A Perceptual Quality Scale for Image-based Relighting of Glossy Surfaces

Olivier Dumont *Vincent Masselus*
Peter Zaenen *Johan Wagemans* *Philip Dutré*

Report CW417, June 2005

Katholieke Universiteit Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

A Perceptual Quality Scale for Image-based Relighting of Glossy Surfaces

Olivier Dumont *Vincent Masselus*
Peter Zaenen *Johan Wagemans* *Philip Dutré*

Report CW417, June 2005

Department of Computer Science, K.U.Leuven

Abstract

We present a method to obtain a perceptual quality scale, relating the number of basis images used in the image-based relighting of scenes containing glossy surfaces to their perceived quality. Our method is based on advanced psychological experiments, requiring observers to make judgments on different relit images. These relit images are constructed with cumulative subsets of basis images. From these judgments, a perceptual quality scale is derived.

We demonstrate this method for three different kinds of glossy materials. From the obtained perceptual quality scales, we can derive a threshold for the number of basis images required to relight the different types of glossy materials to perceptual perfection. Additionally, we determine an approximate threshold function for the number of required basis images to relight to perceptual perfection as a function of glossiness.

Keywords : perception, relighting, light interaction, quality scale.
CR Subject Classification : I.3.7, I.4.8

A Perceptual Quality Scale for Image-based Relighting of Glossy Surfaces

Olivier Dumont† Vincent Masselus† Peter Zaenen‡ Johan Wagemans‡ Philip Dutré†

† Department of Computer Science
Katholieke Universiteit Leuven

‡Department of Psychology
Katholieke Universiteit Leuven

Abstract

We present a method to obtain a perceptual quality scale, relating the number of basis images used in the image-based relighting of scenes containing glossy surfaces to their perceived quality. Our method is based on advanced psychological experiments, requiring observers to make judgments on different relit images. These relit images are constructed with cumulative subsets of basis images. From these judgments, a perceptual quality scale is derived. We demonstrate this method for three different kinds of glossy materials. From the obtained perceptual quality scales, we can derive a threshold for the number of basis images required to relight the different types of glossy materials to perceptual perfection. Additionally, we determine an approximate threshold function for the number of required basis images to relight to perceptual perfection as a function of glossiness.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three dimensional Graphics and Realism Shading and shadowing I.4.1 [Image processing]: Digitization and Image Capture

1. Introduction

Over the recent years, relighting of real objects has been researched to large extent. Published work in this research field allows to faithfully visualize objects, ranging from toys and human faces to buildings and cities, with novel incident illumination. In particular, the image-based relighting techniques are very appealing, since they do not require any geometrical data of the object nor are they restricted to an analytical surface reflection model. Instead, these techniques only require a special set of images of the object, called the basis images. These basis images are photographs of the object, captured from a single fixed viewpoint. For each photograph, the object is illuminated from a different direction. Based on these basis images, the object can be relit with any novel incident illumination [DHT*00]. This is achieved by computing a weighted sum of the basis images. The weights depend on the illumination directions in the used basis images and the novel incident illumination.

Most of the published papers on relighting propose more practical or faster data acquisition methods [HCD01, MDA02], more intelligent techniques to process the basis images [MPDW04], or more efficient storage of all information [WHON97, MGW01]. However, in most applications, such as the movie industry or augmented reality, the relit images do not need to be physically correct. They only need to convince the viewer that the object is illuminated with the novel incident illumination. Furthermore, the number of required illuminant directions, and thus the number of basis images, is seldom discussed.

In this work, we investigate the perceptual effect of the number of basis images on the resulting relit images. More precisely, we research the relation between the number of basis images for relighting different kinds of glossy materials and the perceived quality of the relit images. Based on established psychoperceptual experiments, we present a perceptual quality scale for the number of basis images for relighting surfaces with different gloss. This quality scale denotes how an observer perceives the quality of a relit image when the number of basis images changes in correspondence with other relit images.

