Optimization of freeform lightpipes for light-emitting-diode projectors

Florian Fournier* and Jannick Rolland

CREOL, College of Optics and Photonics, University of Central Florida, Orlando, Florida 32816, USA

*Corresponding author: ffournie@creol.ucf.edu

Received 18 October 2007; revised 10 December 2007; accepted 14 December 2007;
posted 9 January 2008 (Doc. ID 88701); published 28 February 2008

Standard nonimaging components used to collect and integrate light in light-emitting-diode-based projector light engines such as tapered rods and compound parabolic concentrators are compared to optimized freeform shapes in terms of transmission efficiency and spatial uniformity. We show that the simultaneous optimization of the output surface and the profile shape yields transmission efficiency within the étendue limit up to 90% and spatial uniformity higher than 95%, even for compact sizes. The optimization process involves a manual study of the trends for different shapes and the use of an optimization algorithm to further improve the performance of the freeform lightpipe.

OCIS codes: 220.4298, 220.2945, 120.2040, 230.3670.

1. Introduction

With the improvement in solid state lighting, light-emitting diodes (LEDs) are becoming increasingly attractive for displays. LEDs are already being used in monitors, and they have more recently appeared in projection displays. Projection displays have strong constraints on the spatial and angular extents of the beam in their light engine, mainly because of the small size of the microdisplay and the limited numerical aperture required for good contrast. Such systems are said to be étendue limited. In this case, the brightness of the projector depends on the luminance of the source, and not on its luminous flux [1]. Unfortunately, LED luminance is still an order of magnitude lower than the luminance of a regular arc lamp. For this reason it is currently difficult to obtain LED projectors with brightness levels comparable to lamp-based projectors. Nonetheless, the form factor and low power consumption of LEDs have created a research effort toward compact LED projectors that project images the size of a large computer screen [2–4]. If we look at the rate of increase of LED luminance during the past decade, we can predict that luminance levels of approximately $300\,\mathrm{Mcd/m^2}$ will be reached by 2010, thus matching the effective luminance of arc lamps [5]. Currently, state-of-the-art LEDs using photonic lattices emit up to $280\,\mathrm{lm/mm^2}$ [6]. If we assume they are true Lambertian emitters, it is equivalent to a luminance of $90\,\mathrm{Mcd/m^2}$. This enables projectors with lumen outputs over 400 lm with 0.65 in. (16.5 mm) microdisplays. For a 70 in. (178 cm) projection screen with a gain of 1, we can therefore obtain screen luminance levels of about 100 nits, which is appropriate for viewing in dark environments [7].

The goal of the light engine is to efficiently couple the light emitted by the source to the microdisplay and projection optics. Apart from being collected, light from the source also needs to be shaped according to the format of the microdisplay and mixed so that illumination is spatially uniform. In traditional projectors these functions are performed by multiple components: collection is achieved with metallic reflectors and lenses while shaping and mixing are usually performed with mixing rods or lenslet arrays (fly's eye). One of the advantages of freeform lightpipes is that they can be optimized to perform these three functions at once, thus reducing the number of optical components, their associated Fresnel losses,

0003-6935/08/070957-10$15.00/0

and the alignment requirements. They can be directly coupled to the LED and collect all the emitted light. Classic nonimaging components such as compound parabolic concentrators and tapered rods have already been used for light collection [8]. Gradually tapered rods have also been investigated as a means to improve compactness [9]. However, the emphasis has been the efficiency of these devices and there has been little attention paid to spatial uniformity, which is critical for projection displays. In this paper, we studied what lightpipe shapes can provide for both efficiency and uniformity. The analytical design of such components with efficiency and uniformity requirements is extremely difficult with extended sources, hence the use of optimization techniques. We first study the performance of traditional nonimaging components as a benchmark toward optimized freeform shapes. The optimization of the lightpipe profile and output surface is then discussed.

2. Performance of Traditional Nonimaging Devices

A. Light Engine Characteristics

Our goal is to quantify the performance of freeform lightpipes that couple the LED to the projector microdisplay, as depicted in Fig. 1. A transmissive microdisplay can advantageously be directly coupled at the output of the device, thus reducing the light engine to a single freeform component. When using reflective microdisplays such as liquid crystal on silicon (LCoS) and a digital micromirror device (DMD), a relay lens can be added to image the lightpipe output onto the microdisplay.

For clarity we studied a practical example of projector light engine design. The light source is an Ostar Projection LED manufactured by Osram. Its emissive area is 2.1 mm × 2.1 mm and its emission pattern is quasi-Lambertian. The general étendue formula is given in Eq. (1) (integral formulation) and can be approximated for a flat plane with a uniform emission pattern by

$$E = \iint n^2 \mathrm{d}A\mathrm{d}\omega\cos\theta \approx n^2 A\Omega = n^2 A\pi\sin^2\theta_{1/2}, \quad (1)$$

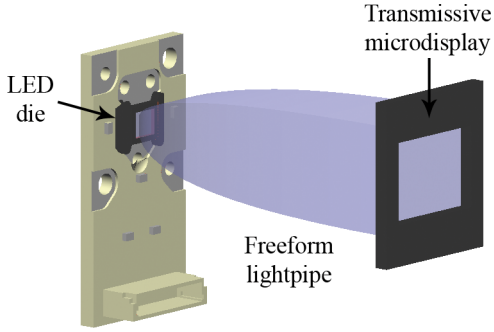

Fig. 1. (Color online) Example of a single-element light engine made of a freeform lightpipe. In this configuration the LED (which is mounted on a board) is directly coupled to the microdisplay.

where A is the emissive area, $\theta_{1/2}$ is the half-angle of the full emission cone, Ω is the projected solid angle of the full emission cone, and n is the refractive index of the material into which light is emitted. Maximum efficiency is achieved when the source étendue is smaller or equal to the étendue limit of the subsequent optical system [7]. For a projector, étendue is often spatially limited by the area of the microdisplay and angularly limited by the numerical aperture of the projection lens. In our case we chose a 0.56 in. (14.2 mm) microdisplay with a 4 : 3 aspect ratio so that the system étendue matches the source étendue for an $f/2.5$ projection lens (NA = 0.2) [1]. A 10% overfill of the microdisplay is included. Higher NAs can be used and will in general make light collection easier, but obtaining good contrast becomes more challenging [7,10,11].

