

The Lessons Learned in the Application of Augmented Reality

Mr Matthew Franklin

QinetiQ Ltd.

Cody Technology Park

Farnborough

Hampshire GU51 0LX

UNITED KINGDOM

Tel: (+44) 1252 393232

Fax: (+44) 1252 396406

Email: MFRANKLIN@QinetiQ.com

Website: www.qinetiq.com

ABSTRACT

The purpose of this paper is to provide an overview of the lessons learned from research conducted on behalf of the UK MoD into the application of Augmented Reality (AR) technologies. The lessons learned include technological and procedural aspects discovered during the preliminary research, user field trials and subsequent feedback. The paper is primarily focused on the application of AR in the training domain and in particular the use of AR to support Forward Air Controller training. However, the lessons learned also apply to the use of AR in operational and test and evaluation domains.

INTRODUCTION

The “Augmented Reality to represent synthetic air assets in the live domain” research was conducted between August 2003 and August 2005 under funding from the UK MOD’s Directorate of Analysis Experimentation and Simulation (DAES). The primary aim of the research was to assess the maturity of the techniques and technologies for the stimulation of the live domain with appropriate representations of simulated air assets.

The impetus for the research arose from studies in which Live, Virtual and Constructive (LVC) techniques were used to supplement live training exercises. A synthetic wrap was produced that included virtual aircraft and virtual representations of live ground assets. The virtual aircraft could see and interact with the live participants. However, the live participants had no representation of the virtual world. Although geospatial information about the location of the virtual entities could be passed to the live domain via the Command Information System (CIS) and other data links, there was no way of presenting that information in the modality that the users were accustomed to i.e. visual, aural, or other platform systems. Fig 1 gives an overview of the interactions and capabilities of each of the LVC elements.

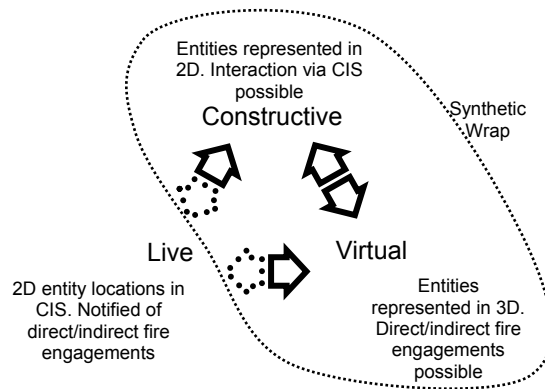


Figure 1: Connections and Capabilities of an Application of Synthetic Wrap.

It was realised that a method of stimulating the live domain was needed to support more complex synthetic wrap architectures. This would be fundamental to the provision of vertical training exercises, such as those required for time sensitive target training.

The research included a number of distinct phases including stakeholder analysis, technical analysis, development, trials and application analysis. The first two phases were required to provide a justification for the long term suitability and feasibility of AR in end user applications. More information on these phases can be found in ref [1] and [2]. The other phases resulted in the development and trialling of a test-bed system that was used to fully explore the technical and procedural issues associated with the use of AR. This paper is primarily concerned with the trials and application analysis phases of the research.

TRIALS

In March 2005 the test-bed system was successfully used by a Forward Air Controller (FAC) to conduct a Close Air Support (CAS) exercise involving synthetic aircraft. The FAC was able to use the test-bed to hear and visualise the synthetic aircraft. In addition the FAC was able to communicate, via voice radio, with the pilot of the flight simulator that was generating the synthetic aircraft. It was therefore possible for the FAC to guide the pilot onto the correct target. As well as providing synthetic air asset stimulus, the test-bed was also able to support the generation of synthetic ground based entities and effects. Since these entities were generated by a semi automated forces application it was possible to control their position and behaviour to suit the needs of the scenario. A limited number of fixed objects in the real world were also represented in the synthetic environment. This meant that the FAC could call for attacks on both live and synthetic objects and visualise the effects of the CAS mission on the targets. The overall system used for the trial is shown in fig 2.

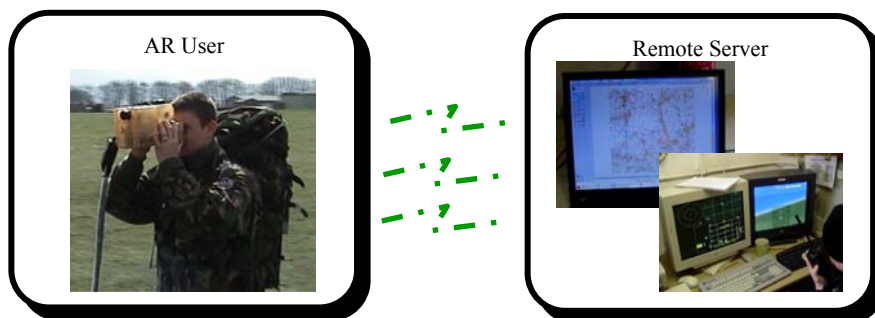


Figure 2: Trial System Components.

