

Two-Dimensional Stochastic Projections for Tight Integration of Optical and Inertial Sensors for Navigation

Major M. Veth, *Air Force Institute of Technology*
J. Raquet, *Air Force Institute of Technology*

BIOGRAPHY

Major Mike Veth is a Ph.D. student in the Department of Electrical and Computer Engineering at the Air Force Institute of Technology. His current research focus is on the fusion of optical and inertial systems. He received his M.S. in Electrical Engineering from the Air Force Institute of Technology, and a B.S. in Electrical Engineering from Purdue University. In addition, Major Veth is a graduate of the Air Force Test Pilot School.

John Raquet is an associate professor in the Department of Electrical and Computer Engineering at the Air Force Institute of Technology, where he is also the Director of the Advanced Navigation Technology Center. He has been involved in navigation-related research for over 15 years.

ABSTRACT

Aircraft navigation information (position, velocity, and attitude) can be determined using optical measurements from imaging sensors combined with an inertial navigation system. This can be accomplished by tracking the locations of optical features in multiple images and using the resulting geometry to estimate and remove inertial errors.

A critical factor governing the performance of optical-inertial navigation systems is the robustness of the feature tracking algorithm. Robust feature tracking research has focused on developing multi-dimensional feature transformations which are invariant to camera pose variations. In addition, significant effort has been placed into algorithms designed to pair features between images from large sets (e.g., RANSAC). This traditional approach requires large computational resources, especially when presented with imaging situations with sparse, partially obscured, or repetitive features.

In this paper, the method of multi-dimensional stochastic constraints is applied to the optical-inertial navigation

problem in two dimensional feature space. The resulting navigation system uses inertial measurements to aid the feature tracking algorithm, which results in improvements in robustness and processing speed. The performance of the optical-inertial navigation system is demonstrated using experimental data.

INTRODUCTION

Motivation

The benefits of tightly integrating navigation sensors, such as inertial measurement units (IMU) and global positioning system (GPS) receivers, is well-known. The complementary characteristics of the two sensors allow the integrated system to perform at levels which are difficult to attain with either sensor alone (see [2]). As a result, integrated IMU/GPS systems have become common, especially in military-grade navigation systems. Unfortunately, GPS signals are not available in all locations, which motivates the development of a non-GPS based navigation reference which can aid an inertial navigation system.

One non-GPS navigation approach is to integrate a camera with an inertial sensor. This technique has some important advantages. Foremost, the sensors can operate in environments where GPS is difficult to receive (e.g., indoors, under trees, underwater, etc.). Secondly, the sensors are completely passive and do not require the transmission (or reception) of radio signals. Finally, optical and inertial sensors are immune to disruptions in the radio spectrum.

The development of low-cost inertial and optical sensors has led to remarkable advances in the field of optical-aided navigation [15, 18–20]. In these systems, digital images are combined with inertial measurements to estimate position, velocity, and attitude. In this paper, a novel, tightly-integrated optical-inertial navigation system is developed, based on the theory of stochastic projections. The system

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the 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 the 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 Washington Headquarters Services, Directorate for Information Operations and Reports, 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 a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 2007		2. REPORT TYPE		3. DATES COVERED 00-00-2007 to 00-00-2007	
4. TITLE AND SUBTITLE Two-Dimensional Stochastic Projections for Tight Integration of Optical and Inertial Sensors for Navigation				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology, 2950 Hobson Way, Wright Patterson AFB, OH, 45433-7765				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES The original document contains color images.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 10	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

is tested with both simulated and experimental data. The performance of the system and potential areas for continued research are discussed. This effort is part of ongoing research into fusion of optical and inertial sensors for long-term autonomous navigation.

Current Methods

It is well-known that optical measurements provide excellent navigation information, when interpreted properly. Optical navigation is not new. Pilotage is the oldest and most natively familiar form of navigation to humans and other animals. For centuries, navigators have utilized mechanical instruments such as astrolabes, sextants, and driftmeters [15] to make precision observations of the sky and ground in order to determine their position, velocity, and attitude.

The difficulty in using optical measurements for autonomous navigation (that is, without human intervention) has always been in the interpretation of the image, a difficulty shared with Automatic Target Recognition (ATR). Indeed, when celestial observations are used, the ATR problem in this structured environment is tractable, and automatic star trackers are widely used for space navigation and ICBM guidance [7, 16, 17]. When ground images are to be used, the difficulties associated with image interpretation are paramount. At the same time, the problems associated with the use of optical measurements for navigation are somewhat easier than ATR. Moreover, recent developments in feature tracking algorithms, miniaturization, and reduction in cost of inertial sensors and optical imagers, aided by the continuing improvement in microprocessor technology, motivates the use of inertial measurements to aid the task of feature tracking in image sequences.

Image-aiding methods are typically classified as either feature-based or optic flow-based, depending on how the image correspondence problem is addressed. Feature-based methods determine correspondence for “landmarks” in the scene over multiple frames, while optic flow-based methods typically determine correspondence for a whole portion of the image between frames. A good reference on image correspondence is [10]. Optic flow methods have been proposed generally for elementary motion detection, focusing on determining relative velocity, angular rates, or obstacle avoidance [5].

