

# Integrated Binocular and Motion Stereo in An Inertial Navigation Sensor-Based Mobile Vehicle

Bir Bhanu, Peter Symosek\*, Scott Snyder\*  
Barry Roberts\* and Subhudev Das

College of Engineering  
University of California  
Riverside, CA 92521-0425

\* Honeywell Systems and Research Center  
3660 Technology Drive  
Minneapolis, MN 55418

## Abstract

Range measurements to objects in the world relative to mobile platforms such as ground or air vehicles are critical for visually aided navigation and obstacle detection/avoidance. This paper presents an approach that consists of a synergistic combination of two types of passive ranging methods: binocular stereo and motion stereo. We show a new way to model the errors in binocular and motion stereo in conjunction with an inertial navigation system and derive the appropriate Kalman filter to refine the estimates from these two stereo ranging techniques. We present results using laboratory images that show that refined estimates can be optimally combined to give range values which are more accurate than any one of the individual estimates from binocular and motion stereo. By incorporating a blending filter, the approach has the potential of providing accurate, dense range measurements for all the pixels in the field of view.

## 1. INTRODUCTION

Range measurements to objects in the world relative to mobile platforms such as ground or air vehicles are critical for visually aided navigation and obstacle detection/avoidance. Active (laser) range sensors can be used to provide such range measurements although they have a limited field of view, suffer from slow data acquisition, and are expensive. Robust passive ranging techniques can be suitable alternative. The passive visual cues of binocular and motion stereo have been the two most popular methods for range estimation. A plethora of algorithms have been proposed to estimate three-dimensional (3D) structure or motion or both, using these two cues individually, but few have been demonstrated to be robust enough for real-world applications such as autonomous mobile robots operating in outdoor scenarios. Some of the real-world factors, such as vibration of the platform on which the cameras are mounted or the wind speed, may prove to be catastrophic to the ranging techniques, such as determination of motion parameters in order to compute range. Incorporation of hardware which can compute stable values of motion parameters under harsh operating conditions, will greatly improve the performance of any motion analysis technique necessary for motion stereo-based ranging. Inertial Navigation System (INS) is one such hardware which is used in many types of land and air vehicles.

The objective of this research is to develop an INS-integrated passive ranging system that utilizes the benefits of binocular and motion stereo. This system is based on the synergistic combination of the two stereo modalities which is achieved by the following sequence of operations: interest point matching, Kalman filtering, and range

measurement blending. The important benefits of the proposed synergistic system are,

- The system is cheap to build (compared to active sensors).
- It is passive (i.e., non-detectable, covert).
- A more dense and more accurate range map is generated than is possible by either passive technique alone which is necessary for obstacle avoidance.
- Negligible motion distortion is caused by the moving platform (i.e., fast data acquisition).

Previous efforts in the derivation of approaches for the synergistic combination of binocular and motion stereo ranging have placed restraints on the problem specification to reduce the complexity of the analysis. To date, no demonstration of a totally general, comprehensive characterization of the ranging problem for multiple binocular stereo frames has been derived.

The emphasis of this paper is on modeling the errors in binocular and motion stereo in conjunction with an INS for a real-world application of the passive ranging system, and deriving the appropriate Kalman filter to refine the estimates from these two stereo ranging techniques. Our particular approach is designed to allow empirical evaluation of the performance and robustness of the passive ranging system for various scenarios. The next section describes in greater detail the background and motivation behind the work reported in this paper. Section 3 presents the technical approach adopted in designing the synergistic system. Section 4 discusses results obtained during an empirical evaluation of the performance of the system with the laboratory data and simulated Inertial Reference Unit data. The last section presents the concluding remarks.

## 2. BACKGROUND AND MOTIVATION

In this section we summarize the past research related to the work reported in this paper, and the motivations that lead to the development of the approach described in the following section.

### 2.1 Background

Features from stereo pairs of images can be matched over time to obtain better accuracy for disparity-based range calculations. Sridhar and Suorsa<sup>9</sup> describe recursive binocular and motion stereo algorithms and compare their performances. However, the confidence factors for each of the range measurements which form the basis of such

comparison are obtained by considering only the errors in image locations of matched feature points. The uncertainty models of their passive ranging techniques are therefore inadequate for a real-world imaging system such as a mobile platform. Several researchers have used Kalman filtering method to estimate range from binocular stereo images<sup>4</sup> and motion sequences.<sup>7</sup>

## 2.2 Motivation

A synergistic combination of binocular and motion stereo is motivated by the following observations about their relative performance as illustrated in Figure 1: binocular stereo-based range computations suffer the greatest error at the edges of the camera's field of view (FOV) where motion stereo-based range is most accurate; the converse scenario holds true in the vicinity of the focus of expansion (FOE) where motion stereo-based range error is very large and binocular stereo-based range error is very small. Thus, a passive ranging system which employs only one of these two methods of range computation, is likely to perform poorly even with the most robust method. On the other hand, a passive ranging system which can successfully employ both, has the advantage of retaining only the best range estimate of a scene point from one of the methods determined by the location of the point in the FOV. This may mean that the visual field can be appropriately segmented to be processed by either binocular stereo or motion stereo, thereby reducing the computational burden. Alternately, range values for distinct points in the visual field can be computed from both binocular and motion stereo and be refined using the statistics of their uncertainties. The refined range estimates for each point can be statistically combined to yield a more precise range value.

