Standards for Augmented Reality: a User Experience perspective

Panagiotis D. Ritsos, Dennis P. Ritsos, Antonis S. Gougoulis
Synthetic Toys, Greece
February 16, 2011

Abstract

An important aspect of designing and implementing Augmented Reality (AR) applications and services, often disregarded for the sake of simplicity and speed, is the evaluation of such systems, particularly from non-expert users, in real operating conditions. We are strong advocates of the fact that in order to develop successful and highly immersive AR systems, that can be adopted in day-to-day scenarios, user assessment and feedback is of paramount importance. Consequently, we also feel that an important fragment of future AR Standardisation should focus on User eXperience (UX) aspects, such as the sense of presence, ergonomics, health and safety, overall usability and product identification. Our paper attempts an examination of these aspects and proposes an adaptive theoretical evaluation framework than can be standardised across the span of AR applications.

Keywords: Augmented Reality (AR), User Experience (UX), Standards, Assessment, Sense of Immersion

1 Introduction

AR has been a subject of research and scientific literature for almost two decades [11]. The high number of different AR paradigms, proposed solutions, applications, frameworks and services present throughout the lifespan of the field depicts the latter's diversity and complexity. Nonetheless, during the last couple of years or so mobile AR drew significant media and public attention [30] when the concept was brought out of the laboratory and in to the smartphone user's hands. Indeed, the current market popularity of Handheld AR [16, 28, 47, 49] hints that this paradigm may be the backbone on which AR overall will become popular.

Although Handheld AR draws much attention these days with some undoubtedly very impressive and well-marketed applications there are many flavours of the concept that remain unexplored, inadequately developed and often hindered by our current technological progress in other sectors. It is characteristic that many research efforts of the past used technologies 'borrowed' from other fields. Say for example, Wearable AR [13, 17, 18, 35–41, 50] was always dependant on the performance of small-form factor units - laptops or single board computers (SBCs). Localisation still remains a challenge and the performance advancement of sensory modalities is often irrelevant to the field of AR. Head Mounted Displays, the definitive icon of Virtual Reality also used for Augmented are often inadequate for the latter as they are primarily designed for indoors, with poor brightness and deficient optics for real-world viewing. Even Handheld AR faces limitations, intrinsic to smartphones such as small screen size.

Nonetheless, if current popularity trends do continue and people start using the 'well-marketed' paradigms of today, maybe we can reinforce our support for the existence, development, research and application of technologies specific and tailor-made to AR. Amidst these developments, AR experts are discussing the need for AR standardisation, to assist larger adoption of the concept and further innovation. AR brings together a community made of various groups with different perspectives of the field, such as researchers, developers, marketeers, science fiction
authors and industry analysis, forming a mosaic of views and opinions that can be an asset in forming such standards for this highly engaging flavour of human computer — and environment — interaction. In this paper we attempt an overview of standardisation from a UX for AR (Ux4AR) perspective.

The remainder of this paper is organised as follows. Section 2 gives an overview of the importance of standards in AR. Section 3 presents the concept of UX within the scope of AR and Section 4 presents our theoretical framework. Finally, Section 5 summarises our conclusions.

2 The Need for Standards in AR

AR Standardisation as a whole can, in principle, provide a schematic on which parties involved can build services and applications, while encouraging innovation through clearly defined practices, maintaining inter-operability, generating market trust and allowing collaborations and efficient information exchange. Moreover, like most interactive technologies of modern times there is need for semantic unification, safety and health regulations, quality assurance and, possibly in the future dedicated legislation — say for integrity, privacy and security.

However, AR currently is a very volatile field. The variety of different devices and platforms currently employing AR services, the undeniable technological challenges, the different research approaches and concepts, coupled with a market hype that may or may not be justifiable are obstacles that need to be overcome. Moreover, the depth of AR as a field, spanning from handhelds to medical AR and military simulations, results in a multidimensional space where requirements are highly dependent on the nature of the offered service. In an attempt to contribute to this effort we look into the standardisation needs following a human-centred approach.

