City Documentation: Creation and Visualization of High Resolution Panoramic Image Mosaics

Mario SORMANN & Gerald SCHRÖCKER & Andreas KLAUS & Konrad KARNER

DI Mario SORMANN & DI Gerald SCHRÖCKER & DI Andreas KLAUS & DI Dr. Konrad KARNER: VRVis Research Center for Virtual Reality and Visualization, Virtual Habitat, Inffeldgasse 16/2, 8010 Graz, sormann@vrvis.at

1 ABSTRACT

We present a novel system for 2D city documentation utilizing panoramic image mosaics. Current panoramic stitching methods often produce poor results and available visualization tools are not able to deal with high resolution panoramic image mosaics. For the construction of a full view high resolution panorama, we introduce a rotational mosaic representation, which includes the rotation parameters, the focal length and the principal point with each input image. A local registration algorithm is developed to quickly align two images exclusively from the image data. In order to reduce accumulated errors a global registration algorithm is performed to simultaneously minimize the misregistration between all overlapping pairs of images. By combining both registration techniques, we improve the quality of full view panoramic image mosaics. Finally we introduce a weighted blending algorithm to generate spherical or cylindrical panoramas.

Furthermore we present an efficient visualization tool to display our produced high resolution panoramas. We developed a method to warp panoramas by utilizing the OpenGL graphic system to overcome the restriction of common panorama viewers, which are only allowing the visualization of small panoramas. The panoramic image is digitally warped on-the-fly to simulate camera panning and zooming. Walking through the virtual environment is currently accomplished by jumping to different panoramic points. The presented documentation system provides the creation as well as the efficient visualization of high resolution panoramic image mosaics to document important areas in urban environments.

2 INTRODUCTION

The automatic construction as well as the visualization of high resolution panoramic image mosaics is an active area of research and can be used in several applications described in Irani [5] et al.. For applications such as virtual travel or architectural walkthrough, it is desirable to have a complete panorama, i. e., mosaics which cover the whole viewing sphere or cylinder.

The first goal of our city documentation system is to introduce an efficient method to construct high resolution panoramic image mosaics, for instance image mosaics with an resolution of 14000x3000 pixels, to obtain as much details from the panorama as possible. In the construction part we additionally avoid manual interaction, thus our stitching method is able to produce high resolution panoramic image mosaics fully automatic. The second purpose is to visualize this large panoramas at interactive framerates. Current available visualization tools, such as QuickTime VR (http://www.apple.com/de/quicktime/) are only suitable for the visualization of small panoramas up to 3000x500 pixels. For high resolution panoramas the rendering performance of such tools is not acceptable. To overcome this restriction, we developed a visualization tool, which can handle these high resolution panoramic image mosaics.

3 RELATED WORK

There are two classes of studies related to our method: creation of panoramic image mosaics and visualization of panoramic image mosaics. Many advanced creation methods have been developed in recent years. A very popular approach is the panoramic image mosaic concept proposed by Shum and Szeliski [6] and its variants. In these approaches a rotation matrix is associated with each input image and to reduce registration errors a local as well as a global registration step is initiated. However, there also exist some commercial and non-commercial products such as the QTVR Authoring Studio (http://www.apple.com/), 3D Vista Studio (www.3dvista.com) or the non-commercial PanoTools by Helmut Dersch [2]. Unfortunately most of these products are limited to cylindrical panoramas obtained with cameras rotating on levelled tripods equipped with a panoramic head to minimize motion parallax. This has limited the use of panoramic building to researchers and professional photographers who can effort such specialized equipment.

The other class of methods represent the visualization of panoramic image mosaics. One of the most important methods in this research field is the image based rendering approach introduced by Chen [1]. This technique has been used in the commercial product QuickTime VR, a virtual reality extension to Apple Computer's QuickTime digital multimedia framework. The system provides a technique which uses 360-degree cylindrical panoramic images to compose a virtual environment. But this and other commercial products are generally restricted to the visualization of small panoramas with a resolution of 3000x500 pixels and are therefore not suitable to deal with our high resolution panoramic image mosaics.

4 CREATION OF HIGH RESOLUTION PANORAMIC IMAGE MOSAICS

Our stitching step is divided into several steps, each dealing with one specific part of the algorithm. At the end we demonstrate the results of applying our method on different types of data-sets.

