

Improving Aviation Safety with Information Visualization: A Flight Simulation Study

Cecilia R. Aragon

University of California, Berkeley
and NASA Ames Research Center
Berkeley, CA 94720-1776
aragon@cs.berkeley.edu

Marti A. Hearst

SIMS
University of California, Berkeley
Berkeley, CA 94720-4600
hearst@sims.berkeley.edu

ABSTRACT

Many aircraft accidents each year are caused by encounters with invisible airflow hazards. Recent advances in aviation sensor technology offer the potential for aircraft-based sensors that can gather large amounts of airflow velocity data in real-time. With this influx of data comes the need to study how best to present it to the pilot – a cognitively overloaded user focused on a primary task other than that of information visualization.

In this paper, we present the results of a usability study of an airflow hazard visualization system that significantly reduced the crash rate among experienced helicopter pilots flying a high fidelity, aerodynamically realistic fixed-base rotorcraft flight simulator into hazardous conditions.

We focus on one particular aviation application, but the results may be relevant to user interfaces in other operationally stressful environments.

Author Keywords

Information visualization, flight-deck displays, human factors in aviation, augmented reality, airflow hazards, presentation of safety-critical information, helicopters.

ACM Classification Keywords

H.5.1. [Multimedia Information Systems]: Artificial, augmented, and virtual realities; H.5.2 [User Interfaces]: User-centered design.

INTRODUCTION

Much research on information visualization has focused on office environments, where it is assumed that the user's attention will be directed exclusively to the visualization interface[30]. However, an area ripe for study is the use of information visualization in safety-related applications, in environments in which the user's primary task is something

other than looking at the computer interface (e.g., emergency response, air traffic control, operating any motor vehicle). For example, there is interest in improving driving safety of automobiles by projecting sensor data, such as heat readings indicating pedestrians or animals ahead of the car on a foggy day, onto the driver's windshield. It is an open question how such sensor data should be presented to be simultaneously useful and safe.

In this work we describe the results of a study in which information visualization of airflow hazards, when presented to helicopter pilots in a highly realistic simulator, dramatically improved their ability to land safely under turbulent conditions. In this case, we find that the kind of visualization needed to improve operational safety is much simpler than that needed for analysis of such hazards.

Below we begin by describing the flight safety problem and a potential solution in the form of new sensor technology. We then discuss previous research relevant to developing a visual hazard display to solve the airflow hazard problem. Next we describe our process of user-centered design and our experimental procedure, discuss the results obtained, and finally give conclusions and directions for further work.

BACKGROUND

Turbulence and other wind-related conditions were implicated in nearly 10% of the over 21,000 aircraft accidents in the U.S. National Transportation Safety Board accident database from 1989-99 [9]. Airflow hazards occurring near the ground can be deadly even to airliners; there have been hundreds of fatalities in the United States in the last two decades attributable to airliner encounters with microbursts and low level wind shear alone [41]. (Microbursts are small, very intense downdrafts that descend to the ground, often associated with thunderstorms [Figure 1], and low level wind shear is defined as a sudden change in wind direction and speed occurring near the surface [7].)

Airflow hazards are challenging to detect simply because air is invisible. Pilots cannot discern airflow patterns unless the air happens to pick up dust, smoke or other aerosols that are visible to the human eye. Being thus unable to detect a

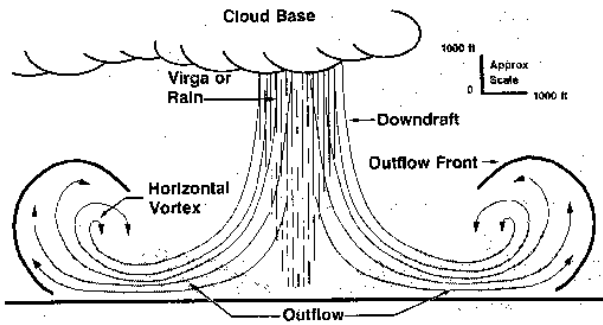


Figure 1. Microburst diagram

factor of potentially great importance to them, pilots learn to use their intuition concerning airflow over obstacles near their takeoff or landing sites, and they learn to pick up visual cues from the surrounding area. However, airflow-related accidents still occur each year. Addressing this issue could be of major benefit to aviation safety.

Focus on Helicopter Pilots

Although the risk of airflow hazards exists for all pilots in all aircraft, for our research we chose to focus on helicopter operations, and specifically on helicopter landings on moving ships. There were several reasons for this choice.

Helicopters are especially vulnerable to airflow disturbances such as vortices, downdrafts, and turbulence from surrounding vegetation or structures [Figure 2]; first, by the nature of the aerodynamic forces involved, and second, because helicopters are often called upon to operate into and out of confined areas or areas that naturally have disturbed airflow. For example, emergency search and rescue may have to operate in mountainous areas and small clearings surrounded by vegetation and cliffs where the winds are frequently high. Helicopters also must land on urban rooftops, offshore oil platforms, or on the decks of ships. A device for detecting airflow hazards therefore has a special utility for helicopter operations.

Operating a helicopter off a moving aircraft carrier is one of the most demanding tasks a helicopter pilot can face [43]. Because the ship is moving, its superstructure will always generate disturbed airflow such as vortices and turbulence. In addition, high seas may cause extreme ship motion, and low visibility may degrade visual cues. The pilot must maneuver the helicopter within very tight tolerances to avoid striking ship structures or other aircraft. It is a task that demands the utmost concentration and skill from the pilot. A system that can deliver even an incremental amount of assistance to the pilot in this high-demand environment could have a significant impact on safety.

Helicopter accidents and incidents that occur on shipboard each year range from incidents such as “tunnel strikes” (where certain wind conditions can cause a helicopter’s rotor blades to spin out of control, damaging the fuselage of the helicopter) to fatal accidents. There have been over 120 tunnel strikes since the 1960s, causing damage ranging

from \$50-\$75K to over \$1M per incident [20]. Analysis of these accidents and incidents frequently finds them to have been caused by unseen airflow hazards where the pilot and ground crew were initially unaware of the danger and the pilot was unable to react in time [9]. Presenting the appropriate information to the pilot or flight deck air boss (shipboard air traffic controller) in advance of the hazard encounter, therefore, could reduce or prevent such accidents.

