Pilot behavior and course deviations during precision flight

Jeffrey B. Mulligan and Xavier L. C. Brolly

NASA Ames Research Center, MS 262-2, Moffett Field, CA 94035

ABSTRACT

In the fall of 2003, a series of flight tests were performed in the Tullahoma, Tennessee area to assess the ability of non-instrument rated helicopter pilots to fly precision routes with the aid of a Global Positioning System (GPS) receiver. The navigation performance of pilot subjects was assessed from GPS recordings of the flight trajectory, while pilot behavior was recorded using four video cameras, two of which were attached to a goggle frame worn by the pilot. This paper describes the processing methods developed for these data, and presents some preliminary results.

Keywords: eye movements, navigation, GPS, rotorcraft

1. INTRODUCTION

Helicopter pilots operating under Visual Flight Rules (VFR) in urban areas do so subject to strict regulations which ensure ample separation from fixed-wing traffic routes. This situation is now being reevaluated due to the widespread addition of Global Positioning System (GPS) receivers to the cockpit. These devices enable a VFR pilot to fly a route (specified as a series of waypoints) within narrow, defined horizontal and vertical airspace limits, while operating under VFR. We refer to this as Precision Visual Flight Rules (PVFR). The Federal Aviation Administration (FAA) has not published PVFR criteria or collected flight data to determine whether a VFR pilot equipped with a GPS can in fact navigate within a restricted flight path while adequately performing all of the aviation, navigation, and communication tasks necessary to operate an aircraft under VFR. Although GPS technology is frequently found in VFR helicopters, the ability to use a GPS receiver to fly within a defined airspace limit has not been validated. Research is needed to enable the FAA to determine airspace requirements, air traffic control procedures, and pilot operational and training considerations.

One area of operations which stands to reap substantial benefits from integration of GPS technology is Emergency Medical Service (EMS) helicopter flight. The mission of an EMS helicopter is to transport injured victims from an accident scene to a medical facility, or transport a critically injured patient between medical facilities. EMS operations are often dangerous due to the nap-of-the-earth flying, especially at night. As a result, controlled flight into terrain is the number one causal factor for EMS helicopter accidents.[1] EMS operations are often flown using Visual Flight Rules (VFR), which require pilots to navigate by visual reference to outside landmarks. VFR pilots are responsible for seeing and avoiding other aircraft and ground obstacles. Pilots are most vulnerable when responding to nighttime traffic accidents at unfamiliar sites that are surrounded by communication and power lines, or other difficult-to-detect ground obstacles. Besides EMS operations, PVFR routes may also enhance helicopter transitions through control zones and congested airspace, and flights through areas with natural or man-made obstacles, such as mountain passes and valleys.

The study described in this paper seeks to answer a series of specific questions: first, what level of performance can be expected from PVFR pilots, and does that performance support revising regulations to permit Simultaneous Non-Interfering (SNI) operations of rotorcraft and fixed-wing aircraft? Second, what sort of gaze behaviors are associated with PVFR flight? Does the task require so much visual attention on the GPS receiver that the pilot is unable to perform the obligatory scanning of the airspace for potential conflicts? Also, do the behaviors associated with superior performance suggest practices that should be incorporated into pilot training?

To address these questions, a series of flight tests were flown in October 2003 in the Tullahoma, Tennessee area. Pilot test subjects flew a route specified by 21 waypoints, some of which corresponded to visible landmarks, and others which were specified only by their latitude and longitude. Video data were collected during each flight

using the Ames portable eye-tracking system.[2] This remainder of this paper describes the data collection in more detail, presents the post-processing techniques developed for the GPS and video data, and concludes with some preliminary results.

2. DATA COLLECTION

A route was designed consisting of a series of waypoints (see figure 1). In some cases the route followed conspicuous terrain features: one series of segments followed the course of the Duck River, while several others followed roads. Detailed specifications of the waypoints are given in table 1; additional details concerning the route and flight test protocol can be found elsewhere.[3]

Table 1. Specification of the flight test route. Latitude and longitude are given in degrees. For each waypoint, the "following segment description" refers to the segment between the given waypoint and the next.

Waypt.	lat.	long.	rel. dist.	turn ang.	following segment description
0	35.3754	-86.2432	0	0	GPS transition from THA to PVFR route
1	35.3332	-86.2484	0.0443177	-92.851	visual flight along a power line
2	35.3294	-86.1953	0.0899592	-12.3893	visual flight along a power line
3	35.3347	-86.1586	0.121876	12.4125	visual flight along a power line
4	35.3312	-86.1098	0.163784	-128.966	GPS segment to avoid simulated restricted area
5	35.3857	-86.1477	0.229333	72.3488	GPS segment to avoid simulated restricted area
6	35.4388	-86.1069	0.294892	44.3114	GPS segment to avoid simulated restricted area
7	35.4472	-86.0482	0.345902	51.2137	GPS segment to avoid simulated restricted area
8	35.4218	-86.021	0.38119	-108.441	GPS route near Hwy 55
9	35.4637	-86.0018	0.427916	-90.5999	GPS route
10	35.5127	-86.1117	0.534997	-94.3728	GPS route following Duck River
11	35.4608	-86.1303	0.591457	46.9548	GPS route following Duck River
12	35.4475	-86.161	0.621243	-7.78002	GPS route following Duck River
13	35.4333	-86.1844	0.646124	111.432	GPS route following Duck River
14	35.4472	-86.1868	0.660735	-45.3991	GPS route over Normandy Lake
15	35.4617	-86.2076	0.684158	-22.4574	GPS route over Normandy Lake
16	35.4702	-86.2463	0.718437	-28.6058	GPS route
17	35.4482	-86.3218	0.786999	-64.2257	GPS route near Hwy 278
18	35.4036	-86.3293	0.834043	-30.733	GPS route
19	35.3352	-86.3028	0.909056	-66.7563	visual flight along a power line
20	35.3332	-86.2484	0.955682	-85.0963	GPS transition from PVFR route to THA
21	35.3754	-86.2432	1	0	n/a

