Eurographics Symposium on Rendering (2006)
Tomas Akenine-Möller and Wolfgang Heidrich (Editors)

Making Radiance and Irradiance Caching Practical: Adaptive Caching and Neighbor Clamping

Jaroslav Křivánek[1] Kadi Bouatouch[2] Sumanta Pattanaik[3] Jiří Žára[1]

[1]Czech Technical University in Prague, Czech Republic [2]IRISA – INRIA Rennes, France [3]University of Central Florida, USA

Abstract

Radiance and irradiance caching are efficient global illumination algorithms based on interpolating indirect illumination from a sparse set of cached values. In this paper we propose an adaptive algorithm for guiding spatial density of the cached values in radiance and irradiance caching. The density is adapted to the rate of change of indirect illumination in order to avoid visible interpolation artifacts and produce smooth interpolated illumination. In addition, we discuss some practical problems arising in the implementation of radiance and irradiance caching, and propose techniques for solving those problems. Namely, the neighbor clamping heuristic is proposed as a robust means for detecting small sources of indirect illumination and for dealing with problems caused by ray leaking through small gaps between adjacent polygons.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism - Rendering, Global Illumination

1. Introduction

One of the practical and widely used algorithms for computing diffuse indirect illumination is irradiance caching [WRC88, WH92, TL04]. The indirect illumination is computed only at a sparse set of points on surfaces, stored in a cache, and then interpolated elsewhere. Radiance caching [KGPB05, Kři05] generalizes irradiance caching to glossy surfaces with low-frequency BRDFs.

To faithfully represent indirect illumination with only a sparse set of values, their density must be proportional to the rate of change of indirect illumination; otherwise, interpolation artifacts may appear. In irradiance caching, the upper bound on the illumination change is estimated based on the scene geometry. Even though the actual illumination conditions are disregarded in such an interpolation error metric, it generates a record density that gracefully follows the indirect illumination changes, and provides good image quality with a relatively low number of cached values.

In radiance caching, however, estimating the rate of change of indirect illumination is more difficult. On glossy surfaces, not only the illumination characteristics, but also the sharpness of the surface BRDF and the viewing direction influence the actual rate of change of indirect illumination. It would be complicated to design an interpolation error criterion that takes all those factors into account, and the resulting formula is likely to be quite computationally involved. We believe that an adaptive algorithm that refines the density of cached values based on a simple perceptual metric is more appropriate.

In this paper we propose an adaptive algorithm for controlling density of cached values in radiance and irradiance caching that we refer to as *adaptive caching*. It starts with an initial set of cached indirect illumination values and then refines their density where necessary to eliminate interpolation artifacts. The main source of artifacts in radiance and irradiance caching are visible discontinuities at a boundary of influence areas of cached values. The proposed error criterion guiding the adaptive refinement is designed to detect those discontinuities. The criterion is, thus, perceptual in the sense that it does not primarily care about the physical correctness of the image, but about the visible artifact elimination.

With adaptive caching, the record density is adapted to the actual illumination conditions better than with the original criterion used in radiance caching [KGPB05]. If indirect illumination is simple, without sudden changes, adaptive caching generates fewer records and rendering is faster. Under complex indirect illumination, adaptive caching can

generate more records than the original criterion would, but without the risk of compromising image quality by interpolation artifacts. Adaptive caching is the most effective in rendering of glossy surfaces with localized indirect illumination features, where it locally increases the record density. With adaptive caching, the record density also automatically adapts to the BRDF sharpness. Additionally, record density is relaxed under strong direct illumination, where the indirect illumination quality does not matter much.

Besides the adaptive caching algorithm, we discuss some issues one must address in a practical implementation of radiance and irradiance caching. We propose a heuristic called *neighbor clamping* and show how it is useful for solving two problems of radiance and irradiance caching: (1) missing small sources of indirect illumination and (2) artifacts caused by the ray leaking through gaps between polygons.

The rest of the paper is organized as follows: Section 2 presents the related work; Section 3 describes a common formulation of irradiance and radiance caching; Section 4 then describes the adaptive caching algorithm; Section 5 discusses some practical problems and describes the neighbor clamping heuristic; Section 6 presents the results and Section 7 concludes the work.

2. Related Work

The techniques proposed in this paper are tightly coupled with irradiance and radiance caching algorithms, both described in detail in Section 3. Ward *et al.* [WRC88] proposed the record density criterion in irradiance caching based on the estimate of the upper bound of the irradiance gradient. This has been adopted by Křivánek *et al.* [KGPB05, Kři05] for radiance caching. The adaptive record density control proposed in this paper is a simple yet practical extension of the Ward *et al.*'s criterion. Tabellion and Lamorlette [TL04] propose practical modifications of the original Ward *et al.*'s irradiance caching algorithm, discussed in detail in Section 5.2. Smyk *et al.* [SM02] increase the density of the cached values based on the gradient magnitude to better accommodate to abrupt changes in indirect illumination, similarly to the way done in the code of the Radiance lighting simulation system [War94] (described in Section 5.1 of this paper). However, this approach is not sufficient for radiance caching due to the view dependence of glossy BRDFs.

The adaptive algorithm we use for steering the density of cached values is closely related to adaptive mesh refinement in radiosity [SP94,CW93], with the difference that our representation of illumination is independent of scene geometry. In particular, our approach is similar to perceptually-driven radiosity [GH97, MPT97], where a perceptual criterion is used to control adaptive mesh refinement in hierarchical radiosity, based on the perceived smoothness of illumination.

In kernel density estimation [WHSG97], the kernel bandwidth is adjusted to reduce variance without introducing too much bias. However, samples of radiance (the photon hits) are given and fixed at the time of the density estimation. On the other hand, we take new samples as needed to perform smooth radiance reconstruction. In density control for photon maps [Suy02], the density of photons is limited by the local target density to avoid storing too many photons. The target density is proportional to pixel importance to equalize reconstruction error over the image.

3. Irradiance and Radiance Caching

In this section we give a common interpolation formula for both irradiance and radiance caching. Irradiance caching [WRC88,WH92] and radiance caching [KGPB05,Kři05] are based on the observation that indirect illumination changes slowly over a diffuse or a glossy surface. Both algorithms compute indirect illumination only at a sparse set of points on surfaces and store it in a cache. The indirect illumination computation is accelerated by reusing the cached indirect illumination through interpolation.

Cached Quantity. In irradiance caching, the cached quantity is irradiance, E, at a point, therefore the directional information about incoming radiance is discarded. In radiance caching, the directional information is retained, which allows to interpolate over surfaces with view-dependent BRDFs. The incoming radiance is represented by a coefficient vector, Λ, with respect to spherical or hemispherical harmonics [GKPB04]. The cached quantity is the only essential difference between the two algorithms; the caching scheme is the same for both.

Caching Scheme. Whenever indirect illumination needs to be computed at a point $\mathbf{p}$, the cache is queried for cached nearby indirect illumination values (called *records*). If no record is found near $\mathbf{p}$, interpolation cannot be used. In such a case the hemisphere above $\mathbf{p}$ is sampled, and irradiance E (in irradiance caching) or the coefficient vector Λ (in radiance caching) is computed and stored in the cache. If, on the other hand, cached records are found near $\mathbf{p}$, the cached indirect illumination is interpolated.