Finally, because shipboard rotorcraft operations are such a demanding environment, the area is very well studied. The Navy has compiled significant amounts of data from shipboard flight tests, wind tunnel tests, and computational fluid dynamics computations studying the airflow around moving ships of all types, and how the airwake changes when helicopters of different makes and models land on the ships. The available data is thus sufficient to support a study on how better to present that data to the pilot.

New Sensor Technology

New advances in sensor technology such as Doppler lidar [22] and other techniques are leading to the development of aircraft-based sensors which can collect large amounts of airflow velocity data in real-time. Within a few years, it is likely that aircraft-mounted hardware will be available that can reliably scan the area a few hundred feet ahead of the aircraft and sample air particle vector velocities at one-foot intervals or less [14]. With the development of such devices, onboard detection systems that can convey detailed, specific information about airflow hazards to pilots in real-time become a possibility. Thus, an interface is required that can present large amounts of data to the pilot in a comprehensive manner in real-time, yet not distract from the pilot’s primary task of flying the aircraft. This is the information visualization task we attempt to address in this paper: how does one best present safety-critical information to a cognitively overloaded user in real-time?

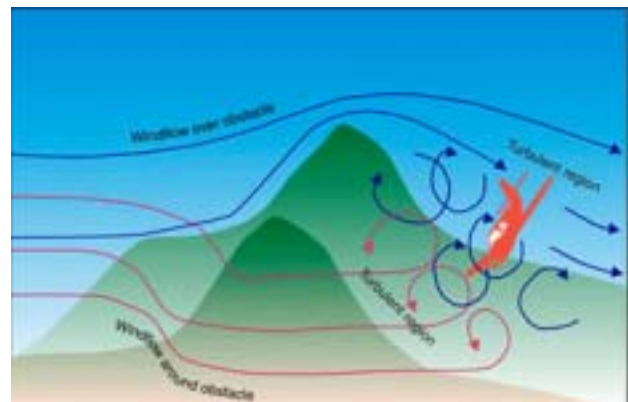


Figure 2. Turbulent flow and vortex formation on leeward side of obstacles

RELATED WORK

There are several significant bodies of research relevant to our current efforts. These include flow visualization, human factors in aviation, augmented reality displays, and Navy shipboard rotorcraft operations.

Flow visualization

Flow visualization systems often consist of detailed imagery of two- and three-dimensional airflow patterns, both static and dynamic, steady and unsteady, all designed to help scientists or engineers understand — and analyze at length — a particular instance of a fluid flow. Examples include streamlines and contour lines for the case of instantaneous flow [6, 33] and streaklines, timelines [15], flow volumes [18] and spot noise [26] for unsteady flow, and terrain and turbulence visualization [37].

We initially investigated the use of flow visualization similar to one of these, with the idea that perhaps smoke trails in the air, or dust devils, could be shown to the pilots. However, once we started applying HCI techniques to the problem, it quickly became apparent that such visualizations would not work in the type of situations we were studying. The imagery is often quite complex and not suitable for rapid glances during time-critical tasks.

Human factors in aviation

It has long been recognized that applying developing technology to improve aviation displays might enhance aviation safety. Significant work in this area includes synthetic vision and augmented-reality displays (terrain in low-visibility environments, navigation aids) [2, 4, 32, 38], weather visualization including NASA's AWIN, TPAWS and AWE [5, 16, 29, 31], and turbulence detection and prediction [36]. Holforty [12] developed a system for wake vortex prediction and display. (As an aircraft passes through the air, it leaves a trail called wake vortices that can be hazardous to following aircraft.) The system provided wake vortex visualization to pilots on a head-down display, during the enroute (least demanding) phase of flight. The focus of this work was on the prediction of the location of the wake, not on the usability of the display, but it was the first study that attempted to display a three-dimensional visualization of any type of airflow hazard to pilots.

There is a large body of relevant work concerning human factors in the cockpit, including the study of attention and cockpit visual displays [8, 17, 28, 42]. Head-up displays (HUDs) provide flight information and guidance to the pilot on a forward field-of-view transparent screen. They have been well studied for use in aviation since they were first developed in the 1950s [27]. The analyses of how attention is divided in the cockpit and the HUD usability studies informed our system design, although none of the work was specifically directed to optimizing airflow hazard display.

It has been known since the 1980s that weather-related airflow phenomena such as microbursts and wind shear have been responsible for airliner accidents [39]. As a

result, a great deal of work has been done to detect, predict, and display this type of information to the pilot [41]. There are commercially available aircraft-based, forward-looking microwave radar and lidar systems that can detect microbursts and wind shear. However, rather than designing new displays to optimally present the new data, the emphasis was placed on integrating the information into existing cockpit displays to reduce time to commercial deployment. Accordingly, no usability studies were focused strictly on the display itself or on whether a three-dimensional head-up display would be helpful in presenting hazard information to the pilot.

Augmented reality displays

Augmented reality has been variously defined as “augmenting natural feedback to the operator with simulated cues” or “a form of virtual reality where the participant's head-mounted display is transparent, allowing a clear view of the real world [21]. HUD technology, as used in military aviation displays, is a relatively mature use of this technology; more recently, a growing body of research has focused on wearable systems, such as those which can be worn by an individual operating in a city. When computer-generated imagery is overlaid on real objects, the technical challenges involved, including information filtering, spatial registration, and how to combine visible and obscured information, are considerable [10, 13]. Much of this research focuses on non-aviation areas; however, developments in this field could benefit and improve the quality of our display.

Navy shipboard rotorcraft operations

Because landing a helicopter on a moving ship deck is hazardous [43], the Navy has long operated a program to perform flight testing in this environment [44] with the stated goal of improving flight safety. In addition to airwake from ship superstructures and the requirement to land in a confined area, aircraft landing on shipboard are plagued by hot exhaust plumes, very powerful shipboard radar that interferes with aircraft systems, inaccurate anemometers, and problems associated with high sea states such as strong, turbulent winds and extreme values of ship pitch, heave, and roll.

For understanding the airwake over the ship, the Navy uses techniques including shipboard flight testing, wind tunnel tests, computational fluid dynamics (CFD) modeling, and sampling of the airflow vector velocities at various points in the flow field behind the superstructure in the helicopter landing zones.