An INS includes an Inertial Reference Unit (IRU) and all necessary hardware for stabilizing and processing the IRU outputs to derive values for the position and velocity (of whatever platform to which INS is attached) in a desired reference frame.<sup>3</sup> IRU measurements are made with gyroscopes, to provide an absolute measure of the rotation difference between the vehicle's coordinate frame and a fixed, geographic, reference frame; such measurements are also made with accelerometers, to provide the vehicle's acceleration relative to the reference frame, time integral of which gives the vehicle's velocity and position. Access to camera's translational and rotational motion information is significantly important. Without this information a motion-based ranging method will be required to explicitly derive the motion parameters before any range value can be computed. The estimation of motion parameters using currently available techniques is extremely noise sensitive.<sup>2</sup> A synergistic combination model for binocular and motion stereo must also include the INS-based variables in addition to the sensory parameters so as to predict and refine the complete error model for the ranging system whose parameters are drawn from both the imaging sensors and the INS. Such considerations make the synergistic combination very effective as will be discussed in the next section.

## 3. INTEGRATED APPROACH

With a two camera system in motion, a stereo ranging system is formed which is capable of *binocular stereo* and *motion stereo* range computations. In the case of binocular stereo two cameras are rigidly mounted on the same fixture such that their optical axes are parallel and yet laterally displaced by a fixed, known distance; the

cameras are longitudinally displaced, due to forward vehicle motion, for motion stereo. On a moving platform, the same two cameras can provide the imagery required to perform one binocular and two motion stereo range calculations.

Our integrated stereo system shown in Figure 2 uses two key elements which constitute the unique features of our approach:

- (1) matching of interesting points in binocular stereo and motion stereo imagery,
- (2) modeling of range errors present in the motion and binocular stereo techniques. These errors are represented as the states of a Kalman filter applied to obtain improved estimates of range values.

The coincident points of interest, i.e., those points for which range is computed by both motion and binocular stereo techniques, are used as measurements to estimate errors in the ranging processes. The points in the range maps which are not coincident can be corrected with these error estimates, improving the overall quality of the composite range map. This can be achieved with the use of a blending filter as shown Figure 3. This filter derives a composite range map for each measurement location as the weighted average of the Kalman filter estimates for the range, where the averaging weights are the current estimates of the measurement noise obtained from each filter. The confidence in each range measurement is inversely proportional to the estimate of the measurement noise, so that when the measurement noise for the binocular stereo algorithm is large, the estimate obtained from the motion stereo Kalman filter is weighted more heavily and vice versa.

### 3.1 Range Error Modeling

The disagreement between the calculated ranges from the motion and binocular stereo algorithms for the coincident points of interest is attributed to the errors in inertial data and geometric alignment of the cameras. The computed discrepancies in the range values are used by a Kalman filter to refine the estimates for the errors in the inertial and system configuration parameters. New estimates could be obtained by adding the updated estimates for the errors to the expected system variable magnitudes. The refined estimates could then be used to calculate improved binocular and motion stereo ranges. Alternatively, the H-matrices for each coincident interesting point could be derived from the dependence of errors for binocular and motion range calculations on errors in the inertial and system configuration data, and the output ranges for these algorithms could be corrected by linear combination of the error states of the Kalman filter. The latter is done in the current implementation of the passive ranging system.

The measurement for the filter's binocular stereo component is the difference of the ranges from binocular and motion stereo; the filter's motion stereo measurement is negative of the filter's binocular stereo measurement. Using the static Gauss-Markov discrete time model, the measurement process is described as follows:

$$y_{mj}(k) = R_{mj} - R_{sj} = H x_m + v_m \quad (1)$$

$$y_{sj}(k) = R_{sj} - R_{mj} = H x_s + v_s, \quad (2)$$

where  $y_{mj}(k)$  is the measurement for the motion stereo component of the Kalman filter for the  $j$ 'th feature point

location at time  $k$  and  $y_{s,j}(k)$  is that for the binocular stereo component;  $R_{m,j}$  is the estimate of the range corresponding to the  $j$ 'th feature point from the motion stereo algorithm and  $R_{s,j}$  is that from the binocular stereo algorithm;  $x_m$  is the error state vector for the motion stereo Kalman filter and  $x_s$  is that for the binocular stereo Kalman filter;  $v_m$  is measurement noise for the motion stereo Kalman filter that is large near FOE and small near periphery and  $E\{v_m^T v_m\} = \sigma_M^2$ ; and  $v_s$  is measurement noise for the binocular stereo Kalman filter that is small near FOE and large near periphery and  $E\{v_s^T v_s\} = \sigma_B^2$ .

