Light-Water Interaction using Backward Beam Tracing

Mark Watt

Digital Pictures, London

Abstract

A new two pass approach is presented based on a variation of backward ray tracing - backward beam tracing. Advantages include its capability of rendering complex, hitherto unattainable, specular to diffuse phenomenon and its easy insertion into standard renderers. The algorithm is applied to aspects of the interaction of light with water. Within this context a variety of first generation effects, including shadowing and light scattering, are both rendered and animated. Results taken from the treatment of caustics within classical optics are included as they provide valuable insights into the precise nature of specular to diffuse transfer .

C.R. Categories I.3.3, I.3.5, I.3.7

Keywords : backward ray tracing, caustics, light beam tracing, specular to diffuse transfer.

Introduction

Computer graphics has now reached the stage where all the basic problems of image synthesis - namely, hidden surfaces, shadowing, reflection, basic diffuse and specular shading, etc. - have been solved. The advent of stochastic ray tracing, producing additional refinements such as motion blur, soft shadowing and depth of field extended the repertoire yet further. The development of the radiosity method enabled the subtle effects of diffuse inter-reflections to be modeled and firmly established the distinction between view dependent and view independent image synthesis techniques. The former are driven from the eye whereas the latter suffer no such constraint.

The search for a more complete global illumination model led to the classification [17] of four mechanisms of light transport between two surfaces. They are as follows :

(i) diffuse to diffuse,
(ii) diffuse to specular,
(iii) specular to specular,
(iv) specular to diffuse.

(i) and (iv) are view independent, (ii) and (iii) view dependent. To some extent this classification is an over simplification as it ignores higher order illumination effects that occur when light arrives at a surface via a path involving several interactions with several surfaces where any of the above transport mechanisms may be involved in any order. Also, cases exist where this approximation of energy interchange breaks down completely - such as the lunar surface [15]. Nevertheless to first order it remains a useful concept and shall be used here.

Both radiosity and ray tracing have attempted to solve for global illumination where any number of these binary interactions are involved. Radiosity is recognized as being the most suitable method for solving the diffuse to diffuse component, whilst ray tracing is best suited solving for the diffuse to specular, and specular to specular components. These associations have given images generated by the separate techniques their unique signatures. Radiosity pictures are usually office interiors which tend to be softly lit; whereas ray tracing pictures have as their hallmark shiny objects, usually spheres, exhibiting sharp multiple reflections.

Either method comes up against difficulties when moving into the other's domain. They are :

(a) Radiosity and the specular component.

Radiosity, by attaining equilibrium within a closed environment, must, by necessity, sample and discretize the whole environment. In general the view dependent specular contribution varies rapidly over adjacent pixels and is not adequately captured by interpolating across relatively wide patch vertices. Subdividing the patches to a level such that when projected onto the image plane they are roughly the same size as a pixel would overcome this problem but is overwhelming computationally.

(b) Ray tracing and the diffuse component.

In order for ray tracing to handle the diffuse to diffuse component, many rays would have to be fired from a ray-object intersection in many directions, to many levels of recursion, resulting in significant computational overheads. A further objection to such an approach lies in the fact that the diffuse contribution typically changes slowly over pixels and a naive ray tracer calculating the diffuse component ab initio at each pixel performs much unnecessary, repetitive work.

By taking advantage of the complementary strengths of the two approaches, work over recent years has led to the evolution of hybrid two-pass algorithms involving a view independent radiosity preprocessing step and a view dependent ray tracing post processing step.

At one end of the spectrum lies the solution of the specular component within the view independent step as proposed by Immel [10]. Here the conventional form factor defining the energy transfer between two patches is extended to include bidirectional reflectance by taking into account all directions of one patch against all directions of the other. Unfortunately this method fails to address the problem outlined in (a) above and consequently it takes weeks to produce even the simplest of pictures.

At the other end of the spectrum lies the solution of the diffuse component within the view dependent step as proposed by Kajiya [11] and subsequently Ward [18]. As noted in (b) however, it is simply too expensive to fire huge numbers of rays into the scene - they must be fired intelligently. To this end, Kajiya employed sophisticated sampling and variance reduction techniques that monitor the progress of ray trees, pruning them if the variation amongst the samples is small, and directing them into regions where they make significant contributions to the final value of the pixel. Nevertheless computation time is still very long .

Within these two extremes fall the remainder of the hybrid algorithms including Shao [14], Wallace [17] and Sillion [16].

Specular to diffuse transport

Notable by its absence so far is a discussion of the fourth mechanism of light transport - specular to diffuse transport. This occurs when light reflecting from, or refracting through, one surface, the specular surface, hits a diffuse surface where, by definition, it is emitted equally in all directions. It is fair to say that, to date, this is the most elusive and neglected of all transport mechanisms - traditional radiosity ignores it completely as do almost all ray tracing techniques. It is the purpose of this work to go some way towards redressing the imbalance.

What little attention this mechanism has received is reviewed below. We consider approaches made from the radiosity and the ray tracing camps separately.

(a) Radiosity and the specular to diffuse component

Wallace et al. [17] treated the specular surface as an additional route by which light leaving one diffuse patch may reach another. The method restricts the specular surface to be a perfect flat mirror so that the extension can be handled by the traditional 'Cornell' approach for from factor determination ie. the hemi-cube. If patch A can see patch B in the mirror then the mirror form factor between patch A and patch B 's reflection is constructed and added to the radiosity equation. The construction entails building a 'virtual world' on the other side of the mirror. The method was able to model the effect of light from a table lamp bouncing off a mirror and hitting the table top.

