PROCEEDINGS OF SPIE

Augmented reality: a class of displays on the reality-virtuality continuum

Paul Milgram, Haruo Takemura, Akira Utsumi, Fumio Kishino

SPIE.

Augmented Reality: A class of displays on the reality-virtuality continuum

Paul Milgram¥, Haruo Takemura¶, Akira Utsumi†, Fumio Kishino‡

ATR Communication Systems Research Laboratories ƒ
2-2 Hikaridai, Seika-cho, Soraku-gun
Kyoto 619-02, Japan

ABSTRACT

In this paper we discuss Augmented Reality (AR) displays in a general sense, within the context of a Reality-Virtuality (RV) continuum, encompassing a large class of "Mixed Reality" (MR) displays, which also includes Augmented Virtuality (AV). MR displays are defined by means of seven examples of existing display concepts in which real objects and virtual objects are juxtaposed. Essential factors which distinguish different Mixed Reality display systems from each other are presented, first by means of a table in which the nature of the underlying scene, how it is viewed, and the observer's reference to it are compared, and then by means of a three dimensional taxonomic framework, comprising: Extent of World Knowledge (EWK), Reproduction Fidelity (RF) and Extent of Presence Metaphor (EPM). A principal objective of the taxonomy is to clarify terminology issues and to provide a framework for classifying research across different disciplines.

Keywords: Augmented reality, mixed reality, virtual reality, augmented virtuality, telerobotic control, virtual control, stereoscopic displays, taxonomy.

1. INTRODUCTION

Our objective in this paper is to review some implications of the term "Augmented Reality" (AR), classify the relationships between AR and a larger class of technologies which we refer to as "Mixed Reality" (MR), and propose a taxonomy of factors which are important for categorising various MR display systems. In the following section we present our view of how AR can be regarded in terms of a continuum relating purely virtual environments to purely real environments. In Section 3 we review the two principal manifestations of AR display systems: head-mounted see-through and monitor-based video AR displays. In Section 4 we extend the discussion to MR systems in general, and provide a list of seven classes of MR displays. We also provide a table highlighting basic differences between these. Finally, in Section 5 we propose a formal taxonomy of mixed real and virtual worlds. It is important to note that our discussion in this paper is limited strictly to visual displays.

---

2. REALITY-VIRTUALITY CONTINUUM

Although the term "Augmented Reality" has begun to appear in the literature with increasing frequency, we contend that this is occurring without what could reasonably be considered a consistent definition. For instance, although our own use of the term is in agreement with that employed in the call for participation in the present proceedings on Telemanipulator and Telepresence Technologies[1], where Augmented Reality was defined in a very broad sense as "augmenting natural feedback to the operator with simulated cues", it is interesting to point out that the call for the associated special session on Augmented Reality took a somewhat more restricted approach, by defining AR as "a form of *virtual reality* where the participant's *head-mounted display* is transparent, allowing a clear view of the real world" (italics added). These somewhat different definitions bring to light two questions which we feel deserve consideration:
- What is the relationship between Augmented Reality (AR) and Virtual Reality (VR)?
- Should the term Augmented Reality be limited solely to transparent see-through head-mounted displays?

Perhaps surprisingly, we do in fact agree that AR and VR are related and that it is quite valid to consider the two concepts together. The commonly held view of a VR environment is one in which the participant-observer is totally immersed in a completely synthetic world, which may or may not mimic the properties of a real-world environment, either existing or fictional, but which may also exceed the bounds of physical reality by creating a world in which the physical laws governing gravity, time and material properties no longer hold. In contrast, a strictly real-world environment clearly must be constrained by the laws of physics. Rather than regarding the two concepts simply as antitheses, however, it is more convenient to view them as lying at opposite ends of a *continuum*, which we refer to as the *Reality-Virtuality (RV) continuum*. This concept is illustrated in Fig. 1 below.
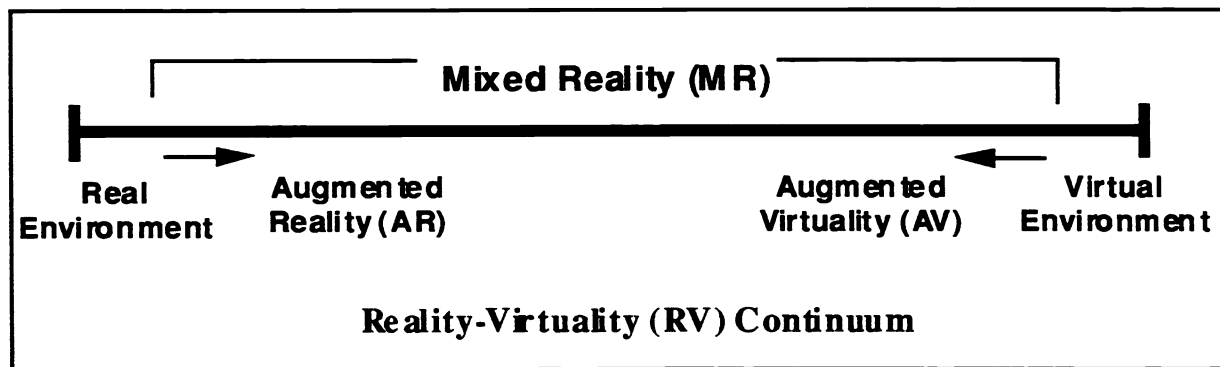

Figure 1: Simplified representation of a RV Continuum.

