International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:7, No:1, 2013

A Case of Study for 3D Stereoscopic Conversion in Visual Effects Industry

Jin Zhi

Abstract**—**This paper covered a series of key points in terms of 2D to 3D stereoscopic conversion. A successfully applied stereoscopic conversion approach in current visual effects industry was presented. The purpose of this paper is to cover a detailed workflow and concept, which has been successfully used in 3D stereoscopic conversion for feature films in visual effects industry, and therefore to clarify the process in stereoscopic conversion production and provide a clear idea for those entry-level artists to improve an overall understanding of 3D stereoscopic in digital compositing field as well as to the higher education factor of visual effects and hopefully inspire further collaboration and participants particularly between academia and industry.

Keywords**—**Clean plates, Mattes, Stereoscopic conversion, 3D projection, Z-depth.

I. Introduction

STEREOSCOPIC is any technique that creates the illusion of depth through the capture of two slightly different views that replicate human binocular vision [1]. These two offset images are then combined in the brain to give the perception of 3D depth.

In order to achieve a 3D illusion from non-stereo traditional 2D images, the process of making stereo images is called stereo conversion. Shooting and working directly in stereo footage is considered to be the best way for producing a result of 3D stereoscopic vision. However, for different purposes even in one shot of stereo film, there may also be a requirement for some footage to be converted.

II. Concepts and Terminologies

A. Principles of 2d to 3d Conversion

The principle of 2D to 3D conversion is based on the principle of binocular disparity, which refers to the difference in image of an object seen by the left and right eyes. Human eyes are physically separated in a horizontal position and the distance between two eyes is about 50 to 75 mm. Binocular disparity causes our brain to extract depth information from the two-dimensional retinal images in stereopsis [2]. Visual binocular disparity is defined as the difference between the point of projection in the two eyes known as visual angle [3]. In computer vision binocular disparity is referenced as coordinate differences of the point between the right and left images, it usually measured in pixels and can be used for the distance calculation. Disparity and distance from the cameras are negatively correlated. As the distance from the cameras increases, the disparity decreases.

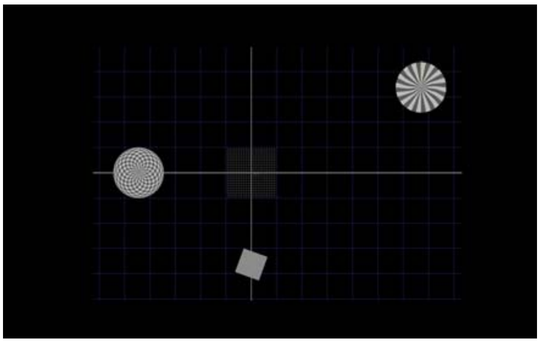

Fig. 1 A spatial relationship of objects in the real world

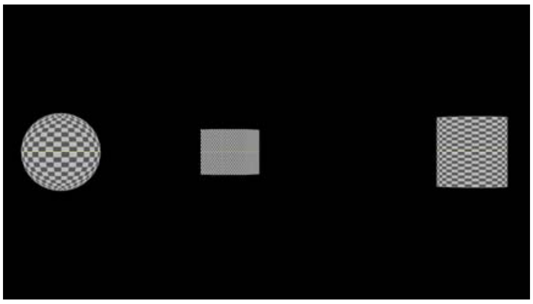

Fig. 2 A result of these objects looking in 2D view

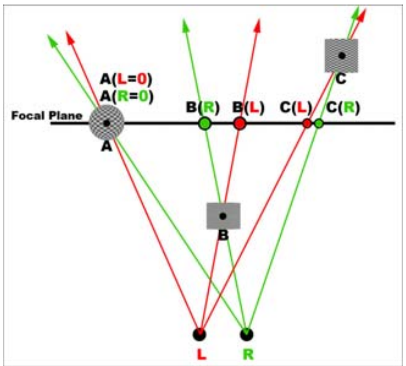

Fig. 3 Illustrates position shifts according to objects layered to focal plane and cameras (L,R at the bottom of this image)

Fig. 3 illustrates a relationship between focal plane and distances of objects layered to cameras (L,R at bottom). In addition, position shifts thereupon can be calculated from it. Focal plane [4] is a critical component in 3D stereoscopic conversion. It presents the area in each scene where key area is focused. In most cases, focal plane is set at the most

