DEPTH MAP QUANTIZATION – HOW MUCH IS SUFFICIENT?

Ianir Ideses, Leonid Yaroslavsky, Itai Amit, Barak Fishbain

Dept. of Interdisciplinary Studies, Tel-Aviv University, Israel

ABSTRACT

With the recent advancement in visualization devices over the last years, we are seeing a growing market for stereoscopic content. In order to synthesize 3D content, one needs to have either a stereo pair or an image and a depth map. Computing depth maps for images is a highly computationally intensive and time-consuming process. In this paper, we describe results of an experimental evaluation of depth map data redundancy in stereoscopic images. In the experiments with computer generated images, several observers visually tested the number of quantization levels required for comfortable and quantization unaffected stereoscopic vision. The experiments show that the number of depth quantization levels can be as low as only a couple of tens. This may have profound implication on the process of depth map estimation and 3D synthesis.

***Index Terms*—** Stereoscopic visualization, Anaglyphs, Quantization, Depth maps

1. INTRODUCTION

As digital imaging techniques gain popularity and acceptance, rendering the race for increased spatial resolution marginal, the field of 3D imaging and visualization remains one of the last uncharted frontiers. Interest in 3D visualization and 3D content has been constantly growing as imaging devices were developed, even though a killer application for 3D visualization is yet to be developed. It is obvious, however, that future imaging devices will not be confined to 2D, but rather will capture 3D images to be displayed either in 2D or 3D at will.

In order to create 3D content, one usually has to acquire 2 images (for still images) or 2 video streams. This has a substantial implication on the process of image/video acquisition. In order to acquire 3D content in this manner, 2 acquisition devices must be used simultaneously and with great care given to synchronization, both temporally and in terms of focal properties (focus, zoom, etc'). In addition such a setup allows only a predefined parallax baseline to be used and does not allow for multi-view displays.

A different approach is to compute for each frame a depth map. This can be done by using a range finder, eg. a LIDAR (Laser RADAR) system, or using 2 cameras and computing depth maps for each stereo pair. Once a depth map is acquired for every frame, it can be used to construct its artificial stereo pair. This approach is superior to using the original stereo pairs in that that it allows transformations on the depth and manipulations on the parallax baseline.

There are many methods to compute depth maps, among them are the works of Lucas and Kanade [1], Horn and Schunck [2] Senthil Periaswamy and Hany Farid [3], Yu-Te Wu et al [4], L. Alvarez et al [5], Jochen Schmidt [6] and Adee Ran and Nir Sochen [7]. This is a greatly advancing field and a lot of effort is guided in this direction.

It is now well known that stereo pairs are extremely redundant. In particular, it has been shown by Yaroslavsky et al [8,9] that one of the images can be greatly defocused before any harm is caused to the 3D perception. Additional work on the effect of compression of depth maps has been conducted by Ulrich Fecker et al [9].

However, no work on the effect of depth map quantization on 3D visualization has been reported in the literature. In this paper we present results of our investigation into this issue.

2. EXPERIMENTAL SETUP

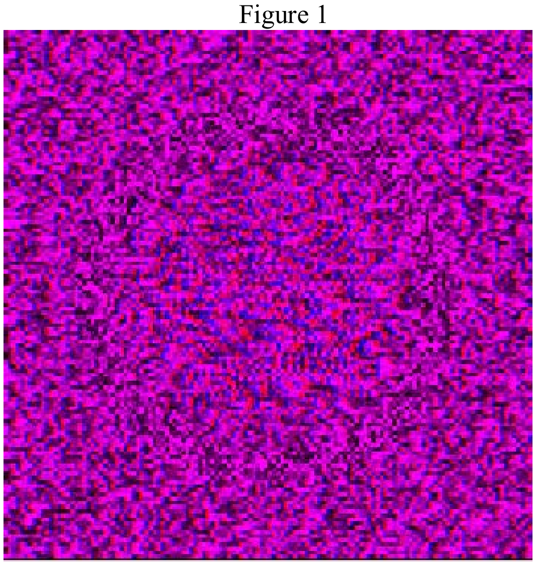
In order to test the effect of depth map quantization on 3D perception one must use a 3D visualization device. In previous experiments (reported in [9]), anaglyphs and stereoscopes were used. Both visualization techniques exhibited similar results. For experiments described in this paper, we used anaglyphs as they are most suited for use on any monitor/projection system. Although anaglyphs usually cause color loss, this was not a drawback in this case as the only measured parameter was depth.

