

RESEARCH ON AN AUTONOMOUS VISION-GUIDED HELICOPTER

Omead Amidi

Yuji Mesaki

Takeo Kanade

Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania

Abstract

We present an overview of the autonomous helicopter project at Carnegie Mellon's Robotics Institute. The goal of this project is to autonomously fly helicopters using computer vision closely integrated with other on-board sensors. We discuss a concrete example mission designed to demonstrate the viability of vision-based helicopter flight and specify the components necessary to accomplish this mission. Major components include customized vision processing hardware designed for high bandwidth and low latency processing and 6-degree-of-freedom test stand designed for realistic and safe indoor experiments using model helicopters. We describe our progress in accomplishing an indoor mission and show experimental results of estimating helicopter state with computer vision during actual flight experiments.

Introduction

Precise maneuverability of helicopters makes them useful for many critical tasks including rescue and security operations, traffic monitoring, mountain fire fighting, and inspection of power transmission lines. The goal of our project is to build a vision-guided helicopter capable of performing these tasks while flying autonomously. In addition to robust helicopter control methods, the development of such a system requires research on vision algorithms for helicopter positioning and object recognition necessary for navigation and tracking tasks, together with real-time hardware for high speed, robust execution of these tasks.

An autonomous helicopter's performance is critically dependent on accurate and frequent estimates of its position and attitude. We focus on methods to provide these estimates using on-board cameras closely integrated with other sensors such as gyroscopes and accelerometers.

We have demonstrated our first results on autonomous helicopter flight. We have built an indoor calibrated testbed that allows free flight experiments with model helicopters. We have custom designed vision hardware which integrates data from on-board sensors with real-time image processing and can now achieve frame-rate (**30 Hz**) vision-based state estimation. Integrating this vision hardware into a stable control system will lead to outdoor autonomous helicopter flight for performing useful, practical missions.

Motivation

A helicopter is an indispensable air vehicle for emergency operations, such as rescuing stranded individuals and spraying fire extinguishing chemicals for fighting forest fires. Uses of helicopters in the electric power industry include inspecting towers and transmission lines for corrosion and other defects. All of these applications demand dangerous flight patterns in close proximity to the ground or other objects which can risk pilot safety. An unmanned helicopter that operates autonomously or is piloted remotely will eliminate these risks and increase the helicopter's effectiveness.

Typical missions of autonomous helicopters require flying at **low** speeds to follow a path or hovering near an object. Positioning equipment such as Inertial Navigation Systems (INS) or Global Positioning Systems (GPS) are well suited for long range, **low** precision helicopter flight and fall short for very precise, close proximity flight. Maneuvering helicopters close to objects requires accurate positioning in relation to the objects. Visual sensing is a rich source of data for this relative feedback.

It is difficult, however, to recover helicopter position and attitude from vision alone. For instance, distinguishing between rotation and translation in a sequence

of images under perspective projection is extremely difficult. On the other hand, the new generation of lightweight gyroscopes and angular rate sensors in the market provide reliable measurement of angular change in an image sequence. For this reason, we concentrate on low-level, close integration of such sensors with vision.

Related Work

The study of the helicopter control problem is not new. Overcoming the inherent instability of helicopters has been the focus of a large body of research, including detailed mathematical models (eg., [10]) for control and Kalman filtering of multiple sensor data for state estimation (eg., [3]). The controller design methods range from linear quadratic (LQ) design to H^∞ design [19] and predictive control [8]. For example, a stable closed loop control system has been formulated [3] by quadratic synthesis techniques for helicopter autoland.

Recently, incorporation of a pilot model has been attempted based on quadratic optimal Cooperative Control Synthesis [17]. This model is used for control augmentation where the control system cooperates with the pilot to increase aircraft performance. The sophisticated pilot model developed by [7] attempts to describe the human's ability to look ahead, which is crucial to precise low-altitude helicopter control. While it is difficult to identify and verify these models, they provide a valuable basis for an intelligent helicopter controller, especially in designing low-level control loops. In this project, we employ a set of low-level controllers which have been designed by using a simplified helicopter dynamics model.

