Arbitrary Viewpoint Rendering from Multiple Omnidirectional Images for Interactive Walkthroughs

Kaname Tomite, Kazumasa Yamazawa and Naokazu Yokoya
Graduate School of Information Science, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0101, Japan, {kaname-t, yamazawa, yokoya}@is.aist-nara.ac.jp

Abstract

Recently, the interactive walkthrough which enables us to look around a virtualized real world has been widely investigated. The walkthrough system can be applied to simulation, telepresence, and so on. This paper describes a method of interactive walkthrough which takes a real scene and enables us to look around in the virtualized scene. Our method is based on acquiring omnidirectional images at multiple points in the real world by using an omnidirectional camera. The method generates an omnidirectional image at an arbitrary location from captured omnidirectional images, and presents a user a converted view-dependent perspective image. In the experiment, we realized the interactive walkthrough using indoor omnidirectional images and confirmed the feasibility of the method.

1. Introduction

Recently, many researches have been made to reconstruct walkthroughs in virtualized real scenes such as city scenes, building rooms and natural environments [1][2][3][4][7][8][9]. These can be applied to virtual tour, driving simulation, disaster simulation, and so on.

Our objective is to develop a walkthrough system which enables a user to look around at arbitrary viewpoints widely using a limited number of input images. For this purpose, the approach is based on using an omnidirectional camera HyperOmni Vision [10]. Omnidirectional images at arbitrary viewpoints are computed from a limited number of omnidirectional images captured at discrete positions. This differs from previous works [1][8][9] that require dense image sequences captured by a moving omnidirectional camera.

This paper is structured as follows. First, we describe a method of generating omnidirectional images at arbitrary viewpoints from sparsely acquired omnidirectional images. Next, we present experiments to reconstruct walkthroughs in an indoor scene from four omnidirectional images. Finally, we summarize the present work.

2. Generating Omnidirectional Image at Arbitrary Viewpoint

2.1. Outline of the method

In our method, we take omnidirectional images at multiple positions with the omnidirectional camera HyperOmni Vision as illustrated in Fig.1. Interactive walkthrough is essentially accomplished by generating an omnidirectional image at an arbitrary position from three omnidirectional images that are acquired around the target position.

The flow of the method is shown in Fig.2. The present work is based on two assumptions: (1) Positions of HyperOmni Vision are known; (2) Partial correspondences among captured omnidirectional images are given. First, the positions of corresponding points in a novel omnidirectional image are computed. Second, the triangles are generated from the computed corresponding points. Third, the novel omnidirectional image is generated by blending corresponding triangular parts of captured omnidirectional images. Blending weights are determined based on the position of the target viewpoint in blending triangular parts. Finally, a common perspective image computed from the generated omnidirectional image is presented to the user.
2.2. Omnidirectional camera: HyperOmni Vision

We use HyperOmni Vision [10] (still version) as an omnidirectional camera in the present work. HyperOmni Vision is composed of a hyperboloidal mirror and a digital still camera as illustrated in Fig.3. The camera acquires an omnidirectional scene reflected by the hyperboloidal mirror. The hyperboloidal mirror is constructed of a hyperboloid of two sheets of revolution, which has two focal points \((O_u, O_c)\). The camera lens center is fixed at the focal point \(O_c\). Given a world coordinate \((X,Y,Z)\) and an image coordinate \((x,y)\) as shown in the Fig.3(b), the shape of hyperboloidal mirror and the two focal points are represented as follows.

\[
\text{Mirror Shape: } \frac{X^2 + Y^2}{a^2} - \frac{Z^2}{b^2} = -1,
\]

Inner focal point \(O_u\) : \((0,0,+c)\), \(1\)

Outer focal point \(O_c\) (Camera lens center): \((0,0,-c)\),

\[
\text{where } c = \sqrt{a^2 + b^2}.
\]

A ray going from the point \(P(X,Y,Z)\) in 3D toward the inner focal point \(O_u\) is reflected by the mirror and passes through the outer focal point \(O_c\) intersecting the image plane at the point \(p(x,y)\). Therefore, the projection of HyperOmni Vision is not common planar perspective, but satisfies the single viewpoint constraint. The relationship between \(P(X,Y,Z)\) and \(p(x,y)\) is given by:

\[
x = \frac{f(b^2 - c^2)X}{(b^2 + c^2)(Z - c) - 2bc\sqrt{X^2 + Y^2 + (Z - c)^2}},
\]

\[
y = \frac{f(b^2 - c^2)Y}{(b^2 + c^2)(Z - c) - 2bc\sqrt{X^2 + Y^2 + (Z - c)^2}}.
\]

By using the above equation, an omnidirectional input image can be converted to a common perspective image at the viewpoint \(O_u\) [6].

In this study, omnidirectional images are obtained at sparsely spaced positions on a single plane perpendicular to camera axes as input data for reconstructing walkthroughs. As mentioned earlier, the camera positions are assumed to be known.

