



CST-RT021-annexe:2016

Annexe sur les transformations couleur

Groupe de travail CST-RT-021-MFFW

Table des matières

Table des matières	2
I. Introduction	3
II. Termes et définitions	3
A. Espace couleur de référence	3
B. Valeur de composante	3
C. Valeur de composante linéaire	4
D. Unités de codage	4
E. Valeur de codage	4
F. Mise à l'échelle en XYZ	4
III. Conversion directe	4
A. Formule de conversion	4
B. Calcul de la matrice de conversion	5
C. Facteur de mise à l'échelle	5
D. Mise à l'échelle pour un blanc hors du « white gamut »	6
IV. Conversion inverse	6
A. Normalisation	7
B. Conversion dans un espace RVB	7
C. Facteur d'échelle	7
V. Précision	7
A. Compression sans pertes et précision	7
B. Précisions requises	7
C. Normalisation des matrices	8
D. Précision des matrices et conditions de comparaison	8
VI. Exemples	9
A. Exemple de calcul de matrice à partir d'un élément de film négatif numérisé (reprise de la RP 177). 9	
B. Exemple de calcul de matrice à partir de données numériques (DSM)	10
VII. Références	11

I. Introduction

Ce document décrit le traitement colorimétrique impliqué par le standard SMPTE ST 2067-40:2016, Interoperable Master Format – Application #4 Cinema Mezzanine. Il concerne la conversion des valeurs colorimétriques obtenues par la numérisation de films cinématographiques ou à l'issue d'une post-production en numérique vers l'espace XYZ linéaire (Note Importante¹).

Cet espace permet de garder les valeurs colorimétriques analysées dans tous les espaces trichromes possibles. Il permet également de convertir les données vers tout espace trichrome de travail présent ou futur.

Afin d'obtenir le meilleur résultat, il est important de garder toutes les informations sur les espaces trichromes de numérisation ou de calcul et sur les fonctions de transfert éventuelles.

Dans le cas de numérisation de films, il est souhaitable que les filtres d'analyse soient adaptés aux caractéristiques couleurs des pellicules analysées.

II. Termes et définitions

A. Espace couleur de référence

Pour l'IMF App #4 l'espace couleur de référence est le même que celui défini pour le cinéma numérique. Il s'agit de l'espace de contrôle de la reproduction. C'est-à-dire que l'apparence du contenu numérique est contrôlée dans le cadre d'une projection cinéma numérique standard.

Le blanc de référence est défini comme dans le document SMPTE 431-1 [1] :

$x = 0,314$; $y = 0,351$

Les couleurs primaires de référence sont définies comme dans le document SMPTE 431-2 [2] :

	x	y
Rouge	0,68	0,32
Vert	0,265	0,69
Bleu	0,15	0,06

B. Valeur de composante

Dans le cas d'un film numérisé, la valeur de composante est un nombre qui est le résultat de l'analyse de la pellicule film à travers les filtres d'analyse de l'appareil de numérisation.

Dans le cas d'une post-production en numérique la valeur de composante est un nombre résultant des calculs.

Cette valeur est exprimée dans l'intervalle défini par la résolution de l'appareil de numérisation ou par la résolution dans laquelle on donne le résultat des calculs en numérique. Par exemple, pour 14 bits, elle est comprise dans l'intervalle [0 ; 16385], et pour 16 bits elle est comprise dans l'intervalle [0 ; 65535].

¹ Attention : ne pas confondre avec les valeurs non linéaires X'Y'Z' retenues pour le cinéma numérique [10]. Le codage XYZ utilise le même système colorimétrique CIE 1931 [9].

C. Valeur de composante linéaire

Avant la conversion de l'espace couleur d'analyse vers l'espace des primaires XYZ on doit obtenir des valeurs linéaires afin d'assurer une conversion couleur correcte.

