Perceptually-Based Compensation of Light Pollution in Display Systems
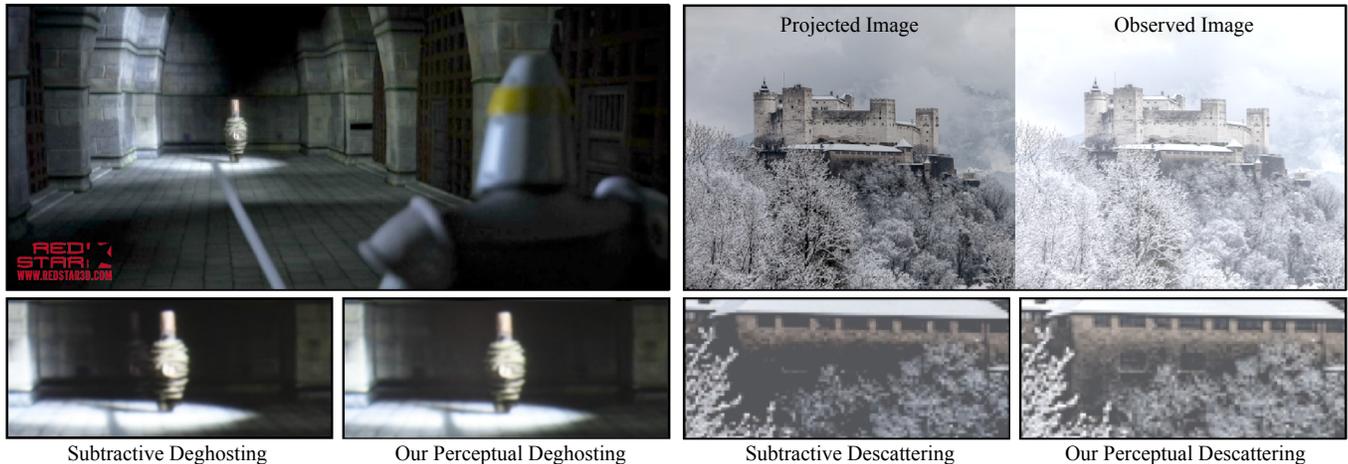
Jeroen van Baar
Disney Research Zürich

Steven Poulakos
ETH Zürich

Wojciech Jarosz
Disney Research Zürich

Derek Nowrouzezahrai
Disney Research Zürich

Rasmus Tamstorf
Walt Disney Animation Studios

Markus Gross
Disney Research Zürich, ETH Zürich

Figure 1: *Using our perceptual framework we can compensate for light pollution due to ghosting in 3D stereo (left) and indirect scattering when projecting onto concave screens (right). Current subtractive solutions fail in areas where the input image is dark (by either retaining some residual ghost or eliminating important surface detail). Our solution automatically solves for a perceptually optimal compensation that diminishes the appearance of ghosts and reveals more detail in these problem regions. Note: All deghosting results in this paper have been captured with a camera through a polarizing filter; All descattering results are simulations.*

Abstract

This paper addresses the problem of unintended light contributions due to physical properties of display systems. An example of such unintended contribution is crosstalk in stereoscopic 3D display systems, often referred to as ghosting. Ghosting results in a reduction of visual quality, and may lead to an uncomfortable viewing experience. The latter is due to conflicting (depth) edge cues, which can hinder the human visual system (HVS) proper fusion of stereo images (stereopsis). We propose an *automatic*, *perceptually-based* computational compensation framework, which formulates pollution elimination as a minimization problem. Our method aims to distribute the error introduced by the pollution in a perceptually optimal manner. As a consequence ghost edges are smoothed locally, resulting in a more comfortable stereo viewing experience. We show how to make the computation tractable by exploiting the structure of the resulting problem, and also propose a perceptually-based pollution prediction. We show that our general framework is applicable to other light pollution problems, such as descattering.

CR Categories: I.4.3 [Image Processing and Computer Vision]: Enhancement—Filtering;

Keywords: Perceptual Framework, Ghosting, Scattering, Optimization, Compensation

1 Introduction

Stereoscopic 3D displays, whereby a different image is presented to the left and right eye, have been around for some time. The recent popularity of 3D movies, TV and games has once again given attention to the problem of crosstalk. Dim copies of the images intended for one eye *leak* to the other eye and are perceived as "ghost" images. This ghosting is a form of unintended light contribution: light that is originally injected as intended light, but by some physical property of the display system results in unintended light contribution to (portions of) the intended image. Ghosting results in a reduction of visual quality due to contrast reduction. More important however, it also interferes with the HVS stereopsis. Unintended (depth) edges conflict with intended (depth) edges, thereby hindering the proper fusion of stereo images [Tsirlin et al. 2011]. These conflicts can result in an uncomfortable viewing experience. It is therefore important to compensate for ghosting.

Current methods for ghosting compensation, or "deghosting", typically rely on some form of subtractive compensation: the anticipated unintended light is subtracted from the images prior to display, e.g. [Konrad et al. 2000]. In areas where the input images have low luminance (Figure 1, left), subtraction would result in negative light values, which are subsequently clamped. Clamping can be avoided by raising the black level. However, this leads to a reduction in contrast and loss of details. We propose an automatic, *perceptually-based* computational framework to compensate for the light pollution. The *error* due to light pollution is distributed to areas of reduced sensitivity for the HVS, where details are better preserved and contrast is reduced only locally. This diminishes the appearance of ghosting and raises the apparent quality of the displayed images.

In this paper we specifically focus on controlled illumination environments, and 3D cinema in particular. Although the focus in this paper
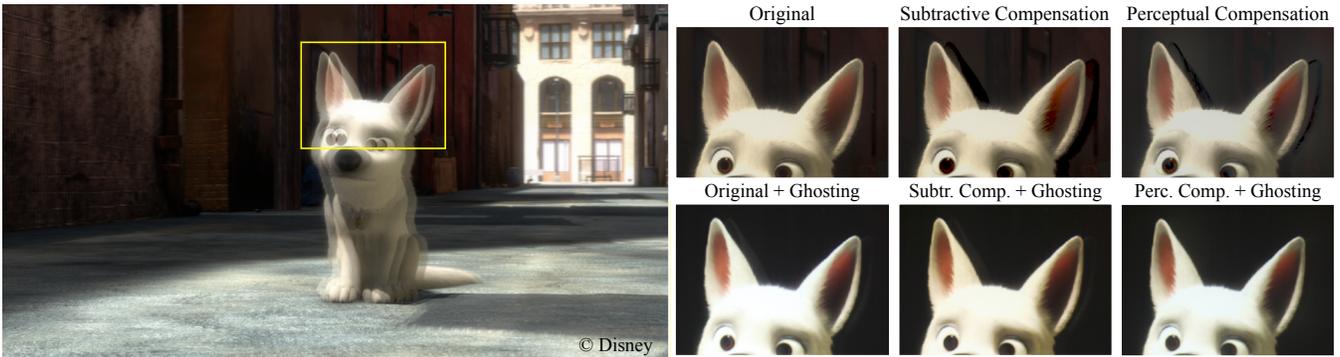
Figure 2: *An example of ghosting and deghosting. Left: To illustrate the disparities the left and right eye images are superimposed. Right: In the top row a comparison between original input, subtractive compensation and perceptual compensation. The bottom row is the same as the top row but now shows observed images acquired with a camera through the eyewear polarizing filter. (We encourage readers to view results electronically, adjusting brightness and gamma settings if images appear too dark.)*

