ELSEVIER

Robotics and Autonomous Systems 28 (1999) 185–193

Robotics and Autonomous Systems

www.elsevier.com/locate/robot

A visual odometer for autonomous helicopter flight

Omead Amidi *, Takeo Kanade, Keisuke Fujita

Carnegie Mellon Robotics Institute, Pittsburgh PA, USA

Received 18 September 1998

Abstract

This paper presents a visual odometer for autonomous helicopter flight The odometer estimates helicopter position by visually locking on to and tracking ground objects. The paper describes the philosophy behind the odometer as well as its tracking algorithm and implementation. The paper concludes by presenting test flight data of the odometer's performance on-board indoor and outdoor prototype autonomous helicopters

Keywords · Helicopter; Vision-based; Odometer; Navigation; Aerial robotics

1. Introduction

The control of an autonomous helicopter is only as good as its positioning. Control response and accuracy is, in essence, dictated by how frequently and how promptly the helicopter's position is determined during flight. Helicopter positioning can be either global or relative depending on the task at hand Global positioning is necessary for long distance flight where the helicopter must reach a predetermined destination. Relative positioning, on the other hand, is necessary for precise flight in relation to objects of interest in the environment. Vision is particularly well-suited for this type of relative positioning. Vision-based positioning frees the system from relying on external positioning devices such as satellites and beacons. Vision can also serve as a resetting mechanism or a substitute for drift-prone inertial sensors.

One way vision can estimate helicopter motion is by tracking stationary objects in the surrounding environment. Objects are displaced in consecutive images as the helicopter moves, and this displacement can be accurately measured by image processing to detect motion. A key concern is the trackability of objects in the field of view Tracking is possible only if visible objects possess distinguishing features which can be consistently identified in image sequences. Such highly contrasting and randomly textured scenery is common in outdoor environments and can provide feature-rich imagery for vision-based motion sensing. On-board vision can take advantage of the abundant natural features to "lock" on to arbitrary objects and track them to sense motion

It is difficult, however, to sense helicopter translation, which is essential for autonomous control, with vision alone since image displacements also occur with helicopter rotation. Distinguishing between rotation and translation in a sequence of images under perspective projection is extremely difficult. For instance, helicopter rolling motion can appear very similar to lateral translation in consecutive images This ambiguity can be resolved by accurately fusing data from angular sensors with image displacement measures

* Corresponding author.
E-mail address: amidi@ri cmu.edu (O Amidi)

PII: S0921-8890(99)00016-0

Light-weight and inexpensive gyroscopes and angular rate sensors available today can reliably measure the angular variations between images to disambiguate the source of image displacement.

This paper presents a visual odometer developed for vision-based helicopter control. The odometer is designed to operate without the aid of high-quality inertial sensors so that it can serve as a redundant position sensor in a guidance system. The odometer maintains position by capturing images from a pair of ground-pointing cameras and visually locking on to and tracking objects in the field of view. The odometer is implemented by a custom-designed vision machine [1] on-board a number of prototype autonomous helicopters [2].

2. Positioning with a visual odometer

The odometer determines the position of objects appearing in the camera field of view relative to the initial helicopter location, and thereafter tracks these objects visually to maintain position. As the helicopter moves, older objects leave the field of view and new objects entering the scene are localized to continue tracking motion.

The visual odometer relies on a "target" template initially taken from the center of an on-board camera image. The location of the objects appearing in the target template is determined by sensing camera range to the object with respect to the current helicopter position and attitude. With the sensed object location, helicopter position is updated as the odometer tracks the object in incoming images with template matching.

Template matching between consecutive images measures lateral and longitudinal image displacement which may result from both helicopter translation and rotation. Template matching in two images, taken simultaneously by a stereo pair of cameras, measures helicopter range. Three-dimensional helicopter motion is then estimated by combining the lateral and longitudinal image displacements and range estimates with helicopter attitude, measured by on-board angular sensors. Several important observations are in order regarding this position tracking approach.

