Automatic Parameter Control for Metropolis Light Transport

Károly Zsolnai[1] and László Szirmay-Kalos[1]

Department of Control Engineering and Information Technology, Budapest University of Technology, Hungary
zsolnai@iit.bme.hu, szirmay@iit.bme.hu

Abstract. Sophisticated global illumination algorithms usually have several control parameters that need to be set appropriately in order to obtain high performance and accuracy. Unfortunately, the optimal values of these parameters are scene dependent, thus their setting requires significant care and is usually based on trial and error. This cumbersome process is unacceptable in production rendering where robust methods are needed that can render a larger set of images, e.g. frames of a movie, without lengthy experimentation on the parameter setting. To address this problem, this paper presents a method to automatically control the large step probability parameter of Primary Sample Space Metropolis Light Transport (PSSMLT). The method does not require extra computation time or pre-processing, and runs in parallel with the initial phase of the rendering method. During this phase, it gathers statistics from the process and the parameters are obtained according to these statistics for the remaining part of sample generation. We show that the theoretically proposed values are close to the manually found optimum for several complex scenes.

1 Introduction

The performance of global illumination rendering algorithms [DBB03] has been dramatically increased recently, thus these methods have become viable alternatives in production rendering and also in real-time applications. The performance increase is the result of sophisticated sampling methods, reuse of information gathered with other samples, and the exploitation of the graphics hardware. However, high performance rendering algorithms often come with a larger set of control parameters that need to be set by the artist before starting the rendering process. For example, when rendering with the photon mapping method [Jen01], the artist should set the number of caustic photons, the number of regular photons, number of final gathering rays, and the number of photon hits that are considered to be in the same local neighborhood. Unfortunately, the optimal values of the control parameters are scene dependent and a careless setting would significantly slow down the convergence or could even result in images of severe artifacts in case of biased methods.

In practice, parameter setting is based on previous experience and a trial and error approach, but this is unacceptable in production rendering where a large number of images should be rendered without visible artifacts under severe time constraints, and also in virtual reality system when the scene may change in time in a way that is not anticipated by the designer. In these cases robust rendering algorithms are needed that require no manual parameter setting and can deliver high quality images in rendering times comparable to manually controlled methods.

Metropolis Light Transport (MLT) [VG97] is known to be a robust approach that can efficiently handle a large variety of lighting effects and scenes. The power of MLT comes from the Metropolis-Hastings sampling that — unlike other importance sampling methods generating samples independently from a prescribed density — explores important light path regions in the scene by mutating previous light paths and evaluating the importance of the mutations. Metropolis sampling establishes a Markov chain on the space of light paths, and by following an almost arbitrarily selected mutation strategy and a carefully defined rejection scheme based on an importance function, it promises that the asymptotic distribution of the Markov chain mimics the prescribed importance function. From the point of view of parameter control in global illumination algorithms, the free choice in MLT is the mutation strategy. There are many possibilities from which the artist should choose, and the decision will be a crucial factor of the performance of the rendering, i.e. the image quality obtained in the available rendering time.

In this paper we build on the Primary Sample Space MLT algorithm and propose an automatic control mechanism that "learns" the properties of the current scene in the initial phase of the algorithm and then controls mutations in the primary sample space in order to maximize the performance. In Section 2, MLT is summarized briefly, then Section 3 presents our simple statistical approach. Section 4 demonstrates that the control parameters proposed by the automatic method are very close to the real optimum.

2 Previous work on MLT

Metropolis sampling was used to solve the global illumination problem first by Veach et al. [VG97]. Global illumination requires a separate quadrature for each pixel:

$$\Phi_i = \int\limits_{\mathcal{P}} W_i(\mathbf{z}) F(\mathbf{z}) \mathrm{d}\mathbf{z},$$

where Φ_i is the power associated with pixel i, $\mathcal{P}$ is the infinite-dimensional path space, $\mathbf{z}$ is a single light path having contribution $F(\mathbf{z})$, and $W_i(\mathbf{z})$ is the measuring function that is 1 if $\mathbf{z}$ contributes to pixel i and zero otherwise.

Monte Carlo methods take M path space samples $\mathbf{z}_1, \ldots, \mathbf{z}_M$ the with probability density $p(\mathbf{z})$ and approximate the integral as a weighted sum of the integrand at these samples:

$$\Phi_i \approx \frac{1}{M} \sum_{i=1}^{M} \frac{W_i(\mathbf{z}_i) F(\mathbf{z}_i)}{p(\mathbf{z}_i)}.$$

According to the concept of importance sampling, this estimate has a low variance if density $p(\mathbf{z}_i)$ is at least approximately proportional to the integrand.

