EG UK Theory and Practice of Computer Graphics (2007)
Ik Soo Lim, David Duce (Editors)

Perception-Based Lighting-by-Example

Hai Nam HA & Patrick Olivier

Newcastle University, UK

Abstract

In computer graphics, easy-to-use tools for configuring lighting for 3D scenes are required by users. Following a perception-based lighting design framework, which models image quality using cognitively inspired objective functions, we present a new approach to lighting design which both: (1) allows the declarative specification of lighting; and (2) uses target scenes and images to facilitate intuitive and natural interactive control of the scene lighting. The LightOPex system enables users to select the desired lighting for a scene using exemplars in the form of 3D scenes and 2D images and uses the perceptual properties of these exemplars as target values in an initial optimization step.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism.

1. Introduction

The problem of finding optimal lighting parameters - positions, directions, colors, and intensities of light sources - to achieve the visual properties required of a 3D scene is referred to as the lighting design process. The means by which the required visual properties are specified, and the degree of manual intervention allowed, are highly application dependent. For example, 3D authoring tools allow full interactive control in the interative design of lighting, whilst in the case of visualization it is necessary that lighting design is a fully automatic process in which the perceptual qualities of real-time generated views of 3D objects are maximized.

Approaches to automated lighting design that assume the existence an ideal configuration of the lighting (i.e. *ideal lighting* approaches) in general aim to somehow optimize the configuration of the lights in order to reveal visual properties of objects in the final 2D images. The visual properties of objects can be characterized according to the different kinds of information that 2D images convey, such as depth information and information about the shape of objects in the scene. Such information is obtained by the viewers from 2D properties of the scene such as the shading gradient, object and feature edges, regions (and degrees) of contrast, and the value and distribution of luminance.

Ideal lighting approaches try to maximize the visual properties of objects in a scene by optimizing an objective function which characterizes such visual properties [HO06][SL01]. In practice, the notion of *ideal lighting* is only meaningful with respect to a small range of real-world graphics applications. Such applications include automatic lighting of 3D visualizations where the number, position and orientation of 3D glyphs or other objects cannot be predicted in advance (and are typically not textured). Although in such visualization applications the color, spatial properties and other physical characteristics of the elements capture all the information to be represented (and the role of lighting is to reveal these), in many domains subtle changes in lighting are used to convey mood, emotion, and factors other than the raw geometric and visual properties of the scene elements. Consequently, the goal of general research program into lighting design research is not an accurate (and empirically verified) objective function for ideally illuminated scenes, but a framework for the specification of lighting using example 3D scenes and even photographs - tools to allow artists to interactively modify scene lighting through inverse design.

Ideal lighting approaches presume the ability to model target scenes in the form of a perceptually meaningful objective function and to optimize source scenes using these objectives. We refer to such a process as *lighting-by-example*. Our approach to *lighting-by-example* is based on a perception-based lighting framework that we initially developed within our ideal lighting framework and with which we can optimize lighting parameters with respect to a set of target values of different components in an objective function. We propose a *lighting-by-example* approach as a result of our recognition of that perceptual optimality is rarely an appropriate or meaningful notion when 3D artists are engaged in lighting design.

Viewers are highly sensitive to the emotional tone of an image arising from its lighting, although non-expert viewers will have little or no insight into the configuration of lights with which such effects are created. Indeed, in photographic and film production the subtleties of scene lighting are often the results of highly artificial configurations of lights on a studio set (and post production editing). In short, we know what we want when we see it, but have little idea of how to reproduce it. This is the observation on which we base the *lighting-by-example* approach - that lighting is best configured for 3D scenes on the basis of existing exemplars of images and not through direct manipulation of lighting types, positions and luminance.