is on ghosting, our perceptually-based framework is more general. As an example, we apply our framework to concave projection-based spherical displays (e.g., IMAX Dome cinemas), which suffer from scattering: unintended indirect illumination due to reflection.

The remainder of our paper is organized as follows: Section 2 discusses related work. Our general perceptually-based compensation framework is explained in Section 3 and its application to deghosting in Section 3.4. We discuss results, including a user evaluation, in Section 4. In Section 5 we propose a model for pollution prediction. Section 6 discusses descattering using our framework. Discussion and avenues for future work are presented in Section 7, followed by conclusions in Section 8.

2   Related Work

Ghosting and Stereopsis   The illusion of stereoscopic depth is dependent on a viewer's ability to fuse corresponding features or edges presented to the two eyes [Howard and Rogers 2002]. Viewing stereoscopic images is a demanding task for the HVS. It requires a decoupling between focus and eye vergence that has been demonstrated to influence not only viewer discomfort, but also hinder visual attention and depth discrimination [Hoffman et al. 2008].

The perception of ghosting can be considered a"binocular noise" that further hinders fusion limits and visual comfort. Yeh and Silverstein demonstrated that crosstalk significantly influences the ability to fuse widely separated images via binocular eye vergence movement [1990]. Ghost images may introduce unintended edges and binocular rivalry making visual processing unstable, unpredictable, and impair guiding visual attention [Patterson 2007]. It has also been found to inhibit the interpretation of depth [Tsirlin et al. 2011].

Use of even minimal crosstalk has been found to strongly affect subjective ratings of display image quality and visual comfort [Yeh and Silverstein 1990; Kooi and Toet 2004]. Although acceptable crosstalk may generally be as high as 5-10%, the detection and acceptability thresholds can be significantly reduced with higher image contrast or larger disparity [Wang et al. 2011]. There is a strong need to remove the detection of crosstalk.

Deghosting.   Subtractive compensation methods for active (time-sequential) and passive (light modulation) stereo display systems, subtract the predicted ghosting contribution prior to display [Konrad et al. 2000; Klimenko et al. 2003]. These methods assume that there is sufficient *signal* to subtract from since physical systems cannot inject negative light. To fully compensate in these cases, the black

level is raised globally (automatically), or locally (manually). Smit et al. [2007] proposed a perceptually motivated extension to subtractive compensation. They perform subtraction in the perceptually uniform *CIE-Lab*, instead of *RGB* color space. This results in less visible ghosting compared to standard subtractive methods. However, in the case of low luminance, their method suffers from the same problem as other subtractive methods, and leaves ghosting as uncorrectable. We specifically want to address these "uncorrectable" cases, and instead propose a perceptually-based distribution of the ghosting error to reduced sensitivity regions of the HVS.

The HVS and Perceptual Models.   Perceptually-based methods have been extensively used and an exhaustive list would be beyond the scope of this paper. We discuss the most relevant works. Perceptual models have been exploited to determine if the texture of objects masks an underlying coarse tessellation [Ferwerda et al. 1997], and in stopping criteria for global illumination [Ramasubramanian et al. 1999; Mantiuk et al. 2006; Longhurst et al. 2006; Sundstedt et al. 2007]. We explicitly exploit perceptual models for formulating an optimization framework.

Tone mapping involves the display of an image with a higher dynamic range on a display with a lower dynamic range [Reinhard et al. 2002; Mantiuk et al. 2008]. The method that we propose is related to local tone mapping in that our goal is to take an image with a higher dynamic range (the intended image), and display it on a (locally) lower dynamic range display (due to the light pollution). We aim to distribute the error smoothly in a local region, and we propose to exploit HVS properties to do this in a perceptually more optimal manner.

Majumder and Stevens [2005] aim to obtain a global smoothly varying luminance in a multi-projector display by incorporating perceptual metrics. Grosse et al. [2010] expoit the CSF to precompute a binary mask, which in turn is used to compute an optimal coded aperture. Our work differs in that our goal is to locally compensate for ghosting only. This requires a different formulation for the optimization problem which we describe in Section 3.

In image processing domain, perceptual metrics are incorporated into a Visible Difference Predictors (VDP), which aims to quantify the perceptual difference between a reference and a test image [Daly 1992; Lubin 1995]. Furthermore, Nadenau et al. [2001] propose to exploit the contrast sensitivity function (CSF) for weighting the coefficients of a wavelet decomposition at different levels [Nadenau et al. 2001]. We propose to incorporate the CSF and components of the VDP directly in our optimization framework.
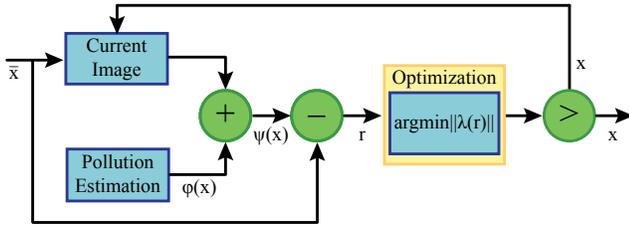
Figure 3: *Dataflow for our perceptual compensation.*

Descattering. Several subtractive compensation methods have been suggested to compensate for indirect scattering by modifying the image before projection [Bimber et al. 2006; Mukaigawa et al. 2006; Wetzstein and Bimber 2007; Dehos et al. 2008]. Bimber et al. [2007] provides a more comprehensive overview of other related work in this field. All these methods suffer from the same problems as subtracive deghosting methods. We propose an analytic subtractive compensation for spherical domes, that operates on full-resolution images, and combine this with our perceptual framework to redistribute this error into visually less important regions.

