Bilateral blue noise sampling

Jiating Chen, Xiaoyin Ge, Li-Yi Wei, Bin Wang, Yusu Wang, Huamin Wang,
Yun Fei, Kang-Lai Qian, Jun-Hai Yong, Wenping Wang

HAL Id: hal-00920673

https://inria.hal.science/hal-00920673

Submitted on 19 Dec 2013

Bilateral Blue Noise Sampling

Jiating Chen[‡][*] Xiaoyin Ge[§][*] Li-Yi Wei[¶][‖] Bin Wang[‡][†] Yusu Wang[§] Huamin Wang[§] Yun Fei[‡]
Kang-Lai Qian[‡] Jun-Hai Yong[‡] Wenping Wang[¶]

[‡] Tsinghua University   [§] The Ohio State University   [¶] The University of Hong Kong   [‖] Microsoft Research

(a) *biological distribution* **(b)** *photon density estimation* **(c)** *point cloud sampling + reconstruction*
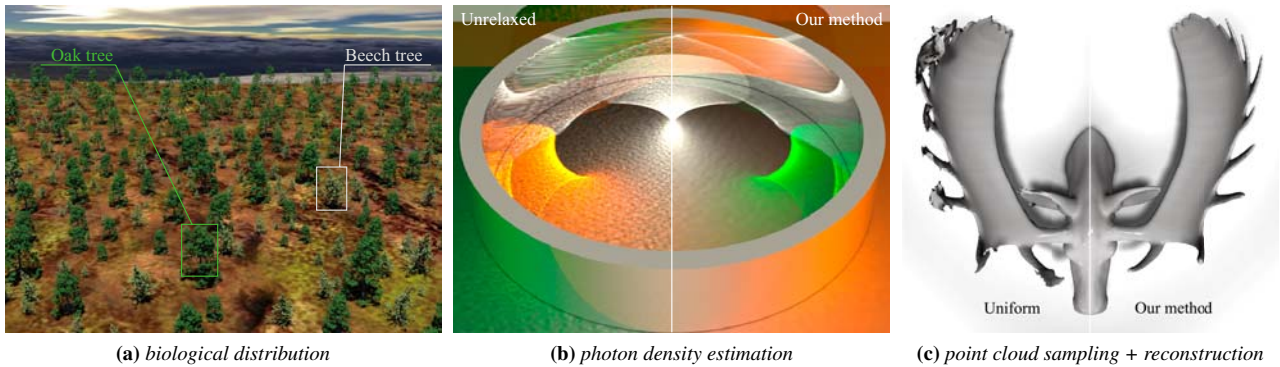
Figure 1: *Analogous to bilateral filtering [Tomasi and Manduchi 1998], our bilateral sampling method considers both spatial-domain and non-spatial-domain properties. Our method can generate distributions with sample attributes that are not direct functions of the underlying domains, such as more natural biological distribution with domain-independent tree type and size (a), less noisy photon density estimation with arbitrary flux and incoming direction (b), and more accurate (hidden) surface reconstruction from point cloud sampling (c).*

Abstract

Blue noise sampling is an important component in many graphics applications, but existing techniques consider mainly the spatial positions of samples, making them less effective when handling problems with non-spatial features. Examples include biological distribution in which plant spacing is influenced by non-positional factors such as tree type and size, photon mapping in which photon flux and direction are not a direct function of the attached surface, and point cloud sampling in which the underlying surface is unknown a priori. These scenarios can benefit from blue noise sample distributions, but cannot be adequately handled by prior art.

Inspired by bilateral filtering, we propose a bilateral blue noise sampling strategy. Our key idea is a general formulation to modulate the traditional sample distance measures, which are determined by sample position in spatial domain, with a similarity measure that considers arbitrary per sample attributes. This modulation leads to the notion of *bilateral* blue noise whose properties are influenced by not only the uniformity of the sample positions but also the similarity of the sample attributes. We describe how to incorporate our modulation into various sample analysis and synthesis methods, and demonstrate applications in object distribution, photon density estimation, and point cloud sub-sampling.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation—Antialiasing; I.3.5 [Computer Graphics]: Computational Geometry and Obje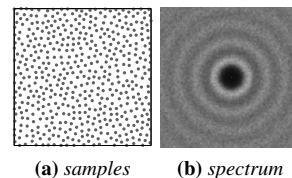ct Modeling—Curve, surface, solid, and object representations; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Color, shading, shadowing, and texture;

Keywords: bilateral measure, blue noise, stochastic sampling, point cloud, object distribution, photon mapping

1 Introduction

(a) *samples* **(b)** *spectrum*

Blue noise refers to sample sets that have random and yet uniform distributions. Due to the spatial uniformity and absence of spectral bias, blue noise has found use in a variety of graphics applications such as rendering [Cook 1986; Sun et al. 2013], halftoning [Ulichney 1987; Chang et al. 2009], stippling [Kopf et al. 2006; Balzer et al. 2009; Fattal 2011], animation [Schechter and Bridson 2012], modeling [Lipman et al. 2007; Bowers et al. 2010], and visualization [Li et al. 2010].

The majority of these methods produce samples whose properties are entirely dependent upon the underlying spatial domain. In particular, the position of each sample determines its spacing with other samples, such as stippling an image. However, many common scenarios do not obey this presumption, such as:

Biological distribution [Pommerening 2002; Illian et al. 2008; Pirk et al. 2012], in which plants often exhibit noise-like arrangements and the pair-wise spacing depends on not only plants' positions but also their own attributes that are not directly dependent on positions, such as heights, trunk diameters, and branch spreads.

Photon mapping [Jensen 1996], in which rendering quality depends on the uniformity of photons with similar attributes (such as flux and incoming direction) that are function of the entire scene instead of just photon spatial locations over the attached surface.

Point cloud sampling [Öztireli et al. 2010], in which the goal is

---

[*]Jiating Chen and Xiaoyin Ge are joint first authors.

[†]Bin Wang is the corresponding author.

to subsample a large collection of initial points so that the remaining points follow geometry features and yet remain uniform and random in smooth regions in order to accurately reconstruct the underlying surface, which is unknown a priori.

These scenarios can also benefit from blue noise sample distributions, e.g., more natural biological distributions, less noisy photon density estimation, or more accurate surface reconstruction. Unfortunately, they cannot be handled by prior sampling methods that are limited to spatial domains.

We propose a generalized form of blue noise whose samples can possess both spatial and non-spatial attributes and consequently whose distributions may not entirely depend upon the underlying domain. Our key idea is a general formulation to modulate traditional distance measures, which are determined by domain positions, with similarity of sample attributes that can be domain independent or unknown a priori. This allows our method to produce *bilateral* blue noise distributions whose quality depends on not only uniformity of sample positions (as in traditional blue noise quality measures [Lagae and Dutré 2008; Wei and Wang 2011]) but also similarity of sample attributes. In particular, unlike traditional blue noise which strives for spatial uniformity, the optimized degree of uniformity of our approach is proportional to the similarity of sample attributes. This is exemplified in Figure 2 and Figure 3, for which the relevant attributes are surface normal directions for point cloud sampling and flux colors for photon density estimation. We propose two forms of modulation, *augmentative* and *multiplicative*, with different properties and suitability for different scenarios. We propose how to analyze bilateral blue noise, and how to synthesize them to not only exhibit blue noise characteristics but also observe the underlying sample properties. We demonstrate potential applications of our approach in object distribution, photon density estimation, and point cloud sub-sampling (as summarized in Figure 1).
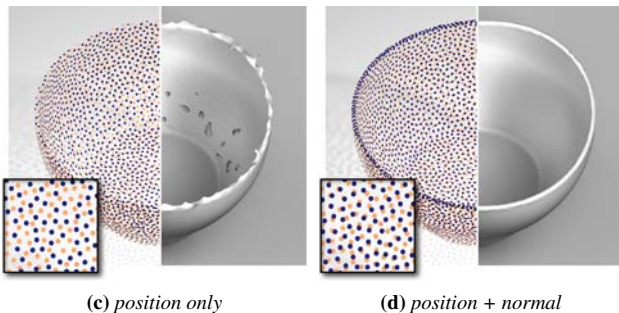

(c) *position only* **(d)** *position + normal*

Figure 2: *Bilateral blue noise with augmentative similarity measure based on normal direction. The goal is to reconstruct the unknown surface (right within each group) from the sub-sampled point cloud (left within each group). Sample point colors indicate whether the normal directions are pointing forward or backward (and thus the corresponding samples lie on the front or back surface). Traditional blue noise methods that consider only position information will uniformly sample the embedded 3D space but not the underlying 2D manifold (c) leading to poor surface reconstruction. Our method, by considering both positions and normal directions, will produce a uniform distribution in the 2D manifold (d) leading to better reconstruction.*

2 Previous Work

Blue noise sampling Coined by Ulichney [1987], blue noise refers to sample distributions that are random and uniform. This is a desirable combination, as its randomness avoids aliasing issues commonly seen in more structured sample patterns, while its uniformity reduces noise compared to other stochastic sampling methods. Consequently, blue noise has found widespread applications in graphics. Most of the prior blue noise methods derive sample
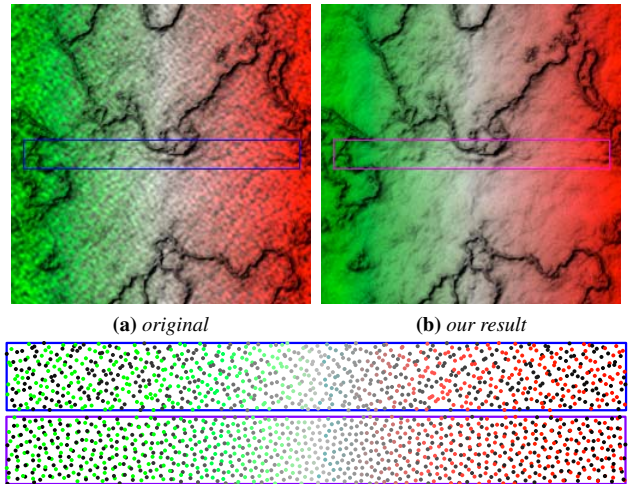