The current method of communicating this information to the pilots consists of publishing pre-computed operational envelopes listing allowable wind conditions for many ship-rotorcraft combinations [44]. The envelope conveys a go/no-go decision, and does not state which safety considerations motivate a given operational limit. Pilots check the published envelope for their helicopter before beginning any approach, and they only fly the approach if

they are within the envelope. This procedure has the advantage of providing clear, simple direction to the pilots under all wind conditions. However, this means that if the winds shift out of the envelope during the approach, or some other event occurs that changes the airflow over the landing site, such as a helicopter on an upwind spot starting up its rotor, a hazardous condition can occur of which the pilot is unaware. This type of situation has been demonstrated to be a causal factor in many incidents [23].

DESIGN OF THE VISUAL HAZARD INDICATOR SYSTEM

In keeping with a user-centered design process, we began by creating a low-fidelity prototype and performing a usability test with domain experts to validate the idea and to refine design choices for visual indicators.

Low-Fidelity Prototype

To construct a low-fidelity prototype, we used Rhino3D, a CAD modeling tool that supports rapid prototyping of 3D objects and simulated animation of the helicopter pilot's view out the cockpit windscreen. A wide selection of different types of hazard indicators were stored in layers in Rhino3D, so that features such as shape, color, texture, transparency, depth cueing, and motion could be selectively turned on and off by the operator. We recruited three highly experienced (>1700 hours) helicopter pilots and flight test engineers as domain experts to evaluate the prototype. (Details of the usability study and its results can be found in [3].) The domain experts were in agreement about the potential value of an airflow hazard visualization system and gave clear indications of the best design choices in producing the hazard visual cues.

This phase of the study showed a strong preference on the part of the pilots for a hazard visualization system in which the hazard indicator appears in the physical scene. During potentially hazardous conditions, the pilot's attention will naturally be focused outside the cockpit during the critical landing moments; he or she will not want to look down at a cockpit instrument display. The pilots strongly preferred an augmented-reality hazard visualization display on a HUD. However, the display must be carefully designed not to distract from the key shipboard visual cues, even when these are degraded during a challenging nighttime or poor-weather landing.

Pilots preferred much simpler imagery than we had initially expected. An evident guiding principle was that helicopter pilots landing on shipboard must focus all their attention to accomplish the landing on a moving ship, and have little spare attention for detailed quantitative information about the hazard. Extensive detail, motion (animation), complex shapes, and too many colors were all stated to be distracting and possibly dangerous in the high-demand environment. The hazard indicators had to be sufficiently transparent so as not to obscure any critical shipboard visual cues that the pilots needed to use as landing aids. The pilots wanted the minimum critical information such as the location of the hazard and its severity as a warning (yellow) or danger (red).

In other words, our domain experts had informed us that they wanted a decision support system, not a scientific visualization system, and the reason had to do with the division of attention in the high-demand environment.

Pilots emphasized the importance of using standard symbology at all times. They warned of the danger a moment of confusion could cause, and strongly recommended that the symbology used in our head-up display conform to current aviation conventions; it was especially important that our symbols not have any chance of being confounded with other types of HUD symbology already in use. Although not completely standardized, current symbology includes items like airspeed and altitude tapes, aircraft reference symbol, flight director, and roll scale pointer [Figure 3]. The results from this low-fidelity prototype study helped us to select a design that was significantly different from any type of HUD symbology.

Design and Implementation Choices

Based on the results from the study of the low-fidelity prototype, we selected a simple, static design for the hazard indicators and used only two colors, yellow (caution) and red (danger). The shape and appearance of the indicators were chosen to indicate the physical location of the hazard

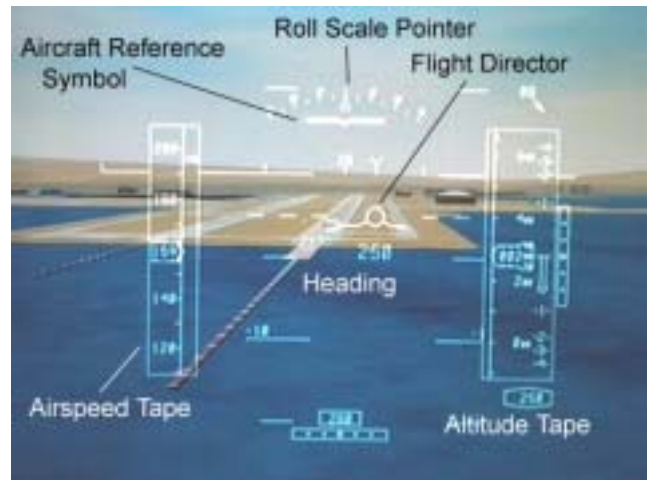


Figure 3. An example of HUD symbology

without undue distraction and without duplicating any symbology used for other purposes, while the color meanings are conventional and widely accepted in the aviation world.

Studies have shown that head-up displays with superimposed symbology may on occasion cause performance problems due to attentional capture by the perceptual grouping of the superimposed symbols [19, 27]. “Scene-linked” head-up displays, or displays where there is no differential motion between the superimposed symbology and the outside scene, can avoid this type of distraction. This study also confirmed the need to develop a head-up display where the hazard indicator is three-dimensional and appears to be physically part of the world.

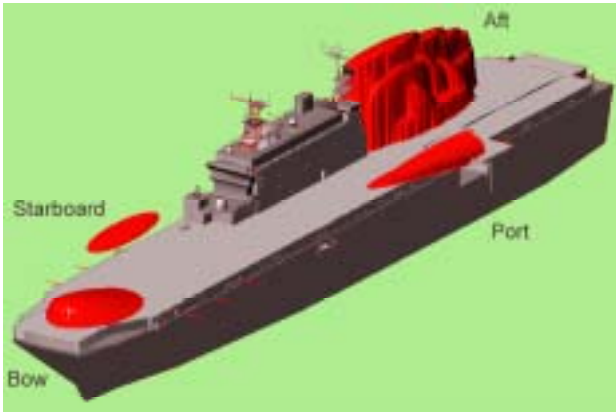


Figure 4. Visual hazard indicators used in the study

We created three-dimensional, translucent red and yellow surfaces that delineated the outer outlines of the areas of hazard on shipboard [Figure 4]. (Actual surfaces were more translucent than pictured in the figure.) The boundaries of the hazardous areas were determined upon extensive review of the archived airflow data from flight tests and consultation with a Navy flight test engineer. The degree of transparency of the hazard indicator objects was set at the level of transparency (around 70% in Rhino3D) preferred by pilots during their evaluation of the low-fidelity prototype. The objects were then imported into the visual subsystem of the simulator and linked to the ship so that they seemed to be part of the simulated outside world; they appeared as clouds or curtains hovering over particular locations on shipboard. This is an accurate model of shipboard airwake; any hazardous areas produced by wind blowing over ship structures will move along with the ship.