Dans le cas de la numérisation d'un film, on doit obtenir à partir de chaque composante (R V B) une valeur linéairement proportionnelle à l'intensité (luminance) de celles-ci. Suivant les cas la valeur de composante linéaire peut être obtenue directement lors de la numérisation ou bien peut provenir d'un calcul de linéarisation à partir des valeurs numérisées non linéaires.

Dans le cas d'une post-production en numérique, si le standard suivi pour la production spécifie un gamma ou une fonction de transfert, la valeur de composante linéaire est le résultat de la conversion dans le domaine linéaire par application de la loi inverse.

D. Unités de codage

Les valeurs de codage sont toujours linéaires et positives. Elles sont le résultat de la conversion du contenu vers les couleurs primaires virtuelles XYZ.

Les spécifications du format IMF Application #4 requièrent que les valeurs de codage XYZ représentent les valeurs physiques de leurs luminances respectives, dans l'espace de référence choisi. Le système international d'unités (SI) spécifie le candela par mètre carré (cd/m²) pour la luminance [3]. En conséquence les valeurs de codage doivent représenter des candelas par mètre carré ou des sous-unités telles que les milli candelas par mètre carré.

E. Valeur de codage

Dans le format actuel la valeur de codage est un nombre compris dans l'intervalle [0 ; 65535].

Les unités pour ces valeurs sont donc des milli candelas par mètre carré. Par exemple les valeurs de codage pour le blanc de référence seront les suivantes :

X=42940 ; Y=48000 et Z=45812

F. Mise à l'échelle en XYZ

Pendant la conversion en composantes XYZ il faut s'assurer que les valeurs finales représentent les valeurs de luminance mesurables lors d'une projection de contrôle. Il est nécessaire de faire une mise à l'échelle pour que les valeurs XYZ représentent ces luminances. Ceci garantit que l'utilisateur du contenu saura ce que celui-ci représente et guidera la conversion vers n'importe quel espace couleur d'exploitation.

III. Conversion directe

A. Formule de conversion

Chaque valeur de composante linéaire est normalisée dans l'intervalle [0 ; 1]. Si les valeurs ont été obtenues dans une résolution de N bits

$$L_R = \frac{R}{(2^N - 1)}$$

$$L_V = \frac{V}{(2^N - 1)}$$

$$L_B = \frac{B}{(2^N - 1)}$$

L_R , L_V et L_B sont les valeurs dans l'intervalle [0 ; 1]

R , V et B sont les valeurs entières linéaires.

On obtient ensuite des valeurs X , Y et Z dans l'intervalle [0 ; 1] en appliquant la matrice de conversion colorimétrique :

$$\begin{bmatrix} V_X \\ V_Y \\ V_Z \end{bmatrix} = |M| \times \begin{bmatrix} L_R \\ L_V \\ L_B \end{bmatrix}$$

V_X , V_Y et V_Z sont les valeurs obtenues dans l'intervalle [0 ; 1]

$|M|$ est une matrice à 3 lignes 3 colonnes.

Les valeurs de codage sont obtenues en faisant une mise à l'échelle adaptée, soit :

$$L_X = \text{Facteur} * V_X$$

$$L_Y = \text{Facteur} * V_Y$$

$$L_Z = \text{Facteur} * V_Z$$

La meilleure stratégie est d'intégrer mise à l'échelle et conversion dans la matrice de conversion en multipliant cette dernière par le coefficient de mise à l'échelle. Soit :

$$\begin{bmatrix} L_X \\ L_Y \\ L_Z \end{bmatrix} = \text{Facteur} * |M| \times \begin{bmatrix} L_R \\ L_V \\ L_B \end{bmatrix}$$

On obtient ensuite les valeurs de codage sur des entiers 16 bits en multipliant par 65535 et en arrondissant :

$$X = \text{INT}[65535 * L_X]$$

$$Y = \text{INT}[65535 * L_Y]$$

$$Z = \text{INT}[65535 * L_Z]$$

B. Calcul de la matrice de conversion

La matrice de conversion $|M|$ se calcule à partir des coordonnées colorimétriques (x,y) (espace Yxy CIE 1931) des trois primaires et du blanc de référence de l'appareil de numérisation.