Indirect Illumination Interpolation. Here, we unify the interpolation formulas of irradiance and radiance caching by interpolating the outgoing radiance contributions:

$$L^o_{\mathscr{S}}(\mathbf{p}) = \frac{\sum_{i\in\mathscr{S}} L^o_i(\mathbf{p}) w_i(\mathbf{p})}{\sum_{i\in\mathscr{S}} w_i(\mathbf{p})}, \qquad (1)$$

where $L^o_i(\mathbf{p})$ is the contribution of the i-th record to the outgoing radiance at the point $\mathbf{p}$. For irradiance caching at a point with diffuse reflectance $\rho(\mathbf{p})$, $L^o_i(\mathbf{p})$ is:

$$L^o_i(\mathbf{p}) = [E_i + (\mathbf{n}_i \times \mathbf{n}) \cdot \nabla_r E_i + (\mathbf{p} - \mathbf{p}_i) \cdot \nabla_t E_i]\rho(\mathbf{p})/\pi.$$

For radiance caching, the outgoing radiance contribution is given by the dot product:

$$L_i^o(\mathbf{p}) = \left[\mathbf{R}_i\left(\Lambda_i + d_x\frac{\partial\Lambda_i}{\partial x} + d_y\frac{\partial\Lambda_i}{\partial y}\right)\right]\cdot C_{(\mathbf{p},\omega_o)}.$$

Various symbols used here are summarized below:

$\mathbf{p}$	point of interpolation.
$\mathbf{n}$	normal at $\mathbf{p}$.
$\mathbf{p}_i$	location of the i-th cache record.
$\mathbf{n}_i$	normal at $\mathbf{p}_i$.
d_x, d_y	displacements of $\mathbf{p}-\mathbf{p}_i$ along the x and y axes of the local coordinate frame at $\mathbf{p}_i$.
$w_i(\mathbf{p})$	weight of the i-th cache record with respect to point $\mathbf{p}$, defined below.
$\mathscr{S}$	set of records used for interpolation at $\mathbf{p}$, defined below.
E_i	cached irradiance value at $\mathbf{p}_i$.
$\nabla_r E_i$	cached rotational irradiance gradient at $\mathbf{p}_i$.
$\nabla_t E_i$	cached translational irradiance gradient at $\mathbf{p}_i$.
Λ_i	coefficient vector for the incoming radiance at $\mathbf{p}_i$.
$\frac{\partial\Lambda_i}{\partial x}$	cached derivative of Λ_i with respect to translation along the local x axis.
$\frac{\partial\Lambda_i}{\partial y}$	cached derivative of Λ_i with respect to translation along the local y axis.
$\mathbf{R}_i$	rotation matrix used to align the coordinate frames at $\mathbf{p}$ and $\mathbf{p}_i$.
$C_{(\mathbf{p},\omega_o)}$	coefficient vector representing the BRDF lobe at $\mathbf{p}$ for the outgoing direction ω_o.

Weighting Function. Weight of the i-th cache record with respect to point $\mathbf{p}$ is given by $w_i(\mathbf{p}) = \left(\frac{\|\mathbf{p}-\mathbf{p}_i\|}{R_i} + \sqrt{1-\mathbf{n}\cdot\mathbf{n}_i}\right)^{-1}$. The harmonic mean distance, R_i, to the objects visible from $\mathbf{p}_i$, is computed from the ray lengths in hemisphere sampling and stored in the cache.

Interpolation Criterion. The definition of the set $\mathscr{S}$ of records used for interpolation is of particular interest for this paper. $\mathscr{S}$ is defined as $\mathscr{S} = \{i | w_i(\mathbf{p}) > 1/a\}$, where a is a user defined maximum allowed interpolation error. The higher the value of a, the bigger the allowance for interpolation and the smaller the accuracy of the final image. The definition of the set $\mathscr{S}$ means that the record i can be used for interpolation in the neighborhood of its location, where $w_i(\mathbf{p}) > 1/a$ holds. This neighborhood is larger in flat open areas, where the indirect illumination changes slowly, and smaller on curved objects and near scene geometry, where the indirect illumination changes more quickly. This adapts the density of the cached radiance values to the expected rate of change of the indirect illumination. The interpolation scheme was derived in [WRC88] for diffuse surfaces. Even though the derivation does not hold for glossy ones, the interpolation gives surprisingly good results even in radiance caching. However, in some situations it fails. This is why we have developed a perceptual error criterion to guide the record density, described in the following section.

4. Adaptive Caching

In this section we present the error criterion used to guide record density and an algorithm that uses this criterion to distribute records on object surfaces.

4.1. Error Criterion and Adaptive Refinement

Adaptive caching uses the original irradiance caching interpolation scheme described above, and an adaptive record density is achieved by *modulating the approximation error value, a, on a per-record basis*, based on our error criterion. So each record i now stores its own value of approximation error, a_i, instead of having one global value for the whole scene.

The goal of the adaptive record density control is to eliminate image artifacts caused by interpolation in places with high rate of change of indirect illumination. Those artifacts exhibit themselves as discontinuities at the boundary of influence areas of two or more records (see the cutout in Figure 7, top). The influence area of a record is defined as $\{\mathbf{p}; w_i(\mathbf{p}) > 1/a\}$. Our error criterion reports a visible discontinuity if two records have a visibly different outgoing radiance contribution somewhere in the overlap of their influence areas. Whenever a visible discontinuity is detected, we decrease a record's value of a_i, effectively locally increasing the record density. In other words, we force the radiance contribution in the overlap areas to be indistinguishable from each other.

Consider a point $\mathbf{p}$ at which the outgoing radiance is computed with the interpolation formula (1) from a set $\mathscr{S}$ of at least two records. We say that the overlap of those records causes a discontinuity if a record i exists in $\mathscr{S}$, such that the outgoing radiance *without* this record's contribution, $L^o_{\mathscr{S}\setminus\{i\}}(\mathbf{p})$, is discernable from the outgoing radiance *with* this record's contribution, $L^o_{\mathscr{S}}(\mathbf{p})$.

If a discontinuity-causing record i is found, its contribution is excluded from the interpolated radiance at $\mathbf{p}$ by decreasing its approximation error to $a_i := \frac{\|\mathbf{p}-\mathbf{p}_i\|}{R_i} - \varepsilon$, where ε is a very small number (see Figure 1). This ensures that the condition $w_i(\mathbf{p}) > 1/a_i$ does not hold anymore and record i does not contribute to the interpolated radiance at $\mathbf{p}$. If there are more than one candidate for exclusion, we choose the one with the lowest weight $w_i(\mathbf{p})$, remove it from $\mathscr{S}$, and reiterate the process. Intuitively, the record with the lowest weight is relatively the farthest from point $\mathbf{p}$ and should be excluded. The metric used to assess the discernibility of two radiance values is discussed in Section 4.3.

