Evaluation of Color Matching Performances for SPEM and Other Models

Masato Tsukada,*[1] Johji Tajima,[1] Hirohisa Yaguchi[2]

[1] Multimedia Research Laboratories, NEC Corporation, Kawasaki, Kanagawa, Japan

[2] Department of Information and Image Science, Chiba University, 1–33 Yayoi-cho, Inage-ku, Chiba 263-8522 Japan

Received 15 April 2002; revised 10 December 2002; accepted 17 January 2003

Abstract: We performed subjective experiments to evaluate color matching performance of the Spectral Properties Estimation Model (SPEM) and six other models (von Kries, CIELAB, LLAB, RLAB, Nayatani, and CIECAM97s) between two CRT monitors whose whites were quite different. Moreover, we evaluated color matching of these models between a CRT monitor and a printed image set in a dark room. The SPEM we developed is a new chromatic adaptation model based on hypothetical spectral properties estimation. This article describes the subjective experiments and the results obtained. The SPEM produced good color matching performance in the experiments. The detailed algorithm of the SPEM is given in the Appendix. © 2003 Wiley Periodicals, Inc. Col Res Appl, 28, 445–453, 2003; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/col.10197

Key words: chromatic adaptation model; color appearance model; color constancy; color management

INTRODUCTION

With recent rapid advances in color imaging, color matching between different devices has become an increasingly important issue. In most CRT displays on the market, the original white is set to bluish white (i.e., 9300K), while printers produce colors under the illumination D50 (i.e., 5003K). An appearance matching method needs to be able to accommodate such differences. If it were possible to confirm on a CRT monitor the exact color appearance to be produced by a printer, it would increase the work efficiency of graphic designers.

In most conventional color management strategies, color transformation is based on colorimetric matching, which is calculated in terms of XYZ, CIELAB, or CIELUV values. Even when these values are matched, however, the colors themselves may not appear to be matched to the human eye.

A number of different color appearance models and chromatic adaptation models have been proposed to replace colorimetric matching.[1–8] In color appearance models, when tristimulus values for a stimulus and data about a viewing environment are given, lightness, hue and chroma, the brightness and colorfulness of the stimulus which are absolute values in human color perception, can be calculated by modeling color appearance phenomena such as the Stevens effect and the Hunt effect.

Color matching performance for these models between different color imaging devices would be of great interest to those who want to solve the color matching issues. Proposed models, however, show a tendency to increase complexity more and more to model various color appearance phenomena on human color vision. Therefore, it is difficult to know which model produces good color matching performance in the practical color management.

To solve the problem, we performed subjective experiments to evaluate the color matching performance of color appearance models and chromatic adaptation models for natural color images displayed on two CRT monitors whose whites were quite different. Moreover, we evaluated the color matching performance of these models between a CRT monitor and a color printer. The experimental conditions can be applied to those for remote color proofing and electronic commerce via the Internet.

In the experiments, we examined six well-known models

* Correspondence to: Masato Tsukada (e-mail: m-tsukada@cj.jp.nec.com).

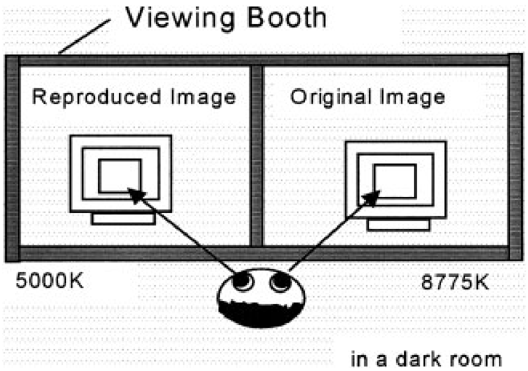

FIG. 1. Viewing booth in the experiments for color matching evaluation.

and a spectral properties estimation model that we developed as a new chromatic adaptation model. The Spectral Properties Estimation Model (SPEM) is based on a new concept in color matching algorithms.

This article describes how we performed the experiments to evaluate color matching performance of each model and discusses the results of the experiments we obtained. The detailed formula for the SPEM is given in the Appendix.

COLOR MATCHING EXPERIMENTS BETWEEN TWO CRT MONITORS

We examined the following six well-known models and a new chromatic adaptation model we developed: 1) von Kries; 2) CIELAB; 3) LLAB;[1] 4) RLAB;[2] 5) Nayatani97;[4,5,8] 6) CIECAM97s;[6] 7) SPEM.[9,10]

Most color appearance models and chromatic adaptation models were derived, based on inductive reasoning, by approximating various color appearance phenomena obtained in psychophysical experiments. On the other hand, it is possible to develop a model based on deductive reasoning, by introducing some assumptions about human color vision. We developed SPEM, which is a chromatic adaptation model, by introducing a new concept in color matching solutions.[9] SPEM is based on spectral properties estimation by using computational theory of color constancy and takes incomplete chromatic adaptation into account. The detailed algorithm is described in the Appendix.

Referring to CIE guidelines,[11] we made a viewing booth with no illumination and set two CRT monitors in it in a dark room as illustrated in Fig. 1. Two CRT monitors were used to easily simulate conditions of remote soft color proofing and electronic commerce via the Internet and those of color matching between softcopy and hardcopy, discarding physical factors other than color appearance such as the difference between color devices' gamuts. Since a CRT monitor is a self-luminous device, it does not exactly apply to the conditions under which we compare colors between softcopy, such as a self-luminous object, and hardcopy, such as a reflecting object surface. We believe, however, that this experiment is the first step in evaluating color matching performance of color appearance models and chromatic adaptation models between cross-media communications. In this experiement we were able to grasp the color matching performances by setting bluish white (e.g., 9300K), which is originally set to CRT monitors, to the one of our monitors and yellowish white (e.g., 5000K) to the other.