According to Eq. (1), the étendue of the source considered is 13.85 $\mathrm{mm}^2\cdot$sr, given that the Ostar Projection LED has a flat cover and emits directly in air ($n = 1$). This is not the case of most LEDs. The vast majority of LEDs are designed with an integrated lens that improves light extraction by preventing total internal reflection at the LED top surface. Such a lens has two major drawbacks: it prevents placing an optical component close to the die (for direct coupling), and it creates a magnified virtual image of the die. In the case of a hemispherical dome centered on the die, the magnification factor is n^2. As the material index of the lens ranges between 1.4 and 1.6, this factor roughly doubles the étendue of the source. Additionally, internal mirrors can also increase the effective size of the die [12]. In the end, the increase in étendue counterbalances the increase in extracted flux with the lens, so luminance (flux per unit of étendue) remains the same. LEDs with flat covers prevent the magnification of the die at the expense of internally reflected light. But the internally reflected light can be partially recycled, thus increasing the extracted flux without increasing etendue, which means increased luminance [13]. Most lighting applications require high flux but have no étendue constraint; therefore in many cases a high flux approach is privileged. In an étendue-limited system, the higher the source luminance, the higher the collected flux. As LED projectors are étendue limited, luminance is the most critical parameter, and therefore LEDs with flat covers are usually more appropriate.

B. Defining Metrics

To compare the performance of different devices and drive the optimization process, we need to define merit functions or metrics relevant to the desired goals. For a projector light engine, the most common requirements are to maximize the amount of light transmitted by the component and to obtain uniform illumination over the microdisplay. These two requirements are quantified as follows.

- *Transmission efficiency* is the percentage of light from the source transmitted by the lightpipe within the microdisplay and the NA set by the projection optics. Rays exiting the lightpipe with higher angles are considered to be wasted. As explained in Subsection 2.A, the NA of the projection optics and the size of the microdisplay are set to match the étendue of the source. Therefore, 100% efficiency means all the light is transmitted within the étendue limit. In this ideal case, the device is said to preserve étendue.
- *Spatial uniformity* is defined as the relative standard deviation (RSD) of illuminance at the plane of interest (in our case the microdisplay). For a discrete set of N illuminance values E_i it is given by

$$\mathrm{RSD} = \frac{\sigma}{\bar{E}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{E_i}{\bar{E}} - 1\right)^2}. \tag{2}$$

In a simulation software sets of rays are traced from the source to the receiver. A "spot diagram" is then obtained at the receiver plane. To be exploited, this spot diagram is converted into an illuminance map (Fig. 2). For this purpose, the receiver is divided into a grid of cells called the mesh. The higher the number of rays in a given bin, the higher the luminance of that bin. Because ray generation is a statistical process, it creates statistical noise on the receiver. If there are few cells in the mesh, the cells are large; a grid with large cells has inherently less statistical error because the cells contain many rays, but its resolution is lower. Conversely, if there are many cells in the mesh, the cells are small; a grid with small cells is statistically less accurate because the cells contain few rays, but its spatial resolution is higher. Consequently, there is a trade-off between resolution and accuracy. To get better accuracy with more cells, more rays must be traced. Therefore, nonuniformities will at best reach the statistical noise floor in all simulations. Without increasing the number of rays (and the computation time), further optimization cannot improve uniformity.

Another parameter to keep in mind while assessing uniformity is the performance of the human visual system. It is difficult to set a general perception threshold for nonuniformities as the sensitivity of the eye to contrast depends on the spatial frequency and the average illuminance. Typically, nonuniformities produced by lightpipes have low spatial frequencies. Let $\upsilon_{\max}$ be the upper bound of spatial frequencies associated with nonuniformities on the microdisplay in cycles/millimeter. To evaluate the contrast sensitivity of the eye, we need to find the corresponding angular frequency in the projected image as viewed by an observer. For a viewer who looks at the magnified image of the microdisplay projected onto a screen, the upper bound of angular frequencies (in cycles/degree) associated with nonuniformities is given by

$$f_{\max} = \frac{\nu_{\max} w_{\mu D}}{\mathrm{FOV}}, \tag{3}$$

where $\nu_{\max}$ is the upper bound of spatial frequencies associated with nonuniformities on the microdisplay (in cycles/millimeter), $w_{\mu D}$ is the width of the microdisplay and FOV is the field of view subtended by the image from the user's viewpoint. The value of reported contrast sensitivity functions [14] at $f_{\max}$ can then be linked to the RSD of illuminance as shown in the Appendix, thus providing a perception threshold for nonuniformities. The number of rays in the simulations is chosen to ensure that the statistical noise is below this perception threshold. This assumes smooth variations, so sharp changes such as edge effects may still be noticeable even below that threshold and need to be assessed separately. In particular, edge effects are important for multiprojector systems where images are stitched together.

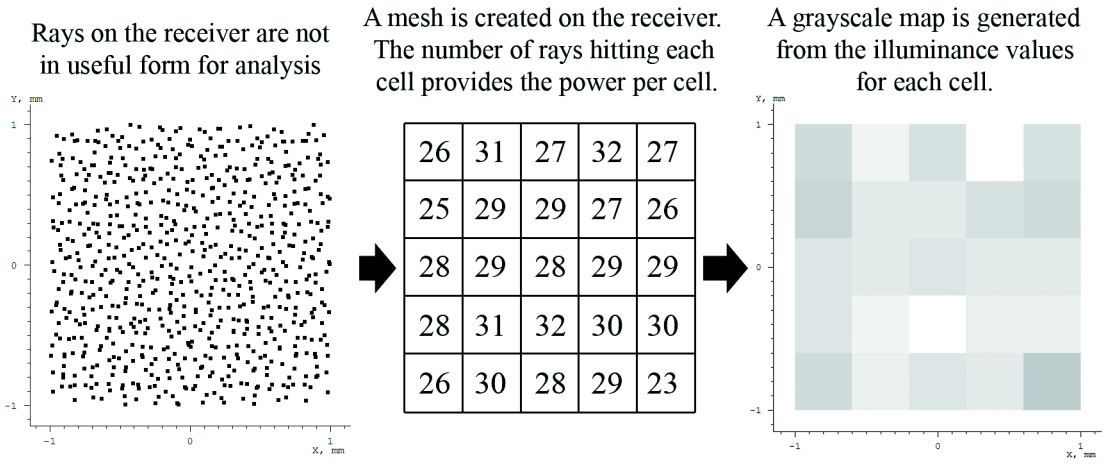

Fig. 2. Rays that hit the receivers in the simulation software are converted to an illuminance map. Because of the statistical nature of the ray generation process, the illuminance map has statistical noise that depends on the number of rays traced on the size of the cells in the mesh.

