Development of XYZ/sRGB–SCID and Color Gamut Mapping

Koichi Sakamoto and Hitoshi Urabe

FUJI PHOTO FILM Co., Ltd., Saitama Japan
E-mail: sakanioto@den.fujifilrn.co.jp

Abstract

The new work items for ISO/TC130/WG2 include preparation of new standard images to cope with the future multi-media prevalence. The new standard images consist of an Lab–SCID set prepared by the Switzerland team and an XYZ/sRGB–SCID by the Japanese team. The paper describes the definition of XYZ/sRGB–SCID color space and the preparation of standard images.

XYZ/sRGB–SCID has been developed particularly for the monitor display of natural and synthetic images. The color space defined by XYZ/sRGB–SCID has a high degree of consistency from image input to monitor output: such a desirable feature has been achieved by adopting the ITU-R BT.709-3 standard as the image capturing norm and sRGB as the monitor display norm. As the two standards are both based on D65 light source, the present standard pictures imply those in a D65 color space.

The image data for a natural picture have been prepared by first shooting an original scene with a color reversal film, converting the recorded XYZ color coordinates back to those of the original scene, then applying image processings such as one making the image displayed on an sRGB monitor subjectively favorable, color gamut compression taking into account the sRGB monitor gamut for reproducible colors, etc. XYZ/sRGB–SCID is described in terms of 16 bit XYZ and 8 bit RGB color data with a resolution of 4k × 3k. The present standard images, comprising stable and reliable high quality image data, are extensively useful for the evaluation of the color reproducing capability of imaging systems, the evaluation of color image output devices, coding techniques associated with the storage and transmission of high resolution image data, etc.
Introduction

Today, personal computers are prevailing not only for business use but into home broadly and deeply. Digital input and output devices such as digital camera, CRT monitor, LCD, scanner, printer, etc., are also rapidly improving, making digital imaging environments a reality where ordinary people can easily handle digital images. At the same time, improvement of image quality is also strongly expected. Under such a trend, standard images play more and more important roles in the related study and development exemplified by the evaluation of image processing techniques including device and algorithm, the evaluation of image quality, etc.

Among the work items of ISO/TC130/WG2, preparation of new standard images succeeding ISO 12640 (CMYK–SCID) is now being discussed. More specifically, two sets of standard images defined in two color spaces, one being Lab–SCID defined in a D50 color space representing the viewing and evaluating condition for reflection prints such as printed matter and photograph, and the other being XYZ/sRGB–SCID defined in a D65 color space representing the viewing and evaluating condition for images on display devices.
As for image preparation, the Swiss team is in charge of the former set while the Japanese team is in charge of the latter (Figure. 1) The present paper describes XYZ/sRGB-SCID.

![Figure. 1 Relationship between D65 and D50 color space.](image)

Features of Image Data of XYZ/sRGB-SCID

**Color space**

XYZ/sRGB-SCID, which aims to be applicable to the standardization of not only printed matters but displayed images, has been developed with a premise of various images such as natural images, computer graphics, business graphics, etc., to be displayed on monitors.

To meet the above premise, the ITU-R BT.709-3[^1] for various video devices has been adopted as the image capturing norm, and sRGB which will soon be standardized as the monitor display norm[^2][^3] to ensure a high degree of consistency from image input to monitor output (Figure. 2).

As ITU-R BT.709-3 and sRGB both assume the color temperature of D65, the present images are defined in a D65 color space; accordingly the XYZ image data are regarded as describing scenes illuminated with D65 light sources. The RGB image data have been derived so as to be interpreted as such obtained by applying...
the ITU-R BT.709-3 transformation to the XYZ image data. It should be noted that the RGB images can be displayed on sRGB monitors without further transformation. It is thus concluded that the RGB images in XYZ/SCID are RGB standard images in the ITU-R BT.709-3/sRGB imaging systems.

**Figure. 2 Relationships among the original scene, XYZ, RGB and sRGB monitor.**

**Image contents and image data specifications**

In order to cope with the future multi-media society in a full use of internet, digital archive, etc., XYZ/sRGB–SCID consists of eight natural scene images and seven synthetic images comprising four computer graphics, one business graph and two color charts.

As for the natural scene images, those containing objects suited for image quality evaluation in a well-balanced manner have been chosen with reference to the notes for image quality evaluation described in CMYK/SCID. On the other hand, the computer graphics are synthetic, pictorial images, some being three-dimensional with shadows and the other being two-dimensional without shadows. The business graph includes a bar chart, a circle graph, letters, etc., all widely used for presentation, etc.