The case at the left of the continuum in Fig. 1 defines any environment consisting solely of real objects, and includes whatever might be observed when viewing a real-world scene either directly in person, or through some kind of a window, or via some sort of a (video) display. The case at the right defines environments consisting solely of virtual objects, examples of which would include conventional computer graphic simulations, either monitor-based or immersive. Within this framework it is straightforward to define a generic Mixed Reality (MR) environment as one in which real world and virtual world objects are presented together within a single display, that is, anywhere between the extrema of the RV continuum.[2]

3. TWO CATEGORIES OF AUGMENTED REALITY DISPLAYS

Within the context of Fig. 1, the above-mentioned broad definition of Augmented Reality – "augmenting natural feedback to the operator with simulated cues" – is quite clear. Also noteworthy in this figure is the corresponding concept of *Augmented Virtuality (AV)*, which results automatically, both conceptually and lexically, from the figure. Some examples of AV systems are given below, in Section 4. In the present section we first contrast two cases of AR, those based on head-mounted see-through displays and those which are monitor based, but both of which comply with the definition depicted in Fig. 1.

3.1 "See-through" AR displays

This class of displays is characterised by the ability to see through the display medium directly to the world surrounding the observer, thereby achieving both the maximal possible extent of presence and the ultimate degree of "realspace imaging".[3] Most commonly display augmentation is achieved by using mirrors to superimpose computer generated graphics optically onto directly viewed real-world scenes. Such displays are already a mature technology in some (mostly military) aviation systems, as either panel-mounted or head-mounted displays (HMD's), but are currently finding new applications as a VR related technology, especially in manufacturing and maintenance environments.[4,5,6] Of special interest are recent efforts to apply such AR displays to medical imaging, by superimposing data acquired via imaging techniques such as ultrasound, CT scanning, etc. conformally onto the actual patient.[7,8]

Several research and development issues have accompanied the advent of optical see-through (ST) displays. These include the need for accurate and precise, low latency body and head tracking, accurate and precise calibration and viewpoint matching, adequate field of view, and the requirement for a snug (no-slip) but comfortable and preferably untethered head-mount.[5,9] Other issues which present themselves are more perceptual in nature, including the conflicting effects of occlusion of apparently overlapping objects and other ambiguities introduced by a variety of factors which define the interactions between computer generated images and real object images.[9] Perceptual issues become even more challenging when ST-AR systems are constructed to permit computer augmentation to be presented stereoscopically.[10]

Some of these technological difficulties can be partially alleviated by replacing the optical ST with a conformal *video*-based HMD, thereby creating what is known as "video see-through". Such displays present certain advantages, both technological and perceptual[9], even as new issues arise from the need to create a camera system whose effective viewpoint is identical to that of the observer's own eyes.[11]

3.2 Monitor based AR displays

We use the term monitor-based (non-immersive), or "window-on-the-world" (WoW), AR to refer to display systems where computer generated images are either analogically or digitally overlaid onto live or stored video images.[12,13,14] Although the technology for achieving this has been well-known for some time, most notably by means of chroma-keying, a large number of useful applications present themselves when this concept is implemented *stereoscopically*.[15,16,17]

In our own laboratory this class of monitor-based AR displays has been under development for some years, as part of the ARGOS (Augmented Reality through Graphic Overlays on Stereovideo) project.[18] Several studies have been carried out to investigate the practical applicability of, among other things, overlaid stereographic *virtual pointers* and *virtual tape measures*[19], *virtual landmarks*[20], and *virtual tethers*[21] for telerobotic control. Current efforts are focused on applying more advanced stereographic tools for achieving *virtual control*[22,23,24] of telerobotic systems, through the use of overlaid *virtual robot simulations*, *virtual encapsulators*, and *virtual bounding planes*, etc.[25]

4 . THE GENERAL CASE: MIXED REALITY ENVIRONMENTS

Thus far we have defined the concept of AR within the context of the reality-virtuality continuum, and illustrated two particular subclasses of AR displays. In this section we discuss how Augmented Reality relates to other classes of Mixed Reality displays.

