Gamut Compression and Extension Algorithms Based on Observer Experimental Data

Byoung-Ho Kang, Ján Morovic, M. Ronnier Luo, and Maeng-Sub Cho

Gamut compression algorithms have traditionally been defined functionally and then tested with deductive methods, e.g., psychophysical experiments. Our study offers an alternative, an inductive method, in which observers judge image colors to represent the original images more accurately. We developed a computer-controlled interactive tool that modifies the color appearance of pictorial images displayed on a monitor. In experiments, observers used the tool to alter color pixels according to the region of color space to which they belonged. We created three different gamut compression algorithms based on the observer experimental data. Observer groups evaluated the performance of the newly-developed algorithms, existing gamut compression algorithms, and an image based on the average observers' results from experiments in this study.

The study of gamut extension is unlike the study of gamut compression in that it mainly deals with the degree of image pleasantness as judged by observers. The results of the gamut extension experiments in this study not only make available worthwhile data but also suggest a methodology for using the observer experimental tool for future gamut extension research.

Manuscript received June 28, 2002; revised Dec. 7, 2002.

Byoung-Ho Kang (phone: +82 42 860 5036, e-mail: bhkang@etri.re.kr) and Maeng-Sub Cho (e-mail: choms@etri.re.kr) are with Computer Graphics Research Team, ETRI, Daejeon, Korea.

Ján Morovic (e-mail: J.Morovic@derby.ac.uk) and M. Ronnier Luo (e-mail: M.R.Luo@derby.ac.uk) are with Color and Imaging Institute (CII), University of Derby, Derby, United Kingdom.

I. Introduction

In color image reproduction across different media, each medium can produce only a subset of the visible color space—the medium gamut. Hence, successful color image reproduction must ensure that all colors can be transformed to colors that can be reproduced. The process of mapping colors from an original medium to fit the gamut of a reproduction medium is called gamut mapping and can be of two major types—gamut clipping and gamut compression. Gamut clipping alters out-of-gamut colors so as to render them reproducible, whereas gamut compression is applied to all original colors and can modify even colors which by themselves would be reproducible so as to distribute gamut differences across the entire range [1].

Numerous gamut mapping approaches have been proposed and examined in the past. In most of these studies, algorithms were first defined and then evaluated by observers judging the algorithms' suitability for a given reproduction intent. An alternative to this approach is to give the observers the opportunity to adjust image colors to better represent the original. Ebner and Fairchild [2] adopted this approach when they asked observers to adjust the color of uniformly colored simple images from a smaller gamut to make them more similar to corresponding original images from a larger gamut. This study resulted in a set of experimental data on how human observers perform gamut clipping. For gamut compression, the closest data of this type comes from a study that analyzed reproductions made by skilled scanner operators and then modeled the results [3]. However, there have not yet been any studies in which observers could interactively adjust reproductions of complex images and thereby explore the possibilities of gamut compression.

Our investigation has two primary aims: first, to achieve a reliable gamut compression algorithm (GCA), and second, to demonstrate the extended applicability of our interactive tool, which modifies the color appearance of images on a monitor.

Five tasks are included in the first aim: i) to develop an interactive experimental tool for acquiring experimental data, ii) to develop GCAs on the basis of the experimental data, iii) to evaluate existing and our newly-derived GCAs, iv) to develop a generally usable algorithm from the best one in the previous experiment, and v) to compare and evaluate the generalized algorithm with other GCAs.

To complete our tasks, we performed three experiments. In the first experiment, observers were asked to match an original image within a CRT gamut by adjusting a second image within a printer gamut, which was much smaller than that of the CRT gamut, using an interactive tool whereby both images were displayed on a CRT. We developed three GCAs based upon the resulting experimental data.

In the second experiment we evaluated the performance of the newly-derived GCAs and six existing algorithms: the LCLIP, which linearly compresses lightness and then clips out-of-gamut colors onto the gamut boundary and leaves the chroma of in-gamut colors unchanged [4]; the LCUSP, which linearly compresses lightness and then maps lightness and chroma towards the cusp simultaneously; the GCUSP, which compresses the lightness of a color in a chroma-dependent way using a function resembling a Gaussian distribution [5]; the SLIN, which linearly compresses lightness and then maps lightness and chroma towards L^*=50 simultaneously; the LLIN, which linearly compresses chroma along lines of constant lightness; and the CLLIN, which first compresses chroma and then maps lightness. We also evaluated an image based on the average observers' results from the first experiment. These algorithms were divided into two groups: sequential and simultaneous. The sequential algorithms consisted of the LLIN and LCLIP. The simultaneous algorithms include the CUSP (GCUSP and LCUSP) and the SLIN. The CUSP algorithms map towards a particular point, the cusp (the color with maximum chroma in a plane of constant hue) of the reproduction gamut, and the SLIN algorithm maps towards L^*=50. From the GCA with the best performance, we created a generalized algorithm to use on any original or reproduction media.

