Chapter 1

A Taxonomy of Real and Virtual World Display Integration

Paul Milgram
Herman Colquhoun Jr.

University of Toronto, Canada

1.1 Definition of Mixed Reality

Our primary objective in this paper is to present a number of fundamental display integration and orientation issues related to the nascent field of Mixed Reality. Our approach is motivated first by the need for a more encompassing term to supplement the existing definition of Augmented Reality (AR), which leads us to propose definitions of the associated concepts of Augmented Virtuality (AV) and then Mixed Reality (MR). Following our discussion of the breadth of Mixed Reality displays in Section 1.1, we discuss the associated issues of viewpoint centricity and control-display mapping in Section 1.2. Finally, in Section 1.3, we present a taxonomy which we hope will be useful for differentiating between several of the issues raised with regard to the different classes of Mixed Reality display systems.

1.1.1 Augmented Reality

An examination of current literature in which the term Augmented Reality (AR) appears will reveal two classes of definition, distinguished from each other in terms of breadth. Most common appear to be those for which AR refers narrowly to the class of display systems comprising some kind of *head-mounted display (HMD)* or *head-up display (HUD)*. In the case of HMD's, the viewer observes a direct *"see-through" view* of the real world, either optically or via video coupling, upon which is superimposed computer generated graphics. A number of reviews which focus on this class of displays have been written, including that by Azuma [1] and that by Fuchs and Ackerman [2] in the present volume. Some of the prominent examples of such displays include systems for assisting in manufacturing [3]–[5] and in medicine [6]. It is important to note that head-up displays (HUD's), which have existed in primarily military aviation environments for several years, clearly fall within the realm of see-through AR as well, in the sense that graphic information is superimposed upon the pilot's direct view of the outside real world [1] [7]. More recently, the same HMD concept has been proposed for use also by combat soldiers on the ground [8] [9].

The second, broader class of definitions in the literature relaxes the constraint of needing the equivalent of a HMD and covers "any case in which an otherwise real environment is 'augmented' by means of virtual (computer graphic) objects" [10] [11], thereby encompassing both large screen and monitor based displays as well. Examples conforming to this broader definition of AR include applications in robotics [12] and medicine [13] [14].

In a sense a third, even broader class of AR displays has been proposed by some in the literature, encompassing those cases involving *any* mixture of real and virtual environments. Consistent with this interpretation, Azuma, in his earlier survey of Augmented Reality, referred to AR as "a variation on Virtual Environments that combines virtual and real" [15]. He later refined this, however, to comprise *any* system that "1) combines real and virtual, 2) is interactive in real time, and 3) is registered in three dimensions" [1]. As we discuss in the following section, extending the realm of AR in this direction brings to light an important issue – whether it is in fact *reality* or *virtuality* which is being enhanced – and, together with this, the need for a broader, more comprehensive set of definitions. It is our contention that any useful definition of AR should definitely encompass the first two classes of displays mentioned here, but that a different term is needed to account for the third class.

1.1.2 Reality vs. Virtuality

Before proceeding, it is imperative that we clarify what we mean by the key terms "real" and "virtual" environments. Following the approach proposed in [10] [11], we first contend that, although both purely real environments (RE's) and virtual environments (VE's) certainly do exist as separate entities, they are not to be considered simply as alternatives to each other, but rather as poles lying at opposite ends of a *Reality-Virtuality (RV) continuum*, as shown in Figure 1.1. The location of any environment, or "world", along this continuum coincides with its location along a parallel *Extent of World Knowledge (EWK) continuum*. In using the latter term, we are referring to the extent of knowledge present within the computer about the world being presented. Figure 1.1 illustrates the parallel nature of the two concepts.

As indicated in the figure, at the right end of the RV continuum are virtual environments, which must necessarily be *completely modelled* in order to be rendered. At the opposite extreme we are taking real environments to be representations of a world, or region, which are *completely unmodelled*. In using the term "model", we are limiting ourselves to quantitative computer models. Thus, in relation to real environments, which have not been modelled, we refer to situations in which the computer does not possess, or does not attribute meaning to, any information about the content of an image.[1] RE's therefore encompass any kind of sampled image data, and include as the primary example video, but also photographic images (visible or infrared), radar, X-ray and ultrasound, as well as laser scanned data (both range and light intensity data). Note that we are not necessarily limiting ourselves in our definition to two dimensional (2D) data, but we include, especially with respect to the latter example of laser range data, also three dimensional (3D) sampled images.

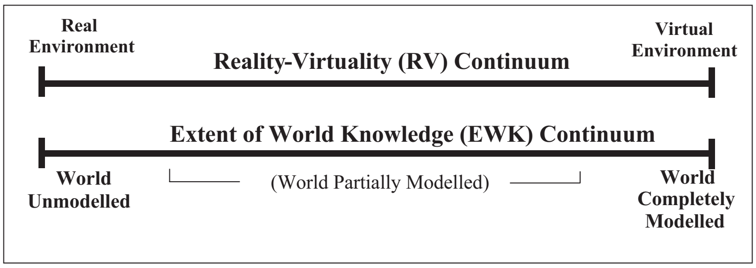


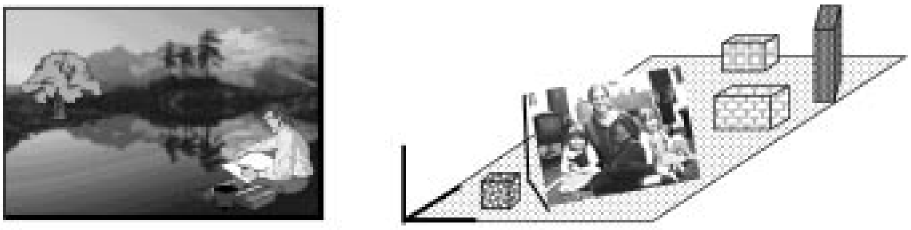
Figure 1.1 Reality-Virtuality (RV) Continuum, in parallel with Extent of World Knowledge (EWK) Continuum.

Returning to the issue of mixing, or combining, real and graphic images, as is the case with AR, this necessarily brings us somewhere towards the middle of the RV spectrum, and concurrently to some region in the middle of the EWK continuum, in which the world is *partially modelled*. This means, for example, that we might know the location of some objects, but nothing about the objects themselves, or we might have elaborate geometric models of some objects, but not know anything about where they are, that is, how they relate to surrounding regions of the image.

As we venture away from the poles of the RV continuum towards the centre, we also eventually begin to encounter the problem of deciding whether in fact what we are doing is augmenting a real world with virtual graphic objects, or whether we are modifying a virtual environment by augmenting it with real data. This issue is discussed further in Section 1.1.4. To the extent that these two cases can be distinguished from each other, in the meantime, it has been proposed that they be labelled Augmented Reality (AR) and Augmented Virtuality (AV) respectively [10] [11].

These concepts are illustrated in Figure 1.2. On the left hand side we have an example of Augmented Reality: a scene comprising a photograph – i.e. a *real image* – of a mountain lake, upon which have been superimposed two computer generated –

[1] It is interesting to point out that Durlach and colleagues have made a similar observation in their recent book on Virtual Reality, a portion of which we cite here for convenience [36], pp.59–60): "When a virtual environment application requires a replica of a real environment, it is generally considered preferable to map the real environment rather than build a model of it. Active mapping techniques, such as scanning laser range finders and light stripes, are used to make three-dimensional measurements directly." We must point out that those authors' use of the term "virtual environment" (VE) does not conform with ours, however, since such a collection of sampled data would continue to be termed a "real environment" (RE) by our definition.



(a) Augmented Reality (AR) (b) Augmented Virtuality (AV)

Figure 1.2 Illustrations of (a) Augmented Reality (AR) and (b) Augmented Virtuality (AV).

i.e. *virtual* – images, of a virtual artist on one side of the lake sketching a virtual tree on the other side. Although we, the viewers of the picture, comprehend the content easily, the computer is presumed to have no model whatsoever of the content of the photographic image, and thus, to the computer, the location of the virtual images relative to the real image is meaningless.

