

Keldysh Institute • Publication search Keldysh Institute preprints • Preprint No. 122, 2018



ISSN 2071-2898 (Print) ISSN 2071-2901 (Online)

#### Ershov S.V., Zhdanov D.D., Voloboy A.G., Sorokin M.I.

Treating diffuse elements as quasi-specular to reduce noise in bi-directional ray tracing

**Recommended form of bibliographic references:** Ershov S.V., Zhdanov D.D., Voloboy A.G., Sorokin M.I. Treating diffuse elements as quasi-specular to reduce noise in bi-directional ray tracing // Keldysh Institute Preprints. 2018. No. 122. 30 p. doi:10.20948/prepr-2018-122-e URL: http://library.keldysh.ru/preprint.asp?id=2018-122&lg=e

Ордена Ленина ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ имени М. В. Келдыша Российской академии наук

S. V. Ershov, D. D. Zhdanov, A. G. Voloboy, M. I. Sorokin

Treating diffuse elements as quasi-specular to reduce noise in bi-directional ray tracing

Mocквa - 2018

#### S. V. Ershov, D. D. Zhdanov, A. G. Voloboy, M. I. Sorokin

#### Treating diffuse elements as quasi-specular to reduce noise in bi-directional ray tracing

In simulation of the light propagation the brightness of each ray is calculated based on the optical properties of the virtual scene objects which it interacts with. According to the optical properties these objects can be roughly subdivided into diffuse and specular. The Monte-Carlo based ray tracing methods which are commonly used in simulation calculate the brightness only for diffuse surfaces. When a ray meets a specular surface it is reflected (or refracted) until it reaches a diffuse surface where the brightness is calculated. In the proposed approach diffuse elements are further divided into "truly diffuse" and "quasimirror". The most natural criterion for the latter is scattering in a narrow cone around a purely specular direction. A scene element can also be a superposition of both types when its scattering function is treated as the sum of a truly diffuse and quasi-mirror parts. The work shows how various components of illumination interact with quasi-mirror objects and describes how this is implemented in the bi-directional Monte Carlo ray tracing. The proposed approach allows to significantly reduce the stochastic noise for a number of scenes including those for which it is not possible to achieve good results with any settings of the standard method. This method is also applicable to the simulation of volume scattering treating the phase function of the medium as quasi-mirror. In this case the selection of quasi-mirror objects is by no means based on the character of the scattering distribution function. For volume scattering the medium is treated as quasi-mirror while the surfaces, even if their scattering distribution functions are narrower, remain "truly diffuse". The article shows the advantage of this approach.

**Key words:** Lighting simulation, Monte Carlo ray tracing, bi-directional ray tracing, noise reduction, bi-directional scattering distribution function.

The work was partially supported by RFBR, Grants No 16-01-00552 and 18-01-00569.

#### С. В. Ершов, Д. Д. Жданов, А. Г. Волобой, М. И. Сорокин

# Обработка диффузных элементов сцены как квази-зеркальных для снижения шума в двунаправленной лучевой трассировке

При моделировании распространения света яркость каждого луча рассчитывается, исходя из оптических свойств объектов виртуальной сцены, с которыми он взаимодействует. По оптическим свойствам объекты могут грубо подразделяться на диффузные и зеркальные. Обычно используемая при моделировании трассировка лучей на основе методов Монте-Карло рассчитывает яркость лишь для диффузных поверхностей. При встрече луча с зеркальной поверхностью он отражается (или преломляется) до тех пор, пока не достигнет диффузной поверхности, где только и рассчитывается яркость. В предлагаемом подходе диффузные элементы разделяются далее на "истинно диффузные" и "квази-зеркальные". Наиболее естественным критерием для последних будет рассеяние в узком конусе вокруг чисто зеркального направления. Элемент сцены может быть также суперпозицией обоих типов, когда её функция рассеяния трактуется как сумма истинно диффузной и квазизеркальной частей. В работе выведено, как различные компоненты освещённости взаимодействуют с квази-зеркальными объектами, и описывается, как это реализуется в двунаправленной стохастической трассировке лучей. Предложенный подход позволяет значительно снизить стохастический шум для ряда сцен, включая и те, для которых не удаётся достичь хороших результатов при любых настройках стандартного метода. Этот метод также применим к моделированию объёмного рассеяния, трактуя фазовую функцию среды как квази-зеркальную. В этом случае выделение квази-зеркальных объектов отнюдь не основано на характере индикатрисы рассеяния. Для объемного рассеяния среда трактуется как квази-зеркальная, в то время как поверхности, даже если их индикатрисы более узкие, остаются "подлинно диффузными". В статье показывается преимущество такого подхода.

**Ключевые слова:** Моделирование освещенности, стохастическая трассировка лучей, двунаправленная трассировка лучей, понижение шума, двунаправленная функция рассеяния.

# Contents

1	Introduction			
<b>2</b>	Global illumination equation			
3	Iterative solution (Neumann series) of the illumination field equation (4) and BMCRT			
4	How the BDD affects the noise level	10		
<b>5</b>	Introducing the quasi-specular method	13		
6	Operator series in presence of quasi-specular BDFs $6.1$ The form to be used for BDD=0 $6.2$ The form to be used for BDD=1 $6.3$ The form to be used for BDD=2 $6.4$ The form to be used for BDD= $N$	<ol> <li>17</li> <li>19</li> <li>20</li> <li>21</li> <li>22</li> </ol>		
7	Integration by paths for BMCRT         7.1 The case of BDD=0	<ul> <li>22</li> <li>23</li> <li>24</li> <li>25</li> <li>25</li> </ul>		
8	Volumetric scattering8.1Standard method8.2"Quasi-specular medium"	<b>26</b> 26 27		
9	Results         9.1       Surface case         9.2       Volumetric case	<b>28</b> 28 28		
10	Conclusion	29		

# Abbreviations

- MCRT = Monte-Carlo ray tracing
- **FMCRT** = forward Monte-Carlo ray tracing. It is tracing of rays from light sources toward scene objects accumulating of illumination on scene objects. Usually it is used to calculate the secondary (indirect) illumination.
- **BMCRT** = backward Monte-Carlo ray tracing. It is tracing of rays from camera through virtual screen toward scene objects. Usually it is used to get the virtual scene image.
- BDF = bi-directional scattering function. It describes surface luminance as a function of the illumination and observation direction
- **BDD** = backward diffuse depth. It is a specific parameter of a hybrid ray tracing, when FMCRT calculates illumination and BMCRT is used to convert it to the observed luminance. In this method the backward ray usually has a limited "length" and terminates after BDD diffuse events.

# 1 Introduction

A powerful method of calculation of a virtual camera image is a hybrid, or bidirectional, Monte-Carlo ray tracing. The self-consistent light field in a scene can be expressed via Neumann series of iterative scattering and transport operators and the Monte-Carlo ray tracing calculates those integrals. Although in principle the *forward* Monte-Carlo ray tracing is enough, it is inefficient for calculation of a camera image because of too low probability of a ray to hit the camera pupil, and long ago a backward Monte-Carlo ray tracing had been proposed [5] that traces rays from camera.

In the "pure" BMCRT a ray is traced from camera until it is absorbed or leaves the scene, and wherever it hits a diffuse surface, it collects direct and caustic illumination. The ray does *not* need to hit light source, so the method can work with point and parallel lights. In many cases this method works well, but regrettably not always. Its problem is that the camera ray must visit scene areas which are directly illuminated, while in some scenes (see e.g. Fig. 1) they are very small and thus are hit too rarely.