The third experiment to achieve our first aim was a psychophysical experiment carried out to verify the generalized GCA.

The second aim of this study was to demonstrate the extended applicability of the experimental tool used in the color imaging manipulation experiment of gamut compression by applying it to the study of gamut extension (GE). Two experiments were conducted: experiments GE1 and GE2. With experiment GE1, we acquired observer data on the extension of gamuts in terms of color pleasantness; hence, only one initial image was shown and then adjusted (extended) by observers using the computer-controlled interactive software. We generated many plots to examine the relationship between the initial and reproduction images; this was the same way we analyzed the data for investigating gamut compression. The reproduced images were then calculated by averaging each observer's data for each image. These images were used to represent the mean visual results.

We made our first GEA, called the GEA-1, to model the observer experimental data from experiment GE1 and derived its coefficients for lightness and chroma extension using the least square technique. Using five variations of the GEA-1 (GEA-2 to GEA-6) and increasing the amount of chroma extension, we verified that i) the amount of extension that the observers applied was indeed sufficient and ii) the observers were too cautious in their modifications in experiment GE1. In experiment GE2, using a pair comparison method, we evaluated the GEA-1 and its five variations as well as the average observer images obtained in experiment GE1.

II. Development of an Interactive Tool

The interactive tool we developed modifies the color appearance of pictorial images displayed on a monitor by altering the colors of their pixels according to the region of color space to which they belonged. Note that in this case the color space used was CIE $L^*a^*b^*$; however, other color spaces could easily be used instead. The tool consists of two principal parts: one for selecting a particular region of color space, designated the color region selector (CRS), and another for modifying the color appearance of pixels from the selected color region via the lightness, chroma, and hue angle controls, designated the color appearance adjuster (CAA). Figure 1(a) illustrates the overall interface of the experimental display (showing the original and user-adjusted reproduction images) and Fig. 1(b) shows the CRS and CAA tools in detail.

At the beginning of each experiment, we obtain a reproduction image using a gamut mapping algorithm and refer to this image as the initial image. Using our interactive tool, observers first inspect the reproduction image to determine which part of it has the most unsatisfactory match. They then select the corresponding colors in CIE $L^*a^*b^*$ using the CRS controls and then adjust all image pixels from that region using the CAA controls. For example, an observer sees that the high key neutral colors in an image are too dark and have a yellow tint. He or she needs to first select the high lightness and low chroma region in CIE $L^*a^*b^*$ and adjust it to

be lighter and bluer, or less colorful.

The CRS tool includes three selectors: lightness, chroma, and hue (Fig. 1). Lightness and chroma are each divided into three color regions: high, medium, and low ranges, i.e., L^* of 0–35, 35–65, and 65–100, and C^* of 0–25, 25–50, and 50 to the maximum achievable chroma, respectively. Hues, on the other hand, are divided into six different sectors: red, yellow, green, cyan, blue, and magenta, with ranges of 356.5–59.0, 59.0–128.0, 128.0–180.5, 180.5–226.0, 226.0–291.0, and 291.0–356.5 degrees, respectively, whereby this division is based on the hue angles of the unique hues [6].

The CAA controls allow subjects to modify colors by increasing or decreasing a particular color by steps of two units for lightness and chroma, and one unit for hue. Observers can select and modify colors using one attribute, e.g., first select a high chroma region and increase or decrease its chroma, or more than one attribute, e.g., select a low chroma and high lightness region and increase or decrease one of the three attributes, lightness, chroma, or hue.

