INERTIAL NAVIGATION SENSOR INTEGRATED MOTION ANALYSIS

Bir Bhanu, Barry Roberts, and John Ming
Honeywell Systems and Research Center
3660 Technology Drive
Minneapolis, MN 55418

ABSTRACT

Land navigation requires a vehicle to steer clear of trees, rocks, and man-made obstacles in the vehicle's path while vehicles in flight, such as rotorcraft, must avoid antennas, towers, poles, fences, tree branches, and wires strung across the flight path. Automatic detection of these obstacles and the necessary guidance and control actions triggered by such detection would facilitate autonomous vehicle navigation. An approach employing a passive sensor for mobility and navigation is generally preferred in practical applications of these robotic vehicles. Motion analysis of imagery obtained during vehicle travel can be used to generate range measurements to world points within the field of view of the sensor, which can then be used to provide obstacle detection. But, state-of-the-art in motion analysis is not robust and reliable enough to handle arbitrary image motion caused by vehicle movement. However, many types of existing vehicles contain inertial navigation systems (INS) which can be utilized to greatly improve the performance of motion analysis techniques and make them useful for practical military and civilian applications. In particular, INS measurements can improve interest point selection, matching of the interest points, and the subsequent motion detection, tracking, and obstacle detection. We review various techniques of ranging (both passive and active) and discuss an inertial sensor integrated optical flow technique for motion analysis to achieve increased effectiveness in obstacle detection during vehicle motion. Our approach to motion analysis for obstacle detection is illustrated by simulated results and the results obtained using land vehicle data.

1. INTRODUCTION

A desired obstacle detection system for many practical applications should exhibit robustness and should not place unduly excessive size, power, or weight demands on the host vehicle. It should work in day/night/adverse weather conditions and should preferentially be covert to minimize the threat to the vehicle and the pilot. The technique used for obstacle detection must also have graceful degradation, instead of total failure, under conditions of limited operability. In recent years, considerable effort has been put toward the detection of obstacles that present themselves primarily to ground vehicles. Using mainly active sensors, such as a laser scanner, obstacles (like fence posts, rocks, vegetation) are detected within the field of view of the vehicle's sensor. Other active sensors such as Millimeter Wave (MMW) can detect obstacles such as wires, but the constant and continuous image of these active sensors betrays vehicle's covertness.

Passive sensors, such as a TV camera, are also being used to detect obstacles for ground vehicles. However, state-of-the-art motion analysis techniques for obstacle detection are not robust and reliable enough for many practical applications. Many of these techniques require that unrealistic constraints be placed on the input data in order to make them work. The largest sources of error are sensor motion and incomplete/ambiguous information in the sensed image data. However, many types of land and air vehicles (e.g. helicopters and military ground vehicles) contain an Inertial Navigation System (INS) whose output can be used for applications beyond the original intent of the system. Such vehicles can use the INS information to greatly simplify some of the functionalities normally provided by computer vision, such as obstacle detection, target motion detection, target tracking, stereo, etc. Figure 1 shows several inertial sensor assembly/packages currently available. In this paper, we make use of INS measurements to enhance the quality and robustness of motion analysis techniques for obstacle detection and thereby provide vehicles with new functionality and capability.

The objective of the work presented in this paper is to describe our maximally passive approach to obstacle detection and to discuss the details of our inertial sensor integrated optical flow analysis technique. In Section 2, we review the principal techniques of passive and active ranging. Section 3 discusses the technical problems associated with motion analysis for passive ranging. We present our new approach to motion analysis in Section 4 and describe the details of the optical flow technique. Section 5 describes the results we have obtained with our optical flow approach. Finally, Section 6 provides the conclusions of the paper. A brief discussion on navigation errors for some ring laser gyro's is provided in Appendix A.
2. OVERVIEW OF PASSIVE AND ACTIVE RANGING TECHNIQUES

2.1 PASSIVE RANGING

There are a large number of obstacle detection techniques proposed in the literature that make use of motion analysis, stereo methods, or other techniques for passive ranging.

Motion Analysis Techniques - These methods can be further broken down into optical flow approaches or structure and motion methods.

Optical Flow Techniques - These techniques utilize information provided by the velocity field that represents the apparent motion of stationary object points through a temporal sequence of images. Most methods for estimation of optical flow can be categorized into two classes, gradient-based methods and displacement-based methods.

Gradient-Based Methods - These methods use relationships involving optical flow and derivatives of the image brightness function coupled with a variety of constraints on the flow field. Although considerable work has been done in this area, significant results have yet to be demonstrated on images of outdoor scenes. One fundamental reason for this is that these methods are highly sensitive to noise. This is a highly undesirable property of any method used on images of outdoor scenes. Furthermore, theoretical analysis has shown that there is a direct conflict between various constraints imposed on the flow field. In particular, it is shown that errors due to instability of solutions of the required systems of equations are inversely related to the size of the neighborhood used for flow smoothness constraints but that increasing the size of the neighborhood increases the error due to violations of flow smoothness.