A more general solution was put forward by Sillione [16] who refinded the traditional form factor definition to be a measure of energy transfer between two patches after any number of intermediary specular reflections of refractions. The form factor of a specular patch now becomes a series of form factors - one for each direction. A given form factor in a given direction distributes itself amongst objects that intersect the ray tracing tree fired in that direction. The method was able to model light refracting through a glass sphere and being focussed on a flat surface below.

Both these implementations still fall foul of the fundamental criticism levelled at radiosity techniques as cited in (a) in the introduction. Consequently, the pictorial examples of both methods show only the softest of specular to diffuse effects varying slowly over pixels and both would prove impractical in capturing higher frequency effects examples of which, as we shall see later, abound and are far the more interesting.

(b) Ray tracing and the specular to diffuse component

Work done in this field, although meeting with similar limited practical success, provides us with a valuable theoretical notion - so we shall consider it in some detail. Let us return to the example of specular to diffuse transport as modeled by Wallace [17] and shown in figure 1. A table lamp illuminates a table directly, along LT, and indirectly by light reflecting off the mirror, along LMT.

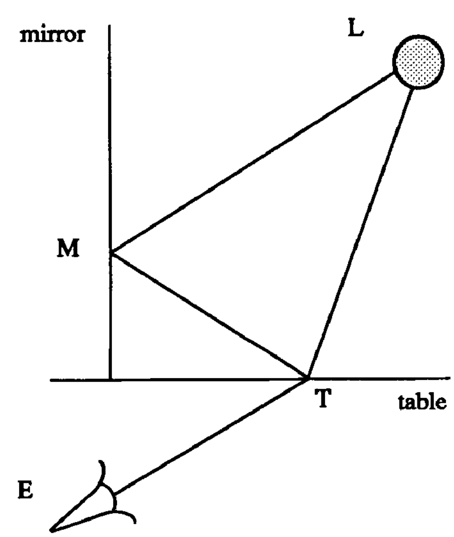

Figure 1.

A standard ray tracer fires rays from the eye through the image plane and into the scene. Suppose one such ray, ET, hits the table at T. Further rays are then spawned at T to determine its illumination. The ray tracer knows that if T can see the light it will contribute to its illumination so a shadow ray TL is sent to the light. Unfortunately, it is completely ignorant of light arriving by the second, indirect route. The only way it can find this second route is by ' accident ' ie. by firing enough rays in enough directions such that one coincides with the direction TM.

Clearly, standard ray tracers, in sampling the whole environment, would have to be made to work very hard to capture the specular to diffuse component and only the most sophisticated of them all [11] has managed to do so.

The problem is fundamental and due to the fact that we are travelling in precisely the reverse direction to that which light is propagating. This applies to any view dependent algorithm. Only light energy transmitted via the reverse path of the ray tracing tree is taken into account. A far more intuitive approach for this transport mechanism at least, and one that this work adopts, would be to move over into the view indeopendent domain and to start from the light following the rays in their direction of propagation as they bounce around the scene - thereby removing our ignorance of indirect illumination at a stroke. Immediately specular to diffuse transport becomes less elusive - our problem above becomes trivial - reducing to the statement that a ray from the light travelling along LM is reflected at M and hits the table at T.

It was just such a consideration of specular to diffuse transfer by Arvo [2] that led to the notion of ' backward ray tracing ' as he coined it , since it is the reverse of conventional ray tracing. Unfortunately a confusion of terminology has evolved as subsequent authors, notably Chattopadhyay [6] and Glassner [7], prefer to label *conventional* ray tracing as backwards since its direction is the reverse of light propagation. For this work we adopt Arvo's historical convention.

Backwards Ray Tracing

Any backward ray tracing technique will be a two pass algorithm - the backward ray tracing from the light forming the view independent step, followed by the forward ray tracing step from the eye. Central to the working of such a strategy will be how information derived during the first pass is communicated to the second or, to borrow a phrase from Arvo, how the rays ' meet in the middle '.

Arvo [2] suggests achieving this via an illumination map, consisting of a grid of data points, which is pasted onto each object in the scene in much the same way as a texture map. A given illumination ray, originating from the light, strikes an object at a point and depositing a certain amount of energy in the surrounding data points, continues on its course, reflecting and refracting through the scene. The first pass consists of ' showering ' the scene with these illumination rays. By ignoring first generation hits , which directly illuminate an object, upon completion of this 'showering' ,the illumination map will provide a measure of the indirect illumination received by the object. During the rendering phase a ray striking the object will pick up a value for this indirect illumination using bilinear interpolation between nearby data points.

The first phase, by using infinitely thin illumination rays to deposit quanta of energy, is point sampling the specular to diffuse energy distribution of the scene. Again, since there is no reason to assume this distribution to be slowly varying, this approach will be highly prone to aliasing problems. This aliasing is compounded further as the discretization of the illumination maps into data points may introduce yet more artifacts. A faithful representation of a specular to diffuse effect of a given frequency, will require constructing an illumination map of greater frequency of data points and, in turn, showering sufficient illumination rays to ensure that hits on the map are many times as dense as these data points.

Other techniques include Zhu ' two-way ray tracing ' [20] and Chattopahyay and Fujimoto ' bi-directional ray tracing ' [6]. If the acid test for shading algorithms is the quality of the image produced - not an unreasonable criterion - neither technique is particularly impressive in that the specular to diffuse effects rendered are, as in the radiosity examples, soft and vague.

It is clear that backward ray tracing as a relative newcomer to computer graphics is still in its infancy. Since the vast bulk of ray tracing methods fall into the forward ray tracing category, there exists a disparity in terms of the effort invested between the two. Much work has been done on bouncing rays around the scene from the eye but little from the light (Heckbert and Hanrahan [8] were the first to mention this asymmetry). We will show that backward ray tracing can produce useful results and suggest it is a powerful idea and a fruitful area for future research.

