Journal of Virtual Reality and Broadcasting, Volume 3(2006), no. 8
JVRB EuroITV 2006 Special Issue, no. 1

Digital Illumination for Augmented Studios

Oliver Bimber, Anselm Grundhöfer, Stefanie Zollmann and Daniel Kolster
Bauhaus-University Weimar,
Bauhausstr. 11, 99423 Weimar, Germany
bimber, grundhoefer, zollmann, kolster@medien.uni-weimar.de
www.uni-weimar.de/medien/AR

Abstract

Virtual studio technology plays an important role for modern television productions. Blue-screen matting is a common technique for integrating real actors or moderators into computer generated sceneries. Augmented reality offers the possibility to mix real and virtual in a more general context. This article proposes a new technological approach for combining real studio content with computer-generated information. Digital light projection allows a controlled spatial, temporal, chrominance and luminance modulation of illumination opening new possibilities for TV studios.

1 Introduction

Many modern TV productions apply virtual studio technology. A good overview can be found at [GAB$^+$98]. Chromakeying is the principle method for superimposing the live captured or recorded video signal of a physical blue screen studio with virtual content. Thereby, the video signal is analyzed and video pixels with a predefined color (e.g. blue or green) are replaced by computer-generated graphics. This allows using image processing techniques to efficiently separate the foreground from the background, and consequently to integrate real objects (such as an actor or a moderator) seamlessly into a purely virtual environment.

Invited article, exclusively created for the TV Conference EuroITV 2006 Special Issue of JVRB

Blue screen techniques, however, limit virtual studio technology to special recording environments. Therefore, recent research initiatives investigate the potential of augmented reality (AR) for TV productions. In contrast to blue screen studios, fully equipped real television studios are augmented with virtual content by superimposing the recorded video stream with computer graphics. According to virtual studios, we want to refer to this as *augmented studios*.

Several groups have already shown the advantages of augmented reality in a studio production context: Yama et al. [YMF$^+$02], for instance, augment 360° ultra high-definition omnidirectional images of artificial backgrounds being distorted in real-time relative to the rotation of a pan-tilt camera, and being occluded by a real actor. An axi-vision camera [KIA$^+$00] is used for simultaneously capturing color and depth information per pixel.

Recent examples are also shown in the context of the EU funded project MATRIS [IST04]: Frahm et al. [FKGK05] use a fish eye camera in addition to a studio camera. While the studio camera records the video content to be augmented, the fish eye camera observes the upper hemisphere to track the installed studio lights. Applying a structure from motion algorithm [KKSES05] to both images makes the estimation of the studio cameras pose possible. Standard stereo algorithms [KPG98] allow reconstructing the depth of the studio setting and consequently enable

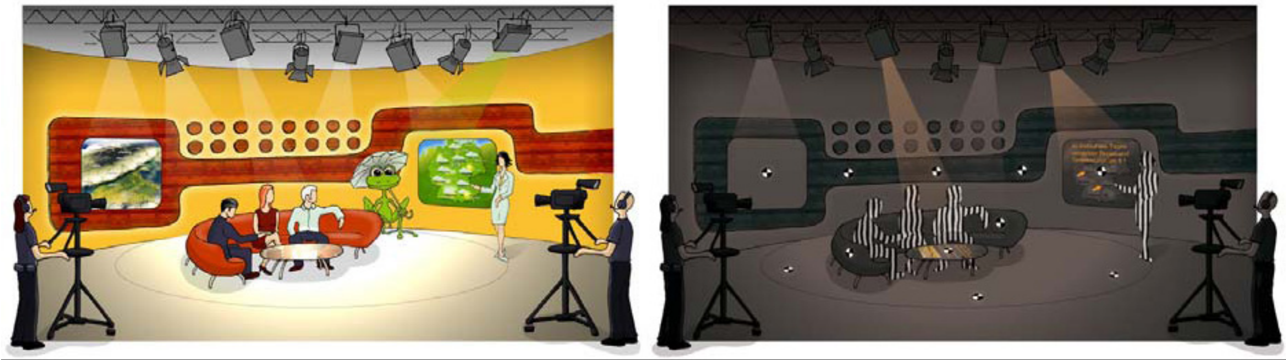

Figure 1: Digital studio illumination: Synthetic illumination of a studio environment with projected spatial augmentations and video augmentations (left). Embedded imperceptible patterns for scene analysis and camera tracking (right).

correct occlusion effects between real and virtual objects. Furthermore, the knowledge of the real studio light sources allows computing a light map that ensures a consistent illumination and shadowing.

Virtual and augmented studio productions have to solve several technical challenges. One of them is the robust and fast *tracking* of the studio cameras [BT03]. Some approaches apply special tracking hardware, while others try to estimate the cameras pose by observing natural features (e.g., ceiling-mounted studio lights or the studio content itself) or artificial tags [TJU97] with additional cameras, as explained above. Optical tracking is becoming more and more popular due to its robustness against most environmental disturbances, speed and precision. Optical tracking approaches can be categorized into *marker-less* and *markerbased* methods. Marker-less techniques, on the one hand, strongly rely on the robust detection of natural scene features [FKGK05]. They will fail for uniformly structured surfaces or under dim lighting conditions. This limits the application of such techniques in TV studios to optimized situations. Marker-based tracking, on the other hand, provides artificial visual features by integrating detectable marker tags. However, these markers should neither be directly visible within the studio environment nor appear in the recorded video stream. Consequently, marker-based tracking is usually restricted to observing out-shot areas such as the ceiling or the floor which are normally covered by studio equipment, like light installations, cables, and mountings. Thus, occlusions and dynamic reconfigurations of the installations cause additional problems for marker-based tracking.

Another problem is the *acquisition of the scene depth*. This is necessary for integrating synthetic 3D objects into the video stream while producing consistent occlusion and illumination effects with the recorded real content. Some approaches reconstruct the scene geometry offline (during a special calibration step) using multiviewpoint stereo from un-calibrated video sequences. The quality of such techniques relies on the quality of feature matching in the stereo pairs. However, finding matchable features to support a high quality depth reconstruction might be difficult not only for real studio environments, but also for virtual studios or embedded blue screens that mainly apply uniformly colored matting surfaces. Besides offline reconstruction of the static studio setting, online depth estimations (e.g., of moving people in the scene) is even more problematic. In virtual studios the uniformly colored background that is required for chromakeying enables the application of fast depth-from-silhouette algorithms or similar techniques (e.g., [GPT04]). This, however, is not possible with a real studio setting.

