Being There: Concepts, Effects and Measurement of User Presence in Synthetic Environments
G. Riva, F. Davide, W.A. IJsselsteijn (eds.)
IOS Press, 2003, Amsterdam, the Netherlands

2 Presence in the past: What can we learn from media history?

Wijnand IJSSELSTEIJN

Abstract. In this chapter we describe the historical development of cinema, television, telerobotics, and virtual environments, with emphasis on their psychological impact, in particular the experience of presence, or sense of 'being there' in a mediated environment. Tracing the roots of today's media technology, we find that media development seems to follow a relentless path towards greater perceptual realism, with current reproductions and simulations of reality having an unprecedented perceptual impact and immediacy. Yet people's responses to media do not appear to be a linear product of the extent of sensory information that the medium provides. Instead, they are very much shaped by people's previous experiences with and expectations towards media – i.e. their media schemata. The current chapter aims to provide a historical context from which lessons can be drawn that may inform our current media research and development efforts.

Contents

2.1 Introduction

> "Parrhasios, it is recorded, entered into a competition with Zeuxis, who produced a picture of grapes so successfully represented that birds flew up to the stage buildings *[in the theater, which served at that time as a public art gallery]*; whereupon Parrhasios himself produced such a realistic picture of a curtain that Zeuxis, proud of the verdict of the birds, requested that the curtain should now be drawn and the picture displayed; and when he realized his mistake, with a modesty that did him honor he yielded up the prize, saying that whereas he had deceived the birds, Parrhasios had deceived him, an artist."
> -- From *Natural History* by Pliny the Elder, referring to a trompe l'oeil painting in Athens, 5th century B.C.

Throughout human history there has been the aspiration to recreate and represent reality through various means, ranging from 'pure arts' such as writing, painting and sculpting, via 'combined arts' such as theater and opera, to relatively modern representational artifacts such as photography, cinema, and television. With rapid progress in the areas of computing power, interactive computer graphics, immersive displays, and digital transmission, we are able to create and experience reproductions and simulations of reality with levels of fidelity that blur the boundary between reality and its representation, rendering the impression of 'being there', or *presence*.

For instance, in the area of broadcast displays, technological advances have been aimed at improving the reproduction and scope for both sound and vision, including wide-screen, high-definition displays, immersive television, stereoscopic television, and directional audio formats [1-4]. Interactive computer graphics, head-slaved visual and auditory displays, as well as computer-generated tactile and kinesthetic interfaces, allow for a direct, multisensory, first-person experience of virtual worlds. When users are exposed to such immersive, interactive, and perceptually realistic media, they report a sense of presence in the mediated environment [5, 6].

Although popular journalism's fevered reactions to new media technologies would have us believe that creating virtual worlds is a new thing, it has in fact been an age-old desire of mankind. Traditionally addressed by the world of art and imagination, this desire has brought us literature, theatre, painting, photography, the movies, which have found their way by distribution and networks into the privacy of one's own home. Current media are more convincing than ever in representing temporally and spatially distant 'real' environments, and synthesizing non-existent ones. However, the desire to render the real and the magical, to create illusory deceptions, and to transcend our physical and mortal existence may be traced back tens of thousands of years, to paleolithic people painting in a precisely distorted, or anamorphic, manner on the natural protuberances and depressions of cave walls in order to generate a three-dimensional appearance of hunting scenes.

These paintings subtly remind us that, in spite of the impressive technological advances of today, our interests in constructing experiences through media are by no means recent. There are important lessons to be learned from the technology successes and failures of the past. Additionally, and more interestingly perhaps, we may gain a deeper insight into the development of media experiences in general, and the presence experience in particular, by attempting to reconstruct *what it was like* to witness the introduction of age-old media technologies. The current chapter investigates this historical context of presence research. We do not intend or claim this to be an exhaustive description, neither in depth nor in breadth, but rather to inform our current endeavors and to place them in a somewhat more humbling perspective.

2.2 The quest for perceptual realism

André Bazin, pioneer of film studies, saw photography and cinema as progressive steps towards attaining the ideal of reproducing reality as nearly as possible. To Bazin, a photograph was first of all a reproduction of objective sensory data and only later perhaps a work of art. In his 1946 essay, *The Myth of Total Cinema* (reprinted in his 1967 essay collection *What is Cinema?*), Bazin [7] states that cinema began in inventors' dreams of reproducing reality with absolute accuracy and fidelity, and that the medium's technical development would continue until that Platonic ideal was achieved as nearly as possible. He states:

> "The guiding myth inspiring the invention of cinema, is the accomplishment of that which dominated in a more or less vague fashion all the techniques of the mechanical reproduction of reality in the nineteenth century, from photography to the phonograph, namely an integral realism, a recreation of the world in its own image, an image unburdened by the freedom of interpretation of the artist or the irreversibility of time. If cinema in its cradle lacked all the attributes of the cinema to come, it was with reluctance and because its fairy guardians were unable to provide them however much they would have liked to." (p. 21)

In what follows, the quest for increasing realism and consequent psychological impact is the guiding theme. I will start somewhat arbitrarily with the development of cinema, including some of its precursors, although it may be argued that the arts predating cinema also clearly exhibited a tendency to reproduce reality as convincingly as possible. This is certainly true for painting, where trompe l'oeil and panorama paintings have been able to engender experiences that many 'modern' media still aspire to produce, as is evidenced by the quote from Pliny the Elder at the start of this chapter.

Although it would be tempting to describe the development of audio-visual technologies as entirely driven by some kind of realism project, this would omit an important and rich heritage of representing the surreal and the extraordinary through audio-visual means. Moreover, representational art, cinematic or otherwise, is inevitably influenced by the subjectivity of the artist, imposing a style and meaning on the medium that go beyond a 'truthful' representation of reality. Nevertheless, much of yesterday's and today's research and development efforts of the audio-visual and telecommunication industries are aimed at creating a more convincing illusion of reality. Given the expectations of current media-aware audiences, attaining an illusion of reality will often require excursions into the *hyperreal*, presenting a more 'vivid' copy of reality than reality itself has to offer.

2.3 Moving images: Early cinema

> "What is the future of the moving image? Ask rather, from what conceivable phase of the future can it be debarred...It is the crown and flower of 19th century magic, the chrystalization of Eons of groping enchantments"
> -- From *History of the Photographic and Scientific Experiments and Developments Leading up to the Perfection of the Vitascope* by William Kennedy Laurie Dickson, 1896.

Cinema was not invented in any 'Eureka' sense of fundamental breakthrough. Rather, it was the steady accumulation of insight and available technology - the pioneering work of many – that allowed Auguste and Louis Lumière to put together a moving image system, the *Cinématographe*, which saw its first public demonstration on 28 December 1895 to an audience of only thirty-three spectators in the Salon Indien, a small hall in the basement of the Grand Café at 14, Boulevard des Capucines in Paris.

The Cinématographe had many precursors. Probably one of the earliest forms of moving image projection are shadow plays which originated in Java and India some thousands of years ago. They used flexible thin leather puppets manipulated by rods in front of a translucent screen, which was lit from behind so that the puppets could be seen in silhouette. Shadow plays are still in use today.

With the development of the magic lantern, an enclosed box with a candle inside and an image drawn on glass between the light source and a lens, the principle of projection that is still used in today's cinema was essentially in place. The invention of the magic lantern is usually attributed to the Jesuit Anthanius Kircher who described the system in his book *Ars magna lucis et ubrae* (1644), or to Christiaan Huyghens, who coined the term ‘lanterne magique’ and had built one by 1659. However, it appears that the Chinese have a better claim at its invention, already by as early as 950 AD [8]. Movement was simulated early on by overlaying the static painting on glass with a second glass which could be moved by means of levers or ratchets. So, when the image was projected, a gentleman could bow from the waist, or a skeleton could playfully remove its skull.

Magic lantern shows often featured goblins and devils. According to an 18th century definition, the magic lantern was “a small machine that serves to make seen on a white wall diverse spectres and frightful monsters so that those who do not know the secret believe that they belong to the magic arts.” In addition to public entertainment, magic lanterns were used in the 18th century as conjuring devices by the pseudo-priests of Rosicrucianism and Freemasonry, who used hidden lanterns to project ghostly images onto smoke or paper-thin walls in order to disorientate and frighten their subjects [9]. Interestingly, these conjurers also used other ‘multimedia’ effects to enhance the impact of their tricks, including sound effects, odors, pyrotechnics, and even ventriloquism.

Robinson [10] notes that it is hard to assess the impact magic lanterns must have had on their 18th century audience. But given the fact that there were hardly any pictures around at that time, the impact of such a novel medium must have been considerable:

> “Even in the cities, the only representational images with which the humbler and greater part of the population could be familiar were the coloured windows and wall paintings of churches and the signs outside inns and shops. Books were inaccessible to most people, and even then rarely illustrated. Paper and ink were costly: the only place where plain folks' children could draw pictures was in the sand or snow. How astonishing then these vivid representations of life, pictures in light and colour, must have seemed.” (p. 34)

In the early 19th century, magic lantern shows became quite sophisticated, using magic lanterns with multiple lenses. A ‘dissolving pair’ could be shown by slowly covering one lens projecting one slide, and simultaneously uncovering a second lens showing a different slide that registered with the first image. At the end of the 19th century, magic lanterns were a popular form of home and public entertainment and could be found everywhere – in homes, churches, schools, large-scale halls, and theaters. Especially for large-size projections, a more powerful light source was needed, and for this ‘limelight’ was used. This was created when oxygen and hydrogen were squirted on a piece of limestone which turned incandescent once the gases were lit, and produced a light as powerful as that in a modern movie projector.

