# Video and Colorspaces

Hans Verkuil
Cisco Systems Norway

# struct v4l2_pix_format (3.18)

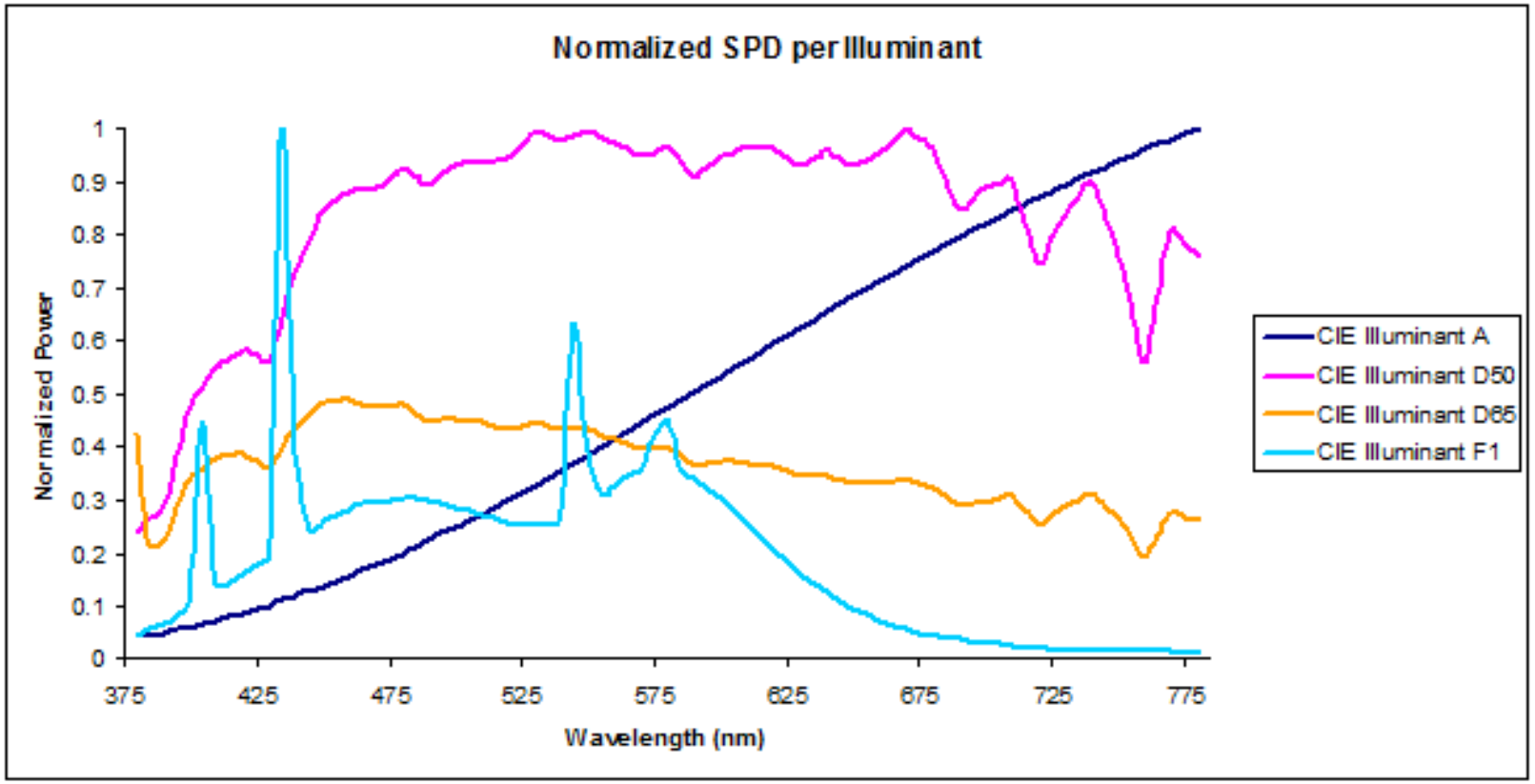
include/uapi/linux/videodev2.h:

```
struct v4l2_pix_format {
        __u32 width;
        __u32 height;
        __u32 pixelformat;
        __u32 field;            /* enum v4l2_field */
        __u32 bytesperline;     /* for padding, zero if unused */
        __u32 sizeimage;
        __u32 colorspace;       /* enum v4l2_colorspace */
        __u32 priv;             /* private data, depends on pixelformat */
        __u32 flags;            /* format flags (V4L2_PIX_FMT_FLAG_*) */
};
```

# What Is Color?

# Spectral Power Distribution

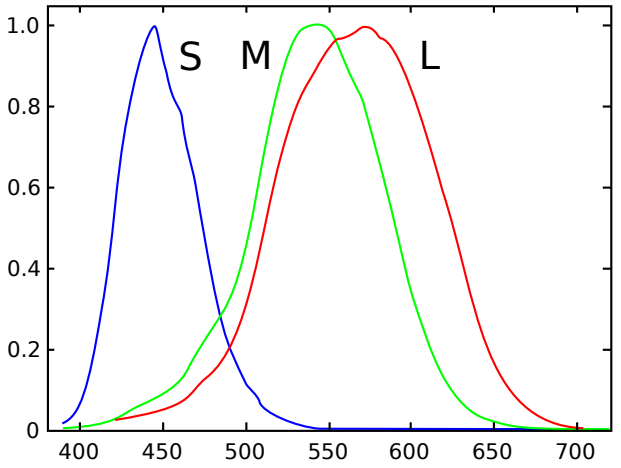
- Power per area per wavelength: i.e. the intensity of the light of different wavelengths.




By SchwartzD (Own work) [CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons

# The Eye

- Three types of cones detect color: S, M and L cones with peak sensitivities around 440 nm, 545 nm and 565 nm. The visible spectrum is about 390 – 700 nm.