2.1. *Effects of rotation*

Helicopter translation is a direct result of its change in attitude, often causing large image displacement. The effects of rotation must be eliminated from the measured image displacement to determine the change in helicopter position. The visual odometer determines these effects by precisely measuring the variation in helicopter attitude between images. This correction is only valid provided that attitude data is captured in precise synchronization with the camera shutter opening.

2.2. *Template matching accuracy*

Helicopters can move rapidly relative to tracked objects. As a result, the template matching process must be consistent and robust to accommodate the quick rotation and distance variations. Fig. 1

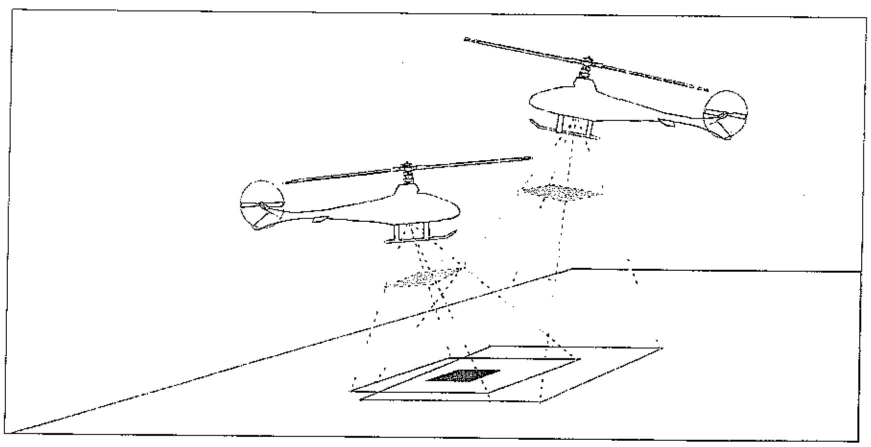

Fig. 1. Template appearance variations.

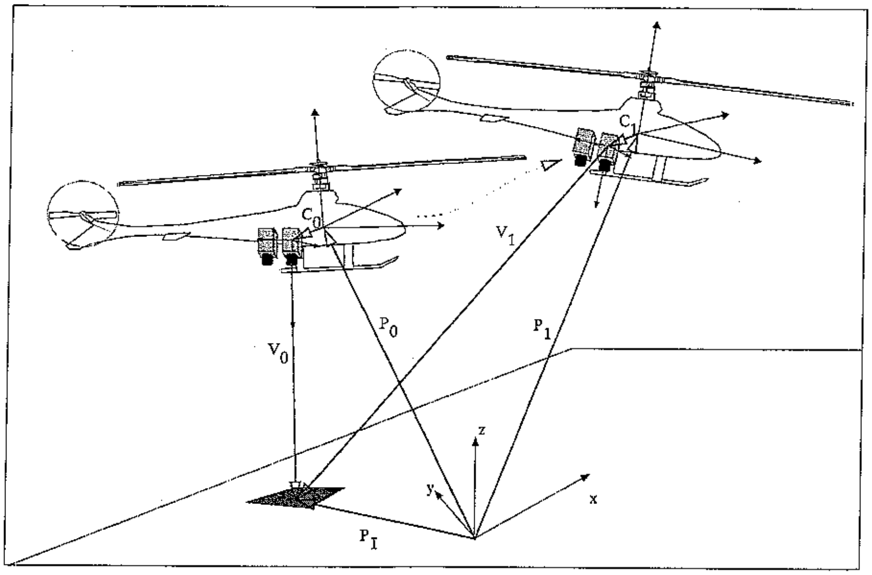

Fig. 2. Visual odometer position tracking.

demonstrates these variations as the helicopter yaws and changes altitude. Accurate template matching to retain visual "lock" on objects requires anticipating how the appearance of the objects changes in future images before performing the matching operation. For accurate and consistent matches, templates must be rotated, scaled, and normalized in intensity. The visual odometer determines incremental template rotation and scaling factors by tracking multiple templates concurrently. By tracking a pair of templates, the odometer estimates the effects of rotation and height variations directly from the image. Observing the direction and magnitude of a vector connecting the templates determines the rotation angle and the scale factor necessary to prepare templates for subsequent accurate matches.

