SMPTE ST 2094-40:2016

SMPTE STANDARD

Dynamic Metadata for Color Volume Transform — Application #4

Table of Contents **Page**

Approved
August 24, 2016

Foreword

SMPTE (the Society of Motion Picture and Television Engineers) is an internationally-recognized standards developing organization. Headquartered and incorporated in the United States of America, SMPTE has members in over 80 countries on six continents. SMPTE's Engineering Documents, including Standards, Recommended Practices, and Engineering Guidelines, are prepared by SMPTE's Technology Committees. Participation in these Committees is open to all with a bona fide interest in their work. SMPTE cooperates closely with other standards-developing organizations, including ISO, IEC and ITU.

SMPTE Engineering Documents are drafted in accordance with the rules given in its Standards Operations Manual.

SMPTE ST 2094-40 was prepared by Technology Committee 10E.

Intellectual Property

SMPTE draws attention to the fact that it is claimed that compliance with this Standard may involve the use of one or more patents or other intellectual property rights (collectively, "IPR"). The Society takes no position concerning the evidence, validity, or scope of this IPR.

Each holder of claimed IPR has assured the Society that it is willing to License all IPR it owns, and any third party IPR it has the right to sublicense, that is essential to the implementation of this Standard to those (Members and non-Members alike) desiring to implement this Standard under reasonable terms and conditions, demonstrably free of discrimination. Each holder of claimed IPR has filed a statement to such effect with SMPTE. Information may be obtained from the Director, Standards & Engineering at SMPTE Headquarters.

Attention is also drawn to the possibility that elements of this Standard may be subject to IPR other than those identified above. The Society shall not be responsible for identifying any or all such IPR.

Introduction

This section is entirely informative and does not form an integral part of this Engineering Document.

Dynamic Metadata for Color Volume Transform Application #4 defines scene-based metadata designed to help to reproduce the original intent of the creator of High Dynamic Range (HDR) and Wider Color Gamut (WCG) image essence on a display having a smaller color volume, even in the case that the mastering display and the targeted system display can both have practical limits on the peak luminance they can produce, these limits being imposed to meet display power restrictions and other constraints.

1 Scope

This standard specifies the metadata for Color Volume Transform Application #4, Scene-based Color Volume Mapping. It is a specialization of the content-dependent transform metadata entries and processing blocks of the generalized color volume transform model defined in the SMPTE ST 2094-1 Core Components standard.

Scene-based Color Volume Mapping consists of scene-based tone mapping and scene-based color saturation mapping processing blocks. These processing blocks make use of the peak luminance behaviors of the display used for mastering the image essence (i.e. the “mastering display”) and the targeted system display as their ability to achieve peak luminance varies with the average brightness level of the pixels and the number of bright pixels within the scene. In addition to these adjustments that can be algorithmically determined from known display and content characteristics, Scene-based Color Volume Mapping also allows creatively approved adjustments.

2 Conformance Notation

Normative text is text that describes elements of the design that are indispensable or contains the conformance language keywords: "shall", "should", or "may". Informative text is text that is potentially helpful to the user, but not indispensable, and can be removed, changed, or added editorially without affecting interoperability. Informative text does not contain any conformance keywords.

All text in this document is, by default, normative, except: the Introduction, any section explicitly labeled as "Informative" or individual paragraphs that start with "Note:”

The keywords "shall" and "shall not" indicate requirements strictly to be followed in order to conform to the document and from which no deviation is permitted.

The keywords, "should" and "should not" indicate that, among several possibilities, one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required; or that (in the negative form) a certain possibility or course of action is deprecated but not prohibited.

The keywords "may" and "need not" indicate courses of action permissible within the limits of the document.

The keyword “reserved” indicates a provision that is not defined at this time, shall not be used, and may be defined in the future. The keyword “forbidden” indicates “reserved” and in addition indicates that the provision will never be defined in the future.

A conformant implementation according to this document is one that includes all mandatory provisions ("shall") and, if implemented, all recommended provisions ("should") as described. A conformant implementation need not implement optional provisions ("may") and need not implement them as described.

Unless otherwise specified, the order of precedence of the types of normative information in this document shall be as follows: Normative prose shall be the authoritative definition; Tables shall be next; then formal languages; then figures; and then any other language forms.

3 Normative References

The following standards contain provisions which, through reference in this text, constitute provisions of this engineering document. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this engineering document are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

SMPTE ST 2086:2014, Mastering Display Color Volume Metadata Supporting High Luminance and Wide Color Gamut Images

SMPTE ST 2094-1:2016, Dynamic Metadata for Color Volume Transform — Core Components

4 Terms and Definitions

SMPTE ST 2094-1, Section 4, Terms and Definitions applies to this document. This section defines additional terms and definitions used in this document.

4.1 actual display peak luminance

peak luminance (in units of cd/m^2) that a display is capable of delivering while rendering the scene

Note: This depends on the spatial distribution of the luminance levels of the pixels in the scene and the power consumption limits of the display

4.2 brightest pixel

pixel with the maximum approximate luminance among the smoothed selected pixels in the scene, with the smoothing carried out by a box filter as defined in Section 10

4.3 input image essence

image essence to which the color volume transform is applied and having three color components labeled R, G and B where the R component corresponds to the Mastering Display Color Volume Display Primary with the largest x chromaticity coordinate, the G component corresponds to the Mastering Display Color Volume Display Primary with the largest y chromaticity coordinate, and the B component corresponds to the remaining Mastering Display Color Volume Display Primary

4.4 linearized maxRGB

maxRGB values as defined in SMPTE ST 2094-1, normalized to the range [0, 1], representing linear light level with 0 representing 0 cd/m^2 and 1 representing 10000 cd/m^2 when the three normalized color components have the same value

4.5 linearized RGB

linearized color component values, normalized to the range [0, 1], with 0 representing 0 cd/m^2 and 1 representing 10000 cd/m^2 when the three normalized color components have the same value

4.6 scene

selected pixels throughout all images in the time interval as defined in SMPTE ST 2094-1

4.7 selected pixels

pixels within a processing window as defined in SMPTE ST 2094-1 or within the intersection of a processing window and the external ellipse (as described in “Calculus and Analytic Geometry”, G.B. Thomas and R. L. Finney) of the elliptical pixel selector if the processing window is extended using the elliptical pixel selector as defined in Section 6

5 Application Identification

The **ApplicationIdentifier** value shall be 4 and the **ApplicationVersion** value shall be 0 to identify this version of Application #4.