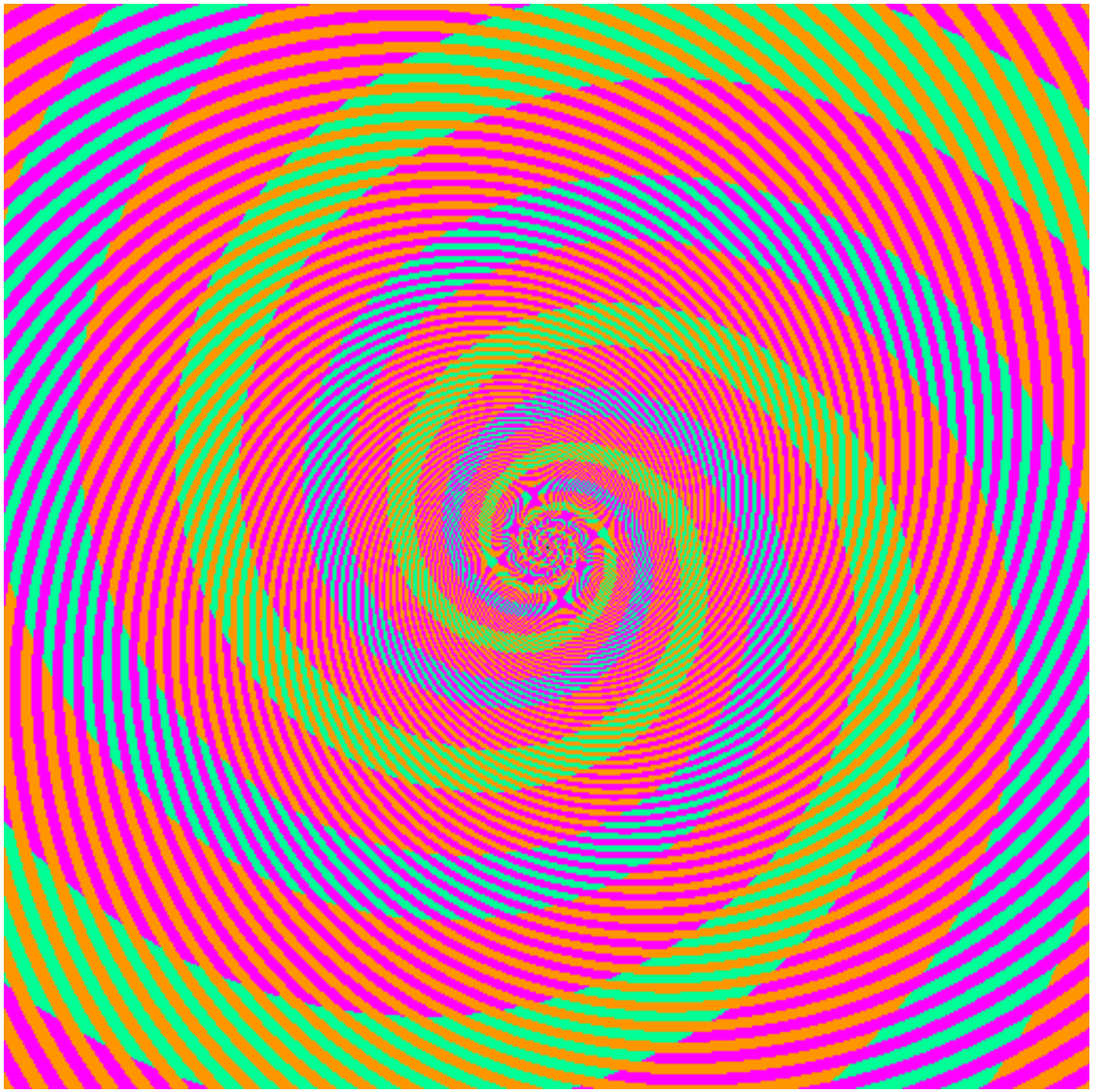



"Cones SMJ2 E" by Vanessaezekowitz at en.wikipedia / Later version uploaded by BenRG. - Based on Dicklyon's PNG version, itself based on data from Stockman, MacLeod & Johnson (1993) Journal of the Optical Society of America A, 10, 2491-2521d http://psy.ucsd.edu/~dmacleod/publications/61StockmanMacLeodJohnson1993.pdf (log E human cone response, via http://www.cvrl.org/database/text/cones/smj2.htm)Transferred from en.wikipedia to Commons by User:Richard001 using CommonsHelper.. Licensed under Creative Commons Attribution-Share Alike 3.0-2.5-2.0-1.0 via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Cones_SMJ2_E.svg#mediaviewer/File:Cones_SMJ2_E.svg