The Metropolis method is a sampling procedure that generates samples mimicking a given importance function $I(\mathbf{z})$, i.e. with probability density

$$p(\mathbf{z}) = \frac{I(\mathbf{z})}{\int\limits_{\mathcal{P}} I(\mathbf{z}) \mathrm{d}\mathbf{z}} = \frac{I(\mathbf{z})}{b}, \quad \text{where} \quad b = \int\limits_{\mathcal{P}} I(\mathbf{z}) \mathrm{d}\mathbf{z}.$$

To do so, it constructs a Markov process whose stationary distribution is proportional to a prescribed importance function $I(\mathbf{z})$ [MRR*53]. The next state $\mathbf{z}_{i+1}$ of this process is found by letting an almost arbitrary *tentative transition function* $T(\mathbf{z}_i \to \mathbf{z}_t)$ generate a *tentative sample* $\mathbf{z}_t$ which is either accepted as the real next state or rejected making the next state equal to the actual state. The decision uses the *acceptance probability* $a(\mathbf{z}_i \to \mathbf{z}_t)$ that expresses the increase of the scalar contribution function:

$$a(\mathbf{z}_i \to \mathbf{z}_t) = \min \left\{ \frac{I(\mathbf{z}_t) \cdot T(\mathbf{z}_t \to \mathbf{z}_i)}{I(\mathbf{z}_i) \cdot T(\mathbf{z}_i \to \mathbf{z}_t)}, 1 \right\}.$$

It would be tempting to define the importance function directly as the integrand, but this is not feasible because of the following three reasons. Normalization constant b involves the integral of the importance function, and if it was the integrand, the computation of the normalization constant would already require the solution of the original problem. The Markov process needs some time to reach its stationary state, which means that it is better to assign a single Markov process to all pixels than to assign a separate process to each pixel. Finally, the contribution of path $F(\mathbf{z})$ is typically not a scalar, but a spectrum, while both the importance function and the probability density must be scalars.

In order to solve these problems, importance function I can be defined as the scalar contribution to any pixel of the screen, that is, I will be the luminance of path contribution $F(\mathbf{z})$ and is independent of the pixel measuring function.

Veach recognized that it is also worth using the rejected samples since they also provide illumination information. Note that a tentative sample is accepted with probability a, while the original sample is kept with probability $1-a$. Replacing this random variable by its mean, both locations can be contributed but the contributions of the tentative sample and the old sample should be weighted with a and $1-a$, respectively.

Summarizing, the pseudo-code of the MLT algorithm is as follows:

Generate path seeds
Approximate $b = \int I(\mathbf{z})\mathrm{d}\mathbf{z}$ from the seeds
Find $\mathbf{z}_1$ from the seeds using $I(\mathbf{z})$
for $i = 1$ **to** M **do**
  Based on $\mathbf{z}_i$, sample a tentative point $\mathbf{z}_t$ using $T(\mathbf{z}_i \to \mathbf{z}_t)$
  $a(\mathbf{z}_i \to \mathbf{z}_t) = \min\left\{\frac{I(\mathbf{z}_t)\cdot T(\mathbf{z}_t \to \mathbf{z}_i)}{I(\mathbf{z}_i)\cdot T(\mathbf{z}_i \to \mathbf{z}_t)}, 1\right\}$
  Select pixel j to which $\mathbf{z}_i$ contributes
  $\Phi_j \mathrel{+}= \frac{b}{M} \cdot \frac{F(\mathbf{z}_i)}{I(\mathbf{z}_i)} \cdot (1 - a(\mathbf{z}_i \to \mathbf{z}_t))$
  Select pixel k to which $\mathbf{z}_t$ contributes
  $\Phi_k \mathrel{+}= \frac{b}{M} \cdot \frac{F(\mathbf{z}_t)}{I(\mathbf{z}_t)} \cdot a(\mathbf{z}_i \to \mathbf{z}_t)$
  // *accept with probability* $a(\mathbf{z}_i \to \mathbf{z}_t)$
  Generate random number r in $[0, 1]$.
  if $r < a(\mathbf{z}_i \to \mathbf{z}_t)$ **then** $\mathbf{z}_{i+1} = \mathbf{z}_t$
  else $\mathbf{z}_{i+1} = \mathbf{z}_i$
endfor

When a particular MLT algorithm is designed, we have to specify the tentative transition function that generates a path by mutating another path. There are several criteria that need to be taken into account during the tentative transition function definition:

– The mutation strategy should guarantee that the Markov chain is *ergodic*, i.e. it has an asymptotic distribution that is independent of the initial state. To satisfy this requirement, all light paths of non-zero contribution should be given a chance to be generated as a tentative sample sooner or later. Veach proposed elementary path modification operators that add and remove rays, or modify interaction points in an already existing path. Pauly et al. [PKK00] extended these operators to handle participating media as well. By giving chance to the full re-generation of a path, the ergodicity condition can be met. From the point of view of parameter control in global illumination algorithms, the free choice in MLT is the mutation strategy from the alternatives meeting the requirement of ergodicity. There are many possibilities from which the artist should choose, and the decision will be a crucial factor of the performance of the rendering, i.e. the image quality obtained in the available rendering time.
– Unlike other Monte Carlo algorithms, the Metropolis algorithm generates not statistically independent, but correlated samples, which can increase the error [SKDP99,APSS01]. Thus, a secondary requirement for mutation strategies is to keep the correlation low, which means that the possibility of rejected samples should be reduced.

