International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-9 Issue-1, October, 2019

3D Stereoscopy Video Production: Z-Depth Extraction and Optimized Rendering using Foundry Nuke

Abhishek Kumar, Achintya Singhal, Jitendra Sheetlani

***Abstract*: This paper proposed an effective and efficient 3D stereo Video production methodology for Stereo conversion of any image or video using the 3D compositing tool Foundry Nuke. For efficient S3D conversion, there are several pipelines uses by video production engineer into the industry. There are several technical issues during the 3D stereo production like spilling, reflection, translucency, occlusion, flickering and noise problems as well due to imperfect calibration.**

This paper presents a theoretical explanation of the principles of stereo vision systems, followed by a quick review of the state of the art. The research paper concludes with validating the assumption of 3D stereoscopy video conversion with film case studies to execute high-end 3d stereo video production quality with roto-based depth map extraction with optimized render conversion methodology.

Keywords: Stereoscopy, Video Production, Rotoscopy, Depth map.

I. INTRODUCTION

The stereoscopic technique was made-up as early as 1838, there is some history behind its development [1]. Charles Wheatstone (English physicist) is the man that discovered the method, and at the same time, he wrote an article. It was published in Philosophical Transactions [2]. This article discusses the different discoveries made by him. The development progressed from his discoveries in several different directions and created many methods, and today we can experience stereoscopic material. In his article, Wheatstone has explained that when we look at objects in a far distance, it does not matter from which eye we look, because the appearance of the objects is exactly same to both eyes [2]. However, if we look at an object much nearer to our eyes, each eye sees the object with a slightly different angle. It is something that we can easily test. If a piece of paper is held so that one of its edges touch our different nose sides of the paper it would naturally be seen when keeping one of our eyes open at a time. Charles Wheatstone however, seemed surprised that no one had documented this phenomenon before him. However, the cause of this, as he has written, might be that any consideration of it was quickly abandoned because they believed that the images presented to each eye were always different, the difference would be too tiny to be taken into account. Although, Charles Wheatstone found an article among the works of Leonardo da Vinci where he observes that objects in a painting can never feel as real as natural objects. To test what da Vinci meant, hold up a finger in front of our eyes and focus on shape or object in the distance, and then move the finger right and left in front of our eyes, so it passes over the object. The finger will never cover any part of that object. This works only when both the eyes are kept open. In painting, everything behind an object is hidden. Charles Wheatstone had established that we see 3D objects because the eyes are presented with little different images. He wanted to know that what the result would be if identical images were presented to both eyes while they were projected on a flat surface. Charles Wheatstone writes after mentioning the complication the eyes can have when trying to comply while each of them is looking at the different images: "These inconveniences are removed by the instrument I am about to describe; two images are put in it at the true concourse of the optic axes, the focal adaptation of the eye preserves its usual adjustment, the appearance of lateral pictures is fully avoided, and a large field of view for both eyes are obtained. The frequent example I shall have an occasion to make to this instrument will render it convenient to provide it a specific name, I, therefore, propose that it be called a stereoscope, to indicate its property of representing solid figures." [2].

II. PROBLEM STATMENT

Several studies has been conducted to investigate the side effect of stereoscopic cinema in relation to eye is visual discomfort, motion sickness and headache. Several researchers has claimed that the visual discomfort caused by conflict between the convergence of eye movement and their functions. Visual discomfort measure via both objective and subjective measures. To receive the objective data so called fusion range measurement was utilized, this is when the maximum screen parallax that is perceived stereoscopically is measured. It is commonly accepted that visual fatigue is related to conflict between the accommodation eye function and the convergence functionality that may occur when stereoscopic content is viewed.