A group of pilot subjects was recruited and brought to the University of Tennessee Space Institute flight research center in Tullahoma, Tennessee. The vehicle was a Bell OH-58A+ helicopter (military version of the Bell Jet Ranger). The vehicle was equipped with a GPS receiver (Bendix-King KLN89B) which was programmed with the flight route and provided guidance to the pilot regarding course deviations and upcoming waypoints. Each pilot received a pre-flight briefing, and then flew with an instructor, who demonstrated the operation of
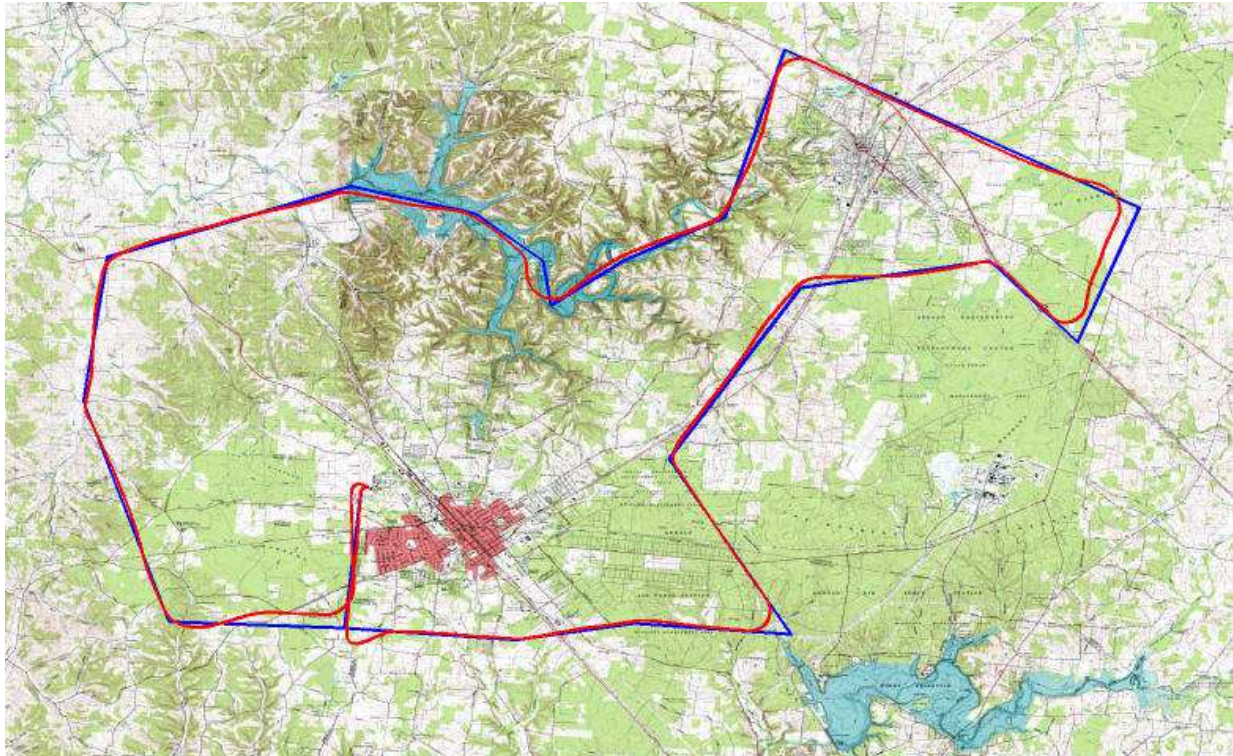
Figure 1. Topographic map of the Tullahoma region, with the test route and a typical flight path superimposed.

the cockpit GPS receiver. After the instructor was satisfied that the subject understood the operation of the GPS system, the route was flown once in the clockwise direction. Data were recorded during a final circuit of the route performed in the counter-clockwise direction.

Deviations from the desired route can arise from two separate sources. "Flight technical error" refers to errors arising from the GPS system itself: the vehicle may not be exactly where the GPS receiver says it is. Sources contributing to the flight technical error may be from both random (e.g. electrical noise in the receiver) and systematic (effects of ambient temperature on the receiver, the number and positions of the satellites used at a given time). If the pilot were able to perfectly adhere to the route indicated by the in-cockpit receiver, this would be the only error. However, a second source of error arises from the finite precision with which a pilot can adhere to the indicated route. This "navigation error" is added to the flight technical error, resulting in the "total system error" (TSE).

In order to obtain independent estimates of the technical and navigation errors, the true aircraft position was obtained using a survey-quality system consisting of a pair of GPS receivers (Ashtech Z-12). One of these receivers was located in the vehicle, while the other remained fixed at a surveyed location. The data streams from both the on-board Bendix and Ashtech receivers were recorded to a laptop computer operated by an engineer seated in the rear of the vehicle during each of the flights. The Ashtech data were later corrected using the data logged by the second stationary Ashtech receiver, using Ashtech software. The error of this "ground-truth" data is estimated to be less than than 1 meter.

