Interner Bericht

Quasi-Monte Carlo Methods in Computer Graphics, Part II: The Radiance Equation

Stefan Heinrich
Alexander Keller

Fachbereich Informatik

Universität Kaiserslautern · Postfach 3049 · D-67653 Kaiserslautern

Quasi-Monte Carlo Methods in Computer Graphics, Part II: The Radiance Equation

Stefan Heinrich
Alexander Keller

Fachbereich Informatik
AG Numerische Algorithmen
Universität Kaiserslautern
Postfach 3049, D-67653 Kaiserslautern
e-mail: heinrich@informatik.uni-kl.de
keller@informatik.uni-kl.de

243/94

April 11, 1994

Abstract

The radiance equation, which describes the global illumination problem in computer graphics, is a high dimensional integral equation. Estimates of the solution are usually computed on the basis of Monte Carlo methods. In this paper we propose and investigate quasi-Monte Carlo methods, which means that we replace (pseudo-) random samples by low discrepancy sequences, yielding deterministic algorithms. We carry out a comparative numerical study between Monte Carlo and quasi-Monte Carlo methods. Our results show that quasi-Monte Carlo converges considerably faster.

1 Introduction

The radiance equation (REQ) describes the global illumination problem in computer graphics: Given a scene and light sources, determine the illumination of the objects, caused by direct light, multiple reflection, refraction etc., i.e. determine the cumulative light radiation from each point on the surface of the objects into each direction. The REQ is an integral equation of the second kind with an unknown function of four independent variables (two coordinates for the location of the point on the scene surface and two for the direction). This high complexity makes classical numerical methods essentially inapplicable. Moreover, due to different optical and geometric properties of the objects of the scene (color, reflectivity, mutual visibility), the involved functions possess many discontinuities. One way of circumventing these difficulties is the simplification of the REQ, as e.g. the radiosity equation (here the radiance is assumed to be independent of direction), but in complex scenes even the solution of that equation is still a highly complicated task.

We are concerned with the general, i.e. four-dimensional case. As it is usual in high dimensional problems, Monte Carlo methods provide a way of solution. For rendering images, this technique has been explored intensively (see the survey [MUD93]), including variance reduction techniques and combinations with other methods like radiosity. There is an alternative, though related approach to such problems - quasi-Monte Carlo methods. This is a technique for multivariate integration, in which the random samples of Monte Carlo are replaced by certain particularly designed, so-called low discrepancy sequences. These have the property of optimally approximating the uniform distribution in high dimensions (see [NIE78], [NIE92]). Quasi-Monte Carlo methods have not yet been explored in computer graphics. In [HEI94] we investigated the relevance of such methods for the purpose of pixel oversampling. Discrepancy as a quality measure for sampling patterns was already discussed in [SHI91]. The use of low discrepancy sequences in particle transport problems was first suggested in [NEU73].

The aim of this paper is to apply quasi-Monte Carlo techniques to the solution of the REQ. We use the Neumann series in a standard way to represent the solution as a sum of integrals of increasing dimension. Then we approximate these integrals numerically by quasi-Monte Carlo methods.

Convergence theory of Monte Carlo methods is based on probabilistic arguments. The sequence of random numbers supplied by a computer is, in fact, a deterministic one - usually based on linear congruences. So one hopes (and this often seems to be the case) that this sequence lies in the large set of "good" realizations for the particular problem in consideration. But no rigorous error bounds can be derived. For quasi-Monte Carlo methods the situation is different: Under certain regularity

assumptions on the integrand (bounded variation) a convergence faster than the expected rate of Monte Carlo can be proved (see [NIE92]). However, these assumptions usually do not hold in scenes of computer graphics. Consequently, theory can neither predict, nor compare the efficiency of both approaches so far. This led us to design and carry out certain numerical experiments in which the efficiency is investigated for a test scene. (Note that comparative studies between Monte Carlo and quasi-Monte Carlo for one-dimensional particle transport through a slab were reported in [FOX86] and [SAR87]).