The white of the right monitor (NEC Multisync KM174Pro2) was set to the default white and that of the left monitor (NEC Multisync KM174Pro) was adjusted to the same chromaticity as D50 in the experiments. The chromaticity of the right monitor's white was $x = 0.281$ and $y = 0.315$ (8775K) and that of the left monitor's white was $x = 0.346$ and $y = 0.359$ (5000K). Luminance levels of the whites were adjusted so that they were nearly the same: 99 cd/m^2 for the right monitor and 94 cd/m^2 for the left monitor. Gamma correction for the CRT monitors was conducted by using ICC profiles[12] and imaging software with color management functions that we developed. We confirmed that the average color difference ΔE_{ab} between an ideal image and a displayed image was less than 1.0 under these experimental conditions. The achromatic background of the image on the monitor was set to a luminance factor of 20%.

Four kinds of natural images were prepared because subjects could be influenced by the contents of the image. N1 (Portrait), N2 (Cafeteria), N3 (Fruit Basket), and N7 (Musician) in ISO/JIS-SCID[13] were used for this evaluation. Images as test stimuli were appropriately converted to *RGB* images because SCID images are supplied as *CMYK* ones. These images were hemmed with a reference white. Viewing angles of the images were 10.5° (horizontal) × 8.3° (vertical) for Fruit Basket and Musician, and the inverse for Portrait and Cafeteria.

A test image as a test stimulus and the reference images reproduced by color appearance models and chromatic adaptation models as reference stimuli were carefully created by taking the following steps.

1. *RGB* values in a test image were converted to tristimulus values *XYZ* under 8775K by a linear transformation.
2. For *XYZ* under 8775K of step 1, the corresponding colors $X'Y'Z'$ under D50 were calculated by using all color appearance models and chromatic adaptation models we evaluated.
3. For the corresponding colors $X'Y'Z'$ of step 2, $R'G'B'$ values that were device-dependent colors were calculated by using a linear transformation.
4. Gamut checking was carried out by judging whether $R'G'B'$ values existed within the CRT monitor's gamut. When there was a pixel out of the *RGB* gamut, step 5 was executed.
5. A simple linear compression of gamut was performed to *RGB* of step 1 in *RGB* space.
6. Steps 1–5 were repeated until the reference image was reproduced.

The procedures to reproduce a test image and the reference image are summarized in Fig. 2. To remove the gamut issue, steps 4 and 5 were introduced so that all colors in a

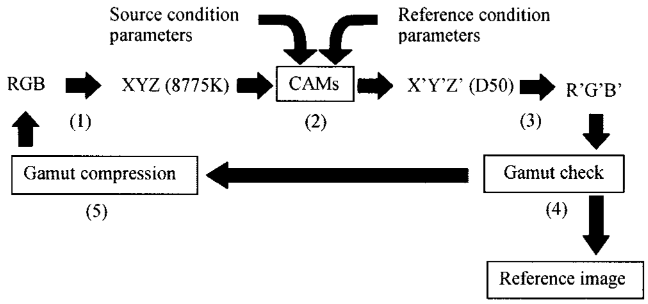

FIG. 2. Test/reference image calculation procedures.

test image and the reference image existed in both gamuts. When the linear compression in step 5 was performed, the test image in step 1 was updated.

Ten subjects, all of whom had normal color vision, evaluated the color appearance of four kinds of natural images displayed on two CRT monitors whose whites were quite different. A test image was displayed on the right monitor and showed the original color appearance as a test stimulus. The reference images reproduced by several kinds of color appearance models and chromatic adaptation models were displayed on the left monitor as reference stimuli. Subjects judged the superiority or inferiority in color matching of the images reproduced by these models. The viewing condition was nearly successive-haploscopic viewing because the subjects could not simultaneously see the two images due to parallax.[14]

We followed the paired comparison method to determine the order of the models' performance for color matching. A test image was displayed on the right monitor and was fixed. The seven reference images reproduced by the above-referenced seven models were displayed on the left monitor. Two images were randomly selected from these seven images and were not simultaneously but alternately displayed by the subjects by clicking a mouse. Subjects were instructed to select the image that was closer in color appearance to the test image displayed on the right monitor from the two images displayed on the left monitor. They were also requested to evaluate whole regions and colors in images and not to concentrate on specific regions or colors. The experiments were repeated twice.

Observation conditions in our experiments are summarized in Table I. Parameters used in each of the models are listed in Table II. In order to ensure fair evaluation, the recommended parameters for each model were selected to match the observation conditions in the experiments.

Interval scales (Z-scales) were calculated from the evaluation results of the ten subjects by using Thurstone's law.[15] Figure 3 shows results of the evaluation experiment. We can see that the SPEM, RLAB, Nayatani97, and CIECAM97s, which take account of incomplete chromatic adaptation, produced good results for two monitors whose whites are quite different, i.e., 8775K-D50, in a dark room. On the other hand, CIELAB, LLAB and von Kries did not produce good results in the experiments.

COLOR MATCHING EXPERIMENTS BETWEEN A CRT MONITOR AND A PRINTER

We performed subjective evaluation of the color matching performance for SPEM and six other models between a CRT monitor and a printed image. We set a printed image and a CRT monitor in the viewing booth in a dark room as illustrated in Fig. 4. The printed image, which was printed by a dye sublimation printer (Victor TruePrint 3500PS) and illuminated with D50, was set in the left booth as a test stimulus. The reproduced image that was a reference stimulus was displayed on the right CRT monitor (NEC Multisync FE70), whose white was about 9300K, with no illumination.

