

Understanding the Role of Phase Function in Translucent Appearance

IOANNIS GKIOULEKAS

Harvard School of Engineering and Applied Sciences
BEI XIAO

Massachusetts Institute of Technology

SHUANG ZHAO

Cornell University

EDWARD H. ADELSON

Massachusetts Institute of Technology

TODD ZICKLER

Harvard School of Engineering and Applied Sciences
and

KAVITA BALA

Cornell University

Multiple scattering contributes critically to the characteristic translucent appearance of food, liquids, skin, and crystals; but little is known about how it is perceived by human observers. This paper explores the perception of translucency by studying the image effects of variations in one factor of multiple scattering: the phase function. We consider an expanded space of phase functions created by linear combinations of Henyey-Greenstein and

von Mises-Fisher lobes, and we study this physical parameter space using computational data analysis and psychophysics.

Our study identifies a two-dimensional embedding of the physical scattering parameters in a perceptually-meaningful appearance space. Through our analysis of this space, we find uniform parameterizations of its two axes by analytical expressions of moments of the phase function, and provide an intuitive characterization of the visual effects that can be achieved at different parts of it. We show that our expansion of the space of phase functions enlarges the range of achievable translucent appearance compared to traditional single-parameter phase function models. Our findings highlight the important role phase function can have in controlling translucent appearance, and provide tools for manipulating its effect in material design applications.

Acknowledgements: I.G. and T.Z. were funded by the National Science Foundation through awards IIS-1161564, IIS-1012454, and IIS-1212928, and by Amazon Web Services in Education grant awards. B.X. and E.A. were funded by the National Institutes of Health through awards R01-EY019262-02 and R21-EY019741-02. S.Z. and K.B. were funded by the National Science Foundation through awards IIS-1161645, and IIS-1011919. We thank Bohnams for allowing us to use the photograph of Figure 0??.

Authors' addresses: I. Gkioulekas (corresponding author), T. Zickler, Harvard School of Engineering and Applied Sciences, Cambridge, MA 02138; email: {igkiou,zickler}@seas.harvard.edu. B. Xiao, Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139; email: beixiao@mit.edu. E.H. Adelson, CSAIL, Massachusetts Institute of Technology, Cambridge, MA 02139; email: adelson@csail.mit.edu. S. Zhao, K. Bala, Computer Science Department, Cornell University, Ithaca, NY 14853, {szhao,kb}@cs.cornell.edu.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or permissions@acm.org.

© YYYY ACM 0730-0301/YYYY/13-ARTXXX \$10.00

DOI 10.1145/XXXXXXXX.YYYYYYY

<http://doi.acm.org/10.1145/XXXXXXXX.YYYYYYY>

Categories and Subject Descriptors: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism; I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—*Physically based modeling*

General Terms: Algorithms, Design, Experimentation, Human Factors, Measurement

Additional Key Words and Phrases: Translucency, subsurface scattering, phase function, human perception, rendering

ACM Reference Format:

Gkioulekas, I., Xiao, B., Zhao, S., Adelson, E. H., Zickler, T., and Bala, K. YYYY. Understanding the Role of Phase Function in Translucent Appearance. *ACM Trans. Graph.* VV, N, Article XXX(MonthYYYY), 19pages.

DOI = 10.1145/XXXXXXXX.YYYYYYY

<http://doi.acm.org/10.1145/XXXXXXXX.YYYYYYY>

1. INTRODUCTION

Light penetrates into translucent materials and scatters multiple times before re-emerging towards an observer. This multiple scattering is critical in determining the appearance of food, liquids, skin, crystals, and other translucent materials. Human observers care greatly about the appearance of these materials, and they are

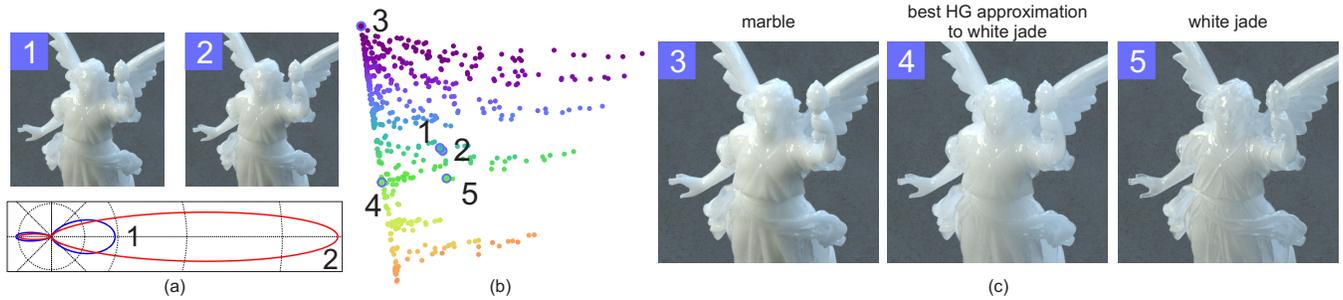


Fig. 1. The role of phase functions. (a) Completely different phase functions can result in very similar appearances, as shown by the two images (top), rendered with the phase functions shown in the polar plot (below). (b) Our analysis identifies a two-dimensional space for phase functions, where neighborhoods correspond to similar translucent appearance. (Numbers indicate the location in the two-dimensional space of the phase functions used to render the corresponding images in (a) and (c).) (c) We derive perceptually uniform axes that we can use to explore this two-dimensional space of translucent appearance. This enables us to design two-lobe phase functions to match the appearance of materials such as marble (left), but also others, such as white jade (right), that single lobe phase functions cannot match well (middle).

skilled at discriminating subtle differences in their scattering properties, as when they distinguish milk from cream, or marble from Formica [?]. The focus in graphics has been to seek realism by accurately simulating the physics of transport in translucent media [?], approximating physics for efficient rendering [?; ?], and acquiring material measurements [?; ?]. However, little is known about how human observers perceive translucency. This lack of understanding severely inhibits our ability to acquire, intuitively design and edit translucent materials.

Simulations of light transport in translucent media are based on radiative transfer [?]. A medium containing particles that scatter and absorb light is represented by its macroscopic bulk properties, typically consisting of two spectral functions (scattering and absorption coefficients) and a spectral-angular “phase function” describing the effective spherical scattering distribution at a point. This approach achieves substantial realism, but the physical parameter space is extremely large and perceptually non-uniform. Without an understanding of the relationship between physical parameters and appearance, choosing parameters to match the appearance of a specific material becomes a tedious (and computationally expensive) process of trial and error.

This paper takes first steps toward understanding the relationship between the perception of translucency and the underlying physical scattering parameters. We focus our attention on the phase function in particular, and we study how changing the shape of this spherical scattering distribution affects appearance. This choice may appear odd at first, as this shape is known to be relatively unimportant for the appearance of thick geometries that are dominated by high-order scattering (allowing, for example, the effective use of similarity theory [?] and the diffusion approximation [?]). However, the phase function can impact appearance in a perceptually important way near thin geometric structures, where light undergoes only a handful of scattering events before being emitted towards the observer. Figure 0??(c) provides an example where, with all other material parameters being equal, certain phase functions induce a soft marble-like appearance (Image 4) while others create a “glassy” effect (Image 5) that has much sharper contrast near surface details, an appearance characteristic of materials such as white jade. (A fine example of white jade is shown in Figure 0??). Therefore, our study draws attention to this, often overlooked, crucial effect that the phase function can have on translucent appearance.

The space of phase functions is the space of probability distributions on the sphere of directions, an infinite-dimensional function space. Typically, we regularize this space by only considering a family of analytic expressions controlled by a small number of parameters. The single-parameter Henyey-Greenstein (HG) distribution is used extensively because of its relative flexibility and sampling efficiency. This choice can, however, be restrictive, and it is known to be a poor approximation for some materials [?; ?], including common wax as shown in Figure 0??. Therefore, we expand the space of analytic phase function models, by including von Mises-Fisher (vMF) distributions in addition to HG, and their linear combinations. In our results, we demonstrate the necessity of such more general phase function models, by showing that they considerably broaden the range of achievable scattering appearance.

Very different phase functions can produce the same visual appearance, even when the other material parameters remain fixed (see Figure 0??(a)). This suggests that the perceptual space of translucency is much smaller than this expanded physical parameter space of phase functions. Our main goals were to obtain an intuitive understanding of this lower-dimensional perceptual space, and derive computational tools for navigating through it. The most direct way to achieve our goals would be to densely sample the space of phase functions, and then design a psychophysical experiment to obtain perceptual distances between images rendered with all pairs of these phase functions. This has provided useful insights into the perception of certain aspects of opaque reflectance [?; ?]. In the present case this approach is not tractable, however: the number of dimensions of the physical parameter space implies that even a coarse sampling would require more than 60 million human judgments. Our solution was to combine psychophysics with computational analysis, and we did so in three steps. First, we densely sampled the physical parameter space of phase functions to compute thousands of images, and employed computational analysis with various image distance measures to find low-dimensional embeddings of this space. These embeddings turned out to be two-dimensional (see Figure 0??(b)), and very consistent across different distance metrics, shapes, and lighting.

Second, we ran a psychophysical study to determine if this computationally-determined embedding is consistent with perception. We sampled from our large image set to obtain a representative subset that was small enough to collect psychophysically exhaustive pairwise perceptual distances. We found that these distances



Fig. 2. Detail from photograph of the Qianlong Emperor’s personal dragon seal, made from exquisite white jade and dated 1793. (Image courtesy of [?].)

are consistent with a two-dimensional embedding that is similar to that found computationally, thereby affirming the perceptual relevance of that embedding.

Third, we performed statistical analysis of the computationally-obtained and perceptually-consistent two-dimensional embedding, to investigate how its dimensions relate to phase function shape. We identified two linearly independent axes that are described as simple analytic expressions of generalized first and second moments of the phase function. We also obtained a learned distance matrix that can be used to compute an efficient approximation to the perceptual distance between any two tabulated phase functions.

Our results can facilitate manipulating the phase function to achieve a target appearance in material design applications. We provide an intuitive characterization of the visual effects achieved at different parts of the derived appearance space. We show what parametric families inside our expanded set of phase functions can be used to reach each part of this space, verifying that models more general than single-lobe HG are necessary in many cases. We validate that the two identified analytic axes are perceptually uniform “knobs” that can be used to select exact physical parameter values.

j

2. RELATED WORK

Perception of opaque materials. Most previous research in the perception of material properties has focused on opaque surfaces, and color and lightness in particular (e.g., [?; ?]). There have also been numerous studies of the perception of gloss, with results suggesting that real-world lighting is important for perception [?]; while simple image statistics correlate with perceived gloss in some situations [?; ?; ?], other situations seem to require explicit reasoning about the three-dimensional scene geometry [?; ?].

In graphics, the perception of materials has focused primarily on gloss perception. Pellacini et al. [?] and Westlund and Meyer [?] developed psychophysical models of gloss perception. Vangorp et al. [?; ?] studied the effect of object shape on gloss perception. Rushmeier et al. [?], Aydin et al. [?], and Ramanarayanan et al. [?] developed image-quality measures to characterize shape, material and lighting perception in CG images and animations [?]; these metrics have been used to assess the fidelity of rendering algorithms for material appearance [?]. Khan et al. [?] leveraged perception to

develop tools for intuitive material editing. Kerr and Pellacini [?] studied the usability of various material design paradigms.

Pellacini et al. [?] used human similarity judgments to derive two perceptually-uniform knobs for the space of Ward bi-direction reflectance distribution functions (BRDF). Wills et al. [?] introduced a novel non-metric multidimensional scaling (MDS) algorithm of human similarity judgments to find similar knobs, but for navigating a much larger space of tabulated isotropic BRDFs. An alternative approach is proposed by Ngan et al. [?], who built a BRDF navigation tool by, instead of using human observers, approximating the perceptual distance between two BRDFs using pixel-wise differences between images they produce for the same scene geometry and lighting. The successes of these important studies provide motivation for both the computational and psychophysical methods that we use in our study of translucency.

Perception of transparent and translucent materials. Traditional research on perception of non-opaque materials has focused on scenes with overlapping thin transparent sheets (e.g., [?; ?]). There has been relatively little research on the perception of translucent three-dimensional objects [?]. A first study was presented by Fleming et al. [?; ?], using synthetic images of three-dimensional objects to identify image cues, such as specular highlights, edge blurring, and luminance histograms, that seem to be correlated with the perception of translucency. The importance of highlights and shading on surfaces of three-dimensional objects has also been studied by Koenderink and van Doorn [?] and Motoyoshi [?]. Cunningham et al. [?] parameterized translucency scattering and absorption, along with gloss, with a focus on aesthetic judgments. None of these approaches attempts to parameterize the space of translucent materials numerically in a way that could be used for rendering, design and acquisition applications in graphics.

