IPT-EGVE Symposium (2007)
B. Fröhlich, R. Blach, and R. van Liere (Editors)

Short Papers

A Unified Multi-Surface, Multi-Resolution Workspace with Camera-Based Scanning and Projector-Based Illumination

T. Johnson and H. Fuchs

University of North Carolina - Chapel Hill, USA

Abstract

Research in the area of projector- and camera-augmented office environments has demonstrated the use of cameras as desktop scanning devices, shown the benefits of using multiple display devices for focus and context information, and advocated display on multiple surfaces for visualization of multi-dimensional data. In this paper, we describe how the implementation of these ideas can be unified to simplify the creation of a workspace that combines them.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Display algorithms

1. Introduction

Research in projector- and camera-augmented office environments has made it possible to create large digital desktops that allow digital documents to be interacted with as if they were real documents. The digital desk [Wel93] blurred the line between real and digital content by allowing digital documents displayed on a desk surface to be marked with a pen as if they were real. It also demonstrated how cameras in the office could be used for desktop scanning, allowing user-selected portions of real documents to be copied and pasted into digital documents directly on the desk.

Other researchers have experimented with combining multiple display devices to create focus plus context displays [BGS01, AR03]. In this type of system, one display provides high spatial resolution for tasks such as text editing, while another display provides context information at a lower spatial resolution, e.g. other documents that may be of interest, but not currently being edited. This context information is useful in many applications.

By extending the digital desktop beyond just the desktop itself and onto other surfaces as well, the workspace can also be used to visualize 3D content. Ashdown et al. [AFSR04] describe how this can be accomplished using a multi-planar surface. In this work, an automatic technique for calculating mutually consistent rectifying homographies between each plane and the projector is described.

We imagine an office environment that combines all of these ideas in a single workspace. A conceptual illustration of our vision for this new type of workspace is shown in Figure 1. In this paper, we describe techniques for calibration

Figure 1: *Multi-surface workspace concept. (Andrei State)*

and rendering that allow these previous ideas from projector- and camera-augmented office environments to be combined to create a multi-surface, multi-resolution workspace with camera-based scanning and projector-based illumination. We show that techniques from multi-projector displays can provide a robust calibration procedure and simplify the task of combining these ideas in a common framework.

2. The Workspace

The workspace we have created is reminiscent of an office cubicle. Our environment, as seen in Figure 3, consists of a flat desktop abutted on two sides by vertical walls. An LCD panel is also embedded within the surface of the desktop. A projector-camera module consisting of a projector and two greyscale cameras is positioned a short distance away to illuminate the surface of the desk as well as the two walls.

3. Calibration

Before imagery is displayed on the workspace, a simple calibration procedure is performed. This determines the information needed to project imagery onto the surfaces of the workspace in a way that compensates for the orientation of the surfaces with respect to the projector. Instead of a homography-based approach, we perform a full 3D calibration, which easily allows content to be displayed across multiple surfaces. The procedure estimates a polygonal model of the workspace geometry as well as a projection matrix for the projector, both cameras, and the LCD panel. This determines an invertible mapping from 2D pixel locations in each device to 3D locations on the surfaces of the workspace.

Our calibration process begins by projecting encoded structured light patterns that are observed by the cameras. Decoding of these structured light patterns allows precise image correspondences between the cameras to be obtained. Since the projector and cameras are rigidly attached, we assume the devices have been pre-calibrated using an existing technique such as [Zha00, RWC*98]. The camera calibration can then be used to triangulate each correspondence into a 3D point, forming a 3D point-cloud representation of the surfaces of the workspace.

To construct a polygonal model from this point-cloud representation, we use the RANSAC-based plane-fitting algorithm described in Quirk [QJS*06], which is robust against noise and outlying points resulting from false stereo matching. The algorithm is used to find the three most dominant planes in the point-cloud data, which are then intersected to form a simple polygonal mesh. This plane intersection process allows the wall corners of the workspace, where calibration errors will be most apparent, to be accurately estimated.

We treat the LCD screen as a projector and display the same structured light patterns. This results in a point-cloud representation of the plane of the LCD screen in the same coordinate system as the projector-camera module and the polygonal model of the workspace. We then calculate a projection matrix relating the 3D points on the surface of the LCD screen to their 2D pixel coordinates in the LCD image. While this mapping could also be accomplished using a homography, a 3D calibration allows the LCD screen to be handled in the same way as the projector during rendering.

The standard technique for calculating projection matrices from 2D-3D correspondences is the direct linear transform (DLT) [AAK71]. In the case of the LCD screen, the 3D points are coplanar, and the DLT algorithm will fail. Knowledge of the center-of-projection (COP) is however sufficient to fully determine a projection matrix. The choice of COP is arbitrary in this case, so we choose it to lie some distance along the normal of the LCD plane.

