Visual Equivalence: an Object-based Approach to Image Quality

James A. Ferwerda [*], Ganesh Ramanarayanan[†], Bruce Walter [°], and Kavita Bala[†]
[*]Munsell Color Science Laboratory, Center for Imaging Science, Rochester Institute of Technology, Rochester, NY, USA
[†]Department of Computer Science, [°]Program of Computer Graphics, Cornell University, Ithaca, NY, USA

a) path traced image	b) environment mapped image	c) visible differences map

Figure 1: *a,b) Computer graphics images rendered with different reflection algorithms and c) the output of a VDP metric showing areas of visible difference. Note that while the images are visibly different, they are similar in quality, and convey equivalent information about object appearance*

Abstract

In this paper we introduce a new approach to characterizing image quality: visual equivalence. Images are visually equivalent if they convey the same information about object appearance even if they are visibly different. In a series of psychophysical experiments we explore how object geometry, material, and illumination interact to produce images that are visually equivalent, and we identify how two kinds of transformations on illumination fields (blurring and warping) influence observers' judgments of equivalence. We use the results of the experiments to derive metrics that can serve as visual equivalence predictors (VEPs) and we generalize these metrics so they can be applied to novel objects and scenes. Finally we validate the predictors in a confirmatory study, and show that they reliably predict observer's judgments of equivalence. Visual equivalence is a significant new approach to measuring image quality that goes beyond existing visible difference metrics by leveraging the fact that some kinds of image differences do not matter to human observers. By taking advantage of higher order aspects of visual object coding, visual equivalence metrics should enable the development of powerful new classes of image capture, compression, rendering, and display algorithms.

Introduction

Measuring image differences is an important aspect of image quality evaluation, and a variety of metrics have been developed for this purpose. *Numerical metrics* measure physical differences between a reference image and test image and characterize quality in terms of the distance from the reference to the test. Well known numerical metrics include mean squared error (MSE) and peak signal to noise ratio (PSNR)[1]. Although these metrics are easy to compute, they often do not correlate well with observers' judgments of image differences. For this reason, *perceptual metrics* have been developed that incorporate computational models of human visual processing. In these metrics visual models are used to represent an observer's responses to the reference and test images and then these responses are compared to identify visible differences. Popular perceptual metrics include Daly's Visible Differences Predictor (VDP)[2] and the Lubin/Sarnoff model[3]. These metrics typically do a better job at predicting the visual impact of common imaging artifacts such as noise and quantization on perceived image quality, and many researchers have successfully applied these perceptual metrics to important problems in digital imaging. However current metrics have an interesting limitation that is illustrated in Figure 1.

Figure 1a and 1b show two computer-generated images of a tabletop scene. Figure 1a was rendered using *path tracing*, a physically accurate but computationally intensive algorithm that can produce faithful simulations of environmental light reflection and transport[4]. It can take hours to render a single image using path tracing. In contrast, Figure 1b was rendered using *environment mapping*, a fast but approximate rendering technique that uses an image of the surround rather than the model of the surround to illuminate the objects on the tabletop. Environment mapping is a standard feature of commodity graphics hardware and can render images like the one shown in Figure 1b at interactive rates. One consequence the approximations used in environment mapping is that illumination features such as surface reflections are warped with respect to the geometrically correct features produced by path tracing. This can be seen by comparing the images reflected by the two teapots.

If we take the path traced image as the reference, and the environment mapped image as the test, and process the images with one of the standard perceptual metrics (in this case an implementation of Daly's VDP[5]), the metric produces the difference map shown in Figure 1c which correctly indicates that the images *are* visibly different (green and red pixels 75% and 95% probability of detection respectively). However an important question is: *are these meaningful image differences?*

When we look at images we don't see pixels. Rather, we see objects with recognizable shapes, sizes, and materials, at specific spatial locations, lit by distinct patterns of illumination. From this perspective the two images shown in Figure 1 are much more similar than they are different. For example, the shapes, sizes, and locations of the objects shown in the images appear the same; the objects appear to have the same material properties; and the lighting in the scenes seems the same. Although the images are *visibly different* they are *visually equivalent* as representations of object appearance. The existence of images like these coupled with the growing range of image transformations used in computer graphics, digital imaging, and computational photography points to the need for a new kind of image difference/quality metric that can predict when different classes of imaging algorithms produce images that are visually equivalent.

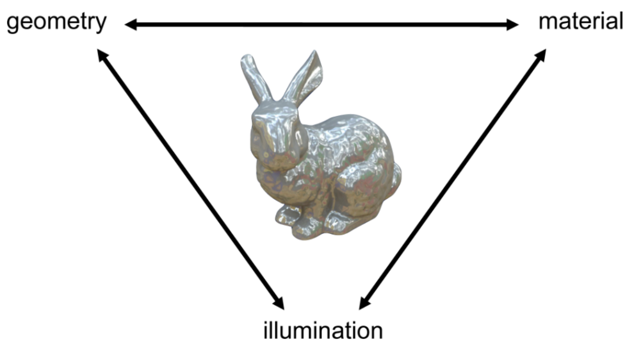

Figure 2*: Factors that affect object appearance. Dynamics and viewpoint can also play significant roles.*

Understanding Object Appearance

The concept of visual equivalence is based on the premise that two visibly different images can convey the same information about object appearance to a human observer. To develop metrics for predicting when images will be visually equivalent we need to understand the factors that influence object appearance.

