1

Introduction

Recent years have seen a reawakening of interest in 3-dimensional (3D) visual technology. 3D, in the form of stereoscopy, has been with us since 1838, when it was first described by Sir Charles Wheatstone. Since then there have been a number of periods when interest in 3D technology has surged and then faded away again. Each resurgence in interest can largely be put down to the development of new technologies, or new marketing initiatives. The constant reawakening of interest also demonstrates the strong desire of the public for immersive 3D experiences. The fading away of interest can largely be put down to the disappointing nature of previous generations of 3D technology.

We are currently at the beginning of another resurgence in interest in 3D, which is likely to be durable. There are a number of reasons why this should be:

- affordable, aesthetically pleasing, 3D displays, which are as capable of displaying high quality 2D colour video, as they are of showing high quality 3D video;
- digital video production techniques to allow correction and optimization of captured 3D video during post-production;
- new developments in the understanding of 3D perception, which enable the production of content which is more comfortable to the eye;
- new formats and standards for the compression of 3D video in digital formats, enabling high quality 3DTV services to become a reality.

This book aims to provide the reader with an overview of the key technologies behind the current generation of 3D technologies, and also to provide a guide to where the technology will head next. It covers the full chain, from capture of 3D video, to display. In between, it examines issues such as 3D video

3DTV: Processing and Transmission of 3D Video Signals, First Edition.
Anil Fernando, Stewart T. Worrall and Erhan Ekmekcioğlu.

compression, assessment of 3D video quality, and transmission of the video over a variety of networks.

In this chapter, Section 1.1 describes the history of 3D video, highlighting the key developments since the nineteenth century. Section 1.2 describes the most common digital 3D video formats currently in use. The motivation for the book, and for the reintroduction of 3D video in general, is outlined in Section 1.4. The most common application scenarios for 3D video are discussed in Section 1.3. Finally, Section 1.5 gives an overview of each of the book chapters.

1.1 History of 3D Video

Before examining the current state-of-the-art in 3D visual related technology, it is instructive to examine the way in which 3D technology has developed, and the reasons for previous failures. In this way it is possible to assess the durability of the current 3D boom, and to consider which of the remaining challenges are the most important to solve.

Figure 1.1 gives a summary of some of the most important milestones over the past 150 years. One aspect that may surprise some readers is the length of the timeline. Many key developments took place either in the nineteenth century or the early twentieth century. 3D was most popular during the 1950s and 1980s, but each 3D boom faded within a few years. The following subsections describe the key technological developments and discuss the reasons for the promotion and subsequent failure of 3D movie technology. Finally, the latest resurgence in popularity in 3D is examined. It is important to note that it is only relatively recently that 3D displays for the home have been available at an affordable price.

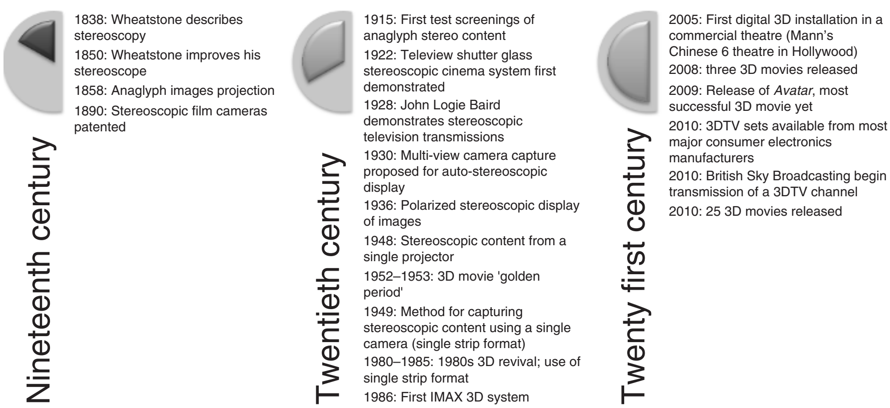


Figure 1.1 Key developments in the history of 3D video

This section of the book describes the history of 3D video-related technology and applications over time, up until the twenty-first century. There follows specific sections on auto-stereoscopic displays, and 3DTV-specific developments. Auto-stereoscopic and volumetric displays have been treated separately because they have been of considerable interest to the research community, but have not yet been commercially exploited.

The history described here is relatively brief, and some details have been left out. For example, a number of significant developments were made in the Soviet Union, which are not covered here. For a more detailed historical overview, readers should consult other references [1, 2].

1.1.1 3D in the Nineteenth Century

Sir Charles Wheatstone is widely considered to be the father of stereoscopy. In 1838, he published a paper describing how each eye sees a slightly different version of the same scene [3]. It seems unlikely that Wheatstone was the first person in history to notice this effect. Wheatstone himself provides a quote in his 1838 paper, which suggests that Leonardo Da Vinci would probably have been aware of the effects of binocular vision. Other authors have noted that Euclid also made certain observations about binocular vision [1]. However, Wheatstone is the first to describe stereoscopy explicitly and in detail. Wheatstone's 1838 paper also described the stereoscope, which allowed viewing of stereoscopic drawings. As shown in Figure 1.2, The stereoscope used mirrors to project the drawings at E′ and E to the position of the viewers' eyes (A′ and A respectively). In 1840, he was awarded the Royal Medal of the Royal Society for his work on stereoscopy.

The next notable development was the introduction of anaglyph images. The anaglyph approach involves the use of glasses where the two lenses are different colours. Red-cyan lenses were often used for viewing stereoscopic

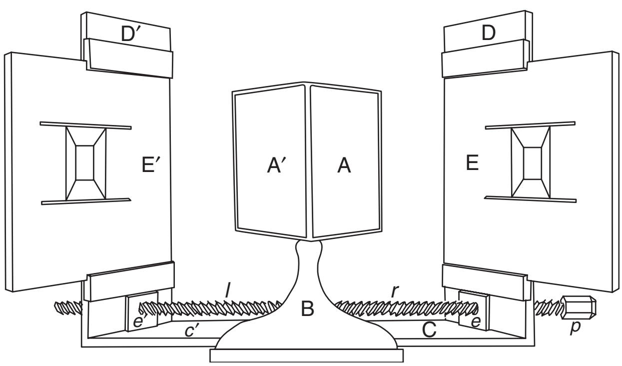


Figure 1.2 Wheatstone's mirror stereoscope, taken from his 1838 paper, which allows viewing of a stereo pair of drawings (3) (Reproduced with permission of the Royal Society Publishing)

images. The two stereoscopic views are superimposed, and the colour lenses are used to filter out one of the views so that the left eye only sees the left view, and the right eye only sees the right view. There are three key milestones during the development of anaglyph image technology. The first milestone is the publication of Wilhelm Rollmann's paper, which described the anaglyph principle from experiments involving red and blue lines, viewed using red and blue lenses [4]. In 1858, Joseph D'Almeida developed a system to project two images, through red and blue lenses, onto a single screen [1]. The images could then be viewed using glasses with red and blue lenses. The third key milestone was reached by Louis Du Hauron, who was at the leading edge of the development of colour photography. Du Hauron proposed a method for combining the two stereoscopic views on to a single print [1]. This eliminated the need for two lenses in projectors, and also enabled printed anaglyph images to be produced.

Stereoscopic images and drawings became quite popular during the nineteenth century. However, the development of technology allowing the capture and display of 3D movies proved to be highly challenging. Of course, much of this can be put down to the primitive nature of motion picture cameras (2D or 3D). Researchers and inventors were still struggling to produce reliable cameras capable of filming at rates greater than a few frames per second. However, some efforts were made, most notably by Frederick Varley and William Friese-Greene. Friese-Greene is a particularly notable figure, given the quotation on his tomb:

> His genius bestowed upon humanity the boon of commercial cinematography of which he was the first inventor and patentee.

Research by Brian Coe has subsequently shown that this claim is something of an exaggeration [5]. Friese-Greene was certainly a very active inventor, filing large numbers of patents. However, many of his inventions proved to be either impractical or unreliable. This description at least partially matches his stereoscopic camera, which he developed in collaboration with Frederick Varley in 1890. By all accounts, the camera was unreliable and only capable of capturing at a rate of a few frames per second, which is not enough to create a true sensation of movement. Furthermore, there is no record of any of the captured content being projected or displayed in a practical manner. Shortly after this development, Friese-Greene was declared bankrupt, and the realization of a commercially viable 3D system was not achieved until the next century.

1.1.2 Early Twentieth-Century Developments

The early twentieth century saw the arrival of a number of key technologies, the basic principles of which are still used in many of today's 3D technologies.

Shutter glasses, and polarized stereoscopic viewing technologies were all initially developed during this period. Shutter glasses and polarized lenses of course form the basis of the stereoscopic systems in use today. However, many initial developments were made using anaglyph systems.

One of the most reliable first appearances of 3D technology in the twentieth century is described by Lynde in a 1915 article in *The Moving Picture World* [6]. Lynde reports a demonstration of a new red–green anaglyphic stereoscopic movie projection system. This example is cited not just because of the reliability of the source, but also because it involved a major Hollywood director. Edwin S. Porter was a well-known movie maker, having directed one of the most important and popular silent movies, namely *The Great Train Robbery*. His presence therefore added a great deal of credibility to the event. The demonstration was made by Porter and the co-inventor of the new system, William E. Waddell, on June 10, 1915, in the Astor Theater, New York City. During the 1915 demonstration, a number of short features were shown, which were filmed as images $2\frac{1}{2}$ inches apart. Two projectors were used to display the left and right views in the red–green anaglyph format. Reaction to the demonstration appears to have been mixed, with Lynde claiming:

> Images shimmered like reflections on a lake and in its present form the method couldn't be commercial because it detracts from the plot.

