



Flow-Guided Warping for Image-Based Shape Manipulation

Romain Vergne, Pascal Barla, Georges-Pierre Bonneau, Roland Fleming

► To cite this version:

Romain Vergne, Pascal Barla, Georges-Pierre Bonneau, Roland Fleming. Flow-Guided Warping for Image-Based Shape Manipulation. ACM Transactions on Graphics, 2016, 34 (4), pp.Article No. 93. 10.1145/2897824.2925937 . hal-01307571

HAL Id: hal-01307571

<https://inria.hal.science/hal-01307571>

Submitted on 29 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Flow-Guided Warping for Image-Based Shape Manipulation

Romain Vergne^{1,2*} Pascal Barla^{2†} Georges-Pierre Bonneau^{1,2‡} Roland W. Fleming^{3§}

Univ. Grenoble Alpes, CNRS¹ Inria² Justus Liebig Universität Gießen³

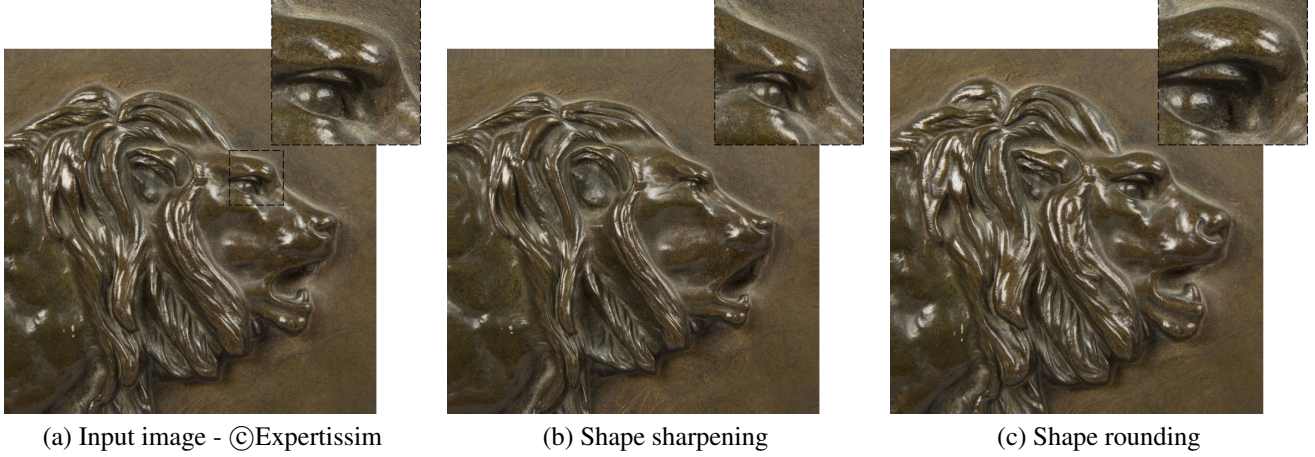


Figure 1: Our warping technique takes as input (a) a single image (Jules Bennes, after Barye: “walking lion”) and modifies its perceived surface shape, either making it sharper in (b) or rounder in (c).

Abstract

We present an interactive method that manipulates perceived object shape from a single input color image thanks to a warping technique implemented on the GPU. The key idea is to give the illusion of shape sharpening or rounding by exaggerating *orientation patterns* in the image that are strongly correlated to surface curvature. We build on a growing literature in both human and computer vision showing the importance of orientation patterns in the communication of shape, which we complement with mathematical relationships and a statistical image analysis revealing that *structure tensors* are indeed strongly correlated to surface shape features. We then rely on these correlations to introduce a flow-guided *image warping* algorithm, which in effect exaggerates orientation patterns involved in shape perception. We evaluate our technique by 1) comparing it to ground truth shape deformations, and 2) performing two perceptual experiments to assess its effects. Our algorithm produces convincing shape manipulation results on synthetic images and photographs, for various materials and lighting environments.

Keywords: Image warping, enhancement, shape perception.

Concepts: •Computing methodologies → Image manipulation;

1 Introduction

When presented with the image of an object as in Figure 1a, we effortlessly perceive its shape, material, and the way it is lit by the surrounding environment. This is a formidable achievement. Accomplishing this feat through inverse optics would be extremely challenging as the necessary inferences are severely under-constrained. However, recent work in visual perception suggests an alternative, and perhaps more reasonable, explanation of how we perceive

and process object appearance. In particular, Fleming et al. [2004; 2011] propose that image properties are indicative of object shape and material properties, simply because they are sufficiently *correlated* together. This opens an exciting avenue of research for computer graphics: by identifying the appropriate correlations, one could elicit perceptually compelling changes in object appearance by mere manipulations of its image, without having to solve the difficult inverse optics problems of reconstructing shape, material or lighting.

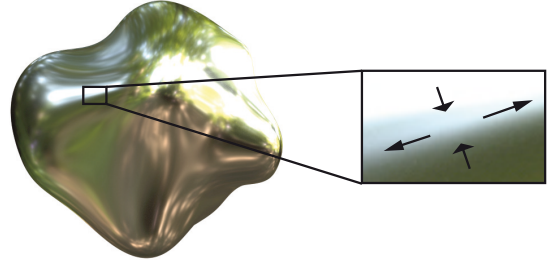


Figure 2: Perception studies suggest that orientation patterns (here due to the compression and stretching of reflections on a shiny 3D object) provide strong visual cues to surface shape.

In this paper, we identify and exploit correlations between image gradient structure and the corresponding object *shape*. The key idea is to consider that orientation patterns in the image are mainly due to surface curvature. This is illustrated in Figure 2, where orientation patterns are due to compression and stretching of the reflected environment. Recent work in human and computer vision stresses the importance of such patterns in the perception of shape (e.g., [Ben-Shahar and Zucker 2001; Fleming et al. 2004; Fleming et al. 2011]). We complement these perceptual findings with mathematical relationships and a statistical analysis in Section 3.

Equipped with these correlations between image and shape properties, in Section 4 we introduce a novel image deformation algorithm that runs in real-time on the GPU. It works in two stages: first

*romain.vergne@inria.fr

†pascal.barla@inria.fr

‡georges-pierre.bonneau@inria.fr

§roland.w.fleming@psychol.uni-giessen.de

the scale of shape-related features are estimated per pixel; then the image is warped, guided by orientation patterns at their respective scale. As shown in Figure 1b,c, such an image warping produces a vivid impression of shape changes, making the object appear either sharper or rounder depending on warping parameters. We evaluate our algorithm by comparing it to ground truth shape deformations in Section 4.3 and by conducting perceptual validation through a pair of psychophysics experiments in Section 5. We demonstrate that our method works for a wide range of object materials and environment lightings in Section 6.

2 Related work

Image manipulation Early methods for the manipulation of content in images were based on partial differential equations, such as anisotropic diffusion [Perona and Malik 1990] for edge-preserving smoothing, or shock filtering [Osher and Rudin 1990] for detail enhancement. Recent methods often make use of multi-scale decompositions to manipulate image details while preserving edges (e.g., [Farbman et al. 2008; Paris et al. 2011]). However, the focus of these methods is on the editing and filtering of low-level image features, rather than on modifying the perceived attributes of the depicted objects, such as their shape or material.

Fattal et al. [2007] specifically target shape details by using a multi-scale decomposition on multi-light image collections. Boyadzhiev et al. [2015] focus instead on material editing in a *single* input image, which is permitted by a careful perceptual analysis of subbands in the multi-scale decomposition. Techniques that magnify nearly imperceptible details in images also work in a multi-scale fashion. They may either magnify deviations through time [Wu et al. 2012; Wadhwa et al. 2013], with respect to ideal geometries [Wadhwa et al. 2015] or relative to an ideally repeated pattern [Dekel et al. 2015].

A parallel line of work stems from the seminal intrinsic decomposition technique of Barrow et al. [1978], which consists in separating an image into its reflectance and shading components. Modern solutions to the intrinsic decomposition problem may require user annotations on a single input image [Bousseau et al. 2009] for increased accuracy. This still assumes that the input image is a simple product between reflectance and illumination components. The work of Carroll et al. [2011] improves on this limitation by further decomposing illumination into direct and indirect components. Other methods propose user-assisted decompositions that address more complex material properties: Dong et al. [2011] propose a workflow for separating diffuse and specular terms in color textures; Yeung et al. [2011] present a method to separate a transparent object from its background, and re-compose it on a different background. Both methods are user-assisted.

The work of Yeung et al. [2011] is most related to ours in that it makes use of image warping to mimic the effect of refraction due to the transparent object on the new background. Our approach instead applies warping directly to the image of a foreground opaque object to modify its perceived shape (without requiring user annotations). It also makes use of a multi-scale analysis to identify the scales of shape features per-pixel.

3D reconstruction An alternative approach to manipulate image content is first to reconstruct 3D information, then modify it and finally re-render (part of) the image.

The technique of Khan et al. [2006] makes the simplifying assumption that surface depth is inversely proportional to image intensity, and computes normals from this surface via differentiation. An environment lighting is recovered by extracting and duplicating

the background image. These approximations prove sufficient for image-based material editing in many practical cases: they permit re-texturing an object or even making it transparent. The method has been extended by Gutierrez et al. [2008] to include caustics by treating the recovered front object surface as a collection of thin lens segments. Once depth and normal buffers become available, other manipulation techniques are possible such as Unsharp Masking [Luft et al. 2006] or Surface Flows [Vergne et al. 2012]. The latter approach specifically extends the method of Khan et al. to directly manipulate highlights in the image, in a way that ensures the deformations conform to the recovered normals.