Video, audio, and GPS data (from the on-board Ashtech unit) were all recorded on an 8mm videocassette tape using a miniature recorder (Sony GVD-200). Images from four individual cameras were multiplexed into a single image using a "quad processor" (Supercircuits QS-25). A time code generator (Horita GPS-3) produced linear time code (LTC), which was recorded on one of the two audio channels. The time code generator also interfaced to the GPS receiver, which allowed the time code to be automatically initialized to the correct time. In addition, the time code generator encoded latitude, longitude and altitude data into the "user byte" portion

of the LTC signal. The remaining audio channel was connected to the cockpit audio system. (An earlier version of the system using an analog tape recorder converted the linear time code to vertical interval time code (VITC), putting the data into unused video lines, and thereby keeping both audio channels free to record audio data. The "digital8" format used by the GVD-200 does not encode the video lines corresponding to the vertical interval, however, and so VITC could not be used.)

The uncorrected Ashtech data recorded along with the video data were used to generate the results presented here. Both the Ashtech correction and the flight technical error produced by the Bendix-King unit are small compared with the observed navigation errors, so it is unlikely that incorporation of the corrections will change the results significantly.

Test subjects were fitted with a goggle-mounted camera assembly. The goggle was based on a racquetball eye-shield (Leader Vision 2), from which the majority of the lens was cut away. A miniature monochrome video camera imaged subject's right eye through a "hot" mirror, having a high near-infrared reflectance while appearing visibly transparent. A monochrome scene camera was mounted on the upper portion of the goggle, and provided a "subject's eye view" of the scene.

The group of subjects varied in the degree of experience, having from 286 to 7000 hours of flight time (see table 2).

Table 2. Qualifications and flight hours of test subjects.

Subject code	rating	VFR hours (helo)	IFR hours (helo)	day/night/both
p1	VFR	7000	-	day
p2	IFR	1800	120	day
a1	VFR	2595 (2445)	-	both
p3	IFR	600	200	both
p5	IFR	3960	190	day
p6	VFR	725 (180)	-	day
p7	VFR	286	-	day
p8	VFR	352	-	both
a2	VFR	865 (645)	-	both
p9	IFR	2061	175	both

3. FLIGHT TRAJECTORY ANALYSIS

3.1. Definitions

To assess the degree to which pilots where able to stay on-route, we computed the "lateral course deviation," which expresses the distance from the vehicle flight path and the desired route in the horizontal plane (ignoring altitude). This is computed by determining, for each point in the flight trajectory, the nearest point on the route. This point is generally found by intersecting the current route segment with the perpendicular line passing through the vehicle position. When the trajectory goes "outside" of a corner, however, (see figure 2), this point does not fall between the waypoints defining the segment, in which case the waypoint itself is the nearest point on the route. The trajectory is processed sequentially, advancing the current segment when the distance to the next segment becomes less than or equal to the distance to the current segment. Thus, for each point in the trajectory, there is an associated segment index which advances monotonically as a function of time from 1 to 22 (the number of linear segments in the route).

The procedure just described associates each point on the trajectory with a point on the route. We define the route distance of a point on the route to be the distance (measured along the route itself) from the start to
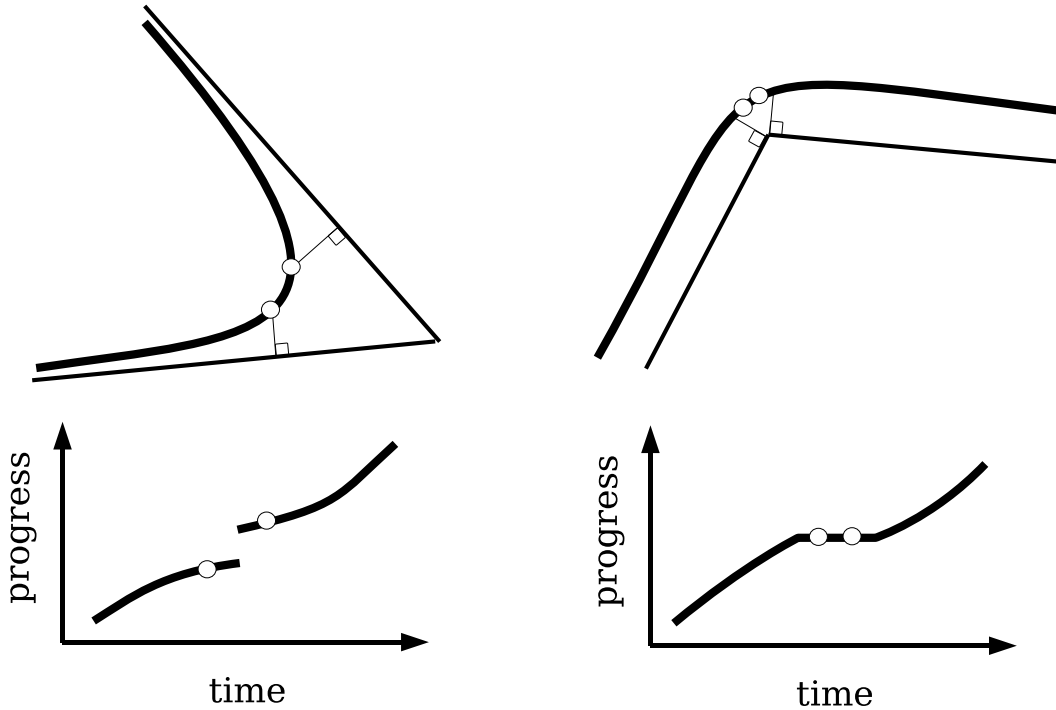
Figure 2. Schematic diagram illustrating the mapping from flight trajectory points to points on the route polyline. Two distinct kinds of singularities are encountered: for turns "inside" a waypoint (left), the nearest point on the route jumps from one leg to the next, producing a discontinuous jump in "progress" (distance along the route). For turns "outside" a waypoint (right), multiple trajectory points are all mapped to the waypoint. .