3. MOTIVATION

We restrict our attention to homogeneous media whose translucent appearance can be determined as a solution of the radiative transfer equation (RTE) [?]:

$$(\boldsymbol{\omega} \cdot \nabla)L(\boldsymbol{x}, \boldsymbol{\omega}) = Q(\boldsymbol{x}, \boldsymbol{\omega}) - \sigma_t L(\boldsymbol{x}, \boldsymbol{\omega}) + \sigma_s \int_{4\pi} p(\boldsymbol{\omega}' \cdot \boldsymbol{\omega}) L(\boldsymbol{x}, \boldsymbol{\omega}') d\boldsymbol{\omega}', \quad (1)$$

where $\boldsymbol{x} \in \mathbb{R}^3$ is a point in the medium; $\boldsymbol{\omega}, \boldsymbol{\omega}' \in \mathbb{S}^2$ are directions; $L(\boldsymbol{x}, \boldsymbol{\omega})$ is the desired light field; and $Q(\boldsymbol{x})$ accounts for emission from light sources. The material properties are described by the *scattering coefficient* σ_s , the *absorption coefficient* σ_a (implicitly via the relationship for the *extinction coefficient* $\sigma_t = \sigma_s + \sigma_a$), and the *phase function* p . For light at a single wavelength, σ_s and σ_a are the amount of light that is scattered and absorbed, and the phase function is a probability distribution function on the sphere of directions that only depends on the angle $\theta = \arccos(\boldsymbol{\omega}' \cdot \boldsymbol{\omega})$ relative to the incident direction. Complementary to the above quantities, the following two parameters are also used for describing scattering behavior: the *mean free path*, equal to $1/\sigma_t$, and the *albedo*, equal to σ_s/σ_t .

3.1 Phase Functions

The phase function describes the angular scattering of light at a small element of volume. Since solving the RTE usually involves drawing samples from this function, convenient analytic forms are desirable, and the form that is most often used is due to Henyey and

Greenstein [?]:

$$p_{\text{HG}}(\theta) = \frac{1}{4\pi} \frac{1 - g^2}{(1 + g^2 - 2g \cos \theta)^{\frac{3}{2}}}, \quad (2)$$

$\theta \in [0, \pi]$. This is proposed as an approximation to the Mie solution [?] and is used widely because of its attractive combination of efficiency and relative flexibility. It is controlled by a single parameter $g \in [-1, 1]$, and can represent both predominantly forward and backward scattering ($g > 0$ and $g < 0$, respectively). The shapes of the HG phase function for three different positive values of g are depicted in Figure 0??(d).

The single g parameter, however, limits the flexibility of the HG model and sacrifices some physical plausibility, including not reducing to the Rayleigh phase function in the small-particle regime [?]. However, it is a common approximation in graphics; renderers, using similarity theory [?], often use $g = 0$ with appropriately scaled coefficients for simulating deep scattering.

To model more complex types of scattering, linear combinations of two HG lobes—one forward and one backward—can be used [?; ?; ?], but there are still materials for which this is a limited approximation. Consider, for example, the photographs of soap and wax in Figure 0??(b)-(c), which we acquired from homogeneous material samples of thickness 4 cm with a viewing direction normal to the surface and laser illumination from an angle of 45° (halfway between the view direction and the surface tangent plane). These photographs have been coarsely quantized to a small number of gray levels (i.e., *posterized*) to emphasize the shapes of the emergent iso-brightness contours. They clearly show that these two materials exhibit very distinct scattering behaviors. We distinguish these patterns as being “teardrop” versus “apple”. The second column of the figure shows the best-fit simulation results that can be obtained when rendering using either a single HG or linear combinations of two HG phase functions, and these are obviously qualitatively very different. In particular, we have found that the HG phase function—or any two-lobe mixture of HG phase functions—is unable to produce the apple-shaped iso-brightness contours that can be observed in the wax photographs. In Figure 0??(a), we also show photographs under natural illumination of the same “frog prince” model made from these two materials, soap and wax; we can observe that the translucent appearance of the two objects is very different.

3.2 Henyey-Greenstein and von Mises-Fisher Lobes

One possible explanation for apple-shaped scattering is a phase function that is forward scattering, but with significant probability mass away from the forward direction, so that enough photons entering the material “turn around” after multiple bounces. HG lobes do not have this property, as different values of parameter g only move between isotropic and strongly-forward lobes.

To augment the usual space of phase functions, we propose adding lobes that are shaped according to the von Mises-Fisher distribution on the sphere of directions [?]. This distribution has many of the properties of the normal distribution for real-valued random variables, and is often used as its analogue for spherical statistics, for example to represent BRDFs [?], or lighting environments [?]. Its probability density function is given as:

$$p_{\text{vMF}}(\theta) = \frac{\kappa}{2\pi \sinh \kappa} \exp(\kappa \cos \theta), \quad (3)$$

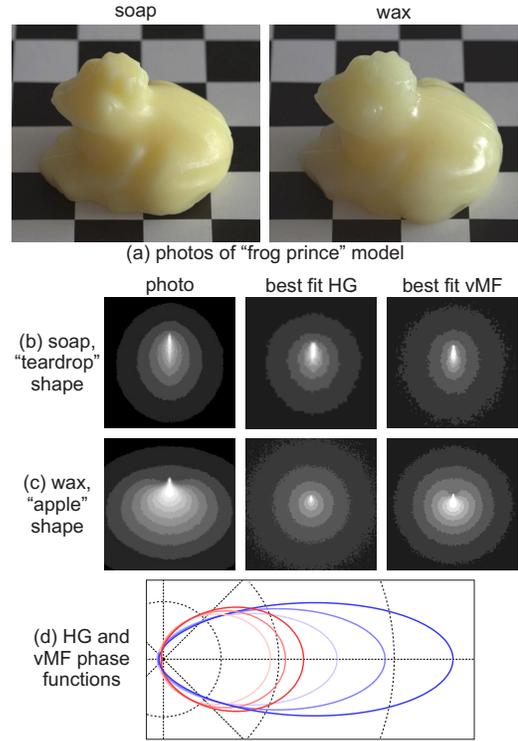


Fig. 3. Phase functions for different materials. (a) Photographs under natural illumination of a “frog prince” model made of two different translucent materials, soap and wax. The middle two rows (b) and (c) show images, and rendered fits, of thick, homogeneous slabs made from these materials, under laser illumination. (b): soap, (c): wax. From left to right: photographs, best fit simulations using single HG lobes or their mixtures, best-fit simulations using single vMF lobes or their mixtures. HG lobes cannot reproduce the “apple” shaped scattering patterns observed in wax and other materials, whereas vMF can. (d) compares HG (blue) and vMF (red) phase functions for increasing values of average cosine, shown as increase in saturation. vMF lobes are wider and not as strongly forward-scattering as HG lobes.

$\theta \in [0, \pi]$, and has a single shape parameter κ :

$$\kappa = 2\bar{C} / (1 - M_C), \quad \text{where}, \quad (4)$$

$$\bar{C} = 2\pi \int_{\theta=0}^{\pi} \cos \theta p(\theta) \sin \theta d\theta, \quad (5)$$

$$M_C = 2\pi \int_{\theta=0}^{\pi} (\cos \theta)^2 p(\theta) \sin \theta d\theta. \quad (6)$$

So, κ is a measure of concentration, being directly related to the average cosine \bar{C} and inversely related to the second moment of the cosine of directions on the sphere. Note that, like the HG distribution, the vMF distribution is suitable for rendering as it can be sampled efficiently (see Appendix 0??).

Figure 0??(d) compares HG and vMF lobes that have the same average cosine \bar{C} —equal in value to the shape parameter g for the HG distribution—, where we observe that the vMF distribution has more mass in off-frontal directions. We found that by using a phase function that is either a single forward vMF lobe or a linear combination where the forward lobe is a vMF lobe, one can successfully reproduce the apple-shaped scattering patterns demonstrated by the wax sample (Figure 0??(c)).

3.3 Towards a Perceptual Space

In this paper we consider the augmented space of analytic phase functions consisting of linear mixtures of two lobes. An instance of a phase function is created by: deciding on the type of each of the forward and backward lobes (vMF or HG); selecting values for their respective shape parameters (κ or g); and selecting a value for the mixing weight between the two lobes. Our space of phase functions is thus a union of three-dimensional subspaces embedded in a five-dimensional ambient space.

The procedure just described involves selecting five parameters. This is inconvenient for design tasks such as adjusting scattering parameters to match a photograph to a desired appearance. The same is true even if one only allows linear mixtures of two HG lobes, which requires adjusting three numerical values. The problem is that there is no intuitive relationship between the physical parameters and the appearance they induce, so one is essentially left to choose parameters blindly through exhaustive search.

While the physical parameter space is daunting, the perceptual space seems more manageable. As shown in Figure 0??(a), different physical parameters can produce the same image, suggesting that the perceptual space of phase functions is smaller. Motivated by this belief, we seek to parameterize this smaller perceptual space, which we do in three steps. In the first step (Section 0??), we design a large set of scenes with physical parameters that characterize the space of translucent appearance, under material, object shape, and lighting variations. The second step (Section 0??) is to perform a large-scale computational analysis to examine the geometry of this space, by rendering thousand of images and using image-driven metrics to compare them. The third step (Section 0??) is to use the findings of the computational analysis to design and perform a psychophysical study involving a manageable subset of scenes and phase functions. We process the results of these experiments in Sections 0?? and 0??, to discover a two-dimensional space of translucent appearance; we find perceptually uniform parameterizations of the two dimensions by analytic expressions of the first and second moments of the phase function, and show that they correspond to intuitive changes in the appearance of a translucent object. Finally, we conclude in Section 0?? with discussion of future directions.

4. CHARACTERIZING TRANSLUCENT APPEARANCE

We begin our study by exploring the space of translucent appearance to determine the image conditions to use for our experiments. This involves examining rendered images corresponding to a large number of combinations of object shape, lighting, and material (i.e., physical parameter values).

To guide our design choices, we ran many pilot studies, each involving 3 to 6 human subjects running a paired-comparison experiment for stimulus sets of 16 images and participating in informal exit interviews. The experimental and data analysis methodologies of these studies are identical to those described in Section 0?? for our large-scale psychophysical experiments. We now describe our design choices based on our findings from these pilot studies.

4.1 Sampling the Physical Parameter Space

We require a sampling of the useful region of the physical parameter space that is dense enough to be visually representative. We formed such a set by considering phase functions with a single HG or vMF lobe, and those with linear mixtures of a forward and a backward lobe, each of which can be one of either

HG or vMF. For the HG lobes, we considered the range of g values often used in the rendering literature, and selected g from $\{0, \pm 0.3, \pm 0.5, \pm 0.8, \pm 0.9, \pm 0.95\}$; for the vMF lobes we found through experimentation that $\kappa > 100$ produced unrealistic results, and selected $\kappa \in \{\pm 1, \pm 5, \pm 10, \pm 25, \pm 75, \pm 100\}$. We also chose values of the mixing weight of the first lobe from $\{0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.99\}$, with the corresponding weight on the second lobe being one minus this value. From the set of phase functions produced by all these combinations, we remove those with negative average directions (i.e., materials that are more backward than forward scattering).

In pilot studies, using images rendered with different subsets from the above set of phase functions, subjects tended to cluster materials into unrealistic glass-like, and more diffuse translucent ones, and this variation dominated their responses. Through experimentation, we found that such appearance corresponded to large average cosine values (\bar{C}). By examining the image sets ourselves, we decided to remove phase functions with $\bar{C} > 0.8$, to arrive at a final set of 753 phase functions. We show an example pruned image in the supplementary material.

As we restrict our attention to phase function effects, we kept the scattering and absorption coefficients, σ_s and σ_a , fixed for the majority of our simulations, except for the cases mentioned in the following paragraph, and constant across color channels. For the values of scattering and absorption, we chose marble as a representative material often used in rendering. The parameters, selected from [?], are those of the green channel of marble: $\sigma_s = 2.62 \text{ mm}^{-1}$ and $\sigma_a = 0.0041 \text{ mm}^{-1}$. The choice of marble was also motivated by the fact that a large enough portion of the phase function space we used produced realistic appearance for the corresponding optical parameters. Subjects in pilot studies expressed preference for images rendered with these parameters than others we obtained through appearance matching experiments for the materials shown in Figure 0??(a) ($\sigma_s = 0.99 \text{ mm}^{-1}$ and $\sigma_a = 0.01 \text{ mm}^{-1}$; and, $\sigma_s = 1.96 \text{ mm}^{-1}$ and $\sigma_a = 0.04 \text{ mm}^{-1}$).

Though the purpose of fixing other scattering parameters is to isolate the effect of the phase function, the obvious issue arises that the conclusions reached from our study may only be applicable to the selected coefficients σ_s and σ_a . For this reason, in some of our simulations we also used perturbations of the scattering parameters around the values for marble we selected. Specifically, we considered an “optically thin material” with $\sigma_s = 1.31 \text{ mm}^{-1}$ and $\sigma_a = 0.00205 \text{ mm}^{-1}$ (same albedo as marble and twice its mean free path); and a “low absorption material” with $\sigma_s = 2.6241 \text{ mm}^{-1}$ and $\sigma_a = 0.00041 \text{ mm}^{-1}$ (same mean free path as marble and absorption decreased by a factor of 10).