There are many different attempts to overcome the limitations of the FM-CRT and BMCRT. One of the most radical is the so-called Metropolis Light Transport (MLT) [9], [8]. This method has a lot of modifications (up to using Hamiltonian dynamics [6]), but all of them differ from the "classic" Monte-Carlo in that they do not trace ray sequentially just *hopping* that it reaches the

camera pupil (in BMCRT — that it hits a light source). Instead, the MLT operates subspace of trajectories which "connect" lights and camera. As a result, all the samples taken are "successful" in the sense that they do bring luminance to an image pixel. In the "classic" MLT versions, these trajectories were generated using the Hastings method which produces a series of random samples with stationary distribution just proportional to their "contribution" (this is called "importance sampling") [5], [9], [8]. There, however, are problems with that. First, although the resulting probability distribution is indeed proportional to the trajectory's "importance", the successive samples (i.e. trajectories) are correlated. In extreme case, the method just *repeats* "successful" samples i.e. the same trajectory is taken several times successively which is not good. Another problem is that the process is not stationary, unlike the classic Monte-Carlo, where both the first and the last trajectories have the same probability distribution. In the Hastings method, the distribution changes in time. Until the density converged to the stationary state (which is proportional to the trajectory "importance"), the result is wrong. In many cases the speed of convergence is good enough and the initial part can be easily discarded, but this is not always. The last problem is that the Hastings method utilizes some "tentative" transition rule. Formally it can be quite arbitrary and all the same the stationary density will be the same (and correct). But practically the speed of convergence (to the stationary probability distribution), correlations etc. depend on it, and a good choice of that "tentative transition rule" is what makes the method to work (or to not work, in case of a bad choice.)

Besides MLT and its relatives, there are less radical changes that improve the speed of Monte-Carlo method, retaining its basics. One of them is bidirectional Mone-Carlo ray tracing, of which, again, a lot of variants exist. The basic idea is to operate trajectories that *always connect* light source and camera (like in MLT). To do so, trajectory is constructed by "concatenating" two its parts: one is from the camera and is obtained by BMCRT and another is from light source and is obtained by BMCRT. Then these pieces are "connected" somehow which creates trajectory that connects light source and camera.

The main modifications of this "connection" are:

- 1. Each (or some subset of them) hit point of a backward piece is connected with each hit point (or, again, a subset of them) of a forward piece by the "shadow ray" segment [2]. There are many advantages in this method, but a serious problem is that how to operate the huge number of connecting segments (of which many have too low importance or are completely shadowed).
- 2. Photon map visualization [4]. Here the FMCRT part populates scene with the hit points which represent the statistic of illumination of the hit point.

Roughly, the density of hits gives the full illuminance, and leaving only the hits with desired direction gives the illuminance in that direction. The backward phase, which can in principle be even deterministic (though BMCRT is applied usually) spreads camera rays over the scene and in each hit point they take illumination from the photon map, "convolve" it with the surface BDF and add resulting luminance to the pixel's. The local illuminance is estimated from the number of FMCRT hits inside (a small) *integration sphere* about the BMCRT hit, and this introduces some blurring and discretization error. However, in most cases this error is inessential as compared to the noise level.

As one can see, the last method also constructs the full trajectory that connects camera and light source, but does this approximately because the end of the forward part (the photon inside integration sphere) deviated slightly from the end of the end of the backward part (BMCRT hit i.e. the centre of that sphere). This is usually inessential though in some cases, especially for caustic photons, may create problems and thus needs special treatment [7].

Once the full trajectory is constructed concatenating the BMCRT and FM-CRT parts, it must be decided how to distribute it among these parts. That is, how long is the BMCRT part. Here one can fix it, or vary at random, or choose from the properties of the trajectory etc.

In principle, all the above decisions can be treated as an average over the "ensemble" of *fixed* choices. So investigation and optimisation of the method with the fixed BMCRT part length is important, and this paper is a step in that direction.

Usually fixed is the number of *diffuse* events, while the number of specular ones is not. This value is termed BDD i.e "backward diffuse depth".

Efficiency of the approach, i.e. the rate of convergence (or noise level, which is more or less the same) strongly depends on that BDD, and its optimal value is specific for each scene.

A better approach is to let it be different across the scene [1] and even mix calculations with different BDD [1]. Frequently it is possible to find it, automatically or manually, so that the calculations are quite efficient. But this is not always and in some cases changing BDD does not help and whichever value we choose the image is highly noisy. Here we also prove why an "adaptive choice" of BDD individually for each ray may result in a biased estimate, i.e. the calculated luminance converges to a wrong value.

# 2 Global illumination equation

Light distribution in a scene is described by a "self-consistent field" equation which can be written in different forms that are named "global illumination equation", "rendering equation" [5] etc. Below we shall describe its form used in this paper and define some necessary terms.

Roughly, this equation describes the following global illumination problem. There is some light field in the scene. It illuminates its surfaces which scatter it. This transformation of the illumination (light that falls onto a surface) into luminance (light that goes away from a surface) is described by the surface BDF:

$$L(oldsymbol{v};oldsymbol{x}) = \int f(oldsymbol{x};oldsymbol{v},oldsymbol{v}') I(oldsymbol{v}',oldsymbol{x}) d^2oldsymbol{v}'$$

Here f is BDF, I is the field of illumination of *diffuse* scene surfaces (i.e. it omits the points of pure specular surfaces) and L is the luminance of *diffuse* scene surfaces (i.e. it omits the points of pure specular surfaces).

Then this light, emitted *from* the surface, propagates over the scene and eventually illuminates its surfaces. This transformation of the luminance (light going away off a surface) into illuminance (light incident onto another surface point) is described by the transport operator:

$$I^{(i)}(\boldsymbol{v}; \boldsymbol{x}) = \int t(\boldsymbol{x}, \boldsymbol{v}, \, \boldsymbol{x}', \boldsymbol{v}') L(\boldsymbol{v}', \boldsymbol{x}') |(\boldsymbol{n} \cdot \boldsymbol{v}')| d\boldsymbol{x}' d^2 \boldsymbol{v}'$$

where  $I^{(i)}$  is the indirect illumination component which does not include light that came "directly" (i.e. without scattering) from a light source. Since our equations explicitly include only *diffuse* scattering, the "direct" component also contains light that underwent pure specular transformation. Again, this equation operates only on the scene points that belong to *diffuse* surfaces.

In the simplest case, when there is no specular surfaces or turbid media etc, this transport operator reduces to transport along the straight ray path:

$$I^{(i)}(\boldsymbol{v}; \boldsymbol{x}) = L(\boldsymbol{v}, \boldsymbol{x} + l\boldsymbol{v})|(\boldsymbol{n} \cdot \boldsymbol{v})|$$

where l is the ray length from x to the hit point. In presence of specular surfaces, the transport operator includes ray splitting and luminance transformation across a boundary etc.

In fact its explicit form is inessential for us and we shall introduce the transport and scattering operators:

$$L = \hat{F}I \tag{1}$$

$$I^{(i)} = \hat{T}L \tag{2}$$

The full illumination consists of indirect and direct components:

$$I = I^{(i)} + I^{(0)} (3)$$

Combining the three equations we arrive at the self-consistent illumination field equation at the form we shall use:

$$I = \hat{T}\hat{F}I + I^{(0)} \tag{4}$$

# 3 Iterative solution (Neumann series) of the illumination field equation (4) and BMCRT

Although it is a common knowledge in the computational optics, we shall produce a basic description in just the form we shall use.

