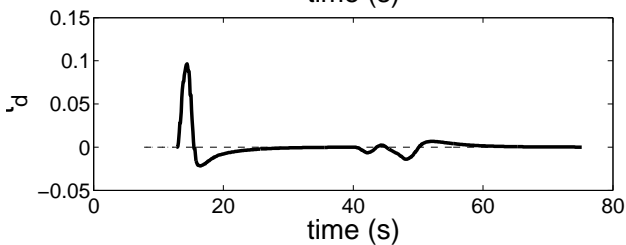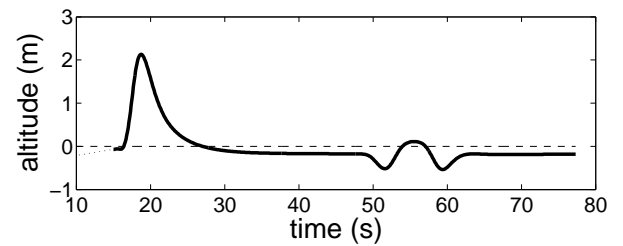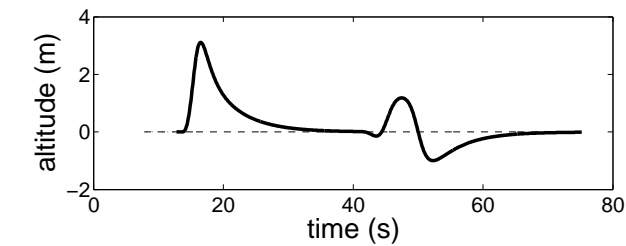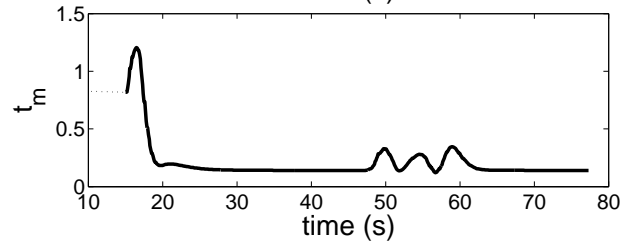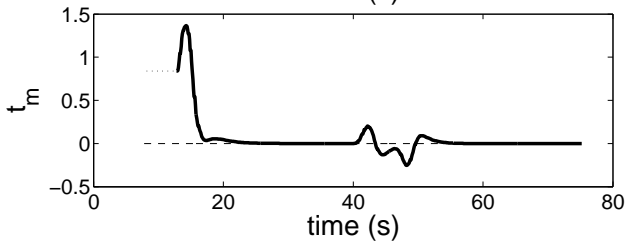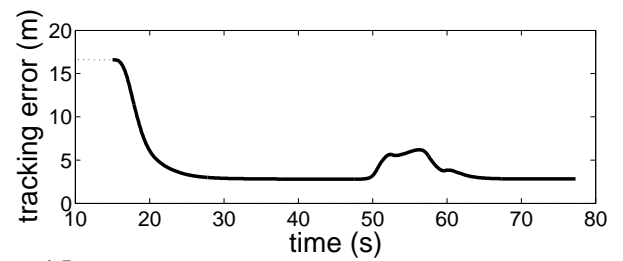
(a) without wind