Figure 2 shows a computer-generated image of a chrome bunny. We perceive the properties of this and other objects based on the patterns of light they reflect to our eyes. For a fixed object and viewpoint these patterns are determined by the *geometry* of the object, its *material* properties, and the *illumination* it receives. The perception of each of these properties is the subject of an extensive literature that will only be briefly reviewed here. More comprehensive introductions on these subtopics are available in the papers cited.

Shape perception: The central problem in shape perception is how the visual system recovers 3D object shape from the 2D retinal images. Many sources of information for shape have been identified including stereopsis, surface shading, shadows, texture, perspective, motion, and occlusion[6,7] Recent work has tried to characterize the effectiveness of these different sources[8-11] and to model how they combine to provide reliable shape percepts[12].

Material perception: Although there is significant interest in industry on the topic of material perception[13], there has been relatively little basic research on the topic[14,15]. This situation is changing with the development of advanced computer graphics techniques that allow the accurate simulation and systematic manipulation of realistic materials in complex scenes. Active research topics include the perception of 3D surface lightness and color[16-18], gloss perception[11,19-22], perception of translucency[26,27] , and 3D texture appearance[28-30].

Illumination perception: Historically, illumination has been regarded as a factor that needs to be discounted to achieve shape and lightness/color constancy[16-18,31], but recently there has been interest in understanding illumination perception itself. Recent studies include the characterization of natural illumination statistics[32] and surface illuminance flow[33], the perception of illumination directionality and complexity[34,35], and tolerance for illumination inconsistencies[36].

So an object's appearance is based on the images it reflects to our eyes, and these images are determined by the object's geometry, material, and illumination properties. How the visual system disentangles the image information to perceive these object properties is one of the great unsolved problems in vision research. Although eventually we would like to understand this process, the goal of this paper is more immediate: to develop metrics that can predict when visibly different images are equivalent as representations of object appearance. To achieve this goal conducted a series of experiments that investigated when different configurations of object geometry, material, and illumination produce visually equivalent images.

Experiments

Even for a single object, the space of images spanned by all possible variations in object geometry, material properties, and scene illumination is vast. To begin to quantify the phenomenon of visual equivalence we had to constrain the scope of our studies. Starting from the proof-of-concept demonstration shown in Figure 1, we decided to study visual equivalence across two kinds of illumination transformations (blurring and warping) for objects with different geometric and material properties. The following sections describe our methods and procedures.

Stimuli

To conduct the experiments we created a set of images that

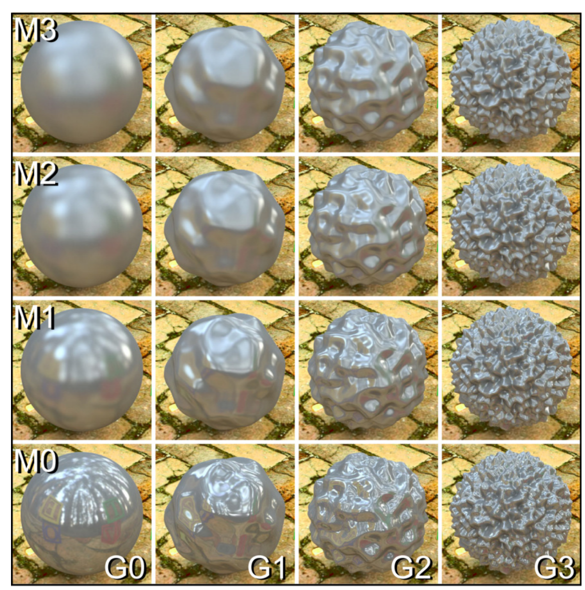

Figure 3*. The geometries and materials of the objects used in the experiments. Parameters were chosen to be perceptually uniform in both surface "bumpiness" and surface reflectance.*

would allow us to systematically explore the interactions between object geometry, material, and illumination. To accomplish this we used computer-generated images. Figure 3 shows representative images from our stimulus set. The scene consisted of a bumpy ball-like test object on a brick patio flanked by two pairs of children's blocks. The following paragraphs describe the parameters we used to generate the images.

Geometry: The four object geometries (G0-G3) shown in the rows of Figure 3 were defined as follows. Object G0 was a geodesically tesselated sphere with 164 vertices. Objects G1 though G3 were generated by placing the sphere in a cube of Perlin noise[37] and displacing the vertices according to the 3d noise function. By varying the size of the cube relative to the sphere (scale factors of $\sqrt{2}$,1,$\sqrt{1/2}$,1/2,$\sqrt{1/8}$ respectively) it was possible to produce random surface displacements that were constant in amplitude but varied in spatial frequency bandwidth. In pre-testing the objects were informally judged to be equally spaced with respect to surface "bumpiness". We chose these geometries for several reasons: 1) their functional definitions facilitated analysis of the effects of geometry on appearance; 2) there is a precedent in the shape perception literature for similar geometries[8,24]; and 3) there are recent studies that point to the importance of mesoscale surface texture in the perception of material and illumination properties[29].

Materials: The columns of Figure 3 show the materials used in the experiments, which represent rolled aluminum with different degrees of microscale roughness. Materials were defined using the Ward[38] light reflection model which has three parameters to describe surface reflectance properties: ρ_d - diffuse reflectance, ρ_s - specular reflectance, and α - specular lobe width). For all materials $\rho_d = 0.19$ and $\rho_s = 0.15$. Alpha values for M0 through M3 were set to 0.01, 0.06, 0.11, and 0.16 respectively. We chose these parameters to produce a set of materials that 1) spanned a significant range of mid to high gloss reflectance and to represent perceptually equal changes in glossy appearance[20].

