

# The Role of Visual Comparison Testing in Material Appearance Modeling

Gary Meyer

Department of Computer Science and Engineering, University of Minnesota, USA

---

## Abstract

*Material appearance modeling should involve end user testing. These tests can be visual comparison experiments where decisions made with real physical samples are compared with choices done using computer graphic simulations. This type of evaluation is necessary to determine whether simulations derived from material appearance models can be used to accomplish such tasks as the design of new products or the sale of manufactured goods. Lighting conditions, viewing circumstances, decision criteria, and user expertise can vary widely in each situation where the results of material appearance modeling are used to make a choice. An experiment is currently underway to perform such tests for automotive paint simulation.*

---

## 1. Introduction

Visual comparison testing should play a greater part in the development of material appearance models. While the physics of light measurement and reflection has been studied intensively to improve the simulation of how things look, the fact that computer graphic pictures are intended to be observed by people has been largely ignored. This oversight is surprising because the potential application of material appearance models in many fields depends upon whether people make the same decision using the computer graphic simulation as they would by looking at real physical samples. People are very good at performing the visual inspections that are necessary to, for example, buy a new garment or select an upholstery fabric, and these choices are typically made in irregular viewing and lighting conditions. These challenging factors make it problematic to assume that a two dimensional synthetic image of a scene will be good enough to allow the individual to make the same decisions. Visual comparison testing can confirm or deny whether this is true.

## 2. Real Material Appearance Design

Although the geometric design of an object is accomplished today using computer aided design tools, the appearance design of that same item is still produced using traditional manual techniques. Physical samples that demonstrate the available material appearances are obtained from suppliers, and these examples are used to create trend boards that illustrate

the types of color appearances that are to be part of the design. In the automotive industry, elaborate exterior and interior prototypes are fabricated to allow designers and decision makers to evaluate how the car will look under different lighting and viewing circumstances. Once the materials have been selected, qualification to be a supplier for the vehicle's paint or trim requires a manufacturer to bring physical samples to the design center so that they can be inspected under controlled lighting conditions. Each of these steps in the material appearance design process involves visual inspection of real material samples.

There are many reasons why visual inspection of actual physical samples still plays such a dominant role in the design of material appearance. While there are promising measurement technologies on the horizon, the most important reason that visual evaluation is still required is that there is no general way to measure all of the spatial, goniometric, and spectral reflectance properties of a material, and, even if there was such a measurement system, there is no universal material appearance metric that can be used to compare the measurement results. The effect of complex lighting and viewing conditions on material appearance also makes it easier to visually compare two samples while standing side by side with someone than it is to specify how the comparison is to be done separately (or by using the yet to be defined material appearance metric). Finally, the visual training of design professionals and the role of traditional design practice cannot be discounted, along with the fact that it always

seems easier and safer to trust a visual evaluation made with the examiner's own eyes.

Although it would appear that the manual material appearance design process is ripe for automation, experience shows that computer graphic material simulations are not readily accepted as part of this procedure. There are certainly technical reasons why this is the case. In the same way that the lack of general material appearance measurement systems makes visual comparison necessary in the traditional design methods, the absence of this measurement tool also means that the data necessary for realistic simulation of material appearance is often not available. The utility of simulations is further limited by the resolution and dynamic range of the display technologies available to display them. But the primary reason for the lack of acceptance is the absence of user testing to determine whether the material appearance simulations can be used to perform the same tasks that are accomplished in the manual process with real physical samples. Without this validation it is impossible to trust the decisions made with the simulation.

Given the important role that visual inspection plays in the current material appearance design process and considering the lack of acceptance that existing computer graphic simulations have received as solutions to this problem, it is time to consider the role that visual comparison testing can play in the development of material appearance simulations. To accomplish this objective it is necessary to identify the situations where people make material appearance judgments, determine the tasks that they are trying to accomplish by making these inspections, and develop visual comparison tests that show whether material appearance simulations can be used to accomplish these same tasks. This process will establish the type of problems that can be solved using material appearance simulations and the level of verisimilitude that is required in each case. These tests can also play an important role in the development of material appearance difference metrics that are expected to predict the same result as the visual comparison tests.

In the remainder of this brief position paper we will attempt to make these ideas more concrete. First we will summarize a range of material appearance observation tasks that are commonly performed. Then we will review early work in appearance evaluation that has been conducted within the field of computer graphics. Finally, we will briefly introduce an experiment in visual evaluation that is currently in progress.

### **3. Material Appearance Observation Tasks**

In addition to the design of new material appearances itself, visual inspection plays a key role in many commercially important situations where computer graphic simulation can be used in place of real physical samples. Examples include the purchase of apparel and furniture, and the design of a new

building or a future automobile. In each of these cases, the viewing and lighting conditions are different and the decision criteria are varied. The purpose of this section is to describe a few of these circumstances and, hopefully, to suggest the uniqueness and complexity of each situation.