In order to eliminate all depth cues but stereopsis, random dot images were used. This is in accordance with the methodology of B. Julesz [11].

1-4244-0722-2/07/$20.00 ©2007 IEEE

The experiments were performed using a standard PC running Matlab. This was done by a program that generates depth maps of different shapes (sphere, cone, etc') quantizes the depth maps using different numbers of quantization levels and displays artificial 3D images created using these depth maps to the viewer side by side**.** The viewer has to specify which image is "better". The quality measure is based on the perception of continuity of depth. Because the basic parallax is a dominant parameter in these tests, it was selected as the highest possible that still enables comfortable viewing. An example of such a 3D image is shown in figure 1.

Figure 1



An example of a 3D image (viewed with anaglyph glasses, red filter for red eye, blue filter for right eye).

In each experiment round, the viewer had to indicate which image has more quantization levels, if correct in his choice, the program would increase the quantization levels until the viewer can not distinguish between the images. If the viewer had been incorrect in his choice, the program would reduce the quantization levels until the viewer would again be able to distinguish between the images. Each experiment is composed of several tens of rounds until the quantization levels converge. In order to increase the reliability of the selection, the viewer was prompted to verify his answer to each round. A screenshot of the program is shown in figure 2.

Because parallax plays a considerable role in 3D perception, we had to fix the parallax level to a comfortable setting. For this, the parallax (expressed in pixels) of 17 pixels was chosen. This value corresponds to comfortable viewing of anaglyphs using a standard 19" CRT screen at a distance of 30 cm and was viewer selected.

The process of synthesizing a synthetic stereo pair from an image and a depth map requires resampling. This resampling is achieved by using the depth map as a map that defines a resampling grid. The only limitation to this simple resampling is that, because of quantization, only integer valued depth map values can be used. We found that the maximum comfortable viewable parallax is of the order of 17 pixels, the maximum levels of quantization can not exceed the parallax value – namely 17. In order to solve this problem, resampling was performed on an interpolated version of the image. We used discrete sinc-interpolation to expand the image, for resampling, by ten fold. This enabled us to use 170 quantization levels. In this fashion we were able to shift pixels according to the depth map with 1/10-th sub pixel accuracy.

3. RESULTS

The experiments were performed by 5 viewers, each viewer ran several tests (5-10), each test comprising of several tens of rounds (30-50). Overall, several hundreds of rounds were performed.

The results of one of these tests are shown in figure 3. The figure axis shows the level of quantization vs round. In each round the viewer was asked to report his perception of continuity. The program reduced the levels of quantization until depth quantization effects were detectable. As can be seen, 20-30 quantization levels are sufficient for visualization of 3D scenes. This range was observed for several viewers as well, with similar results.

These results were repeated for several viewers and several types of objects – cone, sphere, pyramid etc'. We found that for objects that have the largest slope (cone) we need the highest number of quantization levels. For objects that have lower height gradients, even less number of quantization levels are required.

4. DISCUSSION

As shown in the results section, 20-30 quantization levels of the depth map values are sufficient for viewing comfortable representation of 3D images and video. Using this data may have a significant effect on depth map computation and 3D synthesis.

Many modern depth map computation techniques rely on iterative algorithms. Among these are Horn and Schunk [2], Senthil Periaswamy & Hany Farid [3] etc'. Iterative algorithms are bounded by a stopping criteria and have a step size for each iteration. It is therefore possible, to

Figure 2

This figure shows an example of images that were shown to the viewer. The viewer had to indicate which image has a smoother (quantized with more quantization levels) depth map(viewed with anaglyph glasses, red filter for red eye, blue filter for right eye).

compute the required step by dividing the maximal parallax value by the number of quantization levels.

In addition to depth map calculation, depth map quantization may be important for artificial stereo pair synthesis. In our process of stereo pair synthesis we used sub pixel accuracy for the testing. Our results, however, indicate that, for a parallax suitable for visualization, shifting of pixels does not require sub pixel accuracy. This has implications on the computational complexity required for 3D video synthesis.

In related work [12], we have shown that by using compressed video formats, we are able to synthesize 3D video from 2D video in near real-time (15 VGA frames per second). In this algorithm, most of the computation power is spent in the resampling stage, namely in interpolating the image prior to resampling. The results acquired in these new tests shed a new light on this. If interpolation is not necessary for 3D synthesis, we can greatly reduce the computational load in the 3D synthesis stage and achieve above real-time performance. We suggest that similar performance increases are to be expected for other applications as well.

