Projector-Based Augmentation

Oliver Bimber
Bauhaus-University Weimar
Bauhausstr.11, 99423 Weimar, Germany
Phone: +49-(0)3643-583724, Fax: +49-(0)3643-583709
Email: obimber@computer.org

1. INTRODUCTION

Their increasing capabilities and declining cost make video projectors widespread and established presentation tools. Being able to generate images that are larger than the actual display device virtually anywhere is an interesting feature for many applications that cannot be provided by desktop screens. Several research groups discover this potential by applying projectors in unconventional ways to develop new and innovative information displays that go beyond simple screen presentations.

Projector-based displays have clearly replaced head-attached displays for most virtual reality (VR) applications. Immersive surround screen displays and semi-immersive wall-like or table-like configurations are being used for visualizing two-dimensional or three-dimensional graphical content.

Today, the majority of augmented reality applications focuses on mobility. Thus wearable or portable devices have become dominant in this area. However, an increasing trend towards projector-based displays for AR can be noticed. Projector-based augmentation approaches hold the potential of combining the advantages of well-established spatial virtual reality and spatial augmented reality (Bimber, 2005d). Immersive, semi-immersive and augmented visualizations can be realized in everyday environments – without the need for special projection screens and dedicated display configurations. Limitations of mobile devices, such as low resolution and small field of view, focus constrains, and ergonomic issues can be overcome by the application of projection technology. For many applications this requires the abdication of mobility – but not necessarily of portability. Several applications, however, do not require mobility and rather benefit from efficient spatial augmentations. Examples range from edutainment in museums (such as storytelling projections onto natural stone walls in historical buildings) to architectural applications (such as augmentations of complex illumination or surface material simulations in real building structures). The problems, limitations, potentials and details of a variety of existing techniques towards projector based augmentations are described in this chapter.

A variety of stationary, movable, and hand-held projectors has been proposed for displaying graphical information directly on real objects or surfaces instead of performing optical overlays or video compositions.

The Luminous Room (Underkoffler, 1999) for instance, describes an early concept for providing graphical display and interaction on each surface of an interior architecture space. Co-located two-way optical transducers –called I/O bulbs– that consist of projector-camera pairs capture the user interactions and display the corresponding output. With the Everywhere Displays projector (Pinhanez, 2001) this concept has been extended technically by allowing a steerable projection using a pan/tilt mirror. A similar approach is followed by Ehnes, et al. (Ehnes, 2004). Recently, it was demonstrated how context-aware hand-held projectors –so-called iLamps– can be used as mobile information displays and interaction devices (Raskar, 2003).

Another concept called Shader Lamps (Raskar, 2001) attempts to lift the visual properties of neutral diffuse objects that serve as projection screen. The computed radiance at a point of a

non-trivial physical surface is mimicked by changing the bidirectional reflectance distribution function and illuminating the point appropriately with projector pixels. Animating the projected images allows creating the perception of motion without physical displacement of the real object (Raskar, 2002). This type of spatial augmentation is also possible for large, human-sized environments, as demonstrated in (Low, 2001).

Projector-based illumination has become an effective technique in augmented reality to achieve consistent occlusion (Noda, 1999; Bimber, 2002a) and illumination (Bimber, 2003) effects between real artifacts and optically overlaid graphics. Video projectors instead of simple analog light bulbs are used to illuminate physical objects with arbitrary diffuse reflectance. The per-pixel illumination is controllable and can be synchronized with the rendering of the graphical overlays. This also makes the combination of high-quality optical holograms with interactive graphical elements possible (Bimber, 2004a). Using a video projector to produce a controlled reference wave allows reconstructing the hologram's object wave partially – not at those portions that are overlaid by integrated graphical elements.

New optical combiners together with real-time radiometric compensation methods allow superimposing arbitrarily textured, flat surfaces, such as paintings efficiently (Bimber, 2005a).

Other methods allow augmenting arbitrary –flat (Yoshida, 2003; Nayar, 2003; Wang, 2005; Fujii, 2005) or geometrically non-trivial (Grossberg, 2004; Bimber, 2005b)– textured surfaces without the application of special optical combiners, such as transparent projection screens. They scan the surfaces' geometric and reflectance properties and carry out a per-pixel displacement and radiometric compensation before the images are projected.

Beside these examples, a variety of other techniques and applications have been described that utilize projectors and a structured illumination to achieve special effects and augmentations. To describe them all is out of the scope of this chapter. More details can be found in the book *Spatial Augmented Reality: Merging Real and Virtual Worlds* (Bimber, 2005d).

2. FUNDAMENTAL PROBLEMS AND OVERVIEW OF SOLUTIONS

For conventional projection systems, the screen material is optimized for a projection, and its reflectance is uniform across the surface. If images are projected onto everyday surfaces, however, the following fundamental problems arise (cf. figure 1-left):

1. If the surface is not planar, the projected images are geometrically warped and appear distorted to an observer;
2. If the surface is not white but has a colored texture, the projected light is blended with the reflecting surfaces' pigments;
3. Diffused light is scattered from one surface portion to others and is blended together with the direct illumination;
4. Conventional projectors can focus on planar surfaces only. Projections onto geometrically complex surfaces cause a regional defocus in the projected images;
5. Even if problems 1-4 can be avoided by pre-correcting the images before displaying them, slight misregistrations of projected pixels and the corresponding surface pigments can lead to extreme visual artifacts;
6. Multiple projectors can have varying chrominance and luminance parameters, which leads to inconsistent image contributions;
7. If the surface is not Lambertian (i.e., not perfectly diffuse), a projector-based augmentation might not be possible.

These points represent particular problems, if the projected images contain stereo pairs to enable a stereoscopic 3D visualization on screen surfaces that are not optimized for

projection. The reason for this is that the human visual system strongly relies on the extraction of salient structure features (such as edges, corners, etc.) for estimating disparities. Several of the image distortions described above simply wash out these features (4), blend these features with physical features on the surface (2 and 5) or with scattered light from other surface portions (3), or misalign them (1). In the following it is assumed that the projection surface is Lambertian to enable a projector-based augmentation. It is out of the scope of this chapter to describe photometric calibration techniques (such as chrominance mapping and luminance matching) and other multi-projector methods, such as cross-fading, or shadow removal. The interested reader is referred to (Brown, 2005) for an overview of this area. This chapter does also not describe other important issues related to projector-based augmentation, such as projector hardware (electronic and optics), tracking and registration, networked rendering frameworks, projector-based interaction techniques and devices, etc.

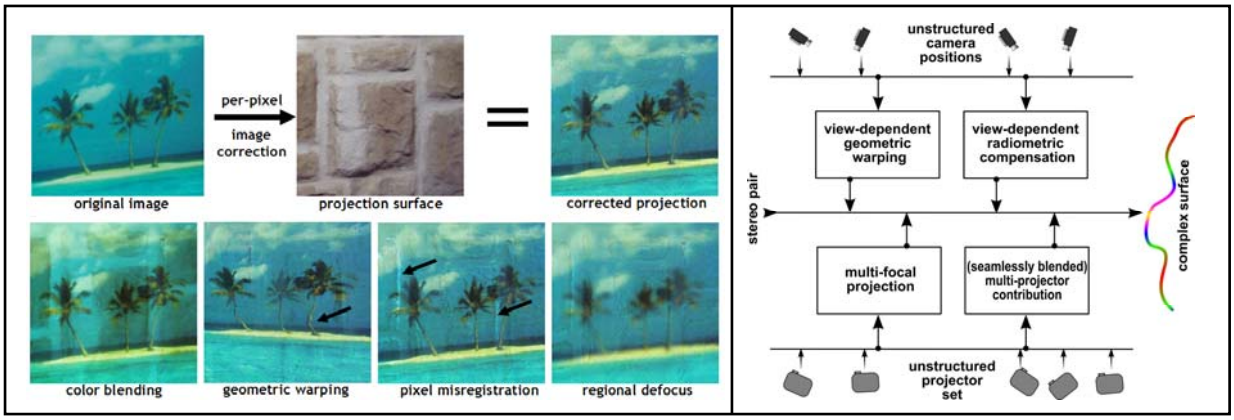

Figure 1: Four exemplary problems resulting from a conventional projection onto a complex surface, and corrected projection (left). Main components for enabling consistent stereoscopic projections on everyday (Lambertian) surfaces (right).

