Embedded Entertainment with Smart Projectors

Essentially video projectors enhanced with sensors to gain information about the environment, smart projectors do not require artificial canvases and allow correct projection of images onto many arbitrary existing surfaces, such as papered walls or curtained windows.

Oliver Bimber
Andreas Emmerling
Thomas Klemmer
Bauhaus University Weimar

Television played a central role in shaping the 20th century and remains the primary entertainment medium for most of us today. TV continues to evolve as innovative display technologies change its look and capabilities at an accelerating rate. The popularity of today's flat-panel liquid-crystal and plasma displays shows that emerging trends favor large-screen displays. Simultaneously, falling prices have led to a booming market for home entertainment technology. However, the physical limitations inherent in these technologies place constraints on maximum screen size, display size, refresh rate, and power consumption.

Another display type may soon conquer the entertainment market, however: *Video projectors* have experienced an enormous metamorphosis during the past decade. The cost reductions and performance increases made in these devices compare favorably with those personal computer manufacturers achieved decades earlier. Video projectors also offer a vital advantage over other display technologies. They can generate images much larger than the devices themselves without being constrained by a traditional TV screen's limitations.

This ability comes at a price, however: The artificial canvas requires a space equal to the size of the image we want displayed. A home theater, for example, might require an entire room. In many situations, the temporary or stationary canvases that projector-based multimedia presentations require also harm the ambience of environments such as a living room or historic site.

Smart projectors, however, do not require an artificial canvas. Instead, they allow a correct projection onto many arbitrary existing surfaces, such as papered walls or curtained windows.

SMART PROJECTORS

Essentially video projectors enhanced with sensors to gain information about the environment, *smart projectors* primarily use cameras to sense their environment. However, other information gathering devices such as *tilt sensors* are also available. Completely calibrated and mounted as a single camera projector unit, or realized with separated components, some smart projectors allow dynamic elimination of shadows the user casts,[1] automatic keystone correction on planar screens,[2] or manually aligned shape-adaptive projection on second-order quadric display surfaces[3] such as cylinders, domes, ellipsoids, or paraboloids.

For projection planes, cameras can help to automatically register multiple projector units based on homographic relationships.[4] In this case, camera feedback also provides the data for intensity blending and color matching[5] of multiple projector contributions. Combining calibrated stereo cameras with projectors allows direct scanning of an arbitrary display surface's 3D geometry, enabling undistorted projection for a known head-tracked observer position.[6]

All these approaches require the calibration of cameras and projectors to determine their intrinsic position—the focal length, principal point, skew

