RANSAC-Assisted Display Model Reconstruction for Projective Display

Patrick Quirk, Tyler Johnson, Rick Skarbez, Herman Towles, Florian Gyarfas and Henry Fuchs

Department of Computer Science
University of North Carolina at Chapel Hill

ABSTRACT

Using projectors to create perspectively correct imagery on arbitrary display surfaces requires geometric knowledge of the display surface shape, the projector calibration, and the user's position in a common coordinate system. Prior solutions have most commonly modeled the display surface as a tessellated mesh derived from the 3D-point cloud acquired during system calibration.

In this paper we describe a method for functional reconstruction of the display surface, which takes advantage of the knowledge that most interior display spaces (e.g. walls, floors, ceilings, building columns) are piecewise planar. Using a RANSAC algorithm to recursively fit planes to a 3D-point cloud sampling of the surface, followed by a conversion of the plane definitions into simple planar polygon descriptions, we are able to create a geometric model which is less complex than a dense tessellated mesh and offers a simple method for accurately modeling the corners of rooms. Planar models also eliminate subtle, but irritating, texture distortion often seen in tessellated mesh approximations to planar surfaces.

CR Categories: I.3.2 [Graphics System]: Distributed/network graphics— [I.3.5]: Computational Geometry and Object Modeling— [I.3.7]: Three-Dimensional Graphics and Realism— Virtual reality I.4.8 [Scene Analysis]: Range data, Stereo, Surface fitting— [I.6.7]: Simulation Support Systems—Environments

Keywords: functional reconstruction, RANSAC plane fitting, 3D-point cloud simplification, surface fitting, projective display

1 INTRODUCTION

To create undistorted imagery on a display surface using casually placed projectors requires the projected imagery to be pre-warped so that the view of the imagery from the user's perspective will appear correct.

If the display surface is a simple plane, the warping function can be described with a simple homography that re-maps the *ideal* image from the display plane back into the projector's image plane. This homography, which can be represented as a collineation matrix, can be concatenated into the graphics transformation stack to render the pre-warped image at no additional computational expense [9, 13, 2].

If the display surface illuminated by a given projector is more complex (e.g. a multi-plane corner or even an arbitrarily shaped curtain), then we and others [12, 10, 15, 1] use a 2-pass rendering algorithm to re-map the desired image so it looks perspectively correct from a given viewpoint. The actual *warping* is achieved in the second pass by using the graphics hardware to projectively texture the ideal image rendered in the first pass onto a 3D model of the display surface.

Many 2-pass display implementations use a calibrated, stereo camera pair in an up-front calibration step to first evaluate 3D points
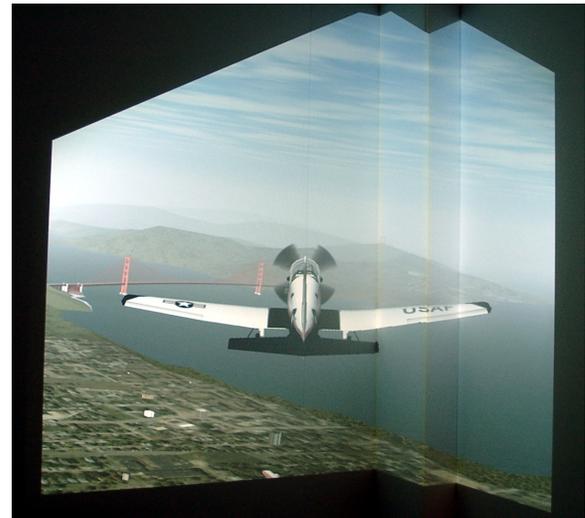

Figure 1: FlightGear displayed into a complex room corner without distortion.

on the display surface based on features that are actively projected. Many types of patterns and codings are commonly used to make feature extraction and the stereo correspondence problem as simple as possible. The resulting 3D-point cloud representation of the display surface is commonly triangulated to produce a tessellated mesh surface description used in the aforementioned 2-pass rendering algorithm.