Instead, this chapter presents a self-contained framework of basic rendering techniques that strive for a projector-based augmentation of everyday environments and surfaces (cf. figure 1-right). The corpus of this chapter describes projector-camera methods and multi-projector techniques that aim at correcting geometry (sections 3 and 4), local and global radiometric effects (section 4), and focus (section 5) properties of projected images to minimize the distortions described above. Based on these fundamental techniques, further methods are explained that enable the view-dependent correction of projected images (section 6). The rendering techniques described in this chapter can all be implemented as hardware-accelerated pixel shaders, and consequently support a pixel-precise correction at interactive frame-rates. Thus, the four modules shown in figure 1-right are projector-individual pixel shaders. Section 7 finally gives a brief overview over current limitations and an outlook of future developments.

3. PROJECTING ONTO OPTIMIZED SURFACES

For surfaces whose reflectance is optimized for projection (i.e., surfaces with a homogenous white reflectance), a geometric correction of the projected images is sufficient to provide an undistorted augmentation for an observer. Slight misregistrations of the images on the surfaces in the order of 2-3 pixels lead to geometric artifacts that -in most cases- can be tolerated. This section gives a brief overview over general geometry correction techniques for single and multiple projectors.

If multiple projectors (P) have to be registered to a planar surface via camera feedback (cf. figure 2-left), collineations with the plane surface can be expressed as camera-to-projector homographies (H). A homography matrix can be automatically determined by correlating a

projection pattern to its corresponding camera image. The homographies are usually extended to homogenous 4x4 matrices to make them compatible with conventional transformation pipelines and consequently benefit from single pass rendering. Multiplied after the projection transformation, they map normalized camera coordinates into normalized projector coordinates. An observer located at the position of the calibration camera (*C*) perceives a correct image in this case. In cases where a head-tracked observer has to be supported the calibration camera can be placed orthogonal to the surface: The virtual scene is rendered from the perspective of the observer with an off-axis projection over the rectangular portion of the plane surface that is visible to the camera. The resulting image is then warped to the perspectives of the individual projectors via their camera-to-projector homographies[1]. This can all be done within a single rendering pass.

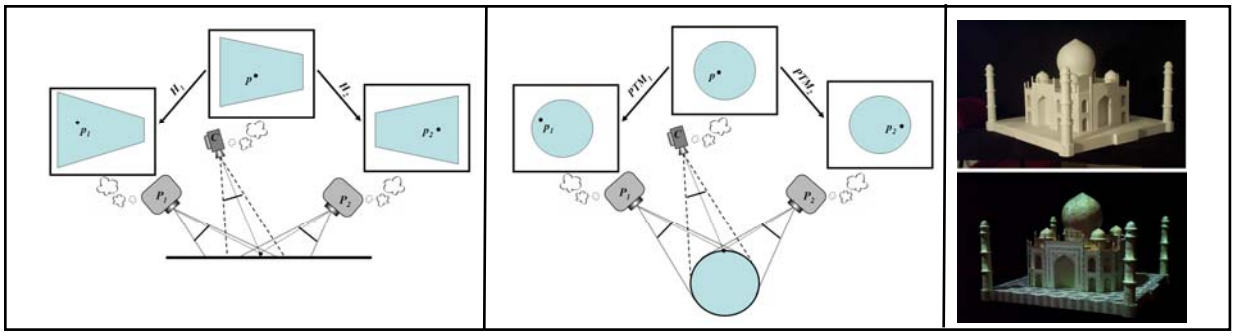

Figure 2: Geometric correction via homographies (left) and projective texture mapping (center). Example for using projective texture mapping (right). Right image reprinted from (Raskar, 2001) © Springer-Verlag.

If the geometry of the projection surface is non-planar but known, a two-pass rendering technique can be applied for projecting the image in an undistorted way[2] (Raskar, 1999b): In the first pass, the image that has to be displayed is off-screen rendered from the perspective of the observer (*C*). This image *O* is then read back into the texture memory. In the second step, the geometry model of the display surface is texture-mapped with *O* while being rendered from the perspective of the projector. For computing the correct texture coordinates that ensure an undistorted view from the perspective of the observer projective texture mapping is applied (cf. figure 2-center). This hardware accelerated technique dynamically computes a texture matrix that maps the 3D vertices (p_i) of the surface model from the perspectives of the projectors (P_i) into the texture space of the observer's perspective (*C*). However, projective texture mapping assumes a simple pinhole camera/projector model and does not take the lens distortion of projectors into account. This can cause misregistrations of the projected images in the range of several pixels – even if other intrinsic and extrinsic parameters have been determined precisely. Projecting corrected images onto textured surfaces with misregistrations in this order causes immediate visual intensity artifacts that can make the fusion of projected stereo-pairs difficult. In section 4, two pixel-precise registration techniques are described. The first method is also being applied by related approaches (Nayar, 2003; Grossberg, 2004). In section 6 this method is extended towards view-dependent rendering. Note, that some of the techniques described below (e.g., compensating secondary scattering and multi-focal projection) are also relevant for projection optimized surfaces.

[1] Consistent depth values have to be ensured (Raskar, 1999a).
[2] The intrinsic and extrinsic parameters of the projectors have to be determined first.

4. PROJECTING ONTO COMPLEX SURFACES

For projections onto colored or textured surfaces the images have to be color corrected in addition to a geometry correction. Recent work on radiometric compensation uses cameras in combination with projectors for measuring the surface reflectance as well as the contribution of the environment light. These parameters are then used for correcting the projected images in such a way that blending artifacts with the underlying surface are minimized.

Nayar et al. (Nayar, 2003), for instance, express the color transform between each camera and projector pixel as pixel-individual 3x3 color mixing matrices. These matrices are estimated from measured camera responses of multiple projected sample images. They can be continuously refined over a closed feedback loop and are used to correct each pixel during runtime. Later, a refined version of this technique was used for controlling the appearance of two- and three-dimensional objects, such as posters, boxes and spheres (Grossberg, 2004). Wang et al. (Wang, 2005) adapt this method to the properties of the human vision system by compressing the contrast of the input images. Fujii et al. (Fujii, 2005) applies a variation of the closed feedback loop method to handle dynamic environments by applying a co-axial projector-camera alignment.

This section begins with a simple implementation of geometric and radiometric image correction (Bimber, 2005a). In contrast to (Nayar, 2003; Grossberg, 2004; Wang, 2005; Fujii, 2005) it uses single disjoint camera measurements of surface reflectance, environment light contribution and projector form-factor components for a per-pixel radiometric compensation using hardware accelerated pixel shaders. But it does not take the color mixing of the individual RGB channels into account. Furthermore, it is described how radiometric compensation can be enhanced by using multiple interplaying projectors for geometrically complex surfaces, or by applying additional transparent film materials overlaid over planar surfaces. Finally, it is explained how to compensate global illumination effects, such as scattering in addition to local ones.

4.1 Radiometric Compensation With A Single-Projector

In its simplest configuration, an image is displayed by a single projector (*P*) in such a way that it appears correct (color and geometry) for a single camera view (*C*). Thereby, the display surfaces must be Lambertian, but can have an arbitrary color, texture and shape.

The first step is to determine the geometric relations of camera pixels and projector pixels over the display surface. Well known structured light techniques (e.g., gray code scanning with phase shift) can be used for measuring the 1-to-n mapping of camera pixels to projector pixels (cf. figure 3-left). This mapping is stored in a 2D look-up-texture having a resolution of the camera, which in the following is referred to as *C2P* map. A corresponding texture that maps every projector pixel to one or many camera pixels can be computed by reversing the *C2P* map. This texture is called *P2C* map. It has the resolution of the projector. The 1-to-n relations (note that *n* can also become *0* during the reversion process) are finally removed from both maps through averaging and interpolation.

Once the geometric relations are known, the radiometric parameters are measured. It can be assumed that a light ray with intensity *I* is projected onto a surface pigment with reflectance *M*. The fraction of light that arrives at the pigment depends on the geometric relation between the light source and the surface. A simple representation of what is known as form factor can be used for approximating this fraction: $F=f*\cos(\alpha)/r^2$, where α is the angular correlation between the light ray and the surface normal and *r* the distance (square distance attenuation) between the light source and the surface. The factor *f* allows scaling the intensity to avoid clipping (i.e., intensity values that exceed the luminance capabilities of the projector) and to consider the simultaneous contributions of multiple projectors.

Together with the environment light *E*, the projected light fraction *I* is blended with the pigment's reflectance *M* (cf. figure 3-center):

$$R = EM + IFM \quad (1)$$

Thereby R is the diffuse radiance that can be captured by the camera. If R, F, M, and E are known, I can be computed with:

$$I = (R - EM)/FM \quad (2)$$

In the single-projector approach, E, F, and M cannot be determined independently. Instead, FM is measured by projecting a white image ($I=1$) and turning off the entire environment light ($E=0$), and EM is measured by projecting a black image ($I=0$) under environment light. Note that EM also contains the black-level of the projector.
Since this holds for every discrete camera pixel, R, E, FM and EM are entire textures and equation 2 can be computed[3] in real-time by a pixel shader.
During runtime pixel displacement mapping is realized by rendering a full-screen quad into the frame buffer of the projector. This triggers fragment processing of every projector pixel. A fragment shader maps all pixels from the projector perspective into the camera perspective (via texture look-ups in the *P2C* map) to ensure a geometric consistency for the camera view. All computations are then performed in camera space. The projection of the resulting image I onto the surface leads to a geometry and color corrected image that approximates the desired image R for the target perspective of the camera.