Approaches based on the recovery of depth or normal buffers suffer from a common limitation: they cannot rotate the object or freely reposition it in space. In contrast, methods that estimate full 3D data [Chen et al. 2013; Kholgade et al. 2014] must restrict the category of objects they recover. In the work of Chen et al. [2013], users assist object extraction by identifying shape components using strokes. Strokes are then automatically adjusted by the algorithm, and used in a sweeping technique to obtain 3D shapes. As a result, recovered objects are restricted to collections of generalized cylinders and cuboids. In the work of Kholgade et al. [2014], 3D shape is recovered with the help of stock 3D models. Recovered objects are thus limited by the availability of similar 3D models, as well as by symmetry priors.

Our approach altogether avoids the reconstruction of 3D information by identifying image properties directly related to shape, which are then manipulated through warping. However, it shares the limitation of methods based on normal or depth buffers as it does not permit 3D object transformations such as rotations.

Shape perception Theories of shape-from-shading (e.g., [Horn and Brooks 1989]) and shape-from-specularities (e.g., [Oren and Nayar 1997; Savarese et al. 2004]) have for the most part relied on computational models. Shape-from-shading models often make the assumption that there is a main light source direction that must be estimated before shape can be recovered. In shape-from-specularities, the hypothesis is similar: the way a nearby object is distorted by shape through reflection cannot be exploited if the neighboring object is unknown. These are drastic constraints that are often not verified in real-world scenarios: the environment may contain multiple strong light sources and reflected objects may be out of sight.

An alternative approach consists in relying on orientation patterns that are directly observable *in the image*. Early perceptual evidence comes from the importance of patterns of isophotes in the perception of shape-from-shading [Koenderink and van Doorn 1980]. This line of research has evolved in the computer vision community as well, such as in the work of Zucker and colleagues (e.g., [Ben-Shahar and Zucker 2001]) who studied the similar concept of “shading flows”. Orientation patterns in images play a similar role in the perception of shape-from-specularities [Fleming et al. 2004]. They are due to compression and stretching of reflections caused by surface curvature. Similar types of patterns may be observed with surface texture markings [Fleming et al. 2011], which suggests a central role of orientation patterns in the perception of shape from various cues.

Orientation patterns are relevant to shape perception because they are sufficiently correlated to physical shape, assuming generic viewpoint and light source configurations [Freeman 1994]. However, in some cases orientation patterns may provide information that departs from physical shape, and perceptual studies show that in these cases perceived shape differs as well. For instance, strong directional lighting may distort the perception of shape from shading [Caniard and Fleming 2007]. The choice of material may also

modulate the perception of 3D shape [Wijntjes et al. 2012; Mooney and Anderson 2014].

We take inspiration from modern models of shape perception and study the correlation between shape and image properties in Section 3. We then use these correlations to design an image deformation algorithm that gives the illusion of shape changes in images. However, our goal is not to manipulate physical, but perceived shape, which we expect to vary with lighting and materials.

3 Image analysis

We begin with a mathematical analysis of the orientation patterns produced by the rendering of a 3D object. Our goal here is to provide a quantitative assessment of the relationships between surface shape and orientation patterns in the image. It is intended to complement the vision literature on the subject, as well as to provide a justification for the warping algorithm of Section 4.

3.1 Definitions

We consider a simplified model of image formation in the neighborhood \mathcal{P} of a pixel of interest \mathbf{p} (see Figure 3 left). We assume \mathcal{P} corresponds to surface points that lie on the same 3D object (i.e., no occluding contour crosses \mathcal{P}). The image intensity (for a single color channel) of a neighboring pixel $\mathbf{q} \in \mathcal{P}$ is then defined by:

$$\begin{aligned} I : \mathcal{P} &\rightarrow [0, 1] \\ \mathbf{q} &\mapsto M(L_r(\mathbf{q}, \omega_o)), \end{aligned}$$

where M is a tone mapping operator and L_r is the radiance reflected from the object surface according to [Kajiya 1986]:

$$L_r(\mathbf{q}, \omega_o) = \int_{\Omega_n} \rho_n(\mathbf{q}, \omega_i, \omega_o) L_i(\mathbf{q}, \omega_i) \max(0, \omega_i \cdot \mathbf{n}(\mathbf{q})) d\omega_i,$$

with ω_o and ω_i the outgoing view and incoming light directions, \mathbf{n} the surface normal at \mathbf{q} , Ω_n the upper hemisphere, ρ_n the material (Bidirectional Reflectance Distribution Function) and L_i the incoming radiance.

We further assume that *spatial* variations of ω_o , ρ_n and L_i are negligible inside \mathcal{P} . In particular, this means that we consider a locally orthographic camera, a homogeneous material and a distant illumination. As a result, the only variable that depends on \mathbf{q} is the normal $\mathbf{n}(\mathbf{q})$, which allows us to write (see Figure 3 right):

$$I(\mathbf{q}) \approx S(\mathbf{m}(\mathbf{q})), \quad (1)$$

where $\mathbf{m} = [m_0 \ m_1]^T = [\frac{-n_0}{n_2} \ \frac{-n_1}{n_2}]^T$ is the surface slope corresponding to the normal $\mathbf{n} = [n_0 \ n_1 \ n_2]^T$ expressed in screen-space; and S is a shading function that encapsulates both reflected radiance and tone mapping. We use surface slopes since they have an exact correspondence with front-facing unit normals, while being easier to relate to surface shape, as explained later on.

Equation 1 thus expresses image intensity I inside \mathcal{P} by a local shading function S applied to surface slopes \mathbf{m} . In the following, we *visualize* such shading functions using Lit Spheres [Sloan et al. 2001]; they are not equivalent though, as they are related by a normal-to-slope transform. In Figure 4 we have rendered a single object with several Lit Spheres exhibiting various shading directions and frequencies. In this case, Equation 1 is exact and the same shading function is applied everywhere. More realistic rendering scenarios using global illumination are considered in Section 3.3.

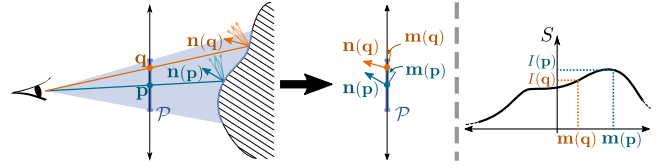


Figure 3: The image intensity I for pixels \mathbf{q} in a neighborhood \mathcal{P} around a pixel of interest \mathbf{p} (left) is locally approximated by a shading function S applied to surface slopes \mathbf{m} (right), with surface slopes being directly related to screen-space normals \mathbf{n} .

3.2 Mathematical relationships

Recent findings in human and computer vision show that orientation patterns in the image constitute visual cues to object shape [Ben-Shahar and Zucker 2001; Fleming et al. 2004; Fleming et al. 2011]. Our goal is to draw explicit relationships between such patterns and object shape in the context of Equation 1. A common approach to identify orientation patterns is to compute structure tensors [Bigun and Granlund 1986; Brox et al. 2006]:

$$\mathbf{T}_I(\mathbf{p}) = \int_{\mathcal{P}} w(\|\mathbf{p} - \mathbf{q}\|) \nabla I(\mathbf{q}) \nabla I(\mathbf{q})^T d\mathbf{q}, \quad (2)$$

where w is usually a Gaussian kernel truncated on \mathcal{P} . An important property of the structure tensor is its positive semi-definiteness. In particular, the maximum eigenvector identifies the dominant local image orientation irrespective of its direction.

The image gradient inside Equation 2 is obtained by differentiating Equation 1, yielding $\nabla I(\mathbf{q}) = \nabla \mathbf{m}(\mathbf{q}) \nabla S(\mathbf{m})$. The surface slope gradient $\nabla \mathbf{m} = [\nabla m_0 \ \nabla m_1]$ is by definition equivalent to the hessian of the surface depth. It is related to surface curvature, but also depends on the positioning of the surface with respect to the viewpoint, as already pointed out by Vergne et al. [2012]. We call *shape features* the properties of $\nabla \mathbf{m}$ across the image. The shading gradient ∇S incorporates angular variations due to both lighting and material. We may thus expect its magnitude to be higher for specular materials and sharp lighting environments.

Replacing ∇I in Equation 2 by its local approximation yields:

$$\mathbf{T}_I(\mathbf{p}) = \int_{\mathcal{P}} w(\|\mathbf{p} - \mathbf{q}\|) \nabla \mathbf{m}(\mathbf{q}) \nabla S(\mathbf{m}) \nabla S(\mathbf{m})^T \nabla \mathbf{m}(\mathbf{q})^T d\mathbf{q}. \quad (3)$$

In order to relate \mathbf{T}_I to shape features, we also compute the structure tensor \mathbf{T}_m on surface slopes:

$$\mathbf{T}_m(\mathbf{p}) = \int_{\mathcal{P}} w(\|\mathbf{p} - \mathbf{q}\|) \nabla \mathbf{m}(\mathbf{q}) \nabla \mathbf{m}(\mathbf{q})^T d\mathbf{q}. \quad (4)$$

The difference between the two tensors lies in the $\nabla S \nabla S^T$ term. By performing an eigen decomposition, we obtain:

$$\nabla S(\mathbf{m}) \nabla S(\mathbf{m})^T = \|\nabla S(\mathbf{m})\|^2 \mathbf{R}_{\phi_S} \mathbf{P}_0 \mathbf{R}_{-\phi_S}, \quad (5)$$

where \mathbf{R} is a 2D rotation matrix, ϕ_S is the angle between $\nabla S(\mathbf{m})$ and $\mathbf{e}_0 = [1, 0]^T$ and $\mathbf{P}_0 = \mathbf{I} - \mathbf{e}_1 \mathbf{e}_1^T$ is the projection on \mathbf{e}_0 .

