Review on 2D to 3D Conversion Techniques and other Dimensions

B. Gnana Priya[1], Dr. M. Arulselvi[2]

[1]Lecturer, Department of Computer Science and Engineering, Annamalai University, Tamil Nadu, India

[2]Assistant Professor, Department of Computer Science and Engineering, Annamalai University, Tamil Nadu, India

ABSTRACT

In modern era there has been a rapid progress in the field of 3D image processing. Many 2D to 3D conversion algorithm have been developed, which works by constructing depth maps. Also we are all familiar about 4d ,5d, 7d ,9d movies, gaming and various shows in malls and theme parks. The measure of spatial extent like length, height and width is called dimension. Though we hear about various dimensions in image processing, mathematically there are only three dimensions in which we can work with. The other dimensions are created with special effects , seat movements ,interactivity and so on.. This paper explores the various ways to create 3d content, explains the 3d file format and present a brief outline behind other dimensional technologies.

Keywords : Image Processing, 3-D Cameras, Local Point Transformation Function, PIC, KNN

I. INTRODUCTION

Nowadays 3D hardware popularity is increased and many of us intent to buy 3D capable hardware such as TVs, Cameras, smart phones ,gaming consoles and much more. 3D content used widely in medicine, engineering, earth science, architecture, video games, movies, printing and much more. The term stereoscopic vision[1] refers to the human ability to view with both eyes in similar, but slightly different ways. Eyes of any person are a few inches apart and this is called the interocular distance. This allows humans to judge distance, which develops their ability to have true depth perception. For people and animals that have stereoscopic vision, each eye sees a similar, but slightly different image. These positional differences are referred to as horizontal disparities or binocular disparities. The brain processes these images in a way that allows the person or animal to slightly see around an obstacle without having to physically move. This process is accomplished by matching the different images and then taking into account the differences in the two images. Although the visual differences are small, the end result is a significantly improved depth perception of one's surroundings. The 3D effect is created by taking two images from two lenses, creating a parallax effect. Parallax occurs from viewing an object along two lines of sight.

II. 3D TECHNOLOGY

3D image can be created using a 3D camera or 2D to 3D image convertor software can be used. 3D camera in current scenario are of two types. One with two lenses and two image sensors are available for simultaneous exposure of a stereo pair. The two lenses are set approximately the same distance apart as a pair of human eyes mimicking the sense of depth that we get when we view things

with our own eyes. The twin lenses capture two images simultaneously similar to how your right and left eye would see two different images and the 3D processor converts these two images into a single 3D image file. By layering the two 2D images on top of one another, the camera is able to create one single stereoscopic 3D image with depth. These cameras are more expensive due to the additional hardware and limited demand. Panasonic and Fujifilm markets double-lens 3-D cameras. Fujifilm FinePix Real 3D W3 and Panasonic Lumix 3D1 with 10 and 12 megapixels of resolution are the most popular of 3D cameras available. Recent trend adds 3-D assist capabilities to conventional cameras. The other type made inexpensively by taking two pictures with the same camera, but moving the camera a few inches either left or right. Nikon and Sony 3D cameras allow the photographer to shoot one scene, then shift the camera assisted by a translucent image on the camera display. After the subject is aligned, the camera exposes the second image, using the same exposure and focus. Sony SLT-A58K and DSLR Samsung NX300 with 2D/3D Lens are examples.

It is not the creation of 3D images that is problematic, but more the displaying of them. Today's 3-D cameras capture the images and do not perform any of the decoding functions needed to view the photos three-dimensionally. Images taken with any 3D camera can be viewed on a 3D TV, using special 3D prints or on the back of the camera as in Fujifilm. Active decoding viewing uses battery-powered glasses to sync polarized images with a 3-D television. Viewing 3-D TV without glasses can be accomplished through the use of lenticular screens, but the field of view for these TVs is limited, and the lenticular screen reduces image resolution. Other method make use of glasses to view the 3d image. Passive decoding technology uses either anaglyphic viewing the red

and cyan lensed glasses or polarized viewing, which uses gray lenses with different polarizing directions. Some of the drawbacks include : 3D TVs are expensive; 3D prints are time consuming to produce and expensive; Its not possible to view 3D image on our camera except Fujis Finepix. Other challenges are : some of the camera will work only in landscape orientation. Images taken with the camera in a vertical orientation are more difficult as your eyes cannot adjust to the parallax . Images with some depth will work better for examples medium distance photo shots. While others like macro shots, night time images and compositions where the subject is near the edge of the frame will not be as effective.