In terms of surface properties, we modeled the material as a smooth dielectric, effectively reducing its BSDF to a Dirac delta function. The motivation for this choice is to have a glossy surface with strong specular highlights, which in the past have been reported to be important for the perception of translucency [?; ?].

4.2 Shape and Lighting

We considered a number of candidate scenes, each using one of four shapes: the Stanford Lucy, Dragon, and Buddha; and our own star-shaped Candle (Figure 0??). This set includes complex and simple geometries. There are thick parts whose appearance is dominated by multiple scattering and is therefore not very sensitive to the shape of the phase function, and there are parts for which dense approximations are not appropriate and low-order scattering is important.

We placed the selected shape with a manually-selected pose in a subset of HDR environment maps from [?]: Campus, Eucalyptus, St. Peter’s, and a rotated version of Dining Room. This set includes the environment maps previously recommended for material perception [?], and represents different lighting conditions such as sidelighting and backlighting. We used a red-and-blue checkerboard as background. The checkerboard gave a patterned background to modulate the rays of light that scatter through the shape, and its use was also motivated from feedback from subjects in pilot studies, reporting difficulty segmenting grayscale objects from a uniform background. We did not use the checkerboard for the backlighting case, as we judged that in that case there was enough contrast with the background due to strong brightness changes at the border of the foreground object (see “Lucy + Dining room” in Figure 0??). Overall, our scenes were designed to feature image cues that have been related to the perception of translucency by others [?; ?] or that we ourselves believe to be related. These cues include fine geometric detail, specular highlights, sharp edges, intensity gradients, color bleeding effects, and strong backlighting.

Based on subject feedback in pilot studies comparing the different objects under the same illumination, we determined that: the Candle and Buddha geometry do not induce a sufficient diversity and number of image cues for translucency; and, the Dragon induces cues that were non-uniformly distributed between the front and rear body, and has confounding features (head versus body parts). The Lucy provided the best compromise between cue diversity and uniform spatial arrangement. As shown in Figure 0??(a), the dress exhibits fine details and sharp edges, specular highlights are present at the chest and the base, and a low-frequency intensity gradient is observed moving from the middle to the edge; furthermore, there is a plethora of thin (hands, wings, folds) and thick parts (body and base), which are very prominent under backlighting conditions.

4.3 Design Choices

Based on the discussion in the previous two subsections, we selected nine combinations of scattering parameters (σ_s and σ_a), geometry, and lighting to characterize translucent appearance, with the intention of studying, for each of these nine scenes, the image variation that is induced by changes in phase function. Our nine scenes can be grouped into three categories, based on which of the three components is variable.

- (1) **Scattering parameter variation.** Includes fixed shape (Lucy) and lighting (Campus), and varying σ_s and σ_a parameters: marble, optically thin material, and low absorption material.
- (2) **Shape variation.** Includes fixed lighting (Campus), σ_s and σ_a parameters (marble), and varying shape: Lucy, Buddha, Dragon, Candle.
- (3) **Lighting variation.** Includes fixed shape (Lucy), σ_s and σ_a parameters (marble), and varying lighting: Campus, Eucalyptus, St. Peter’s, and Dining room.

These are shown in Figure 0??. As discussed above, the “marble + Lucy + Campus” scene is the “base” configuration, and used in all three categories. In the following, whenever not explicitly specified, the σ_s and σ_a parameters of marble are used.

Conducting psychophysical experiments with each of these nine scenes and each of the 753 phase functions discussed in Section 0?? is possible nowadays through the use of crowdsourcing

marketplaces, which facilitate large-scale access to human workers. However, in a number of pilot studies we ran using Amazon MTurk [?], we found that workers routinely failed simple quality control tests. The data we collected from these studies exhibited large inconsistencies across different workers, and because of the amount of noise only provided us with coarse clusterings of images into two groups (roughly, more diffuse and more glass-like ones). On the other hand, doing psychophysical experiments at this scale in a carefully controlled laboratory environment is not tractable. We addressed this problem by combining computation with psychophysics, as described in the next two sections.

5. COMPUTATIONAL EXPERIMENT

We determined the axes of the perceptual space of phase functions in two steps. The first step (this section) was to render thousands of images with different phase functions sampled from our physical parameter space, and to use many instances of MDS with different between-image distance metrics to choose a single metric that provides consistent embeddings across our nine different scenes. The second step (Section 0??) was to use our selected metric to generate a smaller set of stimulus images for a psychophysical study that produces a perceptual embedding.

5.1 Image-based Metrics

The geometry of a space can be described by the distance metric used to compare the objects within it, and one should choose a metric that is representative of the properties considered important. In our setting, objects are phase functions, and we are interested in their effects on translucent appearance. We therefore adopt the following approach, motivated by previous work on opaque BRDFs [?]: we compare phase functions by comparing images they induce for the same scene geometry and lighting.

How do we measure distance between two images? Our choices inevitably introduce bias to our subsequent characterization of the space of phase functions, and if we choose poorly, our conclusions will not generalize to different scenes and metrics. A good metric, then, should be able to isolate translucency variations and characterize translucent appearance in a consistent way across different “nuisance” conditions, to borrow from the statistics literature terminology, such as lighting and shape. At the same time, consistency alone is not enough: a metric collapsing all images to a single point is guaranteed to be consistent across all imaginable variations! A good metric should be descriptive enough to represent the richness of variations in translucent appearance, and prevent such degeneracies.

To ameliorate these issues we took the following approach. We performed our computational analysis multiple times, each with a different combination of scene and image metric, and compared the embeddings produced by all of them. Then, we selected as our final metric that which produces the most consistent embeddings across different scenes. We see later in Sections 0??-0?? that the metric selected by this strategy does indeed possess perceptual relevance. In particular, it leads to predictions that can be validated psychophysically.

This approach required rendering thousands of images, and it was made possible by the availability of utility computing services. Specifically, for each of the scenes we selected in Section 0??, we rendered 753 images, one with each of the phase functions discussed in Section 0??. We used volumetric path tracing and the Mitsuba physically based renderer [?] on Amazon EC2 clusters [?].

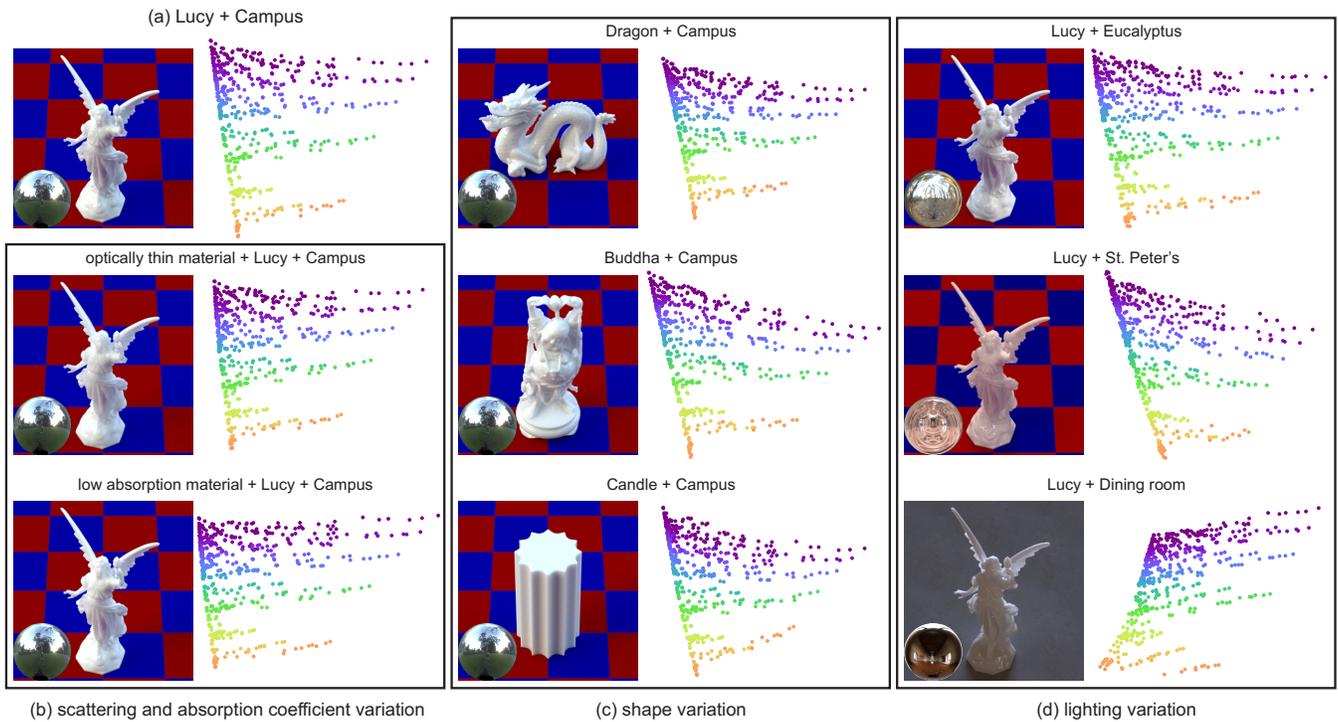


Fig. 4. Two-dimensional embeddings of images rendered with a representative set of phase functions sampled from the physical parameter space, produced using the cubic root metric. The embedded dots, one per rendered image for a total of 753 dots per embedding, are colored according to the square of the average cosine of the phase function used to render them (see Section 0??). To the left of each embedding is shown the scene it corresponds to. Scenes are grouped in terms of the type of change with reference to the “Lucy + Campus” scene shown in (a). Columns left: variation of scattering and absorption coefficients, middle: shape variation, right: lighting variation.

Overall, this produced a final set of 6777 high dynamic range images.

For candidate image metrics, we considered three: pixel-wise L_1 -norm, and L_2 -norm differences of high dynamic range (HDR) images; as well as the perceptually-motivated metric proposed by Ngan et al. [?], the L_2 -norm difference between the cubic root of HDR pixel values. In the following we refer to this last metric as the *cubic root metric*. All of our distance calculations ignored pixels of the background. In the supplementary material, we further report results from small-scale experiments using the image quality (mean-opinion-score) metric proposed by Mantiuk et al. [?], which we found to be consistent with those contained here.

5.2 Embeddings

Choosing a single scene and image metric induces a numerical distance between every pair of phase functions, and these distances can be used to find a low-dimensional embedding of the phase function space by classical MDS [?]. (Experimentation with metric MDS with different normalizations of the stress function produced almost identical embeddings.) We did this separately for each scene/metric pair, producing a set of such embeddings. Each embedding is a projection of the manifold occupied by phase functions in our ambient five-dimensional physical parameter space down to a Euclidean two-dimensional space where distances correspond to appearance differences as captured by the corresponding scene and metric. For all image metrics, the first two eigenval-

ues of the Gramian of the MDS output—the inner product matrix of the feature vectors produced by classical MDS—captured at least 99% of the total energy. This indicates that the space of the phase functions, as described by these metrics, is close to being two-dimensional, so we can compare the different embeddings by visualizing them in two dimensions.

The two dimensional embeddings produced using the cubic root metric for all of the scenes are shown in Figure 0??. For visualization, we first align all of the embeddings with the embedding for the “Lucy + campus” scene shown in (a), through a similarity transformation (rotation, uniform scaling, and translation) computed using *full Procrustes superimposition* [?]. We provide more details regarding Procrustes analysis in Appendix 0??. In Figure 0??(a), we show the embeddings produced from the different computational metrics we considered, for a representative subset of the scenes (the embeddings for all nine scenes produced using the L_1 -norm and L_2 -norm metrics are provided in the supplementary material).