The conduct of the trials consisted of a series of short vignettes in which the synthetic aircraft would approach from varying directions and attempt to bomb targets whose location was described by the FAC. The intention was to measure the performance of the user in their ability to spot the aircraft and engage the correct target. It soon proved obvious that there were clear issues with the operation of the AR system that superseded these observations. These are described in greater detail below.

The trials were also used to provide an initial investigation into the use of AR to provide situational awareness information. Functionality implemented in the test-bed included the labelling of entities with call-signs, representing entities with Allied Publication and Procedures-6 (APP) symbology, associating sounds with entity types, overlaying mapping data on the terrain and providing orientation and location information as shown in Figure 3. Whilst detailed analysis of this functionality was not undertaken there were a number of observations made by users.



Figure 3: Screenshot Showing Tactical Symbology.

LESSONS LEARNED

The first observation made by most users was the overall size, weight and lack of ruggedness of the equipment. This was anticipated as part of the initial analysis phases but was deemed acceptable due to the wide range of requirements that led to its design, see ref [1]. In particular the test-bed was not required to achieve more than Technology Readiness Level (TRL) 3. It was recognised that any production system could be tailored to the particular needs of the end user application and could therefore be reduced in size or integrated into existing platform systems.

The inability to stimulate more than one user at a time was also a major concern. Most real world applications would require the AR solution to provide all systems stimulation, i.e. all participants in an exercise are provided with the same synthetic stimulus through all the systems from which they would normally expect to receive stimulation including sensors, optics, and naked eye. Without such stimulation it is unlikely that the users will be able to perform their tasks in a realistic manner, thus leading to reinforcement of false lessons. Due to the constraints of the programme only one AR system could be developed. Again, it was recognised in the early analysis that this may be a problem. In particular, stimulation of the naked eye as well as a binocular system. To supplement naked eye stimulation and to represent the possible stimulation of multiple users from one synthetic source, a surrogate radar type system was developed that could have been used to cue the user onto the general direction of the synthetic aircraft. For the particular scenario chosen for the trials the surrogate radar was not required since the user

was constantly being told the direction of the incoming aircraft, via radio, and so could orientate themselves such that the aircraft would appear within the field of view of the binoculars. Another approach that was suggested by the users was to emulate the peripheral vision of the user by showing blips at the edges of the binocular display. This approach was not implemented in the test-bed, however, 3D sound was used to provide similar stimulation. The use of 3D sound was found to dramatically improve the ability of the user to orientate themselves correctly. However, it introduced certain false lessons since the sound was generated from the binocular display and required the user to constantly use the binoculars to obtain the 3D effect.

As soon as the system started to be used it became apparent that the representation of the live world in the virtual was equally as important as representing the virtual in the live. Even simple differences in the terrain databases, such as the location and the type of bushes, meant that the FAC could not accurately describe the location of the target to the pilot. This has fair fight implications for the use of AR in a training role. During the course of the trials attempts were made to improve the virtual database with the use of photo realistic textures and increasing the fidelity and number of features. However, significant discrepancies still existed and dynamic elements such as weather, wind, and lighting effects remained problematic. Traditional approaches used to represent dynamic effects in the synthetic environment domain were not sufficiently accurate to replicate the nuances of the real world. i.e. cloud formations and associated shadows. They also required large amounts of manual intervention as these effects could not be automatically generated to match the real world.

Users of the test-bed indicated that a noticeable difference could be perceived when comparing real world entities against synthetic entities at approximately the same range. Whilst this was largely due to environmental effects such as visibility and precipitation it was also attributed to the insufficient resolution of the camera and displays used in the test-bed. This was anticipated in the earlier analysis, see ref [2], however, it was hoped that the resolution of the test-bed would be sufficient to conduct Detection Recognition and Identification (DRI) tasks up to 5km. In practice it was found that whilst detection and recognition of synthetic aircraft versus real aircraft was more difficult at ranges greater than 3km, recognition at shorter ranges was comparatively easier. This was believed to be as a result of the simplified geometry and textures used by the virtual models which made them easier to discriminate. To resolve these problems, increasingly complex synthetic models were used and varying levels of anti-aliasing applied. Increasing the anti-aliasing level effectively blurs the edges of the synthetic image, whilst decreasing it makes the synthetic image more pronounced. Whilst increasing the complexity of the geometry helped, it was found that decreasing the anti-aliasing had a greater effect on detection and recognition at longer ranges. At shorter ranges the use of photorealistic textures and increased anti-aliasing had a greater effect than increasing the complexity of the geometry.