A perceptual quality scale has several benefits: if a relit image, computed with N basis images, is perceived similar

† e-mail:{olivierd,vincentm,phil}@cs.kuleuven.ac.be
‡ e-mail:{peter.zaenen,johan.wagemans}@psy.kuleuven.ac.be

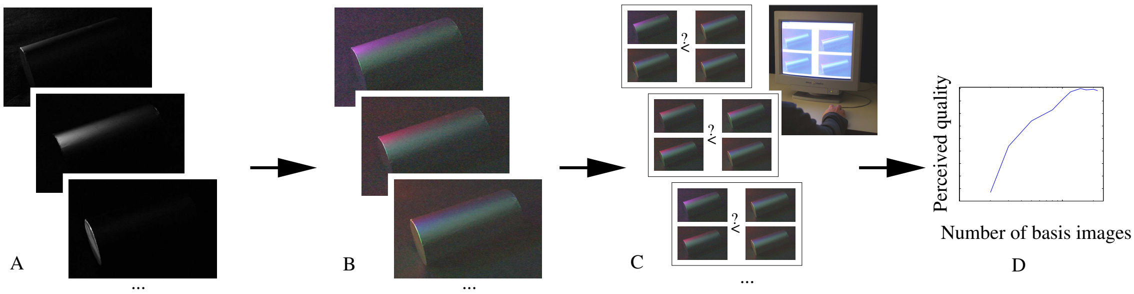


Figure 1: *For a cylinder, a set of basis images are captured (A). Based on a different number of these basis images, different relit images can be renderer, resulting in a series of relit images (B). Different quadruples of the relit images can be selected to be displayed on a CRT (C). From the decisions of observers, a quality scale (D) can be derived, relating the number of basis images for relighting to the perceived quality.*

to a relit image, based on only a small fraction of these N basis images, then only that fraction should be acquired, since the other basis images do not have any added value. Therefore, such a scale would allow to derive a threshold on the number of basis images important for the perceptual quality. Additionally, with this scale one can also predict how many light sources must be sampled to compute the relit images of a user defined perceived quality.

We coated three cylinders with a different type of glossy material. For each of these coated cylinders, we captured a large set of basis images. Then, we used different amounts of these basis images to render relit images, all with the same novel illumination. This resulted in a series of relit images for each cylinder, where only the amount of basis images used to compute the relit images differed. Finally, observers, unaware of our goal, were presented with the relit images of such a series, in a particular order. Based on their opinions, a perceptual quality scale was obtained, using the *Maximum Likelihood Difference Scaling* (MLDS) technique [MY03].

2. Related work

Many studies have been conducted on image-based relighting, but to our knowledge, only the work of Lin et al. [LWS01] takes into account the sampling density of the light sources, and thus the number of basis images. In their work, sampling thresholds for the light source positions are calculated, as a function of the parameters of an analytic BRDF model for the surface and a given tolerable error ε. Therefore, these thresholds for the number of basis functions ensure physical correct relit images up to an error ε. Although of importance, it is hard to find these specific parameters for a real surface. Additionally, these ε-dependent thresholds are to ensure physical correctness and are only a loose upper bound for a perceptual threshold.

Several attempts have already been made to represent the perception of gloss. Pellacini et al. [PFG00] propose a psychophysical model to describe glossy surfaces. This model is based on a two-dimensional space to represent gloss of a surface. These two dimensions are labeled *Contrast* and *Distinctness of Image*. Earlier, Billmeyer and O'Donnell [BO87] found gloss to be a two-dimensional feature as well. However, they found it sufficient to represent gloss in a one-dimensional space.

The approach of Obein et al. [OKV04] is closely related to ours. In the spirit of [BO87], they quantify the perception of glossy surfaces in a one-dimensional space, and build a perceptual difference scale of ten different standard glossy surfaces. In order to achieve this, observers are presented series of two pairs of samples of these materials, for which they have to mark which one of the two pairs is perceived as more different than the other pair. From these tests, a difference scale is obtained using the technique of [MY03].

In this paper, we will follow a similar approach as in [OKV04]. However, we derive a difference scale for the quality of relit surfaces of different gloss in function of the number of basis images used.

3. Overview

First, we select the kinds of glossy materials for which we want a quality scale. For each of these materials, we coat a cylinder with that material and capture a set of basis images (figure 1.A). Relighting with novel illumination but using different amounts of basis images, results in different relit images (figure 1.B). The acquisition of basis images and rendering of different relit images is discussed in section 4.

For each material, based on the series of relit images, we perform experiments in which quadruples of relit images are classified by observers (figure 1.C). From these observations, we derive a quality scale. The details of these experiments are discussed in section 5. In section 6, we take a look at the influence of different novel illumination, scene complexity, as well as texture. Based on the obtained quality scales of the different kinds of materials, guidelines for relighting glossy materials are derived (section 7).

4. Choices and Generation of Data for Experiments

In order to do our experiments, we need to a) choose appropriate materials for which we want a quality scale, and construct scenes with these materials, b) capture a large set of basis images for each of these scenes, and c) generate relit images, based on different numbers of basis images, for the experiments. In the next subsections, we discuss these three issues.

