Improving Helicopter Simulator Fidelity with Augmented Reality

Development and Testing of a Modified Flight Research Simulator where Virtual Windows are made visible through AR Goggles

Alessandro Pardi
8 October 2020

Faculty of Aerospace Engineering
8 October 2020

Improving Helicopter Simulator Fidelity with Augmented Reality

Development and Testing of a Modified Flight Research Simulator where Virtual Windows are made visible through AR Goggles

Master of Science Thesis

For obtaining the degree of Master of Science in Aerospace Engineering at Delft University of Technology

A. Pardi

8 October 2020

Faculty of Aerospace Engineering · Delft University of Technology

Delft University of Technology
Control and Simulation Section

The undersigned hereby certify that they have read and recommend to the Faculty of Aerospace Engineering for acceptance a thesis entitled **"Improving Helicopter Simulator Fidelity with Augmented Reality"** by **A. Pardi** in partial fulfillment of the requirements for the degree of **Master of Science**.

Dated: 8 October 2020

Readers:

prof.dr.ir. M. Mulder

prof.dr. M. M. van Paassen

ir. O. Stroosma

dr. M.A. Mitici

Contents

Part I

Paper

Improving Helicopter Simulator Fidelity with Augmented Reality

A. Pardi*
TU Delft, Faculty of Aerospace Engineering, the Netherlands

Supervised by: O. Stroosma†, M. M. van Paassen‡, M. Mulder§
TU Delft, Faculty of Aerospace Engineering, the Netherlands

ABSTRACT

This project investigates the use of Augmented Reality (AR) in the context of research flight simulators to increase their visual fidelity in rotary-wing applications. The field of view of existing research simulators is often designed after fixed-wing aircraft and can be largely inferior to what is required to simulate helicopters. In particular, visibility downward and to the far side of the cockpit is limited, reducing the field of view of the outside environment available to the pilot. This can be potentially corrected by means of AR goggles adding virtual chin and side windows, supplementing those present in the simulator's cabin. The visual system of the SIMONA Research Simulator (SRS) was complemented with AR goggles generating additional outside visuals, and this prototype was tested in a human-in-the-loop experiment where flight performance, workload and motion sickness parameters were evaluated in a precision piloting task, with and without the AR visuals.

Results confirmed that the prototype was functionally working, with the selected piloting task having been completed both with and without the added AR visuals. However, further analysis found no significant effects of the AR visuals' presence on flight performance, workload and motion sickness. It was concluded that the high levels of noise found in the data were caused by a mismatch between the piloting task and the participants population.

NOMENCLATURE

ADS	=	Aeronautical Design Standard
AR	=	Augmented Reality
DERP	=	Design Eye Reference Point
DM	=	Dependent Measure
IV	=	Independent Variable
MISC	=	MIsery SCale
RMS	=	Root Mean Square
SRS	=	SIMONA Research Simulator
SSQ	=	Simulator Sickness Questionnaire
TLX	=	Task workLoad indeX
SIMONA	=	SImulation, MOtion and NAvigation

INTRODUCTION

Flight simulators have been used in the aerospace field since its early days, with many different users requiring each a different implementation. In the field of scientific research, the ability of a simulator to be versatile and adaptable is of great importance, as most research projects deal with innovation and experimentation on a daily basis. Modifications are then needed for such components as the flight model, the flight controls, the instrumentation and the outside visuals among others. In all cases there are limitations to the extent of such modifications, often dictated by time, money or technology constraints. An example of such limitations is the maximum field of view that a simulator can provide to the pilots, a function of the visual projection technology and of the cockpit configuration.

A certain set of projectors/screens can be completely adequate in a fixed-wing commercial airliner simulator, where the windows arrangement covers an area that is very wide horizontally but not very extensive vertically. If a different type of aircraft is to be simulated, like a helicopter, the same set of projectors becomes insufficient as the field of view of a rotorcraft has additional windows above and below those of a typical fixed-wing aircraft. Upgrading the visual system, apart from being a costly operation, would not remove all limitations as structural components and instruments would still obstruct the line of sight. This difference between fixed-wing aircraft and helicopters can be quickly noticed when comparing the minimum view prescribed by AS-580 (Ref. 1), the visual envelope recommended by Hale (Ref. 2) for observation helicopters and the SIMONA Research Simulator (SRS) field of view (Ref. 3), as presented in Figure 1.

*MSc. Student, Control and Simulation Section, Faculty of Aerospace Engineering, TU Delft, P.O. Box 5058, 2600 GB Delft, the Netherlands; a.pardi@student.tudelft.nl.

†Researcher Flight Simulator, Control and Simulation Section, Faculty of Aerospace Engineering, TU Delft, P.O. Box 5058, 2600 GB Delft, the Netherlands; o.stroosma@tudelft.nl.

‡Associate Professor, Control and Simulation Section, Faculty of Aerospace Engineering, TU Delft, P.O. Box 5058, 2600 GB Delft, the Netherlands; m.m.vanpaassen@tudelft.nl.

§Professor, Control and Simulation Section, Faculty of Aerospace Engineering, TU Delft, P.O. Box 5058, 2600 GB Delft, the Netherlands; m.mulder@tudelft.nl.

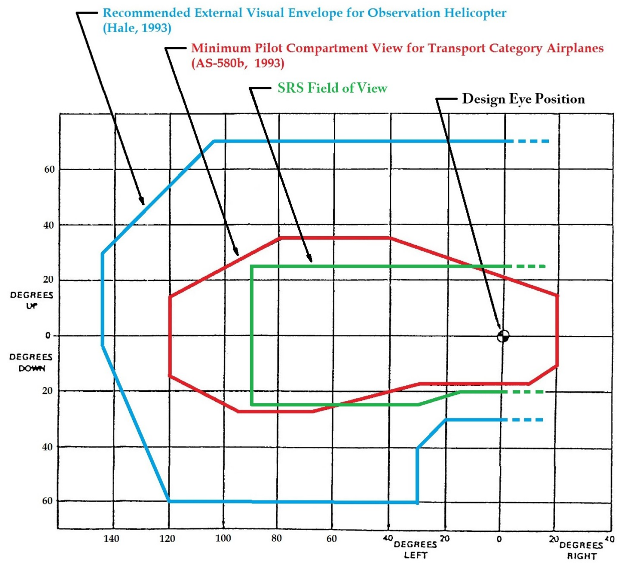


Figure 1. Pilot visual envelopes for transport category airplanes in AS-580b (Ref. 1), observation helicopters recommended by Hale (Ref. 2) and the SRS (Ref. 3), calculated from the Design Eye Reference Point (DERP) of the pilot in the left seat.

The solution evaluated in this paper is to avoid major modifications to the simulator, instead providing the pilot with additional visuals through an augmented reality headset. This would increase objective fidelity in helicopter-based simulations by creating a larger field of view and possibly impact behavioural fidelity by providing pilots with a more realistic experience. Though Augmented Reality (AR) in itself is a mature technology, its introduction in research flight simulators raises questions that need to be answered before its effectiveness can be confirmed. The biggest technical challenge is to achieve full conformality of images across all sources (SRS outside visuals, AR goggles), something that was developed when building the prototype using software tools. Conformality is defined (Ref. 4) as *"The retention of angular relationships at each point on a map projection."* and is interpreted in simulation in a broader way as the degree of coherence that multiple visual information sources achieve.

In the experiment, aspects such as ergonomic compatibility, comfort, practicality and overall usefulness of the AR/SRS prototype were investigated in a human-in-the-loop experiment, where participants piloted a simulated helicopter in a precision maneuver at low speed and altitude. These flight conditions were chosen to create a situation where AR visuals are feasible, meaningful and their use encouraged though not mandated. Low speed and height also mean that focused feature-tracking can be used for spatial awareness and 3D navigation. Current AR technology fully supports this type of vision while other types like peripheral optic-flow based spatial navigation (high velocity) is not possible as it requires a much wider eye-referenced field of view.

Two research questions are formulated, to understand 1) if such precision maneuvers can be executed using AR goggles and 2) whether using AR goggles positively affects the pilots' performance. The first question is addressed by monitoring the correct execution of the maneuvers and analyzing the MIsery SCale (MISC) questionnaire about general health conditions. The second question is answered by analysing the flight controls' activity in terms of cyclic and collective deflections' Root Mean Square (RMS) values, flight parameters (point of zero speed, undershoot, overshoot), motion sickness questionnaires MISC (Refs. 5–7) and Simulator Sickness Questionnaire (SSQ) (Refs. 8–10), Task workLoad indeX (TLX) scores (Ref. 11) and the open comments of the participants.

In the 'Method' section information is provided about the apparatus used in the experiment, the selection criteria for the participants, the piloting tasks and the type of measures recorded. Next in the 'Hypotheses' section the two research questions are presented as null-hypotheses and the eleven measures are listed with indications on how to interpret them. In the 'Results' section all data are reported textually and visually in subsections 'Flight performance', 'MISC - Misery Scale', 'SSQ - Simulator Sickness Questionnaire', 'TLX - Task Workload Index' and 'Open Comments Questionnaires'. Results are then discussed in the 'Discussion' section where statistical analysis is also performed where applicable. In section 'Conclusions' the overall outcome of the experiment is discussed with insight on lessons learned, suggestions for corrections and ideas for future research efforts. Two appendices are present with information on the open comments questionnaire and the reasons behind the choice of the AR goggles.

METHOD

Participants

A total of seventeen participants were selected for the experiment. The requirements for participating were to be aerospace students with knowledge of flight dynamics, have practical experience with computer flight simulation and to be in good health conditions. Efforts were made to simplify the piloting task to a point where it was compatible with the participants and meaningful for the experiment. The fully-coupled 6-DOFs helicopter flight model was changed into a 3-DOF one by preventing lateral movement and rotation around the longitudinal and vertical axes. This way the use of pedals and the lateral movement of the cyclic stick became redundant. Furthermore Control Augmentation System (CAS) loops were added to the flight model to make it easier to control. The effectiveness of such modifications are discussed in the section 'Results'. The experiment was designed as 'within-participants', therefore all students were able to fly in each of the four test conditions following a Latin-square pattern (Ref. 12).

Figure 2. Cockpit of the SRS in typical fixed-wing configuration, with yoke on the left seat and side-stick on the right seat (Ref. 14).

Apparatus

The prototype used was a combination of the SImulation, MOtion and NAvigation (SIMONA) Research Simulator (Ref. 13), an AR workstation and a pair of AR goggles. The SRS is a research flight simulator with the cockpit modeled after a modern passenger aircraft and helicopter-type flight controls. In its original configuration the visual system is collimated, with three projectors covering a total field of view of 180° horizontally by 40° vertically. Cockpit structure and instrument panels reduce the actual field of view, as visible in Figure 2.

In the AR workstation the Unity 3D game engine uses the geometrical model of the Sikorsky S-70 helicopter to calculate where the cockpit windows are from the point of view of the right pilot, and uses them to generate an occlusion mask for the AR visuals. The engine uses full 3D geometry allowing activities such as parallax-based distance evaluation and over-the-edge peeking. Nine windows are available in the S-70: three to the front, two to the sides, two chin and two eyebrows ones. In the experiment the two chin and the left side windows are the only ones made visible through the AR goggles, as the SRS screen already covers front and right windows, while the eyebrow ones are not relevant.

To detect conflicts, misalignments and overlaps with the SRS screen visuals a 3D version of the SRS cockpit is also used in the calculation of the occlusion mask, as shown in Figure 3. The difference in FOVs can be readily appreciated in Figure 4 where the same point of view is presented in either cockpit configuration.

The AR workstation and the AR goggles are used by the pilot according to the test conditions, and extend the field of regard with additional virtual windows. These become visible in the AR goggles when the pilot's head is turned towards the direction where real helicopter windows would be. In all other directions there will be no added AR visuals and, because the AR goggles are *see-through*, flight controls, the cockpit and the ouside visuals on the SRS screen will be visible.

Figure 3. Cockpit 3D models for the (left) and the S-70 (right), with the two models shaded in different colors and separated horizontally for illustration purposes. In the simulation the two models coincide, with windows completely transparent and the cockpits in a black unlit material with no reflective properties (occlusion mask).

The AR goggles used were the "Dream Vision", by DreamGlass Inc., with two Liquid Crystal Display (LCD) screens projecting images on semi-transparent curved mirrors, one for each eye. Conformal images were generated thanks to the CV1 motion capture system by Nolo that provided information about head position and orientation to the AR workstation.

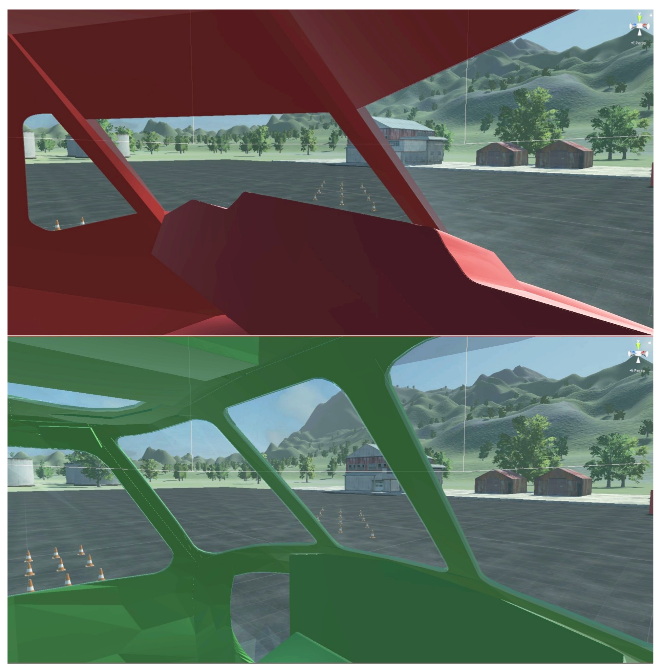


Figure 4. Fields of view comparison between SRS and S-70 cockpits, with the top image showing the outside environment as seen from the pilot's DERP in the SRS and the lower image the one seen in the Sikorsky S-70 under same conditions. Windows and cockpits are made made visible in colour for illustration purposes only.

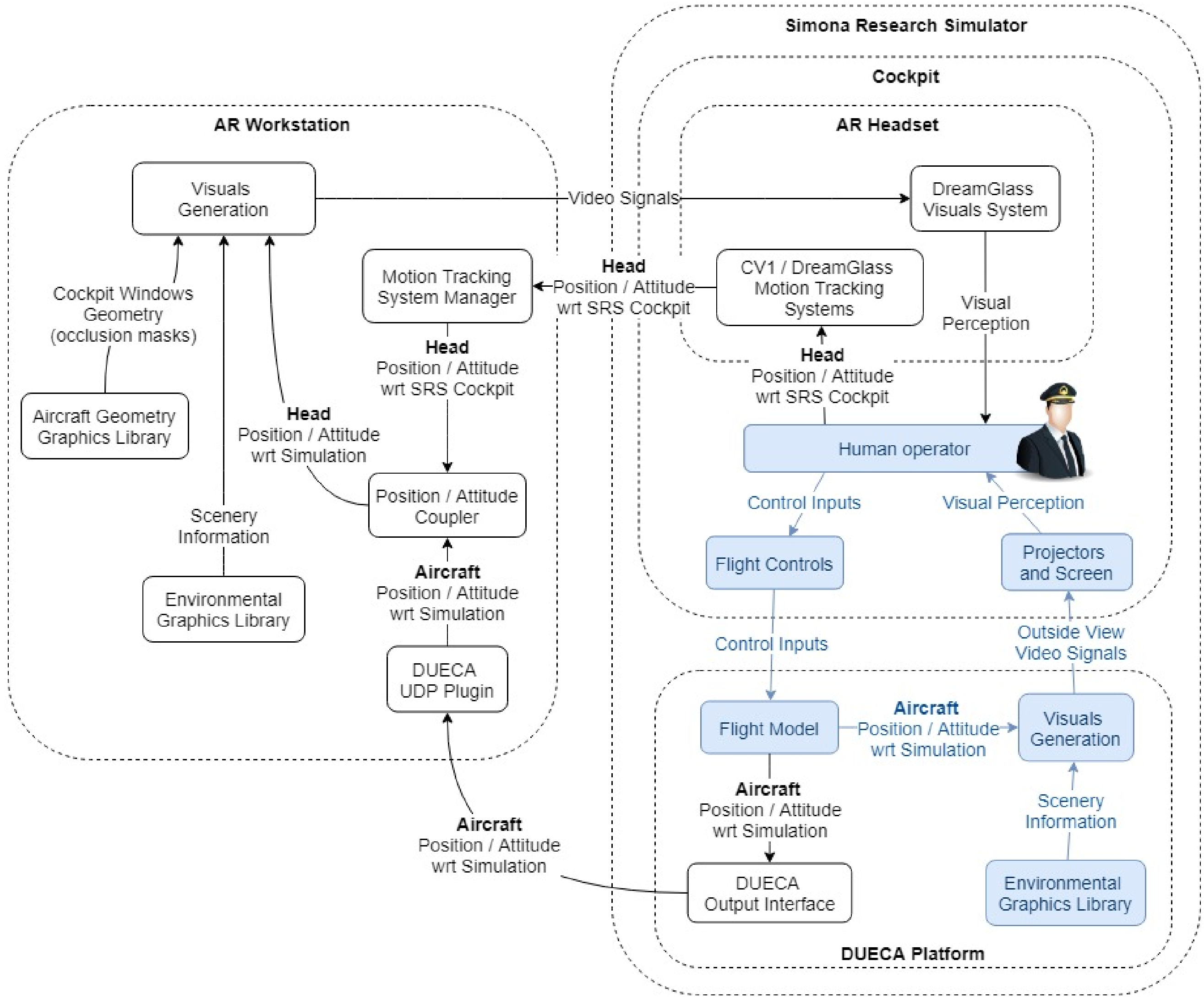


Figure 5. Information flow diagram of the AR/SRS prototype. Light-blue shaded boxes indicate the original configuration of the SRS.

Because of their geometry and the limited numbers of internal reflections, these AR goggles have larger a field of view and better image quality compared to other products such as Epson Moverio (Ref. 15) or the Microsoft HoloLens 1 and 2 (Refs. 16–18). In particular the Dream Vision goggles have a total field of view of 78° horizontally by 48° vertically (Ref. 19). As with other devices, the effective field of regard is unlimited when accounting for the free movement of the head. Though substantial, the field of view of the AR goggles does not cover the peripheral areas of vision, meaning that mostly direct, fixated gaze played a role in the experiment. The piloting task is of precision type at low speed and altitude, promoting direct feature-tracking instead of optic-flow vision.

Two different visual sources are present, the SRS collimated screen and the AR goggles. Conformality of the two sources is achieved in two steps (Ref. 20). Firstly, the Delft University Environment for Communication and Activation (DUECA) platform running the helicopter flight model is used as the source of all flight parameters needed to generate the visuals, such as position and orientation of the helicopter cockpit. Secondly, a head-tracking system is used inside the SRS cockpit, so that the AR workstation can account for position and rotation of the pilot's head in the visuals of the AR goggles. The levels of contrast, saturation and luminosity of the images in the two systems were also regulated to minimize differences in perception. Other geometric characteristics like distortions or 'barrel effects' did not require any adjustment as both systems are based on spherical-type projections without magnification.

The SRS uses three projectors and an optical combiner to fill 180° horizontally and 48.8° vertically (Ref. 13). Each projector, covering a horizontal angle of 73.3°, has a resolution of 1920 by 1200 pixels. This results in a density of 26.2 px/° horizontally and 24.6 px/° vertically in the central areas of projection, or a resolution of 2.29 and 2.44 arcminutes respectively. The DreamGlass, with two LCD screens of 1280 by 800 pixels (Ref. 19), has a density of 16.7° pixels per degree in both axes, or a resolution of 3.59 arcminutes. The standard resolution for the human eye is 1 arcminute in the central region (Ref. 21).

Some visual differences could not be eliminated, but were considered not relevant for the experiment. In particular the single-eye focusing distances are different, with the AR goggles fixed at 2 meters and the SRS projection system at a longer distance. This implies that re-focusing occurs every time the eye switches between the two sources. Another difference is that AR goggles provide stereoscopic vision while the SRS screen is collimated at infinity. The two differences affect the accomodation and convergence aspects of vision respectively, though the type of simulation and the design of the experiment make them less relevant as the objects in the simulation are always several meters away from the observer, many times the eye-to-eye baseline.

The overall structure of the prototype has been summarized in Figure 5, where the information flow diagram highlights the various modules involved. In this diagram, hardware and software entities are grouped functionally in single boxes, and their most relevant data flow relations are expressed with labelled arrows. In this high-level representation the light-blue shaded boxes correspond to the original SRS configuration, with the remaining elements being added for this experiment.

The center of this diagram is the pilot, that operates the flight controls and receives visual information from two sources, as mentioned previously. Inside the DUECA platform the helicopter flight model receives the control inputs and calculates the new position and attitude of the aircraft in the simulated world. This information is used directly to generate the outside visuals for the SRS projectors, as they are fixed to the frame of reference of the cockpit. Through a User Datagram Protocol (UDP) socket interface, this information is also sent to the AR workstation at a frequency of 120 Hz. This frequency is kept constant by the synchronized nodes of the DUECA network, with one packet being sent every 1/120 = 8.3 ms. At the same time, position and attitude of the pilot's head within the physical SRS cockpit is detected by the NOLO CV1 motion tracking system, and sent to the NOLO manager through a Universal Serial Bus (USB) cable connection.

Within the AR workstation, these two spatial information are combined to calculate position and attitude of the pilot's head in the simulation. Once these values are known, the occlusion masks of the S-70 helicopter and that of the SRS cockpit are used to calculate the virtual windows, and generate the outside visuals to be shown in the AR goggles.

In this prototype, the game engine in the AR workstation is used only for visualization purposes, with no physics engine running on it. Because of this, the approach taken was to have the DUECA platform send current spatial information at twice the frequency of the DreamGlass LCD screens (60 Hz) and use the most recent information to calculate the images in the AR goggles. The implementation of a synchronization mechanism was considered but not deemed feasible within the timing of this project.

Piloting Task

The experiment is divided in three sessions for each participant, the first one taking place days in advance to the other two. In the first session a preliminary test is run to make sure that the AR goggles are comfortable to wear and optically compatible with each of the participants. It consists of a passive task where a pre-recorded helicopter flight simulation (with constant pitch, velocity, circular flight path) is presented on the AR goggles alone, outside of the SRS, while the participant moves his/her head around exploring the virtual environment. The duration of the test is purposely extended to about 10 minutes discussing informally about technical aspects and demonstrating the available occlusion masks (S-70, SRS) and their variations.

During and after the test the participant is interviewed about motion sickness and other possible comfort issues. These types of tests were added to the experiment because of earlier episodes, during the development phase of the prototype, when students had problems fitting the AR goggles firmly and comfortably, or had problems focusing their sight. If no problems emerge during such tests the participant is invited to join the second and third sessions of the experiment where the active piloting takes place.