– Metropolis sampling converges to the desired probability distribution, but at the beginning of the process the samples are not selected with the required probability, which introduces some error in the estimation. This error is called as the *start-up bias* [SKDP99]. The original MLT algorithm uses the following solution for the problem: in a preprocessing phase random samples are generated and the initial seed of the Metropolis algorithm is selected from this random population with a probability that is proportional to the scalar contribution function. Since in this case even the first sample follows the desired distribution, the start-up bias problem is eliminated in a statistical sense, i.e. it is converted to noise.

Directly mutating paths in path space is difficult to implement, and the reduction of rejection rate and thus the correlation seems to be not feasible. Furthermore, the application developer has too many control parameters that should be set to define when and what path property must be modified by the tentative transition function.

The large search space of control parameters is effectively limited and the reduction of rejection rate is also successfully addressed by the method of Primary Sample Space MLT (PSSMLT) [KSKAC02], which executes the perturbations in the space of uniformly distributed random numbers, called the *primary sample space*. PSSMLT is based on the recognition that any light path generator algorithm is a mapping between the primary sample space and the path space, and if importance sampling is involved in the path generation then more important regions are represented by larger volume in primary sample space. Thus, making uniform mutations in primary sample space automatically reduces the path perturbation size at important regions and increases it in unimportant regions. We can also say that the involved importance sampling in path generation (e.g. BRDF or light source sampling) does one part of the job of optimal importance sampling, which is further improved by the Metropolis-Hastings scheme by features that cannot be included in BRDF and light source sampling. In this sense PSSMLT is a special type of random number generator that can be plugged in an arbitrary Monte Carlo method. In classical Monte Carlo methods the random number generator provides statistically independent uniformly distributed random numbers, or k-uniform k-dimensional vectors in primary sample space. In PSSMLT a feed-back is introduced that transports back the computed importance and makes the primary samples non-uniform but mimic the importance. In PSSMLT a simple way of ensuring ergodicity is to generate a primary space sample point from scratch, independently of the current paths with some positive probability. Such independent samples are often called *large steps*.

Theoretically, any positive large step probability makes the Markov process ergodic, but its selection affects convergence and thus rendering performance, thus it must be carefully set (Figure 1). Intuitively, we can say that in scenes of difficult light conditions where only a small part of the path space has non-zero contribution, most of the large steps lead to zero or negligible contributions, which are rejected, so large step probability should be kept low. On the other hand, in open scenes having little occlusion and relatively constant contribution

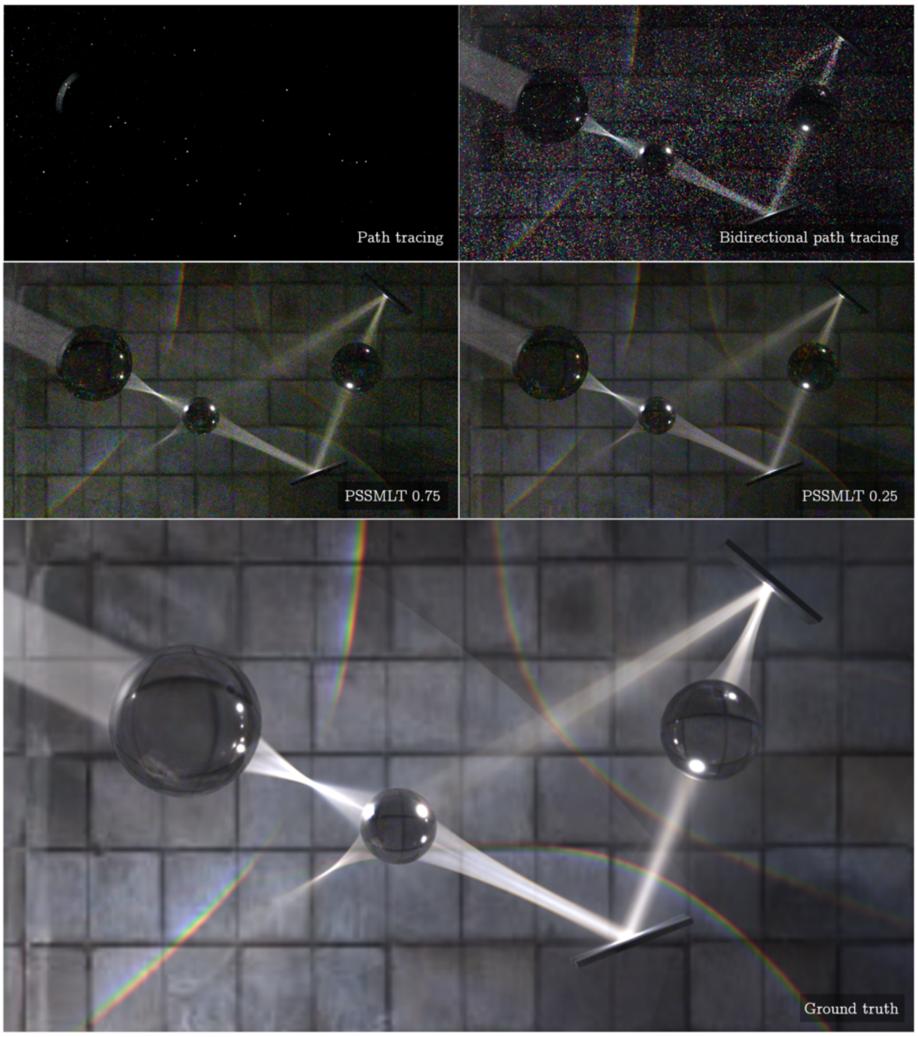

Fig. 1. Images of the Spheres scene after 15 minutes with different algorithms show that Metropolis Light Transport is the most efficient way to overcome this difficult specular transport setting. The choice of the large mutation probability has also a pronounced impact on the render quality.