2. Example-based approaches

There has been a number of approaches which can be considered either examples of, or strongly related to, the example-based approach to lighting design. Schoeneman et al [SD*93] addressed lighting design as an inverse problem. Users are able to configure a set of desired properties that are expected to appear in the final image and the system tries to find out a solution whose properties are closest the set of desired properties. Directly painting on the surfaces of the rendered scene causes a change in surface radiance functions, and these after-painted surface radiance functions are used as target values in the optimization process that follows. Painted surfaces in the rendered image are given more weight, which biases the optimization towards solutions with properties that best match the painted surfaces. In this approach, painted surfaces can be considered as examples affecting the target radiance surface functions, though in this approach Schoeneman et al only addressed the problem for finding matching light intensities and colors for fixed light positions. Design Galleries [MA*97] adopted an approach that was significantly different from that of inverse lighting design through the manipulation of object properties (such as shadows). Here Marks et al's goal was the design of an interactive system to allow a user to interactively reduce the design space for light configurations through the use of a mapping function between an input vector containing light position, light type, and light direction and an output vector containing a set of values that summarizes the perceptual qualities of the final image. During the optimization step lights are moved from one predefined position to another. At each position a light type is selected from a set of light types; and a corresponding image is generated. Final images are then arranged in clusters on the basis of the perceptual distance between images.

Design Galleries [MA*97] can be considered to be in the spirit of an example-based approach despite the fact that there is no specific target used as the basis for an objective function (to be optimized). Through its generation of a wide range of clustered images as examples, Design Galleries presents sets of exemplars for users to perform selection on as part of an render-selection loop. Thus, there is no information about what effects users want to have in the final images, but the user has the opportunity to select good candidates. Image-based lighting can also be considered as another form of example-based approach. Supan and Stuppacher [SS06] presented an approach to lighting augmented environments in which virtual objects fit seamlessly into a real environment. A key challenge in such applications is how to consistently co-ordinate the lighting between virtual objects and a real environment.

Image-based lighting approaches attempt to capture lighting information from a real environment and use it somehow to light virtual objects such that consistency in the lighting of the virtual object and the real-world objects can be obtained. At the heart of this approach is an environment map which represents lighting information of the real environment. To obtain the environment map, a mirrored sphere is set up in real environment such that surrounding scene can be seen as a reflection on the sphere, and the image of the mirrored sphere is captured with a camera. Specular and diffuse sphere maps can be created using a radial blur technique in which image captured from mirrored sphere is mapped to a virtual sphere of unit radius. Shadows are also addressed in this approach and to calculate the shadows cast by virtual objects, light positions in the real environment are identified. Supan and Stuppacher use an intensity distribution-based technique for estimating the light positions from the environment map was used. A high dynamic range image derived by combining several images captured under different lighting exposures was used to enhance accuracy of the light position estimation. The finally rendering process requires only small modifications in the calculation of the lighting (obtained by looking up values in the environment map and using Lambert's Law). This approach can be considered as a class of *lighting-by-example*, that is, a lighting optimization problem in which lighting parameters for virtual environment are optimized on the basis of lighting information captured from real environment.

3. Perception-based approaches

Our proposal for *lighting-by-example* is based on our core perception-based lighting design framework. This in turn is an extension to the approach proposed by Shacked and Lischinski [SL01]. In their perception-based lighting design scheme the position and intensity of light sources (specular and diffuse components of a local illumination model) are optimized using an evaluation function that characterizes separate aspects of low-level processing in the segmentation and recognition of objects. At the heart of this approach is an objective function that is a linear combination of five distinct image properties:

- F_{edge} edge distinctness;
- F_{mean} mean brightness;
- F_{grad} mean shading gradient;
- F_{var} intensity range;
- F_{hist} image intensity distribution.

Thus the objective function $F(\theta_k, \varphi_k, I_{dk}, I_{sk}, R_k)$ is: $F(\theta_k, \varphi_k, I_{dk}, I_{sk}, R_k) = w_e F_{edge} + w_m F_{mean} + w_g F grad + w_v F_{var} + w_h F_{hist}$

Where θ_k is the elevation angle of k^{th} light; φ_k is the azimuth angle of k^{th} light; I_{dk} is the diffuse intensity of k^{th} light, I_{sk} is the specular intensity of k^{th} light; R_k is the distance of k^{th} light (fixed for directional lights); k = 1, 2,... K is the identifier of a light where K is the number of lights; and w_e, w_m, w_g, w_v, and w_h are the weights for different objective function components.