Jin Zhi is with the Duncan of Jordanstone College of Art & Design, University of Dundee, DD1 4HN, United Kingdom (phone: +44(0)1382 385316; e-mail: j.zhi@dundee.ac.uk).

attractive part of a scene for stereo conversion. In image Fig. 3, the focal plane has been set at the position of object A (middle ground). As we can see that the object A has no position shift for both cameras (L=0,R=0) on focal plane. However, there is an obvious position shift of object B in both cameras on focal plane because the distance of object B is closer than the distance of the object A from camera. Two shifted positions are marked as B (R) B (L). Object C has a small amount of position shift on focal plane shows as C (L) C (R) because it is the farthest object from camera positions. To summarise the comparison result from Fig. 3, on one hand, objects positioned closer to cameras (stereo cameras) will have larger position shifted. On the other hand, objects positioned behind focal plane have smaller position shifted on focal plane for both cameras. Objects on the same position of focal plane have zero position shifted.

B. Key Factors of 2D to 3D Conversion

1. Factor one

Most stereo films are shot with an understanding and consideration of the stereo nature of the experience [5]. 2D to 3D conversion ideally requires compositing scripts and converted images to have clear layers, for instance, foreground, middle ground, and background. Many details may have been divided and composited in a complex way which may be extremely valid, however, from the stereoscopic conversion point of view, those details may considerably increase the level of difficulty for stereo conversion.

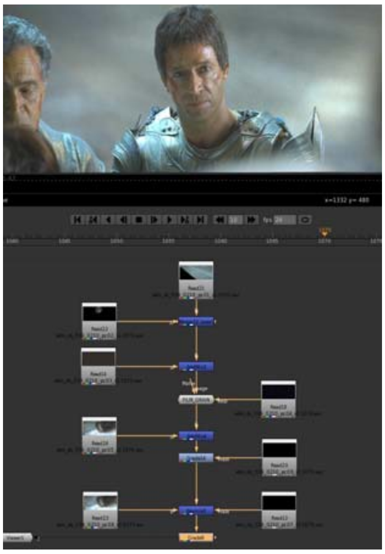

Fig. 4 A good example of where key elements have been divided from a finished compositing master script, and recombined into a reasonable size with a logical order in relation to original objects layered in space

Fig. 4 indicates a simplified version of a master script for one visual effects scene. In general, the simplified scripts are known as mini comps in visual effects field.

Another typical mini comp, which has many elements and layers included. These layers cannot be recombined as the original master script was composited and designed in a very complex and illogical way. The problem in this type of mini comp is that there are many hold-out mattes used in the original mono script. A hold-out matte is a loose-fitting shape which is designed to address specific problem areas, and be used in conjunction with a more exact matte [6]. In the process of stereo conversion, hold-out mattes will cause double edges problem from a newly created view after conversion. In order to deliver uncut elements (elements without holdout mattes) for stereo conversion, those additional elements must be attached into the mini comp. Fig. 5 shows an element (left image), which needs to be cut by a pre-made matte (an image shows in the middle). A result on the right side of the Fig. 5 shows a cut image by the pre-made matte. Mattes are used for combining multiple images into a seamless single, final image in photography and visual effects filmmaking [7].

Fig. 5 An element cut by a pre-made matte

Fig. 6 shows another mini comp, which is more complicated in terms of the consideration of objects layered in real world space towards camera.

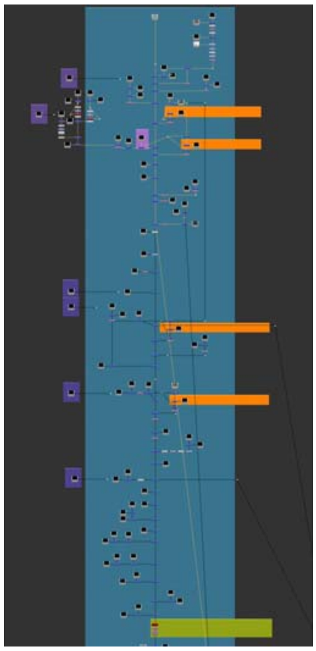

Fig. 6 A complicated mini script with many elements attached

2. Factor two

Roto shapes and Z-Depth files

In the field of visual effects, roto shape is another name of matte [8]. Mattes - in other words, also refers to masks, which is a grey scale image indicates in Fig. 7.

