Robust Sample Budget Allocation for MIS

László Szirmay-Kalos[1] & Mateu Sbert[2]

[1] Budapest University of Technology and Economics, Dept. of Control Eng. and Inf. Tech, Hungary
[2] University of Girona, Spain

Abstract

Multiple Importance Sampling (MIS) combines several sampling techniques. Its weighting scheme depends on how many samples are generated with each particular method. This paper examines the optimal determination of the number of samples allocated to each of the combined techniques taking into account that this decision can depend only on a relatively small number of previous samples. The proposed method is demonstrated with the combination of BRDF sampling and Light source sampling, and we show that due to its robustness, it can outperform the theoretically more accurate approaches.

Keywords: Multiple importance sampling, Monte Carlo.

1. Introduction and previous work

Multiple Importance Sampling (MIS) [VG95, Vea97] has been proven efficient in many Monte Carlo rendering algorithms. It is able to preserve the advantages of the combined techniques and requires only the calculation of the pdf of all methods when a sample is generated with one particular method. The weighting scheme depends on the pdfs of the individual techniques and also on the number of samples generated with each of them.

This paper proposes a simple adaptive technique to automatically determine the sampling budgets of the combined methods based on the statistics of previous samples. Although the proposed method cannot guarantee a theoretical optimum, its advantages are the simplicity and the robustness, and that it has no additional hyper-parameters to tune.

MIS has been applied in a number of rendering algorithms, and its variance is extensively studied [VG95]. Several estimators have been proposed that are better than balance heuristics with equal sample budget [SHSK16, HS14, HKD14, SH17, SHSK18]. Lu et al. [LPG13] proposed an adaptive algorithm for environment map illumination, using the Taylor series approximation of the variance. In [EMLB15a, EMLB15b] strategies and analysis were given assuming equal number of samples. Sbert et al. [SHSKE18] considered the cost associated with the sampling strategies, and in [SHSK19] they obtained the solution by optimizing the variance using the Newton-Raphson method. In [VHH*19] the Kullback-Leibler divergence replaced the variance. The variance of an importance sampling estimator has been shown to be equal to a Chi-square divergence [CMO08, MIG17, MMR*19, SE22], which in turn can be approximated by a Kullback-Leibler divergence up to the second order [NN14] and with gradients differing only in a weighting term [MMR*19]. The optimal sample budget has also been targeted with neural networks [MMR*19, MBPG20]. Recently, a theoretical formula has been elaborated for the weighting functions [KVG*19]. In [SE22] the balance heuristic estimator was generalized by decoupling the weights from the sampling rates, and implicit solutions for the optimal case were given.

These techniques offer lower variance and theoretically outperform MIS with equal number of samples when estimating the integral of the rendering equation. However, equations determining the optimal weighting and sample budget require the knowledge of the integrand and typically numerical methods to solve them. In computer graphics, this integrand is not analytically available, so previous samples should be used for the approximation, which also introduces error in the computation. Thus, it is not guaranteed that a theoretically superior estimator also performs better in practice. In this paper we focus on simple equations and robust estimations, and show that they can outperform more involved approaches.

2. The MIS estimator

Suppose we wish to estimate the value of integral $I = \int f(x)\mathrm{d}x$ and have m proposal pdfs $p_i(x)$ to generate samples in the domain of the integral. The MIS estimator [VG95] has the following expression:

$$F = \sum_{i=1}^{m} \frac{1}{n_i} \sum_{j=1}^{n_i} w_i(X_{ij}) \frac{f(X_{ij})}{p_i(X_{ij})}, \tag{1}$$

where the weights w_i satisfy normalization: $\sum_{i=1}^{m} w_i(x) = 1$.