As stated previously, the binocular stereo and motion stereo range errors are linear combinations of the Kalman filter error states. This linear combination can be expressed as:

$$\delta R_s = H\hat{x} \quad \text{and} \quad \delta R_m = H\hat{x},$$

where  $H$  is the measurement matrix defined by the total differential of binocular stereo range and the total differential of motion stereo range, respectively, and  $\hat{x}$  is the estimated error state vector.

The total differential of motion stereo range is:

$$\begin{aligned} dR_f = & \frac{\partial R_f}{\partial y'} dy' + \frac{\partial R_f}{\partial z'} dz' + \frac{\partial R_f}{\partial y} dy + \frac{\partial R_f}{\partial z} dz + \\ & \frac{\partial R_f}{\partial f_{ov_h}} df_{ov_h} + \frac{\partial R_f}{\partial f_{ov_v}} df_{ov_v} + \frac{\partial R_f}{\partial F} dF + \\ & \frac{\partial R_f}{\partial \Delta\Psi'} d\Delta\Psi' + \frac{\partial R_f}{\partial \Delta\theta'} d\Delta\theta' + \frac{\partial R_f}{\partial \Delta\phi'} d\Delta\phi' + \\ & \frac{\partial R_f}{\partial v_x} dv_x + \frac{\partial R_f}{\partial v_y} dv_y + \frac{\partial R_f}{\partial v_z} dv_z, \end{aligned} \quad (3)$$

where  $(y',z')^\dagger = \text{pixel location of an interest point in the left frame of a motion stereo pair of images that is acquired at time } t_{i+1}$ ;  $(y,z) = \text{pixel location of the interest point in the left frame of a motion stereo pair of images that is acquired at time } t_i \text{ and matches } (y',z')$ ;  $f_{ov_v} = \text{camera vertical field-of-view}$ ;  $f_{ov_h} = \text{camera horizontal field-of-view}$ ;  $\Delta\Psi' = \text{change in yaw angle that occurred in the time interval } t_{i+1}-t_i$ ;  $\Delta\theta' = \text{change in pitch angle that occurred in the time interval } t_{i+1}-t_i$ ;  $\Delta\phi' = \text{change in roll angle that occurred in the time interval } t_{i+1}-t_i$ ;  $(v_x, v_y, v_z) = \text{the velocity of the camera}$ ; and  $F = \text{the focal plane to lens center distance}$ .

The total differential of binocular stereo range is:

$$\begin{aligned} dR_f = & \frac{\partial R_f}{\partial y_l} dy_l + \frac{\partial R_f}{\partial z_l} dz_l + \frac{\partial R_f}{\partial y_r} dy_r + \frac{\partial R_f}{\partial z_r} dz_r + \\ & \frac{\partial R_f}{\partial f_{ov_h}} df_{ov_h} + \frac{\partial R_f}{\partial f_{ov_v}} df_{ov_v} + \frac{\partial R_f}{\partial \Delta\Psi} d\Delta\Psi + \frac{\partial R_f}{\partial \Delta\theta} d\Delta\theta + \\ & \frac{\partial R_f}{\partial \Delta\phi} d\Delta\phi + \frac{\partial R_f}{\partial F} dF + \frac{\partial R_f}{\partial a} da, \end{aligned} \quad (4)$$

where  $(y_l, z_l) = \text{pixel location of an interest point in the left frame of a binocular stereo pair of images acquired at time } t_i$ ;  $(y_r, z_r) = \text{pixel location of an interest point in the}$

† Using a right-handed coordinate system, where x-axis is parallel to the forward direction of travel, y-axis points rightward and z-axis points down. Image coordinates will be denoted by (u,v).

right frame of a binocular stereo pair of images acquired at time  $t_i$  and matches  $(y_r, z_r)$ ;  $\Delta\Psi = \text{the boresight yaw angle}$ ;  $\Delta\theta = \text{the boresight pitch angle}$ ;  $\Delta\phi = \text{the boresight roll angle}$ ; and  $a = \text{camera separation distance}$ .

In the above, we have given only the functional form of range errors. The complete equations for partial derivatives are quite complicated and for clarity we have not presented them here. We have also derived the functional relationships between the variance of range error and the location of an interest point in the field of view. Further details of these steps may be found elsewhere.<sup>1</sup>

An approximation to the range calculation error for the case of motion stereo range computations is,<sup>10</sup>

$$\sigma_M(u_A, v_A) = \sigma_{D_M} \frac{\Delta R_M(u_A, v_A)}{\sqrt{F^2 + u_L^2 + v_L^2}}, \quad (5)$$

where  $\sigma_{D_M} = \text{an initial estimate of the range calculation error due to the error in the motion stereo point matching algorithm}$ ;  $\Delta R_M(u_A, v_A) = \text{is the computed error in range for the world point whose projection onto the image plane is described in three space by } (F, u_A, v_A)$ ; and  $F = \text{is the distance between the lens center and the image plane}$ .