C. Performance of Tapered Rods and Compound Parabolic Concentrators

As a benchmark toward freeform lightpipes, we started by comparing the performance of two commonly used nonimaging components: tapered rods and compound parabolic concentrators (CPCs). These components were originally used for solar concentration. In our case, we use these components "backwards," so that they collimate the light instead of concentrating it. CPCs were designed to achieve performance close to maximum concentration [15,16] (or "maximum collimation" in our case). As we shall see, they do not necessarily achieve good spatial uniformity. Figure 3 shows the transmission efficiency and the spatial uniformity for four different devices: hollow tapered rods, solid tapered rods, hollow CPCs, and solid CPCs, for lengths ranging from 30 to 120 mm. The computer models of these devices are depicted in Fig. 4. All simulations were performed with the illumination software LightTools. The general configuration of the system is depicted in Fig. 5. The LED was modeled by a set of 10^6 rays issued from measurement by an imaging goniometer. The wall reflectivity used for the hollow components was 95%. Solid components were modeled with poly (methyl methacrylate) (PMMA). Material absorption has been taken into account. Scattering losses and Fresnel transmission losses on the input and output faces of the lightpipe were not included as they highly depend on the choice of coating and surface finish.

In our case, the NA of the projection optics is 0.2 and the microdisplay size is 0.56 in. (14.2 mm), which corresponds to the theoretical minimum requirements to transmit all the light from the source. We used a 28×21 mesh (which corresponds to a square bin size of 426 μm), and we traced 10^6 rays. Given symmetry considerations, the corresponding statistical noise creates an uncertainty of approximately $\pm 1.5\%$ for the relative standard deviation of illuminance. In all cases we studied, nonuniformities produced by the lightpipes had spatial frequencies below 0.3 cycles/mm on the microdisplay. If we assume that the field of view of the projected image is 20°, then according to Eq. (3) the corresponding maximum angular frequency for the viewer is approximately 0.2 cycles/deg. At this frequency, and for an average illuminance of 40 cd/m^2, the reported contrast thresholds range from 0.03 to 0.1 [14]. As a "rule of thumb", we therefore use 5% as the perception threshold for nonuniformities. This value is intended for monochrome images. In practice, color shifts can be noticeable for nonuniformities across color channels below 5%. In this case, tighter tolerances can be set or color correction can be achieved by the microdisplay, often at the expense of efficiency.

Let us first have a look at the transmission efficiency within NA = 0.2 in Fig. 3(a). Two different trends appear: hollow components have their peak efficiency at approximately 50 mm; their efficiency then decreases. As length increases, so does the number of reflections of the rays in the devices, therefore increasing the loss from metallic reflection. On the

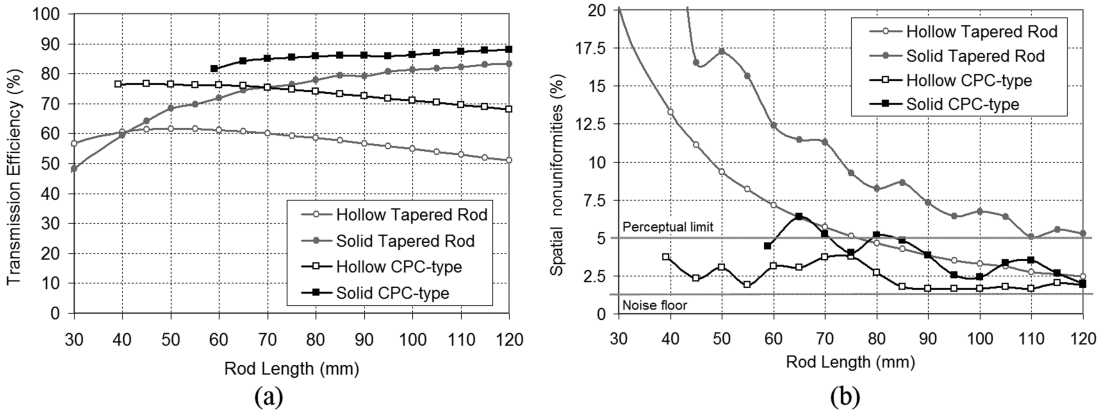

Fig. 3. (Color online) (a) Transmission efficiency within NA = 0.2 and (b) spatial nonuniformities for hollow tapered rods, solid tapered rods, hollow CPCs, and solid CPCs of various lengths. Data for CPC-like devices is not available for the shortest lengths, as the theoretical shape of the CPC requires a minimum length.

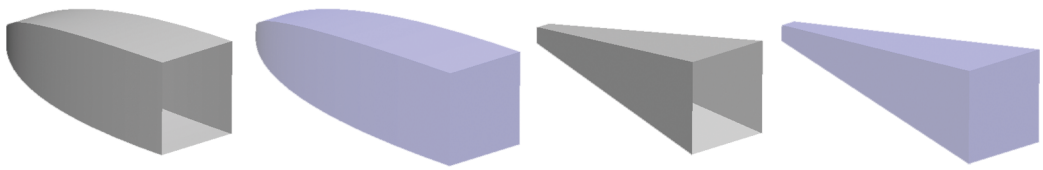

Fig. 4. (Color online) Models of the lightpipes used as benchmarks.

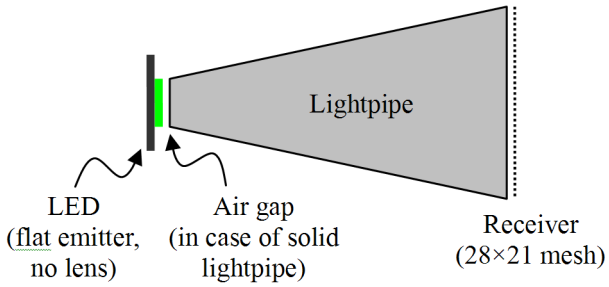

Fig. 5. System configuration. Uniformity and efficiency are measured on a flat plane at the lightpipe output. For a solid lightpipe, an air gap is left between the LED and the lightpipe input face.

other hand, the efficiency of solid components asymptotically increases with length, because total internal reflection yields up to 99.9% reflectivity [17]. Eventually material absorption creates a slow decrease in efficiency for devices longer than 160 mm. CPC-like devices achieve higher efficiency than tapered rods, which comes from the fact that they were originally designed for maximum concentration.

