Ray mapping with surface information for freeform illumination design

CALEB GANNON AND RONGGUANG LIANG*
College of Optical Sciences, University of Arizona, 1630 East University Boulevard, Tucson, Arizona 85721, USA
*riang@optics.arizona.edu

Abstract: A novel approach to incorporate surface information into the ray mapping method is proposed. This method calculates irradiance at the physical optical surface and target plane instead of the usually flat or hemispherical dummy surface, resulting in a mapping relationship which reflects the true geometry of the system. The robustness of the method is demonstrated in an extreme example (60° off axis) where the uniformity is as high as 82%.

© 2017 Optical Society of America

OCIS codes: (080.4298) Nonimaging optics; (080.1753) Computation methods; (220.2945) Illumination design; (080.4225) Nonspherical lens design.

References and links


1. Introduction

The basic concept of freeform illumination design is redirecting light from one region to another as efficiently as possible. Usually this is done through ray targeting, where selected rays from the source are directed towards calculated points in the target plane based on their relative energy and the desired output energy distribution. Constructing a surface to solve this complex inverse problem is a challenging task that usually involves the calculation of two parameters: a mapping relationship between the source and target energy distributions, and a surface capable of enforcing the desired mapping. Simultaneous solutions of these problems using complex differential equations have demonstrated promising results in select cases [1–
However, these methods are not easily generalizable to unique circumstances of geometry or desired energy distributions. One way to overcome this difficulty is to allow the problems to be solved in separate steps by first calculating the mapping relationship based on the source and target distributions and then using that relationship to construct a discrete surface from a normal field calculated by Snell’s law [4–13].

Traditionally the energy distribution of the source used for the mapping relationship is sampled on a dummy surface in front of the lens [4–13]. Although some of the methods listed determine their sampling in angular space [5,8,10,12,13], one can always construct a dummy surface and select points such that the required angular sampling is achieved. For example, one could use a hemispherical surface with points equally distributed along the azimuthal or polar angles. In this way we believe that all approaches of this kind can be equivalently addressed by using dummy surfaces.

In simple cases, the method of sampling source rays at a dummy surface is effective because the resulting real surface and dummy surface do not deviate much from each other. In more extreme cases, however, this method breaks down as the rays sampled at the dummy surface become less reliable as an approximation to the actual ray behavior. This effect can be seen in Fig. 1.

![Fig. 1. Rays from the source hit the real surface in drastically different locations than a sample plane near the input can predict. The resulting mapping from the sampled plane poorly reflects the final spatial location of those rays.](image)

The resulting input energy distribution from these calculations deviates from what is actually happening, causing performance to deteriorate. The severity of this effect can be seen by comparing the irradiance distribution on the dummy surface and the final surface in Fig. 2. Not only can the distribution within the grid be skewed, but the physical extent of the distribution can change as well.

![Fig. 2. Irradiance distributions for an off-axis design: (a) sampled on a hemispherical dummy surface and (b) sampled on the final surface.](image)

Secondly, when surface information is withheld from the mapping calculation, the relative angular relationship between the surface profile and target plane can only be approximated at
best while usually being neglected altogether. Traditionally the input and output distributions are calculated in isolation [4–13], leaving no ability to incorporate the relative geometries between the two surfaces. A mapping relationship calculated in this manner will be invariant of major changes in the geometry of the system, for example moving the target plane laterally away from the source. The resulting compounded approximations lead to a mapping relationship which introduces unnecessary curl into the resulting vector field as shown in Fig. 3.

![Fig. 3. Z component of the curl in the vector field between the physical surface and the target plane for (a) the on axis case and (b) 60° off axis. Notice the curl in the off axis case has increased by more than an order of magnitude.](image)

The requirement for a surface to be both continuous and smooth is called the integrability condition. Essentially, it means that any closed path around the surface needs to end up in the starting location. In order for this to be accomplished, it has been shown that

$$(\nabla \times \mathbf{N}) \times \mathbf{N} = 0$$  \hspace{1cm} (1)$$

must be satisfied, where \(\mathbf{N}\) denotes the vector field of the surface normals [13].

Since the \(\mathbf{N} = 0\) case is not an interesting solution, in order to make a smooth surface we need to reduce \(\nabla \times \mathbf{N}\) as close to zero as possible. By looking at Eq. (3), we can see that the surface normals are directly related to the output vector distribution. In fact, it has been shown that reducing the curl in the outgoing vector field directly impacts the integrability of the surface [14]. To improve both the irradiance distribution and resulting integrability of the vector field, we propose a method which incorporates the actual surface information into the ray mapping calculation. In doing so, we have created a robust and novel method that poses as a significant improvement over previous methods.

