Depicting Procedural Caustics in Single Images

Diego Gutierrez[1*]   Jorge Lopez-Moreno[1]   Jorge Fandos[1]
Francisco J. Seron[1]   Maria P. Sanchez[1]   Erik Reinhard[2]

[1]Universidad de Zaragoza
[2]University of Bristol

Abstract

We present a powerful technique to simulate and approximate caustics in images. Our algorithm is designed to produce good results without the need to painstakingly paint over pixels. The ability to edit global illumination through image processing allows interaction with images at a level which has not yet been demonstrated, and significantly augments and extends current image-based material editing approaches. We show by means of a set of psychophysical experiments that the resulting imagery is visually plausible and on par with photon mapping, albeit without the need for hand-modeling the underlying geometry.

CR Categories: I.3.7 [Computing Methodologies]: Computer Graphics—3D Graphics; I.4.10 [Computing Methodologies]: Image Processing and Computer Vision—Image Representation

Keywords: Image-based material editing, high dynamic range imaging, image processing

1 Introduction

It is a well-known fact that the human visual system is not a simple linear light meter. As a side-effect, this means that in graphics applications we can sometimes get away with imperfect simulations. The challenge is to understand what type of inaccuracies tend to go unnoticed, and which ones are easily spotted. We are interested in extending the set of tools available to artists to effect high level changes in single images, at much reduced labor costs, compared with painstakingly painting over all pixels. We have already seen very interesting advances in this field, such as retexturing objects with arbitrary textures [Fang and Hart 2004; Zelinka et al. 2005; Fang and Hart 2006], creating translucent materials or objects rerendered with arbitrary BRDFs [Khan et al. 2006], or image editing in general [Oh et al. 2001]. We focus in this paper on altering light transport on the basis of a single image which, to our knowledge, has not been attempted before.

We specifically consider the effect some extreme material edits have on their environment and on human visual perception. In particular, changing an object to transparent during an image edit would have an effect on light transport: nearby diffuse surfaces would exhibit caustics. While their exact calculation is expensive, several approaches exist to approximate the solution and obtain faster frame
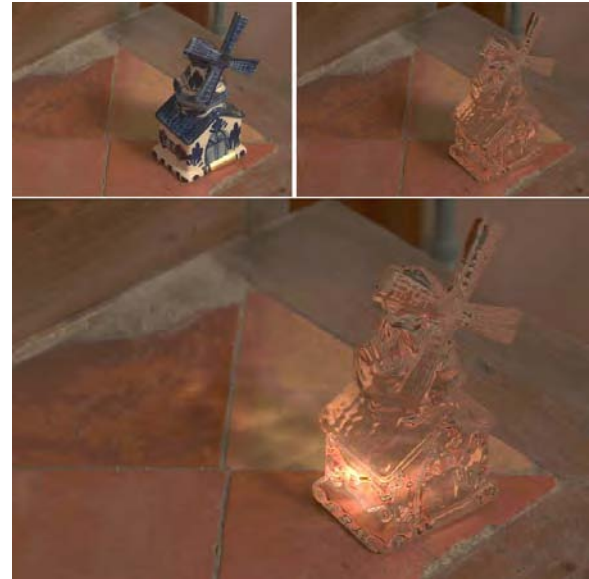
*Contact: diegog@unizar.es

Figure 1: *Example of light transport editing. Top left, original image. Top right, transparent mill following the approach in [Khan et al. 2006]. Notice the absence of caustics. Bottom: final result, with caustics added with our algorithm.*

rates, usually taking advantage of the GPU [Szirmay-Kalos et al. 2005; Shah et al. 2006; Wyman 2005; Wyman 2007]. All existing approaches, however, build a caustics map in 3D space, where the geometry of the objects and the position of the light sources are known. Such caustics maps are generally computed in three distinct steps [Wyman 2008]. In the first step, photons are emitted from the light source, passed through transparent objects, and deposited onto non-transparent surfaces. The second step then uses these photon locations as point primitives, rendering them into the caustic map. The third step projects the caustic map onto the scene. Several different variations have been proposed, including minimizing the number of photons [Szirmay-Kalos et al. 2005], efficient schemes to collect photons in a caustic map [Wyman and Davis 2006], or computing caustics for each object, rather than for each light source [Wei and Kaihuai 2007]. Various techniques which improve quality and performance are also known [Kruger et al. 2006; Wyman and Dachsbacher 2008; Wyman 2008].

In this work we limit ourselves to the more difficult case of single-image inputs. To effectively simulate plausible caustics, the challenge lies in the fact that 3D shape will have to be estimated from the image itself, an inherently under-constrained problem. While multi-camera and video-based solutions would enable us to extract depth more accurately, we envisage our algorithms to find utility in image editing programs such as Photoshop[TM].

To account for the reduced accuracy with which we can estimate the geometry of the environment depicted in the image, we rely heavily on the limitations of human visual perception. By means
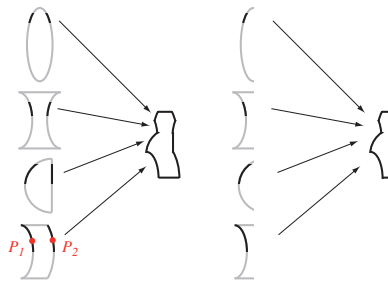
Figure 2: *Left: a simple object constructed from thick lens segments. Right: our thin lens simplification.*

Figure 3: *Perfect symmetry of a theoretical thin lens (left) causes light to converge at the focal point, where a diffuse surface is placed. If the lens were replaced with an arbitrary surface (right), the residual symmetry in the line of interest will contribute to a caustic at the same focal point.*

of a psychophysical study, we show that while humans are adept at detecting caustics, they are very inaccurate at predicting their shape. We therefore follow the rationale that perceptually plausible rather than physically accurate solutions are both desired and sufficient in our case.

