Color Appearance Modeling for Single Color Stimuli in A Simple Viewing Field under A Dim Surround

Ruey-Kuen Perng∗, Mei-Chun Lo∗, Yu Su†, Yui-Liang Chen∗ and Chung-Chi Wang∗

Key Words: Device Characterization, color appearance model, paired comparison, category judgement, color management system.

Abstract: The goal of this study is the evaluation of various color models for achieving accurate color rendition between a self-luminous reproduced color (softcopy) and a original surface color (hardcopy) in a darkened room under a dim surround. This work is emanated from an earlier study, presented in TAGA’99 Vancouver (Lo et al. 1999), in which a similar experiment was carried out in a bright room under average surround. The color models tested were CIELAB, von Kries, ZLAB, LLAB, Hunt and CIECAM97 models. The whole experiment was divided into 4 phases. A set of 36 reflection single stimuli was chosen to be the original colors.

A forced-choice paired-comparison experiment was performed. Observers assessed the color-fidelity quality of each reproduced color, processed using a particular model of interest, in a simple field using a 7-point category scale. The laws of both comparative judgement and category judgement were applied to analyze the visual data. Analysis shows that the CIELAB (also CIEXYZ) system is adequate to implement color matches across media for the similar viewing conditions of softcopy and hardcopy fields. For asymmetric viewing conditions wherein different chromaticities were used for hardcopy and softcopy reference whites, the CIECAM97’s model overall performed better than other models. Moreover, the results obtained using the category ranking method indicate that those good performing models also gave satisfactory color-fidelity quality. Successively, a set of experiments will be carried out to achieve the ultimate goal of establishing a reliable color model, capable of predicting the change of perceived appearance of colors under a wide range of media/viewing conditions.

INTRODUCTION

“What You See Is What You Get” (WYSIWYG), the essence of accurate reproduction color images across a wide variety of media and applications, has already become an urgent demand and recently been substantially researched on in the color imaging industry. It is one of the main concerns in both desktop and pre-press environments.

From a detailed retrospective examination of color-encoding problems, three underlying obstacles however are often encountered in achieving WYSIWYG. These are device dependency, color gamut mismatch among dissimilar color imaging devices, and variations of color appearance under different media/viewing conditions. The practical solution in the process of a state-of-the-art cross-media color image production, dealing with these problems, is to provide software known as a color management system (CMS). Three essential elements, device characterization, color appearance modeling, and gamut mapping should be therefore included in a CMS.

In a practical sense, traditional CIE colorimetry alone, basically only dealing with quantities derived from physical properties of individual stimuli, does not represent color appearance unambiguously. The same

∗ Dept. of Graphic Communications and Technology, Shih Hsin University, Taiwan, R.O.C.
† Graphic Arts Section, Master of Graduate Institute of Paper-making & Graphic Arts, Chinese Culture University, Taiwan, R.O.C.
color stimulus in terms of CIE XYZ tristimulus or CIELAB L*a*b* values could have entirely different appearances under disparate media/viewing environments. As far as the global conditions under which a color is viewed are concerned, some critical factors affecting its appearance include the medium type, the white point, the background, the surround, the adapting stimulus, the luminance level, and cognitive discounting the illuminant etc. This is a critical issue needed concerned due to limitless possible sets of media/viewing conditions involved in practical situations. To deal with these complications just mentioned, the internationally practical solution is to derive appearance-based color-encoding models. Thus, the perceptual color appearance attributes of stimuli of interest can be predicted by taking into account the influence of the environment under which they are viewed.