The arrival of photography in 1839 provided a revolutionary new means of representing reality, giving rise to the famous remark by Paul Delaroche that “from today, painting is dead[1].” From the 1850s photographs were projected via magic lanterns. Early photography (e.g. the Daguerreotype and the Calotype) required long exposure times which inhibited early experiments to capture motion. Early attempts required painstakingly posing and photographing models for each successive phase of movement. Interestingly, the real issue of the day was not to create an illusion of movement but rather to *stop motion* in order to study it, thus using the camera as a scientific instrument [8].

In the 1870s Eadweard Muybridge was the first to succeed in photographing actual movement, through an elaborate setup of 12 equally spaced cameras with electromagnetic shutters that were set to go off sequentially. This way he was able to analyze the movement of the human body and that of animals, answering age-old questions such as how many feet did a galloping horse have on the ground at any one time. His pictures were widely published in still form. They were also put inside a popular parlour toy, a rotating drum known as the *Zoetrope*, or “wheel of life”, which produced an illusion of movement when spun around.

Around the same time, Emile Reynaud developed a similar device, the *Praxinoscope*, an optical toy that used strips of drawings around the circumference of a mirrored drum (see figure 2.1). When the cylinder was spun, the mirrors each reflected a separate drawing so that one could rapidly see each picture in turn, giving the illusion of movement. Reynaud further developed his device by coupling the Praxinoscope to a magic lantern and projecting the moving image onto a screen. By 1880 Muybridge was similarly projecting his motion photographs using the *Zoöpraxiscope*, an adaptation of the Zoetrope. A contemporary observer of Muybridge's demonstration claimed to have seen “living, moving animals”.

Figure 2.1: Reynaud's Praxinoscope, 1876

At this stage, several essential elements of ‘true’ motion pictures still needed to be developed. A mechanism was needed that would enable sequence photographs to be taken with a single camera at regular, rapid intervals, preferably allowing 15 or more separate exposures per second. In addition, a storage medium was required that was capable of storing images for more than the one or two seconds of movement possible from the drums, wheels or disks that were in use at the time.

In 1882 Étienne-Jules Marey, motivated by the desire to analyze the movement of birds in flight, devised a ‘photographic gun’ which exposed a rapid series of images, initially 12 per second, on an

[1] In fact, very few mature representational technologies become obsolete. Contrary to predictions, photography did not kill painting, film did not kill photography (although it did replace magic lantern shows), radio did not replace newspapers, and CDROMs have not replaced books. And despite Hollywood’s fears, television did not kill film, nor did the arrival of the internet or virtual reality replace television, although there are some indications that people are reading less books.

intermittently revolving circular glass plate.[2] He went on to increase the frame rate, and employed strips of sensitized paper and later paper-backed celluloid instead of the fragile, bulky glass.

During an 1888 lecture tour in which he displayed his experiments on animal locomotion using a Zoöpraxiscope, Eadweard Muybridge lectured in West Orange, New Jersey, where he met with one of the world's most prolific inventors of that time, Thomas Alva Edison. In October of that year Edison deposited a caveat with the U.S. patents office, declaring [10]: “I am experimenting upon an instrument which does for the eye what the phonograph does for the ear, which is the recording and reproduction of things in motion” (p.138).

Edison assigned an assistant, William Kennedy Laurie Dickson, to develop the idea. Although Edison provided the facilities and perhaps the vision, there is little doubt today that all the experimental and practical work was Dickson's. The phonograph, Edison's favorite invention, was taken as a point of departure, but the model of a cylinder was doomed to failure. Only after meeting Marey in Europe, and the introduction of celluloid roll film by George W. Eastman (of Eastman Kodak), did Dickson's experiments turn to the use of strips of transparent film.

By 1892 he had developed two apparatuses, one for the recording of images, the *Kinetograph*, and one for viewing them, the *Kinetoscope*. The Kinetoscope was a cabinet with a peephole through which one could view a continuous looped film (see figure 2.2). It did not allow for projection on a screen; it was purely for individual viewing. Within three years the Kinetoscope was commercialized, and Kinetoscope parlors, with coin-operated devices showing their 40-second film, spread across the world. The commercial success of these machines may explain Edison's reluctance to develop a multiviewer projection system. This is certainly true if we accept the following statement that is attributed to Edison [10]:

> “We are making these peep show machines and selling a lot of them at a good profit. If we put out a screen machine there will be a use for maybe ten of them in the whole United States. With that many screen machines you could show the pictures to everyone in the country - and then it would be done. Let's not kill the goose that lays the golden eggs.” (p. 39-40)

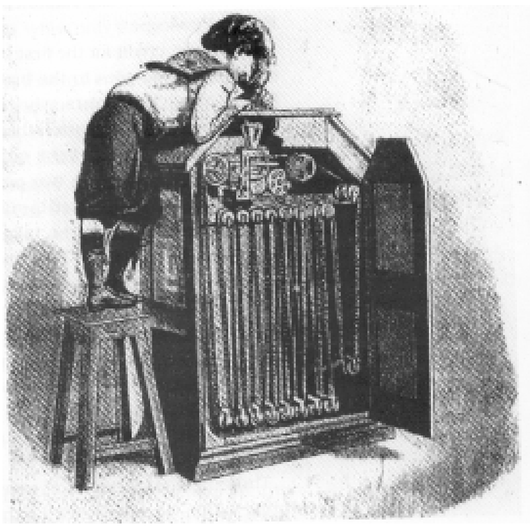

Figure 2.2: The Edison Kinetoscope, 1894

[2] Many of the 19th century inventions mentioned in this chapter can be found, with illustrations, in the wonderful book *Victorian Inventions* by Leonard de Vries [11].

However, it was only a matter of time before Edison was to recognize the ultimate direction and potential of cinema, and he acquired the rights for a projection system, which he renamed and launched as the Edison *Vitascope* in April 1896.

Although Thomas Edison (from an American perspective) or the Lumière brothers (from a European one) are often credited as being the founding fathers of the modern motion picture, it is clear that many pioneers made similar advances around the same time. By the 1890s, the ‘invention’ of cinema was really a question of engineering a practical solution out of readily available devices and concepts [7, 8].

Perhaps the German Skladanowsky brothers, former magic lanternists, should be credited for the first performance before a paying public, as they premiered their *Bioskop* on 1 November 1895 at the Wintergarten theatre in Berlin. However, their projection system, featuring two film strips which ran in parallel and were exposed alternately, proved to be too complicated and limited to survive in what was already becoming a competitive market.

The Lumière invention, patented on 13 February 1895, combined a mechanism for taking pictures and showing them in one single, compact, and lightweight[3] apparatus. It ran at 16 frames per second, and utilized a celluloid 35mm strip, the format selected by Edison and Dickson for their Kinetoscope. An important difference with the latter was that instead of one continuous loop of film, the Cinématographe allowed for the intermittent movement of the film, synchronizing the opening of the shutter with the moment the film was stationary in the aperture. This enabled larger images to be projected onto the screen, yet also produced a visible flicker from the shutter blade.

It is exceedingly difficult for people living in the 21st century to imagine what it must have been like to witness the moving images the Lumière brothers showed to their Parisian public at the dusk of the 19th century. The twenty minute programme consisted of ten short films of approximately one minute each,[4] depicting topics such as a train arriving at a station, or workers leaving through a factory gate. The moving images, projected with considerable flicker, without colour or sound, and with a jerky rapidity of all movement, would today probably strike us as some quaint experiment in minimalist cinema. At the time though, it had a fresh immediacy that went beyond anything audiences were acquainted to; it engendered a sense of realism that superseded preceding realistic media experiences. As magician and film pioneer George Méliès recalled: “We sat there with our mouths open, without speaking, filled with amazement.” A reporter's first impressions of the Cinématographe, recorded for *La Poste*, 30 December 1895, are quite illuminating [12]:

> “Messieurs Lumière, father and son, from Lyons, last night invited the press to the inauguration of a really strange and novel spectacle, whose première was reserved for the Parisian public. (...) Imagine if you will, a screen placed at the back of a vast room. This screen is visible by a crowd. On the screen appears a photographic projection. So far, nothing new. But suddenly, the image - of either natural or reduced size, depending on the scene's dimensions - animates itself and comes to life. It is a factory gate, which opens and releases a flood of workers, male and female, with bicycles, running dogs and carriages – all swarming and milling about. It is life itself; it is movement captured on the spot. (...) Photography no longer records stillness. It perpetuates the image of movement. The beauty of the invention resides in the novelty and ingenuity of the apparatus. When these gadgets are in the hands of the public, when anyone can photograph the ones who are dear to them, not just in their motionless form, but with movement, action, familiar gestures, and the words out of their mouths, then death will no longer be absolute, final” (p. 130).