the point in question. This can be normalized by the total length of the route, giving us a measure of "progress" which advances from 0 to 1. By plotting data as a function of progress (instead of time), we are able to compare data from different subjects at corresponding parts of the route.

Our current definition of the course deviation is deficient in that it does not account for the finite turn radius in the vicinity of a waypoint. The GPS receiver can be programmed for two distinct type of waypoints: "fly-over" waypoints are those for which it is desired that the aircraft pass directly over the waypoint. A waypoint at the runway threshold on an approach route would typically be a "fly-over" waypoint. Most waypoints, however, are "fly-by," meaning that it is not important that the flight path passes directly over the waypoint. GPS receivers programmed for "fly-by" waypoints will typically give an indication to the pilot that it is time to begin the turn some distance before the waypoint. The precise nature of this "anticipation" turn varies from manufacturer to manufacturer, and we are in the process of obtaining data for the Bendix-King unit used in this study. Ideally, the course deviation error should be defined relative to the target flight path indicated by the receiver, and not the line segments linking the waypoints as we have used here. Incorrect computation of the error in the vicinity of the waypoints will produce an overestimate of the navigation error, but is unlikely to interact with other variables.

3.2. A note on units

GPS data have the form of latitude and longitude, whereas we would like to express course deviations in meters. At a local scale, displacements in latitude and longitude are mapped to displacements in meters by a diagonal linear transformation. The diagonal elements have equal values at the equator, and are proportional to the distance from the center of the earth (elevation). For a given land elevation, the diagonal factor is constant for latitude, while for longitude it decreases as the cosine of latitude.

The course deviations computed for this paper were measured assuming a constant elevation corresponding to ground level in Tullahoma. The scaling factors used to convert from angular coordinates to meters were obtained
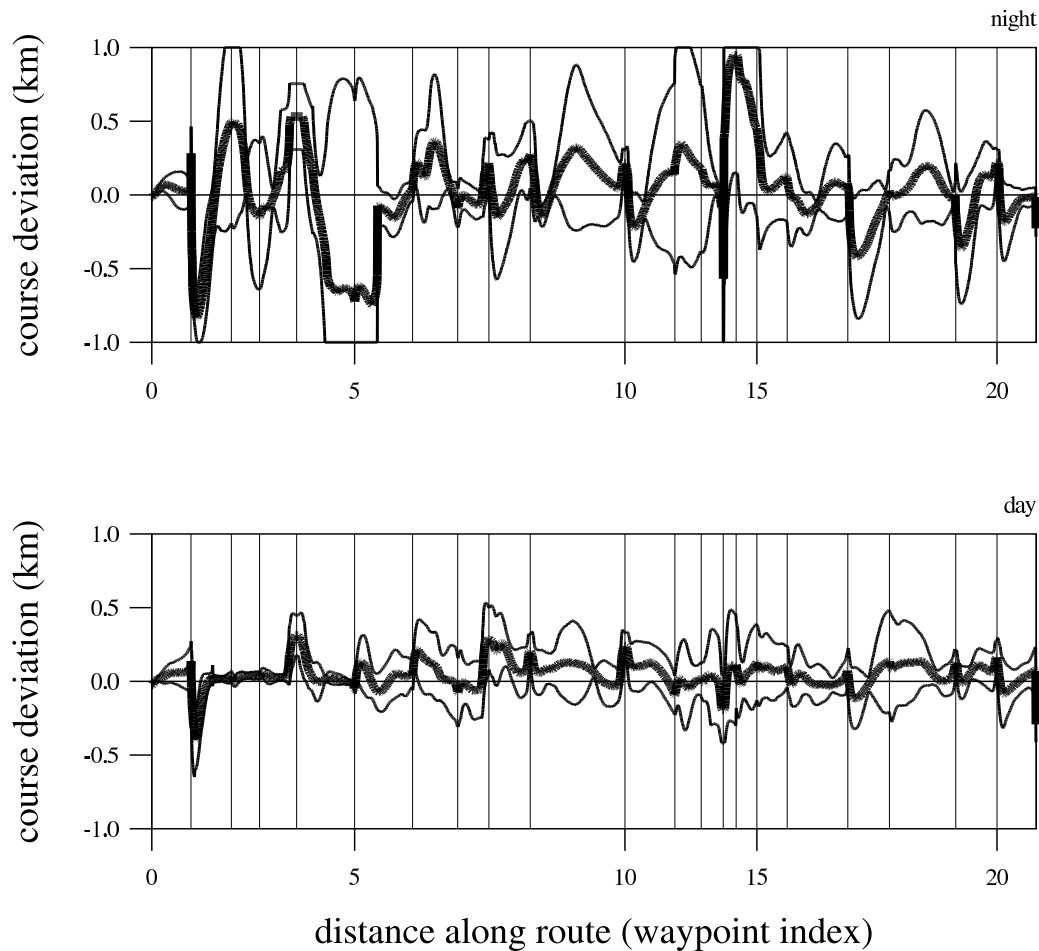
Figure 3. Lateral course deviation as a function of progress along the route. Heavy curves depict the means, while the thinner curves indicate plus and minus one standard deviation. Upper panel: night; lower panel: day.