Published by the IEEE Computer Society

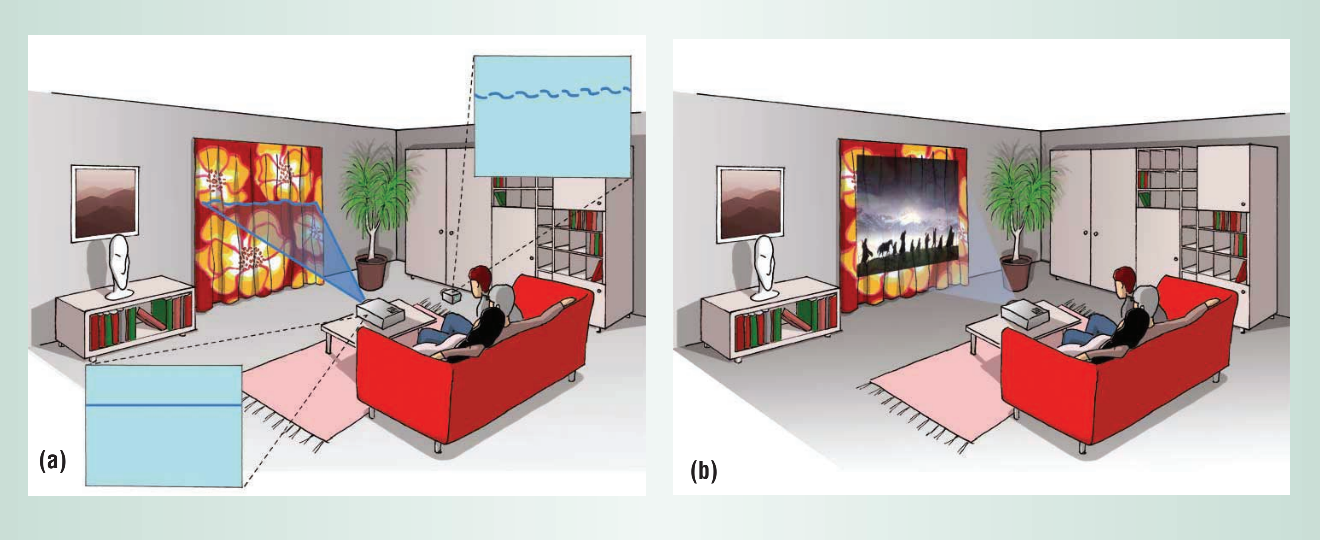


Figure 1. Smart projector concept. The temporarily detached modular camera component (a) calibrates the projector unit by mimicking the observers' target perspective so that the projector can display (b) a geometry- and color-corrected image on a curtained window.

angle, aspect ratio, and field of view—as well as their extrinsic position and orientation parameters. Although some systems can project geometrically predistorted images for a known observer position onto scanned or modeled nonplanar surfaces beforehand, these surfaces are still fairly simple, such as adjacent even walls. Surfaces with fine geometric details represent overkill for the real-time predistortion realized by processing a high-resolution 3D model.

Until now, all projection surfaces, both planar and nonplanar, have required a uniform white texture. Per-pixel color correction, however, becomes feasible with the enhanced capabilities of recent graphics chips. A projection onto arbitrarily textured surfaces has been achieved, for example, with the aid of a special transparent film material that reflects a portion of the incident light.[7]

Although this technique allows superimposing flat paintings onto projected multimedia content, it cannot be employed to display images onto everyday surfaces for two reasons. The technique

- still requires an artificial transparent canvas that can be applied to a plain surface; and
- it needs a precise, manual pixel-to-pigment registration.

The too-low resolution of today's cameras prevents the implementation of an automated calibration process for this special case.

Color- and geometry-corrected projection onto arbitrarily shaped and textured surfaces is possible in real time, with fully automatic, fast, and robust calibration. A compact device, such as that depicted in Figure 1, has yet to be built, however. Instead, the first proof-of-concept prototypes are a combination of off-the-shelf components—such as a consumer LCD video beamer, a CCD camcorder, and a personal computer with a TV card and a pixel-shading-capable graphics board.

CREATING VIRTUAL PROJECTION CANVASES

The smart projector concept combines camera feedback with structured light projection to gain information about the screen surface and the environment. Calibrating the system does not require having information about either the surface geometry or the internal or external parameters of the projector and camera. This makes the system extremely robust and easy to use—crucial attributes for home-entertainment and similar applications.

The modular camera component can be detached from the smart projector's projection unit for calibration. It must be temporarily placed approximately at the observers' optimal viewing location or *sweet spot*—pointing at the screen surface as Figure 1a shows. The projection unit can be placed at an arbitrary location. Its light frustum must also cover the screen surface area.

During calibration, the camera mimics the *target perspective*—the optimal viewing position for which the projection unit will be calibrated. The user can either define the display area by sketching the outlines of a virtual projection canvas over a portion of the camera image or derive it automatically from the margins of the camera's field of view.

The calibration process compensates for camera lens distortion at the start to provide video images without radial distortion. The system then determines all parameters required for real-time *geometric predistortion* and *color correction* of video frames delivered by a PAL/NTCS-compliant device such as a DVD player or game console. The fully automated calibration process takes less than 30 seconds with the chosen hardware configuration.

After the system has been calibrated, the camera module can be removed. Henceforth, the projector unit corrects incoming video signals geometrically and photometrically in real time at no less than 100 frames per second. If the system projects the corrected images onto the nontrivial screen surface, the observer will see them as they would appear if projected onto a plain white canvas. However, this projection canvas is completely virtual and does not exist in material reality, as Figure 1b shows.

For consumer applications, the calibration of smart projectors must be fully automatic, fast, and robust.

VANISHING SHAPES

Smart projector system developers also seek to geometrically predistort the input images so that if they are projected onto a geometrically nontrivial surface and observed from an area close to or at the target perspective, these images appear correct. For consumer applications, the calibration of smart projectors must be fully automatic, fast, and robust.