[3] The Cinématographe weighed less than 20 pounds. By comparison, Edison's camera weighed 100 times as much.

[4] This film length was determined by the fact that all films were 17 meters long at the time, which was merely governed by the capacity of the spoolboxes holding the negative film when the pictures were taken.

Viewers all exhibited similar responses: sceptical or blasé at first when a static photographic projection appeared, which they already knew well from the popular magic lantern shows, then stupefied when the photograph became animated. Audiences reportedly screamed, ducked, or even ran out of the theatre as the train entering the station at La Ciotat seemed to hurl itself at them from the screen. These types of behavioral responses are very interesting and relevant from a presence research point of view, although it is hard to obtain reliable evidence as to the frequency and intensity of their occurrence.[5]

Since the Cinématographe could act as both a projecting and recording device, the operators traveling to all parts of the world not only put on shows for the public, but also produced quite an impressive amount of short films, entering over 1200 films to the Lumière catalog within a few years. Meanwhile, Louis Lumière and others continued to experiment with projection, sound, colour, and stereoscopy.[6]

This research reached a peak at the turn of the century, with several cinematic inventions being displayed at the Paris World Fair of 1900. Film would have a prominent role there, with Lumière's Cinématographe projecting onto a huge screen of about 400 square meters (21 meters wide by 18 meters high) at the fair's huge festival hall, which could accommodate fifteen thousand spectators at any one time. The screen was stored in a tank of water under the floor and was hauled up every evening with powerful winches; the projection's luminosity was enhanced by the dampness of the cloth [12]. Overall, 1.5 million spectators visited the half-hour Cinématographe show at the Paris World Fair, which gave the newborn medium an unprecedented impact, and its inventors widespread recognition.

2.4 The cinema of the 1950s

To André Bazin, the cinematic developments in the first half of the 20th century must have seemed as a logical consequence of our unconscious desire for “a total and complete representation of reality...a perfect illusion of the outside world in sound, color, and relief” [7]. Indeed, the era up until the Great Depression of the 1930s saw several revolutionary developments in cinema. Perhaps the most significant one was the introduction and acceptance of synchronous sound and dialogue in cinema in the late 1920s. One of the early systems, the *Vitaphone*, developed by Warner Brothers and Western Electric in 1926, used a one-sided phonograph record containing the audio track that needed to be kept in synchrony with the visuals on the screen. When first encountering this system, Harry Warner recalled [13]:

> “When I heard the twelve-piece orchestra on the screen at the Bell Telephone Laboratories, I could not believe my own ears. I walked in back of the screen to see if they did not have an orchestra there synchronizing with the picture. They all laughed. The whole affair was in a ten by twelve room. There were a lot of bulbs working and things I know nothing about, but there was not any concealed orchestra” (p. 207).

Later, more practical systems would print the audio track onto the film itself, thereby avoiding the problems of synchronization.

It was also during this time that shooting movies in colour became possible, instead of colouring them by hand, through the use of the Technicolor system. Technicolor was created using a special

[5] The final scene in *The Great Train Robbery* (1903) provides another anecdotal example of a behavioral response that illustrates the difficulty early movie audiences had in distinguishing art from artifice. Here, a gun pointed straight at moviegoers caused them to flinch, duck, or even jump up and run out of the theater.

[6] The Lumières are also credited with exhibiting the first stereoscopic motion picture to the general public in 1903 [13].

camera that ran three strips of film, one in red, one in blue, and one in yellow. When the three strips were consolidated, the resulting image was in full colour.[7] Other developments of the early 20th century included the exploration of various widescreen formats, such as Lorenzo del Riccio's *Magnascope*, introduced in 1924, which consisted of a special screen four times the normal size of a regular theatre screen, as well as experimentation with various stereoscopic techniques.

Thus, many of the technologies and principles behind the boom of the immersive and multisensory cinematic formats of the 1950s already had been in existence since the late 19th and early 20th century. However, it was not until the 1950s that Hollywood had to face dropping box office receipts as a consequence of the increasing popularity of a competing technology: *television*. Thus, Hollywood movie studios turned to 3-D, widescreen, and other formats as the potential 'next big thing', trying to lure people away from their new television sets and back into the cinema.

From 1952 to 1954, with 3-D movies at the height of their popularity, Hollywood produced over 65 stereoscopic films. One of the first 3-D feature films, *Bwana Devil*, released in 1952, became one of the top-grossing films of that time, earning nearly $100.000 in its first week. Producers seemed well aware of the enhanced psychological impact stereoscopic films had on the viewing public, judging from the catch-phrases used on publicity ads, such as: "The flat screen is gone! You – not a camera – but you are there" (*Bwana Devil*, 1952), "It happens to YOU in three dimensions" (*Man in the Dark*, 1953), or the slightly ambiguous "Puts the action right in your lap" (*The Starlets*, 1976), which was used for an X-rated movie [14]. Although such statements were clearly sales pitches of the films' marketing people, they do illustrate the fact that the aim of the cinema experience was to enhance the film's impact and entertainment value by making the viewer feel part of the movie - turning it into a *first-person* experience. Stereoscopic cinema has the ability to generate a compelling sense of physical space, and allows images to emerge from the screen and enter further into the spectator's space than is possible with conventional 2-D or 'flat' cinema. This effect was often exaggerated by throwing or poking objects from the screen at the viewer. As Slater and Wilbur [15] put it, with reference to virtual environments, the "discontinuity between the place of our current reality and the reality showing through the display" seemed to be collapsing.

In 1954, interest in stereoscopic cinema started to wane, in part perhaps because of the discomfort experienced when viewing misaligned and overdone stereoscopic movies. In addition, several new widescreen formats were introduced, such as *Cinerama* and *CinemaScope*, which added to the impact of the viewing experience by stimulating larger portions of the peripheral field of view.

Cinerama,[8] developed by Fred Waller, used three 35mm projections on a curved screen to create a 146 degree wide panorama (see figure 2.3). In addition to the impressive visuals, Cinerama also included a 7-channel directional sound system which added considerably to its psychological impact. Cinerama debuted at the Broadway Theatre, New York in 1952, with the independent production *This Is Cinerama*, containing the famous scene of the vertigo-inducing roller coaster ride. It was an instant success. The ads for 'This is Cinerama' promised: "You won't be gazing at a movie screen - you'll find yourself swept right into the picture, surrounded by sight and sound." The film's program booklet proclaimed [16]:

[7] The colours of these early movies were frequently very exaggerated as can be seen in two such films that were filmed in this manner, *Gone With The Wind* (1939) and *The Wizard of Oz* (1939).

[8] Interestingly, a precursor of the Cinerama system from the late 1930s - a projection system known as *Vitarama*, developed into a forerunner of modern interactive simulation systems and arcade games. Vitarama consisted of a hemispherical projection of eleven interlocked 16mm film tracks, filling the field of vision, and was adapted during the Second World War to a gunnery simulation system. The *Waller Flexible Gunnery Trainer*, named after its inventor, projected a film of attacking aircraft and included an electro-mechanical system for firing simulation and real-time positive feedback to the gunner if a target was hit. The gunnery trainer's displays were in fact already almost identical to the Cinerama system and Waller did not have much work to do to convert it.

> "You gasp and thrill with the excitement of a vividly realistic ride on the roller coaster...You feel the giddy sensations of a plane flight as you bank and turn over Niagara and skim through the rocky grandeur of the Grand Canyon. Everything that happens on the curved Cinerama screen is happening to you. And without moving from your seat, you share, personally, in the most remarkable new kind of emotional experience ever brought to the theater" (p. 189).

As Belton [16] notes, the widescreen experience marked a new kind of relation between the spectator and the screen. Traditional narrow-screen motion pictures became associated, at least from an industry marketing point of view, with passive viewing. Widescreen cinema, on the other hand, became identified with the notion of *audience participation* - a heightened sense of engagement and physiological arousal as a consequence of the immersive wraparound widescreen image and multitrack stereo sound. The type of visceral thrills offered by Cinerama was not unlike the recreational participation that could be experienced at an amusement park, and Cinerama ads often accentuated the audience's participatory activity by depicting them as part of the on-screen picture, such as sitting in the front seat of a roller coaster, "skiing" side by side with on-screen water skiiers, or hovering above the wings of airplanes [16].

Unfortunately however, Cinerama's three projector system was costly for cinemas to install, seating capacity was lost to accommodate for the required level projection, and a large staff was needed to operate the system. In addition, Cinerama films were expensive to produce, and sometimes suffered from technical flaws. In particular, the seams where the three images were joined together were distractingly visible, an effect accentuated by variations in projector illumination [16]. Together these drawbacks prevented Cinerama from capitalizing on its initial success.

