The Visual Computing of Projector-Camera Systems

Oliver Bimber[1], Daisuke Iwai[1,2], Gordon Wetzstein[3] and Anselm Grundhöfer[1]

[1]Bauhaus-University Weimar, Germany
{bimber, iwai, grundhoefer}@uni-weimar.de
[2]Osaka University, Japan
iwai@sens.sys.es.osaka-u.ac.jp
[3]University of British Columbia, Canada
wetzste1@cs.ubc.ca

Abstract
This article focuses on real-time image correction techniques that enable projector-camera systems to display images onto screens that are not optimized for projections, such as geometrically complex, coloured and textured surfaces. It reviews hardware-accelerated methods like pixel-precise geometric warping, radiometric compensation, multi-focal projection and the correction of general light modulation effects. Online and offline calibration as well as invisible coding methods are explained. Novel attempts in super-resolution, high-dynamic range and high-speed projection are discussed. These techniques open a variety of new applications for projection displays. Some of them will also be presented in this report.

Keywords: Projector-camera systems, image-correction, GPU rendering, virtual and augmented reality

ACM CCS: I.3.3 [Computer Graphics]: Picture/Image Generation I.4.8 [Image Processing and Computer Vision]: Scene Analysis I.4.9 [Image Processing and Computer Vision]: Applications.

1. Introduction

The increasing capabilities and declining cost of video projectors them 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.

Today's projectors are able to modulate the displayed images spatially and temporally. Synchronized camera feedback is analyzed to support a real-time image correction that enables projections on complex everyday surfaces that are not bound to projector-optimized canvases or dedicated screen configurations.

This article reviews current projector-camera-based image correction techniques. It starts in Section 2 with a discus-sion on the problems and challenges that arise when projecting images onto non-optimized screen surfaces. Geometric warping techniques for surfaces with different topology and reflectance are described in Section 3. Section 4 outlines radiometric compensation techniques that allow the projection onto coloured and textured surfaces of static and dynamic scenes and configurations. It also explains state-of-the-art techniques that consider parameters of human visual perception to overcome technical limitations of projector-camera systems. In both Sections (3 and 4), conventional structured light range scanning as well as imperceptible coding schemes are outlined that support projector-camera calibration (geometry and radiometry). While the previously mentioned sections focus on rather simple light modulation effects, such as diffuse reflectance, the compensation of complex light modulations, such as specular reflection, inter-reflection, refraction, etc. are explained in Section 5. It also shows how the inverse light transport can be used for compensating all measurable light modulation effects. Section 6 is dedicated to a discussion on how novel (at present mainly experimental)
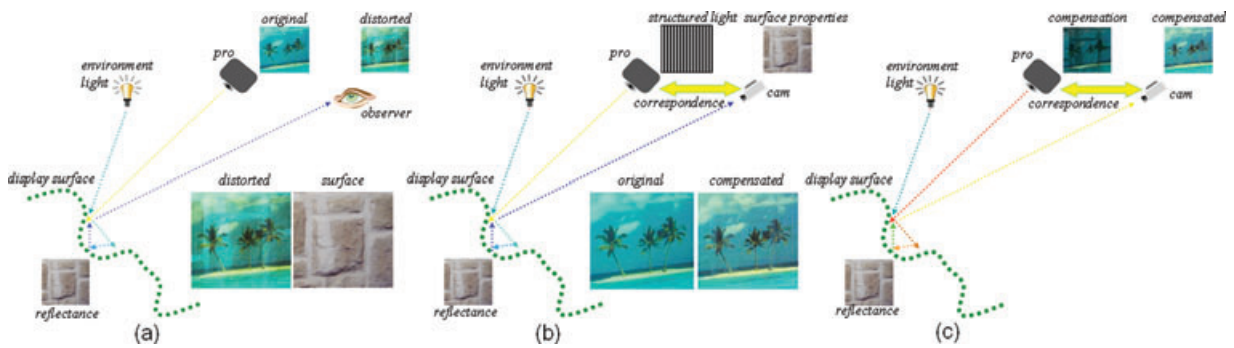
Figure 1: *Projecting onto non-optimized surfaces can lead to visual artefacts in the reflected image (a). Projector-camera systems can automatically scan surface and environment properties (b) to compute compensation images during run-time that neutralize the measured light modulations on the surface (c).*

approaches in high speed, high-dynamic range, large depth of field and super-resolution projection can overcome the technical limitations of today's projector-camera systems in the future.

Such image correction techniques have proved to be useful tools for scientific experiments, but also for real-world applications. Some examples are illustrated in Figures 25–29. They include on-site architectural visualization, augmentations of museum artefacts, video installations in cultural heritage sites, outdoor advertisement displays, projections onto stage settings during live performances and ad-hoc stereoscopic virtual reality (VR)/augmented reality (AR) visualizations within everyday environments. Besides these rather individual application areas, real-time image correction techniques hold the potential of addressing future mass markets, such as flexible business presentations with quickly approaching pocket projector technology, upcoming projection technology integrated in mobile devices – like cellphones, or game-console-driven projections in the home-entertainment sector.

2. Challenges of non-optimized surfaces

For conventional applications, screen surfaces are optimized for a projection. Their reflectance is usually uniform and mainly diffuse (although with possible gain and anisotropic properties) across the surface, and their geometrical topologies range from planar and multi-planar to simple parametric (e.g. cylindrical or spherical) surfaces. In many situations, however, such screens cannot be applied. Some examples are mentioned in Section 1. The modulation of the projected light on these surfaces, however, can easily exceed a simple diffuse reflection modulation. In addition, blending with different surface pigments and complex geometric distortions can degrade the image quality significantly. This is outlined in Figure 1.

The light of the projected images is modulated on the surface together with possible environment light. This leads to a colour, intensity and geometry-distorted appearance (cf. Figure 1a). The intricacy of the modulation depends on the complexity of the surface. It can contain inter-reflections, diffuse and specular reflections, regional defocus effects, refractions and more. To neutralize these modulations in real-time, and consequently to reduce the perceived image distortions is the aim of many projector-camera approaches.

In general, two challenges have to be mastered to reach this goal: first, the modulation effects on the surface have to be measured and evaluated with computer vision techniques and second, they have to be compensated in real-time with computer graphics approaches. Structured light projection and synchronized camera feedback enables the required parameters to be determined and allows a geometric relation between camera(s), projector(s) and surface to be established (cf. Figure 1b). After such a system is calibrated, the scanned surface and environment parameters can be used to compute compensation images for each frame that needs to be projected during run time. If the compensation images are projected, they are modulated by the surface together with the environment light in such a way that the final reflected images approximate the original images from the perspective of the calibration camera/observer (cf. Figure 1c).

The sections below will review techniques that compensate individual modulation effects.

3. Geometric Registration

The amount of geometric distortion of projected images depends on how much the projection surface deviates from a plane, and on the projection angle. Different *geometric projector-camera registration techniques* are applied for individual surface topologies. While simple homographies are suited for registering projectors with planar surfaces, projective texture mapping can be used for non-planar surfaces of
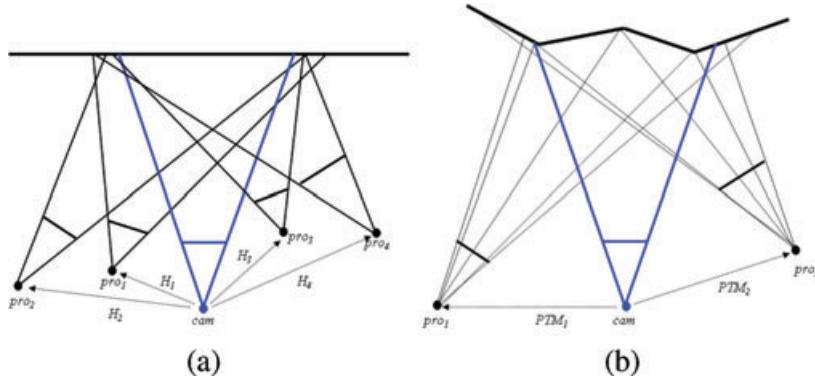
Figure 2: *Camera-based projector registration for untextured planar (a) and non-planar (b) surfaces.*

known geometry. This is explained in Subsection 3.1. For geometrically complex and textured surfaces of unknown geometry, image warping based on look-up operations has frequently been used to achieve a pixel-precise mapping, as discussed in Subsection 3.2. Most of these techniques require structured light projection to enable a fully automatic calibration. Some modern approaches integrate the structured code information directly into the projected image content in such a way that an imperceptible calibration can be performed during run-time. They are presented in Subsection 3.3. Note, that image-warping techniques for parametric surfaces, such as explained in [RvBWR04] are out of the scope of this article.

3.1. Uniformly Coloured Surfaces

For surfaces whose reflectance is optimized for projection (e.g. surfaces with a homogenous white reflectance), a geometric correction of the projected images is sufficient to provide an undistorted presentation to an observer with known perspective. Slight misregistrations of the images on the surface in the order of several pixels lead to geometric artefacts that – in most cases – can be tolerated. This section gives a brief overview over general geometry correction techniques that support a single or multiple projectors for such surfaces.

If multiple projectors (*pro*) have to be registered with a planar surface via camera (*cam*) feedback (cf. Figure 2a), collineations with the plane surface can be expressed as 3x3 camera-to-projector *homography matrix H*:

$$H_{3x3} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}.$$

A homography matrix can be automatically determined numerically by correlating a projection pattern to its corresponding camera image. Knowing the homography matrix H_i for projector pro_i and the calibration camera *cam*, al-

lows the mapping from camera pixel coordinates $cam(x, y)$ to the corresponding projector pixel coordinates $pro_i(x, y)$ with $pro_i(x, y, 1) = H_i \cdot cam(x, y)$. The homographies are usually extended to homogenous 4x4 matrices to make them compatible with conventional transformation pipelines and to consequently benefit from single pass rendering [Ras99]:

$$A_{4x4} = \begin{bmatrix} h_{11} & h_{12} & 0 & h_{13} \\ h_{21} & h_{22} & 0 & h_{23} \\ 0 & 0 & 1 & 0 \\ h_{31} & h_{32} & 0 & h_{33} \end{bmatrix}.$$

Multiplied after the projection transformation, they map normalized camera coordinates into normalized projector coordinates. An observer located at the position of the (possibly off-axis aligned) calibration camera perceives a correct image in this case. Such a camera-based approach is frequently used for calibrating tiled screen projection displays. A sparse set of point correspondences is determined automatically using structured light projection and camera feedback [SPB04]. The correspondences are then used to solve for the matrix parameters of H_i for each projector i. In addition to a geometric projector registration, a camera-based calibration can be used for photometric (luminance and chrominance) matching among multiple projectors. A detailed discussion on the calibration of tiled projection screens is out of the scope of this report. It does not cover multi-projector techniques that are suitable for conventional screen surfaces. The interested reader is referred to [BMY05] for a state-of-the-art overview over such techniques. Some other approaches apply mobile projector-camera systems and homographies for displaying geometrically corrected images on planar surfaces (e.g. [RBvB*04]).

Once the geometry of the projection surface is non-planar but known (cf. Figure 2b), a two-pass rendering technique can be applied for projecting the images in an undistorted way [RWC*98, RBY*99]: In the first pass, the image that has to be displayed is off-screen rendered from a target
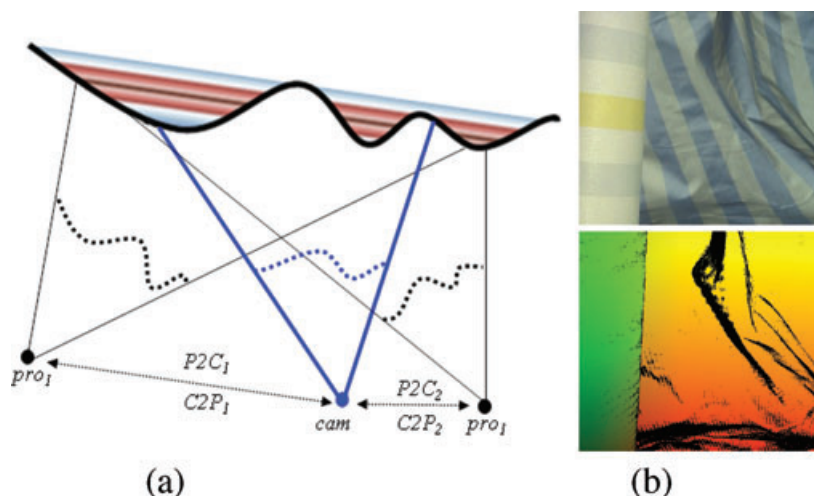
Figure 3: *Camera-based projector registration for textured surfaces (a). The camera perspective onto a scene (b-top) and the scanned look-up table that maps camera pixels to projector pixels. Holes are not yet removed in this example (b-bottom).*

perspective (e.g. the perspective of the camera or an observer). In the second step, the geometry model of the display surface is texture-mapped with the previously rendered image while being rendered from the perspective of each projector *pro*. For computing the correct texture coordinates that ensure an undistorted view from the target perspective *projective texture mapping* is applied. This hardware accelerated technique dynamically computes a texture matrix that maps the 3D vertices of the surface model from the perspectives of the projectors into the texture space of the target perspective.

A camera-based registration is possible in this case as well. For example, instead of a visible (or an invisible – as discussed in Section 3.3) structured light projection, features of the captured distorted image that is projected onto the surface can be analyzed directly. A first example was presented in [YW01] that evaluates the deformation of the image content when projected onto the surface to reconstruct the surface geometry, and refine it iteratively. This approach assumes a calibrated camera-projector system and an initial rough estimate of the projection surface. If the surface geometry has been approximated, the two-pass method outlined above can be applied for warping the image geometry in such a way that it appears undistorted. In [JF07], a similar method is described that supports a movable projector and requires a stationary and calibrated camera, as well as the known surface geometry. The projector's intrinsic parameters and all camera parameters have to be known in both cases. While the method in [YW01] results in the estimated surface geometry, the approach of [JF07] leads to the projector's extrinsic parameters. The possibility of establishing the correspondence between projector and camera pixels in these cases, however, depends always on the quality of detected images features and consequently on the image content itself. To improve their robustness, such techniques apply a predictive

feature matching rather than a direct matching for features in projector and camera space.

However, projective texture mapping in general assumes a simple pinhole camera/projector model and normally does not take the lens distortion of projectors into account (yet, a technique that considers the distortion of the projector for planar untextured screens has been described in [BJM07]). This –together with flaws in feature matching or numerical minimization errors – can cause misregistrations of the projected images in the range of several pixels – even if other intrinsic and extrinsic parameters have been determined precisely. These slight geometric errors are normally tolerable on uniformly coloured surfaces. Projecting corrected images onto textured surfaces with misregistrations in this order causes – even with applying a radiometric compensation (see section 4) – immediate visual intensity and colour artefacts that are well visible. Consequently, more precise registration techniques are required for textured surfaces.