Through the brief discussion in Section 3, it should be evident that the key factors distinguishing see-through and monitor based AR systems go beyond simply whether the display is head mounted or monitor based, which governs the metaphor of whether one is expected to feel egocentrically immersed within one's world or whether is one is to feel that one is exocentrically looking in on that world from the outside. There is also the issue of how much one knows about the world being viewed, which is essential for the conformal mapping needed for useful see-through displays, but much less critical for WoW displays. In addition, there are other, largely perceptual, issues which are a function of the degree to which the fidelity of the 'substratal world' must be maintained. With optical see-through (ST) systems, one has very little latitude, beyond optical distortion, to change the reality of what one observes directly, whereas when video is used as the intermediate medium, the potential for altering that world is much larger.

This leads us back to the concept of the RV continuum, and to the issue of defining the *substratum*:
- Is the environment being observed principally real, with added computer generated enhancements? or
- Is the surrounding environment principally virtual, but augmented through the use of real (i.e. unmodelled) imaging data?

(Of course, as computer graphic and imaging technologies continue to advance, the day will certainly arrive in which it will not be immediately obvious whether the primary world is real or simulated, a situation corresponding to the centre of the RV continuum in Fig. 1).

The case defined by the second question serves as our working definition of what we term "Augmented Virtuality" (AV), in reference to completely graphic display environments, either completely immersive, partially immersive, or otherwise, to which some amount of (video or texture mapped) 'reality' has been added.[2,13] When this class of displays is extended to include situations in which real objects, such as a user's hand, can be introduced into the otherwise principally graphic world, in order to point at, grab, or somehow otherwise manipulate something in the virtual scene[26,27], the perceptual issues which arise, especially for stereoscopic displays, become quite challenging.[28,29]

In order further to distinguish essential differences and similarities between the various display concepts which we classify as Mixed Reality, it is helpful to make a formal list of these:
1. Monitor-based (non-immersive) AR displays, upon which computer graphic (CG) images are overlaid.
2. Same as 1, but using immersive HMD-based displays, rather than WoW monitors.[§]
3. HMD-based AR systems, incorporating optical see-through (ST).
4. HMD-based AR systems, incorporating video ST.
5. Monitor-based AV systems, with CG world substratum, employing superimposed video reality.
6. Immersive or partially immersive (e.g. large screen display) AV systems, with CG substratum, employing superimposed video or texture mapped reality.
7. Partially immersive AV systems, which allow additional real-object interactions, such as 'reaching in' and 'grabbing' with one's own (real) hand.

It is worth noting that additional classes could have been delineated; however, we are limiting ourselves here to the primary factors characterising the most prominent classes of MR displays. One important distinction which has been conspicuously left out of the above list, for example, is whether or not the systems listed are stereoscopic.

[§] This class refers, for example, to HMD-based AR-enhanced head-slaved video systems for telepresence in remote robotic control,[30] where exact conformal mapping with the observer's surrounding physical environment is, however, not strictly necessary.

A summary of how some of the factors discussed thus far pertain to the seven classes of MR displays listed above is presented in Table 1. The first column encompasses the major distinction separating the left and right portions of Fig. 1, that is, respectively, whether the substratum defining the principal scene being presented derives from a real (R) or a computer generated (CG) world. It says nothing, however, about the hardware used to display that scene to the observer. That distinction is made in the second column, where we immediately note that there is no strict correspondence with column 1. A direct view here refers to the case in which the principal world is viewed directly, through either air or glass (otherwise known as "unmediated reality")[3], whereas for the opposite case of non-direct viewing, the world must be scanned by some means, such as a video camera, laser or ultrasound scanner, etc., and then resynthesised, or reconstructed, by means of some medium such as a video or computer monitor.[2]

The question addressed in the fourth column, of whether or not a strict conformal mapping is necessary, is closely related to the exocentric / egocentric distinction shown in the third column. However, whereas all systems in which conformal mapping is required must necessarily be egocentric, the converse is not the case for all egocentric systems.

Class of MR System	Real (R) or CG world?	Direct (D) or Scanned (S) view of substrate ?	Exocentric (EX) or Egocentric (EG) Reference?	Conformal Mapping (1:1), or not (1:k) ?
1. Monitor-based video, with CG overlays	R	S	EX	1:k
2. HMD-based video, with CG overlays	R	S	EG	1:k
3. HMD-based optical ST, with CG overlays	R	D	EG	1:1
4. HMD-based video ST, with CG overlays	R	S	EG	1:1
5. Monitor/CG-world, with video overlays	CG	S	EX	1:k
6. HMD/CG-world, with video overlays	CG	S	EG	1:k
7. CG-based world, with real object intervention	CG	D, S	EG	1:1

Table 1: Some major differences between classes of Mixed Reality (MR) displays.

