ALMA MATER STUDIORUM – UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN:
MECCANICA E SCIENZE AVANZATE DELL'INGEGNERIA
CICLO XXIX

SETTORE CONCORSUALE DI AFFERENZA: 09\A3

SETTORE SCIENTIFICO DISCIPLINARE: ING-IND\15

DESIGN AND DEVELOPMENT OF A RESEARCH FRAMEWORK FOR PROTOTYPING CONTROL TOWER AUGMENTED REALITY TOOLS

PRESENTATA DA:
NICOLA MASOTTI

COORDINATORE DOTTORATO	RELATRICE
PROF. NICOLÒ CAVINA	PROF.SSA FRANCESCA DE CRESCENZIO

ESAME FINALE ANNO 2017

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ACRONYM	DEFINITION
ADM	Arrival and Departure Monitor
AFIS	Aerodrome Flight Information Service
AH	Abstraction Hierarchy
AHP	Analytic Hierarchy Process
AIRMET	Airmen's Meteorological Information
AIXM	Aeronautical Information eXchange Model
ALT	Actual Landing Time
AMEL	Active Matrix Electroluminescent
AOIS	Aeronautical Operational Information System
API	Application Programming Interface
APP	APProach
AR	Augmented Reality
ARTT	Augmented Reality Tower Tools
ASR	Airport Surveillance Radar
ATC	Air Traffic Control
ATCO	Air Traffic COntroller

ATCR	Air Traffic Control Radar
ATCRBS	Air Traffic Control Radar Beacon System
ATIS	Automatic Terminal Information Service
ATM	Air Traffic Management
ATOT	Actual Take Off Time
ATZ	Aerodrome Traffic Zone
BGE	Blender Game Engine
CAT	Category
CAVE	Cave Automatic Virtual Environment
CDM	Collaborative Decision Making
CFMU	Central Flow Management Unit
CFR	Crash Fire Response
CG	Computer Graphics
CPU	Central Processing Unit
CRT	Cathode Ray Tube
CTOT	Calculated Take Off Time
CVS	Combined Vision System
CWP	Controller Working Position
DEL	Delivery

DEM	Digital Elevation Map
DLP	Digital Light Processing
Acronym	Definition
DME	Distance Measuring Equipment
DTD	Distance to Touch-Down
DTED	Digital Terrain Elevation Data
ECW	Enhanced Compression Wavelet
EFVS	Enhanced Flight Vision Systems
EID	Ecological Interface Design
ENAV	Ente Nazionale per Assistenza al Volo
EOBT	Estimated off Blocks Time
ETOT	Estimated Take Off Time
EVS	Enhanced Vision System
FAA	Federal Aviation Administration
FDP	Flight Data Processing
FIS	Flight Information Service
FIXM	Flight Information eXchange Model
FLIR	Forward-Looking InfraRed
FOV	Field Of View

FPS	Flight Plan System
FS	Flight Strip
GCS	Ground Control Station
GGV	Gaze, Gesture and Voice
GIS	Geographic Information System
GML	Geography Markup Language
GND	Ground
GPS	Global Positioning System
GPU	Graphics Processing Unit
HCI	Human-Computer Interaction
HD	High Definition
HF	Human Factors
HMD	Head Mounted Displays
HMS	Helmet-Mounted Sight
HPU	Holographic Processing Unit
HUD	Head Up Display
IAIP	Integrated Aeronautical Information Package
ICAO	International Civil Aviation Organization
IDE	Integrated Development Environment

IEEE	Institute of Electrical and Electronics Engineers
IFR	Instrument Flight Rules
IHP	Intermediate Holding Point
ILS	Instrument Landing System
IMU	Inertial Measurement Unit
IPD	Interpapillary Distance
IRAB	Innovative Research Advisory Board
ISO	International Organization for Standardization
IWXXM	ICAO Weather Information Exchange Model
LCD	Liquid Crystal Display
LCOS	Liquid Crystal Display
LIDAR	Light Detection and Ranging
LOC	Localizer
LOD	Level Of Detail
LVC	Low Visibility Conditions
LVP	Low Visibility Procedures
MET	Meteorological
METAR	METeorological Air Report
MID	Middle

MLAT	MultilLATeration
MMR	Multi-Mode Receiver
NASA	National Aeronautics and Space Administration
NDB	Non-Directional Beacons
NMOC	Network Manager Operations Centre
NUI	Natural User Interface
NVC	Normal Visibility Condition
OFA	Operational Focus Area
OGC	Open Geospatial Consortium
OLED	Organic Light Emitting Diode
PPI	Plan Position Indicators
RETINA	Resilient Synthetic Vision for Advanced Control Tower Air Navigation Service Provision
PSR	Primary Surveillance Radar
PVD	Planar View Display
QFD	Quality Function Deployment
QFE	Query: Field Elevation
QNH	Query: Nautical Height
RGB	Red, Green and Blue
RHP	Runway Holding Point

RPAS	Remotely Piloted Aerial Systems
RSD	Retinal Scanning Display
RT	Remote Tower
RVE	Reconfigurable Virtual Environment
RVR	Runway Visual Range
RWY	Runway
SA	Situation Awareness
SAB	Società Aeroporto di Bologna
SESAR	Single European Sky ATM Research
SID	Standard Instrument Departure
SIGMET	SIGnificant METeorologic information
A-SMGCS	Advanced - Surface Movement Guidance and Control System
SMR	Surface Movement Radar
SPECI	Special Weather Report
SRK	Skills, Rules, Knowledge
SSR	Secondary Surveillance Radar
SV	Synthetic Vision
SVFR	Special Visual Flight Rules
SVS	Synthetic Vision System

SWIM	System Wide Information Management
TAF	Terminal Aerodrome Forecast
TDZ	Touch-Down Zone
TIS	Traffic Information Service
TOBT	Target Off-Block Time
TRACON	Terminal Radar Approach Control
TRL	Technology Readiness Level
TTOT	Target Take Off Time
TWR	Tower
TWY	Taxiway
UCD	User-Centred Design
UI	User Interface
VAC	Vergence-Accommodation Conflict
VCS	Visually Coupled System
VCVS	Verified Combined Vision Systems
VFR	Visual Flight Rules
VHF	Very High Frequency
VOR	VHF Omni Directional Range
VR	Virtual Reality

VRH	Virtual Reality Headset
WCAT	Wake turbulence CATegory
WCS	Web Coverage Service
WDA	Work Domain Analysis
WFS	Web Feature Service
WMO	World Meteorological Organization
WMS	Web Map Service
WMTS	Web Map Tile Service
WXXM	Weather Information Exchange Model
XML	eXtensible Markup Language

Abstract

The purpose of the air traffic management system is to ensure the safe and efficient flow of air traffic. However, the primary goals of safety and efficiency are to some extent conflicting. In fact, to deliver a greater level of safety, separation between aircrafts would have to be greater than it currently is, but this would negatively impact efficiency and vice versa. Therefore, while augmenting efficiency, throughput and capacity in airport operations, attention has rightly been placed on doing it in a safe manner.

In the control tower, many advances in operational safety have come in the form of human-machine interfaces and visualization tools for tower controllers. Advanced Surface Movement Guidance & Control System solutions, such as movement maps, conformance monitoring and conflict detection are a few examples of these tools. But there is a paradox in developing such systems to increase controllers' situational awareness: by creating additional computer displays that show the runway and taxiway layout, the aircrafts and ground vehicles position, and detect actual and foreseen conflicts, the controller's vision is pulled away from the outside view and the time spent looking down at the monitors is increased. This reduces their situational awareness by forcing them to mentally and physically switch between the head-down equipment and the outside view [1].

This research is based on the idea that new developments in the realm of augmented reality may be able to address this issue. Augmented reality differs from virtual reality insofar as it allows users to view the real world along with superimposed, computer-generated information. This concept has become increasingly popular over the past decade and is being proficiently used in many fields, such as entertainment, cultural heritage, aviation, military & defence. In the cockpit, a wide set of virtual and augmented reality systems has been developed and tested to equally operate under visual meteorological conditions and instrument meteorological conditions (e.g., head-up displays, helmet-mounted displays and enhanced/synthetic/combined vision systems). These could be possibly transferred to air traffic control with a relatively low effort and substantial benefits for controllers' situation awareness. Hence, this study focuses on augmented reality tools that support controllers in zero/low visibility conditions and complex traffic situations.

Research on this topic is strongly supported by the Single European Sky Air Traffic Management Research (SESAR), which is the European framework for the development of the future air traffic management system [2]. This is consistent with the objectives of increasing air traffic controllers' situation awareness and enable up to 10 % of additional flights at congested airports while still increasing safety and efficiency [3].

During the Ph.D., a research framework for prototyping augmented reality tools was set up. This framework consists of methodological tools for designing the augmented reality overlays, as well as of hardware and software equipment to test them. Several overlays have been designed and implemented in a simulated tower environment, which is a virtual reconstruction of Bologna airport control tower. The positive impact of such tools on safety and capacity was preliminary assessed by means of the proposed methodology.

1 THESIS OVERVIEW

1.1 MOTIVATION AND PROBLEM STATEMENT

Since the beginning of commercial aviation, the global air traffic rate has exhibited a positive trend, even though economic stagnation, financial crisis and increased security concerns. According to a prevalent opinion, this trend is unlikely to change in the future, although several factors, such as politics, economy, environment, safety and security may affect its actual rate. As a result, the air traffic growth tends to be accepted as a certainty within the industry, especially from a global, long-term perspective [4].

In future scenarios airports are considered as one of the major bottlenecks to increase the capacity of the Air Traffic Management (ATM) system. While augmenting throughput, attention has rightly been placed on doing it in a safe manner. In fact, many of the technological advancements designed to improve airports capacity and safety have come in the form of innovative visualization tools for tower controllers. Advanced - Surface Movement Guidance and Control System (A-SMGCS) based solutions, such as movement maps, conformance monitoring and conflict detection are a few examples of these tools. However, there is a paradox in developing these tools to increase the control-tower air traffic controller's situational awareness. By creating additional computer displays that show the runway and taxiway layout, aircrafts and vehicles position, and detect actual and foreseen conflicts, the controller's vision is pulled away from the out of the window view and his or her 'head-down' time is increased[1]. This reduces their situational awareness by forcing them mentally to repeatedly switch between these two ways of interpreting the working environment. Past studies have already proven that tasks requiring frequent shifts of gaze back and forth between the outside and the inside view may become significantly slow and fatiguing, particularly after the fortieth years of age [5]. In other words, a constant refocusing between the far view and the head-down equipment contributes to the operator's workload and reduces his or her SA. Moreover, the importance of the outside view for controllers' Situation Awareness (SA) has been repeatedly proven [1], [6]–[8].

---

[1] The 'head-down' time is the time spent by the air traffic controller looking at his/her desk equipment or managing flight strips.

Within the ATZ, depending upon weather and lighting conditions, the visual contrast of controlled objects varies substantially, with possible detrimental impact on controllers' performances [9]. When bad weather, fog, smoke, dust or any other kind of environmental occlusion impairs the visibility from the control tower, the airport capacity is reduced and Low Visibility Procedures (LVP) must be applied. In addition, it is also possible for the airport, the surrounding airspace, and the controlled vehicles to be obscured by buildings, high-glare conditions and the cover of night [9]. LVP may include constraints, such as mandatory use of a Surface Movement Radar (SMR), taxiways that cannot be used, block spacing, limitation in pushback operations and use of a predefined runway. Inevitably, if the operational capability is reduced, both carriers and Air Navigation Service Providers (ANSPs) incur in heavy financial losses.

1.2 PROPOSAL

In [10], Shackelf and Karpe refer to large fuel savings and financial benefits if stable rates of airport capacity could be maintained in all visibility and traffic conditions. This also implies a higher arrival and departure rates and a more uniform and productive Air Traffic Flow Management (ATFM). Further, the increased reliability of the surface management service would improve metrics for taxi-times, departure queues, ground-delays, ground-holds and cancellations [10] and, of course, benefit safety.

One way of achieving this is to look at virtual and augmented reality technologies that have been used in both civil and military aviation for years, both for piloting or training purposes (e.g., flight simulators, control tower simulators, HUD, ST-HMDs, etc.). Particularly, the integration of HUD-technologies into modern civil flight decks has demonstrated a lot of advantages. These could be possibly transferred to ATC with relatively low effort and substantial benefits for control tower operations.

Using Virtual/Augmented Reality (V/AR) tools controllers might be able to reduce constraints in airport operations, particularly in Low Visibility Conditions (LVC), when the perspective view of the airport surroundings is (partially) lost. Also, the information that was displayed on the head-down computer screens, such as flight tags, moving aircrafts, MET data, collision warnings, and all sorts of A-SMGCS safety nets could be displayed as superimposed to the outside of the windows view, reducing the head down time and effectively increasing controller situation awareness. Past studies have demonstrated that the interest in doing so is high and some prototypes have been developed with by now out-dated hardware [6], [11]–[15].

1.3 EXPECTED IMPACT

In 2014, within the European Civil Aviation Conference Area (ECACA), an average delay per flight of 9.7 minutes was developed [16]. Further analysis of the rationale behind the delay show that 0.51 min were due to weather, mainly strong wind, snow and LVC, whilst 0.96 min were due to restrictions at the departing or arrival airport, including the ones introduced by LVP. Also, this data does not account for cancelled or redirected flights.

Using AR in the airport tower means that controllers will be no longer be limited by what the human eye can see out of the tower's windows. This is like using the radar without losing the perspective view of the airport.

When relying on visual augmentations, constraints introduced by LVP could be reduced. For instance, an exclusive use of taxiway blocks (a.k.a. block spacing) may not be necessary. Therefore, an aircraft could use a segment of a taxiway before the preceding aircrafts has left such segment. In other words, those tasks that can be negatively affected by poor visibility conditions will become weather-independent, facilitating the maintenance of operational capacity in all weather conditions. This will allow Instrument Landing System (ILS) or SV equipped aircrafts to seamlessly operate under any visibility condition at synthetic vision equipped airports.

AR overlays can also aid users by substantially reducing the amount of visual scanning needed to integrate various sources of information. This contrasts with the current practice of scanning multiple devices (screens, windows, flight strips, etc.), filtering the essential information from data that may not be relevant. An augmented reality HUD would put only the relevant information right in front of controllers. As a result, the head-down time should decrease and controllers' situation awareness should increase.

Considering the recent advancements in camera-based surveillance systems the proposed concept could contribute to establish a satisfactory level of safety at smaller airports, where the traffic volume is simply too low to pay back for the initial investment in a SMR equipment. For further data collection, the System Wide Information Management (SWIM) network can be exploited. SWIM consists of standards, infrastructures and governance enabling the management and exchange of ATM information between qualified parties via interoperable services. It covers all ATM related information, including flight plans, MET data, air traffic flow management and surveillance information [17], [18].

Overall, significant benefits are expected for the entire air traffic system, including (a) increased safety for passengers, (b) financial savings for carriers and ANSPs, (c) environmental pollution reduction, and (d) reduced risk of creating bottlenecks in the traffic flow management system [6], [9], [10], [14], [19].

Side effects, such as the increase of traffic volume at smaller or peripheral airports due to a better level of service must not be neglected. Passengers and couriers could use smaller aircrafts on a more frequent basis, with a positive social impact on tourism and the community living in the airport surrounding.

Finally, the development of AR tools will provide a technology bridge between the current tower systems and the 21st Century 'Remote & Virtual Tower' (R&VT) concept foreseen in both the SESAR and NGATS (Next Generation Air Transportation System) visions. Over the last few years, several concepts for the provision of air traffic service from a distant/remote location have been proposed, including video-surveillance based systems (remote towers), and VR facilities in which a photo-realistic real-time rendering recreates a 360-degree tower view (virtual towers) [8], [9], [20], [21]. The first concept is far advanced in SESAR and has been proven ready for industrialization (leading to operational deployment). As for the second, this may take decades to refine. Nevertheless, there are strong financial reasons to develop this technology.

One of the open issues associated with virtual towers is the assessment of the extent to which the 'digital world' can be trusted to resemble the referenced real world. In this sense, AR may become of critical importance for the R&VT research. If an augmented reality tool became certified and operational in the next several decades, it is expected that the community of tower controllers would generate discrepancy reports each time there is a mismatch between the real world that they observe and the virtual world that is presented via the AR Tower Tool [6]. Conversely, the inability of controllers to detect such discrepancies would become valuable data for the validation, verification and certification of R&VTs [6]. In this sense, AR towers will provide a suitable development path for designing the fully immersive virtual tower of the future [9].

Figure 1. Areas of interest impacted by this research

1.4 OBJECTIVES

In compliance with the idea of using AR for control tower on-the-site operations, this Ph.D. focuses on the design and development of a research framework for prototyping the augmented reality tools. To achieve this goal, four main objectives were defined:

1. The definition of a theoretical framework for designing and evaluating augmented reality tools and overlays.

2. The development of a software/hardware simulation environment for prototyping and testing such tools (a.k.a. 4D airport interactive model)

3. The design and early implementation of a selected sample of AR overlays.

4. The Integration of such overlays with the simulation environment and laboratory equipment.

The simulation environment primarily consists of a CAVE (Cave Automatic Virtual Environment) system but with slight modifications the developed components should be flexible enough to handle other kind of V/AR systems, such as Table Tops or Head Mounted Display.

1.5 INTENDED AUDIENCE

This research was developed primarily for aeronautical stakeholders such as ANSP, regulators, information technologies systems providers and research centres. Other potential users could include entities or projects that are interested in AR systems.

Hopefully by advancing the maturity level of these tools this research will contribute to consolidate the leading role of European companies (ANSP and industries) into the field of air navigation.

2 INTRODUCTION TO AIR TRAFFIC CONTROL

2.1 OVERVIEW

ATC is basically an exercise in flow control where each controller is responsible for a certain portion of airspace [22]. Airspace volumes can be classified into Aerodrome Traffic Zones (ATZs), Terminal Manoeuvring Areas (TMAs) – a.k.a. Terminal Control Areas (TCAs) – and Control Areas (CTAs). CTAs are further subdivided into airspace sectors. Aircrafts may enter the Air Traffic Controller (ATCO) area of responsibility at various points in space and time. Depending on the flight phase, they must be guided through the sector, toward take off or up to landing (parking included). This must occur in an orderly and efficient manner, avoiding the risk of collision. Safety is enforced by agreed standards of separation, specified in terms of minimum permitted distances between aircrafts (both vertically and laterally) [22] and altitude from the ground.

En-route air traffic controllers work in facilities called Area Control Centres (ACC) and control aircrafts from the time they leave an ATZ or a TMA to the time they enter another ATZ or TMA. When managing en-route traffic, controllers work in teams of two: executive and planner. The executive controller (a.k.a. radar controller) is the one that talks to airplanes, i.e. issues instructions (a.k.a. clearances) to pilots, so that they meet altitude and heading restrictions by specific points. The planner controller (a.k.a. coordinator) supports the executive controller by planning time and coordinating with other ATC units. This is done to keep potential conflicts at a minimum. As an aircraft reaches the boundary of a CTA it is 'handed off' or 'handed over' through to the next CTA's ACC. This process either involves transfer of identification and flight details between controllers or can be 'silent' (depending on local agreements). However, for a 'silent' hand over to be performed, the traffic must be handed over in an agreed manner. When the transfer is completed, the pilot is given a frequency change and begins talking to the next ACC. This process repeats until the aircraft is handed off to a TMA control centre (a.k.a. approach control). If a TMA does not exist, the ACC co-ordinates directly with the control tower.

TMAs can be managed by single or multiple ATCOs. However, in the latter case, they do not work in teams. Each of them is responsible for a specific flight phase and manages aircrafts at different flight levels. For instance, a 'feeder' controller is often in charge of lining up and clearing aircrafts for the final ILS approach. As aircrafts move in and out of the TMA, they are handed off to the next control facility, such as a control tower or ACC.

Inside the ATZ, the responsibility of tower controllers typically falls into three main categories: Tower Control, Ground Control and Flight Data/Clearance Delivery. Ground control (a.k.a. Ground Movement Control) is responsible for all the operations taking place on the airport movement area, which comprises by the apron and the manoeuvring area. The manoeuvring area, in turn, comprises taxiways, inactive runways, holding areas and intersections. Tower Control is responsible for the active runways and clears aircraft for take-off and landing. Both Ground Control and Tower Control are also expected to alert airport emergency services in case an aircraft is experiencing difficulties. Clearance Delivery, which is often combined with Flight Data Delivery in controlled airports, issues route clearances and provides pre-flight information to pilots. At busy airports, Clearance Delivery may also plan and issue aircraft push-backs and engine starts. This helps to prevent taxiway and apron gridlock. In this case, Clearance Delivery it is better referred as the Ground Movement Planner (GMP) service. Flight Data Delivery provides pilots with the latest information about weather, traffic, outages, delays, ground stops and other airport restrictions. At busier airports, the Flight Data service may be provided to pilots using a continuous broadcast of a recorded loop message on a specific frequency, which is known as the Automatic Terminal Information Service (ATIS).

Alternate ATCOs' activities include supervisory and redistribution of traffic flows, airspace sectorization, holding stack management and provision of Flight Information Service (FIS).

To perform all ATC tasks, controllers need to extract information from the Planar View Display (PVD), check weather, consult Flight Strips (FS), elaborate long term strategies, detect potential conflicts, radio-communicate with pilots, make tactical decisions, coordinate with each-other and look out of the tower window (if any). Also, controllers need to balance cognitive resources and carefully timetable actions [23].

2.2 THE RADAR

Before the end of the ‘90s, the introduction of Graphical User Interfaces (GUIs) made computing technologies accessible to many professional and amateur users, in addition to computer scientists and programmers. This made Human Computer Interaction (HCI) a subject of a more general interest. Ever since then, a great effort has been put in trying to fill the gap between the user and what is going on in the hidden and intangible parts of computers [24].

In the 90’s, when computers with graphical user interface finally became accessible to many professional and amateur users, the early days when the management of a few planes could be left to little more than the pilots’ eyesight and radio communications had been long passed. Controllers had already moved from ‘procedural’ to ‘radar-based’ ATC, while the oscilloscope-based Radar Bright Display Equipment (RBDE) – a.k.a. Bright Radar Indicator Terminal Equipment (BRITE) – was in the process of being replaced by the raster-scan PVD. Prior to that, Plan Position Indicators (PPIs) were used, especially in military operations rooms. However, they were later substituted by RBDE because of their limited brightness.

The transition from the RDBE to the PVD technology heralded the beginning of the data processing era. Indeed, with oscilloscopes-based interfaces, such as PPI or RBDE, the image was not digitally stored, but only displayed on the screen, i.e. fading away as a function of the cathode ray tube persistence. The screen was updated synchronously with the radar sweep, allowing the controller to see echoes, but only for a few seconds before the image would fade out completely. The radar would then make another sweep and refresh the image. On the contrary, by the time that the PVD technology was mature, the radar data was digitally stored and could be used to generate fully persistent images. That is also when User Interfaces (UIs) started to be populated by data-blocks, labels and DSTs. In a PVD aircrafts are represented as dots moving through the radar screen. For each aircraft, selected information is displayed on the screen by means of symbols, data-blocks and labels. This typically includes the aircraft’s call sign, type, its altitude and its speed. Further information is available upon request through the UI (e.g. flight plan, historical track, forecasted position, etc.). The radar screen also presents information on the airspace itself, such as sectors boundaries, routes, navigational aids, waypoints, fixes, minimum vectoring altitudes and prohibited airspace volumes. As a matter of fact, although the display format is bi-dimensional, a large amount of three-dimensional information is embedded into the interface.

2.3 FLIGHT STRIPS

One of the key aspects of an air traffic controller's job is to make Flight Strips useful within the flow of work [25]. Historically, a Flight Strips a piece of paper about one inch wide and eight inches long that is formatted into 'boxes' and provides information on a single flight. In some cases, old-fashioned strips have been replaced by digital strips (a.k.a. electronic-flight strips or e-strips), which make use of touch screens instead of papers. The information printed on the strip is derived from the Flight Data Processing System (FDPS), which is subject to updates, but not continuously. Thus, the Flight Strip must be considered as a discrete image of the flight progress not a continuous one [25].

'Pending' strips are submitted to ATCOs several minutes in advance and become 'alive' on the receipt of a radio message from the plane entering the controller's area of responsibility. Both are placed in racks in front of the controller. Through a process referred as '*working the strips'* or '*making the strips come to be at hand*', controllers order the strips in such a way that they reflect the work that needs to be done. For example, based on the Estimated Time of Arrival (ETA), controllers order the strips so that the next plane expected at any point is at the top of the rack. In this way, future activities are scheduled and controllers get a sense of what decisions they will have to make in a few minutes [25]. As a matter of fact, each strip becomes a piece of a larger puzzle. Also, ordering the strips contributes to shape controllers' attention in terms of what is likely to happen under their responsibility.

If any problem or situation is spotted, controllers mark out the singularity by slightly lifting the corresponding strips out of the rack, to draw attention onto them. This is also known as '*cooking the strip*'.

Another good practice is to note down on the strip all the information related to the aircraft management, including clearances, ETA, coordination, routes and call sign changes. Attention-drawing symbols and convenient signs, such as arrows, crosses and circles may be also jot down on the strip to remark uncommon routes, highlight crossovers, emphasize destinations or denote actions about to be taken. ATC centres may follow a precise colour-coded protocol, which shows by whom the note has been written (chief, executive or planner). For instance, chiefs usually write coordination agreements on pending strips, whereas planners update ETAs. In this sense, a Flight Strip not only keeps track of the decisions that have been taken, but also indicates by whom those choices have been made.

When an aircraft reaches the final navigation point of the current sector (or is finally parked on the apron) the controller puts a cross through the strip to demonstrate that his work has been properly done and that the strip has not just been thrown away.

Exploiting the Verbal Protocol technique [26], [27] several studies have found FSs to be an essential part of controllers working practice [25]. Further analysis suggest that controllers rely on the strips when trying to obtain a general sense of the traffic situation (e.g. when taking over a working position during the shift change) [25]. As a controller once reported "*it would be an impossible job to sit down and look at the radar, and look at all the different blips, and try to avoid them by putting the aircraft into blank spaces on the radar; so you have got to have this information to tell you what traffic is coming into and out of the sector. From your strips you can find out whether there is or not a possible confliction...and what you can do about it, then you go to your radar and look for that particular aircraft*" [25]. These words demonstrate the use of FSs as a primary resource for the creation and maintenance of a mental model that can be used to shape controllers' attention and organise their activity. Once again, the relation between the working memory and the FSs has been properly remarked by a controller himself when he said: "*the strips are like your memory, everything is there*" [25].

To fully investigate the relationship between FSs and controller's mental model (i.e. 'the picture') is beyond the scope of this document and would be hard to do in absence of further research. However, this aspect should not be neglected in future developments about human-computer interfaces for air traffic controllers.

2.4 THE PICTURE

In ATC, '*to build the picture*' is not a detached expression, but one well understood within controllers' culture. For instance, this phrase is often used to depict the regular habit of incoming controllers to spend anything up to ten minutes watching over the shoulder of their colleagues before taking over the position [25].

In [28], Jeannot, Kelly and Thompson report that both theoretical and empirical studies on the *picture* have been carried out since the late 60s, particularly in France [29], [30]. A synthesis can be found in [31]. They also tried to reshape the Situational Awareness (SA) definition so that it would embrace the concept of the *picture,* resulting in a better fit for the ATC domain. During their study, a controller gave the oddest definition, which, unexpectedly, is also one of the most informing: "*SA is what you need to know not to be surprised*" – he said. In [22] Brown and Slater define the picture as the "*overall awareness* *which enables controllers to carry out their tasks and stay ahead of the game*". Whitfield and Jackson formulated a similar statement, saying that the *picture* is the "*overall appreciation of the traffic situation for which they [controllers] are responsible*" [30]. Further research can be found in [32]–[34]. Overall, the *picture* has been described as the holy grail of the controller, the awareness that he seeks and fears to lose.

Recent developments indicate that even when receiving identical information, each controller shapes his or her own SA. Besides, during interviews with both operative controllers and trainees not everyone reported having experienced the '*picture*' as a vivid mental image [35]. Further, the ones who positively reported about its existence had a hard time in describing it verbally [35]. Between the ones negating its existence, a Swedish trainee stated: *"No, I don't have it [the 3D picture]...at least not me", but later unfolds, in his own words, "I think I work more with blocks of airspace"* [35]. The block (a three-dimensional shape) is a clear reference to the 3D nature of the ATC problem, which involves the simultaneous movements of aircrafts along three spatial axes. Eventually, the problem becomes 4D, if also considering time. In this sense, the trainee checks 'the block' trying to foresee whether a certain airspace volume can be safely used or not. In [35], a training specialist reported that '*3D thinking*' is a peculiar characteristic of every student, specifically tested during the initial selection of the candidates, but not explicitly addressed later in the course. Thus, each trainee seems to be left to work out his or her own way to '*think in 3D*'.

Many ATCOs organise the *picture* in terms of flight levels, foreground and background traffic or inbound and outbound flows [25]. Also, they take advantage of their knowledge of typical routes and procedures and sometimes focus on non-routine flights [25]. Some evidence suggests that Flight Strip play a key ro1e in building and maintaining *the picture* [25], [30]. Other indicates that the mental model is mainly built on top of the PVD image [36], [37].

Overall, the subjective nature of the picture has been found strong and confusing. We would like to endorse Tavanti's definition: "[3D picture is] *a mindful understanding of the spatial-temporal relationships between aircrafts and airspace, referring to the comprehension of both current and potential (i.e. anticipated) spatial configurations"* [35].

2.5 FURTHER WORKING PRACTICE

There are many other aspects of controllers' working practice that are probably worth to be mentioned. However, it is beyond the scope of this document to cover all of them in depth. Below is a summary:

Controllers tend to consider aircrafts in pairs rather than in isolation. In this sense, the information becomes relative (e.g. 'this aircraft is at a higher level than the other one', and not 'this aircraft is at flight level *x*' etc.).

To make decisions, controllers only consider the information they need (e.g. only position and altitude).

Controllers operate in predictive mode. This was confirmed by observing that when they incorrectly report an aircrafts' position (or altitude), most of the time, they are just forecasting the aircraft behaviour.

Functional distortions have been found in both the airspace and the radar map (mental) representations, which seems to be related with the traffic load on those elements [28].

3 INTRODUCTION TO VIRTUAL AND AUGMENTED REALITY

In the following paragraphs, a brief introduction to virtual and augmented reality will be given starting from an historical perspective. The review will continue addressing some basic concepts related to human vision. Finally, several devices will be discussed, together with many techniques to render V/AR contents.

3.1 OVERVIEW AND HISTORY

VR refers to a synthetic environment where one or more sensory systems are engaged. Sight, Hearing and Touch are typically the most commonly reproduced senses. The roots of VR depend upon how much the participatory and immersive nature of the environment is valued. One could go back to the 360o panoramic paintings from the 19th century as a first attempt to immerse the viewer in an historical event. In 1930's, thanks to the View-Master commercialization stereoscopic viewers became popular. These devices allowed the viewer to see stereoscopic 3D images and gave a sense of depth perception and immersion. In 1929, Edward Link created the first commercial flight simulator. While it didn't have any visual representations of the outside environment, it did incorporate flight systems, and sensory input in the form of aircraft motion and was the grandfather of motion based aircraft and spacecraft flight simulators. The first appearance of something like today's VR headset was in the 1930 story Pygmalion's Spectacles, by Stanley G. Weinbaum in which he describes the idea of a pair of goggles that let the wearer experience a fictional world through sight, smell, taste and touch. The first time this vision was brought to reality was in 1960. Morton Heilig invented the Telesphere Mask, which, although not having any motion tracking or interactive capabilities, provided wide screen stereoscopic 3D imagery and stereo sound. The first motion tracking headset was not far behind. In 1961, the Philco Corporation developed the precursor to the Head Mounted Display. It incorporated a video screen for each eye and a magnetic motion tracking system, which was linked to a closed-circuit camera. Developed to allow for immersive remote viewing of dangerous situations by the military, head movements would move a remote camera, allowing the user to naturally look around the environment. The first true VR HMD, was developed in 1968 by Ivan Sutherland, and was called The Sword of Damocles due to its being suspended from the ceiling because of its weight. The computer-generated graphics that were shown were primitive wireframes. It wasn't until 1987 when John Lanier began to popularize the term "virtual reality" to describe the research area as we know it today. His company VPL was the first to sell commercial VR goggles. Various improvements on these types of headsets have been made since then, culminating today in products such as the Oculus Rift and the HTC Vive, which provide a realistic, computer generated, 3D immersive visual environment. Stereoscopic 3D projection has existed since 1915, but has become popular in various segments of the entertainment business only recently thanks to more accessible hardware and standardization. 3D displays have existed since the time of the Cathode Ray Tube (CRT) technology, but 3D monitors and TVs started to massively commercialize only in 2010.

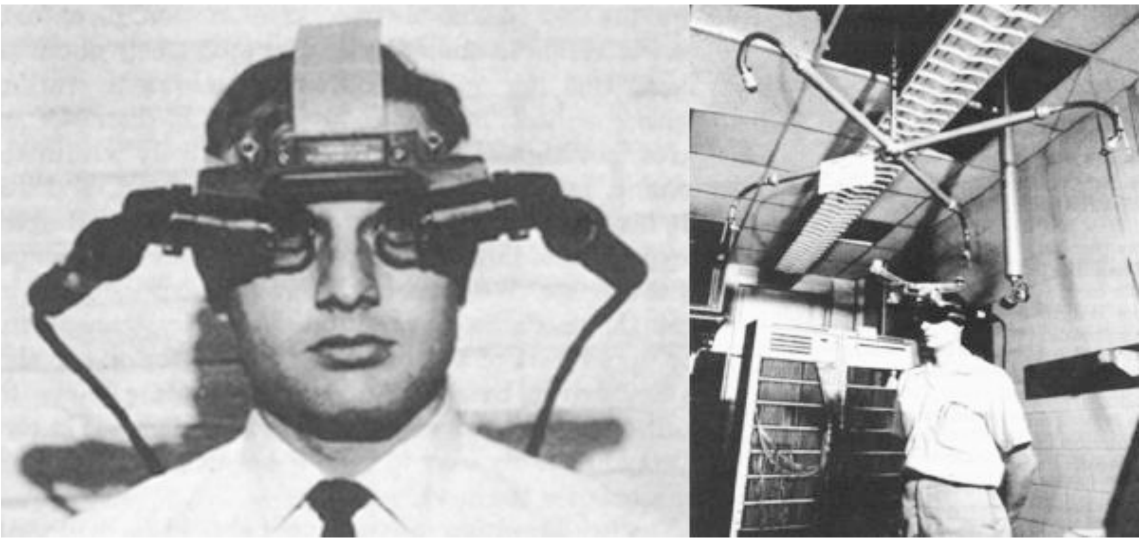


Figure 2. SWORD OF DAMOCLES

AR differs from VR insofar as it allows users to view the ‘real’ world along with superimposed or computer-generated information. It has a similar origin to VR, however, the two begin to diverge in 1975, when Myron Krueger created the Videoplace to allow users to interact with virtual objects using their own shape. In 1980 Steve Mann created the first wearable computer. A computer vision system with text and graphical overlays on a photographically mediated reality. In 1990 the term 'Augmented Reality' is attributed to Thomas P. Caudell, a former Boeing researcher. In 1992 Louis Rosenberg develops one of the first functioning AR systems, called Virtual Fixtures, at the U.S. Air Force Research Laboratory. In 1992 Steven Feiner, Blair MacIntyre and Doree Seligmann present the first major paper on an AR system prototype, KARMA, at the SIGGRAPH conference. In 1999 The US Naval Research Laboratory engage on a decade long research program called the Battlefield Augmented Reality System (BARS) to prototype some of the early wearable systems for dismounted soldier operating in urban environment. The very same year, Hirokazu Kato created ARToolKit, an open-source computer tracking library for the overlay of virtual images. In 2005 the Laster Technologies company develops commercial augmented reality eyewear. In 2006 Ronald Reisman and David Brown, from NASA Ames publish their findings from investigation of an augmented reality prototype for use by airport tower controllers. In 2013 the company Meta announced the Meta 1 developer kit, the first to market augmented reality see-through display that allows multiple users to see and “touch” 3D objects in physical space. The very same year, Google announces an open beta test of its Google Glass augmented reality glasses. In 2015 Microsoft^{TM} announced the HoloLens^{TM} augmented reality headset which utilises various sensors and a processing unit to blend high definition "holograms" with the real world.

3.2 FUNDAMENTALS

3.2.1 REGISTRATION

In the field of AR the concept of spatially matching the real and the virtual objects according to the user perspective is known as registration. Alternate designations include ‘object alignment’, ‘object connectivity’, ‘conformal’ or ‘scene-linked’ symbology and ecological validity of the environment [38]–[40]. Registration is particularly important in panoramic environments, where the augmented reality content should be placed (i.e. perceived) on top of real objects.

In the control tower, augmented reality overlays such as bounding volumes, flight tags and airport layouts should follow this rule. Therefore, several depth cues must be provided to the end user by the AR system so that the perceived depth, shape, dimension and orientation of a virtual object matches that of a real object.

3.2.2 DEPTH CUES

Depth cues are used by the human brain to reconstruct the three dimensions of the space surrounding the viewer and are frequently classified in two categories, i.e. monocular cues and binocular cues. Monocular cues (a.k.a. pictorial cues), are the ones that can be retrieved form a scene by means of a single eye. They are widely used in painting, photography and computer graphics and provide the viewer with a sense of depth and three-dimensionality, to the extent that the content ‘looks like 3D’ even if displayed on a 2D media. The following are the most important monocular cues:

- **Linear perspective**: this is the kind of perspective that projects the world on the human's eye retina, in which parallel lines converge in the distance.

- **Relative size**: large objects are perceived as closer than small ones.

- **Relative height to the horizon** (a.k.a. elevation): objects closer to the horizon are perceived as farther away from the viewer.

- **Lighting and shading**: the way that light falls on objects and reflects off their surfaces, and the shadows that are cast by the same objects provide an effective cue for the brain to determine the shape of objects and their position in space.

- **Occlusion** (a.k.a. interposition): this cue derives from the partial overlap of two objects viewed from a certain perspective. The occluding object appears to be closer than the one that is partially blocked.

- **Texture gradient**: a surface texture gets finer and smoother as it distances the observer.

- **Atmosphere**: the blurrier an object is, the more is perceived as far from the observer.

- **Motion parallax**: far objects seem to move less than nearby objects when the viewer changes his or her viewpoint.

- **Depth from motion**: an object that changes its retinal shape is perceived as moving towards or against the observer. This enables the viewer to estimate the distance from the object in terms of time-to- contact or time-from-contact.

- **Kinetic depth effect**: If a stationary rigid figure (for example, a wire cube) is placed in front of a point source of light so that its shadow falls on a translucent screen, an observer on the other side of the screen will see a two-dimensional pattern of lines. But if the cube rotates, the visual system will extract the necessary information for perception of the third dimension from the movements of the lines, and a cube is seen. This is an example of the kinetic depth effect. The effect also occurs when the rotating object is solid rather than an outline figure.

- **Relative size**: if two objects are known to be the same size (e.g., two trees), even if their absolute size is unknown, the relative size cues can provide information about the separation the two objects.

- **Familiar size**: since the visual angle of an object projected onto one eye’s retina decreases with distance, this information can be combined with previous knowledge of the object's size to determine the absolute depth of the object.

- **Absolute size**: even if the actual size of the object is unknown and there is only one object visible, a smaller object seems further away than a large object that is presented at the same location.

- **Aerial perspective**: due to light scattering by the atmosphere, objects that are at a great distance have lower luminance contrast and lower colour saturation. Because of this, images seem hazy the farther they are from a person's point of view. The colour of distant objects is also shifted toward the blue end of the spectrum (e.g., distant mountains). Some painters (e.g. Cézanne), employ "warm" pigments (red, yellow and orange) and "cool" ones (blue, violet, and blue-green) to make different parts of the painting appear at different depths

- **Curvilinear perspective**: at the outer extremes of the visual field, parallel lines become curved, as in a photo taken through a fisheye lens. Although it is usually eliminated from videos and photos by the cropping or framing of the picture, in real sight, the distortion effect enhances the viewer's sense of being positioned within a real, three-dimensional space.

- **Defocus blur**: selective image blurring is very commonly used in photographic and video for establishing the impression of depth. This contributes to the depth perception also in natural retinal images.

- **Accommodation**: this is the process through which the eye lens reshapes, changing its optical power to focus on a certain point. A depth cue is derived from the kinaesthetic sensations of contracting and relaxing the ciliary muscle.

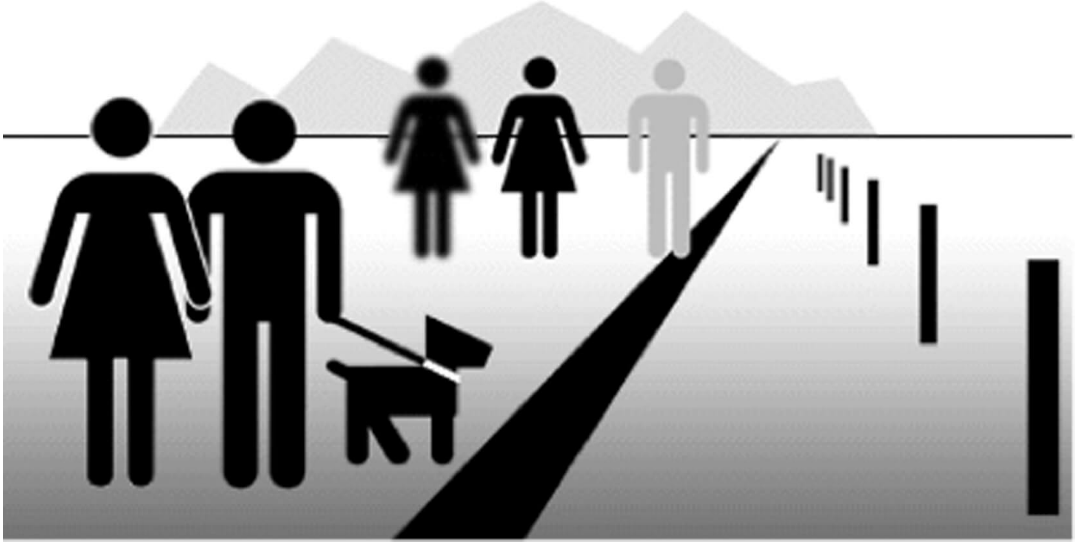


Figure 3. ILLUSTRATION OF SEVERAL DEPH CUES[2].

Binocular cues, namely convergence and stereopsis, are the ones that require the use of both eyes.

- **Convergence** allows the eyes to fixate on objects. Because the two lines of sight converge at a certain point, the angle formed at their intersection will be narrower or wider, depending on the distance between the eyes and the object. Thus, for close objects the angle will be wider, whereas for far objects the angle will be narrower. Depth information is gathered from the kinaesthetic sensation of stretching the extra-ocular muscles in a similar manner to what happens with accommodation.

- **Stereopsis** (a.k.a. retinal, parallax or binocular disparity) is based on the slight difference between the images collected by the eyes. Making use of such disparity the human brain can triangulate the distance between eyes and objects with a relative degree of accuracy.

- **Shadow stereopsis**: A. Medina Puerta demonstrated that retinal images with no parallax disparity but with different shadows are fused stereoscopically, imparting depth perception to the imaged scene. He named the phenomenon "shadow stereopsis" [41]

---

[2] The image illustrates six different monocular depth cues of non-even importance: occlusion, relative size, defocus blur, perspective, aerial perspective and shading

A graphics content that makes use of stereopsis should be referred as 'stereoscopic 3D' or 'stereo 3D' content. On the contrary, a graphic content that does not make use of binocular cues should be labelled as '2.5D'. However, it is common practice to name '3D' what is more precisely a 2.5D render. VR systems typically exploit stereopsis.

It has been demonstrated that the importance of each cue for the perception of depth is relative to the distance between the viewer and the virtual object. For several depth cues, this relationship has been consolidated by Nagata [42] (Figure 4).

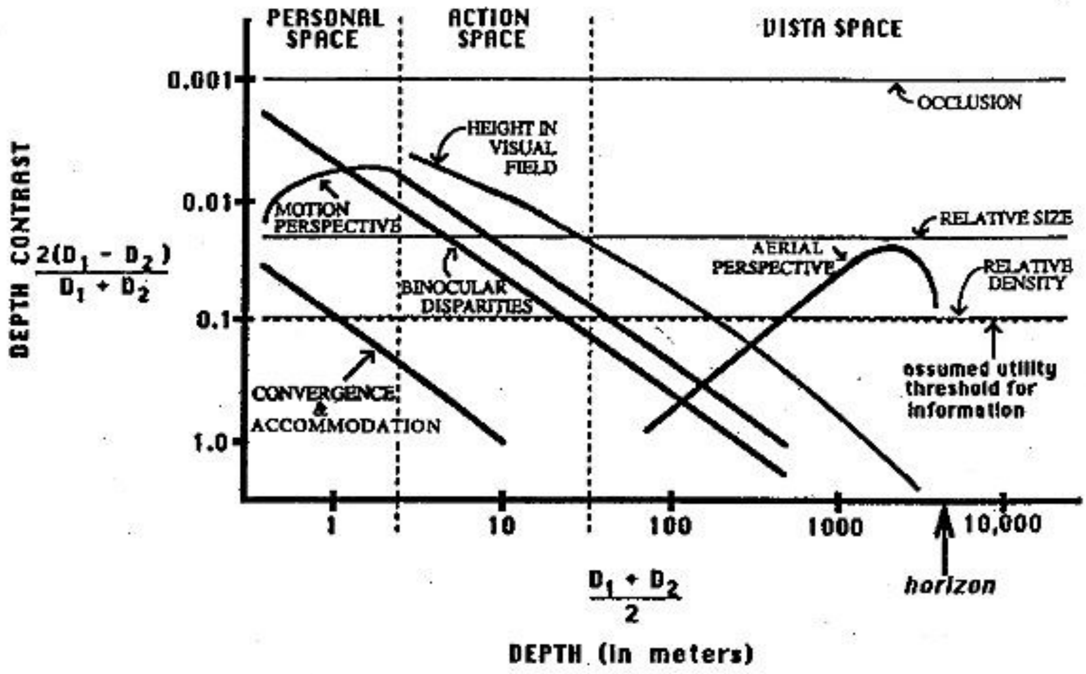
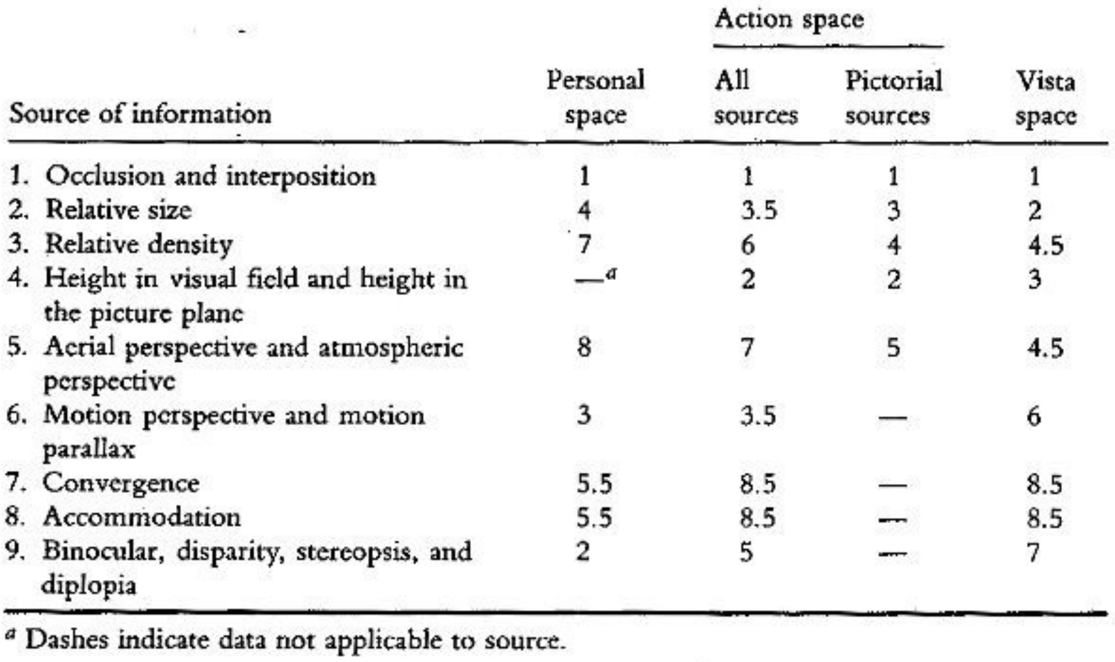


Figure 4. Lower perceivable depth contrast by means of a single depth cue as a function of the mean distance between the viewer and test objects[3]

In 1995, Cutting and Vishton ranked the importance of nine depth cues as a function of the distance between the object and the viewer. Their study distinguishes between three discreet depth intervals (that were already present in Nagata's study): personal space (0,5 - 1,5 m), action space (1.5 – 30 m) and vista space (>30 m) [43].

---

[3] Scale is logarithmic.



Source of information	Personal space	Action space All sources	Action space Pictorial sources	Vista space
1. Occlusion and interposition	1	1	1	1
2. Relative size	4	3.5	3	2
3. Relative density	7	6	4	4.5
4. Height in visual field and height in the picture plane	—[a]	2	2	3
5. Aerial perspective and atmospheric perspective	8	7	5	4.5
6. Motion perspective and motion parallax	3	3.5	—	6
7. Convergence	5.5	8.5	—	8.5
8. Accommodation	5.5	8.5	—	8.5
9. Binocular, disparity, stereopsis, and diplopia	2	5	—	7

[a] Dashes indicate data not applicable to source.

Figure 5. RANKING OF DEPTH CUES IMPORTANCE IN PERSONAL, ACTION AND VISTA SPACES .

Although Cutting and Vishton's chart is a good starting point, other studies do not agree on the importance of every single depth cue. For instance, in [44], Palmisano et al. suggest that binocular disparity has an impact on the vista space as well. This is somehow confirmed by very old studies on human sight [45], [46]. In the first study it is stated that human sight is capable of perceiving depth differences through very low binocular disparity. In the second study the authors conclude that binocular disparity is sufficient for distinguishing a point placed at infinity from a point placed up to 240 m from the user.

In any case, the importance of a depth cue providing information on the depth of an object is always relative to the presence of superior ranking depth cues for the same object. In other words, even if a depth cue provides some minor hint on the positioning such object, that cue is most likely to be overwritten by another having a greater importance in the designated space. For instance, accommodation, vergence and stereopsis can be easily overwritten by occlusion – i.e., even if these cues suggest that an object A is in front of an object B, but B is occluding A, the viewer will perceive B as being closer than A. However, it should not be taken for granted the contemporary presence of all depth cues. In this sense, a low-level ranking cue may become of primary importance in absence of others, which might be exactly the case of the control tower at night or in LVC. During these periods, some of the depth cues that the controller typically relies on are off because of the bad weather or because of the 'light based' visibility (e.g. 1, 2 and 5).

3.2.2.1 THE VERGENCE-ACCOMMODATION CONFLICT

Most V/AR display systems provide some depth cues to perform registration. However, in most cases, there are depth cues missing, in conflict or out of control of the display system. This is one of the primary cause of eye-strain, fatigue and cyber-sickness. For instance, the Vergence-Accommodation Conflict (VAC), is a well-known problem in the realm of virtual/augmented reality and stereoscopic displays. It derives from the fact that the light rays coming from the virtual image source provide an accommodation depth cue that is rarely consistent with the vergence one. This forces the viewer's brain to unnaturally adapt to conflicting cues, increases fusion time of binocular imagery and decreasing accuracy [47]. Also, it contributes to visual fatigue (asthenopia), especially during prolonged use [47]–[50], which, for some people, can even cause serious side-effects even after having used the device [51].

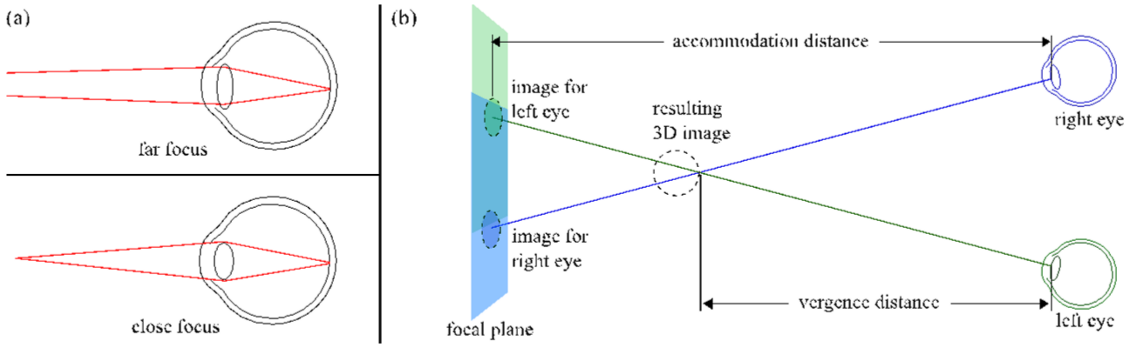


Figure 6. SCHEMATIC OF THE VERGENCE-ACCOMMODATION CONFLICT

The problem is not as acute in some domains, such as 3D TV or cinema viewing, as it is in HMDs (if the content and displays both fit certain constraints). In 3D cinematography, where the light comes from a distant screen and the virtual objects are usually located at a great depth, stereo parameters can be adjusted for each frame prior to viewing. For this reason, several methodologies have been developed on how to tailor the stereo content in order to make the viewer's comfortable [52]–[55]. These are often based on a framework of constraints such as the one from Lambooij et. al in [49]. However, these constraints are hardly applicable to the context of real time VR [56]–[58] and AR applications [59], where content is dynamic and interactive, and must be displayed on the fly, without much post processing.

It should be noted that when the vergence-accommodation conflict occurs, vergence and accommodation are not the only two depth cues conflicting. This is because the accommodation depth cue is probably in conflict with other depth cues as well. However, it has been pointed out that oculomotor cues of consistent vergence and accommodation, which are related to retinal cues of blur and disparity, are critical to comfortable 3D viewing experience. Retinal blur is the actual visual cue driving the oculomotor response of accommodation, which adjusts the eye's lens to focus on the desired depth, thus minimizing the blur. Likewise, retinal disparity is the visual cue that drives vergence. However, there is also a dual and parallel feedback loop between vergence and accommodation, and thus one becomes a secondary cue influencing the other [49], [60], [61]. In fact, Suryakumar et al. measured both vergence and accommodation at the same time during the viewing of stereoscopic imagery, concluding that accommodative response driven from disparity and resultant vergence is the same as the monocular response driven by retinal blur [62]. In a recent review of the topic, Bando et al. summarize some of the literature about this feedback mechanism within the human visual cortex [51].

The practice of providing the viewer with accommodation, vergence and stereopsis depth cues that lead him, or her, into thinking that the object is placed at infinitum is commonly referred to as 'collimation at optical infinity'. Optical infinity is a point in space from which the originating light rays can be considered as if they were parallel (collimated) when reaching the eye. Consequently, beyond optical infinity the eyes' accommodation and vengeance adjustments are negligible. Based on a literature review, Peterson indicates that 6 m can be considered as optical infinity [38]. Others suggest 9 meters [63].

Binocular disparity, which is the distance between a point in the left eye image and the very same point in the right eye image in screen space coordinates, increases with the distance between the viewer and the virtual point in an asymptotic way. If the projection screen is parallel to the segment connecting the viewer's eyes (a.k.a. baseline), the asymptotic value is the viewer's interpapillary distance (IPD)[4], which is typically close to 6 cm. If the projection screen is not parallel to the segment connecting the viewer's eyes the asymptotic value is the IPD multiplied by the cosine of the angle between the eye's segment and the screen plane direction.

As panoramic environments only concern objects more than 30 m away, accommodation, vergence and binocular disparity of the augmented reality content should provide a visual stimulus which is consistent with the one of the real object. For vergence and accommodation this means that the virtual image focal plane should be positioned at least at optical infinity (i.e. at least six meters away from the user). To provide such visual stimulus by means of a transparent screen, either the screen itself must be moved to optical infinity or the emitted light must be collimated beyond that by means of optical lenses. The projection screen must also provide a binocular overlay since the (parallel) light rays from a single point will intersect the display surface at two different

---

[4] This is the distance between the two eyes, measured at the pupils' center.

points before reaching the two eyes. However, if a common projection display surface is positioned in front of the user, it can be seen by both eyes. Therefore, the left eye image in the display must be blocked for the right eye and vice versa. This is usually performed through different multiplexing techniques [64]. In HMDs instead, each eye has its own image source [65].

Since binocular disparity is not effective in panoramic environments it has been suggested that it can be approximated with biocular disparity [15], [38]. Biocular disparity should not be confused with binocular disparity, where two slightly different images are rendered. When a biocular stimulus is used, each eye is provided with the same virtual image slightly translated left or right of a distance which is typically half of the IPD to place the virtual content at infinitum. However, it might not be particularly convenient to use such approximation in a multi-screen non-planar V/AR environment, because this would increase the complexity of seams handling without truly eliminating the need for tracking the viewer' eyes position with respect to the screen position and orientation.

For non-registered information, such as wind direction and speed, temperature, QNH, etc., it might be convenient to place the AR content at optical infinity to minimize refocusing between far and close objects. However, this might depend on the controller's tasks and on the layout of his/her working position. Since little research has been performed on this topic it is still unclear which solution would provide the less eye strain, fatigue and tunnelling effect – i.e. failure to switch between real and superimposed content or even between two synthetic contents (if placed at different depths). Much work has been done on collimation for cockpit HUDs, where some results show that collimation at optical infinity is better [66], while others suggest that the symbols should be displayed at 2 m from the observer [67].

As already mentioned, to achieve registration, one crucial factor is to consider the coupling between the observer's movements and the generation of the VR stimuli. Thus, a major requirement for V/AR systems is to have accurate spatial data of the observed object, display and observer at all instances. This may be obtained by means of depth from stereo, infrared tracking or many other techniques (more about this in the next paragraph). Inaccurate measurements or latency in the tracking methodology lead to registration errors, which can seriously affect the system usability [38]. Tracking is a widely researched topic [68], [69] and will not be further discussed in this document. Eventually, the tracking process must result in the head/eyes coordinates being fed, in real time, to the rendering pipeline. Also, a custom rendering pipeline with a modified projection algorithm is needed to generate the binocular disparity stimuli that are not conflicting with the other depth cues [39].

This kind of behaviour can also be applied to virtual reality environments and synthetic vision systems that can benefit from the application of the fish-tank virtual reality paradigm [70].

At smaller ranges the perspective from each eye is significantly different and the expense of generating two different visual channels for the computer-generated Imagery becomes worthwhile. On the contrary it would be difficult (and not particularly beneficial) to have an Enhanced Vision System (EVS) that follows this rule, given that the camera's optical unit is fixed in space with respect to the parent body (could be an aircraft fuselage or a control tower structure). In any case, for such systems, it is still imperative that the augmented reality content matches the one of the video stream by means of precise calculation of the camera position and rotation with respect to the surrounding environment (which can be derived from the orientation between the camera and the parent body object).

3.2.3 RENDERING IMAGES FOR VIRTUAL AND AUGMENTED REALITY

Most V/AR systems operate on some variant of the pinhole camera metaphor, i.e. a camera object exists in the virtual environment and regularly takes bi-dimensional snapshots of the computer-generated scene. A camera object is characterized by a frustum[5], a position and orientation in space. Whether the camera is fixed or not depends on the implementation as well as the mutable o immutable nature of the frustum shape. The commonest rendering setups are presented below.

3.2.3.1 DESKTOP VIRTUAL/AUGMENTED REALITY (D-V/AR)

Desktop Virtual/Augmented Reality (D-V/AR) is a basic implementation of the pinhole camera model. D-V/AR is ubiquitously supported as the default output mode of nearly every graphics engine or application available today. It is based on a static projection model, which uses symmetrical *frustum*[6]. Thus, it produces a single, camera-centred, perspective image and does not require any special equipment to be used, meaning that any framed or unframed planar display is sufficient. As the simplest form of V/AR, D-V/AR avoids many issues, such as eyestrain, increased computational cost, latency, etc. For this technology to work properly, the user should be positioned relatively to the screen as the camera object is positioned with respect to the near clip plane[3], e.g. head-

---

[5] A *frustum* is a six-sided truncated pyramid, which originates sectioning the shape the virtual camera field of view by means of two user-defined clipping planes. These are known as the 'far clip plane' and the 'near clip plane'. The latter, is the one on which the virtual world is projected as a necessary step of the rendering pipeline.

centred on the screen normal[7]. Given that proportions – i.e. horizontal and vertical field of view – should be kept identical, a *frustum* scale factor is acceptable if the viewer is aware of watching an exaggerated or diminished virtual world. On the contrary, relative movements between the observer and the screen, including back and forward movements, should not be allowed, as they modify the physical *frustum,* whereas the projection model behind the software remains unvaried. In other words, D-V/AR should be considered a 'static' display technique.

3.2.3.2 OFF-AXIS VIRTUAL/AUGMENTED REALITY (O-V/AR)

Off-axis Virtual/Augmented Reality (O-V/AR) comes in handy when the viewing position is not screen-centred, meaning that the straight line from viewer's eyes to the screen, drawn along the screen normal direction, no longer strikes the display in the middle. In this case, an asymmetric *frustum* is used to render the scene. However, the near clip plane stays perpendicular to the camera depth axis, therefore, the same orientation must be used for the physical display. As the frustum shape does not change in time relative movements between the observer and the screen are still forbidden.

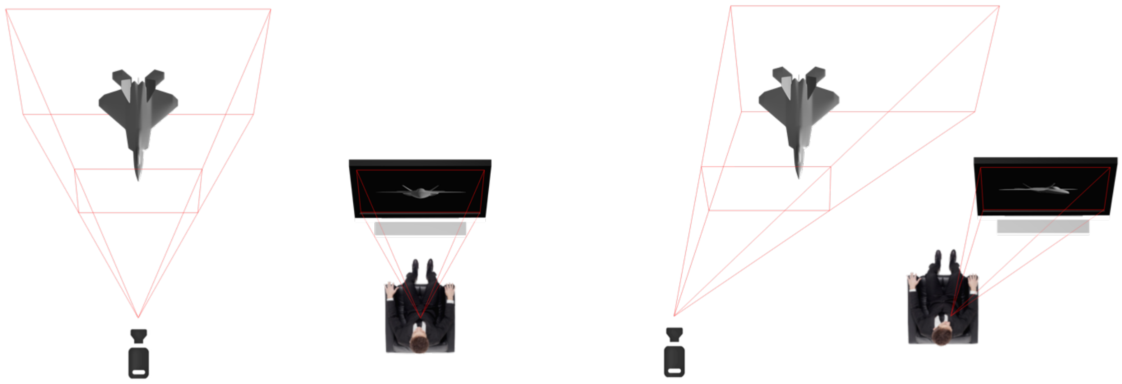


Figure 7. D-V/AR (LEFT IMAGE) AND C-V/AR (RIGHT IMAGE) SCHEMATICS

3.2.3.3 GENERALIZED VIRTUAL/AUGMENTED REALITY (G-V/AR)

Generalized Virtual/Augmented Reality (G-V/AR) is an off-axis projection development that allows the projection plane, therefore the viewing device surface, to be arbitrary oriented. This is achieved by multiplying the standard off-axis projection matrix by a further rotation matrix (more about projection matrixes in section III). Once the viewer standpoint is known (and stays still), the display surface might be arbitrary oriented, i.e. rotated, installed upside down, laid flat on the floor or hung from the ceiling. For all

---

[7] For the sake of readability, here and from now on, we will refer to the straight line being orthogonal to the screen and passing by the center of it simply as the 'screen normal'.

intents and purposes, this makes G-V/AR applicable to a wide range of VE architectures, such as, fixed-viewpoint, non-planar, multi-screen VEs.

