Hypothesis Testing for Progressive Kernel Estimation and VCM Framework

Zehui Lin, Chenxiao Hu, Jinzhu Jia, Sheng Li*, *Member, IEEE*

Abstract—Identifying an appropriate radius for unbiased kernel estimation is crucial for the efficiency of radiance estimation. However, determining both the radius and unbiasedness still faces big challenges. In this paper, we first propose a statistical model of photon samples and associated contributions for progressive kernel estimation, under which the kernel estimation is unbiased if the null hypothesis of this statistical model stands. Then, we present a method to decide whether to reject the null hypothesis about the statistical population (i.e., photon samples) by the F-test in the Analysis of Variance. Hereby, we implement a progressive photon mapping (PPM) algorithm, wherein the kernel radius is determined by this hypothesis test for unbiased radiance estimation. Secondly, we propose VCM+, a reinforcement of Vertex Connection and Merging (VCM), and derive its theoretically unbiased formulation. VCM+ combines hypothesis testing-based PPM with bidirectional path tracing (BDPT) via multiple importance sampling (MIS), wherein our kernel radius can leverage the contributions from PPM and BDPT. We test our new algorithms, improved PPM and VCM+, on diverse scenarios with different lighting settings. The experimental results demonstrate that our method can alleviate light leaks and visual blur artifacts of prior radiance estimate algorithms. We also evaluate the asymptotic performance of our approach and observe an overall improvement over the baseline in all testing scenarios.

Index Terms—progressive photon mapping, vertex connection and merging, bidirectional path sampling, kernel estimation, statistical model, hypothesis testing, F-test, analysis of variance, radius

✦

1 INTRODUCTION

L IGHT transport simulation has been well-studied in photo-realistic rendering and applied to a wide range of digital entertainment industries. The key challenge is to efficiently synthesize a photo-realistic image with various lighting conditions. Considerable advances have been made in the past decades to improve the efficiency of light transport algorithms. Bidirectional path tracing (BDPT) [1], [2] is one of the most successful algorithms. However, it struggles with caustics and other lighting effects caused by specular-diffuse-specular (S-D-S) paths [3]. Photon mapping (PM) [4] can handle S-D-S paths by emitting many photons from light sources and performing density estimation on diffuse surfaces but with bias induced. A large kernel reduces variance by collecting more photons but potentially brings more bias. Progressive photon mapping (PPM) [3] runs multiple iterations and gradually shrinks the support radius of the kernel to alleviate the bias progressively. However, the convergence rate of PPM (optimally $O(N^{-2/3})$) is slower than that of BDPT ($O(N^{-1})$).

Chi-squared progressive photon mapping (CPPM) [5] improves the theoretical convergence rate under ideal conditions by employing hypothesis testing to verify the unbiasedness of kernel estimation. Based on a theory that uniformly distributed photons lead to unbiased estimates, CPPM uses the χ^2-test to verify whether the photons within the kernel are subject to a uniform distribution.

CPPM shrinks the kernel radius only when it finds its test statistics are inconsistent with the hypothesized *statistical model*, where the statistical model usually refers to a set of statistical assumptions concerning the generation of sample data. However, CPPM's statistical assumption about this model is frequently violated in many cases, thus computing inaccurate radii and leading to biased or noisy results.

Combining BDPT with PM technique via multiple importance sampling (MIS) to obtain advantages from both sides, Vertex Connection and Merging (VCM) [6] or Unified Path Sampling (UPS) [7] is widely accepted as a robust and efficient solution for light transport simulation with respect to various lighting conditions. VCM/UPS employs bidirectional path sampling strategy and expresses the probability density functions of BDPT and PPM w.r.t. the same measure. However, PPM will gradually become a short slab of VCM/UPS over iterations and will consequently slow down the overall performance [6]. This is because their MIS weight for PPM is subject to the radius for kernel estimation. Whereas, as the kernel radius of PPM converges to zero, the MIS weights will gradually shrink to zero due to PPM's higher variance than that of BDPT. Consequently, BDPT gradually takes a dominant role in rendering (see Figure 2), and the computational overhead of PPM becomes increasingly worthless, i.e., the resulting algorithm is nearly equivalent to BDPT as the kernel shrinks. This is a key issue of VCM, as pointed out in Figure 7 of VCM [6].

Our method is mainly dedicated to resolving the efficiency problem in these kernel estimation related methods. Identifying an appropriate kernel radius is crucial for both the efficiency of kernel estimation of PPM and the performance of VCM/UPS. As mentioned above, a larger radius tends to induce more bias meanwhile a smaller one may

- *Zehui Lin, Chenxiao Hu, Sheng Li are with the School of Computer Science, Peking University, China.*
  E-mail: {zehui|hineven|lisheng}@pku.edu.cn.
  Jinzhu Jia is with the Dept. of Biostatistics and Center for Statistical Science, Peking University, China.
  Email: jzjia@math.pku.edu.cn.
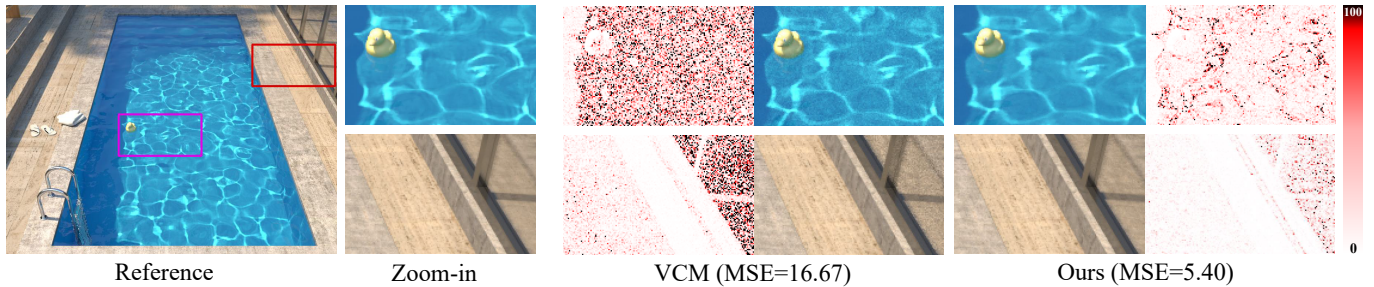- *Sheng Li is the corresponding author.*

Fig. 1: Equal-iteration (10K iterations) comparison on Pool scene rendered by VCM ($2.51h$) and our VCM+ ($2.57h$). Our algorithm is less noisy and has a lower MSE (mean squared error), as shown in the zoom-in regions with error visualization.

lead to higher variance and performance downgrade. We aim to maximize the radius that can minimize variance while keeping the result unbiased. We begin with a large kernel radius and shrink it when the kernel estimation is detected as biased. To determine whether the estimation is biased by photon samples within the kernel, we employ a parametric test that focuses on analyzing and comparing the mean and variance of observations (photon samples), i.e., quantitative analyzing the feature of irradiance distributed on surfaces. In this regard, we propose an efficient statistical testing-based approach to select the kernel radius for unbiased estimation, which tackles the limitation of CPPM. Furthermore, we extend our statistical model with a theoretical formulation applicable to the VCM framework. By hypothesis testing based on this model to help reduce the bias and variance and then optimize the contribution from vertex merging (PPM), our method eventually improves the efficiency of VCM in dealing with diverse scenarios with different lighting settings. We highlight our results and compare our method against VCM, as shown in Figure 1.

To summarize, our main contributions are two-fold:

- We propose a statistical model for hypothesis testing on photon samples along with a testing-based progressive kernel estimation algorithm, which employs the Analysis of Variance (ANOVA) F-test to decide whether the observations follow the null hypothesis of the statistical model or not, and therefore help find an . Our algorithm is more general in handling diverse photon distributions and can synthesize images with better quality.
- In addition, we propose a hypothesis testing-based formulation for the VCM/UPS framework, named VCM+, in which we derive a sufficient condition for a theoretically unbiased estimate. As a reinforcement of VCM, VCM+ can theoretically reach $O(N^{-1})$ convergence. Our algorithm improves the overall performance, wherein PPM can contribute more to the final pixel measurement in those light paths where PPM can handle better than BDPT.

2 RELATED WORK

2.1 Path Tracing

As a classic solution to the rendering equation, the path tracing (PT) algorithm [8] traces random paths from the eye to find light sources. Bidirectional path tracing (BDPT)

[2] traces eye sub-paths from the eye and light sub-paths from light sources synchronously and connects these paths via multiple importance sampling (MIS). Path tracing, as well as BDPT, provides an unbiased estimator of the pixel measurement. BDPT can efficiently sample a wide range of paths, but its ability to handle S-D-S paths is inferior to photon mapping [3]. BDPT also provides a bidirectional path sampling framework in which one eye sub-path and one light sub-path will construct $l - 1$ samples for the path with length l.

2.2 Photon Mapping

Photon mapping (PM) [4] provides an alternative solution but with bias introduced. It also traces eye sub-paths from the eye and light sub-paths from light sources. Unlike BDPT, it first samples a large number of light sub-paths and stores the light vertices (a.k.a photons) on the paths into a range searching data structure (a.k.a photon map). Then it estimates radiance at eye vertices with density estimation using nearby light vertices [9]. Some other methods leave the kernel function simple and lightweight and choose its support radius wisely, such as adaptive progressive photon mapping [10] that leverages the bias and variance with theoretically optimal support radii. Recently, the neural network was employed to predict a kernel function for density estimation [11].

Progressive photon mapping (PPM) [3] diminishes the bias by gradually reducing the support radius of the kernel. Both PM and PPM work with the unidirectional path sampling scheme, i.e., one eye sub-path and one light sub-path construct one sample for each path length. From then onwards, many studies have followed this progressive line. Stochastic progressive photon mapping (SPPM) [12] enables PPM to render more lighting effects like depth-of-field, motion blur, and glossy reflections. Vorba [13] extended PM and presented bidirectional photon mapping that combines the radiance estimated from different vertices on an eye sub-path together via MIS. Although PPM can converge to the correct result, it has a lower convergence rate than BDPT because shrinking the radius leads to fewer samples and resultant higher variance. Hachisuka et al. [14] presented a progressive error estimation framework for PPM. Knaus and Zwicker [15] presented a probabilistic analysis of PPM.