When looking at the spatial nonuniformities plot in Fig. 3(b), the most noticeable trend is that nonuniformities tend to decrease as length increases. Long CPCs even reach the statistical noise floor. This is a classic feature of square and rectangular lightpipes. However, it is important to keep in mind that the rule "longer lightpipe = better uniformity" is only true for specific pipe shapes and specific source emission patterns. A circular pipe will not achieve good uniformity whatever its length is, and a source with a nonrotationally symmetric emission pattern is less likely to yield good uniformity at the pipe output [16]. At a given length, hollow components tend to achieve better uniformity than solid components. In a solid device, refraction at the input face increases the angle of incidence of the rays on the lightpipe walls. Consequently, rays undergo fewer reflections in a solid device than a hollow device of the same length. Light is less "mixed" and therefore uniformity is lower.

Figure 3(b) features another interesting trend: nonuniformities tend to exhibit local minima for specific lengths. For instance, the solid CPC-like devices have local minima at 60, 75, and 100 mm. Figure 6 shows the evolution of the illumination pattern at the device output as length increases. One can take advantage of these local minima to obtain good uniformity at short length, in case compactness is a major constraint. The major drawback is the high sensitivity of these minima with length and source characteristics. Tolerancing such components will therefore be an issue. As we shall see, optimized freeform lightpipes can achieve better uniformity at shorter lengths and with less sensitivity.

CPCs typically achieve better collimation than tapered rods. Figure 7 shows the intensity distribution at nine different locations on the lightpipe output. Each "blob" represents the angular extent of the beam at a given point on the output surface. The full angular extent is ~26° in all cases. CPCs have intensity distributions that are all centered on-axis, whereas tapered rods tend to have rays with higher angles on the edge (the intensity distribution is off-centered). These edge rays do not fit within NA = 0.2, thus creating an illuminance falloff on the edges of the microdisplay. This is why tapered rods achieve poorer uniformity than CPC-like devices. By adding power to the output surface of tapered rods we can minimize this effect and greatly improve both efficiency and uniformity. This is the first step of our optimization process.

3. Optimization of Freeform Lightpipes

Any optimization process requires at least three elements: a model parameterization, a merit function, and an optimization algorithm. In our case, we used the optimization algorithm provided with LightTools, based on the damped least-squares method. Model parameterization and merit function are

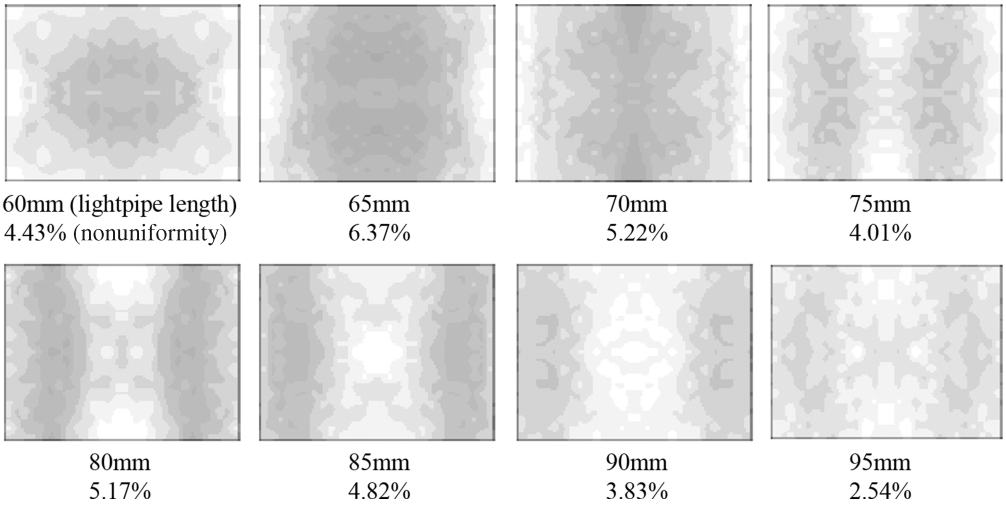

Fig. 6. Evolution of the illuminance pattern at the output of a solid CPC-like lightpipe as length increases from 60 to 95 mm. Local minima in the relative standard deviation of illuminance is observed at 60, 75, and 95 mm.

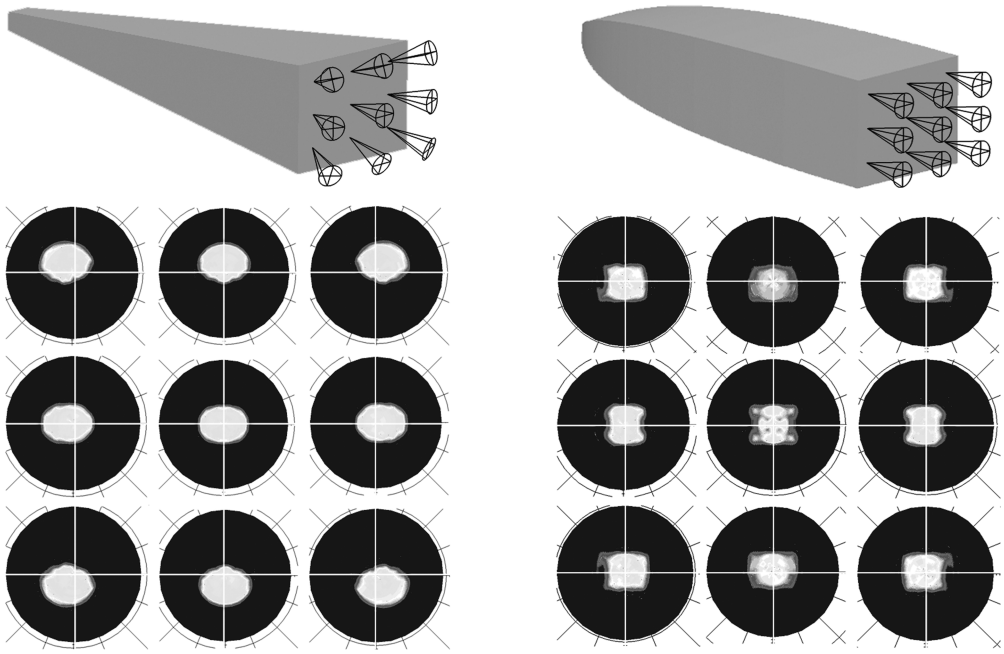

Fig. 7. Intensity distributions at nine different positions on the output face of (a) a 60 mm solid tapered rod and (b) a 60 mm solid CPC-type lightpipe. CPCs exhibit a fairly uniform intensity distribution across the device output, which is not the case for tapered rods. For this reason, the performance of tapered rods can be greatly improved by adding an output lens.

detailed in each section. In general, the optimization process went through two steps. First, a manual investigation was carried out in order to find a starting point and a relevant model parameterization. Then, local optimization was used to further improve the lightpipe performance.