3.2.3.4 Stereoscopic Virtual/Augmented Reality (S-V/AR)

Stereoscopic Virtual/Augmented Reality (S-V/AR) is a dual camera paradigm suited for binocular vision. Stereovision is achieved by rendering the virtual scene twice, once for each eye. Image pairs (a.k.a. stereo pairs) are encoded and filtered so that each single image is only seen by the matching eye. Encoding techniques include colour spectrum decomposition, light polarisation, temporal encoding and spatial encoding. Filtering is most easily attained through special equipment, e.g. polarized eyeglasses, coloured eyeglasses and shutter glasses, but might be also achieved by looking at the screen from a specific position. Encoding and filtering techniques are paired together and named Passive, Active or Auto-stereoscopic techniques, based on their working principles. The need for a triggering system in the filtering equipment – if any – determines whether a technique is Active – or Passive. Following this criterion, passive systems are colour filtering and polarisation, whereas temporal encoding, in combination with shutter glasses, is to be considered Active. Finally, auto-stereoscopic techniques, such as *Parallax Barrier* and *Lenticular Lens*, do not require additional filtering equipment because they encode spatially. In this case, it is the physical distance between the viewer's eyes that filters the images.

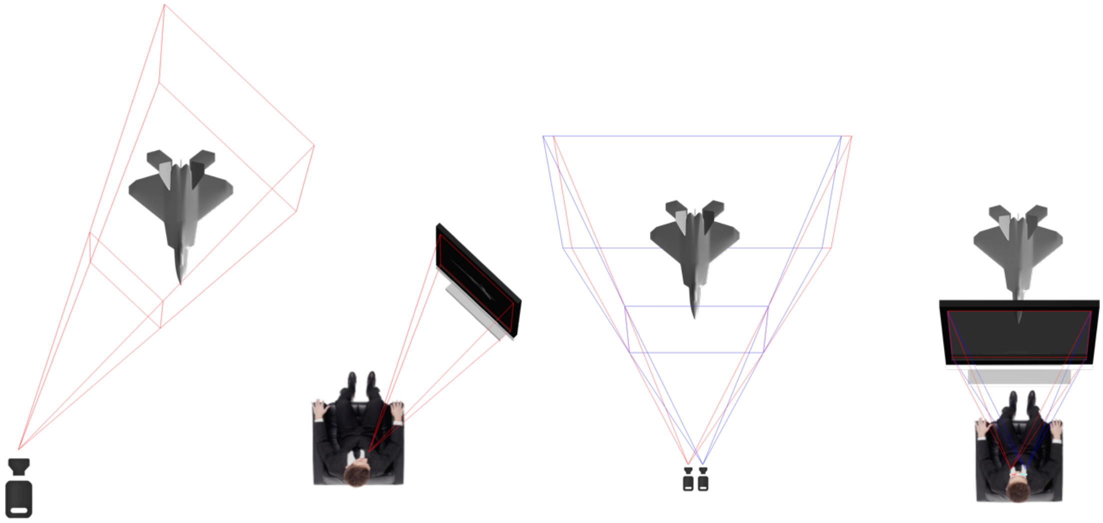


Figure 8. G-V/AR (left image) and S-V/AR (right image) schematics

Even though stereo vision techniques have been around for at least six decades, there are still many widespread misconceptions regarding how stereo pairs should be rendered. As Lang commented[71] on a brief interview with Oliver Keylos: *"It turns out that rendering stereoscopic 3D images is not as simple as slapping two slightly different views side-by-side for each eye. There's lots of nuance that goes into rendering an appropriate 3D view that properly mimics real world vision – and there's lot's that can go wrong if you aren't careful".* This is especially true given that stereoscopic 3D has been pushing hard into the mainstream market segment over the last few years [72]. *"The subtleties of improper 3D rendering could be a major hurdle to widespread consumer adoption of virtual reality, in a way that the everyday first-time VR user won't think: – this is obviously wrong, let me see how to fix it –. They'll say instead: – I guess 3D isn't so great after all; I'll pass. –", says Keylos* [72]. In a nutshell, there are primarily three ways of generating stereo pairs: Parallel stereo, Toe-in stereo and Skewed-*frusta* stereo*.* The latter is correct, whilst the others are not.

3.2.3.4.1 PARALLEL STEREO

Parallel stereo is the easiest and arguably the most common stereovision content creation technique. You simply take two physical or virtual cameras and put them next to each other, with their viewing directions precisely parallel. This is like using the D-V/AR technique twice. Somehow, this symmetric-*frustum* setup will 'work', as the view of the resulting footage will produce a three-dimensional effect. However, after a while, the viewer will realize that the output does not produce the desired effect. Looking carefully, you it is evident that everything in the scene, up to infinity, appears to float in front of your screen. Instead, near objects should be floating in front of the screen, for far objects should be floating behind the screen. Unfortunately, with a parallel set up, there is no way to achieve this. Since the two cameras are 'stereo-focused' at infinity, they can only produce negative horizontal disparity – i.e. one will always perceive objects as if they were in front of the screen. If we want both positive and negative disparity values to result from the rendering process, we must move the stereo-focus plane closer to the cameras set up.

3.2.3.4.2 TOE-IN STEREO

Playing around with cameras set ups, one will find that a symmetrical *frustum* projection model is anything but flexible. In fact, the best we can do to move the stereo-focus plane away from infinity, is to slightly rotate both cameras inwards. This way, the intersection of *frustums'* bisectors defines a closer stereo-focus point (not plane). This approach is often called Toe-in stereo, and, again, it 'sort of' works. Toe-in stereo makes sense intuitively. After all, our eyes rotate inwards when we focus on nearby objects[72]. However, the perspective model lying at the basis of three-dimensional computer graphics should be different from the one used by our own eyes. In real sight, the physical world is directly projected onto our retinas, whilst in computer graphics, an intermediate projection screen is used – i.e. the V/AR display. Therefore, the screen orientation, not the eye orientation, should define the projection. Later, the retinal projection will take care of itself.

Figure 9, right side, shows the render result of the Toe-in rig. As none of the cube's faces is orthogonal to the cameras' viewing directions, a trapezoidal shape is rendered, which looks like a keystoning[8] effect. The two images are then stick together on a single frame, which will be oriented in space perpendicularly to the median line of sight. Since neither the left nor the right view looks as if a real cube was seen through the screen to the naked eye, the Toe-in perspective model fails. This failure leads to serious problems in stereovision, e.g. incorrect depth assessment, inaccurate shape evaluation or three-dimensional illusion breakdown [40]. Looking closely at Figure 9, one can notice that the keystoning effect is more severe towards the left and right edges of the image. This is the reason why Toe-in stereo is considered to work 'well enough' around the centre of the screen, whereas the stereoscopic 3D effect breaks down at the edges of the image. No wonder that in Toe-in stereo cinematography two basic rules of thumb are used: first, one should reduce the amount of eye separation and, second, one should keep the action – therefore the audience's eyes – at the centre of the screen [72]. However, these rules are merely workarounds for a problem that would not exist in first place, if stereo were done properly. All in all, Toe-in stereo is only a rough approximation to correct stereo; thus, it should not be used. Even when the keystoning effect is less severe, our eyes will dart around, trying to make sense of the mismatching images, possibly leading to eye strain and headaches[40], [51], [72]. The fact that Toe-in stereo is seemingly widely used in the three-dimensional industry could (partly) explain the discomfort that many people report while experiencing stereoscopic movies and stereoscopic computer graphics [72]. A good practice against headaches is the one that treats the cause, not the symptoms.

---

[8] Keystoning is a typical video projection unwanted phenomenon due to the use of a projection surface non-orthogonal to the projection beam.

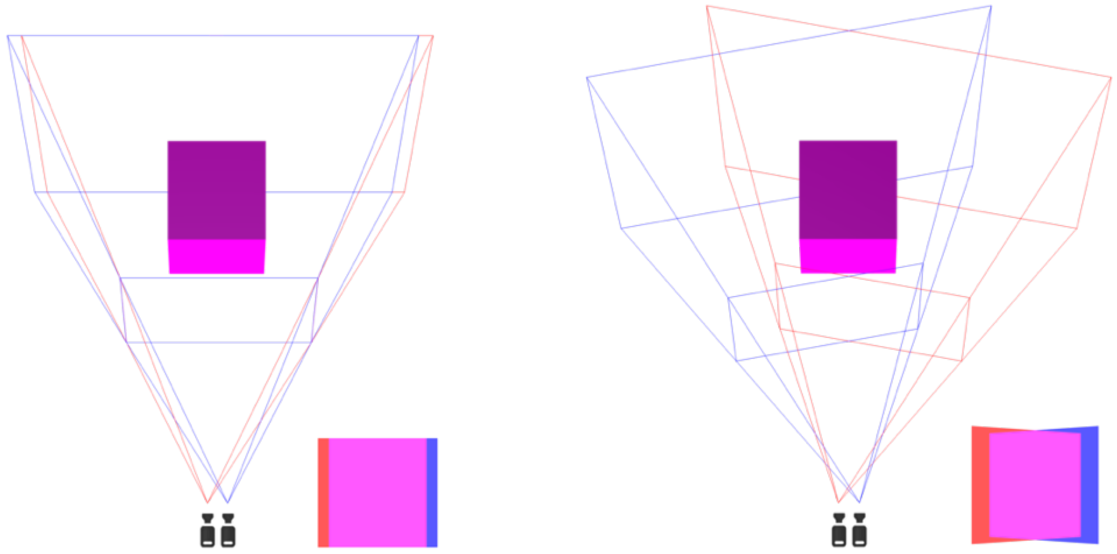


Figure 9. COMPARISON BETWEEN OFF-AXIS STEREO (LEFT IMAGE) AND TOE-IN STEREO (RIGHT IMAGE)

3.2.3.4.3 OFF-AXIS STEREO

So far, we have showed that conventional Toe-in stereo often leads to depth, shape or layout misperception, resulting in three-dimensional illusion breakdown and serious discomfort for the user. Off-axis stereo uses asymmetric-*frusta* (a.k.a. skewed-*frusta*) to solve these issues. Each *frustum* extends from the corresponding eye to the screen corners positions. With this set up a shared plane exists, which is the plane of the screen, a.k.a. the 'stereo-focus' plane or 'zero-parallax' plane. This is oriented in space like the near (or far) clip planes. Nevertheless, two distinct viewpoints exist. Therefore, while keystoning is avoided, separation is ensured. All in all, a physical camera rig should always use lens shift, whereas a virtual cameras setup should always use skewed frusta – i.e. off-axis projection.

It stands to reason that, as the viewpoint changes, the perspective model should be modified as well. This is precisely why eye tracking is the only way toward a consistent projection initialization and maintenance. Moreover, this is how one generates binocular disparity signals that are consistent with depth cues coming from motion parallax, instead of being a further cause of visual fatigue for the observer [40].

3.2.3.5 HEAD-COUPLED VIRTUAL/AUGMENTED REALITY (H-V/AR)

Head-Coupled Virtual/Augmented Reality (H-V/AR), a.k.a. head-coupled perspective, operates on a slightly different principle than D-V/AR. It is the virtual window metaphor, rather than the pinhole camera model, that better fits this technique. A projection surface, representing the physical display, is defined in the virtual environment. Also, the viewer's head position is tracked in space and time. Virtual objects are projected through the so-defined surface, toward the user's head. Thus, the projection outcome depends on the relative position between the viewer's head and the projection surface. Clearly, a perspective projection is still used. Nevertheless, the projection model is not defined *a priori,* but rather computed 'just-in-time'. As a matter of fact, while the observer moves freely in the physical environment, the display becomes a framed window on the virtual world.

A strong limitation of H-V/AR is that any other viewer, looking at the very same display, will perceive a distorted (incoherent) image. This is always true unless multiple perspectives are used (i.e. calculated and displayed).

3.2.3.6 Fish-Tank Virtual/Augmented Reality (F-V/AR)

Fish-Tank Virtual/Augmented Reality (F-V/AR), a.k.a. Eye-Coupled V/AR (E-V/AR) [73] or True Dynamic 3D (TD3D) [40], is an improvement over H-V/AR, which separately considers the viewer's left and right eye position. In this sense F-VAR is a combination of Eye-Coupled Perspective (ECP) and S-V/AR.

3.2.3.7 Hemispherical Virtual/Augmented Reality (H-V/AR)

Hemispherical *VR, a.k.a.* Fish-eye *VR,* has been portrayed as the ultimate technology for VR: a synthetic (computer-based) environment where no frame impinges on the user comprehensive view of the virtual world. A seamless, widescreen, hemispherical display (a.k.a. 'dome') is used, so that the entire user's FOV is engaged. To avoid heavy distortion, a custom rendering pipeline is needed. First, a CAVE-like generalized projection is used. Second, a series of coordinate mappings adapt the result to be projected [74].

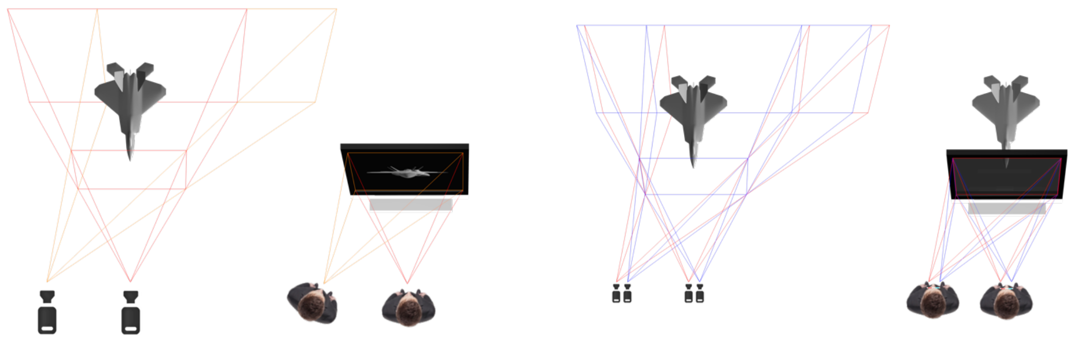


Figure 10. H-V/AR (left image) and F-V/AR (right image) schematics

3.2.3.8 GAZE-DEPENDENT DEPTH OF FIELD V/AR (DOF-V/AR)

In a gaze-dependent V/AR application a scene is rendered and visualised while the viewer's eyes movements and relative position regulate a depth of field simulation. Saccades and fixations are captured by means of an eye tracker and the line of sight direction computed with respect to the virtual environment.

3.3 TAXONOMY

A first classification of AR systems distinguishes between "optical see-through" and "video see-through" systems. When the combination of the real and virtual image is performed by means of lenses, mirrors and/or transparent displays or the system is defined as optical see-through. On the contrary, when the combination is obtained using cameras to transform the real-world view in a video feed that is later merged with the synthetic information the system is defined as video see-through. A third approach, is based on the direct projection of the synthetic information on real objects. Only optically combined displays are considered in this work, since they leave the view of the real world nearly intact and intuitively are thus better suited for panoramic environments.

Aside from the type of device – optical or video see-through – Bimber and Raskar made a classification of AR displays based on where they stand along the optical path between the object and the viewer's eyes [75].

According to this classification three main categories are defined:

1. **Head-attached displays**: worn by users on their head.

2. **Hand-held devices**: hold by users in their hands.

3. **Spatial devices**: placed into the environment.

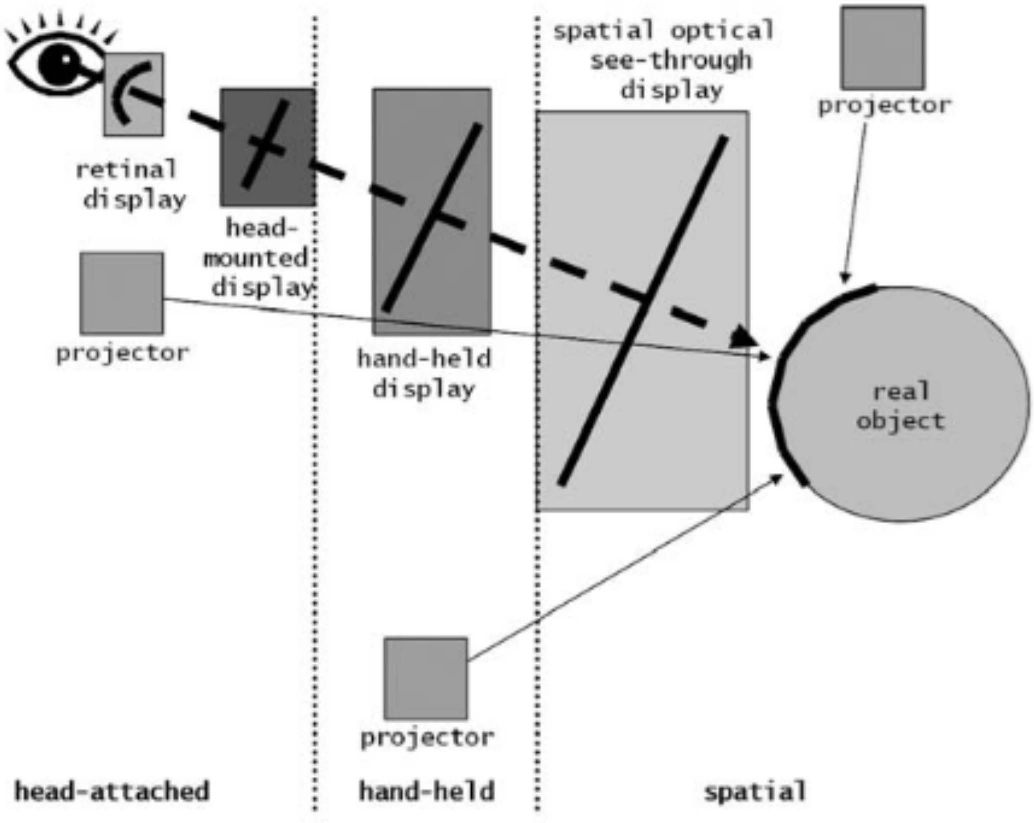


Figure 11. CLASSIFICATION OF THE AUGMENTED AEALITY TECHNOLOGIES BY BIMBER AND RASKAR

1. **Head–attached** devices include three main types of head-wearable displays:

   a) **Retinal** Displays make use of low-power semiconductor lasers to scan modulated light directly on the eye retina.

   b) **Head-mounted Displays** commonly referred to as HMDs consist in a class of devices that make use of very small displays put in front of the user's eyes. They can be either "optical see-through HMDs" or "video see-through HMDs" depending on the way the real and the virtual image are combined.

   c) **Head-mounted projectors** adopt miniature projectors that project images on the surface of a real-world object. Depending on the target surface they can be further distinguished as Head Mounted Projective Displays (HMPDs) or Projective Head Mounted Displays (PHMDs). In the first case the target surface is a retro-reflective combiner in front of the viewer, whereas in the second case it is the diffuse surface of an object. It's worth to remind that the projector based systems are not suitable to those environments where the real objects are located far away from the user. Additionally, the performance of such systems is strongly affected by the environmental lighting conditions.

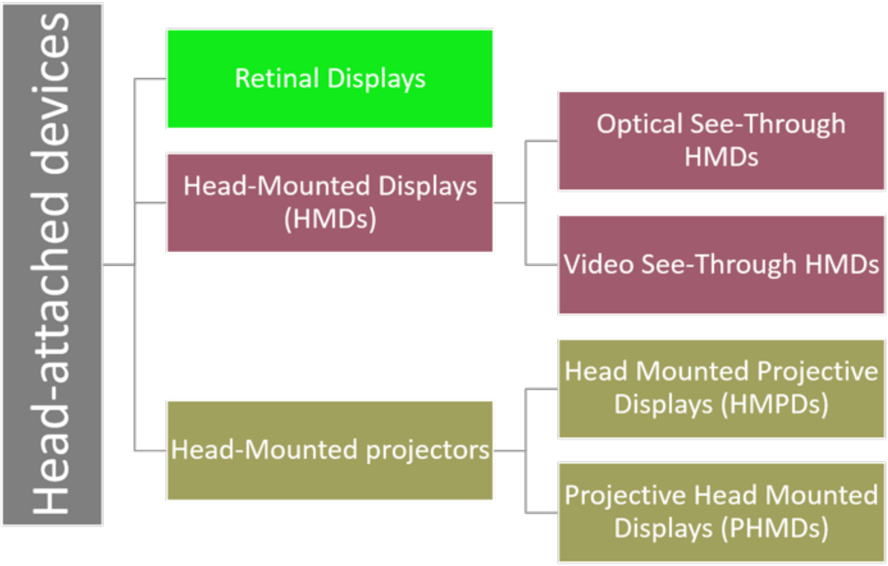


Figure 12. Taxonomy of head-attached devices

2. Hand-held devices include:

   a) **Hand-held displays**. Often embedded within consumer devices, namely Tablet PCs, PDAs (personal digital assistant) or smartphones, as video see-through displays. Alternative solutions based on optical see-through displays are diffused to a lesser extent.

   b) **Hand-held video-projectors**: they depict the synthetic information on the real object by directly projecting it on the object surface.

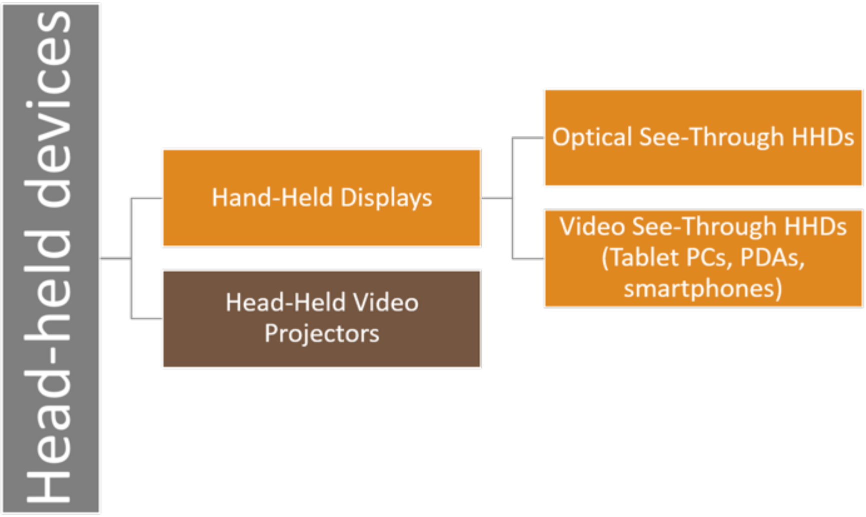


Figure 13. Taxonomy of hand-held devices

3. **Spatial devices** differentiate from head-mounted and hand-held devices as they are not fixed to the user, they are instead linked to the space, e.g. to a desk, the ceiling or the floor. They can be further classified into:

   a) **Screen-based video see-through devices**: they make use of video see-through displays providing the so-called “window on the world” effect.

   b) **Spatial Optical See-Through devices**: they make use of optical combiners to mix the light emitted by the real world with the one produced by an image source that displays the rendered graphics. These are often referred to as head-up displays (HUD).

   c) **Projection based Spatial Displays**: they use front-projection to seamlessly project images directly on physical objects' surfaces.

The taxonomy described above was conceived with the aim of classifying AR devices. Nevertheless, it is possible to derive a similar classification for VR visual devices as well. However, while AR technologies are focused on the vision sensory system, VR may address other sensory systems such as hearing and touch. A comprehensive taxonomy for existing VR technologies can be found in [76].

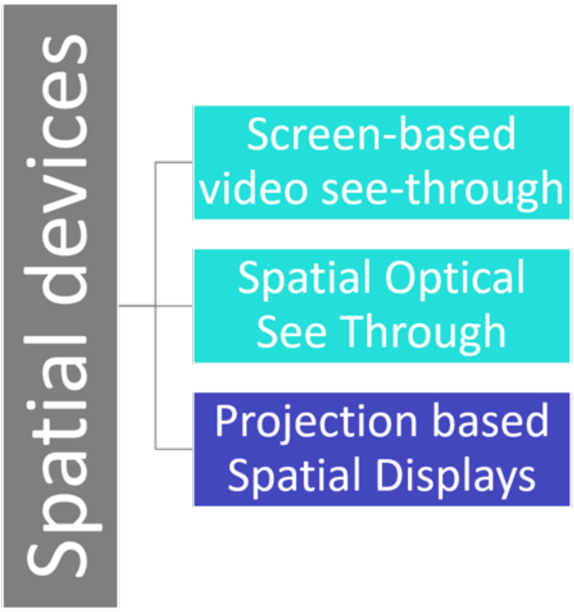


Figure 14. TAXONOMY OF SPATIAL DEVICES

3.4 DEVICES

For the scope of this research, only a subset of V/AR devices will be considered and further discussed. In chapter 6, we compare such devices based on an integrated Quality Function Deployment and Analytic Hierarchy Process approach. Thus, it is important hereafter to describe hereafter the main characteristics of each system.

3.4.1 HEAD MOUNTED DISPLAYS

A head-mounted display, often referred as helmet-mounted display in military applications, both abbreviated HMD, is a single-user V/AR piece of equipment worn on the head or as part of a helmet that provides symbolic or pictorial information by introducing into the user's visual pathway a virtual image. HMDs differ in whether they can display just a computer generated image, show live images from the real world or a combination of both. In the first case, they can be referred as a Virtual Reality Headset (VRH), which are typically opaque and only provide an immersive visual of the virtual environment. In the latter case, they should be referred as a See-Through Head Mounted Displays (ST-HMDs). Amongst ST-HMDs two further categories can be distinguished: video see-through and optical see-through systems. Video see-through displays capture video of the real world and digitally combine it with synthetic imagery before re-displaying it to the user. This can be considered a form of mixed reality. Optical see-through systems let through or propagate light rays from the real world and use semi-transparent combiners (a.k.a. beam-splitters) to combine them with virtual imagery. Helmet-Mounted Sight (HMS) is another term often used referring to an HMD that provides only a simple targeting gunsight.

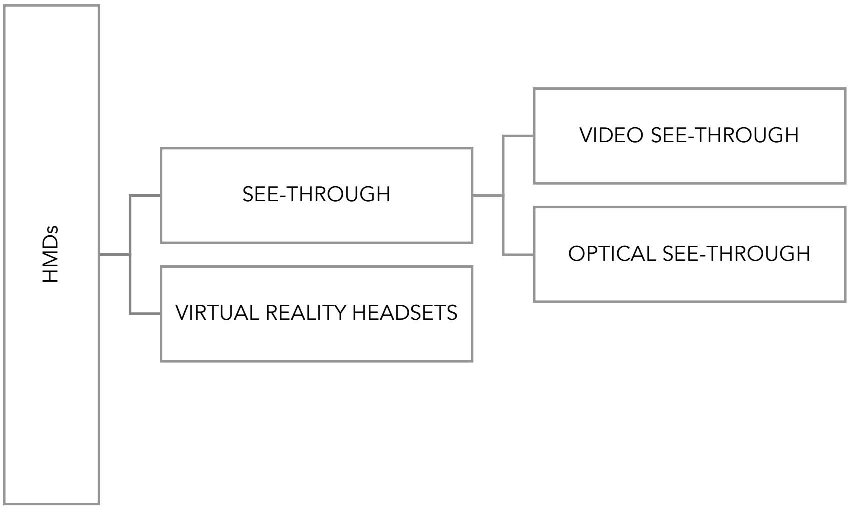


Figure 15. BASIC CLASSIFICATION OF HEAD MOUNTED DISPLAYS

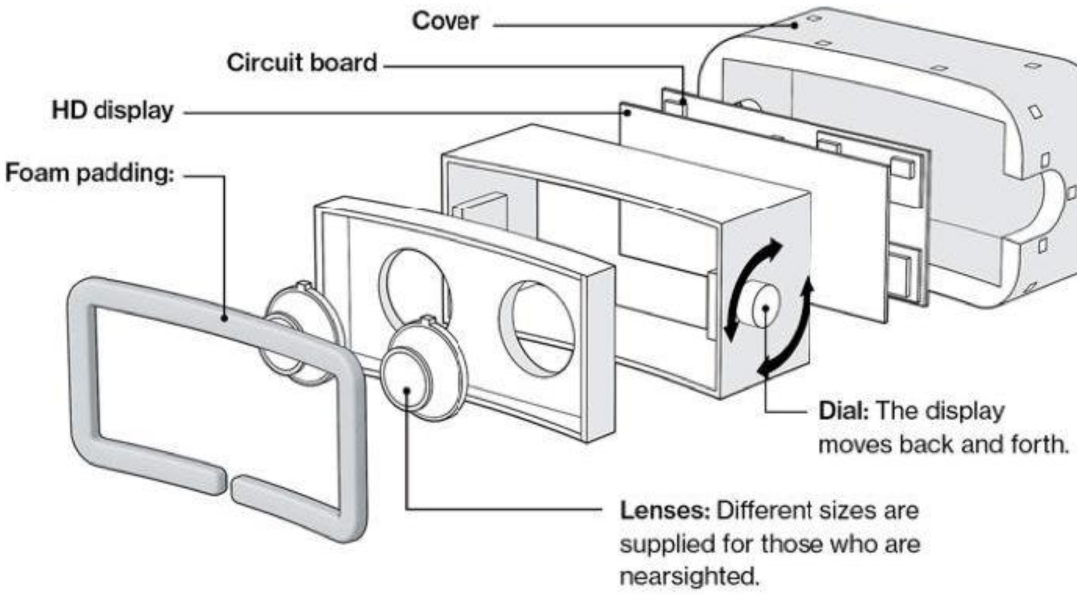


Figure 16. SCHEMATIC OF THE FIRST OCULUS RIFT VIRTUAL REALITY HEADSET[9]

In its simplest form, a HMD consists of an image source and accommodative optics in head mount. However, to understand the more complicated layouts it is useful to classify HMDs based on the type of visual stimuli they provide, i.e. (a) monocular, (b) biocular or (c) binocular, which is sometimes referred as ‘ocularity’.

a) **Monocular** HMDs only provide a single image to a single eye. This is usually the lightest, least expensive, and simplest HMD type. For these reasons, most of current military HMD systems are monocular. A few examples are provided by Melzer in [63]. This particular design often comes associated with an asymmetric centre of gravity and some issues dealing with focus, eye dominance, binocular rivalry, and ocular-motor instability [77].

b) **Biocular** HMDs provide the same image for both eyes (either shared or duplicated). The biocular approach is more complex than the monocular design, but eliminates the ocular-motor instability issues associated with monocular displays. In the end, viewing imagery with two eyes has been shown to yield improvements in detection as well as providing a more comfortable viewing experience [78], [79]. The primary disadvantage of the biocular design is that the image source is usually located in the forehead region, making it more difficult to package. Biocular displays use one or two identical image sources paired with a single set or double set of optics. For this reason, they tend to be larger and heavier than monocular systems, which use a single image source and a single set

---

[9] Image: http://blogs.ucl.ac.uk/digital-education/tag/oculus-rift/

of optics. Finally, binocular HMDs are subject to strict alignment, focus and calibration requirements.

c) **Binocular** HMDs present two slightly different images to the right and left eyes (as in real sight). This is the most complex, expensive, and heaviest of all three options, but one which has all the advantages of a two-eyed system with the added benefit of providing partial binocular overlap (which enlarges the horizontal field of view) and binocular disparity. A binocular HMD is subject to the same alignment, focus and calibration requirements as the biocular design. The simple magnifier or the freeform waveguide designs are two examples of fixed Stereoscopic HMDs. A few examples of military grade Binocular HMDs can be found in [63]. For consumer grade devices and at the time being, the reader can refer to the Microsoft^{TM} HoloLens^{TM} and the DAQRI Smart Helmet for see-though HMDs as well as to the Oculus Rift and the HTC Vive for VRHs.

The following table shows a comparison between the benefits and drawbacks of monocular, biocular and binocular HMDs.

CONFIGURATION	ADVANTAGES	DRAWBACKS
MONOCULAR (SINGLE IMAGE VIEWED BY SINGLE EYE)	• LIGHTWEIGHT • COMPACT • SIMPLE CALIBRATION • CHEAP	• OCULAR MOTOR INSTABILITY • EYE DOMINANCE • FOCUS ISSUES • ASYMMETRICAL CENTRE OF GRAVITY
BIOCULAR (SINGLE IMAGE VIEWED BY BOTH EYES)	• CHEAP • SYMMETRICAL CENTRE OF GRAVITY	• COMPLEX CALIBRATION • BULKY • HEAVY
BINOCULAR (DOUBLE IMAGE VIEWED BY BOTH EYES)	• LARGER FOV • BETTER DEPTH PERCEPTION • SYMMETRICAL CENTRE OF GRAVITY	• COMPLEX CALIBRATION • BULKY • HEAVY • EXPENSIVE

Table I. BENEFITS AND DRAWBACKS OF MONOCULAR, BIOCULAR AND BINOCULAR HEAD MOUNTED DISPLAYS

The HMD itself is often part of a larger system which includes an image generator (i.e. an integrated or separate computer), a head tracker (might be synthesized from multiple sources, such as three-axis gyros, accelerometers and magnetometers), video and/or infrared cameras, depth sensors, audio input and several other input devices. Some HMD vendors offer on-board operating systems (e.g. Android), allowing applications to run locally on the HMD and eliminating the need to be tethered to an external image generator. These are sometimes referred to as Smart Goggles. Other devices perform some calculation locally and continuously exchange information with an external image generator, such as head position, orientation and surrounding space geometry.

The information displayed on a HMD can vary from simple unchanging symbols, through more complex numerical or alphanumerical information, to graphical imagery superimposed on a video image obtained from a sensor or directly linked to the real scene.

Major HMDs applications include military, police, firefighting, medicine, video gaming, sports, etc. In some fields, such as firefighting and infantry, HMDs are often used as a hands-off information source. They display tactical information such as maps or thermal imaging data while viewing the real scene. On the contrary, in aviation, HMDs are increasingly being integrated into helicopters and fighter aircrafts pilot's helmets. In the cockpit, the HMD becomes part of a Visually Coupled System (VCS) that includes the HMD, a head position tracker, and a graphics engine or video source [63]. As the pilot turns his or her head, the tracker relays the orientation data to the mission computer, which updates the displayed information accordingly. This provides the pilot with a multitude of real-time data that is linked to the head orientation. A full description of the potential usage of such system as well as a detailed review of military grade HMDs can be found in [63].

3.4.1.1 SPECIFICATIONS

Field of View (FOV). The FOV can be defined as the aperture of the virtual image at its maximum extents with respect to the viewer’s eyes median point, typically expressed in degrees. No existing HMD achieves the wide FOV of the human visual system, which is about 150°-160° in the horizontal direction and about 110°-120° in the vertical direction for the single eye. The single eye FOV is wider on the temporal side (about 90°-100°) than it is on the nasal side (about 60°) because the nose blocks part of the FOV. The binocular field of view is about 180-200° in the horizontal direction (figure 3) and 110°-120° in the vertical direction [19], [47]. Although both vertical and horizontal FOVs matter, the latter is often emphasized because it is considered more important [80]. Most people do not have a good feel for what a FOV would look like; thus, manufacturers often specify a virtual screen size, viewed from a specified distance, to describe their devices’ FOV. When asked about HMD requirements, users will typically want more of FOV and resolution. Old generation consumer-level HMDs typically offered a FOV of about 30-40° whereas professional HMDs or new generation HMDs offer a field of view up to 150°. However, optical ST-HUDs are typically more FOV limited than video see-through HMDs and VRH. Both intuition and evidence lead to the conclusion that decreasing the FOV size results in a performance loss and compromises the viewer’s sense of immersion and situational awareness. Thus, a wide FOV is highly desirable (but not always necessary). However, for a fixed display, an HMD cannot both simultaneously increase spatial resolution and FOV because these attributes are linked together by the focal length of the collimating optics. Also, increasing the FOV by increasing optical magnification usually provokes some weight increase due to the use of larger optical elements [81].

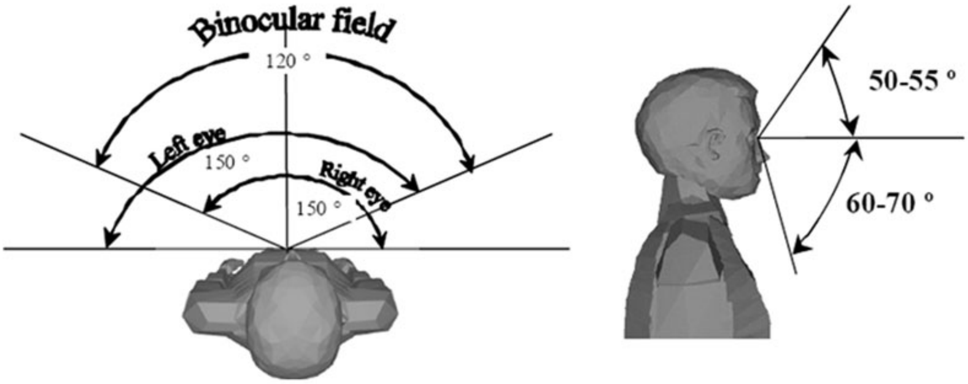


Figure 17. GRAPHICAL REPRESENTATION OF THE HUMAN FIELD OF VIEW

Resolution. HMDs producers usually mention either the total number of pixels or the number of pixels per degree. Listing the total number of pixels (e.g. 1920×1080 pixels per eye) is borrowed from the practice of providing computer monitors specifications. However, the pixel density, usually specified in pixels per degree or in arc-minutes per pixel, is also used to specify visual acuity. 60 pixels/° is usually referred to as eye limiting resolution in the central part of the fovea, above which increased resolution is not noticed by people with normal vision. HMDs typically offer 10 to 20 pixels/°, though advances in micro-displays may help increase this number. While more visual acuity is desirable, FOV and visual acuity (VA) in an HMD are linked by the relationships:

$$VA[pixel/°] = \frac{V[pixel]}{FOV[°]}$$

thus, given a certain image source increasing the FOV also reduces the VA

$$FOV[°] = \tan^{-1}\left(\frac{H[m]}{F[m]}\right)$$

The focal length of the collimating lens (F) determines the relationship between H, the size of the image source and the field of view. Some vendors employ multiple micro-displays to increase total resolution and field of view.

Binocular Overlap. Binocular overlap measures the horizontal field of view that is common to both eyes. This allows the device to create a ‘stereo zone’ where the binocular disparity is provided. Overlap is usually specified in terms of degrees or as a percentage indicating how much of the visual field of each eye is common to the other eye.

Weight. It is highly desirable that the device weights only as little as possible to allow for long usage sessions. Also, the device should be well balances and possibly adjustable to the user’s head geometry. Many reports suggest that past generations devices were very far from being ergonomic.

Size. Sometimes, even if the weight of the device is acceptable, its size can still be an issue. Thus, the design should be as compact as possible except for those applications where the HMD is coupled with a safety or military helmet and some parts of it can be consolidated into that.

Power Consumption. Is the device plugged to a power source? Does it need wired connection for video stream between the optical compartment and the external image generator? If so, power consumption is not a main issue. On the contrary, if the device is completely wireless, which allows for the maximum movement flexibility, a careful design of the image source, sensory components and computing components is needed in order not to carry around too much battery weight. For instance, some image sources, such as CRTs (Cathode Ray Tubes) displays and AMELs (Active Matrix Electroluminescent) displays have considerable power consumption.

Addressability. Raster scan displays are considered infinitely addressable because the imagery is drawn in calligraphic fashion. Pixilated devices such as LCDs (Liquid Crystal Displays), AMELs (Active Matrix Electroluminescent) and OLEDs (Organic Light Emitting Diode) are considered finite addressable displays because the pixel location is fixed. This limits their ability to compensate for optical distortions induced by the optical compartment.

Aspect Ratio. Most miniature CRTs (Cathode Ray Tube) have a circular format, while most of the solid-state pixilated devices such as LCDs and AMELs and OLEDs have a rectangular form factor. For instance, Full HD resolution displays have a 16:9 aspect ratio. This parameter contributes to determine the field of view of the display and the binocular overlap.

Luminance and Contrast. It is important that the image source can provide a display luminance and contrast that is compatible with the ambient backgrounds brightness. Literature proves that text and symbols readability is a major issue in see-through augmented reality displays, and is also influenced by colours and styles [82], [83]. Head

Mounted Displays have been proven to have limited luminance and contrast, especially in bright daylight condition [6], [13], [14]. However, this is expected to be a minor issue with current and future devices.

Colour. Is the image source capable of producing colour imagery? Color-coding has proven to be useful in many situations, however a limited colour spectrum might be sufficient for some AR applications.

Image sources. There are four main categories of image sources (a) transmissive displays, (b) reflective displays (c) emissive displays and (d) scanning displays.

The non-emissive technologies, namely (a) and (b), modulate a separate illumination on a pixel-by-pixel basis to create the desired imagery.

a) **Transmissive** displays use a backlight to illuminate an active matrix of pixels. A modulated electric field controls the transmission of the backlight through the individual pixels (or RGB subpixels). The most common application of such technology is the one of Active Matrix Liquid Crystal Displays (or simply LCD)

b) **Reflective** displays use a backlight to illuminate an active matrix of pixels or micro-mirrors. A modulated electric field controls the reflection of the front light against the individual pixels, RGB subpixels or micro-mirrors. Examples of reflective displays are:

    - **Digital Light Processing (DLP)** displays or projectors. This technology uses tiny mirrors, one for each pixel, to reflect the (coloured) front light either away or into the optical path. Rapidly toggling the mirror between these two orientations produces grayscales, controlled by the ratio of on-time to off-time. Because the driving electronics is placed under the micro-mirrors instead of at their side, DLP technology typically results in a good fill factor, which leads to a reduced screen-door effect.

    - **Reflective Liquid Crystal on Silicon Displays (LCoS)** displays. This is a hybrid technology that combines the idea of LCD and DLP. In LCoS, liquid crystals are applied to a reflective mirror substrate. As the liquid crystals open and close, the light is either reflected from the mirror below, or scattered. This modulates the light and creates the image.

Emissive and scanning technologies, namely (c) and (d), emit light without the need for additional illumination.

a) **Emissive displays** are based on an active pixel matrix where individual pixels or RGB subpixels are turned on/off or partially on. Examples of emissive displays are:

- **Active Matrix Electroluminescent (AMEL)** displays. A thin-film layer of luminescent phosphor is sandwiched between two electrodes, one transparent, in a pixilated array. The subpixels are digitally addressed using high-frequency pulses to achieve grayscale. Recent improvements use a quasi-analog signal to achieve greater grayscale range and improved luminance. These are compact and very rugged devices [84].

- **Vacuum Fluorescent Displays (VFDs)**. VFDs use a vacuum package containing phosphors that are excited by a series of filaments. Currently they are only used as alphanumeric, low-resolution displays. However, their capabilities could be expanded to accommodate higher resolution images.

- **Organic Light Emitting Diodes (OLED)** displays. A thin layer of organic semiconductor material is placed between two electrodes emitting visible light in response to an electric current. It has been demonstrated that this technology can be used to produce very thin, curved and flexible displays.

a) **Scanning displays** do not rely on a pixel matrix to spatially build up the image, but rather on a raster scan path that creates the image in calligraphic fashion. In this sense scanning technologies are both time and space multiplexed. Examples of scanning displays are:

- **Retinal Scanning Display (RSD)**. A light beam (such as a laser) or a line of point sources (such as LEDs) is modulated in space and time using resonance scanners or optoacoustic modulators to produce imagery [85], [86].

- **Cathode Ray Tube (CRT)** displays. These are vacuum tubes with one or more electron-guns at one end and a RGB phosphor screen at the other. The beams from the electron guns are modulated by deflection grids and directed onto the screen. The incident electrons excite the phosphor, emitting visible light [87].

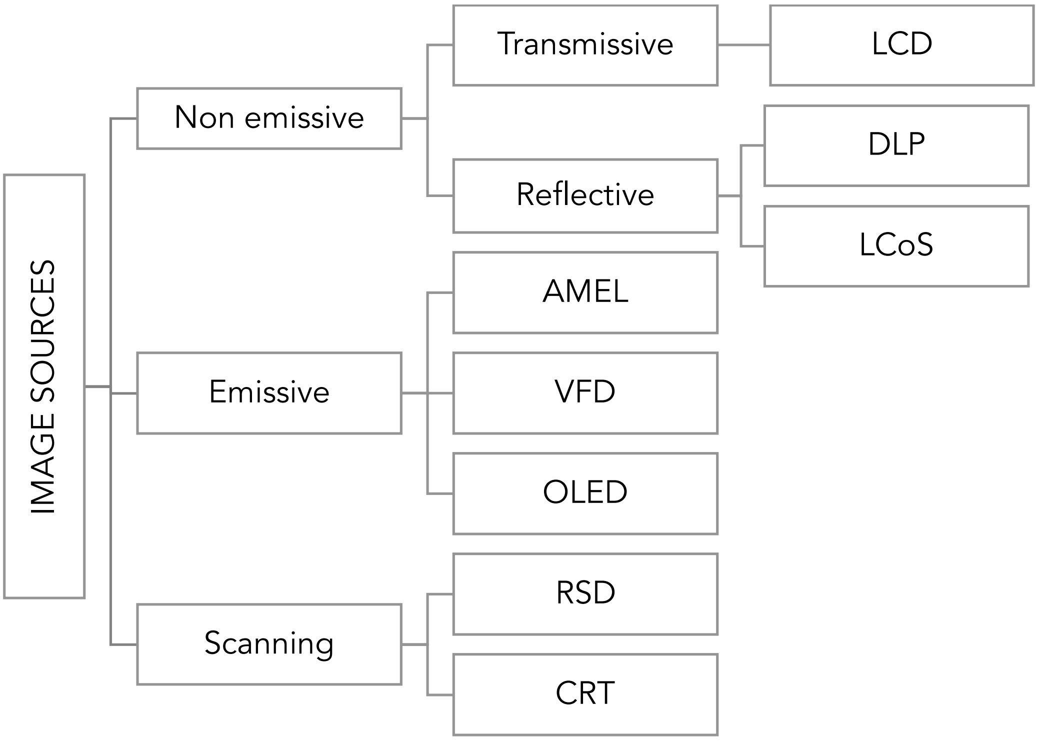


Figure 18. Classification of common head-mounted displays image sources

In the past many HMDs used to have CRT image sources. As of this writing, most HMDs use either LCDs or OLED displays though there is a strong interest in developing new technologies that can further reduce the weight and size of the image source generators.

CATEGORY	TRANSMISSIVE	REFLECTIVE	SELF-EMISSIVE	SCANNING
BENEFITS	LARGELY AVAILABLE ON THE MARKET GOOD IMAGE QUALITY. CHEAP	GOOD CONTRAST FAIR PRICE GOOD FILL FACTOR (DLP)	LOW POWER CONSUMPTION FLEXIBLE THIN LIGHTWEIGHT GOOD RESPONSE TIME	GOOD LUMINANCE AND SATURATION RUGGERIZED DESIGN POSSIBLE EASY DISTORTION HANDLING
DRAWBACKS	POWER CONSUMING LIMITED RESPONSE TIME	HEAT GENERATION POWER CONSUMING	EXPENSIVE LIMITED LUMINANCE	LARGE HEAVY POWER CONSUMING

Table II. Benefits and drawbacks of common head-mounted displays image sources

3.4.1.2 MATURITY LEVEL

Although successful applications can be found in some fields, such as the military, HMD still have a far way to go before they are comfortable enough to be worn by any individual and used for many tasks over extended periods of time. Features such as FOV, resolution, weight and size have improved since their early adoption and are expected to improve even more in the future.

At the time being most commercial see-through HMDs use LCD, OLED or LCOS image sources with fixed focal length optics. However, for eyeglasses-form-factor see-through HMDs, the two solutions that appear to have the highest potential are: (i) the under-explored eye-tracked varifocal optics with liquid crystal lenses and (ii) eye-tracked multiscopic displays. Freeform waveguide stacks with more than two focal planes are another under-explored area [65].

We anticipate that combinations of recent advancements, such as freeform waveguides, micro-lens arrays, DLP mirrors, pinlight displays, eye tracking, pupil tracking, liquid crystal lenses and LCoS displays, will yield much lighter, more ergonomic designs with greater resolution soon, and will also greatly alleviate, if not eliminate, side-effects of the vergence-accommodation conflict.

One of the greatest challenges in HMDs development is their optimization with respect to the user's tasks and needs within the working environment. A good optimization already has proved to be sufficient for the successful integration of such devices in some areas (although the technology was not perfect).

3.4.1.3 BENEFITS AND DRAWBACKS

The main benefit of a HMD is that it is a personal device that follows the user around. In this sense, customized imagery can be shown to each user with a visual efficacy that is irrespective of his or her position. Thus, the application can adapt easily to the context of the user. This has the advantage that the controller will not be distracted with irrelevant information. Also, this will not impair the view of other users which is again an added value.

HMDs can have an ‘unlimited’ field-of-regard, meaning that they can superimpose information on any point of the panoramic view, provided that the viewer is looking in that direction.

HMDs provide multicolour display capability and a colour display is beneficial for organizing and highlighting information using colour codes.

HMDs do not have to be held in the hand or manipulated.

The main drawback of such devices is that, for the time being, they can be rather physically or psychologically cumbersome to wear for extended periods of time.

HMDs could have negative impact on teamwork and communications, because they provoke isolation for each operator that has his or her own situation view.

Same as many other augmented reality devices; HMDs are affected by latency, which means a computer-generated image is lagged with respect to the changes of the background reality. This is caused by the communication time between the image processor, the head movement tracker and the display, as well as by the processing time taken by the AR overlay generation. This represents a bottleneck for future applications unless it can be reduced to an acceptable level.

Clutter is another issue for HMD, which can be reduced taking advantage of the unlimited field of regard.

A preeminent cognitive factor in assessing controllers' performance with HMDs is the phenomenon named "cognitive capture." This effect describes the degradation of responses to external targets due to the processing of information from a HMD image; as such, it principally involves the cognitive operations of selective attention, divided attention, and attention switching. This happens when visual information conveyed via HMDs and visual information from the external driving scene are not processed on separate channels; in other words, it is impossible to process both sources of visual information simultaneously.

Test subjects reported a reduction of peripheral vision when using a HMD [88] . This issue seems to be less relevant with current HMDs designs and hopefully will continue to shrink in the future.

3.4.1.3.1 THE VERGENCE-ACCOMMODATION CONFLICT IN HUDS

The VAC is generally unsolved in modern-day commercial HMDs, contributing to the discomfort, especially for close range tasks. This is because the virtual image is typically focused at a fixed depth, while the depth of the virtual objects, hence the binocular disparity, varies with the content, which ultimately results in conflicting information within the vergence-accommodation feedback loops [47], [89]. Researchers have theorized about potential solutions to the VAC and built prototypes since early 1990s. Since the convergence cue is properly-configured in eye-tracked stereo displays[10], but the accommodation is not, most the scientific effort gears towards adjusting the retinal blur cue to the depth of the in-focus virtual object. This can be done by means of (a) multifocal displays, (b) varifocal displays and (c) multiscopic displays. While (a) and (b) still rely on a stereoscopic virtual camera set up for the image generation, (c) use a different approach.

a) **Varifocal** displays involve adjustable optics which can modify the focal depth of the entire view. Many varifocal display prototypes were built as a proof-of-concept, which could display only simplistic images, such as simple line patterns or wireframe primitives. These either forced the focus information to correspond to the vergence at a single object, or provided some manual input capability to the user to manipulate the coordinate of the focal point, which in turn would tell the system which object to bring into focus. In an effort to improve varifocal designs it has been theorized that the focus of the adjustable optics can also be gaze-driven [90], [91]. In this model the focus of the optics will adapt to the depth of the virtual point where the viewer is looking at any given moment. Authors of several works hypothesized about integrating an eye tracker into an HMD to accomplish this. In fact, some work has been done specifically on designing eye-tracked HMDs (ET-HMDs) in a compact and ergonomic way [91]. So far, several studies have used eye-trackers in conjunction with emulated (i.e. software-rendered) retinal blur, investigating the effects on accommodation. However, to our knowledge, no eye-tracker-driven varifocal design has yet been published. Although gaze-driven emulated blur has been shown to contribute to visual comfort, it was demonstrated, both theoretically and experimentally, that it is incapable of driving accommodation. Indeed, the light rays coming from a display positioned at a given depth still diverge at the same angle before reaching the eye lens [92]–[94].

b) **Multifocal** displays split the view for each eye into depth regions and display each region at a separate, fixed, focal depth, thus emulating a volumetric light field in a discrete fashion. Several multifocal designs with physical stacks of displays were conceived just before the turn of the century, whereas up to now only one space and time multiplexed design exists [47]. Requirements for focal plane stacks have been evaluated based on the criteria of how closely the accommodation response resembles actual live viewing , but fatigue levels haven't been measured for designs that don't adhere to the criteria [47].

---

[10] but not entirely, due to offset between virtual camera and pupil, which should be compensated in HMDs by means of pupil tracking.

In some cases, Multifocal and Varifocal the techniques can be combined [65].

a) **Multiscopic** displays follow the principal of integral imaging, i.e. they try to emulate a contiguous light field within the eye. In doing so, these techniques usually require more complex rendering set-ups than (a) and (b), with several virtual cameras shooting slightly-different viewpoints of the scene and some post-processing going on. The only time-multiplexed multiscopic HMD design known to date relies on a rotating galvanometer scanner and a digital micro-mirror display to generate the needed rays. In contrast, the spatially-multiplexed multiscopic designs use a fine array (or layers of arrays) of microscopic optical elements, such as spatial light modulators, micro-lenses, and/or point light sources ('pinlights').

A fine review of varifocal, multifocal and multiscopic techniques can be found in [65].

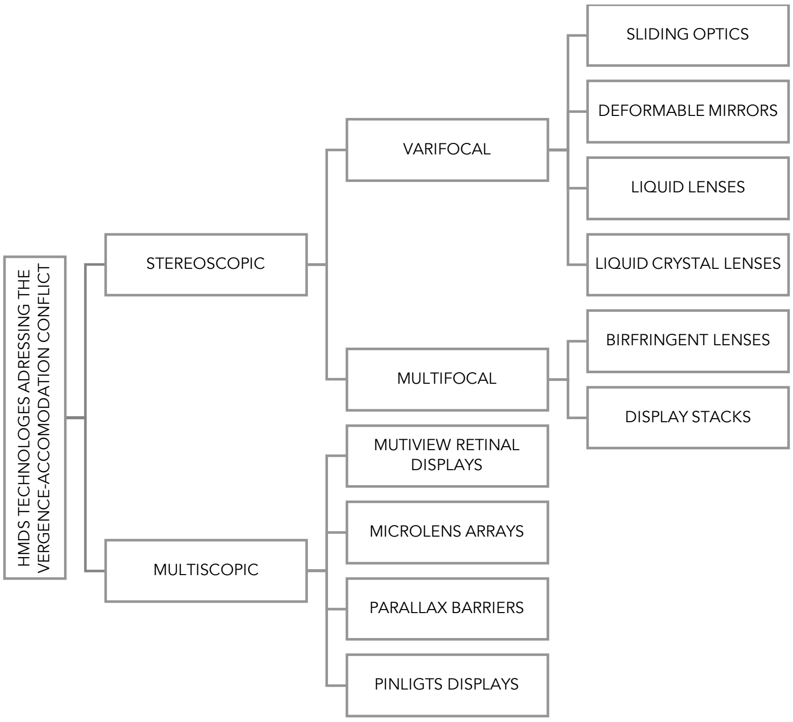


Figure 19. Classification of head mounted display technologies adressing the vergence-accommodation conflict

3.4.2 HAND-HELD DISPLAY

Hand-held AR devices mainly consist of:

a) **Displays that are often embedded within consumer devices**, namely Tablet PCs, PDAs (personal digital assistant), or smartphones, working as video see-through displays. Alternative solutions based on optical see-through hand-held displays are diffused to a lesser extent (Figure 20).

b) **Hand-held video-projectors** which are projector-based systems that depict the synthetic information on the real object by directly projecting it on the object surface.

Figure 20. TYPICAL HAND-HELD DISPLAYS CONFIGURATIONS

3.4.2.1 SPECIFICATIONS

Many of the main features of hand-held AR devices are the same as the HMD devices. To show the differences and similarities, the same structure to describe the features is used.

- **Field of View**. The Field of View in hand-held AR is determined by the device's camera lens. Most of these devices have fixed field lenses, so the field of view does not change as the device is held closer or farther away from the user's eyes. This can cause some vision discomfort if the screen displays a field of view that does not correspond with what has been blocked by the device itself.

- **Resolution.** Phone and tablet producers usually mention the total number of pixels of the display or the number of pixels per degree. Most tablet displays have a pixel density on the order of 350 pixels/inch and an aspect ratio of either 16:9

or 4:3. Smartphone displays have a wider range of pixel densities and aspect ratios, due to the wide variety of screen sizes.

- **Stereoscopic view**. Hand-held displays rarely have the capability to display images in 3D. When they do, the use the parallax barrier technique to filter the left and right image for each eye.

- **Weight and Size**. The lighter weight devices are those, in general, that have a smaller screen. The weight must be balanced with the display size to find an optimal combination of the two.

- **Power Consumption**. Is the device plugged to a power source? Does it need wired connection for video stream between the optical compartment and the external image generator? If so, power consumption is not a main issue. On the contrary, if the device is completely wireless, which allows for the maximum movement flexibility, a careful design of the image source, sensory components and computing components is needed in order not to carry around too much battery weight.

- **Luminance and Contrast**. Tablet and smartphone display luminance and contrast are, in general, compatible with the ambient backgrounds brightness that could be found in a control tower. Holographic displays, due to their transparency, are limited in their luminance and contrast and can often have a "ghostly" aspect to their image.

- **Colour**. The cameras in tablets and smartphones reproduce a full 16M colour range as well as the displays. There are other components that contribute to the overall quality of a screen. Black levels, and colour accuracy are also equally important factors to consider.

- **Image sources**. Tablets and smartphones are emissive displays and most of the newer ones are OLED displays.

3.4.2.2 MATURITY LEVEL

Both tablets and smartphone are fully mature devices, but are continuously being advanced with new technology.

3.4.2.3 BENEFITS AND DRAWBACKS

A benefit of the tablet/smartphone technology is that they are relatively inexpensive and many AR applications are already being developed for these platforms, albeit not for the tower control environment.

One of the drawbacks is that the user has at least one hand occupied. For a tower controller, this can become an inconvenience and a limiting factor to the use of this type of technology. Also, physical fatigue must be considered.

Another drawback is that the screen occupies a small part of the viewing space. This can become a problem with maintaining situational awareness.

3.4.3 SPATIAL DISPLAYS

In contrast to body-attached displays (i.e. head-attached or hand-held), spatial displays detach most of the technology from the user and integrate it into the environment. As for HMDs, spatial displays can be classified based on the type of content they provide. Thus, video see-through, optical see-through and fish-tank VR displays exist. It is also useful to distinguish between monocular, biocular and binocular spatial displays.

Binocular spatial displays are the ones providing binocular disparity in the rendered imagery. When such displays are used in AR, they can provide more immersion, since the real and virtual disparity cues are made to coincide. Thus, the graphics seem to spatially co-exist with the real objects in the physical environment. Image pairs (a.k.a. stereo pairs) are encoded and filtered so that each single image is only seen by the matching eye. Encoding techniques include colour spectrum decomposition, light polarization, temporal encoding and spatial encoding, as further detailed below. Filtering is most easily attained through special equipment, e.g. polarized eyeglasses, coloured eyeglasses and shutter glasses, but might be also achieved by looking at the screen from a pre-defined position.

Monocular and biocular displays, can adequately display heads-up, non-registered graphics in far-field (panoramic) applications, where the graphic content is placed beyond the range of binocular depth cues [95]. Also, a stereo imagery may not be required if only the surface properties (e.g., colour, illumination, or texture) of the real objects are changed by overlaying images. In this case, a correct depth perception is still provided by the physical depth of the objects [75].

Spatial displays can be further divided into desktop (self-emitting) configurations and projection displays. Using desktop monitors as a possible stereoscopic display is the traditional desktop-VR approach. Desktop VR setups are classified as non-immersive since, in contrast to large screens, the degree of immersion is low. Horizontal, workbench-like or vertical wall-like display screens are currently the most common embedded screen displays [96]–[98]. Projection displays currently use CRT, LCD, LCOS, or DLP to beam the stereo images onto single or multiple surfaces, which can be planar (e.g., CAVEs[99]–[101] CABINs [102]) or curve (e.g., domes or panoramic displays [74], [103]. Two types of projections exist: front-projection, where the projectors are located on the same side of the display surface as the observer and rear-projection (or back-projection), the projectors are located on the opposite side of the display surface Thus. When using front projection, the observer might interfere with the projection frustum and cast a shadow onto the display surface. To avoid this problem rear projection is used.

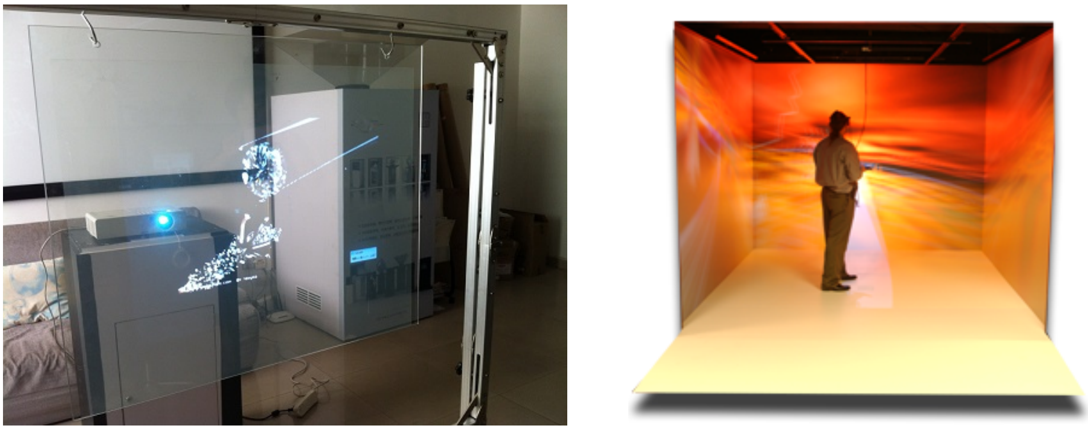


Figure 21. See-through spatial augmented reality display (left image) and the Cave Automatic Virtual Environment^{TM} (right image)

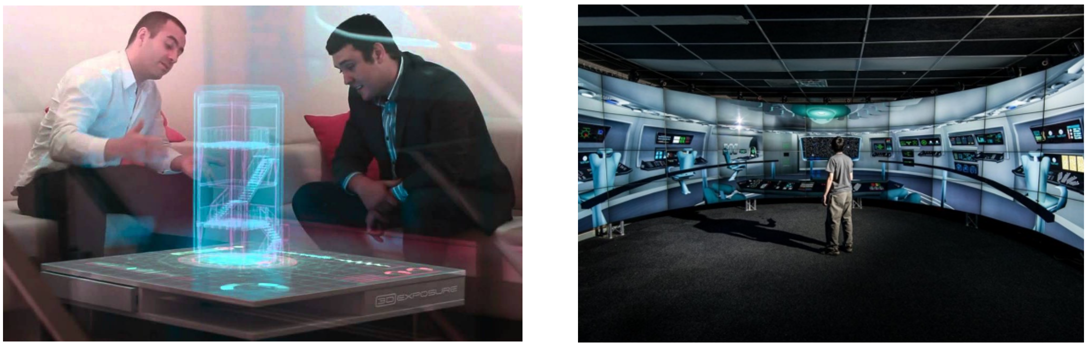


Figure 22. Fish-tank virtual reality table top (left image) and the CAVE^{TM} 2 virtual reality system (right image)

In aeronautics, the use of spatial AR displays goes back to the reflector sight, developed in 1900, and used on fighter aircraft in World War 1. In 1942, the Royal Air Force combined the image from an on-board radar tube with the projection from the gunsight onto a flat area of the windscreen. A key upgrade included an artificial horizon [21] The modern HUD used in instrument flight rule approaches to landing was developed in 1975. HUDs are currently prevalent options on both commercial and private passenger aircraft, and have become standard equipment on Boeing 787s.