In many ray-tracing applications the *illumination* field is calculated by FM-CRT before and independently of rendering of the virtual image. Then these "illumination maps", "photon maps" etc. are used to calculate the surface luminance.

The simplest way is that we trace rays from camera, terminating then at the first diffuse surface. In that hit point we calculate this surface luminance under the full illumination  $(=\hat{F}I)$  and *add* to the pixel luminance. This provides estimate of the surface luminance (1), though the result is not perfect. Indeed, the illumination map (or its analog) i.e. estimation of I from FMCRT is usually subjected to (spatial) filtering to reduce noise. Thus all fine illumination details such as highlights are usually lost.

A better estimation follows from another equation for surface luminance: combining (4) with (1) one has

$$L = (\hat{F}\hat{T})\hat{F}I + \hat{F}I^{(0)}$$
(5)

Notice that it is **the luminance of the scene surfaces**. The luminance of the **camera image** would be  $\mathcal{L} = \hat{S}L$  where the leftmost  $\hat{S}$  describes the *pure* specular transformation between the camera and a scene surface (frequently  $\hat{S} = 1$ ). Here and below we shall only calculate the L. Its transformation into the camera image can be applied after that if needed.

The term  $\hat{F}I^{(0)}$  is the surface luminance under *direct* (including caustic) illumination. The indirect (diffusively scattered) illumination is ignored. This is what the simplest BRT does.

The term  $(\hat{F}\hat{T})\hat{F}I$  is the surface luminance under *indirect* (diffusively scattered) illumination. The integral operator can be calculated using Monte-Carlo ray tracing: we fire rays from camera, they hit a surface, are scattered (operator  $\hat{F}$ ), propagate the scene until reach the 2nd diffuse surface (operator  $\hat{T}$ ) and terminate. Notice that in the *backward* ray tracing the operators are applied in the reverse order) and terminate. In this 2nd diffuse hit point (i.e. just *before* the 2nd diffuse scattering) we calculate this surface luminance under the full illumination  $(\hat{F}I)$ , scale by the luminance transmission factor due to the specular transformation in  $\hat{T}$ , and *add* to the pixel luminance. The average over ensemble of camera rays converges to  $(\hat{F}\hat{T})\hat{F}I$ .

Although mathematically the equation solved is identical to (1), the numerical results differ. The second approach is less sensitive to small distortion of Ibecause the convolution with  $(\hat{F}\hat{T})$  reduces their effect.

The process can be continued further. Substituting (4) for I in (5) gives

$$L = (\hat{F}\hat{T})(\hat{F}\hat{T})\hat{F}I + (\hat{F}\hat{T})\hat{F}I^{(0)} + \hat{F}I^{(0)}$$

Again its terms can be estimated with BMCRT. The first term is estimated taking the full illumination  $(\hat{F}I)$  at the *third* diffuse hit, the term  $(\hat{F}T)\hat{F}I^{(0)}$  by taking the *direct* (and caustic) illumination at the *second* diffuse hit and  $\hat{F}I^{(0)}$  is the *direct* (and caustic) illumination at the *first* diffuse hit. Now the most noisy (or with the strongest artifacts from filtering) term  $\hat{F}I$  is *twice* convolved with  $(\hat{F}T)$  which reduces the noise (or artifacts).

Because of linearity these three terms can be estimated from the *same* backward ray ensemble, taking the full illumination  $(\hat{F}I)$  at the *third* diffuse hit, and the *direct* (and caustic) illumination at the *first and second* diffuse hits.

Continuing the process, one may have BMCRT which takes the direct (and caustic) illumination at the 1st, 2nd, ..., N - 1-th diffuse hit and the full illumination at the N-th diffuse hit for an arbitrary N > 0. The value of N is termed "**backward diffuse depth**", or **BDD** for short.

### 4 How the BDD affects the noise level

As we have seen above, *some* image inaccuracies (noise and/or artifacts of filtering of  $I^{(i)}$ ) decrease with the BDD and one might decide that for the best results the BDD must be as large as possible. Regrettably, the choice is not that simple because there are other sources of inaccuracy.



Figure 1: Diffuse plate illuminated through a diffuse sphere. This is the side view in correct scale. The circle about it is the diffuse lamp shade; the yellow cone inside it mark its light emission (downwards).

For example, let us consider a simple scene consisting of a diffuse plate illuminated by a cone light (which emits downwards) enclosed in a matte (diffuse) spherical shade, see Fig. 1:

We assume that camera sees only the plate while the lamp shade is out of the view area (e.g. it is behind the camera). If the lamp shade is small enough we can neglect interreflections. The plate is pure red Lambert with integral reflectance of 50%.

In this approximation illumination of the plate is entirely *indirect* (light from the diffuse sphere). In case of BDD=0 camera ray collects that illumination in the hit point. There is some moderate noise from the FMCRT (= from photon map), but no more (Fig. 2). In case of BDD=1 the situation becomes much worse (Fig. 2).

Indeed, now in the first hit point (i.e. the plate) we take only the direct (and caustic) illumination which in this model is exactly 0. So *all* contribution to pixel luminance comes only from the *second* diffuse hit, which is only possible when the scattered camera ray hits the lamp shade. But this probability is very low, because the BDF (and thus the scattered light cone) is *wide* while the lamp shade is *small*. Therefore, it is very rare that camera ray brings any luminance to the image pixel, and this means high noise.

The case of BDD>1 is not better because it also requires camera ray hitting the lamp shade.

Therefore, in our model scene the best is BDD=0 while BDD>0 results in *very* high noise. The choice of an optimal BDD is therefore not trivial: while rather frequently increasing it is good, for some scenes the best result is for the



Figure 2: Camera images for scene from Figure 1, calculated during the same time (3000 sec) for BDD=0 (left) and BDD=1 (right).

#### minimal BDD.

Now let us consider another, in a sense the opposite, case by replacing the small lamp shade with something very large. The scene is again a diffuse plate, but now it is surrounded by a large room box with diffuse walls and ceiling which are illuminated by a grid of cone lights (emitting towards the nearest surface and not towards the room centre), so their light cannot reach the plate directly. The camera is inside that room, as shown in Fig. 3.

The room walls and ceiling is Lambert with albedo 0.5 and the plate is rather small as compared to the room. The floor outside it is black. Again, we assume that camera sees only the plate. Like in the previous example, the plate's illumination is purely *indirect*.

In this scene the light from the grid of cone emitters undergoes many interreflections and so creates a rather uniform illumination of the walls and ceiling, though it is not exactly constant and the areas near the light emitters are brighter.

At last, unlike the previous model scene, the bottom plate has now a *sharp* (nearly specular BDF). It is nearly pure green<sup>1</sup> with integral reflectance of 40%.

In case BDD=0 we collect the indirect illumination (from FMCRT!) at the first hit point. Illumination comes from the large box which is illuminated rather uniformly, and thus makes a wide cone. Meanwhile, the BDFs is sharp, so  $\hat{F}I$  effectively senses only illumination in direction *close to the mirror reflection of the view ray*. In other words, most of FMCRT rays will be effectively excluded

<sup>&</sup>lt;sup>1</sup>Its normalized color is (0.1, 1, 0.1)



Figure 3: Diffuse plate illuminated by the room walls and ceiling. This is the side view in correct scale. The yellow cones show emission of lights; the camera view ray and its specular reflection are drawn with red arrows. The thick green strip is the bottom plate.

and only few of them will contribute to the pixel luminance. This means a strong noise (Fig. 4).