A. 3D Conversion Software

There are many 2d to 3d open source convertor software available. The convertor tool adds depth information to the 2d image fed to it. Different algorithms and techniques are used by different convertors to convert a given 2d image into a 3d image. some of the popular software available are Blender, Make3d, 3dShade, 123d catch, Insight3d, ConvertImage, etc. Some of them are fully automatic , have programs that automatically estimates depth for an image. Most of the convertors will get a 2d image captured from a normal camera and stored in PC in standard image format as input and adds depth using some techniques. Others are semi automatic, we have to assign depth to various parts of the image. We need to define few parameters like width , depth, position of object explicitly.

Some software make use of a single 2d image to create its 3d counterpart. The software creates picture of two colours, manipulate the color mixture and overlaps the two images(same image) that gives us the illusion of depth. Some of them calculates color intensities in picture and produce depth information with quite accurate perspective.

The software also allows us to select tones from hundreds of tones available in their database and render it with our image. The second method is to use series of 2d image of the scene, the software will automatically match them and measure position in space. Some of them are design tools, wherein the user draws the image as object in the screen. The camera angle, directional light are set by default and object converted into 3d image. Some of standard features includes allows us to customize shade, view or color special effects like crop, resize, rotate material finishing like wax, sandstone, steel possible to export and import the 3D files.

The software available will convert 2d formats like jpeg, png, gif, bmp, tif, pdf into 3d files. But cannot be guaranteed to achieve complete accuracy.

III. 2D TO 3D IMAGE CONVERSION METHODS

2D to 3D image conversion can be done in two ways. They are automatic and semiautomatic methods. In semiautomatic conversion method we need to assign depth to various parts of the image. Based on this sparse depth assignment, a computer algorithm estimates dense depth over the entire image or video sequence. In automatic methods, no human intervention is needed and a computer algorithm automatically estimates the depth for a single image or video. Automatic methods estimates depth from shading, perspective geometry, structure from motion , depth from models or depth from defocus.

A. Learning Depth from examples

A scene depth can be learned from a large database of images and depth pairs[2]. J.Konard proposed an automatic conversion of 2D to 3D image by machine learning techniques. The method relies on repository of images and depth pairs available. The assumption here is two images that are photo metrically similar are likely to have the same depth. While analyzing the available 3D pairs we may likely get a match for our 2D input image. The k-nearest neighbor algorithm is used to find the closest 3D image for any given input image. Euclidean distance function used to find difference between histograms of oriented gradients. Median filtering used for depth fusion of k images to compute the output. This method yields good depth quality but seldom have distortions.

B. Learning Based 2D to 3D Image Conversion

Janusz Konard[3] came up with other automatic 2D to 3D algorithm which has two methods. (i) Learning by a local point transformation function (ii) Learning by global
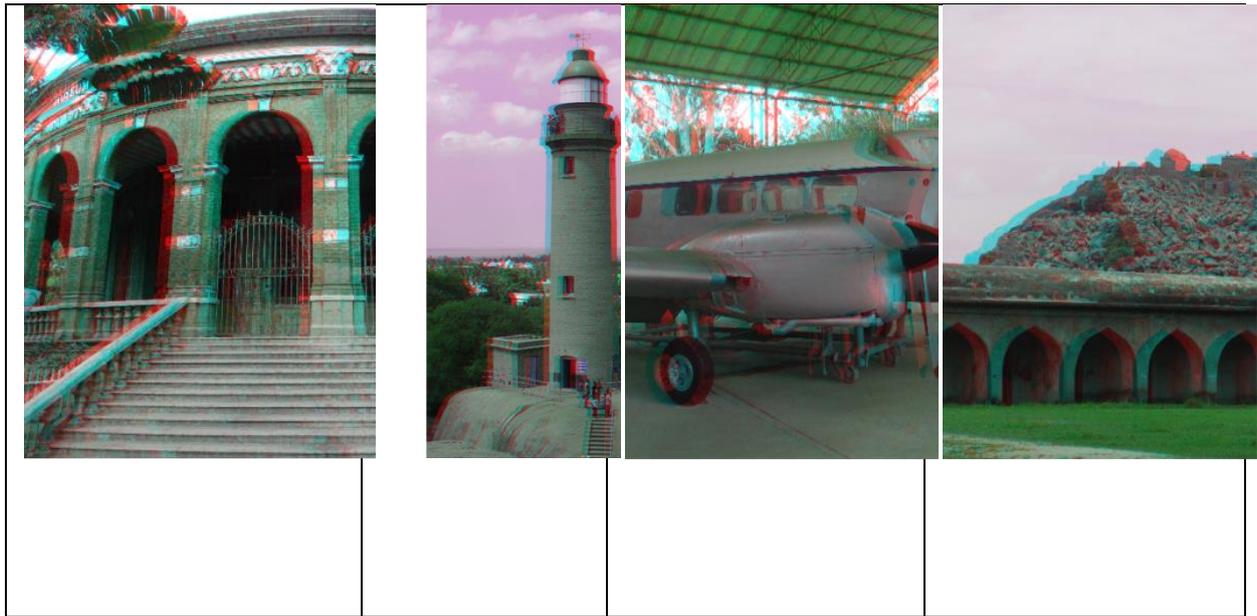