Material appearance evaluation can take place at a large distance at an event like a runway fashion show or an automobile exhibition. The individual making the assessment is typically part of a big crowd watching the show, and the presentation often takes place on a stage under theatrical lighting. The clothing, cars, or other new products are usually viewed one at a time and often the decision to be made is whether to carry an item as a product in a store. Under these circumstances the opportunity to adjust the viewing and lighting conditions is limited, and the evaluation must be made according to how the exhibitor has staged the affair.

Making an individual purchase in a store allows material appearance evaluation to take place at much closer range. This judgment is often made in the company of one or two other people who may be asked to offer their opinions. The objects being considered can be held at arm's length in the hand and repositioned with respect to the light source. The illumination is standard commercial store lighting which ranges from tightly focused spotlights to wide area floodlights and large fluorescent fixtures. Typically, two or three items may be evaluated or compared at the same time. The decision to be made is whether the item should be acquired for the shopper, their family members, or their friends.

Industrial, interior, apparel, and architectural designers must evaluate the appearance of materials that they plan to include as part of their products and buildings. Examples of these visually important components include garment fabrics, wall paint, decorative tiles, upholstery cloth, plastic laminates, and metal trim. Initial decisions about what materials to employ are usually made by an individual in a design studio but opinions from other decision makers may be offered later during a critique. Small material samples are often used to make the selection and they may be spread out throughout the design space. Typical office lighting is used to illuminate the samples but large windows and skylights may be present in a design studio. Decision criteria include appearance matches between disparate materials and the overall color harmony between constituent parts.

Material appearance evaluation becomes more critical as a product design is finalized and once it enters the manufacturing process. Often the question being asked is whether the fabricated example matches a standard sample. To make these comparisons a controlled viewing environment, such as a light booth, is often employed. A light booth is a small enclosure with one open side and a fluorescent fixture in the ceiling. The acceptance or rejection decisions are made by a single individual or by two people standing side by side. The samples may be held in the hand, placed on the floor of the booth, or displayed on an easel. In addition to being per-

formed by designers, evaluations are also done by chemists and engineers who fabricate the materials.

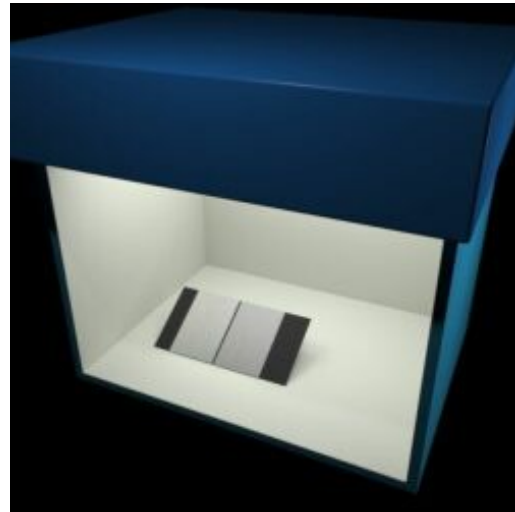
#### 4. Previous Work

There have been attempts in computer graphics to demonstrate, through visual experiments, that simulations can be used as substitutes for reality. Meyer et al. used the newly developed radiosity method in 1986 to set up a side by side comparison between a simulation and a real scene, and they found that most people could not identify which was which [MRC\*86]. Drago and Myskowski did a similar experiment with a more complex environment [DM01] but did not get as compelling of a result. McNamara extended these initial efforts dramatically [McN06]. She used judgments of lightness to demonstrate perceptual equivalence between the real and the simulated scene. Recently Borg et al. [BJTK12] developed an experiment to perform another “Turing Test” between a model and picture of the model viewed under similar restrictions. Their results appear to be similar to those that Meyer et al. obtained using a simple environment. Shortly after Meyer et al. did their experiment, Booth and Cowan [BBC\*87] pointed out that it was important to define a task that was to be performed when determining if the level of realism was appropriate. They showed that at high image resolution the amount of anti-aliasing does not have an effect on how well the subject can determine the orientation of a figure, while at lower resolution, anti-aliasing does improve the subject’s results. Ferwerda [Fer03] refers to this as functional realism and it is the approach that is being advocated in this workshop paper.

#### 5. Example Visual Comparison Test

To provide an example of the type of visual comparison test that should be performed to validate material appearance simulations, this section introduces an experiment that is currently in progress. A complete summary of the issues considered in designing this test will be presented so that the scope and limitations of this type of verification procedure can be better appreciated. However, the final methodology employed and the results obtained will not be discussed. When it is finished, the entire experiment and its outcome will be documented in a separate publication.

