

AFRL-RW-EG-TR-2010-013

Integrated Multi-Aperture Sensor and Navigation Fusion

Andrey Soloviev
Jimmy Touma
Timothy Klausutis
Mikel Miller
Adam Rutkowski
Kyle Fontaine



AFRL/RWGI
101 W Eglin Blvd
Building 13, Suite 205
Eglin AFB, FL 32542

Feb 2010

FINAL REPORT FOR PERIOD FEB 2008 - SEP 2009

DISTRIBUTION A: Approved for public release; distribution unlimited. 96th
ABW/PA Approval and Clearance # 96ABW-2009-0428, dated 09 Oct 2009

AIR FORCE RESEARCH LABORATORY, MUNITIONS DIRECTORATE

■ Air Force Materiel Command ■ United States Air Force ■ Eglin Air Force Base

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 22-02-2010		2. REPORT TYPE FINAL		3. DATES COVERED (From - To) Feb 2008 - Sep 2009	
4. TITLE AND SUBTITLE Integrated Multi-Aperture Sensor and Navigation Fusion				5a. CONTRACT NUMBER Not Applicable	
				5b. GRANT NUMBER Not Applicable	
				5c. PROGRAM ELEMENT NUMBER 61102F	
6. AUTHOR(S) Andrey Soloviev Jimmy Touma Timothy Klausutis Mikel Miller Adam Rutkowski Kyle Fontaine				5d. PROJECT NUMBER 2303	
				5e. TASK NUMBER PM	
				5f. WORK UNIT NUMBER 80	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AFRL/RWGI 101 W Eglin Blvd Building 13, Rm 205 Eglin AFB, FL 32542				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFRL/RWGI 101 W Eglin Blvd Building 13, Suite 205 Eglin AFB, FL 32542				10. SPONSOR/MONITOR'S ACRONYM(S) AFRL-RW-EG	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-RW-EG-TR-2010-013	
12. DISTRIBUTION / AVAILABILITY STATEMENT DISTRIBUTION A: Approved for public release; distribution unlimited. 96 th ABW/PA Approval and Clearance # 96ABW-2009-0428, dated 09 Oct 2009					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT The integration of vision sensors and inertial navigation system (INS) can enable precision navigation capabilities in the absence of GPS. Inspired from biological systems, a multi-aperture vision processing system allows for accurate self-motion (egomotion) estimation by observing optical flow across all apertures. The multi-aperture approach is particularly well suited for resolving motion-ambiguity by providing a wide field of regard for detecting and tracking visual features (optical flow). This paper presents a data fusion approach for multi-aperture sensors that integrates the vision processing into a single unified frame of reference by projecting imagery from each aperture onto the unit sphere centered on the navigation frame. The unit sphere projection allows for the seamless integration of multiple apertures into the more natural angle-angle space of the navigation frame of reference. As a first step in evaluating the multi-aperture processing strategy, algorithms are evaluated based on simulated data. The results presented clearly show the advantage of coupling the inertial system with a multi-aperture optical system.					
15. SUBJECT TERMS Egomotion estimation, multi-sensor fusion, unit sphere, unified reference frame, optical flow, multi-aperture					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UL	18. NUMBER OF PAGES 21	19a. NAME OF RESPONSIBLE PERSON Jimmy E Touma
a. REPORT UNCLASSIFIED	b. ABSTRACT UNCLASSIFIED	c. THIS PAGE UNCLASSIFIED			19b. TELEPHONE NUMBER (include area code) 850-883-0876

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the 96 Air Base Wing, Public Affairs Office, and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) <<http://www.dtic.mil/dtic/index.html>>.

AFRL-RW-EG-TR-2010-013 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

//ORIGINAL SIGNED//

Dr. Robert Murphey
Technical Advisor, RWGI

//ORIGINAL SIGNED//

Dr. Jimmy E Touma
Program Manager, RWGIA

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

Table of Contents

List of Figures.....	ii
Abstract.....	1
Introduction	2
Unit Sphere Representation of Multi-Aperture Imagery	5
Formulation of Motion Constraint Equations in the Unit Sphere Frame	8
Multi-Aperture/Internal Data Fusion	11
Simulation Results	13
Conclusion	18
References.....	19

List of Figures

Figure 1. Example biological multi-aperture vision system: compound eyes of a dragon fly	2
Figure 2. Single-aperture vs. multi-aperture example.....	3
Figure 3. Multi-aperture experimental setup developed by the Alt-Nav team of the Air Force Research Laboratory's Munitions Directorate; the setup includes three video cameras with a 90-deg separation of their optical axes.....	4
Figure 4. Drawbacks of the Cartesian representation of the unified coordinate frame for the multi-aperture camera case	5
Figure 5. Unit sphere frame and unit sphere feature projection	6
Figure 6. Unit sphere feature projection for the case of non-collocated cameras.....	7
Figure 7. Unit circle representation for a 2D, translation only motion case	8
Figure 8. 3D angular motion represented in the unit sphere frame	10
Figure 9. Simulated indoor environment implemented for the evaluation of multi-aperture vision/INS integration	14
Figure 10. Improved range initialization capabilities using multi-aperture vision: large angular change is accumulated over a small distance for the feature observed by the side-looking camera; this feature is initialized and then applied to correct inertial drift over a longer distance that is required to accumulate an angular change sufficient for initializing the feature observed by the front-looking camera.....	15
Figure 11. Simulation results for the stereo-vision/INS integrated case.....	17

ABSTRACT

The integration of vision sensors and inertial navigation system (INS) can enable precision navigation capabilities in the absence of GPS. Inspired from biological systems, a multi-aperture vision processing system allows for accurate self-motion (egomotion) estimation by observing optical flow across all apertures. The multi-aperture approach is particularly well suited for resolving motion-ambiguity by providing a wide field of regard for detecting and tracking visual features (optical flow). This paper presents a data fusion approach for multi-aperture sensors that integrates the vision processing into a single unified frame of reference by projecting imagery from each aperture onto the unit sphere centered on the navigation frame. The unit sphere projection allows for the seamless integration of multiple apertures into the more natural angle-angle space of the navigation frame of reference.

