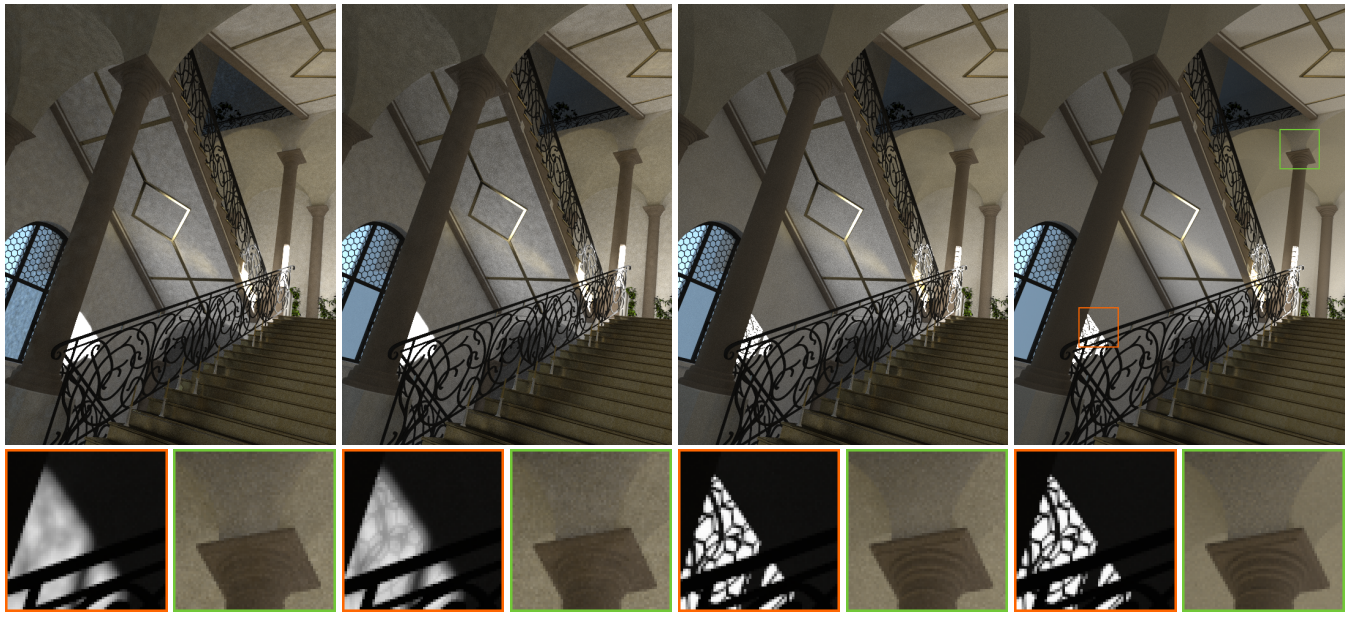
Unbiased Photon Gathering for Light Transport Simulation

Hao Qin* Xin Sun† Qiming Hou*‡ Baining Guo† Kun Zhou*

*State Key Lab of CAD&CG, Zhejiang University †Microsoft Research Asia



(a) SPPM, 1 hour (b) UPS/VCM, 1 hour (c) Our UPG+BDPT, 1 hour (d) Reference, BDPT, 2 days

Figure 1: *A comparison between stochastic progressive photon mapping (SPPM), unified path sampling/vertex connection and merging (UPS/VCM), and our unbiased photon gathering with bidirectional path tracing (UPG+BDPT) after 1 hour of rendering. SPPM utilizes biased photon mapping to produce a low-variance result, at the cost of over-blurring sharp features. UPS/VCM gains extra benefits from BDPT but the vertex merging part is still biased. Our method is both unbiased and robust, producing a result that most resembles the reference. Note that the left inset is set to an exposure of* $1/64$ *to make the HDR shadow details visible.*

Abstract

Photon mapping (PM) has been widely regarded as an efficient solution for light transport simulation, including challenging caustics paths and many-bounce indirect lighting. The efficiency of PM comes from reusing traced photons. However, the handling of photon gathering in existing PM algorithms is universally biased – the expected value of their results does not necessarily agree with the true solution of the rendering equation. We present a novel photon gathering method to efficiently achieve unbiased rendering with photon mapping. Instead of aggregating the gathered photons into an estimated density as in classical photon mapping, we process each photon individually and connect the corresponding light sub-path with the eye sub-path that generates the gather point, creating an unbiased path sample. The Monte Carlo estimate for such a path sample is calculated by evaluating all relevant terms in a strict and unbiased way, leading to a self-contained unbiased sampling technique. We further develop a set of multiple importance sampling (MIS) weights that allow our method to be optimally combined with bidirectional path tracing (BDPT), resulting in an unbiased rendering algorithm that can efficiently handle a wide variety of light paths and that compares favorably with previous algorithms. Experiments demonstrate the efficacy and robustness of our method.

CR Categories: I.3.3 [Computer Graphics]: Three-Dimensional Graphics and Realism—Raytracing

Keywords: global illumination, photon mapping, BDPT

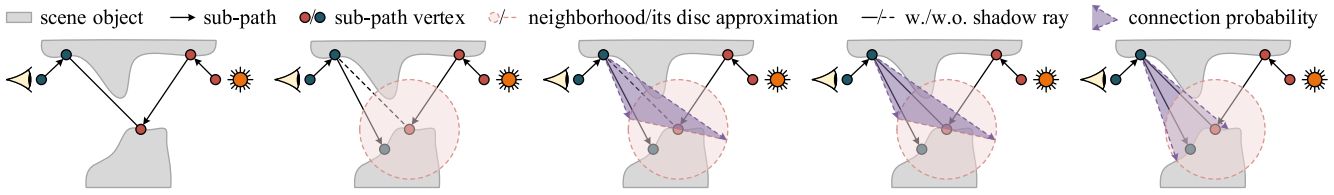
1 Introduction

Extensive study has been devoted to the robust simulation of light transport in virtual scenes. Photon mapping (PM) [Jensen 2001] has been widely regarded as one of the most efficient solutions when handling challenging caustics paths, or when significant illumination comes from many-bounce indirect lighting. However, its estimation of light transport is usually biased by photon gathering, which means that the expected value of rendered results does not necessarily agree with the true solution of the rendering equation [Kajiya 1986]. While progressive variants [Hachisuka et al. 2008; Hachisuka and Jensen 2009; Knaus and Zwicker 2011; Kaplanyan and Dachsbacher 2013a] eventually converge to the correct results, they remain biased at any finite number of samples.

*{qinneo,hqm03ster}@gmail.com, kunzhou@acm.org

†atlas.x.4@gmail.com, bainguo@microsoft.com

‡Corresponding author



(a) BDPT (b) PM (c) UPS/VCM (d) a custom designed PM (e) UPG

Figure 2: *Sub-path connection in five different algorithms. (a) A pair of sub-paths are explicitly connected with a shadow ray in BDPT [Lafortune and Willems 1993; Veach and Guibas 1994]. (b) Photon density estimation [Jensen 2001] does not perform a visibility test. (c) A gathered photon is formulated as an accepted Russian roulette connection in UPS/VCM [Hachisuka et al. 2012; Georgiev et al. 2012], but there is no visibility test for the connection and the connection probability is approximated as if the surface points within the neighborhood form a disc. (d) A custom designed density estimation [Bekaert et al. 2003] traces shadow rays for better accuracy, but the same disc approximation brings bias to the kernel normalization term. (e) UPG is unbiased because it both employs the visibility test and handles the connection probability without approximation.*

Both the efficiency and the bias in photon mapping come from the *photon gathering* step, during which eye sub-paths are traced from the sensor, and a range search is performed at their endpoints to query a photon map traced from the light sources in a previous pass. The efficiency comes from the reuse of sub-paths corresponding to the gathered photons. In classical PM methods, the gathered photons are aggregated into an estimated density that approximates local radiance or irradiance. Such an approximation produces bias, typically in the form of blurred results. As shown in Fig. 1 (a) and (b), the shadow in the zoom-in appears blurred.

In this paper, we propose a novel photon gathering method as a replacement of the density estimation to efficiently achieve unbiased rendering with photon mapping. Specifically, instead of aggregating the gathered photons, we process each photon individually and connect the corresponding light sub-path with the eye sub-path that generates the gather point. This connection can be interpreted as a Russian roulette event decided by the gather point. By removing the gather point itself from the completed path, we obtain a path sample that could contribute to the final radiance estimation. While this path reuse scheme has been formulated in previous work [Bekaert et al. 2003], the roulette probability is approximated in the denominator of the final estimate, resulting in a significant source of bias in the rendering result. Therefore, we introduce an unbiased algorithm to estimate the reciprocal of the roulette probability using a series of Bernoulli trials, making it possible to calculate the Monte Carlo estimate of the path sample by precisely evaluating all its terms. The end result is a self-contained unbiased sampling technique, which can be incorporated into a wide variety of Monte Carlo light transport algorithms. We additionally develop a set of multiple importance sampling (MIS) weights that allow our method to be optimally combined with traditional techniques.

The proposed *unbiased photon gathering* (UPG) can be combined with traditional unbiased techniques for increased robustness. We first combine it with bidirectional path tracing (BDPT) using MIS, in a manner analogous to unified path sampling (UPS) [Hachisuka et al. 2012] and vertex connection and merging (VCM) [Georgiev et al. 2012], resulting in an unbiased rendering algorithm that can efficiently handle a wide variety of light paths. The unbiased nature of our method also makes it easy to serve as an additional sampling technique in primary sampling space Metropolis light transport (PSSMLT) [Kelemen et al. 2002].

For glossy and many-bounce light paths, experimental results demonstrate that our method compares favorably to state-of-the-art alternatives, including multiplexed Metropolis light transport (MMLT) [Hachisuka et al. 2014] and UPS/VCM. While our UPG has a higher per-photon overhead than UPS/VCM, the unbiased nature of our method allows it to maintain a stable sampling speed, i.e., the number of samples generated per unit time, throughout the rendering, which is higher than the effective BDPT sampling speed that UPS/VCM eventually converges to. While UPG loses the ability to render strict specular-diffuse-specular (SDS) paths as a cost of bias-free rendering, such effects can be approximated using highly glossy surfaces, producing plausible results similar to those generated by UPS/VCM (see Fig. 13).

2 Related Work

Photon Mapping (PM). The central idea of PM [Jensen 2001] is to approximate local radiance using photon density estimated in photon gathering.[1] The locality of photon gathering makes it efficient to sample challenging caustics paths and many-bounce indirect lighting. In addition, PM can reuse light sub-paths, which increases the effective sampling rate with minimal cost. The density estimation suppresses sampling noise, but meanwhile, introduces biases. Consequently, PM is predominantly regarded as biased, although state-of-the-art variants like Progressive Photon Mapping (PPM) [Hachisuka et al. 2008] and Stochastic Progressive Photon Mapping (SPPM) are consistent [Hachisuka and Jensen 2009; Knaus and Zwicker 2011; Kaplanyan and Dachsbacher 2013a] and converge to the correct result when the sampling rate approaches infinity. MIS [Veach 1998] has been incorporated into PM [Tokuyoshi 2009; Vorba 2011] approaches to combine multiple techniques, although in a biased manner.

As an attempt to reduce bias, several researchers proposed to remove the gather point and connect its prefix vertex to the photon with a shadow ray [Bekaert et al. 2003]. However, the bias reduction is partial at best since they approximated the essential probability term (corresponding to a normalization factor in their paper). Although the way we generate light paths is technically equivalent to Bekaert et al. [2003], we choose to interpret photon gathering as a stochastic process like UPS/VCM, and provide an unbiased estimator for the created path sample.

The recently proposed particle-guided BDPT [Vorba et al. 2014] shoots photons and importons to compute a full-scene irradiance distribution, which is then used to guide individual ray-surface interactions in a separate unbiased pass. Their method operates on a different principle than PM, as the photons are only used as guides and do not make a direct contribution to the final image. Our

[1] The term "photon gathering" was first proposed as another alternative to photon density estimation which directly connects photons to eye sub-paths with shadow rays [Jensen 2001], though it is occasionally used to refer to the search of nearby photons. In this paper, we take the second meaning to linguistically distinguish the photon search from the traditionally-coupled density estimation.

method is orthogonal to their approach – their guides can be directly applied to improve sub-path sampling in our method. Our preliminary implementation of a combined method has demonstrated considerable benefits (see Fig. 12).

