

1

Introduction

*‘Where shall I begin, please your Majesty?’ he asked.
‘Begin at the beginning,’ the King said, gravely,
‘and go on till you come to the end: then stop.’*

Lewis Carroll

1.1 MOTIVATION

Humans are highly visual creatures. Evolution has invested a large part of our neurological resources in visual perception. We are experts at grasping visual environments in a fraction of a second and rely on visual information for many of our day-to-day activities. It is not surprising that, as our world is becoming more digital every day, digital images and digital video are becoming ubiquitous.

In light of this development, optimizing the performance of digital imaging systems with respect to the capture, display, storage and transmission of visual information is one of the most important challenges in this domain. Video compression schemes should reduce the visibility of the introduced artifacts, watermarking schemes should hide information more effectively in images, printers should use the best half-toning patterns, and so on. In all these applications, the limitations of the human visual system (HVS) can be exploited to maximize the visual quality of the output. To do this, it is necessary to build computational models of the HVS and integrate them in tools for perceptual quality assessment.

The need for accurate vision models and quality metrics has been increasing as the borderline between analog and digital processing of visual information is moving closer to the consumer. This is particularly evident in the field of television. While traditional analog systems still represent the majority of television sets today, production studios, broadcasters and network providers have been installing digital video equipment at an ever-increasing rate. Digital satellite and cable services have been available for quite some time, and terrestrial digital TV broadcast has been introduced in a number of locations around the world. A similar development can be observed in photography, where digital cameras have become hugely popular.

The advent of digital imaging systems has exposed the limitations of the techniques traditionally used for quality assessment and control. For conventional analog systems there are well-established performance standards. They rely on special test signals and measurement procedures to determine signal parameters that can be related to perceived quality with relatively high accuracy. While these parameters are still useful today, their connection with perceived quality has become much more tenuous. Because of compression, digital imaging systems exhibit artifacts that are fundamentally different from analog systems. The amount and visibility of these distortions strongly depend on the actual image content. Therefore, traditional measurements are inadequate for the evaluation of these artifacts.

Given these limitations, researchers have had to resort to subjective viewing experiments in order to obtain reliable ratings for the quality of digital images or video. While these tests are the best way to measure 'true' perceived quality, they are complex, time-consuming and consequently expensive. Hence, they are often impractical or not feasible at all, for example when real-time online quality monitoring of several video channels is desired.

Looking for faster alternatives, the designers of digital imaging systems have turned to simple error measures such as mean squared error (MSE) or peak signal-to-noise ratio (PSNR), suggesting that they would be equally valid. However, these simple measures operate solely on a pixel-by-pixel basis and neglect the important influence of image content and viewing conditions on the actual visibility of artifacts. Therefore, their predictions often do not agree well with actual perceived quality.

These problems have prompted the intensified study of vision models and visual quality metrics in recent years. Approaches based on HVS-models are slowly replacing classical schemes, in which the quality metric consists of an MSE- or PSNR-measure. The quality improvement that can be achieved

using an HVS-based approach instead is significant and applies to a large variety of image processing applications. However, the human visual system is extremely complex, and many of its properties are not well understood even today. Significant advancements of the current state of the art will require an in-depth understanding of human vision for the design of reliable models.

The purpose of this book is to provide an introduction to vision modeling in the framework of video quality assessment. We will discuss the design of models and metrics and show examples of their utilization. The models presented are quite general and may be useful in a variety of image and video processing applications.

1.2 OUTLINE

Chapter 2 gives an overview of the human visual system. It looks at the anatomy and physiology of its components, explaining the processing of visual information in the brain together with the resulting perceptual phenomena.

Chapter 3 outlines the main aspects of visual quality with a special focus on digital video. It briefly introduces video coding techniques and explores the effects that lossy compression or transmission errors have on quality. We take a closer look at factors that can influence subjective quality and describe procedures for its measurement. Then we review the history and state of the art of video quality metrics and discuss the evaluation of their prediction performance.

Chapter 4 presents tools for vision modeling and quality measurement. The first is a unique measure of isotropic local contrast based on analytic directional filters. It agrees well with perceived contrast and is used later in conjunction with quality assessment. The second tool is a perceptual distortion metric (PDM) for the evaluation of video quality. It is based on a model of the human visual system that takes into account color perception, the multi-channel architecture of temporal and spatial mechanisms, spatio-temporal contrast sensitivity, pattern masking and channel interactions.

Chapter 5 is devoted to the evaluation of the prediction performance of the PDM as well as a comparison with competing metrics. This is achieved with the help of extensive data from subjective experiments. Furthermore, the design choices for the different components of the PDM are analyzed with respect to their influence on prediction performance.

Chapter 6 investigates a number of extensions of the perceptual distortion metric. These include modifications of the PDM for the prediction of perceived blocking distortions and for the support of object segmentation. Furthermore, attributes of image appeal are integrated in the PDM in the form of sharpness and colorfulness ratings derived from the video. Additional data from subjective experiments are used in each case for the evaluation of prediction performance.

Finally, Chapter 7 concludes the book with an outlook on promising developments in the field of video quality assessment.