Projection systems sometimes apply wide-field-of-view cameras in a sweet spot position to calibrate multiple overlapping projectors.[8] Projecting pixels and capturing them with the camera results in a projector-to-camera pixel mapping. For performance reasons, only a subset of projector pixels are usually displayed and captured, while the mapping for the remaining ones are interpolated linearly. For arbitrarily shaped surfaces with fine geometric details, however, a high-resolution pixel correspondence must be generated in an acceptable time.

To realize this goal, developers can adapt time-multiplexed line-strip scanning techniques from structured 3D range-finder systems. Placing the camera at the observer's sweet spot matches its view to the target perspective. Then the developer can apply a variation of a column-row coded-pattern projection methods,[9] with phase shifting similar to that proposed by Jens Gühring,[1] to compute a *pixel displacement map*. This creates a lookup table that maps *every* camera pixel to the corresponding projector pixel.

The projector unit displays particular calibration images—for example, the line strips for geometry or a uniform white image for photometry. The camera must capture these images while the projector is displaying them. This means that both must be synchronized and for this we must know the camera's latency in terms of projecting the calibration images long enough. This knowledge helps ensure correct synchronization between projection and capturing during the scanning process.

Hardware image compression and data transfer between the camera and receiving device cause latency. This latency is particularly high for consumer camcorders because they do not target real-time image-processing applications. The smart projector determines the camera's latency automatically at the beginning of the geometric calibration process. The projector does this by sending out sample patterns and measuring the maximum time until it can detect these patterns in the recorded camera images.

For an XGA projector resolution of 1,024 × 768, a PAL camera resolution of 720 × 576, a maximum camera latency of 80 ms for a consumer camcorder delivering an s-video signal over a TV-in channel, and an average image processing duration of 150 ms, the smart projector's total time to generate the displacement map is approximately 28 seconds. The process requires no user intervention.

The different camera and projector resolutions, and their varying distances and perspectives to the screen surface, prevent the displacement map from representing a one-to-one pixel mapping. The mapping might not even be complete because surface portions can lie in shadow areas. Applying multiple projector units can overcome this problem.

Different projected line strips might project on the same camera pixel. To achieve subpixel precision, the system computes and then stores the average values in the displacement map.

If the camera and projector can be calibrated precisely, the system could use the pixel correspondences in the displacement map and triangulation to recover the screen surface's entire 3D geometry. Some 3D scanners function in exactly this way. Given that we do not expect both devices to be located at known positions, the displacement map allows only the mapping of each camera pixel from the target perspective into the projector's perspective. This results in an undistorted perspective even if the screen surface's 3D shape is unknown.

To benefit from hardware-accelerated computer graphics, the displacement map is converted into a texture map—realized with a 32-bit/16-bit P-buffer—that stores a reference for *every* projector pixel and its corresponding video and camera pixels. The system then passes this texture as a parameter to a modern *pixel shader,* which implements real-time image warping via a *pixel displacement mapping*. Standard components today in many consumer graphics cards, pixel shaders enable per-pixel operations. Besides the *displacement texture map*, the system passes several other parameter textures to the pixel shader, such as the uncorrected input image itself.

To execute a geometric-image predistortion, the system need only render a single 2D rectangle into the projector's entire frame buffer. This triggers the rasterization of every projector pixel through the pixel shader before display. The colors of incom-

Figure 2. Projecting images onto environmental surfaces. (a) A scruffy corner serves as the projection surface. (b) The uncorrected image. (c) The projector system corrects the image's geometry and, finally, (d) its color. Displayed content: The Jackal*, Universal Pictures.*

ing pixels are simply overwritten by new colors that result from the corresponding input image pixels. These input image pixels can be found with the aid of the displacement texture map, which has the same effect as actually moving the input image pixels to new positions within the projector frame buffer. The colors are not just copied from input image pixels to projector pixels—they are also modified to enable color correction. This allows warping every pixel of the input image in real time without first acquiring geometric information of the screen's surface.

Viewers can perceive the image as geometrically correct at or near the target perspective. Depending on the screen surface's shape, the observer will detect a more or less extreme distortion. For the walls in Figure 2, a horizontal deviation from the target perspective leads to a larger distortion than a vertical deviation, yet the pitched roof surface in Figure 3 would cause the opposite effect. The latter surface is thus better suited to a group of users sitting or standing next to each other.

