Least Squares Vertex Baking

L. Kavan\textsuperscript{1}, A. W. Bargteil\textsuperscript{2} and P.-P. Sloan\textsuperscript{1}

\textsuperscript{1}Disney Interactive Studios
\textsuperscript{2}University of Utah

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{image.png}
\caption{Ambient occlusion in the Sibenik cathedral model represented at vertices using (a) point sampling, (b) averaging triangle samples and (c) our method. Compare to (d) ground truth per-pixel rendering.}
\end{figure}

Abstract

We investigate the representation of signals defined on triangle meshes using linearly interpolated vertex attributes. Compared to texture mapping, storing data only at vertices yields significantly lower memory overhead and less expensive runtime reconstruction. However, standard approaches to determine vertex values such as point sampling or averaging triangle samples lead to suboptimal approximations. We discuss how an optimal solution can be efficiently calculated using continuous least-squares. In addition, we propose a regularization term that allows us to minimize gradient discontinuities and mach banding artifacts while staying close to the optimum. Our method has been integrated in a game production lighting tool and we present examples of representing signals such as ambient occlusion and precomputed radiance transfer in real game scenes, where vertex baking was used to free up resources for other game components.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Color, shading, shadowing, and texture

1. Introduction

Representation of signals on triangle meshes is a classic problem in computer graphics. The standard solution relies on surface parametrization and texture mapping. While this is an efficient representation for high-frequency signals, the associated memory footprints are often one of the limiting factors in game development. Especially for low-frequency signals it is common to use vertex baking, i.e., representation by linearly interpolated vertex attributes [Gou71]. This is much less expensive, especially for scalar signals such as ambient occlusion [ZIK98, Lan02]—with textures, the uv-coordinates require two scalars per vertex, in addition to the image data. Furthermore, no parametrization or runtime texture memory access is needed. The significantly reduced memory footprints and lower shader complexity are important not only in low-end gaming platforms and mobile devices, but also in high-end games that often combine texture mapping and vertex baking [LH09, Wan11].

The question addressed in this paper is how to determine the vertex values to optimally approximate a given signal $f$ defined on a triangle mesh. The most straightforward solution is to point sample $f$ at vertex locations. While under certain
conditions this is sufficient for exact reconstruction [GW02],
in practice we almost always must accept some inaccuracies. Error is introduced because the vertex density is often irregular, the input signal might not be band limited, and the graphics hardware supports only linear reconstruction. Unfortunately, point sampling does not minimize the approximation error and aliasing may occur, see Figure 2a. A more accurate approach is to densely sample \( f \) inside triangles and calculate the vertex values by taking barycentric-weighted averages [KL06]. This is analogous to low-pass filtering which reduces aliasing, but at the cost of blurring the result, see Figure 2b.

One possibility to improve reconstruction accuracy is to refine and/or optimize the mesh. However, game levels often push the limits of the geometry budgets and the artists are not enthusiastic to compromise on scene size or geometric detail. Moreover, geometry refinement would involve unnecessary replication of all other vertex attributes (vertices, tangent frames, texture coordinates, etc.) and would impact the whole game development pipeline, because geometry also serves purposes other than rendering, such as gameplay and physics.

In this paper, we assume a fixed mesh and we formulate and solve an optimization problem that minimizes the error in a least squares sense, see Figure 2c. However, with more complex geometry, the least squares solution is not always as visually pleasing because high variations in vertex values can be introduced, see Figure 3a. To remedy this, we study regularization techniques that trade off accuracy for smoothness, allowing us to minimize the mach band artifacts. Our experiments indicate that a small amount of regularization improves the visual quality considerably while still staying close to the optimum.

Our techniques have been implemented in a game production lighting tool and we provide examples of vertex baked ambient occlusion and precomputed radiance transfer in real game scenes, demonstrating that our method provides smoother and more accurate approximations than previous methods. From a practical standpoint, vertex baking resulted in significant memory savings compared to textures, freeing up resources for other components of the game while requiring only minimal changes in the development workflows and the game runtime.