These two values identify this document as the defining document for the application-specific metadata specified in Section 9.1.

6 Extension to the Processing Window

6.1 Selection of Pixels for Processing

The Processing Window as defined in SMPTE ST 2094-1 may be extended with an elliptical pixel selector. The elliptical pixel selector shall consist of two aligned, concentric, similar ellipses (internal and external). The ellipse coordinate system shall be the same as the pixel coordinate system. An example of the elliptical pixel selector is shown in Figure 1, where θ, a_1, a_2, b_2 indicate the rotation angle, the semi-major axis of the internal ellipse, the semi-major axis of the external ellipse and the semi-minor axis of the of the external ellipse respectively. (x_0, y_0) denotes the pixel coordinates of the center of the ellipse. The semi-major axes are along the direction of the rotation angle θ. Rotation angle $\theta = 0$ is parallel to the x-axis and it increases in a clockwise fashion.

The pixels between the internal and external ellipses form a transition region to aid in preventing visible boundary artifacts between the local processing region and the background.

The selected pixels shall be the intersection of the external ellipse and the **ProcessingWindow**.

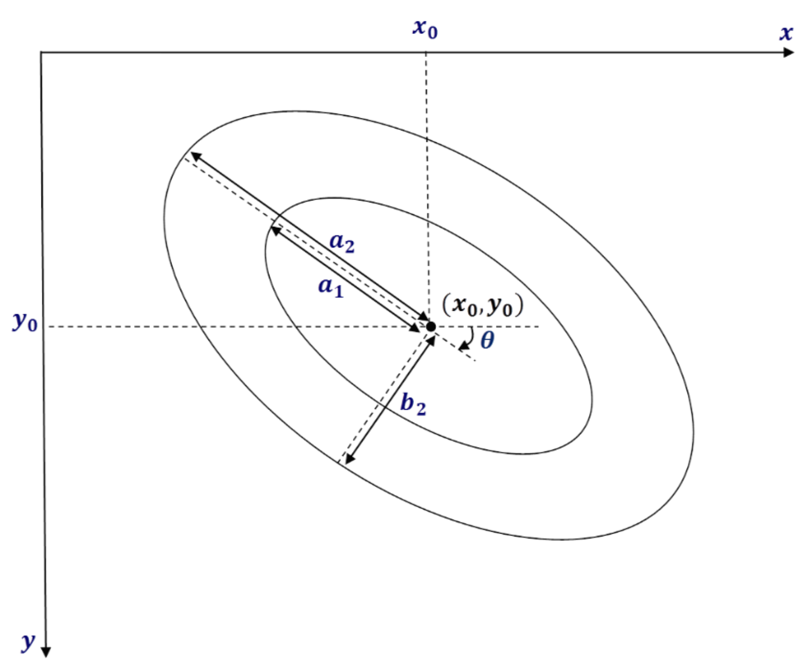

Figure 1 – An elliptical pixel selector consisting of two concentric ellipses: internal and external ellipses

6.2 Center of Ellipse

The **CenterOfEllipse** shall be a vector with two integers. This shall represent the center position (pixel coordinates) of the concentric internal and external ellipses of the elliptical pixel selector. The order of the vector elements shall be {x coordinate, y coordinate}. The x coordinate shall be in the range of [0, image width - 1] and the y coordinate shall be in the range of [0, image height - 1]. The x and y coordinates shall be in multiples of 1 pixel. The **CenterOfEllipse** is denoted by (x_0, y_0) in Figure 1.

6.3 Rotation Angle

The **RotationAngle** shall represent the clockwise rotation angle in degree of arc with respect to the positive direction of the x-axis of the concentric internal and external ellipses of the elliptical pixel selector centered at the **CenterOfEllipse**. **RotationAngle** shall be in the range [0,180] and in multiples of 1. The **RotationAngle** is denoted by θ in Figure 1.

6.4 Semi-Major Axis of the Internal Ellipse

The **SemiMajorAxisInternalEllipse** shall represent the semi-major axis value of the internal ellipse of the elliptical pixel selector in amount of pixels. The **SemiMajorAxisInternalEllipse** shall be in the range of [1, 65535] and in multiples of 1 pixel. The **SemiMajorAxisInternalEllipse** is denoted by a_1 in Figure 1.

6.5 Semi-Major Axis of the External Ellipse

The **SemiMajorAxisExternalEllipse** shall represent the semi-major axis value of the external ellipse of the elliptical pixel selector in amount of pixels. The **SemiMajorAxisExternalEllipse** shall not be less than the **SemiMajorAxisInternalEllipse**. The **SemiMajorAxisExternalEllipse** shall be in the range of [1, 65535] and in multiples of 1 pixel. The **SemiMajorAxisExternalEllipse** is denoted by a_2 in Figure 1.

6.6 Semi-Minor Axis of the External Ellipse

The **SemiMinorAxisExternalEllipse** shall represent the semi-minor axis value of the external ellipse of the elliptical pixel selector in amount of pixels. The **SemiMinorAxisExternalEllipse** shall be in the range of [1, 65535] and in multiples of 1 pixel. The **SemiMinorAxisInternalEllipse** is denoted by b_2 in Figure 1.

6.7 Overlap Process Option

The **OverlapProcessOption** shall be an enumerator that indicates one of the two methods of combining rendered pixels in an image with at least one elliptical pixel selector. The value of **OverlapProcessOption** shall be either 0 or 1, where 0 shall mean method 1 and 1 shall mean method 2. Example methods are described in Annex B. For overlapping elliptical pixel selectors in an image, the **OverlapProcessOption** shall have the same value.

7 Targeted System Display

7.1 Introduction

The **TargetedSystemDisplay** metadata group associated with Application #4 Scene-based Color Volume Mapping contains the metadata item defined in Section 7.2.

7.2 Targeted System Display Actual Peak Luminance (2D LUT)

The **TargetedSystemDisplayActualPeakLuminance** shall be a two-input sampled function as defined in SMPTE ST 2094-1. It takes as x (i.e. first input) **FractionBrightPixels** (as defined in Section 8.6) and y (i.e. second input) **AverageMaxRGB** (as defined in Section 8.4) and outputs the normalized actual peak luminance of the targeted system display. Output values shall be in the range of [0,1] and in multiples of 1/15, with 0 corresponding to 0 cd/m^2 and 1 corresponding to **TargetedSystemDisplayMaximumLuminance.** The default output value shall be 1. The array of the **TargetedSystemDisplayActualPeakLuminance** two-input sampled function shall be limited to a maximum of 25 rows and a maximum of 25 columns.

