3D DISPARITY VISUALIZATION FOR STEREOSCOPIC VIDEOS

Lars Schnyder

Semester project in Multimedia Communications
ETH Zurich

1. INTRODUCTION AND MOTIVATION

3D movie quality gains a lot of importance with today's rising sale of 3D movies and 3D TVs. Unfortunately not everybody has the same experience while consuming 3D content. Complaints about headaches and nausea are very common and ask for a remedy. Due to the fact that viewers get more and more exposed 3D content, it is increasingly important for the producers to improve deficiencies in the viewing experience to keep their customers satisfied. One source for these deficiencies are the additional properties stereoscopic film production introduces to the present monoscopic film-making. While knowledge about the causes is growing, dealing with them often results in technical problems. In order to achieve best 3D video quality all of them have to be considered.

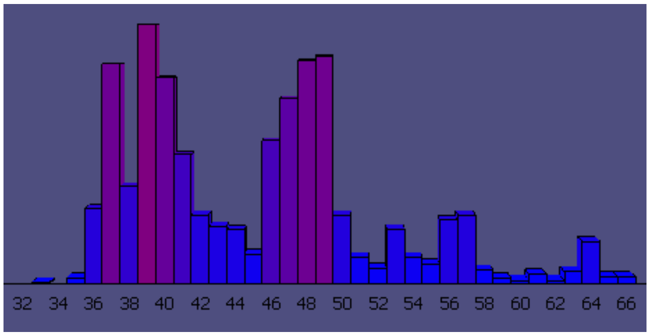

Fig. 1. Depth histogram of one single frame.

The result of this project is a tool that is able to help stereographers create more sophisticated stereo footage by visualizing depth information in stereo video a with the help of a video depth histogram. A frame depth histogram visualizes frame pixels and their corresponding depth in the scene (Figure 1). Visualizing depth information of only one single image is quite simple and very common but does often not satisfy some of the needs since depth information like transitions between frames are not shown. As a result of this I created a 3D visualization which contains the depth histograms of all frames in the video together. This shall enable the editor to see information like the used depth bandwidth and to find and correct undesired temporal effects such as abrupt depth changes.

2. METHOD

The project consists of two steps. At first we need to calculate the single pixel disparities in a frame. This we have to do in a way such that the calculation represent the whole frame depth. The second part is the visualization of this previously gained information.

2.1. Frame Disparities

While the 3D visualization itself is quite straight-forward, the gathering of all information (i.e. frame disparities) is more complex. To retrieve the required frame disparities I used an existing disparity mapping framework. The technique it uses to find the disparities shall be explained in the following paragraphs. In theory feature correspondences from the left to the right stereo image can be used to find matching pairs of pixels which, if subtracted from another, yield the disparity value. Unfortunately it is usually not practicable or efficient enough to calculate the disparities for each pixel in the image due to various problems like movement uncertainty in smooth regions and occlusions/dis-occlusions. Our framework therefore uses a mixture of feature matching and optical flow information to retrieve the disparities of a selected set of pixels.

The first set of disparities is retrieved by evaluating the optical flow information of image pixels on an equally spaced grid. The second set contains disparities calculated from matched features. Any sophisticated feature matching algorithm can be used here. After this step we have a coarse depth map for the current frame. Since the feature matching yields more precise measurements than the optical flow calculation, its disparities are more accurate and should therefore be considered with more weight. For this reason the disparities on the grid are currently set to the median of all surrounding results from both, optical flow and feature matching.

2.2. 3D Visualization

After getting the frame disparities, a depth histogram for each frame can be constructed. Having disparity values d_{p_i} for each pixel p_i in a frame i, this is done by incrementing the histogram bins $b^{(i)}_{\lfloor d_{p_i}+0.5 \rfloor}$ by 1 for each occurrence.

To ensure nice visualization properties the histograms is normalized with $b_j^{(i)} = b_j^{(i)}/d_{max}$ where $d_{max} = max\{b_j^{(i)}\}$ is the maximal depth histogram value in the video for $j = d_{min}^{(i)}, \ldots, d_{max}^{(i)}$ and $i = 1, \ldots, n$ where n is the number of frames.

A stack of all resulting histograms yields a 3D histogram. This representation now visualizes the additional information about the video concerning changes in depth over time and space and can be used to find undesired effects in the video such as hard or fast depth changes.

3. RESULTS

The computed histogram information can be modeled in multiple ways. If each bin is represented as a 3D cuboid, the perspective projection yields a good view over the whole video (Figure 2). With this view, depth transitions to following but also preceding frames are visible.