The challenge with these reconstruction methods is producing an accurate surface model. If sampling is too sparse, the piecewise approximation will not accurately model the display surface shape. This is commonly the case when trying to find the corners of a room. If the sampling is too dense and the tessellated mesh is not simplified, corners may be found with reasonable accuracy but the cost of rendering such a model in the second pass may become overly expensive. In addition, the noise inherent in all image-based stereo reconstruction processes produces 3D points with error. These resulting noisy surface models can produce subtle, but noticeable re-mapping errors in many rendered scenes.

To avoid these difficulties, we have pursued another approach to the task of accurate display surface fitting. It is based on the notion of function reconstruction [6], which turns the task of surface fitting into one of mapping common surface functions (planes, cylinders, spheres) to the surface domain represented by the point cloud samples. In fact, function fitting has been applied by Raskar [11] to reconstruct quadric surfaces for curved-screen projective displays.

In this paper, we describe a method for functional reconstruction of the display surface that takes advantage of the knowledge that most interior display spaces are in fact piecewise planar. Using a RANSAC plane fitting algorithm, we first *recursively* extract individual plane definitions from a complex point sample set originally derived using standard stereo reconstruction methods. The

plane definitions are then processed and converted into simple planar polygon descriptions. The resulting geometric model is significantly less complex than a dense tessellated mesh and more accurately models the corners of rooms, which are computed from the intersection of planes. Using the smooth planar models in the 2-pass projective texture rendering process also eliminates the subtle texture distortion often seen in tessellated mesh approximations to planar surfaces in tiled projective display applications.

In Section 2, we describe our method, in Section 3 we present and discuss some results, and in Section 4 we summarize and offer thoughts on future work.

2 DISPLAY SURFACE MODELING

To re-iterate, display surface modeling along with projector calibration are required in order to pre-warp and thus create perspectively correct imagery for a given user's view. In this section, we detail the three steps our new procedure uses to create an accurate display surface model of complex room geometry.

This room model is then used in the 2-pass application rendering process to achieve the necessary pre-warp of the projected images to create imagery that can wrap around the walls of the room, providing an undistorted immersive experience for the user.

2.1 3D-Point Cloud Generation

A calibrated, stereo camera pair is used to reconstruct the display surface and create the initial 3D-point cloud representation. The basic steps in stereo reconstruction are well established:

- Extract feature points from the stereo camera images.

- Establish pairwise feature point correspondences.

- Triangulate to find the 3D point representing the ray intersection of each corresponding (x,y) feature point pair.

Since ideal display surfaces are without texture, it is most convenient to use the projectors themselves to create the image features needed for surface extraction. Features outside the illuminated region are not needed as it is only necessary in the rendering process to re-map the ideal view image into the image space of each projector.
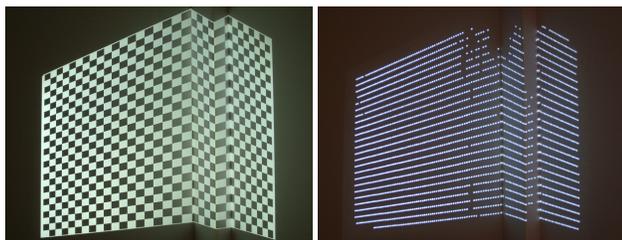

Figure 2: (a) Checkerboard feature pattern. (b) Reprojected 3D feature points.

Only a rather sparse sampling of the piecewise planar surfaces is required for RANSAC plane fitting, so we currently project a checkerboard pattern to establish robust features. The Open Computer Vision library (a SourceForge project) is used to extract the corner features in the stereo images. If the piecewise planar segments of the display surface are small, such as the narrow corner column visible in Figure 2a, we can shift the projected checkerboard pattern horizontally and/or vertically to create more feature points.

To make the task of identifying corresponding features simple and robust between the stereo image pair, we also project a series of images with binary-coded markers that are centered on the checkerboard corners that give each corner feature a unique identification code. Figure 2b shows the resulting 3D points being rendered and reprojected onto the display surface. Missing points may result from checkerboard corner extraction failures or feature point identification decode errors that are generally associated with low image contrast.

Given the list of 2D matching image points and the camera matrix P for each camera, we use a simple linear triangulation method [4] to find the corresponding 3D points. For each point, this involves solving for four homogenous unknowns in four equations using a singular value decomposition method of the DLT algorithm.