in the whole path space, increasing the large step probability will speed up the exploration of the total path space and thus reduces the start up bias. However, the optimal selection of the large step probability is not an easy problem.

This paper addresses this issue and proposes an adaptive method to control the large step probability. The model must be based on parameters that are easy to estimate while starting the algorithm.

3 The proposed mutation control method

The Monte Carlo estimation formula can also be regarded as a piecewise constant approximation of the original integrand, where the path space domain is decomposed to finite regions $\Delta z_1, \ldots, \Delta z_M$ and the integrand is supposed to be constant in each regions:

$$\Phi_i \approx \frac{1}{M} \sum_{i=1}^{M} \frac{W_i(\mathbf{z}_i) F(\mathbf{z}_i)}{p(\mathbf{z}_i)} = \sum_{i=1}^{M} W_i(\mathbf{z}_i) F(\mathbf{z}_i) \Delta z_i$$

where

$$\Delta z_i = \frac{1}{M p(\mathbf{z}_i)}.$$

In this respect, Monte Carlo sampling is the exploration of the integration domain with finite samples and the decomposition of the domain to as many regions as possible. If Metropolis sampling works in the primary sample space, the transformation of the path space to the primary sample space already executes importance sampling offered by the path building strategy, e.g. path tracing or bi-directional path tracing. If the path building strategy is efficient for the particular scene, the importance sampling is close to optimal, which means that the path contribution divided by the density is close to constant where it is non-zero. Note that zero contribution regions keep their zero contribution, so the path building strategy can only flatten the contribution of non-zero regions.

Metropolis sampling should handle the residual variation of the integrand after transforming the domain according to the path building strategy. If the non-zero contribution regions have low variation in the primary sample space, the efficiency of Metropolis depends on how quickly it explores this space since all regions of this space have similar contribution. However, if there is a significant residual variation after transformation, Metropolis should still focus on the high contribution regions, according to the principles of importance sampling. Thus, in the first case the efficiency can be characterized by just the size of the explored domain, but in the second case, the integrand values should also be taken into account.

The Metropolis method is based on rejection sampling, so it can happen that a sample is generated many times, which reduces the number of regions and hence the accuracy of the approximation. The efficiency of the Metropolis method in obtaining samples that are different from the previous ones can be characterized by the following three measures:

- Small step efficiency η_s that is the average probability that a small perturbation is accepted.
- Large step efficiency η_l that is the average probability that a large perturbation is accepted.
- The large step non-zero sample probability η_0 is the average probability that a large perturbation generates a non-zero contribution sample. As large steps are independent and the primary sample space has unit size, this probability is the volume of the path domain where paths have non-zero contribution.

Note that these parameters can be easily estimated in parallel to the running of the algorithm and can be obtained from a relatively few number of samples.

These values are not independent of each other when measured on a particular scene. As Metropolis sampling drives samples towards high importance regions where small mutations may cause smaller importance decrease, the relation $\eta_s \geq \eta_l$ generally holds. On the other hand, the rejection of a large mutation can be due to a mutation to a zero contribution point or to a point having positive but smaller contribution than the current state, thus $\eta_l \leq \eta_0$.

The ratio η_l/η_0 characterizes the efficiency of the importance sampling of the path building strategy. If the integrand is flat in the nonzero contribution regions, then this ratio is close to 1. If it is close to one, the integrand still has a significant variation which needs to be compensated by Metropolis sampling.

To explore the path space efficiently, both small and large perturbations must be applied to the path space walk. Small mutations propose new primary sample points that are in a neighborhood around the point defining the current state. Large mutations lead to an arbitrary point of the unit volume primary sample space. The type for the next mutation is given by the *large step probability* p_l. From a theoretical point of view, any positive value will make the algorithm converge to the exact solution in the limit. So in the stationary case, all positive large step probabilities are equivalent. However, the convergence rate is affected, so, in order to find an optimal large step probability, the dynamic behavior of the Metropolis algorithm requires examination.

3.1 Flattened integrand

First we assume that the path building strategy was good in importance sampling, thus the non-zero contribution regions of the primary sample space have a roughly constant integrand. The convergence is fast if the Markov process explores the integration domain as quickly as possible. The random walks can be decomposed into *local explorations* when only accepted small mutations modify the sample locations. A local exploration phase is terminated when a large step is accepted, and the system starts exploring another part of the integration domain from the seed given by the large step.