Actual flight tests of helicopter controllers have also been done. Notable implemented systems include those at NASA Ames Research Center [17], NASA Langley Research Center [3], and military aircraft manufacturers [5]. Fuzzy controllers have been successfully employed for actual helicopter flight experiments. In Japan, Sugeno's group at Tokyo Institute of Technology [14] has demonstrated fuzzy control of helicopters for crop dusting.

The state feedback for the above helicopter control experiments was primarily provided by on-board INS/GPS or ground-based beacon systems instead of on-board computer vision. Recently, we are beginning to see promising results in real-time vision-based processors, visual servoing of robotic manipulators, and accurate vision-based position estimation systems, some of which are applicable to autonomous helicopter control experiments.

The development of low-cost special-purpose image correlation chips and new multi-processor architectures capable of high communication rates has made a great impact on image processing. Examples of vision systems built from this kind of hardware include transputer-based image hardware for two-dimensional object tracking [4] and real-time tracking and depth map generation using correlation chips [9].

The high rate of image processing has made inclusion of visual feedback in servo loops practical. There is significant development in visual control of manipulators carrying small cameras, eye-in-hand configuration. Researchers at Carnegie Mellon's Robotics Institute [12] demonstrated real-time visual tracking of arbitrary 3D objects traveling at unknown 2D velocities using a direct-drive manipulator arm. The Yale spatial robot juggler [13] demonstrated transputer-based stereo vision for locating juggling balls in real time. Real-time tracking and interception of objects using a manipulator [11] has also been demonstrated based on fusion of the visual feedback and acoustic sensing.

Controlling by vision requires position estimation relative to desired objects and extraction of 3D scene structure based on sequence of images. RAPiD and DROID [6], developed by Roke Manor Research Limited, are systems designed for performing such tasks in unknown environments. RAPiD is a model-based tracker capable of extracting the position and orientation of known objects in the scene. DROID is a feature-based system which uses the structure-from-motion principle for extracting scene structure using image sequences. Real-time implementations of these systems have been demonstrated using dedicated hardware.

Integrating efficient model-based and connectionist techniques with powerful hardware architectures has produced an array of autonomous land and air vehicles. Significant advances in autonomous automobiles has demonstrated vision-based control at highway speeds. Most notable are Carnegie Mellon's Navlab [16] project and the work of Dickmanns at University of Bundeswehr, Munich involved with European PROMETHEUS project [2].

Didmanns applies a 4D approach exploiting spatiotemporal models of objects in the world to autonomous land and air vehicle control [1]. He has demonstrated autonomous state estimation for an aircraft in landing approach using a video camera, inertial gyros and an air velocity meter. Vision-based state estimation is also pursued at NASA Ames Research Center [15] using parallel implementation of multi-sensor range estimation for helicopter flight.

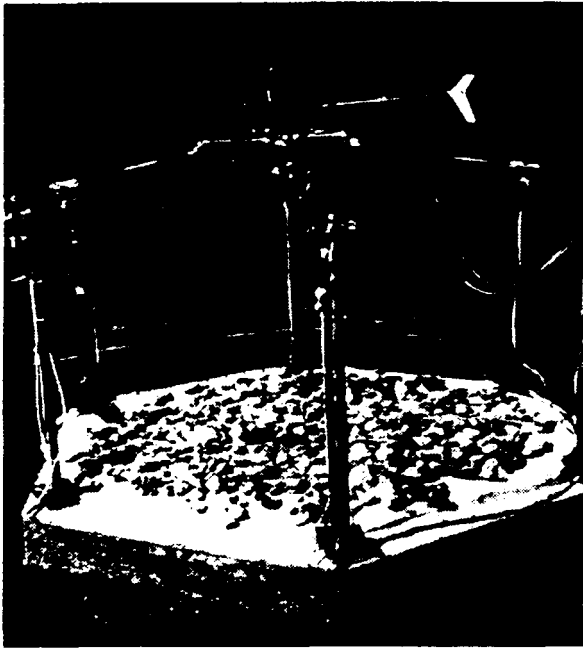


Figure 1: Indoor Testbed

An electrical model helicopter is supported by six light-weight graphite rods. A frictionless air bearing couples each rod with two-degree-of-freedom joints mounted on poles secured to the ground. Ground truth helicopter position is calculated from joint angles measured by shaft encoders.

