ISOMORPHIC SPATIAL VISUAL-AUDITORY DISPLAYS FOR OPERATIONS IN DVE FOR OBSTACLE AVOIDANCE

Dr. Martine Godfroy-Cooper, martine.godfroy-1@nasa.gov, SJSU/ US Army Aviation Development Directorate, Moffett Field, CA, USA; Joel D. Miller, joel.d.miller@nasa.gov, SJSU/ US Army Aviation Development Directorate, Moffett Field, CA, USA; Dr. Edward Bachelder, edward.n.bachelder@nasa.gov, SJSU/ US Army Aviation Development Directorate, Moffett Field, CA, USA; Dr. Elizabeth M. Wenzel, elizabeth.m.wenzel@nasa.gov, NASA ARC, Human System Integration Division, Moffett Field, CA, USA.

Abstract

Helicopter military missions such as combat search and rescue, medical evacuation and landing on unprepared sites can involve operating in hostile, low-altitude, and degraded visual environments (DVE). These conditions may significantly reduce the pilot's capability to use the natural out of the window (OTW) perceptual cues, increase workload and increase the risk of collision with terrain and natural or man-made obstacles. In modern helicopter cockpits, synthetic vision systems (SVSs) can employ conventional non-conformal two-dimensional (2D), egocentric three-dimensional (3D) conformal symbology (CS) and laser detection and ranging (LADAR)/ radio detection and ranging (RADAR)/ forward looking infrared (FLIR) imagery support guidance and control, especially during operations in DVE. Although 3D CS can decrease pilot workload, it can also produce attentional tunneling (cognitive capture) and may not provide maximally effective depiction of the environment around the helicopter. In this context, it is crucial to develop integrated multimodal interfaces that extend the current operational envelope while enhancing flight safety. Several flight simulator studies have investigated the use of spatial auditory displays (SADs) in combination with spatially and temporally congruent visual displays in tasks as diverse as collision avoidance, intruding aircraft detection, or system malfunction warning. In this paper we propose a novel approach to spatial sonification design based on the premises that perception-based synthetic cueing can increase situation awareness (SA), improve overall performance, and allow mental workload to be kept at operationally effective levels. This paper discusses the development, implementation, and evaluation of a sensor-based augmented-reality spatial auditory display (ARSAD) and its visual analog, an integrated collision avoidance display (ICAD) for all phases of flight. Five UH60M Army pilots participated in a low-level flight simulation evaluating the visual and the auditory displays, alone or in combination in low-visibility and zero visibility environments. The results are discussed in the context of pilot cueing synergies for DVE.

1. INTRODUCTION

This study was a joint effort by the U.S. Army Aviation Development Directorate (ADD) and the National Aeronautics and Space Administration Ames Research Center (NASA ARC).

Helicopter pilots' military missions often involve operating in hostile, low-altitude, and in DVE, and can lead to spatial disorientation (SD) and the subsequent loss of SA. In this context, low-altitude flying in DVE exploits the terrain profile to reduce the enemy ability of visual, optical or electromagnetic detection and therefore enhances the survivability. These conditions modify significantly the pilot's capability to use the natural OTW perceptual cues, increase workload and lead to failure to maintain sufficient clearance with the obstacle, and ultimately, collision with terrain controlled flight into terrain (CFIT), natural objects (trees) or erected structures (buildings, poles, towers and wires). According to a recent US Army Aviation accident report [1] from Fiscal Year 2011 through Fiscal Year 2015, 31% of events for class A and 17% of the events for class B were classified as collision-related. Among obstacles, wires represent a specific hazard due to their near invisibility. During the 1994-2003 period, US Army helicopters were involved in 1160 accidents, in which 34 were wire strikes (7 fatalities).

Presented at 44th European Rotorcraft Forum, Delft, The Netherlands, 19-20 September, 2018

In modern helicopter cockpits, synthetic vision systems (SVSs) employing conventional non-conformal 2D, egocentric 3D CS and enhanced vision systems (EVS) such as LADAR/ RADAR/ FLIR imagery support guidance and control, especially during operations in DVE. The primary role of 3D CS and sensor-based imagery is to augment pilot's visual perception to support guidance and control especially during operations in DVE. Although EVS and SVS can improve pilot's SA, thus lowering workload, it can also be misleading, produce clutter and attentional tunneling and may not provide maximally effective depiction of the environment around the helicopter. In this context, it is crucial to develop integrated multimodal interfaces that allow extending the current operational envelope while enhancing flight safety. Aural, tactile and spatial auditory cueing are candidate cueing strategies, although their respective affordances suggest specific uses as a function of the phase of the flight.

Several flight simulator studies have investigated the use of spatial, 3D auditory displays in combination with spatially and temporally congruent visual displays in tasks as diverse as collision avoidance, intruding aircraft detection, or warning for system malfunction. Since 2015, in the context of the DVE-M program [USAARL integrated cueing environment (ICE)], two studies [2], [3] have evaluated the use of 3D spatial auditory cueing for the representation of natural and man-made obstacles.

This paper discusses the evaluation of an augmented-reality spatial auditory display (ARSAD) and its visual analog, an integrated collision avoidance display (ICAD) for obstacle avoidance for all phases of flight.

The two sensor-based displays were prototyped in an ownship simulated environment. The emulated sensor performed 360° short-range radar sweeps in an Earth-horizontal plane below the helicopter center of gravity (COG). After one full azimuth rotation, the sensor determines the two nearest hits. The sensor has a "tolerance" value, which is a +/- value about the first hit to omit from the search for the second hit. The current tolerance is 90°, yielding, in essence, two 180° half-circles. The notion of "safety profile", inclusive of the main rotor footprint was used to determine "security margins". Two rings (caution and warning) were determined by the time to collision (TTC) or distance to obstacle.

For the spatial auditory display (SAD), two sonifications were designed to provide 3D obstacle positional cueing for the two-nearest obstacles using spatial earcons that are *conformal* with the real world. Complex sound synthesis topologies were explored to meet the following requirements for two simultaneous sonifications in the display: identify itself and convey urgency (e.g., nearest obstacle, second-nearest obstacle), belong to the same class of alert (sensor-detected obstacles) and yet be distinct, stand out from the background noise (i.e., avoid masking), stand out in the overall soundscape, spatialize well, not merge, and convey additional meaning, if available (e.g., ownship-obstacle bearing angle). A specific sonification was developed to represent power lines, since power lines still remain a significant source of accidents for low-level operations in DVE.

The ICAD was designed to "complement" visually the SAD, by providing an isomorphic representation that matches the sonification behavior, including caution and warning zones, speed-dependent sensor sweep extent, and speed-dependent safety margin indicators.

A simulation was conducted at the SIL at NASA ARC with five UH60 US Army pilots. Unimodal (visual only, auditory only) and bimodal display presentation (visual + auditory) was tested in low-visibility (LV) and zero visibility (ZV) conditions during a low-altitude flight, including enroute, approach, and hover/landing. The results demonstrated the usability and acceptability of the augmented-reality spatial auditory display and the integrated collision avoidance display to convey intuitive information (re) presentation for obstacle avoidance.

2. AUDITORY DISPLAYS

2.1. Generalities

Auditory displays have been the subject of research for well over two decades [4] and their definition still varies among authors. Here, we refer to an auditory display as any display that uses sound (speech and non-speech) in computational settings to communicate information to users. It was suggested [5] that it should also include the user context (user, task, background sound, constraints) and the application context (aircrafts, automobiles, etc.), since these are all quite essential for the design and implementation.

The rationale and motivation for displaying information using sound (rather than visual information) have been discussed extensively in the literature [6]. Because auditory displays exploit the superior ability of the human auditory system to recognize temporal changes and patterns [7], they may be the most appropriate modality when the information being displayed has complex patterns, changes in time, including warnings, or call for immediate action.

Sonification, using synthesized non-speech sound, is thereby an integral component within an auditory display system, which addresses the rendering of

Presented at 44th European Rotorcraft Forum, Delft, The Netherlands, 19-20 September, 2018

sound signals that depend on data and optional interaction. Sonification is generally defined as the mapping or transformation of data streams onto auditory dimensions for the purposes of facilitating communication or interpretation [8]. Changes in data values are associated with a change in an associated acoustic parameter, such as sound wave frequency or amplitude. Sonifications are built upon the notion of pre-attentive awareness and exploit the auditory modality's ability to recognize patterns or small changes in an auditory event. Sonification's short information units (as compared to speech) make it well suited for conveying rapidly changing data such as relative distance and orientation. Sonifications promote eyes-free continuous monitoring without startling or disrupting attentional focus [9]. Thus, if sonifications are designed and implemented effectively, human operators may effectively monitor complex systems while adhering to additional responsibilities without having to constantly switch attention from one task to another.

Sonification includes auditory icons, earcons, and audification. Auditory icons represent a sound "image" of the object to which it is referring. This is a direct comparison to visual icons. E.g., a heartbeat sound can be used for monitoring pulse information [10]. Earcons are nonverbal abstract audio messages used in the user-computer interface to provide information to the user about some computer object, operation, or interaction [11]. In contrast to auditory icons, earcons are harder to remember and learn because they have no natural link or mapping to the objects or events they represent. On the other hand, they are highly structured and can easily represent families and hierarchies of objects and actions with very simple audio messages. This type of sonification has better results in desktop interfaces, alarms and warning systems such as vehicle collision detection systems, and immersive virtual environments (VEs). Sonification has been used successfully in advanced driver assistance systems (ADAS) with high priority warnings such as forward collision warnings, lane or road departure warnings, and blind spot and back-up warnings. Lastly, audification is a specific type of auditory data analysis in which data samples are isomorphically mapped to time or frequency domain audio data. Audification is the most direct form of sonification, as all data samples are preserved and spectral features within the original data will be present as timbral components in the resulting sound.

2.2. Spatial auditory displays

Spatial auditory displays (SADs) (also referred to as virtual auditory displays) use spatial auditory cues (sounds with spatial positional characteristics) to provide information to a user. SADs create a virtual auditory space where the auditory information can be substitutive or redundant to visual information. Since sensory systems are energy specific, each system provides the organism with characteristic properties that can be either exclusive, i.e. or conversely, amodal, i.e., shared by two or more sensory systems. Indeed, color and timbre are modality-specific, while physical location or duration can be equally conveyed by vision and audition. Different modes for the uses of (spatialized) sound can thus be inferred: substitution, which is the condition in which one modality replaces another modality when the other is not available or degraded, and complementarity [12], where congruent inputs from different sensory channels are combined. For example, spatial auditory displays can be used to alleviate visual workload when the visual channel is saturated [13], [14]. They can also be developed for use in applications for which visual information provides no benefit, in limited field-of-view (FOV) applications, teleoperation [15], [16] or presenting information to the blind [17]. More recently, Beattie [18] investigated the potential application of spatial earcons for presenting primary driving information in automated vehicles.

2.2.1 SADs for Aviation

SADs can be used in complex dynamic tasks such as urban combat simulations, flight simulations, air traffic control, and military command and control. Potential applications include monitoring multiple radio communication channels [19], navigating waypoints, locating threats or system malfunctions, and teleoperation of unmanned vehicles [15]. In cockpit applications, with helmet or head-mounted visual displays with limited field of view, spatial audio can be used to direct the attention to critical events outside the FOV. DVE is another condition in which spatial auditory displays can provide complementarity or substitution to the visual information.

Several flight simulator studies have investigated the use of 3D audio for the aural Traffic alert and Collision Avoidance System (TCAS) warning, which is installed in most commercial aircraft [20]. All studies showed that out-the-window visual search time for the intruding aircraft was reduced with 3D audio, compared to monaural warnings. Bronkhorst [14] examined the application of 3D audio to indicate the location of a target jet in a fighter intercept task. They observed that the fastest target acquisition times were obtained with the combination of the visual head down display (HDD) and the 3D auditory display. No difference was found between the conditions with only the visual display or the 3D auditory display. The application of 3D audio can also be extended to other types of auditory signals in the cockpit. For instance, Haas [21] used 3D audio as a warning display for system malfunctions in helicopters, where the spatial source of the 3D audio warning corresponded to the location of a system malfunction of the aircraft or to the location of a visual indicator light inside the cockpit. The results showed faster warning response times when they were presented with 3D audio (i.e., 3.6 sec on average) compared to the condition when only visual warning signals were present (5 sec). Bastide [22] uses spatial sound to create a multimodal command and control interface for the Rafale aircraft. In critical domains such as low-level flight where unintentional drift, changes in altitude, and sink rates require immediate counteractive measures to avoid flight into terrain, auditory cues have the ability to capture pilot's attention and elicit orientation responses regardless of head position or eye fixation [23]. Novel uses for sonifications have been suggested for the depiction of obstacle location during a simulated helicopter drift during a hover in DVE [2]. Using two earcons (pulsed frequency-modulated waveforms with square-wave modulators, and looming effect), Godfroy-Cooper et al. [2] demonstrated that a single obstacle presented in the frontal hemifield in the horizontal median plane could be localized, under optimal conditions [individualized head-related transfer functions (HRTFs), best sonification type and continuous presentation] with an average accuracy of 3.3° and an average precision of 4.2°.