Dans le cas de films produits en numérique les coordonnées colorimétriques sont définies dans le standard suivi pour la production.

La méthode de calcul est décrite dans la section 3.3 du document SMPTE Recommended Practice RP 177-1993 "Derivation of Basic Television Color Equations" [4].

C. Facteur de mise à l'échelle

L'évaluation du facteur d'échelle est faite en calculant la valeur Y du blanc maximum du contenu. On multiplie pour cela le vecteur composé des valeurs de blanc maximum en RVB par la matrice calculée précédemment :

$$\begin{bmatrix} B_X \\ B_Y \\ B_Z \end{bmatrix} = |M| \times \begin{bmatrix} \text{Blanc}_R \\ \text{Blanc}_V \\ \text{Blanc}_B \end{bmatrix}$$

La valeur B_Y multipliée par 65535 doit donner 48000 (48 cd/m^2)². Le facteur de mise à l'échelle est donc :

$$\text{Facteur} = \frac{48000}{B_Y * 65535}$$

On peut intégrer la mise à l'échelle à la conversion en une fois en multipliant la matrice par ce scalaire (donc chaque terme de la matrice par ce facteur).

Les spécifications de cinéma numérique autorisent cette valeur maximum de luminance de blanc (48 cd/m^2) pour tous les blancs inscrits dans un gamut particulier, c'est-à-dire dans le triangle limité par le blanc D65, le blanc de référence (DCI)³ et le blanc D55.

	x	y	X	Y	Z
white-1 (D65)	0,3127	0,329	45622	48000	52275
white-2 (DCI)	0,314	0,351	42940	48000	45812
white-3 (D55)	0,3324	0,3474	45927	48000	44242

Pour tout blanc en dehors de ce gamut il est recommandé de suivre une procédure particulière.

D. Mise à l'échelle pour un blanc hors du « white gamut »

Le blanc maximum de l'espace couleur d'origine peut se trouver hors du triangle délimité par les trois blancs D65 DCI et D55. Ceci peut se produire essentiellement dans le cas de numérisation d'une pellicule négative ou à la rigueur d'une pellicule intermédiaire. Ce sera le cas si l'appareil de numérisation utilise des primaires non conventionnelles adaptées aux pellicules à numériser.

Le standard Cinéma Mezzanine a pour but de conserver les œuvres cinématographiques telles qu'elles ont été projetées, c'est-à-dire étalonnées. Il est recommandé donc dans le cas considéré de faire la conversion dans l'espace couleur du cinéma numérique afin de vérifier ou réaliser l'étalonnage recherché. La conversion dans le standard se fera à partir des données étalonnées dans le standard cinéma numérique qui est par ailleurs l'espace de référence.

IV. Conversion inverse

Afin d'étalonner les contenus pour une exploitation dans un canal de distribution donné les données XYZ seront converties vers des primaires RVB.

² La valeur maximum est bien 48 (ici 48000) et non 52,37. Les équations décrites dans les documents DCI Digital Cinema System Specifications (v-1.2) et SMPTE RP431-2-2011 utilisent la valeur de 52,37 pour l'encodage X'Y'Z' comme spécifié au chapitre 7.7 du document 431-2. Cette valeur est nécessaire car les valeurs de X et Z peuvent être supérieures à Y lors de l'utilisation d'autres points blancs de mastering. La luminance maximum Y est bien toujours de 48 cd/m^2 . Le codage des millicandela sur 16 bits évite ce problème potentiel.

³ Attention les documents DCI Digital Cinema System Specifications (v-1.2, note 4, page 28) et SMPTE 431-2-2011 (paragraphe 7.7) mentionnent le point blanc D61. Le texte dit précisément qu'il s'agit de pouvoir « inclure » des points blancs de mastering tels que les illuminants D55 D61 ou D65. L'illuminant D61, utilisé parfois comme point blanc de mastering est « sur » la ligne reliant les D55 et D65. Le « White Gamut » est bien défini par le triangle limité par les trois blancs listés dans le tableau, D55, White ref (DCI), et D65.