As a first step in evaluating the multi-aperture processing strategy, algorithms are evaluated based on simulated data. The results presented clearly show the advantage of coupling the inertial system with a multi-aperture optical system.

Introduction

Vision-based navigation techniques serve as a viable option for autonomous, passive navigation and guidance in Global Navigation Satellite Systems (GNSS)-denied environments [1]. This paper discusses extension of the vision-aided inertial navigation approach for multi-aperture camera cases. Inspired from biological systems, a multi-aperture vision processing system allows for accurate motion estimation by observing optical flow across all apertures. The multi-aperture approach is particularly well suited for resolving motion ambiguity by providing a wide field of regard for detecting and tracking visual features. Moreover, as we show in this paper, a multi-aperture system coupled with an inertial system simplifies the problem of obtaining the range to the observed features as our system moves through the environment.

The multi-aperture vision formulation presented herein is inspired by a wide field of regard of insect vision sensors such as compound eyes of a dragonfly shown in Figure 1.



Figure 1. Example biological multi-aperture vision system: compound eyes of a dragon fly

Insects navigate successfully and efficiently in complex environments using optical flow sensors coupled with other sensors such as halteres (inertial sensors). Rather than using a single high-

quality vision device, an eye of an insect can be represented as a combination of multiple low-quality cameras with slightly different fields-of-view and optical characteristics. Moreover, the coupling of vision with other sensors provides more robust egostate estimation than by using vision alone. Therefore, this paper fuses multi-aperture vision sensing with inertial navigation.

Furthermore, the use of multi-aperture vision is also motivated by the enhancement of situational awareness for guidance and mission planning: the capability to look forward is augmented by side-looking and backward-looking capabilities. Figure 2 illustrates the improved situational awareness that is achieved by using multi-aperture vision as compared to a single aperture camera system for an indoor flight mission of a mini-UAV.

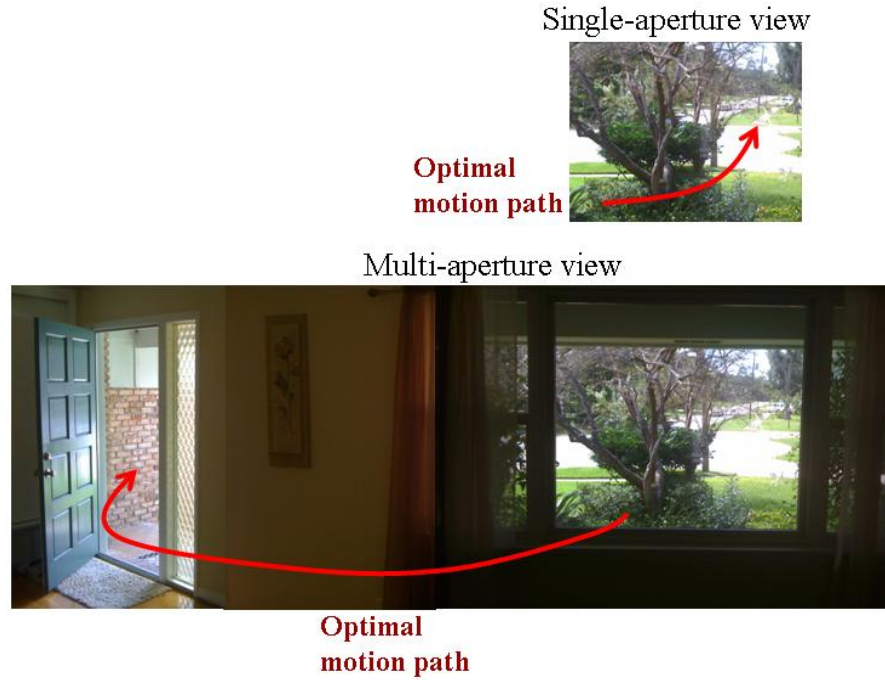


Figure 2. Single-aperture vs. multi-aperture example

In this example, the limited field-of-view of a single aperture results in the mission path that collides into the window. The multi-aperture camera allows for a complete observation of the scene and allows avoiding the collision.

Figure 3 illustrates an example implementation of the multi-aperture camera system.

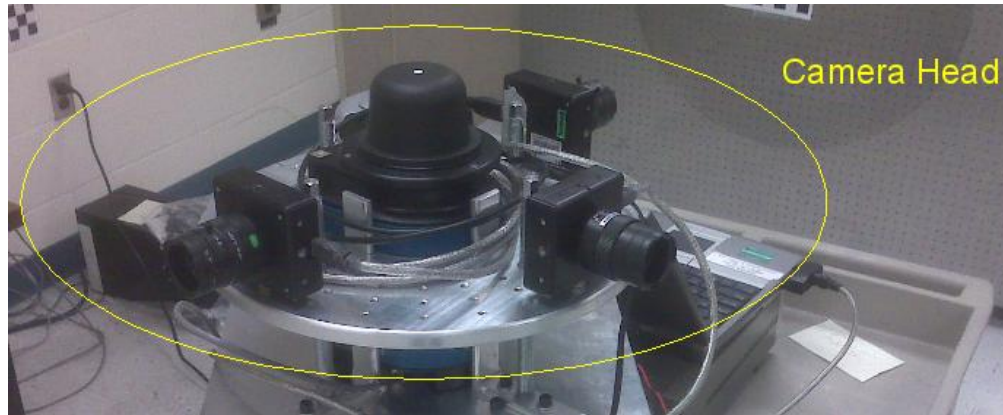


Figure 3. *Multi-aperture experimental setup developed by the Alt-Nav team of the Air Force Research Laboratory's Munitions Directorate; the setup includes three video cameras with a 90-deg separation of their optical axes*

This paper focuses on the benefits of using multi-aperture system for the estimation of vehicle navigation states including position, velocity, and attitude. The main benefits are summarized as follows:

- Increased number of high quality features that can be applied for navigation;
- Improved relative feature geometry, which results in reduced values of dilution of precision (DOP) factors for the estimation of navigation parameters from vision features; and,
- Improved capability to resolve the unknown scale of video images.