3.4.3.1 SPECIFICATIONS

- **Field of View**. The field of view of a spatial display is typically limited and depends on the relative position between the user and the screen. Power-walls and curved Mosaics are the most straightforward way to extend the FOV without decreasing spatial resolution – i.e. PPI (Pixel Per Inch) resolution. These are made of many displays tied together with the only propose to form a larger, wide-view, giant display. The result, either planar or curved, looks as if one was looking through a framed window. Therefore, appropriate bezel correction is used. Sometimes screen edges can be made to overlap with physical frames, such as in car and flight simulators, which is very convenient. A second possibility is to aim for Cave Automatic Virtual Environments (CAVEs), which are well known VEs that fill in the peripheral vision by means of multiple, rear-projected, flat screens. Three to six screens are typically arranged in a 'cubical' configuration, although unconventional architectures, such as the ones in Ref. 5 or Ref. 6, may be used as well. F-V/AR is used for rendering purposes, while edge blending might be needed for a seamless result.

- **Resolution**. High resolution is not an issue for modern spatial displays. Even in consumer applications they can easily reach a 4096×2160 pixels (4K) resolution.

- **Stereo capabilities**. Speaking about spatial displays the concept of conveying two different images to the users' eyes is commonly known as multiplexing [64], [70]. Several techniques can be used to achieve the result:

  1. **Colour Multiplexing**: Anaglyph and Infitec^{TM} Displays. Colour multiplexing techniques use colour filters to separate the left and right eye views, which are rendered simultaneously on a single surface. The user wears a pair of glasses, where each eyepiece accepts a different part of the colour spectrum. The red-blue anaglyph glasses are well-known examples, but newer approaches such as Barco^{TM} Infitec^{TM} subdivide the colour spectrum into finer slices so that each eye view receives apparently similar colour content. The main advantage of this technique is that only lightweight plastic glasses must be worn by the user, however, the colour distortion that this technique introduces can easily make it unsuitable for several V/AR applications.

  2. **Polarization Multiplexing**: Passive Stereo Displays. Polarization multiplexing systems provide two separate images by filtering the light in a polarized way. The user wears a pair of polarized glasses, with corresponding filters. Thus, each eye receives only the light that is let pass by its filter. This type of stereoscopic viewing, with no active parts in the glasses, is commonly referred

to as passive stereo. The main advantage of this technique is that only lightweight plastic glasses must be worn by the user, however, if only one degree of polarization is used (either linear or circular) there is no way of generating more than one perspective. Another drawback of this technique is the loss of contrast and brightness due to polarization.

3. **Time Multiplexing**: Active Stereo Displays. In time-multiplexed systems both left and right eye views are rendered sequentially on a single display surface and transmitted towards the user. The user wears glasses, commonly known as shutter glasses, with a liquid crystal shutter mechanism which is synchronized with the display and continuously blocks (shuts) the incorrect eye view. The main disadvantage for use in AR settings is that the principle of repeatedly blocking the view filters out most of the incident light, and as a result the real-world scene becomes very faint.

4. **Spatial Multiplexing**: Auto-stereoscopic Displays. Another approach to stereoscopic viewing is called spatial multiplexing. Each eye is only provided with the corresponding image by means optical systems. Auto- stereoscopic techniques, such as Parallax Barrier and Lenticular Lens, do not require additional filtering equipment because they encode spatially. In this case, it is the physical distance between the viewer's eyes that filters the images. However, the number of viewing positions is limited: if the viewer's eyes move outside the pre-defined positions (sometimes referred as 'vertical slits), the 3D effect will disappear. If users were tracked and light paths could be (dynamically) adjusted for each pixel, the system would theoretically provide an arbitrary number of viewports and viewing positions.

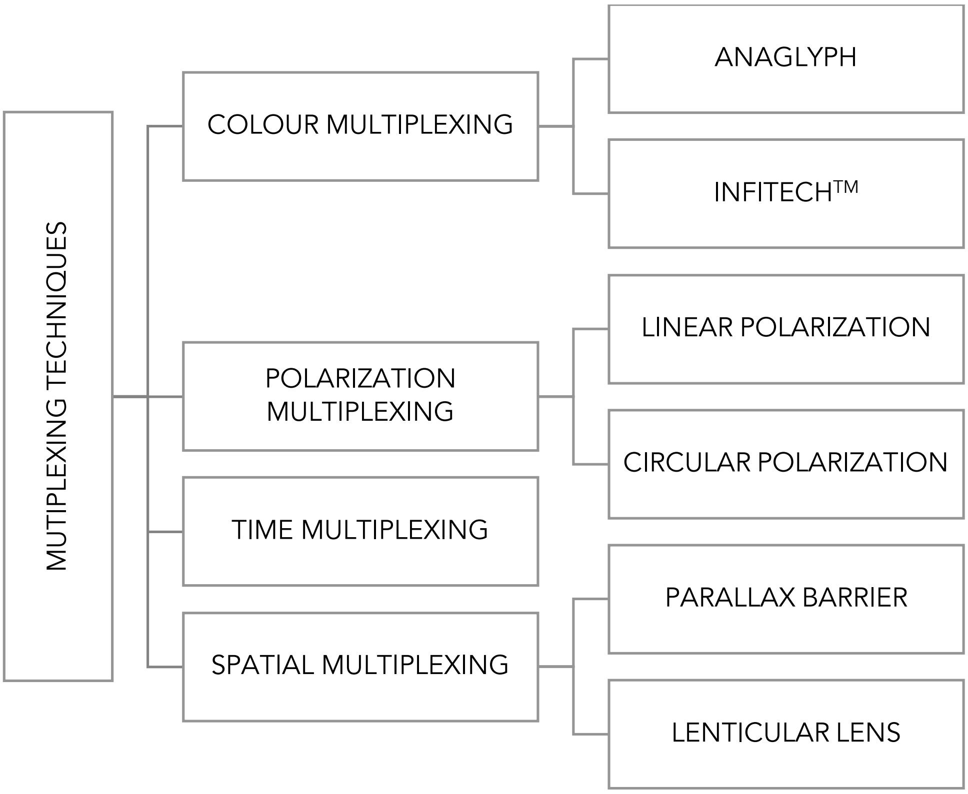


Figure 23. TAXONOMY OF MULTIPLEXING TECHNIQUES

3.4.3.2 MATURITY LEVEL

The maturity level of spatial A/VR displays is high if one does not consider the need for additional equipment to be worn by the user. However, the number of observers that can be supported simultaneously is restricted by the applied optics, which often translates to a single user scenario. A higher maturity level for this technology will be reached if eye-tracked autostereoscopic displays will be released.

3.4.3.3 BENEFITS AND DRAWBACKS

On one hand, spatial displays overcome some of the shortcomings that are related to body-attached displays: an improved ergonomics, a theoretically unlimited field of view and a scalable resolution. On the other hand, they are inclined to many typical V/AR issues such as the frame effect (i.e. virtual objects outside the display area are unnaturally cropped) and the vergence-accommodation conflict (i.e. the coupling of accommodation and vergence is generally not provided in spatial display devices).

There is an increased complexity of maintaining consistent alignment and colour calibration as the number of applied displays increases. Also, there is a risk of showing irrelevant or overlapping information to multiple controllers.

3.4.3.3.1 SYNTHETIC VISION

Synthetic Vision is as a specific application of virtual and augmented reality technologies that is widely used in aeronautics, particularly in aircraft's cockpits and Remotely Piloted Aerial Systems' (RPAS) Ground Control Stations (GCSs). It consists in a real-time, computer-generated image of the topography surrounding the aircraft. This disregard atmospheric occlusion, synthesizing multi-source information into a single, clear and accurate three-dimensional representation of the external environment. Thanks to SV, the pilot maintains excellent ground and airborne situational awareness, even when flying remotely or in adverse conditions – e.g. reduced visibility or darkness. In practice, SV allows the pilot to see through haze, clouds, fog, rain, snow, dust and smoke, while displaying the vehicle's position with respect to the terrain. Advanced SV systems also integrate urban features, obstacles and other significant information, such as flight hazards, flight paths, waypoints, aerodromes, landing points, surrounding facilities (friendly, neutral or hostile), nearby airspace users, runway incursions, taxi navigation and surface guidance maps. These data may be shown on head-down, head-up, helmet-mounted or navigation displays. Because SV is completely artificial, aircraft operations can be monitored from either a 'pilot' perspective (egocentric) or an 'out-of-the-cockpit' perspective (exocentric). The latter, is most commonly used in RPAS GCSs.

The first synthetic vision device was a night vision device developed in the 1930's by Vladimir K. Zworykin for the Radio Corporation of America [14] and was intended for civilian use. Although it didn't achieve commercial success, in 1935 the idea was used by the Air Expeditionary Group for military purposes. As part of advanced cockpit research, NASA and the U.S. Air Force started developing synthetic vision systems in the late 70's to improve situational awareness. In 1993, Loral WDL, with sponsorship from STRICOM, performed the first demonstration combining live AR-equipped vehicles and manned simulators. In 2001 a NASA X-38 was flown using LandForm software video map overlays at the Dryden Flight Research Center, and in 2009, the first FAA certified application of a synthetic vision system was available as part of the Gulfstream PlaneView flight deck [22].

Figure 24. SYNTETIC VISION EQUIPMENT INTEGRATED INTO THE GULFSTREAM PLANEVIEW FLIGHT DECK.

Synthetic Vision devices are application-oriented systems where data coming from different sources is filtered and fused providing the pilot with a comprehensive view of the flying environment in poor visibility conditions. Based on the type of data that is used to reconstruct the external view and the mean used for visualization, Synthetic Vision devices can be classified into three main categories:

1. **Enhanced Vision Systems (EVS) and Enhanced Flight Vision Systems (EFVS)**. An EVS or EFVS is an electronic means to provide a display of the external scene by use of an imaging sensor, such as a Forward-Looking InfraRed (FLIR) or millimetre wave radar. It provides pilots with a clear live video image of the world that s/he could not otherwise see at night, and in poor visibility. As far as technology is concerned, the main difference between EVS and EFVS consists in the alignment of additional information with the external view and the use of head-up displays to show them that are essential features for EFVS.

2. **Synthetic Vision Systems (SVS).** SVS provide situational awareness by placing a 3D geographical image on a cockpit display using terrain, obstacle and other databases. Navigation and positional information is obtained from GPS and Inertial Reference Systems. SVS presents a "clear day" view of the world, but is only as good as the most recent update to the database which can be days, weeks, or even months old.

3. **Combined Vision Systems (CVS) and Verified Combined Vision Systems (VCVS).** CVS is a term applied to the combination of EVS and SVS whereby EVS is used to provide a real-time confirmation (validation) of the SVS environment. In CVS, the pilot is doing the comparison and alignment of the two systems. An evolution of CVS is represented by Verified Combined Vision Systems (VCVS) that perform a smart processing to verify and correct GPS positional error (if any), automatically resolve differences between SVS and EVS and align the images

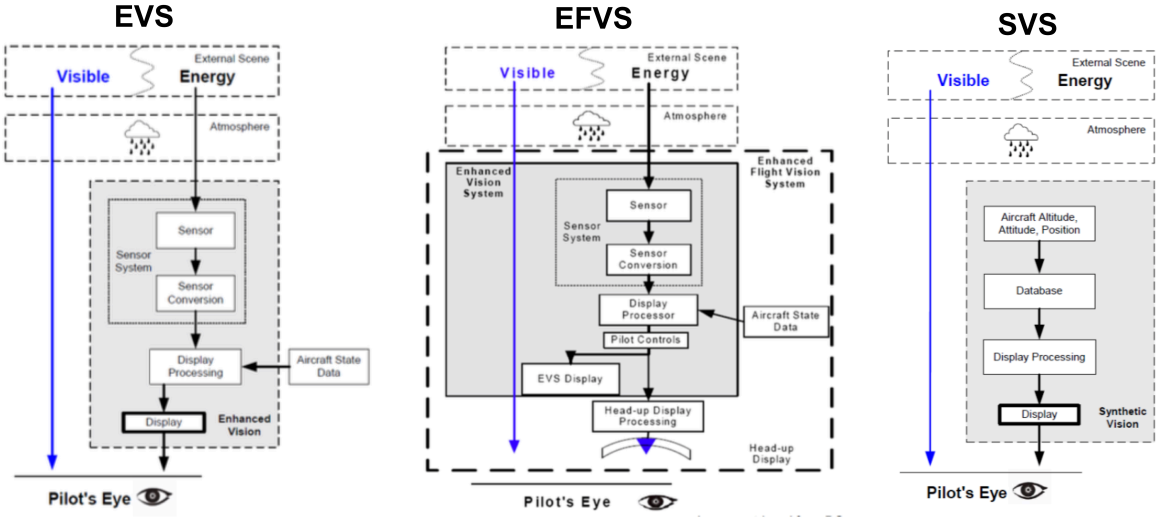


Figure 25. COMPARISON OF EVS, EFVS AND SVS[11]

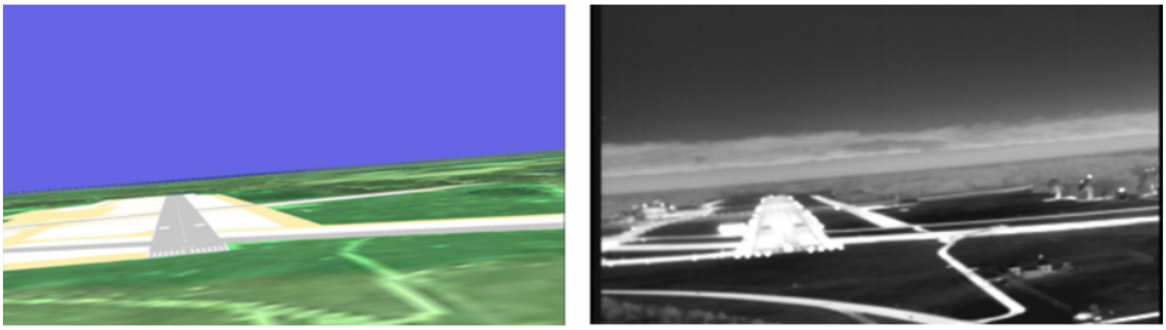


Figure 26. COMPARISON OF SVS (LEFT IMGAE) AND EVS (RIGHT IMAGE)

---

[11] Image: Honeywell™

3.4.4 OBJECT PROJECTED DISPLAYS

Object-projected displays are the ones where the imagery is directly projected on real world objects. In this sense, the object itself becomes the canvas of the V/AR image generator. As already stated, the light source (alias the projector) can be attached to the user's head, held within the hand or positioned in space.

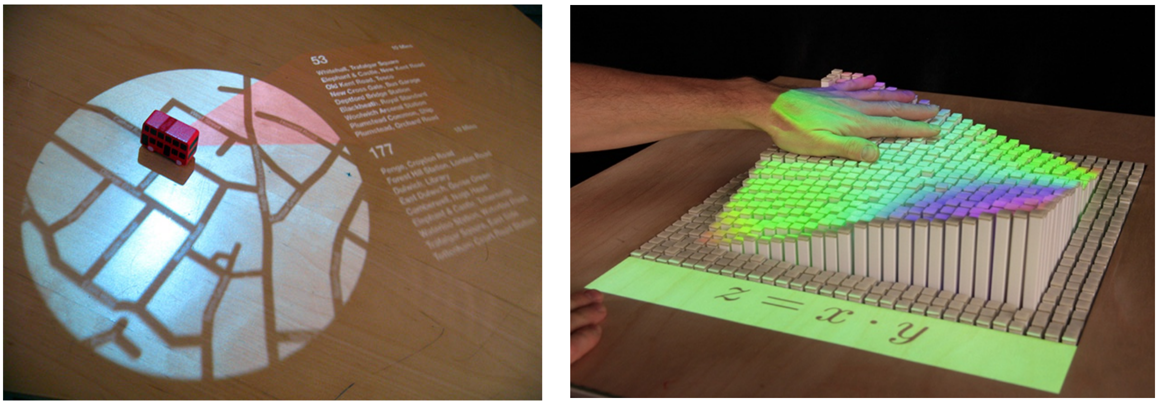


Figure 27. EXAMPLES OF OBJECT-PROJECTED DISPLAYS

3.4.4.1 SPECIFICATIONS

The features of an object projected display cannot be discussed *a priori* as they depend on the projector used, the projection surface and the distance between the two.

3.4.4.2 MATURITY LEVEL

The maturity level of such technology is low. In fact, it is difficult to set up a projection system which can handle at the same time different types of objects and V/AR contents. Therefore, the risk of having to set up a very customized configuration is very high. Also, for hand-held and head-attached object-projected displays, the current hardware may not be miniaturized enough.

3.4.4.3 BENEFITS AND DRAWBACKS

The main benefit of the object-projected technology is a high level of integration with the viewer's tasks within the working environment. This feature makes it perfect for close range and manual applications such as AR maintenance, assembly and installations, as well as for some video)-ludic applications. However, the display area is constrained to the size, shape, and colour of the physical objects' surfaces (for example, no graphics can be displayed beside the objects' surfaces if no projection surface is present) and limited by the capabilities of the projection system. Also, there is no standard procedure for the generation of the AR content. All in all, this technology does not seem to fit complex and far-range applications such as the provision the ATC service by the control tower.

3.4.5 VOLUMETRIC DISPLAYS

Volumetric displays create 3D imagery via the emission, scattering, or relaying of illumination from well-defined regions in (x, y, z) space. In a broader sense, holographic and highly multi-view displays can be considered volumetric displays if they do a reasonable job of projecting a three-dimensional light field within a volume.

There is a volumetric display technology consisting of multiple sandwiched LCDs, where the array of 2D pixel layers defines a larger volume of addressable voxels. This type of display has limited transparency and inability to render at a larger stereoscopic depth than the LCDs themselves.

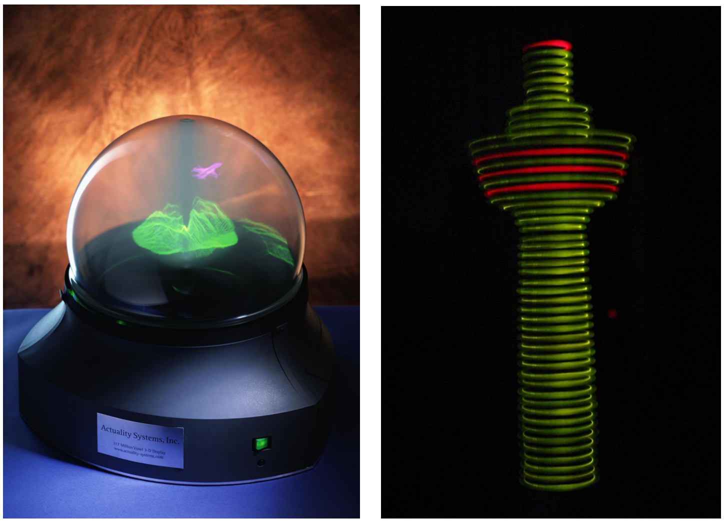


Figure 28. EXAMPLES OF VOLUMETRIC DISPLAYS

3.4.5.1 SPECIFICATIONS

Most information about volumetric displays is proprietary, however it can be safely said that at the current state they are very limited in dimension, resolution and colour spectrum.

3.4.5.2 MATURITY LEVEL

Volumetric displays are still under development, and have yet to reach the general population. With a variety of systems proposed and in use in small quantities – mostly in academia and research labs – volumetric displays remain accessible only to researchers, corporations and the military.

3.4.5.3 BENEFITS AND DRAWBACKS

The benefit of a volumetric display is the ability to see the virtual image in 3D without any additional equipment, as well as to allow more than one user to visualize the data at the same time. The drawback is that the visualization is displayed in a fixed location, usually on a desktop or in a ball like volume.

One other consideration is the amount of bandwidth required. A volumetric display would need to send about three orders of magnitude more information/second to the display hardware to sustain the image. Furthermore, a 3D volumetric display would require two to three orders of magnitude more CPU and/or GPU power beyond that necessary for 2D imagery of equivalent quality, due at least in part to the sheer amount of data that must be created and sent to the display hardware.

4 VIRTUAL AND AUGMENTED REALITY IN AIR TRAFFIC CONTROL

4.1 OVERVIEW

Lloyd Hitchcock introduced the idea of using AR in the control tower in the ‘80s, when the technology was still in the very early stages of its industrialization. At that time, no prototype construction was tried and little was published, though many recall Mr. Hitchcock speculating on several methods that could aid tower controllers fulfilling their tasks [9], [14], [15], [19], [88]. For instance, he suggested that AR displays could provide air traffic controllers with useful status information, such as aircraft identification, barometer settings, wind conditions and runway or gate assignments. More recent studies suggest that other spatially conformal information, such as flight tags, warnings, shapes and layouts can also be presented on AR displays [11], [12], [19], [88], [104]–[109], especially in LVC.

A few years later, several studies speculated about the usefulness of 3D virtual and augmented reality interfaces for Approach control (APP) [24], [110]–[121]. With the current model, the information is embedded in a 2D fashion by means of characters, numbers, colours, shapes, symbols, size and other features [122]–[124]. This practice has been harshly criticised because it takes too long to the ATCO to calculate the spatial-temporal relationship between aircrafts [125], [126]. In fact, this is quite demanding in terms of cognitive resources [125], [126]. Also, controller’s interface as they currently are seem to have reached a threshold in terms of the maximum amount of information that should be made available to the operator at once [127], [128].

Since a substantial amount of the ATCOs work involves three-dimensional problem solving, it has been long suspected that 3D graphics may help controllers to reduce their cognitive workload. If this were true, the operator would be able to manage larger amounts of aircrafts and still reduce fatigue, allowing for longer working sessions to be scheduled [129]. Also, a 3D interface would let designers incorporate additional meteorological and topological information such as 3D weather and terrain orography. However, perception is a delicate matter, which is involved in most of the ATC related accidents [130]. Thus, in order not to jeopardise safety, a three-dimensional interface would need to be carefully designed and validated.

Most recently, the concept of Remote & Virtual Towers (R&VT) was introduced by DLR (*Deutsches Zentrum für Luft- und Raumfahrt e.V.*) and embedded into the SESAR research [9], [106], [131] .

4.2 VIRTUAL AND AUGMENTED REALITY FOR APPROACH CONTROL

Before the end of the ‘90s, the introduction of Graphical User Interfaces (GUIs) made computing technologies accessible to many professional and amateur users, in addition to computer scientists and programmers. This made Human Computer Interaction (HCI) a subject of a more general interest. Ever since then, a great effort has been put in trying to fill the gap between the user and what is going on in the hidden and intangible parts of computers [24]. In the meantime, hardware manufacturers met the requirements for real-time 3D rendering. Very soon, 3D compatible hardware was exploited in many fields, such as simulation, data analysis, computer aided design, engineering, medicine, training, entertainment, cultural heritage and archaeology. By now, thirty years of technological advancement have definitively set the hardware requirements for real time 3D graphics to the level of a regular Personal Computer (PC), which perfectly fits the industry needs. Overall, taking advantage of three-dimensional Computer Graphics (CG) has become less burdensome.

As a matter of fact, Information Visualization (IV) has become a well-established field of study, covering the design of visual systems that enable humans to explore and understand complex data sets [132]. These are sometimes referred as cognitive support systems [132], [133], which improve the viewer’s problem solving capacity by allowing him or her to extract patterns, inspect details, formulate hypotheses and verify theories [134]. This makes it possible to spot relevant correlations across data, making sense of them at a perceptual level and lowering the need for high-level cognitive processing – i.e. the cognitive process involved in thinking, reasoning, planning, and so on (e.g. the one required to interpret numerical and textural representations).

Within this context, several studies – mostly empirical – have been performed to determine whether three-dimensional UIs could be used for the provision of air traffic control services and if there were any point in doing so. Like HCI, the field of UI design is an interdisciplinary subject that draws from existing knowledge in perception, computer graphics and cognitive sciences. A 3D UI may be defined as one in which a sense of depth is created along the line of sight into the display (the effectiveness of this being roughly a weighted additive function of the number of depth cues that the interface integrates [135]).

Generally speaking, 3D graphics is likely to fascinate both end-users and system designers [112]. This is arguably due the ability of easily conveying three-dimensional information [110], "*showing the situation as it really is*" [136] and sparing us (humans) the trouble of collecting and interpreting bi-dimensional information [112].

However, the utility of a graphical interface, either 3D or 2D, cannot be argued *a priori,* nor in absolute terms. Indeed, its pros and cons depend on a number of factors, such as semantics, perception, culture and tasks [24]. Semantics relates to the space representation and the objects' distribution within it. Perception deals with human cognitive abilities. Culture depends on the user familiarity with the interface (including indirect experiences). Tasks relate the user's goals within the current job.

In their review of several human-computer interaction techniques, Andre and Wickens warn system developers about the fact that, occasionally, "*users want what is not best for them*" [137]. In practice, "*they are likely to prefer solutions that hinder rather that hamper performance*" [110]. Therefore, the design of a UI requires a great care and careful optimisation. Non-isolated concerns might include: choice of projection paradigm, choice of depth cues, selection of viewing parameters (e.g. field of view, viewing position, elevation, azimuth and scale), choice of interaction techniques, Human Factors and Ergonomics (HF&E), and further implications for recruitment and training.

4.2.1.1 THE FORESHORTENING EFFECT AND THE HEIGHT CONSTANCY BIAS

One of the major problems associated with 3D is the difficulty of properly estimating distances and dimensions along the line of sight [24], [110], [118]–[120], [138]–[143]. Reliable judgements on the ordinal relationship between the objects are still possible, if enough depth cues are given. However, when it comes to absolute distances estimation, these appear 'distorted' and 'compressed'.

To envisage this behaviour, imagine watching a computer-generated image of a vertical pole that is placed in front of you. Now, let the pole slowly rotate around its centre until you can only see the top of it. The length of the pole will appear shorter and shorter, as if it were being compressed [110]. This effect has been given many names in the literature, including '*foreshortening effect*' [144], '*line of sight ambiguity*' [145] or '*projective ambiguity*' [146]. An example is given in Figure 29. The very same effect exists in bi-dimensional interfaces, where it is clearly impossible to estimate distances along the line of sight. However, both designers and final users are aware of this and do not expect anything different.

Going back to perspective 3D, there are further shortcomings that are worth to be mentioned. For instance, for any object, a change of position in the depth direction will be perceived as a smaller move than an equal amount of lateral or vertical displacement [24]. This is known as the problem of '*display resolution*'. Also, the viewer's ability to correctly judge the height of the objects is compromised [24], [118], [147]. To better understand this, please have a look at You will notice that it is hard to determine which of the aircrafts is flying at the lowest or the highest altitude. This uncertainty is known as the '*height constancy bias*', which is truly a combination of the *foreshortening effect* and the shrink of dimensions along the line of sight. More in depth, the dimensional shrinking *per se* can be calculated mathematically. For instance, if we consider a point of coordinates x_e, y_e and z_e in the *eye-space* coordinate system (a.k.a. *camera-space* coordinate system). For the x coordinate, the result of the shrinking operation is given by:

$$x_p = -\frac{x_e * n}{z_e} \qquad (1)$$

Where n is the distance between the *eye-space* origin (i.e. the field of view origin) and the projection plane, a.k.a. the *near clip plane*. The same calculation applies to the y_p coordinate, but not to the z_p coordinate which is always equal to $-n$.

Figure 29. THE FORESHORTENING EFFECT

In Figure 30, it is hard to determine which of the aircrafts is flying at the lowest or the highest altitude. This is due to the *height constancy bias* which is truly a combination of the *foreshortening effect* and the shrink of dimensions along the line of sight. In truth, all aircrafts are flying at the same altitude. Further, the *height constancy bias* does not only impair decisions on the height of objects, but also prevents reliable judgments on any across-the-line-of-site dimension [148].

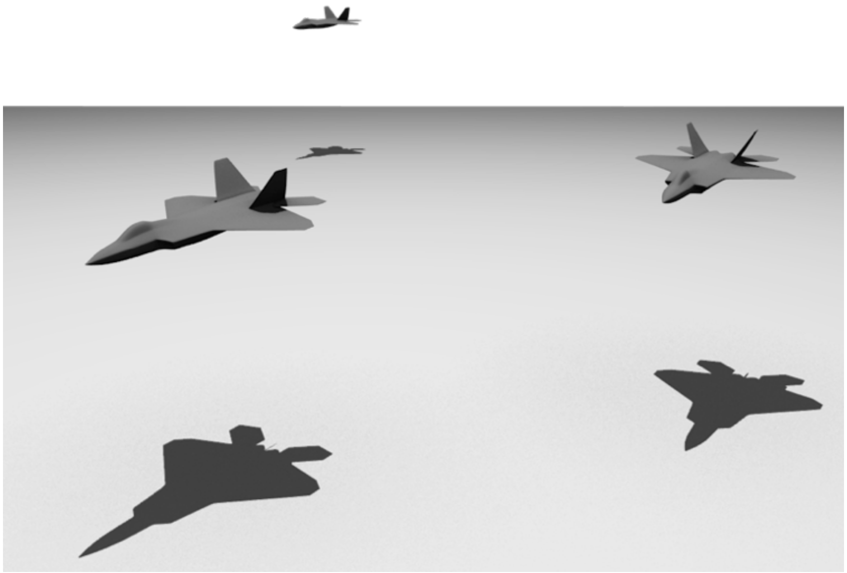


Figure 30. THE HEIGHT CONSTACY BIAS

4.2.1.2 CLUTTERING

Obtaining information from the visual subsystem is often depicted as a stream analysis. Within the analysis, a selective process known as '*attention*' is identified, which can be portrayed as a dynamic allocation of the working memory. In order to prove that visual perception is limited in performance, it has been shown that human beings can hold a maximum of three to seven chunks of information in their working memory while operating under normal workload conditions [149] (chunks are small groups of associated data). When operating outside their comfort zone, humans tend to make mistakes, especially if their attention is split among numerous items on the screen. Because of this, the *attention* tries to capture only the most relevant pieces of information. However, the dynamic changes in objects' size, shape, orientation and behaviour take up large amounts of *attention*, leading to carelessness towards the other areas of the screen [112]. All in all, the amount of information that can be easily perceived and interpreted by human operators is limited. For these reasons, *interface cluttering* (a.k.a. *visual overload*) has always been a major concern for HCI developers [136].

In a perspective 3D interface, when many objects are placed in the distance (i.e. near the far end of the scene), they take up a smaller portion of the screen than if the very same objects were positioned closer to the point of view. This can result in increased clutter. Figure 31 demonstrates this point. In this picture, two aircrafts formations take up different portions of the screen, even if they are truly identical. This results in increased clutter for the 'red' formation.

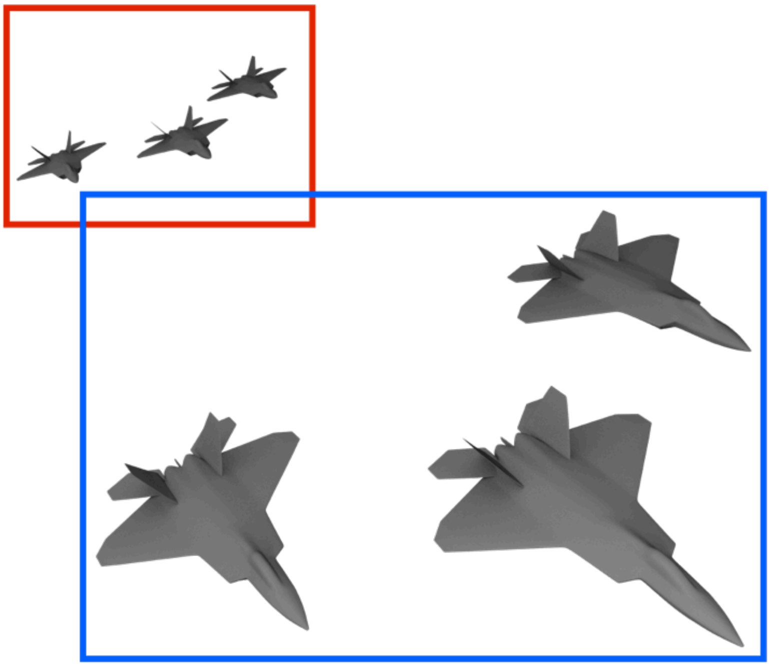


Figure 31. INTERFACE CLUTTERING (INCREASES WITH THE DISTANCE)

4.2.1.3 DESIGN METHODS AND EVALUATION CRITERIA

With the aim of developing and validating the ideal 3D UI for the provision of air traffic control service, many criteria have been used, namely User Centred Design (UCD), technology-driven design and Ecological Interface Design (EID)

UCD focuses on the user and makes use of a cyclic, human in the loop, validation process. Typically, the design process also includes many prototypes, simulations and experiments. However, the ‘experimental approach’ has several limits and has been harshly criticized because it takes no account for the context wherein test subjects act [150]. In other words, experiments ignore elements that are not measurable (e.g. know how, habits, preferences, social interaction, etc.), but would probably affect the subject’s behaviour in a real working environment.

Technology-driven design concentrates on software and hardware prototyping, targeting performance and technological innovation.

EID differs from UCD and technology-driven design insofar it focuses on the analysis of the work domain (a.k.a. Work Domain Analysis - WDA) rather than on the end user or specific tasks. In EID, the *abstraction hierarchy*, which is a 5-level functional decomposition, is used to determine whether a certain piece of information should be displayed on the interface and how this information should be arranged. In doing so the designer attempts to make constraints and complex relationships that are already present in the work environment perceptually evident to the end user (e.g. visible, audible), to free up some cognitive resources that might be used for other cognitive tasks. As an example, the reader can easily refer to the use of *tunnels in the sky* (a.k.a. *highways in the sky*) for aircraft governance.

Overall, UCD has been the leading methodology. A few times, a complementary approach was chosen, which took advantage of both UCD and technology-driven design [151], [152]. EID was only used in [153]. In [24], an in depth analysis of perceptual, contextual and semantic factors was performed in order to categorize surveys and better interpret results.

For the evaluation phase, several criteria have been used, including performance-based techniques, interviews, questionnaires, observations and queries. According to performance-based techniques and queries the success of a prototype depends on the level of performance objectively inferred during or after the experiment. However, interviews, questionnaires and observations may unveil aspects that numerical and fixed-scale assessment techniques cannot detect. For this reason, a subjective metric must be considered a complement rather than a substitute for other metrics [154].

Occasionally, an alternate design criteria was chosen, which is the *cognitive walkthrough* [155]. In [24] and [151] this method allowed for a fast and cost effective validation process relying on theoretical concepts rather than on physical prototypes.

4.2.1.4 RELATED WORK

Ever since the early 90s, many empirical studies have been performed on three-dimensional interfaces for the provision of air traffic control services. The experiments differed in terms of prototypes, equipment and tasks. Also, different groups of people with all kinds of expertise in ATC were involved, including novices, trained ATCOs and pilots.

In an early study from 1991 [156], Burnett and Barfield found that an altitude extraction task, as well as a conflict resolution task, had been performed faster on a (color-coded) 2D interface rather than on a perspective 3D interface. However, for a conflict resolution task, the mean response time was shorter for the 3D interface, but only at low-level traffic density (maximum seven aircrafts *per* simulation). Apparently, the very same interface provided no benefits when the traffic density was increased up to seventeen aircrafts *per* simulation. It is worth noticing that test subjects did know when a conflict was about to occur beforehand. Thus, they were only asked to choose the most appropriate solution among a pre-defined list. In practice, this task entailed the choice of a conflict resolution strategy, rather than the prediction of the conflict itself. *Post-hoc* interviews revealed that controllers did prefer the 3D interface for collecting horizontal position and heading information. However, speed and altitude were still derived from data-blocks. Also, controllers lamented excessive *cluttering* in the 3D interface.

In 1993, Tham and Wikens [118] found little difference between the PVD format and two perspective 3D formats (2.5D and stereo-3D) when tested against selected tasks. The latter led to a higher error rate for speed estimation, whereas the PVD allowed for quicker heading judgments. No difference was found amongst the three formats with respect to a conflict detection task. Overall, the PVD format was favoured over the others. This was more obvious for ATCOs than for pilots. Finally, it is worth noticing that Tham and Wickens do not mention any use of color-coding in their experiments.

In 1993, Haskell and Wickens [138] argued that 3D can be useful "*whenever the tasks to be performed using the display are integrated three-dimensionally*".

In 1994, a study by Wikens and May [119] ascertained the advantages of a PVD over a perspective 3D interface for tasks requiring the vectoring of a few aircrafts around a mountain area. Overall, the PVD allowed for narrower vectors and fewer clearances. Also, decisions were taken faster. The only exception was the judgment of potential penetration threats while flights were not level. In this case, the 3D format brought a little time benefit. The authors ascribed this to "*a more direct spatial extrapolation of changes [that] can be made with the perspective display*". On the contrary, "*a fairly complex extrapolation of changes in digital data tag reading must be performed*" on the PVD. Yet, this benefit was not observed when flights were level because the perspective view disrupted the altitude estimation.

Again in 1994, suspecting that the test subjects' familiarity with a certain interface might influence their performance, Wickens, Miller and Tham, compared six ATC specialists with seventeen pilots trained in ATC [157]. Results confirmed that pilots extracted the information from the PVD as fast as from the perspective 3D interfaces (both 2.5D and stereo-a), whereas controllers were quicker than pilots on the PVD. Also, controllers were erring "on the side of caution", as it happened that they rejected some requests even if these were safe. Finally, results were consistent with the ones obtained in a previous study, involving seven controllers and nine pilots [119].

In 1995, a last experiment was performed by Wickens, Campbell, Liang and Merwin [139]. This time, they focused on bad weather. When subjects were asked to assign a single heading vector to avoid a weather formation, there was no difference in the response time between the PVD and the perspective 3D interfaces. However, test subjects were less conservative with the 2D interface, meaning that they tried to vector the aircrafts closer to the weather formation. Apparently, in this case, they felt more confident in their distance estimation capability. In a second assignment, having a destination point as a target, subjects had to clear a set of vectors to safely avoid a weather formation. This time both changes in altitude and heading were allowed. When relying on 3D formats, test subjects issued a smaller number of vectors. On the contrary, when dealing with the PVD format, the number of vectors was higher and trajectories were less conservative. To be onset, the authors themselves somehow discredited the results of this study. Indeed, they acknowledged having experienced problems with the stereo 3D interface, which "*had not consistently produced true and 'fused' stereo images*". Also, many test subjects said they were unsatisfied with the 3D interface because of the visual clutter. Nevertheless, the spontaneous comment of a controller showed interest toward this technology. He believed it could be used for training purposes. Interestingly, this supposition found empirical evidence when the authors observed what they called the '*asymmetric transfer effect*': test subjects who had begun the experiment with the 3D interface, improved their performance in the subsequent trials with the PVD, whereas the opposite effect was not observed.

In 1995, Wikens declared that the overall results of his research program did not provide enough evidence to fully support three-dimensional interfaces over bi-dimensional PVDs [120]. However, he believed that the 'hamletic' question "*3D or not 3D*?" was not fully answered.

In 1996 Azuma, Daily and Krozel presented a perspective 3D interface prototype featuring:

- An interactive navigation system.
- 3D sound capabilities.
- Advanced decision support tools (e.g. collision avoidance alerts).

Feedbacks were collected from both pilots and controllers. Pilots liked most of the features; especially the use of '*ghost planes*' to represent the aircrafts forecasted position and the visual cues underlining the presence of nearby objects. Reactions to the 3D sound were mixed: some deemed it useful, whereas others judged it irritating. Amongst the military personnel some thought that the 3D format could be useful for sharing information about enemy targets, such as aircrafts, warships, ground troupes, etc. Finally, some suggested using this kind of interface for the '*free flight*' concept. Controllers did not care about the navigation system at all (nor the others advanced features). Most of them chose a fixed viewpoint to watch the traffic from. A few were adamant that 3D interfaces were too confusing, thus, a bad idea. One controller did suggest the use of 3D sound for ground movement guidance. Again, the idea of using the 3D format for training purposes was proposed. Finally, it was hypothesized that the use of a 3D media could possibly shorten the transition time during shift changes.

In 1997, Brown and Slater questioned the controllers' ability of estimating distances and angles in perspective mode. Hence, in a preliminary test [22], they asked several former controllers to judge distances and azimuth angles between pairs of aircrafts in both a perspective 3D and a orthographic 3D interfaces. According to the results, orthographic projection allowed for a greater accuracy in the azimuth angle estimation. On the contrary, the distance evaluation task provided to be independent from the type of projection system. However, further drawbacks of the perspective mode were exposed:

- Objects at a greater depth tend to be displayed closer to each other, resulting in increased clutter (cf. sec. IV)
- The dimensional shrinking along the line of sight partially negated the benefits of aircrafts' drop lines for altitude comparison.

Thus, an orthographic projection was set up in the main study. Surprisingly, no test subject commented on the absence of linear perspective. The authors ascribed this to a first-time experience with the 3D interface. Within the experiment, two groups of test subjects, namely the *expert group* and the *novice group*, were crossed with three types of interfaces: 2D, 2.5D (3D) and stereo 3D. The *expert group* was made up of trained ATCOs, while the *novice group* was composed of novices (i.e. test subjects unfamiliar with ATC). One of the 3D interfaces incorporated stereopsis, whereas the other was built on top of monocular depth cues only. A spatial ability test [158] failed to detect any difference between the skills of the two groups, somehow negating the results of a previous study [119]. However, clear evidence was found, in the form of answers to questionnaires, of a bias towards the use of data-blocks by ATCOs. This was done out of concerns over accuracy or force of habit, even when relying on a 3D display format. Also, a tendency to separately consider vertical and lateral separations was observed. In practice, tasks were not being performed in an integrative manner, i.e. exploiting the integrative nature of the 3D interface. Thus, for tasks entailing angles and distances estimation, 2D yielded to better performance. When the test subjects were asked to identify the two aircrafts flying at the lowest and the highest altitude, no benefits were found for the expert group coupled with the 3D format. Finally, a conflict resolution task showed that many trajectories had been incorrectly perceived on the 3D interface. overall, many questions remained unanswered: had the task been performed in a non-integrative manner because of an inadequate depth perception? Was this related to the task definition or was it due to the subjects' inexperience with the 3D format? Could they be trained to perform those tasks in a way that exploits the integrative nature of a 3D rendering?

In 1999, Van Orden and Broyles compared four types of displays: 2D top-down displays (PVDs), perspective 3D displays, stereo 3D displays, and volumetric displays (laser-based) [159]. For a series of tasks, such as velocity assessment, altitude judgement, vectoring and collision avoidance, performance with the bi-dimensional display was at least as good as with the others display formats. Also, they believed volumetric 3D to be well suited for tasks entailing the perception of "*complex and dynamic information relationships in a confined 3D space*".

In a few of studies between 2000 and 2010, Persiani et al. found 3D features to be useful for specific tasks, such as flight phase recognition and conflict detection [115], [128], [160]. Thus, they conclude: "*the positive acceptance, evidenced in the evaluation phase, shows how the increased readiness of computer graphics and virtual reality technologies can push the adoption of such techniques in the design of innovative interface*".

In 2001 [110], [142], Smallmann, John, Oonk and Cowen compared the effectiveness of representing 3D in a 3D space and 3D in a 2D for an altitude and path determination task:

- **3D in a 3D space**: flight levels and flight paths could be inferred from a 3D icon moving in a 3D space.
- **3D in a 2D space**: the same information was presented by means of 'analogic' symbols 'attached' to the aircraft (e.g., a dynamic bar whose length indicated the flight level).

The '3D in 2D' representation proved to be the quickest source. This demonstrated that the availability of analogical information might have a bigger impact on information acquisition than the choice of the display-format (e.g. 3D). However, the advantages of 3D graphics for tasks entailing 3D thinking and shape understanding remained unquestioned. In a later study, John, Smallman and Cowen found a combination of 2D and 3D to perform better than strict 2D or sole 3D for tactical routing [161]. The task consisted in deploying a chain of antennas across a swath of terrain while avoiding the enemy's sight.

In 2003, Tavanti, Le-Hong and Dang [127], [162] compared the performance of several ATCOs in a series of tasks entailing the use of 2D and stereo 3D display formats. Results showed that participants achieved better performances with the stereo 3D display, with no detriment for accuracy. Yet, it is worth noticing that test scenarios did not display a very deep area. Thus, the *height constancy bias* was not remarkable. Also, the authors expected the acquaintance with the 2D planar interface to negatively affect the performance with the stereo 3D interface. However, this hypothesis was not confirmed.

In 2009, a study by Tavanti and Cooper [35] revealed that ATC trainees did not foresee the use of perspective 3D for tasks entailing typical radar monitoring, such as approach or en-route air traffic control. In general, controllers showed *"distrust and diffidence towards novel and (or) 3D interfaces*". However, they suggested testing 3D against different tasks, such as the provision of air traffic control service from the control tower, the holding stack management, the traffic allocation between sectors and the training of young ATCOs.

In 2010, Cooper, Fridlund, Andel, Bojan and Hardy provided some evidence in favour of the perspective interface concept [129]. In their study, a standard training scenario was more easily solved thanks to the 3D interface. However, for some people, the very same interface was found confusing. Indeed, during the test, a couple of controllers became very bewildered about the lateral separation between two aircrafts and had a hard time trying to solve a situation which was quite simple.

According Cooper et al. the use of a semi-immersive 3D virtual environment resulted in both quantitative and qualitative benefits [163]. In their study from 2005, fourteen former controllers were engaged in a judgment task, i.e. the detection of critical flight levels within a certain air traffic scenario. Results showed that controllers performed tasks faster using the stereo 3D display format rather than the PVD and were at-large satisfied with the 3D rendering system. However, the experiment was based on a low-level traffic density, which critically reduced the possibility of errors. Also, qualitative results were self-esteemed.

In [164], some ATC trainers invited to examine and feedback a stereoscopic 3D interface commented: "*3D visualization could enhance controllers training as these representations are similar to the constructed mental models that the trainee seeks to develop*".

In 2006, Abadir et al. reported that their perspective 3D interface was judged time-consuming: both ATCOs and untrained subjects complained about the presence of too much information, which, according to Sternberg's similarity theory [165], "*made it more difficult to distinguish the important stimulus*" [166]. Most importantly, the interviewed ATCOs argued they "*had no difficulty visualizing the airspace mentally and therefore found 3D unnecessary for operative use*". All in all, the authors suggest focusing on more suitable areas for 3D, such as training and accident/incident analysis.

4.2.1.5 THE EFFORT TOWARD 2D-3D INTEGRATION: AUGMENTED REALITY FOR APPROACH CONTROL

The complexity of the ATC task and the awareness of the pros and cons of a 3D rendering lead some authors to believe that both 2D and 3D views were needed at the same time [110], [139], [161]. A relatively straightforward way to achieve this was to place the conventional PVD alongside with a three-dimensional display. Alternatively, both 2D and 3D views could have been integrated in a multi-frame setup to be displayed on a single screen. In any case, the user would have needed to integrate the information from two different sources and balance his or her attention between them. This is the kind of process that requires a number of fast eye movements known as '*saccades*', which can be a cause of visual fatigue [167]. Also, during saccadic movements, humans are more likely to miss noteworthy events that are displayed on the screen, which is sometimes referred as '*change blindness*' [168]. Thus, they will commit more errors if they divide their attention among several displays (or frames).

A further commitment toward the consolidation of 2D and 3D views was made by the "3D-in-2D Planar View Display" project, sponsored by EUROCONTROL in the framework of the third CARE-INO program (Co-operative Actions of R&D in EUROCONTROL, 2007-2009) [153]. In this project, ten concepts were developed, trying to find the most effective ways of combining the "*3D and 2D views of the air traffic (i) allowing the controller to benefit from the mutual capabilities of the two displays; and (ii) minimizing the effort when moving from one view to the other*"[117]. Concepts were mostly implemented in the Augmented Reality Toolkit (AR-Toolkit) [169], with a maturity level corresponding to the early stages of nine-point NASA Technology Readiness Level (NASA-TRL) framework [170]. Some had been actually drawn from an earlier FP6 project named "AD4 - Virtual Airspace Management" [171]; however, the majority was derived from a two-way 'combination display', in which the project partners correlated numerous 'display techniques', such as *3D in 2D Symbols*, *Multi-Windows*, *Rapid Zooming*, *Distortion*, *Overview Plus Detail*, *In Place*, *Filtering* and *Multiple Coordinated Views*, with several 'display formats', such as *Strict 3D*, *Side By side*, *2D/3D Multi-View*, *Exo-Vis and In Place View* [172], [173, p. 1]. The 'display techniques' determined how the information was rendered, whereas the 'display formats' dictated if and how 3D and 2D blended in. Regrettably, at the end of the first year, the Eurocontrol Innovative Research Advisory Board (IRAB) redirected the project. For this reason, only low-fi prototypes were developed and no time was left for concepts integration.

It is beyond the scope of this document to discuss each concept in depth. However, some early prototypes seemed quite promising and ground-breaking, especially if coupled with specific tasks. On the other hand, aspects that had been initially theorized were not adequately discussed, namely the human mapping capability, the contextual awareness and the limited attention bandwidth. In the absence of new developments, it would be hard to further comment on this study.

4.2.1.6 ANALYSIS AND DISCUSSION

Conventional PVDs use bi-dimensional maps and standard symbology as an interface between the Air Traffic Control Radar Beacon System (ATCRBS) and the ATCO. Except for ground coordinates and heading information, this practice involves a high level of abstraction in the traffic representation. Controllers need to interpret symbolic data, manage flight strips, interact with pilots and possibly coordinate with colleagues. At the same time, they need to stay focused on the job and maintain an adequate SA of the overall traffic situation. For this reason, it has been suspected that the 3D format may provide a more 'natural' and intuitive representation of the airspace, allowing for spatial information to be easily grasped and processed by the operator. Thus, a great effort has been put in creating the ideal three-dimensional UI for the provision of air traffic services. However, empirical studies have shown that fine-grained metric judgments along any particular direction are not possible in perspective 3D, leading to cognitive errors when trying to assess the objects' absolute or relative position [24], [110], [118]–[120], [138]–[143], [147], [148]. A more precise estimate might be given by navigating the virtual environment, looking for the most appropriate viewpoint (which is often the one perpendicular to the distance itself). However, this would increase the complexity of the interaction [114]. Similarly, it has been shown that judging absolute and relative angles may be tricky [22], [119], [119], [139]. With the aim of reducing perceptual biases, corrective measures can be taken, such as the use of widgets, rulers, grids, drop lines and scales [22], [114], [117], [172], [174]. However, this would increase the interface clutter [22], [139], [156]. Finally, the use of perspective cuts out of the display those objects that do not reside within the camera field of view. In other words, a close look to a specific location inhibits the view of the global dataset. As a result, the controller's awareness of the overall traffic situation is reduced, possibly leading to spatial disorientation [175], [176]. Also, because the controller must elaborate both 'focused' and 'contextual' information, the overall cognitive load increases. Of course, a comprehensive view of the scene can be used, which would provide adequate SA. But then again, this may lead to unacceptable levels of *cluttering* and *display resolution* issues [22], [24], [139], [156].

Due to the absence of linear perspective, in a 2D top-down interface the assessment of horizontal distances and angles is a relatively straightforward task. Also, there are no open issues associated with a complex navigation system and all the aircrafts are always in sight. Nevertheless, the decision-making process requires the integration of information that is only exposed by means of non-spatial codes (i.e. alphanumeric data-blocks). This includes the flight level or the altitude information. To convert the data-block information into conceptual knowledge, demanding arithmetic must be performed by the ATCO. This kind of process is typically referred as a '*controlled'* process, because it requires a great care and must be executed sequentially. With a great deal of practice, some *controlled* processes may become ‘*automated’*, i.e. executed unconsciously and performed in parallel. For instance, it has been proven that the establishment of a spatial relationship between the aircrafts becomes a partially automated process when relying on the PVD [35]. This happens as a result of the extensive training to which air traffic controllers are subject and was confirmed several observing how controllers took advantage of the PVD interface over pilots and ATC trainees [114], [118], [157].

As the ‘*Proximity Compatibility Principle*’ by Wickens and Carswell asserts [177], whether or not a 3D interface suits a certain task depends on whether the task itself is integrative or not in space and time [178]. If the task requires the integration of multiple sources of information, performance will be best supported when those sources are displayed in close ‘proximity’ (e.g. closeness in space, resemblance of colour, dimensional integrity, etc.). On the contrary, if the task requires the user to focus on a single source of information, performance will be best supported by a disjoint representation. In this regard, some evidence was found that many tasks in ATC may not require integrated spatial judgements [22], which would somehow negate the advantages of a three-dimensional interface. If this were true, the design of a good interface would be made more difficult by the fact that air traffic control requires the execution of both ‘integrative’ and ‘attention-focused’ tasks at the same time.

Overall, any use of 3D graphics that involves accurate estimates of distances and angles must be rejected, including those tasks that strongly rely on radar monitoring, such as en-route and approach control. For these assignments, 3D will only produce detrimental effects, especially if tested against high traffic conditions. A successful approach could be to combine 3D with those tasks that do not suffer from the perceptual shortcomings associated with it, as proposed in [179].

Concluding, we believe there is no such question as “*3D or not 3D?*” [120]. The utility of a graphical interface, either three-dimensional or bi-dimensional, cannot be argued *a priori*, or in absolute terms. On the contrary, it is relative and dependent on a number of factors, such as rules, goals, perception, culture, semantics and tasks [24]. Hence, a different question must be formulated: given a certain domain, under which rules, circumstances and tasks can a 3D interface contribute to human performance?

4.3 VIRTUAL AND AUGMENTED REALITY FOR TOWER CONTROL

4.3.1 AIRPORT TOWERS

Most of the research on the use of augmented reality in a local tower setting has been performed in the USA.

In [88] a series of head-tracking, see-through, head-mounted display prototypes were developed and evaluated by five controllers at Moffett Field air traffic control tower. The approach described in this paper focused on iterative engineering prototype build-evaluation cycles. Test subjects identified several deficiencies in the initial prototype, such as low optical transmissivity of the display, unacceptable compensation for tower lighting conditions, inadequate symbology and data block information, and unacceptable discomfort caused by wearing the head-attached displays.

The controllers' evaluations confirm that the prototypes are not acceptable for operational use. However, they indicate enthusiasm for augmented reality as an aid for tower operations, providing the technology matures sufficiently. The authors conclude that although many fundamental problems must be understood and solved, there is ample reason to believe that eventually augmented reality tower tools may become operational.

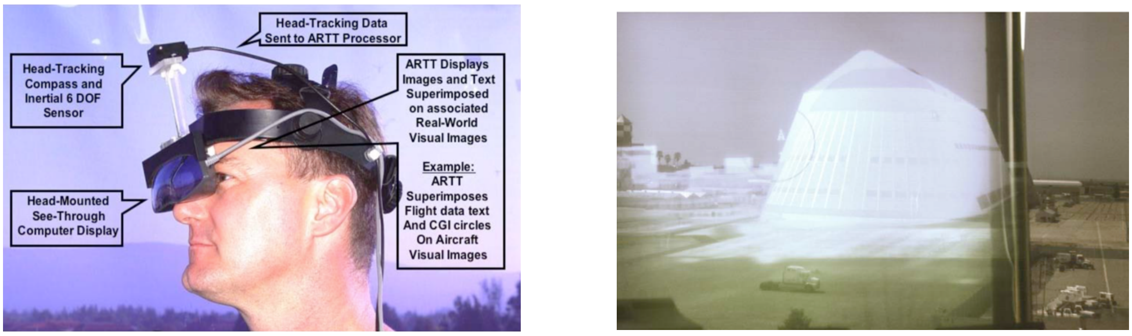


Figure 32. MOFFETT FIELD EXPERIMENT: PROTOTYPE HEAD MOUNTED DISPLAY (LEFT IMAGE) AND AUGMENTED REALITY OVERLAY OF AN HANGAR (RIGHT IMAGE)

In [11], augmented reality technology adapted for control tower applications was used to track an OH-58C helicopter in proximity of an airport. A camera based video see-through display system was used to measure the registration error of static airport features and dynamic objects (aircrafts). The observed registration errors were largely attributable to two terms: 1) Terminal Radar Approach Control (TRACON) ASR-9 surveillance system or ADS-B latency, and 2) registration error of the virtual environment from all sources, including: (i) camera orientation tracking errors, (ii) errors in the visual database of the airfield. The authors state that the best method of eliminating misregistration of static objects would be to have an orientation tracking system without artefacts or errors. Alternatively, it is possible to measure the misregistration in each image and use that information to calculate an improved solution. Manually compensating for registration errors of static features and modelling aircraft behaviour according to tower controllers' recommendations, the authors could reduce registration errors for dynamic objects (aircrafts) to ≤ 2° if the surveillance system latency was ≤ 5 seconds, and to ≤ 1° if the surveillance system latency was ≤ 2 seconds. It was hypnotised that more sophisticated aircraft speed and routing models should yield better results. Performance with the TRACON ASR-9 tracking system deteriorated when the aircraft was < 2 nm from the observation point, and the farther the aircraft was from the observer, the greater was the agreement between the ADS-B and ASR-9 indicators. This discrepancy was largely attributed to the higher update rate of ADS-B (1 second), which compensates for greater angular displacements of closer objects (ASR-9, has a slower update rate of 4.8 second). In addition to the Army's OH-58C, other aircrafts, equipped with ADS-B were also observed for less formalized tests, including a B-767-200. The behaviour of the Augmented Reality Tower Tool (ARTT) prototype with fixed-wing aircrafts was similar to that with the test helicopter.

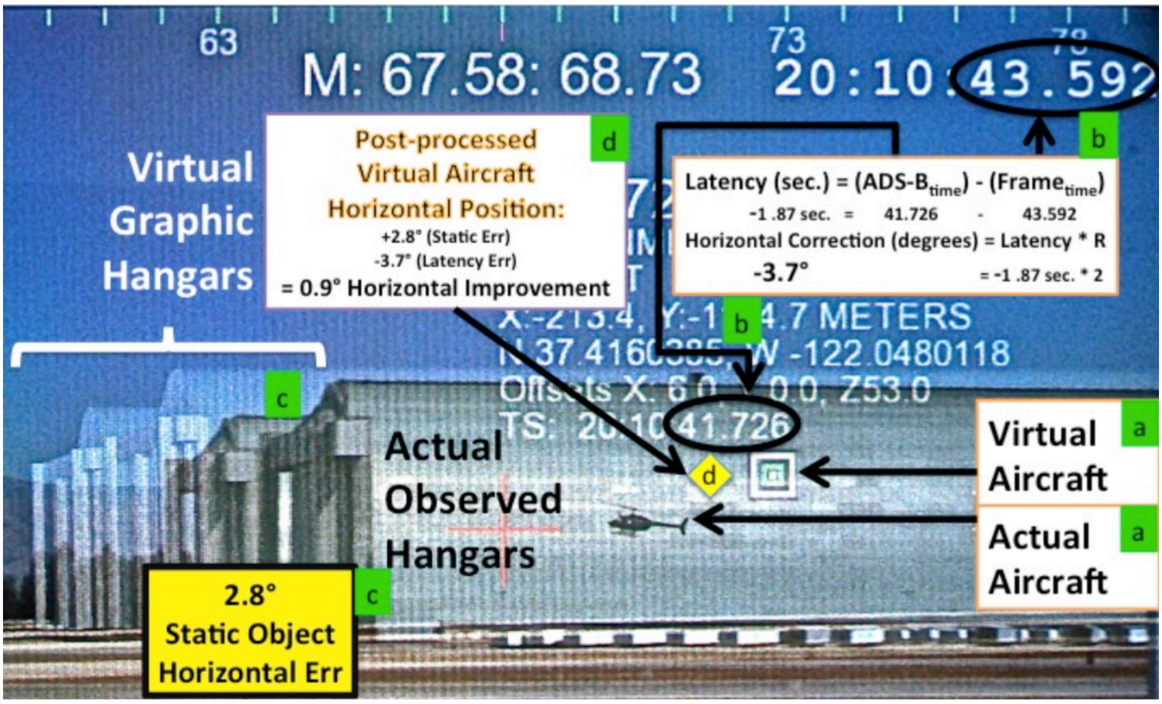


Figure 33. Combined image showing registration errors of both static and dynamic objects

In [13] a field study using a video camera mounted on 'prescription' glasses worn by air traffic controllers was performed to investigate what visual information are acquired, if the scanning pattern depends on the outside view visibility and what are the working methods regarding various information sources. Results indicate that controllers' performance described as 'head-up' and 'head-down' activities depend on traffic load and visibility conditions. Regardless of traffic load, ground controllers prioritize the outside view over other sources of information. However, during high traffic loads controllers show a tendency to use the head-down tools more than during low traffic. Concerning day and night conditions, the results confirmed the expected tendency that during the night the importance of the head-down support tools increases. However, monitoring Runways and scanning Apron by means of aircrafts' and ground's lighting remains a main source of information. The process of scanning the sources of information accessible head-up and head-down is different. Monitoring the airport's surface is a frequent and long duration activity. By comparison, scanning strips and other tools are frequent but relatively short duration activities. The analysis of patterns revealed that the strongest transitions were from looking outside of the tower windows to strips. It was again confirmed that controllers frequently switch attention from the strips on their desk to the view of the far distance through the window. Those two main transitions represent the identification process required to maintain adequate metal picture of the situation. Aircrafts that are visible by the windows are identified using the information provided by the strips such as airline, aircraft type and parking position. Overall, results of this study confirmed that direct observation is of prime importance to the tower controller. Strips are a second source of information, requiring frequent head-down movements and consequently changing the point of gaze and adjusting the focal depth to short distance for a very short time. Presenting the information that currently is available on head-down devices on the head-up display should significantly decrease head-down time. Furthermore, it should eliminate the fixation switch between far and near locations which was reported as being a component of the head-down problem by [180].

In [19] the visual requirements for an augmented or virtual reality display that might be used in real or virtual towers have been reviewed with respect to similar display technologies already used in aircrafts' cockpits. Using an optical see-through head mounted display, different binocular fields of view (14°, 28°, and 47°) were examined to determine their effect on subjects' ability to detect aircraft manoeuvring and landing. The results suggest that binocular fields of view much greater than 50° are unlikely to dramatically improve search performance and that partial binocular overlap is a feasible display technique for augmented reality tower applications. Such a field is easily achievable with existing head-mounted see-through displays. However, the predictability of the traffic pattern probably contributed to this restricted FOV requirement. The subsidiary experiment comparing full with partial overlap systems did not find any consistent performance difference between the 100% and partial overlap conditions, meaning that visual suppression due to 'luning' [181] was compensated by the constant motion of the eye and head.

In [14] a technology assessments of five off-duty controllers who ‘shadow-controlled’ with an augmented reality HMD prototype in their own facility is presented. The HMD prototype used in this study displayed both dynamic and real-time air traffic data (e.g., aircraft location, identity, speed and distance), and computer-generated graphics that correlate to the static environment objects (e.g., horizon line, compass rose). Figure 34 illustrates the symbology used in the study. Computer generated graphical red circles with 2° diameters identified the computed position of each aircraft, based on the airport surveillance real-time data (TRACON). A three-line data block was associated with each of these aircraft marker circles. The first line of the data block contained the aircraft call-sign. The second line contained the horizontal distance (in nautical miles) of the aircraft from the viewer’s position in the control tower. The third line displayed the aircraft’s altitude in feet. Ground reference was represented by a computer generated green horizon line, with compass heading markers at 10°intervals. Each interval had a vertical line with the compass heading displayed above it. The orientation of the viewer’s head was also displayed, expressed as magnetic heading (displayed at the top of the frame), pitch (displayed on the right), and roll (displayed at the bottom of the frame).