Recently, CIE TC1-31 has completed a series of tests to investigate the performance of various previously published color appearance models and data sets. A new model, based on the results of tests, was formulated by incorporating the best features from these existing color models and announced at the 1997 CIE Division 1 meeting in Kyoto, Japan (Hunt 1997). Two versions of this model were derived due to practical and effective needs. These are simple and comprehensive versions of CIECAM97 designated as CIECAM97s and CIECAM97c models respectively. In this study, a work involving color appearance modeling was encompassed. A set of experiments was conducted under dim surround using single stimuli in a simple field. The CIECAM97s was compared with a variety of existing color models for achieving high color fidelity between a reproduced self-luminous color (softcopy) and the original surface color (hardcopy). In a near future, a more advanced study will be performed using complex images, which are practically often used in areas such as graphic arts and desktop publishing. Based on the findings obtained using both single stimuli and complex images, the best performing color model found (assuming CIECAM97s) will be verified. These works are consecutively conducted closely to link to the task of CIE TC1-34, Testing Color Appearance Models, established to test various models for the prediction of the color appearance, under various media/viewing conditions and phenomena, of object colors (Fairchild 1997). The ultimate goal of this research, as expected by CIE TC1-34, is that, at some future day, a more accurate and/or theoretical-based model might be evolved.

COLOR APPEARANCE MODELS TESTED

In this study, color models tested were classified into three categories as follows.

- Uniform Color System: CIELAB (CIE 1986)
- Chromatic-Adaptation-Transforms: von Kries (Helson et al. 1952)

EXPERIMENTAL PREPARATION

Device Characterization

A Barco monitor was used to display screen colors under a dim surrounding. The Barco monitor, internally set white point to the color temperature tested in each particular phase in this study, was characterized using the PLCC (Piecewise Linear interpolation assuming Constant Chromaticity Coordinate) characterization model. (Lo et al. 1998). A characterization data set of 54 color patches was created for each white point of interest. This data set was composed of 18 sample patches produced for each of red, green, and blue channels in terms of 15 Digital to Analog (DAC) values ranging from 0 to 255. The measurement equipment used was a PR650 SpectraColorimeter, manufactured by Photo Research.

Image Preparation and Processing

A set of 36 single color patches, in terms of 3 variations in lightness and chroma approximately at 12 different hues in CIELAB color space (shown in Fig.1), was chosen from Pantone Color Selector to be the original surface colors (hardcopies). Those covered a wide color space and had V values ranging from 5 to 95. The CIE XYZ tristimulus values of the color patches, under the viewing conditions considered in the experiments, were measured using PR650 SpectraColorimeter.
Image processing software was developed to correlate the hardcopy’s XYZ values (in the reference field) to the monitor’s RGB intensities (i.e. DACs) (in the test field) for a particular color model tested on a pixel by pixel basis. The computational process is illustrated in Fig. 2. The XYZ and X’Y’Z’ tristimulus values specifically form a set of corresponding colors, having the same color appearance when viewed under adapting fields of hardcopy (reference) and softcopy (test) viewing conditions respectively. The monitor characterization model described earlier were applied to convert X’Y’Z’ to RGB intensities.

For different categories of color models, the computational procedures are varied. For chromatic adaptation transforms, the XYZ values were directly transformed to X’Y’Z’. As for each uniform color space or color appearance model, the predicted lightness, chroma, and hue are first calculated via its forward model and followed by computing its corresponding X’Y’Z’ using the model’s reverse.

**EXPERIMENT**

**Viewing Conditions**

The experiment was divided into 4 phases according to different light sources and white points used in the viewing reflection hardcopies (single patches in the reference field) and screen softcopies (single patches in the test field). Table 1 summaries the differences among all phases. A panel of 5 observers attended and repeatedly made the assessments twice in each of 4 phases. In total, 21, 600 comparisons were made.

The correlated color temperatures (CCTs), luminances and colorimetric data used for both hardcopy (reference) and softcopy (test) fields in each phase under a dim surround are tabulated in Table 2. The luminance levels of white points used for both screen display and viewing hardcopy are also listed in Table 2.
Fig. 2  Computational procedures of image processing correlating the hardcopy’s XYZ values to the monitor’s RGB intensities for each particular color model tested. The UCS, CAM and CAT processes represent uniform color space, color appearance model and chromatic adaptation transform respectively.