Inserting Equation 5 into Equation 3 and comparing it to Equation 4 sheds light on two important differences between \mathbf{T}_I and \mathbf{T}_m : the $\|\nabla S(\mathbf{m})\|^2$ term *modulates* each surface slope gradient based on the shading gradient magnitude, while the $\mathbf{R}_{\phi_S} \mathbf{P}_0 \mathbf{R}_{-\phi_S}$ term *projects* each surface slope gradient in the shading gradient direction. When combined through integration, the projected and modulated surface slope gradients yield a tensor \mathbf{T}_I that is strikingly similar to \mathbf{T}_m , as shown in the third and fourth rows of Figure 4. This suggests that \mathbf{T}_I contains reliable visual cues to surface

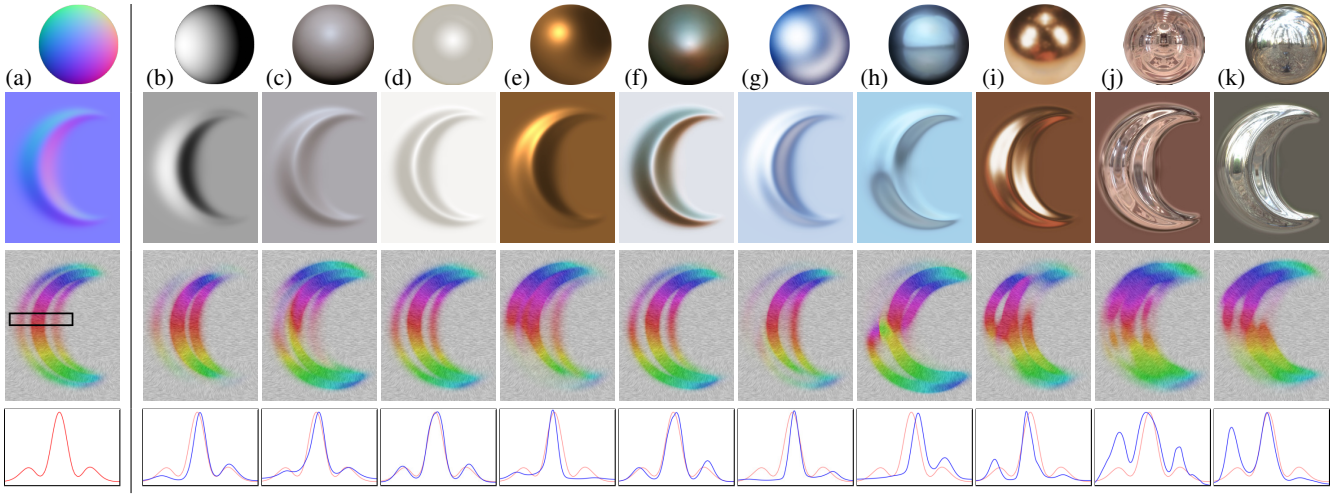


Figure 4: Comparisons between surface-based (a) and image-based (b-k) structure tensors. Rendered images (second row) make use of various *Lit Spheres* (first row). Structure tensors (third row) are visualized using a combination of two techniques. Colors are expressed in HSV space, with their hue angle corresponding to the angle made by the maximum eigen vector with the horizontal axis, and their opacity corresponding to the maximum eigen value. Directional patterns are obtained by applying a line integral convolution [Stalling and Hege 1995] to a noise texture in the direction of the maximum eigen vector. Maximum eigen values along a horizontal scanline (bottom row) help compare surface (red curve) versus image features (blue curves). Masking effects occur due to near-constant regions of the *Lit Spheres*: in (e) due to the brown base shading; in (g) due to the wide highlight; in (h) due to the “sky” region. Directional masking affects the top and bottom of (b) since shading mostly exhibits horizontal variations. Exaggerations are mostly apparent in (h,i,k) where shading exhibits hard edges and high frequencies, while distortions are clearly seen in (h,i) (see the shift of the central peak in the bottom row).

shape, across a wide range of material and lighting. Nevertheless, \mathbf{T}_I and \mathbf{T}_m differ in a few subtle, yet important, details that appear in Figure 4. We describe them below.

Masking In regions corresponding to near-uniform portions of S (such as with an overcast sky or a dim diffuse shading), $\|\nabla S\| \approx 0$ for all \mathbf{m} . As a result, $\mathbf{T}_I \approx \mathbf{0}$ and shape features cannot appear in these image regions. This is called masking in the vision literature, and may be due to either lighting or material [Wijntjes et al. 2012].

Directional masking When the shading gradient is strongly directional (as with a collimated lighting), all surface slope gradients are projected in roughly the same direction. For instance, when $\phi_S \approx 0$ for all \mathbf{m} , we get $\mathbf{T}_I \approx \mathbf{T}_{m_0}$ (i.e., the tensor due to surface slopes in the horizontal direction). This leads to directional masking effects [Fleming et al. 2004] in places where $\nabla m_0 \approx 0$.

Exaggeration and distortion. If instead shading gradients of strong magnitude are roughly aligned with surface slope gradients, shape will tend to be exaggerated. For instance it will happen when most surface variations are due to ∇m_0 and $\phi_S \approx 0$. This exaggeration effect is usually stronger for specular materials [Mooney and Anderson 2014]. It may also lead to distortion effects [Caniard and Fleming 2007] when the distribution of shading gradient magnitudes is shifted from the origin in S .

3.3 Statistical analysis

Even though we are mostly interested in perceived shape, it remains to be shown that the latter is sufficiently similar to physical shape in more general cases than the one in Figure 4. For this purpose, we conducted a statistical analysis on a collection of objects rendered with global illumination using Mitsuba [Jakob 2010]. We used four blobby objects of increasing shape complexity, each made of four types of material and rendered in four different environment lightings. Figure 5 shows a subset of these data; the full dataset is pre-

sented in a supplementary document.

For each object, we computed its surface-based tensor along with image-based tensors for each condition. Next, we assembled a pair of similarity matrices relating all possible conditions, yielding two 17×17 symmetric matrices per object. The first matrix measures orientation similarity averaged over all pixels. It is computed as the dot product between the two maximum eigen vectors of each tensor. The second matrix measures the cross-correlation of the two maximum eigen values (i.e., the best correlation among shifted versions of the image), again averaged over all pixels. We use it instead of correlation to account for the occurrence of the distortion effects, which tend to shift shading patterns. Similarity matrix pairs for individual objects are provided in supplementary material.

Figure 6 summarizes our results using a pair of matrices where similarities have been averaged over all four objects. The average similarity of orientations across all objects and conditions equals 0.87, with a standard deviation of 0.22; while the average cross-correlation of magnitudes across all objects and conditions equals 0.84, with a standard deviation of 0.24 (with a p-value $\ll 0.05$).

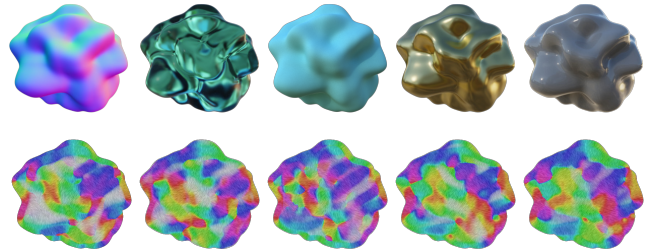


Figure 5: A small subset of our statistical analysis data. Top row: object normals and four rendered images using different materials in the same illumination environment. Bottom row: the corresponding surface-based tensor and four image-based tensors. The full dataset contains sixteen rendered images (four materials, four illumination environments) per object and four objects in total.

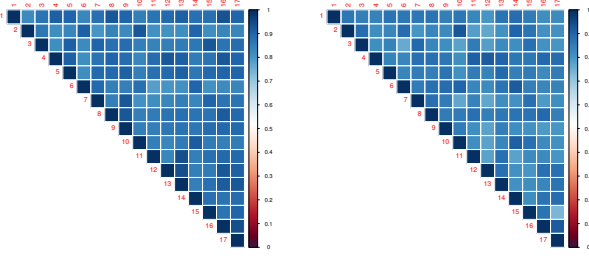


Figure 6: Similarity matrices averaged over all four objects: each entry stores the average orientation similarity (left) and average eigen value cross-correlation (right) over all pixels. The first row of each matrix shows similarities between image-based tensors and the surface-based tensor; the other rows show similarities among image-based tensors in all rendering configurations.

These statistics demonstrate that image-based tensors are significantly correlated not only to the surface-based tensor (first row of each matrix), but also to each other (all other rows).

4 Image deformation

We have shown in the previous section that image-based structure tensors are significantly correlated to shape features, and likely constitute visual cues to perceived shape. Our goal now is to make use of these correlations to deform an image so as to give the illusion of shape changes.

Until now we have implicitly assumed that shape features occur at a single global scale in the image. However, when considering images of complex objects, shape features may occur at different scales in different locations in the image. We thus begin in Section 4.1 by explaining how we identify pertinent scales at which image-based structure tensors will be computed. We next describe our warping algorithm in Section 4.2, compare it to surface-based deformation and provide a geometric interpretation of its effects in Section 4.3. Our GPU implementation is detailed in Section 4.4.