To increase the reliability of the experiments, the original images were created carefully so that physical factors other than color appearance were discarded. To get color data based on measurement and to avoid unnatural contours in the images as much as possible, half a million colors were created by using the linear interpolation based on the measured data of 2951 colors printed by the printer. To remove the problem of the gamut of the printer output being different from that of the CRT monitor, most colors in the original and reproduced images were determined so that they existed in both gamuts.

The comparison of the printer's gamut and the CRT monitor's gamut in CIELAB in Fig. 5 demonstrates that these gamuts are different. From the figure, it can be said that the CRT monitor's gamut is generally bigger than that of the printer. The printer's gamut, however, protrudes from the monitor's one in the area from cyan to green. That is, colors reproduced by the printer include colors that the monitor cannot physically reproduce.

We used the same four SCID images as those used in the experiments between two CRT monitors. These images were hemmed with a reference white. Though SCID images are supplied with *CMYK* data, the *CMYK* data cannot be used in the original images as it is, since physical factors other than color appearance should be removed from the

TABLE I. Experimental conditions.

	Source conditions	Reference conditions
Media type	Softcopy (CRT monitor)	Softcopy (CRT monitor)
White point	8755K (x = 0.281 and y = 0.315)	D50 (x = 0.346 and y = 0.359)
Luminance of white	99 cd/m^2	94 cd/m^2
Luminance of background	20 cd/m^2	19 cd/m^2

TABLE II. Model parameters used in the experiments between two CRT monitors.

	Source condition parameters	Reference condition parameters
LLAB	"Television and VDU displays in dim surrounding"	"Television and VDU displays in dim surrounding"
Nayatani97	Background luminance factor: 20%	Background luminance factor: 20%
	Effective adaptation coefficient: Method II (Luminous-color stimuli)	
CIECAM97s	Viewing condition: Dark surrounding	Viewing condition: Dark surrounding
	Adapting field luminance: 20cd/m^2	Adapting field luminance: 19cd/m^2
RLAB	Media: Softcopy (D = 0.0)	Media: Softcopy (D = 0.0)
	Surrounding: Dark (σ = 1/3.5)	Surrounding: Dark (σ = 1/3.5)
SPEM	Mixing coefficient MC: 0.6	

original images. Original images and the reproduced images were created via the following steps.

1. An appropriate color transformation from *CMYK* of SCID image to *RGB* in the device dependent color of the CRT monitor whose white was set to D50 was carried out. And the *RGB* values were converted to tristimulus values $X_0Y_0Z_0$ by linear transformation.
2. For $X_0Y_0Z_0$ of (1), *CMYK* data and its tristimulus values *XYZ* under D50 were searched from an LUT that consists of half a million colors so that the color difference ΔE_{ab} is minimum. *CMYK* data was saved as the original image that was printed.
3. For *XYZ* under D50 of (2), the corresponding colors $X'Y'Z'$ under 9300K were calculated by using all color appearance models and chromatic adaptation models that we evaluated.
4. For the corresponding colors $X'Y'Z'$ of (3), $R'G'B'$ values that were device dependent colors were calculated by using linear transformation.
5. Gamut checking was carried out by judging whether $R'G'B'$ values existed within the CRT monitor's gamut. When the number of pixels that were out of the gamut exceeded the threshold, saturation compression was performed to $X_0Y_0Z_0$ of step 1 in CIELAB space.
6. Steps 2 through 5 were repeated until the reproduced image was obtained.

The procedures to create an original image and the reproduced image are summarized in Fig. 6. When saturation compression of step 5 was performed, the original image was updated.

The white of a printed image had the same chromaticity as that of the white paper illuminated with D50, and that of a reproduced image was set to bluish white that was originally set to the CRT monitor. The chromaticity of the monitor's white was $x = 0.2841$ and $y = 0.2944$. The backgrounds of the original image and the reproduced image on the monitor were set to gray with a luminance factor of 0.2. Luminance level of the white of the original image was 94.59 cd/m^2 and that of CRT monitor's white was 94.06 cd/m^2.

In the left booth, a black wall with a hole was set between the original image and a subject so that light from the illumination did not directly reach the subject's left eye. The original image appeared not as a self-luminous object but as a reflecting object surface, since the subject could see a gray background around the original image. Viewing angles of the images were 9.54° (vertical) × 11.42° (horizontal) for Musician and Fruit Basket and the inverse for Portrait and Cafeteria.

The SPEM needs basis vectors for the HSPD of the illumination, the HSR of objects and the mixing coefficient *MC*. We used the same parameters as those used in Tsukada *et al.*[9] Gamma correction for the CRT monitor was conducted by using an ICC profile[12] and the same imaging software that we developed. We confirmed that the average color difference (ΔE_{ab}) between an ideal image and an image displayed was less than 1.

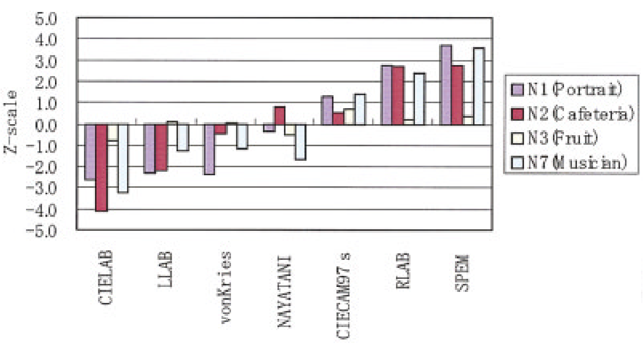

FIG. 3. Result for color matching from 8775K monitor to D50 monitor.

