$2\,nd$ Eurographics Workshop on Rendering, Barcelona, Spain, 1991. Published as "Photorealistic Rendering in Computer Graphics", Springer-Verlag.

Structure-Directed Sampling, Reconstruction, and Data Representation for Global Illumination

George Drettakis
Eugene Fiume

Abstract

The structure of the sampling and reconstruction process, as well as that of the scene and its data representation, can be utilised to effect significant benefits in global illumination algorithms. As an introduction, previous methods that have utilised this structure are reviewed and some possible improvements are suggested. A new framework is then introduced that utilises the concept of an *oracle* as an abstraction to provide information about the complex computational processes of global illumination. This framework is based on the recognition of the relevant aspects of scene structure, the description of the problems to be solved in terms of oracles, the use of structure to create approximations to the oracles, and finally techniques to monitor the cost, error and quality of the approximations. Each aspect of the framework is discussed in detail, presenting certain oracles that are useful for global illumination algorithms. Two examples of its application are presented. The first example is a simple local illumination problem, that allows us to consider several aspects of the oracle methodology. The second example uses the framework to enhance the progressive refinement radiosity based solution to global illumination. Both examples illustrate the benefits of the proposed framework. Finally, some suggestions are made for future research directions.

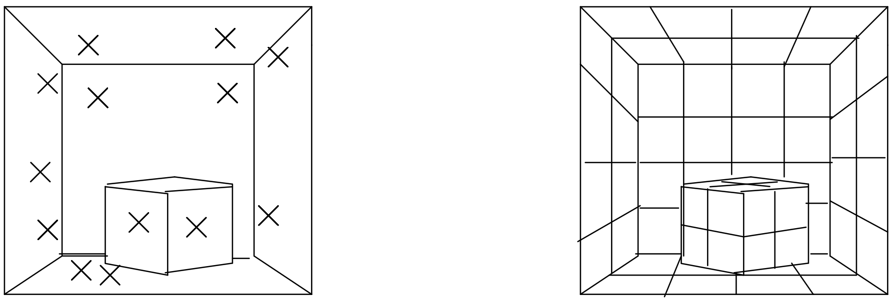
1. Introduction

The sampling theorem for the digital representation of continuous signals places exact constraints on the class of signals that are expressible when using a fixed set of regularly-spaced samples. For our purposes, the ‘‘signal’’ will be an illumination function defined throughout a scene. Many signals can interpolate any finite set of sample values, and the ‘‘least complex’’ of these signals, which is usually an alias for a more desirable signal, is unlikely to be visually satisfactory, unless we are careful to use prefiltering to reduce the class of signals that we intend to represent. The additional facts that, in computer graphics, reconstruction mechanisms in displays are often beyond our control, and that illumination signals inherently have significant high-frequency content, implies that the sampling rate should be even further increased; if this is not possible, then an even smaller class of signals can be faithfully represented. This gloomy news is of course simply a paraphrasing of well-known results from signal theory. However, for several reasons, the news is not nearly as bad as the theory suggests. First, results from texture mapping have shown that careful filtering can dramatically improve image quality using a fixed sampling rate. This suggests that current sampling rates are generally visually adequate. Second, the human visual system appears to be fairly adept at coping with some forms of noise. Third, our *knowledge* of scene structure and of the rendering process can be used to reduce the set of samples required to represent a given illumination signal. If we are careful, therefore, a much smaller set of illumination samples (associated to objects in space) than the set of image samples can be taken. The problem of creating a rectangular set of pixel samples then reduces to performing an appropriate resampling of the set of illumination samples.

Global illumination models and their corresponding algorithms typically require a large amount of storage and computation to simulate light-object interactions, and to finally generate a two-dimensional image. In this paper we introduce the idea of using *structure*-directed rendering to effect significant computational and storage savings. There are two relevant aspects of structure in the rendering process. The first is the structure of the actual scene, and that of its data representation. The second is the structure that exists in, or is imposed on, the sampling and reconstruction process during image generation.

We will propose a framework that utilises these forms of structure, with two main goals in mind. First, to achieve better understanding of global illumination and the rendering process, and second, to develop a set of suitable abstractions and conceptual mechanisms that will lead to better algorithms for rendering. The framework that we will introduce hinges on the following four components: (i) the identification of the aspects of the scene and sampling/reconstruction structure that are relevant for the specific problem being solved, (ii) the development of the *oracle* concept, which is a convenient abstraction allowing us to reason about illumination and rendering, (iii) the development of approximations to these oracles that will lead to algorithms and practical solutions, and (iv) error and quality metrics allowing us to evaluate the quality of the approximations developed.

Before entering a detailed discussion of this framework, we will try to give an intuitive feeling for how knowledge of structure can be used in rendering. The first example is the selection of samples to be used in illumination calculations. If we have a totally diffuse environment, we only require a small number of samples, placed at locations on the surfaces where there is a significant change in illumination. With appropriate reconstruction kernels to interpolate these sample values, accurate pictures can be produced from very little information. If we compare this to the amount of information required in a traditional radiosity-based solution, where the surfaces have to be blindly subdivided into patches, and possibly elements, we see that there are significant savings to be had (see Fig. 1). If a process existed that could choose the best sample locations and the best filters for reconstruction, it would in effect be an *oracle* for image synthesis. In its pure and perfect form, such a process is infeasible, since it would require previous knowledge of the distribution of illumination, a way to minimise the number of samples selected and also an inverse-reconstruction solution to choose the appropriate filters. Nevertheless, a number of meaningful approximations to an ideal oracle can be made, using known properties of a given scene, and structuring the sampling and reconstruction process.



