RECONSTRUCTING VIRTUAL ROOMS FROM PANORAMIC IMAGES

Dirk Farin[a], Peter H. N. de With[a,b]

[a]Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands
[b]LogicaCMG, TSE-2, 5605 JB Eindhoven, The Netherlands
`d.s.farin@tue.nl` and `p.h.n.de.with@tue.nl`

The capturing of panoramic $360°$ *images has become a popular photographic technique. While a complete* $360°$ *view gives an impressive view of the environment, many people have difficulties to understand the spatial scene arrangement from this flat image. For this reason, we present and discuss various visualizations of panoramic images. Furthermore, we propose an alternative visualization for panoramic images, using a simple 3-D reconstruction of the recording environment. Specifically, we reconstruct the 3-D layout of rectangular rooms from a panoramic image. By projecting the panoramic image onto the walls of this virtual room, the scene can be visualized as a 3-D model. The model can be shown either as a scene overview, or it can be displayed with the virtual camera at the position of the original camera, providing realistic views from within the room.*

1. INTRODUCTION

The capturing and display of panoramic $360°$ images has become a standard technique which is used in a variety of applications like the presentation of hotel rooms in the internet for advertisement. The most frequently-used model for storing panoramic images is to project the surrounding environment of a fixed camera location onto a cylindrical surface. The are two commonly used ways to display panoramic images. The first approach shows the unrolled surface of the cylinder in a very wide image. Since this image shows all directions around the camera at the same time, the panoramic image itself can be very confusing to the viewer. Another popular form of presentation is through an interactive viewer. These *panoramic image browsers* (PIB) show a rectified sub-view of the scene in which the user actively controls the view of a virtual camera. The disadvantage of this representation is that it is not possible to get a fast overview of the scene, and it is not possible to see the complete environment on a static medium like a paper copy.

In this paper, we propose a new visualization technique for panoramic images. We concentrate on panoramic images that were recorded inside rectangular rooms,

which is an important special case that covers many application areas like hotel room advertising, or recording of group-meetings. Our visualization is based on an algorithm to reconstruct the 3-D layout of the rectangular rooms from a panoramic image. Once the geometry of the room is known, a 3-D model of the room walls can be synthesized and the wall textures can be added using the image data from the panoramic image. The proposed representation provides a flexible way to visualize the scene. On one hand, the virtual camera can be placed outside of the room, such that the viewer gets an overview of the whole scene appearance and room layout. On the other hand, the virtual camera can also be placed at the position of the original camera. Interactively rotating this virtual camera provides views that equal the output of the PIB technique.

The room reconstruction requires a minimum of user assistance. The user only has to indicate the position of the four room corners in the panoramic image. The reconstruction algorithm converts the positions of the corners into the angle between these corners as observed from the camera position. Subsequently, the room shape and the camera position are determined from these angles, and the textured 3-D model is constructed automatically.

The following two sections introduce the cylindrical model for panoramic images and give an overview of visualization techniques. Section 4 discusses the reconstruction algorithm for the room geometry, enabling a new visualization technique that is described in Section 5.

2. PANORAMIC IMAGES

The most commonly used type of panoramic images are cylindrical panoramic images. The idea is to project the environment onto a vertically aligned cylinder surface with the camera at the cylinder center [4]. If we denote the image coordinates as (x, y) and the cylinder coordinates as (φ, h), we can transform from image coordinates to cylinder coordinates by (Fig. 1)

$$\tan \varphi = x/f \quad \text{and} \quad h = \frac{y}{\sqrt{f^2 + x^2}}, \tag{1}$$

where f is the focal length (the distance of the image plane to the optical center). From these equations, we see that the focal length f of the camera has to be known for the generation of panoramic images.

A technique to generate panoramic images is to take a sequence of images i while rotating the camera around its vertical axis. Because the images are recorded with some change of camera rotation angle, their position on the cylindrical surface is shifted by some amount φ_i. This shift can be determined easily
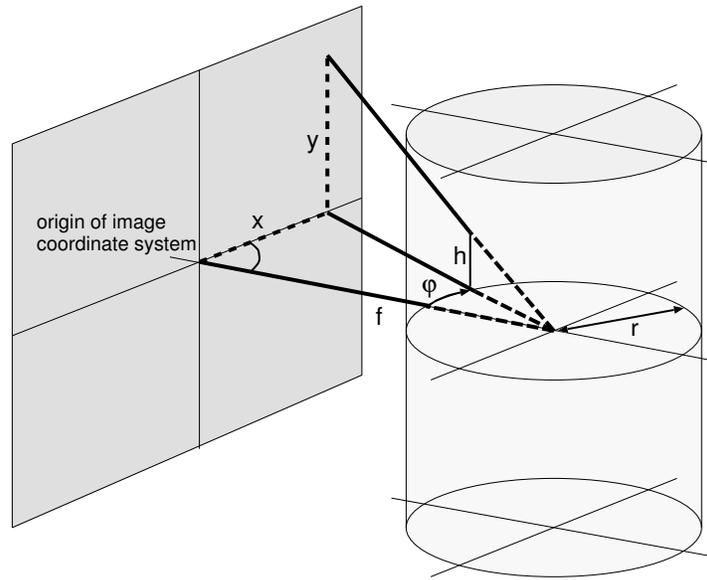
Figure 1: Projection of image coordinates onto cylindrical coordinates.