(a) *original* **(b)** *our result*

Figure 3: *Bilateral blue noise with multiplicative similarity measure based on photon flux. The input scene (a) is illuminated by two directional light sources, one gradually varying from green to gray to dark and another from dark to gray to red, with our bilateral kernel-based optimization result shown in (b). Notice the better quality of (b) versus (a) under the same number of samples. The bottom rows visualize the detailed photon distributions for each case: top for (a) and bottom for (b). The original input has clearly less uniform distribution, causing noisy output in (a). Our result demonstrates a continuously varying degree of bilateral uniformity, ranging from the sides where the two groups of photons have very different attributes and the middle where all photons have similar attributes.*

properties from their spatial positions. In other words, the distributions are determined by sample attributes which are functions of the domain positions, such as importance for adaptive sampling (see e.g., [Balzer et al. 2009; Li et al. 2010; Fattal 2011; de Goes et al. 2012]). Thus, these methods cannot directly consider non-spatial-domain properties, some of which might be domain independent or unknown a priori.

Multi-class sampling Multi-class blue noise sampling [Wei 2010; Schmaltz et al. 2012] provides an exceptional group of methods that consider domain independent attributes. However, these methods are designed for a discrete number of classes, enforcing blue noise properties for each individual class and their total union via an interaction matrix which encodes the spacing between pairs of class samples. However, it remains an open problem on how to extend the notion of discrete classes to handle continuous sample properties, in which an arbitrary collection of samples should maintain a degree of blue noise proportional to their mutual similarity in attributes. (Figure 3 is an example.) Furthermore, these methods assume otherwise known domain properties, such as density for adaptive sampling.

Our method can be considered as a generalization of multi-class sampling [Wei 2010; Schmaltz et al. 2012] in two major aspects: handling general non-spatial sample attributes beyond just sample classes, and enforcing blue noise properties for continuous instead of just discrete combinations of sample groups.

3 Core Ideas

Our core idea is to compute the bilateral differential $\boldsymbol{\xi}(s_i, s_j)$ between pairs of samples s_i and s_j as a modulation $\circ$ of the traditional positional differential $\mathbf{d}(\mathbf{p}_i, \mathbf{p}_j) = \mathbf{p}_i - \mathbf{p}_j$ [Wei and Wang 2011] and a (dis)similarity measure $\mathbf{f}(\mathbf{v}_i, \mathbf{v}_j)$ of non-spatial sample at-

tributes $\mathbf{v}_i$ and $\mathbf{v}_j$:

$$\boldsymbol{\xi}(s_i, s_j) = \mathbf{d}(\mathbf{p}_i, \mathbf{p}_j) \circ \mathbf{f}(\mathbf{v}_i, \mathbf{v}_j). \tag{1}$$

In a nutshell, the $\mathbf{d}$ part can be handled by existing methods (e.g., importance sampling), while the $\mathbf{f}$ part will be handled by our method together with $\mathbf{d}$. Effect-wise, two samples with more different attributes will have larger $\mathbf{f}$ value, allowing them to have closer positions $\mathbf{p}_i$ and $\mathbf{p}_j$ than the spacing criterion r (e.g., Poisson disk conflict radius) would allow under traditional settings. Thus, our method can maintain spatial uniformity for samples with similar attributes while placing more samples in areas with more attribute variation.

Throughout this paper we will use s to represent a sample, and $\mathbf{p}$ and $\mathbf{v}$ to represent a sample's spatial position and non-spatial attribute/feature, respectively. Depending on the circumstances, we indicate the association between $\mathbf{p}/\mathbf{v}$ and s in either the function form (e.g., $\mathbf{p}(s)$ is the position of s) or through subscripts (e.g., $\mathbf{p}_i$ is the position of s_i). We also use p to indicate a point in the domain Ω which might or might not be a sample. We indicate vector quantities in boldface (such as differential $\mathbf{d}$ and $\boldsymbol{\xi}$) and their corresponding scalar quantities (such as distance $\mathrm{d} = |\mathbf{d}|$ and $\xi = |\boldsymbol{\xi}|$) in non-boldface.

We will next present our options for $\circ$ and $\mathbf{f}$.

3.1 Spacing

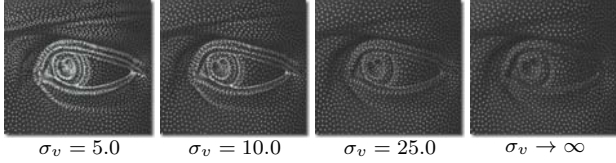We present two specific forms of Equation 1, based on augmentative and multiplicative modulations.

| $\sigma_v = 5.0$ | $\sigma_v = 10.0$ | $\sigma_v = 25.0$ | $\sigma_v \to \infty$ |

Figure 4: *The effect of σ_v for augmentative modulation in Equation 2. Under the same $\sigma_p = 1$, a smaller σ_v can cause more samples to be placed near surface features, such as geometry contours and corners. Setting $\sigma_v = \infty$ reduces our method to traditional blue noise sampling. In this example, we use point cloud sampling with normal direction as $\mathbf{v}$.*

Augmentative In this form, we append $\mathbf{v}$ as additional dimensions of the positional vector:

$$\boldsymbol{\xi}(s_i, s_j) = \left(\frac{\mathbf{d}(\mathbf{p}_i, \mathbf{p}_j)}{\sigma_p}, \frac{\mathbf{d}(\mathbf{v}_i, \mathbf{v}_j)}{\sigma_v} \right), \tag{2}$$

in which σ_p and σ_v are the relative weights between $\mathbf{p}$ and $\mathbf{v}$ differentials. They provide a balance between blue noise and feature preservation. When $\sigma_v \to \infty$, our method is reduced to conventional blue noise sampling, and when σ_v gets smaller, our method puts more emphasis on features than blue noise, as shown in Figure 4. In our experiments, we set $\sigma_p = 1$ and $\sigma_v \in [15, 25]$.

Multiplicative In this form, we divide $\mathbf{d}(\mathbf{p}_i, \mathbf{p}_j)$ with a scalar similarity measure $\tau(\mathbf{v}_i, \mathbf{v}_j) \in (0, 1]$ as follows:

$$\boldsymbol{\xi}(s_i, s_j) = \frac{1}{\tau(\mathbf{v}_i, \mathbf{v}_j)} \mathbf{d}(\mathbf{p}_i, \mathbf{p}_j). \tag{3}$$

By default we use Gaussian for τ, even though other math forms are possible as exemplified in later examples. For notational simplicity, we use $\tau(s_i, s_j)$ interchangeably with $\tau(\mathbf{v}_i, \mathbf{v}_j)$.

Discussion Our augmentative formulation in Equation 2 bears resemblance to bilateral filtering [Tomasi and Manduchi 1998]. In

particular, our $\mathbf{p}$ and $\mathbf{v}$ correspond to their domain position and range information, respectively. (This is also the reason why we name our method *bilateral*.) The main difference is that our method is about sampling while theirs is about filtering.

Our multiplicative formulation in Equation 3 bears resemblance to electrostatic halftoning [Schmaltz et al. 2010; Schmaltz et al. 2012]. In particular, their methods use electric charges to directly scale the desired sample spacing, whereas our method uses a scalar similarity τ for analogous purpose.

Our two formulations have similarities as well as differences. Given the same sample positions $\mathbf{p}_i$ and $\mathbf{p}_j$, they both have larger distance values $\xi(s_i, s_j)$ for less similar attributes $\mathbf{v}_i$ and $\mathbf{v}_j$. Their main difference, as evident, is the augmentative and multiplicative formulation. This provides different trade-offs. The former allows a metric while the latter allows more precise control: two samples with very dissimilar properties can get arbitrarily close. The exact choice can depend on both the algorithms and applications.

We focus on these two formulations due to their simplicity, relationships to prior methods, and adequacy for a variety of applications we have explored. Other forms of formulations are certainly possible but they would be beyond this paper; we leave these as potential future work.

3.2 Formulation

A common method to formulate the uniformity of a sample set S with N samples, $S = \{s_1, s_2, ..., s_N\}$, is to use a variational or energy function (see e.g., [Wei 2010; Schmaltz et al. 2010; Fattal 2011]) with the following general form:

$$E_u(s, S) = \frac{1}{N} \sum_{k=1}^{N} \kappa(s, s_k), \tag{4}$$

where s is a query sample and κ is a kernel function with compact support (exactly or at least approximately, such as Gaussian kernel, $\kappa(s, s_k) = e^{-\frac{\mathrm{d}^2(\mathbf{p}, \mathbf{p}_k)}{\sigma^2}}$). We include all domain information, such as adaptivity and anisotropy, into $\mathbf{d}$, therefore σ remains a global constant for kernel support size.

Following Equation 3, the key modification is to use our bilateral distance instead of the original measure:

$$E_u(s, S) = \frac{1}{N} \sum_{k=1}^{N} e^{-\frac{\xi^2(s, s_k)}{\sigma^2}}, \tag{5}$$

where ξ can come from either the augmentative (Equation 2) or multiplicative (Equation 3) forms. Intuitively, the influence of $\kappa(s, s_k)$ is modulated by the similarity of sample attributes; the less the similarity is, the narrower the kernels are and thus the weaker the influence is.