Indoor Helicopter Testbed

For practical, calibrated experimentation, we have designed and built an indoor helicopter testbed. It consists of an electrical model helicopter mounted on a 6-degree-of-freedom (6-DOF) test stand (see Figure 1). Using the testbed, we can test each critical component necessary for autonomous flight before attempting potentially dangerous outdoor free flight experiments.

Model helicopters provide an inexpensive, safe, and logistically manageable way to experiment with helicopter control. They are faithful reproductions of full size helicopters with respect to the crucial rotor controls and configurations. Control techniques developed for the model helicopters can be directly applied to larger scale helicopters.

The helicopter in our testbed is attached to a frictionless 6-DOF stand as shown in Figure 1. The stand provides ground truth measurement of the helicopter position and attitude, and also works as a safety device preventing crashes and out-of-control flight. The helicopter on the stand can fly freely in a cone-shaped volume six feet wide and five feet tall without major inertial variations from free flight. The helicopter is fastened to six fixed poles by six light-weight graphite

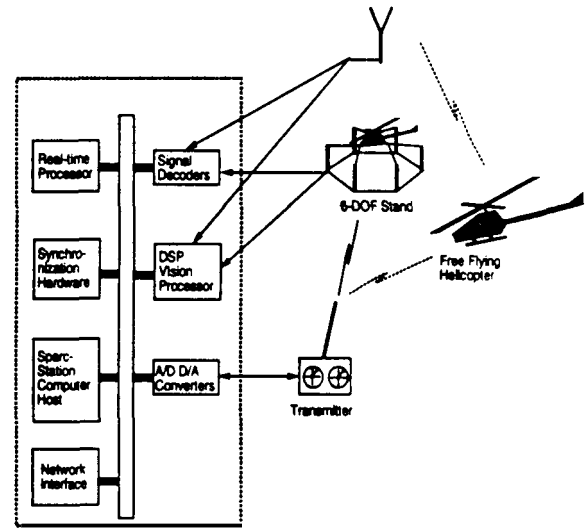


Figure 2: Testbed System Configuration

rods. Each graphite rod is free to move through a frictionless air bearing mounted on a two-degree-of-freedom joint. The joint angles are measured by shaft encoders and used by the computer to calculate the helicopter's ground truth position and attitude for experiment evaluation.

The computer system configuration, shown in Figure 2, consists of a host computer, customized vision processor, a real-time processor, synchronization hardware, and interfacing equipment. A hand-held radio transmitter used by a model helicopter pilot is interfaced to real-time computers. Using this interface, we can send computer control signals to the helicopter. The same interface can be used for free flying helicopters.

With this testbed, we can perform controlled experiments over a wide range of conditions. We can create various wind conditions by using fans, terrain conditions by placing objects, and helicopter setups by adjusting the mechanisms. Because of the safety provided by the testbed, even potentially disastrous situations like the failure of critical helicopter parts can be tested.

Using a simplified helicopter dynamics model we have implemented a control system capable of hovering the helicopter using linear controllers tuned at different operating points. This control system provides us with a stable platform necessary for conducting low-speed and hovering experiments.

One apparent limitation of the test stand is its inability to support larger model helicopters capable of lifting several sensors at once. On the other hand, since the test stand provides ground truth data, we can simulate data from certain sensors by purposely corrupting

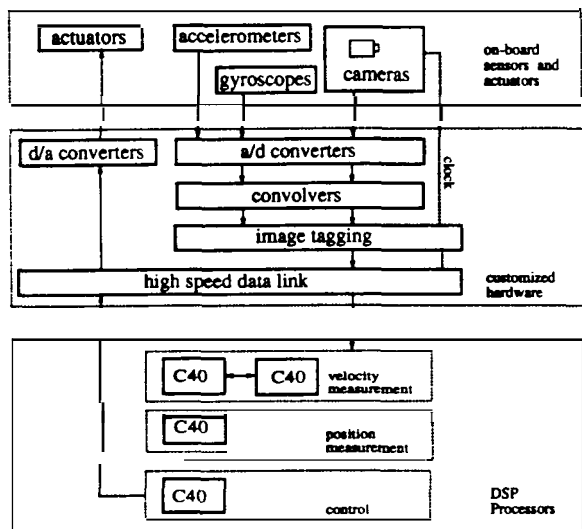


Figure 3: Vision Processor Structure

the stand data before using it. Different sensors can be individually characterized by comparing their response with ground truth data and their presence on-board the helicopter can be simulated during experiments.