The converse case, an example of Augmented Virtuality, is illustrated schematically in Figure 1.2b. Here we see a completely modelled (3D) world, comprising a series of virtual 3D blocks located on a virtual plane. In order to draw these objects, the computer must have a model of all of their dimensions and locations. In addition, a photograph of a group of people, comprising real data, has been added, at a specific location. Although the computer must have knowledge about *where* the real image (photograph) has been placed, we can not assume that any knowledge is held about the *content* of that image.

Figure 1.2 is obviously contrived, to facilitate comprehension of the distinction between the concept of an underlying real world and an underlying virtual world. To show how these concepts manifest themselves in relation to actual practical applications, we present two more examples. The first example, of AR, is shown in Figure 1.3, which is from our own laboratory and illustrates ARTEMIS, our Augmented Reality TEleManipulation Interface System [16] [17]. In the figure we see a real robot situated within a real environment. Although the real environment is completely unmodelled, we do possess a model of the real robot, registered to real-world coordinates. This permits us to superimpose a modelled stereoscopically presented virtual robot on top of the real robot, as depicted in the figure. As described elsewhere [16] [17], this set-up enables an operator to pick up and deposit the real objects depicted in the image – even though they are not modelled – simply by aligning the virtual end effector in the 3D work space with the objects to be manipulated and then transmitting the robot joint coordinates to the remote site at the appropriate moment.

As a converse example, of Augmented Virtuality this time, we present Figure 1.4, which is a screen dump of an image produced with Cyberworld® software (www.cyberworldcorp.com), a commercial product which is used to create realistic "3D web pages". To generate the Christmas scene shown here, a 3D virtual world has been created, comprising a large public square. A miniature plan view of the square is shown at the bottom right corner. However, the buildings, the Christmas tree, Santa Claus, the carollers and all of the other objects in the picture are superimposed 2D photographic images, but with known locations in the 3D virtual world.

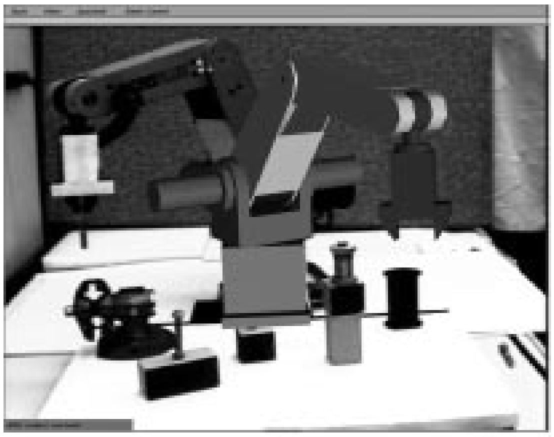


Figure 1.3 ARTEMIS Augmented Reality system: Virtual robot (in this case polygon filled) is overlaid onto a modelled real robot, within an otherwise completely unmodelled world (see color pages).

Figure 1.4 Example of Augmented Virtuality (AV): Superposition of real images and texture mapping onto a virtual 3D world (see color pages). (Composed using Cyberworld® software.)

1.1.3 Exploring the Reality-Virtuality Continuum

Thus far, we have presented only illustrations which purposely emphasise the major distinctions between fundamentally opposing RV mixtures. However, although the two terms Augmented Reality and Augmented Virtuality support these distinctions, in the ensuing discussion we show that it is not always as simple as in the preceding examples to distinguish between AR and AV. We shall thus argue that the more enveloping term, Mixed Reality, becomes necessary, to encompass in a less constrained way all mixtures between the poles of the RV continuum.

To this end, we present Figure 1.5, which illustrates schematically a selection of image composites that could be encountered when one traverses the RV Continuum. On a global level, Figure 1.5 corresponds to the same left-to-right RV Continuum shown along the top part of Figure 1.1. The difference here, however, is that Figure 1.5 highlights the variety of ways in which the real components (R) and the virtual components (V) of an image may be mixed. In terms of our earlier examples, for instance, Figure 1.3 and the left hand side AR example in Figure 1.2a could be considered to correspond to Block #8 or 9 in Figure 1.5, in the sense that we have a predominantly real environment, or background, with "a few" virtual objects superimposed. The AV example in Figure 1.2b and in Figure 1.4, furthermore, would correspond here to Blocks #4 or 6 or 7, for analogous reasons.

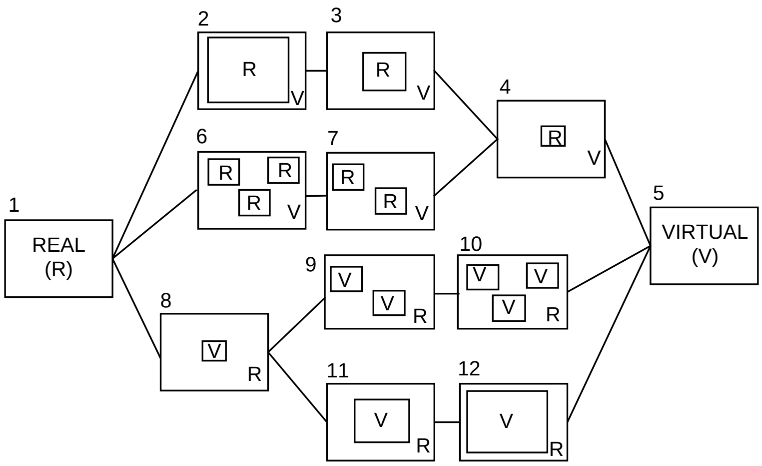


Figure 1.5 Mixed Reality combination space.

In presenting Figure 1.5, we were prompted by a number of cases in the literature for which it is not always obvious whether the primary environment, or "substratum", is real or virtual. One good example of this is the Peloton sports simulator described by Carraro and colleagues [18]. Peloton is a bicycle simulator which simulates a virtual road course for walking, running and bicycle riding. Users stand on a treadmill and locomote through an environment which comprises a computer generated (virtual) surround, in the middle of which is placed a video window, texture mapped onto a large rectangle, or "video screen". In terms of Figure 1.5, the Peloton system therefore corresponds to Block #3. In the SIGGRAPH reference cited [18], the authors describe their simulation of a bicycle path through New York's Central Park, with emphasis placed on their method of blending the internal (real) video window with the surrounding virtual graphics window.

In the following we present an analogous case, but extend the concept somewhat. What we shall do is illustrate what a journey along the RV continuum might look

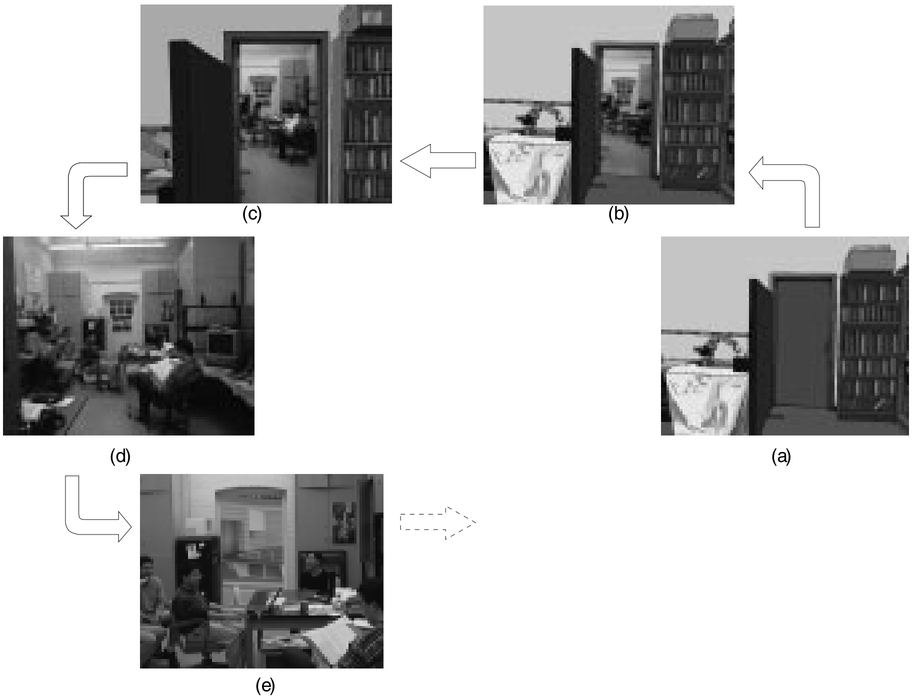


Figure 1.6 University of Toronto ETC Lab, as experienced through a journey along the RV continuum (see color pages).

like, as we travel along the trajectory 5-4-2-1-8 in Figure 1.5. Our point of departure is illustrated in Figure 1.6a, which shows a virtual model of part of our laboratory at the University of Toronto ETC Lab. Because this is a virtual model, the image in Figure 1.6a corresponds to Block #5 above.