Bidirectional Path Tracing (BDPT). The rendering equation [Kajiya 1986] is a physically-based formulation of general light transport, which can be solved by sampling a path integral using Monte Carlo methods. BDPT [Lafortune and Willems 1993; Veach and Guibas 1994] samples paths by tracing light sub-paths and eye sub-paths separately and connecting the endpoints with shadow rays, which is especially efficient for indoor scenes with finite-sized light sources and sensors. With MIS and the path space formulation of BDPT [Veach 1998], a wide variety of path sampling techniques, including ours, can be combined for increased robustness. The final result is unbiased as long as all involved techniques are unbiased.

Unified Sampling. Both PM and BDPT sample paths by connecting light sub-paths and eye sub-paths. Compared to BDPT connection, the locality between gather points and photons in photon gathering improves sampling efficiency especially for caustics paths. Accordingly, PM and BDPT methods generally specialize in mutually complementary light transport types [Hašan et al. 2009; Vorba 2011]. Therefore, an effective combination of both would be significantly more robust than either approach. The key is to integrate BDPT connection and photon gathering with a unified formulation of path integrals.

VCM [Georgiev et al. 2012] redefines photon gathering as a probabilistic bidirectional connection which rejects sub-path pairs with spatially distant endpoints, whereas UPS [Hachisuka et al. 2012] considers BDPT connections as a special case of PM where a photon overlaps the gather point exactly. The two formulations result in an equivalent path-space combination of SPPM [Hachisuka and Jensen 2009] and BDPT, where MIS could be utilized for increased robustness. However, both formulations rely on photon density estimation, and thereby inherit the estimation biases from photon gathering. The final result would remain biased as in photon mapping. In contrast, our unbiased photon gathering formulation excludes the approximated evaluation in density estimation. It directly evaluates the Monte Carlo estimate of the path samples in an unbiased way. Our method can be combined with BDPT in a manner analogous to UPS and VCM, but the final result is unbiased.

Markov Chain Monte Carlo (MCMC). MCMC was introduced to the graphics community as Metropolis light transport (MLT) [Veach and Guibas 1997], which uses the Metropolis-Hastings algorithm [Metropolis et al. 1953] to improve bidirectional path sampling. Since then, a wide variety of mathematical and practical improvements [Kelemen et al. 2002; Cline et al. 2005; Kitaoka et al. 2009; Hachisuka and Jensen 2011; Jakob and Marschner 2012; Kaplanyan and Dachsbacher 2013b; Kaplanyan et al. 2014; Hachisuka et al. 2014] have been proposed in the MCMC framework. However, analogous to other approaches based on path sampling, MCMC algorithms would still become biased when combined with biased methodology [Hachisuka and Jensen 2011; Kaplanyan and Dachsbacher 2013b].

Many-light Methods. Many-light rendering [Keller 1997; Kollig and Keller 2004; Walter et al. 2005; Walter et al. 2006; Hašan et al. 2009; Davidovič et al. 2010; Walter et al. 2012] is another effective alternative to Monte Carlo approaches which is known for its high scalability and low noise level. Drawing parallels between virtual lights and light sub-paths, many-light methods resemble BDPT more than PM as the virtual lights are connected to eye sub-paths with shadow rays rather than local gathering [Segovia et al. 2006; Walter et al. 2012]. Consequently the end-to-end behavior resembles that of BDPT, to which PM remains complementary. This is especially true when handling specific light transport types such as caustics. While this paper focuses on photon mapping, VPL schemes with probabilistic shadow ray tests can potentially employ an unbiased reciprocal estimation scheme analogous to ours.

3 Unbiased Photon Gathering

Light transport simulation is generally formulated as path integration [Veach 1998], and classical BDPT [Lafortune and Willems 1993; Veach and Guibas 1994] is its unbiased estimation by connecting light sub-path / eye sub-path pairs with a shadow ray, as shown in Fig. 2 (a). Photon mapping [Jensen 2001] takes a different approach. Instead of connecting a single light sub-path, the incident radiance at the end point of an eye sub-path is approximated by photon density estimated from nearby photons. As in Fig. 2 (b), it does not perform any visibility test and the bidirectional scattering distribution function (BSDF) term is evaluated for a wrong direction, which creates bias. The equivalent formulations in UPS [Hachisuka et al. 2012] and VCM [Georgiev et al. 2012] unify BDPT connection and photon gathering within the same path integral, however, in a biased way. As illustrated in Fig. 2 (c), they formulate the photon gathering as a hypothetical Russian roulette process without verifying the visibility of accepted connections. In addition, they also approximate the connection probability assuming the surface point within the neighborhood is a perfect disc. Before UPS/VCM, additional visibility tests have been proposed to improve the accuracy of density estimation [Bekaert et al. 2003] (see Fig. 2 (d)), where the UPS/VCM connection probability corresponds to a normalization term. However, they make the same disc approximation, which biases the final estimation. UPG follows the Russian roulette formulation of UPS/VCM, but handles the probability term without any approximation as illustrated in Fig. 2 (e).

In this section, we start with the basic formulations of the path integral and BDPT, and then review the formulation of vertex merging [Georgiev et al. 2012] which interprets photon gathering as a technique for path sampling (Sec. 3.1). Based on the formulation of vertex merging, we analyze the causes of the biases from photon gathering (Sec. 3.2). After that, we introduce our formulation of explicit Russian roulette connection to remove the bias in photon gathering (Sec. 3.3). Finally, we describe an unbiased algorithm to estimate the reciprocal of the sampling probability in photon gathering using a series of Bernoulli trials (Sec. 3.4), and a rendering algorithm using our unbiased photon gathering under the framework of photon mapping (Sec. 3.5).

3.1 Review of Path Integral and Vertex Merging

Path Integral and BDPT. We use the Monte Carlo path integral formulation of light transport:

$$I = \int_{\Omega} f(\bar{\mathbf{x}}) \mathrm{d}\mu(\bar{\mathbf{x}}) \approx \frac{1}{n} \sum_{i=1}^{n} \frac{f(\bar{\mathbf{x}}_i)}{p(\bar{\mathbf{x}}_i)}, \tag{1}$$

where $\bar{\mathbf{x}}$ denotes a path between a light source and a sensor, n is the number of samples, $f(\bar{\mathbf{x}})$ is the energy it carries and $p(\bar{\mathbf{x}})$ is the probability density function used to sample $\bar{\mathbf{x}}$. Classical BDPT connects a light sub-path $y_1 \cdots y_s$ to an eye sub-path $z_t \cdots z_1$ by tracing the shadow ray $y_s z_t$. The resulting full path $\bar{\mathbf{x}}_{s,t} = y_1 \cdots y_s z_t \cdots z_1$ has a length of $k = s + t - 1$. We can define the corresponding Monte Carlo estimator in Eq. (1) as:

$$C^*(\bar{\mathbf{x}}_{s,t}) = \frac{f(\bar{\mathbf{x}}_{s,t})}{p(\bar{\mathbf{x}}_{s,t})} = \alpha^L(\bar{\mathbf{x}}_{s,t}) f^c(\bar{\mathbf{x}}_{s,t}) \alpha^E(\bar{\mathbf{x}}_{s,t}), \tag{2}$$

where $\alpha^L(\bar{\mathbf{x}}_{s,t})$ and $\alpha^E(\bar{\mathbf{x}}_{s,t})$ are the light and eye sub-path terms respectively. For now we will ignore their definitions and focus on

$f^c(\bar{\mathbf{x}}_{s,t})$, which is the measure term of the sub-path connection

$$f^c(\bar{\mathbf{x}}_{s,t}) = f_s(y_{s-1} \to y_s \to z_t) \cdot G(y_s \leftrightarrow z_t) f_s(z_{t-1} \to z_t \to y_s), \quad (3)$$

where f_s is the bidirectional scattering distribution function (BSDF) and G is the geometry term that converts the area measure at z_t to the projected solid angle measure [Veach 1998] centered at y_s, which also takes the mutual visibility into account.

Photon Gathering as Vertex Merging. In photon mapping, the incident radiance at the end point of an eye sub-path is evaluated through photon density estimation, which is computed by gathering photons within a local neighborhood. The neighborhood size is typically specified as a radius d. According to the equivalent formulations of UPS/VCM, photon gathering can be considered as a process that pairs each eye sub-path $z_{t'} \cdots z_1$ with all light sub-paths $y_1 \cdots y_{s'}$ satisfying $\|y_{s'} - z_{t'}\| < d$. Here we follow the *vertex merging* formulation in VCM as it explains our method more elegantly.

Vertex merging requires one of the sub-path endpoints to be removed from the final path. Without loss of generality, we remove $z_{t'}$, and the resulting complete path is $y_1 \cdots y_{s'} z_{t'-1} \cdots z_1$, which we denote as $\bar{\mathbf{x}}_{s',t'-1}$. The estimator corresponding to Eq. (2) is

$$C^*_{VM}(\bar{\mathbf{x}}_{s',t'-1}) = \alpha^L(\bar{\mathbf{x}}_{s',t'-1}) \alpha^c_{VM}(\bar{\mathbf{x}}_{s',t'-1}) \alpha^E(\bar{\mathbf{x}}_{s',t'-1}),$$
$$\alpha^c_{VM}(\bar{\mathbf{x}}_{s',t'-1}) = \frac{f^c_{VM}(\bar{\mathbf{x}}_{s',t'-1})}{p^c_{VM}(\bar{\mathbf{x}}_{s',t'-1})}, \quad (4)$$

where the connection term $f^c(\bar{\mathbf{x}}_{s,t})$ in Eq. (2) becomes $\alpha^c_{VM}(\bar{\mathbf{x}}_{s',t'-1})$, which includes both the measurement contribution $f^c_{VM}(\bar{\mathbf{x}}_{s',t'-1})$ and the connection probability $p^c_{VM}(\bar{\mathbf{x}}_{s',t'-1})$.

3.2 Bias of Photon Gathering

Vertex merging in Eq. (4) explicitly formulates photon gathering as a hypothetical Russian roulette process. The probability p^c_{VM} is the corresponding acceptance probability, where $z_{t'}$ is interpreted as the end point of a random tentative ray. The connection between $y_1 \cdots y_{s'}$ and $z_{t'-1} \cdots z_1$ is only accepted if $z_{t'}$ falls within the neighborhood of $y_{s'}$, which corresponds to the neighborhood search in photon gathering. While this interpretation seemingly goes against engineering intuition, it renders photon gathering as a bidirectional path sampling technique. This allows us to explicitly formulate the bias in classical photon mapping, which is the first step towards unbiased photon gathering.

Comparing Eq. (4) with the unbiased BDPT formulation in Eq. (2), we can see that if $\alpha^c_{VM}(\bar{\mathbf{x}}_{s',t'-1})$ in Eq. (4) could have been an unbiased Monte Carlo estimator of $f^c(\bar{\mathbf{x}}_{s,t})$ in Eq. (2), the estimation $C^*_{VM}(\bar{\mathbf{x}}_{s',t'-1})$ would be unbiased just like $C^*(\bar{\mathbf{x}}_{s,t})$. This requires f^c_{VM} to be equal to f^c and p^c_{VM} to be the real probability for the $z_{t'}$ to accept the Russian roulette connection. However, in vertex merging, C^*_{VM} is simply evaluated using an equivalent form of the density estimation in classical photon mapping, and neither of the two requirements holds true.

Specifically, in vertex merging, f^c_{VM} in Eq. (4) is not directly related to the connection between $y_{s'}$ and $z_{t'-1}$

$$f^c_{VM}(\bar{\mathbf{x}}_{s',t'-1}) = f_s(y_{s'-1} \to y_{s'}(z_{t'}) \to z_{t'-1}) \cdot G(z_{t'} \leftrightarrow z_{t'-1}) f_s(z_{t'-2} \to z_{t'-1} \to z_{t'}), \quad (5)$$

where $f_s(y_{s'-1} \to y_{s'}(z_{t'}) \to z_{t'-1})$ indicates that the BSDF is still evaluated at $y_{s'}$ but the outgoing direction is taken as $z_{t'} \to z_{t'-1}$ instead of $y_{s'} \to z_{t'-1}$.