Other key factors that highlighted the discrepancy between live and virtual entities were the differences in colour balance, contrast, and brightness. In particular the transition between viewing the sky and viewing the ground involved a drastic change in brightness. Whilst the aperture of the camera could be changed to ensure consistent light levels large variations in the brightness across the field of view remained. To solve some of this problem, image analysis techniques were used to adjust the virtual image to suit the properties of the incoming live image. However, these techniques could only be applied to the entire image, rather than small sub areas, resulting in unrealistic synthetic images where steep transitions in contrast or brightness were experienced e.g. when an entity entered an area of shadow caused by a cloud or crossed the horizon.



Figure 4: Colour Balance Problems when Generating Synthetic Imagery.

Prior to the trials it was thought that inaccuracies of the tracking solution would lead to unrealistic relative movement between synthetic entities and the live environment. However, it was found that it was possible to achieve believable levels of correlation by using a combination of high accuracy sensors and software correlation routines. By adjusting the resolution of the tracking solution it was found that a correlation error of about ± 1 mrad could be tolerated prior to virtual objects being perceived as moving relative to the ground. For the kind of head movements that the FAC was conducting, i.e. dismounted role, this equated to approximately 10ms. Another concern that was raised prior to the research was the implications of the tracking system and latency on simulator sickness. This was not found to be a problem with the test-bed despite a latency of 1/25 of a second between the image being captured and displayed to the user.

Absolute errors caused by sensor inaccuracies also caused problems during the trials as synthetic entities appeared to drift across the terrain even though they were actually stationary. This was not a major issue when viewing air entities, which had no fixed frame of reference against which to compare the drift. Even for ground entities, which could be compared against the ground, this drift was not rapid enough to drastically affect the trials. The key area that was affected by the drift was around the horizon where ground entities would appear to hover above the terrain and air entities appeared to pass in-front of hills rather than disappear behind them. For the vignettes used in the trials this problem area was not often crossed so did not significantly detract from the training experience.

Of greater concern to the users were the inaccuracies in the database used to ground clamp ground entities and perform occlusion calculations. Incorrect features or contours in the database would result in objects appearing in front of trees, hills and other objects when they should have been occluded, or visa versa. They also resulted in entities appearing to move across the ground as a result of discrepancies between the projected position on the ground and the actual position on the ground. This later observation was particularly apparent when synthetic objects were closer than 500m to the user. As above attempts were made to improve the fidelity of the databases in line with the virtual database described above. However, as above, this approach was limited by the inability to sense dynamic objects and properly recreate the intricacies of trees and grasses. For the particular vignettes used in the trials these occlusion problems were not perceived to be a major problem. This was primarily due to the barren nature of the trials area but was also down to the lack of objects near to or traversing the horizon.

Another issue that was identified by the users of the equipment were dynamic errors caused by the inability to accurately reproduce the movement of synthetic entities across the limited bandwidth communications that were available. This resulted in synthetic entities moving unrealistically and limiting the ability of the user to track them. This was believed to be mostly due to the high second order accelerations exhibited by the synthetic aircraft, but could also be attributed to interference and latencies introduced by the communications mechanism.

The Lessons Learned in the Application of Augmented Reality

The introduction of call-signs significantly improved the ability for users to detect and discriminate objects at range. When combined with 3D sounds it provided the user with a rapid means to assimilate situational awareness information without requiring them to take their eyes off the primary area of interest. However, concerns were raised as to the feasibility of such stimulation given the restricted bandwidth of current operational communication systems. It was suggested that rather than being transmitted to the user, the information could be generated locally by onboard sensor feeds such as radars, however, concerns as to the timeliness and accuracy of the data still remained. The introduction of static information such as mapping data and orientation data overlaid on the real world were of immediate benefit to the users who rapidly adapted their procedures to exploit the new forms of information. For example users could refer to objects on the ground in terms of map features. Since the map features were common to all participants it avoided the need to interpret descriptions of real world features. This was particularly useful for the FAC scenario as the FAC and pilot have widely different viewpoints.

CONCLUSIONS

Despite the relative success of the trials the application of AR for training will be dependent on the specific requirements of the end user and in particular the realism of the stimulation required. This will be dependent on:

- Means of stimulation required
- Criticality to the primary task and the lessons to be learnt
- Dynamism and complexity of the training environment
- Availability of a common synthetic environment

Stimulating head down displays that are fed by camera images is believed to be the simplest means of providing AR. The platform usually provides a means to host the sensors and communications systems necessary for AR applications. In addition the use of camera images enables the application of image analysis and anti-aliasing needed to ensure the realistic blending of synthetic entities and effects into the live images. Stimulating platform or weapon sights is more difficult because of physical limitations of current display technology. Since the live images are presented to the user in real time it imposes critical constraints on the overall latency of the tracking and image generation processes. From the analysis conducted by the research this would need to be a maximum of 10ms, although it is expected to be even lower with the inherent higher resolution of real world optics. In order to ensure maximum realism it will also be necessary to include a through-sight video camera to enable image analysis to take place as shown in fig 5.