Light Beam Tracing

We present a two pass algorithm, the first pass based on a variation of backward ray tracing called backward beam tracing, that provides a solution of specular to diffuse transfer in polygonal environments including shadowing effects. The algorithm is closest in spirit to a suggestion made by Heckbert and Hanrahan ' Beam tracing polygonal objects ' [8] and it is from here that it takes its name.

We confine our attention largely to first generation specular to diffuse transfer only but, theoretically, there is no restriction on recursion providing the transport is of nature :

specular -> specular -> ..... ->specular -> diffuse

In this respect it suffers from exactly the same problems as the two mainstream global illumination models in that they too are based on one mechanism of light transfer only.

In the algorithm objects in the environment are separated into specular objects which retransmit the light incident upon them, by reflection or refraction, onto the diffuse objects.

(i) first pass - backward beam tracing

For each polygon of each specular object we construct a light beam by casting rays from each vertex to the light. Next we construct the transmitted light beam, by reflection or refraction with the vertex normals, and sweep this beam throughout the entire scene testing for intersections with diffuse polygons. If an intersection occurs we project the transmitted light beam onto the plane of the diffuse polygon. Shown in figure 2 are two such light beams the first of which becomes divergent after refraction, spreading the light over a larger area than that of the specular polygon; the second, convergent, focuses the light onto the diffuse polygon

This projection forms the *caustic* polygon which is allotted an intensity and a tag signifying which diffuse polygon is responsible for its generation. The caustic polygon is added to the polygonal database as a surface detail polygon ie. a polygon which does not affect the shape of an object - only its shading. This step has the usual advantages of view independence namely it need only be computed once providing the spatial relationships between the objects and lights remain unchanged.

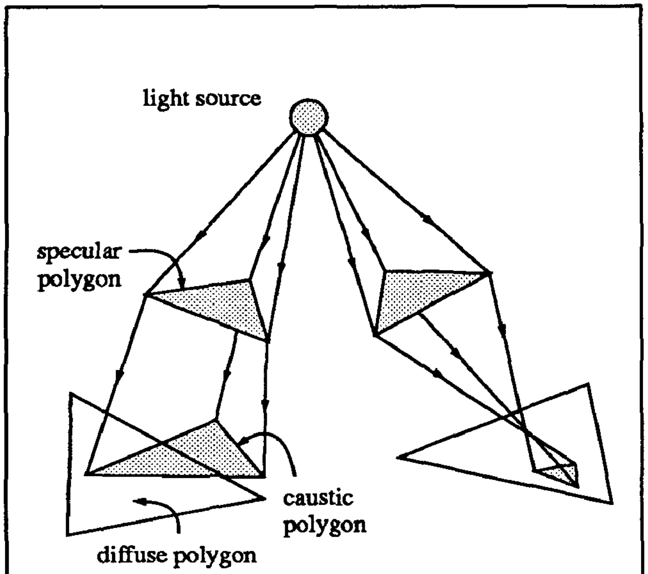

Figure 2. Generation of caustic polygons.

Since, under this scheme, the resolution of the specular polygons dictates the resolution of the transmitted light beam, the size of the specular polygon is critical. A given specular object may be faithfully represented by polygons of a given size. But the resolution required to represent the object may well be inadequate when representing the light this object transmits. In general it was found that the specular polygons had to be much smaller when modeling the light than when modeling the specular object itself.

(ii) second pass - rendering

The rendering phase proceeds largely as it would normally with only one exception. If a diffuse polygon is visible under a pixel we check to see if it has any caustic polygons associated with it. In general a diffuse polygon may intersect more than one light beam and so can have more than one attendant caustic polygon. The intensity of the caustic polygon is simply added to the diffuse component of the final shaded value of the diffuse polygon.

What intensity are we to assign to the caustic polygon ? Let us assume distances between the light and objects are large compared to distances between the objects themselves. If this is not the case an extension to include $1/r^2$ fall off is straightforward . The specular polygon, then, has an intensity, I, incident upon it. The fraction of energy, E, arriving on the specular polygon is the product of I and the area of the polygon ' seen ' by the light, ie. its area projected in the direction of the light L. Referring to figure 3 then, we have:

$$E = I \; N.L \; AreaSpec$$

Where N, the normal associated with the plane of the specular polygon, is usually different from the normals at the individual vertices. Assuming that the specular polygon absorbs no energy, E will thus be the energy incident on the caustic polygon. We can easily include an absorption coefficient into the calculation should this assumption not hold. Providing the specular polygon is small enough such that the variation of distances between points on the specular polygon and points on the caustic polygon is negligible, we can say that this energy will be distributed uniformly over the caustic polygon. The intensity of the caustic polygon Icaustic, is thus :

$$\begin{aligned} Icaustic &= E \,/\, AreaCaus \\ &= I \; N.L \; AreaSpec \,/\, AreaCaus \qquad (I) \end{aligned}$$

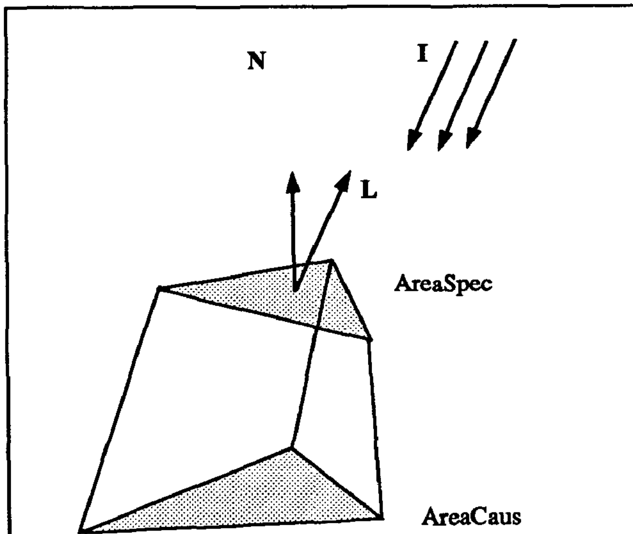

Figure 3. Generation of Icaustic.