Perhaps the most important message to derive from Table 1 is that no two rows are identical. Consequently, even though limited scope comparisons between pairs of Mixed Reality displays may yield simple distinctions, a global framework for categorising all possible MR displays is much more complex. This observation underscores the need for an efficient taxonomy of MR displays, both for identifying the

key dimensions that can be used parsimoniously to distinguish all candidate systems and for serving as a framework for identifying common research issues that span the breadth of such displays.

5. A TAXONOMY FOR MIXING REAL AND VIRTUAL WORLDS

The first question to be answered in setting up the taxonomy is why the continuum presented in Fig. 1 is not sufficient for our purposes as is, since it clearly defines the concept of AR displays and distinguishes these from the general class of AV displays, within the general framework of Mixed Reality. From the preceding section, however, it should be clear that, even though the RV continuum spans the space of MR options, its one dimensionality is too simple to highlight the various factors which distinguish one AR/AV system from another.

What is needed, rather, is to create a taxonomy with which the principal environment, or substrate, of different AR/AV systems can be depicted in terms of a (minimal) multidimensional hyperspace. Three (but not the only three) important properties of this hyperspace are evident from the discussion in this paper:

- *Reality*; that is, some environments are primarily virtual, in the sense that they have been created artificially, by computer, while others are primarily real.
- *Immersion*; that is, virtual and real environments can each be displayed without the need for the observer to be completely immersed within them.
- *Directness;* that is, whether primary world objects are viewed directly or by means of some electronic synthesis process.

The three dimensional taxonomy which we propose for mixing real and virtual worlds is based on these three factors. A detailed discussion can be found elsewhere[2]; we limit ourselves here to a summary of the main points.

5.1 Extent of World Knowledge

Many of the discussions of reality and virtuality in the literature centre strictly on virtual environments, typically on the means by which one can depict virtual objects using graphic techniques which are of sufficiently high quality to make those virtual objects appear 'real'.[3] Others deal with the distinction more multidimensionally, by focusing also on the factors which allow one to feel present within and influence a remote world.[31,32,33] One important related consideration, which is of great practical importance for determining the operational capabilities of many display systems but is nevertheless often overlooked, is the Extent of World Knowledge (EWK). In simple terms, EWK defines how much we actually know about objects and the world in which they are displayed.

The EWK dimension is depicted in Fig. 2. At one extreme, on the left, is the case in which nothing is known about the (remote) world being displayed. This end of the continuum characterises unmodelled data obtained from images of scenes that have been 'blindly' scanned and synthesised via non-direct viewing. It also pertains to directly viewed real objects in see-through displays. In the former instance, even though such images can be digitally enhanced, no information is contained within the knowledge base about the *contents* of those images. The unmodelled world extremum describes for example the class of video displays found in most current telemanipulation systems, especially those that must be operated in such unstructured environments as underwater exploration and military operations. The other end of the EWK dimension, the completely modelled world, defines the conditions necessary for displaying a totally virtual world, in the 'conventional' sense of VR, which can be created only when the computer has complete knowledge about each object in that world, its location within that world, the location and viewpoint of the observer within that world and, when relevant, the viewer's attempts to change that world by manipulating objects within it.
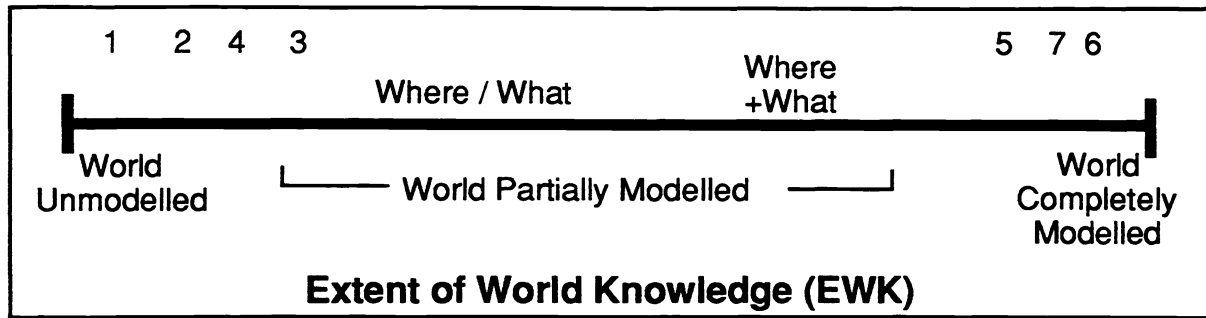
Figure 2: Extent of World Knowledge (EWK) dimension

Although both extrema occur frequently, the region covering all cases in between governs the extent to which real and virtual objects can be merged within the same display. In Figure 2, three types of subcases are shown: Where, What, and Where + What. Whereas in some instances we may know *where* an object is located, but not know what it is, in others we may know *what* object is in the scene, but not where it is. And in some cases we may have both 'where' and 'what' information about some objects, but not about others. This is to be contrasted with the completely unmodelled case, in which we have no 'where' or 'what' information at all, as well as with the completely modelled case, in which we possess all 'where' and 'what' information.