4.2. Full Convergence

The criterion described above can be used to compute indirect illumination at a point $\mathbf{p}$ in a straightforward way:

$\mathscr{S} :=$ look_up_cache$(\mathbf{p},\mathbf{n})$

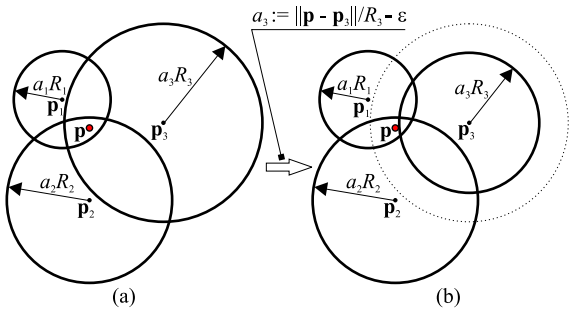


Figure 1: *(a) If the outgoing radiance $L^o_{\mathscr{S}}(\mathbf{p})$, $\mathscr{S} = \{1,2,3\}$ at the point* **p** *is visibly different from $L_{\mathscr{S}\setminus\{i\}}(\mathbf{p})$ for some $i \in \{1,2,3\}$, the overlap area is said to cause a discontinuity. (b) If a discontinuity is detected, the approximation error a_i of the record contributing with the minimum weight $w_i(\mathbf{p})$ is decreased (here record #3), so that point* **p** *be outside the record's influence area.*

If $\mathscr{S}$ is empty, add new record.
If $|\mathscr{S}| = 1$, interpolate from the single record in $\mathscr{S}$.
If $|\mathscr{S}| > 1$, run the adaptive refinement and interpolate from the records that remain in $\mathscr{S}$.

However, this approach does not work well, since a contribution from a record that will be shrunk later might have been mistakenly used earlier for interpolation. For a correct result, the record distribution must *fully converge* with respect to a set of shading points, before any record is used for interpolation at any of those points. We call *shading points* all the points with indirect illumination computed by means of radiance or irradiance caching (either by interpolation or by adding a new record). Typically those are the surface points visible through image pixels either directly or after an ideal specular reflection or refraction.

Definition. Given a set of shading points $\mathscr{P}$, we say that the *record density has fully converged* for points in $\mathscr{P}$ if:

1. each point in $\mathscr{P}$ has at least one contribution from existing records, and
2. for all points having contributions from two or more records, the error criterion does not report a visible discontinuity.

Algorithm 1 describes an efficient procedure for attaining full convergence for a given set of shading points. For the sake of efficiency, we store a list of record contributions with each shading point, so that the (ir)radiance cache does not have to be queried repetitively to find the contributing records for a shading point (line 7). Each contribution in the list stores a pointer to the contributing record, weight $w_i(\mathbf{p})$, outgoing radiance $L^o_i(\mathbf{p})$ (evaluated lazily as requested by the error criterion) and a change counter (described below).

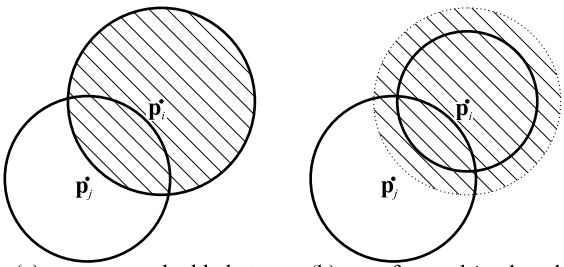
The basic steps of Algorithm 1 are motivated by the two following situations, illustrated in Figure 2: (a) After a new record is added to the cache, all the shading points in its influence area should be checked again by the error criterion, hence are added to pointQueue (line 11). (b) If the approximation error a_i of the record i is decreased (line 13), some shading points might find themselves without any valid contribution. New records must then be added to the cache to cover those points. To efficiently detect the points with no valid contribution, the algorithm works in passes (the outermost while loop of Algorithm 1). At the end of each pass (lines 17-19), we loop over all records whose a_i was decreased in the current pass. For each of the affected records, all the shading points within that record's influence area *as it was at the beginning of the pass* are added to pointQueue. The possible invalid contributions due to the affected records are then removed from those points' contribution lists on

Algorithm 1 Full convergence for a set of shading points.

Input:
The set of shading points $\mathscr{P}$ for which indirect illumination is to be computed by radiance or irradiance caching. Each shading point contains position, local coordinate frame, BRDF pointer, and a *list of record contributions*, empty at start.

Ensures:
All the shading points in $\mathscr{P}$ have at least one contribution from a cache record and for all points having contributions from two or more records, the error criterion does not report a visible discontinuity.

```
1:  pointQueue.push(P)
2:  while pointQueue not empty do
3:    // new pass starts here
4:    while pointQueue not empty do
5:      // for all points to be processed in this pass
6:      p_s := pointQueue.pop()
7:      remove invalid contributions from the shading
        point p_s
8:      if zero contributions in p_s then
9:        add a new record at the position of p_s
10:       add a contribution to each shading point in the
          new record's influence area
11:       add each point in the new record's influence area
          to pointQueue
12:     else if more than one contribution in p_s then
13:       run adaptive refinement the contributing records
14:     end if
15:     // do nothing if p_s has exactly one contribution
16:   end while
17:   for all records affected by refinement in this pass do
18:     add to pointQueue shading points in the record's
        influence area from the beginning of the pass
19:   end for
20:   // go to next pass
21: end while
```



(a) - new record added at $\mathbf{p}_i$ (b) - a_i of record i reduced

Figure 2: *Shading points in the hatched areas are inserted into pointQueue in the following two situations:* ***(a)*** *After a new record is added to the cache.* ***(b)*** *At the end of each pass for each record affected by the adaptive refinement.*

line 7. New records are then possibly added to (ir)radiance cache on lines 9 to 11. The same effect could be achieved by inserting the shading points into pointQueue immediately after a record's a_i-value is decreased. However, this would not be efficient since the a single record's a_i-value is typically reduced more than once in a single pass.

Upon the reduction of a record's a_i value, some of the record's contributions may become invalid. It is thus necessary to check the validity of the entries in the contribution list (line 7) by re-evaluating the weight of a record that has changed since its contribution was last updated. To keep track of the contributions that require the validity check, we use *change counters*. Each record has a counter, incremented every time its a_i is reduced. The value of the change counter is copied to the contribution each time the validity of the contribution is verified on line 7. A contribution remains certainly valid as long as the record's counter and the contribution's counter are equal.

The whole algorithm finishes if no records have been changed in the last pass. Full convergence is ensured for all shading points at its end. In our test, the algorithm always finished in at most eleven passes. The overhead due to this algorithm and due to the evaluation of the perceptual criterion is very small (see Table 1).

There are a few important details to make the algorithm efficient:

- To avoid multiple insertions of one shading point to pointQueue, each shading point has a bit indicating its presence in the queue.
- The initial value of a for a new record is copied from the nearest existing record in the cache and multiplied by 1.4, a multiplier found empirically and tested on a wide range of scenes. Lower values lead to unnecessarily many records after the convergence; higher values do not decrease the number of records but increase the overhead of the convergence algorithm.
- Shading points are stored in a k-D tree [Jen01] to quickly locate the points in a record's influence area (lines 10 and 18).
- The contribution list for a shading point is a static array of length 6. We found that there are almost never more than six contributions for one shading point. Should this be the case, the excess contributions may safely be ignored (six is already enough). Using a dynamic linked list introduces a serious performance penalty because of the constant memory allocation and de-allocation.

This algorithm inverts the classical irradiance and radiance caching algorithm, where the basic operation is: 'given a shading point, find all contributing records and interpolate'. On the other hand, the basic operation in our convergence algorithm is: 'given a record, find all shading points it contributes to and add a contribution'. This is similar to radiance cache splatting [GKBP05] or reverse photon mapping [HHS05].