Besides spatial uniformity, another important criterion for blue noise is randomness, in particular the lack of any structures in the spatial distributions. This can often be best measured in the spectral domain, as discussed in Section 4.

4 Analysis

Here we describe how to extend the differential domain analysis (DDA) method in [Wei and Wang 2011] for bilateral blue noise distribution. Similar procedures can be applied to extending other spatial analysis methods such as the one based on the pair correlation function (PCF) [Öztireli and Gross 2012].

Given a set of N samples $\{s_k\}$ with positions $\{\mathbf{p}_k\}$, its Fourier power spectrum $P(\mathbf{f})$ [Ulichney 1987; Lagae and Dutré 2008] can
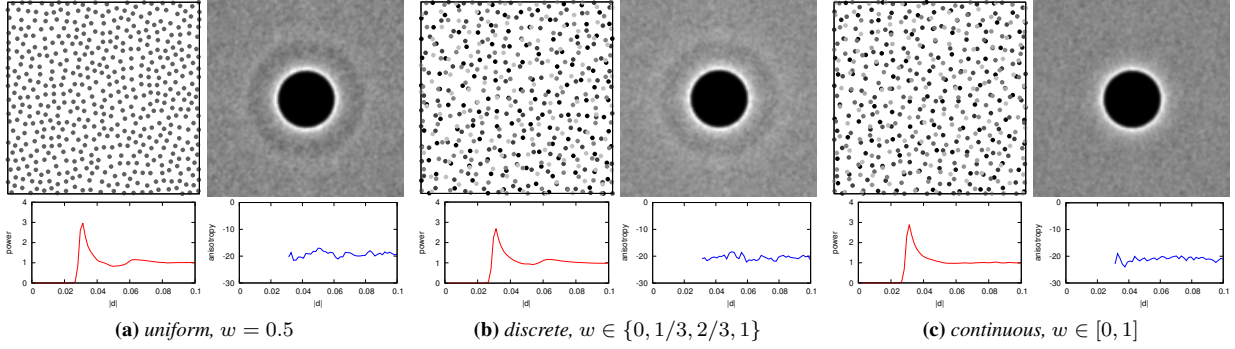
(a) uniform, w = 0.5 (b) discrete, w ∈ {0, 1/3, 2/3, 1} (c) continuous, w ∈ [0, 1]

Figure 5: *Differential domain analysis of bilateral blue noise sampling. Each sample has a weight attribute w (represented by its color). The similarity between any two samples s_i and s_j is defined as $\tau(s_i, s_j) = max(0, 1 - 2|w_i - w_j|)$. Both the synthesized results with discrete (b) and continuous (c) weights exhibit the similar spectral characteristics as that of the typical blue noise distribution (a). Within each group are the spatial samples colored with their weights, the DDA [Wei and Wang 2011] spectrum, and the radial mean and anisotropy. The results are produced by our bilateral dart throwing method.*

be represented as follows:

$$P(\mathbf{f}) = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \cos[2\pi \mathbf{f} \cdot (\mathbf{p}_i - \mathbf{p}_j)] \quad (6)$$

$$\approx \int_{\Omega_d} \cos(2\pi \mathbf{f} \cdot \mathbf{d}) \, p(\mathbf{d}) \, \delta \mathbf{d}, \quad (7)$$

where $\mathbf{f}$ is the frequency vector, and $p(\mathbf{d})$ is called the differential distribution function. It describes the probability distribution of the difference $(\mathbf{p}_i - \mathbf{p}_j)$ between every pair of samples s_i and s_j:

$$p(\mathbf{d}) = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \kappa(\mathbf{d}, (\mathbf{p}_i - \mathbf{p}_j)) \quad (8)$$

$$\approx \int_{\Omega_d} \kappa(\mathbf{d}, \mathbf{d}' = \mathbf{p}_i - \mathbf{p}_j) \delta \mathbf{d}', \quad (9)$$

where κ is a density estimation kernel, usually a Gaussian kernel (with center $\mathbf{d}'$ and standard deviation σ) for robust estimation, $\kappa(\mathbf{d}, \mathbf{d}') = e^{-\frac{\|\mathbf{d} - \mathbf{d}'\|^2}{\sigma^2}}$. We extend Equation 9 for bilateral blue noise by introducing the bilateral distance measure:

$$p(\mathbf{d}) \approx \int_{\Omega_d} \kappa\left(\mathbf{d}, \mathbf{d}' = \boldsymbol{\xi}(s_i, s_j)\right) \delta \mathbf{d}'. \quad (10)$$

See Figure 5 for example results.

For non-uniform sampling, we only need to perform proper transformation of $\mathbf{d}'$ depending on its two end points $\mathbf{p}_i$ and $\mathbf{p}_j$. The transformation may incur changes in length (for adaptive sampling) and/or direction (for anisotropic sampling) that are orthogonal to our $\boldsymbol{\xi}$; details can be found in [Wei and Wang 2011].

5 Synthesis

Various algorithms have been proposed to synthesize blue noise sample sets, such as dart throwing [Cook 1986] and particle/kernel-based optimization [Schmaltz et al. 2010; Fattal 2011]. Our method can be easily incorporated into these methods by replacing the traditional distance measure $\mathbf{d}$ by our bilateral one $\boldsymbol{\xi}$ as in Equation 2 or Equation 3. The net effect is that samples with less similarity will have more wiggle rooms to be spatially closer together under the same original spacing conditions.

5.1 Dart Throwing

The classical dart throwing algorithm [Cook 1986] generates samples one by one. Each time a trial sample s is drawn randomly from

the domain. This trial sample is accepted if it is at least a user-specified distance r away from all existing samples, and is rejected otherwise. This process ends when meeting some termination criteria, e.g., a target number of samples or a maximum number of trials.

Our bilateral dart throwing algorithm follows a similar process, but uses bilateral distance ξ instead of the original distance d for conflict check. Note that since we only change the distance measure ξ but not the conflict measure r, our method can be orthogonally combined with not only uniform (r is a constant) but also adaptive ($r(\mathbf{p})$ depends on spatial sample position) and anisotropic sampling ($r(\mathbf{p}_i, \mathbf{p}_j)$ depends on the relative distance and direction of two sample positions $\mathbf{p}_i$ and $\mathbf{p}_j$).

Since the sample properties might be unknown a priori and can be independent from the underlying domain, in general it is not possible to pre-compute an importance function to guide the generation of trial samples through an adaptive or anisotropic white noise [Li et al. 2010]. We therefore follow the importance sampling strategy for Monte Carlo methods: use known importance/guidance functions if available (e.g., spatial adaptivity or anisotropy), and resort to uniform white noise by default.

5.2 Kernel-based Optimization with Known ρ

Point relaxation such as Lloyds method [Lloyd 1982] and kernel formulation [Fattal 2011] has been used to optimize noise distributions, and here we describe how to combine the latter with our bilateral method. Given a sample set S, the energy $E_u(p, S)$ in Equation 4 can be used to approximate a given density function $\rho(p)$ at any point p in the domain Ω. The energy function that describes the error of this approximation is defined by

$$E(S) = \int_{\Omega} |E_u(p, S) - \rho(p)|^2 \delta p. \quad (11)$$

The goal is to optimize S with low $E(S)$ while maintaining blue noise properties. To apply our bilateral method for non-uniform sampling in an n-dimensional Euclidean space Ω, we factor the density ρ into the distance measure on a Riemannian manifold based on the Nash embedding theorems [Nash 1954], where the density can be treated as a constant. That is, the distance measure $\mathrm{d}(p_i, p_j)$ on the manifold is integrated along the geodesic from p_i to p_j:

$$\mathrm{d}(p_i, p_j) = \int_{p_i}^{p_j} \sqrt[n]{\rho(p)} \delta p. \quad (12)$$

A similar methodology can also be performed for anisotropic sampling [Li et al. 2010].

Having a constant density facilitates our derivations; in Appendix A, we derive that the energy function (Equation 11) with a constant target density function is domain independent,

$$E(S) = \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} E_{i,j} = \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} e^{-\frac{d^2(\mathbf{p}_i, \mathbf{p}_j)}{2\sigma^2}}, \qquad (13)$$

where $E_{i,j} = e^{-\frac{d^2(\mathbf{p}_i, \mathbf{p}_j)}{2\sigma^2}}$ can be considered as an energy between samples s_i and s_j, and the energy $E(S)$ is the sum of the energies between arbitrary two samples in the sample set. When extending for our bilateral formulation, we simply introduce the bilateral distance measure in each energy term $E_{i,j} = e^{-\frac{\xi^2(s_i, s_j)}{2\sigma^2}}$.

We use a gradient descent method to iteratively minimize the energy function $E(S)$, and produce a sample distribution with bilateral blue noise properties. The gradient of $E(S)$ with respect to s_k can be computed by

$$\frac{\partial E(S)}{\partial s_k} = \sum_{i=1, i \neq k}^{N} \frac{\partial E_{i,k}}{\partial s_k} = \sum_{i=1, i \neq k}^{N} \frac{\boldsymbol{\xi}(s_i, s_k)}{\sigma^2} e^{-\frac{\xi^2(s_i, s_k)}{2\sigma^2}}. \quad (14)$$

At each iteration, we move sample s_k locally towards the direction that reduces $E(S)$ fastest by a small step size Δt,

$$\mathbf{p}_k^{t+1} = \mathbf{p}_k^t - \Delta t \frac{\partial E(S)}{\partial s_k}. \qquad (15)$$

Too large step size Δt will cause unstable relaxation, while too small Δt will result in slow convergence. We set the step size $\Delta t = 0.45\sigma^2$ and found that the algorithm converges fast (usually with 10~20 iterations from an initial white noise over ρ).