2.2 RANSAC Plane Fitting for Recursive Plane Extraction

Since the point cloud generated in our system calibration process could represent multiple planar surfaces, it would be necessary to first segment the data into plane-related point clusters if one were to use a standard iterative minimization technique (e.g. least squares) for plane fitting. This segmentation process would basically classify points as inliers and outliers with respect to a given plane fit operation.

To avoid this complex segmentation step, we use a Random Sample Consensus (RANSAC) algorithm [3] to fit planes to the point cloud data. RANSAC is designed to work in the presence of many data outliers. The algorithm begins by fitting a plane to a set of 3 points randomly selected from the input point cloud. These 3 points are then classified as inliers. All other points are then compared to a cost function (distance from the plane) and are reclassified as inliers if they are within an epsilon distance of the plane. These two steps are repeated until the number of inliers converges to a maximum or a count limit is reached. A least-squares fit of the consensus inliers then defines the output plane definition.

To fit planes to all the data in the the point cloud description, the RANSAC plane fitting process is repeated with the candidate points for the next pass coming from the outliers of the previous pass. We terminate the plane generation loop after either a user-specified number of planes are evaluated or the number of outliers remaining to be processed falls below a heuristic threshold.
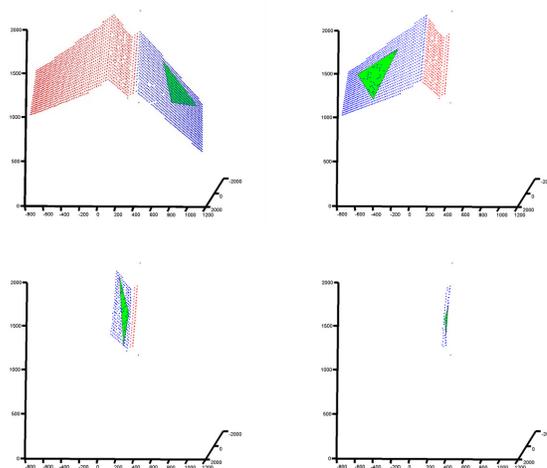

Figure 3: Matlab rendering of successively reconstructed planes.

Figure 3 shows the extraction of four successive planes beginning with a combined 3D-point cloud generated for a display sur-

face looking into a room corner that contains a rectangular column. The green triangle in each image represents the plane extracted on that pass, while the blue points show the inliers and the red points the outliers for that plane. Notice that the number of points to process in each subsequent pass diminishes as inliers for each fitted plane are removed.

Our current plane fitting implementation is in Matlab and we utilize the ransacfitplane.m function authored by Kovesi [7].

2.3 Plane to Polygon Conversion

Given the set of plane definitions generated from the RANSAC plane fitting step, we next create a polygonal representation of the display surface for use in the projective texturing step of our 2-pass rendering algorithm.

The general solution to this task can be a complex issue of which planes to intersect in order to created bounded polygons. To simplify this decision, we currently restrict the display to vertical wall surfaces. This allows us to create a 2D *floor plan* representation of the point cloud data by projecting the inliers of each extracted plane down onto a virtual floor plane of the room, thus creating a line segment for each plane whose endpoints span the projected inliers.

In general, the line segments representing the walls of the room may not overlap due to a lack of point cloud data in the corners of the room. To determine which line segments correspond to contiguous walls, we extend the endpoints of each line segment in order to form intersecting corners. To avoid erroneous corner detection, we consider only those intersections which require the least amount of line segment extension to become intersecting corners.

After extending the line segments in the floor plan to form the corners of the room, it is trivial to create a polygonal display model from the result. We merely extend each line segment in the up-direction by some desired height to turn the line segments into quadrilaterals, forming a simple piecewise planar display model.

3 RESULTS

The ability of our plane fitting method to automatically extract accurate piecewise-planar display surface models is illustrated in Figures 1 and 4b. These images show two different application programs displayed onto a room corner with no perceived distortion. Without an accurate surface model the 2-pass rendering system would simply not be able to create such precise, pre-warped projected imagery in which straight lines remain straight. As a measure of the amount of pre-warping required, Figure 4a shows the actual pre-warped output that projected to create the scene captured in Figure 4b.