3.2. Textured Surfaces

Mapping projected pixels precisely onto different coloured pigments of textured surfaces is essential for an effective radiometric compensation (described in Section 4). To achieve a precision on a pixel basis is not practical with the registration techniques outlined in Section 3.1. Instead of registering projectors by structured light sampling followed by numerical optimizations that allow the computation of projector-camera correspondences via homographies or other projective transforms, they can be measured pixel-by-pixel and queried through look-up operations during runtime. Well-known structured light techniques [BMS98, SPB04] (e.g. grey code scanning) can be used as well for scanning the

1-to-n mapping of camera pixels to projector pixels. 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 (cf. Figure 3). 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 (e.g. via a Delaunay triangulation of the transformed samples in the $P2C$ map, and a linear interpolation of the pixel colours that store the displacement values within the computed triangles). Figure 3b illustrates the perspective of a camera onto a scene and the scanned and colour-coded (red=x,green=y) $C2P$ texture that maps camera pixels to their corresponding projector pixel coordinates. Note, that all textures contain floating point numbers.

These look-up textures contain only the 2D displacement values of corresponding projector and camera pixels that map onto the same surface point. Thus, neither the 3D surface geometry, nor the intrinsic or extrinsic parameters of projectors and camera are known.

During runtime, 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. We want to refer to this as *pixel displacement mapping*. If multiple projectors are involved, a $P2C$ map has to be determined for each projector. Projector-individual fragment shaders will then perform a customized pixel-displacement mapping during multiple rendering steps, as described in [BEK05].

In [BWEN05] and in [ZLB06], pixel-displacement mapping has been extended to support moving target perspectives (e.g. of the camera and/or the observer). In [BWEN05], an image-based warping between multiple $P2C$ maps that have been pre-scanned for known camera perspectives is applied. The result is an estimated $P2C$ map for a new target perspective during runtime. Examples are illustrated in Figures 27 and 28. While in this case, the target perspective must be measured (e.g. using a tracking device), [ZLB06] analyzes image features of the projected content to approximate a new $P2C$ as soon as the position of the calibration camera has changed. If this is not possible because the detected features are too unreliable, a structured light projection is triggered to scan a correct $P2C$ map for the new perspective.

3.3. Embedded Structured Light

Section 3.1 has already discussed registration techniques (i.e. [YW01, JF07]) that do not require the projection of structured calibration patterns, like grey codes. Instead, they analyze the distorted image content, and thus depend on matchable image
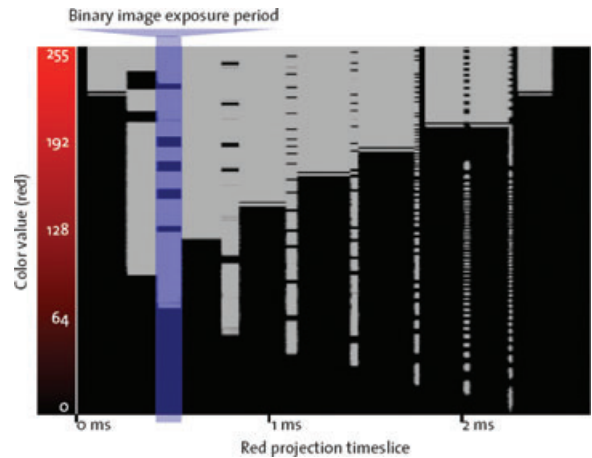

Figure 4: *Mirror flip (on/off) sequences for all intensity values of the red colour channel and the chosen binary image exposure period. © 2004 IEEE [CNGF04].*

features in the projected content. Structured light techniques, however, are more robust because they generate such features synthetically. Consequently, they do not depend on the image content. Overviews over different general coding schemes are given in [BMS98, SPB04].

Besides a spatial modulation, a temporal modulation of projected images allows integrating coded patterns that are not perceivable due to limitations of the human visual system. Synchronized cameras, however, are able to detect and extract these codes. This principle has been described by Raskar *et al.* [RWC*98], and has been enhanced by Cotting *et al.* [CNGF04]. It is referred to as *embedded imperceptible pattern projection*. Extracted code patterns, for instance, allow the simultaneous acquisition of the scenes' depth and texture for 3D video applications [WWC*05, VVSC05]. These techniques, however, can be applied to integrate the calibration code directly into the projected content to enable an invisible online calibration. Thus, the result could be, for instance, a $P2C$ map scanned by a binary grey code or an intensity phase pattern that is integrated directly into the projected content.

The first applicable imperceptible pattern projection technique was presented in [CNGF04], where a specific time slot (called *BIEP*=binary image exposure period) of a digital light processing (DLP) projection sequence is occupied exclusively for displaying a binary pattern within a single colour channel (multiple colour channels are used in [CZGF05] to differentiate between multiple projection units). Figure 4 illustrates an example.

The *BIEP* is used for displaying a binary pattern. A camera that is synchronized to exactly this projection sequence will capture the code. As it can be seen in the selected *BIEP* in
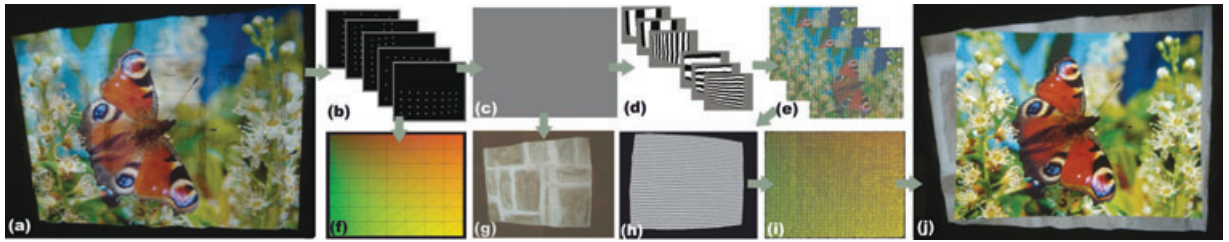
Figure 5: *Imperceptible multi-step calibration for radiometric compensation. A series of invisible patterns (b–e) integrated into an image (a) and projected onto a complex surface (g) results in surface measurements (f–i) used for radiometric compensation (j). © 2007 Eurographics [ZB07].*

Figure 4, the mirror flip sequences are not evenly distributed over all possible intensities. Thus, the intensity of each projected original pixel might have to be modified to ensure that the mirror state is active, which encodes the desired binary value at this pixel. This, however, can result in a non-uniform intensity fragmentation and a substantial reduction of the tonal values. Artefacts are diffused using a dithering technique. A coding technique that benefits from re-configurable mirror flip sequences using the digital micromirror device (DMD) discovery board is described in Section 6.4.

Another possibility of integrating imperceptible code patterns is to modulate the intensity of the projected image I with a spatial code. The result is the code image I_{cod}. In addition, a compensation image I_{com} is computed in such a way that $(I_{cod} + I_{com})/2 = I$. If both images are projected alternately with a high speed, human observers will perceive I due to the slower temporal integration of the human visual system. This is referred to as *temporal coding* and was shown in [RWC*98]. The problem with this simple technique is that the code remains visible during eye movements or code transitions. Both cannot be avoided for the calibration of projector-camera systems using structured light techniques. In [GSHB07] properties of human perception, like adaptation limitations to local contrast changes, are taken into account for adapting the coding parameters depending on local characteristics, such as spatial frequencies and local luminance values of image and code. This makes a truly imperceptible temporal coding of binary information possible. For binary codes, I is regionally decreased ($I_{cod} = I - \Delta$ to encode a binary 0) or increased ($I_{cod} = I + \Delta$ to encode a binary 1) in intensity by the amount of Δ, while the compensation image is computed with $I_{com} = 2I - I_{cod}$. The code can then be reconstructed from the two corresponding images (C_{cod} and C_{com}) captured by the camera with $C_{cod} - C_{com} (< | = | >) 0$. Thereby, Δ is one coding parameter that is locally adapted.

In [PLJP07], another technique for adaptively embedding complementary patterns into projected images is presented. In this work, the embedded code intensity is regionally adapted depending on the spatial variation of neighbouring pixels and their colour distribution in the YIQ colour space. The final code contrast of Δ is then calculated depending on the estimated local spatial variations and colour distributions.

In [ZB07], the binary temporal coding technique was extended to encoding intensity values as well. For this, the code image is computed with $I_{cod} = I\Delta$ and the compensation image with $I_{com} = I(2 - \Delta)$. The code can be extracted from the camera images with $\Delta = 2C_{cod}/(C_{cod} + C_{com})$. Using binary and intensity coding, an imperceptible multi-step calibration technique is presented in [ZB07], which is visualized in Figure 5, and is outline below.

A re-calibration is triggered automatically if misregistrations between projector and camera are detected (i.e., due to motion of camera, projector or surface). This is achieved by continuously comparing the correspondences of embedded point samples. If necessary, a first rough registration is carried out by sampling binary point patterns (cf. Figure 5b) that leads to a mainly interpolated $P2C$ map (cf. Figure 5f). This step is followed by an embedded measurement of the surface reflectance (cf. figures 5c, g), which is explained in Section 4.2. Both steps lead to quick but imprecise results. Then a more advanced 3-step phase shifting technique (cf. Figure 5e) is triggered that results in a pixel-precise $P2C$ registration (cf. Figure 5i). For this, intensity coding is required (cf. Figure 5h). An optional grey code might be necessary for surfaces with discontinuities (cf. Figure 5d). All steps are invisible to the human observer and are executed while dynamic content can be projected with a speed of 20 Hz.

In general, temporal coding is not limited to the projection of two images only. Multiple code and compensation images can be projected if the display frame-rate is high enough. This requires fast projectors and cameras, and will be discussed in Section 6.4.

An alternative to embedding imperceptible codes in the visible light range would be to apply infrared light as shown in [SMO03] for augmenting real environments with invisible information. Although it has not been used for projector-camera calibration, this would certainly be possible.
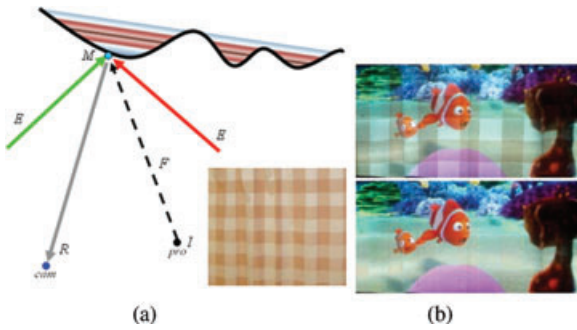
Figure 6: *Radiometric compensation with a single projector (a) and sample images projected without and with compensation onto window curtains (b). © 2007 IEEE [BEK05].*

4. Radiometric Compensation

For projection screens with spatially varying reflectance, colour and intensity compensation techniques are required in addition to a pixel-precise geometric correction. This is known as *radiometric compensation*, and is generally used to minimize the artefacts caused by the local light modulation between projection and surface. Besides the geometric mapping between projector and camera, the surface's reflectance parameters need to be measured on a per-pixel basis before using them for real-time image corrections during run-time. In most cases, a one-time calibration process applies visible structured light projections and camera feedback to establish the correspondence between camera and projector pixels (see Section 3.2) and to measure the surface pigment's radiometric behaviour.

A pixel precise mapping is essential for radiometric compensation since slight misregistrations (in the order of only a few pixels) can lead to significant blending artefacts – even if the geometric artefacts are marginal. Humans are extremely sensitive to even small (less than 2%) intensity variations.

This section reviews different types of radiometric compensation techniques. Starting with methods that are suited for static scenes and projector-camera configurations in Subsection 4.1, it will then discuss more flexible techniques that support dynamic situations (i.e. moving projector-camera systems and surfaces) in Subsection 4.2. Finally, most recent approaches are outlined that dynamically adapt the image content before applying a compensation based on pure radiometric measurements to overcome technical and physical limitations of projector-camera systems. Such techniques take properties of human visual perception into account and are explained in Subsection 4.3.

4.1. Static Techniques

Once the geometric relations are known, the radiometric parameters are measured. One of the simplest radiometric com-

pensation approaches is described in [BEK05]: With respect to Figure 6a, 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 (i.e., the projector) and the surface. A simple representation of the 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 is the distance (considering 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 fraction of I is blended with the pigment's reflectance M: $R = EM + IFM$. Thereby, R is the diffuse radiance that can be captured by the camera. If R, F, M and E are known, a compensation image I can be computed with:

In its most basic configuration (cf. Figure 6a), an image is displayed by a single projector (*pro*) in such a way that it appears correct (colour and geometry) for a single camera view (*cam*). Thereby, the display surfaces must be diffuse, but can have an arbitrary colour, texture and shape. The first step is to determine the geometric relations of camera pixels and projector pixels over the display surface. As explained in Section 3, the resulting $C2P$ and $P2C$ look-up textures support a pixel-precise mapping from camera space to projector space and vice versa.

$$I = (R - EM)/FM. \qquad (1)$$

In a single-projector configuration, E, F and M cannot be determined independently. Instead, FM is measured by projecting a white flood image ($I = 1$) and turning off the entire environment light ($E = 0$), and EM is measured by projecting a black flood 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 (1) can be computed together with pixel displacement mapping (see Section 3.2) in real-time by a fragment shader. Thus, every rasterized projector pixel that passes through the fragment shader is displaced and colour compensated through texture look-ups. The projection of the resulting image I onto the surface leads to a geometry and colour-corrected image that approximates the desired original image $R = O$ for the target perspective of the camera.

One disadvantage of this simple technique is that the optical limitations of colour filters used in cameras and projectors are not considered. These filters can transmit a quite large spectral band of white light rather than only a small monochromatic one. In fact, projecting a pure red colour, for instance, usually leads to non-zero responses in the blue and green colour channels of the captured images. This is known

as the *colour mixing* between projector and camera, which is not taken into account by Equation (1).

Colour mixing can be considered for radiometric compensation: Nayar *et al.* [NPGB03], for instance, express the colour transform between each camera and projector pixel as pixel-individual 3x3 *colour mixing matrices*:

$$V = \begin{bmatrix} v_{RR} & v_{RG} & v_{RB} \\ v_{GR} & v_{GG} & v_{GB} \\ v_{BR} & v_{BG} & v_{BB} \end{bmatrix}.$$

Thereby, v_{RG} represents the green colour component in the red colour channel, for example. This matrix can be estimated from measured camera responses of multiple projected sample images. It can be continuously refined over a closed feedback loop (e.g. [FGN05]) and is used to correct each pixel during runtime. In the case the camera response is known while the projector response remain unknown, it can be assumed that $v_{ii} = 1$. This corresponds to an unknown scaling factor, and V is said to be normalized. The off-diagonal values can then be computed with $v_{ij} = \Delta C_j / \Delta P_i$, where ΔP_i is the difference between two projected intensities $(P_{1i} - P_{2i})$ of primary colour i, and ΔC_j is the difference of the corresponding captured images $(C_{1j} - C_{2j})$ in colour channel j. Thus, six images have to be captured (two per projected colour channel) to determine all v_{ij}. The captured image R under projection of I can now be expressed with: $R = VI$. Consequently, the compensation image can be computed with the inverse colour mixing matrix:

$$I = V^{-1} R. \tag{2}$$

Note, that V is different for each camera pixel and contains the surface reflectance, but not the environment light. Another way of determining V is to numerically solve Equation (2) for V^{-1} if enough correspondences between I and R are known. In this case, V is un-normalized and v_{ii} is proportional to $[FM_R, FM_G, FM_B]$. Consequently, the off-diagonal values of V are 0 if no colour mixing is considered. Yoshida *et al.* [YHS03] use an un-normalized 3x4 colour mixing matrix. In this case, the fourth column represents the constant environment light contribution. A refined version of Nayar's technique was used for controlling the appearance of two- and three-dimensional objects, such as posters, boxes and spheres [GPNB04]. Sections 4.2 and 4.3 also discuss variations of this method for dynamic situations and image adaptations. Note, that a colour mixing matrix was also introduced in the context of shape measurement based on a colour-coded pattern projection [CKS98].