The remainder of the paper is organized as follows. First, a unit sphere coordinate frame that is applied to represent multi-aperture imagery is described. Second, motion constraints are derived in the unit sphere frame. Third, fusion of multi-aperture vision data with measurements of an inertial navigation system (INS) is discussed. Finally, simulation results are presented to demonstrate efficiency of the multi-aperture vision approach.

Unit Sphere Representation of Multi-Aperture Imagery

A unified coordinate frame has to be chosen for multi-aperture image processing. Particularly, features observed by different cameras must be converted into a single frame of reference. This paper assumes that point features are extracted from video images and applied for navigation purposes. Considerations herein are readily extendable for other types of features such as line features and planar surfaces.

Each camera originally represents its pixels by their two-dimensional (2D) Cartesian coordinates resolved in the camera focal plane (these coordinates are generally referred to as homogeneous coordinates [2]). In principle, a Cartesian representation can be also adapted for the unified frame, which, in this case, will be represented by a planar surface. However, this representation has a drawback of singularity cases as shown in Figure 4.

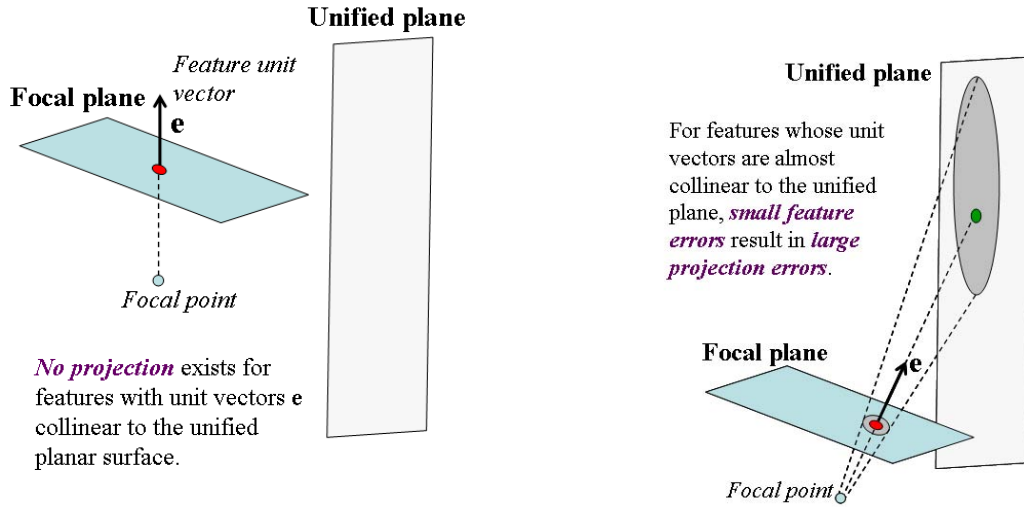


Figure 4. Drawbacks of the Cartesian representation of the unified coordinate frame for the multi-aperture camera case

Homogeneous feature coordinates extracted from an image of a particular camera must be projected onto a planar surface that is chosen for the multi-aperture frame. For those features whose unit vector (i.e., a unit vector pointed from the camera focal point to the feature) is collinear to the unified plane, such a projection does not exist. If the unit vector is almost collinear to the unified plane, then small errors in feature measurements result in large projection errors.

To remove singular cases associated with the Cartesian representation, a unit sphere approach is chosen for multi-aperture cameras. This approach is motivated by nature's multi-aperture vision systems, such as the compound eyes of insects, which process optical flow by projecting it onto a sphere. The unit sphere provides a natural framework for wide field-of-view (>180 degree) sensors, where the traditional focal plane representation is not possible. Figure 5 shows the unit sphere feature projection.

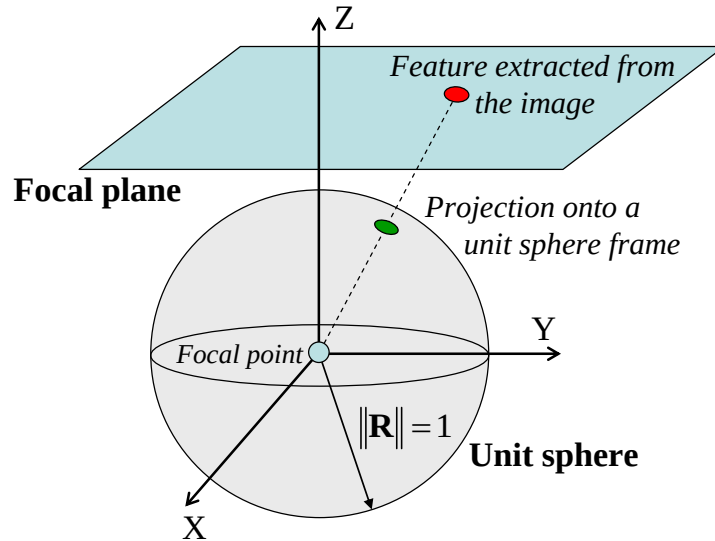


Figure 5. Unit sphere frame and unit sphere feature projection

Features that are extracted from multiple cameras are projected on the surface of a unit sphere and represented by their spherical azimuth and elevation angles. In Figure 5, the unit sphere origin is collocated with the camera's focal point. However, this is not required, and the unit sphere projection can be generalized for cases where cameras' focal points are not collocated with the sphere origin and with each other. Figure 6 illustrates the feature projection for the non-collocated case.