Sparse Sampling   Patch-based subdivision

Fig. 1: Gains from Sparse Sampling

A second example illustrating a similar concept is that of *importance sampling* in Monte Carlo methods. In effect, an ideal importance sampling technique attempts to concentrate the samples taken to approximate an illumination integral, based on visibility and light power information, on those regions that contain the largest amount of information (Hammersly and Handscomb 1964). Again, given an appropriate *oracle*, the computation of an integral over a hemisphere around a point collecting or shooting illumination, could be sped up considerably. The oracle would identify which parts of the hemisphere contribute the most to the integral, and the sampling operation of space would then be limited to these regions. Utilisation of knowledge about scene structure, sample placement information and corresponding reconstruction techniques, can allow an approximation of such an oracle.

The concept of structured rendering has appeared in various forms in previous work. In the following section we review such techniques, and discuss how further use of structure could enhance these methods. In the next section we introduce our proposed framework. This involves a discussion of scene properties used in the rendering process, identifying which of these are more important for global illumination. We then discuss the oracle concept in more detail and how such abstractions can be approximated. Some error and quality measures are then proposed. We continue with two examples. In the first, a fully detailed case-study is presented, in which the framework is applied to solve a simple shading problem. To give an indication of how the framework applies to global illumination problems, the second example shows how a set of approximations can result in certain benefits to the progressive refinement, radiosity-based algorithm. We close with a discussion and suggestions for future research. Throughout the text we refer to the sampling operation performed by a grid superimposed on the image-plane as *image-sampling*; the term *space-sampling* refers to the operation of sampling the hemisphere above a point or area to determine illumination.

2. The Use of Structure in Previous Approaches

Structure, either in the form of the data-representation scheme, or in the sampling and reconstruction process, has been previously utilised to varying degrees. We will concentrate on a limited number of approaches, in which some effort has been made to *combine* more than one aspect of structure to attain a certain goal, thus allowing us to illustrate the importance of the concept of using an integrated approach. We examine cases from both the ray-tracing and the radiosity-based global illumination model.

2.1. Ray-Tracing Approaches

Ward et al. in (Ward et al. 1988) and (Ward and Rubinstein 1988) discuss a sparse sampling technique for the calculation of diffuse illumination in a ray-tracing environment. Four instances of structure are utilised in this method.

Sparse Sample Selection. When a ray intersects a surface, diffuse illumination is calculated in one of two ways: the neighbourhood is searched to determine if there are enough diffuse samples nearby, in which case an interpolation is performed, otherwise a hemispherical integral is approximated using a Monte Carlo technique. An error metric based on distance between samples and a weighting function, derived from this distance, is used for sample selection and interpolation respectively.

Stratified Sampling. The Monte Carlo integral is calculated by segmenting the hemisphere into equal regions to effect faster convergence than a simple ‘‘crude’’ (Hammersly and Handscomb 1964) or ‘‘blind’’ Monte Carlo solution.

Data Structure Utilisation for Sample Location. The sparse samples described above are located using a data-structure that maintains a list of such samples allowing faster access, for interpolation and neighbourhood sample determination.

Simple Reconstruction. For reconstruction, the weights used in the interpolation of the diffuse sample values are based on the distance of the point to be shaded to the various sample locations.

In this method, the scene structure, encoded in the octree, is used in the sample selection and interpolation calculations, resulting in significant computational savings. The sample selection process utilises structure in the following sense: (a) new samples are chosen based on an error approximation metric, and (b) when the new values require an integration to be performed, the nature of the hemisphere sampling operation is taken into account to allow the usage of stratified sampling. In addition, the sample selection metric is used in the specification of the weights for reconstruction, albeit in a somewhat arbitrary fashion.

Thomas et al. (Thomas et al. 1989) introduce the *cover* data representation scheme to enchance the ray-tracing algorithm, that allows efficient determination of shadow boundaries and contours. This is a form of structured data representation for sparse image-sampling.

Cover Data Structure. Each object is covered by an inside and an outside cover that are then intersected with the ray to determine proximity to silhouette contours and shadow boundaries.

Sparse Pixel Sampling. Instead of sampling every pixel in a scan-line, pixels are sampled in a uniformly sparse fashion. When a region of interest is identified by the covers, denser sampling is employed.

The algorithm just described introduces a new data-representation scheme (the covers), specifically built to aid in the determination of the required image sampling rate. This is an example of a new structure being imposed on an existing rendering process (ray-tracing), as opposed to the recognition of underlying or previously existing structure.

2.2. Radiosity-Based Approaches

In the radiosity-based approaches, data-representation structure exists in all scenes, due to the subdivision of the surfaces into planar patches, that are essentially area samples in the rendering process (Goral et al. 1984). Moreover, the hemi-cube (Cohen and Greenberg 1986) defines the way in which the three-dimensional space is sampled. Both these aspects are simple forms of structure, but are not specific attempts to achieve benefit from either recognising, or imposing new, structure in the rendering process of the specified algorithm. Nevertheless, there have been cases that do introduce or utilise such information, and we discuss them in what follows.