4.1 Scale selection

The process of identifying pertinent scales per pixel is called automatic scale selection in the scale space literature [Lindeberg 1998]. We use a simple scale selection mechanism that identifies a single scale σ_I per pixel \mathbf{p} using a weighted combination of the form:

$$\sigma_I(\mathbf{p}, \gamma) = \frac{\int \sigma^{2\gamma} \beta_I(\mathbf{p}, \sigma) d\sigma}{\int \beta_I(\mathbf{p}, \sigma) d\sigma}, \quad (6)$$

where $\sigma \in (0, +\infty)$ stands for the scale dimension, $\gamma \in \mathbb{R}^*$ is a user-controlled parameter and β_I is a scale-space image-based weighting function. The γ parameter directly comes from Lindeberg et al. [1998]: it permits biasing the selected scale toward the smallest scale when it is close to 0 (or towards the biggest scale when greater than 1). We choose $\gamma = 0.5$ as a default value, since Equation 6 then boils down to a linear combination of weights. For the weight, we use $\beta_I(\mathbf{p}, \sigma) = \text{Tr}(\mathbf{T}_I(\mathbf{p}))$ where the structure tensor \mathbf{T}_I from Equation 2 is implicitly computed within a Gaussian window w_σ of standard deviation σ , truncated to $\mathcal{P}(\sigma) = \{\mathbf{q} \mid \|\mathbf{p} - \mathbf{q}\| < 3\sigma\}$. The same approach is easily applied to images of surface slopes: the trace of the tensor \mathbf{T}_m from Equation 4 is then used to build a scale-space weight function β_m , yielding σ_m .

The two resulting scale maps σ_m and σ_I are compared in Figure 7 for various shading functions. As before, we use various Lit

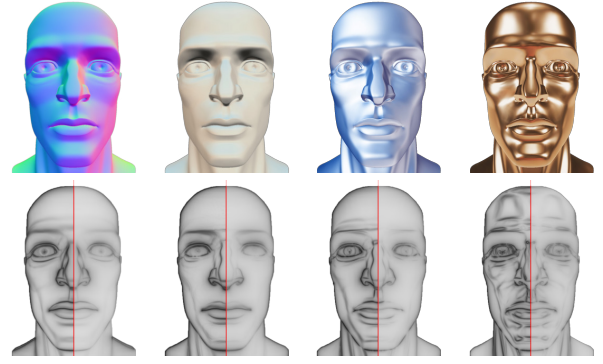


Figure 7: Top: a 3D object rendered with 3 Lit Spheres. Bottom: surface- and image-based scale maps obtained with $\gamma = \{0.1|0.7\}$ (the darker the shade, the smaller the selected scale).

Spheres for rendering the entire object. It is easily shown that for an affine shading function S , we have $\sigma_I = \sigma_m$. For more complex combinations of materials and lighting environments, the two scale maps depart from each other. In particular, high-frequency shading functions exhibit more pronounced differences at low values of γ . Nevertheless, we have found this simple scale selection mechanism to provide reasonable results in practice, as shown in Section 6.

4.2 Deformation algorithm

The general idea of our image deformation algorithm is to exaggerate the orientation patterns found in an input image through warping. For instance, when such patterns are due to either compression or stretching of reflections as in Figure 2, warping should exaggerate compression and stretching so as to give the illusion of shape change. In the following, we use the maximum eigenvalue λ_I^{\max} of \mathbf{T}_I to guide warping since it provides the magnitude of compression or stretching along the dominant local orientation.

Our image deformation procedure is summarized in Algorithm 1 and illustrated in Figure 8. It takes as input an image I , a scale bias γ and a warping parameter α , and returns a warped image. We optionally take a gray-scale mask as input to control warping locally. We describe our algorithm step-by-step below.

Algorithm 1 Image deformation

```

1: procedure DEFORM( $I, \alpha, \gamma$ )
2:    $\sigma_I \leftarrow \text{scaleSelection}(I, \gamma)$ 
3:    $\mathbf{T}_I \leftarrow \text{structureTensor}(I, \sigma_I)$ 
4:    $\mathbf{T}_I \leftarrow \text{enforceContinuation}(\mathbf{T}_I, \sigma_I)$ 
5:    $\lambda_I^{\max} \leftarrow \text{maxEigenValue}(\mathbf{T}_I)$ 
6:    $\lambda_I^{\max} \leftarrow \text{localAdaptation}(\lambda_I^{\max}, \sigma_I)$ 
7:    $\nabla \lambda_I^{\max} \leftarrow \text{guidingFlow}(\lambda_I^{\max}, \sigma_I)$ 
8:   for each pixel  $\mathbf{p}$  do
9:      $W(\mathbf{p}) \leftarrow \text{warp}(\mathbf{p}, \nabla \lambda_I^{\max}, \sigma_I, \alpha)$ 
10:   $W \leftarrow W * w_{\sigma_I}$ 
11:  return  $I \circ W$ 

```

Structure tensor The algorithm starts (lines 2-3) by computing a scale map σ_I using Equation 6, then a structure tensor image \mathbf{T}_I at the selected scales using Equation 2. However, as discussed in the previous section, both shape and shading features act on \mathbf{T}_I . Taking inspiration from Ben Shahr et al. [2001], we enforce orientation continuity (line 4). This is done by smoothing \mathbf{T}_I in the direction of the minimum eigen vector, assuming variations along these directions are mainly due to shading. An illustration is given in the inset figure, where

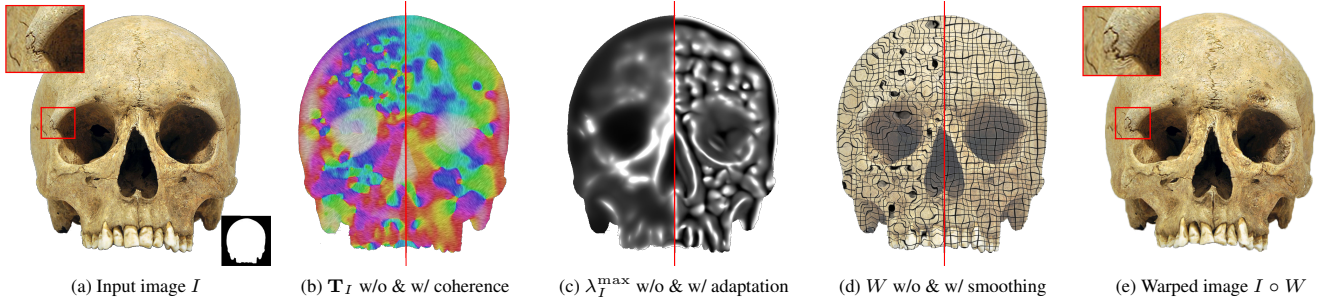
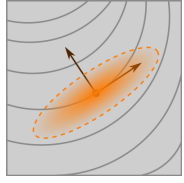


Figure 8: We illustrate our warping algorithm starting from an input image I in (a). Structure tensors \mathbf{T}_I is visualized in (b) before and after continuation. Their maximum eigen values λ_I^{\max} are shown in (c) before and after local adaptation. The warped coordinates W (using $\alpha = -1$ and guided by $\nabla \lambda_I^{\max}$) are visualized using a distorted grid in (d). They are used to produce the final warped image $I \circ W$ in (e).

streamlines (in grey) follow minimum eigen vector directions and the smoothing kernel (in orange) is an elliptic Gaussian. The approach is similar in practice to the image stylization method of Kyprianidis et al. [2008]. We use a filtering kernel of standard deviation σ_I and anisotropy $\frac{\lambda_I^{\max} - \lambda_I^{\min}}{\lambda_I^{\max} + \lambda_I^{\min}}$, where λ_I^{\max} and λ_I^{\min} are the maximum and minimum eigen values of \mathbf{T}_I . Tensor fields before and after continuation enforcement are visualized in Figure 8b.



Coherence enforcement

Guiding flow The next three steps are devoted to computing a unit gradient field that will serve as a guiding flow for warping. We start (line 5) by extracting the maximum eigen value λ_I^{\max} from \mathbf{T}_I . The purpose of local adaptation (line 6) is to make sure all shape features get exaggerated irrespective of their scale. To this end, we center and rescale λ_I^{\max} at a pixel \mathbf{p} based on neighbor eigen values using $\frac{\lambda_I^{\max} - E[\lambda_I^{\max}]}{\text{Std}[\lambda_I^{\max}]}$, where E and Std denote the expectation and standard deviation inside $\mathcal{P}(\sigma_I)$. We call this process local adaptation because of its similarity with lightness adaptation in vision. The non-adapted and adapted maximum eigen value images are visualized in Figure 8c. The guiding flow is then computed (line 7) by convolving λ_I^{\max} with a 1st-order Gaussian derivative of scale σ_I and normalizing the resulting gradient field.

Image warping Now that we are equipped with a guiding flow, we are ready to compute an image W of warped pixel coordinates (lines 8-9). The warping function is detailed in Algorithm 2. It takes an initial pixel position \mathbf{p} , a flow $\nabla \lambda$, a scale map σ and a warping parameter α . The main idea is to walk along $\nabla \lambda$ pixel-by-pixel until a distance d (in pixels) has been traversed. The direction depends on the sign of α , while d is determined both by the magnitude of α and the local scale σ . Using W directly for image warping creates image discontinuities for large α values, which are due to regions where many adjacent pixels are warped to the same coordinate. We avoid such artifacts (line 10) by smoothing out the warped coordinates of W , convolving it with a Gaussian weight function of standard deviation σ_I . We compare non-smoothed vs smoothed coordinates in Figure 8d. The final warped image is obtained by a composition of functions (line 11) and is shown in Figure 8e.