Likewise, an approximation to the range calculation error for the case of binocular stereo range computations is,

$$\sigma_S(u_L, v_L) = \sigma_{D_S} \frac{\Delta R_S(u_L, v_L)}{\sqrt{F^2 + u_L^2 + v_L^2}}. \quad (6)$$

The variances of the measurement noises  $v_m$  and  $v_s$  of (1) and (2) are calculated using these approximations.

In computing range with either the motion stereo or binocular stereo techniques, all range measurements are made relative to the *first of a temporal pair of images (i.e., A of A and B images) and the left image of a stereo pair*, as shown in Figure 4. Hence the subscripts A and L are used for the variables that describe points in three space on the image plane. In our implementation, the A and L images are the same image.

### 3.2 Kalman Filter Implementation

The twenty-nine error states summarized in Table 1 are mechanized in the Kalman filter. The first seven states<sup>6</sup> are based on the level axis "PSI-Angle" IRU error model:

$$\dot{\Psi} = -(\rho + \Omega) \Psi - C \delta\omega \quad (7)$$

$$\dot{\delta V} = C \delta A^B - \Psi \times A^L (\delta R \cdot R/R) (R/R) + \delta g' \quad (8)$$

$$\dot{\delta R} = \delta V - \rho \times \delta R, \quad (9)$$

where  $\Psi = \text{Psi-angle error (states 1, 2, and 3)}$ ;  $\delta V = \text{Psi-angle horizontal velocity error (states 4 and 5)}$ ;  $\delta R = \text{Psi-angle horizontal position error (states 6 and 7)}$ ;  $\rho = \text{local level transport rotation rate (V/R)}$ ;  $\Omega = \text{Earth rate in local level coordinate frame}$ ;  $C = \text{Body to Local Level Direction Cosine Transformation Matrix}$ ;  $\delta\omega = \text{Gyro error states (states 25, 26, 27)}$ ;  $\delta A^B = \text{Accelerometer error states (states 28 and 29)}$ ;  $A^L = \text{Local level acceleration}$ ;  $\omega_s = \text{Shuler frequency } (\sim 0.00125 \text{ rps})$ ;  $R/R = \text{unit vector}$ ; and  $\delta g' = \text{gravity deflection and anomaly errors}$ .

The vertical error states (8, 9, 10) assume an IRU vertical channel damped with a reference altitude from a radar altimeter. The error model implemented in the

Kalman filter can be expressed as

$$\dot{x}_8 = -x_9, \quad \dot{x}_9 = K_1 x_9 + x_{10}, \quad \dot{x}_{10} = x_{24} + K_2 x_9 - K_3 x_8, \quad (10)$$

where  $K_1, K_2, K_3$ , are the vertical channel gains. These gains were selected as 0.6, 0.15, and 0.0156, respectively. The remaining error states are modeled as Gauss-Markov processes with large time constants:

$$\dot{x} = \frac{-1}{\tau} x + \eta \quad (11)$$

where  $\eta$  is a white noise process and  $\tau$  is the time constant. The large time constants effectively model the error sources as constants.

#### 4. IMPLEMENTATION AND RESULTS

In this section, we present details of implementing the synergistic combination of binocular and motion stereo.

##### 4.1 Implementation Details

For the purposes of efficiency, only one Kalman filter is used by the integrated system by stacking the binocular and motion stereo measurements into a single  $2N \times 1$  column vector, where  $N$  is the number of feature points matched by both algorithms for a specific image. The H-matrix is obtained by stacking the total differential of binocular stereo range and the total differential of motion stereo range into a single  $2N \times 28$  matrix, where 28 is the number of states of the integrated system for Kalman filter.

IRU errors are simulated by running an off-line IRU error simulation and adding the errors onto our nominal motion. The simulation used is a monte-carlo simulation of the IRU error equations. The attitude of all experiments is chosen to be level and in a northerly direction cruise at 15 ft/sec. For this cruise scenario, IRU errors are essentially a function of time. Therefore, to formulate IRU errors for our two integrated system stereo cases, the true trajectory is subtracted from the simulated data. The resultant error data is then added to the integrated system trajectory to simulate corrupted IRU data.

The following are the selected camera parameters: horizontal field of view,  $hfov = 0.754160$  rad; vertical field of view,  $vfov = 0.313147$  rad; focal length,  $F = 0.041$  ft; baseline,  $a = 2$  ft.