(b) constant wind

Fig. 7. Lateral tracking: from top to bottom, horizontal tracking error, t_m visual signal, vertical error and t_d visual signal
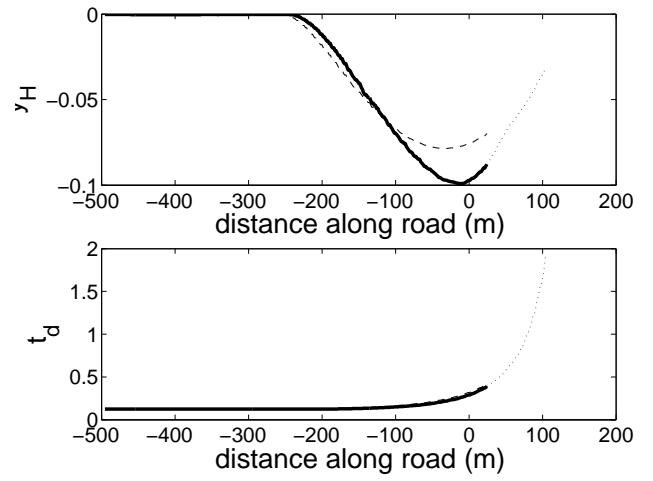
27
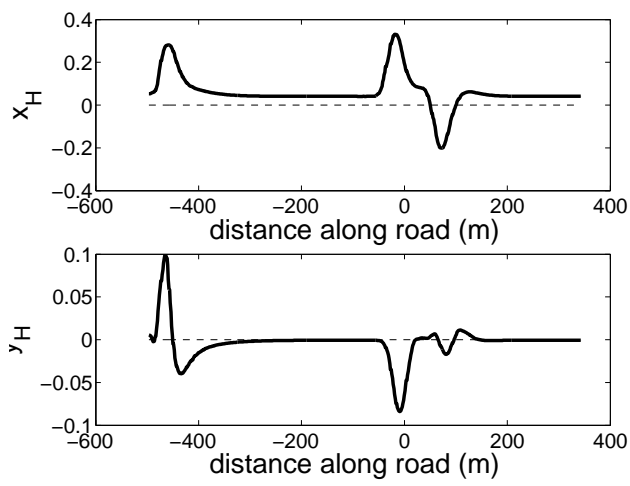
(a) vanishing point coordinates

(b) horizon angle

Fig. 8.   Lateral tracking with wind, angular motion in the image:

(a) with linear descent

(b) with cosine descent

Fig. 9. Longitudinal nominal landing; from top to bottom: altitude, airspeed, pitch angle, thrust demand, elevator deflection
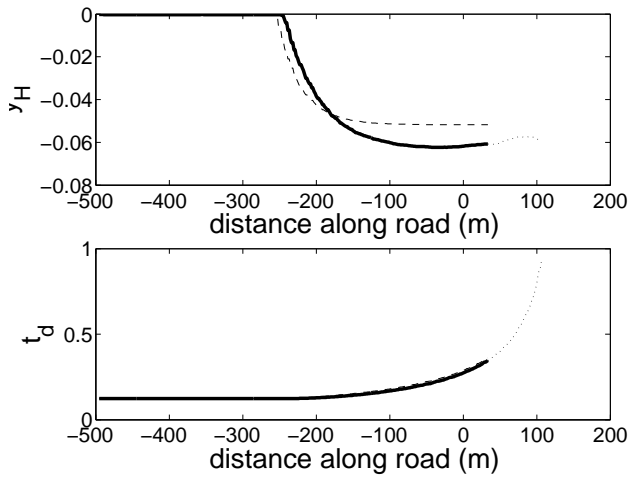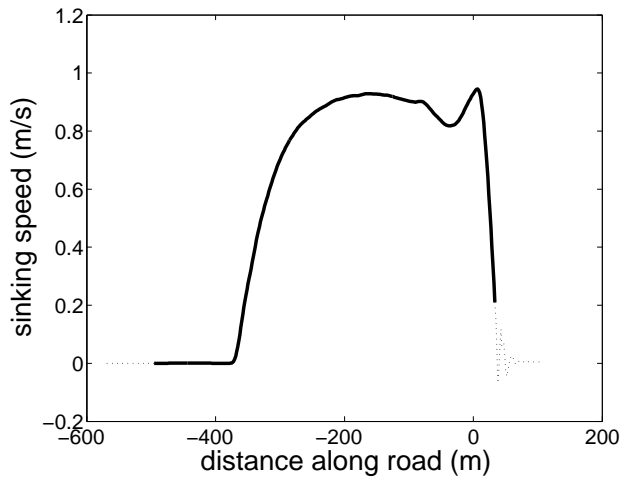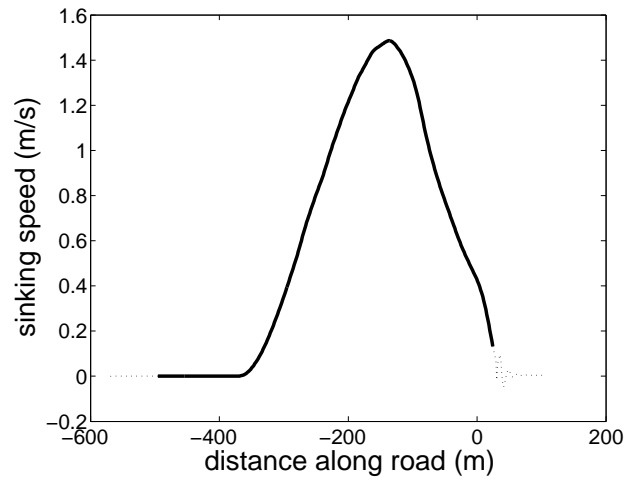
(a) with linear descent