The average probability that Metropolis sampling carries out an accepted large step is then the product of the large step probability and its success ratio, i.e. $p_l\eta_l$. As the number of tried mutations before an accepted large step follows a geometric distribution, the expected length of local exploration is

$$E[N_{local}] = 1 + \frac{1 - p_l \eta_l}{p_l \eta_l}$$

where 1 is the large step that gets the process to start this local exploration and $1 - p_l \eta_l$ is the probability that the next sample will also belong to this local exploration because it is a small step or a rejected large mutation.

During a local exploration, rejected mutations keep the original state while accepted small mutations modify the sample position additively, i.e. the new sample will be in the small neighborhood of the previous sample, where their difference is governed by the probability density of small mutations. As small mutations are independent, the variance of the perturbations caused by accepted small mutations is added, thus the average radius d of the space explored by N_{as} accepted small mutations grows proportional to the square root of the number of new samples. However, we cannot say that only small mutations can explore the full space. The first step of the local exploration, which is an accepted large step, also places a sample. As the perturbation of small steps is set to give the possibility to walk the whole space when the total number of mutations are executed, we can safely assume that the large step starting the local exploration is responsible for the same exploration. Thus, the space visited by a local exploration phase has expected radius

$$E[d] = \sigma \sqrt{N_{as} + 1}$$

where σ is the standard deviation of a single small mutation.

In a local exploration not all N_{local} steps belong to the category of new samples, only those mutations should count that lead to accepted small mutations:

$$E[N_{as}] = \frac{1 - p_l \eta_l}{p_l \eta_l} \cdot \frac{(1 - p_l)\eta_s}{1 - p_l \eta_l} = \frac{(1 - p_l)\eta_s}{p_l \eta_l}$$

since $(1-p_l)\eta_s$ is the probability that a successful small mutation is tried, and $1 - p_l \eta_l$ is the probability that this sample does not terminate the local exploration. If we generate M Metropolis samples, in average $M p_l \eta_l$ local explorations are established, each spreading over a subspace of radius $\sigma\sqrt{N_{as} + 1}$, thus the size D of the total explored space, having projected onto a line is

$$E[D] = M p_l \eta_l \sigma \sqrt{1 + \frac{(1 - p_l)\eta_s}{p_l \eta_l}} = M \eta_l \sigma \sqrt{\eta_l^2 p_l^2 + \eta_s \eta_l p_l (1 - p_l)}.$$

The objective of large step probability selection is to maximize the size of the space explored by M samples, thus we obtain:

$$p_l = \underset{p_l}{\operatorname{argmax}} \ E[D],$$

which can be obtained by finding where the derivative of $E[D]$ with respect to large step probability p_l is zero:

$$\frac{\mathrm{d}}{\mathrm{d}p_l} \left[\eta_l \sqrt{\eta_l^2 p_l^2 + \eta_s \eta_l p_l (1 - p_l)} \right] = 0 \quad \Longrightarrow \quad p_l = \frac{\eta_s}{2(\eta_s - \eta_l)}.$$

If large steps are almost as successful as small mutations, this formula can result in values that are larger than 1. This means that the large step probability should be set to 1, i.e. only large steps should be executed.

3.2 High variation integrand in the non-zero contribution regions of the primary sample space

So far, we assumed that the path building strategy already flattens the integrand transformed to primary sample space in the non-zero contribution regions, and thus the efficiency can be characterized by the speed of space exploration. However, when the transformed integrand still has large variation, Metropolis sampling should focus on the peaks of the integrand, thus the explored size itself is not an appropriate measure for efficiency. Unfortunately, this analysis would also involve the consideration of the integrand, which cannot be robustly estimated with a few samples and in parallel with the sample generation. We can only state that in this case, the small steps must be given higher probability since they will concentrate regions of high contribution while they are poorer in exploring lower contribution regions. The analysis of the previous section would propose $p_l = 0.5$ for this case, which is thus an overestimation. Without robust estimates which upon a theoretical model can be built, we simply propose the application of $p_l = 0.25$ in this case.

The cases of flattened and not flattened integrands can be distinguished by considering η_l/η_0, where we set the threshold to 0.1 above which the integrand is considered as flattened enough. Thus, the general form of the proposed large step probability is:

$$p_l = \begin{cases} \min\left\{\frac{\eta_s}{2(\eta_s - \eta_l)}, 1\right\} & \text{where } \eta_l/\eta_0 > 0.1, \\ 0.25, & \text{otherwise.} \end{cases}$$

4 Results

We have tried a variety of scenes to demonstrate robustness of the new control method and selected bi-directional path tracing with russian roulette termination as the path building strategy, Gauss and Mitchell-Netravali reconstruction filters, and Gauss or Reinhard tone mapping.

The *LuxTime scene* is a typical indoors setup with multiple area light sources where only slight difficulties are present, such as the dial of the watch which can only be illuminated by light that passes through the glass. This scene is lacking complex specular paths and the illumination is mainly diffuse interreflection, which is successfully addressed by bi-directional path tracing, indicating by the higher $\eta_l/\eta_0 \approx 0.4$ factor (Figure 3).