Revised Manuscript Received on October 30, 2019.
* Correspondence Author
Abhishek Kumar*, Assistant Professor, Department of Computer Science, Institute of Science, Banaras Hindu University, Uttar Pradesh, India Email: abhishek.maacindia@gmail.com
Achintya Singhal, Associate Professor, Department of Computer Science, Institute of Science, Banaras Hindu University, Uttar Pradesh, India Email: achintya.singhal@gmail.com
Jitendra Sheetlani, Dean, School of Computer Application, Sri Satya Sai University of Technology & Medical Sciences, Sehore, Madhya Pradesh, India, Email: dr.jsheetlani@gmail.com

III. THE HUMAN VISION – TERMINOLOGY

When viewing the world multiple cues that indicate depth, distance, and shape all are used. Most of those cues are monocular but knowing and understanding the terms is essential to create a convincing 3D scenario. In the following chapter, those terms are discussed briefly, along with a general description of the operation of binocular vision in the natural world on the one hand, and the 3D world digital films can create on the other.

The human eye is a very sophisticated organ. With the brain, it does an astonishing job that helps to determine distance, depth, and location. The vision of the eye is a combination of reflected light on the retina, and then it interpretes into recognizable images by the brain [3]. The area in the back of the brain called occipital lobe processes the information from the eyes. Although the eyes are two and have equally as many points of view that are approximately 6.35 centimeters apart, they are explained as only one image [4]. It is due to particular fusion that happens to the brain where the optic tracts give stimuli from each eye to both visual cortexes. This process is commonly known as binocular vision.

The brain uses several depth cues to localize objects in space. Although binocular depth cues are the most accurate, the brain also uses monocular components to determine depth and location. Monocular vision is provided through one eye only. Monocular depth cues make a traditional 2D motion picture look as it has some depth and not be flat as it necessarily. To easily understand how 2D depth cues calculate depth, there are several terms that should be coined by reviewing the elements, monocular depth cues are clarified. By adding some of those principles into a stereoscopic scene that contains VFX, more depth can be created without necessarily having all elements in stereo.

The first term is occlusion, also known as interposition, it refers to when objects in the foreground overlap the objects in the background [6]. This gives an essential information when determining spatial relations between objects displayed in a 2D image.

Size cue can also give information about how far or close an object is in space [6]. The object size has to be set about another object of known size. For instance, when a man and a house seem to be of the same size in a photograph, the house might be tiny, the man may be a large or most likely, that the house is further far away from the point of view. [4].

The cue size can be aided with a texture gradient, but that pertains to when distance makes pattern and texture appear denser [6]. For instance, when one stands at a meadow filled with flowers the pattern of the flowers appears to be a solid color in the distance. Atmospheric perspective obeys similar laws, meaning that remote objects become less distinguishable. More specifically atmospheric perspective refers to when distant objects appear mistier than objects that are closer [4], again this is illustrated in figure 1.

Both height cues and linear perspective, relate to the horizon. It refers to when two lines that are perpendicular to the horizon and lead to the same spot. For example, when the two edges of a road seem to be meeting at the horizon [6]. The latter pertains to when the distance is determined by the height of objects in the horizon, substance close to the horizon are perceived to be further away than objects that are lower [6] [4].

Depth, depth perception, depth of field and perspective are words we use to describe a construct vital to survival but difficult to describe.

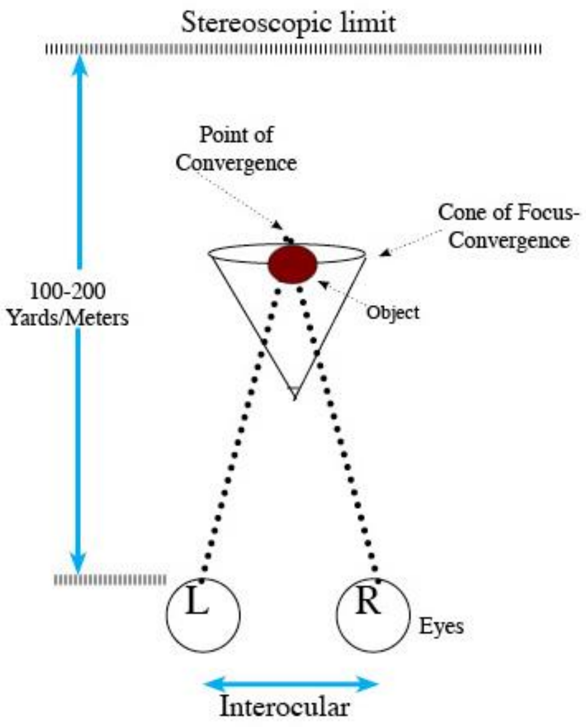

Figure 1: Combination of effects '3-D.