To continue the journey, we open the door in Figure 1.6a and project a real image of the adjoining room onto the doorway, as shown in Figure 1.6b. This picture therefore corresponds to Block #4 of Figure 1.5. As we approach the real portal framed by the doorway, in Figure 1.6c, the composite image becomes proportionately more real and less virtual, corresponding to Block #2 in Figure 1.5. Finally, as we advance all the way through the real portal, in Figure 1.6d, we enter into a completely real environment, corresponding to the Real extremum in Block #1 at the left side of Figure 1.5.

Another important aspect of the schematic representation of Figure 1.5 is the essentially circular nature of the RV continuum. That is, in the description above, we have traversed the continuum, from right to left, corresponding to a transition from a completely virtual environment to a completely real one. In Figure 1.6 we also illustrate, however, that it is not necessary to travel along the same trajectory in reverse to return to the virtual side. Rather, as illustrated in Figure 1.6e, we have turned the next doorway into a virtual portal, depicting an adjoining virtual conference room. Figure 1.6e therefore corresponds to Block #8 in Figure 1.5.

Clearly, it is possible to continue in this manner, enter completely into the virtual room, and thus traverse the bottom path of Figure 1.5 to get back to Block #5.

1.1.4 Defining the Principal Environment

One consequence of the above discussion is that the distinction between AR and AV illustrated in Figure 1.2 is not necessarily as simple as shown there. In the preceding example, we showed that virtual and real environments can "flow" into each other recursively. A somewhat different illustration of this occurs if we re-examine the cases shown in Figure 1.2. Suppose that we continue to add more and more virtual objects to an AR image such as Figure 1.2a, thereby moving effectively from the case of Block #8 to Block #9 and then to Block #10 in Figure 1.5. Eventually, if the entire visible image, or viewport, consisted of virtual objects, one could argue that we had arrived at the completely virtual case of Block #5 in Figure 1.5. This would be true, however, only if we possess complete quantitative information about how all the various virtual objects relate to each other within the (3D) space of the image. Otherwise, this would not fit our definition of a completely virtual environment, presented above.

Analogously, if we were to commence with an AV image such as Figure 1.2b, add more and more sampled data images to it, until the scene appeared totally real, it would give the impression that we had migrated from Block #4 to Block #7 to Block #6 and finally to Block #1 in Figure 1.5. Figure 1.4 is a good example of what such an image might look like. However, as long as we retain quantitative information about the spatial relationship among the various real image components relative to each other and/or to the underlying virtual world, we could not consider the final image to conform to Block #1, which requires that the world be completely *un*modelled.

The conclusion to be drawn from these examples is that, according to the operational definition proposed in Figure 1.1, it is not necessarily true that an environment is *completely virtual* if all of the component visible objects in it are computer generated, and it is also not necessarily true that a world is *completely real* if all of the visible objects in it are sample data images. Furthermore, determining whether an image should be considered Augmented Reality or Augmented Virtuality is also not necessarily a matter of simply summating the respective areas of real and virtual images in order to determine a "majority" portion of real or virtual. A practical example of an extreme case of AR is that in which an underlying scene is created from 3D range image data, but then has computer generated polygons mapped onto it to create the appearance of a continuous modelled surface rather than discretely sampled points [21] [37]. Analogously, we have seen that an image may be considered AV, for instance, even if essentially all visible elements in it are derived from sampled real data, but have been texture mapped onto a completely modelled underlying virtual world.

1.1.5 Definition of Mixed Reality

We conclude this section by presenting Figure 1.7, which essentially repeats the RV spectrum of Figure 1.1, but with the generic cases of Augmented Reality (AR), Augmented Virtuality (AV) and Mixed Reality (MR) indicated explicitly now. The portions of the illustration corresponding to the terms indicated will thus serve as our definitions of AR, AV and MR. Note that the AR segment of the continuum covers a portion of the RV continuum adjacent to, but excluding, the real environment extreme and, similarly, the AV segment lies adjacent to, but excluding, the virtual environment end. Encompassing both AR and AV, the MR portion of the RV continuum covers essentially the entire breadth of the spectrum, but also excludes the end points. In closing, we reiterate that it is our hope that the terms discussed here will serve a useful purpose in distinguishing the various contexts within which diverse research in the field of Mixed Reality is currently being carried out, even though in practice the distinctions are often not always easily recognised.

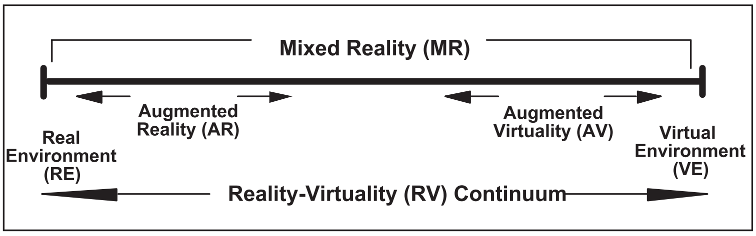


Figure 1.7 Definition of Mixed Reality, within the Context of the RV Continuum.

1.2 Centricity and Control Issues Associated With Mixed Reality

In this section we review some of the issues which may arise when working with complex Mixed Reality worlds, from the point of view of defining an appropriate viewpoint for the observer relative to the objects of interest. One of the fundamental problems which can occur, for example, relates to the fact that virtual environments can generally be presented from any desired viewpoint, whereas the perspective of real data can not ordinarily be changed.[2] This discussion will lead us in a subsequent section to the parallel problems of transitions between virtual and real worlds along the RV continuum and of maintaining suitable control-display relationships when doing so.

1.2.1 Case Study: Remote Mixed Reality Excavation

To illustrate the problem, we shall use as a case study a remotely controlled MR excavator, elements of which have been reported elsewhere, under the acronym VERO (Virtual Environments for Remote Operations) [19] [20]. The original VERO system is currently being followed up by another project dubbed IIRO (Intelligent Interactive Remote Operations), in which we are designing an interface to allow both

[2]Note that this problem pertains not only to 2D images, but also to sampled 3D data images which are ordinarily limited in the amount of viewpoint alteration that is feasible, due to problems of missing data and the occlusion of some objects by others.

teleoperation of the remote excavator, as well as a space robotic arm, and the building of scripts for supervised control of either one of these teleoperators over long time delay communication channels.

A sample display screen related to the IIRO project is given in Figure 1.8, where the image shown comprises a virtual model of an actual excavator situated at a remote work site. The joint angles of the real excavator have been transmitted to the local human interface, thereby permitting repositioning of the model with the proper pose, as would be seen from the remotely located camera external to the excavator. That camera is capable of generating both light intensity images (i.e. video) and 3D laser range images. Because communication with the remote site is poor, we assume that the camera image can not be updated frequently, hence the need for interpolation and enhancement using a virtual model.

In Figure 1.8 a deformable model of terrain has been created by mapping a "skin" of simple polygons onto a field of 3D laser range data obtained from scanning actual terrain elevations [21]. The video image in the centre of the figure has been inserted as a "billboard" display, that is, as a virtual 2D screen upon which have been projected real light intensity data from the laser range camera. Note the barrel in the centre of the image. The portion outside the video window has been created from a prior CAD model and registered to the laser range based image. The portion inside the video window represents a recent update of the remote world. Because the external camera is stationary, the two parts of the same barrel match each other seamlessly.

Four important advantages of this mixed reality image are illustrated in Figure 1.8:

- The light intensity image inside the video window shows much more detail than would be possible with the 3D range image data.
- Because the time taken to update a 3D range image scan of a scene is much longer than a light image (video) scan, the portion inside the window is likely to represent much more recent information.
- The virtual portions of the scene provide the means to simulate intervention operations and interactions, in spite of the large time delay.
- Consistent with the points above, the portion inside the window allows one easily to detect departures from any modelled portions of the scene. In particular, Figure 1.8 shows a second barrel (possibly containing toxic waste, for example) within the video window, which had not been expected and thus had not previously been registered to the image, as was the other barrel.