5.3 Optimization with Unknown ρ

The method introduced above assumes a given target density ρ in a certain form. However, in some scenarios such as photon density estimation, the given sample set was randomly sampled from a target density function ρ, which is unknown to our algorithm. For such cases, using the distance measure in the gradient-descent solver (Equation 15) would introduce diffusion bias, thus breaking the intrinsic density variation. Even if the initial distributions are available for estimating ρ, it is often too noisy for direct use. Here, we describe an extension of our kernel-based optimization method (Section 5.2) to handle unknown target densities. Note that the goal here is to preserve the density underlying a given initial distribution instead of trying to estimate an unknown density from scratch.

To preserve density variation, our basic idea is a robust constraint based on the flow of samples caused by our gradient-descent solver. Specifically, during gradient-descent, the local average flow of samples at s_k, which describes the diffusion bias, can be estimated as

$$\bar{\mathrm{F}}(s_k) = \sum_{i=1}^{N} \varpi(s_k, s_i)(\mathbf{p}_i^t - \mathbf{p}_i^0) \qquad (16)$$

$$= \sum_{i=1}^{N} \varpi(s_k, s_i)\mathbf{p}_i^t - \sum_{i=1}^{N} \varpi(s_k, s_i)\mathbf{p}_i^0, \qquad (17)$$

where $\mathbf{p}_i^0$ is the original position of s_i, and $\varpi(s_k, s_i)$ is a weight function (e.g., Gaussian: $\varpi(s_k, s_i) = e^{-\frac{\xi^2(s_k, s_i)}{2\sigma^2}}$). The first and second terms in Equation 17 are respectively the weighted barycenters of the relaxed and original samples with respect to s_k. In

the uniform areas, these two barycenters tend to remain nearby regardless of relaxation. However in the non-uniform areas, the two barycenters may drift apart; we need to prevent this from happening in order to preserve features of the underlying sample distribution. Our constraint strategy takes advantage of the feedback from the existing flow of samples. The movement of s_k in each iteration considers both the uniformity and the reduction of diffusion bias:

$$\mathbf{p}_k^{t+1} = \mathbf{p}_k^t - \Delta t \left(\frac{\partial E(S)}{\partial s_k} + \alpha \bar{\mathrm{F}}(s_k) \right). \qquad (18)$$

Here α is a user-defined parameter for achieving a trade-off between uniformity and density preservation. Through our experiments, we found that $\alpha = 0.25/\sigma^2$ works well.

5.4 Performance

Our synthesis methods inherit the simplicity of the traditional blue noise sampling approaches, and are ripe for GPU implementation and acceleration. For our dart throwing algorithm, we extend the grid data structure in Wei [2008] for conflict check, and perform parallelization on GPU. Our dart throwing method can produce ~8M samples/sec for 2D synthesis. For the kernel-based optimization running on a GPU (3D synthesis followed by projection onto surfaces for photon mapping), our method can handle ~250K samples/sec using 10 iterations. All the performance data above were measured on a PC with an Intel Core i7 3770 CPU at 3.4GHz, 16GB RAM, and an NVIDIA GeForce GTX580 GPU.

6 Applications

We present applications of bilateral blue noise sampling in object distribution (Section 6.1), photon density estimation (Section 6.2), and point cloud sampling (Section 6.3). They share the common trait of involving both spatial and non-spatial sample properties. They also have different characteristics and thus are suitable for different ξ formulations and synthesis methods: multiplicative ξ + dart throwing for object distribution, multiplicative ξ + kernel-based optimization with unknown ρ for photon density estimation, and augmentative ξ + both dart throwing and kernel optimization for point cloud sampling. We choose the multiplicative ξ for objects and photons to have more direct control of sample spacing, and the augmentative ξ for point clouds since arbitrarily close samples would cause troubles for surface reconstruction.

6.1 Object Distribution

Real world objects are often distributed with particular characteristics. For example, biological individuals tend to be randomly located but remain uniformly separated apart [Illian et al. 2008], which can be characterized by blue noise distributions. Below, we demonstrate cell and tree distributions that can be well modeled by bilateral blue noise, potentially involving both *discrete* and *continuous* properties that can be independent of the underlying spatial domains. These are naturally generated by our dart throwing method.

Discrete properties Cells have been found to follow the blue noise distributions [Yellott 1983]. Figure 6a shows the distribution of rabbit retina amacrine cells of two different types: *on* and *off* [Diggle et al. 2006]. There is a pronounced inhibitory effect between cells of the same type to make their distributions more regular, meaning that no two on (or off) cells can be located arbitrarily close - the reported minimum distances are 21.4μm and 15.8μm for the on and off cells respectively; while the inhibitory effect is much less pronounced between cells of different types - the reported minimum distance is 5μm.

This 2-cell type distribution can be simulated by our method. We use a unified minimum spacing $r(s_i, s_j) = 21.4\mu$m between any

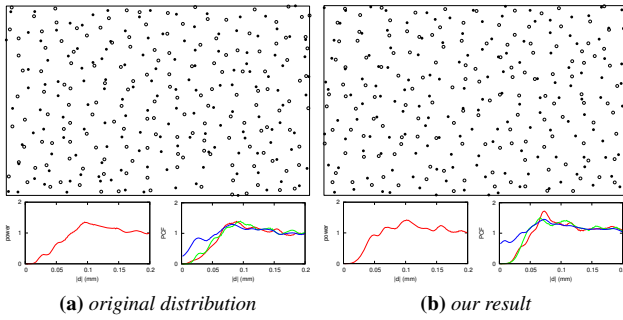**(a)** *original distribution*   **(b)** *our result*

Figure 6: *The distribution of amacrine cells with desired spacing depending on cell types. (a) shows the original cell data reported in Diggle et al. [2006] while (b) is our synthesized result. Within each group are the spatial cell distributions (with 152 on (○) and 142 off (●) cells), DDA radial mean [Wei and Wang 2011], and PCF curves [Öztireli and Gross 2012] (the red/green/blue curves are the PCFs of off/on/all cells).*

two cells s_i and s_j, and synthesize cell distributions based on the dart throwing criterion $\|\xi(s_i, s_j)\| > r(s_i, s_j)$. We modulate the similarity τ as follows to match the reported statistics above:

$$\tau(s_i, s_j) = \begin{cases} 1 & \text{if } s_i \text{ and } s_j \text{ are both of } on \text{ type,} \\ \frac{15.8}{21.4} & \text{if } s_i \text{ and } s_j \text{ are both of } off \text{ type,} \\ \frac{5}{21.4} & \text{if } s_i \text{ and } s_j \text{ are of different types.} \end{cases} \quad (19)$$

Our synthesized result, shown in Figure 6b, is able to match the original distribution (in Figure 6a) quite well, given that the latter is a measured instead of analytical distribution.

Note that even though this discrete setting can also be produced by the method in [Wei 2010], the paper has shown only applications with automatically computed sample spacing. We demonstrated the possibility of user-specified spacing. Next application is of continuous properties, which is beyond Wei [2010].
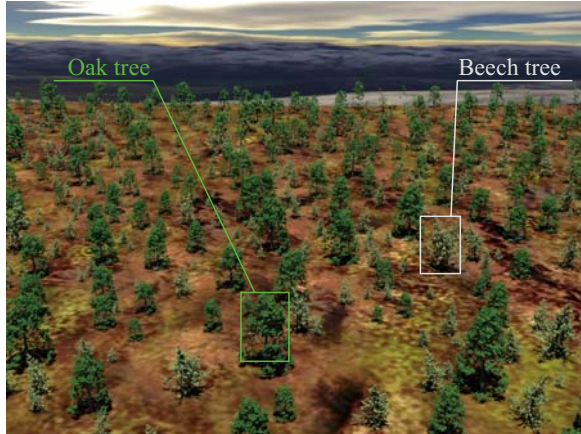

(a) *rendering*

(b) *distribution*

Figure 7: *Simulation of a forest of oak and beech trees with continuously varying crown sizes. (b): the synthesized distribution of two types of trees (● for oak and ● for beech, with symbol sizes indicating the crown sizes). (a): the image rendered using the approach [Bao et al. 2012]. Notice the segregation/mingling patterns for the oak/beech trees.*

Continuous properties Figure 7 provides a simulation of a forest consisting of two discrete types of trees, oak and beech, with continuously varying properties in heights and crown sizes [Pommerening 2002; Illian et al. 2008]. These trees form an interesting mixed pattern: the oaks, which need light, keep segregated characteristics; while the beeches, which can tolerate shade, have mingling characteristics.