All of these techniques support image compensation in real-time, but suffer from the same problem: if the compensation image I contains values above the maximal brightness or below the black level of the projector, clipping artefacts will occur. These artefacts allow the underlying surface structure to become visible. The intensity range for which radiometric
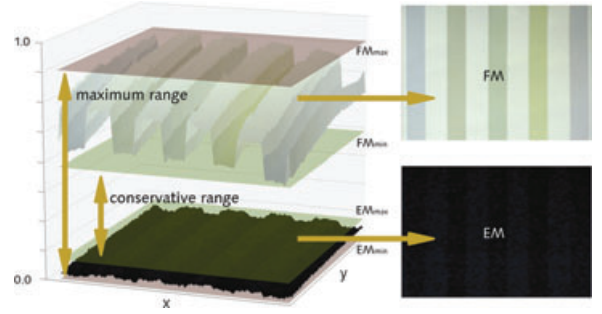

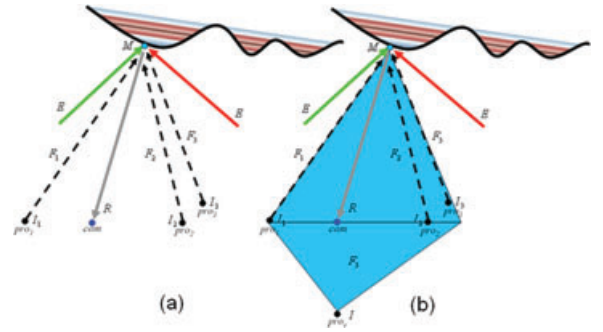**Figure 7:** *Intensity range reflected by a striped wall paper. © 2007 IEEE [GB07].*

Figure 8: *Radiometric compensation with multiple projectors. Multiple individual low-capacity projection units (a) are assumed to equal one single high-capacity unit (b).*

compensation without clipping is possible depends on the surface reflectance, on the brightness and black level of the projector, on the required reflected intensity (i.e. the desired original image), and on the environment light contribution.

Figure 7 illustrates an example that visualizes the reflection properties for a sample surface. By analyzing the responses in both datasets (*FM* and *EM*), the range of intensities for a conservative compensation can be computed. Thus, only input pixels of the desired original image $R = O$ within this global range (bound by the two green planes – from the maximum value EM_{max} to the minimum value FM_{min}) can be compensated correctly for each point on the surface without causing clipping artefacts. All other intensities can potentially lead to clipping and incorrect results. This conservative intensity range for radiometric compensation is smaller than the maximum intensity range achieved when projecting onto optimized (i.e diffuse and white) surfaces.

Different possibilities exist to reduce these clipping problems. While applying an amplifying transparent film material is one option that is mainly limited to geometrically simple surfaces, such as paintings [BCK*05], the utilization of multiple projectors is another option.

The simultaneous contribution of multiple projectors increases the total light intensity that reaches the surface. This can overcome the limitations of equation (1) for extreme situations (e.g. small *FM* values or large *EM* values) and can consequently avoid an early clipping of *I*. Therefore, [BEK05] presents a multi-projector approach for radiometric compensation: If *N* projectors are applied (cf. Figure 8a), the measured radiance captured by the camera can be approximated with: $R = EM + \sum_{i}^{N}(I_i F M_i)$. One strategy is to balance the projected intensities equally among all projectors *i*, which leads to

$$I_i = (R - EM)/\sum_{j}^{N}(I_j F M_j). \qquad (3)$$

Conceptually, this is equivalent to the assumption that a single high-capacity projector (pro_v) produces the total intensity arriving on the surface virtually (cf. Figure 8b). This equation can also be solved in real-time by projector-individual fragment 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. Note however, that the total black level increases together with the total brightness of a multiple projector configuration. Thus, an increase in contrast cannot be achieved. Possibilities for dynamic range improvements are discussed in Section 6.3.

Since the required operations are simple, a pixel-precise radiometric compensation (including geometric warping through pixel-displacement mapping) can be achieved in real-time with fragment shaders of modern graphics cards. The actual speed depends mainly on the number of pixels that have to be processed in the fragment shader. For example, frame rates of >100 Hz can be measured for radiometric compensations using Equation (1) for phase alternating line resolution videos projected in XGA resolution.

4.2. Dynamic Surfaces and Configurations

The techniques explained in Section 4.1 are suitable for purely static scenes and fixed projector-camera configurations. They require a one-time calibration before runtime. For many applications, however, a frequent re-calibration is necessary because the alignment of camera and projectors with the surfaces changes over time (e.g. due to mechanical expansion through heating, accidental offset, intended readjustment, mobile projector-camera systems, or dynamic scenes). In these cases, it is not desired to disrupt a presentation with visible calibration patterns. While Section 3
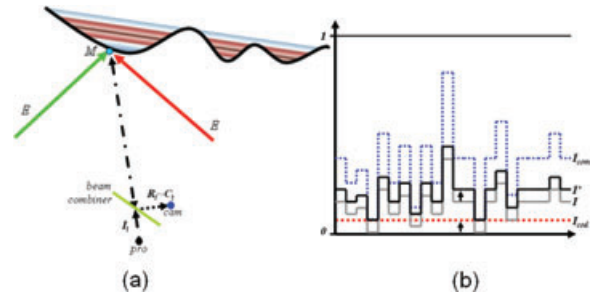

Figure 9: *Co-axial projector-camera alignment (a) and reflectance measurements through temporal coding (b).*

discusses several online calibration methods for geometric correction, this section reviews online radiometric compensation techniques.

Fujii *et al.* have described a dynamically adapted radiometric compensation technique that supports changing projection surfaces and moving projector-camera configurations [FGN05]. Their system requires a fixed co-axial alignment of projector and camera (cf. Figure 9a). An optical registration of both devices makes a frequent geometric calibration unnecessary. Thus, the fixed mapping between projector and camera pixels does not have to be re-calibrated if either surface or configuration changes. At an initial point in time 0 the surface reflectance is determined under environment light ($E_0 M_0$). To consider colour mixing as explained in Section 4.1, this can be done by projecting and capturing corresponding images I_0 and C_0. The reflected environment light E_0 at a pigment with reflectance M_0 can then be approximated by $E_0 M_0 = C_0 - V_0 I_0$, where V_0 is the un-normalized colour mixing matrix at time 0, which is constant. After initialization, the radiance R_t at time *t* captured by the camera under projection of I_t can be approximated with $R_t = M_t/M_0(E_t M_0 + V_0 I_t)$. Solving for I_t results in

$$I_t = V_0^{-1}(R_t M_0/M_{t-1} - E_{t-1} M_0). \qquad (4)$$

Thereby, $R_t = O_t$ is the desired original image and I_t the corresponding compensation image at time *t*. The environment light contribution cannot be measured during runtime. It is approximated to be constant. Thus, $E_{t-1} M_0 = E_0 M_0$. The ratio M_0/M_{t-1} is then equivalent to the ratio C_0/C_{t-1}. In this closed feedback loop, the compensation image I_t at time *t* depends on the captured parameters (C_{t-1}) at time $t-1$. This one-frame delay can lead to visible artefacts. Furthermore, the surface reflectance M_{t-1} is continuously estimated based on the projected image I_{t-1}. Thus, the quality of the measured surface reflectance depends on the content of the desired image R_{t-1}. If R_{t-1} has extremely low or high values in one or multiple colour channels, M_{t-1} might not be valid in all samples. Other limitations of such an approach might be the strict optical alignment of projector and camera that

might be too inflexible for many large scale applications, and that it does not support multi-projector configurations.

Another possibility of supporting dynamic surfaces and projector-camera configurations that do not require a strict optical alignment of both devices was described in [ZB07]. As outlined in Section 3.3, imperceptible codes can be embedded into a projected image through a temporal coding to support an online geometric projector-camera registration. The same approach can be used for embedding a uniform grey image I_{cod} into a projected image I. Thereby, I_{cod} is used to illuminate the surface with a uniform flood-light image to measure the combination of surface reflectance and projector form factor FM, as explained in Section 4.1. To ensure that I_{cod} can be embedded correctly, the smallest value in I must be greater than or equal to I_{cod}. If this is not the case, I is transformed to I' to ensure this condition (cf. Figure 9b). A (temporal) compensation image can then be computed with $I_{com} = 2I' - I_{cod}$. Projecting I_{cod} and I_{com} with a high speed, one perceives $(I_{cod} + I_{com})/2 = I'$. Synchronizing a camera with the projection allows I_{cod} and therefore also FM to be captured. In practice, I_{cod} is approximately 3–5% of the total intensity range – depending on the projector brightness and the camera sensitivity of the utilized devices. One other advantage of this method is, that in contrast to [FGN05] the measurements of the surface reflectance do not depend on the projected image content. Furthermore, Equations (1) or (3) can be used to support radiometric compensation with single or multiple projectors. However, projected (radiometric) compensation images I have to be slightly increased in intensity, which leads to a smaller (equal only if $FM = 1$ and $EM = 0$) global intensity increase of $R = O$. However, since I_{cod} is small, this is tolerable. One main limitation of this method in contrast to the techniques explained in [FGN05], is that it does not react to changes quickly. Usually a few seconds (approx. 5–8 s) are required for an imperceptible geometric and radiometric re-calibration. In [FGN05], a geometric re-calibration is not necessary. As explained in [GSHB07], a temporal coding requires a sequential blending of multiple code images over time, since an abrupt transition between two code images can lead to visible flickering. This is another reason for longer calibration times.

In summary, we can say that fixed co-axial projector-camera alignments as in [FGN05] support real-time corrections of dynamic surfaces for a single mobile projector-camera system. The reflectance measurements' quality depends on the content in O. A temporal coding as in [ZB07] allows unconstrained projector-camera alignments and supports flexible single- or multi-projector configurations – but no real-time calibration. The quality of reflectance measurements is independent on O in the latter case. Both approaches ensure a fully invisible calibration during runtime and enable the presentation of dynamic content (such as movies) at interactive rates ($>= 20$ Hz).

4.3. Dynamic Image Adaptation

The main technical limitations for radiometric compensation are the resolution, frame-rate, brightness and dynamic range of projectors and cameras. Some of these issues will be addressed in Section 6. This section presents alternative techniques that adapt the original images O based on the human perception and the projection surface properties before carrying out a radiometric compensation to reduce the effects caused by brightness limitations, such as clipping.

All compensation methods described so far take only the reflectance properties of the projection surface into account. Particular information about the input image, however, does not influence the compensation directly. Calibration is carried out once or continuously, and a static colour transformation is applied as long as neither surface nor projector-camera configuration changes – regardless of the individual desired image O. Yet, not all projected colours and intensities can be reproduced as explained in Section 4.1 and shown in Figure 7.

Content-dependent radiometric and photometric compensation methods extend the traditional algorithms by applying additional image manipulations depending on the current image content to minimize clipping artefacts while preserving a maximum of brightness and contrast to generate an optimized compensation image.

Such a content-dependent radiometric compensation method was presented by Wang et al. [WSOS05]. In this method, the overall intensity of the input image is scaled until clipping errors that result from radiometric compensation are below a perceivable threshold. The threshold is derived by using a perceptually based physical error metric that was proposed in [RPG99], which considers the image luminance, spatial frequencies and visual masking. This early technique, however, can only be applied to static monochrome images and surfaces. The numerical minimization that is carried out in [WSOS05] requires a series of iterations that make real-time rates impossible.

Park et al. [PLKP06] describe a technique for increasing the contrast in a compensation image by applying a histogram equalization to the coloured input image. While the visual quality can be enhanced in terms of contrast, this method does not preserve the contrast ratio of the original image. Consequently, the image content is modified significantly, and occurring clipping errors are not considered.

A complex framework for computing an optimized photometric compensation for coloured images is presented by Ashdown et al. [AOSS06]. In this method the device-independent CIE L*u*v colour space is used, which has the advantage that colour distances are based on the human visual perception. Therefore, an applied high-dynamic range (HDR) camera has to be colour calibrated in advance. The input images are adapted depending on a series of global and
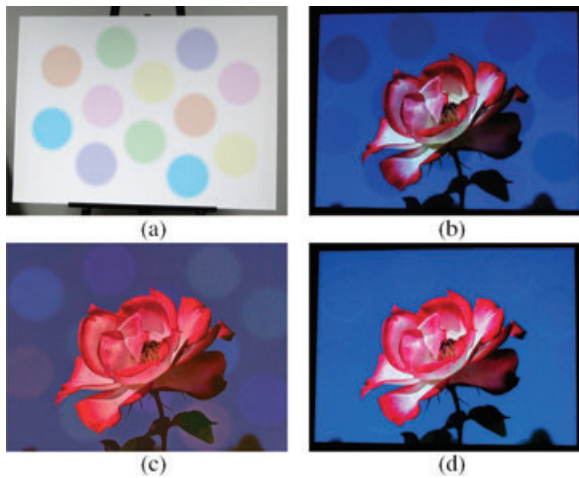
Figure 10: *Results of a content-dependent photometric compensation. The uncompensated image leads to visible artefacts (b) when being projected onto a coloured surface (a). The projection of an adapted compensation image (c) minimizes the visibility of these artefacts (d). © 2006 IEEE [AOSS06].*

local parameters to generate an optimized compensated projection: The captured surface reflectance as well as the content of the input image are transformed into the CIE L*u*v colour space. The chrominance values of all input image's pixels are fitted into the gamut of the corresponding projector pixels. In the next step, a luminance fitting is applied by using a relaxation method based on differential equations. Finally, the compensated adapted input image is transformed back into the RGB colour space for projection.

This method achieves optimal compensation results for surfaces with varying reflectance properties. Furthermore, a compensation can be achieved for highly saturated surfaces due to the fact that besides a luminance adjustment, a chrominance adaptation is applied as well. Its numerical complexity, however, allows the compensation of still images only. Figure 10 shows a sample result: An uncompensated projection of the input image projected onto a coloured surface (a) results in colour artefacts (b). Projecting the adapted compensation image (c) onto the surface leads to significant improvements (d).

Ashdown *et al.* proposed another fitting method in [ASOS07] that uses the chrominance threshold model of human vision together with the luminance threshold to avoid visible artefacts.

Content-dependent adaptations enhance the visual quality of a radiometric compensated projection compared to static methods that do not adapt to the input images. Animated content like movies or TV-broadcasts, however, cannot be compensated in real-time with the methods reviewed above.

