International Journal of Research in Engineering, Science and Management
Volume-3, Issue-1, January-2020
www.ijresm.com | ISSN (Online): 2581-5792

2D-3D Image Conversion

S. R. Ruma
Student, Department of Electronics and Communication Engineering, Marian Engineering College, Thiruvananthapuram, India

***Abstract*: This paper presents an overview on 2D-3D image conversion.**

***Keywords*: Image conversion.**

1. Introduction

2D to 3D image conversion is the process of transforming 2D image into 3D form. The 3D incorporates the third dimension of depth which can be perceived by the human vision in the form of binocular disparity. The 3D display required the depth information which is unavailable in the conventional 2D image. With the development of 3D application, the conversion of existing 2D images to3D images become an important component of 3D content production. The conversion process of existing 2D image to 3D is commercially viable and is fulfilling the growth of high quality stereoscopic images. The dominant technique for such content conversion is to develop a depth map for each frame of 2D material.

2. Definitions

A. Binocular disparity

With two images of the same scene captured from slightly different view point, the binocular disparity can be utilized to recover the depth of an object. Thus is the main mechanism for depth perception. First, a set of corresponding points in the image pair are found. Then by means of triangulation method, the depth information can be retrieved with high degree of accuracy when all the parameters of the stereo systems are known. When only intrinsic camera parameters are available, the depth can be recovered correctly upto a scale factor. In the case, when no camera parameters are known, the resulting depth is correct up to a projective transformation.

B. Depth map

Depth Map in 3D computer graphics is an image channel that contains information relating to the distance of the surfaces of scene objects from the main point.

Fig. 1. 2D image and its depth

3. Conversion methods

The work describes an efficient 2D-to-3D conversion method based on the use of edge information. Importantly, the edge of an image has a high probability as it can be the edge of the depth map. Once the pixels are grouped together, a relative depth value can be assigned to each region. Initially, the block-based image is considered to segment it into multiple groups. Then the depth of each segment is assigned with the help of an initial depth hypothesis. Next, the blocky artifacts have to be removed using cross bilateral filtering. Finally, multi-view images are obtained by the method of DIBR. As a result, the input 2D image is converted into visually comfortable 3D image without the presence of artifacts enhancing the quality of the image in the display.

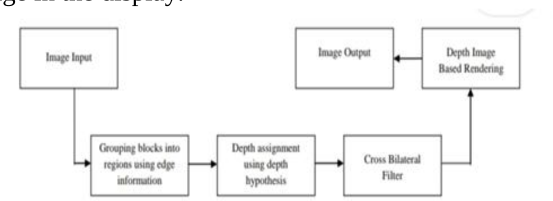

Fig. 2. Block diagram of image conversion system

A. Block-based region grouping

Computational complexity is reduced mainly by block-based algorithm. This implies each pixel in the same block has the same depth value. A 4-by-4 block is used as an example. Each node is a 4-by-4-pixel block, and is four-connected. The value of each link is calculated by considering the absolute difference of the mean of neighboring blocks:

$$\text{Diff}(a, b) = |\text{Mean}(a) - \text{Mean}(b)|$$

where a and b denote two neighboring blocks, respectively, and Mean (a) represents the mean color of a.

A smaller value obtained implies a higher similarity between the two blocks. Following calculation of the absolute difference of the mean of the neighboring blocks, the blocks are then segmented into multiple groups by applying the minimum spanning tree (MST) segmentation. Initially, a minimum spanning tree is constructed. And the multiple grouped regions are generated by removing the links of stronger edges. The MST algorithm is mainly used to identify the coherence among blocks with both the color difference and the connectivity without generating many small groups. This algorithm preserves the link connectivity and also has an excellent result in spatial locality. A situation where MST algorithm possibly fails is when a series of regions, which differ slightly from each other, bridge a large gap in the feature space. As the edge weights are small, the graph will not be cut and so very different objects in terms of feature space characteristics will be merged. Recursive implementations of the MST algorithm, recalculating edge costs, can easily tackle the problem increasing complexity. With the efficient linkage preserving property, the MST segmentation method can generate excellent grouping results. The proposed depth generation method can also be substituted by other automatic or manual segmentation with satisfactory results.