Elsewhere in his article, his descriptions suggest that there were significant problems with synchronization between the two projectors, resulting in significant eye strain among the audience.

The next significant development was presented in December 1922 at the Selwyn Theatre in New York City. Here a new system called *Teleview* was presented by the inventors Laurens Hammond and William F. Cassidy. Their invention made use of what would now be called 'shutter glasses' or 'active stereo' technology. Two reels of film, representing the left and right eye views, were run through two separate, but synchronized projectors. One reel, however, was intended to be one frame behind the other, meaning that the left and right eye views were projected alternately onto the screen. Audience members sat behind their own viewing device, which featured a mechanical shutter, driven by a motor. The motor was synchronized to the projectors. The shutter blocked the left eye when the right eye view was displayed on the screen and vice versa. The 3D effect was reportedly much better than the 1915 demonstration described by Lynde [7]. An article in *The New York Times* stated:

> . . . those who went to the Selwyn last night were surprised, sometimes startled and often delighted with the vividness of the pictures that they saw and the unusual effects obtained by the use of the Teleview device.

However, the high cost of the equipment and the lack of attractive movie content meant that the system was not widely adopted.

Anaglyph filters tend to distort the colours perceived by the viewer, and are therefore not a good way of viewing high quality 3D images of video. The introduction of polarizing filters therefore provided a huge leap forward in stereoscopic video quality. Polarization sheets were first produced by William Bird Herapath, who discovered a method of forming polarizing crystals in 1852 [8]. However, the sheets, formed from Herapathite, proved to be of low quality, which limited their practical use for viewing stereoscopic images. Nevertheless, according to Lipton [1], John Anderton proposed the use of polarizing sheets for stereoscopic display in a patent, filed in 1891. The 1891 British patent is difficult to find. However, a US patent filed in 1893 refers to the original British version [9]. It is also possible to find Anderton's 1898 British patent for stereoscopy using polarizing filters [10]. This 1898 patent states that it provides improvements on the earlier 1891 British patent.

Edwin H. Land improved upon Herapath's original work, and filed the first of his many patents in 1929 [11]. Land was the man who founded the Polaroid company, and was also responsible for the invention of the Polaroid instant camera. In January 1936, he gave the first demonstration of 3D projection using polarizing filters at the Waldorf-Astoria Hotel [12]. During projection, two reels, carrying the right and left eye views respectively, had to be synchronized using an external motor. Another complication was that polarized light would not register on a normal matte white screen, and a silver screen was required to correctly display the two views. Despite these complexities, the quality appears to have been favourable according to the *New York Times* article:

> Observers were ushered into a seemingly living fairyland of forms and colours.

Land followed the first demonstration with another at the New York Museum of Science. It is this kind of polarizing technology that formed the basis for the systems that were used during the first 3D boom in the 1950s.

1.1.3 The 1950s 'Golden' Period

Although there was considerable interest by scientists and researchers in developing 3D movie technology, it was not until the early 1950s that the movie studios started to take a serious interest in the process. Movie studio executives were becoming increasingly concerned about the impact of television on cinema audiences. Between 1946 and 1952, weekly attendance figures in US cinemas had fallen from 82.4 million to 46 million [1]. In addition to this, the House Un-American Activities Committee (HUAC) had blacklisted talented scriptwriters, and directors [13]. The US movie industry was in crisis and needed to find something to attract people back into the cinemas. This dire situation led to experimentation with new technology, so that cinemas could provide an experience far superior to television.

According to Lipton [1], the first American movie to be made in colour and 3D was *Bwana Devil*, which was first screened on November 27, 1952. The movie was a hit, and grossed $100,000 in its first week. This woke the industry up to the potential of 3D, and led to the production of more 3D movies by the major studios. The production process was significantly limited by the fact that most of the major studios had neither the 3D cameras rigs, nor the expertise to shoot movies in 3D. This meant that no more than 45 3D movies were produced in a single year, compared to the typical 2D output of around three hundred per year. It is interesting to note that the entire 3D movie boom only lasted around nine months. By this time, the public had shown a clear preference for seeing 2D movies.

A number of reasons have been put forward for the failure of the 1950s 3D boom. One theory is that the quality of the feature films was particularly poor. Although 3D poor movies were made, there were a number of critically acclaimed movies made, such as *Dial M for Murder*, *House of Wax*, and *Kiss Me Kate*. Table 1.1 shows a selection of 3D movies from the 1950s, along with the average ratings given by users of the Internet Movie Database (IMDB). This reveals that while there were some very poor films made, there are also examples of excellence. It cannot be said that the quality of 3D movies was, on average, any worse than 2D features during the same period.

The reason for the rapid disillusionment by the public is most likely to be the significant eyestrain that many suffered after relatively short periods in the cinema. The eyestrain itself can be put down to a number of factors:

- *A lack of stereoscopic shooting expertise among film crews and inadequacies in understanding of optical problems* – Film crews in the 1950s had to learn about shooting in stereo very quickly. One of the key issues in avoiding eyestrain is to choose the optimal interaxial distance for the lenses. Too great a distance will result in eyestrain. The human interocular distance is typically 2½ inches, so distances greater than this are likely to be too much for viewers. Some rigs had a minimum interaxial distance of 3½ inches, which would have caused problems for some cinemagoers. In addition to this, there were sometimes shifts in the position of the optical axes of the lenses. In the worst cases, this led to vertical parallaxes. Even small vertical parallaxes cause significant muscular strain, as one eye is forced to look slightly upwards compared to the other in order to achieve fusion of the two views.
- *Poor quality control during film processing* – Another factor that can cause eyestrain over time is differences between the left and right views. If the two views are not processed in exactly the same way, then one view may be lighter than the other.
- *Projection systems too complex for most projectionists to handle* – During the 1950s, projectionists were not always able to ensure that the equipment was

Table 1.1 Selection of film titles released during the 1950s 3D boom, with movie ratings taken from IMDB

Title	Year	IMDB Rating
The French Line	1953	5.1/10
Taza, Son of Cochise	1954	5.5/10
Creature from the Black Lagoon	1954	6.9/10
Dial M for Murder	1954	8.1/10
Bwana Devil	1952	5.2/10
House of Wax	1953	7.0/10
Man in the Dark	1953	6.2/10
It Came from Outer Space	1953	6.6/10
Kiss Me Kate	1953	7.2/10
Hondo	1953	7.1/10
Miss Sadie Thompson	1953	6.0/10

correctly set up. Typical problems included differences in projection lens focal length, and differences in illumination of the two views. Furthermore, if films became damaged, it was common practice for projectionists to remove damaged frames. If the damaged frames were not removed from both views, then temporal synchronization would be lost. Also, if we consider that approximately 8% of the population cannot perceive stereo, then there is a risk that the projectionist might not have had stereo vision, making it very difficult for them to configure everything successfully. According to Lipton [1], Polaroid conducted a survey of one hundred stereo-equipped cinemas during 1953. The survey found that 25 of those theatres were poorly set up, causing significant eyestrain among viewers.
- *Corner cutting by cinemas* – The cinemas attempted to reduce their costs by using inferior screen coatings, and by using cheap polarizing glasses. This meant that images were not as bright and sharp as they should have been, and that the viewing experience using the glasses was poor.

When all of these difficulties became apparent, attention switched to alternative technologies for persuading television viewers to come to the cinema. The principal enhancement that the studios considered was changing the aspect ratio, making the screen wider. Examples of this technology include CinemaScope, which changed the standard 1.37:1 aspect ratio to 2.66:1.

1.1.4 The 1980s Revival and the Arrival of IMAX

Although the first 3D boom proved to be disappointingly short-lived, this was not the end of interest in 3D from the public or the movie industry. Certainly by the late 1970s, the movie industry was again concerned about new technology making home viewing more attractive than visiting a cinema. Mass market Video Cassette Recorders (VCR) had arrived, which allowed the public to watch movies at home repeatedly, and at a time of their choosing. This was perceived as a clear threat to movie industry revenues. However, it was clear from the problems described in Section 1.1.3 that more development work was required to make 3D technology commercially viable. Researchers continued to improve the technology, focusing on techniques that would provide a solution to problems such as view synchronization.

The Polaroid Vectograph was one of the candidates offering to provide a solution to view synchronization. It was invented by Edwin Land and Joseph Mahler, and allowed the two views to be placed on one film strip, rather than two [14]. The film was double-sided, and acted as a polarizing filter, as shown in Figure 1.3. This meant that no polarising filters were required for the projectors, resulting in a brighter image. The fact that the whole frame was used for each image also meant that there was no loss of resolution, and combining the two views on one film strip meant that there were no problems in synchronizing the two views. The original Vectograph worked on still images, and further development was needed to achieve motion picture capability. Shortly after the still image Vectograph was patented, Land patented a technique for a Vectograph capable of shooting movies [15]. However, by the time the technology was ready for commercial use, the 1950s boom was over, leaving Land with a promising product, but a non-existent market.