In our experiments we had both methods computing certain integrals based on the global illumination of our scene. It turned out that the low discrepancy sequences performed very well - they provided considerably more precise approximations than the (pseudo-) random sequences. This suggests quasi-Monte Carlo methods as a promising ingredient in algorithms for the global illumination problem.

We plan to explore this algorithmic point further and to study the behaviour of quasi-Monte Carlo methods in greater depth. Theoretical analysis seems to be a very difficult task. Maybe the average behaviour of low discrepancy sequences as studied recently in information-based complexity theory [WOZ91] could give some more insight (here suitable averages over multivariate discontinuous functions should be taken).

2 The Radiance Equation

We consider the illuminance of a scene, which is assumed to be closed and to consist of closed bodies. Their surfaces form the set S. The radiance equation (see [KAJ86], [CHR93]) is formulated as follows:

$$L(y,z) = L_0(y,z) + \int_S \rho(x,y,z)\, G(x,y)\, L(x,y)\, dx \quad (y,z \in S) \tag{1}$$

Here $L(y,z)$ is the unknown function on $S \times S$. It describes the radiance of a given fixed wavelength emitted from $y \in S$ into the direction of $z \in S$ (energy per unit of time per unit spherical angle per unit of projected area - measured in $\left[\frac{W}{m^2 sr}\right]$). $L_0(y,z)$ is the radiance of the light sources. The function $\rho(x,y,z)$ expresses the optical properties of the surfaces. It describes, which portion of light arriving in y from direction of x is scattered into direction of z. If only reflections are considered, ρ is called the bidirectional reflectance distribution function BRDF. Finally, $G(x,y)$ is related to the scene geometry:

$$G(x,y) = V(x,y)\frac{\cos\vartheta_x \cos\vartheta_y}{\|x-y\|^2} \ .$$

where $V(x,y) = 1$ if x and y are mutually visible and $= 0$ otherwise. The angle between the surface normal $\hat{n}_x$ in x and the vector $y - x$ is denoted by ϑ_x, that between $\hat{n}_y$ and $x - y$ by ϑ_y. We abbreviate (1) and write it as

$$L = L_0 + T_\rho L \ , \tag{2}$$

where $T_\rho L$ stands for the integral in (1). Under realistic physical assumptions (no ideal mirrors in the scene), one can prove the convergence of the Neumann series in suitable norms. Hence in this case it is justified to approximate the value of the solution L in $(y,z) \in S^2$ by a finite sum

$$L(y,z) \approx \sum_{m=0}^{M} (T_\rho^m L_0)(y,z) \ . \tag{3}$$

More generally, scalar products of the solution L with a given function Ψ on S^2 can be approximated by

$$< L, \Psi > \approx \sum_{m=0}^{M} < T_\rho^m L_0, \Psi > \tag{4}$$

(inserting the delta function $\delta_{(y,z)}$ in (4) gives (3) as a special case). In the sequel we shall consider (4). Rewriting the powers of T_ρ into integrals gives

$$\begin{aligned} &\sum_{m=0}^{M} < T_\rho^m L_0, \Psi > \\ &= \sum_{m=0}^{M} \int_{S^{m+2}} L_0(x_0,x_1)\ \rho(x_0,x_1,x_2)\ G(x_0,x_1) \cdots \\ &\qquad \cdots \rho(x_{m-1},x_m,x_{m+1})\ G(x_{m-1},x_m)\ \Psi(x_m,x_{m+1})\ dx_0 \cdots dx_{m+1} \ . \end{aligned} \tag{5}$$

Algorithms for the computation of such high dimensional integrals will be described in the next section.

3 Quasi-Monte Carlo Integration

The Monte Carlo method approximates a multidimensional integral as follows

$$\int_{[0,1]^s} f(u)\ du \approx \frac{1}{N} \sum_{i=0}^{N-1} f(u_i) \ . \tag{6}$$

The (u_i) are realizations of independent uniformly distributed on $B = [0,1]^s$ random variables. Quasi-Monte Carlo methods approximate the integral in the same way (6), but with a deterministic pattern (u_i). Such a pattern is designed specifically for integration - the samples should be as close to the uniform distribution on B as possible. This goal is achieved by low discrepancy sequences. As we already mentioned, on the computer Monte Carlo methods are carried out using deterministic samples, usually produced by linear congruential generators (pseudo-random numbers). The difference is the following: Pseudo-random numbers are supposed to mimic realizations of true random variables and as such are designed to serve various statistical purposes and to satisfy corresponding tests. Quasi-random (or low discrepancy) sequences serve the unique purpose of optimally integrating certain classes of multivariate functions.