The contributions of this paper are as follows. First, we introduce a novel algorithm that can produce light transport edits on a single image, in the form of caustics. We show that for simple geometric configurations the caustics obtained with our algorithm are perceptually equivalent to the physically correct solution. Second, with the aid of psychophysics we show that for more complex objects our algorithm produces caustics that are perceived as perceptually equivalent to ground-truth, photon-mapped caustics. Third, we demonstrate that our caustics are on par with output produced by professional artists, but at a fraction of the time.

In the following, we outline the reasoning behind our approach in Section 2. Our algorithm is then described in Section 3, with results shown and validated in Sections 4 and 5. Conclusions are drawn in Section 6.

2　Motivation

Let us consider a homogeneous transparent object, having a constant index of refraction. Since light propagation at our scale of interest is rectilinear, the occurrence of caustics is determined by the shape of the refracting geometry and the placement of light sources. A narrow beam of rays may enter and exit a transparent volume at points P_1 and P_2, causing refraction according to Snell's law.

Assuming that the dielectric boundaries at entry and exit points (P_1 and P_2) are locally smooth, we may view this pair of surface areas to constitute a small segment of a thick lens. Dependent on the orientation of the surface normals at P_1 and P_2, the lens segment will be either converging or diverging according to a limited number of configurations[1].

Similarly, each pair of surface points on the transparent object forms a separate segment of a thick lens. If the local curvature around surface points is consistent with the global curvature, then all surface points form part of the *same* thick lens, resulting in a very simple caustic (see the real sphere in Figure 5). In the limit the global curvature is identical to that of a thick lens.

Conversely, with increasing complexity of surface curvature, the object will cease to resemble a single lens, but can be thought of as a collection of segments belonging to a set of different thick lenses (Figure 2, left). The number of thick lenses that together

would create the same caustic as the object itself, is indicative of the complexity of the caustic. However, we treat here a heavily under-constrained problem, with only the approximate shape of the camera-facing surface of the object available to us (Section 3.1). As a consequence, we have no knowledge of the back-facing surface. Nevertheless, Khan et al [2006] showed that this has little influence on the identification as a transparent object. We assume that this result extends to caustic rendering (an assumption further backed by our psychophysical analysis in Section 5), and therefore ignore the backface in preference of analyzing the frontface of the object only. Thus, we simplify our thick lens approach and interpret the recovered surface as a collection of *thin* lens segments, which refract incoming light and thus generate caustics (Figure 2, right).

A convex thin lens is circularly symmetric, which gives rise to light being focused at a single point, as shown in Figure 3 (left). If the symmetry were broken, for instance by replacing the thin lens with an arbitrary surface, then the amount of residual symmetry would determine how much light is focused along the line of interest, shown in Figure 3 (right), while the remainder of the light diverges into different directions. This is similar to how photons would be refracted by the surface, distributing their energy along the line of interest; in a photon-mapping approach, caustics would then be obtained by estimating radiance. In our method, we obtain a map representing the caustic pattern that would be cast by an object by computing the amount of symmetry present for each point of that object.

Ideally, we would like to detect symmetry with respect to the position of the light source. However, with only one image at our disposal, we are limited to detecting the degree of symmetry with respect to the viewpoint. For a spherically symmetric object our approach will therefore be accurate, while for asymmetric objects the physical error could be large. However, we demonstrate in Section 5 that perceptual equivalence can be maintained even for large discrepancies between the camera and the light positions. We speculate that this is due in part to humans' inability to predict the shape of both caustics and light directions [te Pas and Pont 2005].

Various techniques exist to detect symmetry in images. Morphological approaches such as median-axis transformation or thinning can only be applied to binary images, and the outlines of the object usually need to be smoothed. Intensity gradients tend to be sensitive to contrast in addition to geometry (see Tyler [1996] for a review). We are interested in finding a robust measure which requires no previous knowledge or pre-processing of the image. We find such measure in the frequency domain, where local geometric symmetries can be found in an image by analyzing its phase infor-

---

[1]The three possible converging lenses are biconvex, plano-convex and concave-convex; the three possible diverging lenses are biconcave, plano-concave and convex-concave [Born and Wolf 1999].
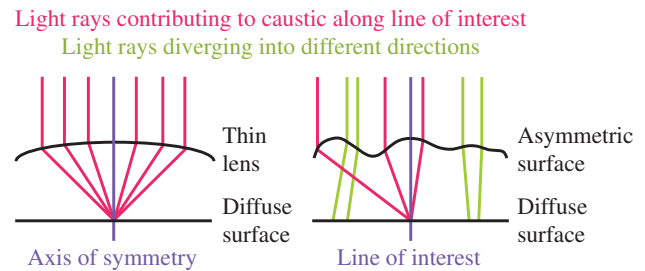
mation [Kovesi 1997; Wu and Yang 2005; Xiao et al. 2005].

Phase symmetry appears to play a role in human vision, which perceives features at points where the phase information is highly ordered [Morrone and Burr 1988; Wichmann et al. 2006], potentially pre-attentatively enhancing the recognition and reconstruction of shapes and objects [Wagemans 1995; Zabrodsky 1993]. Phase symmetry is also used in computer applications ranging from segmentation [Rosenfeld 1986] and feature detection [Kovesi 1996; Yuan and Shi 2005] to image understanding [Openheim and Lim 1981; Piotrowski and Campbell 1982]. On this basis, we argue that phase symmetry may help simulate plausible caustics. The results of our psychophysics tests in Section 5 confirm that this is a viable approach.