8 Color Volume Mapping

8.1 Introduction

The color volume transform for Application #4 is based on scene-based color volume mapping, as described in Annex B. The **ColorVolumeTransform** metadata items associated with Application #4 are defined in Sections 8.2 to 8.8.

8.2 Mastering Display Actual Peak Luminance (2D LUT)

The **MasteringDisplayActualPeakLuminance** shall be a two-input sampled function as defined in SMPTE ST 2094-1. It takes as x (i.e. first input) **FractionBrightPixels** (as defined in Section 8.6) and y (i.e. second input) **AverageMaxRGB** (as defined in Section 8.4) and outputs the normalized actual peak luminance of the mastering display used for mastering the image essence. An output value shall be in the range of [0,1] and in multiples of 1/15, with 0 corresponding to 0 cd/m^2 and 1 corresponding to Maximum Display Mastering Luminance (as defined in SMPTE ST 2086). The default output value shall be 1. The array of the **MasteringDisplayActualPeakLuminance** two-input sampled function shall be limited to a maximum of 25 rows and a maximum of 25 columns.

8.3 Maximum Scene Color Component Levels

The **MaxSCL** shall be a vector with three elements. They shall be the maximum of each component of linearized RGB values in the scene. The elements shall each be in the range [0,1] and in multiples of 0.00001. The order of the elements shall be {R,G,B}.

8.4 Average MaxRGB

The **AverageMaxRGB** shall be the average of linearized maxRGB values in the scene. The value shall be a number in the range [0,1] and in multiples of 0.00001.

8.5 Distribution MaxRGB

The **DistributionMaxRGB** shall be a set of two equal-length vectors. Each element α_i in the first vector shall be an integer value in the range [0,100]. Values in the second vector shall be in the range [0,1] and in multiples of 0.00001. The element at position i in the second vector shall be the value below which α_i percent of the linearized maxRGB values in the scene fall. The maximum length of these vectors shall be 15. The elements of the vectors shall be in ascending order. The first vector should be {5 ,10, 50, 90, 95}.

8.6 Fraction of the Bright Pixels

The **FractionBrightPixels** shall be a number representing the fraction of selected pixels in the image that contains the brightest pixel in the scene. The **FractionBrightPixels** shall be in the range [0,1] and in multiples of 0.001. **FractionBrightPixels** shall be calculated as specified in Section 10.

8.7 Scene Based Tone Mapping

8.7.1 Overview

A tone mapping function shall be composed of a linear part and a curved part. The two parts shall be joined by a knee point (K_s, K_F) as depicted in Figure 2.

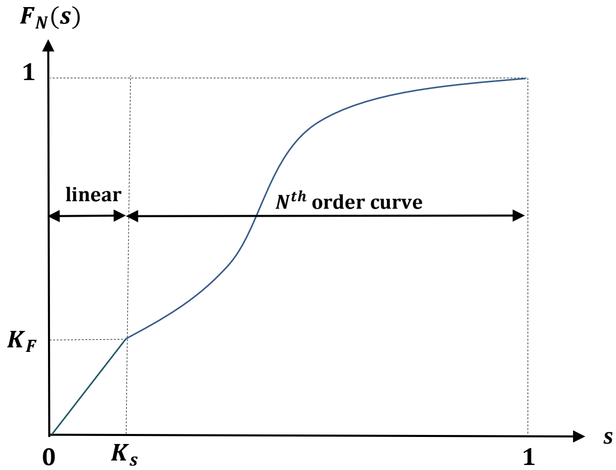

Figure 2 – A tone mapping function with a knee point

The curve part shall be defined by Equation (1):

$$B_N(t) = \sum_{\rho=0}^{N} \binom{N}{\rho} \times t^{\rho} \times (1-t)^{N-\rho} \times P_{\rho}, \qquad \text{for } \ 0 \le t \le 1 \tag{1}$$

where:

$t = \frac{s-K_S}{1-K_S}$,

s = normalized input in the range [0,1] with 1 corresponding to either the output of the metadata item defined in Section 8.2 if available or the maximum of the elements of **MaxSCL** defined in Section 8.3.

N = the number of control points in the range [2,17],

$\binom{N}{\rho} = \frac{N!}{\rho! \times (N-\rho)!}$,

$P_0, \cdots, P_N =$ anchor parameters that characterize the curved part of the tone mapping function with $P_0 = 0$, $P_N = 1$, and where P_i for $i = 1, \ldots, N-1$ shall satisfy the sufficient condition for monotonically increasing the curve part such as

$$\frac{d}{dt} B_N(t) = \sum_{\rho=0}^{N} \binom{N}{\rho} t^{\rho-1} \times (1-t)^{N-\rho-1} \times (\rho - N \times t) \times P_\rho \geq 0, \qquad \forall\, t \in [0,1].$$

Note: The curve given in (1) is an N^{th} order polynomial which can be expressed as

$$B_N(t) = \sum_{\rho=0}^{N} a_\rho \times t^\rho$$

where $a_\rho's$ can be obtained by the binomial expansion of $(1-t)^{N-\rho}$ term in Equation (1).

The tone mapping function $F_N(s)$ shall be as defined in Equation (2).

$$F_N(s) = \begin{cases} \frac{K_F}{K_S} \times s, & 0 \leq s < K_s \\ K_F + (1-K_F) \times B_N\left(\frac{s-K_S}{1-K_S}\right), & K_s \leq s \leq 1 \end{cases} \tag{2}$$

where:

$0 < K_s$,

$0 \leq B_N(t) \leq 1$ for $0 \leq \mathrm{t} \leq 1$.

$F_N(s)$ = normalized output in the range [0,1] with 1 representing either the **TargetedSystemDisplayMaximumLuminance** multiplied by the output of metadata item defined in Section 7.2 if available or the **TargetedSystemDisplayMaximumLuminance.** The default tone mapping function shall be $F_N(s) = s$.

8.7.2 Knee Point

The **KneePoint** shall be a vector with two numbers. It represents the separation point (K_s, K_F) between the linear part and the curved part of the tone mapping function as illustrated in Figure 2. The value of the numbers shall be in the range]0,1] and in multiples of 1/4095.