Illumination: Recent studies have demonstrated the importance of real-world illumination for the accurate perception of shape and material properties[11,22]. For this reason we lit our model using Debevec's[41] "Grove" HDR environment map that captures the illumination field in the Eucalyptus grove at UC Berkeley[41]. We chose this map in particular, because Fleming et al.[22] found that it allowed subjects to most accurately discriminate material properties. Starting with the original "Grove" map, we first generated a *reference illumination map* that incorporated the other components of our scene (i.e. the brick patio and the children's blocks). We then generated two sets of transformed maps. The top row in Figure 4 shows the central panel from the "blurred" map set. Here Blur1-Blur5 represent convolutions of the maps with Gaussian kernels whose widths roughly correspond to Ward model α values of 0.01, 0.035, 0.06, 0.085, and 0.11 respectively. The second row shows these maps reflected in an object with geometry G1 and material M0. The third row shows the central panel from the "warped" map set. Here Warp1-Warp5 represent warps of different magnitudes created using a method analogous to the one described in the "*Geometry*" section. Finally, the bottom row shows these warps reflected by the same G1/M0 object.

Rendering and display: Images were rendered at 484x342 as high dynamic range (HDR) floating point images using a custom-built physically-based Monte Carlo path tracer. Overall 176 images were rendered for the stimulus set (4 geometries x 4 materials x 2 illumination transformations x 5 illumination transformation levels+ 16 reference images). For display, the HDR images were tone mapped using a global sigmoid[39] that was tuned to the characteristics of the display (Dell 2000FP, 20" diagonal LCD, 1600x1200 resolution, sRGB color space, max luminance 200

***Figure 4**. The two classes of illumination transformations used in the experiments (blur and warp). The upper and lower panels show direct views of the blurred and warped illumination maps and their effects on the appearance of a representative object (G1/M0).*

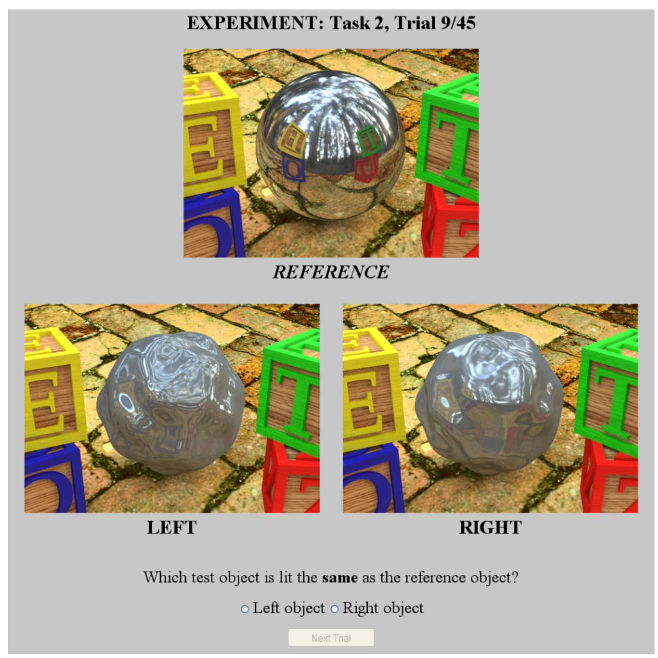

***Figure 5**. Interface used in the experiments: The objects in the LEFT and RIGHT images have the same geometry and material properties. In this condition one is rendered with the REFERENCE illumination map and the other is rendered with a warped map. The observer's task is to identify the object that is lit with the same illumination as the REFERENCE.*

cd/m^2, 60:1 dynamic range, gamma 2.2). The images were viewed under dim office lighting conditions. At a nominal 24" viewing distance each image subtended approximately 12 degrees of visual angle and each test object subtended approximately 7 degrees.

Procedure

The images in the stimulus set were presented to subjects in pairs using the browser-based interface shown in Figure 5. In some conditions a third reference image was shown above the test pair. *In all cases the test pairs showed objects with identical shapes and material properties* (the G/M combinations shown in Figure 3). In each case one of the images was rendered using the reference illumination map, and the other was rendered using one of the transformed maps (Blur1-5 or Warp1-5 as shown in Figure 4). Separate experiments were conducted for the "blurred" and "warped" image sets. An experiment consisted of four related tasks that asked about the image pairs.

Image differences: In this task subjects were shown a reference image, and a pair of test images. All three images showed the same object. The reference image and one of the test images were identical. The other test image was rendered with one of the transformed maps. Subjects were asked: **"Which test image is the same as the reference image?"**. The purpose of this task was to determine when images rendered with the transformed maps were visibly different (in the VDP sense) from images rendered with the reference map.

Shape differences: In this task subjects were shown two images of the same object. One object was rendered with the reference map the other was rendered with an altered map. Subjects were asked: **"Are the left and right test objects the same shape?"**. The intent of this task was to determine if the transformed maps produced illusory changes in the apparent shapes of the objects.

Material differences: In this task subjects were shown two images of the same object. One object was rendered with the reference map the other with a transformed map. Subjects were asked: **"Are the left and right test objects made of the same material?"**. The intent of this task was to determine if the altered maps produced illusory changes in the apparent material properties of the objects.