An alternative implementation of the convergence algorithm without the per-point contribution list uses a (ir)radiance cache look-up on line 7 to gather the contributions for a point p_s and leaves out line 10. Such an implementation consumes less memory, is simpler to implement, but it is significantly slower.

4.3. Discernability Metric

This section describes the metric used to assess discernability of two radiance values. After some experimentation, we have opted for the simplest metric, the Weber law. Weber law says that the minimum perceptible change in a visual signal is given by the fixed fraction of the signal. Although not conservative under all circumstances, the threshold of 2% is widely used in computer graphics (*e.g.* [WFA*05]) and we used it in all our renderings.

The input of the perceptual metric is two radiance values L_1 and L_2. L_1 is the interpolated outgoing radiance contribution from all overlapping records, plus direct illumination at $\mathbf{p}$,

$$L_1 = L^o_{\mathscr{S}}(\mathbf{p}) + L^{o_{direct}}(\mathbf{p}).$$

L_2 does not contain the contribution from the tested record, i,

$$L_2 = L^o_{\mathscr{S}\setminus\{i\}}(\mathbf{p}) + L^{o_{direct}}(\mathbf{p}).$$

The discernability metric based on the Weber law has the following form:

$$L_1 \text{ differs from } L_2 \quad \equiv \quad |Y(L_1) - Y(L_2)| > \Delta_{max},$$

where

$$\Delta_{max} = 0.02\, Y(L_1) + \max\{\sigma(L_1), \sigma(L_2)\}.$$

Y(.) denotes the luminance channel of a tri-stimulus value. Estimates of the standard deviation $\sigma(L_1)$ and $\sigma(L_2)$ are added to compensate for the randomness of L_1 and L_2, stemming from the fact that they are computed from quantities

estimated by Monte Carlo hemisphere sampling. If the standard deviation was not added and only a few rays were used to sample a hemisphere when creating new cache records, it is likely that the irradiance or incoming radiance stored in neighboring records would be very different. In such a case, the criterion would constantly report visible discontinuities and lead to an excessive record density, which is undesirable.

To estimate the standard deviation of irradiance computed by Monte Carlo hemisphere sampling, an irradiance value E is computed from all the sample rays, and another value E' is computed from every second ray. Standard deviation of irradiance, estimated as $\sigma(E) = |E - E'|$ [DS04], is then used to compute the standard deviation of the outgoing radiance. We use the standard deviation of irradiance even in radiance caching with good results.

We experimented with the threshold elevation model of Ramasubramanian *et al.* [RPG99]. We decided not to use it, since visible artifacts were produced near edges, where the model predicts high threshold elevation. This is due to the fact that (ir)radiance caching artifacts have distinct structure that makes them more easily perceptible than unstructured noise for which the model has been designed.

5. Practical Issues

This section describes techniques that help making an implementation of radiance or irradiance caching practical.

5.1. Adapting Record Density by Gradient Magnitude

The source code of Radiance lighting simulation system [War94] contains a (never published) test that avoids a too low record density if the indirect illumination gradient is high. The test consists in comparing the upper bound of the translational irradiance gradient, derived from the "split sphere model" [WRC88], with the actual irradiance gradient, estimated from the hemisphere sampling [WH92]. Should the magnitude of the actual irradiance gradient exceed the supposed upper bound, the actual gradient magnitude will be used in the error metric instead of the upper bound. Technically, this is done by clamping the value of the harmonic mean distance R_i when a new record is being created:

$$\textbf{if} \quad \frac{|\nabla_t E_i|}{E_i} > \frac{1}{R_i}, \quad \textbf{then} \quad R_i := \left(\frac{|\nabla_t E_i|}{E_i} \right)^{-1}. \qquad (2)$$

The use of this test substantially improves the image quality produced by irradiance caching.

We use a similar test in radiance caching:

$$\textbf{if} \quad \Delta_i > 1/R_i, \quad \textbf{then} \quad R_i := 1/\Delta_i,$$

where Δ_i is the ratio of the radiance gradient L2-norm to the L2-norm of the radiance itself:

$$\Delta_i = \frac{\sqrt{\| \frac{\partial \Lambda_i}{\partial x} \|^2 + \| \frac{\partial \Lambda_i}{\partial y} \|^2}}{\|\Lambda_i\|}.$$

This test works aggregately on the whole coefficient vector which represents the result of the whole hemisphere sampling. Therefore, the test is neither view dependent nor it takes the BRDF sharpness into account. It helps to detect the most serious cases of high gradient change, but it leaves much space for improvement to our adaptive radiance caching.

5.2. Neighbor Clamping

In this section we propose *neighbor clamping*, a heuristic used to detect geometrically small sources of indirect illumination.

The radius of the i-th record's influence area $\{\mathbf{p};\ w_i(\mathbf{p}) > 1/a_i\}$ on a flat surface is given by the product $a_i R_i$, which follows from the definition of $w_i(\mathbf{p})$. Here R_i is the harmonic mean of distances to visible surfaces from the record position, $\mathbf{p}_i$, computed from the ray lengths during hemisphere sampling. The closer the surrounding geometry, the smaller R_i, and the higher the record density. However, because the sample rays from which R_i is computed are only a random subset of all directions in the hemisphere, small features are likely to be missed. The computed value of R_i is, then, too large and produces a too low record density. If the missed features are sources of strong indirect illumination, interpolation artifacts in the image result. Because of the randomness of ray directions, the features missed by one record might not be missed by another, which even amplifies the noticeability of the image artifacts.

Examples of features most commonly missed are steps of a staircase, or windowsills on a facade, which may be too small to keep the harmonic mean of the ray lengths low, yet too important in terms of indirect illumination to be missed. The left column of Figure 4 shows the artifacts due to the insufficient record density around a geometry feature (the steps).

Our goal is to make sure that no relevant geometry features are missed. Additionally, a geometry feature detected by the rays from one record should be detected by all nearby records, too.

Tabellion and Lamorlette [TL04] address this problem by computing the R-value of a record as the minimum of the ray lengths during the hemisphere sampling, instead of taking the harmonic mean. This increases the probability of detecting small geometry features and indeed, the step-like features are missed very rarely. However, after some experimentation we decided not to use the minimum ray length for the following reasons. Firstly, using the minimum ray length is overly sensitive to extremely small features that hardly have any importance for indirect illumination. Secondly, and more importantly, the minimum ray length on concave objects is very small, as the rays near the equator are usually extremely short. This produces an excessive record density on those objects. Tabellion solves this problem by excluding

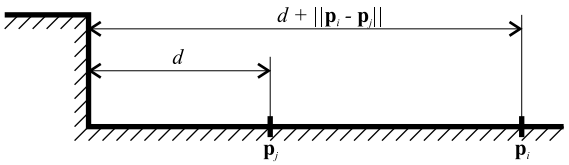


Figure 3: *The maximum possible difference of the two records' distance from the step is given by their mutual distance.*

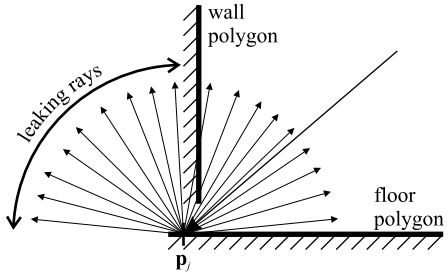


Figure 5: *Inaccurate connection of polygons may result in ray leaking.*

rays near the equator, hitting other surfaces at grazing angles, from the computation of the minimum distance [Tab05]. But even with this modification, we found it difficult to obtain a decent distribution of records on concave surfaces of varying curvature by using the minimum ray length for computing R_i.