Algorithm 2 Warping function

```

1: procedure WARP( $\mathbf{p}$ ,  $\nabla \lambda$ ,  $\sigma$ ,  $\alpha$ )
2:    $d \leftarrow |\alpha| \sigma(\mathbf{p})$ 
3:   while  $d \geq 1$  do
4:      $\mathbf{p} \leftarrow \mathbf{p} + \text{sign}(\alpha) \nabla \lambda$ 
5:      $d \leftarrow \min(d - 1, |\alpha| \sigma(\mathbf{p}))$ 
6:   return  $\mathbf{p} + \text{sign}(\alpha) \nabla \lambda d$ 

```

From a user perspective, our image deformation algorithm offers two controls: the warping amplitude α determines both the magnitude and direction of warping ($\alpha < 0$ leads to sharpening, $\alpha > 0$ to rounding and $\alpha = 0$ has no effect), while the (optional) scale bias γ controls the selected scale as explained in Section 4.1.

4.3 Geometric interpretation

We now give a geometric interpretation of our image deformation algorithm to provide a more intuitive understanding of its effect. We apply two different types of warping on rendered images: one based on the image I , the other based on surface slopes \mathbf{m} . Surface-based warping is easily obtained by replacing I , σ_I , \mathbf{T}_I and λ_I^{\max} in Algorithm 1 by \mathbf{m} , σ_m , \mathbf{T}_m and λ_m^{\max} respectively.

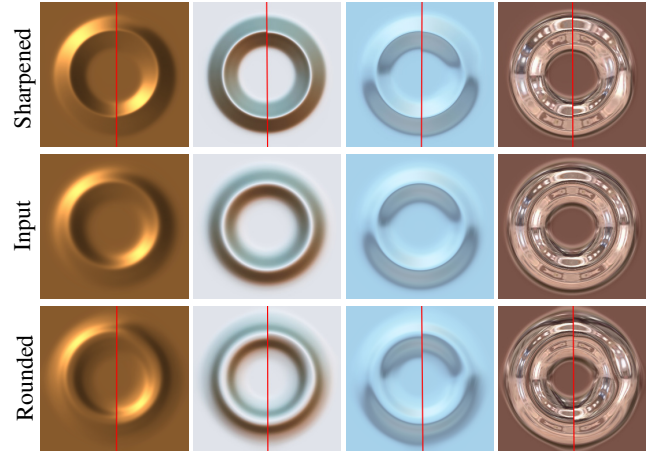


Figure 9: Comparisons between surface-based warping (left halves) and image-based warping (right halves). Middle row: an input height field rendered with 4 different Lit Spheres. Top row: Sharpening of surface shape features ($\alpha = -1$). Bottom row: Rounding of surface shape features ($\alpha = 1$). The two types of warping yield qualitatively similar results overall. Their slight differences become more pronounced with higher-frequency shading.

We first compare image- and surface-based warping in Figure 9 using height fields rendered with Lit Spheres from Figure 4(e,f,h,j). The two types of warping yield qualitatively similar results for various shading functions and warping parameters.

Moreover, since the shading function is applied to the whole image in these examples, warping the image based on surface slopes is strictly equivalent to warping surface slopes and re-rendering. Consequently, we use Poisson reconstruction [Pérez et al. 2003] guided by both initial and warped surface slopes to obtain initial

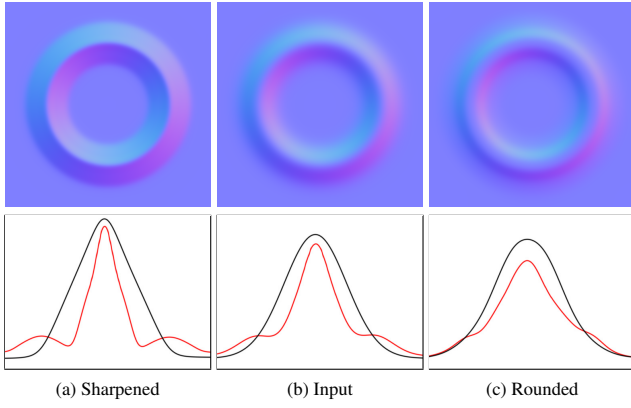


Figure 10: We apply surface-based warping to an input height field (b), either sharpening it (using $\alpha = -1$) in (a) or rounding it (using $\alpha = 1$) in (c). Its effect on surface normals is shown in the top row. A geometric interpretation is given in the bottom row by reconstructing the height field and taking a horizontal image slice. Height values are shown in black and curvature magnitudes in red.

and warped surface depths. As shown in Figure 10, comparing height fields before and after applying warping provides an intuitive geometric interpretation of our algorithm for different signs of α .

4.4 GPU implementation

We have implemented Algorithm 1 in Gratin [Vergne and Barla 2015], using height passes on the GPU. Scale selection requires two passes: one to compute the tensor at multiple scales, and another one to apply Equation 6. In practice, we build a Gaussian pyramid for \mathbf{T}_I using a custom mip map texture, which improves efficiency without sacrificing accuracy. The remainder of Algorithm 1 requires six passes since lines $\{5, 6\}$ and $\{10, 11\}$ can each be implemented in a single pass.

The performance of our warping algorithm is mainly determined by the scale map σ_I : it not only controls the neighborhood size in lines $\{3, 4, 6, 7\}$ of Algorithm 1 but also the traversal distance d in Algorithm 2. However, note that the α parameter is first used at line 9 in Algorithm 1; hence the first steps of the algorithm only have to be recomputed if the γ parameter is modified. Our technique works at interactive rates when modifying γ in most practical cases. This is demonstrated in table 1 where we use a single global scale σ to measure performance independently of the input image. We give timings as a function of the diameter of \mathcal{P} ; hence our smallest neighborhood corresponds to a scale of $1/3$. Note however that when only α is modified, our method runs in real-time (at more than 30fps in all our examples). In practice, for large-scale deformations, we use a low-resolution image as input to lines 2 to 10 in Algorithm 1 to retain interactive performances. The high-resolution image is then used in line 11 to fetch warped coordinates and obtain a high quality result.

$\odot(\mathcal{P})$ in pixels	3	11	21	41	101
800×600	5	60	170	670	4160
1600×1200	10	180	680	2700	10300

Table 1: Performance (in ms) for lines 2 to 7 in Algorithm 1, for various neighborhood sizes ($\odot(\mathcal{P}) = 6\sigma + 1$ in our implementation) at two resolutions, using a Nvidia Quadro M6000.

5 Perceptual evaluation

Since our image deformation technique is intended to manipulate perceived shape in images, its effects must be validated through perceptual experimentation. We have conducted two experiments: one based on the synthetic images used in our Supplemental Material (see also Section 3.3), and another based on photographs. We wrote our experiments in Octave, using the Psychophysics Toolbox extensions [Brainard 1997].

5.1 Experiment 1: synthetic images

Observers 12 subjects (excluding the authors) with normal or corrected-to-normal vision participated in this experiment. All were students in Computer Graphics or Human Vision and were naive with respect to the purpose of the experiment. In particular, no mention of image warping was made.

Procedure Subjects were first shown a simple example of shape sharpening and rounding (see Supplemental material). They were then presented with a pair of images in each trial and asked to pick the one where object shape was sharpest. Each pair contained an original (unmodified) image of a rendered 3D object, and a warped image, their position on the screen (left or right) being randomized on each trial. We rendered a subset of the data used for correlation analysis: 4 objects made of 2 materials in 2 lighting environments. We applied both image- and surface-based warping on all conditions with $\alpha \in \{-1.0, -0.5, 0.5, 1.0\}$ at a fixed scale $\sigma = 4$ for diffuse materials and $\sigma = 5$ for glossy ones (same for both warplings). All combinations were repeated twice and the order of trials randomized, leading to a total of 256 trials. Participants usually completed the experiment in less than 30 minutes.

Results We consider data coming from image-based and surface-based warping separately. A judgment is considered correct when it matches the sign of α . Participants provided correct answers in 86.6% of the trials involving image-based warping (with a std of 26.5%), and in 80.0% of the trials involving surface-based warping (with a std of 32.3%). We plot these percentages as a function of α for all subjects in Figure 12-left.

5.2 Experiment 2: photographs

Observers 12 subjects (excluding authors) participated in experiment 2, of whom 7 had already participated in experiment 1.

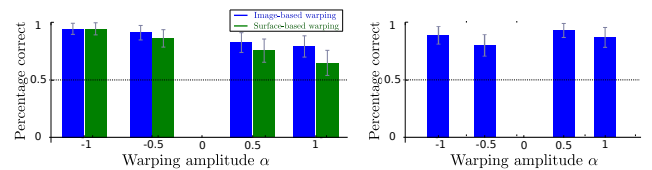


Figure 12: Percentage correct results as a function of the warping amplitude α . Error bars show standard errors of the mean according to the number of participants. In experiment 1 (left), we plot results for image- and surface-based warping separately. In experiment 2 (right), only image-warping results are available. The dashed horizontal line indicates chance level. Subject judgements are well above chance in all conditions in both experiments, which confirms that our warping algorithm produces perceivable shape changes in images.