Colonel Robert Bernier developed techniques for projecting stereoscopic movies from a single projector [16], which is illustrated in Figure 1.4. Later, he developed a method for filming with a single camera, where the two views are projected onto the same frame of a standard camera film [17]. Like the Vectograph, Bernier's system solved the problem of view synchronization loss. As both views were shot onto a single strip of film, problems with film processing could also be avoided. Bernier's system was called SpaceVision, and was used to shoot a number of movies in 3D. In fact, it was systems similar to SpaceVision that were used during the 1980s 3D revival. It is interesting to note a number of disadvantages of systems such as SpaceVision:

- *Loss of image resolution* – The SpaceVision system requires that two views are projected onto a single frame of standard film. Inevitably, the physically smaller area per frame led to a loss in picture fidelity.
- *Loss of image brightness* – Unlike the Vectograph system, SpaceVision requires one set of polarizing lenses in front of the projectors, and another set to the used in the viewing glasses. Two sets of polarizing lenses led to the loss of a significant amount of light.

A number of alternative systems to SpaceVision were developed. For example, the Stereovision system placed the two views side-by-side on a single strip of film. Unlike SpaceVision, a 65mm print size was used,

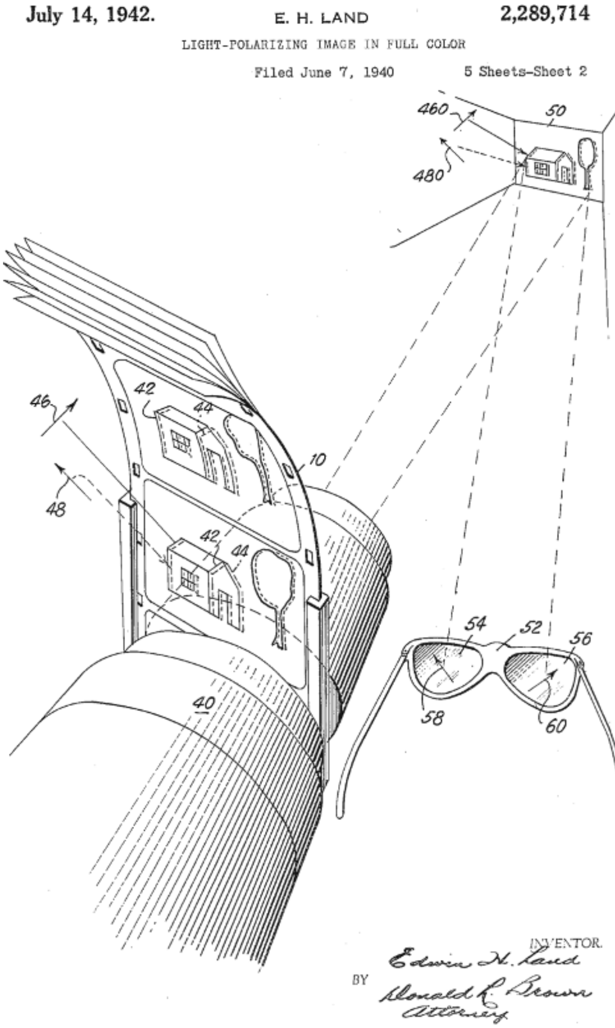


Figure 1.3 Diagram illustrating the basic principle of projection of Vectograph film, taken from Land's 1940 US Patent (14) (Reproduced with permission of the Polaroid Corporation)

meaning higher quality prints. However, for projection, it was often still necessary to reduce print to 35mm, meaning that the quality was not much different from that obtained from SpaceVision.

Although the 3D experience of the 1980s was an improvement over the 1950s' one, the inherent limitations of the SpaceVision type systems meant that 3D pictures were more blurry and were darker than for the same 2D movie. This limited the attraction of 3D for the general public. In addition to these inherent problems with the SpaceVision approach, it is clear from

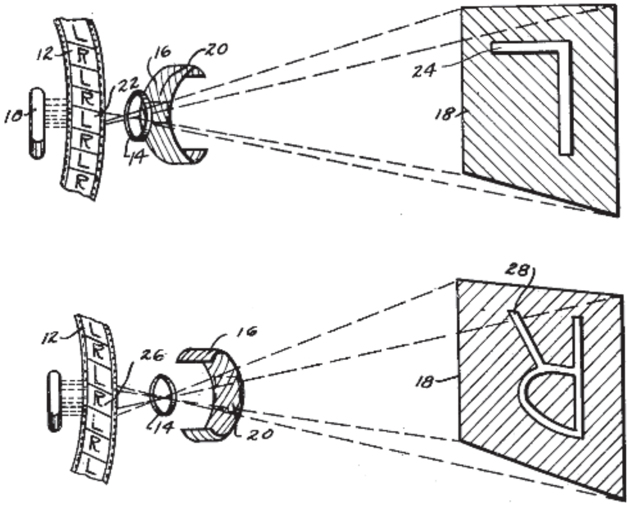


Figure 1.4 Illustration of Bernier's polarized projection system, taken from his 1949 US Patent (16). Left and right frames are contained on the same strip. A polarizing filter is flipped in synchronization with the frames as they pass through the projector (Source: US patent 2478891 (http://www.google.com/patents/US2478891))

expert observers' opinion that some films were poorly shot and processed. This quote from Lipton [1] about one production of the era is damning:

> The summer of 1981 saw the Filmways release of a western shot in Spain, titled *Comin' at Ya*. Production values were low, the acting was terrible, the dialogue moronic, the stereoscopic process, Optrix, was an optical catastrophe, and the filmmakers attempted to place every shot behind the heads of the audience. The stereoscopic system suffers from left and right images of unequal sharpness, severe vertical parallax, and strange watermark-like spots hovering in one or the other image field. The film has apparently been doing good business, and for that reason many major studios and producers are considering employing the three-dimensional medium. The situation brings to mind *Bwana Devil*.

Although not all of the productions were as bad as described by Lipton, in general, the quality problems meant that the new 3D revival was once again short-lived. Only a limited number of 3D movies were shot during the early 1980s, and the technology once again fell out of favour. One conclusion that could be drawn from this revival, is that for 3D to be a successful and enduring fixture, it needs to provide an experience that is not worse than that of 2D.

Development of 3D technology continued. One notable development was the Stereospace system, which was put together by Richard Vetter for United Artists Communications. The system provided a step forward in terms of quality by using two 65mm reels, with the left and right views on separate reels [18]. A magnetic strip with timing information was used to maintain synchronization between the two projectors required for presentation. A similar system to Stereospace was developed by Disney for inclusion in their theme parks. Disney's system was more than a straightforward stereo 3D projection system though. In addition to 3D, in-theatre effects were included, such as laser shows and smoke. One of the most notable 3D Disney productions was *Captain EO*, which starred Michael Jackson, and was directed by Francis Ford Coppola. Both the Stereospace and Disney systems were prone to jitter and vertical parallax. However, Disney's system proved to be more successful and enduring. This was in part because Disney controlled everything from film production to projection, ensuring that quality control was tightly enforced.

The system that became most widespread in the 1980s was IMAX. Special IMAX cinemas have now been set up all around the world by the IMAX Corporation. The company control all aspects of the process, from producing cameras to capture on to 70mm film, to the projection system in each theatre. Early IMAX systems were installed in the 1970s, while the first 3D systems were rolled out in 1986. The IMAX system provides high quality 2D colour picture quality on very large, immersive screens. Filming for IMAX 3D is carried out using special cameras, where lenses 2½ inches apart are used to capture material onto two different films. Two synchronized projectors are used, which have polarizing filters placed over their lenses. The movie is then viewed using glasses with polarized lenses. The only issue with IMAX theatres is that they are not currently as widely available as traditional cinemas. It has managed to carve out its own niche, but movies are still seen by most people in standard format cinema facilities.

1.1.5 The Twenty-first-Century Revival

The twenty-first century has seen a significant revival in 3D technology. Once again, the movie industry has taken the lead by reintroducing 3D into cinemas. Although – thanks to IMAX – 3D technology has never really gone away, the twenty-first century has seen more widespread adoption of the technology by mainstream cinemas. New threats are being faced by the movie industry, in the form of piracy. File sharing technologies have enabled new movies to be shared, and freely downloaded. This can sometimes occur before the official release date. In addition, whereas pirate copies used to be low quality, modern pirate movies can often be obtained in HD. Therefore, providing high quality 3D cinema releases is a way of giving consumers something that cannot be experienced at home. It also acts as an effective method of preventing illegal recordings taking place by bootleggers who take hidden cameras into the cinema.

Technology has improved significantly since the first 1950s boom, although there are still significant gaps in knowledge when it comes to 3D quality (see Chapter 6 for a discussion of 3D quality issues). The new generation of 3D movies have experienced significant commercial success. Currently, three of the top 10 grossing movies of all time have been released in 3D.[1] Of course, these films might have been as successful in 2D as they are in 3D. However, some market research by the International 3D Society has suggested that 3D movies have taken more profits in their 3D form, despite there being fewer 3D screens available [19]. Care should be taken with these figures, given that 3D movie tickets are more expensive than 2D movie tickets. However, the feedback obtained from consumers in the same research is very encouraging, with 74% of 3D movie viewers saying that 3D movies are better than 2D.