We have thus decided to compute R_i as the harmonic mean of the ray lengths, and to prevent random missing of geometry features by imposing an additional constraint on the difference of the R-values of neighboring records. The constraint stems from a basic observation on geometrical coherence in a scene. Consider a record j at position $\mathbf{p}_j$ located on a floor at a distance d from a step (see Figure 3). Now consider another record at position $\mathbf{p}_i$. By triangle inequality, the maximum possible distance of $\mathbf{p}_i$ from the step, in terms of d and the distance between the two records is $d + \|\mathbf{p}_i - \mathbf{p}_j\|$. Motivated by this observation, we never allow the R-values of two nearby records to differ by more than their mutual distance.

Technically, when a new record i is being added to the cache, we locate all nearby records j and clamp the new record's R_i value to $R_i := \min\{R_i, R_j + \|\mathbf{p}_i - \mathbf{p}_j\|\}$. After that we similarly clamp the nearby records' R-values by the new record's R_i-value. A consequence of this clamping is that a too large R-value of a record, caused by missing a geometrical feature, is clamped down by the R-value of some of the neighboring records that did not miss that feature.

This heuristic, which we call *neighbor clamping*, is fully justified when using the minimum ray length for computing R_i, but it gives very good results even for the harmonic mean. The features are almost never missed, and the overall distribution of records in the scene behaves well. Furthermore, we do not experience the problem with excessive record density on concave objects. Figure 4 demonstrates how neighbor clamping (right column) helps to detect small, step-like geometry features. Without neighbor clamping (left column), those features are often missed and artifacts appear in the image (see the detail of the stairs). Both images were rendered using approximately the same number of records (7750). Without neighbor clamping, at least 20,000 records were required to get rid of the image artifacts on the stairs.

5.3. Ray Leaking

Irradiance and radiance caching are quite sensitive to imperfections in scene modeling; a typical example in which they break down is an inaccurate connection of adjacent edges of two polygons. This may be produced *e.g.* by an insufficient number of significant digits when a scene is exported to a text file.

Consider the situation in Figure 5, where there is a small gap between the floor polygon and the wall polygon. If a primary ray hits this gap, its intersection with the floor polygon can be found *behind* the wall polygon. As a consequence, rays that are supposed to hit the wall now *leak* either to the neighboring room or to infinity. The outcome of such an event is quite disastrous:

- The computed irradiance or incoming radiance is wrong.
- The harmonic mean of the ray lengths is much grater than it should be; therefore, the wrong (ir)radiance is extrapolated on a very large area.

The resulting image artifacts are shown in Figure 6 on the left. The following paragraphs discuss two possible solutions to this problem, among which the more successful is to use neighbor clamping.

The first solution is to shift the origin of sampling rays away from polygon edges [TL04, Tab05]. The ray origin shifting suppresses ray leaking in some case, but is not dependable.

A better solution consists in using neighbor clamping. Records suffering from ray leaking have disproportionately bigger R-value than their neighbors not having this problem, thereby breaking the assumption of geometrical coherence used to derive the neighbor clamping heuristic. Therefore, ray leaking is reliably detected by the use of neighbor clamping and its consequences are alleviated by the reduction of the erroneous R-value. Figure 6 on the right shows how neighbor clamping detects and suppresses the effects of ray leaking. Besides small feature detection, ray leaking detection is another purpose that neighbor clamping serves well.

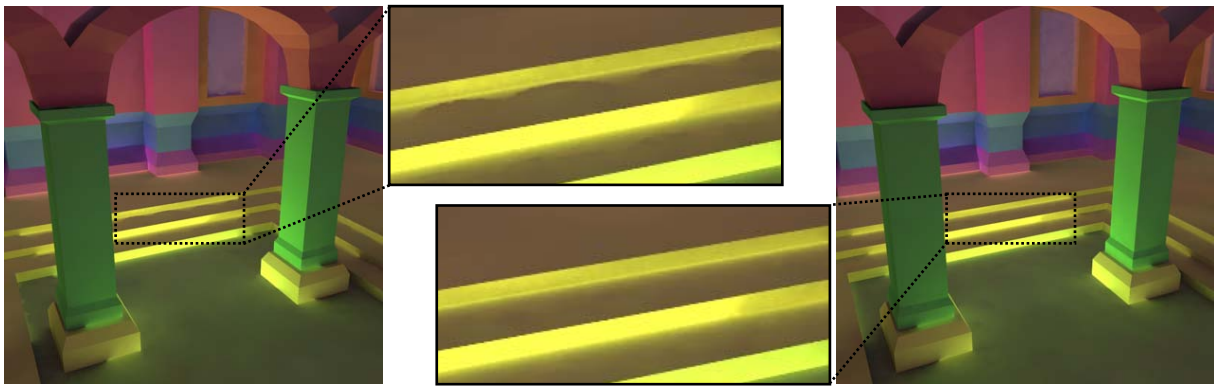


Figure 4: *Without our neighbor clamping, small, step-like geometry features are often missed. Missing geometry features that are strong sources of indirect illumination produces disturbing image artifacts (see the detail of the steps in the left column). Neighbor clamping helps to detect the small geometry features and suppresses the image artifacts (see the right column). The images show only indirect illumination.*

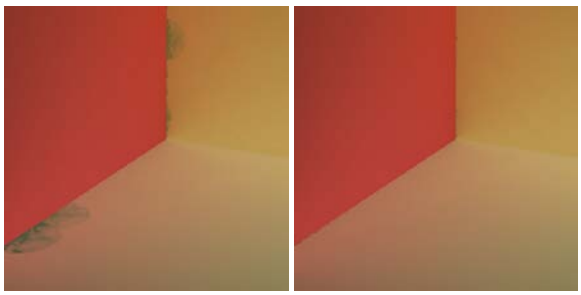


Figure 6: *Serious image artifacts caused by ray leaking (left) are significantly reduced by using the proposed neighbor clamping heuristic (right).*

6. Results and Discussion

Adaptive Radiance Caching. Results of adaptive radiance caching are shown in Figures 7 to 10; the rendering statistics, measured on a PC with Intel Pentium M 1.5GHz and 512MB RAM, are given in Table 1. Notice that the extra overhead introduced by adaptive caching is very small. For each scene, we rendered one image with adaptive caching and another without adaptive caching using approximately the same number of records. We compare the quality of the two renderings. Apart from the rendered images, the figures show the map of the approximation error a and the record locations. Notice how the value of a and the record density adapt to the rate of change of indirect illumination with our adaptive caching.

We do not provide a comparison with a reference solution because our error metric promotes smooth, artifact-free images, but does not ensure any error bound for the image. Our goal is to generate images that look good and are plausible, but not necessarily physically correct.