with a one-dimensional search to maximize the correlation between the overlapping image content.

3. VISUALIZATION OF PANORAMIC IMAGES

A panoramic image represents a complete $360°$ view of the environment around the camera. Hence, it is not an ordinary flat image and a variety of visualizations have been proposed. We briefly introduce the most important in the following, ending with our new proposal.

- **Unwrapped cylinder.** The most common display technique for cylindrical panoramas is to unwrap the cylindrical surface to a flat image. At first glance, this looks like an image with very wide field of view. However, there are two properties that distinguish this image in cylindrical coordinates from a normal, planar image. First, the image shows a complete $360°$ surrounding, which is an unusual experience, since the viewer looks in all directions around him at the same time. Second, straight lines are not preserved. Hence, geometrical concepts like parallel lines and vanishing points, that are important for an intuitive understanding of the scene, cannot be applied easily.

- **Panoramic image browser (PIB).** Another presentation technique is to provide an interactive viewer application which internally uses the cylindrical panorama representation, but which uses the inverse transform to syn-

thesize views for any arbitrary viewing direction.[1] The advantage of this technique is that the generated views look identical to real-world views. The disadvantage is that a static visualization of the complete environment is impossible (e.g., a printout on paper).

- **3-D cylinder projection.** Instead of displaying the panoramic image as a flat stripe, it can also be visualized as a 3-D cylinder model with the panoramic image as texture. This representation combines two advantages. If the virtual camera is placed at the center of the cylinder, the views look similar to the previous PIB approach. However, the camera can also be placed outside, giving a complete, static overview of the scene. While this overview gives some indication of the spatial arrangements in the scene, the intuitive perception of the cylindrical view is often misleading. For example, consider that a panoramic image is captured in a square room, but with the camera position not in the room center. In this case, not all of the walls span the same 90° in the panoramic image.

- **3-D room projection.** In this paper, we propose a new visualization technique which can be regarded an extension of the previous 3-D cylinder projection. Instead of projecting the surrounding on a cylinder surface, we propose to use a 3-D model reconstruction of the room walls with the layout of the real room. From the original camera position, the visualization is equal to the PIB technique. However, for a distant camera, the scene overview provides important information about the scene. First, the scene geometry indicates the aspect ratio of the wall and the camera position during recording. Second, the walls show texture on which straight lines are preserved.

4. RECONSTRUCTION OF RECTANGULAR ROOM GEOMETRY

In this section, we consider the problem of determining the wall sizes of a rectangular room from a cylindrical panoramic image that was captured in this room. Our desire is to provide a reconstruction algorithm that requires only a minimum of user assistance. The idea of our approach is that the user marks the room corners in the image (Fig. 4(a)) and the computer uses this information to calculate the room shape and the camera position.

Since the panoramic image is given in cylindrical coordinates, the horizontal distance between two corners in the panoramic image corresponds to the angle between these corners, measured from the camera position (Fig. 2(a)). Knowing

---

[1]This presentation has become popular with Apple's Quicktime VR standard.
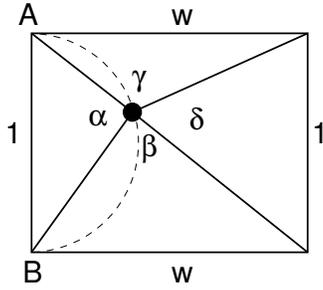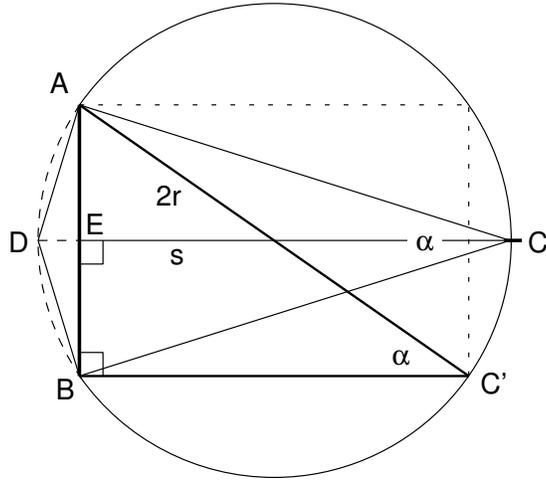
(a) The room geometry. (b) $r = 1/(2\sin\alpha)$ and $s = 1/(2\tan(\alpha/2))$