3 Perceptual Framework

In the following, matrices are denoted with uppercase bold letters, while vectors are denoted with lowercase bold letters.

3.1 Minimization Formulation

Although we focus on light pollution due to ghosting, we will discuss our perceptual framework for the general case of *additive* light pollution, shown in Figure 3. Light pollution is typically addressed by subtractive compensation of the input images followed by clamping. This can mathematically be expressed as a constrained minimization problem if we introduce the observation function $\psi(\cdot)$:

$$\underset{\mathbf{x}}{\operatorname{argmin}}\|\psi(\mathbf{x}) - \bar{\mathbf{x}}\|^2, \quad \text{s.t. } 0 \leq \mathbf{x} \leq 1. \tag{1}$$

Here $\mathbf{x}$ and $\bar{\mathbf{x}}$ denote the compensated and the desired image respectively. The observation function $\psi(\cdot)$, takes into account additive light pollution in the physical system and can be expressed as the sum:

$$\psi(\mathbf{x}) = \mathbf{x} + \varphi(\mathbf{x}), \tag{2}$$

with pollution function $\varphi(\cdot)$. The constraints in Equation 1 are critical to ensure that the pixel values remain within the physically attainable value range.

Our goal is to compute a compensated input image $\mathbf{x}$ resulting in an observed image which is *perceptually* as close as possible to the desired image $\bar{\mathbf{x}}$. To achieve this we define an abstract perceptually based weighting function $\lambda(\cdot)$, which can include any combination of the psychophysical and physiological aspects of the HVS. We then use $\lambda(\cdot)$ for weighting the residual $\mathbf{r} = \mathbf{x} + \varphi(\mathbf{x}) - \bar{\mathbf{x}}$ and rewrite Equation 1 as:

$$\underset{\mathbf{x}}{\operatorname{argmin}}\|\lambda(\mathbf{r})\|^2, \quad \text{s.t. } 0 \leq \mathbf{x} \leq 1. \tag{3}$$

The above formulation applies to a single channel. For color images we solve Equation 3 separately for each channel. This requires the images to be transformed to a color space with independent channels. Sheng et al. [2010] suggest the *YCbCr* color space, which is a good approximation of the perceptually uniform *CIE-Lab* color space.

It would in general be impractical, if not impossible, to solve Equation 3 as $\lambda(\cdot)$ may be highly non-linear. To make Equation 3
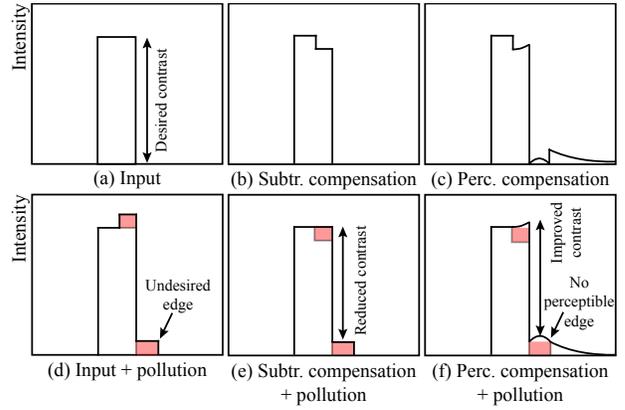

Figure 4: *A 1-D illustration comparing no compensation, subtractive compensation and CSF weighting (perceptual) compensation. (a)-(c) The input images. (d)-(f) The observed images. Pollution is depicted as the red shaded areas. Perceptual compensation smoothes the high contrast edge of the ghost, while simultaneously maximizing the contrast of the original input right-hand edge.*

amenable to efficient computation, we aim to linearize Equation 3. We propose to first apply the CSF as a weighting function directly (3.2). Additional non-linear perceptual models are computed only once incorporated as a weighting mask in the spatial domain (3.3).

3.2 Linear Perceptual Weighting by the CSF

The CSF describes the varying sensitivity of the HVS to spatial frequencies, with a distinct peak and falloff towards low and high frequencies. Weighting the residual $\mathbf{r}$ in Equation 3 with the CSF coefficients distributes the error to frequencies where the HVS is least sensitive. This results in smoothing of edges due to pollution, while non-pollution details are preserved, as illustrated in Figure 4. The figure compares no compensation, subtractive compensation ($\lambda(\cdot)$ is one) and perceptual compensation ($\lambda(\cdot)$ is CSF) of a 1-D signal. The additive pollution is drawn as shaded red areas. For perceptual compensation, Figure 4f, the contrast of the input's right edge is maximized while smoothing the edge due to pollution.

The CSF is most naturally expressed in the frequency domain, but we can represent it in matrix form as:

$$\Lambda_l = \mathcal{F}^{-1} \Omega \mathcal{F}, \tag{4}$$

where Ω is a diagonal matrix of the CSF spectral coefficients and $\mathcal{F} = \mathbf{F_y} \otimes \mathbf{F_x}$ is the two dimensional discrete Fourier transform. We can thus directly apply the CSF as the weighting function, with $\lambda(\mathbf{r}) \equiv \Lambda_l \cdot \mathbf{r}$ we have:

$$\underset{\mathbf{x}}{\operatorname{argmin}}\|\Lambda_l \cdot \mathbf{r}\|^2, \quad \text{s.t. } 0 \leq \mathbf{x} \leq 1. \tag{5}$$

The challenge is that Λ_l is a dense $n \times n$ matrix, with n the number of pixels in the input image. For typical digital cinema quality content, this would result in a matrix with roughly 10^{13} non-zero elements, and would require over a terabyte of memory just to store. Fortunately, by the convolution theorem, we can also express Equation 5 as a convolution:

$$\underset{\mathbf{x}}{\operatorname{argmin}}\|\mathbf{K}_l * \mathbf{r}(\mathbf{x})\|^2, \quad \text{s.t. } 0 \leq \mathbf{x} \leq 1, \tag{6}$$

where $\mathbf{K}_l$ is the spatial convolution kernel corresponding to Λ_l. Equation 6 yields a linear system with a block Toeplitz with Toeplitz

blocks (BTTB) structure. We can exploit this fact to omit the need to store Λ_l explicitly and compute $\Lambda_l \mathbf{r}$ on demand instead. By taking a circulant extension of the BTTB structure, matrix-vector multiplications can be performed using 2D FFTs [Vogel 2002]. This allows us to solve Equation 6 for our large problems using for example the conjugate gradients with gradient projection (GPCG) method [Moré and Toraldo 1991]. We note that Equation 6 is similar in structure to non-blind deconvolution problems [Banham and Katsaggelos 1997], which aim to reconstruct some true signal from corrupted observations given a known corruption kernel.