The figure 1 has only two dimensions ' height and width. Reality has a third dimension ' depth. We call the combination of effects '3-D.'

The perception of depth is how our brain tells us where we are in relation to the rest of the world. Just as it allowed brain to know how far the object was from eyes, it lets the modern shopper reach for a can in the grocery without knocking over the whole stack. Science calls this ability to appreciate depth 'stereopsis'. (Kalloniatis & Luu, 2011).

Although it commonly attribute depth perception to the binocular quality of having two eyes, most theorists say it is the result of a combination of factors called 'cues.' As Majumda (2011) explains.

Cues can be grouped into three categories:

- ***Oculomotor:*** This is the cue our eye gives our brain by sensing the position of the eye and the tension of the eye muscles.
- ***Monocular***: These are cues that work with one eye.
- ***Binocular:*** These are the cues that truly depend on have two eyes.

Stereoscopic cinema is a recently resurrected trend in the film industry. 3D stereo cinema is currently a very emerging trend. This thesis details for producing stereoscopic 3D video using different stereoscopic tools provided by The Foundry Nuke.

At first view it seems sublime to add the 3D but as the technique is becoming more widespread. Stereoscopy 3D cinema has been known for centuries,

Retrieval Number: A1926109119/2019©BEIESP
DOI: 10.35940/ijeat.A1926.109119
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

In 1833 The English scientist Sir Charles Wheatstone invented the way to project two images to create a stereoscopic experience. A patent for a 3D cinema was filed at the end of 19th century, but it was not until the 1st quarter of the 20th century that 3D films or movies were publicly available [5]. The technique continued to progress over the years and around the American Civil War stereoscopic viewing device could be found in many homes. In 1862 inexpensive dual lens cameras came to the market, which enabled even easier access to create a stereoscopic video [5]. The stereoscopic effect was solely used for still images at the time, but the technique was later adapted to create 3D for motion pictures.

The first device to view 3D films was so called peep-show where only one person could watch at a time. In 1915 a more theatre friendly version of the 3D projection was invented, namely the red and green anaglyph technique. This technique consists of two pairs of footage that are printed over each other in two different colors and viewed with glasses in corresponding colors [5]. Although the Stereoscopic method has been around for ages, it took a long time to become a real hype within the cinema industry. It was not until 2007 that several magazines covered speculations about 3D cinema in Hollywood industry that it took off and producers started to invest in full-featured stereoscopic films [8]. The current stage of the 3D cinema and the hype of it is currently creating has some parallels to when the cinema introduced both sound and color which are currently considered as a prominent part of the visual experience. The addition of audio in to the cinema required complete synchronization between two machines, similarly to the stereoscopic projectors used in 3D films. The technology was expensive and resulted in higher ticket prices for cinema goers [8]. Thus it is a viable consideration whether the 3D cinema is not only going to be a hype but become a part of the traditional cinema experience.

IV. DIFFERENT METHODOLOGY

To view images or animation in 3D stereoscopy, cinema, the eyes need to be presented with two s different images one for each eye. There exist several different methods to get this. Through that experience, the stereoscopic effect can be achieved by using our eyes or one of several aids, such as a pair of 3D glasses. For each viewing technique, the substantial stereoscopic needs to be made explicitly for it. There will be a brief detail of several conventional and favorite techniques and a more detailed explanation of anaglyphic stereoscopy. A significant contrast between the methods is the price, which is very important between the methods that use some kind of 3D glasses and are suited for a bigger audience.

