An Algorithm for Real Time Estimation of the Flexible UAV Structural Motions Using a Video-based System

Vincenzo Rosario Baraniello  
GN&C Department  
CIRA  
Italian Aerospace Research Center  
Capua (CE), Italy.  
v.baraniello@cira.it

Marco Cicala  
Modellistic & Simulation Dept.  
CIRA  
Italian Aerospace Research Center  
Capua (CE), Italy.  
m.cicala@cira.it

Luca Cicala  
Payload Sensor and HMI Dept  
CIRA  
Italian Aerospace Research Center  
Capua (CE), Italy.  
l.cicala@cira.it

Abstract - This paper focuses on the application of a vision based system coupled with an Inertial Measurement Unit (IMU) and several accelerometers, distributed along the aircraft structure, to estimate the structural displacements showed by flexible aircrafts. This sensor suite, associated with the proposed data fusion algorithm, represents an innovative approach to solve the problem of structural displacements determination for flexible aircrafts. The knowledge of the aircraft actual shape allows implementing active control techniques, Structural Health Monitoring (SHM) and moreover to determine the true position and attitude of sensors, installed on board. The effectiveness of the proposed approach is demonstrated through numerical simulations.

Keywords: Structural Motions Estimation, UAV, Kalman filtering.

1 Introduction

This paper presents a new approach to problem of structural displacements estimation for flexible aircrafts. The innovative aspects concern both the sensor suite: specifically the application of a vision-based system associated with an IMU and several accelerometers distributed along the aircraft, and the related data fusion algorithm.  

Motivation for this work has been taken from the consideration that the hypothesis of rigid body cannot be always applied for certain categories of aircrafts. This hypothesis is particularly critic for modern aircrafts, because new design methodologies lead to lighter structures and high aspect ratio wings. With such new aircrafts, a new class of flight control problems has emerged, including: minimization of the loads experienced by the aircraft, suppression of flutter, amplitude reduction of the disturbed motions of an aircraft caused by turbulence. These goals can be accomplished using Active Control Technologies (ACT), proposed to meet the demands for more effective and efficient aircrafts [1]. On the other hand, the purpose of SHM is enhancing the operational life of structures and avoiding failures.

In both these fields, it is usually required a knowledge of the actual aircraft shape, either completely or just at some specific locations.  
Depending on the specific application, in literature, several systems have been proposed in order to measure structural displacements.  
A data fusion algorithm, which allows estimation of structural displacements for very flexible aircrafts, has been proposed in [2]. This algorithm is based on a sensor suite composed by one IMU, two GPS antennas and several distributed accelerometers. The aim of [2] was essentially the inclusion of structural elastic effects in estimation of navigation parameters.  
The use of GPS, standalone or not applying carrier phase differences technique, limits the capability to estimate only structural displacements of extremely flexible aircrafts, showing displacements with an order of magnitude of several meters (as for example NASA HELIOS aircraft [3]), because in these operative modes the GPS shows low accuracy. In [18] it is proposed a control system for spatial flexible structures based differential carrier phase measurements obtained using an array of GPS antennas. The differential carrier phase measurements are used to assess deformations.  
Furthermore differential techniques as RTK (Real Time Kinematic) or carrier phase differences present some disadvantages that will be identified in the following.  
Concerning civil engineering, the Structural Health Monitoring is the main field in which the knowledge of the structural shape is required. Some techniques applied in this field could be applied to the aerospace structures.  
Some works propose GPS integrated with accelerometers in order to estimate displacements of civil structures, taking advantage of their complementary dynamic properties and using accelerometers measurements to recover GPS false or missing data. In [9] the results, of some experimental tests conducted on a test rig simulating the typical dynamics of a bridge, are reported. In [10] and [11] a sensors system, composed by GPS and an accelerometer, is applied to monitor a 108 meters high steel tower during typhoons and earthquakes. In [12] Fiber Bragg Grating (FBGs) optical sensors are included in the measurement system, with GPS and accelerometers. In [13] an array of several GPS antennas is installed on the Pacoima Dam (California) in order to measure its
deformation. The measurements obtained in three years have been analyzed and a mean displacement, between the reference points of the dam, of some tenths of millimetres is reported. This analysis has been performed offline averaging GPS measurements (batch processing); this allows reaching a millimetric accuracy in GPS measurements. In [14] another example of GPS and accelerometers integration for bridge monitoring is provided.