By visually comparing the embeddings shown in Figures 0?? and 0??, we can observe that the structure of all of the embeddings is remarkably consistent, in particular across different scattering and absorption coefficients (Figure 0??(b)) and shapes (Figure 0??(c)). Larger differences are observed across lighting, and especially for the backlighting scene (“Lucy + dining room”) where there is a visible shear in the shape of the embedding. This is caused by the large changes in the brightness of the body of the Lucy across the different images; we discuss this effect in greater detail in Sec-

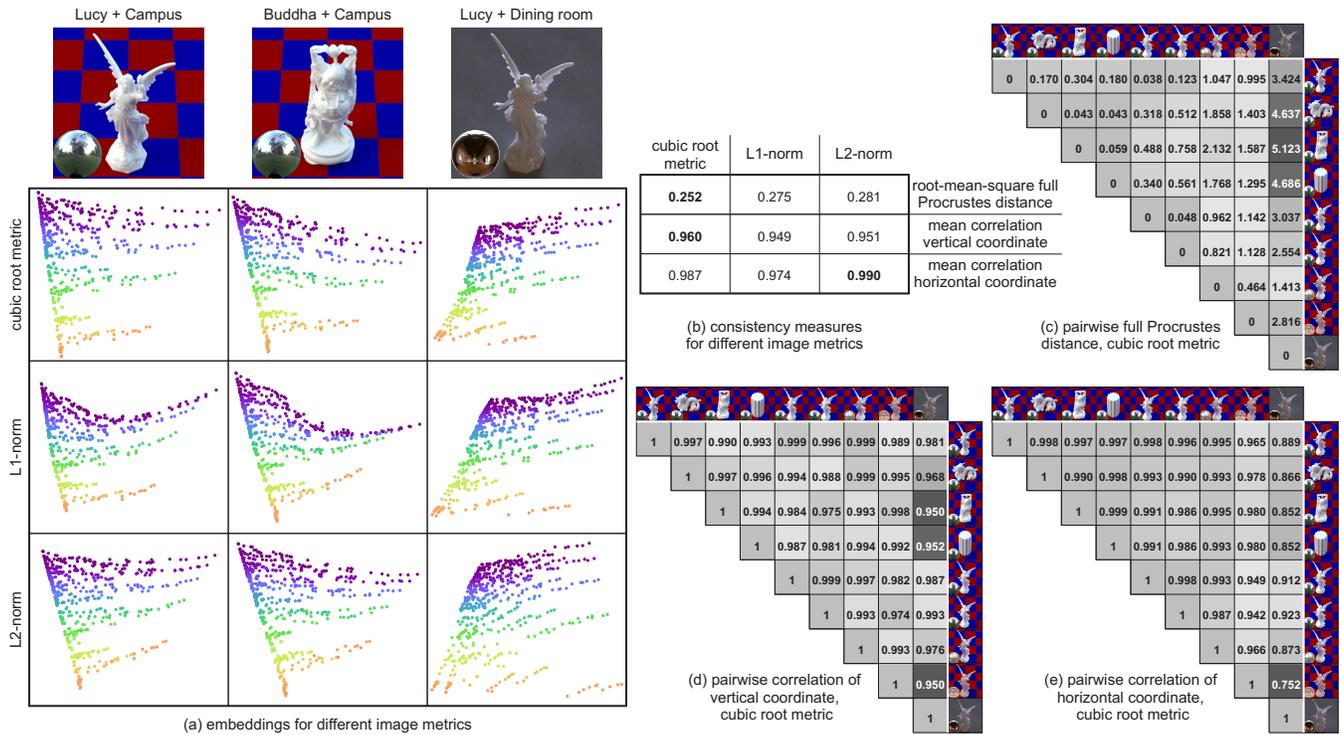


Fig. 5. Computational analysis results. (a) Comparison of two-dimensional embeddings produced using different image-based metrics, for three representative scenes. Points on embeddings are colored according to the square of the average cosine of the corresponding phase function. (b) Comparison of consistency of two-dimensional embeddings across the nine scenes of Figure 0??, for each of the three image metrics. (c)–(e) Measures of two-dimensional embedding similarity for each pair of the nine scenes, when the cubic root metric is used to produce the embeddings.

tion 0??. Even in this case though, the embeddings are quite consistent.

To quantify our observations, we compared embeddings using three measures of similarity: full Procrustes distances, and Pearson’s correlation coefficients between each of the two corresponding dimensions. In Figure 0??(b), we report the root-mean-square of the full Procrustes distances, and average correlation values across the embeddings for the nine scenes of Figure 0??, and for each of the three computational measures. Smaller variability, and therefore better consistency, corresponds to smaller values of the full Procrustes distances and larger correlation values. Based on this table, we found that the cubic root metric produces the most consistent results across scenes, and therefore we chose this metric as our measure of between-image distance.

In Figures 0??(c)–(e), we report numerical values for each of the three embedding similarity measures we considered and all the scene pairs, for the case when the cubic root metric is used to produce the embedding. In agreement with our observations above, the largest differences occur for the backlighting scene. Even for this scene though, we can see in Figures 0??(d)–(e) that there is very strong correlation between the corresponding coordinates of embeddings for different scenes, either vertical or horizontal. For each pair of the nine scenes, we performed a linear regression hypothesis test [?] for their corresponding coordinates, and found that the hypothesis of linear relation between them is statistically significant at the 99% confidence level—the same is true when embeddings are produced using the other two image metrics we considered.

6. PSYCHOPHYSICAL EXPERIMENT

The computational experiment of the previous section provides a consistent metric for measuring distance between same-scene images of different translucent materials. Our next steps were to select a set of stimulus images using this metric, and then recover a perceptual embedding through psychophysical analysis. Our psychophysical analysis parallels that used for the study of gloss by Wills et al. [?]: we collected relative similarity judgments using a paired-comparison approach and then found an embedding through kernel learning.

6.1 Procedure

We designed a paired-comparison experiment in which each observer was shown a series of triplets of rendered images and was required to choose an image-pair that is more similar. An example is shown in Figure 0??(b); the observer was required to indicate whether the center image is more similar to the image on the left or the image on the right. At the beginning of the experiment, observers were told that all images were rendered with the same shape and lighting, but with differing material properties. The exact instructions given to each observer are in the supplementary material.

In each trial (i.e., each triplet), the observer indicated their choice (left or right) by pressing the keys Q or P on a standard keyboard. The observers had unlimited viewing time for each trial but were told in advance that the average viewing time per trial was expected

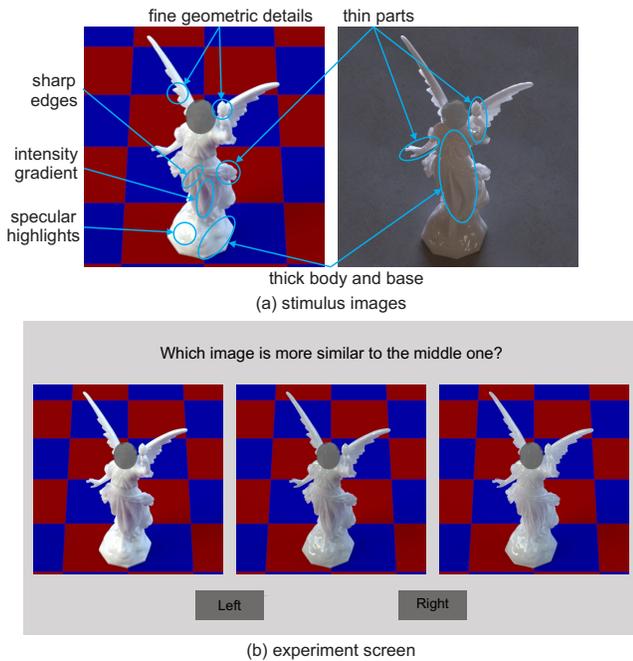


Fig. 6. Setup used for psychophysical experiment. (a) Rendered Lucy images used as stimuli for the sidelighting (left) and backlighting (right) experiments, with the face masked out. (b) Screen capture of the interface used in the experiment.

to be two or three seconds. Each subsequent trial followed immediately after the observer pressed a key to indicate their choice.

6.2 Stimuli

Though the analysis of Section 0?? provides clear motivation for one particular image metric (the cubic root metric), it does not provide clear choices for shape, lighting, and scattering and absorption coefficients. For the reasons described in Sections 0??, we chose the σ_s and σ_t values of marble for generating the stimuli for psychophysical analysis. We also selected the Campus and Dining room lighting environments, corresponding to sidelighting and backlighting conditions, which have been demonstrated to be important for translucency [?].

In terms of shape, we mentioned in Section 0?? that the Lucy has simultaneously the desirable properties of diversity of cues important for translucency, and uniform spatial arrangement. However, in pilot studies such as those used in Section 0??, we found that subjects invariably attended only to the facial region of the model, which alone does not have the full set of cues. We tried both cropping and occluding the face, and selected the latter because it preserved the context of the original scene. Subjects indicated that, as a result, they examined all of the body parts, responding to a more diverse collection of cues, and that they did not consider the blocked-out region distracting.

Based on the above discussion, we settled on two different scenes, shown in Figure 0??(a): the “Lucy + Campus” and “Lucy + Dining room”. We used these to perform two separate experiments of the kind described above, which we refer to as the sidelighting experiment and backlighting experiment, respectively. We emphasize that the consistency in Figures 0?? and 0?? suggests that these

choices of scene are not critical, and that our image-sets can be reasonably well-characterized by any of the scenes.

Having settled on scene geometry and lighting, the number of stimulus images was decided by balancing the competing needs of covering the physical parameter space and reducing our reliance on mixing judgments by different observers. The number of triplets requiring judgments increases combinatorially with the size of the stimulus set, and reducing the set to a point where a single observer can provide judgments for all possible triplets would not allow adequate coverage of the parameter space. Conversely, when the size of the stimulus set is allowed to grow, we end up with many different observers providing judgments for distinct sets of triplets, making inconsistency a real concern. Pilot studies suggested that an observer could complete 2000 trials in approximately two hours, and that increasing this number would result in increased fatigue and a drop in data quality. With this in mind, we set the stimuli set size to 40, implying a total of 30,000 possible triplets for each of the two experiments.

We used a clustering approach to select the 40 stimulus images from the “Lucy + Campus” dataset in Section 0??. Using affinity propagation [?], images were grouped into clusters having large intra-group affinity, and simultaneously, an exemplar image was selected from each cluster. We used, as input to the affinity propagation, pairwise distances between images computed by the cubic root metric, selected based on the results of Section 0??. Figures 0??(a) and 0??(a) show where the resulting exemplars reside in our previous computational embeddings for “Lucy + Campus” and “Lucy + Dining room”; by visual inspection, we observe that they are relatively uniformly spread across the entire embedding. Additionally, we verified that the intra-cluster variation was small by visually inspecting the members of each cluster. This is conveyed in the supplementary material, which includes visualizations of all of the member-images for each of a representative subset of the 40 clusters.

Each observer performed a subset of the 30,000 total triplets of either the sidelighting or the backlighting experiments; to create these subsets, we randomly permuted the 30,000 triplets and divided them into 15 parts with 2000 triplets per part. Each of the parts was assigned to one of 15 different observers for each experiment, for a total of 30 observers.

6.3 Implementation

The stimulus images were tone-mapped using a linear response and clipping at the same value across the dataset. They were presented in a dark room on a LCD display (ViewSonic HDMI 1080P, contrast ratio 1500:1, resolution 1920×1090 , sRGB color space, max luminance 700 cd/m^2 , 60:1 dynamic range, gamma 2.3). At a visual distance of 75 cm, images subtended 11.6° of visual angle.

Overall, 38 naive observers participated (15 observers for each of the two experiments, and 8 more for an inter-subject consistency test described below). All observers had normal or corrected-to-normal vision. Each observer completed 2000 trials, broken into 20 blocks of 100 trials each. Observers had the opportunity to rest between blocks, and before data collection commenced, each observer completed a small practice session of 20 trials.

6.4 Results

Consistency. Each observer in our two experiments provided responses for a unique set of 2000 triplets, and in order to obtain a perceptual embedding, we need to collect all 30,000 responses into a single pool for each experiment. For individual differences not to

factor into the interpretation of the results, it is important that different observers agree on the response for any particular triplet. To measure inter-observer consistency, we performed a separate small-scale study in which each of 8 observers evaluated the same set of 2000 randomly chosen triplets from the sidelighting experiment. By comparing their responses, we found consistency—defined as the number of observers that agree with the majority vote for each triplet—averaged across the 2000 triplets to be 75.94%. Both Pearson’s chi-squared and the likelihood ratio tests [?] show that the consistency of responses is statistically significant at the 99% confidence level, against the null hypothesis that observers decide based on chance.

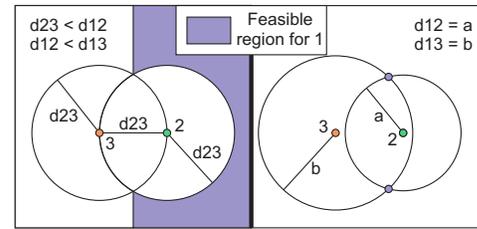
Perceptual embeddings. Having established consistency, we move on to computing a two-dimensional embedding using the pooled responses for each of the two experiments. An additional analysis step is required before we can compute this embedding using a method like MDS, because instead of having direct access to between-image distances as in the computational experiment of Section 0??, we only have relative constraints of the form $d(i, j) < d(i, k)$. Motivated by Wills et al. [?] and Tamuz et al. [?], we addressed this by using the relative constraints to learn a *kernel Gramian matrix* \mathbf{K} for the stimulus set. This is a 40×40 positive semi-definite matrix whose entries $(\mathbf{K})_{ij}$ represent inner-product values between images i and j in some feature space [?]. Once this matrix is learned, the between-image distances needed for computing our embedding are given by

$$d_{\mathbf{K}}(i, j) = \sqrt{(\mathbf{K})_{ii} + (\mathbf{K})_{jj} - 2(\mathbf{K})_{ij}}, \quad (7)$$

in direct analogy to the expression of Euclidean distance in terms of the standard dot product. By learning the kernel Gramian from the observers’ paired comparisons, our intention is that distances given by (0??) will provide a good match to perceptual distances. The learning algorithm we used is inspired by Wills et al. [?] and is described in detail in Appendix 0??.