In terms of the definitions presented in Section 1.1, it is interesting to consider whether Figure 1.8 is an example of AR or of AV. Because the terrain has been recreated from *sampled 3D range data*, it is *unmodelled*, such that the foundation of the image, or substratum, is therefore *real*. By virtue of the addition of the modelled excavator, as well as the surface mesh, or "skin", this figure can therefore be classified as a case of Augmented Reality. Note that such a classification might appear counter-intuitive to many, due to the prominence of the virtual excavator and the "virtual-looking" terrain, such that one might be tempted to classify Figure

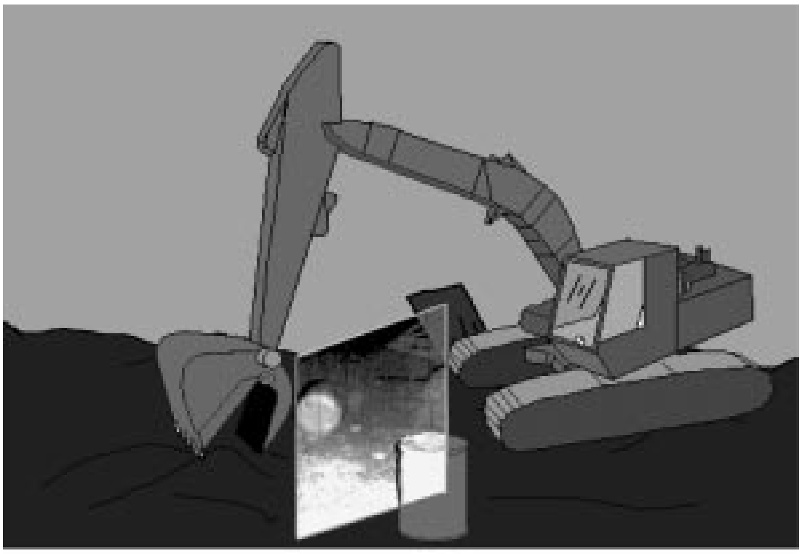


Figure 1.8 Mixed Reality remote excavation example.

1.8 as an AV image, especially if the video window were not present. Conversely in fact, from the point of view of the modelled excavator, one could alternatively contend that the virtual excavator model has been enhanced through addition of the complementary real world data, making the example arguably a case of Augmented Virtuality. As discussed in Section 1.1, therefore, the exact classification could conceivably be presented either way, depending on ones viewpoint, thus providing further justification for the more encompassing term, Mixed Reality.

1.2.2 MR Design Issues

In this section we briefly outline three problems associated with the MR excavator example described above, all related to the issue of defining the human operator's viewpoint relative to the remotely controlled equipment and all of which generalise in some way to a broader class of MR systems.[3] A much more thorough treatment of the considerations outlined here can be found in the various publications of Wickens and his colleagues [22]–[26].

The first problem associated with the IIRO scenario is the "keyhole effect", an excessive narrowing of an observer's field of view, somewhat akin to peeking through a keyhole [27]. Although the view which is generally best for *local guidance and control* of the excavator is the view from the cab, that viewpoint is unfortunately not usually conducive to maintaining global situational awareness. That is, due to the

[3]Note that other technological problems not discussed in detail here must also be overcome, an obvious one being system update rate. Due to the demands of the various levels of detail demanded by the Mixed Reality environment, which includes updating and drawing both modelled objects and real data objects, one of the expected consequences is a slow frame rate, which can clearly have a potentially significant impact on the effectiveness of remote operations.

keyhole effect, the operator is not able to look around the excavator, assess the situation, and plan future operations [23] [24]. This is one of the main reasons for presenting the side view shown in Figure 1.8, rather than a perhaps more conventional through-the-window view.

The second problem is that of distortion. The 2D video image shown inside the window in Figure 1.8 is congruent with the 3D virtual world elements only when the collective viewpoint is located at a station point which corresponds to the location of the remote camera capturing the image. In other words, although the perspective view offered of the virtual screen onto which the video window is mapped is generally correct relative to the surrounding virtual environment, the sampled data content within the video window does not necessarily match that perspective. A view from any angle other than that of the original camera attitude will therefore result in a discontinuity between the virtual environment and the video image contents, resulting in some degree of distortion. It is interesting to note that one of the principal aims of the Peloton system mentioned earlier is to overcome this type of problem [18].

The third problem is that of control reversals, a situation in which perceptual confusion causes an operator to elicit a control action which is opposite in direction to the appropriate action at a particular moment [28]. When presenting the kind of scene shown in Figure 1.8, for example, the potential exists for a mismatch between the outside-in display depicted and conventional inside-out excavator control displays. As discussed further, in Section 1.2.3, this is most likely to have an adverse impact whenever the offset angle between the control action and the observed display is very large.[4]

1.2.3 Effect of Control-Display Compatibility in MR Environments

The principal conclusion to be drawn from the preceding discussion is that, in conjunction with the various advantages associated with mixing real and virtual images in MR, such as improved visualisation of real-world images using RE viewports and added viewpoint flexibility using VE viewports, significant operational problems may result if careful consideration is not given to defining the user's frame of reference relative to the MR display and to the mapping between the user's control actions and the responses of the MR display.

Display centricity

Before delving into these design issues, we must first clarify the meaning of the term "*centricity*", which we use here to refer to the extent to which a human observer's viewpoint is removed from the "*ownship*", that is, from the *nominal viewpoint* with respect to the viewer's own avatar, or own vehicle, or own manipulator within the task space. Referring to the examples presented thus far in the paper, the nominal viewpoint in Figure 1.8 would be within the cab looking out, for local control, and

[4]According to research by Ellis and colleagues, control errors are most likely to increase significantly for angular mismatches between control and display axes in the vicinity of 120 to 180 degrees [29].

from outside looking at the excavator, for global navigation and planning. With respect to Figure 1.4 and Figure 1.6, the nominal viewpoint would be through the eyes of an observer's avatar moving through the various scenes presented. With regard to Figure 1.3, the nominal viewpoint will depend on whether one is controlling all joints of the manipulator independently, in which case the nominal viewpoint might be from behind, consistent with ones view of ones own arm, or whether control is resolved directly through the end-effector, in which case the nominal viewpoint could just as well be from in front of the manipulator, as shown in the figure.

The concept of centricity is generalised schematically in Figure 1.9, which shows yet another continuum, this time a *centricity continuum*, together with illustrations corresponding to different combinations of viewing perspectives of an arbitrary excavator. The point at the left of the figure corresponds to an *egocentric* viewpoint, which occurs when the display viewpoint is *ego-referenced*. On the assumption that the nominal viewpoint of the excavator system in this case is at the driver's seat inside the cab of the excavator and looking out, the egocentric case corresponds to the view which would be seen by that operator. This is depicted in the figure by the camera being mounted within the cab and looking out the window of the cab. The *exocentric* case, in contrast, corresponds to a *world referenced* framework, as shown at the right side of Figure 1.9, where the cameras are fixed with respect to the external world. The prefix "exo" thus refers to the state of being outside of and looking at the "ownship", or "own vehicle", that is, the state of looking at the nominal viewing position.[5]

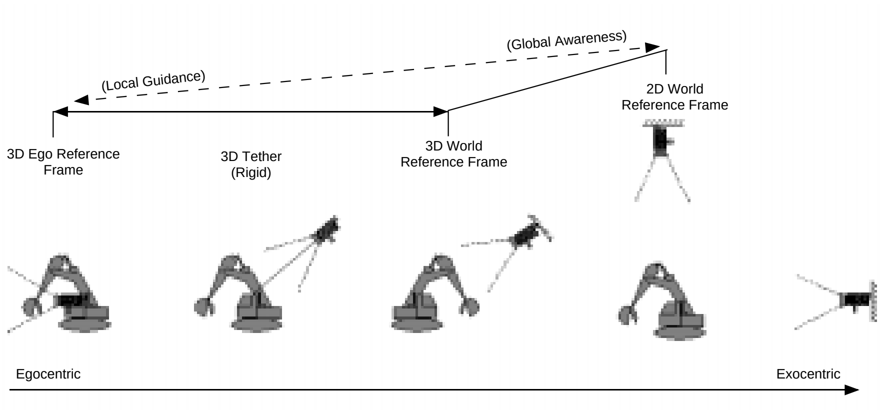


Figure 1.9 Centricity continuum: Illustration of transition from ego- to exocentric viewpoints. (Adapted from [23] [24])

Figure 1.9 conveys the very different effects which would result in the egocentric versus the exocentric cases when the excavator or parts of excavator move. In the egocentric case, movement of the excavator would cause the visual display to change in the same manner as if the operator were inside looking out through the window of

[5]An interesting effort to combine elements of both exocentric and egocentric viewing concurrently, within a single viewport, for large virtual environments, is the work of Kitamura and colleagues [38].

the cab; that is, the *world moves* within the viewport whenever the excavator moves. This view is thus often referred to as an "*inside-out*" view. In the exocentric case, in contrast, we have an "*outside-in*" view and, because the cameras are fixed, we are able to see movements of the "ownship" , as if we were looking at the excavator from above, or from the side, or from some other external viewpoint. In that case, in other words, the *excavator moves* while the rest of the world remains fixed within the viewport.