Displacement-Based Methods - In these methods image features (points, edges, regions, or boundaries) are matched between a temporal sequence of two or more images to derive initial estimates of flow vectors. These methods make use of the flow pattern which is experienced by a moving observer. The motion between frames can be decomposed into translational and rotational components. The final (usually the second) image in the sequence can then be derotated to achieve a relationship approximating pure translation between the first and second image. In the case of pure sensor translation in a stationary environment, every point seems to expand from one particular image location termed the Focus of Expansion (FOE). If the location of the FOE is
known, then the relative depth of matched stationary image points can be found. If, in addition, the velocity of the sensor and the elapsed time between frames are known, then the absolute range can be computed using trigonometric formulas. Range to other image points can be estimated using interpolation procedures. Accurate range estimates utilizing the above approach require long displacement vectors. Issues raised here include accurate frame-to-frame correspondence, accurate FOE location, and the magnitude of interpolation ambiguities. Inertial navigation sensor (INS) integrated methods presented later in this paper alleviate these problems.

**Structure and Motion Methods** -- In these methods, both 3-D structure and motion are computed in one integral step by solving a system of linear or nonlinear equations. These methods, although elegant, reportedly are sensitive to noise, require large amounts of computation, converge slowly, and require many disparate views of the object.

**Stereo Techniques** - These methods are widely studied for determining range passively. In order to use stereo, feature matches must be made between the two images. The accuracy of these matches depends upon the knowledge of relative sensor positions, the displacement between the sensors (the baseline), and the availability of prominent features to match. Once feature correspondences in the two images have been established, the range to the corresponding world points can be computed using trigonometry. Range to other points can then be estimated by using an interpolation procedure. One characteristic property of stereo methods is the fact that, to a first approximation, the error in stereo depth measurements is directly proportional to the positional error of the matches and inversely proportional to the length of the baseline. Thus, as in the case of gradient based techniques, there are sources of error here which are in direct conflict with one another since the longer the baseline is, the harder it is to obtain accurate matches. Potential solutions to this problem include development of highly accurate feature matching techniques and statistical averaging over several views. The statistical averaging method uses the concept of combining motion and stereo, which has the advantage of providing complementary and cooperative information to a passive ranging system.

**Other Techniques** - Although the aforementioned techniques comprise the majority of methods used in passive ranging, various other approaches have also been suggested. A spatio-temporal extension of the Marr-Hildreth edge operator is one method which has been suggested by Buxton and Buxton. The operator is used to locate edges in time varying imagery. This technique can be considered as a type of hybrid between the gradient and displacement methods. This technique suffers from certain disadvantages such as the Aperture Effect as compared to the wavefront region growing techniques developed by Bhanu and Burger. Hollister has developed a technique for passive ranging to point sources using the bearing angles between the sensor line of sight and the point source. This technique is based on the assumption that all motion is in a plane and no results on real data are given. Bowman and Gross have also developed a method for passive ranging to targets using data from two different aircrafts, but it is not applicable to the rotorcraft low-altitude flight scenario. Techniques based on Kalman filtering are also being developed for general motion and passive ranging.

### 2.2 ACTIVE RANGING

A variety of laser and Millimeter Wave (MMW) radar systems exist that can scan relatively large fields-of-view and detect and determine the range to power lines, cables, and terrain obstacles. Currently, a number of 3-D laser scanners using phase detection technology are available. One such sensor, developed for autonomous vehicle navigation, has a field-of-view of ±40° horizontally and covers depression angles from 15° to 45° with a range resolution of 8 centimeters. More advanced systems with multiple lasers operating at multiple frequencies in the visible, near infrared, and shortwave infrared wavelengths are also under development. The multiple wavelengths allow for range reflectance determination with a 60° × 80° field-of-view and a range resolution of 2 centimeters. Another commercially available 3-D laser ranging system with a 60° × 60° field-of-view has similar range resolution. It allows a fixed pattern radar scan with 128 × 128 pixel resolution and a frame rate of 0.8 seconds. Both of the above mentioned systems have phase ambiguities on the order of 40 feet in their range measurements.

Thomson CSF is developing a compact MMW radar system (Romeo 2) which uses a 3-second scan over a 90° sector to detect hazardous objects. Prototype systems have detected 3 millimeter diameter high tension cables at ranges of 1000 meters in foggy weather. The system is designed to detect similar objects with small cross sections.

In general, lasers and MMW radar systems are able to detect and accurately determine the range of terrain obstacles. Studies have demonstrated that both kinds of systems can successfully detect transmission cables at all angles, polarizations, and surface conditions, although transmission line detection at all aspect angles with a MMW sensor requires scanning. This study was conducted using four MMW frequencies (18, 34, 56, and 94 GHz), two laser wavelengths (10.6, 1.06 μm), three polarizations (horizontal, vertical, and cross), various surface conditions (dry, wet, rough, and smooth), five kinds of cables, and several aspect angles. Recent advances in CO₂ laser technology have led to the development of fieldable LADAR system that can provide high resolution imagery suitable for automatic target
recognition and obstacle detection such as wires, poles, etc.

Unfortunately, active systems are subject to threats through automatic detection by the enemy. They undoubtedly provide good obstacle avoidance capability at the price of increased danger to the crew and the vehicle, regardless of whether the active system is based on laser or MMW radar ranging. Consequently, their use should be contingent on the capabilities of passive obstacle detection/avoidance technology in near and far future systems. The report by Bhunu and Roberts\(^9\) presents detailed tradeoffs of passive/active ranging approaches for obstacle detection.