using the United States Geological Survey topographic quad "Tullahoma" (7.5 minute series), by computing the ratio of the length of the map's 1 kilometer scale bar to the length of each side of the rectangular map area. The factors were used for all points on the route, ignoring the small variation with latitude. This introduces an error of approximately 0.2% between the most northern and southern parts of the route, corresponding to an error of 2 meters in a 1 kilometer estimated course deviation.

3.3. Results

Figure 3 shows the lateral course deviation versus progress along the route, averaged across subjects. The abscissa represents distance measured along the route, as described above, with each of the waypoints indicated by a vertical grid line. The heavy curve shows the average error, while the two thin curves represent plus and minus one standard deviation. The upper panel of figure 3 shows the results averaged across all night flights, while the lower panel shows the average of the day flights. A concise summary is obtained by averaging the (unsigned) error magnitude over the entire route: for the day flights, the average error magnitude is 0.123 km, while for the night flights it is 0.280 km.

Figure 4. Sample recorded frame showing quad tiling of images from the four cameras. Upper left: fixed camera looking down and to the right over the subject pilot's left shoulder; the right hand and knee of the safety pilot is seen at the left edge. Upper right: subject pilot's right eye (rotated due to camera mounting). Lower left: image from forward-looking head-mounted scene camera, providing the "subject's eye view." Lower right: view from fixed camera on the instrument panel looking up at the subject pilot's face.

4. VIDEO PROCESSING

Before any processing could be done, the data first had to be transferred from the tapes to a computer. This was done at the University of Tennessee Space Institute (UTSI) campus, using a computer workstation equipped with an analog frame grabber (Matrox Meteor 1). Specially developed software allowed real-time digitization to a pair of dedicated hard disk drives with a capacity of approximately 30 minutes. As each recording had a duration of approximately 1 hour, each recording had to be digitized in two sections. After digitization, the raw images were converted to JPEG sequences, and moved to a conventional file system. The audio and GPS recordings were digitized along with the video. Following this procedure, the files were transferred from UTSI to NASA Ames over the internet.

4.1. Camera demultiplexing

Figure 4 shows a typical video field. Each field consisted of four quadrants, each of which corresponded to one of the four cameras. Camera demultiplexing refers to the process of taking a single movie consisting of the composite frames, and creating four movies corresponding to the individual camera streams. This was accomplished by a straightforward selection of the spatial subregions corresponding to each camera's image. The process was complicated, however, by the fact that the quad processor (which combined the four camera signals into a single signal) sampled the camera signals asynchronously; in other words, each frame put out by the quad processor and captured on tape consisted not of a complete frame from each camera, but was generally made up of portions of two consecutive camera frames. When the objects viewed by the camera were stationary, this could be ignored, but when the objects moved the result was a tearing of the frame. Because each of the four cameras had its own clock, the frame rates were all slightly different, and the tearing artifact occurred at a different position within each subimage.

This artifact was eliminated by first locating the occurrence of the tearing artifact, by looking for image discontinuities between pairs of adjacent scan lines. The vertical position within the frame containing the maximum discontinuity was determined, and plotted as a function of time. Because the artifact was produced by the difference in clock frequencies between the two devices, the discontinuities corresponding to the artifact

fall on a function which is linear in time, resembling a sawtooth. We fit a model to the observed data to reject outliers generated by vertical discontinuities in the image not related to the artifact.

An additional complication arises from the fact that the quad processor uses the interlaced format for its output signal. To reconstruct a camera field, we must deinterlace the recorded video from the quad processor. When the tearing artifact is present, it is only visible in one of the two fields output from the quad processor. Depending on whether it is the first or the second field, we must go forward or backward in time to recover the missing parts of the frame.

4.2. Vibration compensation

In viewing the recordings from the face camera mounted on the instrument panel, non-rigid distortions of the image were observed, which were presumed to result from vibration of the camera. These distortions were corrected as follows: first, a few prominent stationary features (parts of the vehicle visible to the sides of the pilot) were identified and tracked over the entire sequence. The motion of the camera in time was recovered from these displacements by remembering that the video lines are scanned sequentially in time; thus, the time (relative to the start of the frame) at which a feature was imaged was proportional to the vertical distance from the top of the frame. After assigning the proper time to each observation, the motion was well-fit by a simple sinusoid. Using the inferred motion of the camera, the images were then warped to produce a relatively undistorted sequence.

4.3. Eye camera video

Our initial analysis of the eye camera images consisted of localization of the pupil (inner boundary of the iris) and the corneal reflection of the infrared LED's used to illuminate the eye. For the night flights, we obtained images similar to those we routinely gather in the laboratory. The images from the day flights, however, posed some new challenges. Because the ambient daylight illumination was much stronger than that provided by the LED illuminators, these sequences are rife with illumination variations, as the vehicle changed its attitude relative to the sun. During the day the illumination provided by the LED's was generally much less than the ambient illumination, but the reflections of the LEDs themselves were still visible.

