Course Name : Digital Video Signal Processing

Department : Electrical Engineering, IIT Kanpur

Instructor : Prof. Sumana Gupta

Next

The Lecture Contains:

- Introduction
- Topics to be covered
- Video Formation, Perception and Representation
- Color Perception and Specification
- Light and Color
- Human Perception of Color
- The Trichromatic Theory of Color Mixture

Previous Next

Digital Video Signal Processing

INTRODUCTION

From the recent developments in multimedia representation and communication, it has become clear that all aspects of media starting from, representation to transmission, processing to retrieval, and from studio to home are going **"Digital"**. The significant advances in digital multimedia compression and communication algorithms have made it possible to deliver high quality video at relatively low bit rates in today's networks. The advancements in VLSI technologies have enabled sophisticated software to be implemented in a cost effective manner. Lastly, the establishment of several international standards by ISO/MPEG and ITU-T laid the common groundwork for different vendors and content providers.

As multimedia becomes more pervasive, the boundaries between video, graphics, computer vision, multimedia database and computer networking start to blur. This makes video processing an important and exciting field with input from many disciplines. Presently video processing lies at the core of multimedia. Among the many technologies involved, video coding and its standardization are the key enablers of these developments.

In this web based course we intend to cover the fundamental theory and techniques for digital video processing with a focus on video coding for communications. The topics we select are a balance between providing a solid theoretical foundation and presenting complex system issues in real video systems.

Previous Next

Module 1: Digital Video Signal Processing

Lecture 1: Introduction to video formation, Perception and representation

Topics to be covered:

1. Broad overview of video technology, from analog color TV systems to digital video is define of video signal, Video capture & display; analog video raster.

2. Fourier analysis of video signals and frequency response of human visual system (HVS)

3. **Video sampling:** Basics of Lattice Theory; Sampling over lattices; sampling of video signals in 2D and 3D; Aliasing spatial & temporal. Filtering operations in cameras & display devices.

4. **Video sampling rate conversion:** Conversion of signals sampled on different lattices, deinterlacing; conversion between PAL & NTSC signals; motion adaptive interpolation.

5. **Video modeling:** Camera Model, Illumination model; Object Model: Shape motion; Scene Model; 2D motion models

6. 2D Motion Estimation-Optical flow; Pixel based motion estimation, Block matching also; Deformable block matching; Region based motion estimation, Multiresolution based motion estimation.

7. Waveform based video coding-Block based transform coding, Predictive coding, Video coding using temporal prediction and transform coding.

8. **Scalable video coding: Basic modes of scalability:** Quality, Spatial, Temporal, Frequency embedded coding.

9. **Video coding standards:** MPEG 2, MPEG 4, H.264.

Previous Next

Video Formation, Perception and Representation:

In this section we describe, what a video signal is, how it is captured and perceived, how it is stored and transmitted; and what are the important parameters that determine the quality and bandwidth (In turn determines the data rate) of the signal. For proper understanding of video signals we begin with the
underlying physics of color perception and specification.

Color Perception and Specification

A video signal is a sequence of 2D images, projected from a dynamic 3D scene onto the image plane of a video camera. The color value at any point in a video frame records the emitted or reflected light at a particular 3-D point in the observed scene. To understand what the color value means physically, we review the basics of light physics, and describe the attributes that characterize light and its color. We also consider human color perception and different ways to specify a color signal.

Light and Color

It is well known that light consists of an electromagnetic wave with wavelength in the range of 380nm to
780nm in which the human eye is sensitive. Energy of light is measured by **'Flux'** with a unit of watt, which is the rate at which energy is emitted. A light source usually can emit energy in a range of wavelengths, and its intensity can vary in both space and time.

The **"Radiant intensity"** of a light, which is directly related to our perception of the brightness of the light, is defined as the flux radiated into a unit solid angle in a particular direction, and is measured in watt/solid angle. We use the notation $C(\vec{X}, t, \lambda)$ to represent the radiant intensity distribution of light, which specifies the light intensity at wavelength λ , spatial location $\vec{X} = (x, y, z)$ and time t.

The perceived color of light depends on its spectral content (i.e. its wavelength composition) For example; a light that has its energy concentrated near 700nm appears **"Red"**. A light that has equal energy across the entire visible band appears **"White"**. In general, a light that has a very narrow bandwidth is referred to as **"Spectral color"**. On the other hand, a white light is said to be **"Achromatic"**.