FLIGHT SIMULATION STUDY

In order to produce a high-quality usability study of a specialized interface, it is important to select participants who are domain experts. The quality and relevance of the results depend on getting people who actually fly under the demanding conditions that we hope to duplicate in this study. To test our hypothesis that the presence of a visual hazard indicator could improve helicopter flight safety, we recruited sixteen experienced helicopter pilots to participate in the flight simulation study.

The use of a flight simulator in any study immediately raises an issue of realism. Our study was designed for as high a level of technical fidelity as possible, and also to establish the correct mood for the testing. Using a simulator is, of course, not as stressful as the actual flight situation. However, during our pre-flight briefing, we made a special effort to ensure the pilots would use the same judgments they would in the real world. “If you feel the controllability of the aircraft is in question, follow the same safety procedures as you would in the real world.” Although it is impossible to verify whether this proscription was followed absolutely, comments gathered from the pilots during the simulation and observation of the pilots’ behavior during the simulation (the intensity of their gaze,

grip on the controls, sweating, breathing levels, etc.) indicated that they were taking it seriously and not thinking of it as, for example, a video game. Additionally, pilots are generally quite conscious of the fact that lives depend on their proficiency and decision-making during the critical moments of a flight, and take pride in their skills and their ability to consciously marshal their skills even under moments of extreme duress. Although we were clear that the purpose of the study was to test the hazard display system rather than the pilot, the pilots’ awareness that the test was being observed would reasonably be expected to stimulate that pride in their skills. For these reasons, we believe the results of our simulation fairly accurately reflect results that would have been achieved in the real world.

We chose a high fidelity, realistic helicopter flight simulator with accurate aerodynamic models, which we then provided with actual airflow data from shipboard flight tests. The pilots sat in an aircraft seat with full helicopter controls (cyclic, collective, and tail rotor pedals) with force feedback, in front of a cockpit instrument panel, and viewed visuals on three large projection screens [Figure 5]. The pilots flew simulated final approaches to land a Sikorsky H-60 helicopter on a moving ship (an LHA or “Tarawa-class” Navy amphibious assault ship) under different wind conditions, some of which entailed airflow hazards such as vortices, downdrafts, or turbulence on or near the landing site. Four different landing difficulty levels were used. Other than the control approaches (no hazard present), each approach was flown twice by each pilot, once with a hazard



Figure 5. ART flight simulator with pilot in front of projection screen and operator at rear console

indicator present and once without. Data was gathered both objectively from the flight simulator’s recording capability and subjectively from a Likert-scale questionnaire administered to the pilots after the test.

Procedure and Design

The study was a 3 (landing difficulty) x 2 (presence or absence of visual hazard indicator) x 4 (approach type) + 1 x 1 x 4 (control) within-subjects design. Each pilot flew the same 28 simulated approaches, but in different orders. Four different approach scenarios were selected where winds could create a hazard to helicopters landing on the deck of a

Navy ship. A flight test engineer with 17 years of experience with Navy shipboard helicopter flight testing assisted us in designing the four scenarios, selecting various wind speeds and turbulence levels to create approaches with different landing difficulty levels, and determining where hazardous airflow conditions would exist. We then recruited a second experienced Navy helicopter test pilot who flew all the approaches in the simulator and evaluated the correctness of the landing difficulty level and the correct placement of the hazard indicators.

After a pre-flight briefing that explained the structure of the simulation and the use of the controls of the simulator and instructions as to the meaning of the yellow and red hazard indicators, participants performed a series of orientation flights before beginning the actual test. The dual purposes of the orientation flights were to accustom them to the feel of the controls of the simulator, and to determine if they had the skill level to be a credible participant in our experiment. Out of 17 pilots recruited for the study, one was unable to fly the orientation flights and was excused, leaving 16 pilots who then completed the test approaches.

At the outset of each approach, pilots were given wind direction but not wind speed. Revealing wind speed could introduce bias due to the pilots' assumption that wind speed correlates with landing difficulty level, although pilots were briefed that hazards could occur even at low wind speeds.

Approach Description

For each approach or run, the simulator was set to a previously saved checkpoint that positioned the helicopter at 250 feet above mean sea level and 2600 feet back of the stern of the ship. Wind and turbulence conditions that would produce a landing of difficulty 1-4 were pre-programmed into the simulator, and the appropriate hazard indicators were turned on at the beginning of the approach if one was supposed to be present. The simulator flight controls were trimmed to a 30-knot airspeed, and the pilots were given a verbal clearance to land on one of four landing spots and the wind direction. The pilots were asked if they were ready, and then the simulator was set running. Pilots flew until the landing was complete, they verbally called out an aborted approach, or they crashed. Then the simulator was stopped and set up for the next run. Pilots were encouraged to make verbal comments during the test, and the entire test was videotaped for all pilots.

Approach Scenarios

Scenarios were labeled based on which landing spot the pilot would be cleared for and where the airflow hazard would occur under certain wind conditions [Figure 4].

Scenario A (“Aft”): Direct stern approach to landing spot 9, the aft-most landing spot on the LHA. With a direct bow wind, and at high wind speed and turbulence levels, an airflow hazard would occur downwind of the ship superstructure over landing spot 9.

Scenario B (“Bow”): A 45-degree approach to the most forward spot on the bow of the ship, spot 1, and winds directly from the bow. This created an area of heavy downdraft (“suckdown”) directly over spot 1, which was often unexpected as it occurred even at relatively low winds and even in smooth wind conditions.

Scenario P (“Port”): A 45-degree approach to the port side of the ship, to landing spot 7, just forward of the elevator and next to the ship superstructure. Winds from 300 degrees (assuming the ship is moving toward the north or 360 degrees) caused a rotor to form over the deck edge just over landing spot 7. Again, this hazard formed even at relatively low winds.