##### 4.2 Experimental Results

Five frames (each  $512 \times 512$  pixels) of video data were collected in the laboratory at 2 foot intervals. An example of the experimental data is shown in Figure 5. To simulate motion for the motion stereo algorithm, we chose two velocities, 2 ft/sec and 20 ft/sec. From these five frames, the interest points which have the highest promise of repeated extraction throughout multiple frames are extracted using a combination of the Hessian and Laplacian operators.<sup>8</sup> The binocular stereo ranges are calculated to various points using the well-known Marr-Poggio-Grimson algorithm.<sup>5</sup>

To aid the process of interest point matching, each vector,  $(F, y_j, z_j)$  corresponding to the  $j$ 'th interest point in the frame  $m+1$ , is derotated so that the image plane  $m+1$  (sequence  $m+1$ ) appears to be parallel to image plane  $m$  (sequence  $m$ ). The processing of frame  $m$  consists of the following steps: the left and right binocular stereo images ( $L_m$  and  $R_m$ ) are matched; the left image

frames  $L_m$  and  $L_{m+1}$  are matched by motion analysis. 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  $m+1$ . The second pass looks for multiple interest points being matched to a single point in frame  $m$ . The range computations are further improved (for three or more sequential frames) by predicting and smoothing the range to each interest point that can be tracked through multiple frames.

The output binocular stereo and motion stereo range files, and simulated IRU data files are read into the Kalman filter software. The filter software runs a range matching algorithm to detect coincident range points. For each coincident point the corresponding H-matrix and filter measurements are calculated and processed by the filter.

Results from processing "Frame 1" of the sequence with the Kalman filter are shown in Table 2. The results of Table 2 are simulated with IRU noise and a 1 Hz video frame iteration rate. Ground truth measurements for the 12 matched feature point locations of "Frame 1" are presented in Table 3. The center of the image plane is the origin of the pixel coordinates;  $(y_l, z_l)$  in the left image matches  $(y_r, z_r)$  in the right image of "Frame 1", while  $(y'_l, z'_l)$  in the left image of "Frame i+1" matches  $(y_l, z_l)$ .

As shown in Table 2, the corrections added to the binocular stereo range and motion stereo range tend to converge the solutions to a common point as expected, i.e., the corrected range values are in the direction (increasing or decreasing) as that of the ground truth values with respect to the raw range values. In general this behavior can be observed in the results for measurements 1 through 12. There are some exceptions (measurement 8 and 10) which could possibly be due to the measurement weighting. Since the results are for only a pair of frames, the actual convergence of the corrected range values cannot be seen.

Table 4 contains results from processing the first and second "frames" for the 20 ft/sec velocity case (10 Hz video frame iteration rate). Results for "Frame 1" processing are good; the revised range estimates for the binocular and motion stereo ranging algorithms are converging to a unique value with the exception of measurements 4 and 9. Ground truth measurements for the 13 matched feature point locations of "Frame 1" and the 13 matched feature point locations of "Frame 2" are presented in Table 5. It is to be noted that the same scene points matched in "Frame 1" will not necessarily appear in the results of processing "Frame 2." Therefore, the ground truth range values of the corresponding measurements between "Frame 1" and "Frame 2" in Table 5 are not for the same scene point.

#### 5. CONCLUSIONS

We presented the basic concept and results of our binocular and motion stereo synergistic system. These results demonstrate that it is possible to effectively combine binocular and stereo range measurements by incorporating an inertial navigational sensor. The approach has the potential of integrating the two techniques using a blending filter, thus providing accurate range value for any the pixel in the field of view. We plan to do this in the future.

## REFERENCES

1. B. Bhanu and B. Roberts, "Obstacle Detection During Rotorcraft Low-Altitude Flight and Landing," Honeywell Systems and Research Center Report to NASA (July, 1990).
2. J.-Q. Fang and T. S. Huang, "Some Experiments on Estimating the 3-D Motion Parameters of a Rigid Body from Two Consecutive Image Frames," *IEEE Trans. Patt. Anal. Mach. Intell.* PAMI-6 pp. 545-554 (1984).
3. J. A. Farrell, *Integrated Aircraft Navigation*, Academic Press, Inc. (1976).
4. O. D. Faugeras, N. Ayache, and B. Faverjon, "Building Visual Maps by Combining Noisy Stereo Measurements," *Proceedings of the International Conference on Robotics and Automation*, pp. 1433-1438 (April 1986).
5. W. E. L. Grimson, *From Images to Surfaces: A Computational Study of the Human Early Visual System*, MIT Press, Cambridge, MA (1981).
6. M. Ignagni, "Error Analysis of a Strapdown Terrestrial Navigation System," Technical Report, System and Control Sciences (1989).
7. L. Matthies, T. Kanade, and R. Szeliski, "Kalman Filter-based Algorithms for Estimating Depth from Image Sequences," *International Journal of Computer Vision* 3 pp. 209-236 (1989).
8. H. H. Nagel, "Displacement Vectors Derived from Second-Order Intensity Variations in Image Sequences," *Computer Vision, Graphics, and Image Processing* 21 pp. 85-117 (1987).
9. B. Sridhar and R. Suorsa, "Comparison of Motion and Stereo Sensors in Passive Ranging Systems," *IEEE Trans. Aero. Elec. Sys.* 27 pp. 741-746 (July, 1991).
10. P. Symosek, "Error Modeling in Range Measurements from Binocular and Motion Stereo," *Honeywell Systems and Research Center Internal Report*, (1990).