Feature tracking-based navigation methods have been proposed both for fixed-mount imaging sensors or gimbal mounted detectors which “stare” at the target of interest, in a manner similar to the gimballed infrared seeker on heat-seeking, air-to-air missiles. Many feature tracking-based navigation methods exploit knowledge (either *a priori*, through binocular stereopsis, or by exploiting terrain homography) of the target location and solve the inverse

trajectory projection problem [1, 13]. If no *a priori* knowledge of the scene is provided, egomotion estimation is completely correlated with estimating the scene. This is referred as the structure from motion (SFM) problem. A theoretical development of the geometry of fixed-target tracking, with no *a priori* knowledge is provided in [14]. An online (Extended Kalman Filter-based) method for calculating a trajectory by tracking features at an unknown location on the Earth’s surface, provided the topography is known is given in [4]. Finally, navigation-grade inertial sensors and terrain images collected on a T-38 “Talon” were processed and the potential benefits of optical-aided inertial sensors are experimentally demonstrated in [18].

Many methods for solving the correspondence problem have been proposed in the computer vision literature. A popular algorithm is the Lucas-Kanade feature tracker [9], which relies on the premise of the invariance of the intensity field between images. It uses a template correlation algorithm to minimize the sum of squared differences (SSD) between image intensities. The algorithm typically assumes a linear ($x - y$ plane) motion model, but can be extended to optimize over affine or bilinear transformations [9, 21]. Other feature correspondence algorithms have been proposed which are invariant to rotations, scaling or both [8]. More robust feature tracking algorithms are typically computationally expensive, and a designer must trade tracking robustness and accuracy for real-time performance [21].

In this paper, the method of stochastic projections [24] is used as the basis for tightly integrating an inertial and optical sensor using an Extended Kalman Filter (EKF) and an automatic target tracking algorithm. In the following section, the integration architecture is presented, which includes the underlying assumptions, the inertial mechanization algorithms, EKF state model, measurement model, and feature tracking concept.

DEVELOPMENT

The method proposed in this paper employs an extended Kalman filter (EKF) algorithm [11, 12] to recursively estimate the navigation state and associated errors by tracking the pixel locations of stationary objects in an image-aided inertial system.

Assumptions

This method is based on the following assumptions.

- A strapdown inertial measurement unit (IMU) is rigidly attached to one or more cameras. Synchronized raw measurements are available from both sensors.
- The camera produces images of objects which are stationary (or very slowly moving) with respect to the world.
- Some form of range to landmark measurement is available, either through binocular stereopsis or using a statistical terrain model.
- The inertial and optical sensors' relative position and orientation is known (see [23] for a discussion of bore-sight procedures).

Optical Sensor Model

An optical sensor is a device designed to measure the intensity of optical energy (light) entering the sensor through an aperture. Imaging sensors consist of an array of light-sensitive detectors which create a multidimensional light intensity measurement (i.e., image). In this section, the basic physical properties of an optical sensor are presented, and a model representing an optical sensor is given.

For the purposes of this discussion, the *world* is defined as a collection of real objects. Some objects are sources of radiometric illumination or *radiance*. These light sources illuminate the world and interact with the other physical objects through various types of reflection. The amount of light along a certain direction is defined as the *irradiance* [10]. The physical irradiance pattern entering the aperture of the optical sensor is defined as the *scene* and is represented by a continuous array of nonnegative real numbers, $\mathbf{o}(x, y, t)$, projected onto the image plane. For the purposes of this discussion, the irradiance sources are constrained to an arbitrary, piecewise continuous, Lambertian surface in three dimensions.

A digital optical imaging sensor consists of an aperture, lens, detector array, and sampling array. A simple imaging system model is shown in Figure 1. The lens focuses the scene on the detector array. The light pattern focused on the detector array is defined as the *image* and represented by, $\mathbf{i}(x, y, t)$. In statistical terms, the *image* is the mean photon arrival rate, and is well-modeled using a Poisson distribution [3]. The detector array converts the light energy into a voltage or a charge which is converted to a digital value by the sampling array. The sampling array is assumed to be a

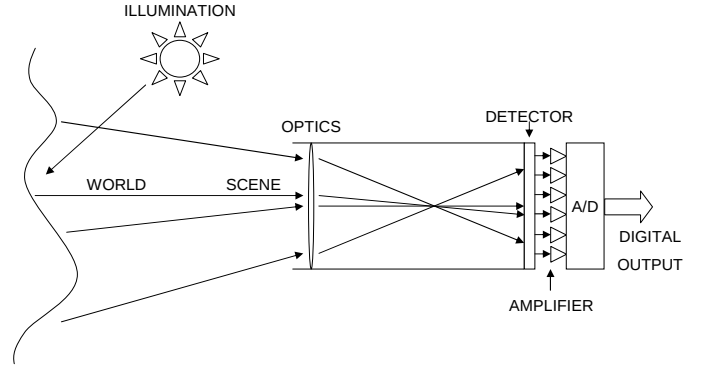


Figure 1: Imaging system model. The imaging system transforms the scene into a digital image. The major components of the camera are the optics, light detector, amplifier, and analog to digital converter.

square grid, although other patterns can be designed (e.g., honeycomb) [6].

