The 3d video production film studies essay

INTRODUCTION

Throughout the last ten years, we have seen a remarkable progress in the field of information and communication technologies. Statements have been made in terms of advantages and disadvantages for these technologies and the affect that will have to the general public. Regardless of the fact that some of us find it challenging to keep up with the fast growth of the technology, most of us take advantage of the features and services that are provided nowadays.

Background

When about a century ago telephony was discovered there was no way to visualize the profound influence it would have on our everyday lives nowadays. Nor was it possible to forecast the future of the television introduced some three decades later. Right from the start, television has gone through a significant revolution due to the introduction of colour in the 1950, and the continuous improvements in camera and display technologies. With the help of the three delivery channels, terrestrial, cable and satellite, respectable quality analogue television is today provided to almost every household. We are also come into contact with a shift to digital television, offering greater image and sound quality, bandwidth efficiency and quality of service requirements. With the use of this technology the transmission of four digital television programs is equivalent with the bandwidth of one analogue channel [1]. But the evolution does not stop there, even if digital television was fully digital, a new even better transmission technology is introduced, the High Definition Television. Moreover since the rebirth of the telephone with the cellular phones, where it was original expected to be used as the business tool, a small number of people forecast the broad use it was to gain in the private market. Currently cellular telephony holds the first place as the fastest growing product of any telecommunications technology developed in the 21th century and also has the broadest worldwide reach. Right from the start of this century the cellular phone networks in some countries have already exceeded the corresponding number in fixed networks [2]. Cellular phones are attractive because of the anywhere and anytime concept. This is to a certain degree is linked to what the Internet provides. Throughout countless networks and with about 30 million users [3], Internet offers access to gigantic amounts of services and information anytime anyplace, as long as there is the ability of connecting to the network. As it is an information environment made by its users lacking the control of a central authority, there are practically no restrictions on what one can find while searching through this jungle of information. This capability has proven exceptionally attractive to the users, and one might say that this interaction possibly is one of the main reasons to the growth of the Internet. The achievements in the areas of digital television, cellular phones and the Internet are all motivated by the conversion from analog to digital format, which makes it easier to complete processing, accumulation and transport of information. Networks that traditionally were planned and built for certain types of traffic such as voice, data, text or video are revolving into digital networks that can be considered as all-purpose networks providing transportation of any type of traffic. This shift from analog to digital visual communication is also known as the digital revolution. The digitalization of information, information processing tools and transport systems are perhaps the most powerful features of development in the whole area of telecommunications. And as the communication networks are one after the other digitized, this process leads to the convergence between telecommunications and other closely related such as computing, broadcasting and publishing. Despite the fact that digital revolution motivates the convergence forward, the necessity for respectable and efficient compression techniques increase. A number of standards have been introduced and broadly accepted, but it was not until 1998, with the standardization of the MPEG-4 standard, that the requirements for interactivity, error resilience and ownership protection, was directly addressed. MPEG-4 audio-visual standard is aiming to establish a universal and effective coding of dissimilar forms of audio-visual data, called audio-visual objects. With revolution also modifications come, and with modifications comes improvements. Even if it might be an area of discussion whether what we now are witnessing should be called a digital revolution or digital evolution, the fact is that we have come to a point where bandwidth problems are resolved and digital equipment have turned into common property. We are reachable anytime and anyplace, and we can access Internet from almost any place where we may find ourselves, and we are experiencing the benefits and potentials that come from the digital revolution. So what is following? Although future rarely plays out the way we imagine it will, the wish for an increased degree of realism in today's two-dimensional imagery has been expressed from various fields. For example, in video conferencing, telemedicine and training of pilots or other groups working with simulators, a perception of depth, so natural in daily life, would greatly enhance the experience.

Motivation

The general public has been enthusiastic about three dimensional images ever since 1833 when the first stereoscope was introduced by Sir Charles Wheatstone. Although the popularity since then has been fluctuating, the technology has experienced peaks for each new development. After its last heights in the 1950s and 1960s, with the introduction of 3D movies and holography, respectively, the interest has been rather low. Nevertheless, after few decades of relative obscurity, three dimensional displays have lately become both more and more popular and practical in numerous areas. Although the perfect setup would be to present the viewer with a depth of dimension without the use of special equipment, e. g. autostereoscopic displays, stereoscopic display, requiring headgear, meets the requirements of today. Taken into consideration the benefits introduced by the perception of depth in areas such as education, therapy and entertainment, the inconvenience of wearing glasses are mainly overshadowed by the advantages [4]. Having been introduced some 150 year ago, the 3D technology is not just an invention of the digital revolution. However, the latest approval of and adoption to the technology is. Due to the inexpensiveness and easier availability of digital equipment, improvements in network bandwidth and compression standards, and improvements in the view separation techniques, 3D is evolving as a viable technology. The addition of a third dimension expected for future visual communications will undoubtedly result in an increase of transmission requirements. Although it can be debated that video compression will become outdated since adequate bandwidth will become available to send uncompressed video, it should be pointed out that as the viewer’s need for ultimate quality images and high resolution stereoscopic video increases, bit rate efficiency will undoubtedly continue to be a major challenge in the future. Together with this desire for more realistic visual communication comes the need for interactivity already accessible at the Internet. Passive watching of a TV program, presentations or video pay-outs are becoming insufficient for many viewers and ways to introduce flexibility and interactivity into such areas are becoming vital.