In contrast, according to Eq. (3) the strict evaluation should be

$$f^c(\bar{\mathbf{x}}_{s',t'-1}) = f_s(y_{s'-1} \to y_{s'} \to z_{t'-1}) \cdot G(y_{s'} \leftrightarrow z_{t'-1}) f_s(z_{t'-2} \to z_{t'-1} \to y_{s'}). \quad (6)$$

Comparing Eq. (5) with Eq. (6), one can see that both the geometry term and the BSDF term at $y_{s'}$ are approximated, because the edge $y_{s'} \leftrightarrow z_{t'-1}$ is replaced by $z_{t'} \leftrightarrow z_{t'-1}$, which is an approximation that originated from classical photon mapping.

The key idea introduced by UPS/VCM is that the photon gathering neighborhood determines the probability of making such a connection. In VCM, p^c_{VM} in Eq. (4) also takes an equivalent form of photon density estimation using a constant disc filtering kernel

$$p^c_{VM}(\bar{\mathbf{x}}_{s',t'-1}) = \pi d^2 \tilde{p}_x(z_{t'-2} \to z_{t'-1} \to y_{s'}), \quad (7)$$

where $\tilde{p}_x(z_{t'-2} \to z_{t'-1} \to y_{s'})$ is the probability density to sample vertex $y_{s'}$ from two preceding vertices $z_{t'-2}$ and $z_{t'-1}$. The tilde indicates that for the purpose of computing $\tilde{p}_x$, $y_{s'}$ is always assumed to be visible from $z_{t'-1}$ even if in fact it is not.

In contrast, the strict evaluation has to properly integrate p_x over every possible $z_{t'}$ in the gathering radius, which is the same as the acceptance probability without approximation in VCM and gives

$$p^c(\bar{\mathbf{x}}_{s',t'-1}) = \int_{S(y_{s'},d)} p_x(z_{t'-2} \to z_{t'-1} \to z) \mathrm{d}z. \quad (8)$$

The integration domain $S(y_{s'}, d)$ is the gathering neighborhood, which consists of all surface points within a sphere of radius d centered at $y_{s'}$. Unlike $\tilde{p}_x$, the integrand p_x has the correct visibility term embedded and $p_x(z_{t'-2} \to z_{t'-1} \to z) = 0$ if z is not visible from $z_{t'-1}$.

Comparing Eq. (7) with Eq. (8), $p^c_{VM}(\bar{\mathbf{x}}_{s',t'-1})$ approximates the integration using one sample at $\tilde{p}_x(z_{t'-2} \to z_{t'-1} \to y_{s'})$ without verifying the visibility, and the area of the neighborhood $S(y_{s'}, d)$ is also approximated as the simple disk area πd^2.

Biases in existing methods. According to the term-wise comparison between the BDPT estimator in Eq. (2) and PM estimator in Eq. (4), we can see that the most essential change photon mapping introduced is the connection probability p^c_{VM}. Existing methods typically use a disc approximation for this term, whether it is interpreted as a Russian roulette probability [Hachisuka et al. 2012; Georgiev et al. 2012] (Fig. 2 (c)) or as a normalization factor in a customized density estimation [Bekaert et al. 2003] (Fig. 2 (d)). The disc approximation is an essential source of estimation biases, and our method is the first to remove it. Details will be introduced in Sec. 3.3 and Sec. 3.4.

Note that the path reuse formulation of photon gathering as well as the connection probability have been described in [Bekaert et al. 2003], and are thus not the contributions of this paper. Our main contribution is an unbiased algorithm to estimate the reciprocal of the connection probability using a series of Bernoulli trials (Sec. 3.3 and Sec. 3.4), making it possible to calculate the Monte Carlo estimate of the path sample for the first time.

3.3 Explicit Russian Roulette Connection

Having analyzed the bias of classical photon mapping, a natural direction is to replace Eq. (5) and Eq. (7) with the corresponding strict evaluations in Eq. (6) and Eq. (8) respectively, which forms an unbiased estimation

$$C^*_{RR}(\bar{\mathbf{x}}_{s',t'-1}) = \alpha^L(\bar{\mathbf{x}}_{s',t'-1}) \alpha^c(\bar{\mathbf{x}}_{s',t'-1}) \alpha^E(\bar{\mathbf{x}}_{s',t'-1}),$$
$$\alpha^c(\bar{\mathbf{x}}_{s',t'-1}) = \frac{f^c(\bar{\mathbf{x}}_{s',t'-1})}{p^c(\bar{\mathbf{x}}_{s',t'-1})}. \quad (9)$$

Eq. (9) is a strict formulation of the hypothetical Russian roulette event in photon gathering. $z_{t'}$ is interpreted as a rejection sample independent of $\bar{\mathbf{x}}_{s',t'-1}$, which conditionally rejects the connection between the two sub-paths $y_1 \cdots y_{s'}$ and $z_{t'-1} \cdots z_1$. The connection is only accepted if $z_{t'}$ falls within the neighborhood $S(y_{s'}, d)$, the probability of which is p^c.

The explicit Russian roulette connection in Eq. (9) differs from vertex merging in Eq. (4) by utilizing the strict evaluations of f^c and p^c. To truly eliminate bias, all terms in the strict formulas Eq. (6) and Eq. (8) must be evaluated in an unbiased manner. As Eq. (6) can be evaluated precisely by tracing the shadow ray between $y_{s'}$ and $z_{t'-1}$, as proposed by Bekaert et al. [2003], the main challenge comes to Eq. (8), which is the connection probability p^c.

Note that $p_x(z_{t'-2} \to z_{t'-1} \to z)$ in Eq. (8) depends on the geometry term $G(z_{t'-1} \leftrightarrow z)$, which additionally depends on a visibility term, making an analytical integration practically impossible. Thus we introduce a separate Monte Carlo estimate to handle p^c. However, since p^c is ultimately used as a denominator in Eq. (9), even an unbiased Monte Carlo estimate of Eq. (8) would still bias the final result. Jensen's inequality [Chandler 1987] leads to:

$$\frac{1}{E\left[p^c\right]} \geq E\left[\frac{1}{p^c}\right]. \tag{10}$$

The equality only holds true when the variance is zero, which clearly does not hold for general sub-path pairs. Therefore, we must develop an unbiased method that directly estimates the *probability reciprocal* $r(\bar{\mathbf{x}}) = 1/p^c(\bar{\mathbf{x}})$.

3.4 Unbiased Probability Reciprocal Estimation

Our probability reciprocal estimation is inspired by Booth's work [2007], which introduces an unbiased solution to the more general problem of estimating the reciprocal of an arbitrary integral. Under Booth's formulation, our method is a special case where the unknown portion of the integrand is a binary function (see details in Appendix C).

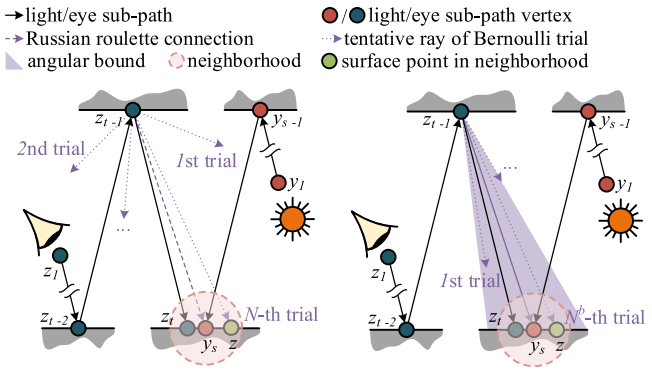
As illustrated in Fig. 3(a), suppose we literally follow through with the hypothetical Russian roulette test described in Sec. 3.1. For each path $\bar{\mathbf{x}}_{s',t'-1}$, we generate the tentative ray $z_{t'-1} \to z$ following the distribution $p_x(z_{t'-2} \to z_{t'-1} \to z)$. Based on the definition in Eq. (8), there is a $p^c(\bar{\mathbf{x}}_{s',t'-1})$ probability for each ray to be accepted by Russian roulette. Repeatedly generating tentative rays for the same path $\bar{\mathbf{x}}_{s',t'-1}$ gives a series of Bernoulli trials. Assuming that the first accepted ray comes after $N(\bar{\mathbf{x}}_{s',t'-1})$ tries, one could see that the random variable $N(\bar{\mathbf{x}}_{s',t'-1})$ follows a geometric distribution

$$\Pr\left(N(\bar{\mathbf{x}}_{s',t'-1}) = i\right) = p^c(\bar{\mathbf{x}}_{s',t'-1})(1 - p^c(\bar{\mathbf{x}}_{s',t'-1}))^{i-1}, \tag{11}$$

and the expectation of N is exactly what we need:

$$\begin{aligned} E\left[N(\bar{\mathbf{x}}_{s',t'-1})\right] &= \sum_{i=1}^{+\infty} i p^c(\bar{\mathbf{x}}_{s',t'-1}) \left(1 - p^c(\bar{\mathbf{x}}_{s',t'-1})\right)^{i-1} \\ &= \frac{1}{p^c(\bar{\mathbf{x}}_{s',t'-1})} = r(\bar{\mathbf{x}}_{s',t'-1}). \end{aligned} \tag{12}$$

In summary, an unbiased estimation of the probability reciprocal can be computed as the number of tentative rays tested until the first successful Bernoulli trial, i.e., up to and including the first ray that hits the gathering sphere $S(y_{s'}, d)$.



(a) Bernoulli trials (b) With angular bound

Figure 3: *Unbiased estimation of the probability reciprocal. (a) illustrates the raw formulation of tracing tentative rays and using the gathering sphere as a Russian roulette test. N, the number of trials required to produce a first hit, is an unbiased estimation of the probability reciprocal. (b) illustrates the efficiency improvement brought by the angular bound.*

Angular Bound for Tentative Rays. An obvious issue of the aforementioned algorithm is the potentially unbounded number of Bernoulli trials N. However, note that the Bernoulli trials are only triggered with a probability of p^c, and N is an estimation of its reciprocal. When amortized over all connections, the *expected* number of Bernoulli trials per connection is simply:

$$E\left[N(\bar{\mathbf{x}}_{s',t'-1})p^c(\bar{\mathbf{x}}_{s',t'-1})\right] = r(\bar{\mathbf{x}}_{s',t'-1})p^c(\bar{\mathbf{x}}_{s',t'-1}) = 1. \tag{13}$$

Thus, on average the total number of required Bernoulli trials is equal to the total number of connections. However, since the Russian roulette formulation in Sec. 3.1 assumes *all* sub-path pairs are tested for connection, the bound is $O(n^2)$ for n sub-paths, which is still prohibitively expensive. To make the Bernoulli trial overhead practical, we must avoid testing potential connections that would be an obvious rejection.

As a solution, we exploit the restricted size of the gathering neighborhood $S(y_{s'}, d)$. Specifically, we project the neighborhood onto the unit sphere centered around $z_{t'-1}$, and only trace tentative rays within the resulting angular bound as illustrated in Fig. 3(b). Consequently, the expected number of Bernoulli trials per connection is reduced by a factor proportional to the squared distance between $y_{s'}$ and $z_{t'-1}$, providing $d \ll \|y_{s'} - z_{t'-1}\|$. Ignoring the effects of BSDF importance sampling, we have:

$$\frac{E\left[N^b(\bar{\mathbf{x}}_{s',t'-1})\right]}{E\left[N(\bar{\mathbf{x}}_{s',t'-1})\right]} \approx \frac{\pi d^2}{\|y_{s'} - z_{t'-1}\|^2}, \tag{14}$$

where $N^b(\bar{\mathbf{x}}_{s',t'-1})$ is the number of Bernoulli trials after applying the angular bounding. The bound may degenerate to the entire upper hemisphere for extremely close pairs of $z_{t'-1}$ and $t_{s'}$. However, such cases only constitute a tiny portion of the Russian roulette connections. Based on our experiments, this bounding reduces the total overhead of our algorithm to approximately $O(n)$.