3. TECHNICAL PROBLEMS WITH MOTION ANALYSIS TECHNIQUES

In this section, we present a critical assessment of some of the problems in the area of motion analysis that demand innovative solution concepts for success. The difficulties involved come from many sources. The general problem areas which we consider important are (not necessarily in order of importance):

Inherent Quantization Error and Noise -- As mentioned earlier, techniques based on discrete differentiation (optical flow with global constraints\(^20\) ) are generally considered to be so sensitive to noise that they are unreliable in outdoor scenes. In addition, the identification of the same world point (interest point) within multiple frames becomes unreliable due to noise and quantization.

Correspondence or Feature Matching Problem -- This problem has been studied quite extensively\(^5,10\). Solution of this problem is necessary in a purely passive, image-based displacement method as well as for stereo techniques. The great deal of effort which has been expended toward solving this problem and the lack of a technique that will insure a very high degree of matching accuracy implies that passive ranging systems which utilize feature matching must be capable of tolerating a certain number of inaccurate (possibly highly inaccurate) matches. One well known approach to increasing the accuracy of matches obtained by the correspondence problem is to use relaxation techniques to maintain certain consistencies in correspondences between neighboring features. Another approach is to use input from a Ring Laser Gyro (RLG)/accelerometer apparatus to determine the motion of the sensor between frames in the monocular case, or the relative motions of the sensors in the stereo case, in order to register the images and locate the FOE without computation. This apparatus can provide all the sensor attitude and velocity information needed to completely describe sensor motion. This information also introduces a number of constraints on the search area required for matching. This approach will be explained in more detail in Section 4.

Determining the Location of the FOE -- This is a major problem which must be addressed if one uses optical flow methods and monocular sequences for passive ranging. Location of the FOE can be approximated by purely image-based methods\(^8\) or by using input from a RLG/accelerometer assembly. When using purely image-based methods, bad correspondences, quantization effects (including roundoff error), and noisy data all contribute to the inherent instability in attempting to solve a set of equations exactly for the FOE. The method of choice is to use the inertial information provided by an RLG/accelerometer to calculate the FOE within each image frame, thereby avoiding the uncertainties involved with passive methods.

Interpolation of the Range Map -- The desire to have range estimates to all points in the field-of-view is another problem which should be addressed in order to develop accurate passive ranging systems. The previous techniques for passive ranging will compute range only as subset of the available points in the field-of-view. In order to solve the obstacle detection/avoidance problem via ranging, a much more dense set of points must have range values available. An interpolation procedure can be used to estimate range over a dense set (possibly all) of pixels that cover the field-of-view. There are some difficulties associated with this process. A standard practice has been to fit a smooth surface, such as a polynomial of two variables or a spline, through the computed range points. Some problems with polynomials are that they are continuous whereas range maps, in general, are discontinuous. Polynomials also have a predetermined shape. To obtain a high degree of variability, one must use fairly high order polynomials, but then the least squares process becomes very computationally intensive. Splines offer a better choice for fitting a smooth function through the points, since, although they are continuous, a high degree of variation can be attained with reliable, less expensive computational methods. Unfortunately, splines are not guaranteed to pass through the given range points. Scene analysis techniques can simplify this process.\(^9\)

Use of Artificial Intelligence and Qualitative Methods -- A final general issue to consider in motion analysis is not a direct problem. The issue is how much, if any, do passive ranging via motion analysis methods need to be augmented with intelligent, or qualitative, techniques. Range is a very concrete concept, and it is easy to understand how and why to use range once it is available. It is true, however, that most humans and animals operate passively with only a very fuzzy sense of absolute range. Motion cues, such as occlusion, a priori knowledge, expectations concerning object sizes and characteristics, and contextual cues may be instrumental in enabling biological entities to detect obstacles while navigating through their environment, and may, in the final analysis, be necessary in order to solve the obstacle detection problem satisfactorily using passive sensors. Encapsulating these types of information, which are based on very abstract concepts, is very difficult and is an active area of
research in artificial intelligence and image and motion understanding. The solution of a number of problems in scene analysis and understanding as well as knowledge representation and utilization need to be solved before a reliable method for obstacle detection using a significant amount of knowledge-based reasoning can be developed.

The contents of section 4 describe our approach to motion analysis and describe some of our methods for dealing with the problems listed above.

4. INERTIAL SENSOR INTEGRATED MOTION ANALYSIS

The purpose of this section is to describe the inertial sensor integrated motion analysis approach we have undertaken. The block diagram of this system is illustrated in Figure 2. The system uses inertial sensor integrated optical flow, scene analysis, and selective applications of active sensors to provide obstacle detection capability. In this paper, we focus on the details of the inertial sensor integrated optical flow algorithm, which computes range to features within the sensor’s field of view. For a pair of image frames, the major steps that are involved within the optical flow algorithm are given below:

1. Input frames, frame A and frame B, are read in along with their associated inertial data.
2. Interest points are extracted from each of the input frames.
3. Location of the focus of expansion (FOE) (in both frames) is computed.
4. FOE and the interest points in frame B are projected onto an image plane that is parallel to the image plane that captured frame A (derotation of frame B).
5. Interest points in frame B are matched to those of frame A based upon four criteria.
6. Range is computed to each interest point in frame B that has a match in frame A.
7. A dense range map is created using context dependent scene analysis and interpolating between the computed range values.