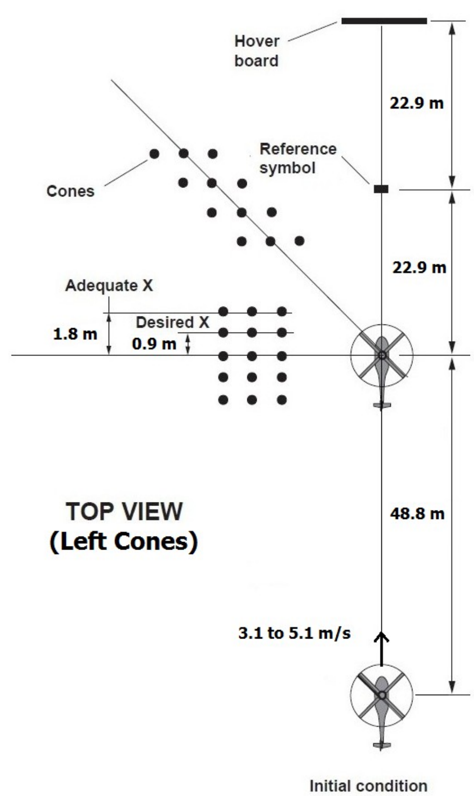


Figure 6. Variation of the Hover MTE found in ADS-33 (Ref. 22), top view of the version with reference traffic cones on the left side, in S.I. units.

Design of the active piloting task was done to make the AR aspect not strictly necessary, so that participants could decide themselves whether to use the additional AR visuals or rely on the SRS screen alone. The task is inspired by the Aeronautical Design Standard 33 (Ref. 22) manual about helicopters' dynamics' evaluation, using a variation of the 'Hover' Mission Task Element (MTE) that is restricted to the longitudinal axis (Figures 6, 8), instead of developing along a diagonal line as originally.

The exploratory nature of the experiment, the emphasis on visual aspects and the naivety of the participants in helicopter piloting skill prompted a reduction of the piloting task difficulty. The fully-coupled 6-DOFs flight model was deemed too difficult to learn for the participants in the available time span and was reduced to a 3-DOFs type with added stability augmentation loops. The allowed DOFs were longitudinal and vertical movements plus rotation around the lateral axis (pitch up/down), removing the need of using the pedals to control the helicopter.

Though simplified, the task remained a challenging one as hover maneuvers are only marginally stable. This is because at very low speeds there are no balancing aerodynamic forces stabilizing the helicopter like those found when moving forward.

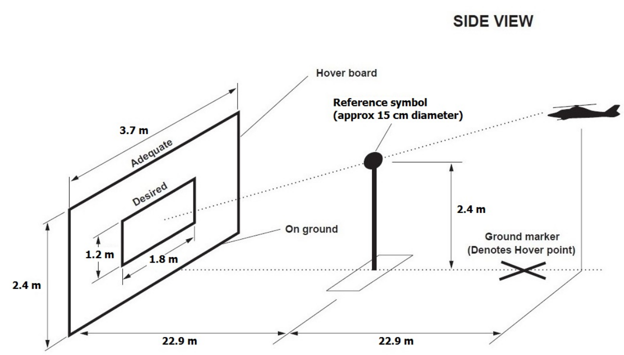


Figure 7. Variation of the Hover MTE found in ADS-33 (Ref. 22), side view of vertical references, in S.I. units.

Another result of this simplification approach was to exclude motion of the simulator's platform. The SRS has a Stewart-type hydraulically-actuated platform supporting the cockpit and the projection system. It was estimated that, considering the participants' starting level, the added motion cueing would have caused more harm than good in the training phase, by causing startle and confusion. Considering the limited time available and the focus of the experiment on visual aspects it was decided to keep the platform still at all times.

The purpose of the experiment was to evaluate feasibility and usefulness of the AR/SRS prototype by measuring certain flight parameters. The interpretation of 'flight performance' is the degree of smoothness and precision of the maneuver and not its aggressiveness and time required for completion. This concept is identical to that of the original 'Hover MTE' in Aeronautical Design Standard 33 (ADS-33) (Ref. 22): and was clearly explained to the participants during the briefing.

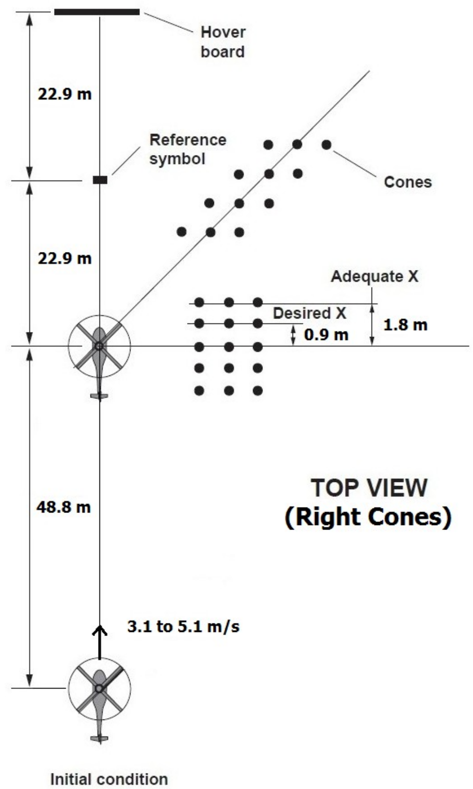


Figure 8. Variation of the Hover MTE found in ADS-33 (Ref. 22), top view of the version with reference traffic cones on the right side, in S.I. units.

This is anticipated in written form and then discussed right before the flights (second and third sessions), with the following being an excerpt: *"Achieving a smooth and precise maneuver is the objective here, and not minimizing the amount of time used to complete it (a maximum limit exists, obviously). As mentioned in the original ADS-33 document: Accomplish the transition in one smooth maneuver. It is not acceptable to accomplish most of the deceleration well before the hover point and then to creep up to the final position"*. It was expected that higher levels of confidence could result in more aggressiveness, however the latter was not stressed as important in the briefing.

Simulation starts with the helicopter at low altitude (10 meters), with small forward velocity (3 m/s) and all the flight controls in trim conditions. The maneuver then requires the pilot to fly forward for about 50 meters until the target area, slow down, stop and hover above the target point for 10 seconds, keeping maximum distances of 2.59 meters horizontally and 1.72 meters vertically from the perfect position set at a height of 3.66 meters, as summarized in Table 1 where also the original ADS-33 values are indicated. Because of the visibility limitations of the cockpit a series of features were added to the scenery, both in front and to the side of the target point.

The scenery shown on both the AR and SRS visual systems is that of a small airport with detailed buildings, surrounded by mountains with a sparse pattern of large, isolated trees and

Figure 9. Scenery setting used in the experiment, a small airport surrounded by mountains with sparse trees and buildings.

a textured sky with distant clouds (Figure 9). The maneuver develops over the parking area in day-light Good Visual Environment (GVE) conditions, with the Sun very high in the sky and all objects casting shadows. These sources of general spatial information are supplemented by others associated with the maneuver. In front of the pilot there is a hoverboard with a reference pole marker to help assessing the vertical position. At an orientation of 45° and 90° to the left and to the right of the target point rows of traffic cones are positioned, as shown in Figures 6, 8, 7, so that the longitudinal position can also be controlled.

The pilot sits in the right seat of the helicopter, so visibility of the outside environment changes according to where the reference points are. To evaluate the different impact of the additional AR visuals, two variations of the MTE tracks are used in the experiment, with the cones being either to the right or to the left of the target point. In Figure 10 the 'Right cones' track course is shown, with the sphere/hoverboard reference clearly visible in front of the pilot and the right-diagonal set of cones partially visible.

Figure 10. Prototype setup with participant wearing the AR goggles. The picture was taken after the end of the experiment and features the cockpit illumination that would normally be maintained off. Also note that information on the instrumentation panel is fixed and not relevant for the experiment.

The AR goggles are worn during all test runs, with the presence (or absence) of added AR visuals creating another type of variation of the piloting task. A total of four test conditions are then obtained mixing the left/right option for the test track with the 'AR on'/'AR off' option for the goggles.

Participants have knowledge of flight dynamics but little to no experience in helicopter controls and piloting technique, therefore a familiarization session is organized before the actual flights begin. This session brings the participants to the level required, introducing one by one the various aspects of the simulation until all four test conditions are practiced. A bi-directional audio channel is used to offer guidance in this phase, while during the actual tests this guidance or correction is not provided.

In the 'within-subjects' design each participant experiences all four test conditions following a Latin-square arrangement. Each condition is simulated until three good test runs are recorded, with a maximum number of seven test runs (regardless of the outcome) to reduce fatigue in the following tests and the onset of motion sickness issues. In order to qualify as 'good maneuver' certain standards need to be fulfilled as presented in Table 1.

Table 1. Adequate performance standards in the Hover MTE of ADS33 (Ref. 22) and those used in the experiment. Values for Cargo/Utility helicopters in GVE conditions, in SI units.

Performance standard	ADS33	Experiment
Attain a stabilized hover within X seconds of initiation of deceleration:	8 sec	n.a.
Maintain a stabilized hover for at least:	30 sec	10 sec
Maintain longitudinal and lateral position within ±X m of a point on the ground:	1.83 m	2.59 m
Maintain altitude within ±X m:	1.22 m	1.72 m
Maintain heading within ±X degrees:	10 deg	n.a.

After each flight the quick MISC questionnaire is filled, while the more lengthy SSQ and TLX ones are only filled after a test condition is completed. When all the conditions are completed a further 'open comments' questionnaire is filled to confirm the validity of the experiment design and visual libraries, assess the usefulness of the added visuals in AR and investigate possible comfort issues.

Independent Variables

Two Independent Variable (IV) were used, each using two values for a total of four test conditions. The first IV is the presence or absence of additional AR visuals in the goggles and the second IV is the location of the reference cones on the ground, either on the left or on the right side.

The variation of the original MTE course in 'Left cones'/'Right cones' is expected to change the difficulty of the piloting task, and motivate the participants to make use of the AR visuals when available. These two levels of difficulty are also expected to act as gain on the results' differences, potentially increasing them in one of the two conditions. This amplification effect also applies to the noise associated with the Dependent Measures (DM), therefore there is no clear expectancy on which difficulty level will provide more significant results.

The vertical reference indicators, namely the reference symbol (a sphere) and the hoverboard, are readily visible in all four test conditions as they lie directly in front of the helicopter Figure 7. However the horizontal reference indicators, the orange traffic cones and the concentric circles marking the hover point, are more or less visible depending on the test conditions.

In the 'Right cones' course variation, the visual contact with the cones can be maintained directly on the SRS screen, and the virtual AR chin windows only offer additional information about the ground markings. This course variation is expected to be easier.

In the 'Left cones' course variation, the horizontal references are all on the far side of the helicopter where outside visibility is more limited. This makes the virtual AR window (the left side one) more important than in the other condition as it increases the field of view in a key area, enabling a more frequent visual contact with the cones.

Dependent Measures

A total of eleven measures are used in this experiment covering the areas of flight performance, pilot workload and motion sickness.

The ideal execution of the maneuver features an initial approach to the target area, followed by a deceleration and a stationary hover above the target point. The absolute distance from the target when zero forward velocity is first achieved, is an indication of the smoothness of the deceleration, with smaller distances from the target indicating better performance. After this point, the pilot tries to fly exactly above the target, however some residual motion of the helicopter is always present, with slight movements back/forth and up/down. In this phase, the displacement from the target is an indication of the precision of the maneuver, therefore the maximum undershoot and overshoot distances are used as measures, with smaller absolute values indicating better performance. The other two DMs about flight performance are the RMS of the control inputs on cyclic and collective, as they depict the amount of effort required to execute the maneuver. In both cases smaller RMS values correspond to lower workload levels.

The raw data about flight performance (Ref. 23) are presented and analyzed here after having been de-trended for inter-participant variability. First the five measures of the flight sessions in each test condition of each participant were averaged, then the mean across conditions was calculated for each measure. A grand mean was also calculated for each measure across all participants in each test condition. Adjustment factors were then obtained by subtracting the means across conditions from the grand means. The original averages of each measure were finally modified summing these adjustment factors.

Motion sickness is evaluated using the MISC (Refs. 5–7) and SSQ questionnaires (Refs. 8–10), the first producing a single value indicating the general health condition of the participant and the second providing three partial scores and a final one. The three partial scores give insight on the type of issue experienced, being it nausea, oculo-motor distress or spatial disorientation. The two questionnaires, though covering the same topic, are not redundant as they have different levels of complexity and analytical power. The MISC is simple and fast, and was filled by the participant after each flight (around every minute) while the SSQ is slower but more powerful, and was used after each test condition was complete (i.e. three to seven flights, around every ten minutes). Pilot workload is measured using the scores of the National Aeronautics and Space Administration (NASA) TLX questionnaire (Ref. 11), with smaller values indicating lower workload levels.

Because of the innovation aspects associated with the AR prototype, an additional open comments questionnaire was used. This is filled at the end of the experiment after all flight tests, with seventeen assertions that each participant can agree (or not) on 5-steps Likert scales (Ref. 24). The complete list of assertions is included in this document, in Appendix A. The questionnaire covers three main areas: confirmation of the experiment design in terms of setup and image generation, potential comfort issues related to the AR goggles and usefulness of the added AR visuals in the current experiment. Written notes can also be added after each question, allowing unexpected issues and new ideas to be documented.

HYPOTHESES

The two research questions mentioned in the introduction are answered by testing the associated null-hypotheses:

1) The selected helicopter-flight maneuvers can not be completed with the AR solution in the SRS.

2) Performance of the selected helicopter-flight maneuvers is not affected when using the AR solution inside the SRS.

In order to disprove the null-hypotheses every DM is expected to show particular trends.

The first null-hypothesis is rejected if the amount of successful maneuvers is adequate and comparable between the AR on/off cases. Furthermore, the health conditions of the participants have to be good at all times, with a maximum MISC rating of "4 - medium dizziness".

The second null-hypothesis is rejected by finding significant differences in the DMs: flight parameters (first point of zero speed, undershoot and overshoot distances), RMS of flight control inputs (cyclic, collective) and the scores calculated in the MISC, SSQ and TLX questionnaires. The various DMs are designed to cover different performance areas and are examined separately during statistical analysis, therefore the presence of multiple DMs linked to the same hypothesis does not increase the possibility of Type-1 errors. Significant differences in either objective flight parameters, motion sickness or workload ratings all result in the rejection of the hypothesis. MISC and SSQ questionnaires do overlap as they deal with motion sickness symptoms, however, they are used together as the SSQ is expected to have more analytical power providing separate ratings for the nausea, oculomotor distress and spatial disorientation areas.

The aircraft modeled visually in the experiment (Sikorsky S-70) has a seat-abreast cockpit configuration, providing each of the two pilots a different field of view of the outside environment. As the pilot is seated on the right-hand side of the helicopter his/her visibility is substantially larger when looking right compared to looking left. By changing the position of the traffic cones on the course track (second IV) the difficulty of the maneuver is expected to change, as the horizontal references in the 'Right cones' option can be more easily seen compared to the 'Left cones' case. The higher difficulty of the latter case is expected to amplify differences in flight performance parameters, RMS of flight control inputs and questionnaire scores. It is expected that with the use of AR visuals the difference between left- and right-cones variations will be reduced. Interpretation of differences, when found, depends on the each particular measure:

- First point of zero forward speed - Measured with respect to the target point in meters, smaller (absolute) values indicate higher precision during the deceleration phase and better performance.
- Undershoot/overshoot distances from target - Measured with respect to the target point in meters, smaller (absolute) values indicate higher precision during the hover phase and better performance.
- RMS of collective and cyclic inputs - Measured as the signal RMS of input-related changes in the angles of the main rotor blades, expressed in degrees. Smaller values indicate lower control activity and are interpreted as higher levels of confidence and smoothness during the maneuver.
- MISC questionnaire - Subjective rating of the general condition as provided by the participants, based on a 11-steps ordinal scale. Higher values indicate less healthy conditions, with a maximum value of '4 - Medium dizziness' allowed for the experiment.
- SSQ questionnaire - Subjective rating of sixteen motion sickness symptoms as provided by the participants, based on 4-steps ordinal scales. Three partial scores are calculated summing groups of seven symptoms' ratings, partially overlapping, in order to give insight into the areas of nausea, oculomotor stress and spatial disorientation. A final score is calculated based on these three values. Higher values indicate less healthy conditions.
- TLX questionnaire - Subjective ratings of seven workload areas as provided by the participants, based on 21-steps ordinal scales. The total score is calculated using the ratings and a series of pair-wise comparisons of the workload areas. Higher values indicate higher levels of workload and stress.

RESULTS

The call for participants in the experiment was initially answered by 19 students, above the threshold of 16 required. The preliminary checks about compatibility with the AR goggles (test session n.1) saw the participation of 17 of them, and all of them were accepted to the next phase. Structure and complexity of the piloting task was refined and corrected with the help of 2 participants, thus excluding them from the experiment. During the second test session (training phase), some participants had difficulties in controlling the helicopter flight model, resulting in a very quick onset of motion sickness and preventing them from continuing. A total of N = 11 participants completed the training phase (test session n.2) and were admitted to the third test session.

Accounting for 3 good maneuvers per participant per test condition, a minimum of 3 x 11 x 4 = 132 flight sessions were required as shown in Table 2. At the end of the experiment a total of 203 flight sessions were executed, as in many cases the maneuvers had to be repeated. Repetitions were decided on-the-fly by monitoring flight data on the screens of the control room, resulting in a total of 127 'good maneuvers'. This latter figure is lower than the required 132 value as in some cases the maximum number of repetitions was reached before 3 'good maneuvers' were achieved.

A script in MATLAB was used in post-processing to check each maneuver, calculating its duration and detecting when the helicopter first enters and first leaves the target area. As motion is only allowed longitudinally and vertically, a simultaneous check of x-position and height was enough to calculate this. At this point it was revealed that the method used initially to judge the goodness of each maneuver was rather optimistic. In particular, 32 flights that were considered good turned out not to fit the criteria of the experiment, lowering the number of 'good maneuvers' from 127 to 95. This reduction was not distributed uniformly across the test conditions, and the initial figures of 31, 32, 33 and 31 good test flights were altered to 17, 25, 26 and 27 flights, respectively, as shown in Table 2. This error of initial judgement had two effects: on the one hand it made the comparison of the numbers of 'good flights' not meaningful and on the other hand it reduced the analytical power of the following statistical analysis done on the DMs by reducing the available data points.

The aspects of simulator sickness and task workload were also examined, and in these cases it was decided to use data from all flights regardless of the outcome, meaning that MISC, SSQ and TLX questionnaires are representative of all 203 flight sessions.

From a statistical standpoint each continuous DM was tested for normality distribution assumptions using the K-

Table 2. Statistics of the amount of flight sessions per test condition.

Test condition #	1	2	3	4	Totals
IV - Augmented Reality	Without AR	Without AR	With AR	With AR	
IV - MTE course track orientation	Left cones	Right cones	Left cones	Right cones	
Maximum available flight sessions	77	77	77	77	308
Flight sessions actually used	52	45	56	50	203
Good maneuvers required	33	33	33	33	132
Maneuvers initially accepted as good	31	32	33	31	127
Maneuvers confirmed good in MATLAB	17	25	26	27	95

S Kolmogorov-Smirnov (Ref. 25) and S-W Shapiro-Wilk (Ref. 26) methods with p = 0.05. The overlap of distributions' tails was inspected visually using box-and-whisker plots. Variance was also tested for homogeneity across conditions. If a DM passed all tests, thus meeting assumptions for parametric testing, the Factorial Repeated Measures ANOVA method was applied to it (Ref. 27) (one outcome, continuous DMs, two categorical IVs, same participants, parametric assumptions compliance). However, this was never the case in the experiment, as one or more assumptions were violated by each DM, prompting a different approach. A series of non-parametric pairwise Friedman's tests were run on the data samples, comparing two conditions at a time for a total of (4-1)! = 6 tests per DM.

The following DMs were considered continuous: the flight performance DMs, the final SSQ scores and the TLX ratings. These other DMs were considered non-continuous due to their limited range of discrete values: MISC and partial SSQ scores. Non-continuous DMs were analyzed directly with pairwise Friedman's tests.

The Likert scores in the Open Comments questionnaire contain information about all four test conditions in an aggregated form. They are presented in textual and graphical forms without further statistical test except for a measure of consensus, the C_t parameter (Ref. 28). This was decided as these scores encompass the whole experiment, cannot be compared based on test conditions and have no expected values.

Flight Performance

Data about the point of first zero forward velocity are presented in Figure 11. Median values closer to the target point are recorded when the AR visuals are available (-1.282 and -1.549 meters) compared to the case without AR visuals (-2.758 and -1.746 meters), suggesting the use of AR increased flight performance. This DM passed the K-S and W-S normality tests but not the visual inspection test due to data overlap. Pairwise Friedman's tests were inconclusive in finding significant differences. No clear trend is visible when comparing data between left and right traffic cones.

After stopping in mid-air for the first time, the helicopter remains close to the target point, recording a minimum and maximum distance from the start point. In Figure 12, data about the minimum distance, or undershoot from the target, are presented. Here median values do not indicate a trend for any of the IVs except for the first column (without AR, left

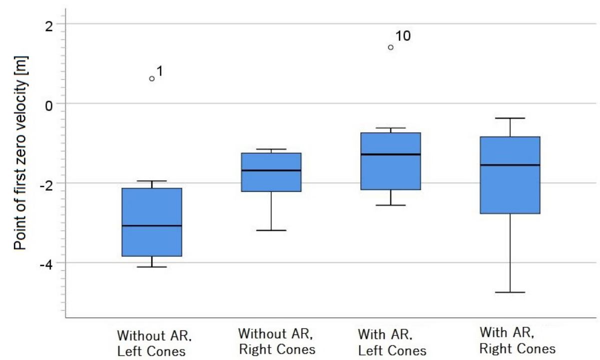


Figure 11. Point of zero speed DM, by test condition

cones) where the lower value of -3.415 meters indicate worse performance than the other three cases (-2.423, -2.795, -2.740 meters). This DM passed the K-S normality test but not the W-S one, the homogeneity of variance test and the visual inspection due to data overlap. Pairwise Friedman's tests were inconclusive in finding significant differences.

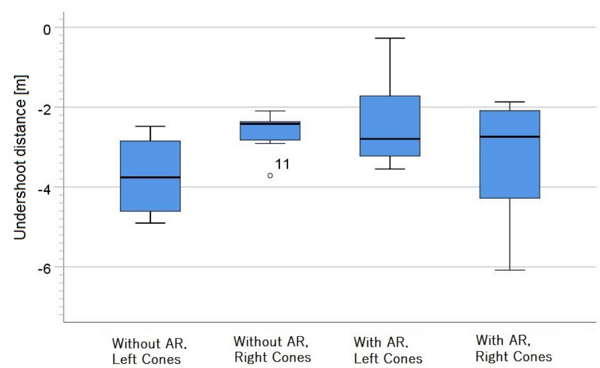


Figure 12. Undershoot distance DM, by test condition

The maximum distance after stopping in mid-air, also called overshoot distance, is presented in Figure 13. In all cases the overshoot median values are below one meter. All test conditions show very close median values (0.990, 0.963, 0.928 meters) except for 'AR visuals, left cones' where the median is 0.466 meters. This suggests that better performance is obtained when AR is used on the left-cones track, however, this test case also shows a higher variance than the other ones, indicating lower performance consistency. This DM passed the K-S and W-S normality tests, the homogeneity of variance test but not the visual inspection, due to data overlap. Pairwise Friedman's tests were inconclusive in finding significant differences.