Figure 2:

these four angles (of which only three are independent, since they sum to 2π), we can determine the ratio of the room dimensions and the camera position. It is not possible to recover the absolute room size, but this absolute size is not required for the visualization, such that we can fix the size of one wall to unity. Note that it is required to know the camera position to generate the texture maps for the walls.

The reconstruction algorithm proceeds in two steps. First, it determines a circular arc on which the camera position must be located. Second, it carries out a binary search to determine the final camera position on this arc. The search in the second step is guided by the preknowledge that opposing walls have equal size.

4.1. Determining the circular arc of possible camera locations

Let us normalize the room size such that the left (and right) walls have unity length and the top (and bottom) walls have length w (Fig. 2(a)). The four walls are observed with the angles $\alpha, \beta, \gamma, \delta$. We first concentrate on the left wall AB of unity length that is observed with an angle α. It is well known (see [1], Book 3, Prop. 21) that all points C for which $\angle ACB = \alpha$ lie on a circular arc ACB (Fig. 2(b)). Obviously, the center of the circular arc must lie on the perpendicular bisection DC of the line AB, but the radius and the horizontal position are unknown. The radius can be obtained by considering the right triangle ABC'. From $AB = 1$ and $\angle AC'B = \alpha$, it follows that $\sin\alpha = 1/2r$. Furthermore, by considering the right triangle AEC, we obtain $\tan\alpha/2 = 1/2s$.
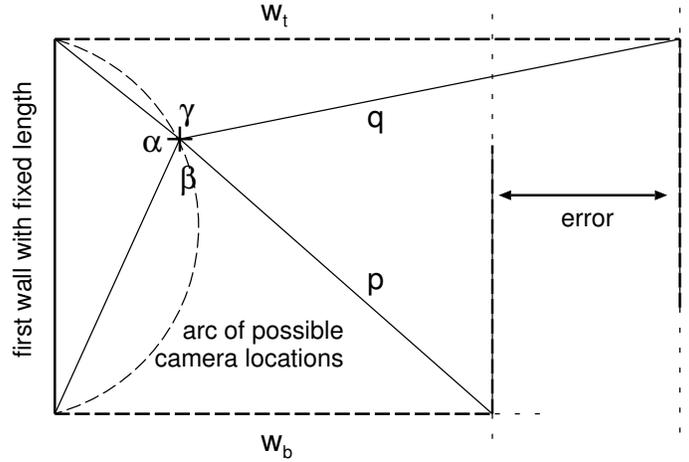
Figure 3: Searching for the correct camera position is limited to the indicated arc. A binary search is applied to find the position for which the error $w_t - w_b$ is zero.

Hence, the camera location is on a circular arc with radius $r = 1/2 \sin \alpha$, and the circle center has a distance of $s - r$ from the wall AB.

4.2. Searching for the camera location

We begin the construction with a wall of unity length, shown on the left side of Figure 3. Since we assumed that the room is rectangular, we know that the top and bottom wall must be perpendicular to this left wall, but we do not know their width yet. But we know that their width must be equal, because the wall on the right side is parallel to the left wall.

Let us choose an arbitrary camera position on the arc and consider this position. Then, the angles β and γ define the direction of rays p, q emanating from the camera position in the direction of the room corners. These rays intersect the top and the bottom walls in a distance w_b and w_t from the left wall, respectively. Because we know that the top and bottom wall should have equal length, w_t should equal w_b. However, if we have chosen the wrong camera position on the circular arc, this will not be true.

Notice that if we move the camera upwards along the arc, the top intersection point moves to the left (w_t decreases), while the bottom intersection point moves to the right (w_b increases). To find the camera position for which $w_t = w_b$, we can exploit this behaviour by applying a binary search for the correct camera position. If $w_t > w_b$, the camera position must be further to the top, while for $w_t < w_b$, the camera position must be lower.

For some camera position, the ray direction of p or q becomes horizontal. For these positions (and the more extreme positions), there is no intersection of the rays with the top or bottom wall. These critical camera positions can be used to determine an initial interval of camera positions for the binary search. Starting the search with this interval not only reduces the number of iterations needed for the binary search, but it also removes the requirement to handle the special case in which the rays p, q do not intersect the top or bottom walls.