Low Latency Vision Hardware for Helicopter Control

Our experience controlling model helicopters using the test stand has shown the necessity of velocity and position feedback rates of **15 to 30 Hz**. Processing image data at these rates requires fast computers capable of acquiring and processing images at frame-rate (30 Hz). There are a number of new cost-effective compact CPU platforms designed for high speed data transfer and processing. Among the most popular are: SGS-Thomson inmos T9000 Transputer, Intel i860, and Texas Instruments TMS320C40 Digital Signal Processor (C40). Our development is based on the **C40** platform primarily for its high speed communication ports each capable of transferring data at **20 MB/s**. Other advantages include: programmable Direct Memory Access (DMA) well-suited for image windowing operations, flexible memory architecture and internal bus structure, and wide availability. The structure of our customized vision processor is shown by Figure 3.

We have achieved close integration of vision with other on-board sensors using customized hardware designed to interface with an array of C40 processors. This low-level integration is key in providing robust velocity and position estimation.

Digitizer Configuration

The helicopter has multiple on-board sensors: two ground-pointing black and white CCD video cameras, vertical and directional gyroscopes, and accelerometers for each translational **axis**. The data from these sensors is digitized using multiple special-purpose digitizers. In particular, our system provides variable sampling rates for image digitization. Typically, the NTSC video signal is sampled at **14.3 MHz** which yields close to 1000 pixels per line. Conventional video digitizers choose 512 or 640 of **these** pixels per line during digitization. Since most CCD cameras have less than 1000 CCDs per line, we directly control digitizer sampling to reduce image data bandwidth and to provide more original image content. The aspect ratio of the image changes with sampling frequency and must be properly calibrated.

Convolution and Image Tagging

Fast convolution is essential for image preprocessing. In addition to edge detection and smoothing, matching and feature extraction can be performed using special convolution masks. We use real-time convolution hardware to perform Gaussian smoothing before processing images. To reduce image data bandwidth, we subsample the image using the digitizer before performing the smoothing operation. For the experiments described in this paper, **8x8** convolution masks were used on images sampled at **6 MHz** pixel frequency.

Using similar convolution hardware, accelerometer and gyroscope data are sampled at **120 Hz** and filtered by **64x1** Gaussian **FIR** filters. The filtered data is sampled and incorporated in the image data stream by an image tagger. Precise temporal matching of this data with the image is performed by using the camera's **60 Hz** field vertical sync clock (VSYNC) and shutter speed. We use **1 millisecond** shutter speed for tagging images accurately and reducing image blurring during helicopter motion.

High Speed Data Link

Because of the camera's **VSYNC** frequency, the processing time period for the sensor-tagged field images can only be multiples of **16.7 milliseconds**. Barely missing an image due to long processing time is expensive since the processor must wait for a new image for proper synchronization. Image field digitization alone requires **16.7 milliseconds**. During this time period the image must be transferred to the processor in order to achieve frame-rate (**30 Hz**) performance. We perform this transfer through a high speed data link designed to communicate with C40 processor comm-ports. This link incor-

porates small hardware buffers to convert the incoming synchronous image stream to the asynchronous comm-port protocol of the **C40**. In addition, since the image data is not directly entering a frame buffer, the high speed link provides proper comm-port synchronization with the camera using an internal state-machine. The comm-port design reduces CPU memory bus traffic by using C40's internal data buses and provides the ability to only transfer regions of interest using C40s versatile **DMAs**. These functions are crucial in improving processor speed.

Search Mission

As a concrete mission for an autonomous vision-guided helicopter, we envision a task of locating a known object in a predetermined outdoor area, for example, a particular car in a parking lot, and tracking the object by controlled helicopter flight.