To model this competition in terms of sunlight, we use the distance-dependent indices, which assume each tree had a circular area of influence expressed as a function of its crown size [Burkhart and Tomé 2012]. For simplicity we also assume the height of a tree is proportional to its crown size. We adopt a simple linear function to define the radius of influence area $R_s = \beta\gamma_s$, where β is a global control parameter proportional to tree spacing, for which we set as 2.0 in our Figure 7 result, and $\gamma_s \in [0.5m,\ 1.5m]$, a uniform random variable independent of the underlying domain, is the crown size of tree s. Thus, the unified minimum distance between samples/trees s_i and s_j is defined as $r(s_i, s_j) = R_{s_i} + R_{s_j}$. Inspired by Pirk et al. [2012], we design the similarity τ between any two trees to incorporate the light/shade properties of oak and beech:

$$\tau(s_i, s_j) = \begin{cases} 1 & \text{if } s_i\ s_j \text{ both } oak, \\ \frac{1}{\beta} & \text{if } s_i\ s_j \text{ both } beech, \\ \min\left(1, \max(\frac{1}{\beta}, \frac{\gamma_{s_{beech}}}{\gamma_{s_{oak}}})\right) & \text{if } s_i\ s_j \text{ mixed types.} \end{cases} \quad (20)$$

Intuitively, we leave sufficient spacing between oaks as they need sunlight, and just enough spacing between beechs so that their crowns will not overlap - the beech-beech condition above translates to $\|d(\mathbf{p}_i, \mathbf{p}_j)\| > \gamma_{s_i} + \gamma_{s_j}$. The oak-beech case is between those two; notice $\tau(s_{oak}, s_{beech})$ is capped above by $\tau(s_{oak}, s_{oak})$ and below by $\tau(s_{beech}, s_{beech})$, and its exact value depends on the relative crown sizes (and thus heights based on our height $\propto$ crown size assumption). When the beech is taller than the oak, we set $\tau(s_{oak}, s_{beech}) = 1$, the same as $\tau(s_{oak}, s_{oak})$, so that the oak will receive enough sunlight. Otherwise, the beech is shorter than the oak, and we set their minimal spacing to be proportional to the beech height to cap the maximal amount of shadow it can cast on the oak trunk, subject to the non-overlapping condition as in the beech-beech case. The final synthesized distribution is able to well exhibit the segregation and mingling properties of the two types of trees, as shown in Figure 7.

6.2 Photon Density Estimation

High-quality density estimation is critical for global illumination, especially in the widely used photon mapping approaches [Jensen 1996; Jensen and Christensen 1998]. Using a smooth kernel can help reduce rendering noise caused by the non-uniformity of photon distributions. However, it also introduces bias; as the number of samples increases, the estimated result converges to the convolution of the kernel with the true density function, rather than the function itself. Hence, there is a trade-off between sampling noise and estimation bias, which is determined by the kernel size. To minimize noise and bias, most existing approaches focus on improving the kernel estimators, such as adaptive bandwidth selection [Schregle 2003] and anisotropic kernel shape [Schjøth et al. 2008].

By observing that rendering noise is mainly caused by the non-uniformity of photon distribution, Spencer and Jones [2009; 2013a] provided state-of-art photon relaxation methods for improving caustics rendering. Due to the improved uniformity in photon distribution, more compact kernel sizes can be used for reducing noise and bias. Figure 8 illustrates a didactic case consisting of a variety of photon distributions including both sharp features and smooth areas. Their [2009] method relaxes all photons equally without considering their individual properties; this can improve global uni-
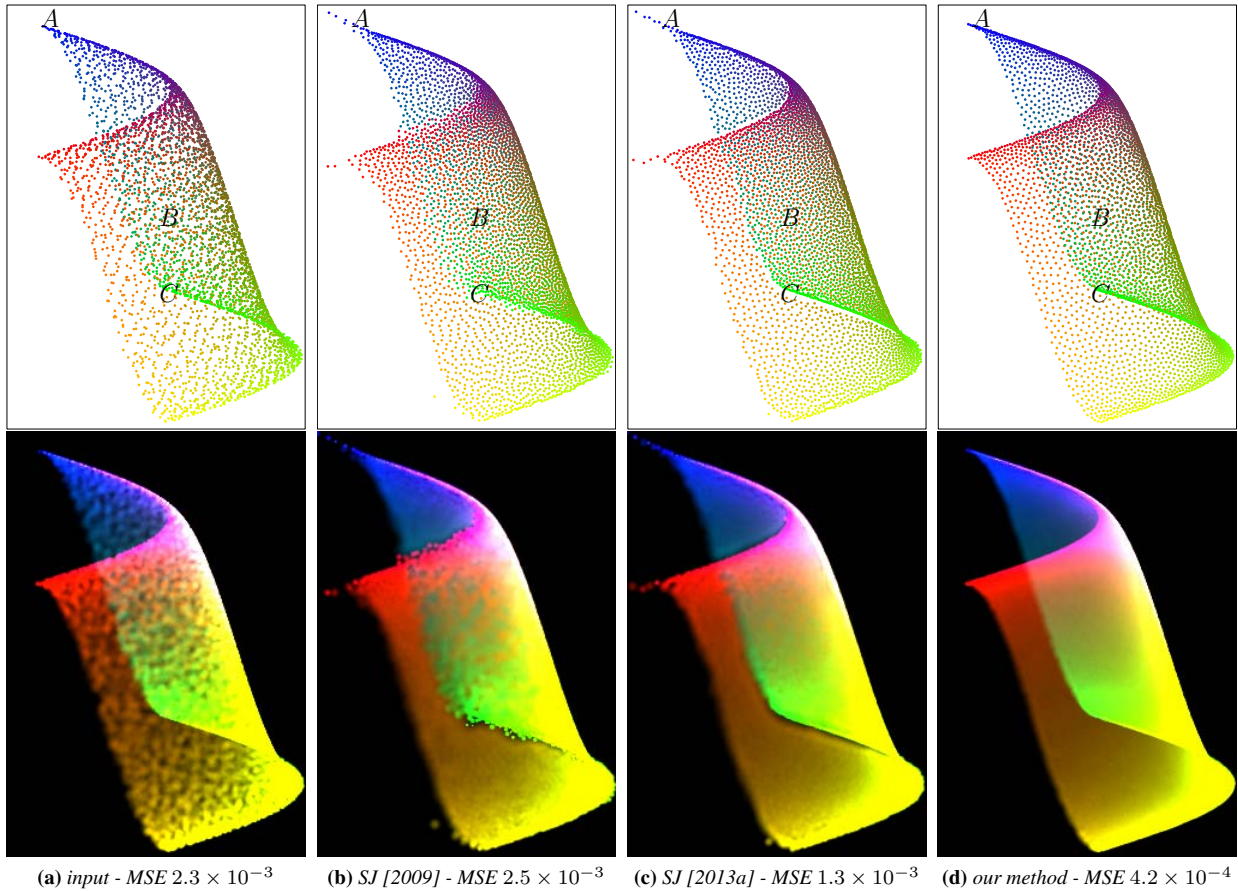
(a) *input - MSE* 2.3×10^{-3} **(b)** *SJ [2009] - MSE* 2.5×10^{-3} **(c)** *SJ [2013a] - MSE* 1.3×10^{-3} **(d)** *our method - MSE* 4.2×10^{-4}

Figure 8: *Comparison of photon relaxation by different methods via sample distribution (top), rendering result (bottom), and mean squared error (MSE) relative to the ground truth. (a): The input distribution is generated by randomly placing photons in a unit square light source and tracing their formation into a planar folded sheet, which consists of a variety of structures including mixed uniform region (B), thin interior feature (C), and sharp corner (A). Each photon is colored by its* (u, v) *coordinate within the light source:* $(r, g, b) = (u, v, \max(0, 1 - u - v))$. *(b): Spencer and Jones [2009] improves quality over the original input in (a) but may not preserve some features well, such as smeared region A, diluted region C, and less uniform region B for samples of similar colors. (c): Spencer and Jones [2013a] is able to maintain some features better than their prior method in [2009] such as regions B and C, but still has problem at sharp corners (see region A) and can also introduce discontinuity (see the gap in region C between green and orange samples). (d): Our method can robustly handle all these scenarios. Notice well preserved thin features in regions A and C and more uniform distributions of similarly colored photons in region B.*

formity compared to original photon distributions (Figure 8a versus Figure 8b) but might not be enough for some features (e.g., sharp regions A and C and smooth region B with mixed photon colors). Their [2013a] method handles features better by considering sample parameterization (e.g., (u, v) coordinates over an area light) analogous to multi-class sampling [Wei 2010; Schmaltz et al. 2012], but still has difficulty handing sharp corners (region A) since it only constrains relaxation in the direction orthogonal to, but not the one parallel to, feature anisotropy. It can also introduce discontinuity (see the gap in region C between green and orange samples) when switching the relaxation from high-dimensional parameter domain to low-dimensional spatial domain.

Our bilateral method directly considers photon properties during relaxation and thus can preserve both smooth regions (e.g., region B) and sharp features (e.g., regions A and C) better than Spencer and Jones [2009; 2013a] as shown in Figure 8. In particular, we use the bilateral kernel relaxation with unknown ρ method (Section 5.3) since we are given the initial photon set but not the underlying density. Figure 9 consists of scenes with complex effects of reflection and refraction. As shown, our bilateral formulation is more general and consistently produces more visually smooth as

well as more accurate rendering results than alternative methods. In addition to surface photon mapping, our method can also be easily applied for volumetric photon mapping [Jensen and Christensen 1998], as shown in Figure 10.