3.3 Non-Linear Perceptual Weighting

We observe that although pollution is physically always present, the actual pollution may be near or below the perceptual threshold of visibility. This could be incorporated into Equation 6 as an additional weighting. For efficiency, non-linear models would be computed once and represented as a diagonal matrix Λ_n. This can be combined with matrix Λ_l and model the perceptual weighting λ of $\mathbf{r}$ as $\lambda(\mathbf{r}) \equiv \Lambda_n \Lambda_l \mathbf{r}$. Equation 3 thus becomes:

$$\underset{\mathbf{x}}{\operatorname{argmin}} \| \Lambda_n \Lambda_l \mathbf{r}(\mathbf{x}) \|^2, \quad \text{s.t.} \ \ 0 \le \mathbf{x} \le 1. \tag{7}$$

For the remainder of this and the next section we set Λ_n to identity. We discuss Λ_n further in Section 5.

3.4 Deghosting

Stereoscopic Input Images To apply our framework to deghosting, we have to take into account that the input is a (stereo) pair of images: left ($\mathbf{x}_L$) and right ($\mathbf{x}_R$). Hence, $\varphi(\cdot)$ is now a function of two input images:

$$\psi(\mathbf{x}_L) = \mathbf{x}_L + \varphi(\mathbf{x}_L, \mathbf{x}_R) \, ; \, \psi(\mathbf{x}_R) = \mathbf{x}_R + \varphi(\mathbf{x}_R, \mathbf{x}_L), \tag{8}$$

where $\psi(\mathbf{x}_L)$ and $\psi(\mathbf{x}_R)$ are the observed images in the presence of ghosting.

We stack $\mathbf{x}_L$ and $\mathbf{x}_R$ into a single vector $\mathbf{x}$ and furthermore assume that during a single iteration of GPCG, $\varphi(\mathbf{x})$ remains constant. We can then write $\mathbf{r} = \mathbf{x} - (\bar{\mathbf{x}} - \varphi(\mathbf{x}))$. Plugging this into Equation 7 we get:

$$\underset{\mathbf{x}}{\operatorname{argmin}} \| \Lambda_n \Lambda_l \mathbf{x} - \Lambda_n \Lambda_l (\bar{\mathbf{x}} - \varphi(\mathbf{x})) \|^2, \quad \text{s.t.} \ \ 0 \le \mathbf{x} \le 1. \tag{9}$$

In practice we precompute $\Lambda_n \Lambda_l (\bar{\mathbf{x}} - \varphi(\mathbf{x})) = \mathbf{x}'$, perform one iteration of GPCG and then use the current solution $\mathbf{x}$ to update $\varphi(\mathbf{x})$, recompute $\mathbf{x}'$, and continue with the next iteration until convergence. Also, whenever we need to compute $\Lambda_n \Lambda_l \mathbf{x}$, we first un-stack $\mathbf{x}$, compute $\Lambda_n \Lambda_l \mathbf{x}_L$ and $\Lambda_n \Lambda_l \mathbf{x}_R$, and re-stack $\mathbf{x}$.

4 Results

To evaluate our perceptual framework for deghosting we superimposed two manually aligned[1] projectors (1280×720 @ 2.5m) using circular polarization, see Figure 5. The projectors' response (in *CIE-XYZ*) are measured using a colorimeter, and we found that $R + G + B = W$, with a gamma value of 2.2. We used the common gamut mapping method [Stone et al. 1988] to correct for color and brightness. We measured the pollution function $\varphi(\cdot)$ by mounting the eyewear's polarizing filter in front of the colorimeter. We measure the *XYZ* values for the *RGB* primaries: once for the *intended eye image*, and once for the *unintended eye image* (for both projectors).

---

[1] alignment is pixel-accurate except for the periphery of the display
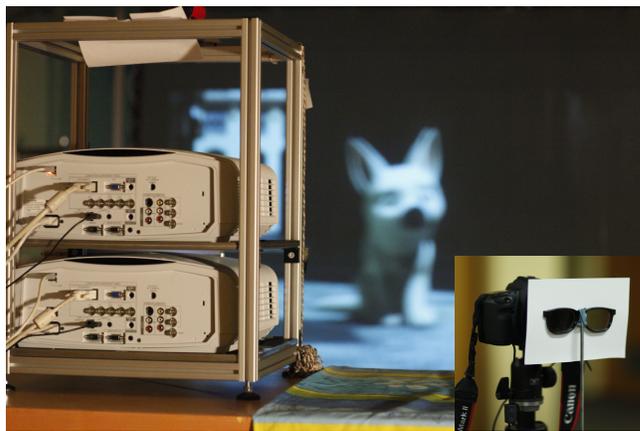
Figure 5: *Our experimental rear-projection setup consisting of two projectors using polarization filtering for generating separate left and right eye images. Results are captured by mounting the eyewear in front of a camera lens (inset).*

We finally gamma correct the *RGB* values and transform $\varphi(\cdot)$ to the $YCbCr$ color space for computing the compensation.

Figure 2 compares no compensation, subtractive compensation and perceptual compensation for a zoomed-in region. The top row on the right-hand side shows the compensated (input) images. The bottom row shows the observed images on our experimental setup. Subtractive compensation cannot compensate for the pollution, and the ghosting is nearly as strong as it is for the original image. The perceptually compensated result shows that the residual has been distributed smoothly to make the ghosting edge imperceptible. The left and right input images are superimposed on the left-hand side to illustrate the disparities. Figure 6 shows additional comparisons. The top row shows the original input image, and the bottom row shows zoomed-in regions for the observed subtractive and perceptual compensation images. The perceptual compensation distributes the residual error smoothly. We point out again that the lack of a visible ghost edge in the perceptually compensated images leads to more comfortable stereoscopic viewing, as verified with our user study (see Section 4.1).

All our results have been generated for a "sweet-spot" location. We use the model proposed by Daly [1992] for generating a 2D CSF. As CSF parameters we use the projection resolution and size, a viewing distance of 3.0 m, light adaptation of 5 cd/m^2, and eccentricity zero.
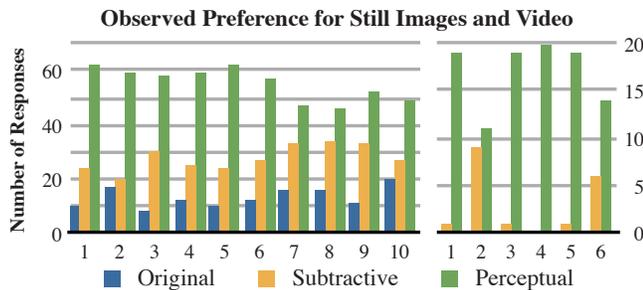
4.1 User Evaluation

We conducted a user evaluation to determine whether perceptual compensation of stereoscopic images and video is more comfortable to view. The experiments apply a single factor, the compensation strategy, with up to three different conditions: no compensation (original image), subtractive compensation, and perceptual compensation. We did not include compensation by raising the black level globally, as this significantly changes the image compared to the intended image. Using a forced-choice, pairwise comparison design, participants were presented balanced trials consisting of two images that differ only in compensation strategy. Participants then chose which one is more comfortable to view.