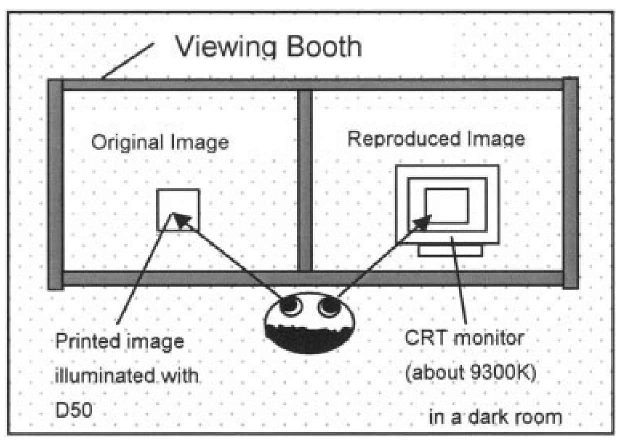

FIG. 4. Color matching between a CRT monitor and a printer.

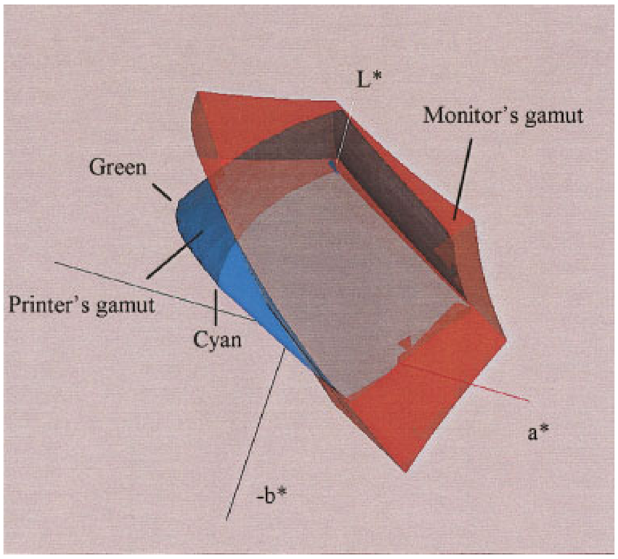

FIG. 5. Comparison of the difference between printer's and CRT monitor's gamuts in CIELAB.

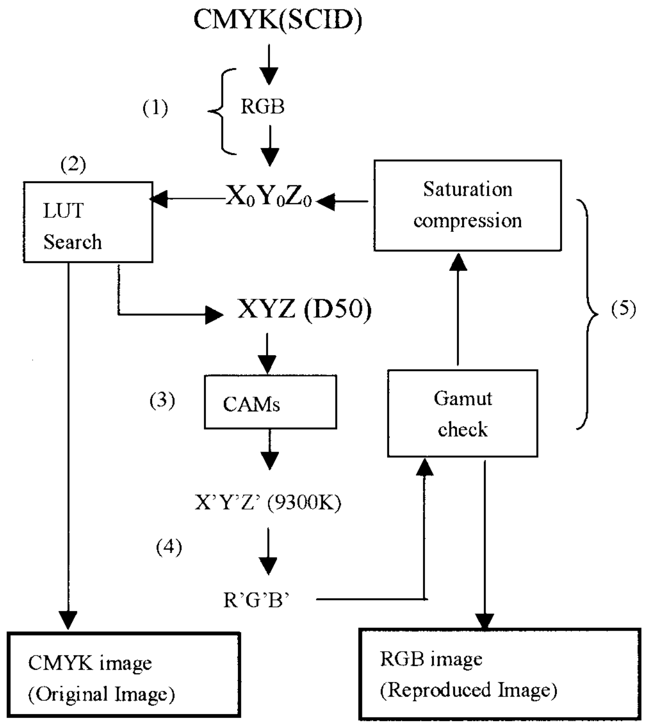

FIG. 6. Procedures for the creation of an original image and a reproduced image.

Ten subjects with normal color vision evaluated the superiority or inferiority in the color matching of images reproduced by these models. The viewing condition was nearly successive-haploscopic viewing because the subjects could not simultaneously see the two images from the parallax.

We followed the paired comparison method to determine the order of the models' performance for color appearance matching. We made reproduced images by using the seven models and displayed two images randomly selected from these seven images on the right monitor. These two images were not simultaneously but alternately displayed on the monitor when subjects clicked their PC mouse.

From the two images displayed on the right monitor, subjects were instructed to select the one which was closer in color appearance to the original image set in the left booth. They were also requested to evaluate whole regions and colors in images and not to concentrate on specific regions or colors. The experiments were repeated twice to improve the accuracy of the data obtained.

Observation conditions in our experiments are summarized in Table III. The parameters used in each of the models are listed in Table IV. In order to ensure fair evaluation, the recommended parameters for each model were selected to match the observation conditions in the experiments.

Interval scales (*Z*-scales) were calculated from the evaluation results for the ten subjects. Figure 7 shows the results of the evaluation experiments. From the figure, it can be seen that SPEM, RLAB, and CIECAM97s, which take account of incomplete chromatic adaptation, produce good results between softcopy (i.e., the white is about 9300K) and hardcopy (i.e., the white is a white paper illuminated with D50). The color appearance matching performance, ranked in order from high to low, is SPEM, RLAB, CIECAM97s, Nayatani, von Kries, CIELAB, and LLAB.

DISCUSSION

In the experiments for two CRT monitors, the ranking of color matching performance from high to low is SPEM, RLAB, CIECAM97s, Nayatani97, von Kries, LLAB, and CIELAB. In the experiment, SPEM produces the highest scores for N1, N2, and N7. RLAB shows the next best performance. Although CIECAM97s produced the highest score for N3, the differences of *Z*-scale between the models were smaller than those of the *Z*-scale for other images. This means that the differences of color appearance between models are slight.

TABLE III. Conditions for color matching between a printed image and a CRT monitor.