8.7.3 Bezier Curve Anchors

The **BezierCurveAnchors** shall be a vector of numbers whose length shall be in the range [0,15] and whose elements represent the intermediate anchor parameters $(P_1, \cdots, P_{N-1})$ in Equation (1), with $P_0 = 0$ and $P_N = 1$. The values of the vector elements shall be in the range [0,1] and in multiples of 1/1023.

8.8 Color Saturation Weight

The **ColorSaturationWeight** shall be a number in the range [0,63/8] and in multiples of 1/8. This can be used to adjust the color saturation mapping gain as described in Annex B.5.3. The default value shall be 1.

9 Application Constraints

9.1 Metadata Set

A metadata set shall contain exactly one of each of the following:

- **ApplicationIdentifier** (= 4)
- **ApplicationVersion** (= 0)
- **TimeInterval**
  - which shall include one of each of the following metadata items defined in SMPTE ST 2094-1:

    TimeIntervalStart and

    TimeIntervalDuration
- **ProcessingWindow**
  - which shall include one of each of the following metadata items defined in SMPTE ST 2094-1:

    UpperLeftCorner,

    LowerRightCorner and

    WindowNumber
  - and may include one **EllipsePixelSelector** metadata group which shall contain one of each of the following metadata items defined in this document:

    CenterOfEllipse

    RotationAngle

    SemiMajorAxisInternalEllipse

    SemiMinorAxisExternalEllipse

    SemiMajorAxisExternalEllipse

    OverlapProcessOption
- **TargetedSystemDisplay**
  - which shall include the following metadata item defined in SMPTE ST 2094-1:

    TargetedSystemDisplayMaximumLuminance
  - and may include the following metadata item defined in this document:

    TargetedSystemDisplayActualPeakLuminance
- **ColorVolumeTransform**
  - which shall include one of each of the following metadata items as defined in this document:

    MaxSCL

    AverageMaxRGB

    DistributionMaxRGB

    FractionBrightPixels
  - and may contain any combination having zero or one of each of the following metadata groups or metadata items as defined in this document:

    MasteringDisplayActualPeakLuminance

BezierCurveToneMapper which shall contain one of each of the following metadata items:

KneePoint

BezierCurveAnchors

ColorSaturationWeight

9.2 Processing Window Constraints

For each targeted system there shall be a maximum of 3 processing windows within one image.

Processing Window 0 shall be always present and shall cover all pixels in an image and shall not be extended with the elliptical pixel selector. Processing Windows 1 and 2, if present, shall be extended with the elliptical pixel selector.

10 Computation of the FractionBrightPixels

The **FractionBrightPixels** value shall be computed using Equation (3)

$$F_{BP} = \frac{1}{H \times W} \sum_z f\left(10000 \times \left(l_{max} - l_{av}(z, k')\right)\right) \times g\left(\Lambda(p', z)\right) \tag{3}$$

where:

F_{BP} = **FractionBrightPixels**,

z = position of a selected pixel at time instance k' ,

$\Lambda()$ denotes the Euclidean distance in amount of pixels,

H = image height,

W = image width,

$$f(\varepsilon) = \begin{cases} 1, & \varepsilon < \mu 1 \\ \left(\frac{1}{\mu 1 - \mu 2}\right) \times (\varepsilon - \mu 2), & \mu 1 \le \varepsilon < \mu 2 \\ 0, & \mu 2 \le \varepsilon \end{cases}$$

$$g(\varepsilon) = \begin{cases} 1, & \varepsilon < d1 \\ \left(\frac{1}{d1 - d2}\right) \times (\varepsilon - d2), & d1 \le \varepsilon < d2 \\ 0, & d2 \le \varepsilon \end{cases}$$

$\mu 1 = \frac{10000}{255}$,

$\mu 2 = \frac{50000}{255}$,

$d1 = \frac{500 \times W \times H}{3840 \times 2160}$,

$d2 = \frac{1000 \times W \times H}{3840 \times 2160}$,

$$l_{av}(p,k) = \frac{1}{A \times E} \sum_{q \in R(p)} Y_{linear}(q,k) \quad (4)$$

p = position of a selected pixel at time instance k

$A = E = 5$,

$R(p) =$ a $(A \times E)$ region centered at pixel p,

q = position of a pixel in region $R(p)$,

If region $R(p)$exceeds the boundaries of an image, $R(p)$ shall include only pixels in the image and $(A \times E)$ in Equation (4) shall be replaced by the number of pixels in region $R(p)$.

$Y_{linear}(q,k) = 0.2627 \times R_{linear}(q,k) + 0.678 \times G_{linear}(q,k) + 0.0593 \times B_{linear}(q,k)$

$(R_{linear}, G_{linear}, B_{linear})$ = linearized RGB components as defined in Section 4.5.

$(p', k') = \text{argmax}_{p,k}\ l_{av}(p,k)$

$l_{max} = l_{av}(p', k')$

Note: The symbols have the following meanings:

l_{max} = maximum luminance of the smoothed pixels,

$p' =$ position of the brightest pixel,

$k' =$ time instance of the brightest pixel,

$k =$ time instance.

Annex A Mapping of Application #4 to the Generalized Color Transform Model (Informative)

The diagram in Figure A.1 describes Application #4 Scene-based Color Volume Mapping in the framework of the Generalized Color Transform Model described in SMPTE ST 2094-1. The processing blocks applied are the Scene-based Tone Mapping and the Scene-based Color Saturation Mapping.