Illumination differences: In this task subjects were shown a reference image, and a pair of test images. The reference image showed a mirror sphere rendered with the reference map. The test images showed identical objects, one rendered with the reference map and the other rendered with a transformed map. Subjects were asked: **"Which test object is lit the same as the reference object?"**. The intent of this task was to determine if subjects can use surface reflection patterns to detect differences in scene illumination.

Each subject performed the image difference task first. The shape, material, and illumination tasks were then delivered in random order. Within each task both the overall order of presentation and left/right positions of the images were randomized across trials. On each trial subjects entered their responses with a keyboard and mouse. The trials were open-ended and subjects could take as much time as they needed. On average subjects took about 45 minutes to complete all four tasks.

Overall 30 subjects participated in the experiments (15 each in the blur and warp conditions). The subjects were university students staff, and faculty (ages 20 to 50). Many had technical backgrounds, but none in imaging. All were naive to the design and purpose of the experiments. All had normal or corrected-to-normal vision.

Note that the four tasks can be divided into two conceptual

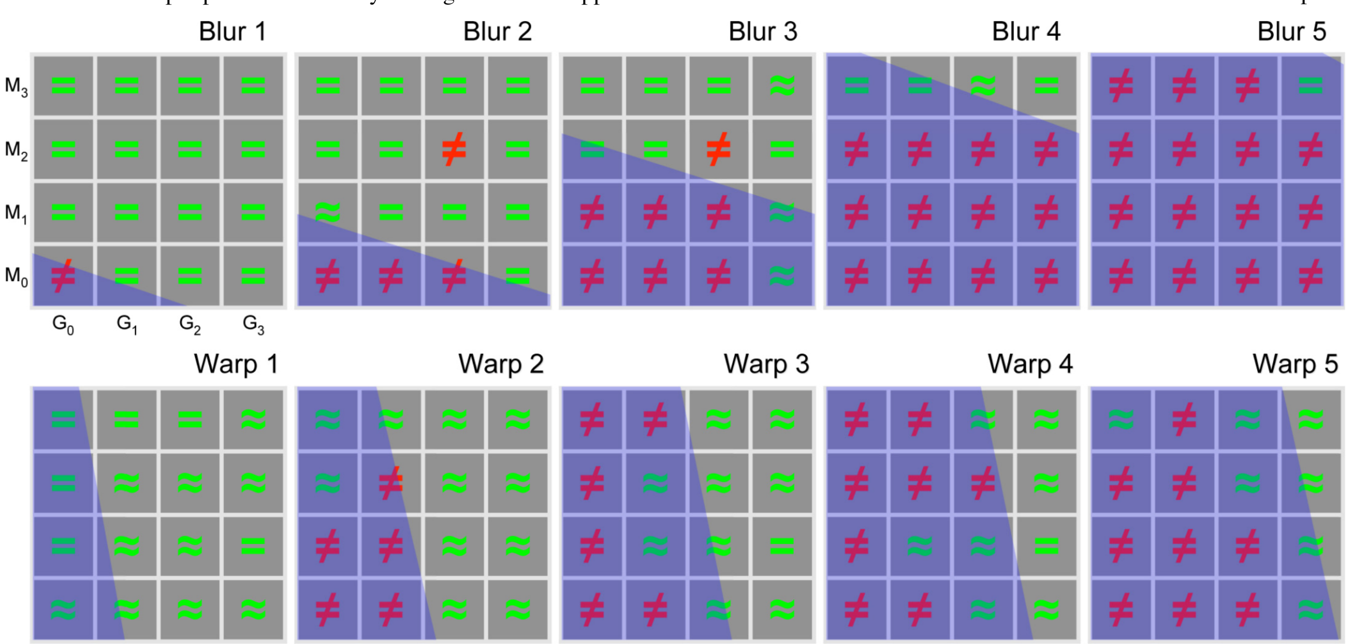

***Figure 6**: Results of the experiments: Each panel represents the objects tested (G0-G3, M0-M3) (see Figure 3). The upper and lower strips show the results for different levels of the blur and warp illumination transformations. Overall the results fall into three categories: equality, non-equality, and equivalence. The blue shaded regions show the metrics for the blur and warp transformations produced by an SVM classifier that separates the equal and equivalent cases (green symbols) from the non-equal cases (red symbols). The metrics can be used as the basis of visual equivalence predictors.*

categories. In the image difference task subjects are being asked to report on *image* differences. In the shape, material, and illumination difference tasks subjects are being asked to report on *object* differences. We chose these tasks because they should allow us to dissociate the effects of image differences on image and object appearance and quantify when different configurations of object geometry, material and illumination produce images that are visually equivalent as representations of object appearance.

Results

The results of the experiments are summarized in Figure 6. Each panel shows the set of objects we tested (G0-G3, M0-M3), and the upper and lower strips show how observer's judgments changed for different levels of illumination map blurring (Blur1-5) and warping (Warp1-5). The results fall into three categories. In general green symbols are good and red symbols are bad.

Equal: When subjects reported that the reference and test images were indistinguishable (chance performance in the image difference task) then we said that the images were equal (green equal signs in Figure 6). Note that for low levels (1,2) of blur the image differences were often undetectable and so the images appeared identical.

Not-Equal: On the other hand when subjects reported that the reference and test images were visibly different, and also reported that the objects looked different (in shape, material or illumination), then we labeled the images as not equal (red signs in the Figure). Note that the number of non-equal cases increases with the magnitude of the blur and warp transformations and appears to vary with material properties for blur (high gloss objects affected first) and with geometry for warp (smooth objects affected first).

