

# 3D Unsharp Masking for Scene Coherent Enhancement

Tobias Ritschel    Kaleigh Smith    Matthias Ihrke    Thorsten Grosch    Karol Myszkowski    Hans-Peter Seidel  
MPI Informatik



**Figure 1:** A naturally illuminated 3D scene (left) and the same scene with 3D unsharp masking enhancement (center). Our enhancement technique is coherent with the scene itself, not simply with each rendered frame, permits arbitrary lighting and is temporally coherent.

## Abstract

We present a new approach for enhancing local scene contrast by unsharp masking over arbitrary surfaces under any form of illumination. Our adaptation of a well-known 2D technique to 3D interactive scenarios is designed to aid viewers in tasks like understanding complex or detailed geometric models, medical visualization and navigation in virtual environments. Our holistic approach enhances the depiction of various visual cues, including gradients from surface shading, surface reflectance, shadows, and highlights, to ease estimation of viewpoint, lighting conditions, shapes of objects and their world-space organization. Motivated by recent perceptual findings on 3D aspects of the Cornsweet illusion, we create scene coherent enhancements by treating cues in terms of their 3D context; doing so has a stronger effect than approaches that operate in a 2D image context and also achieves temporal coherence. We validate our unsharp masking in 3D with psychophysical experiments showing that the enhanced images are perceived to have better contrast and are preferred over unenhanced originals. Our operator runs at real-time rates on a GPU and the effect is easily controlled interactively within the rendering pipeline.

**Keywords:** Enhanced Rendering, Apparent Contrast Enhancement, Temporal Coherence, Visual Perception, Cornsweet Illusion

## 1 Introduction

A basic principle behind image comprehension is that humans are able to mentally reconstruct an original scenario from visual cues such as shading, occlusions, perspective foreshortening, shadows and specularities. Local contrast enhancement emphasizes these

cues, aiding the interpretation of 3D scenes and complex geometry, which is the common task in applications such as medical diagnostics, computer simulations, geographical navigation, game playing and film creation. The main problem is in deciding which cues to emphasize and how to do so with a predictable effect.

The goal of this work is to construct a perceptually founded approach for local contrast enhancement of arbitrary interactive 3D scenes. Such an approach should provide easier shape recognition, better visual separation between objects and a clarification of their spatial arrangement solely by increasing the apparent contrast of specific visual cues. Instead of identifying and modifying cues separately, we look at their common cause, changes (or gradients) in reflected light. These light gradients include all cues caused by variations in surface geometry, material properties, incoming light properties and the spatial arrangement of objects. For example, where a surface receives different amounts of incoming light (possibly in shadow), where reflectance properties change, and where specular highlights occur. In this work we strive to simultaneously increase the contrast of all such gradients without breaking coherence with the depicted scene.

To create a scene coherent enhancement technique for 3D computer graphics, we present a novel form of unsharp masking. In generic unsharp masking, an input signal is enhanced by adding back a contrast signal which is a scaled highpass version of itself. To adapt this principle to 3D, we define the *input signal* as the outgoing radiance from a surface point to a viewpoint and the *contrast signal* as the high-frequencies in this radiance measured over the mesh surface. We determine these high-frequencies by subtracting a smooth version of outgoing radiance from the input signal. Our scene coherent smoothing is applied over the mesh surface itself, so that the contrast signal can be thought of as lying over the surface. In addition, we present an adaptive gain term that accounts for the perceptual effect of texture masking, which makes the enhancement less apparent. Our GPU implementation results in interactive performance and the process fits into the rendering pipeline, allowing simple adjustment of the enhancement strength with an immediately viewable effect.

To justify and analyse our 3D approach, we consider the perceptual effect of unsharp masking; enhancement occurs because the gradient is altered in such a way that its apparent contrast (the perceived magnitude of change over the gradient) is increased beyond its physical contrast. Studies in perception have long analyzed local

contrast, especially in relation to the so-called Cornsweet effect, where a specific shape of gradient creates the illusion of contrast. We introduce recent work by Purves et al. [1999] which shows that the Cornsweet effect is strongest when coherent with a 3D scene and its lighting (see Fig. 2). We generalize this observation for all local contrast to justify our scene coherent approach. To confirm our approach, we also show that our 3D unsharp masking reproduces the Cornsweet illusion and can thus be used as a tool for further perceptual studies.

The outcome of our endeavour is varied; consider the perusal of a complex anatomical model, much like that shown in the right-most image of Figure 1. To ease comprehension of the complex model, the shading and occlusion gradients should be made obvious, but enhancements that appear or disappear when changing viewpoint would create inconsistencies that impede understanding. Our technique exaggerates gradients to provide greater visual separation between regions and enhances the shadows that underscore the geometry, doing so coherently over all viewpoints and transitions, and under any type of illumination (remaining images of Figure 1). Another instance is the exploration of a scene with details that lay in shadow. The desire is for details to remain visible when they are the focus, without forcing the shadows to lighten. By enhancing contrast along the shadow edges, the Cornsweet illusion creates a darker appearance without compromising detail visibility, as we show in Figure 5.

The remainder of this paper is organized as follows. In Section 2 we provide an overview of related work in local contrast enhancement of 3D and 2.5D data. In Section 3, we support our approach with perceptual findings on the apparent contrast of the Cornsweet effect and draw relationships to our task. We present our 3D unsharp masking operator in Section 4, then detail its components and real-time implementation in Section 5. We then show, in Section 6, how the enhancement can be utilized in real applications. We present the major findings from our perceptual study that measures the effect of parameters and analyzes user reaction to our enhancements in Section 7, with additional details contained in the supplemental material entitled “Experimental Validation of the Algorithm”. Finally we discuss limitations of 3D unsharp masking and conclude with our ideas for future directions.

## 2 Related Work

This work falls within 3D computer graphics as a last step in the rendering pipeline. Rendering has succeeded in producing real to life imagery, and the next step is to produce more communicative and efficient imagery from that basis. This is a broad goal with diverse applications, and as such, has generated much interest. Most recent approaches for creating more evocative images perform local enhancement of different aspects, such as incident lighting angle, object features, depth differences. Our work is also a local approach, yet it treats all features simultaneously and is interactive and temporally coherent. Its use is not restricted to the creation of a single effective viewing, but rather to be used in 3D scene exploration and interaction.