3 User Experience in AR

ISO 9241-210[1] defines user experience as "a person’s perceptions and responses that result from the use or anticipated use of a product, system or service". According to the same definition UX includes the users’ emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviours and accomplishments that occur before, during and after use of a system. The term “user experience” is used along with “usability” with varying degrees of relativity. However, we consider UX to be a larger entity, encompassing usability and including both pragmatic and hedonic aspects of a system.

Regarding AR, many definitions [4,5] often imply the use of 3D graphics superimposed on the user’s view of the world, emphasising the visual aspect of AR. However, from a user experience point of view it can be any media — visual, sound, haptic, etc — that enhances the user’s reality and specific context [43], thus addressing the meanings of locality and intentionality [24]. Additionally, the nature and form of the user experience is affected by the number and type of interactions within the synthetic space.

Figure 1: Ux4AR Framework Goals

Our analysis revolves around devising a theoretical framework that would encompass aspects of AR that are important to UX with a twofold purpose (Figure 1). One is to identify those elements that are important to Ux4AR and use them as a roadmap for AR standards. Furthermore, this theoretical framework can be a roadmap for Ux4AR assessments - field surveys questionnaires etc. Thus, any method of assessment based on such a framework can, subsequently, be part of a standardised methodology concerning the evaluation of AR prototypes. Naturally, bearing in mind the aforementioned fluidity of AR, the framework needs to be adaptable to any progress within the field. Moreover, our analysis initiated from previ-
ous research in Wearable AR [39–41], is not exhaustive but more of a preliminary presentation.

4 UX Assessment Framework

The purpose of the framework is to set the grounds for the assessment of AR services and applications. By reviewing research efforts in AR/VR based on questionnaires, such as the ones from Avery et al. [2,3], Gabbard et al. [14,15], Ames et al. [1], Billinghurst et al. [8] and Siegel and Bauer [44] we have identified aspects of AR that can potentially affect user experience. In theory, one should apply this framework of assessment in different scenarios and, through analysis, identify the underlying patterns and correlated factors that affect the user’s overall experience.

However, it must be noted that within each grouping of notions presented below there is great variation on the specific requirements because of the multifarious nature of AR. The ergonomic requirements, for example, for a Military-type AR simulator are different from a handheld AR tour-guide. This is an issue with devising AR standardisation that addresses — as much as possible — the concept globally and does not merely focus on specific flavours. In an attempt to provide a starting point, though, we identified certain ‘core’ concept groupings that we feel can form a foundation layer and be expanded upon a case-by-case basis. These concepts are depicted in Figure 2 and subsequently analysed. The reader must note that, in practice, most concepts intertwine and the underlying correlations must be, ultimately, taken into account.

4.1 Input

Input can be separated in four major categories.

Visual: Visual input usually implies the use of a camera for tracking and context identification. Examples can be marker-based tracking [21] or marker-less [25]. Applications range from simple tracking, in order to place virtual objects on top of those targets to more complex tracking to determine gestures, navigate through interfaces [28] etc. Aspects that could affect UX in this category are the ease of use, responsiveness, accuracy and privacy/social comfort of using them.

Auditory: Sound can also be useful as input, both for direct voice commands as well as inferring the user’s context [31], say from ambient noise levels, patterns or detected frequencies. The same aspects that affect visual apply in this case, too. However, privacy and social comfort of using voice commands is more important for direct commands.

Tactile: By tactile we classify all interfaces that require contact (touch) with a surface, either a keyboard, a touch-screen, a joystick or a mouse. Once again the aspects affecting user experience are similar as above with one added issue involving obtrusiveness for mobile, untethered systems. Touchscreens [13], chord and small-form keyboards [26,32,40] have been used in the past but are often a hindrance and quite tiring to use after a while.

Kinæsthetic: Recent advancements in motion tracking, have lead to various implementations with the aptly-named Kinect® from Microsoft as the currently most popular example. Such interfaces result into another paradigm of interaction where the system recognises — without the need of a body-worn apparatus or sensors — movement and posture and having arguably high potential of utilisation in AR context. Once again, ease of use, responsiveness and accuracy are factors that affect UX.