With N^b Bernoulli trials instead of N, the probability reciprocal estimation in Eq. (12) has to be adjusted accordingly:

$$r(\bar{\mathbf{x}}_{s',t'-1}) = \frac{E\left[N^b(\bar{\mathbf{x}}_{s',t'-1})\right]}{p^b(\bar{\mathbf{x}}_{s',t'-1})}, \tag{15}$$

$$p^b(\bar{\mathbf{x}}_{s',t'-1}) = \int_{\mathbf{\Omega}_b} p_x(z_{t'-2} \to z_{t'-1} \to z)\mathrm{d}z, \tag{16}$$

where $\mathbf{\Omega}_b$ is the set of surface points visible to $z_{t'-1}$ and inside the angular bound. $p^b(\bar{\mathbf{x}}_{s',t'-1})$ is the probability for a tentative ray traced without the angular bound to fall within $\mathbf{\Omega}_b$. As defined in Eq. (13) and Eq. (15), $p^b(\bar{\mathbf{x}}_{s',t'-1})$ is also the ratio by which the angular bound reduces the number of Bernoulli trials: $p^b(\bar{\mathbf{x}}_{s',t'-1}) = E\left[N^b(\bar{\mathbf{x}}_{s',t'-1})\right] / E\left[N(\bar{\mathbf{x}}_{s',t'-1})\right]$.

$\mathbf{\Omega}_b$ can be chosen arbitrarily as long as it includes all points in $S(y_{s'}, d)$ that are visible to $z_{t'-1}$. Note that $p^b(\bar{\mathbf{x}}_{s',t'-1})$ in Eq. (16) is not a random variable and must be evaluated analytically. This type of integration has been well researched in the context of BSDF importance sampling. We simply bound the random number used to generate the tentative ray direction. Details can be found in the supplementary material, which includes an example mathematical derivation for a combined Lambertian and Phong BSDF. The angular bound also significantly reduces the estimation variance of the probability reciprocal r, as discussed in Appendix A.

To prove that the angular bound does not introduce additional bias, we substitute Eq. (15) into Eq. (9) and verify that the expectation of the resulting estimator is still correct

$$
\begin{aligned}
&E\left[\alpha^L(\bar{\mathbf{x}}_{s',t'-1})\frac{f^c(\bar{\mathbf{x}}_{s',t'-1})N^b(\bar{\mathbf{x}}_{s',t'-1})}{p^b(\bar{\mathbf{x}}_{s',t'-1})}\alpha^E(\bar{\mathbf{x}}_{s',t'-1})\right] \\
&= E\left[C^*_{RR}(\bar{\mathbf{x}}_{s',t'-1})\right] E\left[\frac{p^c(\bar{\mathbf{x}}_{s',t'-1})N^b(\bar{\mathbf{x}}_{s',t'-1})}{p^b(\bar{\mathbf{x}}_{s',t'-1})}\right] \\
&= E\left[C^*_{RR}(\bar{\mathbf{x}}_{s',t'-1})\right] \frac{E\left[N^b(\bar{\mathbf{x}}_{s',t'-1})\right]}{r(\bar{\mathbf{x}}_{s',t'-1})p^b(\bar{\mathbf{x}}_{s',t'-1})} \\
&= E\left[C^*_{RR}(\bar{\mathbf{x}}_{s',t'-1})\right], \qquad (17)
\end{aligned}
$$

Notice that $p^c(\bar{\mathbf{x}}_{s',t'-1})$ and $p^b(\bar{\mathbf{x}}_{s',t'-1})$ can be taken out of $E[\cdot]$ as their values are constant for any given path $\bar{\mathbf{x}}_{s',t'-1}$.

We note that our method can be regarded as the first application of Booth's method in global illumination. However, a straightforward application of Booth's formulation is prohibitively expensive. We simplify it as a series of Bernoulli trials and optimize it with angular bounding to make it practical, which can be thought of as the auxiliary function in Booth's framework. This is analogous to the fact that while most unbiased rendering algorithms are applications of the Monte Carlo estimate method, it is the actual sampling technique that is significant in the graphics context.

3.5 Rendering Algorithm

Algorithm 1 summarizes the rendering algorithm using our unbiased photon gathering (UPG) under the framework of photon mapping. Note that the algorithm generates a photon map and performs gathering just like in classical photon mapping. The key difference is in lines 5-16, which generate an unbiased path sample for each photon instead of performing a density estimation.

Compared to PPM-based methods, our algorithm shares the same limitation where the photon gathering in line 4 can only use a fixed-radius search rather than the classical k-nearest neighbor (kNN) search [Jensen 2001]. We require an explicit radius to form a well-defined integration domain in Eq. (8). However, unlike PPM, our method is unbiased, which makes progressive rendering a trivial matter of running Algorithm 1 many times and averaging the results (see the error of UPG+BDPT reducing to zero in Fig. 7). Thus, our neighborhood gathering radius d can be tuned arbitrarily and independently in each rendering session, as d does not affect convergence and only impacts computational efficiency.

Compared to the photon gathering in UPS/VCM, our method uses an equivalent interpretation to formulate photon mapping in the

Algorithm 1 Pseudo code of unbiased photon gathering

1: Generate a fixed number of light sub-paths $\{\bar{\mathbf{y}}\}$
2: Build a photon map using the endpoints $\{y_{s'}\}$
3: **for** each eye sub-path $\bar{\mathbf{z}} = z_{t'} \cdots z_1$ **do**
4: Gather photons in the radius-d neighborhood $S(z_{t'}, d)$
5: **for** the light sub-path $\bar{\mathbf{y}}$ of each gathered photon **do**
6: Project $S(y_{s'}, d)$ to an angular bound Ω_b around $z_{t'-1}$
7: $N^b \leftarrow 0$
8: **do**
9: $N^b \leftarrow N^b + 1$
10: Sample a tentative ray $z_{t'-1} \rightarrow z$ from Ω_b
11: **while** $z \notin S(y_{s'}, d)$
12: Create a complete path $\bar{\mathbf{x}} = y_1 \cdots y_{s'} z_{t'-1} \cdots z_1$
13: Compute $p^b(\bar{\mathbf{x}})$ using Eq. (16)
14: $C^*_{RR}(\bar{\mathbf{x}}) = \alpha^L(\bar{\mathbf{x}}) \left(f^c(\bar{\mathbf{x}})\frac{N^b}{p^b(\bar{\mathbf{x}})}\right) \alpha^E(\bar{\mathbf{x}})$ as in Eq. (17)
15: Accumulate $C^*_{RR}(\bar{\mathbf{x}})$ to the pixel corresponding to $\bar{\mathbf{x}}$
16: **end for**
17: **end for**

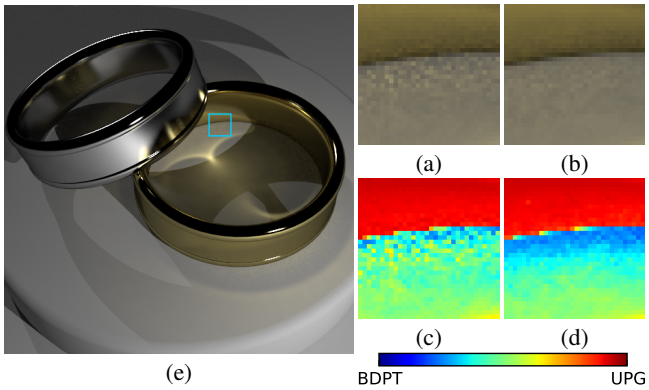
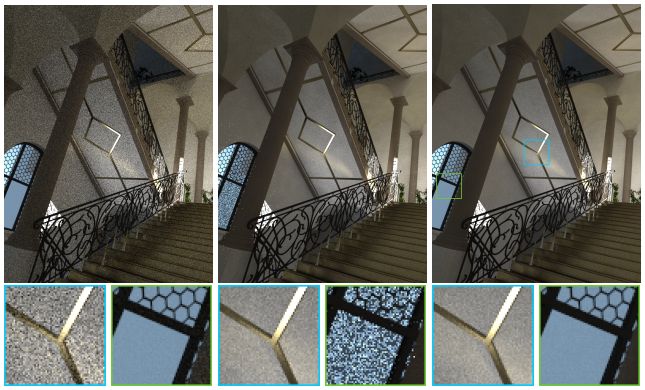


Figure 4: *The effect of clamping the approximated p^c to 1. (a) is rendered without clamping. (b) is rendered with clamping. (c) and (d) visualize the per-pixel relative contribution of UPG in (a) and (b) respectively. (e) is the complete image rendered with clamping enabled, and the rendering takes 10 minutes.*

BDPT path space. However, those approaches focus on combining PM with BDPT in general, and bias is not a major concern. Hence, they utilize approximations for their equivalent terms of $p^c(\bar{\mathbf{x}})$. While such approximations are biased, they are simpler and more efficient than our unbiased estimation. This is not surprising as removing bias would unavoidably incur some additional cost.

Compared to BDPT, our method has to process a non-trivial number of light and eye sub-paths in batch to be effective. While technically Algorithm 1 remains correct for a single pair of sub-paths, doing so hardly has any practical value due to the additional Bernoulli trials required. Our scheme retains the inherent behaviors of PM-based approaches, and the most critical factor in efficiency is the reuse of sub-paths [Hachisuka et al. 2012; Georgiev et al. 2012]. Significant benefits are only possible when each eye sub-path can be connected to multiple light sub-paths and vice versa.

Finally, two implementation details should be noted. First, when our method is used in conjunction with BDPT using formulations like UPS/VCM, it is possible to reuse the same set of light and eye sub-paths in both algorithms. As previously discussed, when connecting the same pair of sub-paths $y_1 \cdots y_s$ and $z_t \cdots z_1$, BDPT generates a path $y_1 \cdots y_s z_t \cdots z_1$ which has $s + t - 1$ edges, whereas our method generates a path $y_1 \cdots y_s z_{t-1} \cdots z_1$ which has $s+t-2$ edges with z_t excluded. The complete paths do not overlap



(a) BDPT (b) UPG (c) UPG+BDPT

Figure 5: *The effect of UPG in BDPT. (a) is rendered using classical BDPT. (b) is rendered using pure UPG. (c) is rendered by combining both methods. The images are rendered with the same computation time (1 hour).*

in the path space, and they can be treated as independent in practice.

Second, in the previous sections we have only discussed about the path $\bar{\mathbf{x}}_{s',t'-1}$ generated by removing $z_{t'}$. In practice one could also consider the alternative that removes $y_{s'}$ instead and generates $\bar{\mathbf{x}}_{s'-1,t'}$. While it is possible to evaluate both paths, in practice we found that doing so is not cost-effective. The more efficient approach is to evaluate one of the two candidates and discard the other, where we make the decision heuristically by removing the vertex generated from the less glossy reflection. A more detailed explanation is provided in Appendix B.

Our UPG can also be taken as an individual unbiased sampling technique and combined with other sampling techniques under the unified framework of MIS. This is analogous to UPS/VCM, where a biased SPPM is combined instead. We follow the VCM framework for combination, since we choose to exclude $z_{t'}$ from the complete path rather than including both endpoints in an extended path space. In Sec. 4 and Sec. 5, UPG is combined with BDPT and PSSMLT respectively. The combination of UPG with another unbiased method is, of course, still unbiased.