The procedure for applying the changes indicated by observers can be written using the following notation, which shows lightness changes as an example:

$$L^{*\prime} = f(L^*_0, P_{L^*Low}, P_{L^*Medium}, P_{L^*High})$$
$$L^{*\prime\prime} = g(L^{*\prime}, P_{C^*Low}, P_{C^*Medium}, P_{C^*High})$$
$$L^{*\prime\prime\prime} = h(L^{*\prime\prime}, P_{hRed}, P_{hYellow}, P_{hGreen}, P_{hCyan}, P_{hBlue}, P_{hMagenta}).$$

Here L^*_0 is the lightness from the initial gamut-mapped image, P_{ij} is the observer-selected lightness adjustment parameter for interval j of color attribute i. Analogous equations apply to chroma and hue angles as well. These transformations are always applied to the initial gamut-mapped image and the transformation coefficients (e.g., P_{ij}) are refined by the observers. The nature of the iterative adjustment is shown in the following example: if an observer changes the lightness of the low lightness region (P_{L^*Low}) by +1, +2, +10, and –2, then the resulting change applied to the initial gamut-mapped image is a simple sum of these, i.e., 11.

The resulting change not only affects the desired color region but also has a slight effect on colors in neighboring regions by linear interpolation. This is done intentionally to avoid discontinuities between color regions defined by CAA controls. Once the operating procedure for modifying a color region is completed, observers press the *PROCESS* button. This leads to the initial gamut-mapped image being processed and displayed to replace the reproduction image. If the colors selected are out-of-gamut, they are automatically clipped onto the printer gamut boundary with the same gamut mapping algorithm that was used for the initial image. The control menu, which includes SNAP SHOT, UNDO, and INITIALIZATION, allows observers to restore previous versions of the transformed images so that they can undo any mistakes previously made.

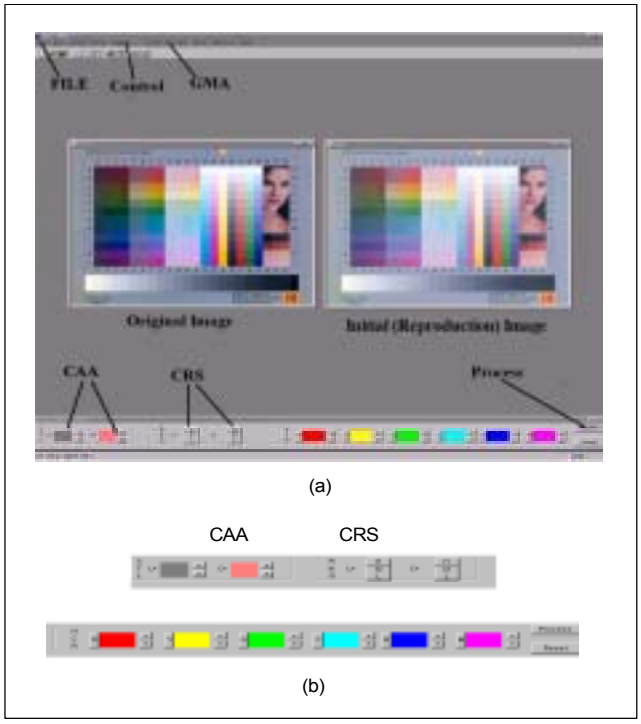

Fig. 1. Experimental display: (a) tool, (b) CRS controls and CAA controls.

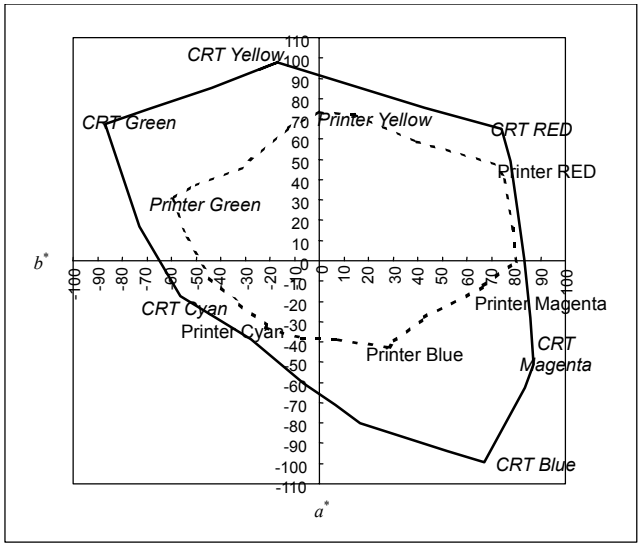

Fig. 2. Color gamut difference between original (CRT) and reproduction (printer) devices.