	Source conditions	Reference conditions
Media type	Hardcopy (Printed image)	Hardcopy (CRT monitor)
White point	White paper illuminated with D50 (x = 0.3457 and y = 0.3585)	about 9300K (x = 0.2841 and y = 0.2944)
Luminance of white	94.59 cd/m^2	94.06 cd/m^2
Luminance of background	20% of white luminance	20% of white luminance

TABLE IV. Model parameters used in the experiments between a CRT monitor and a printed image.

	Source condition parameters	Reference condition parameters
LLAB	"Relection samples in average surround. Subtending $>$ 4°"	"Television and VDU displays in dim surrounding"
Nayatani97	Background luminance factor: 20%	Background luminance factor: 20%
CIECAM97s	Viewing condition: Dark surrounding	Viewing condition: Dark surrounding
	Adapting field luminance: 20cd/m^2	Adapting field luminance: 19cd/m^2
RLAB	Media: Hardcopy (D = 1.0)	Media: Softcopy (D = 0.0)
	Surrounding: Dark (σ = 1/3.5)	Surrounding: Dark (σ = 1/3.5)
SPEM	Mixing coefficient MC: 0.6	

These results show that human color recognition has a tendency to fall into incomplete chromatic adaptation when we compare colors reproduced by color imaging devices whose whites are quite different, in the absence of information on illumination color. Therefore, SPEM, RLAB, and CIECAM97s, which take account of incomplete chromatic adaptation, produce good results.

Four kinds of natural images were selected in the experiments. From Fig. 3, it is obvious that the results obtained for N3 (Fruit) were quite different from those obtained for the other images. We assume that this is because colors of images affect human chromatic adaptation.

As mentioned above, the subjects in our experiments were instructed to select the image that was closer in color appearance by evaluating whole regions and colors in images and not to concentrate on specific regions or colors. Some subjects, however, commented that prominent area/color combination in the images affected their judgment. For N7 (Musician) for example, a slightly bluish gray background and blue clothing were prominent area/color combinations, as were yellowish fruit and brown basket for N3 (Fruit).

The distributions of hue and chroma in the evaluated images were investigated. The four graphs in Fig. 8 show the distributions of these for each image. The X-axis is chroma $C=\sqrt{a^2+b^2}$ divided into segments of 20 in CIELAB space. The number "0" on the X axis includes pixels for $0 \le C < 20$, "1" includes $20 \le C < 40$ and so on. The Y-axis is hue angle ($H = \arctan(b/a)$) divided into 30-degree segments. The Z-axis shows the number of pixels in an image.

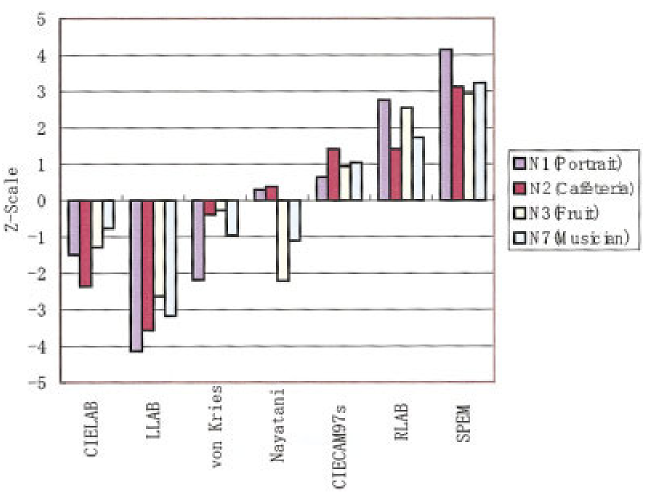

FIG. 7. Result for color matching between a printed image and a CRT monitor.

N1 (Portrait), N2 (Cafeteria), and N7 (Musician) consist of many colors having hue from 240 to 300 and chroma from 0 to 40. On the other hand, N3 (Fruit) includes colors having hue from 0 to 90 and chroma from 0 to 40. The distribution of hue for N3 is obviously different from those of the other images. Colors having hue from 240 to 300, which are prominent in N1, N2, and N7, are bluish colors. Colors having hue from 0 to 90, which are prominent in N3, are colors from red to yellow (i.e., warm colors).

Human chromatic adaptation from higher to lower correlated color temperatures can easily occur for warm colors. For example, this phenomenon was observed in the experimental data obtained by Hunt.[16] N3 mainly comprises warm colors, and subjects' color recognition for N3 in this experiment may reflect a tendency of complete chromatic adaptation.

In the experiments for color matching between a printed image and a CRT monitor, the results for N3 were not different from the other images. We believe that it is due to low saturation of evaluated images that most colors in the images exist in both color gamuts.

The performance of color matching may also depend on the degree of incomplete chromatic adaptation. In RLAB, the degree of incomplete chromatic adaptation is determined by calculating the difference between the illuminant in a viewing condition and the equi-energy illuminant that is the reference illuminant incorporated in the model. That is, the reference white is defined as equi-energy white ($X = Y = Z = 1.0$). The farther the white in a viewing condition is from the reference white, the stronger the degree of incomplete chromatic adaptation becomes.

In SPEM, the mixing coefficient MC is introduced to determine the degree of incomplete adaptation in order to match the white displayed on the left monitor (D50) with the white on the right (8775K). The MC value applied to the environment in this experiment, 0.6, was the same value as that we derived in another experiment[9]; we did not newly select the value to match it the current environment.

Adaptation white points in the color matching experiment from 8775K to D50 by using two CRT monitors are calculated. Figure 9 shows a comparison of adaptation white points for each of the models in terms of uv coordinates.