4 MIS Combination with BDPT

Now we describe how to combine our UPG with BDPT using MIS. With multiple sampling techniques, MIS formulates the Monte Carlo path integral in Eq. (1) as

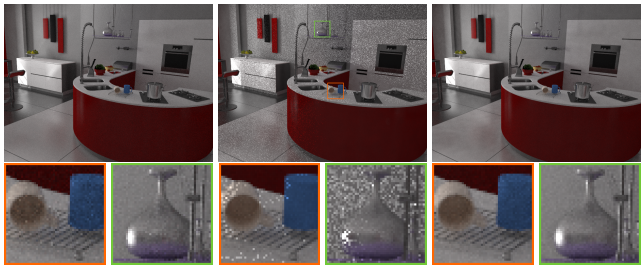
$$I \approx \sum_{i=1}^{m} \frac{1}{n_i} \sum_{j=1}^{n_i} w_i(\bar{\mathbf{x}}_{i,j}) \frac{f(\bar{\mathbf{x}}_{i,j})}{p_i(\bar{\mathbf{x}}_{i,j})}, \tag{18}$$

where m is the number of sampling techniques, n_i is the sample count taken for technique i. $p_i(\bar{\mathbf{x}})$ and $w_i(\bar{\mathbf{x}})$ are the corresponding path probability density values and MIS weights. For the combination to be unbiased, the weights $w_i(\bar{\mathbf{x}})$ must be consistent when the same $\bar{\mathbf{x}}$ is generated from different techniques. In addition, w_i must meet the normalization and non-negativity requirements:

$$\sum_{i=1}^{m} w_i(\bar{\mathbf{x}}) = 1, w_i(\bar{\mathbf{x}}) \geq 0. \tag{19}$$

We use the widely-accepted power heuristic

$$w_i(\bar{\mathbf{x}}) = \frac{(n_i p_i(\bar{\mathbf{x}}))^\beta}{\sum_{j=1}^{m} (n_j p_j(\bar{\mathbf{x}}))^\beta}, \tag{20}$$



(a) PSSMLT (b) UPG (c) UPG+PSSMLT

Figure 6: *The effect of UPG in PSSMLT. (a) is rendered with PSSMLT. (b) is rendered with pure UPG and (c) is rendered with UPG+PSSMLT. The images are rendered within the same computation time (30 min). We can see that UPG helps PSSMLT to improve sampling quality for many-bounce light paths and glossy materials, and combining with PSSMLT down-weights low probability UPG paths, which eliminates the bright spike artifacts.*

where $\beta > 0$ is a predefined parameter. We take $\beta = 2$ as recommended by Veach [1998]. Since Eq. (20) satisfies Eq. (19) by definition, the main implementation challenge lies in computing the relevant probability density terms p_j for paths sampled using a different technique i.

Given a path $\bar{\mathbf{x}}$ with k edges, there are $k+1$ BDPT sampling techniques and $k-1$ UPG techniques, corresponding to different choices of the pre-connection sub-path lengths. The number difference comes from the fact that BDPT allows zero-length sub-paths. Regardless of the technique chosen to generate $\bar{\mathbf{x}}$, it is straightforward to compute the probability density of BDPT techniques directly from the complete path:

$$p_{s,t}^{\mathrm{BDPT}}(\bar{\mathbf{x}}) = p^L(\bar{\mathbf{x}}_{s,t}) p^E(\bar{\mathbf{x}}_{s,t}), \tag{21}$$

where p^L and p^E are the classical sub-path probability density terms [Veach 1998] included in Eq. (2)

$$\alpha^L(\bar{\mathbf{x}}_{s,t}) = \frac{f^L(\bar{\mathbf{x}}_{s,t})}{p^L(\bar{\mathbf{x}}_{s,t})}, \quad \alpha^E(\bar{\mathbf{x}}_{s,t}) = \frac{f^E(\bar{\mathbf{x}}_{s,t})}{p^E(\bar{\mathbf{x}}_{s,t})}, \tag{22}$$

where $f^L(\bar{\mathbf{x}}_{s,t})$ and $f^E(\bar{\mathbf{x}}_{s,t})$ are the measurement contribution terms corresponding to the light and eye sub-paths respectively.

On the other hand, the probability density of unbiased photon gathering is defined as

$$p_{s',t'}^{\mathrm{UPG}}(\bar{\mathbf{x}}) = p^L(\bar{\mathbf{x}}_{s',t'-1}) p^c(\bar{\mathbf{x}}_{s',t'-1}) p^E(\bar{\mathbf{x}}_{s',t'-1}), \tag{23}$$

in which the connection probability p^c cannot be evaluated analytically. While an unbiased estimation is technically possible, it would not serve the purpose here, since deterministic approximation of MIS weights would not introduce bias as long as Eq. (19) is still satisfied. Hence we revert to the approximation of Eq. (7) used by UPS/VCM when computing p^c for MIS weights.

However, Eq. (7) may disagree with our real connection probability in Eq. (8). Therefore, a sufficiently large error can lead to suboptimal combination, which is most pronounced when $y_{s'}$ and $z_{t'-1}$ are spatially close. We thus propose a heuristic clamping to alleviate this problem

$$p^c(\bar{\mathbf{x}}_{s',t'-1}) \approx \min(\pi d^2 p_x(z_{t'-2} \to z_{t'-1} \to y_{s'}), 1). \tag{24}$$

The intuition behind Eq. (24) is simple. Note that p^c is a unit-less probability of a hypothetical Russian roulette event and its correct value should never exceed 1. However, p_x is a probability density per unit area and by definition $p_x(z_{t'-2} \to z_{t'-1} \to y_{s'})$

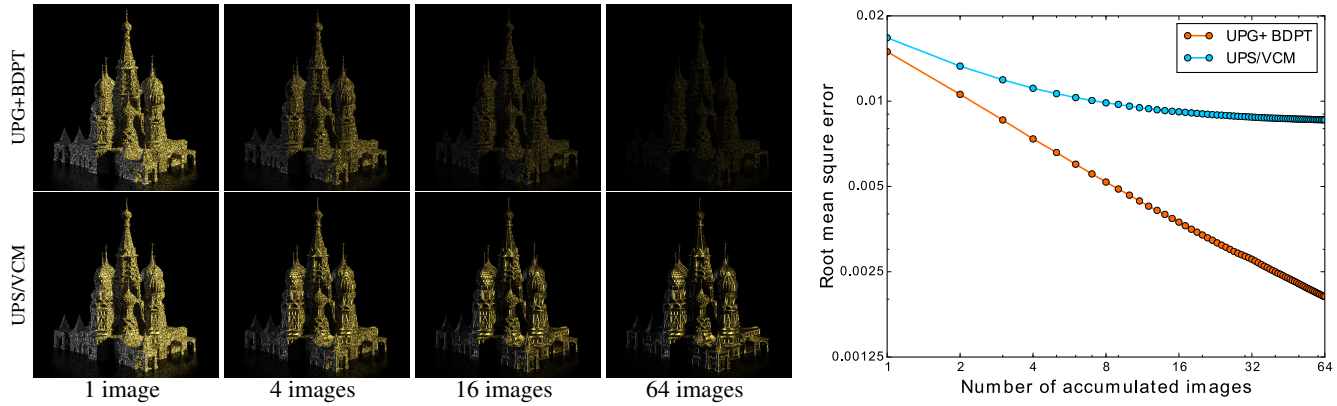


Figure 7: *Bias experiment comparing UPG+BDPT and the biased UPS/VCM. Each algorithm renders 64 images independently, then we average an increasing number of them and compare the results with the ground truth. The left side shows difference images and the right side is a log-log plot of RMSE against the number of averaged images. As illustrated, UPG+BDPT can converge this way but UPS/VCM cannot. The rendering results are provided in the supplementary material.*

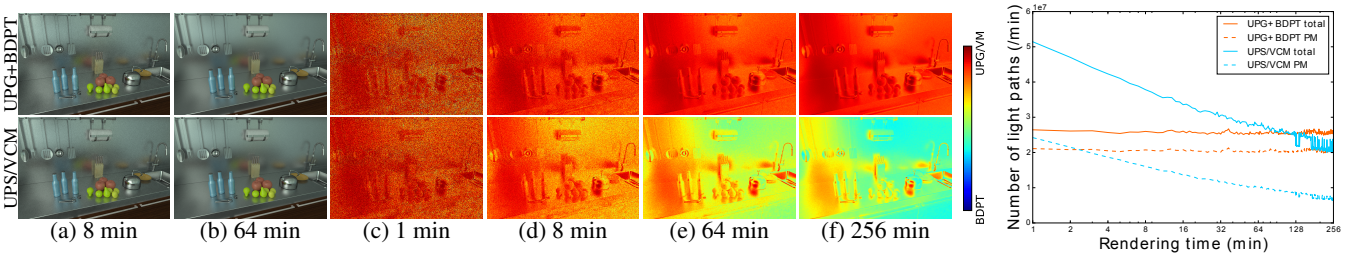


Figure 8: *Relative contribution of BDPT and PM in UPS/VCM and our UPG+BDPT. (c)-(f) visualize the relative contribution for different rendering times. We can see that the PM sampling technique consistently contributes to the image in our UPG+BDPT, while UPS/VCM's progressive shrinking radius causes the contribution to diminish. The plot on the right shows the sampling speed of the two methods as a function of time. UPS/VCM is initially more efficient, but its sampling speed diminishes over time as the more effective PM technique gradually falls out of use. On the other hand, UPG+BDPT starts slow, but its stable speed eventually takes the lead with enough computation time. Therefore, UPS/VCM provides lower estimation variance in the first several minutes but UPG+BDPT overtakes it after* 1 *hour of rendering, as shown in (a) and (b).*

is inversely proportional to $\|z_{t'-1} - y_{s'}\|^2$. By assuming p_x to be constant within a radius of d, Eq. (7) implicitly assumes that $d \ll \|z_{t'-1} - y_{s'}\|$. When the assumption fails, the resulting approximation could produce a value significantly larger than 1. Clamping the value back to 1 restricts this error, as shown in Fig. 4. The clamped result (b) computed using Eq. (24) is noticeably less noisy than the naive MIS result in (a) computed using Eq. (7).

Note that such probability clamping is not required in UPS/VCM. Their approaches use photon density estimation to approximate the path integral and produce a smooth result regardless of how small $\|z_{t'-1} - y_{s'}\|$ is. The "probability" greater than 1 implicitly makes the smoother PM result dominate the MIS combination, which is usually desirable in a biased method.

Fig. 5 illustrates the effect of combining BDPT and UPG. Note that UPG (b) converges significantly better for many-bounce indirect paths, while BDPT (a) is more efficient for diffuse direct lighting. Like in UPS/VCM, the MIS combination (c) gets the best of both methods and is robust for both caustics and direct lighting.

5 MCMC Integration

MCMC methods sample light transport by sampling a Markov Chain, the state of which typically represents one or more paths. The Metropolis-Hastings algorithm is applied to ensure that the sampling probability density of each state is proportional to a Monte Carlo estimator computed from the corresponding paths. A path sampling technique can either be added as a new type of state mutation or as an extra term in the per-state estimator. We take the later approach and build upon PSSMLT [Kelemen et al. 2002].

The classical PSSMLT state is a conceptually infinite sequence of random numbers, which is used to sample an eye sub-path with s vertices and a light sub-path with t vertices. $s \times t$ complete paths are generated by connecting sub-path prefixes as in classical BDPT. The MIS combination of the respective path integral estimators is taken as the MCMC sampling probability density. To add unbiased photon gathering, we first pre-trace a fixed set of light sub-paths as in lines 1-2 of Algorithm 1. When evaluating each PSSMLT state, we take the original eye sub-path, and execute the gathering process in lines 4-16 on all its $t - 1$ prefixes. The resulting estimators are inserted to the per-state MIS combination, which is then used in the original PSSMLT mutations. The pre-traced light sub-paths are regenerated periodically to minimize correlation artifacts.

The simplicity of combining UPG with MCMC methods stems from its unbiased nature. Adding UPG to the set of sampling techniques of PSSMLT will not affect the target distribution of the Markov Chain process. The combination would be no more than combining UPG with BDPT. While the combination of PPM with MCMC methods needs to either design another target distribution [Hachisuka and Jensen 2011] or carefully deal with the gather radius shrinkage rate [Kaplanyan and Dachsbacher 2013b].

Fig. 6 illustrates how UPG improves the PSSMLT result. Classical PSSMLT (a) does not converge very well for the many-bounce indirect lighting, which is a behavior inherited from BDPT. Combining

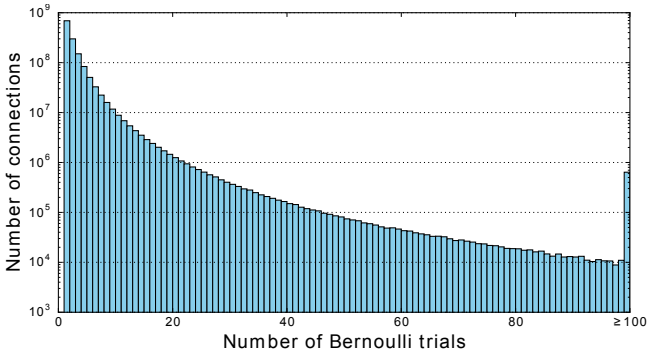


Figure 9: *A histogram of the number of per-path Bernoulli trials for Fig. 1. On average 2.81 trials are performed for each gathering. The last bin represents 100 or more trials.*

PSSMLT with UPG (b) results in a noticeable improvement. However, currently the UPG+PSSMLT (c) combination does not perform as well as UPG+BDPT, as the fixed per-state estimator does not allow flexible adjustment of relative sampling rates. Therefore, unless otherwise stated, we default to UPG+BDPT in comparisons.