Obviously the batch approach cannot be used in real time application, as for example for the aircraft active control, because the accuracy of GPS standalone in real time is considerably worst than the one available with a post-processing of measurements. RTK technique (which allows obtaining an accuracy of some centimetres) is not always suitable to be used in dynamical environments and moreover the maximum distance between the rover and the base station is limited to 10 km [15]. Furthermore a GPS device, enabled to receive RTK corrections, is more expensive (the price increase depends on the corrections update rate and can reach some thousands of US dollars).

Concerning carrier phase measurements, the resolution of initial integer ambiguity is required. Moreover GPS signal could be lost during some manoeuvres and consequently compromise the control or monitoring action.

To overcome limitations due essentially to GPS and previously presented, integration of a vision-based system with accelerometers and/or GPS has been widely examined and applied above all in civil engineering, specifically for monitoring of bridges and tall buildings. A vision-based system, alone or coupled with accelerometers, provides compactness and affordable price.

Like GPS, a vision-based system allows obtaining a direct measurement of displacement but with a higher accuracy. Furthermore videometry offers the same level of accuracy of contact devices as optical fibers, but at the same time, it is more flexible in usage [5].

In [5] and [7] it is analyzed the application of a vision based sensor system for direct determination of displacement time histories at selected locations on a bridge undergoing to ambient oscillations. The sensor architecture consists in a video-camera coupled with targets composed by a two high resolution low-power light emitting diodes (LEDs) spaced with a known distance. The two targets are installed under the bridge deck near the mid span, while the camera is placed at one of the bridge columns. Measuring the relative motions between the two targets is possible to determine the bridge displacements. The results of a test, performed on Vincent Thomas Bridge in Los Angeles, are reported. Specifically, a displacement time history is reproduced and the frequencies, correspondent to the first two modes, are calculated. According to the authors, the measurements obtained with the vision based system are coherent with the same information obtained using the accelerometer array installed on the bridge. In order to detect modifications in structural properties (specifically dynamic features of a structure) for SHM purposes, in [8] it is presented another application of a vision-based system composed only by a camera. No optical targets are mounted on a bridge, because the applied image processing technique requires the selection of some reference points in the first frame (their coordinates must be provided to the algorithm with a sub-pixel accuracy) and consequently it follows their displacements in the following video frames.

In the aerospace field there are few examples of videometry applications to detect structural displacement. An application of a vision-based system for identification of structural modes of vibration in the aeronautical field is presented in [16]. In this paper two techniques are examined: specifically an online technique based on a multi sensor vision system and an offline technique based on a single camera. Offline measurements based on batches of data are suitable to obtain higher accuracy. The online technique is more suitable to be used in dynamic environments but the different cameras need to be synchronized and the HW is more complex.

Both the approaches presented in [16] are related to experimental identification of wing structural modes through laboratories tests and not to true flights. In the aeronautical field, videometry is also extensively used in wind tunnel tests, to evaluate both aerodynamic and structural properties.

In [4] an application of range imaging sensors is proposed in order to estimate elastic displacements of spatial structures; the proposed system is based on a vision based systems installed on robots orbiting around the space structure. Such a kind of system is not suitable for aeronautical applications in which the video-based system has to be necessarily installed on the aircraft.

In [17] an application of a stereo-vision system for deformation determination of large flexible satellite is presented. In case of satellite the parts, more sensitive to deformations, are the solar arrays or large phased array radars. The application of accelerometers, as only means to determine shape variations, is not suitable due to bias integration and high sensibility to disturbances present in the spatial environment, thus a system which can directly determine displacements could be more robust to these problems and, moreover, a video system is nonintrusive (in the sense that the cameras could be installed on the body of the satellite and not directly on the deformable structure) except for the presence of LED on the structure. A multi-view vision system is composed at least by a pair of cameras opportunely arranged or more cameras depending on the structure dimensions and cameras field of view. Obviously a multi-view vision system induces a greater increment to the overall mass of the system than a single camera system.

The aim of this paper is to extend the possibility to estimate structural displacements to aircrafts which are not extremely flexible, using as observations the output of a vision based system with a target LED positioned in a suitable point of the aircraft. In fact an element of
innovation provided by this paper is the analysis of feasibility and verification of effectiveness, through numerical simulations, of such a system for aeronautical applications.

This paper has been organized in three macro-areas: a description of the system architecture, a description of the algorithm and finally numerical simulations, with whom it will be demonstrated the effectiveness of the proposed approach.

2 Vision-based System Architecture

The vision-based system consists of an emitter-receiver pair operating at visible or infrared wavelength.

In the proposed setup the receiver is a digital camera, the emitter is a Light-Emitting-Diode (LED). The system output consists of the LED image coordinates in the Camera acquired frames.

Both Camera and LED follow linear and rotational deformations of the mounting points.