Yet another challenge for virtual and augmented studios is the question of how to *display direction information* to moderators, actors or participants during a live broadcast or a recording. Teleprompters or fixed screens offer limited possibilities since they do not allow bringing the presented information into a spatial context. Step sequences, for instance, are usually marked statically on the floor ground.

In this article, we propose the application of digital light projection for studio illumination either exclusively or in combination with an existing analog lighting (cf. figure 1). This can solve several of the problems that are mentioned above, but also opens new possibilities for modern television productions.

The remainder of this article is organized as follows: Chapter 2 presents the technical key concept of digital studio illumination, while chapter 3 presents early laboratory examples. Chapter 4 finally outlines open problems and challenges that have to be addressed when transferring the concept and the presented technological approaches to real studio environments.

2 Digital Studio Illumination

Projectors allow a spatial and temporal modulation of light that can be computer controlled and synchronized with the recording process of studio cameras. Our technical key concept is visualized in figure 2: Multiple projectors are used exclusively, or in combination with analog light sources for illuminating the entire, or parts of the studio environment.

Physically, projectors represent point light sources. Their capability of spatially modulating luminances and chrominances on a per-pixel basis, however, allows for computing and creating almost arbitrary shading effects within the studio synthetically. This is called *projector-based illumination* and is described in more detail in section 3.4. Compared to a conventional analog illumination, a projector-based illumination allows re-illuminating the studio on the fly without changing the physical light sources. It can be combined with a similar technique that we refer to as *screen-based illumination*, which is also described in section 3.4.

Besides a spatial modulation, a temporal modulation of the projected light enables displaying different portions of the illumination time-sequentially. Figure 2 shows an example of two sequentially projected images l and r) that carry different parts of the illumination. Due to the lack of the human visual system, the average illumination of $l + r$ will be perceived when images are projected fast enough. The projection of more than two images is also possible, but requires fast display and capturing rates. This allows integrating coded patterns into one or several of these time slices in such a way that the sum of all images will result in the desired illumination or pictorial content. This principle has been described by Raskar et al. [RWC+98]. However, since video projectors were too slow and access to low-level image generation was not available at this time, a complete imperceptible code projection was not achieved and flickering was still well visible. Synchronizing the studio cameras to the projection and capturing all slices separately, however, will make the coded patterns visible to the camera but not inside the studio. Summing all slice images computationally after recording will lead to the same image that would be integrated optically by the camera sensor during the shutter time that equals the duration required for projecting all slices (i.e., the fully illuminated studio). This is referred to as *embedded imperceptible pattern projection* and is explained in section 3.1. The extracted coded patterns can be used for camera pose estimation (section 3.2) or for depth reconstruction (section 3.3).

Embedded imperceptible pattern projection has recently been realized for 3D video capturing and continuous projector calibration applications: Cotting et al. [CNGF04], for instance, synchronize a camera to a well-selected time-slot during the modulation sequence of a DLP projector. Within this time-slot the calibration pattern is displayed and detected by the camera. Since such an approach requires modifying the original colors of the projected image, a loss in contrast is an undesired side effect. Other techniques rely on a fast projection of images that cannot be perceived by the observer. This makes it possible to embed calibration patterns in one frame and compensate them with the following frame. Capturing altering projections of colored structured light patterns and their complements allows the simultaneous acquisition of the scenes depth and texture without loss of image quality [WWC+05, VVSC05]. Using structured light, depth reconstruction becomes easier since natural feature detection is not necessary because correspondences are provided by the projected codes. This also applies for camera tracking (section 3.2). Once the studio scene and the poses of the cameras are reconstructed, computer generated graphics can be augmented consistently into the recorded video stream.

Furthermore, a temporal modulation of light supports the visualization of information that is visible in the studio, but not in the camera image. Fukaya et al. [FFY+03], call this *invisible light projection* with respect to invisibility to a camera rather than to an observer. They project an image onto a blue screen located in a real TV studio. The projected image is alternately blocked and transmitted by an LCD shutter mounted in front of the projector lens. A separate shutter control unit synchronizes projection and exposure time of a studio camera in such a way that images are only captured when the projection is blocked. Chro-

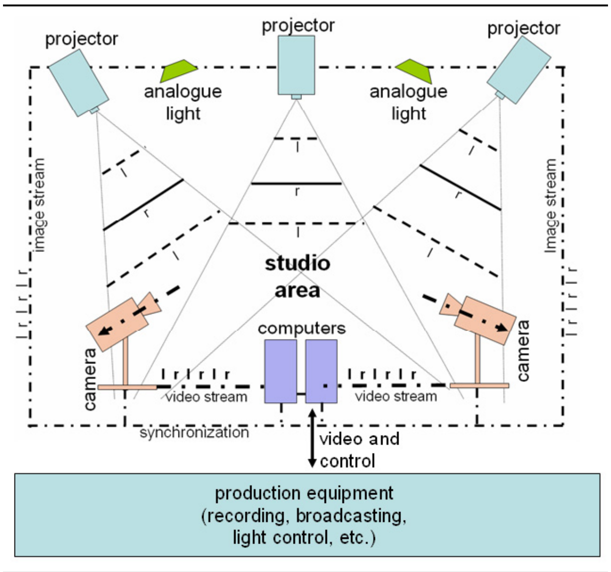

Figure 2: Modulating digital light in TV studios.

makeying can then be applied in a conventional way.

Shirai et al. [STK+05] apply chrominance and luminance keying instead of a shuttered projection for solving the same problem. Projecting an image onto a blue screen enables computing both a luminance key and a chrominance key that allow masking out the blue screen area for a final composition. This is possible only if the projected image is not brighter than the studio illumination.

If a virtual studio contains retro-reflective instead of blue or green diffuse screens, projectors can also display direction information that is invisible to the recording camera [GPT04]. This is achieved by mounting a blue or green light source near or around the camera lens. Most of the key-colored light is re-directed directly towards the camera lens by the retro-reflective coating, while the projected images are mainly reflected back towards the displaying projectors. The blue illumination and the projected images, however, are also partially diffused by the retro-reflective material. Consequently, moderators or actors inside the studio will perceive an equal mixture of both components. Since for the perspective of the camera the reflected key-color is much stronger than projected images, chromakeying becomes possible. Shape-from-silhouette techniques are used in combination with such a method by Grau et al. [GPT04] for fast scene acquisition and projection control.

Fukaya's simple concept can be extended to combine the presentation of arbitrary information that is visible only in the studio together with other data (e.g., the synthetic studio illumination) that is recorded simultaneously (section 3.1). This, for instance, makes it possible to display direction, moderation and other information dynamically and spatially anywhere within the studio not being limited to fixed locations, like teleprompters or static screens. Furthermore, limitations that are due to chromakeying in virtual studios, such as dim or color-distorted projections, or the application of uniformly colored or retro-reflective screens can be overcome.