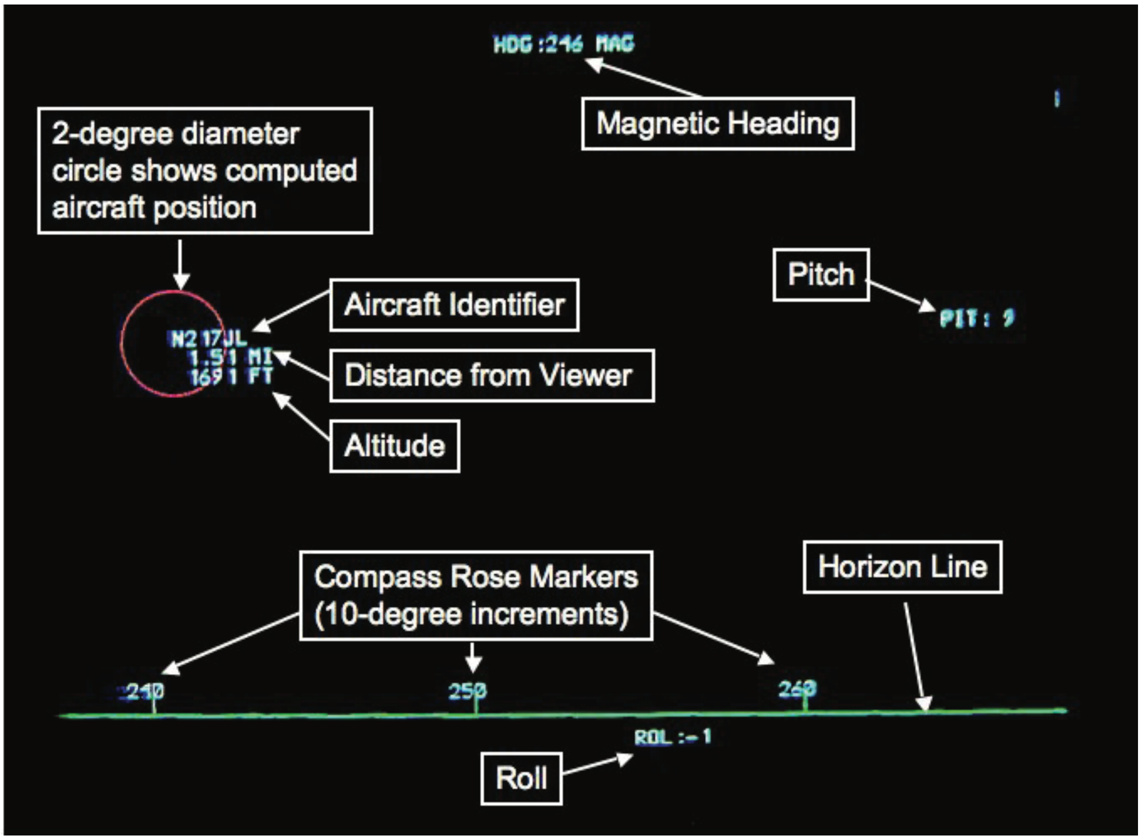


Figure 34. Symbology used in the prototype display[12]

---

[12] The callout labels (in white) are illustrative and were not shown in the actual prototype. The areas that are black in the above illustration were transparent when viewed through the see-through head mounted display.

A subset of 79 questions addressing the potential of AR technology to benefit well-defined controller's tasks and duties was submitted to test subjects. Another subset of 26 questions addresses potential information acquisition benefits, including complementary functionality with the radar display. The research prototype apparatus was evaluated for comfort, displayed data block utility and impact on situation awareness. Head tracking performance was also assessed. Results indicate unanimous agreement that this technology is potentially beneficial, though the prototype used in the study was not adequate for operational use. Some controllers agreed that augmented reality technology improved situational awareness, had potential to benefit clearance, control, and coordination tasks and could be very useful for acquiring aircraft and weather information, particularly aircraft location, heading, and identification. The strongest objections to the prototype used in this study were directed at aircraft registration errors, unacceptable optical transparency, insufficient display performance in sunlight, inadequate representation of the static environment and insufficient symbology.

After some early studies on stereoscopic labels placement [95], Peterson proposed the use of a transparent projection screen for controller's standing in a fixed position and observing distant objects [38]. The transparent projection screen is similar to the head-up displays used in aircraft cockpits, since they both superimpose graphics on objects in the background. However, in cockpits' HUDs, there is usually no registered overlay of real objects, just symbols and data at fixed screen locations (i.e. no object connectivity). Registered AR instead, strives to perfectly overlay the real with virtual objects (i.e. full object connectivity). Peterson hypnotises that since binocular disparity is not effective in panoramic environments it should not be necessary. Instead it should be sufficient to provide a biocular overlay to the user (biocular should not be confused with binocular, where two different images are produced that respect the disparity.). However, this was not experimentally proven.

In a following study, Peterson performed some experiments to determine the need for collimation in spatial AR devices [15]. A slight overestimation (8.4 cm) was found in the matching of virtual objects with real ones using biocular rendering over short and medium ranges (3-10 m). However, the results did not show any statistically significant effects between the independent measured variables, namely screen distance (1 or 2.5 m) and target distance (4, 6 or 7.5 m). The authors conclude that the three variables have no effect on depth-matching accuracy (which would be contrary to theory) or the statistical power and precision of his experiment was insufficient. He recommends performing a follow-up experiment with modified target shapes and broader ranges in the experiment variables. Especially the target distance range should be substantially increased to determine the effects of collimation and depth rendering in the distance.

Figure 35. Dual projector set up (left image) used in Peterson's depth matching test (right image)

4.3.2 Remote Towers

The concept of placing AR overlays into the control tower outside view has been further developed in SESAR Operational Focus Area (OFA) 06.03.01 (Remote Tower), which relates to ICAO Block-Module B1-81 (Remote Operated Aerodrome Control Tower) and is currently undertaking Operational Improvement Step SDM-0201 (Remotely Provided Air Traffic Service for Single Aerodrome). Being contributed by work package 06.02 (Coordination and consolidation of operational concept definition and validation), project 06.09.03 (Remote & virtual TWR) and several Enablers, SDM-0201 is expected to finish by the end of 2017. The Remote Tower concept enables the provision of the Air Traffic Control Service and the Aerodrome Flight Information Services (AFIS) where such services are either currently unavailable, or where it is too difficult or expensive to implement and staff a conventional manned facility.

Figure 36. Experimental remote tower facility by Saab

The idea of using augmented vision in the RT concept was firstly introduced in the “Virtual Tower“ (ViTo, 2002 - 2004) project. In a follow-up project named “RapTOr” (Remote Tower Operation with Augmented Vision Videopanorama, 2005-2007) the real-time aircraft position obtained from the multilateration system as well as the wind speed and direction were integrated into the video-panorama view of the Braunschweig airport. The concept was further explored and refined in the RAiCe (Remote Airport Traffic Control Center) project.

Figure 37. Flight tag overlay at Braunschweig airport

Since then, many visual aids have been proposed, including flight tags on top of aircrafts, bounding boxes, airport layout and buildings' reconstructions and several types of warnings [182], [183]. However, in the remote tower application, the AR overlays are placed on top of a 2.5D video-surveillance feed of a remote location instead of the actual tower' view. This poses a number of technical and ergonomics challenges which are very different from the ones that arise from the use of such technology in on-the-site control towers [39].

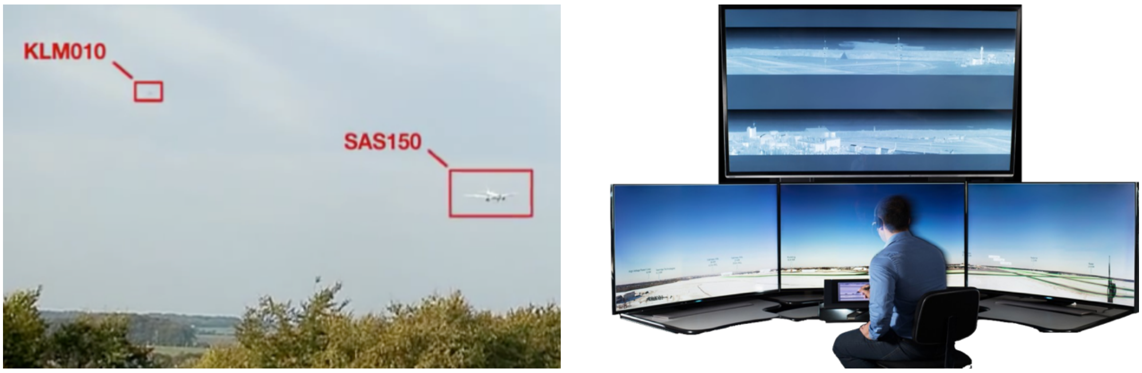


Figure 38. Augmented reality bounding boxes and call signs provided by a Saab remote tower facility (left image) and remote tower controller's working position by Searidge technologies (right image)

Figure 39. Remote towers visual augmentation by Kongsberg Gallium

4.4 HUMAN FACTORS & ERGONOMICS

The application of human factors and ergonomics methods is a key part of the system design, evaluation, and timely implementation. Human Factors and Ergonomics are concerned with designing for human use, and are essentially composed of data, principles and techniques. The data concern human attributes which determine how to achieve good performance, e.g. anthropometric data on body dimensions, or visual data on colour perception, both of which are useful when designing interfaces to ‘fit’ people and help them make sense of what the interface is trying to tell them. The Human Factors and Ergonomics professionals’ main activity is therefore applying generalised data, principles and techniques to the specific context being studies, in this case ATM. The Human Factors and Ergonomics specialist must therefore adapt these data and carry out detailed analysis on human performance in the specific context under analysis.

As an example, fast time or model based simulations and real time human-in-the-loop experiments are frequently used with the objective to assess workload, situation awareness and team-working. Prototyping tools then allow early testing of the concept, via simulation methods predicting controller interactions and workload, and small-scale simulation prototyping exercises allow on-line evaluations with samples of real prospective users. Both approaches allow insights into the degree of usability and performance with the new system concept before detailed design, and for sure are of help to improve the overall concept development.

The main Human Factors and Ergonomics areas that should be investigated are the followings:

- Situational Awareness
- Workload
- Teamwork
- Acceptability
- Usability.

Hereafter it is provided a description of these areas.

4.4.1.1 SITUATIONAL AWARENESS

Situational awareness can be defined as the continuous extraction of environmental information, the integration of this information with previous knowledge to form a coherent mental picture, and the use of that picture in directing further perception and anticipating future events.

In this regard, situational awareness can be considered a mental state consisting of three phases:

1. Perception of the situation (perception of important elements in the environment);

2. Comprehension of the situation (integration of different pieces of data to determinate their relevance);

3. Anticipation of future states of the current situation.

To improve the comprehension of the situation for the user, it is important that he is provided with only the information that is needed for his role and/or location and/or the active procedure (i.e. the context) in the tower control environment. If all available information from all services and data sources where to be provided to every user, there is a risk that the data becomes incomprehensible. Therefore, it's important that the application can adapt itself based on the current context of the user.

4.4.1.2 WORKLOAD

There are two main types of perceived workload: (a) physical workload and (b) cognitive workload.

a) Physical workload is related to the actions required to interact with the system in performing tasks (e.g. clicking, making a phone call, moving head to switch form a monitor to another, etc.). In ATM, the mission is to keep controller's global workload in a range where they are kept (at least mentally) stimulated without going to the point where they become overloaded and start to postpone tasks.

b) Cognitive workload can be defined as the degree of processing capacity that is used to perform certain tasks.

4.4.1.3 TEAMWORK

Teamwork and communication refers to the allocation of tasks between team members and the way information is exchanged between them. Changes in team structure can include changes to the composition of a team in terms of roles, as well as, changes to the way in which tasks are allocated between the team members. Such changes may impact the communication flow within a team and the way tasks are performed.

4.4.1.4 ACCEPTABILITY

There is a causal relationship between system design features and the user's attitude toward using a system and actual usage behaviour. Among this are the perceived usefulness (i.e. the extent to which a person believes that using a technology will enhance productivity) and perceived ease of use (extent to which a person believes that using a technology will be free of effort). A lack of user acceptance can compromise performance and become a serious impediment to the success of new systems and technology. Thus, measuring subjective acceptance is a valuable component in the evaluation of systems.

4.4.1.5 USABILITY

Usability can be defined as the extent to which a software can be used by specified consumers to achieve quantified objectives with (a) effectiveness, (b) efficiency and (c) satisfaction in a quantified context of use [184]. It can be measured only by considering the context of use of the system — i.e., who is using the system, what they are using it for and the environment in which they are using it. Effectiveness, efficiency and satisfaction are defined as follows:

- Effectiveness: the extent to which users can successfully achieve their objectives
- Efficiency: the measure of the effort and resources allocated to achieve the objectives
- Satisfaction: the extent to which the experience was satisfactory.

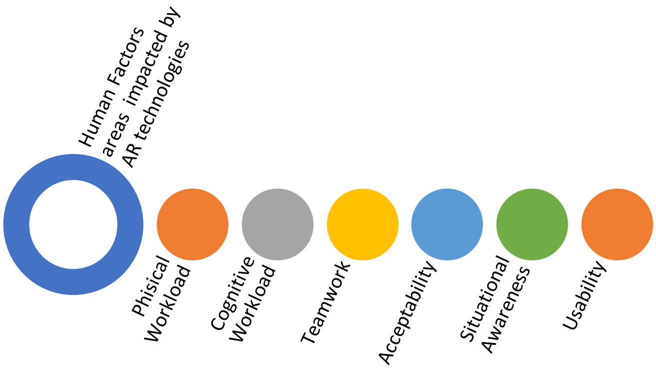


Figure 40. HUMAN FACTORS AREAS IMPACTED BY THE USE OF AUGMENTED REALITY TECHNOLOGIES

5 REFERENCE SCENARIO

This charter describes the reference scenario and the available data sources that have been or can be used in the ARRT development process.

5.1 INTRODUCTION

For early experimentation with AR tools an airport shall meet some basic requirements. These are related to the equipment, lay-out, traffic and procedures:

To provide the AR systems with the position and identification of aircrafts, the airport shall be equipped with Primary and Secondary Surveillance Radar (PSR/SSR) and with Surface Movement Radar (SMR).

Airports with moderate complexity in term of layout have some strong benefits for a first implementation of the ARTT. For instance, a single runway layout should be easy to model and AR tools should impact safety and efficiency in a more effective manner. Moreover, as a first implementation, a too big layout could be confusing and dispersive.

The airport shall be able to confront LVC by means of Low Visibility Procedures. This is very important to show the benefits of the AR tools as the visibility decreases. CAT II/III approach and Low Visibility Take Off Operations shall be available. In terms of equipment this means that the airport shall be ILS CAT 3B equipped.

It is important that certain procedures for the apron management are available and implemented. Such procedures are based on slots and timings often displayed on video (e.g. calculated take off time) and entail head down operations. Exposing such information by means of the ARTT has several benefits.

Resuming, to be eligible as a target scenario an airport shall have at least the following features:

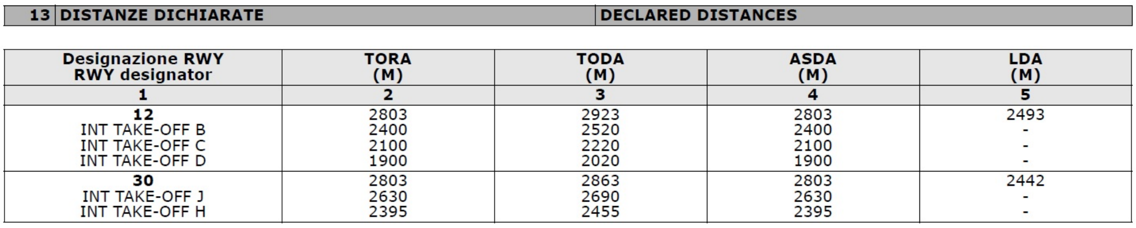
- Primary Surveillance Radar and Secondary Surveillance Rada (PSR/SSR) equipped;
- Surface Movement Radar (SMR) equipped;
- Low Visibility Procedures able to manage more than one aircraft at the same time implemented;
- ILS CAT 3B equipped;
- Moderate complexity (one runway, several taxiways, more than one apron)
- Moderate traffic: volume of 200/300 movement per day;
- Implemented Apron Management Procedures.

For these reasons, Guglielmo Marconi international airport (ICAO code: LIPE) has been chosen as reference scenario for the implementation phase. Bologna Airport meets all the requirements mentioned above moreover the Control Tower is quite big and can easily host future real time experiments.

5.2 GUGLIELMO MARCONI INTERNATIONAL AIRPORT (LIPE)

5.2.1 LAYOUT

Guglielmo Marconi international airport is a single Runway airport with several taxiways, aircraft stands and aprons. The runway orientation is 12/30° with an asphalt strip of 2803x45 m. In Table III, the declared distances are reported for both runway directions.



13 DISTANZE DICHIARATE **DECLARED DISTANCES**

Designazione RWY RWY designator	TORA (M)	TODA (M)	ASDA (M)	LDA (M)
1	2	3	4	5
12	2803	2923	2803	2493
INT TAKE-OFF B	2400	2520	2400	-
INT TAKE-OFF C	2100	2220	2100	-
INT TAKE-OFF D	1900	2020	1900	-
30	2803	2863	2803	2442
INT TAKE-OFF J	2630	2690	2630	-
INT TAKE-OFF H	2395	2455	2395	-

Table III. BOLOGNA G. MARCONI AIRPORT: DECLARED RUNWAY DISTANCES

In the pictures below the aerodrome layout is reported.

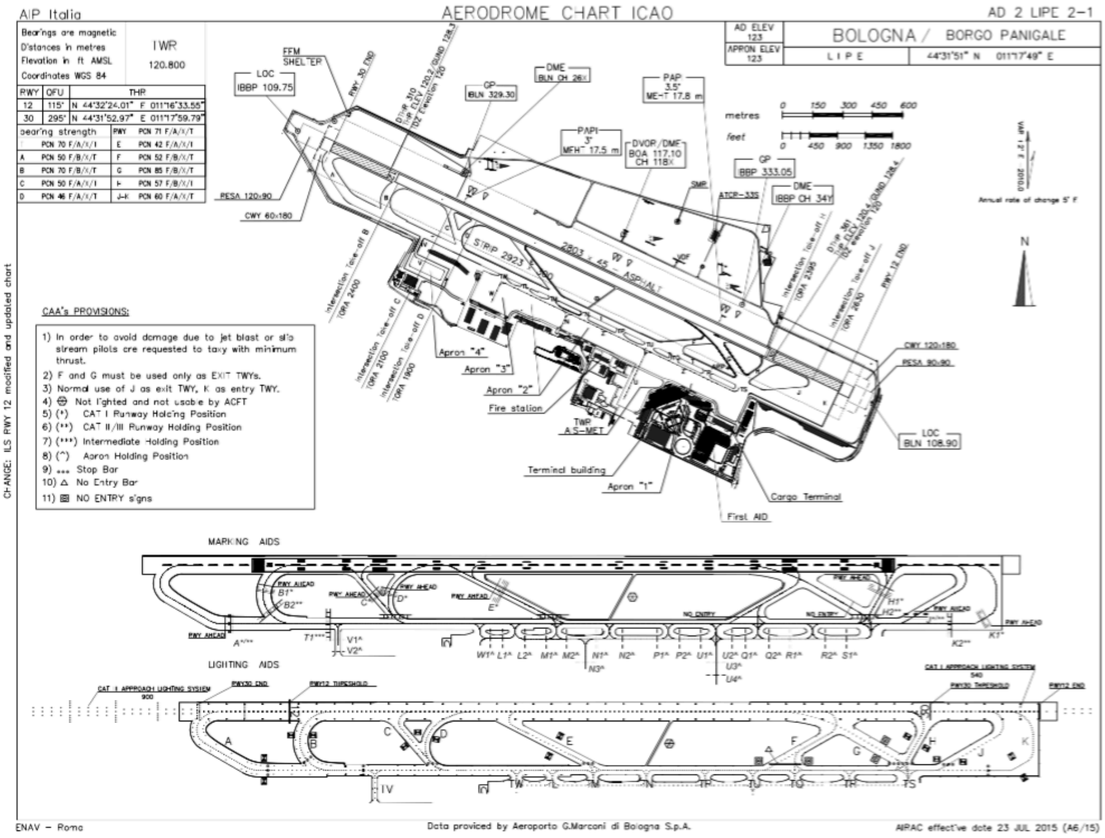


Figure 41. Bologna G. Marconi Airport: aerodrome chart

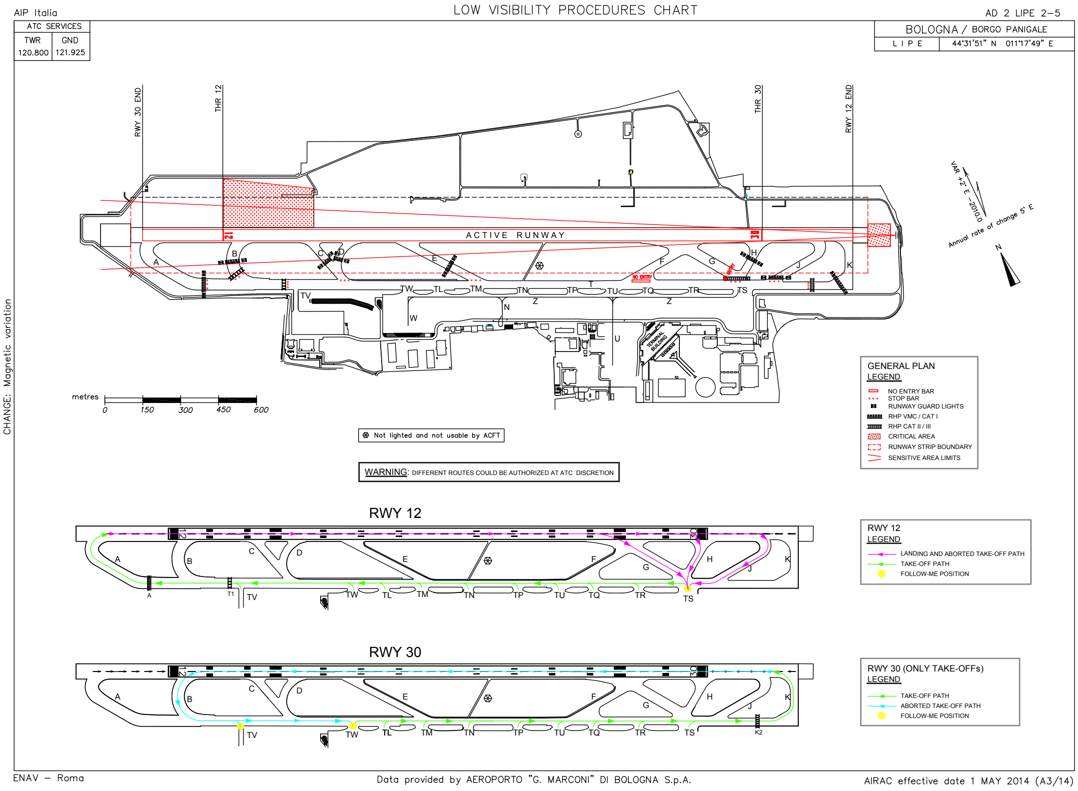
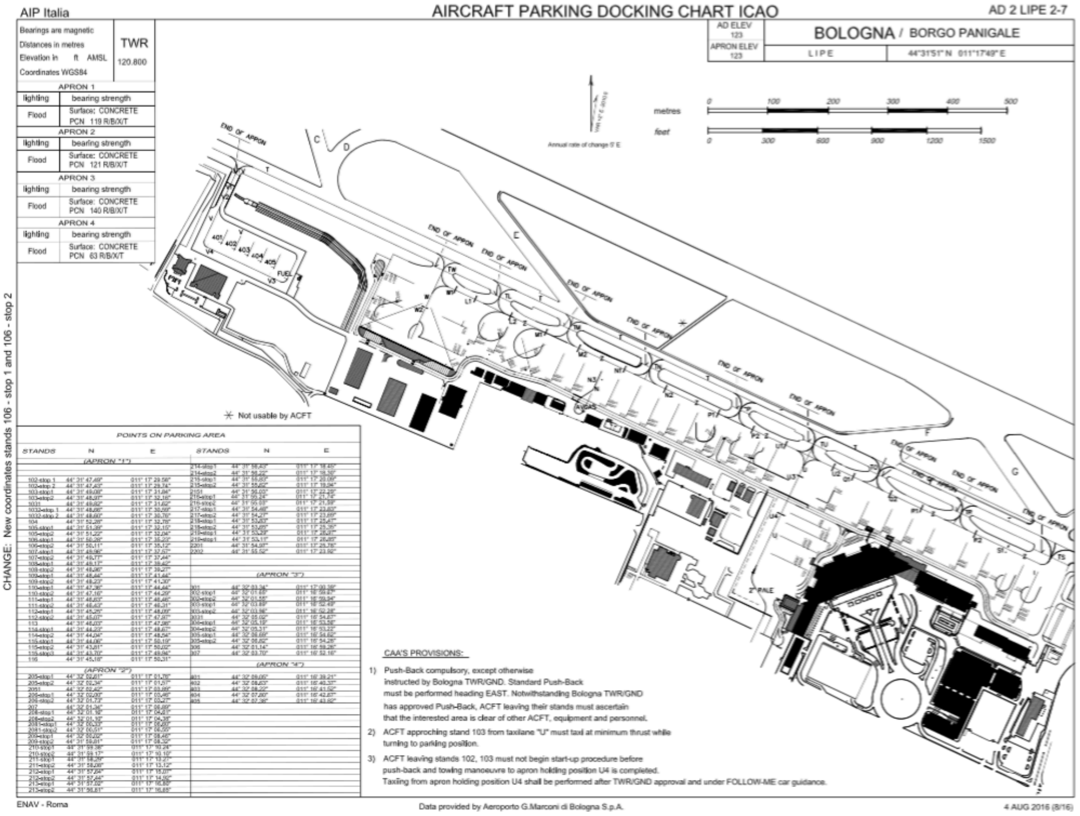


Figure 42. Bologna G. Marconi: low visibility chart



CAA'S PROVISIONS:

1) Push-Back compulsory, except otherwise instructed by Bologna TWR/GND. Standard Push-Back must be performed heading EAST. Notwithstanding Bologna TWR/GND has approved Push-Back, ACFT leaving their stands must ascertain that the interested area is clear of other ACFT, equipment and personnel.

2) ACFT approching stand 103 from taxilane "U" must taxi at minimum thrust while turning to parking position.

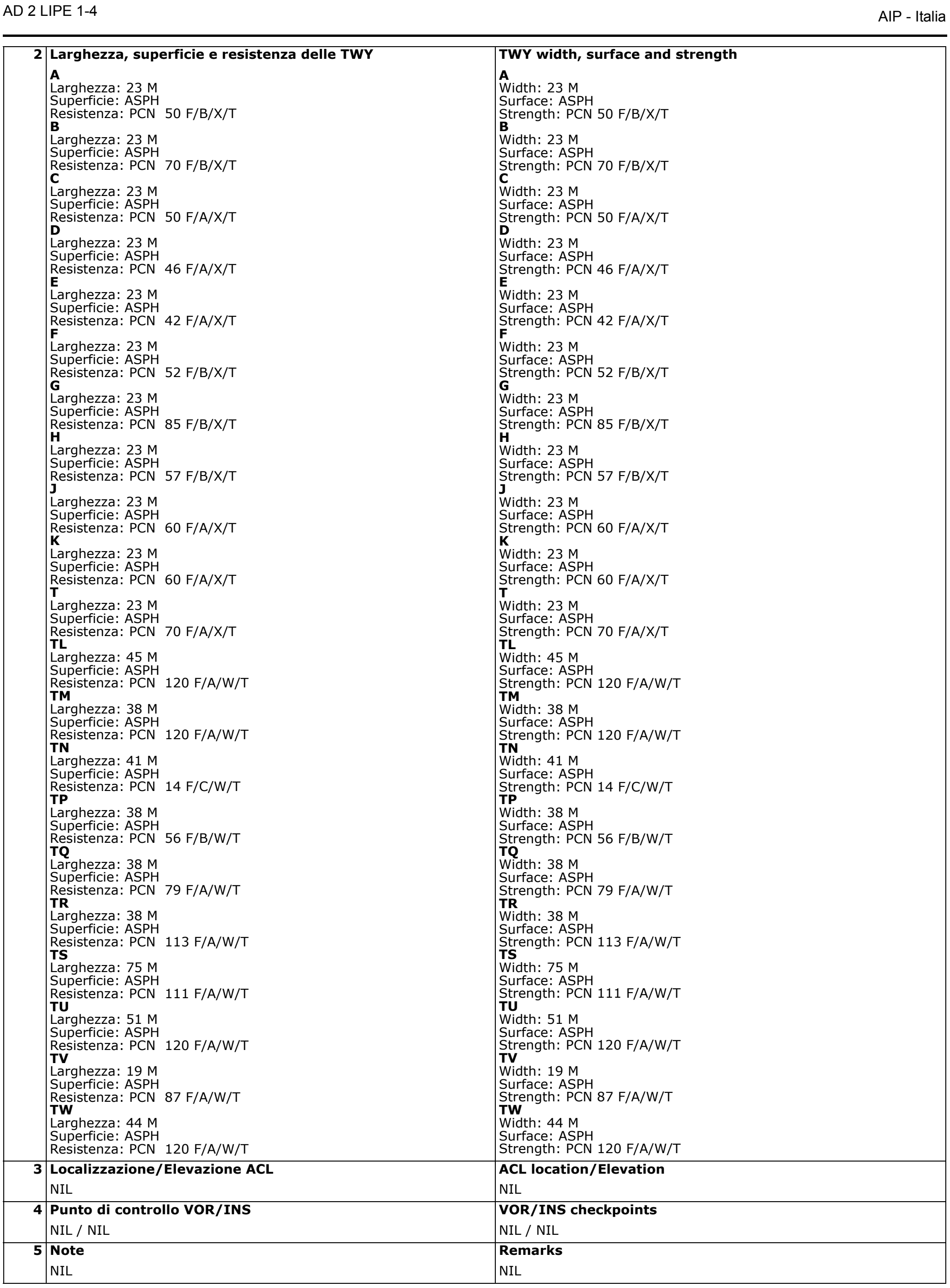
3) ACFT leaving stands 102, 103 must not begin start-up procedure before push-back and towing manoeuvre to apron holding position U4 is completed. Taxiing from apron holding position U4 shall be performed after TWR/GND approval and under FOLLOW-ME car guidance.

Data provided by Aeroporto G.Marconi di Bologna S.p.A.

Figure 43. Bologna G. Marconi: parking docking chart

The main taxiway T is parallel to the runway. The two are linked via taxiways A, B, C, D, E, F, G, H, J and K. Four apron are available; Apron 1 in front of the terminal and of the Control Tower, Apron 2 on left in front of firefighting area and hangars, Apron 3 which is the cargo area and Apron 4 for general aviation. Apron 1, 2 and 3 are linked to taxiway T with short taxiway TW, TL, TN, TM, TP, TU, TQ, TR, and TS. Apron 4 is separated from the other apron and it is linked to the main taxiway T with taxiway TV. The stands are grouped in blocks: all the stands belonging to a block have the same Apron Holding Position (i.e. a position where the aircraft are pushed back and where they start up the engines).

Taxiway characteristics including width, surface and strength are reported below:



AD 2 LIPE 1-4

AIP - Italia

2	**Larghezza, superficie e resistenza delle TWY** **A** Larghezza: 23 M Superficie: ASPH Resistenza: PCN 50 F/B/X/T **B** Larghezza: 23 M Superficie: ASPH Resistenza: PCN 70 F/B/X/T **C** Larghezza: 23 M Superficie: ASPH Resistenza: PCN 50 F/A/X/T **D** Larghezza: 23 M Superficie: ASPH Resistenza: PCN 46 F/A/X/T **E** Larghezza: 23 M Superficie: ASPH Resistenza: PCN 42 F/A/X/T **F** Larghezza: 23 M Superficie: ASPH Resistenza: PCN 52 F/B/X/T **G** Larghezza: 23 M Superficie: ASPH Resistenza: PCN 85 F/B/X/T **H** Larghezza: 23 M Superficie: ASPH Resistenza: PCN 57 F/B/X/T **J** Larghezza: 23 M Superficie: ASPH Resistenza: PCN 60 F/A/X/T **K** Larghezza: 23 M Superficie: ASPH Resistenza: PCN 60 F/A/X/T **T** Larghezza: 23 M Superficie: ASPH Resistenza: PCN 70 F/A/X/T **TL** Larghezza: 45 M Superficie: ASPH Resistenza: PCN 120 F/A/W/T **TM** Larghezza: 38 M Superficie: ASPH Resistenza: PCN 120 F/A/W/T **TN** Larghezza: 41 M Superficie: ASPH Resistenza: PCN 14 F/C/W/T **TP** Larghezza: 38 M Superficie: ASPH Resistenza: PCN 56 F/B/W/T **TQ** Larghezza: 38 M Superficie: ASPH Resistenza: PCN 79 F/A/W/T **TR** Larghezza: 38 M Superficie: ASPH Resistenza: PCN 113 F/A/W/T **TS** Larghezza: 75 M Superficie: ASPH Resistenza: PCN 111 F/A/W/T **TU** Larghezza: 51 M Superficie: ASPH Resistenza: PCN 120 F/A/W/T **TV** Larghezza: 19 M Superficie: ASPH Resistenza: PCN 87 F/A/W/T **TW** Larghezza: 44 M Superficie: ASPH Resistenza: PCN 120 F/A/W/T	**TWY width, surface and strength** **A** Width: 23 M Surface: ASPH Strength: PCN 50 F/B/X/T **B** Width: 23 M Surface: ASPH Strength: PCN 70 F/B/X/T **C** Width: 23 M Surface: ASPH Strength: PCN 50 F/A/X/T **D** Width: 23 M Surface: ASPH Strength: PCN 46 F/A/X/T **E** Width: 23 M Surface: ASPH Strength: PCN 42 F/A/X/T **F** Width: 23 M Surface: ASPH Strength: PCN 52 F/B/X/T **G** Width: 23 M Surface: ASPH Strength: PCN 85 F/B/X/T **H** Width: 23 M Surface: ASPH Strength: PCN 57 F/B/X/T **J** Width: 23 M Surface: ASPH Strength: PCN 60 F/A/X/T **K** Width: 23 M Surface: ASPH Strength: PCN 60 F/A/X/T **T** Width: 23 M Surface: ASPH Strength: PCN 70 F/A/X/T **TL** Width: 45 M Surface: ASPH Strength: PCN 120 F/A/W/T **TM** Width: 38 M Surface: ASPH Strength: PCN 120 F/A/W/T **TN** Width: 41 M Surface: ASPH Strength: PCN 14 F/C/W/T **TP** Width: 38 M Surface: ASPH Strength: PCN 56 F/B/W/T **TQ** Width: 38 M Surface: ASPH Strength: PCN 79 F/A/W/T **TR** Width: 38 M Surface: ASPH Strength: PCN 113 F/A/W/T **TS** Width: 75 M Surface: ASPH Strength: PCN 111 F/A/W/T **TU** Width: 51 M Surface: ASPH Strength: PCN 120 F/A/W/T **TV** Width: 19 M Surface: ASPH Strength: PCN 87 F/A/W/T **TW** Width: 44 M Surface: ASPH Strength: PCN 120 F/A/W/T
3	**Localizzazione/Elevazione ACL** NIL	**ACL location/Elevation** NIL
4	**Punto di controllo VOR/INS** NIL / NIL	**VOR/INS checkpoints** NIL / NIL
5	**Note** NIL	**Remarks** NIL

5.2.2 RADIO AIDS AND SURVEILLANCE SYSTEMS

LIPE is equipped with Primary and Secondary Surveillance Radar and with Surface Movement Radar (SMR). The PSR/SSR version is ATCR 33/S, which is Mode-S equipped. The range of the PSR covers about 65 nm and the range of the SSR is about 110 nm; the antennas are located together with and rotate every 4 seconds. Mode S information is displayed in a specific window of the Controller Working Position (CWP) and includes several data, including aircrafts' Call-sign, Indicated Air Speed, Heading, Flight Level, etc. The SMR provides controllers with aircraft and vehicles position on the manoeuvring area. Specific labelling is available on the CWP for identification. The SMR has a range of 3.5 nm and provides raw video information. The SMR is also able to detect foreign objects and flocks of bird on the runway.

Both runways are equipped with ILS. Runway 12 is ILS CAT III B and runway 30 is ILS CAT I.

5.2.3 METEOROLOGICAL SYSTEMS

Meteorological information is available to air traffic controllers through tower equipment. This includes:

- Wind (direction and intensity, both average and instant value)
- Pressure (QNH, QFE)
- Temperature
- Dew Point
- General visibility and Runway Visual Range (RVR) at Touch Down Zone (TDZ), Midpoint (MID) and STOP/END.
- Cloud base

The same information is also available to pilots via the ATIS.

5.2.4 DATA SUPPORTING SYSTEMS

Data supporting systems, such as FDP (Flight Data Processing) and AOIS (Aeronautical Operational Information system) provide the controller with a set of information related to scheduled times, including:

- EOBT/TOBT: Estimated (Target in case of A-CDM) Off-Block Time.

- ETOT/TTOT: Estimated (Target in case of A-CDM) Take Off Time.

- CTOT: Calculated take off time, that is provided by the CFMU (Central Flow Management Unit)

From a controller perspective, the most significant scheduled times are the EOBT and the CTOT.

The FDP also provides controllers with route and clearance information for all IFR flights.

NOTAMs (NOTice to Air Man) are available to controllers through the AOIS, together with ALT (Actual Landing Time) and ATOT (Actual Take Off Time).

5.2.5 LOCAL RULES

The use of taxiways is regulated via some restrictions:

- TWY F and G shall be used only as exit taxiways
- TWY B and D shall not be used to enter the RWY 12 and perform backtrack
- TWY G is a rapid exit taxiway: max speed 93 km/h
- Minimum thrust requested to pilots on all taxiways/taxilanes.
- RWY 30 shall be used only if RVR (TDZ, MID and STOP/END) is equal or greater than 550m.

Also, some restrictions apply depending on the ICAO code of the aircraft:

- Aircraft with ICAO Code F shall use only taxiway A, J and K to enter the runway: A to enter runway 12, J as preferential to vacate runway 12 and K to enter runway 30.
- Taxilane Z shall be used by aircraft up to ICAO code C between TQ and TS
- aircrafts with ICAO code letter D can simultaneously taxi on TWY T and aircraft stand taxilane Z only together with aircrafts with ICAO code letter A
- aircraft with ICAO code letter E shall not taxi on aircraft stand taxilane Z. Taxiing on TWY T and aircraft stand taxilane Z simultaneously with any other aircraft is forbidden
- aircraft with ICAO code letter F shall not taxi on aircraft stand taxilane Z. Taxiing on TWY T and aircraft stand taxilane Z simultaneously with any other aircraft is forbidden
- aircraft with ICAO code D, E, F parked on stand 114 or 115 shall be pushed-back on TWY T through TWY TS
- aircraft with ICAO code E and F parked on apron 3 shall be pushed-back on TWY T through TWY TW
- Use of taxilane N is allowed only for aircraft up to ICAO code B included

- aircraft with ICAO code letter D shall not taxi on the aircraft stand taxilane Z between apron holding points Q2 and S1
- aircraft with ICAO code letter E shall use TWY TU/TS/ TW as exit/entry TWY from/to aprons
- aircraft with ICAO code letter F shall use TWY TS as exit/entry TWY from/to stands 114 and TWY TW as exit/entry from/to Apron 3.

5.2.6 LOCAL RULES IN LOW VISIBILITY CONDITIONS

In case of poor visibility, a reduced airport capacity can be expected due to the increased spacing between arriving/departing aircrafts and/or restrictions applied to ground movements. From a pilot perspective, Low Visibility Procedures activate at the following conditions:

a) RVR TDZ is 550 m or below

b) Cloud base height/ceiling is below 200ft according to the meteorological local report

c) When there is a rapid deterioration of weather conditions.

Pilots will be informed by ATIS and/or by appropriate frequencies when LVP are in force.

From an ATCO perspective, three cases are distinguished:

1. **Low Visibility Condition 1 (LVC 1)**: the control tower can visually monitor all the traffic on the manoeuvring area and pilots are able to taxi maintaining visual reference with other traffic.

2. **Low Visibility Condition 2 (LVC 2)**: all or a part of the manoeuvring area cannot be visually monitored from the TWR but aircraft are able to taxi maintaining visual reference with other traffic.

3. **Low Visibility Condition 3 (LVC 3)**: the RVR at TDZ, MID or STOP/END is less than 400 m.

It is important to note that LVC 1 and 3 are defined by measured values while the shift from LVC 1 and LVC 2 is decreed by the controller.

5.2.6.1 LOW VISIBILITY CONDITION 1 (LVC 1)

- Runway 12 is used preferentially and it is mandatory if RVR is less than 550m.
- Arriving aircraft vacate runway 12 only via taxiway G, H and J and runway 30 only via B.
- Departing aircraft enter runway 12 only via A and runway 30 via J.
- The stopbar at the Runway Holding point CAT II and III are activated.
- Minimum spacing between arriving aircraft is 10 nm in case of no departures and 12 nm in case of departure.
- To ensure that the radio path of the ILS is free, the TWR controller will clear for take-off a departure only if it will overfly the Localizer (LOC) antenna before the arriving aircraft is 4 nm on final.

5.2.6.2 LOW VISIBILITY CONDITION 2 (LVC 2)

- Only runway 12 is used.
- Intermediate holding point (IHP) T1 on main taxiway is activated
- The follow-me is positioned on the taxiway T abeam TS on TWR request in case of arrival.
- Departing aircraft taxi to IHP T1 initially and then to RHP A. Further departures start taxi only once the previous one is between T1 and RHP A.
- Arriving aircraft vacate the runway only via J and follow the follow-me until the parking.
- Push back operations are allowed only from stand belonging to not contiguous blocks.
- Minimum spacing between arriving aircraft is 15nm in case of no departures and 16nm in case of departure.
- To ensure that the radio path of the ILS is free, the TWR controller will clear for take-off a departure only if it will overfly the LOC antenna before the arriving aircraft is 4 nm on final.

5.2.6.3 LOW VISIBILITY CONDITION 3 (LVC 3)

- Only runway 12 is used.
- Intermediate holding point (IHP) T1 on main taxiway is activated

- The follow-me is positioned on the taxiway T abeam TS on TWR request in case of arrival.
- Departing aircraft taxi to IHP T1 initially and then to Runway Holding Point (RHP) A. Further departures start taxi only once the previous one is between T1 and RHP A.
- Arriving aircraft vacate the runway only via J and follow the follow-me until the parking.
- Push back operations are allowed only from stand belonging to not contiguous blocks.
- Minimum spacing between arriving aircraft is 15 nm in case of no departure and 16 nm in case of departure.
- To ensure that the radio path of the ILS is free, the TWR controller will clear for take-off a departure only if it will overfly the LOC antenna before the arriving aircraft is 4nm on final.

5.3 DATA SOURCES AND SENSING SYSTEMS

The data requirements for the implementation of the ARTTs vary depending on the tool and the prototyping environment (real or simulated). This research has tried to identify a set of sensing technologies and data provisioning standards on which the ARTT could be built. These have been grouped into three major categories:

1. Operational data. This is information used by the controller as part of his or her tasks, such as aircraft positions and identifications, flight schedule information or weather data.

2. User data. This is the input provided to the ARTT about the user's position, gaze or actions. This information is typically obtained from sensors worn by the user or installed in the working/prototyping environment.

3. Auxiliary data: information which is not necessarily of direct use from an operational perspective, but can provide context to other information. This also includes the information necessary to build the airport digital model.

For each category, several data sources are reviewed, with a focus on their applicability in the ARTT research. Also, the maturity of each technology will be discussed. While most of these could be said to have a Technology Readiness Level (TRL) of 9, many have not been used in the ATC environment or in the proposed manner. For the fully developed sensing technologies, the maturity level will focus on these aspects.

For simplicity reasons, only some of the data sources and tools listed hereafter have been used during the implementation phase of this thesis.

5.3.1 OPERATIONAL DATA

5.3.1.1 SKY RADAR SYSTEMS (PSR, SSR, MULTILATERATION)

Radar systems, or more generally the use of radio waves, are widely used as the main source for aircraft positional information (surveillance). These systems can be found in many forms such as Primary Surveillance Radar, Secondary Surveillance Radar, Mode A, C and S transponders and Multilateration (MLAT). They all serve to show the location of the aircraft, and some, such as Mode A, C and S, contain information related to the specific aircraft, such as the identification code. This information is necessary to properly place the aircraft related information and overlays in the correct location within the controller's field of view.

Using radio wave signals for surveillance is one of the oldest remote sensing technologies and is at a very mature state as it is used all around the world. The main benefit of these technologies is that they are already in use at most medium and all large airports and would not require any additional investment for their use. The drawbacks are that they are expensive to maintain, and the ground systems are not available in many smaller airports.

The Eurocontrol standard for exchanging ATM surveillance data (including RADAR) is ASTERIX[13]. ASTERIX category 240 defines the transmission format for raw RADAR video. Visualizations of ASTERIX cat 240 data typically mimic a Plan Position Indicator (PPI) display, which is of limited use to the V/ARTT. Other ASTERIX data categories such as 010 and 048 provide interpreted data which is more suitable to future research.

---

[13] http://www.eurocontrol.int/services/asterix

5.3.1.2 GROUND RADAR SYSTEMS (SMR, A-SMGCS)

The SMR provides controllers with positional information of both static and moving objects on the airport surface, i.e. apron and manoeuvring area.

A-SMGCS[14] (Advanced Surface Movement Guidance and Control System) adds to the capabilities of the SMR the identification of transponder equipped aircrafts and provides alerts for runway incursions and other safety related events. It has 4 different levels of implementation:

- **A-SMGCS Level 1** (improved Surveillance) makes use of improved surveillance. The controllers are given traffic position and identity information.

- **A-SMGCS Level 2** (Surveillance + Safety Nets) adds safety nets which protect runways and designated areas. Appropriate alerts are generated for the controllers in case of conflicts between all vehicles on runways and the incursion of aircraft onto designated restricted areas.

- **A-SMGCS Level 3** (Conflict Detection) involves the detection of all conflicts on the movement area as well as improved guidance and planning to be used by controllers.

A-SMGCS Level 4 (Conflict Resolution, Automatic Planning & Guidance) provides resolutions for all conflicts and automatic planning and automatic guidance for the pilots as well as the controllers.

A-SMGCS is currently in the process of deployment throughout Europe. Level 1 is seeing delays in deployment, so it is safe to assume that level 2 deployment will also be delayed, but it is difficult to predict how long the delays will be. The incorporation of A-SMGCS functions on the outside of the tower view is highly desirable. However, it is expected that the research on ARTT tools will be firstly look at already available and unprocessed data before considering advanced safety nets ad data processing systems. A drawback of considering A-SMGCS inputs in this research would be that such systems are not being planned for installation in smaller airports.

---

[14] http://www.skybrary.aero/index.php/Advanced_Surface_Movement_Guidance_and_Control_System_(A-SMGCS)

5.3.1.3 LIGHT DETECTION AND RANGING (LIDAR)

LIDAR (Light Detection and Ranging) is a measurement of distance obtained by scanning an object (even small particles) or an area with a laser. It has been used to profile clouds, measure winds, and study atmospheric contamination. It can do this by measuring the backscatter in the atmosphere or the scattered reflection on the ground. Doppler LIDAR can be used to measure wind speed, turbulence and wind shear, all of which can be useful to the tower controller, especially the turbulence, which cannot be obtained through SWIM. While the specific implementation as a controller tool is yet to be applied, Doppler LIDAR has been used for years to measure wind and turbulence data. The data supporting the RECAT15 wake turbulence re-categorisation of aircraft was obtained from Doppler LIDAR systems at airports both in the US and in Europe. Doppler LIDAR could potentially give the controllers a view of where the wake turbulence actually is behind an aircraft, providing a possible safety and efficiency benefit. While the technology behind Doppler LIDAR is mature and in use in many areas, including at some airports, the specific implementation as a sensing technology for controller tools has yet to be done.

5.3.1.4 AUTOMATIC DEPENDENT SURVEILLANCE – BROADCAST (ADS-B)

ADS-B (Automatic Dependent Surveillance – Broadcast) is a system that uses transmissions from aircraft to provide other actors geographical position, pressure altitude data, flight identity, 24 bit transponder address, velocity and other data which have been determined by airborne sensors. Typically, the airborne position sensor is the output a GPS receiver or the output of a Multi-Mode Receiver (MMR). Integrated GPS and inertial systems are also used. ADS-B also provides integrity data that indicates the containment bound on positional errors. The altitude sensor is typically the same barometric source used for SSR (Secondary Surveillance Radar). Currently inertial only sensors do not provide the required integrity data although these are likely to be provided in the future. An ADS-B ground system uses a non-rotating antenna positioned within a coverage area, to receive messages transmitted by aircraft. The ADS-B ground system does not necessarily transmit anything. ADS-B receiver ground stations are the simplest and lowest cost installations of all options to provide air-ground surveillance, although costs may increase if ADS-B transmitter (to broadcast or rebroadcast ADS-B data, e.g. TIS-B[16], ADS-R[17] or FIS-B[18]) capabilities are deemed necessary. An ADS-B receiver is typically less than six inches high by nineteen inches wide and a duplicated site consumes less than 200 watts of electricity. An ADS-B ground station can normally

---

[15] http://www.skybrary.aero/index.php/RECAT_-_Wake_Turbulence_Re-categorisation

[16] Traffic Information Service – Broadcast

[17] Automatic Dependent Surveillance – Re-broadcast

[18] Flight Information System – Broadcast

be installed in an existing VHF (Very High Frequency) communications facility. ADS-B is becoming a mandatory piece of equipment for new aircraft as part of Single European Sky regulation and it is being used in Canada as part of air traffic control in certain areas. Its use is also mandatory in parts of Australia. The most important information provided by ADS-B to a Control Tower environment is the aircraft position and identification.

Like RADAR data, ADS-B information can be supplied to the ARTT in ASTERIX format (category 021).

The main benefits and drawbacks of the ADS-B system are summarized in table below:

BENEFITS	DRAWBACKS
➢ Simple ground station receiver design, low cost (but highly variable ADS-B transmitter avionics fitment cost). ➢ Can be installed at sites together with other equipment. ➢ Very high update rate and resolution. ➢ High accuracy and integrity of transmitted data (airborne measurements). ➢ Accuracy not dependent on range ➢ Facilitates future provision of innovative ATM services based on air-to-air ADS-B.	➢ Collaborative surveillance system ➢ Dependent on aircraft avionics ➢ Equipage rates are relatively low at this stage ➢ Requires unobstructed view to aircraft.

Table IV. BENEFITS AND DRAWBACKS OF THE ADS-B SYSTEM

5.3.1.5 VISIBLE LIGHT OR INFRARED LIGHT CAMERAS

From an ATC perspective, visible and infrared light cameras have no operational application in local control towers. They are basically used for monitoring apron activities such as boarding, refuelling, aircraft loading, etc. On the contrary, the use of cameras becomes mandatory in the remote towers applications where the out of window's view is replaced by environmental video. In this case, the air traffic control service is provided using cameras. However, in some areas the camera output is inferior to the human eye capabilities and suffers from drawbacks that have a negative impact on the ability to provide air traffic services. One example is the ability to provide stereoscopic 3D visualization, which is currently absent in remote towers applications. One of the situations in which it is hard for a camera to perform is when it's faced with different light conditions in the image (e.g. a bright sky and dark ground). Another problem derives from the fact that each camera in an array produces different results since they all are faced with different light conditions. Future research should strive for automation to achieve good image quality; however, it is foreseen to be a need for manual intervention. This must be comprehensible to the ATCO; otherwise it will not be used. This semi-automatic configuration should be better than an automatic approach as there is a direct relation between the manual settings for the image and the awareness of applied enhancements to the picture. For example, if obligated to switch between day and night settings, the ATCO might be more aware of the current visual condition. Contrariwise, in the automatic approach, where a digital camera has automation for exposure times, shutter, ISO, gain etc. the ATCO could be not aware of the current visual condition. The risk otherwise is that the ATCO would make decisions based on *too good* information and/or give confusing directions. "Behind landing aircraft, line up runway..." is not helpful to the pilot if the landing aircraft can only be seen by the ATCO. To switch between enhanced and normal image on the press and release of a button is one possible way to gain from the benefits while still being aware of the actual visual conditions. Visible light cameras can be used in conjunction with AR technologies to improve the human eye control, particularly to avoid blind spots covered by airport features or barely visible due to long distances.

Infrared imaging provides a thermo-graphic representation of the focused. Our eyes are detectors that can only see light in some parts of the light spectrum. These parts are hence defined as the visible spectrum. There are other forms of light that the human eye cannot see. At one end of our visible range is the ultraviolet light, and in the other end is infrared light. An infrared imaging camera produces an image based on the differences in infrared radiation intensity that an object emits. All objects with a temperature above absolute zero, emit radiation visible by an infrared camera. Therefore, an infrared camera isn't affected by a dark environment. An example of the image provided by an infrared camera is shown in Figure 44.

Figure 44. LEFT IMAGE: INFRARED CAMERA VIEW. RIGHT IMAGE: INFRARED CAMERA VIEW MERGED WITH THE OUTSIDE OF THE TOWER VIEW IN LOW VISIBILITY CONDITIONS (FOG).

The main benefit of an infrared camera is to increase the ATCO situational awareness during night time, and in fog, to increase the overall safety and stretch the LVP boundaries. Infrared camera usage in fog is especially useful since it can increase visibility. The position of vehicles can be visually confirmed, unauthorized movements can be detected and wildlife incidents can be avoided. By using this technology all airports could expect quicker and more efficient runway checks during LVC.

Ideally, cameras could be placed in such a way that the view could be switched between regular cameras and infrared vision. Both could be attached to a manoeuvrable tilt-zoom mechanism. The resulting view could be presented either on a separate screen, or overlaid in the out of the window view, at the position where the camera is pointed.

Technology for the streaming of live video to multiple clients is mature and commonplace. There are no specific technical caveats when providing video input to the ARTT. The main benefit of these technologies is that they could be adopted by airports with no impact on the current infrastructure but only by adding an additional infrastructure to be integrated in the current system. The main drawback of this solution is that it may suffer from poor image quality.

5.3.1.6 METEOROLOGICAL DATA

Met data provision and access is an essential aspect of the SESAR System Wide Information Management (SWIM[19]) implementation. SWIM is a SESAR initiative which creates an all-encompassing set of data exchange standards for the ATM domain. By sharing information between stakeholders in a uniform and standardized way, SWIM strives to improve the safety and cost efficiency of ATM operations. By enabling optimized air traffic movement and infrastructure usage, SWIM aims to reduce the

---

[19] http://www.eurocontrol.int/swim

environmental impact of ATM operations. In the context of the ARTT research SWIM will be considered to ensure optimal interoperability between the AR tools and external systems or services.

Meteorological data can be provisioned using IWXXM[20] (ICAO Weather Information Exchange Model). This format is one of the primary candidates for Met data exchange in a SWIM-enabled environment and is maintained by the World Meteorological Organization (WMO) and by ICAO. It is based on ICAO's meteorological requirements with respect to METAR (METeorological Air Report), SPECI (Special Weather Report), TAF (Terminal Aerodrome Forecast) and SIGMET (SIGnificant METeorologic information) weather data messages. Version 2.0 adds AIRMET, Tropical Cyclone Advisory and Volcanic Ash Advisory data products. A modified version of IWXXM is WXXM, a standard jointly developed and maintained by Eurocontrol and the FAA. This format extends IWXXM capabilities adding supplementary types of weather information not covered in IWXXM.

Being based on OGC GML (Open Geospatial Consortium - Geography Markup Language), IWXXM and WXXM are well-suited for distribution through OGC web services, also embraced by SWIM, including WFS (Web Feature Service), WMS (Web Map Service) and WMTS (Web Map Tile Service). By definition, each OGC web service type has its own characteristics and suitable data model exchange types. WMS and WMTS can be used to access rendered versions of the data, using bitmap formats such as JPEG and PNG. WFS on the other hand focuses on the exchange of GML-based vector data such as IWXXM, WXXM.

To ease the discovery of actual Met (and other aviation-related) data and services in a SWIM environment, SESAR deployed an online catalogue, called the SWIM Registry[21].

The provision of MET data to the control tower is a mature service. What is not mature yet, is providing this service via SWIM, or the integration of this data into a visualization tool that is not a head down screen such as an overlay of the radar screen. To put MET data such as QNH, wind (speed and direction), ceiling (type and altitude), RVR, wake vortexes, etc. could be helpful for reducing controllers' head down time. The drawback of having this type of information in the controller's field of view is that it may crowd out more important information. This trade off would have to be validated.

---

[20] http://www.wxxm.aero/

[21] http://eur-registry.swim.aero

5.3.1.7 Flight Data Systems (FDP, AOIS)

Flight data systems, such as the Flight Data Processing System (FDP) and the Aeronautical Operational Information System (AOIS) provide the controller with information related to the scheduled times and routes. Data supporting system information is available via the CWP and typically requires controllers head down operations. The integration of this information in the ARTT is very useful to reduce the head down operations and improve the information accessibility. For example, the provision of the CTOT on the identification label of a flight should reduce mistakes and clearly inform the controller about the time constraints.

FIXM22 (Flight Information eXchange Model) is a cross-domain standard for the exchange of flight information. Eurocontrol has defined an A-CDM extension23 to the standard, which covers the scheduled times information. FIXM defines a conceptual model as well as an eXtensible Markup Language (XML) based physical representation of it.

5.3.2 User Data

User input data is a valuable resource for interactive augmented reality. First, it is used to let the user interact with the application, e.g. through mouse, keyboard or gesture input. Second, it contains some information on the user state, including position in space, head tracking and ideally, eyes vergence and accommodation. This information is essential for augmented and virtual reality application, and will be briefly discussed in the next paragraphs.

5.3.2.1 User Location

The physical location of the user can be either a geospatial location or a relative location in a room. For the research on ARTT it is convenient to consider the location to be the location of the user inside the control tower. Location data can have multiple purposes. First, it can be used by the application to adapt itself to the changing context depending on where the user is currently positioned. For instance, depending on which side of the control tower the viewer is located, different runway information might be required. Second, it can be used to avoid visualization conflicts when two or more viewers are looking at the same AR (spatial) device.

---

[22] https://www.fixm.aero/

[23] https://www.fixm.aero/eurocontrol_extension_10.pl

5.3.2.2 HEAD TRACKING

Head tracking provides information about the orientation and position of the user's head. Head tracking typically has six degrees of freedom, three for translation by the axis and one for each rotation around the different axis. The rotations are typically indicated as pitch, yaw and roll angles. Two important properties that should be considered when selecting the head tracking technology are the latency and the accuracy. If there is a significant latency between an action of the user (e.g. looking to another direction) and the result of that action it will break the AR overlay registration and may impact the user's performance. On a technical level, most head tracking solutions provide direct access to the position of the head via either positional vector, transformation matrix or another mathematical representation. When dealing with HMDs, this data can be used directly by the rendering engine to position the cameras correctly, while the head orientation is used to orient it based on the direction in which the viewer is looking. When dealing with spatial displays the projection model must be continuously updated based on the relative position between the screens and the viewer's eyes. For this reason, it would be best to know the position of the viewer's eyes instead of the head. If this data is not directly available it is typically derived by considering the user head orientation and separating the left and right eyes by the standard interpapillary distance.

If necessary, the position and orientation of the head can be combined with the location data described above.

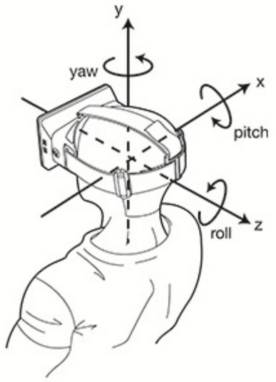


Figure 45. DEGREES OF FREEDOM OF THE HEAD MOTION

5.3.2.3 VERGENCE AND ACCOMMODATION

A measure of the vergence and accommodation distances of the eyes is rarely made available to the AR developer. This is mostly because of the difficulty of retrieving this data. Nevertheless, this information would be very useful, as it could be used to position the AR content next to the point the viewer is looking at and generally speaking inside his or her stereo comfort zone [185].

5.3.3 AUXILIARY DATA

Because the use cases for the auxiliary data sources are strongly intertwined, this section will discuss integration of the data into the ARTT in a separate paragraph at the end, rather than for each data source individually.

5.3.3.1 3D CONSTRUCTIONS MODELS

Many landmarks, public buildings and infrastructures have been modelled in 3D and are freely or commercially available from various online sources. Two popular examples of such sources are 3D Warehouse[24] and Turbosquid[25]. However, the quality of models obtained from these services is highly variable.

Some cities provide access to CityGML[26] data. CityGML is an OGC standard for the modelling of urban environments and typically cover larger areas (i.e. many buildings) within a single data set. CityGML supports models with multiple levels of detail (LOD). The highest detail level, if available, also includes building interiors. Larger-scale models such as CityGML are more commonly reconstructed from other data (e.g. LiDAR or aerial photography) using automated processes. The main downside to CityGML is that it is that public availability is still very limited.

Many file formats exist for the exchange of 3D models. Popular but dated formats for software interoperability include WaveFront OBJ[27] (.obj) and VRML[28] (.wrl). X3D[29] (.x3d) is another XML-based format which was built with the aim of extending VRML capabilities. At present, however, Collada[30] (.dae), 3DS[31] (.3ds) and FDX (.fdx) are

---

[24] https://3dwarehouse.sketchup.com

[25] http://www.turbosquid.com

[26] http://www.citygml.org

[27] https://en.wikipedia.org/wiki/Wavefront_.obj_file

[28] https://it.wikipedia.org/wiki/VRML

[29] http://www.web3d.org/x3d-vrml-most-widely-used-3d-formats

[30] https://www.khronos.org/collada

widely perceived as the industry standard. Collada also integrates with KML[32] (.klm), that allows for precise geolocation of 3D models. KML has been popularized by Google Earth and later adopted as an OGC standard[33]. The authoring of 3D models in any of the aforementioned formats can be done using a variety of different tools. Manual authoring of models is commonly done using one of the many free and commercial 3D modelling packages, such as SketchUp[34], Blender[35] or 3DS Max[36].

5.3.3.2 ELEVATION MODELS

Elevation models are widely and freely available from various sources. Popular data sets with worldwide coverage include ETOPO[37] and SRTM[38]. Many national or local governments have GIS (Geographic Information System) portal sites, through which they provide access to geospatial data for their region. Such portal sites will often carry DEM data that is more detailed than the aforementioned sources.

Popular file formats for elevation data include DEM (Digital Elevation Model), DTED (Digital Terrain Elevation Data) and GeoTIFF. In addition to file-based distribution, elevation data can be served using an OGC Web Coverage Service (WCS)[39].

5.3.3.3 AERIAL AND SATELLITE IMAGERY

Like elevation data, aerial and satellite imagery is widely available from various sources. NASA publishes data from its various satellite missions such as LandSat[40]. Again, government GIS portals will often provide more detailed regional data.

GeoTIFF, JPEG2000, ECW are some of the formats that can be used to store imagery. The OGC Web Map Service (WMS)[41], Web Map Tile Service (WMTS)[42] and Web Coverage Service can all be used to distribute imagery to multiple clients. Services such

---

[31] https://en.wikipedia.org/wiki/.3ds

[32] https://developers.google.com/kml/

[33] http://www.opengeospatial.org/standards/kml

[34] http://www.sketchup.com/

[35] https://www.blender.org/

[36] http://www.autodesk.com/products/3ds-max/overview

[37] https://www.ngdc.noaa.gov/mgg/global/global.html

[38] http://www2.jpl.nasa.gov/srtm/

[39] http://www.opengeospatial.org/standards/wcs

[40] http://landsat.gsfc.nasa.gov

[41] http://www.opengeospatial.org/standards/wms

[42] http://www.opengeospatial.org/standards/wmts

as Bing Maps[43] and Google Maps[44] aggregate imagery from various sources and provide a unified access point.