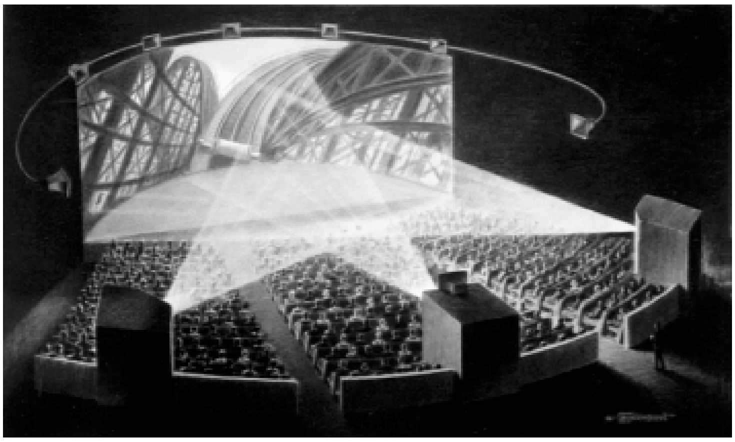

Figure 2.3: The Cinerama three projector system, 1952.

A less complex and cheaper alternative was Twentieth Century-Fox's CinemaScope, which used anamorphic[9] lenses to squeeze a (single-strip) image with a factor of 2. Thus, with a full frame camera aspect ratio (i.e. the width:height ratio) of 1.33:1 this gave a projected aspect ratio of 2.66:1, close to that of Cinerama.[10] However, CinemaScope's innovation also came at a cost. The quality of the picture suffered and projectionists had more difficulty in maintaining sharp focusing.

[9] The principle of anamorphosis, squeezing a picture during painting or photography and subsequently expanding it for viewing, was well-known and it had been applied to films as early as 1898.

[10] Eventually the aspect ratio was reduced to 2.55:1 to accommodate for the four sound tracks (left, centre, right, and surround) which were included on the film itself.

Following Cinerama and CinemaScope, numerous other film formats have attempted to enhance the viewer's cinematic experience by using immersive projection and directional sound. Today, the change to a wider aspect ratio - 1.65:1 or 1.85:1 - has become a cinematic standard. In addition, some very large screen systems have been developed, of which *Imax*, introduced at the World Fair in Osaka, Japan in 1970, is perhaps the best-known. When projected, the horizontally-run 70mm Imax film, the largest frame ever used in motion pictures, is displayed on screens as large as 30 x 22.5 m, with outstanding sharpness and brightness. By seating the public on steeply sloped seats quite close to the slightly curved screen, the image becomes highly immersive. As the ISC publicity says, “Imax films bring distant, exciting worlds within your grasp...It's the next best thing to being there” [17]. A reporter for the *Philadelphia Inquirer* writes: “It’s hard to beat the thrills of ‘Grand Canyon: The Hidden Secrets,’ the Imax film that rides the Colorado River rapids slicing through nature’s most majestic chasm….So effective are the film’s you-are-there perspectives that you wipe away imagined white-water from your brow when you’re not holding on to your armrest for dear life” (C. Rickey, June 10, 1994). Imax has also introduced a stereoscopic version, *3D-Imax*, and a hemispherical one known as *Omnimax*. Imax and other large-format theaters have been commercially quite successful, despite its auditoria being relatively few and far apart.[11]

In addition to stereoscopic and widescreen films, a number of other interesting and sometimes exotic multisensory technologies were introduced to movie theaters in the 1950s and early 1960s in order to attract viewers [18]. For instance, tactile stimulation was introduced with *Percepto* which was used for the 1959 movie *The Tingler*, and employed vibrating devices attached to theater seats. Another gimmick, called *Emergo*, sent flying skeletons down a wire and over the audience. In the 1970s, with the return of the gimmick film, a new sound effects speaker system, advertised as *Sensurround*, introduced sub-sonic rumbles that allowed the audience to ‘feel’ the theatre shake, most famously used in the 1974 movie *Earthquake.* Such gimmicks can be viewed as precursors of today's motion-based rides and enhanced theme park theaters, collectively known as *location-based entertainment* (LBE). For example, *Showscan* simulators move, tilt and shake seats in the auditorium in synchrony with what is happening on the large-format (70mm) screen. Chittock [19] describes an interesting installation in a Las Vegas hotel which clearly illustrates the ongoing tendency to increase the sensation of actually being there. It also illustrates how the careful structuring and constraining of audience participation, in combination with a seamless transition from a real to a mediated environment can greatly enhance the illusion of non-mediation [19]:

> “Arriving guests enter a bogus lift which seems to have in the far wall another set of lift doors which are actually a back-projected image from a large format film. The view through the rear ‘doors’ starts to move as the lift appears to ascend, with the floor vibrating slightly in synchronism. The speed of ascent then increases as the image of the passing floors speeds up, with the floor vibration also increasing. Suddenly, the lift cable ‘breaks’ and the moving image of the passing floors is reversed as the lift appears to plunge downwards, the floor rocking and shaking violently.” (p. 223)

Around 1960 two competing olfactory technologies, *Smell-O-Vision* and *Aromarama*, introduced smells as an added sensory dimension to sight and sound. As with stereoscopic and widescreen formats, adding smells to the motion picture had also been tried earlier. Perhaps the oldest example dates back to 1906, when S.L. Rothafel dipped cotton in a rose essence and put it in front of an electric fan, thereby introducing a rose scent into his silent film theatre in Forest City, Pennsylvania, during the newsreel of the Pasadena Rose Bowl Game [13].

[11] There are some 350 theaters worldwide that are able to project large-format movies.

Both Smell-O-Vision and Aromarama used a scent track to trigger the film's odors. Problems that any olfactory technology needs to address are the synchronization of the odors to the appropriate audio-visual scenes (basically a delay problem) and the efficient removal of the odors after they have had their impact. Smell-O-Vision used an expensive distribution system which dispensed scents from under each theater seat. Aromarama spread its odors through the theater's air conditioning system using Freon gas to diffuse the smells. However, this didn't work well, often resulting in an unpleasant mixture of aromas. When reviewing the China travelogue production *Behind the Great Wall* (1959), which featured 31 different odors, one critic from *Time* wrote: “A beautiful old pine grove in Peking smells rather like a subway rest room on disinfectant day”. Although olfactory interfaces weren't particularly successful in the cinema of the 1960s, they are today being seriously explored for their potential to enhance the experience of virtual environments [20], television [2, 3] and internet [21].

Of particular interest to presence research today is the visionary paper *El Cine del Futuro*, ‘The Cinema of the Future’, published in 1955 by Morton Heilig [22] in which he states his vision for multisensory films. Heilig, who was impressed and fascinated by the Cinerama system, went on to work out a detailed design for an *Experience Theater* in 1959 that integrated many of the ideas previously explored, and expanded on them considerably. His goal was to produce total cinema, a complete illusion of reality engendering a strong sense of presence for the audience. To Heilig, Cinerama was only a promising start, not a conclusion. He wrote [23]:

> “If the new goal of film was to create a convincing illusion of reality, then why not toss tradition to the winds? Why not say goodbye to the rectangular picture frame, two-dimensional images, horizontal audiences, and the limited senses of sight and hearing, and reach out for everything and anything that would enhance the illusion of reality?” (pp. 343-344)

This is exactly what he aimed to do by designing the Experience Theater and subsequently the *Sensorama Simulator* and *Telesphere Mask*, possibly the first head-mounted display. He went on to actually build the Sensorama Simulator [24] which stimulated as much as possible all the different senses of the observers through coloured, wide-screen, stereoscopic, moving images, combined with directional sound, aromas, wind and vibrations. The patent application for the Experience Theater explicitly mentions the presence-evoking capacity of this system [25]:

> “By feeding almost all of man's sensory apparatus with information from the scenes or programs rather than the theater, the experience theater makes the spectator in the audience feel that he has been physically transported into and made part of the scene itself.”

An interesting point illustrated by this quotation is the importance of receiving little or no sensory information that conflicts with the mediated content, such as incongruent information from one's physical surroundings (in this case the theater). This signals the mediated nature of the experience and thus acts as a strong negative cue to presence. Because of Sensorama's ability to completely immerse the participant in an alternate reality, the system is often cited as one of the precursors of modern virtual environment (VE) systems (see e.g. [26, 27]). However, despite the considerable accomplishments of Heilig's prototypes, they were still based on a relatively passive model of user perception, lacking the possibility of user action within the mediated environment but rather offering completely predetermined content.

2.5 Television

A week before the 1939 New York World's Fair opened, television pioneer David Sarnoff dedicated the RCA pavilion with a speech, entitled *Birth of an Industry*. In this first news event to be carried by television, Sarnoff predicted that television one day would become an important medium [28]: “It is with a feeling of humbleness that I come to this moment of announcing the birth in this country of a new art so important in its implications that it is bound to affect all society” (pp.167-168). A week later, on 30 April 1939, president Roosevelt's opening speech at the World's Fair would mark the start of regular commercial TV broadcasts in the United States. RCA's featured TV, the Phantom Teleceiver, attracted a great deal of attention at the Fair. The TV, normally housed in a large wooden case, was exhibited in a clear glass cabinet which enabled viewers to look inside, thus removing any doubts in viewers' minds that magic or trickery was involved in obtaining the pictures. Its small 20 x 25 cm black and white (or more precisely, black and pale green) picture tube pointed straight up, with the image reflecting off a tilted mirror hinged to the top of the case. To watch TV visitors had to watch the mirror (see figure 2.4).