Scenario S (“Starboard”): A 45-degree approach from starboard to landing spot 3A just forward of the ship superstructure. When winds are from 60 degrees, a vortex forms just at the deck edge and beside landing spot 3A.

Landing Difficulty Level

We used four different landing difficulty levels based on the Navy's Pilot Rating Scale of landing difficulty [44]. Each pilot flew each approach scenario at all landing difficulty levels. For each of LD 2 through 4, each pilot flew one approach with and one without a visual hazard indicator. For LD 1, each pilot flew one approach without a hazard indicator. Thus, each pilot flew 7 approaches in each of the 4 landing scenarios, a total of 28 approaches each. The approaches were designed to take about 1-2 minutes each; thus, the entire simulation took about 1 hour per pilot; this time was intentional to prevent pilot fatigue.

Landing difficulty 1 (LD 1) – Control. These approaches showed how well the pilot could operate the simulator in the absence of particular hazards, and also provided periods of rest to the pilots to reduce fatigue and avoid discouragement (since the test consisted of an abnormally high percentage of very challenging landing conditions).

Landing difficulty 2 (LD 2) – Testing for negative effects of the hazard indicator. This difficulty level required moderate pilot effort. The hazard indicator (if present) was a translucent yellow object outlining the area where turbulent flow could be found. Because the conditions at LD 2 are considered to be within normal pilot abilities, we would expect few crashes even without the hazard indicator. The hypothesis tested at LD 2 was that the hazard indicator would not *increase* the crash rate (e.g. by distracting the pilot). Pilots were instructed that the yellow hazard represented caution and that they could continue the approach.

Landing difficulty 3 (LD 3) – Testing for benefit of hazard indicator. This difficulty level required maximum pilot effort. The hazard indicator was the same type as for the LD 2 approaches. Pilots were told that yellow represented caution and they were to continue the approach. A higher

Order 1 (4 pilots)				Order 2 (4 pilots)			
A	P	B	S	A	P	B	S
2-	4-	2H	4H	2H	3H	2-	3-
3-	2-	4H	3H	3H	4H	4-	2-
4-	3-	3H	2H	4H	2H	3-	4-
4H	2H	3-	4-	4-	3-	3H	2H
3H	4H	4-	2-	3-	2-	4H	3H
2H	3H	2-	3-	2-	4-	2H	4H
Order 3 (4 pilots)				Order 4 (4 pilots)			
A	P	B	S	A	P	B	S
4H	2H	3-	4-	4-	3-	3H	2H
3H	4H	4-	2-	3-	2-	4H	3H
2H	3H	2-	3-	2-	4-	2H	4H
2-	4-	2H	4H	2H	3H	2-	3-
3-	2-	4H	3H	3H	4H	4-	2-
4-	3-	3H	2H	4H	2H	3-	4-

Table 1. Simulated Approach Orders

crash rate was expected at LD 3 commensurate with the more challenging conditions compared with LD 2. We hypothesized that the hazard indicator would reduce this crash rate – ideally, to a rate comparable to LD 2.

Landing difficulty 4 (LD 4) – Testing for benefit of hazard indicator with pilot instructional procedure. At LD 4, safe landings were not probable. Fifteen pilots were told that if they detected a red hazard indicator along their approach path, standard operating procedure (SOP) was to abort the landing immediately. (The 16th pilot, who was not initially given this instruction, spontaneously proposed that it should be standard operating procedure.) These approaches test whether the same hazard indication methodology used for reducing the crash rate in marginal conditions will also operate reasonably in extreme conditions.

Order of Presentation

To compensate for possible learning effects, half the pilots flew scenarios A and P without the hazard indicators and scenarios B and S with the hazard indicators during the first half of the test, and then conversely for the second half. The other pilots flew scenarios A and P with hazard indicators and scenarios B and S without indicators during the first half of the test. This was accomplished by defining an approach order randomly within these constraints, then reversing it to create a second order, then switching the first and second halves to create a third and fourth order. It was chosen so that the most difficult approaches would not all follow one another, to reduce the likelihood of pilot fatigue. Table 1 lists the approach orders. Rows indicate order of presentation. Within the rows, the order is randomized. In the cells, numbers represent landing difficulty, followed by presence (H) or absence (-) of hazard indicator, e.g. “3-“ in

column “B” indicates an approach to the Bow spot at LD 3 and no hazard; “2H” in column “A” indicates a run to the Aft spot at LD 2 with hazard indicator present. LD 1 (control) runs were scattered randomly through the series so each pilot flew 28 runs.

Dependent Variables

During the simulation, 50 variables such as velocity and position of aircraft in x, y, z, control stick position both lateral and longitudinal, collective and pedal positions, landing gear forces, etc., were collected by the flight simulator at 10 Hz and stored in data files labeled for each run and pilot. However, our primary dependent measure was the crash rate. A “crash” was defined as an impact with the ship’s deck with a vertical velocity of 12 feet per second or greater as measured by the simulator. In order to be certified for shipboard use in the US Navy, rotorcraft must be able to withstand an impact of 12 fps at touchdown [40].

We also gathered subjective pilot opinions from a 21-probe Likert-scale (1-5) questionnaire administered to the pilots at the end of the simulation. For each probe, the pilots had to circle one of “Strongly Disagree” (1), “Disagree” (2), “Neither Agree Nor Disagree” (3), “Agree” (4), and “Strongly Agree” (5).

Hypotheses

We tested four hypotheses:

1. Crash rate will be reduced by the presence of hazard indicator (LD 3).
2. Crashes will be eliminated by red hazard indicator if a standard operating procedure is given to pilots (LD 4).
3. Hazard indicator won’t cause distraction or degradation in performance in situations where adequate performance is expected without indicator (LD 2).
4. Pilots will say they would use the visualization system.

Participants

We recruited 17 military and civilian helicopter pilots by word-of-mouth and through emailed requests for volunteers. Sixteen pilots flew the orientation flights successfully and completed the simulation test. This group of pilots had no previous experience on the simulator used in the experiment and had not seen or heard of any type of visual hazard indicating system before. Pilot experience ranged from 200 to 7300 helicopter flight hours with the median number of hours being 2250, from 2 to 46 years of experience as a helicopter pilot with the median 13 years, and were from 25 to 65 years old, with a median age of 36. All pilots had normal or corrected-to-normal eyesight and were not color-blind. The study took about two hours, of which about one hour was spent in the simulator, and pilots were not paid for their participation.