Lin et al. [5] proposed chi-squared progressive photon mapping (CPPM). It gives a better estimation of support radius that minimizes bias and variance with chi-squared

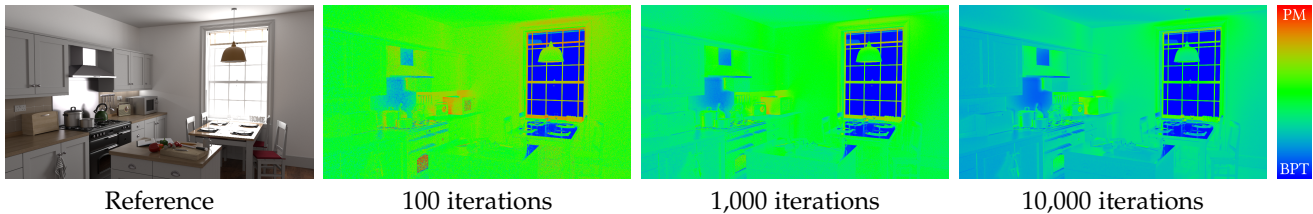| Reference | 100 iterations | 1,000 iterations | 10,000 iterations |

Fig. 2: Contribution visualization of BDPT and PM in VCM, respectively, taking Kitchen scene as an example. The relative contribution from PM gradually diminishes over iterations, and drags down VCM's overall performance consequently.

tests. An assumption about the correlation between photon distribution and radiance distribution is required, and this assumption is further relaxed as the *location-independent assumption (abbr. independence assumption)* that the contribution of a light vertex is independent of the relative position to the eye vertex, wherein a full path is constructed by this light vertex and eye vertex. χ^2-test gives good estimations of support radius when this assumption holds. However, this independence assumption may still be violated in practice. χ^2-test malfunctions in those circumstances, resulting in such artifacts as noise, light leaks, and visual blur. result. To improve the generality and thereby promote efficiency, we investigate a new model for hypothesis testing that removes this less commonly held assumption.

2.3 Vertex Connection and Merging

Vertex Connection and Merging (VCM) [6] and Unified Path Sampling (UPS) [7] are bidirectional frameworks that both combine the BDPT and PM techniques by putting their probability density functions (pdf) under the same measure. MIS leverages the relative contribution between BDPT and PPM to minimize the variance of this combined estimator. However, the variance of PPM gradually increases with the kernel shrinking. Whereas, the variance of BDPT will eventually be much lower than that of PPM for the paths that BDPT and PPM can both sample. Consequently, the MIS weight for PPM gradually reduces to zero, making the rendering less efficient. Because VCM and UPS share the same key technical features, we use VCM as a proxy of VCM/UPS for simplicity in the rest sections.

Thereafter, unbiased photon gathering [16] based on the bidirectional framework was proposed to process each photon individually to create an unbiased path sample. Still, it cannot handle S-D-S paths well. As an orthogonal technique to the VCM, Markov chain Monte Carlo (MCMC) methods were combined with VCM to improve the rendering efficiency [17].

As a combination of BDPT and PPM, reducing the variance from PPM would reduce the overall variance of VCM. Our algorithm expects to reduce the variance brought by the inappropriate radius of kernel estimation, elevate the contribution from PM with low variance and present an improved estimator under the VCM framework, thereby speeding up the rendering process.

2.4 Statistical Hypothesis Testing

Statistical hypothesis testing is a method of statistical inference [18]. It first sets up a *null hypothesis* that is testable based on observations. Hypothesis testing requires constructing a *statistical model* of the shape the data would be in. The hypothesis that the statistical divergence of the observations is due to chance alone is called the *null hypothesis*. Next, it computes the test statistic from these observations. If the resultant statistic is unlikely to occur when the null hypothesis is true (according to a predetermined threshold probability, i.e., a significance level), it rejects the null hypothesis. It's possible for a statistical hypothesis testing to incorrectly reject a true null hypothesis or incorrectly not reject a false null hypothesis. The probability of the former event is controlled by the significance level and is usually equal to a scalar value of the significance level. For the latter, increasing the sample size is always a feasible solution to decrease its probability with a fixed significance level [19].

CPPM [5] assesses whether the photons (or light vertices) are uniformly distributed by Pearson's χ^2-test, which can be used to evaluate the difference between sets of categorical data and a hypothetical frequency distribution [20]. ANOVA [21] developed by Ronald Fisher [22] is a collection of statistical tools that partitions the observed variance into different sources to analyze the differences of means. The ANOVA F-test is used to assess whether the expected values of a quantitative variable within several groups differ from each other. Therefore, it can be used to assess whether the radiance (or the contribution of full paths) is uniform within the kernel, so it can work for situations where the independence assumption is invalid. Compared with non-parametric ANOVA alternatives such as Kruskal–Wallis test [23], which requires ranking the samples, the F-test is more lightweight. We propose a null hypothesis for unbiased estimation concerning the contribution of full paths, and then use the Analysis of variance (ANOVA) F-test for hypothesis testing.

3 HYPOTHESIS TESTING FOR KERNEL ESTIMATION

Quantitative analysis of the feature of irradiance distributed on surfaces can help improve the efficiency of radiance estimation. CPPM focuses on analyzing the distribution of photon samples, and its statistical model assumes independence of the populations (i.e., photon samples) and uses the χ^2-test to verify the unbiasedness [5]. However, in practical scenarios, various lighting conditions such as multiple area light, chromatic light/textured light, and glossy reflection may vary the contribution of each sample during light transport, introducing correlations between the position and the contribution of photons, as illustrated in Figure 3. This violates the independence assumption and leads to biases in some local regions, resulting in blurry images, as shown
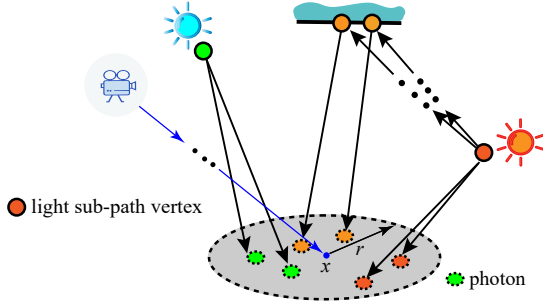
Fig. 3: Illustration of photon samples with different contributions within a searching area. In general, different light sources or light sub-paths during transport varies the associated contribution of each sample. These photons should not be treated equally.

in Figure 4 (left). Meanwhile, CPPM tends to underestimate the kernel radius for radiance estimation, leading to noisy results frequently.

Our statistical model is designed to eliminate the independence assumption in kernel estimation; herein, our approach can guarantee unbiasedness when a sufficient condition is satisfied, regardless of the scene settings. The key to achieving this goal is to identify a statistical model with generality and its corresponding unbiased conditions for kernel estimation (as detailed in Sec. 3.1), as well as a method to verify the unbiasedness of radiance within a kernel radius (as described in Sec. 3.2). Based upon the proposed statistical model (Sec. 3.1) as a theoretical foundation, our hypothesis testing methodology works as follows: firstly, the irradiance mean values for various regions can be calculated using the samples within a given kernel; next, hypothesis testing is employed to analyze whether there is a statistically significant difference in irradiance values among these regions; lastly, whether the estimation is unbiased or not can be determined based on the statistical results of difference analysis. The ANOVA F-test is employed for hypothesis testing, and the principle behind verifying unbiasedness will be explained in detail in Sec. 3.2.

3.1 Statistical Model for Unbiased Estimation in PPM

Given that the accuracy of any conclusion drawn from a statistical inference depends on the validity of the *statistical assumptions*, the assumptions are expected to be met practically. We derive a formulation without extra assumptions that can support the hypothesis of our statistical model as follows.

We start with the integral expression of radiance I in PM by Kaplanyan and Dachsbacher [10]:

$$I \approx \int_{\mathcal{M} \times \mathcal{M}} K_r(\mathbf{x} - \mathbf{x}^*) \Psi(\mathbf{x}, \mathbf{x}^*) d\mathbf{x} d\mathbf{x}^*, \qquad (1)$$

where K_r is a 2D normalized kernel function with support radius r; $\mathcal{M}$ is the scene surface; and $\Psi(\mathbf{x}, \mathbf{x}^*)$ is the contribution of all possible full paths constructed by the eye vertex $\mathbf{x}$ and the light vertex (i.e., photon) $\mathbf{x}^*$.

To align the light vertices from multiple passes associated with different eye vertices, we map these vertices to a
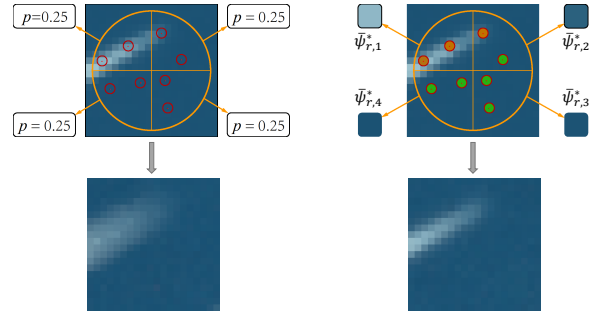

Fig. 4: Illustration of a simple case that the independence assumption is violated. The photons have different contributions within a kernel (upper-right). Here, photons on the top are likely to have more contributions. However, CPPM detects no bias through χ^2-test due to equal contribution assumption (left), whereas our method observes variational contributions across photons and detects bias through F-test for clearer result (right).

unified 2D space similar to CPPM [5], and the map function and its inverse function are

$$\begin{cases} f_{\text{map},\mathbf{x}}(\mathbf{x}^*) = (\langle \mathbf{x} - \mathbf{x}^*, \mathbf{u}(\mathbf{x}) \rangle, \langle \mathbf{x} - \mathbf{x}^*, \mathbf{v}(\mathbf{x}) \rangle) \\ f_{\text{map},\mathbf{x}}^{-1}(\mathbf{y}) = \mathbf{x} + \mathbf{u}(\mathbf{x})\mathbf{y}_u + \mathbf{v}(\mathbf{x})\mathbf{y}_v \end{cases}, \quad (2)$$

where $\mathbf{x}$ is the location of an eye vertex; $\mathbf{x}^*$ is a light vertex in the scene space; $\mathbf{u}(\mathbf{x})$ and $\mathbf{v}(\mathbf{x})$ are two orthogonal directions on the tangent plane at $\mathbf{x}$; $\langle \cdot, \cdot \rangle$ is the inner product; $\mathbf{y} = (\mathbf{y}_u, \mathbf{y}_v)$ is a point in the unified 2D space. In fact, this mapping aligns multiple gather points onto the unified space, and thus it adapts to multiple eye paths.

Then, one of the domains of integration in Equation 1 can be substituted by the unified space as

$$I \approx \int_{\mathcal{M} \times \Omega_r} K_r(\mathbf{y}) \Psi(\mathbf{x}, f_{\text{map},\mathbf{x}}^{-1}(\mathbf{y})) d\mathbf{x} d\mathbf{y}$$

$$\approx \int_{\Omega_r} K_r(\mathbf{y}) \int_{\mathcal{M}} \Psi(\mathbf{x}, f_{\text{map},\mathbf{x}}^{-1}(\mathbf{y})) d\mathbf{x} \, d\mathbf{y}, \qquad (3)$$

where Ω_r is the region within a radius r from the origin in this unified 2D space. For simplicity, we define a contribution function $\Psi^*(\mathbf{y})$ in the unified space, which is

$$\Psi^*(\mathbf{y}) = \int_{\mathcal{M}} \Psi(\mathbf{x}, f_{\text{map},\mathbf{x}}^{-1}(\mathbf{y})) d\mathbf{x} .$$

Then we can simplify the estimation as

$$I \approx \int_{\Omega_r} K_r(\mathbf{y}) \Psi^*(\mathbf{y}) d\mathbf{y} .$$

According to the principle of kernel estimation of PM, if the kernel function K_r is substituted by a Dirac delta function, I in Equation 3 will be unbiased and

$$I = \Psi^*(\mathbf{0}) \approx \int_{\Omega_r} K_r(\mathbf{y}) \Psi^*(\mathbf{y}) d\mathbf{y} .$$

Thereby, we propose the unbiased condition we hope to hold, which is a simple and sufficient condition for unbiased kernel estimation:

$$\forall \mathbf{y} \in \Omega_r, \Psi^*(\mathbf{y}) \equiv \Psi^*(\mathbf{0}) . \qquad (4)$$

That is, the contribution function $\Psi^*(\mathbf{y})$ is a constant function within Ω_r. The intuition behind the condition is that

the estimate is unbiased when the illumination is evenly distributed within the disk Ω_r. Therefore, the null hypothesis for our PPM algorithm is that the unbiased condition mentioned above holds. In other words, the observations of the contribution function are subject to a constant function, which can be tested using our hypothesis testing method in the following Sec. 3.2.