3 Simulating Caustics

The problem of adding a caustic to an image can be split into several stages. First, the image is preprocessed to obtain a depth map, serving as a rough representation of the object's geometry. Second, the recovered geometry is analyzed to establish likely caustic patterns that such an object may cast. As previously mentioned, this analysis takes the form of symmetry detection, for which we employ an algorithm that works in frequency space and makes minimal assumptions on its input. Finally, the luminance channel of the image is varied according to the projected caustic patterns. These steps are discussed in the following sub-sections.

3.1 Depth Recovery

Given that global illumination is an inherently three-dimensional process, we must first approximate the 3D object depicted in the image. We rely on the depth-map recovery algorithm by Khan et al [2006]. Depth recovery starts by applying a bilateral filter [Tomasi and Manduchi 1998] to the luminance values of the object's pixels, obtaining the signal $D(x, y)$. This signal is then reshaped to produce the final depth values [Khan et al. 2006].

This approach is based on the idea of "dark-is-deep" which can be seen as one (of possibly several) components of human depth perception [Langer and Bülthoff 2000]. We demonstrate here that it can also be used to produce procedural, perceptually-plausible caustics, relying on two key insights. First, we will produce a caustic from the perspective of the view-point, given that this is the only view available from a single image. While physically inaccurate, statistical symmetries of the transparent object ensure that for our purposes, in most cases this is a reasonable approximation. Second, with this approach, the depth map is both created and used from the same perspective, so that systematic errors introduced by the depth extraction algorithm do not become perceptually distracting.

3.2 Phase Symmetry

To detect symmetries in the recovered depth map, we follow the approach of Kovesi [Kovesi 1996; Kovesi 1997], which has the desirable property that no assumptions on the input are required. However, while Kovesi uses the intensity values of the image as input, thus providing a low-level view of symmetry, we use the depth map instead. This allows us to identify higher level structures based on the recovered geometry. The phase of the depth map at each location is obtained by decomposing it into its different frequency components: we convolve it by even-symmetric (sine) and odd-symmetric (cosine) wavelet filters operating at different scales. We use log Gabor filters, which have the desirable property of having a Gaussian transfer function on the logarithmic frequency scale, consistent with the characteristics of our visual system. Symmetry

Figure 4: *From left to right: segmented mill from Figure 1, recovered depth map [Khan et al. 2006] and two maps with 1 and 20 orientations respectively.*

appears as large absolute values of the even-symmetric filter and small absolute values of the odd-symmetric filter [Kovesi 1997]. A weighted average combines information over multiple scales n and multiple orientations θ_i, yielding the following symmetry map $S(x, y)$:

$$S(x, y) = \frac{\sum_i \sum_n \lfloor A_{n,\theta_i}(x, y)\, B - T_{\theta_i} \rfloor}{\sum_i \sum_n A_{n,\theta_i}(x, y) + \epsilon} \quad (1)$$

$$B = |\cos(\Theta_{n,\theta_i}(x, y))| - |\sin(\Theta_{n,\theta_i}(x, y))| \quad (2)$$

where A and Θ represent amplitude and phase respectively and T is an estimate of the signal noise. Details of the implementation are provided in the appendix.

The two parameters in this equation are the angular interval between filter orientations θ_i (which defines the number of directions d where symmetry is searched for) and the angular spread of each filter (which is a Gaussian with respect to the polar angle around the center). Ideally, we seek the minimal necessary angular overlap to achieve approximately even spectral coverage [Kovesi 1999]; angular overlap is given by the ratio of the angular interval between filter orientations and the standard deviation of the angular Gaussian function used to construct filters in the frequency plane θ/σ_θ. Our experience indicates that good results are achieved with $\theta/\sigma_\theta = 1.2$, which is the value used for all the images in the paper. The number of directions d varies between 1 and 20 (see Table 4), and is the only user-defined parameter of the symmetry detection algorithm. Direction $d = 1$ is defined as the direction yielding the highest symmetry for a given object, for which an initial search is performed at one-degree increments over the full angular space, a process that takes only a few seconds. Successive directions specified by the user are then defined according to this reference.

Intuitively, increasing the number of search directions will create a progressively more complex pattern, given that more symmetries will be detected, thus yielding more complex combined patterns. The degree to which this happens depends on the geometrical complexity of the object. Very simple objects like the sphere in Figure 5 are relatively invariant to changes in d, but the resulting caustics are very similar to the physically-correct ones. The influence of d on more complex objects will be analyzed in Section 4.

3.3 Luminance Adjustment

To apply the caustic map $S(x, y)$, we first obtain its projection $S'(x, y)$ onto a user-defined quadrilateral projection area. This is achieved by means of a simple perspective transform. In general, shadows cast by the opaque object provide a reasonable first indi-
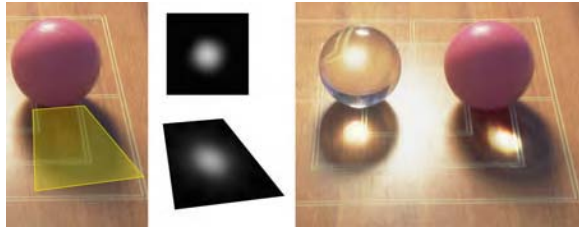
Figure 5: *From left to right: Detail of the original picture, with user-defined projection area. Original focused caustic, and its projected version. Final result, shown next to a real transparent sphere for comparison purposes.*

Object	Maps	d_1	d_2	Object	Maps	d_1	d_2
Mill	2	1	20	Phone	2	1	4
Can	1	2		Sphere	1	4	
Horse	1	4		Skull	2	4	20
Elephant	1	20		Vertebrae	2	4	20
Vase	2	1	12	Dolphin	2	4	20
Doll	2	1	12	Bull	2	4	20
Car	2	1	12				

Table 1: Number of caustics maps and directions d for the images in the paper.

cator of a suitable quadrilateral projection region (see Figure 5, left and middle).