These are two extreme cases, however. The geometric distortion of, for example, a window curtain and a natural stone wall are relatively low in any direction if viewers are observing the projected image from an adequate distance.

Even though viewers observe the predistorted projected image as geometrically correct on a nontrivially shaped surface, its uncorrected colors will still be blended with the screen surface's texture.

NEUTRALIZED TEXTURES

When light strikes a surface, only a fraction of its original intensity and color reflects back; the surface absorbs the rest. For Lambertian (completely diffuse) surfaces, the amount and color of reflected light depends on several parameters, such as the surface's material color (M), the light color and intensity that leaves the source (I), as well as the distance (r) and the incidence angle (α) of light rays with respect to the surface—together called the *form factor* (F). For perfectly diffuse surfaces, Lambert's law approximates the diffuse reflection of light for each spectral component with $R = IFM$, where $F = \cos(\alpha)/r^2$.

In addition to the light that a video projector projects, the environment light is color blended with the surface texture in the same way. Assuming additive color mixing, we can extend Lambert's law to take this into account: $R = EM + IFM$, where E is the environmental light's intensity and color. Environmental light differs from projected light in that the latter can be controlled.

Smart projectors seek to neutralize this natural blending effect by projecting an image (I) in such a way that its blended version on the screen surface appears to observers in its known original colors (R). Given that we consider only diffuse Lambertian screen surfaces—most other surface types are improper for a video projection—we must simply solve the equation for I: $I = (R - EM)/FM$.

Because we do not require information about the

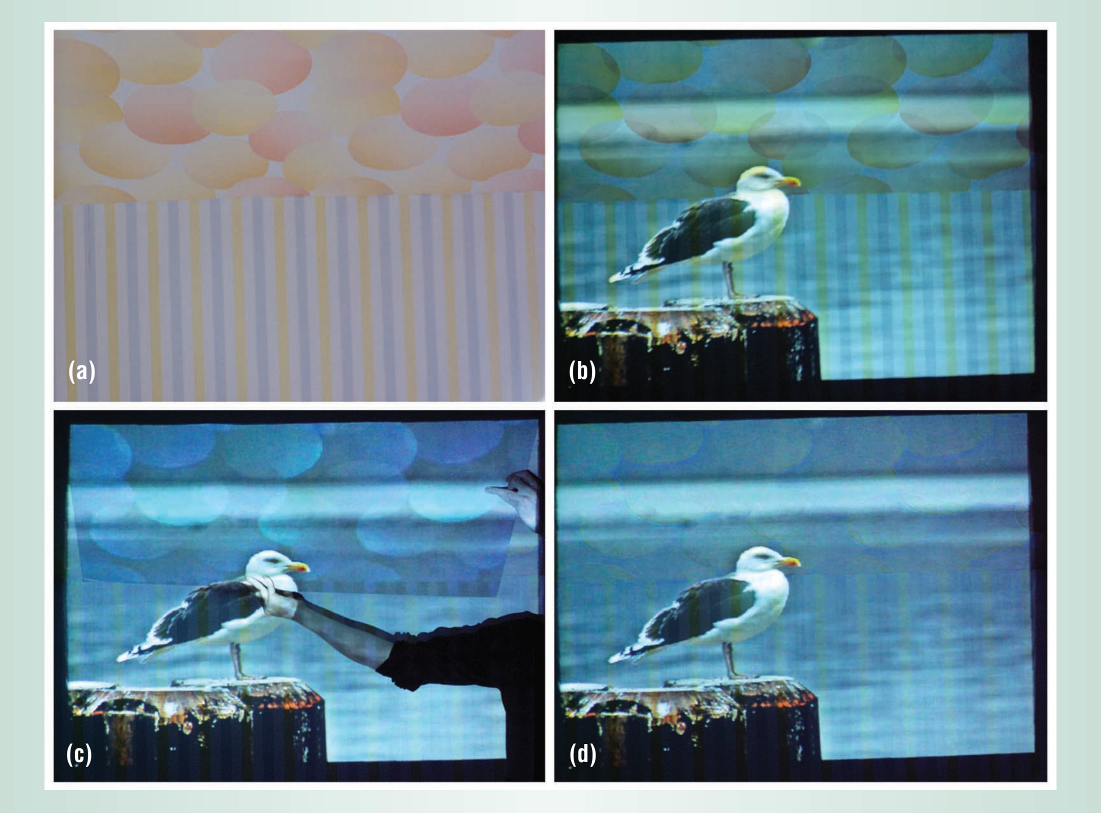


Figure 3. Projection onto a pitched roof area. The image sequence shows (a) the wallpapered surface, (b) the projection with uncorrected colors, (c) color correction projected onto a white piece of paper, and (d) the color-corrected image on wallpaper. All projections are geometry corrected.