Equivalent: Finally, when subjects reported that the reference and test images were visibly different but also said that the objects represented by the images appeared the same (same geometry and material, no clear differences in illumination) we labeled the images as equivalent. Note that while there are few equivalent cases for the blur transformation, there are cases of equivalence at all levels of the warp transformation, even the most severe.

What these results show is that there is a significant class of conditions (indicated by green symbols) where the images rendered with the transformed illumination maps are either equal or equivalent as representations of object appearance. While existing visible difference metrics (VDPs) could predict the cases of equality, they would not identify the much larger set of visually equivalent images. To take advantage of this new found latitude in perceptually acceptable image distortions we can use these results to develop a new kind of image metric: *visual equivalence predictors* (VEPs)

Defining the predictor

To turn the results of the experiment into a metric that can be used to predict whether images will be visually equivalent we need to classify the results into "good" and "bad" categories. We used Support Vector Machines[40] to fit planes that separate the equal and equivalent cases from the non-equal cases (in our final metric the planes were linearly shifted by a small amount to be fully conservative). The cutting planes are described by Equations 1 and 2 in terms of the geometric (G) material (M) and illumination (I) parameters used to generate the test images. These planes are illustrated in Figure 6 where each of the panels can be thought of as a level in a three-dimensional space (where blur or warp level is the third dimension) and the blue shaded regions show slices through the plane at different levels.

$$Blur: 0.181G + 0.546M - 0.728I + 1.027 = 0 \quad (1)$$
$$Warp: 0.772G + 0.128M - 0.456 - 0.299 = 0 \quad (2)$$

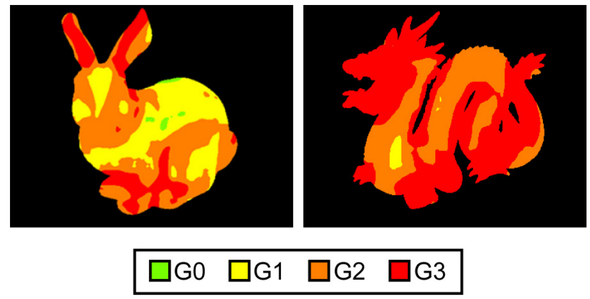

Figure 7: Generalizing the predictor. The scale indicates different degrees of normal variation in the G0-G3 objects used in the experiment. The images show the Stanford bunny and dragon models mapped into this space. This technique can be used to apply the VEP metric to novel geometries.

Generalizing the predictor

We now have a metric that can predict visual equivalence for the images in our test set. While this is interesting, to be useful we need to generalize the metric so it can predict equivalence for images of novel objects and scenes. To generalize to novel geometries we characterize the average surface normal variation for the objects in our test set (G0-G3) and map the normal variations of new objects into this space. Figure 7 illustrates this method, where the green to red scale indicates increasing levels of normal variation, and the images show how different regions of the Stanford bunny and dragon models correspond to the G0-G3 surfaces. To generalize to novel materials, we can fit surface reflectance data with the Ward model and cast the parameters into our M0-M3 material space. While this is not a comprehensive solution the materials we tested were high gloss metals that showed sharp, high contrast surface reflections. These materials represent a worst case, so our metric should be conservative with respect to this dimension. Finally with respect to illumination, we only require that the illumination field have "natural" image statistics ($\sim 1/f^2$ power spectrum), which is a modest constraint that

#	Scene	Predicted/ ExpResult	#	Scene	Predicted/ ExpResult
0	Bunny M1 Blur2 Campus	● / ●	7	G1 M0 Warp4 Galileo	■ / ■
1	Bunny M1 Blur4 Campus	■ / ■	8	Dragon M2 Warp3 Grove	● / ●
2	Dragon M1 Blur2 Grove	● / ●	9	Dragon M2 Warp5 Grove	■ / ●
3	Dragon M1 Blur4 Grove	■ / ■	10	Bunny M1 Warp2 StPeters	● / ●
4	Dragon M2 Blur2 Galileo	● / ●	11	Bunny M1 Warp5 StPeters	■ / ■
5	Dragon M2 Blur4 Galileo	■ / ■	12	Dragon M1 Warp2 Galileo	● / ●
6	G1 M0 Warp1 Galileo	● / ●	13	Dragon M1 Warp5 Galileo	■ / ■

Table 1: Results of the validation experiment: Fourteen novel scenes with the indicated geometry, material, and illumination properties were tested. The green circles indicate that the reference and test images were visually equivalent. The red squares indicate that the images were not-equal. In 13 of 14 cases the VEP accurately predicts the observer's judgments. In case #9 it was overly conservative.