4.1. Choice of Scenes

In the work of [OKV04], the gloss of materials is categorized as either *Matte*, *Intermediate*, or *High Gloss*. This categorization is based on the size of the highlight induced by these materials: a matte material will have a highlight covering the complete surface, an intermediate material will have a clearly distinct highlight and a high glossy material will have a highlight in which the shape of the light source is clearly recognizable. Based on this subdivision of gloss, we choose three materials representing each of these categories: white glossy paper for matte material, glossy aluminum foil for intermediate material and nearly specular metal foil for the high gloss material.

In order to relight these materials, we choose to coat cylinders in each of the chosen materials. The choice for cylinders is based on the fact that these have a varying surface normal with relation to the camera while it is still an easy task to coat a cylindrical surface, as opposed to a sphere.

Figure 2: *The three coated cylinders.*

4.2. Acquisition of Basis Images

For each cylinder, we need to record a set of basis images. This set needs to be large because we are interested in the effect of the number of basis images with relation to the resulting relit images. However, we do not need to sample the complete hemisphere for the light directions, only a part will suffice. Afterwards, our results can be extrapolated for light directions over the complete hemisphere. Therefore, for each type of material, we will record an abundant amount of images over a small section of the light source position domain: in our case, 640 light directions are uniformly sampled on only a part of the hemisphere above the cylinder. Out of practical reasons and repeatability, the data acquisition of the 640 basis images is equal for all materials.

We want to create different relit images based on subsets of basis images. These subsets need to be *cumulative*. If the subsets would not be cumulative, we could not fairly compare relit images based on different number of basis images. With each basis image corresponds a light source position. To find a correct quality scale for the number of basis images for relighting, we want *the light source positions for each subset to be optimally sampled* over the light source position domain. Again, if some subsets would not be optimally sampled, the comparison would not be fair. Therefore, we want:

$$\forall \quad \text{subsets} \quad S_i, S_j : \begin{cases} S_i \subset S_j & \text{if} \quad i < j \\ S_i & \text{is optimally sampled.} \end{cases} \tag{1}$$

However, equation 1 rules out the possibility of extending a subset by one basis image and sampling uniformly for each subset. A solution to cumulative subsets with individual optimal sampling is to create a hierarchical Poisson disk sampling – presented in [MF92] – of 640 samples on the domain of light source positions (i.e. the part of the sampled hemisphere above the object) and position the light sources at these sample positions.

However, implementing the hierarchical Poisson disk distribution approach directly is not practical since this would require us to place a light source at every sampled non-grid point. In our case, this would also be very costly to build, since it would require us to have 640 light sources. Practical approaches do exist to capture basis images with only a few light sources. Sampling light sources positions at grid points on a plane, as proposed by [LWS01], or the use of a Light Stage, as proposed by [DHT*00] are very common and practical. However, these approaches sample the light directions at grid points in cubic or polar parameterization.

We have chosen to build a custom Light Stage, similar to [HCD01, MPDW04]. The cylinder and the camera are placed on a turntable which rotates in 32 steps over 90 degrees. However, in our setup, instead of a semi-circular brace on which the light sources are mounted, we mount a set of 20 light sources on a vertical column. Thus, the sampling of light source positions is a grid pattern of 20×32 on a cylinder around the scene. This is close to optimal for a section of the hemisphere close to the equator and satisfies our needs. Due to this setup, our sampling domain of the light sources on the hemisphere covers 90^o of azimuth angle and 45^o of tilt angle, being a solid angle of $\sqrt{2}\pi$ sr. Note that any other Light Stage approach could be used here as well.

With our Light Stage approach, cumulative subsets of the acquired basis images can easily be produced: given the set of 640 basis images, we can create cumulative subsets for these 640 images by numbering the images. This numbering then defines the subsets: a subset of $N(\leq 640)$ basis images consists of the basis images with numbers 1 through N. Although this allows to obtain cumulative subsets of basis images, these subsets may not be individually ideally sampled.

To ensure optimally sampled subsets, the numbering of the light sources is of importance. With an intelligent numbering scheme, each subset can be nearly optimally sampled: since the 640 basis images correspond to a 20×32 sampling grid, we compute 640 samples with a modified hierarchical Poisson disk sampling method on the light source position domain. For each iteration step, we compute the next sample in the hierarchical Poisson disk sampling. Since the true light source positions are grid points, we snap this sample to the nearest grid point and tag this grid point with the number of the iteration step.