Problematic for most of the techniques described above is the fact that the images are projected onto complex surfaces of a real studio rather than onto geometrically and radiometrically uniform surfaces, as it is the case for blue-screen studios. Projected images will be modulated (e.g., color blended or geometry distorted) by the underlying surfaces. Projected code patterns or direction information will not appear correct neither directly in the studio nor in the captured images. Appropriate correction techniques have to be applied to compensate for these effects before the images are projected. Some of these techniques will be outlined in section 3.3.

It should be noted at this point that the projected images do not necessarily have to contain pure illumination information. They can carry an arbitrary content, such as projected pictures or special effects. Thus, it is interesting to investigate the combination of *projector-based spatial augmentation* [BR05] and conventional *video augmentation* for future virtual and augmented studio productions.

3 Bits and Pieces in a Small Scale

This chapter presents several proof-ofconcept realizations for different technological components that have been mentioned in chapter 2. These components have not yet been transferred to a real studio environment but rather give an indication for the feasibility of our concept. Chapter 4 will discuss the remaining challenges when putting these pieces together in the large scale.

3.1 Embedding Imperceptible Patterns and Projecting Invisible Light

Digital Micro Mirror Devices (DMDs) that are used in DLP projectors control the pixel intensities by modulating the time in which the mirrors reflect light towards the projection surface. The micro mirrors can switch between their on/off states within approximately 10 μs. Colors are modulated in addition by synchronizing a rotating color wheel to the DMD. Thus, the light projected per pixel by a DLP projector can be seen as a time-multiplexed sequence of a discrete number of color bands (e.g., for red, green and blue). Each color slot itself is intensity modulated by a sequence of on/off states of the DMD that is encoded by the bit chain which represents the corresponding intensity value.

Imperceptible binary patterns can be embedded into a small time segment of this sequence. They are visible only to a camera which is synchronized to the same time segment. Cotting et al. [CNGF04] occupies specific time slots exclusively for displaying a binary pattern. However, pixels of the original image that have colors and intensities which are modulated within these time slots cannot be displayed if their corresponding code pattern is turned off. They must be modified in such a way that they do not fall into these code slots which results in the reduction of tonal resolution of the projected image. Furthermore, the mirror flip sequences have to be measured precisely for each projector. Thus, an individual calibration of each projector is required. The advantage of this approach, however, is that it can be used together with off-the-shelf projectors.

Instead of allocating time slots of a DLP projector's modulation sequence, code patterns can also be displayed within an entire projection frame. To ensure that they are not visible, the complementary pattern has to be projected immediately after, as suggested in [RWC$^+$98]. If this is being done fast enough, a white image will be perceived. A synchronized camera can capture both images the code image and its complement. Adding both images computationally results in the image of the scene illuminated with projected white light. This technique is used for capturing scene colors and geometry in the context of 3D video applications [VVSC05, WWC$^+$05]. One main drawback of this approach is that the code patterns are projected with the limited frame rate of the projector (e.g. 60Hz-70Hz for conventional projectors). This results in a well perceivable flickering and is consequently not appropriate for a real studio illumination.

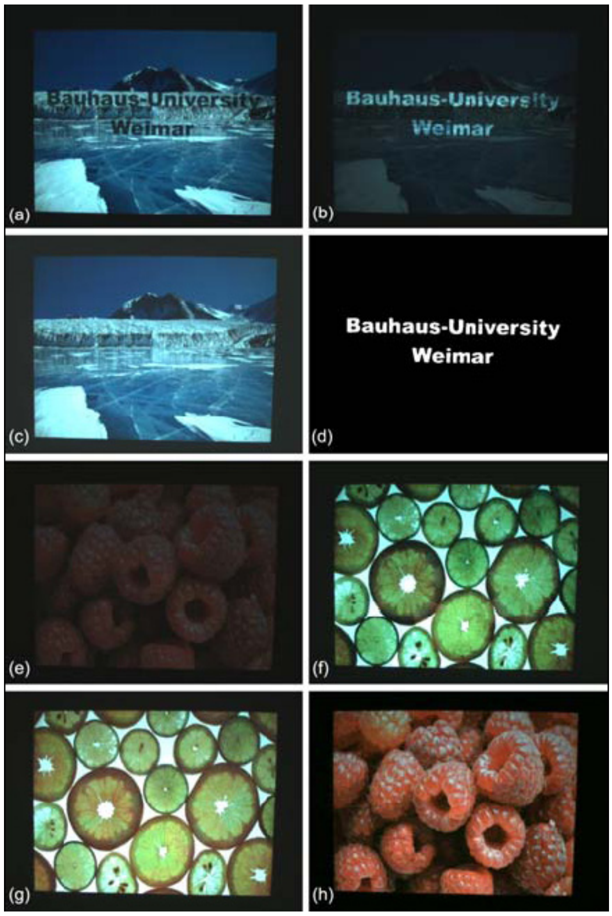

Figure 3: Embedded imperceptible patterns (a-d) and projected invisible light (e-h).

In our approach, we combine the advantages of both techniques discussed above.With modern projectors, this leads to a method that embeds imperceptible patterns into projected images as in [RWC$^+$98] - but without perceived flickering, reductions in contrast, or the need for a projector individual mirror flip calibration. Furthermore, it is neither limited to DLP technology, nor does it require extremely short exposure times of cameras. However, we do apply special stereoscopic projectors that are capable of a high native frame rate (120Hz in our case).

As discussed earlier, projecting two complementary images (l and r) with a high frequency makes them appear as the average sum $(l + r)/2$ of the two images (cf. figure 3c). A synchronized camera, however, can capture both images individually (cf. figure 3a,b). Subtracting or dividing both images, normalizing and binarizing the result allows extracting the embedded code (cf. figure 3d). Both images can also be added computationally to determine the image that is actually visible. This image can then be recorded (it equals figure 3c).

One challenge is to avoid visible flickering even if the code pattern is exchanged during the projecting sequence. In our solution, we smoothly fade new code patterns in and out within a short sequence of consecutive projection frames. Dynamic content, such as videos or interactive applications have to be synchronized to this process to ensure that corresponding frames contain the same visible content. Currently, we achieve a frame rate of 20Hz for displaying dynamic content and simultaneous code extraction (while projecting with 120Hz). This speed is mainly limited by the capturing rate of our camera[1]. By applying a faster camera[2] we estimate to double this frame rate approximately.