We first consider related work that enhances the shape and surface details of a single 3D object. Given a 3D model, Cignoni et al. [2005] shift object normals, performing normal enhancement to emphasize geometric discontinuities in a single rendered image. In this approach, enhancement strength is controlled by how much the geometry of the mesh is sharpened. Inspired by cartographic illustration, Rusinkiewicz et al. [2006] introduce a new shading model to expose shape features and surface details by positioning a local light per-vertex to achieve maximum contrast. Enhancement strength is controlled by adaptively combining multiple scales of renderings.

The enhancement of complex geometry and multi-object scenes addresses the task of communicating the spatial arrangement of various objects and their relationships to each other. Rendering of molecules [Tarini et al. 2006] or medical volumetric data [Bruckner and Gröller 2007] can benefit from image space halos and enhanced gradients. Ambient occlusion [Landis 2002] introduces gradients towards concavities of a surface. Our approach achieves a similar effect with lower overhead, and increases gradients that are not only due to simple concavities.

Our work is most closely related to Luft et al. [2006], who focus on enhancing depth perception by unsharp masking the depth buffer. Depth between objects is emphasized by darkening or lightening the original image at locations of discontinuities in its accompanying depth map, however, places without depth changes (like object and ground intersections) are not enhanced. A trivial extension of this work to 3D scenes would filter additional 2D buffers to construct the contrast signal. However, this would be unsatisfactory, due to temporal coherence problems, having parameters that do not reflect scene geometry and viewpoint, and the precision is not high enough to enhance small mesh features. For these reasons, and to provide an interactive approach that handles arbitrary geometry, a different approach, like ours, is required.

One important cue for spatial understanding is object and ground plane intersections, however, there is no enhancement because these areas have no depth discontinuity. This work was also presented solely for single frames, and its frame-by-frame application could lead to temporal incoherence when the depth map changes drastically.

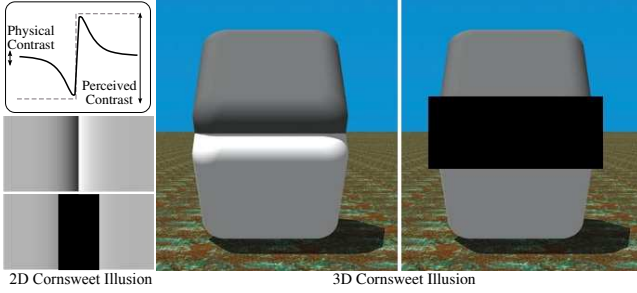
Most of the approaches mentioned above do not enhance lighting gradients that occur where geometry is smooth or where there is no depth difference. These are common locations of shadows and specular highlights, which in addition to being necessary cues for proper scene comprehension, can help by underscoring geometry and clarifying spatial relationships [Cavanagh and Leclerc 1989] and communicating shape and material properties [Fleming et al. 2004]. Recent work by DeCoro et al. [2007] shows that shadows can benefit from modification. To achieve their goal, shadows in image space undergo a recovered perspective transformation so that filtering is correct in 3D, whereas our approach modifies shadows directly in 3D, treating them as a natural part of the whole surface lighting.

Existing 2D and 2.5D approaches applied in a frame-to-frame manner would introduce image space artifacts, such as popping and the effect of enhancements stuck within the image (shower door effect), instead of within the scene itself. In the field of expressive rendering, there exist techniques for creating many forms of coherent stylized renderings that address the problem of temporal coherence, for example line drawings of 3D scenes [DeCarlo et al. 2004]. Their findings show that the contours depicted (and by analogy, the contrasts we add) can not simply exist in the 2D projection and slide through image space (creating a shower door effect), but should follow the content and movement of the underlying 3D scene. To our knowledge, we are the first to explicitly construct a 3D enhancement that satisfies temporal coherence constraints.

## 3 Local Contrast and the Cornsweet Illusion

While apparent contrast is most directly related to differences between adjacent regions, it is also impacted by the separating gradient’s shape and magnitude [Kingdom and Moulden 1988]. If a gradual gradient blends between regions, an enhanced gradient does the opposite by adding distinction, thus increasing our perception of contrast. Contrast studies show that a sharp gradient can increase apparent contrast to the point where an equiluminant patch appears

to have two contrasting regions purely on account of the edge. Most evocative is the Cornsweet illusion, where neighbouring regions are filled with illusory brightness, giving a sense of contrast where there is nearly none [Cornsweet 1970], as can be seen in Figure 2. This shows that local features can increase the apparent dynamic range far beyond that which is physically present. Recent research in computer graphics has demonstrated the Cornsweet illusion’s contrast enhancement abilities in 2D: one, for restoring original high dynamic range (HDR) contrast to low dynamic range (LDR) tone-mapped photographs [Krawczyk et al. 2007], and another for restoring colour contrast to greyscale converted images [Smith et al. 2008].



**Figure 2:** Illusions from Purves et al. [Purves et al. 1999] showing that a Cornsweet profile in a 3D scene produces a much stronger perceptual contrast than simply showing the brightness profile in 2D. The top face of the cube appears much darker than the bottom face when in fact they are nearly identical.

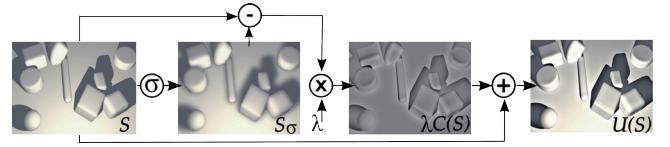
Purves et al. [1999] explain that the Cornsweet illusion arises because we are predisposed to interpreting the stimulus as a result of lighting or reflectance properties. For instance, the gradient could arise from changes in surface reflectance/texture variation, gradients of illumination: penumbra (occlusion), curvature, attenuation with distance, transmittance (partial occlusion). This is why the illusion is weakened if the edge seems painted on, or if it becomes discordant with other aspects in the image. It also explains why the illusion is strengthened when it is reinforced by compatible visual cues. In particular, the illusion’s salience is increased when it implies depth by incorporating perspective projections. The gradient’s orientation should be consistent with lighting, as well as other cues like texture and background, things that would indicate normal viewing conditions. These findings tell us that although local contrast is interpreted from 2D visuals, its interpretation is in terms of real-world lighting or reflections. For this reason, we argue that enhancement should be coherent with the entire scene and its viewing.

