N94- 25505

Multiple–Camera/Motion Stereoscopy for Range Estimation in Helicopter Flight

Phillip N. Smith, Banavar Sridhar, and Raymond E. Suorsa
NASA Ames Research Center

Abstract

Pilot aiding to improve safety and reduce pilot workload to detect obstacles and plan obstacle-free flight paths during low-altitude helicopter flight is desirable. Computer vision techniques provide an attractive method of obstacle detection and range estimation for objects within a large field of view ahead of the helicopter. Previous research has met considerable success by using an image sequence from a single moving camera in solving this problem. The major limitations of single camera approaches are that no range information can be obtained near the instantaneous direction of motion or in the absence of motion. These limitations can be overcome through the use of multiple cameras. This paper presents a hybrid motion/stereo algorithm which allows range refinement through recursive range estimation while avoiding loss of range information in the direction of travel. A feature-based approach is used to track objects between image frames. An extended Kalman filter combines knowledge of the camera motion and measurements of a feature's image location to recursively estimate the feature's range and to predict its location in future images. Performance of the algorithm will be illustrated using an image sequence, motion information, and independent range measurements from a low-altitude helicopter flight experiment.

1 Introduction

To increase safety and improve mission effectiveness during low-altitude helicopter flight, NASA Ames Research Center in conjunction with the U.S. Army has been developing automation tools to assist pilots in detecting obstacles and planning obstacle-free flight paths. The most challenging mode of low-altitude flight is Nap-of-the-Earth (NOE) flight, characterized by lateral maneuvers below tree-top level in order to conceal the helicopter behind available terrain or man-made objects. An on-line sensor to gather obstacle information is required for pilot-aiding during NOE flight because existing *a priori* terrain data such as digital maps (1) suffer from inaccuracies larger than the vehicle's altitude, (2) have insufficient resolution to show obstacles such as trees and buildings, and (3) cannot easily account for changes in the terrain such as the growth of new trees or the construction of new buildings. Vision sensors are desirable for obtaining the online obstacle information due to their passive nature and relatively large field of view.

The classification of obstacles is unnecessary for accomplishing the obstacle avoidance task because it is sufficient to avoid *all* obstacles regardless of identity. It is therefore required only that the vision system provide position information for each object in the field of view. In practice the vision system attempts to compute a range map depicting the distance to the terrain for each point in the field of view.

A common approach to this problem makes use of an image sequence collected from a single moving camera and in some cases the camera's motion information. Small regions of interest (called features) are identified in an image, the feature's location is tracked in successive images, and a recursive filter is used to estimate range and/or camera motion [1, 2, 3]. The authors have previously developed an algorithm of this class and evaluated its performance with helicopter flight data as described in [4, 5, 6]. A major limitation of this approach is that range information cannot be obtained along the instantaneous direction of motion and, in practice, reliable range information cannot be obtained even for objects lying near the direction of motion. This limitation can be overcome through the use of multiple cameras mounted so their baseline is roughly normal to the motion direction [7, 8]. A hybrid motion/stereo algorithm is presented in this paper which allows range refinement through recursive range estimation while avoiding loss of range information in the direction of travel.

The extended Kalman filter provides a convenient structure for the implementation of motion/stereo range estimation. The Kalman filter allows for range refinement through recursive estimation. Furthermore, the range prediction generated during the time update serves to constrain the search area required to locate the feature in future images.

A low-altitude helicopter flight experiment has been conducted to obtain realistic data for evaluating the motion/stereo algorithm. The flight experiment provides video imagery from two monochrome video cameras, helicopter motion data, and camera calibration information. True range measurements have been obtained using a laser tracker to allow evaluation of the algorithm's performance.

The purpose of this paper is to describe a Kalman filter based motion/stereo ranging algorithm and to present preliminary results obtained using data from a helicopter flight experiment. Section 2 will discuss the Kalman filter implementation of the motion/stereo ranging algorithm. Section 3 will describe the helicopter flight experiment and calibration of the camera system. In Section 4, preliminary results obtained using the experimental data and the motion/stereo algorithm will be presented. Finally, Section 5 will complete the paper with a brief discussion and concluding remarks.