The development of the indoor test stand allows us to conveniently simulate search mission scenarios using a variety of objects and terrain for visual tracking experiments. By carefully choosing these indoor experiments, we expect similar performance outdoors. The differences in flight altitude and terrain illumination can be resolved by small modifications to camera lenses, shutter speeds, and digitizing hardware.

Our mission is to search for a small car stranded somewhere in rough terrain. Performing this task requires object recognition to find the car, and visual measurement of position and velocity for autonomous flight. We have covered the stand base with gravel collected from the outdoor mission site to provide a realistic scene for our vision algorithms.

Velocity and Position Measurement

To measure helicopter velocity or position based on image data, we must first determine the displacement between consecutive images. This displacement in camera pixel coordinates is a function of camera attitude and distance relative to objects in the scene and camera calibration parameters such as focal length. For the indoor search experiments, camera attitude is estimated by gyroscopes and camera distance from the ground is estimated using the test stand. Performing outdoor experiments without the test stand requires altitude measurement by stereo vision possibly integrated with a laser rangefinder or microwave radar system.

The apparent displacement between consecutive images is a result of camera translation and rotation. Disambiguating rotation from translation is especially important for helicopter control since helicopter transla-

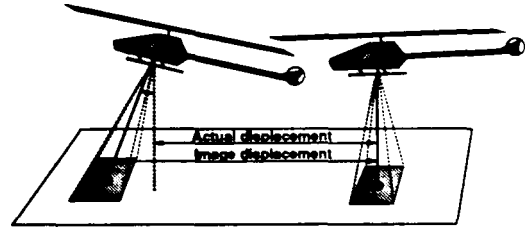


Figure 4: Effect of helicopter rotation

tion is directly a result of its change in attitude. Figure 4 shows the significance of this effect while the helicopter flares for reducing forward speed or stopping.

By carefully measuring the angular change between templates and images, we can estimate the effect of rotation and correct the image displacement to only reflect translational motion. This correction is useless without precise synchronization of gyroscope data with images. The drift common to all gyroscopes is not a problem here, since only the change in attitude is necessary from frame to frame.

Image Displacement Measurement

We use template matching to measure the displacement between consecutive images. We use sum-of-squared-differences (SSD) as our matching criteria. Each template is an $m \times n$ window of image intensities selected from the previous image. The best match of the template in the image can be determined by minimizing the SSD of the template and image pixels. To reduce the amount of computation, we restrict our search area to a small window around the template's neighboring pixels. The size of this search area is determined by helicopter altitude and anticipated worst case change in helicopter motion within one frame period. As the helicopter altitude decreases, the same translational motion causes a larger displacement in the image. The minimum altitude of the test stand is 1 meter and the on-board camera lens has 7.8 mm focal length. If we allow maximum helicopter velocity to be 1.5 meters per second during hover, our maximum image template displacement is 32 pixels per frame.

A coarse to fine strategy further improves search area and speed. We begin by using every fourth pixel to produce a coarse match for narrowing the search to 64 possible pixel locations. This estimate is improved to subpixel accuracy by fitting a parabolic surface to the SSD of the 64 match candidates. Figure 5 shows an example of a fitted parabola. A good parabola fit will refine the best single pixel match within ± 1 pixel. The parabola minimum is disregarded if it is not within one pixel of the single pixel match.

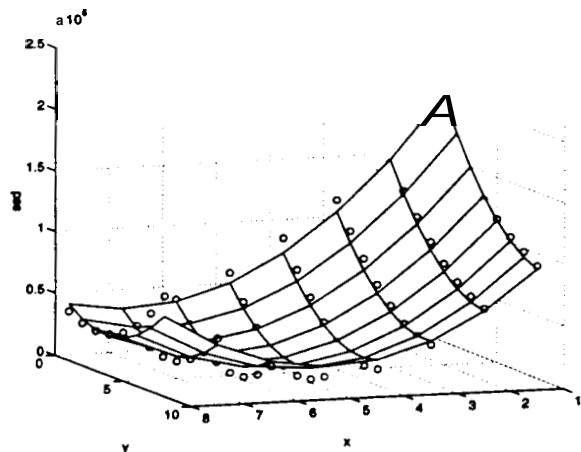


Figure 5: SSD Parabola Fit

In addition to subpixel accuracy] the fitted parabola provides match uncertainty information. A steep parabola versus a shallower one signals a more accurate match. Covariance matrices constructed from parabola coefficients will allow us to combine data from each template using a Kalman filter to produce the best estimate of image displacement.