On the other hand the editor might still be interested in single frame disparities while looking at the visualization. A cut through the model at a given frame yields this information (Figure 3). Like this we achieve an augmented view of existing 2D histograms.

Not using a 3-dimensional representation of the model, such a full video disparity histogram is usually represented as a color-coded 2D histogram. Since the bins in the current implementation are already color-coded, a simple orthogonal projection leads to the same result without any additional computations (Figure 4). In this view, depth information can be evaluated without the perspective distortion in the direction of time.

4. CONCLUSIONS AND FUTURE WORK

Whereas the tool itself works well, the integration into the existing framework, where it takes its information from, still requires some work. Some tweaks in the visual appearance like colors and viewing angles or handling flexibility like scaling shall be incorporated.

Setting pixel disparities of the grid according to the optical flow and feature information around it results in a good approximation of the total frame depth. We can only improve accuracy by either increasing the number of grid lines or finding more features. Unfortunately this approach still suffers from problems like (dis-)occlusions or large smooth regions.

With the knowledge about undesired effects in the video and their representation in the 3D histogram, an automatic recognition of these could be possible. Further research in this area could lead to a better understanding and further visualization options.

Since the result of the project was initially designed for stereographers in 3D film production, a plug-in for existing stereo auditing software is a main goal for the future.

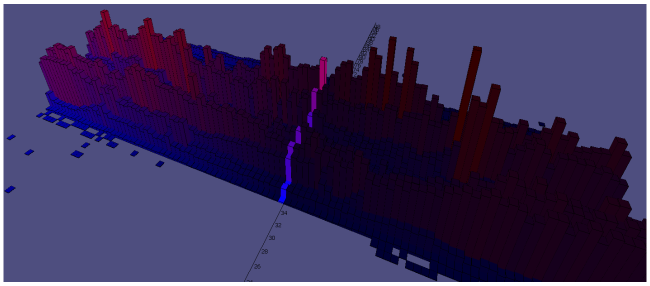

Fig. 2. Depth histogram overview.

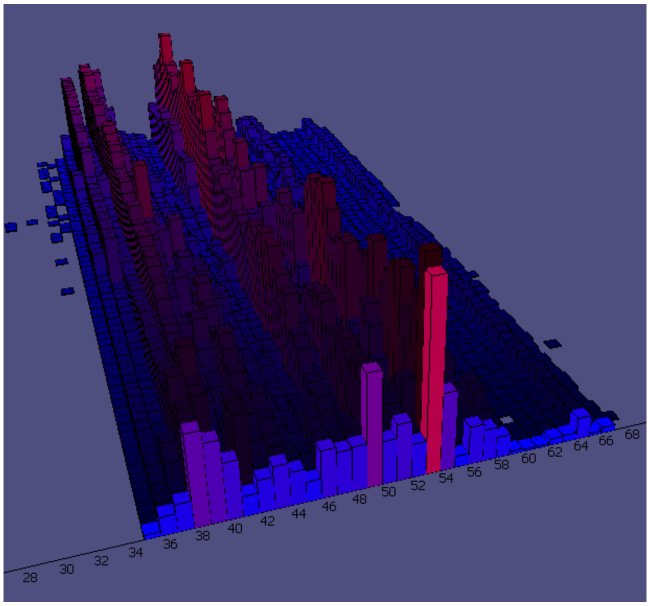

Fig. 3. Perspective view of a cut through the histogram.

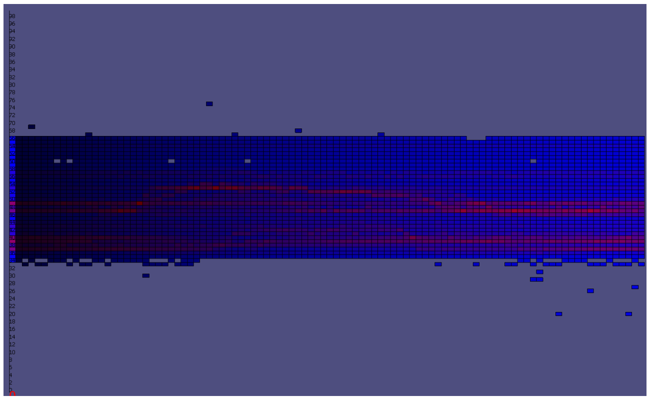

Fig. 4. Orthogonal histogram projection.