B. Depth from prior hypothesis

The extraction of depth is the crucial one in the conversion process. The greatest difference between 2D and 3D image is the depth information. The object can jump out of the screen and look like a real life due to the depth information. If we extract these depth signals and integrate them together, we will build a strong foundation to make 3D images of better and higher quality. The depth generation algorithms are roughly classified into three categories which utilize different kinds of depth cues: the binocular, monocular and pictorial depth cues. Each signal represents different depth information. In this conversion process, following the generation of the block groups, the corresponding depth for each block is then assigned by the depth gradient hypothesis. The process includes the generation of gradient planes, depth gradient assignment, consistency verification of the detected region, and finally the depth map generation. When each scene change is detected, the linear perspective of the scene can be analyzed with the help of line detection algorithm using Hough transform.

C. Bilateral filtering

The bilateral filter is non-iterative and also achieves satisfying results with only a single pass. This makes the filters parameters relatively intuitive as their effects are not cumulated over several iterations. The bilateral filter has proven to be much useful although it is slow. It is nonlinear and also its evaluation is computationally expensive because the traditional accelerations like performing convolution after an FFT, are not applicable. Nonetheless, solutions have been proposed later in order to speed up the evaluation of the bilateral filter. Unfortunately, these methods seem to rely on approximations that are not grounded on firm theoretical foundations. Among the variants of the bilateral filter, this conversion method has selected the cross bilateral filtering. In some applications like computational photography, it is often useful to decouple the data to be smoothed to define the edges to be preserved. The chosen cross bilateral filter is a variant of the classical bilateral filter. This filter is used to smoothen the image to locate the edges to preserve. The depth map generated by block-based region grouping contains blocky artifacts. Here, the blocky artifacts are removed by using the cross bilateral filter.

D. Depth image based rendering

The filtered depth map has a comfortable visual quality because the cross bilateral filter generates a smooth depth map inside the smooth region with similar pixel values and preserves sharp depth discontinuity on the object boundary. Following filtering by the cross bilateral filter, the depth map is then used for the generation of the left/right or multi-view images using depth image-based rendering (DIBR) for 3D visualization

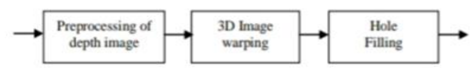

Fig. 3. Block diagram of DIBR

DIBR for advanced 3D TV System can be illustrated by the following block diagram. This system includes three parts, pre-processing of depth map, 3D image Warping and Hole Filling. Smoothing filter is first stage applied to smooth the depth image. Then the 3D image warping generates the left and right view according to the smoothed depth map and also intermediate view. If there are still holes in the image, hole filling is then applied to fill color into these holes.

1) Pre-processing of depth image

Pre-Processing of depth image is usually a smoothing filter. Because depth image with that of the horizontal sharp transition may result in big holes after warping, smoothing filter is applied to smooth sharp transition to reduce the number of big hole. However, if we blur the depth image, we will not only reduce big holes but also degrade the warped view because the depth map of non-hole area is smoothed.

2) 3D image warping

3D image warping maps the intermediate view pixel by pixel to left or right view according to the pixel depth value. In the other words, 3D image warping transforms the location of pixels according to depth value.

3) Hole filling

Average filter interpolation method is a common method for Hole-Filling in DIBR. However, using average filter would result in artifacts at highly-textured areas. Besides, hole-size in DIBR is so huge such that it is needed to using average filter with large window size. At the same time, the average filter with large window size is unable to preserve edge information for the reason that edge information is blurred.