Fig. 7 A traditional matte

A simple mathematical equation (1) is applied in compositing in order to let mattes isolate parts of an image and make modifications, such as cutting unwanted parts of images through the mattes, and colour correction. In general, a colour image has three channels (R,G,B) contented. However, when the image is laid on top of other images, visual effects artist usually are asked to get rid of those unwanted parts of the top image and obtain a seamless composition with the background image, this process requires an extra channel – matte.

Composition result = (A x M) + [(1 – M) x B] (1) [9]

In equation (1), A means image A, B indicates image B and M means a matte (alpha or mask), (1- M) tells an inverted matte. The white area of the image in Fig. 7 indicates 1 and black area of the image means 0. In math, if any number multiplied by 1 the result will remind the number unchanged. However, if any value multiplied by 0 the result is 0. If we apply this equation into compositing, a specific area can be isolated by using its matte meaning a colour correction can be done through the matte and the result will only affect the area where has its matte attached. In addition, any number in between 0 and 1 can be thought as semitransparent areas.

Mattes for 3D conversion are required to have separated inner mattes for different parts of an image indicated in Fig. 8 in order to generate different depth values, which are correspondent with the distance to cameras. Fig. 9 indicates one form of z-depth files. Z-depth with a higher value means it is closer to the camera and appears to be brighter. Lower value z-depth indicates that the distance is further away from the camera and looks darker [10]. Another form of z-depth is that the nearer surfaces are darker; further surfaces are lighter. Depth files provide accurate values to determine a correct distribution of characters in the space for the second camera. Z-depth value can be stored in R,G,B alpha channels separately.

Fig. 8 Series of mattes created for 3D stereoscopic conversion purpose

Fig. 9 One form of z-depth, which is stored in R channel of a projected image through a 3D stereoscopic camera

3. Factor three

Parallax and clean plates

Parallax is a displacement or difference in the apparent position of an object viewed along two different lines of sight, and is measured by the angle or semi-angle of inclination between those two lines [11]. A new generated view in stereo will see things the original mono view is not able see, thus a clean plate is required in order to replace gaps created by the offset image. The Clean plate [12] is an essential element for 3D stereo conversion shots. Without clean plates, converted stereo images will show double edges of characters. In most cases, the double edges will appear on the newly created image through the offset view. See Fig. 10, Fig. 11 and Fig. 12.

Fig. 10 An issue of double edges from right view after stereo conversion. It caused by an improperly created clean plate

Fig. 11 A correct clean plate with no characters on it

Fig. 12 A final correct result in right view (as a newly offset view in this shot) after stereo conversion

III. 3D Stereoscopic Projection

3D projection is a successfully developed method, which has been using in many 3D stereo conversion films in visual effects industry. This hybrid method contains a series of key components.

3D projection requires intensively created internal and separated roto shapes for each object of a shot; clean plates; an in-house stereo conversion pipeline setup for re-projecting mono footage over 3d models, which were roughly built however with accurately animated movement for all moving characters for every scene. As we can see from Fig. 13 that all green nodes highlighted in the white rectangle indicate separated roto shapes for every part of the characters in one scene. Breaking down nodes in this way allows stereo artists to modify each part of the character shapes if needed.

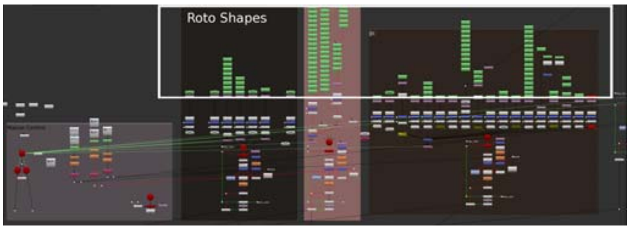

Fig. 13 Illustrates a group of separated roto shapes used in a 2D to 3D stereoscopic conversion script

Creating geometries in 3D applications is also a vital part component of stereo conversion. Meshes of the characters don't need to be built particularly accurately for the most part but must cover or even be slightly larger than the actual characters in the scene. However, it is important that the movement of characters and the camera is as accurate as possible even with subtle changes, otherwise the result after conversion will be that it will not stick to the surface of the projected characters. It's a time consuming part of the process, but additional data will also contribute to the final result, such as good matchmoving data.