The images are transformed into the feature space, which is described in the next section. The pixel location of feature m in the image at time t_i is given by the vector, $\mathbf{z}_m(t_i)$. This pixel location corresponds to a line-of-sight vector which extends outward from the camera focus. This homogeneous line-of-sight vector, $\underline{\mathbf{s}}_m^c(t_i)$ is calculated using the linear projection determined during camera calibration,

$$\underline{\mathbf{s}}_m^c(t_i) = \mathbf{T}_c^{pix} [\mathbf{z}_m(t_i)] \quad (1)$$

where $\mathbf{z}_m(t_i)$ has been corrected for non-linear distortion effects of the imaging system.

Feature Model

To ensure robust, long-term feature tracking, much effort is placed upon extracting features which are invariant to camera motion. This motion is observed as changes in scale, rotation, translation, and affine transformation of the image. It is interesting to note the coupled nature between camera motion and image changes. When estimating the camera motion, it is the “change” in features which must be observed. Lowe’s scale-invariant (SIFT) features are fully invariant to both rotation and scale and partially invariant to affine motion [8]. These invariant properties make this feature transformation attractive for image-aided navigation.

Algorithm Description

The system parameters (see Table 1) consist of the navigation parameters (position, velocity, and attitude), inertial sensor biases, and a vector describing the location of landmarks of interest ($\mathbf{y}$). The navigation parameters are calculated using body-frame velocity increment ($\Delta \mathbf{v}^b$) and angular increment ($\Delta \theta_{ib}^b$) measurements from the inertial

navigation sensor which have been corrected for bias errors using the current filter-computed bias estimates. These measurements are integrated from an initial state in the navigation (local-level) frame using mechanization algorithms described in [22].

Table 1: System Parameter Definition

Parameter	Description
$\mathbf{p}^n$	Vehicle position in navigation frame (northing, easting, and down)
$\mathbf{v}^n$	Vehicle velocity in navigation frame (north, east, down)
$\mathbf{C}_b^n$	Vehicle body to navigation frame DCM
$\mathbf{a}^b$	Accelerometer bias vector
$\mathbf{b}^b$	Gyroscope bias vector
$\mathbf{t}_m^n$	Location of landmark m in the navigation frame (one for each landmark currently tracked)
$\mathbf{d}^b$	Camera-to-IMU lever arm in body frame
$\mathbf{C}_c^b$	Camera-to-IMU orientation DCM

An Extended Kalman Filter was constructed to estimate the errors in the calculated system parameters. In order to minimize the effects of linearization errors, the system parameters were periodically corrected by removing the current error estimate [11]. A block diagram of the system is shown in Figure 2.

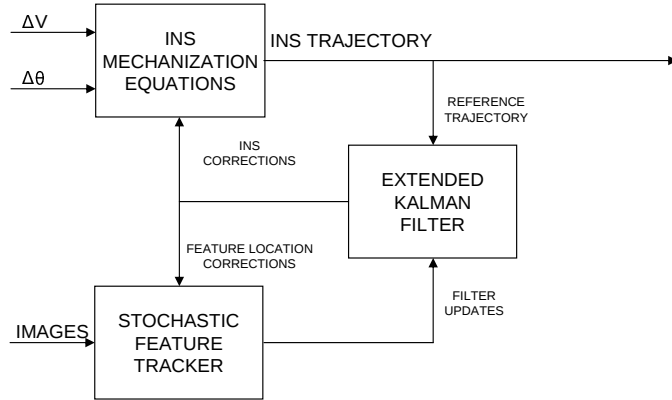


Figure 2: Optical-inertial navigation filter block diagram. In this filter, the location of stationary objects are tracked and used to estimate and update the errors in an inertial navigation system. The inertial navigation system is, in turn, used to support the feature tracking loop.

The Kalman filter state vector, $\hat{\mathbf{x}}$, is defined as

$$\hat{\mathbf{x}} = \begin{bmatrix} \delta \hat{\mathbf{p}}^n \\ \delta \hat{\mathbf{v}}^n \\ \hat{\psi} \\ \delta \hat{\mathbf{a}}^b \\ \delta \hat{\mathbf{b}}^b \\ \delta \hat{\mathbf{y}} \end{bmatrix} \quad (2)$$

where $\delta \hat{\mathbf{p}}^n$ is the estimated position error vector, $\delta \hat{\mathbf{v}}^n$ is the estimated velocity error vector, and $\hat{\psi}$ is the estimated body-to-navigation frame attitude error (defined as small-angle rotations about the north, east, and down axes). The accelerometer and gyroscope bias errors are represented by $\delta \hat{\mathbf{a}}^b$ and $\delta \hat{\mathbf{b}}^b$, respectively. The landmark position error vector ($\delta \mathbf{y}$) is defined as a collection of the errors ($\delta \mathbf{t}_m^n$) in the currently tracked landmark positions. The perturbation error states are defined as

$$\delta \mathbf{p}^n = \tilde{\mathbf{p}}^n - \mathbf{p}^n \quad (3)$$

$$\delta \mathbf{v}^n = \tilde{\mathbf{v}}^n - \mathbf{v}^n \quad (4)$$

$$\tilde{\mathbf{C}}_b^n = [\mathbf{I} - \psi \times] \mathbf{C}_b^n \quad (5)$$

$$\delta \mathbf{a}^b = \tilde{\mathbf{a}}^b - \mathbf{a}^b \quad (6)$$

$$\delta \mathbf{b}^b = \tilde{\mathbf{b}}^b - \mathbf{b}^b \quad (7)$$

$$\delta \mathbf{y} = \tilde{\mathbf{y}} - \mathbf{y} \quad (8)$$

where the tilde ($\tilde{\cdot}$) represents the nominal (estimated) trajectory.