Sensory Modalities: Another form of input is using sensors to detect the user’s and the environment’s context. Researchers have investigated the use of various schemes in order to increase accuracy of sensory modalities, most notably combining more than one (sensor fusion) to what is often referred to hybrid sensors. Active (sensor-emitter) tracking technologies require powered-device installation and are often susceptible to interference, whereas inertial sensors, although completely passive, exhibit drift. In addition, the aforementioned vision-based sensors are

\[\text{http://www.xbox.com/en-us/kinect}\]
computationally demanding and often unusable (occlusion). Combining more than two sensors into complementary fashion is an approach taken by many researchers. One issue affecting UX in this case, apart from the overall accuracy of the arrangement, is seamless switching between modalities when that is needed.

### 4.2 Output

Output, likewise, can be separated in three major categories.

**Visual:** Visual output is probably the most important aspect of augmented reality, ever present through the history of the field and in almost all incarnations and concepts. However, the type of visual information ranges from simple informational annotations to complex 3D architectural and humanoid modelling, resulting in different requirements for effective UX. It is also important to note that high fidelity and accurate representation of modelled entities is not always needed and depends on the application and the intended abstraction level.

However, there are various implications on health and safety [22, 23] from the use of certain output systems, such as HMDs for example. The presentation media issues are of paramount importance to Ux4AR as issues like narrow field-of-view, inadequate depth perception, low display brightness and poor ergonomics [6, 12, 27] can hinder any sense of presence.

Moreover, visual output encompasses one of the most acute and blocking problems in AR, that of registration, also related to Context Awareness presented further on. Humans have an extremely sensitive perceptual system, able to detect small anomalies and irregularities such as mis-registrations and delays [39]. In principle the ideal solution for accurate registration and localisation is a positional error of 1 mm and angular error of less than 0.5° with no drift [5]. A more realistic aim is to achieve positional stability so that a user is able to negotiate doorways in an AR reconstruction without difficulty — say 0.2 m. Angular errors however need be of the aforementioned level, otherwise, large discrepancies occur as the distance increases [4, 5].

Overall, things that should, primarily, be assessed from a UX perspective regarding visual output include:

- presentation media quality (field of view, brightness, contrast, depth perception, ease of use etc.)
• content quality (realism, abstraction, frame rates etc)
• synthetic world consistency and stability (registration, temporal and spatial stability).

Auditory: Auditory output is somewhat simpler to implement with reasonable quality as available technology can provide high-fidelity, directional sound with small cost and with non-obtrusive gear (headphones etc). However, sound can be disruptive towards the user and his or her environment, having noteworthy safety and privacy implications. Therefore, apart from an appraisal of how sound feedback does contribute to UX, health and social aspects of using sound as an augmentation method should be included.

Haptic: Although synthetic visual and audio paradigms are present throughout AR, examples of haptic interfaces, where the user is able to touch and feel objects are limited [29] and are usually found in the form of force-feedback [9, 51] systems, such as the ones found in gaming console controllers and in medical applications [7, 19]. Although our experience in the field is limited, we feel that the ultimate AR system would induce more than just visual or audio stimulation and potentially enhance even further the sense of immersion within a user’s synthetic (natural combined with digital) environment. From a user experience point of view ease of use, accuracy, feel, obtrusiveness and overall ergonomics are important for utilising such systems in AR context.

4.3 Context Awareness
AR’s contextual and in situ nature demonstrates how the concept is intertwined with Context Awareness (CA). Extracting information about a user’s location, posture, intentions as well as environmental features is an inherit function and has direct consequences to how synthetic information is utilised and ‘placed’ spatially and temporally. Although we have briefly touched the subject from the interfaces point of view, we feel it is important to focus on the mechanics themselves as they are the source of the discrepancies described before. The main aspects of context are: where you are, who you are with, and what resources are nearby [42]. It is comprised of more than the user’s location, including, lighting, ambient noise, tethering, bandwidth and social conditions. Detecting these ‘features’ accurately, in real-time remains one of the fundamental challenges in AR. Properly placing and registering synthetic information spatially and temporally is the next. Arguably, the level of accuracy required by AR [5] is not achievable with current tracking technology [4] that employs more than mere annotations and multi-modal (hybrid) schemes have to be used to enhance accuracy and speed.