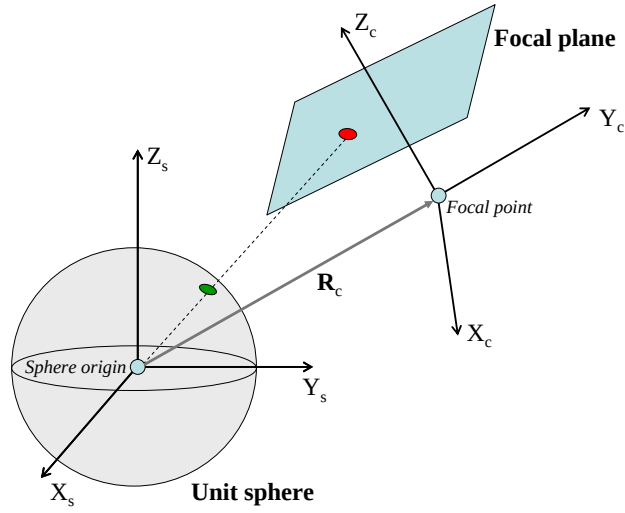


Figure 6. Unit sphere feature projection for the case of non-collocated cameras

Equation (1) defines unit sphere projections:

$$\begin{aligned} \mathbf{R}_f &= \mathbf{C}_c^{us} \cdot [\tilde{m}_x \quad \tilde{m}_y \quad F]^T - \mathbf{R}_c \\ \mathbf{e}_f &= \mathbf{R}_f / \|\mathbf{R}_f\| \\ \varphi &= \arctan(e_{fy}, e_{fx}) \quad \varphi = \arcsin(e_{fz}) \end{aligned} \quad (1)$$

where:

$\mathbf{C}_c^{us}$ is the coordinate transformation matrix from the camera $\{X_c, Y_c, Z_c\}$ -frame into the $\{X_s, Y_s, Z_s\}$ -frame attached to the unit sphere; $\tilde{m}_x, \tilde{m}_y$ are homogeneous feature coordinates extracted from the image;

F is the camera focal length;

$\mathbf{R}_c$ is the vector that originates from the unit sphere origin and ends at the camera focal point; and,

$\arctan(\cdot)$ is the four-quadrant arctangent function.

The main advantage of the unit sphere formulation is the removal of singularity projection cases that are present if Cartesian representation is utilized (see Figure 4 above).

Formulation of Motion Constraint Equations in the Unit Sphere Frame

The use of the unit sphere feature projections requires the transformation of motion equations (i.e., relationships between feature parameters and motion parameters) from the Cartesian frame into the spherical frame. This transformation is illustrated below for a 2D motion case where only a translation motion component is present. In this case, the unified frame is represented by a unit circle as shown in Figure 7.

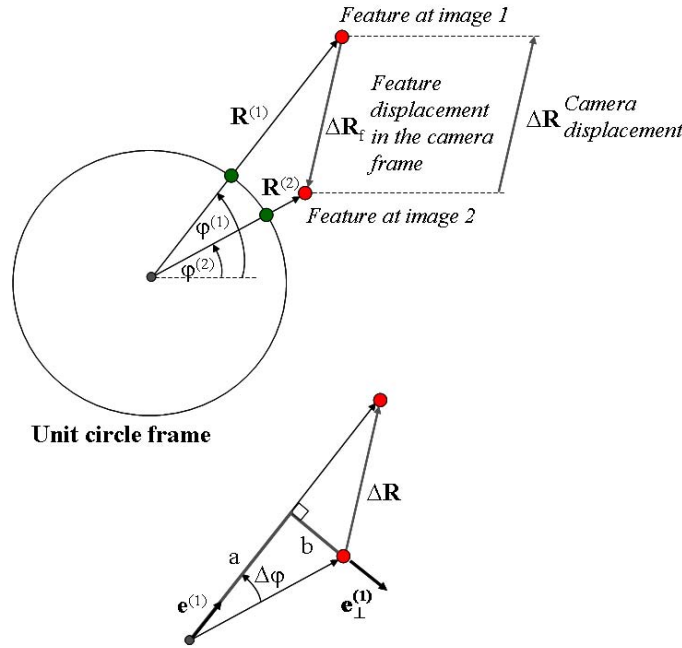


Figure 7. Unit circle representation for a 2D, translation only motion case

From the geometry shown in Figure 3:

$$a / \cos \Delta\varphi = b / \sin \Delta\varphi \quad (2)$$

where:

$$a = \rho^{(1)} - (\Delta\mathbf{R}, \mathbf{e}^{(1)});$$

$$b = -(\Delta\mathbf{R}, \mathbf{e}_{\perp}^{(1)});$$

$$\Delta\varphi = \varphi^{(1)} - \varphi^{(2)};$$

$(\cdot)$ is the vector dot product;

$\rho^{(1)} = \|\mathbf{R}^{(1)}\|$ is the distance to the feature for image 1;

$\mathbf{e}^{(1)}$ is the feature unit vector for image 1; and, $\mathbf{e}_{\perp}^{(1)}$ is the unit vector perpendicular to $\mathbf{e}^{(1)}$.

