Semi-Automatic Method for Creating Virtual Reality Environments

UNIVERSITY OF TARTU
Institute of Computer Science

Andres Traumann

Institute of Computer Science, University of Tartu, Estonia
andres.traumann@ut.ee

PROJECT WEBSITE: https://andrestraumann01.wixsite.com/vrenvironments

Study IT in .ee
sponsored by Skype

Introduction

Traditional virtual reality environments are based on rendering all the surroundings using 3D models of objects. This requires graphically modelling the whole environment, which is usually a long and tedious task. 360-degree cameras that have recently become commercially available might be able to change that situation, as image-based virtual reality requires no graphical modelling and therefore has the promise of being faster with the results resembling the real world more accurately. The downside of this approach is the requirement to create a separate image for each possible location in the scene that is being captured, thus resulting in a large number of images that need to be labelled, so that navigation between images would become possible. This project investigates possible methods to create such maps semi-automatically and provides an interface to display the virtual reality environments so created.

360-degree Imaging

In this project the Samsung Gear 360 camera (Figure 1) was used for capturing 360-degree content. The camera has two ultrawide lenses that both have a view angle of 190 degrees, thereby covering the whole 360-degree view together.

Fig. 1: Samsung Gear 360 camera seen from two different view points

The output of this camera is shown in Figure 2. It consists of two spherical images that need to be stitched together before they can be used in practice.

Fig. 2: Output of the camera: two spherical images

360-degree images need to be handled slightly differently compared to images from ordinary cameras. Most conventional cameras output images that more or less correspond to the way we see the world. The view angle is at most 120 degrees in most cases, so an image looks similar to what we would see with our own eyes. 360-degree images, however, show the world differently: all the surroundings are visible on the same image at once. For a human observer that kind of situation is hard to comprehend and it would be much more meaningful to show only a subset of that image corresponding to some narrow view angle.

Stitching

The spherical images need to be stitched together before they can be used for virtual reality. This can be achieved by creating correspondences between the overlapping parts of the spherical images, and generating a new final image based on those correspondences. The result is called the equirectangular image.

Fig. 3: Equirectangular image, result of the stitching process

Rendering

After the image has been stitched together, the equirectangular image can be used for generating views from it, this process is illustrated on Figure 4. An equirectangular image is shown along with two images generated from it given two slightly different view angles. Notice, how the equirectangular image itself is not suitable for image display: some parts of the image are distorted, wheareas some straight lines are shown very curved on that image. The generated images, on the other hand, are much more natural to look at, as they correspond to a view angle that resembles the view angle of human vision.

Fig. 4: Two different views generated from the same equirectangular image

Generating these views makes use of the standard computer graphics pipeline. The equirectangular image is used as a texture and mapped onto a 3-dimensional sphere. The camera is placed inside that sphere and the whole scene is rendered. This enables to generate different views by simply changing the view angle, the view angle determines, which part of the sphere becomes visible, and this creates an impression of looking around in that scene.

Map Generation

Fig. 5: Remote controlled car with a 360-degree camera used for capturing video data for the virtual reality system.

In this work, the 360-degree image data was captured using a small remote controlled car that had the Samsung Gear 360 camera mounted on top of it, shown in Figure 5. The car was used to capture image sequences of roughly parallel movement pattern as visualised in Figure 6.

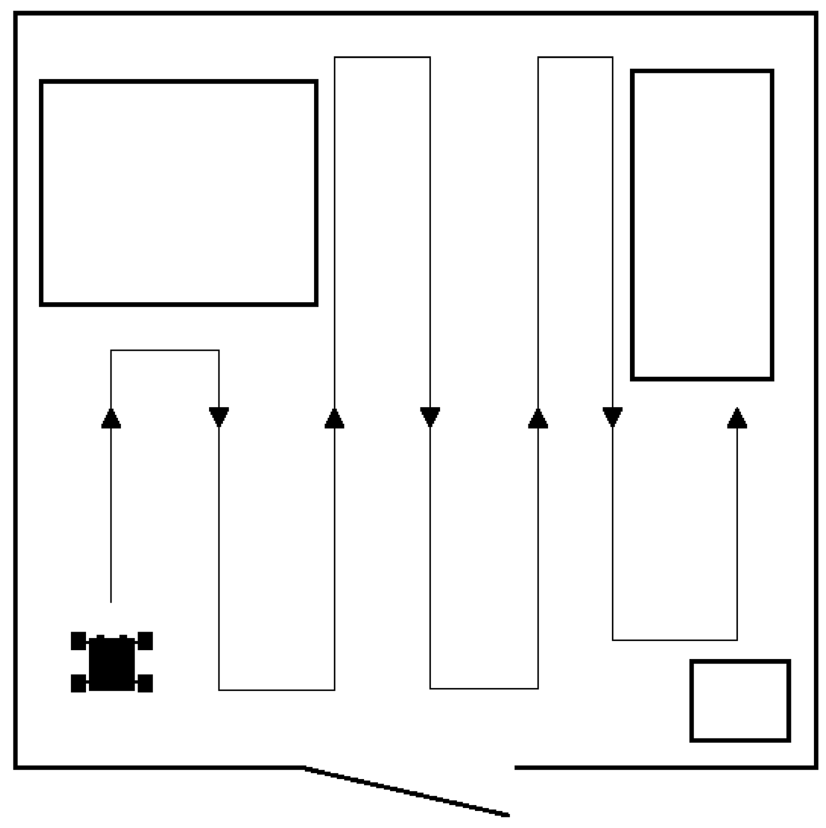

Fig. 6: Equirectangular image, result of the stitching process

Extracted image sequences are mapped using a dynamic time warping based approach that places the sequences side by side. The image similarity is used to determine the cost of mapping the sequences to each other. The similarity is computed by finding the pixel-wise sum of differences between the images, while this approach might sound naive, it performed very well in determining the closest neighbours, and as a result it could be used as a reasonably good similarity metric.

The map generation algorithm outputs an implicit map that can be used for navigating the scene. At each point on that map the user can look around freely and according to the connections described by the map, the user can also move around between the different points of the map.

Conclusion

This project investigated methods for automatically creating maps for image-based virtual reality environments. A graphical user interface for displaying such environments was also created. Some existing methods were studied and tested, along with a novel method designed based on the specifics of our image capturing system.