For extension experiments, at the beginning of each experiment, we obtain the reproduction image using a gamut clipping algorithm. The resulting image is referred to as the initial image. Using the interactive tool, observers first determine which part of the initial image appears to be the most

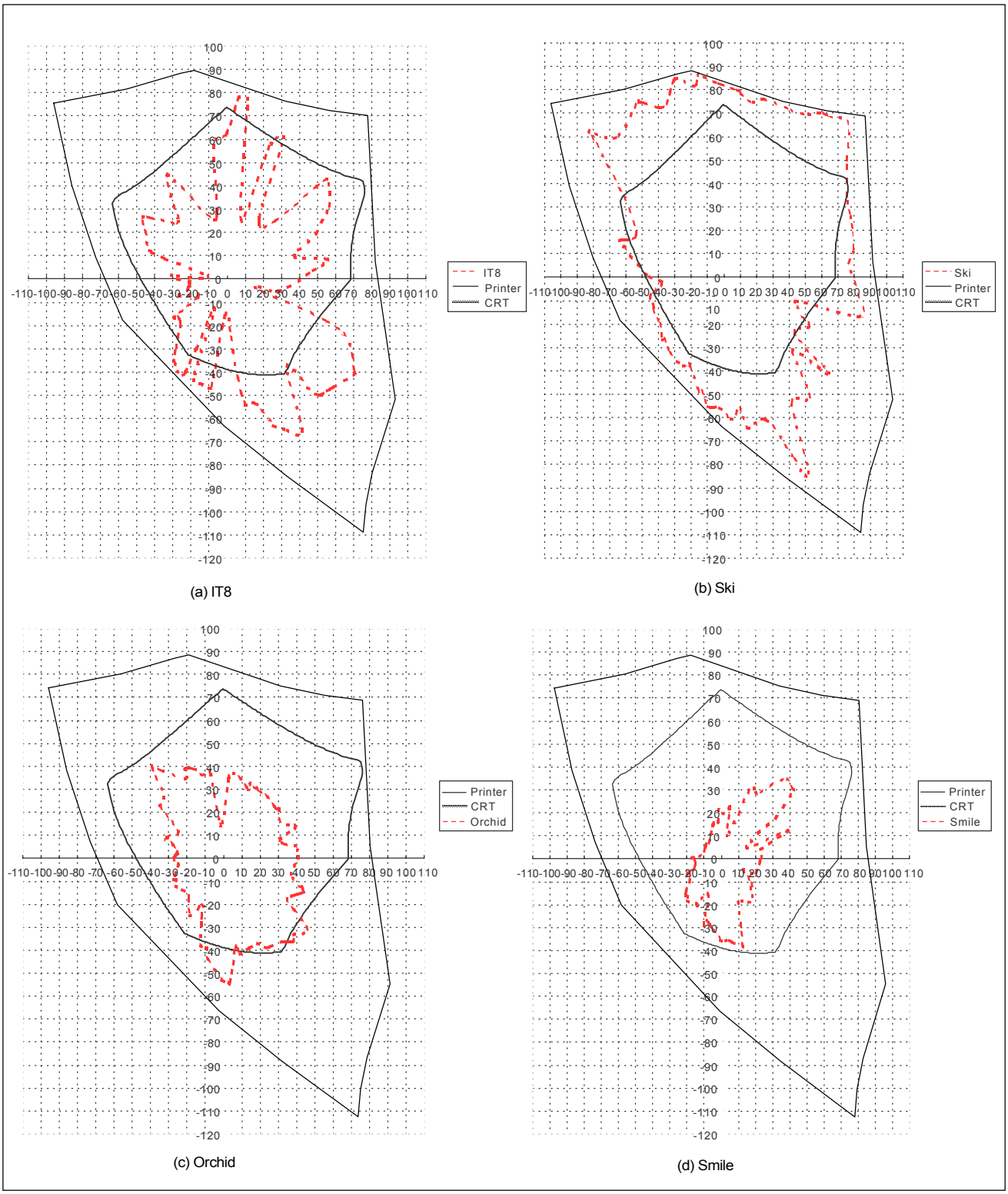

Fig. 3. Image gamut boundaries on CIE a^*-b^*.

unpleasant. They then select their color region in CIE $L^*a^*b^*$ with the CRS controls and adjust these colors using the CAA controls. Unlike in the compression experiment, in this experiment, just one reproduction image is shown to the observers at one time. They can extend the initial image colors towards the monitor gamut boundary. If an observer

extends a color outside the medium gamut boundary, the tool automatically clips it onto the surface of the gamut boundary.