2.3. Template matching speed

The computationally complex matching operations must be performed frequently to ensure accurate template matching and to provide sufficient rate of feedback for helicopter stabilization. Searching for a matching position in the entire image is not always necessary. To reduce computational requirements, the odometer searches for templates in a small window surrounding the previous match. The search window size is chosen based on the matching frequency, helicopter proximity to the objects appearing in the template, and anticipated helicopter movement based on the previous match displacement. It is worth noting that the search area can be reduced with increasing processing frequency and that high processing frequency may be achievable only if the search area is smaller. Therefore, it is beneficial to carry out the matching operation as fast as possible limited only by the camera image acquisition frequency.

3. Position sensing algorithm

To analyze the odometer's position sensing algorithm, let us examine a situation where the helicopter moved from ground navigation frame position P_0 to P_1 as depicted in Fig. 2. Based on the current estimated position, P_0, the algorithm estimates the new position, P_1, by estimating the target's position, P_I, view vectors V_0 and V_1, and camera translation vectors C_0 and C_1 in the ground navigation frame.

The algorithm bootstraps by locating the ground target visible at the image center. This task requires sensing the target location in the camera frame, which defines the view vector, and the camera translation vector to the ground frame based on measured helicopter attitude. The view vector is simply camera range to the target which is sensed by stereo image processing.

With a localized template to lock on to, the algorithm is ready to track the helicopter's position. The algorithm measures the new view vector V_1 by image processing and transforms this vector along with the camera translation vector to the navigation frame based on measured helicopter attitude. These two vectors and the target location are all the algorithm needs to determine the new helicopter position, P_1. Concurrent to this process, the algorithm continuously selects and locates a new potential target in case the current target is about to leave the field of view.

4. Image processing

The odometer's positioning quality is directly affected by the accuracy and robustness of image template tracking. Templates must be identified consistently as they vary in size and rotation as the helicopter gains or loses altitude and banks or yaws. The odometer must also handle changes in image intensity as lighting conditions change. The odometer tracks a pair of templates, referred to the main and auxiliary templates, to measure image rotation and scaling as well as normalizing target templates before matching. Changes in angle and size of the baseline between the templates measures the required image rotation and scaling. Fig. 3 depicts the odometer's image processing flow chart. The odometer chooses an auxiliary template near the main template at the image center. The auxiliary template is offset by a nominal distance from the main template. This x direction offset is the initial horizontal baseline representing zero image rotation.

The odometer locates and stores the two templates from the initial image and commences matching them in incoming images. The templates are calibrated using the baseline and image intensities of the previous match. Image rotation, scaling, and intensity variation in one cycle is assumed to be insignificant relative to the algorithm tracking frequency. To accommodate changes in template intensity, the template pixel intensities are normalized to correspond to the most recent image match. In addition, a scale factor is determined by comparing the intensity within the calibrated templates and the intensity within the matched areas.