The color charts are test images designed for the quantitative evaluation of output devices, consisting of primary, secondary and tertiary color step tablets and vignettes. The gradation depth of the individual image data is 8 bits/color for RGB and 16 bits/channel for XYZ. Each image has a unified aspect ratio of 4:3 (or 3:4) except the color charts. In particular, the natural scene images are of high resolution.
consisting of 4096 × 3072 pixels.

**Applications**
The RGB image data provided with such characteristics can be widely used for the evaluation of the color reproduction characteristics of imaging systems, the evaluation of color image output devices, and the evaluation of coding techniques associated with the storage and transmission of high resolution image data, etc., while the XYZ image data can be used for experiments on color appearance models, etc., since the both image data are consistent and reliable as high resolution digital data sources.

Table 1 Specifications of image data.
<table>
<thead>
<tr>
<th>File name</th>
<th>Descriptive name</th>
<th>Height (pixels)</th>
<th>Width (pixels)</th>
<th>Color Space</th>
<th>Bits depth</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>Woman with glass</td>
<td>4096</td>
<td>3072</td>
<td>XYZ</td>
<td>16-bits</td>
<td>Human skin tone</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RGB</td>
<td>8-bits</td>
<td></td>
</tr>
<tr>
<td>N2</td>
<td>Flowers</td>
<td>3072</td>
<td>4096</td>
<td>XYZ</td>
<td>16-bits</td>
<td>Color reproduction</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RGB</td>
<td>8-bits</td>
<td></td>
</tr>
<tr>
<td>N3</td>
<td>Fishing goods</td>
<td>4096</td>
<td>3072</td>
<td>XYZ</td>
<td>16-bits</td>
<td>Image sharpness,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RGB</td>
<td>8-bits</td>
<td></td>
</tr>
<tr>
<td>N4</td>
<td>Japanese goods</td>
<td>3072</td>
<td>4096</td>
<td>XYZ</td>
<td>16-bits</td>
<td>Wide color gamut</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RGB</td>
<td>8-bits</td>
<td></td>
</tr>
<tr>
<td>N5</td>
<td>Field fire</td>
<td>3072</td>
<td>4096</td>
<td>XYZ</td>
<td>16-bits</td>
<td>Color reproduction</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RGB</td>
<td>8-bits</td>
<td></td>
</tr>
<tr>
<td>N6</td>
<td>Pier</td>
<td>3072</td>
<td>4096</td>
<td>XYZ</td>
<td>16-bits</td>
<td>Fine geometrical patterns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RGB</td>
<td>8-bits</td>
<td></td>
</tr>
<tr>
<td>N7</td>
<td>Threads</td>
<td>3072</td>
<td>4096</td>
<td>XYZ</td>
<td>16-bits</td>
<td>Wide color gamut</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RGB</td>
<td>8-bits</td>
<td></td>
</tr>
<tr>
<td>N8</td>
<td>Silver</td>
<td>4096</td>
<td>3072</td>
<td>XYZ</td>
<td>16-bits</td>
<td>Gray reproduction</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RGB</td>
<td>8-bits</td>
<td></td>
</tr>
<tr>
<td>S1</td>
<td>Teapot</td>
<td>360</td>
<td>480</td>
<td>XYZ</td>
<td>16-bits</td>
<td>3D-CG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RGB</td>
<td>8-bits</td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>Fukusuke</td>
<td>1536</td>
<td>2048</td>
<td>XYZ</td>
<td>16-bits</td>
<td>3D-CG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RGB</td>
<td>8-bits</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>Cat</td>
<td>1536</td>
<td>2048</td>
<td>XYZ</td>
<td>16-bits</td>
<td>2D-CG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RGB</td>
<td>8-bits</td>
<td></td>
</tr>
<tr>
<td>S4</td>
<td>Sports</td>
<td>2048</td>
<td>1536</td>
<td>XYZ</td>
<td>16-bits</td>
<td>2D-CG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RGB</td>
<td>8-bits</td>
<td></td>
</tr>
<tr>
<td>S5</td>
<td>Business graph</td>
<td>1536</td>
<td>2048</td>
<td>XYZ</td>
<td>16-bits</td>
<td>Business graph</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RGB</td>
<td>8-bits</td>
<td></td>
</tr>
<tr>
<td>S6</td>
<td>Color chart</td>
<td>1332</td>
<td>2736</td>
<td>XYZ</td>
<td>16-bits</td>
<td>Step tablets</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RGB</td>
<td>8-bits</td>
<td></td>
</tr>
<tr>
<td>S7</td>
<td>Color vignettes</td>
<td>2608</td>
<td>4256</td>
<td>XYZ</td>
<td>16-bits</td>
<td>Vignettes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RGB</td>
<td>8-bits</td>
<td></td>
</tr>
</tbody>
</table>