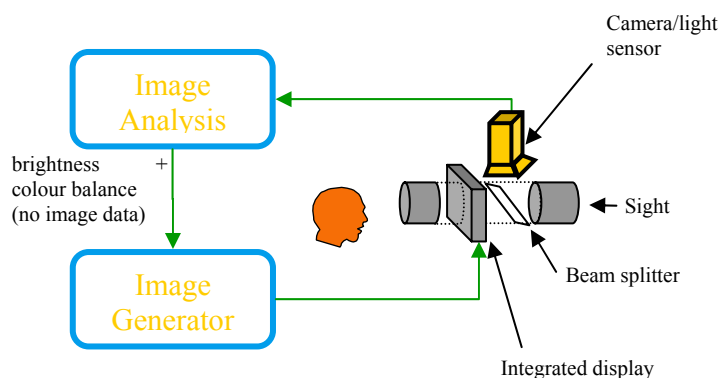


Figure 5: Integrating Image Analysis into a Direct Optical AR Display.

Stimulating the naked eye remains technically challenging because of the resolution required and the accuracy at which the display device can be positioned in relation to the eye. It also imposes limitations on the ability to easily use other optical systems such as binoculars and weapon sights. It was found that stimulation of the naked eye could potentially be avoided through careful selection of the scenario or by providing stimulation via another mechanism. Stimulating other platform systems such as sensors and C4I systems can help provide some of this stimulation but is also likely to be an integral part of any AR application. All systems stimulation is seen as a key element for enabling many tasks to be undertaken. However, this is likely to require embedded functionality in operational systems that does not currently exist.

The criticality of realistic stimulation will also depend entirely on how the synthetic stimulation will be used. It has been recognised that AR could be used to enhance already existing live exercises by providing additional clutter to the tactical picture. The clutter adds to the cognitive load of the users but does not form part of their primary task. Even for stimulation that forms part of the primary task of the users the inability to generate completely realistic synthetic entities and effects may not detract from the lessons to be learned. E.g. the location of the fall of shot of indirect fire only needs to convey the information about the location and not the intricacies of the propagation of blast effects.

The dynamism and complexity of the training environment impacts on both the realism of synthetic image provided to the user but also on the fair fight issues associated with recreating the real world in the synthetic environment. Completely realistic occlusion databases could be pre-generated for many range based applications where the firing position and range environment is largely static. For more dynamic applications that involve platforms it may be possible to mount sensors to generate occlusion databases on the fly. It has been identified, that despite improvements in the database generation process discrepancies between the live and the virtual environments are still likely to occur and limit the suitability of AR to particular end user applications for the foreseeable future e.g. training within woodland.

Many AR training applications will also be dependent on the availability of a communications mechanism with sufficient bandwidth to support real time communications between AR users and a common synthetic environment. Without such a mechanism maintaining realism and achieving a fair fight is unlikely to be feasible. The bandwidth requirements will be dependent on the rates of movement of the AR users and the numbers of synthetic entities and effects. It will also be dependent on the level at which other communications systems need to be emulated e.g. voice radio communications. Some applications may not require AR users to interact with other systems therefore the common synthetic environment can be generated locally to the user with limited dependency on a communications capability.

The use of AR to aid in the provision of situational awareness information is also dependent on the end user's application. The same technical and procedural constraints apply as in the training domain. However, instead of being critical to the realism of the imagery it is important that the information is accurate. Of particular concern is the latency of the information and avoiding adding to the cognitive workload of the user. By presenting the information in the primary modality of the user e.g. mapping data, it was observed that information could be assimilated more rapidly and easily. In fact entirely new means of operation were made possible by the introduction of tactical information that was common to all participants. It should also be noted that since realism is no longer a critical element the technical requirements from many of the system components is believed to be less than for a training application.

The primary obstacle to the use of AR in defence applications remains the trade-off between the cost and the benefit. It has been shown that AR can enable new ways of training and facilitate new operational procedures that were previously not possible. AR is also able to enhance existing training and support existing ways of operating at much lower technical and human risk. However, the cost associated with producing systems for these applications is prohibitively high until clear and costed benefits can be

identified. The AR test-bed is an ideal platform for investigating these benefits and driving the need for further technical development.

ACKNOWLEDGEMENTS

Thanks go to MOD Directorate of Analysis Experimentation and Simulation (DAES) for sponsoring this work, and to the Joint Air Support Organisation (JASO) and Joint Forward Air Control Training Support Unit (JFACTSU) for their support.

REFERENCES

- Franklin M. (2005), *Augmented Reality for live training*, ITEC 2005.
Franklin M. (2005), *Synthetic Air Assets in the Live Domain*, IITSEC 2005.