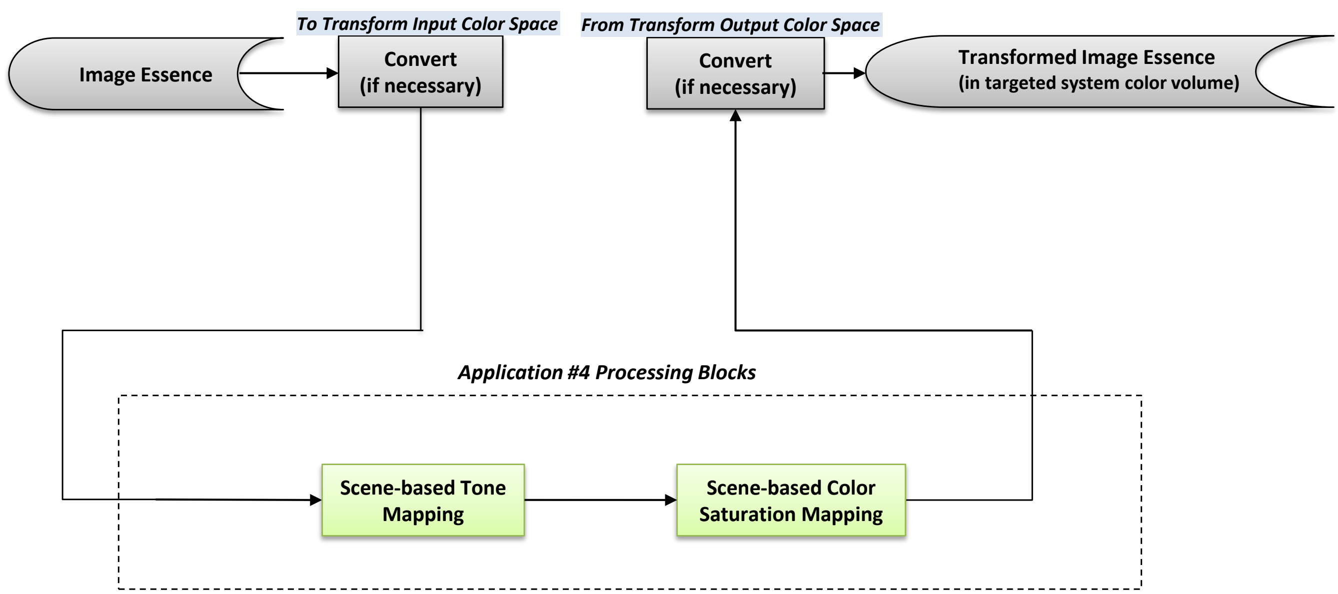

Figure A.1 – Processing blocks used by Application #4

Annex B Scene-based Color Volume Mapping Method Description (Informative)

B.1 Introduction

The scene-based color volume mapping for Application #4 is applied to the linearized color components of the source content. The proposed method uses a metadata set defined in Section 9.1 and successively processes the linearized RGB components, $R_{linear}, G_{linear}, B_{linear}$ which are in the mastering display color volume.

B.2 Source Normalized Actual Peak Luminance and Target Normalized Actual Peak Luminance

Source Normalized Actual Peak Luminance and Target Normalized Actual Peak Luminance are used to compute the metadata items defined in this document. They are defined and computed as follows.

B.2.1 Source Normalized Actual Peak Luminance

Source Normalized Actual Peak Luminance is the normalized peak level of the image essence as displayed on the mastering display used for mastering the image essence. This value is normalized by the Maximum Display Mastering Luminance (from SMPTE ST 2086) and given by Equation (5):

$$M_P = M_{RP} \times S_{MC} \times 10000 / M_{ML} \tag{5}$$

where:

M_P = Source Normalized Actual Peak Luminance,

$M_{RP} = MLUT\ (F_{BP}, S_{AMC})$,

$MLUT =$ **MasteringDisplayActualPeakLuminance**,

$S_{MC} =$ maximum of the elements of **MaxSCL** defined in Section 8.3, in the range [0,1],

M_{ML} = Maximum Display Mastering Luminance from SMPTE ST 2086,

F_{BP} = **FractionBrightPixels**, in the range [0,1],

$S_{AMC} =$ **AverageMaxRGB**, in the range [0,1].

B.2.2 Target Normalized Actual Peak Luminance

Target Normalized Actual Peak Luminance is the normalized displayed peak level of the Targeted display. This value is normalized by the **TargetedSystemDisplayMaximumLuminance** and given by Equation (6):

$$T_P = T_{RP} \times S_{MC} \times 10000 / T_{ML} \tag{6}$$

where:

T_P = Target Normalized Actual Peak Luminance,

$T_{RP} = TLUT(F_{BP}, S_{AMC})$,

$TLUT$ = **TargetedSystemDisplayActualPeakLuminance** ,

T_{ML} = **TargetedSystemDisplayMaximumLuminance**.

and where S_{AMC}, M_{ML} and F_{BP} are as defined in Annex B.2.1 above.

B.3 Color Components Normalization

Linearized RGB components are normalized using Equation (7):

$$\begin{bmatrix} R_{norm} \\ G_{norm} \\ B_{norm} \end{bmatrix} = \begin{cases} \mathrm{VecMin}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{M_{RP}} \times \begin{bmatrix} R_{linear} \\ G_{linear} \\ B_{linear} \end{bmatrix} \right), & if\ M_{RP}\ available \\ \frac{1}{S_{MC}} \times \begin{bmatrix} R_{linear} \\ G_{linear} \\ B_{linear} \end{bmatrix}, & otherwise \end{cases} \tag{7}$$

where:

R_{norm}, G_{norm}, B_{norm} = Normalized color components,
R_{linear}, G_{linear}, B_{linear} = Linearized RGB components as defined in Section 4.5,
S_{MC} = maximum of the elements of **MaxSCL** defined in Section 8.3,
$\mathrm{VecMin}()$ = is a vector min() function working component-wise,

B.4 Scene Adaptive Tone Mapping

B.4.1 Tone Mapping with the N-th Order Curve

Normalized color components are tone mapped using Equation (8):

$$\begin{bmatrix} R_{stm} \\ G_{stm} \\ B_{stm} \end{bmatrix} = \mathrm{VecMin}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, r(s) \times \beta \times \begin{bmatrix} R_{norm} \\ G_{norm} \\ B_{norm} \end{bmatrix} \right) \tag{8}$$

where:

$$r(s) = \begin{cases} \frac{F_N(s)}{s}, & s > 0 \\ 0, & s = 0 \end{cases} \tag{9}$$

$(R_{stm}, G_{stm}, B_{stm}) =$ the tone mapped normalized color components,

$s = \max(R_{norm}, G_{norm}, B_{norm})$,

$F_N(s)$ = the tone mapping function as defined in Section 8.7,

$\beta =$ dimming adjusting factor as shown in Figure B.1 and calculated as follows

$$\beta = \begin{cases} 1 + \left(\frac{\Delta - 1}{T-1}\right) \times (\delta - 1), & 1 \leq \delta < T \\ \Delta, & T \leq \delta \end{cases}$$