For both still images and video, we observed strong statistical evidence that the perceptually-based compensation is more comfortable to view. For still images we collected from 960 total balanced trials across 16 participants, while for video we collected across 120
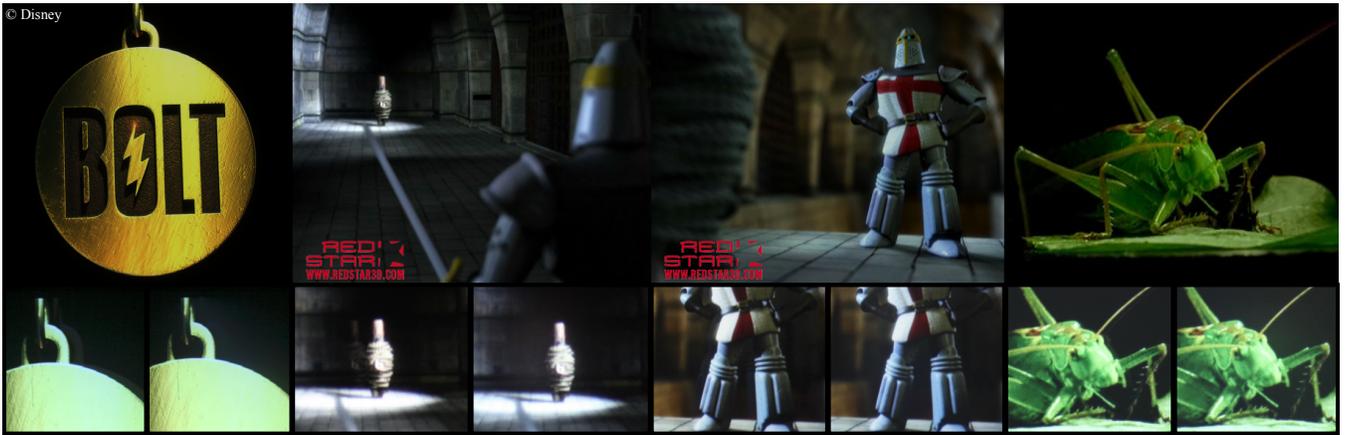
Figure 6: *Examples of ghosting and deghosting on the badge, indoor, knight and grasshopper scenes. The top row shows the non-compensated input image. The bottom row shows a side-by-side comparison of observed images acquired with a camera for subtractive compensation and perceptual compensation, for a selected area.*

Figure 7: *Significant preference for perceptual compensation is present in all ten images (left) and five of six videos (right) of our two user evaluations. The outlier, video #2, reaffirms the value of ghosting prediction (see section 4.1)*

total balanced trials from 10 participants, see Figure 7. We applied Pearson's chi-squared goodness-of-fit test to analyze the participant preferences [Sheskin 2007]. Perceptual compensation is significantly preferred over the original image and subtractive compensation for both still images and video ($p = 0.01$ and $\chi^2(2, 960) = 283.0$, $p = 0.01$ and $\chi^2(1, 120) = 58.8$ resp.). Video #2 was the only non-significant result. The ghosting for this case is in a relatively small and visually less important region of the image (included as Bolt Indoor in the accompanying video). This result reaffirms the value of using ghosting prediction.

5 Pollution Prediction

Due to the size of the problem in Equation 6, performance and convergence is slow. The size of the optimization problem can be reduced by restricting only to areas where pollution is noticeable. The diagonal matrix Λ_n can be exploited as a predictor by turning this into a *binary* weighting mask in the spatial domain. For prediction we propose to use the threshold-vs-intensity (TVI) and visual masking in Λ_n [Ferwerda et al. 1996; Ferwerda et al. 1997].

The TVI describes the minimum contrast required to distinguish between foreground and background intensities. A per-pixel test is performed to check whether the residual $\psi(\mathbf{x}) - \bar{\mathbf{x}}$, averaged over some area, is above a threshold:

$$\delta(\mathbf{x}) = (\psi(\mathbf{x}) - \bar{\mathbf{x}}) > \Delta_{TVI}. \qquad (10)$$

For mesopic luminance level settings, e.g. cinema, Ferwerda et

al. [1996] suggest to blend between values computed by photopic and scotopic models.

Mechanisms in the visual system are tuned to different frequency and orientations bands, and visual masking describes the reduction in contrast sensitivity due to interactions between image components within mechanism bands. We use the model from Ferwerda et al. [1997] with $\bar{\mathbf{x}}$ and $\psi(\mathbf{x})$ as the reference and test images to compute per-pixel masking values:

$$\upsilon(\mathbf{x}) = \sum_i \sum_\theta \Delta R_{i,\theta}^2(\mathbf{x}), \qquad (11)$$

where $\Delta R_{i,\theta}$ is the difference in response of a mechanism with frequency band i and orientation θ, to a reference and test image.

To determine the spatial domain weighting mask Λ_n we normalize and combine maps $\delta(\mathbf{x})$ and $\upsilon(\mathbf{x})$ using component-wise multiplication:

$$\Lambda_n = \delta(\mathbf{x}) \odot \upsilon(\mathbf{x}). \qquad (12)$$

To turn Λ_n into a binary mask we compare each pixel against a threshold t_{Λ_n}. The predictor can be further extended by considering only ghosting in visually important regions, e.g., Harel et al.[2007] proposed an MRF-based approach to predict object regions. Saliency masking $\gamma(\mathbf{x})$ is incorporated in the prediction as:

$$\Lambda_n = \gamma(\mathbf{x}) \odot (\delta(\mathbf{x}) \odot \upsilon(\mathbf{x})). \qquad (13)$$

With deghosting, the pollution is due entirely to unintended light contribution from the *other* eye image. To better predict the noticeability of the ghosting we use the saliency of that eye's image. Thus, for the left eye we have:

$$\Lambda_n^{\text{left}} = \gamma(\mathbf{x})_{\text{right}} \odot (\delta(\mathbf{x})_{\text{left}} \odot \upsilon(\mathbf{x})_{\text{left}}), \qquad (14)$$

and similarly for Λ_n^{right}.