In this combination scheme, sampling method i uses probability density function $p_i(x)$ to generate n_i number of random samples X_{ij}, $(j = 1,\dots,n_i)$. The total number of samples is $\sum_{i=1}^{m} n_i = N$. Integral estimator F is unbiased, as its expected value is equal to integral I:

$$E[F] = \sum_{i=1}^{m} \frac{1}{n_i} \sum_{j=1}^{n_i} \int \frac{w_i(x)f(x)}{p_i(x)} p_i(x)\mathrm{d}x = \int f(x)\mathrm{d}x. \quad (2)$$

To express the variance of the estimator, let us define F_{ij} as

$$F_{ij} = w_i(X_{ij}) \frac{f(X_{ij})}{p_i(X_{ij})}. \quad (3)$$

For a fixed method i, the estimators F_{ij} are independent and identically distributed random variables with expected value μ_i:

$$\mu_i = E_{p_i}[F_{ij}] = \int \frac{w_i(x)f(x)}{p_i(x)} p_i(x)\mathrm{d}x = \int w_i(x)f(x)\mathrm{d}x. \quad (4)$$

The variance of F_{ij} is

$$\sigma_i^2 = E[F_{ij}^2] - E^2[F_{ij}] = \int \frac{w_i^2(x)f^2(x)}{p_i(x)}\mathrm{d}x - \mu_i^2. \quad (5)$$

2.1. Balance heuristics

The variance of the estimator depends on the combination weights $w_i(x)$, so we find them minimizing the variance. One step into this direction is to heuristically propose the algebraic form of the weight functions. Balance heuristic states that the weight of method i at domain point x should be proportional to the density of samples generated by method i in this point:

$$w_i(x) = \frac{\alpha_i p_i(x)}{p(\alpha, x)} \quad (6)$$

where $\alpha_i = n_i/N$ is the fraction of the samples allocated to method i, and $p(\alpha, x)$ is the mixture density:

$$p(\alpha, x) = \sum_{k=1}^{m} \alpha_k p_k(x).$$

Substituting this weighting function into the MIS estimator formulas, we obtain the balance heuristics estimator:

$$F = \frac{1}{N} \sum_{i=1}^{m} \sum_{j=1}^{n_i} \frac{f(X_{i,j})}{p(\alpha, X_{i,j})}. \quad (7)$$

The variance of the balance heuristics estimator is

$$V[F] = \frac{1}{N} \sum_{i=1}^{m} \left(\int \frac{\alpha_i p_i(x) f^2(x)}{p^2(\alpha, x)}\mathrm{d}x - \frac{\mu_i^2}{\alpha_i} \right) = \frac{1}{N} \sum_{i=1}^{m} \frac{\sigma_i^2}{\alpha_i} \quad (8)$$

where σ_i^2 and μ_i are also functions of fractions $\alpha_1, \ldots, \alpha_m$.

Having fixed the algebraic form of the weight function, the task is to find the close to optimal fractions α_i. The variance of the estimator depends on the allocation of samples to different methods, α_i, in a complicated way. If the integrals of μ_i and the variances σ_i^2 could be easily computed as functions of α_i, then this would be a vector valued optimization process with the constraint that the sum of samples must be equal to the total sample budget, i.e. $\sum_{k=1}^{m} \alpha_k = 1$. However, there are a few additional issues that need to be considered:

- The integrals cannot be computed analytically, but must be estimated from the available samples. This uncertainty may significantly affect the goodness of the final sample numbers.
- The optimization process should be fast, free from hyper-parameters, and should not introduce too high overhead.

Unfortunately, the direct optimization for the variance does not meet these requirements. We need heuristics that may not guarantee the lowest variance but lead to simple and robust estimation.