IMAGE DATA PREPARATION FOR XYZ/sRGB-SCID
As for the natural scenes of XYZ/sRGB-SCID, each original scene was shot with a 4 × 5 inches color reversal film, which was then scanned with a drum scanner of a high S/N to obtain 4K x 3K image data quantized to 12bits. Further, these data were used to calculate the XYZ values of the recorded film image.

On the other hand, XYZ values are regarded as those describing scenes illuminated with D65 light sources according to the definition of ITU-R BT.709-3. If the distribution of light reflected by such a scene could be captured with an imaging device having the ideal spectral sensitivities defined by ITU-R BT.709-13, then it would be possible to obtain the scene XYZs via the transformation between XYZ and RGB. However, from a practical point of view, production of an electronic imaging device provided with the ideal spectral sensitivities and a sufficiently high resolution is almost impossible. It is also unrealistic to measure the reflection spectra of individual points throughout a real scene.

**Figure. 3 Image data processing flow**
Thus, to obtain XYZ data representing real scenes, the following approach was adopted which can deal with natural images of high resolution under the precisely defined color space of the present standard.

According to the definition of ITU-R BT.709-3, the linear rgb values are linearly related to the scene brightness. Such linear rgb values can be obtained as follows. First, the colors recorded in a color transparency are transformed to XYZs linear to the scene brightness by using the $\gamma$ characteristics of the reversal film. After the XYZs thus obtained are subjected to an exposure correction treatment so as to achieve an optimized display lightness on monitors, they are used to derive linear rgb values defined by ITU-R BT.709-3. Though the final XYZs are different from those of the original scene illuminated with D65 light sources, they can be regarded as those corresponding well to the scene, as they are linearly related to the scene brightness.

However, it must be kept in mind that the device-dependent color characteristics of the color reversal film has been modified by an operation to impart a subjectively favorable appearance to the display image on the sRGB monitor, and further by gamut mapping.

In the next section, details of the data derivation will be described. Figure 3 illustrates such an image data processing flow.

**Image scanning**

Each of the eight, 4”x 5” color transparencies having captured natural scenes is scanned with a drum scanner described in Table 2 to store the images in digital form.

<table>
<thead>
<tr>
<th>Manufacturer, Product Type</th>
<th>Dainippon Screen Co.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scanning Aperture</td>
<td>25 microns square</td>
</tr>
<tr>
<td>Density Resolution</td>
<td>0.001 at D&lt;2.5</td>
</tr>
<tr>
<td>A/D Conversion</td>
<td>Analog density signals were quantized to 12bits/channel</td>
</tr>
</tbody>
</table>

The resulting data consist of transmission density values quantized to 12bits/channel.

**Transformation of integral density to analytical one**

Via the inverse transform of the $\gamma$ characteristics of the reversal film, the data gathered with the scanner are converted to those linearly related to the scene brightness (Figure 4)
Figure 4 γ characteristics of the color reversal film used for image capture.

The resulting integral density data, $D_i$, $r$, $D_i$, $g$ and $D_i$, $b$ are first converted to analytical ones, $D_a$, $r$, $D_a$, $g$ and $D_a$, $b$.

\[
\begin{pmatrix}
D_{a,r} \\
D_{a,g} \\
D_{a,b}
\end{pmatrix} = A_{k,1}
\begin{pmatrix}
D_i, r \\
D_i, g \\
D_i, b
\end{pmatrix}
\]  

\hspace{1cm} (1)

The conversion matrix ($A_{k,1}$) has been calculated in advance for each kind of reversal film.

Calculation of spectral transmittance distribution
By using the Lambert–Beer's law, the spectral transmittance of each pixel is calculated:

\[
T(\lambda) = 10^{-\int S(\lambda) \times \text{DDi}(\lambda) d\lambda}
\]  

\hspace{1cm} (2)

wherein $i = r$, $g$ or $b$, and $\text{DDi}(\lambda)$ represents the spectral density distribution exhibited by the individual dyestuff existing alone.