We implemented this prediction model for our experimental setup. We extracted a 1D CSF from the generated 2D CSF in Section 4, and this CSF is used in the masking model. Adaptation luminance for the TVI is computed over a small area as proposed by Ramasubramaniam et al. [1999] (using the *XYZ* measurements obtained for our experimental setup). Figure 8 shows an example of ghosting prediction, with $t_{\Lambda_n} = 0.05$.

Table 1 shows the timing results for our algorithm implemented in CUDA. The first row shows the performance for a full resolution
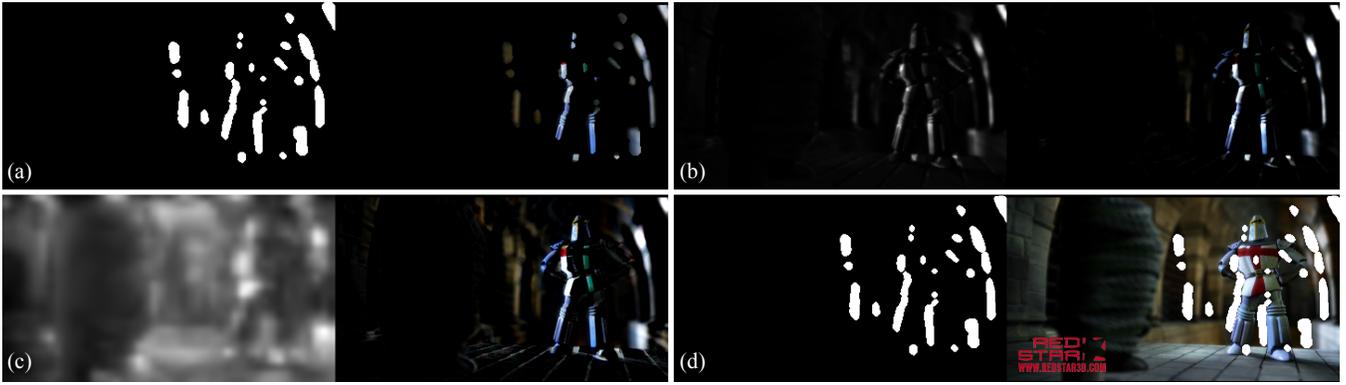
Figure 8: *Example of the ghosting prediction map. For each of the four images in this figure the computed map is shown on the left-hand side, and the modulation of the map with the residual is shown on the right-hand side: (a) TVI map, (b) visual masking map, (c) saliency map, and (d) thresholded prediction map.*

(2K) input image. Subsequent rows show performance for examples of areas determined by the pollution prediction (Section 5). As compensation is only required for these areas, the runtimes for smaller areas greatly reduce. Even with prediction, Table 1 shows that our method is not suitable for time-critical applications, but rather for applications which allow the compensation to be performed in an offline preprocess.

Table 1: *Performance of CUDA implementation.*

Resolution	#FFTs	#Iterations	Runtime (secs)
2028×1080	3772	159	248.5557
534×844	3992	167	128.4280
435×505	3684	154	34.3573
196×232	2636	109	4.1634

6 Descattering

Our perceptual framework can also be applied to the problem of descattering for immersive displays, specifically IMAX Dome projections. In this case, the intended image is degraded due to light pollution caused by multiple bounces of indirect scattering. With descattering, simply evaluating the pollution term $\varphi(\cdot)$ is a complex and computationally expensive process which, in the general case, involves solving the rendering equation [Kajiya 1986]. However, the screens in IMAX Dome systems are low-gain, resulting in Lambertian reflectance [Lantz 1995; Scott 2008]. This allows us to obtain closed-form solutions.

6.1 Efficient Closed-Form Pollution Estimation

To compute the pollution term $\varphi(\cdot)$, we exploit the relatively unknown fact that the point-to-point form factor within a sphere is a constant. The consequence of this is that the indirect illumination within a Lambertian sphere is spatially uniform, regardless of the projected illumination. This fact was used by Szirmay-Kalos [2000] to obtain a closed-form solution to the one-bounce light transport operator within a closed, perfectly Lambertian sphere, as well as by Hawkins et al. [Hawkins et al. 2005]. We generalize this result to partial spherical sections and multiple bounces. This results in the following analytic expression for the pollution term:

$$\varphi(\mathbf{x}) = \frac{\rho \pi r^2}{4 - \Omega_\mathbf{x} \rho \pi r^2} (\mathbf{a} \cdot \mathbf{x}), \qquad (15)$$

where $\rho \in (0,1)$ is the screen gain (diffuse albedo), $\mathbf{a}$ is a vector specifying the projected area of each pixel onto the screen, $\Omega_\mathbf{x}$

is the projected area of the entire image, and r is the radius of the sphere. For simplicity of notation, we omit the projector-to-screen form factors, but incorporate these in our implementation. Equation 15 computes all bounces of indirect illumination, for all pixels, using a single dot product in $O(n)$ time where n is the number of image pixels. Hence, this computation can be performed efficiently within the inner-loop of perceptual compensation, without down-sampling, even for high-resolution input images typical of IMAX Dome projection.

Figure 1 shows a simulation of an image projected onto an IMAX Dome and compares it to the image observed by the audience once indirect scattering is added by the physical system. The observed image is simulated using the radiosity solution described in Section 6.1. We compare subtractive descattering to our perceptual descattering solution for this image, as well as in Figure 9. In each case, our perceptual compensation increases the relative contrast without the loss of details arising from the light clamping necessary in subtractive compensation approaches.

7 Discussion & Future Work

As discussed, ghosting introduces local error in the form of edges, which is typically compensated for by subtraction prior to display. For low background luminance areas this error can only be hidden by reducing the contrast. We propose an automatic method for local contrast reduction. Under the non-negative light constraint, we apply a weighted L2 norm to the residual according to perceptual metrics. Residual error is then distributed into regions of lower sensitivity for the HVS. The formulation of light pollution compensation as an optimization problem is a generalization of the subtractive compensation approaches. It is specifically aimed to handle areas which previous work classifies as uncorrectable.