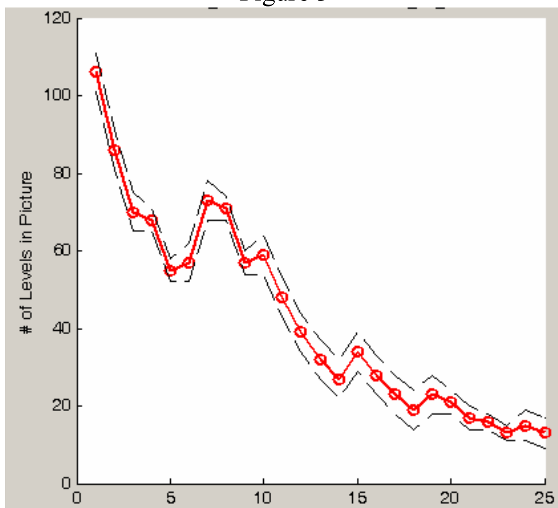
5. CONCLUSION

In this paper we have tested the effect of quantization of depth maps on 3D perception. In order to test this, we wrote a program that ran several tens of rounds for each test, where, in each test, the viewer was prompted for his evaluation of the depth map quantization.

The results of these tests show that, for depth map quantization, a relatively low number of about 20 quantization levels of depth map are sufficient for 3D synthesis. This number was acquired for shapes with high height gradients and is lower for other shapes.

The obtained results can be utilized in different applications, and especially in iterative algorithms of depth map computation and in the process of generating artificial stereo pairs from an image and a depth map.

Figure 3



An example of the resulting graph of a quantization test. The X axis shows the number of the round, the Y axis shows the quantization levels. The envelope surrounding the graph shows the quantization level difference between the images that are evaluated as having the same quality.

As can be seen, the program started with 110 quantization levels of depth map and the viewers were unable to accurately tell which image had more quantization levels so the program reduced the quantization level for each round.

The plateau was reached at around 20 quantization levels.

11. REFERENCES

[1] Lucas, B., and Kanade, T. "An Iterative Image Registration Technique with an Application to Stereo Vision*", in Proceedings of 7th International Joint Conference on Artificial Intelligence* (IJCAI), pp. 674-679, 1981.

[2] B. Horn and B. Schunck, "Determining Optical Flow", *Artificial Intelligence*, 17:185-203, 1981.

[3] Senthil Periaswamy, Hany Farid, "Elastic Registration in the Presence of Intensity Variations", *IEEE Transactions on Medical Imaging*, , Volume 22, Number 7, July 2003.

[4] Yu-Te Wu, Takeo Kanade, Ching-Chung Li and Jeffrey Cohn, "Image Registration Using Wavelet-Based Motion Model", *International Journal of Computer Vision*, July 2000.

[5] L. Alvarez, R. Deriche, J. Sanchez, and J. Weickert, "Dense Disparity Map Estimation Respecting Image Discontinuities: A PDE and Scalespace Based Approach. Technical Report" *RR-3874, INRIA*, January 2000.

[6] Jochen Schmidt, Heinrich Niemann, and Sebastian Vogt. "Dense disparity maps in real-time with an application to augmented reality. Orlando", FL USA,. *IEEE Computer Society*, December 3-4 2002.

[7] Adee Ran, Nir A. Sochen, "Differential Geometry Techniques in Stereo Vision" *Proceedings of EWCG*, pp 98-103, 2000.

[8] L. P. Yaroslavsky, "On Redundancy of Stereoscopic Pictures," *Acta Polytechnica Scandinavica*, n. 149. Image Science '85. Proceedings. Helsinki, Finland., V. 1, p. 82-85, 11-14 June 1985.

[9] L. P. Yaroslavsky, J. Campos, M. Espínola, and I. Ideses, "Redundancy of stereoscopic images: Experimental evaluation," Opt. Express 13, 10895-10907, 2005.

[10] U. Fecker, A. Guenegues, I. Scholz, A. Kaup "Depth Map Compression for Unstructured Lumigraph Rendering", J. G. Apostolopoulos, A. Said, editors, Proceedings Visual Communications and Image Processing, pages 60770B-1 - 60770B-8, San Jose, USA, January 2006.

[11] B. Julesz, "Foundations of Cyclopean Perception", The University of Chicago Press, p. 96, 1971.

[12] Ianir Ideses, Leonid Yaroslavsky, Roni Vistuch and Barak Fishbain, "3D from compressed 2D video", Proceedings of Stereoscopic Displays and Applications XVIII, SPIE and IS&T, San Jose California, 2007 (In print).