This results in a numbering for the grid points of the light sources and thus also for the accompanying basis images. Due to this numbering, we ensure that all subsets have the best possible discrete sampling over the domain, while keeping the subsets cumulative. The sampling for the subsets with 10, 50 and 100 basis images is displayed in figure 3.

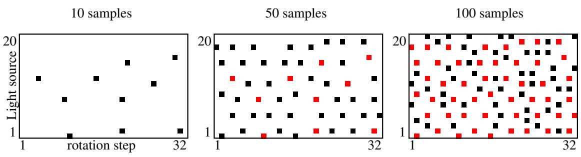


Figure 3: *The subsets of light source positions with 10, 50 and 100 samples on the sampling grid. As can clearly be seen, all subsets have a nearly optimal individual sampling of the domain.*

4.3. Generating Relit Images

To generate the series of relit images, we first decide on the different subsets for which a relit image is to be rendered. This decision depends on the material. For a chosen subset of basis images, we create a weighted sum of these basis images. These weights are defined by the novel incident illumination and the light source positions corresponding to the basis images in the subset. This relighting approach has proven to be a good method for relighting with static illumination, as demonstrated by [MPDW04].

The novel illumination will have an impact on the relit images. In our experiments, we choose to relight the cylinders with three colored soft spots (see figure 4). This allows us to visualize three different highlights while the illumination is not too high frequent.

5. Experiments

5.1. Experiment Methodology

We used the MLDS method presented in [MY03] to obtain a quality scale for each of the test setups. In each test, the observer was presented with all possible combinations of two

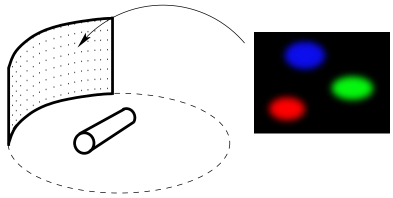


Figure 4: *The lighting originates from the sampled domain of light sources. In our setup, this is a quarter of a cylinder around the object. To relight, a weighted sum of basis images is computed. The novel illumination is decomposed into weights depending on the used basis images.*

couples of relit images and was asked to select the couple with the largest perceived difference according to the criterion requested. In our experiments, we asked the observers to judge the difference in perceived realism between all the images.

This method has several advantages over previous approaches, that required the observer to sort or make pairwise comparisons between the stimuli themselves([Gui54]). This class of methods – introduced by [Thu27] – relies on the fact that observers behave stochastically in their choices between stimuli; thus it follows that the stimuli may only differ by a few *just noticeable differences* (JND's). By using the distance between two images itself as stimulus, this restriction is overcome and a larger perceptual range can be studied.

The two pairs of images were presented simultaneously on a CRT monitor in a slightly darkened environment. The observers were unaware of the goal of the tests, and all received the same instructions. All participants had normal, or corrected to normal, vision. Each test took on average 10 minutes to complete.

5.2. Experiment Setup

For each scene, we selected 10 relit images that cover the entire perceptual range that we wish to study. The observers were presented with the $C_{10}^4 = 210$ quadruples of images for each of the 3 scenes. For the matte and intermediate gloss materials, not all 640 basis images are required to generate realistically relit images. For the matte scene, the presented stimuli were constructed with relit images based on 2, 3, 4, 5, 6, 7, 8, 9, 10 and 12 basis images. The stimuli for the intermediate gloss scene used 2, 3, 5, 8, 10, 12, 15, 17, 20 and 22 basis images. For the high gloss scene, substantially more basis images are required for realistic shading reproduction, resulting in the choice of 3, 5, 12, 25, 50, 100, 150, 200, 400 and 640 basis images for the stimuli. The resulting relit images are shown in figure 5.