Automotive paint is the material appearance simulation problem involved in this particular visual comparison study. In the design and production of a new car, paint sample comparisons are made by a variety of different individuals, including paint company stylists who create new paint colors, automotive color and trim designers who pick the paint color for the car, and paint engineers who must figure out how to apply the paint and make sure that it lasts for the lifetime of the car. Visual inspection of car paint is complicated because it must take into account the variation of the color with reflectance angle (the travel of the color), the textural appearance and glints from the flakes in the paint, and even surface



**Figure 1:** Computer graphic rendering of a simple light booth. Sample is placed on a fixed easel under an overhead light source.

artifacts, such as orange peel, due to spray gun application. All of these factors make visual paint inspection a time consuming process that extends the length of the automotive design cycle.

There are a number of factors to consider when designing visual comparison tests to evaluate the material appearance simulation of automotive paint. Because of the complex reflectance properties of car paint, the most important consideration is the viewing circumstance to be employed. While simple light booths are often used when inspecting real paint samples (Figure 1), it may be better to employ a more specialized booth that keeps the lighting direction constant while allowing the viewing direction to be adjusted (Figure 2). The selection of the monitor on which to display the simulation is important, and must account for the dynamic range and calibration of the device. Finally, putting the viewing circumstances for the light booth and the monitor on an equal footing is an important consideration. Factors that must be taken into account include eye point position and viewing distance relative to the light booth and the monitor, and whether a binocular or monocular comparison is to be made.

It is necessary to design an experiment with human subjects to demonstrate that visual comparisons made using the car paint simulations are equivalent to those performed using real paint samples. If it can be shown that people make the same decisions when looking at computer graphic renderings, this result validates the material appearance simulation and suggests that color appearance decisions currently made with real samples could be replaced by tests performed with simulations. Given a set of real paint samples, a task must

be performed with the samples in the real light booth and then the task should be repeated using the computer graphic simulation displayed on the monitor. The task could be as simple as making comparisons between paint samples and determining whether they match. If the same results are obtained when the task is performed using the material appearance simulation then this result can be taken as evidence of the equivalence between the real and simulated paint samples.

## 6. Conclusions

If material appearance modeling is to achieve its potential to assist designers in the creation of new products and to help consumers in the purchase of the resulting commodities, the research must acknowledge the importance of the end user and visual comparison testing should be done to validate the effectiveness of the resulting simulations. For practical and cultural reasons, visual inspection of material samples still plays a major role in the color appearance design of new products. Visual comparison testing is therefore necessary for any computer aided color appearance design tool to gain acceptance in this environment. Shopping for objects that have important material appearance properties also involves careful visual examination of the product un-



**Figure 2:** Simulation of a light booth with an adjustable easel. The light source direction remains constant with respect to the sample as the viewing angle changes.

der complex lighting and viewing conditions. Visual comparison testing is again required to determine if people make the same material appearance decisions when using a simulation. Both of these situations have aspects of a user interface problem, and a successful outcome depends on closing the loop and performing tests with human subjects.

## 7. Acknowledgements

The automotive paint experiment is funded by USCAR and is being conducted with the assistance of colleagues at Ford (Chris Seubert), General Motors (Charles Fan), and Dassault Systemes (Curtis Evey, Jan Meseth, and Ryan Schnackenberg). Figure 2 was created by Ashley Nieman and Dave Toennies at Dassault Systemes.

## References

- [BBC\*87] BOOTH K. S., BRYDEN M. P., COWAN W. B., MORGAN M. F., PLANTE B. L.: On the parameters of human visual performance: An investigation of the benefits of antialiasing. *IEEE CG&A* 7, 9 (1987), 34–41.
- [BJTK12] BORG M., JOHANSEN S. S., THOMSEN D. L., KRAUS M.: Practical implementation of a graphics turing test. *Lecture Notes in Computer Science 7432* (2012), 305–313. (International Symposium on Visual Computing).
- [DM01] DRAGO F., MYSZKOWSKI K.: Validation proposal for global illumination and rendering techniques. *Computers & Graphics* 25, 3 (2001), 511–518.
- [Fer03] FERWERDA J.: Three varieties of realism in computer graphics. *Proceedings of the SPIE 5007* (2003), 290–297. (Human Vision and Electronic Imaging VIII Conference).
- [McN06] MCNAMARA A.: Exploring visual and automatic measures of perceptual fidelity in real and simulated imagery. *ACM TAP* 3, 3 (2006), 217–238.
- [MRC\*86] MEYER G. W., RUSHMEIER H. E., COHEN M. F., GREENBERG D. P., TORRANCE K. E.: An experimental evaluation of computer graphics imagery. *ACM TOG* 5, 1 (1986), 30–50.