Human perception is extremely sensitive to spatial and temporal discrepancies. From a user experience point of view it is important to gauge the opinion of users on sensory modalities identifying what they perceive as accurate and properly positioned spatially. Moreover, it is also interesting to assess the temporal positional stability and consistency of the synthetic environment, for example if a synthetic chair is in the ‘same place’ when entering and exiting a room after some time.

4.4 Use Cases
One recent contribution from the community working on AR standards is the introduction of “use cases” [2]. Use cases essentially describe an application by classifying it in three major categories — Guide, Create and Play — and must meet the criteria of augmented reality, as described by Azuma [4]. It is reasonable to expect that each use case may have specific requirements and underlying dynamics, in terms of UX. Moreover, use case mapping can help address UX4AR as a whole and introduce unification that spans across the aforementioned flavours of AR.

4.5 Health and Safety
Health and safety concerns are of paramount importance, in all AR scenarios. Although well-thought AR can enhance one’s reality, unwise implementations can be potentially disruptive, cause accidents and in extreme cases have health implications. HMDs for example have long been...
in the centre of investigation for ocular and non-ocular symptoms of use [22,23]. We feel any form of standardisation in AR should include these aspects both in terms of regulations and guidelines as well as the focus of assessments [33]. Feedback from healthcare and medical field experts would be immensely helpful on the subject.

4.6 Integrity, Privacy and Security

AR currently remains a 'personal' experience to great extent. However, much like users currently tend to 'meet' in various shared spaces – essentially domains of specific content, like social networks, massive multiplayer online (MMO) games etc. – with enormous implications regarding interaction, we can expect a change of this paradigm. The ultimate incarnation of the 'Play' use case, described above is a shared synthetic environment. Where participants can explicitly and implicitly interact with each other and with their real and digital environment.

Nonetheless, shared environments have intrinsic integrity, security and privacy implications. Just as Vernon Vinge describes in Rainbows End [46], sharing or accessing someone's 'view' of things may or may not always be desirable. Also, hiding information — or indeed true identities — behind synthetic 'cloaks', or allowing 'virtual' access to otherwise protected areas and information are examples of security breaches. Granted, this level of augmentation may appear to some as technologically distant, but implications of the notion of shared environments have been examined for some time [10,20]. It is only wise to include these aspects in any AR standardisation.

4.7 Sense of Immersion

All of the above concepts contribute to varying degrees to what we call “sense of immersion”, otherwise knows as “presence”. Many assessments, concerning both VR and AR environments try to quantify immersion [2,34,45,52] denoting its importance. One would say that sense of immersion is the integration of attitudes towards a system, evaluated by the user in terms of importance. For some people the poor quality of the fidelity of the synthetic world is restrictive while others find registration problems and spatially instability is more disruptive. In any case, we could say that, to a large extent, sense of immersion is the archetype of user experience in AR.

5 Conclusion

Augmented Reality is a technology that has been around as a concept for almost two decades. During the last couple of years it received a great deal of publicity through various implementations of the notion on smartphones, mainly involving image and textual annotations as well as simple marker-positioning of 3D models. Nonetheless, AR has many different flavours with different requirements and goals. Any form of AR standardisation must cater for the multi-dimensionality of the field and incorporate requirements, suggestions, practices and regulations proportionate to the level of immersion and specific needs of each AR use case.

User experience is about how a user feels about using a system encompassing feelings, motivation, satisfaction and overall attitude. In a sense, in AR context, UX can be parallelised to the feeling of immersion, as far as usage of a system is concerned. However, UX as a whole also includes branding, marketing image, standards compliance support and overall quality of service offered.

We have attempted an overview of the technical aspects that we feel can affect UXAR, and tried to present the theoretical foundation for devising related standards. Although our list is not exhaustive, we feel it can form the basis of further discussion and experimentation by assessment in AR. We have classified UX content in the categories of input (visual, auditory, tactile and kinesthetic), output (visual, auditory and haptic), context awareness, sense of immersion, health, safety and integrity, privacy and security. These broad categories encompass various sub-concepts that need careful and dedicated consideration.

AR is foremost a human-centred technology. It is a concept whose sole purpose is to enhance one's — or a group's — reality. We feel that a human-centred approach is of paramount importance to AR standardisation and an excellent starting point to enhance the field's technological and marketing reality.
References