In case BDD=1 the situation improves. At the first camera hit we collect only the direct and caustic illumination, which is exactly 0 there. After diffuse reflection by the plate, the camera ray hits the box and there it takes the full (in fact, again indirect) illumination. Most of camera rays and most of FMCRT rays contribute here. So the noise is low (Fig. 4).

The larger values of BDD work *less good*. Since illumination is merely *indirect* (save for the small area near above the lights which we neglect), the contribution to pixel luminance comes only from the *third* hit point. It can be in the box or in the bottom plate with more or less close probabilities. When it is in the box, it is like for BDD=1 i.e. all works good. But when it hits a bottom plate, all is like for BDD=0 i.e. we take a wide cone illumination of a narrow BDF which results in strong noise.

# 5 Introducing the quasi-specular method

The model scenes from Section 4 at least admit an optimal BDD. So, if the BDD is chosen automatically by a method which is intellectual enough, or by an experienced human operator, a good quality image can be obtained. But there are scenes in which whatever BDD we choose the image is highly noisy.

A very simple example is a *combination* of the two model scenes from Section 4: the bottom plate now has BDF which is a sum of a wide (nearly Lam-



Figure 4: Camera images for scene from Fig. 3, calculated during the same time (3000 sec) for BDD=0 (left) and BDD=1 (right).

bert) and a sharp (nearly specular component). To see which BDF does which contribution to the image, the Lambert component is pure red and the sharp component is nearly pure green<sup>2</sup>. Their integral reflectance is 50% and 40%, respectively. Indirect illumination comes both the room walls and ceiling (illuminated by an array of cone lights), and also by a small "central" lamp shade with a cone light inside (Fig. 5).

Unlike the first example (Fig. 1), the light from the lamp shade also results in interreflections. But when the bottom plate is small (recall the light which hits the room floor outside it is absorbed), it is inessential.

Repeating the analysis from Section 4 one finds that BDD=0 is bad because of high noise from illumination of a narrow BDF part from the box and BDD=1 is bad because of high noise from illumination of a Lambert BDF from a small lamp shade, (Fig. 6). Larger BDDs are not better.

There is a temptation to make the BDD to be individual for a ray, so that it is 0 for the camera ray which was scattered by a Lambert component of the floor and it is 1 for the camera ray that was (first) scattered by the nearly-specular BDF component. That is, denote the angle between the scattered camera ray and its ideally specular reflection as  $\vartheta$ , then if  $\vartheta \ge \Theta$  then BDD=0 (and ray terminates here), while if  $\vartheta \le \Theta$  then BDD=1 (and ray propagates further).

One must recall that BDD affects not only the ray length but also which illumination component is taken at the point. For BDD=0 in the first hit point we take full illumination. For BDD=1 in that same point we take only

<sup>&</sup>lt;sup>2</sup>Its normalized color is (0.1, 1, 0.1)



Figure 5: Combined model scene: indirect illumination both from the lamp shade and the room walls and ceiling. This is the side view in correct scale. The yellow cones show emission of lights. The camera view ray and its specular reflection are drawn with red arrows. The thick green strip is the bottom plate. The circle about it is the diffuse lamp shade. The yellow cone inside it mark its light emission (downwards).



Figure 6: Camera images for scene from Fig. 5, calculated during the same time (3000 sec) for BDD=0, BDD=1 and quasi-specular method with BDD=0. For the area of 100x100 pixels around the centre, the average RGB color is: (150, 27, 2.71); (150, 31.6, 3) and (150, 31.4, 3), respectively. The noise level (rel-ative to the photometric luminance) is 1000%, 200% and 76% respectively.

direct+caustic.

However this does not work[1]. Indeed, let us consider a model scene from Figure 3, but let now the bottom plate has BDF which is a sum of Lambert and a nearly specular narrow lobe:

$$\hat{F} = \hat{F}_{ns} + a\hat{F}_1$$

where a is the albedo (integral) of the Lambert's component,  $\hat{F}_1$  is operator for Lambert with unit albedo and  $\hat{F}_{ns}$  is the nearly-specular part of the BDF.

Those camera rays that were scattered in off-specular direction  $\vartheta \geq \Theta$  stop here and collect the full illumination. Those which were scattered in nearspecular direction  $\vartheta \leq \Theta$  collect only direct+caustic and propagate further. Then they hit the box and their collect the full illumination.

Now let us assume that interreflections between the floor plate and the room can be neglected (e.g. their integral scattering is low, or instead of the "plate" there is only a *small* rectangle seen by camera while outside the surface is entirely black). In this case what happens with the nearly-specular rays and the luminance they bring to the image is *independent* from the Lambert component. If we set the latter to 0, all works *exactly* as in the "standard mode" for BDD=1. Therefore, the luminance brought by these near-specular rays coincides with  $\hat{F}_{ns}I$  where I is the full illumination of the floor and  $\hat{F}_{ns}$  is the nearly-specular part of the BDF.

As to the off-specular rays, their fraction  $\lambda$  approximately equals the fraction of rays scattered by Lambert. In fact it is slightly lower because some rays scattered by Lambert go near-specular direction. But their fraction is very small and we neglect it. So,  $\lambda \approx \frac{a}{a+b}$  where b is the integral scattering of the nearly-specular lobe. Each that ray behaves like for BDD=0, i.e. it takes the full illumination I and convolves with the *full* BDF. Therefore the contribution to the image luminance from these rays is  $\lambda \hat{F}I$ .

The total image luminance is thus

$$\hat{F}_{ns}I + \lambda \hat{F}I \approx \hat{F}_{ns}I + \frac{a}{a+b}\hat{F}I$$

$$= \hat{F}_{ns}I + \frac{a}{a+b}(\hat{F}_{ns} + a\hat{F}_1)I$$

$$= \left(1 + \frac{a}{a+b}\right)\hat{F}_{ns}I + \frac{a^2}{a+b}\hat{F}_1I$$

which *differs* from the exact value  $\hat{F}I$ .

Therefore, just making the near-specular and off-specular rays to collect illumination according to the ray's BDD produces wrong luminance, and one must apply a more sophisticated selection of the illumination components collected in each hit point.

The key idea of the approach though remains the same: the nearly specular narrow BDF is treated specially, not as a "genuine diffuse"; but in a sense closer to specular ones. Such narrow BDFs (and the whole method of their treatment in BMCRT) are thus named "quasi-specular". The result of calculation of the same model scene with this method is shown in the rightmost panel of Figure 6, see Section 9.1 for more explanations.

One must realize that since BDF is a linear operator, so we can treat a surface BDF as either a single function (= the sum over all sub-BDFs) or as a sum which gives exactly the same result.

Here and below we treat the diffuse BDF as a sum of *two* parts: the "genuine diffuse" and "quasi-specular". We shall prove that the separation can be *arbitrary*. Roughly, we can take a sum of a Lambert and a Ward lobe so that... the Lambert is quasi-specular while the narrow lobe is "genuine diffuse"! Another question is whether *such* a separation is advantageous in terms of noise. To this end, *usually* the separation must be "natural", i.e. the wide component is "genuine diffuse" while the narrow one is "quasi-specular". But it is *not* always so e.g. in scenes with turbid media (i.e. volumetric diffuse scattering). Although the volumetric scattering is rather specific, it has much common with surface one and even can be approximated by a series of *surface* scattering events. And, it happens that in many cases it is advantageous to treat the volumetric scattering as quasi-specular, i.e. not add this event to the "BDD counter", and also not to take their indirect illumination. This improves results even if the medium scattering (i.e. phase function) is *wider* than the surface BDFs!