By analysing the silhouette of the shadow, in combination with the silhouette of the shadow-casting object, it may be possible to infer the orientation of the underlying plane. However, we are not aware of robust solutions to this problem. Moreover, in the case of non-planar surfaces, further depth map extraction would be required to determine how the caustic map should be projected.

To avoid these complications, we assume the caustic to be mapped onto a planar surface, adopting a simpler user-assisted approach similar to Mohan et al's [2007], whereby the user specifies the vertices of the projection region by just clicking four points located approximately around the shadow region. An additional advantage to this solution is that the user implicitly and naturally accounts for the fact that the transparent object may be some distance away from the surface that exhibits the caustic.

We then modify the original image according to the following operation on the luminance channel:

$$L_c(x,y) = L(x,y) + \alpha S'(x,y) \qquad (3)$$

where α represents a weighting factor to control its apparent brightness, and $L_c(x,y)$ is the luminance channel of the final image (see Figure 5 (right)).

4 Results

The choice of the number of search directions in the phase symmetry has an impact on the appearance of the resulting caustic, as shown in Figure 6. Fewer directions in general yield simpler, more focused caustics, whereas increasing the number of directions creates more complex patterns. Note that the apparent degree of sharpness in the mapped caustics w.r.t. the number of directions analyzed depends on the specific object and the corresponding ratio defining $S(x)$ in Equation 10. Usually, it is desirable to have a mixture of both focused and complex patterns to better simulate the appearance of real-world caustics. Several caustics maps can be combined in those cases using:

$$L_c(x,y) = L(x,y) + \sum_i \alpha_i S'_i(x,y) \qquad (4)$$

However, our experiments revealed that combining up to two symmetry maps usually suffices in producing plausible imagery. Table 4 shows the number of caustics maps and directions d for each image in this paper.

Figure 7 shows three real-world objects and their caustics computed with our algorithm. The real objects have not been changed to transparent for demonstration purposes. It can be seen that, for simple
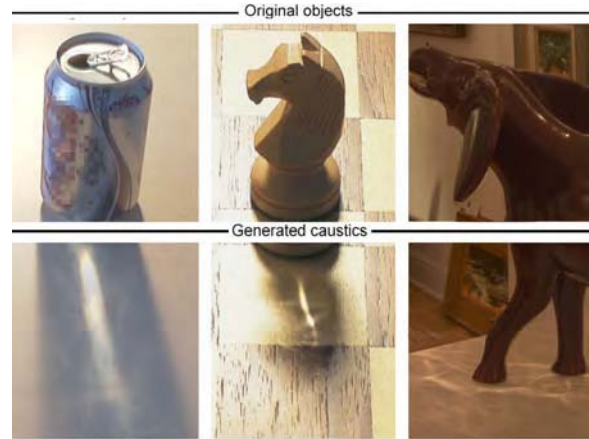

Figure 7: *Real objects with the caustics obtained with our algorithm. For simple objects such as the soda can, the caustics obtained accurately resemble those that would occur in real transparent objects. For more complicated objects, it starts diverging from the real solution but still produces plausible results.*

objects such as the soda can, the algorithm yields results very similar to those obtained in real life (as in the case of the sphere in Figure 5 and the vase in Figure 8 (top)). As the object becomes progressively more complex, like the chess piece and the elephant figurine, the caustics become more complicated and less predictable for an observer. Nonetheless, the caustics produced by our algorithm continue to be commensurate with the expected visual complexity, thereby remaining plausible (Figure 8 (bottom)). This will be validated by means of psychophysical studies in Section 5, while further results are shown in Figure 9.

5 Psychophysics

We claim that the human visual system cannot reliably predict caustics for relatively complex geometries. A very simple test suggests that this is so: Figure 10 shows two images of crystal figurines. One image has photon-mapped caustics, which we take as ground-truth; the other has caustics painted by a digital artist. We then asked 36 participants which one they thought was real. Even though both images present clear differences in the shape and concentration of caustics, none was chosen above chance: 17 people chose the photon-mapped image, compared to 19 people who chose the artist's impression.

Does our algorithm perform as well as this artist? To find out, we performed two experiments, described below. The first assesses the level of tolerance that humans exhibit with respect to errors in caustics, while supporting our choice of algorithm to simulate them. The second experiment is then a ranking of our algorithm against

Figure 6: *The influence of the number of directions. From left to right, caustics obtained searching for symmetries in 1, 2, 4, 12 and 20 directions respectively. The complexity of the caustic pattern increases accordingly.*

Figure 8: *Two full results, showing transparent objects casting caustic patterns near their base (transparency achieved using [Khan et al. 2006]). The shape of the caustic for the vase is relatively simple due to the high degree of symmetry of the object, whereas for the elephant is more complex. Both produce perceptually plausible results. Insets: original images.*

Figure 9: *Additional results adding caustics to the doll, car and phone images.*

Figure 10: *Computer generated crystal figurines. Left: photon-mapped caustics. Right: caustics painted by an artist.*

several images on which artists have painted their impression of caustics. We have taken this specific approach since the only way to produce caustics in existing images is currently by painting over pixels.

A set of 44 participants took part in our first study, and 87 different observers partook in the second, all of them having reported normal or corrected to normal vision. They were naïve as to the design and goals of the experiments, and included computer graphics graduate students as well as non-experts.

5.1 Experiment 1: Validation against 3D Rendering

In this experiment, the first question answered is whether our algorithm produces images which are visually as plausible as a full 3D photon mapping simulation. For this, we employ four different

3D opaque objects of increasing geometric complexity: skull, vertebrae, dolphin and bull (Figure 11). For each one, on the one hand, the algorithm described in this work was applied: phase symmetry was computed in image-space from the opaque renders, then composited into a similar image with a transparent version of the object, thus simulating caustics. Note that no 3D information was used to derive the caustics at this stage. On the other hand, regular photon mapped caustics were rendered for the transparent versions, taken advantage of the true 3D information of the objects. The stimuli were then used in a paired comparisons test.