Objectives

This master thesis has several different areas of importance. The main focus of this thesis is to evaluate the different techniques available in 3D video production. In this context it is necessity to determine the best method to obtain a realistic stereoscopic video. A great amount of research and publications has been done in the area of 3D video production, and some of these have been studied in order to decide upon one particular method for the capture and display process. The publications studied mostly refer to ideal conditions where the cameras are adjustable; hence due to the given equipment for this project several approximations will be made in the production procedure in order to reach acceptable results. Therefore, one area of emphasis of this project will be to describe the possible techniques available in order to produce and project good quality stereoscopic video, both in the ideal situation and in a situation with limited means. Finally a case study, such as e-learning, will be selected in order to demonstrate potential developments using the 3D technology.

3D Video Production

Several current and emerging applications in the areas of entertainment, remote operations, manufacturing industry and medicine can benefit from the depth perception offered by stereoscopic video systems. A stereoscopic video system uses two views of a scene imaged under the limitations given by the human visual system. However, there are many challenges to be overcome for practical understanding and extensive use of 3D or stereoscopic systems. The challenges are decent systems for3D video capture, display and efficient techniques for digital compression of the massive amounts of data, while keeping compatibility with normal video decoding and display systems. This chapter will cover the basic techniques used in production of stereoscopic video as well as techniques used in the display procedure. In order to understand the principles in 3D video production, an overview of the human visual system is included. Followed by some essential terminology which is presented before the different stereoscopic techniques for 3D viewing. An ideal stereoscopic video system configuration is defined and so are some of the potential sources of visual distortions.. To round off this chapter a short introduction to the general principles in compression of stereoscopic video is given.

Introduction to 3D

The history of stereoscopy can be traced back to 1833 when Sir Charles Wheatstone designed a mirror device that enabled the spectator to fuse two slightly different views of the same drawing into one stereoscopic image, resulting in a three-dimensional perception of the original picture. Regardless of the compelling side of a third dimension in pictures and drawings and the several new techniques developed during the following years, it was not until the 1950s when the first 3D motion picture was presented, that the 3D technology had a new revolution. During the years from 1952 to 1954 Hollywood produced more than 65 stereoscopic films, some of them making huge incomes. Still the big break-through was yet to come, not even the astonishment of holography presented in the 1960 did it [4]. However, since most of us observe a depth dimension in our everyday life, the area has always been tempting, and thus also under active research. Due to evolution in digital technology and the potentials that follows with, three-dimensional displays is again becoming gradually popular and for the first time also practical. Stereoscopic displays are being utilized in many application areas such as simulation systems, medical systems, telerobotics, computer-aided design, telecommunications and entertainment. Stereoscopic systems strengthen the perception of physical space in both camera recorded and computer-generated environments.

The Human Visual System

A significant aspect in understanding the 3D imaging mechanics is to first understand how humans see things, or in other words, how the human visual system works. Far from being a complete report of the human visual system, this section aims to help the reader understand how we identify the objects that surrounds us. The human visual system consists of a pair of eyes, where each eye has a retina. In short, the retina collects information and by optic nerves transmits this to a region called the lateral geniculate body, and further via pathways to the visual cortex, sees figure 2. 1. In order to go more thoroughly over the process of seeing the area the light reflecting off external objects must be explained. The light enters the eye through the cornea and the lens. Eventually it will reach the retina, which is the light sensitive part of the eye. The retina consists of photoreceptors called rods and cones. The rods lever vision of low light and the cones lever colour vision and detail. Light is absorbed by these receptors and electrochemical reactions are initiated leading to a generation of electrical pulses that are passed through to the optic nerve. The pictures shaped at each of the retinas are upside down, and as it is shown in figure 2. 1, the fibres of each optic nerve are split at the optic chiasm and cross over to the opposite sides. In other words, the fibres from the right eye go over to the left lateral geniculate body and fibres from the left eye go over to the right lateral geniculate body. As the visual information are passed on to the visual cortex and processed, we'll end up with one single upright image. As the human eyes are separated with about 6-8 cm, each eye will perceive a slightly different image [19]. It is the mixture of these two images into one final picture that gives us the capability to perceive depth. This ability to perceive three-dimensional depth due to the distance among a person's two eyes is called stereopsis. Stereopsis can be defined as the perception of depth produced by binocular retinal disparity; hence, it is simply a member of a set of properties frequently mentioned as visual-depth cues. These properties are often classified as being psychological or physiological or monocular or binocular. The latter classification scheme will be used in this thesis [19]. C: UsersalexDesktopfigure 3. 1. jpg

Figure 2. 1: The human visual system

Complex, natural scenes cover an extensive variety of visual cues to determine depth. Although stereopsis cannot occur monocularly, our visual system still uses several monocular, one eye, cues to appreciate relative location of objects in a scene. Such cues are [4]:• Motion Parallax: a depth cue that result from our motion. As we move, objects that are closer to us move further across our field of view than the objects that are in the distance.• Relative Size: Retinal image size allows us to judge distance based on our past and present experience and familiarity with similar objects.• Linear Perspective: When objects of known distance subtend a smaller and smaller angle, it is interpreted as being further away. Parallel lines converge with increasing distance such as roads, railway lines, electric wires, etc.• Aerial Perspective: Relative colour of objects due to scattering and contrast gives us some clues to their distance. Also contrast of objects also provides clues to their distance. When the scattering of light blurs the outlines of objects, the object is perceived as distant.• Colour and shading: a light object implies closer objects and a dark object implies distant objects.• Shadowing: shadows can provide information about an object's dimensions and depth.• Occlusion: occur when there is overlapping of objects. The overlapped object is considered further away.• Texture Gradient: the objects with the most details are considered closest. Most essential are still the binocular, or two-eye, depth cues. The cue binocular disparity, already stated, delivers alone enough evidence for the visual system to extract depth. The binocular disparity exists because the human eyes are horizontally separated which provides each eye with a single viewpoint unto the world. Two other important binocular cues are accommodation and convergence. Convergence involves the focusing on an object by rotation the eyeballs and adjusting the optical axis for each of the eyes so that these two axes can meet in the same point. Accommodation is the process of adjusting the focal length of the lens so that the eye adapts to bring near and distant objects into focus by use of the surrounding muscles, in other words, focuses at a particular depth. Experience is also an important factor in the discussion of stereopsis. Experience helps us calculate the relative size, form, distance and colour of an object. It is a built in characteristic in most humans that for example a bus is bigger than a bike, if they appear to be of the same size, the bus must be located at a greater distance. To complete our discussion of the human visual system, it must be stated that not all of us possess the ability to perceive stereo images. About 10% of the population suffers from a so-called stereo-blindness. This can be the result of colour blindness, night and snow blindness, lazy eye or astigmatism, a condition in the cornea. For more information about these factors, please refer to [19].