Tristimulus value
Tristimulus values $X_Y_Z_0$ are calculated by using a D65 light source and the color matching functions. In the following formulae, $S(\lambda)$ is the spectral energy
distribution of the light source[4], and X(λ), Y(λ) and Z(λ) are the color matching junctions[5].

\[ X_0 = 100 \left( \Sigma S(\lambda) T(\lambda) X(\lambda) / \Sigma S(\lambda) Y(\lambda) \right) \]

\[ Y_0 = 100 \left( \Sigma S(\lambda) T(\lambda) Y(\lambda) / \Sigma S(\lambda) Y(\lambda) \right) \]  \hspace{1cm} (3)

\[ Z_0 = 100 \left( \Sigma S(\lambda) T(\lambda) Z(\lambda) / \Sigma S(\lambda) Y(\lambda) \right) \]

**Gamut mapping**

As all the colors of the images must be reproduced on sRGB monitors without any further transformation, the color gamut thereof has been compressed to such that sRGB monitors can reproduce; in other words, the XYZ values representing each standard image must be within the color gamut of ITU-R BT.709-3. Since the color gamut of color transparencies differs from that of ITU-R BT.709-3 in general, the XYZ to RGB transformation yields certain colors not reproducible with sRGB monitors if the image data are not subjected to a gamut mapping. Among various methods treating the data outside the reproducible color gamut, we have adopted one to be described below.

**Outline of gamut mapping**

In order to carry out a gamut mapping in the CIELAB color space the data XYZ₀ are transformed to the CIELAB. Separately, the following two regions are provided within the ITU-R BT.709-3 color gamut (Figure 5).

![Figure 5 Conceptual illustration of gamut mapping.](image)

(1) An original color region where no color chance takes place with the present gamut mapping, and
(2) A compressed color region where colors outside the ITU−R BT.709−3 color gamut as well as close to its boundaries are present as the results of the present gamut mapping. When the Lab data lie in the original color region, they are conserved as they are. When the Lab data lie in the compressed color region, chroma is compressed 50% to keep lightness and hue unchanged.

**Algorithm**

**Step1:** The XYZ₀ data are transformed to the CIELAB data \( V_{lab}(i,j) \) where \( i \) and \( j \) represent the spatial coordinates of pixel.

\[
V_{lab}(i,j) = (L₀^*, a₀^*, b₀^*) \tag{4}
\]

Further, the following two formulae are used to derive chroma \( C^* \) and hue angle \( \theta^* \).

\[
C₀^*(i,j) = \left( a₀^*(i,j)^2 + b₀^*(i,j)^2 \right)^{1/2}
\]

\[
\theta₀^*(i,j) = \tan^{-1}\left( \frac{b₀^*(i,j)}{a₀^*(i,j)} \right) \tag{5}
\]

Then, the CIELAB data \( V_{lab}(i,j) \) is expressed in terms of chroma and hue as,

\[
V_{lab}(i,j) = (L₀^*, C₀^*, \theta₀^*) \tag{6}
\]

**Step2:** Using the lightness data \( L₀^* \) and the hue data \( \theta₀^* \) of the pixel in concern, the chroma data \( C''k \) of the boundary for ITU−R BT.709−3 with the same \( L₀^* \) and \( \theta₀^* \) is calculated. Then, the chroma data \( C'o \) of the pixel in concern is normalized with the calculated chroma \( C''k \)

\[
P(i,j) = \frac{C₀^*(i,j)}{C''k(i,j)} \tag{7}
\]
Figure. 6 Examples of data lying outside and inside of the reproducible color gamut.

Step 3: The boundary point $P_x$ of the original color region and compressed one as described outline is illustrated in Figure 7.

The histogram of the normalized chroma is obtained by performing the above calculation or all the pixels.

A peak is identified that lies below but closest to 1.0 in the histogram. Such a peak defines the mapping boundary (Figure 7).

Figure. 7 Example of mapping function.
Step 4: The mapping function $F(P)$ must satisfy the following conditions, where $P_{\text{max}}$ is maximum data of normalized chroma:

$F(P_{\text{max}}) = 1.0$

At the mapping boundary $P_x$,

$$\frac{\partial F(P_x)}{\partial P_x} = 1.0$$

For $0 \leq P \leq P_x$, $F(P) = P$, and

For $P_x < P \leq P_{\text{max}}$, $F(P)$ increases smoothly and monotonically.

By making use of the mapping function defined above, the normalized chroma of the pixel in concern after mapping $P'$ is calculated from its normalized chroma $P$.

Then, the chroma value $C_1'$ after mapping can be obtained by

$$C_1'(i,j) = P' C_k(i,j)$$

By the transformation of $C_1'$ to XYZ, one finally completes color mapping onto the reproducible gamut of ITU-R BT.709-3 color space.

