

MIT2000-02

The Recognition of Material Properties for Vision and Video Communication

Edward H. Adelson

When a future domestic robot cleans the kitchen table, it should know the difference between spilled milk, smeared cream cheese, and scattered sugar. All are white materials, but all must be treated differently. We can divide visual problems into those dealing with "things" (objects) or with "stuff" (materials). Stuff is half of vision, but has received relatively little attention in the vision community. The importance of material appearance for humans is evidenced by the large amount of effort expended in computer graphics toward getting materials to "look right." Understanding material perception would be useful in manipulating image data. For instance, when transmitting faces, greasy, blemished, or wrinkled skin could be recognized and modified for better appearance. In addition, advanced video systems use computer graphics models to efficiently code images. Rather than sending a the pixels of a silver goblet, for example, one can send a 3-D model of the goblet. However, one must recognize the material as well as the object. Recently, there has been an increasing interest in material perception both in human psychophysics (e.g., Nishida & Shinya, *JOSA A*, 1998) and in machine vision (e.g., Wolff, Nayar, and Oren, *IJCV*, 1998); in addition, there has been progress in understanding surface lightness (e.g., Adelson, 1999, in *The New Cognitive Neurosciences*) and texture (e.g., DeBonet, *SIGGRAPH*, 1997). We propose to combine techniques from these areas to develop a new understanding of material perception. We have evidence that material perception requires joint information about structure and statistics, i.e., that is involves a blend of the algorithms from object recognition and texture analysis. We will characterize the conditions under which humans can recognize materials, and will develop machine-vision models to emulate this ability.

This research requires equipment for the digital capture of images and image sequences, and support for one graduate student.