3. The heuristics for the weights

To optimize the variance in Eq. 8 under the constraint that the sum $\sum_k \alpha_k$ of weights must be equal to 1, the following derivatives must be made equal to zero according to Lagrange multipliers:

$$\frac{\partial}{\partial \alpha_k} \left(V[F] - \lambda \left(1 - \sum_{i=1}^{m} \alpha_i \right) \right) = 0, \quad (9)$$

which leads to the following equations for $k = 1, \ldots, m$ [SE22]:

$$\frac{\sigma_k^2}{\alpha_k^2} + 2\frac{\mu_k^2}{\alpha_k^2} - 2\sum_{i=1}^{m} \mu_i \int \frac{p_i(x)p_k(x)f(x)}{p^2(\alpha, x)}\mathrm{d}x = \lambda. \quad (10)$$

Thus, the expression on the left hand side is the same for different $k = 1, \ldots, m$ indices. Our heuristic is that here the dominant term is μ_k^2/α_k^2 and the other terms do not significantly modify the selection of α_k parameters. This heuristic leads to the requirement that

$$\frac{\mu_k}{\alpha_k} = \int \frac{p_k(x)f(x)}{p(\alpha, x)}\mathrm{d}x = E_{p_k}\left[\frac{f(X)}{p(\alpha, X)}\right] \quad (11)$$

should be constant. If the mixture density is fair enough, then $p(\alpha, X)$ is approximately proportional to integrand $f(X)$, thus their ratio is constant everywhere. The average of these roughly constant ratios has only a small fluctuation no matter what density is used to draw the samples.

4. Combination of light source sampling and BRDF sampling

In order to demonstrate the proposed scheme, we consider the combination of light source sampling and BRDF sampling using probability densities $p_1(X)$ and $p_2(X)$, respectively. The mixture of the two original proposal pdfs is:

$$p(\alpha, x) = \alpha p_1(x) + (1 - \alpha)p_2(x). \quad (12)$$

The primary MIS estimator is:

$$F = \frac{f(X)}{p(X)} = \frac{f(X)}{p(\alpha, X)}. \quad (13)$$

Equivalently, the weighting scheme is

$$w_1(X) = \frac{\alpha p_1(X)}{p(\alpha, X)}, \quad w_2(X) = \frac{(1-\alpha)p_2(X)}{p(\alpha, X)}.$$

Finally, the μ_i/α_i terms are

$$\frac{\mu_1(\alpha)}{\alpha} = E_p\left[\frac{f(X)p_1(X)}{p^2(\alpha, X)}\right], \quad \frac{\mu_2(\alpha)}{1-\alpha} = E_p\left[\frac{f(X)p_2(X)}{p^2(\alpha, X)}\right].$$

According to our heuristics, the optimal α can be obtained by solving the following equation:

$$C(\alpha) = \frac{\mu_1(\alpha)}{\alpha} - \frac{\mu_2(\alpha)}{1-\alpha} = E_p\left[\frac{f(X)(p_1(X) - p_2(X))}{p^2(\alpha, X)}\right] = 0. \quad (14)$$

Note that this is similar to the gradient of the method aiming at minimizing the Kullbach-Leibler divergence [VHH*19], which terminates when the gradient is zero, i.e. when this equation is solved. It means that our heuristics is as justified as the variance can be replaced by the Kullbach-Leibler divergence in optimum budget allocation. However, the two methods have different iteration schemes and convergence speeds. We use the Newton-Raphson method for the numerical solution, when we also need the derivative:

$$\frac{\mathrm{d}C}{\mathrm{d}\alpha} = -2E_p\left[\frac{f(X)(p_1(X)-p_2(X))^2}{p^3(\alpha,X)}\right].$$

Expectation values are estimated from previous samples generated with mixture density $p(x)$.

5. Results

In order to test the proposed method, we render the classic scene of Veach with combined light source and BRDF sampling. The shiny rectangles have streched-Phong BRDF [NNSK99] with shininess parameters 500, 1000, 5000, and 10000, respectively. The four spherical light sources emit the same power.

We allocated 100 samples per pixel organized in 10 batches. The process starts with 5 BRDF and 5 light source samples per pixel, and the per-pixel α weights are updated at the end of each batch. Figures 2 and 1 show the rendered images together with the obtained α maps. In addition to the original sampling techniques and equal count MIS, we also compared the proposed approach to [SHSK19], [LPG13], and to the replacement of the variance by the Kullbach-Leibler divergence as proposed in [VHH*19]. Figure 3 depicts the RMSE values obtained as averages of 30 rendering runs.