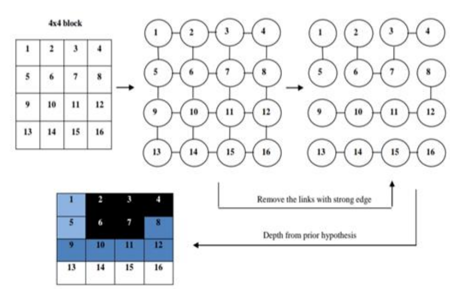

Fig. 4. Flow diagram of image conversion system

International Journal of Research in Engineering, Science and Management
Volume-3, Issue-1, January-2020
www.ijresm.com | ISSN (Online): 2581-5792

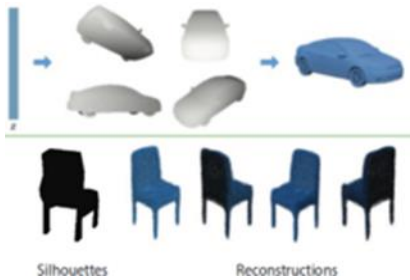

Fig. 5. Reconstruction of the 3D image

4. How to convert 2D image into 3D image

Step-1: Choose an image


Fig. 6. Choose an image

Upload an image only in JPG and PNG formats. The image should have a clear solid background.

Step-2: Choose proper settings

After uploading the image, a relief will be automatically generated. If we don't like the result, we can experiment with the following features:

Select relief direction: This feature works based on the light in the image. It we want the light areas of the image to be sunken and dark areas to be raised, then we should select 'Light areas sunken'. If we want the light areas of the image to be raised and dark areas to be sunken, then you should select 'Light areas raised'.

Select mode: This feature allows us to select between monochrome and colored. Monochrome uses only the one color and is recommended for standard desktop 3D printers and CNC machines.

Denoise factor: The denoise factor can reduce visual distortions or the appearance of noise in images and make it more accurate for manufacturing.

Smoothness factor: The smoothness factor can be compared to a blur tool in photo editing. It smooths edges and roughness, but some details may disappear as well. The greater the smoothness factor, the smoother the relief will be.

Quality: This feature allows us to choose the quality and size of the file: Small, Recommended and Extra. A small file will be ready instantly. The extra file will take more time to prepare, but it will also give the best quality as well.

Step-3: Manufacturing settings

Plate size: We should keep in mind that 3D printers and CNC machines have limits on the maximum size of an object they can produce. Don't overdo it.

Plate thickness: The thickness of the plate can be as thick as we want but similar to the plate size, it can't be too thick or else it can't be manufactured.

Relief depth: The relief depth determines the depth of the relief from the rear of the plate. The greater the depth, the deeper our relief will be. Minimum setting is 1mm.

Step-4: Have it manufactured

Once we are happy with the appearance of the relief, we turn the 3D model into a real object using manufacturing services.


Fig. 7. 3D model manufactured

5. Importance and applicability

With the increase of films released in 3D, 2D to 3D conversion has become more common. The majority of non-CGI stereo 3D blockbusters are converted fully or atleast partially from 2D footage. Even avatar contains several scenes shot in 2D and is converted to stereo in post –production. The reason for shooting in 2D instead of stereo are financial, technical and sometimes artistic:

Stereo post production workflow is much more complex and not as well established as 2D workflow, requiring more work and rendering. Professional stereoscopic rigs are much more expensive and bulky than customary monocular cameras. Some shots, particularly action scenes, can be only shot with relatively small 2D cameras.

Stereo cameras can introduce various mismatches in stereo image that should be fixed in post-production anyway because they ruin the 3D effect. This correction sometimes may have complexity comparable to stereo conversion.

Stereo cameras can betray practical effects used during filming. The scene filmed in stereo would reveal that the actors were not the same distance from the camera.