The star-discrepancy is a measure of the uniform distribution of a given (deterministic) point set $P_N = \{u_0, \ldots, u_{N-1}\}$. It is defined as

$$D^*(P_N) = \sup_A \left| \lambda_s(A) - \frac{m(A, P_N)}{N} \right| , \tag{7}$$

where the supremum is taken over all sets A of the form $A = \prod_{i=1}^{s} [0, a_i]$, with $0 \leq a_i \leq 1$, λ_s is the Lebesgue measure on B, and

$$m(A, P_N) = \sum_{i=0}^{N-1} \chi_A(u_i) \tag{8}$$

is the number of elements of P_N which fall into A.

The Koksma-Hlawka inequality for integration on B states that for any set $P_N = \{u_0, \ldots,$
$u_{N-1}\} \subset B$

$$\left| \int_B f(u)\, du - \frac{1}{N} \sum_{i=0}^{N-1} f(u_i) \right| \leq V(f) \cdot D^*(P_N) . \tag{9}$$

Here $V(f)$ is the total variation of f in the sense of Hardy and Krause (for further details see [NIE92]). Due to the discontinuities of ρ and L_0 we cannot apply (9) directly, since the total variation $V(f)$ may become infinite. Instead we have to integrate over many subregions of the unit cube, and we need an analogue of (9) for this situation, which is subject of our further investigations. So for the moment let us assume that we are integrating functions of bounded total variation, i.e. $V(f) < \infty$. Then the error depends on the discrepancy $D^*(P_N)$.

Let us now recall two standard types of low discrepancy sequences. First, we

define Φ as the radical inverse function:

$$\Phi_b(i,\sigma) = \sum_{j=0}^{\infty} \sigma(a_j(i))\, b^{-j-1} \text{ when } i = \sum_{j=0}^{\infty} a_j(i)\, b^j$$

where the natural number $b > 1$ is the base, $i \in I\!N$, and σ is a permutation of the set $\{0, 1, \ldots, b-1\}$. The values Φ_b are always in the unit interval $[0, 1)$. To see how Φ acts on a natural number i we assume σ to be the identity. Then $\Phi_2(i)$ is simply the binary representation of i mirrored at the decimal point [1].

The Halton and the Hammersley sequence are s-dimensional vectors built from radical inverse functions in relatively prime bases b_j. For $0 \leq i < N$ we have:

$$\begin{aligned} \text{Halton points:} \quad u_i &= (\Phi_{b_1}(i), \ldots, \Phi_{b_s}(i)) \\ \text{Hammersley points:} \quad u_i &= (\frac{i}{N}, \Phi_{b_1}(i), \ldots, \Phi_{b_{s-1}}(i)) \end{aligned}$$

The points all lie inside the s-dimensional unit cube $B = [0, 1]^s$ by the definition of the radical inverses. If the permutations are not identities, the sequences are called scrambled. The discrepancies $D^*(P_N)$ for both sets are bounded by (see [NIE92]):

$$\begin{aligned} D^*_{Halton} &< \frac{s}{N} + \frac{1}{N} \prod_{j=1}^{s} \left(\frac{b_j - 1}{2 \log b_j} \log N + \frac{b_j + 1}{2} \right) \\ &\Rightarrow D^*_{Halton} \in O\left(\frac{\log^s N}{N} \right) \\ D^*_{Hammersley} &< \frac{s}{N} + \frac{1}{N} \prod_{j=1}^{s-1} \left(\frac{b_j - 1}{2 \log b_j} \log N + \frac{b_j + 1}{2} \right) \\ &\Rightarrow D^*_{Hammersley} \in O\left(\frac{\log^{s-1} N}{N} \right) . \end{aligned}$$

The Halton sequence is an incremental pattern, meaning that increasing the number of samples is possible without discarding the samples already drawn. On the contrary the Hammersley sequence is not incremental since increasing the sampling rate results in discarding all samples computed so far. In consequence this pattern is not useful for adaptive sampling, although it has an asymptotically smaller discrepancy. No values of either sets need to be precomputed, since all values are directly computable.