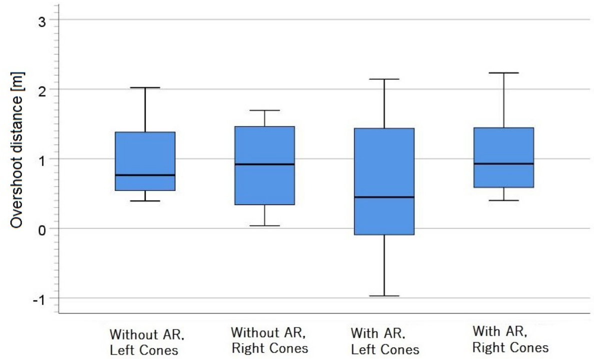


Figure 13. Overshoot distance DM, by test condition

The control activity of the pilot is evaluated with the RMS of cyclic and collective inputs. Data about blades' deflections due to cyclic input are presented in Figure 14, in degrees. No clear trend is noticeable but data suggest an interaction effect: the addition of AR increased the RMS for the left cones scenario, while it decreased for the right cones scenario. These two cases show RMS of 0.790 and 0.802 degrees respectively, while the other two have a RMS of 0.777 degrees (without AR, left cones) and 0.785 (with AR, right cones). Higher cyclic control activity is associated with overcompensation, and is considered an indication of lower confidence when executing the piloting task. This DM passed the K-S normality test but not the W-S one and the visual inspection, due to data overlap. The pairwise Friedman's tests found one significant difference, between the 'Without AR, Left Cones' condition and the 'Without AR, Right Cones' (exact Sig. = 0.012). Friedman's pairwise tests were inconclusive in finding significant differences.

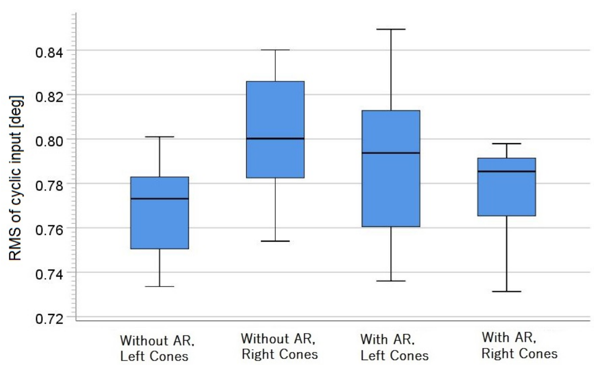


Figure 14. RMS of cyclic input DM, by test condition

Deflections of the collective are shown in Figure 15, where the RMS of the main rotor blades' deflections due to collective inputs is shown in degrees. The median values are closely grouped (1.479, 1.470, 1.483, 1.481 degrees) and no clear trends are noticeable. As for the cyclic input also for the collective a higher control activity would denote overcompensation, and would indicate lesser confidence when executing the maneuver. This DM did not pass the K-S and W-S normality tests, the homogeneity of variance test and the visual inspection, due to data overlap. Pairwise Friedman's tests did not find significant differences.

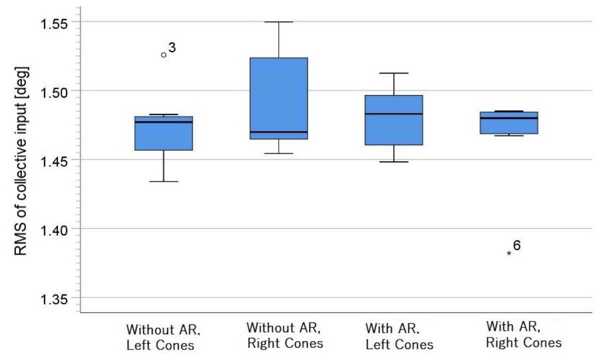


Figure 15. RMS of collective input DM, by test condition

MISC - Misery Scale

After each test flight, the Misery Scale questionnaire was filled, regardless of the success of the maneuver. This results in an over-representation of pilots with higher failure rates. However, no compensation was made to account for this, as the focus of the MISC is to capture the health status of the participants and not the performance levels. Because of this choice, between 45 and 55 scores are available per test condition, for a total of 203 data points (see Table 2). The MISC scores have discrete values ranging from 0 to 10, and are presented in Figure 16 by counting their occurrences and showing them as percentages within each test condition.

Data reveal that the zero-score ('No problem') was chosen in more than 75% of the cases, and that the amount of non-zero scores is higher when AR visuals are not available. It must be noted that all the scores are well within the maximum values allowed in the experiment ('5 - severe dizziness') and that all scores '2 - vague dizziness' were selected by a single participant.

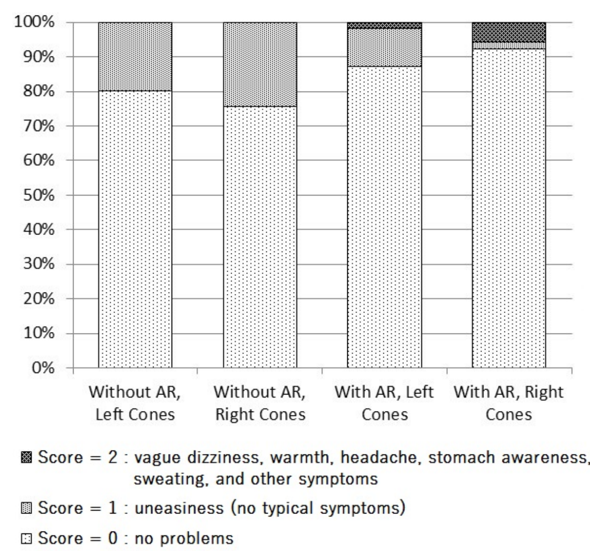


Figure 16. MISC score DM, by test condition

SSQ - Simulator Sickness Questionnaires

The Simulator Sickness Questionnaire results are presented in Figures 17, 18, 19 and 20, where the three partial scores and the total one are shown. Similarly to the MISC results the three partial SSQ scores produce discrete values and are thus presented by counting their occurrences and showing them as percentages within each test condition. The total score, being calculated by applying weights to the three partial scores, can be treated as a continuous variable and thus examined with boxplots.

The SSQ partial score about nausea is shown in Figure 17, with the range of possible values going in 22 steps from zero to 200.34, this corresponding to all seven nausea symptoms at level 'severe'. Symptoms featured in this partial score include increased salivation and stomach awareness. As no score above 28.62 is present, only the first four steps of the scale are depicted, with no noticeable difference between test conditions. It must be noted how in roughly half of the cases some nausea-related symtom was experienced by the participants.

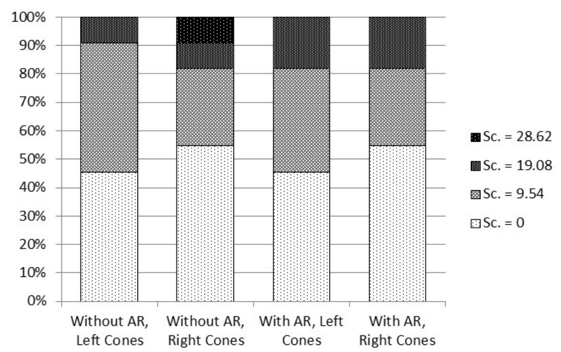


Figure 17. SSQ Nausea score DM, by test condition

The second SSQ partial score is about oculomotor distress and investigates symptoms such as eyestrain and blurred vision. In this case the possible values range from zero to 159.18 in 22 steps. In the experiment only the first seven levels of the scale are present, therefore the data are presented in Figure 18 features seven categories. A clear trend can be observed as the test condition 'With AR, Left cones' shows higher scores than the other ones.

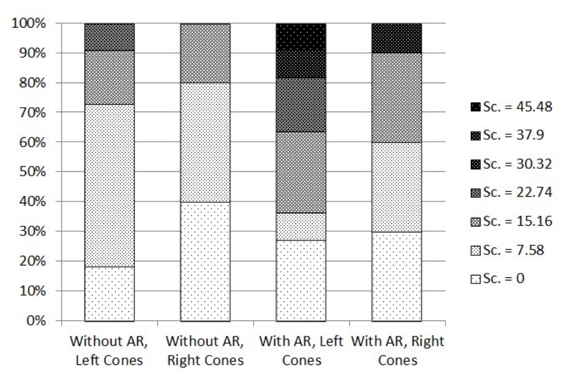


Figure 18. SSQ Oculomotor score DM, by test condition

The third SSQ partial score investigates the spatial disorientation experienced by the pilot, looking at symptoms such as vertigo and fullness of head. The scores, ranging from zero to 292.32 in 22 steps, are presented in Figure 19. By looking at the data it can be noted how the scores are higher when the AR visuals are available, and also how absolute values remain confined to the first four steps of the scale.

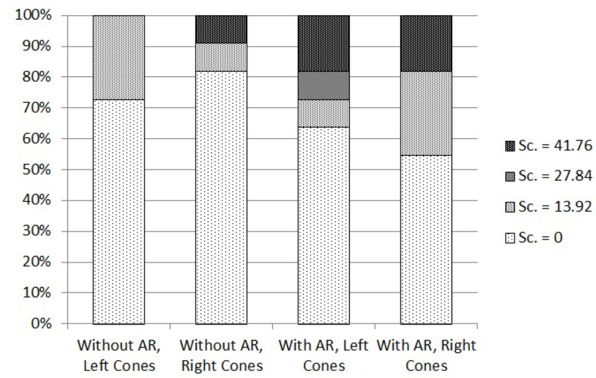


Figure 19. SSQ Disorientation score DM, by test condition

The total SSQ score is calculated using a linear combination of the three partial scores, producing values from zero to 120.54. Because of its granularity a box-and-whiskers representation is chosen to present them in Figure 20, with the maximum recorded value being 20.22. Most apparent is the larger variance of the condition 'With AR, Left cones', this being caused by higher oculomotor distress scores. Further analysis of the SSQ scores is not advised as the scores are too close to the lower end of the scale to be meaningful (Ref. 8).

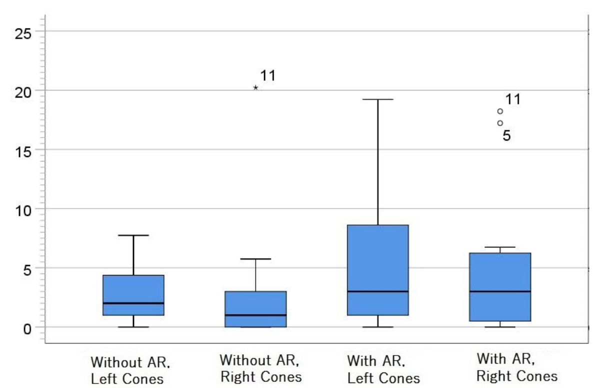


Figure 20. SSQ total score DM, by test condition

TLX - Task Workload Index

The workload experienced by the pilot is measured by the TLX scores, based on subjective scale ratings and pairwise evaluations between the scales themselves. In Figure 21 the data are presented in box-and-whisker format. A trend for lower workload scores can be observed for the test condition 'With AR visuals, Right cones', with a median of 47.00 when the others have 55.00, 55.67 and 56.33, respectively.

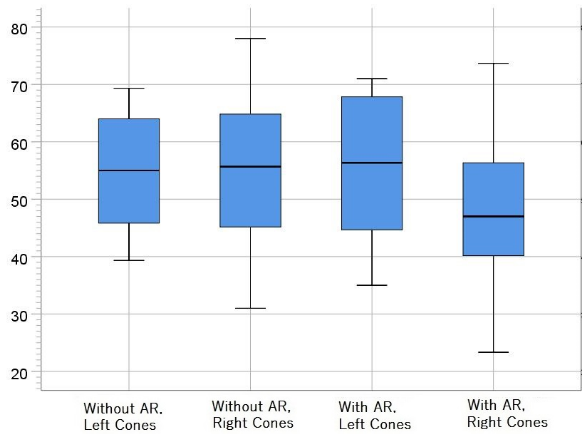


Figure 21. TLX score DM, by test condition

Open comments questionnaires

The Likert ratings provided in the open comments questionnaire, reported in detail in (Ref. 23), are summarized here. In most cases more than half of the participants shared the same general opinion, being it positive ('Agree', 'Strongly Agree') or negative ('Disagree', 'Strongly Disagree'). In other questions there is no general consensus and the five possible ratings were widely distributed, making the analysis of written comments more useful to understand the opinions of the participants. The complete list of the assertions is included in the Appendix. The measure of consensus is calculated using the C_t parameter (Ref. 28) after having simplified data by grouping the two positive and the two negative categories between them, resulting in a 3-scale rating system made of 'Agree', 'Neutral' and 'Disagree'. The C_t parameter ranges from 0 (full dissent, uniform distribution of ratings between the three categories) to 1 (full consensus in one category).

All four questions about the experiment design and the visual libraries were answered with consensus. The 3D simulation environment was considered to have enough objects and references to determine longitudinal position (Q1, C_t = 1.00), vertical position (Q15, C_t = 1.00) and angle of pitch (Q11, C_t = 0.60), while the absence of flight instruments was not considered a big inconvenience for this experiment (Q9, C_t = 0.55).

A group of ten questions focused on the possible comfort issues of the AR headset and its integration with the SRS. In four of them the participants expressed consensus, about the quality of images (Q5, C_t = 0.73) and the ability to focus on virtual objects (Q12, C_t = 0.60).

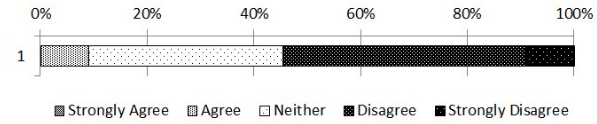


Figure 22. Likert ratings' distribution for Question 7 of the open comments questionnaire: *"The AR headset was very comfortable to wear".*

From an ergonomic standpoint the temperature (Q13, C_t = 0.87) of the AR headset was considered not excessive or too high, and the cables connecting the AR headset were found not to be obstructing head movements (Q14, C_t = 0.60).

In the other five cases the consensus parameter was below 0.50: latency/lag of images in the AR headset (Q2, C_t = 0.40), presence of noticeable jumps in the AR images (Q4, C_t = 0.46) and excessive weight of the AR goggles (Q10, C_t = 0.46).

When asked about comfort in a more general way the AR headset was not considered very comfortable to wear (Q7, C_t = 0.40, Figure 22), with the following comments:

"At first it is comfortable, but after sustained wear it becomes uncomfortable", *"Slightly pushing on forehead to make sure it stay where it should be"*, *"Strain on forehead made it a little uncomfortable"*, *"OK by itself, but right ear started to hurt in combination with headset"*, *"I had to strap the AR very tightly to focus the images"*, *"It felt tight and heavy around the back of my head. Otherwise it was comfortable. Maybe could have a stop over the head for support"*, *"It pushed into my [own] glasses, adding a bit of pressure, It did not get uncomfortable, however."*, *"Was comfortable but for long periods of time would be noticeable"*, *"Difficult to get first correct position. Tends to be somewhat more heavy after some time"*, *"I could not have kept it on 30 minutes extra ->headache"*.

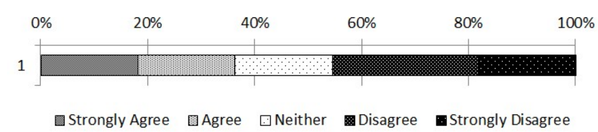


Figure 23. Likert ratings' distribution for Question 6 of the open comments questionnaire: *"The overlap of virtual and screen visuals in the left side window was very noticeable".*

Opinions were again divided in Questions 17 and 6: participants did not show consensus about the tiredness of the eyes when refocusing between the AR visuals and the projectors' screens (Q17, C_t = 0.48), and the same happened about the overlap of AR and SRS visuals on the left side (Q6, C_t = 0.24, Figure 23). In the latter case the following comments were added:

"They were slightly off (not fully calibrated) but good enough", *"Mostly ignored the screen visual and only looked at the virtual window"*, *"I barely looked so much to the left"*, *"The left windows were not aligned, and sometimes the AR left window was even rotated by plus/minus 30 degrees"*,

"I noticed it but it did not affect my flying", *"The left side cones became unusable because of it"*, *"Occasionally noticeable but not to an effort that ruined it"*, *"Due to aligning not exactly there is a little shift. Moreover the AR is [a] bit more blurry / pixelated"*, *"Did not notice. Brain filters it out. The trim displays [in the instrument panel] were also reflected on the outside visuals"*,

Three questions dealt with the overall usefulness of the added visuals in AR, first asking if the AR visuals were needed to execute the piloting task in all its variations and then addressing each of the two virtual windows. In the first question (Q8, C_t = 0.87) most participants agreed that AR visuals are not strictly needed in the experiment, confirming the design choice of leaving the participants free to decide whether to use the AR visuals or not. However, when asked about the two AR virtual windows, opinions were more divided.

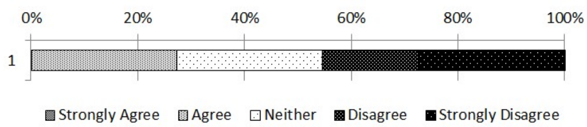


Figure 24. Likert ratings' distribution for Question 3 of the open comments questionnaire: *"Observing the right chin window in AR helped me maintaining the hover position".*

The assertion about the usefulness of the AR right chin window (Q3, C_t = 0.18, Figure 24) recorded the most diverse distribution of answers in the experiment, with the following comments added:

"The wires of the AR headset made it very cumbersome to move the head", *"Only looking at board and cones"*, *"Not significantly used"*, *"Main problem: small FOV - up and down makes it impossible to quickly glance in corner"*, *"I never used the chin window because it would take too long to move my head back to the pole reference"*, *"It did [help] maybe once or twice but mostly did not use it"*, *"Only with right cones, only in peripheral vision"*, *"Did not use this window much"*, *"Seeing circle [ground marker] helps in knowing when you need to really stop"*, *"Whenever I tried to use it the position got unstable (for my feeling)"*.

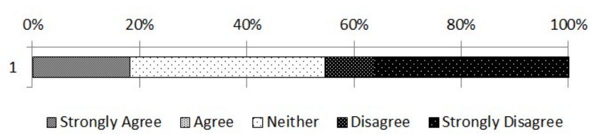


Figure 25. Likert ratings' distribution for Question 16 of the open comments questionnaire: *"Watching the left side window in AR helped me maintaining the hover position".*

The statement about the usefulness of the AR left side window (Q16, C_t = 0.24) was rated as shown in Figure 25, with the following comments added:

"Only when flying past the cones it was really useful otherwise it could be done without but made it easier", *"Although I tried to keep looking ahead a quick look to the cones on the left really helped determining position"*, *"I did not use it at all"*, *"I never used the left side AR window"*, *"Did not use it"*, *"The latency on AR made me see everything twice, making it almost impossible [to use]"*, *"Did not help at all"*, *"I mostly used diagonal cones as less head tilt was required."*, *"Too blurred and too delayed to be a real help. I only looked out of curiosity, but that distracted."*.

DISCUSSION

Analysis of results associated with the first research question uses data from Table 2 regarding initiated and completed maneuvers, with good or bad outcomes. A "good maneuver" is one that complies with the requirements indicated in Table 1, while any violation makes it a "bad" one. During the experiment the test coordinator instructed participants to repeat maneuvers when a clear violation of the requirements was noticed. As shown in Table 2 when AR was not used a total of 97 flights were needed, and when AR was used a total of 106 sessions were done. In post-processing it was possible to re-assess these figures with computer scripts, finding out that only 53 were actually "good maneuvers" when using AR, and 42 when not using AR.

Because of the test repetition criteria it is difficult to compare such figures using statistical tests, and they are only useful in order to reject the first null-hypothesis, being them non-zero and of the same order of magnitude. The absence of significant differences in the MISC ratings and their low values also support this conclusion. The associated research question (*Can helicopter pilots execute such particular maneuvers with AR goggles?*), can thus be answered positively.

The remainder of the data is used to answer the second research question, about performance differences in various areas.

The 'point of first zero speed' data show violations of assumptions about sample normality and homogeneity of variance. Some effects can be still noted. In particular in the 'Right cones' variation the presence or absence of AR visuals seems not to change performance, while in the 'Left cones' condition the absence of AR results in worse performance, with the stop point being more far away from the target.

Also the 'undershoot distance' data violate distribution assumptions, and the only effect is again a notable worse performance in the 'Left cones' variation when AR visuals are not in use. In the case of the 'overshoot distance' normality is satisfactory but the samples are largely overlapping, showing no appreciable trends apart from a larger variance in the 'Left cones' variation with AR visuals.

The cyclic control input RMS data have roughly normal distributions, but they too largely overlap. A cross effect is notable with conditions 'Without AR, Right cones' and 'With AR, Left cones' resulting in the larger control activities, an indication of less piloting confidence. Data regarding the collective input is heavily skewed and overlapping with no clear trends visible.

The only statistically significant effect was found in the cyclic RMS measure using Friedman's test between two particular conditions, both not using the AR visuals. Positioning the traffic cones on the left side was found to cause a significant increase in cyclic control activity, when compared to having the cones on the right side (Sig. = 0.012).

Questionnaires about simulator sickness show in general very low results with a single exception. In the case of the MISC ratings the most chosen answer by far is 'no problems' with only few other cases of 'uneasiness' and 'vague dizziness, [...]' answers. No clear trend is noticeable across test conditions.

The SSQ method does not indicate problems in the areas of nausea and disorientation, with low scores in general and no visible trends. On the other hand in the oculomotor distress area the test conditions 'With AR', in particular that with the 'Left cones' variation, have higher scores than when AR is not available. In the calculated total scores skewedness and tails overlap prevent further analysis. This latter issue is also present in the TLX results, where only a trend for lower workload ratings is visible in the 'With AR, Right cones' test condition.

In general the results connected to the second research question ("*Flight performance of the selected helicopter-flight maneuvers is not affected when using the AR solution inside the SRS*") are noisy and inconclusive, preventing the formulation of a clear answer. Some trends can be observed, like the cross-effect in the cyclic input RMS or the SSQ oculomotor distress, however they never translate into statistically significant effects.

CONCLUSION

The experiment showed that it is possible to introduce AR in an already existing research flight simulator, but the benefits of such choice were not conclusively demonstrated and require further research efforts. While the technical challenges found during the development phase were overcome with few exceptions, other aspects and design choices can be revised and improved.

The adoption of helicopter simulation as context for the experiment was made to exploit a much increased fields of regard with the use of AR. This choice entailed the use of helicopter flight models that were difficult to learn for inexperienced participants, resulting in large performance variances, high error rates and test repetitions. On the other hand the choice of young aerospace students as population sample was made to address the ergonomic uncertainties associated with AR technology, and to avoid dropping out actual helicopter pilots on the day of the experiment due to previously undetected physical issues.

To avoid these problems future research efforts should rely on piloting tasks that are fully compatible with the population sample, so that the observed measures will not suffer from the same noise levels.