2. Method

The approach is performed in two steps. First, a mapping relationship is calculated using the Optimum Transport ray mapping method described in [11]. Irradiance sampled on a dummy surface above a Lambertian point source is used as the input, while the output is a uniform rectangular distribution.

Using this mapping relationship, a discrete input vector field \(\mathbf{r}_{ij} = \{x_{ij}, y_{ij}, z_{ij}\}\) is created according to the calculated x and y coordinates in the mapping relationship and their corresponding z locations interpolated on the dummy surface. The surface points, denoted by \(\mathbf{P}_{i,j}\), are a distance \(r_{ij}\) radially along the input vectors such that \(\mathbf{P}_{i,j} = r_{ij}\mathbf{r}_{i,j}\). An output vector field \(\mathbf{O}_{ij}\) can be established according to the differences between the surface and target points \(\mathbf{t}_{i,j}\).
Finally, using the vector form of Snell’s law
\[ \hat{N}_{ij} = \frac{\hat{O}_{ij} - n_{\text{low}} \hat{r}_{ij}}{\| \hat{O}_{ij} - n_{\text{low}} \hat{r}_{ij} \|}, \] (3)
the surface normal field can be evaluated using \( \hat{O}_{ij} \) and \( \hat{r}_{ij} \) scaled by the index of refraction of the lens. Unfortunately, as the magnitude of the input vector field \( r_{ij} \) is unknown, these parameters cannot be solved for directly. To overcome this problem, we used the iterative surface construction method proposed by Zexin et al [11].

Then the resulting freeform surface is used as the dummy surface for the next iteration of the same process. The irradiance distribution on the freeform surface can be found according to the inverse square law, with two cosine factors arising due to the projection between the surface and the source.

\[ E = \frac{LA}{r^2} \cos \theta_1 \cos \theta_2. \] (4)

In general \( L \) represents the source radiance and \( A \) is the source area. Since the total power incident on the surface needs to be normalized for the optimum transport calculation, these scalar values can usually be ignored. The factor \( \cos \theta \) can be computed according to the dot product between the radial unit vector, \( \hat{r}_{ij} \), and the normal on the target or surface such that
\[ \cos \theta_1 = \hat{r}_{ij} \cdot \hat{N}_{\text{surface}_{ij}}, \quad \cos \theta_2 = \hat{r}_{ij} \cdot \hat{N}_{\text{source}}. \] (5)

Figure 4 shows this in more detail below.

Fig. 4. Vector definitions and surface locations used to calculate the output vectors and the irradiance distribution on the surface. Although \( \hat{r}_{ij} \) is a unit vector, it has been extended to the surface point for greater clarity on its direction.

Inserting the values for \( \cos \theta_1 \) and \( \cos \theta_2 \) from Eq. (5) into Eq. (4), the resulting irradiance distribution on the surface becomes
\[ E_{ij} = \frac{LA}{r_{ij}^2} (\hat{r}_{ij} \cdot \hat{N}_{\text{source}})(\hat{r}_{ij} \cdot \hat{N}_{\text{surface}_{ij}}) \] (6)
Using this irradiance calculated at the surface of the lens as the input for the optimal transport calculation guarantees a unique mapping relationship which is specifically tailored to the geometry of the system. Additionally, we shift the boundary conditions of the Monge-Ampere calculation over to the physical location of the surface. This way, the mapping calculation can take into account not only the spatial information of the surface, but also the offset between the input and output planes. As one can see in Fig. 5, the resulting mapping relationship is very different than the original.

Fig. 5. (a) A typical mapping relationship between a square sample of a Lambertian source and a uniform square target plane, shown here as the deformation of a unit grid. (b) The new mapping relationship calculated between the energy distribution on the actual surface and the target plane for an off-axis case, shown as the deformation of a unit grid.

After calculating the mapping relationship from the new distribution, a surface is constructed in the same manner as before. This cycle of irradiance sampling and surface construction can be repeated as many times as is necessary, but in every case we studied it converged after one iteration.

3. Design example

To demonstrate the robust nature of the proposed method, we examined the performance in an extreme case. Starting with a rectangular surface inscribing a ±55° collection angle we measured the performance as a function of the angle off axis, keeping the source pointed directly away from the target plane. A Lambertian point source was used for all of the simulations. A diagram of the test geometry is shown in Fig. 6.