As mentioned in chapter 2, it is also possible to display information that is visible in the studio but is not recorded by the camera. An example is illustrated in figure 3. Assuming a desired visible studio image v, we know that it will be composed from two or more sequentially projected slice images (e.g., $v = (l+r)/2$ for two slices). Knowing the image that should be captured by the camera in addition (e.g., l), we have to compute a compensation image r in such a way that that $r = 2v - l, 2v >= l$. Figure 3e illustrates the image l that is only visible to the camera when capturing it at the corresponding time slot. Figure 3f presents the computed compensation image r for the desired image v that is visible as projection in the studio (cf. figure 3g). Figure 3h illustrates a digitally contrast enhanced version of l. This image is actually being recorded simultaneously (instead of the low contrast version shown in figure 3e).

Note that the methods described in this section are widely independent of the image content. Thus, the projected images do not necessarily have to contain pictorial data as in the example shown with figure 3. They can also carry synthetic illumination information instead. This will be described in section 3.4.

3.2 Camera Pose Estimation with Embedded Markers

Embedding imperceptible patterns into projected images that carry pictorial or illumination information allows supporting a marker-based in-shot camera tracking. This means that markers can be displayed directly within the studio environment. However, they are not directly visible within the studio itself and will neither be recorded. In contrast to out-shot approaches (e.g., when tracking markers or studio lights that are attached to the ceiling), this does not cause conflicts with other studio equipment and does not require an additional camera or other devices.

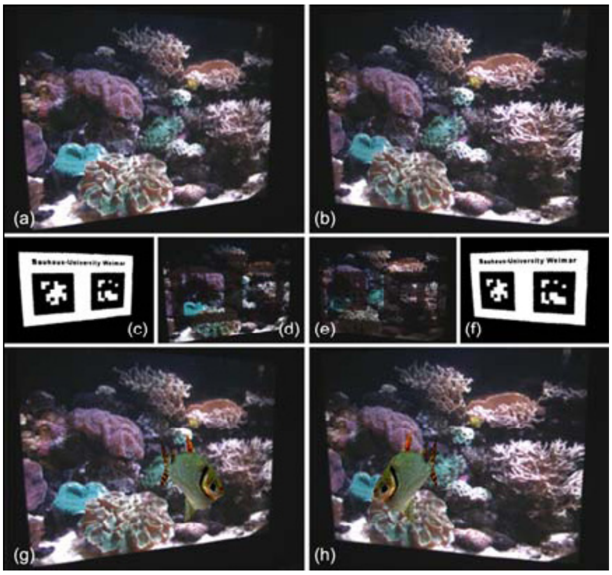

Figure 4: Embedded imperceptible markers for inshot camera tracking and video augmentation.

Figures 4a and 4b show a projected image as it can be perceived when observing it from two different perspectives. The projected image contains an embedded imperceptible marker that is extracted with the technique explained in section 3.2. This is illustrated in figures 4c-4f for both perspectives. After binarizing the marker the position and orientation of the camera relative to the markers origin can be determined[3]. The two captured images l and r that are required for separating the embedded code are computationally added. A virtual object (the fish) is then rendered on top of the combined image from the estimated camera perspective (cf. figures 4g, 4h). This leads to a live video augmentation of the captured environment that contains arbitrary projections. The augmentation, however, will not appear in the real environment.

Once again it has to be noted that the projected images do not necessarily have to contain pictorial information, as shown in figure 4. They can carry the

[1]A Point Grey Dragonfly 2.

[2]Such as a Dragonfly Express.

[3]ARTag [Fia05] was used for marker tracking.

studio illumination instead (see section 3.4) making it possible to augment the illuminated studio environment with exactly the same method.

3.3 Projecting onto Everyday Surfaces and Acquisition of Scene Depth

The application of projection technology in virtual or real studios is usually restricted to the projection of images onto special screen surfaces, such as white diffuse screens or blue screens. The concept that is proposed in this article, however, requires a projection onto the real surfaces of the existing studio setting and possibly onto dynamic content, such as moderators and actors. As discussed above, these images can carry pictorial information making it possible to display visible direction information dynamically onto arbitrary studio surfaces without being limited to static screens. The image can display imperceptible embedded patterns that are used for camera tracking, depth acquisition, or online calibration processes. Finally, they can also contain the synthetic studio illumination (see section 3.4). In any case, the projected images are distorted when being reflected from non-optimized surfaces. This does not only lead to visible errors in projected pictorial content which are well perceivable in the studio or in the recorded video stream, but also to problems when extracting embedded code patterns. Thus, real-time image correction techniques are required that are capable of compensating for any image distortion that is caused by a projection onto arbitrary surfaces. Furthermore, multiple projectors have to be calibrated in such a way that a single consistent image can be presented from multiple individual projector contributions.

Geometric calibration techniques for multi-projector systems, such as tiled screens, widely use camera feedback to support automatic registration. Geometric image registration approaches for simple planar surfaces determine homography matrices when warping images from a reference perspective to the perspectives of the projectors [YGH$^+$01]. For projection surfaces with a non-trivial but known geometry, intrinsic and extrinsic projector parameters have to be estimated to enable image warping via projective texture mapping operations [RBY$^+$99].

These conventional techniques alone will fail in the case of displaying images with projectors in a complex studio environment, because the surfaces available in a real studio are usually not optimized for projections. They can be geometrically complex, colored, textured, nondiffuse, and can cover a large depth range. This results in geometric distortions, color blending, intensity variations, and regional defocus effects in the projected images.

With coded structured light techniques, the surface geometry, reflectance as well as the global and local illumination behavior can be automatically determined by evaluating the corresponding camera feedback. As explained above, the structured code can be embedded seamlessly into the projected content and remains imperceptible in the studio or in the recorded content.

This makes it possible to determine the relation between all pixels of each projector with respect to the parameters of the scene points onto which they project (i.e., their geometric position, reflectance, local and global illumination parameters).

Knowing these parameters, many distortion effects can be compensated on a per-pixel basis and in real-time when being carried out on modern GPUs. Besides pixel-precise geometric warping [BEK05], photometric [BMY05] and radiometric calibration [NPGB03, GPNB04, BEK05, FGN05, AOSS06, GB06, WB06] techniques ensure chrominance and luminance consistency, as well as the compensation of color and intensity artifacts when projecting onto colored and textured surfaces. Multi-focal projection [BE06] and defocus compensation [ZN06] methods can be used for increasing the focal depth of single or multiprojector systems. Reflection highlights on specular surfaces can also be eliminated with appropriate multi-projector techniques [PLKP05]. Even global illumination effects, such as surface-to-surface scattering [BGZ$^+$06] or complex physical light modulations like inter-reflections and refractions [WB06] can be neutralized.