The second most important issue to be addressed in the future is the latency of the system that detects position and rotation of the participant's head. In the used apparatus the Nolo CV1 was not fast enough, so the Unity 3D game engine could not generate time-conformal images, and the perceived delay discouraged the use of AR visuals by the participants. By choosing a different system or designing a new one the latency aspect could be corrected. With this issue resolved, a synchronization mechanism could also be implemented between the AR and SRS visual sources. For example, this could use high-speed camera sensors, pulse images and control step inputs to model the total delays in the graphical pipelines and account for them, thus achieving better time conformality.

The third issue is the optionality of AR. The idea of having redundant visual sources of spatial information and to let the pilots decide which one to use turned out to be a confound, as the effective use of AR visuals could not be objectively determined and assessed. It is therefore recommended that future studies make sure if and when the participant is using the AR visuals, for example by displaying key information on the AR goggles alone.

APPENDIX - OPEN COMMENTS

In the open comments questionnaire a series of 17 assertions was presented to the participants. Each statement was answered by choosing a value on a 5-scale Likert scale from 'strongly disagree' to 'strongly agree' and by optionally adding a hand-written comment. Assertions deal with three main topics and were presented in a randomized order, as indicated by the numbers next to the 'Q' letter in the following list.

Experiment design confirmation
Q1: There were enough objects and references in the environment to determine my longitudinal position.
Q11: There were enough objects and references in the environment to determine my angle of pitch.
Q15: The environment had enough objects and references for me to determine my vertical position.
Q9: Not having flight instruments (e.g. the attitude indicator) was a big inconvenience in this experiment.

Usefulness of AR visuals
Q3: Observing the right chin window in AR helped me maintaining the hover position.
Q16: Watching the left side window in AR helped me maintaining the hover position.
Q8: The piloting task can be completed without the added AR visuals.

Comfort issues of the AR goggles
Q2: There was an excessive lag (latency) in the images seen in the AR headset.
Q4: When moving my head the images in AR changed smoothly, with no noticeable jumps.
Q5: The image quality in the AR headset (contrast, luminosity) was very similar to that of the simulator screens.

Figure 26. Front-left view of the DreamGlass AR goggles (Ref. 19).

Q6: The overlap of virtual and screen visuals in the left side window was very noticeable.
Q7: The AR headset was very comfortable to wear.
Q10: The weight of the AR headset was excessive.
Q12: I could always focus precisely on objects inside the AR visuals.
Q13: The AR headset temperature was often too high.
Q14: The connecting cables of the AR headset allowed me to turn my head around with ease.
Q17: Refocusing between the AR visuals and the simulator screens was tiring for the eyes.

APPENDIX - CHOICE OF AR GOGGLES

Selection of the AR goggles was made following precise technical criteria and at the same time accounting for the limited space available inside the SRS cockpit. Mandatory characteristics are summarized in the following list with all items starting with “The AR device must ...”:

- be of the Head-Mounted Display, see-through type
- have six Degree Of Freedom (DOF) motion tracking sensors
- have an unobstructed look-down field of view
- be ergonomically compatible with the SRS cockpit
- be visually compatible with the SRS projection system
- have motion tracking sensors technically compatible with the SRS
- be available on the market and affordable (Refs. 29, 30)

Once a list of compatible devices was put together three other criteria were considered on a best-performance basis, namely having the largest field of view, the best visual quality and the most friendly Interactive Development Environment (IDE) software.

Most of the points in the above list are self-explanatory, e.g. compatibility, however some need a more detailed description. The requirement about unobstructed look-down field of view is needed to make sure pilots can continuously see instrumentation, flight controls and their own hands so that piloting realism is not affected and no confound is created in the experiment. The presence of 6-DOFs tracking is needed so that the visuals in the AR headset can be adjusted according

Figure 27. Top view of the DreamGlass AR goggles (Ref. 19). Retention on the head is achieved by two padded strips pressing on the back of the head and on the front, with pressure regulated by the wheel inside the back support.

to the head position in the SRS cockpit, to achieve conformality and allow visual effects such as parallax to be experienced seamlessly.

The three performance-based criteria were considered for all devices satisfying the mandatory requirements, and a trade-off matrix was not needed as one product was superior in all of them. The DreamVision “Dream Glass” AR headset, supplemented by the Nolo CV1 motion tracking system, provided the largest field of view, best visual quality and had very friendly IDE (Ref. 19). This device uses a so-called ’bug-eye’ configuration with two separate LCD screens projecting light onto an optical combiner in the form of two curved semi-transparent mirrors, as seen in Figures 26, 27 and 28. The more famous Microsoft HoloLens 1 was considered first, but then excluded due to an inferior field of view when compared to the Dream Glass.

It must be said that mandatory criteria about SRS compatibility could not be verified without first building the prototype, therefore an estimation had to be made based only on previous experience and reported data. Ergonomic compatibility in the SRS is achieved if the AR headset does not interfere physically with parts of the cockpit, in particular with the overhead panel, and this was assessed examining pictures of the AR devices.

Visual compatibility with the SRS projection system is about the two different light sources in the prototype and their interaction. Even though the AR systems use the principle of additive light, the possibility of flickering and colour distortion could not be excluded. To deal with the issue and mitigate risks it was decided to exclude all AR headsets with visual update frequencies below 60 Hz.

The issue of sensors’ interference was also considered, as there can be incompatibilities originating from space limitations and/or unwanted interactions of Infra-Red (IR), light and radio signals. The motion tracking systems of AR products are typically designed for use in a room with free space around, and their technical sheets often cite a minimum dis-

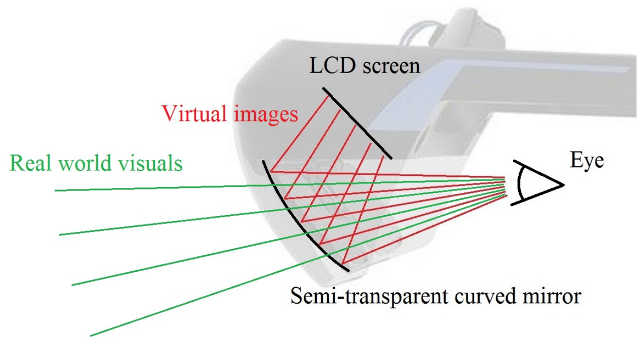


Figure 28. Projection technology of the DreamVision Dream Glass (Ref. 19). Images are generated on the LCD screens and reflected by a single lens before entering the eye. The lens is a semi-transparent curved mirror that acts as single-step image combiner for virtual and real-world images.

tance between sensors and trackers, creating uncertainties on their behaviour if that distance is reduced. In the SRS cockpit the amount of space around the head of the pilot is extremely limited, in some cases less than 20 cm, and this could have caused the motion tracking to produce erratic data.

During prototype development all the technical issues initially hypothesized were examined in dedicated test sessions and dismissed, with a single exception. The Nolo CV1 mocap system is used to calculate the head position and generate the images in the Dream Glass AR goggles, and the latency times of such system were measured to be around 120 ms.

The 120ms delay affects the update of the point of view of the AR images, resulting in a perceived latency when the head is turned. As long as the direction of the head is kept constant this effect is not present. Time figures were measured using 50 frames-per-second (FPS) video recordings of the AR goggles and the output screen: in particular the frame count difference was measured between sudden jerks applied to the headset and the resulting change in images on the screen. Results were averaged over several video recordings, with resolution being 20 ms due to the 50 FPS frame-rate (1/50 = 0.02 s).

REFERENCES

1. Anon., *AS580B - Pilot Versatility From the Flight Deck Design Objectives for Commercial Transport Aircraft*, SAE International, Troy, Michigan, USA, November 1978.

2. Hale, S., *Helicopter External Vision Requirements and Visual Display Characteristics*, Aberdeen, Maryland, USA, October 1987.

3. Ellerbroek, J., Stroosma, O., Mulder, M., and Van Paassen, M. M., "Role Identification of Yaw and Sway Motion in Helicopter Yaw Control Tasks," *Journal of Aircraft - J AIRCRAFT*, Vol. 45, 07 2008, pp. 1275–1289. DOI: 10.2514/1.34513

4. Anon., *Dictionary of Scientific & Technical Terms*, The McGraw-Hill Companies, Inc., 2003.

5. Bos, J., Mackinnon, S., and Patterson, A., "Motion Sickness Symptoms in a Ship Motion Simulator: Effects of Inside, Outside and No View," *Aviation, space, and environmental medicine*, Vol. 76, 01 2006, pp. 1111–8.

6. Bles, W., et al., *Experiments On Motion Sickness Aboard The M.V. Zeefakkel*, TNO Human Factors Research Institute, Soesterberg, The Netherlands, Report IZF-1991-A-34, 1991.

7. De Graaf, B., et al., *Exploratory Study With A Sea Sickness Questionnaire On Hr. Ms. Frigate Tromp*, TNO Human Factors Research Institute, Soesterberg, The Netherlands, Report IZF-1992-A-1, 1992.

8. Kennedy, R.S., et al., "Simulator Sickness Questionnaire : An Enhanced Method for Quantifying Simulator Sickness," *The International Journal of Aviation Psychology*, Vol. 3, (3), 1993, pp. 203–220. DOI: 10.1207/s15327108ijap0303

9. Brunnström, K., et al., "Quality Of Experience For A Virtual Reality Simulator," *IS and T International Symposium on Electronic Imaging Science and Technology*, March 2018. DOI: 10.2352/ISSN.2470-1173.2018.14.HVEI-526

10. Balk, S.A., et al., "Simulator Sickness Questionnaire: Twenty Years Later," *Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, 2013, pp. 257–263. DOI: 10.17077/drivingassessment.1498

11. Hart, S.G., Staveland, L.E., "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," *Advances in Psychology - Human Mental Workload*, Vol. 52, 1988, pp. 139–183.

12. Field, A., and Hole, G., *How to Design and Report Experiments*, Sage Publications Ltd, Newbury Park, California, USA, 2003.

13. Stroosma, O., Van Paassen, M., and Mulder, M., "Using the SIMONA Research Simulator for Human-machine Interaction Research," *AIAA Modeling and Simulation Technologies Conference and Exhibit*, 2003. DOI: 10.2514/6.2003-5525

14. Anon., *Interior view of the SRS cockpit*, `https://flightsafety.org/wp-content/uploads/2019/01/humanfactors_dec18-jan19_fig3.jpg`, 2019.

15. Anon., *Epson MOVERIO Augmented Reality Smart Glasses*, `https://epson.com/moverio-augmented-reality`, 2019.

16. Microsoft Corporation, *HoloLens Hardware Details*, https://developer.microsoft.com/en-us/windows/mixed-reality/hololens_hardware_details, 2017.

17. Saarikko, P., et al., *Waveguide*, Google Patents, US Patent 9,513,480, 2016.

18. Ashley, J., *How Hololens Displays Work*, John Wiley and Sons, New York, USA, http://www.imaginativeuniversal.com/blog/2015/10/18/how-hololens-displays-work/, 2015.

19. Anon., *DreamGlass Developer Edition*, https://www.dreamworldvision.com/product-page/dreamglass-headset, 2019.

20. Viertler, F., and Hajek, M., "Evaluation of Visual Augmentation Methods for Rotorcraft Pilots in Degraded Visual Environments," *Journal of the American Helicopter Society*, Vol. 62, (1), 2017, pp. 1–11. DOI: 10.4050/JAHS.62.012005

21. Yanoff, M., and Duker, J.S., *Ophthalmology*, MOSBY Elsevier, Amsterdam, The Netherlands, third edition, 2009, p. 54.

22. Baskett, B., *Aeronautical Design Standard Performance Specification Handling qualities Requirements for Military Rotorcrraft*, Defense Technical Information Center, Fort Belvoir, Virginia, USA, 2000.

23. Pardi, A., *Improving Helicopter Simulator Fidelity with Augmented Reality - Final Report*, Delft University of Technology, Faculty of Aerospace Engineering, 2020.

24. Likert, R., *A Technique for the Measurement of Attitudes*, Vol. 140, 1932, p. 1–55.

25. Massey, F., "The Kolmogorov-Smirnov test for goodness of fit," *Journal of the American statistical Association*, Vol. 46, (253), 1951, pp. 68–78.

26. Shapiro, S., and Wilk, M. B., "An Analysis of Variance Test for Normality (Complete Samples)," *Biometrika*, Vol. 52, (3/4), 1965, pp. 591–611.

27. Field, A., *Discovering Statistics Using SPSS*, Sage Publications Ltd, Newbury Park, California, USA, 2005.

28. Claveria, O., "A New Metric of Consensus for Likert Scales," *SSRN Electronic Journal*, 09 2018. DOI: 10.2139/ssrn.3255555

29. Anon., *New F-35 Helmets Cost £260,000 Each*, Telegraph Media Group Limited, London, United Kingdom, https://www.telegraph.co.uk/news/uknews/defence/11963873/New-F-35-helmets-cost-260000-each.html, 2015.

30. Davenport, C., *Meet The Most Fascinating Part Of The F-35: The $400,000 Helmet*, Washington Post, Washington D.C., USA, https://www.washingtonpost.com/news/checkpoint/wp/2015/04/01/meet-the-most-fascinating-part-of-\the-f-35-the-400000-helmet/, 2015.

Part II

Book of Appendices

A | Introduction to Augmented Reality

In the definition given on Wikipedia, Augmented Reality is "*an interactive experience of a real-world environment where the objects that reside in the real world are enhanced by computer-generated perceptual information, sometimes across multiple sensory modalities, including visual, auditory, haptic, somatosensory and olfactory*". This definition is very wide, covering a large number of uses and technical implementations, however the immersive nature of the experience is the common denominator of all these options. The visual augmentation of reality is a concept that can be traced back to the ghost illusion first described in the 16th century by Giambattista della Porta, now commonly known as Pepper's Ghost. By using a large angled glass pane and by tuning the levels of illumination it is possible to make the viewer perceive objects in the scene that are not actually present. This illusion was made available to large audiences in theatres as shown in Figure A.1. The underlying principle of manipulating light to "build" virtual objects did not change in time, to the point that even the most recent AR products can be compared technically to the first theatrical uses, as can be seen in Figure A.6 for the Microsoft HoloLens. Modern AR makes use of CGI and optical image combiners to achieve this effect, merging the added visual information to that of the environment before presenting the result to the viewer. Better results are achieved if the virtual objects appear to belong with the real ones in the visual scene, a result possible only if characteristics such as size, position, orientation, luminance and contrast are finely tuned.

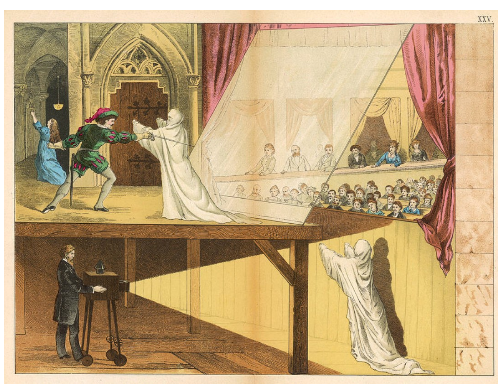


Figure A.1: Pepper's Ghost Illusion from Mary Evans Picture Library [1], illustrating the 19th century theatrical application of the AR principle to create the illusion of human figures on the stage.

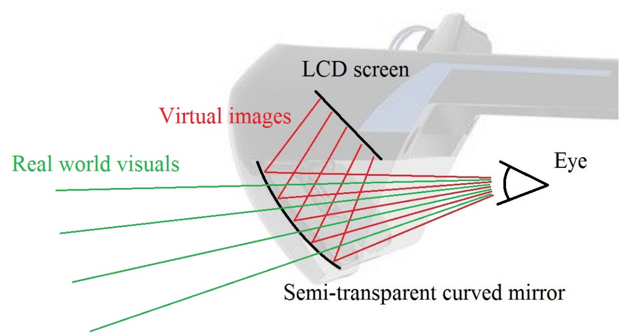


Figure A.2: DreamGlass single optical combiner AR design [2]. Image generated on the LCD screens is reflected by a single lens before entering the eye. The lens is a semi-transparent curved mirror that acts as image combiner for virtual and real-world images.

A major boost in AR research and technology occurred in the second half of the 20th century, as the HUDs (Head-Up Display) of military fighter aircraft became more and more advanced. What started in WW2 as a simple fixed aiming point for machine guns, focused at infinity, evolved to include initially flight parameters like altitude and velocity, and later navigation and radar targeting information. The last step was the shift from an airplane-mounted device to its helmet-mounted equivalent version, making augmented information available regardless of where the pilot was turning his head. Because of low production numbers and exclusively military application AR remained a very expensive technology, as exemplified by the F-35 helmet (see Figure A.3) reported to cost between 300 and 400 thousand euros apiece [3]. In more recent years investments in mobile phone and computer gaming industries have lowered the costs of both development and production, allowing its use in automotive industry, medical sciences (see Figure A.4) and educational fields. A simultaneous growth in computational power improved the initially limited graphics (text, pictograms, simple shapes) to the point that entire photo-realistic virtual environments can now be shown inside the AR headsets. With low costs, good CGI quality and portability achieved, AR today is being used in a vast number of fields, as virtual personal assistants, help in medical surgery, oil industry maintenance, industrial design and many more.

Figure A.3: Composite image of the Lockheed Martin F-35 Gen.I helmet demonstrating AR capabilities in night flight mode [4].

Figure A.4: Composite image of a medical application where AR is used while performing surgery [5].

Augmented Reality is closely related to Virtual Reality, a similar technology that instead of augmenting the real-world with virtual content isolates the user and brings him/her inside a completely virtual one, for immersive experiences that do not directly relate with the immediate surroundings. In VR the user has the outside vision blocked by the headset, and can only see virtual CGI overlaid on a black background. By getting rid of the real-world VR solutions can create virtual environments in a much more consistent way, without having to deal with changing environments and light levels. Furthermore the additive light principle, used in projection systems of both AR and VR, produces the best results in zero-light conditions, exactly those found inside VR headsets. The virtual worlds in VR can be inspired by and also accurately reproduce real-world locations and contexts, making the distinction between AR and VR less neat. One of the most clear differences remaining is about body proprioception when interacting with the computer program.

Human-computer interaction has been relegated to traditional textual input and visual output for decades, later followed by pointing devices like mouse, joysticks and electronic drawing pads. As new controllers were developed, a special interest was dedicated to the most complex controller of the human body, the hand. The level of dexterity and precision that human hands routinely achieve in real-life is very difficult to replicate when interacting with a computer, and virtual object manipulation is a classical example of this. One of the many factors contributing to such levels of dexterity is proprioception of our own limbs, in this case the hands. By observing our own hands we can more easily and more precisely coordinate their movement in space, e.g. achieving a good grip of an object in front of us. If something similar wants to be obtained with virtual objects the task of AR is much easier than that of VR. Both AR and VR solutions have to understand the complete geometry of the hand, by means of cameras and other sensors as exemplified by the HoloLens 2, but only VR has the added task of generating a 'virtual hand' for visualization purposes, as the AR user can already see his/her hands directly.

From a technical point of view there are three main types of projection systems in AR: the 'birdbath', the single combiner and the waveguide type. Examples of these methods are illustrated in Figure A.5 for the Magic Leap One, in Figure A.2 for the DreamGlass and in Figure A.6 for the Microsoft HoloLens. Each type has different limitations in terms of ambient light loss, achievable FOV range, image quality and cost. In the first case the number of internal reflections and lens coatings results in a loss of up to 95% of light, meaning that the real-world environment will appear dark and that very bright projectors will have to be used for the virtual images. The width of the FOV is limited to ≈ 20○ in the waveguide-based products because of light physics preventing further evolution (Figure A.6). Image quality can be lost in many ways some of which cannot be avoided, like the prismatic diffusion in waveguide gratings creating glowing effects or the chromatic aberrations due to ambient light sources.

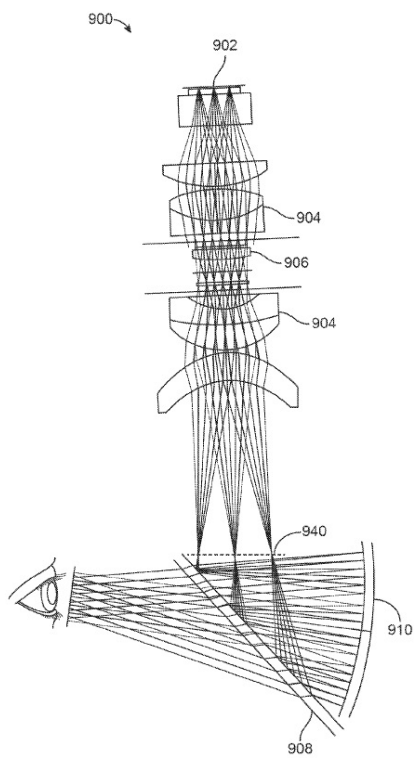


Figure A.5: MagicLeap drawing of 'birdbath' AR projection method, US patent 2015/0346495 [6]. Images generated at the top of the figure (902) pass through optical lenses (904-906) and are reflected to the right by a flat semi-transparent mirror (908), called beam-splitter. Then they are reflected back to the left by a semi-transparent curved mirror (910), a meniscus lens that also provides focusing power.

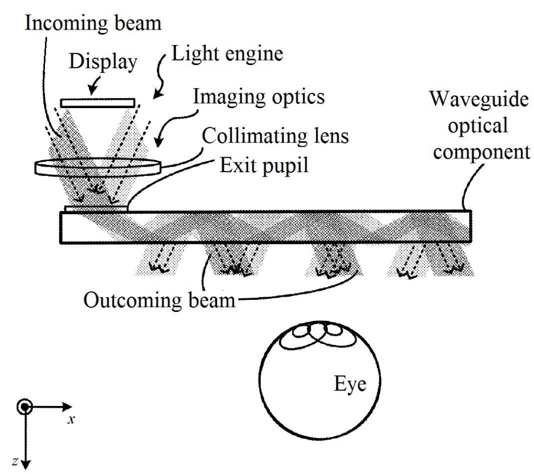


Figure A.6: Microsoft HoloLens waveguide US patent 9,513,480 [7]. Technical drawing showing of the Microsoft HoloLens waveguide principle based on diffraction gratings. Images generated by the display pass through a series of optics before entering the waveguide component. Here the light rays are reflected by various gratings levels by bending them at angles close to 90○. Result is the eye perceiving images based on its orientation and not its lateral position.

Each AR implementation can be different depending on the field and exact use, with designer's choices being affected by other factors and restrictions such as indoor/outdoor usage, local health laws, workplace regulations, acceptance by insurance companies as professional tools, acceptance of the market as entertainment devices. Before building the working prototype for this thesis project knowledge was gathered on the more technical aspects of AR. Krichenbauer studied [8] such AR/VR difference and found out that user interaction tasks were completed in AR test sessions faster compared to VR. The use of AR in automotive traffic context was studied by Abdi [9] to determine the graphical characteristics of the symbology presented to the user, e.g. size, contrast, transparency. In another research Aragon [10] studied how AR increases safety in helicopters by providing airflow information in hazardous and stressful situations, for example when landing on ships with turbulent winds, through the use of adequate symbology.

AR was also studied in general aviation by Beringer [11], where landing approaches were examined in flat and challenging terrain during a pilot-in-the-loop experiment. The use of different augmentation systems (head-up and head-down displays) and varying symbology confirmed the increase in confidence for experienced pilots with HMD AR. A review of modern AR devices for aviation applications by Bailey [12] summarized the current limitations, constraints and future developments of Head-Worn Displays (HWD) as part of the FAA NextGen initiative.