\* The authors are grateful to Honeywell Inc. where this work was supported by an initiative grant.

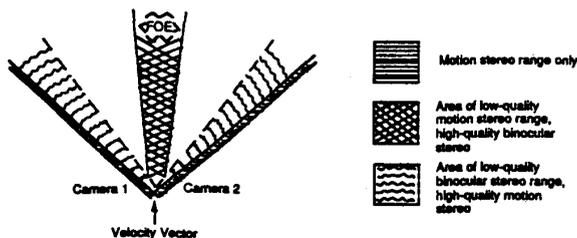


Figure 1: Modalities for passive ranging.

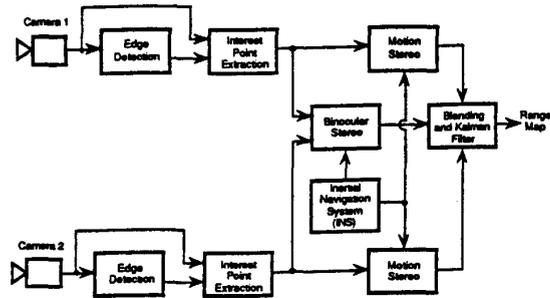


Figure 2: System for integration of binocular and motion stereo.

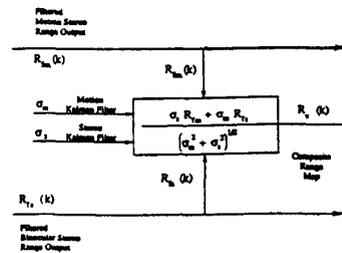


Figure 3: Composite range map/blending filter.

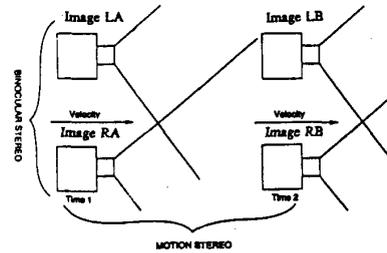


Figure 4: Binocular and motion stereo data acquisition.

Table 1: Error states used in Kalman filtering.

| Error State | Description   |
|-------------|---|
| 1           | IRU psi 1 angle error                                       |
| 2           | IRU psi 2 angle error                                       |
| 3           | IRU psi 3 angle error                                       |
| 4           | IRU x velocity error  |
| 5           | IRU y velocity error  |
| 6           | IRU x position error  |
| 7           | IRU y position error  |
| 8           | Vertical channel acceleration error                         |
| 9           | Vertical channel velocity error                             |
| 10          | Vertical channel position error                             |
| 11          | Horizontal FOV error ( $fov_x$ )                            |
| 12          | Vertical FOV error ( $fov_y$ )                              |
| 13          | Camera focal plane to lens center distance (F)              |
| 14          | $y_l$ left camera Y optical axis offset error               |
| 15          | $z_l$ left camera Z optical axis offset error               |
| 16          | $y_r$ right camera Y optical axis offset error              |
| 17          | $z_r$ right camera Z optical axis offset error              |
| 18          | Camera yaw angle bore-sight error                           |
| 19          | Camera pitch angle bore-sight error                         |
| 20          | Camera roll angle bore-sight error                          |
| 21          | Camera separation distance (s)                              |
| 22          | $y_l'$ left camera Y optical axis offset error (past frame) |
| 23          | $z_l'$ left camera Z optical axis offset error (past frame) |
| 24          | Z accelerometer bias  |
| 25          | X gyro bias error   |
| 26          | Y gyro bias error   |
| 27          | Z gyro bias error   |
| 28          | X accelerometer bias error                                  |
| 29          | Y accelerometer bias error                                  |



(a)



(b)



(c)

Figure 5: Laboratory image database. (a) Frame 1 image obtained from left camera of stereo pair. (b) Frame 1 image obtained from right camera of stereo pair. (c) Frame 2 image obtained from left camera of stereo pair.