Table 1  Summary of experimental phases.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Hard-copy</th>
<th>Soft-copy</th>
<th>VTa</th>
<th>MDb</th>
<th>No. of observers</th>
<th>Repetition</th>
<th>Pairs</th>
<th>No. of comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>D50</td>
<td>D65</td>
<td>BSM</td>
<td>SS</td>
<td>5</td>
<td>2</td>
<td>1,080</td>
<td>5,400</td>
</tr>
<tr>
<td>2</td>
<td>D50</td>
<td>D93</td>
<td>BSM</td>
<td>SS</td>
<td>5</td>
<td>2</td>
<td>1,080</td>
<td>5,400</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>D65</td>
<td>BSM</td>
<td>SS</td>
<td>5</td>
<td>2</td>
<td>1,080</td>
<td>5,400</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>D93</td>
<td>BSM</td>
<td>SS</td>
<td>5</td>
<td>2</td>
<td>1,080</td>
<td>5,400</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td>4,320</td>
<td>21,600</td>
</tr>
</tbody>
</table>

VT: viewing technique; MD: monitor display; aBSM : binocular simultaneous matching
bSS : simultaneous display.
Table 2  Colorimetric data and luminances used for both hardcopy (reference) and softcopy (test) fields in each phase under an average surround.

<table>
<thead>
<tr>
<th>Phase</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hardcopy field</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White point</td>
<td>D50</td>
<td>D50</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>CCT</td>
<td>4716K</td>
<td>4716K</td>
<td>2420K</td>
<td>2420K</td>
</tr>
<tr>
<td>L (cd/m²)</td>
<td>65.89</td>
<td>65.89</td>
<td>55.57</td>
<td>55.57</td>
</tr>
<tr>
<td>X</td>
<td>96.04</td>
<td>96.04</td>
<td>116.36</td>
<td>116.36</td>
</tr>
<tr>
<td>Y</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Z</td>
<td>74.35</td>
<td>74.35</td>
<td>22.71</td>
<td>22.71</td>
</tr>
<tr>
<td>X</td>
<td>0.3552</td>
<td>0.3552</td>
<td>0.4867</td>
<td>0.4867</td>
</tr>
<tr>
<td>Y</td>
<td>0.3698</td>
<td>0.3698</td>
<td>0.4183</td>
<td>0.4183</td>
</tr>
<tr>
<td>u'</td>
<td>0.2112</td>
<td>0.2112</td>
<td>0.2763</td>
<td>0.2763</td>
</tr>
<tr>
<td>v'</td>
<td>0.4948</td>
<td>0.4948</td>
<td>0.5343</td>
<td>0.5343</td>
</tr>
</tbody>
</table>

| Softcopy field |         |         |         |         |
| White point | D65 | D93 | D65 | D93 |
| CCT     | 6214K  | 8474K  | 6214K  | 8474K  |
| L (cd/m²) | 65.74 | 62.70 | 65.74 | 62.70 |
| X       | 94.72  | 94.40  | 94.72  | 94.40  |
| Y       | 100.00 | 100.00 | 100.00 | 100.00 |
| Z       | 103.61 | 136.00 | 103.61 | 136.00 |
| X       | 0.3175 | 0.2857 | 0.3175 | 0.2857 |
| Y       | 0.3698 | 0.3027 | 0.3352 | 0.3027 |
| u'      | 0.2112 | 0.1886 | 0.1988 | 0.1886 |
| v'      | 0.4948 | 0.4495 | 0.4723 | 0.4495 |

Note: CCT: correlated color temperature

**Viewing Configuration and Viewing Techniques**

Fig. 3 illustrates the display arrangements configured in the experiment. The experiment was carried out in a darkened room. The experimental viewing pattern basically included a test color being surrounded by 12 color patches, which were randomly selected from the Pantone Color Paper Selector, along with the “reference white”. Two reproduced viewing patterns of softcopies in the test field were displayed side by side. Each reproduced pattern had equal size to that of the original hardcopies illuminated in a VeriVide viewing cabinet.
The binocular simultaneous matching (BSM) technique was applied for all experimental phases. This technique is the most common method used in industry. Each observer sat in a darkened room under a dim surrounding, at approximately 100 cm from the monitor and original considered. The softcopy images and the illuminated reflection hardcopy were located side by side and coplanar. Each observer was instructed to use both eyes to look at either the test or the reference fields at a time, but could switch between two fields at any time.