## 4 3D Unsharp Masking

Unsharp masking in the two-dimensional image domain proceeds first by smoothing the entire image and then adding back the scaled difference between the original and the smooth image, as depicted in Figure 3 [Badamchizadeh and Aghagolzadeh 2004]. More generally, unsharp masking  $\mathcal{U}(S)$  of a signal  $S$ , with strength  $\lambda$  and smoothness  $\sigma$  is defined as:

$$\mathcal{U}(S) = S + \lambda(S - S_\sigma).$$

In our 3D adaptation,  $S$  is the outgoing radiance from a surface location to a viewpoint, and  $S_\sigma$  is that radiance *smoothed over the surface of the mesh*. For simplicity, we denote the difference between the sharp and smooth signals as the contrast signal  $\mathcal{C}(S) = S - S_\sigma$ . The width of contrast signals are controlled by a user chosen smoothness parameter  $\sigma$  and  $\lambda_{\text{Adapt}}$  is a locally adaptive function



**Figure 3:** The process of unsharp masking sharpens a signal  $S$  by adding back a  $\lambda$ -scaled contrast signal  $\mathcal{C}(S)$ , the difference between the original and a smoothed version  $S_\sigma$ .

of a user chosen constant gain  $\lambda$ . By its definition in terms  $S$  and  $S_\sigma$ , the highpass signals contained in  $\mathcal{C}(S)$  lie over the surface, and are thus intrinsically coherent with the scene itself. Details on its calculation and coherence are given in Section 5.

All 3D unsharp masking operations are performed in the CIELAB colour space because it provides a perceptually uniform lightness channel  $L^*$  that is decorrelated from its  $a^*$  and  $b^*$  chromaticity channels. By doing so, we ensure that contrast is measured and modified as apparent lightness and that the hue angle  $H_{ab} = \tan^{-1}(\frac{b^*}{a^*})$  is unchanged. Formally, the per-channel operations are:

$$\mathcal{U}(S)_{\text{LAB}} = [S_{L^*} + \lambda_{\text{Adapt}}\mathcal{C}(S)_{L^*}, kS_{a^*}, kS_{b^*}]$$

Where  $S_{a^*}$  and  $S_{b^*}$  are scaled by  $k = \mathcal{U}(S)_{L^*}/S_{L^*}$  to prevent over- or under-saturation during the lightness adjustment. After our enhancement, the new saturation  $S_{ab}(\mathcal{U}(S))$  is equal to the original  $S_{ab}(S)$ , since saturation is defined as  $S_{ab} = \sqrt{a^{*2} + b^{*2}}/L^*$ .

The gain value is usually a user-chosen constant, however, better results are achieved with a gain value that locally adapts to the scene. With  $\lambda$ , the user balances the amount of enhancement to be applied against overshooting, which appears as visible halos. The halo effect is caused by a deficit of the low frequency signal in  $S$  with respect to the high frequencies amplified by adding  $\mathcal{C}(S)$ . The visual system is less sensitive to halos when other frequencies, like those created by a background texture, are present. Thus, in regions where texture masking occurs,  $\lambda$  can be greater without introducing visible halos. To adapt to this perceptual effect, we increase  $\lambda$  in textured regions using a simple model inspired by the quantization strategy in JPEG 2000 [Zeng et al. 2001]. Given a user chosen  $\lambda$ , we define our adaptive gain  $\lambda_{\text{Adapt}}$  as:

$$\lambda_{\text{Adapt}} = \lambda\sqrt{1 + \text{Var}(S)} \quad (1)$$

where  $\text{Var}(S)$  is the variance in the reflected light over a small neighbourhood on the surface of the mesh (its 1-ring neighbourhood).

## 5 Contrast Signal Calculation

We now describe our calculation of the smooth lighting signal  $S_\sigma$ , that gives rise to the contrast signal  $\mathcal{C}(S)$ .  $S_\sigma$  is the view-dependent reflected light  $S$  smoothed in 3D over the mesh surface. The surface is given as a triangular mesh, consisting of vertices and facets. We use *Laplacian smoothing* [Taubin 1995], which replaces the intensity of each vertex  $v_i$  by the average intensity of its neighbours:

$$v_i = \frac{1}{n_i} \sum_{j=1}^{n_i} r_{i,j}$$

where  $n_i$  is the number of neighbours and  $R_i = \{r_{i,j}\}$  are the direct neighbours of  $v_i$ . We call  $n_i$  the *valence* and  $R_i$  the *1-ring* of  $v_i$ . Iterating this filter results in increasingly smooth versions of  $S_\sigma$ .

We render  $S_\sigma$  for use in the per-pixel calculation of the contrast signal  $\mathcal{C}(S)$ , however it is sufficient to calculate and store  $S_\sigma$  per-vertex since it is band-limited. When  $\sigma$  is small, the contrasts in

$\mathcal{C}(\mathcal{S}) = \mathcal{S} - \mathcal{S}_\sigma$  are narrow (the higher frequencies); as  $\sigma$  increases, the lower frequency bands become included in the contrast signal.

As opposed to 2D smoothing operations, our smoothing automatically adapts to the surface’s orientation and location with respect to the viewpoint, thus undergoing correct perspective foreshortening. If we were to consider a hypothetical 2D filter kernel enacting a similar convolution over an image, it would have to change its 2D orientation, shape and size according to the extracted 3D information. Because our smoothing is applied directly over the surface, the resulting contrasts are coherent with the scene itself, and the method is often more efficient than a 2D approach attempting to achieve a similar coherence.