Fig. 6. Test geometry for the 60° off-axis case shown in LightTools Simulation software. The surface was placed 10mm above the source and pointed directly away from the target plane. The angular offset was measured as the angle between the surface normal at the center of the target plane and the source.

Although the method is fully capable of handling arbitrarily shaped apertures, our examples were performed with a rectangular aperture. This is due to the fact that at extreme angles, error due to skew mismatch between circular and rectangular geometries is
intensified. In order to isolate errors arising from the mapping method it seemed best to reduce inherent geometrical problems such as skew mismatch as much as possible. While the solution of that problem is imperative, as we discuss in the conclusion this method is best suited for scalar problems such as distortion in the irradiance patterns. Vector distortion will be addressed in future research.

On axis, where the final and dummy surfaces don’t deviate much from a flat plane the performance of both systems is quite good: 95% uniformity for the original method and 98% for the new method, with resulting irradiance patterns shown in Fig. 7.

On axis, where the final and dummy surfaces don’t deviate much from a flat plane the performance of both systems is quite good: 95% uniformity for the original method and 98% for the new method, with resulting irradiance patterns shown in Fig. 7.

![Irradiance patterns on target surface](image)

Fig. 7. Irradiance patterns on the target surface for (a) traditional method and (b) the proposed method for the on-axis case. The corresponding uniformities are 95% and 98% respectively. The irradiance uniformity was calculated as 100%(1-RMSdeviation/mean) on the entire 101x101 grid with a 3 pixel smoothing kernel to reduce statistical error from the raytracing.

The performance of each method was evaluated every 10° off axis by moving the target laterally away from the source and is plotted below in Fig. 8 with the resulting irradiance patterns shown in Fig. 9. The results demonstrate the capabilities of the proposed method in handling both extreme and traditional geometries.

![Performance measured in terms of uniformity](image)

Fig. 8. Performance measured in terms of the uniformity within the target region as a function of angular offset for both the traditional method and the proposed method.
Fig. 9. Irradiance patterns at 60° off axis for (a) the traditional method (b) the proposed method. The corresponding uniformities within the 200mmx200mm target are 56% and 82% respectively. The Irradiance uniformity was calculated as \(100\%(1\text{RMS}\text{deviation/mean})\) on the entire 101x101 grid with a 3 pixel smoothing kernel to reduce statistical error from the raytracing.

4. Conclusion

The proposed method is shown to be robust and improve the performance of illumination systems, especially in extreme cases such as the 60° off axis design. In particular, it removes distortions in the irradiance pattern which would normally be introduced by such difficult geometries. The main limitation of the method is that the Optimum Transport calculation is performed on scalar quantities (energy distribution in space), whereas light rays are actually vectors. By dealing with scalar quantities we can only ever approximate the true behavior of light rays at an optical surface. Although this approximation provides an improvement over traditional methods, in Fig. 10 we can see there is still some residual curl in the calculated ray-field. This curl introduces surface error on the lens, which causes the observed warping in the final irradiance pattern.

Fig. 10. Residual Z component of the curl in the vector field between the physical surface and the target plane using the proposed method.

Another limitation is that ray mapping requires a point source so that a one-to-one mapping relationship can be achieved. When a real source is introduced, the ray mapping relationship begins to lose validity. This is because instead of one ray corresponding to one point on the surface, each point on the surface has to direct multiple rays. This problem is illustrated below in Fig. 11.
If the source is relatively small compared to the lens surface, designs using a mapping relationship can still be made by approximating the source as a point. Although this works relatively well on-axis, as the target is moved off-axis the projected size of the source on the target can be quite large and lead to blurring in the final irradiance pattern. In Fig. 12 we see this effect, where the drop off region in the shifted direction is quite large even though the source is only 1mmx1mm.

The next steps in this design will be to augment the method to work with vector fields so that the integrability condition can be satisfied and to establish a mapping relationship and surface construction which incorporate the physical extent of the source. By doing so we hope to overcome the two major remaining limitations of the ray mapping method.

Fig. 12. Irradiance pattern produced when a 1mmx1mm LED is used instead of a point source with the proposed method at 60° off axis. The uniformity is slightly worse at 80%, largely due to the drop off regions at the edges of the pattern.

Funding

Research reported in this publication was supported by the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health under award number UH2EB022623.
Acknowledgments

The authors would like to thank Zexin Feng and Brittany Froese [11] for help with the Optimum Transport code, Rengmao Wu for posing the challenging design problem of off-axis ray mapping and John Koshel for providing helpful discussions.