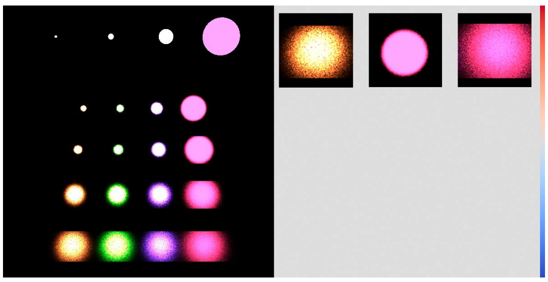

Figure 1: *Equal sample count MIS rendering with 100 rays per pixel used as a baseline. The average RMSE is 318.*

6. Conclusions

In this paper we investigated the problem of determining the sample budgets for techniques combined by Multiple Importance Sampling, with a special focus on the simplicity and robustness. The proposed method calculates the weighting factors of two combined techniques iteratively in parallel with sample generation. We have shown that such adaptation can outperform methods directly aiming at the optimization of the variance because the total error contribution of both the variance and the unreliability of the optimal parameter estimation is reduced in our case.

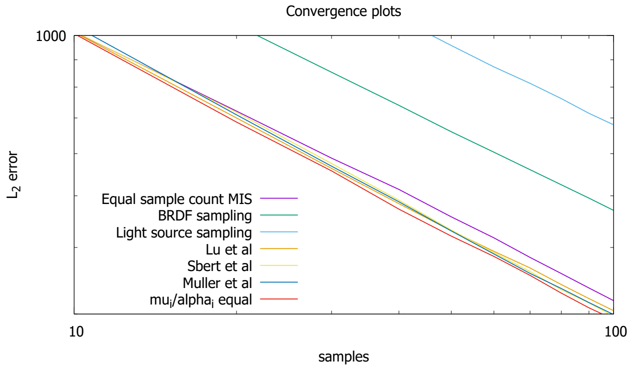

Figure 3: *RMSE as functions of the number of samples per pixel.*

Acknowledgements

This work was supported by OTKA K-124124 and by PID2019-106426RB-C31 funded by MCIN/AEI/ 10.13039/5011.

References

[CMO08] CORNEBISE J., MOULINES E., OLSSON, J.: Adaptive methods for sequential importance sampling with application to state space models. *Stat. Comput.* **2008**, *18*, 461–480. 1

[EMLB15a] ELVIRA V., MARTINO L., LUENGO D., BUGALLO M. F.: Efficient multiple importance sampling estimators. *IEEE Signal Process. Lett. 22*, 10 (2015), 1757–1761. doi:10.1109/LSP.2015.2432078. 1

[EMLB15b] ELVIRA V., MARTINO L., LUENGO D., BUGALLO M. F.: Generalized Multiple Importance Sampling. *ArXiv e-prints* (Nov. 2015). arXiv:1511.03095. 1

[HKD14] HACHISUKA T., KAPLANYAN A. S., DACHSBACHER C.: Multiplexed metropolis light transport. *ACM Trans. Graph. 33*, 4 (2014), 100:1–100:10. doi:10.1145/2601097.2601138. 1

[HS14] HAVRAN V., SBERT M.: Optimal Combination of Techniques in Multiple Importance Sampling. In *VRCAI '14* (2014), pp. 141–150. doi:10.1145/2670473.2670496. 1

[KVG*19] KONDAPANENI I., VEVODA P., GRITTMANN P., SKŘIVAN T., SLUSALLEK P., KŘIVÁNEK J.: Optimal multiple importance sampling. *ACM Trans. Graph. 38*, 4 (2019). doi:10.1145/3306346.3323009. 1