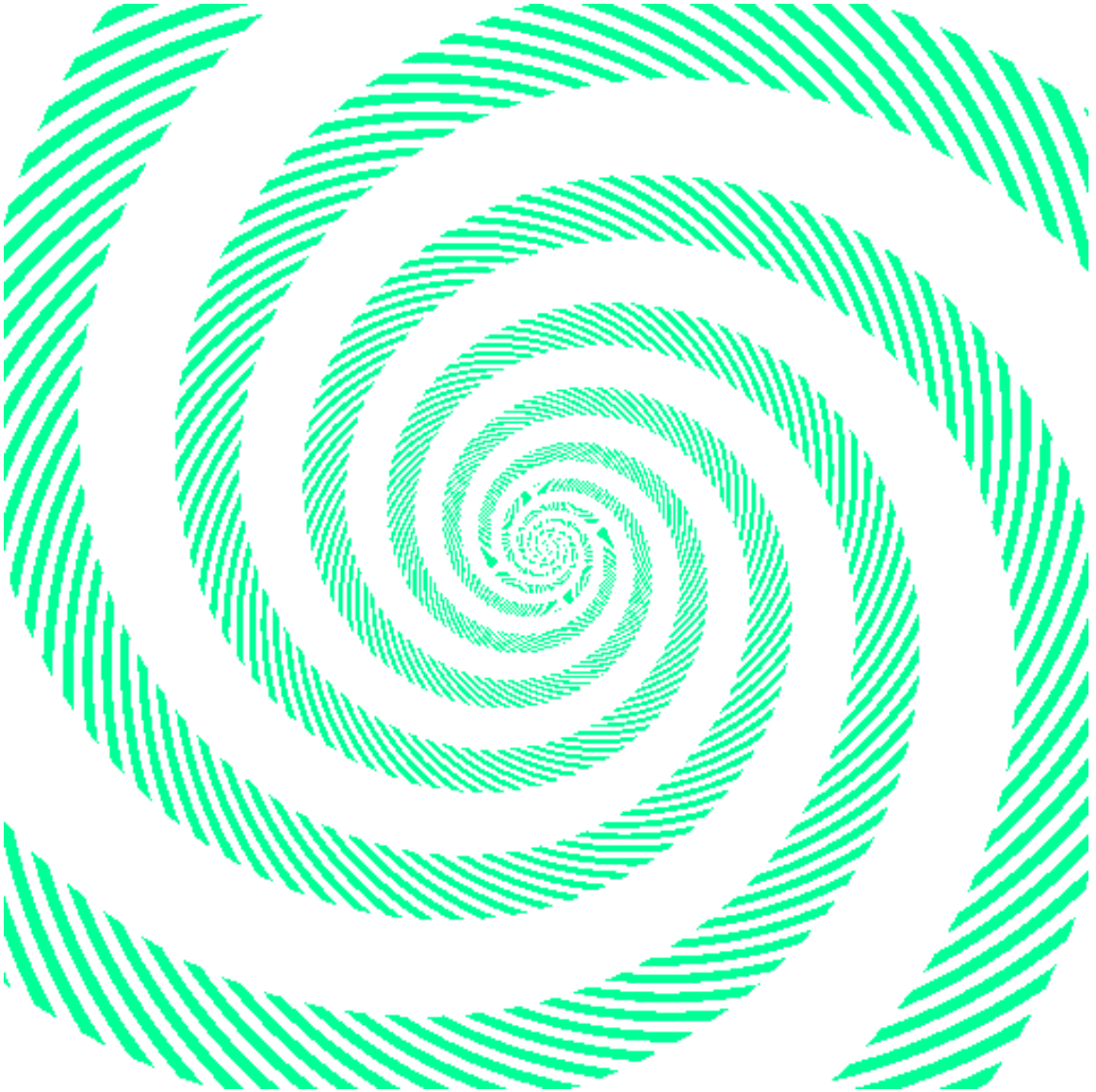
# An Illusion

- Light from an object with a certain Spectral Power Distribution hits the cones in the eye and the SPD is reduced to three 'values', one from each cone type.
- These three values are interpreted as colors by the brain. Many different SPDs will all map to the same three values.
- To reproduce colors all you have to do is create an SPD that results in the same three values coming from the cones.
- Displays and photos do not recreate the original image, they just recreate the cone impulses to the brain.
- Three cones, so you need three light sources with wavelengths around the cone peak sensitivities to recreate the illusion.
- The mantis shrimp is sensitive to 12 different wavelengths, and so would need displays with 12 light sources to recreate the illusion.

# An Illusion



# An Illusion



# CIE XYZ

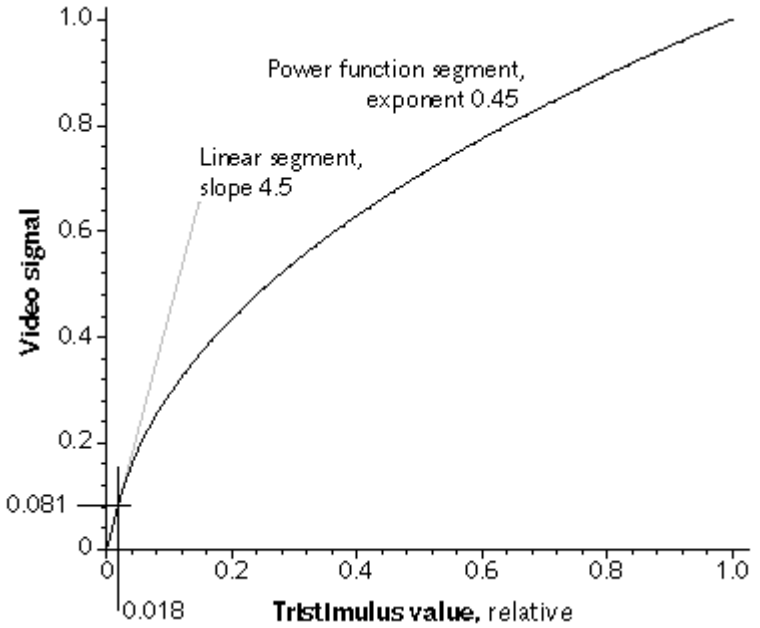
- Studies done in the 1920s resulted in data sets that determined how a specific wavelength can be reproduced using three primary light sources of 435.8, 546.1 and 700 nm. Also known as the CIE RGB colorspace.
- CIE: *Commission internationale de l'éclairage* or *International Commission on Illumination*.
- In practice the CIE XYZ colorspace is used, which is a linear transformation of RGB to an all positive set of color matching functions using imaginary primaries (i.e. light sources that are not physically possible).
- CIE XYZ is the foundation of all colorspaces, based from measurements taken from a pool of 17 people over 85 years ago!
- Note: Y describes luminance.

# Colorspaces

- CIE xyz: $x = X / (X + Y + Z)$,
  $y = Y / (X + Y + Z)$,
  $z = Z / (X + Y + Z) = 1 - x - y$
- CIE Yxy: Y = luminance, xy = chromaticity.
- A colorspace defines which three primaries are used to recreate colors. These are usually defined as a (x, y) tuple. These tuples represent the color vectors (1, 0, 0), (0, 1, 0), (0, 0, 1).
- A colorspace also defines the white point (x, y) tuple, which is the color obtained by the maximum of all three primaries or the color vector (1, 1, 1). It effectively defines the relative power or energy of the primaries.
- In the past these were based on phosphors used in TVs.
- Any colorspace can be derived from the CIE XYZ colorspace through a 3x3 matrix multiplication.

# Transfer Function

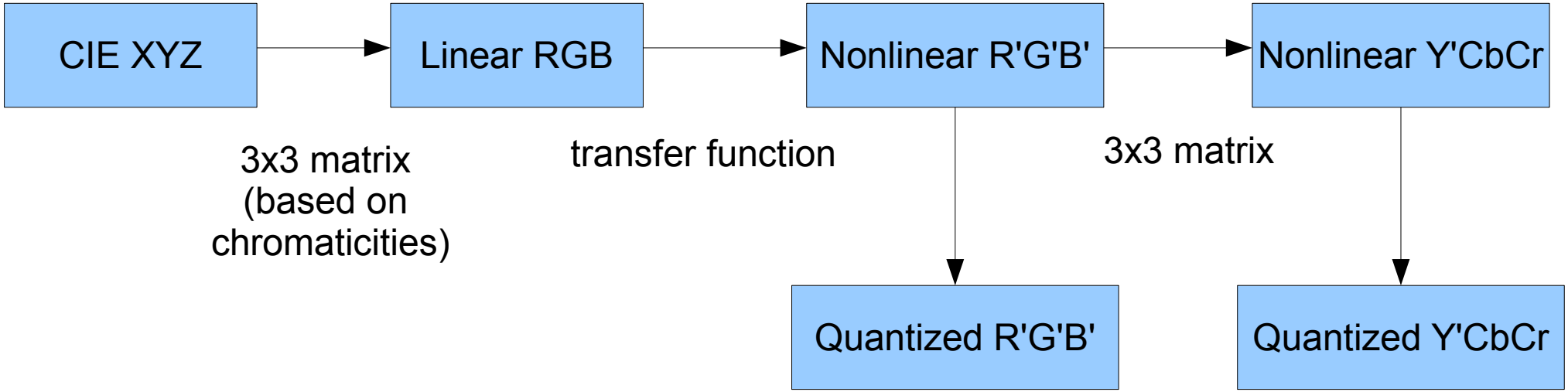
- Colorspaces are linear but early CRTs weren't and neither is human vision: double the luminance will not actually look twice as bright.
- Colorspaces define a transfer function to go from linear to a non-linear color representation. Denoted as R'G'B' where each component has gone through the transfer function. The transfer function is sometimes called 'gamma' which is usually incorrect.



# Y’CbCr Conversion & Quantization

- Y'CbCr (aka YUV) is in turn derived from R'G'B' via a matrix multiplication. Colorspaces can define one or more possible matrices for this. Note that Y'CbCr is *not* a colorspace, it is just a different way of representing colors within a colorspace.
- All colorspaces can represent all colors (although with values outside the 0-1 range), but not after quantization, i.e. when values < 0 or > 1 are cut off. This defines the gamut or extent of a colorspace.
- Take care when using textures in openGL: openGL expects linear RGB by default, not R'G'B'!