The *Spheres scene* includes dispersion and heavy volume scattering, and the only light that enters the scene is from the upper left, with a path through multiple thick glass-like surfaces making it a very difficult light transport situation.

Scene	η_l	η_s	η_0	$\frac{\eta_s}{2(\eta_s-\eta_l)}$	p_l	p_l^{opt}
LuxTime	0.377	0.783	0.985	0.95	0.95	(0.65-1)
Spheres	0.005	0.394	0.487	0.51	0.25	0.25
Chess Day	0.061	0.641	0.886	0.55	0.25	(0.2-0.4)
Cherry Splash	0.126	0.487	0.911	0.67	0.67	(0.4-0.5)
Cornell	0.004	0.438	0.022	0.50	0.50	(0.5-0.6)
Glass Ball	0.088	0.489	0.87	0.61	0.61	(0.5-0.6)

Table 1. Scene statistics with the proposed and optimum p_l values.

Standard bi-directional path tracing with random walk is unable to render this scene efficiently, thus there is a significant residual variation of the integrand even in primary sample space, which is also shown by the very low $\eta_l/\eta_0 \approx 0.01$ factor.

The *Chess Day scene* looks simple, but there are significant glossy interreflections that cannot be mimicked effectively by the deterministic connection rays of bi-directional path tracing, and the illumination comes partly through glass windows (Figure 2). As a result $\eta_l/\eta_0 \approx 0.07$.

In the *Cherry Splash scene* we can also observe complex specular paths, but they have significantly smaller total contribution than in the Chess and Spheres scenes.

The *Cornell Box scene* is a pathological case, where illumination comes in a light tube that allows the light to enter the box only after very many specular interreflections (Figure 3). In this scene, the size of the non-zero contribution primary sample space domain is 0.022, i.e. even after emphasizing important paths, non-zero contribution paths occupy only 2% of the space of paths. However, within this small domain, the integrand is relatively constant, thus $\eta_l/\eta_0 \approx 0.18$.

Finally, the *Glass Ball scene* is a typical outdoors setup with glossy interreflections, depth of field and caustics.

5 Conclusions

In this paper we analyzed the convergence properties of Primary Sample Space MLT, and proposed a method to control its large step probability parameter. Our method is based on a few statistical parameters, including the success ratios of small perturbations, large perturbations, and perturbations leading to non-zero contribution tentative samples. These parameters can be obtained easily and robustly at the beginning of the rendering process. We also showed that these parameters tell us a lot about the scene and its suitability for sampling by the given path building strategy. As Metropolis is responsible to do importance