The floor of the box in Figure 7 uses the Lafortune BRDF [LFTG97] fitted to measured metallic BRDF data [Wes00]. Only direct illumination and first bounce indirect illumination for the floor is used. Non-adaptive radiance caching has problems capturing the change of indirect illumination caused by the edges of the box and the pyramid. On the other hand, the record density behind the two objects is unnecessarily high. Adaptive caching produces a record density that reproduces all the details of the indirect illumination. To obtain an artifact-free image with non-adaptive caching, at least 3,300 records were required, which is 3 times more than with adaptive caching.

The scenes in Figures 8 and 9 use four indirect bounces and both irradiance and radiance caching. The teapot in Figure 8 has a Phong BRDF with the exponent of 16. Non-adaptive caching runs into troubles capturing the reflection of the spout on the teapot body, which adaptive caching renders faithfully. In this scene, adaptive caching also increases record density on the teapot rim, where there are no noticeable artifacts even with the non-adaptive algorithm. This suggests that a better perceptual metric, which possibly takes the projected record size into account, would probably allow even lower number of records without deteriorating image quality. The teapot scene required at least 2,800 records for an artifact-free image without adaptive caching, that is 1.7 times more than with adaptive caching.

In the Walt Disney Hall scene (Figure 9), the adaptive record density is useful not only for radiance caching (cutouts 1 and 2), but also for irradiance caching. Adaptive irradiance caching nicely captures the caustics formed on the floor at the entrance to the building (cutout 3), and also delivers a better image quality on the staircase, where the non-adaptive caching leaves some smudges. Notice the slight visible discontinuity on the metal wall at the very right of the image. This is actually a discontinuity of first order, caused

	Cornell Box 800×800		Teapot 800×800		Walt Disney Hall 1280×800	
	adaptive ($a = 0.55$ - 0.05)	non-adaptive ($a = 0.23$)	adaptive ($a = 0.6$ - 0.05)	non-adaptive ($a = 0.15$)	adaptive ($a = 0.35$ - 0.05)	non-adaptive ($a = 0.18$)
# records	990	1010	1589	1695	4078	4144
Cache filling [s]	68.6	68.7	94.3	94.8	195	196
Overhead [s]	2.18	1.6	4.18	2.81	11.2	9.34

Table 1: *Rendering statistics for adaptive radiance caching example scenes.* Cache filling *is the time spent on creating radiance cache records (hemisphere sampling).* Overhead *is the time spent on interpolating from cache records and evaluating the perceptual error criterion. Notice that the extra overhead introduced by adaptive caching is very small.*

by the inversion of the indirect illumination gradient and by the change of the tint from brownish to bluish. Our current discernability metric does not capture those aspects of discontinuity perception. For an image quality similar to that produced by adaptive radiance caching, the non-adaptive radiance caching required as many as $29,000$ records (7 times more than adaptive caching). For non-adaptive *irradiance* caching, $17,500$ records were sufficient, which is only 1.3 times more than with adaptive irradiance caching.

Figure 10 demonstrates how record density automatically adapts to the sharpness of a BRDF. The floor is assigned the isotropic Ward-Duer BRDF [Due05, War92] with $\rho_d = 0$, $\rho_s = 0.9$ and $\alpha = 0.60$, 0.30 and 0.15 (from left to right). With increasing BRDF sharpness, the record density automatically grows to reproduce the sharper features of the indirect illumination, without requiring any user intervention. The number of records is 335, 516, and 993, respectively.

Discussion of the Radiance Caching Results. Figures 7, 8 and 9 show that adaptive radiance caching mostly pays off if there is an indirect illumination feature that requires a locally high record density. In the non-adaptive caching, the only solution to render such a feature is to increase the record density in the whole image by manually altering the global value of the approximation error a. This is inefficient, since most of the image already has a high enough record density. We would like to point out that on glossy surfaces, the localized features almost always appear; therefore, adaptive radiance caching is very profitable.

In a walkthrough animation with adaptive caching, one might observe slight popping artifacts as refinements occurs at new locations. To suppress the artifacts in high quality rendering, we first let the Algorithm 1 converge for each frame, and then perform the actual image synthesis.

Adaptive Irradiance Caching Results and Discussion. To test our adaptive algorithm with irradiance caching, we rendered a number of scenes, of which two are shown in Figure 11. We have found that the number of records and the rendering times for both adaptive and non-adaptive irradiance caching are similar. This is caused by the fact that localized indirect illumination features rarely appear on diffuse surfaces. (A notable exception are caustics, usually handled by other means than irradiance caching.) Thus, the adaptive record density does not bring as much profit to irradiance caching as it does to radiance caching.

The irradiance caching results demonstrate that extending the combination of the original Ward's error criterion [WRC88] and the adaptation by the gradient magnitude (Section 5.1) with the proposed neighbor clamping heuristic (Section 5.2) results in a robust, dependable interpolation error criterion. Within the irradiance caching framework, it would probably be difficult to find an error criterion that computes the indirect illumination with a significantly fewer records, without missing any illumination features.

7. Conclusion and Future Work

In this paper we described an adaptive algorithm for controlling record density in radiance and irradiance caching. The algorithm is based on detecting visible discontinuities caused by inadequate sampling of indirect illumination. If a discontinuity is detected, record density is locally increased in order to produce smooth, artifact free images. As a result, record density automatically adapts to the complexity of indirect illumination—fewer records are used for smoothly varying illumination and more records are generated in areas of abrupt illumination changes. Moreover, the record density adapts to the BRDF sharpness and the viewing direction. The adaptation is automatic, without the need for user intervention. We demonstrated on a number of scenes that the image quality produced by adaptive caching is superior to that generated by the error criterion of radiance caching.

We also proposed the neighbor clamping heuristic, which significantly reduces the probability of missing small geometry features. Neighbor clamping also helps to detect, and alleviate the consequences of, ray leaking—a serious problem of irradiance and radiance caching.

Adaptive caching in combination with neighbor clamping represent a step forward towards more robust radiance and irradiance caching—the algorithms are now more reliable and much less tweaking is needed for high quality results.

In future work, it would be interesting to use the results of Durand *et al.*'s frequency analysis of light transport [DHS*05] to design a more complete interpolation error criterion for glossy surfaces. Using wavelets in radiance caching for the incoming radiance representation would allow interpolation on surfaces with higher-frequency BRDFs.