A. Optimization of the Output Shape

By adding power to the output surface of a solid tapered rod, both efficiency and spatial uniformity can be improved. We studied two cases: a spherical lens and a Fresnel lens with four grooves per millimeter and a 5° draft angle (Fig. 8). It appeared that using more complex surfaces (aspheric or XY polynomial) brought little improvement. The simulation results for lengths ranging from 20 to 120 mm are shown in Fig. 9. For each length the radius of curvature was optimized in order to obtain the highest efficiency. The optimum output radius is the same for the spherical lens and the Fresnel lens. It can be seen in Fig. 10 that the optimum radius follows a linear relationship with the lightpipe length. For both the spherical and the Fresnel lenses, efficiency reaches a plateau at 90% for lengths over 60 mm. Spatial nonuniformities are more than halved, and the statistical noise floor is reached at a local minimum at 50 mm with the Fresnel lens. A "stable" uniformity is reached for both devices over 75 mm. For a 30 mm tapered rod, a ± 0.5 mm variation of the output radius produces a 0.4% efficiency drop, and an approximately 0.5% change in nonuniformity. The sensitivity of the device on the output radius is therefore low and becomes even lower for longer lightpipes.

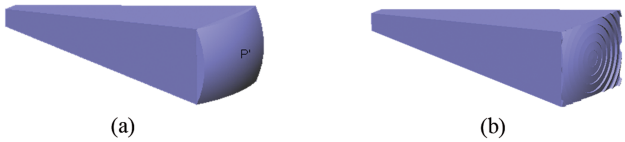

Fig. 8. (Color online) Tapered rods with (a) a spherical and (b) a Fresnel lens.

The spatial nonuniformity values for Fresnel lenses can be misleading: in fact, if we increase significantly the size of the receiver mesh and the number of rays used in the simulation, the pattern of the Fresnel grooves becomes clearly visible as shown in Fig. 11. Total internal reflection on the vertical face of the Fresnel lens grooves is responsible for this effect. One possible solution is to "smear" the ring pattern by defocusing the receiver. The result with a 1 mm defocus is also shown in Fig. 11. Defocusing, however, produces a slight edge falloff similar to what is observed with a spherical lens for small radii. Therefore, unless flatness of the lightpipe output is a requirement, Fresnel lenses are likely to be more difficult to implement. The edge falloff observed with spherical lenses stems from the departure of the output surface from flatness. For long lightpipes, optimum radii are larger and thus edge falloff is minimized. In cases where the lightpipe is used with a reflective microdisplay, the edge falloff can be compensated by adjusting the field curvature or the focus plane of the relay lens. As we can see in Fig. 11, good overall uniformity can be obtained for long lengths. However, uniformity is still an issue when compact-

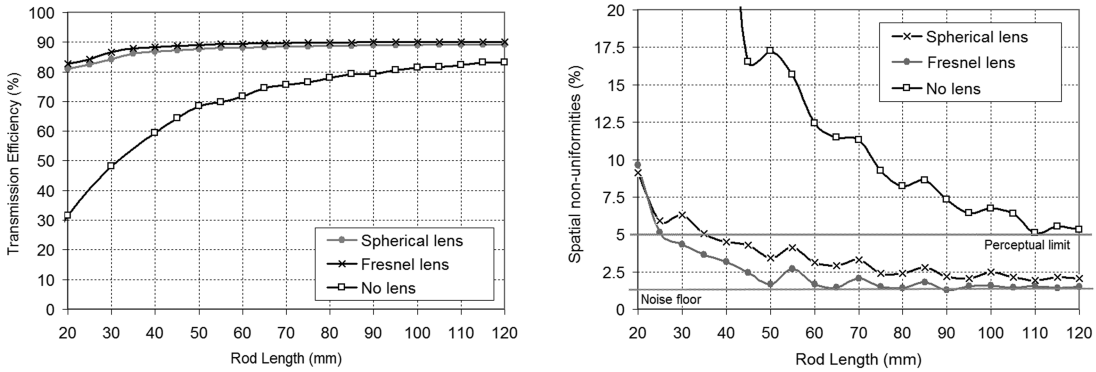

Fig. 9. Transmission efficiency within NA = 0.2 and (b) spatial nonuniformities for solid tapered rods with no lens, a spherical lens, and a Fresnel lens at the rod output.

ness is required. For a 30 mm tapered rod, a dip in illuminance is observed in the center. Optimization of the profile shape can help to improve uniformity for short lengths.

B. Optimization of the Profile Shape

Model parameterization is critical when dealing with freeform surfaces. Optimizing a nonuniform rational B-spline (NURBS) surface with dozens of nodes greatly increases the parameter space, the computation time, and the likelihood of getting stuck in a local minimum. One possible approach is to increase incrementally the complexity of the surface description. As a first step, we thus define the lightpipe profile as a rational Bezier curve with only three control points. Two of these points are fixed by the location and size of the input and output apertures. The third point position is left as a variable for optimization. Bezier curves are good candidates for surface parameterization because they require few parameters and can describe theoretical shapes such as CPCs and other classic conic concentrators. Also, control points can easily be added to the profile without perturbing an existing shape (degree elevation) [18]. To get a better understanding of the trends involved with different shapes, we placed the control point at a medium position on a tapered rod and simply moved it up and down. When the control point is moved upward, the lightpipe shape becomes parabolic (CPC-like); when the control point is moved downward it becomes hyperbolic (like the shape of a compound hyperbolic concentrator). As it can be seen in Fig. 12, a slight hyperbolic shape can provide good uniformity. The shape of the resulting component may look familiar. In fact, it is similar to the so-called dielectric totally internally reflecting concentrator (DTIRC) [19]. Traditionally, the main advantage of DTIRCs compared to CPCs is their compactness. It appears that this type of shape can also provide superior spatial uniformity.