Figure 0??(b) visualizes the final two-dimensional perceptual embedding of the stimulus set, obtained by first learning the kernel matrix and then using it as input to kernel PCA [?], for the sidelighting experiment (“Lucy + Campus”). Similarly, Figure 0??(b) shows the perceptual embedding for the backlighting experiment (“Lucy + Dining room”). In both cases, the perceptual embedding can be directly compared to the two-dimensional embedding of the same 40 images that was obtained computationally in Section 0??. (We provide, in the supplementary material, full-page versions of these perceptual and computational embeddings, with all 40 stimulus images overlaid on their corresponding locations.) The geometry of the perceptual and computational embeddings is very similar. In particular, we observe through close inspection that the overall ordering of the images is very consistent, though there are some notable differences. We note, for instance, that the ordering of images 2 and 6 in both embeddings is reversed. These two stimulus images appear significantly different than the rest of the dataset. This was reflected in subject responses for triplets including both images: we observed that when one image was in the middle, the two were very consistently judged as similar; whereas when neither of them was in the middle, there was no consistent trend in the responses. As a consequence, in the perceptual embedding they were both correctly placed at the upper-right corner of the embedding, but their relative position was inferred reversed. Differences are also seen in the upper-left corner, where all stimulus images have a comparatively similar appearance resembling marble. It is important to note that, due to the use of relative con-

straints, the ordering of the points of the embedding is more important than their absolute locations. We use the drawing in the inset figure to the left to provide some intuition for why this is true:



2 and 3) on the plane. The right and left part of the drawing shows all the admissible configurations for compatible absolute and relative constraints, respectively. Absolute constraints permit only two isolated locations as solutions (shown as blue circular points, on the right), whereas with relative constraints possible solutions span almost an entire half-plane (shown in blue, on the left).

In Figures 0??(c) and 0??(c), a subset of the embedded images are shown arranged in a grid consistent with their positions in both the computational and perceptual embedding. In both figures, we also report correlation values between corresponding coordinates of the perceptual and computational embeddings, for each of the two scenes. In all four cases, a linear regression hypothesis test showed that the hypothesis of linear relation between corresponding coordinates is statistically significant at the 99% confidence level.

As we mentioned at the beginning of this section, we found that there was significant consistency in the responses provided by the study subjects. In order to further examine the stability of the perceptual embeddings to noise in the responses, we performed a bootstrapping analysis: we randomly selected a percentage of the 30,000 triplets from the sidelighting experiment equal in size to the percentage of inconsistent responses, switched the (binary) response provided by the subjects for those, and then learned an embedding using the perturbed data. We report the results of this analysis in the supplementary material, where we observe that the embeddings learned from such perturbed data sets are very similar to the embedding learned from the actual user data. We believe that this robustness to noise is due to the strong and correlated constraints the full set of relative comparisons imposes on admissible geometric configurations, as well as the generalization properties of the regularization term we use in the learning algorithm (we discuss this point in greater detail in Appendix 0??).

7. ANALYSIS

In this section, we characterize the geometry of the two-dimensional perceptual embedding, and we derive computational tools for translucent material design, including analytic expressions for two distinct perceptual axes and a perceptual metric for comparing tabulated phase functions. First, though, it is worth discussing the significant implications of the consistency between the perceptual embeddings of the previous section and the computational embeddings of Section 0??.

Consistency between perceptual and computational embeddings implies that any geometric structure we discover in the computational embedding—including, for example, expressions for axes and learned metrics—is likely to exist in the perceptual space as well. This is a very useful property, as it means that we can leverage powerful data mining techniques to discover geometric structure in the large computational datasets of Section 0??. and then interpret this as structure in the perceptual space. Without consistency be-

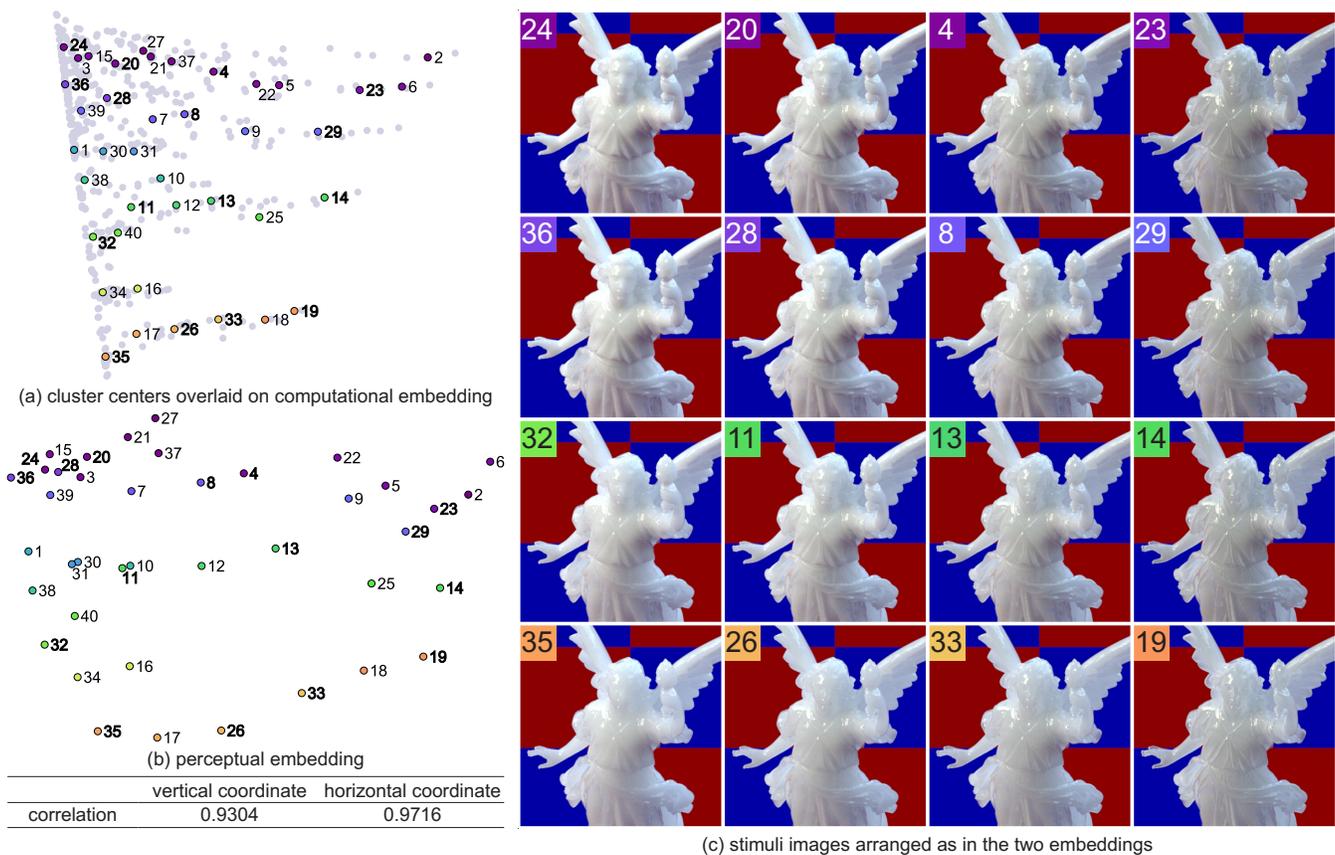


Fig. 7. Comparison of computational and perceptual embeddings for the sidelighting experiment. (a) Locations of the stimulus images selected by the clustering algorithm (overlaid on the two-dimensional computational embedding) are well-distributed. (b) Two-dimensional embedding learned from the data collected from the psychophysical experiment. The inset table shows values for Pearson’s correlation coefficient between corresponding coordinates of the cluster centers on two embeddings. (c) A subset of the stimulus images, arranged in a grid according to their positions in the computational embedding. The same grid is also consistent with the positions of the images on the perceptual embedding. Points on embeddings and images on the grid are colored according to the square of the average cosine of the corresponding phase function, and consistently labeled across (a)-(c).

tween embeddings, such tools would not be applicable, because it is not feasible to run psychophysical paired comparison experiments at the scale required to generate embeddings of a sufficiently large set of images. (Perceptually embedding a stimulus set of even 500 images would require 60 million human judgments.) Without this consistency, we would be left to draw conclusions from dozens of embedded points instead of thousands, and we would be limited to using a much weaker set of tools.

A similar benefit comes from the consistency of computational embeddings across different scenes (Section 0??). These results suggest that structure we discover in one of our two-dimensional embeddings is likely to generalize, at least to some degree, to different scene geometries and lighting conditions.

In light of this discussion, most of the analysis in this section is based on the computational two-dimensional embedding obtained with the “Lucy + Campus” scene (Figure 0??(a)), with the expectation that results also apply to the perceptual space of phase functions. In the supplementary material, we further report the results of the same analysis for the full Procrustes mean of the nine embeddings of Figure 0??.

7.1 Parameterization of the 2D Perceptual Space

We begin by characterizing the two dimensions of the computational embedding in terms of functionals of the phase function. We seek uniform parametric representations of the two dimensions, meaning that: 1) a fixed increment to any of the two parameter values induces an equal-length movement in the two-dimensional space regardless of the initial parameter values before the increment was applied; and 2) by adjusting both parameters, we can get from any point in the two-dimensional space to any other. Parameters defined in this way facilitate designing a phase function to correspond to any location on the embedding.

To identify our two parameters, we considered a large candidate set of statistical functionals of probability distributions, such as mean, variance, kurtosis, skewness, moments of different orders, and different analytic functions of these quantities. We then examined the linear dependence between these candidate functionals and the two coordinates of the embedding, using a combination of intuition and exhaustive computational search. To quantify the linear dependence, we used Pearson’s correlation coefficient [?], which takes values $r \in [-1, 1]$. Larger absolute values indicate stronger

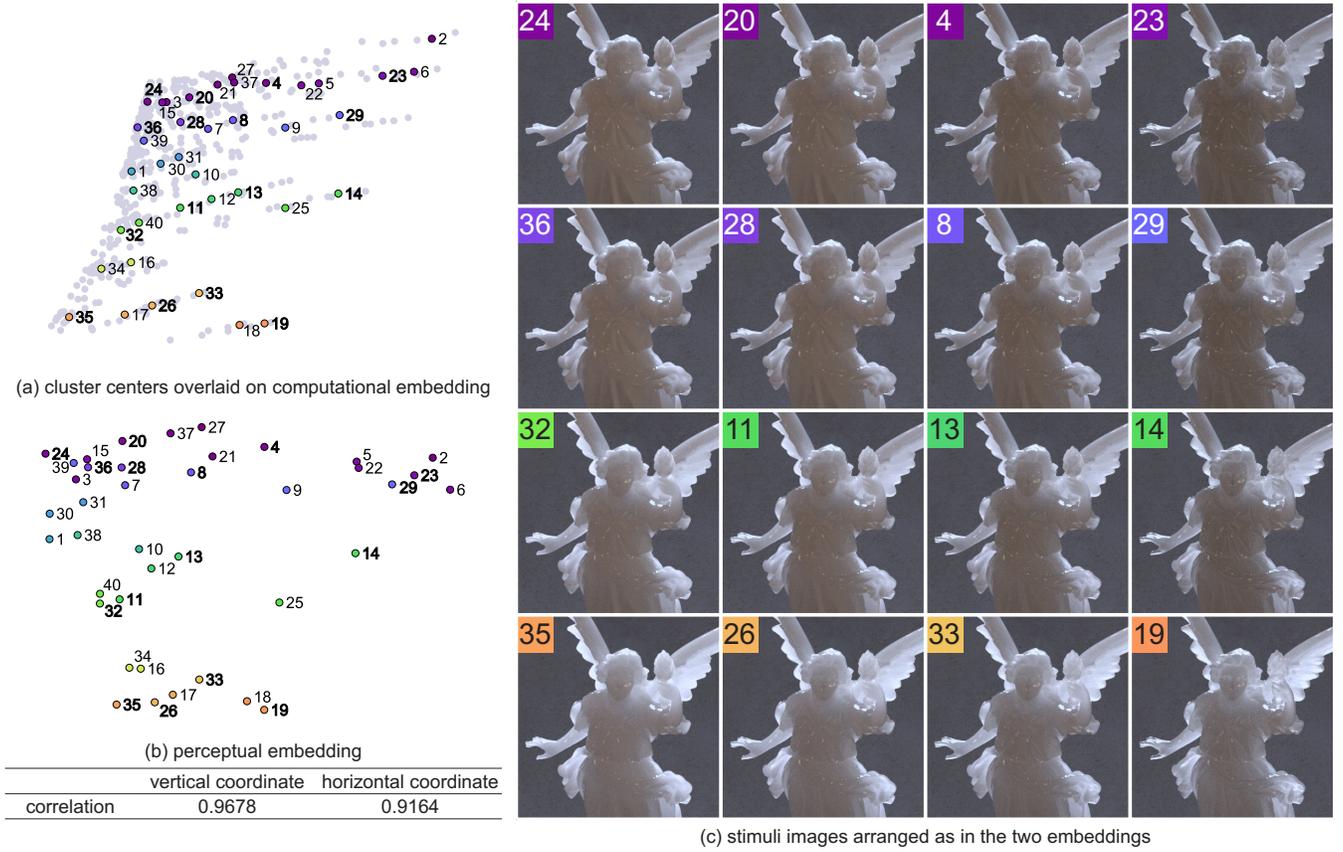


Fig. 8. Comparison of computational and perceptual embeddings for the backlighting experiment. (a) Locations of the stimulus images selected by the clustering algorithm, overlaid on the two-dimensional computational embedding. (b) Two-dimensional embedding learned from the data collected from the psychophysical experiment. The inset table shows values for Pearson’s correlation coefficient between corresponding coordinates of the cluster centers on the two embeddings. (c) A subset of the stimulus images, arranged in a grid according to their positions in the computational embedding. The same grid is also consistent with the positions of the images on the perceptual embedding. Points on embeddings and images on the grid are colored according to the square of the average cosine of the corresponding phase function, and consistently labeled across (a)-(c).

linear dependence, while lower ones imply either lack of dependence or absence of a linear relationship. Since any functional can be negated, we are not concerned with the sign of the coefficient, and for the remainder of this section reported values correspond to the absolute value $|r|$.