The second question is whether a simpler algorithm would also produce plausible caustics. If so, then this would indicate that our proposed algorithm is overly complicated, and a simpler solution would suffice. In particular, one might reconstruct approximate geometry from the image, and then render them directly with photon mapping. One of the simplest approaches to generate geometry is

Figure 11: *The four objects used in our first psychophysical test. From left to right: skull, vertebrae, dolphin and bull.*

Scene	ξ	u	Angle	ξ	u
Skull	0.790	-0.068	0	0.903	0.040
Vertebrae	0.903	-0.047	60	0.903	0.044
Dolphin	0.966	0.240	120	0.909	0.021
Bull	0.972	0.249	180	0.914	0.020

Table 2: Coefficient of consistency ξ and coefficient of agreement u per scene and per angle.

to assume that objects are globally convex, thus enabling their silhouettes to be revolved. This approach was added to the paired comparisons test.

Finally, we assess whether knowledge of the light direction in the scene is important for constructing a believable caustic. To this end, each stimulus was recreated for 4 different light positions, with one of the light directions coinciding with the viewpoint. This test allows us to determine if the error introduced by our algorithm (it generates the caustic from the viewpoint, rather than from the light source) in any way harms visual impression. Figure 12 shows the complete set of stimuli for the skull and bull objects.

For each object and light position, we employed a balanced paired comparison test, for a total of 48 pairs (4 scenes × 4 light positions × 3 rendering algorithms), shown side-by-side in random order. The display is a calibrated 21" TFT LCD monitor (1800 × 1600 resolution, 60 Hz refresh rate) with an approximately 150:1 contrast ratio. The participants had to perform a two-alternative forced-choice (2AFC) to answer the question *Which image contains the caustics that look more real to you?*. Upon request, the concept of caustics was explained to each participant individually. All the participants were informed that all the images were computer generated, and that there was not a right or wrong answer. They were also told that the images in each pair were identical except for the caustics. They were previously trained with a different set of images until they felt confident with both the question and the procedure.

As a paired comparisons test is an indirect way to infer a rank order of the three algorithms, it is possible that circular triads occur. For instance a participant may indicate the following preference order: $A_1 > A_2 > A_3 > A_1$, which signifies an inconsistency. The presence of inconsistencies can be measured with the coefficient of consistency ξ [Kendall and Babington-Smith 1940]. Its value will tend to 1 the more consistent the results are. Values for each scene and for each light direction (angle) are given in Table 2, showing that consistency is overall very high.

The coefficient of agreement u, also shown in Table 2, measures whether the three algorithms received equal preference (low scores) or could be discerned based on preference (high scores). We see that for simple geometries (Skull, Vertebrae), participants found it difficult to indicate a preferred algorithm, whereas complicated geometries, with associated complex caustics, lead to more pronounced viewer preference.

These results are consistent over all angles tested, showing that the position of light sources is of little influence, as evidenced by the low values of u shown on the right side of Table 2. We therefore

Skull:	K PM R	K	= Kovesi Phase Symmetry
Vertebrae:	K PM R	PM	= Photon Mapping
Dolphin:	K PM R	R	= Revolution Method
Bull:	K PM R		

Table 3: Grouping of algorithms per scene.

conclude that the error we make by computing the caustic with respect to the viewpoint, rather than with respect to the light source, does not impair our ability to generate a plausible caustic.

Finally, as complicated geometries lead to larger differences in preference ratings, we carried out a significance test of score differences, which allows us to assess which algorithms belong to the same group. Two algorithms belong to different groups if the difference in scores R is below $\lceil R_c \rceil$. Thus, we would like to compute R_c such that:

$$P\left(R \geq \lceil R_c \rceil\right) \leq \alpha \tag{5}$$

where α is the significance level. It can be shown that in the limit R will be identical to the distribution of the range $W_{t,\alpha}$ of a set of t normally distributed random variables with variance $\sigma = 1$ [David 1988]. This enables us to compute R_c using [Setyawan and Lagendijk 2004; Ledda et al. 2005]:

$$P\left(W_{t,\alpha} \geq \left(2R_c - 0.5\right)/\sqrt{nt}\right) \tag{6}$$

where n is the number of participants (44 in our case) and t is the number of algorithms we compare ($t = 3$). The value of $W_{t,\alpha}$ can be interpolated from tables provided by Pearson and Hartley [1966]. For $\alpha = 0.01$, we find that $W_{3,0.01} \approx 4.125$, so that $\lceil R_c \rceil = 24$. The resulting groupings per scene are given in Table 3. At the 0.01 confidence level, our algorithm is always in the same group as the photon mapping approach, and can therefore not be distinguished from the ground truth. For simple geometric shapes this is true also for the method which revolves the silhouette. However, for more complex geometries, this technique is too simple and is reliably distinguished from the ground truth. We therefore conclude that in cases where true 3D geometry is unavailable, our phase symmetry approach can be effectively employed.

This experiment provides insight into our algorithm as compared with a full 3D simulation, showing that the results are visually equivalent. Moreover, for complex geometry an obvious simpler approach falls short, whereas the phase symmetry algorithm continues to produce plausible caustics.

5.2 Experiment 2: Validation against Direct Painting

In addition to assessing the performance of our algorithm with respect to 3D rendering, which establishes a ground truth, we are interested whether direct painting using an image editing program (such as Adobe Photoshop^{TM}) would produce visually comparable results. We expect that the success of direct painting depends on the skill of the artist, as well as the amount of time expended to generate the image.