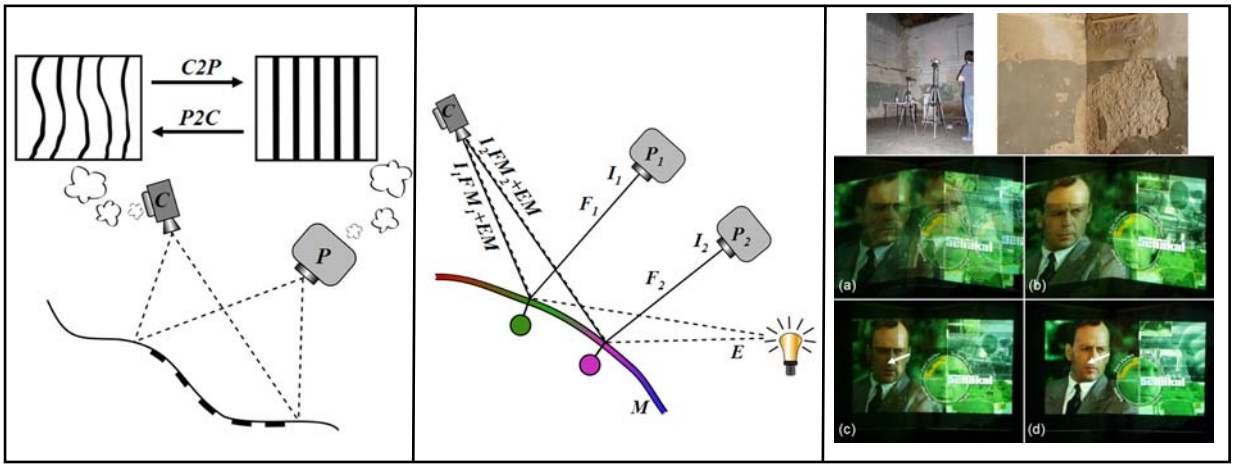

Figure 3: Pixel-precise geometric correction (left) and radiometric compensation (center). Example of a projection onto a scruffy room corner (right): (a) two projectors uncorrected, (b) registered projectors, (c) geometry corrected projections, (d) radiometric compensated projections. Right image reprinted from (Bimber, 2005c) © IEEE.

4.2 Radiometric Compensation With Multi-Projectors

The simultaneous contribution of multiple projectors increases the total light intensity that arrives at the surface. This can overcome the limitations of equation 2 for extreme situations (e.g., small FM values or large EM values) and can consequently avoid an early clipping of I.
If N projectors are applied, equation 1 extends to (cf. figure 3-center):

$$R = EM + \sum_i^N I_i FM_i \quad (3)$$

[3] For each color channel separately.

One strategy is to balance the projected intensities equally among all projectors *i* which leads to:

$$I_i = \frac{(R - EM)}{\sum_j^N FM_j} \tag{4}$$

This equation can also be solved in real-time by projector-individual pixel shaders (based on individual parameter textures FM_i, $C2P_i$ and $P2C_i$ – but striving for the same final result *R*). Note that *EM* also contains the accumulated black-level of all projectors.
If all projectors provide linear transfer functions (e.g., after a linearization) and identical brightness, a scaling of $f_i=1/N$ used in the form factor balances the load among them equally. However, f_i might be decreased further to avoid clipping and to adapt for differently aged bulbs.

4.3 Amplification Through Transparent Projection Screens

A clipping of *I* can also be minimized if the projection surface is coated with a transparent projection screen (e.g., a flexible transparent film material that diffuses a portion *d* and transmits a fraction *t* of the light projected onto it). This is a practical option for augmenting pictorial artwork, such as paintings or sketches (Bimber, 2005c).

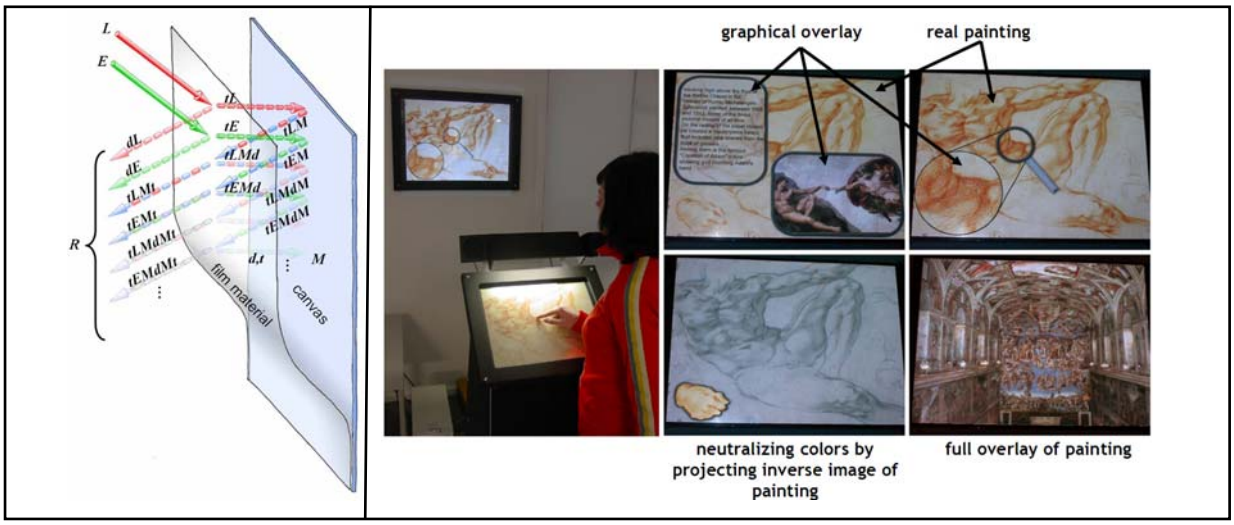

Figure 4: Interaction of projected light and environment light with the screen canvas and the transparent projection film - sequence diagram (right). Examples of interactive projections onto real painting (right). Images reprinted from (Bimber, 2005a) © IEEE.

If a light beam with incident radiance *L* is projected onto the transparent film material that is located on top of the surface (e.g. the canvas of a painting), a portion *d* of *L* is directly diffused from the film while the remaining portion *t* of *L* is transmitted through the film. The transmitted light *tL* interacts with the underlying pigment's diffuse reflectance *M* on the canvas, and a color blended light fraction *tLM* is diffused. The portion *tLMt* is then transmitted through the film, while the remaining part *tLMd* is reflected back towards the canvas where it is color blended and diffused from the same pigment again. This ping-pong effect between film material and canvas is repeated infinitely while for every pass a continuously decreasing amount of light is transmitted through the film that contributes to the resulting radiance *R*. Mathematically, this can be expressed as an infinite geometric series that converges towards a finite value. The same is true for the environment light with incident radiance *E* that is emitted from uncontrollable light sources. Since these light sources also

illuminate the canvas and the film material, the environment light's contribution to *R* has to be considered as well.
Figure 4 describes this process in form of a sequence diagram. Note that in contrast to this conceptual illustration, there should be no physical gap between film material and canvas, and that the light interaction occurs at the same spot.
If all parameters (*L*, *E*, *M*, *t*, and *d*) are known the resulting radiance *R* that is visible to an observer in front of the canvas can be computed:

$$R = (L+E)d + (L+E)t^2 M \sum_{i=0}^{\infty} (Md)^i = (L+E)\left(d + \frac{t^2 M}{1-Md}\right) \quad (5)$$

Since *R* (image that is expected to be seen) is known, the above equation needs to be solved for *L*:

$$L = \frac{R}{\left(d + \frac{t^2 M}{1-Md}\right)} - E \quad (6)$$

This allows computing the incident radiance *L* that needs to be projected onto the film and the canvas to create the known result *R*. The radiant intensity *I* of the projector that is required to create *L* is related to a discretized pixel value and is given by:

$$I = L\frac{1}{F}f \quad (7)$$

where $F=f*\cos\alpha/r^2$ is the form factor of the projector (see section 4.1).
As described earlier, the contribution of multiple (*N*) projectors allows minimizing the clipping problem with:

$$L = \sum_{i=1}^{N} L_i \quad , \quad I_i = L_i \frac{1}{F_i} \quad (8)$$

4.4 Compensating Secondary Scattering Through Reverse Radiosity
The techniques described above can compensate only radiometric effects that result from a direct illumination of the surface. However, diffuse surface portions scatter a fraction of light to other surface portions. This amount of indirect illumination adds to the direct illumination and has to be compensated as well (Bimber, 2006b) (cf. figure 5). Problematic is the fact that the amount of scattering depends on the projected compensated image (*I*) and vice versa. In computer graphics, the reverse situation (i.e., the global radiosity of each surface element) is determined based on a known illumination (*I*) and is computed by numerically solving a linear equation system. Some modern approaches implement this finite element model –which is commonly known as radiosity rendering– with pixel shaders and achieve interactive frame rates.