Vertical dimension. We observed through experimentation that the vertical dimension of the computational embedding is parameterized well by the average cosine \bar{C} of the phase function, defined in (0??). Based on this observation, we performed an exhaustive search over powers of the average cosine, \bar{C}^a , for $a \in [1.5, 2.5]$ and found that correlation is maximized by $a = 2.05$, where $|r| = 0.9939$. For the average cosine and its square, $a = 1$ and 2, the correlation is equal to 0.9621 and 0.9939, respectively. The square of the average cosine is very strongly correlated to the vertical coordinate, and the average cosine slightly less so. A linear regression test showed that the hypothesis of linear relation between \bar{C}^2 and the vertical coordinate is statistically significant at the 99% confidence level. Figure 0?? shows the two-dimensional embedding, colored using \bar{C}^2 . This result is the reason why we use \bar{C}^2 to

color the majority of two-dimensional embeddings throughout the paper, most notably Figure 0??.

In Table 0??, we report the Pearson’s correlation coefficient values between powers of the average cosine \bar{C} and the vertical coordinate of the cubic-root-metric embeddings for all nine scenes of Figure 0??. We observe that, in all cases, the square of the average cosine is strongly correlated with the vertical coordinate, and the best power $a \in [1.5, 2.5]$ is close to 2.

Horizontal dimension. By visually examining phase functions at different points across horizontal lines of the embedding, we observed that they differ in terms of the spread of their probability mass. We therefore considered a number of functions of the second moment of the cosine M_C , defined in (0??), as well as its variance,

$$V_C = 2\pi \int_{\theta=0}^{\pi} (\cos \theta - \bar{C})^2 p(\theta) \sin \theta d\theta. \quad (8)$$

Table 0?? shows the correlation values of the horizontal coordinate with different functions of these quantities, the largest being

Table I. Parameterization of the vertical dimension of the embeddings of Figure 0???. (“Correlation” refers to the absolute value of Pearson’s correlation coefficient.)

scene	correlation with \bar{C}	correlation with \bar{C}^2	best power $a \in [1.5, 2.5]$	correlation with \bar{C}^a
“Lucy + Campus”	0.9621	0.9939	2.05	0.9939
“Lucy + Dragon”	0.9480	0.9871	2.22	0.9879
“Lucy + Buddha”	0.9365	0.9766	2.24	0.9774
“Lucy + Candle”	0.9388	0.9765	2.15	0.9768
“optically thin material”	0.9701	0.9950	1.89	0.9952
“low absorption material”	0.9699	0.9958	1.93	0.9959
“Lucy + Eucalyptus”	0.9586	0.9917	2.07	0.9918
“Lucy + St. Peter’s”	0.9399	0.9772	2.17	0.9776
“Lucy + Dining room”	0.9658	0.9913	1.96	0.9914

Table II. Absolute value of the Pearson’s correlation coefficient between functions of variance and the second moment of cosine, and the horizontal coordinate of the two-dimensional embeddings from Figure 0???(a).

Expression	$ r $	Expression	$ r $
$1/\sqrt{1-M_C}$	0.9122	$1/\sqrt{M_C}$	0.6421
$\sqrt{1-M_C}$	0.8209	V_C	0.6102
M_C	0.7354	$\bar{C}/(1-V_C)$	0.6094
$\sqrt{M_C}$	0.7064	$1/M_C$	0.6082
$\sqrt{(1-M_C)\bar{C}}$	0.6895	$\sqrt{V_C}$	0.5590

Table III. Parameterization of the horizontal dimension of the embeddings of Figure 0???. (“Correlation” refers to the absolute value of Pearson’s correlation coefficient.)

scene	correlation with $1/\sqrt{1-M_C}$
“Lucy + Campus”	0.9122
“Lucy + Dragon”	0.9020
“Lucy + Buddha”	0.9172
“Lucy + Candle”	0.9181
“optically thin material”	0.9129
“low absorption material”	0.8912
“Lucy + Eucalyptus”	0.9256
“Lucy + St. Peter’s”	0.8757
“Lucy + Dining room”	0.9101

achieved by

$$D_C = 1/\sqrt{1-M_C}, \quad (9)$$

which is related inversely with the second moment of the cosine M_C . The correlation value for this case, $|r| = 0.9122$, indicates that D_C is strongly correlated with the horizontal coordinate. Figure 0?? shows the two-dimensional embedding, colored using D_C . We observe that the axes corresponding to changes in D_C are not exactly horizontal, but instead diagonal with small slope. As before, a linear regression test showed that the hypothesis of linear relation between D_C and the horizontal coordinate is statistically significant at the 99% confidence level.

In Table 0??, we report the Pearson’s correlation coefficient values between the quantity $D_C = 1/\sqrt{1-M_C}$ and the horizontal coordinate of the cubic-root-metric embeddings for all nine scenes of Figure 0???. We found that, for all scenes, the correlation with this quantity is the largest among all the other quantities we show in Table 0??, though the order after the first was not identical across scenes.

7.2 Phase Function Distance Metric

The image-driven metric adopted in Section 0?? is very cumbersome because comparing two phase functions requires first rendering two images and then comparing those. In this section, we learn an efficient distance metric that operates directly on tabulated phase functions and provides an approximation to this image-driven metric.

For two phase functions p_1, p_2 , we considered metrics of the form

$$d_w(p_1, p_2) = \int_{\theta_1=0}^{\pi} \int_{\theta_2=0}^{\pi} (p_1(\theta_1) - p_2(\theta_2))^2 w(\theta_1, \theta_2) d\theta_1 d\theta_2, \quad (10)$$

with symmetric and non-negative weight function w . Note that if

$$w(\theta_1, \theta_2) = \begin{cases} 1, & \theta_1 = \theta_2 \\ 0, & \theta_1 \neq \theta_2 \end{cases}, \quad \theta_1, \theta_2 \in [0, \pi], \quad (11)$$

the metric d is the usual squared distance. Even though the squared distance is a straight-forward way to compare two phase functions, it does not predict differences in translucent appearance well. To demonstrate this, we show in Figure 0??(b) the two-dimensional embedding obtained using squared distance between phase functions as input to MDS. The geometry of the obtained embedding is very different from the embedding produced using the image-driven metric that we are trying to approximate.

To improve this approximation, we learned the weight function w by finely discretizing the phase function domain $\theta \in [0, \pi]$ and optimizing the weight matrix representation of w of the resulting Mahalanobis distance. We used the metric learning algorithm of Davis et al. [?] with the objective of having distances $(p_1 - p_2)^T w (p_1 - p_2)$ between discretized phase functions p_1, p_2 be close to those given by the image-driven metric. We found that the learned weight function had large values on its diagonal, that is when $\theta_1 = \theta_2$, and much smaller values otherwise. Figure 0??(a) provides a visualization, in the form of a polar plot, of a smoothed version of the diagonal of the learned weight function (shown in blue), and how it compares with the squared distance from (0??) (shown in red). (A visualization of the full two-dimensional function as a matrix is provided in the supplementary material.) The metric emphasizes differences in the backward scattering lobe. To assess the quality of the approximation to the image-based metric afforded by this learned weight function, we also show a two-dimensional embedding of the phase functions produced by MDS with distances given by the learned metric in Figure 0??(c). This embedding indeed can be observed to have qualitatively similar structure to that obtained using the image-driven metric (Fig-

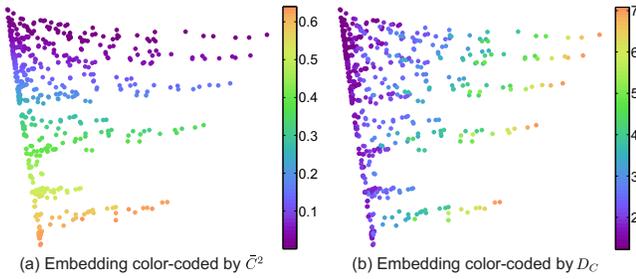


Fig. 9. Parameterization of the computational two-dimensional embedding. (a) Embedded points are colored using the square of the average cosine of the corresponding phase function, which parameterizes the vertical dimension of the embedding. (b) Embedded points are colored using function (0??), which parameterizes well the horizontal dimension.

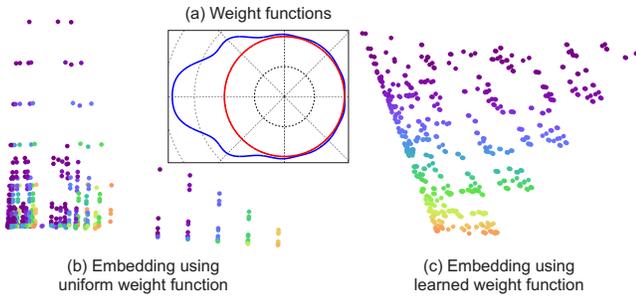


Fig. 10. Results of metric learning for phase functions. (a) compares the polar plots of the diagonal of the learned weight function (shown in blue, smoothed to 15 first Legendre polynomial coefficients for visualization), and the uniform one, corresponding to the L_2 -distance (shown in red). (b) and (c) show the embeddings of phase functions using the L_2 -distance and the learned metric, with points colored according to the square of the average cosine of the corresponding phase function (compare with Figure 0??(a)).

ure 0??(a). Quantitatively, their full Procrustes distance is equal to 0.3381; the absolute value of the Pearson’s correlation coefficient between corresponding dimensions of the two embeddings is equal to 0.96 and 0.89, for the vertical and horizontal coordinates respectively; and in both cases the hypothesis of linear relation between the corresponding coordinates is statistically significant at the 99% confidence level. For comparison, the full Procrustes distance of the embedding of Figure 0??(b), produced using the simple squared distance, from that of Figure 0??(a) is equal to 0.7760; and the Pearson’s correlation coefficient between corresponding coordinates is 0.3044 and 0.5729, for the vertical and horizontal coordinates respectively.

8. PHASE FUNCTION AND MATERIAL DESIGN

Our analysis so far has shown that our expanded space of phase functions can produce a two-dimensional perceptual appearance space, with the two axes being controlled by functionals of the first and second moments of the phase function. In this section, we explore this space further with material design applications in mind, and seek to answer the following questions: what types of visual effects can be produced and with what phase function in different

parts of the appearance space; and how to select a phase function for achieving a target material appearance.

8.1 Role in Translucent Appearance

Single-lobe phase function models that consist of either one HG or one vMF lobe provide a single-parameter family of distributions. Thus, by varying their shape parameter (g or κ respectively) one can only access materials that lie on a one-dimensional curve within the two-dimensional embedding. Moreover, we find that for both the single-HG and single-vMF cases, the resulting one-dimensional curve of materials is almost vertical and near the left-most part of the embedding. (Dense samplings of these one-dimensional curves are depicted in Figure 0??(a)-(b).) This shows that using only these single-lobe phase function models limits the set of reachable translucent appearances to those on the left side of the embedding.

