

The Recognition of Material Properties for Vision and Video Communication MIT2000-02

Progress Report: July 1, 2002—December 31, 2002

Edward H. Adelson

Project Overview

How can we tell that an object is shiny or translucent or metallic by looking at it? Humans do this effortlessly, but machines cannot. Materials are important to humans and will be important to robots and other machine vision systems. For example, if a domestic cleaning robot finds something white on the kitchen floor, it needs to know whether it is a pile of sugar (use a vacuum cleaner), a smear of cream cheese (use a sponge), or a crumpled paper towel (grasp with fingers).

An object's appearance depends on its shape, on the optical properties of its surface (the reflectance), the surrounding distribution of light, and the viewing position. All of these causes are combined in a single image. In the case of a chrome-plated object, the image consists of a distorted picture of the world; thus the object looks different every time it is placed in a new setting. Somehow humans are good at determining the "chromeness" that is common to all the chrome images.

Progress Through December 2002

Psychophysics of Reflectance Estimation

In previous work we measured the accuracy with which humans match surface reflectance properties across variations in illumination. We found that subjects could match surface gloss (specular reflection) reliably and accurately across variations in illumination, as long as the statistics of the illumination are characteristic of the real world. Critically, however, performance declined considerably under illuminations such as single point light sources and random noise patterns, which do not have realistic statistics.

Our theory is that real-world illumination has certain well-conserved statistical properties, which lead to reliable diagnostic features in images of glossy materials. The key research question is to identify the statistical properties of real-world illumination that allow subjects to estimate surface reflectance properties.

In the last 6 months we have successfully identified a number of properties of real-world illumination that are important for surface reflectance estimation. Illumination in the real world has many regularities, ranging from simple properties (such as a pixel histogram that is skewed to low intensities); to higher order regularities (such as a ground plane and the presence of recognizable objects such as buildings and trees). Which regularities are important for surface reflectance estimation?

We have found that certain statistics of intermediate complexity lead to a good impression of glossiness. Histograms of pixel intensities in real-world illumination are heavily skewed to low intensities. This appears to be an important characteristic for surface reflectance estimation. Intuitively the skew results from the fact that most of the surfaces in the world reflect, rather than emit light. Although these surfaces structure the patterns of reflection by providing visible objects in the reflection, they are responsible for only a small proportion of the

incoming light. By contrast, direct light sources, such as candles and the sun, are generally rather compact and sparsely distributed, although they provide the majority of the incident light. This finding is consistent with previous research (S. Nishida and M. Shinya. Use of image-based information in judgments of surface-reflectance properties. *J. Opt. Soc. Am. A*, **15**:2951–2965, 1998), which has found that image pixel histograms play an important role in surface reflectance matching.

We found that a real-world illumination that is robbed of its characteristic skew, makes glossy surfaces look dull, matte and unrealistic. Conversely, random noise patterns can lead to vivid impressions of gloss when endowed with skew. This suggests that the pixel histogram skew, or sparseness of illumination, is an important characteristic of real-world illumination for surface reflectance estimation. We have also found that endowing random noise patterns with realistic wavelet statistics (i.e. those that are characteristic of real-world illumination), causes the illumination to yield glossy-looking surfaces. Examples of this are shown in Figure 1. Wavelet statistics capture local information about scale and orientation, which is important for representing the clumps, blobs and edges that are found in real-world illumination.

This research has been accepted (with minor revisions) for publication in the Journal of Vision.