A sixth component of the image quality function originally proposed by Shacked & Lischinski biases the optimization of a key light to a particular elevation and orientation above and in front of the object (relative to the viewpoint). Although this is a standard practice in photography and might be explained in terms of evolutionary psychology [Gro94] - that our perceptual system evolved for scenes lit by the sun or moon - we instead simply constrain the light position to a quarte sphere, in front of, and above, the centre of the scene.

In their original proposal Shacked Lischinski formulate the optimization problem such that the lower values of $F(\theta_k, \varphi_k, I_{dk}, I_{sk}, R_k)$ correspond to configurations with desired visual characteristics and a greedy gradient descent minimization algorithm is utilized in the discovery of appropriate lighting configurations.

We have extended their approach by: (1) analyzing the shape of the objective function and applying a more powerful optimization technique; and (2) by adding new perceptually motivated components to the objective function. We have implemented both genetic algorithm and simulated annealing in addition to the steepest decent technique originally deployed by Shacked & Lischinski [SL01]. We have also incorporated a number of features that were not apparent in previous perception-based lighting design systems:

(a) *contrast*: contrast between different surfaces of an object is important in conveying information about shape and depth of objects.

(b) *back-lighting*: a well established feature of cinematic and photographic practice is to back-light the subject, this has the effect of maximizing the gradient in the intensity between the subject and the background.

(c) *Perceptually uniform color spaces*: standard approaches in lighting design implement metrics over standard RGB (or equivalent) color spaces, despite the fact that such spaces are highly non-uniform with respect to human judgments of color.

For a complete description of our extensions and an discussion of the evaluation of different optimization schemes see [HO06].

4. Lighting-by-Example

The aim of the *lighting-by-example* approach to lighting design is to provide users with a means to express their desired characteristics through the use of pre-lit exemplars. A user selects target image from a set of examples provided by the lighting system in the form of 3D scenes or 2D images. Properties of the selected example are used in an initial optimization step. The optimization process seeks to discover a configuration of the lighting parameters such that the components of the objective function extracted from rendered scene have values that are close to that of the target example.

4.1. Target property extraction

The first step of *lighting-by-example* pipeline is the computation of target values of the six objective function properties: edge distinctness; mean brightness; mean shading gradient; intensity range; image intensity distribution; and contrast. The key data structure for this process is the pixel type map which records the type of pixels in a rendering. A pixel type map for renderings of 3D scenes is extracted by applying an edge detection operator to the depth buffer [ST90]. Sobel and Laplace edge detection operators are combined to enhance the accuracy of edge detection. Points on surfaces of objects in the scene are detected by an algorithm using color-coded polygon identifiers [HO106]. The resulting pixel type map contains 3 types of pixels: EDGE: pixels on the edge of an object; SURFACE: pixels on the surface of an object; and BACKGROUND: pixels corresponding to objects and scene elements that are considered to be in the background. The pixel type map is used in the calculation of different components in the objective function.

Edges

According to psychological research on the human visual system [Gro04][SJ90], edges convey significant information about shape of objects. Therefore, under the ideal lighting view they should be clearly apparent in the image. The edge component (F_{edge}) measures the prominence of edges in an image. The prominence of edges is estimated by computing the ratio of the number of pixels detected by a pixel-based edge detection operator applied to the rendered image, and the total number of actual edge pixels computed for the pixel type map.