While movies could be pre-corrected frame-by-frame in advance, real-time content like interactive applications cannot be presented.

In [GB07], a real-time solution for adaptive radiometric compensation was introduced that is implemented entirely on the graphics processing unit (GPU). The method adapts each input image in two steps: First it is analyzed for its average luminance that leads to an approximate global scaling factor, which depends on the surface reflectance. This factor is used to scale the input image's intensity between the conservative and the maximum intensity range (cf. Figure 7 in Section 4.1). Afterwards, a compensation image is calculated according to equation (1). Instead of projecting this compensation image directly, it is further analyzed for potential clipping errors. Errors are extracted and blurred in addition. In a final step, the input image is scaled globally again depending on its average luminance and on the calculated maximum clipping error. In addition, it is scaled locally based on the regional error values. The *threshold map* explained in [RPG99] is used to constrain the local image manipulation based on the contrast and the luminance sensitivity of human observers. Radiometric compensation (Equation (1)) is applied again to the adapted image, and the result is finally projected. Global, but also local scaling parameters are adapted over time to reduce abrupt intensity changes in the projection, which would lead to a perceived and irritating flickering. This approach does not apply numerical optimizations and consequently enables a practical solution to display adapted dynamic content in real-time and in increased quality (compared to traditional radiometric compensation). Yet, small clipping errors might still occur. However, especially for content with varying contrast and brightness, this adaptive technique enhances the perceived quality significantly. An example is shown in Figure 11: Two frames of a movie (b,e) are projected with a static compensation technique [BEK05] (c,f) and with the adaptive real-time solution [GB07] (d,g) onto a natural stone wall (a). While clipping occurs in case (c), case (f) appears too dark. The adaptive method reduces the clipping errors for bright images (d) while maintaining details in the darker image (g).

5. Correcting Complex Light Modulations

All image correction techniques that have been discussed so far assume a simple geometric relation between camera and projector pixels that can be automatically derived using homography matrices, structured light projections or co-axial projector-camera alignments.

When projecting onto complex everyday surfaces, however, the emitted radiance of illuminated display elements is often subject to complex lighting phenomena. Due to diffuse or specular inter-reflections, refractions and other global illumination effects, multiple camera pixels at spatially distant regions on the camera image plane may be affected by a single projector pixel.
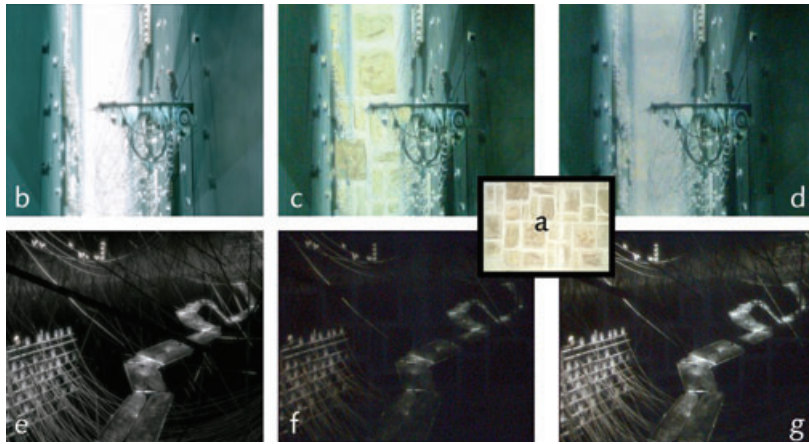
Figure 11: *Two frames of a movie (b,e) projected onto a natural stone wall (a) with static (c,f) and real-time adaptive radiometric compensation (d,g) for bright and dark input images. © 2007 IEEE [GB07].*

A variety of projector-camera-based compensation methods for specific global illumination effects have been proposed. These techniques, as well as a generalized approach to compensating light modulations using the inverse light transport will be discussed in the following subsections. We start with discussions on how diffuse inter-reflections (Subsection 5.1) and specular highlights (Subsection 5.2) can be compensated. The inverse light transport approach is introduced as the most general image correction scheme in Subsection 5.3.

5.1. Inter-reflections

Eliminating diffuse inter-reflections or scattering for projection displays has recently gained a lot of interest in the computer graphics and vision community. Cancellation of inter-reflections has been proven to be useful for improving the image quality of immersive virtual and augmented reality displays [BGZ*06]. Furthermore, such techniques can be employed to remove indirect illumination from photographs [SMK05]. For compensating global illumination effects, these need to be acquired, stored and processed, which will be discussed for each application.

Seitz *et al.* [SMK05], for instance, measured an *impulse scatter function* (ISF) matrix B with a camera and a laser pointer on a movable gantry. The camera captured diffuse objects illuminated at discrete locations. Each of the samples' centroid represents one row/column in the matrix as depicted in Figure 12.

The ISF matrix can be employed to remove inter-reflections from photographs. Therefore, an *inter-reflection cancellation operator* $C^1 = B^1 B^{-1}$ is defined that, when multiplied to a captured camera image R, extracts its direct illumination. B^{-1} is the ISF matrix's inverse and B^1 contains only direct illumination. For a diffuse scene, this can easily
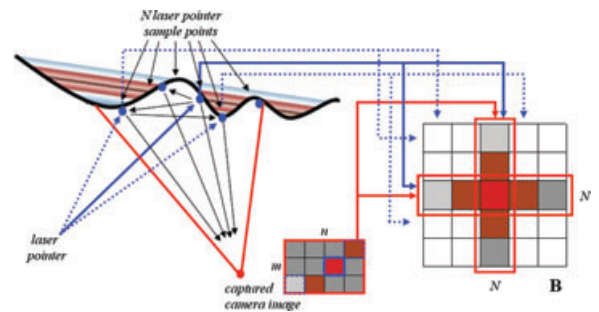

Figure 12: *A symmetric ISF matrix is acquired by illuminating a diffuse surface at various points, sampling their locations in the camera image and inserting captured colour values into the matrix.*

be extracted from B by setting its off-diagonal elements to zero. A related technique that quickly separates direct and indirect illumination for diffuse and non-diffuse surfaces was introduced by Nayar *et al.* [NKGR06].

Experimental results in [SMK05] were obtained by sampling the scene at approx. 35 locations in the camera image under laser illumination. Since B is in this case a very small and square matrix it is trivial to be inverted for computing B^{-1}. However, inverting a general light transport matrix in a larger scale is a challenging problem and will be discussed in Section 5.3.

Compensating indirect diffuse scattering for immersive projection screens was proposed in [BGZ*06]. Assuming a known screen geometry, the scattering was simulated and corrected with a customized reverse radiosity scheme. Bimber *et al.* [Bim06] and Mukaigawa *et al.* [MKO06] showed that a compensation of diffuse light interaction can be
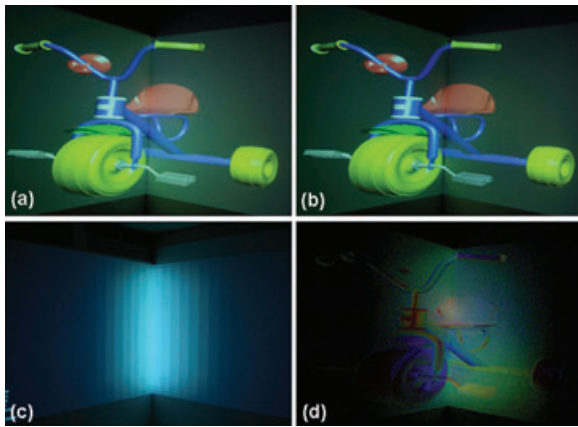
Figure 13: *Compensating diffuse scattering: An uncompensated (a) and a compensated (b) stereoscopic projection onto a two-sided screen. Scattering and colour bleeding can be eliminated (d) if the form factors (c) of the projection surface are known. © 2006 IEEE [BGZ*06].*

Figure 14: *Radiometric compensation in combination with specular reflection elimination. Projection onto a specular surfaces (a) – before (b) and after (c) specular highlight compensation. © 2006 IEEE [PLS*06].*

performed in real-time by reformulating the radiosity equation as $I = (1 - \rho F)O$. Here O is the desired original image, I the projected compensation image, 1 the identity matrix and ρF the pre-computed form-factor matrix. This is equivalent to applying the inter-reflection cancellation operator, introduced in [SMK05], to an image O that does not contain inter-reflections. The quality of projected images for a two-sided projection screen can be greatly enhanced as depicted in Figure 13. All computations are performed with a relatively coarse patch resolution of about 128×128 as seen in Figure 13(c).

While the form factor matrix in [Bim06], [MKO06] was pre-computed, Habe *et al.* [HSM07] presented an algorithm that automatically acquires all photometric relations within the scene using a projector-camera system. They state also that this theoretically allows specular inter-reflections to be compensated for a fixed viewpoint. However, such a compensation has not been validated in the presented experiments. For the correction, a form-factor matrix inverse is required, which again is trivial to be calculated for a low-patch resolution.

5.2. Specular Reflections

When projecting onto non-diffuse screens, not only diffuse and specular inter-reflections affect the quality of projected imagery, but a viewer may also be distracted by specular highlights. Park *et al.* [PLKP05] presented a compensation approach that attempts to minimize specular reflections using multiple overlapping projectors. The highlights are not due to global illumination effects, but to the incident illumination that is reflected directly towards the viewer on a shiny sur-
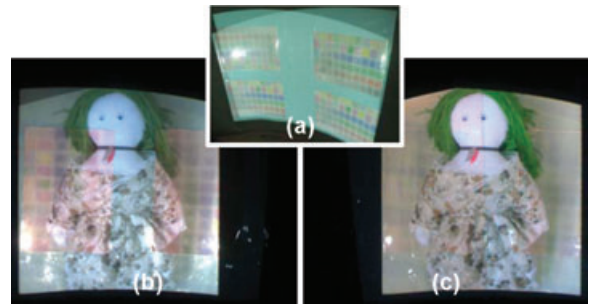
face. Usually, only one of the projectors creates a specular highlight at a point on the surface. Thus, its contribution can be blocked while display elements from other projectors that illuminate the same surface area from a different angle are boosted.

For a view-dependent compensation of specular reflections, the screen's geometry needs to be known and registered with all projectors. Displayed images are pre-distorted to create a geometrically seamless projection as described in Section 3. The amount of specularity for a projector i at a surface point s with a given normal n is proportional to the angle θ_i between n and the sum of the vector from s to the projector's position p_i and the vector from s to the viewer u:

$$\theta_i = \cos^{-1}\left(\frac{-n \cdot (p_i + u)}{|p_i + u|}\right). \tag{5}$$

Assuming that k projectors illuminate the same surface, a weight w_i is multiplied to each of the incident light rays for a photometric compensation:

$$w_i = \frac{\sin(\theta_i)}{\sum_{j=1}^{k} \sin(\theta_j)}. \tag{6}$$

Park *et al.* [PLS*06] extended this model by an additional radiometric compensation to account for the colour modulation of the underlying projection surface (cf. Figure 14). Therefore, Nayar's model [NPGB03] was implemented. The required one-to-one correspondences between projector and camera pixels were acquired with projected binary grey codes [SPB04].

5.3. Radiometric Compensation Through Inverse Light Transport

Although the previously discussed methods are successful in compensating particular aspects of the light transport between projectors and cameras, they lead to a fragmented
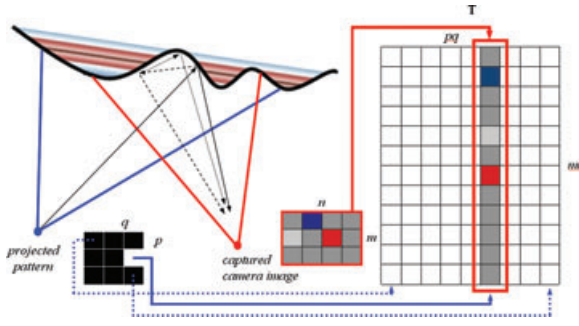
Figure 15: *The light transport matrix between a projector and a camera.*

understanding of the subject. A unified approach that accounts for many of the problems that were individually addressed in previous works was described in [WB07]. The full light transport between a projector and a camera was employed to compensate direct and indirect illumination effects, such as inter-reflections, refractions and defocus, with a single technique in real-time. Furthermore, this also implies a pixel-precise geometric correction. In the following subsection, we refer to the approach as performing radiometric compensation. However, geometric warping is always implicitly included.

In order to compensate direct and global illumination as well as geometrical distortions in a generalized manner, the full light transport has to be taken into account. Within a projector-camera system, this is a matrix T_λ that can be acquired in a pre-processing step, for instance as described by Sen *et al.* [SCG*05]. Therefore, a set of illumination patterns is projected onto the scene and recorded using HDR imaging techniques (e.g. [DM97]). Individual matrix entries can then be reconstructed from the captured camera images. As depicted in Figure 15, a camera image with a single lit projector pixel represents one column in the light transport matrix. Usually, the matrix is acquired in a hierarchical manner by simultaneously projecting multiple pixels.

For a single-projector-camera configuration the forward light transport is described by a simple linear equation as

$$
\begin{bmatrix} r_R - e_R \\ r_G - e_G \\ r_B - e_B \end{bmatrix} = \begin{bmatrix} T_R^R & T_R^G & T_R^B \\ T_G^R & T_G^G & T_G^B \\ T_B^R & T_B^G & T_B^B \end{bmatrix} \begin{bmatrix} i_R \\ i_G \\ i_B \end{bmatrix}, \quad (7)
$$

where each r_λ is a single colour channel λ of a camera image with resolution $m \times n$, i_λ is the projection pattern with a resolution of $p \times q$, and e_λ are direct and global illumination effects caused by the environment light and the projector's black level captured from the camera. Each light transport matrix $T_{\lambda_c}^{\lambda_p}$ (size: $mn \times pq$) describes the contribution of a single projector colour channel λ_p to an individual camera channel λ_c. The model can easily be extended for k projectors

and l cameras:

$$
\begin{bmatrix} {}_1 r_R - {}_1 e_R \\ {}_1 r_G - {}_1 e_G \\ \vdots \\ {}_l r_B - {}_l e_B \end{bmatrix} = \begin{bmatrix} {}_1^1 T_R^R & {}_1^1 T_R^G & \cdots & {}_1^k T_R^B \\ {}_1^1 T_G^R & {}_1^1 T_G^G & \cdots & {}_1^k T_G^B \\ \vdots & \vdots & \ddots & \vdots \\ {}_l^1 T_B^R & {}_l^1 T_B^G & \cdots & {}_l^k T_B^B \end{bmatrix} \begin{bmatrix} {}^1 i_R \\ {}^1 i_G \\ \vdots \\ {}^k i_B \end{bmatrix}. \quad (8)
$$

For a generalized radiometric compensation the camera image r_λ is replaced by a desired image o_λ of camera resolution and the system can be solved for the projection pattern i_λ that needs to be projected. This accounts for colour modulations and geometric distortions of projected imagery. Due to the matrix's enormous size, sparse matrix representations and operations can help to save storage and increase performance.