Mixtures of single-lobe and isotropic phase function models are often used as an extension of the simple single-lobe model [?]. Such phase functions are parameterized by two parameters, one for the shape of the forward scattering lobe (g or κ), and another for the weight of the lobe in the linear mixture. Despite this added flexibility, the set of translucent appearances that can be reproduced is still restricted to the left side of the embedding, as shown in the top row of Figure 0??.

Double-lobe phase function models are necessary, therefore, for accessing the complete two-dimensional space of appearances. To show which parts of the two-dimensional space can be accessed using different combinations of HG and vMF lobes, the second row of Figure 0?? highlights materials that are created when the forward lobe of the mixture is one of either HG (Figure 0??(c)) or vMF (Figure 0??(d)) type. While most of the embedding can be reached using either an HG or vMF forward lobe, there is a region on the right that can only be accessed using a vMF forward lobe. The importance of the backward lobe is also evident from Figure 0??(a), where we can see that the learned weights for comparing phase functions place strong emphasis on the backward scattering part.

Visualization of the appearance corresponding to the middle and right parts of the embedding, can be achieved by looking at Figures 0??(c) and 0??(c). These correspond to phase functions that reside in an approximately-uniform grid of positions within the two-dimensional embedding, and the position of each image within Figures 0??(c) and 0??(c) is indicative of where the corresponding phase function exists in that grid.

We see that moving from top to bottom (across the vertical dimension parameterized by \bar{C}^2) results in more diffusion, an effect similar to that achieved by increasing the mean free path. Moving from left to right (across the horizontal dimension parameterized by D_C) results in greater surface detail and more “glassy” appearance. By selecting different points in the embedding, we can achieve different trade-offs between diffusion and sharpness. In the middle of the embedding we see the emergence of very distinctive appearance, part translucent, part glassy with sharp details, that is characteristic of materials such as high quality white jade (see Figure 0??).

To demonstrate the importance of the middle part of the embedding, Figure 0??(c) shows renderings of the “Lucy + Campus” scene (with the checkerboard replaced by a concrete background), where we have matched the appearances of marble and white jade. For the marble image (Figure 0??(c)-(3)), a broad single-lobe phase function (HG with $g = 0$) is used to produce the characteristic dif-

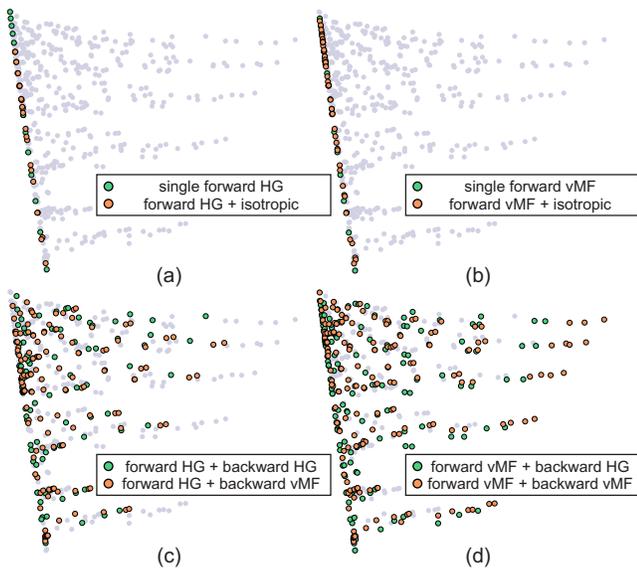


Fig. 11. Locations of points corresponding to different types of phase functions on the two-dimensional computational embedding. Embedded points corresponding to the specific type of phase function are shown colored according to the squared average cosine of the phase function, whereas the rest of the embedding is faded out. Top row: single lobe and single lobe + isotropic combinations are limited to the left part of the embedding. Bottom row: two lobe combinations achieve appearance towards the right in the embedding.

fusion in the body of the object. In contrast, matching white jade requires much less diffusion and greater preservation of surface detail. Attempting to match this appearance with only HG, that is by staying at the left part of the embedding only, produces the middle image in Figure 0??(c)-(4), which has most detail blurred out. On the other hand, by using a double-lobe phase function to move towards the middle of the embedding, we achieve the result shown at the right of Figure 0??(c)-(5), which captures both of the desired characteristics. High resolution versions of these renderings are provided in the supplementary material.

Thus, a result of our study is that including a secondary small backward lobe in the phase function is necessary to achieve glassy translucent appearance, like that of white jade. This is a visually interesting region of translucent materials that is neither highly diffusive nor highly transparent.

The effect of material, lighting, and shape. Our results in Section 0?? show that the geometry of the appearance space is consistent across some variations of material (in the sense of different scattering and extinction coefficients), shape and lighting. This indicates that our conclusions in the above discussion generalize to the nine scenes of Figure 0??, and to novel scenes. The computational embedding for the “Lucy + Dining room” scene, corresponding to backlighting, is the one exhibiting the largest structural differences from the other embeddings, and for this reason we discuss it in greater detail. In particular, the embedding shows a skewing, being “stretched” across the diagonal from bottom left to top right. This effect can be understood by examining the grid of images in Figure 0??(c), and specifically all the diagonals from bottom left to top right. We observe that this diagonal direction corresponds to a decrease in brightness, both in the thin parts (wings, hands)



Fig. 12. Images of the same materials as in Figure 0??(c), in a backlighting environment.

and the thick body. In a lighting environment as dramatic as back-lighting, this change in overall brightness may be an important cue that strongly affects the responses of observers, confounding other cues. Such large changes in overall brightness also strongly affect the image metrics used to produce the computational embedding.

This strong change in brightness when moving from bottom to top, and from left to right on the embedding, can be related, by considering the physics of subsurface scattering, to the parameterization we derived in Section 0?? for the vertical and horizontal dimensions of the embedding. Moving from bottom to top on the embedding results in less forward scattering and more isotropic phase functions (smaller \bar{C}), implying that less light reaches from the back of the Lucy, where it enters the medium, to the front, where the camera observes it. Moving from left to right on the embedding results in phase functions with larger variance (larger M_C and D_C), implying that more light gets scattered towards side-directions, and therefore less light exits from the front for the camera to observe. Strong backlighting conditions emphasize this effect of phase function in brightness, confounding its other effects in translucent appearance. Geometrically, this manifests itself as a shear stress of the two-dimensional embedding, suppressing to an extent its two dimensions to a single diagonal dimension.

Despite this strong confounding effect, our numerical analysis in Sections 0?? and 0?? demonstrates that the two-dimensional structure of the embedding is very consistent with those for the other scenes. Furthermore, by looking at Figure 0??(c), we can observe the same qualitative changes in translucent appearance for vertical and horizontal movements on the embedding (changes in amount of diffusion and surface sharpness, respectively), as those we described earlier in this section for the “Lucy + Campus” scene (Figure 0??(c)). To highlight this consistency, we show in Figure 0?? renderings of the same three materials as in Figure 0??(c), using the backlighting environment. We observe again that, using the same navigation tools as before to utilize the full two-dimensional embedding, we can control the trade-off between diffusion and detail, and achieve the part translucent, part glassy with high sharpness, appearance of white jade. High resolution versions of these renderings are provided in the supplementary material.

The importance of phase function. We can summarize the above discussion as a list of salient points relating the phase function to translucent appearance:

- (1) Our expanded phase function space, of double-lobe models with HG and vMF distributions, creates a perceptual two-dimensional space of translucent appearance.
- (2) Our expansion of the phase function space is non-trivial; the expanded space is *necessary* for producing all of the appearance space, and simpler single-lobe or forward-isotropic mixture models can only produce a one-dimensional subspace.

- (3) The expansion of the appearance space itself is physically important, allowing to achieve realistic appearance corresponding to materials such as white jade.
- (4) The two dimensions of the appearance space can be intuitively described as controlling diffusion and sharpness.
- (5) The appearance space remains relatively unaffected by changes in scattering parameters other than the phase function, shape, or illumination, the last having the strongest effect.

The above points highlight the fact that the phase function can crucially affect the appearance of translucent objects. In the past, this effect has often been overlooked, in favor of emphasizing the role of the other scattering parameters (mean free path, albedo). This approach is justified when simulating higher-order scattering, which according to similarity theory can be modeled accurately using the isotropic phase function. In turn, this implies that phase function is less important for thick objects and optically dense materials, where higher-order scattering dominates appearance. However, most interesting real objects have both thin and thick parts, and many materials exist that are not as dense. In these cases, the contribution of the low-order scattering component, which is strongly affected by the shape of the phase function, to the overall appearance becomes non-negligible and, as our results demonstrate, perceptually important.

8.2 Phase Function Design

As we established in the previous subsection, selecting the right phase function can be critical for achieving a desired appearance in a material design application. This can be formulated as a dual task of selecting a convenient parametric family of phase functions, and then finding appropriate values for the corresponding parameters.

The qualitative observations of Section 0?? and quantitative analysis in Section 0?? can facilitate this process in a two-fold way. First, by considering the intuitive description of vertical and horizontal movements on the two-dimensional translucent appearance space (changes in amount of diffusion and sharpness, respectively), we can approximately localize the target appearance on this space. This, in turn, allows us to use Figure 0?? to guide our selection of parametric family of phase functions.

Second, having settled on a parametric family, we can reparameterize it in terms of the analytic expressions \bar{C}^2 and D_C of the first and second moment of the phase function, respectively, that we derived in Section 0??. These quantities uniformly parameterize the vertical and horizontal dimensions of the appearance-based two-dimensional embedding of phase functions. This further implies that they can be used as quantitative “knobs” to control the appearance effects of movements on the translucent appearance space in a perceptually uniform way: specifically, \bar{C}^2 controls amount of diffusion, and D_C amount of sharpness. Writing the selected parametric family in terms of these quantities enables us to do perceptually uniform interpolation directly in the phase function space, which can be helpful for fine-tuning the parameter values.

As an instructive example, consider the case of single-lobe phase functions. Based on the discussion of the previous subsection, these families span an almost vertical one-dimensional subspace of the appearance space, which is uniformly parameterized by \bar{C}^2 . For the case of HG lobes, where the shape controlling parameter g is equal to \bar{C} , this suggests that a perceptually uniform reparameterization can be obtained by simply using g^2 instead of g . Similarly we can interpolate appearance in a more perceptually uniform way for the vMF family, through the equation,

$$\bar{C}^2 = (-1/\kappa + \coth \kappa)^2. \quad (12)$$

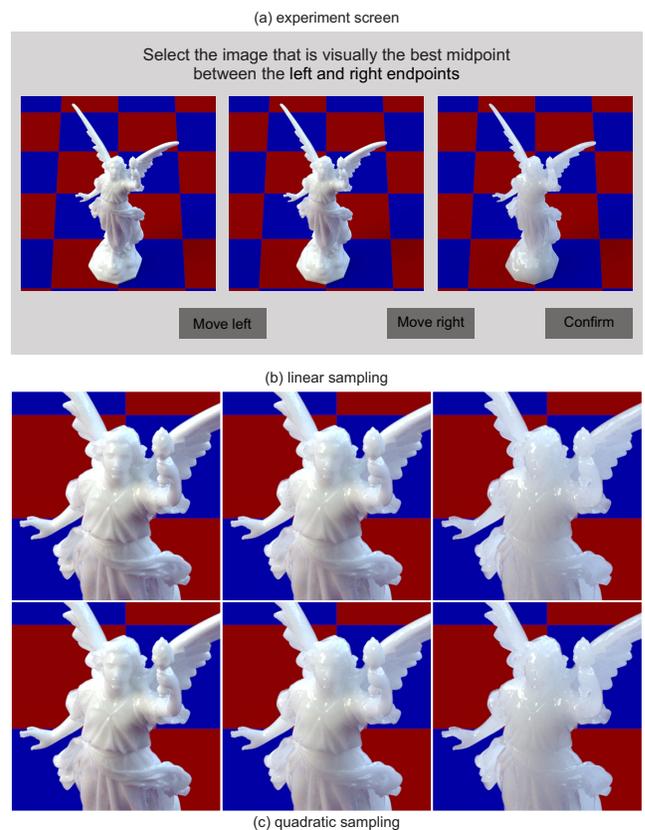


Fig. 13. Perceptually uniform parameterization of a single-lobe HG. (a) Graphical user interface used to validate perceptually uniform interpolation. (b)-(c) Both rows: left image ($g = 0$), right image ($g = 0.9$). Middle image rendered using g corresponding to the midpoint between these values: by linear sampling ($g = 0.45$, top row), or quadratic sampling according to g^2 ($g = 0.6364$, bottom row). By looking at each row from the left to the right, we observe that the appearance of the middle image in the bottom row is perceptually a better “middle point” than in the top row.