This term resembles the form factor of radiosity. This is not surprising since its derivation starts from similar considerations of energy transfer, projected areas etc. Indeed, removing the assumptions made about distances in the derivation would take us even closer to a form factor definition. This suggests that the step of constructing caustic polygons from specular polygons could be construed as constructing special, simplified adaptive form factors that take full advantage of an a priori knowledge of light propagation in the environment. Compare this to the traditional method of building form factors which entails a somewhat arbitrary dicing up of surfaces completely independently of the light distribution. Many radiosity algorithms quote this independence/ignorance of lights as an advantage - while it may be good for the algorithm's generality it certainly does nothing for the rendering of complex images efficiently.

The algorithm has the following advantages over existing techniques for rendering specular to diffuse effects :

(i) ease of implementation - since the caustic polygons translate into surface detail polygons during the second rendering phase, standard depth buffer or ray tracing renderers need only minor modification to implement the algorithm.

(ii) efficiency - above all the algorithm is efficient. As we shall see later, frame times are sufficiently small to enable animated caustic sequences to be produced. Reasons for this efficiency include the following :

The reduction of caustic polygons to surface detail polygons in the second pass removes the need to clip the caustic polygon to its diffuse polygon in the first pass. During the rendering a caustic polygon is only rendered if its diffuse polygon is behind it. In the case of depth buffer renderers, regions where the caustic polygon exists and the diffuse polygon do not, are masked out by the pixel mask of the diffuse polygon. In the case of ray tracers these regions are not even considered since ray-caustic polygon intersections are only tested for after a ray-diffuse polygon intersection occurs.

Optimization techniques developed for ray-object intersections in forward ray tracers, such as space partitioning or bounding volumes, can be adopted when testing for transmitted beam-diffuse polygon intersections during the first pass.

(iii) By replacing rays, which point sample the environment, with beams we avoid aliasing problems associated with the former. By using light beams we are effectively tracing a *bundle* of rays with an intensity varying as the density of rays in the bundle changes through the optical system. As shown in (I) this density is inversely proportional to the bundle's cross section or, in our terminology, Areacaus, the area of the caustic polygon.

(iv) The intricacy of the light pattern produced on the diffuse surface is directly related to the geometric complexity of the specular surface. The greater the variation in curvature of the specular surface the greater the directions over which incident light is dispersed. By driving the method from the specular surface, regions over which the curvature varies rapidly can be sampled more intensively . Such adaptive sampling enables sharply varying specular to diffuse phenomenon to be represented efficiently thereby, extending the range over existing techniques which, hitherto, as we have seen, have confined themselves to rendering only the vaguest of effects.

We now turn our attention to the term caustic . Derived from the fact that a lens can focus sunlight to burn a hole in a surface, it is a term taken from classical optics and is completely synonymous and interchangeable with our term specular to diffuse transfer. We include a discussion of the treatment of caustics within classical optics here [5], since it provides us with valuable insights into the precise nature of specular to diffuse transfer.

Caustics

We start with some definitions. If a curve exists such that it is tangent to family of curves but is not itself a member of that family, then it is the *envelope* of that family. Let the family of curves be those rays transmitted from a specular surface then

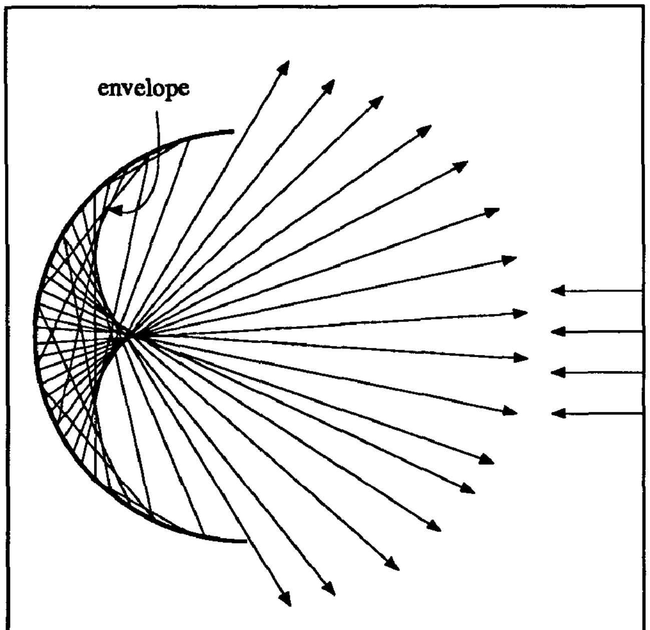

Figure 4. Reflected caustic in spherical mirror.

the caustic surface is the envelope of the transmitted rays.

Figure 4 shows an example of a caustic surface formed after reflection, created when light travelling parallel to the axis of revolution of a spherical mirror is reflected by it. The caustic can be shown to be an epicycloid, the cusp of which is at the principal focus of the mirror. An everyday example is the bright line on the surface of a cup filled with liquid - due to intersections between rays of light reflected from the cylindrical wall of the cup and the liquid.