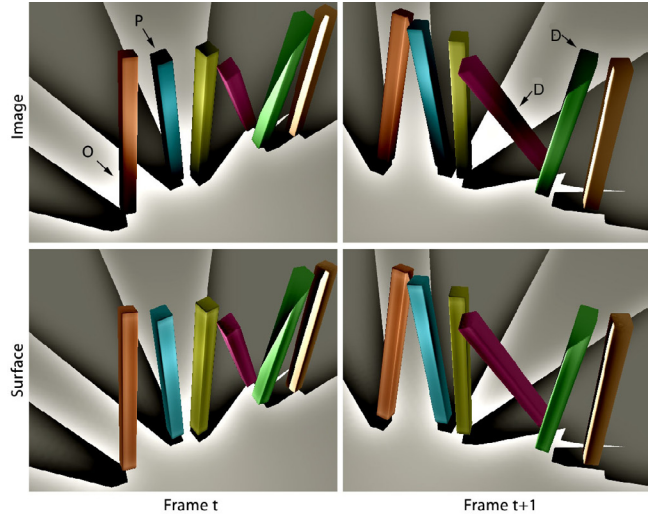
**GPU Implementation** Our method is implemented using a GPU to perform lighting then smoothing. We first compute the initial outgoing light with an arbitrary, sufficiently fast rendering method for every vertex. Instead of rasterizing triangles in the framebuffer from those vertices, we store every lit vertex into individual texels of an output texture. This texture encodes  $\mathcal{S}$  and has as many texels as there are vertices. For proper anti-aliasing, we optionally supersample the 1-ring of each vertex with a low number (e.g., 3) samples per triangle, chosen with the same barycentric coordinates to maintain temporal coherence. For smoothing, we apply a fragment program to every texel (that represents a lit vertex in  $\mathcal{S}$ ) replacing it with the average of all texels in its 1-ring (other lit vertices in  $\mathcal{S}$ ). The 1-ring for every vertex  $v_i$  is stored using two *connectivity* textures: Recall that the valence is different for every  $v_i$ . Therefore, a first texture stores at texel  $i$  the sum of all valences *before* vertex  $v_i$ :  $\sum_{j=1\dots i} n_j$ . A second texture stores the 1-rings for all vertices as a concatenation of the individual indices. To replace a  $v_i$  with the average 1-ring, first the  $i$ -th and the following texel are read from the first texture. Their difference is  $n_i$  and the first texel’s value is used to index into the second texture where  $n_i$  consecutive texels form  $R_i$ . We circulate  $R_i$ , in a small loop, divide by  $n_i$  and store back the increasingly smooth lighting into texel  $i$  of an output texture.

## 5.1 Coherence of the Contrast Signal

Let us now consider the origin of highpass signals contained in  $\mathcal{C}(\mathcal{S})$  and their temporal coherence. The signals we detect need not originate from gradients with specific causes such as geometric discontinuities or depth discontinuities. Instead, they include all gradients in reflected light: surface geometry – variations due to curvature and surfaces facing towards or away from light; reflectance of surface material – variations due to texture and highlights; and incident light – variations due to different amounts of incoming light from different directions and light blocked by occluders creating shadows. The contrast signal calculation is said to be scene coherent and temporally coherent because it is not affected by view-dependent occlusions (whether a surface point is visible or not in the image), nor is it limited by image space undersampling especially at grazing angles.

The effect of our contrast signal is best understood through a comparison to a 2D unsharp masking contrast signal. In our case, the signals are continuous along the surface of the mesh and are not interrupted by distant objects that may be acting as occluders, as would occur in a 2D approach, Figure 4(O). This prevents the presence of continuous signals around disconnected objects that change as viewpoint changes. Our contrast signal calculation respects gaps between objects to be consistent with their spatial grouping, which remains temporally consistent. This also prevents brightness changes to a surface on account of its contrast with other distant surfaces appearing to its front or back, leading to brightness inconsistency over time, Figure 4(D).

In 2D, the detected high-frequencies have the same width in image space, so that near and far gradients are treated in the same scale, when in fact they have different sizes. So a wide gradient that is distant, may be enhanced as a very narrow gradient, Figure 4(P). Lastly, since our calculation of contrast is not affected by what is visible or not, we do not introduce popping due to visibility changes and we do not perform spurious enhancement at image borders. Please refer to the accompanying video to see the coherence of our enhancements, and a comparison between 2D unsharp masking and our 3D approach.



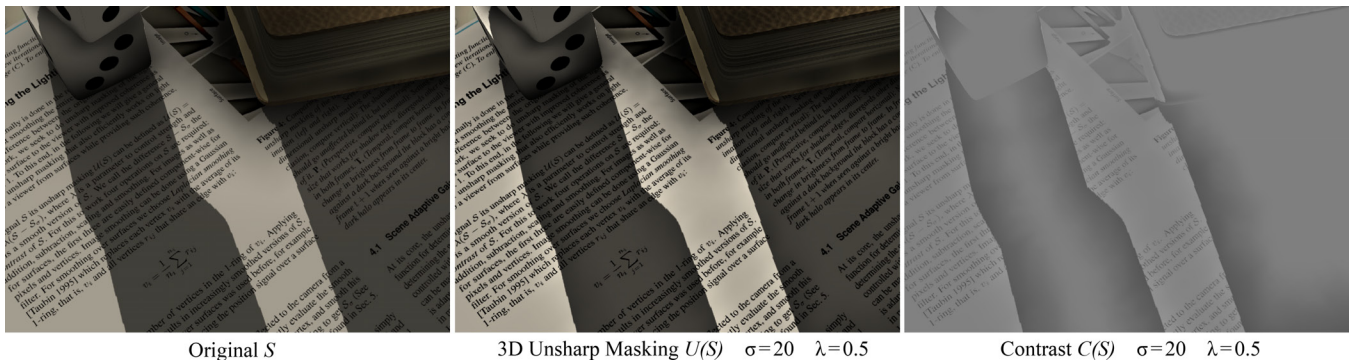
**Figure 4:** Comparing 2D (top) and 3D (bottom) unsharp masking over two successive frames (left and right). **O** (Occlusion, compare vertical): The shadow enhancement should not be interrupted by the occluding stick. **D** (Distant, compare horizontal): Note the strong change in brightness of surfaces between frames because the sticks are darker or lighter than distant background. **P** (Perspective, compare horizontal): Setting a constant filter size that works for shadow edges, disregards discontinuities on the sticks.