2. Related Work

Vertex baking is common on low-end gaming platforms like the Nintendo Wii, cell phones and tablets, where memory and power consumption considerations require that compromises be made. On high-end platforms, texture memory consumption and shader complexity are still issues and the savings resulting from vertex baking can significantly improve the quality of the final products [LH09, Wan11]. Another problem is that texture mapping requires parametrization, often carefully authored to avoid undesirable uv-seams. In recent years, texturing techniques that do not require an explicit parametrization have been proposed [BL08, YKH10], but efficient implementations to date have only been presented on high-end graphics hardware.

The problem of representing signals on triangle meshes occurs in many areas of computer graphics. Much of the early work in Radiosity [CW93] deals with representing lighting signals by point sampling. More accurate results can be achieved by adapting the mesh to conform to the signal [HSA91] or by using Galerkin methods [Zaf93]. Discontinuity meshing [LTG92] refines the mesh at shadow boundaries to more accurately represent the signal before or during the solution. While these methods are important, they are less useful in games, where geometry refinement or higher-order reconstruction functions are typically not an option due to resource limitations and hardware capabilities.

A related approach studied in the literature is to sample a fine mesh and then simplify it to better represent a given signal [SWH’95, Hop96]. However, the final signal used in these techniques is computed using a greedy optimization process which does not come with any optimality guarantees. In this paper, we leverage advances in linear system solvers over the past decades [BBK05] to efficiently compute optimal solutions for a fixed mesh, resulting in higher quality.
Precomputed radiance transfer [SKS02] is another technique that usually stores data at the vertices of a mesh. In the literature, these meshes are typically finely tessellated and some work has been done toward computing the solution more efficiently [HR10] or adapting the mesh [KP04]. Precomputed radiance transfer, especially if only used for indirect lighting, is another example of a signal that can be efficiently represented using our technique.

Probably the closest work to ours [KL06] reduces aliasing artifacts when storing volumetric irradiance by performing the volumetric equivalent of the barycentric-weighted averages described in this paper. Note that in spite of the aliasing issues, point sampling at vertex locations is still one of the most common solutions, and in fact is the default in many applications, including Autodesk Maya. The artifacts are either resolved during manual post-processing or are masked by another layer of high-frequency textures. Our method results in more visually pleasing approximations, eliminating the need for any post-processing or masking.

3. Vertex Baking

As input, we have a triangular mesh with vertices \(v_1, \ldots, v_N\) representing a piecewise linear surface \(S \subset \mathbb{R}^3\). We assume the vertices at creases have been duplicated, i.e., each connected component of the mesh corresponds to a smooth patch (smoothing group). We further assume there is an integrable function \(f : S \to \mathbb{R}\) that assigns each surface point a scalar value (vector functions are handled per-component). The task is to find vertex values \(x = (x_1, \ldots, x_N)^T \in \mathbb{R}^N\) such that their piecewise linear interpolant \(g_x : S \to \mathbb{R}\) is as close as possible to \(f\) in the least squares sense, i.e., minimizes

\[
E(x) = \int_S (f(p) - g_x(p))^2 dp
\]

where \(p\) is an arbitrary surface point. The function \(g_x\) can be formally expressed as

\[
g_x(p) = \sum_{i=1}^N h_i(p) x_i
\]

where \(h_i : S \to \mathbb{R}\) are piecewise linear hat functions (nodal shape functions from linear finite elements) such that \(h_i(v_i) = 1\) and \(h_i(v_j) = 0\) for \(i \neq j\). In the following we drop the subscript from \(g_x\) for improved readability.

3.1. Optimal Fit

Equation (1) represents a continuous linear least squares problem. To solve it, we substitute Equation (2) into Equation (1) and re-arrange the terms, obtaining

\[
E(x) = x^T Ax - 2x^T b + c
\]

where \(A \in \mathbb{R}^{N \times N}, b \in \mathbb{R}^{N \times 1}\) and \(c \in \mathbb{R}\) are defined as follows

\[
A_{i,j} = \int_S h_i(p) h_j(p) dp
\]

\[
b_i = \int_S h_i(p) f(p) dp
\]

\[
c = \int_S f^2(p) dp
\]

The matrix \(A\) is sparse, symmetric positive definite and is also known as the mass or Gram matrix. Its entries can be calculated analytically: \(A_{i,j}\) is the total area of the triangles incident to vertex \(i\) divided by 6 and \(A_{i,j}\) is the total area of the triangles incident to edge \(ij\) divided by 12 (zeros elsewhere). The vector \(b\) and scalar \(c\) typically cannot be calculated analytically, because we usually do not have a closed-form expression for \(f\) and, therefore, we apply Monte Carlo integration, see Section 3.3.