1.1.6 Auto-Stereoscopic

One of the inherent problems with many 3D display technologies is that they require the viewer to wear special glasses. Some viewers find such glasses off-putting, and therefore researchers have put considerable efforts into developing display technologies that do not require glasses. Such displays usually fall into the auto-stereoscopic category, which is described in this subsection.

Two main classes of auto-stereoscopic have been deployed:

- *Parallax barrier* – where a barrier is placed in front of the display. The barrier features a series of regularly spaced slits, which ensure that each eye sees a different area of the screen. The left and right views are carefully spliced together before display, so that when viewed, a 3D binocular effect is obtained. The problem with this technology is that it is extremely sensitive to head movements. Even a small head movement can lead to unpleasant effects, such as reverse stereo where the left eye sees the right eye view and vice versa.
- *Lenticular lens* – where the display is coated with an array of semi-cylindrical lenses. An image strip sits behind each lenticule, which contains a succession of partial views of the subject, from the extreme left camera position to the extreme right. Lenticular lenses are more expensive to produce than parallax barriers, but allow greater head movement by the viewer before unpleasant stereo artefacts can be seen.

These display types are discussed in more detail in Section 5.4.4. Here, we are concerned mainly with the historical progress of these technologies.

[1] As published by http://www.boxofficemojo.com on 21 February 2011.

One of the earliest publications describing the parallax barrier concept was produced by Auguste Berthier in 1896 [20]. Further developments were reported by Frederick Ives in 1902 [21, 22]. Clearly, at this stage, many of the developments were made using drawings, or still images. Motion picture cameras during this period were still relatively limited in their capability.

The credit for proposing lenticular lenses is often given to Gabriel Lippmann for his paper published in 1908 [23]. Once again, the technology was mainly limited to use on still images at this point. In 1930, Herbert Ives filed his patent describing a system for auto-stereoscopic projection of motion pictures, based on lenticular lenses [24]. The system principles are illustrated in Figure 1.5. As can be seen from Figure 1.5, the arrangement made use of multiple cameras and multiple projectors, which all required precise synchronization. It is difficult to find evidence of a successful implementation of the ideas put forward in this patent, and it seems likely that this was an idea that remained on the drawing board.

Over the years, many proposals for auto-stereoscopic displays have been put forward [25–29]. The result has been a steady and continual improvement in the quality of the displayed video, with increases in resolution, and the number of available viewpoints. Of course, the technology has not been fully exploited for cinema, because of the restrictions of seating position and head movement required for viewing stereoscopic video. There are now a large number of companies offering auto-stereoscopic displays, including Alioscopy, Dimension Technologies, and NewSight. The quality of these displays is very high, but they cost more than similar-sized displays that make use of shutter glasses. This and the head movement restrictions have so far prevented significant commercial success for this type of technology.

1.1.7 3D Television Broadcasts

Many of the historical developments described in this section have been targeted at providing 3D cinema experiences. Of course, there is a significant amount of overlap between cinema and television. However, television pictures need to be transmitted, and then shown on displays that are affordable for the average consumer. This section therefore describes some of the developments towards providing actual 3D television services for consumers.

One of the earliest examples of a 3D television was put together and demonstrated by John Logie Baird. He filed his first stereoscopic television patent in 1926 [30], and demonstrated the system in 1928 [31]. Transmission of two views of a person's face were demonstrated within the laboratory. The two views were alternately shown on the left and right sides of a single neon tube. A stereoscope device was used by viewers, which deployed a prism to direct the views to the eyes and therefore allow fusion. Of course, this system was a long way from being a practical solution. There were still many quality issues to solve with non-stereoscopic television, and

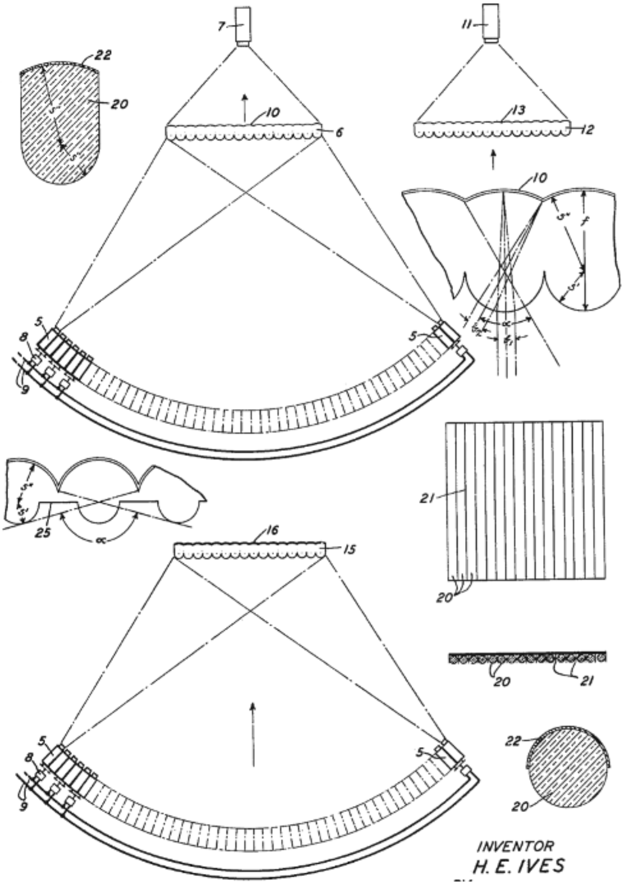


Figure 1.5 Drawings taken from Ives' patent for auto-stereoscopic projection (24). The drawings assume that fifty cameras were originally used during filming. In Fig. 1, one projector per original camera is used to project the views onto a translucent lenticular array. Light passes through this array to be captured by a single camera. In Fig. 2, material captured by the single camera is projected onto another lenticular array for viewing. Fig. 3 shows an alternative viewing scenario, where the fifty projectors are used in conjunction with a lenticular array backed with a white screen (Reprinted with the permission of Alcatel-Lucent USA Inc.)

the stereoscope viewing method was only suitable for a single viewer at a time.

The next 3DTV event of note was the experimental broadcast in the US on April 29, 1953. A trial broadcast of the series *Space Patrol* was conducted in Los Angeles at the National Association of Radio and Television Broadcasters 31st Annual Gathering. The experimental broadcast used polarizing technology to display the stereoscopic content [2]. The trial was conducted at the same time as the 1950s boom, to which television broadcasters were almost certainly paying attention. The problem with the technology based on the polarizing glasses was that it required consumers to purchase special expensive new television sets. This, and the rapid demise of 3D cinema in the 1950s, meant that interest in 3DTV died out.

Further broadcasts were made in the 1980s, following the introduction of a new system developed by Daniel Symmes, founder of the 3D Video Corporation. This new system allowed anaglyph format stereoscopic video to be broadcast and displayed on standard television systems. The system is described in patents filed in various countries, including Britain [32]. This particular patent includes diagrams demonstrating how it can be used with a National Television System Committee (NTSC) television system. Use of standard television sets is the key advantage of this approach. Consumers only need to obtain anaglyph glasses, which can be produced relatively cheaply. Although broadcasts in the US and Europe were initially popular, interest died out as consumers realized the limitations of anaglyph approaches (i.e. poor colour rendition).

During the 1980s, the Institut für Rundfunktechnik (IRT) worked on broadcast systems that relied on polarizing glasses. Systems were demonstrated at a number of trade fairs [33], and received very good responses from those who saw the technology. At the trade fairs, two projectors were used for the two polarized stereoscopic views. In addition to this system, IRT produced a smaller polarized display using standard television monitors, mirrors, and polarizing filters. Sand claimed that one of the limitations was a lack of satisfactory 3DTV cameras [33]. The viewing equipment would also have been significantly more expensive for consumers than a standard television set.

Some attempts were made in the 1980s and early 1990s to exploit the Pulfrich effect. The Pulfrich effect involves the use of special glasses, where one lens is darker than the other [34]. This reduced illumination seems to cause a delay in signal transmission to the brain, and therefore the difference between the two eyes results in a perceived depth disparity. The effects are fairly limited, as 3D can only be perceived for objects moving left of right across the screen. The direction of their depth change depends on the direction of movement of the object *and* the eye with the darkened lens. 3D effects are not seen if movement is vertical, with respect to the camera, or if the object moves towards or away from the camera. This effect has limited value for true 3D reproductions, and its use has largely been restricted to short sections of television programmes or advertising. The restriction on movement is too great for practical widespread use in 3DTV production.

Digital television standards were introduced during the 1990s, which largely relied upon the MPEG-2 video codec for compression. Although this codec could not originally handle stereoscopic video, it was updated to include the Multi-View Profile (MVP) [35]. However, this profile has not seen commercial usage so far, and is unlikely to do so. More recent standards are capable of coding multiple views with much greater efficiency than this MPEG-2 extension (see Section 3.2 for more details).

Significant 3DTV work has been carried out in Japan over the years. Sand's (1992) paper describes research at the University of Tokyo, where an auto-stereoscopic system with up to 24 channel multiple-views was displayed [33]. Nippon Hōsō Kyōkai (NHK) developed a 3D HDTV relay system, which was used in 1998 to transmit stereoscopic live images of the Nagano Winter Games via satellite to a large-scale demonstration venue in Tokyo [36].