2 Kalman Filter

The proposed algorithm uses a feature based method in which the image is treated as a collection of tokens or features and information (such as range) is computed only for the individual features rather than for every point in the image. Currently, features are defined to be 11×11 square pixel image patches which exhibit a sufficiently high intensity variance. A feature's location in another image is determined by correlation of the feature's intensity surface with the intensity surface of the other image. The correlation surface is then interpolated in the region near its peak, and the location of the resulting peak is taken to be the feature's location to subpixel accuracy. Features can be born with each new image, and old features die when they fail to be tracked between images. Further discussion of feature detection and tracking can be found in Ref. [4, 9].

In our implementation, a Kalman filter is associated with each feature for determining the location of the object which gives rise to the feature. The motion/stereo Kalman filter is an extension of the monocular range estimation Kalman filter derived in an earlier work [10]. Both filters rely on the assumptions that all objects of interest are stationary in an Earth-fixed frame, and that measurements of the camera's linear and angular velocities are available (from an inertial navigation system, for example). The resulting state equation is an expression of the Coriolis equation:

$$\dot{X} = -[\omega_c]X - V_c \tag{1}$$

where

$$[\omega_c] = \begin{bmatrix} 0 & -\omega_{z_c} & \omega_{y_c} \\ \omega_{z_c} & 0 & -\omega_{x_c} \\ -\omega_{y_c} & \omega_{x_c} & 0 \end{bmatrix}$$

$X = [x_c, y_c, z_c]^T$ is the object position relative to the camera, $\omega_c = [\omega_{x_c}, \omega_{y_c}, \omega_{z_c}]^T$ is the camera's angular velocity, and V_c is the linear velocity. The measurement equation accounting for perspective projection of the object onto the image plane is given below

$$Z = h(X) = [fy_c/x_c, fz_c/x_c]^T$$

where $Z = [u, v]^T$ is the location of the object on the image plane and f is the camera's focal length. Here the camera axes have been defined with the x_c axis passing through the focal point and perpendicular to the sensor array, and y_c and z_c in the direction of the rows and columns of the sensor array, respectively. The extended Kalman filter is formed by linearizing $h(X)$ about the current state estimate yielding

$$\begin{aligned} Z &= HX \\ H &= \partial h(X)/\partial X \\ &= f\begin{bmatrix} 1/x_c & 0 & -y_c/x_c^2 \\ 0 & 1/x_c & -z_c/x_c^2 \end{bmatrix} \end{aligned}$$

To extend the Kalman filter characterization for two cameras we need additional measurement equations relating $Z' = [u', v']^T$, the image location of the same object in the second camera. Let X' be the object position relative to the second camera. The relationship between the cameras is of the form

$$X' = RX + T$$

where R is a 3×3 matrix and T is a vector representing the relative rotation and translation, respectively, between the two cameras' coordinate systems and centers of reference. Then the measurement Z' can be written as follows

$$Z' = h(X') = [fy'_s/x'_s, fz'_s/x'_s]^T$$

As above, we can derive a linearized measurement equation of the following form

$$\begin{aligned} Z' &= H'X \\ H' &= \partial h(X')/\partial X \end{aligned}$$

The Kalman filter can be computed for the system using the state equation (1) and the composite linearized measurement

$$Z_c = \begin{bmatrix} H \\ H' \end{bmatrix} X$$

Thus, the Kalman filter measurement update may be performed based on the obstacle location in any imaging sensor provided the location and orientation of the sensor are known relative to the reference sensor system. The stereo system has four measurements and the same state equations as the monocular system. Based upon the given state and measurement equations, the full discrete time extended Kalman filter equations can be derived in the standard manner. This method can be extended in the same way to any number of cameras.

The range estimation process begins when a feature is identified in the image from one camera. A stereo match is determined by searching an area in the image from the second camera which is constrained by *a priori* values of the minimum and maximum range of interest. The resulting stereo range estimate is used to initialize the Kalman filter. The initial value of the Kalman filter's state covariance matrix may also be estimated or chosen *a priori*. The range estimate is then propagated forward in time by the Kalman filter, and the predicted state vector and state covariance matrix give rise to a search area to be traversed in locating the feature in the next image [9]. The Kalman filter uses the matched feature locations to perform its measurement update. As the Kalman filter converges, the value of the state covariance matrix decreases leading to smaller search areas and reduced computational effort. Given images from the two cameras over time, a variety of tracking schemes are possible. The currently implemented approach is to match each feature (1) from the left camera at the current time to the left camera at the next time and (2) from the left camera at the current time to the right camera at the next time. The above procedure is repeated for each feature until such time as the feature fails to be matched.