Equipment

The study was performed at Advanced Rotorcraft Technology, Inc. in Mountain View, California, a small flight simulation company specializing in rotorcraft non-linear dynamics modeling. ART’s aerodynamic models

have been verified by the US Navy via stability and control techniques and frequency domain validation [11,34], and Navy flight test engineers and pilots have stated that they are more aerodynamically accurate than other rotorcraft flight simulators currently available [34].

The study was performed in a high fidelity helicopter flight dynamics simulator with a single seat configuration, flight controls with force feedback, instrument panel, and a three-channel projection outside world visual system utilizing 3D Perception projectors to provide 1024 x 768 resolution at 1000 ANSI lumens. Visual rendering is done using ART software that supports rendering on OpenGL graphics cards using OpenFlight format visual databases. Image generation is done on PCs with graphic acceleration hardware that provides a 60 Hz update rate with full-screen anti-aliasing and a 188" horizontal by 54" vertical field of view on a 6.5-ft radius cylindrical screen.

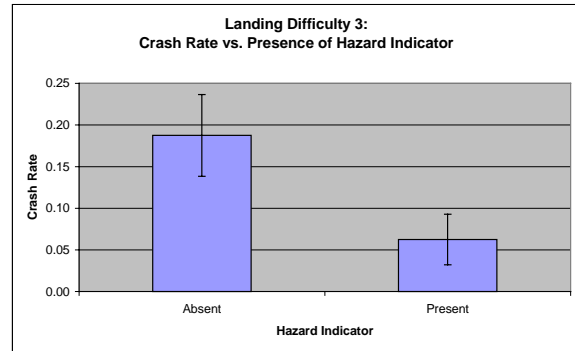
An operator console provides full simulator control, monitoring of the visual system and instrumentation displays, initialization to saved reset points and arbitrary test conditions. Control loaders for the pilot's controls are electric and are driven by software that interfaces the flight dynamics model to the control loaders and edits the force feel characteristics. Four sets of control loaders are used to drive the longitudinal cyclic, lateral cyclic, collective and pedal controls. Computer generated images of the instrument panel are rendered. A dual 1.9GHz AMD processor computer with two graphics boards, located in the operator console, is used to drive a flat panel display that is mounted behind instrument panel overlays.

RESULTS

Choice of tests for statistical analysis. The statistical test we applied to most of our crash-rate hypotheses was the "paired two-sample t-test for means." The study does not have the multiple categories or varying populations that would make a t-test inappropriate and require the more general ANOVA procedure. The test data has been structured so that the assumptions required for a t-test are valid. Additionally, within each landing difficulty, the data is so structured that it is possible to pair the samples where in each pair, all variables are held constant except the presence or absence of the hazard indicator. This paired design allows the application of a paired-sample t-test. Conversely, ANOVA is inappropriate for a paired design, because it assumes independence of the sample populations.

Hypothesis 1 confirmed. For the test at landing difficulty 3, there were 12 crashes out of 64 approaches without the hazard indicator (crash rate .19, standard error .049) and 4 crashes out of 64 with the hazard indicator (crash rate .063, standard error .030) [Table 2]. A t-test for paired samples shows that the hypothesis that the presence of the hazard indicator reduces the frequency of crashes during simulated shipboard helicopter landings is confirmed ($t=2.39$, $df=63$, $p<0.00985$).

Table 2. Landing Difficulty 3 - Crash Data



These strong results indicate the system should improve helicopter flight safety under hazardous conditions. During the tests, pilots remarked several times that the indicators were helpful warnings; that they were able to modify their flight path or power settings to counteract the known hazardous conditions, or make appropriate safety decisions based on knowledge gained from viewing the hazard indicators. Additionally, in the approaches without hazard indicators, pilots commented on several occasions that they were surprised by the wind conditions as they entered the hazardous areas. In a few of these runs where the pilot made such a comment, the approach terminated in a crash.

Hypothesis 2 confirmed. At landing difficulty 4 (beyond the capacity of the aircraft), there were 0 crashes in 64 approaches with the hazard indicator as opposed to 15 crashes out of 64 without the indicator, for crash rates of 0% and 23% respectively. (Standard errors were 0 and .053.) A t-test for paired samples shows that this hypothesis—that the presence of the red hazard indicator combined with appropriate instructions to the pilot prevents crashes—is strongly confirmed ($t=4.39$, $df=63$, $p<0.00022$). What this means is that although pilots may sometimes continue into a situation that is beyond the capacity of the aircraft if they do not have sufficient knowledge of the danger of the situation, giving them the appropriate information in a clear and simple manner during the approach can prevent accidents. This is an improvement over the current envelope system because, as one pilot noted, it would be very helpful in case the winds shifted during the approach. If he suddenly saw a red hazard area appear on deck, he would know immediately to abort the approach.

For pilot 16, we experimented with not giving the pilot the standard operating procedure (SOP) of a mandatory go-around upon detection of the red indicator. This pilot continued each approach with a red hazard indicator until he got close to the red zone, then he aborted the approach. In other words, he took the same actions the other pilots had, just a little later during the approach. Interestingly, during the post-flight debrief, this pilot stated that with the red hazard we should have given pilots an SOP of an automatic abort upon detection. Otherwise, he said, pilots might be tempted to go on and "test the waters." Although this variation with the last pilot could not produce any

statistically valid results due to the small sample size, it suggested that our results might have been the same had we not had an SOP for the red hazards for the first 15 pilots.

Hypothesis 3. No negative effect of hazard indicator. It appears that the hazard indicators did not distract the pilots. The crash rate at LD 2 was the same with and without the indicator. Crash rate for both was identical, 7.8% or 5 crashes out of 64 for each set of approaches. (Standard error was .034.) However, because the crash rate was low, with a sample of this size it is not possible to conclusively state that the hazard indicator made no difference in crash rate. On the other hand, the pilots did not feel the hazard indicators were distracting. On our simulation evaluation questionnaire, probe 6 was, “The airflow hazard visualization distracted me from the task of flying the aircraft.” The pilots disagreed with this statement: 94% of the pilots answered “Strongly Disagree” (1) or “Disagree” (2) with the median “Disagree” (2).