Substructuring. The substructuring technique (Cohen et al. 1986) introduces an additional level of subdivision that is used to sample the patches in the hemi-cube computation. The additional subdivision, in essence a hierarchical sampling approach, occurs when deemed necessary by examining the variation of the radiosity values calculated on a surface. An iterative approach is used until desirable convergence is achieved. Space supersampling is thus achieved by the introduction of subsamples, called *elements*. One way of viewing the substructuring process, is that the knowledge of scene structure conveyed by the radiosity gradient is used to determine the sampling rate and to improve image quality.

Ray-Tracing Sampling. In (Wallace et al. 1989) an alternative to the hemi-cube sampling approach is introduced, as applied to the progressive refinement solution (Cohen et al. 1988, Chen 1989). All elements in the scene receive power, but an adaptive sampling approach is used when such sampling of a source (an emitting patch), is required. This adaptive technique is based on the magnitude of the contribution of each region of the source to the form-factor, that depends on the geometrical configuration. The use of adaptive sampling imposes a different structure in the sampling process, utilising knowledge of scene geometry information.

In these methods, there has been some interaction of the structure of the underlying data representation (patches or elements) and the process of space sampling to determine form-factors. Here, as in the ray-tracing approaches, reconstruction has received little attention. This is due to the fact that simple interpolation schemes usually suffice for the purposes of the diffuse environments modelled, at least for the image-sampling case. Finally the view-independence constraint has limited the introduction of new data representation schemes (such as space subdivision) that could be utilised in the space-sampling process.

2.3. Additional Structure Utilisation in Existing Techniques

We now discuss some possible enhancements to the methods described above, that will give a broader view of how the knowledge of structure can be beneficial. For illumination models based on ray-tracing or on radiosity, significant benefits can be achieved, either with the introduction of different data representation, or by taking advantage of existing aspects of the rendering process. We discuss a few such possible improvements to existing methods.

2.3.1. Ray Tracing

Sparse Sampling. In Ward et al.'s approach, space-sampling for diffuse illumination could be enhanced by taking geometry information into account. This information could come from the octree structure, or an enhanced hierarchical representation. The stratified sampling approach for the Monte-Carlo space sampling could be improved by utilising what would amount to importance sampling methods. Techniques similar to the shadow acceleration for octrees (Woo and Amanatides 1990) could be used to determine which part of the hemisphere contributes the most significant part of illumination. This would be especially useful for the case of occluded complex environments, where the existing method is noted to have problems (Ward and Rubinstein 1988). Additional scene information, pertaining to geometry and shadows could be added to the criteria for placing samples, or calculating weights for interpolation. This would enhance the image quality. These operations would constitute a first order approximation to a sample selection oracle (see also Section 3.2).

Covers. Another approach to selecting the sparse grid from which to start shooting rays could be based on a scan-conversion first pass (in the spirit of (Weghorst et al. 1984)), augmented with a shadow volume calculation. The shadow boundaries and first-level contours could then be immediately located. This would allow the avoidance of all shadow casting operations for diffuse objects. In addition, the information recovered from the cover intersection operations could determine the choice of the different filters to achieve better effects, linking the cover structure to reconstruction (see also Section 3.2).

2.3.2. Radiosity-Based Approaches

Hemicube. The hemicube calculations can benefit significantly from a different approach to patch subdivision. A pyramid of varying resolution could be used instead of maintaining one or two levels (patches and elements), similar to the idea of MIP maps (Williams 1983). A related method has been proposed in (Houle 1991) to sample linear and area light sources in an adaptive fashion. If the view independence constraint is relaxed, then an octree-type structure can be utilised to sample adaptively, concentrating again on the regions of the hemisphere that contain the most powerful and large light sources.

Ray-Tracing Solution. Similarly to the hemicube approach, hierarchical structures could be used to reduce computation and storage. The pyramid approach can drastically reduce the number of elements visited by each patch emission operation, since the elements will be in effect adaptively collapsed depending on projected size (onto the light-source).

In both cases, data representation would be enhanced and linked more closely to the sampling and reconstruction process. Additional information can be used in the image-sampling phase, by utilising the tagged sample concept introduced in Section 3.3.

3. Oracles and Structure: a New Framework for Global Illumination

The discussion on previous work gave some intuition on what is meant by identification of scene structure, and by the imposition of structure in the sample/reconstruction processes. In addition, indications were given on how the knowledge of such structure can be beneficial to global illumination and rendering algorithms.

In this section we attempt to establish a methodology that will allow the development of systematic structure-directed approaches. Initially, we examine some of the scene properties with respect to illumination and the rendering process. We attempt to isolate those aspects of scene structure that are important, and identify appropriate methods for their representation. We view this as an initial proposal that requires future research, that will enhance our understanding of global illumination and rendering by isolating the salient structural properties of scene representation as well as sampling and reconstruction. We proceed with a discussion of oracles, that are a convenient conceptual device allowing us to reason about global illumination and rendering. By discussing ways to approximate these oracles, we show how the framework can lead to practical implementations. Finally we present some initial thoughts on determining the quality and measuring the error of these approximations.

3.1. Aspects of Scene Structure

Many properties of scene structure influence the complexity of image synthesis. They can be grouped into four major categories: geometry, viewing parameters, light sources and local illumination.