When one color component is significantly larger than other color components in the scene, the resulted tone mapped scene can get dimmer compared to creative intent (or image essence). This can be adjusted using the dimming adjusting factor. In Figure B.1, Δ= 2 and *T =16*, and

$$\delta = \frac{\max(\mathbf{MaxSCL})}{\min(\mathbf{MaxSCL})} \qquad (10)$$

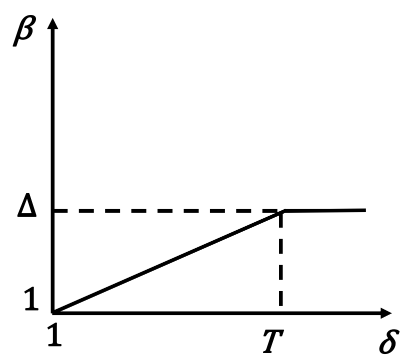

Figure B.1 – Dimming adjusting factor

B.4.2 Example: Tone Mapping with the 4th Order Bezier Curve

In this example, the tone mapping function uses a 4th order Bezier curve and defined in Equation (11)

$$F_4(s) = \begin{cases} \frac{K_F}{K_s} s, & 0 \leq s < K_s \\ K_F + (1 - K_F) B_4\left(\frac{s - K_s}{1 - K_s}\right), & K_s \leq s \leq 1 \end{cases} \qquad (11)$$

where:

$0 < K_s$,

$B_4(t) = 4 \times t \times (1-t)^3 \times P_1 + 6 \times t^2 \times (1-t)^2 \times P_2 + 4 \times t^3 \times (1-t) \times P_3 + t^4.$

The three anchor parameters P_1, P_2, P_3 determine the behavior of the Bezier curve. For instance, Figure B.2 shows the curve for $P_1 = P_2 = P_3 = 1$ and $P_1 = 0,\ P_2 = P_3 = 1$.

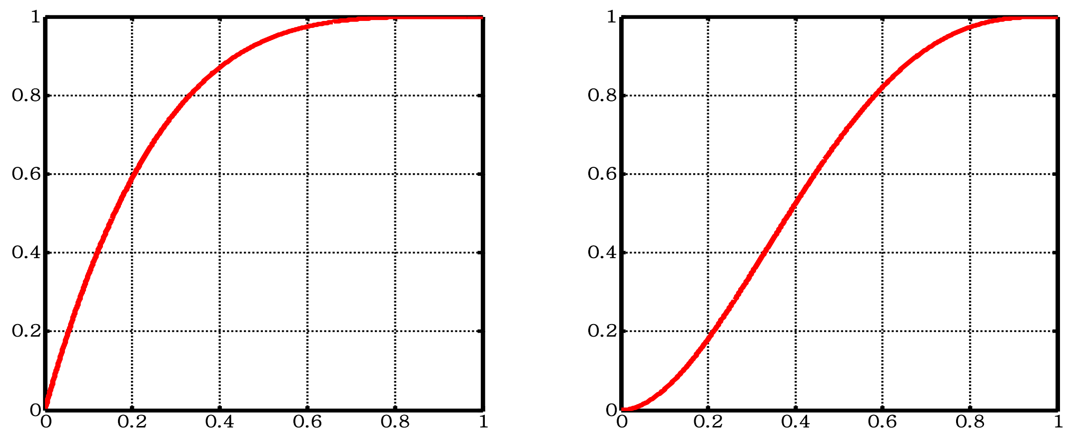

Figure B.2 – Examples of the 4th Oder Bezier curve for $P_1 = P_2 = P_3 = 1$ and $P_1 = 0,\ P_2 = P_3 = 1$

B.5 Scene-based Color Saturation Mapping

B.5.1 Color Saturation Mapping

The saturation mapping compensates the decrease in color saturation in the targeted system display with a smaller color volume. The color saturation is compensated according to the actual luminance difference between the source image and the tone-mapped image.

B.5.2 Color Saturation Mapping Process

Prior to color saturation mapping, color components are converted from RGB to $Y'C_B'C_R'$. After color saturation mapping, they are converted back to RGB.

RGB Conversion to $Y'C_B'C_R'$ is performed using Equations (12) and (13)

$$\begin{bmatrix} R_{stm}' \\ G_{stm}' \\ B_{stm}' \end{bmatrix} = I_{1886}\left(\begin{bmatrix} R_{stm} \\ G_{stm} \\ B_{stm} \end{bmatrix}\right) \tag{12}$$

$$\begin{bmatrix} Y_{stm}' \\ C_{B,stm}' \\ C_{R,stm}' \end{bmatrix} = Q_{2020} \times \begin{bmatrix} R_{stm}' \\ G_{stm}' \\ B_{stm}' \end{bmatrix} \tag{13}$$

where:

Y'_{stm}, $C'_{B,stm}$, $C'_{R,stm}$= Tone mapped color components,

I_{1886} = inverse of component-wise EOTF defined in Rec. ITU-R BT.1886-0,

Q_{2020} = RGB to $Y'C'_BC'_R$ transformation matrix from Rec. ITU-R BT.2020-2, given by

$$Q_{2020} = \begin{bmatrix} 0.2627 & 0.6780 & 0.0593 \\ -0.1396 & -0.3604 & 0.5000 \\ 0.5000 & -0.4598 & -0.0402 \end{bmatrix}.$$

Color saturation mapping is performed using Equation (14):

$$\begin{bmatrix} Y'_{scsm} \\ C'_{B,scsm} \\ C'_{R,scsm} \end{bmatrix} = \begin{bmatrix} Y'_{stm} \\ S_{scsm} \times C'_{B,stm} \\ S_{scsm} \times C'_{R,stm} \end{bmatrix} \tag{14}$$

where:

$$S_{scsm} = \min(f_{SCSM}(s), \eta), \tag{15}$$

$f_{SCSM}()$ = color saturation mapping function (as defined in Annex B.5.3),

$\eta =$ maximum color saturation mapping gain (as defined in Annex B.5.4).