We therefore asked five digital artists with different backgrounds and styles to paint caustics in two images which were manipulated to create transparency without caustics using Khan et al's method [2006]. One image has a highly symmetric object (a vase) which presumably would yield a symmetric caustic that may be predicted more easily. The other contains an asymmetric object (an elephant figurine) which would produce more complicated caustics. Some of the results are shown in Figure 13, whereas the output of our algorithm is given in Figure 8. One of the artists failed to deliver
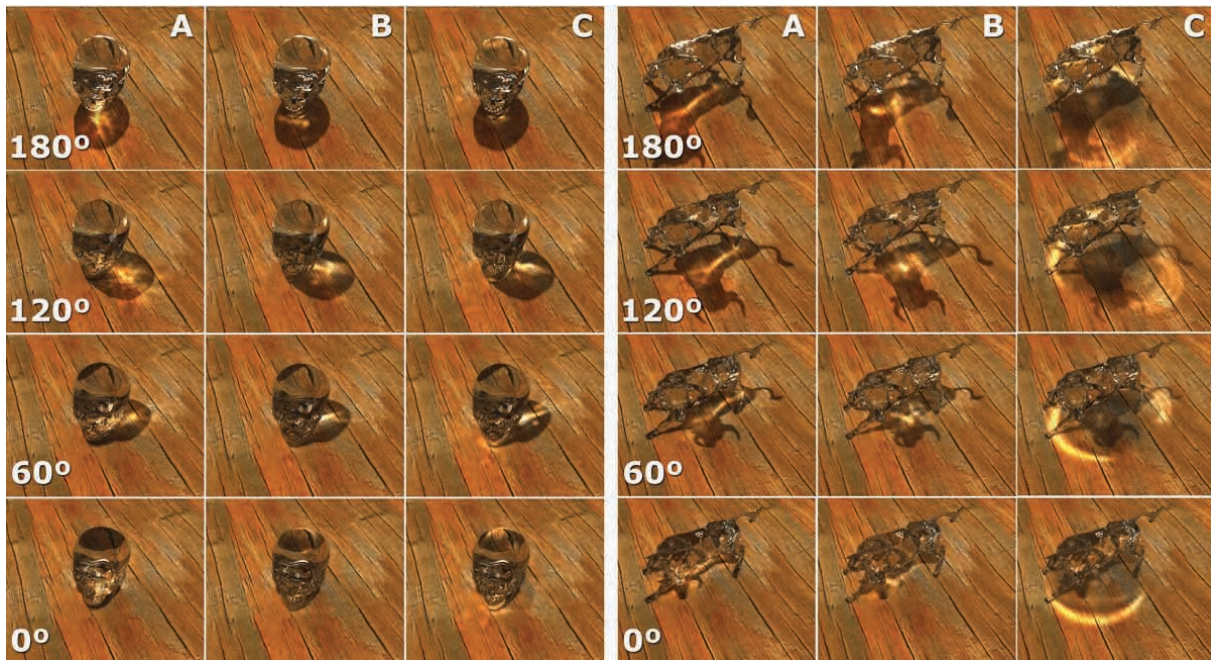
Figure 12: *The complete set of stimuli for the skull and bull objects. Columns A, B and C show the results of our algorithm, photon mapping and the alternative algorithm respectively. Rows indicate light position (degrees) w.r.t the camera. Details are given in the text.*

Figure 13: *Detail of the some of the artists' depictions of the caustics for the vases and elephant images (images 1, 2 and 4 for the vases; 2, 3 and 4 for the elephant, as numbered in the tests).*

Figure 14: *Mean intervals for all the images in the vase and elephant series, along with the 0.95 confidence interval.*

the vase image. Each of the eleven resulting images was printed using a professional sublimation printer at 20×15 cm.

Each participant was informed that the only variation between each set of images were the caustics, and was asked to order the images from more to less real (from 1 to 5 in the vase image; 1 to 6 in the elephant image), according to his or her own standards. No previous training was performed, other than an explanation of what caustics are. The order of the images was randomized within each set for each subject.

Since our goal is to determine if our algorithm produces results comparable to what can be achieved by using image-editing software, rank data is sufficient for our analysis. Figure 14 shows mean rankings for all the images in each series (lower ranking means higher perceived realism) with $p < 0.05$. Our algorithm performed slightly better than the best of the artists images in the case of the vase series, and significantly better in the elephant series.
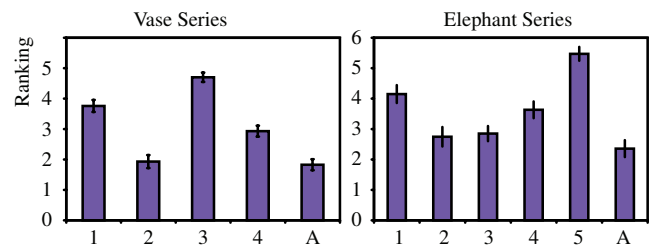
Tables 4 and 5 show normal fit data for all images. Our algorithm has the lowest mean (higher perceived realism) of all the tested images. The artists had no time limitations to paint the caustics. They ended up spending between five and fifty minutes to produce the images, while our algorithm runs in approximately two minutes for the images shown in this paper. We therefore conclude that our algorithm produces results significantly faster than an artist, while obviating the need for skilled input. Moreover, our results are perceived to be more realistic than artists' best efforts.

6   Conclusions

In this paper we have demonstrated the feasibility of rendering perceptually plausible caustics into existing images. We have shown that although humans find it easy to detect the presence of caustics, they are much less adept at predicting the shape of caustics. We have leveraged this feature of human vision to produce an image editing tool that enables, for the first time, aspects of global illumination to be simulated on the basis of a single photograph. There are several advantages to this approach. First, the required user input

Image	1	2	3	4	A
Mean	3.759	1.931	4.701	2.931	1.828
StDev	0.939	1.021	0.733	0.860	0.865

Table 4: Normal fit data (vase series).