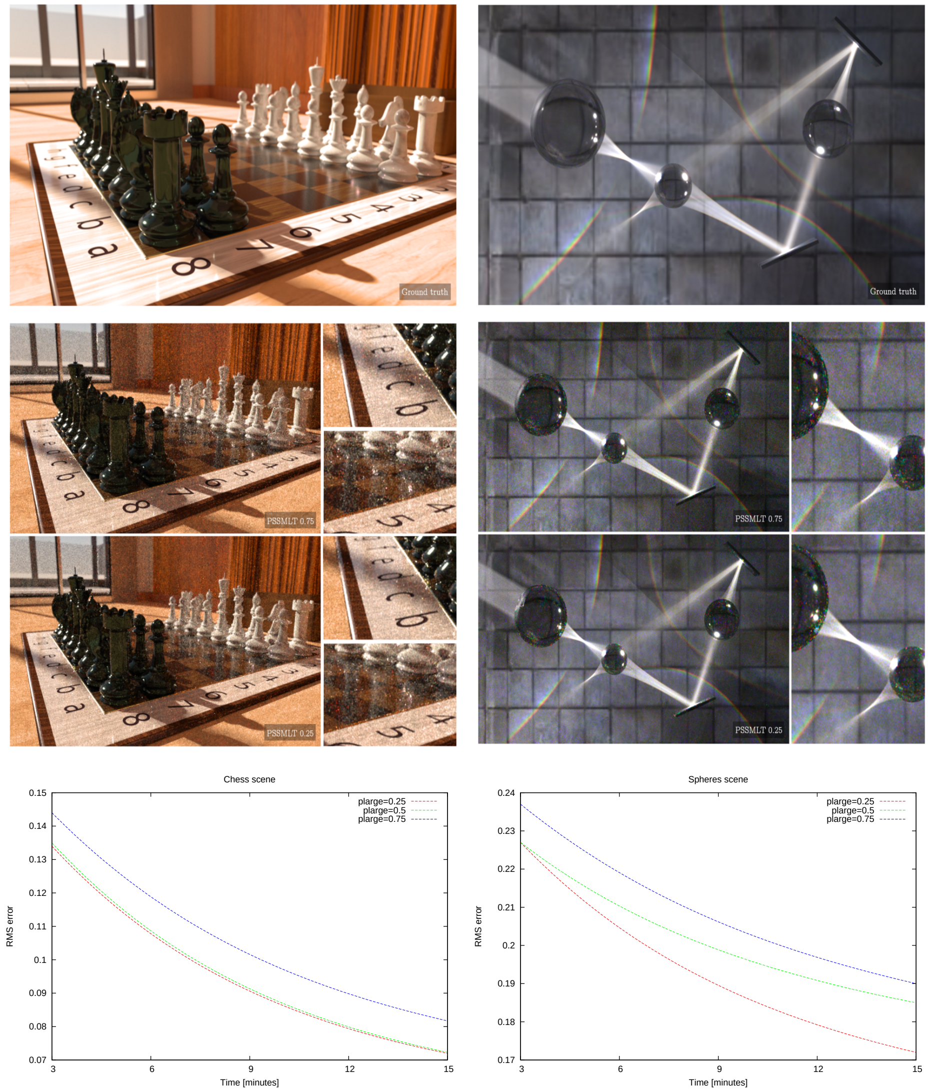

Fig. 2. Analysis of the Chess Day scene (left) and the Spheres scene (right). We show the rendering results with three large step probabilities and also the L2 error curves in terms of the rendering time.

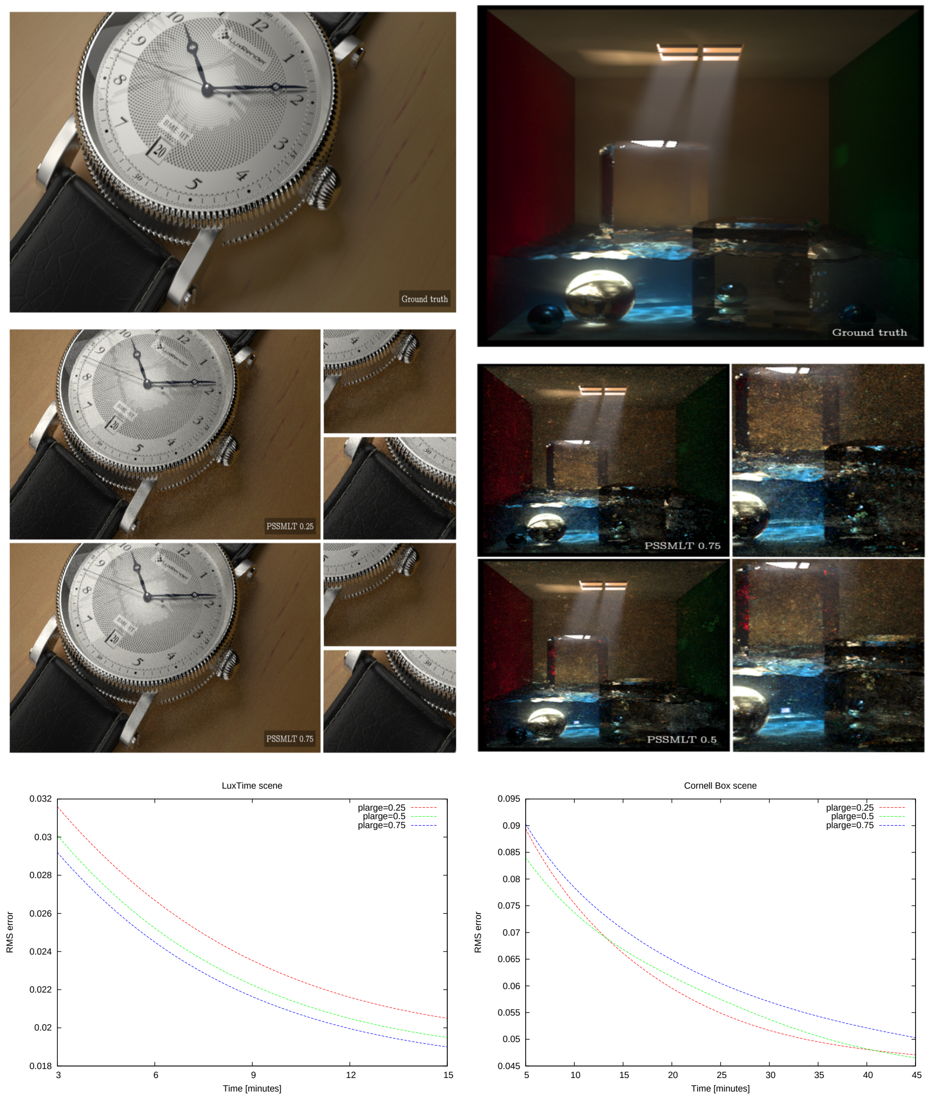

Fig. 3. Analysis of the LuxTime scene (left) and the Cornell Boxl scene (right). We show the rendering results with three large step probabilities and also the L2 error curves in terms of the rendering time.