In order to further validate these findings, we designed a simple graphical user interface, shown in Figure 0??(a). The three images shown are all rendered with a single-lobe HG phase function; those to the left and right correspond to values $g = 0$ and $g = 0.9$, respectively, whereas the g value for the middle image can be increased or decreased by the user, by pressing the corresponding buttons. This is achieved internally using pre-rendered images for a dense set of g values in the interval $(0, 0.9)$. We then used this interface to perform the following psychophysical experiment: subjects were asked to adjust the middle image until they find the one that is, in their opinion, the best visual midpoint between the images shown to the left and right, at which point their selection was recorded. Our previous findings predict that this middle image should correspond to a g value different from the arithmetic mean 0.45 and near the quadratic mean 0.6364 of the endpoints 0 and 0.9.

We used two different scenes for the experiment, “Lucy + Campus” and “Buddha + Campus” from Section 0??. For selecting the g values used to create the interpolation sequence, we considered both linear (uniformly in g) and quadratic (uniformly in g^2) sampling from the interval $(0, 0.9)$. Then, each subject performed the

Table IV. Mean value and 99% confidence interval for the image midpoint experiment described in Figure 0??(a). In all cases, the mean response is close to the value 0.6364 predicted by our analysis.

Parameterization	“Lucy + Campus”	“Buddha + Campus”
Linear	0.6583 ± 0.1390	0.6786 ± 0.0476
Quadratic	0.6507 ± 0.0900	0.6945 ± 0.0612

experiment using either only the linear or only the quadratic sampling setting for both scenes. We emphasize here that the reason for considering both parameterization settings was not to compare their utility for design applications, which was not the purpose of the experiment; rather, we wanted to make sure that the results would not be biased by the choice of sampling. In Table 0??, we report the mean g values selected by the subjects for each sampling and scene, as well as 99% confidence intervals. We observe that, in all cases, the mean of the selected g values is significantly skewed towards the quadratic midpoint of 0.6364 predicted by our analysis, and the exact numerical value is always inside the confidence intervals; using ANOVA [?], we found that there is no statistically significant difference between the four sample means. To the contrary, using two-sided t-tests [?], we found that in all cases the sample mean is statistically significant from the value 0.45 predicted by linear scaling. These results are in agreement with the predictions of our analysis.

For visualization, in Figure 0??(b)-(c), we compare renderings of the “Lucy + Campus” scene of Section 0??, with single-lobe HG functions. The left-most and right-most renderings of each row correspond to a phase function with $g = 0$ and $g = 0.9$, respectively. Renderings in between use g values sampled from this interval, either uniformly in the parameter space (top row, $g = 0.45$), or uniformly according to g^2 (bottom row, $g = 0.6364$). We observe that transitions from left to right are more perceptually uniform in the second case. We include in the supplementary material dense interpolation sequences in video form, to demonstrate this effect.

9. CONCLUSIONS

In this paper, we focused on the effect of phase functions in the appearance of translucent materials, and highlighted that for many realistic geometries and materials this effect is perceptually important. We explored an expanded space of phase functions, double-lobe models with both HG and vMF distributions, and showed that it can produce a much richer set of realistic translucent appearance. We also took first steps towards parameterizing the perceptual dimensions of translucency, by deriving computational tools for describing and navigating this expanded space of phase functions in a perceptually meaningful way. We achieved our results using a new methodology that combines computational analysis and psychophysical experiments, and we found strong evidence that our results generalize to novel scene shapes and lighting environments. That said, additional tests will be required to assess generalization more completely.

Our findings can be used for the development of interfaces for design of translucent materials by novice users, without the need for understanding the underlying physical models. Inspiration for how to proceed towards that direction can be drawn from several analogous studies for BRDF editing [?; ?; ?; ?]. The utility of perceptually intuitive instead of physical parameters in interfaces for translucent appearance editing needs to be further evaluated, as has been done in the past for BRDF editing [?]. Such design tools should also correctly convey appearance to the user, and the

appearance navigation tools we derived can potentially be used to discover canonical objects and lighting conditions suitable for this purpose.

Our expanded space of phase functions adds richness, but it is not complete, and there are surely important translucent materials whose appearance cannot be satisfactorily reproduced by it. Identifying extensions to our phase function model that improve completeness is an important direction for research, and a part of this may be exploring new systems for acquisition. Indeed, many existing systems for measuring scattering parameters are predicated on the single-lobe HG model, and given the limitations of single-lobe or single-lobe plus isotropic formulations revealed by our analysis, acquisition systems that fit to single lobes could limit the range of appearance they can capture.

Additionally, our perceptual parameters only control aspects of translucent appearance that are due to variations in phase function, and this is only one of many factors. Exploring other parameters of multiple scattering, including scattering coefficient, albedo, and subtle spectral dependencies, will be crucial for advancing our understanding of translucency. Another important aspect of the appearance of many translucent materials is heterogeneity, which requires further considering spatial variations of all these parameters.

These future research directions will involve analyses of physical parameter spaces that are much larger than the one considered in this paper. Faced by these challenges, methods for combining computational and psychophysics like the one proposed here, are likely to become increasingly important.

APPENDIX

A. SAMPLING FROM THE VON MISES-FISHER DISTRIBUTION

The analytic form (0??) of the probability density function of the vMF distribution allows for efficient importance sampling. Specifically, using inverse transform sampling, sample directions distributed according to a vMF with parameter κ can be generated from the expression

$$\cos \theta = \frac{\log(\xi \exp(\kappa) + (1 - \xi) \exp(-\kappa))}{\kappa}, \quad (13)$$

where ξ is a random variable uniformly distributed in $[0, 1]$.

B. PROCRUSTES ANALYSIS

We provide some discussion of the planar Procrustes analysis techniques that we use in Section 0??. For more details, we refer to Dryden and Mardia [?].

We will use the term *configuration* to refer to a set of landmark points in \mathbb{R}^2 , denoted by

$$\mathcal{E} = \{\mathbf{x}_n \in \mathbb{R}^2, n = 1, \dots, N\}. \quad (14)$$

In the following, all configurations we consider will have the same number of points N , and there will be a correspondence between points in different configurations with the same index n . In the context of Sections 0?? and 0??., each two-dimensional embedding is a configuration, with one embedded point per rendered image for a total of $N = 753$ images; corresponding points across different embeddings represent images of different scenes rendered with the same phase function.

Given two configurations \mathcal{E}_1 and \mathcal{E}_2 , we can align \mathcal{E}_2 with \mathcal{E}_1 by finding a translation, rotation, and uniform scaling transform

such that corresponding points in the two configurations optimally match, under some matching criterion. When the matching criterion is the l_2 distance between points, this corresponds to solving the optimization problem

$$\min_{s \in \mathbb{R}, \mathbf{T} \in SO(2), \mathbf{c} \in \mathbb{R}^2} \sum_{n=1}^N \|\mathbf{x}_{1,n} - (s\mathbf{T}\mathbf{x}_{2,n} + \mathbf{c})\|_2^2, \quad (15)$$

over all scale factors s , rotation matrices \mathbf{T} , and translation vectors \mathbf{c} . This alignment problem is called the *full Procrustes superimposition* of \mathcal{E}_2 onto \mathcal{E}_1 , and has a closed-form solution that we denote by $\mathbf{P}(\mathcal{E}_1, \mathcal{E}_2)$. Note that \mathbf{P} is not symmetric in its two arguments, that is, the transformation for aligning \mathcal{E}_2 onto \mathcal{E}_1 is different from the transformation for aligning \mathcal{E}_1 onto \mathcal{E}_2 .

Given any configuration \mathcal{E} , we can *normalize* it through a uniform scaling and translation such that

$$\frac{1}{N} \sum_{n=1}^N \|\mathbf{x}_n\|_2^2 = 1, \quad \sum_{n=1}^N \mathbf{x}_n = \mathbf{0}. \quad (16)$$

For two configurations \mathcal{E}_1 and \mathcal{E}_2 , one way to compare their similarity is first to normalize them, and then compute the minimum of the objective function of (0??). The square root of this minimum value, which we denote by $d_P(\mathcal{E}_1, \mathcal{E}_2)$, is called the *full Procrustes distance* between \mathcal{E}_1 and \mathcal{E}_2 . Due to the normalization step, this is a scale-independent measure of shape similarity. Furthermore, $d_P(\mathcal{E}_1, \mathcal{E}_2)$ is symmetric in its two arguments, unlike the corresponding full Procrustes superimpositions transformations

Given a set of configurations $\{\mathcal{E}_1, \dots, \mathcal{E}_K\}$, we can define their *full Procrustes mean* as the solution of the optimization problem

$$\min_{\mathcal{E}} \frac{1}{K} \sum_{k=1}^K d_P^2(\mathcal{E}, \mathcal{E}_k). \quad (17)$$

This problem also has a closed-form solution. The square root of the minimum of the objective function of (0??) is called the *root-mean-square of full Procrustes distances* for this set of embeddings. The full Procrustes mean configuration and root-mean-square of full Procrustes distances are measures of central tendency and variability for sets of shapes analogous to sample arithmetic mean and standard deviation for sets of real numbers. The procedure of solving the optimization problem (0??) is called *generalized Procrustes analysis*.

C. LEARNING FROM PAIRED COMPARISONS

The task of kernel learning from paired comparisons has received significant recent attention, both in graphics [?] and other areas [?], and a number of algorithms have been proposed. Here, following [?], we learn \mathbf{K} by solving the problem

$$\min_{\mathbf{K} \succeq \mathbf{0}} \lambda \|\mathbf{K}\|_* + \frac{1}{S} \sum_{s=1}^S L(d_{\mathbf{K}}(i_s, k_s) - d_{\mathbf{K}}(i_s, j_s) + b), \quad (18)$$

where summation is over the set $\mathcal{S} = \{(i_s, j_s, k_s), s = 1 \dots S\}$ of triplets for which a human observer has judged that $d(i_s, k_s) < d(i_s, j_s)$. In this equation, $\|\cdot\|_*$ is the *nuclear norm* of a matrix, equal to the sum of its singular values; $L(\cdot)$ is a loss function, described below; $d_{\mathbf{K}}(i_s, k_s)$ and $d_{\mathbf{K}}(i_s, j_s)$ are shorthand for entries of \mathbf{K} via (0??); and b is a margin parameter, set equal to 1, used to resolve the scale ambiguity in relative comparisons.

The objective of (0??) provides a balance between an empirical loss term penalizing violations of the constraints in \mathcal{S} , and a regularization term to prevent overfitting the data. The use of the nuclear

Algorithm 1 Kernel learning from paired comparisons.

Input: $\mathcal{S} = \{(i_s, j_s, k_s), s = 1 \dots S\}, \lambda$.

$\mathbf{K}_0 = \mathbf{0}; \mathbf{K}_{-1} = \mathbf{0}; t_0 = 1; t_{-1} = 1; \lambda_0 = \delta\lambda$.

while not converged **do**

{Update low-rank kernel Gramian.}

$\mathbf{Y}_k = \mathbf{K}_k + \frac{t_{k-1}-1}{t_k} (\mathbf{K}_k - \mathbf{K}_{k-1})$.

$\mathbf{G}_k = \mathbf{Y}_k - \frac{1}{\rho_k} \frac{1}{S} \sum_{s=1}^S \nabla_{\mathbf{Y}_k} L(d_{\mathbf{Y}_k}(i_s, k_s) - d_{\mathbf{Y}_k}(i_s, j_s) + b)$.

$(\mathbf{Q}, \mathbf{e}) = \text{eig}(\mathbf{G}_k)$.

$\mathbf{K}_{k+1} = \mathbf{Q} \text{diag}\left(\max\left(e - \frac{\lambda_k}{\rho_k}, 0\right)\right) \mathbf{Q}^T$.

{Prepare for next iteration.}

$t_{k+1} = \frac{1 + \sqrt{4t_k^2 + 1}}{2}$.

$\lambda_{k+1} = \max(\eta\lambda_k, \lambda)$.

$k = k + 1$.

end while

Output: $\mathbf{K} = \mathbf{K}_k$.

norm for regularization is to encourage low-rank solutions \mathbf{K} [?], and also due to its provable generalization and robustness to noise properties [?]. In terms of the loss function, a natural choice would be to use the *hinge loss*,

$$L(x) = \max(x, 0). \quad (19)$$

In this case, (0??) is the Lagrangian formulation of the optimization problem of (0??), with the slack variables and linear constraints of the latter replaced by the empirical loss term. In our implementation, we use a modification proposed by Rosset et al. [?] called *Huberized squared hinge loss*,

$$L(x) = \begin{cases} -4 + 4x, & x > 2, \\ [\max(x, 0)]^2, & \text{otherwise.} \end{cases} \quad (20)$$

This loss function shares the attractive properties of the hinge loss—convexity, linear penalization of violated constraints, and zero loss for satisfied constraints; but is also smooth. Together with the Lagrangian formulation, these modifications to the optimization problem proposed by Wills et al. [?] provide attractive scalability and speed, by avoiding the need for semidefinite programming solvers. In our implementation, we solve (0??) using the accelerated proximal gradient algorithm of [?], described in Algorithm 0??. We set λ through cross-validation as in [?], and ρ and δ (parameters controlling the speed of convergence of the algorithm) empirically.