Before starting a detailed discussion of the major steps in the algorithm, let us first describe the coordinate systems that are used. The digitized imagery contains pixels addressed by row and column with the origin of the 2-D coordinate system located in the upper left corner of the image. The horizontal axis, c, points to the right and the

![Diagram](image)

*Figure 2: Inertial sensor integrated optical flow and scene analysis using both passive and selective applications of active sensors provide robust image analysis useful for obstacle detection/avoidance by a robotic land vehicle or helicopter.*
vertical axis, \( r \), is in the downward direction. This image plane is perpendicular to the \( x \) axis of a 3-D coordinate system and is located at a distance of the focal length, \( F \), from the origin with the \( z \) axis in the downward direction. Therefore, the pixels in the image plane can be described in the 2-D coordinate frame as \( (c, r) \) and in the 3-D coordinate frame by the vector \( (F, y, z) \). The geometry described above is graphically illustrated in Figure 3.

As shown in Figure 4, the data input to the obstacle detection algorithm consists of a sequence of digitized video or FLIR frames that are accompanied by inertial data consisting of rotational and translational velocities. This information, coupled with the temporal sampling interval between frames, is used to compute the distance vector, \( d \), between each pair of frames and the roll, pitch and yaw angles, \((\phi, \theta, \psi)\), of each frame. Both \( d \) and \((\phi, \theta, \psi)\) are crucial to the success of the algorithm described in the following section. There are several other possible variations which we do not discuss in this paper.

**Figure 3:** The coordinate system geometry of the sensor's image plane is perpendicular to the \( x \) axis, located at the distance of the focal length, \( F \), from the origin of the coordinate system.

**Figure 4:** Inertial sensor integrated optical flow technique.
4.1 DISTINGUISHED FEATURES

The features within the imagery (TV or FLIR) that are most prominent and distinguished, mark the world points to which range measurements will be made. These prominent world points, known as interest points, are easy to extract from the imagery and have the highest promise of repeated extraction throughout multiple frames. The interest points within the field-of-view of the monocular sensor are of fundamental and critical importance to optical flow calculations. In the following subsections, the extraction and subsequent use of interest points is described in detail.

4.1.1 Interest Point Selection

The computation of distinguishable points is accomplished by passing a Moravec operator over each frame of imagery. The operator is applied to each image pixel (within a desired offset from the image border) which was identified as a strong edge pixel by a Sobel edge operator. The interest operator examines all pixels within a square window, of side length L, that surrounds each edge pixel and computes the relative variance between pixel values. As each pixel within the window is examined, the square of the difference between its value and the values of its neighboring pixels is computed and summed. Actually, four different sums are recorded which correspond to the same 4 neighbors relative to each pixel within the window; there is a sum for the square of the difference between the current pixel and its neighbor to the right and likewise for three other neighbors (below, below & right, below & left). After each pixel under the window has contributed to the 4 sums, the smallest of the sums, S, is selected and stored as the pixel’s value. A pixel is deemed an interest point if its assigned value of S is greater than the corresponding sum generated at each pixel within a square window of side length K, centered on the pixel in question. In the discussion that follows, a pixel’s value of S will be referred to as its interestingness.

Our implementation of the Moravec operator ranks the detected interest points (pixels with a value of S which is a local maximum) in the order of their computed interestingness. This interest point extraction routine divides the image into M uniform regions and returns only the N points within each region which have the highest values of S, where N and M are inputs to the program. The result of returning only the best interest points (in terms of S) in each region is that the processed scene is more uniformly covered with interest points. If this were not the case, a small number of occasionally adjacent regions will lay claim to the major portion of interest points.

Unfortunately, not all regions within a scene can contain reliable interest points (e.g. wave crests on a body of water are not good interest points). Scene analysis techniques are used to ascertain the goodness of regions prior to interest point selection. Moreover, interest point selection can be further improved by incorporation of Kalman filtering techniques, which use inertial sensor data to track and predict interesting point features.

4.1.2 Interest Point Derotation

To aid the process of interest point matching, we must make it seem as though image plane B is parallel to image plane A. If this is done, the FOE and pairs of interest points in frames A and B that match, would ideally be collinear should the image planes be superimposed (see Figure 5). To make the image planes parallel, derotation is performed for each vector, \( \hat{F}, \hat{y}, \hat{z} \) that corresponds to each interest point in frame B. The equation for the derotation transformation and projection (in homogeneous coordinates) is

\[
\begin{bmatrix}
F' \\
y' \\
z'
\end{bmatrix} = P R_{\Phi}^{-1} R_{\Psi}^{-1} R_{\Phi} R_{\Psi} R_{\Theta} R_{\Phi} \begin{bmatrix}
F \\
y \\
z
\end{bmatrix} = P C_{\text{NED}}^{\Phi} C_{\text{NED}}^{\Psi} P
\]

where

\[
R_{\Psi} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \cos\phi & \sin\phi & 0 \\
0 & -\sin\phi & \cos\phi & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

and where NED (north, east, down) is the coordinate frame in which inertial measurements are made. Use of the NED frame assumes that vehicle motion is "local" to a patch of Earth.

