A Huge Spectral Characteristics Database and its Application to Color Imaging Device Design

Johji Tajima
C&C Media Research Laboratories, NEC Corporation
Kawasaki, Japan

Abstract

A huge spectral characteristics database for color reproduction evaluation is being developed. As its first application, properties of data contained in the database were analyzed by principal component analysis with respect to color imaging device design. From results of the analysis, it was found that colors reproduced by photographic materials and graphic prints are distributed in low-dimensional spaces and most color scanners may not necessarily satisfy the Luther condition, whereas camera-type image input devices should satisfy the condition.

Introduction

It is well known that the Luther condition should be satisfied by sensitivities of sensors which are used in color image input devices. However, the author made it clear that it is not necessary to satisfy this condition if the spectral reflectance of objects can be restored by the sensor output [1]. Figure 1 shows the relation. In this figure, two 3-dimensional subspaces are depicted in one dimension. As object color distribution is depicted as an area, object color distribution is considered to spread out over three dimensions. However, Fig.1 shows that extent in the 4th dimension is smaller than that in the former three dimensions. An object color **P** is sensed by human vision as $\mathbf{P}_e$ (perceived color) and sensed by sensors as $\mathbf{P}_f$. If we know the object color distribution and its three principal axes, the color **P** is approximately restored as **P'**. Projecting **P'** to human visual subspace, an approximated color **P**e' is obtained. This back-projection and re-projection process can be summarized as a 3x3 matrix application. The usual 3x3 matrix color correction process, where the matrix is obtained using the least square method on a set of color samples, is equivalent to the process.

For this reason, it has become very important to understand the spectral object color distribution. Color reproduction evaluation has been carried out on the Luther condition basis. However, sensors in a color image input device, which images objects whose spectral reflectances are confined in a three dimensional subspace, need not satisfy the condition. Sensors with good color reproduction may be developed without consideration of the Luther condition. A project to construct a huge spectral characteristics database was proposed and carried out to solve this problem.

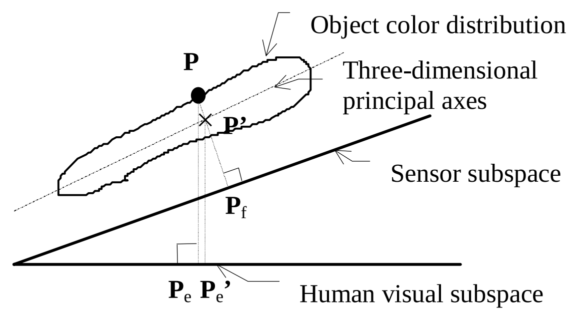

Figure 1 Color correction by spectral characteristic restoration.

Spectral Characteristics Database

Spectral data were collected in a systematic manner according to the following object categories.

a. Photographic materials (Transparencies / Reflection prints) produced by four major vendors: 2,304 colors.
b. Graphic prints (offset / gravure) including 33 combinations of eleven types of ink and three types of paper: 30,624 colors.
c. Color computer prints produced by 21 models in five sub-categories, i.e., dye-sublimation, ink-jet, electrostatic, silver halide and others: 7,856 colors.
d. Paints – water color , oil paints and other two sub-categories: 229 colors.
e. Flowers: 148 colors.
f. Leaves: 92 colors.
g. Human skin – several positions on face and others: 8,049 colors
h. Krinov data (historical data for outdoor objects): 370 colors.

In total, 49,472 colors were collected. Objects in categories a and b are conventional subjects for color scanners. Though objects in category c have not been such subjects so far, they will likely be considered as such in future, since their printing quality is being rapidly improved. Objects in categories e, f and g are especially important subjects for digital still and video cameras. Category d will be a very important category in extending digital archive applications. Category h is a historical database for outdoor objects. It should be very useful for digital camera design, if the data are reliable.

We consider this to be the largest spectral characteristics database ever constructed. This database is to be used as a standard to evaluate color reproduction quality in image input devices. We proposed the database to be a technical report with CD-ROM to ISO/TC130 so that any engineer who develops an image input device will be able to access the database and evaluate the spectral sensitivities for the device. Content of this database is explained in Ref.[2] in detail.

Distribution Analysis

As the first application of the database, spectral characteristics distribution in each category or sub-category was analyzed by the principal component analysis. As described in Ref.[1], the spectral characteristic of an object is well restored (approximated) by a small number of principal components, if the distribution is confined in a low dimensional subspace. In this case, good color reproduction is achieved, even if the sensor does not satisfy the Luther condition.