III. Experimental

1. Test Images

In this study, we used four test images: IT8, Ski, Orchid, and Smile. The original and reproduction images were arranged side by side on a CRT display (Fig. 1). Figure 2 shows their very different image gamuts.

Figure 3 depicts the image gamut boundaries for each image, and Table 1 describes the statistical information about the image gamut volume. To develop a gamut compression algorithm capable of giving good results for a wide range of original images, the test images chosen should cover a variety of possibilities in the color space. The four test images we chose satisfy this requirement (Fig. 3). The same four images were used to obtain experimental data in all the following test experiments, though the observer groups were different in each experiment. In future studies, more images should be used to further test our gamut compression algorithms.

These image statistics give useful information for analyzing experimental data. In Table 1, Mean Chroma is the average chroma of the image pixels. Out of Printer Gamut indicates that the number of image pixels is beyond the gamut boundary of the printer. Relative Image Chroma Range to CRT (or Printer) Gamut compares the chroma value of the image pixel to that of the corresponding media (CRT or printer) gamut. The average values for the whole image are represented in columns (3) and (4) of Table 1. Column (5) of Table 1 represents the percentage of pixels for each image that are out of the CRT gamut. Figure 2 shows the gamut boundaries for the original (Sony Trinitron CPD) and reproduction (IBM Lexmark InkJet 4097) gamuts. Our analysis showed that the gamuts of the two devices are quite different, not only in volume but also in shape. The CRT gamut was much larger than that of the printer, except for the cyan region around L^*=50.0.

2. Device Characterization

An LMT C1210 colorimeter was used for measuring the CRT, which was modeled using the Gain-Offset-Gamma model developed by Berns et al. [7], in which the white point was set to D65. We then compared the measured results and the results predicted from the model based upon 729 testing colors that covered the entire color gamut. The results showed an average color difference of 0.6 with a maximum of 3.7 $\Delta E_{CMC(1:1)}$ units. An Xrite 938 spectrophotometer measured the tristimulus values to characterize the printer. A third-order masking equation was derived between CMY and XYZ values of the printer. Another test using 729 colors showed a mean prediction error of 2.6 with a maximum of 7.1 $\Delta E_{CMC(1:1)}$ units under D65 and the CIE 1931 standard colorimetric observer. We consider this a reasonable degree of accuracy, especially as the printer characterization model was only used for calculating the simulated reproduction gamut boundary. The printer device characterization in this study focused on the masking model.

3. Experiments

We made a small number of compression experiments in a darkened room to test our interactive tool; eleven observers, who were ETRI staff, received training prior to the real experiment, and they all passed the Ishhara Color Deficiency Test [8]. In the real experiment, each observer made the adjustments twice for each image. These results were used to investigate observer repeatability. On average, each image took about 30 to 60 minutes to complete. For each image, there was an initial image, which was gamut compressed using the LCLIP gamut mapping algorithm. Table 3 shows the experimental viewing conditions.

In the first experiment, 11 observers used the tool to adjust four different test images twice. Finally, four averaged images were obtained through 22 experiments averaging the device-

Table 1. Image statistics in terms of CIE $L^*a^*b^*$.

	(1) Mean Chroma	(2) Out-of-Printer Gamut	(3) Relative Image Chroma Range to CRT Gamut	(4) Relative Image Gamut Chroma Range to Printer Gamut	(5) Out-of-Monitor Gamut
IT8	22.94	22.57%	37.98%	76.38%	0.87%
Ski	33.35	37.36%	46.07%	108.65%	2.07%
Orchid	23.70	42.11%	50.79%	111.62%	1.38%
Smile	20.67	3.31%	32.07%	50.29%	0 %

Table 2. Information on observers.

	Age (Average, Range)	Duration	Sex (M/F)
First experiment	28 (25–37)	30–60 min/image	11/0
Second experiment	23 (21–33)	30–40 s/image	6/6

Table 3. Experimental viewing conditions.