# 6 Operator series in presence of quasi-specular BDFs

Now let us come to the formal derivation of what to do with a backward ray when the diffuse BDF is subdivided into the "genuine diffuse" and "quasispecular" components:

$$\hat{F} = \hat{F}_d + \hat{F}_{qs}$$

We assume that if the backward ray underwent scattering by the quasi-specular component, this does *not* increment the diffuse event counter. So the ray does not terminate and shall derive which illumination components must be taken at which hit points so as "to be compatible" with the above behavior, i.e. so that the mathematical expectation of the BMCRT luminance to match the exact value.

Luminance of a surface point  $\boldsymbol{x}$  is:

$$L(\boldsymbol{v}, \boldsymbol{x}) = (\hat{F}_d I)(\boldsymbol{v}, \boldsymbol{x}) + (\hat{F}_{qs} I)(\boldsymbol{v}, \boldsymbol{x})$$

where

$$I = I^{(0)} + I^{(qc)} + I^{(i)}$$

is *illuminance* of that point which consists of **direct** (that was not scattered at all or scattered by pure specular surfaces), **quasi-caustic** (scattered at least once by a specular BDF and never by a diffuse BDF) and **indirect** (scattered at least once by a diffuse BDF and any times by a specular BDF or quasi-specular) components.

The above luminance expression can be written as

$$L = \hat{F}I = \hat{F}_{qs} \left( I^{(0)} + I^{(qc)} \right) + \hat{F}_{qs}I^{(i)} + \hat{F}_dI$$
(6)

Substituting our decomposition of I into the global illumination equation (4) one arrives at

$$I^{(qc)} + I^{(i)} = \hat{T}\hat{F}_{qs}\left(I^{(0)} + I^{(qc)}\right) + \hat{T}\hat{F}_{qs}I^{(i)} + \hat{T}\hat{F}_{d}I$$

The right hand side term  $\hat{T}\hat{F}_{qs}I^{(i)} + \hat{T}\hat{F}_dI$  comprises light that underwent at least one *diffuse* scattering while the term  $\hat{T}\hat{F}_{qs}\left(I^{(0)} + I^{(qc)}\right)$  comprises light that underwent only specular scattering. In view of the decomposition into diffuse and quasi-caustic illumination, this means

$$I^{(i)} = \hat{T}\left(\hat{F}_d I + \hat{F}_{qs} I^{(i)}\right) \tag{7}$$

$$I^{(qc)} = \hat{T} \left( \hat{F}_{qs} I^{(qc)} + \hat{F}_{qs} I^{(0)} \right)$$
(8)

Here,

$$(\hat{F}_{d}I)(\boldsymbol{v},\boldsymbol{x}) \equiv \int f_{d}(\boldsymbol{x};\boldsymbol{v},\boldsymbol{v}')I(\boldsymbol{v}',\boldsymbol{x})d^{2}\boldsymbol{v}'$$
$$(\hat{F}_{qs}I)(\boldsymbol{v},\boldsymbol{x}) \equiv \int f_{qs}(\boldsymbol{x};\boldsymbol{v},\boldsymbol{v}')I(\boldsymbol{v}',\boldsymbol{x})d^{2}\boldsymbol{v}'$$

where  $f_d(\boldsymbol{x}; \boldsymbol{v}, \boldsymbol{v}')$ ,  $f_{qs}(\boldsymbol{x}; \boldsymbol{v}, \boldsymbol{v}')$  are the diffuse and quasi-specular BDFs at the point  $\boldsymbol{x}$  and the result of the action of operator is the local luminance.

Obviously, Eq. (7) implies

$$I^{(i)} = \left(1 - \hat{T}\hat{F}_{qs}\right)^{-1}\hat{T}\hat{F}_dI\tag{9}$$

and similarly Eq. (8) yields

$$I^{(qc)} = \left(1 - \hat{T}\hat{F}_{qs}\right)^{-1}\hat{T}\hat{F}_{qs}I^{(0)}$$
(10)

### 6.1 The form to be used for BDD=0

Now we substitute  $I^{(i)}$  from Eq. (9) into Eq. (6) which gives:

$$L = \hat{F}_{qs} \left( I^{(0)} + I^{(qc)} \right) + \hat{F}_{qs} \left( 1 - \hat{T}\hat{F}_{qs} \right)^{-1} \hat{T}\hat{F}_{d}I + \hat{F}_{d}I$$
  
$$= \hat{F}_{qs} \left( I^{(0)} + I^{(qc)} \right) + \left( 1 + \hat{F}_{qs} \left( 1 - \hat{T}\hat{F}_{qs} \right)^{-1} \hat{T} \right) \hat{F}_{d}I$$

Obviously,

$$F_{s}\left(1-\hat{T}\hat{F}_{qs}\right)^{-1}\hat{T} = F_{s}\left(\sum_{m=0}^{\infty}(\hat{T}\hat{F}_{qs})^{m}\right)\hat{T}$$
  
=  $\hat{F}_{qs}\hat{T} + (F_{s}\hat{T})^{2} + (F_{s}\hat{T})^{3} + \cdots$   
=  $\left(1-\hat{F}_{qs}\hat{T}\right)^{-1} - 1$ 

 $\mathbf{SO}$ 

$$L = \hat{F}_{qs} \left( I^{(0)} + I^{(qc)} \right) + \left( 1 - \hat{F}_{qs} \hat{T} \right)^{-1} \hat{F}_d I$$
(11)

There is an **alternative** form. From Eq. (10) it follows that

$$I^{(0)} + I^{(qc)} = I^{(0)} + \left(1 - \hat{T}\hat{F}_{qs}\right)^{-1}\hat{T}\hat{F}_{qs}I^{(0)}$$
  
=  $I^{(0)} + \left(\hat{T}\hat{F}_{qs} + (\hat{T}\hat{F}_{qs})^2 + (\hat{T}\hat{F}_{qs})^3 + \cdots\right)I^{(0)}$   
=  $\left(1 - \hat{T}\hat{F}_{qs}\right)^{-1}I^{(0)}$ 

so the term  $\hat{F}_{qs}(I^{(0)} + I^{(qc)})$  in Eqs. (11), (14) and similar below is

$$\hat{F}_{qs}\left(I^{(0)} + I^{(qc)}\right) = \hat{F}_{qs}\left(1 - \hat{T}\hat{F}_{qs}\right)^{-1}I^{(0)} 
= \left(\hat{F}_{qs} + \hat{F}_{qs}(\hat{T}\hat{F}_{qs}) + \hat{F}_{qs}(\hat{T}\hat{F}_{qs})^{2} + \cdots\right)I^{(0)} 
= \left(1 + \hat{F}_{qs}\hat{T} + (\hat{F}_{qs}\hat{T})^{2} + \cdots\right)\hat{F}_{qs}I^{(0)} 
= \left(1 - \hat{F}_{qs}\hat{T}\right)^{-1}\hat{F}_{qs}I^{(0)}$$
(12)

which is nothing but "direct visibility" (all specular reflections) of light source. Substituting Eq. (12) for the first term of Eq. (11) gives:

Substituting Eq. (12) for the first term of Eq. (11) gives:

$$L = \left(1 - \hat{F}_{qs}\hat{T}\right)^{-1} \left(\hat{F}_{qs}I^{(0)} + \hat{F}_dI\right) = \sum_{m=0}^{\infty} \left(\hat{F}_{qs}\hat{T}\right)^m \left(\hat{F}_{qs}I^{(0)} + \hat{F}_dI\right)$$
(13)