# Color Transformations



# Standards

- Rec. 709: HDTV colorspace. Standard: Rec. ITU-R BT.709-5.
- Best known: sRGB. Used for computer graphics. Standard: IEC 61966-2-1:1999. Chromaticities identical to Rec. 709, but different transfer function.
- SMPTE 170M (aka SMPTE C): defines the SDTV colorspace. The transfer function is identical to Rec. 709, but the chromaticities are different. The Y'CbCr encoding uses the encoding defined in BT.601. Note: BT.601 only defines the Y'CbCr encoding, not a colorspace. BT.601 is sometimes – incorrectly – used as an alias for SMPTE 170M.
- BT.2020: used for UHDTV with deep color (>= 10 bits per component). The transfer function is the same as Rec. 709. It defines two different Y'CbCr encodings.

# Limited and Full Range

- Typically 8-bit R', G' and B' values are quantized to the range [0-255]. This is full range quantization.
- Typically 8-bit Y' values are quantized to the range [16-235] and Cb and Cr values to the range [16-240]. This is limited range quantization.
- But limited range R'G'B' (values in the range [16-235]) and full range Y'CbCr variants exist as well. HDMI can signal both variants.

# Guidelines

- SDTV: SMPTE 170M.
- HDTV: Rec. 709.
- Graphics: sRGB.

# Problems

- Highly confusing colorspace names.
- Inconsistent conversion matrices: often buggy, never the same.
- Applications often ignore colorspace information.
- Incorrect handling of limited/full range quantization.
- Hardware does not take different transfer functions into account. E.g. the adv7604 driver produces the wrong results when converting from Rec. 709 Y'CbCr to sRGB R'G'B'.

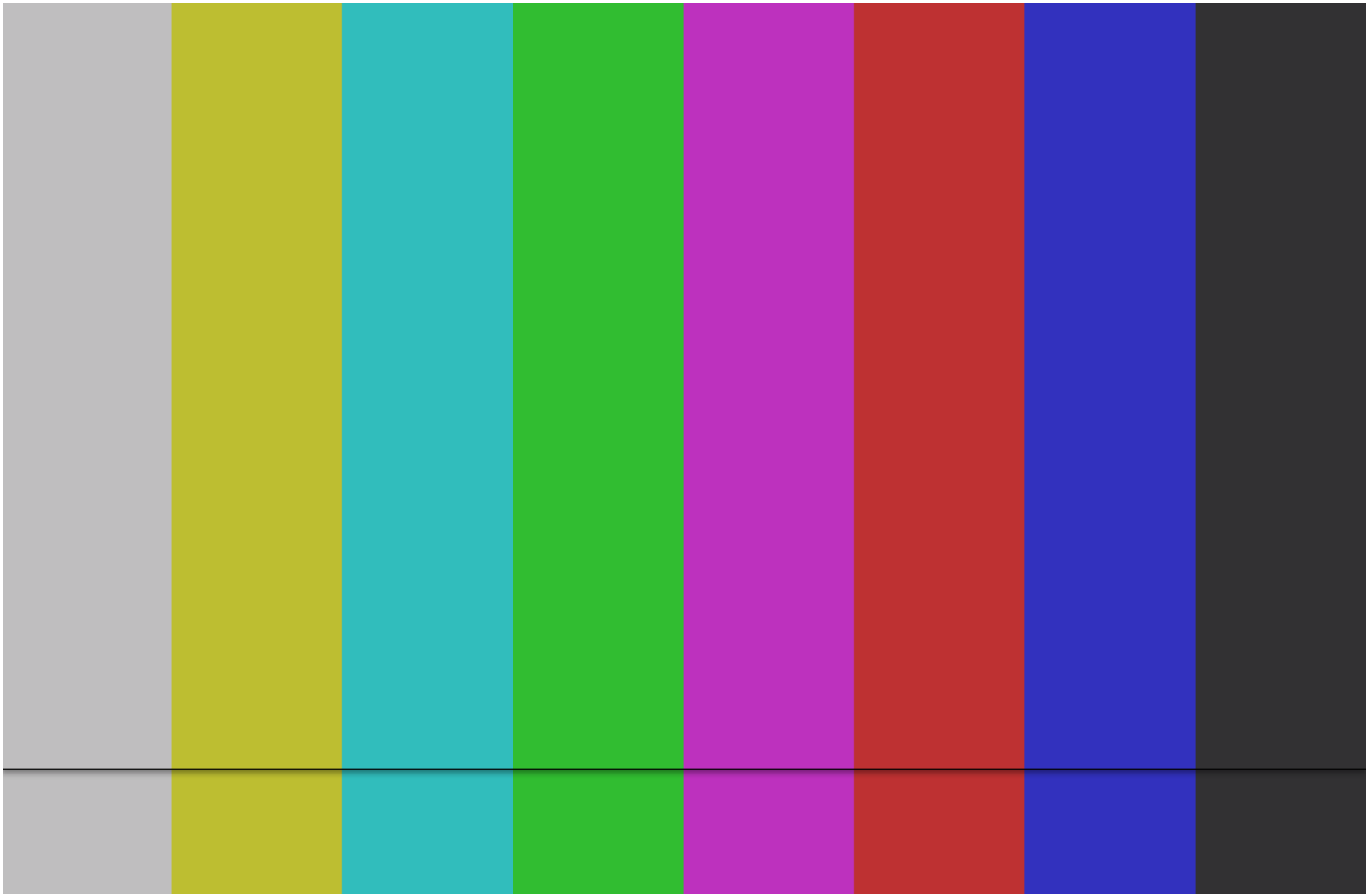
# struct v4l2_pix_format (today)

include/uapi/linux/videodev2.h:

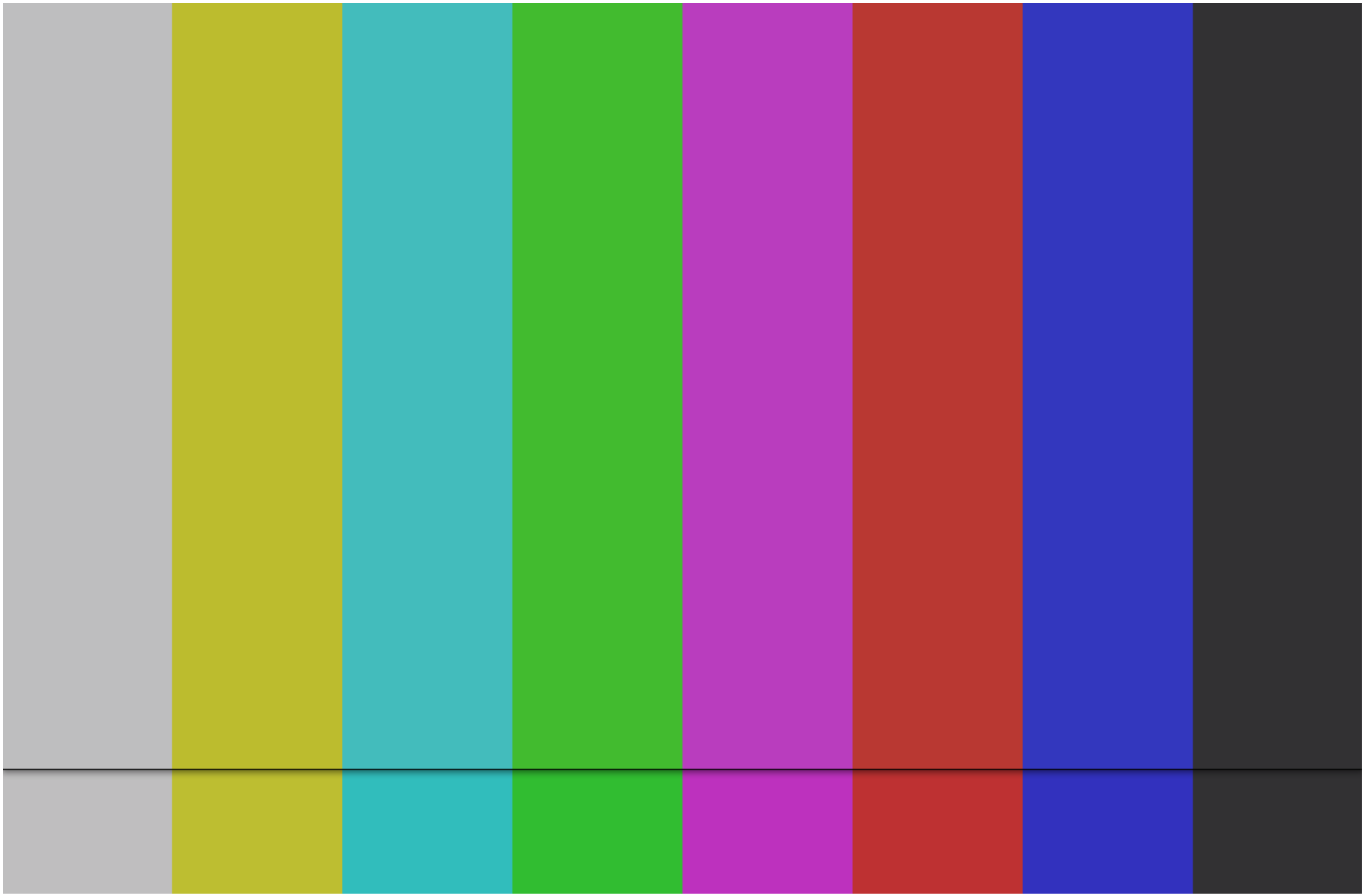
```
struct v4l2_pix_format {
        __u32 width;
        __u32 height;
        __u32 pixelformat;
        __u32 field;            /* enum v4l2_field */
        __u32 bytesperline;     /* for padding, zero if unused */
        __u32 sizeimage;
        __u32 colorspace;       /* enum v4l2_colorspace */
        __u32 priv;             /* private data, depends on pixelformat */
        __u32 flags;            /* format flags (V4L2_PIX_FMT_FLAG_*) */
        __u32 ycbcr_enc;        /* enum v4l2_ycbcr_encoding */
        __u32 quantization;     /* enum v4l2_quantization */
        __u32 xfer_func;        /* enum v4l2_xfer_func */
};
```