Because \(A\) is positive definite, \(E(x)\) is convex and we can find its global minimum by solving \(\partial E(x) / \partial x_i = 0\) for \(i = 1, \ldots, N\), which reduces to

\[
Ax = b
\]

A common approximation is to replace \(A\) with a diagonal matrix \(D\) such that \(D_{i,i} = \sum_{j=1}^N A_{i,j}\) a technique known as mass-lumping. Note that the mass lumped solution \(\tilde{x}\) is

\[
\tilde{x}_i = \frac{\int_S h_i(p) f(p) dp}{\int_S h_i(p) dp}
\]

which correspond to a weighted average (with \(h_i\) being the weighting function [KL06]). However, better results can be achieved by solving Equation (7) exactly, see Figure 2b,c. We discuss the numerical solution methods in Section 3.3.

3.2. Regularization

While the least squares formulation from Equation (1) works very well for simple scenes as in Figure 2, in more complex scenes the objective \(E(x)\) becomes overly eager to exploit all available degrees of freedom. For example, Figure 3a shows the exact solution of Equation (7) applied to a game scene. The piece-wise linear approximation \(g\) exhibits large gradient discontinuities, visually exacerbated by the mach band effect.

To obtain a better behaved \(g\), we augment the linear system from Equation (7) with a smoothness (or regularization) term \(R \in \mathbb{R}^{N \times N}\), resulting in

\[
(A + \alpha R)x = b
\]

where \(\alpha \geq 0\) is a user-tunable parameter. The simplest choice, inspired by Tikhonov regularization, i.e., \(R := A\), is inadequate because it corresponds to reducing brightness by pushing \(x\) to zero. A more suitable \(R\) should not affect the absolute values of \(x\). This motivates gradient-based regularization,
which can be expressed as
\[
\int_S (f(p) - g(p))^2 \, dp + \alpha \int_S \|\nabla g(p)\|^2 \, dp \tag{10}
\]
To obtain the regularization matrix, we note that \(\nabla g\) is piecewise constant which allows us to split the smoothness term per triangle. For triangle \(t\) with vertices \(v_i, v_j, v_k\),
\[
\nabla g_t = (x_j - x_i) \frac{(v_i - v_k) \perp}{2a_t} + (x_k - x_i) \frac{(v_j - v_k) \perp}{2a_t}
\tag{11}
\]
where \(a_t \in \mathbb{R}\) is the area of triangle \(t\) and \(\perp\) denotes a counterclockwise rotation by 90° in the plane of the triangle [BKP*10]. One subtle issue is that the units of \(\nabla g\) may not match the units of \(f(p) - g(p)\) (the former is inversely proportional to the units of distance, while the latter is independent of them). This can be fixed by scaling the gradient by the square root of the triangle area, i.e., introducing \(\nabla g_t = R \nabla g_t\). Employing \(\nabla g_t\) instead of \(\nabla g_t\) makes the units of the data and smoothness terms match and corresponds to giving more weight to the gradients of larger triangles. Because the gradient is linear in \(x\), we can write \(\nabla g_t = F_t x\) where \(F_t\) is a \(3 \times N\) sparse matrix. This allows us to express the regularization term as
\[
\int_S \|\nabla g(p)\|^2 \, dp = \sum_i a_i \|\nabla g_t\|^2 = \sum_i a_i x^T F_t^T F_t x
\tag{12}
\]
where \(R_{grad} \in \mathbb{R}^{N \times N}\) is symmetric positive definite. It can be shown \(R_{grad}\) is closely related to the standard discretization of the Laplace-Beltrami operator, i.e., the familiar cotangent formula [LZ10].