Given a 2D rendered image, edges are detected by combining first and second order derivatives of image intensity function I(x,y). The first derivative of I(x,y), the gradient, detects local changes in the luminance at a pixel p(x,y). The second order derivative aims to address discontinuities of the first derivative. This approach to edge detection is widely exploited in computer vision [ST90]. Specific implementation details can be referred in [HO106]

Shading gradient

Perceptual psychologists consider shading as one important depth cue [Gro04][SJ90]. In our lighting design framework, the shading gradient component of the objective function serves to enhance perception of depth. The shading gradient at each pixel p(x,y) is calculated and the final shading gradient is derived by averaging shading gradients over the whole image. This mean value for the target image is used as the shading gradient component in the optimization of the scene to be lit.

$$T_{grad} = \sqrt{\frac{1}{N_s} \sum_{p(i,j) \in S} |\nabla_{i,j}|^2}. \tag{1}$$

T_{grad}: the target shading gradient component extracted from an example;

$p(i,j)$: a pixel in the ith row and jth column in an image;

$\nabla_{i,j}$: the shading gradient of the image function I(x,y) at a pixel p(i,j);

S: a set of surface pixels derived from the pixel type map;

N_s: the number of surface pixels.

Luminance

The mean luminance component has a global impact on the luminance of the rendered image. The optimization process attempts to set the mean luminance of the rendered image close to the target value. Thus the mean luminance of the example is used as the target value for mean luminance component in the optimization process. The mean luminance of an example is calculated as follows:

$$T_{mean} = \frac{1}{N_s} \sum_{p(i,j) \in S} I(i,j) \tag{2}$$

T_{mean}: the target mean component extracted from an example scene;

$p(i,j)$: a pixel in the ith row and j_{th} column in an image;

$I(i,j)$: the value of the image function at a pixel p(i,j);

S: a set of surface pixels derived from the pixel type map; N_s: the number of surface pixels.

Luminance variance

The human visual system is particularly sensitive to a narrow range of luminance around a certain average luminance value. The luminance variance component aims to constrain the overall brightness of the rendered image to an appropriate luminance range. LIGHTOPEX computes this variance for the exemplar and aims to set the value for the scene to be lit accordingly. The variance component is calculated as follows:

$$T_{var} = \sqrt{\frac{1}{N_s} \sum_{p(i,j) \in S} (I(i,j) - T_{mean})^2} \tag{3}$$

T_{var}: target luminance variance component extracted from an example;

$p(i,j)$: a pixel in the i^{th} row and j^{th} column in an image;

$I(i,j)$: the value of the image function at a pixel p(i,j);

S: a set of surface pixels derived from the pixel type map;

N_s: the number of surface pixels;

T_{mean}: the target mean component extracted from an example scene.

Brightness histogram

A brightness histogram is used to represent the distribution of brightness value over pixels. The histogram comprises 256 bins, and the value of each bin represents the number of pixels within a certain brightness range. The histogram is normalized by dividing value of every bin by the total number of pixels used in the histogram.

$$T^k_{k \in [0,255] hist} = \frac{N_k}{N_t} \tag{4}$$

T^k_{hist}: the target histogram value of bin k^{th} extracted from an example.

N_k: the number of pixels at brightness level k

N_t: the total number of pixels used for calculating the histogram. If the example is a 3D model, this is the total number of edge and surface pixels. If the example is a 2D image, this is the total number of pixels in the 2D image.

Contrast component

Empirical studies of visual cognition have also demonstrated that object perception depends on both the absolute amount of luminance and the difference of the object's luminance from its background [Gro04]. We extend this notion through the provision of a means of evaluating in luminance between adjacent parts of an object and incorporating this in our objective function as follows: Contrast between two parts of an object is given by:

$$C_{ij} = \frac{Y_i - Y_j}{Y_j} \tag{5}$$

C_{ij}: the contrast between part i and part j;

Y_i: the mean luminance of part i.

The mean luminance of a part is calculated as follows:

$$Y_i = \frac{1}{N_i} \sum_{p(x,y) \in P_i} I(x,y) \tag{6}$$

P_i: part i of an object;

N_i: the number of pixels in the image corresponding to part i;

$p(x,y)$: a pixel in the x^{th} row and y^{th} column in an image;

$I(x,y)$: Value of image function at a pixel p(x,y) Edges in the pixel type map correspond to boundaries between parts of an object. With this assumption, we developed an algorithm for calculating the contrast between adjacent parts of a 3D object using the pixel type map [HO06].