Today, the truth in Sarnoff's words is evident and the impact of television on our society can hardly be overestimated. It has deeply influenced our social behavior, and the way we spend our leisure time.[12] Individuals from industrialized societies spend, on average, an astonishing three hours a day watching television. Television viewing has had a considerable impact on our taste for things visual and has shaped our conceptions and expectations of mediated as well as unmediated experiences.

As with cinema, the invention of television was the creative work of many people in many countries over many years, with critical innovations dating back to the late 19th and early 20th century. Its precise technical development is beyond the scope of this chapter, however a number of trends are relevant with respect to the presence-evoking capacity of the medium.

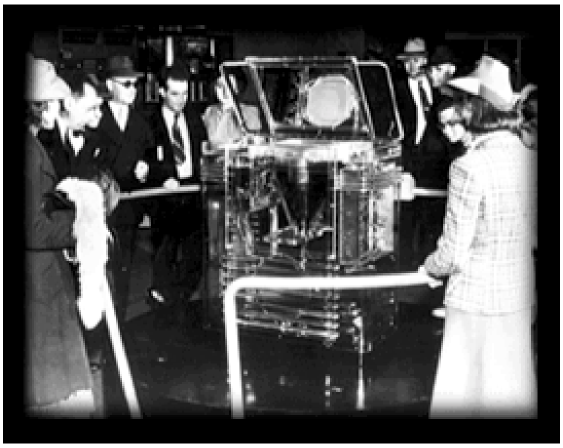

Figure 2.4: The RCA Phantom Teleceiver at the 1939 New York World's Fair

The developments of the 1950s showed that, when properly motivated, new technologies in cinema can be introduced at a fairly rapid rate. This is facilitated by the fact that cinemas do not necessarily need to adhere to any one format, and experimental immersive and multisensory formats can be

[12] In the Netherlands, a recent survey by the Social and Cultural Planning Bureau (SCP) showed that in the year 2000, people watched on average about 12.1 hours of television each week, compared to about 4 hours of reading (newspapers, books, magazines) and 2 hours of computer use for leisure purposes. While time spent on reading has been steadily decreasing over the last 15 years and computer use is showing, not surprisingly, a sharp increase, TV viewing time has remained relatively constant since 1985.

tried out in any number of cinemas before achieving broader acceptance (or failing abysmally). However, this is not the case with television, where the same signal is broadcast to millions of TV sets. As new standards emerge, the need to support the older format, either by downward compatibility or through simultaneously supporting both old and new formats, puts a constraint on the rate of introduction and acceptance of new technologies in the home (mainly since the late 1980s). In addition, with regulation authorities and industrial consortia each wanting to secure part of this huge global consumer market - at least on their own continent - negotiations on standards tend to be very complex and time-consuming.

There have been a number of experiments within existing broadcast formats to enhance the viewing experience. Stereoscopic TV, for instance, has been experimentally broadcast on several occasions using the anaglyphic (or colour-multiplexed) method to separate left and right eye views. Smells have been 'broadcast' by using *scratch'n'sniff* cards with which people could release the hidden smells at the specially-chosen moments. For example, Cartoon Network (USA) recently presented the animated series *Cow and Chicken* in 'glorious Smelly Telly', by handing out over 200.000 scratch'n'sniff cards to fans of the show.

More serious attempts at making television more immersive and realistic include the development of high-definition television (HDTV) and stereoscopic television (3-D TV). An HDTV picture is made up of about four to five times the information needed for a conventional PAL, SECAM or NTSC picture. With the extra bandwidth HDTV promises to offer a wider picture[13] at a higher resolution[14] and with more colours than current formats [29]. In addition, HDTV broadcasts will include Dolby Digital audio, featuring six separate audio tracks, each of which can be sent to a different speaker, allowing detailed and realistic surround sound. The vision behind HDTV, originally formulated in the late 1960s by researchers at NHK (Nippon Hoso Kyokai, Japan's broadcasting company) and SONY, is to match the picture quality and viewing experience of 35mm film. More than three decades later, however, HDTV has still not been introduced at any significant scale, with the notable exception of Japan.[15] It is expected however that with the introduction of digital TV transmission this process will pick up speed, and may well have an impact comparable to the change from black and white television to colour.

Although stereoscopic television has received considerable attention in the past, the advent and increased acceptance of digital TV transmission [30] reinforces its relevance. Digital transmission enables broadcasters to transmit two synchronized digital channels (one for the left and one for the right eye) in bandwidths smaller than those utilized by one analogue TV channel. In addition, stereoscopic display technology has evolved greatly over the past decades [31-33], and 3-D TV is considered to be the logical next step following HDTV. Recently, with the growing interest in stereoscopic broadcast television services, a number of laboratories, most notably in Japan, Germany, France, Canada, UK, and The Netherlands have been developing the technology and investigating the human factors requirements for high quality stereoscopic television systems. In addition, the European Commission has stimulated the development of the standards, technology, signal processing and content production facilities needed for stereoscopic broadcast services

[13] The width-to-height ratio for HDTV will be 1.78:1 (or 16:9), compared with today's TV broadcasts at 1.33:1 (or 4:3). Most motion pictures are shot at a 1.85:1 ratio.

[14] Proposals range from 1050 (HDTV U.S.A.) to 1250 (HDTV Europe) lines of resolution rather than today's 525 (NTSC in the U.S.) or 625 (PAL/SECAM in Europe). The temporal resolution of full 60 frames per second is twice the resolution NTSC images offer today.

[15] On 1 December 2000 digital HDTV broadcasts were launched in Japan. The seven-channel service is being provided by NHK, five major Tokyo-based commercial broadcasters, and WOWOW (Japan's first commercial fee-charging broadcaster). The channels have been organized to carry three conventional (standard definition) TV programs for each HDTV program. Programs are being aired 24 hours a day with the catch phrase of "Migotaesugi" (television that is worth watching). Around the time of its launch about 1.8 million Japanese households were viewing digital HDTV, with experts expecting this number to increase to 10 million households within 3 years.

through several EU-funded projects (e.g. COST 230, RACE-II DISTIMA, ACTS MIRAGE, ACTS TAPESTRIES, IST ATTEST).

A recent development that is of particular relevance to us here is the combination of a number of cinematic, LBE and VR technologies within a broadcasting framework to create *Immersive Television* [1, 3]. Aimed at broadcasting live events (e.g. a rock concert, football match, or theater performance), Immersive Television will broadcast wide-angle, high definition pictures combined with directional audio, and can optionally incorporate smells, motion feedback, wind, or temperature cues. The display can be either a magnetically tracked head-mounted look-around display for individual viewing, or a large projection display for shared viewing situations [3]. Although the technology is clearly still in a prototype stage, early evaluations already demonstrate its presence-evoking capacity.

With the arrival of digital technology in the broadcasting arena, the possibilities for user interactivity are growing fast. Even without a return signal to the broadcaster, interactivity can be (and has been) implemented on a limited scale. Through local processing, e.g. included in the set-top box, the user is able to interact with content that is already included in the broadcast signal. Taking the example of Immersive Television, local processing would allow viewers to select their own particular viewpoints from a (compressed) 360 degree broadcast signal, using a head-mounted look-around display (see also [34]).

At the time of this writing, emerging trends suggest the convergence of different platforms (television, PC, game machine, wireless phones, and PDAs), different content (audio, video, graphics, and data) and different distribution channels (cable and broadcast TV, telephone networks, satellite, and fixed wireless) into a number of integrated digital entertainment systems and services [35]. Such systems will allow users much greater choice and control, incompatible with the 'traditional' broadcasting model where the same signal is sold to millions of viewers. Enabled by digital technology and integrated broadband distribution channels, consumers need not be satisfied with whatever broadcasters choose to serve up at a given time, but can choose themselves, anytime and anywhere, from the world's collection of games, music, television programmes and movies. More importantly however, new forms of interactive entertainment will allow users to control the content to a certain extent - to become part of the story.

2.6 Interactivity and user control

It is clear from the preceding paragraphs that non-interactive media can become highly immersive and realistic, engendering a compelling sense of presence for the participant. However, no matter how engrossing non-interactive systems may become, they still assume a rather passive role for the observer. Perception, of course, is more than mere 'sensing' of the environment through our various sense organs and subsequently matching these sensations against passive representations or templates of stored information. Rather, perception is a highly activity-dependent and context-dependent process (both embodied and environmentally embedded) that integrates multimodal sensory data, ongoing actions and intentions, and memory processes. Perception serves the individual's need to control behaviour (action) within an environment. By allowing real-time action at a distance or in virtual space the participant is able to control certain aspects of the mediated environment, and, as a consequence, his/her perception of the environment. In this way, the participant will become aware that he or she is an *actor* within the environment and it is likely that this experience of willful control or *feeling of doing* will greatly enhance the feeling of actually being there within the mediated environment – the sense of presence.