The colorimetric data and luminance of the gray background used in the hardcopy field are listed in Table 3. The color of this gray background was transformed into the softcopy’s background color for each phase. Different models may produce different background colors. In our earlier study (Lo et al. 1999), one background color was used, by averaging those of the background colors predicted from all the models in each particular phase, to avoid incomplete adaptation. In this study, the background color was processed for each model. Therefore, two background colors for a pair of reproduced softcopies may be different. The purpose of this arrangement is to test the model’s capability for predicting the effect of incomplete adaptation. Table 4 lists the parameters used in each of color appearance models tested, accounting for effects of such as cognitive or sensory chromatic mechanism, surroundings etc. It is noted that the D parameter in LLAB model and the F parameter in ZLAB and CIECAM97s models are adjusted according to the viewing condition for the hardcopy field and the white point for the softcopy field. Different parameter values represent different states of adaptation. The larger value of the parameter means the higher degree of adaptation predicted by the model.

Table 3  Colorimetric data, luminance of gray background used in the reference (hardcopy) field.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Source</th>
<th>X</th>
<th>Y</th>
<th>L</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phases 1 &amp; 2</td>
<td>D50</td>
<td>0.3493</td>
<td>0.3638</td>
<td>16.29</td>
<td>23.74</td>
<td>24.72</td>
<td>19.50</td>
</tr>
<tr>
<td>Phases 3 &amp; 4</td>
<td>A</td>
<td>0.4892</td>
<td>0.4140</td>
<td>14.31</td>
<td>30.43</td>
<td>25.75</td>
<td>6.02</td>
</tr>
</tbody>
</table>
### Table 4  Parameters used for Hunt, LLAB, ZLAB and CIECAM97s color appearance models in the experiment.

<table>
<thead>
<tr>
<th>Phase</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hunt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hardcopy Field</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$L_A$</td>
<td>16.29</td>
<td>16.29</td>
<td>14.31</td>
<td>14.31</td>
</tr>
<tr>
<td>Flas</td>
<td>0.92</td>
<td>0.92</td>
<td>0.59</td>
<td>0.59</td>
</tr>
<tr>
<td>$N_c$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$N_b$</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>Soft Field</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$L_A$</td>
<td>16.29</td>
<td>16.29</td>
<td>14.31</td>
<td>14.31</td>
</tr>
<tr>
<td>Flas</td>
<td>1.049</td>
<td>1.198</td>
<td>1.049</td>
<td>1.198</td>
</tr>
<tr>
<td>$N_c$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$N_b$</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
</tbody>
</table>

$L_A$: Photopic luminance of adapting field  
Flas: Scotopic luminance level conversion factor  
$N_c$: Chromatic surround induction factor  
$N_b$: Brightness surround induction factor

<table>
<thead>
<tr>
<th>LLAB</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hardcopy Field</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D$</td>
<td>0.747</td>
<td>0.373</td>
<td>0.41</td>
<td>0.41</td>
</tr>
<tr>
<td>$F_S$</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>$F_L$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$F_C$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Soft Field</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D$</td>
<td>0.747</td>
<td>0.373</td>
<td>0.41</td>
<td>0.41</td>
</tr>
<tr>
<td>$F_S$</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
</tr>
<tr>
<td>$F_L$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$F_C$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

$D$: Discounting-the-illuminant factor  
$F_S$: Surround induction factor  
$F_L$: Lightness induction factor  
$F_C$: Chroma induction factor
Table 4  Parameters used for Hunt, LLAB, ZLAB and CIECAM97s color appearance models in the experiment. (Continued)