4.2. 3D vs 2D examples

A 3D example is created by rendering a sample model with lighting parameters previously specified - where the model has been created either for a separate purpose or for the specific purpose of being used in as an exemplar in LIGHTOPEX. Separated components of the objective function are extracted and saved in a configuration file. The rendered image is also saved to a bitmap file. In practice the 3D model can be rendered multiple times with different configurations of lighting parameters in order to create different examples of the same model (as in *Design Galleries*).

2D examples are actually 2D images created by rendering 3D models or photographic images of real scenes. Other than the edge and contrast components, which are not used in 2D-exemplars, the remaining target components are extracted in the same way for both 3D and 2D examples. Target components are calculated over all pixels of the 2D image example. For 3D examples, the pixel type map is used to decide which pixels should be taken in to consideration in the calculation of the target components. Note that for 2D image examples, S is computed using a set of pixels of the whole 2D image excluding only the set of edge pixels which can be derived by applying a standard edge detection technique to the example image.

4.3. Shadow processing using a shadow-map

Shadows have not been considered in most previous interactive lighting design approaches, probably because of the complexity of the implementation and the rendering times required. In practice, shadows are one of the key ecological features of real-world vision that needs to be incorporated in the lighting design process. However, shadows are problematic as they actually contribute regions of low luminance to the rendered scenes. For 2D image-based optimization techniques in lighting design, shadows tend to result in what would traditionally be considered as 'non-ideal' brightness levels in rendered images. Specifically, shadows significantly impact on the edge (F_{edge}), mean luminance (F_{mean}) and luminance variance (F_{var}) components of the objective function used in our approach.

With respect to the mean luminance component, shadows give rise to low brightness regions in the image that result on lowered values of the mean luminance component. For the same reason shadows also tend to drive the luminance variance into a lower range of intensity since the target value of intensity variance constrains the intensity of the rendered image to a predefined width of intensity range. During the optimization process lighting parameters are thus influenced such that of the mean brightness of the rendered image increases to reach the target value component during optimization process. This leads to a defect whereby object surfaces are overlit, and the final image of the optimization process is too bright. Edges are also significantly weakened in shadow regions and need to be enhanced using particular techniques [Bra04].

We propose a solution to these problems whereby for the brightness and intensity variance components (F_{mean} and F_{var}) in the objective function, the pixels in shadow regions are not used in the calculation of the mean brightness and intensity variance components. For the edge component (F_{edge}), edge pixels are weighted differently for two types of edge pixels, that is, edge pixels are in and out of shadow regions. The edge pixels in shadow regions are given a weight whose value is higher than that of edge pixels not in shadow regions. As a consequence edges pixels in shadow regions have significantly more influence on the objective function.

A data structure called shadow map is used in the implementation of approach. The shadow map is actually an array whose elements show which pixels in shadow and which are not in shadow. The shadow map is derived by casting shadows into the scene with a depth buffer-based technique.

5. Results

Figures 1 and 2 show examples of results of the approach for different classes of exemplars (2D, 3D and photograph). The recognizable properties of exemplars such as overall luminance and shading gradient are captured to optimized images.

Figure 1 shows the results in which target 2D images are used, and figure 2 shows the results in which target 3D images and target photo are used . Images in the first row are targets, and those in the other rows are corresponding results.

Results in columns 1 and 2 of figure 1 and figure 2 have luminances that are quite similar to those of targets. The ranges of brightness of results in columns 1 and 2 of figure 1 and figure 2 are recognizably equivalent to those of targets, and that means the luminance variance and histogram components have significant impacts on optimization process. Looking at results in columns 1 and 2 of figure 1, the average luminances of those results are pretty the same but the shading effects are significantly different due to the impacts of shading gradient component on optimization process.