As with the other continua discussed earlier, the centricity continuum also encompasses a variety of interesting intermediate cases. Wickens and his colleagues have treated this topic very thoroughly, together with the implications of centricity considerations with respect to interface design, most prominently in relation to aviation displays [23] [24] [26], but also as it relates to scientific visualisation [30]. One of the important metaphors introduced by Wickens and colleagues in this context is that of a *tether* joining the virtual camera with the nominal viewpoint, to cover the range of cases along the centricity continuum between the ego-referenced and world-referenced extrema. In Figure 1.9 this is depicted as a rigid tether. The reason for using the tether metaphor to represent intermediate cases between pure egocentric and pure exocentric, as indicated in Figure 1.9, is that movement of the excavator will cause corresponding movement of the world as the camera is dragged along by the tether, as with an egocentric (inside-out) display; however, the observer will also have a more encompassing view of his/her own excavator (or aeroplane or automobile, as the case may be) and its surroundings, as with an exocentric (outside-in) display.

It is interesting to note that, although the concept of exocentricity and a world-referenced frame are highly related, they are not equivalent concepts. This is because world referencing is essentially an absolute concept, relating to *whether* the (virtual) camera is fixed within the world, while exocentricity can be considered a relative concept, relating to *where* the virtual camera is fixed relative to the nominal viewpoint. This becomes evident as one extends the length of the tether. As the view corresponding to the position of the (virtual) camera moves farther and farther away from the nominal viewpoint within the "ownship", one gets the sense of being more and more outside of, or exocentric to, the vehicle. Using this analogy, therefore, the location of the observer along the ego-exo centricity continuum can be considered to correspond to the effective length of the tether.

Control-display mapping

In conjunction with establishing the user's viewpoint, it is critical to take into account the mapping between that viewpoint and the user's ability to manipulate objects when designing a MR system. A practical discussion of these considerations is provided in the following section, based in large part on earlier survey literature [31] [32], primarily by Zhai [33]. In the present section we define some of the basic concepts.

As illustrated in Figure 1.10, the *congruence* of mapping a user's input actions to responses in the display space can also be regarded as a continuum. The basic idea is that, depending on the means provided and the circumstances, a user can effect changes in the observed scene either congruently with or, to varying degrees, incongruently with respect to the form, position and orientation of the device(s) provided, as shown across the top of the figure. Ordinarily, a highly congruent control-display relationship will correspond with a natural, or *intuitive*, control scheme, whereas an incongruent relationship will compel the user to perform a number of *mental transformations* in order to use it. The degree of congruence depends on a number of factors, depicted by the individual arrows beneath the all-encompassing congruency continuum shown in the figure.

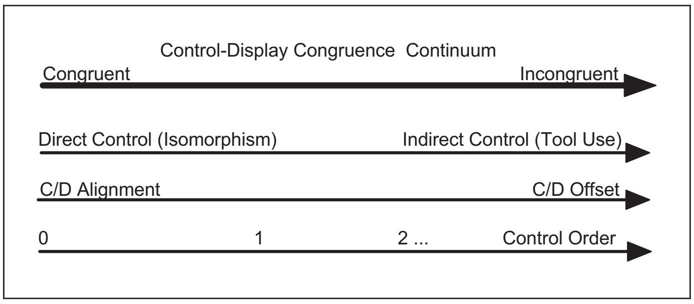


Figure 1.10 Control-Display (C/D) Congruence Continuum.

The most encompassing factor, *directness*, relates to whether the user's control actions map directly (or isomorphically) onto the display space or whether some real or metaphorical device lies between the user and the environment. The former case is naturally quite relevant to the field of Mixed Reality, which includes the broad category of see-through AR display environments, in which the user can interact with the environment with maximal directness, by using her own hands or feet. As one departs from isomorphism, the metaphor is that one is using some kind of a "tool" to manipulate the environment. Physically, this may comprise a variety of manipulanda, such as mice, joysticks, steering wheels, gloves, wands, etc., with the degree of directness of those tools affected by the other factors shown in the figure.

The next factor refers to the *alignment*, or relative location and/or orientation of the control device relative to the display space. A control/display (C/D) offset refers to a displacement between the location of the control device and the corresponding controlled object, and/or to a difference between the orientation of these. A completely aligned mapping therefore corresponds to direct control, in the line above. Conversely, as mentioned earlier, research has shown that performance degrades significantly as the size of the C/D offset increases [29] [39]. As we shall see in the following section, this factor is especially important when dealing with exocentric displays.

The bottom line in Figure 1.10 refers to the transformation between input commands to the control device and the resulting responses of the system being controlled. At a basic level this corresponds to the *control order*, that is, whether the controller has a zero, first, second or higher order transfer function. The zero order control case, at the left or congruent side of the continuum, corresponds to position control, in which case there is a simple gain factor relating control and display response. First order, or rate, control is less direct in the sense that all inputs to the control device are first integrated before control is effected. Second order control involves passing all inputs through two stages of integration, and so forth.

Example of control-display compatibility effects

In Section 1.2.2 we mentioned some of the practical issues involved in designing a particular MR system, the IIRO remotely controlled excavator, where the display involves a mixture of real data and modelled virtual objects. Thus far in Section 1.2.3 we have defined some of the basic concepts which contribute to those design issues. In the present section we bring these concepts together and discuss how they affect each other within a unified framework and how familiarity with these factors can be used to determine an appropriate viewpoint in MR systems. To simplify the discussion, as well as emphasise the generality of these considerations beyond the narrow context of remotely controlled excavators, we present the discussion within the framework of remote vehicle control, or a generic vehicle simulator.

In Figure 1.11 we present four cases of how a simple vehicle simulator might be implemented, from the point of view of real and virtual images, display centricity and control display mapping. In the two quadrants on the left (1 and 3) we have the case of an egocentric (out-the-window) display of a roadway, showing real data. In the two quadrants on the right (2 and 4) we have the same vehicle being controlled, but this time using an exocentric, or world referenced, map display. In both of the displays on the right, the "own vehicle" to be controlled is depicted by an arrow superimposed on the map. Although it is not necessary to determine in this case precisely to what extent the entire map being displayed has been created from modelled or unmodelled data, it is important to note that both the scale of the map and its orientation are modelled, in that the map is presented in a canonically conventional *north-up* fashion.

Turning to the rows, the control device for the top two quadrants is a standard steering wheel, in addition to an accelerator and brake pedals. Because the metaphor with a steering wheel is that one is situated within ones own vehicle and steering it, the nominal control mapping with the steering wheel is considered egocentric, or ego referenced. Note that this is defined to be the case also for the map display in quadrant 2, even though one is looking down on the vehicle (the small arrow shown pointing southwest in the figure). The control device shown in the bottom row (quadrants 3 and 4) is a simple computer mouse, which is exocentrically world-referenced, in a superordinate north-up manner, as indicated in the figure. This means that, by moving the mouse to the right, for example, the vehicle being controlled would be forced in an eastward direction relative to the real world, regardless of its current heading.[6]

In spite of the real vs virtual and the two ego-exo-centric distinctions shown in Figure 1.11, classifying the four cases shown in relation to the various dimensions of the MR taxonomy presented thus far is not straightforward. Although it is evident that neither of the controls are completely isomorphic (refering to Figure 1.10), since both involve the intervention of some kind of a device between the operator and the

[6]Note that the effect of using other control devices, such as a fixed joystick, either isotonic or isometric, would be very similar to that of the mouse presented here.