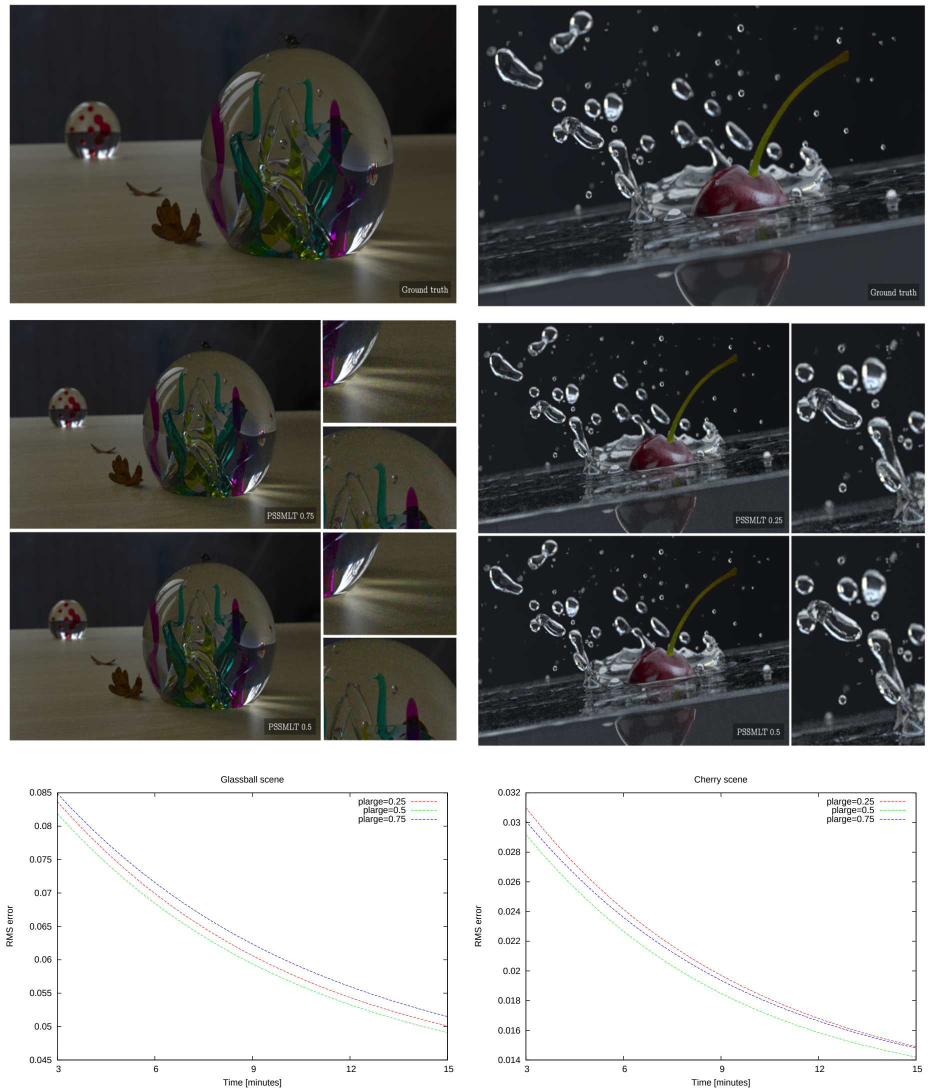

Fig. 4. Analysis of the Glass Ball scene (left) and the Cherry Splash scene (right). We show the rendering results with three large step probabilities and also the L2 error curves in terms of the rendering time.

sampling on the function that is already flattened by the path building strategy, we can build our proposed large step value on these parameters.

6 Acknowledgements

This work has been supported by TÁMOP-4.2.2.B-10/1-2010-0009 and OTKA K-104476. We would like to thank Kai Schwebke for providing LuxTime, Vlad Miller for the Spheres, Giulio Jiang for the Chess, Aaron Hill for the Cornell Box, Andreas Burmberger for the Cherry Splash and Glass Ball scenes. We are also grateful for the LuxRender community for their help with rendering the ground truth image for the Spheres and Cornell Box scenes, which required tens of thousands of samples per pixel to get a completely noise-free result.

References

[APSS01] Ashikhmin M., Premoze S., Shirley P., Smits B.: A variance analysis of the Metropolis light transport algorithm. *Computers & Graphics 25*, 2 (2001), 287–294.

[DBB03] Dutre P., Bekaert P., Bala K.: *Advanced Global Illumination.* A K Peters, 2003.

[Jen01] Jensen H. W.: *Realistic Image Synthesis Using Photon Mapping.* AK Peters, 2001.

[KSKAC02] Kelemen C., Szirmay-Kalos L., Antal G., Csonka F.: A simple and robust mutation strategy for the Metropolis light transport algorithm. In *Eurographics '02* (2002), pp. 531–540.

[MRR*53] Metropolis N., Rosenbluth A., Rosenbluth M., Teller A., Teller E.: Equations of state calculations by fast computing machines. *Journal of Chemical Physics 21* (1953), 1087–1091.

[PKK00] Pauly M., Kollig T., Keller A.: Metropolis light transport for participating media. In *Rendering Techniques* (2000), pp. 11–22.

[SKDP99] Szirmay-Kalos L., Dornbach P., Purgathofer W.: On the start-up bias problem of Metropolis sampling. In *Winter School of Computer Graphics '99* (Plzen, Czech Republic, 1999), pp. 273–280.

[VG97] Veach E., Guibas L.: Metropolis light transport. *SIGGRAPH '97 Proceedings* (1997), 65–76.