3 Flight Experiment

The helicopter flight experiment conducted to provide raw data and independent truth measurements for development and validation of passive ranging algorithms is illustrated in Figure 1. The resulting data set includes video imagery from two monochrome video cameras, helicopter motion data from an onboard inertial navigation system (INS), true range measurements obtained with a

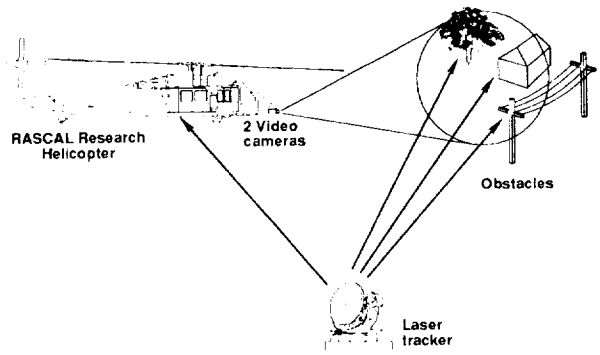

Figure 1: Flight Experiment Overview

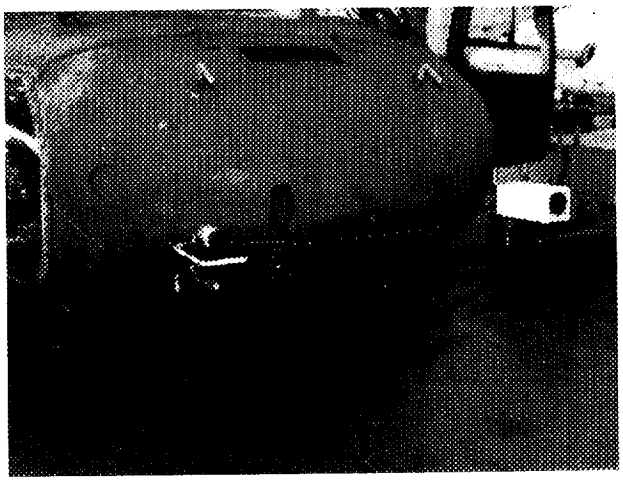

Figure 2: Camera Installation

laser tracker, and experimentally determined camera calibration parameters which characterize the geometry and imaging properties of the camera system.

The test apparatus consists of two Cohu 6410 monochrome interlaced video cameras mounted 1 meter apart on a horizontal bar attached to the nose of a UH-60 Blackhawk helicopter as shown in Figure 2. The cameras have a focal length of 6 mm, a field of view of 58 × 45 degrees and they are electronically shuttered with a 1/1000 sec exposure time to reduce image smear due to camera motion. The video imagery from each camera is time-tagged using a Datum 9550 video time inserter unit and recorded using a Sony VO-9600 U-matic SP video recorder onboard the helicopter. The images are acquired at the rate of 30 frames/sec per camera. The helicopter's motion state is measured by a Litton LN93 inertial navigation system (INS) and also recorded onboard the helicopter. A laser tracker measures the helicopter's position during flight and also measures the location of the (stationary) obstacles of interest. Synchronization of the various data sources is accomplished by recording a master time index along with each element of the data set.

Post-flight processing consists of digitizing the recorded video data into 512 × 512 pixel images with 256 levels of gray. In addition, INS-derived motion data and laser-tracker-derived position data are processed together using a forward-backward filtering technique [11] to ensure kinematic consistency and to identify and correct for any sensor bias or scale factor errors. The resulting uncertainty in the motion data is approximately ± 2 ft in position, ± 0.01 deg in orientation, ± 0.25 ft/sec in velocity, and ± 0.3 deg/sec in angular velocity. Filtered motion data is desirable for development of the ranging algorithm, but in an operational system the motion state would be acquired directly from the INS.