4.1 Overview

Image mosaics are collections of overlapping images that are transformed in order to result in a complete image of a wide angle scene. The transformations can be seen as simple relations between coordinate systems. By applying the appropriate transformations and merging the overlapping regions of an image, it is possible to construct a single image covering the entire visible area of the scene. Normally, those coordinate transformations are not known beforehand, unless the camera parameters are tracked with high precision. Therefore the central problem of panoramic image mosaicing is to compute the transformation parameters exclusively from the image data. Such a problem is called image registration.

However, we restricted our method to a single viewpoint constraint approach, which means that all images are taken from the same viewpoint without zooming. Consequently the eight transformation parameters are reduced to three angles (ω, θ, ϕ).

4.1.1 Cylindrical and Spherical Panoramas

Cylindrical panoramas are commonly used because of their ease of construction. To build a cylindrical panorama a sequence of images is projected onto a cylinder. The x-axis is proportional to a full rotation and the y-axis can be interpreted as the elevation. One disadvantage of cylindrical panoramas is that they do not cover the zenith (straight up) or nadir (straight down). In contrast to cylindrical panoramas, spherical panoramas cover the entire sphere, i.e. 360° horizontally and 180° vertically. One can see straight up and straight down. Consequently both axis of the panorama image are interpreted as angles. To build a cylindrical panorama, we map world coordinates p = (X, Y, Z) to 2D cylindrical screen coordinates (θ, ν) using

$$\theta = \tan^{-1}(X / Z), \nu = Y / \sqrt{X^2 + Z^2}$$

where θ is the panning angle and ν is the scanline. Similarly we map world coordinates into 2D spherical coordinates (θ, ϕ) using

$$\theta = \tan^{-1}(X / Z), \phi = \tan^{-1}(Y / \sqrt{X^2 + Z^2})$$

4.2 Work Flow

Our workflow, illustrated in Figure 1, comprises four consecutive steps where the first step consists of calibrating the camera and correcting the lens distortions. Detailed information about geometric camera calibration can be found in Heikkilä [4]. The second task is known as local registration of image sequences where a complete initial panoramic mosaic is assembled sequentially by computing local registration parameters between each image pair. In the third step we need to deal with accumulated misregistration errors, which are always present in any large image mosaic. For example, if we register a sequence of images using pairwise alignments, there is usually a gap between the last image and the first one. Therefore the third step includes a global optimization technique to find the optimal overall registration. The last step activates the stitching algorithm to stitch all image pairs together and to create a coherent panoramic image.

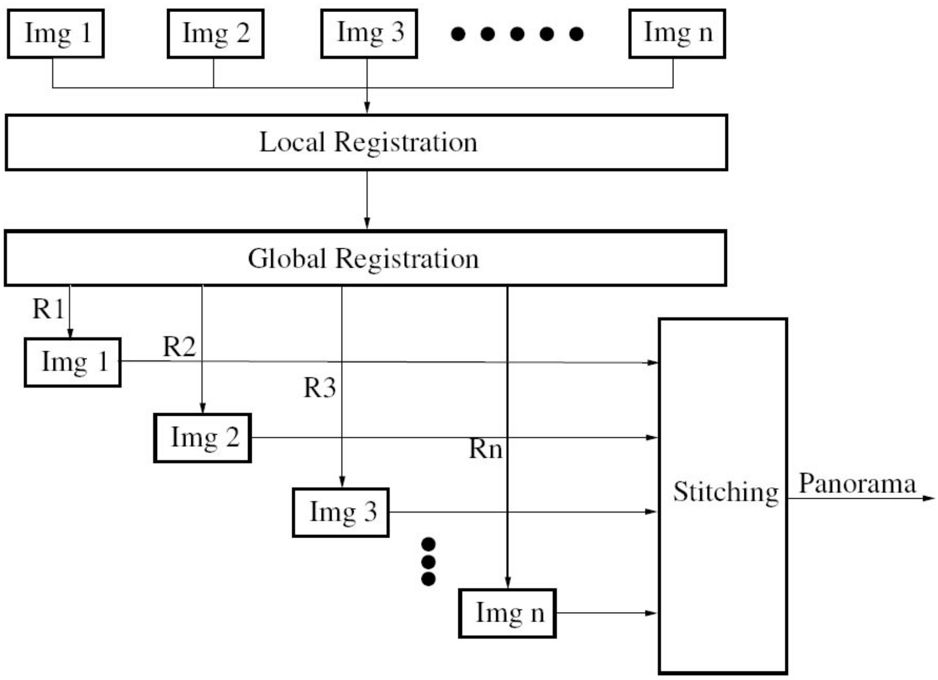

Figure 1: Outline of the major components of our panoramic image mosaicing approach.