Terminology

Since not one but two capturing devices are needed in order to create the essential two images with a slightly different perspective, it has to be done in such a way that the human visual system can extract those images into one acceptable image. The two images will be joined into one in the visual cortex and depth information will be extracted as it is done in common viewing. If the pair of images is created with a conflict of depth cues, the final image will not be perceived correctly in the brain. This might result in an excess or reduce in depth perception leading to an uncomfortable viewing or without merging of the stereo pair at all. The actual camera and display geometry required to accomplish this will be explained later, since this section covers the human visual system and the vital terminology for the perception of a stereoscopic image. As outlined before both monocular and binocular signals are used to perceive depth. A stereoscopic image is simply a combination of two almost similar images where both binocular disparity and convergence signals will work properly. Accommodating those signals on the other hand will be inconsistent. This inconsistency appears because each eye focuses on a flat image, and not on a particular depth. This conflicting accommodation will be compensated for in the visual system, but only to a certain extent. The area of fusion is therefore is limited. Figure 2. 2 shows the area of exact fixation, termed as the horopter. In this area no disparity will occur, that is, the exact point in space where eyes meet where there will be no dissimilarity in the views observed by the left and right eye. The shaded area around the horopter shows the Panums Fusional Area, which expresses the limits of " close" and " far" objects that can be fused. The further away from the horopter the greater the angle of disparity gets, and when outside of the Panums area, diplopia (double images) will occur [21]. C: UsersalexDesktopfigure 3. 2. jpg

Figure 2. 2 The area of fixation, the horopter, and Panums are where fusion of two images will occur

C: UsersalexDesktopfigure 3. 3. jpg

Figure 2. 3 Shows the exact fixations point and examples of uncrossed and crossed disparity

As shown on figure 2. 3 two types of disparity can occur: crossed and uncrossed. Crossed disparity causes objects to seem closer to the spectator compared to the fixation point, while for uncrossed disparity the object seems to be further away than the fixation point. Uncrossed disparities cause positive parallaxes, while crossed disparities cause negative parallaxes, figure 2. 4, 2. 5 and 2. 6 shows examples of positive, negative and zero parallax. Figure 2. 4 demonstrates the situation where the object is behind the display screen, called positive parallax. In this case the projection of the left eye is on the left, and the projection of the right eye is on the right, and the distance between the projections is called horizontal parallax. The maximum positive parallax takes place when the object is at infinity at this point the horizontal parallax is equivalent to the distance between the eyes. The points where the eyes cross the screen are called the homologous points and the distance between these are as a maximum equal to the eye separation.

Figure 2. 4 Positive parallax [22]

By negative parallax it is meant that the point seems to be in front of the display screen, as shown figure 2. 5, the point is called the convergence point or object point. The distance between the left and right eye projections is called the negative horizontal parallax. The projection for the right eye is on the left and the projection for the left eye is on the right. The negative horizontal parallax is equivalent to the distance between the eyes when the object is half way between the screen and the centre of the eyes.

Figure 2. 5 Negative parallax [22]

The last case, figure 2. 6 illustrates the event of an object that lies exactly on the projection plane, which is called zero parallax.

Figure 2. 6 Zero parallax [22]

2. 1. 3. Techniques for 3D viewing

There are numerous different methods to categorize the various types of 3D displays. One way is to group the types based on the type of 3D images they produce. The types could be real 3D, true 3D or not-true 3D [23]. Real 3D produces actual three dimensional images. True 3D produces two sets of images and by different means presents these to the viewer in such a way that a three-dimensional effect will happen. While not-true 3D present the same image to both eyes, such as three-dimensional images in computer games. But there is also an additional and more common way of distinguishing between 3D display types. This arrangement separates between the ones where the viewer requires wearing special headgear, usually glasses, to perceive the third dimension, and the types of display where such viewing aids are not required. These 3D display types are generally mentioned to as being stereoscopic and autostereoscopic, respectively [20]. As this thesis focuses on stereoscopic techniques for production and viewing of 3D only this type of techniques will be defined in the rest of this chapter.

2. 2. Stereoscopic techniques for 3D viewing

All stereoscopic display techniques are based upon the same principle. The necessity is to project two images with a slightly different perspective into each eye of a viewer, with a small amount of visual distortion. This means that the left and the right eye will receive two images, join them together and then produce a 3D effect. This is accomplished by projecting two images onto a projection screen or displaying two images on a monitor while leading them to either eye by filters or optics. As stated in the previous section, stereoscopic display techniques produce images of true 3D and are not created in the actual three-dimensional space. This implies that there will not be a way of looking " around" images produced with a stereoscopic method. However, the images created will be of good quality, if necessary, real-time and some of the methods give rise to good representation of the original colours. Within the stereoscopic techniques for 3D viewing there are alternative ways used to drive the left and right images to the suitable eye. The following section is an introduction to the essential filters and optics available for separating two images and presenting them to the viewer. But in order to understand the terminology first the general display principles are analysed.