The recent arrival of affordable 3D television sets means that 3DTV broadcasts are once again of interest to broadcasters. Many experimental broadcasts are taking place across the world, particularly using the frame compatible stereoscopic format (see Section 1.2.1). The frame compatible format can already be deployed using existing digital broadcast technology, such as DVB-S2. A commercial 3DTV channel was started by Sky Television in the United Kingdom in October 2010.

1.2 3D Video Formats

This section provides an overview of the most common video formats currently in use, or currently being considered by the research community. The formats are summarized in Table 1.2, which also describes the respective advantages and disadvantages. The frame compatible and service compatible formats are currently the most commonly used for commercial applications, and are supported by most consumer 3DTVs. The stereoscopic video formats are relatively straightforward, as they consist of only two colour video views. Therefore, most space in this section is given over to the description of the depth-based formats, and multi-view formats.

1.2.1 Frame Compatible and Service Compatible Stereoscopic Video

The first formats to be used in modern 3DTV systems will be stereoscopic, as this type of video does not require excessive transmission bit-rates, and is easier to capture than the other formats described in this section. All stereoscopic formats suffer from a lack of flexibility in terms of rendering. It is very difficult to change the users viewpoint, and to change the disparity between the views presented to the user. This prevents users from adjusting the presentation of the 3D video for comfortable viewing.

The *frame compatible* stereoscopic video format has been selected for use in the first generation of 3DTV systems. Its principal advantage is that the two views are packed together within a single video frame. This means that the format is compatible with most existing digital television systems, as no change of resolution or frame rate is required to support 3D.

Table 1.2 Summary of 3D video formats and their respective advantages and disadvantages

Format name	Description	Advantages	Disadvantages
Frame compatible stereoscopic video	Two views are combined into a single view by down-sampling horizontally and placing side-by-side	Compatible with existing 2D transmission systems, video decoders and 3DTVs	Involves loss of spatial resolution
Service compatible stereoscopic video	Two stereo views are transmitted jointly, but in separate frames, e.g. in consecutive frames	Compatible with existing 2D transmission systems, video decoders and 3DTVs. No spatial resolution loss	Double the bit-rate of frame compatible stereo
Colour-plus-depth	The depth is used to render two stereo views for the display	Depth can be compressed to a fraction of the size of a colour view	A single colour view means that occluded areas are unavailable during rendering
Multi-view video	Multiple views are captured usually from multiple cameras arranged in an array	Provides support for multi-view and holographic displays, in addition to free-viewpoint applications	Large bit-rate requirements. Depth information must be estimated at the renderer
Multi-view video plus depth	Similar to multi-view video, but with added depth information	Depth information makes rendering easier, and can mean fewer colour views are needed	Accurate depth information may be hard to generate
Layered depth video	Similar to colour plus depth, but with additional colour and depth data to compensate for occlusions	Improved quality over colour plus depth video	Additional band-width required

The disadvantage is that some loss of resolution is implicit, as the original views must be downsampled so that they may be packed into a single frame. Figure 1.6 shows some of the most common frame compatible formats. The side-by-side format was used in the first 3D broadcasts by Sky in the United Kingdom, and most 3DTVs support the format (i.e. they are able to accept side-by-side video as an input and render it as stereoscopic 3D video).

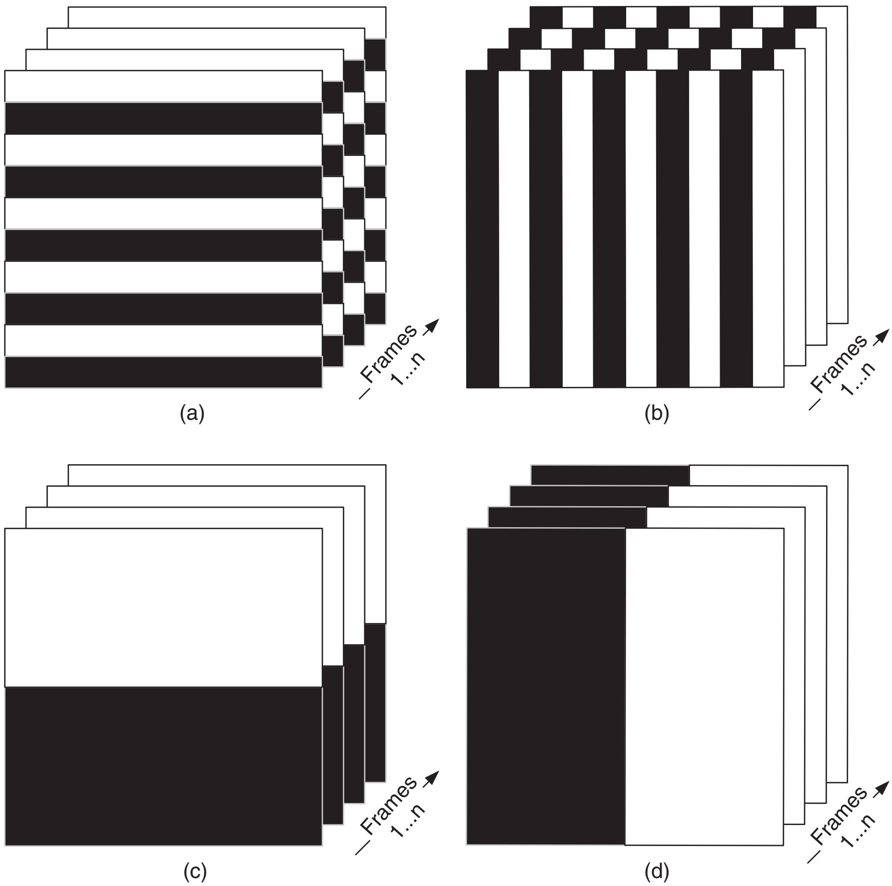


Figure 1.6 Formats for frame compatible stereoscopic video, where black represents pixels from the left view, and white represents pixels from the right view. All arrangements require downsampling of the original left and right views for packing within one frame. (a) Left and right view pixel lines are spliced together. (b) Left and right view pixel columns are spliced together. (c) Left and right views are placed at the top and bottom of the frame. (d) Left and right views are placed side-by-side within the frame

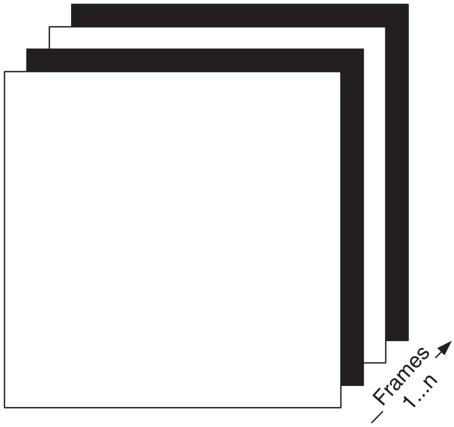


Figure 1.7 Example of frame compatible stereoscopic video, with the left and right frames packed into sequential frames

Service compatible stereoscopic video will be used in the second generation of 3DTV broadcasts, and has been specified for use in the 3D Blu-Ray standard. This format removes the limitations of frame compatible video by presenting the frames separately, as shown in Figure 1.7. Full High Definition resolution can therefore be achieved with this format. Its disadvantage is that it requires changes to the system used to transport the video, requiring users to be supplied with new set-top boxes or disc players.

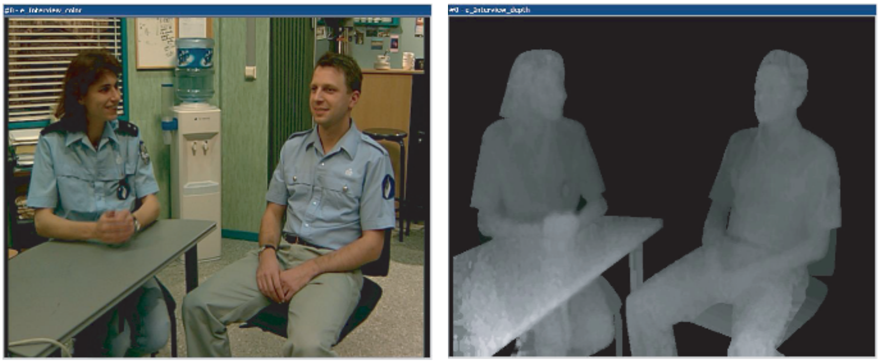
1.2.2 Colour-Plus-Depth

Depth maps (also known as range images) have been of considerable interest to 3D video researchers in recent years. Depth map sequences usually have similar spatio-temporal resolution as the colour image sequence with which they are associated. The depth maps can be stored as 8-bit per pixel grey values, where grey value 0 specifies the furthest position from the camera, and a grey level of 255 specifies the closest position to the camera. This depth map representation can be mapped to real, metric depth values. To support this, the grey levels are normalized into two main depth clipping plains. The near clipping plane Z_{near} (grey level 255), represents the smallest metric depth value Z. The far clipping plane Z_{far} (grey level 0), represents the largest metric depth value Z. In case of linear quantization of depth, the intermediate depth values can be calculated using the following equation:

$$Z = Z_{far} + v\left(\frac{Z_{near} - Z_{far}}{255}\right) \text{ with } v \in [0, \ldots, 255] \qquad (1.1)$$

where v specifies the respective grey level value.