(b) with cosine descent

Fig. 10.   Longitudinal landing, visual signals: y_H (above) and t_d (below)

(a) with linear descent

(b) with cosine descent

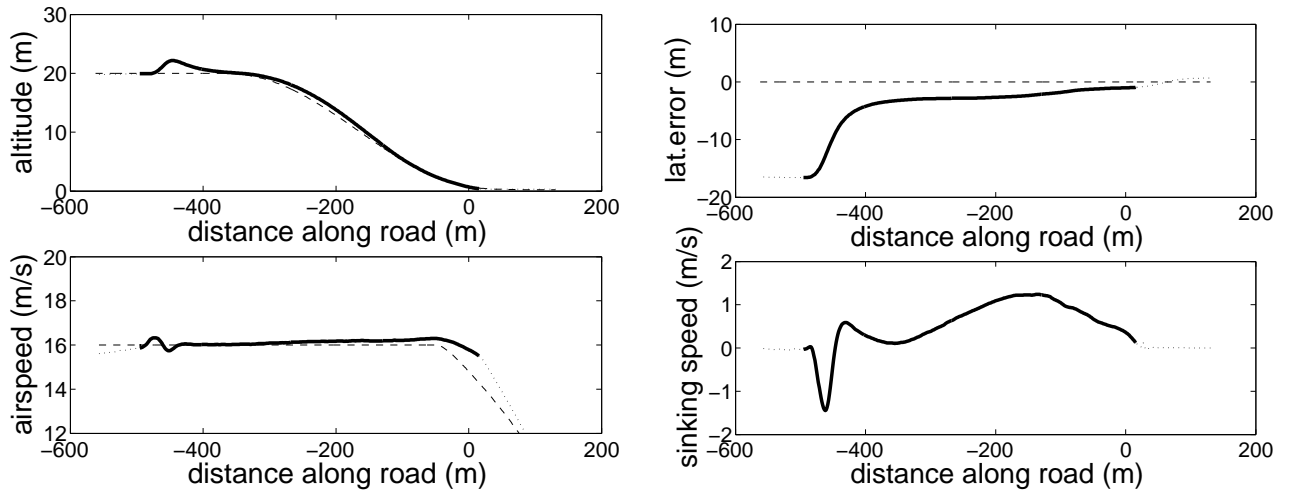Fig. 11.   Longitudinal nominal landing: sinking speed