Correspondingly, the following constraint equation is formulated:

$$(\rho^{(1)} - (\Delta\mathbf{R}, \mathbf{e}^{(1)})) \cdot \sin \Delta\varphi = -(\Delta\mathbf{R}, \mathbf{e}_{\perp}^{(1)}) \cdot \cos \Delta\varphi \quad (3)$$

This equation can be rearranged as follows:

$$\rho^{(1)} \cdot \sin \Delta\varphi = -(\Delta\mathbf{R}, \mathbf{e}_{\perp}^{(2)}) \quad (4)$$

Note that $\Delta\varphi$ is directly calculated from camera measurements. Also, equation (4) only includes the unknown range to the feature at the first image. The knowledge of range information for other images (second, third, etc.) is not required. Once the initial range is estimated, the motion constraint (4) can be applied to formulate Kalman filter vision/inertial measurement observables for other images without the need to know (or measure) their feature ranges. As compared to a complete motion formulation that includes feature range variables for the current image, the motion constraint formulation is beneficial for a monocular camera case, as well as for stereo cases with a limited baseline, where direct measurements of the range value are not possible or inefficient.

The unit sphere formulation is particularly well suited for representing cases of the rotational motion only. Figure 8 exemplifies the rotational motion case.

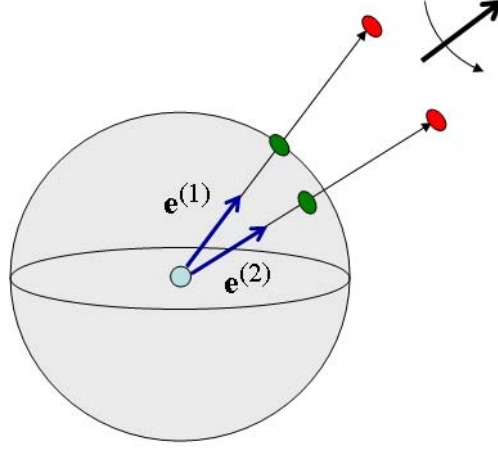


Figure 8. 3D angular motion represented in the unit sphere frame

In this case, camera rotation directly translates into the rotation of the feature unit vector. This is expressed as follows:

$$\Delta \mathbf{C}_b^N \cdot \mathbf{e}^{(2)} = \mathbf{e}^{(1)} \quad (5)$$

For a general 3D case that includes translational and rotational motion components, two motion constraint equations are derived in a manner similar to the 2D consideration above. The first constraint is obtained by considering the horizontal component of the translational motion, and the second constraint is formulated by adding the vertical motion to it. Constraint equations are then modified to incorporate the rotational motion component. The final constraint equations are as follows:

$$\begin{aligned} \left(\mathbf{e}^{(2)} \right)^T \cdot \Delta \mathbf{C}_N^b \cdot \mathbf{B} \cdot \Delta \mathbf{R} &= \left(\mathbf{e}^{(1)} \right)^T \cdot \mathbf{B}^T \cdot \Delta \mathbf{C}_b^N \cdot \mathbf{e}^{(2)} \cdot \rho^{(1)} \\ \left(\mathbf{e}^{(2)} \right)^T \cdot \Delta \mathbf{C}_N^b \cdot \mathbf{D} \cdot \Delta \mathbf{R} &= \left(\mathbf{e}^{(2)} \right)^T \cdot \Delta \mathbf{C}_N^b \cdot \mathbf{e}_{\perp}^{(1)} \cdot \rho^{(1)} \end{aligned} \quad (6)$$

where $\Delta \mathbf{R}$ and $\Delta \mathbf{C}_b^N$ are position and orientation changes between images 1 and 2; and, matrices

B and D are defined as:

$$\mathbf{B} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} 0 & 0 & -\cos(\varphi^{(1)}) \\ 0 & 0 & -\sin(\varphi^{(1)}) \\ \cos(\varphi^{(1)}) & \sin(\varphi^{(1)}) & 0 \end{bmatrix} \quad (7)$$

Multi-Aperture/Internal Data Fusion

Multi-aperture/INS data fusion is formulated in the feature domain using the complementary Kalman filter methodology [3]. In this case, the multi-aperture vision/inertial Kalman filter estimates inertial error states. A fifteen-state INS error model is used (including position errors, velocity errors, attitude errors, gyro drifts and accelerometer biases) and INS error propagation equations [4] are applied to represent the dynamic state model. This section formulates the filter measurement observables.

Motion constraint equations (6) are applied to formulate Kalman filter observables η_p for the multi-aperture vision inertial aiding:

$$\eta_p^{(n)} = \begin{bmatrix} \left(\tilde{\mathbf{e}}_p^{(n)} \right)^T \cdot \Delta \tilde{\mathbf{C}}_N^b \cdot \mathbf{B} \cdot \Delta \tilde{\mathbf{R}}_{INS} - \left(\tilde{\mathbf{e}}_p^{(1)} \right)^T \cdot \mathbf{B}^T \cdot \Delta \tilde{\mathbf{C}}_b^N \cdot \tilde{\mathbf{e}}_p^{(n)} \cdot \hat{\rho}_p^{(1)} \\ \left(\tilde{\mathbf{e}}_p^{(n)} \right)^T \cdot \Delta \tilde{\mathbf{C}}_N^b \cdot \tilde{\mathbf{D}} \cdot \Delta \tilde{\mathbf{R}}_{INS} - \left(\tilde{\mathbf{e}}_p^{(n)} \right)^T \cdot \Delta \tilde{\mathbf{C}}_N^b \cdot \tilde{\mathbf{e}}_{\perp p}^{(1)} \cdot \hat{\rho}_p^{(1)} \end{bmatrix} \quad (8)$$

$p = 1, \dots, P$

where:

$\tilde{\mathbf{e}}_p^{(s)}$, $p=1, \dots, P$, $s=1, 2$ is the unit vector of the feature p that is extracted from image s ; $\hat{\rho}_p^{(1)}$ is the estimated feature range for image 1, and, $\Delta \tilde{\mathbf{R}}_{INS}$ and $\Delta \tilde{\mathbf{C}}_b^N$ are INS position and orientation changes between images. Image 1 is generally the image where the feature was first observed and image n is the current image.