[1] For fast algorithms for computing Φ we refer to [STR93] or [HAL64]

4 Quasi-Monte Carlo Methods for the Radiance Equation

In order to apply the technique described above we have to transform the integrals in (5) to the unit cube. There are, of course, various ways. We shall pass to local spherical coordinates (φ, ϑ). Given $x \in S$ and

$$(\varphi, \vartheta) \in Q = [0, 2\pi] \times [0, \frac{\pi}{2}]$$

we define $h(x, \varphi, \vartheta)$ to be the closest point to x in which a ray from x in direction (φ, ϑ) intersects S. Next we specify, which type of functions Ψ will be used in our numerical experiments: Given a subset $A \subset S$, we compute the total radiance that reaches A. It is given by the integral

$$\int_{S^2} L(y,z)\, G(y,z)\, \chi_A(z)\, dy\, dz\ .$$

Thus, we consider functions Ψ of the form

$$\Psi(y,z) = G(y,z)\, \chi_A(z)\ . \tag{10}$$

Now we transform $< T_\rho^m L_0, \Psi >$. We start with $m = 0$:

$$\begin{aligned} &< T_\rho^0 L_0, \Psi > \\ &\quad = \int_{S^2} L_0(x_0, x_1)\, G(x_0, x_1)\, \chi_A(x_1)\, dx_0\, dx_1 \\ &\quad = \int_0^{\frac{\pi}{2}} \int_0^{2\pi} \int_{S_0} L_0(x_0, h(x_0, \varphi_0, \vartheta_0))\, \cos\vartheta_0\, \sin\vartheta_0\, \chi_A(h(x_0, \varphi_0, \vartheta_0))\, dx_0\, d\varphi_0\, d\vartheta_0\ . \end{aligned}$$

Here S_0 is the surface of light sources. Proceeding similarly with $m > 0$, we get

$$\begin{aligned} &< T_\rho^m L_0, \Psi > \\ &\quad = \int_{Q^{m+1}} \int_{S_0} g(x_0, \varphi_0, \vartheta_0, \ldots, \varphi_m, \vartheta_m)\, dx_0\, d\varphi_0\, d\vartheta_0 \cdots d\varphi_m\, d\vartheta_m \end{aligned} \tag{11}$$

where

$$\begin{aligned} &g(x_0, \varphi_0, \vartheta_0, \ldots, \varphi_m, \vartheta_m) \\ &\quad = L_0(x_0, x_1) \left(\prod_{j=0}^{m-1} \rho(x_j, x_{j+1}, x_{j+2}) \right) \left(\prod_{j=0}^{m} \cos\vartheta_j\, \sin\vartheta_j \right) \chi_A(x_{m+1}) \end{aligned} \tag{12}$$

with

$$x_k = h(x_{k-1}, \varphi_{k-1}, \vartheta_{k-1}) \ . \tag{13}$$

The terms $\cos\vartheta_j \ \sin\vartheta_j = \frac{1}{2}\sin 2\vartheta_j$ can be eliminated by the substitution

$$\tau_j = \frac{1}{2}(1 - \cos 2\vartheta_j) \ ,$$

as we do in our computations. Let us now for simplicity assume that the light sources are rectangular. (If they are not, a transformation has to be found in such a way, that the parameter domain is a rectangle, look e.g. [GRE90] or [PAT93].) Now we reached an integration domain being - up to scaling - the unit cube. Let

$$z_i = (z_{i1}, \ldots, z_{i(2M+4)}) \in [0,1]^{2M+4}$$

$(i = 0, \ldots, N-1)$ be a low discrepancy sequence. The quasi-Monte Carlo algorithm for the computation of (5) and (14) is the following:

1. For each i, compute the values of g in (12) and (13) using the first (scaled) $(2m+4)$ components of z_i.

2. Sum over $m = 0, \ldots, M$ and average over $i = 0, \ldots, N-1$.

Now let us look at standard Monte Carlo: By choosing N realizations of the vector $(x_0, \varphi_0, \vartheta_0, \ldots, \varphi_m, \vartheta_m)$ with components independent and uniformly distributed in their respective domains, we obtain a version of direct particle simulation: The "photon" starts at a random position on the light surfaces and moves through the scene having a fixed number of reflections. (A portion of radiance is assigned to it in the beginning, which is diminished at each reflection according to the BRDF). Viewed from this point, quasi-Monte Carlo replaces the usage of the random number generator by the components of certain deterministic sequences, but the principal nature of the simulation (random walk) is preserved.

5 Numerical Results

Our goal is the comparison of quasi-Monte Carlo with Monte Carlo. We therefore take the approximation (4) as our starting point, that is, we investigate the precision, with which the random or quasi-random samples approximate

$$\sum_{m=0}^{M} < T_\rho^m L_0, \Psi > \ . \tag{14}$$

(precision of the truncation of the Neumann series is a different topic and will not be discussed here). We consider a test scene consisting of 276 triangular elements (see table 1, in the experiments the test scene was used without textures, so that the results reflect the basic global illumination quantities). We consider light of the usual three basic colors red, green and blue, that is, the radiance $L = (r, g, b)$ consists now of three components. We fix $M = 4$ and let A in (10) run through all the triangles of the scene. That is, we check, how precisely the radiance arriving at A is computed by different methods. This gives some quality measure for the precision of the particle simulation, an idea which goes back to [NEU73]. Since, of course, no explicit exact solution to (14) is known, we carry out master computations with a very high number of particles (in both Monte Carlo and quasi-Monte Carlo, until a satisfactory coincidence is reached). These master values are then compared with results of the different kinds of approximations using lower numbers of particles.

A measurement $\mathcal{M}$ is the set of radiances L_i cumulated for the surface elements A_i. For two measurements $\mathcal{M}$ and $\mathcal{M}'$ we define the mean square deviation by

$$\|\mathcal{M} - \mathcal{M}'\|_2 = \sqrt{\frac{\sum_i |A_i| d^2(L_i, L_i')}{\sum_i |A_i|}}$$

where the distance d of two radiances L_i and L_i' is defined by

$$d(L_i, L_i') = \sqrt{(r_i - r_i')^2 + (g_i - g_i') + (b_i - b_i')}$$

and r, g and $b \in [0, 1]$ are the color intensities of the color base red, green and blue. We performed the measurements for $N = 10^4, 10^5, 10^6$ and 10^7 particles and for all patterns mentioned before. In tables 2 to 5 the convergence rates of the different sampling strategies are listed. A superiority of the deterministic patterns can be observed. Their convergence is faster than the convergence of random sampling. Also, it can be seen, that the Hammersley point set, which has the least order of discrepancy, performs best.

To illustrate that the sampling patterns calculate the same result, we compared the deviations of different patterns for the master computations $N = 10^7$ in table 6.

6 Conclusion and Further Work

In this article we proposed the use of quasi-Monte Carlo methods for computer graphics, especially for the solution of the radiance equation. By numerical evidence it is shown, that these methods promise a better estimation of the true solution than Monte Carlo methods. Since quasi-Monte Carlo methods use low discrepancy point sets, deterministic upper error bounds hold for functions of bounded variation.