Employing the regularization term \(R_{grad}\) leads to a smoother \(g\), see Figure 3b. While \(R_{grad}\) preserves brightness, the side effect is a loss of contrast, because as \(\alpha \to \infty\), the function \(g\) converges to the average of all samples on each connected component of the mesh. This is obviously suboptimal—for example, there is no reason to reduce the gradient of an isolated triangle. Therefore, we propose a better regularization term that explicitly penalizes large gradient discontinuities across internal mesh edges. Specifically, for each pair of triangles \(t\) and \(u\) that share an internal (i.e., non-crease and non-boundary) edge, we sum
\[
\sum_{(t,u)} (a_t + a_u) \|\nabla g_t - \nabla g_u\|^2 = \sum_{(t,u)} (a_t + a_u) x^T (F_t - F_u)^T (F_t - F_u) x = x^T R_{edge} x \tag{13}
\]
where \(R_{edge} \in \mathbb{R}^{N \times N}\) is a symmetric positive definite edge-based regularization term. Employing this new regularization term instead of \(R_{grad}\) results in solutions that better preserve contrast; for example, notice the effect on the background wall in Figure 3. This is because the matrix \(R_{edge}\) encourages smoother gradients over connected parts of the mesh without penalizing gradient magnitudes at boundaries and creases.

3.3. Numerical Solution

So far we have described how to calculate the system matrix. It remains to evaluate the signal-dependent right hand side \(b\) from Equation (5). Because an explicit formula for \(f\) is typically unavailable, we use Monte Carlo integration. Specifically, given approximately uniformly distributed samples \(p_1, \ldots, p_M \in S\), we estimate \(b_i\) as
\[
b_i \approx \frac{\mu_i}{|X_i|} \sum_{j \in X_i} h_i(p_j) f(p_j) \tag{14}
\]
where \(X_i\) is the set of samples in the triangles incident to vertex \(i\) and \(\mu_i\) is the sum of the areas of the triangles incident to \(i\). For ambient occlusion, \(f(p_j)\) is calculated by importance sampling the hemisphere of rays starting at \(p_j\) and ray casting against the scene geometry using 256 directions. For precomputed radiance transfer, this process is modified to account for spherical harmonics basis functions and 1-bounce indirect lighting [SKS02]. Monte-Carlo integration can also be used to estimate the coefficient \(c\) from Equation (6), but this is only necessary if the actual numerical value of the error term \(E(x)\) is needed.

The system matrix \(A + \alpha R\) for \(\alpha \geq 0\) is sparse, symmetric.
positive definite and depends only on the mesh geometry. Typically, this class of linear systems can be most efficiently solved using direct sparse solvers [BBK05]. In our implementation, we used PARDISO [SG06]. Even though the numerical factorization needs to be recomputed for different values of $\alpha$, the system matrix has simple structure and permits solves fast enough for interactive feedback, facilitating parameter tuning. For ambient occlusion, one complication is that while the input values are within $[0, 1]$, the resulting vertex values $x_i$ can under/overshoot (an effect similar to the Gibbs phenomena). An accurate constrained solution could be obtained with quadratic programming, but we found simple clamping to $[0, 1]$ is sufficient for our purposes. In our experiments, only about 5% or fewer of the values $x_i$ are outside $[0, 1]$ and often violate the bounds only slightly. Alternatively, we could also store arbitrary values of $x_i$ and perform the clamping after interpolation (i.e., in the pixel shader).

### 4. Results and Discussion

Our main practical motivation was to efficiently represent ambient occlusion. Aside from the Sibenik cathedral in Figure 1 and the didactic example in Figure 2, we tested our algorithm on three levels from a shipped game, see Figure 4 and Figure 9. To facilitate comparisons, the images are also available as supplementary materials. Note that the game meshes have occasional artifacts such as unintended penetrations or inconsistent winding orders which show up in the result but do not cause our algorithm to malfunction. We compare the results of our method to point sampling, weighted averaging, and ground truth (raytraced using 1024 rays per pixel). Point sampling requires fewer samples, but produces noticeable aliasing artifacts. Weighted averaging uses more samples and produces higher quality results, but mesh dependent mach band artifacts are apparent. Moreover, in large coarse triangles some contrast is lost (see Figure 2b) due to excessive blurring. Our method employing edge-based regularization with a small $\alpha$ uses the same samples as weighted averaging but produces smoother results, better preserves contrast and results in lower approximation error.