projector's or camera's internal and external parameters, we cannot determine each *E*, *M*, and *F* component individually. Rather, we can measure the products *EM* and *FM* while *R* is the given input image—a video frame, for example.

If the video projector displays a bright white image ($I = 1$) onto the screen surface within a dark environment ($E = 0$), the camera captures an image proportional to *FM*. Further, if we turn off the projector ($I = 0$), the screen surface image captured under environmental light is proportional to *EM*. These assumptions imply a color- and intensity-adjusted projector and camera, with automatic brightness control, focus, and white-balancing turned off. These simple approximations let us determine the required parameters robustly, without performing complicated measurements or using additional special-purpose devices.

We can use modern *pixel shader* hardware to perform the final correction computations in real time—no less than 100 fps on an Nvidia GeForce FX6800GT—and represent all these parameters as textures. Today, many consumer graphics cards use pixel shaders as standard components and allow per-pixel operations.

The system warps pixels of images taken from the camera view (*EM* and *FM*), as well as the input image *R*, to the projector view via pixel-displacement mapping. This ensures a correct concatenation of corresponding pixels. The resulting image *I* is finally displayed from the perspective of the projector. These computations are performed on all three RGB color channels separately. In addition, the pixel shader allows fine-tuning of the output images by considering manually set color, brightness, and gamma correction parameters. It also clips out extreme-intensity situations to avoid visible artifacts.

Figure 3 shows an example of a geometry-corrected projection onto a wallpapered, pitched-roof area. If the input image is not color corrected, the projected colors blend with the colors of the screen surface, as Figure 3b shows.

The wallpaper texture interferes with the video image, which results in a disturbing effect. The color-corrected image (*I*) is partially shown in Figure 3c by projecting it onto white cardboard. Blending *I* with the screen surface results in the image shown in Figure 3d, which closely approximates the original input image *R*. In this case, the screen surface becomes almost invisible. All figures show freeze images of movie frames that the smart projector normally corrects continuously, on the fly, during playback.

BEYOND THE MEANS

Obviously, both geometry correction and color correction will fail if the screen surface's material absorbs the light entirely. Failure will also occur if the surface completely absorbs even part of the spectrum that is visible to the camera. Fortunately, absorbent materials such as velvet are comparatively uncommon in everyday environments, and most diffuse materials produce fairly acceptable results. If the surface can reflect a certain fraction