2. 2. 1. Display Techniques

To begin with, let's recall the definition of disparity and parallax described earlier. Zero disparity or parallax means that an object is lying in the area of fixation; this means that if the scene projected is in true scale representation, the maximum uncrossed disparity will be the distance between the observer's eyes. This can be explained by the fact that if an object moves further behind the scene, the projection lines from the object to the left and right eyes become more parallel. The result of this movement will be that with a limit at a distance of infinity the disparity on the display screen will be equal to the separation between the centres of two eyes. Given an viewer at some distance from the screen presented with a slightly different perspective on each eye, the perception of images in front of and behind the screen will be as shown in figure 2. 7. The viewer will focus on the screen while using monocular and binocular to converge its eyes on objects in front or behind the screen. This is thought to be the primary reason for the tiredness in stereo displays. As the disparity of an object on the screen increases so does the discomfort, hence, the only way to reduce the visual discomfort is to reduce the disparity on the screen. C: UsersalexDesktopfigure 3. 7. jpg

Figure 2. 7: The disparity and projection of a stereo image on a screen [24]

2. 2. 2. Colour Separation

The colour separation technique, also called anaglyph method, is possibly the easiest way for producing a 3D image. It produces the two sets of images necessary for the perception of 3D. One set is coloured red and superimposed on the other set, which is coloured green or occasionally blue. When viewed through glasses with these coloured filters, a 3D effect will be occurred, see figure 2. 8. The red and green filters mask of the opposite eyes colours. One problem with this method is that the real colours occasionally will be hard to reproduce, and the images may appear rather dark [20]. C: UsersalexDesktopfigure 3. 8. jpgFigure 2. 8: Separation of red and green coloured stereo images by the use of glasses with red and green glass [25]

2. 2. 3. Passive Separation

Passive separation, or also known as polarization separation, exploits the concept of polarized light to separate the left and right eye images. Polarized light is light with an electric field in only one direction, see figure 2. 9. This can be accomplished by filtering normal white light, where the light waves have electric fields in all directions. A polarizer is a material that lets only light with a specific angle of vibration to pass through. The filter used in such a setup is made up of very thin stripes. If two polarizers are set up in series so that their optical axes are parallel, light passes through both. Nevertheless, if the axes are set up 90 degrees apart, the polarized light from the first will blocked by the second, and the outcome will be of no light passing through. As the angle rotates from 0 to 90 degrees, the amount of light that is transmitted decreases. Using this fact in a 3D display it is required to have two output channels, one for each eye. In front of these orthogonally oriented filters can be places. The viewer then needs to use polarized glasses in order to separate the different views. When the right polarization directions are used corresponding to the direction on the glasses, the viewer will experience a decent quality stereoscopic image. C: UsersalexDesktopfigure 3. 9. jpg

Figure 2. 9: Polarization of natural light [26]

The polarization method defined so far is named linear polarization and is simply a special case of circular polarization technique, see figure 2. 10. When using circular polarization a smaller amount of light will disappear in the polarization procedure. Also the problem of misalignment when the viewer moves its head is addressed by this method. It is based on the use of two polarizations, one clockwise and one counter clockwise, produced by the use of right and left circular polarizing filters. The outcome is an electric field that rotates around the origin as the wave propagates [20]. C: UsersalexDesktopfigure 3. 10. jpg

Figure 2. 10: Circular Polarization [27]

2. 2. 4. Active separation

Active separation is a technique used in active glasses synchronized with a display such as a TV monitor or a computer monitor in order to produce a 3D effect. The left and right eye views are showed in rapid alternation and synchronized with a liquid crystal shuttering system which opens in turn for one eye while blocking the other. This form of system makes use of the fact that the human visual system is capable of integrating the components of a stereo pair with a delay of up to 50 ms [20]. If the shuttering speed is not fast enough, the resultant image might still be a stereoscopic image, but there will also be some annoying flickering. This technique is called active since the LCD-glasses contain some sort of electronics to lever the shuttering effect, e. g. alternately blocking the left and right eye. At any point in time only one eye must be open. Such glasses can either be wired to a monitor, or controlled by an infrared controller.

2. 2. 5. Active and Passive Separation

As the name implies this method, it uses both the active and passive separation techniques. This method is achieved by placing an alternating polarization plate in front of a monitor. Although this method is rarely used, but in order to describe all the different stereoscopic display techniques, a short overview will be given. The polarization plate used must use circular polarization, and so must the spectator. The alternating plate can then switch at the frame rate of the monitor among left and right - circular polarization. If the glasses worn by the spectator use passive circular polarization, an alternating stream of images for the left and right eye will ultimately reach the spectator [20].

2. 3. Stereoscopic Video System Configuration

Stereoscopic video can be created by using a stereoscopic camera consisting of two separate regular cameras, each equivalent to one eyes view. It should be clear that to create decent quality stereoscopic images, the two cameras need to be calibrated in terms of contrast, brightness, colour and sharpness of focus. It is also very vital that they are calibrated in terms of geometry, that is, without any vertical displacement. The two types of camera configurations called parallel and converging camera configurations will be covered in this section. The parallel camera setup uses two cameras with parallel optical axes, while the second setup, the converging camera configuration, uses cameras that rotate somewhat inwards [28]. The degree of stereo effect depends in both cases on the distance from the camera to the projection plane and the separation of the left and right camera. In addition, when using the converging camera configuration, the angle of convergence must be measured. Too much separation can be difficult to resolve for the visual system and is known as hyperstereo. To attain a good outcome the filming process needs to be wisely planned and executed. The stereoscopic parameters need to be calculated and adjusted in such a way that the disparities a spectator receives does not cause any visual discomfort and at the same time offers good stereoscopic images. It is also essential to certify the negative parallax does not exceed the eye separation. Depending on the stereoscopic encoding format chosen the left and right video signals can be kept separate through the entire image recording, encoding and storage and display. Or they can be stored on a single recording and played back from a single device.