For experiments reported here, we use four image templates for velocity and one template for position estimation as shown by Figure 6. The four velocity templates are 40×40 pixels in size and are positioned in each image quadrant. After each matching operation, the displacement of each template is calculated and the templates are updated with new image data from the same location.

Actual velocity measurement during flight is shown by Figure 7. This figure compares ground truth lateral and longitudinal velocity measurement (solid line) from the test stand with vision-based velocity estimates. The dashed and dotted lines in each figure represent vision-based velocity estimates with and without attitude correction. The correction was performed by measuring the attitude change between each template-image pair. Assuming images are taken from a locally flat surface, we can construct a transform, based on helicopter altitude and camera focal length, to convert the attitude change to a correction vector on the image plane for each template location. The effect of this correction is significant: **33 cm/s RMS error** in lateral velocity measurement without attitude correction versus **5 cm/s** after correction.

The position estimation template is 64×64 pixels in size and its location varies as the helicopter moves. This template is updated with image data from the best match in order to compensate for changes in helicopter

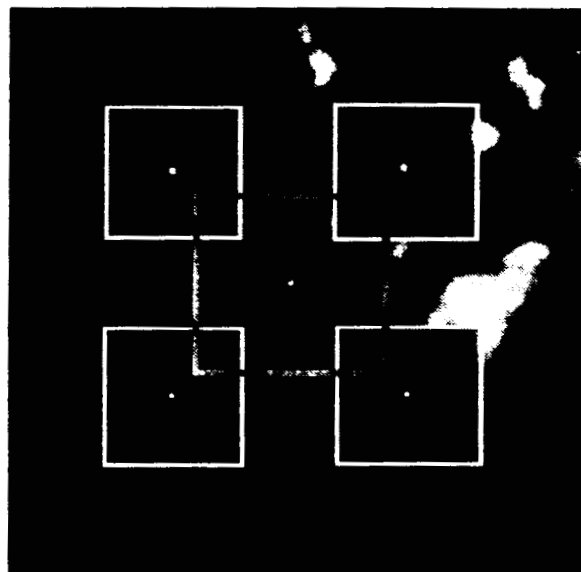


Figure 6: Image Templates

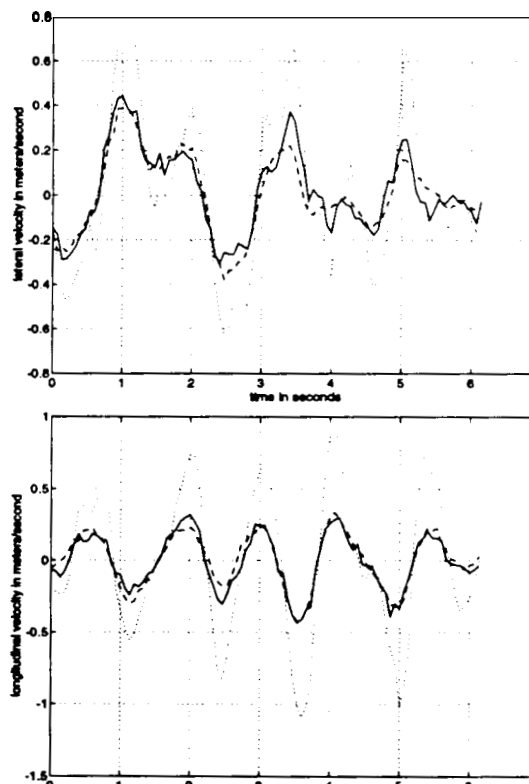


Figure 7: Vision-Based Velocity Measurement

The solid lines represent ground truth helicopter velocity from the test stand. The dotted lines show velocity based on image displacement alone and the dashed lines represent vision-based velocity with attitude correction.