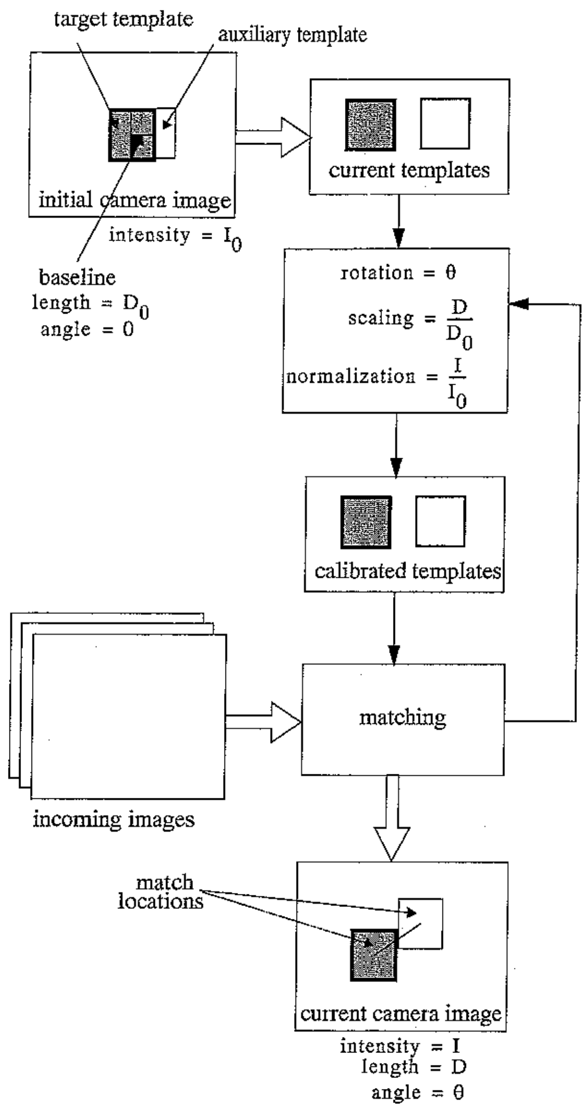

Fig 3 Odometer processing flow chart

Finally, the stored templates are rotated by the baseline angle before locating the templates in the next image.

From time to time, the odometer must reacquire new templates. The odometer captures new templates in every cycle while it is locked on to current templates in preparation of this event. The odometer switches to a new set of templates under the following conditions:

- *Either one of the current templates is about to go out of view.* Since the templates are searched for in the neighboring area of the last successful match, they must be replaced if the neighboring search area reaches an image border.

- *The current template baseline length, D, is above a threshold value.* A large baseline indicates a significant reduction in helicopter altitude or large attitude changes which degrade the resolution of calibrated templates, and leads to poor matches. Large template separations also signal a potential mismatch of one of the templates; switching both templates improves the matching for upcoming images.
- *The current baseline length, D, is below a threshold value.* A short baseline is a result of a significant gain in altitude or a large change in helicopter attitude. The angular resolution of image rotation angle reduces with baseline length and new templates are necessary to restore this resolution.

The odometer measures object range by stereo image processing. It interpolates one range estimate at the image center to detect the range arbitrary image templates. This interpolation assumes the area in the camera field of view is locally flat.

5. Template matching

The odometer employs the sum of the squared differences (SSD) matching criteria to locate templates in incoming camera images. The SSD criteria, one of a large class of image comparison strategies [3], is the traditional choice because of its proven effectiveness in many object tracking applications. In each image, the odometer searches for the location yielding the minimum SSD of image and template pixels to locate a matching area. To reduce the matching computational cost, the odometer restricts the search area to a small window around the previous successful template match. This search area is experimentally determined so that it can accommodate the maximum change in template location within one processing period. As the helicopter altitude decreases, the same translational motion causes a larger displacement in the image and the entire image may need to be searched to locate the target template.

The odometer employs a coarse-to-fine strategy to further improve the matching processing complexity. Initially, the template and image are subsampled by calculating the SSD of every fourth pixel to narrow the search to a 9×9 pixel area. The subsampled match is then improved by computing the SSD at the unexamined pixels within the subsampling neighborhood.

Image subsampling can be susceptible to mismatches, especially in images with highly contrasting intensities. Typically, search algorithms employ a multi-resolution image pyramid constructed by interpolating adjacent pixels. This interpolation has been shown to improve match consistency and reduce computational cost. However, for the helicopter application, images must be filtered due to the significant inherent noise of the power plant and on-board electronics which lowers image contrast and eliminates the need for such pixel interpolation. In fact, by smoothing images with an 8×8 Gaussian convolution mask, the odometer produced consistent matches of high contrasting natural vegetation by subsampling alone. There was no need for pixel interpolation.

The odometer improves the template match location to subpixel accuracy by fitting a two-dimensional parabolic surface to the SSD error of the pixel match candidates using the following equation:

$$SSD(x, y) = ax^2 + by^2 + cxy + dx + ey + f, \quad (1)$$

where (x, y) are the match candidate locations and (a, b, c, d, e, f) represent the parabola coefficients. The odometer employs a least squares parabolic fit to determine the subpixel match location. The least square parabola coefficients are determined by:

$$[a \quad b \quad c \quad d \quad e \quad f]^{T} = (A^{T}A)^{-1}A^{T}e, \quad (2)$$

where A is an $n \times 6$ matrix with rows representing one of n match candidates and e is a vector of SSD errors for the corresponding pixel. Each row of A consists of the six parabola variables evaluated at the particular integer pixel coordinates. The matrix A can be stored as a constant to reduce computational cost.