Our approach has certain limitations. First, computational complexity is high due to the fact that Λ_l is dense. Improved performance may be obtained by formulating an approximation to the perceptual metrics with more desirable properties for optimization. Second, our method relies on a controlled illumination environment, and results are computed for a sweet-spot location. SMPTE D-Cinema specifications [2008] were developed to ensure uniformity among digital cinemas and offline pre-computed deghosting material would thus be valid for most digital cinemas. Furthermore, measurements in a real cinema show an increase in ghosting from 1.1% at the sweet-spot, to 2.0% at the periphery. We evaluated this on our setup. Figure 10 shows subtractive and perceptual compensation for locations left off-axis, sweet-spot and right off-axis on our experimental
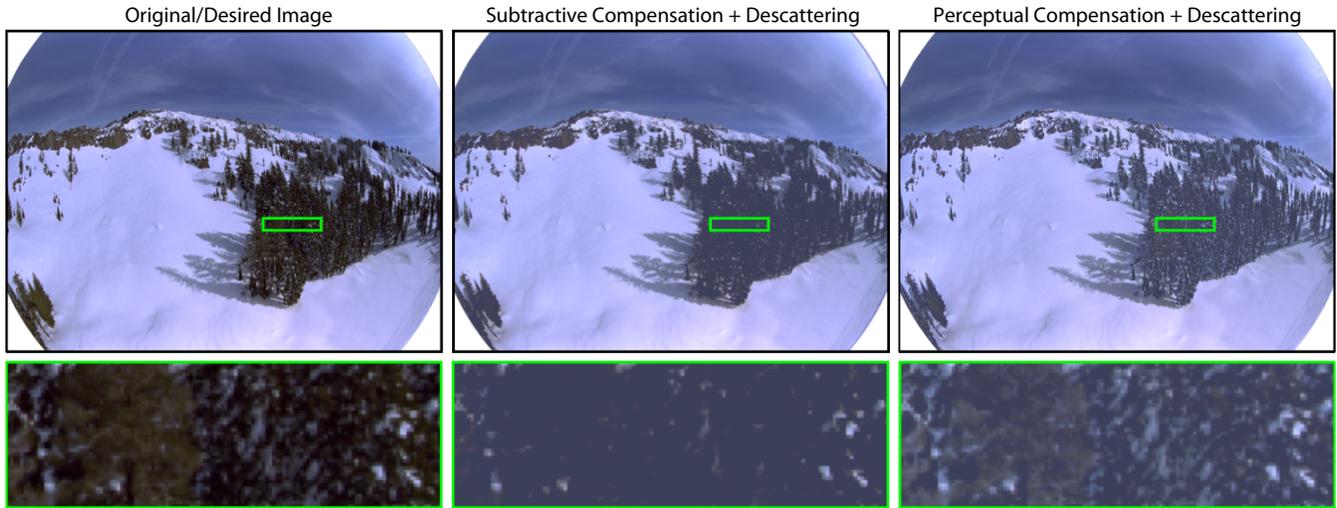
Figure 9: *When projecting onto a spherical dome, the projected image (left) is corrupted by indirect scattering. Subtractive methods (middle) can only correctly compensate the projected image in image regions with enough signal and negative values are clamped to black, leading to loss of detail in the observed image. Our perceptual compensation (right) retains more of these dark details while maintaining a final observed image that is perceptually closer to the unpolluted original image.*
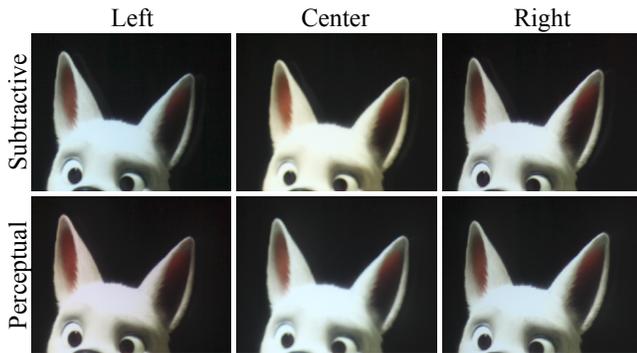
Figure 10: *We compare compensations for left off-axis, center (sweet-spot) and right off-axis locations. The top and bottom rows compare subtractive and perceptual compensations. Although the ghosting contribution increases for off-axis locations, the perceptual compensation still increases viewing comfort.*

setup. Although an increase in ghosting for left and right off-axis can be observed, the off-axis locations still benefit from the perceptual compensation.

We only validated the prediction (Section 5) experimentally. For a truly effective prediction, a more rigorous validation should be performed. Recent work explores subjective detection or acceptability thresholds for ghosting [Wang et al. 2011]. Future work to obtain a more thorough understanding of discomfort and fatigue for viewing stereoscopic imagery in general will be necessary. Further research in the spatio-temporal domain and the general concept of stereoscopic saliency could help improve this understanding. Reducing any discomfort for stereoscopic viewing will be especially important with the increasing popularity of 3D television, games and mobile devices.

8 Conclusion

We have descrived a framework for compensation of light pollution, formulated as a perceptually-based optimization problem. Light pollution is defined as the unintended light contribution onto (por-

tions of) an intended image. The residual error of light pollution is distributed to regions with less sensitivity of the HVS. Our formalization is a generalization of exisiting subtractive compensation methods. We described two compensation applications: deghosting in stereoscopic 3D display systems, and descattering for concave projection-based displays. Most importantly the perceptibility of conflicting edge cues is reduced for perceptual deghosting, which makes watching stereoscopic 3D displays more comfortable. For spherical display surfaces often used in immersive displays, we have developed an analytic formulation for the scattering pollution.

Acknowledgements

References

BANHAM, M., AND KATSAGGELOS, A. 1997. Digital Image Restoration. *IEEE Signal Proc. Mag. 14*, 2 (March), 24–41.

BIMBER, O., GRUNDHOFER, A., ZEIDLER, T., DANCH, D., AND KAPAKOS, P. 2006. Compensating indirect scattering for immersive and semi-immersive projection displays. In *Proc. of the IEEE Conf. on Virtual Reality*, 151–158.

BIMBER, O., IWAI, D., WETZSTEIN, G., AND GRUNDHÖFER, A. 2007. The visual computing of projector-camera systems. In *STAR Proc. of Eurographics 2007*, 23–46.

DALY, S. J. 1992. The visible differences predictor: an algorithm for the assessment of image fidelity. In *Human Vision, Visual Processing, and Digital Display III*, SPIE, vol. 1666, 2–15.

2008. Digital cinema system specification v1.2. Digital Cinema Initiatives, LLC, http://www.dcimovies.com.

DEHOS, J., ZEGHERS, E., RENAUD, C., ROUSSELLE, F., AND SARRY, L. 2008. Radiometric compensation for a low-cost immersive projection system. In *Proc. of VR Software and Tech.*, 130–133.