An example for a projection onto a nonoptimized surface is illustrated in figures 5a-e. The pixels of an uncompensated image projected onto a colored and textured surface will be color and intensity blended with the underlying surface pigments (cf. figures 5a,d). After measuring the surface parameters via structured light projection (cf. figures 5b,c), a real-time and pixel-precise radiometric compensation can be applied. This minimizes these artifacts directly on the surfaces and consequently in the captured images (cf. figure 5e). Note that in both cases the image is warped (also in real-time and on a per-pixel basis) to compensate for geometric distortions caused by the

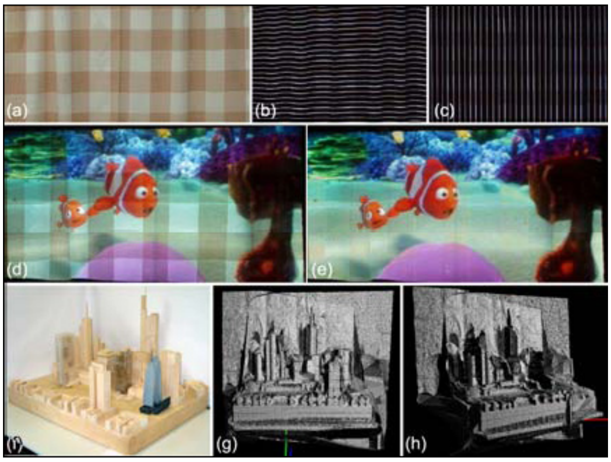

Figure 5: Projection onto a colored window curtain with and without radiometric compensation (a-e), ©2005 IEEE [BEK05]. Acquired scene depth via structured light projection (f-h).

non-planar surface. Thus, it appears as being projected onto a planar surface from the perspective of the camera.

All of these image correction techniques ensure that the desired images can be displayed and captured when projecting synthetic illumination, pictorial content or code patterns directly onto real studio surfaces. The parameters that are required for carrying out the compensation computations can be determined continuously and unnoticeably with the techniques described in section 3.2. This, for example, implies that the coded patterns shown in figures 5b and 5c are not visible in the studio or in the recorded video. Besides radiometric and geometric distortions, other effects, such as global inter-reflections, specular highlights or defocus effects can be compensated as described earlier. To explain theses techniques is out of the scope of this article.

Furthermore, being able to project corrected code patterns onto arbitrary surfaces allows applying structured light projection for a fast depth acquisition more efficiently. This is shown in figures 5f-h. The depth map has been computed for a number of unstructured positions of a marker-tracked camera. Line-strip patterns (such as the ones shown in figures 5b,c) have been projected onto the scene for providing sufficient artificial features.

3.4 Projector-Based and Screen-Based Illumination

Synthetic re-illumination has been an active topic in computer graphics and computer vision for many years. We can differentiate between methods that re-illuminate recorded image or video content, or approaches that physically re-illuminate a real scene or object with controlled lighting. Only the latter category is of interest for our concept.

A technique called *virtual re-illumination* has been introduced by Paul Debevec [DWT$^+$02]. It was used to create special effects in recent Hollywood movies, such as Spiderman or King Kong. A special recording environment, called *LightStage*, allows producing a variety of different basic lighting situations in a high speed with analog light bulbs surrounding an actor. They are captured with a synchronized fixed camera. Having recorded the discrete basis functions of a 5D slice of a 12D reflectance field, the illumination in the video content can be altered after recording.

Other recording environments surround actors with diffuse rear-projection screens instead of light bulbs. This makes a direct re-illumination before or during recording possible by displaying appropriate environment maps on the screens [MFKY05]. We want to refer to this as *screen-based illumination*.

Both approaches, virtual and screen-based re-illumination, require specialized recording environments and are so far not suitable for real studios. They also focus on the re-illumination of actors or moderators, rather than on the re-illumination of an entire studio environment.

Projector-based illumination techniques re-illuminate a physical environment synthetically (before or during recording) using a discrete number of unstructured aligned projectors. Thereby, the projectors illuminate the real environment directly and on a per-pixel basis, rather than indirectly by projecting onto diffuse screens that scatter light into the environment.

The application of projectors for simple shadow casting in a virtual studio context has been demonstrated by Wojdala et al. [WGO00]. Thereby, projectors have been located at the position of the studio lights for creating synthetic shadows on real objects cast by virtual objects. Each projector displays an individual shadow mask based on the virtual scenery that is dynamically computed for its particular position. The shadow casting, however, is constrained to the static configuration of the projectors. This means,

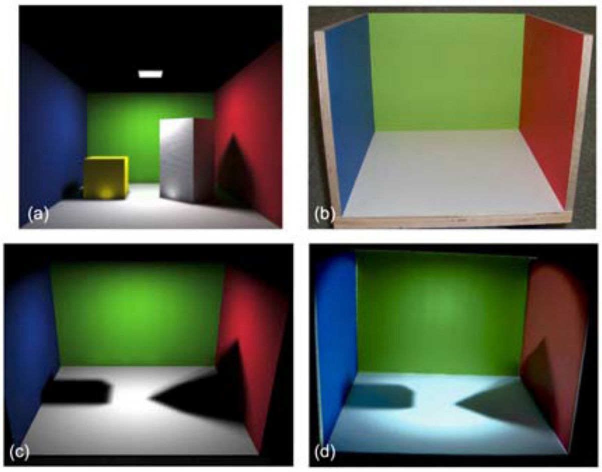

Figure 6: Projector-based illumination for creating global illumination effects synthetically, ©2003 IEEE.

for instance, that shadows from virtual objects on real objects cannot be cast from other directions without physically changing the position of the projectors. In addition, shadows on virtual objects (either from other virtual objects or from real ones) must be cast from the same positions to ensure consistency. Furthermore, shadows on real objects cast from other real objects must also be created from the same positions for the same reason. A true projector-based illumination holds the potential of solving this problem and consequently offers more flexibility. It allows computing and creating synthetic shading and shadowing for real components that matches an arbitrary illumination of the virtual scenery regardless of the physical position of the projectors.

The challenge of projector-based illumination is to produce a defined lighting situation within a real environment without the necessary light sources. The only available light sources are the projectors themselves that represent static point lights. Consequently, images have to be computed for each projector with the following objectives: First, they must neutralize the physical illumination effects that are caused by each projector as a real point-light source. Second, they have to produce the defined virtual lighting situation synthetically.