An example colour-plus-depth representation can be seen in Fig. 1.8.



(a) (b)

Figure 1.8 Example colour-plus-depth video sequence commonly used by video coding experts for testing video compression algorithms, called Interview, produced in the course of ATTEST Project (a) Colour sequence. (b) Depth sequence (Reproduced with permission of the ATTEST Consortium)

Depth-Image-Based (DIBR) can be used to synthesize two views for the left and right eyes using colour image sequences, and the corresponding per-pixel depth map [37, 38]. This process requires two main steps: re-projection of original image point into 3-D space using depth information, and projection of the 3-D space points into the image planes of the left and right views.

The advantages of using colour-plus-depth map representation of stereoscopic video compared to video generated with a stereo camera pair can be summarized as follows:

- 3D rendering can be optimized for different stereoscopic displays and viewing scenarios to yield a disparity that is comfortable to the eye.
- Head-Motion Parallax (HMP) may be supported, which provides an additional 3D depth clue. This format also partially overcomes the viewing angle limitation of traditional stereoscopic camera set-ups.
- Most depth information does not have high frequency components. Thus, the depth sequence can be efficiently compressed with existing compression standards [39], and will require only limited space and bandwidth compared to that required by the colour image sequence.
- Photo-metrical asymmetries (e.g. in terms of brightness, contrast or colour) between the left and the right eyes, will be eliminated. Thus, the associated eyestrain problems will be avoided.
- Depth can be used in 3D post production (e.g. fine tuning of depth to eliminate stereoscopic artefacts that may occur during filming).

However, there are a number of drawbacks associated with this representation. The disadvantages and possible solutions are as follows:

- The quality of the rendered stereoscopic views depends on the accuracy of the per-pixel depth values. Therefore, the effects of compression and transmission of depth maps on the perceived quality need to be carefully considered (see Chapter 6).
- Objects that are visible in the rendered left and right views may be occluded in the original view. An example of this is shown in Figure 1.9. This phenomenon is also known as exposure and dis-occlusion in Computer Graphics (CG) [38]. This effect can be minimized using Layered Depth Video (LDV), where more than one pair of colour-plus-depth sequences is transmitted depending on the requirements of the expected quality. However, this approach demands more storage and bandwidth to be used in communication applications. In addition, different hole-filling algorithms (e.g. linear interpolation of foreground and background colour, background colour extrapolation, mirroring of background colour information) can be utilized to recover the occluded areas of the original image sequence [39]. Moreover, pre-processing/smoothing of depth maps (e.g. use of a Gaussian filter) will avoid this occlusion problem. However, this approach will lead to some distortions of the rendered 3D video scene.
- Certain atmospheric effects (e.g. fog, smoke), and semi-transparent objects are often poorly handled with this approach.

1.2.3 Multi-View Video

Multi-View Video (MVV) is a format required to support many promising 3D video applications, such as FVV and holographic displays (rendered as stereoscopic video in the simplest case, or multi-view auto-stereoscopic video in more advanced scenarios). Section 2.2.3 describes the camera arrays used

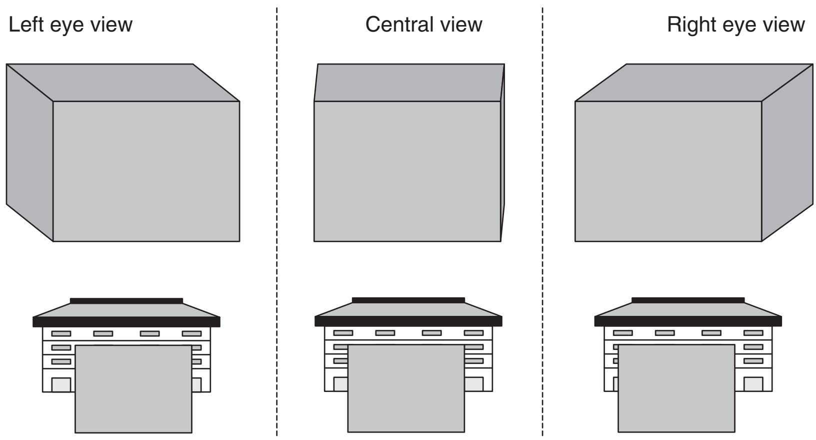


Figure 1.9 Illustration of occlusion problems with the colour-plus-depth format, where only a central viewpoint is stored, rather than the left and right views

to capture MVV in detail. The arrays provide several calibrated viewpoints of the same 3D scene, captured simultaneously. Multi-view video application, unlike conventional video applications (e.g. 2D HDTV or stereoscopic 3DTV), promise to offer a high level of interaction between the user and the content. FVV allows the user to navigate throughout the video scene.

Following extensive investigations into 3D video (mainly within the MPEG 3DV Ad Hoc group), multi-view video was widely recognized as a powerful video format and the need for a standardization activity for the compression of multi-view videos was identified. The results of these standardization activities are discussed in more detail in Section 3.3.2.

1.2.4 Multi-View Plus Depth Video

The colour-plus-depth format introduced in Section 1.2.2 provides very limited free-viewpoint functionality. The head motion parallax in the colour-plus-depth format can adjust the rendered video within a very narrow navigation range [40]. The multi-view video explained in Section 1.2.3 can theoretically provide a much larger scene navigation range. However, rendering using only colour views means that depth estimation must be performed at the receiver. Depth estimation errors can cause significant visual artefacts in the rendered video. Therefore, multi-view video itself cannot always provide good quality free-viewpoint rendering, particularly if the inter-camera baselines are large.

Multi-View plus Depth (MVD) is an extension of both the multi-view video format and the video-plus-depth format. It consists of N colour texture viewpoints and N corresponding per-pixel depth map sequences. The inclusion of depth extends the free-viewpoint navigation range [41, 42], as well as smoothness of scene navigation. Using the MVD format is it possible to render intermediate virtual viewpoints at any position, provided the spatial position vectors (rotational and translational matrices) and camera parameters of the target virtual viewpoint are present. MVD is a very powerful scene representation format enabling full feature 3DTV supporting multiple simultaneous viewers. Figure 1.10 depicts the multi-view plus depth representation, and shows the synthesis of arbitrary viewpoints with this representation.

Multi-view plus depth does not necessarily carry excessive additional transmission overhead compared to multi-view video, as the multi-view depth map sequences can be treated as colour texture videos and can be efficiently compressed using multi-view coding tools. These multi-view coding tools are described in Section 3.3.3. Extensive coding results can be found in [40] and [41].

1.2.5 Layered Depth Video

The concept of Layered Depth Image (LDI) or Layered Depth Video (LDV) was introduced in [43], and can facilitate efficient image-based 3D scene rendering. A layered depth image represents a scene using an array of pixels

Figure 1.10 Multi-view plus depth format video allows synthesis of virtual views using colour and depth information from nearby viewpoints (Reproduced with permission of Microsoft)

viewed from a single camera position [44]. Each pixel in the array consists of colour, luminance, depth, and other data that assists 3D scene rendering. According to [44], three key characteristics of LDV are:

- It contains multiple layers at each pixel position.
- The distribution of pixels in the back layer is sparse.
- Each pixel has multiple attribute values.

As shown in Figure 1.11, intersecting points between the rays emanating from a reference camera, and an object are stored [44]. The information stored for each intersecting point consists of colour and depth information. The conceptual diagram of a layered depth image is depicted in Figure 1.11 [43, 44]. In Figure 1.11, the intersecting points closest to the reference camera form the first layer of the LDV. The second nearest are used in the second layer, and so on. Consequently, each Layered Depth Pixel (LDP) has a different number of Depth Pixels (DPs), each of which contain colour, depth, and auxiliary information for reconstruction of arbitrary viewpoints of the 3D scene [44]. For the specific example shown in Figure 1.11, LDP 4 has four layers, which contain data pertaining to the intersecting points between Ray D and the object. A single LDI is created per time instant (i.e. the LDI at time t is composed of the images and depth map values belonging to the same time instant from all different views).

Conventional video codecs, such as MPEG-4 AVC, are optimized to remove temporal and spatial redundancies, which are common for most 2D video sequences. However, the layers of LDV do not exhibit the same characteristics, in terms of spatial and temporal correlation. Conventional video coding tools are inefficient when coding the layers of an LDV. Therefore, although LDV is an excellent format for rendering, the lack of an efficient compression scheme is likely to hold it back.

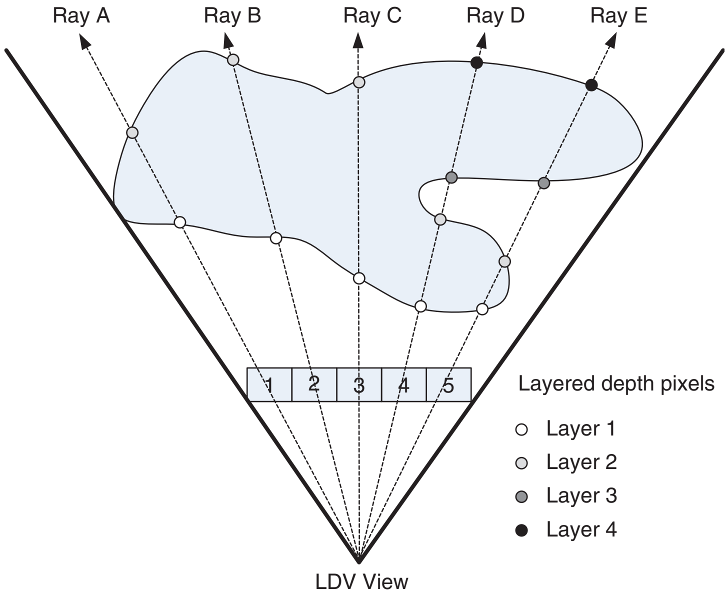


Figure 1.11 Concept of layered depth video

1.3 3D Video Application Scenarios

Although this book focuses mainly on 3DTV, the techniques and technologies described in this book are relevant to a range of application scenarios. This section describes some of the most interesting applications, including broadcast, mobile 3DTV, 3D video streaming, immersive video-conferencing, and remote applications (e.g. remote surgery, control of robots).