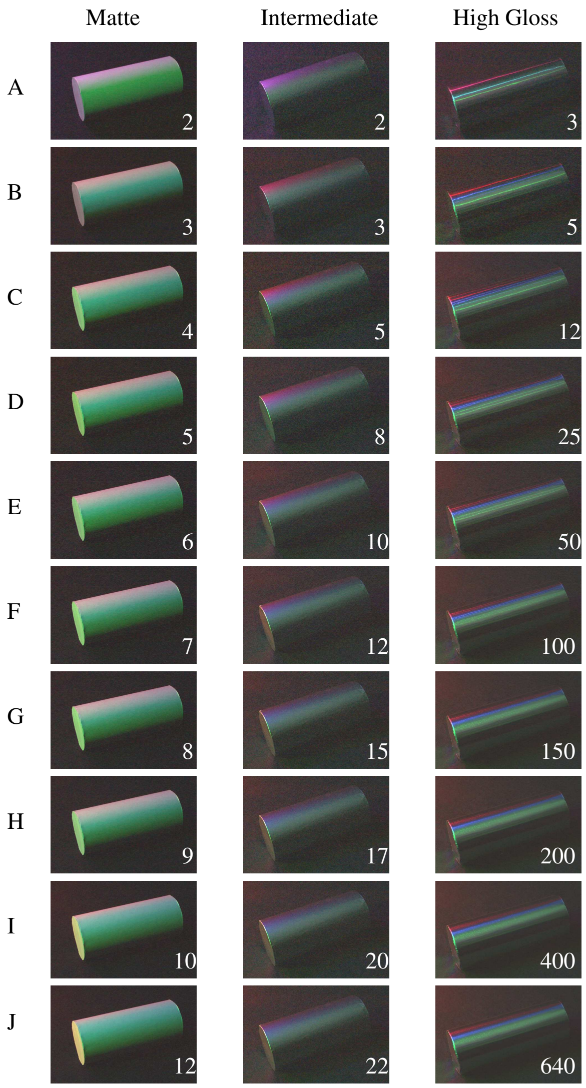


Figure 5: *The relit images for each of the three cylinders. From left to right: matte, intermediate and high gloss. The number of used basis images is depicted in the lower right bottom of each relit image.*

5.3. Processing of Data

The obtained data was fitted on an extended version of the model introduced by Maloney and Yang [MY03]. A lapse factor and observer dependent variance factors were added to more accurately describe the measured data. We used the robust optimization procedure *fminsearch* in the MATLAB environment as a basis to implement the maximum likelihood method.

The first of these extensions, the lapse factor, was necessary because several observers reported having lapses during the long test. Due to the sensitivity of a maximum likelihood estimation to lapses (see [WH01]), we extended the MLDS model to take this into account. A lapse factor was estimated in an unrestricted way by the fitting process and always remained within the interval [0,0.06], which was also reported by [WH01] to be a very reasonable interval.

The second extension, observer dependent variance factors, was necessitated by the complexity of the stimuli presented. As mentioned by [MY03], sometimes the data for one individual observer does not result in a unique solution. This required us to perform a global optimization over the data of all observers instead of one per observer. In order to allow for observer-dependent factors (e.g. the attention they invested in the test) we introduced an observer-dependent variance factor per observer.

The resulting quality scales are shifted so that the perceptually perfect stimulus corresponds to zero-quality and stimuli perceived as worse correspond to negative values. The distances between all of these values was also normalized in correspondence to the square mean of the variances. The larger the distance into the negative, the greater the distance was perceived by the "average observer" in our tests.

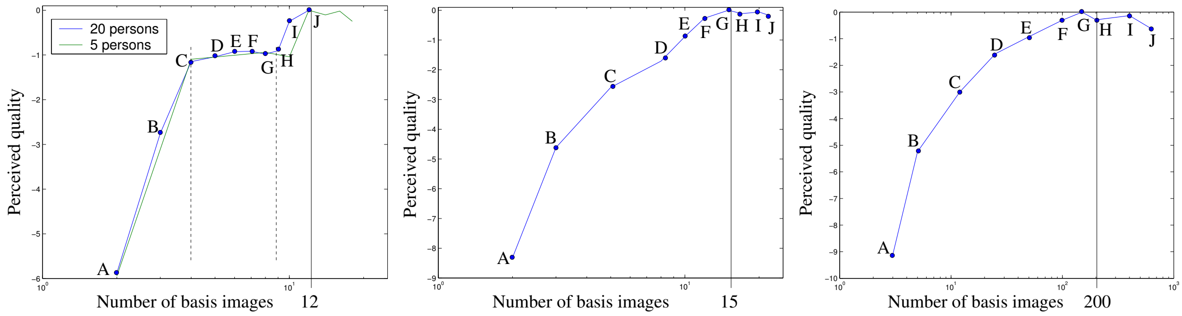
5.4. Results of tests

In this subsection, we discuss the resulting scales for the different surface materials: matte, intermediate and high gloss.

Matte In figure 6(a), the resulting quality scale for our matte material is depicted for our original test with 20 observers along with an additional test with only 5 observers but a larger test domain. A strong increase in image quality for the first few basis images is evident, with a transition into a plateau and a subsequent second strong improvement around the stimulus with 10 basis images (I). This second improvement is the result of the side of the cylinder undergoing an above average change in lighting by this 10th basis image. We performed an additional test with up to 18 basis images and 5 observers (see figure 6(a), green graph), that confirms that there is a negligible quality increase after the stimulus with 12 basis images.

