Comparing Real & Synthetic Scenes using Human Judgements of Lightness

Ann McNamara Alan Chalmers
Department of Computer Science

Tom Troscianko Iain Gilchrist
Department of Experimental Psychology
University of Bristol Bristol
mcnamara@cs.bris.ac.uk

Abstract. Increased application of computer graphics in areas which demand high levels of realism has made it necessary to examine the manner in which images are evaluated and validated. In this paper, we explore the need for including the human observer in any process which attempts to quantify the level of realism achieved by the rendering process, from measurement to display. We introduce a framework for measuring the perceptual equivalence (from a lightness perception point of view) between a real scene and a computer simulation of the same scene. Because this framework is based on psychophysical experiments, results are produced through study of vision from a *human* rather than a *machine* vision point of view. This framework can then be used to evaluate, validate and compare rendering techniques.

1 Introduction

The aim of realistic image synthesis is the creation of accurate, high quality imagery which faithfully represents a physical environment, the ultimate goal being to create images which are perceptually indistinguishable from an actual scene. Rendering systems are now capable of accurately simulating the distribution of light in an environment. However, physical accuracy does not ensure that the displayed images will have authentic visual appearance. Reliable image quality assessments are necessary for the evaluation of realistic images synthesis algorithms. Typically the quality of an image synthesis method is evaluated using numerical techniques which attempt to quantify fidelity using image to image comparisons (often comparisons are made with a photograph of the scene that the image is intended to depict).

Several image quality metrics have been developed whose goals are to predict the *visible* differences between a pair of images. It is well established that simple approaches, such as mean squared error (MSE), do not provide meaningful measures of image fidelity, more sophisticated techniques are necessary. As image quality assessments should correspond to assessments made by humans, a better understanding of features of the Human Visual System (HVS) should lead to more effective comparisons, which in turn will steer image synthesis algorithms to produce more realistic, reliable images. Any feature of an image not visible to a human is not worth computing. Results from psychophysical experiments can reveal limitations of the HVS. However, problems arise when trying to incorporate such results into computer graphics algorithms. This is due to the fact that, often, experiments are designed to explore a single dimension of the HVS at a time under laboratory conditions. The HVS comprises

B. Péroche et al. (eds.), *Rendering Techniques 2000*

many complex mechanisms, which rather than function independently, often work in conjunction with each other, making it more sensible to examine the HVS as a whole. Rather than attempting to reuse results from previous psychophysical experiments, new experiments are needed which examine the complex response HVS as a *whole* instead of than trying to isolate features for individual investigations. In this work we study the ability of the HVS to perceive albedo and the impact of rendering quality on *this task*. Rather than deal with atomic aspects of perception, this study examines a complete task in a more realistic setting.

Human judgements of lightness are compared in real scenes, and synthetic images. Correspondence between these judgements is then used as an indication of the fidelity of the synthetic image.

1.1 Lightness Perception

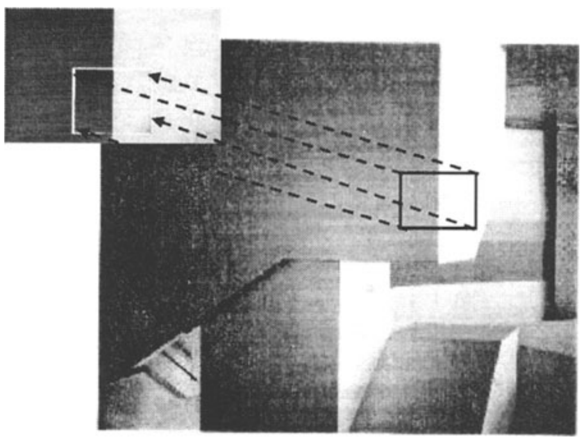

Fig. 1. Importance of depth perception for lightness constancy

Lightness is apparent reflectance, brightness is apparent intensity of the illuminant. Reflectance is the proportion of light falling on an object that is reflected to the eye of the observer. Reflectance (albedo) is constant, the perception of lightness depends of reflectance [1]. Gilchrist [8] showed that the perception of the degree of "lightness" of a surface patch (i.e. whether it is white, gray or black) is greatly affected by the perceived distance and orientation of the surface in question, as well as the perceived illumination falling on the surface - where the latter were experimentally manipulated through a variety of cues such as occlusion, or perspective.

Perception of the lightness of patches varying in reflectance may thus be a suitable candidate for the choice of visual task. It is simple to perform, and it is known that lightness constancy depends on the successful perception of lighting and the 3D structure of a scene, for example figure 1. When viewed in isolation the patches on the top left hand corner appear to be of different luminance. However, when examined in the context of the entire scene, it can be seen that the patches have been cut from the edge of the stairwell, and is perceived as an edge where the entire stairwell has the same luminance. Eliminating the depth cues means the patches are perceived as different, demonstrating