5.3.3.4 AIRPORT LAYOUT

The aforementioned data sources can all contribute to a faithful 3D reproduction of the airport environment, but they do not provide any semantic information that may be valuable to programmers and controllers. For instance, a software system cannot readily identify buildings described in a 3D model, nor can it easily distinguish taxiways, runways, or other important airport features in an aerial photograph. Also, there are no dynamic features or variable information associated to 3D objects. To implement that, a developer would have to leverage the capabilities of a 3D engine and manually script objects behaviours and assign properties. Another possibility is to use the industry standard exchange format for dynamic, semantically rich aeronautical data, which is AIXM 5.1[45] (Aeronautical Information eXchange Model). AIXM combines a conceptual model of the aeronautical information domain with a GML[46]-based storage format. The AIXM conceptual model supports aerodrome mapping, obstacle modelling, digital NOTAMs and so on. All features described by AIXM are dynamic features, which use "time slices" to describe the changes of the AIXM feature over time (e.g. a temporary runway closure).

---

[43] https://www.microsoft.com/maps/

[44] https://www.google.it/maps

[45] http://aixm.aero/

[46] http://www.opengeospatial.org/standards/gml

6 IMPLEMENTATION

This chapter describes the starting datasets and the modelling techniques that have been used to reconstruct the reference scenario. It presents the methodology that was used for the selection of the best ARTT and the outcomes of this process. It also describes the first implemented overlays and the undelaying methodology that was used to conceive them. The same methodology was used to highlight the benefits of using AR, which were initially theorized by this research. Finally, it describes the simulation environments were the AR overlays have been integrated and tested.

6.1 MODELLING THE REFERENCE SCENARIO

6.1.1 AIRPORT STATIC FEATURES

For modelling the target scenario, the data contained in LIPE Integrated Aeronautical Information Package (IAIP) was used. This data can be obtained from ENAV[47] (Ente Nazionale per l'Assistenza al Volo) website and contains the geographic coordinates for all the relevant points within Bologna airport.

Satellite imagery was also used, which was downloaded from Google Maps[48] and tiled together using Google Satellite Maps Downloader[49]. The image was later converted in grayscale using Gimp[50].

The data relative to the height of buildings was not available. Therefore, this had to be estimated from Google Earth[51] 3D reconstruction.

---

[47] ENAV is the Italian Air Navigation Service Provider

[48] https://www.google.it/maps

[49] http://www.allmapsoft.com/gsmd/

[50] https://www.gimp.org

[51] https://www.google.it/intl/it/earth/

Several static airport features have been modelled including: the runway, the taxiways, the taxilanes, the apron, the control tower, Terminal 1 and 2, Car parking P2 e P3, Firefighter facilities, SAB[52] operative centre, 21 hangars and several other buildings.

The scenario also includes the runway ground signs, the taxiway centrelines and aircraft parking stands ground signs.

Figure 46. COMBINED SATELLITE IMAGE OF BOLOGNA G. MARCONI AIRPORT: 32800 × 13216 PIXELS

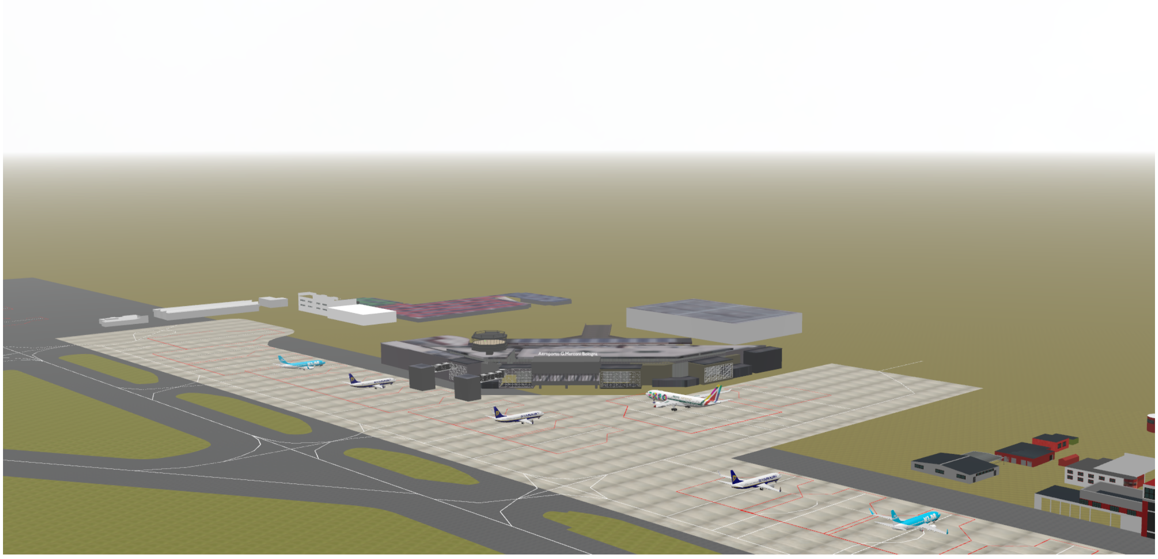


Figure 47. BOLOGNA AIRPORT DIGITAL RECONTRUCTION

---

[52] Società Aeroporto di Bologna. This is the company that manages the Bologna Airport

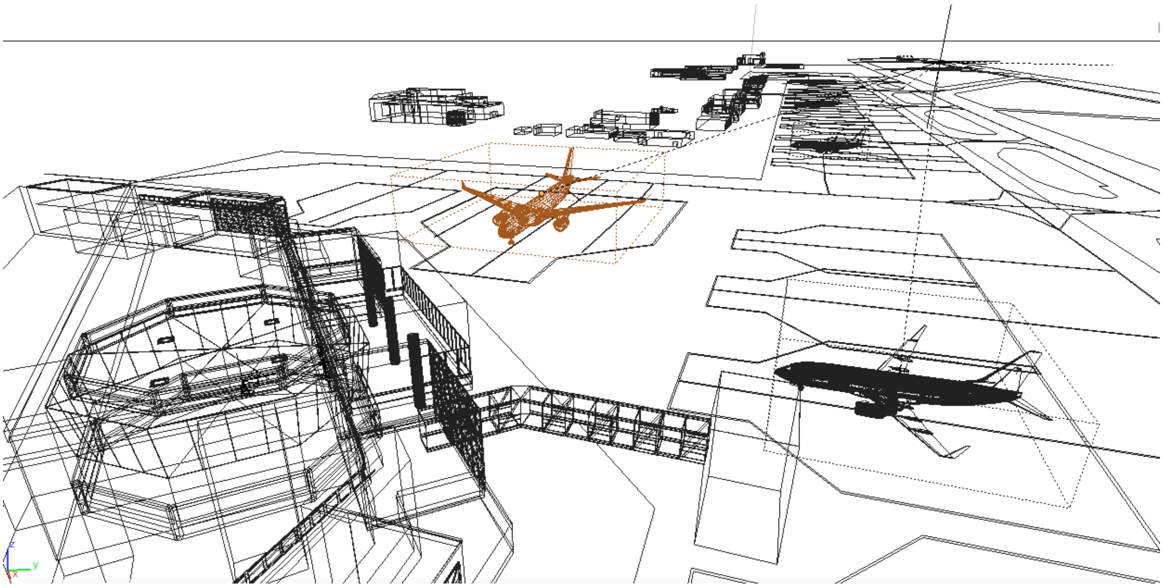


Figure 48. BOLOGNA AIRPORT DIGITAL RECONTRUCTION: WIREFRAME VIEW

6.1.2 AIRCRAFTS

A small library of aircrafts was set up, including:

- Airbus A330-200 Alitalia
- ATR72-600 Jettime
- Boeing 737-100 KLM
- Boeing 737-800 Ryanair
- APM20 Lionceau
- Piper PA18

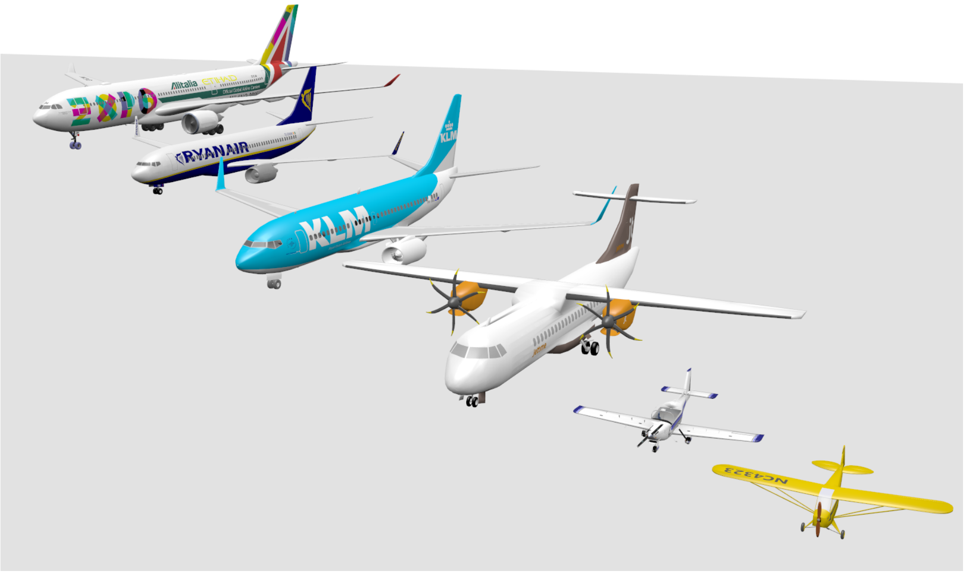


Figure 49. AVAILABLE AIRCRAFT MODELS

6.1.3 ENVIRONMENT AND TERRAIN

The presence of the sky is simulated by means of a skydome superimposed to a background colour. The background controls the colour of the sky while the skydome adds clouds to the scene.

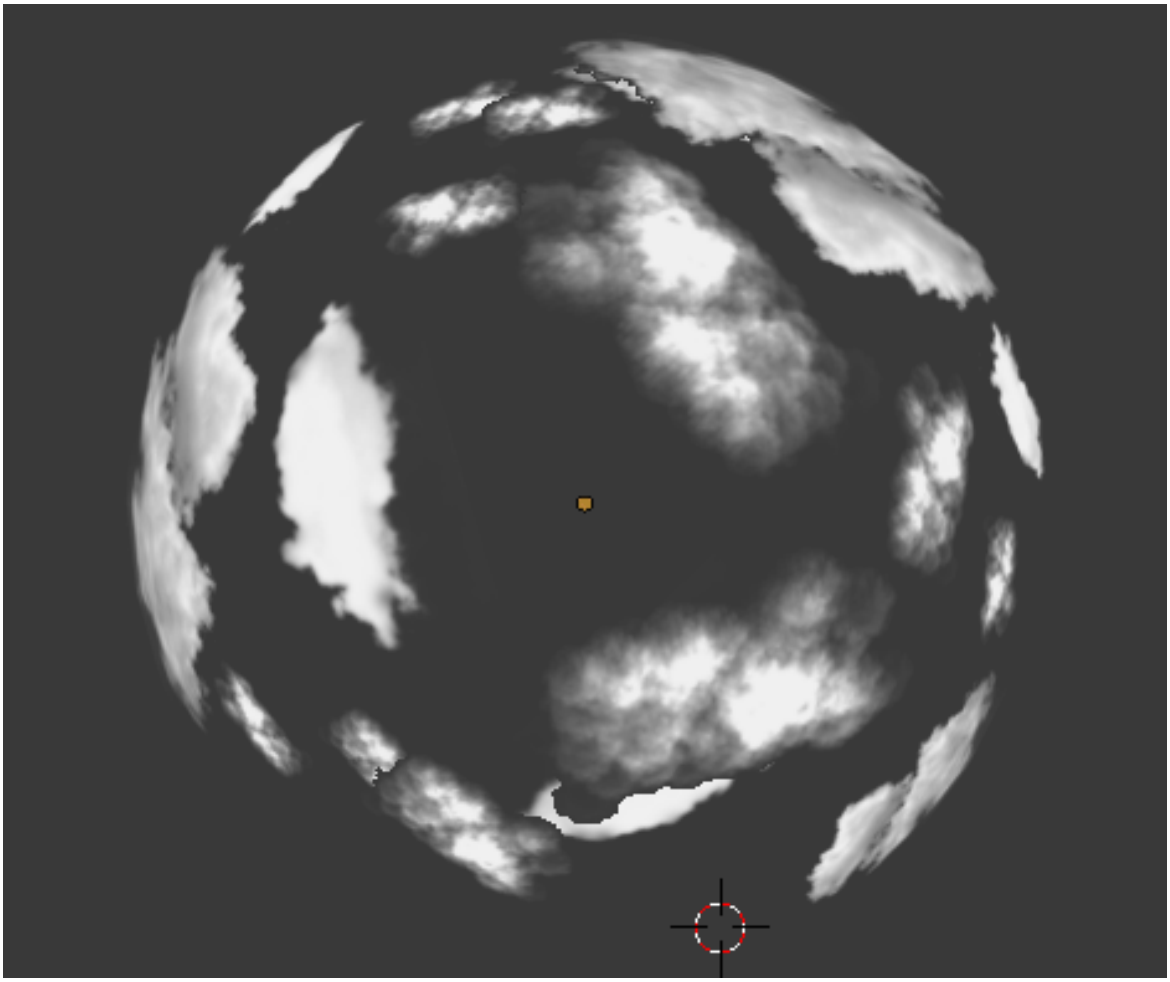


Figure 50. SIMULATED REPRESENTATION OF CLOUDS

LVC, particularly fog, have been implemented using the integrated mist functionality of BGE. Table V shows the appropriate settings for each visibility condition. These can be edited through the Blender user interface.

CONDITION	START (km)	END (km)
LVC 1	0,01	3
LVC 2	0,01	1,5
LVC 3	0,01	0,5

Table V. MIST SETTINGS ACCORDING TO THE LOW VISIBILITY CONDITION

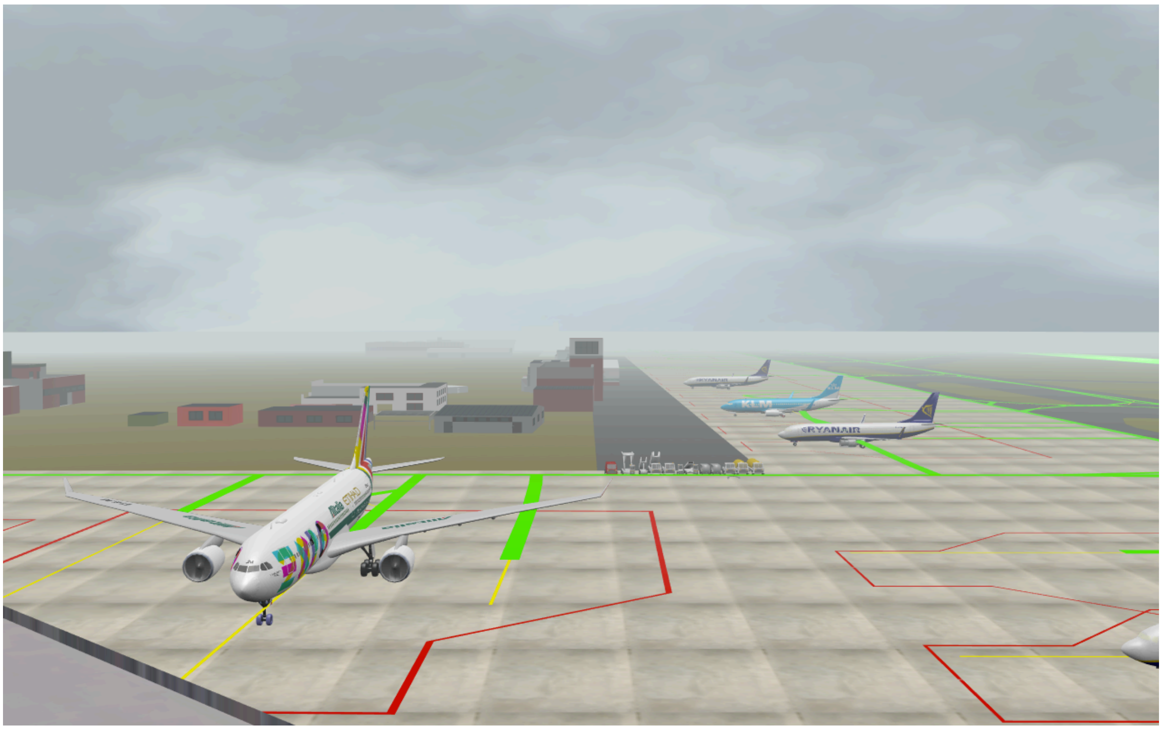


Figure 51. SIMULATED LOW VISIBILITY CONDITION 2

The surrounding orography was derived from a 50x50 km ASTER[53] (Advanced Spaceborne Thermal Emission and Reflection Radiometer) image with 30m resolution. The image was imported firstly imported in Blender and later used to create a vertex displacement in the terrain mesh.

---

[53] https://asterweb.jpl.nasa.gov/gdem.asp

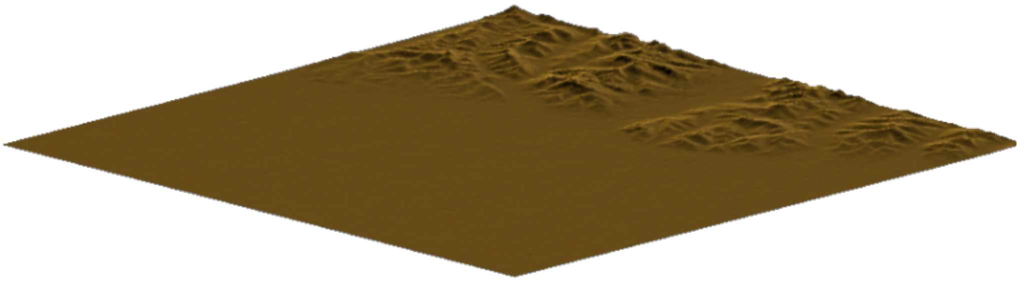


Figure 52. TERRAIN OROGRAPHY SURROUNDING THE BOLOGNA AIRPORT

6.1.4 GROUND MOVEMENTS

A "point and click" interface for managing aircrafts and ground vehicles, i.e. assign taxi routes ad clear take-offs and landings, was developed. The ground movements are based on a custom implementation of the A* pathfinding algorithm which uses designated objects (e.g. taxiways centrelines) as a navigation mesh. A graph of potential waypoints is derived from the navigation mesh ad used as a basis to determine the most convenient path. From a pseudo-pilot perspective, the interface lets you select aircrafts and ground vehicles and assign them with a destination and several intermediate waypoints just by clicking on the airport map.

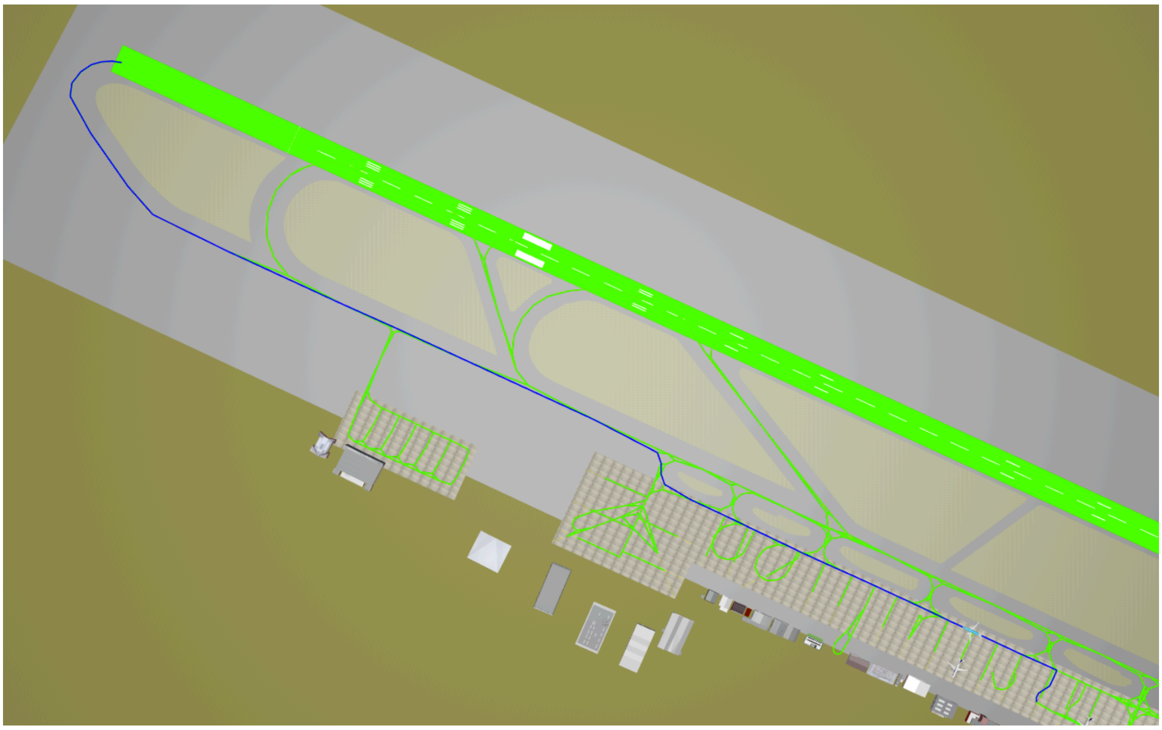


Figure 53. GROUND MOVEMENT SYSTEM: VISUALIZATION OF POTENTIAL (GREEN) AND CALCULATED (BLUE) WAYPOINTS

6.1.5 TAKE-OFF AND LANDING

Take-offs and landings have been implemented through standard animations and blended with the ground movement system by means of the BGE visual programming (logic bricks).

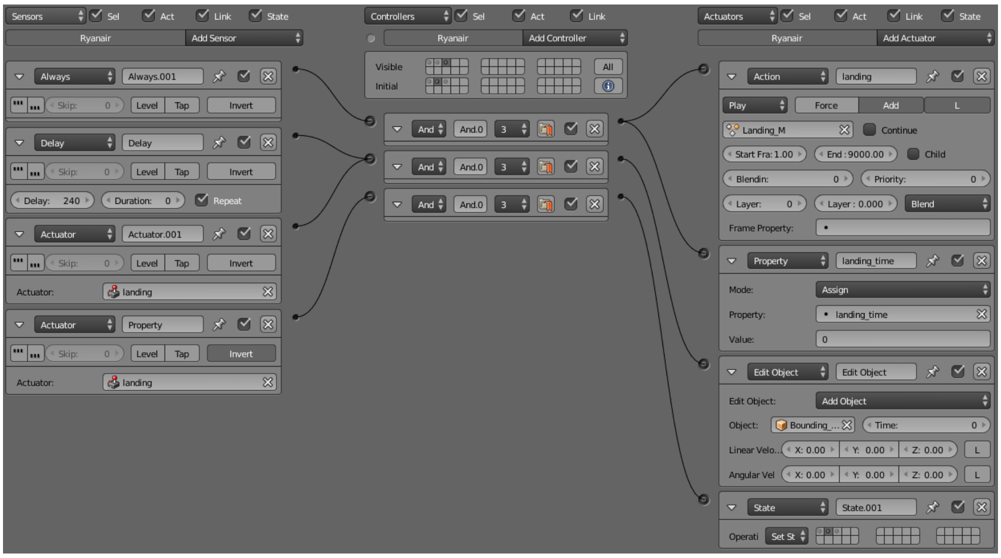


Figure 54. BLENDER GAME ENGINE VISUAL PROGRAMMING (LOGIC BRICKS)

The take-off path was designed considering the approximate lift off distance and lift-off time of a short-medium range aircraft such as the A320 and B737-800. The landing path was designed according to the ILS glide slope of Bologna airport.

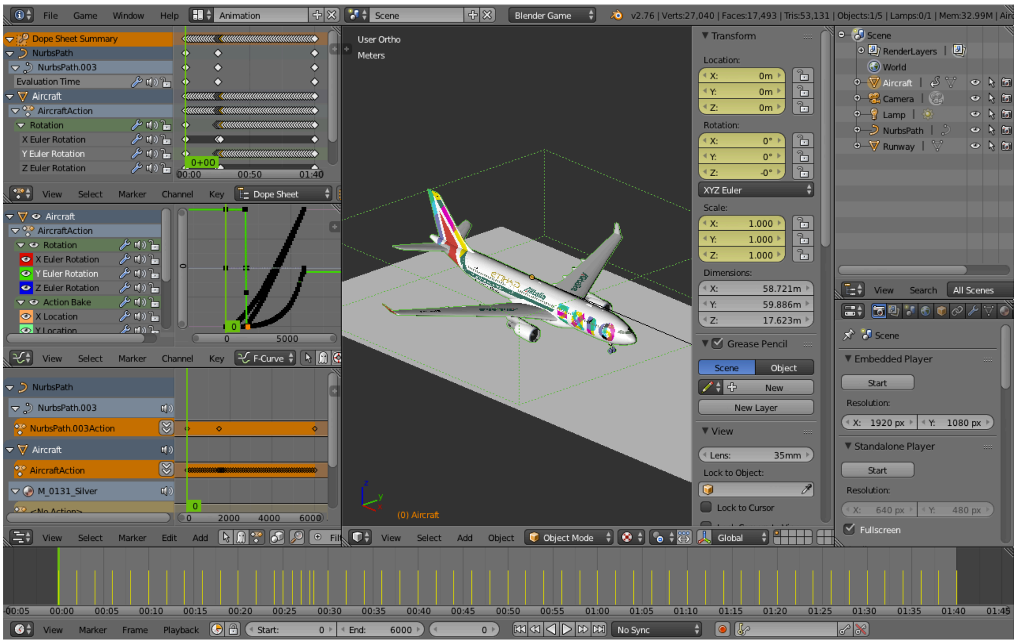


Figure 55. AIRCRAFT TAKE OFF ANIMATION

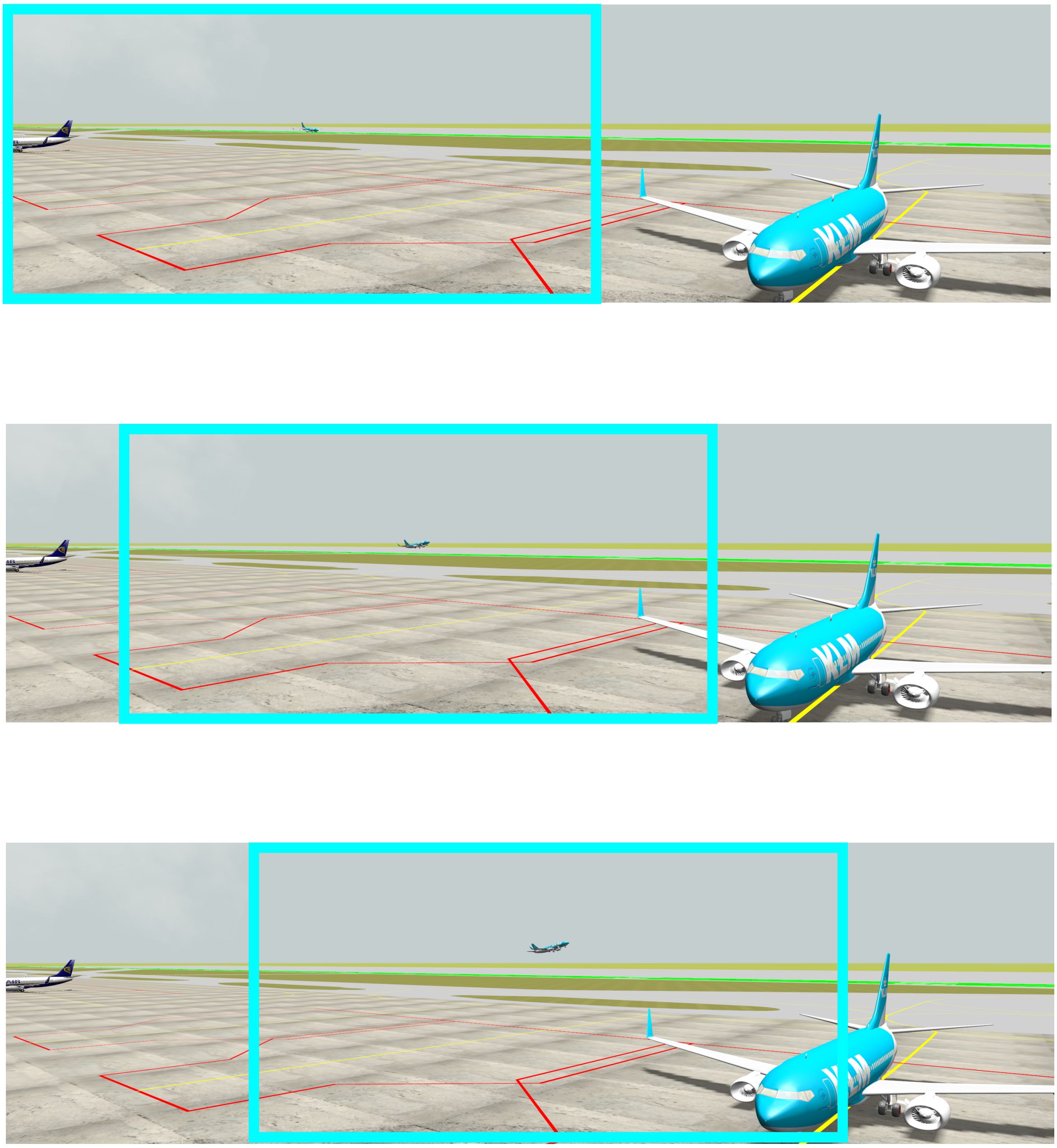


Figure 56. Take-off animation (highlighted by the light blue box), sequence of frames

6.2 SELECTING THE AUGMENTED REALITY TOOLS: AN INTEGRATED QUALITY FUNCTION DEPLOYMENT AND ANALYTIC HIERARCHY PROCESS ANALYSIS

Within the framework of the RETINA[54] Project an integrated Quality Function Deployment (QFD) and Analytic Hierarchy Process (AHP) analysis was performed to select the best AR tools to be used in a control tower [186].

The analysis had two main purposes:

1. Identify and improve a methodology for the selection of the best AR tools to be used in a specific working environment.
2. Apply the methodology to the control tower environment and to available AR devices at their current state.

The aim of the study is obviously related to the scope of this research; therefore, its results will be presented hereafter.

Five different AR technologies were compared against a set of weighted specifications chosen by interface designers (Figure 57 and Table VI).

The weight of each specification ultimately derives from a House of Quality comparison between the very same specifications and a list of requirements provided by air traffic controllers (Table VII).

Figure 57. AUGMENTED REALITY TECHNOLOGIES CONFRONTED IN THE QUALITY FUNCTION DEPLOYMENT - ANALYTIC HIERARCHY PROCESS ANALYSIS

---

[54] http://www.retina-atm.eu

SPECIFICATIONS	MEANING	WEIGHTS[55] (ROUNDED)
PERFORMANCE IN DEPTH CUE PROVISION	monocular, binocular, biocular	13%
OVERLAY SEPARATION	The system can provide separated overlays for different users (each user has its own overlays)	11 %
FOV	Field of view	10 %
RESOLUTION	The ability of an imaging system to resolve detail in the object that is being imaged	10 %
WEARABILITY	Intrusiveness of the device	10%
LATENCY	virtual image delay with respect to real image (including refresh rate, tracking system latency, pixel latency). Very strict hard real-time constraint in digital image processing is mandatory!	9%
WEIGHT	Weight of the device (if not spatial)	9%
BRIGHTNESS, CONTRAST AND LIGHT COMPENSATION	Compensation for tower lighting conditions. They could change from window to window	6%
DISPLAY TRANSMISSIVITY	Display opacity. If transmissivity increases, opacity decreases.	6%
FOV ASPECT RATIO	Ratio between vertical FOV and horizontal one	5%
LAYOUT ADAPTABILITY	adaptability to airport layout changes	5%
CONFIGURATION TIME	configuration time at the beginning of each task	5%

Table VI. SPECIFICATIONS FOR THE AUGMENTED REALITY TOOLS SELECTION

---

[55] Weights are calculated based on a House of Quality confrontation between specifications and requirements, see [186] for further details.

REQUIREMENTS	MEANING	WHEIGHTS [56] (ROUNDED)
Precision	monocular, binocular, biocular	18%
OVERLAY SEPARATION	The system can provide separated overlays for different users (each user has its own overlays)	18%
Clear vision	Field of view	17%
RESOLUTION	The ability of an imaging system to resolve detail in the object that is being imaged	9%
Reactivity	Intrusiveness of the device	9%
LATENCY	virtual image delay with respect to real image (including refresh rate, tracking system latency, pixel latency). Very strict hard real-time constraint in digital image processing is mandatory!	7%
Comfortable (physical comfort)	Weight of the device (if not spatial)	7%
BRIGHTNESS, CONTRAST AND LIGHT COMPENSATION	Compensation for tower lighting conditions. They could change from window to window	6%
No overlapping images for controllers performing different tasks	Display opacity. If transmissivity increases, opacity decreases.	5%
FOV ASPECT RATIO	Ratio between vertical FOV and horizontal one	3%
Easy to setup at each use	adaptability to airport layout changes	3%

Table VII. REQUIREMENTS PROVIDED BY AIR TRAFFIC CONTROLLERS

Table VIII shows the comparison of each technology with the others, performed following the AHP approach. In the comparison matrixes, a negative value shows that the i-th element of the first column (e.g. Head Mounted Display) performs worse than the j-th element of the first row (e.g. Spatial Displays) in terms of element [1;1] (e.g. Resolution). Vice versa, a positive value shows that the i-th element of the first column (e.g. Head Mounted Display) performs better than the j-th element of the first row (e.g. Volumetric Displays) in terms of element [1;1] (e.g. Resolution). The difference in performance is quantified by the absolute value of the registered number, on a scale from -9 to 9.

---

[56] Weights are calculated based on a AHF confrontation between requirements, see [186] for further details.

Resolution	Head Mounted Displays	Spatial Displays	Hend Held Displays	Object-Projected Displays	Volumetric Displays
Head Mounted Displays		-7	-4	-5	5
Spatial Displays			3	6	9
Hend Held Displays				2	8
Object-Projected Displays					6
Volumetric Displays	Incon: 0.10				
FOV	Head Mounted Displays	Spatial Displays	Hend Held Displays	Object-Projected Displays	Volumetric Displays
Head Mounted Displays		-7	3	-5	5
Spatial Displays			9	5	9
Hend Held Displays				-5	2
Object-Projected Displays					5
Volumetric Displays	Incon: 0.10				
FOV Aspect Ratio	Head Mounted Displays	Spatial Displays	Hend Held Displays	Object-Projected Displays	Volumetric Displays
Head Mounted Displays		-7	3	6	7
Spatial Displays			7	9	9
Hend Held Displays				4	4
Object-Projected Displays					2
Volumetric Displays	Incon: 0.10				
Display Transmissivity	Head Mounted Displays	Spatial Displays	Hend Held Displays	Object-Projected Displays	Volumetric Displays
Head Mounted Displays		-5	3	-6	-7
Spatial Displays			8	-2	-3
Hend Held Displays				-7	-8
Object-Projected Displays					-2
Volumetric Displays	Incon: 0.05				
Brightness, contrast and light compensation	Head Mounted Displays	Spatial Displays	Hend Held Displays	Object-Projected Displays	Volumetric Displays
Head Mounted Displays		-7	-3	4	3
Spatial Displays			5	6	7
Hend Held Displays				2	3
Object-Projected Displays					2
Volumetric Displays	Incon: 0.10				
Performaces in depth cue provision	Head Mounted Displays	Spatial Displays	Hend Held Displays	Object-Projected Displays	Volumetric Displays
Head Mounted Displays		-3	5	5	6
Spatial Displays			7	7	9
Hend Held Displays				1	2
Object-Projected Displays					2
Volumetric Displays	Incon: 0.02				
Latency	Head Mounted Displays	Spatial Displays	Hend Held Displays	Object-Projected Displays	Volumetric Displays
Head Mounted Displays		1	1	4	3
Spatial Displays			1	4	3
Hend Held Displays				4	3
Object-Projected Displays					-2
Volumetric Displays	Incon: 0.00				
Wearability	Head Mounted Displays	Spatial Displays	Hend Held Displays	Object-Projected Displays	Volumetric Displays
Head Mounted Displays		-9	-2	-9	-9
Spatial Displays			7	1	1
Hend Held Displays				-7	-7
Object-Projected Displays					1
Volumetric Displays	Incon: 0.01				
Weight	Head Mounted Displays	Spatial Displays	Hend Held Displays	Object-Projected Displays	Volumetric Displays
Head Mounted Displays		-9	1	-9	-9
Spatial Displays			7	1	1
Hend Held Displays				-7	-7
Object-Projected Displays					1
Volumetric Displays	Incon: 0.00				
Layout Adaptability	Head Mounted Displays	Spatial Displays	Hend Held Displays	Object-Projected Displays	Volumetric Displays
Head Mounted Displays		4	1	8	5
Spatial Displays			-4	4	2
Hend Held Displays				8	5
Object-Projected Displays					-3
Volumetric Displays	Incon: 0.02				
Overlay Separation	Head Mounted Displays	Spatial Displays	Hend Held Displays	Object-Projected Displays	Volumetric Displays
Head Mounted Displays		7	3	9	5
Spatial Displays			-3	3	-2
Hend Held Displays				7	3
Object-Projected Displays					-4
Volumetric Displays	Incon: 0.03				
Configuration Time	Head Mounted Displays	Spatial Displays	Hend Held Displays	Object-Projected Displays	Volumetric Displays
Head Mounted Displays		-3	-2	-3	-3
Spatial Displays			2	1	1
Hend Held Displays				1	1
Object-Projected Displays					1
Volumetric Displays	Incon: 0.01				

Table VIII. COMPARISON OF AUGMENTED REALITY TECHNOLOGIES BY SPECIFICATIONS

The overall results show that the Spatial Displays are the preferred choice for the current state of the art of the AR technologies. However, Head Mounted Displays also perform well and are growing fast in terms of TRL.

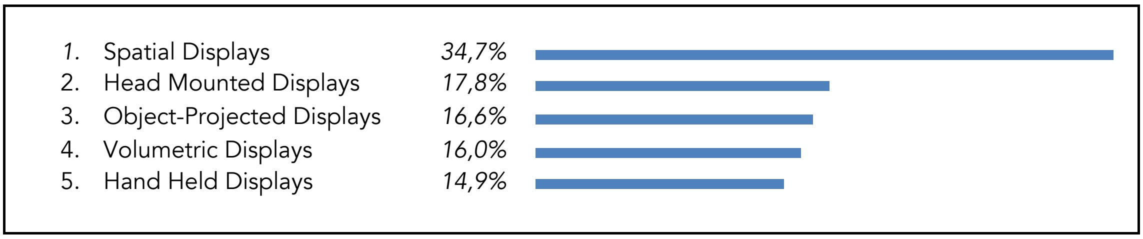


Figure 58. FINAL RANKING OF AUGMETED REALITY TECHNOLOGIES

6.3 Rendering for Virtual Environments and Augmented Reality Tools

6.3.1 Overview

The focus of this research, is the placement of information over the actual control tower outside view. Therefore, the relationship between synthetic information and the user's perspective is a major issue. This subject has been widely studied in other fields of application of virtual and augmented reality, such as cultural heritage, entertainment and medicine. In general, a V/AR system can be made of multiple displays, of any size, position and orientation. Below, is a modified rendering pipeline that was designed to obtain registration in any condition and later customized to our laboratory equipment.

6.3.2 Standard Projection (Uncoupled Perspective)

As already stated in chapter 4, most V/AR applications operate on some variant of the pinhole camera metaphor. According to this model, programmers may simply select a horizontal FOV, specify an aspect ratio, declare the distances from the near clipping plane and the far clipping plane, and build the projection matrix. For instance, the OpenGL[57] function *gluPerspective* [193] sets up a perspective projection matrix based on four user specified parameters (r, t, n and f). This entails the use a symmetrical frustum[58] such as the one represented in Figure 7 (left side).

---

[57] OpenGL (Open Graphics Library) is a cross-language, multi-platform Application Programming Interface (API) for rendering 2D and 3D vector graphics

[58] A *Fi* is a six-sided truncated pyramid that originates sectioning the shape the virtual camera FOV by means of two user-defined clipping planes. These are known as the 'far clipping plane' and the 'near clipping plane'. The latter is the one on which the scene must be projected as a necessary step of the rendering pipeline.

Extensive information about the OpenGL projection matrix and *gluPerspective* input parameters can be found either on the Internet [193], [194] or in the OpenGL Programming Guide, *alias* The Red Book [194], [195]. Also, alternate graphical Application Programming Interface (API) exist [196]. When *gluPerspective* is invoked, it builds a projection matrix that looks like this:

$$P = \begin{bmatrix} \frac{n}{r} & 0 & 0 & 0 \\ 0 & \frac{n}{t} & 0 & 0 \\ 0 & 0 & \frac{n+f}{n-f} & -\frac{2fn}{f-n} \\ 0 & 0 & -1 & 0 \end{bmatrix} \qquad (1)$$

Where r and t represent half of the horizontal and vertical near clip plane extents respectively, while n (near-Val) and f (far-Val) refer to the distances between the viewpoint (i.e. the eye-space origin) and the near and far clipping planes respectively.

When (1) is used, a few underlying assumptions have been made, and that is that (a) the viewer is positioned in front of the screen, (b) facing perpendicular to it and (c) looking at the centre of it. This is also known as the 'on-axis' projection model (Figure 7, left side). If the projection matrix does not change, relative movements between the eyes and the screen (e.g. back and forward movements) are theoretically forbidden, as they modify the physical FOV whereas the projection model (i.e. the projection matrix) does not change. To free the viewpoint position from the screen normal[59], OpenGL provides a second function (*glFrustum*) that sets up the projection matrix as follows:

$$P= \begin{bmatrix} \frac{2n}{r-l} & 0 & \frac{r+l}{r-l} & 0 \\ 0 & \frac{2n}{t-b} & \frac{t+b}{t-b} & 0 \\ 0 & 0 & \frac{n+f}{n-f} & -\frac{2fn}{f-n} \\ 0 & 0 & -1 & 0 \end{bmatrix} \qquad (2)$$

Where l, r, b and t denote the distances between the near clipping plane edges and the straight line that goes from the camera origin to the plane itself (in a perpendicular manner). Again, extensive information about *glFrustum* input parameters – namely l (left), r (right), b (bottom), t (top), n (nearVal) and f (farVal) – can be found either on the Internet [197] or in the in the OpenGL Programming Guide [195].

A projection matrix such as (2) allows for asymmetric *frusta* to be used. In other words, the viewpoint position is freed from the screen normal. This is known as "off-axis projection" (Figure 7, right side). As a matter of fact, the projection model delivered by (2) is much more flexible than the one provided by (1). For instance, the extents of each

---

[59] For the sake of readability, we refer to the straight line being orthogonal to the screen and passing by the centre of it simply as the screen normal.

frustum can be separately determined for each eye-screen pair, effectively implementing stereovision [70]. Still, the near clipping plane must be orthogonal to the virtual camera depth axis (i.e. the eye-space coordinate system e_n axis) and relative movements between the screen and the viewpoint are forbidden (unless accounted for by a tracking system).

Eventually, the field of AR introduces circumstances under which the assumptions of both *glFrustum*() and *gluPerspective*() fail and the resulting incorrectness is not tolerable [70], [101], [198]. Hence, a more general model is needed.

6.3.3 GENERALISED PROJECTION (EYE-COUPLED PERSPECTIVE)

To generate registered V/AR contents, the programmer needs formulas to compute the parameters of a perspective projection matrix (*l*, *r*, *b* and *t*) based on the relative position between the viewer's eyes and the screen.

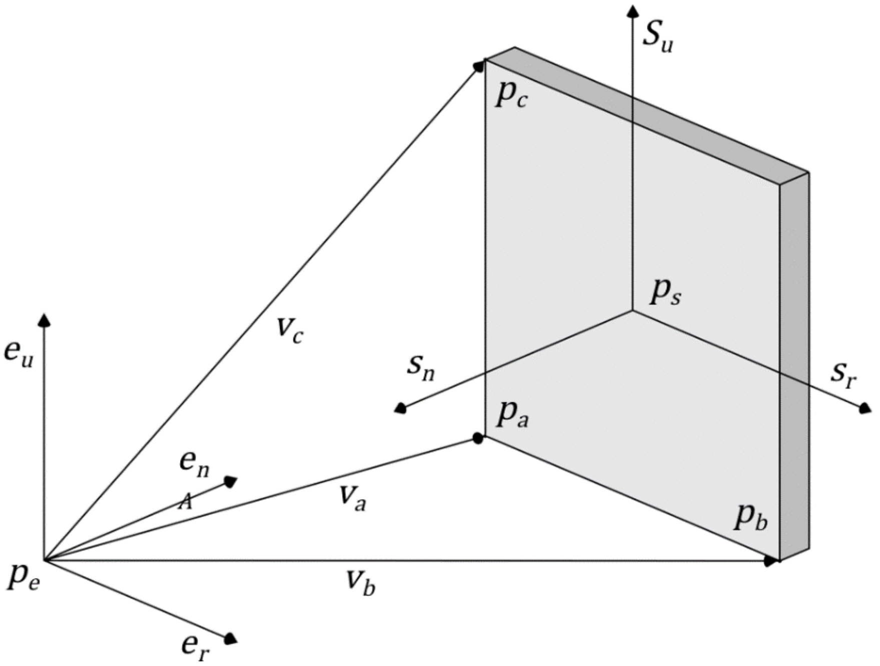


Figure 59. THE EYE-SPACE COORDINATE SYSTEM (ORIGIN P_E), THE SCREEN-SPACE COORDINATE SYSTEM (ORIGIN P_S) AND THE SCREEN CORNERS VECTORS V_A, V_B AND V_C.

The main characteristics of the V/AR system are depicted in Figure 59. These are the coordinates of the display corners, the origin of screen-space coordinate system, and the distance from the eye-space coordinate system origin to the screen.

The coordinates of the head-up display corners, namely p_a (lower left corner), p_b (lower right corner), and p_c (upper left corner) are expressed with respect to world-space coordinate system. If a flat screen is used, the position of the fourth point is implicit. Together these points encode the size of the screen, its aspect ratio, its position and orientation. Also, they can be used to compute an orthonormal basis for the screen-space coordinate system[60], which is the triad of vectors composed by s_r (the vector toward the right), s_u (the vector pointing up), and s_n (the vector normal to the screen, pointing in front of it).

$$s_r = \frac{p_b - p_a}{|p_b - p_a|} \tag{3}$$

$$s_u = \frac{p_c - p_a}{|p_c - p_a|} \tag{4}$$

$$s_n = \frac{s_r \times s_u}{|s_r \times s_u|} \tag{5}$$

The origin of the screen space coordinate system is the intersection between the perpendicular line drawn from p_e to the screen, and the plane of the screen itself. Since neither $\underline{p_e}$ nor p_s are fixed in space, when the viewer moves with respect to the screen, the screen-space origin changes accordingly. If s/he moves far to the side of the screen, then the screen space origin may not fall within the screen at all.

The coordinates of the eye-space origin $\underline{p_e}$ are typically available as a vector in the tracker-space frame of reference, and must be converted in the world-space coordinates. Thus, it is mandatory to know the exact transformation between the tacker-space frame of reference and the world space frame of reference. In most cases, this can be easily determined once the location and the orientation of the tracking device(s) is(are) fixed and known.

The distance from the eye-space origin p_e and the screen-space origin p_s can be computed by taking the dot product of the screen normal v_n with any of the screen vectors. However, because these vectors point in quite opposite directions, their product must be negated.

$$d = -(s_n \cdot v_a) \tag{6}$$

To compute the extents of the *frustum*, the vectors from the camera space origin (p_e) to the screen corners are needed. Once again, these can be easily calculated using the screen corners.

$$v_a = p_a - p_e \tag{7}$$

$$v_b = p_b - p_e \tag{8}$$

$$v_c = p_c - p_e \tag{9}$$

---

[60] In linear algebra an orthonormal basis for an inner product space is a basis whose vectors are all unit vectors orthogonal to each other.

Frustum extents may be interpreted (and computed) as distances from the screen-space origin to the edges of the screen (Figure 60). However, because these are not specified at the near clipping plane, they must be scaled from their value at the plane of the screen, d units away from the eye-space origin, to their value at the near clipping plane, n units away from the eye-space origin.

$$l = \frac{(s_r \cdot v_a)\, n}{d} \tag{10}$$

$$r = \frac{(s_r \cdot v_b)\, n}{d} \tag{11}$$

$$b = \frac{(s_u \cdot v_a)\, n}{d} \tag{12}$$

$$t = \frac{(s_u \cdot v_c)\, n}{d} \tag{13}$$

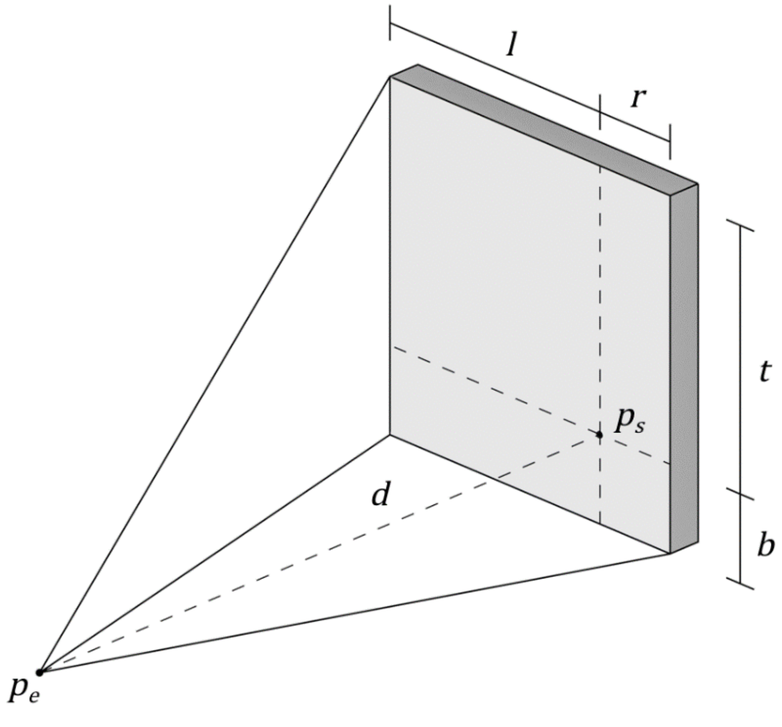


Figure 60. THE LENGTH OF THE FRUSTUM EXTENTS (*L, R, B* AND *T*) AT THE PLANE OF THE SCREEN.

Inserting these values in (2) allows the programmer to build a head-tracked (or eye-tracked), off–axis projection (Figure 7, right side). In other words, a skewed frustum for an arbitrary screen viewed by an arbitrary 'eye' can be used. However, here is one final limitation that must be overcome, and that is that the near clipping plane is orthogonal to the eye-space depth axis (e_n), whereas the plane of the screen may not be. In practice, the projection plane must be freed from the orientation of the e_u-e_r plane. Unfortunately, the way the perspective projection matrix is built simply disallows this. Instead, it is possible to rotate the virtual world to line up the desired projection plane with the e_u-e_r orientation. As far as the projection outcome is concerned this is equivalent to rotating the viewing frustum aligning the near clipping plane to the plane of the screen. Note that this operation does not affect the frustum extents calculation.

The transformation matrix that rotates the eye-space coordinate system so that its standard axis e_r, e_u and e_n match the orientation of the screen-space coordinate system is this:

$$R = \begin{bmatrix} s_{rx} & s_{ux} & s_{nx} & 0 \\ s_{ry} & s_{uy} & s_{ny} & 0 \\ s_{rz} & s_{uz} & s_{nz} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{14}$$

As can be easily deduced from:

$$R \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = s_r \tag{15}$$

$$R \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = s_u \tag{16}$$

$$R \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = s_n \tag{17}$$

If something lies on the e_u-e_r plane, this transformation will align it to the plane of the screen.

The inverse mapping is produced by:

$$R^{-1} s_r = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{18}$$

$$R^{-1} s_u = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \tag{19}$$

$$R^{-1} s_n = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \tag{20}$$

Resulting in:

$$R^{-1} = \begin{bmatrix} s_{rx} & s_{ry} & s_{rz} & 0 \\ s_{ux} & s_{uy} & s_{uz} & 0 \\ s_{nx} & s_{ny} & s_{nz} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{21}$$

Also, because R is orthogonal:

$$R^{-1} = R^T \tag{22}$$

Applying (21) to all objects in the scene will rotate the virtual world until the plane of the screen lines up with the e_u-e_r plane.

Multiplying the projection matrix P by the rotation matrix R^T will solve the issue of the arbitrary oriented projection screen.

$$M = R^T P \tag{23}$$

With this matrix, the programmer is finally able to render the scene and generate stereoscopic contents for a multi-display virtual/augmented reality system of any size, configuration, position and orientation.

The C#/python code to implement this algorithm in the blender game engine can be found in appendix A. Within the software, the complexity of defining the user's viewport is handled by giving a physical representation to the VR system within the model itself. In this way, the screens shape, position, orientation and dimension can be adjusted through the user interface in order to resemble the virtual environment setup. This is shown in Figure 61, Figure 62 and Figure 63.

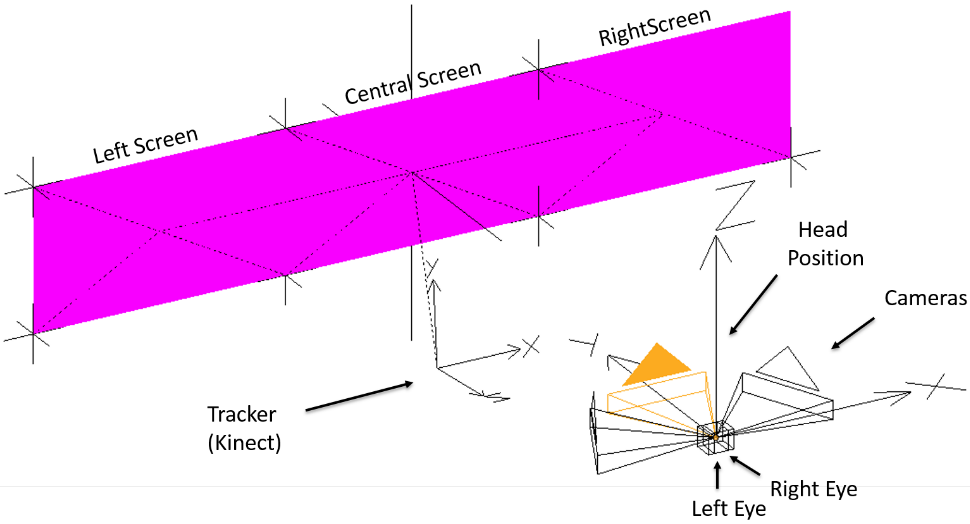
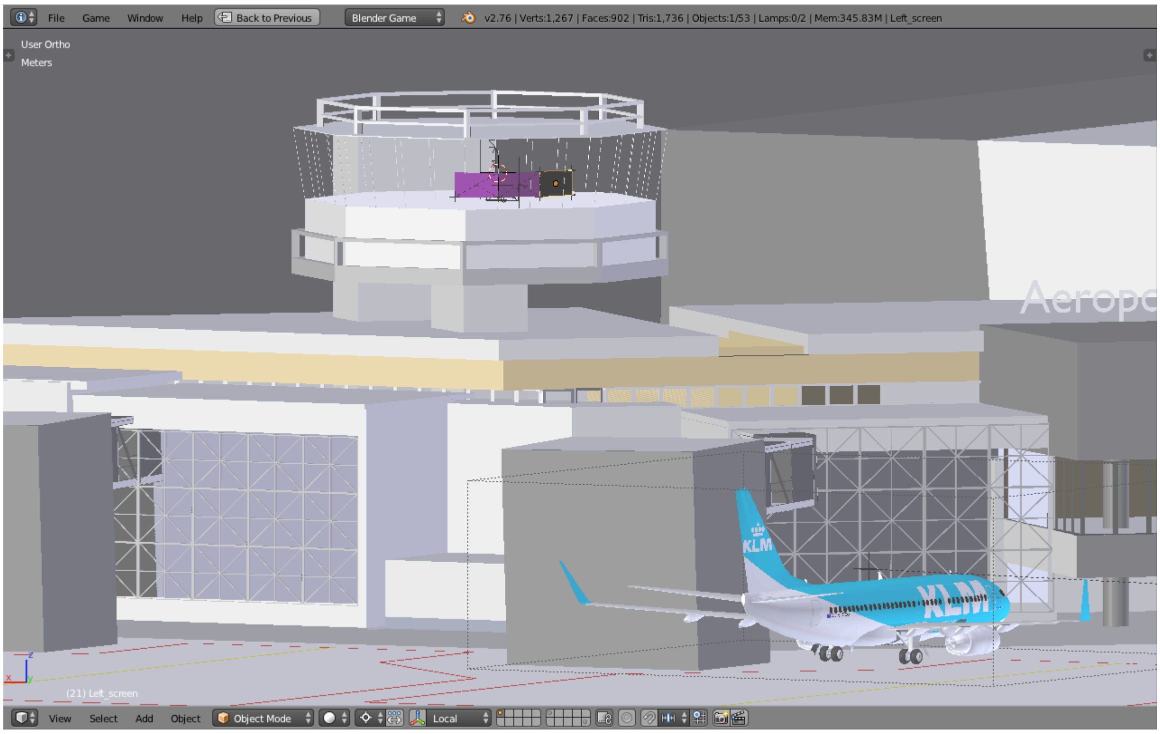


Figure 61. VIEWPORT SETUP: DIGITAL REPRESENTATION OF THE VIRTUAL ENVIRONMENT COMPONENTS



v2.76 | Verts:1,267 | Faces:902 | Tris:1,736 | Objects:1/53 | Lamps:0/2 | Mem:345.83M | Left_screen

Figure 62. VIEWPORT SETUP: REPRESENTATION OF THE VIRTUAL ENVIRONMENT COMPONENTS WITHIN THE CONTROL TOWER (FRONT VIEW)

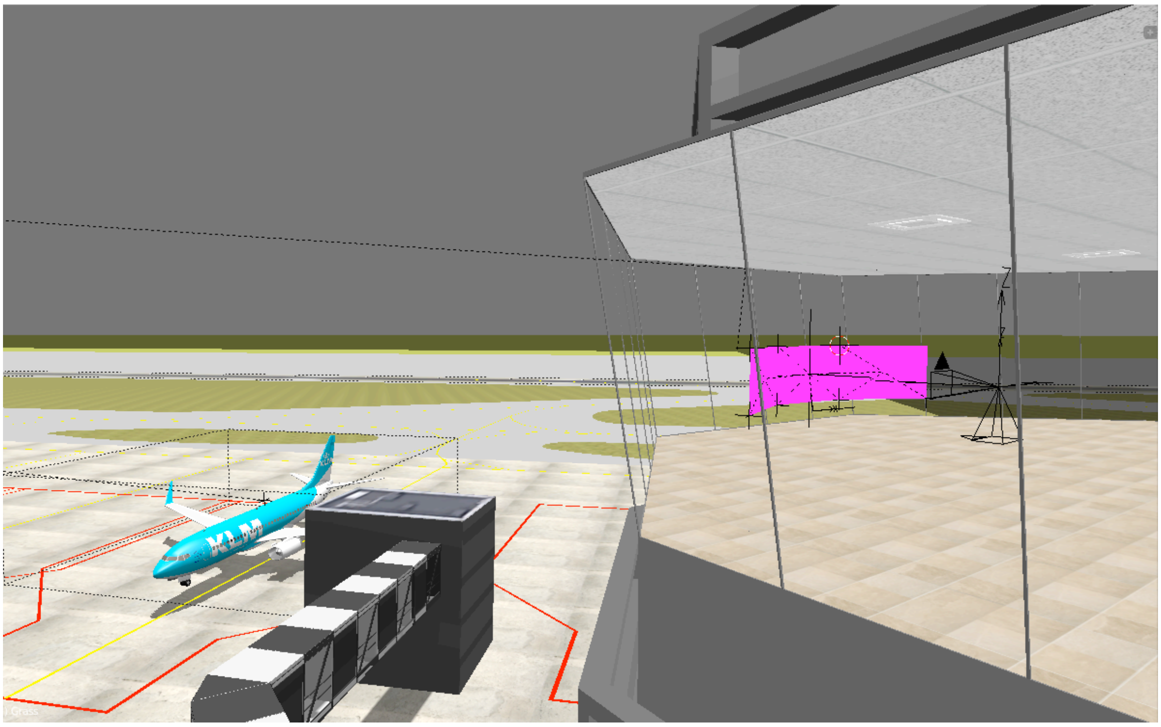


Figure 63. VIEWPORT SETUP: REPRESENTATION OF THE VIRTUAL ENVIRONMENT COMPONENTS WITHIN THE CONTROL TOWER (SIDE VIEW)

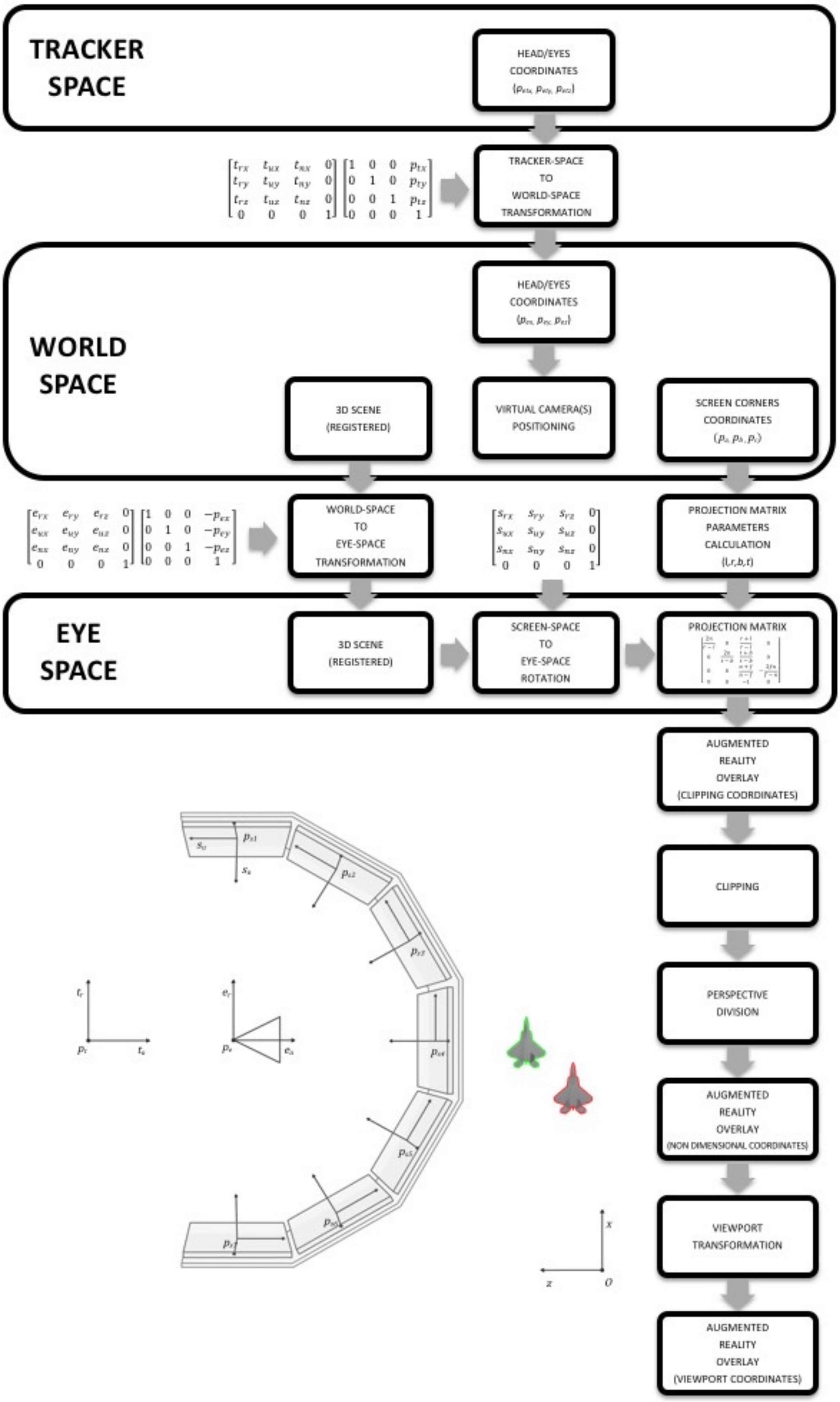


Figure 64. OVERALL SCHEMATICS OF THE RENDERING PIPELINE

6.4 DESIGNING THE OVERLAYS

AR overlays to be used by air traffic controllers have been conceptually developed using an ecological interface design approach and taking into consideration the reference scenario described in paragraph 5. The output of this process is detailed in the overlay chart. Some of the overlays contained in the chart have been implemented as proof of concept in simulated environments.

