

Detection of Stereo Window Violation in 3D Movies

Yasin Nazzar, Jonathan Bouchard, James J. Clark*
Centre for Intelligent Machines, McGill University

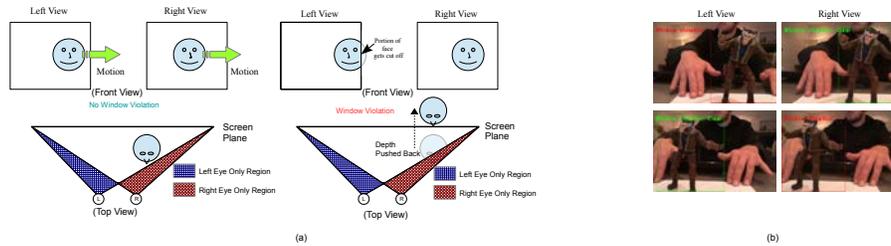


Figure 1: (a) Schematic view of Stereo Window Violation (b) Detection of Stereo Window Violation by proposed method

1 Introduction

The aim of 3D movies is to create a more immersive experience for the audience. However, if these movies are not created properly it can end up being a strenuous experience for the audience rather than an enjoyable one. The discomfort can be caused by a variety of artifacts that are produced when creating a 3D movie. Being a relatively new paradigm, the causes and existence of these artifacts may not be fully understood by cinematographers or post production crew that perform 2D to 3D conversion. We can analyse these 3D movies and its artifacts by employing techniques from stereo computer vision. In this work we identify and detect one such artifact called stereo window violation [Mendiburu 2009]. The screen of a 3D movie can be thought of as a window. A stereo window violation occurs when an object in front of the screen crosses the edge of this window and stays there. A portion of the details around the edge is present to one eye only. The effect of such a violation causes a sudden jump in depth of the object in front of the screen to the screen plane. Additionally, it becomes more difficult to fuse such an image. Both effects can cause a great deal of discomfort to the viewers in the form of headaches and eye strain.

2 Our Approach

Our first step to flagging stereo window violation is to determine the depth of a scene by finding correspondences between the stereo image pairs. The offset of corresponding points from one image to another is referred to as binocular disparity. The disparity is inversely proportional to depth. The amount and direction of the disparity affects our depth perception of a scene. For disparity estimation, we implemented the method in [Hosni et al. 2011], mainly because it does well around boundaries of an object. In a 3D movie the point of convergence (of the cameras and not the viewers eyes) has zero disparity and defines the screen plane. When an object in the left image shifts to the left in the right image it creates an out-of-screen perception. This is used to threshold the disparity map

and removes pixels that are behind the screen plane. A thresholded disparity map is not sufficient as we cannot just track each pixel that approaches the screen edge. We need to locate objects within the scene. Real objects tend to be piecewise smooth with true discontinuities at the boundary of the objects. So by clustering pixels based on their disparity we would be able to segment objects. The clustering was done using the mean-shift algorithm. The features used for the mean shift clustering were image intensity values, the thresholded disparity values and a measure of the amount of blur for each pixel. The motivation for using the blur information is that our brain tries less to fuse images of objects that are out of focus and these objects should not flag a window violation problem. We can then track the objects outside the screen by drawing bounding boxes around their segments. Objects small enough to be not noticed are not tracked. When objects, that are not blurry, cross the edges a stereo window violation is flagged. This tool can then be used by film crew on set to analyze their stereo quality as well as post production companies doing 2D to 3D conversions. The film crew can re-adjust their shots and post production can employ corrective measures that yields a 3D movie free of window violation artifacts leading to a more comfortable viewing.

3 Implementation and Future Work

We filmed a short video sequence containing stereo window violations with a stereo camera setup. OpenCV's stereo calibration method was used to rectify the images. The video sequence consisted of 862 frames at a resolution of 511×366 (with downscaling and cropping from rectification). The proposed method was quite successful in detecting the window violations, with 131 frames positively identified as window violations, 5 frames with false detections and 1 frame with missed detection. There was also 1 single frame with a false warning for approaching a violation. Our ongoing work is focused on speeding up the implementation using GPU based computing.

References

- HOSNI, A., BLEYER, M., RHEMANN, C., GELAUTZ, M., AND ROTHER, C. 2011. Real-time local stereo matching using guided image filtering. In *Multimedia and Expo (ICME), 2011 IEEE International Conference on*, IEEE, 1–6.
- MENDIBURU, B. 2012. *3D movie making: stereoscopic digital cinema from script to screen*. CRC Press.

*e-mail: {ynazzar, jonathan, clark}@cim.mcgill.ca