6 Experimental Results

We conduct a series of experiments to validate our method as well as to compare with alternative light transport algorithms on a computer with four Intel i7 3.40GHz CPU cores and 16GB memory. All methods are implemented in the Mitsuba renderer [Jakob 2010]. We choose to compare with multiplexed Metropolis light transport (MMLT) [Hachisuka et al. 2014] as an unbiased alternative and UPS/VCM as a biased alternative.

Bias Experiment. We experimentally validate the unbiased nature of our method by averaging images generated from independent runs. As illustrated in the top-left row of Fig. 7, the estimation error of UPG+BDPT approaches zero as the number of images increases. In contrast, the consistent-yet-biased UPS/VCM does not converge to the ground truth in such a setup. The RMSE (root mean squared error) curve of UPG+BDPT shows up as a line of slope -0.5 in a log-log plot (Fig. 7 right), which follows the classical square root convergence of unbiased Monte Carlo estimate. In contrast, UPS/VCM cannot reduce the error beyond a certain threshold, which is a classical behavior of a biased algorithm. Here we should note that this experiment setup uses *independent runs* to expose bias behavior, in which UPS/VCM has to restart at a fixed initial radius when generating each new image. This should not be confused with the *parallel runs* used in practical UPS/VCM rendering, where participating machines are not fully independent – they have to communicate initially to give each a different initial radius.

Relative Contribution and Overhead. Compared with UPS/VCM, our UPG formulation does not need a progressively shrinking radius d. Therefore, in UPG the relative contribution of photon mapping does not diminish as the number of iterations increases. This is illustrated in Fig. 8. Consequently, we have more freedom when fine-tuning d, as it is purely an efficiency issue and is free of convergence-imposed constraints. In Fig. 14, d is chosen to compare with UPS/VCM at a similar level of variance.

Compared with traditional photon mapping, UPG introduces two sources of overhead. The first source is the shadow ray traced between $z_{t'-1}$ and y_s, plus the follow-up BSDF evaluations. The second source is the Bernoulli trials. The average number of Bernoulli trials per connection is approximately the same across our test

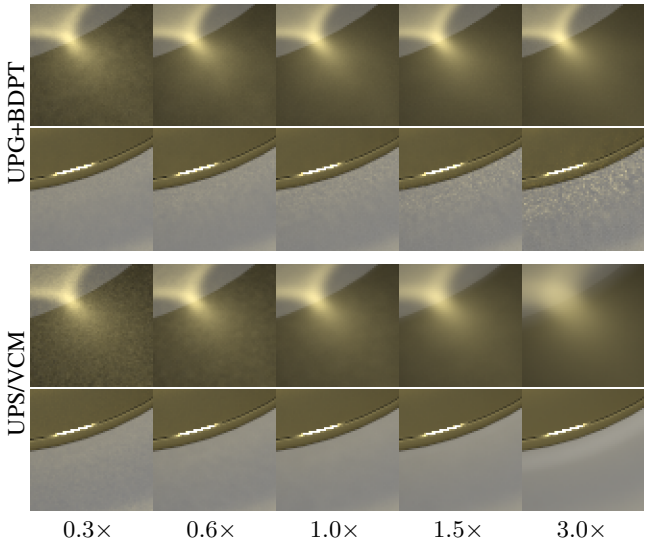


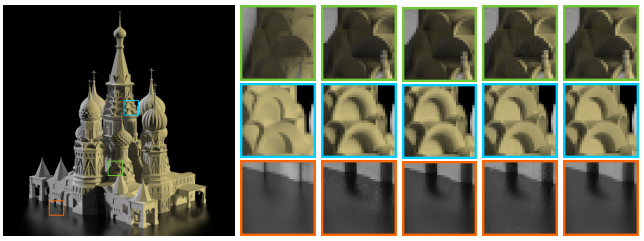
Figure 10: *UPG+BDPT and UPS/VCM rendering with different gather radiuses. All insets are chosen from Rings and are rendered with the same computation time (10min). When the gather radius increases, our algorithm puts greater emphasis on the photon mapping sampling technique, and the area mainly rendered by BDPT shows more variance. For UPS/VCM, the increasing gather radius will reduce variance and inevitably increase bias at the same time.*

scenes, ranging from 2.60 to 2.87. As illustrated in Fig. 9, most pixels of Fig. 1 require fewer than 3 trials, indicating that the angular bound reduces the complexity of Bernoulli trials from $O(n^2)$ to $O(n)$ in practice. Although there are a few long-tail cases triggered at a very low probability, the corresponding performance overhead is acceptable when amortized over the entire image.

We quantitatively evaluate the UPG overhead using the kitchen scene in Fig. 8 as an example. Compared with UPS/VCM, UPG takes approximately $4\times$ time to process each photon. However, as shown in Fig. 8 (c)-(f), the shrinking radius in UPS/VCM gradually reduces the effectiveness of photon mapping during each iteration, gradually shifting the dominance to the less efficient BDPT. Here the initial radius of UPS/VCM is $1.5\times$ greater than the radius of UPG. As shown on the right chart of Fig. 8, this allows UPS/VCM to generate light paths at a significantly higher speed than UPG+BDPT in the first few minutes, which leads to the lower estimation variance in Fig. 8 (a). As the rendering progresses, however, the sampling speed of UPS/VCM drops rapidly. On the other hand, UPG+BDPT maintains a constant speed and takes the lead after approximately 2 hours. Even well before the turning point, the cumulative effect of having more PM samples already allows UPG+BDPT to provide superior rendering quality, as shown in Fig. 8 (b). This indicates that photon gathering is a more effective sampling technique than BDPT for the kitchen scene, which is dominated by indirect illumination.

Although UPG+BDPT maintains a stable sampling speed, the speed still depends on d and some fine tuning is required for optimal rendering efficiency. A larger d generates more connections while potentially consuming more computation time on low-contribution light paths. A comparison between different values of d is shown in Fig. 10. While our method is not sensitive to small changes of d, excessively large or small values can still be problematic.

Comparisons with Biased Alternatives. Fig. 14 compares our UPG with state-of-the-art biased alternatives. As shown in the figures and zoom-ins, the UPS/VCM results exhibit bias artifacts from



Reference (a) (b) (c) (d) (e)

Figure 11: *Bias artifacts cannot be completely removed without our unbiased probability reciprocal estimation. (a)-(b) are rendered with UPS/VCM. (b) uses the correct f^c in place of f^c_{VM} but leaves disc-approximated probability untouched, and thus employs the same connection probability approximation as in [Bekaert et al. 2003]. The artifacts of biases are reduced but not removed. (c) follows up with another 50 minutes of rendering for (b), and still fails to remove the artifacts. (d) is rendered using UPG and is thus free of such artifacts. (e) is the reference. The initial radius of UPS/VCM and the gathering radius of UPG are the same. Rendering time is* 10 *minutes except for the reference. The full images are provided in the supplementary material.*

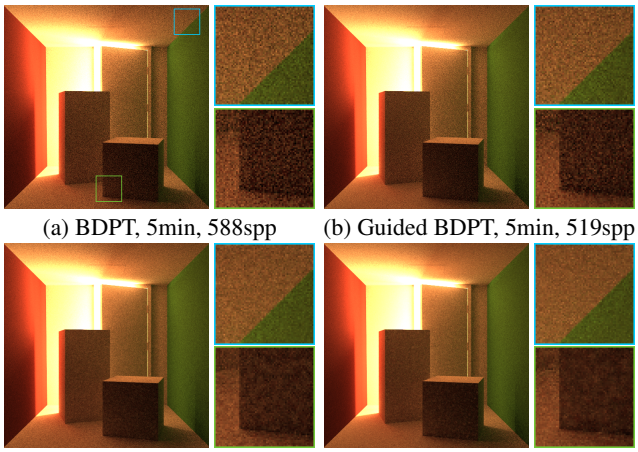
the reference images even after significant rendering time. For example, the caustics and shadows in the rings scene are blurred. The detailed geometry of the palace looks flattened. Light leaks around the wall corner in the living room scene. In the kitchen scene with many glossy reflections, the kettle reflects a highlight that does not exist in the reference result.

Unbiased connection probability reciprocal estimation is critical when removing bias artifacts. Simply replacing f^c_{VM} with f^c reduces the artifacts of incorrect evaluations of visibility and shadow. However, the incorrect probability term causes legitimate light paths to become under-weighted, which manifests as artificial darkening at sharp features or increased estimation variance, as illustrated in Fig. 11 (b). Such artifacts remain in the image even after an hour of rendering, as illustrated in Fig. 11 (c). It requires the unbiased probability reciprocal in UPG to completely eliminate such artifacts while keeping the low variance estimation of photon gathering, as illustrated in Fig. 11 (d).

Comparisons with Unbiased Alternatives. We also compare UPG+BDPT with the state-of-the-art unbiased alternative of MMLT in Fig. 14. For the same computation time, MMLT can also produce a low variance estimation. However, our method does not produce spike artifacts even with challenging paths, and generates less noise for deeper indirect lighting paths such as the living room counter. That is an advantage inherited from photon mapping.

We also compare our method with particle-guided BDPT [Vorba et al. 2014]. As shown in Fig. 12, UPG+BDPT (c) outperforms guided BDPT (b) in general. We further made a preliminary attempt of using guided BDPT for sub-path sampling in UPG+BDPT. Although our unoptimized implementation is not efficient enough for a same-time comparison yet, it demonstrates some improvements in (d) when compared with plain UPG+BDPT at the same number of paths per pixel.

Although our method does not have the low frequency filtering effect of traditional density estimation techniques, UPG still performs well for a variety of light paths including many-bounce diffuse inter-reflections, generating a smoother result when compared with other unbiased methods. This indicates that the fundamental efficiency of photon gathering goes beyond its filtering effect, and is not limited to caustics paths.



(a) BDPT, 5min, 588spp (b) Guided BDPT, 5min, 519spp

(c) UPG+BDPT, 5min, 503spp (d) Guided UPG+BDPT, 503spp

Figure 12: *Comparison between particle-guided BDPT and UPG+BDPT. (a)-(c) are rendered in five minutes. (d) is rendered with the same sampling rate as (c).*

7 Discussion and Future Work

Limitations. The most important limitation of our unbiased photon gathering is the Bernoulli trial overhead, which puts our method at a disadvantage for shallow paths that are already easy to sample. As shown in Table 1, our method is inefficient for direct lighting paths due to the higher per-sample cost. This is the low variance problem discussed by Veach [1998]. If a specific type of light transport has already been well-sampled by existing techniques, combining with more techniques would only introduce additional variance. We currently work around this issue by heuristically setting a lower sampling rate for shallow paths and disabling direct lighting paths entirely in our method. This is consistent with the practical usage of PM methods, which often recommend excluding direct lighting [Jensen 2001].

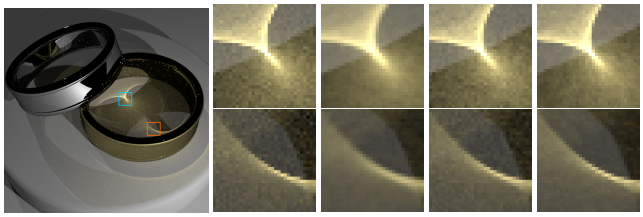
Table 1: *RMSE comparison for different path depths. This experiment is performed on the stairs scene in Fig. 14.*

Depth	$k \leq 2$	$k = 3$	$k = 4$	$k > 4$
BDPT	0.0074	0.0614	0.0733	0.1570
UPG	0.0123	0.0333	0.0310	0.0418

UPG cannot handle strict specular-diffuse-specular (SDS) paths, where the BSDF terms are Dirac delta functions. This is an open problem for unbiased path sampling which we do not solve. Existing biased estimations such as UPS/VCM approximate such specular bounces by smoothing out the sampled specular photons without fully generating the relevant light paths in a physically based manner, which cannot be easily adapted to unbiased rendering.