A customized clustering scheme that allows the light transport matrix's pseudo-inverse to be approximated is described in [WB07]. Inverse ISFs or form-factor matrices had already been used in previous algorithms [SMK05, Bim06, MKO06, HSM07], but in a much smaller scale, which makes an inversion trivial. Using the light transport matrix's approximated pseudo-inverse, radiometric compensation reduces to a matrix-vector multiplication:

$$
i_\lambda = T_\lambda^+ (o_\lambda - e_\lambda). \quad (9)
$$

In [WB07], this was implemented on the GPU and yielded real-time frame-rates.

Figure 16 shows a compensated projection onto highly refractive material (f), which is impossible with conventional approaches (e), because a direct correspondence between projector and camera pixels is not given. The light transport matrix (cf. Figure 16b) and it's approximated pseudo-inverse (visualized in c) contain local and global illumination effects within the scene (global illumination effects in the matrix are partially magnified in b).

It was shown in [WB07] that all measurable light modulations, such as diffuse and specular reflections, complex inter-reflections, diffuse scattering, refraction, caustics, defocus, etc. can be compensated with the multiplication of the inverse light transport matrix and the desired original image. Furthermore, a pixel-precise geometric image correction is implicitly included and becomes feasible – even for surfaces that are unsuited for a conventional structured light scanning. However, due to the extremely long acquisition time of the light transport matrix (up to several hours), this approach will not be practical before accelerated scanning techniques have been developed.

6. Overcoming Technical Limitations

Most of the image correction techniques that are described in this report are constrained by technical limitations of projector and camera hardware. An insufficient resolution
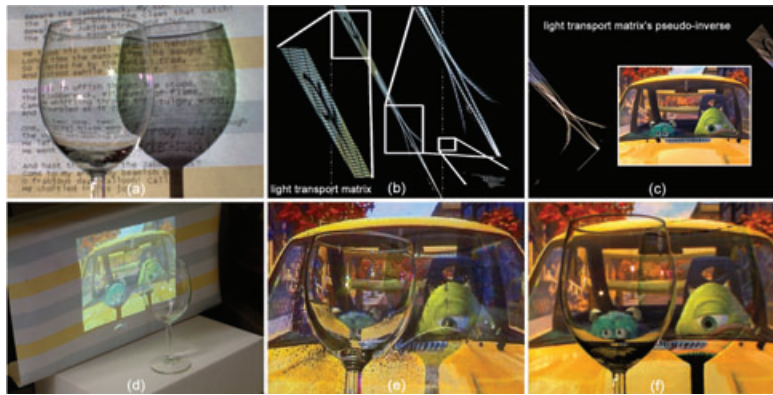
Figure 16: *Real-time radiometric compensation (f) of global illumination effects (a) with the light transport matrix's (b) approximated pseudo-inverse (c).*

or dynamic range of both devices leads to a significant loss of image quality. A too short focal depth results in regionally defocused image areas when projected onto surfaces with an essential depth variance. Slow projection frame-rates will cause the perception of temporally embedded codes. This section is dedicated to giving an overview of novel (at present mainly experimental) approaches that might lead to future improvements of projector-camera systems in terms of focal depth (Subsection 6.1), high resolution (Subsection 6.2), dynamic range (Subsection 6.3) and high speed (Subsection 6.4).

6.1. Increasing Focal Depth

Projections onto geometrically complex surfaces with a high-depth variance generally do not allow the displayed content to be in focus everywhere. Common DLP or LCD projectors usually maximize their brightness with large apertures. Thus, they suffer from narrow depths of field and can only generate focused imagery on a single fronto-parallel screen. Laser projectors, which are commonly used in planetaria, are an exception. These emit almost parallel light beams, which make very large depths of field possible. However, the cost of a single professional laser projector can exceed the cost of several hundred conventional projectors. In order to increase the depth of field of conventional projectors, several approaches for deblurring unfocused projections with a single or with multiple projectors have been proposed.

Zhang and Nayar [ZN06] presented an iterative, spatially-varying filtering algorithm that compensates for projector defocus. They employed a coaxial projector-camera system to measure the projection's spatially varying defocus. Therefore, dot patterns as depicted in Figure 17a are projected onto the screen and captured by the camera (b). The defocus kernels for each projector pixel can be recovered from the captured images and encoded in the rows of a matrix B. Given the environment light EM including the projector's
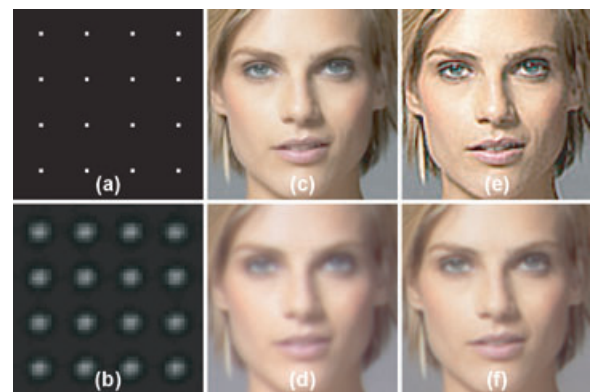

Figure 17: *Defocus compensation with a single projector: An input image (c) and its defocused projection onto a planar canvas (d). Solving Equation (10) results in a compensation image (e) that leads to a sharper projection (f). For this compensation, the spatially varying defocus kernels are acquired by projecting dot patterns (a) and capturing them with a camera (b). © 2006 ACM [ZN06].*

black level and a desired input image O, the compensation image I can be computed by minimizing the sum-of-squared pixel difference between O and the expected projection $BI + EM$ as

$$\arg\min_{I,\ 0 \le I \le 255} \|BI + EM - O\|^2, \qquad (10)$$

which can be solved with a constrained, iterative steepest gradient solver as described in [ZN06].

An alternative approach to defocus compensation for a single projector setup was presented by Brown *et al.* [BSC06]. Projector defocus is modelled as a convolution of a projected original image O and Gaussian PSFs as $R(x, y) = O(x, y) \otimes H(x, y)$, where the blurred image that can be cap-

tured by a camera is R. The PSFs are estimated by projecting features on the canvas and capturing them with a camera. Assuming a spatially invariant PSF, a compensation image I can be synthesized by applying a Wiener deconvolution filter to the original image

$$I(x, y) = \mathcal{F}^{-1} \left\{ \frac{\tilde{H}^*(u, v)\, \tilde{O}(u, v)}{|\tilde{H}(u, v)|^2 + 1/SNR} \right\}. \quad (11)$$

The signal-to-noise ratio (SNR) is estimated a priori, $\tilde{O}$ and $\tilde{H}$ are the Fourier transforms of O and H, respectively, and $\tilde{H}^*$ is $\tilde{H}$'s complex conjugate. $\mathcal{F}^{-1}$ denotes the inverse Fourier transform. Since the defocus kernel H is generally not spatially invariant (this would only be the case for a fronto-parallel plane) Wiener filtering cannot be applied directly. Therefore, basis compensation images are calculated for each of the uniformly sampled feature points using Equation (11). The final compensation image is then generated by interpolating the four closest basis responses for each projector pixel.

Oyamada and Saito [OS07] presented a similar approach to single projector defocus compensation. Here, circular PSFs are used for the convolution and estimated by comparing the original image to various captured compensation images that were generated with different PSFs.

The main drawback of these single projector defocus compensation approaches is that the quality is highly dependent on the projected content. All of the discussed methods result in a pre-sharpened compensation image that is visually closer to the original image after being optically blurred by the defocused projection. While soft contours can be compensated, this is generally not the case for sharp features.

Inverse filtering for defocus compensation can also be seen as the division of the original image by the projector's aperture image in frequency domain. Low magnitudes in the Fourier transform of the aperture image, however, lead to intensity values in spatial domain that exceed the displayable range. Therefore, the corresponding frequencies are not considered, which then results in visible ringing artefacts in the final projection. This is the main limitation of the approaches discussed above, since in frequency domain the Gaussian PSF of spherical apertures does contain a large fraction of low Fourier magnitudes. As shown above, applying only small kernel scales will reduce the number of low Fourier magnitudes (and consequently the ringing artefacts) – but will also lead only to minor focus improvements. To overcome this problem, a coded aperture whose Fourier transform has initially less low magnitudes was applied in [GB08]. Consequently, more frequencies are retained and more image details are reconstructed (cf. Figure 18).

An alternative approach that is less dependent on the actual frequencies in the input image was introduced in [BE06]. Multiple overlapping projectors with varying focal depths
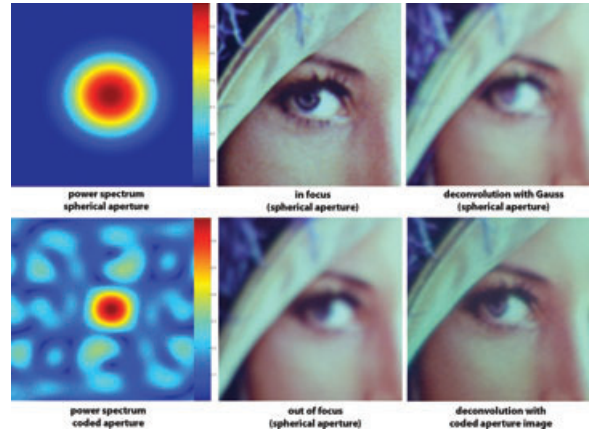

Figure 18: *The power spectra of the Gaussian PSF of a spherical aperture and of the PSF of a coded aperture: Fourier magnitudes that are too low are clipped (black), which causes ringing artefacts. Image projected in focus, and with the same optical defocus (approx. 2 m distance to focal plane) in three different ways: with spherical aperture – untreated and deconvolved with Gaussian PSF, with coded aperture and deconvolved with PSF of aperture code.*

illuminate arbitrary surfaces with complex geometry and reflectance properties. Pixel-precise focus values $\Phi_{i,x,y}$ are automatically estimated at each camera pixel (x, y) for every projector. Therefore, a uniform grid of circular patterns is displayed by each projector and recorded by a camera. In order to capture the same picture (geometrically and colour-wise) for each projection, these are pre-distorted and radiometrically compensated as described in Sections 3 and 4.

Once the relative focus values are known, an image from multiple projector contributions with minimal defocus can be composed in real-time. A weighted image composition represents a trade-off between intensity enhancement and focus refinement as:

$$I_i = \frac{w_i(R - EM)}{\sum_j^N w_j FM_j}, \quad w_{i,x,y} = \frac{\Phi_{i,x,y}}{\sum_j^N \Phi_{j,x,y}}, \quad (12)$$

where I_i is the compensation image for projector i if N projectors are applied simultaneously. Display contributions with high focus values are up-weighted while contributions of projectors with low focus values are down-weighted proportionally. A major advantage of this method, compared to single projector approaches, is that the focal depth of the entire projection scales with the number of projectors. An example for two projectors can be seen in Figure 19.

6.2. Super-Resolution

Super-resolution techniques can improve the accuracy of geometric warping (see Section 3) and consequently have the
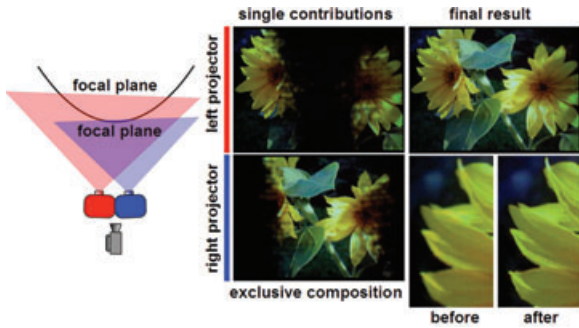
Figure 19: *Defocus compensation with two overlapping projectors that have differently adjusted focal planes. © 2006 IEEE [BE06].*

Figure 20: *Super-resolution projection with a multi-projector setup (a), overlapping images on the projection screen (b) and close-up of overlapped pixels (c).*

potential to enhance radiometric compensation (see Section 4) due to a more precise mapping of projector pixels onto surface pigments. Over the past years, several researches have proposed super-resolution camera techniques to overcome the inherent limitation of low-resolution imaging systems by using signal processing to obtain super-resolution images (or image sequences) with multiple low-resolution devices [PPK03]. Using a single camera to obtain multiple frames of the same scene is most popular. Multi-camera approaches have also been proposed [WJV*05].

On the other hand, super-resolution projection systems are just beginning to be researched. This section introduces recent work on such techniques that can generally be categorized into two different groups. The first group proposes super-resolution rendering with a single projector [AU05]. Other approaches achieve this with multiple overlapping projectors [JR03],[DVC07].

In single projector approaches, so-called *wobulation* techniques are applied: Multiple sub-frames are generated from an original image. An optical image shift displaces the projected image of each sub-frame by a fraction of a pixel [AU05]. Each sub-frame is projected onto the screen with slightly different positions using an opto-mechanical image shifter. This light modulator must be switched fast enough so that all sub-frames are projected in one frame. Consequently, observers perceive this rapid sequence as a continuous and flicker-free image while the resolution is spatially enhanced. Such techniques have been already realized with a DLP system (SmoothPicture®, Texas Instruments Incorporated).

The goal of multi-projector super-resolution methods is to generate a high-resolution image with the superposition of multiple low-resolution sub-frames produced by different projection units. Thereby, the resolutions of each sub-frame differ and the display surfaces are assumed to be diffuse. Super-resolution pixels are defined by the overlapping sub-frames that are shifted on a sub-pixel basis as shown in Figure 20. Generally, the final image is estimated as the sum
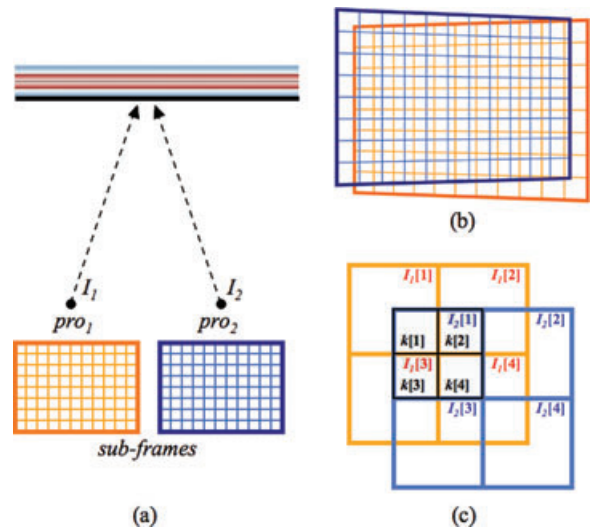
of the sub-frames. If N sub-frames $I_{i=1\ldots N}$ are displayed, this is modelled as:

$$R = \sum_{i}^{N} A_i V_i I_i + EM. \tag{13}$$

Note that in this case, the parameters R, I_i and EM are images, and that A_i and V_i are the geometric warping matrix and the colour mixing matrix that transform the whole image (in contrast to Sections 3 and 4, where these parameters represent transformations of individual pixels).

