Fourth International Conference on Virtual Systems and Multimedia, Gifu, Japan. November 1998

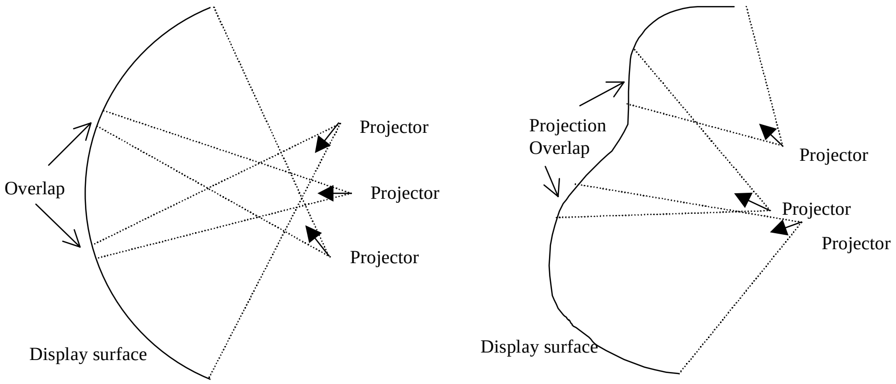
Seamless Projection Overlaps using Image Warping and Intensity Blending

Ramesh Raskar, Greg Welch, Henry Fuchs
{raskar,welch,fuchs}@cs.unc.edu
University of North Carolina,Chapel Hill, CB# 3175 Sitterson Hall, U of North Carolina, Chapel Hill NC 27599-3175 USA

Abstract. High-resolution Spatially Immersive Displays (SID) generally involve wide field of view (WFOV) image generation using multiple projectors. This paper describes a robust calibration and rendering method for projector based seamless displays using a video camera. It solves the basic problem of registering and blending overlap of two projections at a time. It is applicable even when the displays are not flat walls or projection axes are not orthogonal to the displays. Projectors' intrinsic or extrinsic parameters are not required.

1. Introduction

High-resolution Spatially Immersive Displays (SID) involve wide field of view (WFOV) image generation. Traditionally, this has been achieved by using non-overlapping projections (e.g. CAVE [Cruiz93]), by using side-by-side overlapping projections with precise electro-mechanical setup (e.g. flight simulators) or with a WFOV projector with expensive optics and computer hardware (e.g. domes [Bennett98]). The displays need to create an impression of a continuous and seamless image.



Typical seamless image generation using overlapping projections, with well defined display surface and projection configuration

Our method to create seamless images for complex display surface geometry and projector configuration. The generated images are geometrically registered and intensity blended in the overlap region.

Figure 1: Seamless Image Generation

We present here a robust calibration and rendering method for projector based seamless displays. We assume the projectors are used in a front-projection or rear-projection setup and projections of atmost two neighboring projectors can overlap. The method solves the basic problem of pairwise registering and blending overlap of projections. It is applicable even when the displays are not flat walls or projection axes are not orthogonal to the displays. This allows various geometrical configurations of the display surfaces such as cylindrical surround screens, hemispherical domes [Bennett98], [Jarvis97], CAVE like cubicles or other arbitrary shapes as encountered in the Office of the Future [Rask98a]. Projectors' intrinsic or extrinsic parameters are not required. Projectors with varying resolutions (number of pixels), zoom or density of pixels on the display surface can be used. However, we assume that the radiance at display surface due to any projector is equal. We further assume that the dynamic response curve i.e. function which maps (*value in the framebuffer/ maximum value in framebuffer*) to the ratio of corresponding radiant intensities, is known. For example, when gamma correction is used, this function is a simple mathematical formula. The display surfaces are assumed to be lambertian.

Since this method allows creating seamless displays using a set of projectors, it eliminates the need of precise electro-mechanical setup and expensive optics. Using more sophisticated rendering methods, wide field of view display environments can be created without using geometrically well-defined display surfaces, as described in [Rask98a], [Rask98b].

2. Calibration

Calibration stage uses a wide-field-of-view camera in two steps. First, we compute warping functions from desired image to each of the projected images to achieve geometric registration between them. During the second step, we compute the intensity weighting basis functions (IWBS) for projector pixels to achieve photometric normalization so that the perceived intensity across the seams is constant.

For displays with well-defined geometric shapes, a few projected pixels can be observed to parameterize the warping and blending functions using a least squares method. For example, consider the display surface made up of a flat wall on which projections of two projectors overlap. The images in the two framebuffers are related by a single affine transformation $\mathbf{A}*\mathbf{x} = \mathbf{x}'$, where $\mathbf{A}$ is the unknown 3x2 matrix, $\mathbf{x}$ is homogeneous 3-vector $[x,y,1]^T$ for coordinates of pixels in the first projector and $\mathbf{x'}$ is 3-vector for coordinates of pixels in the second projector.