Another source of illumination variations was the vehicle rotor: because the clear windshield extended back over the pilot's head, a shadow was cast as the rotor passed overhead. Because this was a brief event, it only affected a few video scan lines, producing a dark band in the image (see Figure 5, left panel). The band is vertical in figure 5 because the image has been rotated to put the eye in the proper orientation.

In addition to the illumination variations, there are a number of other features of the daylight eye images which have made robust tracking problematic. Because of the high ambient light levels, the pupil tends to be constricted, making it a smaller target. Similarly, the resting pose of the eyelids tends to be more closed, as if the pilots were squinting. This is problematic for two reasons: first, the eyelids hide more of the eye when they are partially closed; second, the upper eyelashes move in front of the pupil as the lid is closed, obscuring the features we are trying to detect.

Because of all these factors, our initial efforts to track the eye in the daylight videos have been only partially successful, with estimates obtained for only about 40% of the frames in the two flights processed. To overcome this shortcoming, we plan to redo the analysis, introducing a number of new techniques. In frames where the eye is visible, we will track the limbus (outer margin of the iris) in addition to the pupil. In addition to providing an additional feature, localization of the limbus will also provide a check on the pupil localization, because these two features should share a common center. (Refraction by the cornea makes them have slightly different centers for eccentric gaze directions, but this can be taken into account.)

We also plan to introduce methods to estimate gaze direction when the eye itself is hidden by the upper eyelid. We expect that the vertical component of gaze will be especially easy to recover, because the lid moves with the eye, and therefore the vertical position of the lid is monotonically related to the vertical component of gaze. The horizontal component may be more difficult to extract, but we note that because of the fact that the cornea is a small dome rising out of the roughly spherical eyeball, its lateral motion causes a change in the shape of the covering eyelid, and in particular the form of the margin of the lid. Accuracy using this technique may
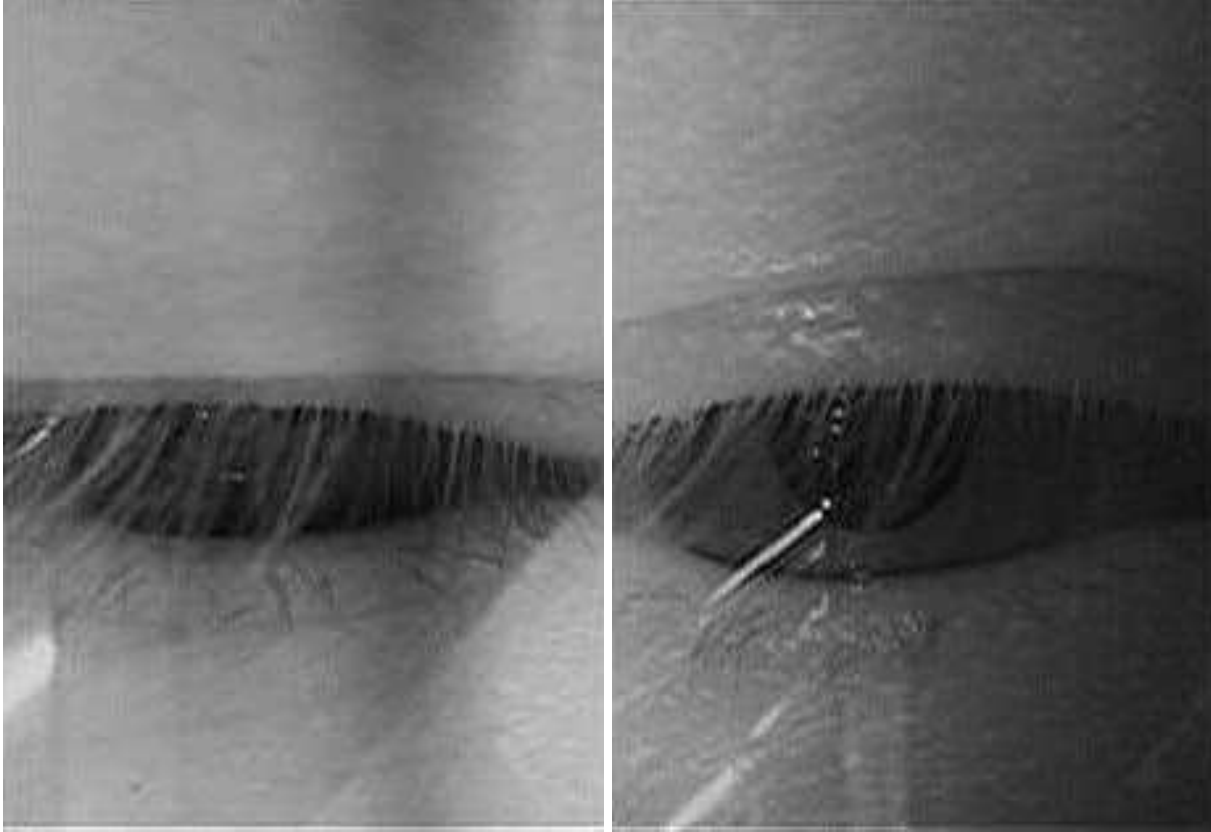
Figure 5. Typical images of eye captured during day flight (left) and night flight (right). Tracking of the eye during the day is complicated by the fact that most of the illumination comes from uncontrolled environmental sources, the pupil is generally constricted and the eyelids tend to be more closed.

suffer for two reasons: first, the measure itself is likely to be less sensitive than direct measurement of the pupil position; and second, we may not have calibration data for the extreme down-gaze positions for which the lid entirely hides the eye. However, these gaze directions do not correspond to those of most interest for this study (i.e., the GPS receiver and out-the-window landmarks), but rather correspond to the instruments at the bottom of the panel, and charts in the pilot's lap. Therefore, we believe that degraded accuracy for these gaze targets will be acceptable.