Geometry. A great deal of scene information resides in the geometric models of the individual objects in the scene, as well as in their subsequent interaction for the purposes of determining occlusion and visibility. Geometric primitives usually include objects such as polyhedra, polygons, free-form surfaces, algebraic surfaces, etc.; the "normal form" for many of these objects at rendering time is a polygonal decomposition. Object descriptions sometimes include an explicit, or have a natural, parameterisation, as is the case for bicubic surfaces or patches. Others can be assigned an implicit parameterisation, to aid in such applications as texture mapping. Surface information is usually encoded in the object structure, that may include textures or complex shading model parameters. One important aspect of geometrical representation is the storage of bounding volume information, and an optional hierarchical "level of detail" structure (for example that in (Rubin and Whitted 1980)). Finally, the geometry of the scene can be enhanced by structuring information such as relative visibility or object proximity relationships, that are crucial, especially in light transfer calculations.

Viewing Parameters. The set of viewing parameters that define the contents of the final image are of extreme importance in any rendering model. A fundamental structure representing this information is the viewing pyramid, based the definition of the viewpoint, the screen size and the near and far planes. This volume allows clipping to be performed when appropriate, provides projected screen size information for objects in the scene, and allows optimisations to be performed, such as backface removal. The use of these parameters has found less use in global illumination models, such as ray-tracing or radiosity-based methods, than it has in traditional scan-line visible surface determination algorithms (Sutherland et al. 1974).

Light Sources. The description of light sources for a computer synthesized scene is extremely important. Primary illumination accounts for a very large number of the important visual effects that we perceive, and it has the benefit that it is somewhat simpler to calculate than secondary lighting. There are three main light source characteristics that need to be described: physical geometry, spatial and spectral light power distributions. There are a number of ways that have been proposed to represent these, concentrating mainly on the latter two. Storage of simple descriptions has been used in (Verbeck and Greenberg 1984) while recent work attempts more compact representations that capitalise on scene structure (including relative light-source size, surface properties, etc.)(Houle 1991).

Local Illumination. Despite the increasing focus on realistic global illumination models, it is the local illumination models that characterise much of the interaction of light with objects. A local illumination model is responsible for providing low-level, often high-complexity detail. Such models typically simulate several classes of surface microstructure, anisotropy, surface self-shadowing, as well as a wide variety of isotropic spectral and spatial reflectance functions, and many forms of texture detail (such as intensity-, bump-, displacement-, environment-, and reflectance-mapping). At first blush, the problem of exploiting the structure of local illumination models may seem impossibly difficult, but this is not so. Many aspects of local surface variation are represented in compact analytic, tabular, or probabilistic forms. This suggests that approaches based on multiresolution or levels-of-detail techniques, low-pass (or scale-space) filtering, or vector quantisation would be effective.

Recognising that the above scene properties are influential is only a small part of the battle. It is now critical to develop representations that facilitate efficient evaluation where possible, and that identify problem areas of the scene. In most of the above cases, we have pointed to possibilities for efficient representations, and one goal of our research is to demonstrate how these representations are useful to the implementation of oracles.

3.2. Oracles for Global Illumination

In theoretical computer science, oracles provide an elegant way of relating problems to one another. For our purposes, an *oracle* is a fictional black box that given a request, produces an exact answer for certain aspects of the sampling and reconstruction process. They are used as decision making processes that give answers to (usually complex and difficult) questions. Oracles allow us to encapsulate complex decision making into isolated objects or entities. These objects have the ability to access information on scene structure and all aspects of the rendering process, and can therefore reach approximate or sometimes exact decisions. This abstraction, as we shall see in Section 4.1, need not be limited solely to global illumination, but can also be used in other traditional problems in computer graphics.

We present three oracles directly related to the global illumination problem, thus giving an indication of how such a mechanism can encapsulate some of the difficult selection procedures.

Optimal Sample Selection. The most obvious oracle that would be useful in a global illumination algorithm is one that selects the minimal set of samples located on the visible surfaces of scene S, from which all the pixel colours could then be satisfactorily reconstructed. Since we are determining a set of samples, a reconstruction method r is implied. As we shall see, the selection of r can be the result of a different oracle, but for the purposes of our definition, we assume that r is given. We informally define oracle O_{SaSel}:

$$O_{SaSel}\,(S,\, r,\, q) \;=_{df}\; \textit{a sample set}\ \{S_i\}\ s.t.$$
$$\|\{S_i\}\| = \textit{minimum, w.r.t. image quality metric } q.$$

Optimal Space Sampling. A second important oracle would allow us to determine the minimal set of illumination samples in the 3D environment required to accurately compute the illumination at a given point, from the surrounding hemisphere of space. Given an area A for which illumination is being calculated, a hemi-sphere of integration above A, h, we define oracle O_{SpSa}:

$$O_{SpSa}\,(S,\,h,\,A,\,\tau) =_{def} \textit{a sample } \{S_i\} \textit{ s.t.} \\ \|\{S_i\}\| = \textit{minimum and calcluated} \\ \textit{illumination value is within tolerance } \tau$$

Optimal Image Reconstruction. A third oracle permits efficient reconstruction, and chooses the correct reconstruction kernel for each specific pixel. In this fashion the final image, reconstructed from the minimal sample set (created for example by oracle O_{SaSel}), is equivalent to the image that would be generated by a given, much more dense, image sampling operation. We define oracle O_{IR}, given sample set $\{S_i\}$ and pixel intensity set $\{P_i\}$ (to be calculated):

$$O_{IR}(S,\,\{S_i\},\{P_i\},\,\tau) =_{def} \textit{a filter set } \{f_i\} \textit{ s.t.} \\ \textit{the time to generate } \{P_i\} \\ \textit{to tolerance } \tau \textit{ is a minimum.}$$

Simple brute-force implementations of these oracles either require the solution of the problem beforehand, or are simply impossible. However, even brute-force approaches may lead to the synthesis of effective approximate oracles based on observations of an ensemble of similar problems. One reason that the formulation in terms of these oracles is actually useful is that it allows us to reason about the complex processes of illumination and rendering in a meaningful way. As such it allows us to isolate the parts that require complicated calculation and knowledge, thus enhancing our understanding of the underlying process. In addition this formulation permits the development of approximations to these oracles in an elegant fashion, achieving modularity and encapsulation of certain complex aspects of the procedures modelled, since in this way only the oracle need have access and knowledge of structure.