The Camera Field of View (FOV) is selected to keep the LED image within the frame in the case of maximum deformation.

The operative condition is assumed fine weather and good visibility. Perturbative phenomena of weather conditions able to create refractions or occlusions, as fog or precipitation of various kinds, are neglected.

The model of vision system includes the image projection model and the measurement error model.

In the first one the equations of projective geometry define the transformation of the LED position from the three dimensional space to its projection in the bidimensional frame.

Let be the origin of the Camera reference frame in the optical centre of the camera system, the x-axis (left-right direction) and y-axis (down-up direction) parallel to the sensor surface, the z-axis orthogonal to the sensor surface and with back-front direction with respect to the image plane. The orientation of such reference frame follows the elastic rotation of the Camera mounting point.

Named $x_c$, $y_c$, $z_c$ the LED coordinates in the Camera Reference Frame, the LED coordinates $u$, $v$ respect to the 2D video frame (with origin in the top left corner and positive indices for rows and columns), measured in pixels, may be written as in formulas (1) and (2):

\[
\begin{align*}
    u &= \frac{f_x x_c}{z_c} + u_o, \quad (1) \\
    v &= -\frac{f_y y_c}{z_c} + v_o, \quad (2)
\end{align*}
\]

where $u_o$, $v_o$ are constant off-sets and $f_x$, $f_y$ are the positive values, measured in pixel, of the focal length normalized respectively with respect to the vertical and the horizontal size of a generic photosensitive element of the camera sensor.

\[\text{Figure I Image Reference Frame}\]

In the system model two sources of error have been considered.

The first one is the truncation error due to the pixel quantization in the digital acquisition.

The second one is the approximation error on the 3D position of the LED: $P_c = (x_c, y_c, z_c)$, ideally coincident with the centre of the brightest part of the light spot, that could be not clearly distinguishable in the image. In fact many pixels around the projection $(u, v)$ of $P_c$ can have very similar values. Usually the projection of the light spot centre, under a small error, can be estimated with the barycentre of the pixels with the highest response. Under the previous hypothesis it can be supposed that the coordinates founded by the image processing algorithms are the projection of an immaterial 3D point belonging to an uncorrelated Gaussian distribution centred in $P_c$ and with a standard deviation comparable to the physical dimensions of the LED, identical for every component.

Because the wing range of deformation is known (by experimental evidence or by mechanical considerations), the optical zoom of the camera can be tuned with aim to minimize the field of view.

Special filters can be mounted on the optic in order to increase the sensitivity to the wavelengths emitted by the LED and decrease the sensitivity to other undesired wavelengths.

The area, in which the LED is installed, has to be painted with a low refraction index varnish at the selected wavelengths.

Under the previous setup conditions, the light spot is easily distinguishable due to its high contrast with the background. On the base of projective geometry considerations, the shape and the dimensions in pixels of the light spot projection are approximately known so that it is easy to design a bidimensional digital filter able to detect the LED in the image [20]. Because the aspect of the image patch around the light spot projection, at sufficiently high frame rates, is slowly variant with its location in the image, a simple patch comparison strategy based on block matching algorithms [21] can be implemented. A rectangular patch representing the projection of the light spot in the image is taken as reference template. At the beginning, during the initialization phase, when the wing is still undeformed, the 3D LED position is exactly known, therefore the reference patch can be easily localized around its 2D projection.
During the operation phase, at the end of every update step, the previously used template can be replaced with an updated version, that is obtained by the cut of the current image around the predicted 2D location of the LED projection. These coordinates can be obtained projecting the last 3D position of the LED estimated by the sensor fusion algorithm presented in the paper or otherwise using an independent visual tracking algorithm that estimates the apparent 2D motion of the visual objects in the scene (optical flow) [22]. The reference patch is then compared with small, overlapped, portions of the image that have exactly the same dimensions of the template patch and that belong to a limited search window around the predicted 2D LED coordinates; the search window, thanks to the use of tracking, is usually smaller than the whole image so that the computational burden can be strongly reduced. The minimum admissible dimensions of the search window depend on the projected area of the spot and on the prediction uncertainty. The best match in terms of the reference metric, for example the mean square error, is used to determine the current 2D location of the spot.

The algorithms described above are the same used in video coding [23], currently able to work at 30-60 Hz on Central Processing Units (CPUs) or Digital Signal Processors (DSPs) for consumer electronics applications.

3 Algorithm Description

The final purpose of the proposed approach is the displacement estimation of a certain number of points on the aircraft structure. If the number of points is high, direct estimation of displacements might be unfeasible, because it would require installation of several LEDs with complication in images processing (increasing the computational burden). Moreover, depending on the structure dimensions and interesting points’ distribution along it, one camera could be not sufficient due to limitation in its field of view and consequently more cameras could be required. To overcome this problem, one could exploit the information included in a modal decomposition.