Equation (8) is linearized to support the linear formulation of the complementary Kalman filter:

$$\begin{aligned} \eta_p^{(n)} = & \mathbf{H}_{p,\Delta\mathbf{R}} \delta\mathbf{R}_{\text{INS}} + \mathbf{H}_{p,\alpha} \delta\boldsymbol{\psi}_{\text{INS}} \\ & + \mathbf{H}_{p,f} \begin{bmatrix} \delta\phi_p^{(1)} \\ \delta\theta_p^{(1)} \\ \delta\rho_p^{(1)} \end{bmatrix} + \mathbf{H}_{p,\varepsilon} \begin{bmatrix} \delta\phi_p^{(n)} \\ \delta\theta_p^{(n)} \end{bmatrix} \quad (9) \\ p = & 1, \dots, P \end{aligned}$$

where $\delta\mathbf{R}_{\text{INS}}$ and $\delta\boldsymbol{\psi}_{\text{INS}}$ are INS position and attitude errors; $\delta\phi_p^{(1)}, \delta\theta_p^{(1)}, \delta\rho_p^{(1)}$ are feature state errors (including initial angular errors and initial range errors); and, $\delta\phi_p^{(n)}, \delta\theta_p^{(n)}$ are vision measurement errors for the current update epoch. The filter state vector includes both inertial error states and feature error states. As mentioned previously, the inertial error state vector is a fifteen-state vector comprised of position error states (three states), velocity error states (three states), attitude error states (three states), gyro drift states (three states), and accelerometer bias states (three states). Incorporation of vision feature error states into the Kalman filter state vector is similar to simultaneous localization and mapping (SLAM) formulations (see, for example, [5]).

The initial feature range $\rho^{(1)}$ is estimated by observing the feature from two different locations of the platform and using inertial displacement and orientation change measurements. This estimation is derived from Equation (8) and is formulated as follows:

$$\begin{aligned} \hat{\rho}_p^{(1)} = & (\mathbf{H}_p^T \cdot \mathbf{H}_p)^{-1} \cdot \mathbf{H}_p^T \cdot \begin{bmatrix} (\tilde{\mathbf{e}}_p^{(2)})^T \cdot \tilde{\mathbf{C}}_N^b \cdot \mathbf{B} \cdot \Delta\tilde{\mathbf{R}}_{\text{INS}} \\ (\tilde{\mathbf{e}}_p^{(2)})^T \cdot \tilde{\mathbf{C}}_N^b \cdot \mathbf{D}_p \cdot \Delta\tilde{\mathbf{R}}_{\text{INS}} \end{bmatrix} \\ \mathbf{H}_p = & \begin{bmatrix} (\tilde{\mathbf{e}}_p^{(1)})^T \cdot \mathbf{B}^T \cdot \Delta\tilde{\mathbf{C}}_N^b \cdot \tilde{\mathbf{e}}_p^{(2)} \\ (\tilde{\mathbf{e}}_p^{(2)})^T \cdot \Delta\tilde{\mathbf{C}}_N^b \cdot \tilde{\mathbf{e}}_{\perp p}^{(1)} \end{bmatrix} \quad (10) \end{aligned}$$

Essentially, the image depth is initialized using synthetic stereo-vision: camera motion is applied to synthesize a stereo-vision baseline, which is measured by the INS. Note that in this case a

correlation between depth estimation errors and inertial position and orientation errors is introduced. This correlation must be taken into account by the Kalman filter design. Particularly, the filter states include INS error states, and errors in initial range estimates. Hence, the correlation between INS position and orientation errors and range estimation errors must be incorporated into the state covariance matrix of the Kalman filter. An approach for computing the INS/vision correlation for the multi-aperture case is similar to the computation of INS/vision correlation for a single aperture case, which is described in [1].

Simulation Results

Performance of multi-aperture/vision inertial aiding was assessed in a simulated indoor environment. Figure 9 shows a 2D projection of the simulation environment. The simulation scenario included motion in an indoor hallway. Hallway walls were simulated as vertical with a 2.5-m height. For each wall, four point features were uniformly distributed over its length and height. The simulated motion trajectory was two-dimensional. However, the vision/inertial system estimated all six degrees of motion freedom. Inertial measurement unit was simulated as a lower-grade unit with an accelerometer bias of 1 mg and a gyro drift stability of 50 deg/hr. Specifications of video cameras were simulated as follows: 640x480 resolution, 10 Hz update rate, 40 deg by 30 deg field-of-view (FoV). Errors of extracting a point feature from camera images were simulated as a camera quantization error plus a random Gaussian noise error with a standard deviation of one pixel.

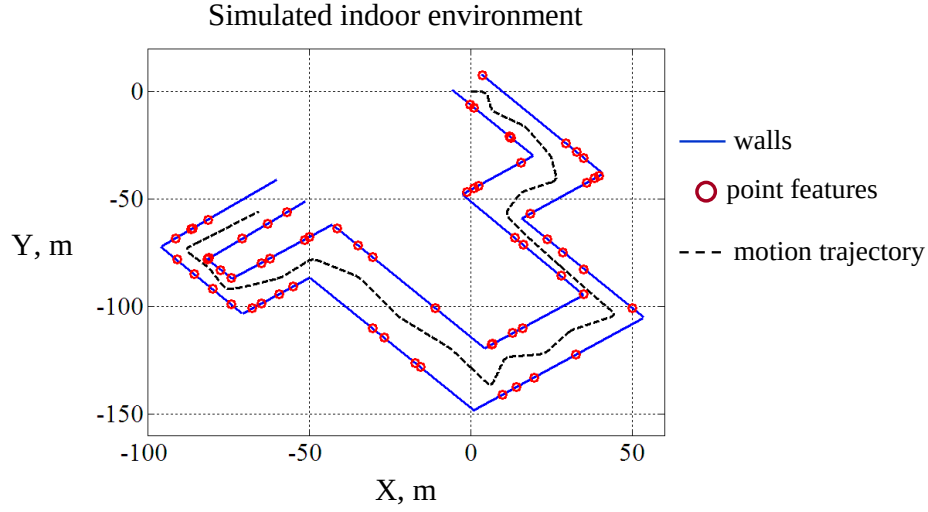
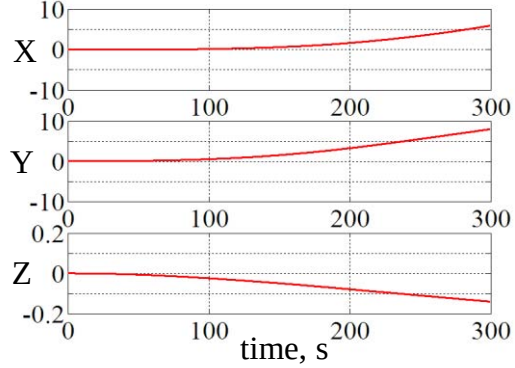