2.3. Ecological Psychoacoustics

The accurate and precise determination of the spatial location and path of objects in the environment is crucial for navigation and object interaction. Unlike the visual system, for which there is a relatively isomorphic correspondence between spatial position in the environment, position in the retina (retinotopic coding), and organization along the visual pathway, auditory spatial information is not directly represented at the level of the sensory receptor. Instead, the sound source location is estimated by integrating neural binaural properties interaural level differences (ILDs) and interaural time differences (ITDs) (for azimuth, defined by the angle between the source and the sagittal plane) and frequency-dependent pinna (external part of the ear) filtering (for elevation, defined by the angle between the source and the horizontal plane containing the listener's ears) [24]. As a result of these differences in coding spatial information in the visual and auditory systems, vision spatial resolution is superior by up to two orders of magnitude [1 min of angle (minute of arc, MOA)] [25], compared to the auditory domain [minimum audible angle (MAA): 1° to 2° for frontal positions, 6-7° for rear] [26], [27]; while the temporal resolution of the auditory system exceeds that of the visual system [28]. Thus, the two systems *complement* each other. Some of the deficits of the visual system due to environmental or physiological factors, for example, not being able to perceive 360° of azimuth simultaneously, or to sense through obstacles are *compensated* for by the auditory system (in contrast to light, sound is generally able to travel around and/ or through occluding objects). Furthermore, audition plays a key role in guiding locomotion by the central nervous system (CNS) when vision is not available, for which an accurate internal representation of the distance between the organism and the target is essential. The two principal dimensions of egocentric (observer as origin) auditory spatial perception are direction and distance of the sources.

2.3.1 Direction

The localization of an auditory stimulus in the horizontal dimension (azimuth) results from the detection of left-right ITDs and ILDs [29]. To localize a sound in the vertical dimension (elevation) and to resolve front-back confusions [30], the auditory system relies on the spectral cues provided by the detailed geometry of the pinnae. Pinna features cause acoustic waves to diffract and undergo direction-dependent reflections [27], [31]. The two different modes of indirect coding of the position of a sound source in space (as compared to the direct spatial coding of visual stimuli) result in differences in spatial resolution in these two directions. Indeed, auditory localization performance is "direction-dependent". Localization precision and accuracy is greater in azimuth (horizontal median plane, HMP) than in elevation. For a sound source located on the sagittal median plane (SMP), precision varies between 2º and 3º in azimuth, 4º to 9º in elevation. For accuracy, Makous & Middlebrooks [32] found similar variations: 1.5º in azimuth, 2.5º in elevation. Auditory localization precision is maximum in the SMP and remains relatively constant outside this plane. Auditory localization accuracy is the greatest for sound sources located 10º to 23º ("auditory horizon") above the visual horizon (0º elevation) and is characterized by a symmetrical undershoot around this plane, resulting in a compression of the auditory space in this dimension.

2.3.2 Distance

Auditory distance perception plays a major role in spatial awareness, enabling location of objects and avoidance of obstacles in the environment. Sound localization in this third dimension is not nearly as accurate as that in the horizontal and vertical dimensions. Much as with the perception of

Presented at 44th European Rotorcraft Forum, Delft, The Netherlands, 19-20 September, 2018

visual distance, there are several sources that listeners can use to determine the distance of sound sources. Two of the most informative cues are intensity change (i.e., sound level arriving at the listener's ears), and direct-to-reverberant (D/R) energy ratio [33], [34]. The relative importance of these cues varies widely across conditions.

The intensity cue arises from the physical attenuation of a sound with distance. Given a point sound source in anechoic conditions, sound intensity arriving at the listener will decay by 6 dB with every doubling of the distance; the rate of decay is lower in reflective surroundings or if the source is directional. The range over which distance cues are operable varies, and some cues are only useful within peripersonal space (sounds that are within reaching and grasping distance, <1 m from the listener), a region where internal representations of distance are based on both auditory and tactile information [35]. Listeners may be particularly sensitive to auditory distance for near sources, potentially because nearby auditory events may require immediate motor responses, especially if the signal is threatening or particularly interesting. Also, there are spectral cues for near-field sources that don't occur for far-field sources as a function of distance. Note listeners tend to underestimate distances that are greater than 1.5 m and tend to overestimate distances that are less than 1.5 m [34]. Distance judgments are also generally more accurate for lateral sounds than for sounds near the median plane, both for far and for nearby sources [36]. Finally, non-perceptual factors, including the importance of the auditory event to the listener, also can affect perceived distance.

2.3.3 Auditory Looming and Time to Contact (TTC)

Visual looming refers to the rate of change in the size of an approaching object's retinal image. A corresponding auditory "looming effect" [37] exists supported by monaural loudness changes, interaural time differences, and to a lesser extent, Doppler effect. Both ITDs and monaural intensity change have salient physical characteristics that mark the point of closest passage for a sound source. The Doppler shift (change in frequency) has not such salient characteristics.

For a constant-frequency approaching source, the frequency that arrives at the observation point (perceived as pitch by the listener) is initially higher than the frequency that is emitted by the source. The observed frequency remains initially constant, rises at a successively increasing rate as the source approaches, and finally drops at a successively decreasing rate as the source recedes. The magnitude of the frequency drop depends on the speed of the source. However, despite the drop in frequency, listeners tend to report hearing a rise in pitch as acoustic sources approach. The apparent paradox between falling frequency and perception of a rising pitch has been termed the "Doppler illusion" [37]. The pattern of perceived rising intensity produced by an approaching sound source is particularly salient information on source approach. It has been termed "acoustic tau" [38] in reference to the visual tau variable that specifies the time to contact by the optical expansion pattern produced by visual approaching objects. Interestingly, humans systematically underestimate the source location and generally underestimate the TTC, expecting contact before the source actually arrives [39]. This tendency may provide sufficient time to initiate an appropriate behavior to avoid the object. This primary warning role of the auditory system is also at work in the estimation of auditory distance judgments by a listener in motion toward an auditory object [40].

2.4. Virtual Acoustics and Head-Related Transfer Functions (HRTFs)

A virtual auditory space (VAS) is created through the use of loudspeakers or headphones and designed to replace or augment the natural listening environment. An anechoic individualized VAS can be generated by simulating the wave pattern at the eardrum of an external sound source in the free field [41]. For each sound source location in space relative to a listener's head, a unique spectral and temporal pattern is imposed on the sound by the head, pinnae, and torso. These patterns are termed HRTFs in the frequency domain (Head-Related Impulse Responses, HRIRs, in the time domain) and can be captured and reproduced to create a purely virtual simulation. Alternatively, the simulation can overlay the listener's existing environment to create an augmented-reality display.

In a static anechoic environment, filtering of a source signal with the HRTFs for a given direction delivers to the listener's eardrums the same acoustic pressure wave as the true source in the same environment. By including reverberation and motion cues due to ego-motion of the listener, one can synthesize more realistic environments [42]. Unfortunately, individual differences in anatomy, especially the shape of the pinnae, means one HRTF dataset does not fit all. Pinnae, head, and torso sizes can vary greatly from one person to the next. Thus, spectral characteristics can also vary greatly so that the HRTFs of one individual can yield significant perceptual distortions when used for another.

Presented at 44th European Rotorcraft Forum, Delft, The Netherlands, 19-20 September, 2018

3. VISUAL DISPLAYS

3.1. Enhanced, Synthetic and Combined vision systems

The function of enhanced, synthetic and combined vision systems is to provide a supplementary view of the external scene thereby delivering the crew with an awareness of terrain, obstacles and relevant man-made features such as buildings, towers and wires. Alerting functions can be added.

3.1.1 Enhanced Vision Systems (EVS)

An EVS is real time "electronic means of displaying a sensor-derived or enhanced real-time image of the external scene through the use of external sensor such as forward looking infrared (FLIR), millimeter wave radiometry, millimeter wave radar and/ or low-light level image intensifying" [43].
The image is displayed to the pilot *conformal* to the outside scene, i.e. the pilot sees the displayed elements the same relative size and aligned with objects outside the aircraft.
Conformal symbology (CS) [44] can be superimposed on the display image, such as the locus of the landing zone (see Figure 1 Top).

3.1.2 Synthetic Vision Systems (SVS)

An SVS is an aircraft cockpit display technology that presents the pertinent and critical features of the environment external to the aircraft through computer-generated image of the external scene topography from the egocentric perspective of the flight deck (egocentric) or from an exocentric perspective (the aircraft position symbol is placed on the terrain and obstacle map).
SVS are usually displayed in a track-up orientation (rather than north-up) to avoid circular mental rotation and translation cognitive operations required to align the egocentric reference frame (ERF) and the world reference frame (WRF) [45], [46].
The displayed information is derived from aircraft attitude, altitude, position and a coordinate-referenced database [47]. Enhanced intuitive views, precise navigation guidance, and hazard detection displays are key elements of SVS. Enhanced awareness is achieved by employing a "look-ahead" function (forward looking terrain avoidance warning), also referred to as enhanced ground proximity warning system (EGPWS), or automatic ground collision avoidance system (auto-GCAS). Helicopter terrain and warning systems (HTAWS) displays provide the pilot with alerts (usually color-contouring and aural advisories) of potential wires, terrain and obstacle conflicts along the flight path [48], [49], [50].

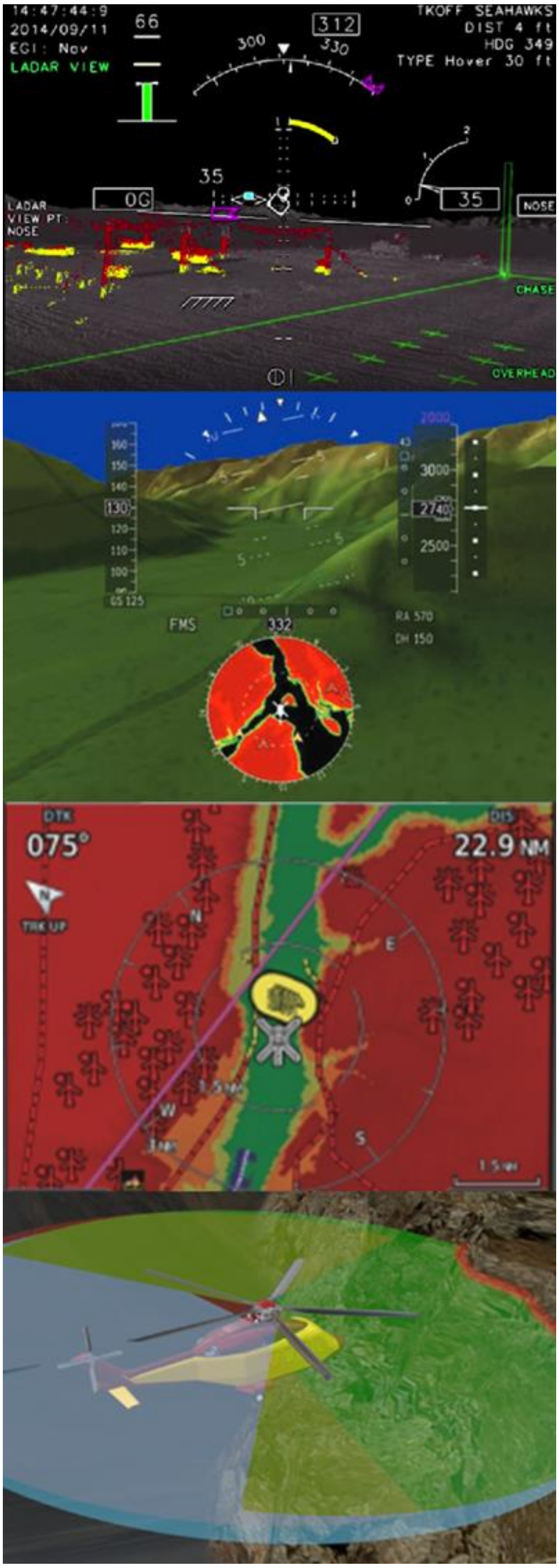


Figure 1. From top to bottom: EVS: ADD LADAR, SVS: Rockwell Collins, Garmin (with WireAware power line avoidance), and Augusta Westland OPLS.

Presented at 44th European Rotorcraft Forum, Delft, The Netherlands, 19-20 September, 2018

Some of the commercially available systems are: Honeywell's HTAWS, Sandel Avionics' HeliTAWS featuring a "WireWatch" capability to provide advance warning of transmission wires whether they are powered or not, Garmin WireAware Wire-Strike Avoidance Technology that graphically overlays comprehensive power line location and altitude information on the moving map and AgustaWestland Obstacle Proximity light detection and ranging (LIDAR) System (OPLS) [51] ((see Figure 1). These are designed to help the crew avoid main and tail rotor strikes against peripheral obstacles which jeopardize the aircraft's safety during low speed hovering maneuvers in confined spaces.