However, in general, the display surfaces are difficult to define globally due to local variations or desired flexibility. For example, a cylindrical display surface may have small deviations from its design. Hence, the geometric and intensity blending functions need to be computed per-pixel.

The goal is to geometrically predistort and blend the image for every projector such that, when projected and viewed from the position of an ideal viewer, the complete image appears undistorted with normalized intensity.

2.1 Geometric registration

For geometric registration as well as intensity normalization, a wide-field-of-view video camera is located at the ideal viewer's location. The view of the camera is the final desired view for the user. We will call the image from this desired view as 'desired image'. A subset of projector pixels are projected one by one and observed by the video camera.

This gives a mapping from projected pixel to coordinates in desired image. Using bilnear interpolation, we compute mapping from all visible projector pixels to desired image pixels coordinates. Inverse of this mapping allows us to warp desired image into an image that will be projected by the projector.

The projected pixels used to compute this mapping are on an adaptively subsampled grid. The grid is dense in the region where projections overlap and where the display surface is not flat (planar). Typically, only one projected pixel is observed during one video frame. However, the projector pixels can be binary coded and observed in camera in $\log_2$*(number of projected pixels)* frames.

The observed locations of all visible projector pixels in camera image are unstructured set of points. This set is unstructured because (a) the projected pixels may come from adaptively sampled grid and (b) some projected pixels may not be reliably detected in the camera due to occlusion, low intensity or noise. Now, we have the desired image coordinates to projector coordinates mapping at discrete points in the camera image. We use piecewise bilinear interpolation to compute the mapping for rest of the pixels. This assumes local monotonicity within the projected image and the camera image i.e. pixels inside the triangle formed by three projected points in projector framebuffer maintain the same order when projected and viewed by camera. The bilinear interpolation for camera pixels is performed using the three nearest samples in the camera image. The discrete set of points is tessellated using delunay triangulation and the bilinear interpolation is achieved by affine texture mapping. To reduce the number of textured triangles, one can also do triangle simplification in 2D. A mapping vector for a camera pixel is vector between camera pixel location and corresponding projector pixel location, assuming they are in the same coordinate system. A simple simplification strategy is to remove active samples (vertices) if the mapping vector at that pixel is similar to the mapping vector of its neighboring samples. Simplification, however, can introduce visible perspective distortion after texture mapping.

2.2 Intensity Normalization

After the projections from multiple projectors are geometrically registered, the regions of display surfaces where projections overlap, the perceived intensity is higher than where the projections do not overlap. This increase is primarily dependent on the number of projectors active in the overlap region. The image generated by projections from multiple projectors is said to be *intensity normalized* if the intensity at any pixel in the observing camera is equal to the intensity due to the same image generated by a single projector. To create spatially immersive displays, the projectors are usually placed so that a maximum of 2 projections overlap side-by-side. We will focus on two projection overlap in the following discussion but this can be extended to the case of multiple projections.

The projection overlap is observed in the camera image space (CIS). To find the extent of projection of a projector in the wide-field-of-view camera image space, all the pixels in the projector are turned on and the resultant camera image is segmented by thresholding. The projection overlap due to multiple projectors in CIS is the intersection of the set of pixels in the extent of each projector in CIS.

The projection overlap is used to determine the intensity weighting basis functions (IWBF). Determining IWBF that are most insensitive perceptually and to electro-mechanical variations is a non-trivial task. We use linear intensity ramps, assuming that the display surface at overlap is oriented approximately at the same angle with respect to the two projectors i.e. ignoring the sign, the angles between display surface normal in the overlap region and light vectors to the two projector light sources is approximately the same.

In the overlap region in CIS, the intensity contribution from one projector decreases linearly with an intensity weight of 1.0 at one end and 0.0 at the other end. For the other projector active in this overlap, in CIS, the weight changes linearly from 0.0 to 1.0 in the same direction. Outside the overlap region and where only one projector is active, the intensity weight for all pixels in CIS for that projector is 1.0. Thus, the sum of intensity weights of all projectors at any pixel in CIS is always 1.0. If all the projectors have identical and linear dynamic response curves, the resultant image on display surfaces will be intensity normalized.