Compared our unbiased condition (Equation 4) with the one used in CPPM that the samples of $\mathbf{y}$ are assumed to be uniformly distributed in Ω_r, we can infer that CPPM actually relies on the correlation of the contribution function $\Psi^*(\mathbf{y})$ and the distribution of $\mathbf{y}$. Therefore, its high efficiency benefits from the scene settings and the sampling strategy of $\mathbf{y}$, and it may fail if the distribution of $\mathbf{y}$ is inconsistent with the contribution function $\Psi^*(\mathbf{y})$. The independence assumption can be removed. In contrast, this condition used to verify unbiasedness will be more reliable than CPPM. The estimate will be theoretically unbiased when the observations follow our proposed statistical model. Note that the samples need not be uniformly distributed for unbiased estimation with regard to our statistical model.

3.2 Hypothesis Testing to Verify Unbiasedness

A lightweight hypothesis testing method is required to infer the properties of samples based on our statistical model for kernel radiance estimate. In this subsection, we present a method to compute the test statistics within a kernel radius and obtain the inference through hypothesis testing. Therefore, an appropriate radius can be determined by shrinking the radius and testing over iterations until the observations within this radius are deemed unbiased.

3.2.1 Domain Partition and Integration

The hypothesis test requires grouping the observations into categories, i.e., the random samples should be classified into several mutually exclusive classes. Accordingly, we can specify the equal-areal sectors as categories, and their associated light vertices being collected are thereby grouped. By partitioning Ω_r into sectors and assuming that $\Psi^*(\mathbf{y})$ is constant in any sector, a straightforward solution for hypothesis testing is to gather samples of $\Psi^*(\mathbf{y})$ in each sector and compare each sector's sample mean. However, a sector's sample mean will be unavailable when it receives no light vertex, which may frequently occur after sample collection during each pass of photon mapping. We notice that the kernel estimation in photon mapping based on Monte Carlo integration can also be expressed as sample means. Therefore, we use the samples of Monte Carlo integration in estimating the integration of $\Psi^*(\mathbf{y})$ over each sector instead of the grouped light vertices in the statistical tests to compare these sectors' integration.

We first partition Ω_r into n sectors $\Omega_{r,1} \ldots \Omega_{r,n}$. The integration of $\Psi^*(\mathbf{y})$ over sector $\Omega_{r,i}$, i.e. $\int_{\Omega_{r,i}} \Psi^*(\mathbf{y}) d\mathbf{y}$, can be written as

$$\int_{\Omega_{r,i}} \Psi^*(\mathbf{y}) d\mathbf{y} =$$
$$\int_{\mathcal{M}^2} \Psi(\mathbf{x}, \mathbf{x}^*) \mathbf{1}_{\Omega_{r,i}}(f_{\mathrm{map},\mathbf{x}}(\mathbf{x}^*)) dA(\mathbf{x}) dA(\mathbf{x}^*) , \quad (5)$$

where $\mathbf{1}_{\Omega_{r,i}}(\mathbf{x})$ is the indicator function whose value is 1 if $\mathbf{x}$ is in $\Omega_{r,i}$ and 0 otherwise. Equation 5 shows that the integral of $\Psi^*(\mathbf{y})$ over $\Omega_{r,i}$ can be estimated like photon mapping as

$$\int_{\Omega_{r,i}} \Psi^*(\mathbf{y}) d\mathbf{y} \approx \overline{\psi}_{r,i}^* \approx \frac{1}{m} \sum_{j=1}^{m} \psi_{r,i,j}^* , \quad (6)$$

where $\overline{\psi}_{r,i}^*$ is the average of totally m samples, and $\psi_{r,i,j}^*$ is the j-th sample of the Monte Carlo integration. With this interpretation, for one pair of light sub-path and eye sub-path, we can obtain one sample for each sector in which only the indicator function of the samples is different. Therefore, for m pairs of light sub-paths and eye sub-paths, we can obtain m samples for any of the n sectors to compare their sample means.

3.2.2 ANOVA F-test

Since ANOVA F-test can assess whether the expected values of a quantitative variable within n groups differ from each other, it is applied to test our null hypothesis that the expected value of $\psi_{r,i,j}^*$ in each sector is the same. The Monte Carlo samples used to estimate the integration of $\Psi^*(\mathrm{y})$ over the same sector are regarded as samples of the same group in the ANOVA F-test. The F-test statistic is the ratio of between-group variability and within-group variability, and its equation for groups with equal sample size is

$$F = \frac{m \sum_{i=1}^{n} (\overline{\psi}_{r,i}^* - \overline{\psi}_r^*)^2 / (n-1)}{\sum_{i=1}^{n} \sum_{j=1}^{m} (\psi_{r,i,j}^* - \overline{\psi}_{r,i}^*)^2 / (n(m-1))} , \quad (7)$$

where m is the number of samples of each sector, and $\overline{\psi}_r^*$ is the average of all the nm samples. Intuitively, if the mean values are close across these sectors, then the between-group variability should be close to the within-group variability, and the statistic should be greater than and approaching 1. Otherwise, the between-group variability should be greater than within-group variability, and the statistic should be significantly greater than 1.

We use a significance level α_F (typically 0.01, 0.05 or 0.10 in statistics) to obtain a critical value F_c, which is a quantile of $1 - \alpha_F$ of the F-distribution with degrees of freedom $d_1 = n - 1$ and $d_2 = n(m - 1)$ in our case. If F is greater than F_c, our statistical model is considered violated by the real population, and the null hypothesis is rejected. We decrease the radius for kernel estimation in this circumstance to reduce potential bias. The ANOVA F-test is conservative in rejecting the null hypothesis. It rejects the null hypothesis only when the between-group difference is statistically significant.

The F-test assumes that the samples within a group are independent and normally distributed, and all groups have the same standard deviation, which is rather strict. The violation of these assumptions will decrease the accuracy of the F-test. However, the F-test for ANOVA is generally robust and the accuracy increases as sample size grows [24]. Also, F-test has shown its availability when the samples follow a wide range of distributions [25]–[27]. F-test works well in our application because we keep accumulating samples when the F-test does not reject the null hypothesis, which will be described in our algorithm in Sec. 5.1.

3.2.3 Analysis F-test vs. χ^2-test

χ^2-test is used to assess whether samples are uniformly distributed in n equi-areal sectors. Like F-test, if the statistic χ^2 is greater than a critical value χ_c calculated using the significance level α_χ, the null hypothesis is rejected. The χ^2-test statistic can only evaluate the difference between a real population and the hypothesized distribution. While according to Equation 7, the ANOVA F-test statistic evaluates the difference between the population with its contribution and the hypothesized distribution function. Therefore, the ANOVA F-test is more accurate in making statistical inferences for this type of sample, which can alleviate the potential bias, as illustrated in Figure 4 (right).

3.2.4 Test Samples

Since Equation 5 has a similar formulation to PM (as shown in Equation 1), we can reuse the samples of PM to get $\psi_{r,i,j}^*$. According to the definition of $\Psi(\mathbf{x}, \mathbf{x}^*)$, each sample $\psi_{r,i,j}^*$ is the summation of the contributions from all full paths constructed by one eye sub-path and one light sub-path. Therefore, one eye sub-path and one light sub-path make up a sample for each sector, though most samples $\psi_{r,i,j}^*$ equal zero because their values of indicator function are zero. By convention, we use a range searching data structure to obtain non-zero contribution eye and light vertices pairs, and therefore we can not directly get the value of $\psi_{r,i,j}^*$. We have to identify the sub-paths that the vertices belong to and sum up the contributions from the same sub-paths, and this procedure is computationally expensive.

To reduce this heavy computational overhead, we assume that one eye sub-path and one light sub-path can only construct no more than one full path with a non-zero contribution. This is based on an observation that a random eye vertex and a random light vertex generally have a very low probability of constructing a non-zero-contribution path since the kernel radius is relatively small compared with the scene size. Consequently, each non-zero-contribution full path is treated as an individual sample to get $\psi_{r,i,j}^*$.

4 HYPOTHESIS TESTING FOR VCM+

PPM algorithm works in unidirectional sampling way, that is one eye sub-path and one light sub-path construct one sample for each path length. While in the VCM framework, one eye sub-path and one light sub-path construct multiple samples for each path length using the MIS technique, and each sample is weighted to minimize the total variance [13]. Consequently, the statistical model used in PPM formulation is incompatible with VCM's bidirectional framework. In this section, we first deduce the unbiased condition for bidirectional PPM of VCM without BDPT involved. Thereby, we present an unbiased VCM estimator provided that the unbiased condition is satisfied, and we obtain VCM+ by hypothesis testing for unbiasedness.

4.1 Unbiased Condition for Bidirectional PM of VCM

We start with the estimate of bidirectional PM as an ingredient of VCM [6], which can be written as

$$I \approx \frac{1}{N} \sum_{i=1}^{N} \langle I \rangle_i \quad , \tag{8}$$

where N is the number of eye sub-paths; and $\langle I \rangle_i$ is the estimate using the i-th eye sub-path as:

$$\langle I \rangle_i \approx \frac{1}{J} \sum_{j=1}^{J} \sum_{l=2}^{\infty} \sum_{s=2}^{l-1} w_{s,r}(\overline{\mathbf{x}}_{i,j,l,s}^*) \frac{K_{s,r}(\overline{\mathbf{x}}_{i,j,l,s}^*) \Psi_s(\overline{\mathbf{x}}_{i,j,l,s}^*)}{p_s(\overline{\mathbf{x}}_{i,j,l,s}^*)}. \tag{9}$$

In this equation, J is the number of light sub-paths for each eye sub-path; r is the support radius of the kernel; $\overline{\mathbf{x}}_{i,j,l,s}^*$ is the full path of length l constructed by the s-th vertex of the i-th eye sub-path and the (l-s+2)-th vertex of the j-th light sub-path; $K_{s,r}$ is the kernel function over the two vertices; $w_{s,r}$ is the MIS weight function concerning sampling probability; p_s is the probability density function; Ψ_s is the full path contribution function, which includes the emission, geometry, bidirectional scattering distribution function (BSDF), and eye importance terms.

To derive a theoretical formulation, we use path integral framework [28], and the expected value of I in an integral form is:

$$I = \sum_{l=2}^{\infty} \sum_{s=2}^{l-1} \int_{\mathcal{M}^{l+2}} w_{s,r}(\overline{\mathbf{x}}_l^*) K_{s,r}(\overline{\mathbf{x}}_l^*) \Psi_s(\overline{\mathbf{x}}_l^*) d\mu(\overline{\mathbf{x}}_l^*), \tag{10}$$

where $\mathcal{M}$ is the scene surfaces in $\mathbb{R}^3$; $\overline{x}_l^* = (\mathbf{x}_0 \dots \mathbf{x}_{l+1})$ is a full path of length l constructed by l+2 vertices; $d\mu(\overline{x}_l^*) = dA(\mathbf{x}_0) \dots dA(\mathbf{x}_{l+1})$ is the differential product area measure.