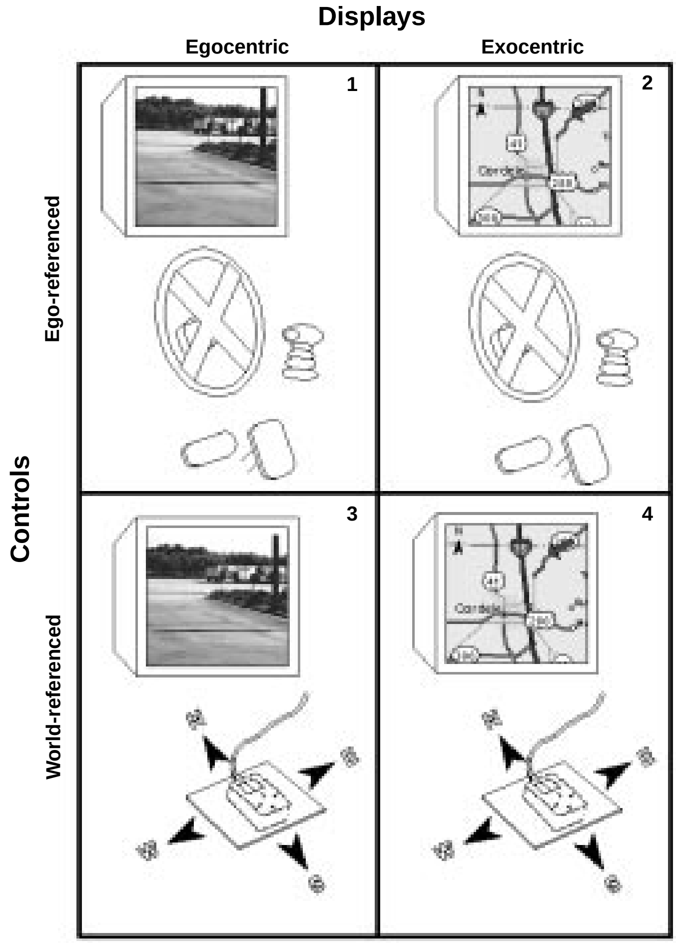


Figure 1.11 2x2 space of control-display mappings for vehicle simulator.

system being controlled, the overall degree of congruence depends not just on the device itself but on the context in which it is used, especially in relation to the associated display. For example, from a control point of view, although both devices influence direction of travel, they could each be programmed differently, as either zero-order or first order input devices.

Considering in turn the overall pros and cons of the cases shown in Figure 1.11 [23] [24] [26], in quadrant 1 we have an egocentric display together with a highly congruent control device, in the sense that a leftward control input will cause the simulated vehicle to turn to the left (as the visual display rotates to the right). Such a setup is very effective for *local guidance*, since it enables the operator/driver readily to follow an established trajectory to get from one point to another. In terms of Figure 1.9, this case lies at the ego-reference frame side of the continuum. This kind of display does not, however, encompass “flying up” above ones vehicle to survey the scene globally,[7] thus increasing the potential for keyholing, as mentioned earlier.

In quadrant 4, the keyholing problem is greatly reduced, because we now have an exocentric display, which affords a great degree of *global awareness* about where one is situated relative to the world. This case is indicated at the right hand side of Figure 1.9, at the world-reference end of the continuum. Although the display shown in quadrant 4 allows one efficiently to plan a route from A to B, traversing the route may not be as efficient as with the out-the-window view at ground level provided in quadrant 1. With respect to the control device shown in quadrant 4, a relatively high degree of control/display congruency is achieved, due to the fact that all mouse movements map directly onto the direction of motion of the vehicle. In this case, mouse input is thus highly consistent with the direct manipulation metaphor for which the mouse is best known [34].

In quadrant 2 we have an exocentric (world-referenced) display combined with an ego-referenced control device. In one sense control with the steering wheel should not be very different from input with the mouse, so this combination should work satisfactorily, but only as long as the direction of travel is generally northward. For southward driving, on the other hand, as indicated by the south-west orientation of the vehicle icon in the figure, confusion will be more likely to occur because the operator would have to perform a significant mental rotation to figure out, for example, that a rightward turn of the steering wheel would cause the vehicle icon to turn towards the exocentric operator's left.

In quadrant 3, finally, we have an egocentric display combined with an exocentric control device. Although clearly not optimal for local guidance, it might nevertheless be possible to "drive" the vehicle using the controller shown – but only as long as the vehicle is headed in a generally northward direction. The larger the deviation from a northerly heading, however, the greater the expected degree of confusion. For example, when heading eastwards, control input would be to the right in order to travel straight forward. Whenever the vehicle is headed in a southerly direction, west would be on the right and east on the left of the display. It would thus be necessary to apply a rightward force on the mouse controller in order to make the vehicle travel eastwards, to the left, and vice versa. The likelihood of control reversals would therefore increase greatly as the heading deviates more and more from a northerly direction.

1.2.4 Summary and Implications of Control-Display Issues

In the preceding discussion a number of tradeoffs among the four cases presented are outlined. These involve tradeoffs between egocentric versus exocentric displays and between ego-referenced versus world-referenced control inputs. Although the cases of quadrants 1 and 4 in Figure 1.11 are clearly more "natural" than quadrants 2 and 3, the latter have been included here not only for the sake of completeness of the

[7] Although this statement most often pertains to cases where the information in the display is comprised of real unmodelled image data, as depicted in Figure 1.11, it is nevertheless conceivable that one could also fly above a scene made up of real (3D) data which had been gathered from a number of viewpoints and integrated into a comprehensive 3D database ideally incorporating image interpolation as well. In spite of this, such a fly-through capability would certainly be more tractable with a virtual model rather than using real data.

discussion, but also because they represent realistic cases which can actually occur with Mixed Reality teleoperation systems (such as IIRO). That is, as discussed in Section 1.1, and as illustrated in the discussion of Figure 1.8, Mixed Reality display systems provide users with the opportunity to move back and forth between real world and virtual world scenes. In general, as discussed in Sections 1.2.1 and 1.2.2, real-world images typically provide increased detail, due to higher resolution relative to modelled objects, but they are also often provided only from an egocentric, out-the-window viewpoint. This is especially true for systems which move through their environment, and thus for which an external camera is not usually feasible. As indicated in Figure 1.9, such views are most effective for local guidance, that is, the task of maintaining an accurate trajectory from A to B. Conversely, exocentric displays are more conducive to global situational awareness, which comprises such tasks as landmark recognition, path planning, and obstacle avoidance, since they allow one to view the world from a number of exocentric viewpoints, either tethered or world-referenced. Together with the flexibility which accompanies the capability provided by MR to transit back and forth between real and virtual worlds, however, comes the issue of preserving control-display compatibility, as the advantages of one control scheme with one viewpoint metaphor become disadvantageous with another viewpoint. Because the consequences of such incongruencies include increased mental workload (due to the need to perform more mental transformations) and increased probability of control reversals, as well as other errors, the issues outlined in this section are expected to become increasingly relevant in future MR system research.[8]

1.3 Global Taxonomy of Mixed Reality Display Integration

In this final section we return to our original discussion of the meaning of Mixed Reality in Section 1.1 and discuss how the various issues of centricity and control presented in Section 1.2 pertain to a selection of different types of MR systems. In other words, for each class of MR system, our objective is to place the various factors presented in this paper into a single unified framework, defined within a space determined by its location along the RV continuum, the centricity continuum and the control-display congruency continuum. Our summary is presented in Figure 1.12.

1.3.1 HMD Based AR

In this section we consider the important class of AR displays based on optical or video see-through, that is, all cases in which the user "wears" the display system.