3.1.3 Display Integration

The number, quality and interaction between sensory modalities are key to the realism of the simulated environment and ultimately, to its usefulness. Modality combination should support thematic congruent manageable information loading, complementarity, consistency (spatial temporal and semantic congruency), viewpoints (shared reference frames and map orientation), and redundancy, i.e. the use of several modalities for processing identical information [52], [53], [54].
A natural basis for sensory substitution or complementarity is the *isomorphism* of the perceptual representations created by two senses. Under a range of conditions, visual and auditory perception result in nearly isomorphic perceptual representations. The similar representations are likely the basis both for cross modal integration, where two senses cooperate in sensing spatial features of an object, and for the case with which subjects can perform cross-modal matching, i.e. hearing an object and then recognizing it visually. *Spatial isomorphism* between representations from two modalities ensures that parameters extracted from one will match those of the other, without systematic bias. Spatial and temporal register between the sensory inputs is a pre-requisite for an integrated user experience. It supports the semantic congruency and unity assumption (i.e. a dog's image and a barking sound) [55].

4. THE PRESENT RESEARCH

The ARSAD for obstacle detection and avoidance was designed in the context of the ADD DVE-M program (a preliminary version was described extensively by Miller et al., [3]) and more specifically, as a component of the integrated cueing environment (ICE) effort. As such, its usability was contingent to the development of a visual analog, ICAD. The two displays share the conceptual representation of speed-dependent sensor sweep extent, safety margin indicators (caution and warning zones), and nearest and second-nearest obstacle representation. They both share a new Power Line warning system as well, that includes a distance-based sonification and an altitude-to-go visualization. The experimental design allowed the evaluation of each display individually, thereby providing a proof of concept for ARSAD as a substitutive display in case of visual display malfunction/ unavailability and as an integrated multimodal display suite. It was also designed to determine how the egocentric PFD and exocentric ICAD were used when merged with the information provided in the auditory dimension. Lastly, the simulation allowed one to evaluate how the degree of information provided by the OTW scene modulated the relative distribution of attention between the visual and the auditory displays.
The results of these observations will be presented in a forthcoming paper. Pilot Bedford ratings were collected after each individual run. Comments and responses to a tailored questionnaire were gathered at the end of the experiment to determine the usability and acceptability of the displays.

5. METHOD

5.1.1 ARSAD

The Sensor Model and Obstacle Sensor Sonifier Logic

The Obstacle Sensor Sonifier responds to obstacle locations provided by a short-range LIDAR/RADAR-like Sensor Model. As such, the logic for the overall sensor-detected obstacle sonification is split between the Sensor Model and the Obstacle Sensor Sonifier.
Similar to the real-world sensor probe of physical objects using radio waves or lasers, simulators typically have the capability of performing object polygon hit testing with virtual lasers fired in the virtual environment [57].
The Sensor Model sweeps 360° in an Earth horizontal plane below helicopter COG, performing 720 hit tests per rotation for a resolution of 0.5° at a 2 Hz update rate. The helicopter can pitch and roll relative to the sensor frame, but sensor frame yaw remains fixed to helicopter yaw. The sensor descends 2.5 meters after lifting off. Because two obstacle sonifiers can be presented at the same time, it was necessary to determine a minimum spatial separation between the two sensor hits to ensure that 1) they do not depict the same object (for example, a wall) and 2) that they are perceived as representing two separate obstacles.

Presented at 44th European Rotorcraft Forum, Delft, The Netherlands, 19-20 September, 2018

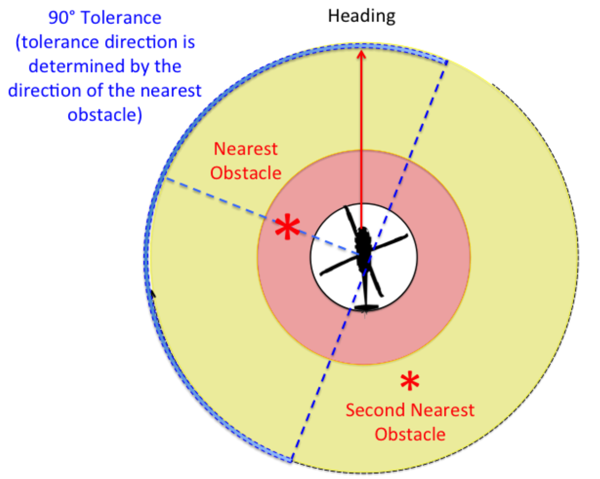


Figure 2: Sensor Model "tolerance" of 90° about the nearest hit to omit from search for second-nearest hit.

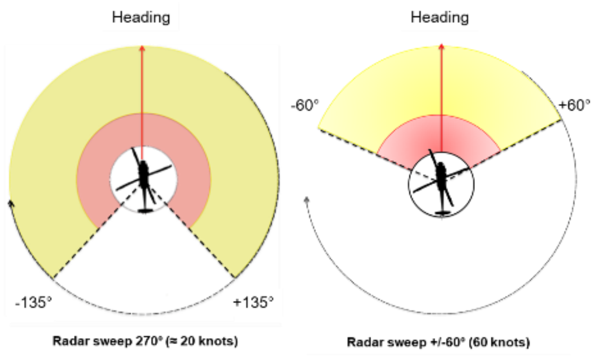


Figure 3. Speed-Dependent Sensor Sweep Extent. Left: at 20 knots, the sensor sweeps 270° (+/- 135°). Right: at 60 knots, the sensor sweeps its minimal value of 120° (+/- 60°).

This resulted in the development of a "tolerance" parameter, which is a +/- angle value about the first hit to omit from the search for the second hit.
In the present configuration, the tolerance is set to 90° yielding a 180° pie slice centered on the "Nearest" hit and a second remaining 180° pie slice containing the "Second-Nearest" hit (Figure 2). However, smaller tolerance values could be considered for low-speed phases of flight, when greater spatial resolution is required.

Speed-dependent sensor sweep extent

As ownship speed increases, sensor sweep extent decreases linearly from 360° below 6 knots to +/- 60° at 60 knots. The speed-dependent sweep extent method was used to eliminate non-threatening obstacles, in particular obstacles located to the sides and behind the back of the aircraft during the enroute phase of flight (see Figure 3).

Sonification Mappings

Obstacle Urgency to Earcons

For the Obstacle Sensor sonification, the two nearest obstacle hits are identified by two unique spatial earcons, "Nearest" and "Second-Nearest", mapping urgency to timbre. The earcons are synthesized in real time using three signal oscillators and two modulators, a topology based on traditional hardware synthesizers [3].
The "Nearest" earcon is composed of a sawtooth and two triangle waveforms with a pitch reference of 622 Hz while the "Second-Nearest" is composed of two square waves and one sawtooth with a pitch reference a minor third (3 semitones) below the "Nearest" pitch reference. The result is the "Nearest" earcon sounding slightly higher in pitch and a bit harsher than "Second-Nearest" to denote urgency. To convey the same class of alert, identical modulation values were used, a 5 Hz depth with modulation frequencies of 64 Hz and 87.5 Hz (details in [3]).

Obstacle Location to Augmented-Reality Display Location

Given that obstacle locations are presented using an augmented-reality display, obstacle location is mapped to the acoustic model parameters azimuth, elevation, and range relative to the listener. Azimuth and elevation are implemented via HRTF indexing and interpolation, and range via a spherical-spreading loss gain model.

Obstacle Azimuth to Pitch Scaling

To accentuate obstacle azimuth angle relative to ownship and to reduce front-back reversal, a Sonifier "Pitch Scaling" algorithm was developed. Inspired by HRTF head and pinna shadowing (a darkening of the sound due to the head and pinna's obstruction of high frequencies for rear-incident sources), it reinforces HRTF shadowing when the pilot is looking forward.
Eight azimuth pie slices (four for each hemisphere) are used to decrease earcon pitch 40 cents per slice, front-to-back (Figure 4). Since the pitch scaling is performed relative to ownship, the pitch remains unchanged with head-tracked head motion.

Dynamic Obstacle Range to Looming Effect

Visual looming refers to the rate of change in the size of an approaching object's retinal image. A corresponding auditory "Looming Effect" occurs with an oncoming sound's increase in intensity over time.

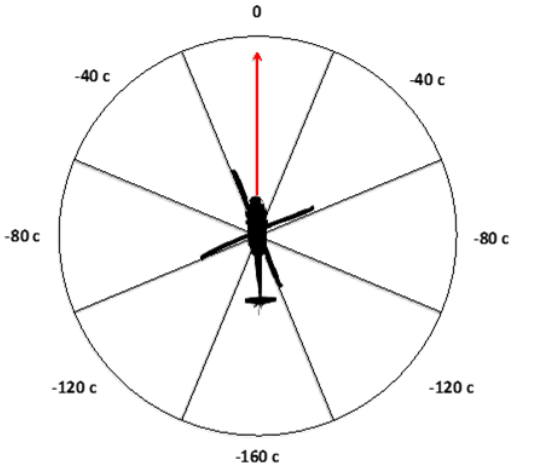


Figure 4. Obstacle-Ownship azimuth pie slices for sonifier earcon pitch scaling. Implemented to accentuate obstacle azimuth angle relative to ownship and to reduce front-back reversals.

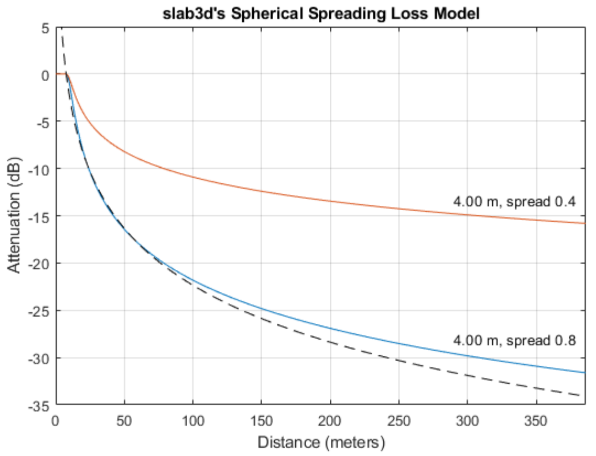


Figure 5. Looming Effect implemented via slab3d's source-listener distance gain model for a source radius of 4 m, spread exponents of 0.4 (Obstacle Sensor) and 0.8 (Power Line), and a 0 dB reference at a helicopter blade radius of 25 ft (7.6 m). Also shown is a dashed inverse-distance gain curve referenced to 25 ft.

Therefore, it is advantageous for a visual object's sonification to share an overall stimulus energy profile with the visual object (when visible).

The audio engine used for implementing the sonifications was the Open-Source slab3d real-time virtual acoustic environment rendering system (http://slab3d.sonisphere.com/) [56]. Its sound-source spherical spreading loss model was used to fine tune the Obstacle Sensor and Power Line sonification Looming Effects. The model computes a spreading loss gain attenuation of: $(1 + d^2 / r^2)^{-s/2}$, where d is the source-listener (obstacle-listener) distance, r is the source radius, and s is a spread exponent (Figure 5).

When s is 1, this characteristic closely approximates that of a planar baffled cylindrical piston of radius r [58]. If the distance d is made relative to a helicopter blade radius of 25 ft., the model parameters source radius 4 m and s 0.8 approximate point-source inverse-distance gain behavior (Figure 5, dashed curve). This yielded a good dynamic range for the Power Line sonification with the warning ring extending to 386 m at 60 knots.

To keep the Obstacle Sensor sonification gain comfortable at short distances while being audible at longer distances (e.g., 232 m at 60 knots) required a reduction of the spread exponent to 0.4. Thus, the Power Line sonification closely follows a physically realistic Looming Effect, while the Obstacle Sensor sonification follows a similar physical exponential gain profile, but with less attenuation.

Obstacle Range to Pulse Period

For non-speech audio, Blattner [59] proposed an approach to construct (warning) signals based on the musical qualities of auditory information using relatively simple tones. Patterson [60] and Edworthy [61] stated that temporal aspects are critical in distinguishing between sounds and that speed is probably the strongest influence on perceived urgency. Later work by Brewster [62] showed that rhythm and tempo variations (i.e., speeding up or slowing down the patterns) are an effective method for differentiating earcons.

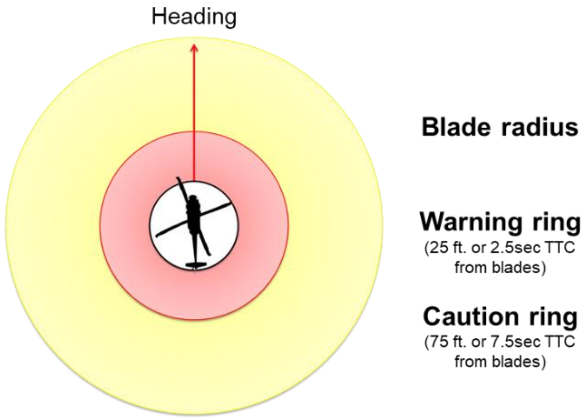


Figure 6. The caution and warning rings were set to 75 ft and 25 ft, respectively, from the helicopter's blades. The corresponding TTC values were set at 7.5 sec for the caution ring and 2.5 sec for the warning ring.