The matrix \( P \) projects a world point onto an image plane and is used to compute the FOE, \( \text{FOE} = P \hat{\alpha} \), where \( \hat{\alpha} = \alpha \hat{v} \). The matrix \( C_{\text{NED}}^{\Phi} \) converts points described in the NED coordinate frame into an equivalent description within a coordinate frame parallel to the A coordinate frame. Likewise, the matrix \( C_{\text{NED}}^{\Psi} \) converts the descriptions of points in the B coordinate frame into descriptions in a coordinate frame parallel to NED.
4.1.3 Interest Point Matching

The matching of interest points is performed in two passes. The goal of the first pass is to identify and store the top three candidate matches for each interest point in frame B, \((F, y_B, z_B)\). The second pass looks for multiple interest points being matched to a single point in frame A. Hence, the result of the second pass is a one-to-one match between the interest points in the two successive frames. For our application, a one-to-one match of interest points is necessary. We acknowledge that the projection onto the sensor's image plane of an object in the world will grow in size as the sensor moves toward the object. This situation might imply that a one-to-one match does not make sense since what was one pixel in size in frame A might become two or more pixels in size in frame B. In this work, we assume that the growth of objects, in terms of pixel size, is negligible in the passive ranging for obstacle detection scenario. All objects are assumed to be at certain safe distances for vehicle maneuvering and one pixel (of interest point quality) in two frames is all that is required of an object's surface for the range to the object to be computed.

Pass One:

To determine the candidate matches to \((F, y_B, z_B)\), each of the interest points in frame A is examined with the successive use of four metrics. The first metric makes certain that candidate matches lie within a cone shaped region bisected by the line joining the FOE and the interest point in frame B. This metric limits candidate matches to lie within the cone with apex at the FOE, as shown in Figure 6(a). If an interest point in frame A, \((F, y_A, z_A)\), passes the first metric, then the second metric is applied to it. The second metric requires that the interestiingness of candidate matches is close to the interestingness of the point that we are trying to match.

The third metric restricts all candidate matches in frame A to lie closer to the FOE than the points in frame B (as physical laws would predict for stationary objects). This metric involves the computation of the distances of the interest points from the FOE, which can be computed in two different ways:

---

Figure 5: An illustration of the sensor geometry that records two perspective views of a scene at two positions separated by a distance \(l_1 \Delta t = l_2 \Delta t\) (with no rotation of the sensor between positions). When there is no rotational change between image frames, there is a special property of the perspective projection of a world point onto the two image planes; the FOE and the projections of the world point are all colinear.
(1) The direct euclidean distance, $d_1$, between $(F_{y_{A}}, z_{A})$ and $(F_{y_{B}}, z_{B})$, and
(2) the distance $d_2$ which is the projection of $d_1$ onto the line joining $(F_{y_{B}}, z_{B})$ and the FOE.

The distance measures are graphically illustrated in Figure 6(b). Regardless of the way that the distance measure is computed, it can be used to identify the closest candidate matches to $(F_{y_{B}}, z_{B})$.

The fourth metric constrains the distance between an interest point and its candidate matches. For an interest point in frame $A$, $A_i$, to be a candidate match to point $B_i$, it must lie within the shaded region of Figure 6(a). The depth of the region is determined by this fourth metric while the width of the region is fixed by an earlier metric. By limiting interest points, $A_i$, to lie in the shaded region, we have effectively restricted the computed range of resulting matches to lie between $R_{\text{max}}$ and $R_{\text{min}}$. The reasoning behind this restriction is that world objects of range less than $R_{\text{min}}$ should not occur due to autonomous or manual navigation of the vehicle, thus avoiding potential collisions. Likewise, objects at a range greater than $R_{\text{max}}$ are not yet of concern to the vehicle.

The result of the first pass of interest point matching is a list, for each $(F_{y_{B}}, z_{B})$, of three or fewer candidate matches that pass all metrics and have the smallest distance measures of all possible matches.

Pass Two:

The goal of the second pass of the matching process is to take the matches provided by the first pass and generate a one-to-one mapping between the interest points in frames $A$ and $B$. Initially, it can be assumed that the best match to $(F_{y_{B}}, z_{B})$ will be the stored candidate match which has the smallest distance measure. Unfortunately, there may be multiple points, $(F_{y_{B}}, z_{B})$, which match to a single $(F_{y_{A}}, z_{A})$. Hence, the recorded list of best matches is searched for multiple occurrences of any of the interest points in frame $A$. If multiple interest points in frame $B$ have the same best match, then the point, $B_i'$, which is at the minimum distance from the $A_i$ in question, will retain this match and is removed from the matching process. The remaining $B_i'$'s are returned to the matching process for further investigation after having $A_i$ removed from their lists of best matches. This process continues until all of the interest points in frame $B$ either have a match, or are determined to be unmatchable by virtue of an empty candidate match list. Hence, the final result of the matching process is a one-to-one mapping between the interest points in frames $A$ and $B$.

