ATSC Candidate Standard: Amendment to A/341:2019 – ST 2094-40

Doc. S34-614r2
28 August 2019

Advanced Television Systems Committee
1776 K Street, N.W.
Washington, D.C. 20006
202-872-9160

The Advanced Television Systems Committee, Inc., is an international, non-profit organization developing voluntary standards and recommended practices for digital television. ATSC member organizations represent the broadcast, broadcast equipment, motion picture, consumer electronics, computer, cable, satellite, and semiconductor industries. ATSC also develops digital television implementation strategies and supports educational activities on ATSC standards. ATSC was formed in 1983 by the member organizations of the Joint Committee on Inter-society Coordination (JCIC): the Electronic Industries Association (EIA), the Institute of Electrical and Electronic Engineers (IEEE), the National Association of Broadcasters (NAB), the National Cable Telecommunications Association (NCTA), and the Society of Motion Picture and Television Engineers (SMPTE). For more information visit www.atsc.org.

Note: The user's attention is called to the possibility that compliance with this standard may require use of an invention covered by patent rights. By publication of this standard, no position is taken with respect to the validity of this claim or of any patent rights in connection therewith. One or more patent holders have, however, filed a statement regarding the terms on which such patent holder(s) may be willing to grant a license under these rights to individuals or entities desiring to obtain such a license. Details may be obtained from the ATSC Secretary and the patent holder.

This specification is being put forth as a Candidate Standard by the TG3/S34 Group on Applications and Presentation. This document is a revision of the Working Draft (S34-614r1) dated 17 July 2019. All ATSC members and non-members are encouraged to review and implement this specification and return comments to cs-editor@atsc.org. ATSC Members can also send comments directly to the TG3/S34 Group. This specification is expected to progress to Proposed Standard after its Candidate Standard period.

Implementers with feedback, comments, or potential bug reports relating to this document may contact ATSC at https://www.atsc.org/feedback/.

Revision History

Version	Date
Candidate Standard approved	28 August 2019
Amendment approved	[date]

Table of Contents

Index of Figures

ATSC Candidate Standard:
Amendment to A/341:2019 – ST 2094-40

1. OVERVIEW

This document describes technology documented in SMPTE ST 2094-40 "Dynamic Metadata for Color Volume Transform — Application #4" [1] which is a technology for the use of dynamic metadata for HDR content. If approved by the ATSC, A/341:2019, "Video – HEVC," ("A/341") would be amended according to the edits described herein.

2. REFERENCES

The following references would be added to A/341.

2.1 Normative References

[1] SMPTE: "Dynamic Metadata for Color Volume Transformation – Application #4," Doc. ST 2094-40 (2016), Society of Motion Picture and Television Engineer, White Plains, NY.

[2] Delta: "Manufacturers Codes for H.32X Terminal" (2017), Delta Information Systems, http://www.delta-info.com/DeltaWeb/Manufacturer_codes/Manucode.pdf.

[3] CTA: "A DTV Profile for Uncompressed High Speed Digital Interfaces", Doc. CTA-861-G (November 2016), Consumer Technology Association, Arlington, VA.

[4] CTA: "Dynamic HDR and ICTCP", Doc. CTA-861.4 (March 2019), Consumer Technology Association, Arlington, VA.

2.2 Informative References

None.

3. DEFINITION OF TERMS.

No new acronyms, abbreviations or terms would be added to A/341.

4. CHANGES TO A/341

In this section of this document, "[ref]" indicates that a cross reference to a cited referenced document that is listed in A/341 would be inserted (or as otherwise described within the square brackets). An actual cross reference to a referenced document listed in this document would be updated with the reference number of the newly added references that would be incorporated into A/341.

4.1 Add a Bullet to Section 6.3.2.2

- The bitstream may contain SEI messages with payloadType value equal to 4. This allows for the optional transmission of the ST 2094-40 metadata message described in [ref to new subsection described below].

4.2 Add a New Subsection under Section 6.3.2.2

The text below would be added to A/341 as a new subsection under Section 6.3.2.2 "PQ Transfer Characteristics." The new subsection would be entitled Section 6.3.2.2.2 "Encoding and Transport of ST 2094-40 Metadata Messages."