### 6.2 The form to be used for BDD=1

Substituting  $I = I^{(0)} + I^{(qc)} + I^{(i)}$  in Eq. (11)

$$L = \hat{F}_{qs} \left( I^{(0)} + I^{(qc)} \right) + \left( 1 - \hat{F}_{qs} \hat{T} \right)^{-1} \left( \hat{F}_d (I^{(0)} + I^{(qc)}) + \hat{F}_d I^{(i)} \right)$$

and evaluating  $I^{(i)}$  from Eq. (9) we have

$$L = \hat{F}_{qs} \left( I^{(0)} + I^{(qc)} \right) + \left( 1 - \hat{F}_{qs} \hat{T} \right)^{-1} \left( \hat{F}_d \left( I^{(0)} + I^{(qc)} \right) + \hat{F}_d \left( 1 - \hat{T} \hat{F}_{qs} \right)^{-1} \hat{T} \hat{F}_d I \right) = \hat{F}_{qs} \left( I^{(0)} + I^{(qc)} \right) + \left( 1 - \hat{F}_{qs} \hat{T} \right)^{-1} \left( \hat{F}_d \left( I^{(0)} + I^{(qc)} \right) + \hat{F}_d \hat{T} \left( 1 - \hat{F}_{qs} \hat{T} \right)^{-1} \hat{F}_d I \right)$$
(14)

since

$$\left(1 - \hat{T}\hat{F}_{qs}\right)^{-1}\hat{T} = \left(1 + \hat{T}\hat{F}_{qs} + (\hat{T}\hat{F}_{qs})^2 + (\hat{T}\hat{F}_{qs})^3 + \cdots\right)^{-1}\hat{T} = \hat{T} + \hat{T}\hat{F}_{qs}\hat{T} + \hat{T}\hat{F}_{qs}\hat{T}\hat{F}_{qs}\hat{T} + \hat{T}\hat{F}_{qs}\hat{T}\hat{F}_{qs}\hat{T}\hat{F}_{qs}\hat{T} + \cdots = \hat{T}\left(1 - \hat{F}_{qs}\hat{T}\right)^{-1}$$
(15)

There is an **alternative** form. We substitute Eq. (12) in the first term of Eq. (14)

$$L = \left(1 - \hat{F}_{qs}\hat{T}\right)^{-1}\hat{F}_{qs}I^{(0)} + \left(1 - \hat{F}_{qs}\hat{T}\right)^{-1}\left(\hat{F}_{d}\left(I^{(0)} + I^{(qc)}\right) + \hat{F}_{d}\hat{T}\left(1 - \hat{F}_{qs}\hat{T}\right)^{-1}\hat{F}_{d}I\right) = \left(1 - \hat{F}_{qs}\hat{T}\right)^{-1}\left(\hat{F}I^{(0)} + \hat{F}_{d}I^{(qc)} + \hat{F}_{d}\hat{T}\left(1 - \hat{F}_{qs}\hat{T}\right)^{-1}\hat{F}_{d}I\right) = \sum_{m=0}^{\infty} \left(\hat{F}_{qs}\hat{T}\right)^{m}\left(\hat{F}I^{(0)} + \hat{F}_{d}I^{(qc)}\right) + \sum_{m=0}^{\infty} \left(\hat{F}_{qs}\hat{T}\right)^{m}\hat{F}_{d}\hat{T}\sum_{n=0}^{\infty} \left(\hat{F}_{qs}\hat{T}\right)^{n}\hat{F}_{d}I$$
(16)

#### 6.3 The form to be used for BDD=2

Substituting  $I = I^{(0)} + I^{(qc)} + I^{(i)}$  in the above expression and evaluating  $I^{(i)}$  from Eq. (9) we have

$$L = \hat{F}_{qs} \left( I^{(0)} + I^{(qc)} \right) + \left( 1 - \hat{F}_{qs} \hat{T} \right)^{-1} \hat{F}_d \left( \left( I^{(0)} + I^{(qc)} \right) + \left( 1 - \hat{T} \hat{F}_{qs} \right)^{-1} \hat{T} \hat{F}_d \left( I^{(0)} + I^{(qc)} \right) + \left( \left( 1 - \hat{T} \hat{F}_{qs} \right)^{-1} \hat{T} \hat{F}_d \right)^2 I \right)$$

With the help of Eq. (15) this becomes

$$L = \hat{F}_{qs} \left( I^{(0)} + I^{(qc)} \right) + \left( 1 - \hat{F}_{qs} \hat{T} \right)^{-1} \hat{F}_d \left( \left( I^{(0)} + I^{(qc)} \right) + \hat{T} \left( 1 - \hat{F}_{qs} \hat{T} \right)^{-1} \hat{F}_d \left( I^{(0)} + I^{(qc)} \right) + \left( \hat{T} \left( 1 - \hat{F}_{qs} \hat{T} \right)^{-1} \hat{F}_d \right)^2 I \right) (17)$$

There is an **alternative** form. Substituting Eq. (12) into the 1st term of Eq. (17) we obtain

$$L = \left(1 - \hat{F}_{qs}\hat{T}\right)^{-1} \left(\hat{F}I^{(0)} + \hat{F}_{d}I^{(qc)}\right) + \left(1 - \hat{F}_{qs}\hat{T}\right)^{-1} \hat{F}_{d}\hat{T} \left(1 - \hat{F}_{qs}\hat{T}\right)^{-1} \hat{F}_{d} \left(I^{(0)} + I^{(qc)}\right) + \left(1 - \hat{F}_{qs}\hat{T}\right)^{-1} \hat{F}_{d}\hat{T} \left(1 - \hat{F}_{qs}\hat{T}\right)^{-1} \hat{F}_{d}\hat{T} \left(1 - \hat{F}_{qs}\hat{T}\right)^{-1} \hat{F}_{d}I$$
(18)

#### 6.4 The form to be used for BDD=N

Continuing the process we obtain by induction,

$$L = \hat{F}_{qs} \left( I^{(0)} + I^{(qc)} \right) + \sum_{k=0}^{N-1} \left( \left( 1 - \hat{F}_{qs} \hat{T} \right)^{-1} \hat{F}_d \hat{T} \right)^k \left( 1 - \hat{F}_{qs} \hat{T} \right)^{-1} \hat{F}_d \left( I^{(0)} + I^{(qc)} \right) + \left( \left( 1 - \hat{F}_{qs} \hat{T} \right)^{-1} \hat{F}_d \hat{T} \right)^N \left( 1 - \hat{F}_{qs} \hat{T} \right)^{-1} \hat{F}_d I$$
(19)

There is an **alternative** form. Substituting Eq. (12) into the 1st term of Eq. (19) we obtain

$$L = \left(1 - \hat{F}_{qs}\hat{T}\right)^{-1} \hat{F}_{qs}I^{(0)} + \sum_{k=0}^{N-1} \left(\left(1 - \hat{F}_{qs}\hat{T}\right)^{-1}\hat{F}_{d}\hat{T}\right)^{k} \left(1 - \hat{F}_{qs}\hat{T}\right)^{-1}\hat{F}_{d} \left(I^{(0)} + I^{(qc)}\right) + \left(\left(1 - \hat{F}_{qs}\hat{T}\right)^{-1}\hat{F}_{d}\hat{T}\right)^{N} \left(1 - \hat{F}_{qs}\hat{T}\right)^{-1}\hat{F}_{d}I$$
(20)

# 7 Integration by paths for BMCRT

The terms like  $\left(\hat{F}_{qs}\hat{T}\right)^m \left(\hat{F}_d\hat{T}\right) \left(\hat{F}_{qs}\hat{T}\right)^n \hat{F}_d I^{(i)}$  mean that the scene surface luminance is created in the following way:

- 1. *indirect* illumination is scattered by the *diffuse* BDF, then
- 2. it travels the scene, undergoing n quasi-specular events (and an arbitrary number of pure specular); then

- 3. it undergoes *diffuse* scattering, then
- 4. it travels the scene, undergoing m quasi-specular events (and an arbitrary number of pure specular)

Notice that although we write a uniform n-th power of an operator, each its term means scattering by a *specific* i.e. generally *different* (although always quasi-specular) BDF.