Spatial mapping is an important characteristic in AR and can be present in various levels, depending on the exact implementation, from 'no mapping' to full 'conformal mapping' [10] [13]. The term describes the degree of coupling between virtual and real objects, by analyzing the spatial relationship between them.

In the case of 'no mapping' the virtual content presented to the user has no relation with the real world, for example in the case of an oil industry worker visualizing a maintenance manual on the AR goggles while operating machinery on-site. The other end of the scale is full 'conformal mapping', where from the user's perspective every object in the virtual scene matches its counterpart in the real world. Continuing with the example of the oil industry worker, he/she would now see information such as operating instructions, or actions to be taken, directly overlaid on the related items like valves or levers, with spatial consistency maintained also when the head is turned in different directions.

Intermediate levels of 'partial mapping' exist where virtual objects appear in a fixed position of the user's field of view, that are only related but not overlayed to real-world objects, as shown in Figure A.7. It is clear how more conformal spatial mapping levels correspond to higher degrees of technical complexity in the AR solution, as the environment needs to be scanned continuously for features, and the head of the user tracked in 3D space. In Figure A.8 the real-world scenery is used as canvas for the virtual objects, represented by the blue traffic lane CGI exactly matching the real one. Achieving full conformal mapping could be viewed as the default goal of all AR applications, however this is questionable, as use cases exist where augmented content is better perceived and understood by the user through 'partial mapping'. An example of this is when the real-world objects used as anchor for the virtual content are outside the field of view of the user or is moving erratically or too fast, making tasks such as reading instructions or more difficult.

Figure A.7: Composite image of AR application with partial spatial-mapping. [14] The worker in the scene is using a pair of Epson Moverio BT-200 to access virtual content related to the real object in front of him but not conformally overlayed to it [14].

Figure A.8: [15]

From a technical stand point, in order to properly render objects in the AR headset, a calibration of the device is often needed taking into account the position of the user's eyes and their interpupillary distance. In his work Grubert [16] reviewed the various procedures used in the industry in the last twenty years, and their requirements. The optical quality of the AR displays was addressed by Hess [17], starting from assessment methods for the fidelity of flight simulators. In his work Hess investigated how visual degradation in the display affects human pilots performing a continuous control task.

The ability of AR devices to provide the user with a high-fidelity experience is affected by many factors, some more closely related to the technological side and some others following human biology. Typically AR aims at blending real-world experiences with additional virtual information, therefore the reference condition is the operator moving in the real world. In the case of this project the reference condition is twofold, as the pilot in the SRS is interacting with a physical cockpit and flying in a simulated environment at the same time. This duality requires the AR solution to maintain relations with both entities at all times, on a hardware and/or software level, so that AR visuals are consistent and do not to trigger motion-sickness and other discomforting feelings in the operator.

A.1 The Human Factor

From a biological point of view we navigate the world using a collection of sensory information, of which vision is only a part. In order to maintain an erect posture the direction of the force of gravity needs to be continuously and accurately calculated, so that muscles can be used to maintain equilibrium in a smooth fashion. The semi-circular canals in the inner ears evolved to provide this information, but in the following stage a form of complex sensor fusion occurs, where discrepancies are detected between the ears and with respect to other senses. Semi-circular canals themselves are not ideal sensors, as they have detection thresholds and a drifting behaviour that needs to be corrected from time to time. Visual information coming from the eyes is used for this purpose, as a stationary environment offers many reference points to check against. The relation between vestibular and visual systems is bi-directional, with the first helping for example in the compensatory movement of the eyes during feature tracking, a task linked to evolution by the need to hunt prey.

As in all closed-loop control systems their stability and output quality depends on the components and their interaction. Conflicting information during sensor fusion, even if managed by e.g. ignoring one of the sources, can result in incorrect outputs. Within the vestibular system inconsistency can produce two main effects, an incorrect position/orientation determination and nausea symptoms. Depending on the situation the first can have serious consequences: a standing person can lose equilibrium and fall to the ground, while a flying pilot could suffer from spatial disorientation and possibly put the aircraft in a dangerous attitude without realizing it. The feelings of nausea are believed to be connected with the scare of food poisoning, as this was the most frequent case during our evolution phase, when disorientation was the main tell-tale of poisoning and quickly emptying the stomach (emesis) most often saved the day [18].

In more modern times these two effects can be experienced by VR users when motion and linear accelerations, such as those found in car-racing rides, are only presented visually as the user sits still in a room. The sensory discrepancy is perceived as the brain fails to detect the onset of real accelerations, while the VR system is showing visual information that is consistent with their presence, e.g. while taking a curve in a car. The eyesight is suggesting a body-balancing reaction but all the other sensors (vestibular, kidney, large blood vessels) suggest otherwise. The food-poisoning defence mechanism is then triggered and the user becomes quickly uncomfortable and motion-sick. A comparison between the same game running on a computer monitor and a VR headset offers more insight on the topic.

In the first case the central vision of the user is filled by the game visuals, as the computer monitor shows the images of the roller-coaster ride. However the peripheral areas of his/her vision are filled with objects that provide static points of reference. It is then reasonable to believe that, as long as the amount of this peripheral information is greater than that the one in the central vision, the sense of vision will solve the incongruity favoring the former. In other words there would be intra-sensory but not inter-sensory discrepancy, and this would be resolved before the sensor fusion stage, preventing in general the onset of motion-sickness symptoms.

By using larger and larger monitors the amount of static peripheral information is conversely reduced, with VR headsets bringing this to the final step, removing completely the vision of the user's real surroundings, and resulting in faster onset of motion-sickness and discomfort. Indications confirming this can be found in the lack of success of dynamically-oriented VR games and the existence of "pause safe modes" in VR headsets like the Oculus Quest [19], using IR sensors and camera (normally used for 3D positioning) to see the real surroundings and postpone/avoid the onset of motion-sickness.

One of the common features of AR solutions is to let the user see the real environment around him/her, either by direct vision or panoramic cameras projecting inside the goggles. It must be said that current projection technology is limited in its field of view from the eye perspective, covering a much smaller cone compared to the eye field of regard, this intended as the sum of all possible directions the eye can point at. This technical limitation implicitly guarantees the persistence of static environment visual information in the peripheral area of vision, preventing any interference by the added CGI. At some point, as technology evolves, this assumption will cease to be valid and AR headsets will exist where projection areas cover the full field of view, thus requiring some care to avoid the pitfall of motion-sickness.

A.2 Frames of reference

Sensory discrepancies can arise from other sources than the human body, and AR solutions (intended as headset and computer program) are not exempt from motion-sickness and discomfort in the user. A series of steps needs to be working in a coordinated way, to accomplish consistency of visual information between the headset and the outside world. A hitch in any of these steps can degrade the user experience in the same ways as described above for the VR case. Perfect consistency between generated imagery and real world visuals means that geometric characteristics of the user's eyesight should be calculated continuously and without flaws, so that a corresponding computer image can be calculated and presented in the headset.

The position of the eyes inside the head is assumed not to change in time, even though this is not completely true as described later. The next step is to calculate the relation between the frame of reference of the head and that of the headset. This can be made constant in time by achieving a firm fit of the goggles on the head of the user, however this is not easy as it seems because in the field of ergonomics the requirements of 'tight fit' and 'wearer comfort' are often conflicting ones. This can be observed in the automotive industry where many devices exist to tension safety belts only if and when a crash is occurring, not before. In AR and VR headsets the wide range of skull bone geometries poses a challenge to the designers of retention systems. Thanks to the lightweight construction of most recent AR headsets this challenge is less demanding, as larger accelerations of the user's head are required to dislodge the device and misalign the images. The addition of fittings such as nose-rests further helps in holding fast the goggles and remove any initial angular difference, so that the whole transformation step can be simplified and treated as a static transformation.

Once the geometric relations between eyes, head and headset have been determined the next step is to find similar relations between the latter and the environment. AR systems can feature a mix of inertial and non-inertial sensors, therefore the presence of additional transformation layers, e.g. the Stewart platform of a full-motion simulator, requires additional care.

In the most simple case of the user standing in a room, detection of Earth's magnetic and gravitational fields (inertial references) can provide for orientation/attitude angles while IR/sound markers and sensors (non-inertial references) produce the 3D coordinates. Other systems for motion capture like the OptiTrack use light sensors for all calculations [20]. Most AR products have linear and/or rotational accelerometers inside the HMD, while others have cameras to detect the relative motion of the captured images and calculate, by inversion, the motion of the HMD in the environment.

The wide range of sensing technologies available requires an adequate selection in the design phase, and the development of good fusion algorithms, as imperfections can have negative effects on the chain of frames' transformations and even introducing instabilities. Some of these imperfections are: delay, imprecision and inaccuracy of the raw measurements, tendency to drift in time, noise, jitter, analog-to-digital sampling issues and sensitivity to external signals like sources of strong light or EM-waves. Often the electronics inside sensors perform some type of filtering, smoothing, interpolation or extrapolation, with resulting added delays, lags and leads to the signal. If these imperfections are not minimized or compensated for the result will be a time/space disalignment between the images projected inside the headset and the outside environment, and depending on its magnitude the user will start to perform worse and be ultimately subject to disorientation and nausea.

This integration process is completed during the development phase and implemented with embedded systems and software drivers in the AR solution. Because of the complexity of this process, its sensitivity to modifications and the need to protect intellectual property it is often impossible to inspect/modify its internal mechanisms. An example is the attitude/positioning suite is the Microsoft HoloLens, where only the final 6-variable output is available to developers.

If eyesight were a perfect, point-like mechanism the work would be over. Instead more aspects need to be dealt with. On a larger scale simplification of human vision can be justified, as even the binocular feature can be neglected above certain distances, measured in multiples of the Inter-Pupillary Distance baseline. On one hand the small perceived rotation of distant objects produces images in the left and right eyes that are not distinguishable, on the other hand the angular difference between the direction of the pupils can be neglected. In AR however the close distance between projection devices and the pupils requires more understanding and precision. The front part of the eye deflects incoming light rays to form a consistent and focused image in the rear half of the eye, on the retina. When the eyeball rotates in its socket the center of rotation does not correspond with the retina, causing a slight shift of the viewpoint. If this shift is not taken into account by the AR projection technology a misalignment will occur between virtual images and real-life environment and tasks requiring high angular precision would be affected. Solutions based on waveguides like the HoloLens are not affected by this shift of the retina as their holographic lenses generate the same image when looked at from different viewpoints, resulting in the retina seeing the same visuals regardless of its rotation state. The waveguide technology however has important drawbacks, such as the see-through display not covering more than 20° of the field of view of the pupil (no peripheral vision possible), chromatic aberrations of most ambient light sources and a substantial reduction of the amount of light passing through the display (even by 85%) due to the gratings in the lenses [21].

Each eye focuses on an object by adjusting the geometry of the front part, where the cornea is housed (accomodation), and at the same time by having both eyes aim at the same point to achieve stereoscopic vision (vergence). This poses a problem as the AR projection system should allow the same 'real-life' viewing experience also in the computer-generated imagery, with the user able to focus closer and farther 'virtual objects'. This capability has not been achieved in full by any AR technology. Having two separated projection systems allows to provide each eye with virtual images calculated from different view-points, solving the vergence aspect of the problem. Achieving eye accomodation also for virtual images has currently been achieved only in prototypes as the common stereoscopic vision methods are unable to provide for it. More complex varifocal, multifocal, light-field and holographic technologies can be used, often needing high computation power or constant eye-tracking. One example of multifocal application is the 3D wearable AR display from LightSpace Technologies [22], where extremely thin screens (6–12 micrometers) are stacked on top of each other, creating a miniature volumetric display that is then expanded through a magnifying lens.

The solution adopted by most AR manufacturers is to have a single focusing plane for the virtual images and, if needed, to simulate out-of-focus objects by blurring their virtual representation, the same method used in movies to manipulate the attention of the audience. Choice of the exact focusing distance is different across manufacturers, as it must fit the expected/designed scope of use of the AR solution. Head-up displays as found on military airplanes, though not being stereoscopic, all have the same focusing distance. This is set to infinity because the real-world objects visible behind the HUD normally are at distances of hundreds of meters (takeoff/landing excepted). The HoloLens instead, having been designed for more office- and desk-oriented activities, has a fixed focusing distance of the virtual images of 85 cm. A different choice was made in the DreamGlass HMD, where a 200 cm focusing distance was adopted. This is in line with 'industrial design' usage expectations, possibly featuring a standing AR user and 'room-sized' virtual objects like cars and house furniture.

B | Methods to Assess Simulator Sickness

The first documented accounts of Motion Sickness (MS) date back to the ancient Greeks, with the word 'nausea' itself stemming from 'naus', greek for 'ship'. It was only in late 19th century that scientific studies [23] started connecting MS with the semicircular canals, the level of familiarity with the type of motion and the functionality of the vestibular system in general. The list of the motion environments where MS can take place includes *"ships, small boats, life rafts, coaches, trains, automobiles, aircraft, spacecraft, carnival rides, swings, tilted rooms, rotating rooms, rotating chairs, camel riding, elephant riding, earthquakes, and newly prescribed eye glasses"* [24]. The last example is of particular interest as it implies that alterations in the visual system by means of deformations of the visual field can cause temporary MS.

In the 20th century the logistics involved in moving soldiers by ships and trains in WW1 and WW2 resulted in more MS research efforts, also to help build the first simulators for aircraft and helicopters. Johnson [24] studied sensory conflict and postural instability theories in the context of helicopter simulation, starting from the puzzling results of the first rotorcraft simulators built in the 1950s. He found out that helicopter instructors were more affected by MS than the students, and that the symptoms shown in the simulator corresponded only in part to those normally associated to MS. Different rating scales were then developed to assess motion sickness in flight simulation, using both objective and subjective criteria. Depending on testing conditions it is possible to measure the physiological indicators associated with MS such as changes in heartbeat rates, ocular activity or changes in skin tone, however determining the reference cases and interpreting data for each participant can be time-consuming. Subjective criteria, on the other hand, rely on the judgement of the participants themselves, and take the form of questionnaires to be filled during the tests or right after them. This self-assessment process must be completed on-the-spot or in a short period of time, as long delays alter the perception of the symptoms evaluated in the questionnaire, making the data less accurate.

B.1 MISC method

The so-called 'MIsery SCale', or MISC [25] [26], is a rating method originally developed at TNO in Soesterberg (NL) to measure motion sickness in ships, and later adopted in automotive, aviation and simulation contexts. The method uses a single scale with eleven discrete values describing the level of discomfort experienced by the test participant, as presented in Table B.1. One of the underlying assumptions is that MS can be measured along the given scale, excluding for example different symptoms for head and stomach. This simplification follows inevitably from selecting a single entry in the MISC scale, and while this could be considered a loss of analytical power it must be said that the overall health status of the participants is still captured. This also allows the questionnaire to be filled with very little effort and without time delays, making repeated measurements easier.

Table B.1: Misery Scale for assessment of Motion Sickness symptoms [25] [26], translated in English.

Symptom Description	**Rating**
No Problems	0
Stuffy or Uneasy Feeling in the Head	1
or	2
Stomach Discomfort	3
or	4
Nauseated	5
or	6
Very Nauseated	7
or	8
Retching	9
Vomiting	10

B.2 SSQ method

An other wide-spread method to assess MS in simulation is the Simulator Sickness Questionnaire, or SSQ [27], first outlined by Kennedy and others in 1993. After more than 20 years the method has proved to be effective [28] and is used for scientific research also in AR/VR fields [29].

The main purpose of SSQ was to revise the previous MSQ method, accounting for the symptoms' differences between simulation and actual ships, cars and airplanes [24]. The original MSQ listed between 25 and 30 separate symptoms to be evaluated by the test participants, each on a discrete severity scale going from zero to a maximum value. Lessons learned in the 1980s showed that new simulators caused MS in a smaller population percentage, and that symptoms were less severe than in the past [24]. The MSQ method was producing less significant values in these new conditions and a more powerful method was needed. An additional role was also devised for the SSQ, to help trouble-shooting simulators with poor MS performance.

First the SSQ criteria were selected from the original MSQ list based on their actual occurrence in simulators, with 16 symptoms being retained out of 28. Then a three-factor solution was adopted, grouping the symptoms in three clusters labeled 'Nausea', 'Oculomotor', and 'Disorientation' (NOD) as presented in Table B.2. The purpose of these clusters is to provide analytical insight on which technical aspect of the simulator is responsible for its poor MS performance. Because the focus is on MS symptoms the SSQ is of no value to compare simulators that are not problematic.

Table B.2: Computation of scores in the SSQ method using weighted sums of symptom-related rating scores and conversion formulas, adapted from [27]. The '-' has been added for better readability in place of '0' or an empty field.

SSQ Symptom	**Nausea**	**Oculomotor**	**Disorientation**
General discomfort	1	1	-
Fatigue	-	1	-
Headache	-	1	-
Eyestrain	-	1	-
Difficulty focusing	-	1	1
Increased salivation	1	-	-
Sweating	1	-	-
Nausea	1	-	1
Difficulty concentrating	1	1	-
Fullness of head	-	-	1
Blurred vision	-	1	1
Dizziness (eyes open)	-	-	1
Dizziness (eyes closed)	-	-	1
Vertigo	-	-	1
Stomach awareness	1	-	-
Burping	1	-	-
Weighted sum	N-sum	O-sum	D-sum
Nausea score =	N-sum x 9.54		
Oculomotor score =	O-sum x 7.58		
Disorientation score =	D-sum x 13.92		
Total score =	N-sum + O-sum	+ (D-sum x 3.74)	

Each symptom can be featured in more than one cluster, and is assessed on a subscale of four values, from zero to three. The exact composition of each cluster was decided following a varimax-rotated matrix calculated using data from hundreds of simulator sessions. The resulting weighting matrix is made only by ones and zeroes, to indicate the presence or absence of that particular symptom in the cluster. This weighted sum is calculated over scores within each cluster and multiplied by coefficients (9.54, 7.58 and 13.92 respectively), so that three distinct scores are obtained for 'Nausea', 'Oculomotor', and 'Disorientation'. An overall total score (TS) is then computed with a second weighted sum over the NOD scores, with weights being 1, 1 and 3.74.

C Answers to Open Comments Questionnaire

This appendix contains all the data from the open comment questionnaires used in the experiment. For every one of the seventeen statements the Likert scale results are presented graphically using a different shade of grey for each possible rating ('Strongly agree', 'Agree', 'Neither', 'Disagree' and 'Strongly disagree'). Spontaneous comments are added in *italic* with a reference to the participant number.

The questionnaire provides insight in three main areas: confirmation of the experiment design in terms of setup and visual libraries (section C.3), possible comfort issues related to the AR goggles (section C.2) and finally usefulness of the added AR visuals for the current experiment (section C.1). Each question is formulated as a positive or negative statement about a clearly identifiable quantity or feature, e.g. *"The weight of the AR headset was excessive"*. With the use of 5-step Likert scales a general idea of the consensus (if any) can be obtained about each topic, while written comments give more depth into new, potentially unexpected issues.

Usefulness of the added AR visuals (section C.1)

- Statement 3 - Observing the right chin window in AR helped me maintaining the hover position. (subsection C.1.1)
- Statement 8 - The piloting task can be completed without the added AR visuals. (subsection C.1.2)
- Statement 16 - Watching the left side window in AR helped me maintaining the hover position. (subsection C.1.3)

Comfort issues of AR goggles (section C.2)

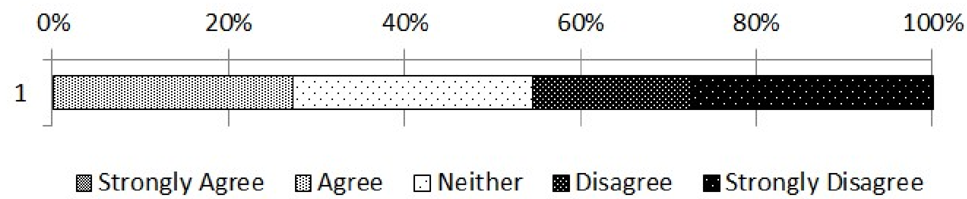
- Statement 2 - There was an excessive lag (latency) in the images seen in the AR headset. (subsection C.2.2)
- Statement 4 - When moving my head the images in AR changed smoothly, with no noticeable jumps. (subsection C.2.3)
- Statement 5 - The image quality in the AR headset (contrast, luminosity) was very similar to that of the simulator screens. (subsection C.2.4)
- Statement 6 - The overlap of virtual and screen visuals in the left side window was very noticeable. (subsection C.2.5)
- Statement 7 - The AR headset was very comfortable to wear. (subsection C.2.6)
- Statement 10 - The weight of the AR headset was excessive.(subsection C.2.7)
- Statement 12 - I could always focus precisely on objects inside the AR visuals. (subsection C.2.8)
- Statement 13 - The AR headset temperature was often too high. (subsection C.2.9)
- Statement 14 - The connecting cables of the AR headset allowed me to turn my head around with ease. (subsection C.2.10)
- Statement 17 - Refocusing between the AR visuals and the simulator screens was tiring for the eyes. (subsection C.2.11)

Confirmation of the experiment design (section C.3)

- Statement 1 - There were enough objects and references in the environment to determine my longitudinal position. (subsection C.2.1)
- Statement 9 - Not having flight instruments (e.g. the attitude indicator) was a big inconvenience in this experiment. (subsection C.3.1)
- Statement 11 - There were enough objects and references in the environment to determine my angle of pitch. (subsection C.3.2)
- Statement 15 - The environment had enough objects and references for me to determine my vertical position. (subsection C.3.3)

C.1 Usefulness of the added AR visuals

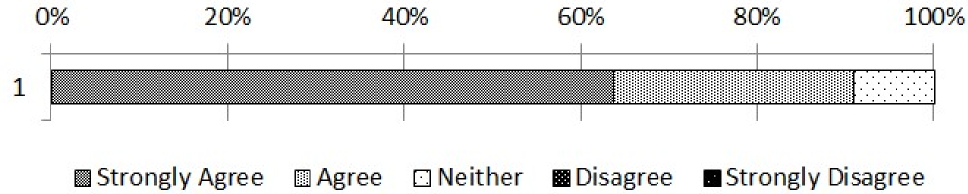
C.1.1 Statement 3: Observing the right chin window in AR helped me maintaining the hover position.



Comments:

- Participant 1: *"The wires of the AR headset made it very cumbersome to move the head"*
- Participant 3: *"Only looking at board and cones"*
- Participant 8: *"Not significantly used"*
- Participant 10: *"Main problem: small FOV - up and down makes it impossible to quickly glance in corner"*
- Participant 11: *"I never used the chin window because it would take too long to move my head back to the pole reference"*
- Participant 13: *"It did [help] maybe once or twice but mostly did not use it"*
- Participant 14: *"Only with right cones, only in peripheral vision"*
- Participant 15: *"Did not use this window much"*
- Participant 16: *"Seeing circle [ground marker] helps in knowing when you need to really stop"*
- Participant 19: *"Whenever I tried to use it the position got unstable (for my feeling)"*

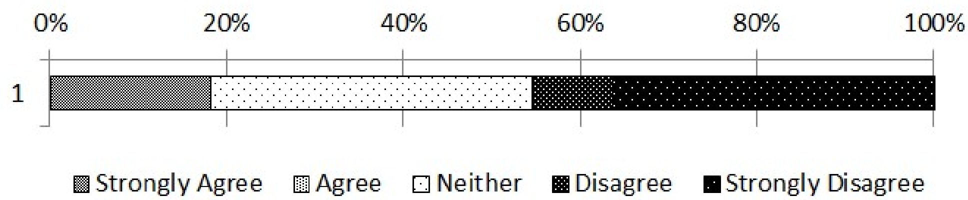
C.1.2 Statement 8: The piloting task can be completed without the added AR visuals.