Because we sampled a solid angle of $\sqrt{2}\pi$ sr, one can see that for a matte material, a uniform sampling where every light source occupies 9.25e-2 sr is sufficient to reproduce perceptually perfect images. This implies that 70 basis images – corresponding to uniformly distributed light sources over the hemisphere – is sufficient to reconstruct the shading properties of a matte material similar to ours.

Intermediate Gloss The quality increase of the perceptual scale for the intermediate glossy cylinder (figure 6(b)) is less pronounced than for the matte material. Near the stimulus with 15 basis images (G), the quality increase is below statistical significance. A solid angle of 7.4e-2 sr is therefore a sufficiently dense sampling for the light sources. With 120



(a) Matte material. (b) Intermediate glossy material. (c) High glossy material.

Figure 6: *The quality scales for the number of basis images for relighting of our three tested materials.*

basis images from light sources over the hemisphere, the shading of an intermediate glossy material can be realistically reproduced.

High Gloss The quality scale of the high gloss cylinder (figure 6(c)) confirms the trend that the perceptual quality increase tends to require more basis images as the surface is able to reflect more high-frequent details of the illumination. The stagnation of quality occurs at about 200 basis images. A dense sampling of the hemisphere with a solid angle of 5.5e-3 sr is adequate to faithfully relight scenes with high gloss materials present. We conclude that by using 1150 light sources – uniformly distributed over the hemisphere – one can reconstruct the shading of our high gloss material to perceptual perfection.

5.5. Validation of fitted models

To validate our results, we applied the procedure proposed by [MY03] to our extended model. [KLST71] state the necessary and sufficient conditions for which an observer can be said to judge stimuli according to a difference scaling model. In [MY03] one of these conditions is reformulated as the *Six-Point Property*, providing a method to validate the fitted model. The Six-Point Property states the transitivity property for specific combinations of 3 quadruple tests. Specifically, when the observer judges the difference $d(i, j)$ to be greater than $d(i', j')$ and $d(j,k) > d(j',k')$, then he/she should also judge $d(i, j) > d(i',k')$. When deriving a difference scale for n stimuli, there exist C_n^6 such relations. Because of the stochastic nature of the observer's decisions, there is a nonzero probability of violating this property. Subsequently, we can compute the probability of observing the violations and compliances of a single participant.

The validation of the fit is achieved by comparing the proportion of Six-Point Property failures of an observer's test results to the distribution of those failures, generated by Monte-Carlo techniques. The observed data for one participant is rejected if the probability of observing such Six-Point Property failures as present in the data lies below the α-percentile of the histogrammed Monte-Carlo values. If our datasets were generated by our fitted models, α percent would lie below this α-percentile. The number of rejections of the observed data by the fitted model was not above the expected number of rejections. Therefore, we conclude that the observed data is consistent with the fitted models presented.

6. Extensions

In this section we consider other dimensions that influence the perceptual realism of relit scenes: lighting, texture and scene composition.

6.1. Influence of Frequency of Lighting

The perceived quality of relit images is dependent on the type of illumination with which it is relit. Thus far, we employed a fairly high-frequent illumination (figure 4). In order to ascertain the sensitivity of the quality scale for a change of illumination, we performed the same type of experiments (only 6 observers) for all three materials with a different illumination. The images were relit using a homogeneous illumination (i.e. white light from the complete sampled domain).

The quality scales of the three materials resulting from using the homogeneous low frequent illumination as well as the high frequent illumination are shown in figure 7. We found that the frequency content of the lighting indeed has an effect on the perceived quality scale.

For a high glossy material, the two scales are almost identical, while the intermediate and certainly the matte material show significant deviations. For lower frequency reflectance surfaces (i.e. matte and intermediate), every estimated quality scale for the high frequency illumination tests has a larger domain than the corresponding low frequency tests, and lie mostly below them. This means that the difference in quality is more easily seen by the observers as the illumination has a larger high frequency component.