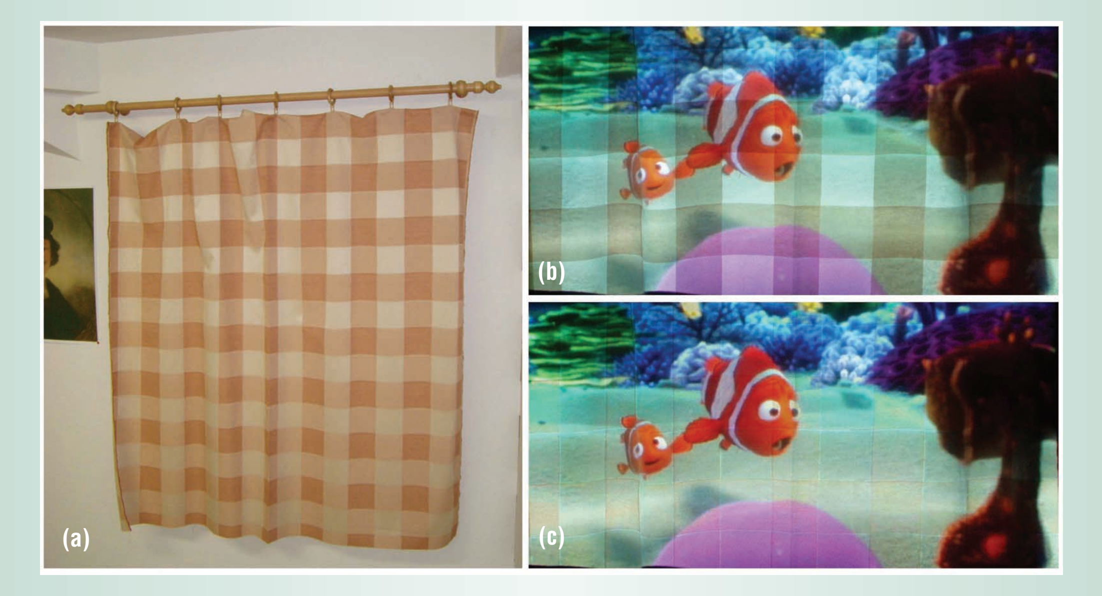


Figure 4. Image projection onto a curtain. (a) The display surface, a checkered curtain, shows the differences between a projected image with (b) uncorrected colors and (c) one with corrected colors. Both projections are geometry corrected. Displayed content: **Finding Nemo**, *Disney/Pixar Animation Studios.*

of the desired light color, the projector only needs to determine how much incident light it must project to produce the desired output. If one projector cannot generate the necessary light, multiple projectors can do so by complementing each other.

Additionally, several technical limitations lower the quality of a current smart projector, such as the limited resolution of consumer camcorders that use PAL or NTSC. If it is necessary to place the camera far away from a large screen surface to capture the entire display area, the projector cannot detect and correct fine surface details. Higher-resolution cameras, such as megapixel digital cameras, can provide better-quality images. In fact, most camcorders already combine two devices in one: a high-resolution digital camera and a video camera that delivers a live video stream. This combination facilitates both fast geometry correction and high-quality color correction.

On the projector side, the limited resolution, low dynamic range, small color space, and high black level of consumer devices represent the main restrictions. A too-low projector resolution causes overly large pixel projections that cannot cover smaller pigments on the screen surface precisely. In particular, the inability to control the black level contributes to the environmental light. Even in a completely dark room, a projector's black level causes the screen surface to be visible. As occurs with a normal projection onto a regular canvas, the black level and the environmental light make displaying dark colors difficult.

However, the human visual system adapts well to local contrast effects. Dark areas surrounded by brighter ones appear much darker than they actually are. Even though researchers will solve these problems in future projector generations, one general problem will remain: Their limited depth focus prevents conventional projectors from displaying images on extremely curved screen surfaces. Because laser projectors, which can focus on nonplanar surfaces, remain far too expensive for the consumer market, using multiple projectors offers a promising solution. The realization of a multifocal projection extension is part of our current research efforts.

One issue remains when projecting onto nonplanar surfaces: A single projector can cast shadows on the screen surface that, from the target perspective, appear as cuttings in the presented output image. However, other projectors that contribute from different directions can cover these shadow areas, as Figure 4 shows.