2. 3. 1. Stereo-Camera Terminology

Before the actual geometry of the two cameras is discussed it is essential to present some of the terminology used in this aspect. There are six basic factors that uniquely characterise a stereoscopic camera and display system, their part will become obvious during the next two sections. A camera system configuration is determined by [29]:• The camera separation (t)• The convergence distance (C). E. g. the distances of the optical axes of the two cameras intersect.• The field of view of the cameras (α)The display system is determined by:• The viewing distance of the observer from the display (V)• The size of the display (Ws)• The distance between the viewers eyes (e)• The field of view of the viewer (θ)In order to develop a mathematical model from the geometry of stereoscopic camera configurations, together parallel and converging and some more variables need to be defined [20]:• CCD Sensor Width (Wc): the width of the CCD chip used as the imaging sensor for the imaging in the CCD camera• CCD Sensor Shift (h): the distance by which the CCD sensor is shifted laterally in order to achieve convergence in a parallel camera configuration• Focal Length (f): The distance from the optical centre of the lens to the image sensor when the lens is focused on infinity. The focal length is usually expressed in millimetres (mm) and defines the angle of view (how much of the scene can be fitted in the picture) and the size of objects in the image. The longer the focal length, the narrower the angle of view and the more that objects are magnified.• The Camera Convergence Angle (β): The angle where the optical axes in the convergent camera configuration intersects• Image Parallax (P): The horizontal distance between homologous points on the screen• Frame Magnification (M): The ration screen width to camera sensor width. The entire procedure of determining the image and display geometry is very complex. This has to do with the fact that one has to think through numerous coordinate transformations, that is, from capturing images in a 3D object space a transformation to 2D CCD coordinates is essential for left and right cameras, moreover, these coordinates need to be altered into 2D screen coordinate and lastly back to a 3D image space. In this respect it is worthwhile with some formulas that relate the different coordinates to each other. This can be visually described as [29]: Object space→2D CCD Coordinates→2D Screen Coordinates → Image Space(Xo, Yo, Xo) → (Xcl, Ycl), (Xcr, Ycr) → (Xsl, Ysl), (Xsr, Ysr) → (Xi, Yi, Zi)

2. 3. 2. The Ideal Camera Configuration

The ideal camera configuration has been described in several studies [20][24][29], and in this sector the main principles will be recited. If we consider a screen of width s at a distance Z from the spectator and two CCD cameras consisting of an imaging sensor, a CCD chip of width w and a lens of focal length f, the following ratios should be maintained [20]: Additionally, figure 2. 11 demonstrations the geometry of the camera configuration and display system reproduced in [24], from where more associations in an ideal configuration can be assumed. The conditions represented in the figure are [24]: Equation 2. 2 simply indicates that the separation among the spectator's eyes simply has to be equal to the separation of the cameras. While equation 2. 3 states that the field of view of the cameras has to be equal to the horizontal extent of the screen from the eye. The field of view of a camera can be adjusted using the zoom function. In this feature the significance lays in adjusting the focus in such a way the both the cameras have the exact same field of view. C: UsersalexDesktopfigure 3. 11. jpg

Figure 2. 11: (a) stereoscopic camera system and (b) stereoscopic display system [29]

The focal length must also match each other in a stereo camera setup. This is essential in order to guarantee that both cameras have the same object in or out of focus. There are three methods to accomplish this:• Set both cameras into auto-focus. The disadvantage with this method is that in auto-focus the system will provide greater weight to the centre of the image, which will be to some extent different for each camera in a stereo-camera setup.• Manually regulate the focal length. This technique would work if the iris could be fixed to a small slit while exposure is measured automatically• Use IEEE-1394 (FireWire) and AV/C commands to create a master-slave system, where one of the cameras uses autofocus and the camera follows the master's adjustments. This may result in blurring in the slave as the master camera has full control over the focal length. One more significant issue is the exposure level. Similarly in this situation a master-slave system could solve the problem, but the AV/C commands controlling the exposure manually are not so flexible. One could also alter the automatic gain control (AGC) but also this would be relative among the cameras. Thus the easiest method might be to place both cameras in auto-exposure and simply accept the small image differences. Perhaps the most significant feature in the generation of stereo video is synchronization. It is clear that images taken at different times will have a vertical or horizontal disparity when displayed. Synchronization can be achieved by setting the time-code on both cameras. Afterward, the two kinds of camera setups used in stereoscopic video production, i. e. converging and parallel configurations are discussed. Formulas that allow the calculation of the range of distances that can be stereoscopically fused based on the camera separation that will be taken. The discussion will start with the case of a converging configuration, as the parallel configuration is just an unusual case where the optical axes do not converge.