V. COMMON APPROACH

There are several methods for show the 3D stereoscopic effect in addition to the anaglyphic. Although this thesis is based on the anaglyphic method, there is a short explanation of a few other common and popular methods, SBS and Autostereogram below to provide some insight to different methods of experiencing stereoscopic material. Maximum of these are suited for only one viewer at a time, while there are different types of 3D-glasses that work for a more extensive audience.

A. *Side by side* – A very simple method to experience stereoscopic 3D cinema that has two separate images placed side by side (SBS), where the perspective is set for one of the images represent the right eye and the perspective for the other image represent the left eye. These pictures can usually be viewed as they are, without the use of any additional tools. The two pictures can be placed on either side of each other. There is a little difference in the way of viewing them, if the image that is made for the right eye is placed on the right side of the two, they are meant to look at by keeping the eyes in parallel. If the right eye image is instead placed on the left side, the two images are meant to view by crossing the eyes. Using the wrong viewing method (for example the parallel method when the image is made for cross-eyed viewing) will reverse the stereoscopic effect. Things that are for example to look like they are nearer to the viewer will instead look further far away. These pictures are also for use with a stereoscope. If we are looking through a proper stereoscope it creates very less strain for the eyes, and there will be less distraction as well for the viewer.

B. Autostereogram – The autostereogram images show a 3D depth when the right viewing technique is used to look at them. They can sometimes be hidden behind what seems to be an entirely random mess of colors. This “mess” is actually not so random for it is a repeated pattern that conceals the three-dimensional objects [Auto stereogram] by using the parallel viewing technique, it will create a stereoscopic effect as the adjacent pieces merge together. By using the z-depth or camera depth image (where objects closer to the camera is lighter and objects further away darker) of a 3D object or scene, the final image is generated and hidden behind either random dots or a pattern. [Shannon’s Autostereograms Page., 2009].

C. Polarized 3D-glasses – Polarized stereoscopy we use when two projectors project images onto a aluminum or silver covered screen, where the orientation of the electric field of the light from the projectors has been specified [Polarization. http://en.wikipedia.org/wiki/Polarization (2009)]. When we use this method to show stereoscopic material, there are two methods how the light is polarized: circularly and linearly. The electric field of the linear polarized light is directed vertically and horizontally from the projectors, and the viewer wears 3D-glasses which passes through light from only one direction for each eye. If the glasses are tilted their polarization, it will no longer match the polarization of the light and a bleeding effect between the two images will appear. This is avoided when using circularly polarized light and glasses. Then the light is directed in a circular motion, either in rightwards or in leftwards, which remains if the glasses are tilted to another side.

D. Liquid crystal shutter glasses – By alternately Liquid shutter glasses shutting the view at a high rate for the right and left eye, the two eyes can be given images with different perspectives and through that create the stereoscopic effect. These glasses contain liquid crystal and a polarizing filter which blocks the view by turning black when we apply a voltage or power to them [LCD shutter glasses, 2009].

The glasses start to synchronize the update frequency of the viewing source for continuously provide the correct perspective for our eyes. If the update frequency is not available as required, a certain amount of flickering or jerk might be noticed when one of the eyes sees a bit of the image meant for the other eye.

E. Infitec glasses – "INFITEC is a new technique to display stereoscopic images where the image information is transmitted in different wavelength triplets of the visible spectrum of light" [Jorke, Helmut & Fritz, Markus, 2003]. The technique divides the visible color spectrum into six parts, two for each of the three primary colors; red, green and blue [Stereoscopy. http://en.wikipedia.org/wiki/Stereoscopy (2009)]. One part from each of the colors is meant for the right eye while the other part is in the left eye. It will provide us slightly different colors and visual experience to each eye. Where the primary colors have a very narrow bandwidth, but the eyes are not sensitive enough to notice the tiny eerie differences this display method is needed to be used and it can show full-color images.