The practical importance of these considerations to Augmented Reality systems is great. Usually it is technically quite simple to superimpose an arbitrary graphic image onto a real-world scene, which is either directly (optical-ST) or non-directly (video-ST) viewed. However, for practical purposes, to make the graphical image appear in its *proper* place, for example as a wireframe outline superimposed on top of a corresponding real-world object, it is necessary to know *exactly* where that real-world object is (within an acceptable margin of error) and what its orientation and dimensions are. For stereoscopic systems, this constraint is even more critical. This is no simple matter, especially if we are dealing with unstructured, and completely unmodelled environments. Conversely, if we presume that we do know what and where all objects are in the displayed world, one must question whether an augmented reality display is really the most useful one, or whether a completely virtual environment might not be better.

In our laboratory we view Extent of World Knowledge considerations not as a limitation of Augmented Reality technology, but in fact as one of its strengths. That is, rather than succumbing to the constraints of requiring exact information in order to place CG objects within an unmodelled (stereo)video scene, we instead use human perception to "close the loop" and exploit the virtual interactive tools provided by our ARGOS system, such as the virtual stereographic pointer[19], to make quantitative measurements of the observed real world. With each measurement that is made, we are therefore effectively increasing our knowledge of that world, and thereby migrating away from the left hand side of the EWK axis, as we gradually build up a *partial model* of that world. In our prototype *virtual control* system[24,25], we also create partial world models, by interactively teaching a telemanipulator important three dimensional information about volumetrically defined regions into which it must not stray, objects with which it must not collide, bounds which it is prohibited to exceed, etc.

To illustrate how the EWK dimension might relate to the other classes of MR displays listed above, these have been indicated across the top of Fig. 2. Although the general grouping of Classes 1-4 towards the left and Classes 5-7 towards the right is reliable, it must be kept in mind that this ordering is very approximate, not only in an absolute sense, but also ordinally. That is, as discussed above, by using AR to interactively specify object locations, progressive increases in world knowledge can be obtained, so that a Class 1 display might be moved significantly to the right in the figure. Similarly, in order for a Class 3 display to

provide effective conformal overlays, for example, a significant amount of world knowledge is necessary, which would also move Class 3 rightwards in the figure.

5.2 Reproduction Fidelity

The remaining two dimensions also deal with the issue of realism in MR displays, but in different ways: in terms of image quality and in terms of immersion, or presence, within the display. It is interesting to note that this approach is somewhat different from those taken by others. Both Sheridan's[31] and Robinett's[33] taxonomies, for example, focus on the feeling of presence as the ultimate goal. This is consistent as well with the progression in "realspace imaging" technology outlined in Naimark's taxonomy[3], towards increasingly more realistic displays which eventually make one feel that one is participating in "unmediated reality". In our taxonomy we purposely separate these two dimensions, however, in recognition of the practical usefulness of some high quality visual displays which nevertheless do not attempt to make one feel *within* the remote environment (e.g. Class 1), as well as some see-through display situations in which the viewer in fact is already physically immersed within the actual real environment but may be provided with only relatively low quality graphical aids (e.g. Classes 3 and 4).

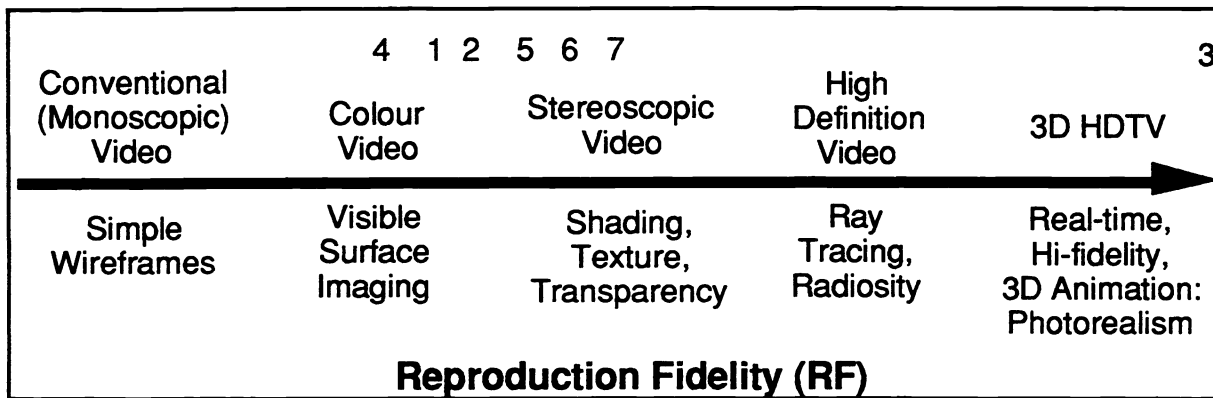|	4 1 2	5 6 7		3
Conventional (Monoscopic) Video	Colour Video	Stereoscopic Video	High Definition Video	3D HDTV
Simple Wireframes	Visible Surface Imaging	Shading, Texture, Transparency	Ray Tracing, Radiosity	Real-time, Hi-fidelity, 3D Animation: Photorealism
Reproduction Fidelity (RF)				

Figure 3: Reproduction Fidelity (RF) dimension.