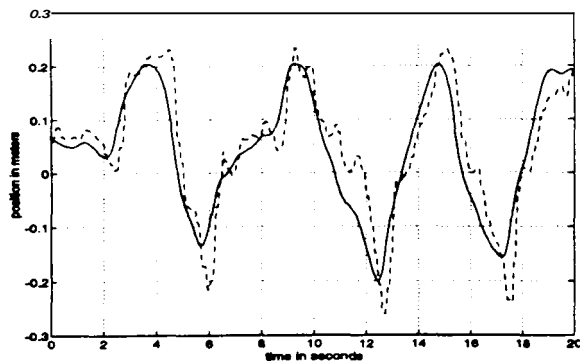


Figure 8: Position Measurement

The solid line represents ground truth helicopter position from the test stand. The dotted line shows position based on image displacement.

altitude and heading. If the best template match is close to leaving the camera view, the position template is loaded from the image center. A larger search area of 64 pixel displacement is used due to longer processing time. Figure 8 shows vision-based (dashed line) and ground truth (solid line) lateral position with respect to camera starting point. Attitude correction is more complicated in this case since the template changes position in the image plane. The figure shows uncorrected position estimation.

Position and Velocity Data Flow and Synchronization

We can not overemphasize the role of accurate synchronization in integration of on-board sensor data with high speed image processing. As observed above, attitude correction by synchronizing image and gyroscope data produces a significant improvement on position and velocity measurement accuracy. Figure 9 shows the data flow and synchronization we are performing for above helicopter motion estimation.

The solid vertical lines represent the camera VSYNC from the second image field (B). For high speed performance, only one image field (A) is used for motion estimation. The process begins with opening the camera shutter for 1 millisecond prior to VSYNC. Filtered gyroscope and accelerometer data is sampled with VSYNC and included in the image data stream by the image tagger. The tagged image is transferred to C40-1 which partitions field A for other C40s. The top half of field A is used by C40-1 and the bottom half by C40-2 for velocity estimation. In addition, field A is transferred to C40-3 for position estimation. Due to the high band-width of connections between C40s, it is

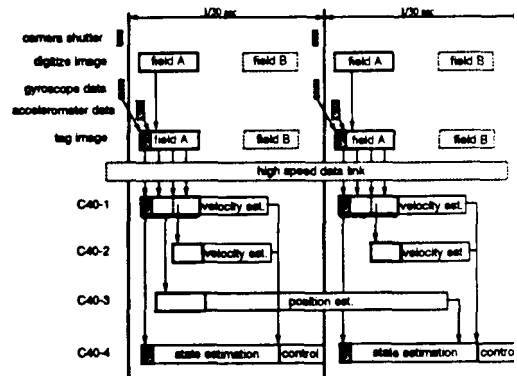


Figure 9: Data Flow and Synchronization

possible to start image processing during image transfer. The transferring is performed by DMAs which do not interfere with data processing. C40-1 also transfers synchronized gyroscope and accelerometer data to C40-4 which is responsible for state estimation and control. The state estimation is performed by transforming image displacement data from other C40s to helicopter translational motion. The estimated translational velocity and position in conjunction with accelerometer and gyroscope data are used by linear control loops to control the helicopter.

Object Search

The vision-based velocity and position estimation provides the basic capability for hovering and low-speed flight necessary for our indoor search mission. Locating the object of interest is the next step. We use template matching to perform this search. A major difficulty in this approach is that object orientation is unknown. This requires templates of the object in all possible orientation for matching. Methods such as K-L expansion [18] can be used to reduce computational complexity and storage of necessary templates. Another problem stems from varying helicopter altitude which will change the size of the object in the image. Close regulation and measurement of helicopter altitude is necessary to further reduce the complexity of the search.

We are conducting the search using a set of twenty 32 x 32 templates. These twenty templates, generated by K-L transform techniques, are sufficient for locating objects with $\pm 40^\circ$ orientation discrepancy as accurately as one degree resolution. The processing frequency for searching the entire image is 3 Hz using one C40 processor. Upon locating the object, the position estimator can now use the object in the image as its template providing relative helicopter position necessary for object tracking.

