P17

The ideal spectral response curves of colored photodiodes

Guy Meynants and Bart Dierickx
Imec vzw, Kapeldreef 75, B-3001 Leuven, Belgium

1. Introduction

It is not possible to render exactly all colors of a scene without making color errors. Hue and/or saturation errors will appear for certain colors. In this article, we present a method to acquire an image with an optimal approximation for the color information, so that it can be displayed without any further color transformation on a display. This article is about color image acquisition, but there is a very close relation to the way a color image is reproduced.

The most important parameters for the acquisition of a color image are the spectral response curves of the colored photodiodes (or pixels). These are the products of the color filter transmittances and the spectral response of the photodiodes. Normally, these response curves are considered as fixed and unchangeable. Imperfections in the response curve are corrected by data processing. The RGB data are corrected by a color matrix that is obtained from the response of the sensor to a standard color chart. The color matrix is calculated by minimizing the color error [1].

In this article, we look towards the spectral response curves from another point of view. We assume that we can adapt the color filters or the spectral response curves to our needs. We will try to find the curves that are optimal for a human observer. When these response curves are used, a color image is acquired that can be displayed without any color transformation on a color display. Unlike the spectral response curves, we consider the display characteristics as unchangeable. They will play an important role in finding the optimal spectral response curves.

2. Motivation and definition of a 'critical human observer'

A critical human observer wants that at least the colors of the rainbow are correctly reproduced. These are the monochromatic colors. They only contain one wavelength, but if these colors are correctly reproduced, mixtures of these colors will also correctly be rendered.

Color rendering on a display is done by mixing together light of three standardized phosphors. Each of these phosphors generates light of a certain color (red, green or blue). But these light sources are not purely monochromatic. They are contaminated with other wavelengths. Furthermore, only mixtures of these 3 colors can be rendered. Some colors are not part of this color gamut. This can be seen on the CIE chromaticity diagram (figure 1). This represents the colors in the CIE x-y color space. The elliptic curve represents the monochromatic colors. The triangle connects the 3 EBU primary colors. Displays can only generate colors inside this triangle, by linear mixing of these 3 primaries. Colors outside this gamut cannot be generated.

The CIE chromaticity diagram is normally only used for color reproduction, and not for analysis of color acquisition. The acquisition is fully described by the spectral response curves of the three colored photodiodes. They indicate for each monochromatic color what the output of the pixel will be, and how the color mixing ratio will be. But if we want that this mixing ratio is correct, without any transformation or correction step, we will need the CIE diagram in our analysis.

3. Construction of the response curves

We know from the CIE diagram that it is impossible to generate the monochromatic colors exactly, as all monochromatic colors are located outside the triangle of primaries. The best solution is to approximate the monochromatic colors by the colors on the outer border of the triangle. If we mix the monochromatic colors with the reference white, we obtain a point on one of the edges of the triangle, that is the best approximation of the monochromatic color. The color is less saturated, but with the same hue. We propose to render this less

1

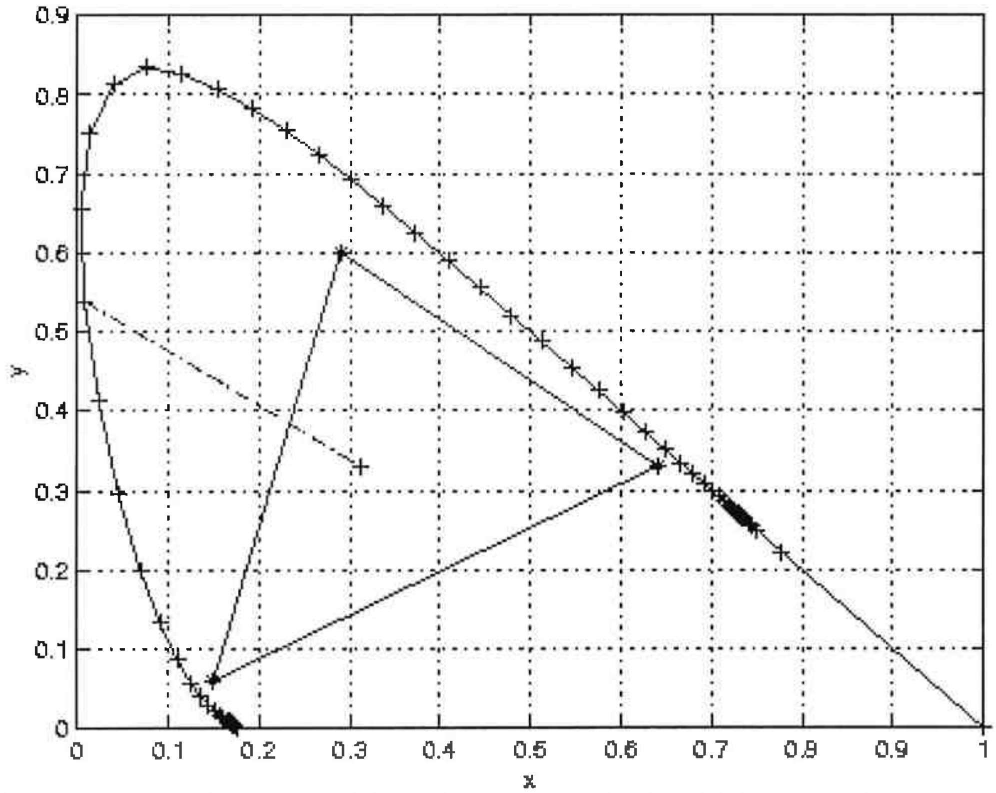

Figure 1: CIE chromaticity diagram, location of EBU primary colors (*) and reference white (+), and method to determine the blue/green mixing ratio for reconstruction of a monochromatic color of 500 nm.