The essential characteristic of any interactive system is that it will allow the user some measure of control over the media form and/or content. In this context, it is useful to distinguish between two different types of user-system interaction: navigation and manipulation. *Navigation* will allow the user to explore a given computer-generated or distant real environment, perhaps having the ability to look around in the environment (for instance enabled through a head-tracked display) and/or navigate through the environment via one of the many input devices currently available, ranging from simple cursor keys to sophisticated full-body motion trackers (see e.g.[36]). Although navigation allows the user to dynamically change the current view onto the environment, it leaves the environment itself essentially unaffected. *Manipulation* on the other hand, allows the user to affect a meaningful change in the real or virtual environment itself, e.g. cutting through tissue in endoscopic surgery, disposing of nuclear waste material using a robotic arm, or pouring virtual tea from a virtual teapot.

Importantly, the overwhelming majority of systems that support real-time interactive behavior only allow one user at a time to have control over the environment (i.e. one-to-one interactive systems). Individual control is clearly problematic for media that are designed to be experienced by more than one person at the same time (i.e. one-to-many or broadcast type systems), which is the case with cinema, television, or multi-person theme-park rides. It also raises the issue of how the quality of the content (the plot) can be assured. It's already hard to do a story well with just one plot, so one can imagine the difficulties that one would encounter when a room full of people would have control over plot development in a nontrivial manner [37]. Even when creative cohesion can be maintained, the number of choice options offered to the public needs to be kept limited, since the number of necessary pre-filmed segments grows exponentially. For example, for an interactive film with only 5 choice points, each with two choices, already 63 prerecorded sequences are needed.

Alternatively, in networked interactive environments (i.e. many-to-many type systems) multiple participants can control the outcome of a common media experience to a certain extent, an ability that trades off against a predetermined storyline. This type of shared control over a common environment is seen in networked gaming and online virtual worlds. A special case of many-to-many interactive experiences, at the boundary between theater and shared virtual environments (SVEs), is known as *participatory immersive drama*, where multiple participants are enabled to each act out their own dramatic personae within a controlled virtual story context [38-40]. In this case, the participants essentially *are* the content.

2.7 Teleoperation

Interactive systems that allow users to control and manipulate real-world objects within a remote real environment are known as *teleoperator systems*. Remote-controlled manipulators (e.g. robot arms) and vehicles are being employed to enable human work in hazardous or challenging environments such as space exploration, undersea operations, minimally invasive surgery, or hazardous waste clean-up. The design goal of smooth and intuitive teleoperation triggered a considerable research effort in the area of human factors (see e.g. [41-44]), which can be considered as one of the roots of today's presence research.

In fact, the term *telepresence* was first used in the context of teleoperation by Marvin Minsky (suggested to him by his friend Pat Gunkel) in his classic 1980 paper on the topic [45]. Minsky's paper was essentially a manifesto to encourage the development of the science and technology necessary for a remote-controlled economy that would allow for the elimination of many hazardous, difficult or unpleasant human tasks, and would support beneficial developments such as the creation of new medical and surgical techniques, space exploration, and tele-working. He writes [45]:

> "The biggest challenge to developing telepresence is achieving that sense of "being there." Can telepresence be a true substitute for the real thing? Will we be able to couple our

artificial devices naturally and comfortably to work together with the sensory mechanisms of human organisms?" (p. 48)

These questions are still valid today. Although the remote-controlled economy didn't arrive in the way Minsky envisioned, the development of telepresence technologies has significantly progressed in the various areas he identified. In addition, the arrival and widespread use of the internet brings us remote access to thousands of homes, offices, street corners, and other locations where webcameras have been set up [46]. In some cases, because of the two-way nature of the internet, users can log on to control a variety of telerobots and manipulate real-world objects.

2.8 Virtual environments

Whereas telerobotics enable the manipulation of remote real-world environments, virtual environments[16] allow users to interact with synthetic or computer-generated environments, by moving around within them and interacting with objects and actors represented there. In its most well-known incarnation, VEs are presented to the user via a head-mounted display (HMD) where the (often stereoscopic) visual information is presented to the eyes via small CRTs or LCDs, and auditory information is presented using headphones. The HMD is usually fitted with a position tracking device which provides the necessary information for the computer to calculate and render the appropriate visual and auditory perspective, congruent with the user's head and body movements[17] [48-50]. An alternative interface to the HMD is the *BOOM* (Binocular OmniOriented Monitor) where the display device isn't worn on the head but mounted onto a flexible swivel arm construction so that it can be freely moved in space. Moving a BOOM needs to be done manually however, thereby occupying the hands. Tactile and force feedback is also sometimes provided through various devices, ranging from inflatable pressure pads in data gloves or body suits to force-feedback arms or exoskeleton systems. Although there is an increasing interest in engineering truly multisensory virtual environments [51], such systems are still rather the exception.

A second common design of immersive virtual environments is through multiple projection screens and loudspeakers placed around the user. A popular implementation of such a projection system is known as the *CAVE* [52], a recursive acronym for CAVE Automatic Virtual Environment, and a reference to *The Simile of the Cave* from Plato's *The Republic*, in which he discusses about inferring reality from projections (shadows) thrown on the wall of a cave [53]. The CAVE consists of three stereoscopic rear-projection screens for walls and a down-projection screen for the floor. Participants entering this room-like display are surrounded by a nearly continuous virtual scene. They can wear shutterglasses in order to see the imagery in stereo, and wearing a position tracker is required to calculate and render the appropriate viewer-centered perspective.

Other less immersive implementations of virtual environments are gaining in popularity because they do not isolate the user (like an HMD) or require a special room (like a CAVE) and are thus

[16] Virtual environments are sometimes also referred to as *virtual reality* or VR. Although we feel that both terms are essentially synonymous, we agree with Ellis [47] who notes that the notion of an *environment* is in fact the appropriate metaphor for a head-coupled, coordinated sensory experience in three-dimensional space.

[17] The main advantage of an HMD - its immersive characteristics, sealing off the physical environment and presenting the senses with an inclusive virtual environment - also accounts for the disadvantage of perceptually isolating the user from other participants in the real world, as well as from visual feedback of his/her own body. This property makes it more difficult for multiple participants to communicate effectively, as it is hard to share a common frame of reference, and virtually impossible to communicate non-verbally (e.g. establishing eye contact or perceiving gestures) when wearing an HMD. Ways to overcome this limitation include using projection displays or see-through displays (*augmented reality*), although alternatively social actors can also be modeled and represented as virtual actors within the VE. However, this latter approach still poses significant hardware and software challenges.

more easily integrated with daily activities. Such systems include stationary projection desks, walls, or head-tracked desktop systems [54]. The latter is sometimes referred to as *fish-tank* virtual reality [55].

The first virtual environment display system where a totally computer-generated image was updated according to the user's head movements and displayed via a head-referenced visual display was introduced in 1968 by Ivan Sutherland, who is now generally acknowledged as one of the founding fathers of virtual reality. The helmet device Sutherland built was nicknamed *Sword of Damocles* as it was too heavy to wear and had to be suspended from the ceiling, hanging over the user's head. In his classic 1965 paper *The ultimate display*, he describes his ideas of immersion into computer-generated environments via new types of multimodal input and output devices. He concludes his paper with the following vision [56]:

> "The ultimate display would, of course, be a room within which the computer can control the existence of matter. A chair displayed in such a room would be good enough to sit in. Handcuffs displayed in such a room would be confining, and a bullet displayed in such a room would be fatal. With appropriate programming such a display could literally be the Wonderland into which Alice walked." (p.508)

Figure 2.5: A U.S. Postal stamp depicting a prototypical virtual reality setup with HMD and datagloves

Although today we are still clearly a long way removed from such an ultimate display, current virtual environments, through their multisensory stimulation, immersive characteristics and real-time interactivity, can approximate the stimulation a user would receive if moving about in a comparable real physical environment. Consequently, participants often experience a strong sense of presence within the virtual environment.[18] To date, much of the research on the various factors

[18] In fact, Steuer [57] argues for a definition of virtual reality as "a real or simulated environment in which a perceiver experiences telepresence", thus approaching VR from an entirely human experience point of view, without reference to any particular hardware system. Such a definition of VR covers a much wider range of media than just interactive computer-generated environments, as it includes all media capable of engendering a sense of presence in the user.

that can influence the presence experience has been performed in relation to virtual environments (see e.g. [58-61]). This research has been reviewed elsewhere in detail (see e.g. [62, 63]).

2.9 What can we learn from media history?

> "When anything new comes along, everyone, like a child discovering the world, thinks that they've invented it, but you scratch a little and you find a caveman scratching on a wall is creating virtual reality in a sense. What is new here is that more sophisticated instruments give you the power to do it more easily. Virtual reality is dreams."
> -- Morton Heilig, quoted in [64], p.57.

The history of cinema and VR reflects our ongoing aspiration to create a "total and complete representation of reality". It follows a relentless path towards greater perceptual realism, with current technologies enabling more realistic reproductions and simulations than ever before. But as Morton Heilig's quote illustrates, the vision behind these developments is age-old. The search for the 'Ultimate Display', to use Sutherland's [56] phrase, has been motivated by a drive to provide a perfect illusory deception, as well as the ancient desire for physical transcendence, i.e. escaping from the confines of the physical world into an 'ideal' world dreamed up by the mind [65].