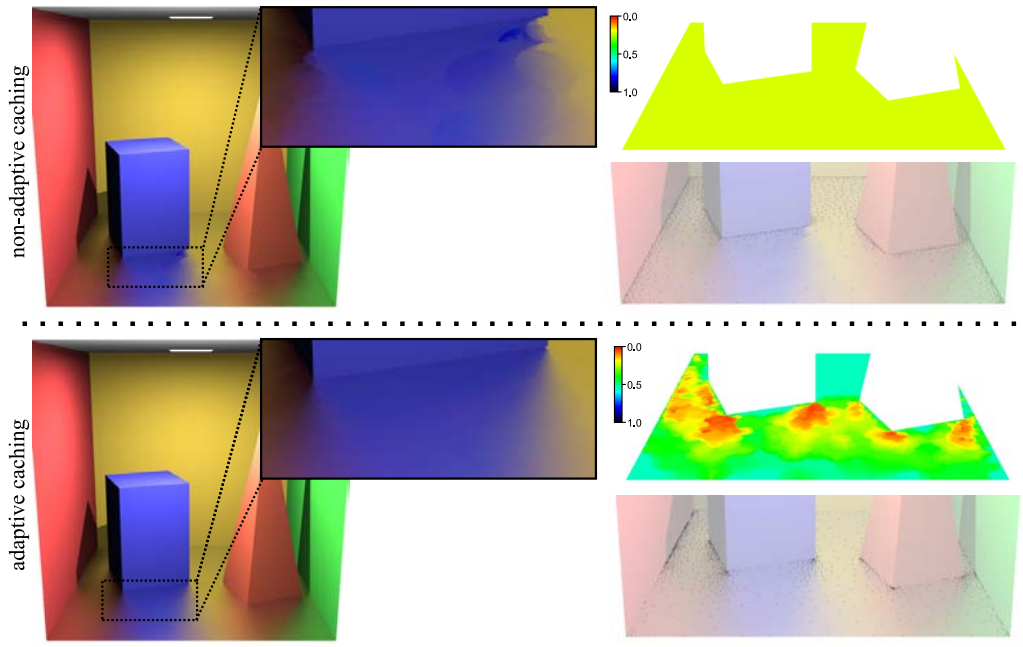


Figure 7: *Comparison of non-adaptive (top) and adaptive (bottom) caching for the Cornell box scene. The figure shows the rendered images, cutouts from the rendered images, color coded values of the approximation error a and record positions.*

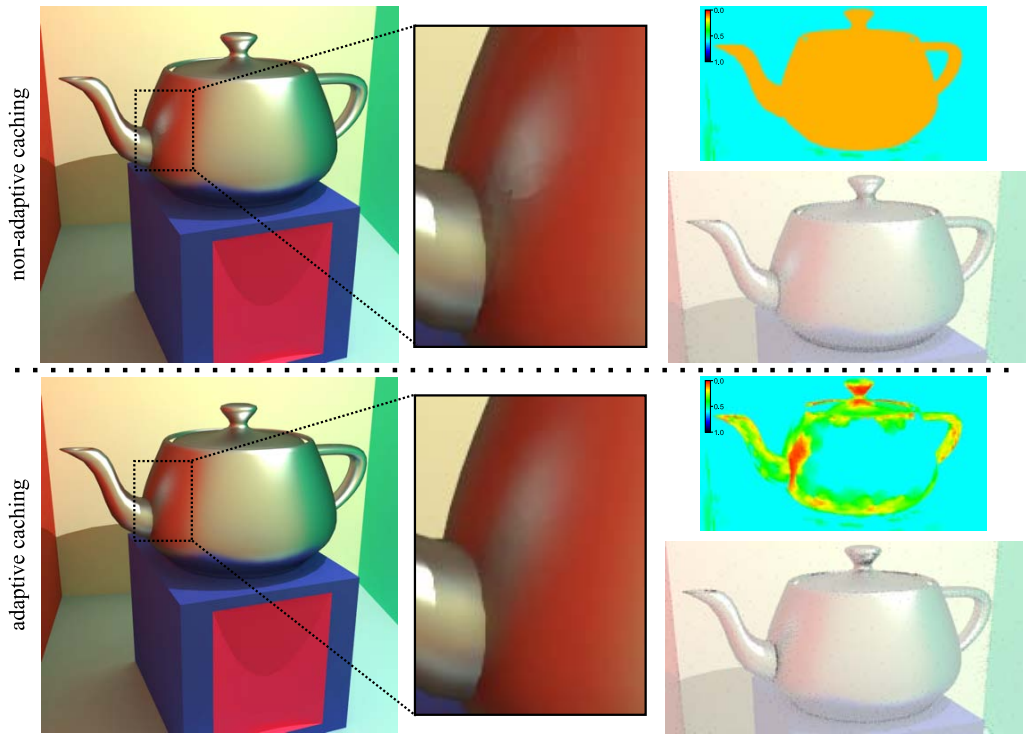


Figure 8: *Comparison of non-adaptive (top) and adaptive (bottom) caching for the Teapot scene. The figure shows the rendered images, cutouts from the rendered images, color coded values of the approximation error a and record positions.*

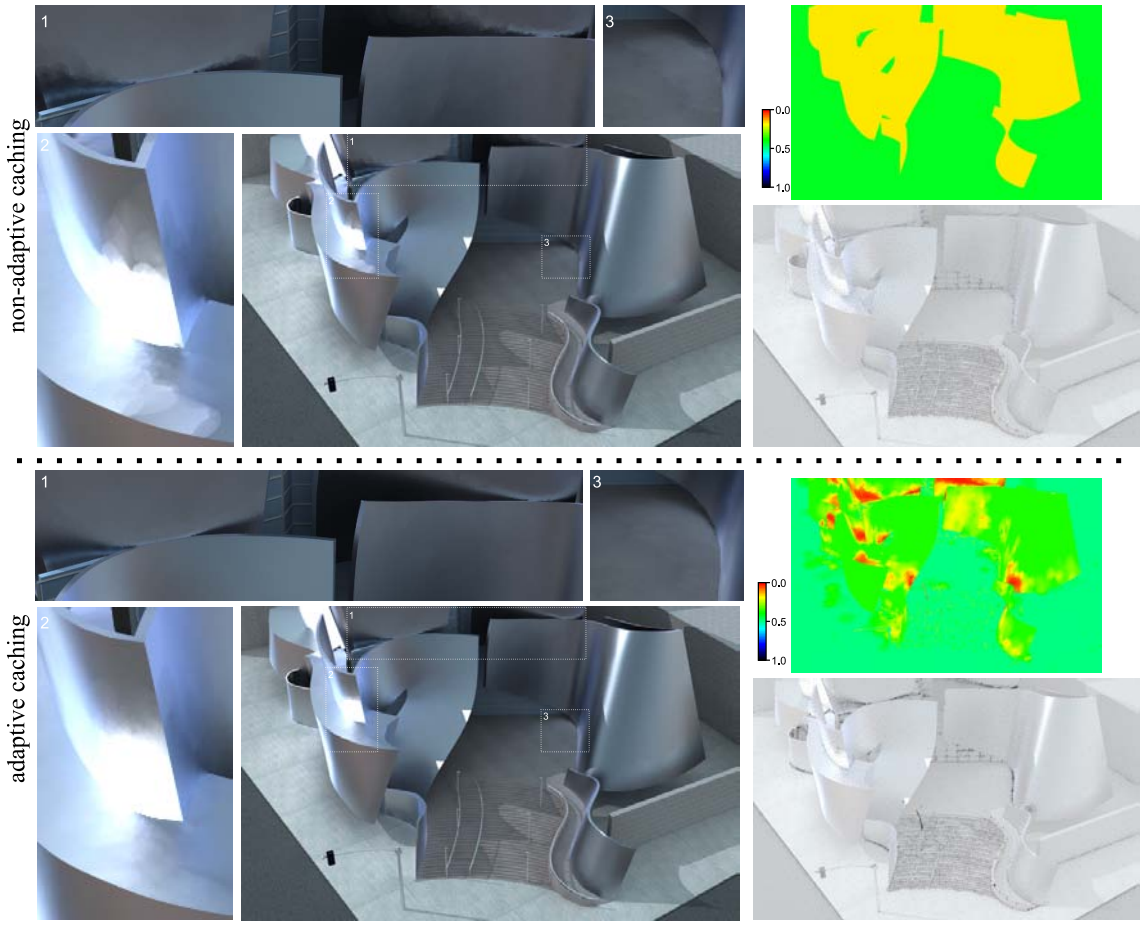


Figure 9: *Comparison of non-adaptive (top) and adaptive (bottom) caching for the Walt Disney Hall. The figure shows the rendered images, cutouts from the rendered images, color coded values of the approximation error a and record positions.*

Acknowledgements

This work has been supported by France Telecom R&D, Rennes, France, the Ministry of Education of the Czech Republic under the research programs LC-06008 (Center for Computer Graphics), and MSM 6840770014.