Table 2: Kalman filter-compound range errors for 1 Hz processing rate (2 ft/sec velocity).

| Time = 1.0 sec (Frame 1) |                     |                  |                               |                            |                           |                        |
|--------------------------|---------------------|------------------|-------------------------------|----------------------------|---------------------------|------------------------|
| Measurement              | Raw Binocular Range | Raw Motion Range | Kalman Filter Binocular Error | Kalman Filter Motion Error | Corrected Binocular Range | Corrected Motion Range |
| 1                        | 23.470947           | 13.229049        | 9.257078                      | -1.906973                  | 14.213869                 | 15.160212              |
| 2                        | 15.250125           | 19.987539        | -2.715582                     | -1.591671                  | 17.965708                 | 21.579220              |
| 3                        | 23.286278           | 11.497955        | 5.687111                      | -4.775131                  | 17.599167                 | 16.273087              |
| 4                        | 17.710770           | 22.075123        | -1.643602                     | 1.591987                   | 19.354372                 | 20.483137              |
| 5                        | 13.850588           | 20.908190        | -3.369545                     | 7.472088                   | 17.220133                 | 13.436102              |
| 6                        | 15.973729           | 21.540310        | -3.716971                     | 2.922868                   | 19.350700                 | 18.617441              |
| 7                        | 16.092087           | 17.760782        | 1.714687                      | 2.406360                   | 17.806774                 | 15.354422              |
| 8                        | 16.151932           | 14.471107        | -2.965745                     | 2.350610                   | 19.117678                 | 12.120497              |
| 9                        | 21.183895           | 22.302511        | 2.563853                      | 2.367056                   | 18.620041                 | 19.935455              |
| 10                       | 15.358275           | 14.538172        | -3.013412                     | 1.480840                   | 18.371687                 | 13.057332              |
| 11                       | 18.167021           | 20.215263        | -0.081987                     | 1.600948                   | 18.083033                 | 18.614315              |
| 12                       | 15.797955           | 21.457747        | -1.679191                     | 1.594449                   | 17.477146                 | 19.863297              |

Table 3: Ground truth measurements for 1 Hz processing rate.

| Time = 1.0 sec (Frame 1) |                 |                 |                |                |                 |            |  |
|--------------------------|-----------------|-----------------|----------------|----------------|-----------------|------------|--|
| Measurement              | $y'_1$ (pixels) | $z'_1$ (pixels) | $y_r$ (pixels) | $z_r$ (pixels) | $\eta$ (pixels) | $R_f$ (ft) |  |
| 1                        | -152            | -59             | -182           | -53            | -128            | 13.9764    |  |
| 2                        | -52             | 122             | -131           | 114            | -46             | 18.8474    |  |
| 3                        | -47             | -50             | -94            | -38            | -38             | 12.2229    |  |
| 4                        | -12             | 113             | -83            | 106            | -9              | 19.1300    |  |
| 5                        | 41              | 65              | -58            | 37             | 60              | 19.2826    |  |
| 6                        | 49              | 202             | -39            | 189            | 44              | 20.2861    |  |
| 7                        | 92              | -61             | 0              | -49            | 82              | 18.0941    |  |
| 8                        | 95              | 167             | 0              | 149            | 82              | 14.0358    |  |
| 9                        | 124             | -15             | 51             | -8             | 113             | 22.3442    |  |
| 10                       | 157             | 152             | 48             | 143            | 134             | 14.2818    |  |
| 11                       | 170             | -8              | 81             | -1             | 153             | 20.1257    |  |
| 12                       | 170             | -16             | 71             | -9             | 154             | 21.4648    |  |

Table 4: Kalman filter-compound range errors for 10 Hz processing rate (20 ft/sec velocity).

| Time = 0.2 sec (Frame 1) |                     |                  |                               |                            |                           |                        |
|--------------------------|---------------------|------------------|-------------------------------|----------------------------|---------------------------|------------------------|
| Measurement              | Raw Binocular Range | Raw Motion Range | Kalman Filter Binocular Error | Kalman Filter Motion Error | Corrected Binocular Range | Corrected Motion Range |
| 1                        | 23.470947           | 12.980159        | 9.617279                      | -1.099180                  | 13.853668                 | 14.079339              |
| 2                        | 15.250125           | 19.499459        | -2.733061                     | 0.434156                   | 17.983187                 | 19.025303              |
| 3                        | 23.286278           | 10.937357        | 5.861856                      | -2.925393                  | 17.424423                 | 13.862750              |
| 4                        | 15.922381           | 14.836158        | -1.956218                     | -4.332531                  | 17.878599                 | 19.368689              |
| 5                        | 17.710770           | 22.263815        | -1.671963                     | 2.354116                   | 19.382732                 | 19.909698              |
| 6                        | 13.850588           | 21.918085        | -3.384527                     | 9.700621                   | 17.235115                 | 12.217464              |
| 7                        | 15.973729           | 21.828383        | -3.465710                     | 6.287485                   | 19.439438                 | 15.541098              |
| 8                        | 16.092087           | 18.314566        | -1.629888                     | 0.328105                   | 17.751974                 | 17.986462              |
| 9                        | 16.151932           | 14.771038        | -3.026224                     | 2.168971                   | 19.178156                 | 12.600267              |
| 10                       | 21.183895           | 22.844698        | 2.680561                      | 1.348993                   | 18.503334                 | 21.495705              |
| 11                       | 15.358275           | 14.761926        | -3.051116                     | 0.234672                   | 18.409391                 | 14.527253              |
| 12                       | 18.167021           | 20.570257        | 0.180036                      | 0.792287                   | 17.986984                 | 19.777969              |
| 13                       | 15.797955           | 21.834696        | -1.613573                     | 0.830073                   | 17.411528                 | 21.004623              |