Whereas arts and entertainment enterprises were behind much of the developments in cinema and television, the initial development of virtual environments, teleoperation and simulation systems has mainly been driven by industrial and military initiatives. Despite these different historical roots, a major connection between work in virtual environments, aircraft simulations, telerobotics, location-based entertainment, and other advanced interactive and non-interactive media, is the need for a thorough understanding of the human experience in real environments [47]. It is the perceptual experience of such environments that we are trying to create with media displays. When comparing real space to virtual space, limiting ourselves to visual media for the time being, we find that real world perception has several critical features [66]:

- Static depth information is provided via several independent mechanisms (e.g. linear perspective, interposition, texture density gradients, binocular disparity) that are consistent with each other and the observer's viewpoint.
- The resolution and intensity of the image is only limited by the sensitivities of our visual system.
- The effective image size fills our entire field of view, limited only by our facial structures, but without an externally imposed frame.
- Dynamic depth information (i.e. motion parallax) is coupled to observer movement

As we move towards increasingly realistic media, each development in visual media can be viewed as a gradual buildup of perceptual cues that simulate natural perception and enhance the experience of presence [65]. Early perspective paintings, dating back to the middle ages, only included static monocular depth information which violated most of the critical features of real-world perception mentioned above. The end of the 18th century saw the introduction of panorama paintings, which stimulated large portions of the visual periphery, a principle that was also applied to great effect in the cinema of the 1950s (Cinerama, CinemaScope), and in more recent large film formats. The stereoscope of the 19th century allowed each eye to view the same scene from a slightly different perspective (i.e. stereoscopically), contributing greatly to the perception of egocentric distance and exocentric depth within an image.

With the introduction of cinema, motion has been added to high-resolution photorealistic imagery as a fundamental perceptual cue. The visual system is highly motion-sensitive, and the onset of motion cannot be ignored - it demands attention [67] and automatically elicits an orienting response. Certain camera movements provide motion parallax as a cue to depth, although it is important to note that observer movement does not transform the image appropriately. In the case of head-mounted virtual environments this viewpoint-dependent transformation is possible in real-time, although with the current state of technology real-time interactivity trades off against photorealism.[19] Importantly, it is not clear at present how much each feature or perceptual cue contributes to the perceived realism of media, or to eliciting a sense of presence for the participant, nor is it clear how these cues interact with each other. As Heeter [68] states, "the alchemy of presence in VR is in part a science of tradeoffs". It is of clear theoretical and practical value to establish what the optimal mix of cues might be for different application contexts, or, if the optimum is unattainable, which elements are most critical to the experience of presence? These are empirical questions worth investigating.

As we are approaching the stage at which visual media can simulate the majority of critical features of natural visual perception mentioned above, the media experience is transformed from looking at a picture, to experiencing a space, to eventually visiting it as a place and being able to act within it. Despite the rapid developments, it is still a considerable challenge to design a realistic multisensory environment with which we can interact intuitively and where participants can not only look, but also hear, touch, smell, taste, walk, run, pickup objects, etc.

From the anecdotal evidence accumulating throughout media history, it becomes clear that people's responses to media are not a linear product of the extent of sensory information that the medium provides, but are very much shaped by people's previous experiences with and expectations towards media. It would seem a little odd to us now if people should panic and run out of a movie theatre at the sight of an approaching train on the screen. This is because our *media schemata*, or knowledge representations of what media are, and are capable of, tell us what to expect from mediated experiences, including the many perceptual tricks that cinema or VR can play on us. Through our many encounters with different forms of media, most notably television, we have learned what media can do or, at least, what we *think* they can do.[20] When suddenly confronted with sensory stimulation that up to that point in time had only been possible in the realm of unmediated reality, it comes as little surprise that people respond to it in similar ways as they would respond to unmediated stimuli. Thus, media schemata may act as an attenuating factor on our initial response to take the stimulus at face value and act accordingly.

Despite this clear inhibitory effect of media schemata, there are numerous examples where we still exhibit a tendency to respond to media in much the same way as we would to reality [69]. At a non-cognitive response level, our perceptual system has simply not evolved to deal with media as something separate from reality. As Reeves and Nass [69] put it:

---

[19] At the time of this writing, the computational resources that photorealistic real-time graphics rendering require are still too demanding for current systems. Significant simplifications need to be incorporated in the virtual world models in order to make them run interactively in real-time without perceptible lags.

[20] Sometimes, our expectations of media technology can become unrealistically high. For example, movies that feature virtual reality technology (e.g. *Lawnmower man*) often depict the computer-generated scenes as highly perceptually realistic (typically photorealistic, with significant auditory and haptic feedback), responsive in real-time to our actions (i.e. without a perceptible lag), and with limitless worlds to explore. This combination of features is clearly beyond the current state-of-the-art in VR. In addition, the photographic quality of graphics seamlessly integrated in current movies (e.g. *Titanic* or *The Matrix*) may also heighten the expectations of what is possible, hiding the fact that some scenes may take days to render. Thus, chances are that people may initially be disappointed by the level of realism in a typical VR experience.

"During nearly all of the 200.000 years in which *Homo Sapiens* have existed, anything that acted socially really was a person, and anything that appeared to move toward us was in fact doing just that. Because these were absolute truths through virtually all of human evolution, the social and physical world encouraged automatic responses that were, and still are, the present-day bases for negotiating life. Acceptance of what only *seems* to be real, even though at times inappropriate, is automatic" (p. 12).

This fact can provide a basis for measuring presence responses in a behavioral way, corroborating subjective rating methods. For instance, Freeman, IJsselsteijn and colleagues [70-72] have investigated observers' automatic postural responses to moving video (a sequence of a rally car traversing a curved track at speed) and found that more substantial lateral postural responses occurred when the video was projected stereoscopically than when it was presented monoscopically, thus corroborating subjective ratings of presence.

Finally, one lesson that history seems to be teaching is that increasing the breath and depth of sensory experience – increasing the perceptual bandwidth, if you like - will automatically improve the media experience. It is important to keep in mind, however, that *intensity does not equal quality*. The basic appeal of media still lies in its content, the storyline, the ideas and emotions that are being communicated. We can be bored in VR and moved to tears by a book. Nevertheless, the psychological impact of content, both good and bad, exciting and boring, depends to a large extent on the form in which it is represented.

2.10 Acknowledgement

I would like to thank dr. Nick Lodge, whose views on the history of cinema and the future of television have been inspirational, and who made useful comments on the first draft of this chapter.

2.11 References

[1] Freeman, J., Lessiter, J. & IJsselsteijn, W. (2001). Immersive television, *The Psychologist* 14: 190-194.
[2] Isono, H., Komiyama, S. & Tamegaya, H. (1996). An autostereoscopic 3-D HDTV display system with reality and presence, *Society for Information Display 96 Digest* 135-138.
[3] Lodge, N. (1999). Being part of the fun - Immersive television!, Keynote address at the Conference of the Broadcasting Engineering Society, New Delhi, India, 2-4 February 1999.
[4] Meares, D. (1991). Sound systems for high definition television, *Applied Acoustics* 33: 229-243.
[5] Barfield, W., & Weghorst, S. (1993). The sense of presence within virtual environments: A conceptual framework, in G. Salvendy & M. Smith (eds), *Human computer interaction: Software and hardware interfaces*, Elsevier, Amsterdam, 699-704.
[6] Freeman, J., & Avons, S.E. (2000). Focus group exploration of presence through advanced broadcast services. *Proceedings of the SPIE*, 3959: 530-539.
[7] Bazin, A. (1967). *What is cinema? Vol.1*, University of California Press, Berkeley.
[8] Winston, B. (1996). *Technologies of seeing: Photography, Cinematography and Television*, British Film Institute, London.
[9] Heard, M. (1996). The magic lantern's wild years, in C. Williams (ed.), *Cinema: The beginnings and the future*, University of Westminster Press, London, 24-32.
[10] Robinson, D. (1996). Realising the vision: 300 years of cinematography, in C. Williams (ed.), *Cinema: The beginnings and the future*, University of Westminster Press, London, 33-40.
[11] de Vries, L. (1971). *Victorian inventions*, John Murray, London.
[12] Toulet, E. (1995). *Cinema is 100 years old*, Thames and Hudson Ltd., London.
[13] Limbacher, J. (1968). *Four aspects of the film: 3-D, Color, Sound, Widescreen*, Brussel and Brussel, New York.
[14] Hayes, R. (1989). *3-D Movies. A History and Filmography of Stereoscopic Cinema*, McFarland & Company, Jefferson, NC.

[15] Slater, M. & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments, *Presence: Teleoperators and Virtual Environments* 6: 603-616.

[16] Belton, J. (1992). *Widescreen Cinema*, Harvard University Press, Cambridge, MA.

[17] Wollen, T. (1993). The bigger the better: From CinemaScope to Imax, in P. Hayward & T. Wollen (eds), *Future visions: New technologies of the screen*, British Film Institute, London, 10-30.

[18] Lodge, N. (2000). What can we learn from the cinema of the 1950's?, Invited paper presented at PRESENCE 2000, 3rd International Workshop on Presence, Delft, The Netherlands, 27-28 March 2000.