Presented at 44th European Rotorcraft Forum, Delft, The Netherlands, 19-20 September, 2018

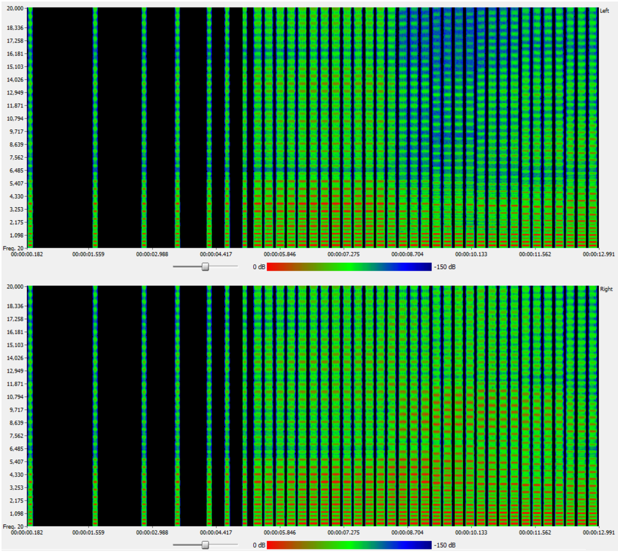


Figure 7. A spectrogram of a "Nearest" obstacle earcon illustrating the time and frequency sonification mappings (top: left ear, bottom: right ear). The obstacle was approached head on and then the helicopter yawed 180°. The urgency is conveyed by the "Nearest" earcon harmonic pattern (timbre). The obstacle location in azimuth and elevation is shown by the HRTF's selective shaping of the earcon frequency content. The azimuth Pitch Scaling is indicated by the vertical frequency shift downward of like harmonics during the 180° rotation. Obstacle range is denoted by both a Looming Effect gain increase as the obstacle nears as well as an increase in pulse rate (caution zone) and pulse duration (warning zone) (horizontal time axis).

The range of signal rates generally applied is based on the standard work of Hellier et al. [63], who used a 200 ms tone and inter-pulse intervals ranging from 9 to 475 ms (i.e., pulse rates of 1.5 to 4.8 Hz). Small pulse durations (<80 ms for complex and <30 ms for simple earcons) decrease perception and should be avoided [62].

Both pulse-duration and pulse-period ownship-obstacle distance sonification mappings were explored with time-to-collision and fixed-distance thresholds. Given pilots were already accustomed to the fixed-distance pulse-period collision indicators provided in some modern automobiles, a design mimicking that behavior was selected. The Obstacle Sensor sonification uses complex tone pulses 80 ms in duration with an inclusive fade in and out of 30 ms.

The sonification obstacle-ownship range mapping consists of a range-to-pulse period ring at 75 ft for Caution and 25 ft for Warning with the pulse period linearly scaled 2000 ms (½ Hz) to 250 ms (4 Hz) between them (Figure 6). If the range is lower than 25 ft (i.e., in the Warning zone) the pulse period remains a constant 250 ms, while the pulse duration doubles to 160 ms. If the range is greater than 75 ft, the sonification is muted.

A summary of all Obstacle Sensor sonification mappings is listed in Table 1. A spectrogram time-frequency illustration of the mappings for a "Nearest" obstacle are shown in Figure 7.

Table 1. Sonification Mappings.

Urgency (nearest and second-nearest obstacles)	Two unique sonifier earcon timbres
Obstacle Location	Augmented-reality spatial auditory display (HRTFs, spherical spreading loss gain)
Ownship-Obstacle Azimuth	40 cents per 4 slice pitch reduction front-to-back (to reduce front-back reversals)
Ownship-Obstacle Range (dynamic)	Auditory Looming due to dynamic spherical spreading loss gain changes
Ownship-Obstacle Range	Pulse period and pulse duration

Power Line Sonification

The Power Line sonification uses a recording of a power line [64] as a spatial auditory icon (a sound "image" of the object to which it is referring) that is swept up and down a power line at a rate of 100 ft./sec. The nearest power line segment in the terrain database is found with the closest point chosen as the central location from which to sweep 33 ft. on either side.

For a perpendicular approach and a constant sweep extent, the perceived spatial extent increases as the pilot approaches the power line (Figures 7a and 7b), in a fashion similar to how it appears visually.

For a parallel heading (Figure 7c), the sonification's relative center point remains fixed and produces a longitudinal (front-back) sweep.

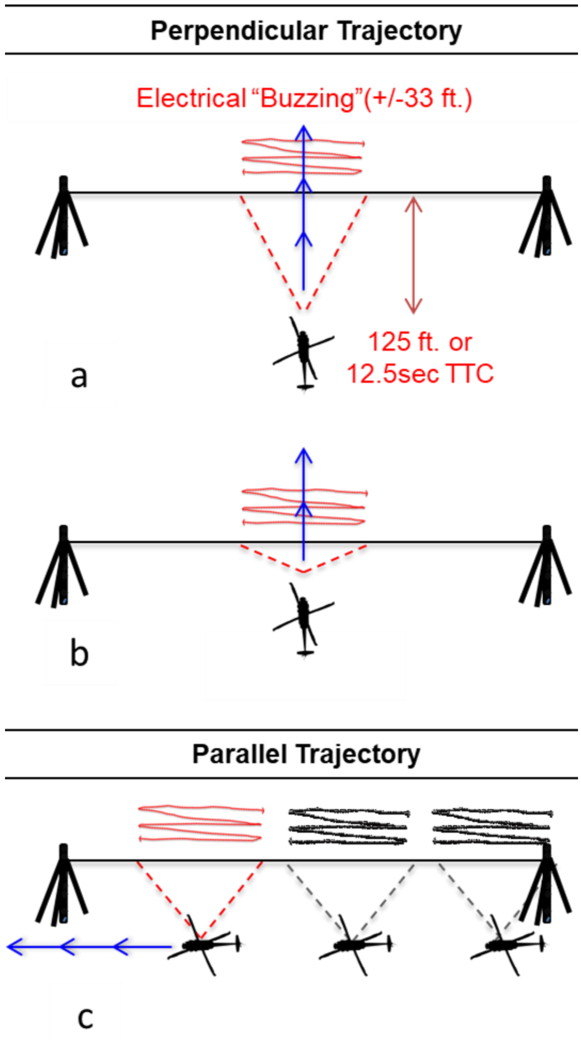


Figure 7. Power Line Database Sonification Concept: fixed 125 ft. sonification below 6 knots, 12.5 sec TTC above 6 knots. a and b: Perpendicular trajectory: c: Parallel trajectory.

Similar to the Obstacle Sensor sonification, the Power Line sonification is fixed-distance below 6 knots, enabling at a distance of 125 ft. from blades, and time to collision above 6 knots, with a TTC time of 12.5 seconds from blades.

5.1.2 ICAD

ICAD can be considered as a Helicopter Terrain Awareness and Warning System (HTAWS) that matches the sonification behavior.
The exocentric 2D top-down orthogonal projection is presented in a track-up configuration.
The display includes (see Figure 8):

- The sensor-detected obstacle locations relative to the helicopter.
- The safety margin indication: caution and warning zones with speed-dependent radii and sensor sweep extent.
- Numerical range of the caution and warning zones (bottom of the display).
- The ground track (magenta line).
- A series of concentric 50-foot wide white rings ranging from 50 ft. to 200 ft. centered on ownship COG as scale indicators.
- A compass (inside 200 ft. ring).

Auditory Display matching elements

Track-up

The exocentric 2D top-down orthogonal projection is displayed in a track-up configuration to facilitate the mapping with ARSAD, the PFD and the OTW egocentric reference frames (ERFs).

Obstacle Range to Color Coding

Obstacle sensor hits are color-coded based on their distance/ TTC to the ownship and superimposed onto the transparent yellow caution zone or red warning zone. If the obstacle is outside the caution and warning zones, and sensed by the sensor, the hits appear green. The layout of the hits allows the pilot to rapidly identify the nature of the obstacles: for example, the urban environment in Figure 8, center, versus the natural environment in Figure 8, bottom. Two vectors are drawn from the helicopter to the two nearest obstacles to match ARSAD behavior: a thick cyan line for the nearest obstacle, and a thin dark blue line for the second-nearest obstacle (Figure 8).

Speed-dependent sensor sweep extent

As ownship speed increases, sensor sweep extent decreases linearly from 360° below 6 knots to +/- 60° at 60 knots. This represents a one-to-one mapping with ARSAD.

Speed-dependent safety margin indicators

Below 6 knots, the zone radii remain constant, yellow Caution at 75 ft. from blades, red Warning at 25 ft. from blades. Above 6 knots, the zone radii change as a function of TTC, yellow Caution at 7.5 sec from blades, red Warning at 2.5 sec from blades. Again, this represents a one-to-one mapping with ARSAD.

Power Line

The Power Line represents a special case of an obstacle in that it is a database object versus a sensed object (though it can be sensed as well). As such, visually, it is attached to the moving terrain map. By default, the nearest power line segment itself is orange, but it also provides altitude-to-go clearance information that is linearly mapped to a red or green region that radiates outward from the power line at 2 Hz (Figure 9). When the helicopter is below clearance altitude, this region is red, when above, green.

Presented at 44th European Rotorcraft Forum, Delft, The Netherlands, 19-20 September, 2018

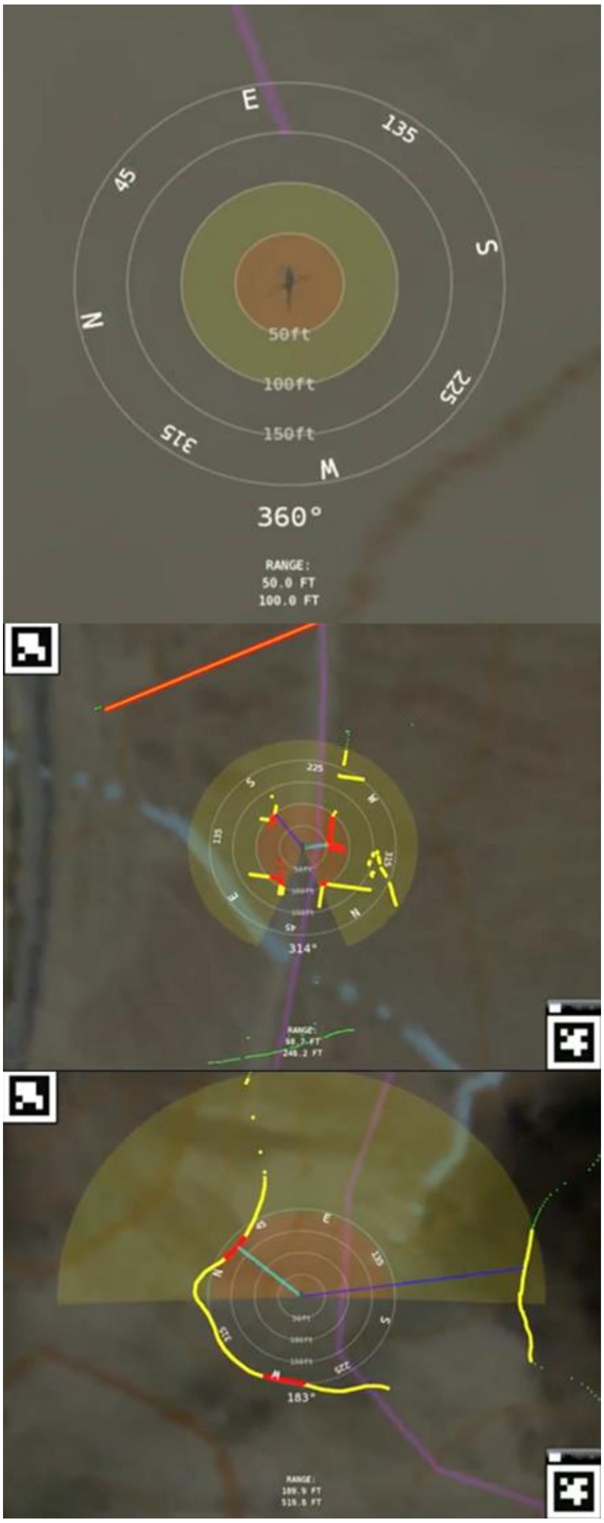


Figure 8. ICAD. Top: speed ≤ 6 knots: 360° sweep extent, fixed-distance safety margin indicators, yellow Caution 75 ft. from blades, red Warning 25 ft. from blades. Center and Bottom: Speed ≤ 6 knots: sweep extent decreases as speed increases, TTC-based safety margin indicators where ring radii increase as a function of speed, Caution 7.5 s from blades, Warning 2.5 s from blades.

Figure 9. Power Line: clearance status is color-coded (red: ownship altitude below line, green: ownship altitude above line). The power line width expands outward from its nominal configuration at a 2Hz cycling rate to capture the pilot's attention. Note that this clearance behavior is not matched auditorily other than by the perception of elevation cues (above or below).

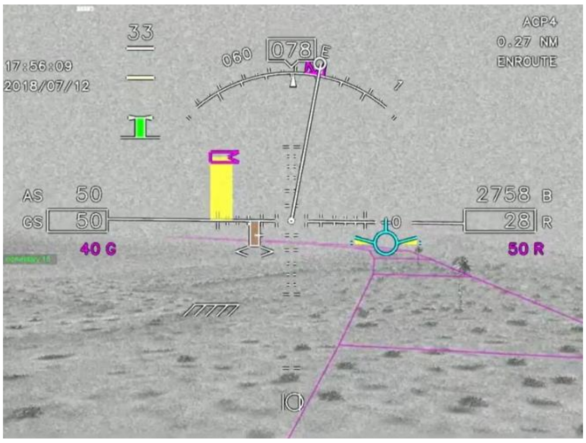


Figure 10. ICE modified PFD symbology superimposed on FLIR imagery.