Implementation There are two main parameters we need to consider here. The first is the Gaussian kernel size σ_k for each sample s_k (used in Equation 14). In our implementation, we take $n = 40$ nearest neighbors into consideration using our bilateral distance measure. Assuming R_k is the radius of the disc center at s_k and enclosing exactly its n nearest neighbors, we set the kernel size $\sigma_k = R_k/3$. (The evaluations of local samples beyond $3\sigma_k$ can be ignored.) Since all photons will not move far from their original positions in our method, we use the same local sample set for each sample during relaxation in an approximate manner for efficiency considerations. The second parameter is the similarity $\tau(s_i, s_j)$ of photons s_i and s_j, which represents the similarity of photon attributes involved in radiance estimation. Usually, the photon flux Φ is involved, and $\tau(s_i, s_j) = \kappa_\Phi(\Phi_{s_i}, \Phi_{s_j})$. When the radiance estimation is performed on non-Lambertian surfaces, the incoming direction of photon ω is also involved, and
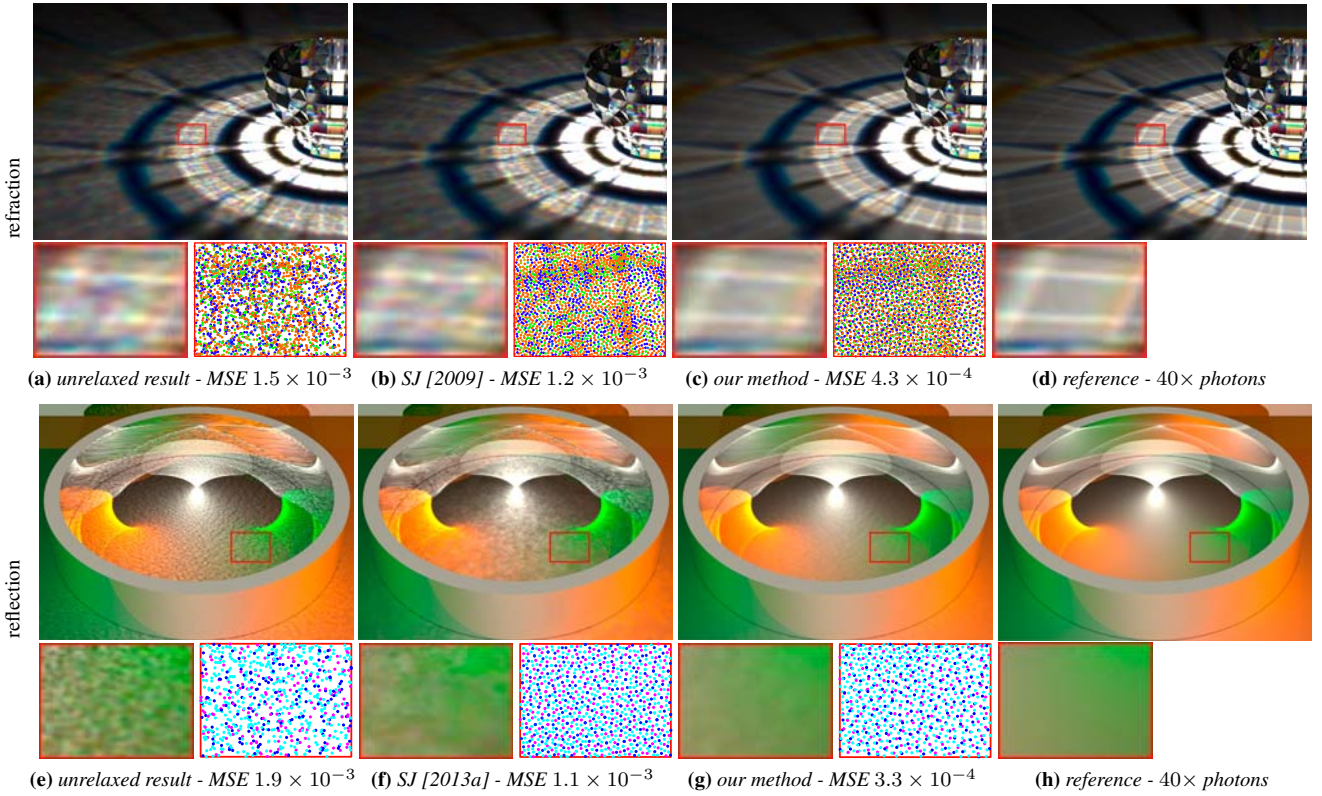
(a) *unrelaxed result - MSE* 1.5×10^{-3} (b) *SJ [2009] - MSE* 1.2×10^{-3} (c) *our method - MSE* 4.3×10^{-4} (d) *reference -* $40\times$ *photons*

(e) *unrelaxed result - MSE* 1.9×10^{-3} (f) *SJ [2013a] - MSE* 1.1×10^{-3} (g) *our method - MSE* 3.3×10^{-4} (h) *reference -* $40\times$ *photons*

Figure 9: *Surface photon density estimation. Top row: A complex pattern of dispersion caused by the candle light through a glass bowl. Photons exhibit different fluxes due to spectral sampling, and light of different wavelengths is refracted differently. The similarity of each sample pair depends on the photon flux. Bottom row: A scene containing a metal ring illuminated by three directional light sources with hybrid caustics on a glossy floor. The similarity relates to both photon flux and incoming direction. Our bilateral method produces more visually smooth as well as more accurate rendering results compared with the state-of-the-art photon relaxation method in Spencer and Jones [2009; 2013a]. We use [2009] for the refraction case since it is not clear how to apply or even extend [2013a] for different wavelengths. The reference images are rendered with* $40\times$ *photons without any relaxation.*
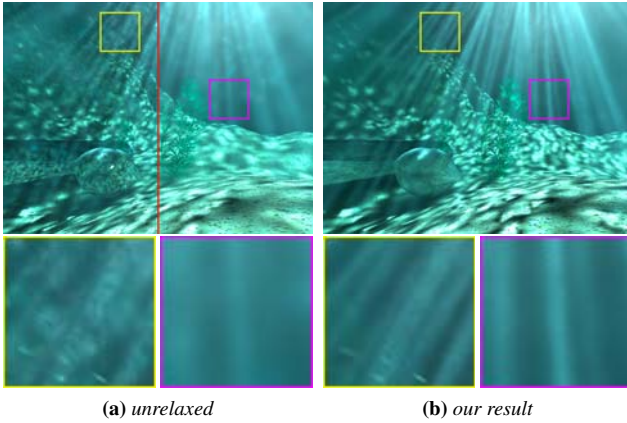

(a) *unrelaxed*   (b) *our result*

Figure 10: *Seabed rendered by volumetric photon mapping. (a): Images rendered without sample relaxation using a compact (left part) and a larger kernel size (right part). (b): Image rendered with our method using the same compact kernel size. Our method is able to smooth out the rendering noise while preserving the sharpness of volumetric caustics.*

$\tau(s_i, s_j) = \kappa_\Phi(\Phi_{s_i}, \Phi_{s_j})\kappa_\omega(\omega_{s_i}, \omega_{s_j})$. The functions $\kappa(.,.)$ can be Gaussian kernels, whose values reduce from 1 to 0 as the difference between these two photon attributes increases. The standard deviation σ_ω of Gaussian kernel κ_ω for incoming direction is set

to be 0.3 ($\sigma_\omega \in [0.15, 0.45]$), while σ_Φ for κ_Φ is set to be 0.15 ($\sigma_\Phi \in [0.05, 0.3]$). Here we normalized the photon flux Φ by the area of radiance estimation for the computation of similarity.

6.3 Point Cloud Sub-sampling

Geometry sampling is important for a variety of applications and can benefit from sample distributions that preserve both features and blue noise properties. In particular, allocating samples at features area such as tips and creases are critical in preserving the original geometry proprieties[Lévy and Liu 2010], while distributing samples with blue noise properties can avoid undesirable aliasing or bias in computations involving geometry samples [Turk 1992; Bolander and Saito 1998].

These prior methods mainly deal with known surfaces. Scenarios involving unknown surfaces are less explored but can have important applications. Here, we use point cloud sub-sampling [Öztireli et al. 2010] as a specific example. Given an input point cloud set that is often over dense and redundant, the goal is to select a subset of samples while maintaining quality of the reconstructed surface. The main challenge is to place enough samples around features while having blue noise distributions at smooth regions without explicitly knowing the underlying surface.