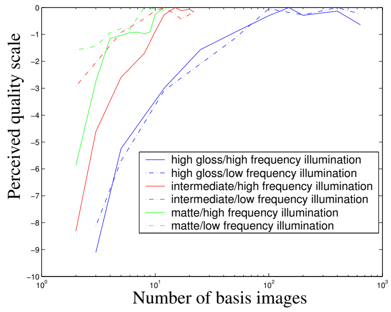


Figure 7: *Quality scales for the three glossy materials with two different illuminations.*

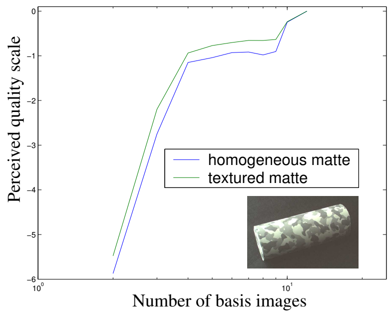


Figure 8: *In the right bottom, the textured matte cylinder is depicted. The two quality scales for the non-textured matte cylinder and the textured matte cylinder.*

6.2. Influence of Texture on the Surface

In addition to the three categories of surface reflection types, we also investigated the effect of the presence of surface texture on the perceptual image scale for a matte reflection type. The importance of textures in computer graphics has already been studied in [FSPG97]. In order to investigate, we manufactured an extra matte cylinder with texture (see right bottom of figure 8). With this textured cylinder, we performed the same experiment as with the other cylinders, and derived a quality scale. Because of the already steep increase of image quality for the matte cylinder, only a small effect is noticeable. The presence of texture does not seem to influence the threshold at which there is no quality difference to be perceived for matte materials.

6.3. Influence of complex scene composition

All tests thus far were performed on simple cylinders. In order to investigate the impact of scene composition on the perceptual quality scale, we presented 6 observers with 2 additional tests. The setup for the first of these tests can be seen in figure 9(a), depicting a highly glossy mug, a stapler and a slightly glossy ball. The second test setup was constructed by just showing the highly glossy mug to the observers (as in figure 9(b)).

(a) Complete Scene Setup

(b) Partial Scene Setup

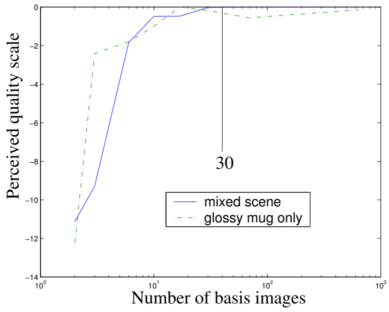


Figure 9: *Perceptual quality scales of complete scene and only the high gloss mug*

The results of these two tests is shown in figure 9. The threshold for perceptual perfection remains approximately the same for both setups (30 basis images).

7. Discussion

For each of the tested materials, we obtained a quality scale for the relighting of the material based on the number of basis images. As can be seen in the graphs in figure 6, there exists a saturation point at which adding more basis images is irrelevant to the observer. For the matte material, this saturation occurs at 70 basis images, corresponding with 70 uniformly sampled light source positions on the hemisphere. Similarly, for the intermediate and high gloss tested materials, we derive a saturation point for the number of light sources on the hemisphere at 120 and 1150, respectively. The glossy mug from subsection 6.3 was added as an additional material sample, with a minimal sampling of 170 light sources over the hemisphere.

If we plot all these thresholds as a function of gloss, we can obtain a graph defining the threshold for the number of basis images for relighting any kind of glossy material. In order to do this, we need to know the gloss of each sampled material. For each of the materials, the nearest stimulus used by [PFG00] was matched and their perceptual contrast gloss dimension used as gloss unit. We normalized this dimension so that 0 corresponds to perfect matte and 1 to completely specular. These findings are shown in figure 10, defining an approximate lower bound for the number of basis images re-

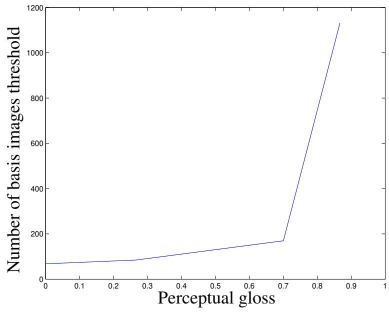


Figure 10: *Minimal sampling of the hemisphere for perceptual gloss units*

quired to relight glossy surfaces in a perceptually acceptable fashion.

All these thresholds were found to be reasonably independent of the type of illumination used, the presence of textures and the complexity of the scene.

8. Conclusion and Future Work

We studied the minimal sampling required to relight scenes consisting out of materials of different gloss. A perceptual quality scale for each of these was constructed, providing a more substantial basis to perform relighting of scenes than the previously employed method of relying on know-how or just capturing a lot more images than strictly necessary. By fitting our stimuli on the perceptual scale of [PFG00], we provide a minimal relighting configuration that will provide perceptual convincing results. For our matte, intermediate and highly glossy materials we found 70, 120 and 1150 basis images respecively to be sufficient to realistically relight scenes. After establishing the glossy perceptual contrast gloss dimension of the materials to be relit, one can use figure 10 to get an upper bound for the number of basis images that will provide realistic relit images.