Previous Next

There are two types of light sources:

Illuminating sources emit an electromagnetic wave. Such sources include the sunlight bulbs, TV monitors & so on. The perceived color of an illuminating light source depends on the wavelength range in which it emits energy. Illuminating light follows an additive rule: i.e., the perceived color of several mixed illuminating light sources depends on the sum of the spectra of all the light sources.

For example, Combining red, Green and Blue lights in the right proportions creates the color White.

Reflecting sources reflect an incident wave. Reflecting light sources that reflect an incident light could
itself be a reflected light. When a light beam hits an object, the energy in a certain wavelength range is
absorbed, while the rest is reflected. The color of a reflected light depends on the spectral content of the incident light and the wavelength range that is absorbed. The most notable reflecting light sources are color dyes and paints. A reflecting light source follows a **"Subtractive rule"** i.e. the perceived color of several mixed reflecting light sources depend on the remaining, unabsorbed wavelengths.

For example, if the incident light is white, a dye that absorbs the wavelength near 700nm (Red) appears as cyan. In this sense we say that cyan is the complement of red (i.e. white minus red). Similarly magenta and yellow are complements of green and blue. Mixing cyan, magenta and yellow dyes produces black, which absorbs the entire visible spectrum.

Previous Next

Human Perception of Color

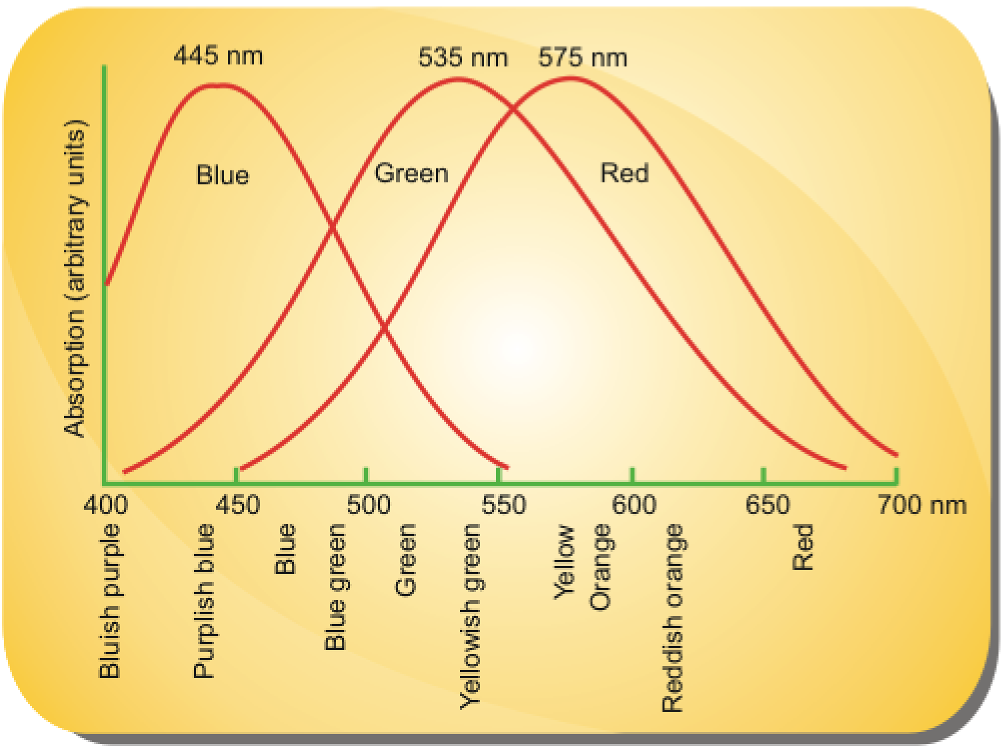
The perception of light by a human being starts with the photoreceptors located in the retina (Rear surface of the interior of the eyeball).There are two types of receptors:

Cones: Which function under bright light & can perceive color tone, and

Rods: Which work under low ambient light and can only extract luminance information.

Visual information from retina is passed via optic nerve fibers to the brain area called the visual cortex where visual processing and understanding is a accomplished.