The intensity weights are computed only at samples in CIS where projected dots were observed. Using the camera pixel coordinates to projector pixel coordinates mapping, the intensity weights are assigned to pixels in the projector image space. The intensity weights are treated as *alpha* values during rendering using APIs such as OpenGL. Due to tessellation and smooth shading with the triangles, the alpha values are bilinearly interpolated for other projector pixels.

3. Rendering

Rendering stage uses a two-pass method for every frame. The desired image is computed by rendering from the desired viewpoint (or a video image is received). The desired image is stored in the texture memory.

In the second pass, the image is warped using 2D texture mapping. In projector image space, the vertices are basically locations of dots projected during calibration. Since, not all the projected dots were reliably observed, the triangles for these vertices were created using delunay triangulation in CIS. The texture coordinates are locations in CIS where these dots were observed.

All the vertices are also assigned alpha values for intensity normalization. Thus the texture color value assigned to a given triangle is multiplied by the bilinearly interpolated alpha values in the triangle. If the dynamic response curve is not linear but known, the alpha values are accordingly modified. For example, non-linear behavior of projectors due to gamma distortion with exponent γ, can be rectified by changing alpha values with a gamma function with exponent $1/\gamma$.

4. Results

We used a 3 projector setup and a video camera with 110^0 field of view in a front-projection display environment of size 12x6x10 (width1 x width2 x height) feet. The projectors were 800x600 resolution and the camera 640x480. Two of the projectors overlap along a flat surface creating 12x6 (width x height) ft image. The third projection 8x6 ft is along a connected perpendicular wall.

The digital light projectors used [Hornbeck95] have a gamma distortion with $\gamma = 2.2$. The projector image space was sampled in a grid of size 100x100 pixels. Grid size was 10x10 pixels in the projection overlap region and where the display surface was not flat.

During rendering a single SGI Onyx machine with a framebuffer large enough to drive all three projectors was used. However, three individual image generators can also be used. In the first pass of the rendering an image is generated in one of the three framebuffers and the colors values in framebuffer are stored in texture memory. During the second pass, the texture is warped into three different views. The rendering is done at 30Hz.

Geometrically, the image appeared to be a generated from a single projector when viewed from the ideal user's location. The overlap region common in the first and second projection also showed correct intensity blending and no intensity seams.

A curved object with footprint of 3x6 (width x height) ft was kept in the overlap region and the system was recalibrated. The recalibration took approximately 3 minutes, projecting single dot at a time, to compute camera-to-projector pixel mapping. With the new warping and intensity weighting functions, the curved object 'disappeared'. The generated images were again geometrically correct from the ideal viewpoint and also photometrically normalized.

5. Conclusion

We have described a robust method to calibrate and render seamless images for a multi-projector display environment where atmost two projections overlap. Using a video camera during calibration, the necessary geometric warping function and intensity weighting function for projector pixels can be calculated. We are currently working on more general-purpose methods to handle arbitrary overlap circumstances.

The described completely automatic method can be used for one-time calibration of display environments or for periodic maintenance. Using a cluster of off-the-shelf projectors, one can avoid the use of precise electro-mechanical setup and expensive optics. Wide field of view display environments can be created without geometrically well-defined display surfaces.

References

[Bennett98] David T. Bennett. Chairman and Co-Founder of Alternate Realities Corporation, 215 Southport Drive Suite 1300, Morrisville, NC 27560, USA. Cited 29 March 1998, available at http://www.virtual-reality.com.

[Cruz93] Carolina Cruz-Neira, Daniel J. Sandin, and Thomas A. DeFanti. 1993. “Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE,” SIGGRAPH 93 Conference Proceedings, Annual Conference Series, ACM SIGGRAPH, Addison Wesley.

[Hornbeck95] Hornbeck, Larry J., "Digital Light Processing for High-Brightness High-Resolution Applications," http://www.ti.com/dlp/docs/business/resources/white/hornbeck.pdf, 1995 [cited 21 April 1998].

[Jarvis97] Kevin Jarvis, “Real Time 60Hz Distortion Correction on a Silicon Graphics IG,” in Real Time Graphics, Vol. 5, No. 7, pp. 6-7, February 1997.

[Raskar98a] Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs. 1998. “The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays,” to appear in SIGGRAPH 98 Conference Proceedings, Annual Conference Series, Addison-Wesley, July 1998.

[Raskar98b] Raskar, Ramesh, Matt Cutts, Greg Welch, Wolfgang Stürzlinger. "Efficient Image Generation for Multiprojector and Multisurface Displays," Rendering Techniques ‘98, Drettakis, G., Max, N. (eds.), Proceedings of the Eurographics Workshop in Vienna, Austria, June 29-July1, 1998