Projector-based illumination has been demonstrated on the small scale: for uniform white surfaces simulating only local lighting effects [RWLB01], for colored and textured surfaces simulating local and global lighting effects [BGWK03], and even for optical holograms [Bim04]. An image-based technique to re-light real objects illuminated by a 4D incident light field has also been described by Masselus et al. [MPDW03].

Figure 6 demonstrates an example for projector-based illumination. The global illumination for a model of a simple room environment, consisting of walls, boxes and an area light source is computed. Figure 6a shows a screenshot of the results after several radiosity iterations. Shading, soft-shadows and color-bleeding caused by the area light source and the models geometry are clearly visible. A real miniature mock-up of the room without the boxes is illustrated in figure 6b. Using projector-based illumination, the computed lighting on the wall surfaces can be created syntactically. While figure 6c shows another screenshot of the computationally illuminated scene without boxes, figure 6d presents a photograph of the real mock-up from the same perspective, but illuminated synthetically with a single projector. Similar shading effects that are computed for the virtual area light can be produced by the projector (i.e., a physical point light being located at a completely different position). Minor differences between the computed illumination (figure 6c) and the synthetically created illumination in the real mock-up (figure 6d) are due to technical limitations of the projector, such as its high black-level and its low contrast.

The rendering of the intensity images that create the synthetic illumination is fully compatible with the techniques presented in section 3.1 and 3.2. The only difference is that projected images do not carry pictorial content only, but also the synthetic studio illumination.

Note that the defined virtual illumination can be arbitrary. Besides global and view-independent effects, such as the ones shown in figure 6, it can also contain view-dependent effects, such as synthetic specular highlights or refractions. The latter case requires the knowledge of the cameras poses. This can be determined as described in section 3.2.

However, for a projector-based illumination it is essential that physical light modulations within the real environment are canceled out. This was discussed in section 3.3. Projector-based and screen-based illumination can also be combined by projecting through a shuttered diffuse screen (cf. figure 7). Such screens contain a phase dispersed liquid crystal layer that can be electronically switched to a diffuse or a transparent state. In the transparent state images are projected straight through the screen (cf. figures 7a,c),

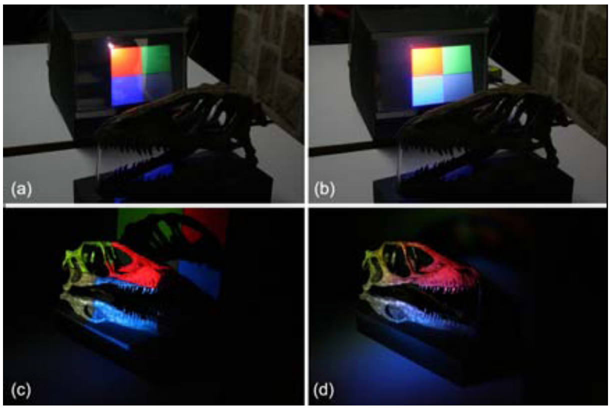

Figure 7: Combination of projector-based illumination (a,c) and screen-based illumination (b,d).

and projector-based illumination or displaying pictorial content is supported. In the diffuse mode, the projected image is scattered at the screen (cf. figures 7b,d), and a screen-based illumination in a real recording environment becomes possible. The sizes of the diffusers affect the possible illuminating effects. They can reach from small screens that are rigidly mounted in front of each projector lens (as shown in figure 7) to large screens, as used for existing recording environments that currently apply passive projection screens [MFKY05].

Since the screens can be switched between transparent and diffuse modes with a high speed (e.g., 50Hz-100Hz), a simultaneous activation of both illumination schemes can be perceived and recorded. Synchronizing the camera to the shutter signal of the screen allows separating both illumination types and combining them computationally, if necessary. If analog light sources are used in addition, they have to be shuttered and synchronized with projectors, screens and cameras.

For projector-based and screen-based illumination, an adequate number of projectors is required to lighten the entire studio environment from multiple directions. Multi-projector techniques, such as the ones outlined in section 3.3 can be used for calibration and for displaying consistent and undistorted images. Multi-projector techniques can also be used for the removal of shadows that are cast by static or dynamic content [SCS01, JWS$^+$01].

4 The Big Picture

With the concept proposed in this article, we envision a technical extension to existing virtual and augmented studio technology that enables new effects and control possibilities.

Modulated digital light projection opens new possibilities for modern television studios:

- dynamic re-illumination of studio settings, moderators and actors without physical modification of the lighting equipment;
- marker-based in-shot tracking of studio cameras without visible markers;
- dynamic presentation of un-recorded direction, moderation and other information spatially anywhere within the studio;
- integration of imperceptible coded patterns that support continuous online-calibration, camera tracking, and acquisition of scene depth.

Most of these points can be addressed individually with different technological approaches. Digital light projection, however, holds the potential of offering a unified solution.

Although the examples presented in chapter 3 proof the feasibility of individual techniques in the small scale and under laboratory conditions, several problems have to be addressed before they can be applied in real studio environments. Up-scaling these techniques clearly represents the main challenge.

Installing a large number of projectors in a studio makes the investigation of a robust hardware infrastructure necessary. This includes adequate cooling for ceiling mounted projectors and analog lighting equipment, networked PCs for distributed rendering, synchronization electronics for projectors, cameras and shutter screens, and interfaces to existing production equipment (cf. figure 2).

Besides the hardware infrastructure, a scalable software framework has to be implemented that supports the following points:

- driving multi-projector cluster over a grid of networked PCs (including calibration);
- supporting fast frame grabbing and video processing;

- realizing real-time rendering techniques for multi-projector systems, such as synthetic relighting and image compensation;
- offering computer vision techniques for projector-camera systems, such as code extraction, pose-estimation, depth acquisition, and scene analysis;
- implementing consistent video and projector-based spatial augmentation techniques;
- interfacing to existing production software and equipment.

Automatic calibration methods will register the projectors to the studio environment continuously if embedded into online presentations. Once the projectors are registered, imperceptible structured light projection can be used together with the tracked studio cameras to acquire scene depth of static (e.g., furniture) as well as dynamic (e.g., moving people) content. Techniques have to be developed that automatically extract the dynamic from the static content. While the static content can be scanned during an offline stage, dynamic components have to be scanned online. The reconstructed depth information is required for creating consistent occlusion and illumination effects for graphical augmentations that are integrated into the recorded video stream. They are also essential for the synthetic re-illumination process.

Having geometric information of the static studio content allows analyzing its topology. This, for instance, enables the search for planar subsurfaces that are suitable for displaying projected tracking markers. Doing this continuously makes it possible to dynamically reconfigure marker positions on the fly to ensure an optimal visibility and to avoid occlusions. This is clearly not possible with static markers [TJU97].