The HEVC video bitstream may contain ST 2094-40 metadata messages in order to provide dynamic information about the video signal. ST 2094-40 metadata messages, when present, can provide statistical information about the scene as well as basis OOTF (optical-optical transfer function) data that can guide the tone mapping of displays with lower peak luminance capabilities than the peaks present in the video signal. The information conveyed in the ST 2094-40 metadata message defined in [ref to new Annex A.2 described below] provides carriage for metadata elements defined in SMPTE ST 2094-1 [27] and SMPTE ST 2094-40 [1].

ST 2094-40 metadata messages, when present, shall be as specified in SMPTE ST 2094-40 [1] and encoded and transported as User data registered by a Recommendation ITU-T T.35 Supplemental Enhancement Information (SEI) message per registration authority codes [2] and CTA-861-G [3] Section S.3 as modified by CTA-861.4 [4].

The syntax, semantics, and additional constraints for ST 2094-40 metadata messages shall be as specified in Annex A. Where present, the corresponding NAL unit type shall be set equal to PREFIX_SEI_NUT.

If an ST 2094-40 metadata message is present, the following constraints shall apply:

- The ST 2094-40 metadata message shall be associated with every access unit of the bitstream. If this message is present, it shall only be present once per access unit.
- Mastering Display Color Volume SEI messages (containing SMPTE ST 2086 [23] static metadata) shall be present in the bitstream.

4.3 Add a New Annex to A/341

The below text comprises a new Annex that would be added to A/341. The Annex would be entitled "Metadata Based on SMPTE ST 2094-40."

Annex A Metadata Based on SMPTE ST 2094-40

A.1 GENERAL

This annex specifies the syntax, semantics, and additional constraints of ST 2094-40 metadata messages for use in ATSC 3.0 services as well as the theory of operation for metadata based on SMPTE ST 2094-40. The syntax and semantics are specified in A.2, additional constraints are specified in A.3, and the theory of operation is described in A.4.

A.2 SYNTAX AND SEMANTICS (NORMATIVE)

The syntax of the SEI message carrying the metadata based on SMPTE ST 2094-40 shall be as specified in CTA-861-G [3] Table 148 and the paragraph above Table 148 as modified by CTA-861.4 [4]. The semantics of the metadata based on SMPTE ST 2094-40 shall be as specified in CTA-861-G [3] Section S.3 and Table 149 as modified by CTA-861.4 [4].

> *Note*: Definitions of the metadata items and terms referred to in this section of the document are provided in SMPTE ST 2094-1 [27] and SMPTE ST 2094-40 [1]. A color volume transform method is described in Annex B of SMPTE ST 2094-40 [1].

A.3 ADDITIONAL CONSTRAINTS (NORMATIVE)

The additional constraints specified in CTA-861.4 [4] Section S.4 shall apply.

A.4 THEORY OF OPERATION (INFORMATIVE)

A.4.1 Metadata

The metadata based on SMPTE ST 2094-40 can be considered in three groups. The first is for identification purposes. The second describes a basis OOTF (optical-optical transfer function) for a producer specified target peak luminance. This basis OOTF can be used to construct the guided OOTF, which defines the adaptation to the presentation display. The third describes statistical characteristics of the video signal. Other elements are constrained so that they are, or effectively are, unused.

Metadata elements associated with identification are not used for image processing. These elements include:

- itu_t_t35_country_code
- itu_t_t35_terminal_provider_code
- itu_t_t35_terminal_provider_oriented_code
- application_identifier
- application_mode

Metadata elements associated with the basis OOTF include:

- targeted_system_display_maximum_luminance
- tone_mapping_flag[w]

- knee_point_x[w]
- knee_point_y[w]
- num_bezier_curve_anchors[w]
- bezier_curve_anchors[w][i]

Metadata elements associated with statistical characteristics include:

- maxscl[w][i]
- average_maxrgb[w]
- num_distributions[w]
- distribution_index[w][i]
- distribution_values[w][i]
- fraction_bright_pixels[w]

A.4.2 Basis OOTF

Prior to emission, the metadata associated with the basis OOTF is set. The basis OOTF is the transfer function for presentation at a single peak luminance point. This point is set within the peak luminance range of television sets that are available to viewers and is represented by targeted_system_display_maximum_luminance.

The basis OOTF curve depends on the content within the video images. In general, the basis OOTF compresses the dynamic range in signal ranges with lower information density and preserves contrast and details in signal ranges with higher information density. A well-constructed basis OOTF is reasonably faithful to the look of the original scene and has no discontinuities or sharp slope changes that might introduce visible artifacts into the images. Specific algorithms for creating a basis OOTF are out of scope of this document.