[19] Chittock, J. (1996). Back to the future: The cinema's lessons of history, in C. Williams (ed.), *Cinema: The beginnings and the future*, University of Westminster Press, London, 217-227.

[20] Barfield, W. & Danas, E. (1996). Comments on the use of olfactory displays for virtual environments, *Presence: Teleoperators and Virtual Environments* 5:109-121.

[21] Platt, C. (1999). You've got smell!, *Wired* 7(11): 256-263.

[22] Heilig, M. (1992). El cine del futuro: The cinema of the future, *Espacios*, reprinted in *Presence: Teleoperators and Virtual Environments* 1: 279- 294.

[23] Heilig, M. (1998). Beginnings: Sensorama and the Telesphere Mask, in C. Dodsworth Jr. (ed.), *Digital Illusion*, ACM Press, New York, 343-351.

[24] Heilig, M. (1962). Sensorama simulator, U.S.Patent # 3050870.

[25] Heilig, M. (1971). Experience theater, U.S.Patent # 3628829.

[26] Rheingold, H. (1991). *Virtual Reality*, Martin Secker & Warburg Ltd., London.

[27] Coyle, R. (1993). The genesis of virtual reality, in P. Hayward & T. Wollen (eds), *Future visions: New technologies of the screen*, British Film Institute, London, 148-165.

[28] Gelernter, D. (1995). *1939 - The lost world of the fair*, The Free Press, New York.

[29] Renaud, J.-L. (1993). Towards higher definition television, in P. Hayward & T. Wollen (eds), *Future visions: New technologies of the screen*, British Film Institute, London, 46-71.

[30] Meyer, L. & Fontaine, G. (2000). Development of digital television in the European Union, Final Report to the European Commission DC XIII, June 2000, Institut de l'audiovisuel et des Telecommunications en Europe (IDATE).

[31] Okoshi, T. (1980). Three-dimensional displays, *Proceedings of the IEEE* 68:548-564

[32] Pastoor, S. & Wöpking, M. (1997). 3-D displays: A review of current technologies, *Displays* 17: 100-110.

[33] Sexton, I. & Surman, R (1999). Stereoscopic and autostereoscopic display systems, *IEEE Signal Processing Magazine* 85-99.

[34] Hirose, M., Yokoyama, K. & Sato, S. (1993). Transmission of realistic sensation: Development of a virtual dome, *Proceedings of IEEE VRAIS '93* 125-131.

[35] Forman, P. & Saint John, R. (2000). Creating convergence, *Scientific American* 283(5): 34-40.

[36] Subramanian, S. & IJsselsteijn, W. (2000). Survey and classification of spatial object manipulation techniques, in C. Paris, N. Ozkan, S. Howard & S. Lu (eds), *Interfacing reality in the new millennium*. OZCHI 2000 Conference Proceedings, CHISIG, Sydney, 330-337.

[37] Kamberg, M. (1998). Seafari: An expedition into motion based ride filmmaking, in C. Dodsworth Jr. (ed.), *Digital illusion*, ACM Press, New York, 391-404.

[38] Murray, J. (1997). *Hamlet on the Holodeck. The future of narrative in cyberspace*, MIT Press, Cambridge, MA.

[39] Springel, S. (1999). Participatory immersive drama - a media paradigm for the next century, Paper presented at the 2nd International Workshop on Presence, University of Essex, UK, 6-7 April 1999.

[40] Davenport, G. (2000). Your own virtual storyworld, *Scientific American* 283(5):61-64.

[41] Johnson, E. & Corliss, W. (1971). *Human factors applications in teleoperator design and operation*, Wiley-Interscience, New York.

[42] Bejczy, A. (1980). Sensors, controls, and man-machine interface for advanced teleoperation, *Science* 208: 1327-1335.

[43] Sheridan, T. (1992). *Telerobotics, automation and human supervisory control*, MIT Press, Cambridge, MA.

[44] Stassen, H. & Smets, G. (1995). Telemanipulation and telepresence, in T. Sheridan (ed.), *Analysis, Design and Evaluation of Man-Machine Systems 1995*, Elsevier Science Ltd., Oxford, 13-23.

[45] Minsky, M. (1980). Telepresence, *Omni* 45-51

[46] Campanella, T. (2000). Eden by wire: Webcameras and the telepresent landscape, in K. Goldberg (ed.), *The Robot in the Garden*, MIT Press, Cambridge, MA, 22-46.

[47] Ellis, S. (1991). Nature and origins of virtual environments: A bibliographical essay, *Computing Systems in Engineering* 2: 321-347.

[48] Kalawsky, R. (1993). *The science of virtual reality and virtual environments*, Addison-Wesley, Reading, MA.

[49] Burdea, G. & Coiffett, P. (1994). *Virtual reality technology*, Wiley & Sons, New York.

[50] Biocca, F. & Delaney, B. (1995). Immersive virtual reality technology, in F. Biocca & M. Levy (eds), *Communication in the age of virtual reality*, Lawrence Erlbaum, Hillsdale, NJ, 57-124.

[51] Dinh, H., Walker, N., Song, C., Kobayashi, A. & Hodges, L. (1999). Evaluating the importance of multi-sensory

input on memory and the sense of presence in virtual environments, *Proceedings of IEEE VR 1999* 222-228.
[52] Cruz-Neira, C., Sandin, D. & DeFanti, T. (1993). Surround-screen projection-based virtual reality: The design and implementation of the CAVE, *Computer Graphics: Proceedings of SIGGRAPH* 135-142.
[53] Plato (380 B.C., translation 1955). *The Republic*, Penguin Books, London.
[54] Fisher, S. (1982). Viewpoint dependent imaging: An interactive stereoscopic display, *Proceedings of the SPIE 367*.
[55] Ware, C., Arthur, K. & Booth, K. (1993). Fish tank virtual reality, *Interchi '93 Conference Proceedings* 37-42.
[56] Sutherland, I. (1965). The ultimate display, *Proceedings of the IFIP Congress* 2:506-508.
[57] Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence, *Journal of Communication* 42(4): 73-93.
[58] Hendrix, C. & Barfield, W. (1996). Presence within virtual environments as a function of visual display parameters, *Presence: Teleoperators and Virtual Environments* 5: 274-289.
[59] Welch, R., Blackmon, T., Liu, A., Mellers, B. & Stark, L. (1996). The effects of pictorial realism, delay of visual feedback, and observer interactivity on the subjective sense of presence, *Presence: Teleoperators and Virtual Environments* 5: 263 - 273.
[60] Witmer, B. & Singer, M. (1998). Measuring presence in virtual environments: A presence questionnaire, *Presence: Teleoperators and Virtual Environments* 7: 225 - 240.
[61] Nichols, S., Haldane, C. & Wilson, J. (2000). Measurement of presence and its consequences in virtual environments, *International Journal of Human-Computer Studies* 52: 471-491.
[62] Stanney, K. M., Salvendy, G., et al. (1998). Aftereffects and sense of presence in virtual environments: Formulation of a research and development agenda, *International Journal of Human-Computer Interaction* 10: 135-187.
[63] IJsselsteijn, W.A., de Ridder, H., Freeman, J., & Avons, S.E. (2000). Presence: Concept, determinants and measurement, *Proceedings of the SPIE* 3959: 520-529.
[64] Hamit, F. (1993). *Virtual reality and the exploration of cyberspace*, SAMS Publishing, Carmel, IN.
[65] Biocca, F., Kim, T. & Levy, M. (1995). The vision of virtual reality, in F. Biocca & M. Levy (eds), *Communication in the age of virtual reality*, Lawrence Erlbaum, Hillsdale, NJ, 3-14.
[66] Avons, S. (1996). Mechanisms of visual perception: A selective review with reference to VR applications, Unpublished document, University of Essex, UK.
[67] Reeves, B., Thorson, E., Rothschild, M., McDonald, D., Hirsch, J. & Goldstein, R. (1985). Attention to television: Intrastimulus effects of movement and scene changes on alpha variation over time, *International Journal of Neuroscience* 27: 241-255.
[68] Heeter, C. (1992). Being There: The Subjective Experience of Presence, *Presence: Teleoperators and Virtual Environments* 1: 262-271.
[69] Reeves, B. & Nass, C. (1996). *The media equation: How people treat computers, television and new media like real people and places*, Cambridge University Press, New York.
[70] Freeman, J., Avons, S.E., Meddis, R., Pearson, D.E. & IJsselsteijn, W.A. (2000). Using behavioural realism to estimate presence: A study of the utility of postural responses to motion-stimuli, *Presence: Teleoperators and Virtual Environments* 9: 149-164.
[71] IJsselsteijn, W.A., de Ridder, H., & Freeman, J. (2000). Postural responses to real-world moving video - an objective measure of presence? *Proceedings of Measuring Behavior 2000*: 154-156.
[73] IJsselsteijn, W.A., de Ridder, H., Freeman, J., Avons, S., & Bouwhuis, D. (2001). Effects of stereoscopic presentation, image motion and screen size on subjective and objective corroborative measures of presence, *Presence: Teleoperators and Virtual Environments* 10: 298-311.