The development of scalable re-illumination techniques for creating synthetic illumination effects is as challenging and exciting as realizing scalable image compensation methods for neutralizing the physical modulation of light in real environments, such as TV studios. Furthermore, the investigation of efficient mixtures between video augmentations and projector-based spatial augmentations in a studio production context is yet another interesting topic of research.

Although this article focuses on spatial and temporal light modulation in TV studios, alternative modulated methods, such as wavelength modulation (e.g., on an invisible infrared basis), will also have to be investigated. Finally, we believe that these techniques will not only find their applications in television studios only. They will offer new possibilities to similar domains, such as photo studios (in the context of computational photography [LND+06]) or live stage performances.

Our group at the Bauhaus-University Weimar is actively working on realizing this concept and transferring our current small-scale implementations to the large scale.

5 Remark

The investigation and realization of concepts and techniques outlined in this article have been proposed to the Deutsche Forschungsgemeinschaft (DFG) under reference number BI 835/2-1.

References

[AOSS06] Mark Ashdown, Takahiro Okabe, Imari Sato, and Yoichi Sato, *Robust Content-Dependent Photometric Projector Compensation*, Proceedings of the Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06), Projector-Camera Systems (PROCAMS), 2006, p. 6.

[BE06] Oliver Bimber and Andreas Emmerling, *Multifocal Projection: A Multiprojector Technique for Increasing Focal Depth*, IEEE Transactions on Visualization and Computer Graphics **12** (2006), no. 4, 658–667, ISSN 1077-2626.

[BEK05] Oliver Bimber, Andreas Emmerling, and Thomas Klemmer, *Embedded entertainment with smart projectors*, Computer **38** (2005), no. 1, 48–55, ISSN 0018-9162.

[BGWK03] Oliver Bimber, Anselm Grundhöfer, Gordon Wetzstein, and Sebastian Knödel, *Consistent illumination within optical see-through augmented environments*, Proceedings of the Second IEEE and ACM International Symposium on

Mixed and Augmented Reality, 2003, ISBN 0-7695-2006-5, pp. 198–207.

[BGZ+06] Oliver Bimber, Anselm Grundhöfer, Thomas Zeidler, Daniel Danch, and Pedro Kapakos, *Compensating Indirect Scattering for Immersive and Semi-Immersive Projection Displays*, Proceedings of the IEEE Virtual Reality Conference (VR 2006), 2006, pp. 151–158.

[Bim04] Oliver Bimber, *Combining Optical Holograms with Interactive Computer Graphics*, Computer **37** (2004), no. 1, 85–91, ISSN 0018-9162.

[BMY05] Michael Brown, Aditi Majumder, and Ruigang Yang, *Camera-Based Calibration Techniques for Seamless Multiprojector Displays*, IEEE Transactions on Visualization and Computer Graphics **11** (2005), no. 2, 193–206, ISSN 1077-2626.

[BR05] Oliver Bimber and Ramesh Raskar, *Spatial Augmented Reality: Merging Real and Virtual Worlds*, A. K. Peters, Ltd., Natick, MA, USA, 2005, ISBN 1-568-81230-2.

[BT03] M. Bernas and M. Teisler, *Determination of the camera position in virtual studio*, Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology, 2003. ISSPIT 2003, 2003, ISBN 0-7803-8292-7, pp. 535–538.

[CNGF04] Daniel Cotting, Martin Näf, Markus H. Gross, and Henry Fuchs, *Embedding Imperceptible Patterns into Projected Images for Simultaneous Acquisition and Display*, Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'04), 2004, ISBN 0-7695-2191-6, pp. 100–109.

[DWT+02] Paul Debevec, Andreas Wenger, Chris Tchou, Andrew Gardner, Jamie Waese, and Tim Hawkins, *A Lighting Reproduction Approach to Live-Action Compositing*, ACM Transactions on Graphics **21** (2002), no. 3, 547–556, Special Issue: Proceedings of ACM SIGGRAPH 2002, ISSN 0730-0301.

[FFY+03] T. Fukaya, H. Fujikake, Y. Yamanouchi, H. Mitsumine, N. Yagi, S. Inoue, and H. Kikuchi, *An effective Interaction Tool for Performance in the virtual Studio - Invisible Light Projection System*, NHK Science & Technical Research Laboratories, Japan, 2003.

[FGN05] Kensaku Fujii, Michael D. Grossberg, and Shree K. Nayar, *A projector-camera system with real-time photometric adaptation for dynamic environments*, IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2005, Volume 1, 2005, ISBN 0-7695-2372-2, pp. 814–821.

[Fia05] Mark Fiala, *ARTag, a Fiducial Marker System Using Digital Techniques*, IEEE Computer Vision and Pattern Recognition or CVPR'05, 2005, Volume 2, ISBN 0-7695-2372-2, pp. 590–596.

[FKGK05] Jan-Michael Frahm, Kevin Koeser, Daniel Grest, and Reinhard Koch, *Markerless Augmented Reality with Light Source Estimation for Direct Illumination*, Conference on Visual Media Production CVMP, London, December 2005, 2005.

[GAB+98] Simon Gibbs, Costas Arapis, Christian Breiteneder, Vali Lalioti, Sina Mostafawy, and Josef Speier, *Virtual studios: an overview*, IEEE Multimedia **5** (1998), no. 1, 18–35, ISSN 1070-986X.

[GB06] Anselm Grundhöfer and Oliver Bimber, *Real-Time Adaptive Radiometric Compensation*, ACM SIGGRAPH 2006 (Poster) Submitted to IEEE Transactions on Visualization and Computer Graphics (TVCG), 2006.

[GPNB04] Michael D. Grossberg, Harish Peri, Shree K. Nayar, and Peter N. Belhumeur, *Making One Object Look Like Another: Controlling Appearance Using a Projector-Camera System*, 2004 IEEE Computer Society Conference on

Computer Vision and Pattern Recognition (CVPR'04) - Volume 1, 2004, ISBN 0-7695-2158-4, pp. 452–459.

[GPT04] Oliver Grau, Tim Pullen, and Graham A. Thomas, *A combined studio production system for 3-D capturing of live action and immersive actor feedback*, IEEE Transactions on Circuits and Systems for Video Technology **14** (2004), no. 3, 370–380, ISSN 1051-8215.

[IST04] IST, *MATRIS - Markerless real-time Tracking for Augmented Reality Image Synthesis*, http://www.ist-matris.org/, 2004, Last visited August 31st, 2006.