The basis OOTF allows those responsible for content to view the baseline tone mapped results for the target peak luminance level on a mastering monitor for quality control purposes. By providing this same basis OOTF to ST 2094-40-capable displays, these displays receive the same baseline as a common starting point for their individual tone mapping. OOTF curves based on SMPTE ST 2094-40 are composed by a linear part starting from (0, 0) and ending at a knee point (k_x, k_y) and a second part which is the Bezier curve with its anchor points starting with (k_x, k_y) . These two parts are linked together as a smooth and continuous curve to avoid banding artefacts. Figure 4.3.1 shows an example OOTF.

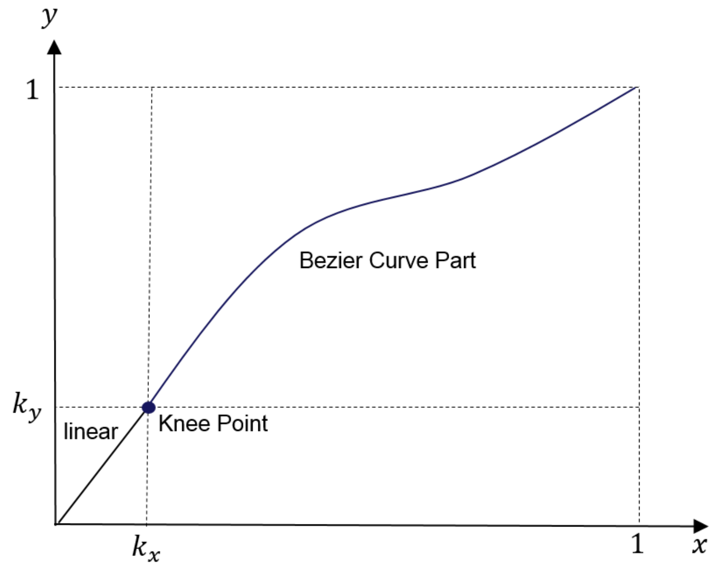

Figure 4.3.1 Example OOTF.

See SMPTE ST 2094-40 [1] for more information regarding the basis OOTF.

A.4.3 Reference Method for Receiver-side Tone Mapping using ST 2094-40 Metadata

A.4.3.1. General

Figure 4.3.2 shows a block diagram of a tone mapping function based on ST 2094-40 metadata.

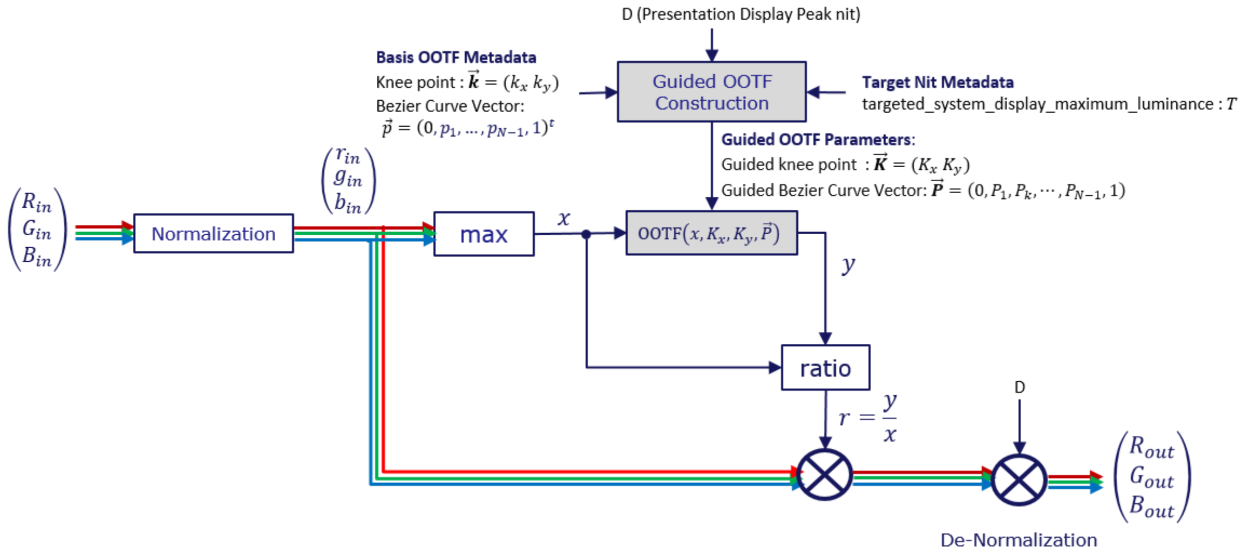

Figure 4.3.2 Tone-mapping system of an ST 2094-40-based device.