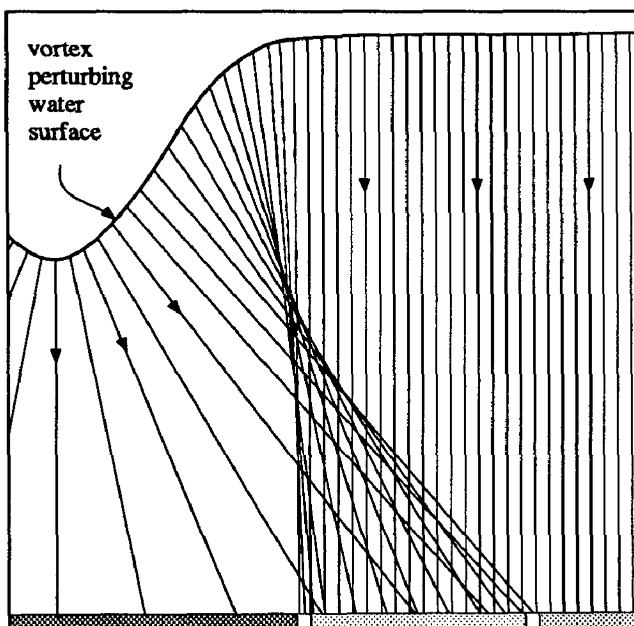

Figure 5. Refracted caustic through water vortex.

Figure 5 shows a caustic formed by refraction in water. The surface of the water is deformed by a vortex shed from an object, eg. an oar, moving through the water.. A section through the caustic, such as would be seen at he bottom of a tank, consists of two concentric circles bounding a bright ring. A radial slice of this section in relation to the caustic surface is shown diagrammatically at the foot of the figure. The circles on the edges of the ring, caused by intersection with the caustic surface, are brightest; the interior of the ring where light is refracted away from the axis of the vortex is darkest. Studies of such patterns have enabled Berry and Hajnal in 'The shadows of floating objects and dissipating vortices' [4] to predict analytically the shape of the vortex to be a blend of a parabolic core surrounded by a hyperbolic surface.

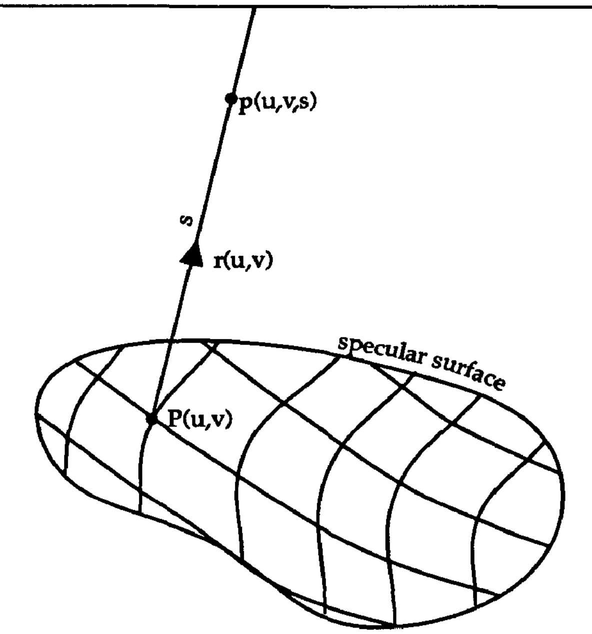

Figure 6.

Let us look at the caustic surface more closely. Consider the specular surface described by curvilinear coordinates (u,v) as shown in figure 6. Each point **P**(u,v) on the surface has associated with it a transmitted ray in the direction of **r**(u,v). Any point p on this ray is given by :

$$\mathbf{p}(u,v,s) = \mathbf{P}(u,v) + s\mathbf{r}(u,v)$$

where s is the length along **Pp**. Foci occur where rays emanating from the surface intersect, or to translate into terms of differential geometry, assuming the surface to be differentiable, where points on these rays are separated by a distance that is of second or higher order ie.

$$\mathbf{p}(u,v,s) = \mathbf{p}(u+du,v+dv,s+ds)$$

which, expanding out, gives :

$$\mathbf{p}_u du + \mathbf{p}_v dv + \mathbf{p}_s ds = 0$$

where $\mathbf{p}_u$,$\mathbf{p}_v$ and $\mathbf{p}_s$ are the partial derivatives with respect to u,v, and s. This implies that the three vectors $\mathbf{p}_u$,$\mathbf{p}_v$ and $\mathbf{p}_s$ are coplanar which is equivalent to saying that their scalar triple product vanishes ie.

$$[\mathbf{p}_u, \mathbf{p}_v, \mathbf{p}_s] = 0$$

Now p is linear in s, and $\mathbf{p}_s$ is independent of s, so the above reduces to quadratic in s. This means that there are two foci on each light ray. As u and v vary over the specular surface the foci traces out a corresponding caustic surface which comprises of two separate caustic sheets.

This is a remarkable result, although the transmitted light may create structures rivaling a pinnacled cathedral in complexity, each light ray will be tangent to the two sheets.

These caustic sheets meet at cusps or folds, which are actually sections of three-dimensional catastrophes, and are the proper study of catastrophe optics [3]. Consider figure 4, one of these sheets is a trumpet shaped surface of revolution, whose cross section is an epicycloid, the other degenerates to a line segment lying on the axis of revolution. In the case of figure 5 the caustic sheets consist of two roughly cylindrical sheets joined at a cusp edged ring.

Returning to our algorithm, the diffuse polygon can be interpreted as sampling slices of this partially concealed structure. The intersection of the plane of the diffuse polygon with this structure provides us with a cross sectional view of it - a picture made up of caustic polygons. This picture will depend largely on the orientation and distance between the diffuse polygons and the caustic surface.