In BMCRT, each operator is simulated by transformation of a backward ray. Naturally the leftmost operator in the product is "the closest" to camera. So the operators are read "left to right". That is, camera ray first undergoes m quasi-specular scattering events (with subsequent ray transport to the next hit which may include pure specular transform), then *one* diffuse scattering events (with subsequent ray transport to the next hit), then n quasi-specular scattering events (with subsequent ray transport to the next hit), then n quasi-specular scattering events (with subsequent ray transport to the next hit which may include pure specular to the next hit), then n quasi-specular scattering events (with subsequent ray transport to the next hit which may include pure specular transform) here we take BDF luminance under the "indirect illumination".

Notice that a sum of several such operators can be estimated from the same camera rays, just in different points. That is, we trace backward rays and taking *indirect* illumination after m quasi-specular scattering events. Then *one* diffuse event and after n quasi-specular scattering events we accumulate the estimate for

$$\left(\hat{F}_{qs}\hat{T}\right)^{m}\left(\hat{F}_{d}\hat{T}\right)\left(\hat{F}_{qs}\hat{T}\right)^{n}\hat{F}_{d}I^{(i)}$$

while taking direct+quasi-caustic illumination after k quasi-specular scattering events (of the same ray, if its length is long enough) we accumulate the estimate for

$$\left(\hat{F}_{qs}\hat{T}\right)^k\hat{F}_d\left(I^{(0)}+I^{(qc)}\right)$$

#### 7.1 The case of BDD=0

Rewriting Eq. (11)

$$L = \hat{F}_{qs} \left( I^{(0)} + I^{(qc)} \right) + \hat{F}_{d}I + \sum_{m=1}^{\infty} \left( \hat{F}_{qs}\hat{T} \right)^{m} \hat{F}_{d}I$$
$$= \hat{F} \left( I^{(0)} + I^{(qc)} \right) + \hat{F}_{d}I^{(i)} + \sum_{m=1}^{\infty} \left( \hat{F}_{qs}\hat{T} \right)^{m} \hat{F}_{d}I$$

which means that:

- before (and including) the first not pure specular event we take  $\hat{F}(I^{(0)} + I^{(qc)}) + \hat{F}_d I^{(i)}$  or, which is the same,  $\hat{F}_{qs}(I^{(0)} + I^{(qc)}) + \hat{F}_d I$ .
- after the first quasi-specular event and up to the first diffuse event we take  $\hat{F}_d I$ ;
- at the first diffuse event we stop

Alternatively, Eq. (13) means that:

- from the **first** hit of camera ray and up to the first diffuse event we take  $\hat{F}_{qs}I^{(0)} + \hat{F}_dI$ ;
- at the first diffuse event we stop

#### 7.2 The case of BDD=1

Rewriting Eq. (14)

$$L = \hat{F} \left( I^{(0)} + I^{(qc)} \right) + \sum_{m=1}^{\infty} \left( \hat{F}_{qs} \hat{T} \right)^m \hat{F}_d \left( I^{(0)} + I^{(qc)} \right)$$
$$+ \sum_{m,n=0}^{\infty} \left( \hat{F}_{qs} \hat{T} \right)^m \hat{F}_d \hat{T} \left( \hat{F}_{qs} \hat{T} \right)^n \hat{F}_d I$$

That is,

- before (and including) the first not pure specular event we take  $\hat{F}(I^{(0)} + I^{(qc)});$
- after the first quasi-specular event and up to the first diffuse event we take  $\hat{F}_d \left( I^{(0)} + I^{(qc)} \right)$ ;
- *after* the first diffuse event we take  $\hat{F}_d I$
- at the second diffuse event we stop

Alternatively, Eq. (4alt) means that

- up to the 1st diffuse event we take  $\hat{F}I^{(0)} + \hat{F}_d I^{(qc)}$ ;
- after the 1st diffuse event we take  $\hat{F}_d I$
- at the 2nd diffuse event we stop

#### 7.3 The case of BDD=2

Rewriting Eq. (17)

$$L = \hat{F} \left( I^{(0)} + I^{(qc)} \right) + \sum_{m=1}^{\infty} (\hat{F}_{qs}\hat{T})^m \hat{F}_d \left( I^{(0)} + I^{(qc)} \right) + \sum_{m,n=0}^{\infty} \left( \hat{F}_{qs}\hat{T} \right)^m \hat{F}_d \hat{T} \left( \hat{F}_{qs}\hat{T} \right)^n \hat{F}_d \left( I^{(0)} + I^{(qc)} \right) + \sum_{k,m,n=0}^{\infty} \left( \hat{F}_{qs}\hat{T} \right)^k \hat{F}_d \hat{T} \left( \hat{F}_{qs}\hat{T} \right)^m \hat{F}_d \hat{T} \left( \hat{F}_{qs}\hat{T} \right)^n \hat{F}_d I$$

that is,

- before (and including) the first not pure specular event we take  $\hat{F}(I^{(0)} + I^{(qc)});$
- after the first quasi-specular event and up to the second diffuse event we take  $\hat{F}_d \left( I^{(0)} + I^{(qc)} \right)$ ;
- after the second diffuse event we take  $\hat{F}_d I$
- at the third diffuse event we stop

Alternatively, Eq. (18) means that

- up to the first diffuse event we take  $\hat{F}I^{(0)} + \hat{F}_dI^{(qc)}$ ;
- after the first diffuse event and up to the 2nd diffuse event we take  $\hat{F}_d \left( I^{(0)} + I^{(qc)} \right);$
- after the second diffuse event we take  $\hat{F}_d I$
- at the third diffuse event we stop

#### 7.4 General case of BDD = N

By induction,

• before (and including) the first not pure specular event we take  $\hat{F}(I^{(0)} + I^{(qc)});$ 

In which hit points and for which illumination component the "genuine diffuse" BDF part  $\hat{F}_d$  is used instead of the full  $\hat{F}$ .

	Main variant	Alternative variant
Direct+caustic	After the 1st quasi-specular event	After the 1st diffuse event
Quasi-caustic		Always
Indirect	Always	

- after the first quasispecular event and up to the N-th diffuse event we take  $\hat{F}_d (I^{(0)} + I^{(qc)});$
- after the N-th diffuse event we take  $\hat{F}_d I$
- at the (N+1)-th diffuse event we stop

#### Alternatively,

- up to the first diffuse event we take  $\hat{F}I^{(0)} + \hat{F}_dI^{(qc)}$ ;
- after the first diffuse event and up to the N-th diffuse event we take  $\hat{F}_d \left( I^{(0)} + I^{(qc)} \right);$
- after the N-th diffuse event we take  $\hat{F}_d I$
- at the (N+1)-th diffuse event we stop

Therefore, in both variants camera ray stops at the (N+1)-th diffuse event. In the intermediate hit point, if the surface is not pure specular, it takes a *part* of the local luminance. Table 1 summarizes which one.