Similar to [Öztireli et al. 2010], we assume the input consists of a point cloud with per vertex position and normal direction information, the later one can be estimated through various methods as discussed in [Öztireli et al. 2010; Huang et al. 2013]. Our method can use geometry measures other than surface normal **n** (which tends to
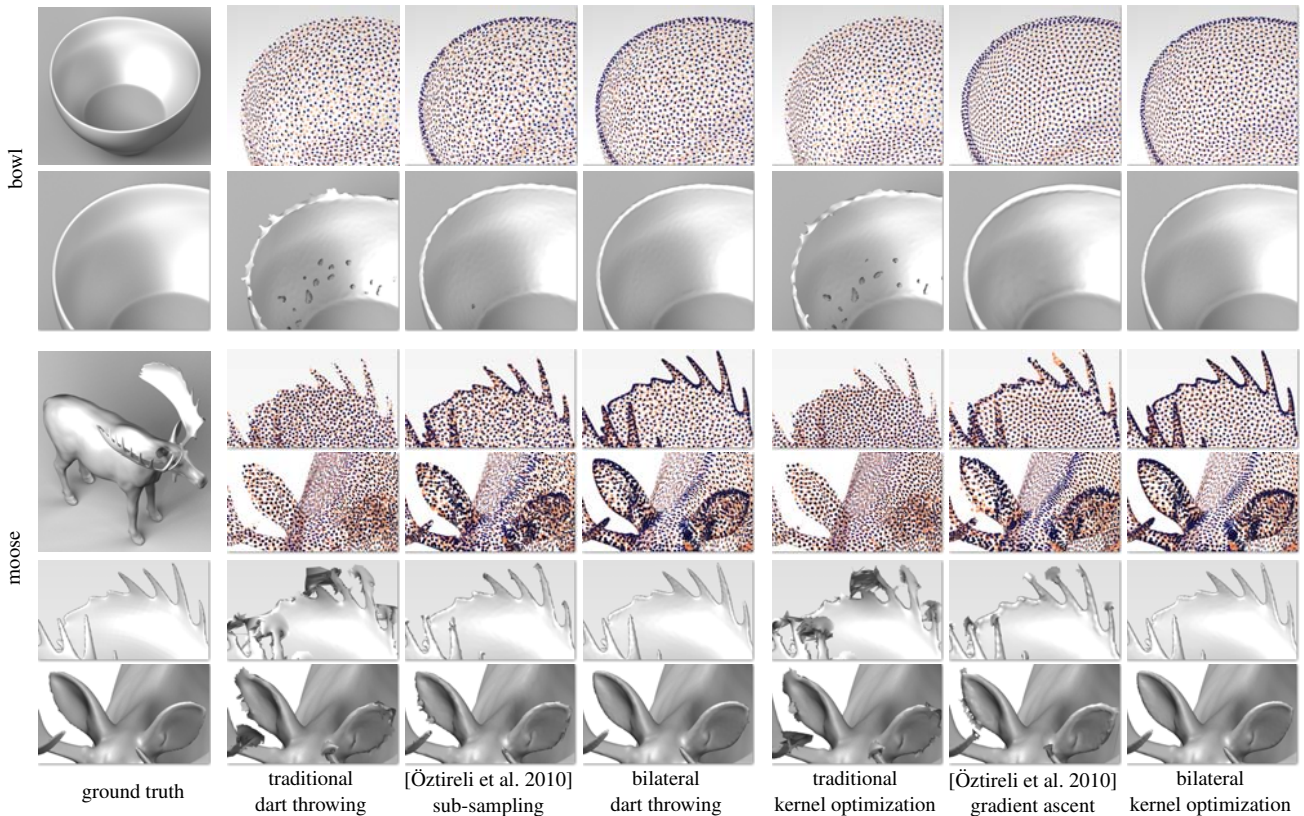
Figure 11: *Point cloud resampling results. Here, we compare traditional dart-throwing/kernel-optimization, [Öztireli et al. 2010] sub-sampling/gradient-ascent, and our bilateral dart-throwing/kernel-optimization. Each kernel-optimization/gradient-ascent result is produced from the corresponding dart-throwing/sub-sampling result. Within each bowl/moose case are the point clouds (top row) and reconstructed surfaces (bottom row), except for the first column which shows the original models in full (top) or in the demonstrated subset (bottom). As shown, bilateral sampling methods achieve a more uniform sample distributions in smooth areas and more accurate feature preservation in sharp areas, translating to better surface reconstruction. For the bowl/moose case, the input point cloud has $\sim 400K$/$800K$ samples and each output result has $\sim 6K$/$30K$ samples. All optimization results are generated via 15 iterations.*

be noisy and unstable) as the feature $\mathbf{v}$, but we will focus on $\mathbf{n}$ for easy presentation and comparison with [Öztireli et al. 2010]. Both [Öztireli et al. 2010] and our method involve a 2-step process. The initial selection step: sub-sampling in [Öztireli et al. 2010] and bilateral dart-throwing in our method; and the subsequent refinement step: gradient-ascent in [Öztireli et al. 2010] and bilateral kernel-optimization in our method.

Without knowing the underlying surface, traditional blue noise methods, which consider only sample positions, only can promise a uniform sample distribution in the 3D Euclidean space. Thus, they might incorrectly handle samples that have close spatial proximity but differ sufficiently in normal directions, such as samples lying on the opposite sides of a thin feature (as shown in Figure 2 and 11). The spectrum method in [Öztireli et al. 2010] considers both position and normal and thus improves the result quality. Our method, using $\mathbf{n}$ as $\mathbf{v}$ in Equation 2, is much simpler and produces better results. As shown in Figures 11 and 12, our methods can produce more effective distributions, in terms of both feature preservation and blue noise properties, than the corresponding methods in [Öztireli et al. 2010], resulting in better surface reconstruction quality. Notice that we need both blue noise uniformity in smooth regions and feature preservation in sharp regions; the holes on the bowl are caused by lacking of the former while the jagged bow lip, moose horn, and ear are caused by lacking of the latter.

Analysis The sub-sampling method in Öztireli et al. [2010] accepts a sample if it brings enough improvement to the current sub-

set. Thus, it behaves similarly to the soft disk sampling algorithm in Wei [2010] and tends to produce less uniform distributions than bilateral hard disk sampling. Our bilateral kernel-optimization can outperform gradient-ascent in Öztireli et al. [2010] mainly due to our robust diffusion constraint for feature preservation.

Implementation Since the underlying surface is unknown, we use Euclidean instead of geodesic distance for $\mathbf{d}(\mathbf{p}_i, \mathbf{p}_j)$ in Equation 2. This is usually adequate with sufficiently dense input samples [Lipman et al. 2007; Huang et al. 2013; Öztireli et al. 2010]. We restrict all samples to lie on the tangent planes of the points in the input point cloud. Thus, after moving each sample (via Equation 18), we project it back to the tangent plane of the nearest point in the input point cloud. We set σ_v to be 25 for Figure 1c, 2 and 11. Note we rescaled all input point cloud into a unit bounding box.

7 Limitations and Future Work

Since our bilateral distance measure allows closer sample pairs with less similarity, our dart throwing method can no longer take advantage of the gap processing techniques for accelerating maximal sampling [Ebeida et al. 2011; Yan and Wonka 2013]. A potential future direction is to explore data structures suitable for tracking such gaps for maximal sampling.

Our bilateral distance measure is designed for sample pairs. This is sufficient for dart throwing and kernel-based optimization, but not methods such as Lloyd relaxation which also need to compute the
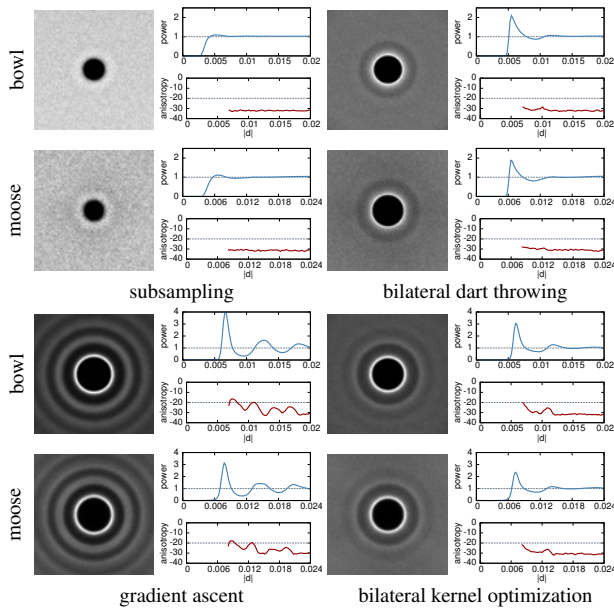
Figure 12: *Differential domain analysis for surface sampling. Here we show the DDA [Wei and Wang 2011] spectrum, radial mean, and radial anisotropy for the point sets in Figure 11. Compared to the methods in Öztireli et al. [2010] (left), our methods (right) have better blue noise properties, such as more uniform distributions in dart throwing versus sub-sampling (upper group) and the lower amount of anisotropy in general.*

distance between a sample and an arbitrary non-sample point in the domain. Our formulation can be directly applied if all attributes are derived from the domain so that points and samples share the same set of attributes. Otherwise, alternative schemes have to be considered and we leave them as potential future work.

Currently we empirically determine the parameters σ_p and σ_v in Equation 2, and would like to pursue more rigorous analytical methods. In addition, our current results are produced with uniform σ_p and σ_v, and we believe varying them spatially may benefit certain applications.

The current similarity measure τ in Equation 3 is a scalar. We would like to explore anisotropic τ (e.g., tensor) that would allow not only length but also direction change for positional difference.

To clearly demonstrate sample distributions, we use a relatively sparse spacing for the forest example in Figure 7. It is straightforward to tune our similarity function for producing denser forests or considering alternative biological attributes. As shown in [Zhou et al. 2012] biological phenomena can be characterized by other noise distributions such as pink noise for clustered bushes. Our bilateral formulation can be easily combined with the approach in [Zhou et al. 2012] for general (non-blue) *bilateral noises*.

Regarding photon mapping, our current implementation does not re-sample the fluxes and incoming directions of photons while optimizing their spatial positions. We have found this adequate since our method constrains the movements of photons around their original positions, but better accuracy could be obtained through direct high dimensional sampling as described above. Analogous to Spencer and Jones [2013b], we would also like to extend our bilateral method to progressive photon mapping [Hachisuka et al. 2008] for further quality and performance enhancement.

We have applied our method to 2D surfaces and 3D volumes; additional applications, such as nonlinear filtering as well as stippling videos and dynamic 3D objects, are described in [Ge et al. 2013]. We would also like to explore synthesis in higher dimensional spaces, such as treating light-paths instead of just photons as samples.

References

BALZER, M., SCHLÖMER, T., AND DEUSSEN, O. 2009. Capacity-constrained point distributions: A variant of Lloyd's method. *ACM Trans. Graph. 28*, 3, 86:1–86:8.

BAO, G., LI, H., ZHANG, X., AND DONG, W. 2012. Large-scale forest rendering: Real-time, realistic, and progressive. *Comput. Graph. 36*, 3, 140–151.

BOLANDER, J., AND SAITO, S. 1998. Fracture analyses using spring networks with random geometry. *Engineering Fracture Mechanics 61*, 5, 569–591.

BOWERS, J., WANG, R., WEI, L.-Y., AND MALETZ, D. 2010. Parallel Poisson disk sampling with spectrum analysis on surfaces. *ACM Trans. Graph. 29*, 6, 166:1–166:10.