4.2 RANGE CALCULATION AND INTERPOLATION

Given the result of interest point matching, which is the optical flow, range can be computed to each match. Given these sparse range measurements, a range or obstacle map can be constructed. The obstacle map can take many forms, the simplest of which consists of a display of bearing versus range. In what follows, range calculation is described and the important issue of range interpolation is discussed.

Given pairs of interest point matches between two successive image frames and the translational velocity between frames, it becomes possible to compute the range to the object on which the interest points lie. One approach to range,
R, computation is described by the equation
\[ R = \frac{\Delta Z}{\Delta t} \frac{x' - x_f}{x' - x} \]
where
- \(x_f\) = the distance between the FOE and the center of the image plane,
- \(x\) = the distance between the pixel in frame A and the center of the image plane,
- \(x'\) = the distance between the pixel in frame B and the center of the image plane,
- \(\Delta Z = \frac{\Delta t}{\cos \alpha}\) = the distance traversed in one frame time, \(\Delta t\), as measured along the axis of the line of sight,
- \(\alpha\) = the angle between the velocity vector and the line of sight,
- \(x' - x_f\) = the distance in the image plane between \((F, y_{B'}, z_{B'})\) and the FOE, and
- \(x' - x\) = the distance in the image plane between \((F, y_{B'}, z_{B'})\) and \((F, y_{A'}, z_{A'})\).

These variables are illustrated in Figure 7.

An alternate approach involves the calculation of the angles \(\alpha_A\) and \(\alpha_B\) between the translational velocity vector and the vectors that describe the matched pair of interest points in frames A and B,
\[ R_A = \frac{\Delta Z \sin \alpha_B}{\sin(\alpha_B - \alpha_A)} \]
as indicated in Figure 8. Both of the range calculating techniques compute the distance to a world point relative to the lens center of frame A (similar equations would compute the distance from the lens center of frame B). The accuracy of the range measurements that result from either approach is very sensitive to the accuracy of the matching process as well as the accuracy of the inertial measurement unit (IMU) data.

The task of range interpolation is the last processing step required of the passive ranging system (this ignores any postprocessing of the range that may be required before it gets passed to the automatic vehicle control and display systems). The purpose of this task is to create, by means of interpolation between the sparse range samples generated from the optical flow measurements, a dense range map representing the objects within the field of view. Essentially, this task is one of surface fitting to a sparse, nonuniform set of data points. To obtain an accurate surface fit that

Figure 7: The geometry involved in the first approach to range calculation is illustrated here. The figure shows the imaged world point in motion rather than the sensor, thus simplifying the geometry.
physically corresponds to the scene within the field of view, it is necessary that the sparse set of range samples be as uniformly spread throughout the field of view as possible. This will require processing steps hinted at in previous sections; scene understanding/segmentation must be used to create regions from which a desired number of interest points are extracted.

The type of surface fitting is important because the resulting surface (i.e. the range map) must pass through each of the range samples. It would be especially dangerous if the surface passed under any range samples. There are many techniques of surface fitting available to our task. To date, we have explored a method of bivariate interpolation over irregular spaced samples proposed by Akimo. This technique uses 5th degree polynomials to interpolate over the triangular regions formed by the range samples. The major drawback associated with this approach is its assumption that all of the given points fall within a convex region. A solution to this problem is to use an improved Delaunay-based triangulation of the range samples, proposed by DeFloriani et al., which works over arbitrarily shaped regions of interest.

A less elaborate technique of range interpolation consists of a fitting of planes to the available range samples. This approach gets the job done quickly and efficiently and does succeed in passing through each range sample. All techniques of range interpolation should be careful to avoid interpolation over discontinuities that occur between range samples on the surface under investigation. With the use of scene analysis/segmentation, the smoothing of discontinuities can be avoided by interpolating only over smooth regions or segments of the scene. Techniques of joining the regions after interpolation have yet to be developed.

Finally, there is some concern as to the purpose of interpolation. Surely, interpolation will aid an operator/pilot in the interpretation of the results of optical flow measurements, but its use by automatic vehicle control is questionable. Also, a large number of interest points can be selected and matched, so there may not be any need for interpolation. These issues are being explored further.

5. RESULTS

Our inertial navigation sensor integrated optical flow algorithm has been used to generate range samples using both synthetic data and real data (imagery and INS information) obtained from a moving vehicle. In this section, we describe the conditions under which the data was created/collected and provide images illustrating the results of the major steps in the optical flow algorithm.
The synthetic interest points were generated from a file containing the 3-D coordinates of 15 world points. Table 1 shows the 3-D locations of these world points. In the same coordinate system as the interest points are located, Table 2 lists the location, roll, pitch, and yaw of the camera at the two instances of time at which frames A and B were acquired. The time between frame acquisition is 0.2 seconds. Figure 9(a) shows the locations (circles) of the projection of the world points onto the first location of the image plane where the field of view of the synthesized camera model is 52.0° x 48.75° with a focal length of 9 mm. Figure 9(b) shows the locations (squares) of the projections of the world points onto the second location of the image plane and shows the new locations (diamonds) of those projections after derotation. Figure 9(c) shows the results of the matching process in which circles are connected to their corresponding diamond with a straight line and the POE is labeled and marked with an X. The final frame, Figure 9(d), shows the computed range to each point resulting from each of the matches.