VI. ELEMENT BASED CONVERSION

Element based compositing work with multiple elements to

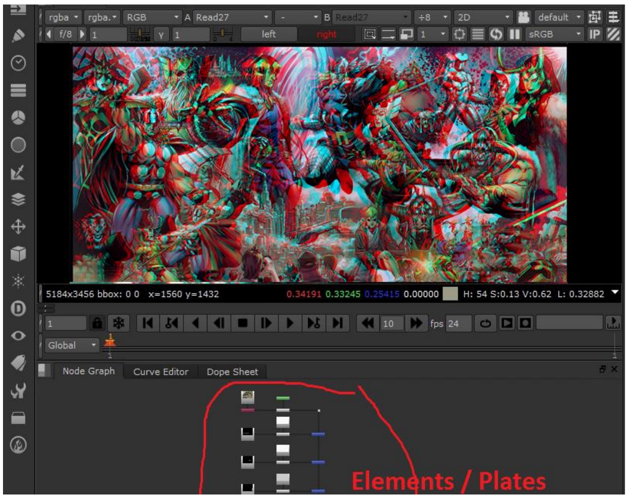

Figure 2: Nuke interface for stereo conversion

Combine together to create a final output of shot. In this proposed framework need to extract all the images one on one to getting the elements of the shot, and add them in sequential order and reassemble them into a stereo depth or stereo image.

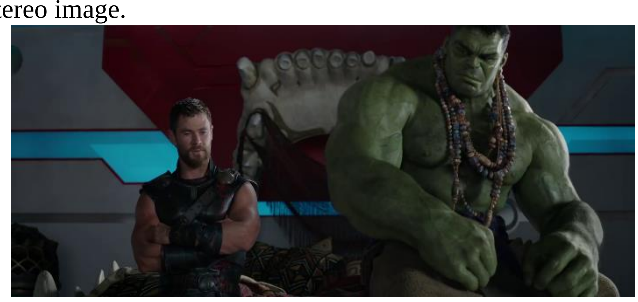

Figure 3: Test Shot for Stereo conversion

Figure *3*, having two characters Thor and Hulk talking with each other and the camera angle is low. So here there is need to split the shot in atleast 4 element to better Depth vison in 3D Stereo shot. So, first extract the Thor character as a separate element and Hulk as well also separate the hulk hand as a separate elements so hand look like more closer to camera. Then extract the midground of tub and at last extract the background wall in a separate channel using foundy nuke nodes.

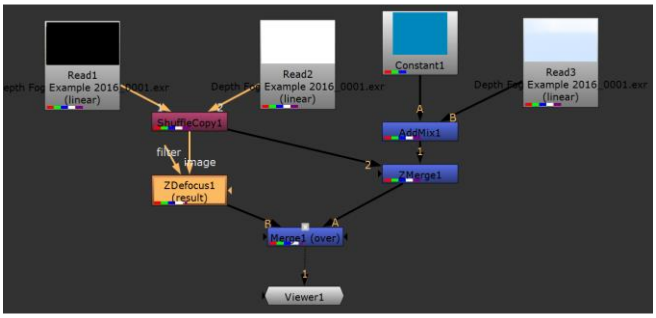

Figure 4: Framework of 3D stereo conversion in Nuke

After extraction of the elements in minimum 6 parts and Z-depth preparation. Now it's time to get connect the element, figure 4. First load both color and depth image using the read node, Read 1 and read 2 then use the merge node with shuffle copy node to blend the Z-depth with color pass then adjust the attribute and finally if require add the noise to realistic and natural view.

Figure 5: 3D Stereo image

This is the output of anaglyph 3D stereo image (For better vision of the depth use the anaglyph glasses).

VII. STEREO VISION CORRESPONDENCE ISSUE & ANALYSIS

The camera separation becomes significant difficulties arise in corresponding the two camera photographs and measure the depth of a point it must be visible to each camera, and we must be able to identify the point in each image. As camera distance increases, so it makes the differences in the scene as recorded by each camera. Thus it becomes difficult to match corresponding pixels in the images to produce the accurate result. Traditional z-based compositing operations are fairly limited in their use. This has one simple reason. They don't account for transparency. In a Z-depth pass resulting from a render engine's depth buffer, the distance from the camera is represented as a greyscale image. Thus, floating point or integer values, that are often scaled to accommodate scene scale and their limits.