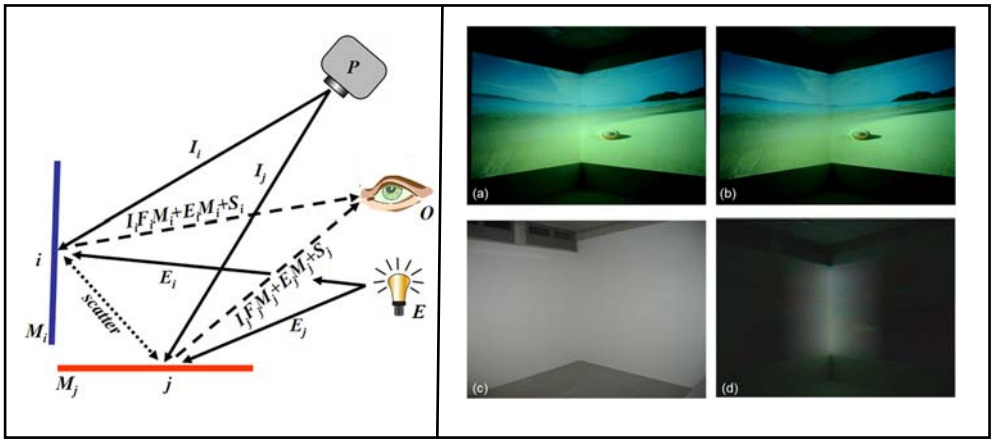

Figure 5: Scattering of light between two surface patches (left). Example of compensated scattering (right): (a) uncorrected projection onto two-sided projection screen (c), (b) corrected projection, (d) corrected amount of scattering (subtraction of images a and b). Images reprinted from (Bimber, 2006b) © IEEE

Assume that the surface is subdivided into A discrete patches (e.g., in camera space or in parametric space). The form factor F_{ji} describes the amount of radiance that is scattered from patch j to patch i. This can either be measured or be computed (if the surface geometry is known). The amount of radiance (S_i) that is reflected by patch i which is a result of indirect scattering of light from all other patches can be described as:

$$S_i = \sum_{j,j\neq i}^{A}\left(I_jF_jM_j + E_jM_j\right)F_{ji}M_i + \sum_{j,j\neq i}^{A}\sum_{k,k\neq j}^{A}\left[\left(I_kF_kM_k + E_kM_k\right)F_{kj}M_j\right]F_{ji}M_i + \ldots \qquad (9)$$

Note, that equation 9 presents only the first two scattering levels to a patch i (i.e., the radiance that is scattered directly from the patches j to patch i, and the radiance that is scattered from all patches k over all patches j to patch i). This can be rewritten into a recursive from:

$$S_i^1 = \sum_{j,j\neq i}^{A}\left(I_jF_jM_j + E_jM_j\right)F_{ji}M_i \quad , \quad S_i^h = \sum_{j,j\neq i}^{A}S_j^{h-1}F_{ji}M_i \qquad (10)$$

Note, that S_i^h indicates the h-th scattering level at patch i (i.e., the reflected radiance at patch i that can be contributed to scattering which arrives over h-1 intermediate patches). To compensate the direct illumination and the scattering, the following computation has to be performed:

$$I_i = \frac{1}{F_iM_i}\left(R_i - E_iM_i - \sum_{h=1}^{H}S_i^h\right) \qquad (11)$$

As mentioned above, the difficulty in this case is that I_i depends on S_i^h and vice versa. This linear equation system can be solved numerically by approximating I_i through several iterations:
Initially, I_i is computed for the direct case (equation 2). The result is used in the second iteration to compute a first estimate of scattering, and to compensate the direct and the

approximated indirect illumination (equation 11). The result of this iteration is then used in a third iteration, and so on. This process is repeated until I_i converges (i.e., the results of two consecutive iterations do not reveal significant differences).
While equations 10 and 11 can be implemented as pixel shaders, the iterations can be realized as individual rendering passes. While frame-buffer objects offer an efficient way of exchanging the resulting textures I_i between different rendering passes, occlusion queries support the hardware accelerated comparison of two consecutive results I_i. Both techniques avoid the time consuming read-back of textures from the graphics card. The patches' form factors F_{ji} are efficiently compressed into a single parameter texture to keep the memory requirements on the graphics card at a minimum. This can be done offline. Note that only patches with form factors larger than a predefined threshold are selected. This leads to additional resources and performance optimizations and does not affect the outcome's quality significantly. Thus, the form-factor matrix F is a sparse matrix.
A more efficient analytical solution to equations 10 and 11 exists: It was recently shown by Seitz et al. (Seitz, 2005) that global interreflections can be removed from photographs of unknown scenes under unknown illuminations by applying an interreflection cancellation operator. It can also be shown that equation 10 equals a geometric series that converges when the number of scatter levels h approaches infinity. Thus the analytical complement to equation 11 is:

$$I_i = \frac{1}{F_i M_i}\left(R_i - E_i M_i - \sum_{j, j \neq i}^{A} R_j F_{ji} M_j \right) \tag{12}$$

This is similar to Seitz's approach – but with a main difference: The interreflection cancellation operator is not applied to an image that contains interreflections to remove them, but to an image that does not contain interreflections to compensate them when projecting the compensated image. This also proofs that the amount of indirect scattering which is produced by the compensated image throughout all scatter levels equals the amount of indirect scattering that is produced by the original image in the first level!
It is clear that equations 10-12 cannot be computed in real-time if A equals the resolution of the projector (or a camera with similar resolution). Consequently, all parameter textures have to be down-sampled to an acceptable patch resolution (e.g., A=128x128). The final compensated image can then be computed as follows:

$$I'_B = I_B - \uparrow\left(\downarrow I_B - I'_A\right) \tag{13}$$

Thereby, I_B is the compensated image for direct illumination (computed as described in equations 2 or 4) in the projector resolution (B), I'_A is the compensated image in patch resolution A (solved as described above), and $\uparrow$ and $\downarrow$ indicate up- and down-sampling image operators (from B to A and vice versa). The result I'_B contains the compensation of the direct illumination and the scattering, and is finally being displayed.
The extensions of equations 10-12 to support multi-projector configurations based on the balancing strategy described in section 4.2 are simply derived by replacing the single projector form factors by the sum of form factors of all projectors (see equations 3 and 4). As described above, the projected intensities are then balanced equally among all projectors. Thus I_i is the same for each projector, and can easily be computed.

5. MULTI-FOCAL PROJECTION

Today's consumer projectors are designed and engineered to focus images on planar display surfaces. The Schleimpflug principle describes how to offset the focal plane by an off-axis

configuration of the optical system. However, plane-focused images are partially blurred if projected onto surfaces with substantial depth differences. Special lenses, such as f-theta lenses, allow generating focused images on spherical surfaces. Planetariums and some cylindrical projection displays (Biehling, 2004) apply laser projectors to overcome this problem. Direct-writing-scanning-laser-beam projectors scan almost parallel beams of laser light onto the projection screen. Thereby, the laser beams remain constant in diameter over a substantial depth range. This results in a large focal depth and in the possibility to display sharp images in large dome-like or cylindrical theatres. The cost of a single laser projector, however, can quickly exceed the cost of several hundred conventional projectors. But the development of low-cost laser-diodes is promising and can overcome these drawbacks in future.

If the projection surface is multi-planar, multiple projectors can be arranged in such a way that they focus on individual planar sections (e.g. Low, 2001). This, however, becomes inefficient and sometimes impossible the more complex the surface becomes.

In this section a multi-focal projection technique (Bimber, 2006a) is described that projects images with minimal defocus onto geometrically and radiometric complex surfaces. This is essential to enable, for instance, stereoscopic projections supporting disparity-based depth perception on arbitrary surfaces (Bimber, 2005b).