6.4.1 THE ECOLOGICAL INTERFACE DESIGN APPROACH

In ATC, operators must deal with easy tasks and familiar events, as well as with unfamiliar, time consuming and unexpected events. Besides talking to pilots, controllers need to extract information from the head down displays, check weather, consult Flight Strips (FS), elaborate long term strategies, detect potential conflicts, make tactical decisions, coordinate with each-other and look out of the tower window (if any) [25], [187]. In addition, controllers need to balance cognitive resources and carefully timetable actions [25], [187]. Under these circumstances, the design of human-computer interfaces cannot only focus on the user wishes but must consider the complexity of the work domain.

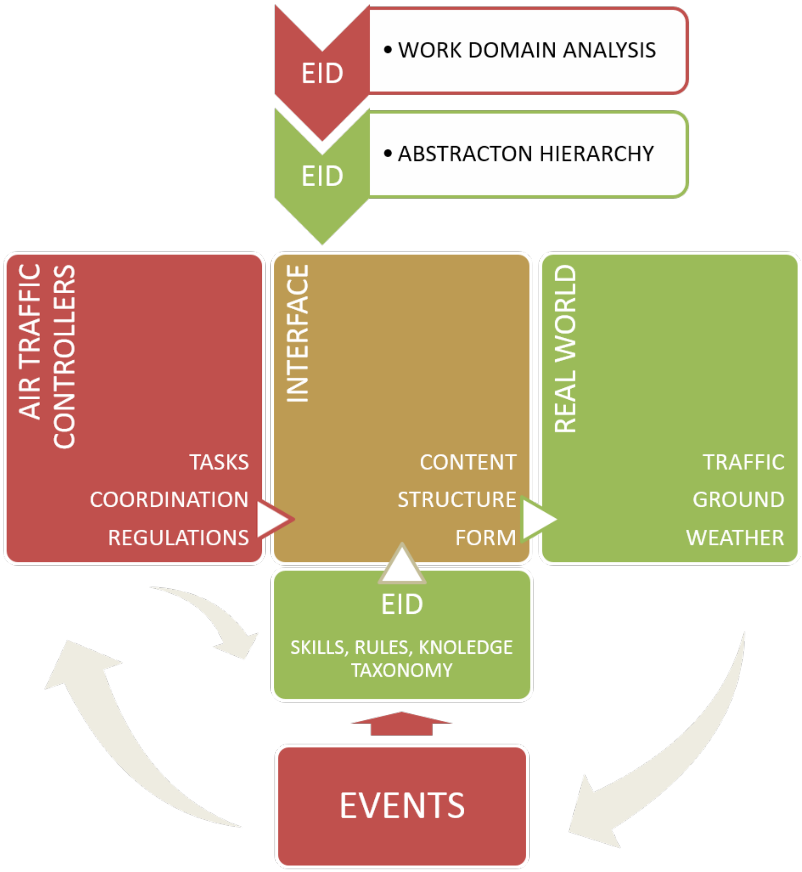


Figure 65. THE ECOLOGICAL INTERFACE DESIGN APPROACH APPLIED TO AIR TRAFFIC CONTROL

Ecological Interface Design (EID) is a theoretical framework for designing human-machine interfaces in complex, real-time and dynamic systems. This methodology differs from User-Centred Design (UCD) insofar it focuses on the analysis of the work domain (a.k.a. Work Domain Analysis - WDA) rather than on the end-user wishes. EID attempts to provide the operators with the necessary tools and information to become active problem solvers as opposed to passive monitors, particularly during the development of unforeseen events [188]. Interfaces designed following the EID approach aim to decrease the mental workload when dealing with unfamiliar and unexpected events, which are attributed to increased psychological pressure [188]. The EID approach makes constraints and complex relationships in the work environment perceptually evident (e.g. visible, audible) to the user. By doing this, the interface allows more of users' cognitive resources to be devoted to higher cognitive processes such as problem solving and decision making. By reducing mental workload and supporting knowledge-based reasoning, EID aims to improve user performance and overall system reliability for both anticipated and unanticipated events in a complex system.

EID makes use of three pillars of cognitive engineering research

- the Work Domain Analysis (WDA)
- the Abstraction Hierarchy (AH)
- the Skills, Rules, Knowledge (SRK) taxonomy.

6.4.1.1 THE WORK DOMAIN ANALYSIS

The EID process commences with a work domain analysis that comprises one or more of the following sources: (a) document analysis, (b) real world observations and (c) interviews with domain experts. The procedure depends on the domain in question and the availability of data. Within this research mostly (a) and (c) were used.

6.4.1.2 THE ABSTRACTION HIERARCHY

The information gathered with the WDA can be organized in a 5-level decomposition chart called the Abstraction Hierarchy. The AH describes a system at different levels of abstraction using how and why relationships. Moving down the levels shows how certain elements in the system are achieved, whereas moving up reveals why certain elements exist. Elements at highest level of the model define the purposes and goals of the system. Elements at the lowest levels of the model indicate and describe the physical components (i.e. equipment) of the system. The how and why relationships are shown on the AH as means-ends links. It is not uncommon for a Work Domain Analysis to yield multiple AH models; each examining the system at a different level of detail. The AH was not used within this research.

6.4.1.3 THE SKILLS, RULES, KNOWLEDGE TAXONOMY

According to the framework proposed by Rasmussen, the terms Skill, Rule and Knowledge (S-R-K) refer to "the degree of conscious control exercised by the individual over his or her activities, depending on the degree of familiarity with the task and the environment" [189]. In other words these are types of behaviour or psychological processes present in operators' information processing [190]. The three categories describe three possible ways in which information is extracted, understood and processed by human beings. The behaviour can shift from a level to another or different cognitive behaviours can be active in parallel. The SRK framework was developed by Rasmussen to help designers combine information requirements for a system and aspects of human cognition [188].

6.4.1.3.1 SKILL-BASED BEHAVIOUR

A skill-based behaviour is a type of behaviour that requires very little or no conscious control to perform or execute an action once an intention is formed. Rasmussen defines it as "the smooth execution of highly practiced, largely physical actions in which there is virtually no conscious monitoring" [189]. At the skill-based level, the behaviour is regulated by the lowest level of conscious involvement and is characterized by highly routinized and automated activities. This automaticity allows operators to free up cognitive resources, which can then be used for higher cognitive functions like problem solving. For example, bicycle riding is considered a skill-based behaviour in which very little attention is required for control once the skill is acquired. In brief, a skill based behaviour involves:

- High Automated processes involving long term memory (procedural)
- Low Executive control (i.e. low attention and working memory)
- No Problem solving
- No Decision-making

6.4.1.3.2 RULE-BASED BEHAVIOUR

A rule-based behaviour is characterised by the use of rules and procedures to select a course of action in a familiar work situation [191]. The rules can be a set of instructions acquired by the operator through experience or given by supervisors and former operators. Operators are not always required to know the underlying principles of a system, to perform a rule-based control. For example, hospitals have highly-procedurals instructions for fire emergencies. Therefore, when one sees a fire, one can follow the necessary steps to ensure the safety of the patients without any knowledge of fire behaviour. Rule-based behaviour is also activated in familiar work situations, but it is distinguished from skill-based behaviour, as "it requires some degrees of conscious involvement and attention. Situation assessment leads to recognition of which procedures apply to particular familiar situations". In brief, a rule based behaviour involves:

- Less automated processes and long term memory (procedural) than Skill level
- More executive control (i.e. more attention and working memory) than Skill level
- No Problem solving
- No Decision-making

6.4.1.3.3 KNOWLEDGE-BASED BEHAVIOUR

A knowledge-based behaviour represents a more advanced level of reasoning. This type of control must be employed when the situation is novel and/or unexpected and no rule-based solutions are available by the book. At this level, the user "carries out a task in an almost completely conscious manner. This would occur in a situation where a beginner is performing the task (e.g. a trainee at the beginning of its training) or where an expert is facing with a completely novel situation. In either such cases, the user would have to exert considerable mental effort to assess the situation and his or her responses are likely to be slower. Also, after each control action, the user might need to review its effect before taking further action, which would probably further slow-down the responses to the situation" [192]. Since operators need to form explicit responses based on their current analysis of the system, cognitive workload is typically greater than when using skill or rule-based behaviours. In brief, a rule based behaviour involves:

- No automated processes and long term (procedural) memory
- Executive control (high attention and working memory)
- Problem solving
- Decision-making

6.4.2 THE CONTROL TOWER WORK DOMAIN ANALYSIS

A work domain analysis of the control tower environment has been performed by means of exiting documentation review, visual observations and domain experts consultation. A previous study performed within the context of Integrated Tower Working Position development was used, which included on site visits to Stockholm Arlanda, London Gatwick, Rome Fiumicino and Naples Capodichino. In this sense, this analysis is representative of many control towers typical operations and does not imply any SESAR forthcoming solution or Collaborative Decision Making (CDM) messaging and communication systems. In the case of CDM airports local procedures may be somewhat adapted in accordance with the CDM implementation manual. Also, the technological environment may or may not include electronic strips or various controller support tools.

The output of the WDA is presented hereafter (in both written and graphical forms) divided by four typical control tower roles and positions:

1. Clearance Delivery Controller

2. Ground Delivery Controller

3. Tower Controller

4. Tower Supervisor

The WDA covers both standard and low visibility conditions.

6.4.2.1 CLEARANCE DELIVERY CONTROLLER

Receives all data via the Flight Plan System (FPS) and ensures local strips (paper or electronic) are generated in correct and complete format. The following are crosschecked and verified; NMOC (Network Manager Operations Centre) restrictions, wake turbulence category and aircraft type, and stand number. Once the aircraft calls for a clearance with ATIS and start-up request (prior to the EOBT - Estimated off Blocks Time), the Clearance Delivery controller activates the flight plan in the Flight Data Processing System. The ATC clearance can then be issued, this includes; SID (Standard Instrument Departure), climb level, local transponder squawk, departure runway, and ground control frequency. Finally, the Clearance Delivery controller advises or coordinates with the Ground controller and Tower Controller

6.4.2.2 GROUND CONTROLLER

The Ground controller has active control and responsibility for surveillance on the entire airport platform except for the active runway(s). The controller issues push back clearance and taxi clearances to both aircrafts and vehicles. He checks taxiway usage against aircraft type to ensure that wing span and PLR (Pavement Load Ratings) are respected. In the event of an emergency, The Ground controller stops all moving traffic and coordinates with CFR (Crash Fire Response) as well as with the Tower Controller. In case of runway crossing, the Ground controller coordinates directly with the Tower controller for all active runway crossing clearances (aircraft and vehicles). In case of LVP, the Ground controller applies the related restrictions in accordance with the weather conditions. For example, in case of no visibility on the Apron, the Ground Controller approves the push back only for one aircraft if multiple push back requests come from the same area. In case of severe LVP (RVR<400m), the Ground Controller provides taxi clearance using block spacing to ensure that an adequate distance is maintained between the flights that are moving on the manoeuvring area.

6.4.2.3 TOWER CONTROLLER

The Tower controller (a.k.a. runway controller) is responsible for the active runway(s) and all airborne traffic arriving, departing and overflying the airport. He or she is also responsible for all the air traffic in the ATZ. The Tower controller ensures that landing aircraft use the correct exit runways and no unknown traffic enters the ATZ. If present, the Tower controller monitors safety nets and warning systems (e.g. for runway incursions). For IFR aircraft, transfer of control (and frequency) from APP to the TWR is typically made once ILS intercept is obtained and for departures immediately after take-off (or as soon as possible after take-off). For (Visual Flight Rules) VFR aircraft, the Tower controller ensures that VFR or SVFR (Special Visual Flight Rules) conditions allow appropriate operations. A clearance can be then given to a point within the ATZ or to a pre-defined point on the traffic pattern (e.g. downwind). Regarding helicopter traffic, the Tower controller is responsible for landing pads and entry/departure routes. In case of Runway Crossing, the Tower controller coordinates with the Ground controller to issue runway crossing clearances (aircraft or vehicles). The Tower controller is responsible for runway separation notably between arrivals and departures or subsequent arrivals or departures. Separation is adapted as to whether the aircraft is flying IFR (SID dependant) or VFR as well as according to the Wake Turbulence Category. Once runway separation is obtained, the Tower controller can issue take-off or landing clearances. The Tower controller coordinates with the Ground controller, clearance delivery controller and, of course, with the Tower supervisor if a runway change is required. The Tower controller coordinates with the APP controller the gaps in the arrival flow to sequence VFR traffic. The Tower controller coordinates with Approach control if gaps in the arrival flow are required to release pending IFR departures. In case of LVP, the TWR controller is responsible to apply the related restriction. In particular, he or she is responsible to ensure that the critical and sensitive areas of the ILS are always free during the approach. If present, the A-SMGCS systems can be used to ascertain aircraft positions particularly during times of darkness or LVC.

6.4.2.4 TOWER SUPERVISOR

The Tower supervisor monitors the weather through the hourly weather reports. By doing so, he or she ensures that aircraft operations can be safely performed and chooses what runway configuration shall be used. The Tower supervisor coordinates with the Ground controller and with airfield maintenance to ensure that runway inspections are performed and in the case of contaminated runways he or she takes appropriate actions. The Tower supervisor coordinates with the Tower controller, Clearance Delivery controller and the Approach control if a runway change is required. The Tower supervisor coordinates with approach control to identify airport hourly arrival and departure rates and ensures compliance with regulations issued by the Network Manager

Operation Centre. In case of LVP, the Tower Supervisor is responsible to coordinate flow restrictions in accordance with the weather and visibility conditions. In case of LVP, the Tower Supervisor is responsible to inform the flight of the current weather and visibility condition by continuously updating ATIS messages.

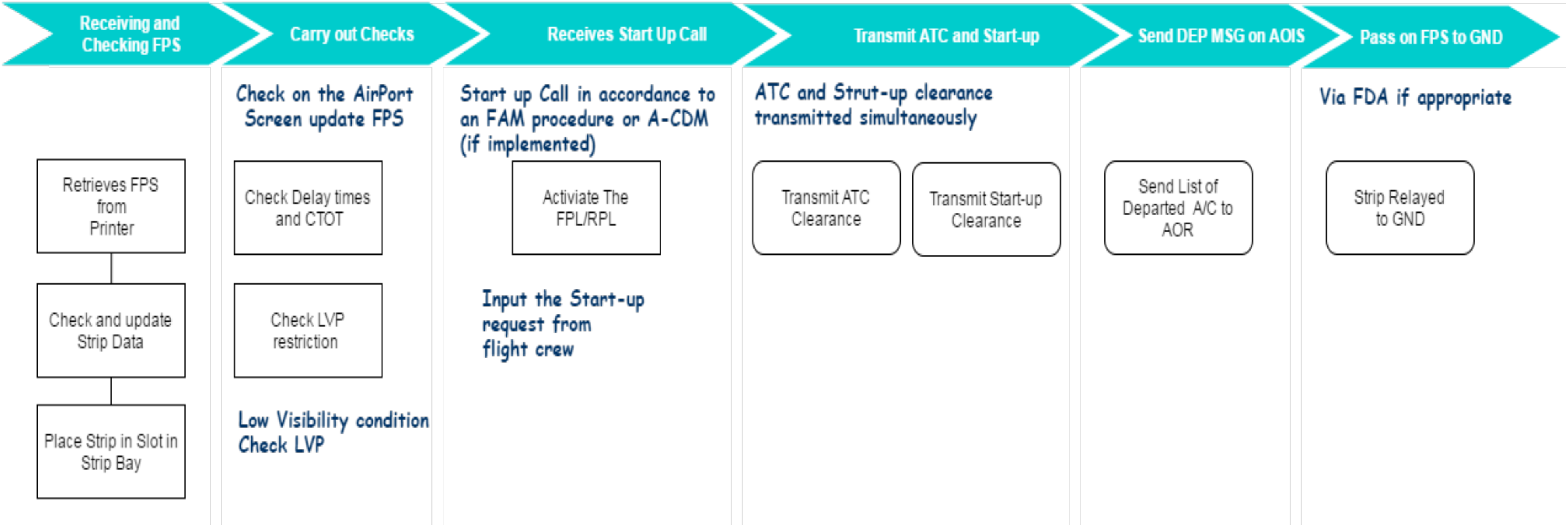


Figure 66. Control tower work domain analysis - delivery controller

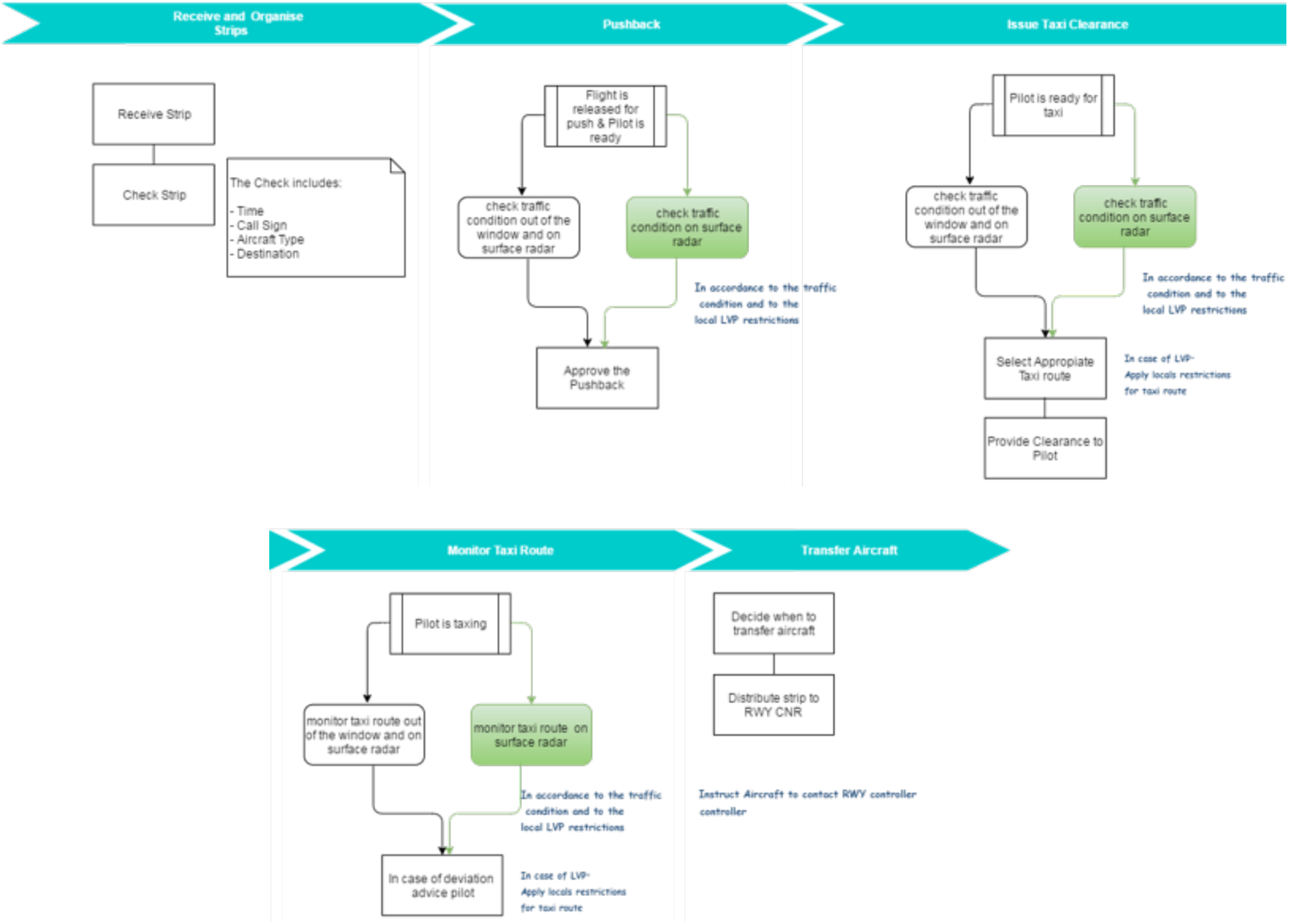


Figure 67. Control tower work domain analysis - ground controller: departures (above), arrivals (below)

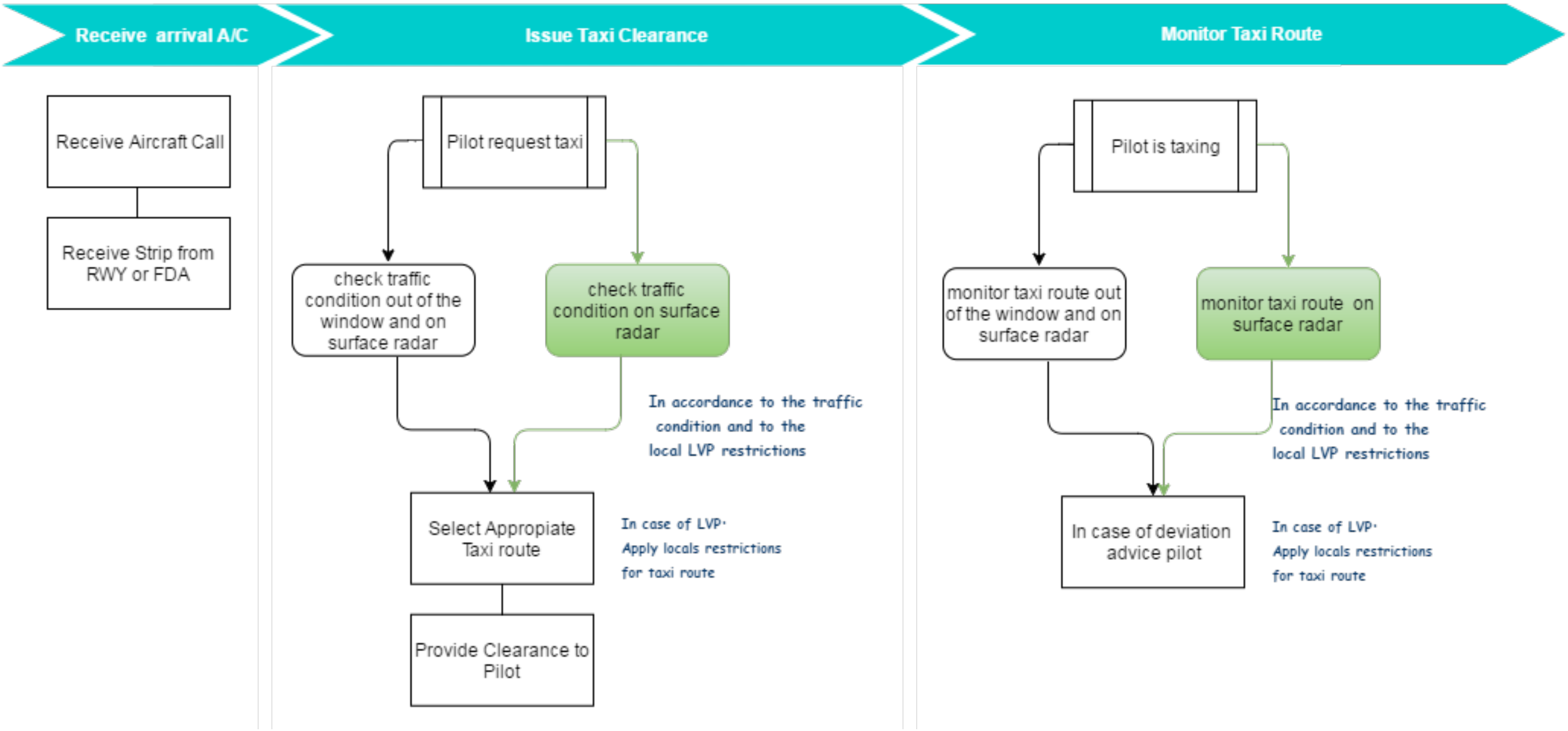


Figure 68. CONTROL TOWER WORK DOMAIN ANALYSIS - GROUND CONTROLLER: ARRIVALS

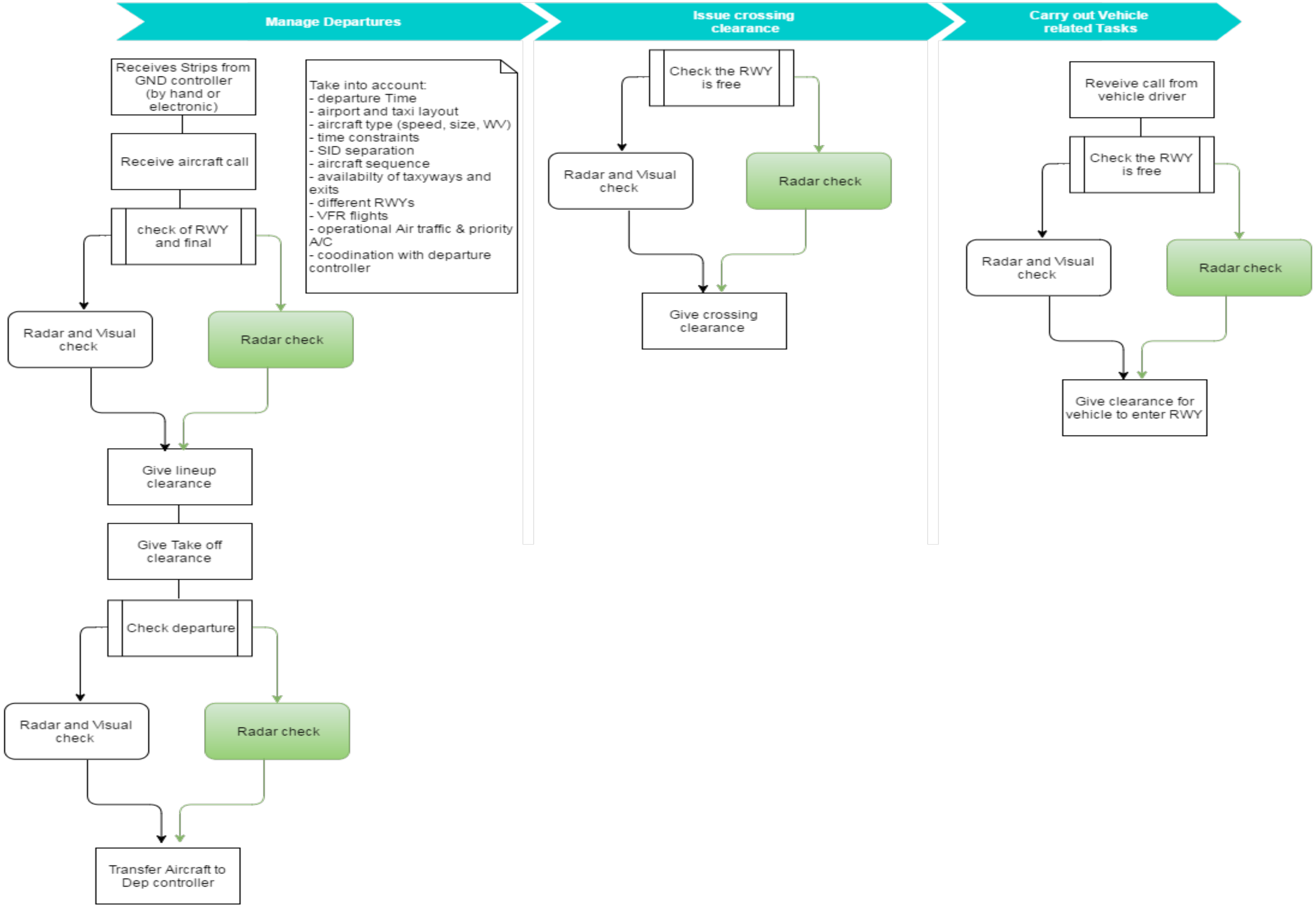


Figure 69. CONTROL TOWER WORK DOMAIN ANALYSIS - TOWER CONTROLLER: ARRIVALS

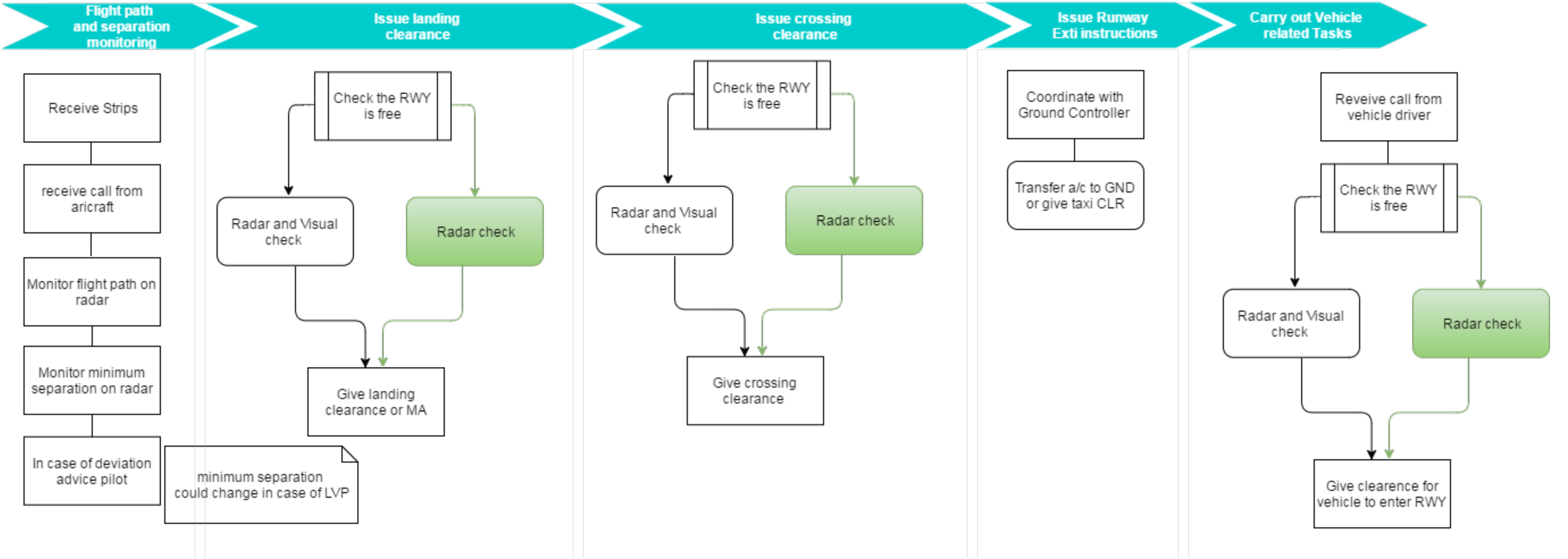


Figure 70. Control tower WDA - tower controller: departures

6.4.2.5 THE CASE OF BOLOGNA AIRPORT

In Bologna airport control tower a couple of controllers, namely the Tower (TWR) and Ground/Delivery (GND/DEL) controller, work together to ensure the safe and expedite flow of traffic.

The GND/DEL controller is responsible to provide the Aerodrome Control Service and the Flight Information Service on the manoeuvring area except the runway. The image below shows the GND/DEL controller CWP. Information from supporting systems are displayed in the screens on the left and on the right. The screen on the right displays AOIS information and the screen on the left displays ADM and FDP. The central screen is the SMR and provides the ATCO with the position information of all the traffic in the manoeuvring area.

Figure 71. BOLOGNA AIRPORT CONTROL TOWER: GROUND/DELIVERY CONTROLLER WORKING POSITION

The TWR controller is responsible to provide the Aerodrome Control Service, the Flight Information Service and the Alert Service to all the traffic in the Aerodrome Traffic Zone and on the Runway. The image below shows the TWR controller CWP. The most important systems used by the ATCO in his tasks are the Radar (air and ground), the Communication system, the Light Control and the Strips. The radar screen is in front of the ATCO and provides aircrafts position and identification. The SMR (Ground radar) screen is positioned in higher position (not visible in the picture) linked to the room ceiling in front of the ATCO. The screen on the right is used by the ATCO to control the aerodrome lights, stopbars included. Specific buttons are available to set the light in accordance to the visibility conditions and the approach category (ILS CAT II and III).

On this screen is also displayed to the ATCO a warning system that informs him of her of aerodrome de-categorization in case of system failure. Between the radar and the light screen the communication control panel is available. Via this panel, the ATCO manages frequencies and phone calls. On the right, a strip printer prints the arrival strips 20 minutes before estimated landing time (departure strips are provided to the TWR ATCO by the GND ATCO).

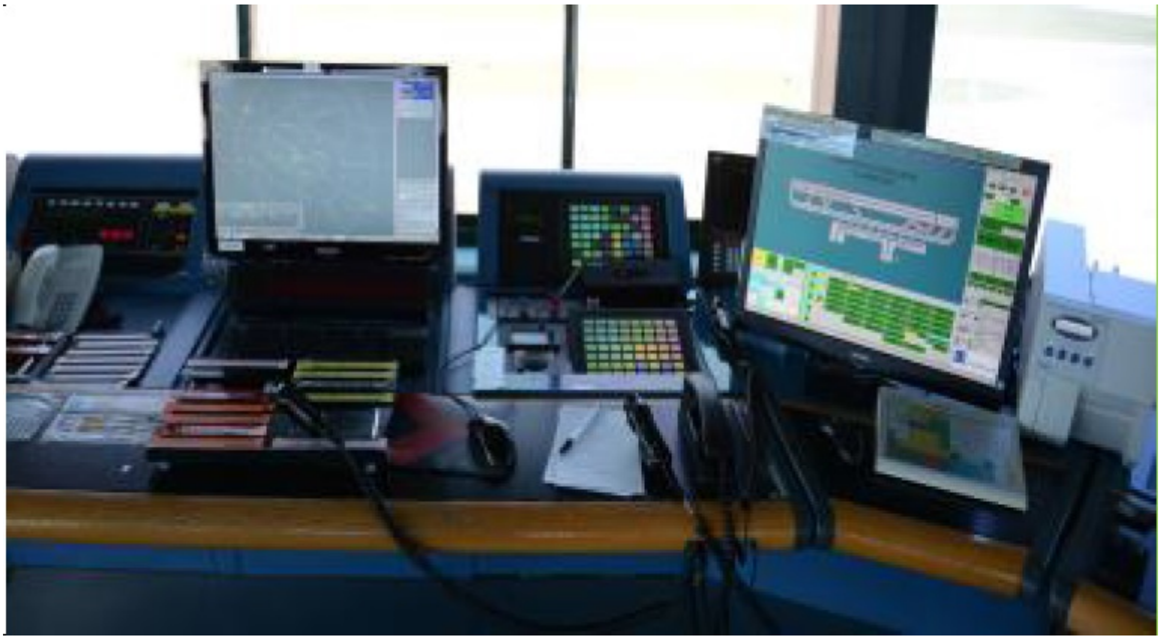


Figure 72. BOLOGNA AIRPORT CONTROL TOWER: TOWER CONTROLLER WORKING POSITION

A side position is also available next to the TWR CWP for a coordinator controller. This position is equipped with the approach radar providing information for all inbound and outbound flights within an area of about 100 nm. The screen on the right provides the ATCO with access to all the supporting systems (FDP, AOIS, ADM). A Communication panel is also available to manage frequencies and telephone.

Figure 73. BOLOGNA AIRPORT CONTROL TOWER: COORDINATOR WORKING POSITION

6.4.3 THE SKILL-RULE-KNOWLEDGE ANALYSIS

Amongst bologna tower controllers' tasks a subset of nine was selected for further analysis based on their relevance to this research. For these tasks a SRK analysis was performed to examine controller's cognitive behaviour. The selected tasks are summarized in Table IX.

TASK CODE	TASK DESCRIPTION
GND/DEL 1	ISSUE ATC CLEARANCE
GND/DEL 2	ISSUE START UP CLEARANCE
GND/DEL 3	APPROVE PUSH BACK
GND/DEL 4	ISSUE TAXI CLEARANCE
GND/DEL 5	MONITOR TAXI ROUTE
TWR 1	ISSUE LANDING CLEARANCE
TWR 2	ISSUE TAKE OFF CLEARANCE
TWR 3	MONITOR TAKE OFF AND LANDING OPERATIONS
TWR 4	ISSUE CLEARANCE TO VEHICLE FOR RUNWAY INSPECTIONS/OPERATIONS

Table IX. SELECTED CONTROL TASKS

The S, R, K analysis has been performed in four visibility conditions:

1. **NVC** (Normal Visibility Condition): visibility equal or greater than 5km and ceiling equal or greater than 1500ft

2. **LVC 1**: the control tower can visually monitor all the traffic on the manoeuvring area and pilots are able to taxi maintaining visual reference with other traffic.

3. **LVC 2**: all or a part of the manoeuvring area cannot be visually monitored from the TWR but aircraft are able to taxi maintaining visual reference with other traffic.

4. **LVC 3**: the RVR at TDZ, MID or STOP/END is less than 400 m.

For each type of behaviour (skill, rule and knowledge) three intervals were considered, for a total of nine. Given a certain task, the more the assigned interval is shift to the right, the more the task is deemed complex. Also, most of the tasks could not be strictly classified as skill, rule or knowledge. In fact, their complexity depends on several other factors, including the traffic condition (low or high).

A more detailed analysis of the selected control tasks is reported in Appendix B.

NVC

TASK CODE	TASK DESCRIPTION	S			R			K		
GND 1	ISSUE ATC CLEARANCE			▨	▨	▨				
GND 2	ISSUE START UP CLEARANCE				▨	▨				
GND 3	APPROVE PUSH BACK					▨	▨	▨		
GND 4	ISSUE TAXI CLEARANCE						▨	▨		
GND 5	MONITOR TAXI ROUTE					▨	▨			
TWR 1	ISSUE LANDING CLEARANCE						▨	▨		
TWR 2	ISSUE TAKE OFF CLEARANCE						▨	▨		
TWR 3	MONITOR TAKE OFF AND LANDING OPERATIONS					▨	▨			
TWR 4	ISSUE CLEARANCE TO VEHICLE FOR RUNWAY INSPECTIONS / OPERATIONS						▨	▨		

Table X. S-R-K ANALYSYS - NORMAL VISIBILITY CONDITIONS

<table>
<tr><th colspan="11">LVC 1</th></tr>
<tr><th>TASK CODE</th><th>TASK DESCRIPTION</th><th colspan="3">S</th><th colspan="3">R</th><th colspan="3">K</th></tr>
<tr><td>GND 1</td><td>ISSUE ATC CLEARANCE</td><td></td><td></td><td>▨</td><td>▨</td><td>▨</td><td></td><td></td><td></td><td></td></tr>
<tr><td>GND 2</td><td>ISSUE START UP CLEARANCE</td><td></td><td></td><td></td><td>▨</td><td>▨</td><td></td><td></td><td></td><td></td></tr>
<tr><td>GND 3</td><td>APPROVE PUSH BACK</td><td></td><td></td><td></td><td></td><td>▨</td><td>▨</td><td>▨</td><td></td><td></td></tr>
<tr><td>GND 4</td><td>ISSUE TAXI CLEARANCE</td><td></td><td></td><td></td><td></td><td></td><td>▨</td><td>▨</td><td></td><td></td></tr>
<tr><td>GND 5</td><td>MONITOR TAXI ROUTE</td><td></td><td></td><td></td><td></td><td>▨</td><td>▨</td><td></td><td></td><td></td></tr>
<tr><td>TWR 1</td><td>ISSUE LANDING CLEARANCE</td><td></td><td></td><td></td><td></td><td></td><td>▨</td><td>▨</td><td></td><td></td></tr>
<tr><td>TWR 2</td><td>ISSUE TAKE OFF CLEARANCE</td><td></td><td></td><td></td><td></td><td></td><td>▨</td><td>▨</td><td></td><td></td></tr>
<tr><td>TWR 3</td><td>MONITOR TAKE OFF AND LANDING OPERATIONS</td><td></td><td></td><td></td><td></td><td>▨</td><td>▨</td><td></td><td></td><td></td></tr>
<tr><td>TWR 4</td><td>ISSUE CLEARANCE TO VEHICLE FOR RUNWAY INSPECTIONS / OPERATIONS</td><td></td><td></td><td></td><td></td><td></td><td>▨</td><td>▨</td><td></td><td></td></tr>
</table>

Table XI. S-R-K ANALYSYS - LOW VISIBILITY CONDITIONS 1

LVC 2

TASK CODE	TASK DESCRIPTION	S			R			K		
GND 1	ISSUE ATC CLEARANCE			▨	▨	▨				
GND 2	ISSUE START UP CLEARANCE				▨	▨	▨			
GND 3	APPROVE PUSH BACK					▨				
GND 4	ISSUE TAXI CLEARANCE						▨	▨		
GND 5	MONITOR TAXI ROUTE						▨	▨		
TWR 1	ISSUE LANDING CLEARANCE						▨	▨		
TWR 2	ISSUE TAKE OFF CLEARANCE						▨	▨		
TWR 3	MONITOR TAKE OFF AND LANDING OPERATIONS						▨	▨		
TWR 4	ISSUE CLEARANCE TO VEHICLE FOR RUNWAY INSPECTIONS / OPERATIONS							▨	▨	

Table XII. S-R-K analysys - low visibility conditions 2

LVC 3

TASK CODE	TASK DESCRIPTION	S			R			K		
GND 1	ISSUE ATC CLEARANCE				▨	▨				
GND 2	ISSUE START UP CLEARANCE					▨	▨			
GND 3	APPROVE PUSH BACK						▨			
GND 4	ISSUE TAXI CLEARANCE						▨	▨		
GND 5	MONITOR TAXI ROUTE						▨	▨		
TWR 1	ISSUE LANDING CLEARANCE						▨	▨		
TWR 2	ISSUE TAKE OFF CLEARANCE						▨	▨		
TWR 3	MONITOR TAKE OFF AND LANDING OPERATIONS						▨	▨		
TWR 4	ISSUE CLEARANCE TO VEHICLE FOR RUNWAY INSPECTIONS / OPERATIONS							▨	▨	

Table XIII. S-R-K ANALYSYS - LOW VISIBILITY CONDITIONS 3

6.4.4 OVERLAY CHART

Based on the control tower WDA and S-R-K task analysis, several AR overlays were conceived having in mind the following goals:

- Increase controllers' situation awareness in NVC
- Restore controllers' situation awareness in LVC
- Reduce head down time
- Limit physical and mental refocusing between the outside view and head down equipment
- Simplify controllers' cognitive behaviour (favour rule-based behaviour against knowledge based behaviour)

A first classification is given in the overlay chart below:

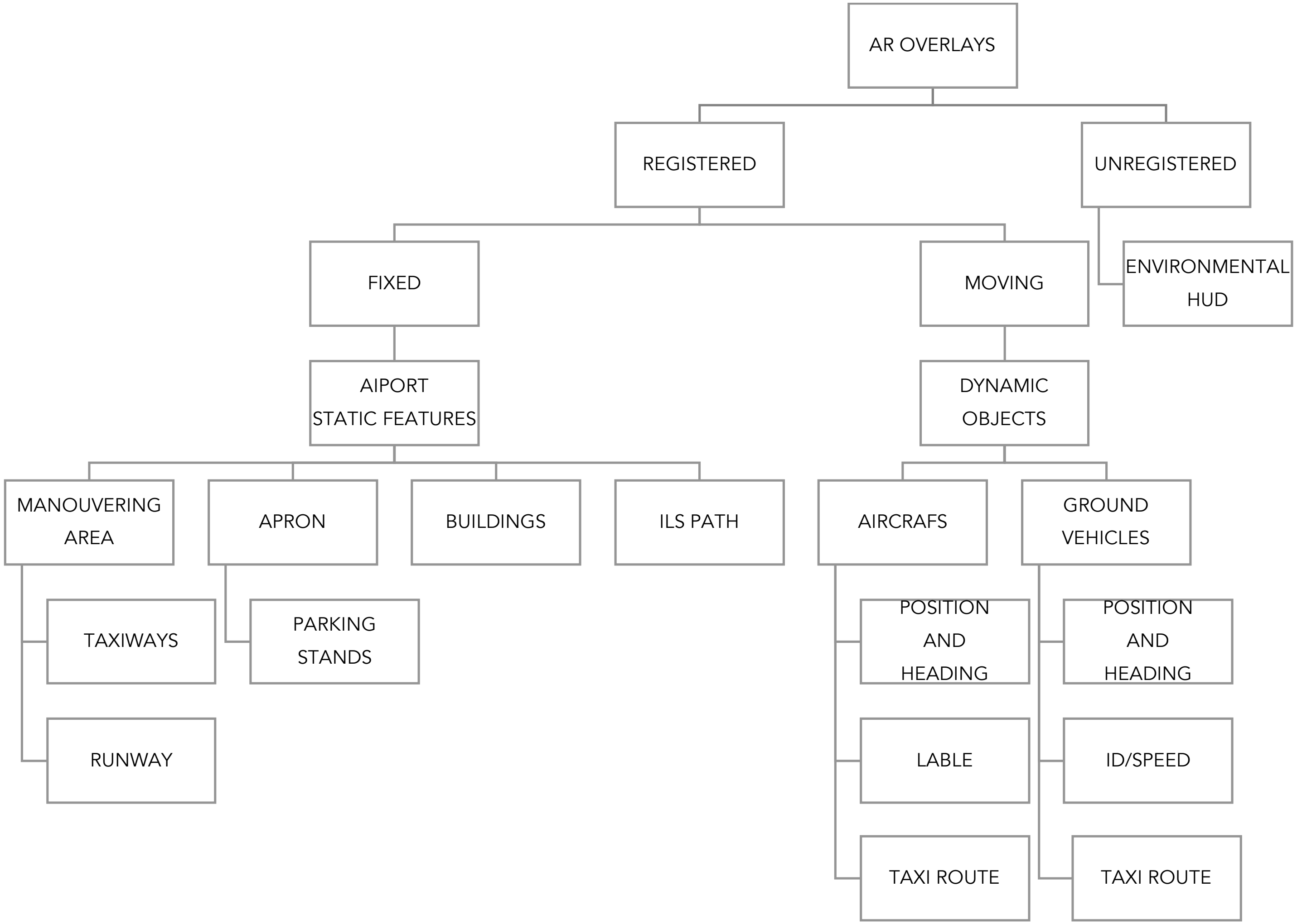


Figure 74. Augmented reality overlays chart

6.4.4.1 AIRPORT STATIC FEATURES OVERLAYS

Airport Static Features Overlays are registered and fixed overlays overlapping the airport main infrastructures. Using simple shapes and primitive geometries they will allow controllers to maintain adequate situation awareness on the airport layout in LVC. They can also communicate status information on certain areas and provide context to other overlays.

Ideally, the static features overlays include a synthetic representation of the main airport buildings (visible by the control tower), the runways, the taxiways, the apron, the taxilanes, the stop bars and all the principal ground signs such as runway ground signs, taxiways centrelines and aircraft parking stand borderlines. Taxiway names and parking stands numbers could be also traced by the AR tools; however, there is a risk to overload the interface with too much information.

By changing the colour of the overlay, relevant status information on the specific feature such as temporary closure or occupancy can be conveyed. On the contrary, it is not recommended to change the shape of the overlay in order not to create misperception on the airport layout.

Depending on the visibility condition, overlaid static features will include taxiways borderlines, parking stands, stop-bars and restricted areas.

A standard colour coding, such as the one depicted in Table XIV, can be used to highlight runway and restricted areas closure/occupancy. Parking stands occupancy can be signalled in the same way.

	COLOUR CODING	
FEATURE	RED	GREEN
RWY	Closed/Occupied	Open/Available
TWY	Closed/Occupied	N/A
PARKING STAND	Occupied	Free

Table XIV. SEMANTIC MEANING OF OVERLAYS COLOUR CODING

The ideal path of an ILS approach could be shown to TWR controllers to improve monitoring of landing operations.

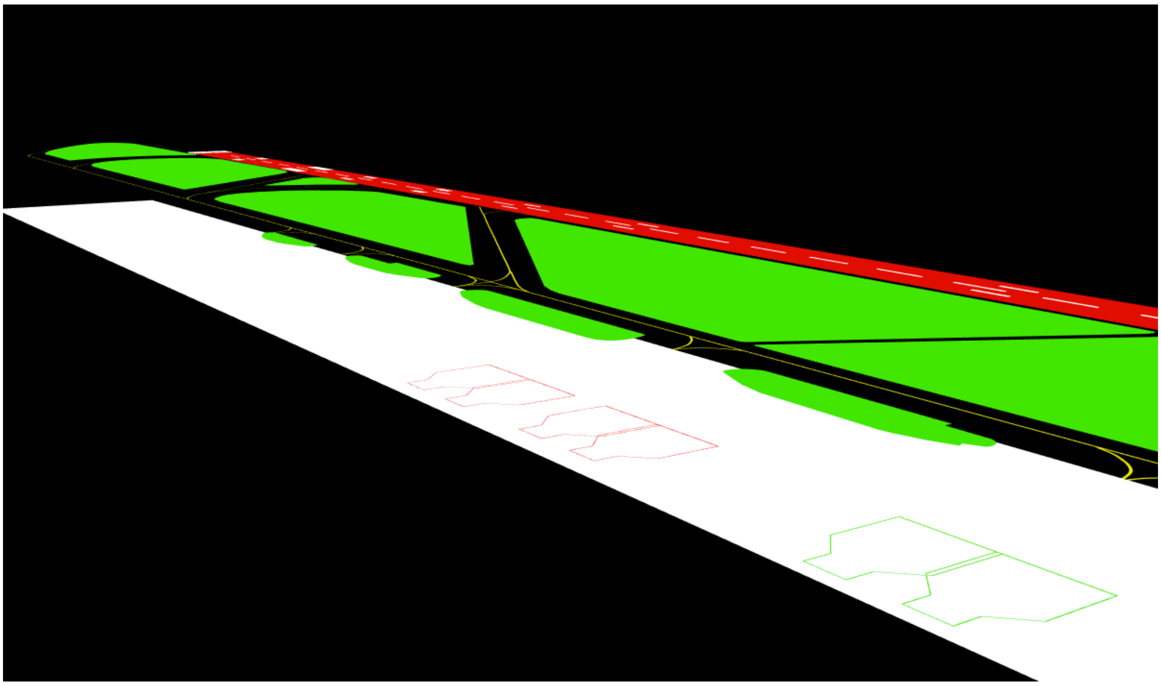


Figure 75. AIRPORT STATIC FEATURES OVERLAYS

6.4.4.1.1 EXPECTED IMPACT

Airport static features overlays will impact controllers' working practice by providing them with easy to understand, updated information on the runway, taxiways and taxilanes status, particularly in LVC. Thus, controllers' will rely less on working memory and on head down equipment and look outside the tower windows to easily retrieve ground related information.

6.4.4.2 DYNAMIC OBJECTS OVERLAYS

Moving objects overlays are the ones 'attached' to airport dynamic features such as aircrafts and ground vehicles.

When looking at far aircrafts (e.g., >1 nm from the control tower) the AR equipment will shows a bounding volume that draws controllers attention and helps them retrieving the aircraft position and heading. This concept can also be applied also to closer aircrafts and ground vehicles when LVP are in force. In this case, primitive shapes projected on the ground will signal aircrafts presence and taxi direction (if on the move). Another way of providing taxi direction is to keep track of the recent historical position and show it to controllers. Assigned taxi routes can be shown on the airport layout in a green fashion.

Alphanumeric text labels can be displayed near active aircrafts that are inside the AR device FOV. Labels will provide controllers with aircrafts' identification (call sign), altitude, speed, type/WCAT, CTOT/EOBT, distance from touch down (only arrival) and ready message (only departure at stand). The displayed information will depend on the aircraft flight phase (departure or arrival). The selected colour coding is depicted in Figure 76.

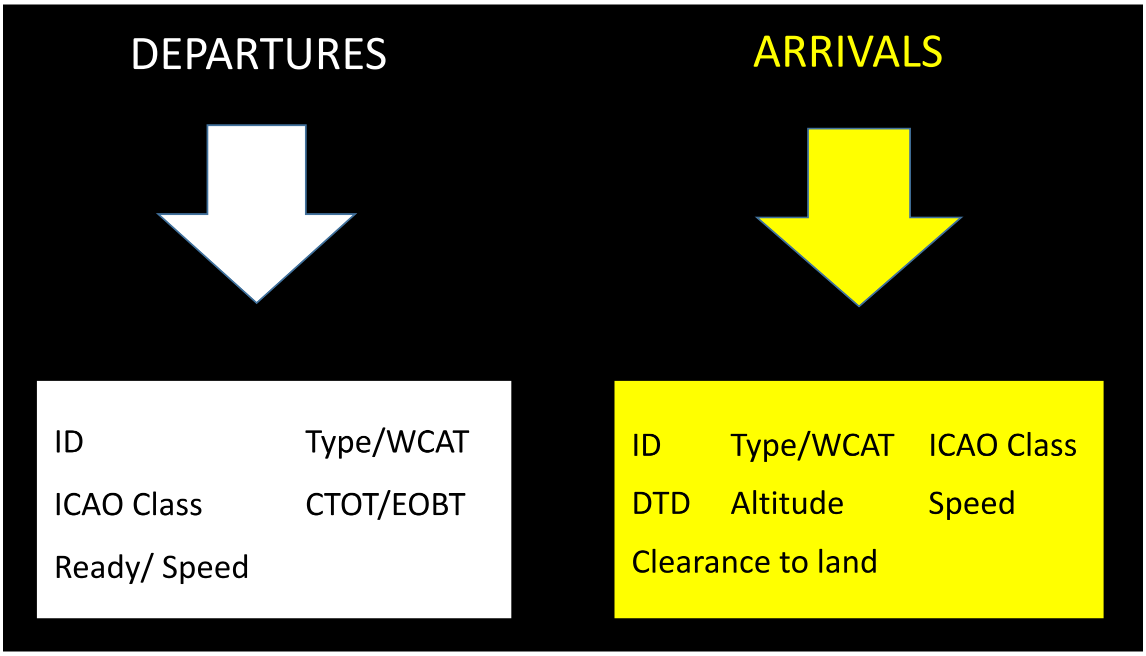


Figure 76. Information displayed by the alphanumeric text labels

Similar text labels could be attached to ground vehicles providing only identification and speed.

6.4.4.2.1 Expected Impact

Billboards will provide controllers with vehicles related information that is currently available only via the integration of many sources, such as radio communication, flight strips and radars. This will simplify controllers' cognitive behaviour when dealing with aircrafts or ground vehicles.

Bounding volumes will draw controller's attention toward aircrafts that are still far from the tower's view and thus can be barely seen by the naked eye. This will increase controllers' situation awareness without forcing them to look at the radar to confirm the position and heading of such aircrafts.

Alphanumeric labels will identify aircrafts by call sign, type/WCAT (Wake turbulence CATegory)

) and provide controller's with CTOT/EOBT, DTD and ready message. This feature will reduce the visual scanning needed to locate specific aircrafts on the apron and manoeuvring area. Also, the time spent looking down at the radar and flight strips to retrieve aircrafts related information should decrease.

Showing the taxi route will allow controllers' to easily double-check the cleared path against taxi blocs that might be closed or restricted to specific aircraft categories. This is expected to simplify controller's cognitive behaviour when performing this task.

6.4.4.3 ENVIRONMENTAL HEAD UP DISPLAY

The HMD will show a semi-transparent display that provides ATCOs with the most relevant environmental information based on the current visibility condition. The displayed information is summarized as follows (see also Table XV):

- NVC: RWY in use, Wind, QNH, RWY surface condition, NAVAIDS status.
- LVC 1: RWY in use, Wind, QNH, RWY surface condition, NAVAIDS status, ceiling.
- LVC 2: RWY in use, wind, QNH, RWY surface condition, NAVAIDS status, RVR (if visibility $<$ 2000 m).
- LVC 3: RWY in use, wind, QNH, RWY surface condition, NAVAIDS status, RVR.

INFORMATION	EXAMPLE
RWY in use	RWY 12
WIND (direction and speed)	60° - 04 Kts
QNH	1024 hPa
RWY surface condition (in colour coding)	▭
CEILING (only if BKN or OVC)	050 BKN
NAVAIDS status (ILS, VOR, NDB, DME)	OK
RVR TDZ (MID, END only if <TDZ)	TDZ 1200

Table XV. INFORMATION DISPLAYED BY THE ENVIRONMENTAL HEAD UP DISPLAY

6.4.4.3.1 EXPECTED IMPACT

Since the display will be positioned on the outside view the controller will be able to retrieve basic environmental information by simply looking thought the control tower windows. This is expected to limit the time the controller will spend looking at the head down equipment and will reduce the number of head movements and attentional shifts between the outside and inside view.

6.4.5 THE SKILL-RULE-KNOWLEDGE ANALYSIS WITH AUGMENTED REALITY TOWER TOOLS

Considering the AR overlays described in paragraph 6.4.4, a further S-R-K analysis was performed to show theoretical shifts in controllers' cognitive behaviour due to the use of AR tools. This time, three scenarios were considered:

- **Scenario 1**: use of AR tools together with local rules[61]. This is represented by the green frames.

- **Scenario 2:** use of standard tools with limited restrictions (NVC rules). This is represented by the combination of grey and red squares. Red squares show the detrimental effects of using NVC rules in LVC.

- **Scenario 3**: use of AR tools with limited restrictions (NVC rules). This is basically the overlap of two effects and it is represented by the combination of green frames with red squares. When the effect of using NVC rules in LVC is counteracted by the ARTT this is shown with a green X. In such cases, the red square must be ignored.

As further discussed in section 0, the analysis of the fist scenario shows the benefits of the ARTT in terms of increased safety, but equal efficiency, capacity and throughput. On the contrary, the analysis of the third scenario shows the benefits in terms of increased efficiency, capacity and throughput, but equal safety.

A more detailed analysis of the selected control tasks is reported in Appendix B.

---

[61] As described in section 5.2.6.