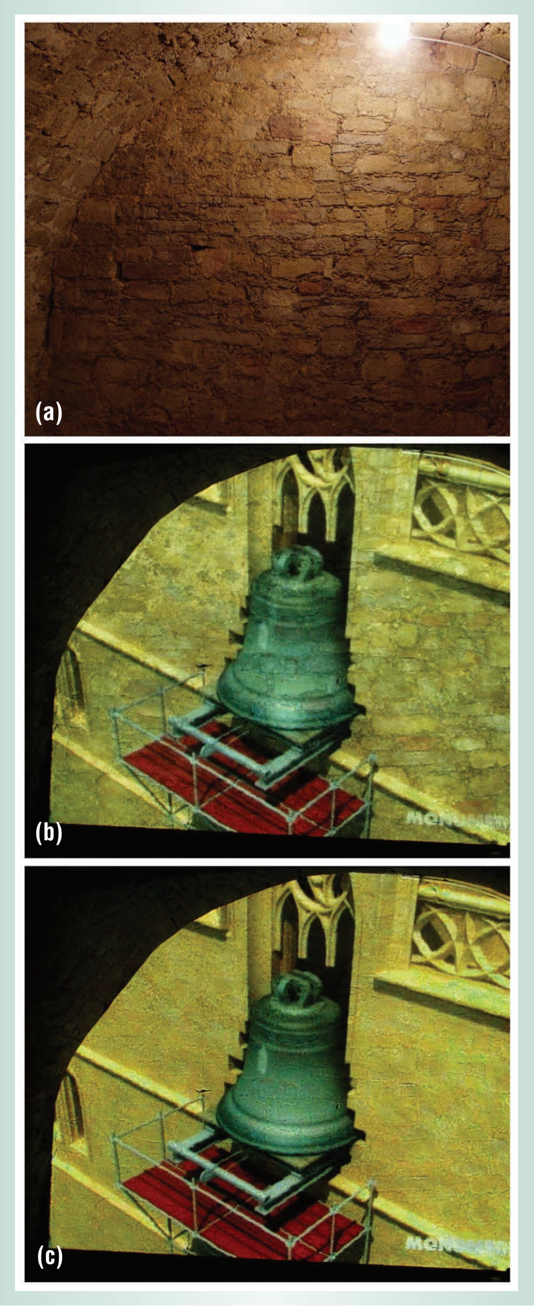
JOINT FORCES

Multiple projectors can enhance the final quality of the output image by complementing each other to achieve increased light intensities and canceling out individual shadow regions. Their output images can fully or partially overlap or can be completely independent. In addition, using multiple projectors allows covering large screen surfaces with high-resolution image tiles. These configurations, known as *tiled screen displays*, provide an overall resolution that a single projector cannot achieve.

In the example in Figure 2, two partially overlapping projectors generate a high-resolution 16:9 format. A smart projector must be scalable so that the end user's calibration effort does not increase with the number of applied projector units. As with the single projector configuration, multiple projectors can be aligned arbitrarily—the calibration process remains fully automatic.

Realizing a second proof-of-concept prototype with two projector units required using a dual-output graphics card to synchronize them. During the geometry calibration, the system generates two dis-

Figure 5. (a) Stenciled projection onto a natural stone wall inside a castle vault. Both projections have undergone geometry correction, while (b) is color uncorrected and (c) is color corrected. Displayed content: The Recovery of Gloriosa, _Bennert-Monumedia GmbH._



placement maps sequentially, one for each projector. Consequently, the graphics card can map pixels from the camera view into each projector's perspective so that all pixels display at exactly the same spot on the screen surface. Thus, for N projectors the individual light intensities add up to $R = EM + I_1F_1M + I_2F_2M + \ldots + I_NF_NM$.

We can achieve a balanced load among all projectors by assuming that $I_i = I_1 = I_2 = \ldots = I_N$. This implies that $R = EM + I_i(F_1M + F_2M + \ldots + F_NM)$, and we can solve for $I_i = (R - EM)/(F_1M + F_2M + \ldots + F_NM)$. This is equivalent to the assumption that a single high-capacity projector produces the total intensity arriving on the screen surface virtually. Physically, however, the intensity is evenly distributed among multiple low-capacity units.

Although each projector sends the same output intensity, the potentially varying form factors cause different fractions to arrive at the screen surface. The smart projector mixes these fractions on the surface, leading to the final result of $R = EM + I_iF_1M + \ldots + I_iF_NM = EM + (R - EM)(F_1M + \ldots + F_NM)/(F_1M + \ldots + F_NM) = R$.

As for a single projector, a pixel shader that receives the parameter textures EM, and $F_1M \ldots F_NM$ computes I_i in real-time. The form factor components F_iM can be determined in two ways:

- by sequentially sending out a white image ($I = 1$) from each projector and capturing each component one by one, or
- by capturing a single image proportional to $F_1M + \ldots + F_NM$ by sending each projector's maximum contribution simultaneously.

Although the second method is conceptually more compact, the first method prevents the system from overmodulating the camera's CCD/CMOS sensor. The system captures shadow regions that individual projectors cause in the form factor components. Consequently, the system cancels out the shadows that individual projector units automatically create as a side effect. This, however, implies that the projectors are placed so that at least one projector can reach each surface portion. The cross-fading techniques common to multiprojector setups can achieve smooth transitions among different contributions.

The total duration for calibration increases linearly with the number of projectors. Thus, the two-projector setup can be geometry- and color-calibrated in less than one minute.