The position, velocity, and attitude errors were modeled as a stochastic process based on the well-known Pinson navigation error model [22]. The accelerometer and gyroscopic bias errors were each modeled as a first-order Gauss-Markov process [11], based on the specification for the inertial measurement unit (IMU). The landmarks are modeled as stationary with respect to the Earth. A small amount of process noise is added to the state dynamics to promote filter stability [12].

Landmark Track Maintenance

In a practical system, the number of Kalman Filter states is limited by available computer resources. As a result, the number of landmarks actively tracked must be constrained. This inherent limitation motivates the implementation of a track maintenance algorithm.

The general concept for the track maintenance algorithm is to add and prune landmark tracks in order to provide the “best” navigation information to the filter. Although the optimal landmark choices are highly dependent on the trajectory and scene, some general “rules-of-thumb” were used in the track maintenance algorithm.

In general, features which can be easily and accurately tracked for long periods of time provide the best navigation information. This implies choosing features which

are: strong and easily identified (to help maintain track), locally distinct (to eliminate false correspondence), and well-separated in image space (to maximize filter observability). Thus, when Kalman Filter landmark track states are available, the feature space of the current image is searched and new landmarks are added based on the above criteria. The filter states are augmented in accordance with the stochastic projection algorithm defined in [24].

In order to maintain only the best tracks, stale landmark tracks (i.e., no successful correspondence available for a given period of time) are pruned by removing the associated filter states. Other track maintenance approaches are possible which could theoretically improve the track performance (e.g., semi-causal, multiple model, or velocity prediction), however these approaches will not be pursued in this paper.

Stale Landmark Revisitation

As proposed by Strelow in [21], the navigation performance can be improved considerably when the system can revisit landmarks which were previously tracked. This approach could theoretically constrain the growth of navigation errors indefinitely; however, it would not be of benefit for long-distance navigation. This approach will not be considered in this paper.

Measurement Model In order to exploit the synergistic properties of optical and inertial sensors, the navigation and feature tracking algorithms are tightly-coupled. This results in a slight modification to the standard Kalman filter update and propagation cycles in order to incorporate the feature tracking loop. The tracking loop is responsible for: incorporating new landmark tracks, using stochastic projections to predict and match features between images, providing filter measurements, and deleting stale landmarks from the filter.

The Kalman filter assists the tracking algorithm by maintaining and propagating the minimum mean-square error state estimate. This provides the stochastic projections which help improve the speed and robustness of the tracking loop.

The tracking loop incorporates new landmark tracks when necessary by determining an initial estimate of the landmark location (using either a terrain model or binocular stereopsis combined with the current navigation state vector). This estimate, along with the calculated covariance and cross-correlation matrices are augmented into the Kalman filter state vector and covariance matrix. Details on the mathematics involved in the calculation of the above process are based on the stochastic projection model described in [24].

After the landmark tracks are properly augmented into the Kalman filter, the standard propagation algorithms are used to predict the augmented state to the time of the next image. The location of each landmark (along with arbitrary uncertainty ellipsoid) can then be projected into the feature space of the new image. In this paper, the feature space corresponds to a two-dimensional pixel location and associated uncertainty ellipse. The tracking algorithm then searches this uncertainty ellipse for a feature which has similar characteristics to the reference feature and is distinct. In this paper, a $2 - \sigma$ ellipse was used. An example of feature prediction is shown in Figure 3. The reader is referred to [24] for more details regarding the feature prediction algorithm.

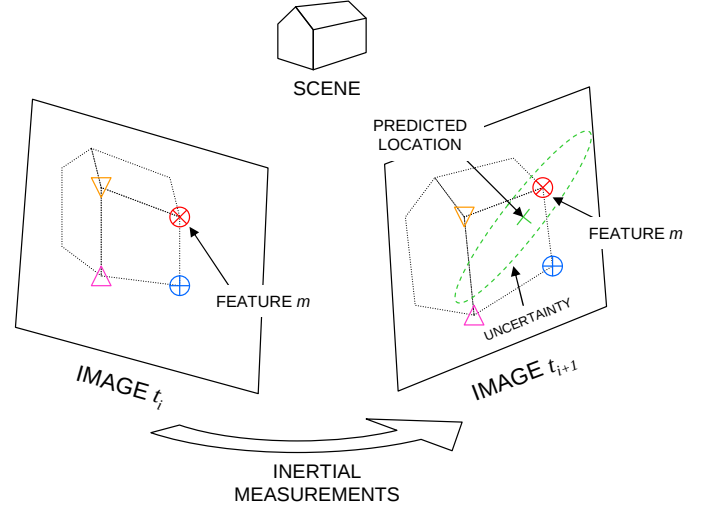


Figure 3: Stochastic feature projection. Optical features of interest are projected into future images using inertial measurements and stochastic projections.