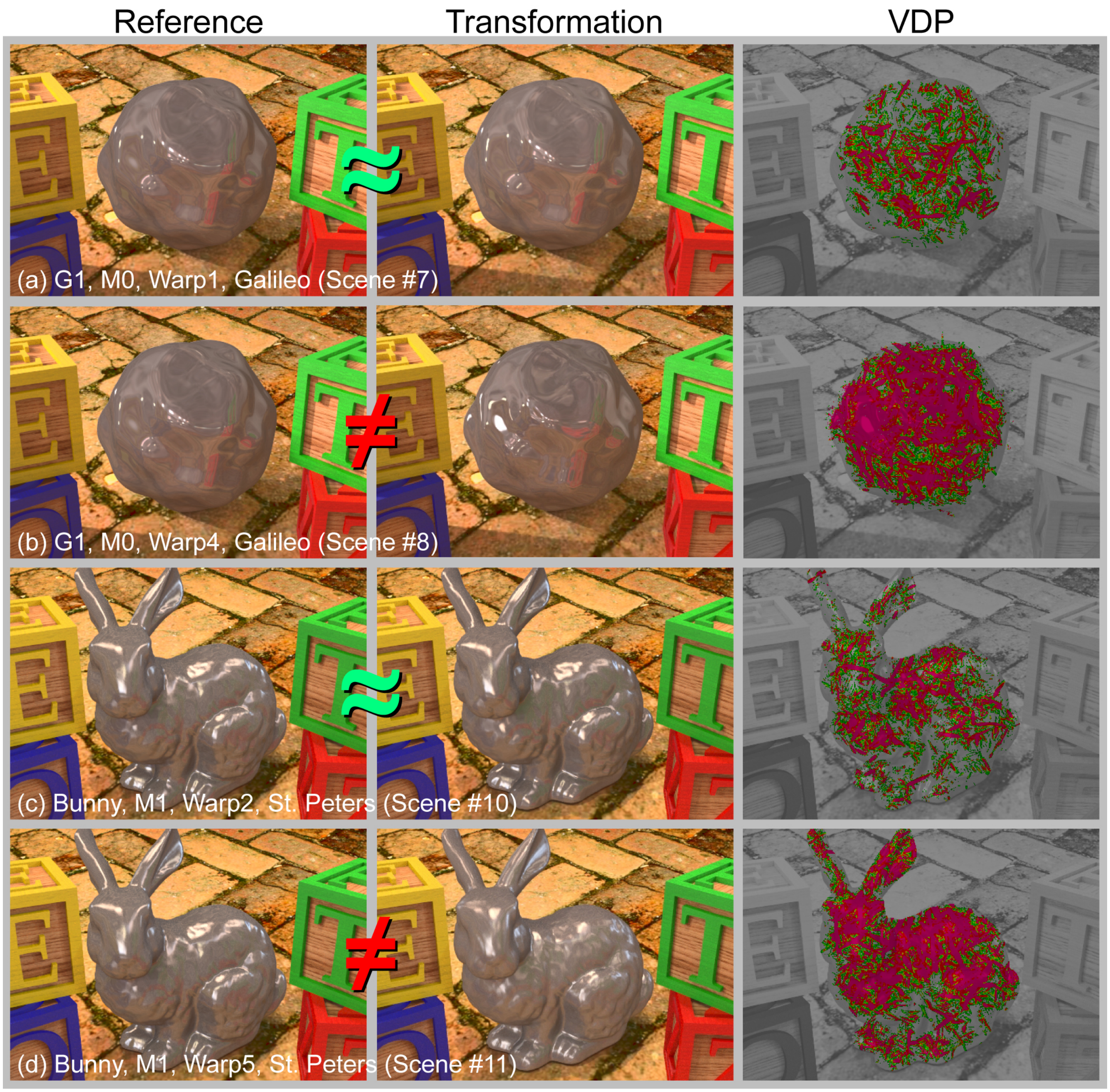

***Figure 8**. Selected images from the validation experiment. The reference images are on the left, test images are in the middle, visible difference (VDP) maps on the right. Symbols on each image pair indicate whether they were seen/predicted to be visually equivalent despite being visibly different. Note that the visual equivalence predictor (VEP) correctly predicts both equivalent and non-equivalent cases.*

relieves the metric from trying to predict equivalence under degenerate (and unrealistic) illumination cases such as point light sources.

Validating the predictor

To test the predictive power of the metric, we ran a confirmatory experiment where we created reference and test images of 14 novel scenes, ran them through the metric and also had subjects judge them using the same procedure they had in the main experiment. The geometric (G1, bunny, dragon), material (M0-2), and illumination (Debevec's Grove, Campus, Galileo, and St. Peters maps) properties of the scenes are listed in Table 1, and partially illustrated in Figure 8. Ten new subjects with the approximately the same demographics as those tested in the main experiment were tested here.

The results are summarized in Table 1 where the green circles indicate cases of equivalence and the red squares indicate non-equality. The first column shows the result predicted by our metric and the second column shows the subjects' actual judgments. The metric correctly predicted the judgments in 13 out of 14 cases (being overly conservative in one case (#9)), and was able to predict both equivalence and non-equality.

Visual equivalence and illumination statistics

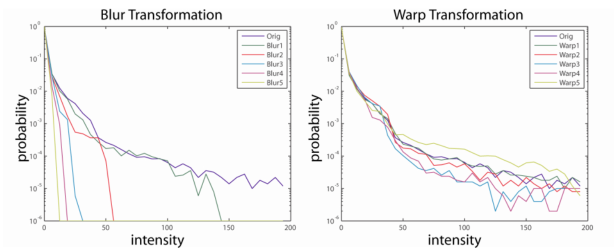

***Figure 9**. Illumination intensity histograms for the blur and warp transformations. Note that blur significantly reduces the original intensity dynamic range, while warp leave it relatively intact*

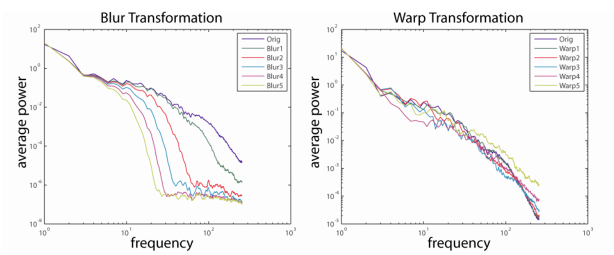

***Figure 10**. Frequency power spectra for the blur and warp transformations. Note that blur effectively low pass filters the original spectrum while warp leave it relatively constant.*