6. Conclusion

LIGHTOPEX implements a novel *lighting-by-example* approach to lighting design. A perception-based objective function was used at the core of the approach to capture the properties over an image function. Extended features as well as shadow processing have been incorporated into the processing pipeline. The potential application of research to the design of lighting includes the specification of lighting in which photographs are used as targets, though future tools should incorporate controls that help graphics designers interactively modifying scene lighting through inverse design. Such approaches hypothesize that target scenes can be modeled in the form of a per ceptually meaningful objective function, and the lighting of source scenes can be optimized using these objectives. One of the drawbacks of this system is that the spatial frequency information in the target images is not captured to the final image other than information about edge prominence, average luminance, shading gradient, histogram, and contrast. In other words, spatial distribution of pixels at certain luminance level is not captured. An extension to this system that enables the *lighting-by-example* to capture spatial frequency information is the topic of ongoing work.

References

[Bra04] BRAJOVICH V.: Brightness Perception, Dynamic Range and Noise: A Unifying Model for Adaptive Image Sensors. In *Proc. IEEE (CVPR '04)*, Vol. 2, 189-196, 2004.

[Gro94] Gross, M. *Visual Computing. Springer-Verlag*, 1994.

[HO106] HA, H.N., OLIVIER, P.: Peception-based lighting design. In *Pro. TPCG '06*, 2006.

[HO06] HA, H.N., OLIVIER, P.:Explorations in Declarative Lighting Design. *6th International Sympo-sium on Smart Graphics*, 160-171, 2006.

[MA*97] MARKS, J., ANDALMAN, B., BEARDSLEY, P. A., FREEMAN, W., GIBSON, S., HODGINS, J., KANG, T., MIRTICH, B., PFISTER, H., RUML, W., RYALLl, K., SEIMS, J.,SHIEBER, S. Design galleries: a general approach to setting parameters for computer graphics and animation. In *Proc. SIGGRAPH '97*,389-400, 1997.

[SD*93] SCHOENEMAN, C., DORSEY, J., SMITS, B., ARVO, J., GREENBURG., D.: Painting with light. In: *Proc. SIGGRAPH '93*, 143-146, 1993.

[SJ90] SPILLMANN, L., JOHN, S., W.: *Visual Perception the Neurophisiological Foundations*, Aca-demic Press Inc., 1990.

[SL01] SHACKED, R., LISCHINSKI, D.: Automatic Lighting Design using a perceptual quality met-ric. In *Proc. Eurographics '01*, vol. 20, no.3, 2001.

[SS06] STUPPACHER, I., SUPAN, P.: Image Based Lighting in Augmented Reality. In *Proc.CESCG*, 2006.

[ST90] SAITO, T., TAKAHASHI, T.: Comprehensible rendering of 3-D shapes. In Proc. *SIGGRAPH'90*, 197-206, 1990.