2. 3. 2. 1. Converging Camera Configuration

In a converging camera setup the optical axes of the cameras are rotated in the direction of each other at angle β so that their optical axes interconnect at a convergence point. In figure 2. 12 and 2. 13, we assume that two identical cameras are positioned along the x-axis and that their optical axes point in the direction of the z-axis. In the case of a converging camera setup the optical axes of the cameras will cross the z axis at some point, whereas in the parallel setup the optical axes will be parallel to the z-axis. The depth of the captured scene is in the range of the closest and farthest visible objects. The plane behind the lens of the camera, shown in the figure, is called the imaging plane. These planes are not to be shifted or rotated and hence their y-axis should be exactly the same. In the case of a converging camera setup the usual vectors of these planes are crossing with an angle, in the parallel setup they are parallel. C: UsersalexDesktopfigure 3. 12. jpg

Figure 3. 12: Converging camera configuration [29]

For a convergent camera system, disparity values can be either negative or positive dependent on whether the depth of the object is smaller or larger than the convergence distance. The convergence distance in the convergent setup is found by applying simple trigonometry and is given by: The alterations from object space to CCD and screen space and then to images space will be clarified afterward. First, the coordinate transform from a 3D object space to2D CCD coordinates is shown in equations 2. 5 to 2. 8.• First transformation: Object space → 2D CCD CoordinatesNote that by subtracting equation 2. 5 from 2. 6 we attain the disparity subsequent from the separation of the cameras. The maximum crossed and uncrossed disparities that can be stereoscopically are 4. 93º and 1. 57º, respectively [24]. When the stereoscopic images are displayed on a television or a computer monitor, the CCD coordinates must be altered into screen coordinates. This is accomplished by multiplying the CCD coordinates by the screen magnification factor M, which is the ratio of the horizontal screen width over the camera sensor width, which is perceived in equations 2. 9 to 2. 12.• Second transformation: 2D CCD Coordinates→2D Screen CoordinatesThe final transformation is from screen coordinates to image space and it is presented in the equations 2. 13 to 2. 15.• Third transformation: 2D Screen Coordinates → Image SpaceEquations 2. 13 and 2. 15 can be modified in a simplest form in order to get an expression for the screen parallax, P. As stated before screen parallax should not exceed eye separation and will occur when Zi is huge, related to the viewing distance, v. When the parallax is precisely equal to the eye separation the spectator is looking straight ahead. If the eyes need to diverge to fuse an image, this requires an unusual muscular effort which may cause discomfort [28]. So the parallax is given by:

2. 3. 2. 2. Parallel Camera Configuration

In a parallel camera setup the optical axes of both monoscopic cameras where are recording parallel, corresponding as observing a point at infinity. This point at infinity will have zero disparity since projections of this point will concurrent on the screen. The outcome will be an image perceived only at screen distance away from the spectator. This means that the rest of the point will have negative disparities and consequently fill the space between the spectator and the screen. A parallel camera setup is basically a special case of the convergent setup where the cameras are not rotated toward each other. The axes point in the positive z-direction, therefore β equals zero. Hence, the equations that were describing in the convergent camera setup also apply in the case of a parallel camera setup by simply removing the term β [20]. In figure 2. 13 the video captures variables that have an influence on the horizontal image disparity captured at the camera's imaging sensors included. C: UsersalexDesktopfigure 3. 13. jpg

Figure 2. 13: Parallel camera configuration [29]

2. 3. 3. Possible Sources of Visual Distortions

Presence the potential visual distortion that will arise due to the stereoscopic geometry that is going to be used is necessary in order to minimize their effect. Visual distortion in a way is the difference between the in a stereoscopic image of a scene with the actual viewing the scene. Distortions will constantly be a problem when dealing with stereoscopic imaging and display, but there are methods to decrease them significantly. Besides of being uncomfortable and irritating such distortions may also destroy the whole 3D effect. Consequently, this section will briefly describe the different types of distortion. For more information on distortions refer to [20] and [29].

2. 3. 3. 1. Depth Plane Curvature

Depth Plane Curvature means that there is a curvature of the depth plane, as shown on figure 2. 14. This form of distortion is a result of using a converging camera configuration. Figures a, b display the same convergence distance achieved by a convergent and parallel camera configuration, correspondingly. The curvature in the depth planes presented by the convergent configuration stands in contrast to the depth planes of the parallel setup, which are parallel to the surface of the screen. C: UsersalexDesktopfigure 3. 14. jpg

Figure 3. 14: The convergence distance achieved by a converging camera configuration (a) and a parallel configuration (b) [29]

Such a curvature of the depth plane can lead to incorrectly perceive of the relative depth, as objects in the corners of an image will appear to be further away than the objects in the centre of an image. Image motions subsequent from a panning of the camera will also be affected.

2. 3. 3. 2. Keystone Distortion

Keystone distortion is one more type of distortion that only happens in configurations with a convergent camera setup. This causes vertical parallax in a stereoscopic image because of the difference in the planes of the imaging sensors of the two cameras. The outcome of this distortion can be seen on the screen as the image from the right camera will be slightly larger on the right side that the left side, while the image from the left camera will be slightly larger on the left side than the right side. This results to a vertical parallax whose value is greater at the edges of the screen and increase when the camera distance increases, and also when convergence distance and focal length decease. Additionally to the vertical parallax, a minor horizontal parallax will also be presented at the edges.

2. 3. 3. 3. Shear Distortion

One recognized property of stereoscopic images is that they look like following the viewer while the viewer alters its position. A sideways movement of the spectator results in a sideways shear of the stereoscopic image on the surface of the screen. This may result in faultily observing the relative distance among objects and in some cases an incorrect perception of motion in a scene.

2. 3. 3. 4. Depth Nonlinearity

As shown in figure 2. 14, the depth is stretched among the spectator and the screen and it is compressed among the monitor and infinity. This non-linearity of depth on the display can lead to false observation of depth on the screen. It may also lead to wrong estimation of speed in cases where the camera is moving. It has been shown that a linear association amongst image depth and object depth can only be achieved by configuring the stereoscopic video system in a way that infinity of an object is presented at image infinity on the stereoscopic display.

2. 3. 3. 5. Lens Radial Distortion

Lens radial distortion is caused by using spherical lens. The use of a spherical lens means that the focal length will be changed at various radial distances from the centre of the lens, where it will be highest at the centre and lower at the edges. The outcome will be an image with curvature in the corner, producing vertical parallax. The vertical parallax will be subject upon the radius of the lens and increases with the reduction of the focal length. Aspherical lenses will decrease the extent of radial distortion.