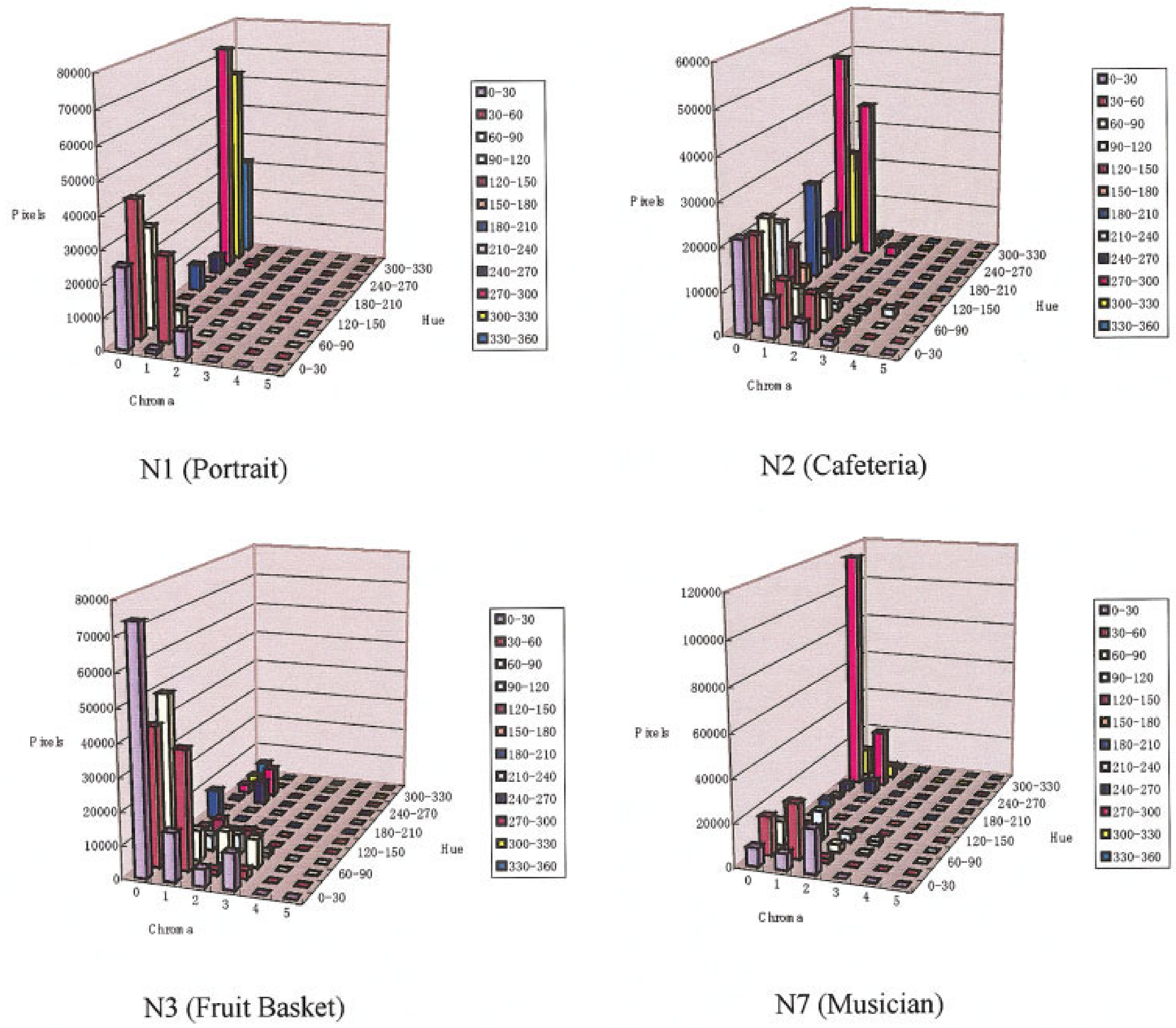

FIG. 8. Distributions of hue and chroma in the evaluated images.

Though CIECAM97s produced the best result for N3 (fruit), SPEM and RLAB produced good scores for the other three images. From Fig. 9, we can see that the adaptation white in CIECAM97s is very near to D50. This means that chromatic adaptation advances toward the target white (i.e., D50) too much. We believe that this may be one of the reasons that CIECAM97s did not produce good results for images other than N3. From the adaptation white points we can see that

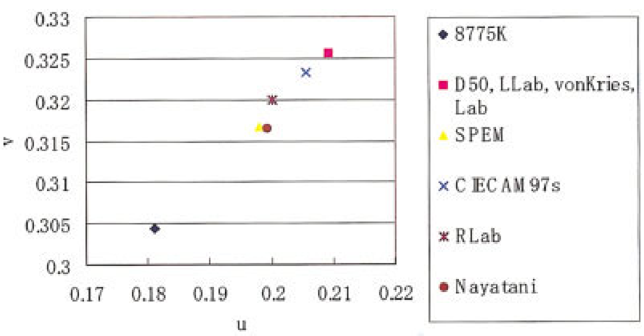

FIG. 9. Comparison of adaptation white points.

states of chromatic adaptation in RLAB and SPEM are less complete than those in CIECAM97s. Therefore, their color matching performances were better than that of CIECAM97s. In the Nayatani97 model, although the adaptation white is the same as that of SPEM, the performance was not especially good. Failure to estimate corresponding colors other than white may have a harmful influence on color matching performance.

Through our experiments, we found that the adaptation white points were mostly in the middle of the straight line that links 8775K and D50 in the *u, v* coordinates in color matching from about 9000K to D50 of natural images displayed on CRT monitors. Though human chromatic adaptation is affected by prominent colors and areas in natural images, SPEM and RLAB provide good practical models for human chromatic adaptation in color matching and produce good results.