To focus on the kernel estimation, we apply Fubini's theorem to the above equation (Equation 10) to change the order of integration and then integrate over $d\mu(\overline{\mathbf{x}}_l^*)$ except the (s-1)-th and the s-th differential into $F_{l,s,r}$, then we have:

$$I = \int_{\mathcal{M}^2} \sum_{l=2}^{\infty} \sum_{s=2}^{l-1} K_r(\|\mathbf{x} - \mathbf{x}^*\|) F_{l,s,r}(\mathbf{x}, \mathbf{x}^*) dA(\mathbf{x}^*) dA(\mathbf{x}). \tag{11}$$

Let $\overline{\mathbf{x}}'$ be a path of length l with $l + 2$ vertices having $\mathbf{x}$ for the fixed $(s - 1)$-th vertex and $\mathbf{x}^*$ for the fixed s-th vertex, then it lives within $\mathcal{M}^l$ and $d\mu(\overline{\mathbf{x}}')$ equals to $dA(\mathbf{x}_0) \dots dA(\mathbf{x}_{s-2}) dA(\mathbf{x}_{s+1}) \dots dA(\mathbf{x}_l)$. $F_{l,s,r}(\mathbf{x}, \mathbf{x}^*)$ can then be written as:

$$F_{l,s,r}(\mathbf{x}, \mathbf{x}^*) = \int_{\mathcal{M}^l} w_{s,r}(\overline{\mathbf{x}}') \Psi_s(\overline{\mathbf{x}}') d\mu(\overline{\mathbf{x}}'). \tag{12}$$

Same with the way discussed in subsection 3.1, we change one of the domains of integration in Equation 11 to the unified space as

$$I \approx \int_{\Omega_r} K_r(\|\mathbf{y}\|) \int_{\mathcal{M}} \sum_{l=2}^{\infty} \sum_{s=1}^{l-1} F_{l,s,r}^*(\mathbf{x}, \mathbf{y}) dA(\mathbf{x}) \, d\mathbf{y}, \tag{13}$$

where $F_{l,s,r}^*(\mathbf{x}, \mathbf{y}) = F_{l,s,r}(\mathbf{x}, f_{\mathrm{map},\mathbf{x}}^{-1}(\mathbf{y}))$.

In principle, the bias of PM algorithm is produced by the kernel function K_r, while the term $F_{l,s,r}^*$ will not introduce any bias. Therefore, by substituting K_r with Dirac delta function δ, we can obtain an unbiased integral as

$$I = \int_{\Omega_r} \int_{\mathcal{M}} \sum_{l=2}^{\infty} \sum_{s=1}^{l-1} F_{l,s,r}^*(\mathbf{x}, \mathbf{y}) \delta(\|\mathbf{y}\|) dA(\mathbf{x}) d\mathbf{y}$$

$$= \int_{\mathcal{M}} \sum_{l=2}^{\infty} \sum_{s=2}^{l-1} F_{l,s,r}^*(\mathbf{x}, \mathbf{0}) dA(\mathbf{x}). \tag{14}$$
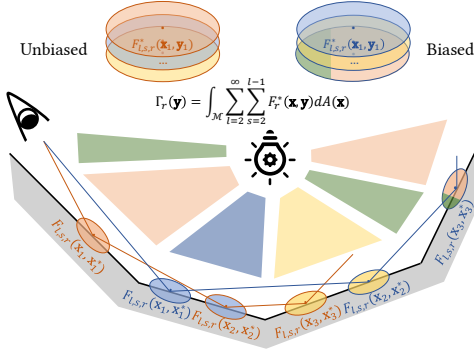
Fig. 5: Illustration of unbiased condition for a bidirectional estimator. Two eye sub-paths are traced from the eye individually. The kernel estimation of the eye sub-path in red will be unbiased, while the blue ones will be biased. The discs at light vertices $\mathbf{x}$ indicate the contribution function $F_r(\mathbf{x}, \mathbf{x}^*)$ in path space. The contribution $F_r(\mathbf{x}, \mathbf{x}^*)$ around $\mathbf{x}$ are aligned as $F_r^*(\mathbf{x}, \mathbf{y})$ and further integrated to obtian function $\Gamma_r(\mathbf{y})$. The estimator is unbiased if $\Gamma_r(\mathbf{y})$ is a constant.

For simplicity, we define function $\Gamma_r(\mathbf{y})$ as

$$\Gamma_r(\mathbf{y}) = \int_{\mathcal{M}} \sum_{l=2}^{\infty} \sum_{s=1}^{l-1} F_{l,s,r}^*(\mathbf{x}, \mathbf{y}) dA(\mathbf{x}), \qquad (15)$$

and we can obtain a simple and sufficient condition for an unbiased estimator as

$$\forall \mathbf{y} \in \Omega_r, \Gamma_r(\mathbf{y}) \equiv \Gamma_r(\mathbf{0}), \qquad (16)$$

i.e., function $\Gamma_r(\mathbf{y})$ concerning contributions should be constant within a disk domain with a support radius r, as illustrated in the top of Figure 5. Equation 16 is served as the unbiased condition for bidirectional PM (VCM without BDPT). Therefore, the null hypothesis for the bidirectional PM algorithm is that the unbiased condition mentioned above holds.

Compared to the unbiased condition for PM (Equation 4), the intuition of Equation 16 is also evenly distributed illumination, but Equation 16 takes multiple sampling strategies together with MIS weights into consideration. It implies the overall bias is also the MIS-weighted average of the bias from different sampling strategies. Note that multiple estimates with bias can theoretically be unbiased when their bias happens to cancel out. Therefore, the MIS weights affect bias, and the unbiased condition (Equation 16) should implicitly involve the MIS weights to safeguard the unbiasedness tightly.

4.2 Unbiased VCM Estimator

In the bidirectional framework, the VCM estimator considers one eye sub-path each iteration and combines multiple sampling techniques from VC (i.e., BDPT) and VM (i.e., PM) using the MIS technique. The eye vertices on the eye sub-path are used to construct regular paths with n_{VC} independent light sub-paths and extended paths with n_{VM} independent light sub-paths using radius r. VCM estimator can then be written as:

$$\langle I \rangle = \langle C \rangle_{\mathrm{VC}} + \langle C \rangle_{\mathrm{VM}}, \qquad (17)$$

where

$$\langle C \rangle_{\mathrm{VC}} = \frac{1}{n_{\mathrm{VC}}} \sum_{l=1}^{n_{\mathrm{VC}}} \sum_{k=1}^{\infty} \sum_{s=0}^{k+1} w_{\mathrm{VC},s,r}(\overline{\mathbf{x}}_{k,s,l}) \frac{f(\overline{\mathbf{x}}_{k,s,l})}{p_s(\overline{\mathbf{x}}_{k,s,l})}$$

$$\langle C \rangle_{\mathrm{VM}} = \frac{1}{n_{\mathrm{VM}}} \sum_{l=1}^{n_{\mathrm{VM}}} \sum_{k=2}^{\infty} \sum_{s=1}^{k-1} w_{\mathrm{VM},s,r}(\overline{\mathbf{x}}_{k,s,l}^*) \frac{f_{s,r}(\overline{\mathbf{x}}_{k,s,l}^*)}{p_s(\overline{\mathbf{x}}_{k,s,l}^*)}. \qquad (18)$$

In this equation, $\overline{\mathbf{x}}_{k,s,l}$ is the regular path constructed by k-s+1 vertices from the eye sub-path and s vertices from the l-th light sub-path among n_{VC} light sub-paths; $\overline{\mathbf{x}}_{k,s,l}^*$ is the extended path constructed by k-s+1 vertices from the eye sub-path and s+1 vertices from the l-th light sub-path among n_{VM} light sub-paths; f is the measurement contribution function measuring the contribution along the path; $f_{s,r}$ is the measurement contribution function for VM; p_s is the probability density function; and $w_{v,s,r}$ is the power heuristic function of MIS for corresponding techniques where v^* denotes VC or VM:

$$w_{v^*,s,r}(\overline{\mathbf{x}}_k) = \frac{n_{v^*}^{\beta} p_{v^*,s,r}^{\beta}(\overline{\mathbf{x}}_k)}{\Upsilon},$$

$$\Upsilon = n_{\mathrm{VC}}^{\beta} \sum_{s=0}^{k+1} p_{\mathrm{VC},s,r}^{\beta}(\overline{\mathbf{x}}_k) + n_{\mathrm{VM}}^{\beta} \sum_{s=1}^{k-1} p_{\mathrm{VM},s,r}^{\beta}(\overline{\mathbf{x}}_k). \qquad (19)$$

The balance heuristic corresponds to β=1. In the above equation, $p_{\mathrm{VM},s,r}$ is proportional to r^2, while $p_{\mathrm{VC},s,r} = p_{\mathrm{VC},s}$ is independent of r. That is, the kernel radius is the key to determining the contribution from VM.

Then, the progressive estimator is the average of N iterations:

$$\langle I \rangle = \frac{1}{N} \sum_{i=1}^{N} \langle I \rangle_i, \qquad (20)$$

where $\langle I \rangle_i$ is the result computed by Equation 17 using the i-th radius r_i, the i-th set of eye sub-paths and the i-th set of light sub-paths.

Besides the MIS weight, VM ingredient in the estimate shares the same components of Equation 16. Therefore, we can integrate the MIS weight of both VM and VC into the full paths' contribution. Then the unbiased condition for bidirectional PM is also available for the VCM framework.

In our VCM+ algorithm, we use ANOVA F-test to test whether the unbiased condition holds or not in the same way as discussed in Sec. 3.2. We see the null hypothesis as an equivalent of the unbiased condition for VM in VCM+ in practice. We will reject/not reject the null hypothesis given the testing of observations of the contribution. If not rejecting the null hypothesis, the radius r will be identified for unbiased estimation. VCM+ will find the right radius as early as possible during the iterations, and hold this radius unchanged for kernel estimation. That is, the kernel radius needs not shrink progressively to mitigate the bias of VM. Thus we reduce variance by collecting more photons with larger kernels.

5 ALGORITHM

In this section, we will outline our algorithm based on hypothesis testing. As no modifications have been made to the BDPT component, we will focus solely on the updates made

to the PPM algorithm, which involve general enhancements to VCM under our theoretical model. The statistical models proposed for both PPM and VCM frameworks only affect the sample collection and radius reduction module solely, so that they can share the same working pipeline thereof.