Retrieval Number: A1926109119/2019©BEIESP
DOI: 10.35940/ijeat.A1926.109119
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Since integer values heavily delimit the precision of the stored information, they are outdated and hopefully not used by others for storage of depth values.

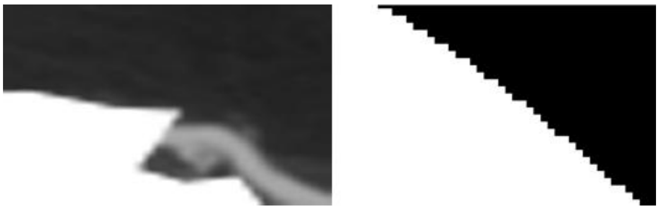

Figure 6: Antialiased (left) and aliased (right) between object A and B in-depth, where left output is right & right is wrong.

In stereo conversion deep images save multiple samples per pixel, one performance issue is clearly deductible. Saving much information is directly connected to bigger file sizes. Without compression, the relation between sample count and file size is proportional.

The Figure 11,12,13,14 showing the trend analysis of sample sizes and render time, CPU usage, and threads and ram consumption of the computer in comparison to traditional research methodology which is 30% more optimized than the traditional methodology. The data was taken from a 3D conversion rendering sampled with Thor Ragnarock figure 7 and figure 9 different step sizes and therefore the content is approximately the same.

Figure 7: 2D Video screenshot of Thor (Valkarie)

Figure 8: 3D stereoscopic screenshot of Thor (Valkarie)

Figure 9: 2D Video screenshot of Thor (Hulk Shot)

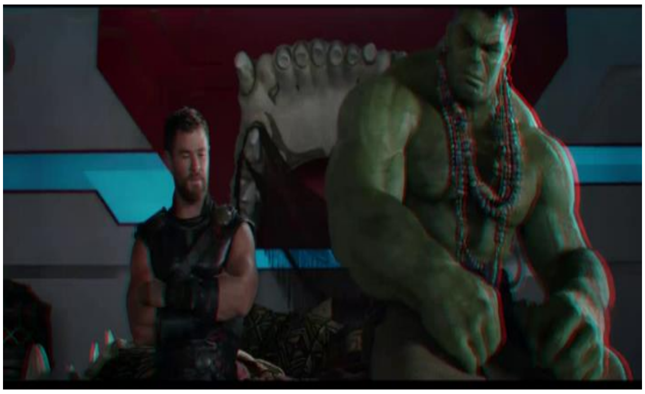

Figure 10: 2D Video screenshot of Thor (Hulk Shot)

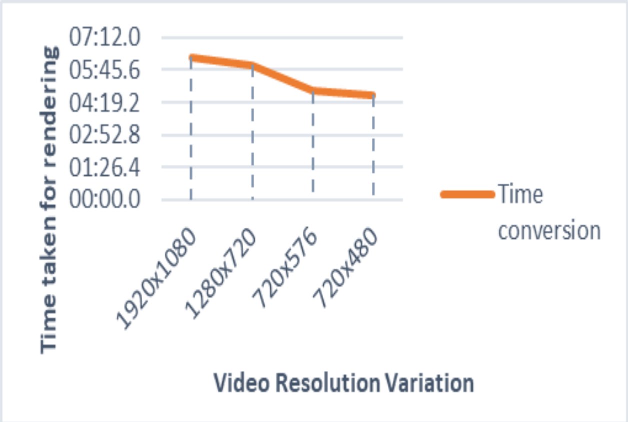

Figure 11: Relation of video resolution and render time

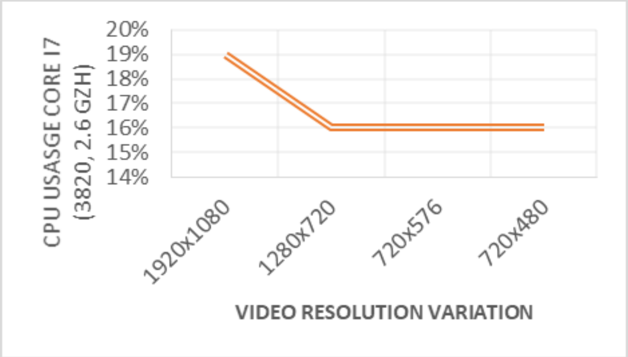

Figure 12: Relation of video resolution and CPU usage

Retrieval Number: A1926109119/2019©BEIESP
DOI: 10.35940/ijeat.A1926.109119
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