# What Happens When...

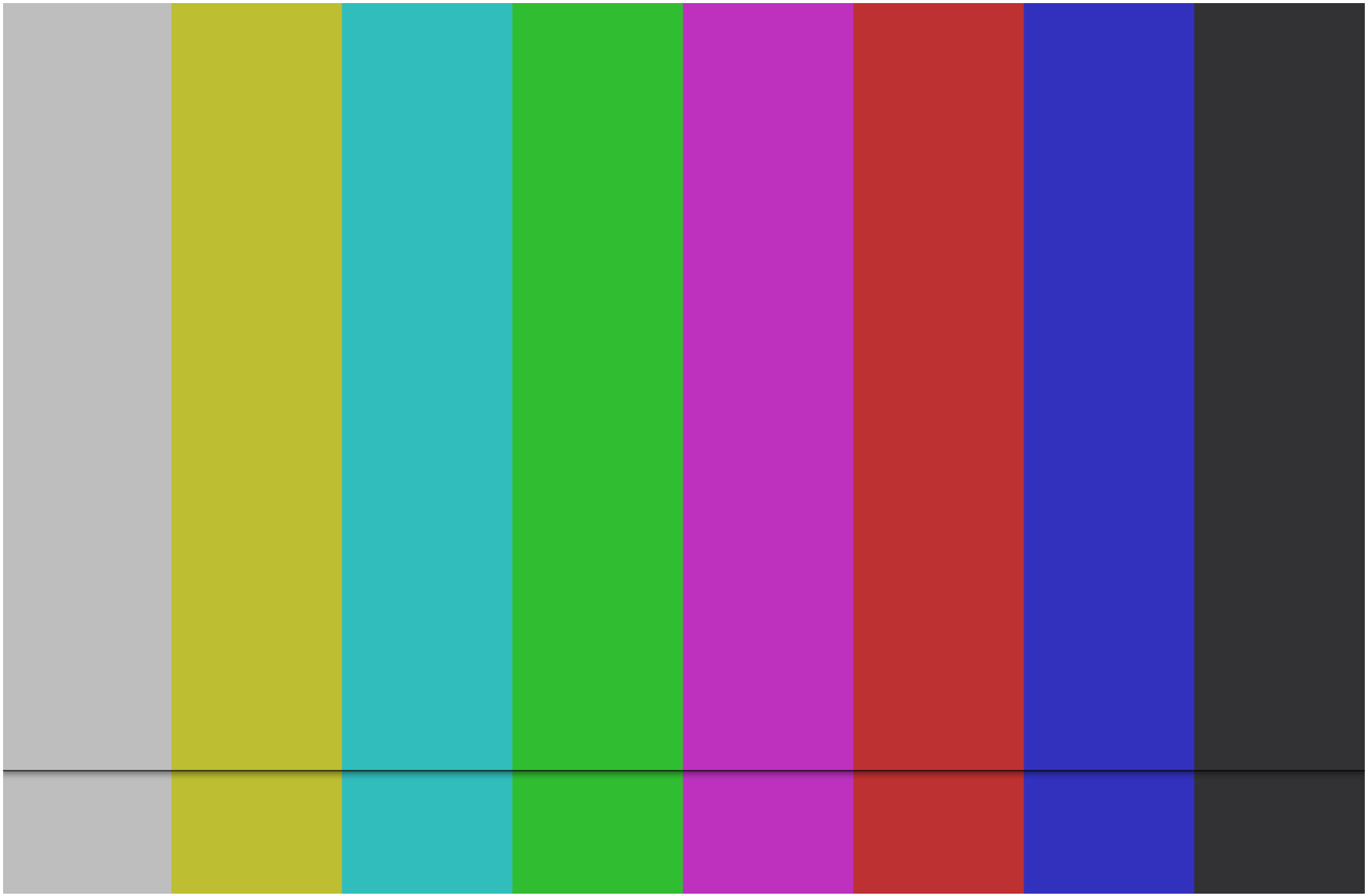
# SMPTE-170M vs Rec. 709



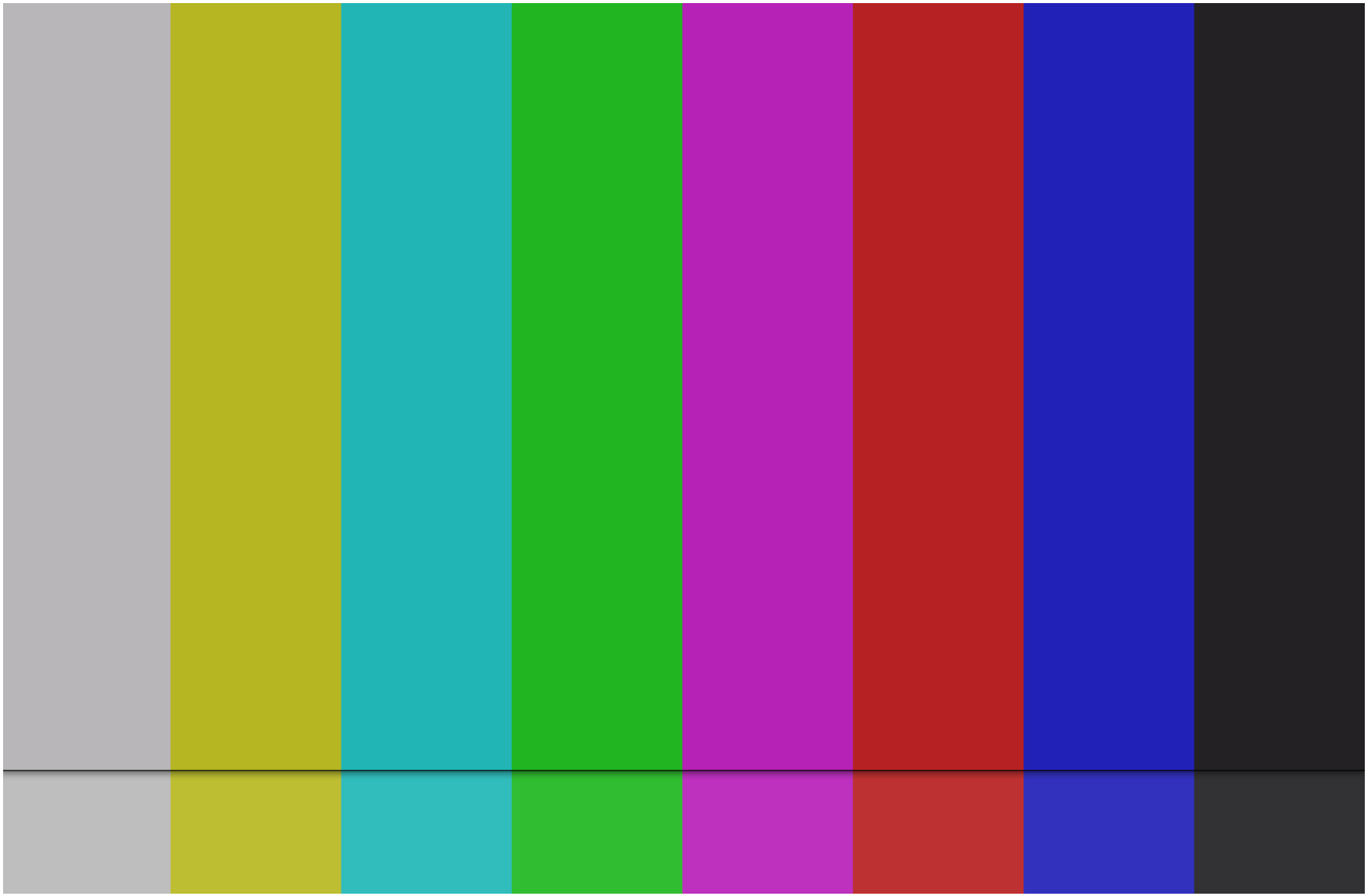
# SMPTE-170M vs Rec. 709



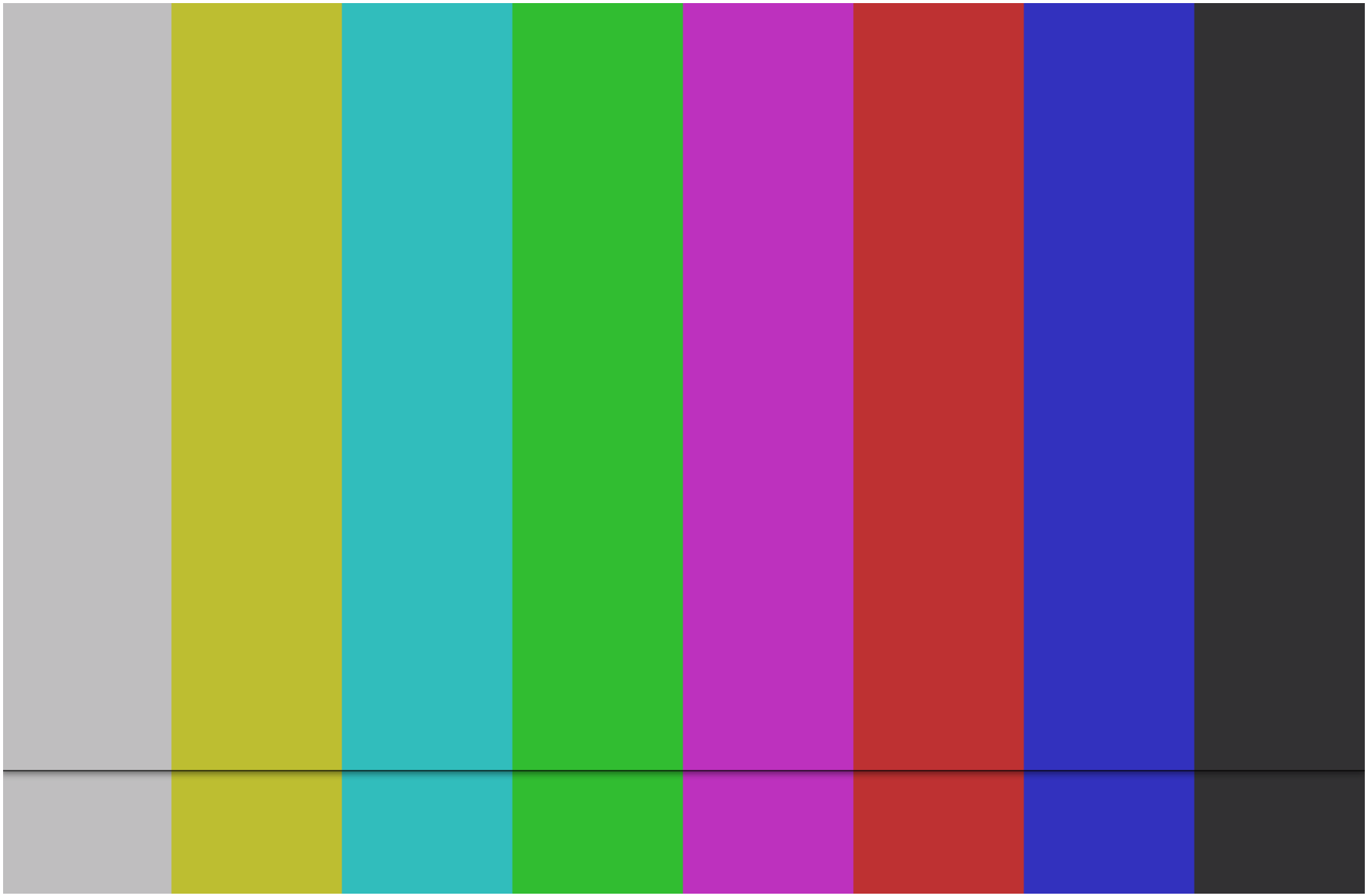
# Rec. 709 vs sRGB Transfer Function



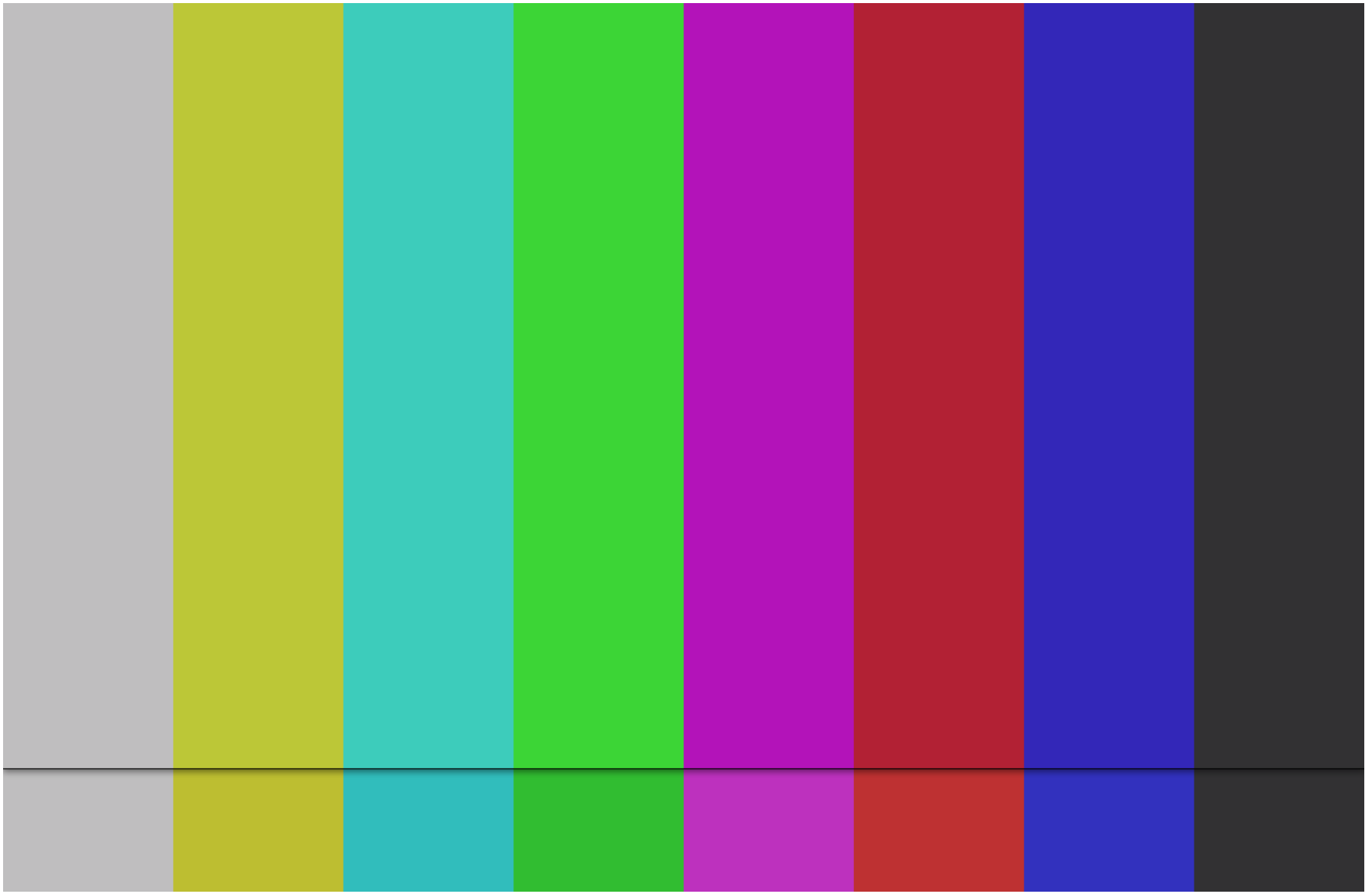