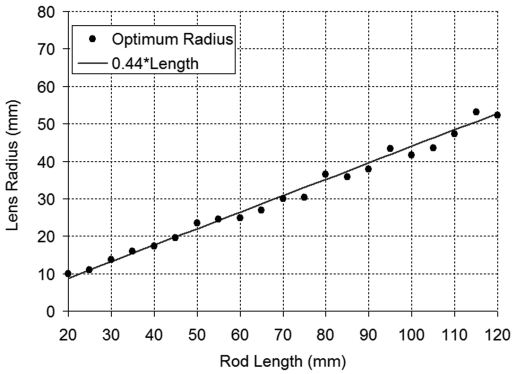

Fig. 10. Optimum output radius for rod lengths ranging from 20 to 120 mm.

The horizontal position of the control point also has a significant impact. We plotted in Fig. 13 the transmission efficiency and spatial nonuniformities when moving the horizontal and vertical positions of the control point. The gray-scale maps in Fig. 13 therefore represent the landscape of each merit function (transmission efficiency and spatial uniformity) in the parameter space of the control point location. Dark regions correspond to low transmission efficiency and low nonuniformities; whereas light regions correspond to high transmission efficiency and high nonuniformities. It appears that there is a local maximum in efficiency when the control point is placed close to the source. Two local minima exist for spatial nonuniformities when the control point is placed close to the source, either for a slight parabolic or a slight hyperbolic shape. In this case the best compromise to improve both efficiency and uniformity is therefore the hyperbolic shape. It may not be surprising that the optimum position of the control point happens to be close to the source. The walls of the lightpipe close to the source subtend a larger solid angle than the walls close to the output. In other words, a small change in the slope of the "input zone" of the lightpipe tends to have a greater effect than in the "output zone."

A final optimization of a 30 mm lightpipe (including both the weight and the position of the control point) yields 86.7% transmission efficiency within NA = 0.2 and 2.5% nonuniformities. In general,

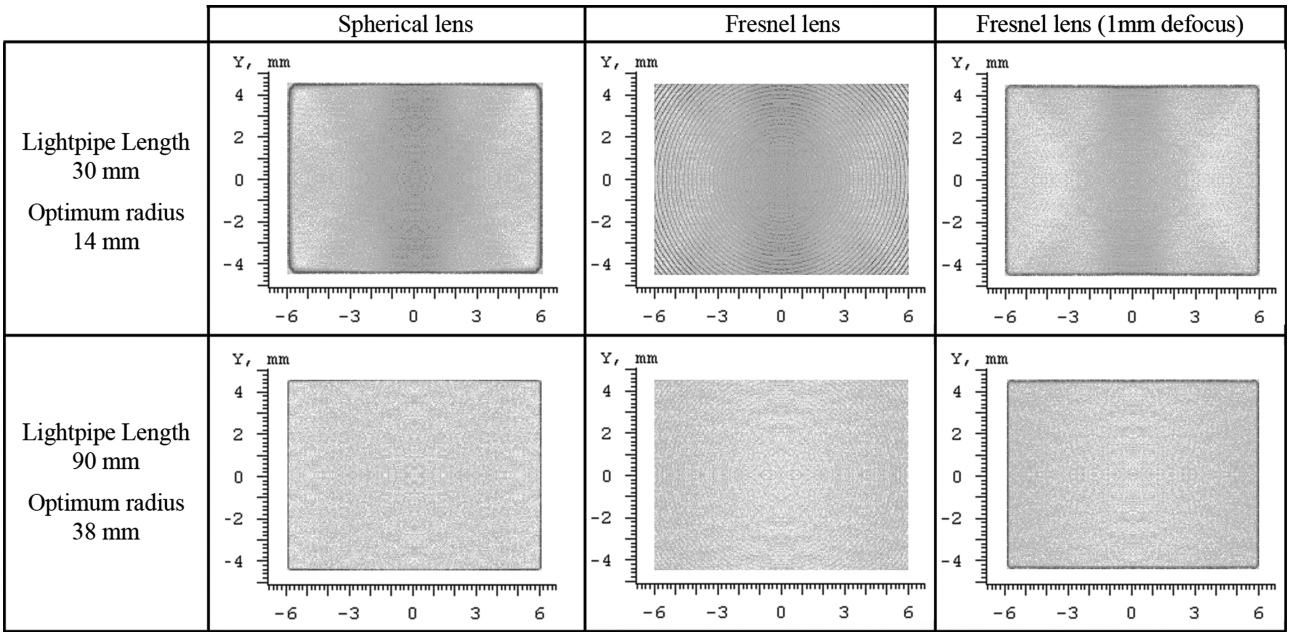

Fig. 11. Illuminance pattern at the lightpipe output with a spherical lens, a Fresnel lens, and a Fresnel lens with 1 mm defocus. The images have been generated with a larger receiver mesh (200 × 150) and 10^7 rays to resolve the pattern caused by the Fresnel lens grooves. Defocus partially gets rid of the ring pattern but produces a slight illuminance falloff on the edges. In a configuration using a reflective microdisplay, the edge falloff observed with a spherical lens can be compensated by the relay lens.

trying to improve further the device performance results in a trade-off between efficiency and uniformity. Depending on the constraints or specifications, the merit function items can be weighted to selectively improve one or the other.

C. Optimization of the Input Shape

The input face of the lightpipe ensures the collection of the light emitted by the LED. The input size of the lightpipe can be set in order to match the size of the LED die, i.e., 2.1 mm × 2.1 mm. However, this proves not to be the most efficient solution. It can be seen in Fig. 14 that light on the top LED surface "spills" beyond the actual size of the die, creating a gradual edge falloff (because of internal reflections and scattering). A large lightpipe input size would collect most of the emitted light, at the expense of a lower average luminance. In contrast, a small input size increases the average luminance but collects less flux. Additionally, the input size affects the taper angle of the lightpipe, which in turn impacts the performance of the device. Figure 15 shows the transmission efficiency within NA = 0.2 for a 30 mm lightpipe with a lens for various input widths and heights. The optimum size of the lightpipe input is in this case 2.24 mm × 2.04 mm, the larger side being oriented in the same direction as the larger side of the output face.