If a diffuse polygon is close to one of the caustic sheets and more or less tangential to it , it will appear approximately uniformly lit - the intensity increasing the closer the diffuse polygon gets to the sheet. If the diffuse polygon is more or less normal to a caustic sheet the variation of light intensities across it will increase the closer it gets to the sheet - until it actually intersects whereupon it will have a bright line running through it. The farther the diffuse polygon gets from the caustic surface the more negligible and uniform the effect becomes as the light rays become more dispersed ie. as the specular to diffuse effect tends to diffuse to diffuse. The latter consideration provides us with a bound beyond which specular to diffuse effects can be ignored.

For more details on caustics and their treatment within computer graphics, including an expanded version of this paper, the reader is referred to [19].

Shadowing

The algorithm also solves for shadowing effects on the diffuse surface within the context of first generation specular to diffuse transfer. There are two distinct cases depending on whether the shadowing object is :

(i) between the light and the specular surface, or
(ii) between the specular surface and the diffuse surface.

These cases have to be treated separately, but both produce shadowing effects on the diffuse surface that differ from traditional computer graphic shadows. In case (i) a straight silhouette edge may produce a shadow on the diffuse surface whose edge is not, in general, straight, as the light gets 'bent' at the specular surface. In case (ii) the shadowing object, embedded in this region of 'bent' light, occludes light rays that are travelling in directions other than the traditional direction, ie. directly from the light.

Case (i) has to be dealt with in the view independent phase. Each ray of the light beam is tested for hits with objects between the specular polygon and the light. If a specular polygon is completely occluded it is passed over. If it is partially obscured we recursively subdivide the specular polygon - processing the smaller ones that can see the light as normal. The subdivision stops when the area of the specular polygon falls below a predefined level. An alternative approach here would be to clip out the silhouette of the obstruction from the beam cross section and to continue to process the remainder similar to Heckbert [8]. We adopt the recursive subdivision strategy - wary of fragmented, non convex polygons that can be produced by recursively clipping.

Case (ii) can be handled in either the view independent or view dependent phase. If it is solved in the former we proceed largely as in case (i) except each ray of the transmitted light beam is tested for hits with objects between the specular surface and the diffuse surface. To solve in the view dependent phase we carry over the rays of the transmitted light beam from the view independent phase as the normals of caustic polygon. A caustic polygon underneath a pixel can only make a contribution to that pixel then, if the rays fired in the direction of its normals hit the specular surface before they hit anything else.

Light - Water Interaction

As stated previously, we confine our application to model a specific subset of first generation specular to diffuse transfer namely the simulation of light - water interaction where the water surface acts as the specular surface. Our principal justification lies in the fact that no computer graphics model of these phenomena has previously been rigorously presented and the results achieved are realistic in terms of rendering, animation and demands made upon the computer - thus extending the class of natural phenomena capable of being effectively simulated.

Strong sunlight incident on water that is gently perturbed, by wind say, will produce familiar sinuous shifting patterns of light on objects beneath the water's surface as they intersect the caustic. Such patterns were studied and painted by Hockney eg [9]. Figure 7 shows two frames taken from an animated sequence showing these effects as viewed underwater when a shaft of light, coming through a window say, is incident on the surface of an indoor pool. These images were generated using an enhanced depth buffer renderer. Note, as predicted in the section on caustics, the sides of the pool, which are roughly tangential to the caustic, are more uniformly lit than the floor, which being roughly normal to the caustic contains the greater variation of light intensities.

As the pattern is driven directly by the water surface, animating the pattern entails animating the water surface.The water surface consists of a polygonal mesh made up of triangles displaced by a height field.The height field is a supersition of a distribution of sine waves of varying frequency and amplitude as used by Max [12] to which the reader is referred since the details of the surface modeling are not essential to the application - more sophisticated water models may easily be substituted. Since, under this model, the wave speed is a function of its frequency and gravity - animating the water surface consists of deciding upon an appropriate frequency/amplitude distribution and fine tuning the time interval over consecutive frames to achieve the desired effect.

Under certain conditions particles or impurities in the water that are within the transmitted beams may become visible by scattering the light - enabling us in effect to see the transmitted light beam as opposed to cross sections through it. As the water surface changes shape regions that previously dispersed the light may now focus it and vice versa - causing the beams themselves to change shape accordingly.

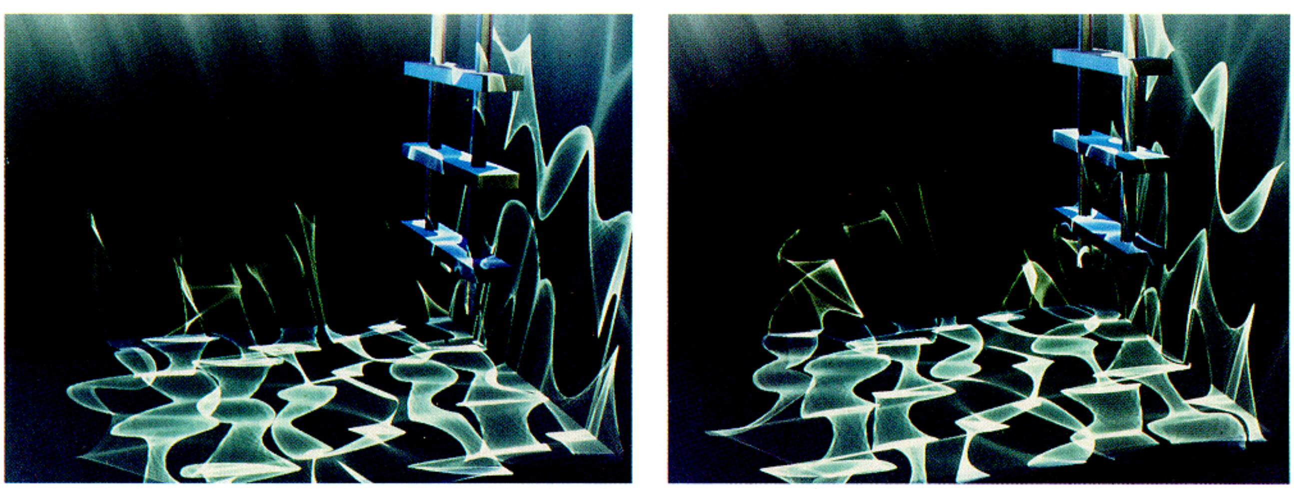

Figure 7. Two frames from an animated underwater caustic sequence.