# Rec. 709 vs sRGB Transfer Function



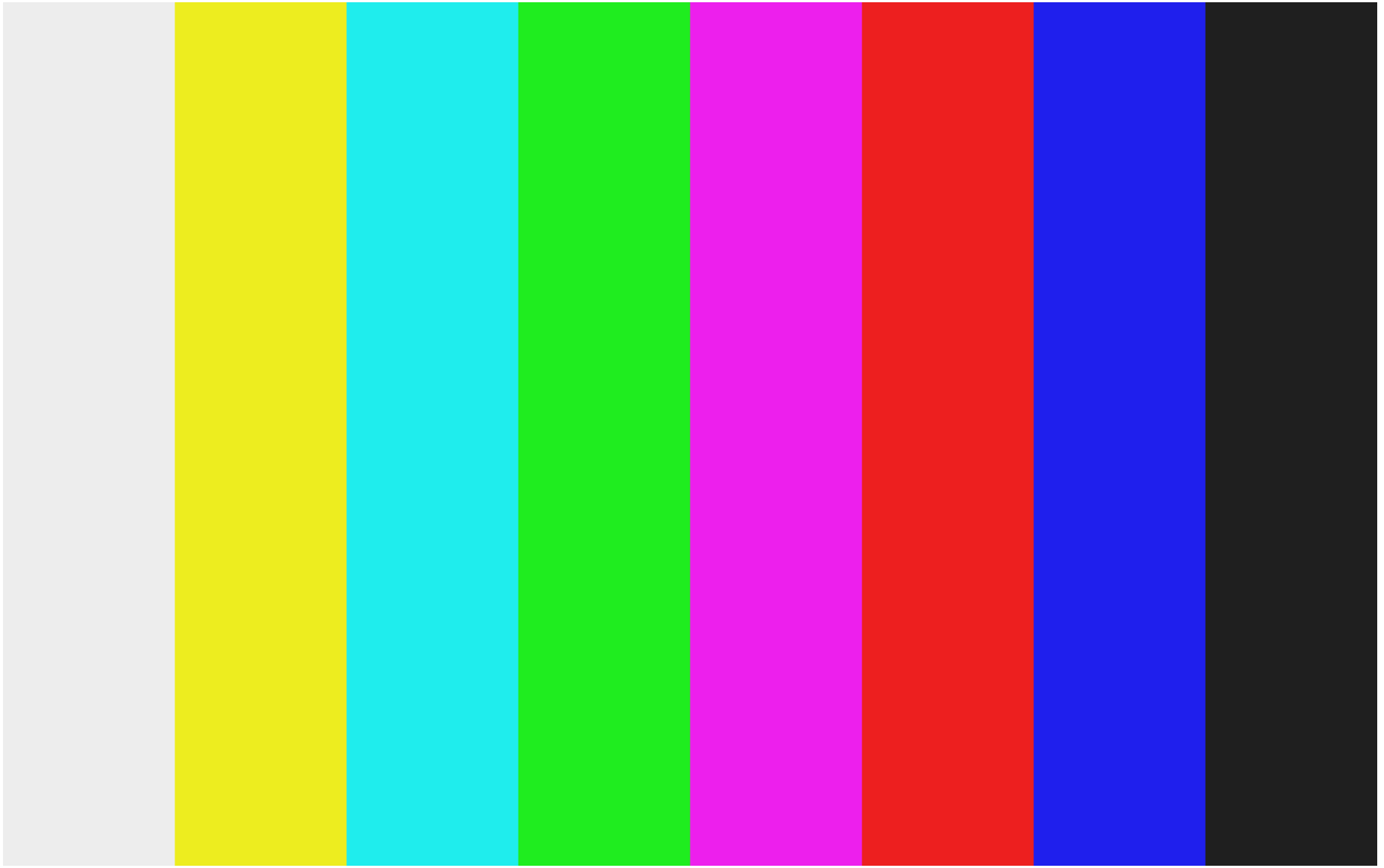
# Rec. 709 vs BT.601 Y´CbCr Encoding



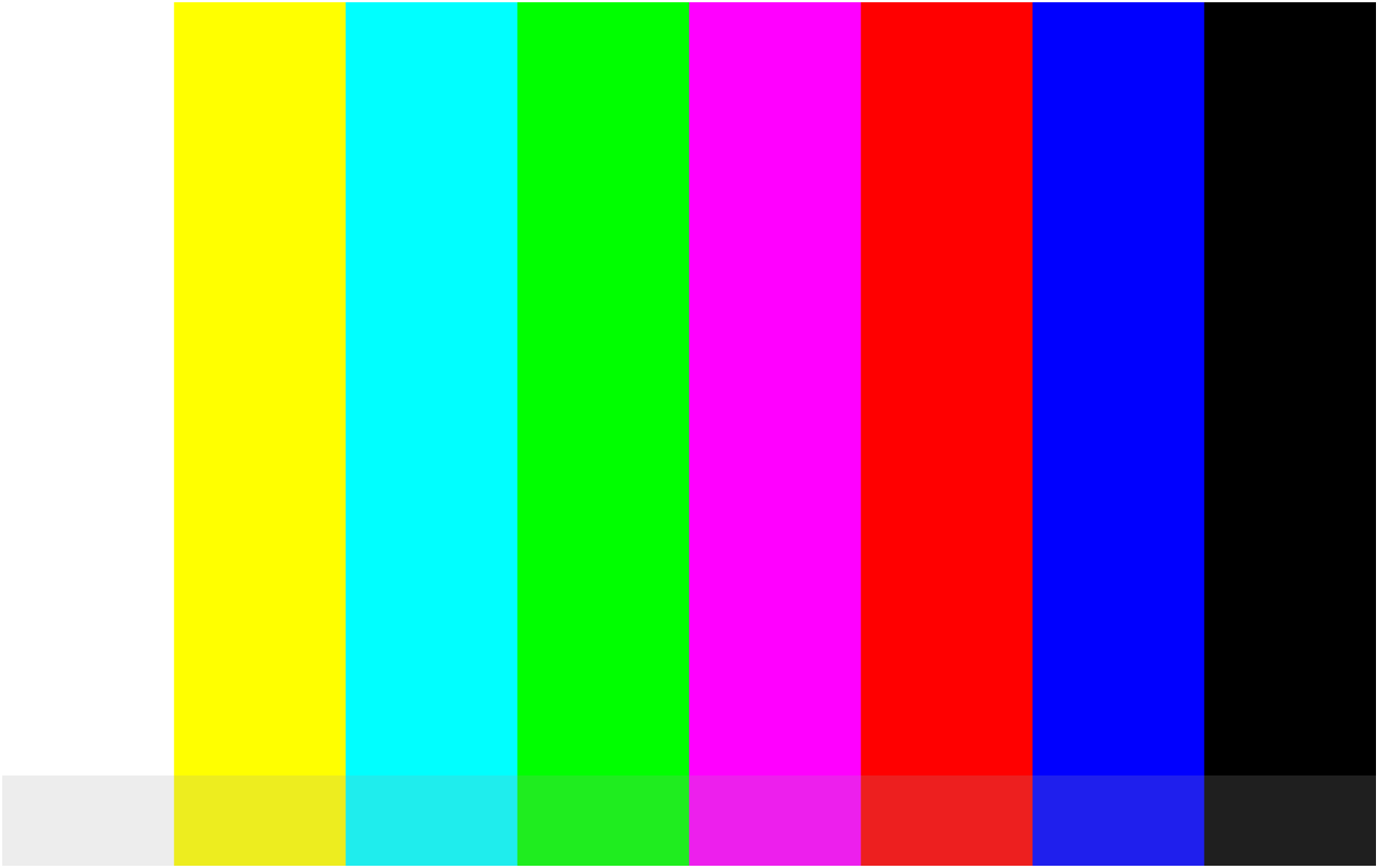
# Rec. 709 vs BT.601 Y´CbCr Encoding



# Limited vs Full Range Quantization



# Limited vs Full Range Quantization



# Resources

- Email: hverkuil@xs4all.nl
- Color Imaging – Fundamentals and Applications, Erik Reinhard et. al.
- Digital Video and HDTV – Algorithms and Interfaces, Charles Poynton.
- http://www.brucelindbloom.com
- https://linuxtv.org/downloads/v4l-dvb-apis-new/media/uapi/v4l/colorspaces.html

# Questions?