Figure 20c shows a close-up of overlapping pixels to illustrate the problem that has to be solved: While $I_1[1\ldots 4]$ and $I_2[1\ldots 4]$ are the physical pixels of two projectors, $k[1\ldots 4]$ represent the desired 'super-resolution' pixel structure. The goal is to find the intensities and colours of corresponding projector pixels in I_1 and I_2 that approximate k as close as possible by assuming that the perceived result is $I_1 + I_2$. This is obviously a global optimization problem, since k and I have different resolutions. Thus, if O is the desired original image and R is the captured result, the estimation of sub-frame I_i for projector i is in general achieved by minimizing $||O - R||^2$:

$$I_i = \arg\min_{I_i} ||O - R||^2. \tag{14}$$

Jaynes *et al.* first demonstrated resolution enhancement with multiple superimposed projections [JR03]. Homographies are used for initial geometric registration of multiple sub-frames onto a planar surface. However, homographic transforms lead to uniform two-dimensional shifts and

sampling rates with respect to the camera image rather than to non-uniform ones of general projective transforms.

To reduce this effect, a warped sub-frame is divided into smaller regions that are shifted to achieve sub-pixel accuracy. Initially, each such frame is estimated in the frequency domain by phase shifting the frequencies of the original image. Then, a greedy heuristic process is used to recursively update pixels with the largest global error with respect to Equation (14). The proposed model does not consider V_i and *EM* in Equation (13) and a camera is used only for geometric correction. The iterations of the optimization process are terminated manually in [JR03].

Damera-Venkata *et al.* proposed a real-time rendering algorithm for computing sub-frames that are projected by superimposed lower-resolution projectors [DVC07]. In contrast to the previous method, they use a camera to estimate the geometric and photometric properties of each projector during a calibration step. Image registration is achieved on a sub-pixel basis using grey code projection and coarse-to-fine multi-scale corner analysis and interpolation. In the proposed model, A_i encapsulates the effects of geometric distortion, pixel reconstruction point spread function (PSF) and resample filtering operations.

Furthermore, V_i and *EM* are obtained during calibration by analyzing the camera response for projected black, red, green, and blue flood images of each projector. In principle, this model could be applied to a projection surface with arbitrary colour, texture and shape. However, this has not been shown in [DVC07]. Once the parameters are estimated, equation (14) can be solved numerically using an iterative gradient descent algorithm. This generates optimal results but does not achieve real-time rendering rates.

For real-time sub-frame rendering, it was shown in [DVC07] that near-optimal results can be produced with a non-iterative approximation. This is accomplished by introducing a linear filter bank that consists of impulse responses of the linearly approximated results, which are pre-computed with the non-linear iterative algorithm mentioned above. The filter bank is applied to the original image for estimating the sub-frames.

In an experimental setting, this filtering process is implemented with fragment shaders and real-time rendering is achieved. Figure 21 illustrates a close-up of a single projected sub-frame (a) and four overlapping projections with super-resolution rendering enabled (b). In this experiment, the original image has a higher resolution than any of the sub-frames.

6.3. High-Dynamic Range

To overcome the contrast limitations that are related to radiometric compensation (see Figure 7), high-dynamic range (HDR) projector-camera systems are imaginable. Although
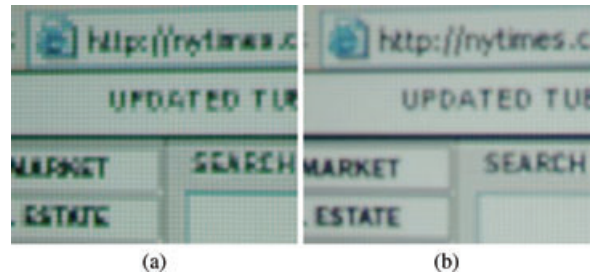

Figure 21: *Experimental result for four superimposed projections: Single sub-frame image (a) and image produced by four superimposed projections with super-resolution enabled. © 2007 IEEE [DVC07].*

there has been much research and development on HDR camera and capturing systems, little work has been done so far on HDR projectors.

In this section, we will focus on state-of-the-art HDR projector technologies rather than on HDR cameras and capturing techniques. A detailed discussion on HDR capturing/imaging technology and techniques, such as recovering camera response functions and tone mapping/reproduction is out of the scope of this report. The interested reader is referred to [RWPD06].

Note, that for the following we want to use the notation of *dynamic range* (unit decibel, dB) for cameras, and the notation of *contrast ratio* (unit-less) for projectors.

The dynamic range of common charge-coupled device or complementary metal oxide semiconductor (CMOS) chips is around 60 dB while recent logarithmic CMOS image sensors for HDR cameras cover a dynamic range of 170 dB (HDRC®, Omron Automotive Electronics GmbH, Stuttgart, Germany). Besides special HDR sensors, low-dynamic rage (LDR) cameras can be applied for capturing HDR images.

The most popular approach to HDR image acquisition involves taking multiple images of the same scene with the same camera using different exposures, and then merging them into a single HDR image.

There are many ways for making multiple exposure measurements with a single camera [DM97] or with multiple coaxially aligned cameras [AA01]. The interested reader is referred to [NB03] for more information. As an alternative to merging multiple LDR images, the exposure of individual sensor pixels in one image can be controlled with additional light modulators, like an LCD panel [NB03] or a DMD chip [NBB04] in front of the sensor or elsewhere within the optical path. In these cases, HDR images are acquired directly.

The contrast ratio of DMD chips and liquid crystal on silicon (LCoS) panels (without additional optics) is about 2000:1 [DDS03] and 5000:1 (SXRD®, Sony Corporation),

respectively. Currently, a contrast ratio of around 15000:1 is achieved for high-end projectors with auto-iris techniques that dynamically adjust the amount of the emitting light according to the image content. Auto-iris techniques, however, cannot expand the dynamic range within a single frame. On the other hand, a laser projection system achieved the contrast ratio of 100,000:1 in [BDD*04] because of the absence of light in dark regions.

Multi-projector systems can enhance spatial resolution (see Section 6.2) and increase the intensity range of projections (see Section 4.1). However, merging multiple LDR projections does not result in an HDR image. Majumder *et al.*, for example, have rendered HDR images with three overlapped projectors to demonstrate that a larger intensity range and resolution will result in higher-quality images [MW01]. Although the maximum intensity level is increased with each additional projector unit, the minimum intensity level (i.e., the black level) is also increased. The contrast of overlapping regions is never greater than the largest one of each individual projector.

Theoretically, if the maximum and the minimum intensities of the ith projector are I_i^{max} and I_i^{min}, its contrast ratio is I_i^{max}/I_i^{min} : 1. If N projectors are overlapped, the contrast ratio of the final image is $\sum_i^N I_i^{max} / \sum_i^N I_i^{min}$: 1. For example, if two projectors are used whose intensities are $I_1^{min} = 10$, $I_1^{max} = 100$ and $I_2^{min} = 100$, $I_2^{max} = 1000$ (thus both contrast ratios are 10 : 1), the contrast ratio of the image overlap is still 10 : 1 ($10 = (I_1^{max} + I_2^{max})/(I_1^{min} + I_2^{min})$).

Recently, HDR display systems have been proposed that combine projectors and external light modulators. Seetzen *et al.* proposed an HDR display that applies a projector as a backlight of an LCD panel instead of a fluorescent tube assembly [SHS*04]. As in Figure 22a, the projector is directed to the rear of a transmissive LCD panel. The light that corresponds to each pixel on the HDR display is effectively modulated twice: first by the projector and then by the LCD panel. Theoretically, the final contrast ratio is the product of the individual contrast ratio of the two modulators. If a projector with a contrast ratio of c_1 : 1 and an LCD panel with a contrast ratio of c_2 : 1 are used in this example, the contrast of the combined images is $(c_1 \cdot c_2)$: 1. In an experimental setup, this approach achieved a contrast ratio of 54,000 : 1 using an LCD panel and a DMD projector with a contrast ratio of 300 : 1 and 800 : 1, respectively. The reduction of contrast is due to noise and imperfections in the optical path.

The example described above does not really present a projection system since the image is generated behind an LCD panel, rather than on a projection surface. True HDR projection approaches are discussed in [DRW*06, DSW*07]. The basic idea of realizing an HDR projector is to combine a normal projector and an additional low-resolution light modulating device. Double modulation decreases the black level of the projected image, and increases the dynamic range as
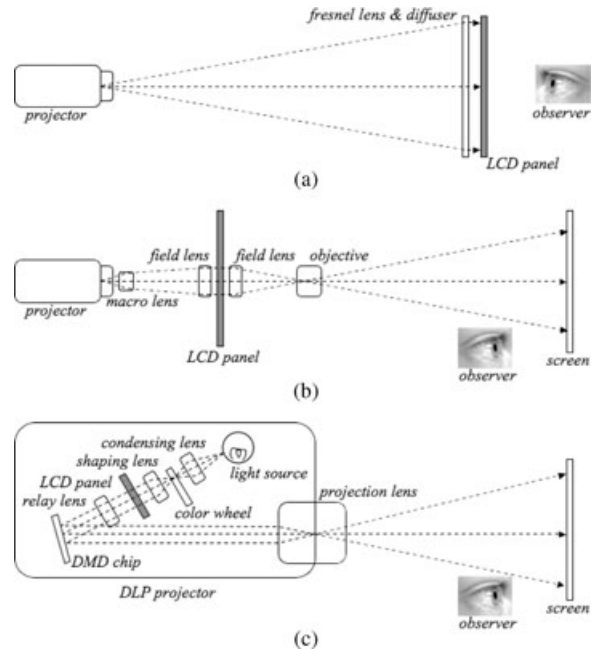

Figure 22: *Different HDR projection setups: using a projector as backlight of an LCD (a), modulating the image path (b) and modulating the illumination path (c).*

well as the number of addressable intensity levels. Thereby, LCD panels, LCoS panels, DMD chips can serve as light modulators.

HDR projectors can be categorized into systems that modulate the image path (cf. Figure 22b), and into systems that modulate the illumination path (22c). In the first case, an image is generated with a high-resolution light modulator first, and then modulated again with an additional low-resolution light modulator. In the latter case, the projection light is modulated in advance with a low-resolution light modulator before the image is generated with a high-resolution modulator.

In each approach, a compensation for the optical blur caused by the low-resolution modulator is required. The degree of blur can be measured and can be described with a PSF for each low-resolution pixel in relation to corresponding pixels on the higher-resolution modulator. A division of the desired output image by the estimated blurred image that is simulated by the PSF will result in the necessary compensation mask, which will be displayed on the high-resolution modulator.

Pavlovych *et al.* proposed a system that falls into the first category [PS05]. This system uses an external attachment (an LCD panel) in combination with a regular DLP projector (cf. Figure 22b). The projected image is resized and focused first on the LCD panel through a set of lenses. Then it is
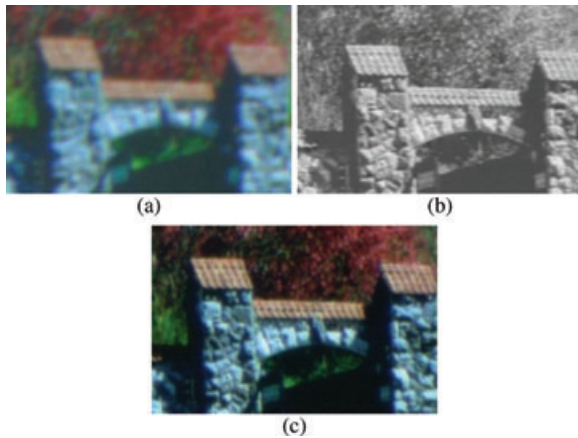
Figure 23: *Photographs of a part of an HDR-projected image: image modulated with low-resolution chrominance modulators (a), image modulated with a high-resolution luminance modulator (b), output image (c). © 2006 ITE/SID [KKN*06].*

modulated by the LCD panel and projected through another lens system onto a larger screen.

Kusakabe *et al.* proposed an HDR projector that applies LCoS panels that falls into the second category [KKN*06]. In this system, three low-resolution (RGB) modulators are used first for chrominance modulation of the projection light. Finally, the light is modulated again with a high-resolution luminance modulator which forms the image.

The resolution of the panel that is applied for chrominance modulation can be much lower than the one for luminance modulation because the human visual system is sensitive to a relatively low chrominance contrast. An experimental result is shown in Figure 23. The proposed projector has a contrast ratio of 1,100,000 : 1.

6.4. High Speed

High-speed projector-camera systems hold the enormous potential to significantly improve high-frequent temporal-coded projections (see Sections 3.3 and 4.2). They enable, for instance, projecting and capturing imperceptible spatial patterns that can be efficiently used for real-time geometric registration, fast shape measurement and real-time adaptive radiometric compensation while a flicker-free content is perceived by the observer at the same time. The faster the projection and the capturing process can be carried out, the more information per unit of time can be encoded. Since high-speed capturing systems are well established, this section focuses mainly on the state-of-the-art of high-speed projection systems. Both together, however, could be merged into

future high-speed projector-camera systems. For this reason, we first want to give only a brief overview of high-speed capturing systems.

Commercially available single-chip high-speed cameras exist that can record 512x512 colour pixels at up to 16,000 fps (FASTCAM SA1, Photron Ltd., Tokyo, Japan). However, these systems are typically limited to storing just a few seconds of data directly on the camera because of the huge bandwidth that is necessary to transfer the images. Other CMOS devices are on the market that enable a 500 fps (A504k, Basler AG, Ahrensburg, Germany) capturing and transfer rates.

Besides such single-camera systems, a high capturing speed can also be achieved with multi-camera arrays. Wilburn *et al.*, for example, proposed a high-speed video system for capturing 1,560 fps videos using a dense array of 30 fps CMOS image sensors [WJV*04]. Their system captures and compresses images from 52 cameras in parallel. Even at extremely high frame-rates, such a camera array architecture supports continuous streaming to disk from all of the cameras for minutes.

In contrast to this, however, the frame-rate of commercially available DLP projectors is normally less than or equal to 120 fps (DepthQ®, InFocus Corporation, Wilsonville, OR, USA). Although faster projectors that can be used in the context of our projector-camera system are currently not available, we want to outline several projection approaches that achieve higher frame-rates – but do not necessarily allow the projection of high-quality images.

Raskar *et al.*, for instance, developed a high-speed optical motion capture system with an LED-based code projector [RNdD*07]. The system consists of a set of 1-bit grey code infrared LED beamers. Such a beamer array is effectively emitting 10,000 binary grey-coded patterns per second, and is applied for object tracking. Each object to be tracked is tagged with a photosensor that detects and decodes the temporally projected codes. The 3D location of the tags can be computed at a speed of 500 Hz when at least three such beamer arrays are applied.

In contrast to this approach which does not intent to project pictorial content in addition to the code patterns, Nii *et al.* proposed a visible light communication (VLC) technique that does display simple images [NSI05]. They developed an LED-based high-speed projection system (with a resolution of 4x5 points produced with an equally large LED matrix) that is able to project alphabetic characters while applying an additional pulse modulation for coding information that is detected by photosensors. This system is able to transmit two data streams with 1 kHz and 2 kHz, respectively at different locations while simultaneously projecting simple pictorial content. Although LEDs can be switched with a high speed (e.g., the LEDs in [NSI05] are temporally modulated at 10.7 MHz), such simple LED-based projection systems offer a too low spatial resolution at the moment.