Table 1: Test scene

N	10^4	10^5	10^6
10^5	0.00430451	-	-
10^6	0.00361428	0.00130049	-
10^7	0.00373137	0.00136592	0.000339878

Table 2: Mutual deviation of different sampling rates for the random sequence

N	10^4	10^5	10^6
10^5	0.00111829	-	-
10^6	0.00101867	0.000326033	-
10^7	0.00101063	0.000347712	0.0000798317

Table 3: Mutual deviation of different sampling rates for the Halton sequence

N	10^4	10^5	10^6
10^5	0.00111391	-	-
10^6	0.00115895	0.00042524	-
10^7	0.00114342	0.000421544	0.0000678782

Table 4: Mutual deviation of different sampling rates for the scrambled Halton sequence

N	10^4	10^5	10^6
10^5	0.00111804	-	-
10^6	0.00112252	0.000296641	-
10^7	0.00114025	0.000323288	0.0000636086

Table 5: Mutual deviation of different sampling rates for the Hammersley sequence

$\mathcal{M}$	$\mathcal{M}'$	$\|\mathcal{M} - \mathcal{M}'\|_2$
Random	Halton	0.000100943
Random	Scrambled Halton	0.000103805
Random	Hammersley	0.000110742
Halton	Scrambled Halton	0.0000358037
Halton	Hammersley	0.0000294866
Scrambled Halton	Hammersley	0.0000300549

Table 6: Deviation of different sampling patterns for the master computations. $N = 10^7$

For the particular situation of global illumination, a theoretical foundation and error bounds are still an open problem. Progress might be possible by combining results of quasi-Monte Carlo theory [NIE92] and information-based complexity [TRA88].

Finally, stimulated by the good practical behaviour of low discrepancy sequences, we are designing and testing new illumination algorithms based on quasi-Monte Carlo techniques.

References

[CHR93] P.H.Christensen, D.H.Salesin, T.D.DeRose: A continuous adjoint formulation for radiance transport, in Fourth Eurographics Workshop on Rendering, 1993, pp. 95-104.

[FOX86] B.L.Fox: Algorithm 647: Implementation and relative efficiency of quasi-random sequence generators, ACM Trans. Math. Software, 12(1986), pp. 362-376.

[GRE93] G.Turk: Generating random points in triangles, in Graphics Gems, Ed. A.Glassner, Academic Press, 1990.

[HAL64] J.H.Halton, G.Weller: Algorithm 247: Radical-inverse quasi-random point sequence [G5], Comm. ACM, No.12, 7(1964), pp. 701-702.

[HEI94] S.Heinrich, A.Keller: Quasi-Monte Carlo methods in computer graphics, Part I: The QMC-Buffer, Technical Report, University of Kaiserslautern, to appear.

[KAJ86] J.T.Kajiya: The rendering equation, Computer Graphics, No.4, 20(1986), pp. 143-150.

[MUD93] S.P.Mudur, S.N.Pattanaik: Monte Carlo methods for computer graphics, in State of the Art Reports, Eurographics, 1993.

[NEU73] H.Neunzert, J.Wick: Die Theorie der asymptotischen Verteilung und die numerische Lösung von Integrodifferentialgleichungen, Numer. Math., 21(1973), pp. 234-243.

[NIE78] H.Niederreiter: Quasi-Monte Carlo methods and pseudo-random numbers, Bull. Amer. Math. Soc., 84(1978), pp. 957-1041.

[NIE92] H.Niederreiter: Random Number Generation and Quasi-Monte Carlo Methods, SIAM Philadelphia, Pennsylvania 1992.

[PAT93] S.N.Pattanaik: Computational Methods for Global Illumination and Visualization of Complex 3D Environments, Ph.D. Thesis, Birla Institute of Technology & Science, Pilani, 1993.

[SAR87] P.K.Sarkar, M.A.Prasad: A comparative study of pseudo and quasi random sequences for the solution of integral equations, J. Comp. Physics, 68(1987), pp. 66-88.

[SHI91] P.Shirley: Discrepancy as a quality measure for sample distributions, Eurographics '91, F.H.Post and Barth (Editors), Elsevier Science Publishers B.V. (North-Holland), pp. 183-195.

[STR93] J.Struckmeier: Fast generation of low-discrepancy sequences, Berichte der Arbeitsgruppe Technomathematik, Nr.93, Fachbereich Mathematik, Universität Kaiserslautern.

[TRA88] J.F.Traub, G.W.Wasilkowski, H.Woźniakowski: Information-Based Complexity, Academic Press, 1988.

[WOZ91] H.Woźniakowski: Average case complexity of multivariate integration, Bull. Amer. Math. Soc., 24(1991), pp. 185-194.