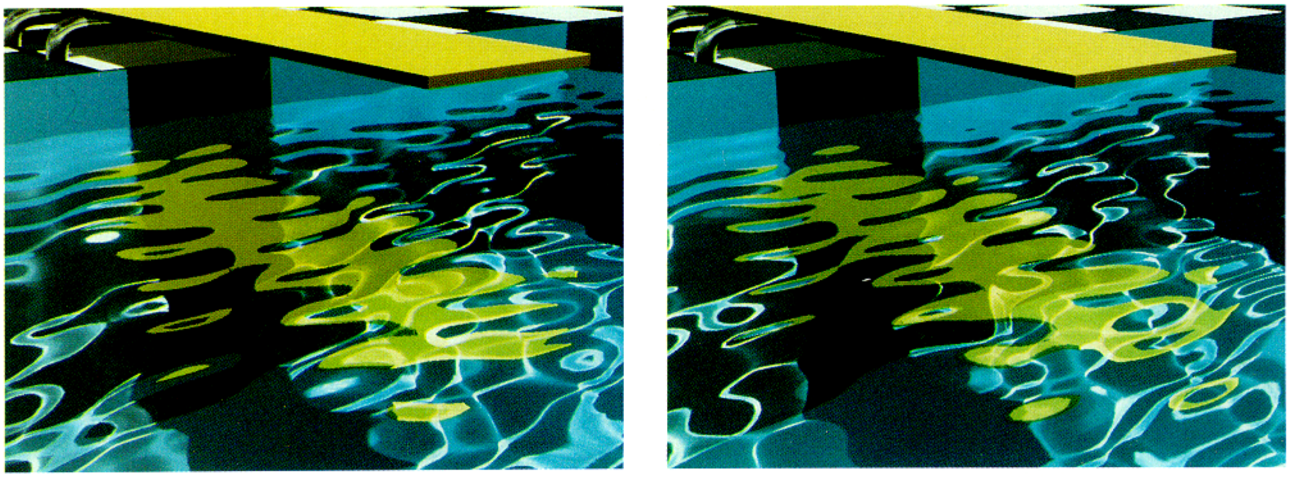

Figure 8. Two frames from an animated water caustic sequence seen through the water.

If the triangles in the polygonal mesh are sufficiently small, the variation of the refracted rays along an edge of a triangle can be ignored. This approximation enables us to represent the pencil beams as polygonal illumination volumes. These beams are shown in figure 7, and are rendered using a modified version of the light volume rendering technique as proposed by Nishita [13]. As in that paper, assuming uniform particle density, we integrate intensities of the scattered light along segments of rays that lie within the illumination volumes. We also include an additional term in the integration to account for the concentration, or dispersion, of light within the beam.

Figure 8 shows two frames taken from an animated sequence of an outdoor swimming pool seen from above. These images were generated using a standard ray tracer. Strictly speaking, the additional refraction of rays from the diffuse surface, through the water, to the eye means this is an example of specular to diffuse to specular transport. The animation is particularly effective since it is really the combined effect of two separate animations - one imposed on the other. The underwater animating pattern is seen through an animating water surface. These figures also show the shadow cast by the diving board on the pool - an example of case (i) in the section on shadowing. Figure 9 shows the shadow with the water surface taken away, showing that straight silhouette edges can produce shadows with curved edges.

Since the water surface changes frame for frame, the view independent phase consisting of generating the caustic polygons must be calculated likewise. Since we are considering first generation effects only, it follows that a transmitted beam can only contribute to an image if it is itself within the image. By taking advantage of this, significant improvements in efficiency were obtained by first testing for intersections between the transmitted beam and the viewing frustrum of the image, rejecting those that fall outside.

Finally, we consider a rather subtle example of light - water interaction caused by a meniscus. A pencil partially immersed in water under an overhead light will produce a bifurcated shadow separated by a whitened gap which has been termed the 'shadow-sausage effect' [1]. Both the dry and submerged

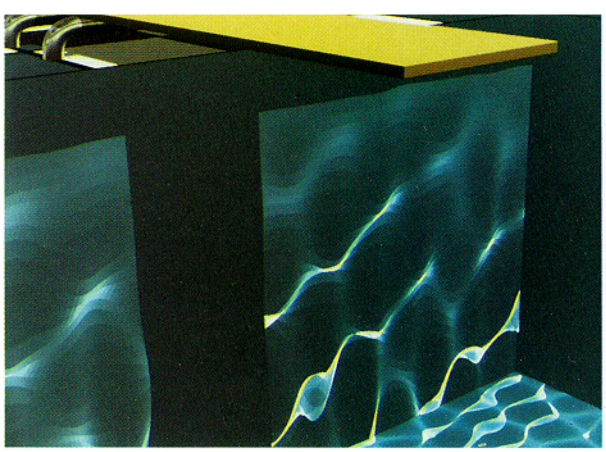

Figure 9. Pool caustics with water removed.