In principle, binary frame-rates of up to 16,300 fps can currently be achieved with DMDs for a resolution of 1024x768. The DMD discovery board enables developers to implement their own mirror timings for special purpose application [DDS03]. Consequently, due to this high binary frame-rate some researchers utilized the Discovery boards for realizing high-speed projection techniques. McDowall *et al.*, for example, demonstrated the possibility of projecting 24 binary code and compensation images at a speed of 60 Hz [MBHF04]. Viewers used time-encoded shutter glasses to make individual images visible.

Kitamura *et al.* also developed a high-speed projector based on the DMD discovery board [KN06]. In their approach, photosensors can be used to detect temporal code patterns that are embedded into the mirror flip sequence. In contrast to the approach by Cotting *et al.* [CNGF04] that was described in Section 3.3, the mirror flip sequence can be freely re-configured.

The results of an initial basic experiment with this system are shown in Figure 24a: The projected image is divided into 10 regions. Different on/off mirror flip frequencies are used in each region (from 100 Hz to 1000 Hz at 100 Hz intervals), while a uniformly bright image with a 50% intensity appears in all regions – regardless of the locally applied frequencies. The intensity fall-off in the projection is mainly due to imperfections in applied optics. The signal waves are received by photosensors that are placed within the regions. They can detect the individual frequency.

Instead of using a constant on-off flip frequency for each region, binary codes can be embedded into a projected frame. This is illustrated in Figure 24b: For a certain time slot of T, the first half of the exposure sequence contains a temporal code pattern (modulated with different mirror flip states) that is compensated with the second half of the exposure sequence to modulate a desired intensity. Yet, contrast is lost in this case due to the modulated intensity level created by the code pattern. Here, the number of on-states always equals the number of off-states in the code period. This leads to a constant minimum intensity level of 25%. Since also 25% of the off states are used during this period, intensity values between 25% and 75% can only be displayed.

All systems that have been outlined above, apply photosensors rather than cameras. Thus, they cannot be considered as suitable projector-camera systems in our application context. Yet, McDowall *et al.* combined their high-speed projector with a high-speed camera to realize fast range scanning [MB05]. Takei *et al.* proposed a 3,000 fps shape measurement system (shape reconstruction is performed off-line in this case) [TKH07].

In an image-based rendering context, Jones *et al.* proposed to simulate spatially varying lighting on a live performance based on a fast shape measurement using a high-
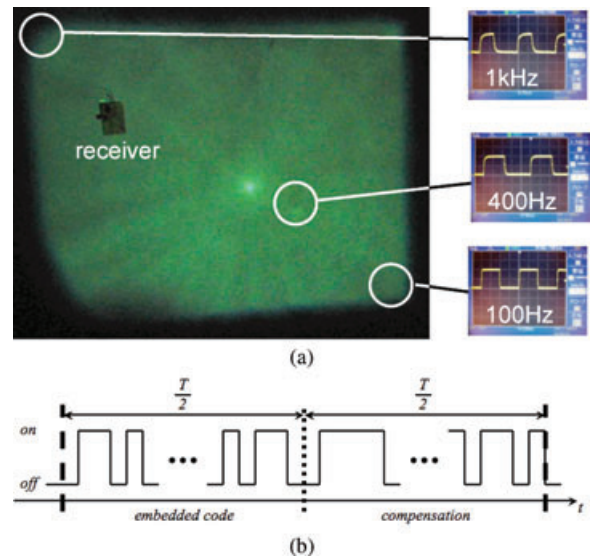

Figure 24: *Regionally different mirror flip frequencies and corresponding signal waves received by photosensors at different image areas. The overall image appears mostly uniform in intensity (a). Binary codes can be embedded into the first half of the exposure sequence while the second half can compensate the desired intensity (b). © 2006 IPSJ [KN06].*

speed projector-camera system [JGB*06]. However, all of these approaches do not project pictorial image content, but rather represent encouraging examples of fast projector-camera techniques.

The mirrors on a conventional DMD chip can be switched much faster than alternative technologies, such as ordinary LCD or LCoS panels whose refresh rate can be up to 2.5 ms (= 400 Hz) at the moment.

LEDs are generally better suited for high-speed projectors than a conventional ultra high performance (UHP) lamp (we do not want to consider brightness issues for the moment), because three or more different LEDs that correspond to each colour component can be switched at a high speed (even faster than a DMD) for modulating colours and intensities. Therefore, a combination of DMD and LED technologies seems to be optimal for future projection units.

Let's assume that the mirrors of a regular DLP projector can be switched at 15 μs (= 67000 binary frames per second). For projecting 256 different intensity levels (i.e., an 8 bit encoded grey scale image), the grey scale frame rate is around 260 Hz (= 67000 binary frames per second/256 intensity levels). Consequently, the frame rate for full colour images is around 85 Hz (= 260 grey scale frames per second/3 colour channels) if the colour wheel consists of three filter segments.
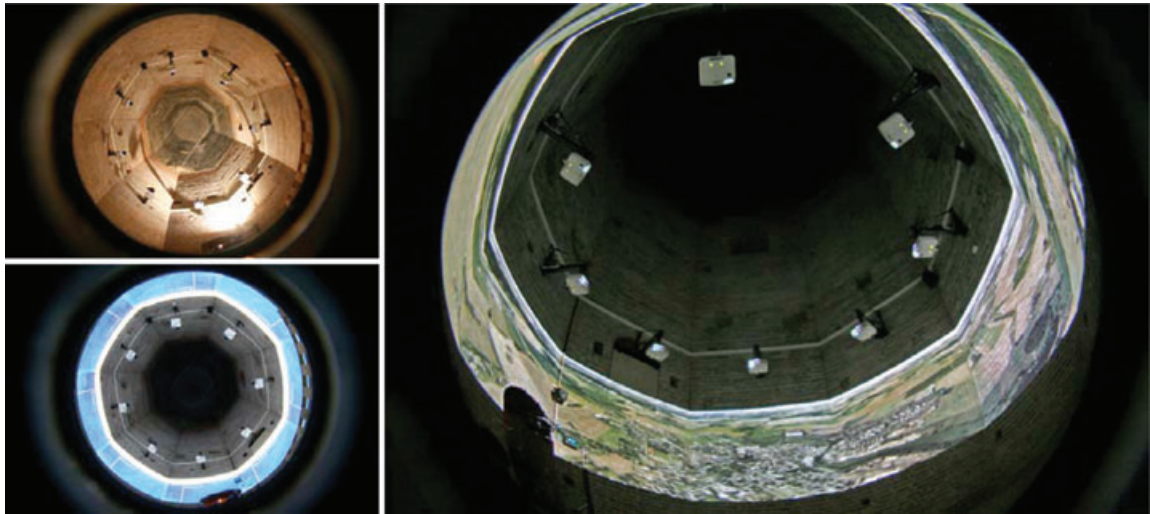
Figure 25: *360 degree surround projection onto the natural stone walls of castle Osterburg in Weida, Germany.*

Now, let's consider DLP projectors that apply LEDs instead of a UHP lamps and a colour wheel. If, for example, the intensities of three (RGB) colour LEDs can be switched between eight different levels (1,2,4,8,16,32,64,128) at a high speed, a full colour image can theoretically be projected at around 2,800 Hz (= 67000 binary frames per second/8 (8-bit encoded) intensity levels/3 colour channels).

To overcome the bandwidth limitation for transferring the huge amount of image data in high speed, the multi use light engine projector adopts a custom-programmed field programmable gate array (FPGA)-based circuitry [JMY*07]. The FPGA decodes a standard DVI signal from the graphics card. Instead of rendering a colour image, the FPGA takes each 24 bit colour frame of video and displays each bit sequentially as separate frames. Thus, if the incoming digital video signal is 60 Hz, the projector displays $60 \times 24 = 1,440$ frames per second. To achieve even faster rates, the refresh rate of a video card is set at 180–240 Hz. At 200 Hz, for instance, the projector can display 4,800 binary frames per second.

7. Conclusion

This article reviewed the state-of-the-art of projector-camera systems with a focus on real-time image correction techniques that enable projections onto non-optimized surfaces. It did not discuss projector-camera-related areas, such as camera-supported photometric calibration of conventional projection displays (e.g. [BMY05], [JM07], [BM07]), real-time shadow removal techniques (e.g. [STJS01], [JWS*01], [JWS04]), or projector-camera-based interaction approaches (e.g. [Pin01], [EHH04], [FR06]).

While most of the presented techniques are still on a research level, others found already practical applications in

Figure 26: *Outdoor projection onto a building front in Leinefelde, Germany.*

Figure 27: *Projecting onto the stage settings during live performance of the Karl May festival in Elspe, Germany.*
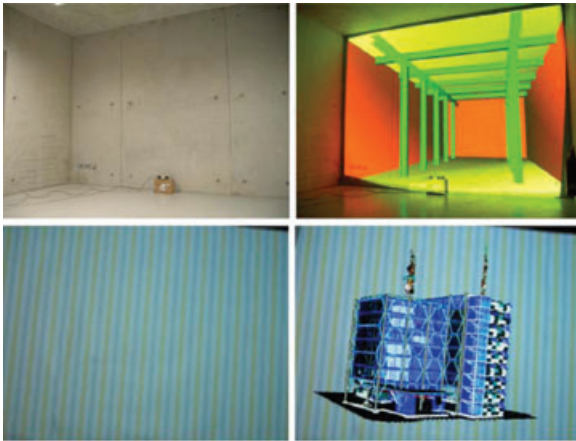

Figure 28: *View-dependent stereoscopic projection onto concrete walls (top) and onto wall paper (bottom). © 2005 IEEE [BWEN05].*

Figure 29: *Projection onto an iron plate in the Henrichshütte in Hattingen, Gemany (top). View-dependent stereoscopic projection onto a natural stone wall (bottom). © 2005 IEEE [BWEN05]*

Imagining projector-camera technology to be integrated into, or coupled with mobile devices, such as cellphones or laptops, will support a truly flexible way for presentations. There is no doubt that this technology is on its way. Yet, one question needs to be addressed when thinking about mobile projectors: What can we project onto, without carrying around screen canvases? It is clear that the answer to this question can only be: Onto available everyday surfaces. With this in mind, the future importance of projector-camera systems in combination with appropriate image correction techniques becomes clear.

theatres, museums, historic sites, open-air festivals, trade shows and advertisement. Some examples are shown in Figures 25–27.

Future projectors will become more compact in size and will require little power and cooling. 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. They will integrate GPUs for real-time graphics and vision processing. While resolution and contrast will keep increasing, production costs and market prizes will continue to fall. Conventional UHP lamps will be replaced by powerful LEDs or multi-channel lasers. This will make them suitable for mobile applications.
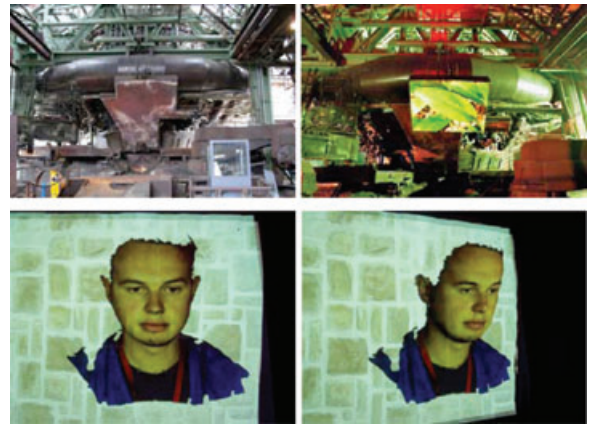
References

[AA01]   AGGARWAL M., AHUJA N.: Split Aperture imaging for high dynamic range. In *Proc. of IEEE International Conference on Computer Vision (ICCV)* (2001), vol. 2, pp. 10–17.

[AOSS06]   ASHDOWN M., OKABE T., SATO I., SATO Y.: Robust content-dependent photometric projector compensation. In *Proc. of IEEE International Workshop on Projector-Camera Systems (ProCams)* (2006).

[ASOS07] ASHDOWN M., SATO I., OKABE T., SATO Y.: Perceptual photometric compensation for projected images. *IEICE Transaction on Information and Systems J90-D*, 8 (2007), 2115–2125 (in Japanese).

[AU05] ALLEN W., ULICHNEY R.: Wobulation: Doubling the addressed resolution of projection displays. In *Proc. of SID Symposium Digest of Technical Papers* (2005), vol. 36, pp. 1514–1517.

[BCK∗05] BIMBER O., CORIAND F., KLEPPE A., BRUNS E., ZOLLMANN S., LANGLOTZ T.: Superimposing pictorial artwork with projected imagery. *IEEE MultiMedia 12*, 1 (2005), 16–26.

[BDD∗04] BIEHLING W., DETER C., DUBE S., HILL B., HELLING S., ISAKOVIC K., KLOSE S., SCHIEWE M.: LaserCave – some building blocks for immersive screens. In *Proc. of International Status Conference Virtual and Augmented Reality* (2004).

[BE06] BIMBER O., EMMERLING A.: Multifocal projection: A multiprojector technique for increasing focal depth. *IEEE Transactions on Visualization and Computer Graphics (TVCG) 12*, 4 (2006), 658–667.

[BEK05] BIMBER O., EMMERLING A., KLEMMER T.: Embedded entertainment with smart projectors. *IEEE Computer 38*, 1 (2005),56–63.

[BGZ∗06] BIMBER O., Grundhöfer A., ZEIDLER T., DANCH D., KAPAKOS P.: Compensating indirect scattering for immersive and semi-immersive projection displays. In *Proc. of IEEE Virtual Reality (IEEE VR)* (2006), pp. 151–158.

[Bim06] BIMBER O.: Projector-based augmentation. In *Emerging Technologies of Augmented Reality: Interfaces and Design*, Haller M., Billinghurst M., Thomas B., (Eds.). Idea Group, 2006, pp. 64–89.

[BJM07] BHASKER E. S., JUANG R., MAJUMDER A.: Registration techniques for using imperfect and partially calibrated devices in planar multi-projector displays. *IEEE Transactions on Visualization and Computer Graphics. 13*, 6 (2007), 1368–1375.

[BM07] BHASKER E., MAJUMDER A.: Geometric modeling and calibration of planar multi-projector displays using rational Bezier patches. In *Proc. of IEEE International Workshop on Projector-Camera Systems (ProCams)* (2007).

[BMS98] BATLLE J., MOUADDIB E. M., SALVI J.: Recent progress in coded structured light as a technique to solve the correspondence problem: a survey. *Pattern Recognition 31*, 7 (1998), 963–982.

[BMY05] BROWN M., MAJUMDER A., YANG R.: Camera based calibration techniques for seamless multi-projector displays. *IEEE Transactions on Visualization and Computer Graphics (TVCG) 11*, 2 (2005), 193–206.

[BSC06] BROWN M. S., SONG P., CHAM T.-J.: Image preconditioning for out-of-focus projector blur. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2006), vol. II, pp. 1956–1963.

[BWEN05] BIMBER O., WETZSTEIN G., EMMERLING A., NITSCHKE C.: Enabling view-dependent stereoscopic projection in real environments. In *Proc. of IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR)* (2005), pp. 14–23.