There are 3 types of cones, which have overlapping pass-bands in the visible spectrum with peaks at red (Near 570 nm) green (Near 535 nm) and blue (Near 445 nm) wavelengths as shown:



(Figure 1)

Previous Next

The responses of these receptors to an incoming light distribution $C(\lambda)$ can be described by the following equation:

$$C_i = \int_\lambda c(\lambda)a_i(\lambda)d\lambda, \qquad i = r, g, b$$

Where $a_r(\lambda)$, $a_g(\lambda)$ and $a_b(\lambda)$ are referred to as the relative absorption functions of the red, green & blue cones respectively.
The combination of these 3 types of cone receptors enables a human being to perceive color. This implies that the perceived color depends only on the three numbers C_r, C_g, C_b , rather than on the complete spectrum $C(\lambda)$.

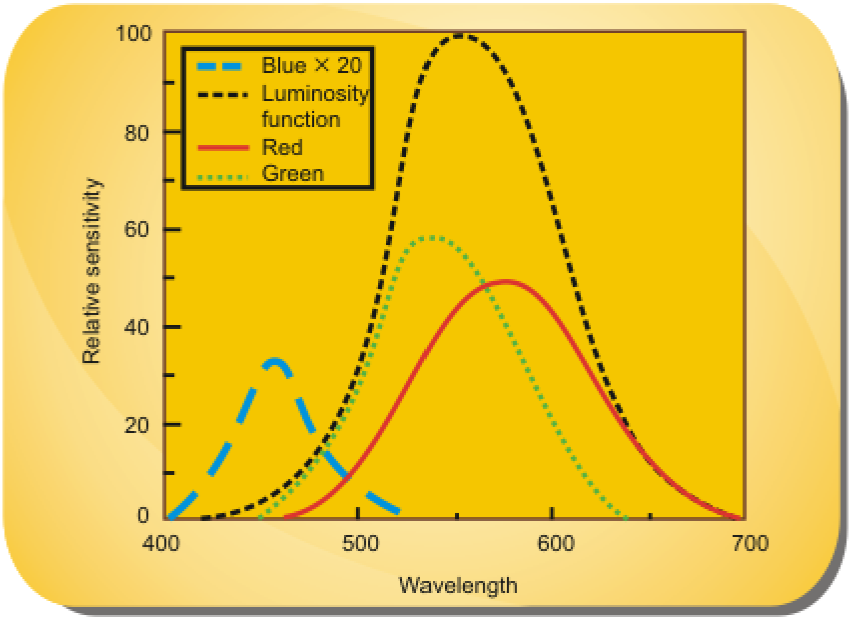
This is known **As tri-receptor theory of color vision**. This is shown in fig (2) .
There are two attributes that describe the color sensation of a human being: luminance and chrominance.
Chrominance in turn is characterized by attributes such as, hue and saturation .Hue specifies the color
tone which depends on the peak wavelength of the light. Saturation describes how pure the color is and
depends on the spread or bandwidth of the light spectrum. We will use the word “color” to refer to both luminance & chrominance attributes of light.
Experiments have shown that there exists a secondary processing stage in HVS. This stage converts the 3 color values obtained by the cones into one value that is proportional to the luminance and two other values that are responsible for the perception of chrominance. This is known as the **Opponent color model of HVS.**

It has also been found that the same amount of energy produces different sensation of brightness at different wavelengths. This wavelength dependent variation of brightness sensation is characterized by **“Relative luminous efficiency function $a_y(\lambda)$” of the HVS**. This is essentially the sum of the frequency responses of all 3 types of cones. We can see from the plot shown in Figure 2 below, that the green wavelength contributes the most to the perceived brightness, the red wavelength the second-most and the blue the least. The luminance (Denoted by Y) is related to the incoming light spectrum by

$$Y = \int_\lambda C(\lambda)a_y(\lambda)d\lambda$$

In the preceding equation we omitted the time and space variables, since we are only concerned with the perceived color or luminance at a fixed spatial and temporal location.

Previous Next



(Figure 2)

The Trichromatic Theory of Color Mixture:

An important finding in color physics is that most colors can be produced by mixing three properly chosen primary colors. This is known as trichromatic theory of color mixture first demonstrated by Maxwell in 1855.

Let $C_k, K = 1,2,3$ represent the colors of three primary color sources and C a given color. There the theory essentially says that

$C = \sum_{K=1,2,3} T_K C_K$ where $T_K's$ are the amounts of the 3 primary colors required to match the color C. The $T_{K's}$ are known as "tristimulus values". In general, some of the $T_K's$ can be negative. Assuming only T_1 is negative; this means that one cannot match color C by mixing $C_1, C2, C3$. However, one can match $C + |T_1| C_1$ with $T_2 C_2 + T_3 C_3$.