[LPG13] LU H., PACANOWSKI R., GRANIER X.: Second-Order Approximation for Variance Reduction in Multiple Importance Sampling. *Computer Graphics Forum 32*, 7 (2013), 131–136. doi:10.1111/cgf.12220. 1, 3, 4

[MIG17] MÍGUEZ, J.: On the performance of nonlinear importance samplers and population Monte Carlo schemes. In Proc. of the 22nd International Conference on Digital Signal Processing (DSP), 2017. 1

[MBPG20] MURRAY D., BENZAIT S., PACANOWSKI R., GRANIER X.: On Learning the Best Local Balancing Strategy. In *Eurographics 2020 - Short Papers* (2020), doi:10.2312/egs.20201009. 1

[MMR*19] MÜLLER T., MCWILLIAMS B., ROUSSELLE F., GROSS M., NOVÁK J.: Neural importance sampling. *ACM Trans. Graph. 38*, 5 (2019). doi:10.1145/3341156. 1

[NN14] NIELSEN F., NOCK R.: On the chi square and higher-order chi distances for approximating f-divergences. *IEEE Signal Processing Letters 21*, 1 (2014), 10–13. doi:10.1109/lsp.2013.2288355. 1

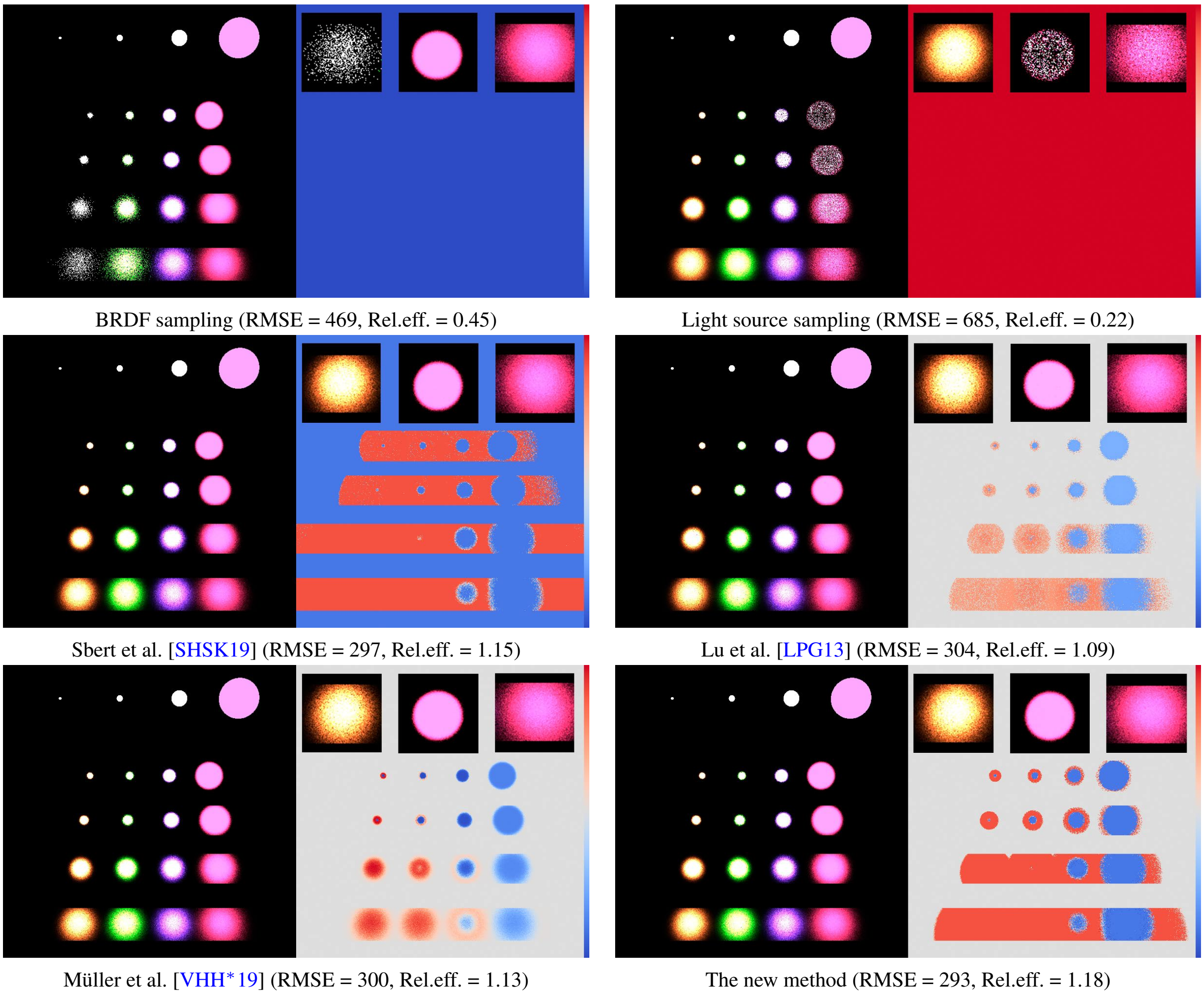

Figure 2: *Comparison of MIS weighting schemes. The left part is the image rendered with 100 rays per pixel, the right part is weight α of the light source sampling. The RMSE and relative efficiency with respect to the equal sample count MIS are computed as the average of 30 independent executions. The relative efficiency is the square of the ratio of equal sample count MIS RMSE and the method's RMSE. The [0,1] interval of possible α values is visualized by the color bar.*

[NNSK99] NEUMANN L., NEUMANN A., SZIRMAY-KALOS L.: Compact metallic reflectance models. *Computer Graphics Forum 18*, 3 (1999), 161–172. 3