Due to the high number of factors involved in relighting, this paper addresses the most prominent one: type of materials. Some preliminary investigations were performed as to how the type of illumination, texturing and scene composition relate to the perceived relit images. None of these seem to have a significant impact on the minimum number of basis images. Yet these and many other factors remain to be considered in future research. Interreflection in a complex scene, self-occlusion, and geometrical properties of the relit objects are among factors that remain to be addressed.

References

[BO87] BILLMEYER F. W., O'DONNELL F. X. D.: Visual gloss scaling and multidimensional scaling of painted specimens. *Color research and applications 12*, 6 (1987), 315–326. 2

[DHT*00] DEBEVEC P., HAWKINS T., TCHOU C., DUIKER H.-P., SAROKIN W., SAGAR M.: Acquiring the reflectance field of a human face. In *SIGGRAPH 2000, Computer Graphics Proceedings* (July 2000), Addison Wesley, pp. 145–156. 1, 3

[FSPG97] FERWERDA J. A., SHIRLEY P., PATTANAIK S. N., GREENBERG D. P.: A model of visual masking for computer graphics. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1997), ACM Press/Addison-Wesley Publishing Co., pp. 143–152. 7

[Gui54] GUILFORD J. P.: *Psychometric Methods*. McGraw-Hill, New York, 1954. 4

[HCD01] HAWKINS T., COHEN J., DEBEVEC P.: A photometric approach to digitizing cultural artifacts. In *In 2nd International Symposium on Virtual Reality, Archaeology, and Cultural Heritage, Glyfada, Greece, November 2001.* (2001). 1, 3

[KLST71] KRANTZ D. H., LUCE R. D., SUPPES P., TVERSKY A.: *Foundations of measurement*, vol. 1: Additive and polynomial representations. Academic Press, New York, 1971. 6

[LWS01] LIN Z., WONG T.-T., SHUM H.-Y.: Relighting with the reflected irradiance field: Representation, sampling and reconstruction. *IEEE Computer Vision and Pattern Recognition 1* (Dec. 2001), 561–567. 2, 3

[MDA02] MASSELUS V., DUTRÉ P., ANRYS F.: The free form light stage. In *Rendering Techniques EG 2002* (2002), Annual Conference Series, Eurographics Association, pp. 247–256. 1

[MF92] MCCOOL M., FIUME E.: Hierarchical poisson disk sampling distributions. *Graphics Interface '92 1* (1992), 94–105. 3

[MGW01] MALZBENDER T., GELB D., WOLTERS H.: Polynomial texture maps. In *SIGGRAPH 2001, Computer Graphics Proceedings* (2001), Fiume E., (Ed.), ACM Press / ACM SIGGRAPH, pp. 519–528. 1

[MPDW04] MASSELUS V., PEERS P., DUTRÉ P., WILLEMS Y. D.: Smooth reconstruction and compact representation of reflectance functions for image-based relighting. In *Rendering Techniques EG 2004* (2004), Annual Conference Series, Eurographics Association, pp. 287–298. 1, 3, 4

[MY03] MALONEY L. T., YANG J. N.: Maximum likelihood difference scaling. *Journal of Vision 3* (2003), 573–585. 2, 4, 5, 6

[OKV04] OBEIN G., KNOBLAUCH K., VIÉNOT F.: Difference scaling of gloss: Nonlinearity, binocularity, and constancy. *Journal of Vision 4* (2004), 711–720. 2, 3

[PFG00] PELLACINI F., FERWERDA J. A., GREENBERG D. P.: Toward a psychophysically-based light reflection model for image synthesis. In *SIGGRAPH* (2000), pp. 55–64. 2, 7, 8

[Thu27] THURSTONE L. L.: The method of paired comparisons for social values. *Journal of Abnormal and Social Psychology 21* (1927), 384–400. 4

[WH01] WICHMANN F., HILL N.: The psychometric function: I. fitting, sampling and goodness-of-fit. *Perception and Psychophysics 63 (8)* (2001), 1293–1313. 5

[WHON97] WONG T.-T., HENG P.-A., OR S.-H., NG W.-Y.: Image-based rendering with controllable illumination. In *Eurographics Rendering Workshop 1997* (June 1997), Springer Wien, pp. 13–22. 1