In addition to subpixel accuracy, the fitted parabola provides match uncertainty information. A steep parabola versus a shallower one signals a more accurate match. For instance, the fitted parabola of Fig. 4 is shallow, indicating a poor match. The parabola coefficients can be used to construct a covariance matrix describing template match uncertainty in two dimensions. This uncertainty measure is essential for data fusion from multiple template matches or external sensors to improve image displacement estimation.

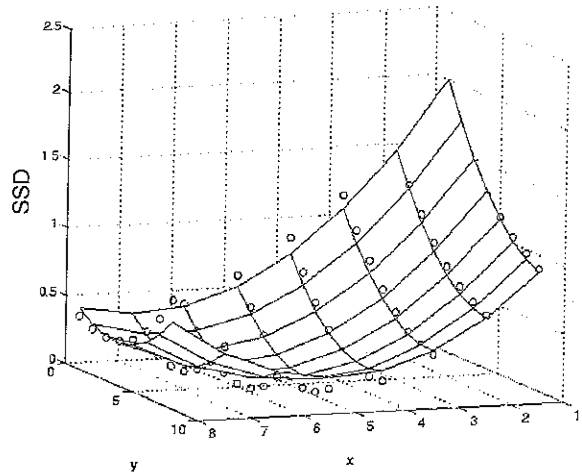

Fig. 4. Parabola fit.

6. Implementation

The visual odometer poses a number of difficult demands on an on-board vision system. The system must process images in real time with minimum latency. It must possess high computational power and data throughput, and must allow close integration of external sensing with vision. Furthermore, the system must be compact enough in size, and efficient in power usage to fit on-board a small helicopter.

The lack of commercially available vision systems capable of meeting these requirements motivated the development of a new reconfigurable architecture to implement a visual odometer machine. The machine is composed of a number of modules including: image A/D and D/A converters, image convolvers, powerful digital signal processing (DSP) elements, an image tagging and synchronization module, and external communication bridge modules. Fig. 5 shows how these modules are interconnected to realize a prototype visual odometer machine.

The visual odometer machine captures images at field rate (60 Hz) from two ground-pointing b/w video cameras. An 8×8 convolver (GEC Plessy) smooths the images before they are transferred to two DSP (TI C44) through dedicated data communication links. The DSPs track a pair of 32×32 pixel templates with a 16 pixel neighboring search area. The initial template baseline is 20 pixels. A sensor interface or bridge, triggered by the central synchronization generator, records and transfers helicopter attitude data from gyroscopes (gyration vertical and directional gyro-engines) to another DSP. Image processing results and filtered and synchronized attitude data are transformed to sensed helicopter motion by a fourth DSP before they are sent to a real-time controller (Motorola 68040) which controls the helicopter's actuators. The visual odometer positioning frequency is 60 Hz with 20 ms latency.

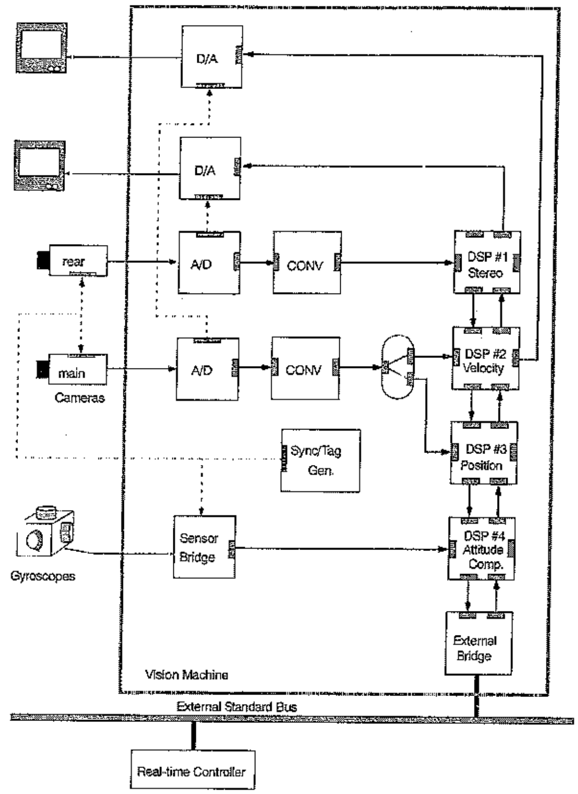

Fig. 5. Visual odometer machine.