[SE22] SBERT M., ELVIRA V.: Generalizing the balance heuristic estimator in multiple importance sampling. *Entropy 24*, 2 (2022). doi:10.3390/e24020191. 1, 2

[SH17] SBERT M., HAVRAN V.: Adaptive multiple importance sampling for general functions. *The Visual Computer* (2017), 1–11. doi:10.1007/s00371-017-1398-1. 1

[SHSK16] SBERT M., HAVRAN V., SZIRMAY-KALOS L.: Variance analysis of multi-sample and one-sample multiple importance sampling. *Computer Graphics Forum 35*, 7 (2016), 451–460. doi:10.1111/cgf.13042. 1

[SHSK18] SBERT M., HAVRAN V., SZIRMAY-KALOS L.: Multiple importance sampling revisited: Breaking the bounds. 13. URL: https://asp-eurasipjournals.springeropen.com/articles/10.1186/s13634-018-0531-2. 1

[SHSK19] SBERT M., HAVRAN V., SZIRMAY-KALOS L.: Optimal Deterministic Mixture Sampling. In *Eurographics 2019 - Short Papers* (2019) doi:10.2312/egs.20191018. 1, 3, 4

[SHSKE18] SBERT M., HAVRAN V., SZIRMAY-KALOS L., ELVIRA V.: Multiple importance sampling characterization by weighted mean invariance. *The Visual Computer* (2018). doi:10.1007/s00371-018-1522-x. 1

[Vea97] VEACH E.: *Robust Monte Carlo Methods for Light Transport Simulation*. PhD thesis, Stanford University, 1997. 1

[VG95] VEACH E., GUIBAS L. J.: Optimally Combining Sampling Techniques for Monte Carlo Rendering. In *SIGGRAPH '95* (1995), pp. 419–428. doi:10.1145/218380.218498. 1

[VHH*19] VORBA J., HANIKA J., HERHOLZ S., MUELLER T., KRIVANEK J., KELLER A.: Path guiding in production. In *ACM SIGGRAPH 2019 Courses* (2019), doi:10.1145/3305366.3328091. 1, 3, 4