A Survey of Data Driven Methods in Realistic Image Synthesis

J. Vorba, O. Karlík
Charles University, Faculty of Mathematics and Physics, Prague, Czech Republic.

Abstract. Producing photo-realistic images in computer graphics is a computationally demanding task based on physically plausible simulation of light transport in the artificial scene. Many articles published in the last 20 years try to increase the efficiency of this simulation by obtaining and exploiting a partial knowledge about the importance of light transport paths in the scene. We present a survey of these methods, which is the first step of our research in this area.

Figure 1. Scene descriptions (left) and photorealistic images created from them (right).

Introduction

The goal of realistic image synthesis (Dutré et al. [2006]) is to create visually plausible images of virtual scenes, indistinguishable from real world photographs (see Figure 1). Its applications span a wide range of fields, including film, games, architecture, and marketing industries. High-quality images are obtained by performing physically based simulation of light transport in the scene. This transport was formalized in form of the rendering equation by Kajiya [1986], but in this paper we use its path formulation by Veach [1998], which simplifies the original recursive integral formulation to simple non-recursive integral over the space of all

Figure 2. An example of a light path, starting on a light source \( x_0 \), bouncing two times in the scene \( x_1 \) and \( x_2 \), and ending on the eye/sensor \( x_3 \). The emission of light is described by light emission function \( L_e \), response of camera by the measurement function \( W_e \). Reflective properties of surfaces at all scene interactions (inner vertices of the path) are described by the bidirectional reflectance distribution function \( f_s \), and the transport of light between all adjacent vertices is described by the geometry term \( G \). The measurement contribution function \( f_j \) is a product of all these terms. The image is courtesy of Veach [1998].
VORBA AND KARLÍK: DATA-DRIVEN METHODS IN REALISTIC IMAGE SYNTHESIS

light paths Ω. A light path is a trajectory of hypothetical particle emitted from a light source, that bounces off the scene surfaces, and ends in the camera. Energy transmitted by each light path from light sources to camera is given by the measurement contribution function $f_j$ which is a product of few terms (see Figure 2). The color of each pixel $I_j$ is then simply an integral over the space of all light paths $Ω$:

$$I_j = \int_{Ω} f_j(\bar{x})d\mu(\bar{x}).$$  \hspace{1cm} (1)

Solving the rendering equation analytically is impossible in all but the simplest scenes (Dutré et al. [2006]). Monte Carlo integration is used instead: $N$ random samples (light paths) are generated according to some probability density function ($pdf$) $p$, evaluate the measurement contribution function $f_j$ for them, and estimate the result as

$$\langle I_j \rangle = \frac{1}{N} \sum_{i=1}^{N} f_j(\bar{X}_i)p(\bar{X}_i).$$  \hspace{1cm} (2)

$\langle I_j \rangle$ is estimator of the true value of $j$-th pixel color $I_j$. Because evaluating the samples is computationally expensive, high-quality images can take hours, or even days to compute. Furthermore, a class of difficult scenes exists, that take very long to render even with specialized advanced algorithms, and cannot be rendered with the simple ones at all. This is because it is difficult to find the important light paths in them (paths that contribute the most to the result). Examples are shown in Figure 3.

![Figure 3. Two examples of difficult scenes, where it is difficult to find the important light paths. In the left image, all lighting to the room comes through a narrow door slit. The right image contains lots of sharp illumination peaks – caustics – produced by small spots of light shining on the mirror balls. The images are courtesy of Veach [1998] and Hachisuka et al. [2008].](image)

To increase the efficiency of the process, it is advisable to use importance sampling, i.e. to construct samples with $pdf$ $p$ as similar to $f_j$ as possible, which decreases the error greatly. There is zero error when $p(\bar{X})$ is perfectly proportional to $f_j(\bar{X})$. This is however impossible to achieve and importance sampling according to only some of the $f_j$ terms is used instead. It is also possible to sample using multiple strategies ($pdf$s) corresponding to different terms of $f_j$ and then combine the results optimally via Multiple Importance Sampling (MIS) (Veach [1998]).

**Classification of methods**

In our survey we classify methods into three groups. Each group utilizes the importance for light path construction in different way.