2.3. Estimation of 3D position of corresponding points

We estimate a 3D coordinate \(T(X,Y,Z)\) from corresponding points \(p_i(x_i,y_i)\), \(p_j(x_j,y_j)\), and \(p_k(x_k,y_k)\) in three omnidirectional images for generating a novel omnidirectional view. The method is actually a variant of omnidirectional trinocular stereo.

Figure 4(a) shows a bird’s-eye view of straight lines \(l_1\), \(l_2\), and \(l_3\) connecting centers of projection and the points \(p_i(x_i,y_i)\), \(p_j(x_j,y_j)\), and \(p_k(x_k,y_k)\), respectively. The lines \(l_1\), \(l_2\), and \(l_3\) do not always cross at a single point because of quantization errors. We first determine a 3D point \(T'\) by minimizing the sum of distances between the point \(T'\) and the lines \(l_1\), \(l_2\), and \(l_3\). We then determine the points \(T_1\), \(T_2\), and \(T_3\) which are the nearest to the point \(T'\) on the lines \(l_1\), \(l_2\), and \(l_3\) (see Fig.4(b)). Finally, we compute the 3D point \(T(X,Y,Z)\) by solving the following equations as shown in Fig.4(c).

\[
P_i = w_1P_{i1} + w_2P_{i2} + w_3P_{i3},
\]

\[
w_1 + w_2 + w_3 = 1,
\]

\[
T = w_1T_1 + w_2T_2 + w_3T_3.
\]

Note that points \(P_i(X_i,Y_i)\), \(P_j(X_j,Y_j)\), and \(P_k(X_k,Y_k)\) represent omnidirectional camera positions and point \(P_i(X_i,Y_i)\) is a novel viewpoint. Parameters \(w_1\), \(w_2\), and \(w_3\) are weights to generate the same image as the captured image when the viewpoint \(P_i(X_i,Y_i)\) is at one of the original camera positions. The point \(p_i(x_i,y_i)\) in the omnidirectional image at the novel viewpoint can be obtained from the point \(T(X,Y,Z)\) using Eq. (2).
Fig. 4 Estimation of 3D point $T$ from points $p_1$, $p_2$, and $p_3$; (a) Relationship between omnidirectional camera positions and estimated point $T'$, (b) Enlarged version of (a), (c) 3D point $T$ estimated based on weights $w_1$, $w_2$, and $w_3$.

2.4. Generating omnidirectional image at arbitrary viewpoint

We generate triangulated patches in a novel viewpoint omnidirectional image based on two-dimensional coordinates of computed points $p_i$ using Delaunay’s triangulation [5]. It should be noted that these triangle patches are generated supposing that the novel viewpoint is located at the center of triangle constructed of viewpoints of input omnidirectional images. This reduces the discontinuity in view changes when the novel viewpoint moves continuously. The omnidirectional image at the novel viewpoint is generated by blending the parts of input omnidirectional images which correspond to the generated triangles. The blending is achieved by computing the weights $w_1$, $w_2$, and $w_3$ at each viewpoint. A viewpoint-dependent common perspective image can be easily generated from the blended omnidirectional image using Eq.(2) [6].

3. Experiments

We took four omnidirectional images (1024x768 pixels) using the HyperOmni Vision composed of a digital still camera (Nikon CoolPix 990). Omnidirectional camera positions are illustrated in Fig. 5. Figure 5(a) shows a room to be captured, and Figure 5(b) shows the enlarged version of the area containing camera positions. We divided the area into triangles (I, II, IV) and (I, III, IV). Next, we manually selected 120 points such as corners of objects in each omnidirectional image as corresponding points. In the experiment, the viewpoint is assumed to move from the point A through E.

Figure 6(a) shows a sampled sequence of generated omnidirectional images at the points from A to E. User’s view-dependent perspective images are also shown in Fig. 6, in which stereoscopic images are generated at the points B, C, and D. It took about 0.1 second to generate each view-dependent common perspective image from three captured omnidirectional images using a PC (CPU: Dual Pentium4 Xeon 1.7GHz, Mem: 2GB RIMM). Therefore, we can confirm that our method realized the interactive walkthrough.

However, we have encountered some problems in the experiments; for example, objects close to the viewpoint are not correctly reconstructed, and the discontinuous change is observed at the point C.

4. Conclusions

We have proposed a new method of interactive walkthrough which takes the omnidirectional images at discrete positions and enables us to look around at arbitrary viewpoints. Our method uses positions of HyperOmni Vision, omnidirectional images and partial correspondences among omnidirectional images as input data. The omnidirectional images at arbitrary viewpoints are generated by blending captured images. User’s view-dependent common perspective images computed from generated omnidirectional images are presented to the user. The proposed method is characterized by easy acquisition of images and capability of looking around the virtualized scene at arbitrary viewpoints. Automatic finding of correspondences among original omnidirectional images should be investigated in future work.
References


Fig.6 Experimental results with an indoor scene; (a) Generated omnidirectional images, (b) Generated common perspective images of front view, (c) Generated common perspective images of right side view.