The cycling rate of the power line was chosen to roughly match the sonification longitudinal sweep rate of 0.8 Hz. Note that the Power Line visual display conveys height information while the auditory display conveys distance information.

5.1.3 Primary Flight Display

The PFD contains actual and modified ICE elements (see Figure 10) with a FLIR image:

- The recommended ground speed and altitude (magenta) as a function of the phase of flight.
- The recommended heading displayed by the *Target Heading Bug* (magenta).
- Altitude information provided by the *Target Altitude Bug* (magenta), error bar, and ground symbol.
- The ground track (magenta) displayed by a flat tunnel-in-the-sky.
- An instantaneous *Flight Path Marker* (FPM).
- Green Landing Zone Beacon (V-like) and Target Landing Zone (X-like).
- ICE aural cues: “vertical speed excessive”, “pull-up”, “10 feet” (too low), “radar tracking” (too high).

5.1.4 The Simulation Facility

The experiment was performed in the SIL NUH-60FS Black Hawk helicopter flight simulator at NASA Ames Research Center.

The RIPTIDE visual image generator was used to simulate natural helicopter environment surroundings.

The selected region, the Fort Irwin National Training Center (NTC, 35°16’17.1’’N, 116°41’32.66’’W), consisted of a series of hilly and flat terrain and two landing zones (LZs), one in an urban area and one in a compound (Figure 11).

The area was populated with natural (trees, canyon walls) and man-made (buildings, towers, high-tension power lines) features that provided ideal obstacle configurations to make use of the 3D auditory cueing. In particular, three power lines were inserted, including two preceding the LZs, which made them more difficult to negotiate.

5.1.5 The Equipment

Pilots were eye-tracked for a forthcoming more extensive data analysis. Head-tracking was required to render the augmented-reality spatial auditory display.

The Ergoneers Dikablis Glasses 3 and a custom prototype Visor Embedded Eye-Tracking System, VEETS, developed for the experiment, were used to track the subject’s pupils.

VEETS is a non-intrusive, non-disruptive device mounted on the optics of an SA photonics HMD. The VEETS eye cameras, infrared LEDs, scene camera, and SA photonics HMD are mounted on a GENTEX HGU-56/P helmet (Figure 12).

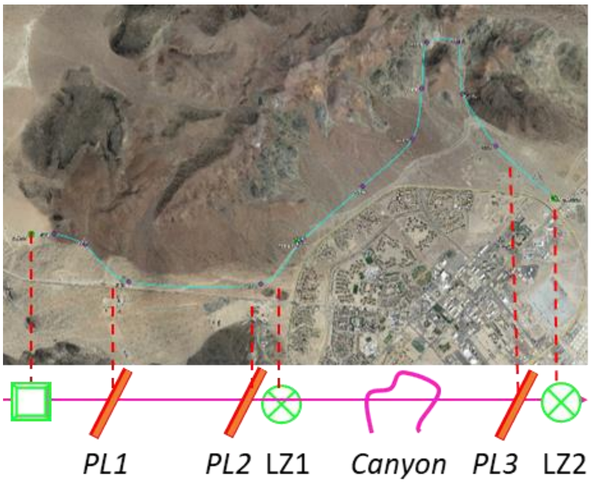


Figure 11. The Experimental Route. Two landing zones, three power lines and a canyon constituted the critical elements of the flight. Trees, hangars and towers were randomly distributed along the path.

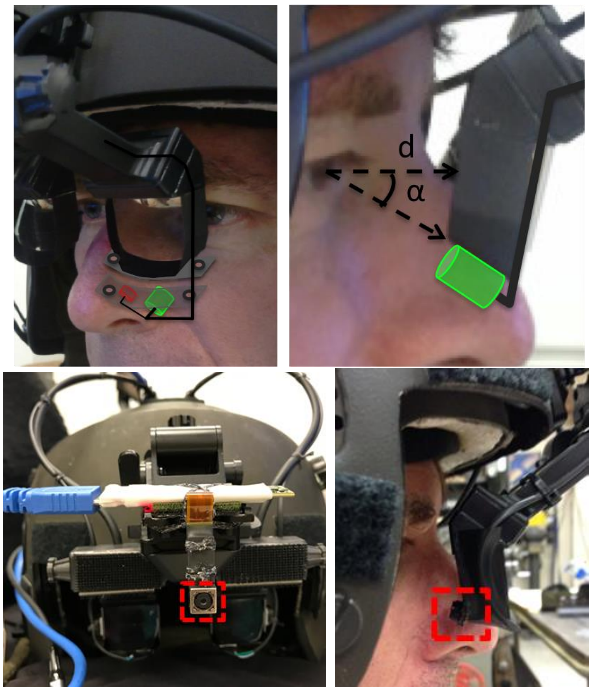


Figure 12. The Visor Embedded Eye-Tracking System (VEETS). Top and bottom-right: eye-cameras and IR LEDs mounted on the optics of the SA Photonics HMD. Bottom-left: the scene camera mounted on a GENTEX HGU-56/P flight helmet.

The audio was presented via Sennheiser HD 700 headphones in the “glasses” condition and via CEP508-SR stereo Communication Ear Plugs (CEPs) in the VEETS condition.

Presented at 44th European Rotorcraft Forum, Delft, The Netherlands, 19-20 September, 2018

The “glasses” condition used the Dikablis Glasses 3 eye tracker with a Polhemus Fastrak head tracker. The VEETS condition used the helmet’s Thales Scorpion Hybrid Optical-based Inertial Tracker (HObIT) head tracker. Both head trackers were sampled at 60 Hz.

5.1.6 The stimuli

Auditory Condition

The nearest, second-nearest, and power line sonifications were displayed in the presence of UH60 cockpit background noise (binaural in-flight recording), ICE aural cues, and pre-recorded cockpit communication chatter.

Visual Condition

The ICAD was presented in conjunction with the PFD and the OTW scene (displayed on three LCD screens) when applicable (low-visibility condition, with a fog level allowing a .25-mile visibility). In the Zero Visibility condition, the OTW scene was black (screen displays off).

5.1.7 The scenario

A single route was selected to provide different cueing conditions as a function of the environment and the phase of the flight.
The scenario can be coarsely decomposed into four segments. During the first phase of the flight, the route is flown at a recommended altitude of 50 ft (±10 ft) and a 40 knots (±10 knots) ground speed (GS). During this phase, the obstacles are essentially towers, palm trees, a hangar, and two power lines. In the second phase, the pilot’s task is to proceed to a landing in an urban area, in close proximity to buildings. During the third phase of the flight, the pilot follows a narrow canyon back and forth at a recommended altitude of 50 ft above ground level (AGL) (±10 ft) and a 20 knots (±10 knots) ground speed (GS). The canyon walls and a few towers constitute essentially the obstacles. The last segment of the flight involves an approach and landing at a compound. A power line is encountered prior to the approach. The obstacles were positioned in such a way that they would require the pilot to eventually take action to avoid them by deviating trajectory or altitude.

5.1.8 The participants

Five UH60 (as a primary aircraft) US Army pilots (all male), aged 32 to 39 (average age 37) participated in the experiment. Mean flight time was 2430 hours (SD=976).

5.1.9 The Experimental Design

All pilots flew a total of 8 runs, with a random presentation order of the dependent variables (Modality: Visual, Auditory, Bimodal, Visibility: Low-Visibility, Zero-Visibility). Conditions were counterbalanced between participants (Table 2). The VEETS/CEPs runs were always performed last.

Table 2. Experimental Design.

		Modality		
Eye-Tracker/ Transducer	Visibility Level	A	V	AV
Glasses/ Headphones	Low-Visibility	1	1	1
	Zero-Visibility	1	1	1
VEETS/ CEPs	Zero-Visibility	NA	NA	2

5.1.10 The Training

A pre-flight briefing was provided to familiarize the pilots with the 3D audio technology, the sonification logic, and sensor model behavior.
For ARSAD, the two virtual obstacle sonification sounds were first demonstrated rotating about the pilot’s head at ear level. The experimental scenario was then experienced in an audio-video recording as flown by another pilot. In particular the critical elements of the route (power lines, LZs, and canyon) were discussed in detail. Pilots were instructed to notice the behavioral analogies between the auditory and the visual displays. They were ultimately briefed about the nature of the mission and how they should interact with the environment:
“You will free-fly in a good visual environment (GVE) without ARSAD/ICCAD.
The PFD will be displayed, but the ground track does not need to be followed.
Once you are sufficiently familiar with the simulator dynamics, you will do an entire GVE run with ARSAD (A) and ICAD (V). A second run is possible if desired.
Your mission is to follow the magenta ground track and maintain the speed and altitude displayed on the PFD. The first landing is planned in an urban area. The mission completes with a landing in a compound. Remain clear of obstacles by deviating from the ground track, speed, and/or altitude. Please report to the experimenter any difficulty in localizing the sonifications or any other issue.
The experimental conditions are randomized. After each run, you will rate your workload using the Bedford Scale. Please feel free to take breaks between runs at your convenience. The experiment completes with the administration of a questionnaire and a debriefing.”
After the briefing, the pilots flew a first time without

Presented at 44th European Rotorcraft Forum, Delft, The Netherlands, 19-20 September, 2018

the visual and the auditory displays to focus on the simulator behavior. Once comfortable with the simulator, they were requested to fly an entire run with ARSAD and ICAD in a high-visibility condition. They were allowed a second run if needed.

6. RESULTS

6.1.1 The Measures of Performance

Objective data

Flightpath Tracking Performance, Vertical Position Error, Lateral Position Error and eye-tracking data have been collected but are not presented in this paper.

Subjective data: Questionnaire

A questionnaire was designed that aims to evaluate the displays dimensions and suggested features. The pilots specified their degree of agreement/disagreement to a total of 44 statements (28 for ARSAD, 8 for ICAD and 8 general items) using a five-level Likert rating scale [Strongly Disagree (1), Disagree (2), Neither Agree nor Disagree (3), Agree (4), Strongly Agree (5)].

Table 3. Post-Demonstration Questionnaire. Display Cueing Dimensions.

	Scale rating 1 (strongly disagree) to 5 (strongly agree)				
	1	2	3	4	5
ARSAD					
Intuitiveness	0%	0%	0%	20%	80%
Realism	0%	0%	6.7%	40%	53.3%
Detectability	0%	0%	0%	0%	100%
Interpretability	0%	0%	20%	40%	40%
Directionality	0%	0%	0%	20%	80%
Discriminability	0%	0%	0%	40%	60%
ICAD					
Intuitiveness	0%	10%	10%	20%	60%
Interpretability	0%	0%	0%	20%	80%
Discriminability	0%	0%	0%	20%	80%

Displays Dimensions

Table 3 contains pilot opinions regarding the “dimensions” of the display cueing elements, such as Intuitiveness, Realism, Detectability, Interpretability, Directionality and Discriminability for ARSAD and Intuitiveness, Interpretability, and Discriminability for ICAD.

One can see from Figure 13, that all the ratings were very high, homogeneous between dimensions, and showed very low within-variability (Friedman rank tests: ARSAD: $X_5^5 = 9.54, p =.08$, ICAD: $X_2^5 = 8, p =.01$), The highest rank was obtained for Detectability and Intuitiveness for ARCAD and Interpretability and Discriminability for ICAD. The ratings were not statistically different between displays (Wilcoxon signed ranks tests: ICAD, ARSAD: Intuitiveness: $Z = -1.28, p = .16$, Interpretability: $Z = -1.73, p = .08$, Discriminability: $Z = -1, p = .31$).

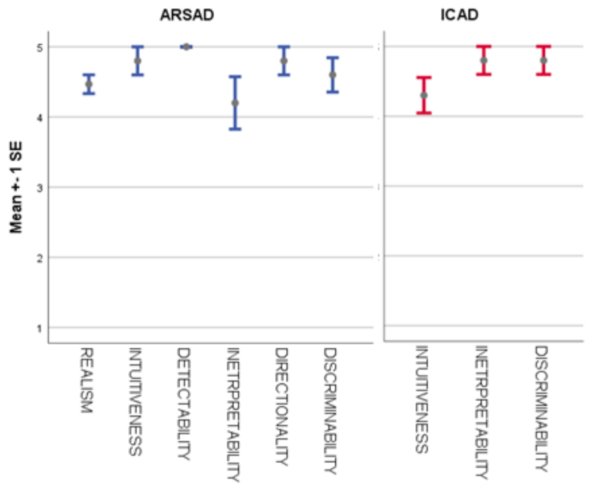


Figure 13. Display Dimensions. Left: ARSAD, Right: ICAD. Vertical Axis: Mean scale rating (1 = strongly disagree, 5 = strongly agree).