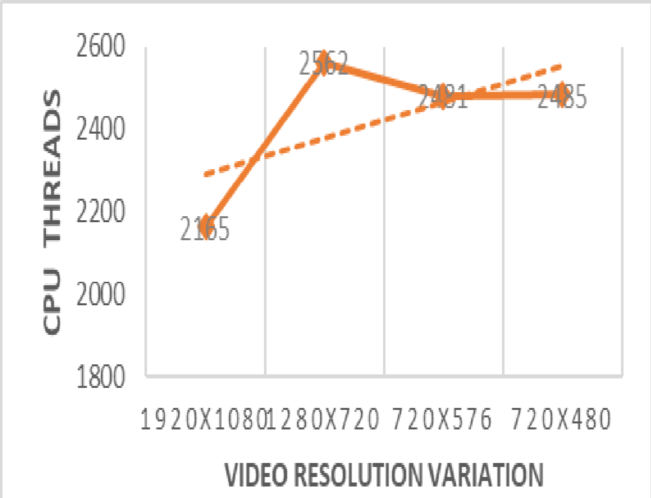

Figure 13: Relation of video resolution and CPU threads

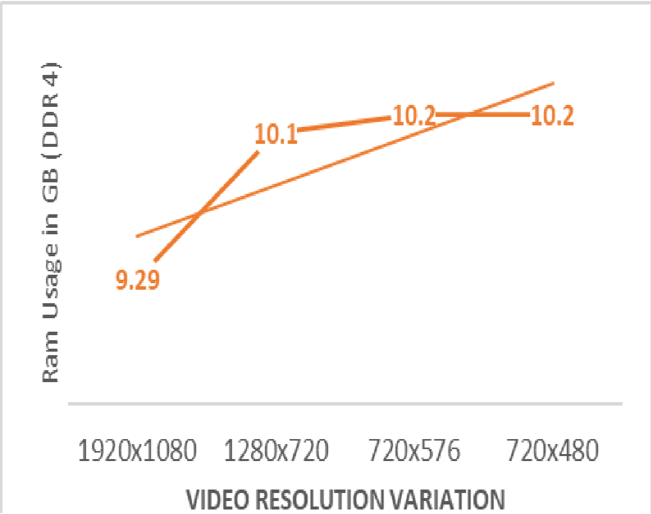

Figure 14: Relation of video resolution and RAM usage

Figure 11, 12, 13 and 14 is showing how the processing time, CPU threads and ram usages depends upon the screen resolution. The performance result analysis to convert the 2:26 min of Thor Ragnarock video Sequences. The following performance tests were done with a batch script launching the Nuke for composite rendering process of Thor Ragnarock Sequence.

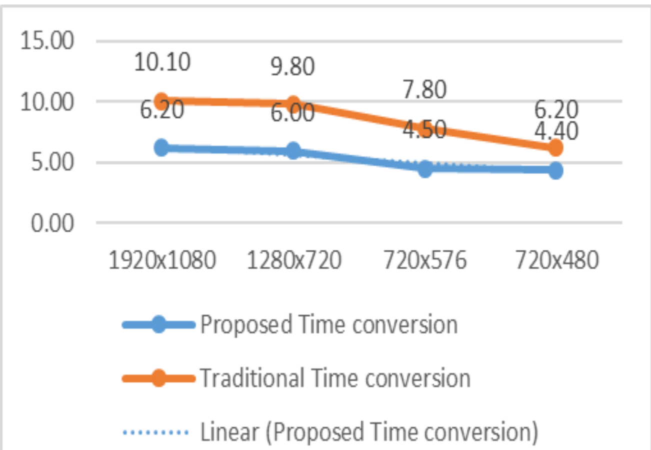

Figure 11: Proposed methodology time taken for rendering vs traditional rendering methodology