The camera calibration parameters which characterize the camera system consist of two sets: the external parameters which include the geometrical description of the camera system, and the internal parameters which describe the imaging properties of the cameras. The external parameters allow the motion state measurements to be transformed from the helicopter body axes (as defined by the INS) to the sensor axis system (as defined by the cameras) for input to the Kalman filter. Similarly, using the external parameters, range estimates can be transformed back from sensor axes to body axes where they are more useful to the pilots or to an obstacle-avoidance guidance system. In addition, the external parameters provide the cameras' relative orientation, which is required for the stereo component of the ranging algorithm. The internal parameters define the mapping from points in the sensor axis system to pixel row and column coordinates in a digitized image. Internal parameters include the focal length, the pixel location where the x_s axis passes through the image plane, the effective dimensions of the pixels including any stretching effects caused by the recording and digitization process, and any distortion effects. There are a total of six external parameters and 5 internal parameters (assuming no distortion) for each camera. We have not yet found it necessary to model distortion terms with the ranging algorithms we have tested.

A separate experiment has been performed to determine the calibration parameters. Camera calibration has not received much attention in the literature but plays a central role in the performance of operational vision system. Some treatment of calibration techniques can be found in [12, 13, 14]. The approach taken here has been to (1) place a grid of target points within the cameras' field of view, (2) measure the locations of target points relative to the helicopter body axes, (3) determine the pixel locations of the target points in a digitized image taken with the camera, and (4) estimate the camera calibration parameters relating the two sets of measurements by solving a nonlinear cost minimization problem.

The calibration procedure uses a grid of horizontal and vertical lines, the 99 intersections of which serve as the calibration targets. A surveyor's transit is used to determine the target locations in the helicopter body axis system with an accuracy of approximately ± 3 mm. Five target points are measured directly, from which the remaining target locations can be interpolated. The entire grid assembly is stationed at four different distances in front of the cameras ranging between eight and 22 feet.

From a digitized image, the target pixel locations are found with subpixel accuracy by computing the intersections of curves fit to each of the grid lines. First, the intensity distribution perpendicular to one of the grid lines at some station is examined. The intensity peak, which is determined by locally fitting the intensity distribution with a parabola, defines one point on the grid line. The

process is repeated for several stations along each grid line, and the resulting points are fit with a curve (a line or a higher-order polynomial depending on the significance of image distortion). The curves' intersections are determined mathematically to give the target locations to subpixel accuracy.

In the final step, the parameters are estimated by minimizing a cost function which is a sum of squared errors terms. Two general approaches were taken: estimating the parameters for each camera separately and estimating the parameters for both cameras simultaneously. In the first case the cost function is the sum of errors in distance between the measured target pixel locations and the estimated pixel locations based on the measured body-axis locations and postulated parameter values. In a variation of this cost function, penalty terms were included for violation of Tsai's radial alignment constraint [12]. This calibration procedure resulted in RMS errors of approximately 0.4 pixel. However, using the resulting calibration parameters with the measured target pixel locations to estimate the corresponding body axis locations using stereo leads to large errors. By estimating the calibration parameters for both cameras simultaneously the stereo ranging errors can be reduced through augmentation of the cost function. Several variations of the cost function were implemented, but little difference was observed in the result so long as terms were included for errors in the location of target points in the image plane and in the body axes. Weighting an error of 0.5 pixel in the image plane equivalently with a 0.25 inch error in the body axes leads to an RMS error of approximately 0.5 pixel and 0.5 inch, respectively.

4 Results

The image sequence used in generating the results given in this section was taken with the helicopter following a nominally straight flight path at a velocity of about 25 knots (42 ft/sec) 20 feet above a runway. Six trucks were positioned along the runway to serve as obstacles, initially ranging between 500 and 1400 feet from the helicopter. Figure 3 shows the first and last images in a sequence of 180 frames taken with the left camera. It is noted that in spite of the nominally straight line flight path, the FOE (depicted by crosshairs in Figure 3) travels 30 pixels in both the horizontal and vertical directions throughout the image sequence.