The elements of the Reproduction Fidelity (RF) dimension are illustrated in Fig. 3. The term "Reproduction Fidelity" refers to the relative quality with which the synthesising display is able to reproduce the actual or intended images of the objects being displayed. It is important to point out that this figure is actually a gross simplification of a complex topic, and in fact lumps together several classes of factors, such as display hardware, signal processing and graphic rendering techniques, etc., each of which could in turn be broken down into its own taxonomic elements.

In terms of the present discussion, it is important to realise that the RF dimension pertains to reproduction fidelity of *both* real and virtual objects. The reason for this is not only because many of the hardware issues, such as display definition, are related. Even though the simplest graphic displays of virtual objects and the most basic video images of real objects are quite distinct, the converse is not true for the high fidelity extremum. In Fig. 3 the ordering above the axis is meant to show a rough progression, mainly in hardware, of video reproduction technology. Below the axis the progression is towards more and more sophisticated computer graphic modelling and rendering techniques. At the right hand side of the figure, the 'ultimate' video display, denoted here as 3D HDTV, might be just as close in quality as the 'ultimate' graphic rendering, denoted here as "real-time, hi-fidelity 3D animation", both of which approach photorealism, or even direct viewing of the real world.

The importance of Reproduction Fidelity for the MR taxonomy goes beyond having world knowledge for the purpose of superimposing modelled data onto unmodelled data images, or vice versa, as discussed

above. The ultimate ability to blend CG images into real-world images or, alternatively, to overlap CG and real images while keeping them distinct, will depend greatly on the fidelity both of the principal environment and of the overlaid objects. (It will also depend on whether CG images must be blended with directly viewed real objects, or with non-directly viewed images of real objects.) Although it is difficult to distinguish the seven MR display classes as clearly as in Fig. 2, since the location of any one system on this axis will depend on the particular technical implementation, a (very) approximate ordering is nevertheless indicated on Fig. 3. Note that Class 3 has been placed far to the right in the figure, since optical see-through represents the ultimate in fidelity: directly viewed reality.

5.3   Extent of Presence Metaphor

The third dimension in our taxonomy, depicted in Fig. 4, is the Extent of Presence Metaphor (EPM) axis, that is, the extent to which the observer is intended to feel "present" within the displayed scene. In including this dimension we recognise the fact that Mixed Reality displays can range from immersive environments, with a strong presence metaphor, such as Class 2, 3, 4 and 7 displays, to important exocentric Class 1 AR- and Class 5 AV-type displays.
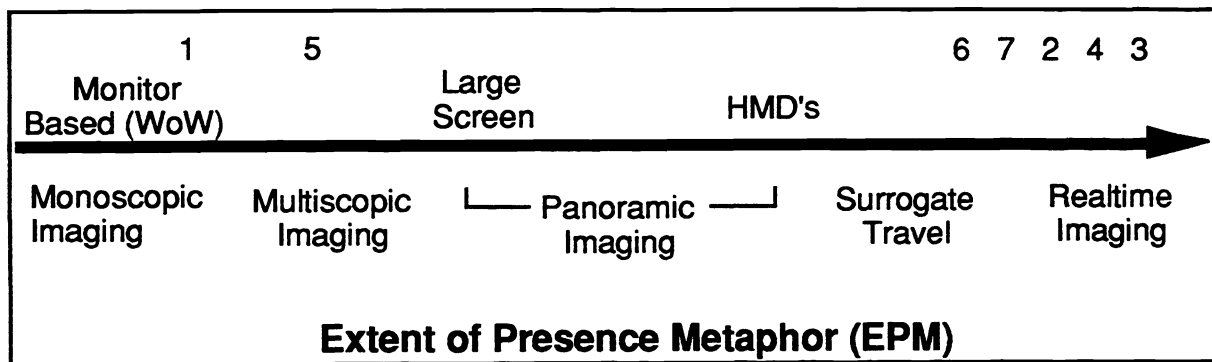

Figure 4: Extent of Presence Metaphor (EPM) dimension

In some sense the EPM axis is not entirely orthogonal to the RF axis, since each dimension independently tends towards an extremum which ideally is indistinguishable from viewing reality directly. In the case of EPM the axis spans a range of cases extending from the metaphor by which the observer peers from outside into the world from a single fixed monoscopic viewpoint, up to the metaphor of "realtime imaging", by which the observer's sensations are ideally no different from those of unmediated reality.[3] Above the axis in Fig. 4 is shown the progression of display media necessary for realising the corresponding presence metaphors depicted below the axis.