By their very nature, stereo cameras have restrictions on how far the camera can be from the filmed subject and still provide acceptable stereo separation. However, while the zoom lens would provide acceptable image quality, the stereo separation would be virtually nil over such a distance

Even in the case of stereo shooting, conversion can frequently be necessary. Besides the mentioned hard-to-shoot scenes, there are situations when mismatches in stereo views are too big to adjust and it is simpler to perform 2D to stereo conversion, treating one of the views as the original 2D source.

6. General problems

Without respect to particular algorithms, all conversion workflow should solve the following tasks:

Allocation of 'depth budgets'- defining the range of permitted disparity or depth, what depth value corresponds to the screen position, the permitted distance ranges for out- \of-the-screen effects and the behind-the -screen background objects. If an object in stereo pair is in exactly the same spot for both eyes, then it will appear on the screen surface and it will be in zero parallax. Objects infront of the screen are said to be negative parallax and background imagery behind the screen is in positive parallax. There are the corresponding negative or positive offsets in object positions for left and right eye images. Control of comfortable disparity- depending on scene type and motion – too much parallax or conflicting depth cues may cause eye-strain and nausea effects.

Filling of uncovered areas – left or right view images show a scene from a different angle, and parts of objects or entire objects covered by the foreground in the original 2D image should become visible in a stereo pair. Sometimes the background surfaces are known or can be estimated, so they should be used for filling uncovered areas. Otherwise the unknown areas must be filled in by an artist or in painted, since the exact reconstruction is not possible. High quality conversion methods should also deal with many typical problems including:

- Translucent objects
- Reflections

Fuzzy semitransparent object borders – such as hair, fur, foreground out-of-focus objects, thin objects, Film grain (real or artificial) and similar noise effects, Scenes with fast erratic motion. Small particles – rain, snow, explosions and so on.

7. Automatic conversion

A. Depth from motion

It is possible to automatically estimate depth using different types of motion. In case of camera motion, a depth map of the entire scene can be calculated. Also, object motion can be detected and moving areas can be assigned with smaller depth values than the background. Occlusions provide information on relative position of moving surfaces.

B. Depth from focus

Approaches of this type are also called "depth from defocus" and "depth from blur". On "depth from defocus" (DFD) approaches, the depth information is estimated based on the amount of blur of the considered object, whereas "depth from focus" (DFF) approaches tend to compare the sharpness of an object over a range of images taken with different focus distance in order to find out its distance to the camera. DFD only needs two or three at different focus to properly work, whereas DFF needs 10 to 15 images at least but is more accurate than the previous method. If the sky is detected in the processed image, it can also be taken into account that more distant objects, besides being hazy, should be more desaturated and more bluish because of a thick air layer.

C. Depth from perspective

The idea of the method is based on the fact that parallel lines, such as railroad tracks and roadsides, appear to converge with distance, eventually reaching a vanishing point at the horizon. Finding this vanishing point gives the farthest point of the whole image. The more the lines converge, the farther away they appear to be. So, for depth map, the area between two neighboring vanishing lines can be approximated with a gradient plane.

Fig. 8. An example of depth map

8. Conclusion

This work has presented a 2D-to-3D conversion algorithm. The proposed algorithm utilizes edge information to group the image into coherent regions. A simple depth hypothesis is adopted here to assign the depth for each region and a cross bilateral filter is subsequently applied for removing the blocky artifacts. The proposed algorithm is quality-scalable as it depends on the block size. Smaller block size will result in better depth detail and that of the larger block size will have lower computational complexity. Capable of generating a comfortable 3D effect, this proposed algorithm is highly promising for 2D-to-3D conversion in case of 3D applications

References

[1] H. Murata et al., "Conversion of Two-Dimensional Images to Three dimensions", SID Digest of Technical Papers, vol. 26, pp. 859-862, 1995.
[2] H. Murata, X Mori, S. Yamashita, A. Maenaka, S. Okada, K. Oyamada, S. Kishimoto, "A Real-Time 2-D to 3-D Image Conversion Technique Using Computed Image Depth," Sanyo Electric Co., Ltd., Osaka, Japan.