Upper image in Fig. 14 shows a carefully created 3D modeling. Lower image indicates the modeling is set to match its original footage. As we can see from the lower image that the 3D modeling is slightly larger than the original footage in order to leave some extra space for receiving 3D projected image on it.

Fig. 14 A 3D modeling, and a set with the 3D modeling on top of its original footage

A. 3D Projection Setup

The 3D projection setup may be different in different conversion companies but key principles will remain the same.

An original camera will be used as a hero eye (as an original untouched camera), another new shifted camera will be created based on the hero camera. Fig. 15 shows a typical setup of a camera rig for stereo conversion using 3D projection approach. There are three cameras, the top one is the original camera (mono camera) for the film. Two cameras down below have been driven by the original camera with expressions. The left camera maintains the same position of the original camera but the positions and rotations on the right camera have been shifted by using the following python codes:

```
translate x: 0.1625
StereoCameraAxis.principalView
=1?0:StereoCameraAxis.interaxial
```

Above equation shows that the position of the right camera in X axis is shifted to 0.1625 against the left camera, if the original camera value equals 1, the position of right camera will be back to where the original camera is without position shift at all.

```
Rotate y: 0
StereoCameraAxis.mode
==2?StereoCameraAxis.toeInAngle:0
```

A similar setting has been applied to the camera rotation in Y axis on the right camera.

Fig. 16 indicates a master control of the stereo projection setup. All individual parts of the projections have been connected with a control centre: scene, as well as left and right cameras. Conversion artists can adjust all individual controls from the scene attributes, for instance; interocular distance, convergence point (focal plane), near plane and far plane of the camera, etc.

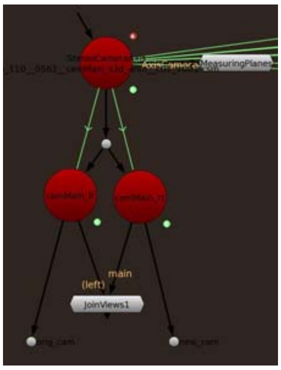

Fig. 15 A stereo conversion camera rig setup

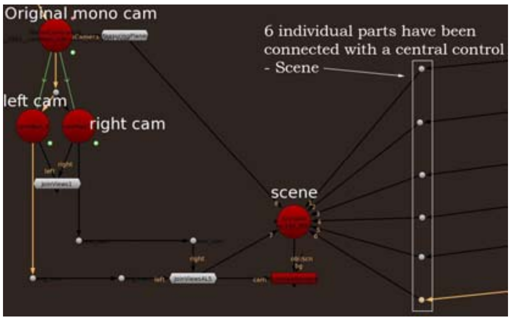

Fig. 16 A master control of the stereo projection setup

Fig. 17 An example of stereo conversion script

Fig. 17 indicates an example of a stereo conversion script. On the right side of Master Control section there are two independent projecting sections (the middle and right part). These two sections use a series of roto shapes to isolate necessary parts of the mono footage and re-project the mono footage carefully over the animated 3D models through the hub stereo camera. This process produces a number of newly created images with z-depth values included for both cameras which then recombine all parts of the images together, seamlessly producing an identical image of the mono footage and with all newly created z-depth files.

