# An Error Bound for Decoupled Visibility with Application to Relighting

Karsten Schwenk[1], Johannes Behr[1] and Dieter W. Fellner[1,2,3]

[1]Fraunhofer IGD, Germany    [2]TU Darmstadt, Germany
[3]Institut für ComputerGraphik & Wissensvisualisierung, TU Graz, Austria

## Abstract

*Monte Carlo estimation of direct lighting is often dominated by visibility queries. If an error is tolerable, the calculations can be sped up by using a simple scalar occlusion factor per light source to attenuate radiance, thus decoupling the expensive estimation of visibility from the comparatively cheap sampling of unshadowed radiance and BRDF. In this paper we analyze the error associated with this approximation and derive an upper bound. We demonstrate in a simple relighting application how our result can be used to reduce noise by introducing a controlled error if a reliable estimate of the visibility is already available.*

Categories and Subject Descriptors (according to ACM CCS):   Computer Graphics [I.3.7]: Three-Dimensional Graphics and Realism—

## 1. Introduction

Accurately computing direct illumination from an area light source requires the evaluation of an integral over the visible part of the light. In practice Monte Carlo integration by means of stochastic ray tracing is often used to evaluate this integral. Determining the visibility of samples on the light source by tracing shadow rays is usually a major part of the overall computation time. For some applications, especially in interactive computer graphics, the overall look of ray-traced soft shadows is desirable, but a completely accurate solution is not necessary. A well-behaved error may be tolerable if rendering time can be decreased by introducing it.

In this paper we study the consequences of factoring the visibility function out of the local reflectance integral and replacing it by a scalar visibility factor. Once factored out, this visibility factor can be estimated independently of the remaining terms in the integrand. By decoupling the relatively expensive estimation of visibility from the relatively cheap sampling of radiance and BRDF an application gains more freedom to optimize. For example one estimate can be (partially) reused if only the other changes, or samples can be selectively invested in the estimator with the highest variance to reduce noise faster.

Our main contribution is the analysis of the error associated with this approximation and the derivation of an upper bound for the error. In addition we describe how our result can be used in practice to enhance a simple relighting application. We believe that especially upcoming GPU-based progressive renderers using stochastic ray tracing with focus on interactivity can benefit from our results.

## 2. Related Work

Methods for reducing the number and costs of visibility tests have received much attention by the computer graphics community.

Many approaches use a preprocessing step to build acceleration structures. Lehtinen et al. [LLA06] reconstruct the visibility function from a list of silhouette edges that is quickly fetched from a BSP tree. Xie et al. [XTP07] accelerate soft shadows by tracing rays against a multilayer transparent shadow map instead of the scene geometry. Clarberg and Akenine-Möller [CAM08] exploit coherence in the visibility function by using a visibility cache to construct a control variate for the local reflectance integral. Soler and Sillion [SS96] use the visibility error in form factors as a subdivision criterion for hierarchical radiosity. Walter et al. [WFA*05] analyze and bound the error associated with a cluster in their *Lightcuts* approach, but they assume each light is fully visible.

In addition to these methods there are techniques for adaptive and importance sampling [CPF10]. Ghosh and Heidrich [GH06] obtain an initial energy estimate and detect partially occluded regions in a first pass, and use this information in a second pass to explore penumbrae using Metropolis sampling. Donikian et al. [DWB*06] iteratively construct an optimized PDF per pixel over multiple passes.

In real-time rendering approximations based on ambient occlusion, shadow maps and PRT are most common, we refer to [HLHS03] and [Ram09] for thorough surveys on these techniques. A notable exception is Nichols et al. [NPW10], who resolve visibility by ray marching a screen-space voxelization of the scene.

With our work we hope to further bridge the gap between fast but very approximate real-time methods and the accurate but rather involved approaches used in offline rendering.

## 3. An Error Bound for Decoupled Visibility

We begin our derivation with the local reflectance integral in surface form:

$$L_o(x', x'') = \int_A L(x, x') f_r(x, x', x'') G(x, x') V(x, x') \mathrm{d}x. \quad (1)$$

$L_o$ is the radiance leaving $x'$ towards $x''$, $A$ is the surface of the light source, $L$ is the radiance leaving a point $x$ on the light source towards $x'$, $f_r$ is the BRDF, and $G$ is the geometry factor $G(x, x') = \cos\theta \cos\theta' r^{-2}$, where $r$ is the distance between $x$ and $x'$; $\theta$ and $\theta'$ are the angles between the line segment connecting $x$ and $x'$ and the surface normals at those points.