4.3 Local Registration

The local registration consists of considering each pair of image and finding the best transformation that maps the first image on the second one. A search for the best match for all possible image parameters can be computationally extremely expensive. Therefore we implemented a feature based method relying on the accurate detection of image features. As a consequence of our wide angle lens, it is necessary to resample all images in a preprocessing step to compensate for any lens distortions. The lens distortions and intrinsic parameters are determined with an automatic calibration process described by Heikkilä [4]. Our feature based matching algorithm can be split into three major parts: extract the feature points, find the initial registration and iterative refine the initial solution.

4.3.1 Feature Point Extraction

To locate common points of interests in the images we combine a generalization of a gray image Harris corner detector introduced by Harris and Stephens [3] with a matching algorithm to detect accurately feature points.

4.3.2 Initial Registration

Considering one image pair, it is possible to match the features from the first image to some of the features of the second image. From those point pairs, called homologues points, it is possible to derive the transformation that map all points of the first image into the second image. The approach for our initial registration is based on the following steps:

- Select all feature points in both images.
- Set an approximate range for all rotation angles (ω, θ, ϕ).
- Compute the sum of point to point distances of all possible rotation angles.

4. Keep that rotation angles (ω, θ, ϕ) which provide the minimum sum of distances.

So far we obtain from each image pair the initial 3D rotation parameters (ω, θ, ϕ).

4.3.3 Iterative Refinement

For a more precise registration of the image pairs it is necessary to initiate a simple iterative refinement step. The refinement is based on the fact that the initial rotation angles, obtained from the second step, are approximately correct. Thus we select the distributed feature points in the overlapping area and iteratively refine the initial rotation angles utilizing the sum of point to point distances.

4.4 Global Registration

Even with the best algorithms for recovering rotations and eliminating lens distortions, when a complete panoramic sequence is stitched together, there will be either a gap or an overlap, due to accumulation errors in the rotation estimations. Thus it is essential to reduce accumulated errors by simultaneously minimizing the misregistration between all overlapping pairs of images. A common way to solve such a problem is the following:

- Convert the misregistration into a gap angle θ_g.
- Distribute the gap angle θ_g over all image pairs.
- Compute the optimal rotation angles using a gradient descent optimization.

4.5 Stitching

Once we have determined the absolute rotation matrix for every single source image, it is possible to create one panoramic image by utilizing one of the traditional projection methods. The most common mapping techniques are the spherical and the cylindrical mapping. The color associated with each pixel is computed by first converting the pixel address to a 3D ray, and then mapping this ray into each input image through our known transformations. Further we introduce a weighted blending function to obtain a smooth transition between each image pair. This blending algorithm has to weigh each pixel in the overlap region with a factor such that the edge pixels weight is zero. Hence we implemented a simple algorithm, which measures the distance of a given point to the closest edge in a given image. The pixel weight is then proportional to this distance to some power of n. The overall stitching algorithm for a spherical panorama is the following:

1. For each pixel (ω, θ) in the spherical map calculate the corresponding point p on a unit sphere.

- For each p determine its mapping into each image utilizing the computed transformation parameters.
- Render a composite blended panorama image from the single source images.

As additional improvement we integrated a kind of filtering algorithm into the stitching process to filter disturbing objects (pedestrians, biker) from the final high resolution panoramic image mosaic.

4.6 Results

Two different locations were chosen to demonstrate our creation approach: the main square in front of the city hall and the square at the entry side of the Kaiser Mausoleum. Note that all results illustrated in this report were obtained on a 1,6 GHz Intel processor with 512 MB main memory and GeForce4 with 64MB graphics memory.

Figure 2: Collections of 18 source images (4064x2704) taken in front of the city hall with an angular separation of 20°.

Figure 3: High resolution spherical panorama generated from illustrated source images covering the main square in front of the city hall. We obtain a geometric resolution of 14000x2704.

Figure 4: Spherical panorama generated from 18 input images covering the entry side of the Kaiser Mausoleum.

5 VISUALIZATION OF HIGH RESOLUTION PANORAMIC IMAGE MOSAICS

Our visualization approach is divided into several subsections where we will give an brief overview of the visualitzation process and present some results.

5.1 Overview

As mentioned before, our panoramic stitching algorithm produces very large panoramas with a resolution of 14000x3000 pixels. State of the art visualizations tools are not able to visualize such high resolution panoramic image mosaics and therefore to overcome this restriction our method warp panoramas with the use of the OpenGL graphic system.