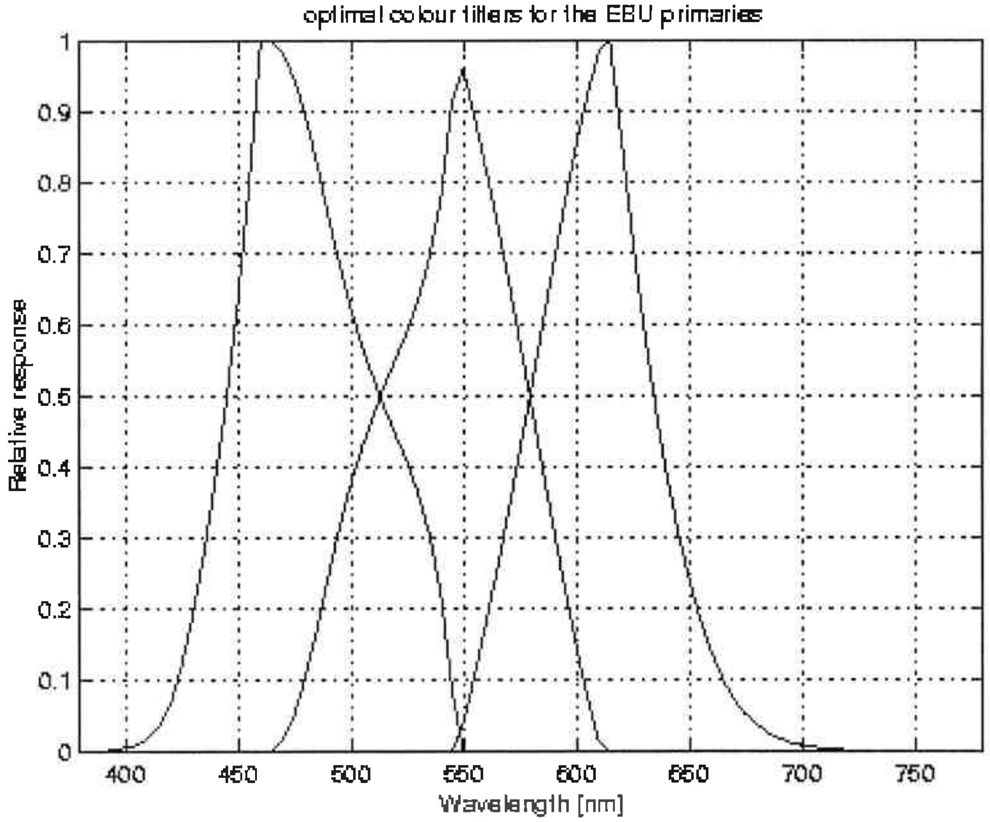

Figure 2: Optimal response curves for colored photodiodes

saturated color when the sensor is illuminated with this monochromatic color. So, the sensor should output the mixing ratio needed to render this color.

The dashed line on figure 1 illustrates how it works. The dashed line connects monochromatic light of 500 nm with the reference white. Where this line crosses the triangle, is the less saturated color. This color should be reproduced by the display. The mixing ratio between the primaries can be determined from this point. It can be read from the position of the point along the line between the blue and green primary. The response curve of the colored pixels should directly output this mixing ratio. If this is the case, no further color transformation is necessary, and the data obtained from the sensor can directly be used to display the image.

This method can be repeated for the other monochromatic wavelengths. We can reconstruct the ideal response curves for the colored photodiodes by calculating the mixing ratios needed to reconstruct each wavelength using the above procedure. In this way, the filter characteristics of figure 2 are obtained.

Wavelengths below blue (453.9 nm) or above red (700 nm) can not be reconstructed by this procedure. These colors cannot be rendered because they are located outside the color gamut: there are no 2 primaries available to mix. There are two options. The blue/red filters can pass all light below/above this wavelengths (below 453.9 nm for blue & above 700 nm for red). Or the response curve can follow the sensitivity of the human eye (as was done in figure 2). In any case, the other two filters must have zero response.

4. Interpretation of the response curves

The response curves of figure 2 have the following properties:

1) The maximal response of a filter is at the wavelength of its primary. The two other filters have zero response at this wavelength. Indeed, the primary monochromatic color is best reproduced by the phosphor that corresponds to it.
2) At any wavelength, one of the filters has a zero response. The responses of the two other filters record a less saturated approximation of monochromatic light of this wavelength.
3) Within the range between the blue (453.9 nm) and the red (700 nm) light, the sum of the overal spectral responses of the pixels must be one. This is a normalization step. Of course, in real life, white balancing is necessary: the different color channels will each be amplified with a different gain to cope with variations in the global illumination. Then, the sum is not one any more.
4) The curves have a rather sharp peak (especially the green curve). This is in contrary to most real-life color filters, which have a smooth peak.

For other display media (e.g. LCD displays, color printers, etc.), the ideal response curves will be different. However, the principle of constructing the curves remains the same.

5. Conclusions

In this article, we presented the 'ideal' spectral response curves for the colored pixels. When the pixels would have such color response, the color information in an image would be acquired directly, and no color correction would be necessary to render the image.

6. Reference

[1] Henry Kang, "Color Technology for Electronic Imaging Devices", SPIE press, 1996