4. Conclusion and Open Problems

We showed that a single freeform lightpipe can efficiently couple a high-brightness LED to a microdisplay. While simple, this approach still gives flexibility to adapt the device to the various require-

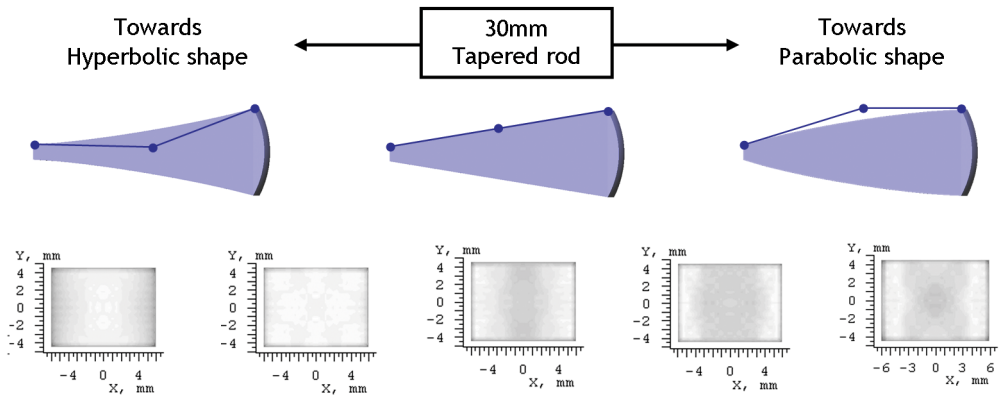

Fig. 12. (Color online) Variation of the illuminance map as the central control point is moved up and down for a 30 mm lightpipe. A slight hyperbolic shape can provide good uniformity.

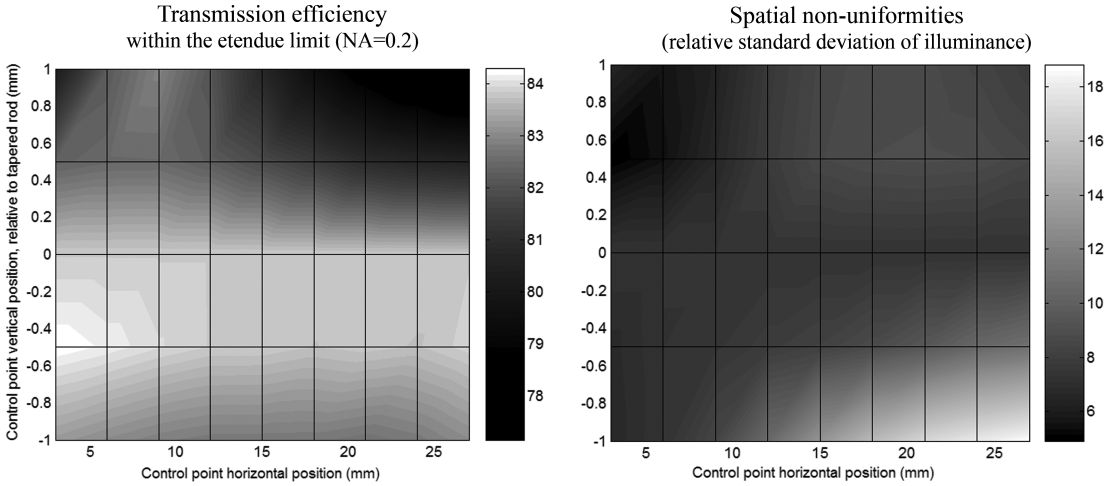

Fig. 13. Transmission efficiency (%) and spatial nonuniformities (%) for a 30 mm tapered rod with various profile shapes. The x axis corresponds to the horizontal position of the middle control point of the Bezier curve, as depicted in Fig. 12. The y axis is the vertical displacement of the control point relative to the tapered rod position. $y = 0$ corresponds to a tapered rod shape; $y > 0$ corresponds to a parabolic shape; $y < 0$ corresponds to a hyperbolic shape.

ments of the design problem. Adding a lens at the output surface greatly improves transmission efficiency compared to classical nonimaging concentrators, thus enabling compact and highly efficient components. A hyperbolic profile was found to overcome the spatial nonuniformities resulting from the aspect ratio mismatch between the source and the microdisplay.

Optimization for nonsequential systems remains a challenge as the choice of the best model parameterization and optimization algorithm highly depends on the design constraints and how they are translated into metrics. Spatial uniformity is a good example of this type of problem. In this study we used the relative standard deviation as a metrics for uniformity. Figure 6 shows an example of two different illuminance maps that have a similar relative standard deviation. However, it gives us no information on the nonuniformity pattern, for instance whether variations are smooth or not. The illuminance map itself could be used as a merit function (all cells in the bin are assigned a target illuminance); however it becomes much more difficult to evaluate the performance of an illumination map compared to another. Therefore, more sophisticated application-specific metrics related to human visual perception need to be defined to drive the optimization process.

Tolerancing has been briefly mentioned in this study. This critical aspect of the design should be integrated in the optimization process as a way to only select robust solutions. One possible solution is to perform parameter sensitivity analyses. The merit function landscape gives an idea of the sensitivity of the device. However, it is still difficult to tolerance freeform surfaces and the model parameterization used for optimization may not be meaningful for mechanical tolerances. A classical example of the phenomenon is the tolerances of aspheric

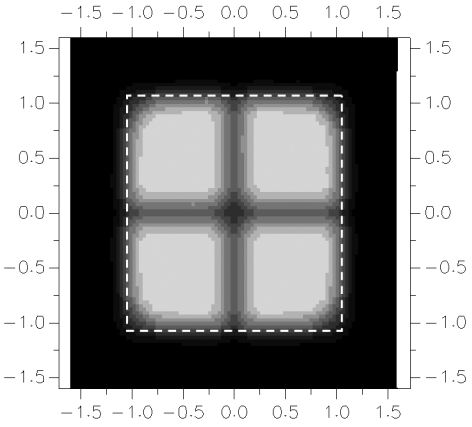

Fig. 14. Illuminance map at the top surface of the Osram Ostar LED. The four-die structure of the emissive area is clearly visible. Light falloff on the edges is gradual, and occurs beyond the actual physical size of the die indicated by the white dashed line (2.1 mm × 2.1 mm).

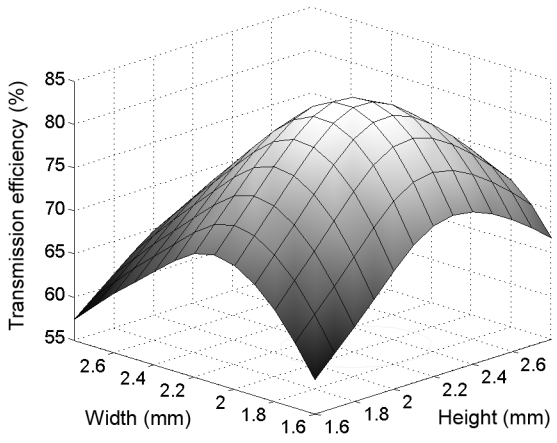

Fig. 15. Transmission efficiency (%) within NA = 0.2 for a 30 mm tapered rod with a lens for varying lightpipe input height and width. Maximum transmission efficiency is obtained in this case for a 2.24 mm × 2.04 mm input size.

surfaces. A small change in a single aspheric coefficient can result in a tremendous decrease in performance [20]. However, this worst case scenario does not happen in practice: deviations from the prescribed shape owing to manufacturing processes typically affect all aspheric coefficients at the same time, thus averaging the effects of the variations of each individual coefficient. In this case, slope tolerances prove to be more relevant. Applying tolerances to aspheric coefficients is therefore not meaningful, thus the importance of choosing a relevant model parameterization for freeform surfaces. In this area, little work has been done so far.