The current obstacle sonifications

Overall, the pilots strongly appreciated the sonifications developed for this experiment (see Figure 14, left). The consensus was even higher for the power line than for the obstacle sonifications (Nearest and Second-Nearest), although the difference didn’t reach significance (Obstacle, Power line: $Z = -1.73, p = .08$). Meanwhile, pilots strongly disagreed (80% of the cases) that they *“would have liked to have different types of sonifications representing different types of obstacles”*. This result validates the concept behind representing urgency (Nearest, Second-Nearest) rather than providing information regarding the type of the obstacle (identification). Although the level of the sonifications was considered adequate (60% strongly agreed for the obstacles, 80% strongly agreed for the power line), the pilots were unanimous regarding the fact that a volume and/or ON/OFF switch is mandatory (Figure 14, right),

Presented at 44th European Rotorcraft Forum, Delft, The Netherlands, 19-20 September, 2018

such as in existing Passive Radar Detection system (APR39). They stated "…*obstacle earcons could get tuned out over time…too loud and too frequent...too much beeping"* and commented that an *"acknowledge turn off would be useful – how other verbal alerts handled".* The level of annoyance provided by the sonifications was moderate (a pilot commented the *"sonifications were slightly annoying …as warnings should be!"*). The sonifications were not perceived overly overwhelming, however, a pilot reported that when he *"became saturated, the sound was tuned out except for the power lines since it was more rare".*

Integrated Visual-Auditory Display

When presented in combination, the visual and the auditory information was perceived as congruent (80% strongly agreed, 20% agreed). Pilots essentially disagreed (80% strongly disagreed, 20% disagreed) to the statement that they perceived spatial or temporal discrepancy between the two sensory inputs. Four pilots over five strongly agreed (one strongly disagreed) that obstacle detection was easier in the bimodal (B) than either in the auditory (A) or in the visual (V) conditions. Meanwhile, pilots in majority agreed that obstacle detection was easier with ICAD alone than with ARSAD alone (two pilots strongly disagreed). One pilot who strongly agreed stated that he "*could see obstacles relative to route. Didn't tune out visual but did tune out audio because of alerts too frequent".*

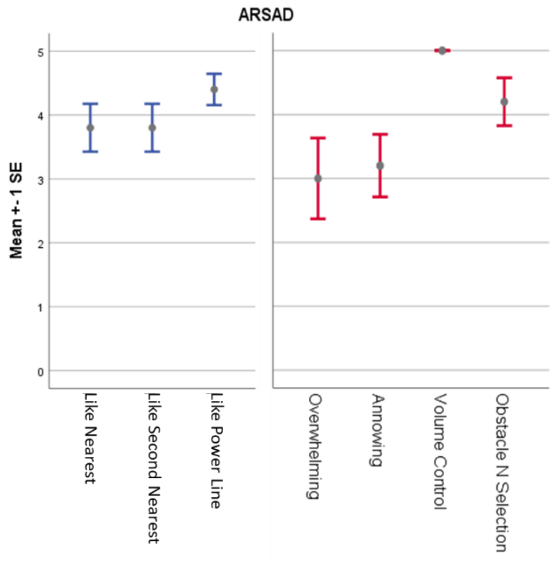


Figure 14. ARSAD Behavior. Left: Obstacle Sonification Evaluation. Right: Workload and proposed workload reduction strategies. Vertical Axis: Mean scale rating (1 = strongly disagree, 5 = strongly agree).

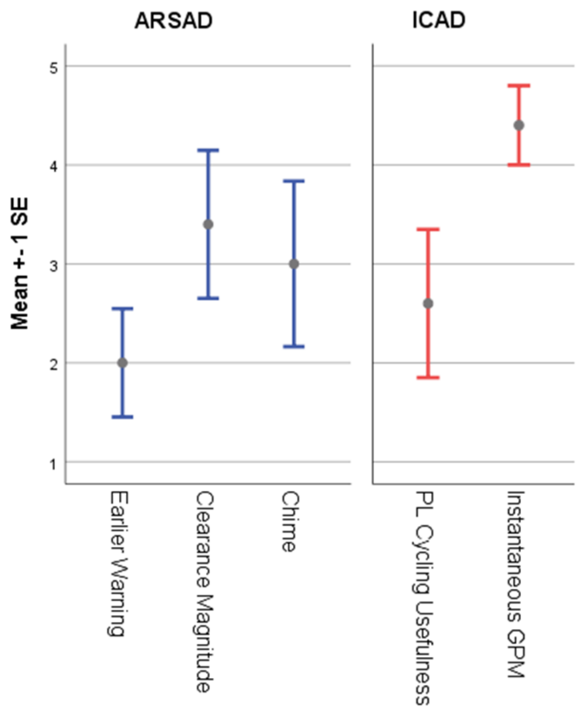


Figure 15. Suggested functionalities for ARSAD, left (I would like to hear the obstacle caution/warning/power line earlier, the power line TTC should include clearance magnitude, I would like to hear a chime prior to the power line earcon), and ICAD, right (the power line cycling was useful, the power line behavior should include instantaneous glide path angle). Vertical Axis: Mean scale rating (1 = strongly disagree, 5 = strongly agree).

Lastly, pilots agreed in 40% of the cases and strongly agreed in 60% of the cases that the combined presentation of ARSAD and ICAD provided a greater sense of immersion. One pilot who agreed to that statement reported that "*audio helped but I wasn't dependent on it – visual helped with route maintenance – didn't need to hear mountains – liked the audio ICE alerts. Spatial audio helped N/W/S/E but not so much distance". "Like combo of A/V, doesn't think A work alone".*

ARSAD

The power line sonification was considered as intuitive and attracted attention ("*caught my attention)".* Most of the pilots reported they *"felt the power line was the only obstacle that needed its own sonification".* Pilots comments were not in agreement regarding whether the power line needed to be perceived earlier ("…*if massive or in a hill*") or later (see Figure 15, left). Most pilots reported they liked hearing the power line before seeing it, as it attracted his attention and directed it

Presented at 44th European Rotorcraft Forum, Delft, The Netherlands, 19-20 September, 2018

toward the visual display. A suggested additional feature was an audio signal (chime or text to speech, TTS) to provide rate of climb and clearance status. Another pilot also commented that, while not of interest in this scenario, *"an aural cueing such as 'wire at 2 o'clock' could be useful"*.

ICAD

The power line behavior was considered less intuitive than with ARSAD (see Figure 13 right, Display Dimensions). Pilots reported they *"couldn't tell thickness well enough to judge the degree of clearance"*. They suggested using *"a yellow intermediate zone like going through an intersection"*. Pilots also recommended (see Figure 15, right) that the display could include the *"clearance angle, rate of climb/ change to clear the obstacle, as an alternative to or in addition to current altitude-to-go"* or a *"glide-slope Indicator, 500 ft/minute vertical-velocity indicator (VVI)"*. Lastly, the cycling feature was not considered an important feature of the power line behavior. A potential explanation is that the cycling rate may have been too low to properly "grab" pilot's attention.

Caution and warning rings, speed-dependent sweep extent (ARSAD and ICAD).

For ICAD, it appears that "*the range of Nearest/ Second-Nearest rings was too big for fidelity at 40 knots"* and that *"there is a "need to zoom in closer to the aircraft".* Overall, pilots would prefer distance to TTC thresholds to be set at 20 knots for visual ("*Below 6 knots distracting", "…at low speed looking for info that was not available"*). For ARSAD, a pilot reported that he liked "*6 knots for the audio* and felt that *"both different* (speeds) *okay"*).
Regarding speed-dependent sweep extent, one pilot reported that there was *"too much information presented from the sides"* and that he *"would like the decrease of zone to occur more quickly"*, adding that *"at 40 knots, everything on the side is distracting".* Accordingly, "*narrower audio rings"* in the left/right axis were suggested. At high-speeds, pilots recommended an earlier obstacle caution, and setting TTC threshold further out for steep terrain, set. One pilot reported that the warning time for an obstacle in the direction of the velocity vector was far too short when a temporarily nearer obstacle to the side was present. This was determined to be due to the behavior of the Tolerance window blocking the display of the further obstacle. This pilot reported that he cared much more about obstacles in the direction of the velocity vector than those to the sides and even recommended a possible change of earcon urgency mapping from Nearest and Second-Nearest to a mapping based on velocity vector regions (see his drawings in Figure 16. Lastly, multiple pilots recommended that the display should include the tail rotor in the distance and TTC thresholds.

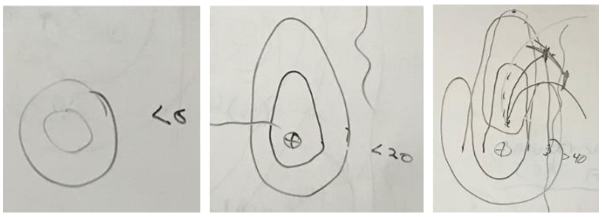


Figure 16. Pilot drawing of a potential revision to the Caution and Warning Zones concept.

Bedford rating scale

Bedford ratings were compared as a function of pilot, display modality (A, V, AV), visibility level (LV, ZV) and eye-tracking system (in the AV display modality and the Zero-Visibility condition) (Figure 17).
Pilot rating variability was highly significant (Friedman Ranks Test: $X_8^3 = 19.53, p < .0001$) and reflected both the differences in magnitude of the perceived workload and its variability. Overall, the perceived workload was relatively high, with little to very little spare capacity. For display modality, workload rating is lower in the AV than in the V or A condition, although the difference didn't reach statistical significance (A: $\mu = 6.75, SD = 1.90$, V: $\mu = 6.38, SD = 1.06$, AV: $\mu = 5.75, SD = 1.48$, Friedman Test: $X_8^2 = 3.39, p = .18$).

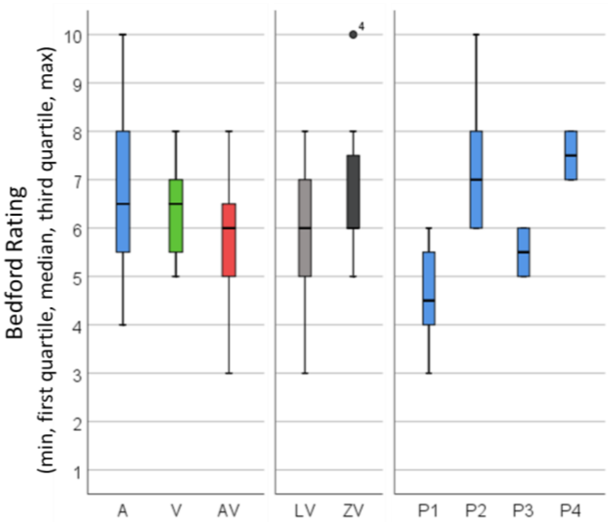


Figure 17. Left: Bedford rating as a function of Display type (A=ARSAD, V=ICAD, AV=ARSAD + ICAD). Center: Bedford rating as a function of Visibility Level, Low vs. Zero. Right: Interindividual variability (one pilot's data missing).

Bedford rating was significantly higher for ZV than for LV environment (LV: $\mu = 5.92, SD = 1.56$, ZV: $\mu = 6.67, SD = 1.43$, Wilcoxon Signed Ranks Test: ZV, LV: $Z = -2.12, p = .03$). Lastly, there was no significant effect of eye-tracking condition (Glasses: $\mu = 6.25, SD = 1.25$, VEETS: $\mu = 6, SD = 1.68$, Wilcoxon Signed Ranks Test: Glasses, VEETS: $Z = 0, p = 1$).

7. FUTURE WORK

Towards a new topology of the Obstacle Sensor Model and sonification logic

Based on trial observations, questionnaire data and pilot feedback, a variety of improvements have been suggested for the obstacle sensor model and sonification logic.
The primary pilot concerns were that the sides were too responsive and that direction of flight alerts were, at times, too late. It was also observed that the tolerance window could have a deleterious effect when treating distances equally about ownship, even with reduced sweep extent. An obstacle to the side of the direction of flight could preempt and omit a close but further away obstacle in the direction of flight.
Lastly, the current design's sweep extent is centered on heading, which can introduce obstacle consideration errors when the heading and the velocity-vector direction differ.
It follows that taking the velocity vector into account would help address these issues. We will use a threat scale from 0 to 1, where 0 corresponds to the outside caution ring in the distance-based model and 1 to the blade radius (Figure 18 A). A simple threat mapping for angular threat is to then take the cosine of the ownship-obstacle vector velocity-vector projection with the ownship-obstacle vector. This threat space is then scaled by the distance model to decrease with distance (Figure 18 C).
Since, at lower speeds, an omnidirectional paradigm might still be advantageous, a weighting of the two methods can be used (Figure 18 B). Notice that the negative cosine values can create a threat rejection region similar to the previous sweep extent algorithm, albeit illustrated at low speeds in this example (fixed distance versus time to collision). To make the threat space better follow the contour of the velocity-vector cosine-based threat space, thresholding can be applied (Figure 18 D). Time to collision and velocity vector magnitude extend the threat space in the direction of the velocity vector (Figure 18 E).
Lastly, the threat level of obstacles close to the velocity vector can be emphasized by warping orthogonal distance calculations (Figure 18 F).