Following color saturation mapping, Y'_{scsm}, $C'_{B,scsm}$, $C'_{R,scsm}$ are converted back to RGB components using Equations (16) and (17).

$$\begin{bmatrix} R_{scsm}' \\ G_{scsm}' \\ B_{scsm}' \end{bmatrix} = M_{2020} \times \begin{bmatrix} Y'_{scsm} \\ C'_{B,scsm} \\ C'_{R,scsm} \end{bmatrix} \tag{16}$$

$$\begin{bmatrix} R_{scsm} \\ G_{scsm} \\ B_{scsm} \end{bmatrix} = E_{1886}\left(\begin{bmatrix} R_{scsm}' \\ G_{scsm}' \\ B_{scsm}' \end{bmatrix}\right) \tag{17}$$

where:

E_{1886} = component-wise EOTF defined in Rec. ITU-R BT.1886-0,

M_{2020} = inverse of Q_{2020} matrix given by

$$M_{2020} = \begin{bmatrix} 1.0000 & 0.0000 & 1.4746 \\ 1.0000 & -0.1646 & -0.5714 \\ 1.0000 & 1.8814 & 0.0000 \end{bmatrix}.$$

B.5.3 Color Saturation Mapping Weight

Color saturation mapping is shown in Figure B.3 and defined by Equation (18):

(18)

$$f_{SCSM}(s) = 1 + \mathbf{ColorSaturationWeight} \times \max\left(0, \frac{\log(M_p \times M_{ML} \times s)}{\log\left(T_p \times T_{ML} \times F_N(s)\right)} - 1\right),$$

where:

$s = \max(R_{norm}, G_{norm}, B_{norm})$,

M_P = Source Normalized Actual Peak Luminance as defined in Annex B.2.1,

T_P = Target Normalized Actual Peak Luminance as defined in Annex B.2.2,

M_{ML} = Maximum Display Mastering Luminance from SMPTE ST 2086,

T_{ML} = **TargetedSystemDisplayMaximumLuminance**,

$F_N(s)$ = Tone mapping function as defined in Section 8.7,

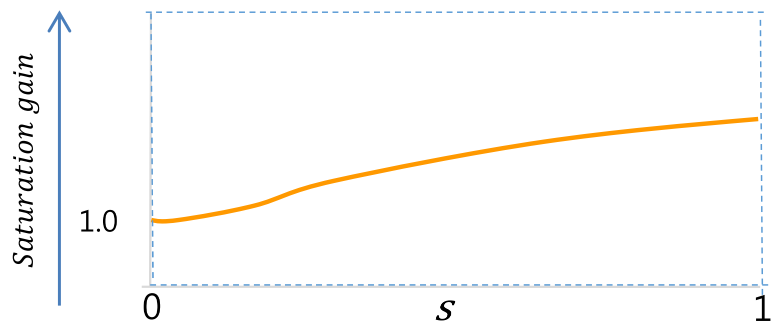

Figure B.3 – Color saturation mapping function

B.5.4 Color Saturation Mapping Maximum Gain

Color saturation gain is limited to a maximum value to avoid clipping when compensated RGB values become negative or larger than unity. The maximum gain, η, satisfies the following conditions:

$$\begin{gathered} 0 \le M_{11} \times Y'_{stm} + \eta(M_{12} \times C'_{B,stm} + M_{13} \times C'_{R,stm}) \le 1 \\ 0 \le M_{21} \times Y'_{stm} + \eta(M_{22} \times C'_{B,stm} + M_{23} \times C'_{R,stm}) \le 1 \\ 0 \le M_{31} \times Y'_{stm} + \eta(M_{32} \times C'_{B,stm} + M_{33} \times C'_{R,stm}) \le 1 \\ 0 \le \eta \le 4. \end{gathered} \tag{20}$$

where:

$[M_{ij}]$ = the elements of M_{2020}.

B.6 Local Pixel Processing in Elliptical Pixel Selector

B.6.1 Introduction

In a Processing Window extended with the elliptical pixel selector, the tone mapping function, as described in Annex B.4.1, operates on the pixels in the intersection of the elliptical pixel selector and the corresponding Processing Window. Local pixel processing primarily targets the pixels inside the internal ellipse and is not intended for the pixels outside of the external ellipse.

To prevent any visible boundary between the pixels selected by the elliptical pixel selector and the region outside of the external ellipse, we introduce a transition weighting function $w(x, y)$ at the position (x, y) by evaluating the normalized sum of distances to the foci of the external ellipse. The transition weighting function is 1 inside the internal ellipse, 0 outside of the external ellipse, and gradually change from 1 to 0 as the considered position moves from the internal ellipse to the external ellipse.

B.6.2 Transition Weighting Function

For pixels selected by the elliptical pixel selector, we rotate the pixel position, (x, y), in a clockwise direction, over the ellipse center by θ. That is, we get

$$\begin{bmatrix} X \\ \mathrm{Y} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \times \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \qquad (21)$$

Then, with simple mathematical manipulation, it can be shown that the normalized summation of the distances from the position (x, y) to the external foci is found as

$$D(x, y) = \sqrt{\frac{(X - x_0)^2}{a_2^2} + \frac{(Y - y_0)^2}{b_2^2}}. \qquad (22)$$

Then the transition weighting function $w(x, y)$ can be designed as shown in Figure B.4. The transition weighting function is composed of a unity part and a curve part, which are separated by the point ($\frac{a_1}{a_2}$,1) as depicted in Figure B.4. The curve part of the transition weighting function can be any monotonically descending curve with continuous derivatives at the end points. $D(x, y) > 1$ when (x, y) is outside the external ellipse, $D(x, y) = 1$ when (x, y) is on the external ellipse, $\frac{a_1}{a_2} < D(x, y) < 1$ when (x, y) is between the internal and the external ellipses, $D(x, y) = \frac{a_1}{a_2}$ when (x, y) is on the internal ellipse, and $D(x, y) < \frac{a_1}{a_2}$ when (x, y) is inside the internal ellipse. An example equation for $w(x, y)$ is given by

$$w(x, y) = \begin{cases} 1, & D(x, y) < \dfrac{a_1}{a_2} \\ 0.5 + 0.5 \times \cos\left(\dfrac{(a_2 \times D(x, y) - a_1) \times \pi}{a_2 - a_1}\right), & \dfrac{a_1}{a_2} \le D(x, y) < 1 \\ 0, & 1 \le D(x, y) \end{cases}$$

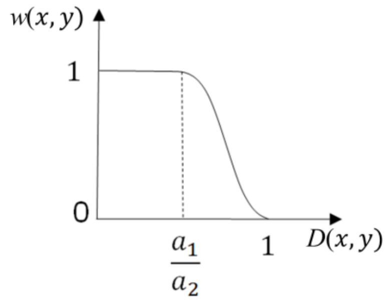

Figure B.4 – Transition Weighting Function

B.6.3 Output Pixel Values for up to Two Elliptical Pixel Selectors

B.6.3.1 Example Methods

Two example methods of pixel processing for up to two elliptical pixel selectors in one image are described in this section.