5.1 Kernel Radius and Radiance Estimation

The PPM part of our algorithm overall follows the working flow similar to CPPM: In each iteration, we trace J light sub-paths, build a range search data structure of the light vertices, trace multiple eye sub-paths for each pixel, collect the light vertices around the eye vertices for kernel estimation, update the searching radius by hypothesis testing, update the power heuristic function of MIS for VCM+ according to Equation 19, and update the estimate of the radiance finally (Equation 17). The key steps with regard to hypothesis testing at each gather point are as the following:

1) Disk Ω_r is partitioned into sectors (Sec. 3.2.1), using the same strategy as CPPM.
2) The statistics of the light vertices that fall in the sectors are accumulated. By applying light path tracing, each light vertex that falls inside the disk produces one sample for each sector. For each sector, we maintain the sample count (m), the sums of sample contributions ($\overline{\psi}_{r,i}^*$) and the sums of their squares ($\sum_{j=1}^m (\psi_{r,i,j}^*)^2$) individually. They correspond to the terms presented in Sec. 3.2.2. These statistics are collected to compute an F-statistic using Equation 7 for the ANOVA F-test.
3) ANOVA F-test (Sec. 3.2.2) is used to check whether the null hypothesis can be rejected. The null hypothesis is that the unbiased condition holds itself (see Equation 16) as mentioned in Sec. 3.1. We compare the computed F-statistic against a threshold based on a fixed confidence level of the statistical test (typically a global parameter) to reject or not reject the null hypothesis. In this way, we determined if the observations within the current kernel radius can lead to unbiased estimation.
4) Update the kernel radius according to Equation 21.
5) Update the power heuristic function of MIS for VCM+ according to Equation 19.

The radius is updated according to the F-test as

$$r_{i+1} = \begin{cases} r_i, & \text{if } F \leq F_c \\ \max\left(kr_i, r_{\min}\sqrt{(i+1)^{\alpha-1}}\right), & \text{if } F > F_c \end{cases} . \quad (21)$$

Here, the parameter $\alpha \in (0,1)$ bounds the rate of radius reduction, $k \in (0,1)$ is a user-defined shrinking ratio, and $r_{\min}\sqrt{(i+1)^{\alpha-1}}$ serves as a lower bound to ensure convergence. This is effective because the F-test is conservative in rejecting the null hypothesis, and a small sample size does not significantly increase the probability of wrongly rejecting a true null hypothesis [20], [29]. We remove the unnecessary parameter (e.g., β) designed for tuning in CPPM pipeline, which can improve the flexibility of our algorithm based on the convergence analysis in the later Sec. 5.2.

When performing the ANOVA F-test, J new samples are collected in one iteration for each sector. As known to all, a large sample size can decrease the probability of wrongly not rejecting a false null hypothesis. To increase the sample size, we retain sample statistics of former iterations where the radius does not shrink. Therefore, denoting t as the number of iterations for which the current radius remains fixed, the F-test has a total of tJ samples for each sector. According to Equation 7, with an infinite number of samples, the F-test statistic goes to infinity if bias still exists. Our proposed F-test will finally reject the null hypothesis with sufficient samples. The only exception is, by chance, that differences in contribution occur not between sectors but within. It is a rather rare circumstance and can hardly occur in practice. Generally, our algorithm performs F-test only on the luminance channel. In this way, the probability of wrongly rejecting an unbiased estimator equals $\alpha_F = 0.01$. When handling the chromatic light sources, we perform F-test on RGB channels individually.

5.2 Convergence Analysis

In this subsection, we analyze the convergence rate of our algorithm. In the worst case where the hypothesis testing always rejects the null hypothesis, i.e., the condition for unbiased estimation is never satisfied, the radius will always shrink once the algorithm gets a non-zero sample. Without the lower bound of the radius, we can deduce that k does not affect the convergence of the kernel radius, and the radius convergence rate is $O(N^{-1/2})$ in this case [15]. This allows our algorithm to shrink the kernel radius faster before reaching an appropriate radius compared to CPPM and PPM. However, applying Knaus and Zwicker's probabilistic framework [15], we learn that the variance of PPM does not converge in this case. Therefore, with a lower bound, the convergence rate of PPM in our algorithm is asymptotically identical to PPM under the worst situation. The optimal mean squared error (MSE) convergence rate $O(N^{-2/3})$ is achieved by setting $\alpha = 2/3$ in this situation. In practice, the lower bound can hardly be reached since most pixels can pass the hypothesis testing with a relatively large kernel and stop reducing the kernel radius. Assuming the hypothesis testing produces the correct result, the convergence rate becomes $O(N^{-1})$. With this convergence rate, we can deduce that our algorithm also obtains the optimal convergence rate of $O(N^{-1})$ for VCM [6]. That is, the VCM+ algorithm can reach a convergence rate of $O(N^{-1})$ under ideal conditions.

The parameter k only affects initial bias and variance. A larger k generally introduces more bias but less variance. Therefore, the optimal k may vary from scene to scene. We will provide some empirical values from our experiments in Sec. 6.2.2.

6 EXPERIMENTS AND DISCUSSIONS

We implement our algorithm under both the PPM framework and the VCM framework. In this section, we provide experiments and asymptotic performance on diverse scenarios, and we also discuss the effect of different parameter values related to our algorithm in the experiments. The results of all comparative experiments are obtained under the same environment on the same machine. All the algorithms are developed based on the Mitsuba Renderer [30].
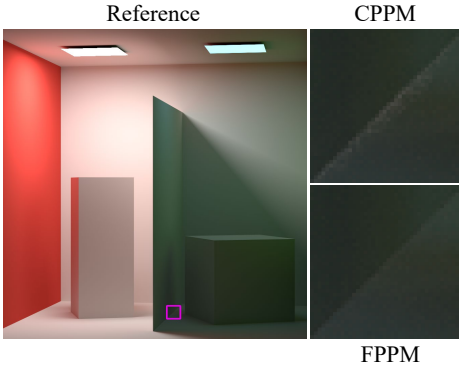
Fig. 6: FPPM (ours) vs. CPPM after 1,000 iterations, using Split Cornell Box as an example. Our algorithm helps alleviate the light leakage artifacts.

6.1 Results and Comparisons

Firstly, we compare two progressive photon mapping algorithms: FPPM (our F-test-based PPM) and CPPM (state-of-the-art method). We then conduct various experiments and compare VCM+ with VCM (BDPT+SPPM).

6.1.1 FPPM vs. CPPM

Settings: Both CPPM and FPPM (our algorithm) are based on hypothesis testing. The common parameters used by the CPPM and ours are the same unless otherwise stated. To perform the statistics test, we partition Ω_r into n_a equiareal annuli and each annulus into n_s equi-areal sectors. Specifically, we generally set $n_a = 2$, $n_s = 6$, $k = 0.7$ and significance level $\alpha_\chi = \alpha_F = 0.01$. The critical value for the χ^2-test (χ_c) for CPPM is 24.724, and the critical value for the F-test (F_c) update in run time according to the sample size, and is approximately 2.25 when the total sample count approaches infinity.

We first validate that our algorithm can minimize the common artifacts compared to CPPM. We test a special case, Split Cornell Box. As shown in Figure 6, our algorithm alleviates the light leaks and obtains more accurate results in equal-iteration comparison. Moreover, we conducted comparisons on three scenes, as shown in Figure 7 with highlighted details. These three scenes showcase diverse lighting conditions. The Kitchen contains numerous objects, small area lights above the oven, and a large area light from the window; the Hall scene is a two-storey building with glossy floors, and has many area lights as well as a large light source outside; the Staircase scene features a textured spotlight, which is a special case.

In all tested scenes, our algorithm (FPPM) shows superior performance over CPPM in terms of MSE. Moreover, our algorithm can alleviate the noise, as highlighted in the zoom-in region of Kitchen and Hall; and it can alleviate the visual blur, as highlighted in the zoom-in region of Staircase. In CPPM, χ^2-test works well relying on the independence assumption, a special case where unweighted photon statistics can model the local radiance. It is not always the case under complex indoor lighting conditions. χ^2-test will malfunction with either glossy materials, indirect illumination with multiple area lights, or textured lighting. Where it incorrectly underestimated the kernel radius, noise

may occur due to the lack of photons within a small kernel. Where the kernel size is overestimated, noticeable bias will arise. In contrast, our algorithm correctly determines the kernel radius to balance noise and bias by removing the independence assumption of CPPM, leading to better results. We use greyscale to visualize the kernel radius of each zoom-in region in the middle of each block. White indicates a relatively large radius while black indicates a relatively small one. Our algorithm can clearly indicate the boundary at illumination discontinuities, implying that small radii should be used as expected to reduce bias.

We also highlight the asymptotic performance of FPPM compared with CPPM, using SPPM as a baseline, as the convergence plots shown in Figure 8. Both our algorithm and CPPM significantly outperform the baseline. Since our hypothesis-testing model is more general for samples with contributions attached under various lighting conditions, our algorithm shows superior performance and converges fast.

6.1.2 VCM+ vs. VCM

Settings: As a reinforcement of VCM, VCM+ shares the same setting with VCM on common parameters unless otherwise stated. Specifically, initial radius r_1 is set to $0.1\% R_{bb} \sim 0.3\% R_{bb}$ for all algorithms, where R_{bb} is the radius of the bounding box of visible objects. Lower bound ratio of the radius r_{min} is set to $0.1\% r_1$; and α is set to 0.75 for all algorithms, as recommended in VCM [6]. Balance heuristic for MIS is used. Eight benchmarks (including the three scenes previously used to test PPM algorithms) are used for testing. We collect the number of iterations and time consumption required to reach a specified MSE, and show the convergence plot over iterations. As seen from Table 1 and Figure 9, our VCM+ achieves the best performance with the lowest MSE over VCM in all testing scenarios.

The Pool is a typical scene that includes many complex caustics, and it is very hard to converge to a correct result. Our VCM+ shows more than 2× speedup over VCM with far fewer iterations, as shown in Table 1; the speedup is going larger asymptotically, as shown in Figure 9. We also highlight the significant improvement of our algorithm in Figure 1. VCM produces a much noisier result (MSE=16.67) than ours (MSE=5.54) after 10K iterations. When compared with VCM, VCM+ has fewer iterations and less time consumption (Table 1). Noticed that VCM+ spends more time on each iteration than VCM because it needs around 10% extra computational overhead for hypothesis tests. However, this cost is worthwhile for prominent performance gain.