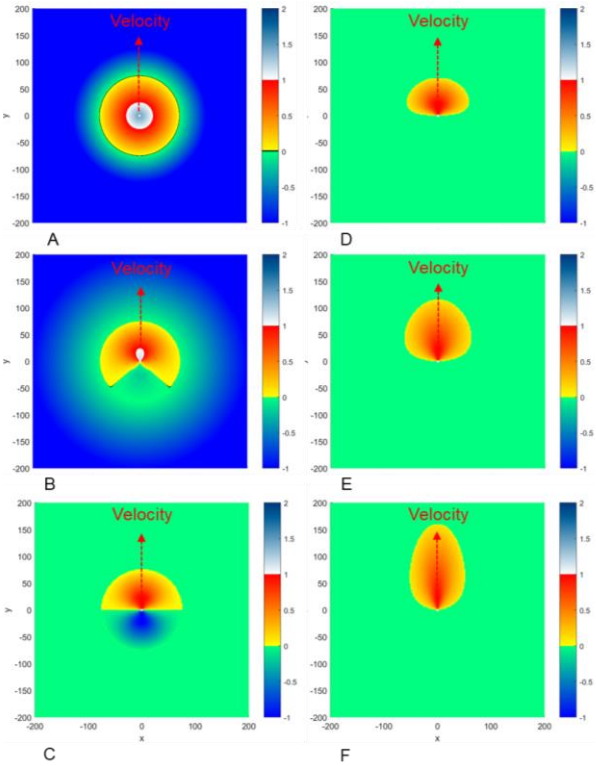


Figure 18. Horizontal slices through threat space with the y axis in the direction of the velocity vector, x and y axes in feet, and the color axis in threat magnitude. (A) current omnidirectional distance-based threat magnitude, (B) sum of weighted (A) and (C) methods, (C) distance-scaled velocity-vector cosine threat magnitude, (D) with threat-magnitude thresholding, (E) with time to collision increasing the distance threshold, and (F) with increased speed and orthogonal warping.

Note, that although these techniques are illustrated in a horizontal plane, they naturally extend to three dimensions. The development of this algorithm is in progress. Areas for refinement include the mapping of velocity-vector magnitude to the various parameters discussed (weights, thresholds, and warping) and the reexamination of tolerance and earcon mapping (e.g., distance to ownship versus distance to flight path).

8. CONCLUSIONS

This paper documents the development, implementation, and evaluation of an augmented-reality spatial auditory display (ARSAD) and its visual analog, ICAD, to render a hybrid sensor-data-based augmented environment. This research is part of the ADD degraded visual environment mitigation (DVE-M) program where

Presented at 44th European Rotorcraft Forum, Delft, The Netherlands, 19-20 September, 2018

the objective is to provide an integrated (visual, auditory, tactile) cueing solution to extend the current operational envelope for all visual environments while enhancing flight safety.

As part of this work, significant effort was devoted to the development of an Aviation Auditory Display Engine (AvADE) and the experimental study of advanced auditory display concepts [3]. AvADE was used to develop the ARSAD earcons and sonifications that address the specific needs of auditory cueing for obstacle avoidance. For the purpose of the obstacle avoidance task, earcons were preferred to auditory icons to direct the pilot's focus on the obstacle's relative position to ownship rather than its identification. The pilot questionnaire results and comments clearly validate this approach. The two selected obstacle sonification earcons were designed to enable the concurrent presentation of the two nearest obstacles and maximize their perceptual separability. To further accentuate obstacle azimuth angle relative to ownship and relative to one another, a sonifier pitch scaling algorithm and a sensor model tolerance window were developed, respectively. Because pitch scaling reinforces HRTF shadowing, it aims to reduce the occurrence of front/back reversal [65], a phenomenon caused by the roughly spherical shape of the head and the primary role of ITD and ILD localization cues in azimuth. The sensor model tolerance window also strives to present the two nearest obstacles instead of two "hits" belonging to the same obstacle. In the present demonstration, the 90° tolerance window allowed the pilots to successfully identify the respective direction of the obstacles when two were present at the same time.

The other part of the work was the development of an isomorphic visual display, ICAD, a necessary component of the ICE suite. Indeed, ARSAD is not intended to be used without its visual counterpart, except in off-nominal situations where the visual information would be momentary unavailable. In its present configuration, ICAD positions itself as a head-down MFD *(JDM – MFD and PFD need defs)*, but could ultimately be integrated with the PFD, in a head-up display configuration. As evaluated in this experiment, ICAD accomplished its role of complementary display, in the sense that it provided synergistic information with ARSAD. The most recommended display configurability enhancements relate to the Obstacle Sensor Model and sonification logic with the velocity vector being emphasized for integration.

From Bimodal to Trimodal

The last phase of the DVE-M integration program concerns the integration of visual, aural, spatial auditory, and tactile information. The inherent structural and functional differences between vision, audition and touch have important implications for multisensory integration, even more in virtual environments where system latency is susceptible to produce spatio-temporal incongruencies. Indeed, to lead to a unified percept, the sensory inputs must be spatially (Stein & Meredith, 1993) and temporally congruent [66]; [67], i.e. respectively perceived at the same place and at the same time. With complex, dynamic stimuli, the tolerance to visual-auditory spatial discrepancy is relatively high., up to ~ ±15° horizontal separation. The variety of the loci for tactile stimuli makes it more difficult to determine a unique fusion limit with auditory and visual information.

Humans are much more sensitive to temporal discrepancy or asynchrony. Asynchrony thresholds, however, are more difficult to determine than their spatial counterparts. Auditory-tactile delays (25ms) appear to be more critical than auditory-visual delays (ranging from 50 to 250 ms as a function of the nature of the stimulus).

A variety of other factors influence the perception of unity between the sensory inputs. For example, the frequency of the sound and the frequency of the vibration are coupled to each other by physical laws. Similarly, in regards to perception, human response to vibration (or to tactile feedback) and sound is strongly dependent on the frequency of the stimulus.

Of greater importance is the congruence between the informational content of the modalities. Indeed, at the cognitive level, it is important that the presented information in one modality finds its analog in another modality. The power line visual representation on ICAD, its display on the PFD and its associated sonification provide congruent information about the same object.

Visual and auditory integration was successfully tested in the present experiment. The relative contribution and potential combination of auditory and tactile information remains to be determined to prevent information overload and inadequate pairing. Gaps in the knowledge of how tactile information can be useful to determine egocentric direction persist.

Two forthcoming experiments will evaluate 1) the new velocity vector-based sonification logic (NASA UH-60 SIL) and 2) a trimodal visual-auditory-tactile obstacle avoidance display (USAARL immersive, full-motion, enhanced brownout NUH-60FS Black Hawk helicopter flight simulator). The intended logic behind this multimodal ICE is complementarity rather than redundancy, at least for the auditory and tactile components of the display. Currently, the tactile component of the display would represent an "ultimate" warning, after visual and auditory warnings would have failed to correct the pilot's obstacle avoidance trajectory.

Presented at 44th European Rotorcraft Forum, Delft, The Netherlands, 19-20 September, 2018

This sequential rather than parallel presentation mode has been selected to reduce the potential workload resulting from the division of attention between the different sensory modalities, in particular for the auditory and tactile components which may not provide congruent information regarding the obstacle position in relation to the ownship. Indeed, in the auditory modality, the obstacle position in relation to the ownship (for example, an obstacle located to the left of the ownship) would indicate to the pilot to adopt a trajectory opposite to that of the obstacle, information which is congruent with that of an identically-oriented visual display (track-up or North-up). Meanwhile, it is still matter of debate whether the tactile display should indicate the direction of the obstacle or the direction to maneuver to avoid the obstacle. These two options need to be tested both in serial and parallel presentation with the visual and/or auditory displays.

The results of these two experiments, with two different simulator environments, are expected to support the last phase of the ICE DVE-M program (scheduled for 2020), which will culminate with the integration of the multimodal integrated display suite in real flight conditions. *To this end, the experimental UH-60 helicopter will be equipped with three short-range radars that will capture the obstacles in the three dimensions of space (left/right, up/down, forward, backward), allowing for an exhaustive depiction of the threats/ wingmen in the helicopter operational environment. Ultimately, the short-range radar data should be merged with EVS data, such as LIDAR/LADAR*, to maximize obstacle detection and avoidance behavior, while keeping the workload at its minimum level.

Acknowledgments

This work is a joint effort between the U.S. Army and NASA Ames Research Center. Special thanks to the pilots for their time and their very valuable inputs, and to Nathan Mielcarek, Nelson Iwai and John Archdeacon for their support.

Biographies

Dr. Martine Godfroy-Cooper is a senior Research Psychologist at NASA ARC, CA. Her research interests encompass multisensory integration, advanced controls and displays and perception-action coupling.

Mr. Joel Miller is a senior Research Engineer at NASA ARC, CA, with an emphasis in real-time audio signal processing, virtual acoustic environment modeling, and immersive multimodal displays.

Dr. Edward Bachelder is a senior Research Engineer at ADD, Moffett Field, CA. His domain of research includes pilot controls and workload modelling.

Dr. Elizabeth Wenzel is a senior Research Psychologist at NASA ARC, CA. Her current research interests include spatial auditory and multimodal displays.

References

[1] Feltman, K. A., Kelley, A. M., Curry, I. P., Boudreaux, D. A., Milam, L., Mathews, C., & Russell, D. (2017). *Review of US Army Aviation Accident Reports: Prevalence of Environmental Stressors and Medical Conditions* (No. 2018-02). Army Aeromedical Research Lab, Fort Rucker, AL.

[2] Godfroy-Cooper, M., J. D. Miller, Z. Szoboszlay, and E. M. Wenzel. "3D-Sonification for Obstacle Avoidance in Brownout Conditions." (2017). Proceedings of the 73th American Helicopter Society Annual Forum & Technology Display, Fort Worth, Texas, USA, May 9-11, 2017.

[3] Miller, J. D., Godfroy-Cooper, M., Wenzel, E.M. (2017). ARSAD: An Augmented-Reality Spatial Auditory Display for Obstacle Avoidance for all Phases of Flight, Proceedings of the 74th of the American Helicopter Society Annual Forum & Technology Display, Phoenix, Arizona, USA, May 14-17, 2018.

[4] Sanderson P., Anderson J., & Watson M., "Extending ecological interface design to auditory displays," Proceedings of the 10th Australasian Conference on Computer-Human Interaction, Dec. 4th 2000, pp. 259-266.

[5] Hermann, T. (2008). Taxonomy and definitions for sonification and auditory display. Proceedings of the 14th International Conference on Auditory Display (ICAD 2008).

[6] Nees, M. A., & Walker, B. N., *Auditory interfaces and sonification*, The universal access handbook, 2009, pp. 507-521.

[7] Bregman, A. S. *Auditory Scene Analysis: The perceptual organization of sound*, MIT Press, Cambridge, MA, 1994.

[8] Kramer, G., Walker, B., Bonebright, T., Cook, P., Flowers, J. H., Miner, N., & Neuhoff, J. (2010). Sonification report: Status of the field and research agenda.

[9] Watson, M., & Sanderson, P. (2004). Sonification supports eyes-free respiratory monitoring and task time-sharing. *Human factors*, *46*(3), 497-517.

[10] Sanderson, P. M., Watson, M. O., & Russell, W. J. (2005). Advanced patient monitoring displays: tools for continuous informing. *Anesthesia & Analgesia*, *101*(1), 161-168.

[11] Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. *Human–Computer Interaction*, *4*(1), 11-44.