NVC

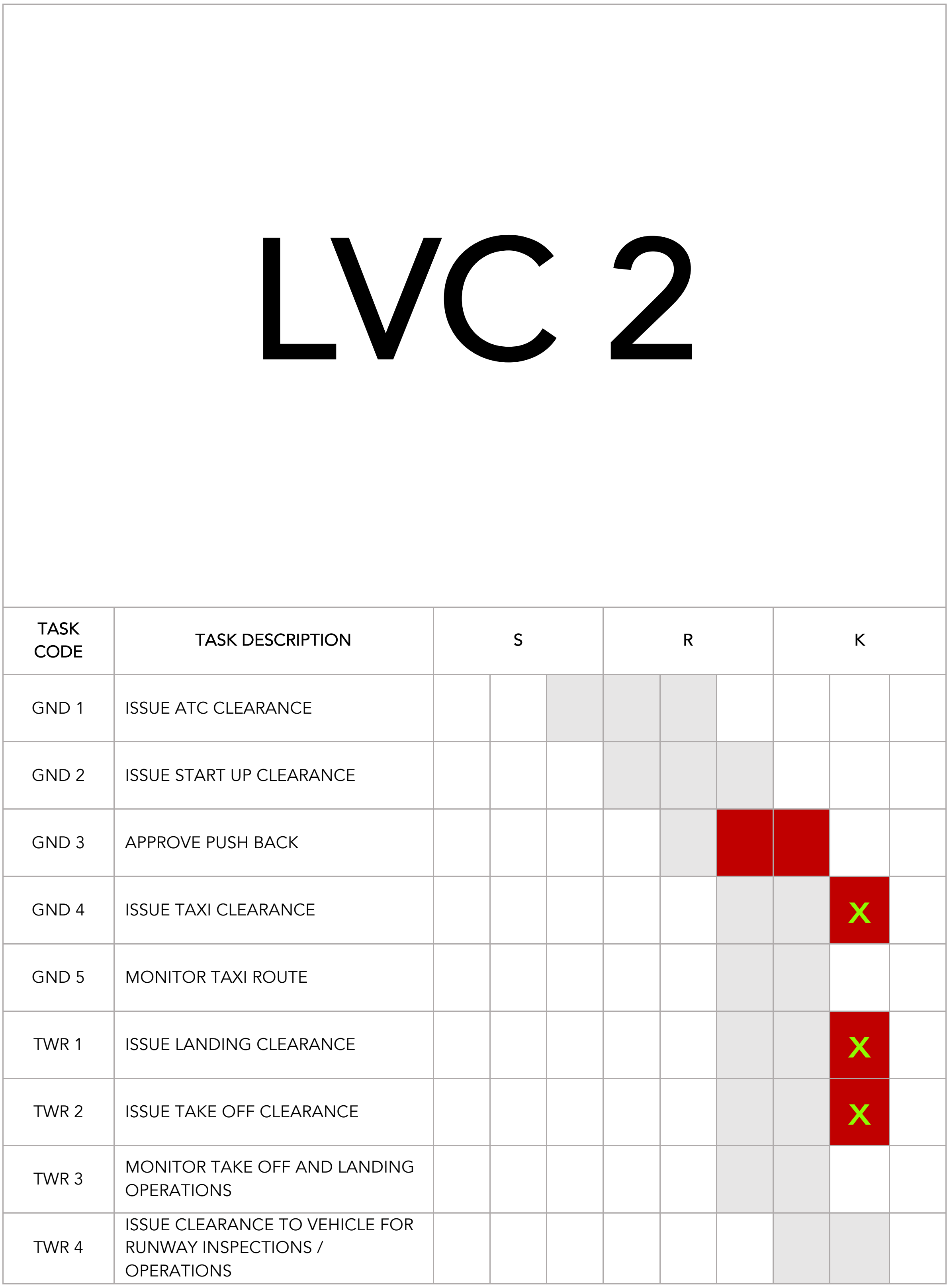
TASK CODE	TASK DESCRIPTION	S			R			K		
GND 1	ISSUE ATC CLEARANCE			▨	▨	▨				
GND 2	ISSUE START UP CLEARANCE				▨	▨				
GND 3	APPROVE PUSH BACK					▨	▨	▨		
GND 4	ISSUE TAXI CLEARANCE						▨	▨		
GND 5	MONITOR TAXI ROUTE					▨	▨			
TWR 1	ISSUE LANDING CLEARANCE						▨	▨		
TWR 2	ISSUE TAKE OFF CLEARANCE						▨	▨		
TWR 3	MONITOR TAKE OFF AND LANDING OPERATIONS					▨	▨			
TWR 4	ISSUE CLEARANCE TO VEHICLE FOR RUNWAY INSPECTIONS / OPERATIONS						▨	▨		

Table XVI. S-R-K ANALYSYS WITH AUGMENTED REALITY - NORMAL VISIBILITY CONDITIONS

LVC 1

TASK CODE	TASK DESCRIPTION	S			R			K		
GND 1	ISSUE ATC CLEARANCE			▨	▨	▨				
GND 2	ISSUE START UP CLEARANCE				▨	▨				
GND 3	APPROVE PUSH BACK					▨	▨	▨		
GND 4	ISSUE TAXI CLEARANCE						▨	▨		
GND 5	MONITOR TAXI ROUTE					▨	▨			
TWR 1	ISSUE LANDING CLEARANCE						▨	▨		
TWR 2	ISSUE TAKE OFF CLEARANCE						▨	▨		
TWR 3	MONITOR TAKE OFF AND LANDING OPERATIONS					▨	▨			
TWR 4	ISSUE CLEARANCE TO VEHICLE FOR RUNWAY INSPECTIONS / OPERATIONS						▨	▨		

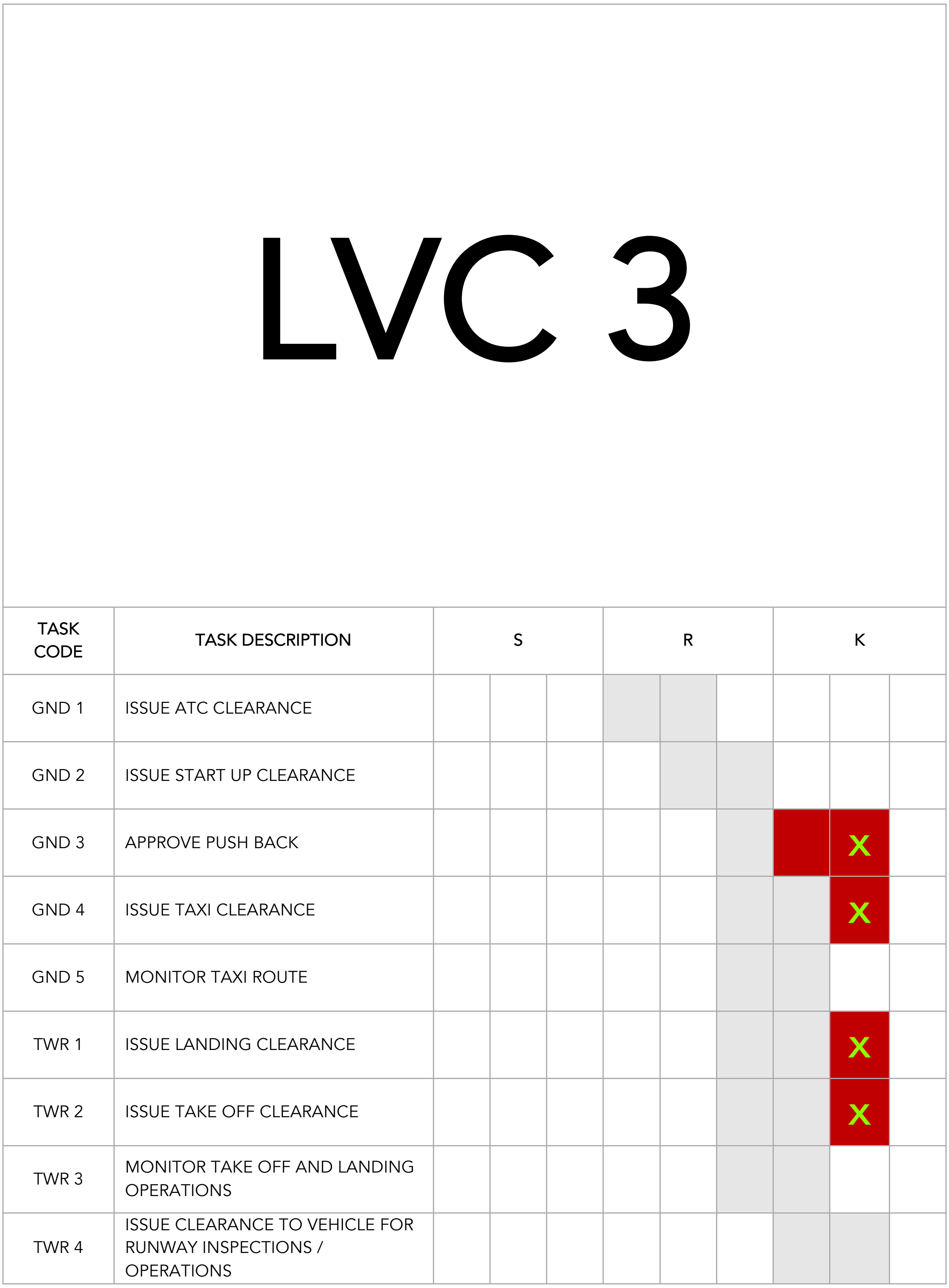
Table XVII. S-R-K analysys with augmented reality - low visibility conditions 1



LVC 2

TASK CODE	TASK DESCRIPTION	S			R			K		
GND 1	ISSUE ATC CLEARANCE			▨	▨	▨				
GND 2	ISSUE START UP CLEARANCE				▨	▨	▨			
GND 3	APPROVE PUSH BACK					▨	🟥	🟥		
GND 4	ISSUE TAXI CLEARANCE						▨	▨	X	
GND 5	MONITOR TAXI ROUTE						▨	▨		
TWR 1	ISSUE LANDING CLEARANCE						▨	▨	X	
TWR 2	ISSUE TAKE OFF CLEARANCE						▨	▨	X	
TWR 3	MONITOR TAKE OFF AND LANDING OPERATIONS						▨	▨		
TWR 4	ISSUE CLEARANCE TO VEHICLE FOR RUNWAY INSPECTIONS / OPERATIONS							▨	▨	

Table XVIII. S-R-K ANALYSYS WITH AUGMENTED REALITY - LOW VISIBILITY CONDITIONS 2



LVC 3

TASK CODE	TASK DESCRIPTION	S			R			K		
GND 1	ISSUE ATC CLEARANCE				▨	▨				
GND 2	ISSUE START UP CLEARANCE					▨	▨			
GND 3	APPROVE PUSH BACK						▨	🟥	X	
GND 4	ISSUE TAXI CLEARANCE						▨	▨	X	
GND 5	MONITOR TAXI ROUTE						▨	▨		
TWR 1	ISSUE LANDING CLEARANCE						▨	▨	X	
TWR 2	ISSUE TAKE OFF CLEARANCE						▨	▨	X	
TWR 3	MONITOR TAKE OFF AND LANDING OPERATIONS						▨	▨		
TWR 4	ISSUE CLEARANCE TO VEHICLE FOR RUNWAY INSPECTIONS / OPERATIONS							▨	▨	

Table XIX. S-R-K analysys with augmented reality - low visibility conditions 3

6.5 Implementing the Overlays

A selected subset of AR overlays has been implemented in the Blender Game Engine as a part of the Bologna airport scenario. Following the eye-coupled approach already discussed in paragraph 6.3.3, and by means of active stereoscopy, these can be rendered and visualized in the Reconfigurable Virtual Environment (see paragraph 6.6.1 for details on the RVE).

From a technical perspective, the AR overlays have been put in a linked scene from where they can be triggered when certain events occur (e.g. runway occupancy by aircrafts or vehicles). Thus, the VR scene and the AR scene are rendered separately and later overlaid. This approach is more complex than rendering only one scene, but is more flexible and allows the developer to extract the AR overlay for post processing or use with a separate device.

Prototyped overlays include the RWY status overlay, a dynamic aircraft bounding box, aircrafts 'attached' labels, visualization of taxi route and the environmental HUD.

6.5.1 The Environmental Head Up Display

The environmental HUD is shown in Figure 78 and fetches data from a selected METAR. The data is interpreted according to the standard METAR schema (Figure 77) and later injected into the HUD.

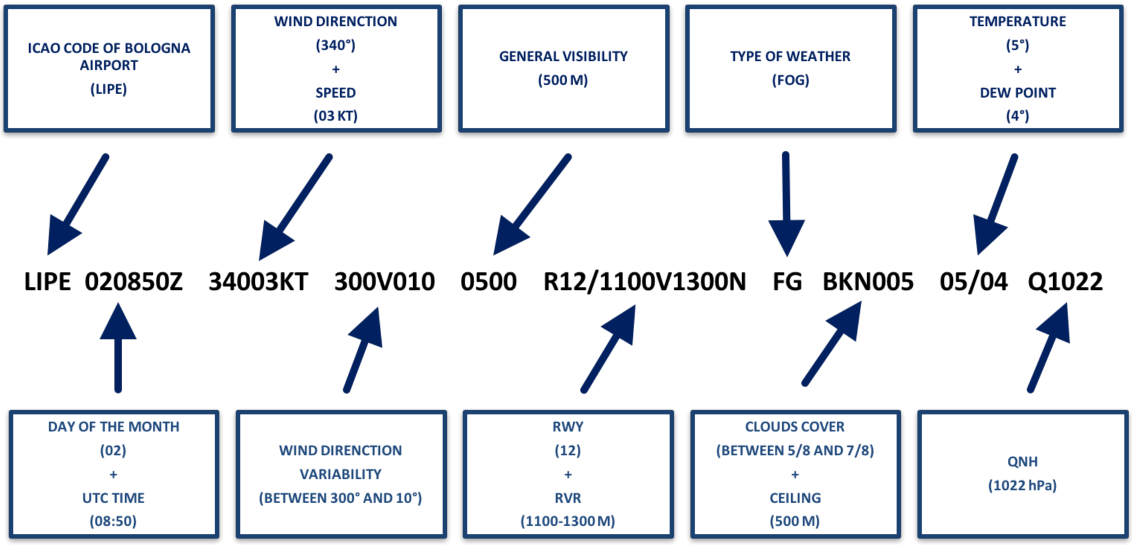


Figure 77. Interpretation of a standard METAR schema

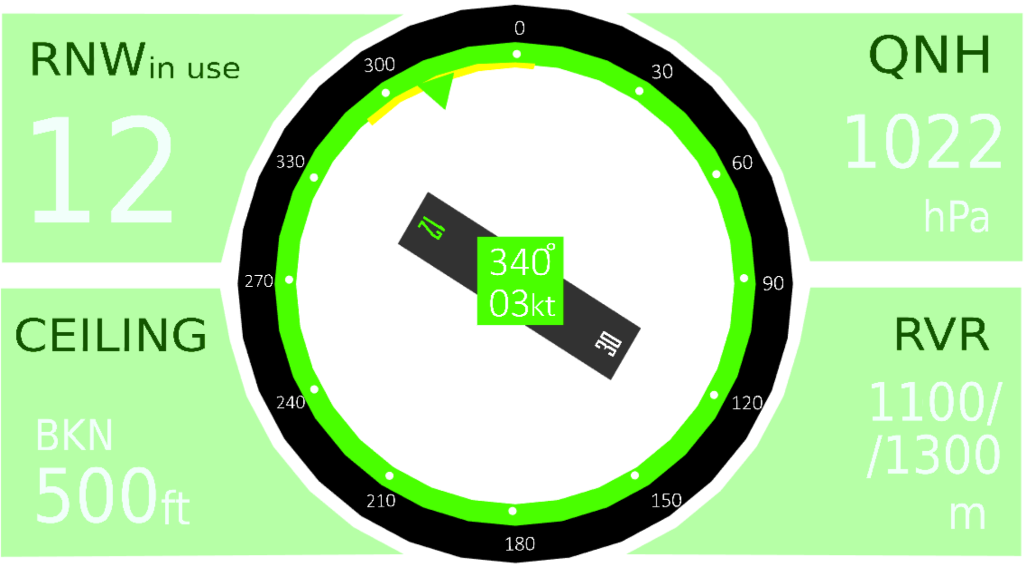


Figure 78. DRAFT DESIGN OF THE ENVIRNOMRNTAL HEAD UP DISPLAY

It can be positioned within the inside or outside view depending on controllers' needs. It can be scaled, rotated and made transparent.

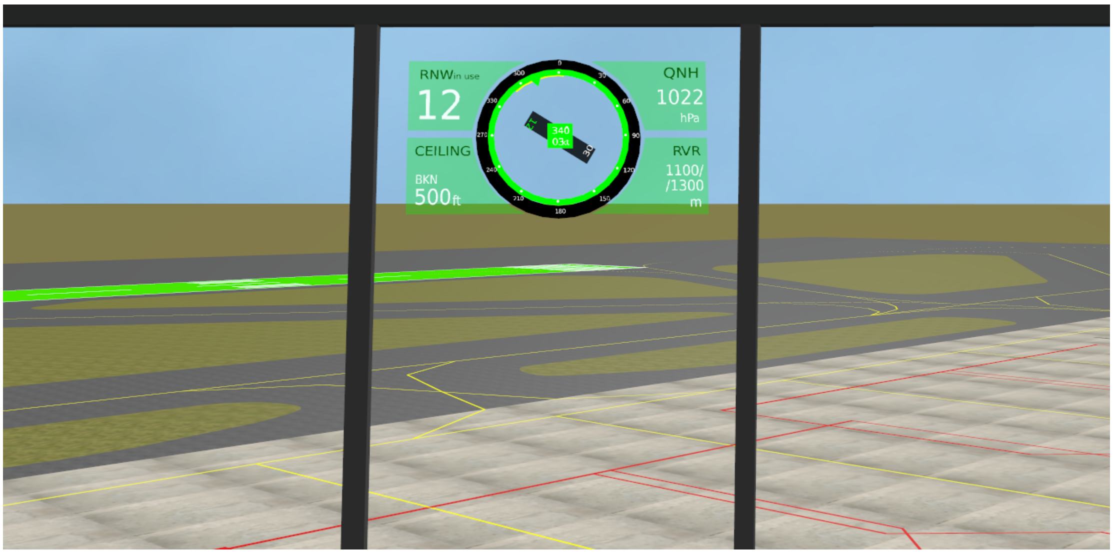


Figure 79. ENVIRNOMRNTAL HEAD UP DISPLAY IN THE SIMULATED CONTROL TOWER

6.5.2 THE RUNWAY STATUS OVERLAY

The RWY status overlay (Figure 85) follows the colour coding depicted in Table XIV and turns red when occupied by aircrafts or vehicles. This provide controllers with easy to access status information on runway occupancy and fasten their reasoning.

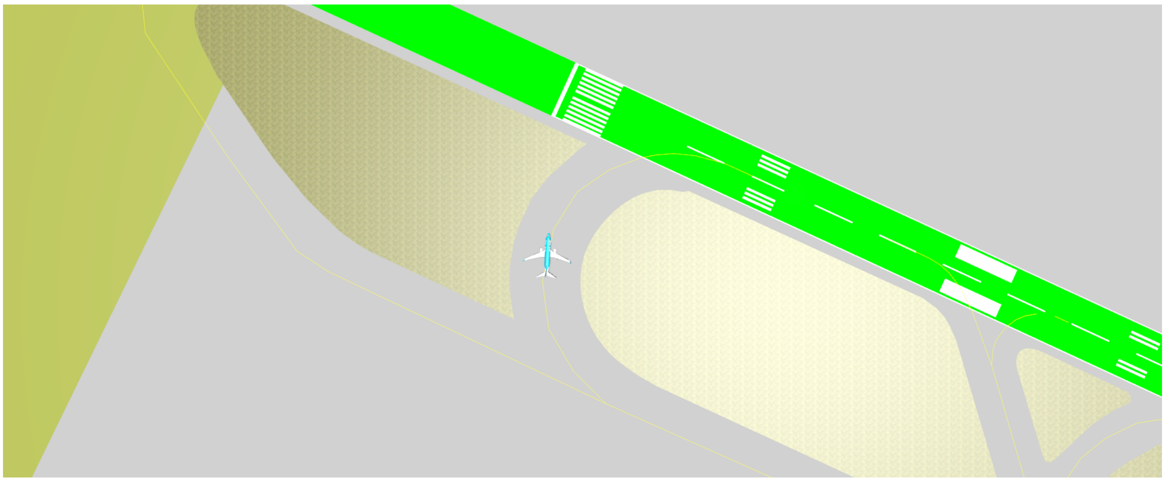


Figure 80. Runway status overlay, top view: runway is free

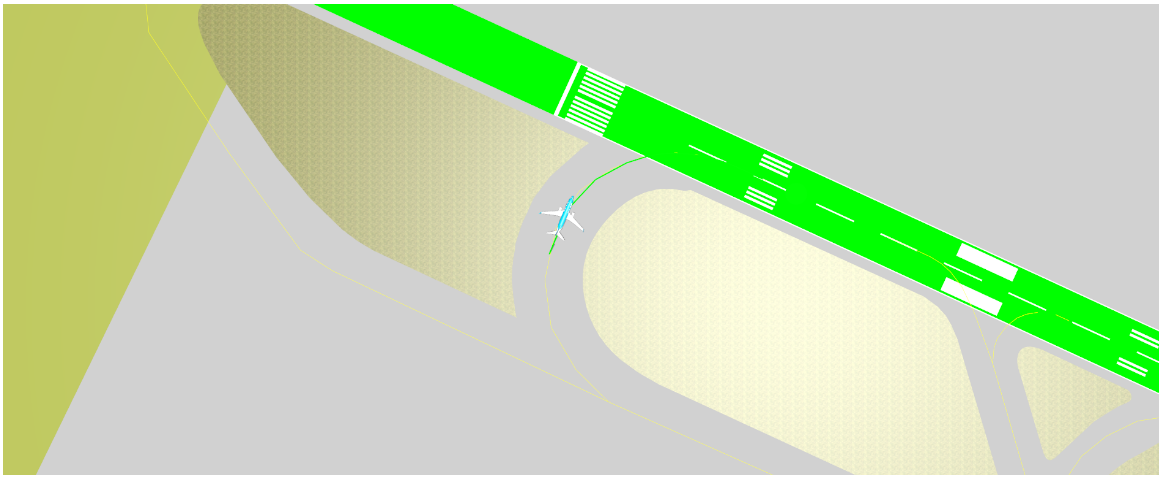


Figure 81. Runway status overlay, top view: taxi phase, runway is still free

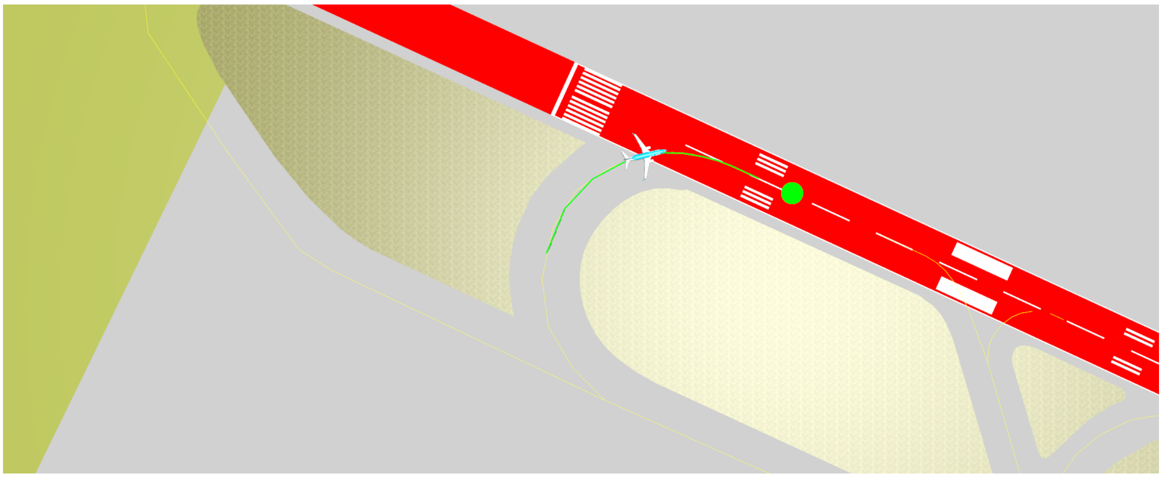


Figure 82. Runway status overlay, top view: runway is occupied

Figure 83. Runway status overlay, control tower view: runway is free

Figure 84. Runway status overlay, control tower view: aircraft is taxiing, runway is still free

Figure 85. Runway status overlay, control tower view: aircraft is ready for take off, runway is occupied

6.5.3 THE DYNAMIC BOUNDING VOLUME

The dynamic bounding volume overlay highlights far aircrafts by means of shrinking concentric circles that are spawned in the main scene with a frequency that resembles the one of the radar signal (~ 0,25Hz). They also provide directional information by keeping track of vehicles historical positions (circles become more and more transparent in time).

6.5.4 AIRCRAFTS ATTACHED LABELS

Aircrafts 'attached' labels follow aircrafts around the scene and always face the control tower direction. They can either be set to scale with the distance or keep a standard dimension. The background colour of the label changes according to the flight phase (yellow: arrivals, white: departures). The information contained in the label changes according to what already described in Figure 76.

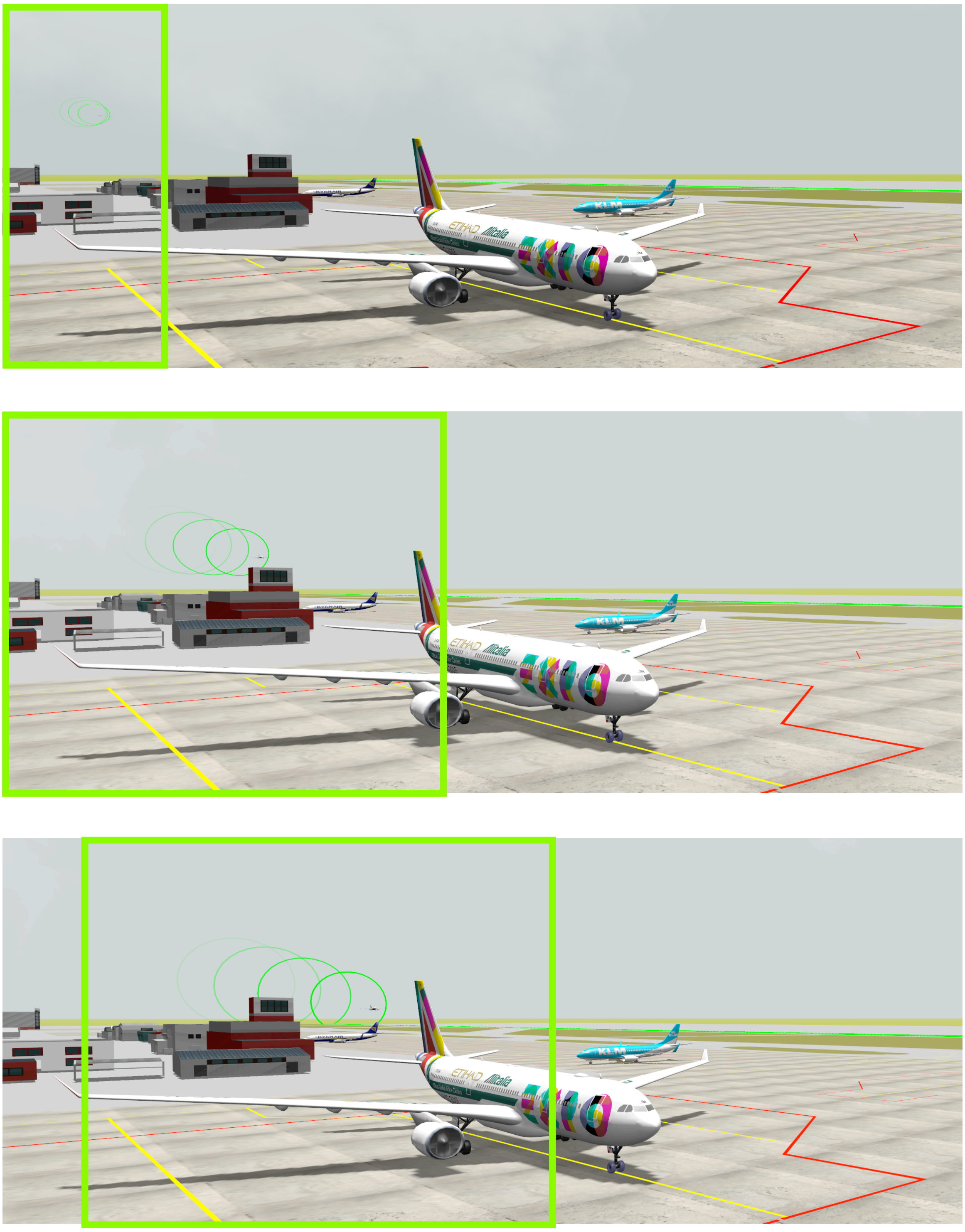


Figure 86. DYNAMIC BOUNDING VOLUME OVERLAY (HIGHLIGHTED BY THE GREEN BOX)

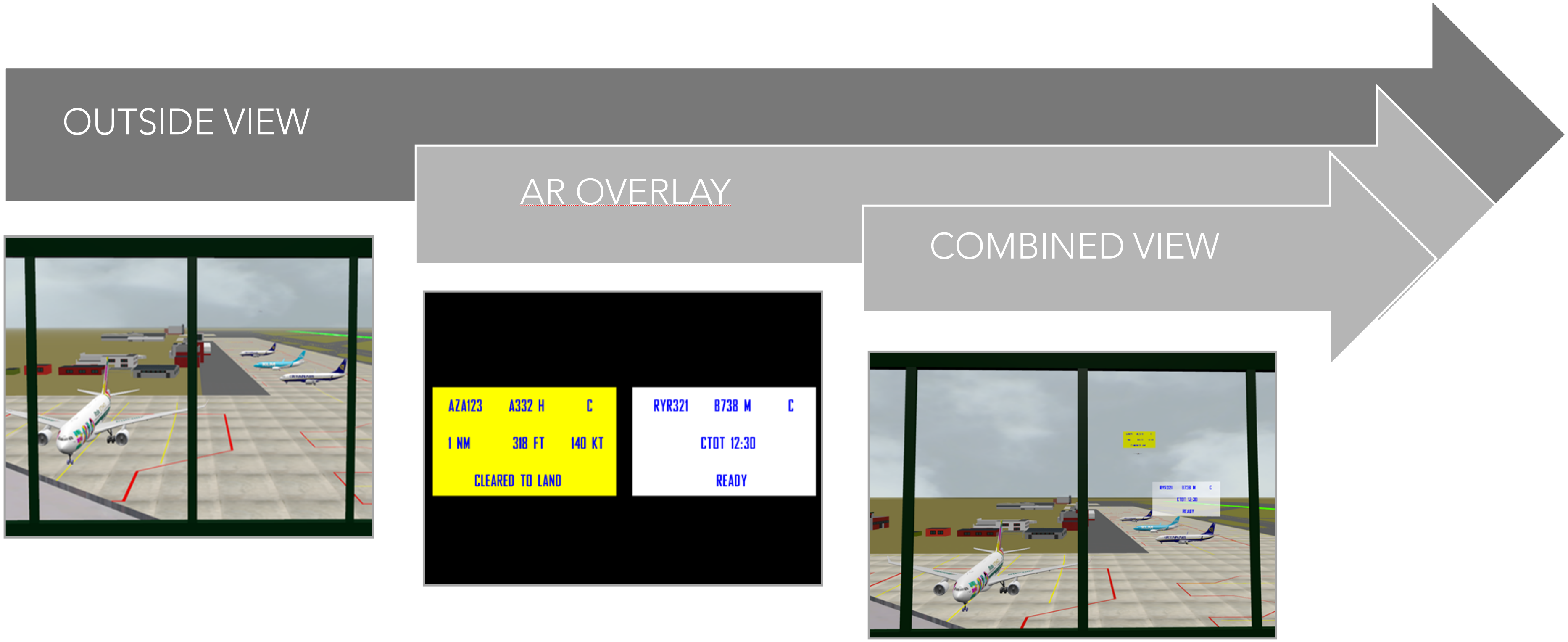


Figure 87. Aicraft lables overlay

6.5.5 THE TAXI ROUTE OVERLAY

A taxi route overlay can be visualized in green once the aircraft path has been cleared. This is useful to controllers as they can check the real path against the cleared route and the cleared route against restricted areas (which could be also overlaid, in red).

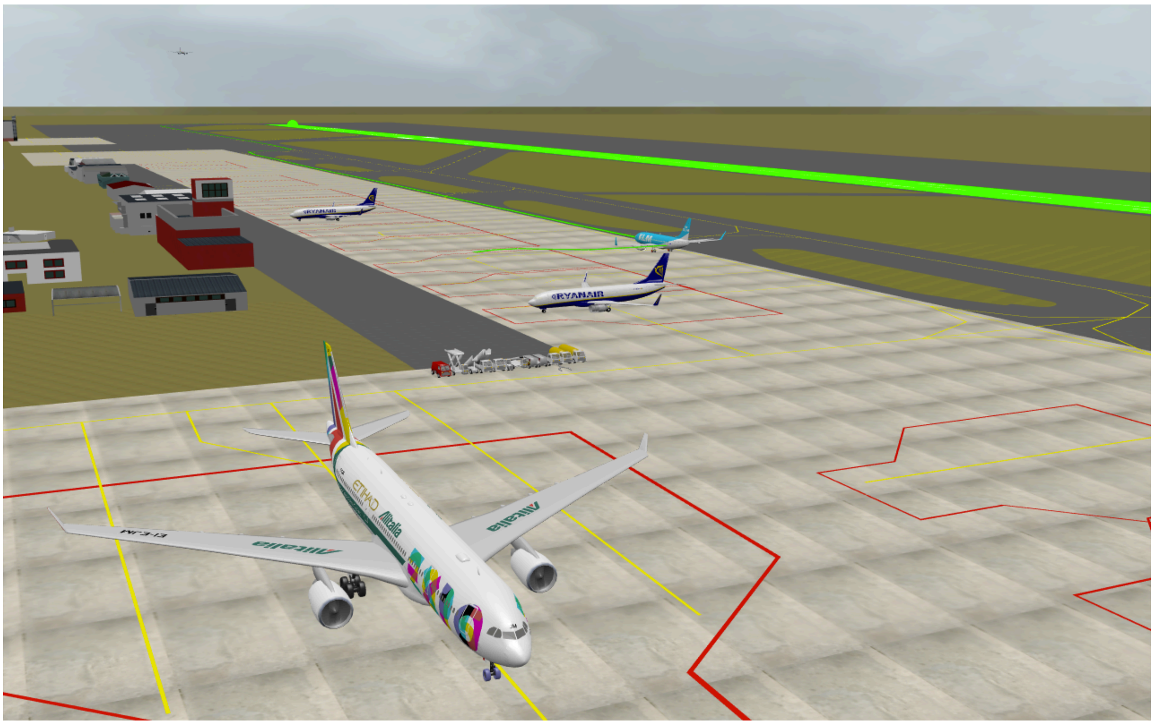


Figure 88. TAXI ROUTE OVERALY

6.5.5.1 OVERLAYS BEHAVIOUR IN LOW VISIBILITY CONDITIONS

Because of the way overlays are treated inside the model they are insensitive to LVC, particularly fog. This is a desired behaviour as the overlays will be provided by an AR layer that resides inside the control tower (regardless of the hardware that will allow this).

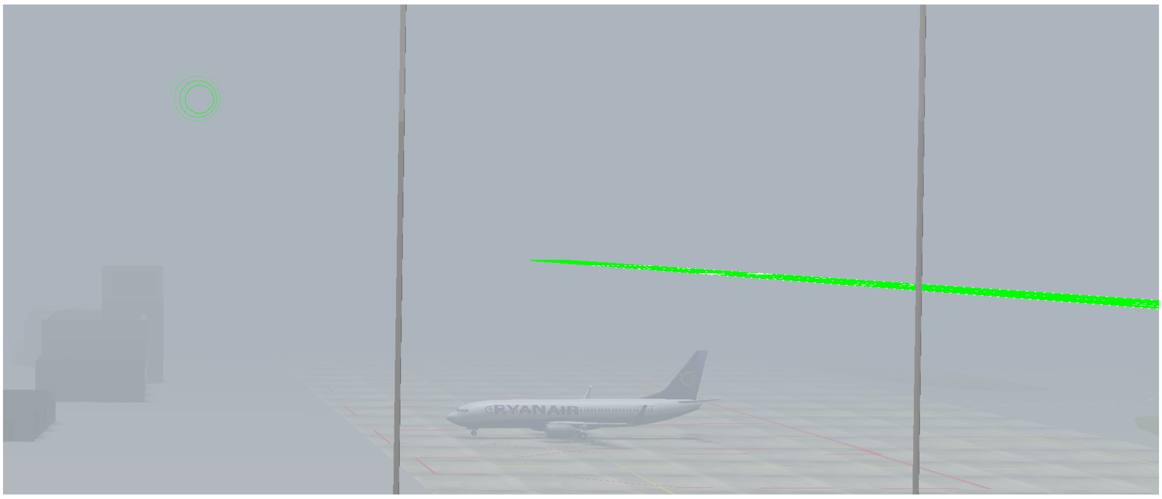


Figure 89. THE RUNWAY OVERLAY AND THE BOUNDING VOLUME OVERLAY IN LOW VISIBILITY CONDITIONS

6.6 SIMULATION ENVIRONMENTS AND AUGMENTED REALITY TOOLS

Hereafter we describe the simulation environments where the AR overlays have been deployed and tested. This equipment is available at the University of Bologna's Virtual Reality and Simulation Laboratory; however, similar facilities can be found all over the world particularly in research centres, universities and academies.

6.6.1 RECONFIGURABLE VIRTUAL ENVIRONMENT

The Reconfigurable Virtual Environment is a CAVE-like virtual environment designed to recreate a sense of immersion by means of three, rear-projected, flat screens. The screens can be arranged in three different configurations, closed, semi-closed and wide open. A stereoscopic 3D effect is obtained by means of active shutter glasses (NVIDIA 3D Vision) and compatible projectors. Head tracking is obtained by means of a Microsoft^{TM}. Kinect^{TM} sensor. A custom rendering pipeline generates images based on the viewer's position as further detailed in chapter 6.3.

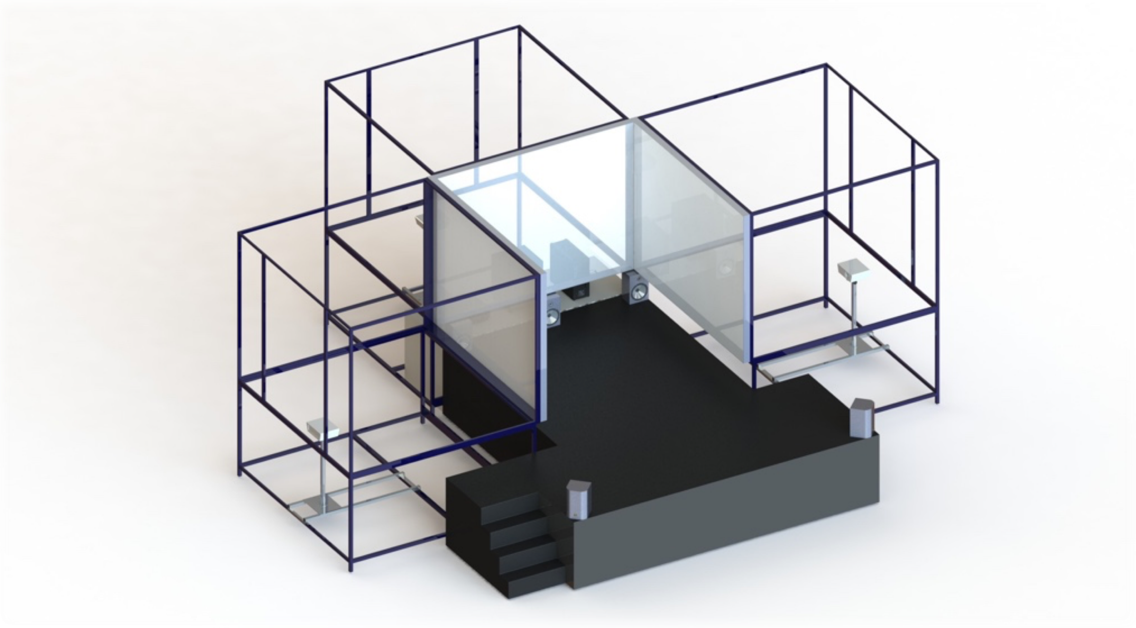


Figure 90. Reconfigurable Virtual Environment: closed configuration

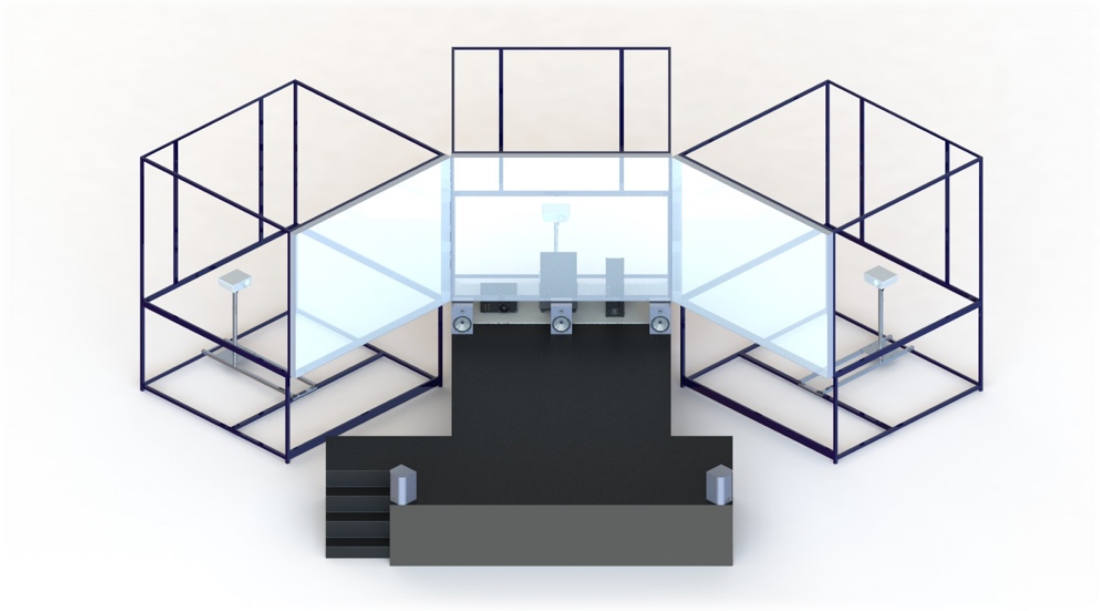


Figure 91. Reconfigurable Virtual Environment: semi-closed configuration

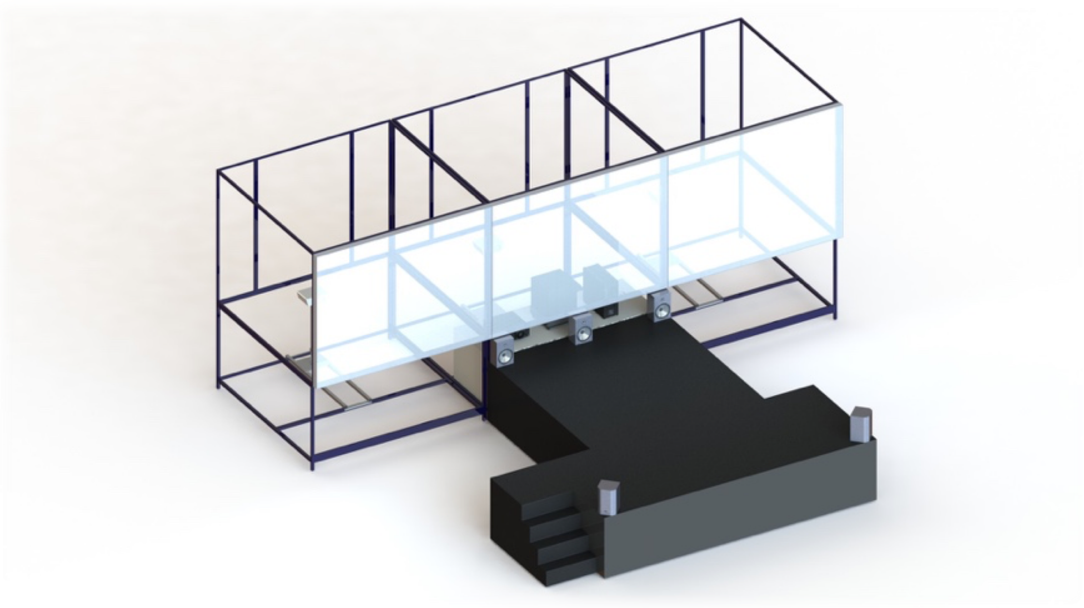


Figure 92. RECONFIGURABLE VIRTUAL ENVIRONMENT: WIDE-OPEN CONFIGURATION

Within this research, the RVE has been used to run the control tower scenario and prototype the augmented reality overlays as if they were shown by a see-through spatial AR display. It was also used as a background scene when prototyping with the Microsoft^{TM} Hololens^{TM}.

Figure 93. THE RECONFIGURABLE VIRTUAL ENVIRONMENT INSIDE THE VIRTUAL REALITY LABORATORY (VLAB)

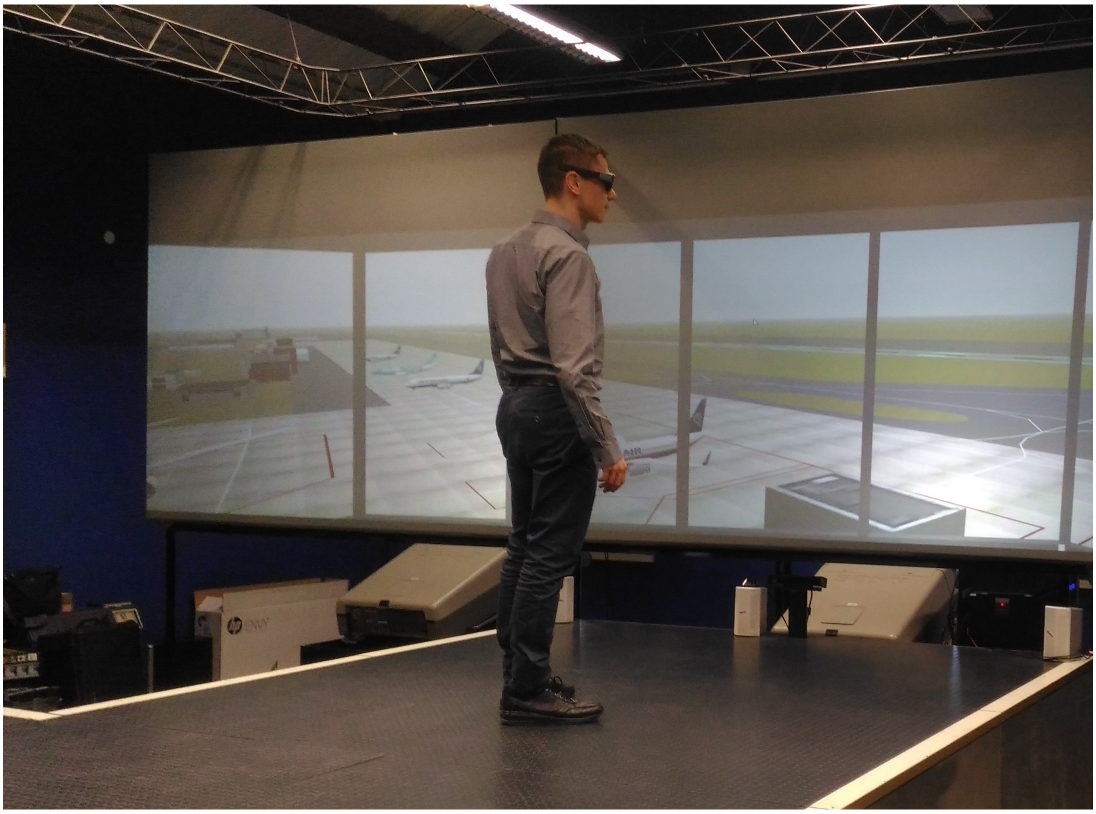


Figure 94. The Reconfigurable Virtual Environment running the bologna airport scenario[62]

---

[62] Photo was taken in monoscopic mode

6.6.2 MICROSOFT HOLOLENS

The Microsoft^{TM} HoloLens^{TM} is an optical see-through head mounted display developed and manufactured by Microsoft^{TM}. The device can trace its lineage to the Kinect, a previous device that let Microsoft^{TM} experiment with ambient and body tracking sensors and gestures. The pre-production version of HoloLens^{TM}, a.k.a. Development Edition, shipped on the 30th of March 2016, and was targeted to developers in the United States and Canada. On October 12, 2016, Microsoft^{TM} announced global expansion of HoloLens^{TM} and publicized that HoloLens^{TM} would be available by the end of the year in Australia, Ireland, France, Germany, New Zealand and the United Kingdom.

From a hardware perspective, much of the sensors and related hardware is enclosed in the front part of the device. The HoloLens^{TM} features: (a) an Inertial Measurement Unit (IMU) which includes an accelerometer, gyroscope, and a magnetometer, (b) four grayscale “environment understanding" cameras (two on each side), (c) a depth camera with a 120°×120° FOV, (d) a 2.4-megapixel video camera, (e) a four-microphone array and (f) an ambient light sensor. A "light engine" projects images into a pair of combiner lenses enclosed in a visor piece. The HoloLens^{TM} uses optical waveguide to deliver blue, green, and red colours separately. Like many other optical HMDs, the display projected by the HoloLens^{TM} occupies a limited portion of the user's FOV which is about 30°×17.5°. The device provides the possibility to customize images generation based on the viewer’s interpapillary distance (IPD). Near the user's ears, are a pair of small, red 3D audio speakers. Compared against typical headsets these speakers do not obstruct external sounds, allowing the user to hear computer generated sounds, along with environmental sounds. Using head-related transfer functions, the HoloLens^{TM} can simulate spatial sound; meaning that the user can perceive and locate a sound, as though it is coming from a virtual pinpoint or location. In addition to an Intel Cherry Trail SoC containing the CPU and GPU. HoloLens^{TM} features a custom-made Microsoft^{TM} Holographic Processing Unit (HPU), a coprocessor manufactured specifically for the HoloLens^{TM} by Microsoft^{TM}. This integrates data from the sensors and handles tasks such as spatial mapping, gesture recognition, voice and speech recognition. The internal battery allows the device to have an average autonomy of 2–3 hours of active use, or about two weeks of standby time. The device can also be operated while charging. As for wireless connectivity HoloLens^{TM} features IEEE 802.11ac Wi-Fi and Bluetooth 4.1 Low Energy (LE).

The device comes bundled with a thumb-sized input device named “the Clicker” that can be used for selecting and scrolling, via tilting and panning or clicking. For further interaction, Natural User Interface (NUI) commands such as gaze, gesture and voice commands can be used. These are sometimes referred to as GGV (Gaze, Gesture and Voice) inputs. Gaze commands, such as head-tracking, allows the user to bring focus to whatever he or she is looking at. Elements or buttons are selected using an air tap gesture. The tap can be held to drag or resize elements. A "bloom" gesture, which consists of opening one's hand fingers spread with the palm facing up, can be performed to access the main menu. This function is like pressing a Windows key on a Windows^{TM} equipped PC or tablet. Virtual elements such as windows or menus can be "pinned" to locations, physical structures or objects within the environment; or can be "carried" around following the user.

Microsoft^{TM} recommends using Visual Studio^{TM} IDE to develop applications (both 2D and 3D) for HoloLens^{TM}, which can be tested using HoloLens^{TM} emulator (included into Visual Studio 2015 IDE). HoloLens^{TM} can run most of Universal Windows Platform apps. and the same tools that are used to develop applications for Windows PC or Windows Phone can be used to develop a HoloLens^{TM} app. 3D applications use Windows^{TM} Holographic APIs. Microsoft^{TM} recommends the Unity Game Engine and Vuforia to create 3D apps for the HoloLens^{TM}. However, it is also possible for a developer to build his or her own engine using DirectX and Windows APIs.

In April 2016 Microsoft^{TM} Created the Microsoft^{TM} HoloLens^{TM} App for Windows^{TM} 10 PC's and Windows 10 Mobile devices, that allows developers to run apps, use his or her phone or PC's keyboard to type text, view a live stream from the HoloLens^{TM} user's point of view, and remotely capture mixed reality photos and videos.

Within this research, the HoloLens^{TM} was used as an experimental HUD. As the device was made available only in early 2016 there was no time to implement complex overlays, (although there are plans to do so in the future). For the time being, most of the effort was focused on customizing the environmental HUD that had been already prototyped in the BGE.

Figure 95. Microsoft™ HoloLens™

Figure 96. Draft design of the environmental head up display on the Microsoft™ HoloLens™

Figure 97. HEAD-UP DIPSLAY OVERLAY WITH THE MICROSOFT™ HOLOLENS™

7 CONCLUSION AND FUTURE DEVELOPMENTS

Various analysts have estimated the benefits of using AR tools in the control tower. However, it has been rarely specified how such tools should be designed and operated. Most of the prototypes have been based on iterative design-test cycles and our review confirms that many problems must be addressed before these tools become operational.

In this research, a theoretical framework for the selection of the best augmented reality tools has been designed. This is based on an integrated Quality Function Deployment – Analytic Hierarchy Process approach. Such methodology has been applied to the selection of the best augmented reality technologies for the control tower, showing that head-mounted displays are probably the easiest-for-development but not the best-for-use choice. Although their usability is expected to improve, I would recommend continuing the research on both types of devices (spatial and head-mounted). A combined use of such technologies seems also reasonable.

The same method that was used in this research for ranking augmented-reality technologies could be used for choosing the best equipment within a single category of devices. Conversely, it could be used to define specifications for augmented reality devices that will be developed in the future.

A reference scenario was chosen, i.e. Bologna G. Marconi international Airport (ICAO code LIPE). For this scenario, a 4D model (3D + time) was developed in Blender and tested in the Blender Game Engine. The model includes most of the airport static features and ground signs. Also, a small library of aircrafts and ground vehicles was set up.

A “point and click” interface for managing aircrafts and ground vehicles, assigning taxi routes ad clearing take-offs and landings was developed. The ground movements are based on a custom implementation of A* pathfinding algorithm. The algorithm uses designated objects (e.g. taxiways centrelines) as a navigation mesh. Take-offs and landings have been implemented through animations and blended with the ground movement system by means of BGE visual programming (logic bricks).

For the viewport, a custom rendering pipeline was designed and implemented. This incorporates eye-coupled perspective and allows the designer to deal with multiple, arbitrary oriented displays. This is extremely useful when rendering stereoscopic virtual/augmented reality contents (for any type of display) which would otherwise contain unacceptable registration errors. Thanks to the use of a Microsoft^{TM} Kinect^{TM} sensor, there is no need for the user to wear additional head tracking equipment.

The design of the augmented reality overlays was based on an 'ecological' approach and was performed in close partnership with air traffic controllers. Several overlays where conceived and prototyped. A selected sample was implemented in the Blender Game Engine and integrated within the airport 4D model. The simulation environment primarily consists of a Reconfigurable Virtual Environment system, but with slight modifications the developed components should be flexible enough to handle other kind of virtual/augmented reality systems, such as table tops, see-through spatial displays or head mounted displays.

The Skill-Rule-Knowledge analysis makes it possible to show how the augmented reality tools support the controller in performing his or her tasks. Overall, there is a theoretical shift in the skill-rules-knowledge paradigm that favours rule-based behaviour over the knowledge based behaviour. This becomes more relevant in LVC: as the visibility decreases, restoring controllers' situation awareness by means of augmented-reality tools leads to growing benefits. The benefits in normal visibility conditions are more linked to the reduction of head down operations, to the less refocusing between the inside and outside view and to the availability of contextualized (registered) information.

A future development will test the prototyped overlays by means of a validation campaign performed in a laboratory setting. This will be done within the RETINA[63] project. In the future, the experiment could be taken to the control tower for a shadow control session[64].

Futuristic usage of augmented reality could lead to the capability of superimposing information and instructions not only over the outside view, but also over control panels and instruments of controllers’ working positions. This could be used for training and early practice in a new tower environment.

---

[63] http://www.retina-atm.eu

[64] During a shadow control session, a controller pretends to control the traffic, while a colleague is actually in charge.

APPENDIX A

This section presents the code that was developed to exploit the Blender Game Engine functionalities with University of Bologna's Reconfigurable Virtual Environment. The code implements eye-coupled perspective and manages the airport scenario. The system also accepts voluntary user inputs, such as mouse and keyboard, and involuntary user inputs such as head movement.

Most of the code was written in Python[65] 3 and tested against Blender v2.78 and the Blender Game Engine version that comes bundled with it. Our programming takes advantage of the standard BGE API[66] and a few other Python packages and modules, including more functions and modules programmed by the author of this theses.

It shouldn't pose to much of an issue to convert such code in other languages or modify its parameters to handle other types of V/AR equipment.

A Visual C# program handles the Microsoft™ Kinect™ and sends the retrieved data over the network.

---

[65] https://www.python.org

[66] https://docs.blender.org/api/

APPENDIX A-1

This section presents the code for the viewport management.

viewport.py

```
##########################################################################

# This script handles the viewport for a 3 screen V/AR system
# the position of the screens corners are retrieved from the Blender
Scene

from bge import logic, types, render
from mathutils import Matrix, Vector

# Global variables

scene = logic.getCurrentScene()
objects = scene.objects
gD = logic.globalDict
skt = None
tracker = None
head = None
left_eye = None
right_eye = None


# Global constants

W = render.getWindowWidth()
H = render.getWindowHeight()
LEFT_EYE = render.LEFT_EYE
RIGHT_EYE = render.RIGHT_EYE
EYE_SEPARATION = 0.06


##########################################################################

# Classes

# a class that handles (stereo) pre-draw
class Stereo:

  @staticmethod
  def get_current_eye_object():

    try:
      if render.getStereoEye() == 1:
        return left_eye
      elif render.getStereoEye() == 2:
```

```
        return right_eye
    except:
      raise Exception('Stereo is not active')

  @staticmethod
  def pre_draw_setup():

    for obj in objects:
      if type(obj) == Camera:
        obj.worldPosition =
Stereo.get_current_eye_object().worldPosition

    for eye in left_eye, right_eye:
      if eye.side == render.getStereoEye():
        for n in range(1, 4):
          if 'projection' + str(n) in eye:
            eye['projection' + str(n)].projection_cycle()


# A class that represents the tracking system
class Tracker(types.KX_GameObject):

  def __init__(self, empty):
    self.t = self.worldPosition # This is the translation vector from
tracker-space to world-space
    self.R = self.orientation.transposed()


# A derived class for Blender Camera objects
class Camera(types.KX_Camera):

  def __init__(self, camera):
    # this is for the triple screen config
    self.number = self['number']
    self.left = int(round(W * (self.number - 1) * 1 / 3))
    self.bottom = 0
    self.right = int(round(W * self.number * 1 / 3))
    self.top = H
    self.setViewport(self.left, self.bottom, self.right, self.top)
    self.useViewport = True


# a derived class representing the V/AR screens
class Screen(types.KX_GameObject):

  def __init__(self, plane):
    self.number = self['number']
    for child in self.children:
      if 'screen' not in child:
        setattr(self, child.name[0:2], child.worldPosition)
        # stands for self.p = p.worldPosition, for p in [pa,pb,pc]


# a derived class representing the viewer's head.
class Head(types.KX_GameObject):
```

```
  def __init__(self, empty):
    pass

  def set_head_position(self):
    if gD:
      self.worldPosition = (gD['head_position_in_tracker_space'] *
tracker.R) + tracker.t
      print("Head worldPosition is: {}".format(self.worldPosition))
    else:
      print("GlobalDict is empty, not updating head position")


# a derived class representing the viewer's eyes
class Eye(types.KX_GameObject):

  def __init__(self, empty):
    if 'left' in self:
      self.worldPosition -= Vector((EYE_SEPARATION/2, 0, 0))
      self.side = 1

    elif 'right' in self:
      self.worldPosition += Vector((EYE_SEPARATION/2, 0, 0))
      self.side = 2


# Each instance of this class handles the projections for a single eye-
screen pair
class Projection:

  def __init__(self, camera, screen):
    self.camera = camera
    self.screen = screen

    # set the near and far clipping planes distances for the current
projection
    # keep the ones defined in Blender UI
    self.n = camera.near
    self.f = camera.far

    self.va = None
    self.vb = None
    self.vc = None

    self.sr = None
    self.su = None
    self.sn = None

    self.l = None
    self.r = None
    self.b = None
    self.t = None

    self.d = None
```

```
    self.M = None
    self.P = None

  def projection_cycle(self):
    self.update_frustum_edges()
    self.compute_screen_orthonormal_basis()
    self.update_distance_to_screen()
    self.update_screen_extents()
    self.update_projection_matrix()

  def update_frustum_edges(self):
    # convert screen corners' world coordinates into eye-space
coordinates
    # these are the frustum edges in eye space
    self.va = self.camera.world_to_camera * self.screen.pa
    self.vb = self.camera.world_to_camera * self.screen.pb
    self.vc = self.camera.world_to_camera * self.screen.pc

  def compute_screen_orthonormal_basis(self):
    # compute an otho-normal basis that defines the screen's local
coordinate system orientation
    self.sr = (self.vb - self.va)
    self.sr.normalize()
    self.su = (self.vc - self.va)
    self.su.normalize()
    self.sn = self.sr.cross(self.su)
    self.sn.normalize()

    # compute the transformation matrix that maps the screen space
coordinate system
    # onto the camera space coordinate system transforming the otho-
normal basis (sr,su,sn)
    # into camera space basis (x,y,z)

    self.M = Matrix.Identity(4)
    self.M.zero()

    self.M[0][0] = self.sr[0]
    self.M[0][1] = self.sr[1]
    self.M[0][2] = self.sr[2]
    self.M[1][0] = self.su[0]
    self.M[1][1] = self.su[1]
    self.M[1][2] = self.su[2]
    self.M[2][0] = self.sn[0]
    self.M[2][1] = self.sn[1]
    self.M[2][2] = self.sn[2]
    self.M[3][3] = 1

  def update_distance_to_screen(self):
    # compute the distance from eye to screen plane
    self.d = -self.sn.dot(self.va)

  def update_screen_extents(self):
    # find the screen extents of the perpendicular off-axis perspective
projection and scale them
```

```
    # to the near clipping plane
    self.l = self.sr.dot(self.va) * self.n / self.d
    self.r = self.sr.dot(self.vb) * self.n / self.d
    self.b = self.su.dot(self.va) * self.n / self.d
    self.t = self.su.dot(self.vc) * self.n / self.d

  def update_projection_matrix(self):
    # build the standard 3D perspective projection matrix for the
current frustum
    self.P = Matrix.Identity(4)
    self.P.zero()

    self.P[0][0] = 2 * self.n / (self.r - self.l)
    self.P[0][2] = (self.r + self.l) / (self.r - self.l)
    self.P[1][1] = 2 * self.n / (self.t - self.b)
    self.P[1][2] = (self.t + self.b) / (self.t - self.b)
    self.P[2][2] = -(self.f + self.n) / (self.f - self.n)
    self.P[2][3] = -2 * self.f * self.n / (self.f - self.n)
    self.P[3][2] = -1

    # set the final projection matrix as the composition of everything
    self.camera.projection_matrix = self.P * self.M


########################################################################


# Module execution entry point

def main(cont):

  own = cont.owner

  if 'viewport_init' not in own:
    init(cont, own)
  else:
    head.set_head_position()
    # for eye in left_eye, right_eye:
      # print("{} worldPosition is: {}".format(eye, eye.worldPosition))


########################################################################

# initialization function that runs once

def init(cont, own):

  try:
    render.setEyeSeparation(0.0)

    try:
      for obj in objects:

        if 'tracker' in obj:
          global tracker
```

```
          tracker = Tracker(obj)

        elif 'head' in obj:
          global head
          head = Head(obj)

        elif 'eye' in obj:
          if 'left' in obj:
            global left_eye
            left_eye = Eye(obj)

          elif 'right' in obj:
            global right_eye
            right_eye = Eye(obj)

        elif 'screen' in obj:
          Screen(obj)

      if '__default__cam__' not in scene.cameras: # if in 'Active
Camera' view
        for camera in scene.cameras:
          Camera(camera) # re-instancing KX_Camera objects
    except:
      raise Exception('Failed to re-instance some Blender object')

    for obj in objects:

      if type(obj) == Eye:

        eye = obj
        for n in range(1, 4):

          screen = None
          camera = None
          for object in objects:

            if type(object) == Screen:

              if object['number'] == n:

                screen = object

            if type(object) == Camera:

              if object['number'] == n:

                camera = object

          if screen is not None and camera is not None:

            eye['projection' + str(n)] = Projection(camera, screen)
            # may be set as an attribute as well
            # setattr(eye, 'projection'+str(n), Projection(camera,
screen))
```

```
    scene.pre_draw_setup.append(Stereo.pre_draw_setup)

    own['viewport_init'] = True

  except:
    raise Exception("Viewport initialization failed")
```

Appendix A-2

This section presents the code for receiving data over the network.

network.py

```
#########################################################################

# This script receives head coordinates data from the network

import socket
import bge
from bge import logic
from mathutils import Vector


# Global variables

gD = logic.globalDict # shorten the syntax
skt = None


# Classes

# a class that defines a custom socket
class Socket(socket.socket):

  def __init__(self, *args):
    super().__init__(args[0], args[1])
    self.setsockopt(socket.SOL_SOCKET, socket.SO_REUSEADDR, 1)
    # reuse a local socket in TIME_WAIT state without waiting for    its
natural timeout to expire
    self.setblocking(False) # set socket to non-blocking mode
    self.bind(("127.0.0.1", 9999))
    print('Socket bind to IP 127.0.0.1 and port 9999')

  def receive(self):
    print("Trying to receive data from network")
    try:
      data, addr = self.recvfrom(1024)
      message = data.decode('Utf8').splitlines()
      current_head_position_in_tracker_space =
Vector((float(message[0].replace(',', '.')),
                               float(message[1].replace(',', '.')),
                               float(message[2].replace(',', '.'))))

      # Filtering incoming data
      if gD:
        previous_head_position_in_tracker_space =
gD['head_position_in_tracker_space']
```

```
        if (current_head_position_in_tracker_space -
previous_head_position_in_tracker_space).magnitude > 0.3:
          print('Data received but invalid')
        else:
          gD["head_position_in_tracker_space"] =
current_head_position_in_tracker_space
          print('Data received and valid')
      else:
        gD['head_position_in_tracker_space'] =
current_head_position_in_tracker_space
        print('Data received and valid')

    except socket.error:
      print('No data received')


########################################################################

# Initialization function that runs only once

def init(own):

  # create a socket using IPV4 address family and UDP protocol
  try:
    global skt
    skt = Socket(socket.AF_INET, socket.SOCK_DGRAM)

  except:
    raise Exception('!!! Unable to create socket, networking
initialization failed')
    pass

  own['networking_init'] = True


########################################################################

# Module execution entry point

def main(cont):

  own = cont.owner

  # Initialize networking if not already done
  if 'networking_init' not in own:
    init(own)

  # exit Game Engine if ESC key is pressed
  elif (logic.keyboard.events[bge.events.ESCKEY] ==
logic.KX_INPUT_JUST_ACTIVATED or
      logic.keyboard.events[bge.events.ESCKEY] == logic.KX_INPUT_ACTIVE
or
      logic.keyboard.events[bge.events.ESCKEY] ==
bge.logic.KX_INPUT_JUST_RELEASED):
```

```
  # close the socket before exiting
  if skt:
    skt.close()

  # end the Game Engine
  logic.endGame()

else:
  skt.receive()
```

APPENDIX A-3

This section presents the code that creates custom classes used by other scripts.

myTypes.py

```
########################################################################


# This script defines custom classes for the airport scenario

import math
import mathutils
import bgeutils
import manager
import mouse_utils
import py_utils
import draw_utils
from bge import logic, types, events

########################################################################

# This is a custom class for non-static objects

class Mobile(types.KX_GameObject):

  def __init__(self, game_obj):
    self.ground_position = mathutils.Vector((self.worldPosition[0], 
self.worldPosition[1], 0))
    self.waypoints = []
    self.taxi_route = []
    self.states = {'taxi': False}

  def taxi(self):
    self.ground_position = mathutils.Vector((self.worldPosition[0], 
self.worldPosition[1], 0))
    delta = self.taxi_route[0] - self.ground_position
    self.setLinearVelocity([self['linear_velocity'], 0, 0], True)
    self.alignAxisToVect(delta, 0, self['angular_velocity'])

  def has_reached(self, point):
    if math.fabs((self.ground_position - point).magnitude) < 3:
      return True
    else:
      return False


# This is a custom class for Aircrafts objects. Derives from Mobile
```

```
class Aircraft(Mobile):

  def __init__(self, game_obj):
    super(Aircraft, self).__init__(game_obj)


class Scene(types.KX_Scene):

  def __init__(self, scene_obj):
    self.taxi_objects = []
    self.selected_object = None

  def run(self):

    Interface.get_user_input(self)

    for obj in self.objects:
      if type(obj) is Aircraft:

        if obj.waypoints:
          nearest_node_key =
bgeutils.get_nearest_node(obj.worldPosition, manager.nav_mesh.nav_dict)
          nearest_node = manager.nav_mesh.nav_dict[nearest_node_key]
          obj.waypoints = [nearest_node.location] + obj.waypoints
          obj.taxi_route = RouteFinder.search_taxi_route(obj)
          py_utils.clear_list(obj.waypoints)

        if obj.states['taxi']:
          if obj.taxi_route:
            obj.taxi()
            if obj.has_reached(obj.taxi_route[0]):
              obj.taxi_route.pop(0)


# findes the taxi route for the selected object

class RouteFinder():

  @staticmethod
  def search_taxi_route(obj):
    taxi_route = []
    for i in range(len(obj.waypoints)-1):
      start_waypoint = obj.waypoints[i]
      end_waypoint = obj.waypoints[i+1]
      start_node = bgeutils.get_nearest_node(start_waypoint,
manager.nav_mesh.nav_dict)
      end_node = bgeutils.get_nearest_node(end_waypoint,
manager.nav_mesh.nav_dict)
      taxi_route.extend(bgeutils.a_star(manager.nav_mesh.nav_dict,
start_node, end_node, obj.waypoints[-1]))
      for key in manager.nav_mesh.nav_dict:
        manager.nav_mesh.nav_dict[key].reset()
    taxi_route.append(obj.waypoints[-1])
    taxi_route = [obj.waypoints[0]] + taxi_route
```

```
    if manager.debug_mode:
      draw_utils.draw_taxi_route(taxi_route, time=1000, color=(0, 0, 1,
1), offset=(0,0,1))

    return taxi_route


class NavMesh(types.KX_GameObject):

  def __init__(self, game_obj):
    self.nav_dict = bgeutils.get_nav_mesh()
    if manager.debug_mode:
      draw_utils.display_nav_mesh(self)


# handles the interaction between the pseudo-pilot and the scene

class Interface:

  clicked_points = []

  @staticmethod
  def get_user_input(scene):

    hit_object = mouse_utils.mouse_sensor.hitObject
    if hit_object:

      # Handle left click
      if mouse_utils.left_click():
        if "mobile" in hit_object and hit_object is not
scene.selected_object:
          if scene.selected_object:
            py_utils.clear_list(scene.selected_object.waypoints)
py_utils.clear_list(scene.selected_object.taxi_route)
          scene.selected_object = hit_object
        else:
          if scene.selected_object:
            hit_position = mouse_utils.mouse_sensor.hitPosition
bgeutils.move_target_to(scene, hit_position)
            Interface.clicked_points.append(hit_position)

      # Handle right click
      elif mouse_utils.right_click():

              if scene.selected_object:
          py_utils.clear_list(scene.selected_object.waypoints)
          scene.selected_object = None

        if 'mobile' in hit_object: # if it is a mobile
          hit_object.states.update({'taxi': False})
          py_utils.clear_list(hit_object.waypoints)
          py_utils.clear_list(hit_object.taxi_route)
          if hit_object in scene.taxi_objects:
            scene.taxi_objects.remove(hit_object)
```

```
    # Handle the ENTERKEY press event
    if logic.keyboard.events[events.ENTERKEY] ==
logic.KX_INPUT_JUST_ACTIVATED:
      print("ENTERKEY")
      if scene.selected_object:
        scene.selected_object.waypoints.extend(Interface.clicked_points)
        Interface.clicked_points = []
        if scene.selected_object.waypoints:
          scene.selected_object.states.update({'taxi': True})
        scene.selected_object = None
```

APPENDIX A-4

This section presents the code that is used to initiate mobile objects and the navigation meshes in the airport scene. After the initialization is complete it runs the scene every logic tic.

manager.py

```
#######################################################################
# This script initializes the navigation mesh and manages mobile objects
in the scene

import my_types
from bge import logic

#######################################################################

debug_mode = None # this is set through a property of the script owner
nav_mesh = None

#######################################################################


class Manager:

  @ staticmethod
  def manage():
    logic.getCurrentScene().run()

#######################################################################

# Module execution entry point


def main(cont):

  own = cont.owner

  if 'init' not in own:
    try:
      init(own)
      own['init'] = True

    except UserWarning:
      print("Error: Manager initialization failed")
  else:
    Manager.manage()

#######################################################################
```

```
def init(own):

  # try to set the debug flag
  try:
    global debug_mode
    debug_mode = own['debug']
  except KeyError:
    print('KeyError: script owner has no debug property')

  # re-instancing the current Scene
  my_types.Scene(logic.getCurrentScene())

  # re-instancing objects in the scene
  for obj in logic.getCurrentScene().objects:

    if "type" in obj:
      if obj["type"] == "aircraft":
        my_types.Aircraft(obj) # re-instancing all aircrafts objects

    elif "navmesh" in obj:
      global nav_mesh
      nav_mesh = my_types.NavMesh(obj)
```

APPENDIX A-5

This section presents the code that is used to track the viewer's head with the Miscrosoft™ Kinect™ and send the head coordinates data over the network.