FERWERDA, J. A., PATTANAIK, S. N., SHIRLEY, P., AND GREEN-

BERG, D. P. 1996. A Model of Visual Adaptation for Realistic Image Synthesis. In *Proc. of SIGGRAPH '96.*

FERWERDA, J. A., SHIRLEY, P., PATTANAIK, S. N., AND GREENBERG, D. P. 1997. A Model of Visual Masking for Computer Graphics. In *Proc. of SIGGRAPH '97.*

GROSSE, M., WETZSTEIN, G., GRUNDHÖFER, A., AND BIMBER, O. 2010. Coded aperture projection. *ACM Trans. Graph. 29* (July), 22:1–22:12.

HAREL, J., KOCH, C., AND PERONA, P. 2007. Graph-based visual saliency. In *Advances in Neural Information Processing Systems*, MIT Press, Cambridge, MA, B. Schölkopf, J. Platt, and T. Hoffman, Eds., vol. 19.

HAWKINS, T., EINARSSON, P., AND DEBEVEC, P. 2005. A dual light stage. *Rendering Techniques.*

HOFFMAN, D. M., GIRSHICK, A. R., AKELEY, K., AND BANKS, M. S. 2008. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *J. Vis. 8*, 3 (3), 1–30.

HOWARD, I. P., AND ROGERS, B. J. 2002. *Seeing in Depth: Volume 2: Depth perception.* Oxford University Press, NY, USA.

KAJIYA, J. T. 1986. The rendering equation. In *Comp. Graph. (Proceedings of SIGGRAPH 86)*, ACM, vol. 13, 143–150.

KLIMENKO, S., FROLOV, P., NIKITINA, L., AND NIKITIN, I. 2003. Crosstalk reduction in passive stereo-projection systems. In *Proceedings of Eurographics*, 235–240.

KONRAD, J., LACOTTE, B., AND DUBOIS, E. 2000. Cancellation of image crosstalk in time-sequential displays of stereoscopic video. In *IEEE Trans. on Image Processing*, vol. 9, 897–908.

KOOI, F. L., AND TOET, A. 2004. Visual comfort of binocular and 3d displays. In *Displays*, vol. 25, 99–108.

LANTZ, E. 1995. Graphics design and production for hemispheric projection. ACM SIGGRAPH 1995 Course #2 Notes, ch. Spherical Image Representation and Display: A New Paradigm for Computer Graphics.

LONGHURST, P., DEBATTISTA, K., AND CHALMERS, A. 2006. A GPU based Saliency Map for High-Fidelity Selective Rendering. In *Proceedings of AFRIGRAPH '06.* 21–29.

LUBIN, J. 1995. A visual discrimination model for imaging system design and evaluation. In *Vision Models for Target Detection and Recognition*, World Scientific, 245–283.

MAJUMDER, A., AND STEVENS, R. 2005. Perceptual Photometric Seamlessness in Projection-Based Tiled Displays. *ACM Trans. on Graph. 24*, 1 (January), 118–139.

MANTIUK, R., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2006. A perceptual framework for contrast processing of high dynamic range images. *ACM Trans. Appl. Percept. 3*, 3, 286–308.

MANTIUK, R., DALY, S., AND KEROFSKY, L. 2008. Display adaptive tone mapping. *ACM Trans. Graph. 27* (August), 68:1–68:10.

MORÉ, J. J., AND TORALDO, G. 1991. On the solution of large quadratic programming problems with bound constraints. *SIAM Journal of Optimization 1*, 1 (February), 93–113.

MUKAIGAWA, Y., KAKINUMA, T., AND OHTA, Y. 2006. Analytical compensation of inter-reflection for pattern projection. In *Proc. of VR Software and Tech.*, ACM, New York, NY, USA.

NADENAU, M. J., REICHEL, J., AND KUNT, M. 2001. Wavelet-based color image compression: Exploiting the contrast sensitivity function. In *IEEE Trans. on Image Proc.*, vol. 12, 58–70.

PATTERSON, R. 2007. Human factors of 3-d displays. *Society for Information Display 15*, 861–871.

RAMASUBRAMANIAN, M., PATTANAIK, S. N., AND GREENBERG, D. P. 1999. A Perceptually Based Physical Error Metric for Realistic Image Synthesis. In *Proc. of SIGGRAPH '99.*

REINHARD, E., STARK, M., SHIRLEY, P., AND FERWERDA, J. 2002. Photographic Tone Reproduction for Digital Images. In *Proc. of SIGGRAPH.*

SCOTT, K. 2008. *Planetarium Development Guide.* Int. Planetarium Society, ch. Theater Configuration.

SHENG, Y., YAPO, T. C., AND CUTLER, B. 2010. Global illumination compensation for spatially augmented reality. *Computer Graphics Forum 29*, 2.

SHESKIN, D. J. 2007. *Handbook of Parametric and Nonparametric Statistical Procedures.* Chapman & Hall/CRC.

SMIT, F. A., VAN LIERE, R., AND FROEHLICH, B. 2007. Three extensions to subtractive crosstalk reduction. In *Eurographics Symp. on Virtual Env.*, 85–92.

STONE, M. C., COWAN, W. B., AND BEATTY, J. C. 1988. Color Gamut Mapping and the Printing of Digital Color Images. *ACM Trans. on Graph. 7*, 4, 249–292.

SUNDSTEDT, V., GUTIERREZ, D., ANSON, O., BANTERLE, F., AND CHALMERS, A. 2007. Perceptual rendering of participating media. *ACM Trans. Appl. Percept. 4*, 3, 15.

SZIRMAY-KALOS, L., 2000. Monte-carlo methods in global illumination. unpublished.

TSIRLIN, I., WILCOX, L. M., AND ALLISON, R. S. 2011. The effect of crosstalk on depth magnitude in thin structures. In *Proc. of SPIE-IS&T Electronic Imaging*, vol. 7863.

VOGEL, C. R. 2002. *Computational Methods for Inverse Problems.* Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.

WANG, L., TEUNISSEN, K., TU, Y., CHEN, L., ZHANG, P., ZHANG, T., AND HEYNDERICKX, I. 2011. Crosstalk evaluation in stereoscopic displays. *J. of Display Technology 7*, 4, 208–214.

WETZSTEIN, G., AND BIMBER, O. 2007. Radiometric compensation through inverse light transport. In *Pacific Conf. on Comp. Graph. and Appl.*, IEEE Computer Society, 391–399.

YEH, Y., AND SILVERSTEIN, L. 1990. Limits of fusion and depth judgment in stereoscopic color displays. *Human Factors 32*, 1.