Another issue is that UPG is inefficient for connections with specular bounces. The problem is more pronounced for refraction effects where typical light paths have more specular bounces than non-specular bounces. If one of the sub-path endpoints to be connected is on a specular surface, the corresponding delta function BSDF evaluates to zero, which prevents the connected path from making any contribution. On the other hand, non-SDS specular paths can still be generated by making a connection at a pair of non-specular vertices.

In addition, UPG does handle highly glossy materials efficiently, which can approximate specular effects to a level of sharpness comparable with UPS/VCM. Fig. 13 compares specular caustics gen-



Reference (a) (b) (c) (d)

Figure 13: *Caustics rendering. All results are from 10 minutes of rendering except for the reference. (a) specular rings rendered with BDPT. (b)-(c) specular rings rendered with UPS/VCM. The initial radius in (c) is a half of the initial radius in (b). (d) Highly glossy rings rendered with UPG+BDPT, whose gathering radius is the same as the initial radius in (b). (c) and (d) look similar but (d) is less noisy. (b) is more blurry compared with (c) and (d). The full images are provided in the supplementary material.*

erated using a variety of methods. (a) is the reference generated using BDPT rendering, which is still noisy but illustrates the true "sharpness" of the specular caustics. In (b), the biased UPS/VCM produces a smooth-but-blurry result. A smaller initial radius improves the sharpness at the cost of a noisy image, as illustrated in Fig. 13 (c). Replacing the specular input with a highly glossy BSDF, UPG+BDPT in Fig. 13 (d) produces a sharpness similar to Fig. 13 (c) but is much less noisy.

Discussion. Accuracy and efficiency are the endless pursuit of light transport simulation. Recently, plenty of efficient methods have been proposed but accompanied with biases. The error of Monte Carlo estimate comes from both variance and bias. When variance is significantly reduced to a low level, the error naturally results from the bias. Variance and bias are like two sides of a coin, without considering each we cannot have a comprehensive understanding. Biased-but-consistent methods like PPM/SPPM and UPS/VCM strike a nice balance by providing a low variance initially and progressively diminishing the bias to zero. UPG takes another approach for which we clarify the sources of bias and remove them at a mathematical level. The elimination of bias restricts the estimation error to variance only and makes our method natively progressive. This is practically significant as parameters like d become significantly simpler to tune than in biased methods.

Photon mapping has been widely used for efficient light transport simulation for decades. Extensions like PPM/SPPM and UPS/VCM pushed the state-of-the-art for high quality rendering. However, photon mapping has always been considered a biased algorithm from the time of its introduction, with speculation that the method is intrinsically biased. UPG takes a critical step to disprove this speculation. In doing so, we decouple photon mapping from traditional biased evaluations, which fully integrates photon mapping into the family of path sampling techniques by allowing photons to be evaluated in an unbiased manner.

Future Work. It would be interesting to extend our method to volume rendering where the gathering neighborhood or virtual lights can take more flexible shapes such as rays and beams [Sun et al. 2010; Jarosz et al. 2011; Novák et al. 2012; Křivánek et al. 2014]. The radius can be also chosen adaptively and dynamically for better performance. Another interesting direction is to extend the Russian roulette formulation to light path construction mechanisms other than nearest neighborhood search, such as VPL with the probabilistic shadow test. Finally, our current MCMC integration is rather straightforward, and more advanced solutions could be developed.

Acknowledgements

We would like to thank the anonymous reviewers for their helpful comments, and Steve Lin for proofreading this paper. This research was supported in part by NSFC (No. 61272305 and No. 61472352) and the National Program for Special Support of Eminent Professionals of China.

References

BEKAERT, P., COOLS, R., AND SLUSSALEK, P. 2003. A custom designed density estimator for light transport. Tech. Rep. MPI-I-2003-4-004, Max-Planck Institut fr Informatik (Sarrebruck, DE), Saarbrcken.

BOOTH, T. E. 2007. Unbiased Monte Carlo estimation of the reciprocal of an integral. *Nuclear Science and Engineering 156*, 3, 403–407.

CHANDLER, D. 1987. *Introduction to Modern Statistical Mechanics*. Oxford University Press.

CLINE, D., TALBOT, J., AND EGBERT, P. 2005. Energy redistribution path tracing. *ACM Trans. Graph. (Proc. SIGGRAPH) 24*, 3 (July), 1186–1195.

DAVIDOVIČ, T., KŘIVÁNEK, J., HAŠAN, M., SLUSALLEK, P., AND BALA, K. 2010. Combining global and local virtual lights for detailed glossy illumination. *ACM Trans. Graph. (Proc. SIGGRAPH [Asia]) 29*, 6 (Dec.), 143:1–143:8.

GEORGIEV, I., KŘIVÁNEK, J., DAVIDOVIČ, T., AND SLUSALLEK, P. 2012. Light transport simulation with vertex connection and merging. *ACM Trans. Graph. (Proc. SIGGRAPH [Asia]) 31*, 6 (Nov.), 192:1–192:10.

HACHISUKA, T., AND JENSEN, H. W. 2009. Stochastic progressive photon mapping. *ACM Trans. Graph. (Proc. SIGGRAPH [Asia]) 28*, 5 (Dec.), 141:1–141:8.

HACHISUKA, T., AND JENSEN, H. W. 2011. Robust adaptive photon tracing using photon path visibility. *ACM Trans. Graph. 30*, 5 (Oct.), 114:1–114:11.

HACHISUKA, T., OGAKI, S., AND JENSEN, H. W. 2008. Progressive photon mapping. *ACM Trans. Graph. (Proc. SIGGRAPH [Asia]) 27*, 5 (Dec.), 130:1–130:8.

HACHISUKA, T., PANTALEONI, J., AND JENSEN, H. W. 2012. A path space extension for robust light transport simulation. *ACM Trans. Graph. (Proc. SIGGRAPH [Asia]) 31*, 6 (Nov.), 191:1–191:10.

HACHISUKA, T., KAPLANYAN, A. S., AND DACHSBACHER, C. 2014. Multiplexed Metropolis light transport. *ACM Trans. Graph. (Proc. SIGGRAPH) 33*, 4 (July), 100:1–100:10.

HAŠAN, M., KŘIVÁNEK, J., WALTER, B., AND BALA, K. 2009. Virtual spherical lights for many-light rendering of glossy scenes. *ACM Trans. Graph. (Proc. SIGGRAPH [Asia]) 28*, 5 (Dec.), 143:1–143:6.

JAKOB, W., AND MARSCHNER, S. 2012. Manifold exploration: A Markov chain Monte Carlo technique for rendering scenes with difficult specular transport. *ACM Trans. Graph. (Proc. SIGGRAPH) 31*, 4 (July), 58:1–58:13.

JAKOB, W., 2010. Mitsuba renderer. http://www.mitsuba-renderer.org.

JAROSZ, W., NOWROUZEZAHRAI, D., THOMAS, R., SLOAN, P.-P., AND ZWICKER, M. 2011. Progressive photon beams. *ACM*

Trans. Graph. (Proc. SIGGRAPH [Asia]) 30, 6 (Dec.), 181:1–181:12.

JENSEN, H. W. 2001. *Realistic Image Synthesis Using Photon Mapping*. A. K. Peters, Ltd., Natick, MA, USA.

KAJIYA, J. T. 1986. The rendering equation. *SIGGRAPH Comput. Graph. 20*, 4 (Aug.), 143–150.

KAPLANYAN, A. S., AND DACHSBACHER, C. 2013. Adaptive progressive photon mapping. *ACM Trans. Graph. 32*, 2 (Apr.), 16:1–16:13.

KAPLANYAN, A. S., AND DACHSBACHER, C. 2013. Path space regularization for holistic and robust light transport. *Computer Graphics Forum (Proc. of Eurographics 2013) 32*, 2, 63–72.

KAPLANYAN, A. S., HANIKA, J., AND DACHSBACHER, C. 2014. The natural-constraint representation of the path space for efficient light transport simulation. *ACM Trans. Graph. (Proc. SIGGRAPH) 33*, 4 (July), 102:1–102:13.

KELEMEN, C., SZIRMAY-KALOS, L., ANTAL, G., AND CSONKA, F. 2002. A simple and robust mutation strategy for the Metropolis light transport algorithm. *Comput. Graph. Forum 21*, 3, 531–540.

KELLER, A. 1997. Instant radiosity. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, SIGGRAPH '97, 49–56.

KITAOKA, S., KITAMURA, Y., AND KISHINO, F. 2009. Replica exchange light transport. *Comput. Graph. Forum 28*, 8, 2330–2342.

KNAUS, C., AND ZWICKER, M. 2011. Progressive photon mapping: A probabilistic approach. *ACM Trans. Graph. 30*, 3 (May), 25:1–25:13.

KOLLIG, T., AND KELLER, A. 2004. Illumination in the presence of weak singularities. In *Monte Carlo and Quasi-Monte Carlo Methods*, 245–257.

KŘIVÁNEK, J., GEORGIEV, I., HACHISUKA, T., VÉVODA, P., ŠIK, M., NOWROUZEZAHRAI, D., AND JAROSZ, W. 2014. Unifying points, beams, and paths in volumetric light transport simulation. *ACM Trans. Graph. (Proc. SIGGRAPH) 33*, 4 (July), 103:1–103:13.

LAFORTUNE, E. P., AND WILLEMS, Y. D. 1993. Bi-directional path tracing. In *Proceedings of CompuGraphics*, vol. 93, 145–153.

METROPOLIS, N., ROSENBLUTH, A. W., ROSENBLUTH, M. N., TELLER, A. H., AND TELLER, E. 1953. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics 21*, 6 (June).

NOVÁK, J., NOWROUZEZAHRAI, D., DACHSBACHER, C., AND JAROSZ, W. 2012. Virtual ray lights for rendering scenes with participating media. *ACM Trans. Graph. (Proc. SIGGRAPH) 31*, 4 (July), 60:1–60:11.

SEGOVIA, B., IEHL, J. C., MITANCHEY, R., AND PÉROCHE, B. 2006. Bidirectional instant radiosity. In *Proceedings of the 17th Eurographics Conference on Rendering Techniques*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, EGSR '06, 389–397.

SUN, X., ZHOU, K., LIN, S., AND GUO, B. 2010. Line space gathering for single scattering in large scenes. *ACM Trans. Graph. (Proc. SIGGRAPH) 29*, 4 (July), 54:1–54:8.

TOKUYOSHI, Y. 2009. Photon density estimation using multiple importance sampling. In *ACM SIGGRAPH ASIA 2009 Posters*, ACM, New York, NY, USA, SIGGRAPH ASIA '09, 37:1–37:1.

VEACH, E., AND GUIBAS, L. J. 1994. Bidirectional estimators for light transport. In *Proceedings of the Fifth Eurographics Workshop on Rendering*, Eurographics, 147–162.

VEACH, E., AND GUIBAS, L. J. 1997. Metropolis light transport. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, SIGGRAPH '97, 65–76.

VEACH, E. 1998. *Robust Monte Carlo Methods for Light Transport Simulation*. PhD thesis, Stanford, CA, USA. AAI9837162.

VORBA, J., KARLÍK, O., ŠIK, M., RITSCHEL, T., AND KŘIVÁNEK, J. 2014. On-line learning of parametric mixture models for light transport simulation. *ACM Trans. Graph. (Proc. SIGGRAPH) 33*, 4 (July), 101:1–101:11.

VORBA, J. 2011. Bidirectional photon mapping. In *Proc. of the Central European Seminar on Computer Graphics (CESCG 11)*.

WALTER, B., FERNANDEZ, S., ARBREE, A., BALA, K., DONIKIAN, M., AND GREENBERG, D. P. 2005. Lightcuts: A scalable approach to illumination. *ACM Trans. Graph. (Proc. SIGGRAPH) 24*, 3 (July), 1098–1107.

WALTER, B., ARBREE, A., BALA, K., AND GREENBERG, D. P. 2006. Multidimensional lightcuts. *ACM Trans. Graph. (Proc. SIGGRAPH) 25*, 3 (July), 1081–1088.