Normalization is performed as follows:

The linear RGB signals $(R_{in}, G_{in}, B_{in})^t$ fed to the tone mapping system represent absolute luminance, and must be converted to a single channel input x before applying the guided OOTF, which processes a single input parameter.

The absolute luminance values $(R_{in}, G_{in}, B_{in})^t$ are normalized into the values between 0 and 1 by

$$\begin{pmatrix} r_{in} \\ g_{in} \\ b_{in} \end{pmatrix} = \begin{pmatrix} \min(1, R_{in}/NORM\ ) \\ \min(1, G_{in}/NORM\ ) \\ \min(1, B_{in}/NORM\ ) \end{pmatrix} \qquad (1)$$

where $NORM$ is the normalization factor given by

$$NORM = \max(D, H_M) \qquad (2)$$

in which D is the peak luminance of the presentation display and H_M = **distribution_values[0][8].**

For each pixel, the maximum value of r_{in}, g_{in}, and b_{in} is determined, as represented by x. The value of x is applied to the guided OOTF, producing the resultant value, y. The values of r_{in}, g_{in}, and b_{in} are each scaled by the ratio of y/x.

At the end of the process, the signal is de-normalized based on the peak luminance of the presentation display.

A.4.3.2. Guided OOTF Construction

A.4.3.2.1. General

The guided OOTF is based on the peak luminance of the presentation display and is derived from the basis OOTF. There are three cases:

1) The peak luminance of the presentation display (D) is equal to the target peak luminance (T), in which case the basis OOTF can be used directly;
2) the peak luminance of the presentation display (D) is greater than the target peak luminance (T), in which case the basis OOTF is effectively interpolated with a linear transfer function to create the guided OOTF;
3) the peak luminance of the presentation display (D) is less than the target peak luminance (T), in which case the basis OOTF is extrapolated to create the guided OOTF.

In each case, the goals are to create a transfer function appropriate for the presentation display, be reasonably faithful to the nature of the guided OOTF, and to avoid introducing level or slope discontinuities that might introduce visible artifacts.

A reference method of guided OOTF generation is as follows. This method is shown to produce good results; however, it is possible that other implementations can improve the interpolation and extrapolation processes to produce results with improved detail and faithfulness to the intent of the basis OOTF.

In general, guided OOTF construction is composed of the following two parts with the inputs T (peak luminance of the target display that is obtained with the basis OOTF) and D (peak luminance of the presentation display):

- Guided Knee Point
- Guided Bezier Curve Anchors

They are described in the subsequent sections.

A.4.3.2.2. Guided Knee Point Construction

The construction of the guided knee point can be classified into two cases.

Case I: When $D \leq T$

The guided knee point, $\vec{K} = \begin{pmatrix} K_x \\ K_y \end{pmatrix}$, can be obtained by

$$\vec{K} = (w, \quad 1 - w) \cdot \begin{pmatrix} \vec{k} \\ \vec{K}_0 \end{pmatrix} \tag{3}$$

where $\cdot$ represents the dot product of two vectors, $\vec{k} = \begin{pmatrix} k_x \\ k_y \end{pmatrix}$ is the knee point of the given basis OOTF, $\vec{K}_0$ is a pre-defined constant vector such as $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, and w is the guided knee point mixing parameter as a function of D. There are various ways to design w; however, a linear method is simple and effective. The reference, linear method is as shown in Figure 4.3.3, where D_L is a pre-defined low luminance level, where $D_L \leq T$.

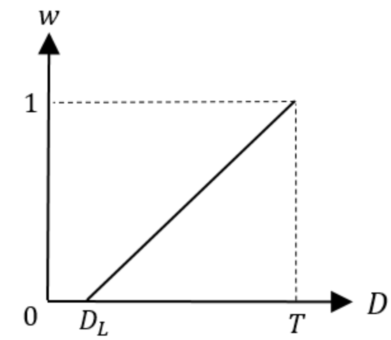

Figure 4.3.3 Example of guided knee point mixing parameter function when $D \leq T$.