The importance of the EPM dimension in our MR taxonomy is principally as a means of classifying exocentric vs egocentric differences between MR classes, while taking into account the need for strict conformality of the mapping of augmentation onto the background environment, as shown in Table 1. As indicated in Fig. 4, a generalised ordinal ranking of the display classes listed above might have Class 1 displays situated towards the left of the EPM axis, followed by Class 5 displays (which more readily permit multiscopic imaging), and Classes 6, 7, 2, 4 and 3, all of which are based on an egocentric metaphor, towards the right.

6. CONCLUSION

In this paper we have discussed Augmented Reality (AR) displays in a general sense, within the context of a Reality-Virtuality (RV) continuum, which ranges from completely real environments to completely virtual environments, and encompasses a large class of displays which we term "Mixed Reality" (MR). Analogous, but antithetical, to AR within the class of MR displays are Augmented Virtuality (AV) displays, into whose properties we have not delved deeply in this paper.

MR displays are defined primarily by means of seven (non-exhaustive) examples of existing display concepts in which real objects and virtual objects are displayed together. Rather than relying on the comparatively obvious distinctions between the terms "real" and "virtual", we have probed deeper, and posited some of the essential factors which distinguish different Mixed Reality display systems from each other: Extent of World Knowledge (EWK), Reproduction Fidelity (RF) and Extent of Presence Metaphor (EPM). One of our main objectives in presenting this taxonomy has been to clarify a number of terminology issues, in order that apparently unrelated developments being carried out by, among others, VR developers, computer scientists and (tele)robotics engineers can now be placed within a single framework, depicted in Fig. 5, which will allow comparison of the essential similarities and differences between various research endeavours.
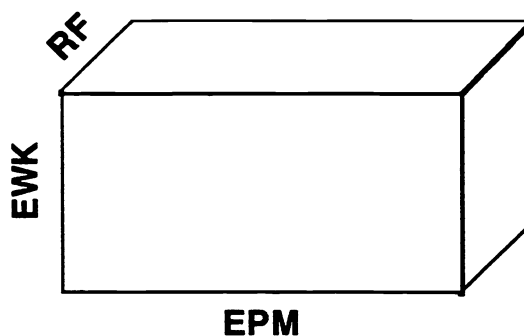

Figure 5: Proposed three dimensional taxonomic framework for classifying MR displays.

7. REFERENCES

1. H. Das (ed). *Telemanipulator and Telepresence Technologies*, SPIE Vol. 2351 Bellingham, WA, 1994.
2. P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays", *IEICE (Institute of Electronics, Information and Communication Engineers) Transactions on Information and Systems*, Special issue on Networked Reality, Dec. 1994.
3. M. Naimark, "Elements of realspace imaging: A proposed taxonomy", Proc. SPIE Vol. 1457, *Stereoscopic Displays and Applications II.*, 1991.
4. T.P. Caudell and D.W. Mizell, "Augmented reality: An application of heads-up display technology to manual manufacturing processes". *Proc. IEEE Hawaii International Conf. on Systems Sciences*, 1992.
5. A.L. Janin, D.W. Mizell, and T.P. Caudell, "Calibration of head-mounted displays for augmented reality". *Proc. IEEE Virtual Reality International Symposium (VRAIS'93)*, Seattle, WA, 246-255, 1993.
6. S. Feiner, B. MacIntyre, and D. Seligmann, "Knowledge-based augmented reality". *Communications of the ACM, 36(7)*, 52-62, 1993.
7. M. Bajura, H. Fuchs, and R. Ohbuchi, "Merging virtual objects with the real world: Seeing ultrasound imagery within the patient". *Computer Graphics, 26(2)*, 1992.
8. H. Fuchs, M. Bajura, and R. Ohbuchi, "Merging virtual objects with the real world: Seeing ultrasound imagery within the patient." *Video Proceedings of IEEE Virtual Reality International Symposium (VRAIS'93)*, Seattle, WA, 1993.
9. J.P. Rolland, R.L. Holloway and H. Fuchs, "Comparison of optical and video see-through head-mounted displays". *Proc. SPIE Vol. 2351-35, Telemanipulator and Telepresence Technologies*, 1994.

10. S.R. Ellis and R.J. Bucher, "Distance perception of stereoscopically presented virtual objects optically superimposed on physical objects by a head-mounted see-through display". *Proc. 38th Annual Meeting of Human Factors and Ergonomics Society*, Nashville, TE, 1300-1304, 1994.

11. E.K. Edwards, J.P. Rolland & K.P. Keller, "Video see-through design for merging of real and virtual environments". *Proc. IEEE Virtual Reality International Symp. (VRAIS'93)*, Seattle, WA, 223-233, 1993.