[JWS+01] Christopher Jaynes, Stephen Webb, R. Matt Steele, Michael Brown, and W. Brent Seales, *Dynamic Shadow Removal from Front Projection Displays*, Proceedings of the 12th IEEE Visualization 2001 (VIS'01), 2001, ISBN 0-7803-7200-X, pp. 175–182.

[KIA+00] Masahiro Kawakita, Keigo Iizuka, Tahito Aida, Hiroshi Kikuchi, Hideo Fujikake, Jun Yonai, and Kuniharo Takizawa, *Axi-vision Camera: A Three-dimensional Camera*, Proceedings of the Conference on Three-Dimensional Image Capture and Applications III, Volume 3958, 2000, ISBN 0-8194-3576-7, pp. 61–70.

[KKSES05] R. Koch, K. Koeser, B. Streckel, and J.-F. Evers-Senne, *Markerless Image-based 3D Tracking for Real-time Augmented Reality Applications*, WIAMIS, April 13-15, Montreaux, Switzerland, 2005.

[KPG98] Reinhard Koch, Marc Pollefeys, and Luc J. Van Gool, *Multi Viewpoint Stereo from Uncalibrated Video Sequences*, ECCV'98 5th European Conference on Computer Vision, Lecture Notes in Computer Science, 1998, ISBN 3-540-64569-1, pp. 55–71.

[LND+06] M. Levoy, S. Nayar, P. Debevec, M. Cohen, R. Szeliski, and O. Bimber, *IEEE Computer, Special Issue on Computational Photography*, August Issue, 2006.

[MFKY05] H. Mitsumine, T. Fukaya, S. Komiyama, and Y. Yamanouchi, *Immersive Virtual Studio*, ACM SIGGRAPH'05 Sketch, 2005.

[MPDW03] Vincent Masselus, Pieter Peers, Philip Dutre, and Yves D. Willems, *Relighting with 4D incident light fields*, ACM Transactions on Graphics **22** (2003), no. 3, 613–620, ISSN 0730-0301.

[NPGB03] Shree K. Nayar, Harish Peri, Michael D. Grossberg, and Peter N. Belhumeur, *A Projection System with Radiometric Compensation for Screen Imperfections*, Proceedings of the IEEE International Workshop on Projector-Camera Systems (PROCAMS), 2003.

[PLKP05] Hanhoon Park, Moon-Hyun Lee, Sang-Jun Kim, and Jong-Il Park, *Specularity-Free Projection on Nonplanar Surface*, Advances in Mulitmedia Information Processing - PCM 2005, Lecture Notes in Computer Science, Volume 3767, 2005, ISBN 978-3-540-30027-4, pp. 606–616.

[RBY+99] Ramesh Raskar, Michael S. Brown, Ruigang Yang, Wei-Chao Chen, Greg Welch, Herman Towles, Brent Seales, and Henry Fuchs, *Multi-Projector Displays Using Camera-Based Registration*, Proceedings of the conference on Visualization '99, 1999, ISBN 0-7803-5897-X, pp. 161–168.

[RWC+98] Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs, *The office of the future: a unified approach to image-based modeling and spatially immersive displays*, Proceedings of the 25th annual conference on Computer graphics and interactive techniques, 1998, ISBN 0-89791-999-8, pp. 179–188.

[RWLB01] Ramesh Raskar, Greg Welch, Kok-Lim Low, and Deepak Bandyopadhyay, *Shader Lamps: Animating Real Objects With Image-Based Illumination*, Proceedings of the 12th Eurographics Workshop on Rendering Techniques, London, UK, June 25-27, 2001, 2001, ISBN 3-211-83709-4, pp. 89–102.

[SCS01] Rahul Sukthankar, Tat-Jen Cham, and Gita Sukthankar, *Dynamic Shadow Elimination for Multi-Projector Displays*, 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'01) - Volume 2, 2001, ISBN 0-7695-1272-0, p. 151.

[STK+05] Akihiko Shirai, Masafumi Takahashi, Kiichi Kobayashi, Hideki Mitsumine, and Simon Richir, *Lumina Studio: Supportive Information Display for Virtual Studio Environments*, IEEE VR 2005 Workshop on Emerging Display Technologies, 2005, pp. 17–20.

[TJU97] G.A. Thomas, J. Jin, and C. A. Urquhart, *A versatile camera position measurement system for virtual realityTV production*, International Broadcasting Convention, 1997, ISBN 0-85296-694-6, pp. 284–289.

[VVSC05] M. B. Vieira, L. Velho, A. Sa, and P. C. Carvalho, *A Camera-Projector System for Real-Time 3D Video*, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Volume 3, 2005, ISBN 0-7695-2372-2-3, p. 96.

[WB06] G. Wetzstein and Oliver Bimber, *Radiometric Compensation of Global Illumination Effects with Projector-Camera Systems*, ACM SIGGRAPH 2006 (Poster), submitted to IEEE Transactions on Graphics (TOG), 2006.

[WGO00] Andrzej Wojdala, Marek Gruszsewski, and Ryszard Olech, *Real-time shadow casting in virtual studio*, MGV (Machine Graphics and Vision) **9** (2000), no. 1/2, 315–329, ISSN 1230-0535.

[WWC+05] Michael Waschbüsch, Stephan Würmlin, Daniel Cotting, Filip Sadlo, and Markus H. Gross, *Scalable 3D video of dynamic scenes*, The Visual Computer **21** (2005), no. 8-10, 629–638, ISSN 1432-2315.

[YGH+01] Ruigang Yang, David Gotz, Justin Hensley, Herman Towles, and Michael S. Brown, *PixelFlex: A Reconfigurable Multi-Projector Display System*, Proceedings of the 12th IEEE Visualization 2001 (VIS'01), 2001, ISBN 0-7803-7200-X, pp. 167–174.

[YMF+02] Yuko Yamanouchi, Hideki Mitsumine, Takashi Fukaya, Masahiro Kawakita, Nobuyuki Yagi, and Seiki Inoue, *Real space-based virtual studio seamless synthesis of a real set image with a virtual set image*, Proceedings of the ACM symposium on Virtual reality software and technology, Hybrid VR, 2002, ISBN 1-58113-530-0, pp. 194–200.

[ZN06] Li Zhang and Shree Nayar, *Projection defocus analysis for scene capture and image display*, ACM Transactions on Graphics **25** (2006), no. 3, 907–915, ISSN 0730-0301.

Citation

Oliver Bimber, Anselm Grundhöfer, Stefanie Zollmann and Daniel Kolster, *Digital Illumination for Augmented Studios*, Journal of Virtual Reality and Broadcasting, 3(2006), no. 8, November 2006, urn:nbn:de:0009-6-6302, ISSN 1860-2037.