B.6.3.2 Weighted Average Method

As a first method of rendering, we describe a way of weighted averaging for two pixel selectors in one image. The output of the weighted average method, $\Omega_{WA}(x,y)$, is given as

$$\Omega_{WA}(x,y) = \begin{cases} \Omega_{scsm,0}(x,y), & if\ \sum_{l=1}^{Q} w_l(x,y) = 0 \\ \sum_{j=1}^{Q}\left[w_j(x,y) \times \Omega_{scsm,j}(x,y) + \left(1 - w_j(x,y)\right) \times \Omega_{scsm,0}(x,y)\right] \times \gamma_j(x,y), & otherwise. \end{cases} \quad (24)$$

where:

Q = number of the elliptical pixel selectors in the image ∈ {1,2},

$w_j(x,y)$= transition weighting function for Processing Window j ∈ {1,2},

$\Omega_{scsm,0}(x,y) = (R_{scsm}, G_{scsm}, B_{scsm})$ output of Equation (17) for Processing Window 0 at pixel position (x,y),

$\Omega_{scsm,j}(x,y) = (R_{scsm}, G_{scsm}, B_{scsm})$ output of Equation (17) for Processing Window j ∈ {1,2} at pixel position (x,y).

$\gamma_j(x,y) = \frac{w_j(x,y)}{\sum_{l=1}^{2} w_l(x,y)}$.

B.6.3.3 Layering Method

As a second method of rendering, we describe a way of layering for up to two pixel selectors in one image. Layering gives priority to the outputs of the Processing Windows extended with the elliptical pixel selector. Suppose the priority of the elliptical pixel selectors is given in accordance with the number. That is, Processing Window 1 has the higher priority (top layer) and Processing Windows 2 has the lower priority (bottom layer).

The output of the layering method, $\Omega_L(x,y)$, is given as

$$\Omega_L(x,y) = \begin{cases} w_1(x,y) \times \Omega_{scsm,1}(x,y) + \left(1 - w_1(x,y)\right) \times \Omega_{scsm,0}(x,y) & if\ w_1(x,y) \neq 0,\ else \\ w_2(x,y) \times \Omega_{scsm,2}(x,y) + \left(1 - w_2(x,y)\right) \times \Omega_{scsm,0}(x,y) & if\ w_2(x,y) \neq 0,\ else \\ \Omega_{scsm,0}(x,y) & otherwise. \end{cases} \quad (25)$$

Annex C Measurement of Actual Peak Luminance (Informative)

The actual peak luminance value can be measured by changing the background gray level and inserting a variable-size white patch on the background. Figures C.1 and C.2 show actual peak level of a display while changing the background gray level and the white patch size, and Figure C.3 shows the resulting two-dimensional LUT generated from the measurements.

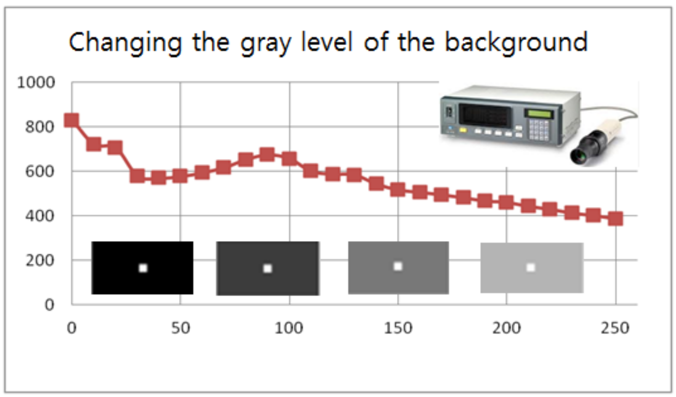

Figure C.1 – Actual peak (in nits) while changing the background gray level

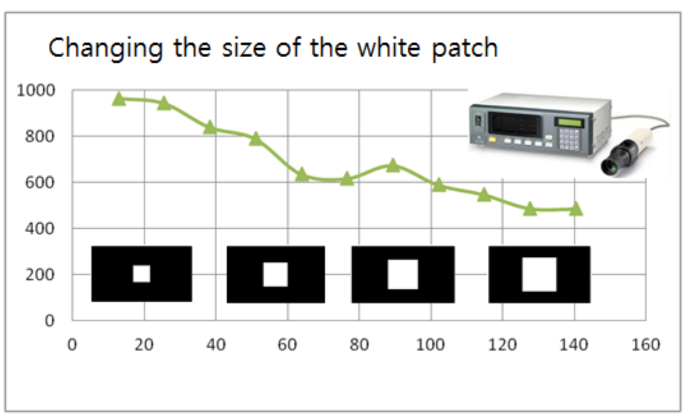

Figure C.2 – Actual peak (in nits) while changing the white patch size

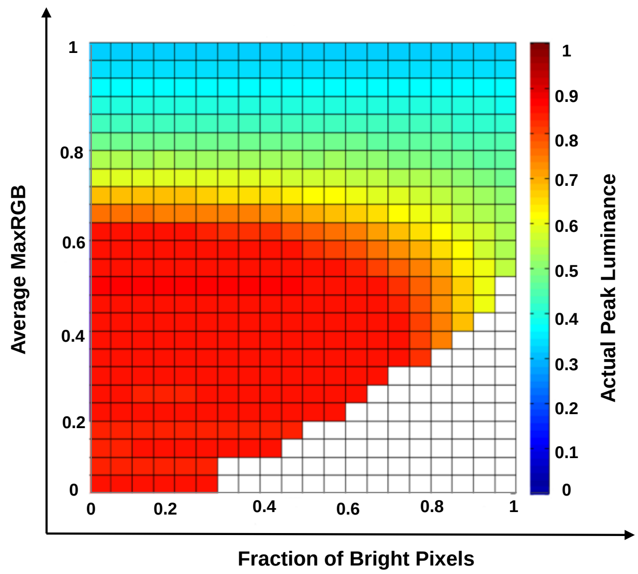

Figure C.3 – Two-dimensional LUT of actual peak luminance

The white squares in Figure C.3 indicate areas where the bright pixels would raise the average MaxRGB level above the corresponding level shown on the Y axis.

Bibliography (Informative)

Calculus and Analytic Geometry, G.B. Thomas and R. L. Finney, Addison-Wesley, 1988.

Recommendation ITU-R BT.1886-0 (3/2011), Reference Electro-optical Transfer Function for Flat Panel Displays used in HDTV Studio Production

Recommendation ITU-R BT.2020-2 (10/2015), Parameter Values for Ultra-high Definition Television Systems for Production and International Programme Exchange