By fixing $x'$ and $x''$ and setting $LB(x) = L(x) f_r(x) G(x)$ we can write this compactly as

$$L_o = \int_A LB(x) V(x) \mathrm{d}x. \quad (2)$$

The idea is to sample $LB$ independently from $V$ and use a scalar visibility factor, $\widehat{V} \in [0, 1]$, as an approximation:

$$L_o = \int_A LB(x) V(x) \mathrm{d}x \approx \int_A LB(x) \mathrm{d}x \cdot \widehat{V}. \quad (3)$$

We choose $\widehat{V} = E[V]$ to be the *expected* visibility of the samples drawn for $LB(x)$. (This is not necessarily the average visibility of the light source.) This choice guarantees that our derivation is correct in the presence of importance sampling, as will become clear in the following section.

### 3.1. Derivation

The error associated with the approximation in Eq. 3 is

$$\mathrm{Err}(LB, V) = \int_A LB(x) V(x) \mathrm{d}x - \int_A LB(x) \mathrm{d}x \cdot \widehat{V}. \quad (4)$$

Given the problem in this form it is relatively straightforward to bound the error using probability theory. By introducing a valid probability density function $p(x)$ for sampling

$x$ we can relate the integrals to the expected values of random variables and their covariance:

$$\begin{aligned}\mathrm{Err}(LB, V) &= \int_A \frac{LB(x) V(x)}{p(x)} p(x) \mathrm{d}x \\ &\quad - \int_A \frac{LB(x)}{p(x)} p(x) \mathrm{d}x \cdot \int_A V(x) p(x) \mathrm{d}x \\ &= E\left[\frac{LBV}{p}\right] - E\left[\frac{LB}{p}\right] E[V] \\ &= \mathrm{Cov}\left(\frac{LB}{p}, V\right). \end{aligned} \quad (5)$$

If the correlation coefficient is defined, i.e. the variances of both variables are finite and not zero, we have

$$\mathrm{Cov}\left(\frac{LB}{p}, V\right) = \mathrm{Corr}\left(\frac{LB}{p}, V\right) \sqrt{\mathrm{Var}\left(\frac{LB}{p}\right) \mathrm{Var}(V)}, \quad (6)$$

and since the correlation coefficient always is in $[-1, 1]$ we can bound the error as

$$|\mathrm{Err}(LB, V)| \le \sqrt{\mathrm{Var}\left(\frac{LB}{p}\right) \mathrm{Var}(V)}. \quad (7)$$

A special case worth noting is if $p(x) = 1/|A|$, i.e. if uniform sampling is used. Then the error bound is

$$|\mathrm{Err}(LB, V)| \le |A| \sqrt{\mathrm{Var}(LB) \mathrm{Var}(V)}. \quad (8)$$

### 3.2. Interpretation

We will use $\langle \cdot \rangle$ to refer to estimated quantities. For the sake of brevity we will refer to samples used to estimate $\langle LBV \rangle \approx \int LB(x) V(x) \mathrm{d}x$, $\langle LB \rangle \approx \int LB(x) \mathrm{d}x$, and $\langle V \rangle \approx \widehat{V}$ as *LBV*-samples, *LB*-samples, and *V*-samples, respectively.

Eq. 5 is mainly interesting from a theoretical point of view. It expresses in mathematical form what one would intuitively expect: In a scenario where *LB*-samples are strongly correlated to *V*-samples the error introduced by decoupling *V*-samples is large. An example of such a situation is a light source with a very bright spot and an occluder that only occludes this spot. The error is smaller if *LB*-samples and *V*-samples are weakly correlated. An example is a moderately sized diffuse white area light source (uniformly sampled) shining on a diffuse surface. Due to the low frequency content of *LB* it does not matter much exactly which parts of the light are occluded.

Eq. 7 is the main result of this paper, as it can be used to determine *a priori* whether the error will be tolerable, if reliable estimates of $Var(LB/p)$ and $Var(V)$ are available. Note that we are only required to sample *LB* and *V* with the same PDF for Eq. 7 to hold, not with the same samples. This means we can decouple the relatively cheap evaluation of *LB* from the relatively expensive evaluation of *V* and reuse variance estimates if only one function is changing. One application that immediately comes into mind is relighting and

material editing. The following section describes how such a system may take advantage of Eq. 7.

## 4. Application to Relighting

The basic idea of our relighting approach is to treat noise for error – more precisely, to reduce the variance in the estimator for direct lighting (and thus to reduce noise in the image) by introducing bias. In an environment where the lighting or BRDFs often change, but visibility stays constant, a controlled error may be preferable to an unbiased but noisy result.