Figure 4.3.4 shows the graphical illustration of how the guided knee point is constructed as a function of D as expressed in (3), where the dotted red arrow is the Guide Knee Point trajectory.

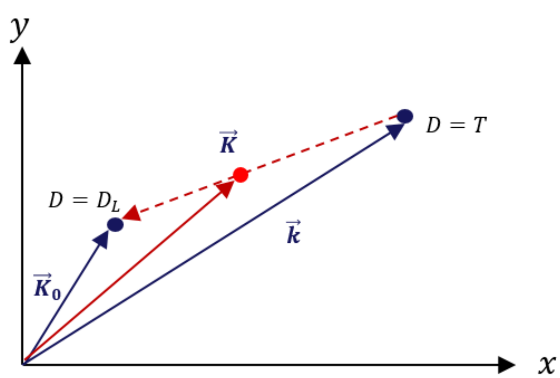

Figure 4.3.4 The guided knee point when $D \leq T$.

Case II: When $T \leq D$

Similar to $D \leq T$, in the case that $T \leq D$, the guided knee point , $\vec{K} = \begin{pmatrix} K_x \\ K_y \end{pmatrix}$ can be obtained by

$$\vec{K} = (w, \quad 1 - w) \cdot \begin{pmatrix} \vec{k} \\ \vec{K}_1 \end{pmatrix} \tag{4}$$

where $\vec{K}_1 = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$, and w is the guided knee point mixing parameter as a function of D which can be designed as shown in Figure 4.3.5.

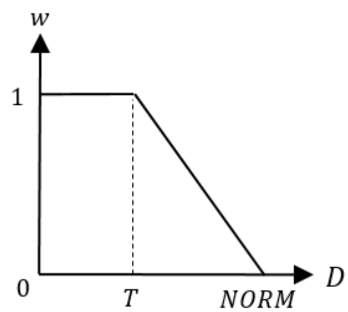

Figure 4.3.5 Example of guided knee point mixing parameter function when $T \leq D$.

Figure 4.3.6 shows the graphical illustration of how the guided knee point is constructed as a function of D as expressed in (4), where the dotted red arrow is the guide knee point trajectory.

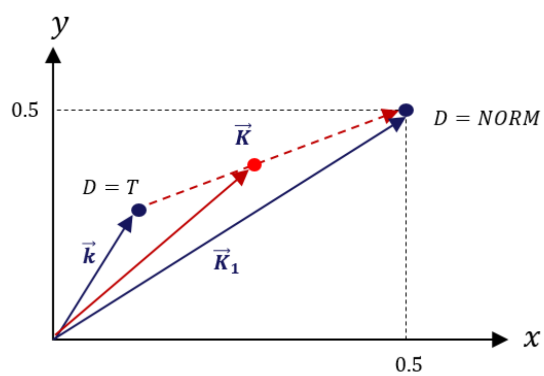

Figure 4.3.6 The guided knee points when $T \leq D$.

Note that

$$\vec{K} = \vec{K}_1 = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \tag{5}$$

when $NORM \leq D$.

A.4.3.2.3. Guided Bezier Curve Vector Construction

Following two properties of the Bezier curve are fundamental to understand the notion behind the guided Bezier curve vector construction.

Property 1: If

$$\vec{P}_L = \begin{pmatrix} 0 \\ \frac{1}{N} \\ \vdots \\ \frac{N-1}{N} \\ 1 \end{pmatrix} \quad (6)$$

then the normalized explicit Bezier curve becomes an identity line. That is, $B_N(\vec{P}_L, t) = t$ and $\vec{P}_L$ is referred to as the identity Bezier curve vector.