Once the landmark tracker has determined a corresponding match, the pixel location of the feature, corrected for optical distortion, is used to update the navigation state. The measured pixel location for feature m is described as a function of the system parameters by

$$\mathbf{z}^b(t_i) = \mathbf{h} [\mathbf{p}^n, \mathbf{C}_b^n, \mathbf{t}_m^n, \mathbf{C}_c^b, \mathbf{d}^b] + \mathbf{v}(t_i) \quad (9)$$

or, more specifically,

$$\mathbf{z}^b(t_i) = \mathbf{T}_c^{pix} \underline{\mathbf{s}}_m^c(t_i) + \mathbf{v}(t_i) \quad (10)$$

where $\mathbf{T}_c^{pix}$ is the linear projection matrix (see [10]) and $\underline{\mathbf{s}}_m^c(t_i)$ is the homogeneous (i.e., unit normalized by s_z component) line-of-sight vector to landmark m , expressed in the camera frame. The measurement is corrupted by $\mathbf{v}$, a zero-mean, white, Gaussian noise process with covariance

$$E [\mathbf{v}(t_i) \mathbf{v}^t(t_j)] = \begin{cases} \mathbf{R} & t_i = t_j \\ \mathbf{0} & t_i \neq t_j \end{cases} \quad (11)$$

All measurements are modeled as independent. The line-of-sight vector is defined as

$$\mathbf{s}_m^c = \mathbf{C}_b^c \mathbf{C}_n^b [\mathbf{t}_m^n - \mathbf{p}^n] - \mathbf{C}_b^c \mathbf{d}^b \quad (12)$$

where $\mathbf{C}_b^c$ and $\mathbf{d}^b$ are the camera boresight parameters.

The linearized observation matrix, $\mathbf{H}$, is the Jacobian of the nonlinear measurement function, $\mathbf{h}[\cdot]$, linearized about the reference trajectory, $\hat{\mathbf{x}}$:

$$\mathbf{H} = \left. \frac{\partial \mathbf{h}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}} \quad (13)$$

The partial derivative with respect to position is expressed as

$$\frac{\partial \mathbf{h}}{\partial \mathbf{p}^n} = \mathbf{T}_c^{pix} \frac{\partial \mathbf{s}_m^c}{\partial \mathbf{p}^n} \quad (14)$$

where,

$$\frac{\partial \mathbf{s}_m^c}{\partial \mathbf{p}^n} = \frac{\frac{\partial \mathbf{s}_m^c}{\partial \mathbf{p}^n} - \mathbf{s}_m^c \frac{\partial \mathbf{s}_m^c}{\partial \mathbf{p}^n}}{s_{m_z}^c} \quad (15)$$

and,

$$\frac{\partial \mathbf{s}_m^c}{\partial \mathbf{p}^n} = -\mathbf{C}_b^c \mathbf{C}_n^b \quad (16)$$

The partial derivative with respect to body-to-navigation frame misalignment angle vector, ψ , is

$$\frac{\partial \mathbf{h}}{\partial \psi} = \mathbf{T}_c^{pix} \frac{\partial \mathbf{s}_m^c}{\partial \psi} \quad (17)$$

where

$$\frac{\partial \mathbf{s}_m^c}{\partial \psi} = \frac{\frac{\partial \mathbf{s}_m^c}{\partial \psi} - \mathbf{s}_m^c \frac{\partial \mathbf{s}_m^c}{\partial \psi}}{s_{m_z}^c} \quad (18)$$

and

$$\frac{\partial \mathbf{s}_m^c}{\partial \psi} = -\mathbf{C}_b^c \mathbf{C}_n^b [(\mathbf{t}^n - \mathbf{p}^n) \times] \quad (19)$$

The skew symmetric operator, $(\cdot) \times$, is defined for a generic 3×3 vector, $\mathbf{v}$ as

$$\mathbf{v} \times = \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} \times = \begin{bmatrix} 0 & -v_z & v_y \\ v_z & 0 & -v_x \\ -v_y & v_x & 0 \end{bmatrix} \quad (20)$$

Finally, the partial derivative with respect to the landmark position vector, $\mathbf{t}_m^n$, is

$$\frac{\partial \mathbf{h}}{\partial \mathbf{t}_m^n} = \mathbf{T}_c^{pix} \frac{\partial \mathbf{s}_m^c}{\partial \mathbf{t}_m^n} \quad (21)$$

where

$$\frac{\partial \mathbf{s}_m^c}{\partial \mathbf{t}_m^n} = \frac{\frac{\partial \mathbf{s}_m^c}{\partial \mathbf{t}_m^n} - \mathbf{s}_m^c \frac{\partial \mathbf{s}_m^c}{\partial \mathbf{t}_m^n}}{s_{m_z}^c} \quad (22)$$

and

$$\frac{\partial \mathbf{s}_m^c}{\partial \mathbf{t}_m^n} = \mathbf{C}_b^c \mathbf{C}_n^b \quad (23)$$



Figure 4: Data collection system. The data collection system consisted of a tactical-grade IMU and monochrome digital cameras.

All other partial derivatives are zero.

The resulting observation matrix is

$$\mathbf{H} = \begin{bmatrix} \frac{\partial \mathbf{h}}{\partial \mathbf{p}^n} & \mathbf{0}_{3 \times 3} & \frac{\partial \mathbf{h}}{\partial \psi} & \mathbf{0}_{3 \times 6} & \cdots & \frac{\partial \mathbf{h}}{\partial \mathbf{t}_m^n} & \cdots \end{bmatrix} \quad (24)$$

calculated for all successful feature correspondence matches at the current time.