4.4. Face camera processing

We obtain an estimate of the pose of the pilot's head by analysis of the images from the fixed camera mounted on the instrument panel. Our procedure is a mix of automatic and manual procedures. First, a set of conspicuous features on the head are selected, such as the headset earphones, the microphone, etc. Next, a training set of 150 frames is selected. For each frame in the training set, an operator manually indicates the position of each feature using the mouse. At this point, we have 150 views of each feature, stored as small subimages. The various appearances of a feature can be efficiently described using a small number of parameters by applying a Principal Components Analysis (PCA) to the set of feature appearances, a technique first applied to entire face images by Turk and Pentland.[4]

We next obtain an approximate 3-dimensional configuration of the features from a pair of mug-shot views, that is by picking a view which is close to frontal and another which is nearly profile. The positions of the features in the frontal view give us the approximate x (side-to-side) and y (vertical) coordinates of the features, while the profile view provides approximate z (fore-and-aft) and y. We then refine the the estimates by alternately

Figure 6. Image from "face-cam" with superimposed rendering of rigid 3-D space curve linking tracked features.

optimizing the structure and pose parameters over all 150 training images. This procedure stabilizes after 2 or 3 iterations, at which point we have estimates of both the 3-D structure of the features, and the pose of the head in each of the training images.

The next step is to derive the relationship between the pose and the appearance of each of the features (as described by the eigen-feature coefficients). For each training frame, we have a set of pose parameters (angles) and a set of coefficients describing the appearance of the features. We derive an algebraic relation between the pose angles and each of the coefficients, which allows us to predict the appearance of each feature for an arbitrary pose, including poses which we may not have seen before.

We are now ready to describe the pose estimation process for an arbitrary new frame: we first make a guess about the pose, either recycling the final pose estimate from the previous frame, or assuming a frontal view for the first frame. Using this guess, we predict the corresponding appearance of each of the features. Using the expected feature appearances, we then search for each of the features in the image using cross-correlation. From the locations of the features, we estimate the pose. If the new estimate of the pose differs from our initial guess, we recompute the appearance of the features using the new pose estimate, and repeat the process until the estimate is stable (usually 2-3 iterations). Typical results are shown in figure 6.

4.5. Scene camera processing

The images gathered by the head-mounted scene camera provide a second source of information about the position and orientation of the head. Structure-from-motion refers to a technique by which both the camera pose and the 3-D locations of scene features can be computed from a series of images. While we ultimately hope to apply this technique, here we present a simpler method in which we approximate the camera motion by a pure rotation about the camera's optical nodal point. This simplification affords two advantages: first, we do not have to solve for the 3-D structure (or construct an accurate model of the cockpit interior); second, instead of identifying and tracking individual features, we can simply solve for the camera pose parameters which provide the best overall registration of the image with the previous image or a template formed by mosaicking a set of images.

To register images related by large rotations, we must take into account the effect of the perspective projection performed by the camera-lens system. Because the camera-lens system projects the sphere of viewing directions onto a flat image plane, it is necessary to apply a complex non-rigid warp to bring two images into correspondence. We address this problem by adopting a cylindrical coordinate system to which we transform all the images.

To derive the transformation from the image sensor coordinates to the global coordinate system, we assume a generic pinhole camera model. But this is a poor approximation to our actual camera, which has a short focal
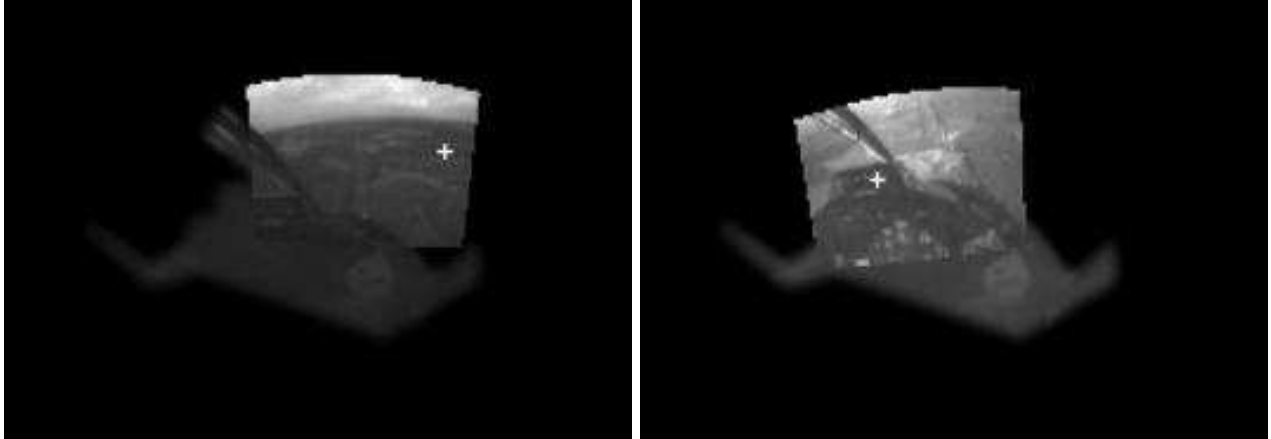
Figure 7. Images from the head-mounted scene camera corresponding to the frames in figure 6, inset into the cockpit mosaic. The position of the inset indicates the orientation of the head, while the small white cross indicates the direction of gaze.

length wide-angle lens which introduces considerable lens distortion. This distortion is evident in the appearance of the horizon, which usually appears curved in the raw video images. We apply an approximate correction for lens distortion by assuming a generic lens distortion model, and adjusting its single parameter to produce a linear horizon in a small number of representative frames.