1.3.1 3DTV Broadcast Systems

Broadcast remains the most efficient way of delivering high quality video content to large numbers of people. In Europe, the digital television standards are being revised to enable support for 3DTV channels. Frame compatible format video is already supported by the various Digital Video Broadcasting (DVB) standards, while work is under way to enable support for service compatible video. On 1 October 2010, British Sky Broadcasting launched a 3D channel that could be decoded using their standard High Definition (HD) DVB-S2 set-top box. Compatibility with the existing technology is made possible by the use of frame compatible 3D, where the two views are downsampled in the horizontal direction, so that they can both be placed within one HD frame.

Figure 1.12 A Sony 3DTV, along with a consumer-level stereoscopic camcorder (Courtesy of Sony)

Unfortunately, the range of content currently available on 3DTV channels is somewhat limited. Much of the existing content is made up of 3D Hollywood movies. However, filming of live sports events has also been performed using stereoscopic camera rigs.

3D broadcasts have also taken place in other countries, including Japan, the USA, and Australia. As more cost-effective 3DTV sets are produced by mainstream manufacturers (such as the set shown in Figure 1.12), broadcasters are likely to introduce new premium 3DTV channels to increase their revenues. An interesting point to note about 3DTV services is that the success of 3D movies may be a significant driver of 3DTV sales. While the movie industry may have started 3D productions to keep people away from their televisions, the end result could be the successful establishment of 3D services for the home. In the 1950s and 1980s, this would have been unthinkable due to the costs. But good quality 3D televisions prices are dropping to prices low enough for the average consumer.

1.3.2 Mobile 3DTV

One of the original ideas of stereoscopic cinema was that it would provide a greater sense of realism and immersion. It may therefore appear to be a strange choice to try and introduce 3D services on to small mobile devices. Also, the use of 3D glasses in many mobile scenarios is likely to be inconvenient for users of the service. However, because mobile devices are typically viewed by an individual, rather than by a group of people, they are uniquely suited to many auto-stereoscopic approaches (see Section 5.4.4 for a more detailed description of this technology), which often require that the viewer positions themselves within a ‘sweet spot’ to view the video in 3D.

Commercial mobile devices with 3D displays are already starting to find their way to the market. Although not a mobile phone, Nintendo’s 3D DS portable games platform is an interesting example of a small device with a 3D display. Manufacturers such as Samsung and Nokia have already begun to produce experimental mobile phones with 3D capability. In addition to this, research funded by the European Union has brought together coalitions of industry and academia to work on the following two projects:

- *3D PHONE* – examines various topics relevant to future 3D mobile phones, such as auto-stereoscopic displays, stereoscopic capture of video, 3D user interfaces, and compression of 3D video. The project partners are Bilkent University, Fraunhofer Heinrich-Hertz-Institute (HHI), Holografika, TAT (The Astonishing Tribe), Telefonica and University of Helsinki.
- *MOBILE 3DTV* – investigates how to deliver 3DTV services to a mobile hone with an auto-stereoscopic display, over DVB-H. Special attention is paid in the project to quality issues. The project partners are Tampere University of Technology, Technische Universitt Ilmenau, Middle East Technical University, Fraunhofer HHI, Multimedia Solutions Ltd and Tamlink Innovation-Research-Development Ltd.

Both of the above-named projects have successfully demonstrated 3D mobile phone technology, and have carried out user trials showing strong user acceptance.

1.3.3 3D Video on Demand

3D Video on Demand involves the delivery of video over Internet Protocol (IP)-based networks, when requested by the user. In the United Kingdom, Video on Demand has become popular, and has been branded for the public using names such as iPlayer, and Catch-Up TV. Many television sets are now being produced that are capable of being connected to the Internet to view VoD. The Consumer Electronics Show (CES) in 2011 featured a large number of so-called ‘connected TVs’, and Forrester Research predicts that [45]:

> Internet connected TVs will continue their steady penetration into consumers’ homes, in large part due to retailers’ commitment to only sell connected TVs in the future.

Forrester estimates that 58 per cent of all TVs sold will be ‘connected’ by 2015. Of course, if these connected TVs are also capable of displaying 3D, then the extension of these 2D video services to 3D is relatively trivial. Service providers can either use the frame compatible formats currently in use for

European DVB broadcasts, or frame sequential service compatible format (see Section 1.2). Problems with 3D VoD are likely to arise, when more views are required to support multi-view displays, which is clearly more demanding in terms of network bandwidth requirements. These transmission issues are discussed in more detail in Section 4.3.2.

1.3.4 3D Immersive Video-Conferencing

A number of companies have already produced high-end video-conferencing facilities, which attempt to provide greater sense of the remote participants being in the same room. Hewlett Packard and Cisco are two of the main proponents of such high-end facilities. Both companies offer systems that exploit High Definition video-conferencing, rendered on multiple large screens. Cicso's Telepresence 3000 system is shown in Figure 1.13, which demonstrates the basic configuration of such high end systems. People within the room are made to feel that the remote participants are present in the same room.

The introduction of 3D to such high-end video-conferencing systems would seem like a natural step forward in terms of immersion. However, there are some issues that currently limit the effectiveness of 3D for video-conferencing. These issues become apparent when we consider the currently available 3D display technology:

- *Passive or active stereo displays* – require that glasses are used. This may cause problems for participants to interact effectively, as it may inhibit eye contact with remote meeting participants.
- *Auto-stereoscopic displays* – typically feature sweet spots. Although some displays can be created with multiple sweet spots, viewers movement is

Figure 1.13 Cisco's Telepresence 3000 system, featuring high definition video-conferencing (Courtesy of Cisco Systems Inc.)

still heavily constrained, making it difficult to interact naturally with both remote participants and colleagues in the same room. There are also likely to be significant complexities in setting up multiple 3D displays, so that the sweet spots all converge on particular seating spaces.

The answers to these problems will likely either arise come from improving auto-stereoscopic displays, or from image processing techniques that are capable of making 3D glasses invisible to remote participants. However, this last image processing solution does not solve the problems with interacting with people in the same room.

1.3.5 Remote Applications

The medical community has conducted a number of studies into the use of 3D video for surgery. Stereoscopic video can be used for surgery that requires cameras (e.g. remote surgery, or keyhole surgery). 3D video can provide more accurate perception of depth, allowing the surgeons to make more accurate incisions. This in turn can lead to fewer complications arising from surgery, and faster recovery times.

1.4 Motivation

It is clear from the history of 3D video (see Section 1.1), that while public popularity of the technology has waxed and waned, development has continued. The repeated revivals point to real public interest in 3D technology. Previous failures have been caused by significant quality problems in the production and projection of 3D movies, and a lack of good affordable 3D technology for the home.

While significant improvements in camera and projection technology have been made over the years, as well as the introduction of affordable 3DTVs for the home, one of the most important developments has been the introduction of digital video processing technology. Many of the problems associated with previous 3D booms can be put down to production and projection problems that are difficult to spot and to fix with the naked eye. Digital technology allows analysis and correction of stereoscopic errors to sub-pixel accuracy. Examples of the benefits that can be made by digital technology are as follows:

- *Colour correction* – differences in contrast, and colour saturation between views may be precisely adjusted to ensure that stereoscopic views match (see Section 2.3.1). This reduces eye fatigue issues.
- *Rectification* – any problems with camera viewpoints can be corrected, as described in Section 2.3.1. An example of such a problem is when one camera, from a stereoscopic pair, points slightly upwards or downwards compared to the other.

- *Extraction of depth* – can be achieved using the kind of algorithms described in Section 2.3.2. Depth allows better free-viewpoint rendering, and also fine-tuning of disparities between viewpoints so that the content can be optimized to viewing position and screen size.

There are therefore many reasons to believe that the current revival in 3D is one that will prove more enduring than previous booms.

1.5 Overview of the Book

This book is split into seven chapters, most of which can be directly related to blocks in the 3DTV chain shown in Figure 1.14. Following this introduction, the book works its way through the 3DTV chain, starting with the capture of 3D video in *Chapter 2*. This chapter discusses capture of the various different formats for 3D video (see Section 1.2), and the camera technology currently available to perform the capture. Following capture, some image processing steps are often required. First the videos must be rectified and colour matched to prevent visual fatigue and to enhance the efficiency of multi-view video compression. Finally, depth must be extracted. Approaches for depth extraction are discussed in Section 2.3.2.