<table>
<thead>
<tr>
<th>Phase</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ZLAB</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hardcopy Field</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$L_A$</td>
<td>16.29</td>
<td>16.29</td>
<td>14.31</td>
<td>14.31</td>
</tr>
<tr>
<td>$F$</td>
<td>0.9</td>
<td>0.45</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>$D$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$1/2\sigma$</td>
<td>0.345</td>
<td>0.345</td>
<td>0.345</td>
<td>0.345</td>
</tr>
<tr>
<td>Soft Field</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$L_A$</td>
<td>16.29</td>
<td>16.29</td>
<td>14.31</td>
<td>14.31</td>
</tr>
<tr>
<td>$F$</td>
<td>0.9</td>
<td>0.45</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>$D$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$1/2\sigma$</td>
<td>0.295</td>
<td>0.295</td>
<td>0.295</td>
<td>0.295</td>
</tr>
<tr>
<td><strong>CIECAM97s</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hardcopy Field</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$L_A$</td>
<td>16.29</td>
<td>16.29</td>
<td>14.31</td>
<td>14.31</td>
</tr>
<tr>
<td>$F$</td>
<td>0.8</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>$C$</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
</tr>
<tr>
<td>$F_{LL}$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$N_c$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$D$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Soft Field</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$L_A$</td>
<td>17.42</td>
<td>16.72</td>
<td>6.767</td>
<td>6.767</td>
</tr>
<tr>
<td>$F$</td>
<td>0.8</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>$C$</td>
<td>0.59</td>
<td>0.59</td>
<td>0.59</td>
<td>0.59</td>
</tr>
<tr>
<td>$F_{LL}$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$N_c$</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$D$</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
</tr>
</tbody>
</table>

$L_A$ : Luminance of adapting field  
$F$ : Factor for degree of adaptation  
$C$ : Impact of surround  
$F_{LL}$ : Lightness contrast factor  
$N_c$ : Chromatic-induction factor  
$D$ : Degree of adaptation  
$\sigma$ : Impact of surround
DATA ANALYSIS

A forced-choice paired comparison method was employed in this experiment. It was based on the judgements made for the color-fidelity quality of test colors reproduced by color models of interest. A panel of 5 observers viewed a paired of reproduced softcopies, and judged which of the two gave a better match (i.e. color fidelity) to the original hardcopy (as shown in Fig. 3). In addition, they also assessed the degree of match of the softcopy against the hardcopy using a predetermined equi-interval of 7-point category scale for overall color fidelity. Each observer entered visual results via radio buttons numbered from 1 to 7, which were located underneath each corresponding viewing pattern. The resulted visual data were then analyzed using the laws of both category judgement (Togerson 1958) and comparative judgement (Thurstone 1927, Bartleson and Grum 1984).

The ranking method of the category judgement yields ordinal data. Theoretically, an ordinal scale involves assigning stimuli on a limited usually numerical scale correlating with their magnitude for a specified attribute. The rule for assigning number on an ordinal scale is that the ordinal position (rank order) of numbers on the scale must represent the rank order of psychological attributes of the stimuli of interest. In this study, the 7-point category scale for color fidelity is defined from 1 (exact match), through 4 (acceptable match) to 7 (awful match) as below.

<table>
<thead>
<tr>
<th>Category Rank</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Category</td>
<td>Exact Match</td>
<td>Good Match</td>
<td>Moderate Match</td>
<td>Acceptable Match</td>
<td>Poor Match</td>
<td>Bad Match</td>
<td>Awful Match</td>
</tr>
</tbody>
</table>

Note that the law of category judgement relates the relative position of a specified attribute, of the stimuli considered with respect to category boundaries on the psychological continuum rather than with respect to one another as Thurstone’s law of comparative judgement does. The results obtained using this approach would, hence, serve as a basis to locate the models tested on the absolute scales of color-fidelity quality. The models tested are considered to have similar performances if they were located on the same category of color fidelity. Therefore an identical category rank number would designate these models.