After correcting the raw video for lens distortion, we proceed to construct a mosaic of the cockpit as follows: we initialize the mosaic using an image filled by the instrument panel. Successive images are processed by first making an initial guess concerning the camera orientation (usually the orientation estimated for the previous frame). We then use the estimated orientation to warp the image to the common image space. The quality of the resulting registration is assessed by computing the normalized cross-correlation. The STEPIT optimization routine[5] is used to adjust the rotation parameters to optimize the fit.

Typical images from the scene camera contain both fixed features of the cockpit, and moving terrain features seen out the window. Because of the motion of the aircraft, these terrain features are not useful in determining the pose of the head, and we therefore wish to exclude them from the registration process. This is done by hand-construction of a mask which selects the portion of the mosaic image corresponding to the vehicle instrument panel and frame. The two images in figure 7 each show the masked mosaic, with an input frame registered and overlaid.

5. DETERMINATION OF GAZE

The ultimate goal of the video analyses described in the preceding sections is to determine the direction of gaze of the subject pilot. Here we describe the methods used to accomplish this. At this writing, while we have successfully localized the eye in tens of thousands of frames, this is still a small fraction of the total data set, and we therefore defer presentation of actual results to a subsequent publication.

Determination of the target of gaze is accomplished in two steps: first, analysis of the eye image allows us to compute head-relative gaze; second, this head-relative gaze direction is transformed to a cockpit-relative direction using an estimate of head pose, obtained from either the face-cam video or the head-mounted scene camera.

The relation between the position (in the image) of the eye and the direction of gaze (relative to the head) depends on a number of factors, including the size and shape of the subject's eye, the alignment of the eye camera and mirror, and of course the position of the goggle on the head. Determination of the relation is referred to as calibration. The calibration procedure consist of having the subject maintain fixation on a known point while making slow head movements. This produces a set of poses of the eye and head, all of which correspond to

a single target in the world. This procedure is simplest when the target is visible in images from the head-mounted scene camera. In this case, we identify (either manually or automatically) the position of the target in each frame. Because the head-mounted scene camera is located as close as possible to the eyes, there is an approximate correspondence between the scene camera pixels and the head-relative gaze directions. We can then solve for regression coefficients relating the eye measurements to the scene camera locations. For small-to-moderate excursions away from primary position, a linear mapping is generally adequate, while the use of higher-order polynomial terms provide a good approximation for larger gaze angles. An alternative approach[6,7] is to use an optical model of the eye and camera system, but it is not clear that the added complexity produces an increase in accuracy, and we feel that our method has the advantage of automatically compensating for any abnormalities in the subject's anatomy.

Once the target of gaze has been localized in the scene camera image, all that remains is to register the elements of the image with our model of the world. This is easily obtained from the procedure described in section 4.5 above. The registration parameters tell us how to map each pixel into the cockpit mosaic, and in the mosaic image we can define fixed regions corresponding to each instrument, the windows, and so on.

6. SUMMARY

We have described a number of image processing procedures which have been applied to video data collected in the 2003 Tullahoma data collection flights. Our most reliable data has been obtained from the face-camera-based head pose estimation, with estimates obtained for approximately 85% of all frames, while the least reliable has been the day flight eye camera measurements, with estimates obtained for only 40% of all frames. We hope to improve the reliability and accuracy of all measures in the coming year.

Preliminary examination of the flight tracks reveals significant performance differences between night and day flights, demonstrating that subjects utilize visual terrain features to augment navigation guidance provided by the GPS receiver.

ACKNOWLEDGMENTS

REFERENCES

1. *National Transportation Safety Board Accident Database*, http://www.ntsb.gov/ntsb/query.asp.
2. R. P. Darken, J. A. Sullivan, and J. B. Mulligan, "Progress on the simulator and eye-tracker for assessment of PVFR routes and SNI operations for rotorcraft," in *FAA FY03 program review, Human Factors Vertical Flight*, http://www.hf.faa.gov/docs/508/docs/VF - SNI PVFR Darken.pdf, 2003.
3. S. M. Hickok and E. D. McConkey, "Flight test plan to assess of [sic] PVFR routes and SNI operations for rotorcraft," in *FAA FY03 program review, Human Factors Vertical Flight*, http://www.hf.faa.gov/docs/508/docs/VF - SNI PVFR STI.pdf, 2003.
4. M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience* **3**, pp. 71–86, 1991.
5. J. P. Chandler, "Subroutine stepit - finds local minima of a smooth function of several parameters," *Behavioral Science* **14**, pp. 81–82, 1969.
6. D. Beymer and M. Flickner, "Eye gaze tracking using an active stereo head," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. II–451 – II–458, 2003.
7. T. Ohno, N. Mukawa, and A. Yoshikawa, "Freegaze: A gaze tracking system for everyday gaze interaction," in *Proc. Eye Tracking Research and Applications Symposium*, pp. 125–132, 2002.