We report the quantitative results in Table 1. Note that the reported vertex counts include duplicates created on creases (i.e., vertices that have the same spatial location but different sample values). The error is calculated according to Equation (1), divided by the total mesh area and multiplied by 100. The times were measured on a 6 core dual 2.9GHz Intel X5670 CPU with 24GB memory. The number of samples is a user specified parameter; in our scenes we used the rather conservative number of $10^6$ samples. Increasing the sample count further has only negligible impact on the final vertex signal (see Figure 5). While fewer than $10^6$ samples would probably also suffice, the raytracer performance makes this question less practically important. The additional time and memory complexity of the linear solve required by our method is negligible, see Table 1.

A useful feature of our technique is that the amount of smoothing is controllable by the parameter $\alpha$. For lower values of $\alpha$, we obtain sharp gradients aggressively approximating the input signal, while higher values result in a softer look with more prominent creases, see Figure 10. For even more flexible art control, we could also consider combining the gradient- and edge-based regularization terms.

While our main motivation was efficient representation of
ambient occlusion, our method can be successfully applied to other input signals, including vector ones. For example, we consider precomputed radiance transfer [SKS02] represented with 6th order spherical harmonics. Each of the 36 dimensions is vertex-baked separately using our method. In Figure 6, we compare the results of weighted averaging to our method ($\alpha = 0.1$ edge-based regularization). Similarly to ambient occlusion, our method produces a smoother and more accurate approximation resulting in more visually pleasing dynamic relighting.

Our technique outperforms previous methods especially with coarse and irregularly tessellated meshes, but more subtle differences can be visible even with finer geometry. We illustrate this on a simple uniformly tessellated L-shaped mesh with varying resolution, see Figure 7. While the deficiencies of weighted averaging begin to disappear with increased resolution, the costs of additional geometry can be prohibitive in large scenes.

While vertex baking is best suited for low-frequency lighting signals, our method is general and applicable even in the presence of sharp edges and noise. To illustrate this, we execute progressive meshes to represent a 298×298 image using a 2000 triangle mesh with a vertex baked RGB signal [Hop96]. Because progressive meshes optimize vertex attributes in a greedy way without any optimality guarantees, our method results in a more accurate approximation for the same coarse mesh, see Figure 8.

5. Conclusion and Future Work

We present a mathematical framework for vertex-level representation of signals on triangle meshes, explicitly exposing the trade-offs between accuracy and smoothness. Compared to previous methods (point sampling and weighted averaging), our technique generates more accurate results, helps to eliminate the mach band artifacts and better preserves contrast of the input signal. Compared to using textures, vertex baking requires less storage and memory bandwidth, resulting in significant savings for real-time applications, such as games. Our method is best suited to low-frequency lighting signals such as ambient occlusion or precomputed radiance transfer. While our approach is applicable to any signal, for higher frequencies and/or repeated patterns, textures are often a better choice. In future work, it would be interesting to couple our technique with mesh optimization [HDD+93] to further improve reconstruction accuracy, consider non-linear regularization terms [ROF92] and combine our method with
emerging texturing schemes [YKH10]. An extension of our method to higher order basis functions [SSD03] could be derived by replacing our piecewise linear hat functions $h_i$ (Section 3) with higher order ones. The increased cost of runtime interpolation would be balanced by higher approximation quality for the same number of degrees of freedom, especially with strictly $C^1$-conforming schemes [CO01].

Acknowledgements. We thank the anonymous reviewers for their feedback and helpful comments. Many thanks to Jeff Grills, Robert Kovach and Tony Bratton for help with integration in production tools and Peter Shirley for careful proofreading.

References


[LLH09] Larsson D., Halen H.: The unique lighting of Mirror’s Edge. In Game Developers Conference (2009). 1, 2


© 2011 The Author(s)
© 2011 The Eurographics Association and Blackwell Publishing Ltd.
Figure 9: Three levels from our game with vertex ambient occlusion values calculated by (left to right): (a) point sampling, (b) averaging triangle samples, (c) our method and (d) ground truth per-pixel rendering.

Figure 10: Varying the regularization parameter (left to right): (a) $\alpha = 0$ (no regularization), (b) $\alpha = 0.01$, (c) $\alpha = 0.1$ (recommended), (d) $\alpha = 1$. 