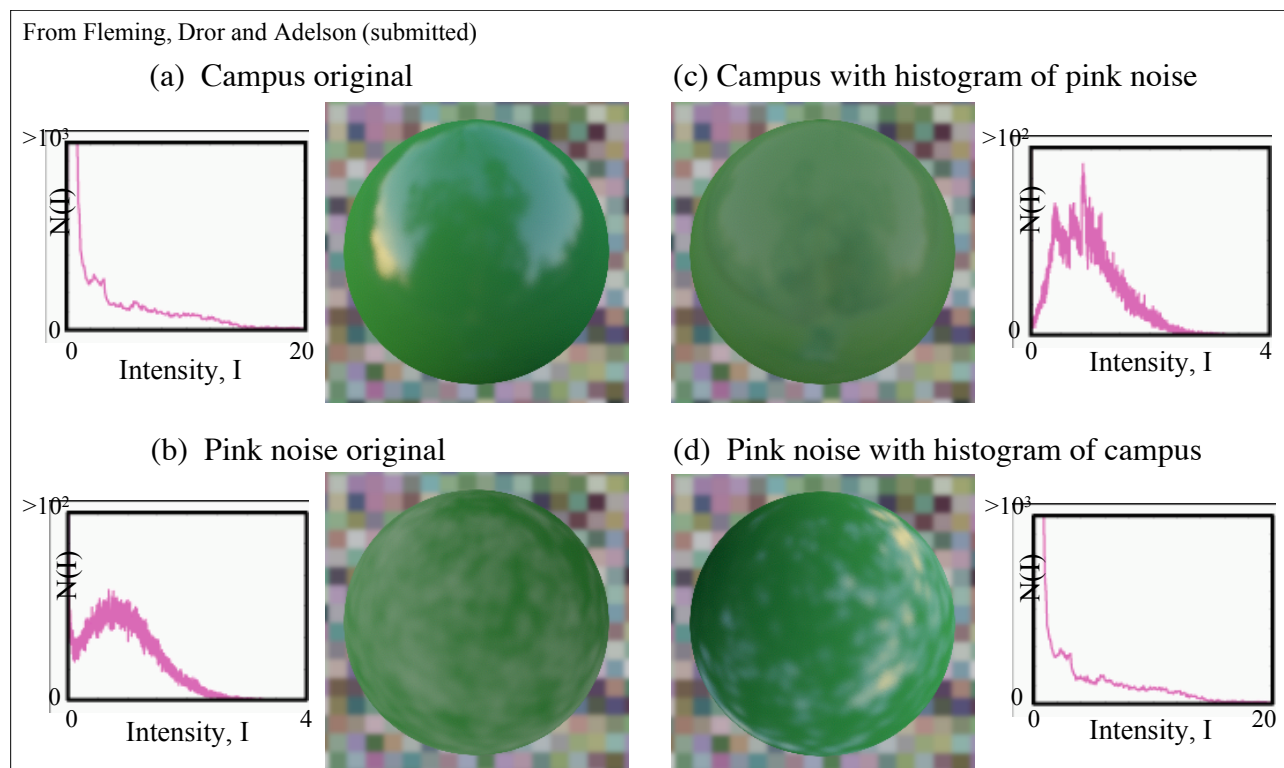


Figure 1 - Spheres rendered under illuminations with modified pixel histograms. The sphere in (a) was rendered under the original Campus illumination, which has a heavily skewed pixel histogram. The sphere in (b) was rendered under pink noise illumination with a truncated gaussian pixel histogram. These illuminations were then modified using histogram matching. The sphere in (c) was rendered under modified Campus illumination with an approximately gaussian histogram derived from (b). The sphere in (d) was rendered under modified noise illumination with a histogram derived from (a). Note the difference in scale on the pixel histogram plots; the original Campus is considerably more skewed than the original Noise.

Recovering Intrinsic Images from a Single Image

Previously, we had developed a system for decomposing a single image into two images, one representing the shading of the scene and a second image containing the reflectance of each point in the scene. This system works by calculating horizontal and vertical derivatives of the image, then classifying each derivative as being caused by either shading or by a reflectance change. This classification is made using local gray-scale and color image information.

In many cases, the local evidence that we use is not sufficient to correctly classify derivatives. Figure 1 shows an example of this situation. When compared to the example shading and reflectance change, the center of the mouth is equally well classified with either label. However, the corners of the mouth can be classified as being caused by a reflectance change with little ambiguity. Since the derivatives in the corner of the mouth and the center all lay on the same image contour, they should have the same classification. In the last six months, we have refined our techniques for propagating information from the corners of the mouth, where the correct classification is clear, into the center, where the local evidence is ambiguous.

We propagate information by treating the classification of each derivative as a random variable in a Markov Random Field, or MRF. In this model, the label of each derivative only directly depends on the labels of the neighboring derivatives. We set up the MRF so that derivatives along an image contour are strongly influenced to have the same label as their neighbors. The most probable label for each derivative is found using the Generalized Belief Propagation algorithm. As part of this work, we created and verified an efficient C++ implementation of the Generalized Belief Propagation algorithm.

This work was presented at the Neural Information Processing Systems Conference.

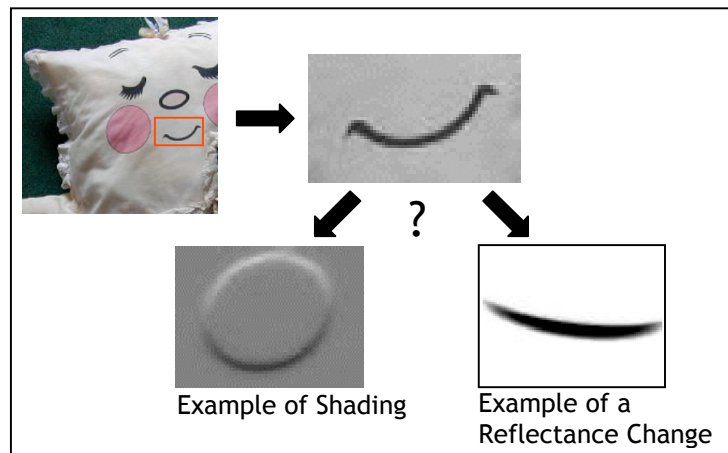


Figure 2- Locally, the center of the mouth is equally well explained as shading or a reflectance change. Local evidence cannot be used alone to classify it

Research Plan for the Next Six Months

Psychophysics of Reflectance Estimation

In direct collaboration with Shin'ya Nishida and colleagues we will replicate and extend our experiments with surface shapes other than spheres. We also hope to perform experiments with surfaces that have empirically measured BRDFs, rather than parametric approximations.

Recovering Intrinsic Images

We will investigate decompositions involving different types of visual characteristics, such as transparency or material properties. We are also interested in extending the technologies developed in this project into an interactive computer graphics tools for synthetically relighting images.