## 6 Results

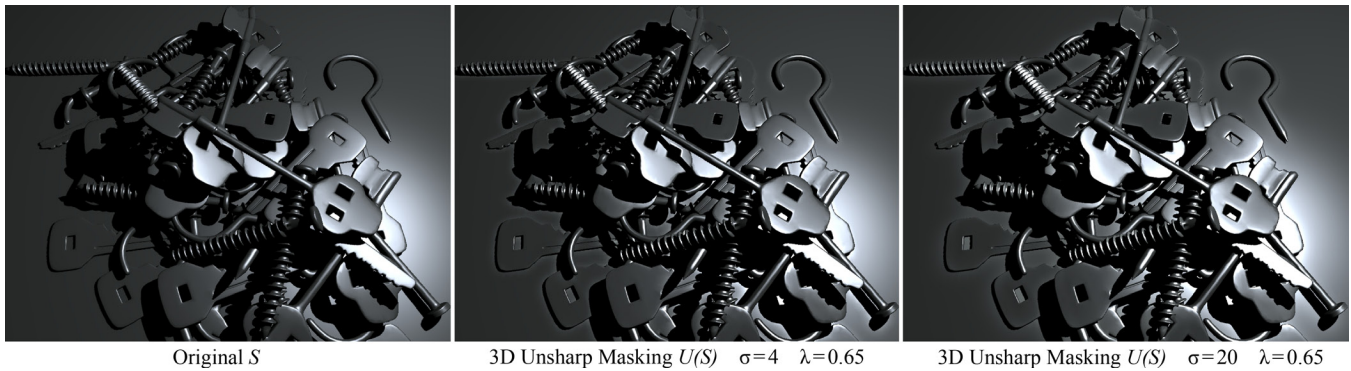
We begin by showing that our method achieves similar high-level enhancements as normals sharpening [Cignoni et al. 2005], exaggerated shading [Rusinkiewicz et al. 2006] and depth unsharp masking [Luft et al. 2006] in Figure 7. Using similar viewpoints and models, we show that the results of 3D unsharp masking are indeed comparable to, or improve over the other methods, and ours is a comprehensive approach that performs the diverse enhancements with a single simple technique (please see images in cited papers for comparison). In addition to making shapes more obvious, depth also appears enhanced due purely to the enhanced shading gradients.

Rusinkiewicz [2006] shows that varying enhancement strengths according to importance may be desired. They do so over the space of the rendered image with a multiscale approach. We show that the same can be done within the rendering pipeline, without need for rendering at multiple scales and recombining. We simply paint rough gain controls over the 3D object, in this case to indicate importance with both positive and negative values, and use these as  $\lambda$ . As depicted in Figure 10, the effect both enhances the face, torch and hands while deemphasizing the robe and pedestal.

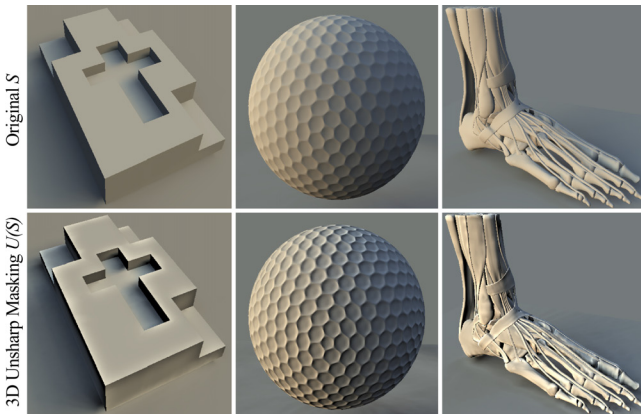
In response to our perceptual motivation from the Cornsweet illusion, we show in Figure 11, that 3D unsharp masking of a uniformly grey cube, lit from below, creates perceived contrast where there is only



**Figure 5:** The apparent dynamic range of the original (left) is increased by 3D unsharp masking (middle), which emphasizes the shadow edges.  $\lambda_{Adapt}$  ensures that the contrast signal is slightly higher at text areas to account for texture masking (right).



**Figure 6:** 3D unsharp masking enhances cast shadows, emphasizing occlusions. The  $\sigma$  value changes the width of enhancement effects.



**Figure 7:** Using meshes from Cignoni, Rusinkiewicz and Luft (left to right) under natural lighting, we show that our 3D unsharp masking enhances the renderings in the proper regions and additionally accounts for shadows and contrasts that do not have depth differences.

slight actual contrast. Returning to our introductory scenario about detail in shadows, we show that 3D unsharp masking creates the impression of darker shadows while keeping the text readable in Figure 5. The shadows appear darker because of the Cornsweet illusion, yet do not appear “sharpened”.

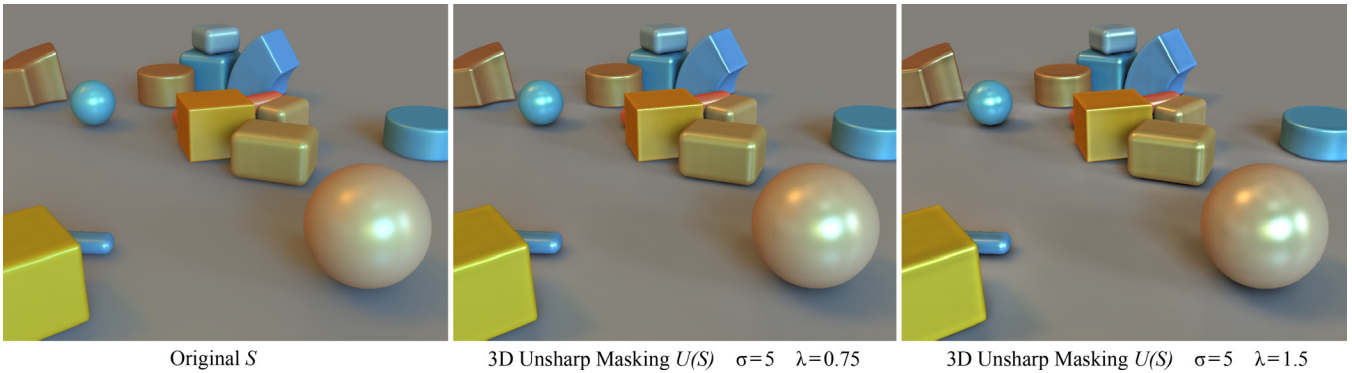
A cluttered scene of similarly shaped keys and screws are shown in Figure 6. Our enhancements (for two different  $\sigma$  values) ease the

cognitive task of object distinction and separation from ground by enhancing the cast shadows, especially those between objects and the ground, and the enhanced speculars aid shape recognition. The *Chamfer Plane* scene in Figure 8 shows that a naturally lit scene with complex materials and spatial arrangement is seamlessly enhanced. For two different  $\lambda$  values, its soft shadows remain soft, yet slightly more prominent. Occlusions become more obvious, object silhouettes are emphasized. The glossy highlights become more bright, helping with shape comprehension and curvature understanding. Objects appear more grounded on the plane, since their intersection with the ground is also emphasized.