The image sequence is processed with the motion/stereo algorithm of Section 2 giving the range estimates to approximately 300 features in each image. To evaluate the algorithms performance, the average of the range estimates for all features belonging to each truck is computed. These preliminary results for the five closest trucks are given in Table 1 along with the true range at frame numbers 1, 60, 120, and 180. For reference, the corresponding results obtained with the earlier monocular ranging algorithm are also shown in Table 1. The preliminary results show that the initial range estimates are significantly better using the stereo method as expected since the trucks are both far away and close to the FOE. Over time, the additional measurements lead to improved range estimates and the results of both

Table 1: Preliminary Range Results

Truck	Frame	Range, ft		
		Truth	Monocular	Motion/Stereo
A	1	488	171	489
	60	399	405	431
	120	316	335	350
	180	235	227	247
B	1	614	270	785
	60	525	568	587
	120	443	462	463
	180	363	364	341
C	1	741	267	739
	60	650	519	498
	120	568	606	565
	180	487	514	486
D	1	860	138	N/A
	60	770	618	594
	120	688	653	799
	180	609	534	671
E	1	991	122	955
	60	899	995	813
	120	817	594	698
	180	736	863	722

methods converge toward the true range. Note that the motion/stereo case sometimes produces less accurate results, potentially due to the following characteristics of the currently-implemented algorithm. Range estimates are not always available using the stereo motion method. In fact there are only half as many features resulting from the motion/stereo method as from the monocular method, indicating fewer (though hopefully stronger) feature matches. Sometimes even apparently strong features may fail to match in both cameras which on further examination is attributed to small-scale differences between the images from the two cameras due to image noise and the differences in the cameras themselves. A modification of the tracking scheme to match only between images taken with the same camera or between images taken at the same time may lead to better matching. Even if matching cannot be improved, the motion/stereo results could be enhanced by allowing range estimates to be propagated based on monocular motion only rather than killing the feature in the event that a stereo match cannot be made. In this way, the motion/stereo algorithm gracefully degrades to the monocular algorithm when stereo matches cannot be obtained, but stereo information is utilized when it is available.

5 Concluding Remarks

A hybrid motion/stereo range estimation algorithm has been described which combines the strengths of stereo methods (i.e., ranging without motion and ranging to objects near the FOE) and monocular methods (i.e., recursive range refinement). This motion/stereo algorithm has been implemented as a Kalman filter. A helicopter flight experiment was conducted to collect data for validation of the algorithm. Preliminary results indicate that initial motion/stereo range estimates are an improvement over initial monocular estimates and that both methods give range results which generally approach the true range over time. It was noted that some improvement in the robustness of the motion/stereo algorithm could be obtained by