SYSTEM TESTS

The data collection system consisted of an IMU and two digital cameras (see Figure 4). The IMU was a Honeywell HG-1700 tactical-grade unit which measured acceleration and angular rate at 100 Hz. The digital cameras were both Pixelink A-741 machine vision cameras which incorporated a global shutter feature and a Firewire interface. The lenses were wide-angle Pentax lenses with approximately 90 degrees field of view. The sensors were mounted on an aluminum plate and calibrated using procedures similar to those described in [23]. Images were captured at approximately 1 Hz.

The algorithm was tested experimentally using two navigation profiles designed to examine the sensitivity and robustness of the feature tracking system. The first profile consisted of a closed path over an outdoor parking area. The path was traversed forward and backwards with the camera pointed toward the outside of the path. This trajectory resulted in seven segments that presented a scene change and forced the filter to search for new features. This outdoor scene consisted of a combination of man-made features (buildings, fences, roads, etc.) and natural features such

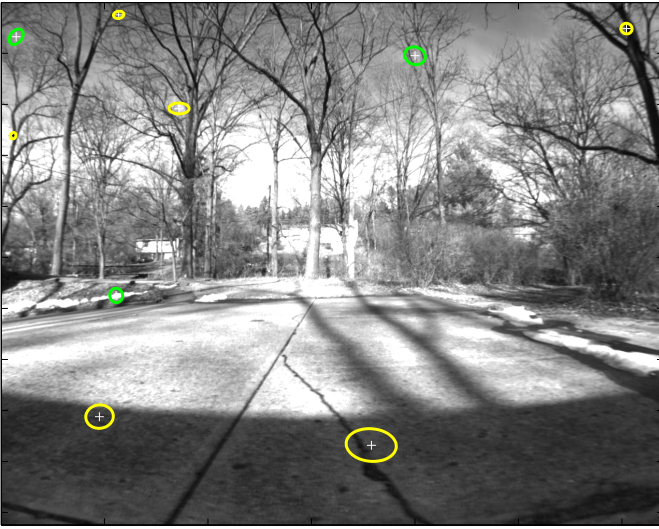


Figure 5: Sample image from outdoor data collection. The outdoor data collection presented the filter with a combination of man-made and natural features. The crosses and ellipses indicate the locations and uncertainty of currently tracked landmarks.

as grass and trees. The profile consisted of a 10-minute stationary alignment period, followed by four minutes of navigation using only images and inertial measurements. No prior knowledge was used with any feature. The filter was limited to a maximum of ten features at any time. A sample image from the outdoor profile is shown in Figure 5.

This profile presented the algorithm with a challenging feature tracking environment due to the high-contrast lighting conditions, large variation in feature distance (zero to infinite), and complicated images with semi-transparent objects overlapping at different ranges (e.g., multiple layers of tree limbs).

The filter successfully utilized inertial measurements to predict and constrain the image correspondence search during the entire profile. In return, the feature correspondence updated and corrected the inertial measurement errors and significantly reduced the resulting drift in the navigation solution. Over the four-minute non-stationary profile, the navigation errors were estimated to be less than 1 m in the horizontal plane and less than 3 m in the vertical. Typical free-inertial performance for this inertial sensor is estimated to be on the order of hundreds of meters horizontal. The unstable nature of the vertical channel would require external aiding in order to maintain stability. These initial results clearly show the benefits of the described method.

The second profile consisted of a closed path in an indoor environment. The path began and ended at the same location and orientation in the Advanced Navigation Technology Center laboratory. As in the previous profile, the data



Figure 6: Sample image from indoor data collection. The indoor data collection presented the filter with man-made features in an office environment. The crosses and ellipses indicate the locations and uncertainty of currently tracked landmarks.

collection began with a 10-minute stationary alignment period. After the alignment period, the sensor was moved in a 10-minute loop around the hallways of the building. In contrast to the previous profile, the sensor was pointed primarily in the direction of travel. No prior knowledge was provided to the algorithm regarding the location of features or structure of the environment. A sample image from the indoor profile is shown in Figure 6.

The indoor profile presented the algorithm with different challenges from a feature tracking perspective. The indoor environment consisted of repetitive, visually identical features (e.g., floor tiles, lights, etc.), which could easily cause confusion for the feature tracking algorithm. In addition, reflections from windows and other shiny surfaces would not be interpreted properly by the filter and could potentially result in navigation errors. Finally, the lower average light intensity levels and large areas with poor contrast (e.g., smooth, featureless walls) presented a relatively stark feature space.