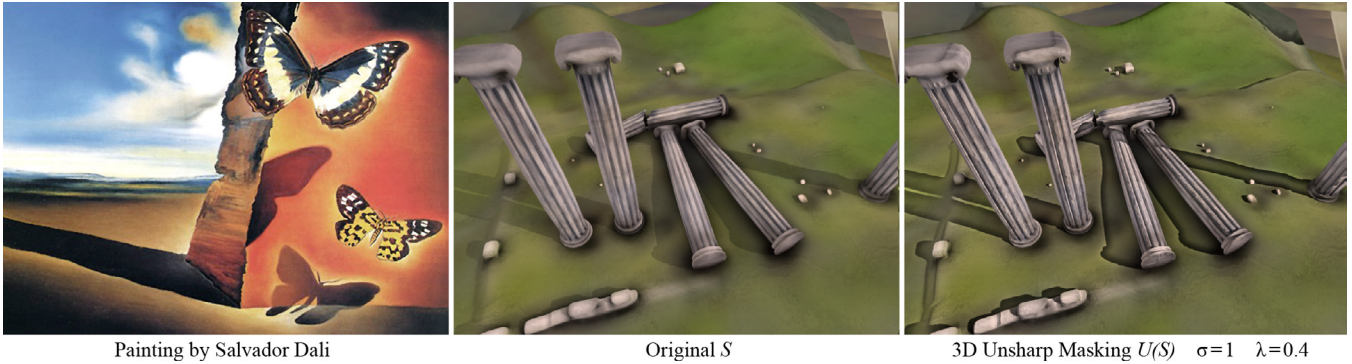
The same effects we create can be observed in art, where contrasting shades are juxtaposed to exaggerate local lighting discontinuities. For instance, Salvador Dali extends the apparent dynamic range beyond what is physically possible by counter-shading along the shadow edges. As in our 3D unsharp masking, such counter-shading effects occur within the context of the depicted scene, not in the context of the 2D painting. Games could opt for approximations of our technique that have neglectable overhead and enhance gradients at strong contrast shadow edges, making game environments that look like those in Figure 9.

The temporal coherence of these enhancements is shown in the accompanying video. We also show that for simple deforming scenes without topology changes, our approach can be successfully used on deforming models.

A main benefit of our approach is its integration in the rendering pipeline and its allowance for interactive parameter adjustment, which is real-time when the original can be real-time rendered. In Table 1 we present timing numbers from our implementation us-



**Figure 8:** Our approach coherently enhances soft shadows from natural lighting, object shading changes and specular highlights. Two  $\lambda$  values lead to a subtle or more obvious enhancement; both help to emphasize the objects shapes and their spatial arrangement.



**Figure 9:** In a stylized game-like environment, enhancement can be visible to emphasize column shape and shadow direction.

ing an Nvidia GeForce 8800 GTX, on an Intel Core 2 Duo 6300 1.8 GHz. The main bottleneck is the speed of shadow rendering, not our smoothing step, as detailed in the supplemental material. In game scenarios, this is not the case, and with an optimized shader, we anticipate much higher framerates.

Scene & Vertices	Lighting	3D Unsharp	Original	Overhead
Cross (8 k)	Natural	10.9 FPS	12.4 FPS	16 %
Golfball (127 k)	Natural	17.9 FPS	31.3 FPS	43 %
Feet (57 k)	Natural	10.2 FPS	15.2 FPS	33 %
Chamfer (39 k)	Natural	8.3 FPS	10.7 FPS	22 %
Keys (152 k)	Point	15.2 FPS	63.0 FPS	76 %
Dice (74 k)	Point	15.6 FPS	63.0 FPS	75 %
Columns (119 k)	Point AO	28.3 FPS	63.2 FPS	55 %

**Table 1:** Timing numbers for 3D unsharp masking and percent of additional computation required for enhancement.

## 7 Perceived Effect of 3D Unsharp Masking

We have designed a psychological user study to confirm that 3D unsharp masking leads to superior and preferred renderings of greater contrast, and that our technique is easily controlled by the  $\lambda$  parameter. This study does not measure the effect in terms of easier shape or spatial organization understanding, but measures whether a general enhancement is seen and preferred. 15 naïve subjects without prior experience in computer graphics were presented the following diverse scenes: *Chamfer Plane*, *Dice*, *Feet* and *Keys* (Figures 8, 5, 1 and 6). For each scene and its set viewpoint, we generated stimuli at three  $\sigma$  values (small, medium and large) over the range

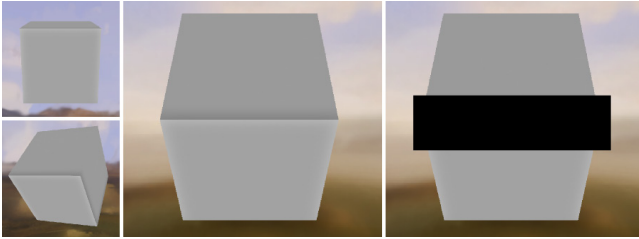
$\lambda = [0.0, 3.0]$ , sampled at stepsizes below visible threshold. To determine user response, we presented two images side-by-side: the unenhanced image and an adjustable enhanced image controlled by arrow keys to increase or decrease  $\lambda$ . Users were asked to select the 3 images that best responded to the following three questions: first, *When is a difference first visible?* ( $\lambda_{Low}$ ), second, *When is the difference too strong or objectionable?* ( $\lambda_{High}$ ), and third, *Which image has the best contrast?* ( $\lambda_{Best}$ ).