[8] It is important to take note of the fact that we have intentionally omitted consideration here of the issue of fixed versus rotating maps in our discussion of the exocentric displays in Figure 1.11. Much of the theory behind the concept of rotating displays, which corresponds to the central range of the centricity continuum depicted in Figure 1.9, can be reviewed in the writings of Wickens and his colleagues [22] [23] [26]. Research in our laboratory is currently focussing on exploring this concept experimentally, as a means of addressing some of the issues introduced in the present section.

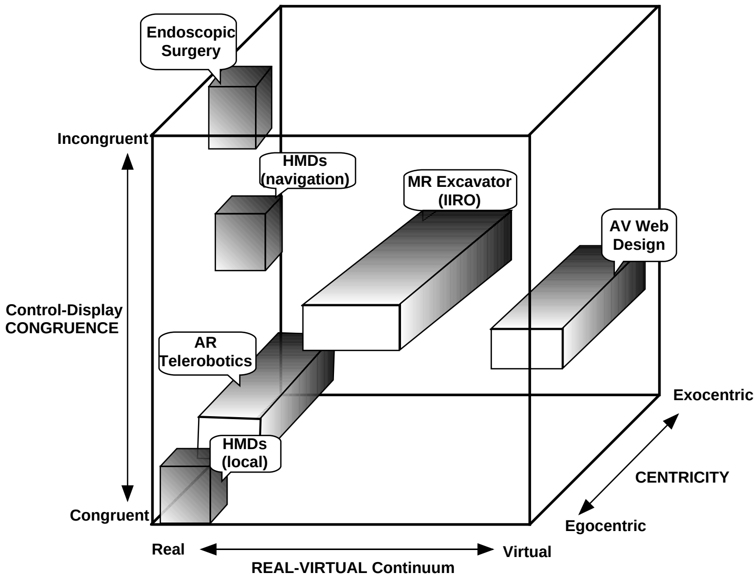


Figure 1.12 Global Taxonomy of MR Display Integration.

As discussed above, however, a distinction must be made between tasks that require information which promotes local task execution and those which require support for global situational awareness. Whereas the former class is the subject of most AR developments, for applications in manufacturing, maintenance, medicine, entertainment, etc., the latter is receiving increasing interest from the military, for example, where HMD's have been proposed for tactical displays for soldiers [8].

Consider first the more conventional AR displays for local task execution, labelled "HMD's (local)" in Figure 1.12. Because such displays are by definition based on a real-world background, this class, must lie very close to the real end of the RV Continuum, as shown. For the same reason, most such displays are also very close to the Egocentric end of the Centricity axis. Finally, for most cases one would expect to encounter good control-display congruency, since such displays are frequently designed for the user to interact directly with his immediate environment.

The classification is different for the second class of see-through AR displays, labelled "HMD's (navigation)" in Figure 1.12. Here we are referring to AR displays which attempt to provide global navigational information in a head-up form [3] [9]. From the point of view of defining AR, such displays also lie very close to the real end of the RV continuum. However, because global navigational is by definition exocentric, such displays will lie closer to the middle of the Centricity axis, in the sense that the graphic information is by nature usually top-down. Finally, for the reasons discussed above, this class of displays should be placed relatively far away from the origin, in the direction of low congruence between the outside world display and the superimposed control related navigation information.

1.3.2 Endoscopic Surgery

In many ways the field of endoscopic surgery resembles the case study presented in conjunction with Figure 1.11. In endoscopic surgery, a video camera and a set of specialised instruments are introduced into a patient's body, in lieu of direct viewing of the surgical field. With this relatively recent capability, however, comes a new set of problems involving control-display compatibility due to variations in the orientation of the intraoperative camera. For example, depending on the circumstances, a movement of the surgeon's hand, which might cause the instrument to move from left to right in her normal visual field, might cause the instrument to move from right to left on the video monitor [35]. Augmented Reality offers a number of potential advantages in this area, both for presenting navigational information and for providing the means to estimate absolute 3D distances and dimensions which would not otherwise be possible [13].

This promising area of application is represented in Figure 1.12, at the Real end of the RV Continuum, in light of the fact that the primary display medium is simple video. It is also depicted as being somewhere in the middle of the Centricity continuum. In using the term Centricity here, we are referring to the user's view of that which is being manipulated, i.e. the instruments and the surgical site, both of which are removed from the camera looking in at them. Furthermore, adding computer generated graphics containing navigational information, for example, could displace the cube even farther from the Egocentric end of the Centricity continuum, similar to the adjacent military HMD's. Finally, in light of the large obstacles which currently remain to be overcome in providing the means to map control movements unambiguously onto corresponding displayed responses, the cube has been placed fairly close to the maximum level of incongruence along the C/D Congruence dimension.

1.3.3 AR Telerobotics

The block labelled "AR Telerobotics" in Figure 1.12 refers to the system illustrated in Figure 1.3. Once again this block lies at the Real end of the RV Continuum. As discussed earlier, depending on the user's nominal viewpoint with respect to the overlaid graphics and the nature of the control laws, the Centricity could be anywhere along that axis. In using the term Centricity this time, we are referring to how the camera is located relative to the objects being manipulated. For the particular resolved control robot shown in Figure 1.3 this is an exocentric view; however for other cases, it could also have been more egocentric. For that reason, the AR Telerobotics block is shown stretching across most of the Centricity axis. However, it is shown fairly close to the highly congruent end of the Congruence axis, by virtue of the fact that the controller is fairly well matched with the display in this particular case [16] [17].

1.3.4 MR Excavator

The block labelled "MR Excavator (IIRO)" in Figure 1.12 refers to the case study presented in Section 1.2.1. Because this is truly an example of many levels of Mixed Reality, it is shown covering a large part of the centre of the RV Continuum. It is also shown stretching across the Centricity dimension, in light of the fact that users are able to alter their view flexibly between egocentric and exocentric. Finally, the IIRO block is shown in the middle of the Congruence axis, since the control display matching can be considered neutral in this case for the reasons discussed earlier.

1.3.5 AV Web Design

The block labelled "AV Web Design" refers to the example shown in Figure 1.4, which is a case of Augmented Virtuality applied towards the design of 3D web pages. Consequently, this block is located at the Virtual end of the RV Continuum. Because users are able to browse such pages either egocentrically, by moving their mouse through the central portion of the image shown in Figure 1.4, as well as exocentrically, by navigating through the plan view at the lower right, we have extended this block also to cover the whole Centricity continuum. Finally, although the control viewpoint is generally consistent with traversing the environment, the only means available for traversing the 3D web page at the present time is a 2D mouse, which can operate in a number of modes, including both position control and velocity control, the C/D Congruence has been rated neutral here.

1.3.6 Conclusion

From the examples presented in this section, a number of observations emerge. It is interesting to note how, in general, the blocks tend to spread out across the Centricity axis as the different systems vary from mostly real (AR) to mostly virtual (AV). This reflects the great flexibility offered by MR displays, where, as with the case of the IIRO excavator, users are able to exploit the advantages of both the real components and the virtual components of their system. Another message to be derived is that current applications of Mixed Reality are not occupying only one corner of the space defined in Figure 1.12, but rather are spread out over most of the taxonomy space. It is important to realise, however, that the classifications presented here are not definitive, but are subject to interpretation, as well as modification in response to changes in specific operational contexts. With this in mind, it is our hope that this graphic representation will serve both as an indicator of the great diversity of activity in the field of MR as well as a useful framework for permitting researchers to understand the similarities and differences characterising their respective endeavours.

Acknowledgements

The authors gratefully acknowledge Spar Aerospace Ltd (under the "Interactive Intelligent Remote Operations (IIRO)" Project) and the Institute for Robotics and Intelligent Systems (IRIS) (under the "Effective Display and Tele-Control Technology Integration for Real and Virtual Environments" project) for their generous support of the work reported here. We would also like to acknowledge the cooperation of the various members of the University of Toronto ETC Lab for their contributions, and especially Dr. Haruo Takemura, on leave from the Nara Institute of Science and Technology, for his generous willingness and exceptional capacity to assist throughout the composition and editing of this paper.

References

[1] R. T. Azuma: "A survey of augmented reality," *Presence*, vol.6, no.4, pp.355–385, 1997.

[2] H. Fuchs and J. Ackerman: "Displays for augmented reality: Historical remarks and future prospects," in (Y. Ohta and H. Tamura, eds.) *Mixed Reality – Merging Real and Virtual Worlds*, Ohmsha(Tokyo)-Springer Verlag(Berlin), 1999.