Screen	Width: 40 cm, Height: 30 cm
Viewing Distance	50–55 cm
Angular Subtense	15°–20° (for observing two images simultaneously)
Ambient Lighting	x: 0.31 y: 0.28 Y: 0.03

Table 4. Observer accuracy and repeatability.
(the first experiment) $\Delta E_{CMC(1:1)}$

	Observer Accuracy	Observer Repeatability
Minimum	1.43	1.35
Median	2.83	3.86
Mean	2.94	3.88
Maximum	6.55	7.09
Standard Deviation	0.99	1.43

independent space (CIE $L^*a^*b^*$). We averaged these four images to obtain one image representative of all the experimental results, the average observer image, and used it in the further algorithm development. To analyze observer behavior, observer accuracy and repeatability were calculated (Table 4). Observer accuracy was measured by the mean color difference between the individual images and the average observer image. Repeatability was measured by the mean color difference between two experiments carried out by one observer.

IV. Development of Compression Algorithms

We developed three new algorithms using the data from our experiments. We devised the GCA-1 by modifying the CLLIN, so we could apply non-linear chroma compression.

By observing the L^*-C^* plots to find the relationship between the original and observer data, three systematic patterns in terms of their converging vectors towards the L^* axis emerged. The GCA-2 algorithm consists of three steps: lightness mapping, determining the convergent point, and finally non-linear chroma compression. From the significant shortcomings we found when testing the GCA-2, we arrived at the GCA-3, in which we modified the methods for determining the convergent point and chroma compression to fit the experimental results more accurately. We then evaluated the three new GCAs together with some existing GCAs and the average observer images from the observer data acquisition experiment, the first experiment in section III. The GCA-3 generally outperformed the others and was even better than the average observer images in some images and color regions. Among the existing GCAs, the GCUSP outperformed the other five.

The GCA-3 algorithm was developed based upon the device gamuts of the original and reproduction images used in the data acquisition experiments in this study, and thus had gamut dependent parameters. We then developed a generalized model, the generally usable gamut compression algorithm (GUGCA), to take into account other possible device conditions.

We conducted a third and final experiment to verify the GUGCA. Twelve observers conducted pair comparisons of the four reproductions used in the first and second experiments. The images used in the experiment consisted of reproductions made with the GCUSP, GCA-1, GCA-2, GCA-3, GCA-3-h, and GUGCA, and the average observer images. The GCA-3-h is an extension of the GCA-3, which included a hue-shifting module, modeling the changes of hue seen in the experimental results. The results showed that the GUGCA gave a similar performance to the other five. The hue-shifting reproduction did not give better results than the non-hue-shifting reproduction. In summary, we used observer experimental data to develop GCAs and then applied various statistical or optimization techniques, which are described in the following sections.

1. Algorithm 1 (GCA-1) Development

Our first attempt at developing an algorithm was a modification of the CLLIN algorithm [5]. The chroma mapping procedure of the CLLIN is given below:

$$C_r^* = \begin{cases} C_o^* \times \left(\dfrac{C_{cusp,r}^*}{C_{cusp,o}^*} \right), & \text{if } C_{cusp,o}^* \geq C_{cusp,r}^* \\ C_o^*, & \text{if } C_{cusp,o}^* < C_{cusp,r}^* \end{cases}, \qquad (1)$$

where here and hereafter subscripts *o* and *r* denote the original and reproduction gamuts, respectively. Comparing the chroma of the average observer images and data transformed via the CLLIN algorithm revealed that the resulting relationship was not linear, while lightness compression followed a linear function. Hence the core for deriving the GCA-1 was to include a non-linearity chroma compression as given in (2).

$$C_r^* = \begin{cases} (1-\alpha)\times C_o^* + \alpha \times [C_{cusp,r}^* - (C_{cusp,o}^* - C_o^*)\times \left(\frac{C_{cusp,r}^*}{C_{cusp,o}^*}\right), \\ \qquad\qquad \text{for } C_{cusp,o}^* \geq C_{cusp,r}^* \\ C_o^*, \qquad \text{otherwise} \end{cases} \tag{2}$$

with $\alpha = 1.0-(1-\alpha_{min})\times e^m$,
where

e: Euler's number ($\ln^{-1}1$),

α_{min}= 1.84 × (C_α/ C^* of original medium cusp), and

$$m = -\left(\frac{L^* - L_\alpha^*}{\delta}\right)^2 - \left(\frac{C^* - C_\alpha^*}{\delta}\right)^2,$$

where

L_α= 0.797 × L^* of reproduction medium cusp,

C_α= 0.563 × C^* of reproduction medium cusp, and

δ : 22 (arbitrary disperse parameter).