To figure out the rationale behind our algorithm's better performance, we conduct more comparisons between our VCM+ and the baseline VCM. We demonstrate the results Figure 10 in and Figure 11, respectively. The Glass scene (Figure 11) has many caustics, and the Kitchen scene (Figure 10) has few S-D-S paths but many glossy materials. Specifically, we illustrate the relative contribution between BDPT and PM in false color. In our VCM+, PPM has a large relative contribution than BDPT in the regions that include S-D-S paths as expected, whereas PPM contributes less as it supposed to in other regions. VCM+ can compute appropriate contributions from photon mapping for most pixels. Consequently, VCM+ achieves steady gains at the
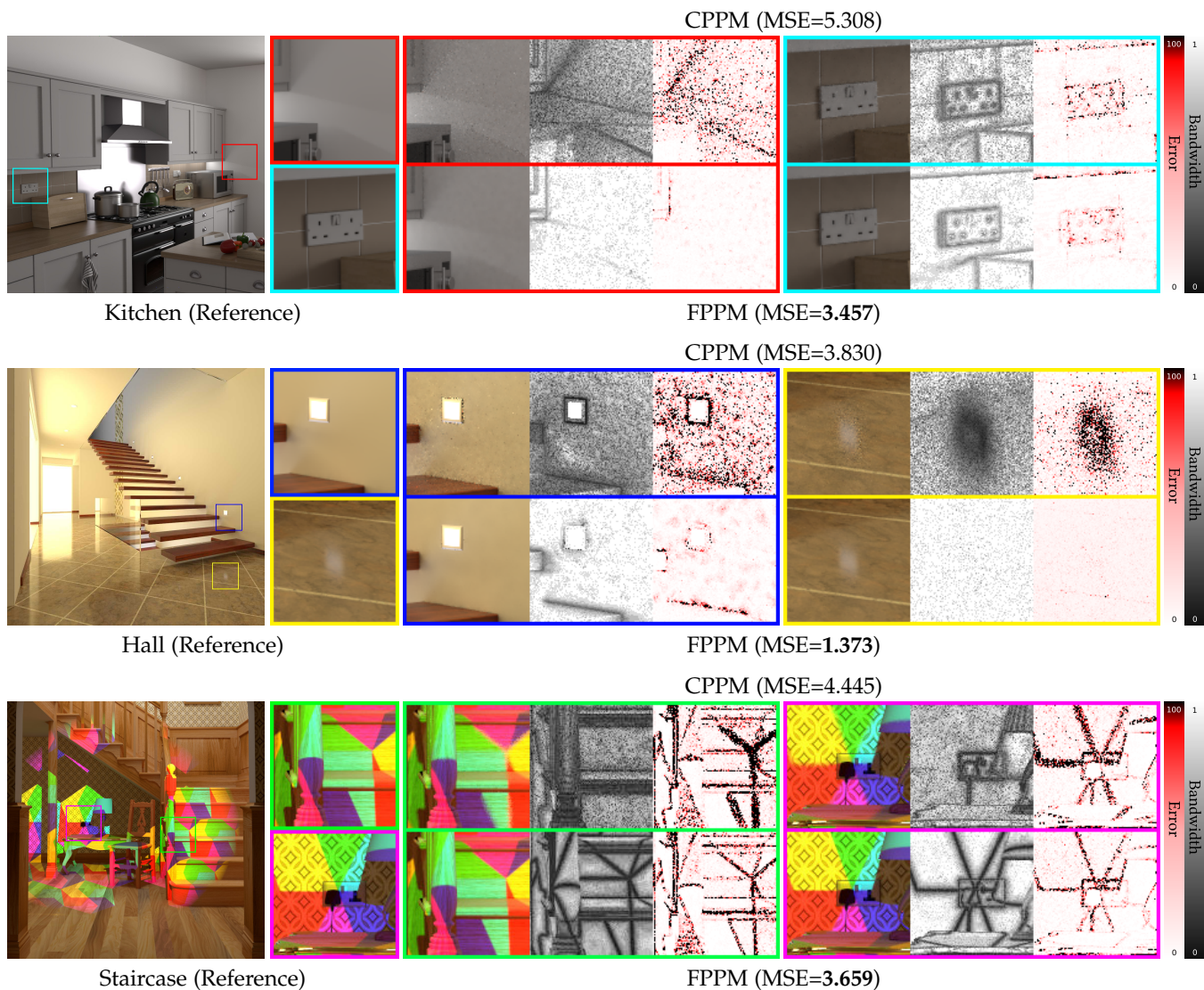
Fig. 7: Our algorithm (FPPM) outperforms CPPM in equal-pass comparisons on different test scenes (10K iterations for Kitchen and Hall, and 5K iterations for Staircase). We visualize the kernel radius in greyscale (middle of each zoom-in block) and the squared error (right of each zoom-in block). Our algorithm yields less noisy results and correctly handles false rejections produced by χ^2-test where emitted photons failed to capture the actual irradiance distribution. Moreover, F-tests on different spectral channels help improve the quality of textured lights.
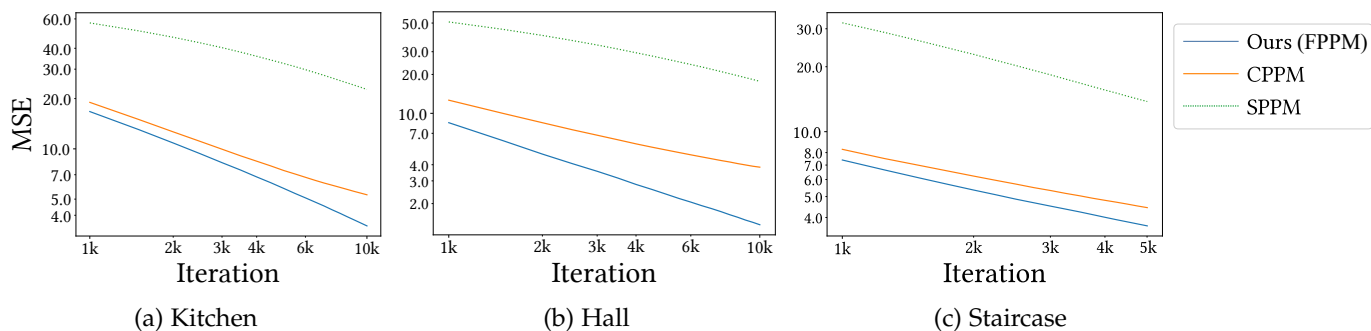
Fig. 8: MSE over iterations of SPPM, CPPM, and Ours (FPPM). FPPM shows superior performance, converges faster than CPPM, and significantly outperforms the baseline.

Scene		Pool	Hall	Bathrooom	Dragon	Glass	Ball	Kitchen	Staircase
MSE $\leq$		50.0	20.0	10.0	10.0	10.0	10.0	10.0	10.0
VCM	Iter.	501	650	184	226	1361	149	808	573
	Time	435.48s	2358.52s	807.03s	287.39s	2244.49s	115.85s	3284.59s	2541.59s
VCM+	Iter.	222	397	153	183	807	98	599	101
	Time	214.16s	1496.16s	732.68s	241.83s	1445.52s	81.31s	2627.52s	586.70s

TABLE 1: Iterations and time budgets required to reach the target MSE for VCM+ and VCM on eight benchmarks. Our VCM+ shows superior performance in terms of iterations and time consumption.
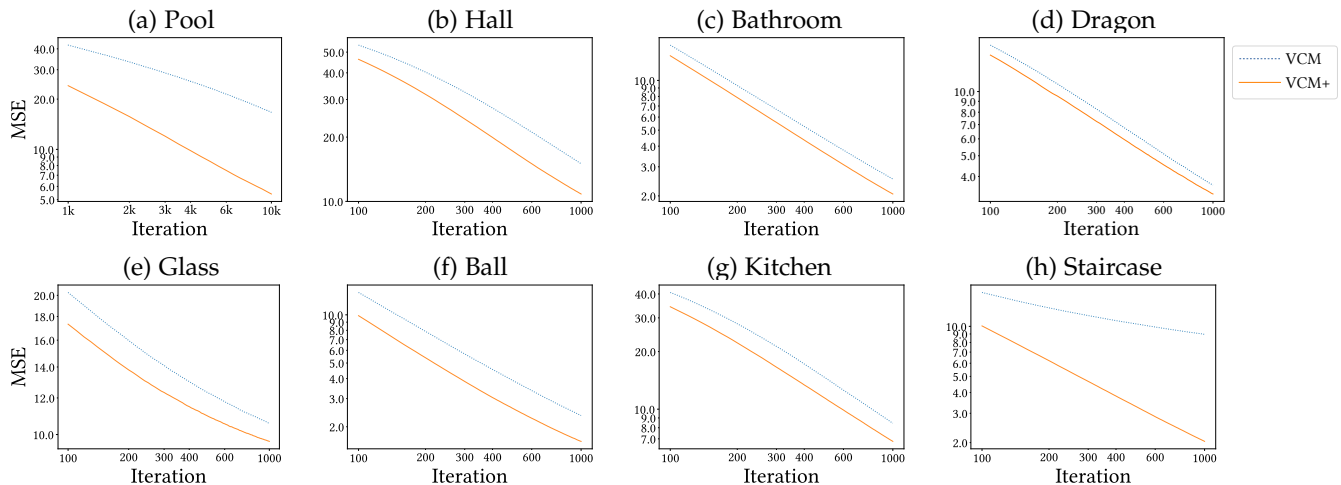

Fig. 9: MSE over iterations of eight test scenes by different methods. Our VCM+ algorithm using hypothesis testing has lower MSE and shows superior performance over VCM.
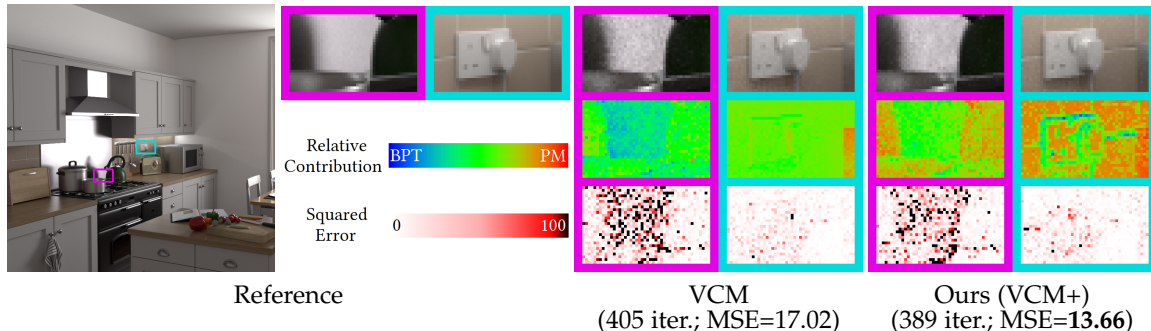

Fig. 10: Equal-time (35 minutes) comparison on Kitchen scene. We visualize the relative contribution and the squared error. PM plays a more important role in VCM+ and compensates for the high variance brought by BDPT. VCM+ can compute appropriate contributions from BDPT and PM individually and obtains better results.

expense of computing photon mapping (i.e., vertex merging, VM). And for the pixels initially estimated to be biased, PM in VCM+ has less contribution than that of VCM, thereby alleviating the PM-induced bias. As a result, our algorithm has lower variance and bias overall and produces less noisy images with sharper edges. Our results also indicate that the initial bias is typically lower in the regions where illumination changes smoothly, and higher at the boundaries of objects and caustics.