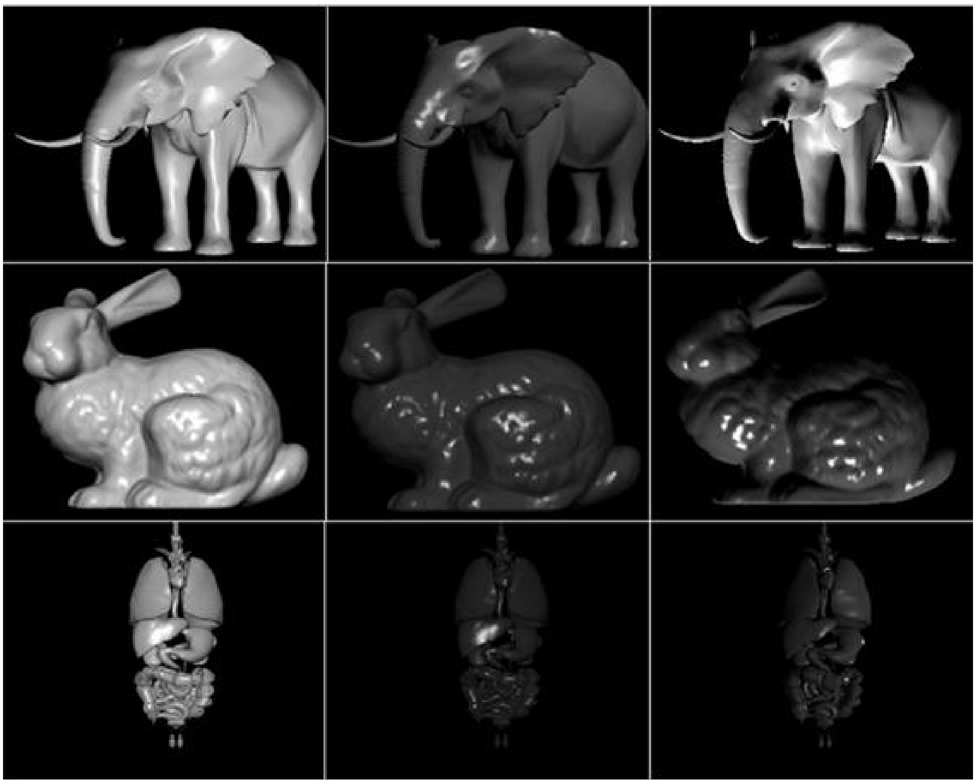

Figure 1: *Lighting by example with 2D targets. Images in the first row are target 2D images. Images in the other rows are results derived by optimizing different 3D scenes with corresponding targets in the first rows.*

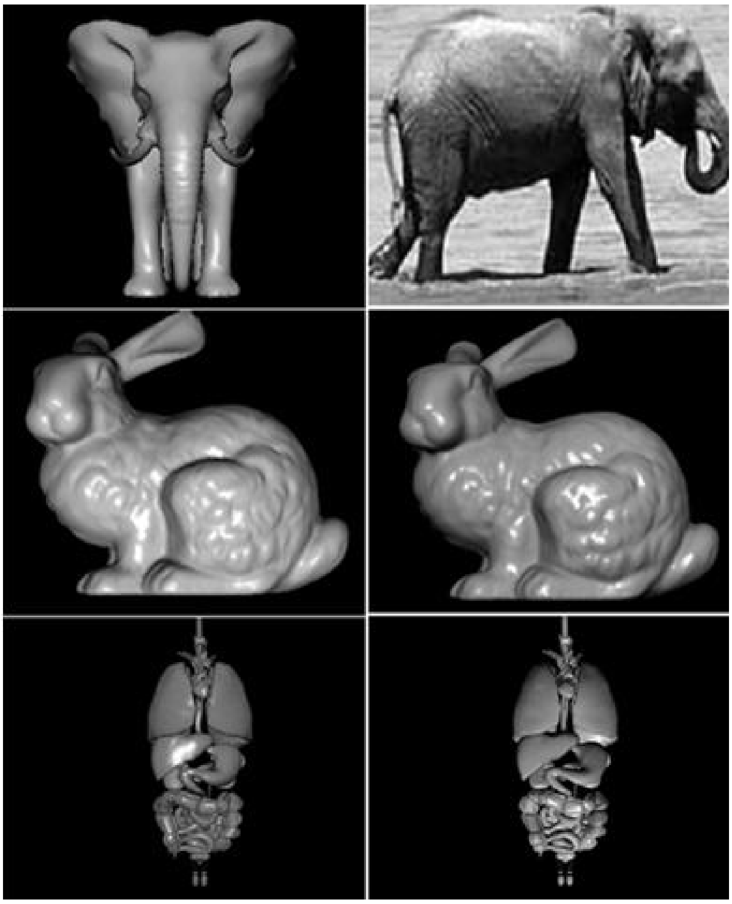

Figure 2: *Lighting by example with 3D targets and photos. Images in the first row are target 3D image and target photo. Images in the other rows are results derived by optimizing different 3D scenes with corresponding targets in the first rows.*