3.3. Approximating Oracles Using Scene Structure

When defining a problem in terms of our framework, one would typically first define the set of oracles that suit the solution. The next step would be the identification of the scene and rendering process structure that is relevant, and the salient aspects of this structure. One would attempt to identify these aspects and quantify their effect on the selection or decision making process that the oracle is designed to perform. Based on these observations and this formulation an algorithm can be created that approximates the ideal oracle, which we call an *oracle approximation*. For the purposes of discussion, from now on we will use the terms *oracle* and *oracle approximation* interchangeably, since it will be clear from the context which is meant. To clarify the concept of oracle approximations we present a high-level formulation of approximations for the oracles discussed in the previous section.

To approximate optimal sample selection oracle O_{SaSel}, information from all the scene properties should be used. Light source and geometry information can be used to efficiently predict where higher sampling rates are required due to primary illumination phenomena, such as shadow boundaries or variations due to geometry or local illumination characteristics. Sample selection for secondary illumination is a much more difficult problem and must necessarily be related to the techniques for optimal space-sampling, since the space-sampling approach will determine where in space illumination will be distributed to, or collected from. This can be achieved in part by the adaptive methods described in Section 3.4.

To approximate the space-sampling oracle O_{SpSa}, geometry and viewing information must be utilised. View dependent rendering is critical in achieving low space-sampling rates, since numerous optimisations can be made due to specific geometrical configurations, relative visibility and projected size relationships between objects and the screen. "Level of detail" hierarchical structures as well as proximity information can result in large savings by avoiding significant computation (and storage) for parts of the scene that have little or no visible effect. It must be emphasized here that sacrificing pure view independence does not mean that the ability for motion or animation is forfeited. To the contrary, a necessary addition to the structured approach to global illumination would be that of temporal structure. Such a discussion will not be included in this paper however.

For a efficient reconstruction oracle O_{IR} we introduce the notion of *tagged samples*. When the samples are chosen (by O_{SaSel}) during the sample selection stage, this choice is made due to a scene property, such as geometry, or a preprocessing step, such as shadow boundary determination. The reason causing this choice is stored with the samples, so that it can be accessed in the image reconstruction process. Efficient access mechanisms can be built to determine which samples need to be consulted, and by utilising the tags for each sample, appropriate filters can be chosen, that approximate the optimal. For example, for a set of samples resulting from slowly varying diffuse illumination, wide-support blurring filters can be used, while for boundaries, smaller-support, anisotropic, or even nonlinear filters can be used to maintain, or possibly enhance, the boundaries.

There is an important property of oracles that must be noted here: the computational cost of the oracle (i.e., the decision making process), itself. To make our proposed framework worthwhile, the computation time of the oracle itself must be very small compared to the computation time of the task we are attempting to improve. It may be possible to precompute the cost of the oracle, in which case it can be determined beforehand whether the method has any benefit. Alternatively, if the cost can not be precomputed, then it must be monitored during execution. In either case the implementation must always include a cost calculation mechanism that will allow the system to determine whether use of the oracle is beneficial.

3.4. Error and Quality Assessment for Oracles and their Approximations

A fourth important aspect of the proposed framework is the assessment of oracle quality. The quality of the approximation and of the benefits achieved by using the oracle must be continuously monitored. This checking procedure can then be utilised to tune the selection/decision criteria appropriately, achieving higher quality results. This is in effect a true adaptive technique, since the decision logic adapts, depending on performance. Error and quality measures can be derived from the selection criteria and the approximation techniques used. Typically a problem will have an exact solution, and the approximation can be defined in terms of the error resulting from the use of the approximation. The error can then be calculated during execution, and it can be determined whether or not certain selection criteria appear to be performing badly, contributing to higher error rates. If such a case is found, the criteria can be tuned to deliver better results. This error/quality assessment code is an integral part of any oracle's functionality.

Apart from the adaptive capabilities that the error/quality assessment provides for, it allows us to define oracles in terms of a *tolerance*. Such tolerance measures are usually defined in terms of the mean error μ and variance σ^2 (other statistics may be used too). This is very useful for many aspects of global illumination and rendering, since a requirement of exact equality (e.g., of pixel intensities) would be in many cases impossible to meet. In addition, a tolerance requirement can be relaxed, allowing us to define "levels of quality" of various oracles.

4. Two Examples: a Case Study and Global Illumination

This section attempts to solidify the proposed framework described in the previous section through the presentation of two examples. The first is a case-study that, though not a global illumination problem, is sufficiently simple to allow development of a number of details of the approach, and indicate how the framework leads to useful, practical solutions. In the second example, we discuss a more relevant global illumination problem, relating it to some of the oracles described in Section 3.2, albeit on a more general and abstract level than for the case study.