Definitions in ITU-R BT.709-3 and XYZ/RGB transformation

The primary colors and the white of ITU-R BT. 709-3 are defined by the following color coordinates in Table 3:

<table>
<thead>
<tr>
<th></th>
<th>Red</th>
<th>Green</th>
<th>Blue</th>
<th>D65</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>0.6400</td>
<td>0.3000</td>
<td>0.1500</td>
<td>0.3127</td>
</tr>
<tr>
<td>$y$</td>
<td>0.3300</td>
<td>0.6000</td>
<td>0.0600</td>
<td>0.3290</td>
</tr>
<tr>
<td>$z$</td>
<td>0.0300</td>
<td>0.1000</td>
<td>0.7900</td>
<td>0.3583</td>
</tr>
</tbody>
</table>

Then, the transformation from XYZ(D65) to linear rgb values is performed by

$$\begin{bmatrix} r \\ g \\ b \end{bmatrix} = \begin{bmatrix} 3.2470 & -1.5374 & -0.4986 \\ -0.9692 & 1.8760 & 0.0416 \\ 0.0556 & -0.2040 & 1.0570 \end{bmatrix} \begin{bmatrix} 0 \\ 0.01 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Further, the opto-electronic characteristic function (OECF) of video cameras is
defined as follows:

\[
V' = \begin{cases} 
1.099 \times V^{-0.45} - 0.099 & 0.018 \leq V' \leq 1.0 \\
4.50 \times V' & 0.0 \leq V' < 0.018 
\end{cases}
\]

\[
V' = R, G, B \\
I' = r, g, b
\]

wherein \( r \), \( g \) and \( b \) are the linear rgb values of the original scene, and \( R \), \( G \) and \( B \) are the non-linearly transformed RGB video signals.

Data processing to make the appearance of each image displayed on an sRGB monitor subjectively favorable

As the present standard is aiming to be used in a broad range of applications not only for high-end but also low-end users, each image is expected to have a subjectively favorable quality. Thus, it must satisfy the requirement of looking subjectively favorable on monitors as well as in the form of hardcopy (Looking good).

For that purpose, we have taken the following steps: transformation of the XYZ values compressed with in ITU-R BT.709-3 color gamut onto the RGB color space defined by ITU-R BT.709-3, "Looking good" operation of the resulting data to secure a subjectively favorable appearance under sRGB monitor display conditions, and inverse transformation of ITU-R BT.709-3 to derive XYZ values.

The "Looking good" processing mainly consists of lightness (exposure) correction, color saturation enhancement, etc.; the processing parameters have been determined by subjective evaluation of each image.

**XYZ/RGB data description**

**XYZ image data**

The XYZ image data were normalized by the XYZ values for D65 white, and expressed as 16bit data:

\[
X_{16bit} = 65535 X_{65} \\
Y_{16bit} = 65535 Y_{65} \\
Z_{16bit} = 65535 Z_{65}
\]

(12)

wherein, \( X_{65} \), \( Y_{65} \) and \( Z_{65} \) represent the tristimulus values of D65.

**RGB image data**
The RGB image data were obtained by first converting the XYZ, via ITU-R BT.709-3 transformation, to non-linear RGB data, which were multiplied by 255 to obtain 8bit data without sign:

\[ R_{8bit} = 255 R \]
\[ G_{8bit} = 255 G \]
\[ B_{8bit} = 255 B \]

R8bit, G8bit and B8bit all were rounded to the nearest integers.

**Ending remarks**

The image data preparation for XYZ/sRGB-SCID now under study in ISO/TC-130/WG2 has been described.

The present standard images are widely applicable to the evaluation of the color reproduction characteristics of imaging systems, the evaluation of color image output devices, and the evaluation of coding techniques associated with the storage and transmission of high resolution image data, etc., as they are of guaranteed consistent quality.

Now, a DIS (Draft international Standard) of the present Standard is being made. In Japan, to facilitate its early and extensive use, a part of the XYZ/sRGB-SCID images was distributed in the August of 2000 from JSA (Japanese Standard Association).

**NOTE 1**

The present paper modified a part of reference [7] and was described.

**NOTE 1**

The JIS version of XYZ/sRGB-SCID does not include N6 to N8 images. Those who are interested in them may refer to Standard High Precision Pictures (SHIPP)[6] which are based on the same color space as the present standard and which include N6 to N8.

**References**


http://www.w3.org/pub/WWW/Graphics/Color/sRGB.html