References

[CW93] COHEN M. F., WALLACE J. R.: *Radiosity and Realistic Image Synthesis*. Morgan Kaufmann, San Francisco, CA, 1993.

[DHS*05] DURAND F., HOLZSCHUCH N., SOLER C., CHAN E., SILLION F. X.: A frequency analysis of light transport. *ACM Trans. Graph. (Proceedings of SIGGRAPH 2005) 24*, 3 (2005).

[DS04] DMITRIEV K., SEIDEL H.-P.: Progressive path tracing with lightweight local error estimation. In *Vision, modeling, and visualization 2004 (VMV-04)* (2004).

[Due05] DUER A.: On the ward model for global illumination. Unpublished material, 2005.

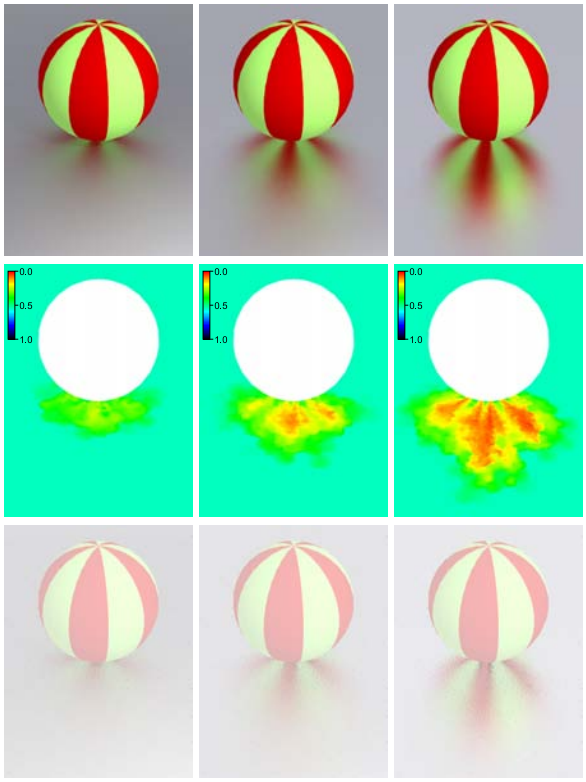
[GH97] GIBSON S., HUBBOLD R.: Perceptually-driven radiosity. *Computer Graphics Forum 16*, 2 (June 1997).

[GKBP05] GAUTRON P., KŘIVÁNEK J., BOUATOUCH K., PATTANAIK S. N.: Radiance cache splatting: A GPU-friendly global illumination algorithm. In *Rendering Techniques 2005* (June 2005).

[GKPB04] GAUTRON P., KŘIVÁNEK J., PATTANAIK S. N., BOUATOUCH K.: A novel hemispherical basis for accurate and efficient rendering. In *Rendering Techniques 2004* (June 2004), pp. 321–330.

[HHS05] HAVRAN V., HERZOG R., SEIDEL H.-P.: Fast final gathering via reverse photon mapping. *Computer Graphics Forum 24* (2005).

[Jen01] JENSEN H. W.: *Realistic Image Synthesis Using Photon Mapping*. A. K. Peters Ltd., Natick, MA, 2001.



$\alpha = 0.60$, #records = 335 | $\alpha = 0.30$, #records = 516 | $\alpha = 0.15$, #records = 993

Figure 10: *Automatic adaptation of record density to the sharpness of the BRDF. The floor is assigned the Ward-Dür BRDF with* $\rho_d = 0$, $\rho_s = 0.9$ *and (left to right)* $\alpha = 0.60$, 0.30, *and* 0.15.

Figure 11: *Scenes used to test adaptive irradiance caching.*

[KGPB05] KŘIVÁNEK J., GAUTRON P., PATTANAIK S., BOUATOUCH K.: Radiance caching for efficient global illumination computation. *IEEE TVCG 11*, 5 (2005).

[Kři05] KŘIVÁNEK J.: *Radiance Caching for Global Illumination Computation on Glossy Surfaces*. PhD thesis, Université de Rennes 1 and Czech Technical University, December 2005.

[LFTG97] LAFORTUNE E. P. F., FOO S.-C., TORRANCE K. E., GREENBERG D. P.: Non-linear approximation of reflectance functions. In *SIGGRAPH '97* (1997).

[MPT97] MARTIN I., PUEYO X., TOST D.: An image-space refinement criterion for linear hierarchical radiosity. In *Graphics Interface '97* (1997), pp. 26–36.

[RPG99] RAMASUBRAMANIAN M., PATTANAIK S. N., GREENBERG D. P.: A perceptually based physical error metric for realistic image synthesis. In *SIGGRAPH '99* (1999).

[SM02] SMYK M., MYSZKOWSKI K.: Quality improvements for indirect illumination interpolation. In *Proceedings of the International Conference on Computer Vision and Graphics* (2002).

[SP94] SILLION F., PUECH C.: *Radiosity and Global Illumination*. Morgan Kaufmann, 1994.

[Suy02] SUYKENS - DE LAET F.: *On Robust Monte Carlo Algorithms for Multi-pass Global Illumination*. PhD thesis, Katholieke Universiteit Leuven, September 2002.

[Tab05] TABELLION E.: E-mail communication, 2005.

[TL04] TABELLION E., LAMORLETTE A.: An approximate global illumination system for computer generated films. *ACM Trans. Graph. 23*, 3 (2004).

[War92] WARD G. J.: Measuring and modeling anisotropic reflection. In *SIGGRAPH '92* (1992).

[War94] WARD G. J.: The Radiance lighting simulation and rendering system. In *SIGGRAPH '94* (1994).

[Wes00] WESTIN S. H.: Lafortune BRDF for RenderMan. http://www.graphics.cornell.edu/~westin/lafortune/lafortune.html, 2000.

[WFA*05] WALTER B., FERNANDEZ S., ARBREE A., BALA K., DONIKIAN M., GREENBERG D. P.: Lightcuts: a scalable approach to illumination. *ACM Trans. Graph. 24*, 3 (2005).

[WH92] WARD G. J., HECKBERT P. S.: Irradiance gradients. In *Proceedings of the Third Eurographics Workshop on Rendering* (1992), pp. 85–98.

[WHSG97] WALTER B., HUBBARD P. M., SHIRLEY P., GREENBERG D. P.: Global illumination using local linear density estimation. *ACM Trans. Graph. 16*, 3 (1997).

[WRC88] WARD G. J., RUBINSTEIN F. M., CLEAR R. D.: A ray tracing solution for diffuse interreflection. In *Proceedings of SIGGRAPH '88* (1988), pp. 85–92.