4.1. A Case Study: an Oracle for Scan Conversion

In a traditional scan-conversion system, one may use either Gouraud shading, in which a simple interpolation of colour computed at vertices is performed, or Phong shading, in which the normals are interpolated and the local illumination model is applied at every pixel. The difference in the renderings can be significant, mainly with respect to the highlights and the banding artifacts. Phong shading is obviously much more expensive computationally, so it is desirable to use it as rarely as possible. For the purpose of this discussion we will assume that the desired quality of rendering is the one generated purely by the application of Phong shading throughout (though it is questionable whether or not Phong shading always gives the best results). For a given scene, and within a certain tolerance (defined in terms of the mean μ and the variance σ^2), we can create the same image by selectively using Gouraud shading for certain cases. We represent this requirement in terms of the previously described framework as follows. Let $O(\Pi, S, \tau)$ be an oracle with respect to a primitive Π (e.g., polygon, scan-line) a scene S and a tolerance τ, that answers the question: “To generate an image of the desired quality, which shading rule should I use ?” as follows:

$$\mathrm{O}(\Pi, S, \tau) = \begin{cases} \text{Phong} & \textit{when quality of alternatives exceeds tolerance} \\ \text{Gouraud} & \textit{when satisfactory} \\ \text{Neither} & \textit{when only simple shading required (e.g., } \Pi \textit{ is in shadow)} \end{cases}$$

The general benefits of the oracle framework has been discussed above. One of these is that of software engineering encapsulation that is provided for free. The oracle can be seen as an “object”, that encapsulates the decision making logic for the shading rule selection. The specific oracle O would be used in the following way in a real rendering system:

```
foreach Polygon
    foreach scan-line span S_p of polygon
        answer = O(S_p, S, τ)
        if answer == ``Gouraud''
            ScanConvGouraud(S_p)
        else if answer == ``Phong''
            ScanConvPhong(S_p)
        else
            ScanConvConst(S_p)
```

This oracle is now well defined, and we wish to see how to take advantage of scene structure to develop an approximation, as well as the oracle cost and error/quality evaluation metrics.

4.1.1. Relevant Aspects of Scene Structure

Due to the limited nature of the specific problem we are able to create an extensive list of structural aspects that affect the choice of shading rule. We will construct the list assuming a polygon primitive, since the scan-line case can be derived.

The purely geometric scene properties that are relevant are the following: polygon area, polygon shape, planarity, vertex normals and their variation, rate of change of the angle between the normal and the viewer, the distance of the polygon from the viewer, and the angle between the polygon normal and the viewing direction. Other aspects that are of importance are texture, surface properties and the local illumination model used, as well as light and shadow information. Each of these properties of the scene contributes to the decision of whether it is possible to use Gouraud shading, without significant degradation in image quality.

The properties above convey information that is relevant to the result of the applied shading rule. As an example, if the polygon is planar, and there is no variation in vertex normal, and in addition the angle between the viewer and the normal changes very little, Gouraud can be used without a problem. Obviously one can construct arbitrary combinations of cases that result in situations with varying degrees of certainty as to whether the shading rule can be substituted.

To allow a decision to be made, one can assign a weight to each of the properties, and construct a weighted sum that is then used as the decision variable (e.g., if the sum is larger than .5 one can use Gouraud). In addition one can rank each aspect so that the more important ones are those examined first (for example whether or not the polygon is in shadow), and apply increasingly complex heuristics to detect easy reject/accept cases.

The determination of the weights is initially a static, precomputation phase, followed by dynamic alteration discussed in Section 3.4. The static determination is based on a careful analysis of the effect each parameter or its variation has on the calculated (mean) intensity. Even for a simple case like this, one has to see how each of the properties listed above affects the final result, and express the effect each has by assigning appropriate weights.

4.1.2. Oracle Cost

The issue of oracle cost is of course of major importance. The case we have here is straightforward, since we have a yes/no decision to make. Suppose we have a sequence $Opt_1, Opt_2, \cdots$ of rendering options ordered by (increasing) costs. In our specific example, the rendering options are none, Gouraud, and Phong. We can also define a hierarchy of possible oracles as follows:

$O^1(S, S, \tau)$ *Choose shading rule for entire scene S*

$O^2(O, S, \tau)$ *Choose shading rule for object O*

$O^3(P, S, \tau)$ *Choose shading rule for polygon P*

$O^4(S_{p,} S, \tau)$ *Choose shading rule for span S_p*

We will also write Opt^i_j to denote that option j is being used at level i of this hierarchy. A number of cost constraints are evident:

$$Cost(Opt^i_{j+1}) > Cost(O^i) + Cost(Opt^i_j),$$

where $Cost(O^i)$ indicates the computational cost of the oracle O^i itself, while $Cost(Opt^i_j)$ is the cost of using rendering option j at level i. For example, if $i = 4$ and $j = 2$, then the above constraint states that the use of Phong shading across a scan-line span must be more expensive than the cost of evaluating the oracle for this primitive together with the cost of employing Gouraud shading for the span. If it is more expensive to use the oracle to determine that one can shade scan-lines with Gouraud than it is to compute directly with Phong the advantage of the whole technique is lost.

In our case the cost of the computation for the oracle is fixed, and the cost of each option is also easily calculable since the cost of the application of illumination model and of the incremental operations is known, given a set number of pixels †.

† Depending on the algorithm used for visibility, the cost comparison between O^1 and O^2 may prove problematic, but we will assume for now that this is not the case.