Figure 1. Original 2D Image

Figure 2. 3D Image using PIC 3D Software

Figure 3. 3D Image using Convert Image Software

nearest neighbor (KNN) algorithm. Local attributes like color, spatial position are used for learning a point transformation. We can estimate the transformation function by training the database. We can have separate transformation for each attribute and finally they are linearly combined to find the final depth. Advantage here is computational efficiency and reduced complexity.

C. Multiple Depth Cues

An image have various attributes and the depth perception arises from a variety of depth cues. Wang[4] proposed an automatic 2D to 3D conversion algorithm that uses multiple depth cues. Perspective geometry, defocus and visual saliency(subjective perceptual quality that makes an object distinct from background and other object in an image) are depth cues used here. The given color image is converted into grayscale. Then depth is detected from any of the three steps. (i) Vanishing point detection: If image are of perspective geometry, intersection tend to aggregate to one cluster. If predominant cluster exists

the scene contains an vanishing point. In this case a cone depth model and saliency map used to estimate depth. If no vanishing point we can estimate depth by defocus. Here two dimensional discrete cosine transformation used and results remapped to depth range 0-255. If image does not have vanishing point or defocus, depth extraction done using depth models. Spherical, Cylindrical and plane model used in this method.

D. Random Walks and Graph Cuts

This is a semi automatic method wherein user assigns a rough estimate of depth values to a few points in the image. The depth values for rest of the image determined by the system[5]. With the resulting depth map we can produce a 3D image. This method is based on the Random walker algorithm for image segmentation. The algorithm works by assigning small number of pixels called seeds with labels. The unlabelled pixels have a random walker and probability computed to find which seed it reaches first and is labelled. Graph cuts have hard segmentation and smoothness provided by random walks. Depth maps using graph cuts are generated with user defined depth values to generate depth prior. The depth prior and user stroke are integrated into Random walks as additional feature to determine edge weights. To produce quality depth maps Random walks are combined with Graph cuts.

IV. 3D CONTENT

What does a 3D file contains? is an important thing we need to know before handling them. The content of a 3D file can be classified into three categories- such as geometry, appearance and scene information.

A. Geometry

The geometry of a model is stored as a set of 3D points. There are various ways to represent surface information of a 3D image. A set of Polygons may be used to represent the surface of the model . These polygonal meshes are convenient and quick to render 3D content. This method is very simple and is widely used in 3D printing. Here use of more vertices yields better results but this may lead to larger storage space. Use of polygons not feasible for smooth surfaces. We can use Non uniform Rational B-spline patches to create smooth surfaces. Here, a number of control points each with its own weights and parameters known as knots are stored. We can easily edit the 3D model because we need to work only with control points to achieve transformation in a particular part of an image. But, the method is quiet time consuming. Now the latest method is to use constructive solid geometry, CSG. All the image elements are not treated as the same. Instead assign different 3D shapes to different objects based on their appearance. Simple primitive shapes like cylinder, sphere and cubes are taken and Boolean operations like addition and subtraction are performed to built 3D shapes. We must save each of the operations and shape information for future editing. Conversion from CSG format to other mesh type formats is a challenging task and is a subject under study. Constructive solid geometry is a user friendly and time saving method.

B. Appearance

Assigning texture to the surface of the model is more important, as in 3D we need to have a feel of the model. The texture and color for each vertices (or) knots (or) shapes need to be assigned by mapping each 3D vertex to a corresponding point in the 2D image. Light source plays a vital role in the appearance of an object. Colour of the object we perceive depends on the light source. Most of 3D file formats support texture mapping. The extra information may be stored in the same file or in a separate file.

C. Scene

Layout of the 3D model with respect to camera, light source and nearby objects is referred as the scene information of the 3D image. The direction of camera, its magnification, principal point and up

direction are needed to define the view of the camera. In order to define the scene the position of the light source, its intensity and colour need to be known. A single 3D image may have several parts. In such case we need to know how to layout the several parts and separate transformation matrix for each part need to be stored. Most of 3D file formats currently available does not support scene information.