In addition, we analyze the radius of each iteration in PPM, which determines the contributions between PPM and BDPT. We show the results of a Ball scene with lights behind textured glasses (Figure 12), and all illuminations come from

this type of light source. The results show that our VCM+ can find appropriate radii for unbiased kernel estimation and synthesize superior images over VCM. We visualize the relative radius (actual radius / VCM's radius) in this scene after 400 iterations in Figure 13, which can explain the rationale for better performance by our statistical model with hypothesis testing using F-test. Our method can find an appropriate radius suitable for the minimization of variance and reduction of bias at different locations, especially a relatively small radius at the boundary or discontinuity of illumination while a relatively large radius at the region with a smooth lighting transition. Overall, our algorithm leverages the contributions from VC (BDPT) and VM (PM)
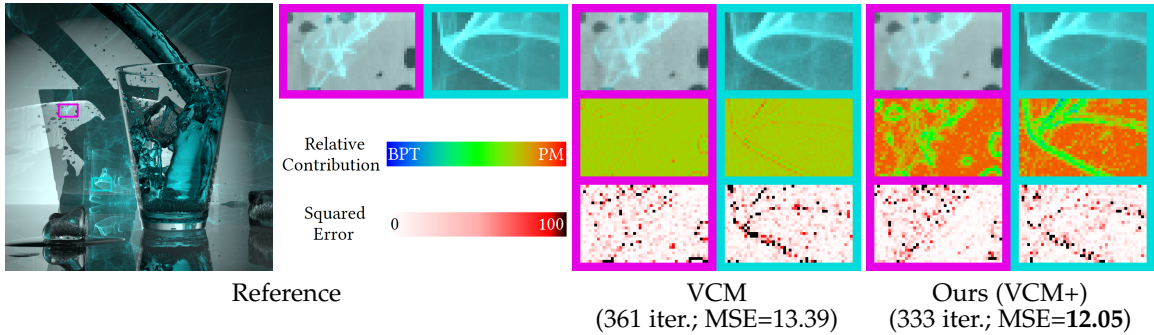
Fig. 11: Equal-time (10 minutes) comparison on Glass scene. We visualize the relative contribution and the squared error. We highlight the contribution from PM that plays an important role in the region of caustics in our VCM+. Whereas, VCM is less efficient in handling these regions and leads to blurry or noisy results.

by finding the appropriate radius and hereby achieves better performance.

6.2 Parameters

We investigate the robustness of our VCM+ by tuning the values of different parameters used in the experiments that may affect the performance. The values of parameters vary, but we still observe improvements over VCM.

6.2.1 Initial Radius

The initial radius may affect the performance of a progressive algorithm. Therefore, we test two algorithms (VCM and VCM+) with different initial radii. Three values including $1\%, 0.1\%$, and $0.01\% R_{\mathrm{bb}}$ are used for the test of two scenes, Glass as the caustics representative and Dragon as the diffuse representative. These values range from coarse to fine-grained, and 0.01% may be even pixel-level smaller. The asymptotic performance of these two algorithms using different initial radii are shown in Figure 14.

From the results, the initial radii have impact on both VCM and VCM+ performance. However, VCM+ outperforms VCM in all settings and exhibits good convergence behavior. A large initial radius (i.e. 1%) favors lower MSE in the early stage when variance dominates the error. As bias gradually accounts for a higher proportion of MSE, a relatively larger radius will slow down convergence. This can be manifested in that VCM with 0.1% radius outperforms VCM with 1% radius at around 1000 iterations in both scenes, with a steeper downward trend. our VCM+ benefits from the adaptive kernel radius selection scheme and hence receives a smaller impact from an overly large initial radius when compared to VCM. However, an aggressively small initial radius (i.e. 0.01%) makes the contribution from PM always trivial to VCM via MIS weights, resulting in a shift of VCM towards BDPT in the entire process. Obviously, the smallest initial radius leads to the worst performance for both VCM and VCM+, and their performances are very similar. Therefore, we suggest using a larger initial radius for VCM+ compared to the values suggested by [6], because our test-driven unbiased radiance estimate can handle the bias well in the late rendering stages.

6.2.2 Radius Shrinkage k

As discussed in section 5, we should select a moderate k. We compare different values of k on a variety of scenes,

TABLE 2: MSE obtained under different radius shrinkage values of k. Different scenarios are tested by our approach and VCM (baseline). Our approach using different k values always outperforms the baseline; meanwhile, different k values do not show significant discrepancies in the results.

Bathroom (100 iter.)	k	0.5	0.6	0.7	0.8	0.9	VCM
	MSE	14.09	14.08	**14.06**	14.20	14.42	16.36
Ball (100 iter.)	k	0.5	0.6	0.7	0.8	0.9	VCM
	MSE	10.25	10.03	9.81	9.66	**9.56**	13.79
Hall (1,000 iter.)	k	0.5	0.6	0.7	0.8	0.9	VCM
	MSE	11.02	10.92	10.83	10.70	**10.60**	15.01
Pool (1,000 iter.)	k	0.5	0.6	0.7	0.8	0.9	VCM
	MSE	24.40	24.30	24.05	23.87	**23.71**	41.99

as listed in Table 2. For the Bathroom and Ball scenes, 100 iterations are tested. For the Hall and Pool scenes that are difficult to converge, 1,000 iterations are tested. Our approach consistently outperforms the baseline obtained by VCM, regardless of the value of k used. From this table, $k = 0.9$ obtains the best performance in most scenes, albeit different values do not introduce significant differences between each other. This may imply that a conservative shrinkage strategy may be appropriate for VCM+.

6.2.3 F-test Significance Level α_F

If a smaller α_F is used, the probability of wrongly rejecting an unbiased estimator will be lower. If a relatively larger α_F is used, the number of samples required to reject a biased estimator will be fewer. We test three typical values 0.01, 0.05, and 0.10 in statistics, and find that $\alpha_F = 0.01$ obtains better performance than the other two, as shown in Table 3. This implies that a lower rejection rate is a better setting. Our algorithm is conservative about kernel radius reduction in this case. However, most pixels can still converge to unbiased radii quickly. Overall, our VCM+ shows better performance than VCM regardless of the parameters used.

7 CONCLUSION, LIMITATION, AND FUTURE WORK

In this paper, our improvement improves the efficiency of the PPM and VCM estimators. Our key is a novel statistical model along with a hypothesis-testing method for an unbiased condition that is feasible for both PPM and VCM frameworks. From a statistical view, F-test, as one of the
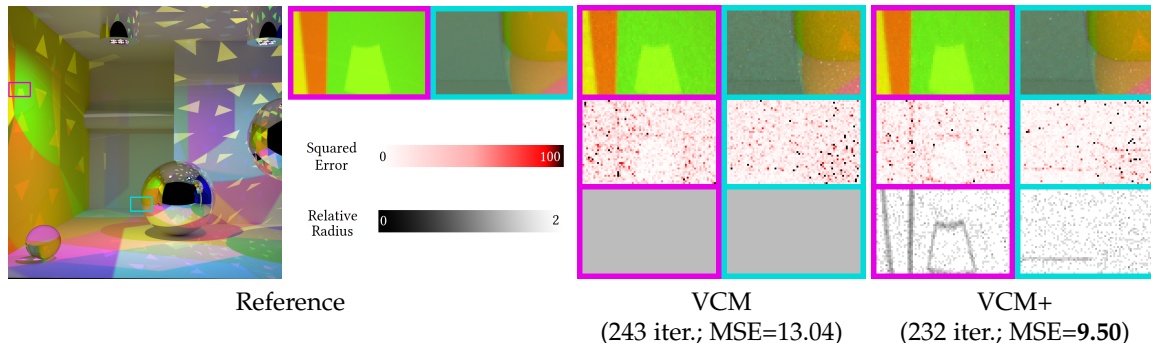
Fig. 12: Equal-time (10 minutes) comparison on the modified Ball scene with lights behind textured glasses. We visualize the squared error and relative radius (actual radius / VCM's radius). We highlight that our VCM+ can also handle special lighting setting well.
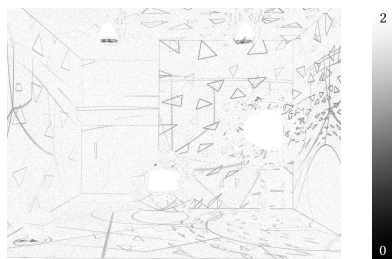

Fig. 13: Relative radius (VCM+/VCM) visualization. We highlight that VCM+ can find the appropriate radius that leverages the contributions from VC (BDPT) and VM (PM).

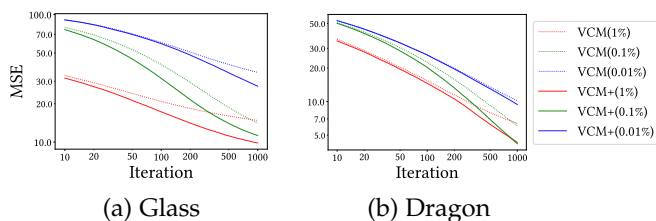(a) Glass   (b) Dragon

Fig. 14: Asymptotic performance of VCM+ vs. VCM over iterations using different initial radius. VCM+ always exhibits better performance than VCM.

parametric tests, is useful when comparing statistical models to observations and can compute the reliable test statistic. Our algorithm uses the F-test to detect potential bias and can find the appropriate radius for kernel estimation for most pixels. Consequently, our algorithm has lower variance and bias overall in practice and generates less noisy images with sharper boundaries. Our algorithm can help alleviate the light leakage artifacts as well.

Our algorithm has some limitations. First, the F-test requires sufficient samples to reject a false null hypothesis. In those regions where samples are inadequate, the F-test may take a few iterations to reject a biased estimation, resulting in some initial bias. Next, our algorithm is more suitable for those complex scenes with multiple bounces during the light transport because this is apt to produce light vertices with different contributions within a kernel radius (as Ψ_s of a path contribution function discussed in Equation 9). For those scenes where the light setup is simple and centered on one single object, which may imply the discrepancy of photon contribution is tiny, our algorithm no longer has an

TABLE 3: MSE value obtained under different α_F. Kitchen and Glass are rendered with 100 iterations. The MSE obtained by VCM is the baseline. VCM+ with $\alpha_F = 0.01$ obtains the best performance.

Kitchen	α_F	0.01	0.05	0.10	VCM
	MSE	**34.30**	35.58	37.05	40.90
Glass	α_F	0.01	0.05	0.10	VCM
	MSE	**17.44**	18.29	19.70	22.93

obvious advantage. Our algorithms only slightly improve over the baseline, as shown in the Car scene in Figure 15.

There are many venues for future work to extend our work. Combination with Markov chain Monte Carlo methods (MCMC) [17], [31] or online learning methods [32]–[35] to improve the sampling probability of high contribution paths may be able to alleviate the problem of initial bias. The MCMC methods introduce severe violations of the assumption of independence (discussed in Sec. 3.2.2), which needs further evaluation. Our unbiased condition used for VCM+ involves MIS weights that contain pdf values (Equation 16), which implicitly assumes pdfs do not change during the iterations, but a path in the online learning methods may have a different pdf at each iteration. Therefore, seamless combination of online learning methods also needs further investigation. In addition, the way of finding a better MIS strategy such as optimal MIS [36], or taking correlation of samples into consideration [37], or subspace MIS-aware sampling [38] that can combine highly efficient BDPT with PM technique is also an interesting topic.