7. Position estimation experiments

A six-degrees-of-freedom (6-DOF) testbed was developed for evaluating various helicopter position estimation and control systems. As shown in Fig. 6, the testbed supports an electrical model helicopter, Kalt Whisper, attached to poles by graphite rods through frictionless air bearings. The rods provide safety and measure ground-truth helicopter position. The testbed allows unobtrusive helicopter free flight in a

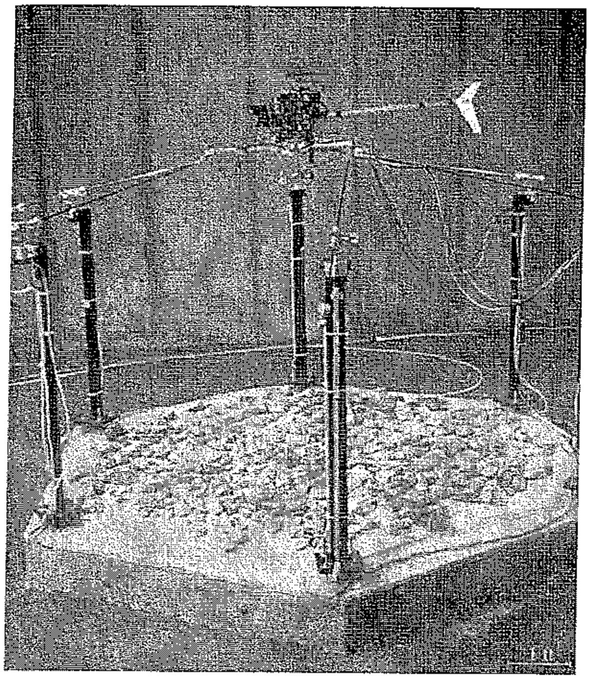

Fig. 6 6-DOF testbed

cone-shaped area with mechanical stops preventing the helicopter from crashing or flying away.

The testbed helicopter was outfitted with the odometer machine's two cameras and gyroscopes. The level area underneath the helicopter was covered with gravel to simulate outdoor rough terrain. The visual odometer machine, implemented by off-board computing (not shown), was evaluated using the 6-DOF testbed before integration on-board a larger and more capable helicopter. Fig. 7 shows the outdoor helicopter, Yamaha R50, which was stabilized and guided by the on-board version of the visual odometer machine.

8. Position estimation results

Fig. 8 compares ground truth lateral and longitudinal position (dashed lines) measured by the testbed, with vision-based estimates (solid lines), which were collected during flight tests of the 6-DOF testbed helicopter. The graph at the bottom of each plot shows the absolute value of the positioning error.

Helicopter maneuvers were performed under computer control during the flight tests to observe the positioning accuracy under abrupt (1–3 Hz, 5–10° amplitude) attitude oscillations In spite of the constant attitude oscillation and camera vibration, the lateral and longitudinal position estimates are accurate within 1.5 cm with 1/50 s latency Errors in the longitudinal direction are 50–60% larger due to the lower image field resolution in this direction. The lower resolution reduces the template match location accuracy. Other errors stem from small position displacements introduced each time a target template is reinitialized from the image center The lens imperfections, which were not carefully modeled, are also a significant source of error. Templates leaving the image are near the border and largely affected by lens distortion compared