The modal decomposition allows describing displacements as a linear combination of spatial functions “mode shapes”: Φ(xb, yb, zb), which could be determined using the structural analysis or with suitable tests, and time varying functions: “generalized coordinates” η, which can be estimated using the data fusion algorithm, which is proposed in this paper.

We suppose that the mode shapes functions are constant in time, from a physical perspective this implies that mass and the stiffness of the aircraft will not be subject to changes.

Moreover we consider that they will be perfectly known. The estimation of generalized coordinates η associated with the knowledge of the relative mode shapes matrix Φj allows determination of displacement dj, in the body reference frame, for the jth point of the structure.

\[ d_j(t) = \Phi_j \cdot \eta(t) \]  

The data fusion algorithm is based on the Extended Kalman Filter (EKF) technique, because dynamic and observation equations are non-linear in their original form. We consider, as state variables, the generalized coordinates η and their first derivatives \( \dot{\eta} \); while as observations we considered the displacements \( u \) and \( v \) of the LED centre in the images provided by the vision based system.

The expression relating second derivative of generalized coordinates \( \ddot{\eta} \) with \( \eta \) and \( \dot{\eta} \) is non linear and is obtained comparing linear acceleration sensed by the generic accelerometers distributed along the aircraft with the output of a reference one.

In this application, the reference accelerometer corresponds to the accelerometer triad included in the IMU used for navigation purposes, as well as the gyroscope triad which measurements are used for lever arm corrections [2].

The linearized system to which the EKF will be applied is reported in formula (4)

\[
\begin{bmatrix}
\delta \eta \\
\delta v \\
\delta \omega
\end{bmatrix} = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix} \begin{bmatrix}
\delta \eta \\
\delta v \\
\delta \omega
\end{bmatrix} + \begin{bmatrix}
G_{\eta \omega} & G_{v \eta} & G_{\omega v}
\end{bmatrix} \begin{bmatrix}
\nu \\
v \\
\omega
\end{bmatrix}
\]

In formula (4) \( \nu_\omega \) represents the noise on angular speed measurements, the second one \( \nu_v \) is related to linear accelerations measurements sensed by the generic distributed accelerometer and the third one \( \nu_{f,imu} \) is the noise on acceleration measurements sensed by the reference accelerometer; while \( \begin{bmatrix}
\nu_v \\
\nu_\omega
\end{bmatrix} \) is the vector of noises on the output of the vision based system. All noises are considered as white Gaussian Random noises.

In the observation equations it appears the dependence of the LED position in the image on the generalized coordinates.

Specifically this dependence derives from the following sources:

- the actual LED position in the body reference frame;
- the actual position of the origin of the Camera reference frame respect to the Body reference frame;
- the actual axis orientation of the Camera reference frame.

Analyzing the observability matrix, you can figure out that the system is completely observable.

Using directly structural displacements as state variables, the dimension of state vector depends on the number of
points of interest. On the other hand using the modal decomposition, the dimension of state vector depends on model order, independently by the number of points.

The best choice between these two approaches derives from a trade-off analysis, which considers the number of points of interest, the number of generalized coordinates required to correctly describe the structural dynamic.

Furthermore in modal decomposition approach, the number of generalized coordinates influences the number of distributed accelerometers to be installed along the aircraft structure, while using directly structural displacements, the number of LEDs to install depends on the number of points of interest.

It useful to remark that a high number of points of interest could require the installation of several cameras, due to limits of the camera field of view and possible occlusions.

The proposed approach, based on modal decomposition, offers enormous advantages in terms of reduced computational burden, for low order system and for which the number of points of interest is high.

4 Simulation Results

To prove the effectiveness of the proposed data fusion algorithm, several simulations have been carried out in Matlab Simulink® environment. An aircraft simulation model, including an elastic dynamic of customizable order, has been used to generate values of rigid and elastic state variables, consequence of imposed maneuvers. The mode shapes are provided by the FEM structural analysis. The sensors suite includes, a camera with the related LED, four triaxial accelerometers and an IMU. It is useful to remark that the optimal arrangement depends on the model order which you have to estimate; consequently before executing any simulation, the selection of the best arrangement for the four triaxial distributed accelerometers and for the video-based system has been performed. The LED is placed on the wing tip while the Camera is placed on the tail empennage to centre the LED in the undeformed condition.

In Figure II a schematic view of the sensors arrangement on the modeled aircraft has been reported.