The experiments by using two CRT monitors can apply to the case of color matching in a remote color proofing using CRT monitors. They also simulate color matching between an image displayed on a CRT monitor whose white was

8775K and a printed image illuminated with D50. Since two CRT monitors set in a dark room with no illumination were used in the experiment, the viewing conditions did not exactly apply to color matching between softcopy and hardcopy in a real-world situation. We believe, however, that the conditions were appropriate for the evaluations of color appearance models and chromatic adaptation models, because stimuli were presented under well-controlled viewing conditions without influences of surrounding environment such as surrounding light. The parameters for each model we evaluated were appropriately selected and given in the experiments to ensure fair evaluations. Moreover, the results reflect more practical performances of models than can be obtained with color patches, since natural images instead of color patches were used as subjects in the experiments.

We performed practical experiments for color matching between a printed image and a CRT monitor. The ranking of color matching performance in the experiments is SPEM, RLAB, CIECAM97s, Nayatani97, von Kries, CIELAB, and LLAB. SPEM produces the highest scores for all images and RLAB follows SPEM.

As for SPEM, we do not argue that human beings have complete color constancy. The results we obtained indicate that a model based on spectral properties estimation produces good color matching performance in evaluation experiments. Further study is required, however, to ascertain the validity of the model. Systematic calculation of the mixing coefficient *MC* in the model is one subject for future work.

CONCLUSION

We performed subjective experiments to evaluate color matching performance of five color appearance models (CIECAM97s, Nayatani97, RLAB, LLAB, and CIELAB) and two chromatic adaptation models (von Kries and SPEM) in practical use. We found that the adaptation white points greatly affect color matching performance between color imaging devices whose whites are quite different. We also found that our chromatic adaptation has a tendency to be influenced by prominent area and color combination in the evaluated images.

In many cases, SPEM and RLAB provided a good practical model for human chromatic adaptation in color matching and produced good results.

APPENDIX: CHROMATIC ADAPTATION MODEL BASED ON SPECTRAL PROPERTIES ESTIMATION

The detailed algorithm of SPEM is described in this appendix.

Computational theories of color constancy have been proposed.[12–19] The conventional methods are designed so that machine vision, not human vision, can identify an object under different illuminations by estimating spectral properties of the object and the illumination in a scene. That is, the goal of the methods is to realize a mechanism of complete color constancy in a computer system.

We try to apply the basic concept of color constancy to color matching in human vision by introducing the following assumptions.

Assumption 1 In human vision, surface reflectance of an object color in a scene is inferred under the recognition that white in a scene is perceived as its nearest CIE daylight illuminant.

Assumption 2 Since most spectral properties of objects and illuminations show comparatively smooth curves, we can model them as the weighted sum of a small number of vectors.

Assumption 3 White in the image is equal to the illumination in the scene.

As to the assumption 1, unfortunately the exact physiological mechanism of human color recognition has not been elucidated yet since it is extremely complex. However, we feel that this assumption might be appropriate when we consider that human brains learn from experience that objects' colors look similar under different colors of daylight by recovering the surface reflectance that an object fundamentally has. That is, it has been one hundred years at most since artificial illuminations such as the fluorescent lamp were developed, and daylight color is very important for us to recognize objects' colors. In fact, the results under 3400, 6000K and 30 000K illuminants, which are daylight colors, showed almost perfect color constancy in Kuriki's experiments.[20]

Assumptions 2 and 3 are needed to make the color constancy problem solvable. It is well known that spectral properties of daylights and objects can be represented as the weighted sum of a small number of basis vectors.[21,22]

Since the detailed algorithm we propose has been previously described,[9] it is mentioned only briefly here. The chromatic adaptation transform from an input color under an original condition to a corresponding color under a reference condition is described below.

(1) The hypothetical spectral power distribution (HSPD) of the illumination in a scene is calculated by using the prediction equation for the spectral power distribution of a daylight illuminant and the correlated color temperature or chromaticity of the white of a color device reproducing the scene.[23] HSPD under an original condition and that under a reference condition are, respectively, $I(\lambda)$ and $I(\lambda)'$.

(2) The hypothetical surface reflectances (HSR) $O(\lambda)$ of all objects under an original condition are obtained by solving Eq. 1 for $O(\lambda)$.

$$\begin{aligned} X &= \int I(\lambda)O(\lambda)\bar{x}(\lambda)d\lambda, \\ Y &= \int I(\lambda)O(\lambda)\bar{y}(\lambda)d\lambda, \\ Z &= \int I(\lambda)O(\lambda)\bar{z}(\lambda)d\lambda, \end{aligned} \tag{1}$$

where X, Y and Z are tristimulus values of an input color under an original condition and, $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ are color matching functions. To solve Eq. 1, HSR $O(\lambda)$ can be modeled as Eq. 2 by introducing a finite dimensional linear model.

$$O(\lambda) = o_0(\lambda) + a_1 o_1(\lambda) + a_2 o_2(\lambda) + a_3 o_3(\lambda), \quad (2)$$

where $o_0(\lambda)$ is the mean vector and $o_i(\lambda)$ s ($i = 1, 2, 3$) are basis vectors. They are derived from a number of surface reflectances of objects and known parameters. The weighted coefficients a_i ($i = 1, 2, 3$) are unknown parameters representing the color of an object.

An observation equation for HSR can be made by substituting Eq. 2 for $O(\lambda)$ in Eq. 1.