Quality Criteria

First of all, criteria for color reproduction quality evaluation were investigated. In many papers, where distributions are analyzed using principal component analysis, percentage of variance was used as a criterion; e.g. spectral reflectance can be sufficiently represented by the first four components, since 99% of the entire variance is expressed by the first four principal components (PCs). In color science, percentage does not directly relate to color reproduction quality. Therefore, this paper introduces a more practical criterion to judge quality. Quality that is comparable with the stability of graphic prints or high-quality color computer prints is defined as class A, while quality that is comparable with the stability of consumer-oriented color computer prints is defined as class B. If the class A criterion is met, professional users will be satisfied by the quality. If the class B criterion is met, non-professional users will be satisfied. The stability, or deviation, of graphic prints and color computer prints was evaluated as the first step. Figure 2 shows standard deviation per wavelength σ_w and the maximum deviation d_{max} per wavelength in spectral reflectances in color patches which were repeatedly printed.

- ✧ Patch 'Null': σ_w and d_{max} for repeatedly measured calibration color patches. These deviations also include errors caused by measurement noises.
- ✧ Patch 'Offset': σ_w and d_{max} by measurements for color patches repeatedly printed by offset printing.
- ✧ Patch 'Gravure': σ_w and d_{max} by measurements for color patches repeatedly printed by gravure printing.
- ✧ Patch 'Printer-SH': σ_w and d_{max} by measurements for color patches repeatedly printed by a silver halide computer color printer.
- ✧ Patch 'Printer-DS': σ_w and d_{max} by measurements for color patches repeatedly printed by a dye-sublimation computer color printer.
- ✧ Patch 'Printer-IJ': σ_w and d_{max} by measurements for color patches repeatedly printed by an ink-jet computer color printer.
- ✧ Patch 'Printer-IJD1' and 'Printer-IJD2': σ_w and d_{max} by measurements for color patches printed by two products of two models among ink-jet computer color printers.

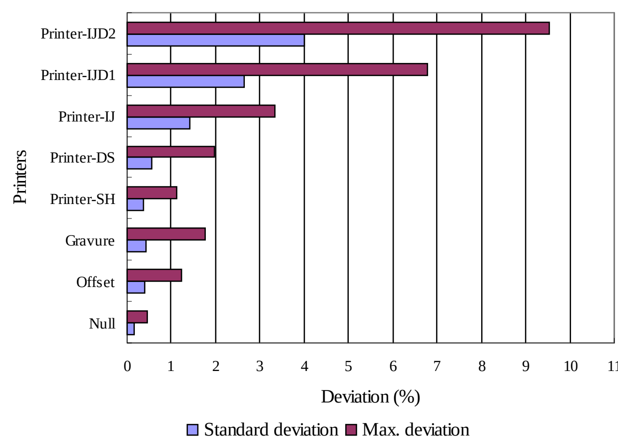

Fig.2 Reflectance deviation caused by printing repetition.

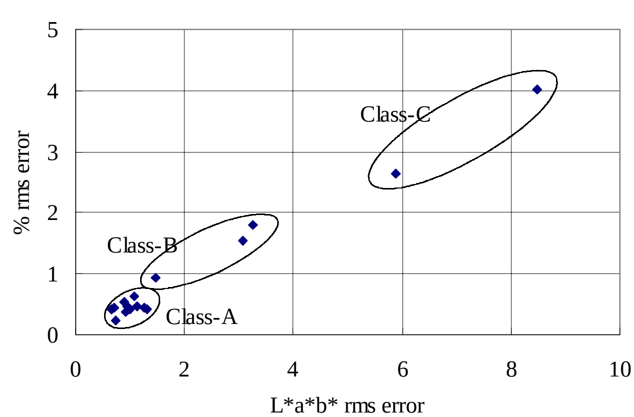

Fig.3 Correlation between reflectance deviation (%) and CIELAB deviation.

Silver-halide printers and dye-sublimation printers are high quality printers which are accepted for professional use, while ink-jet printers are consumer-oriented products. As color differences cannot be evaluated by the standard deviations in spectral reflectances, reproduced colors for the spectral reflectances under the D65 illuminant were also calculated and represented in CIELAB uniform color space. Standard deviations in CIELAB space σ_{ab} were computed and plotted in relation to σ_w in Fig.3. They are almost linearly correlated and it is understood that $\sigma_{ab} \cong 2$ and $\sigma_w \cong 1$ are criteria for professional standard (Class-A). In the same way, $\sigma_{ab} \cong 4$ and $\sigma_w \cong 2$ are criteria for consumer products (Class-B). Color difference of color patches by two models of the same consumer product is too large to be allowed even by non-professionals (Class-C).