Image	1	2	3	4	5	A
Mean	4.149	2.747	2.851	3.632	5.471	2.356
StDev	1.402	1.527	1.186	1.313	1.087	1.329

Table 5: Normal fit data (elephant series).

is unskilled, making the algorithm straightforward to apply. Second, the results are at least on a par with those produced by skilled artists, as evidenced by the second validation study reported in this paper. Third, the time required to render a caustic is only a fraction of the time that a skilled artist would need to paint over all pixels. Our approach could potentially be used in combination with a traditional 3D rendering algorithm, avoiding the need to compute costly caustics and approximating them in image-space. Accurate object depth could be used instead of shape-from-shading information.

Extending this work to video is also possible. For the simplest case of camera movement only, the caustics shape is not expected to change, given that the light is fixed with respect to the object. The projected caustics map for the first frame simply needs to be tracked over successive frames. For more general dynamic scenes with moving objects and/or lights, we can leverage the fact that the shape from shading approach used (from which phase symmetries are obtained) does not introduce temporal artifacts [Khan et al. 2006].

7 Acknowledgments

References

BORN, M., AND WOLF, E. 1999. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th ed. Cambridge University Press, Cambridge, UK.

DAVID, H. A. 1988. *The Method of Paired Comparisons*. Charles Griffin & Company, London.

FANG, H., AND HART, J. C. 2004. Textureshop: Texture synthesis as a photograph editing tool. *ACM Transactions on Graphics 23*, 3, 354–359.

FANG, H., AND HART, J. C. 2006. Rototexture: Automated tools for texturing raw video. *IEEE Transactions on Visualization and Computer Graphics 12*, 6, 1580–1589.

KENDALL, M. G., AND BABINGTON-SMITH, B. 1940. On the method of paired comparisons. *Biometrica 31*, 3/4, 324–345.

KHAN, E. A., REINHARD, E., FLEMING, R., AND BÜLTHOFF, H. 2006. Image-based material editing. *ACM Transactions on Graphics 25*, 3, 654–663.

KOVESI, P. 1996. *Invariant measures of image features from phase information*. PhD thesis, The University of Western Australia.

KOVESI, P. 1997. Symmetry and asymmetry from local phase. In 10th *Australian Joint Converence on Artificial Intelligence*, 2–4.

KOVESI, P. 1999. Image features from phase congruency. *Videre: Journal of Computer Vision Research 1*, 3, 2–26.

KRUGER, J., BURGER, K., AND WESTERMANN, R. 2006. Interactive screen-space accurate photon tracing. In *Proceedings of the Eurographics Symposium on Rendering*, 319–329.

LANGER, M., AND BÜLTHOFF, H. H. 2000. Depth discrimination from shading under diffuse lighting. *Perception 29*, 6, 649–660.

LEDDA, P., CHALMERS, A., TROSCIANKO, T., AND SEETZEN, H. 2005. Evaluation of tone mapping operators using a high dynamic range display. *ACM Transactions on Graphics 24*, 3, 640–648.

MOHAN, A., TUMBLIN, J., AND CHOUDHURY, P. 2007. Editing soft shadows in a digital photograph. *IEEE Computer Graphics and Applications 27*, 2, 23–31.

MORRONE, M. C., AND BURR, D. C. 1988. Feature detection in human vision: A phase-dependent energy model. *Proceedings of the Royal Society of London B 235*, 1280, 221–245.

OH, B. M., CHEN, M., DORSEY, J., AND DURAND, F. 2001. Image-based modeling and photo editing. In *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer Graphics and Interactive Techniques*, 433–442.

OPENHEIM, A. V., AND LIM, J. S. 1981. The importance of phase in signals. *Proceedings of the IEEE 69*, 5, 529–541.

TE PAS, S. F., AND PONT, S. C. 2005. Estimations of light source direction depend critically on material brdfs. *Perception. Supplement ECVP05 34*, 212.

PEARSON, E. S., AND HARTLEY, H. O. 1966. *Biometrika Tables for Statisticians*, 3rd ed., vol. 1. Cambridge University Press.

PIOTROWSKI, L. N., AND CAMPBELL, F. W. 1982. A demonstration of the visual importance and flexibility of spatial-frequency amplitude and phase. *Perception 11*, 3, 337–346.

ROSENFELD, A. 1986. Axial representations of shape. *Computer Graphics and Image Processing 33*, 2, 156–173.

SETYAWAN, I., AND LAGENDIJK, R. L. 2004. Human perception of geometric distortions in images. In *Proceedings of SPIE, Security, Steganography and Watermarking of Multimedia Contents VI*, 256–267.

SHAH, M., KONTTINEN, J., AND PATTANAIK, S. 2006. Caustics mapping: an image-space technique for real-time caustics. *IEEE Transactions on Visualization and Computer Graphics 13*, 2, 272–280.

SZIRMAY-KALOS, L., ASZÓDI, B., LAZÁNYI, I., AND PREMECZ, M. 2005. Approximate ray-tracing on the GPU with distance impostors. *Computer Graphics Forum 24*, 3, 695–704.

TOMASI, C., AND MANDUCHI, R. 1998. Bilateral filtering for gray and color images. In *Proceedings of the IEEE International Conference on Computer Vision*, 836–846.

TYLER, C. W., Ed. 1996. *Human Symmetry Perception and its Computational Analysis*. VSP International Science Publishers, Utrecht.

WAGEMANS, J. 1995. Detection of visual symmetries. *Spatial Vision 9*, 1, 9–32.