For analysis, we scale each set of  $\lambda$ s to JND (just-noticeable difference) units based on  $\lambda_{Low}$ . The overall ANOVA test revealed a significant main effect of the three questions ( $F(2, 26) = 69.63, p < .001$ ). Demonstrating the usability of our approach, we find that the  $\lambda$  parameter permits a wide range of acceptable enhancements ( $\approx 4$  JND units on average). These results are depicted in Figure 12, which shows the choices of  $\lambda_{Low}$ ,  $\lambda_{High}$  and  $\lambda_{Best}$  in JND. Additionally, we observe that the  $\sigma$  parameter did not significantly affect the choice of  $\lambda_{Best}$  ( $F(2, 28) = 0.69, p = .51$ ), however, it did impact  $\lambda_{High}$  for two of the scenes, see Figure 13. Holm-corrected pairwise statistical contrasts proved that users preferred visible enhancements ( $\lambda_{Best} > \lambda_{Low}: t(14) = 3.06, p < .01$ ) and that they avoided objectionable artifacts ( $\lambda_{Best} < \lambda_{High}: t = 5.53, p < .001$ ).

We confirm that as with all enhancement techniques, the preferred result is subjective. For most of the scenes, users preferred a strong enhancement of approximately twice the visibility threshold (*Feet*:  $\lambda_{Best} = 1.74, p < .01$ , *Dice*:  $\lambda_{Best} = 1.97, p < .01$ , *Chamfer Plane*:  $\lambda_{Best} = 1.84, p < .01$ ). For the *Keys* scene, users chose relatively subtle enhancements that were statistically indistinguishable from the lower threshold ( $t(14) = 1.28, p = .21$ ). Through a cluster-analysis (Ward, euclidean distance), we identify



**Figure 10:** The original rendering (left) is 3D unsharp masked using  $\lambda$  encoded by vertex colour to represent importance, producing imagery (right) that directs attention away from the statue’s robe.



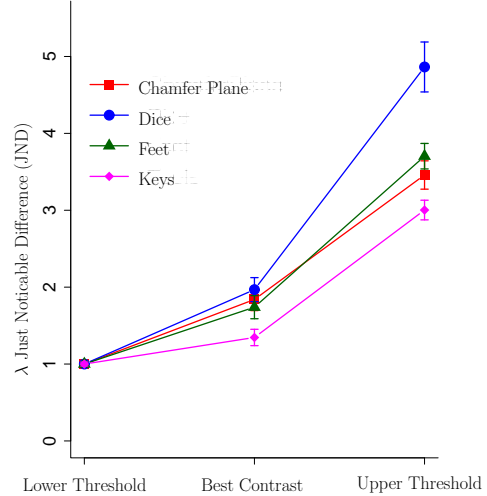
**Figure 11:** Our version of the Cornsweet illusion, made by 3D unsharp masking a uniformly grey cube lit from below.

two main groups of users: those who prefer weak enhancement (33%,  $\lambda_{\text{Best}} \approx 1.4$  JND) and those who prefer stronger (66%,  $\lambda_{\text{Best}} \approx 1.9$  JND). In general, we observe that an enhancement of double  $\lambda_{\text{Low}}$  in JND is considered to have “best contrast”, and that users tolerate lightness enhancements of up to four times the JND. The full set-up and analysis is detailed in supplemental material entitled “Experimental Validation of the Algorithm”.

## 8 Discussion

Without a proper  $\lambda$ , adverse percepts may occur for some viewers. The most common unintended effect is when the enhancement is perceived as a halo. However, this occurs only for extreme choices for  $\lambda$ . Material characteristics may also be enhanced. For instance, shiny objects may appear slightly shinier. We see this as an advantage that makes cues for material recognition more obvious. Our technique may enhance artifacts, for instance in soft shadows, by bringing below threshold errors above visible threshold. To prevent this if it occurs, one could consider separate rendering passes, with weaker  $\lambda$  for shadows than meshes. However, we find that this is not necessary when high quality rendering techniques are used.

Our smooth lighting, described in Section 5, is sensitive to the mesh tessellation and especially to the uniformity of length of edges.



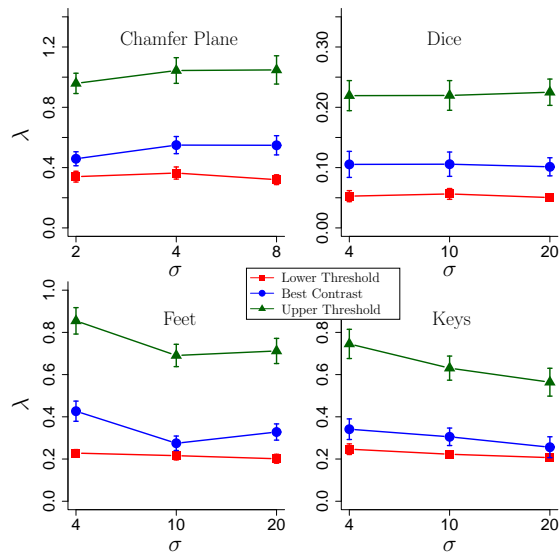
**Figure 12:** User chosen  $\lambda$ ’s in JND, averaged over  $\sigma$  for each scene. An enhancement of double  $\lambda_{\text{Low}}$  in JND has “best contrast”, tolerable enhancements are up to four times the JND. The error bars are standard errors of the mean (SEM).

Therefore, we apply mesh cleaning operations to all surfaces to refine their tessellation. However, the results from our perceptual study (Section 7) indicate that users do not prefer smaller  $\sigma$  values over larger, so very refined tessellation is not necessary. Our method does not smooth lighting across disconnected components, so our approach does not properly handle meshes that are not manifold where it is expected. Also, different mesh representations will result in slightly different enhancements.

We considered two technical alternatives to overcome mesh imposed limitations, a kernel method and a smooth volume method. We could reconstruct a smooth light signal from a cloud of lit points [Lehtinen et al. 2007], and a kernel smoothing method could be used to reconstruct smooth lighting. However, this method requires an acceleration structure like a  $kd$ -tree or a hash table to work efficiently. Creating such acceleration structures for animated scenes in realtime is notoriously difficult, whereas our method works for animated surfaces without modifications. Another option would be to smooth over a volume by rendering the lighting into the voxels of a 3D texture and building its MIP map. Such an approach requires a voxel resolution that can not be rendered at the rates our method achieves using surfaces, wasting many samples on empty space. Both approaches ignore surface constraints, meaning that light may bleed through thin structures or at discontinuities unless a high number of samples is used. We found these approaches to have severe drawbacks that did not outweigh their benefit of working over disconnected meshes.