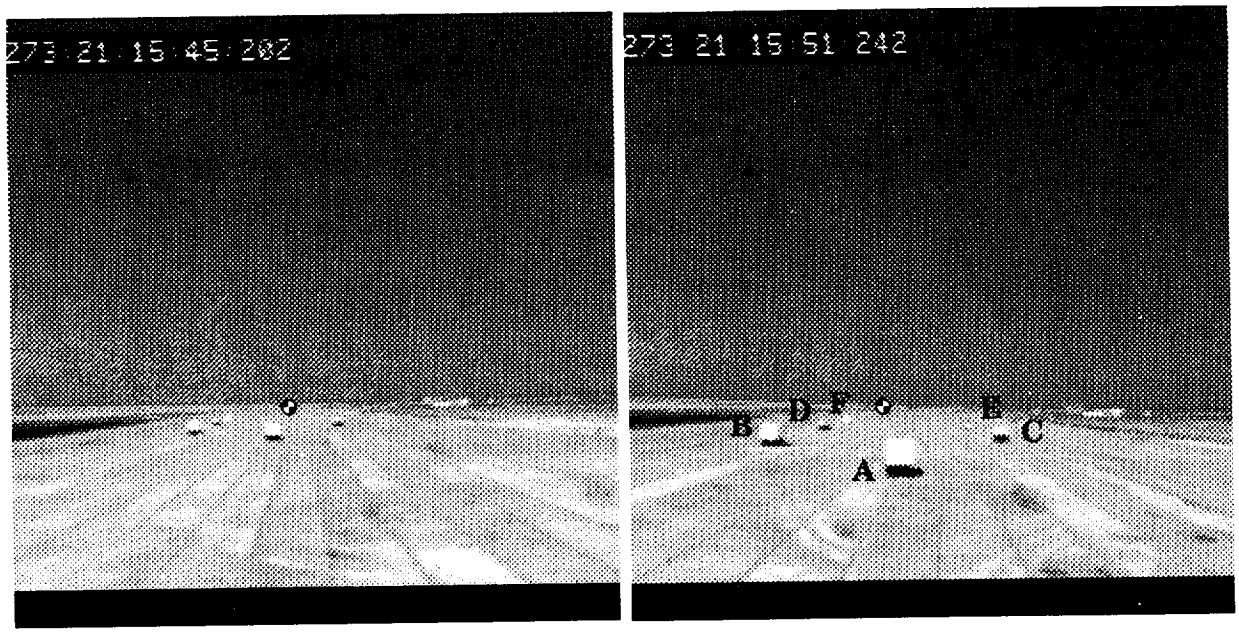

Figure 3: First and Last Images of Helicopter Sequence

allowing it to degrade to the monocular algorithm for a given feature when a stereo match cannot be established. In the future we plan to continue refinement of the motion/stereo algorithm and to test it with flight sequences having curvilinear motion and images of natural terrain.

References

[1] Chellappa, R., and Shekar, C., "Passive Ranging Using a Moving Camera," *Journal of Robotic Systems*, vol. 9, 1992.

[2] Bhanu, B., Roberts, B., and Ming, J., "Inertial Navigation Sensor Integrated Motion Analysis," *Proceedings of the DARPA Image Understanding Workshop*, Palo Alto, CA, May 1989.

[3] Kumar, R., and Hansen, A.R., "Model Extension and Refinement Using Pose Recovery Techniques," *Journal of Robotic Systems*, vol. 9, 1992.

[4] Sridhar, B., Suorsa, R., and Hussien, B., "Passive Range Estimation for Rotorcraft Low-altitude Flight," NASA TM, October 1990; to appear in *Journal of Machine Vision and Applications*, March 1993.

[5] Smith, P.N., "Flight Data Acquisition for Validation of Passive Ranging Algorithms for Obstacle Avoidance," *Journal of the American Helicopter Society*, vol. 37, No. 4, pp. 32-37, October 1992.

[6] Smith, P.N., Sridhar, B., and Hussien, B., "Vision-Based Range Estimation Using Helicopter Flight Data," *Proceedings of The 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp 202-208, Champaign, IL, June 1992.

[7] Sridhar, B. and Suorsa, R., "Integration of Motion and Stereo Sensors in Passive Ranging Systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 27, No. 4, pp. 741-746, July 1991.

[8] Kriegman, D.J., Triendl, E., and Binford, T.O., "Stereo Vision and Navigation in Buildings for Mobile Robots," *IEEE Transactions on Robotics and Automation*, vol. 5, pp. 792-803, 1989.

[9] Suorsa, R.E., and Sridhar, B., "A Parallel Implementation of a Multisensor Feature-Based Range-Estimation Method," NASA TM 103999, January 1993.

[10] Sridhar, and Phatak, A., "Simulation and Anaysis of Image-Based Navigation system for Rotorcraft Low-Altitude Flight," *American Helicopter Society Specialists' Meeting on Automation Applications of Rotorcraft*, Atlanta, GA, April 1988; *IEEE Transactions on Systems, Man and Cybernetics*, vol. 22, no. 2, pp. 96-101, March/April 1992.

[11] Bach, R.E., Jr., "State Estimation Applications in Aircraft Flight-Data Analysis (A User's Manual for SMACK)," NASA RP 1252, March 1991.

[12] Tsai, R.Y., "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses," *IEEE Journal of Robotics and Automation*, vol. RA-3, no. 4, Aug. 1987, pp. 323-344.

[13] Slama, C.C. (editor), *Manual of Photogrammetry, Fourth Edition*, American Society of Photogrammetry and Remote Sensing, Falls Church, Virginia, 1980.

[14] Beyer, H.A., "Accurate Calibration of CCD-Cameras," *Proceedings of The 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp 202-208, Champaign, IL, June 1992.