After the chroma compression, the lightness was linearly mapped along the lines of a constant chroma using (3),

$$L_r^* = L_{r2}^* + (L_o^* - L_{o2}^*) \times (L_{r1}^* - L_{r2}^*)/(L_{o1}^* - L_{o2}^*), \tag{3}$$

where the subscripts *o*1 and *o*2 denote the maximum and minimum, respectively, of the original gamut, and *r*1 and *r*2 denote those of the reproduction. Derivation of the chroma compression formulae depends on the assumption that chroma compression follows a non-linear way relative to the gamut boundaries and this non-linear function follows the normal distribution determined by three different parameters, i.e., chroma, lightness, and α. Formulae for determining L_α, C_α, and α_{min} were derived by observing the relationship between α and the cusp's lightness (or chroma) in the qualitative plots. The primary shortcoming of this algorithm is that the procedure used for deriving it was too heuristic and complex.

2. Algorithm 2 (GCA-2) Development

Further observing the L^*-C^* plots between the original and observer data revealed that compression patterns can be grouped into three in terms of their converging vectors towards the L^* axis. Figure 4 shows the modeling of this phenomenon. The GCA-2 was developed according to Fig. 4. Its procedure is described below.

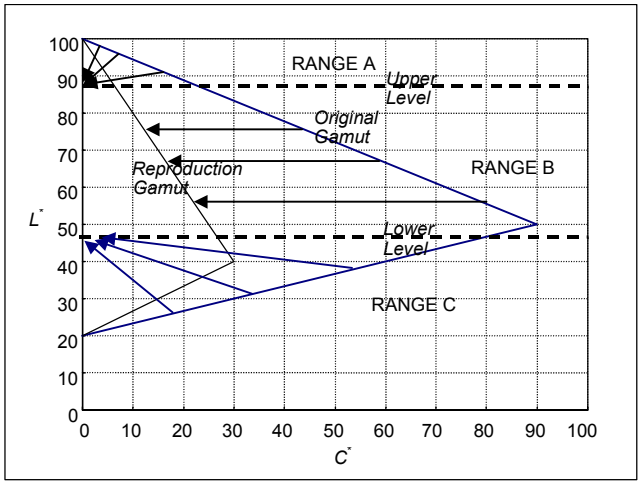

Fig. 4. Three different patterns and two convergent points.

[STEP 1] Lightness mapping

We derived the lightness mapping function by modeling the relationship between the lightness of the original and observer data. This takes into account the media used in the experiment by having as parameters the minimum L^* point, which is the point where the slope is changed, and the maximum L^* point. A piecewise linear interpolation function connecting these three points was then derived to model the relationship as follows.

If $L^*_{min,o} \leq L^*_o < L^*_{min,o} + ((16.9/A)\,(L^*_{max,o} - L^*_{min,o}))$,

$$L_r^* = \left[L_{min,r}^* + \frac{1.8}{A}\left(L_{max,r}^* - L_{min,r}^*\right)\right] - \frac{10.65}{A}\left[\left(L_{min,o}^* - L_o^*\right) + \frac{16.9}{A}\left(L_{max,o}^* - L_{min,o}^*\right)\right]\left(\frac{\left(L_{max,r}^* - L_{min,r}^*\right)}{\left(L_{max,o}^* - L_{min,o}^*\right)}\right)$$

else if $(L^*_{min,o} + (16.9/A)\,(L^*_{max,o} - L^*_{min,o})) \leq L^*_o \leq L^*_{max,o}$

$$L_r^* = L_{max,r}^* - 1.18\left(L_{max,o}^* - L_o^*\right)\frac{\left(L_{max,r}^* - L_{min,r}^*\right)}{\left(L_{max,o}^* - L_{min,o}^*\right)}. \tag{4}$$

[STEP 2] Determining the convergent point

Investigating 24 L^*-C^* plots (6 hues × 4 test images) revealed three patterns in terms of their converging vectors on the L^* axis. Two convergent lightness points were determined by the L^* of the original medium's cusp. The two convergent points are named upper and lower level points, which are calculated as follows.

□ Lower convergent point

If the original medium cusp's $L^* \leqq 30$

then the lower convergent point = 31.54 (L^*).

Otherwise,

the lower convergent point

= 19.42 + 0.404 × *the original cusp's* L^*.

□ Upper level convergent point

If the original medium cusp's $L^* \leqq 49$

then the upper convergent point = 70.12 (L^*).