Hypothesis 4 confirmed. When pilots were asked to report their level of agreement with the statement, “I would use this system if it were available on my aircraft,” eight pilots chose “Strongly Agree” (5), five chose “Agree” (4), one chose “Neither Agree Nor Disagree” (3) and two chose “Disagree” (2). Median response was 4.5, between “Strongly Agree” and “Agree.” This indicates confirmation of Hypothesis 4, that pilots would use the system. Pilot buy-in to new systems is very important as pilots have been known to ignore safety mechanisms if they feel they are “stupid” or intrusive, and override or turn off systems that they feel interfere with their handling of the aircraft [24].

However, although it was the first time the pilots were flying the simulator and the first time they had seen a visual hazard warning system, they were quite enthusiastic about the system. Several of the pilots wanted to know how far the system was from implementation in aircraft. Even two of the three pilots who did not agree with wanting to use the system were not completely negative; of the three, one said he needed more time with the system before he could make a decision, one pilot liked the idea for less experienced pilots but felt he already knew where all the hazards were. Only one felt his HUD was already too visually cluttered and he did not want anything else displayed on it; he preferred an auditory warning. (HUDs have display modes so pilots can select the amount of information displayed—such modes are called “clutter modes.”) Given how resistant pilots can be to externally imposed changes, that 13 out of 16 pilots say they would use the system after one hour with it is a very strong positive result.

Control group (LD 1). Because conditions in the simulator are somewhat different than in a real helicopter, and visual and proprioceptive feedback is reduced (no chin bubble through which helicopter pilots can look down past their feet and see how close they are to the deck, no depth perception in the visuals, no bump when the landing gear contacts the deck, etc.), and especially because pilots are

flying it for the first time without any training with an instructor (the usual procedure when transitioning to a new aircraft), a certain number of crashes in the simulator are to be expected. For this reason we included a set of low-hazard approaches in the study to serve as a control (LD 1).

The crash rate at landing difficulty 1 was 9.4% (6 out of 64, standard error .037), which is not significantly different from LD 2 or LD 3’s crash rates (5 out of 64, std. err. .034 and 4 out of 64, std. err. .030, respectively; t-test, $p=0.38$ and $p=0.26$) when the hazard indicator is present.

Learning effects. For the first half of the simulator test, the pilots crashed 25 times out of 224 approaches flown for a crash rate of 11.2%, while in the second half of their tests, the pilots crashed 22 times out of 224 approaches, for a crash rate of 9.8%. This is not a significant difference (t-test, $t=0.46$, $df=445$, $p=0.32$), although the pilots did state that they believed they performed better as they flew the simulator longer. (Probe 17: “It became easier over time to fly because my experience on the simulator increased.” Eight pilots answered “Strongly Agree” (5), six pilots chose “Agree” (4), and two pilots chose “Neither Agree or Disagree” (3). Median response was 4.5.) This appears to indicate that learning effects did not bias our study, as was intended in its construction.

CONCLUSIONS AND FUTURE WORK

Results of a study where information visualization of airflow hazards was presented to helicopter pilots in a highly realistic simulator showed a significant improvement in their ability to land safely under turbulent conditions when supplied with the visualization interface. In this experiment, we discovered that the type of visualization needed to improve operational safety was much simpler than that needed for analysis of airflow hazards, providing an example in which the appropriate visualization differs for analysis vs. presentation.

This study also validated the use of HCI techniques and user-centered design. By providing for usability testing early in the design and obtaining feedback from domain experts, we were able to avoid the potentially costly mistake of developing an overly elaborate interface based on existing flow visualization techniques which could have interfered with pilot operations.

The enthusiastic response we received from the pilots and the strong positive results from our simulation study indicate that such an airflow hazard visualization system could improve aviation safety, especially because it appears that pilots would actually use the system in the cockpit.

There are many opportunities for future research in this area. Several pilots expressed a desire for an adaptive display—one that presented more detailed flow information at the start of the landing approach, but changed to the existing visualization as the approach progressed and pilot workload increased. An open question is how and when to change the display.

ACKNOWLEDGMENTS

This work was funded by the NASA Ames Full-Time Graduate Study Program. The research could not have taken place without the support of Kurtis R. Long of the Navy Dynamic Interface Flight Test Group, who generously shared his extensive knowledge. Heartfelt thanks are also due to Advanced Rotorcraft Technology, Inc., for the use of their high-fidelity simulator, as well as to the ART staff members who spent many hours assisting in the development of the system and simulation study setup.