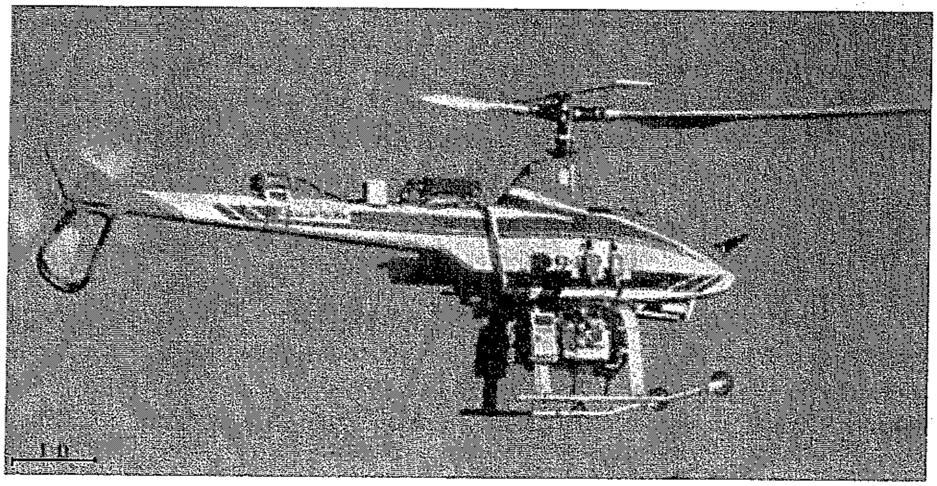

Fig 7 Autonomous helicopter

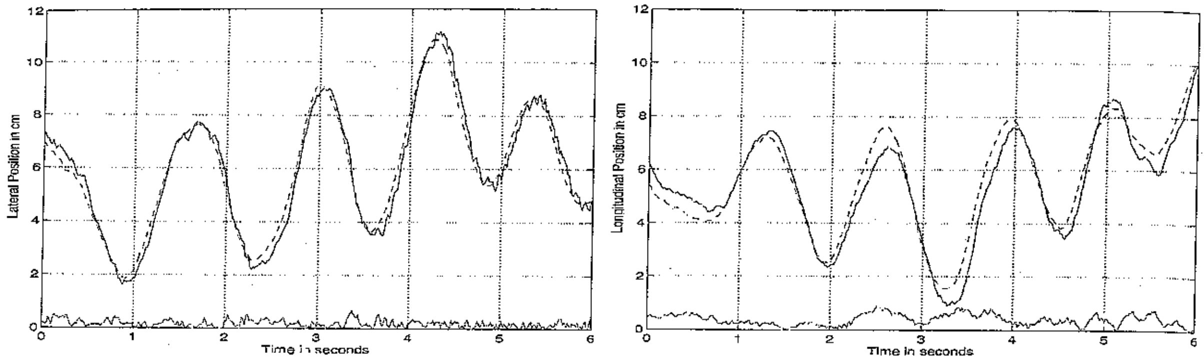

Fig 8 Indoor odometer positioning [solid: vision, dashed: ground truth]

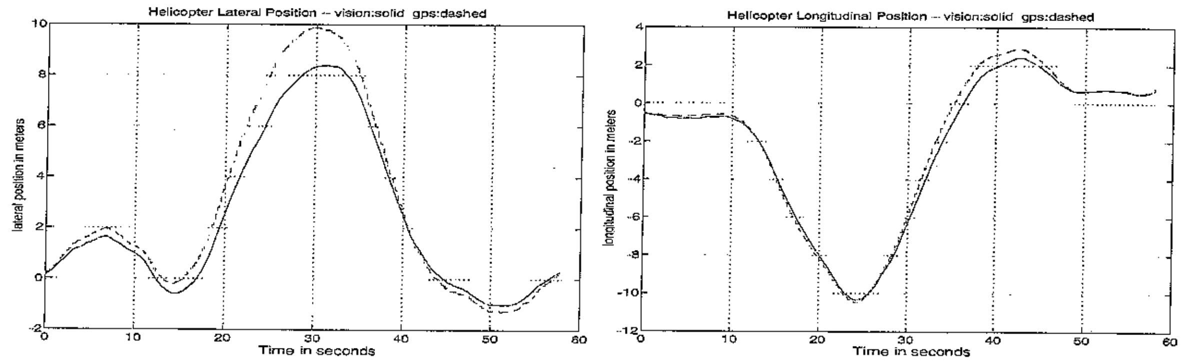

Fig 9 Outdoor odometer positioning [solid: vision, dashed: ground truth]

to new templates taken from the image center. A discrepancy of up to three image rows was observed near the image border using a 6 mm wide-angle lens. This pixel error translated to 0.54 cm error in helicopter lateral positioning for the helicopter height ($\sim$ 1.1 m) in the experiments. The same 3 pixel error produced a maximum of 1.2 cm error in the longitudinal direction each time a template was reinitialized. The drift rates of 2–3 cm were observed for 1 min hovering flight tests.