Assume we already have a converged visibility solution and have estimates of $Var(V)$ and $E[V] \approx \langle V \rangle$ per pixel, e.g. by using a method for incremental variance calculation. We can then change the $LB$ term for any pixel, i.e. change the light's radiance distribution or the BRDF of the surface point, quickly collect some relatively cheap samples of $LB$, estimate $Var(LB/p)$ and apply Eq. 7 to estimate the error we introduce by simply reusing $\langle V \rangle$ as a scalar visibility factor. The noise in the image can be reduced by blending an unbiased estimator with the approximation of Eq. 3:

$$\langle L_o \rangle = (1-t)\langle LBV \rangle + t\langle LB \rangle \langle V \rangle. \qquad (9)$$

A reliable (already converged) $\langle V \rangle$ will in general reduce the variance in the second term, because the variance in $V$ disappears. In addition, $\langle LB \rangle$ will usually converge faster than $\langle LBV \rangle$, simply due to the fact that one can evaluate more samples in the same time. The estimator given in Eq. 9 is biased (it is not even consistent), but if we use Eq. 7 to choose $t$, we can bound the expected error.

Let $\langle \text{Err} \rangle$ be the error estimated by the procedure described above, i.e. by applying Eq. 7 with estimates for $Var(LB/p)$ and $Var(V)$. We define the estimated relative error as

$$\langle \text{Err}_{\text{rel}} \rangle = \frac{\langle \text{Err} \rangle}{\langle LB \rangle}. \qquad (10)$$

The blending factor $t$ in Eq. 9 is set to

$$t = \text{clamp}\left( \frac{\text{Err}_{\text{relThres}}}{\langle \text{Err}_{\text{rel}} \rangle}, 0, 1-\varepsilon \right), \qquad (11)$$

where $\text{Err}_{\text{relThres}}$ is the maximum relative error we want in the final estimator, and $\varepsilon$ is a small factor to accommodate the fact that the relative error is only estimated and to guarantee that the unbiased part of the estimator contributes at least a small fraction. We use the color channel that gives the highest relative error to determine $t$. In regions with very low estimated luminance levels or very low estimated variance, the denominator in Eq. 10 or Eq. 11 can be (close to) zero. We set $t = 1 - \varepsilon$ in these cases. If the relative error approaches zero $t = 1 - \varepsilon$ is the natural limit of Eq. 11, and if the luminance level is almost zero a potential relative error of 1 still results in a low absolute error.

**Figure 2:** *Top: $Var(V)$ and $Var(LB)$. Bottom: Estimated and actual absolute error for Fig. 1.*

### 4.1. Results and Discussion

Fig. 1 shows a relighting scenario using our technique. After obtaining a converged solution with a light source of the same geometry, we switched to the red and green pattern. Then we used the procedure outlined in the previous section to sample the radiance distribution, i.e. we reused already gathered visibility samples while respecting the given error threshold. Note that Fig. 1 is a very difficult case for decoupled visibility, because there are areas that will receive only green or only red light. A completely decoupled solution would light these areas with an attenuated yellow.

As the main purpose of our relighting application is to demonstrate Eq. 7 in practice, not to provide a state-of-the-art relighting system, we will primarily discuss issues related to Eq. 7 here.

**Error bound.** Fig. 2 shows that our upper bound is reasonably close to the actual error. However, overestimation occurs in regions where either $Var(LB/p)$ or $Var(V)$ is very large (glossy surfaces or penumbrae), due do the assumption of perfect correlation. Please see the supplemental material for a discussion of this issue. Also note that Eq. 7 delivers more than just a penumbra-detection, because it accounts for the variance in $LB$, too (as can be seen on the glossy table).

**Adaptive sampling.** The idea of mixing an unbiased with a biased estimator can be taken one step further by implementing an adaptive sampling in the sense that the number of samples an estimator receives is proportional to the blending factor. This is similar to adaptive sampling using the per-pixel variance of $L_o$, but there is an important difference: We can get a quick-start if only $LB$ changes, because we already have an estimate of $Var(V)$ and an estimate of $Var(LB/p)$ can be quickly obtained in one frame. However, the images in this paper did *not* use this extension.