Figure 5 shows that, by using multiple smart projectors, we can project images onto surfaces that are neither plain nor white and need not have a rectangular shape. Instead, we can convert many existing surfaces to a display screen by projecting color- and geometry-corrected images onto them. In the consumer context, this capability offers the advantage of fast, fully automatic, and robust calibration, and it allows the correction of video signals in real time. It isn't necessary to know either geometry information or projector and camera parameters. Instead, the projector system performs the entire calibration and correction on a per-pixel level.

Video projectors will play a major role in future home entertainment and edutainment applications—ranging from movies and television to computer games and multimedia presentations. Smart video projectors have the potential of sensing the environment and adapting to it. This promotes a seamless embedding of display technology into our everyday life.

Future hardware improvements will pave the way for further smart-projector advancements. Upcoming graphics chips, for instance, will be more powerful than ever before. Successive projector generations will continue to feature enhanced quality factors such as brightness, resolution, dynamic range, and black-level. Simultaneously, prices will continue to drop as the smart-projector market share increases. The development of digital cameras and camcorders will follow a similar pattern. Thanks to ongoing miniaturization, all these components could soon be integrated into compact and mobile devices as inconspicuous as light bulbs.

To make immersive 3D visualizations possible within arbitrary environments, we are currently extending the smart-projector concept toward large-scale stereoscopic and multifocal projection. When hardware and software developments have improved, converting a bookshelf into a TV screen or turning a child's entire room into an interactive virtual playground could become possible. ■

Acknowledgments

We thank Gordon Wetzstein and Anselm Grundhöfer for their help on the pixel shader implementation.

References

1. J. Gühring, "Dense 3D Surface Acquisition by Structured Light Using Off-the-Shelf Components," *Proc. Videometrics and Optical Methods for 3D Shape Measuring*, SPIE, 2001, vol. 4309, pp. 220-231.
2. R. Raskar, G. Welch, and H. Fuchs, "Seamless Projection Overlaps Using Image Warping and Intensity Blending," *Proc. 4th Int'l Conf. Virtual Systems and Multimedia*, IEEE Press, 1998, www.cs.unc.edu/~fuchs/publications/VSMM_seamless.pdf.
3. R. Yang et al., "PixelFlex: A Reconfigurable Multi-Projector Display System," *Proc. IEEE Conf. Visualization*, IEEE Press, 2001, pp. 68-75.
4. A. Majumder et al., "Color Calibration of Projectors for Large Tiled Displays," *Proc. IEEE Conf. Visualization*, IEEE Press, 2000, pp. 102-108.
5. R. Raskar et al., "iLamps: Geometrically Aware and Self-Configuring Projectors," *Proc. ACM Siggraph*, ACM Press, 2003, pp. 809-818.
6. R. Raskar et al., "Multiprojector Displays Using Camera-Based Registration," *Proc. IEEE Conf. Visualization*, IEEE Press, 1999, pp. 161-168.
7. O. Bimber et al., "Superimposing Pictorial Artwork with Projected Imagery," to appear in *IEEE MultiMedia*, Jan.-Mar. 2005.
8. R. Sukthankar, T.J. Cham, and G. Sukthankar, "Dynamic Shadow Elimination for Multi-Projector Displays," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 2, IEEE CS Press, 2001, pp. 151-157.
9. R. Sukthankar, R. Stockton, and M. Mullin, "Automatic Keystone Correction for Camera-Assisted Presentation Interfaces," *Proc. Int'l Conf. Multimodal Interfaces;* http://www.ri.cmu.edu/pubs/pub_3396.html.

Oliver Bimber *is a junior professor of Augmented Reality at Bauhaus University, Weimar, Germany. His research interests include display technologies, computer graphics, and human-computer interaction. Bimber received a PhD in engineering from Darmstadt University of Technology, Germany. Contact him at obimber@computer.org.*

Andreas Emmerling *is a graduate student at Bauhaus University. His research interests include projection technology, computer graphics, and computer vision. Contact him at andreas.emmerling@medien.uni-weimar.de.*

Thomas Klemmer *is a graduate student at Bauhaus University. His research interests include augmented reality, robotics, and real-time rendering. Contact him at thomas.klemmer@medien.uni-weimar.de.*