Appendix

In this paper, we use the relative standard deviation of the illuminance distribution as a metric for spatial nonuniformity. The ability for the eye to perceive nonuniformities depends on their spatial frequency. Contrast sensitivity functions (CSFs) reported in the literature measure the minimum contrast required for a periodic pattern at a given spatial frequency to be perceived. The contrast C is usually defined using the "Michelson definition"

$$C(\nu) = \frac{E_{\max} - E_{\min}}{E_{\max} + E_{\min}}, \tag{A1}$$

where ν is the spatial frequency of the pattern, and $E_{\max}$ and $E_{\min}$ are the maximum and minimum illuminance of the pattern. This contrast definition for a periodic pattern can be linked to the RSD of the illuminance distribution. If the periodic pattern used to measure the CSF is a sinusoidal function, the illuminance at a given point x is given by

$$E(x) = \frac{1}{2}[E_{\max}(1 + \sin x) + E_{\min}(1 - \sin x)]. \tag{A2}$$

The mean $\bar{E}$ and the variance σ^2 of the illuminance distribution can then easily be calculated:

$$\bar{E} = \frac{1}{2\pi}\int_0^{2\pi} E(x)\mathrm{d}x = \frac{E_{\max} + E_{\min}}{2}, \tag{A3}$$

$$\sigma = \sqrt{\frac{1}{2\pi}\int_0^{2\pi} (E(x) - \bar{E})^2 \mathrm{d}x} = \frac{E_{\max} - E_{\min}}{2\sqrt{2}}. \tag{A4}$$

Therefore, we obtain

$$\mathrm{RSD}(\nu) = \frac{\sigma}{\bar{E}} = \frac{C(\nu)}{\sqrt{2}}. \tag{A5}$$

The authors acknowledge Optical Research Associates for the educational license of LightTools. This project was motivated by related research with VDC Display Systems. It was initiated under support by the Office of Naval Research (ONR) Virtual Technologies and Environments (VIRTE) program and was completed under a Fellowship to Florian Fournier from Optical Research Associates. We thank William Cassarly for insightful conversations about illumination.

References

1. F. Fournier and J. Rolland, "Design methodology for high brightness projectors," J. Display Technol. **4**, 86–91 (2008).
2. J. M. Teijido, F. Ludley, O. Ripoll, M. Ueda, Y. Oshima, T. Yoshida, K. Toyota, K. Yamamoto, T. Nagara, Y. Kato, A. Wajiki, and S. Umeya, "73.2: compact three panel LED projector engine for portable applications," SID Int. Symp. Digest Tech. Papers (2006).
3. E. Geissler, "Meeting the challenges of developing LED-based projection displays," presented at the Photonics in Multimedia Conference, Strasbourg, France, 3–7 April 2006.
4. M. H. Keuper, G. Harbers, and S. Paolini, "26.1: RGB LED illuminator for pocket-sized projectors," SID Int. Symp. Digest Tech. Papers **35**, 943–945 (2004).
5. M. R. Krames, O. B. Shchekin, R. Mueller-Mach, G. O. Mueller, Z. Ling, G. Harbers, and M. G. Craford, "Status and future of high-power light-emitting diodes for solid-state lighting," J. Display Technol. **3**, 160–175 (2007).
6. R. Karlicek, "Photonic Lattice LEDs are new class of light-emitting device," LEDs Magazine, 20–23 (August 2007).
7. E. H. Stupp and M. S. Brennesholtz, *Projection displays*, Wiley SID series in display technology (Wiley, 1999), pp. xviii, 418.
8. M. P. Krijn, B. A. Salters, and O. H. Willemsen, "LED-based mini-projectors," presented at the Photonics in Multimedia Conference, Strasbourg, France, 3–7 April 2006.
9. H. Murat, D. Cuypers, and H. De Smet, "Design of new collection systems for multi LED light engines," presented at the Photonics in Multimedia Conference, Strasbourg, France, 3–7 April 2006.
10. S. Ohuchi, T. Miyoshi, T. Imahase, T. Nakashima, and K. Shikita, "Ultra portable LCOS projector with high-performance optical system," IEEE Trans. Cons. Electron. **48**, 388–393 (2002).
11. J.-W. Pan, C.-M. Wang, W.-S. Sun, and J.-Y. Chang, "Portable digital micromirror device projector using a prism," Appl. Opt. **46**, 5097–5102 (2007).
12. I. Moreno, "Spatial distribution of LED radiation," Proc. SPIE **6342**, 634216 (2006).
13. W. Falicoff, J. Chaves, and B. Parkyn, "PC-LED luminance enhancement due to phosphor scattering," presented at the Nonimaging Optics and Efficient Illumination Systems II Conference, San Diego, Calif. 31 July–1 August 2005.
14. E. Peli, "Contrast sensitivity function and image discrimination," J. Opt. Soc. Am. A **18**, 283–293 (2001).
15. R. Winston, J. C. Minano, P. Benitez, and W. T. Welford, *Nonimaging Optics* (Elsevier Academic, 2005), pp. xi, 497.
16. W. J. Cassarly, "Nonimaging optics: concentration and illumination," in *Handbook of Optics*, 2nd ed. (McGraw-Hill, 1995).
17. L. A. Whitehead and M. A. Mossman, "Off the beaten path with total internal reflection," Proc. SPIE **6342**, 63420U (2006).
18. L. A. Piegl and W. Tiller, *The NURBS Book*, 2nd ed. (Springer, 1997), pp. xiv, 646.
19. X. Ning, R. Winston, and J. O'Gallagher, "Dielectric totally internally reflecting concentrators," Appl. Opt. **26**, 300–305 (1987).
20. J. Rodgers, "Slope error tolerances for optical surfaces," presented at the SPIE Optifab Conference, Rochester, N.Y., 15–17 May 2007.