**Overhead.** Without adaptive sampling the runtime overhead of our implementation for the scene in Fig. 1 was 16% for the first frame (when $Var(LB/p)$ needs to be estimated) and 5% for the following frames. Storage requirements are $2 \times 3$ floats per pixel for the running mean and variance of
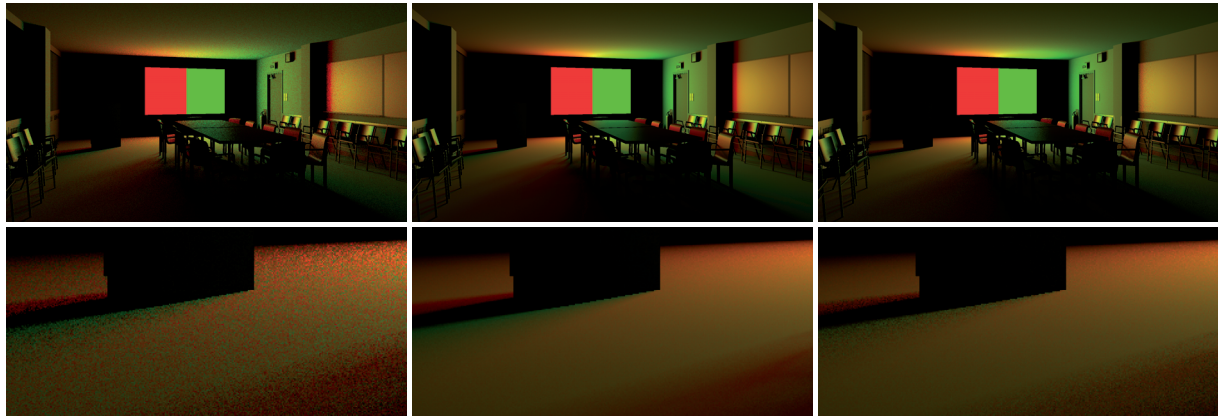
**Figure 1:** *Conference room model (by Grynberg and Ward) after a lighting change. Top row from left to right: Classical stochastic ray tracing after 9 V-samples; reference (1024 V-samples); our method with a permitted relative error of 0.25 after 8 V-samples (approx. same time as classical solution, $\varepsilon = 0.1$, Var($LB$) was estimated with 64 samples after the lighting change). Bottom row: Area around the lectern enlarged. Please note that the images are gamma-encoded and the relative error was calculated in linear RGB space.*

$LB$, and 2 floats per pixel for the mean and variance of $V$. (These costs may accrue per light source, see *Multiple light sources* below.)

**Other optimizations.** Our technique is compatible with general Monte-Carlo optimizations like stratification and basic importance sampling. Advanced variance reduction techniques and caching structures may conflict with our approach, but we consider those to be uncommon in the class of applications we target.

**Fixed view, direct light.** Since we are storing the estimates per pixel, we can only apply our method to direct lighting and have to discard the cached information if geometry or view are changing. If, however, the information was stored in a world-space data structure, the method should be usable with indirect illumination and varying viewpoints.

**Multiple light sources.** Eq. 7 works best for a single, moderately sized light source. Multiple light sources can be handled, but with a performance penalty. They can either be modeled as a single disconnected light source (which will possibly increase the variances in Eq. 7 and lead to an overly conservative error estimate) or as multiple light sources (which will require a separate estimation of means and variances and lead to additional storage costs).

## 5. Conclusion

We have analyzed the error associated with replacing the visibility function by a single scalar visibility factor. An upper bound was derived that allows an a priori estimation of the error introduced by such a decoupled sampling. We also demonstrated how our result can be used in practice in a simple relighting application.

## References

[CAM08]  CLARBERG P., AKENINE-MÖLLER T.: Exploiting visibility correlation in direct illumination. *Computer Graphics Forum 27*, 4 (2008), 1125–1136.

[CPF10]  COLBERT M., PREMOZE S., FRANCOIS G.: Importance sampling for production rendering course. SIGGRAPH 2010. http:\\sites.google.com/site/isrendering, 2010.

[DWB*06]  DONIKIAN M., WALTER B., BALA K., FERNANDEZ S., GREENBERG D. P.: Accurate direct illumination using iterative adaptive sampling. *IEEE TVCG 12* (2006), 353–364.

[GH06]  GHOSH A., HEIDRICH W.: Correlated visibility sampling for direct illumination. *The Visual Computer 22* (2006), 693–701.

[HLHS03]  HASENFRATZ J.-M., LAPIERRE M., HOLZSCHUCH N., SILLION F.: A survey of real-time soft shadows algorithms. *Computer Graphics Forum 22*, 4 (2003), 753–774.

[LLA06]  LEHTINEN J., LAINE S., AILA T.: An improved physically-based soft shadow volume algorithm. *Computer Graphics Forum 25*, 3 (2006), 303–312.

[NPW10]  NICHOLS G., PENMATSA R., WYMAN C.: Interactive, multiresolution image-space rendering for dynamic area lighting. *Computer Graphics Forum 29*, 4 (2010), 1279–1288.

[Ram09]  RAMAMOORTHI R.: Precomputation-based rendering. *Foundations and Trends in Computer Graphics and Vision 3*, 4 (2009), 281–369.

[SS96]  SOLER C., SILLION F.: Accurate error bounds for multiresolution visibility. In *7th Eurographics Workshop on Rendering* (1996), pp. 61–70.

[WFA*05]  WALTER B., FERNANDEZ S., ARBREE A., BALA K., DONIKIAN M., GREENBERG D. P.: Lightcuts: a scalable approach to illumination. In *SIGGRAPH '05* (2005), pp. 1098–1107.

[XTP07]  XIE F., TABELLION E., PEARCE A.: Soft shadows by ray tracing multilayer transparent shadow maps. In *Eurographics Symposium on Rendering* (2007), pp. 265–276.