There are many possible surfaces upon which perspective projections can be mapped. One possibility is a set of six planar projections in the form of a cube with the projection center in the middle. While this representation can be easily stored and accessed by a computer, it is difficult to achieve the precise camera position and orientation. Also the planar cubic mapping does not represent a uniform sampling, it is considerably oversampled at the edges and corners. It is complicated to avoid artefacts from discontinuities at the image borders. In contrast the most natural projection is a sphere centered around the viewpoint. One problem of the spherical projection, is the representation of the surface of the sphere in a form which is suitable for storage and fast access on a computer. This is particularly difficult because a uniform discrete sampling is desirable. A trade-off is a projection on the surface of a cylinder. One advantage of the cylinder is, that it can be easily unrolled into a simple planar map, making computer access easy. Most of the commercial stitching tools support this class of panoramic images. A drawback is the limited vertical field of view.

5.2 Virtual Camera

With the virtual camera all known rotations can be emulated. A full rotation about the vertical axis is possible as well as a limited rotation about the horizontal axis. A roll motion can also be simulated, but this rotation of the image is not reasonable, because such a motion is not common in conventional photography. By modifying the field of view of the artificial camera, a zooming effect can be achieved. But through this image magnifying no new details can be seen.

5.3 Geometry

Due to the curved projection surface large distortions are inescapable in panoramic images (see Figure 3 and 4). New views are generated by mapping the panoramic image onto a cylinder or sphere which are viewed through a central projection from the center. Thereby the distortions are corrected.

In other systems (e.g. QuickTime VR by Chen [1]) a custom image warping algorithm has been used for this task. However the goal of this work is to use the OpenGL graphic system and therefore a standard rendering pipeline. One can produce arbitrary image distortions by texturing a uniform polygonal mesh and transforming the vertices appropriately. To warp the panoramic image, a cylindrical or spherical surface is approximated with a triangular mesh and the synthetic camera is placed in the center as illustrated in Figure 5.

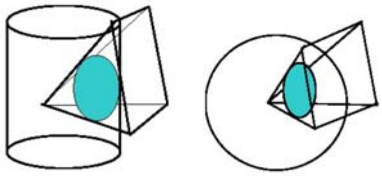

Figure 5: Camera positions in the approximated surface

5.4 Recording and Reproduction

By taking a point in 3D world coordinates (X, Y, Z) and mapping this point using a central projection onto the approximated surface, we obtain the transformed coordinates.

In order to warp the panorama for viewing, the image is projected from the approximated surface to a plane, which is normal to the optical axis and tangents at a defined point. Consequently a point on the surface is projected over a known transformation into a point on the plane. The process of warping can now be accomplished by an algorithmic operation, called warp operation. Detailed information about this algorithm can be found in Chen [1].

5.5 Texturing

The polygon mesh, whose shape was derived in the previous section, has to be textured with the panoramic image. Therefore we divide the large panoramic image mosaic into suitable parts and sent these parts as single textures to the OpenGL rendering pipeline.

6 RESULTS

Figure 6 shows a screenshot of the current implementation. To illustrate the quality of the high resolution panoramic image mosaic Figure 7 presents a magnified view of the city hall during the visualization. In Figure 8 we illustrate the underlying geometry of the textured cylinder overlapped to the warped image.

Figure 6: Visualization of a 14000x2704 high resolution panorama covering the main square in front of the city hall. We obtain a frame-rate of 70-100 frames/sec on a 1,6 GHz Intel processor with 512MB main memory and a GeForce4 with 64 MB graphics memory.