Property 2: Linearly adding two Bezier curves in the same order is equivalent to linearly adding their Bezier curve vectors. That is, if we let $B_N(\vec{\alpha}, t)$ and $B_N(\vec{\beta}, t)$ two Bezier curves in N^{th} order characterized by the Bezier curve vectors $\vec{\alpha}$ and $\vec{\beta}$, respectively, such as

$$\begin{cases} B_N(\vec{\alpha}, t) = \sum_{k=0}^{N} C_N^k \, t^k (1-t)^{N-k} \cdot \alpha_k \\ B_N(\vec{\beta}, t) = \sum_{k=0}^{N} C_N^k \, t^k (1-t)^{N-k} \cdot \beta_k \end{cases}, \quad (7)$$

then the following can be easily shown

$$(a, b) \cdot \begin{pmatrix} B_N(\vec{\alpha}, t) \\ B_N(\vec{\beta}, t) \end{pmatrix} = B_N(\vec{\gamma}, t) \quad (8)$$

where

$$\vec{\gamma} = (a, b) \cdot \begin{pmatrix} \vec{\alpha} \\ \vec{\beta} \end{pmatrix} \quad (9)$$

Similar to the guided knee point construction, the guided Bezier curve vector construction from a given basis Bezier curve vector can also be classified into two cases. The basic idea behind guided Bezier curve construction is to interpolate Bezier curve with the Bezier curve of the basis OOTF and a pre-determined boundary Bezier curve as a function of D.

Case I: When $D \leq T$

In the case of $D \leq T$, the guided Bezier curve, $B_N(\vec{P}, t)$, can be found as

$$B_N(\vec{P}, t) = (u, 1-u) \cdot \begin{pmatrix} B_N(\vec{p}, t) \\ B_N(\vec{P}_0, t) \end{pmatrix} \quad (10)$$

where $\vec{p} = \vec{p} = (0,\ p_1,\ \ldots,\ p_{N-1},\ 1)^t$ is the Bezier Curve vector of basis OOTF, and $\vec{P}_0$ is a pre-defined Bezier curve vector such as, but not limited to, $(1,1,\cdots,1)^t$, and u is the control parameter as a function of D which can be designed as shown in Figure 4.3.7 but not limited to.

By Property 2, the guided Bezier curve vector, $\vec{P}$, of $B_N(\vec{P}, t)$ can be calculated as

$$\vec{P} = (u, \quad 1-u) \cdot \begin{pmatrix} \vec{p} \\ \vec{P}_0 \end{pmatrix} \tag{11}$$

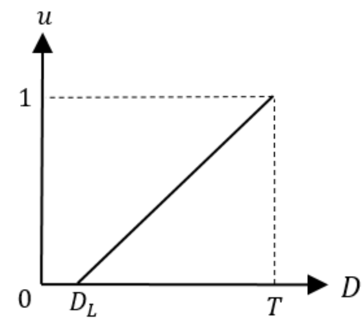

Figure 4.3.7 Example of Bezier curve vector mixing coefficient function for $D \leq T$.

Case II: When $T \leq D$

In the case of $T \leq D$, the guided Bezier curve, $B_N(\vec{P}, t)$, can be found as

$$B_N(\vec{P}, t) = (u, 1-u) \cdot \begin{pmatrix} B_N(\vec{p}, t) \\ B_N(\vec{P}_L, t) \end{pmatrix} \tag{12}$$

where $\vec{p} = \vec{p} = (0,\ p_1,\ \ldots,\ p_{N-1},\ 1)^t$ is the Bezier Curve vector of the basis OOTF, and $\vec{P}_L$ is Identity Bezier curve introduced in (6), and u is the mixing parameter as a function of D which can be designed as shown in Figure 4.3.8 but not limited to.

By Property 2, the guided Bezier curve vector, $\vec{P}$, of $B_N(\vec{P}, t)$ can be calculated as

$$\vec{P} = (u, \quad 1-u) \cdot \begin{pmatrix} \vec{p} \\ \vec{P}_L \end{pmatrix} \tag{13}$$

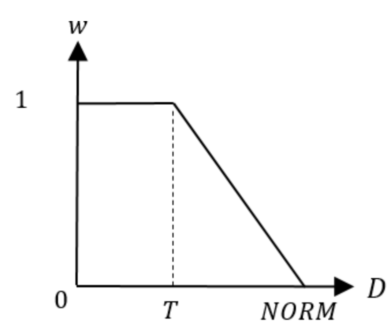

Figure 4.3.8 The Bezier curve anchor control determining function for $T \leq D$.

Note that we have

$$\vec{P} = \vec{P}_L \tag{14}$$

when $NORM \leq D$.

Note: By merging (5) and (14), the guided OOTF becomes the identity line as depicted in Figure 4.3.9 when $NORM \leq D$, which is the case when the dynamic range of the presentation display is larger than that of the incoming HDR scene.