[3] D. Curtis, D. Mizell, P. Gruenbaum, and A. Janin: "Several devils in the details: Making an AR app work in the airplane factory," *Proc. First Int'l Workshop on Augmented Reality (IWAR'98)*, 1998.

[4] D. Reiners, D. Stricker, G. Klinker, and S. Muller: "Augmented Reality for construction tasks: Doorlock assembly," *Proc. First Int'l Workshop on Augmented Reality (IWAR'98)*, 1998.

[5] J. Molineros, V. Raghavan, and R. Sharma: "AREAS: Augmented Reality for evaluating assembly sequences," *Proc. First Int'l Workshop on Augmented Reality (IWAR'98)*, 1998.

[6] M. Bajura, H. Fuchs, and R. Ohbuchi: "Merging virtual objects with the real world: Seeing ultrasound imagery within the patient," *Computer Graphics (Proc. SIGGRAPH'92)*, vol.26, no.2, pp. 203–210, 1992.

[7] A. Stokes, C. Wickens, and K. Kite: *Display Technology: Human Factors Concepts*, Society of Automotive Engineers (SAE), 1990.

[8] W. O. Blackwood, T. R. Anderson, C. T. Bennett, J. R. Corson, M. R. Endsley, P. A. Hancock, J. Hochberg, J. E. Hoffman, and R. B. Kruk: *Tactical Displays for Soldiers – Human Factors Considerations*, National Academy Press, 1997.

[9] M. M. Glumm, W. P. Marshak, T. A. Branscome, M. McWesler, D. J. Patton, and L. L. Mullins: *A Comparison of Soldier Performance Using Current Land Navigation Equipment With Information Integrated on a Helmet-Mounted Display (ARL-TR-1604)*, Army Research Laboratory, 1998.

[10] P. Milgram and F. Kishino: "A taxonomy of mixed reality visual displays," *IEICE Trans. on Information and Systems (Special Issue on Networked Reality)*, vol.E77–D, no.12, pp.1321–1329, 1994.

[11] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino: "Augmented reality: A class of displays on the reality-virtuality continuum," *Proc. SPIE Conf. Telemanipulator and Telepresence Technologies,*, vol.2351–34, pp.282–292, 1994.

[12] P. Milgram, S. Yin, and J. J. Grodski: "An augmented reality based teleoperation interface for unstructured environments," *Proc. ANS 7th Topical Meeting on Robotics and Remote Systems*, pp.966–973, 1997.

[13] M. Kim, P. Milgram, and J. Drake: "Computer assisted 3D measurements for micro-surgery," *Proc. 41st Annual Meeting of Human Factors and Ergonomics Society*, pp.787–791, 1997.

[14] B. L. K. Davey, R. M. Comeau, P. Munger, L. Pisani, D. Lacerte, A. Olivier, and T. M. Peters: "Multimodal interactive stereoscopic image-guided neurosurgery," *Proc. Visualization in Biomedical Computing (VBC) 94.*, Sep. 1994.

[15] R. T. Azuma: "A survey of augmented reality," *SIGGRAPH'95 Course Notes*, pp.1–38, 1995.

[16] P. Milgram, A. Rastogi, and J. J. Grodski: "Telerobotic control using augmented reality," *Proc. IEEE Int'l Workshop on Robot-Human Communication (RO-MAN'95)*, pp.21–29, Jul. 1995.

[17] A. Rastogi, P. Milgram, D. Drascic, and J. J. Grodski, "Telerobotic control with stereoscopic augmented reality," *Proc. SPIE Conf. Stereoscopic Displays and Virtual Reality Systems III*, vol.2653, pp.115-122, 1996.

[18] G. U. Carraro, J. T. Edmark, and J. R. Ensor: "Techniques for handling video in virtual environments," *Proc. SIGGRAPH'98,* pp.353–360, 1998.

[19] J. Ballantyne, M. Greenspan, and M. Lipsett: "Virtual environments for remote operations." *Proc. ANS 7th Topical Meeting on Robotic and Remote Systems*, pp.545–549, 1997.

[20] P. Milgram and J. Ballantyne: "Real world teleoperation via virtual environment modelling," *Proc. 7th Int'l Conf. on Artificial Reality and Tele-Existence (ICAT'97)*, pp.1–9, 1997.

[21] E. M. Gagnon, M. Greenspan, and J. W. Ballantyne: "Real-time surface following for teleoperation in virtual environments" *Proc. SPIE Conf. Telemanipulator and Telepresence Technologies IV*, vol.3206, pp.66–73, 1997.

[22] C. D. Wickens: *Engineering Psychology and Human Performance (Second Edition)*, Harper-Collins, 1992.

[23] E. L. Faye and C. D. Wickens: "Strategies for display integration in navigational guidance and situation awareness," ARL-95-4/NASA-95-1, Aviation Research Laboratory, Institute of Aviation, University of Illinois at Urbana-Champaign, 1995.

[24] C. D. Wickens: "Integration of navigational information for flight," ARL-95-11/NASA-95-5, Aviation Research Lab, Institute of Aviation, University of Illinois at Urbana-Champaign, 1995.

[25] C. D. Wickens, C. Liang, T. Prevett, and O. Olmos: "Egocentric and exocentric displays for terminal area navigation," ARL-94-1/NASA-94-1, Aviation Research Lab, Institute of Aviation, University of Illinois at Urbana-Champaign, 1994.

[26] C. D. Wickens: "Frame of reference for navigation," in (D. G. A. Koriat, ed.) *Attention and Performance*, Academic Press, vol.16, 1998(in press).

[27] D. D. Woods: "Visual momentum: A concept to improve the cognitive coupling of person and computer," *Int'l J. Man-Machine Studies*, vol.21, pp.229–224, 1984.

[28] S. Roscoe: "Horizon control reversals and the graveyard spiral," *CSERIAC (Crew System Ergonomics Information Analysis Center) Gateway, VII,* pp.1–4, 1997.

[29] S. R. Ellis, M. Tyler, W. S. Kim, and L. Stark: "Three-dimensional tracking with misalignment between display and control axes," *Proc. 21st Int'l Conf. on Environmental Systems*, SAE Technical Paper Series 911390, 1991.

[30] E. P. McCormick, C. D. Wickens, R. Banks, and M. Yeh: "Frame of reference effects on scientific visualization subtasks," *Human Factors*, vol.40, no.3, pp.443–451, 1998.

[31] K. Hinckley, R. Pausch, J. C. Goble, and N. F. Kassell: "A survey of design issues in spatial input," *Proc. ACM Symp. on User Interface Software and Technology (UIST'94)*, pp.213–222, 1994.

[32] D. A. Bowman, D. Koller, and L. F Hodges: "Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques," *Proc. IEEE Virtual Reality Annual Int'l Symp. (VRAIS'97)*, pp.45–52, 1997.

[33] S. Zhai and P. Milgram: "Quantifying coordination in multiple DOF movement and its application to evaluating 6 DOF input devices," *Proc. ACM Conf. on Human Factors in Computing Systems (CHI'98)*, pp.320–327, 1998.

[34] A. Dix, J. Finlay, G. Abowd, and R. Beal: *Human-Computer Interaction*, Prentice Hall Europe, 1998.

[35] J. G. Holden, J. M. Flach, and Y. Donchin: "Perceptual-motor coordination in an endoscopic surgery simulation," *Surgical Endoscopy*, vol.13, 1999(in press).

[36] N. I. Durlach and A. S. Mavor (eds.): *Virtual Reality: Scientific and Technological Challenges*, National Academy Press, 1995.

[37] D. Terzopoulos and K. Fleischer: "Deformable models," *The Visual Computer*, vol.4, no.6, pp.306–331, 1988.

[38] S. Fukatsu, Y. Kitamura, T. Masaki, and F. Kishino: "Intuitive control of "Bird's Eye" overview images for navigation in an enormous virtual environment," *Proc. ACM Symp. on Virtual Reality Software and Technology (VRST'98)*, pp.67–76, Nov. 1998.

[39] M. W. McGreevy and S. R. Ellis: "Format and basic geometry of a perspective display of air traffic for the cockpit," NASA Technical Memorandum 86680, Jun. 1991.