WALTER, B., KHUNGURN, P., AND BALA, K. 2012. Bidirectional lightcuts. *ACM Trans. Graph. (Proc. SIGGRAPH) 31*, 4 (July), 59:1–59:11.

A Variance Reduction with Angular Bound

The angular bound introduced in Sec. 3.4 also reduces the estimation variance of the probability reciprocal r. Without the angular bound, the estimation variance can be computed based on Eq. (11) and Eq. (12). Since all relevant terms are computed for the same path $\bar{\mathbf{x}}_{s',t'-1}$, we drop the function notations for brevity.

$$\begin{aligned} V[N] &= E\left[N^2\right] - E^2[N] \\ &= \sum_{i=1}^{+\infty} i^2 p^c (1-p^c)^{i-1} - (\frac{1}{p^c})^2 = \frac{1-p^c}{(p^c)^2}. \end{aligned} \tag{25}$$

The angular bound factors p^b out of N as N^b/p^b, the variance of which can be also evaluated based on Eq. (25)

$$\begin{aligned} V\left[\frac{N^b}{p^b}\right] &= \frac{1}{(p^b)^2} V\left[N^b\right] \\ &= \frac{1}{(p^b)^2} \frac{1-\frac{p^c}{p^b}}{\left(\frac{p^c}{p^b}\right)^2} = \frac{1-\frac{p^c}{p^b}}{(p^c)^2}. \end{aligned} \tag{26}$$

Note that if we ignore the shared $(p^c)^2$ term, the remaining $1-p^c$ and $1-p^c/p^b$ terms are the respective rejection probabilities of the Russian roulette process. Applying a tight angular bound usually allows most sampled tentative rays to hit the gathering neighborhood successfully. This reduces the rejection probability $1-p^c/p^b$, and by extension the estimation variance, to nearly zero. In practice, the bound is constrained by the analytical integration requirement and is not always tight. We found that without bounding, the

original rejection probability $1 - p^c$ is almost 1 as p^c is usually in the range between 10^{-4} and 10^{-3}, deduced from the number of required Bernoulli trials. With bounding, the probability $1 - p^c/p^b$ can be reduced to less than 0.4, which corresponds to a more than 60% reduction of variance.

B Alternative Connection Candidate

Given two sub-paths $y_1 \cdots y_{s'}$ and $z_{t'} \cdots z_1$ paired by photon gathering, generating a full path requires removing one of the sub-path endpoints (i.e., $y_{s'}$ or $z_{t'}$). There are two candidates for removal, which would generate two different paths $\bar{\mathbf{x}}_{s'-1,t'}$ and $\bar{\mathbf{x}}_{s',t'-1}$. However, the two path candidates are spatially close and they are unlikely to make a significantly different contribution to the final image. Therefore, as a cost-motivated optimization, we only evaluate one of them and discard the other.

Since the bottleneck in evaluating a path is the Bernoulli trials, we do not want the selection criterion to depend on p^c. In addition, we prefer a deterministic criterion, as correlating the final Monte Carlo estimate with dependent random variables may bias the result. Finally, MIS-based theories do not apply here, as fundamentally this is a decision between different paths, not a decision between different sampling techniques creating the same path. Therefore, we elect to make the choice heuristically.

Note that $y_{s'}$ is importance-sampled from $f_s(y_{s'-2} \to y_{s'-1} \to y_{s'})$ and $z'_{t'}$ is importance-sampled from $f_s(z_{t'-2} \to z_{t'-1} \to z_{t'})$. Intuitively, removing one of the two endpoints corresponds to reconnecting the corresponding importance-sampled edge to a slightly shifted location. One could formulate the resulting change as

$$C^*_{RR}(\bar{\mathbf{x}}_{s',t'-1}) \propto \frac{f_s(z_{t'-2} \to z_{t'-1} \to y_{s'})}{f_s(z_{t'-2} \to z_{t'-1} \to z_{t'})}, \tag{27}$$

$$C^*_{RR}(\bar{\mathbf{x}}_{s'-1,t'}) \propto \frac{f_s(y_{s'-2} \to y_{s'-1} \to z_{t'})}{f_s(y_{s'-2} \to y_{s'-1} \to y_{s'})}. \tag{28}$$

Since a less glossy BSDF is less likely to result in a significant change in C^*_{RR}, we heuristically test the surface properties at $y_{s'-1}$ and $z_{t'-1}$, then we remove the endpoint corresponding to the surface with the lower glossiness. If $y_{s'-1}$ is the light source, we assign an effective glossiness based on its type, which is positive infinity for directional and point sources, or a low value comparable to diffuse surfaces for area light sources. If $z_{t'-1}$ is the camera, we assign a high effective glossiness corresponding to a surface that produces mirror-like reflections. As shown in Table 2, our heuristic significantly reduced the overall variance level. Note that the high effective glossiness for the camera is required to exclude paths that directly connect the camera to a photon on a diffuse surface, which is the main source of error for $\bar{\mathbf{x}}_{s',t'-1}$-type connections. Also note that the probability to sample the abandoned path is still non-zero because it can be generated by connecting $y_1 \cdots y_{s'-1} z_{t'}$ and $y_{s'} z_{t'-1} \cdots z_1$ instead.

Table 2: *RMSE comparison for different connection choices at the same computational time. The experiment is performed on the living room scene in Fig. 14. Note that evaluating both connections produced a worse result due to the computational cost wasted on less effective candidates.*

$\bar{\mathbf{x}}_{s',t'-1}$ only	$\bar{\mathbf{x}}_{s'-1,t'}$ only	Average both	Our heuristics
0.0438	0.0172	0.0274	0.0148

C Reciprocal of an Integral

In this section, we briefly introduce the unbiased algorithm for estimating the reciprocal of a general integral proposed by Booth [2007], and discuss its relationship with our method proposed in Sec. 3.4.

Given a non-negative function $f(x)$ defined in a domain Ω, Booth [2007] aimed to evaluate the reciprocal of its integral

$$I = \frac{1}{\int_\Omega f(x)\mathrm{d}x}. \tag{29}$$

Assuming $\int_\Omega f(x)\mathrm{d}x \in (0, 1)$, Eq. (29) can be evaluated as a series expansion:

$$I = \frac{1}{1-g} = \sum_{i=0}^{+\infty} g^i, \tag{30}$$
$$g = 1 - \int_\Omega f(x)\mathrm{d}x.$$

Reorganizing the infinite series and taking an independent Monte Carlo estimation for each occurrence of g, we have

$$\hat{I} = 1 + \sum_{i=1}^{+\infty} \prod_{j=1}^{i} \hat{g}_j, \tag{31}$$
$$\hat{g}_j = 1 - f(x_j),$$

where $\{x_j\}$ is an infinite set of stochastic samples in Ω, and we assume $\int_\Omega 1\mathrm{d}x = 1$.

The series can be terminated using Russian roulette once $\prod_{j=1}^{i} \hat{g}_j$ falls below a pre-specified threshold. However, in our case, f is the binary outcome of a Bernoulli trial, and $\hat{g}_j$ always takes a value of 0 or 1. Therefore, the series in Eq. (31) degenerates into a counting process. Assuming the series of Bernoulli trials succeeds for the first time after N tries, we have:

$$\hat{g}_j = \begin{cases} 1, & j < N \\ 0, & j \geq N, \end{cases}, \tag{32}$$
$$\hat{I} = 1 + \sum_{i=1}^{N} \prod_{j=1}^{i} \hat{g}_j = N,$$

which is identical to our algorithm in Sec. 3.4.

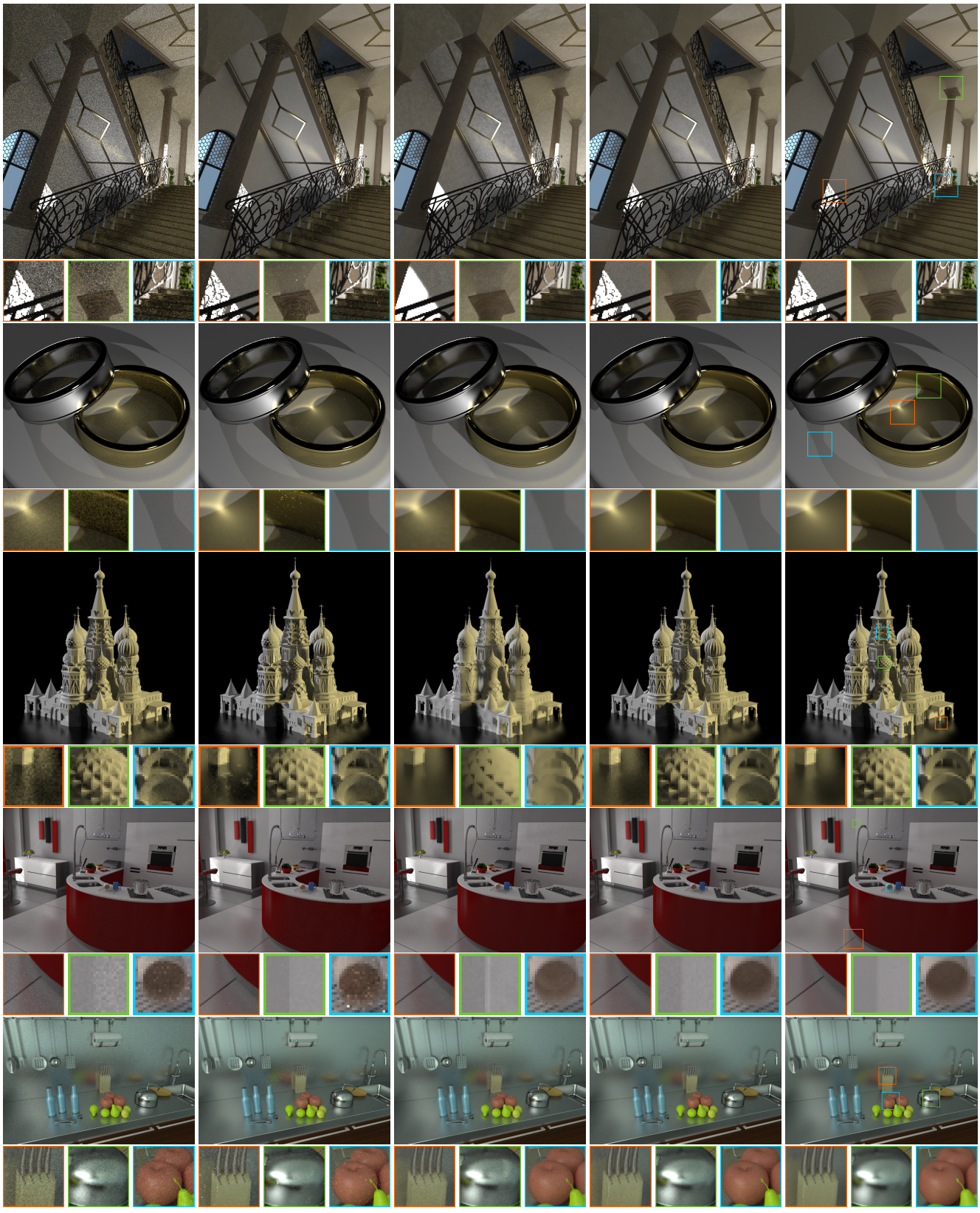


Figure 14: *Same-time quality comparison between different methods. The stairs scene in the first row is primarily illuminated by many-bounce indirect lighting. The rings scene in the second row is a classical caustics setup. The palace has detailed geometry and produces intricate occlusions and inter-reflections. The living room scene in the third row combines many glossy objects with many-bounce lighting. The kitchen scene is dominated by glossy-glossy inter-reflection. UPS/VCM and UPG can take advantage of photon mapping, and achieve a lower variance level than classical BDPT. MMLT delivers an overall variance level comparable to ours, but challenging light paths create bright spike artifacts. Our method has a noticeable advantage for deep light paths and glossy materials. UPS/VCM also achieves a variance level comparable with our algorithm, but it over-blurs thin features and leaks light at high frequency occlusions. The rendering times of the 5 scenes except for the reference are 1 hour, 10 minutes, 10 minutes, 1 hour and 30 minutes respectively. The corresponding ratios of radius between UPS/VCM and UPG are* $1\times$, $1.25\times$, $1\times$, $1\times$ *and* $1.5\times$ *respectively.*