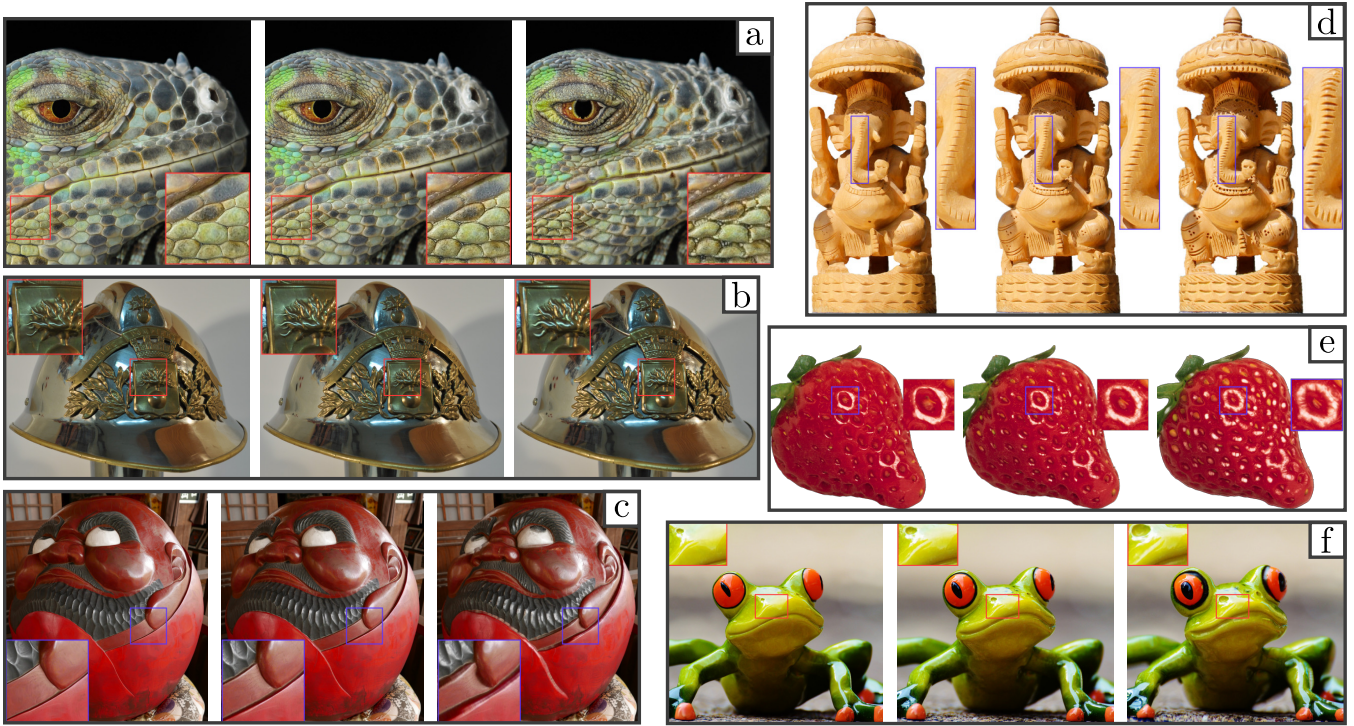


Figure 11: Image-based warping applied to photographs exhibiting different materials and lighting conditions. Each result shows the initial image in the middle, a sharpening result ($\alpha = -1$) on the left and a rounded result ($\alpha = 1$) on the right. Warping is only applied to foreground objects. (a) Macro of a lizard with an intricate scale pattern shot in studio lighting. (b) Metallic helmet photographed indoors. (c) Painted wooden statue shot at the entrance of a budhist temple (partially indoor and outdoor). (d) Wooden statue photograph taken in natural outdoor lighting. (e) Strawberry photographed in studio (collimated) lighting. (f) Macro of a glazed ceramic frog in natural lighting.

Procedure Subjects were first shown a simple example, this time using a warped photograph (see Supplemental material). Then, on each trial, they were asked to identify in which of two images the object shape appeared sharper. The difference with experiment 1 is that input images were photographs; hence only image-based warping was applied in this case. We used 9 photographs showing objects of different shapes and made of different materials. We kept the same parameter values for α and we adapted σ in each image to match their relative feature scales. We repeated all combinations twice, leading to a total of 72 trials. Participants usually completed the experiment in less than 10 minutes.

Results Participants provided correct answers in 85.8% of the trials (with a std of 28%). We plot these percentages as a function of α for all subjects in Figure 12-right.

5.3 Discussion

Experiment 1 shows that subjects could readily perceive the effect of both image-based and surface-based warping on synthetic input. Surprisingly, image-based warping gave slightly higher percentage correct results, which might be partly explained by exaggeration effects (see Section 3.2). Experiment 2 shows that subjects were equally good at perceiving the effect of image-based warping on input photographs. Figure 12 shows that in both experiments, a warping amplitude as small as $\alpha = \pm 0.5$ is enough to elicit changes in shape perception. We also observe a general decrease in performance for $\alpha = 1.0$. This might be explained by the effect of our warping algorithm, which tends to “flatten” perceived shape in some image regions for high positive values of α .

These experiments confirm that our flow-guided warping algorithm is effective at manipulating perceived surface shape in color images. Figure 12 hints that image-based warping may even produce a stronger effect compared to surface-based warping, although the difference is not statistically significant, so a more refined analysis is required. If this observation turns out to be significant, one explanation might be that shape perception is essentially based on directly available image properties, which includes masking, exaggeration and distortion effects not present in surface-based warping results.

6 Results

We now present warping results showing the robustness of our technique to different types of material and lighting, along with detailed parameter control and comparisons with related work. The supplemental video shows results with animated warping parameters, which vividly conveys the effect of our warping technique.

Material and lighting Our method works for a variety of matte to shiny materials in natural as well as artificial lighting. Figure 11a demonstrates warping of small shape features on a lizard in studio lighting. The spaces between the scales that occur at various sizes are either reduced or enlarged. Figure 11b applies warping to a shiny metallic helmet in artificial indoor lighting. Observe in particular how the intricate details of the front part are either sharpened or rounded. Results on aged metallic objects are shown in Figures 1 and 14. Our method works equally well on materials exhibiting a mix of diffuse and specular effects, as shown in Figure 11c. The image shows the wooden statue of a face painted in red, black and

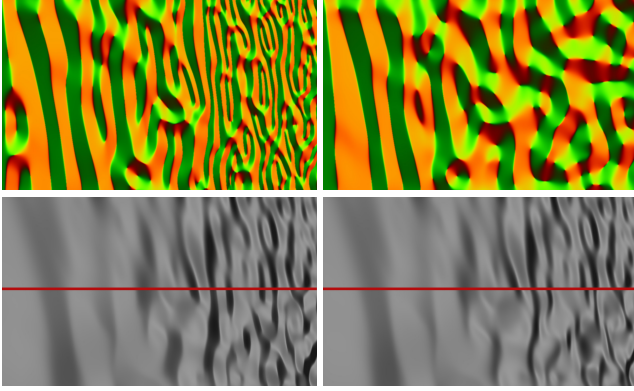


Figure 13: Comparison between local scale selection (left) and a single global scale (right) on a synthetic example. The guiding flow (top row) better adapts to the size variation of features with our local scale. This has a significant impact on the warped image (bottom row) when compared to the input image (shown above the red line): using a single global scale tends to distort shape features.

white, photographed partly indoors and outdoors. Our image warping method makes surface creases either sharper or rounder depending on chosen parameters. It is applied to another wooden statue in Figure 11d, shot in natural outdoor lighting. Observe how the holes in the statue appear either reduced or enlarged, in a way similar to Figure 11a. Our method also works on translucent (natural or man-made) objects. Figure 11e shows a strawberry shot in collimated lighting. Image warping makes it look more or less ‘inflated’. Figure 11f shows a macro photograph of a glazed ceramic frog, exhibiting a mix of specular reflection and subtle translucency. The effect of our method is best seen on the mouth and nostrils, which become appropriately reduced or enlarged. Additional results with varying materials and lighting properties are shown in Figure 18 and in the supplemental document.

Warping control Until now, we have focused on the effect of the warping amplitude α on the final result. Figure 13 shows how our local scale selection process makes it possible to adapt warping to the size of shape features on a synthetic example. In our experience, photographs typically contain features at various scales, and considering a single global scale will inevitably distort some of these features as shown at the bottom right of the figure. Figure 14 further demonstrates the effect of the scale bias γ on image warping: it provides an effective control over the scale of exaggerated shape features. We have found that when pushed toward high scales (i.e., $\gamma > 1$), our warping technique produces compelling caricatures by strongly exaggerating face features, as seen in Figure 15. We show these examples in animation in the supplemental video, along with a brush-based interface that we use to control α and γ locally.

Comparisons Our method follows a different approach compared to related work: instead of modifying pixel intensities, it modifies pixel *locations*. This is shown in Figure 16, where we compare our method to two classic image manipulations: shock filtering [Osher and Rudin 1990] and unsharp masking [Luft et al. 2006]. It is clearly apparent that the effects on the final image are different; in particular, neither the material or lighting appear to change in our results, as opposed to previous work. We believe that the combination of methods that modify pixel intensities *and* locations is an interesting direction of future work, as discussed next.