Figure 4: Piazza San Marco, Venice. (a) Pre-warped frame buffer output. (b) Undistorted projected view.

Figure 5a shows the actual room corner *without* projector illumination. It is a complex 4-plane structure that includes a 5" x 12" offset column. Figure 5b shows the reconstructed computer model

of this corner with room-aligned coordinate system axes. Beginning with 3618 reconstructed vertices, our RANSAC plane-fitting, surface reconstruction algorithm produced an output display model represented by just 4 quadrilaterals defined by ten 3D vertices. The plane fit tolerance specified was 5 millimeters.
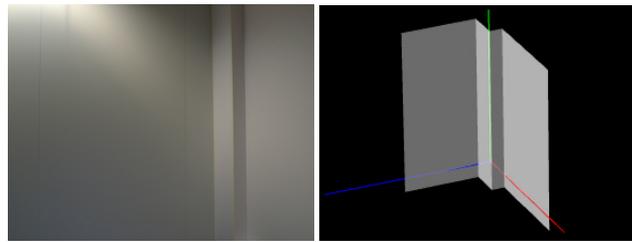

Figure 5: (a) Real-world room corner. (b) Reconstructed display surface model.

One measure of comparing the real room with the extracted model is to compare the dimensions of the offset column. In doing so, we found that the offset column feature was modeled with less than 2.3% dimensional error.

As another measure of the overall display system modeling accuracy, we can also render the display model with each plane distinctly colorized. If the model and projector matrix computation are accurate, then the projected colorized planes will map back precisely onto the their respective planes in the real-world. Figure 6 shows just such a re-projection. Close examination shows a very accurate remapping with color transitions occurring as expected at the surface intersections with one minor exception. In the lower part of the image, the yellow plane overlaps onto the real-world red plane a few millimeters.
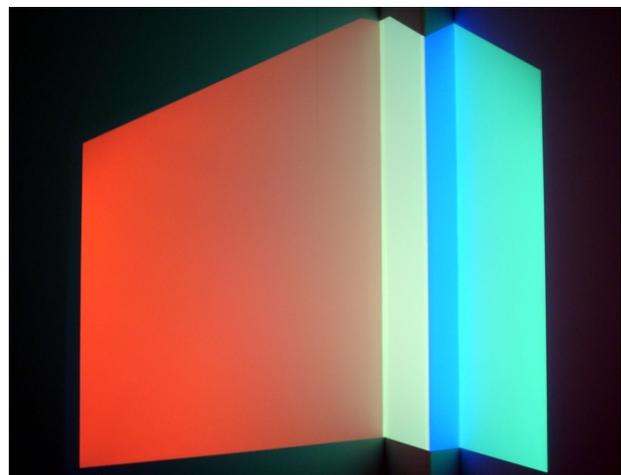

Figure 6: Extracted display model re-projected onto the real-world corner.

Our testing to date has involved configurations of one and two projectors (ProjectionDesign F1 SXGA) illuminating a corner of a room. We have explored two different corner configurations - a standard 2-plane corner and the more complex 4-plane corner. The material presented in this paper is based on a single projector configuration illuminating the complex wall corner.

Our calibration toolset is a combination of C/C++ code and Matlab scripts running on Windows XP. Beginning with a pre-calibrated stereo camera pair, the script-driven process runs automatically taking less than 5 minutes to complete the full rendering system cal-

ibration from checkerboard image capture to display model polygons and projector matrices out.

4 SUMMARY AND FUTURE WORK

We have demonstrated a new method for functional reconstruction of projective display surfaces that takes advantage of the knowledge that most interior display spaces are piecewise planar. Using a RANSAC algorithm to recursively fit planes to a 3D-point cloud sampling of the surface, this new solution clearly generates geometric models that are both simple and accurate in their representation of the display surface shape. As a result, we are now able to more precisely create geometrically correct and pleasing immersive imagery on the walls of any room with casually placed projectors than previously possible when modeling the surface as a complex tessellated mesh.

There are a number of areas for future research including:

- Scaling up to handle more walls and projectors and thereby create a more immersive, wide-area visual display system.

- Improve the robustness of the current plane to polygon conversion algorithm to handle without restrictions an arbitrary configuration of extracted plane definitions.