The captured and post-processed video is incredibly large, and therefore compression, as discussed in *Chapter 3*, is essential. The basics of video compression are briefly covered, before looking specifically into compression of the formats described in Section 1.2. Existing standards are covered, and the chapter also discusses current and future standardization efforts.

Chapter 4 examines approaches for transmitting 3D video. Transmission of stereoscopic video over broadcast or IP-based networks can be achieved using relatively straightforward extensions of existing 2D transmission schemes. The real problems occur when multi-view or holographic display support is required. In such cases, multiple HD views must be transported, placing a significant strain on communication networks. Some holographic displays require between eight and 64 views, which requires a bandwidth greater than can be provided by existing networks.

After delivery and decompression of the video, rendering and display must be performed. *Chapter 5* examines a variety of 3D rendering approaches.

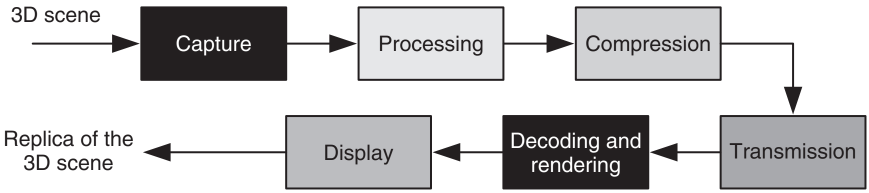


Figure 1.14 Basic building blocks for 3DTV

In addition, several adaptation types for 3D and multi-view video are also outlined here, such as network adaptation, context adaptation and adaptation to users preferences. Finally, the principles behind existing stereoscopic displays and more advanced displays technologies, which may become popular in the future (such as light-field and holographic displays), are looked at in this chapter.

Chapter 6 discusses quality and human perception of 3D video. This has become an increasingly important topic for 3D video, and understanding of these issues requires discussion of the physiology and psychological elements of the Human Visual System (HVS). The chapter discusses how to carry out subjective viewing tests in order to obtain reliable quality results. It also discusses methods for evaluating 3D video using objective metrics. This kind of approach is significantly faster than carrying out subjective testing, but may not always produce reliable results.

The final section of the book, *Chapter 7*, summarizes the book, and provides pointers to which improvements are necessary, and which improvements are likely to be made, in future years. It also makes a brief assessment concerning the continued success of 3D video applications.

References

[1] Lipton, L. (1982) *Foundations of the Stereoscopic Cinema: A Study in Depth*, Van Nostrand Reinhold, Amsterdam.

[2] Hayes, R.M. (1998) *3-D Movies: A History and Filmography of Stereoscopic Cinema*, McFarland & Company, Jefferson, NC.

[3] Wheatstone, C. (1838) 'Contributions to the physiology of vision. Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision', *Philosophical Transactions of the Royal Society of London*, **128**, 371–394, June.

[4] Rollmann, W. (1853) 'Notiz zur Stereoskopie', *Annalen der Physik*, **165** (6), 350–351.

[5] Coe, B. (1969) 'William Friese Greene and the origins of cinematography', *Screen*, **10** (2), 25–41.

[6] Lynde, D. (1915) 'Stereoscopic pictures screened', *The Moving Picture World*, p. 2072, 26 June.

[7] Screen (1922) 'Vivid pictures startle', *The New York Times*, 28 December.

[8] Herapath, W.B. (1852) 'XXVI. On the optical properties of a newly-discovered salt of quinine, which crystalline substance possesses the power of polarizing a ray of light, like tourmaline, and at certain angles of rotation of depolarizing it, like selenite', *Philosophical Magazine Series 4*, **3** (17), 161–173.

[9] John Anderton, 'Method by which pictures projected upon screens by magic lanterns are seen in relief', Patent number 542321, July 1895.

[10] John Anderton, 'An improved method or system by means of which pictures projected upon a screen from an optical or magic lanterns or lanterns are seen in relief or with stereoscopic effect', Patent number GB189801835 (A), November 1898.

[11] Edwin Herbert Land, 'Polarizing refracting bodies', Patent number US1918848 (A), July 1933.
[12] *NY Times* (1936) 'New ''glass'' cuts glare of light; aid to movies and science seen; inventor shows polarizing substance which, with spectacles of same material, dims brilliance of auto lamps – may be used in beauty treatments.', January.
[13] *Encyclopedia Britannica* (2011) 'House Un-American Activities Committee (HUAC)' available at: www.britannica.uk.
[14] Land, E.H. (1940) 'Vectographs: Images in terms of vectorial inequality and their application in three-dimensional representation', *Journal of Optical Society of America*, **30** (6), 230–238.
[15] Edwin Herbert Land, 'Light-polarizing image in full color', Patent number 2289714, July 1942.
[16] Robert V. Bernier, 'Three-dimensional adapter for motion-picture projectors', Patent number US2478891 (A), August 1949.
[17] Robert V. Bernier, 'Stereoscopic camera', Patent number CA787118 (A), June 1968.
[18] *BOXOFFICE Magazine* (1982) 'UACs stereospace', November.
[19] International 3D Society (2010) '3D movie fans expand box office says international 3D society study', March.
[20] Berthier, A. (1896) 'Images stéréoscopiques de grand format', *Cosmos*, **34**, 205–210.
[21] Ives, F.E. (1902) 'A novel stereogram', *Journal of the Franklin Institute*, **153** (1), 51–52.
[22] Frederick E. Ives, 'Parallax stereogram and process of making same', Patent number 725567, April 1903.
[23] Lippmann, M.G. (1908) 'Épreuves réversibles. Photographies intégrales', *Comptes Rendus de l'Académie des Sciences*, **9** (146), 446–451.
[24] Herbert E. Ives, 'Projection of stereoscopic pictures', Patent number 1883290, October 1932.
[25] Clarence W. Kanolt, 'Stereoscopic picture', Patent number 2140702, December 1938.
[26] D. Gabor, 'System of photography and projection in relief', Patent number 2351032, June 1944.
[27] F. Savoye, 'Stereoscopic motion-picture', Patent number 2441674, May 1948.
[28] E.H.V. Noaillon, 'Art of making cinematographic projections', Patent number 1772782, August 1930.
[29] R.B. Collender, 'Three dimensional unaided viewing method and apparatus', Patent number 3178720, April 1965.
[30] John Logie Baird, 'Improvements in or relating to stereoscopic pictures', Patent number GB292365 (A), June 1928.
[31] Tiltman, R.F. (1928) 'How ''stereoscopic'' television is shown', *Radio News*, November.
[32] James Frank Butterfield, Willard Stanton Alger, and Daniel Leslie Symmes, 'Stereoscopic television system', Patent number GB2114395 (A), August 1983.
[33] Sand, R. (1992) '3-DTV-a review of recent and current developments', in *IEE Colloquium on Stereoscopic Television*, London, October, p. 1, IRT stereo system.

[34] Terry D. Beard, ‘Low differential 3-D viewer glasses and method’, Patent number EP0325019 (A1), July 1989.
[35] Imaizumi, H. and Luthra, A. (2002) ‘Stereoscopic video compression standard: MPEG-2 multiview profile’, in *Three-Dimensional Television, Video, and Display Technologies*, Springer-Verlag, New York, pp. 169–181.
[36] Yuyama, I. and Okui, M. (2002) ‘Stereoscopic HDTV’, in *Three-Dimensional Television, Video, and Display Technologies*, Springer-Verlag, New York, pp. 3–34.
[37] McMillan, L. (1997) ‘An image-based approach to three-dimensional computer graphics’, PhD thesis, University of North Carolina at Chapel Hill, Chapel Hill, NC.
[38] Mark, W.R. (1999) ‘Post-rendering 3D image warping: visibility, reconstruction, and performance for depth-image warping’, PhD thesis, University of North Carolina at Chapel Hill, Chapel Hill, NC, April.
[39] Fehn, C. (2004) ‘Depth-image-based rendering (DIBR), compression and transmission for a new approach on 3DTV’, *Proceedings of SPIE Conference on Stereoscopic Displays and Virtual Reality Systems XI*, **5291**, 93–104.
[40] Merkle, P., Smolic, A., Müller, K. and Wiegand, T. (2007) ‘Multi-view video plus depth representation and coding’, *IEEE International Conference on Image Processing*, **1**, 201–204.
[41] Kauff, P., Atzpadina, N., Fehna, C., Miller, M., Schreer, O., Smolic, A. and Tanger, R. (2007) ‘Depth map creation and image-based rendering for advanced 3DTV services providing interoperability and scalability’, *Signal Processing: Image Communication*, **22** (2), 217–234.
[42] Zitnick, C.L., Kang, S.B., Uyttendaele, M., Winder, S. and Szeliski, R. (2004) ‘High quality video view interpolation using a layered representation’, *ACM Transactions on Graphics*, **23**, 600–608.
[43] Shade, J., Gortler, S.J., He, L.W. and Szeliski, R. (1998) ‘Layered depth images’, *SIGGRAPH*, **July**, 231–242.
[44] Yoon, S. and Ho, Y.S. (2007) ‘Multiple color and depth video coding using a hierarchical representation’, *IEEE Transactions on Circuits and Systems for Video Technology*, special issue on multi-view video coding and 3DTV, **17** (11), 1450–1460.
[45] Warman, M. (2011) ‘CES 2011: tablets and televisions set to dominate a connected show’, *The Daily Telegraph*, 5 January.