| Time = 0.3 sec (Frame 2) |           |           |           |           |           |           |
|--------------------------|-----------|-----------|-----------|-----------|-----------|-----------|
| 1                        | 15.395482 | 15.375104 | -2.823521 | -1.199930 | 18.219004 | 16.573033 |
| 2                        | 15.055490 | 18.953318 | -3.251050 | -3.732748 | 18.306541 | 22.680666 |
| 3                        | 15.658957 | 19.364252 | -2.859818 | -3.794225 | 18.318774 | 23.158478 |
| 4                        | 14.771465 | 16.280893 | -3.150485 | -4.380288 | 17.921949 | 20.661182 |
| 5                        | 16.188807 | 18.075811 | -3.643293 | 6.040530  | 19.852100 | 12.035282 |
| 6                        | 16.690779 | 18.732043 | -2.552827 | 4.724038  | 19.253507 | 14.028026 |
| 7                        | 15.174622 | 18.209389 | -3.081228 | 3.214266  | 18.255850 | 14.995123 |
| 8                        | 19.333355 | 24.704124 | -1.267558 | 6.274734  | 20.609912 | 18.429390 |
| 9                        | 14.181705 | 17.482733 | -3.954892 | 1.894178  | 18.136597 | 15.588555 |
| 10                       | 16.496565 | 16.344479 | -3.117268 | 0.850061  | 19.613832 | 15.494417 |
| 11                       | 14.511797 | 16.740181 | -4.223449 | 1.269790  | 17.752247 | 15.470390 |
| 12                       | 15.358172 | 18.337030 | -2.874749 | 2.921894  | 18.232922 | 15.615136 |
| 13                       | 13.100904 | 19.790014 | -4.306006 | 2.210015  | 17.406910 | 17.580000 |

Table 5: Ground truth measurements for 10 Hz processing rate.

| Time = 0.2 sec (Frame 1) |                 |                 |                |                |                 |            |  |
|--------------------------|-----------------|-----------------|----------------|----------------|-----------------|------------|--|
| Measurement              | $y'_1$ (pixels) | $z'_1$ (pixels) | $y_r$ (pixels) | $z_r$ (pixels) | $\eta$ (pixels) | $R_f$ (ft) |  |
| 1                        | -152            | -59             | -182           | -53            | -128            | 13.9764    |  |
| 2                        | -52             | 122             | -131           | 114            | -46             | 18.8474    |  |
| 3                        | -47             | -50             | -94            | -38            | -38             | 12.2229    |  |
| 4                        | -12             | 113             | -83            | 106            | -9              | 19.1300    |  |
| 5                        | 41              | 65              | -58            | 37             | 60              | 19.2826    |  |
| 6                        | 49              | 202             | -39            | 189            | 44              | 20.2861    |  |
| 7                        | 92              | -61             | 0              | -49            | 82              | 18.0941    |  |
| 8                        | 95              | 167             | 0              | 149            | 82              | 14.0358    |  |
| 9                        | 124             | -15             | 51             | -8             | 113             | 22.3442    |  |
| 10                       | 157             | 152             | 48             | 143            | 134             | 14.2818    |  |
| 11                       | 170             | -8              | 81             | -1             | 153             | 20.1257    |  |
| 12                       | 170             | -16             | 71             | -9             | 154             | 21.4648    |  |

| Time = 0.3 sec (Frame 2) |      |     |      |     |      |     |         |
|--------------------------|------|-----|------|-----|------|-----|---------|
| 1                        | -158 | -26 | -218 | -22 | -137 | -21 | 15.9601 |
| 2                        | -85  | -7  | -161 | 3   | -76  | -3  | 20.0576 |
| 3                        | -34  | -71 | -115 | -53 | -32  | -59 | 22.9181 |
| 4                        | -33  | -95 | -118 | -86 | -30  | -81 | 19.5996 |
| 5                        | 79   | 205 | -11  | 183 | 71   | 180 | 16.4573 |
| 6                        | 85   | -19 | -3   | -8  | 76   | -13 | 17.7225 |
| 7                        | 103  | -73 | 5    | -56 | 92   | -61 | 17.7835 |
| 8                        | 125  | -43 | 47   | -30 | 115  | -54 | 23.7157 |
| 9                        | 143  | 56  | 33   | 60  | 126  | 54  | 16.6384 |
| 10                       | 145  | 138 | 46   | 125 | 126  | 126 | 15.4135 |
| 11                       | 172  | 168 | 59   | 158 | 150  | 152 | 15.9483 |
| 12                       | 204  | -50 | 96   | -35 | 181  | -40 | 17.9549 |
| 13                       | 207  | 50  | 85   | 52  | 185  | 49  | 19.1903 |