Conclusions

We have successfully developed the key components necessary for vision-guided autonomous flight. As our experimental results demonstrate, we are achieving real-time low latency image processing at suitable rates to stably fly helicopters. The major elements in our development have been custom designed vision hardware and indoor testbed. In addition to high speed processing, customized hardware provides flexible integration of on-board sensors which significantly improves vision-based state estimation. The indoor testbed provides convenient calibrated experimentation which is essential in building real autonomous systems.

References

- [1] E.D. Dickmanns. A general dynamic vision architecture for UGV and UAV. *Journal of Applied Intelligence*, 2:251-270, 1992.
- [2] E.D. Dickmanns and B.D. Mysliwetz. Recursive 3-D road and relative ego-state recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(2):199-213, Feb. 1992.
- [3] David R. Downing and Wayne H. Bryant. Flight test of a digital controller used in a helicopter autoland system. *Automatica*, 23(3):295-300, 1987.
- [4] H. Ekerol and D.C. Hodgson. A machine vision system for high speed object tracking using a moments algorithm. *Mechatronics*, 2(6):555-565, Dec. 1992.
- [5] Dale F. Enns. Multivariable flight control for an attack helicopter. *IEEE Control Systems Magazine*, April 1987.
- [6] C. Harris and C. Stennett. Rapid - a video rate object tracker. *BMVC90 Proceedings of the British Machine Vision Conference*, pages 73-77, Sept. 1990.
- [7] R. A. Hess and K. K. Chane. Preview control pilot model for near-earth maneuvering helicopter flight. *Journal of Guidance*, 11(2):146-152, April 1988.
- [8] Ronald A. Hess and Yoon C. Jung. An application of generalized predictive control to rotorcraft terrain-following flight. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(5), 1989.
- [9] H. Inoue, T. Tachikawa, and M. Inaba. Robot vision system with a correlation chip for real-time tracking, optical flow and depth map generation. *Proc. 1992 IEEE Int. Conf. on Robotics and Automation*, 2:1621-1626, May 1992.
- [10] Jurgen Kaletka and Wolfgang von Grunhagen. Identification of mathematical derivative models for the design of a model following control system. *Vertica*, 13(3):213-228, 1989.
- [11] R.C. Luo and R.E. Mullen Jr. Combined vision / ultrasonics for multi-dimensional robotic tracking. *Sensor Fusion: Spatial Reasoning and Scene Interpretation; Proceedings of the SPIE*, 1003:113-122, 1989.
- [12] N.P. Papanikolopoulos, P.K. Khosla, and T. Kanade. Visual tracking of a moving target by a camera mounted on a robot : a combination of control and vision. *IEEE Trans. Robot. Autom.*, 9(1):14-35, Feb. 1993.
- [13] A.A. Rizzi, L.L. Whitcomb, and D.E. Koditschek. Distributed real-time control of a spatial robot juggler. *Computer*, 25(5):12-24, May 1992.
- [14] M. Sugeno, T. Murofushi, J. Nishino, and H. Miwa. Helicopter flight control based on fuzzy logic. *Fuzzy Engineering toward Human Friendly Systems*, 1991.
- [15] R.E. Suorsa and B. Sridhar. A parallel implementation of a multisensor feature-based range-estimation method. *Proc. 1993 IEEE Conference on Computer Vision and Pattern Recognition*, pages 379-385, 1993.
- [16] Charles E. Thorpe. *Vision and Navigation: the The Carnegie Mellon Navlab*. Kluwer Academic Publishers, 1990.
- [17] Barbara K. Townsend. The application of quadratic optimal cooperative control synthesis to a CH-47 helicopter. *Journal of the American Helicopter Society*, pages 33-44, January 1987.
- [18] M.A. Turk and A.P. Pentland. Face recognition using eigenfaces. *Proc. 1991 IEEE Conf. on Computer Vision and Pattern Recognition*, pages 586-591, June 1991.
- [19] A. Yue, I. Postlethwaite, and G. Padfield. H^∞ design and the improvement of helicopter handling qualities. *Vertica*, 13(2):119-132, 1989.