## 9 Conclusion

We have presented a new approach for increasing the local contrast of a 3D scene. Our holistic approach is perceptually motivated and preserves coherence with the entire depicted scene. By unsharp masking light gradients over the mesh surface we avoid temporal artifacts introduced by image-based approaches. Our enhancement better depicts shape details, provides clearer separation between objects and deepens the impression of a scene’s overall dynamic range.



**Figure 13:** The user chosen lower threshold ( $\lambda_{Low}$ ), preferred contrast ( $\lambda_{Best}$ ), and upper threshold ( $\lambda_{High}$ ) as a function of  $\sigma$ . Notice that  $\sigma$  has only a small effect on  $\lambda_{Best}$ .

Compared to existing approaches, we enhance gradients that occur without depth or geometry discontinuities. We automatically adapt the contrast signal to the 3D orientation and perspective foreshortening in the scene. By calculating contrast in 3D, we prevent image space and view-dependent incoherencies that are suffered by image space enhancement and using buffered information like depth. Additionally, we take care to produce pleasing images by adjusting lightness only and correcting colour saturation to prevent the effect of whitening or blackening. We also take advantage of texture masking to add more contrast where it is less visible. Our approach is integrated so that adjustment is simple and responsive, and does not require separate processing, as is the case with techniques that manipulate the mesh or collections of renderings.

The applications of 3D unsharp masking are diverse and include visualization, model analysis, presentation of 3D designs, education, filmmaking and game rendering. Related to the field of perception, 3D unsharp masking could be used to create stimuli for psychophysical experiments on the Cornsweet illusion and other local contrast effects. For other forms of enhancement,  $\lambda$  can be controlled by external 2D or 3D data, such as HDR contrast, lighting intensity, etc. As an NPR tool, our approach could achieve contrast effects, and could be used for enhancing colour, texture or other attributes instead of lightness.

In the future, we would like to test the coherence of 3D unsharp masking on deformable meshes, especially under topological changes and self-intersections. We would investigate a multiscale approach that allows the smoothing to adapt locally by varying  $\sigma$  over the scene. We foresee that the approach could be applied to volume rendering to enhance visualizations by smoothing over a lit volume instead of a lit surface. Lastly, we would like to measure the psychophysical effect of 3D unsharp masking on shape and spatial organization comprehension.

## References

BADAMCHIZADEH, M. A., AND AGHAGOLZADEH, A. 2004. Comparative study of unsharp masking methods for image en-

hancement. In *International Conference on Image and Graphics (ICIG'04)*, 27–30.

BRUCKNER, S., AND GRÖLLER, M. E. 2007. Enhancing depth-perception with flexible volumetric halos. *IEEE Transactions on Visualization and Computer Graphics* 13, 6, 1344–1351.

CAVANAGH, P., AND LECLERC, Y. 1989. Shape from shadows. *Journal of Experimental Psychology: Human Perception and Performance* 15, 1, 3–27.

CIGNONI, P., SCOPIGNO, R., AND TARINI, M. 2005. A simple normal enhancement technique for interactive non-photorealistic renderings. *Computer & Graphics* 29, 1.

CORNSEWET, T. 1970. *Visual Perception. Chapter II: The Experiment of Hecht, Schlaer, and Pirenne*. Academic Press.

DECARLO, D., FINKELSTEIN, A., AND RUSINKIEWICZ, S. 2004. Interactive rendering of suggestive contours with temporal coherence. In *International Symposium on Non-Photorealistic Animation and Rendering (NPAR 2004)*, 15–145.

DECORO, C., COLE, F., FINKELSTEIN, A., AND RUSINKIEWICZ, S. 2007. Stylized shadows. In *International Symposium on Non-Photorealistic Animation and Rendering (NPAR 2007)*, 77–83.

FLEMING, R., TORRALBA, A., AND ADELSON, E. H. 2004. Specular reflections and the perception of shape. *Journal of Vision* 4, 9, 798–820.

KINGDOM, F., AND MOULDEN, B. 1988. Border effects on brightness: A review of findings, models and issues. *Spatial Vision* 3, 255–262.

KRAWCZYK, G., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2007. Contrast restoration by adaptive countershading. In *Computer Graphics Forum (Proceedings of Eurographics 2007)*, vol. 26.

LANDIS, H. 2002. RenderMan in Production. In *ACM SIGGRAPH 2002 Course 16*.

LEHTINEN, J., ZWICKER, M., KONTKANEN, J., TURQUIN, E., SILLION, F., AND AILA, T. 2007. Meshless finite elements for hierarchical global illumination. Tech. Rep. TML-B7, Helsinki University of Technology, may.

LUFT, T., COLDITZ, C., AND DEUSSEN, O. 2006. Image enhancement by unsharp masking the depth buffer. *ACM Trans. Graph.* 25, 3, 1206–1213.

PURVES, D., SHIMPI, A., AND LOTTO, R. B. 1999. An empirical explanation of the cornsweet effect. *J. of Neurosci* 19, 8542–8551.

RUSINKIEWICZ, S., BURNS, M., AND DECARLO, D. 2006. Exaggerated shading for depicting shape and detail. *ACM Trans. Graph. SIGGRAPH* 25, 3, 1199–1205.

SMITH, K., LANDES, P.-E., THOLLOT, J., AND MYSZKOWSKI, K. 2008. Apparent greyscale: A simple and fast conversion to perceptually accurate images and video. *Computer Graphics Forum (Proceedings of Eurographics 2008)* 27, 2.

TARINI, M., CIGNONI, P., AND MONTANI, C. 2006. Ambient occlusion and edge cueing for enhancing real time molecular visualization. *IEEE Transactions on Visualization and Computer Graphics* 12, 5, 1237–1244.

TAUBIN, G. 1995. A signal processing approach to fair surface design. In *ACM Trans. Graph. SIGGRAPH*, 351–358.

ZENG, W., DALY, S., AND LEI, S. 2001. An overview of the visual optimization tools in JPEG 2000. *Signal Processing: Image communication Journal* 17, 1, 85–104.