5. VISUALIZATION

Once we have obtained the sizes of the room walls and the camera position in the room, we can create a virtual 3-D model of the room and generate textures for the room walls from the panoramic image by using the inverse of Eq. (1). This room model is then rendered using an OpenGL based viewer application. The user can control the rotation of the scene around the x and y axes, as well as control the distance between the virtual camera and the original camera position. If this distance is zero, the generated views are similar to PIB. For large distances, a scene overview is shown (Fig. 4(b)).

6. CONCLUSIONS

This paper presented an algorithm to reconstruct the geometry of rectangular rooms from panoramic images. The reconstruction is based on the position of the four room corners, that are marked by the user. Once the room geometry is known, a virtual model of the room is generated and the panoramic image is used to obtain texture maps for the walls. This 3-D model of the captured environment provides an intuitive visualization that gives a scene overview as well as virtual views from the original camera position.
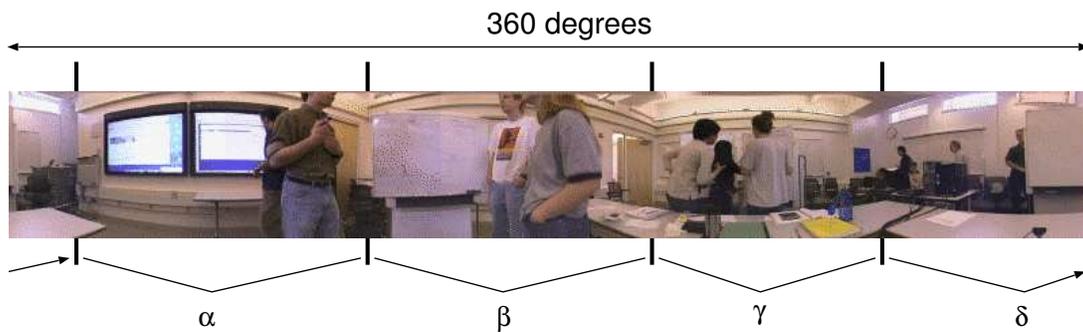
(a) panoramic input image with marked room corners
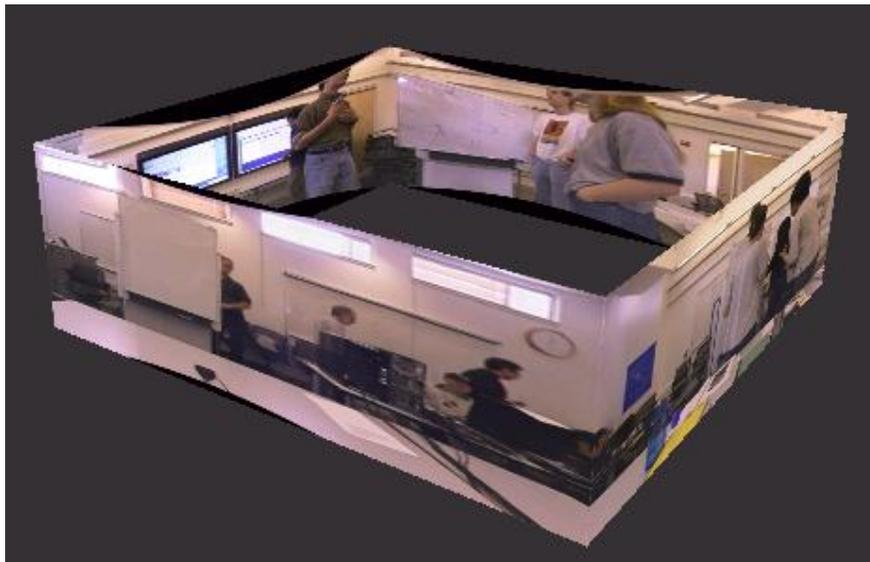
(b) view from above

Figure 4: Sample reconstruction of a rectangular room.

REFERENCES

[1] Euclid. *Elements, Book III: Theory of Circles*. 400 b.c.

[2] R. Pea, M. Mills, J. Rosen, K. Dauber, W. Effelsberg, and E. Hoffert. The diver project: Interactive digital video repurposing. *IEEE Multimedia*, 11(11):54–61, 2004.

[3] H.-Y. Shum, M. Han, and R. Szeliski. Interactive construction of 3D models from panoramic mosaics. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'98)*, pages 427–433, June 1998.

[4] R. Szeliski and H.-Y. Shum. Creating full view panoramic image mosaics and environment maps. In *SIGGRAPH*, pages 251–258. ACM Press/Addison-Wesley Publishing Co., 1997.