Comments:

- Participant 3: *"I almost did not feel like I needed them except sometimes with cones on left when past cones. Without AR falling back to diagonal cones"*
- Participant 10: *"It was very challenging with and without"*
- Participant 11: *"I did not use any of the added AR visuals"*
- Participant 13: *"I was mainly using the projected screens and not much of the chin windows"*
- Participant 14: *"The cones and hoverboard were always my main reference points"*
- Participant 15: *"Helps to some degree, but I can complete without AR"*
- Participant 16: *"Can be completed without [AR] but seeing circle [ground marker] helped, as turning head down felt less disorienting / distracting than sideways"*

C.1.3 Statement 16: Watching the left side window in AR helped me maintaining the hover position.

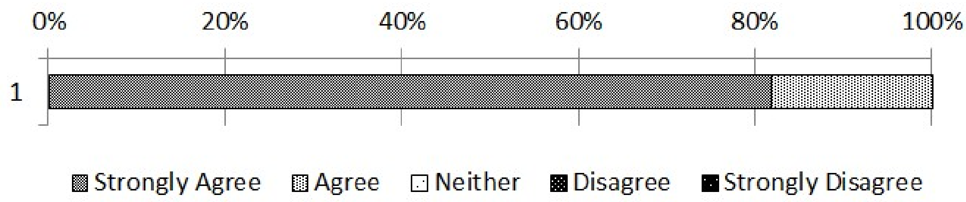


Comments:

- Participant 3: *"Only when flying past the cones it was really useful otherwise it could be done without but made it easier"*
- Participant 8: *"Although I tried to keep looking ahead a quick look to the cones on the left really helped determining position"*
- Participant 10: *"I did not use it at all"*
- Participant 11: *"I never used the left side AR window"*
- Participant 13: *"Did not use it"*
- Participant 14: *"The latency on AR made me see everything twice, making it almost impossible [to use]"*
- Participant 15: *"Did not help at all"*
- Participant 16: *"I mostly used diagonal cones as less head tilt was required."*
- Participant 19: *"Too blurred and too delayed to be a real help. I only looked out of curiosity, but that distracted."*

C.2 Comfort aspects of the prototype

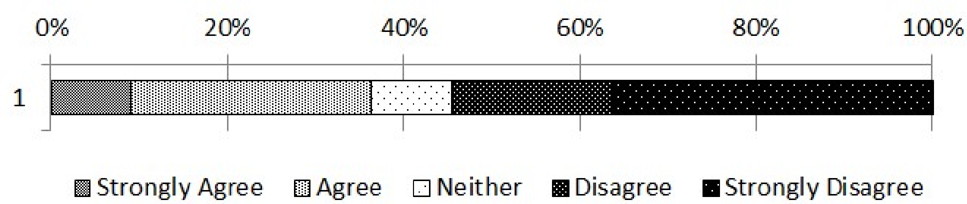
C.2.1 Statement 1: There were enough objects and references in the environment to determine my longitudinal position



Comments:

- Participant 3: *"Cones are enough"*
- Participant 11: *"I used the pole and the cones left and right to determine the longitudinal position"*
- Participant 13: *"I mainly used the sphere an hoverboard with the cones diagonally in front of me. I did not use the side cones and only the ground marker once or twice"*
- Participant 14: *"Especially the diagonal cones were useful. Chin window only useful when already in peripheral vision"*
- Participant 15: *"Referred objects to [the] right, cones at 45 degrees and straight forward best [?]"*
- Participant 19: *"Texture close to the helo to assess forward speed / I mainly used the cones"*

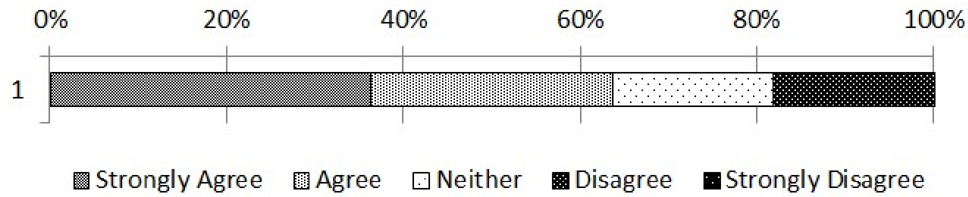
C.2.2 Statement 2: There was an excessive lag (latency) in the images seen in the AR headset.



Comments:

- Participant 3: *"When moving / scanning front board to cones there was a slight lag which can be found annoying"*
- Participant 10: *"Mostly noticeable with complete overlay"*
- Participant 11: *"I had to adjust my control gains to compensate the AR lag"*
- Participant 13: *"No lag I could notice"*
- Participant 14: *"Never noticed it (during flight) since I only used it as a very loose reference"*
- Participant 15: *"Smooth as butter"*
- Participant 19: *"My subjective feeling is that the delay was too much for helicopter control. Without the 'SIMONA less delayed image[s]' -> unstable"*

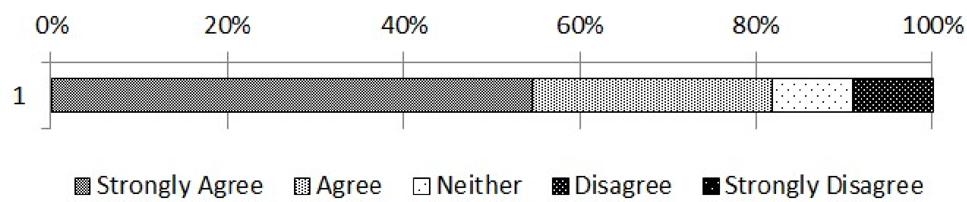
C.2.3 Statement 4: When moving my head the images in AR changed smoothly, with no noticeable jumps.



Comments:

- Participant 1: *"Some jumps. During flight it became less apparent"*
- Participant 3: *"Although slight lag but smoothly"*
- Participant 13: *"When my head was still (or felt still) there was some small wobble in the AR images"*
- Participant 14: *"Only slightly noticed it during flight"*
- Participant 15: *"AR headset very smooth"*
- Participant 19: *"Smooth, but very delayed"*

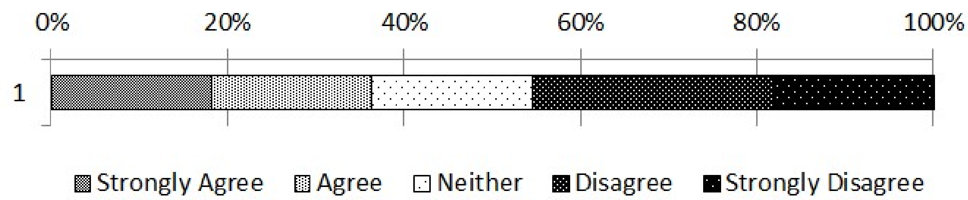
C.2.4 Statement 5: The image quality in the AR headset (contrast, luminosity) was very similar to that of the simulator screens.



Comments:

- Participant 8: *"Contrast, luminosity: yes. resolution: no"*
- Participant 11: *"Maybe the AR had slightly less contrast, but it was not distracting"*
- Participant 13: *"Yes I thought they were the same"*
- Participant 14: *"[AR] It was slightly more dimmed"*
- Participant 15: *"Very smooth transition from screen to AR"*
- Participant 16: *"[AR] bit more blurry / less sharp. [AR glasses] were bright enough"*
- Participant 19: *"Luminosity was perfect. Image looked identical, but with lower resoution, delay and focus problems."*

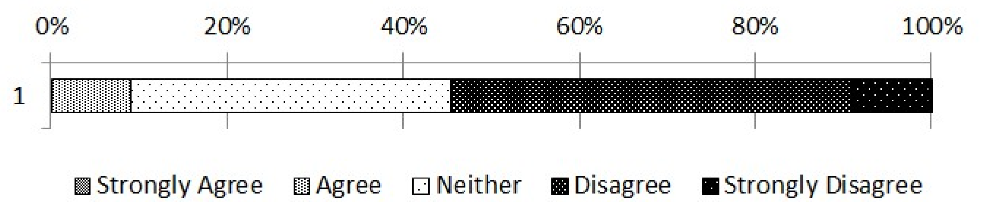
C.2.5 Statement 6: The overlap of virtual and screen visuals in the left side window was very noticeable.



Comments:

- Participant 3: *"They were slightly off (not fully calibrated) but good enough"*
- Participant 8: *"Mostly ignored the screen visual and only looked at the virtual window"*
- Participant 10: *"I barely looked so much to the left"*
- Participant 11: *"The left windows were not aligned, and sometimes the AR left window was even rotated by plus/minus 30 degrees"*
- Participant 13: *"I noticed it but it did not affect my flying"*
- Participant 14: *"The left side cones became unusable because of it"*
- Participant 15: *"Occasionally noticeable but not to an effort that ruined it"*
- Participant 16: *"Due to aligning not exactly there is a little shift. Moreover the AR is [a] bit more blurry / pixelated"*
- Participant 19: *"Did not notice. Brain filters it out. The trim displays [in the instrument panel] were also reflected on the outside visuals"*

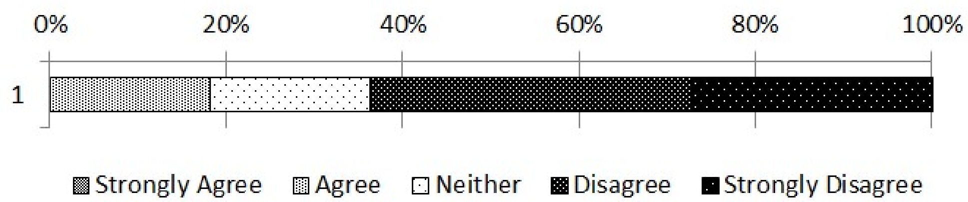
C.2.6 Statement 7: The AR headset was very comfortable to wear.



Comments:

- Participant 1: *"At first it is comfortable, but after sustained wear it becomes uncomfrortable"*
- Participant 3: *"Slightly pushing on forehead to make sure it stay where it should be"*
- Participant 8: *"Strain on forehead made it a little uncomfortable"*
- Participant 10: *"Ok by itself, but right ear started to hurt in combination with headset"*
- Participant 11: *"I had to strap the AR very tightly to focus the images"*
- Participant 13: *"It felt tight and heavy around the back of my head. Otherwise it was comfortable. Maybe could have a stop over the head for support"*
- Participant 14: *"Pushed on my [own] glasses, adding pressure, Did not get uncomfortable, however."*
- Participant 15: *"Was comfortable but for long periods of time would be noticeable"*
- Participant 16: *"Difficult to get first correct position. Tends to be somewhat heavier after some time"*
- Participant 19: *"I could not have kept it on 30 minutes extra -> headache"*

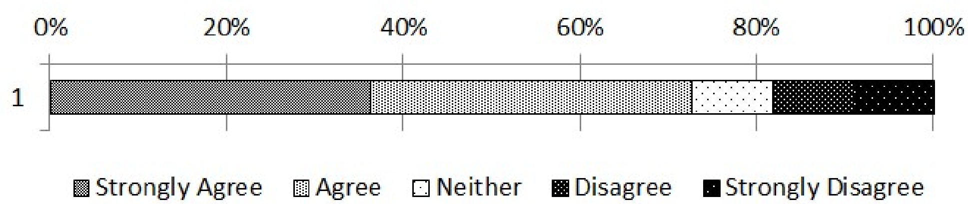
C.2.7 Statement 10: The weight of the AR headset was excessive.



Comments:

- Participant 1: *"Too heavy for comfort after sustained use"*
- Participant 3: *"Not too heavy"*
- Participant 13: *"It is quite heay and did start to weigh on the back of my head"*
- Participant 14: *"Did not notice it"*
- Participant 15: *"Weight was never an issue"*
- Participant 16: *"Could be lighter, but cable pull at [the] back sometimes made it move."*

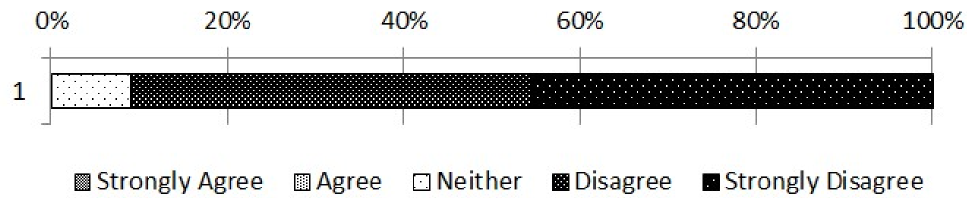
C.2.8 Statement 12: I could always focus precisely on objects inside the AR visuals.



Comments:

- Participant 3: *"If I looked at them yes"*
- Participant 11: *"I could clearly distinguish the objects from the environment, but not as clearly as the SRS visuals"*
- Participant 13: *"No problems focusing"*
- Participant 14: *"The circles [ground marker] didn't feel blurry"*
- Participant 15: *"Have to adjust at beginning but usually was fine afterwards"*
- Participant 16: *"No problems"*
- Participant 19: *"I had a hard time to get the AR image focused. The top text [during centering] was never in focus. Then when I looked as Simona [screens], AR was blurred again"*

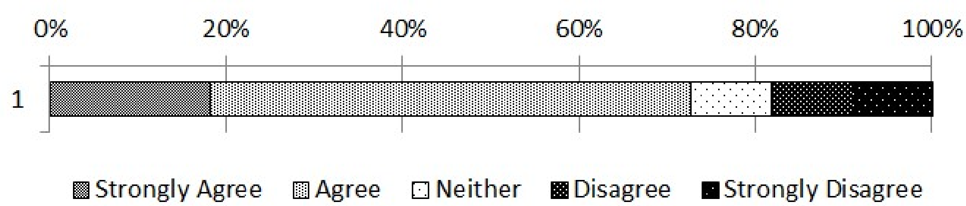
C.2.9 Statement 13: The AR headset temperature was often too high.



Comments:

- Participant 10: *"Too busy to notice"*
- Participant 13: *"Temperature was ok"*
- Participant 14: *"Didn't notice anything"*
- Participant 15: *"Was noticeable but not very warm"*
- Participant 16: *"Maybe [a] bit warmer on forehead but not discomforting"*
- Participant 19: *"Mainly pressure"*

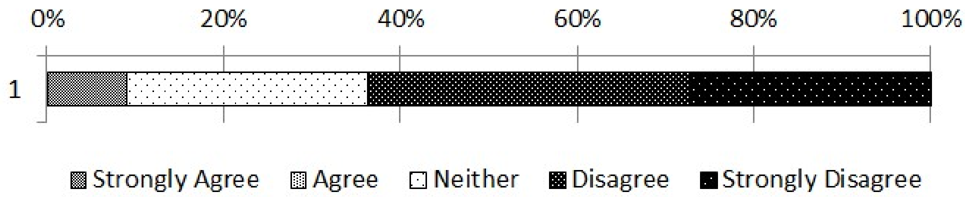
C.2.10 Statement 14: The connecting cables of the AR headset allowed me to turn my head around with ease.



Comments:

- Participant 1: *"The cables made it very cumbersome to turn my head"*
- Participant 3: *"Cables are ok, only when tilting forward and looking left it was a little bit short"*
- Participant 8: *"Only felt the cable pulling once during all experiments"*
- Participant 13: *"No restrictions of movement"*
- Participant 14: *"Once we moved them into place properly"*
- Participant 15: *"Would have to move the cables sometimes, but would eventually adjust them to have freedom"*
- Participant 16: *"As they quite often got [a] bit stuck behind my back and the chair"*
- Participant 19: *"I fixed the cables with my back to not give tension while allowing about 30 degrees motion"*

C.2.11 Statement 17: Refocusing between the AR visuals and the simulator screens was tiring for the eyes.

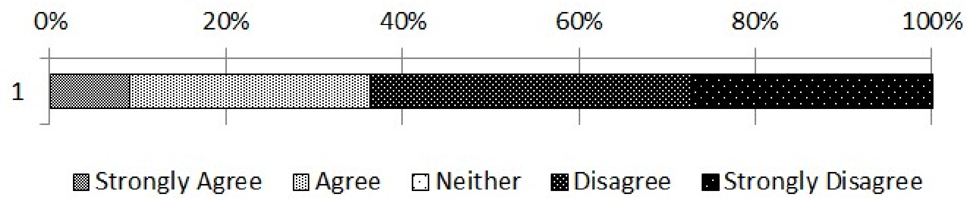


Comments:

- Participant 3: *"It was easy"*
- Participant 8: *"Not noticeable"*
- Participant 10: *"Barely used AR visuals"*
- Participant 11: *"I mainly focused on the SRS screens, not on the AR visuals"*
- Participant 13: *"Was no problem"*
- Participant 14: *"I never noticed anything"*
- Participant 15: *"Maybe only occasionally, [it] was not a real issue"*
- Participant 16: *"As long as not both screens and [AR] glasses showed the world everything was fine"*
- Participant 19: *"I just focused on SIMONA [screen] to get the job done and accepted blurry AR images"*

C.3 Confirmation of experiment design

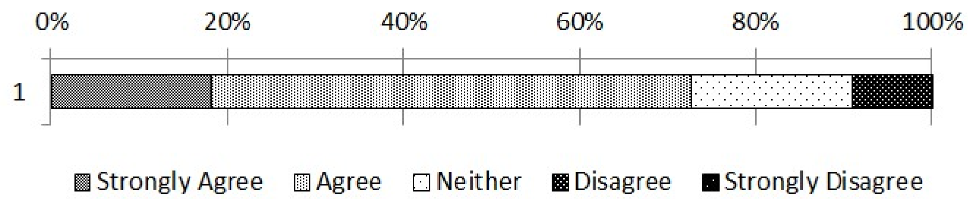
C.3.1 Statement 9: Not having flight instruments (e.g. the attitude indicator) was a big inconvenience in this experiment.



Comments:

- Participant 1: *"A horizon indicator would definetly be helpful"*
- Participant 3: *"Not needed, all states can be retrieved from outside visual"*
- Participant 10: *"Pitch can only be estimated by resulting velocity, when it was already too late"*
- Participant 11: *"A vertical speed indicator and a virtual horizon would have made it much easier to hover"*
- Participant 13: *"For me it would have been helpful to have flight instruments"*
- Participant 14: *"I never missed ithem, you get enough information from the runway or the ramp. An artificial horizon may have helped, especially at first."*
- Participant 15: *"Would have been nice to see an artificial horizon at least"*
- Participant 16: *"Maybe pitch attitude [would have] helped. But environment and especially cones helped."*
- Participant 18: *"It's a visual maneuver, so no need to look at instruments"*
- Participant 19: *"In this phase a pilot would look at altitude, speed, and power"*

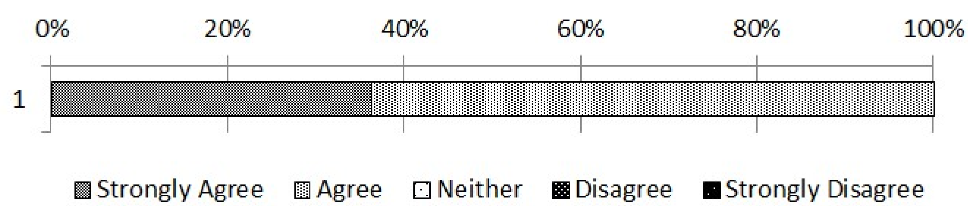
C.3.2 Statement 11: There were enough objects and references in the environment to determine my angle of pitch.



Comments:

- Participant 1: *"Hard to determine pitch"*
- Participant 10: *"The angles are too small to judge"*
- Participant 11: *"I could determine my angle of pitch from objects in the environment, but it would have been more accurate with an attitude indicator."*
- Participant 13: *"It would have been helpful to have an attitude indicator. Near the beginning it was hard to know my pitch andle but as I hot more used to it, I could know it better."*
- Participant 14: *"It is possible with the runway and the ramp. It did take some [time?] getting used to."*
- Participant 15: *"Sometimes would be confused [confusing?], only when pitch was plus/minus a few degrees from zero"*
- Participant 16: *"Cones helped, especially [the] right ones as those were more easy to see."*
- Participant 19: *"The lack of motion made a difference and felt weird"*

C.3.3 Statement 15: The environment had enough objects and references for me to determine my vertical position.



Comments:

- Participant 3: *"Board + mountain / trees, good"*
- Participant 10: *"The board was very helpful"*
- Participant 11: *"Especially the pole was very helpful"*
- Participant 13: *"Again would have been useful to have an altimeter maybe"*
- Participant 14: *"The hoverboard was fine, I never really used AR"*
- Participant 15: *"The ball and hoverboard helped a lot"*
- Participant 16: *"Pole helped and [also] looking at buildings"*
- Participant 19: *"Texture and stereo-vision normally help with height. Here only the gold ball was used."*

D | Contents of the Compressed Archive

All the experiment material has been gathered in a single compressed archive file, including raw data files, SPSS and spreadsheet statistical analysis files, documents and the likes. The archive is organized in several folders:

- 01 - Prototype Source Code
- 02 - Raw Data Files
- 03 - Statistical Analysis Files
- 04 - Images Library
- 05 - Experiment Setup Instructions
- 06 - Experiment Forms
- 07 - CAD Models

Due to anonymization each participant is identified solely by the number assigned to him/her during the initial registration process. Because some participants did not complete the experiment some numbers do not appear in the archives. The list of participants' numbers featured in the archive is the following: 1, 3, 8, 9, 10, 11, 13, 14, 15, 16, 18 and 19.

Folder '01 - Prototype Source Code'

This folder a single compressed file of about 3 gigabytes containing the Unity 3D folder of the project. This archive contains about 160,000 files and sub-folders, and is designed to be opened with Unity 3D 2019 after having been de-compressed. Files are of many different types: images, sound, 3D shapes, shaders, graphical assets, projects settings, XML configuration files and source code in C, C++ and C#. Because of the large number of files it is recommended to open the archive on a fast SSD-type hard drive.

Writing, editing and de-bugging of the source code were accomplished with Microsoft Visual Studio 2017. The project was designed to be built both for Windows and Linux platforms, as the former is required by the AR goggles while the latter is used by the three SRS image servers.

Folder '02 - Raw Data Files'

This folder contains twelve compressed files, each with multiple timetrack files in HDF5 format, as produced by the DUECA platform. The twelve files are named after the participant's number, e.g. "P11.rar".

The names of the timetrack files includes a timestamp to facilitate order-by-time and matching with the test run documents, e.g. "P01-20200226_103006.hdf5". Files with dimensions below 10 kilobytes correspond to the calibration sessions occurring between the flights, and can be ignored.