2. 4. Stereoscopic Imaging and Display

The method of obtaining two 3D projections is known as stereoscopic imaging. It serves as the origin for capturing the relative depth information in a 3D scene as seen from a specific couple of viewpoints. As such, the stereoscopic imaging imitates the human optical system in the sense that two 2D projections are achieved by imaging the 3D scene onto two properly placed CCD sensors, through two imaging elements (lenses) that are separated by a distance called the camera baseline. The applicable camera baseline, when the stereoscopic image couples are intended for viewing, is the separation among the two human eyes

2. 4. 1. Understanding 3D projection systems

No 3D projection system is neutral as far as image quality is concerned. Image quality is the key factor in 3D perception, and the color artist is on the front lines in the battle for image quality. Understanding how projection systems split the original 2D resolution, lightness, and color density into two even shares is key to making it produce a good 3D image. The process of stereoscopic projection relies on an encoding and decoding process that assumes the function of exposing each eye exclusively to its intended image. This encoding, also called multiplexing, can be in time, light, or color.

Time multiplexing

Active stereoscopy separates eyes in time, with alternative projection of the left and right images on the screen. Viewers wear liquid crystal shutter (LCS) glasses that alternately blind each eye in synch with the projection. It is mostly deployed in Europe.

Light polarization

Polarized stereoscopy uses neutral gray filters that orient the light waveforms at the projector output. The screen should be silver-coated to preserve the light polarization. The viewers wear sunglass like plastic glasses that block the light that is not in the appropriate orientation. Linear polarization is more discriminating, but the viewer needs to keep the eyes level. Circular polarization systems are generally less discriminating and leaning your head will only slightly shift colors.

Wavelength multiplexing

In wavelength multiplexing, the projector’s light source is divided into six narrow bands, organized in two discrete sets of RGB lights. The left eye uses one RGB set, while the right eye uses the second one. No special screen is needed, and special glasses filter the needed wavelengths.

Dual projection

Nothing seems simpler than one projector per eye. That’s not that simple, for it requires perfect synchronization and geometric alignment of the projectors, just like pairs of cameras. It was the favorite projection system in the 1950s, with polarized filters and glasses. Most IMAX 3D movies use dual 70 mm film projector setups.

Beam splitters

The beam-splitter projector is the equivalent of a camera 3D lens. Images are stacked together on the imaging device, and the optical attachment un-squeezes and overlaps them on screen. This method was often used in the 3D systems of the 1970s. The beam splitter is making a comeback as a light efficiency optimizer for RealD and on Sony's 4 K projectors.

Mixes and matches

You can mix and match 3D projection technologies. Active stereoscopy is used with polarization or wavelength filters for single-projector 3D. There are more unexpected systems, like active stereoscopy using dual projectors with mechanical shutters in some large-format venues.

Color multiplexing

The most infamous color filtering is the anaglyph encoding discussed at length in this book. Despite all its faults, anaglyph encoding is still a distribution format, especially for the home market under its modern green-magenta and yellow-blue incarnations. It is not a projection system by itself, but we want to mention it. Despite being crude and ugly, it works all around the world, from the richest home cinema installations to the poorest venues with legacy 16-mm projectors.

2. 4. 2 Stereoscopy glasses

The vast majority of stereoscopic display methods (primarily 3D TV and cinema) require the viewer to wear some form of 3D glasses. As described above there are many technologies used to display a stereoscopic image, and each of them requires a certain type of glasses. Generally, these can be classified in to Active or Passive 3D glasses.

2. 4. 2. 1 Active 3D Glasses

As the name suggests, these types of 3D glasses have an active component to them and can either be Shutter Glasses or Display Glasses.

Shutter Glasses

These are the most common type of active glasses used to create stereo images, as they can be used with most types of existing display technologies. Shutter technology works by producing a pair of glasses with a pair of solid LCDs as the viewing lens. The LCD is not the usual kind you see in a TV, but appears as a sheet of glass that when activated will block light. Shutter glasses use this effect to alternate the image received at the left and right eyes. This means that the display can then show alternating angles of the stereoscopic image, so that each frame is seen at one eye, then the other. Using this technique will efficiently half the refresh rate of a display, as two frames are essential in order to build a stereo image. So, if an LCD TV operates at 100 hz, the stereoscopic image is displayed at 50hz. This procedure needs to be completed fast enough so that the brain is fooled into seeing the separate images as a unbroken series of motion, rather than a series of stills. Consequently, the higher the refresh rate, the more comfortable the effect will be. Shutter glasses are nowadays reasonably compact devices, often wireless with rechargeable batteries. However, the image viewing isn't as comfortable as passive glasses for many people. The advantages are that they can often work with existing hardware, reducing the total cost of the solution.

Display Glasses

Often used in high end applications, usually in the form of Head Mounted Displays (HMDs) in virtual reality environments or military hardware. They consist of two full LCD screens that can display separate images to each eye, often providing a high-resolution and quality image to each eye. This is a much more expensive option, though the outcome is often a more effective stereo image. A prime advantage is that the display device is built in to the glasses, so they are a much more compact display solution.

2. 4. 2. 2. Passive 3D Glasses

These type of 3D glasses are more commonly used on specialized stereoscopic hardware, such as high-end 3D monitors and 3D cinemas. There are no active components, just a method of filtering a single image in to two channels (providing a stereo image).