A. Normalisation

On normalise les valeurs XYZ dans l'intervalle [0 ; 1]

$$L_X = \frac{X}{65535}$$

$$L_Y = \frac{Y}{65535}$$

$$L_Z = \frac{Z}{65535}$$

B. Conversion dans un espace RVB

On utilise en suite une matrice de conversion adaptée aux composantes RVB avec lesquelles le travail d'étalonnage sera effectué avec le facteur d'échelle correspondant. Cette matrice de conversion est l'inverse de la matrice calculée précédemment uniquement dans le cas où les primaires RVB d'analyse et d'exploitation coïncident. Dans le cas contraire une matrice de conversion directe doit être calculée qui correspond à l'espace final RGB et sa matrice inverse doit être utilisée ici.

C. Facteur d'échelle

La façon la plus simple de calculer le facteur d'échelle pour cette opération inverse est de partir de la conversion directe. Cela conduit à calculer la matrice de conversion « directe » correspondant au passage de l'espace couleur d'arrivée à l'espace XYZ (voir III-B), puis à calculer le facteur d'échelle correspondant (voir III-C). L'inverse de ce nombre est alors le facteur d'échelle à appliquer.

V. Précision

A. Compression sans pertes et précision

Le standard Cinéma Mezzanine utilise une compression sans pertes afin de minimiser l'espace nécessaire au stockage des données de chaque image tout en préservant la qualité de l'original. Néanmoins la conversion vers l'espace XYZ et le retour dans un espace RVB d'exploitation impliquent une possibilité de distorsion des valeurs numériques. Il est rappelé que cet espace XYZ permet de garder les valeurs colorimétriques analysées dans tous les espaces trichromes possibles. Il permet également de convertir les données vers tout espace trichrome de travail présent ou futur.

Il n'y a pas de possibilité de comparaison pixel à pixel si les espaces couleurs RVB de départ et d'exploitation sont différents. Une différence peut être constatée en comparant pour chaque pixel les valeurs numériques d'une image originale et de l'image convertie au préalable dans le standard et restaurée ensuite dans le même espace couleur que celui de départ.

Le raisonnement sur l'estimation des erreurs relatives appliquée à ces transformations ainsi que des exemples sont détaillés dans une conférence SMPTE de 2014 [5].

B. Précisions requises

Dans les conversions, directe et inverse, on passe par des opérations de linéarisation, normalisation et des multiplications ou divisions par 65535. L'utilisation, pour ces calculs, de variables flottantes avec la plus grande précision disponible permet de minimiser les différences. Les calculs de l'erreur relative, détaillés dans la conférence de 2014, montrent que celle-ci est fixe pour le codage en virgule flottante et variable pour le codage en entiers. Dans ce dernier cas l'erreur relative est plus élevée pour les faibles valeurs. Une simulation numérique montre en effet que seules les très faibles valeurs sont susceptibles d'être modifiées.

La plus grande incertitude, au sens statistique, provient des matrices de conversion. La meilleure approche est donc de calculer les matrices de conversion avec une très grande précision.

Le document SMPTE RP 177, qui est la référence pour le calcul des matrices de conversion couleur, recommande de mener les calculs avec une précision de 10 digits⁴ et d'arrondir les résultats à 4 places décimales ou 4 digits selon ce qui est le plus précis.

Comme le stockage intermédiaire des valeurs dans le standard fait appel à des entiers sur 16 bits, il est plutôt raisonnable de chercher une plus grande précision. Une spécification de 6 (six) places décimales pour le résultat des matrices est recommandée ; le raisonnement conduisant à cette recommandation est détaillé dans la suite. Les calculs devront être menés avec une plus grande précision. Les variables flottantes en simple précision⁵ (format IEEE 754 32 bits [6]– aussi appelé **float** en C/C++⁶) couvrent 7 places décimales ; il est recommandé de mener les calculs en variables décimales à double précision (format IEEE 754 sur 64 bits – aussi appelé **double** en C/C++), et d'arrondir les résultats sur 6 places décimales.