4.1.3. Error/Quality Assessment and Adaptivity

As discussed in the previous section an important aspect of the oracle approximation process is determining the success of the selection strategy, which is the choice of shading rule in our case. A first order approach would be for the oracle to perform "random tests" in which when it selects Gouraud it also computes some pixels of Phong and computes the real mean error μ and variance σ^2. A second order approach is to identify those parts of the shading rule equations that are most influential in the determination of the intensity, which can be determined by the weights used in the selection process, and compute an approximation of the error metrics, thus saving computation. When the oracle decides that it is performing poorly, it re-examines the weight assignments. By keeping track of selection history, it will be able to determine which weights are incorrect, and modify them so that they are closer to the correct decision.

4.2. An Example of Structural Diffuse Illumination

In this section we will present an augmented, progressive refinement radiosity-based solution to global illumination. Since this a much more involved and complicated problem than that of the shading rule, the discussion is necessarily less detailed and at a more abstract level. We formulate the problem in terms of two main oracles. The first is an approximation to the sample selection oracle O_{SaSel}, treating patches as illumination samples. The approximation uses knowledge of the scene structure to determine patch size. The second is an approximation to oracle O_{IR}, that, given a set of pixels associated to a patch or set of patches, selects a filter to be used.

4.2.1. Approximating the Sample Selection Oracle

The oracle implementation bases its choice of patches and patch size on a two-step selection process. The first step takes into account the effect of viewing parameters, and the second performs a preprocessing step to determine the influence of geometry and shadow boundary information.

As a first step we allow patch sizes to vary, depending on the size of the projection of each object onto the screen. In this fashion, objects that have a small screen projection are subdivided into a smaller number of patches. The area of each of these patches is calculated at subdivision time, since this is necessary in form-factor calculation. It must be noted that initial patch size, even for large objects, is taken to be significantly larger than in a normal radiosity-based approach.

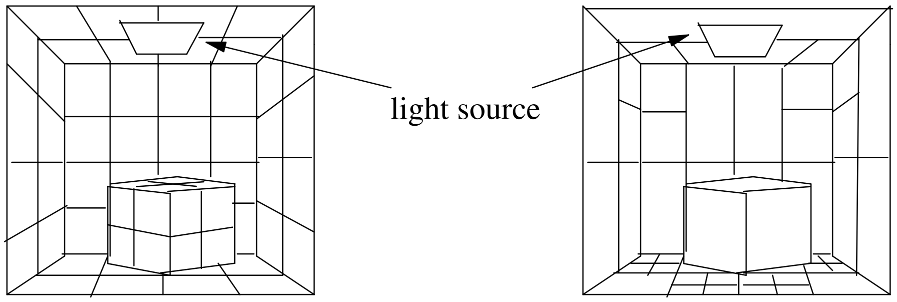
In a second step, shadow boundary computation is performed in the spirit of (Atherton et al. 1978), from light cast from the initial emitters, onto the remaining objects in the scene. Object proximity is considered in these calculations, since diffuse emitters do not produce sharp shadows beyond a certain distance (or solid angle). These boundaries are then utilised to subdivide the patches that they intersect, so that a finer meshing exists along shadow boundaries. Other considerations, such as geometry information, and proximity, especially to light sources, can be used to further refine the meshing approach. All additional subdivision is subject to the projected screen size constraint: i.e., if the projected object is very small, subdivision will not be performed.

4.2.2. Approximating the Reconstruction Oracle

The reconstruction phase implements a simple form of the "tagged sample" concept. For patches that have not undergone additional subdivision, and thus do not have shadow boundaries crossing them, wide area filters are used to effect smooth blurring. An additional criterion in the kernel selection can depend on surface properties such as the existence of texture. Fast approximations based on table lookups or pyramidal techniques can be used (Feibush et al. 1980, Williams 1983). For more highly subdivided patches, blurring will not be attempted, allowing shadow boundaries, and intensity variations due to geometry to be more pronounced than in the standard solution.

4.2.3. Using the Oracles in the Enhanced Algorithm

The enhanced algorithm can now utilise these two oracles as black boxes. Instead of the initial static subdivision of the scene that typically occurs in traditional radiosity-based implementations, the first oracle is used, resulting in the generation of an ‘‘intelligent’’ selection of patches and patch sizes. This set of patches is subsequently used in the traditional progressive refinement solution. The number of patches will have been reduced compared to a straightforward patch subdivision (see Fig. 2). As a result the actual computation time for a converged solution will be significantly reduced, since the hemi-cube operation complexity is dependent on the number of patches in the environment.



Simple Patch-based subdivision — Adaptive subdivision

Fig. 2: Comparison of Adaptive and Simple Subdivision

At the final rendering stage, the filter selection oracle will be used instead of the usual bilinear interpolation (Goral et al. 1984). This will lead to better defined boundaries for shadows and corners, at a fraction of the cost of higher subdivision that would be required to achieve the same result using the traditional approach.

5. Discussion and Future Directions

As we have seen, structure for the rendering process, both in the scene representation and in sampling/reconstruction, is an important factor that can significantly contribute to global illumination algorithms. The discussion of previous work indicated that, perhaps unknowingly, previous authors have capitalised on certain aspects of structure. In this paper we discussed how conscious and increased structure utilisation can benefit previous approaches.