As with the previous profile, the filter performed well. The filter's estimate of the trajectory was overlaid on a floor plan of the building in Figure 7 along with the inertial-only trajectory calculated using the Novatel Black Diamond System and an image-aided inertial trajectory with stochastic constraints disabled. The estimated trajectory corresponds well to the building's hallways. The inertial-only trajectory quickly develops large errors, even though the Novatel filter is applying an altitude hold correction and using a vehicle motion model in an attempt to constrain the inertial drift. With stochastic constraints pur-

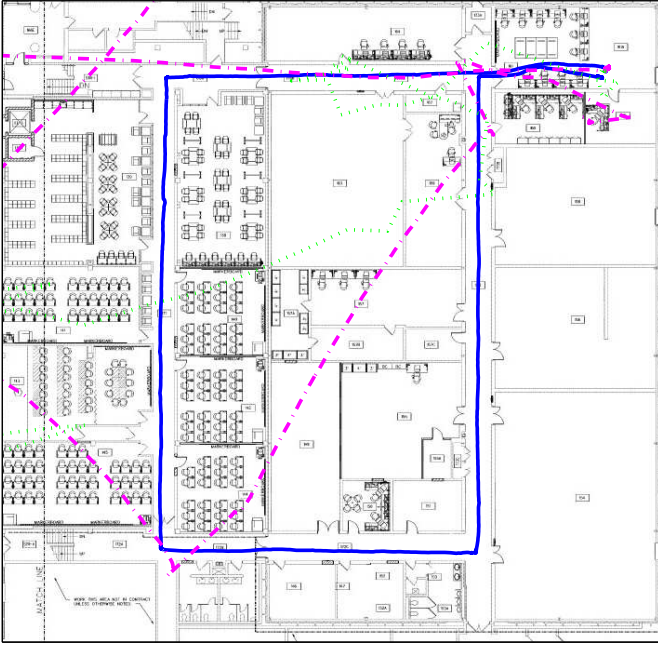


Figure 7: Estimated path from indoor data collection. The filter's estimate of the path (indicated by the solid line) agrees well with the known floor plan. The inertial-only best estimate of trajectory (indicated by the dashed line) and optical-inertial with stochastic constraints disabled (indicated by the dotted line) show large errors in position and heading.

posely disabled, the trajectory estimate quickly diverges due to false correspondence matches. This illustrates the catastrophic effects of incorporating false updates into an Extended Kalman Filter with inertial feedback and highlights the inherent strength of applying robust correspondence methods, and in particular stochastic constraints.

The filter's estimated trajectory and estimated location of features used for tracking are shown in Figure 8. More detail of the start/stop area is shown in Figure 9. The difference in the estimated start and stop locations shows the accumulated errors in the filter over the path. Over the 10-minute profile, the navigation errors were less than 1 m in the horizontal plane and less than 3 m in the vertical. Again, this was a significant improvement over the Novatel free-inertial performance. The results are particularly impressive as the solution was calculated using an online algorithm, with only raw inertial and image data. No *a-priori* knowledge of the environment was provided to the system.

CONCLUSIONS

In this paper, an algorithm is presented which integrates inertial and optical measurements to provide an enhanced navigation solution. Stochastic algorithms are applied

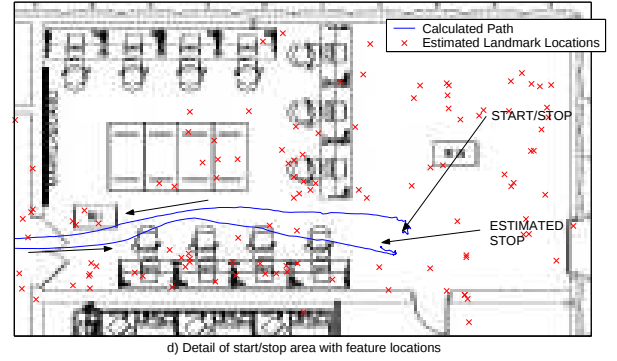
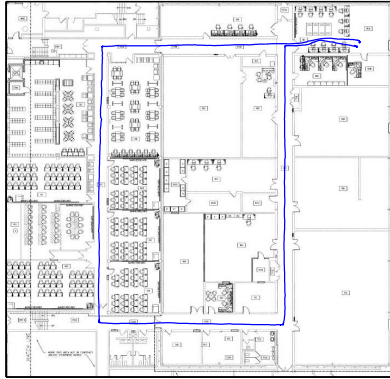


Figure 9: Enhanced detail of the start/stop area illustrating the estimated trajectory and feature locations. The difference between the estimated start and stop location illustrates the accumulated position error.

which utilize the natural strengths of each sensor in a synergistic manner. Results from two data collection scenarios show the effectiveness of the stochastic projection method for tightly-integrating optical and inertial sensors. The filter successfully utilized inertial measurements to enhance feature tracking while simultaneously correcting for inertial errors. These tests demonstrate the viability of our approach for navigation without external signals.

DISCLAIMER

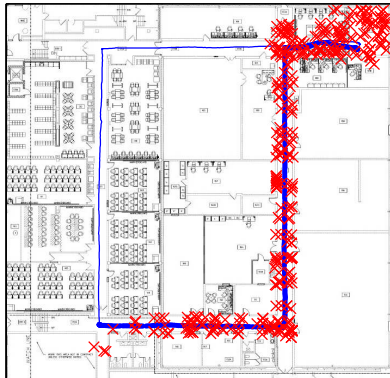
The views expressed in this article are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S Government.