[12] Shepard, R. N. (1981). Psychological relations and psychophysical scales: On the status of “direct” psychophysical measurement. *Journal of Mathematical Psychology*, *24*(1), 21-57.
[13] Begault, D. R. (1993). Head-up auditory displays for traffic collision avoidance system advisories: A preliminary investigation. *Human factors*, *35*(4), 707-717.
[14] Bronkhorst, A. W., Veltman, J. A., & Van Breda, L. (1996). Application of a three-dimensional auditory display in a flight task. *Human factors*, *38*(1), 23-33.
[15] Simpson, B. D., Bolia, R. S., & Draper, M. H. (2013). Spatial audio display concepts supporting situation awareness for operators of unmanned aerial vehicles. *Human Performance, Situation Awareness, and Automation: Current Research and Trends HPSAA II*, 61.
[16] Wenzel, E. M., Godfroy, M., & Miller, J. D. (2012, October). Prototype spatial auditory display for remote planetary exploration. In *Audio Engineering Society Convention 133*. Audio Engineering Society.
[17] Loomis, J. M., Marston, J. R., Golledge, R. G., & Klatzky, R. L. (2005). Personal guidance system for people with visual impairment: A comparison of spatial displays for route guidance. *Journal of visual impairment & blindness*, *99*(4), 219.
[18] Beattie, D., Baillie, L., & Halvey, M. (2017). Exploring How Drivers Perceive Spatial Earcons in Automated Vehicles. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *1*(3), 36.
[19] Wenzel, E. M. & Begault, D. R. (2009), “CrewSound: An Auditory Display System for the Study of Advanced Cockpit Communications”, report for NASA Intelligent Integrated Flight Deck project.
[20] Begault, D. R., & Pittman, M. T., “Three-dimensional audio versus head-down traffic alert and collision avoidance system displays,” *The International journal of aviation psychology*, Vol. 6(1), Jan. 1st 1996, pp. 79-93.
[21] Haas, E. C., “Can 3-D auditory warnings enhance helicopter cockpit safety?” Human Factors and Ergonomics Society Annual Meeting Proceedings, Oct. 1st 1998, Vol. 42, (15), pp. 1117-1121). SAGE Publications.
[22] Bastide, R., Navarre, D., Palanque, P., Schyn, A., & Dragicevic, P. (2004, October). A model-based approach for real-time embedded multimodal systems in military aircrafts. In *Proceedings of the 6th international conference on Multimodal interfaces* (pp. 243-250). ACM.
[23] Vidulich, M. A., Wickens, C. D., Tsang, P. A., & Flach, J. M. (2010). Information processing in aviation. *Human factors in aviation*, 175-215.
[24] Yost, W. A., & Dye, R. H. (1997, November). Fundamentals of directional hearing. In *Seminars in Hearing* (Vol. 18, No. 04, pp. 321-344). Copyright© 1997 by Thieme Medical Publishers, Inc.
[25] Buser, P., & Imbert, M. (1987). Vision, éd. *Hermann, Paris*.
[26] Mills, A. W. (1958). On the minimum audible angle. *The Journal of the Acoustical Society of America*, *30*(4), 237-246.
[27] Blauert, J. (1997). *Spatial hearing: the psychophysics of human sound localization*. MIT press.
[28] Warren, R. M. (2013). *Auditory perception: A new synthesis* (Vol. 109). Elsevier.
[29] Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual review of psychology*, *42*(1), 135-159.
[30] Wightman, F. L., & Kistler, D. J. (1999). Resolution of front–back ambiguity in spatial hearing by listener and source movement. *The Journal of the Acoustical Society of America*, *105*(5), 2841-2853.
[31] Hofman, P.M., & Van Opstal, A.J. (2003). Binaural weighting of pinna cues in human sound localization. *Experimental brain research*, 148(4), 458-470.
[32] Makous, J. C., & Middlebrooks, J. C. (1990). Two-dimensional sound localization by human listeners. *The journal of the Acoustical Society of America*, *87*(5), 2188-2200.
[33] Bronkhorst, A. W., & Houtgast, T. (1999). Auditory distance perception in rooms. *Nature*, *397*(6719), 517.
[34] Zahorik, P. (2002). Auditory display of sound source distance. In *Proc. Int. Conf. on Auditory Display* (pp. 326-332).
[35] Serino, A., Canzoneri, E., & Avenanti, A. (2011). Fronto-parietal areas necessary for a multisensory representation of peripersonal space in humans: an rTMS study. *Journal of cognitive neuroscience*, *23*(10), 2956-2967.
[36] Kopčo, N., & Shinn-Cunningham, B. G. (2011). Effect of stimulus spectrum on distance perception for nearby sources. *The Journal of the Acoustical Society of America*, *130*(3), 1530-1541.
[37] Bach, D. R., Neuhoff, J. G., Perrig, W., & Seifritz, E. (2009). Looming sounds as warning signals: The function of motion cues. *International Journal of Psychophysiology*, *74*(1), 28-33.
[38] Shaw, B. K., McGowan, R. S., & Turvey, M. T. (1991). An acoustic variable specifying time-to-contact. *Ecological Psychology*, *3*(3), 253-261.
[39] Rosenblum, L. D., Wuestefeld, A. P., & Saldana, H. M. (1993). Auditory looming perception: Influences on anticipatory judgments. *Perception*, *22*(12), 1467-1482.
[40] Neuhoff, J. G. (2001). An adaptive bias in the perception of looming auditory motion. *Ecological Psychology*, *13*(2), 87-110.
[41] Carlile, S. (1996). The physical and psychophysical basis of sound localization. In

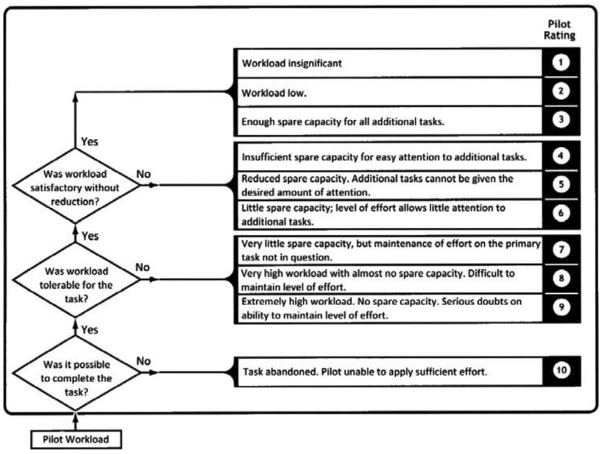
Virtual auditory space: Generation and applications (pp. 27-78). Springer, Berlin, Heidelberg.
[42] Wenzel, E. M., Arruda, M., Kistler, D. J., & Wightman, F. L., “Localization using non-individualized head-related transfer functions,” *The Journal of the Acoustical Society of America*, Vol. 94(1), Jul. 1st 1993, pp.111-23.
[43] Kramer, L. J., Harrison, S. J., Bailey, R. E., Shelton, K. J., & Ellis, K. K. (2014, June). Visual advantage of enhanced flight vision system during NextGen flight test evaluation. In *Degraded Visual Environments: Enhanced, Synthetic, and External Vision Solutions 2014* (Vol. 9087, p. 90870G). International Society for Optics and Photonics.
[44] Peinecke, N., Schmerwitz, S., Döhler, H. U., & Lüken, T. (2017). Review of conformal displays: more than a highway in the sky. *Optical Engineering*, *56*(5), 051406.
[45] Levine, M., Jankovic, I. N., & Palij, M. (1982). Principles of spatial problem solving. *Journal of Experimental Psychology: General*, *111*(2), 157.
[46] Aretz, A. J. (1991). The design of electronic map displays. *Human Factors*, *33*(1), 85-101.
[47] Prinzel, L.J. & Kramer, L.J. Synthetic vision systems. Retrieved from https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20090007635_2009006413.pdf
[48] Mämpel, M., Münsterer, T., Kramper, P., Kress, M., & Wegner, M. (2011). A new approach to a smart helicopter terrain awareness and warning system.
[49] Vreeken, J. (2013). Helicopter flight in a Degraded Visual Environment.NLR-TP-2013-559.
[50] Viertler, F. X. (2017). *Visual Augmentation for Rotorcraft Pilots in Degraded Visual Environment* (Doctoral dissertation, Technische Universität München).
[51] Brunetti, M. (2014). The Guardian project: the reasons, concept and advantages of a novel obstacle proximity LIDAR system.
[52] Sutcliffe, A. (2003). *Multimedia and virtual reality: designing multisensory user interfaces*. Psychology Press.
[53] Sarter, N. B. (2006). Multimodal information presentation: Design guidance and research challenges. *International journal of industrial ergonomics*, *36*(5), 439-445.
[54] Dettmann, A., & Bullinger, A. C. (2017). Spatially distributed visual, auditory and multimodal warning signals–A comparison. *Proceedings of the Human Factors and Ergonomics Society Europe*, 185-199.
[55] Vatakis, A., & Spence, C. (2007). Crossmodal binding: evaluating the “unity assumption” using audiovisual speech stimuli. *Perception & Psychophysics, 69*(5), 744–756.
[56] Miller, J.D., Abel, J.S. & Wenzel, E.M. (1999) Implementation issues in the development of a real-time, Windows-based system to study spatial hearing. Journal of the Acoustical Society of America, 105, 1193.
[57] http://www.urs-simulation.com/products/xig_image_generator.aspx
[58] Pierce, A. D., & Beyer, R. T. (1990). Acoustics: An Introduction to Its Physical Principles and Applications. 1989 Edition.
[59] Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. *Human–Computer Interaction*, *4*(1), 11-44.
[60] Patterson, R. D. (1982). *Guidelines for auditory warning systems on civil aircraft*. Civil Aviation Authority.
[61] Edworthy, J., Loxley, S., & Dennis, I. (1991). Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Human factors*, *33*(2), 205-231.
[62] Brewster, S. A., Wright, P. C., & Edwards, A. D. (1995). Experimentally derived guidelines for the creation of earcons. In *Adjunct Proceedings of HCI* (Vol. 95, pp. 155-159).
[63] Hellier, E. J., Edworthy, J., & Dennis, I. A. N. (1993). Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency. *Human factors*, *35*(4), 693-706.
[64] https://freesound.org/people/felix.blume/sounds/169942/, 169942__felix-blume__electricity-going-trough-a-post-in-the-desert.wav.
[65] Wenzel, E. M., Arruda, M., Kistler, D. J., & Wightman, F. L., “Localization using non-individualized head-related transfer functions,” *The Journal of the Acoustical Society of America*, Vol. 94(1), Jul. 1st 1993, pp.111-23.
[66] Diederich, A., & Colonius, H. (2015). The time window of multisensory integration: Relating reaction times and judgments of temporal order. *Psychological review*, *122*(2), 232.
[67] Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: a tutorial review. *Attention, Perception, & Psychophysics*, *72*(4), 871-884.

Annexes

Annex 1. Bedford Rating Scale

The Bedford rating scale is a three-rank ordinal structure used to assess pilot workload defined as: “… the integrated mental and physical effort required to satisfy the perceived demands of a specified flight task” (Roscoe, 1984). The concept of spare capacity is used to help define levels of workload.

Bedford Workload Rating Scale



Annex 2. Post Experimental Questionnaire

	1	2	3	4	5	NA
	Strongly Disagree	Disagree	Neither Disagree nor Agree	Agree	Strongly Agree	Not applicable
SPATIAL AUDITORY DISPLAY						
The spatial audio was realistic (sounds sounded external, responded well to head motion, etc…)	○	○	○	○	○	○
The SAD was intuitive	○	○	○	○	○	○
The SAD provided a realistic perception of space	○	○	○	○	○	○
The audio level of the obstacle sonifications was adequate	○	○	○	○	○	○
The audio level of the power line sonification was adequate	○	○	○	○	○	○
I liked the nearest obstacle earcon sound	○	○	○	○	○	○
I liked the second-nearest obstacle earcon sound	○	○	○	○	○	○
I liked the power line auditory icon sound	○	○	○	○	○	○
The sonifications were clearly noticeable	○	○	○	○	○	○
The sonifications were noticeable during motion	○	○	○	○	○	○
I could immediately interpret the sonifications	○	○	○	○	○	○
The direction of the obstacle(s) was clear	○	○	○	○	○	○
The distance behavior of the obstacles was realistic	○	○	○	○	○	○
I would like to hear the obstacle caution earlier	○	○	○	○	○	○
I would like to hear the obstacle warning earlier	○	○	○	○	○	○
I would like to hear the power line earlier	○	○	○	○	○	○
The power line TTC should include clearance magnitude	○	○	○	○	○	○
I would like to hear a chime prior to the power line earcon	○	○	○	○	○	○
The sonifications were overwhelming	○	○	○	○	○	○
The sonifications were annoying	○	○	○	○	○	○
I would prefer to hear only the nearest obstacle	○	○	○	○	○	○
I would like to be able to turn the spatial audio display on/off	○	○	○	○	○	○
I would like to be able to select the number of obstacles presented (e.g. 1 or 2)	○	○	○	○	○	○
I liked the speed-dependent caution and warning rings (distance-based below 6knts and TTC-based above 6knts)	○	○	○	○	○	○
I would like to be able to select the distance values of the obstacle caution and warning thresholds	○	○	○	○	○	○
I would like to be able to select the TTC time values of the obstacle caution and warning thresholds	○	○	○	○	○	○
I could easily discriminate two obstacles when presented at the same time	○	○	○	○	○	○
I would like to have different sonifications representing different types of obstacles (a pole, a vehicle, etc.)	○	○	○	○	○	○
VISUAL DISPLAY						
Overall, the visual display was intuitive	○	○	○	○	○	○
The colors used for the caution and warning zones were adequate	○	○	○	○	○	○
The transparency used for the caution and warning zones allowed me to clearly discern the obstacle zones, the obstacle hits and the terrain	○	○	○	○	○	○
The power line clearance behavior was intuitive (thickness and color change)	○	○	○	○	○	○
The power line cycling was useful	○	○	○	○	○	○
The power line cycling rate was too low	○	○	○	○	○	○
The power line behavior should include instantaneous glide path angle	○	○	○	○	○	○
I liked the speed-dependent caution and warning zones (distance-based below 6knts and TTC-based above 6knts)	○	○	○	○	○	○
BIMODAL AND GENERAL						
The visual display and the auditory display provided congruent information	○	○	○	○	○	○
I perceived spatial discrepancy between the visual and the auditory information	○	○	○	○	○	○
I perceived asynchrony between the visual and the auditory information	○	○	○	○	○	○
I liked the speed-dependent scan extent	○	○	○	○	○	○
The obstacle detection was easier in the bimodal than in the visual condition	○	○	○	○	○	○
The obstacle detection was easier in the bimodal than in the auditory condition	○	○	○	○	○	○
The obstacle detection was easier in the visual than in the auditory condition	○	○	○	○	○	○
The bimodal display gave me a greater sense of immersion	○	○	○	○	○	○

Presented at 44th European Rotorcraft Forum, Delft, The Netherlands, 19-20 September, 2018