In this paper we report the results of a simulation in which the elastic dynamic of the aircraft has been represented with the first two symmetric generalized coordinates $[\eta_1, \eta_2]$ and elevator deflection is imposed in order to excite structural motions (Figure III).

<table>
<thead>
<tr>
<th>Sensors Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camera Field of View</td>
</tr>
<tr>
<td>Camera Resolution</td>
</tr>
<tr>
<td>LED Radius</td>
</tr>
<tr>
<td>Focal Length/ photosensitive element dimension Ratio</td>
</tr>
<tr>
<td>Frame Rate</td>
</tr>
<tr>
<td>Accelerometers Bias</td>
</tr>
<tr>
<td>Accelerometers Noise @ 30 Hz (Standard Deviation)</td>
</tr>
<tr>
<td>Gyrosopes Bias</td>
</tr>
<tr>
<td>Gyrosopes Noise @ 40 Hz (Standard Deviation)</td>
</tr>
</tbody>
</table>

Table I

In Table I, we have summarized the principal parameters describing the features and the error of the simulated sensors. The error model of inertial sensors, used in simulations, is typical of Micro Electro-Mechanical Systems (MEMS) sensors.
In Figure VI it is reported the variance (the diagonal elements of the variance matrix estimated by the Kalman Filter) for both the estimated generalized coordinates and their first derivatives. The variance is a means to assess the accuracy of estimation performed by the filter. A regime the variance for the first and the second generalized coordinate is respectively equal to 0.00042 rad² and 0.0023 rad²; while for the relative derivatives is equal to 0.0012 (rad/sec)² and 0.0078 (rad/sec)². The first generalized coordinate is estimated with greater accuracy, this is due essentially to its greater observability.

In Figure IX, Figure X and Figure XI we have reported the results of LED displacements estimation. This estimation has been performed applying the formula (1) and substituting the generalized coordinates with their estimation. A video-based system composed by one camera, used standalone, cannot provide any information about three dimensional displacements of a generic point, because the number of equations (bidimensional displacements in the image plane) is less than the problem unknowns.

Using only a video-based system, in order to estimate three dimensional displacements, at least two cameras (stereo-vision) have to be included in the setup, with an increase in costs and weights.
Applying the technique proposed in [19], we have evaluated the degree of observability of the proposed state variables. Considering that the number of state variables is equal to four, eigenvalues are comprised in the range $[0 \div 4]$. The smaller eigenvalues is equal to $0.0154$, the corresponding eigenvector represents the direction of higher observability, its components are $[-0.9207; -0.3523; 0.1610; 0.048]$. This eigenvector is direct predominantly along the first state variable, consequently the first generalized coordinate has the greatest observability. This result was expected because of the particular choose of the camera and led installation points, in fact motions of this points are predominantly influenced by the first mode shape. In Figure XII we have reported the eigenvalues of the normalized covariance matrix, in order to show their rapid convergence.

### 5 Conclusion

In this paper an innovative technique for estimation of structural elastic motions of an aeronautical structure has been proposed. The innovative aspects of this technique concern, first of all, the setup, which includes, as aiding device in a sensor fusion algorithm, a vision-based system composed by a Camera-LED pair. This system is extremely compact, thus it allows avoiding complex sensors installation along the aircraft as for example in the application of optical fibers or strain gauges. A certain number of MEMS accelerometers, depending on the model order to be estimated, are distributed along the aircraft. An IMU is also required, but it is already installed onboard for navigation and control purposes. An Extended Kalman Filter has been implemented, in order to estimate generalized coordinates and their first derivatives. The estimation of generalized coordinates, associated with the knowledge of mode shapes, allows estimation of three-dimensional displacements of any point of the structure. This estimation is unfeasible with one standalone pair camera-target, but also with the accelerometers standalone, because of rapid divergence of the solution due to biases integration.

The variance associated with the estimation of both the generalized coordinates and their first derivative has been also analyzed. It depends on the degree of observability of each one of the state variables, and consequently of the generalized coordinates we would like to estimate. The observability is determined essentially by the relative position of the Camera and the LED. The application of the vision-based system, as aiding measurement in a Kalman Filter, allows estimation of displacements having a centimetre order of magnitude, unlike the application of GPS, which, operating in standalone, instead requires displacements with higher magnitude. The application of the proposed approach is suitable to estimate the structural displacements of flexible aircrafts. This information is required for determination of precise orientation of sensors and devices installed onboard, as for example, pointing devices and in general for active control and structural health monitoring.

### References


X. Li, “Integration of GPS, Accelerometer and Optical Fiber Sensors for structural deformation monitoring”