Fig. 12. Realistic landing with wind disturbances, altitude (top-left), lateral position error (top-right), airspeed (bottom-left) and sinking speed (bottom-right)
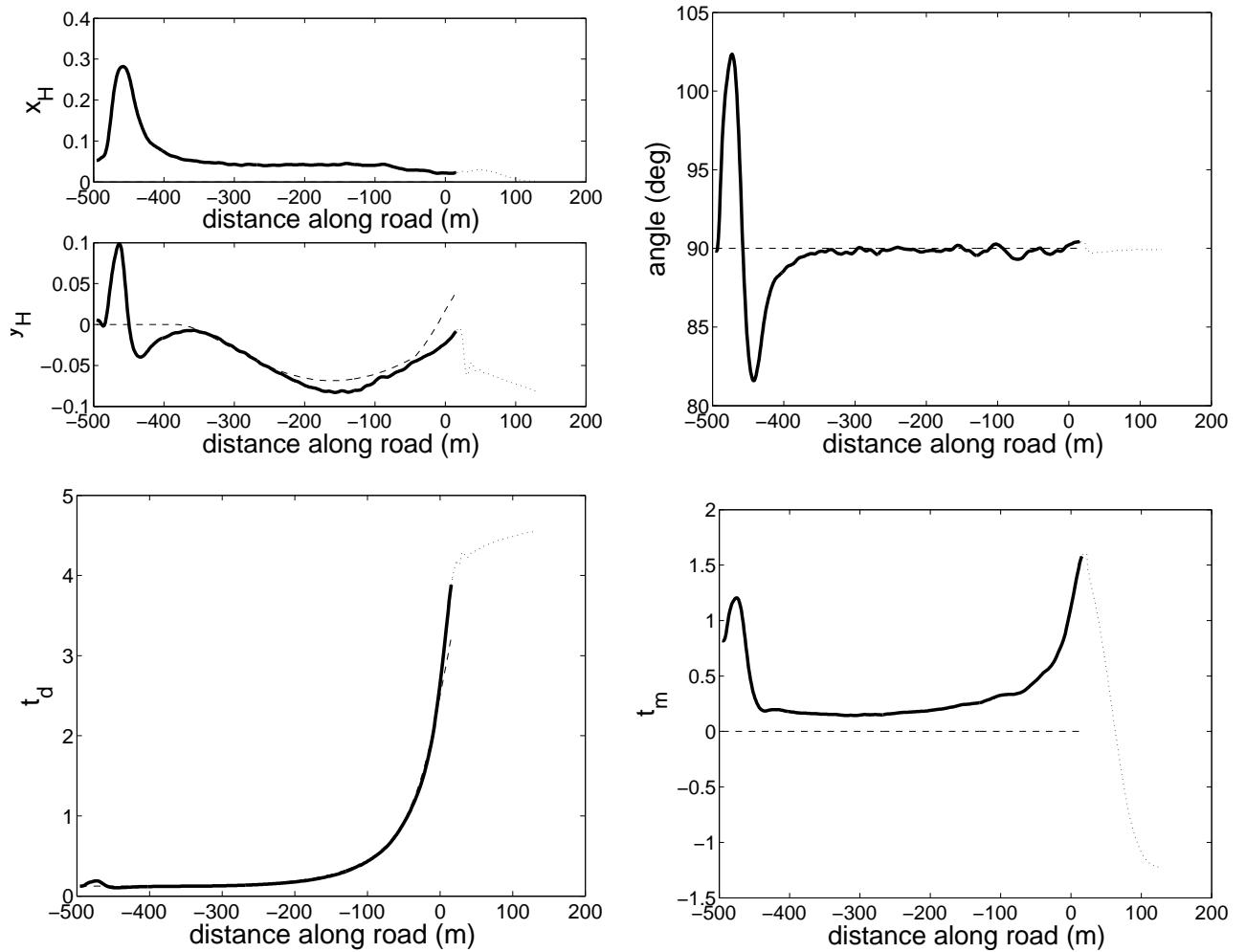

Fig. 13. Realistic landing, visual signals: vanishing point coordinates x_H and y_H, horizon line angle θ_H, t_d and t_m