[CKS98] CASPI D., KIRYATI N., SHAMIR J.: Range imaging with adaptive colour structured light. *IEEE Transactions on Pattern Analysis and Machine Intelligence 20*, 5 (1998), 470–480.

[CNGF04] COTTING D., NÄF M., GROSS M. H., FUCHS H.: Embedding imperceptible patterns into projected images for simultaneous acquisition and display. In *Proc. of IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR)* (2004), pp. 100–109.

[CZGF05] COTTING D., ZIEGLER R., GROSS M. H., FUCHS H.: Adaptive instant displays: Continuously calibrated projections using per-pixel light control. In *Proc. of Eurographics* (2005), pp. 705–714.

[DDS03] DUDLEY D., DUNCAN W. M., SLAUGHTER J.: Emerging digital micromirror device (DMD) applications. In *Proc. of SPIE* (2003), vol. 4985, pp. 14–25.

[DM97] DEBEVEC P. E., MALIK J.: Recovering high dynamic range radiance maps from photographs. In *Proc. of ACM SIGGRAPH* (1997), pp. 369–378.

[DRW∗06] DEBEVEC P., REINHARD E., WARD G., MYSZKOWSKI K., SEETZEN H., ZARGARPOUR H., MCTAGGART G., HESS D.: High dynamic range imaging: Theory and applications. In *Proc. of ACM SIGGRAPH (Courses)* (2006).

[DSW∗07] DAMBERG G., SEETZEN H., WARD G., HEIDRICH W., WHITEHEAD L.: High-dynamic-range projection systems. In *Proc. of SID Symposium Digest of Technical Papers* (2007), vol. 38, pp. 4–7.

[DVC07] Damera-Venkata N., CHANG N. L.: Realizing super-resolution with superimposed projection. In *Proc. of IEEE International Workshop on Projector-Camera Systems (ProCams)* (2007).

[EHH04] EHNES J., HIROTA K., HIROSE M.: Projected augmentation – augmented reality using rotatable video

projectors. In *Proc. of IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR)* (2004), pp. 26–35.

[FGN05] FUJII K., GROSSBERG M., NAYAR S.: A projector-camera system with real-time photometric adaptation for dynamic environments. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2005), vol. I, pp. 814–821.

[FR06] FLAGG M., REHG J. M.: Projector-guided painting. In *Proc. of ACM Symposium on User Interface Software and Technology (UIST)* (2006), pp. 235–244.

[GB07] GRUNDHOEFER A., BIMBER O.: Real-time adaptive radiometric compensation. *Transactions on Visualization and Computer Graphics (TVCG) 14*, 1 (2008), 97–108.

[GB08] GROSSE M., BIMBER O.: Coded aperture projection. In *Proc. of Emerging Display Technologies and Immersive Projection Technologies 2008 (EDT IPT08)* (2008).

[GPNB04] GROSSBERG M., PERI H., NAYAR S., BELHUMEUR P.: Making one object look like another: Controlling appearance using a projector-camera system. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Jun 2004), vol. I, pp. 452–459.

[GSHB07] GRUNDHÖER A., SEEGER M., HÄNTSCH F., BIMBER O.: Dynamic adaptation of projected imperceptible codes. *Proc. of IEEE International Symposium on Mixed and Augmented Reality* (2007).

[HSM07] HABE H., SAEKI N., MATSUYAMA T.: Inter-reflection compensation for immersive projection display. In *Proc. of IEEE International Workshop on Projector-Camera Systems (ProCams) (poster)* (2007).

[JF07] JOHNSON T., FUCHS H.: Real-time projector tracking on complex geometry using ordinary imagery. In *Proc. of IEEE International Workshop on Projector-Camera Systems (ProCams)* (2007).

[JGB∗06] JONES A., GARDNER A., BOLAS M., MCDOWALL I., DEBEVEC P.: Simulating spatially varying lighting on a live performance. In *Proc. of European Conference on Visual Media Production (CVMP)* (2006), pp. 127–133.

[JM07] JUANG R., MAJUMDER A.: Photometric self-calibration of a projector-camera system. In *Proc. of IEEE International Workshop on Projector-Camera Systems (ProCams)* (2007).

[JMY∗07] JONES A., MCDOWALL I., YAMADA H., BOLAS M., DEBEVEC P.: Rendering for an interactive 360◦ light field display. In *Proc. of ACM SIGGRAPH* (2007).

[JR03] JAYNES C., RAMAKRISHNAN D.: Super-resolution composition in multi-projector displays. In *Proc. of IEEE International Workshop on Projector-Camera Systems (ProCams)* (2003).

[JWS∗01] JAYNES C., WEBB S., STEELE R., BROWN M., SEALES W.: Dynamic shadow removal from front projection displays. In *Proc. of IEEE Visualization* (2001), pp. 175–555.

[JWS04] JAYNES C., WEBB S., STEELE R. M.: Camera-based detection and removal of shadows from interactive multiprojector displays. *IEEE Transactions on Visualization and Computer Graphics (TVCG) 10*, 3 (2004), 290–301.

[KKN∗06] KUSAKABE Y., KANAZAWA M., NOJIRI Y., FURUYA M., YOSHIMURA M.: YC-separation type projector with double modulation. In *Proc. of International Display Workshop (IDW)* (2006), pp. 1959–1962.

[KN06] KITAMURA M., NAEMURA T.: A study on position-dependent visible light communication using DMD for ProCam. In *IPSJ SIG Notes. CVIM-156* (2006), pp. 17–24 (in Japanese).

[MB05] MCDOWALL I. E., BOLAS M.: Fast light for display, sensing and control applications. In *Proc. of IEEE VR 2005 Workshop on Emerging Display Technologies (EDT)* (2005), pp. 35–36.

[MBHF04] MCDOWALL I. E., BOLAS M. T., HOBERMAN P., FISHER S. S.: Snared illumination. In *Proc. of ACM SIGGRAPH (Emerging Technologies)* (2004), p. 24.

[MKO06] MUKAIGAWA Y., KAKINUMA T., OHTA Y.: Analytical compensation of inter-reflection for pattern projection. In *Proc. of ACM Symposium on Virtual Reality Software and Technology (VRST) (short paper)* (2006), pp. 265–268.

[MW01] MAJUMDER A., WELCH G.: Computer Graphics Optique: Optical superposition of projected computer graphics. In *Proc. of Immersive Projection Technology – Eurographics Workshop on Virtual Environment (IPT-EGVE)* (2001).

[NB03] NAYAR S. K., BRANZOI V.: Adaptive dynamic range imaging: Optical control of pixel exposures over space and time. In *Proc. of IEEE International Conference on Computer Vision (ICCV)* (2003), vol. 2, pp. 1168–1175.

[NBB04] NAYAR S. K., BRANZOI V., BOULT T. E.: Programmable imaging using a digital micromirror array. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2004), vol. I, pp. 436–443.

[NKGR06] NAYAR S., KRISHNAN G., GROSSBERG M. D., RASKAR R.: Fast separation of direct and global components of a scene using high frequency illumination. In *Proc. of ACM SIGGRAPH* (2006), pp. 935–944.

[NPGB03] NAYAR S. K., PERI H., GROSSBERG M. D., BELHUMEUR P. N.: A projection system with radiometric compensation for screen imperfections. In *Proc. of IEEE International Workshop on Projector-Camera Systems (ProCams)* (2003).

[NSI05] NII H., SUGIMOTO M., INAMI M.: Smart light-ultra high speed projector for spatial multiplexing optical transmission. In *Proc. of IEEE International Workshop on Projector-Camera Systems (ProCams)* (2005).

[OS07] OYAMADA Y., SAITO H.: Focal pre-correction of projected image for deblurring screen image. In *Proc. of IEEE International Workshop on Projector-Camera Systems (ProCams)* (2007).

[Pin01] PINHANEZ C.: Using a steerable projector and a camera to transform surfaces into interactive displays. In *Proc. of CHI (extended abstracts)* (2001), pp. 369–370.

[PLJP07] PARK H., LEE M.-H., JIN B.-K. S. Y., PARK J.-I.: Content adaptive embedding of complementary patterns for nonintrusive direct-projected augmented reality. In *HCI International 2007* (2007), vol. 14, pp. 132–141.

[PLKP05] PARK H., LEE M.-H., KIM S.-J., PARK J.-I.: Specularity-free projection on nonplanar surface. In *Proc. of Pacific-Rim Conference on Multimedia (PCM)* (2005), pp. 606–616.

[PLKP06] PARK H., LEE M.-H., KIM S.-J., PARK J.-I.: Contrast enhancement in direct-projected augmented reality. In *Proc. of IEEE International Conference on Multimedia and Expo (ICME)* (2006).

[PLS∗06] PARK H., LEE M.-H., SEO B.-K., SHIN H.-C., PARK J.-I.: Radiometrically-compensated projection onto non-Lambertian surface using multiple overlapping projectors. In *Proc. of Pacific-Rim Symposium on Image and Video Technology (PSIVT)* (2006), pp. 534–544.

[PPK03] PARK S. C., PARK M. K., KANG M. G.: Super-resolution image reconstruction: A technical overview. *IEEE Signal Processing Magazine 20*, 3 (2003), 21–36.

[PS05] PAVLOVYCH A., STUERZLINGER W.: A high-dynamic range projection system. In *Proc. of SPIE* (2005), vol. 5969.

[Ras99] RASKAR R.: Oblique projector rendering on planar surfaces for a tracked user. In *Proc. of ACM SIGGRAPH (Sketches and Applications)* (1999).

[RBvB∗04] RASKAR R., BEARDSLEY P., van Baar J., WANG Y., DIETZ P., LEE J., LEIGH D., WILLWACHER T.: RFIG Lamps: Interacting with a self-describing world via photosensing wireless tags and projectors. In *Proc. of ACM SIGGRAPH* (2004), pp. 406–415.

[RBY∗99] RASKAR R., BROWN M., YANG R., CHEN W., WELCH G., TOWLES H., SEALES B., FUCHS H.: Multi-projector displays using camera-based registration. In *Proc. of IEEE Visualization* (1999), pp. 161–168.

[RNdD∗07] RASKAR R., NII H., de Decker B., HASHIMOTO Y., SUMMET J., MOORE D., ZHAO Y., WESTHUES J., DIETZ P., INAMI M., NAYAR S. K., BARNWELL J., NOLAND M., BEKAERT P., BRANZOI V., BRUNS E.: Prakash: Lighting aware motion capture using photosensing markers and multiplexed illuminators. In *Proc. of ACM SIGGRAPH* (2007).

[RPG99] RAMASUBRAMANIAN M., PATTANAIK S. N., GREENBERG D. P.: A perceptually based physical error metric for realistic image synthesis. In *Proc. of ACM SIGGRAPH* (1999), pp. 73–82.

[RvBWR04][RvBWR04] RASKAR R., van Baar J., WILLWACHER T., RAO S.: Quadric transfer for immersive curved screen displays. In *Proc. of Eurographics* (Aug 2004).

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

[RWPD06] REINHARD E., WARD G., PATTANAIK S., DEBEVEC P.: *High Dynamic Range Imaging - Acquisition, Display and Image-Based Lighting*. Morgan Kaufmann, 2006.

[SCG∗05] SEN P., CHEN B., GARG G., MARSCHNER S. R., HOROWITZ M., LEVOY M., LENSCH H. P. A.: Dual photography. In *Proc. of ACM SIGGRAPH* (2005), pp. 745–755.

[SHS∗04] SEETZEN H., HEIDRICH W., STUERZLINGER W., WARD G., WHITEHEAD L., TRENTACOSTE M., GHOSH A., VOROZCOVS A.: High dynamic range display systems. In *Proc. of ACM SIGGRAPH* (2004), pp. 760–768.

[SMK05] SEITZ S. M., MATSUSHITA Y., KUTULAKOS K. N.: A Theory of inverse light transport. In *Proc. of IEEE International Conference on Computer Vision (ICCV)* (2005), vol. 2, pp. 1440–1447.

[SMO03] SHIRAI Y., MATSUSHITA M., OHGURO T.: HIEI Projector: Augmenting a real environment with invisible information. In *Proc. of Workshop on Interactive Systems and Software (WISS)* (2003), pp. 115–122. (in Japanese).

[SPB04] SALVI J., PAGÈS J., BATLLE J.: Pattern codification strategies in structured light systems. *Pattern Recognition 37*, 4 (2004), 827–849.

[STJS01] SUKTHANKAR R., Tat-Jen C., SUKTHANKAR G.: Dynamic shadow elimination for multi-projector displays. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2001), vol. II, pp. 151–157.

[TKH07] TAKEI J., KAGAMI S., HASHIMOTO K.: 3,000-fps 3-D Shape measurement using a high-speed camera-projector system. In *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2007), pp. 3211–3216.

[VVSC05] VIEIRA M. B., VELHO L., SA A., CARVALHO P. C.: A camera-projector system for real-time 3D video. In *Proc. of IEEE International Workshop on Projector-Camera Systems (ProCams)* (2005).

[WB07] WETZSTEIN G., BIMBER O.: Radiometric compensation through inverse light transport. *Proc. of Pacific Graphics* (2007).

[WJV∗04] WILBURN B., JOSHI N., VAISH V., LEVOY M., HOROWITZ M.: High-speed videography using a dense camera array. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2004), vol. II, pp. 294–301.

[WJV∗05] WILBURN B., JOSHI N., VAISH V., TALVALA E.-V., ANTUNEZ E., BARTH A., ADAMS A., HOROWITZ M., LEVOY M.: High performance imaging using large camera arrays. In *Proc. of ACM SIGGRAPH* (2005), pp. 765–776.

[WSOS05] WANG D., SATO I., OKABE T., SATO Y.: 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)* (2005).

[WWC∗05] WASCHBÜSCH M., WÜRMLIN S., COTTING D., SADLO F., GROSS M. H.: Scalable 3D video of dynamic scenes. *The Visual Computer 21*, 8-10 (2005), 629–638.

[YHS03] YOSHIDA T., HORII C., SATO K.: A virtual colour reconstruction system for real heritage with light projection. In *Proc. of International Conference on Virtual Systems and Multimedia (VSMM)* (2003), pp. 161–168.

[YW01] YANG R., WELCH G.: Automatic and continuous projector display surface calibration using every-day imagery. In *Proc. of International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG)* (2001).

[ZB07] ZOLLMANN S., BIMBER O.: Imperceptible calibration for radiometric compensation. In *Proc. of Eurographics (short paper)* (2007), pp. 61–64.

[ZLB06] ZOLLMANN S., LANGLOTZ T., BIMBER O.: Passive-active geometric calibration for view-dependent projections onto arbitrary surfaces. *Proc. of Workshop on Virtual and Augmented Reality of the GI-Fachgruppe AR/VR 2006 (re-print to appear in Journal of Virtual Reality and Broadcasting 2007)* (2006).

[ZN06] ZHANG L., NAYAR S. K.: Projection defocus analysis for scene capture and image display. In *Proc. of ACM SIGGRAPH* (2006), pp. 907–915.