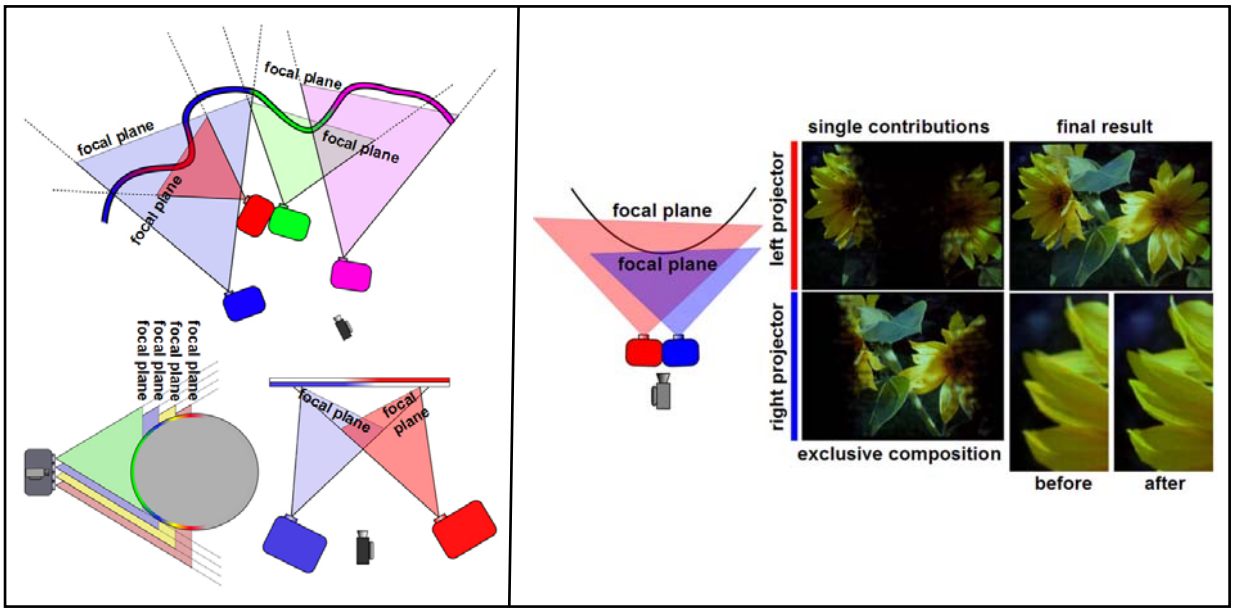

Figure 6: Multi-focal projection concept (left): unstructured projector set, multiple projection units integrated in one projector, and two projectors projecting on plane. Example of multi-focal projection onto a curved screen (right). Images reprinted from (Bimber, 2005b) © IEEE.

The multi-projector technique works as follows: Multiple conventional projectors with differently adjusted focal planes, but overlapping image areas are used (cf. figure 6). They can be either arbitrarily positioned in the environment, or can be integrated into a single projection unit. The defocus created on the surface is estimated automatically for each projector pixel via camera feedback and structured light projection. If this is known, a final image with minimal defocus can be composed in real-time from individual pixel contributions of all projectors.

The technique is independent of the surfaces' geometry, color and texture, the environment light as well as of the projectors' position, orientation, luminance and chrominance. In the following subsections it is explained how intensity spreads are measured, how relative focus values are estimated from the measured intensity spreads, and how the final images are composed from the simultaneous contributions of multiple projectors.

5.1 Measuring Intensity Spreads

The goal of a multi-focal projection is to estimate relative focus values caused by each projector pixel on all portions of an arbitrary display surface. Having this information, a final image with minimal defocus can be composed from multiple projector contributions.

This sub-section describes the approach of measuring the intensity spreads of projected sample points. They are proportional to the defocus of the sampling projector on the corresponding surface portion.

Like for the geometric and radiometric image correction (section 4), a structured light projection is applied in combination with camera feedback to estimate the focus values (cf. figure 7-left). However, instead of displaying horizontal and vertical scan lines, a uniform grid pattern of circular sample points is used. Applying points instead of strips also allows determining the defocus parameters in multiple directions simultaneously.

Displaying a sample point from the projector's view leads to a color and geometry distorted image of it in the camera's view. This situation makes it impossible to estimate focus values for the corresponding surface area. The reason for this is that it is not detectable in this case whether the recorded intensity spreads of the sample point can be contributed to defocus or to external factors (e.g., geometric or radiometric distortion, or blending with the environment light on the surface, the projector's position or orientation).

To overcome this problem, the surface is sampled from the perspective of the camera instead of from the perspective of the projector. Measuring relative focus values within the same space and under the same conditions, enables a qualitative comparison between individual projector contributions. Thus, a sample point is initially defined in the camera perspective and is then geometrically warped into the perspective of the projector with the beforehand determined *C2P* map (section 3.2). In addition, the sample point is color corrected with equation 2 to compensate the surface reflectance, the environment light, and the projector's form factor contribution. Thereby, $R>0$ for pixels that belong to the sample point and $R=0$ for all other pixels.

If a perfectly sharp image is assumed (perfectly focused by the projector and by the camera) the image recorded by the camera contains the initially defined and undistorted sample point. In this case, it has retained its original circular shape in the defined size and appears in a uniform intensity that approximates the defined intensity and color R.

However, due to blur effects (caused either by the projector or by the camera) the shape and the intensity of the sample point are no longer uniform. The resulting intensity spread and intensity loss can be measured in the camera image. They are proportional to the relative defocus of the projector at this point. Measuring the defocus caused by different projectors at the same sample point enables a qualitative and relative comparison and finally an optimal image composition. The camera's parameters must not be changed during the measurements for multiple projectors that need to be compared.

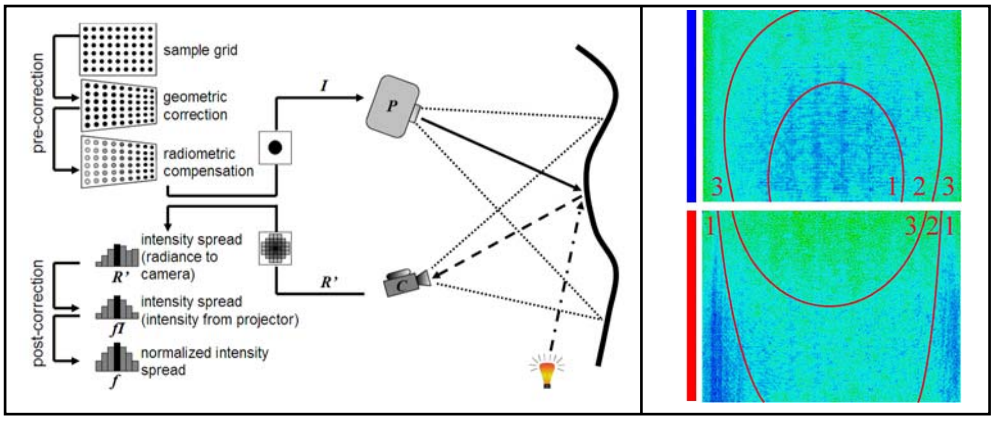

Figure 7: Measuring intensity spreads on geometric and radiometric complex surfaces (left). Example of color coded focus estimations for the example shown in figure 6-right (right): blue to green = best to worst focus. Images reprinted from (Bimber, 2006a) © IEEE.

5.2 Relative Focus-Estimation

As explained above, a defocused sample point creates an intensity spread on the surface that is captured in the camera image. The intensities in the blurred area are not projected in a controlled way. They are also blended with the underlying surface reflectance. To estimate the focus values of a sample point consistently, the intensity spread has to be normalized in such a way that it becomes independent of the surface reflectance, the environment light and the projectors' form factor contributions at the spread areas. Note that this normalization cannot be carried out during the pre-correction step because the intensity spread has to be measured first.

The pre-correction applies equation 2 to compute and project the corrected sample point in such a way that it appears at coordinate x,y in the camera image:

$$I_{x,y} = (R_{x,y} - EM_{x,y}) / FM_{x,y} \tag{14}$$

As outlined above, $R_{x,y}>0$ for pixels that belong to the sample point, and $R_{x,y}=0$ for all other pixels.