Polarizing Glasses

Most people who have seen a 3D film at the cinema will have used some type of polarizing glasses. IMAX and Real3D both use this technology, but implement it in a different way. IMAX uses glasses with linear polarization and Real3D use circular polarization. This means that the glasses can filter two dissimilar images from the screen by the direction that the light is travelling in, as the lens in front of each eye is polarized in different directions. Linear polarization filters light in to vertical and horizontal light wave components, which has the weakness of requiring the eyes to be almost level. Circular polarization can filter light waves based on the clockwise or counterclockwise direction, meaning that the stereo image is maintained even if the head is tilted. The drawback with these glasses is that the brightness of the source is reduced, as only half of the overall light from the screen reaches each eye. The benefits are that it is each and cheap to display the same stereo image to many users, as passive glasses are cheap to create. Each of the 3D display technologies uses a different method of presenting the image, which is beyond the scope of this article. Some use multiple/single projectors, LCD screens and many other display methods.

Anaglyph Glasses

These are the older color filter 3D glasses that were seen in children's books and early cinemas. Usually red and cyan filters are placed over each eye to allow two images to be formed from a single source. The stereo image doesn't retain the original colors and is generally considered an inferior method of stereo viewing.

2. 5. Compression of stereoscopic video

As a result of low bandwidth speeds and high bit rate of raw 3D video, efficient compression tools are essential in order to satisfy the average TV viewer, addressing the progressively more demanding requirements for interactivity and a sense of realism. In stereoscopic video the raw bandwidth of the video signal is twice that of standard video, since stereoscopic video has two channels, each carrying a different viewpoint. When compressing a stereoscopic video it is required that the large correlation among left and right views is exploited. This scenario is very similar to that of exploiting the similarities of temporally close frames in motion compensation. The dissimilarity among left and right can be considered as a type of motion, and thus motion estimated. For the specific case of stereoscopic compression, there are some vital notes to be made. First of all, if the two cameras used for capturing stereoscopic sequences are set up horizontally, searching for block matches need to be done horizontally only. The search region in the y-direction can be narrowed down to about zero dependent on how well the cameras have been aligned. Normally, large regions of the two formats will be similar. Therefore, larger block size could improve performance, because that would involve fewer motion vectors. Since two frames are used for the prediction when encoding a B-frame, the reference frames can be chosen so that one frame is taken from the opposite view, and one frame from the same view, and as a result both temporal and spatial redundancy can be exploited. There are three types of methods that can be used for coding of stereoscopic video [20]:• Simulcast Stereoscopic Coding: This methodology codes the left and right views individually, which means that any of the two views may be decoded and shown on a normal TV.• Compatible Stereoscopic Coding: This method codes one view, say the left view, individually and only the relative differences for the right view. This means that the view coded individually can be shown on normal TV sets.• Joint Stereoscopic Coding: This approach codes both views at the same time, which means that neither one of the views can be decoded by a normal TV set. The simulcast stereoscopic coding method is the least effective approach due to the fact that two independent views must be coded separately while compatible stereoscopic coding independently encodes only one view and exploits the correlations among the two views in order to code the other view dependently. The joint stereoscopic coding method is clearly the most effective approach, due to the fact that it does not implicate any independent coding at all, hence allowing even greater compression efficiency. This increased compression efficiency has the disadvantage of compatibility, hence the compatible stereoscopic coding seems to be the most practical method as it provides a reasonable compromise among the conflicting requirements of higher compression and compatibility with normal TV. There are numerous techniques for compatible stereoscopic video coding which all are based on present methods for normal video coding. A technique based on the disparity among the views are the most common, however several other methods are available and constantly under development.

3. Digital Video Production

Digital video production in simple words means, using cameras in order to record pictures and sound into digital form. A digital camera records about 50 pictures per second and combines every two consecutive pictures into one frame. Those pictures are named field, hence using the right terminology, two fields are mixed into one frame. This mixing is called interlacing and it is a unique feature of the digital camera as opposed to the traditional analogue camera. The digital storage, which makes the editing procedure easier, is another characteristic feature with the digital camcorder. The digital pictures and the digital sound can be transferred, as files, to a computer where the editing process takes place using programs, and then the video is ready for screening. The production process used in digital video production has evolved from the film and television industry. Many of the same terms are used, such as filming and shooting, which can reflect the evolution. The digital technology was first introduced in video and audio production to improve picture and sound quality. But it was also quickly adopted for editing because it offers a greater degree of flexibility and efficiency than traditional methods. The recordings can be done a lot cheaper since expensive film is no longer essential and the equipment is becoming low-cost and more available. Nevertheless, the digital video production process involves more or less the same basic steps as the traditional production.

3. 1. Digital Video Production Basics

No matter the type of the video or the equipment that is used, the same basic steps must be followed each time. The process of producing a digital video is divided into three stages (see figure 3. 1):• Pre-production: this phase includes everything that must be planned or done before filming and editing can take place. During pre-production the storyboard, script and shooting schedule are made. The team who will make the production is brought together, and all the resources needed are assembled.• Production: this is the actual recording, or " shooting", of the audio and video, which will then be edited together into the final product. This stage includes all location filming, indoor and outdoor, plus any extra " footage" (shots) which might be required for the titles or credit sequences. The storyboard, script and shooting schedule which have been produced in the pre-production stage, are the bible for this stage. The quality of the material will determine the quality of the finished video, as these are the raw materials or basic building blocks of the production. As there is not much to do in post-production to improve the quality of poor video or audio, it is vital to get it right at this stage.• Post-production: is the stage of transferring and editing the audio and video clips into a sequence. Titles captions and music are added and the finished production is output. This includes many different things, from logging the shots to edit, recording voice-overs, mixing the audio and music, to creating and distributing the final product. The post-production is dependent on the two previous stages, and a satisfactory product will be impossible to achieve if the pre-production and production stages are done poorly. C: UsersalexDesktopfigure 4. 1. jpg

Figure3. 1: Overview of the digital video production process [33]