- Implement the optimal triangulation algorithm [4] to replace the less precise homogeneous DLT method used now.

- More fully review the literature for the applicability of other surface fitting strategies that can model a wider variety of shapes and sharp features. Certainly the research of [5, 14, 8] and others is of relevance. Also, the multiRANSAC algorithm [17] for detecting multiple planar homographies in parallel is of importance.

Longer term, we are especially interested in developing techniques whereby the calibration process is not just a pre-process step to rendering, but one [16] that continuously and successively refines and improves display quality while the user application is running.

REFERENCES

[1] M.S. Brown and W.B. Seales. A practical and flexible tiled display system. In *Proceedings of the 10th Pacific Conference on Computer Graphics and Applications*, pages 194–203, New York, NY, USA, 2002. IEEE.

[2] Han Chen, R. Sukthankar, G. Wallace, and Kai Li. Scalable alignment of large-format multi-projector displays using camera homography trees. In *Proceedings of Visualization 2002*, pages 339–346, New York, NY, USA, 2002. IEEE.

[3] Martin A. Fischler and Robert C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, 1981.

[4] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in computer vision*. Cambridge University Press, first edition, 2002.

[5] Hugues Hoppe, Tony DeRose, Tom Duchamp, Mark Halstead, Hubert Jin, John McDonald, Jean Schweitzer, and Werner Stuetzle. Piecewise smooth surface reconstruction. In *SIGGRAPH '94: Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 295–302, New York, NY, USA, 1994. ACM Press.

[6] Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, and Werner Stuetzle. Surface reconstruction from unorganized points. In *SIGGRAPH '92: Proceedings of the 19th annual conference on Computer graphics and interactive techniques*, pages 71–78, New York, NY, USA, 1992. ACM Press.

[7] Peter Kovesi. Ransacfitplane - a robust matlab function for fitting planes to 3d data points. http://www.csse.uwa.edu.au/(tilde)pk/research/matlabfns/robust/ransacfitplane.m, June 2003.

[8] Maxime Lhuillier and Long Quan. A quasi-dense approach to surface reconstruction from uncalibrated images. *Transactions on Pattern Analysis and Machine Intelligence, IEEE*, 27(3):418–433, 2005.

[9] Ramesh Raskar. Immersive planar display using roughly aligned projectors. In *Proceedings of Virtual Reality 2000*, pages 109–115, New York, NY, USA, 2000. IEEE.

[10] Ramesh Raskar, M. Brown, R. Yang, W.-C. Chen, G. Welch, H. Towles, B. Seales, and H. Fuchs. Multi-projector displays using camera-based registration. In *Proceedings of Visualization '99*, pages 161–168, 522, New York, NY, USA, 1999. IEEE.

[11] Ramesh Raskar, Jeroen van Baar, Thomas Willwacher, and Srinivas Rao. Quadric transfer for immersive curved screen displays. *Computer Graphics Forum*, 23(3):451–460, 2004.

[12] Ramesh Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs. The office of the future: a unified approach to image-based modeling and spatially immersive displays. In *SIGGRAPH '98: Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 179–188, New York, NY, USA, 1998. ACM Press.

[13] Rahul Sukthankar, Robert Stockton, and Matthew Mullin. Smarter presentations: exploiting homography in camera-projector systems. In *Proceedings of International Conference on Computer Vision, 2001*, 2001.

[14] George Vosselman, B.G.H. Gorte, G. Sithole, and T. Rabbani. Recognising structure in laser scanner point clouds. *International Archives of Photgrammetry, Remote Sensing and Spatial Information Sciences*, 46(8/W2):33–38, 2004.

[15] Ruigang Yang, D. Gotz, H. Towles J. Hensley, and M. Brown. Pixelflex: a reconfigurable multi-projector display system. In *Proceedings of Visualization 2001*, pages 167–174, 554, New York, NY, USA, 2001. IEEE.

[16] Ruigang Yang and Greg Welch. Automatic projector display surface estimation using every-day imagery. In *9th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2001*, 2001.

[17] M. Zuliani, C. S. Kenney, and B. S. Manjunath. The multiransac algorithm and its application to detect planar homographies. In *IEEE International Conference on Image Processing*, Sep 2005.