The intensity spread $R'_{x',y'}$ in the blurred area x',y' is measured by the camera as well. It results from defocus and is a fraction f of the sample original point's intensity that is blended with the surface reflectance in the blur area:

$$R'_{x',y'} = fI_{x,y} FM_{x',y'} + EM_{x',y'} \tag{15}$$

To normalize the intensity spread, f has to be determined. For this, the intensity the projector would have to produce in terms of creating the same radiance $R'_{x',y'}$ by a direct illumination has to be estimated:

$$fI_{x,y} = \frac{(R'_{x',y'} - EM_{x',y'})}{FM_{x',y'}} \tag{16}$$

Finally, f can be computed by comparing the results of equations 14 and 16:

$$f = \frac{fI_{x,y}}{I_{x,y}} \tag{17}$$

If the pixels in the captured camera image are normalized, a normalized intensity distribution of the sample point is received. This is independent of the surface reflectance of the projector's form factor. Due to limitations of the projectors (brightness and dynamic range) and the camera (response and noise) their maxima of the intensity spreads might not equal one. A second normalization that lifts their maxima intensities to one (and the other intensities accordingly) enhances the quality of focus estimation that does not analyze the intensity loss of the spread.

Each sample point in the camera view corresponds to a distinct area on the display surface. The focus of every projector is estimated within the same surface area (i.e., the same sample point in the camera view). To do this for the entire surface, multiple sample points are shifted within two-dimensional scan windows in the entire camera view. A pixel-by-pixel shift results in exact focus estimations for every single camera pixel. Thereby the projector-individual focus values computed for a sample point that is larger than a camera pixel are mapped to its center pixel. Alternatively, the same focus values can be mapped to all pixels of the sample point. This allows shifts in the size of the sample points' radii and leads to shorter scanning times with lower precision. A third alternative is to map the focus values to the center pixel while performing larger shifts. The resulting voids can then be interpolated. This leads to even shorter scanning times but also to a further reduction of the precision. A planar surface, for example, requires capturing only one image with a coarse grid of sample points.

Thus, the normalized intensities f are used to estimate relative focus values. Common focus operators, such as intensity-based and frequency-based techniques, or point-spread methods can be applied for this. Frequency-based focus operators, such as Laplacian, discrete cosine transformation, or fast Fourier transformation, are often referred to in related literature. However, an even better result can be received by using the focus operator described in (Tsai, 2003), which is originally used to measure the blur on CRT screens. Applying the momentum preserving principle, this operator segments any discrete geometry of the intensity spread inside the corresponding search window into its foreground and background. The proportion of foreground relative to the background is used as focus value.

The focal plane of the camera causes additional blur and influences the absolute focus values. However, since it remains constant for all projectors the same amount of defocus is added in each measurement. This applies also to secondary scattering effects. Note that if the defocus of the camera is too high, small focus variations between the projectors might not be detectable with resolution, intensity response and dynamic range provided by the camera. Consequently, the camera's focus should be adjusted adequately to avoid extreme blur effects in the camera image. This allows computing relative focus values Φ that enable a quantitative comparison of corresponding projector pixels.

The focus values are first determined for each color channel separately and the results are averaged. The momentum preserving operator in (Tsai, 2003) proved also to be stable among the RGB channels. The operator is fast and robust against camera noise.

5.3 Image Compositions

The relative focus values $\Phi_{i,x,y}$ of each projector i that reaches a surface area which is visible in the camera pixel x,y are now known. An image from multiple projector contributions with minimal defocus can then be composed in real-time. Two general techniques are imaginable: An exclusive composition or a weighted composition.

5.3.1 Exclusive Composition

An exclusive image composition allows only one projector (the one with the largest focus value) to cover a surface area which appears at pixel *x,y* in camera space:

$$I_i = w_i(R\text{-}EM)/FM_i,\ w_i = \begin{cases} 1 & \Phi_{i,x,y} \geq \Phi_{j,x,y} \\ 0 & else \end{cases} \tag{18}$$

The binary weights w_i are determined from the focus values Φ offline, coded into a single texture map and are passed to the projector-individual pixel shaders. Alternatively, a stencil mask can be computed for each projector. The mapping from camera space to projector space is given by the *C2P* maps.

Neighboring pixels of different projectors might not be correctly aligned on the surface. This is due to their potentially unequal sizes and orientations, as well as due to imprecision of the geometric calibration. The resulting overlaps and gaps can lead to visible artifacts. To reduce these artifacts, the weight texture must be smoothened using a low-pass filter. This results in soft edges and in non-binary weight values. Since neighboring pixel-contributions from different projectors can now overlap partially, a different composition method has to be applied to ensure a correct radiometric compensation. The weighted composition described in equation 19 can be used for this. But instead of computing normalized weights directly from the focus values, the softened weights of the exclusive weight texture are normalized.

5.3.2 Weighted Composition

Another disadvantage of the exclusive composition method is that the total light intensity that arrives at the surface cannot be larger than that produced by a single projector. This causes visual artifacts at surface pigments with extremely low reflectance or bright environment light.

As explained in section 4, the simultaneous contribution of multiple projectors can overcome this problem. A weighted image composition represents a tradeoff between intensity enhancement and focus refinement:

$$I_i = \frac{w_i(R - EM)}{\sum_j^N w_j FM_j},\ w_{i,x,y} = \frac{\Phi_{i,x,y}}{\sum_j^N \Phi_{j,x,y}} \tag{19}$$

The intensity contribution and the form factor component of each projector *i* that covers the same surface area *x,y* in camera space is weighted. The weights w_i are derived from the focus values and are normalized. They are not binary in this case. Projector contributions with high focus values are up-weighted, and contributions of projectors with low focus values are down-weighted. Together, however, all contributions will always produce the correct result *R* when being reflected by the surface.

The weighted composition allows also to scale the focus values Φ_i up or down during runtime. This makes it possible to amplify or attenuate the contribution (and consequently the focus properties) of an individual projector *i* under retention of a correct radiometric compensation.

Similarly as for the exclusive composition, a static alpha mask can be used alternatively for blending each projector's output instead of weighing the result inside the pixel shaders.

6. VIEW-DEPENDENT PIXEL-PRECISE AUGMENTATIONS

For view-dependent applications (e.g., head-tracked stereoscopic visualizations), however, a single sweet spot (i.e., a single camera view) as assumed in sections 4.1 and 4.2 is not sufficient. The extension towards a view-dependent correction of geometry and radiometric measurements is described below in two variations: as an image-based and as a geometry-based approach.

6.1 Image-Based Approach

To create a radiometrically compensated and geometrically corrected projection for a single camera perspective, the geometric mapping between camera and projector(s), as well as the radiometric parameters have to be measured during calibration. This results in the parameter textures *C2P*, *P2C*, *EM*, and *FM* that are used by projector-individual pixel-shaders to perform a pixel-precise geometric and radiometric correction for the actual perspective in real-time. This process was described in section 4.

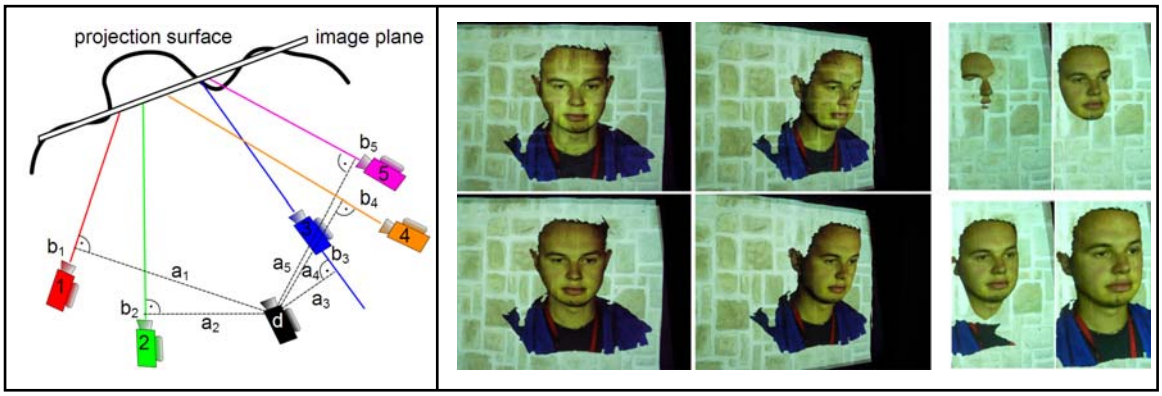

Figure 8: Image-based rendering approach (left): Five source cameras (1-5), one destination camera (d). Example of stereoscopic and head-tracked projection onto complex surface (right): with and without radiometric compensation, and occlusion effects. Images reprinted from (Bimber, 2005b) © IEEE.

Figure 8-left illustrates how this approach can be extended towards view-dependent image-based rendering (Bimber, 2005b) that was inspired by unstructured Lumigraphs (Buehler, 2001):

The set of parameters for multiple, unstructured source camera positions are measured: P_i2C_j, FM_{ij}, and EM_{ij}, where i is the projector index and j is the camera index. The camera can be tracked and its position and orientation can be stored together with each corresponding parameter set. Figure 8-left illustrates this for five unstructured source camera positions. For rendering the image correctly, it has to be defined where the image plane will appear in space. This can be done once before calibration by interactively aligning a 3D model of the image plane at the desired position in the real environment[4]. If the camera is moved, the registered image plane has to be rendered according to the new camera perspective. Two different image plane types can be supported: An on-axis image plane remains at a fixed position in space but its orientation is updated in such a way that it is perpendicular to the vector that is spanned by the camera's position and the central position of the image plane. The orientation and position of an off-axis image plane remains constant in space – no matter where the camera is located.