A close look at the experimental results (Figure 6) reveals that the blur transformation has many more "not-equal" cases (39 out of 80) than the warp transformation (27 out of 80). Why do some illumination maps lead to more or less visual equivalence than others? One way to gain insight into this question is to look at illumination statistics. Recent work by Dror[32] has shown that natural illumination maps, such as the one we used, exhibit many statistical regularities. How much are its statistics affected by the transformations we applied? In Figures 9 and 10, we plot two standard statistical measures from Dror's paper, illumination intensity and frequency power spectra, for the blur and warp transformations. As blur increases, both the number of high intensity locations in the illumination map (Figure 9) and the average power at higher frequencies (Figure 10) decrease; this is to be expected from the effective low pass filtering. However, with the warp, this is not so; both intensity and frequency power spectra are stay relatively constant regardless of the magnitude of the warp. From this we conclude that the blur transformation causes more deviation from natural illumination statistics than the warp, which is a possible explanation for why warp produces more equivalence. Further investigation along these lines may help us to identify other computationally useful classes of image transformations that produce images that are visually equivalent to reference renderings.

Conclusion

In this paper we have introduced a new foundation for image difference/quality metrics: visual equivalence. Images are visually equivalent if they convey the same information about object appearance even if they are visibly different. We have explored the phenomenon of visual equivalence in a series of psychophysical experiments using realistic computer generated images of objects with different geometries, materials, and illumination and have identified conditions where transformations of the illumination fields (blurring and warping) produce images that are equivalent and non-equivalent to reference images. Using the results of these experiments we have derived metrics for predicting visual equivalence (VEPs) and we have generalized these predictors so that it can predict equivalence for novel scenes. We have validated these VEPs in a confirmatory study and show that they can reliably predict observer's judgments of image differences and object appearance. We believe that visual equivalence is a novel approach to quantifying image quality that goes significantly beyond existing metrics by taking advantage of the limits of visual coding of object appearance and leveraging the fact that some kinds of image differences do not matter to human observers.

We are beginning to explore possible applications of VEPs. In related work[42] we have demonstrated how VEPs can be used to accelerate advanced image synthesis algorithms. In one case we have used a VEP to set limits on the refinement of illumination calculations in the Lightcuts[43,44] algorithm and have cut rendering time in half while maintaining image fidelity. In another case we have used a VEP to pre-warp the illumination maps used in Precomputed Radiance Transfer[45] so they can be more efficiently compressed using Wavelet techniques.

We are also working to extend the concepts introduced here and are currently exploring several research areas. In computer graphics we are investigating how to generalize VEPs to dynamic and non-rigid objects, moving viewpoints, and a larger classes of geometries, materials and illuminations. In digital imaging we are studying how the metrics can apply to optically captured images and how they might be used to identify new efficient yet high fidelity algorithms for image coding, compression, and display. Finally, in computational photography we are investigating how the metrics might be used to guide methods for image-based modeling and rendering, and to specify perceptually acceptable samplings and transformations of object light fields.

References:

1 WU, H.R. AND RAO, K.R. (Eds.) . 2005. *Digital Video Quality and Perceptual Coding*. Taylor and Francis Group. London

2 DALY, S. 1993. The visible differences predictor: an algorithm for the assessment of image fidelity. In Digital Images and Human Vision, A. B. Watson, Ed. MIT Press, 179–206.

3 LUBIN, J 1997. A visual discrimination model for imaging system design and evaluation. Ch. 10 in E. Peli (Ed.) Vision Models for Target Detection and Recognition. 245-283.

4 SHIRLEY, P. 2005. *Fundamentals of Computer Graphics*. A.K. Peters. Ltd., Wellesley, MA.

5 MANTIUK, R., DALY, S., MYSZKOWSKI, K., AND SEIDEL, H.P. 2005. Predicting visible differences in high dynamic range images - model and its calibration. In Proceedings of the SPIE:Human Vision and Electronic Imaging X, vol. 5666, 204–214.

6 GIBSON, J. 1979. The Ecological Approach to Visual Perception. Houghton Mifflin, Boston, MA.

7 ROCK, I. 1983. The Logic of Perception. MIT Press, Cambridge, MA.

8 TODD, J. T., AND MINGOLLA, E. 1983. Perception of surface curvature and direction of illumination from patterns of shading. Journal of Experimental Psychology: Human Perception and Performance 9, 4, 583–595.

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

10 CUTTING, J. E., AND MILLARD, R. T. 1984. Three gradients and the perception of flat and curved surfaces. Journal of Experimental Psychology: General 113, 2, 198–216.

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

12 KNILL, D. C., AND RICHARDS, W. 1996. Perception as Bayesian Inference. Cambridge University Press, New York, NY.

13 HUNTER, R.S. 1987. *The measurement of appearance*. 2nd Edition. Wiley Inter-Science, New York.

14 BECK, J. 1972. Surface Color Perception. Cornell University Press, Ithaca, NY.

15 O'DONNELL, F.X.D., and BILLMEYER, F.W. 1986. Psychometric scaling of gloss. ASTM special technical publication No. 914, 14-32. ASTM Press. Philadephia PA.

16 GILCHRIST, A., KOSSYFIDIS, C., BONATO, F., AGOSTINI, T., CATALIOTTI, J., LI, X., SPEHAR, B., ANNAN, V., AND ECONOMOU, E. 1999. An anchoring theory of lightness perception. Psychological Review 106, 4, 795–834.