The key of our radiance estimate to dealing with the samples of light vertices is orthogonal to those signal sampling and adaptive methods [39], [40]. Nevertheless, stochastic sampling of each pixel is employed in our method during the distributed ray tracing stage to generate multiple eye paths, same with SPPM [12]. On the one hand, it may be possible to replace the unbiased condition for kernel estimation derived in section 3 with a simpler criterion used in adaptive sampling. It may save computational overhead and make a possible improvement in the accuracy of bias detection. However, empirical criteria will not safeguard the unbiasedness of radiance estimation [40]. On the other hand, an extension of our hypothesis-testing-based methodology to guide the sampling of a pixel for distributed ray tracing process may be an alternative solution to improve the efficiency further. However, a reasonable statistical model
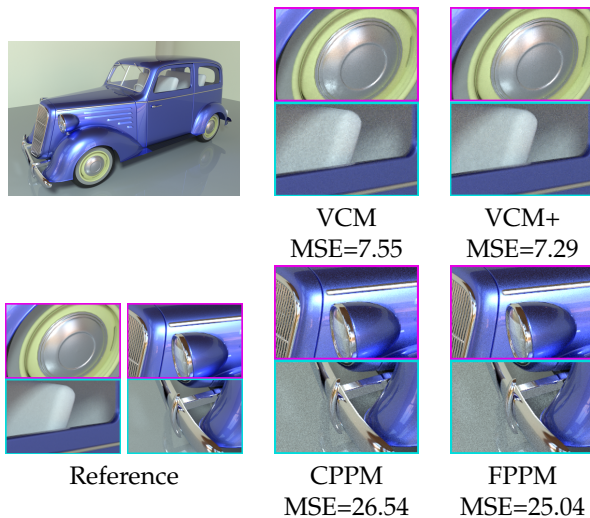
Fig. 15: Equal-time (20 minutes) rendering of Car scene, VCM (829 iterations) vs. VCM+ (811 iterations), and CPPM (5619 iterations) vs. FPPM (4742 iterations). VCM+ and FPPM exhibit only marginal improvements over VCM and CPPM, respectively.

and assumption based on the prior about the morphology of each pixel should be established in advance for a parametric test. We can turn to nonparametric tests, but a larger sample size is required to conclude with the same degree of confidence as parametric tests, i.e., it may have less power. An interesting idea will be to use unbiased samples to backward guide the adaptive sampling process in distributed ray tracing. These still need further investigation in the future.

REFERENCES

[1] E. P. Lafortune and Y. D. Willems, "Bi-directional path tracing," in *Proceedings of Third International Conference on Computational Graphics and Visualization Techniques (Compugraphics '93)*, Alvor, Portugal, December 1993, pp. 145–153.

[2] E. Veach and L. Guibas, "Bidirectional estimators for light transport," in *Photorealistic Rendering Techniques*. Berlin, Heidelberg: Springer, 1995, pp. 145–167.

[3] T. Hachisuka, S. Ogaki, and H. W. Jensen, "Progressive photon mapping," *ACM Transactions on Graphics (TOG)*, vol. 27, no. 5, p. 130, 2008.

[4] H. W. Jensen, "Global illumination using photon maps," in *Rendering Techniques' 96*. Vienna: Springer, 1996, pp. 21–30.

[5] Z. Lin, S. Li, X. Zeng, C. Zhang, J. Jia, G. Wang, and D. Manocha, "Cppm: Chi-squared progressive photon mapping," *ACM Trans. Graph.*, vol. 39, no. 6, Nov. 2020. [Online]. Available: https://doi.org/10.1145/3414685.3417822

[6] I. Georgiev, J. Křivánek, T. Davidovič, and P. Slusallek, "Light transport simulation with vertex connection and merging," *ACM Trans. Graph.*, vol. 31, no. 6, pp. 192:1–192:10, Nov. 2012. [Online]. Available: http://doi.acm.org/10.1145/2366145.2366211

[7] T. Hachisuka, J. Pantaleoni, and H. W. Jensen, "A path space extension for robust light transport simulation," *ACM Trans. Graph.*, vol. 31, no. 6, Nov. 2012. [Online]. Available: https://doi.org/10.1145/2366145.2366210

[8] J. T. Kajiya, "The rendering equation," ser. SIGGRAPH '86. New York, NY, USA: Association for Computing Machinery, 1986, p. 143–150. [Online]. Available: https://doi.org/10.1145/15922.15902

[9] T. Hachisuka, W. Jarosz, G. Bouchard, P. Christensen, J. R. Frisvad, W. Jakob, H. W. Jensen, M. Kaschalk, C. Knaus, A. Selle *et al.*, "State of the art in photon density estimation," in *Acm Siggraph 2012 Courses*. ACM Siggraph, 2012, p. 6.

[10] A. S. Kaplanyan and C. Dachsbacher, "Adaptive progressive photon mapping," *ACM Trans. Graph.*, vol. 32, no. 2, Apr. 2013. [Online]. Available: https://doi.org/10.1145/2451236.2451242

[11] S. Zhu, Z. Xu, H. W. Jensen, H. Su, and R. Ramamoorthi, "Deep kernel density estimation for photon mapping," *Computer Graphics Forum*, vol. 39, no. 4, pp. 35–45, 2020.

[12] T. Hachisuka and H. W. Jensen, "Stochastic progressive photon mapping," *ACM Transactions on Graphics (TOG)*, vol. 28, no. 5, p. 141, 2009.

[13] J. Vorba, "Bidirectional photon mapping," in *Proc. of the Central European Seminar on Computer Graphics (CESCG'11)*, 2011.

[14] T. Hachisuka, W. Jarosz, and H. W. Jensen, "A progressive error estimation framework for photon density estimation," *ACM Trans. Graph.*, vol. 29, no. 6, Dec. 2010. [Online]. Available: https://doi.org/10.1145/1882261.1866170

[15] C. Knaus and M. Zwicker, "Progressive photon mapping: A probabilistic approach," *ACM Transactions on Graphics (TOG)*, vol. 30, no. 3, p. 25, 2011.

[16] H. Qin, X. Sun, Q. Hou, B. Guo, and K. Zhou, "Unbiased photon gathering for light transport simulation," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 6, pp. 1–14, 2015.

[17] M. Šik, H. Otsu, T. Hachisuka, and J. Křivánek, "Robust light transport simulation via metropolised bidirectional estimators," *ACM Trans. Graph.*, vol. 35, no. 6, pp. 245:1–245:12, Nov. 2016. [Online]. Available: http://doi.acm.org/10.1145/2980179.2982411

[18] K. Calder, "Statistical inference," *New York: Holt*, 1953.

[19] B. Everitt, "The cambridge dictionary of statistics," *Cambridge University Press*, 1998.

[20] W. G. Cochran, "The chi-square test of goodness of fit," *The Annals of Mathematical Statistics*, vol. 23, no. 3, pp. 315–345, 1952.

[21] R. G. Miller Jr, *Beyond ANOVA: basics of applied statistics*. CRC press, 1997.

[22] R. A. Fisher, "On the 'probable error' of a coefficient of correlation deduced from a small sample," *Metron*, vol. 1, pp. 1–32, 1921.

[23] W. H. Kruskal and W. A. Wallis, "Use of ranks in one-criterion variance analysis," *Journal of the American statistical Association*, vol. 47, no. 260, pp. 583–621, 1952.

[24] D. S. Moore, G. P. McCabe, and B. A. Craig, *Introduction to the Practice of Statistics*. Macmillan, 2016.

[25] M. J. Blanca, R. Alarcón, J. Arnau, R. Bono, and R. Bendayan, "Non-normal data: Is anova still a valid option?" *Psicothema*, vol. 29, no. 4, pp. 552–557, 2017.

[26] B. J. Feir-Walsh and L. E. Toothaker, "An empirical comparison of the anova f-test, normal scores test and kruskal-wallis test under violation of assumptions," *Educational and Psychological Measurement*, vol. 34, no. 4, pp. 789–799, 1974.

[27] A. Bathke, "The anova f test can still be used in some balanced designs with unequal variances and nonnormal data," *Journal of Statistical Planning and Inference*, vol. 126, no. 2, pp. 413–422, 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378375803002787

[28] E. Veach, *Robust Monte Carlo methods for light transport simulation*. Stanford University PhD thesis, 1997, vol. 1610.

[29] S.-L. Jan and G. Shieh, "Sample size determinations for welch's test in one-way heteroscedastic anova," *British Journal of Mathematical and Statistical Psychology*, vol. 67, no. 1, pp. 72–93, 2014.

[30] W. Jakob, "Mitsuba renderer," 2010, http://www.mitsuba-renderer.org.

[31] T. Hachisuka and H. W. Jensen, "Robust adaptive photon tracing using photon path visibility," *ACM Trans. Graph.*, vol. 30, no. 5, Oct. 2011. [Online]. Available: https://doi.org/10.1145/2019627.2019633

[32] J. Vorba, O. Karlík, M. Šik, T. Ritschel, and J. Křivánek, "On-line learning of parametric mixture models for light transport simulation," *ACM Trans. Graph.*, vol. 33, no. 4, Jul. 2014. [Online]. Available: https://doi.org/10.1145/2601097.2601203

[33] T. Müller, M. Gross, and J. Novák, "Practical path guiding for efficient light-transport simulation," in *Computer Graphics Forum*, vol. 36, no. 4, 2017, pp. 91–100.

[34] T. Müller, B. Mcwilliams, F. Rousselle, M. Gross, and J. Novák, "Neural importance sampling," *ACM Trans. Graph.*, vol. 38, no. 5, Oct. 2019. [Online]. Available: https://doi.org/10.1145/3341156

[35] T. Müller, F. Rousselle, A. Keller, and J. Novák, "Neural control variates," *ACM Trans. Graph.*, vol. 39, no. 6, Nov. 2020. [Online]. Available: https://doi.org/10.1145/3414685.3417804

[36] I. Kondapaneni, P. Vevoda, P. Grittmann, T. Skřivan, P. Slusallek, and J. Křivánek, "Optimal multiple importance sampling," *ACM*

Trans. Graph., vol. 38, no. 4, Jul. 2019. [Online]. Available: https://doi.org/10.1145/3306346.3323009

[37] P. Grittmann, I. Georgiev, and P. Slusallek, "Correlation-aware multiple importance sampling for bidirectional rendering algorithms," *Computer Graphics Forum*, vol. 40, no. 2, pp. 231–238, 2021.

[38] F. Su, S. Li, and G. Wang, "Spcbpt: Subspace-based probabilistic connections for bidirectional path tracing," *ACM Trans. Graph.*, vol. 41, no. 4, jul 2022. [Online]. Available: https://doi.org/10.1145/3528223.3530183

[39] R. L. Cook, "Stochastic sampling in computer graphics," *ACM Trans. Graph.*, vol. 5, no. 1, p. 51–72, jan 1986. [Online]. Available: https://doi.org/10.1145/7529.8927

[40] J. Rigau, M. Feixas, and M. Sbert, "Refinement Criteria Based on f-Divergences," in *Eurographics Workshop on Rendering*. The Eurographics Association, 2003.

Zehui Lin obtained a Master Degree from Graphics & Interactive Technology Lab., Peking University in 2021. His research interests include computer graphics and virtual reality technology.

Chenxiao Hu is currently an undergraduate student at School of Computer Science, Peking University. His research interests include realistic rendering and real-time rendering.

Jinzhu Jia is currently an Associate Professor of Dept. of Biostatistics and Center for Statistical Science, Peking University. His research interests include high-dimensional statistical inference, statistical machine learning, causal inference and biostatistics.

Sheng Li is currently a Professor at the School of Computer Sciences, Peking University. He serves as a member of the Graphics & Interactive Technology Lab. and has published over 40 refereed papers in prestigious journals and conferences, including ACM TOG, IEEE TVCG, Computer Graphics Forum, etc. His research interests include virtual reality, rendering, physical simulation and animation. He is a member of ACM and IEEE.