V. 3D FILE FORMATS

3D file formats are used to store information about 3d models. Unlike 2d image files ,a 3D image file not only stores pixel values but some additional information are also added. This includes geometry, appearance and information about the scene. Table list some of popular 3D image formats available. Most of the 3D formats are optimized for use on the specific software alone. Different formats stores different attributes about the image. Therefore difficult to exchange files and view them on different software. 3D image created using one software cannot be viewed using another software. .BLEND file created by Blender cannot be viewed using AutoCad software. Some open source formats to solve this interoperability problem are also available. Neutral formats like STL and CoLLADA(.STL and .DAE extension respectively) may be used to exchange and view the 3D models. First convert 3D file into .DAE file and export it. Use other 3D software and import the .DAE file and convert them into the software respective format. Some of the 3D file formats are : .3D, .3D2, .3D4,.3DM, .3DX, .ANIM, .BLEND, .C3D, .D3D, .MESH, .P3L, .P3M, .STC, .U3D , .X3G.

VI. OTHER TECHNOLOGIES

A. 4D Technology

4D technology involves the use of already existing 3D tech with an additional dimension. For the most part 4D applications are simply a matter of marketing. For now a real 4D object, as defined in mathematics or physics, does not exist. 4D in 4D cinemas is simply 3D plus an additional dimension or element. This additional element is usually a simulation or group of simulations of environmental occurrences based on what is being projected. With 3d videos we use special seat effects. Movement to any direction as per the film, sways ,wave motion , tilts, jumping, bending, racing, tumbling all in sync with screen activity. In addition to these leg ticklers, back ticklers, neck blast are also added.

B. 5D Technology

5D techniques include

i) Seat effects

All the effects like 4D are used. A high grade chair with a driver unit and air tank aids in seat movement.

ii) Special effects

Effects like smoke, rainfall, snow, bubble, wind, and lightning are added.

iii) Sound effects

Surround sound components, speakers with special processors and amplifiers are used.

iv) Screen and projection

Two high quality projectors are used for viewing the 3D movie on display which can be viewed by special 3D glasses that offer passive polarization effect. This 3D projection is controlled by a special 3D computer that keeps streaming both the left and right eyes simultaneously for the effect. We also need customized silver screens that offer an image bouncing effect which is necessary for creating the 3d effect.

v) Control system

A motion control personal computer that syncs the chairs with the screen action is used. The control system also syncs the video and the special effects.

C. 7D Technology

7D is an interactive cinema where we can use the toy gun to shoot the target. The equipment included in 7d theatre are a Master System , projection system , sound system ,6 DOF dynamic seat system , environment , interactive simulation system . The main control system is mainly in the computer software and it controls the operation of the entire theatre, Projection systems, including the screen and projector.

D. 9D Technology

Now, the 9D is the newest technology in the world. Like our 9d virtual reality, viewers just need to wear VR glasses and then go into the world of the movies. It breaks the limit of virtual and real world and takes viewers amazing experience. It consists of three revolutionary new products: Interactive cinema with 360° rotation function, Immersive Glasses, Breakthrough VR entertainment content platform.

VII. CONCLUSION

3D technology gives us a realistic view of things and in near future will be a very common application in many fields. Glasses free 3D display using plastic films help us to view 3D content without much difficulty. The other dimensional technologies are becoming much popular and most of us excited to view such shows. This paper reviews the basic concepts in 3D processing and rather explains the techniques behind how other dimensions are implemented in real time.

VIII. REFERENCES

[1] https://en.wikipedia.org/wiki/Stereoscopy
[2] J. Konrad, M. Wang, and P. Ishwar, "2D-to-3D image conversion by learning depth from examples," in Proc. IEEE Computer. Soc. CVPRW, Jun. 2012, pp. 16-22.
[3] J. Konrad, M. Wang, and P. Ishwar,C.Wu,D. Mukharjee, "Learning based, automatic 2D-to-3D image and video conversion," in Image Processing IEEE Trans on,vol.22,no.9,pp.3485-96,Sept. 2013 Jun. 2012, pp. 16-22.
[4] P.Ji,L.Wang,D. Li, M. Zhang, "An automatic 2D to 3D conversion algorithm using multi-depth cues," IEEE Conf.Audio,Language and Image Processing,pp.546-50,July 2012.
[5] R. Phan, R. Rzeszutek, and D. Androutsos, "Semi-automatic 2D to 3D image conversion using scale-space random walks and a graph cuts based depth prior," in Proc.18th IEEE Int. Conf. Image Process., Sep. 2011, pp. 865– 868.
[6] Kenton McHenry and Peter Bajcsy "An Overview of 3D Data Content and Viewers", National Centre for Supercomputing Applications, 2008.
[7] http://edutechwiki.unige.ch/en/3D_file_format
[8] http://convertimage.net
[9] http://pic3d.egedsoft.com