Principal Component Analysis

Spectral characteristic distributions were then analyzed according to these criteria. PCs were calculated for each category and all spectral data for a color patch **p** was approximated using the first m (=3~11) PCs ($\mathbf{p}_1$, $\mathbf{p}_2$, … , $\mathbf{p}_m$) as Eq.(1).

$$\mathbf{p} = \sum_{i=1}^{m} a_i \mathbf{p}_i \qquad (1)$$

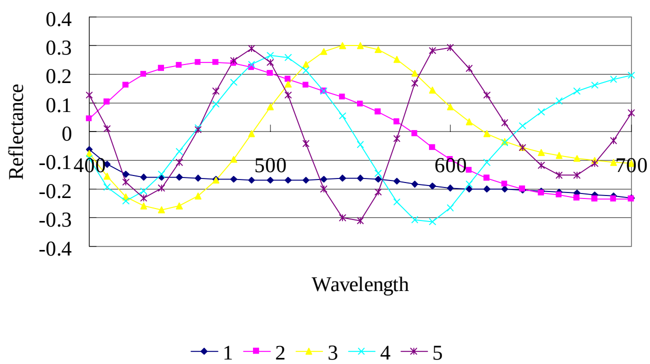

Fig.4 Five PCs for photo color prints.

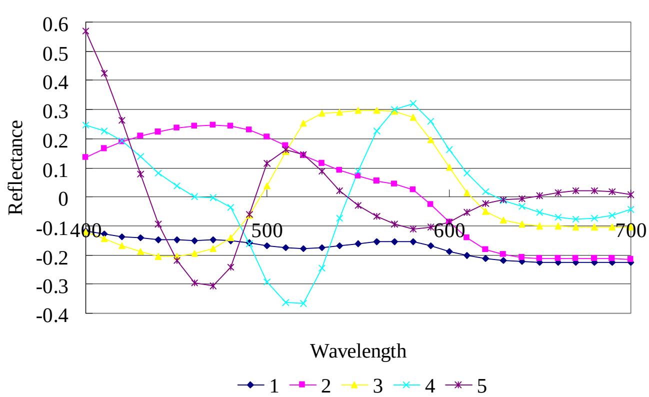

Fig.5 Five PCs for offset color printing.

As examples, figure 4 shows the first five PCs for color transparencies and figure 5 shows those for offset color printing. We can see that the first three components have different but similar shapes.

Figure 6 shows how many PCs can represent the dimensionality of object colors in each category. It is apparently observed that graphic prints (offset and gravure) colors are distributed in especially low dimensional subspaces. Using only three PCs, approximation satisfies the Class-B criterion. Even the Class-A criterion can be satisfied by using four PCs. Photographic materials (reflection prints and transparencies) are also approximated very well by three PCs. On the other hand, object colors in other categories (ex. flowers & leaves or paints) are distributed in high-dimensional subspaces. Colors of paints satisfy the Class-B criterion when at least six PCs are used for approximation.

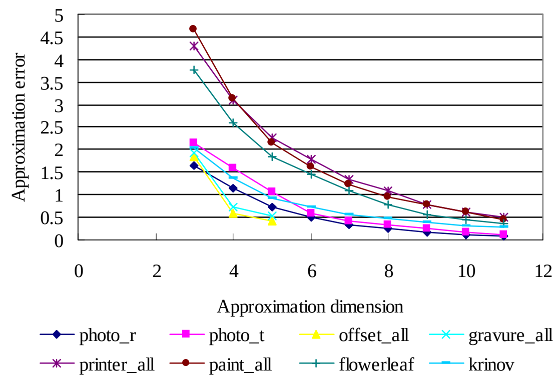

Fig.6 Approximation error using PCs for each category.

The PCs for offset prints were calculated for 30 combinations of ten types of ink and three types of paper. Using these PCs, ten color patch sets printed by ten types of ink were approximated and approximation errors were evaluated in σ_w. The result is depicted in Fig.7. Colors printed with each ink are sufficiently approximated by the common offset PCs and the Class-B criterion can be satisfied by three PCs. In this experiment, colors in gravure prints and photo reflection prints are also approximated, using the common offset PCs. It is shown that colors in gravure prints are approximated by the PCs rather well, while photo reflection prints are not.

From the result, we can conclude as follows.

a. Because scanner users can easily distinguish graphic prints, photo color prints and photo color transparencies, scanner sensor sensitivities can be designed without much consideration on the Luther condition, if color correction matrices can be switched by the users.

b. If the sensor sensitivities should be designed so that one color correction matrix (e.g. for offset color prints) should be used for all scanned targets, the Luther condition should be considered for the sensitivity design.

On the other hand, spectral reflectance distributions of color computer prints (especially those output by dye sublimation and ink-jet printers), paints, and flowers are not described well by low-dimensional spaces. For reference, PCs for flowers & leaves are depicted in Fig. 8. Shapes of PCs look very different from those for hardcopies. Camera-type color image input devices often make pictures which contain flowers and leaves as well as human faces. Therefore, such analyses for individual categories are not sufficient. More detailed evaluation for color reproduction of each object should be investigated.

From this preliminary analysis, it may be concluded that color sensors for camera-type image input devices do need to satisfy the Luther condition.