Folder '03 - Statistical Analysis Files'

The timetrack data files were analyzed with the help of various software packages. All the scripts and intermediate data formats used in the process are gathered in this folder, organized in subfolders 'MATLAB', 'SPSS' and 'Spreadsheet'.

In folder 'SPSS' all files start with 'DraftSimAR_' followed by the name of the associated DM. Each dependent measure has its SPSS syntax file (with extension '.sps') containing the commands used in the statistical analysis process. These files, when run inside SPSS, load the corresponding data set (with extension '.sav') and produce output tables and graphs in a new Output Viewer Window that can be saved (with extension '.spv'), customized in format format and exported as desired. A template file with extension '.sgt' was used to generate plots in the same style. Please note that the plots in the final report were refined afterwards with the addition of properly sized axis labels.

The folder 'Spreadsheet' contains seven files generated in Microsoft Excel 2010:

- 01 - Number of Flights.xlsx
- 02 - Flight Performance and Detrend.xlsx
- 03 - SPSS Tests Output.xlsx
- 04 - SSQ.xlsx
- 05 - MISC.xlsx
- 06 - TLX.xlsx
- 07 - Open Comments Likert.xlsx

Information about the number of flight sessions completed by the participants and their outcome is contained in file '01 - Number of Flights.xlsx'.

In file '02 - Flight Performance and Detrend.xlsx' the figures about the five DMs related to flight performance (calculated in the MATLAB script) are used as input and de-trended for inter-participant variability. The result of this process is then used in SPSS as input for statistical analysis.

The output of all statistical tests run in SPSS is saved in file '03 - SPSS Tests Output.xlsx'. In particular the first three tabs contain information about the K-W and S-W normality tests, the homogeneity of variance tests and the pairwise Friedman's tests. The fourth tab contains the calculation of consensus coefficients about the Likert scores of the 'Open Comments' questionnaire.

The remaining four files contain the ratings given by the participants in the four questionnaires used in the experiment (MISC, SSQ, TLX and Open Comments ones), and are used to calculate percentages and generate the plots used in the final report.

Folder '04 - Images Library'

This folder contains all the pictures and figures used in the final report and additional material documenting all the experiment phases. The files are loosely organized in sub-folders by topic.

Folder '05 - Experiment Setup Instructions'

In the development phase a series of procedures were formalized in documents to keep track of all the steps needed to re-create the experiment. All these documents, after proper anonymization, have been gathered in this folder. Please note that some information could become outdated as it refers to particular software versions and workstations that could be modified or become unavailable in time.

File 'Accounts on TUD278713 workstation.docx/.pdf' document the user accounts on the indicated workstation, including information on privileges and type of account.

File 'Catia to Unity conversion procedure.docx/.pdf' describe how to bring 3D graphical elements from Catia CAD software into the Unity 3D engine using MeshLab as stepping stone. This simple process can be used to generate 3D occlusion masks of any cockpit geometry available in Catia with millimeter accuracy.

In file 'Headset setup instructions.docx/.pdf' the correct procedure to connect the DreamGlass and the NOLO tracker to the workstation is described.

At some point in the development phase the Unity 3D engine, by default running on Windows 10, was installed on Ubuntu 18.04.3 LTS to generate Linux-compatible executables. After a series of successful tests, done using a ported version of the original project, it was decided to use a single instance of Unity on Windows to generate all executables (for Windows and for Linux). The procedure used to setup the Ubuntu workstation is included for future projects in ‘Setup of Ubuntu development workstation.docx/.pdf’.

A similar document is provided for the Windows 10 workstation where the Unity 3D project was developed, named ‘Setup of Windows development workstation.docx/.pdf’, with information provided also on the Windows accounts to use. The procedure used to create the new project under Unity, valid in both Windows 10 and Ubuntu 18.04.3 LTS, is documented in ‘Setup of Unity project.docx/.pdf’, where the DreamGlass and NOLO instructions are also referenced.

Folder ‘06 - Experiment Forms’

All the forms written for the experiment are gathered in this folder. Four of the documents were provided in hard-copy to the participants, namely the consent form (‘01 - Informed Consent Form.docx/.pdf’), the briefing (‘02 - Participant Briefing.docx/.pdf’) and the various questionnaires (‘03 - MISC SSQ TLX Questionnaires.docx/.pdf’ and ‘04 - Open Comments Questionnaire.docx/.pdf’).

The other two documents were used by the test coordinator to operate the SIMONA simulator (‘05 - SRS operator guide.docx/.pdf’) and to keep track of the flight sessions (‘06 - Experiment run tables.xlsx’). For convenience four of the files are also included in appendices F, G, H and J of this document.

Folder ‘07 - CAD Models’

This folder contains all files developed in CATIA V5 R20, with 3D models of the S-70 helicopter and of the SRS cabin in various configurations. Because of the conversion process described in Appendix N, every part is available as CATIA part, MeshLab and Collada formats.

The complete geometry of the S-70 helicopter is provided in folder ‘S-70 Model’, in particular in its MH-60 Jayhawk variant in service with the U.S. Coast Guard [30]. The model is made by multiple parts: the cabin section, the tail section, the main rotor, the tail rotor and the various separate windows. To account for inaccuracies between frames and windows the windows are available in their original size and in a 5% larger area. The same approach is used in folder ‘SRS Cabin Model’ for the geometry of the SRS cockpit and its windows. The SRS cockpit geometry was built from scratch using data points from a 3D laser scan, where the windows’ borders were measured from the pilot’s DERP. In both models the cockpit interiors, like flight controls and pilot seats, have been discarded as they are not relevant for the experiment.

Within the game engine, various materials, colors and opacities levels are assigned to these parts from time to time, so that the AR occlusion masks can be assembled. Throughout the CAD files the term “Positive Cockpit” is used to identify the ‘opaque cockpit / transparent windows’ combination, while the term “Negative Cockpit” identifies the ‘transparent cockpit / opaque windows’ setup. The latter can be used to assess conformality of images between AR goggles and SRS cockpit, as the two visual sources would ideally complement each other without overlapping.

E | Specifications of other AR goggles

In this appendix a brief presentation is given of the other AR devices considered for this thesis project in the preliminary stage, with the most relevant technical parameters summarized in Table E.1.

Table E.1: Relevant technical parameters of AR devices considered for the thesis project.

Device	**Resolution (pixels)**	**Horiz. FOV**	**Vert. FOV**	**Refresh rate**
Human eye [31] [32]	≈ 15M per eye*	≈ 210°	≈ 130°	n.a.
Microsoft HoloLens 1 [21] [33]	1268 x 720	29°	16°	60 hz
DreamGlass [2]	1280 x 800	78°	48°	60 hz
Meta 2 AR	855 x 850**	85°	50°	60 hz
Magic Leap One	n.a.	50°	n.a.	60 hz
Epson Moverio BT-350 [14]	1280 x 720	23°	n.a.	30 Hz

* In variable-density eye regions
** Due to warping, nominally 1280x1440

Figure E.1: Microsoft HoloLens 1 [33]

Figure E.2: Meta 2 AR

Figure E.3: Magic Leap One

Figure E.4: Epson Moverio BT-350

F | Form for Informed Consent

Experiment Consent Form – *Flight Tests*

Improving Helicopter Simulator Fidelity with Augmented Reality

I hereby confirm that:

1. I volunteer to participate in the experiment conducted by the researcher **Alessandro Pardi** under supervision of **Ir. Olaf Stroosma** from the Faculty of Aerospace Engineering of the TU Delft. I understand that my participation in this experiment is voluntary and that I may withdraw and discontinue participation at any time, for any reason.
2. I have read the experiment briefing. Also, I affirm that I understand the experiment instructions and have had all remaining questions answered to my satisfaction.
3. I understand that my participation involves performing a simple flying task in a non-moving flight simulator setup.
4. I agree that the researcher collect the following data: previous flight experiences, answers to MISC, SSQ, TLX and general opinions questionnaires.
5. I agree that the following data is recorded during the experiment: pilot control inputs, aircraft states, specific forces, pilot head position and orientation.
6. I confirm that the researcher has provided me with detailed safety and operational instructions for the hardware used in the experiment.
7. I understand that the researcher will not identify me by name in any reports or publications that will result from this experiment, and that my confidentiality as a participant in this study will remain secure.
8. I have been given a copy of this consent form.

______________________________ My Signature	______________________________ Date
______________________________ My Printed Name	______________________________ Signature of Researcher

Contact Information Researcher	**Contact Information Research Supervisor**
Alessandro Pardi	Ir. Olaf Stroosma
--- email address ---	*--- email address ---*
--- telephone number ---	*--- telephone number ---*

G Experiment Briefing

Experiment Briefing – *Flight Tests*

Improving Helicopter Simulator Fidelity with Augmented Reality

This document serves as briefing for the second part of the experiment called "Improving Helicopter Simulator Fidelity with Augmented Reality", organized by the researcher Alessandro Pardi under the supervision of Ir. Olaf Stroosma.

After completing with success the AR Compatibility Test you will have the chance of joining the flight simulation test runs inside the SIMONA Research Simulator.

In this document you will find information on how the experiment is organized, and what the researcher expects from you. As your participation is entirely voluntary you can decide to stop the experiment at any point by informing the researcher, without providing any explanation.

Because of the innovative design of the AR goggles (Dream Glass by Dreamvision USA Inc.) some users can experience problems in wearing and regulating the goggles, and the participant will be assisted by the researcher in wearing them correctly and firmly.

Time schedule

The whole test session is made by four main parts: a briefing, a training session, a series of flight sessions and a debriefing. The approximate times for these activities are summarized here:

Start	End	Activity
00:00	00:15	Welcome and briefing
00:15	00:35	Training session
00:35	00:50	1st Test session – 3 flight test runs
00:50	01:05	2nd Test session – 3 flight test runs
01:05	01:20	3rd Test session – 3 flight test runs
01:20	01:35	4th Test session – 3 flight test runs
01:35	01:45	Debriefing

The experiment

As you may already know in this experiment I am trying to determine if enhancing flight simulators with Augmented Reality increases their behavioural fidelity (or not). In other words I am trying to 1) know if it works and 2) if it helps [the pilot] achieve a better performance.

Your role in the experiment is to fly a helicopter in a low-speed, low-altitude precision maneuver in four different test conditions. After every test run, each lasting less than two minutes, you will fill a quick questionnaire about your general condition, called the MISC rating, or MIsery SCale.

After completing the 3 test runs there will be time to answer questionnaires on different topics. The SSQ rating (Simulator Sickness Questionnaire) and the NASA TLX (Task Load indeX) focus on comfort and workload. At the very end of the experiment your opinion will be asked also on other aspects unique to this prototype, including AR.

Before starting with the actual test runs a training session will let you familiarize with the research simulator, the flight controls, the physics model of the helicopter and the Augmented Reality headset. The approach taken is in steps, starting with the helicopter stationary and no AR and gradually adding features and degrees of freedom. After adequate proficiency is achieved the training will stop and the actual flight tests will begin.

Flight maneuver

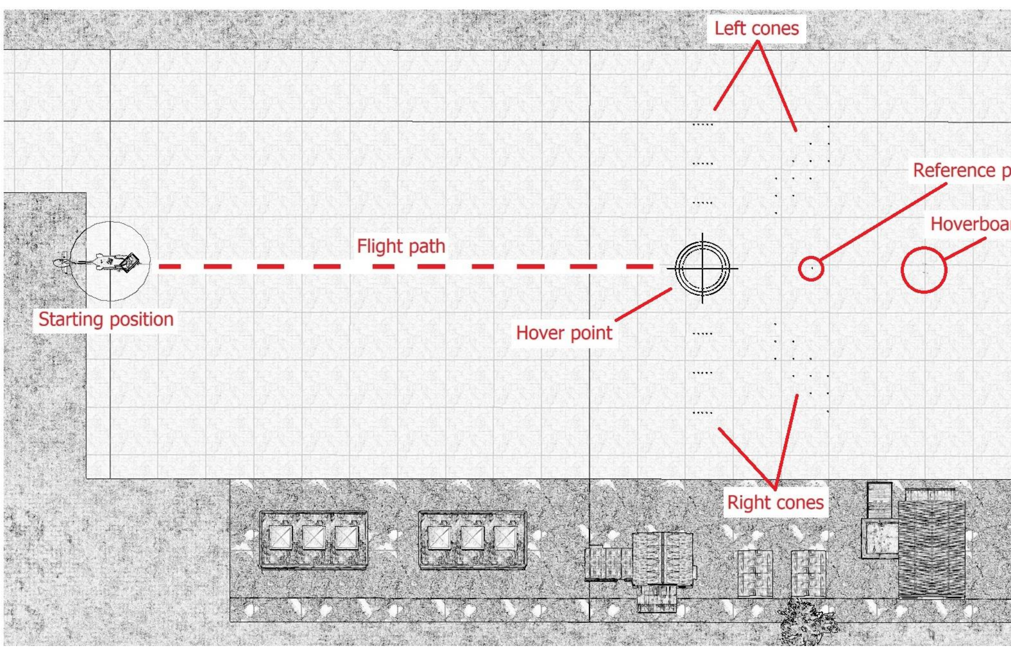
It is a variation of the ADS33 mission task element "Hover", where the pilot brings the helicopter above a target hover point and stabilizes it for a certain amount of time.

Achieving a **smooth and precise maneuver** is the objective here, and not minimizing the amount of time used to complete it (a maximum limit exists, obviously). As mentioned in the original ADS33 document: *"Accomplish the transition in one smooth maneuver. It is not acceptable to accomplish most of the deceleration well before the hover point and then to creep up to the final position".*

As presented in the following top-down drawing you will fly from a starting position to a final hover position, where you will stop the helicopter and keep it centered using the external references provided.

At the starting position the helicopter will be in trimmed conditions and with a small forward velocity, and you are encouraged to keep this configuration until you get close to the target area. At that point decelerate the helicopter smoothly and stop above the markings on the ground.

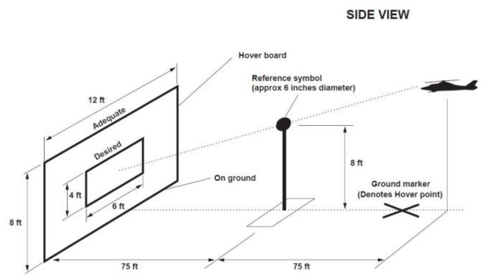
After 30 seconds of hover the maneuver is considered complete, and the test run is repeated.



Reference points

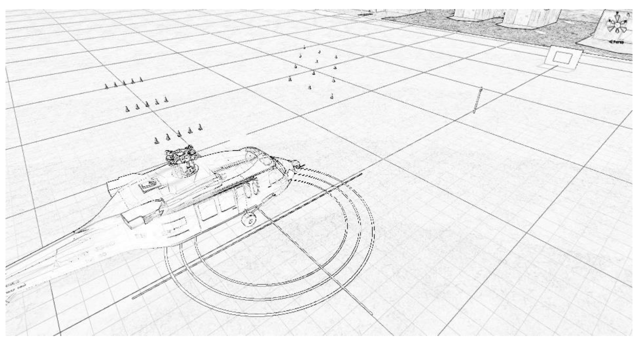
A series of objects have been placed in the environment to give reference points to the pilot, namely a realistic airport environment, a hoverboard with reference pole, two sets of traffic cones and a ground marker.

The next picture shows how the **hoverboard** can be used to determine your position above the hover point, aligning the reference symbol on top of the pole with the squares of the hoverboard.

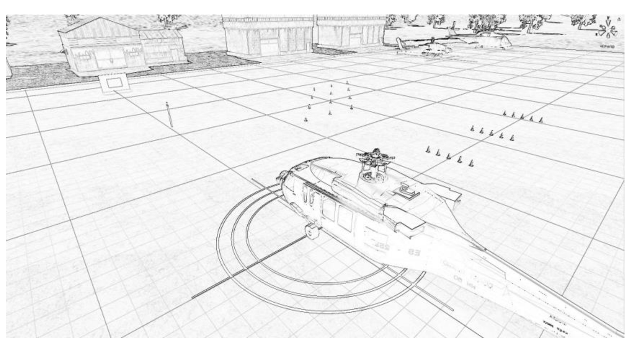


The circular **ground marker** seen in the two pictures below is designed to be seen through the chin windows of the helicopter, the right one in particular, with portions of the three circles all visible when hovering in the correct position.

The following picture shows the "Cones Left" test condition where two groups of **traffic cones** are visible to the left of the helicopter. When hovering in the correct position the central row of cones appears vertical.



A variation of the test track is obtained by moving the cones to the right (condition "Cones Right") as shown in the next picture. The hoverboard, reference pole and ground marker remain unchanged between the "Cones Left" and "Cones Right" conditions.



Questionnaires

The helicopter flight states and the position and attitude of the AR headset will be recorded, however a series of questionnaires will also be used.

After each test run (one test flight) the MISC (Misery Scale) rating will be filled, to monitor your general health condition quickly.

After the 3 flights making up one test condition the SSQ (Simulator Sickness Questionnaire) and TLX (NASA Task Load Index) questionnaires will be filled to investigate also the workload you are experiencing.

At the end of the experiment a final "general opinions" questionnaire will allow you to rate other aspects unique to this experiment including your opinion on the AR headset.

Participation and anonymity

Participation to this experiment is voluntary. You can in any moment decide to stop the experiment and leave it without providing any explanation. Also the data gathered is anonymized right after it, making it impossible to link your identity to the data.

In case of doubts or other questions the researcher will be happy to answer and give you feedback.

Thank you for joining !

Contact Information Researcher	**Contact Information Research Supervisor**
Alessandro Pardi	Ir. Olaf Stroosma
--- email address ---	*--- email address ---*
--- telephone number ---	*--- telephone number ---*

H MISC, SSQ and TLX Questionnaires

Participant # .... Test session # .... Date/time.............................

MISC Ratings

After each test run (after every flight) mark the box corresponding to your general condition during the test run you just completed. Columns 6 and 7 are included in case one or two test runs need to be repeated.

General condition		**Value**	**1st test**	**2nd test**	**3rd test**	**4th test**	**5th test**	6th test	7th test
No problems		0							
Slight discomfort but no specific symptoms		1							
Dizziness*, warm feeling, headache, stomach awareness, sweating, etc.	Vague	2							
	Some	3							
	Medium	4							
	Severe	5							
Nausea	Some	6							
	Medium	7							
	Severe	8							
	Retching**	9							
Vomiting		10							

SSQ Ratings

Fill this questionnaire after all 5 test runs, marking one box in each row according to how you feel now:

#	**Symptom**	**None**	**Slight**	**Moderate**	**Severe**
1	General discomfort				
2	Fatigue				
3	Headache				
4	Eyestrain				
5	Difficulty focusing				
6	Increased salivation				
7	Sweating				
8	Nausea				
9	Difficulty concentrating				
10	Fullness of head***				
11	Blurred vision				
12	Dizziness (eyes open)*				
13	Dizziness (eyes closed)*				
14	Vertigo				
15	Stomach awareness				
16	Burping				

*** Dizziness:** *"the state of being dizzy; the sensation of instability, of whirling, with a tendency to fall."*

**** Retching:** "*to make an unsuccessful effort to vomit; to strain, as in vomiting."*

***** **Fullness of head:** associated to higher-than-normal blood pressure, e.g. as experienced when lying supine on the back and raising the legs in the air.

TLX Ratings

NASA Task Load Index - Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Fill this questionnaire after the SSQ ratings one, marking the vertical bar that better describes your workload in each different area. At the end of the experiment you will determine which of these scales was more important in your opinion.

Mental Demand - How mentally demanding was the task?

Very Low | Very High

Physical Demand - How physically demanding was the task?

Very Low | Very High

Temporal Demand - How hurried or rushed was the pace of the task?

Very Low | Very High

Performance - How successful were you in accomplishing what you were asked to do?

Perfect | Failure

Effort - How hard did you have to work to accomplish your level of performance?

Very Low | Very High

Frustration - How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low | Very High

I Open Comments Questionnaire

Participant # .... Date/time ....

Opinion ratings

In this questionnaire select on the Likert scale how much you agree with each given statement. In the same box there is space for further comments in case you want to add them.

Statement	① Strongly Agree	② Agree	③ Neither	④ Disagree	⑤ Strongly Disagree
Q1 - There were enough objects and references in the environment to determine my longitudinal position.	①	②	③	④	⑤
Q2 - There was an excessive lag (latency) in the images seen in the AR headset.	①	②	③	④	⑤
Q3 - Observing the right chin window in AR helped me maintaining the hover position.	①	②	③	④	⑤
Q4 - When moving my head the images in AR changed smoothly, with no noticeable jumps.	①	②	③	④	⑤
Q5 - The image quality in the AR headset (contrast, luminosity) was very similar to that of the simulator screens.	①	②	③	④	⑤

Page **1** of **3**

Q6 - The overlap of virtual and screen visuals in the left side window was very noticeable.	1 Strongly Agree	2 Agree	3 Neither	4 Disagree	5 Strongly Disagree
....					
Q7 - The AR headset was very comfortable to wear.	1 Strongly Agree	2 Agree	3 Neither	4 Disagree	5 Strongly Disagree
....					
Q8 - The piloting task can be completed without the added AR visuals.	1 Strongly Agree	2 Agree	3 Neither	4 Disagree	5 Strongly Disagree
....					
Q9 - Not having flight instruments (e.g. the attitude indicator) was a big inconvenience in this experiment.	1 Strongly Agree	2 Agree	3 Neither	4 Disagree	5 Strongly Disagree
....					
Q10 - The weight of the AR headset was excessive.	1 Strongly Agree	2 Agree	3 Neither	4 Disagree	5 Strongly Disagree
....					
Q11 - There were enough objects and references in the environment to determine my angle of pitch.	1 Strongly Agree	2 Agree	3 Neither	4 Disagree	5 Strongly Disagree
....					

	①	②	③	④	⑤
Q12 - I could always focus precisely on objects inside the AR visuals.	Strongly Agree	Agree	Neither	Disagree	Strongly Disagree
Q13 - The AR headset temperature was often too high.	Strongly Agree	Agree	Neither	Disagree	Strongly Disagree
Q14 - The connecting cables of the AR headset allowed me to turn my head around with ease.	Strongly Agree	Agree	Neither	Disagree	Strongly Disagree
Q15 - The environment had enough objects and references for me to determine my vertical position.	Strongly Agree	Agree	Neither	Disagree	Strongly Disagree
Q16 - Watching the left side window in AR helped me maintaining the hover position.	Strongly Agree	Agree	Neither	Disagree	Strongly Disagree
Q17 - Refocusing between the AR visuals and the simulator screens was tiring for the eyes.	Strongly Agree	Agree	Neither	Disagree	Strongly Disagree

J | Experiment Control: Run Tables

Experiment Run Table

Date	
Time	
Participant	01

Test Conditions Sequence	4 - AR on, cones Right
	1 - AR off, cones Left
	2 - AR off, cones Right
	3 - AR on, cones Left

	Start	End	Notes
Briefing			
Training			

Test Condition	Test Run	Start	End	Result	MISC	Notes
4 - AR on, cones Right	1					
	2					
	3					
	4					
	5					
	[6]					
	[7]					

Test Condition	Test Run	Start	End	Result	MISC	Notes
1 - AR off, cones Left	1					
	2					
	3					
	4					
	5					
	[6]					
	[7]					

Test Condition	Test Run	Start	End	Result	MISC	Notes
2 - AR off, cones Right	1					
	2					
	3					
	4					
	5					
	[6]					
	[7]					

Test Condition	Test Run	Start	End	Result	MISC	Notes
3 - AR on, cones Left	1					
	2					
	3					
	4					
	5					
	[6]					
	[7]					

	Start	End	Notes
[Break 1]			
[Break 2]			
[Break 3]			
Debriefing			

K | Headset Setup Instructions

DreamGlass AR headset and NOLO CV1 tracking system

Headset Setup Instructions

Version 1.1 - Document originally written by A. Pardi in March 2019 and last updated by A. Pardi on 24 April 2019.