VIII. CONCLUSION

For effective S3D conversion there are several pipeline use by video production engineer into the industry. In this paper explain to develop the robust technique to extract the photorealistic stereo view of scenes. There are several issues we face during the stereo production like spilling, reflection, translucency and occlusion. Sometime we face the flickering and noise issues as well due to imperfect calibration. Stereo 3D have representation, objective function and optimization technique algorithm.

In conclusion, the 3D stereoscopy cinema is audience demand and a vital part of the entertainment. But current stage further development required in this technology for future success.

ACKNOWLEDGMENT

The image sequence, Video clip is taken for research study purpose from the movie Thor: Ragnarok and it was Produced by Marvel Studio and Distributed by Walt Disney Studios Motion Pictures.

REFERENCES

1. https://en.wikipedia.org/wiki/David_Brewster
2. Wheatstone, Charles. "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. Vol. 128 (1838) pp. 371-394. http://www.stereoscopy.com/library/wheatstone-paper1838.html (2009-05-27).
3. Seuntiens, P. J. H., Meesters, L. M. J., & IJsselsteijn, W. A. (2006). Perceived quality of compressed stereoscopic images : effects of symmetric and asymmetric JPEG coding and camera separation. ACM Transactions on Applied Perception, 3(2), 95-109. DOI: 10.1145/1141897.1141899
4. 3D Movie Making Stereoscopic Digital Cinema from Script to Screen By Bernard Mendiburu, https://doi.org/10.4324/9780080877969
5. Hayes, R. M. (1989). 3-D movies: A history and filmography of stereoscopic cinema. Londres: McFarland.
6. Foley, M.A., Foley, H.J., Durso, F.T. et al. Mem Cogn (1997) 25: 140. https://doi.org/10.3758/BF03201108
7. Kim, W.J., Leigh, S., Hur, N., & Choi, J. (2012). Feature-Based Detection of Inverted-Stereo for Stereoscopic 3D Viewing Comfort. IEEE Transactions on Broadcasting, 58, 296-300.
8. The global burden of headache: a documentation of headache prevalence and disability worldwide, Stovner Lj, Hagen K, Jensen R, Katsarava Z, Lipton R, Scher A, Steiner T, Zwart JA, https://doi.org/10.1111/j.1468-2982.2007.01288.x
9. Crary, Jonathan."Techniques of the Observer."The MIT Press45 (1988): 3–35.JSTOR. Web. doi:10.2307/779041. 30 Aug. 2018.
10. Zone, Ray. 2007. Stereoscopic Cinema and the Origins of 3-D Film, 1838–1952. Kentucky: The University Press of Kentucky.
11. https://en.wikipedia.org/wiki/Thor:_Ragnarok

AUTHORS PROFILE

Abhishek Kumar is an Assistant Professor at Department of Computer Science, Banaras Hindu University, Varanasi. He is Apple Certified Associate, Autodesk Certified. His research interests is Stereoscopy, 3D Animation, Image Processing, VR & AR, Multimedia, Game Technology, MOOCs, Computer Graphics & Visual Effects. He has a decade of experience in Motion Picture Science industry.

Dr. Achintya Singhal is an Associate Professor of Department of Computer Science, Banaras Hindu University, Varanasi. He has 42+ numerous research publications in national and international journal and conference proceeding. He has also visited Sri Lanka and Thailand in 2006. Research interest includes Digital Watermarking, Image Processing, Database; Programming, etc. He has over 11 years of teaching experience and good programming skill.

Dr. Jitendra Sheetlani having 14 years of teaching experience in Computer application & science. He is working as a dean school of computer application at Sri Satya Sai University of Technology and medical sciences. His research area is computer graphics, image processing, Cloud computing and Data mining.