WEI, H., AND KAIHUAI, Q. 2007. Interactive approximate rendering of reflections, refractions, and caustics. *IEEE Transactions on Visualization and Computer Graphics 13*, 3, 46–57.

WICHMANN, F. A., BRAUN, D. I., AND GEGENFURTNER, K. R. 2006. Phase noise and the classication of natural images. *Vision Research 46*, 8/9, 1520–1529.

WU, J., AND YANG, C.-X. 2005. Detecting image symmetry based on phase information. In *Proceedings of the Fourth International Conference on Machine Learning and Cybernetics*, 5150–5153.

WYMAN, C., AND DACHSBACHER, C. 2008. Improving image-space caustics via variable-sized splatting. *Journal of Graphics Tools 13*, 1, 1–17.

WYMAN, C., AND DAVIS, S. 2006. Interactive image-space techniques for approximating caustics. In *Proceedings of the ACM Symposium on Interactive 3D Graphics and Games*, 153–160.

WYMAN, C. 2005. An approximate image-space approach for interactive refraction. *ACM Transactions on Graphics 24*, 3, 1050–1053.

WYMAN, C. 2007. Interactive refractions and caustics using image-space techniques. In *ShaderX5: Advanced Rendering Techniques*, W. Engel, Ed. Charles River Media, 359–371.

WYMAN, C. 2008. Hierarchical caustic maps. In *Proceedings of the ACM Symposium on Interactive 3D Graphics and Games*, 163–171.

XIAO, Z., HOU, Z., MIAO, C., AND WANG, J. 2005. Using phase information for symmetry detection. *Pattern Recognition Letters 26*, 13, 1985–1994.

YUAN, X., AND SHI, P. 2005. Iris feature extraction using 2d phase congruency. In *Third International Conference on Information Technology and Applications*, vol. 2, 437–441.

ZABRODSKY, H. 1993. *Computational Aspects of Pattern Characterization. Continuous Symmetry*. PhD thesis, Hebrew University in Jerusalem.

ZELINKA, S., FANG, H., GARLAND, M., AND HART, J. C. 2005. Interactive material replacement in photographs. In *Proceedings of Graphics Interface*, 227–232.

A Phase symmetry

The phase symmetry algorithm is based on a log Gabor filter bank. We present the phase symmetry algorithm in 1D first, and then show how it is applied to the 2D signal. In 1D, a signal $D(x)$ is convolved by even-symmetric (cosine) wavelet filters M_n^e and odd-symmetric (sine) wavelet filters M_n^o which operate at scale n. The even-symmetric and odd-symmetric responses to such a quadrature pair of filters at scale n is given by $e_n(x)$ and $o_n(x)$ respectively [Kovesi 1999]:

$$(e_n(x), o_n(x)) = (D(x) \otimes M_n^e, \; D(x) \otimes M_n^o) \quad (7)$$

where $\otimes$ denotes a convolution. Wavelets have a limited spatial extent, which is determined by the chosen scale n. A filter bank

analyzing different frequencies can therefore be constructed by repeating this computation for different scales. The $e_n(x)$ and $o_n(x)$ values represent the real and imaginary components of the local frequencies present in the signal around the location of interest x. The amplitude $A_n(x)$ and phase $\Theta_n(x)$ are then given by[2]:

$$A_n(x) = \sqrt{e_n^2(x) + o_n^2(x)} \quad (8)$$

$$\Theta_n(x) = \tan^{-1}\left(\frac{e_n(x)}{o_n(x)}\right) \quad (9)$$

Given that symmetry appears as large absolute values of the even-symmetric filter and small absolute values of the odd-symmetric filter, we can subtract both values and produce a weighted average to combine information over multiple scales. This measure of symmetry $S(x)$ corresponds to [Kovesi 1997]:

$$S(x) = \frac{\sum_n \lfloor A_n(x)\,(|\cos(\Theta_n(x))| - |\sin(\Theta_n(x))|) - T \rfloor}{\sum_n A_n(x) + \epsilon} \quad (10)$$

Here, ϵ is a small constant to avoid division by zero (we use 0.01), and T is an estimate of the signal noise, and is included to remove spurious responses. This estimate can be computed by first considering the energy vector $E(x)$:

$$E(x) = \sqrt{\left(\sum_n e_n(x)\right)^2 + \left(\sum_n o_n(x)\right)^2} \quad (11)$$

Assuming that the noise has a Gaussian distribution with random phase and a standard deviation of σ_G, then it can be shown that the noise distribution of the magnitude of the energy vector has a Rayleigh distribution with mean μ_R and variance σ_R^2 given by [Kovesi 1999]:

$$\mu_R = \sigma_G \sqrt{\frac{\pi}{2}} \quad (12)$$

$$\sigma_R^2 = \frac{4 - \pi}{2}\sigma_G^2 \quad (13)$$

With a scale factor k chosen to be 2 or 3, a good value for T is then:

$$T = \mu_R + k\,\sigma_R \quad (14)$$

The one-dimensional symmetry computation $S(x)$ can be extended to two dimensions by repeating Equation (10) for different directions in the frequency domain. Using polar coordinates, the filter in the radial direction is given by $S(x)$, whereas in the angular direction θ filters $G(\theta)$ with Gaussian cross-sections are chosen:

$$G(\theta) = \exp\left(-\frac{(\theta - \theta_0)^2}{2\sigma_\theta^2}\right) \quad (15)$$

Here, θ_0 is the orientation angle of the filter, and σ_θ is the standard deviation chosen for the Gaussian filter. In addition to summing over all scales, we now have to sum over all orientations θ_i as well, yielding Equation (1).

---

[2]Note that to determine in which quadrant $\Theta_n(x)$ lies, it is effectively computed with $atan2()$.