A pair of real images was selected to test the capabilities of the optical flow algorithm using real imagery. Table 3 indicates the location, roll, pitch, and yaw of the camera associated with the pair of real image frames that were used. The field of view of the camera for the real images is 52.1° x 40.3° and the focal length = 9 mm. The elapsed time between the two frames for this experiment was 0.2 seconds. Figure 10(a) shows the locations of the extracted interest points obtained from the first frame, drawn as circles. Similarly, Figure 10(b) indicates the location of extracted interest points (squares) and the corresponding derotated locations (diamonds). Since the vehicle undergoes very little rotation between frames, the derotated locations are nearly coincident with the original point locations. The results of the point matching process for the real imagery is shown in Figure 10(c). Finally, the computed range to each of the matched points is displayed in Figure 10(d).

### 6. CONCLUSIONS

We have presented initial work for INS integrated motion analysis. Future work will involve incorporating context dependent qualitative scene analysis, knowledge-based sensor management, and incorporation of Kalman filtering into our approach, as shown in Figure 2. Our ultimate goal is to develop the complete, fieldable system for obstacle detection during rotorcraft low altitude flight. We are also applying this technology for land vehicle applications to achieve robust obstacle detection, target motion detection, and target tracking.

<table>
<thead>
<tr>
<th>x (ft)</th>
<th>y (ft)</th>
<th>z (ft)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>95</td>
<td>-30</td>
</tr>
<tr>
<td>3</td>
<td>90</td>
<td>-10</td>
</tr>
<tr>
<td>4</td>
<td>85</td>
<td>-5</td>
</tr>
<tr>
<td>5</td>
<td>80</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>75</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>70</td>
<td>-8</td>
</tr>
<tr>
<td>8</td>
<td>65</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>60</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>55</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>50</td>
<td>-15</td>
</tr>
<tr>
<td>12</td>
<td>35</td>
<td>10</td>
</tr>
<tr>
<td>13</td>
<td>30</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>25</td>
<td>-5</td>
</tr>
<tr>
<td>15</td>
<td>20</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 1: Locations of interest points.

<table>
<thead>
<tr>
<th>x (ft)</th>
<th>y (ft)</th>
<th>z (ft)</th>
<th>roll (deg)</th>
<th>pitch (deg)</th>
<th>yaw (deg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame A</td>
<td>0</td>
<td>-7</td>
<td>0</td>
<td>-15</td>
<td>0</td>
</tr>
<tr>
<td>Frame B</td>
<td>5</td>
<td>1</td>
<td>-6</td>
<td>5</td>
<td>-11</td>
</tr>
</tbody>
</table>

Table 2: Location, roll, pitch, and yaw of the camera for synthetic frames A and B.
Figure 9: Optical flow results using synthetic data. (a) Locations of interest points in the first image, indicated by circles. (b) Locations of interest points in the second image, shown using squares. Diamonds indicate the derotated interest point locations. (c) Matching process results in displacement vectors between circles and diamonds. The FOE is indicated by a cross. (d) Computed range values to the interest points.

<table>
<thead>
<tr>
<th></th>
<th>x (ft)</th>
<th>y (ft)</th>
<th>z (ft)</th>
<th>roll (deg)</th>
<th>pitch (deg)</th>
<th>yaw (deg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame A</td>
<td>-230.3</td>
<td>-20.72</td>
<td>6.43</td>
<td>0.959</td>
<td>-1.179</td>
<td>-176.737</td>
</tr>
<tr>
<td>Frame B</td>
<td>-231.7</td>
<td>-20.83</td>
<td>6.44</td>
<td>1.222</td>
<td>-1.231</td>
<td>-176.852</td>
</tr>
</tbody>
</table>

Table 3: Location, roll, pitch, and yaw of the camera for two frame of real imagery.
Figure 10: Optical flow results using real data. (a) Locations of interest points in the first image, indicated by circles. (b) Locations of interest points in the second image, shown using squares. Diamonds indicate the derotated interest point locations. (c) Matching process results in displacement vectors between circles and diamonds. The FOE is indicated by a cross. (d) Computed range values to the interest points.

ACKNOWLEDGEMENTS

This material is based upon work supported by NASA under contract NAS2-12800 and DARPA/U.S. Army ETL contract DACA 76-86-C-0017. The authors would like to thank Scott Snyder, Dave Duncan, and Paul Gader for discussions and their input to this work.
REFERENCES


**APPENDIX A**

**NAVIGATION ERROR CHARACTERISTICS**

The passive ranging technique relies on attitude and position information provided by the inertial navigation system (INS). The purpose of this appendix is to supply error models which are representative of a land vehicle/helicopter navigation system. These error models can then be used to assess the impact of navigational errors on the passive ranging performance. A short term navigation error model of a strapdown inertial navigation system has been developed. The INS derives attitude, velocity, and position based on inputs from orthogonal triads of gyros and accelerometers. Assuming that the only error of concern to the passive ranging technique is the drift between samples, Table A1 provides error estimates for the GG1328 Ring Laser Gyro (RLG) INS and the GG1342 RLG INS systems. The error estimates in the table are based on the following assumptions:

<table>
<thead>
<tr>
<th>30 Hz Sample Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>45 Degree Latitude and Heading</td>
</tr>
<tr>
<td>10 Minute Gyrocompass Alignment</td>
</tr>
<tr>
<td>0.5 g X and Y Helicopter Acceleration</td>
</tr>
<tr>
<td>1.0 g Z Helicopter Acceleration</td>
</tr>
<tr>
<td>100 deg/sec Roll, Pitch, and Yaw Rates</td>
</tr>
<tr>
<td>H-764 Output Quantization Levels</td>
</tr>
</tbody>
</table>

As evident from Table A1, the most significant one sigma error source is the output quantization error. The output quantization is a function of output requirements as internally the attitudes, velocities, and positions are all double precision quantities. Thus, the small errors in attitude and velocity vectors obtained from the INS allow the location of
the focus of expansion (FOE) in the original and derotated images to be accurate enough to perform robust motion analysis. Further details of this analysis are given in the report by Bhanu and Roberts.  

<table>
<thead>
<tr>
<th>Error Term</th>
<th>Level Error (mrad)</th>
<th>Heading Error (mrad)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GG1328</td>
<td>GG1342</td>
</tr>
<tr>
<td>Gyro Bias</td>
<td>0.1600e-04</td>
<td>0.1600e-05</td>
</tr>
<tr>
<td>Gyro Random Walk</td>
<td>0.1585e-02</td>
<td>0.2642e-03</td>
</tr>
<tr>
<td>Gyro Scalefactor</td>
<td>0.1728e-02</td>
<td>0.8639e-03</td>
</tr>
<tr>
<td>Gyro Misalignment</td>
<td>0.9647e-04</td>
<td>0.6431e-04</td>
</tr>
<tr>
<td>Gyro Misalignment</td>
<td>0.9647e-04</td>
<td>0.6431e-04</td>
</tr>
<tr>
<td>Heading Error</td>
<td>-1.404e-07</td>
<td>-1.790e-08</td>
</tr>
<tr>
<td>RSS Total</td>
<td>0.9080e-03</td>
<td>0.9034e-03</td>
</tr>
<tr>
<td>Output Quantization</td>
<td>0.0141e-00</td>
<td>0.0141e-00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Error Term</th>
<th>X - Y Velocity Error (fps)</th>
<th>Z Velocity Error (fps)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GG1328</td>
<td>GG1342</td>
</tr>
<tr>
<td>Accel Bias</td>
<td>0.1063e-03</td>
<td>0.5313e-04</td>
</tr>
<tr>
<td>Accel Scalefactor</td>
<td>0.1063e-03</td>
<td>0.5313e-04</td>
</tr>
<tr>
<td>Accel Misalignment</td>
<td>0.2576e-04</td>
<td>0.1288e-04</td>
</tr>
<tr>
<td>Accel Misalignment</td>
<td>0.1030e-03</td>
<td>0.5152e-04</td>
</tr>
<tr>
<td>Level Error</td>
<td>0.1063e-03</td>
<td>0.5314e-04</td>
</tr>
<tr>
<td>RSS Total</td>
<td>0.2125e-03</td>
<td>0.1063e-03</td>
</tr>
<tr>
<td>Output Quantization</td>
<td>0.1409e-03</td>
<td>0.1409e-03</td>
</tr>
<tr>
<td>RSS Total</td>
<td>0.2550e-03</td>
<td>0.1765e-03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Error Term</th>
<th>X - Y Position Error (feet)</th>
<th>Z Position Error (feet)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GG1328</td>
<td>GG1342</td>
</tr>
<tr>
<td>Accel Bias</td>
<td>0.3507e-05</td>
<td>0.1753e-05</td>
</tr>
<tr>
<td>Accel Scalefactor</td>
<td>0.3507e-05</td>
<td>0.1753e-05</td>
</tr>
<tr>
<td>Accel Misalignment</td>
<td>0.8500e-06</td>
<td>0.4250e-06</td>
</tr>
<tr>
<td>Accel Misalignment</td>
<td>0.3400e-05</td>
<td>0.1700e-05</td>
</tr>
<tr>
<td>Level Error</td>
<td>0.3508e-05</td>
<td>0.1754e-05</td>
</tr>
<tr>
<td>RSS Total</td>
<td>0.7013e-05</td>
<td>0.3506e-05</td>
</tr>
<tr>
<td>Output Quantization</td>
<td>0.2528e-00</td>
<td>0.2528e-00</td>
</tr>
<tr>
<td>RSS Total</td>
<td>0.2528e-00</td>
<td>0.2528e-00</td>
</tr>
</tbody>
</table>

Table A1: One Sigma, Short Term Navigation Errors
Image Understanding Workshop

Proceedings of a Workshop
Held at
Palo Alto, California

May 23-26, 1989

Sponsored by:
Defense Advanced Research Projects Agency
Information Science and Technology Office

This document contains copies of reports prepared for the DARPA Image Understanding Workshop. Included are Principal Investigator reports and technical results from both the basic and strategic computing programs within DARPA/ISTO sponsored projects and certain technical reports from selected scientists from other organizations.

APPROVED FOR PUBLIC RELEASE
DISTRIBUTION UNLIMITED

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the United States Government.