BURKHART, H., AND TOMÉ, M. 2012. Indices of individual-tree competition. In *Modeling Forest Trees and Stands*. Springer Netherlands, 201–232.

CHANG, J., ALAIN, B., AND OSTROMOUKHOV, V. 2009. Structure-aware error diffusion. *ACM Trans. Graph. 28*, 5, 162:1–162:8.

COOK, R. L. 1986. Stochastic sampling in computer graphics. *ACM Trans. Graph. 5*, 1, 51–72.

DE GOES, F., BREEDEN, K., OSTROMOUKHOV, V., AND DESBRUN, M. 2012. Blue noise through optimal transport. *ACM Trans. Graph. 31*, 6, 171:1–171:11.

DIGGLE, P., EGLEN, S., AND TROY, J. 2006. Modelling the bivariate spatial distribution of amacrine cells. In *Case Studies in Spatial Point Process Modeling*, Lecture Notes in Statistics 185. Springer New York, 215–233.

EBEIDA, M. S., DAVIDSON, A. A., PATNEY, A., KNUPP, P. M., MITCHELL, S. A., AND OWENS, J. D. 2011. Efficient maximal Poisson-disk sampling. *ACM Trans. Graph. 30*, 4, 49:1–49:12.

FATTAL, R. 2011. Blue-noise point sampling using kernel density model. *ACM Trans. Graph. 30*, 4, 48:1–48:12.

GE, X., WEI, L.-Y., WANG, Y., AND WANG, H. 2013. Bilateral blue noise sampling: Additional algorithms and applications. Tech. Rep. OSU-CISRC-8/13-TR17, The Ohio State University - Department of Computer Science and Engineering. ftp://ftp.cse.ohio-state.edu/pub/tech-report/2013/TR17.pdf.

HACHISUKA, T., OGAKI, S., AND JENSEN, H. W. 2008. Progressive photon mapping. *ACM Trans. Graph. 27*, 5, 130:1–130:8.

HUANG, H., WU, S., GONG, M., COHEN-OR, D., ASCHER, U., AND ZHANG, H. R. 2013. Edge-aware point set resampling. *ACM Trans. Graph. 32*, 1, 9:1–9:12.

ILLIAN, J., PENTTINEN, A., STOYAN, H., AND STOYAN, D. 2008. *Statistical Analysis and Modelling of Spatial Point Patterns*. John Wiley & Sons, Ltd.

JENSEN, H. W., AND CHRISTENSEN, P. H. 1998. Efficient simu-

lation of light transport in scenes with participating media using photon maps. In *SIGGRAPH '98*, 311–320.

JENSEN, H. W. 1996. Global illumination using photon maps. In *Proceedings of the Eurographics Workshop on Rendering Techniques '96*, 21–30.

KOPF, J., COHEN-OR, D., DEUSSEN, O., AND LISCHINSKI, D. 2006. Recursive Wang tiles for real-time blue noise. *ACM Trans. Graph. 25*, 3, 509–518.

LAGAE, A., AND DUTRÉ, P. 2008. A comparison of methods for generating Poisson disk distributions. *Comput. Graph. Forum 27*, 1, 114–129.

LÉVY, B., AND LIU, Y. 2010. L_p centroidal Voronoi tessellation and its applications. *ACM Trans. Graph. 29*, 4, 119:1–119:11.

LI, H., WEI, L.-Y., SANDER, P. V., AND FU, C.-W. 2010. Anisotropic blue noise sampling. *ACM Trans. Graph. 29*, 6, 167:1–167:12.

LIPMAN, Y., COHEN-OR, D., LEVIN, D., AND TAL-EZER, H. 2007. Parameterization-free projection for geometry reconstruction. *ACM Trans. Graph. 26*, 3, 22:1–22:6.

LLOYD, S. 1982. Least squares quantization in PCM. *IEEE Transactions on Information Theory 28*, 2, 129–137.

NASH, J. 1954. C^1–isometric imbeddings. *Annals of Mathematics 60*, 3, 383–396.

ÖZTIRELI, A. C., AND GROSS, M. 2012. Analysis and synthesis of point distributions based on pair correlation. *ACM Trans. Graph. 31*, 6, 170:1–170:10.

ÖZTIRELI, A. C., ALEXA, M., AND GROSS, M. 2010. Spectral sampling of manifolds. *ACM Trans. Graph. 29*, 6, 168:1–168:8.

PIRK, S., STAVA, O., KRATT, J., SAID, M. A. M., NEUBERT, B., MĚCH, R., BENES, B., AND DEUSSEN, O. 2012. Plastic trees: Interactive self-adapting botanical tree models. *ACM Trans. Graph. 31*, 4, 50:1–50:10.

POMMERENING, A. 2002. Approaches to quantifying forest structures. *Forestry 75*, 3, 305–324.

SCHECHTER, H., AND BRIDSON, R. 2012. Ghost SPH for animating water. *ACM Trans. Graph. 31*, 4, 61:1–61:8.

SCHJØTH, L., SPORRING, J., AND OLSEN, O. F. 2008. Diffusion based photon mapping. *Comput. Graph. Forum 27*, 8, 2114–2127.

SCHMALTZ, C., GWOSDEK, P., BRUHN, A., AND WEICKERT, J. 2010. Electrostatic halftoning. *Comput. Graph. Forum 29*, 8, 2313–2327.

SCHMALTZ, C., GWOSDEK, P., AND WEICKERT, J. 2012. Multiclass anisotropic electrostatic halftoning. *Comput. Graph. Forum 31*, 6, 1924–1935.

SCHREGLE, R. 2003. Bias compensation for photon maps. *Comput. Graph. Forum 22*, 4, 729–742.

SPENCER, B., AND JONES, M. W. 2009. Into the blue: Better caustics through photon relaxation. *Comput. Graph. Forum 28*, 2, 319–328.

SPENCER, B., AND JONES, M. W. 2013. Photon parameterisation for robust relaxation constraints. *Comput. Graph. Forum 32*, 2, to appear.

SPENCER, B., AND JONES, M. W. 2013. Progressive photon relaxation. *ACM Trans. Graph. 32*, 1, 7:1–7:11.

SUN, X., ZHOU, K., GUO, J., XIE, G., PAN, J., WANG, W., AND GUO, B. 2013. Line segment sampling with blue-noise properties. *ACM Trans. Graph. 32*, 4, 127:1–127:14.

TOMASI, C., AND MANDUCHI, R. 1998. Bilateral filtering for gray and color images. In *ICCV '98*, 839–839.

TURK, G. 1992. Re-tiling polygonal surfaces. In *SIGGRAPH '92*, 55–64.

ULICHNEY, R. A. 1987. *Digital Halftoning*. MIT Press, Cambridge, MA.

WEI, L.-Y., AND WANG, R. 2011. Differential domain analysis for non-uniform sampling. *ACM Trans. Graph. 30*, 4, 50:1–50:10.

WEI, L.-Y. 2008. Parallel Poisson disk sampling. *ACM Trans. Graph. 27*, 3, 20:1–20:9.

WEI, L.-Y. 2010. Multi-class blue noise sampling. *ACM Trans. Graph. 29*, 4, 79:1–79:8.

YAN, D.-M., AND WONKA, P. 2013. Gap processing for adaptive maximal Poisson-disk sampling. *ACM Trans. Graph.*, to appear.

YELLOTT, J. I. J. 1983. Spectral consequences of photoreceptor sampling in the rhesus retina. *Science 221*, 382–385.

ZHOU, Y., HUANG, H., WEI, L.-Y., AND WANG, R. 2012. Point sampling with general noise spectrum. *ACM Trans. Graph. 31*, 4, 76:1–76:11.

A Kernel-based Optimization

Under a constant $\rho(p) \equiv C$, the energy function in Equation 11 is minimized by finding an optimal (uniform) sample set S^*,

$$S^* = \arg\min_S E(S) \tag{21}$$

$$= \arg\min_S C^2 \int_\Omega \delta p - 2C \int_\Omega (\sum_{k=1}^N \kappa(p, s_k))\delta p \tag{22}$$

$$+ \int_\Omega (\sum_{k=1}^N \kappa(p, s_k))^2 \delta p. \tag{23}$$

Suppose the domain Ω has no boundary, we can see that the first two terms are constant, and only the third term depends on the sample set. Thus, the above equation can be simplified as

$$S^* = \arg\min_S \int_\Omega (\sum_{k=1}^N \kappa(p, s_k))^2 \delta p \tag{24}$$

$$= \arg\min_S \sum_{i=1}^N \sum_{j=1}^N e^{-\frac{\|\mathbf{P}_i - \mathbf{P}_j\|^2}{2\sigma^2}} \int_\Omega e^{-\frac{2}{\sigma^2}\|\mathbf{P}(p) - \frac{\mathbf{P}_i + \mathbf{P}_j}{2}\|^2} \delta p \tag{25}$$

$$= \arg\min_S \sum_{i=1}^N \sum_{j=1, j\neq i}^N e^{-\frac{\|\mathbf{P}_i - \mathbf{P}_j\|^2}{2\sigma^2}}. \tag{26}$$

It shows that the energy function is domain independent,

$$E(S) = C' + \sum_{i=1}^N \sum_{j=1, j\neq i}^N e^{-\frac{\mathrm{d}^2(\mathbf{P}_i, \mathbf{P}_j)}{2\sigma^2}}, \tag{27}$$

where $E_{i,j} = e^{-\frac{\mathrm{d}^2(\mathbf{P}_i, \mathbf{P}_j)}{2\sigma^2}}$ can be considered as an energy between samples s_i and s_j, and the energy $E(S)$ is the sum of the energies between arbitrary two samples in the sample set (let the constant $C' = 0$).