If it is assumed to render a correct image only for one of the calibrated source cameras j, the following step is performed: The projection of the image plane into the camera's perspective

[4] Visual feedback for this process can be provided by rendering the image plane perspectively correct into the video stream of one source camera.

has to be computed first. This is done by off-screen rendering the registered image plane model from the perspective of this camera. The image plane is shaded with texture coordinates that allow a correct perspective mapping of the original image *O* onto it. These texture coordinates range from u=0..1 and v=0..1 for addressing and displaying the entire image *O*. In the following this image plane texture is referred to as *IP*. Projector-individual pixel shaders can then carry out the following tasks: For each pixel of projector *i* find the corresponding radiometric parameters in FM_{ij} and EM_{ij} using P_i2C_j. Then find the corresponding pixel of the original image *O* by referencing P_i2C_j first to look-up the texture coordinate of *O* in *IP*. Using this texture coordinate, perform a look-up in *O*. Having all parameters, the color correction is performed (see section 4) and the pixel is displayed.

For a novel destination camera position that does not match any of the source camera positions, however, all parameters have to be computed rather than being measured: The geometric and radiometric parameter textures, as well as the direction vector for this novel camera perspective are interpolated from the measured parameters of the source cameras. A new image plane texture *IP* is then rendered from this interpolated perspective. For a correct interpolation, the position of the destination camera is projected onto the direction vectors of the source cameras. Two distances can now be computed for each source camera *j* (cf. figure 8-left): The distance from the destination camera to its projection points on each source cameras' direction vector (a_j). And the distance from the destination camera's projection points to each source cameras' position (b_j). These distances are used for computing penalty weights for each source camera with:

$$p_j = \alpha a_j + (1-\alpha) b_j \tag{20}$$

Note that all distances a_j and b_j have to be normalized over all source cameras before computing the penalties. The factor α allows weighting the contribution of each distance. Since a shift of the destination camera along a source direction vector causes less distortion than a shift away from it, a_j must be weighted higher than b_j. One possibility is to choose $\alpha = 0.75$. Note that neither the orientation of the destination camera, nor the intrinsic parameters of source or destination cameras have to be taken into account for computing the penalty weights.

From all source cameras, a subset of *k* cameras with the smallest penalties is selected. This has to be done due to memory restrictions of current graphics cards. Only these *k* source cameras are considered for sampling the destination camera's new parameter textures. For each of the *k* source cameras, a weight factor can be computed with:

$$w_j = \left(1 - \frac{p_j}{\max_{pk}}\right)\frac{1}{p_j} \tag{21}$$

where $\max_{pk}$ is the maximum penalty among the *k* selected source cameras. Note that all weights have to be normalized after being computed. This implies that the source camera with the largest penalty ($\max_{pk}$) among the *k* selected ones is weighted with 0. A source camera with close-to-zero penalty is first weighted with a value approaching infinity, but is then mapped to 1 after normalization.

The parameter textures and interpolated direction vector for the destination camera can now be computed, rather than being measured. This is performed with the pixel shaders by interpolating each parameter entry t_j of P_i2C_j, FM_{ij}, EM_{ij}, and the original direction vectors among the *k* selected source cameras as follows:

$$t_d = \sum_j^k w_j t_j \quad (22)$$

Note that look-ups in FM_{ij} and EM_{ij} have to be carried out with the original (non-interpolated) P_i2C_j map while look-ups in IP have to be done with the interpolated projector-to-camera map. This allows the computation of the geometric warping, the image plane projection, and the radiometric parameters (surface reflectance, environment contribution and black-level) for a novel destination camera. For completely diffuse surfaces, the radiometric parameters do not change. Weak specular effects, however, are taken into account with this method as well. To handle a flexible number of source cameras, the pixel shaders are not hard-coded, but dynamically generated and loaded onto the graphics card during runtime. This happens only if k is modified.
The view-independent parameter textures, such as the scattering described in section 4.4 or the focus values described in section 5 do not have to be interpolated.

6.2. Geometry-Based Approach

As explained in section 3.2, projective texture mapping and two-pass rendering can be applied for image warping if the surface geometry is known. It is easy to see that the precision of this method strongly depends on the quality of the surface model and on an adequate registration between surface model and projectors. Misregistrations of 2-3 pixels that lead to geometric image distortions can be tolerated in case of simple non-textured surfaces. Performing the radiometric compensation with wrong parameters, however, leads to immediately visible color and intensity artifacts. They make the fusion of stereo pairs difficult. Below a variation of this geometric rendering method that is applicable for view-dependent radiometric compensation is described:
Knowing the pixel correspondence between two or more camera positions allows computing each surface point's 3D position in space which is stored in projector individual geometry maps (GM_i). A definite mapping of 3D surface points to projector pixels is provided implicitly through indexing GM_i. The radiometric parameters (EM_{ij} and FM_{ij}) are measured for each camera-projector combination. The parameter textures that belong to the same projector are then averaged and C_j2P_i-mapped to projector individual look-up textures (EM_i and FM_i) that correspond to the indexing of GM_i.
A texture matrix (TM) that transforms the 3D surface points into the perspective of the observer camera can then be computed. This matrix is a composition ($TM = N*I*E$) of extrinsic (E: position and orientation transformations) and intrinsic (I: perspective projection) parameters of the observer camera followed by a transformation from normalized device coordinates into normalized texture space (N=translate[0.5,0.5,0.0]*scale[0.5,0.5,1.0] for OpenGL). Note that the same matrix is also applied for texturing by conventional projective texture mapping methods. The rendering of the geometry from the perspective of the projector(s), however, is different in the described approach:
For every projector pixel, the corresponding surface point SP_i is looked up in GM_i and is mapped into the perspective of the observer camera with $TM*SP_i$. Consistent occlusion effects can be achieved by performing a depth test with the transformed scene points and the depth map of the virtual scene. Being in the camera space, the pixel of the original image O can be referenced in the corresponding look-up textures, as described in section 6.1. Remember that for performing a look-up in O, the texture coordinates of the defined image plane have to be referenced. Thus the image plane texture IP has to be computed for the perspective of the observer camera and passed to the pixel shader exactly as described in section 6.1. The radiometric parameters can be looked-up in EM_i and FM_i. Note that EM_i, FM_i, and GM_i have projector resolution.
In contrast to conventional projective texture mapping approaches, this variation ensures that the look-up of the radiometric parameters for each projector pixel always matches with its

corresponding surface pigment. Only the mapping into the observer's camera perspective depends on the quality of the estimated geometry map and the precision of the tracking device. This, however, can only cause a geometric misalignment of the image – but no color or intensity artifacts. Intrinsic and extrinsic parameters of the projectors do not have to be acquired. Non-linear projector distortions are corrected by this method as well.

7. POTENTIALS AND LIMITATIONS

Projector-based augmentation has a great potential in many areas. However, there are also several technical limitations that have to be reported on current projector-camera approaches that are used for augmented reality: The fixed resolution of both – cameras and projectors– prevent from measuring and correcting small geometric details and colored pigments that fall below their resolutions. The solution to this problem is to ensure a higher spatial resolution (projector and camera) on a smaller surface area. This can be achieved by a larger number of stationary projectors and cameras[5] or by an interactively tracked projector/camera device. Furthermore off-the-shelf projectors and cameras suffer from a low dynamic range, which makes the capturing and compensation of a large color space impossible. Potential solutions are multi-channel projectors that provide a high-dynamic range in combination with multi-spectral imaging technology (Hill, 2002). High-dynamic range or dynamic range increase techniques represent further software solutions that can enhance the camera measurements.

The high black-level of conventional projectors also makes it difficult to produce absolutely dark areas. For multi-projector configurations the black-level of each projector is being added. This prevents current projector-based AR configurations from using a large number of overlapping projectors for creating very bright images with a high contrast that can compensate all possible pigment colors. Optical filters can reduce the black-level – but they will also reduce the brightness. In some situations, local contrast effects (dark areas surrounded by bright areas are perceived darker than they actually are) reduce this problem on a perceptual level.