4. Rendering

In this section, we describe how it is possible to provide the user with content displayed on only a single surface with content displayed on multiple surfaces in a unified way using a single technique - two-pass rendering [RWC*98]. We also describe how this same technique can be used to correct image distortion when using a camera for desktop scanning.

4.1. Two-Pass Rendering

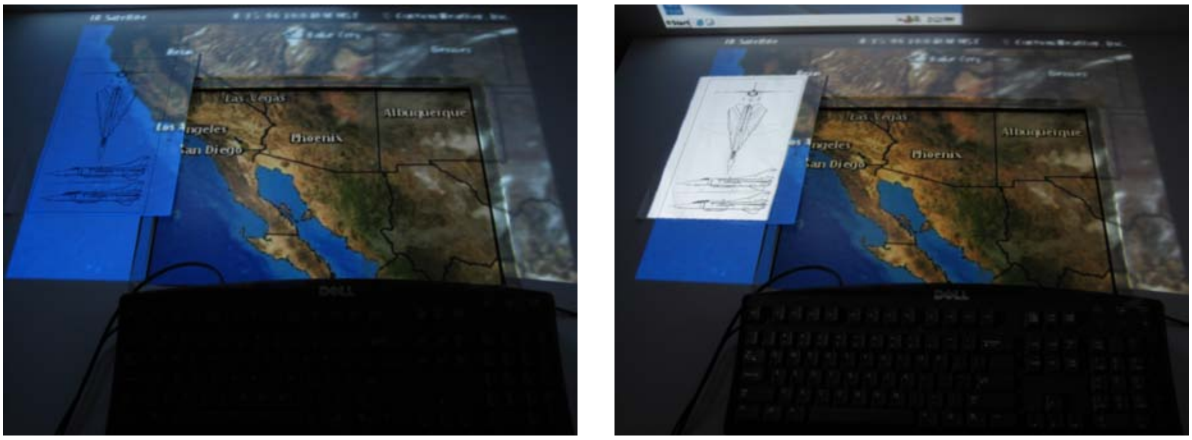
The basic approach behind two-pass rendering is to take a desired image to be observed on a projection surface and determine the necessary warping of this desired image to compensate for the geometry of the projection surface.

The desired image is rendered in pass one. This can be an image generated from a 3D graphics application or simply an image of a 2D windowed application such as a word processor. The warping of this image occurs in the second pass using projective texturing, where the image is projected onto the geometry of the projection surface from the user's viewpoint. The textured surface geometry is then rendered from the perspective of the projector using its projection matrix.

4.2. Single-Surface Content

Some types of imagery, such as that of a word processor, are best displayed to the user on a single surface of the workspace. While this content can be made to look geometrically correct across multiple surfaces from the user's viewpoint, the notion that this type of content is flat can conflict with the stereo cues the user receives and be disturbing.

Using two-pass rendering, we can easily determine how to display a window in order to align it with the natural coordinate system of one of the workspace surfaces. When a window is opened on one of the surfaces, the four coplanar corners of the window are used to calculate a projection matrix with a COP located a short distance from the center of



a) b)

Figure 2: ***a)*** *Undesired projection on a real document.* ***b)*** *Neutral illumination on a document and keyboard.*

the window along the plane normal. This is identical to the calibration process we described for the LCD screen.

During rendering, this matrix is used as the projective texturing matrix in pass-two to project the window contents onto the geometry of the workspace model before it is rendered from the perspective of the projector. Note that this process for displaying content on a single surface is independent of the location of the viewer.

4.3. Multi-Surface Content

The two-pass rendering algorithm also supports the display of imagery across multiple surfaces. The workspace can then also be used as a convenient medium for visualizing content such as 3D models or virtual data sets. The rendering process used to display this content is identical to that used to display single-surface content except that the COP of the projective texturing matrix is constrained to be the viewpoint of the user, which can be maintained by a separate tracking system.

4.4. Desktop Scanning

Using the projector-camera module, it is also possible to create a desktop scanner, allowing the user to copy from real documents and paste into digital ones without the need to leave his desk. This technology was first demonstrated as part of the digital desk [Wel93]. Here, we describe how this technology can be implemented in the context of a multi-surface workspace to allow documents located on any surface of the workspace to be copied.

Using a selection rectangle, which is handled as a single-surface window that follows the motion of the mouse, the user selects the portion of a document he wishes to copy. The projector then illuminates the selected area to improve its brightness while the camera captures an image. Two steps now remain - the user-selected portion of the camera image must be segmented from the rest of the image, and the distortion caused by the orientation of the camera with respect to the surface must be removed.

Both of these steps can be accomplished simultaneously using two-pass rendering. We set the desired image to be the captured camera image and use the camera's projection matrix as the projective texturing matrix. The geometry that is projectively textured and rendered is a quad formed from the corners of the selection rectangle. This quad is then rendered with the projection matrix of the selection rectangle window.