REFERENCES

- Advanced Rotorcraft Technology, Inc., 1685 Plymouth St., Suite 250, Mountain View, CA, USA 94043, <http://www.flightlab.com>.
- Alter, K.W., Barrows, A.K., Jennings, C., Powell, J.D. 3-D Perspective Primary Flight Displays for Aircraft. *Proceedings of Human Factors and Ergonomics Society*, (2000).
- Aragon, C. Usability Evaluation of a Flight-Deck Airflow Hazard Visualization System. 23rd Digital Avionics Systems Conference (2004).
- Bailey, R.E., Parrish, R.V., Kramer, L.J., Harrah, S.D., Arthur, J.J. III. Technical Challenges In the Development of a NASA Synthetic Vision System Concept. NATO RTA (2003).
- Brita, C. L., Kelly, C. W. User's Guide for an Airborne Doppler Weather Radar Simulation (ADWRS), NASA TPAWS (2002).
- Buning, P. Numerical algorithms in CFD post-processing, Computer Graphics and Flow Visualization in Computational Fluid Dynamics, von Karman Institute for Fluid Dynamics Lecture Series 1989-07.
- Chambers, J. Concept to Reality, NASA SP-2003-4529, http://oea.larc.nasa.gov/PAIS/Concept2Reality/wind_shear.html.
- Endsley, M.R. Designing for Situation Awareness in Complex Systems, *Proceedings of the Second International Workshop on symbiosis of humans, artifacts and environment*, Kyoto, Japan (2001).
- FAA National Aviation Safety Analysis Center, <https://www.nasdac.faa.gov/>.
- Furmanski, C., Azuma, R. and Daily, M., Augmented-reality visualizations guided by cognition: Perceptual heuristics for combining visible and obscured information. *Proceedings of IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2002)*, 2002, 215-224.
- He, C. and Goericke, J., Army ART TR 2011 - DAAH10-03-C-0001, FLIGHTLAB UH-60L Simulation Model Validation, 2004.
- Holforty, W. Flight-Deck Display of Neighboring Aircraft Wake Vortices, Ph.D. dissertation, Stanford University (2003).
- Julier, S., Lanzagorta, M., Baillet, Y. and Brown, D. Information Filtering for Mobile Augmented Reality. *IEEE Computer Graphics and Applications*, 22 (5), 2002, 12-15.
- Kao, C.-Y. J., Cooper, D.I., Reisner, J.M., Eichinger, W.E., and Ghil, M. Probing Atmospheric Turbulence with High-Resolution Lidar and Models, *J. Geophys. Res.* (2002) 107(D9): 7 1-7.10.
- Lane, D.A. Visualizing Time-Varying Phenomena in Numerical Simulations of Unsteady Flows, NASA NAS Technical Report NAS-96-001 (1996).
- Latorella, K., Chamberlain, J. Graphical Weather Information System Evaluation: Usability, Perceived Utility, and Preferences from General Aviation Pilots, NASA AWIN technical report (2002).
- Lee, A.T. Aircrew Decision-Making Behavior in Hazardous Weather Avoidance, *Aviation, Space, and Environmental Medicine* (1991).
- Max, N., Becker, B., Crawfis, R. Flow Volumes for Interactive Vector Field Visualization, *Proceedings of Visualization 93*, IEEE Computer Society Press (1993), 19-23.
- McCann, R.S., Foyle, D.C. Superimposed symbology: Attentional problems and design solutions, *SAE Transactions: Journal of Aerospace*, 103 (1994), 2009-2016.
- Mewhinney, M. Flow over helicopter carriers, *NASA Ames Astrogram*, October 30, 1998.
- Milgram, P. and Kishino, F. A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information Systems*, E77-D (12), 1994.
- NASA Marshall Space Flight Center Lidar Tutorial, http://www.ghcc.msfc.nasa.gov/sparcle/sparcle_tutorial.html.
- Naval Safety Center Aviation Data Analysis Division, <http://www.safetycenter.navy.mil/aviation/aviationdata/default.htm>.
- Orlady, H.W. The professional airline pilot of today: All the old skills--and more. International Airline Pilot Training Seminar, Caracas, Venezuela (1989).
- Particle Image Velocimetry at NASA Glenn, <http://www.grc.nasa.gov/WWW/OptInstr/piv/background.htm>.
- Post, F., van Wijk, J. Visual representation of vector fields: recent developments and research directions, *Scientific Visualization Advances and Challenges*, Academic Press (1993), 367-390.
- Prinzel, L. Head-Up Displays and Attention Capture, NASA Technical Memorandum 2004-213000.
- Roscoe, S.N. Airborne displays for flight and navigation, *Human Factors*, 10 (1968), 321-322.
- Shaw, C.D., Jiang, F.T., Parry, R.M., Plale, B., Wasilewski, A., Ribarsky, W., Faust, N.L. Real-Time Weather Data on Terrain, NASA AWIN, SPIE Vol. #4368, *Radar Sensor Technology and Data Visualization* (2001).
- Spence, R. *Information Visualization*, Addison-Wesley/ACM Press (2001).
- Spirkovska, L., Lodha, S.K. AWE: Aviation weather data visualization environment, *Computers and Graphics*, 26 (2002), 169.
- Spitzer, C. R., Parrish, R.V., Baize, D.G., Lewis, M.S. *The Avionics Handbook: Synthetic Vision*, NASA Aviation Safety Program (2001).
- Strid, T., Rizzi, A., Ooppelstrup, J. Development and use of some flow visualization algorithms, Computer Graphics and Flow Visualization in Computational Fluid Dynamics, von Karman Institute for Fluid Dynamics Lecture Series 1989-07 (1989).
- Strope, K., Borden, C. and Harding, J., Verification and Validation of a UH-60 FLIGHTLAB model in support of the UH-60M Limited User Test. American Helicopter Society 60th Annual Forum (2004).
- Sujudi, D., Haimes, R. Identification of Swirling Flow in 3D Vector Fields, AIAA Paper 95-1715, San Diego, CA, USA (1995).
- Switzer, G., Britt, C. Performance of the NASA Airborne Radar with the Windshear Database for Forward-Looking Systems, NASA CR 201607 (1996).
- Turner, B., Leclerc, M.Y., Gauthier, M., Moore, K., and Fitzjarrald, D. Identification of Turbulence Structures above a Forest Canopy using the Wavelet Transform, *J. Geophys. Res.* 99 D1 (1994), 1919-1926.
- Uenking, M.D., Hughes, M.F. The Efficacy of Using Synthetic Vision Terrain-Textured Images to Improve Pilot Situation Awareness, SAE World Aviation Congress (2002).
- United States Department of Commerce, 1986, The Crash of Delta Flight 191 at Dallas-Fort Worth International Airport on 2 August 1985: Multiscale Analysis of Weather Conditions, National Oceanic and Atmospheric Administration, Boulder, CO, USA.
- United States Navy, A1-H60BB-NFM-000, NATOPS Flight Manual – Navy Model SH-60B Aircraft (2004).
- Wallace, L. Airborne Trailblazer, NASA Langley Wind Shear Program (1993), <http://oea.larc.nasa.gov/trailblazer/SP-4216/>.
- Wickens, C.D., Carbonari, R., Merwin, D., Morphey, E., O'Brien, J. Cockpit Displays to Support Hazard Awareness in Free Flight. NASA Technical Report ARL-97-7/NASA-97-4 (1997).
- Wilkinson, C.H., Zan, S.J., Gilbert, N.E., Funk, J.D. Modelling and Simulation of Ship Air Wakes for Helicopter Operations. Fluid Dynamics Problems of Vehicles Operating Near or in the Air-Sea Interface RTO-MP-15 (1999).
- Williams, S.L., Long, K.R. Dynamic Interface Flight Tests and the Pilot Rating Scale. American Helicopter Soc. 53rd Annual Forum (1997).