Introduction

This document describes the current best practices regarding the setup of DreamGlass and NOLO CV1 tracking system on the Dell Precision 5820 Tower computer. The procedure described needs to be followed by competent personnel accustomed to Windows 10 Enterprise and Unity 2018.f3, starting from an already working situation, i.e. where all needed drivers and programs have been installed and tested (Unity, Dreamglass drivers/SDK, NOLO drivers/SDK, etc.). A slightly different procedure can be used on the laptop Asus N56, as described later on.

Main items featured

- Dell Precision 5820 Tower computer with Windows 10 Enterprise and Unity 2018.f3, including DP/HDMI adapter
- Dreamworld Dreamglass augmented reality headset
- NOLO CV1 motion tracking system (head tracker, base station and two controllers)

General procedure

The main reason behind this document is to handle the latent incompatibilities between Dreamglass and NOLO CV1, summarized in the official Dreamworld documents as *"plug in NOLO first and Dreamglass second"*. Things are more complex as there are Windows drivers and NOLO processes involved, generating a scenario that sometimes is difficult to troubleshoot. To observe possible problems the following three items should be kept open and checked often: Windows task manager, Windows display settings and NOLO Home. For more information refer to the later section about issues.

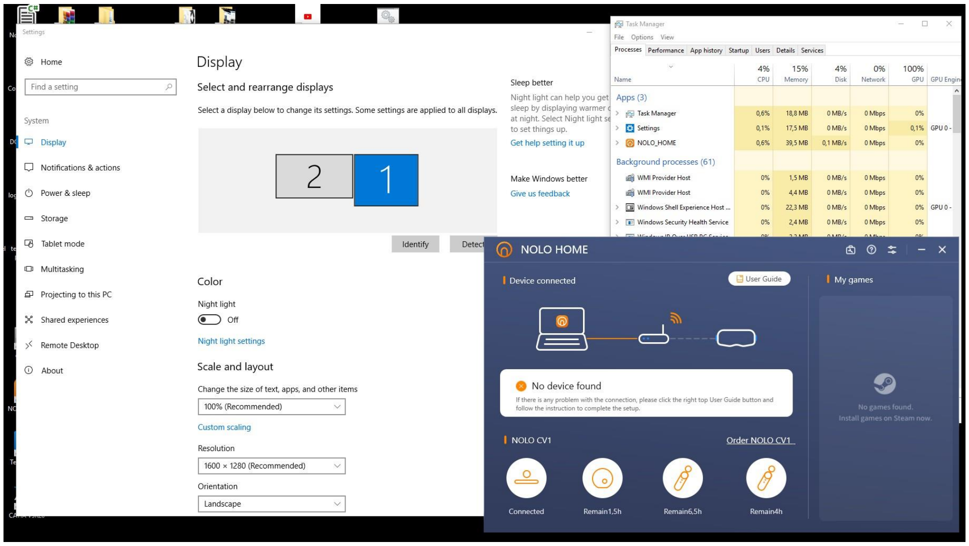
Detailed procedure

Starting conditions are: computer powered off, DP/HDMI adapter plugged in, all other devices not plugged.

1. Power on the computer and log in Windows
2. Unwind the cables of the DreamGlass and make sure they lie comfortably on your left shoulder
3. **Connect** the Nolo head tracker USB
4. Optional (can be postponed to any later time): power on the Nolo base station and the two controllers
5. Open the Windows display settings (right-click on the Desktop, "`Display settings`")
6. Open the Windows task manager (right-click on the taskbar, "`Task Manager`")
7. Launch Nolo Home (`Nolo_Home.exe`) and enter the Windows credentials of a local admin. (i.e. "`.\localadmin`")
8. Verify that the orange dials come to life accordingly to the actual switched-on devices, with the Headset Marker being the only really mandatory item (see screenshot on following page)
9. **Connect** the DreamGlass HDMI cable
10. **Connect** the DreamGlass USB (connector with two cables attached)
11. Wait 15 seconds: if USB drivers are being installed wait until all devices are "ready to use" (pen camera, audio pnp)
12. Verify in Windows Display Settings that a second monitor appears with resolution 1600x1280
13. Verify that its presence is constant in time, i.e. the new display does not "appear and disappear" every few seconds from the display settings window (see screenshot on following page)
14. Do not attach the secondary USB connector of the DreamGlass (the one with only one cable attached)
15. Procedure is complete: launch the Unity executable from the builds folder

TU Delft

Delft University of Technology



Shortened procedure for disappearing DG display

This procedure is applicable when, after some time measurable in tens of minutes, the DreamGlass display starts to show the behaviour described at **step n.13** of the standard procedure, namely the "appears and disappears" one. This is expected to be caused by the NOLO drivers, and can be often overcome by these steps:

1. Unplug DreamGlass HDMI and USB connectors and the Nolo head tracker USB.
2. Plug in Nolo head tracker USB connector
3. Plug in DreamGlass HDMI connector
4. Plug in DreamGlass USB connector (still the one with two cables)

Variation for particular laptops

This procedure is applicable to the Asus N56 laptop (i7-3630QM, Windows 8.1, GT-750M), the only differences being:

5. Both DreamGlass USB connectors are used at **step n.10** due to older USB 1.1 ports providing less energy
6. No DP/HDMI adapter is needed on the Asus laptop as HDMI is already present

Observed issues

USB drivers of the DreamGlass are installed again when the connector is plugged in. This seem to be related to Windows and does not have any impact if **step n.11** is followed. Exactly four USB devices get installed when plugging in the DreamGlass USB connector for the first time: "Pen Camera", "USB2.0 camera device", "USB pen PnP audio device" and "DREAM GLASS HID".

Multiple Nolo_WPFAssistant.exe processes are consistently present, Nolo Assistant GUI window shows no connected device (sometimes a flickering controller shows up momentarily). Normally this happens after the Unity application crashes or is closed in an unclean way. It is suggested to have Application.Quit() triggers inside the Unity game to close it in an explicit, clean way. Impact is total, i.e. no application will work unless the setup procedure is repeated.

The DreamGlass display appears and disappears every few seconds in Display settings. This is the main observable result of the conflict between DreamGlass and NOLO if the procedure is not followed. Impact is total, i.e. no application will work unless the setup procedure is repeated.

When any display goes in sleep mode the DreamGlass display starts to "appear and disappear" in Display settings. This is the same problem as mentioned before, and can be addressed with the shortened setup procedure. It is recommended to extend the display timer in the registry as described in the workstation setup document (`DreamGlass setup of development workstation.pdf`).

System interrupts in the Task Manager are consistently observed (15% cpu time) above normal levels. Conditions and causes of this problem have not been yet identified, apart from general considerations about multiple processes and drivers. Impact is total, i.e. no application will work unless the setup procedure is repeated.

(More technical information about interrupts: closing the following items did not stop the interrupts flood: stopped Unity executable, powered off Nolo base station, stopped the nolo assistant process. Disconnecting the DreamGlass USB actually caused interrupts to go to zero istantly)

TUDelft

Delft University of Technology

L | Procedure to Create a Unity Project

DreamGlass AR headset and NOLO CV1 tracking system

Setup of a Unity project

Version 1.3 - Document originally written by A. Pardi in March 2019 and last updated by A. Pardi on 5 December 2019

Introduction

This document describes the creation of new projects in Unity 2018.f3 that use the DreamGlass headset and the NOLO CV1 motion tracking system. Before using this document the preparation of the workstation must be completed, as indicated in the document "`DreamGlass setup of development workstation.pdf`". Some instructions are different when using a Linux Ubuntu workstation to build the project, as detailed later in this document.

General procedure

The main steps are:

1. Create a new Unity project,
2. Import the DreamgGlass and Nolo SDKs,
3. Follow the instructions in document "`Dreamworld SDK documentation_12_31_18.docx`" and
4. Build the project to check everything is fine.

Detailed procedure

1. Open Unity 2018.f3 and log in with your personal account
2. Create a new project in Unity using the template "`3D`"
3. *NOTE: Because of the internal dynamics of Unity renaming old project folders (and their contents) will not work. Copy-pasting folders just messes things up as internal objects and cloud entities continue to have the old identifiers, both human-readable and binary-encoded ones.*
4. **Import** in Unity the DreamGlass file `DreamWorld_2018.3.6.unitypackage` under menu "Assets" -> "Import Package" -> "Custom Package…", with all items
5. **Ignore** for now the Android-related error "`Unloading broken assembly Assets/DreamWorld/ML-Agents/Plugins/Android/TensorFlowSharp.Android.dll, this assembly can cause crashes in the runtime`"
6. **Import** in Unity the NOLO file `UnitySDK_v1.1.unitypackage.unitypackage` under menu "Assets" -> "Import Package" -> "Custom Package…", with all items
7. Follow instructions in the document "`Dreamworld SDK documentation_12_31_18.docx`", summarized here:
   a. **Delete** the `Main Camera` object in the scene
   b. From the folder "Assets\DreamWorld\Prefabs" **drag** the `DWCameraRig` prefab into your scene
   c. Add (if missing) a new resolution option to the Game's view, with name "`DreamGlass`" and resolution 1600x1280
   d. From the folder "Assets\NoloVR\Prefabs" **drag** the `NoloManager` prefab into the scene
   e. **Drag** the `DWCameraRig` object under `Hmd(camera)` to make it a child of it
   f. In the `Features` menu of the `DWCameraRig` object **change** the `Tracking` variable to `NOLO_6DOF`
   g. Focus the `NoloManager` object and **drag** `DWCameraRig` object onto the variable "`VR Camera`" field

TUDelft

Delft University of Technology

h. In folder "Assets\NoloVR\Scripts\Nolo_Unity" **comment** lines number 46-55 of the script `NoloVR_TrackedDevice.cs`
i. In their place **add** the 2 lines "`transform.localPosition = pose.pos;`" and "`transform.localRotation = Quaternion.Euler(pose.rot.eulerAngles.x, pose.rot.eulerAngles.y, pose.rot.eulerAngles.z);`", right below the commented lines 46-55.
j. In folder "<project_folder>\Assets\DreamWorld\ML-Agents\Plugins\Android" **delete** the two files `TensorFlowSharp.dll` and `TensorFlowSharp.Android.dll`
k. **Save** the scene with CTRL-s
l. **Open** the build menu with CTRL-SHIFT-b
m. **Click** "Add Open Scenes"
n. Click build, then **create** a new folder called "`Builds`" in the project folder, then select that folder
o. Verify that the building process is successful and that the build folder pops up
p. Connect the DreamGlass headset and run the executable.

8. Close Unity and make a backup of the Unity project folder.
9. In the Unity Store find and install the "Package Uninstaller" asset
10. In the Unity Store find and install the "Standard Assets" asset

Variations under Linux Ubuntu

When the target environment is Linux it is possible to use the same or another Linux machine to host Unity and build the project. Instructions contained in document '`Setup of Ubuntu development workstation.pdf`' describe how to create such machine. Only difference for the current procedure is at point 7, from item 'l', that is amended as follows:

7. *Follow instructions in the document "`Dreamworld SDK documentation_12_31_18.docx`", summarized here:*
   a. b, c, d, e, f, g, h, I, j, k [...]
   l. **Open** the build menu with CTRL-SHIFT-b
   m. **Select** target platform as 'PC, Mac & Linux standalone', x86_64, all boxes unselected except 'development build' (set to Yes), and compression set to 'default'
   n. **Click** "Add Open Scenes"
   o. Click build, then **create** a new folder called "`Builds`" in the project folder, then select that folder
   p. **Click** 'Save' on the top/right corner of the window
   q. **Confirm** 'Replace' existing files
   r. Verify that the building process is successful and that the build folder pops up
   s. Connect the DreamGlass headset and run the executable.

M Procedure to Setup the AR Development Workstation

DreamGlass AR headset and NOLO CV1 tracking system

Setup of Windows dev. workstation

Version 1.1 - Document originally written by A. Pardi in March 2019 and last updated by A. Pardi on 5 December 2019.

Introduction

This document describes the installation steps followed when setting up the development workstation for DreamGlass on the Dell Precision 5820 Tower computer. The procedure described needs to be followed by competent personnel accustomed to Windows 10 Enterprise and Unity 2018.f3.

General procedure

Creation of local non-administrator users is recommended because the remotized environment in DASTUD (Eduroam) has limitations, including a space limit of 6 GB, the steps are :

1. Creation of all user accounts
2. Modification of the lock timer of the screen
3. Setup of accounts for Microsoft and Unity
4. Download of the SDKs for DreamGlass and NOLO CV1

Detailed procedure

Starting conditions are: computer installation completed by Dell and TU Delft technicians.

1. Power on the computer with id `TUD278713`
2. Change the "`.\localadmin`" user password and try out the login to test the new password
3. Create a new local non-administrator user (e.g. "`.\Pardi_local`" or "`.\Meima_local`") on domain `TUD278713`.
4. Install a new program from internet: Unity 2018.3.4f1 (edition personal, 64-bit)
5. Install a new program from internet: Visual Studio 2017 Community edition (Microsoft account required, 32bit or 64bit versions are equivalent, do not install together with Unity)
6. Install a new program from internet: Irfan Viewer 4.52 32bit
7. Install a new program from internet: VLC player (vlc-3.0.6-win32)
8. Install a new program from internet: Winrar 5.61 (40 days trial)
9. Verify that the program F-secure is updated
10. Checked the system configuration to be Windows 10 Enterprise, ram 16 GB, hard-disk 476 GB, graphics card P4000
11. [If needed create a Microsoft personal account on Microsoft website using the TU Delft mail address as username]
12. [If needed create a Unity personal account on the Unity website using the TU Delft mail address as username]
13. Open the Windows registry editor (Start menu' -> regedit) and find the key `Computer\HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Power\PowerSettings\7516b95f-f776-4464-8c53-06167f40cc99\8EC4B3A5-6868-48c2-BE75-4F3044BE88A7`
14. Change the value in "Attributes" from 1 to 2 and close the registry editor (this step and the previous one to be done only on the Admin account)
15. Open the Windows sleep settings (Start menu' -> Power & sleep settings)
16. Change the timeout to 5 hours (Tab "Power & sleep", Screen turn off timer)
17. Log in Unity with your personal account attached to the TU Delft email address
18. Download the whole SDK folder from DreamWorld website

TU Delft

Delft University of Technology

19. Download the Nolo SDK (please note the confusing filename "`UnitySDK_v1.1.unitypackage.unitypackage`")
20. Follow the document "`DreamGlass setup of Unity project.pdf`" about creating a new Unity project to verify everything is working
21. End of the procedure

N | CATIA to Unity File Conversion

DreamGlass AR headset and NOLO CV1 tracking system

CATIA to Unity parts conversion procedure

Version 1.1 - Document originally written by A. Pardi in May 2019 and last updated by A. Pardi on 20 November 2019.

Introduction

This document describes a working method to bring objects from CATIA to Unity. Once the object is inside Unity more features can be added as material, finish and other light-handling properties as reflectivity, making it closer to its original CATIA counterpart. The CATPart in CATIA is first exported in STL format, then the program MeshLab is used to convert the STL into CAE format, and finally the CAE file is imported in Unity.

Programs needed

- Dassault Systemes CATIA V5.20 (CAD platform)
- ISTI/CNR MeshLab 2016.12 (mesh converter)
- Unity 2018.f3 (game engine)

File formats used

- `.CATPart` (CAD file format)
- `.STL` (stereolithography file format)
- `.CAE` (Collada file format)

Procedure

1. **Open CATIA V5.20**
2. Create a new part with a solid element, for example a 'thick surface'
3. Save the CATProduct file
4. `File -> Save As... -> Save as type 'STL'`
5. Close CATIA V5.20
6. **Open MeshLab 2016**
7. Drag and drop the 'STL' file in the MeshLab window
8. Confirm `Unify Duplicated Vertices` in the dialog box 'Post-Open Processing'
9. Scaling: `Filters -> Normals, Curvature and Orientation -> Transform: Scale, Normalize`
10. Scaling: in all axes use 0.001 (use 0.08 for the HH60 model, it is in 1:80 scale), then `Apply` and `Close`
11. Exporting: `File -> Export Mesh As... -> Files of type 'Collada File Format (*.DAE)'`
12. Confirm `Normal` in the dialog box, if asked
13. Close MeshLab *(ignore the green advancing bar in the lower right corner of the MeshLab window)*
14. **Open Unity 2018.3.4f1 (64-bit)**
15. Drag and drop the 'DAE' file in a folder inside the 'Project' window
16. Drag the object the 'Project' to the 'Hierarchy' window *(under the appropriate master object if applicable)*
17. Save the scene

Observed issues

Scaling. The original CATIA part has dimensions in millimeters, however these are interpreted as meters by MeshLab, so a scaling is required by a 1,000 factor. The scaling steps 9 and 10 solve this issue and are applicable to already scaled models as long as the scale is known as in the case of the HH60 model.

Sub-parts. In order to have different subparts in Unity make solids show/hide in CATIA when exporting to STL. The positioning coordinates will be carried on so in Unity they will be in the right position wrt the other sub-parts.

Thick surface. Exporting CATIA surfaces will make them visible in Unity only from one side. It is recommended to generate solids (thick surfaces of 1mm or less) before exporting in STL.

TU Delft

Delft University of Technology

Bibliography

[1] Anon. *Pepper's Ghost Illusion*. "M. Evans Picture Library", 1882. https://pixels.com/featured/the-peppers-ghost-illusion-1882-mary-evans-picture-library.html.

[2] Anon. *DreamGlass Developer Edition*. 2019. https://www.dreamworldvision.com/product-page/dreamglass-headset.

[3] Anon. *New F-35 Helmets Cost £260,000 Each*. Telegraph Media Group Limited, London, United Kingdom, 2015. https://www.telegraph.co.uk/news/uknews/defence/11963873/New-F-35-helmets-cost-260000-each.html.

[4] C. Kopp. *Lockheed-Martin F-35 Lightning II - Assessing the Joint Strike Fighter*. 2012. http://www.ausairpower.net/APA-JSF-Analysis.html.

[5] J. Vittori. *Applications of Augmented Reality In the Healthcare Industry and Its Impact on Surgical Outcomes*. 2018. https://www.healthworkscollective.com/augmented-reality-in-healthcare-its-impact-on-surgical-outcomes/.

[6] W.H. Welch et al. *Methods And System For Creating Focal Planes In Virtual And Augmented Reality*. Google Patents, 2015. US Patent 2015/0346495.

[7] P. Saarikko et al. *Waveguide*. Google Patents, 2016. US Patent 9,513,480.

[8] M. Krichenbauer et al. *Augmented Reality versus Virtual Reality for 3D Object Manipulation*, volume 24. 2018.

[9] L. Abdi and A. Meddeb. *In-vehicle augmented reality system to provide driving safety information*, volume 21. Springer Berlin Heidelberg, 2018.

[10] C.R. Aragon and M.A. Hearst. *Improving aviation safety with information visualization*. 2005.

[11] D.B. Beringer. An exploratory evaluation of general aviation pilot performance and preferences using synthetic vision displays for approaches and missed approaches in flat and challenging terrain. *Proceedings of the Human Factors and Ergonomics Society*, pages 76–80, 2016.

[12] R.E. Bailey et al. *Head-Worn Displays for NextGen*, volume 8041. 2011.

[13] M. Billinghurst et al. *A survey of augmented reality*, volume 8. 2014.

[14] Anon. *Epson MOVERIO Augmented Reality Smart Glasses*. 2019. https://epson.com/moverio-augmented-reality.

[15] C. Merenda. *Augmented Reality Interface Design Approaches for Goal-directed and Stimulus-driven Driving Tasks*, volume 24. Nov 2018.

[16] J. Grubert et al. *A Survey of Calibration Methods for Optical See-Through Head-Mounted Displays*, volume 24. IEEE, 2018.

[17] R.A. Hess. *Modeling the Effects of Display Quality upon Human Pilot Dynamics and Perceived Vehicle Handling Qualities*, volume 25. 1995.

[18] M. Treisman. *Motion sickness: An evolutionary hypothesis*, volume 197. 1977.

[19] A. Willame. Oculus quest - health & safety warnings. `https://securecdn.oculus.com/sr/oculusquest-warning-english`. Accessed: 2019-10-09.

[20] Anon. *OptiTrack - Motion Capture for Virtual Reality*. 2019. `https://www.optitrack.com/motion-capture-virtual-reality/`.

[21] J. Ashley. *How Hololens Displays Work*. John Wiley and Sons, New York, USA, 2015. `http://www.imaginativeuniversal.com/blog/2015/10/18/how-hololens-displays-work/`.

[22] R. Zabels et al. *AR Displays: Next-Generation Technologies to Solve the Vergence–Accommodation Conflict*, volume 9. 2019.

[23] R.S. Kennedy, L.H. Frank. *A Review Of Motion Sickness With Special Reference To Simulator Sickness*. Naval Training Equipment Center, Orlando, FL, 1985.

[24] D.M. Johnson. *Introduction to and Review of Simulator Sickness Research*. Number April. Arlington, Virginia, USA, 2005.

[25] W. Bles et al. *Experiments On Motion Sickness Aboard The M.V. Zeefakkel*. TNO Human Factors Research Institute, Soesterberg, The Netherlands, Report IZF-1991-A-34, 1991.

[26] B. De Graaf et al. *Exploratory Study With A Sea Sickness Questionnaire On Hr. Ms. Frigate Tromp*. TNO Human Factors Research Institute, Soesterberg, The Netherlands, Report IZF-1992-A-1, 1992.

[27] R.S. Kennedy et al. Simulator Sickness Questionnaire : An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, 3(3):203–220, 1993.

[28] S.A. Balk et al. Simulator Sickness Questionnaire: Twenty Years Later. *Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, pages 257–263, 2013.

[29] K. Brunnström et al. Quality Of Experience For A Virtual Reality Simulator. *IS and T International Symposium on Electronic Imaging Science and Technology*, March 2018.

[30] A. Cerrudo. *Black Hawk Rescue Helicopter with Catia V5*. 2015. `https://grabcad.com/library/black-hawk-rescue-helicopter-with-catia-v5-1`.

[31] M.F. Deering. *The Limits of Human Vision*. University of Leicester, Leicester, UK, 1998.

[32] A.R. Tilley. *The Measure of Man and Woman - Human Factors in Design*. John Wiley and Sons, New York, USA, 2002.

[33] Microsoft Corporation. *HoloLens Hardware Details*. 2017. `https://developer.microsoft.com/en-us/windows/mixed-reality/hololens_hardware_details`.