Diffuse materials that perfectly absorb light in one or more bands of the spectrum are not well suited for a radiometric compensation approach. Fortunately, such materials are not very common in the real world. Most diffuse surfaces scatter at least a small portion of the light being projected onto them. Thus this challenge reduces to the question of how much light can be projected for achieving the desired result. Strongly specular surfaces, however, will make a projector-based augmentation fail in general. To find a solution to this problem remains a challenge in the future of projector-based augmentation. Alternative concepts, such as mobile and spatial optical or video see-through technologies might be preferable in these situations.

Future projectors will become compact in size and require little power and cooling. Conventional lamps will be replaced by powerful LEDs. This makes them suitable for mobile augmented reality applications or spatial augmented reality configurations (Bimber, 2005d) that apply large number projectors which are seamlessly integrated into everyday environments. Reflective technology (such as DLP or LCOS) will more and more replace transmissive technology (e.g., LCD). This leads to an increased brightness and extremely high update rates that will easily support multi-user stereoscopic visualizations with a single projection unit. Such enabling technologies make the implementation of novel imaging and display techniques possible, such as a simultaneous acquisition and display (Cotting, 2004), or a programmable imaging directly over the digital micromirror array of the projector (Nayar, 2004).

Another advantage of projection technology over traditional display technology applied by the AR community is that its technological development is strongly driven by a large consumer

[5] Or a single mobile camera covering a larger number of sample positions sequentially.

electronics market. This will also lead in future to a fast technical progress and to continuously falling prices.

ACKNOWLEDGEMENTS

I would like to thank and acknowledge all students and colleagues who contributed to the work summarized above: Andreas Emmerling, Gordon Wetzstein, Anselm Grundhöfer, Christian Nitschke, Daniel Danch, Petro Kapakos, Thomas Klemmer, Franz Coriand, Alexander Kleppe, Stefanie Zollmann, Tobias Langlotz, Erich Bruns, and Thomas Zeidler.

REFERENCES

Bell, I.E. (2003). Neutralizing Paintings with a Projector. *Proc. of SPIE/IS&T*, 5008, 560-568.

Biehling, W., Deter, C., Dube, S., Hill, B., Helling, S., Isakovic, K., Klose, S., & Schiewe, K. (2004). LaserCave – Some Building Blocks for immersive Screens. *Proc. of Int. Status Conference on Virtual - and Augmented Reality*.

Bimber, O. & Fröhlich, B. (2002). Occlusion Shadows: Using Projected Light to Generate Realistic Occlusion Effects for View-Dependent Optical See-Through Displays. *Proc. of International Symposium on Mixed and Augmented Reality (ISMAR'02)*, 186-195.

Bimber, O. Grundhöfer, A., Wetzstein, G., & Knödel, S. (2003). Consistent Illumination within Optical See-Through Augmented Environments. *Proc. of IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR'03)*, 198-207.

Bimber, O. (2004) Combining Optical Holograms with Interactive Computer Graphics. *IEEE Computer*, 37(1), 85-91.

Bimber, O., Emmerling, A., & Klemmer, T. (2005a). Embedded Entertainment with Smart Projectors. *IEEE Computer*, 38(1), 56-63.

Bimber, O., Wetzstein, G., Emmerling, A., & Nitschke, C. (2005b). Enabling View-Dependent Stereoscopic Projection in Real Environments. *Proc. of IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR'05)*, 14-23.

Bimber, O., Coriand, F., Kleppe, A., Bruns, E., Zollmann, S., & Langlotz, T. (2005c). Superimposing Pictorial Artwork with Projected Imagery. *IEEE MultiMedia*. 12(1), 16-26.

Bimber, O. & Raskar, R. (2005d). *Spatial Augmented Reality: Merging Real and Virtual Worlds*. A K Peters LTD (publisher), ISBN: 1-56881-230-2.

Bimber, O. & Emmerling, A. (2006a). Multi-Focal Projection: A Multi-Projector Technique for Increasing Focal Depth. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*.

Bimber, O., Grundhöfer, A., Zeidler, T., Danch, D., & Kapakos, P. (2006b). Compensating Indirect Scattering for Immersive and Semi-Immersive Projection Displays. *Proc. of IEEE Virtual Reality (IEEE VR'06)*, 2006.

Brown, M., Majumder, A., and Yang, R. (2005). Camera-Based Calibration Techniques for Seamless Multi-Projector Displays. I*EEE Transactions on Visualization and Computer Graphics*, 11(2), 193-206.

Buehler, C., Bosse, M., McMillan, L., Gortler, S.J., & Cohen, M.F. (2001). Unstructured Lumigraph Rendering. *Proc. of ACM Siggraph'01*, 425-432.

Cotting, D., Naef, M., Gross, M., & Fuchs, H. (2004). Embedding Imperceptible Patterns into Projected Images for Simultaneous Acquisition and Display. *Proc. of IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR'04)*, 100-109.

Ehnes, J., Hirota, K., & Hirose, M. (2004). Projected Augmentation – Augmented Reality using Rotatable Video Projectors. *Proc. of IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR'04)*, 26-35.

Fujii, K., Grossberg, M.D., & Nayar, S.K. (2005). A projector-camera system with real-time photometric adaptation for dynamic environments. *Proc. of Computer Vision and Pattern Recognition (CVPR'05)*, 2, 20-25.

Grossberg, M.D., Peri, H., Nayar, S.K., & Bulhumeur, P. (2004). Making One Object Look Like Another: Controlling Appearance Using a Projector-Camera System. *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'04)*, 1, 452-459.

Hill, B. (2002). (R)evolution of Color Imagining Systems. *Proc. of 1st European Conference on Color in Graphics, Imagining and Vision (CGIV'02)*, 473-479.

Low, K-L., Welch, G., Lastra, A., & Fuchs, H. (2001). Life-Sized Projector-Based Dioramas, *Proc. Symp. Virtual Reality Software and Technology (VRST'01)*, 93-101.

Nayar, S.K., Peri, H., Grossberg, M.D., & Belhumeur, P.N. (2003). A Projection System with Radiometric Compensation for Screen Imperfections. *Proc. of International Workshop on Projector-Camera Systems (ProCams'03)*.

Nayar, S.K. & Boult, T.E. (2004). Programmable Imaging using a Digital Micromirror Array, *Proc. of Computer Vision and Pattern Recognition (CVPR'04)*, 436-443.

Noda, S., Ban, Y., Sato, K., & Chihara, K. (1999). An Optical See-Through Mixed Reality Display with Realtime Rangefinder and an Active Pattern Light Source. *Transactions of the Virtual Reality Society of Japan*, 4(4), 665-670.

Pinhanez, C. (2001). The everywhere displays projector: A device to create ubiquitous graphical interfaces, *Proc. of Ubiquitous Computing (UbiComp'01)*, 315-331.

Raskar, R. (1999a). Oblique Projector Rendering on Planar Surfaces for a Tracked User. *Proc. of ACM Siggraph'99*, sketch.

Raskar, R., Brown, M.S., Yang, R., Chen, W., Welch, G., Towles, H., Seales, B., & Fuchs, H. (1999b). Multi-projector displays using camera-based registration, *Proc. of IEEE Visualization* (IEEE Viz'99), 161-168.

Raskar, R., Welch, G., Low, K.L. & Bandyopadhyay, D. (2001). Shader Lamps: Animating real objects with image-based illumination. *Proc. of Eurographics Rendering Workshop*, 89-102.

Raskar, R., Ziegler, R. & Willwacher, T. (2002). Cartoon Dioramas in Motion. *Proc. of Int. Symposium on Non-photorealistic Animation and Rendering*, 7-ff.

Raskar, R., van Baar, J., Beardsly, P., Willwacher, T., Rao, S. & Forlines, C. (2003). iLamps: Geometrically Aware and Self-Configuring Projectors. *Proc. of ACM Siggraph'03*, 809-818.

Seitz, S.M., Matsushita, Y., & Kutulakos, K.N. (2005). A Theory of Inverse Light Transport. *Proc. of IEEE International Conference on Computer Vision (IEEE ICCV'05)*, 1440-1447, 2005.

Tsai, D.M. and Chou, C.C. (2003). A fast focus measure for video display inspection. *Machine Vision and Applications*, 14(3), 192–196.

Underkoffler, J., Ullmer, B. & Ishii, H. (1999). Emancipated pixels: real-world graphics in the luminous room. *Proc. of ACM Siggraph*, 385-392.

Wang, D., Sato, I., Okabe, T., & Sato, Y. (2005). Radiometric Compensation in a Projector-Camera System Based on the Properties of Human Vision System. *In Proc. of IEEE International Workshop on Projector-Camera Systems (ProCams'05)*.

Yoshida, T., Horii, C. & Sato, K. (2003). A Virtual Color reconstruction System for Real Heritage with Light Projection. *Proc. of Virtual Systems and Multimedia*, 158-164.