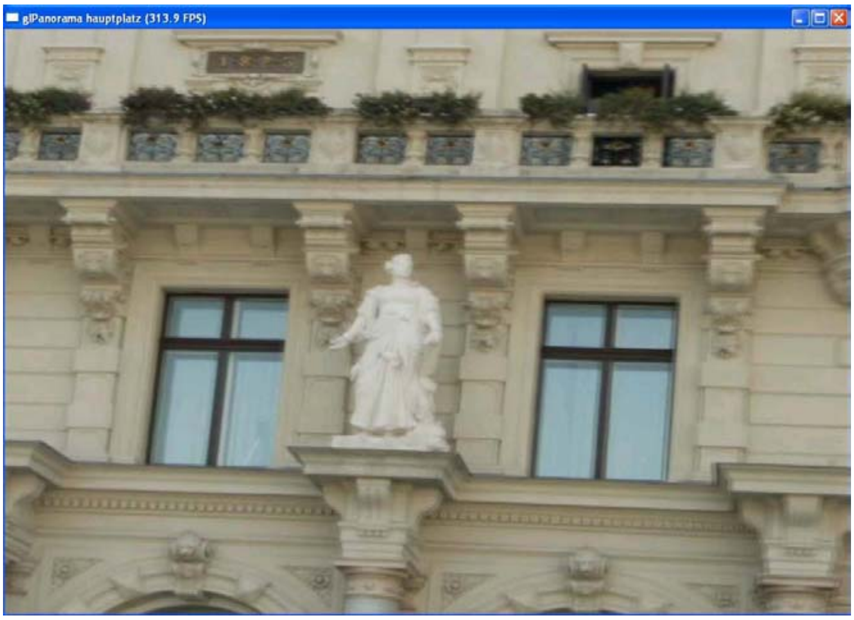

Figure 7: Magnified view of the main square in front of the city hall to illustrate the quality of our panoramic image mosaic.

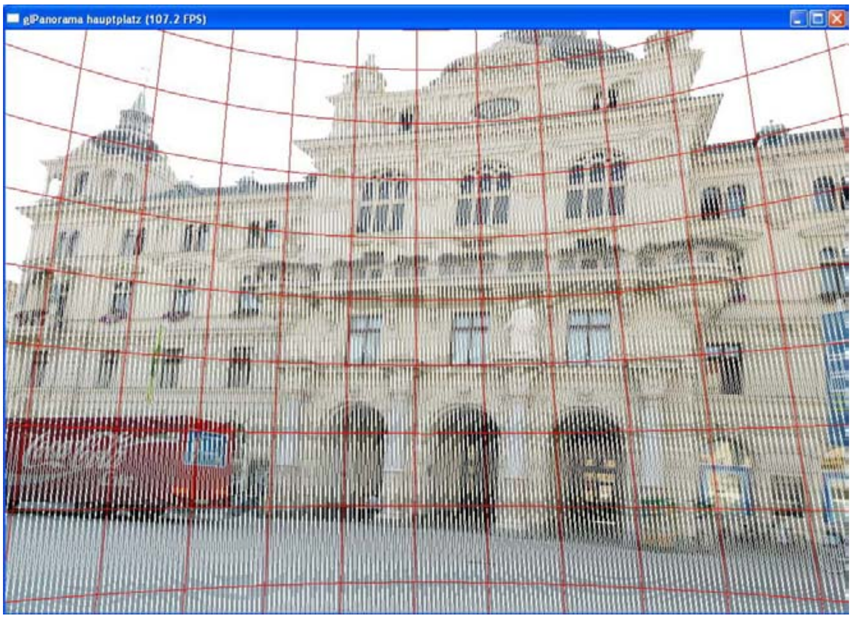

Figure 8: Underlying geometry of the approximated cylinder during the visualization process

7 CONCLUSION AND FUTURE WORK

We have presented a 2D city documentation system based on panoramic image mosaics. Moreover a stitching method was shown to create cylindrical as well as spherical high resolution panoramic image mosaics from image sequences fully automatic. Furthermore we illustrated some results of applying our method on different types of datasets. The second part of our framework demonstrates a visualization tool, which is able to visualize high resolution panoramic image mosaics by utilizing the OpenGL graphic system. Navigation through the environment is accomplished by jumping to different panoramic points.

One possible extension of the presented framework would be the creation and visualization of stereo panoramas to give the user a realistic sense of depth. Furthermore it would be useful to integrate metainformation about the presented environment into the final panoramic image mosaic.

8 ACKNOWLEDGEMENTS

Project Virtual Heart of Central Europe, Towers, Wells, and Rarities 3D On-line, has been co-funded by European Commission in Culture 2000 Framework Programme, Agreement No. 2003 - 1467/ 001 / 001 CLT CA12.

9 REFERENCES

[1] Chen S. E. and Williams L.: QuickTime VR – An image based approach to virtual environment navigation; *Proc. of Siggraph*, pages 29-38, August 2000.

[2] Dersch H.: Panorama Tools; *World Wide Web*, http://www.path.unimelb.edu.au/~dersch.

[3] Harris C.G. and Stephens M.: A combined corner and edge detector; *Proc. 4th Alvey Vision Conference*, pages 147-151, 1988.

[4] Heikkilä J.: Geometric camera calibration using circular control points; *Pattern Analysis and Mashine Intelligence*, pages 1066-1077, October 2000.

[5] Irani M. Anandan P. and Hsu S.: Mosaic based representations of video sequences and their applications; *Fifth International Conference on Computer Vision*, pages 605-611, June 1995.

[6] Shum H. and Szeliski R.: Creating full view panoramic image mosaics and environment maps; *Proc. of Siggraph*, pages 251-258, 1997.