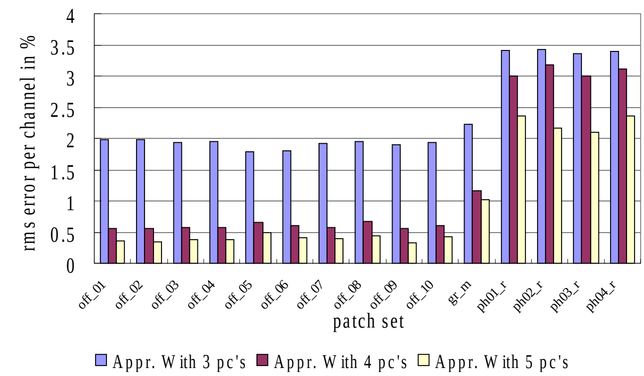

Fig.7 Approximation errors by 3, 4 and 5 PCs obtained using 30 combinations of ink and paper.

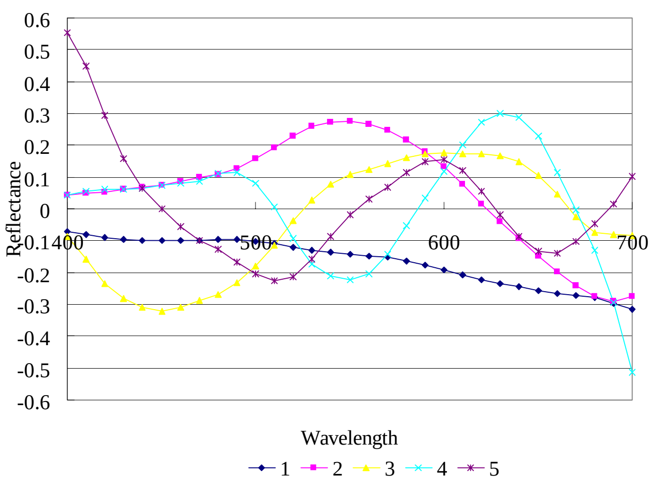

Fig.8 Five PCs for flowers & leaves.

Conclusion

A huge spectral characteristics database was constructed. It is expected to be used as a standard database for color reproduction evaluation in color image input devices. The next key issue is to establish an evaluation method based on the use of this database. As the first preliminary study, data distribution in spectral measurement space was analyzed by principal component analysis. It was shown that object colors which are conventionally input by color scanners are distributed in low dimensional subspaces. This may be the reason normal color scanners can reproduce good colors, though normal color scanner sensors do not necessarily satisfy the Luther condition.

However, even for such object colors, only class B quality can be realized by using only three PCs. More deliberate device design for high-quality color scanners is essential. In addition, digital cameras will need to input objects whose colors are distributed in higher dimensional subspace. For such device design, this database will be very useful, since colors, which cannot be well reproduced, can be searched in advance. More detailed discussion and further work are essential to define and standardize an evaluation method for color image input devices which uses this database.

Acknowledgement

This project to construct a huge database of spectral characteristics (reflectances / transmittances) of major objects was sponsored by the New Energy and Industrial Technology Development Organization (NEDO) and the Japan Standard Association (JSA).

The database development was carried out through collaboration of the following committee members:

[Spectral Characteristics Database Construction WG]

Chair:	Johji Tajima	NEC Corporation
Advisor:	Yoichi Miyake	Chiba University
Members:	Yoshihiko Azuma	Tokyo Inst. of Polytechnics
	Hideaki Haneishi	Chiba University
	Tetsuo Iga	Toyo Ink Mfg. Co., Ltd.
	Masao Inui	Konica Corp
	Masayuki Nakajima	Tokyo Inst. of Technology
	Noboru Ohta	Fuji Photo Film Co.
	Nobutoshi Ojima	Kao Corp.
	Sei Sanada	Toppan Printing Co.,Ltd.
	Masato Tsukada	NEC Corporation
	Norimichi Tsumura	Chiba University
	Yutaro Horikoshi	Ministry of International Trade and Industry
	Hajime Kawanago	Japan Standard Association

References

[1] J.Tajima, *Proc. of 4th Color Imaging Conference*, pp.25-28 (1996).

[2] J. Tajima, M. Tsukada, Y. Miyake, H. Haneishi, N. Tsumura, M. Nakajima, Y. Azuma, T. Iga, M. Inui, N. Ohta, N. Ojima and S. Sanada, *Proc. of International Symposium on Electronic Image Capture and Publishing (SPIE-3409)*, (to be published) (1998)

Biography

Johji Tajima received his B.S. degree in Physics and a doctorate in Information Science from the University of Tokyo in 1971 and 1990, respectively. Dr. Tajima joined NEC Corporation in 1971 and is now a senior principal researcher in C&C Media Research Laboratories. He has been engaged in research on image processing and pattern analysis, especially color image processing and 3D vision.