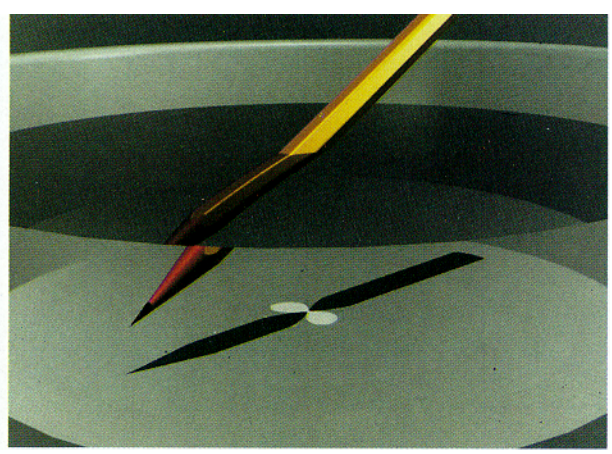

Figure 10. Caustics from a meniscus.

parts of the pencil cast typical shadows (cases (i) and (ii) of the shadowing section respectively) but the meniscus generates caustics which concentrate light in the gap area thereby washing out the shadow one would expect to find.

Figure 10 shows a good approximation of this phenomenon with the front section of the bowl removed for clarity. The meniscus is represented as an elliptical annular region over which the normals are varied from the unperturbed vertical state on the exterior boundary to the contact angle with the shaft at the interior. This contact angle changes round the shaft being greatest where the pencil forms an acute angle with the water and least where the angle is most obtuse. Both the shadows and caustics were calculated at the view independent phase by adaptively subdividing the water surface - performing more work around shadow edges and the meniscus.

Conclusions and future directions

A method of computing specular to diffuse transfer using backward beam tracing has been presented. It is shown to have advantages over radiosity and ray tracing approaches. The use of beams originating from the light ensures adequate and efficient adaptive sampling. Its application to the interaction of light with water is particularly effective.

Extensions include applying the algorithm to model other instances of specular to diffuse transfer and removing the restriction to first generation effects by recursively tracing the beams. This would enable the modeling of, say, the caustics produced on a tablecloth when light is refracted through a wine glass. More generally, it is hoped that this work draws attention to the pertinence of algorithms driven from the eye, for this transport mechanism at least, and it is suggested that work could be done exploiting this notion further.

Acknowledgements

Thanks to Dr. Alan Watt of Sheffield University for background research, reviewing early drafts and continued moral support. Thanks also to David Lomax of Digital Pictures for help in the initial stages of the development of the algorithm and to Digital Pictures for their support.

References

[1] Adler, Cyrus. Shadow-sausage effect. American journal of physics, Vol. 35, No. 8, (August 1967), 774-776.
[2] Arvo, James. Backward ray tracing. Developments in ray tracing, siggraph course notes, Vol. 12 . (August 1986).
[3] Berry & Upstill. Catastrophe optics : morphologies of caustics and their diffraction patterns. Progress in optics, Emil Wolf ed., Elsevier North-Holland, Vol. 18, (1980).
[4] Berry & Hajnal. The shadows of floating objects and dissipating vortices. Optica Acta, Vol. 30, No. 1, (January 1983), 23-40.
[5] Born & Wolf. Principles of optics, Pergamon Press, (1964).
[6] Chattopadhyay & Fujimoto. Bi-directional ray tracing. Computer graphics : proceedings of CG International 1987, Tosiyasu Kunii ed., Springer Verlag, Tokyo, (1987), 335-343.
[7] Glassner, Andrew S. An overview of ray tracing. An introduction to ray tracing, Andrew Glassner ed., Academic Press, (1989), 1-31.
[8] Heckbert & Hanrahan. Beam tracing polygonal objects. Computer graphics, Vol. 18, No. 3, (July 1984), 119-127.
[9] Hockney, David. Portrait of an artist (pool with two figures). [painting] (1971).
[10] Immel, Cohen, & Greenberg. A radiosity method for non-diffuse environments. Computer graphics, Vol. 20, No. 4, (August 1986), 133-142.
[11] Kajiya, James T. The rendering equation. Computer graphics, Vol. 20, No. 4, (August 1986), 143-150.
[12] Max, Nelson L. Vectorized procedural models for natural terrain : waves and islands in the sunset. Computer graphics, Vol. 15, No. 3, (August 1981), 317-324.
[13] Nishita, Miyawaki & Nakamae. A shading model for atmospheric scattering considering luminous intensity distribution of light sources. Computer graphics, Vol. 21, No. 4, (July 1987), 303-308.
[14] Shao, Peng, & Liang. A new radiosity approach by procedural refinements for realistic image synthesis. Computer graphics, Vol. 22, No. 4, (August 1988), 93-101.
[15] Siegel & Howell. Thermal radiation heat transfer, second edition, Mcgraw-Hill, (1981), 142-145.
[16] Sillione & Puech. A general two pass method integrating specular and diffuse reflection. Computer graphics, Vol. 23, No. 3, (July 1989), 335-344.
[17] Wallace, Cohen, & Greenberg. A two-pass solution to the rendering equation : a synthesis of ray-tracing and radiosity methods. Computer graphics, Vol. 21, No. 4, (July 1987), 311-320.
[18] Ward, Rubinstein, & Clear. A ray tracing solution for diffuse interreflection. Computer graphics, Vol. 22, No. 4, (August 1988), 85-92.
[19] Watt, Alan & Watt, Mark. Advanced rendering and animation techniques : theory and practice. Addison - Wesley. To be published Autumn 1990.
[20] Zhu, Peng & Liang. PERIS : a programming environment for realistic image synthesis, computers and graphics, Vol. 12, No. 3/4, (1988), 299-307.