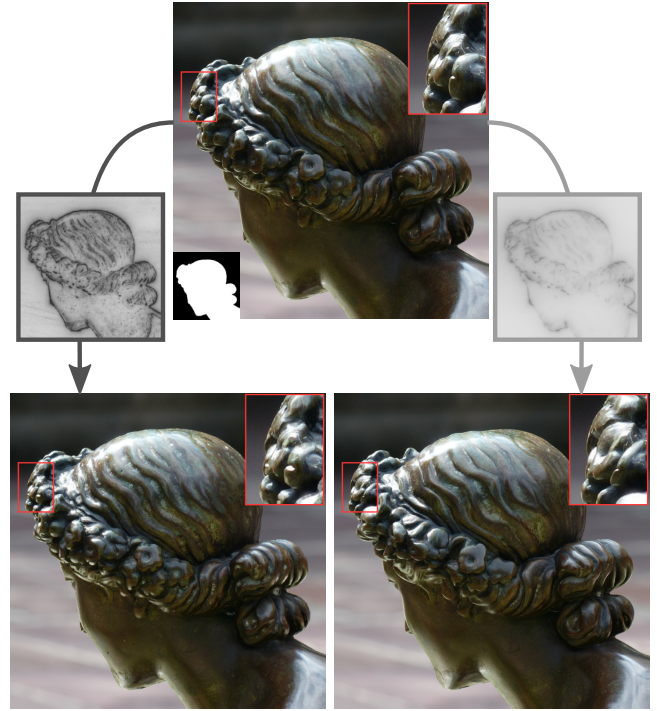


Figure 14: Variations of the scale bias γ on the photograph of a bronze statue of a woman. The top image shows the input image with its mask. Each column shows rounding ($\alpha = 1$) at two different scales ($\gamma = \{0.1, 0.6\}$).

7 Limitations and Future Work

Our warping algorithm presents a few technical limitations. The scale selection mechanism considers a single scale per pixel. However, objects with complex shapes may exhibit multiple scales at a single image location. Applying warping with a bias for large scales might distort small shape details, as can be seen by zooming in Figure 8e for instance. We thus plan to extend our technique to work directly in scale space to manipulate such complex surface structures. Warping may also be restricted by the limited information available in images. For instance, shape features may appear faintly because of a locally smooth shading. In such cases, warping will still work but produce rather subtle effects, as seen in Figure 11e for example. Another limitation occurs with extremely thin highlights, which when rounded lead to flat regions of uniform intensity. Addressing these issues will likely require incorporating additional filtering operators to our approach.

From a more general perspective, our method is restricted by the



Figure 15: With large scale bias values (here $\gamma = 1.0$), our warping method produces striking face caricatures.

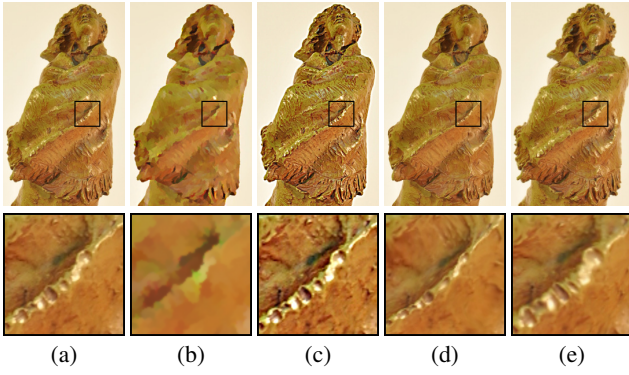


Figure 16: Comparisons of our approach with previous work. (a) Input image; (b) shock filter; (c) unsharp masking; (d) our method (sharpening, $\alpha = -1$); (e) our method (rounding, $\alpha = 1$).

hypothesis made in Equation 1. In particular, images containing many occluding contours (e.g., natural landscapes), optical effects (depth of field) or high-frequency variations due to textures or shadows usually lead to artifacts, as shown in Figure 17. In our current system, we rely on a user-provided mask to apply warping only to a foreground, in-focus and not strongly textured object, hence avoiding these artifacts. A challenging issue will be to deal with the occurrence of multiple *layered* orientation patterns due to textures, shadows or transparency (ideally without requiring user annotations). As shown in Figure 17, our current approach already produces reasonable results with transparency, but tends to create artifacts in the presence of textures. Another potential improvement would be to analyze and manipulate object shape with a more global approach. This is necessary not only to disambiguate concave shape features from convex ones, but also to preserve object symmetries through manipulation. Finally, our method is not temporally coherent as shown in the supplemental video. In future work, we would like to explicitly take into account the additional shape cues available in motion for the exaggeration of shape in videos.

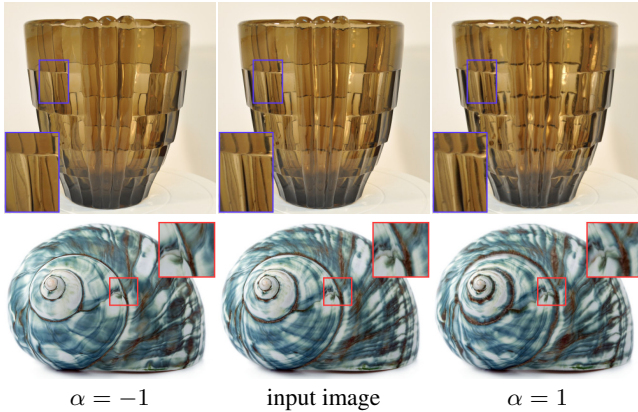


Figure 17: Two example photographs that violate our hypothesis. Top: a transparent object. Bottom: a textured object.

8 Conclusions

We have presented a real-time image warping method that gives compelling illusion of shape changes, as demonstrated by our perceptual evaluation. It relies on correlations between surface shape features and image orientations, which provide additional evidence

for the central role played by orientation patterns in the perception of shape. Similarly to recent work [Wu et al. 2012; Wadhwa et al. 2013; Dekel et al. 2015; Boyadzhiev et al. 2015], our method specifically targets one scene property (surface shape) in one family of input images (objects with homogeneous materials). An exciting direction of research will be to combine and improve on these separate approaches to deal with arbitrary images and scene properties.

Acknowledgements

We would like to thank Pierre B  nard and the anonymous reviewers for their insightful comments on the paper. We are also grateful to Jean-Pierre and Marion Vergne for providing the photographs used in Figures 15, 16 and 17 (top). We thank Paul Debevec for his environment maps used in Figure 4. The photograph used in Figure 1 is courtesy of   Expertissim. Images of Figures 11 (a,e,f), 14, 17 (bottom) and 18 come from *pixabay.com* (released under Creative Commons CC0). This work was partly supported by the EU Marie Curie Initial Training Network “PRISM” (FP7-PEOPLE-2012-ITN, Grant Agreement: 316746) and by the DFG SFB-TRR-135 “Cardinal Mechanisms of Perception”.

References

- BARROW, H. 1978. Recovering intrinsic scene characteristics from images. *Computer Vision Systems*, 3–26.
- BEN-SHAHAR, O., AND ZUCKER, S. 2001. On the perceptual organization of texture and shading flows: from a geometrical model to coherence computation. In *CVPR 2001.*, vol. 1, 1–1048–I–1055 vol.1.
- BIGUN, J., AND GRANLUND, G. H. 1986. Optimal orientation detection of linear symmetry. Tech. Rep. Report LiTH-ISY-I-0828, Computer Vision Laboratory, Link  ping University, Sweden.
- BOUSSEAU, A., PARIS, S., AND DURAND, F. 2009. User assisted intrinsic images. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2009)* 28, 5.
- BOYADZHIEV, I., BALA, K., PARIS, S., AND ADELSON, E. 2015. Band-sifting decomposition for image-based material editing. *ACM Trans. Graph.* 34, 5 (Nov.), 163:1–163:16.
- BRAINARD, D. H. 1997. The psychophysics toolbox. *Spatial Vision* 10, 4, 433–436.
- BROX, T., VAN DEN BOOMGAARD, R., LAUZE, F., VAN DE WEIJER, J., WEICKERT, J., AND KORNPROBST, P. 2006. Adaptive structure tensors and their applications. In *Visualization and image processing of tensor fields*, J. Weickert and H. Hagen, Eds., vol. 1. Springer, Jan., ch. 2, 17–47.
- CANIARD, F., AND FLEMING, R. W. 2007. Distortion in 3d shape estimation with changes in illumination. In *Proceedings of the 4th Symposium on Applied Perception in Graphics and Visualization*, ACM, New York, NY, USA, APGV ’07, 99–105.
- CARROLL, R., RAMAMOORTHY, R., AND AGRAWALA, M. 2011. Illumination decomposition for material recoloring with consistent interreflections. *ACM Trans. Graph.* 30, 4 (July), 43:1–43:10.
- CHEN, T., ZHU, Z., SHAMIR, A., HU, S.-M., AND COHEN-OR, D. 2013. 3sweep: Extracting editable objects from a single photo. *ACM Trans. Graph.* 32, 6 (Nov.), 195:1–195:10.
- DEKEL, T., MICHAELI, T., IRANI, M., AND FREEMAN, W. T. 2015. Revealing and modifying non-local variations in a sin-