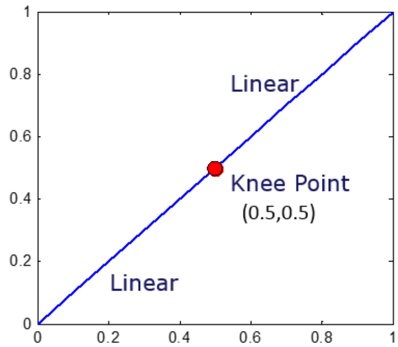

Figure 4.3.9 The guided Bezier curve when $NORM \leq D$.

A.4.3.2.4. Slope Continuity Condition at the Knee Point

After the guided Knee point and Bezier curve vector are constructed, one last step is required to adjust the Bezier curve vector to ensure the slope continuity at the guided knee point. Note that the linear part and the Bezier curve part of a guided OOTF curve are jointed by the knee point. Discontinuity between the two parts at the knee point may lead to banding artifacts in the tone-mapped signal. In any implementations of the ST 2094-40 tone mapping system, the condition for continuity of the slopes at the knee point must be satisfied. Namely, the tangent of the Bezier curve part at the knee point should be equal to the slope of the linear part, which can be expressed as

$$\frac{d}{dx}\left(K_y + \left(1 - K_y\right) \cdot B_N\left(\vec{P}, \frac{x - K_x}{1 - K_x}\right)\right)\Bigg|_{x=K_x} = \frac{K_y}{K_x} \tag{15}$$

By solving (15), the condition for continuity of the slopes at the knee point is given as

$$P_1 = \frac{1}{N} \cdot \frac{K_y}{K_x} \cdot \frac{1 - K_x}{1 - K_y} \tag{16}$$

Figure 4.3.10 shows some examples of various guided OOTF curves that satisfy the condition for continuity of the slopes at the knee point. As can be seen in Figure 4.3.10, the OOTF curves are continuous and will not lead to banding artifacts.

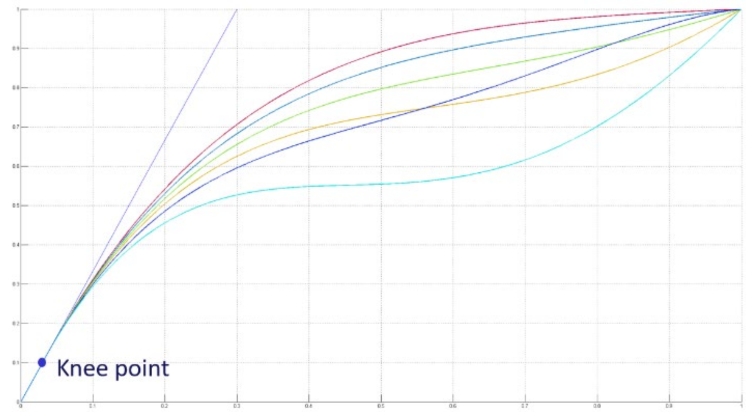

Figure 4.3.10 Examples of the continuity of different guided OOTF curves at the knee point.

A.4.4 Statistical Characteristics

Statistical measurements of each video picture is a straightforward mathematical process that is performed before emission. See tsconSMPTE ST 2094-40 [1] for details.

While it is possible for a receiver to make these same measurements on a frame-by-frame basis, it is more efficient to perform this process upstream. These video picture-based measurements cannot be completed until the full picture is available, potentially requiring frame buffers if calculated in the receiver. However, when measured before the emission, the calculations do not necessarily add latency. For instance, the measurement can be calculated in parallel with HEVC encoding, which necessarily adds latency of more than one picture of video, and the metadata is then inserted after encoding is completed. In addition, the receiver would not be able to make forward-looking, scene-based measurements.

Television sets, even with extremely high peak luminance capabilities, do not necessarily employ a neutral transfer function. The user might have selected a picture preset (Dynamic Mode) with high contrast based on personal preferences. An ambient light sensor might indicate a very bright environment, in which case low- and mid-tones might be lifted in order to ensure that the image is clearly visible. The manufacturer might also employ an adjustment to compensate for characteristics of the physical display device.

The statistical characteristic metadata can aid the television set in making the above adjustments. In general, a good implementation would avoid compressing the dynamic range in signal ranges where there is high information density and would instead compress the dynamic range in signal ranges with low information density. How this is achieved is left to the manufacturer and is not specified in this document.

End of Document