Rochester Institute of Technology

Optical Flow Estimation and Cube Projections
Facebook Surround 360

Nick Franco

PHPS.401.01 - PS Capstone I

Prof. Michael Peres

20 October 2016

Optical Flow and Cube mapping are the two best ways to keep file sizes and processing to their optimal levels when working with 360 3D. While traditional 360 images use an equirectangular format, we face the issue that has always plagued normal video, file size and speed. Cube mapping is great for a normal 360 video and 3D 360 by reducing size by 25%. Optical flow is the extra step that allows for stereo from a 360 video and more frames for a seamless playback and stitching for motion. This makes the need for 120 fps seamless video to the eye to only need standard 60 or 30 fps playback.

The Surround 360 Facebook project is very large containing aspects of software hardware image testing and psychological physics. For my research my main problem is pertaining to software and more specifically to how depth is made in this video capture device. To start I will go over why I chose this as my project and what I plan to do. 360 is becoming one of the newest ways to view content. For years we have been tied to 2D images because of how we view our content. With the advent of virtual reality we now have a new way to view images. No longer are we tied to photographic paper and computer screens. But 3D causes some new challenges creative aspects become lost in the huge amount of data available in this capture, the photographer becomes a part of his image, and the complexity of our capture devices becomes immense. The hardware in the surround 360 is very impressive, whats more impressive is how we store this media. At 17gb/s we need to make use of the newest technologies available today connecting through VR is an amazing, immersive experience, and we want to do our part to make sure more people have a chance to experience it in the future. Like all new technologies the beginning, first adopters must pay a hefty price, but these are the people who pave the way for future iterations of the devices. A great example of this is the iPad and how the first generation of users made use of the technology in ways never thought of. This forced apple to add features to better suit these uses and made the iPad a truly universal device. Being part of a new technology is extremely expensive and this is why our budget has grown to the large amount of $60,000.00 in the meantime however software is very easy to obtain from Facebook since it was made open source. This is where I do my part of make this project impact the future. By using the resources I have I started to test how this device works. Using a dslr and two lenses I was able to replicate the images the surround uses and then tweak the software to use a DSLR as input. While capture goes great and the stitching works as planned I still run into one fundamental problem. The file sizes are massive and the stitching requires hours of manual selection. While an easy fix for this has been compression and blurring its not the best approach for 360. Zooming and retaining data are what we must have in a cinematic professional grade device. This is where my work with optical flow happens.

Optical flow is a mathematically rigorous approach that produces superior results. What this does is compute left-right eye stereo disparity. We leverage this ability to generate seamless stereoscopic 360 panoramas. Optical flow estimation is a very advanced artificial intelligence that can predict motion. A very simple form of optical flow is computing difference between two images of a similar object. But before we get a difference we also need object detection in place. In the software I am able to find things in the frame such as people, the tripod/pole and trees/poles. By using this object detection and finally the optical flow not only can we make virtual cameras to make depth, but we can do image corrections live. This fully solves our issues of seamless stitching but then there is an added bonus. Making stereo can be done with optical flow. Since both eyes have there own 8k display the difference between eyes would not be as noticeable as an anaglyph. This means we can take the virtual cameras from our data and make a 3D file from the monoscopic video itself. But this only tackles some of the file size issue. While using processing power to relieve the amount of data needed at capture is great there is another area where our processors can aid in the files. This is at the upload to Facebook itself where we take the equirectangular frames and project from a cube. This reduces each frame by 25% of its pixels and is lossless. But how do we delete data and still stay lossless sounds impossible. To start a equirectangular image is extremely warped so this is just wasted pixel space to begin with. By using a cube projection we simply unwarp and then make a cube then rewarp. This means we don't save a file with warped pixels we just let the computer warp in real time. Cube warping unlike 3D 360 is not new, video games have been using this for a long time because of its simplicity. Cube warping is very low on processor resources and it works very well for our application. Back in 1986 cube mapping was first proposed by Ned Greene but it could not be used due to hardware limitations. In 1999 the Nvidia GeForce 256 was the first capable device and even advertised this ability.

Currently my work with 360 is all very new. Every day software is updated and I have already left my mark on the git. Some of my work is already implemented and my feedback is helping make the surround and 360 viewing better every day. Because this is such a new area history is hard to come by for 3D 360 and the advanced software

that is a part of it. Normal 360 and large panoramic images have a history but they have very little to do with 3D 360 due to the fact this is very computationally dependent and closer to video cgi. Currently my problems I have been facing include the masking of the tripod. Fish eye lenses use the same sensors as the wide angles on the surround and the fish eyes have much less resolution in favor of their larger FOV. Because of this we try and keep all content to the wide angles since the software has to work extremely hard to correct on the fish eye plane.

The goals I want to accomplish with my work is to help with the final release of the software and to keep contributing to the git Surround 360. Currently the surround 360 is open source and a global effort to create the best experience for 360. Being a part of that is incredibly meaningful, after building the entire device the amount we can help the community outside of RIT is incredible and what I want to accomplish is a set of standards for 360 to test quality and provide the best 3D 360 scenes from the software stitches.

Mentor Choices
Nitin Sampat
J A Stephen Viggiano PhD
Nanette Salvaggio

Works Cited

"Cube Mapping." Wikipedia. Wikimedia Foundation, n.d. Web. 20 Oct. 2016. <https://en.wikipedia.org/wiki/Cube_mapping>.

Facebook. "Facebook/Surround360." GitHub. N.p., 19 Oct. 2016. Web. 20 Oct. 2016. <https://github.com/facebook/Surround360>.

"Isometric Projection." Wikipedia. Wikimedia Foundation, n.d. Web. 20 Oct. 2016. <https://en.wikipedia.org/wiki/Isometric_projection>.

Lazyriver555. "Optical Flow Test: 360 Shuv It." YouTube. YouTube, 14 Jan. 2012. Web. 20 Oct. 2016. <https://www.youtube.com/watch?v=oT00aVKhshM>.

"Optical Flow." Wikipedia. Wikimedia Foundation, n.d. Web. 20 Oct. 2016. <https://en.wikipedia.org/wiki/Optical_flow>.

Stavens, David. An Introduction to OpenCV and Optical Flow (n.d.): n. pag. Web.

"Surround 360 Is Now Open Source." Facebook Code. N.p., n.d. Web. 20 Oct. 2016. <https://code.facebook.com/posts/265413023819735/surround-360-is-now-open-source/>.

"Under the Hood: Building 360 Video." Facebook Code. N.p., n.d. Web. 20 Oct. 2016. <https://code.facebook.com/posts/1638767863078802/under-the-hood-building-360-video/>.