12. M. Tani, K. Yamashi, K. Tanikohsi, M. Futakawa, and S. Tanifuji. "Object-oriented video: Interaction with real-world objects through live video". *Proc. CHI '92 Conf on Human Factors in Computing Systems*, 593-598, 1992.

13. P.J. Metzger, "Adding reality to the virtual". *Proc. IEEE Virtual Reality International Symposium (VRAIS'93)*, Seattle, WA, 7-13, 1993.

14. L.B. Rosenberg, "Virtual fixtures: Perceptual tools for telerobotic manipulation", *Proc. IEEE Virtual Reality International Symposium (VRAIS'93)*, Seattle, WA, 76-82, 1993.

15. P. Milgram, D. Drascic and J.J. Grodski, "Enhancement of 3-D video displays by means of superimposed stereographics". *Proceedings of Human Factors Society 35th Annual Meeting*, San Francisco, 1457-1461, 1991.

16. D. Lion, C. Rosenberg and W. Barfield, "Overlaying three-dimensional computer graphics with stereoscopic live motion video: Applications for virtual environments". *SID Conf. Proceedings*, 1993.

17. D.J. Wenzel, S.B. Seida, and V.R. Sturdivant, "Telerobot control using enhanced stereo viewing". *Proc. SPIE Vol. 2351-29, Telemanipulator and Telepresence Technologies* (this volume), 1994.

18. D. Drascic, J.J. Grodski, P. Milgram, K. Ruffo, P. Wong and S. Zhai, "ARGOS: A Display System for Augmenting Reality", *ACM SIGGRAPH Tech Video Review, Vol 88: InterCHI '93 Conf on Human Factors in Computing Systems*, (Abstract in *Proceedings of InterCHI'93*, p 521), Amsterdam, April 1993.

19. D. Drascic & P. Milgram, "Positioning accuracy of a virtual stereographic pointer in a real stereoscopic video world", *Proc. SPIE Vol. 1457, Stereoscopic Displays & Applications II*, 302-313, 1991.

20. P. Milgram and M. Krüger, "Adaptation effects in stereo due to on-line changes in camera configuration", *Proc. SPIE Vol. 1669-13, Stereoscopic Displays and Applications III*, 1992.

21. K. Ruffo and P. Milgram, "Effect of stereographic + stereovideo "tether" enhancement for a peg-in-hole task", *Proceedings of Annual Conference of IEEE Systems, Man & Cybernetics Society*, Oct. 1992.

22. S. Zhai and P. Milgram, "A telerobotic virtual control system", *Proc. SPIE, Vol. 1612, Cooperative Intelligent Robotics in Space II*, Nov. 1991.

23. S. Zhai and P. Milgram, "Human-robot synergism and virtual telerobotic control", *Proc. Annual Meeting of Human Factors Association of Canada*, Oct. 1992.

24. P. Milgram, S. Zhai and D. Drascic, "Applications of augmented reality for human-robot communication", *Proc. IEEE/RSJ Intn'l Conf. on Intelligent Robots and Systems (IROS'93)*, Yokohama, Japan, July 1993.

25. A. Rastogi and P. Milgram, D. Drascic, and J.J. Grodski, "Virtual telerobotic control", *Proc. DND Knowledge-Based Systems & Robotics Workshop*, Ottawa, Nov. 1993.

26. H. Takemura and F. Kishino, "Cooperative work environment using virtual workspace". *Proc. Computer Supported Cooperative Work (CSCW'92)*, 226-232, 1992.

27. M. Kaneko, F. Kishino, K. Shimamura and H. Harashima, "Toward the new era of visual communication". *IEICE Transactions on Communications*, Vol. E76-B(6), 577-591, June 1993.

28. A. Utsumi, P. Milgram, H. Takemura and Kishino, "Investigation of errors in perception of stereoscopically presented virtual object locations in real display space", *Proc. 38th Annual Meeting of Human Factors and Ergonomics Society*, Boston, Oct. 1994.

29. A. Utsumi, P. Milgram, H. Takemura & F. Kishino, "Effects of fuzziness in perception of stereoscopically presented virtual object locations", *Proc. SPIE Vol. 2351-39, Telemanipulator and Telepresence Technologies*, 1994.

30. S. Tachi, "Virtual reality and tele-existence – Harmonious integration of synthesized worlds and the real world", *Proc. Industrial Virtual Reality Conf. (IVR'93)*, Makuhari Messe, Japan, June 23-25, 1993.

31. T.B. Sheridan, "Musings on telepresence and virtual reality". *Presence 1(1)*, 120-126, 1992.

32. D. Zeltzer, "Autonomy, interaction and presence". *Presence 1(1)*, 127-132, 1992.

33. W. Robinett, "Synthetic experience: A proposed taxonomy". *Presence, 1(2)*, 229-247, 1992.