We then proposed a new structure-directed framework based on the concept of oracles. We discussed initial approaches to approximating these ideal procedures. A discussion of their cost followed, together with simple suggestions on quality and error assessment, their adaptive nature and their definition in terms of tolerance factors.

We presented two examples: a detailed case study on how the framework can be of practical use, and an example for a radiosity-based progressive refinement approach that gives an indication of how this method relates to global illumination.

Further research is being pursued departing from existing global illumination techniques, that will take full advantage of the principles outlined in this paper. As such, algorithms developed will not be retrofitted to accommodate a structural approach but will be designed from the outset with it in mind.

To conclude, we are convinced that a structural approach to global illumination is a worthwhile area for future research. By carefully considering information given in scene descriptions, appropriately tuning data representation, and correctly structuring sampling and reconstruction methods in terms of abstractions such as oracles, it is quite possible that global illumination algorithms will see dramatic improvements both in terms of computational speed and storage requirement reduction.

Acknowledgements

The authors gratefully acknowledge the financial support of National Science and Engineering Research Council of Canada, the Information Technology Research Centre of Ontario, and the University of Toronto, for this research and for travel to the workshop. Thanks also to Pierre Poulin for reviewing an early version of this paper, Gavin Miller for fruitful discussions and also Greg Ward and the reviewers for their useful comments that significantly improved the presentation of this work.

References

Hammersly, John M. and Handscomb, David C. (1964) *"Monte Carlo Methods*", Methuen, London .

Ward, Gegory J., Rubinstein, Francis M., and Clear, Robert D. (1988) "A Ray Tracing Solution for Diffuse Interreflection," *ACM Computer Graphics (SIGGRAPH '88 Proceedings)*, vol. 22, no. 4.

Ward, Gregory J. and Rubinstein, Francis M. (1988) "A New Technique for Computer Simulation of Illuminated Spaces," *Journal of the Illuminating Engineering Society*, vol. 17, no. 1, pp. 80-91.

Thomas, David, Netravali, Arun N., and Fox, D. S. (1989) "Anti-aliased Ray-Tracing with Covers," *Computer Graphics Forum*, vol. 8, pp. 325-336.

Goral, Cindy M., Torrance, Kenneth E., Greenberg, Donald P., and Battaile, Bennet (1984) "Modelling The Interaction Of Light Between Diffuse Surfaces," *ACM Computer Graphics (SIGGRAPH '84 Proceedings)*, vol. 18, no. 3.

Cohen, Michael F. and Greenberg, Donald P. (1985) "The Hemi-Cube, A Radiosity Solution For Complex Environments," *ACM Computer Graphics (SIGGRAPH '85 Proceedings)*, vol. 19, no. 3.

Cohen, Michael F., Greenberg, Donald P., Immel, David S., and Brock, P. J. (1986) "An Efficient Radiosity Approach for Realistic Image Synthesis," *IEEE Computer Graphics & Applications*, vol. 6, no. 3.

Wallace, John R., Elmquist, Kells A., and Haines, Eric A. (1989) "A Ray-Tracing Algorithm for Progressive Radiosity," *ACM Computer Graphics (SIGGRAPH '89 Proceedings)*, vol. 23, no. 3.

Cohen, Michael F., Chen, Eric S., Wallace, John R., and Greenberg, Donald P. (1988) "A Progressive Refinement Approach to Fast Radiosity Image Generation," *ACM Computer Graphics (SIGGRAPH '88 Proceedings)*, vol. 22, no. 4.

Chen, Shenchang Eric (1989) "A Progressive Radiosity Method and its Implementation in a Distributed Processing Environment," *Masters of Science Thesis, Cornell University.*

Woo, Andrew and Amanatides, John (1990) "Voxel Occlusion Testing: A Shadow Determination Accelerator for Ray Tracing," *Proceedings of Graphics Interface '90.*

Weghorst, Hank, Hooper, Gary, and Greenberg, Donald P. (1984) "Improved Computational Methods for Ray-Tracing," *ACM Transactions on Graphics*, vol. 3, no. 1.

Williams, Lance (1983) "Pyramidal Parametrics," *ACM Computer Graphics (SIGGRAPH '83 Proceedings)*, vol. 17, no. 3, pp. 1-11.

Houle, Caroline (1991) "Light Source Modelling," *M.Sc. Thesis, Department of Computer Science, University of Toronto*.

Rubin, Steve M. and Whitted, Turner (1980) "A 3-Dimensional Representation for Fast Rendering of Complex Scenes," *ACM Computer Graphics (SIGGRAPH '80 Proceedings)*, vol. 14, no. 3, pp. 110-116.

Sutherland, Ivan E., Sproull, Robert F., and Schumacker, Robert A. (1974) "A Characterization of Ten Hidden-Surface Algorithms," *Communications of the ACM*, vol. 6, no. 1.

Verbeck, Channing P. and Greenberg, Donald P. (1984) "A Comprehensive Light-Source Description for Computer Graphics," *IEEE Computer Graphics & Applications*, vol. 3, no. 7, pp. 66-75.

Atherton, P., Weiler, K., and Greenberg, Donald P. (1978) "Polygon Shadow Generation," *ACM Computer Graphics (SIGGRAPH '78 Proceedings)*, vol. 12, no. 3.

Feibush, Eliot, Levoy, Marc, and Cook, Robert (1980) "Synthetic Texturing Using Digital Filters," *ACM Computer Graphics (SIGGRAPH '80 Proceedings)*, vol. 14, no. 3, pp. 294-301.