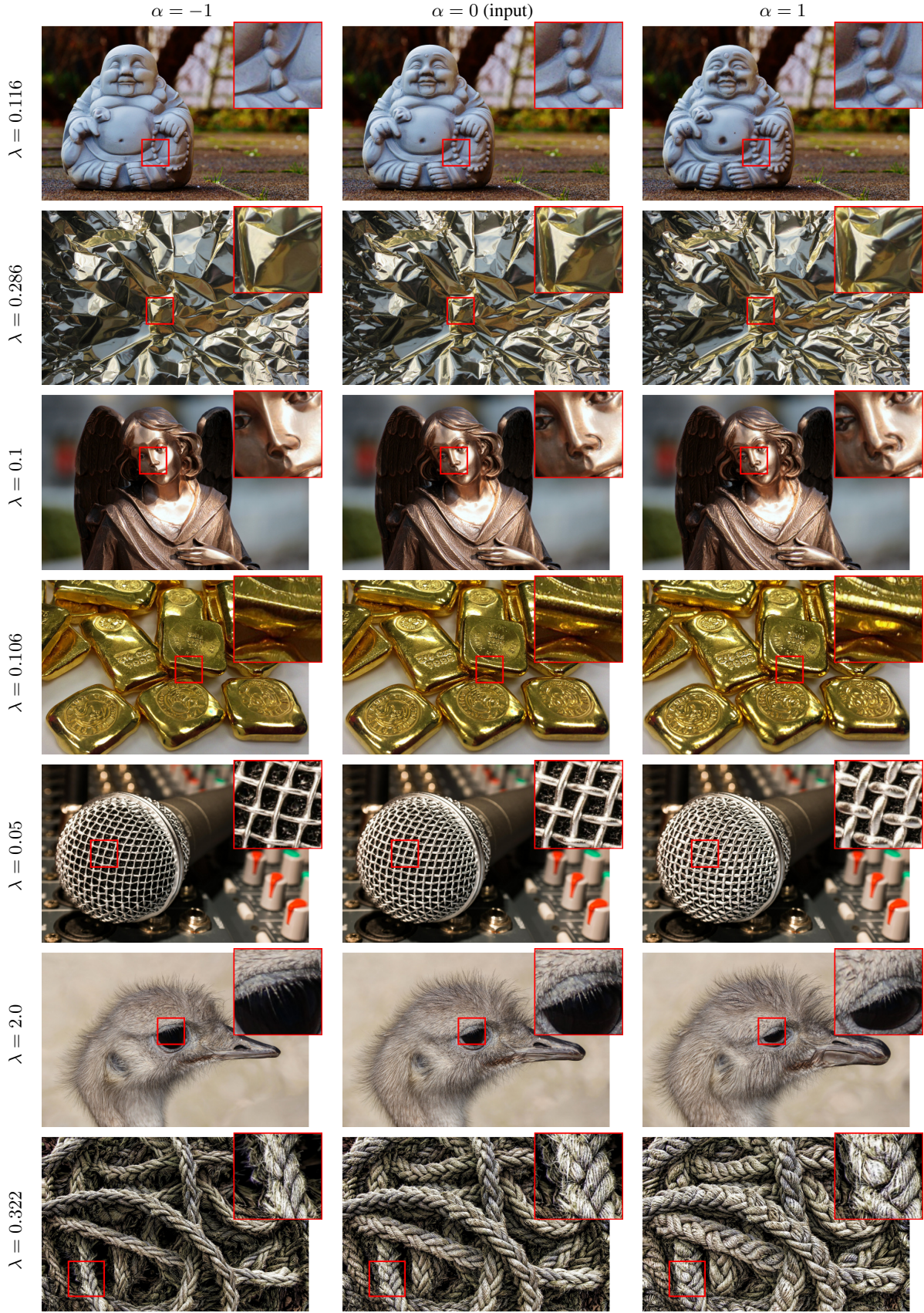


Figure 18: More results with varying materials, and scale choices (see supplemental document).

- gle image. *ACM Transactions on Graphics (Proc. SIGGRAPH Asia)*.
- DONG, Y., TONG, X., PELLACINI, F., AND GUO, B. 2011. App-gen: Interactive material modeling from a single image. *ACM Trans. Graph.* 30, 6 (Dec.), 146:1–146:10.
- FARBMAN, Z., FATTAL, R., LISCHINSKI, D., AND SZELISKI, R. 2008. Edge-preserving decompositions for multi-scale tone and detail manipulation. In *ACM SIGGRAPH 2008 Papers*, ACM, New York, NY, USA, SIGGRAPH '08, 67:1–67:10.
- FATTAL, R., AGRAWALA, M., AND RUSINKIEWICZ, S. 2007. Multiscale shape and detail enhancement from multi-light image collections. *ACM Trans. Graph.* 26, 3 (July).
- FLEMING, R. W., TORRALBA, A., AND ADELSON, E. H. 2004. Specular reflections and the perception of shape. *J. Vis.* 4, 9 (9), 798–820.
- FLEMING, R. W., HOLTMANN-RICE, D., AND BÜLTHOFF, H. H. 2011. Estimation of 3d shape from image orientations. *Proceedings of the National Academy of Sciences* 108, 51, 20438–20443.
- FREEMAN, W. T. 1994. The generic viewpoint assumption in a framework for visual perception. *Nature* 368, 6471 (Apr.), 542–545.
- GUTIERREZ, D., SERON, F. J., LOPEZ-MORENO, J., SANCHEZ, M. P., FANDOS, J., AND REINHARD, E. 2008. Depicting procedural caustics in single images. *ACM Trans. Graph.* 27, 5 (Dec.), 120:1–120:9.
- HORN, B. K. P., AND BROOKS, M. J., Eds. 1989. *Shape from Shading*. MIT Press, Cambridge, MA, USA.
- JAKOB, W., 2010. Mitsuba renderer. <http://www.mitsuba-renderer.org>.
- KAJIYA, J. T. 1986. The rendering equation. *SIGGRAPH Comput. Graph.* 20, 4 (Aug.), 143–150.
- KHAN, E. A., REINHARD, E., FLEMING, R. W., AND BÜLTHOFF, H. H. 2006. Image-based material editing. *ACM Trans. Graph.* 25, 3 (July), 654–663.
- KHOLGADE, N., SIMON, T., EFROS, A., AND SHEIKH, Y. 2014. 3d object manipulation in a single photograph using stock 3d models. *ACM Trans. Graph.* 33, 4 (July), 127:1–127:12.
- KOENDERINK, J., AND VAN DOORN, A. 1980. Photometric invariants related to solid shape. *Optica Acta: International Journal of Optics* 27, 7, 981–996.
- KYPRIANIDIS, J. E., AND DÖLLNER, J. 2008. Image abstraction by structure adaptive filtering. In *Proc. EG UK Theory and Practice of Computer Graphics*, 51–58.
- LINDBERG, T. 1998. Feature detection with automatic scale selection. *Int. J. Comput. Vision* 30, 2 (Nov.), 79–116.
- LUFT, T., COLDITZ, C., AND DEUSSEN, O. 2006. Image enhancement by unsharp masking the depth buffer. *ACM Trans. Graph.* 25, 3 (July), 1206–1213.
- MOONEY, S. W., AND ANDERSON, B. L. 2014. Specular image structure modulates the perception of three-dimensional shape. *Current Biology* 24, 22, 2737 – 2742.
- OREN, M., AND NAYAR, S. K. 1997. A theory of specular surface geometry. *Int. J. Comput. Vision* 24, 2 (Sept.), 105–124.
- OSHER, S., AND RUDIN, L. I. 1990. Feature-oriented image enhancement using shock filters. *SIAM J. Numer. Anal.* 27, 4 (Aug.), 919–940.
- PARIS, S., HASINOFF, S. W., AND KAUTZ, J. 2011. Local laplacian filters: Edge-aware image processing with a laplacian pyramid. *ACM Trans. Graph.* 30, 4 (July), 68:1–68:12.
- PÉREZ, P., GANGNET, M., AND BLAKE, A. 2003. Poisson image editing. *ACM Trans. Graph.* 22, 3 (July), 313–318.
- PERONA, P., AND MALIK, J. 1990. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* 12, 7 (July), 629–639.
- SAVARESE, S., FEI-FEI, L., AND PERONA, P. 2004. What do reflections tell us about the shape of a mirror? In *Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization*, ACM, New York, NY, USA, APGV '04, 115–118.
- SLOAN, P.-P. J., MARTIN, W., GOOCH, A., AND GOOCH, B. 2001. The lit sphere: A model for capturing npr shading from art. In *Proceedings of Graphics Interface 2001*, Canadian Information Processing Society, Toronto, Ont., Canada, Canada, GI '01, 143–150.
- STALLING, D., AND HEGE, H.-C. 1995. Fast and resolution independent line integral convolution. In *Proceedings of the 22Nd Annual Conference on Computer Graphics and Interactive Techniques*, ACM, New York, NY, USA, SIGGRAPH '95, 249–256.
- VERGNE, R., AND BARLA, P. 2015. Designing Gratin, A GPU-Tailored Node-Based System. *Journal of Computer Graphics Techniques* 4, 4, 17.
- VERGNE, R., BARLA, P., FLEMING, R. W., AND GRANIER, X. 2012. Surface flows for image-based shading design. *ACM Trans. Graph.* 31, 4 (July), 94:1–94:9.
- WADHWA, N., RUBINSTEIN, M., DURAND, F., AND FREEMAN, W. T. 2013. Phase-based video motion processing. *ACM Trans. Graph.* 32, 4 (July), 80:1–80:10.
- WADHWA, N., DEKEL, T., WEI, D., DURAND, F., AND FREEMAN, W. T. 2015. Deviation magnification: Revealing departures from ideal geometries. *ACM Trans. Graph.* 34, 6 (Oct.), 226:1–226:10.
- WIJNTJES, M. W. A., DOERSCHNER, K., KUCUKOGLU, G., AND PONT, S. C. 2012. Relative flattening between velvet and matte 3d shapes: Evidence for similar shape-from-shading computations. *Journal of Vision* 12, 1, 2.
- WU, H.-Y., RUBINSTEIN, M., SHIH, E., GUTTAG, J., DURAND, F., AND FREEMAN, W. 2012. Eulerian video magnification for revealing subtle changes in the world. *ACM Trans. Graph.* 31, 4 (July), 65:1–65:8.
- YEUNG, S. K., TANG, C.-K., BROWN, M. S., AND KANG, S. B. 2011. Matting and compositing of transparent and refractive objects. *ACM Transactions on Graphics* 30, 1, 2.