RESULTS AND DISCUSSION

The results in terms of z score obtained using the law of comparative judgement are summarized in Table 5. A model tested is considered not to be significantly different from another if its z score is within the 95% confidence limit (CL, i.e. ±2 units of standard deviations) of the other. Hence, these would be ranked in the same order. The overall results are also depicted in Fig. 4 for Phases 1 to 4. T height of the rectangle shown in Fig. 4 represents the z score of a model of interest, and a line drawn indicates its 95% confidence limit (CL).

From the overall results shown in Table 5 and Fig. 4, the CIECAM97s model outperformed the others in each of all 4 phases. Both LLAB and ZLAB models has the same performance with rank 2 for phase 2 shown in Table 5. In phase 3 and 4 the performance of ZLAB model was slightly better that of LLAB model, and both model performed superior to other models except the CIECAM97s. The Hunt model was judged better than other models except the CIECAM97s model for phase 1. It is noted that the CIELAB model got the lowest rank for each phase.

As mentioned earlier, the raw visual data from this experiment were also transformed to an interval scale using a method described by Torgerson (1958). Empirical estimates of both the scale values of the color models tested and category boundaries of 7-point scale were obtained and summarized in Table 6. The data for category scales are not normally distributed as those are obtained on an ordinal scale as mentioned above. Nevertheless, it is assumed that the data would be forced to fit the statistical normal distribution, accomplished by dealing with the cumulated distribution, if they were transformed in Torgenson’s method.
Table 5  Paired comparison results in terms of z-score scale based on the judgements made for the overall color-fidelity accuracy obtained from all 36 test colors combined.  (95% CL = ±0.4383 for all phases)

<table>
<thead>
<tr>
<th>Phase</th>
<th>Hard-copy</th>
<th>Soft-copy</th>
<th>Von Kries</th>
<th>CIELAB</th>
<th>ZLAB</th>
<th>LLAB</th>
<th>Hunt</th>
<th>CIECAM97s</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>D50</td>
<td>D65</td>
<td>-0.2016</td>
<td>-3.1062</td>
<td>-0.1634</td>
<td>-1.0388</td>
<td>1.3198</td>
<td>3.1903</td>
</tr>
<tr>
<td></td>
<td>Rank</td>
<td></td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>D50</td>
<td>D93</td>
<td>-3.3040</td>
<td>-5.3089</td>
<td>2.9114</td>
<td>2.6717</td>
<td>-1.5491</td>
<td>4.5789</td>
</tr>
<tr>
<td></td>
<td>Rank</td>
<td></td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>D65</td>
<td>-5.0164</td>
<td>-5.3486</td>
<td>3.5282</td>
<td>2.4645</td>
<td>-3.6046</td>
<td>7.9768</td>
</tr>
<tr>
<td></td>
<td>Rank</td>
<td></td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>D93</td>
<td>-4.2963</td>
<td>-4.6962</td>
<td>3.7093</td>
<td>2.5764</td>
<td>-3.8647</td>
<td>6.5714</td>
</tr>
<tr>
<td></td>
<td>Rank</td>
<td></td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: The underlined figure indicates the best performing model in a particular phase, and CL represents confidence limit.

Fig. 4  Color models’ performance evaluated using the law of comparative judgement for the overall color-fidelity accuracy of all 36 test colors from Phases 1 to 4 (including 95% confidence limit).
It is shown that the CIECAM97s model, performing the best found in the results obtained using the law of comparative judgement, also had top rank in those obtained using the law of category judgement for each of 4 phases tested. The CIECAM97s, LLAB, and ZLAB models were judged with category scale of “moderate match” or “acceptable match” for each of 4 phases tested. The results are consistent with those of our earlier study (Lo et al. 1999). The results show that the CIECAM97s, LLAB, and ZLAB models gave better prediction than other models when the white points for the softcopy and the hard copy fields are dissimilar. The results listed in Table 6 also show that all models in phase 1 were located at least on the category scale of “acceptable match”. The white point of the softcopy field is D65 and the light source for the hardcopy is D50 approximately for phase 1, these similar viewing conditions imply the chromatic and lightness adaptation become more complete as compared with those in other phases. Therefore, a satisfactory visual match could be provided by the von Kries and the CIELAB models, in which complete adaptation is assumed. Moreover, all models outperforming or giving average predictions, were located on the category scale of “moderate match” or “acceptable match”.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Hard Copy</th>
<th>Soft Copy</th>
<th>Boundary estimates (T_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>T_1</td>
</tr>
<tr>
<td>1</td>
<td>D50</td>
<td>D65</td>
<td>0.0000</td>
</tr>
<tr>
<td>2</td>
<td>D50</td>
<td>D93</td>
<td>0.0000</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>D65</td>
<td>0.0000</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>D93</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phase</th>
<th>Hard Copy</th>
<th>Soft Copy</th>
<th>Category scales</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Von Kries</td>
</tr>
<tr>
<td>1</td>
<td>D50</td>
<td>D65</td>
<td>2.2391</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Category Rank</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>D50</td>
<td>D93</td>
<td>1.7429</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Category Rank</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>D65</td>
<td>1.0977</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Category Rank</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>D93</td>
<td>0.7363</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Category Rank</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
</tbody>
</table>