C. Normalisation des matrices

Le document RP 177 cité plus haut mentionne (paragraphe 3.3.8) une normalisation éventuelle de la matrice de conversion après l'application de l'arrondi à 4 places décimales. Ceci est nécessaire pour que la somme donnant la luminance après l'opération d'arrondi soit égale à 1,0. Comme les facteurs d'échelle sont calculés en fonction de la luminance de référence et combinés avec la matrice cette normalisation n'est pas nécessaire ici.

D. Précision des matrices et conditions de comparaison

Les valeurs XYZ doivent être codées sur des entiers 16 bits avec des valeurs maximales de l'ordre de 48000. Une précision de 5 places décimales permet donc de réduire suffisamment les erreurs d'arrondi. Cependant la récupération des séquences implique que la conversion soit faite en sens inverse. Il est donc important que la multiplication de la matrice directe par la matrice inverse donne des résultats qui soient précis à 5 places décimales. Il est recommandé par conséquent d'arrondir les matrices directes et inverses à 6 places décimales.

Il est possible de comparer pour chaque pixel les valeurs numériques d'une image originale et de l'image convertie au préalable dans le standard et restaurée ensuite dans le même espace couleur que celui de départ. Si toutes les précautions sont prises cette vérification ne pourra donner des valeurs identiques que pour les valeurs les plus élevées. Il est donc recommandé de faire cette vérification pour des valeurs RGB élevées (par exemple supérieures à la moitié du maximum). Il est important dans ce cas que les trois composantes aient des valeurs élevées ce qui exclut les couleurs saturées.

⁴ Digit : puissance de 2.

⁵ La précision des variable float, appelée « unit round-off » dépend de la taille de la mantisse. Celle-ci est de 23 bits pour les float simple précision ce qui donne 2^{-24} , environ 6×10^{-8} , c'est-à-dire 7 places décimales seulement.

⁶ La taille exacte des variables **float** (32 ou 64 bits) dépend du compilateur.

VI. Exemples

A. Exemple de calcul de matrice à partir d'un élément de film négatif numérisé (reprise de la RP 177).

Cet exemple reprend le déroulement des opérations décrit dans le document SMPTE RP 177. Il détaille les opérations de calcul de la matrice de conversion dans un cas de primaires non conventionnelles (voir III-D). Ceci n'inclut pas la vérification ou conformation de l'étalonnage.

On suppose ici que l'illuminant du scanner est conforme à la spécification « status M » de l'ISO 5-3 :2009 [7] pour la mesure des densités de matériaux photographiques négatifs.

Un calcul d'après les données spectrales de cet illuminant donne les coordonnées colorimétriques du blanc de référence $x = 0,4231$ et $y = 0,4172$.

On suppose également que les filtres d'analyse de la négative sont conformes à la recommandation SMPTE RP180 « Conditions spectrales définissant les densités de tirage pour les négatives et les films intermédiaires » [8]. Les coordonnées colorimétriques des filtres sont calculées d'après les données spectrales de ces filtres définies dans la recommandation RP180 :

	x	y
Rouge	0,7248	0,2752
Vert	0,2013	0,7567
Bleu	0,1613	0,0144

Le calcul de la matrice de conversion est obtenu d'après la recommandation SMPTE RP 177 :

On calcule d'abord les valeurs de z ($z = 1 - x - y$) ce qui donne la matrice des primaires |P| suivante :

	Rouge	Vert	Bleu
X	0,7248	0,2013	0,1613
Y	0,2752	0,7567	0,0144
Z	0,0	0,042	0,8243

On calcule l'inverse de cette matrice, ce qui donne la nouvelle matrice |Pinv| :