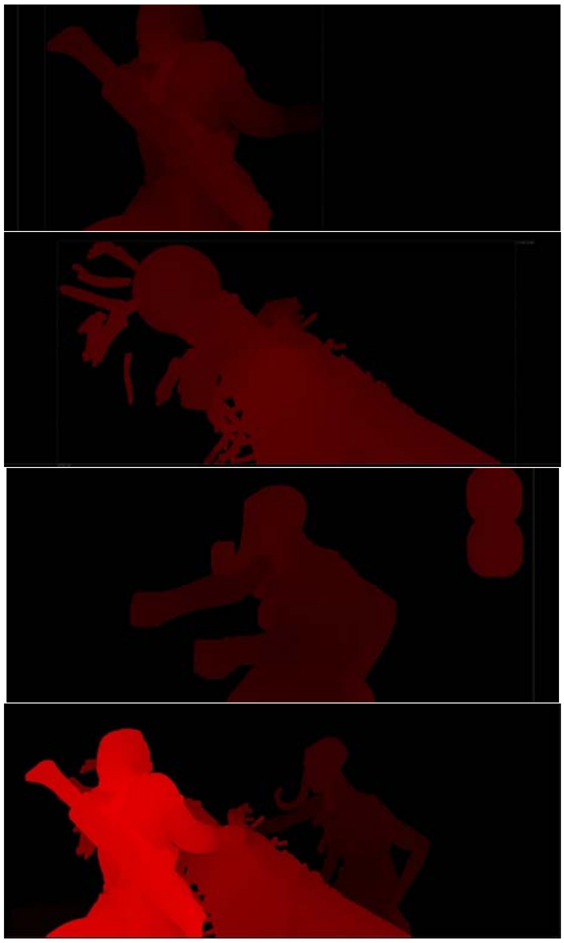

Fig. 18 Series of newly created z-depth files for the mono footage

In this example scene, the convergence point (focal plane) has been set in the middle of two characters so there are no obvious position shifts between these two characters. However those elements with a long distance from focal plane on the background have relatively big position shifts. As a result shown below in Fig. 19, the focal plane is set on the object between the two living characters. According to the principles of focal plane in 3D stereoscopic conversion, the object will be on the screen, however, the character on left side pops up in front of the screen and therefore the character on right side has been pushed to the back of the screen. We also notice that there is a black floating window, which is set for covering the entire image. The left and right side of this window can be position shifted in order to cover some parts of characters so they don't look like they are popping forward too much and potentially causing uncomfortable feelings in the audience.

Fig. 19 A final converted stereoscopic image (viewing through an anaglyph mode)

IV. Conclusion

The purpose of this study was to demonstrate a clear and good understanding of the connections between existing theories and professional practices in 2D to 3D stereo conversion. There are several approaches used for 2D to 3D conversion, most notably depth-based methods. This paper revealed a series of being successfully used methods, and furthermore, a stereo conversion pipeline setup in those major professional visual effects companies was also investigated in this study.

Several key principles and theories of stereoscopic and stereo conversion were laid out at the beginning of the paper. A theory which is to shift object positions in X axis in order to produce an illusion of stereoscopic can be seen in the first section of this study. However, professional visual effects companies have used a more complex hybrid conversion method and pipeline to create 2D to 3D stereo conversion shots. This pipeline involves several key factors, which also were separately discussed through the study.

In the key factors section of this study, two examples of simplified scripts for preparing stereo conversion were compared in factor one. Factor two explained the importance of roto shapes and z-depth in stereo conversion. Parallax and clean plate factors were studied in detail in factor three. An explanation of how the combination of four key factors of stereo conversion worked in a real world project was also demonstrated. Lastly, a detailed setup of a 3D projection script was reviewed, including virtual camera setup, driven python expressions on cameras, conversion script layouts and details as well as functions of each section of the script. A series of persuadable final results have been presented at the end of this study.

References

[1] Harper, D. (2001). Online etymological dictionary. Retrieved April 2, 2008.
[2] http://en.wikipedia.org/wiki/Binocular_disparity
[3] Qian, N., Binocular Disparity and the Perception of Depth, Neuron, 1997, 18, pp. 359-368,
[4] http://en.wikipedia.org/wiki/Focal_plane#Focal_planes
[5] Mike Seymour, Art of Stereo Conversion: 2D to 3D. Fxguide, 2012.
[6] Ron Brinkmann, The Art and Science of Digital Compositing. Elsevier Inc., 2008, pp. 223-224.
[7] Mark Cotta Vaz and Craig Barron, The Invisible Art: The Legends of Movie Matte Painting, Chronicle Books, 2002, pp. 33.
[8] Lee Lanier, Professional Digital Compositing. Wiley Publishing, Inc., Indianapolis, Indiana, 2010, pp. 206.
[9] Ron Brinkmann, The Art and Science of Digital Compositing. Elsevier Inc., 2008, pp.158.
[10] W.K. Giloi, J.L. Encarnação, W. Straßer. "The Giloi's School of Computer Graphics". Computer Graphics 35 4:12–16.
[11] Shorter Oxford English Dictionary. 1968. "Mutual inclination of two lines meeting in an angle".
[12] Ron Brinkmann, The Art and Science of Digital Compositing. Elsevier Inc., 2008, pp.353.