Figure 9. Simulated indoor environment implemented for the evaluation of multi-aperture vision/INS integration

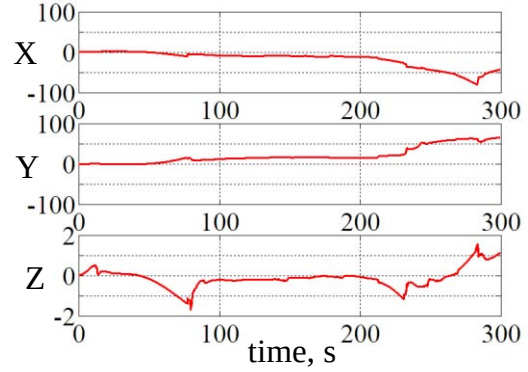
Figure 10 exemplifies simulation results. Position accuracy performance is shown for cases of a free-running inertial, a single camera/INS integration, and a multi-aperture vision/INS integration including cases of three and eight apertures. For the multi-aperture cases, the angular separation between individual cameras was implemented as 120 deg and 45 deg for the three-aperture and eight-aperture cases, accordingly.

Simulation results presented demonstrate that for the simulation scenario implemented the use of multi-aperture INS aiding allows reducing position errors by two orders of magnitude as compared to the single-aperture case: position drift is reduced from a 100-m level to a 1-m level. The accuracy performance is improved due to the following factors. Firstly, the number of available features generally increases by increasing the camera FoV using multiple apertures. Secondly, and more importantly, the use of multiple apertures enhances the feature range initialization capabilities. Figure 11 illustrates this effect.

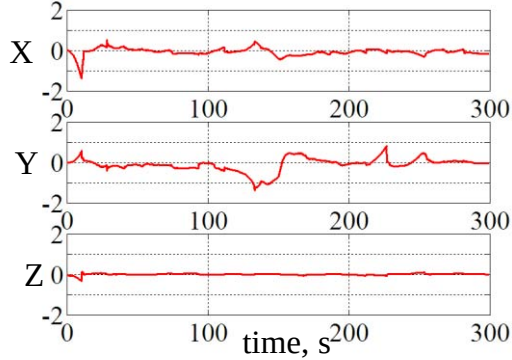
Free-running INS: position errors, km



Single aperture/INS: position errors, m



Three apertures/INS: position errors, m



Eight apertures/INS: position errors, m

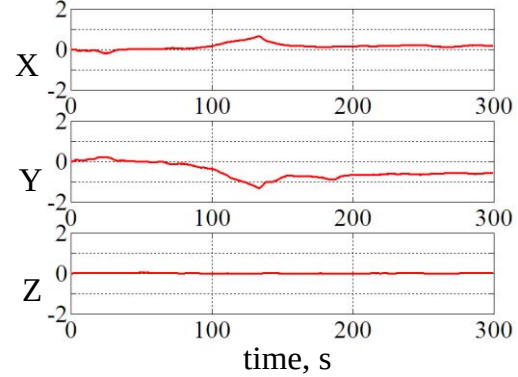


Figure 10. Example simulation results for the indoor simulation environment

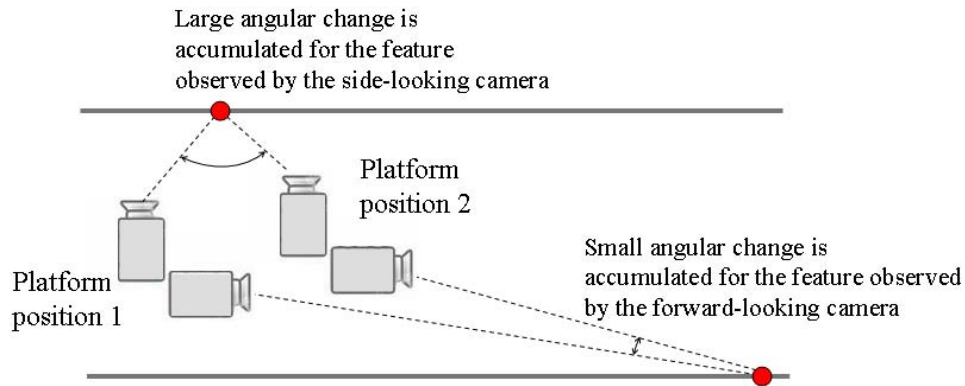


Figure 10. Improved range initialization capabilities using multi-aperture vision: large angular change is accumulated over a small distance for the feature observed by the side-looking camera; this feature is initialized and then applied to correct inertial drift over a longer distance that is required to accumulate an angular change sufficient for initializing the feature observed by the front-looking camera

The range estimation accuracy is primarily influenced by the feature angular change between images that are used for the estimation. The determinant of the $\mathbf{H}_p^T \cdot \mathbf{H}_p$ matrix in the pseudo-inverse formulation of equation (10) is directly related to changes in spherical angles. An increased angular change increases the determinant value and thus improves the estimation accuracy. The maximum angular change is accumulated when passing over a feature and observing this feature with a side-looking camera as shown in Figure 11. As a result, this feature is efficiently initialized after a small platform displacement. Kalman filter observables for this feature are then applied to correct inertial drift over longer displacement intervals that are required in order to accumulate an angular change that is sufficient to initialize the feature range for the front-looking camera.