1,527857845	-0,390230255	-0,292155956
-0,556197395	1,46486872	0,083247034
0,028339549	-0,074638464	1,208908922

On calcule le vecteur résultant de la multiplication de cette matrice par le vecteur blanc normalisé |Bn| obtenu comme suit :

x blanc / y blanc	1,014141898
y blanc / y blanc = 1	1,0
z blanc / y blanc	0,382790029

Le produit |Pinv| * |Bn|, de la matrice primaires inverse par le vecteur blanc normalisé donne un nouveau vecteur à partir duquel on construit la matrice diagonale suivante |Mdiag| :

1,047400014	0,0	0,0
0,0	0,932671772	0,0
0,0	0,0	0,416860141

La matrice finale qui permet la conversion des valeurs RVB vers les valeurs XYZ est le résultat du produit de la matrice des primaires par cette matrice diagonale $|P| * |Mdiag|$.

Ceci donne la matrice de conversion :

0,7591555298	0,1877468278	0,0672395408
0,2882444837	0,7057527302	0,0060027860
0,0000000000	0,0391722144	0,3436178143

Soit en arrondissant à 6 places décimales :

0,759156	0,187747	0,067240
0,288244	0,705753	0,006003
0,000000	0,039172	0,343618

B. Exemple de calcul de matrice à partir de données numériques (DSM)

On suppose que l'étalonnage de sources numériques a été fait en utilisant le projecteur de référence de cinéma numérique.

Le blanc de référence est défini dans le document SMPTE 431-1 :

$x = 0,314$, $y = 0,351$

Les primaires sont définies dans le document SMPTE 431-2 :

	x	y
Rouge	0,68	0,32
Vert	0,265	0,69
Bleu	0,15	0,06

On suppose également que les données sont exprimées en 16 bits linéaires.

Calcul de la matrice de conversion d'après RP 177 :

0,4451698156	0,2771344092	0,1722826698
0,2094916779	0,7215952542	0,0689130679
0,0000000000	0,0470605601	0,9073553944

Soit en arrondissant à 6 places décimales :

0,445170	0,277134	0,172283
0,209492	0,721595	0,068913
0,000000	0,047061	0,907355

Le facteur d'échelle doit être tel que la valeur Y correspondant au blanc maximum soit égale à 48000. La valeur, affichée ici avec 9 décimales, avant multiplication par 65535 est la suivante :

0,73243305.

VII. Références

- [1] SMPTE, «SMPTE 431-1-2006 Screen Luminance Level, Chromaticity and Uniformity,» SMPTE, 2006.
- [2] SMPTE, «RP 431-2-2011 D-Cinema Quality Reference Projector and Environment,» SMPTE, 2011.
- [3] Bureau International des Poids et Mesures, «The International System of Units - 8th Edition,» BIPM, 2006.
- [4] SMPTE, «RP 177-1993 Derivation of Basic Television Color Equation,» SMPTE, 1993.
- [5] F. Helt et V. La Torre, «Quality Assessment Framework for Color Conversions and Perception,» chez *SMPTE 2014 Annual Technical Conference*, Los Angeles, 2014.
- [6] «IEEE 754,» 3 Juin 2016. [En ligne]. Available: https://fr.wikipedia.org/wiki/IEEE_754. [Accès le 18 Octobre 2016].
- [7] ISO, «ISO 5-3:2009 Photography and graphic technology -- Density measurements -- Part 3: Spectral conditions,» ISO, 2009.
- [8] Society of Motion Picture and Television Engineer, «SMPTE RP 180 Spectral Conditions Defining Printing Density in Motion-Picture Negative and Intermediate Films,» SMPTE, revised 2006.
- [9] International Commission on Illumination (CIE), «CIE 15-2004 - Colorimetry 3rd Edition,» CIE, 2004.
- [10] Digital Cinema Initiatives, LLC, «Digital Cinema System Specifications Version 1.2,» Digital Cinema Initiatives, LLC, 2008.