**Incremental path construction from the camera**

The simplest approach to constructing the light paths is unidirectional local construction: all paths are constructed incrementally from one end. First vertex and direction where to
continue are chosen either on camera (path tracing) or light source (light tracing). Then, the intersection with the scene is found (the first inner vertex, see Figure 2). Now, new vertices are added incrementally using the same process, until the path is terminated. The paths are constructed via series of local decisions (choosing the direction where a ray is sent to find the next path vertex). The direction is usually chosen by importance-sampling local bidirectional reflectance distribution function (BRDF) $f_s$. While this approach works well in simple scenes with mostly uniform luminance, it fails to find the important light paths in difficult scenes, resulting in overall poor performance.

To improve the performance in such scenes, importance sampling of the incident illumination needs to be incorporated at the inner vertices during the path construction. Sampling according to the incident illumination would allow us to generate paths from a probability density very similar to $f_j$. However, to implement it, we would already need to know the illumination in the scene, which is what we are trying to compute in the first place.

Probably the first attempt to solve this dilemma was made by Lafortune and Willems [1995]. They use an ordinary path tracer, but store the information from already constructed and evaluated light paths in a 5D radiance tree (with 3 dimensions for position and 2 for direction). The tree represents approximation of the incident illumination for all points in the scene. This allows them to perform the incident illumination importance sampling. However, the chosen representation enables only coarse reproduction without the ability to capture high-frequency details. The method is also very slow and memory intensive, because large amount of samples has to be stored.

Additionally, there is a conceptual problem, that the data source for importance sampling is the same algorithm as the one for rendering. This means, that the method cannot efficiently render images containing light paths which are hard to find for this single algorithm.

The problem was addressed by many authors (e.g. Jensen [1995], Szirmay-Kalos et al. [1999]), who use the photon map as the source of data for importance sampling. Photon mapping is originally a biased technique for rendering developed by Jensen [1996]. It precomputes a set of light paths starting on light sources (photons), and stores them in a global structure called the photon map.

When using the photon map for importance sampling, nearby photons are splatted into a grid, and an approximate incident illumination representation is reconstructed from them. This improves the robustness by combining different algorithms for rendering and for obtaining importance sampling data source, but other problems, namely large memory footprint, inadequate capture of high frequencies, and large overhead of pdf creation and sampling, remain.

Hey and Purgathofer [2001] developed Hemispherical Particle Footprints, a method for better sampling of the high frequencies. It takes into the account the directional photon density, which causes the piecewise-constant representation of the pdf to be denser where the value of pdf is higher. This leads to superior sampling quality compared to previous approaches, but the memory footprint and sampling overhead is considerably increased.

**Direct lighting computation**

The direct lighting (the last segments of light paths connecting the scene geometry to lights) substantially contribute to the final result in most scenes. It is natural that specialized methods for its handling exist. A common approach to direct lighting calculation is to explicitly connect a light to the current path via a shadow ray. In the simplest form each path vertex is connected to all lights in the scene. This is however computationally expensive as the intersections with the scene geometry are included. Several methods have been developed that aim to decrease the number of used shadow rays without increasing an unpleasant noise in the image.

The main idea of Ward [1991] and Shirley and Wang [1991] is to devote expensive shadow

---

1 Starting from the camera is usually preferred because of its higher efficiency in most scenes.
rays only to those light sources which might potentially have a substantial contribution to the current path. Light source intensity decreases with a square of distance and is also dependent on the orientation of a light source. Thus both methods first estimate the potential contribution of light sources without taking the visibility into account and afterwards they decide where to use the shadow rays.

Ward suggested to use a shadow ray for every light source which has potentially high contribution. The contribution from the rest is computed by multiplication of potential contribution by rough estimate of visibility. This estimate is cheap in comparison with testing visibility by using shadow rays and the introduced error should be small since the estimate is used only for light sources with small potential contribution.

On the other hand, the method of Shirley and Wang is unbiased because shadow rays are used for computing contributions of all light sources. Their method constructs a probability distribution defined over all light sources in the scene which is proportional to light source potential contributions. This distribution is eventually used for sampling one or just a few shadow rays per shading point. To construct such a distribution light sources are divided into two groups. Light sources with substantial potential contribution to the shading point (bright group) and the rest (dim group). To construct the distribution the potential contribution of each light source in the bright group is estimated using one sample. Total potential contribution of all light sources in the dim group is estimated using just one sample whose value is multiplied by the number of dim light sources. After normalization we obtain the desired probability density.