Note: The underlined figure indicates the best performing model in a particular phase.
CONCLUSION

This study closely followed up the task of CIE TC1-34, Testing Color Appearance Models, established to test various models for the prediction of the color appearance of object colors. The experiment proceeded in a darkened room under a dim surround, and was an extended study emanated from our earlier work (Lo et al. 1999) in which the experiment was carried out in a bright room under average surround. The experiment was divided into 4 phases according to different light sources and white points used in the viewing reflection hardcopies and screen softcopies. Both the paired comparison and the category ranking techniques were applied to analyze the raw visual results, obtained by a 7-point category scale for color-fidelity. The evaluation of results shows that the CIECAM97s, LLAB, ZLAB models gave satisfactory predictions for all experimental phases. It also suggests that the CIELAB (also CIEXYZ) system should be accurate enough to fulfill color matches across media for the similar viewing conditions (e.g. showing in Phase 1 in this study). This is also agreed with the results found in the earlier study (Lo et al.). In other words, a simple colorimetric match among media is practically adequate for the majority of color imaging applications wherein similar viewing conditions are used. For asymmetric viewing conditions wherein different chromaticities were used for the hardcopy’s and the softcopy’s reference whites, the CIECAM97’s model overall performed better than other models. Moreover, the color fidelity of colors produced using the best performing model found in each of 4 phases, i.e. the CIECAM97s model, was all at the category scale of “moderate match”.

The work in the test of color appearance models’ performance will be further extended to an advanced testing using complex images practically often used. The findings obtained will be compared and/or conformed to those found using single stimuli and based on to further verify the best performing color model (assuming CIECAM97s). These anticipated scientific insights will, at some time in the future as expected by CIE TC1-34, allow the derivation of more theoretically correct models, capable of predicting the change of perceived attributes of color appearance under various media/viewing conditions and phenomena.

ACKNOWLEDGMENTS

This research has been mainly financed by the Industrial Technology Research Institute, Taiwan, R. O. C.

REFERENCE

Bartleson, C. J. and Grum, F.

Braun, K. and Fairchild, M. D.

CIE
1986 Colorimetry. 2nd edition, CIE Publication No. 15.2, Central Bureau of the CIE.

Fairchild M. D.


Helson, H., Judd, D. B. and Warren, M. H.

Hunt, R. W. G.

Hunt, R. W. G.
1997 “Comparison of the structures and performances of colour appearance models,” AIC Color 97, Kyoto, pp. 171-174.

Lo, M. C.

Lo, M. C., Chen, Y. L. and Lee, T. T.
1998 “The characterization models and quantification of color appearance on
HDTV,” *IPC ’98 (International Photonics Conference), Taiwan*, pp. 461-463.

Lo, M. C., Luo, M. R. and Rhodes, P. A.  

Lo, M. C., Perng, R. K., Su, Y. and Chiang, J. C.  

Luo, M. R., Lo, M. C. and Kuo, W. G.  

Thurstone, L. L.  

Torgerson, W. S.  