MainWindow.xaml.cs

```
using System;
using System.Collections.Generic;
using System.Linq;
using System.Text;
using System.Threading.Tasks;
using System.Windows;
using System.Windows.Controls;
using System.Windows.Data;
using System.Windows.Documents;
using System.Windows.Input;
using System.Windows.Media;
using System.Windows.Media.Imaging;
using System.Windows.Navigation;
using System.Windows.Shapes;
using Microsoft.Kinect;
using System.Net;
using System.Net.Sockets;

namespace Head_tracking_WPF_application
{
  public partial class MainWindow : Window
  {
    KinectSensor sensor;
    BodyFrameReader bfr;
    Body[] bodies = new Body[6];

    Socket socket = new Socket(AddressFamily.InterNetwork,
SocketType.Dgram, ProtocolType.Udp);
    IPAddress IP = IPAddress.Loopback;
    IPEndPoint ep;

    public MainWindow()
    {
      InitializeComponent();

      ep = new IPEndPoint(IP, 9999);

      this.Loaded += MainWindow_Loaded;
    }

    void MainWindow_Loaded(object sender, RoutedEventArgs args)
    {
      text_block_1.Text = "HEAD COORDINATES";
      text_block_2.Text = String.Format("Sending data to IP {0}", IP);
      sensor = KinectSensor.GetDefault();
      sensor.Open();
      bfr = sensor.BodyFrameSource.OpenReader();
      bfr.FrameArrived += bfr_FrameArrived;
    }
```

```
    void bfr_FrameArrived(object sender, BodyFrameArrivedEventArgs args)
    {
      using(BodyFrame bf = args.FrameReference.AcquireFrame())
      {
        if (bf != null)
        {
          bf.GetAndRefreshBodyData(bodies);
          foreach (Body body in bodies)
          {
            if (body.IsTracked)
            {
              Joint headJoint = body.Joints[JointType.Head];
              if (headJoint.TrackingState == TrackingState.Tracked)
              {
                Single X = headJoint.Position.X;
                Single Y = headJoint.Position.Y;
                Single Z = headJoint.Position.Z;

                string x = X.ToString("0.00000000");
                string y = Y.ToString("0.00000000");
                string z = Z.ToString("0.00000000");

                text_block_1.Text = String.Format("HEAD COORDINATES: \n
X: {0} \n Y: {1} \n Z: {2} \n", x, y, z);

                string message = x + "\n" + y + "\n" + z;
                byte[] stream = Encoding.UTF8.GetBytes(message);
                socket.SendTo(stream, ep);
              }
            }
          }
        }
      }
    }
  }
}
```

APPENDIX B

This section presents a detailed analysis of the selected control tasks. The analysis is based on the feedback obtained by experienced ATCO and was used as a basis to classify controllers' working behaviour in chapters 6.4.3 and 6.4.5.

To better assess the level of significant S-R-K indicators, such as automation, executive control, decision making and problem solving, each task was subdivided into multiple subtasks.

Similar to the one reported in chapter 6.4.3 and 6.4.5., this analysis is organized in four scenarios.

- **Scenario 0**: use of standard tools together with local rules[67]. This is represented by the grey letters (first line of each subtask)
- **Scenario 1**: use of AR tools together with local rules. This is represented by the green letters (second line of each subtask)
- **Scenario 2:** use of standard tools with limited restrictions (NVC rules). This is represented by the red letters (third line of each subtask)
- **Scenario 3**: use of AR tools with limited restrictions (NVC rules). This is represented by the orange letters (fourth line of each subtask)

Also, four different visibility conditions were considered: **NVC, LVC 1, 2** and **3**

---

[67] As described in section 5.2.6.

Legend	NMC				LVC 1				LVC 2				LVC 3			
H = High M = Medium L = Low N = Not present	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making

TASK GND 1 – ISSUE ATC CLEARANCE:																
1. Identify aircraft and activate electronic strip on FDP	H	M	N	N	H	M	N	N	M	M	N	N	M	M	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	M	N	N	H	M	N	N	H	M	N	N	M	M	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
2. Check if SID is congruent to RWY in use	M	L	L	L	M	L	L	L	M	L	L	L	M	L	L	L
	M	L	L	L	M	L	L	L	M	L	L	L	M	L	L	L
	M	L	L	L	M	L	L	L	M	L	L	L	M	L	L	L
	M	L	L	L	M	L	L	L	M	L	L	L	M	L	L	L
3. Assign initial flight level: (local procedure request to assign 5000ft to every flight)	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
4. Transmit ATC clearance and hear-back confirmation of correct receipt	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N

Legend	NMC				LVC 1				LVC 2				LVC 3			
H = High M = Medium L = Low N = Not present	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making

TASK GND 2 – ISSUE START UP CLEARANCE:																
1. Check EOBT and CTOT (if any)	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
2. Check traffic condition	M	L	N	N	M	L	N	N	M	M	N	N	M	M	N	N
	M	L	N	N	M	L	N	N	M	L	N	N	M	L	N	N
	M	L	N	N	M	L	N	N	M	M	N	N	M	M	N	N
	M	L	N	N	M	L	N	N	M	L	N	N	M	L	N	N
3. If necessary, ask for start-up approval to approach unit	M	L	L	L	M	L	L	L	M	L	L	L	M	L	L	L
	M	L	L	L	M	L	L	L	M	L	L	L	M	L	L	L
	M	L	L	L	M	L	L	L	M	L	L	L	M	L	L	L
	M	L	L	L	M	L	L	L	M	L	L	L	M	L	L	L
4. Estimate any delay	L	M	M	L	L	M	M	L	L	M	M	L	L	M	M	L
	L	M	M	L	L	M	M	L	L	M	M	L	L	M	M	L
	L	M	M	L	L	M	M	L	L	M	M	L	L	M	M	L
	L	M	M	L	L	M	M	L	L	M	M	L	L	M	M	L
5. Transmit clearance	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N

Legend	NMC				LVC 1				LVC 2				LVC 3			
H = High M = Medium L = Low N = Not present	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making

TASK GND 3 – APPROVE PUSH BACK																
1. Identify aircraft on apron	M	M	N	N	M	M	N	N	M	M	N	N	L	H	N	N
	M	M	N	N	M	L	N	N	M	L	N	N	M	M	N	N
	M	M	N	N	M	M	N	N	M	M	N	N	L	H	N	N
	M	M	N	N	M	L	N	N	M	L	N	N	M	M	N	N
2. Check if there are push back conflicts between stand and apply local regulation	M	M	L	L	M	M	L	L	M	M	L	L	L	M	L	L
	M	M	L	L	M	M	L	L	M	M	L	L	M	M	L	L
	M	M	L	L	M	M	L	L	M	M	L	L	L	H	H	H
	M	M	L	L	M	M	L	L	M	M	L	L	L	M	M	M
3. Estimate any delay	L	M	M	L	L	M	M	L	L	M	M	L	L	M	M	L
	L	M	M	L	L	M	M	L	L	M	M	L	L	M	M	L
	L	M	M	L	L	M	M	L	L	M	M	L	L	M	M	L
	L	M	M	L	L	M	M	L	L	M	M	L	L	M	M	L
4. Transmit pushback clearance	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N

Legend	NMC				LVC 1				LVC 2				LVC 3			
H = High M = Medium L = Low N = Not present	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making

TASK GND 4 – ISSUE TAXI CLEARANCE																
1. Identify aircraft on apron	M	M	N	N	M	M	N	N	M	M	N	N	L	H	N	N
	M	M	N	N	M	L	N	N	M	L	N	N	M	M	N	N
	M	M	N	N	M	M	N	N	M	M	N	N	L	H	N	N
	M	M	N	N	M	L	N	N	M	L	N	N	M	M	N	N
2. Identify taxiway closed or not allowed, choose correct holding point	M	M	N	N	N	M	N	N	M	M	N	N	M	M	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	M	M	N	N	M	M	N	N	M	M	N	N	M	M	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
3. Assess potential conflicts between moving aircrafts/vehicles	L	M	M	M	L	M	M	M	L	H	H	H	L	H	H	H
	L	M	L	L	L	M	L	L	L	M	L	L	L	M	L	L
	L	M	M	M	L	M	M	M	L	H	H	H	L	H	H	H
	L	M	M	L	L	M	M	L	L	M	M	L	L	M	M	L
4. Choose best path	M	M	M	M	M	M	M	M	H	M	L	L	H	M	L	L
	M	M	M	M	M	M	M	M	H	M	L	L	H	M	L	L
	M	M	M	M	M	M	M	M	L	H	M	M	L	H	M	M
	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
5. Transmit taxi clearance	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N

Legend	NMC				LVC 1				LVC 2				LVC 3			
H = High M = Medium L = Low N = Not present	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making

TASK GND 5 – MONITOR TAXI ROUTE																
1. Identify aircraft on maneuvering area	M	M	N	N	M	M	N	N	L	H	N	N	L	H	N	N
	M	M	N	N	M	L	N	N	M	M	N	N	M	M	N	N
	M	M	N	N	M	M	N	N	L	H	N	N	L	H	N	N
	M	M	N	N	M	L	N	N	M	M	N	N	M	M	N	N
2. Monitor if aircraft is following the assigned taxi route	M	M	N	N	M	M	N	N	M	H	N	N	M	H	N	N
	M	L	N	N	M	L	N	N	M	M	N	N	M	M	N	N
	M	M	N	N	M	M	N	N	L	H	N	N	L	H	N	N
	M	L	N	N	M	L	N	N	M	H	N	N	M	H	N	N
3. Assess potential conflicts between moving aircrafts/vehicles	L	M	M	M	L	M	M	M	L	H	H	H	L	H	H	H
	L	M	L	L	L	M	L	L	L	M	L	L	L	M	L	L
	L	M	M	M	L	M	M	M	L	H	H	H	L	H	H	H
	L	M	M	L	L	M	M	L	L	M	M	L	L	M	M	L
4. Be sure that aircraft stops at assigned holding point	M	M	N	N	M	M	N	N	M	H	N	N	M	H	N	N
	M	L	N	N	M	L	N	N	M	M	N	N	M	M	N	N
	M	M	N	N	M	M	N	N	L	H	N	N	L	H	N	N
	M	L	N	N	M	L	N	N	M	H	N	N	M	H	N	N
5. If needed, choose appropriate corrective action	L	H	M	M	L	H	M	M	L	H	M	M	L	M	M	M
	L	H	M	M	L	H	M	M	L	H	M	M	L	M	M	M
	L	H	M	M	L	H	H	M	L	H	H	H	L	H	H	H
	L	H	M	M	L	H	H	M	L	H	H	H	L	H	H	H

Legend	NMC				LVC 1				LVC 2				LVC 3			
H = High M = Medium L = Low N = Not present	Automation	Executive control	Decision-making	Problem solving	Automation	Executive control	Decision-making	Problem solving	Automation	Executive control	Decision-making	Problem solving	Automation	Executive control	Decision-making	Problem solving

TASK TWR 1 – ISSUE LANDING CLEARANCE																
1. Identify aircraft position	M	M	N	N	M	M	N	N	L	H	N	N	L	H	N	N
	M	M	N	N	M	L	N	N	M	M	N	N	M	M	N	N
	M	M	N	N	M	M	N	N	L	H	N	N	L	H	N	N
	M	M	N	N	M	L	N	N	M	M	N	N	M	M	N	N
2. Check runway status (free/occupied)	H	M	N	N	H	M	N	N	H	H	N	N	H	H	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	M	N	N	H	M	N	N	H	M	N	N	H	M	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
3. Check wind, runway surface status, NAVAIDS, visibility and ceiling	H	M	N	N	H	M	N	N	H	M	N	N	H	M	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	M	N	N	H	M	N	N	H	M	N	N	H	M	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
4. Check traffic condition and choose the right RWY exit	M	M	M	M	M	M	M	M	H	M	L	L	H	M	L	L
	M	M	M	M	M	M	M	M	H	M	L	L	H	M	L	L
	M	M	M	M	M	M	M	M	L	H	M	M	L	H	M	M
	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
5. Transmit landing clearance and required information	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N

Legend	NMC				LVC 1				LVC 2				LVC 3			
H = High M = Medium L = Low N = Not present	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making

TASK TWR 2 – ISSUE TAKE OFF CLEARANCE																
1. Identify aircraft position	M	M	N	N	M	M	N	N	L	H	N	N	L	H	N	N
	M	M	N	N	M	L	N	N	M	M	N	N	M	M	N	N
	M	M	N	N	M	M	N	N	L	H	N	N	L	H	N	N
	M	M	N	N	M	L	N	N	M	M	N	N	M	M	N	N
2. Check runway status (free/occupied)	H	M	N	N	H	M	N	N	H	H	N	N	H	H	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	M	N	N	H	M	N	N	H	M	N	N	H	M	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
3. Check wind, runway surface status, NAVAIDS, visibility and ceiling	H	M	N	N	H	M	N	N	H	M	N	N	H	M	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	M	N	N	H	M	N	N	H	M	N	N	H	M	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
4. Check runway availability and if departure route is free of traffic	M	M	M	M	M	M	M	M	H	M	L	L	H	M	L	L
	M	M	M	M	M	M	M	M	H	M	L	L	H	M	L	L
	M	M	M	M	M	M	M	M	L	H	M	M	L	H	M	M
	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
5. Transmit take-off clearance and required information	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N

Legend	NMC				LVC 1				LVC 2				LVC 3			
H = High M = Medium L = Low N = Not present	Automation	Executive control	Decision-making	Problem solving	Automation	Executive control	Decision-making	Problem solving	Automation	Executive control	Decision-making	Problem solving	Automation	Executive control	Decision-making	Problem solving

TASK TWR 3 – MONITOR TAKE OFF AND LANDING OPERATION																
1. Identify aircraft position	M	M	N	N	M	M	N	N	L	H	N	N	L	H	N	N
	M	M	N	N	M	L	N	N	M	M	N	N	M	M	N	N
	M	M	N	N	M	M	N	N	L	H	N	N	L	H	N	N
	M	M	N	N	M	L	N	N	M	M	N	N	M	M	N	N
2. Check if landing or take off occurred	M	M	N	N	M	M	N	N	L	H	N	N	L	H	N	N
	M	M	N	N	M	L	N	N	M	M	N	N	M	M	N	N
	M	M	N	N	M	M	N	N	L	H	N	N	L	H	N	N
	M	M	N	N	M	L	N	N	M	M	N	N	M	M	N	N
3. Identify unexpected events (aborted take-off, go around, RWY incursion)	L	M	N	N	L	M	N	N	L	H	N	N	L	H	N	N
	L	M	N	N	L	M	N	N	L	M	N	N	L	M	N	N
	L	M	N	N	L	M	N	N	L	H	N	N	L	H	N	N
	L	M	N	N	L	M	N	N	L	M	N	N	L	M	N	N
4. Choose appropriate action to be taken, if any	L	H	M	M	L	H	M	M	L	H	M	M	L	M	M	M
	L	H	M	M	L	H	M	M	L	H	M	M	L	M	M	M
	L	H	M	M	L	H	H	M	L	H	H	H	L	H	H	H
	L	H	M	M	L	H	H	M	L	H	H	H	L	H	H	H
5. If necessary, transmit new instruction to avoid conflict	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N

Legend	NMC				LVC 1				LVC 2				LVC 3			
H = High M = Medium L = Low N = Not present	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making	Automation	Executive control	Problem solving	Decision-making

TASK TWR 4 – ISSUE CLEARANCE TO VEHICLE FOR RUNWAY INSPECTION/OPERATIONS																
1. Identify vehicle position	M	M	N	N	M	M	N	N	M	M	N	N	L	H	N	N
	M	M	N	N	M	L	N	N	M	L	N	N	M	M	N	N
	M	M	N	N	M	M	N	N	M	M	N	N	L	H	N	N
	M	M	N	N	M	L	N	N	M	L	N	N	M	M	N	N
2. Check runway status free/occupied	H	M	N	N	H	M	N	N	H	H	N	N	H	H	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	M	N	N	H	M	N	N	H	M	N	N	H	M	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
3. Check traffic condition	M	L	N	N	M	L	N	N	M	M	N	N	M	M	N	N
	M	L	N	N	M	L	N	N	M	L	N	N	M	L	N	N
	M	L	N	N	M	L	N	N	M	M	N	N	M	M	N	N
	M	L	N	N	M	L	N	N	M	L	N	N	M	L	N	N
4. Check if any restriction must be applied	M	M	N	N	N	M	N	N	M	M	N	N	M	M	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	M	M	N	N	M	M	N	N	M	M	N	N	M	M	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
5. Transmit clearance and required information	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N
	H	L	N	N	H	L	N	N	H	L	N	N	H	L	N	N

REFERENCES

[1] Ella Pinska, “An Investigation of the Head-up Time at Tower and Ground Control Positions,” in *Proceeding of the 5th Eurocontrol Innovative Research Workshop*, 2006, pp. 81–86.

[2] SEAR JU, “SESAR 2020 Exploratory Research: First Call for Research Projects.” 2015.

[3] SESAR JU, “European ATM Master Plan.” 2015.

[4] *Challenges of Air Transport 2030 Survey of experts’ views*. Brétigny-sur-Orge , France: EUROCONTROL, 2009.

[5] C. Neveu, T. Blackmon, and L. Stark, “Evaluation of the Effects of a Head-mounted Display on Ocular Accommodation,” *Presence*, vol. 7, no. 3, pp. 278–289, Jun. 1998.

[6] R. Reisman and D. Brown, “Design of Augmented Reality Tools for Air Traffic Control Towers,” in *Proceedings of 6th AIAA Aviation Technology, Integration and Operations Conference (ATIO)*, Wichita, KS, 2006.

[7] Monica Tavanti, “Control Tower Operations: A Literature Review of Task Analysis Studies,” EEC Technical/Scientific Report No. 2006-013, EEC Technical/Scientific Report No. 2006-013, Jul. 2006.

[8] M. Schmidt, M. Rudolph, B. Werther, C. Möhlenbrink, and N. Fürstenau, “Development of an Augmented Vision Video Panorama Human-Machine Interface for Remote Airport Tower Operation.,” in *HCI (9)*, 2007, vol. 4558, pp. 1119–1128.

[9] S. R. Ellis, “Towards Determination of Visual Requirements for Augmented Reality Displays and Virtual Environments for the Airport Tower,” in *Virtual Media for Military Applications RTO-MP-HFM-136*, Neuilly-sur-Seine, France, 2006, pp. 31–1-- 31–10.

[10] Shackelford, R and Karpe, P., “Low/Zero Visibility Tower Tools (L/ZVTT) Benefits Assessment,” NASA Ames Research Center Technical Reports, Moffett Field, CA 94035, Sep. 1998.

[11] Reisman, R.J., Feiner, S.K., and Brown, D.M., “Augmented Reality Tower Technology Flight Test,” in *International Conference on Human-Computer Interaction in Aerospace (HCI-Aero)*, Santa Clara, CA, 2014.

[12] R. J. Reisman and S. R. Ellis, “Air Traffic Control Tower Augmented Reality Field Study,” in *ACM SIGGRAPH Posters*, New York, NY, USA, 2005.

[13] E. Pinska and C. Tijus, “Augmented Reality Technology for Control Tower - Analysis of Applicability Based on the Field Study,” presented at the 1st CEAS European Air and Space Conference, 2007.

[14] Ronald J. Reisman and David M. Brown, “Augmented Reality Tower Technology Assessment,” NASA Ames Research Center Technical Reports, NASA/TM–2010-216011, Oct. 2010.

[15] S. Peterson, “Very large format stereoscopic head-up display for the airport tower,” in *In Proceedings of the Virtual Images Seminar, number 16. CNRS/Renault*, 2007.

[16] EUROCONTROL, “All-Causes Delay and Cancellations to Air Transport in Europe,” 2014. [Online]. Available: https://www.eurocontrol.int/publications/coda-digest-annual-2014.

[17] SESAR JU, "System Wide Information Management (SWIM)." [Online]. Available: http://www.sesarju.eu/sesar-solutions/swim.

[18] SESAR JU, "SWIM Concept of Operations." .

[19] Stephen R. Ellis, Bernard D. Adelstein, Ronald J. Reisman, Joelle R. Schmidt-Ott, Jonathan Gips, and Jimm Krozel, "Augmented Reality in a Simulated Tower Environment: Effect of Field of View on Aircraft Detection," NASA Ames Research Center Technical Reports, NASA/TM-2002-211853, NAS 1.15:211853, IH-025, Sep. 2002.

[20] N. Fürstenau, M. Schmidt, M. Rudolph, C. Möhlenbrink, A. Papenfuß, and S. Kaltenhäuser, "Steps Towards the Virtual Tower: Remote Airport Traffic Control Center (RAiCe)," ENRI International Workshop on ATM/CNS, Tokyo, Japan, 2009.

[21] D. Schulz-Rückert, "Future of Aerodrome Traffic Control," presented at the Aachen Aviation Convention, Aachen, Germany, 2007.

[22] M. A. Brown and M. Slater, "Some experiences with three-dimensional display design: an air traffic control visualisation," in *Proceedings of the 6th IEEE International Workshop on Robot and Human Communication. RO-MAN '97.*, Sendai, Japan, 1997, pp. 296–301.

[23] A. Esgate, K. Baker, and D. Groome, *An introduction to applied cognitive psychology*. Hove, UK ; New York, NY: Psychology Press, 2005.

[24] M. Tavanti, *On the Relative Utility of 3D Interfaces*. Uppsala, Sweden: Acta Universitatis Upsaliensis, 2004.

[25] R. H. R. Harper and J. A. Hughes, "'What a f-ing system! Send 'em all to the same place and then expect us to stop 'em hitting': Making Technology Work in Air Traffic Control," in *Technology in Working Order: Studies of Work, Interaction, and Technology*, G. Button, Ed. London: Routledge, 1993, pp. 127–144.

[26] L. Bainbridge, "Analysis of Verbal Protocols from a Process Control Task," in *The Human Operator in Process Control*, E. Edwards and F. P. Lees, Eds. London: Taylor & Francis, 1974, pp. 146–158.

[27] J. Rasmussen and A. Jensen, "Mental Procedures in Real-Life Tasks: A Case Study of Electronic Trouble Shooting," *Ergonomics*, vol. 17, no. 3, pp. 293–307, 1974.

[28] E. Jeannot, C. Kelly, and D. Thompson, "The Development of Situation Awareness Measures in ATM Systems," EUROCONTROL, Brussels, HRS/HSP-005-REP-01, 2003.

[29] A. Bisseret, "Analysis of Mental Processes Involved in Air Traffic Control," *Ergonomics*, vol. 14, no. 5, pp. 565–570, 1971.

[30] D. Whitfield and A. Jackson, "The Air Traffic Controller's 'Picture' as an example of a Mental Model," in *Analysis, Design and Evaluation of Man–Machine Systems: Proceedings of the IFAC/IFIP/IFORS/IEA Conference*, Baden-Baden, Federal Republic of Germany, 1983.

[31] A. Bisseret, *Représentation et décision experte: Psychologie cognitive de la décision chez les aiguilleurs du ciel*. Toulouse: Octares Editions, 1995.

[32] D. Ochanine, "Rôle de l´image opérative dans la saisie du contenu informationnel des signaux. Questions de psychologie," *Quest. Psychol.*, vol. 4, pp. 209–224, 1969.

[33] D. Ochanine, *L'image Opérative: Actes d'un Seminaire (1-5 juin 1981) et Recueil d'Articles de D. Ochanine*. Paris: Universite de Paris I (Pantheon-Sorbonne), Centre d'Education Permanente, Department d'Ergonomie et d'Ecologie Humaine, 1981.

[34] S. D. Gronlund, O. D. Daryl, M. R. P. Dougherty, J. L. Perry, and C. A. Manning, "Aircraft Importance and its Relevance to Situation Awareness Federal Aviation

Administration," FAA Civil Aeromedical Institute, Oklahoma City, OK, Final Report DOT/FAA/AM-98/16, May 1998.

[35] M. Tavanti and M. Cooper, "Looking for the 3D Picture: The Spatio-temporal Realm of Student Controllers," in *First International Conference, HCD 2009, Held as Part of HCI International 2009, Proceedings*, San Diego,CA, 2009, vol. 5619, pp. 1070–1079.

[36] Geoffrey G. Strutt, "The Human Computer Interface in Air Traffic Control: Feasibilty Study for a Three Dimensional Radar Picture," Master's Thesis, Middlesex Polytechnic, London, 1991.

[37] M. S. Burnett, "An Analysis of the Effects of Multiple Colors, Information Density, and Traffic Complexity on Plan View and Perspective Air Traffic Control Radar Displays," Master's Thesis, University of Washington, Deptartment of Industrial Engineering, London, 1991.

[38] Peterson, Stephen, and Ella Pinska, "Human Performance with simulated Collimation in Transparent Projection Screens," in *Proceedings of the Second International Conference on Research in Air Transportation (ICRAT)*, 2006.

[39] N. Masotti, F. De Crescenzio, and S. Bagassi, "Augmented Reality in the Control Tower: A Rendering Pipeline for Multiple Head-Tracked Head-up Displays," in *Augmented Reality, Virtual Reality, and Computer Graphics: Third International Conference, AVR 2016, Lecce, Italy, June 15-18, 2016. Proceedings, Part I*, T. L. De Paolis and A. Mongelli, Eds. Cham: Springer International Publishing, 2016, pp. 321–338.

[40] F. Solari, M. Chessa, M. Garibotti, and S. P. Sabatini, "Natural perception in dynamic stereoscopic augmented reality environments," *Displays*, vol. 34, no. 2, pp. 142–152, 2013.

[41] A. M. Puerta, "The power of shadows: shadow stereopsis," *J. Opt. Soc. Am. A*, vol. 6, no. 2, pp. 309–311, Feb. 1989.

[42] S. Nagata, "Visual depth sensitivities of various cues for depth perception.," *NHK Lab. Note*, vol. 266, pp. 1–11, 1981.

[43] J. E. Cutting and P. M. Vishton, "Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth," *W Epstein Rogers Eds Handb. Percept. Cogn.*, vol. Vol 5; Perception of space and motion, pp. 69–117, Academic press, San Diego (CA 1995.

[44] S. Palmisano, B. Gillam, D. G. Govan, R. S. Allison, and J. M. Harris, "Stereoscopic perception of real depths at large distances," *J. Vis.*, vol. 10(6):19, pp. 1–16, 2010.

[45] H. J. Howard, "A Test for the Judgment of Distance," *Trans. Am. Ophthalmol. Soc.*, vol. 17, pp. 195–235, 1919.

[46] H. Helmholtz, *Treatise on Physiological Optics(1909/1925)*, vol. 3: Visual Perception. James P.C. Southall, Optical Society of America, 1925.

[47] G. Kramida and A. Varshney, "Resolving the Vergence-Accommodation Conflict in Head Mounted Displays," *IEEE Trans. Vis. Comput. Graph. Issue 07*, vol. 22, pp. 1912–31, Jul. 2016.

[48] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence-accommodation conflicts hinder visual performance and cause visual fatigue," *J. Vis.*, vol. 8, no. 3, p. 33.1–30, Jan. 2008.

[49] M. Lambooij, W. Ijsselsteijn, M. Fortuin, and I. Heynderickx, "Visual Discomfort and Visual Fatigue of Stereoscopic Displays: A Review," *J. Imaging Sci. Technol.*, vol. 53, no. 3, p. 30201, 2009.

[50] S. Reichelt, R. Häussler, G. Fütterer, and N. Leister, “Depth cues in human visual perception and their realization in 3D displays,” in *SPIE Defense, Security, and Sensing,* 2010, p. 76 900B–76 900B–12.

[51] T. Bando, A. Iijima, and S. Yano, “Visual fatigue caused by stereoscopic images and the search for the requirement to prevent them: A review,” *Displays*, vol. 33, no. 2, pp. 76–83, 2012.

[52] A. Shamir and O. Sorkine, “Visual Media Retargeting,” in *ACM SIGGRAPH ASIA 2009 Courses*, New York, NY, USA, 2009, p. 11:1–11:13.

[53] M. Lang, A. Hornung, O. Wang, S. Poulakos, A. Smolic, and M. Gross, “Nonlinear Disparity Mapping for Stereoscopic 3D,” in *ACM SIGGRAPH 2010 Papers*, New York, NY, USA, 2010, p. 75:1–75:10.

[54] Y. Liu, L. Sun, and S. Yang, “A retargeting method for stereoscopic 3D video,” *Comput. Vis. Media*, vol. 1, no. 2, pp. 119–127, 2015.

[55] P. Didyk, T. Ritschel, E. Eisemann, K. Myszkowski, H.-P. Seidel, and W. Matusik, “A Luminance-contrast-aware Disparity Model and Applications,” *ACM Trans Graph*, vol. 31, no. 6, p. 184:1–184:10, Nov. 2012.

[56] A. Sherstyuk, A. Dey, C. Sandor, and A. State, “Dynamic Eye Convergence for Head-mounted Displays Improves User Performance in Virtual Environments,” in *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, New York, NY, USA, 2012, pp. 23–30.

[57] T. Oskam, A. Hornung, H. Bowles, K. Mitchell, and M. Gross, “OSCAM - Optimized stereoscopic camera control for interactive 3D,” *ACM Trans Graph. Proc SIGGRAPH*, vol. 30, no. 6, p. 189:1–189:8, 2011.

[58] U. Celikcan, G. Cimen, E. B. Kevinc, and T. Capin, “Attention-Aware Disparity Control in interactive environments,” *Vis. Comput.*, vol. 29, no. 6, pp. 685–694, 2013.

[59] J. P. Rolland, M. W. Krueger, and A. A. Goon, “Dynamic focusing in head-mounted displays,” 1999, vol. 3639, pp. 463–470.

[60] S. Reichelt, R. Häussler, G. Fütterer, and N. Leister, “Depth cues in human visual perception and their realization in 3D displays,” in *SPIE Defense, Security, and Sensing,* 2010, p. 76 900B–76 900B–12.

[61] K. J. Ciuffreda, B. Wang, and D. Wong, “Central and near peripheral retinal contributions to the depth-of-focus using naturalistic stimulation,” *Vision Res.*, vol. 45, no. 20, pp. 2650–2658, Sep. 2005.

[62] R. Suryakumar, J. P. Meyers, E. L. Irving, and W. R. Bobier, “Vergence accommodation and monocular closed loop blur accommodation have similar dynamic characteristics,” *Vision Res.*, vol. 47, no. 3, pp. 327–337, Feb. 2007.

[63] J. E. Melzer, “Head-Mounted Displays,” Kaiser Electro-Optics Inc., 2001.

[64] D. F. Mc Allister, “Display Technology: Stereo & 3D Display Technologies.” Mar-2003.

[65] G. Kramida and A. Varshney, “Resolving the Vergence-Accommodation Conflict in Head Mounted Displays,” *IEEE Trans. Vis. Comput. Graph. Issue 07*, vol. 22, pp. 1912–31, Jul. 2016.

[66] R. J. Sojourner and J. F. Antin, “The Effects of a Simulated Head-Up Display Speedometer on Perceptual Task Performance,” *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 32, no. 3, pp. 329–339, 1990.

[67] H. Kato, H. Ito, J. Shima, M. Imaizumi, and H. Shibata, “Development of a hologram head up display,” 1992.

[68] J. P. Rolland, Y. Baillot, and A. A. Goon, “A Survey of Tracking Technology for Virtual Environments,” *Fundam. Wearable Comput. Augment. Real.*, p. 67–112 (chap. 3), 2001.

[69] F. Zhou, H. B.-L. Duh, and M. Billinghurst, “Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR,” in *Mixed and Augmented Reality, 2008. ISMAR 2008. 7th IEEE/ACM International Symposium on*, 2008, pp. 193–202.

[70] Nicola Masotti and Franco Persiani, “Gaze-Coupled Perspective for Enhanced Human-Machine Interfaces in Aeronautics,” presented at the Conferences in Air Transport & Operations, Delft University of Technology, Delft, The Netherlands., 2015.

[71] Ben Lang, “VR Expert to Oculus Rift Devs: Make Sure You’re Doing 3D Right,” *http://www.roadtovr.com*, 25-Apr-2013. [Online]. Available: http://www.roadtovr.com/vr-expert-to-oculus-rift-devs-make-sure-youre-doing-3d-right/.

[72] “Good stereo vs. bad stereo,” *Doc-Ok.org*, 2012. .

[73] E. Peek, B. Wünsche, and C. Lutteroth, “Virtual Reality Capabilities of Graphics Engines,” presented at the 21st International Conference on Computer Graphics, Visualization and Computer Vision, 2013.

[74] P. D. Bourke, S. Hwy, and D. Q. Felinto, *Blender and Immersive Gaming in a Hemispherical Dome*. .

[75] Bimber Oliver and Ramesh Raskar, “Spatial augmented reality: merging real and virtual worlds,” 2005.

[76] C. Anthes, R. García-Hernández, M. Wiedemann, and D. Kranzlmüller, “State of the Art of Virtual Reality Technology,” in *IEEE Aerospace Conference 2016*, 2016.

[77] M. Velger, “Helmet-Mounted Displays and Sights,” *Disp. 21139–40*, Mar. 2000.

[78] C. E. Rash and R. W. Verona, *The human factor considerations of image intensification and thermal imaging systems*, vol. chap. 16. New York, Marcel Dekker: Karim, M. A., 1992.

[79] K. W. Moffitt, “Ocular responses to monocular and binocular helmet-mounted display configurations,” in *SPIE, Helmet-Mounted Displays, 1116*, 1989, p. 142.

[80] D. E. Thorpe, *Visual Requirements in HMDs: What can we see and what do we need to see?* New York: Mc Graw-Hill.: J.E. Melzer & KMoffit (Eds.), 1997.

[81] M. J. Wells and M. Venturio, “Performance and head movements using a helmet-mounted display with different fields-of-view.,” *Opt. Eng.*, vol. 29 (8), p. 870–877., 1990.

[82] M. Fiorentino, S. Debernardis, A. E. Uva, and G. Monno, “Augmented reality text style readability with see-through head-mounted displays in industrial context,” *Presence Teleoperators Virtual Environ.*, vol. 22, no. 2, pp. 171–190, 2013.

[83] S. Debernardis, M. Fiorentino, M. Gattullo, G. Monno, and A. E. Uva, “Text Readability in Head-Worn Displays: Color and Style Optimization in Video versus Optical See-Through Devices,” *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 1, pp. 125–139, Jan. 2014.

[84] L. Arbuthnot, B. Aitchison, C. Carlsen, C. N. King, T. Larsson, and T. Nguyen, “High-luminance active matrix electroluminescent (AMEL) display for see-through symbology applications,” 1999, vol. 3689, pp. 260–268.

[85] H. Urey, “Optical advantages in retinal scanning displays,” 2000, vol. 4021, pp. 20–26.

[86] H. Urey, N. Nestorovic, B. S. Ng, and A. A. Gross, “Optics designs and system MTF for laser scanning displays,” 1999, vol. 3689, pp. 238–248.

[87] J. P. Sauerborn, “Advances in miniature projection CRTs for helmet displays,” 1992, vol. 1695, pp. 102–116.

[88] R. Reisman and D. Brown, “Design of Augmented Reality Tools for Air Traffic Control Towers,” in *Proceedings of 6th AIAA Aviation Technology, Integration and Operations Conference (ATIO)*, Wichita, KS, 2006.

[89] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, “Vergence-accommodation conflicts hinder visual performance and cause visual fatigue,” *J. Vis.*, vol. 8, no. 3, p. 33.1–30, Jan. 2008.

[90] S. Liu, H. Hua, and D. Cheng, “A novel prototype for an optical see-through head-mounted display with addressable focus cues,” *IEEE Trans. Vis. Comput. Graph.*, vol. 16, no. 3, pp. 381–393, 2010.

[91] H. Hua and C. Gao, “A compact eyetracked optical see-through head-mounted display,” 2012, vol. 8288, p. 82881F–82881F–9.

[92] M. Vinnikov and R. S. Allison, “Gaze-contingent Depth of Field in Realistic Scenes: The User Experience,” in *Proceedings of the Symposium on Eye Tracking Research and Applications*, New York, NY, USA, 2014, pp. 119–126.

[93] M. Day, D. Seidel, L. S. Gray, and N. C. Strang, “The effect of modulating ocular depth of focus upon accommodation microfluctuations in myopic and emmetropic subjects,” *Vision Res.*, vol. 49, no. 2, pp. 211–218, 2009.

[94] L. O’Hare, T. Zhang, H. T. Nefs, and P. B. Hibbard, “Visual discomfort and depth-of-field,” *-Percept.*, vol. 4, no. 3, pp. 156–169, 2013.

[95] S. D. Peterson, “Stereoscopic Label Placement: Reducing Distraction and Ambiguity in Visually Cluttered Displays,” Dissertations, No. 1293, Linköping University, Division of Visual Information Technology and Applications (VITA) Department of Science and Technology, Linköping University SE-601 74 Norrköping, Sweden, 2009.

[96] W. Krueger and B. Froehlich, “The Responsive Workbench,” *IEEE Comput Graph Appl*, vol. 14, no. 3, pp. 12–15, May 1994.

[97] Persiani F., De Crescenzio F., Fantini M., and Bagassi S., “A Tabletop-Based Interface to Simulate Air Traffic Control in a Distributed Virtual Environment,” in *AA.VV., Atti del XVI Congresso ADM e XIX Congreso Ingegraf, PI, ETS, 2007*, Perugia, Italy, 2007.

[98] F. Persiani and A. Liverani, “A Semi-Immersive Synthetic Environment for Cooperative Air Traffic Control,” in *Proceedings of the 22nd Congress of International Council of the Aeronautical Sciences*, Harrogate, UK, 2000.

[99] C. Cruz-Neira, D. J. Sandin, and T. A. DeFanti, “Surround-screen Projection-based Virtual Reality: The Design and Implementation of the CAVE,” in *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, New York, NY, USA, 1993, pp. 135–142.

[100] T. A. DeFanti *et al.*, “The StarCAVE, a third-generation {CAVE} and virtual reality OptIPortal,” *Future Gener. Comput. Syst.*, vol. 25, no. 2, pp. 169–178, 2009.

[101] A. Liverani, F. Persiani, and F. De Crescenzio, “An Immersive Reconfigurable Room (I.R.R.) for Virtual Reality Simulation,” in *Electronic Proceedings of 12th International Conference on Design Tools and Methods in Industrial Engineering*, Rimini, Italy, 2001, pp. E1-9-- E1-16.

[102] M. Hirose, T. Ogi, S. Ishiwata, and T. Yamada, “A Development of Immersive Multiscreen Displays (CABIN),” in *Proceedings of the VRSJ 2nd Conference*, 1997, pp. 137–140.

[103] P. Bourke, “iDome: Immersive gaming with the Unity game engine,” in *Proceedings of the Computer Games & Allied Technology 09 (CGAT09), Research Publishing Services, ISBN: 978- 981-08-3165-3, pp 265-272.*

[104] Herman Rediess, “An augmented reality pilot display for airport operations under low and zero visibility conditions,” in *Guidance, Navigation, and Control Conference*, 1997.

[105] J. W. Ruffner, D. M. Deaver, and D. J. Henry, “Requirements analysis for an air traffic control tower surface surveillance enhanced vision system,” presented at the SPIE-The International Society for Optical Engineering, 2003, vol. 5081, pp. 124–135.

[106] N. Fürstenau, “Virtual Tower,” presented at the 5th ATM R&D Symposium, Braunschweig, 13.10-2005.

[107] R. Azuma, “A Survey of Augmented Reality,” in *Presence: Teleoperators and Virtual Environments 6*, 1997, vol. 6, 4, pp. 355–385.

[108] De Crescenzio F., Bagassi S., Fantini M., and Lucchi F., “Virtual Reality based HUD (Head Up Display) to simulate 3D conformal symbols in the design of future cockpits,” presented at the Council of European Aerospace Societies, Venice, Italy, 24-28 October, 2.

[109] Sara Bagassi, Francesca De Crescenzio, Francesca Lucchi, and Franco Persiani, “Innovation in Man Machine Interfaces: Use of 3D Conformal Symbols in the Design of future HUDs (Head Up Displays),” presented at the 28th International Congress of The Aeronautical Sciences, Brisbane, Australia, 2012.

[110] M. St. John, M. B. Cowen, H. S. Smallman, and H. M. Oonk, “The Use of 2D and 3D Displays for Shape-Understanding versus Relative-Position Tasks,” *Hum. Factors*, vol. 43, no. 1, pp. 79–98, 2001.

[111] S. Rozzi, P. Woodward, P. Amaldi, B. Fields, and W. Wong, “Evaluating Combined 2D/3D Displays for ATC,” in *Proceedings of the 5th EUROCONTROL Innovative Research Workshop & Exhibition*, Brétigny-sur-Orge , France, 2006, pp. 173–180.

[112] S. Rozzi, A. Boccalatte, P. Amaldi, B. Fields, M. Loomes, and W. Wong, “D1.1: Innovation and Consolidation Report Version 0.4,” EUROCONTROL, 3D-in-2D Display Project, Bretigny, Technical Report, 2007.

[113] R. Azuma, I. Neely H., M. Daily, and R. Geiss, “Visualization tools for Free Flight air-traffic management,” *Comput. Graph. Appl. IEEE*, vol. 20, no. 5, pp. 32–36, Sep. 2000.

[114] R. Azuma, M. Daily, and J. Krozel, “Advanced human-computer interfaces for air traffic management and simulation,” in *Proceedings of Flight Simulation Technologies Conference 1996*, San Diego,CA, 1996.

[115] S. Bagassi, F. D. Crescenzio, and F. Persiani, “Design and evaluation of a four-dimensional interface for air traffic control,” *Proc. Inst. Mech. Eng. Part G J. Aerosp. Eng.*, vol. 224, no. 8, pp. 937–947, 2010.

[116] M. A. Brown, “Displays for air traffic control: 2D, 3D and VR A Preliminary Investigation,” University of London, London, QMW Computer Science Technical Report 675, Dec. 1994.

[117] B.L. William Wong *et al.*, “3D-in-2D Displays for ATC.” .

[118] M. Tham and C. D. Wickens, “Evaluation of Perspective and Stereoscopic Displays as Alternatives to Plan View Displays in Air Traffic Central,” University of Illinois, Aviation Research Laboratory, Savoy, IL, Final Technical Report ARL-93-4/FAA-93-1, Jun. 1993.

[119] C. D. Wickens, S. Todd, and K. Seidler, “Terrain Representation for Air Traffic Control: A Comparison of Perspective with Plan View Displays,” University of Illinois, Aviation Research Laboratory, Savoy, IL, State of the Art Report CSER IAC-SOAR-89-001/AAMRL-TR-90, Oct. 1994.

[120] C. D. Wickens, “Display Integration of Air Traffic Control Information: 3D Displays and Proximity Compatibility,” University of Illinois, Aviation Research Laboratory, Savoy, IL, Final Technical Report ARL-9 5-2/FAA-9 5-2, Jan. 1995.

[121] A. Boccalatte, A. Monteleone, and S. Rozzi, “4d Visualization Constructs for Human-Machine Interfaces in the Control Tower Domain,” in *Proceedings of the 5th EUROCONTROL Innovative Research Workshop & Exhibition*, Brétigny-sur-Orge , France, 2006, pp. 165–166.

[122] S. G. Hart and L. L. Loomis, “Evaluation of the Potential Format and Content of a Cockpit Display of Traffic Information,” *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 22, no. 5, pp. 591–604, 1980.

[123] M. K. Kaiser and D. R. Proffitt, “Using the Stereokinetic Effect to Convey Depth: Computationally Efficient Depth-from-Motion Displays,” *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 34, no. 5, pp. 571–581, 1992.

[124] C. S. Narborough-Hall, “Recommendations for applying colour coding to air traffic control displays,” *Displays*, vol. 6, no. 3, pp. 131–137, 1985.

[125] D. Whitfield, “A preliminary study of the air traffic controller’s ‘picture’. Journal of the Canadian Air Traffic Controllers’,” *J. Can. Air Traffic Controll. Assoc.*, vol. 11, no. 1, pp. 19–28, 1979.

[126] R. H. Mogford, E. D. Murphy, R. J. Roske-Hofstrand, G. Yastrop, and J. A. Guttman, “Research Techniques for Documenting Cognitive Processes in Air Traffic Control: Sector Complexity and Decision Making,” Federal Aviation Administration Technical Center, Atlantic City International Airport, Technical Note DOT/FAA/CT-TN94/3, Jun. 1994.

[127] H. Le-Hong, “3D Visualization In Air Traffic Control,” in *Actes de la Première Conférence Internationale RIVF’03*, Hanoi, Vietnam, 2003, pp. 65–70.

[128] S. Bagassi, F. De Crescenzio, and F. Persiani, “Design and Development of an ATC Distributed Training System,” in *Proceedings of the 26th International Congress of the Aeronautical Sciences including the 8th AIAA 2008 ATIO Conference*, Anchorage, AL, 2008.

[129] M. Cooper, A. Fridlund, M. Andel, C. Bojan, and J.-L. Hardy, “Educational Benefits of 3D Displays in Early Controller Training,” in *Proceedings of the 27th International Congress of the Aeronautical Sciences including the 8th AIAA 2008 ATIO Conference*, Nice, France, 2010.

[130] S. T. Shorrock, “Errors of perception in air traffic control,” *Saf. Sci.*, vol. 45, no. 8, pp. 890–904, 2007.

[131] SESAR, “06.09.03 Remote & virtual TWR.” [Online]. Available: https://www.atmmasterplan.eu/projects/16484.

[132] C. Ware, *Information Visualization: Perception for Design*. Elsevier Science, 2004.

[133] M. Tory and T. Moller, “Human factors in visualization research,” *Vis. Comput. Graph. IEEE Trans. On*, vol. 10, no. 1, pp. 72–84, Jan. 2004.

[134] R. Kosara, H. Hauser, and D. L. Gresh, “An Interaction View on Information Visualization,” in *Proceedings of EUROGRAPHICS 2003*, Granada, Spain, 2003, pp. 123–137.

[135] C. D. Wickens, S. Todd, and K. Seidler, "Three-Dimensional Displays: Perception, Implementation and Applications.," University of Dayton, Research Institute, Dayton, OH, State of the Art Report CSER IAC-SOAR-89-001/AAMRL-TR-90, Oct. 1989.

[136] H. S. Smallman and M. St. John, "Naive Realism: Misplaced Faith in Realistic Displays," *Ergon. Des. Q. Hum. Factors Appl.*, vol. 13, no. 3, pp. 6–13, 2005.

[137] A. D. Andre and C. D. Wickens, "When Users Want What's not Best for Them," *Ergon. Des. Q. Hum. Factors Appl.*, vol. 3, no. 4, pp. 10–14, 1995.

[138] I. D. Haskell and C. D. Wickens, "Two- and Three-Dimensional Displays for Aviation: A Theoretical and Empirical Comparison," *Int. J. Aviat. Psychol.*, vol. 3, no. 2, pp. 87–109, 1993.

[139] C. D. Wickens, M. Campbell, C.-C. Liang, and D. H. Merwin, "Weather Displays for Air Traffic Control: The Effect of 3D Perspective," University of Illinois, Aviation Research Laboratory, Savoy, IL, Final Technical Report ARL-95-1/FAA-95-1, Jan. 1995.

[140] N. Naikar, "Perspective Displays: A review of Human Factors Issues," Department of Defence, Defence Science and Technology Organisation, Australia, DSTO-TR-0630, Feb. 1998.

[141] C. D. Wickens and R. Banks, "Commanders' Display of Terrain Information: Manipulations of Display Dimensionality and Frame of Reference to Support Battlefield Visualization," Army Research Laboratory, Aberdeen Proving Ground, MD, Technical Report ARL-CR-445, Sep. 1999.

[142] H. S. Smallman, M. St.John, H. M. Oonk, and M. B. Cowen, "Information availability in 2D and 3D displays," *Comput. Graph. Appl. IEEE*, vol. 21, no. 5, pp. 51–57, Sep. 2001.

[143] M. Lind, G. P. Bingham, and C. Forsell, "Metric 3D Structure in Visualizations," *Inf. Vis.*, vol. 2, no. 1, pp. 51–57, 2003.

[144] J. A. Perrone, "Visual Slant Underestimation: A General Model," *Perception*, vol. 11, no. 6, pp. 641–654, 1982.

[145] B. S. Boyer and C. D. Wickens, "3D Weather Displays for Aircraft Cockpits," University ofIllinois, Aviation Research Laboratory, Savoy, IL, Technical Report ARL-94-11, 1994.

[146] H. A. Sedgwick, "Space Perception," in *Handbook of Perception and Human Performance*, vol. Vol. 1: Sensory processes and perception, K. R. Boff, L. Kaufman, and J. P. Thomas, Eds. New York, NY: Wiley, 1986, p. 21.1-21.57.

[147] Q. Zhang, "Theoretical Review on a 3D Based Air Traffic Control System," Uppsala University, Department of Informatics and Media, Human-Computer Interaction., Uppsala, Sweden, 2012.

[148] A. Stokes, C. D. Wickens, and K. Kite, *Display Technology: Human Factors Concepts*. Warrendale, PA: Society of Automotive Engineers, 1990.

[149] R. Westrum, "Automation, Information and Consciousness in Air Traffic Control," in *Automation and Systems Issues in Air Traffic Control*, vol. 73, J. Wise, V. D. Hopkin, and M. Smith, Eds. Springer Berlin Heidelberg, 1991, pp. 367–380.

[150] L. Vygotsky, *Mind and Society*. Cambridge, MA: Harvard University Press, 1930.

[151] N.-T. Dang, "A Stereoscopic 3d Visualization Environment for Air Traffic Control: An Analysis of Interaction and a Proposal of New Interaction Techniques," Ph.D., Ecole Pratique des Hautes Etudes, Paris, 2005.

[152] M. Lange, T. Dang, and M. Cooper, "Interactive Resolution of Conflicts in a 3D Stereoscopic Environment for Air Traffic Control," in *Proceedings of the 4th International*

Confernce on Computer Sciences: Research, Innovation and Vision for the Future, Ho Chi Minh City, Vietnam, 2006, pp. 32–39.

[153] EUROCONTROL, "3D-in-2D Planar View Display," *eurocontrol.it*, 2015. [Online]. Available: http://www.eurocontrol.int/eec/public/standard_page/proj_CARE_INO_III_3D_2D.html. [Accessed: 17-Jan-2015].

[154] D. G. Jones, "Subjective Measures of Situation Awareness," in *Situation Awareness Analysis and Measurement*, M. R. Endsley and D. J. Garland, Eds. Mahwah, NJ: Lawrence Erlbaum Associates, 2000, pp. 101–114.

[155] P. G. Polson, C. Lewis, J. Rieman, and C. Wharton, "Cognitive walkthroughs: a method for theory-based evaluation of user interfaces," *Int. J. Man-Mach. Stud.*, vol. 36, no. 5, pp. 741–773, 1992.

[156] M. S. Burnett and W. Barfield, "Perspective versus Plan View Air Traffic Control Displays: Survey and Empirical Results," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, San Francisco, CA, 1991, vol. 35, 2, pp. 87–91.

[157] C. D. Wickens, S. Miller, and M. Tham, "The Implications of Datalink for Representing Pilot Request Information on 2D and 3D Air Traffic Control Displays," University of Illinois, Aviation Research Laboratory, Savoy, IL, Technical Report, 1994.

[158] P. Smith and C. Whetton, *GAT: General Ability Tests*. Windsor, UK: ASE, A Division of NFER-NELSON, 1988.

[159] K. F. V. Orden and J. W. Broyles, "Visuospatial task performance as a function of two- and three-dimensional display presentation techniques," *Displays*, vol. 21, no. 1, pp. 17–24, 2000.

[160] F. Persiani and A. Liverani, "A Semi-Immersive Synthetic Environment for Cooperative Air Traffic Control," in *Proceedings of the 22nd Congress of International Council of the Aeronautical Sciences*, Harrogate, UK, 2000.

[161] M. John, H. S. Smallman, T. E. Bank, and M. B. Cowen, "Tactical Routing Using Two-Dimensional and Three-Dimensional Views of Terrain," *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 45, no. 18, pp. 1409–1413, 2001.

[162] M. Tavanti, L. Hong-Ha, and D. Nguyen-Thong, "Three-Dimensional Stereoscopic Visualization for Air Traffic Control Interfaces: A Preliminary Study," in *The 22nd Digital Avionics Systems Conference: Proceedings*, Indianapolis, IN, 2003, vol. 1, p. 5.A.1-5.1-7.

[163] M. Cooper, M. Cooper, V. Duong, J. Hjalmarsson, M. Lange, and A. Ynnerman, "Interactive and Immersive 3D Visualization for ATC," presented at the 6th USA/Europe ATM 2005 R&D Seminar, Baltimore, MD, 2005.

[164] M. Lange, M. Cooper, V. Duong, and M. Ynnerman, "3D VR Air Traffic Management Project," EUROCONTROL Experimental Center, Brétigny-sur-Orge , France, Innovative Research: Activity Report 2004, 2004.

[165] R. J. Sternberg, *Cognitive psychology*. Fort Worth: Harcourt Brace College Publishers, 1996.

[166] G. Abadir, D. Karlsson, F. Molander, M. Schylström, H. Tjusberg, and K. Vegelius, "The future of Air Traffic Control? Digitalizing the mental image," *Unpublished*, 2010.

[167] Somervell, Jennifer, "Role of awareness of cognitive style in hypermedia," *Int. J. Educ. Technol.*, 1999.

[168] R. A. Rensink, “Internal vs. External Information in Visual Perception,” in *Proceedings of the 2nd International Symposium on Smart Graphics*, Hawthorne, NY, 2002, pp. 63–70.

[169] A. Boccalatte and W. Wong, “D1.3: Architecture and Common Simulation Engine Version 0.2,” EUROCONTROL, 3D-in-2D Display Project, Bretigny, 2007.

[170] NASA, “NASA Technology Readiness Level,” *http://www.nasa.gov*, 2015. [Online]. Available: http://www.nasa.gov/content/technology-readiness-level/#.VMZDucZ95rM.

[171] NEXT Ingegneria dei Sistemi SpA, “AD4: 4D Virtual Airspace Management System,” European Commission, Publishable Final Activity Report FAR, 2007.

[172] S. Rozzi, A. Boccalatte, P. Amaldi, B. Fields, M. Loomes, and W. Wong, “D1.2: Innovative Concepts and Their Design Rationale,” EUROCONTROL, 3D-in-2D Display Project, Bretigny, 2007.

[173] W. Wong, S. Gaukrodger, and F. Han, “D2.2: Year 1 Prototypes Evaluation,” EUROCONTROL, 3D-in-2D Display Project, Bretigny, Lot No. 1, WP 2, 2008.

[174] W. Wong, S. Gaukrodger, and F. Han, “Year 1 Prototypes Evaluation (Lot No. 1, WP 2),” EUROCONTROL, Bretigny, D2.2, 2008.

[175] S. Rozzi *et al.*, “Developing visualisations to support spatial-temporal reasoning in ATC,” in *Proceedings of the 2nd International Conference on Research in Air Transportation*, Belgrade, Serbia and Montenegro, 2006.

[176] C. D. Wickens, M. Vincow, and M. Yeh, “Applications of Visual Spatial Thinking: The Importance of Frame Reference,” in *The Cambridge Handbook of Visual Spatial Thinking*, P. Shah and A. Miyake, Eds. Cambridge, UK: Cambridge University Press, 2005, pp. 383–425.

[177] C. D. Wickens and A. D. Andre, “Proximity Compatibility and Information Display: Effects of Color, Space, and Objectness on Information Integration,” *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 32, no. 1, pp. 61–77, 1990.

[178] C. D. Wickens and C. M. Carswell, “The Proximity Compatibility Principle: Its Psychological Foundation and Relevance to Display Design,” *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 37, no. 3, pp. 473–494, 1995.

[179] N. Masotti and F. Persiani, “On the history and prospects of three-dimensional human–computer interfaces for the provision of air traffic control services,” *CEAS Aeronaut. J.*, vol. 7, no. 2, pp. 149–166, 2016.

[180] B. Hilburn, “Head Down Time in Aerodrome Operations: a Scope Study,” Center for Human Performance Research, 2014.

[181] V. Klymenko *et al.*, “Factors affecting the perception of luning in monocular regions of partial binocular overlap displays,” Aug. 1994.

[182] M. Schmidt, M. Rudolph, B. Werther, C. Möhlenbrink, and N. Fürstenau, “Development of an Augmented Vision Video Panorama Human-Machine Interface for Remote Airport Tower Operation.,” in *HCI (9)*, 2007, vol. 4558, pp. 1119–1128.

[183] K. Fält, “SAAB - Remote Tower,” 2012.

[184] “ISO 9241-11:1998: Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability.” Mar-1998.

[185] T. Shibata, J. Kim, D. M. Hoffman, and M. S. Banks, “The zone of comfort: Predicting visual discomfort with stereo displays,” *J. Vis.*, vol. 11, no. 8, pp. 11–11, 2011.

[186] “RETINA D2.1 - Conceptual Design.” 2017.

[187] N. Masotti and F. Persiani, "On the history and prospects of three-dimensional human–computer interfaces for the provision of air traffic control services," *CEAS Aeronaut. J.*, pp. 1–18, 2016.

[188] Vicente, Kim J. and Jens Rasmussen, "Ecological interface design: Theoretical foundations.," *Syst. Man Cybern. IEEE Trans.*, vol. 22, no. 4, pp. 589–606, 1992.

[189] J. Rasmussen, *Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering*. New York, NY, USA: Elsevier Science Inc., 1986.

[190] Vicente, K. J., "Ecological Interface Design: Supporting operator adaptation, continuous learning, distributed, collaborative work," in *Proceedings of the Human Centered Processes Conference*, 1999, pp. 93–97.

[191] J. Rasmussen, "Mental Models and the Control of Action in Complex Environments," in *Selected Papers of the 6th Interdisciplinary Workshop on Informatics and Psychology: Mental Models and Human-Computer Interaction 1*, Amsterdam, The Netherlands, The Netherlands, 1990, pp. 41–69.

[192] D. Embrey, "Understanding human behaviour and error." Human Reliability Associates, 2005.

[193] OpenGL, "gluPerspective," *www.opengl.org*. [Online]. Available: https://www.opengl.org/sdk/docs/man2/xhtml/gluPerspective.xml.

[194] Song Ho Ahn, "OpenGL Projection Matrix," *http://www.songho.ca*. [Online]. Available: http://www.songho.ca/opengl/gl_projectionmatrix.html.

[195] M. Woo, J. Neider, T. Davis, and D. Shreiner, *OpenGL Programming Guide: The Official Guide to Learning OpenGL, Version 1.2*, 3rd ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1999.

[196] Microsoft, "DirectX Graphics and Gaming," *www.msdn.microsoft.com*. [Online]. Available: https://msdn.microsoft.com/en-us/library/windows/desktop/ee663274(v=vs.85).aspx.

[197] "glFrustum - OpenGL," *www.opengl.org*. [Online]. Available: https://www.opengl.org/sdk/docs/man2/xhtml/glFrustum.xml.

[198] R. Kooima, "Generalized Perspective Projection," *J. Sch. Electron. Eng. Comput. Sci.*, 2009.