Neither of the methods take the visibility into account for importance sampling. The disadvantage of the latter method is also dividing light sources into two groups with respect to the surface points of the scene geometry. Complicated algorithm involving space partitioning in combination with light source bounding volumes is used. Peter and Pietrek [1998] take the visibility partially into account and also simplify the classification of light sources into groups by using photon maps.

Metropolis light transport

A conceptually different approach was taken by Veach [1998], who applied the Metropolis sampling algorithm (Hastings [1970]), known from computational physics, to light transport simulation. The resulting Metropolis light transport algorithm (MLT) is an unbiased Markov chain Monte Carlo algorithm. It works by continuously mutating a sample (light path) by following these steps:

1. Propose new sample $\hat{X}_{\text{new}}$ as a mutation of current sample $\hat{X}_{\text{current}}$. Transition function $T(\hat{X}_{\text{new}} \rightarrow \hat{X}_{\text{current}})$ gives the probability density of the mutation.

2. Compute acceptance probability $p_{\text{acc}}$ as $\min\left\{ \frac{T(\hat{X}_{\text{new}} \rightarrow \hat{X}_{\text{current}})}{T(\hat{X}_{\text{current}} \rightarrow \hat{X}_{\text{new}})}, 1 \right\}$.

3. With probability $p_{\text{acc}}$ set $\hat{X}_{\text{current}} = \hat{X}_{\text{new}}$.

4. Go to 1.

When certain assumptions are met, this algorithm generates a chain of random samples, whose distribution becomes perfectly proportional to $f_j$ as number of iterations approaches infinity.

Designing optimal mutation technique in step 1 is crucial to MLT performance. Veach originally proposed a rather complicated set of mutation strategies (e.g. caustics and lens perturbations). Kelemen and Szirmay-Kalos [2001] simplified this by computing the mutations not in the space of light paths, but in the space of unit hypercube of random numbers used for generating the light paths. This formulation, called Primary Sample Space MLT, is simpler and more efficient. It has just two simple mutation strategies (small and large perturbations).

Kitaoka et al. [2009] improved the algorithm further by applying the Replica Exchange algorithm, originally used in computational physics and statistics. In this version, the single
Markov chain is replaced with multiple ones, called replicas, each one mutated using a different strategy. Additionally, there is a chance that two neighbouring chains swap their current samples (exchange of internal states). Samples from all replicas are combined using multiple importance sampling (Veach [1998]). In some very difficult scenes the algorithm performs much better than previous formulations.

Another important deficiency of MLT, the inability to mutate specular paths, was addressed by Jakob and Marschner [2012]. Specular paths cannot be mutated by the original algorithm, because it is not possible to change only one of their vertices at a time. Jakob and Marschner describe the set of all light paths with nonzero $f_j$ in a scene as a manifold embedded in higher-dimensional space, and devise a method for traversing this manifold, lifting this restriction. This allows for better exploration of light paths in scenes with difficult specular transport.

Unfortunately, even with all these improvements, the Metropolis light transport algorithm is still inefficient in simple scenes, has problems in certain difficult scenes, and remains notoriously hard to implement.

Conclusion and future work

The list of an importance sampling related literature stated in this survey is far from being comprehensive. We have made an effort to select methods so that we show various different approaches to importance sampling in realistic image synthesis. We also took into account the relevance to our own research.

According to our survey there is no universal rendering method which is able to effectively locate highly contributing light paths in the scene. Existing methods are not applicable to the wide range of scene types either because they simply do not work in difficult scenarios (e.g. building interiors containing highly glossy surfaces and illuminated by small windows) or they are too slow in simple scenes.

This is motivation for our own research in the area. According to our classification the methods which we aim to develop belong into the first group. We will construct the light paths incrementally from the camera, using pdf constructed with respect to both incident illumination and the BRDF. The initial information about the incident illumination will be pre-computed by photon mapping and then it will be adaptively corrected during rendering. Unlike existing methods our approach will be able to precisely target even very small bright areas, handle both direct and indirect illumination and will be designed to have small memory and computational time overhead.

Acknowledgments. This work was supported by the Grant Agency of Charles University, Prague (project number 580612).

References