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} M(x,o_1) & M(x,o_2) & M(x,o_3) \\ M(y,o_1) & M(y,o_2) & M(y,o_3) \\ M(z,o_1) & M(z,o_2) & M(z,o_3) \end{pmatrix}^{-1} \begin{pmatrix} X - M(x,o_0) \\ Y - M(y,o_0) \\ Z - M(z,o_0) \end{pmatrix} \quad (3)$$

where $M(x,o_i)$ ($i = 0 \sim 3$) represents an integral term $\int I(\lambda)o_i(\lambda)\bar{x}(\lambda)d\lambda$. We can recover HSRs of all objects in the scene by calculating characteristic parameters a_i for all pixels in the image.

(3) Two HSRs of the complete white are obtained, $O_w(\lambda)$ under an original condition and $O_w'(\lambda)$ under a reference condition. In order to obtain the HSR of the white $O_{wm}(\lambda)$ under a reference condition, which matches the color appearance of the white under an original condition, the mixture of an object's surface reflectances recovered under different illuminations is introduced. We assume that this would be a reasonable mechanism to describe incomplete chromatic adaptation in human color recognition. $O_{wm}(\lambda)$ can be calculated by using the equation below.

$$O_{wm}(\lambda) = MC \times O_w(\lambda) + (1 - MC) \times O'_w(\lambda), \quad (4)$$

where MC is the mixing coefficient ($0.0 \leq MC \leq 1.0$).

(4) In order to calculate the HSR for a color other than white, we define an adjusting function for surface reflectance on wavelength in visible light. Let $rf_{ad}(\lambda)$ denote an adjusting function for surface reflectance.

$$rf_{ad}(\lambda) = \frac{O_{wm}(\lambda)}{O_w(\lambda)} \quad (5)$$

(5) A HSR $O'(\lambda)$ under a reference condition corresponding to a HSR $O(\lambda)$ of an arbitrary input color under an original condition can be calculated by multiplying $O(\lambda)$ by $rf_{ad}(\lambda)$.

$$O'(\lambda) = O(\lambda) \times rf_{ad}(\lambda) \quad (6)$$

(6) The tristimulus values of a color under a reference condition corresponding to an arbitrary input color under an original condition are obtained by Eq. 7.

$$X' = \int I'(\lambda)O'(\lambda)\bar{x}(\lambda)d\lambda$$

$$Y' = \int I'(\lambda)O'(\lambda)\bar{y}(\lambda)d\lambda \quad (7)$$

$$Z' = \int I'(\lambda)O'(\lambda)\bar{z}(\lambda)d\lambda$$

1. Luo MR, Lo MC, Kuo WG. The LLAB(l:c) colour model. Color Res Appl 1995;21:412–429.
2. Fairchild MD. Refinement of the RLAB color space. Color Res Appl 1996;21:338–346.
3. Hunt RWG. An improved predictor of colourfulness in a model of colour vision. Color Res Appl 1994;19:23–26.
4. CIE technical report: CIE109; A method of predicting corresponding colours under different chromatic and illuminance adaptations. Vienna, CIE Central Bureau, 1994.
5. Sobagaki H, Nayatani Y. Field trials of the CIE chroma-adaptation transform. Color Res Appl 1998;23:78–91.
6. Luo MR, Hunt RW. The structure of the CIE 1997 colour appearance model (CIECAM97s). Color Res Appl 1998;23:138–146.
7. Katoh N. Practical method for appearance match between soft copy and hard copy. Vol. 2170. SPIE; 1994. p 170–181.
8. Nayatani Y. Simple estimation method for effective adaptation coefficient. Color Res Appl 1997;22:259–268.
9. Tsukada M, Tajima J, Yoshikawa H, Yaguchi H. Color appearance matching based on color constancy theory. Proc IS&T 1999 PICS. 1999; p 348–353.
10. Tsukada M, Tajima J. Chromatic adaptation model based on spectral property estimation and its color matching performance. Proc IS&T 1st CGIV. 2002; p 11-15.
11. CIE guideline for coordinated research on evaluation of colour appearance models for reflection print and self-luminous display image comparisons. Color Res Appl 1994;19:49–58.
12. International color consortium: Specification ICC.1:1998-09 file format for color profiles. http://www.color.org/profiles.html.
13. Graphic technology—Prepress digital data exchange—Standard colour image data (SCID). ISO/JIS-SCID, JIS X 9201-1995. Japanese Industrial Standard; 1995.
14. Fairchild MD, Pirrotta E, Kim T. Successive-Ganzfeld haploscopic viewing technique for color-appearance research. Color Res Appl 1994;19:214–221.
15. Ohkushi K, Nakayama T, Fukuda T. Evaluation techniques for picture and sound qualities. Shokodo; 1991, p 34–47.
16. Hunt RWG. Measurement of color appearance. J Opt Soc Am. 1965; 55:1540–1551.
17. Maloney LT, Wandell BA. Color constancy: a method for recovering surface spectral reflectance. J Opt Soc Am A. 1986;3:29–33.
18. Gershon R, Jepson AD, Tsotsos JK. From R, G, B to surface reflectance: Computing color constancy descriptors in images. Proc 10th Int Joint Conf on Artificial Intelligence. 1987; p 755–758.
19. Tsukada M, Ohta Y. An approach to color constancy using multiple images. Proc IEEE 3rd Int Conf on Computer Vision. 1990; p 385–389.
20. Kuriki I, Uchikawa K. Limitations of surface-color and apparent-color constancy. J Opt Soc Am A 1996;13:1622–1636.
21. Parkkienen JPS, Hallikainen J, Jaaslainen T. Characteristic spectra of munsell color chips. J Opt Soc Am A 1989;6:318–322.
22. Judd DB, MacAdam DL, Wyszecki G. Spectral distribution of typical daylight as a function of correlated color temperature. J Opt Soc Am 1964;54:1031–1040.
23. Handbook of color science, 2nd ed. University of Tokyo Press; 1998; p 69–71.