Fig. 9 shows odometer performance on-board the outdoor autonomous platform. Vision and GPS data was collected while the helicopter flew in a circular test path. The two dimensions are the X and Y axes of the local navigation frame, with Y pointing north and X east. Vision and GPS estimates matched accurately in the Y dimension but there was a consistent 20% difference in the X dimension. This discrepancy was attributed to the downhill grade of the terrain along this direction. This significant grade violates the visual odometer's flat ground assumption and adds a systematic bias to position measurement. No significant drift was detected in the circular test run.

9. Summary and discussion

This paper presented a visual odometer for helicopter positioning. The odometer incrementally maintains helicopter position by sensing image displacements. The odometer measures these displacements by visually locking on to and tracking ground objects. The odometer disambiguates rotation from translation in sensed image displacement by tagging images with measured helicopter attitude. The disambiguated image displacement is then transformed to determine helicopter motion. The odometer relies on two main assumptions. The first assumes that the helicopter flies over locally flat ground and the second assumes that the objects appearing in the field of view are rich in features.

The locally flat ground assumption simplifies range measurement by allowing the odometer to interpolation one range measurement for arbitrary image template range estimation. Extending the odometer's capability to handle non-flat areas requires matching of the main target template in two cameras. This approach provides accurate range to the target objects regardless of the ground shape. This extension of the algorithm was not pursued in building the first autonomous vision-guided helicopter prototype. Flight experiments presented in this paper were performed over locally flat or gently sloping farm land.

The assumption regarding the availability of image features can be made less restrictive by intelligent selection of high texture image regions. The statistical distribution of template matches can be applied to the entire image to determine areas of high contrast suitable for target template selection. This approach requires a 10-fold increase in the computational power currently realizable on-board the helicopter and was not pursued. However, the odometer was augmented with principal component analysis systems using K-L expansion [5,6] to help select objects of interest in the image. The augmented system successfully carried out a number of tasks including barrel label identification and tracking for the 1997 Unmanned Aerial Robotics Competition [4].

References

[1] O. Amidi, An autonomous vision-guided helicopter, Ph.D. Thesis, Electrical & Computer Engineering Department, Carnegie Mellon, August 1996.

[2] O. Amidi, T. Kanade, R. Miller, Autonomous helicopter research at Carnegie Mellon Robotics Institute, in: Proceedings of Heli Japan '98, Gifu, Japan, April 1998.

[3] P. Anadan, A computational framework and an algorithm for the measurement of visual motion, Int. J. Comp. Vision 2 (1989) 283–310.

[4] R. Michelson, Aerial robotics competition rules, Technical Report, Georgia Technical Research Institute, 1995.

[5] M. Uenohara, T. Kanade, Vision-based object registration for real-time image overlay, in: Computer Vision, Virtual Reality and Robotics in Medicine, CVR Med'95, France, 1995.

[6] S. Yoshimura, T. Kanade, Fast template matching based on the normalized correlation by using multiresolution eigenimages, in: IEEE/RSJ International Conference on Intelligent Robotics and Systems, Germany, 1994.

Omead Amidi is a research faculty member at Carnegie Mellon University (CMU) Robotics Institute (RI). He is the principal investigator of the CMU Autonomous Helicopter project and the co-principal investigator of the Reconfigurable Vision Machine project at CMU RI. His research focuses on aerial robotic vehicles which solely rely on on-board resources for autonomous flight. He has developed a number of autonomous helicopters capable of precise flight maneuvers using on-board vision. He received his Masters and Ph.D. degrees in Electrical & Computer Engineering from CMU in 1990 and 1996.