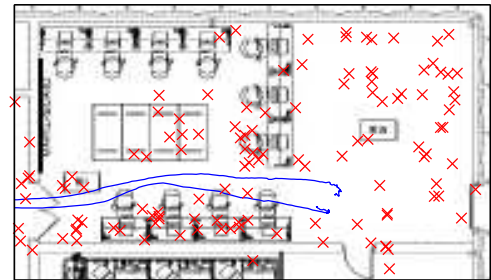
a) Entire path estimate



b) First half of estimated path with tracked feature locations



c) Second half of estimated path with tracked feature locations



d) Detail of start/stop area with tracked feature locations

Figure 8: Estimated path and feature locations from indoor data collection. Pane (a) shows the entire path estimate. Pane (b) shows the first half of the path along with the estimated location of the features (indicated by x symbols). Pane (c) shows the last half of the path and estimated feature locations. Pane (d) provides detail of the start/stop area.

References

- [1] H. Adams, S. Singh, and D. Strelow. An Empirical Comparison of Methods for Image-based Motion Estimation. In *Proceedings of the 2002 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems*, volume 1, pp. 123–128, September 2002.
- [2] R. G. Brown and P. Y. Hwang. *Introduction to Random Signals and Applied Kalman Filtering*. John Wiley and Sons, Inc., New York, NY, 1992.
- [3] J. W. Goodman. *Introduction to Fourier Optics*. McGraw Hill, Boston, Massachusetts, 1996.
- [4] E. Hagen and E. Heyerdahl. Navigation by Optical Flow. In *Proceeding of the 11th IAPR International Conference on Pattern Recognition*, volume 1, pp. 700–703, 1992.
- [5] S. Hrabar and G. S. Sukhatme. A Comparison of Two Camera Configurations for Optic-Flow Based Navigation of a UAV Through Urban Canyons. In *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 3, pp. 2673–2680, September 2004.
- [6] A. K. Jain. *Fundamentals of Digital Image Processing*. Prentice Hall, Inc., Upper Saddle River, New Jersey 07458, 1989.
- [7] C. Liebe. Star trackers for attitude determination. *Aerospace and Electronic Systems Magazine, IEEE*, 10(6):10–16, June 1995.
- [8] D. G. Lowe. Object Recognition from Local Scale-Invariant Features. In *Proc. of the International Conference on Computer Vision*, volume 2, pp. 1150–1157, September 1999. Corfu, Greece.
- [9] B. D. Lucas and T. Kanade. An Iterative Image Registration Technique with an Application to Stereo Vision. *Proc. DARPA Image Understanding Workshop*, pp. 121–130, 1981.
- [10] Y. Ma, S. Soatto, J. Kosecka, and S. S. Sastry. *An Invitation to 3-D Vision*. Springer-Verlag, Inc., New York, New York, 2004.
- [11] P. S. Maybeck. *Stochastic Models Estimation and Control, Vol I*. Academic Press, Inc., Orlando, Florida 32887, 1979.
- [12] P. S. Maybeck. *Stochastic Models Estimation and Control, Vol II*. Academic Press, Inc., Orlando, Florida 32887, 1979.
- [13] C. F. Olson, L. H. Matthies, M. Schoppers, and M. W. Maimone. Robust Stereo Ego-motion for Long Distance Navigation. In *Proceedings of the IEEE Conference on Advanced Robotics*, volume 2, pp. 453–458, June 2000.
- [14] M. Pachter and A. Porter. Bearings-only Measurements for INS Aiding: The Three-Dimensional Case. In *Proceedings of the 2003 AIAA Guidance, Navigation and Control Conference*, 2003. AIAA paper number 2003-5354.
- [15] M. Pachter and A. Porter. INS Aiding by Tracking an Unknown Ground Object - Theory. In *Proceedings of the American Control Conference*, volume 2, pp. 1260–1265, 2003.
- [16] P. Pissavin, J. Krebs, P. LeLong, P. Vidal, and R. Navoni. Improved Star Tracker for ODIN Satellite. In *Proceedings of ESA International Conference on Spacecraft Guidance, Navigation and Control Systems*, pp. 611–616, November 1997.
- [17] D. Purll, N. Gradmann, and M. Bollner. The ROSAT Star Tracker: Flight Experience. *ESA, Spacecraft Guidance, Navigation and Control Systems*, N92-2443215-18:551–556, 1991.
- [18] J. F. Raquet and M. Giebner. Navigation Using Optical Measurements of Objects at Unknown Locations. In *Proceedings of the 59th Annual Meeting of the Institute of Navigation*, pp. 282–290, June 2003.
- [19] D. Strelow and S. Singh. Optimal Motion Estimation from Visual and Inertial Measurements. In *Proceedings of the Workshop on Applications of Computer Vision*, December 2002.
- [20] D. Strelow and S. Singh. Online Motion Estimation from Visual and Inertial Measurements. In *Proceedings of the Workshop on Integration of Vision and Inertial Sensors (INERVIS 2003)*, June 2003.
- [21] D. W. Strelow. *Motion Estimation from Image and Inertial Measurements*. PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, November 2004.
- [22] D. Titterton and J. Weston. *Strapdown Inertial Navigation Technology*. Peter Peregrinus Ltd., Lavenham, United Kingdom, 1997.
- [23] M. J. Veth and J. F. Raquet. Alignment and Calibration of Optical and Inertial Sensors Using Stellar Observations. In *Proceedings of ION GNSS 2005*, pp. 2494–2503, September 2005.
- [24] M. J. Veth, J. F. Raquet, and M. Pachter. Stochastic Constraints for Robust Image Correspondence Search. *IEEE Transactions on Aerospace Electronic Systems*, Publication forthcoming.