# 8 Volumetric scattering

# 8.1 Standard method

The "standard" implementation of volume scattering is that volumetric scattering event is processed like a surface one. For example, suppose that camera is inside a turbid medium and BDD=1. Then camera ray propagates in the medium and when it undergoes the first volumetric scattering. It takes *direct* and *caustic* illumination and "convolves" it with the phase function. In the second volumetric scattering it takes *full* illumination and "convolves" it with the phase function. Let us consider a model scene in which camera looks through a layer of turbid medium onto some illuminated object. Let also the turbid medium is absorption-free and has a large scattering coefficient, so that camera ray undergoes many (volumetric) scattering events before it penetrates the medium and reaches the object. Let the phase function be sharp enough so that each scattering changes the ray direction only slightly.

If we apply the above "standard" method with BDD=0 to this example, then camera ray terminates at the first volumetric scattering and collects the full illumination. Because of high scattering coefficients, this happens close to the ray origin (i.e. to the camera), and the ray does not reach the object. Since the phase function is narrow, its convolution with that illumination which came from the scene surfaces results in strong noise, cf. Section 4. For BDD=1 the situation does not change much.

We must set a *large* BDD so that the camera ray *leaves* the medium, reaches a scene surface behind and also not to collect indirect illumination in the volumetric scene points which due to sharp phase function creates strong noise. But then we must keep in memory all the volumetric scattering points (because they still collect direct and caustic illumination). This is usually too expensive.

#### 8.2 "Quasi-specular medium"

Applying the quasi-specular approach to the volumetric scattering greatly improves the situation with the above scene.

Suppose that the scene *surfaces* are not quasi-specular. Let also the *whole* phase function be treated as quasi-specular, i.e. its "genuine diffuse" part  $\hat{F}_d = 0$ .

Then, , it undergoes only quasi-specular events until the camera ray travels inside the medium and thus does not increment the "diffuse counter". As a result it penetrates the medium layer and reaches for a scene surface.

*Indirect* illumination now is the light after diffuse scattering by scene surfaces. *Quasi-caustic* illumination is the light which underwent at least one volumetric scattering and any specular events, but no diffuse surface scattering.

Since  $\hat{F}_d = 0$ , the "main variant" from Table 1 implies that:

- the *indirect* illumination (=from the scene surfaces, since the volumetric scattering is completely quasi-caustic) is effectively *ignored* inside the medium
- the *direct, caustic* and *quasi-caustic* illumination is taken only up to the *first* volumetric scattering, i.e. in fact for the *first* volumetric event only.

As a result, for *any* BDD the camera ray reaches the scene surfaces behind the medium layer. Only *one* ray event (the first volumetric) is remembered. There is no noisy convolution of the sharp phase function with the indirect illumination from scene surfaces.

Although there is the convolution of phase function with the direct, caustic and scattered by the medium light. But they all do not have a *wide* angular distribution that would result in strong noise.

Now let us look at the "alternative" variant from Table 1. Here, the direct and caustic illumination *will be* taken "up to the first *diffuse* event". Since the volumetric scattering is treated as quasi-specular, it happens only *after* the camera rays leaves the medium. In other words, we must remember *all* the ray scattering events inside the medium to take direct and caustic in that points. This is expensive in memory. So advantageous is only the "main variant".

# 9 Results

### 9.1 Surface case

Calculations were done for the model scene from Fig. 5. They were performed for BDD=0 and the same other conditions, including run time etc. as those for the "standard mode" (without quasi-specular).

The results are presented in the right panel of Fig. 6. One can see the noise level is about threefold lower than for the best case without the quasi-specular method. And what is more essential, the calculated image has a better estimation of color:

- for BDD=0 the image looks mainly red with rare bright green dots;
- for BDD=1 it looks mainly green with rare bright red dots;
- and only with th quasi-specular method we see a "mixture" of red and green.

In the meanwhile the *averages* obtained in the three simulations are about the same. It means that the result e.g. for the "standard BDD=1" has very bright red dots which "on average" would give the correct value. That is, in the standard method achieves the correct average by very rare very bright peaks, i.e. the worst sort of noise.

#### 9.2 Volumetric case

The calculation had been performed for the model scene which consists of a plane parallel plate of thickness 3 mm laid upon a paper sheet with a chessboard-



Figure 7: Camera images for the scene with a plate of turbid medium laid upon a chessboard-like texture. The left image is the "standard mode" of volumetric scattering, when it is treated as diffuse. The right panel is when the volumetric scattering is purely quasi-specular. In both cases BDD=1.

like texture, illuminated by a self-emitting sphere above it. The plate's medium has refraction index 1.5 and scattering coefficient 7.5  $mm^{-1}$ . The phase function is the Henyey-Greenstein one [3] with g = 0.9, i.e. it is rather directional scattering in a narrow cone about the incident direction. The images calculated during the same time (and with the same settings) are presented in Fig. 7. One can see that in the "standard mode" the boundaries of the texture squares, seen through the plate, are *sharp* while in reality they must blur. Meanwhile in the image calculated with the quasi-specular representation of the phase function these boundaries are smoothed.

# 10 Conclusion

Proposed method of treating some scene elements (surface or volumetric) as "quasi-specular" is a flexible tool. The separation can be done using arbitrary criterion which can be own for each BDF. For example, if a BDF contains a sharp peak in an *off-specular direction*, this can be nevertheless declared quasispecular. For the selected separation the method admits at least two variants (see Table 1) which differ in processing direct, caustic and quasi-specular illumination components. Apparently there are more such variants which differ in BMCRT implementation. And in different situations advantageous are different methods. We demonstrate how the quasi-specular method can reduce noise for several scenes which cannot be calculated efficiently with the "standard" method for any value of the BDD parameter.

### References

- [1] Д.Д. Жданов, С.В. Ершов, А.Г. Волобой. Адаптивный выбор глубины трассировки обратного луча в методе двунаправленной стохастической трассировки лучей. Proceedings of 25-th International Conference on Computer Graphics and Vision, September 22-25, 2015, pages 44-49.
- [2] Yves D. Willems, Eric P. Lafortune. Bi-directional path tracing. Proceedings of Third International Conference on Computational Graphics and Visualization Techniques (Compugraphics '93), Alvor, Portugal, December 1993, pages 145– 153, 1993.
- [3] Pomraning G.C. On the henyey-greenstein approximation to scattering phase functions. Journal of Quantitative Spectroscopy and Radiative Transfer, 39(2):109 - 113, 1988.
- [4] Henrik Wann Jensen. Global illumination using photon maps. Proceedings of the Eurographics Workshop on Rendering Techniques '96, pages 21–30, London, UK, 1996. Springer-Verlag.
- [5] James T. Kajiya. The rendering equation. Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '86, pages 143–150, New York, NY, USA, 1986. ACM.
- [6] Tzu-Mao Li, Jaakko Lehtinen, Ravi Ramamoorthi, Wenzel Jakob and Frédo Durand. Anisotropic gaussian mutations for metropolis light transport through Hessian-Hamiltonian dynamics. ACM Trans. Graph., 34(6):209:1– 209:13, October 2015.
- [7] Ryusuke Villemin, Sebastien Speierer, Christophe Hery, Wenzel Jako. Caustic connection strategies for bidirectional path tracing. Pixar technical memo #18-01, 2018.
- [8] Eric Veach. A dissertation: Robust Monte-Carlo methods for light transport simulation, 1997.
- [9] Eric Veach, Leonidas J. Guibas. Metropolis light transport. Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '97, pages 65–76, 1997.