17 ADELSON, E. H. 2000. Lightness perception and lightness illusions, 2nd ed. The MIT Press, Cambridge, MA, 339–351.

18 BRAINARD, D. H., AND MALONEY, L. T. 2004. Perception of color and material properties in complex scenes. Journal of Vision 4, 9, 2–4.

19 BLAKE, A., AND BULTHOFF, H. H. 1990. Does the brain know the physics of specular reflection? Nature 394, 165–168.

20 FERWERDA, J. A., PELLACINI, F., AND GREENBERG, D. P. 2001. A psychophysically-based model of surface gloss perception. In Proceedings of the SPIE: Human Vision and Electronic Imaging VI, vol. 4299, 291–301.

21 WESTLUND, H. B., AND MEYER, G. W. 2001. Applying appearance standards to light reflection models. In SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques, 501–51.

22 FLEMING, R. W., DROR, R. O., AND ADELSON, E. H. 2003. Real-world illumination and the perception of surface reflectance properties. Journal of Vision 3, 5, 347–368.

23 HARTUNG, B., AND KERSTEN, D. 2002. Distinguishing shiny from matte. Journal of Vision 2, 7, 551–551.

24 TODD, J. T., NORMAN, J. F., AND MINGOLLA, E. 2004. Lightness constancy in the presence of specular highlights. Psychological Science 15, 1, 33–39.

25 XIAO, B., AND BRAINARD, D. H. 2006. Color perception of 3D objects: constancy with respect to variation of surface gloss. In APGV '06: Proceedings of the 3rd symposium on Applied perception in graphics and visualization, 63–68.

26 ROBILOTTO, R., KHANG, B.-G., AND ZAIDI, Q. 2002. Sensory and physical determinants of perceived achromatic transparency. Journal of Vision 2, 5, 388–403.

27 FLEMING, R. W., AND BULTHOFF, H. H. 2005. Low-level image cues in the perception of translucent materials. ACM Transactions on Applied Perception 2, 3, 346–382.

28 DANA, K. J., VAN GINNEKEN, B., NAYAR, S. K., AND KOENDERINK, J. J. 1999. Reflectance and texture of real world surfaces. ACM Transactions on Graphics 18, 1, 1–34.

29 TE PAS, S. F., AND PONT, S. C. 2005. A comparison of material and illumination discrimination performance for real rough, real smooth and computer generated smooth spheres. In APGV'05: Proceedings of the 2nd symposium on Applied perception in graphics and visualization, 75–81.

30 HO, Y.-X., LANDY, M. S., AND MALONEY, L. T. 2006. How direction of illumination affects visually perceived surface roughness. Journal of Vision 6, 5, 634–648.

31 WANDELL, B. A. 1993. Color appearance: The effects of illumination and spatial pattern. In Proceedings of the National Academy of Sciences, vol. 90, 9778–9784.

32 DROR, R. O., WILLSKY, A. S., AND ADELSON, E. H. 2004. Statistical characterization of real-world illumination. Journal of Vision 4, 9, 821–837.

33 PONT, S. C., AND KOENDERINK, J. J. 2004. Surface illuminance flow. In 3DPVT '04: Proceedings of the 2nd international symposium on 3D data processing, visualization and transmission, 2–9.

34 KOENDERINK, J. J., VAN DOORN, A. J., KAPPERS, A. M. L., TE PAS, S. F., AND PONT, S. C. 2003. Illumination direction

35 KHANG, B.-G., KOENDERINK, J. J., AND KAPPERS, A. M. L. 2006. Perception of illumination direction in images of 3-D convex objects: Influence of surface materials and light fields. Perception 35, 5, 625–645.

36 OSTROVSKY, Y., CAVANAGH, P., AND SINHA, P. 2005. Perceiving illumination inconsistencies in scenes. Perception 34, 11, 1301–1314.

37 PERLIN, K. 2002. Improving noise. ACM Transactions on Graphics 21, 3 (July), 681–682.

38 WARD, G. J. 1992. Measuring and modeling anisotropic reflection. In SIGGRAPH '92: Proceedings of the 19th annual conference on Computer graphics and interactive techniques, vol. 26, 265–272.

39 TUMBLIN, J., HODGINS, J.K., GUENTER, B. 1999. Two methods for display of high contrast images. ACM Transactions on Graphics. 18(1), 56-94.

40 JOACHIMS, T. 1999. Making large-scale support vector machine learning practical. In Advances in kernel methods: support vector learning. MIT Press, Cambridge, MA, 169–184.

41 DEBEVEC, P. E., AND MALIK, J. 1997. Recovering high dynamic range radiance maps from photographs. In SIGGRAPH '97, 369–378.

42 [Reference suppressed for anonymity]

43 WALTER, B., FERNANDEZ, S., ARBREE, A., BALA, K., DONIKIAN, M., AND GREENBERG, D. P. 2005. Lightcuts: a scalable approach to illumination. ACM Transactions on Graphics 24, 3, 1098–1107.

44 WALTER, B., ARBREE, A., BALA, K., GREENBERG, D. P. 2006. Multidimensional Lightcuts. ACM Transactions on Graphics, 25, 3, 1081-1088

45 NG, R., RAMAMOORTHI, R., AND HANRAHAN, P. 2004. Triple product wavelet integrals for all-frequency relighting. ACM Transactions on Graphics, 477–487.