To illustrate the efficacy of using the multi-aperture vision approach for the range initialization, position accuracy of the multi-aperture/INS system was compared to performance of a stereo-vision-aided inertial. For the stereo case, a two-camera system was implemented. Images of both cameras are converted into a unit sphere frame and process of that frame. The stereo-vision principle is applied to initialize unknown ranges of video features. Performance of the stereo-vision system with a 1-m baseline was accessed in the simulated indoor environment shown in Figure 9.

Figure 12 shows position performance of the stereo-vision-aided inertial. In this case, position errors stay at a meter-level and the maximum error in horizontal position components is 4 m, approximately.

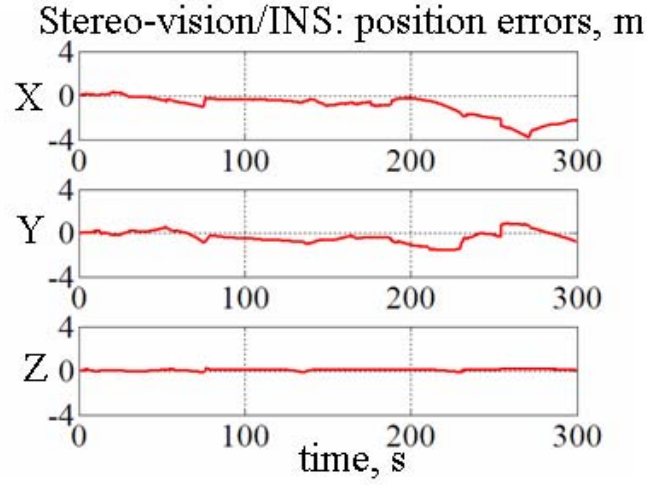


Figure 11. *Simulation results for the stereo-vision/INS integrated case*

Simulation results shown in Figures 10 and 12 demonstrate that similar to the multi-aperture case, the use of stereo-vision allows for meter-level positioning accuracy. It is noted that multi-aperture vision still delivers a better performance: maximum position error is about 2 m for multi-aperture case and about 4 m for the stereo-vision case. Hence, it can be inferred that the use of multi-aperture imagery allows for range initialization capabilities that are at the same level or better as compared to stereo-vision systems with long baselines. The main advantage of the multi-aperture approach is its scalability and direct applicability for small platform application scenarios where large baseline cannot be utilized.

Overall, simulation results presented in this section demonstrate the efficiency of the multi-aperture vision/INS data fusion methodology that is developed in this paper.

Conclusion

This paper proposes the use of multi-aperture imagery integrated with inertial measurements for navigation in GNSS-denied environments. A unit sphere representation of features that are extracted multi-aperture images was developed. Motion constraints that relate feature parameters to the motion of cameras were represented in the unit sphere frame. The unit sphere motion constraints were then applied to formulate multi-aperture vision/INS data fusion. Performance of the integrated solution was assessed in the simulated indoor environment.

Future efforts will evaluate performance characteristics of the multi-aperture/inertial integration using experimental data including urban ground data and flight test data.

References

- [1] M. J. Veth “Fusion of Imaging and Inertial Sensors for Navigation,” Ph.D. Dissertation, Air Force Institute of Technology, September, 2006.
- [2] Y. Ma, S. Soatto, J. Kosecka, S. Sastry, *An Invitation to 3-D Visio*, Springer-Verlag Inc., New York, 2004.
- [3] R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering, Third Edition*, John Wiley & Sons, Inc., New York, 1997.
- [4] J. L. Farrell, “GPS/INS-Streamlined,” NAVIGATION, Journal of the Institute of Navigation, Vol. 49, No. 4, Summer 2002.
- [5] J.A. Castellanos and J.D. Tardós, *Mobile Robot Localization and Map Building: A Multisensor Fusion Approach*, Kluwer Academic Publishers, Boston, 1999.

DISTRIBUTION LIST
AFRL-RW-EG-TR-2010-013

DEFENSE TECHNICAL INFORMATION CENTER
ATTN: DTIC-OCA (ACQUISITION)
8725 JOHN J. KINGMAN ROAD, SUITE 0944
FT. BELVOIR VA 22060-6218

NAWC/WD (BRENT HEDMAN)
CODE 472000D
1900 NORTH KNOX ROAD, MS 6611
CHINA LAKE CA 93555-6106

NAWC/WD (RON SCOTT)
CODE 472100D
1900 NORTH KNOX ROAD, MS 6612
CHINA LAKE CA 93555-6106

NAWC/WD (SAM GHALEB)
CODE 474300D
1900 NORTH KNOX ROAD
CHINA LAKE CA 93555-6106

COMDR, U.S. ARMY AVIATION AND MISSILE COMMAND
ATTN: AMSAM-RD-MG-SP (DR. PAUL RUFFIN)
REDSTONE ARSENAL AL 35898-5241

EGLIN AFB OFFICES:
AFRL/RWOC-1 (STINFO Office)
AFRL/RWG (2 copies)
AFRL/RWGG
AFRL/RWGI
AFRL/RWGN
AFRL/RWGS
AFRL/RWA
AFRL/RWM
AFRL/RW/CA-N (Provided with Notice of Publication Only)

ONE COPY OF EACH UNLESS OTHERWISE SPECIFIED