4.5. Projector-Based Illumination

There are often objects present in the office environment on which it may not be desirable to have imagery projected. As shown in Figure 2a, when documents have imagery projected on them, it can be difficult to distinguish the document from the projected imagery. This is especially troubling for documents containing text, which are made difficult to read.

We have created a simple interface that allows the user to select the corners of a quad where uniform white light should be displayed by the projector. In Figure 2b, the imagery projected on the document has been replaced with uniform white light from the projector, allowing the contents of the document to be clearly seen. The keyboard has also been illuminated by the projector in this image.

5. Results

We have implemented an application allowing windows containing various types of content to be opened on any surface of the workspace. The windows are simply static frames used to demonstrate the calibration and rendering and are not tied to any application. Figure 3 is an image of the application running. Using our rendering process, when a window overlaps the LCD screen, the overlapping window contents are automatically displayed in high resolution on the LCD screen and registered to the surrounding projected imagery. The satellite in the image demonstrates the ability to simultaneously view content on multiple surfaces.

Figure 4 shows a hard-copy document and the resulting digitized version selected by the user to be scanned. The distortion involved in capturing an image of the document from the perspective of the camera has been eliminated.

Figure 3: *A multi-surface, multi-display workspace combining both single- and multi-surface content.*

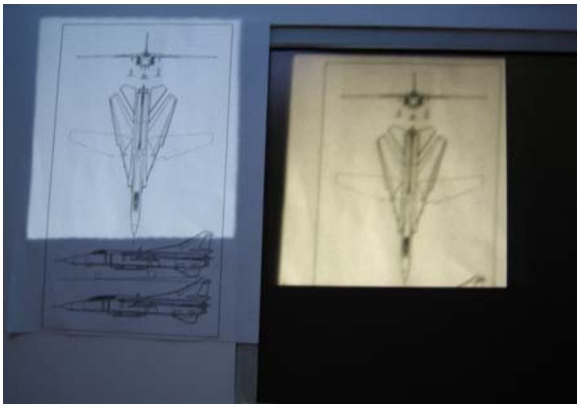


Figure 4: *Desktop scanning using a projector-camera module.*

6. Conclusions and Future Work

We have described a framework that allows previously separate ideas from projector- and camera-augmented office environments to be combined into a single workspace. Our system includes a multi-planar surface, a projector-camera module, and an LCD screen and allows single- and multi-surface content to be displayed in a unified way. Using this framework, the LCD screen is calibrated to the projector and its imagery automatically registered to the projected imagery to provide an area for tasks requiring high resolution. We also describe how the cameras of the projector-camera module can be used for desktop scanning within the same framework to further simplify implementation.

We have focused solely on rendering and calibration issues in this paper, but the potential for new interaction techniques should also be investigated. We think this type of display has great potential for remote collaboration with the wall surfaces being used as a window to a remote environment. In the future, we hope to replace the need for manual selection of objects that should only be illuminated by white light with an automatic technique that detects objects occluding projected imagery.

7. Acknowledgments

The authors would like to thank Herman Towles and John Thomas for building the workspace used in this project, as well as Andrei State for his conceptual artwork. This work was partially supported by the VIRTE program under Office of Naval Research award number N00014-03-1-0589.

References

[AAK71] ABDEL-AZIZ Y., KARARA H.: Direct linear transformation into object space coordinates in close-range photogrammetry. In *Symposium on Close-Range Photogrammetry* (1971), pp. 1–18.

[AFSR04] ASHDOWN M., FLAGG M., SUKTHANKAR R., REHG J.: A flexible projector-camera system for multi-planar displays. In *CVPR* (2004).

[AR03] ASHDOWN M., ROBINSON P.: The escritoire: A personal projected display. In *11th International Conference in Central Europe on Computer Graphics* (2003), pp. 33–34.

[BGS01] BAUDISCH P., GOOD N., STEWART P.: Focus plus context screens: combining display technology with visualization techniques. In *UIST* (2001), pp. 31–40.

[QJS*06] QUIRK P., JOHNSON T., SKARBEZ R., TOWLES H., GYARFAS F., FUCHS H.: Ransac-assisted display model reconstruction for projective display. In *Emerging Display Technologies* (2006).

[RWC*98] RASKAR R., WELCH G., CUTTS M., LAKE A., STESIN L., FUCHS H.: The office of the future: a unified approach to image-based modeling and spatially immersive displays. In *SIGGRAPH* (New York, NY, 1998), ACM Press, pp. 179–188.

[Wel93] WELLNER P.: Interacting with paper on the DigitalDesk. *Communications of the ACM 36*, 7 (1993).

[Zha00] ZHANG Z.: A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence 22*, 11 (2000), 1330–1334.