Otherwise,

the upper convergent point

= 49.297 + 0.425 × *the original cusp's* L^*.

[STEP 3] Chroma Compression

Comparing the chroma as predicted by the LLIN [5] with the chroma of the observer data, we found that the maximum ratio (chroma of the observer data/chroma of the linear compression method) was C^*=11.93 (of the linear chroma compression method), which is 17% of the original cusp's chroma and about 1.90 times greater than that of the LLIN. The chroma compression of the GCA-2 follows this non-linear way as shown in Fig. 4. The convergent point is determined by whether the lightness of the color is higher than the upper level convergent point or lower than the lower level convergent point. Between the two convergent points, the color is mapped along the lines of constant lightness.

In summary, the GCA-2 consists of these three steps: i) piecewise linear lightness mapping, ii) the determination of convergent points, and iii) lightness-chroma compression along either constant lightness lines or lines towards convergent points. The GCA-2 outperforms the GCA-1 in accurately modeling observer experimental data. However, the GCA-2 does have some shortcomings. Unwanted artifacts were found particularly in the darker blue areas in the Ski image. If a lower convergent point is determined at a higher level by the algorithm, the mapped lightness of the reproduction might be unnecessarily high. This indicates that the lower level convergent point needs to be further improved. In addition, the chroma compression function does not fit precisely to the experimental data, especially for the higher chromatic region.

3. Algorithm 3 (GCA-3) Development

The procedure for the GCA-3 was identical to that of the GCA-2 except for modifications to steps 2 and 3.

[STEP 2]

To circumvent the shortcomings in the GCA-2, the distance between the lower and higher level convergent points was extended, thus making the higher convergent point higher and the lower point lower. A new set of equations was defined:

□ Lower convergent point

If the original medium cusp's $L^* \leqq 30$

then the lower convergent point is 28 (L^*).

Otherwise,

the lower convergent point

= 45 × (*the original cusp's* $L^*/100)^2$

–2.8 (*the original cusp's* $L^*/100$) + 24.8.

□ Higher convergent point

If the original medium cusp's $L^* \leqq 30$

then the lower convergent point= 75 (L^*).

Otherwise,

the lower convergent point

= 67 × (*the original cusp's* $L^*/100)^2$

– 51.97 (*the original cusp's* $L^*/100$) + 84.8.

[STEP 3]

For chroma compression, a new equation was derived:

$$C_{compressed} = -1 \times \left(\frac{7.8}{B}\right) \times (C^*)^2 + \left(\frac{122.78}{B}\right) \times (C^*) + 3.409, \quad (5)$$

where B=1,000 (constant).

We derived the parameters of the equations, e.g., lightness mapping and convergent points, by linear regression and optimization techniques. For example, the optimization procedures for determining the convergent points are as follows: i) Combine the original and reproduction data of the test images, ii) divide the data by hue areas and then plot 72 (360°/10×2) L^*–C^* diagrams, iii) divide the data in the same hue plane into 3 different L^* areas, iv) select the initial two convergent points from the function and calculate the difference between points from the function and observer data, and v) determine the minimum difference by adjusting the parameters of the function for determining the convergent point .

4. The Second Experiment

Twelve observers took part in the second experiment using the pair comparison method and the four images used in the first experiment. These images were mapped through the three newly-derived algorithms and six existing algorithms (the LCLIP, LCUSP, GCUSP, LLIN, CLLIN, and SLIN) with an image based on the average observers' results in the first experiment. In the experiment, an original image was shown to an observer; after 20 to 30 seconds, two images from different compression algorithms were given to the observer and he or she was asked to select which of the two was closer to the original. If an observer could not make a decision immediately, then the original was shown again. Based on Thurstone's law of comparative judgment [9], the overall accuracy results are shown in Fig. 5 in terms of the z-score. The newly-derived algorithms and the average observer image significantly outperform the existing algorithms. Among the new algorithms, the GCA-3 was the best and even better than the average observer image. For the existing algorithms, the GCUSP outperformed the other five. The results of the overall case were very similar to those of individual images.

5. Development of Generally Usable Algorithm

From the second experiment, we determined that the GCA-3 performed the best. However, this algorithm was developed according to the media gamuts of the original and reproduction images used in the first experiment. Hence, the GCA-3 might not give consistently good color matching results for other reproduction media. To overcome this problem, algorithm generalization was performed as follows.

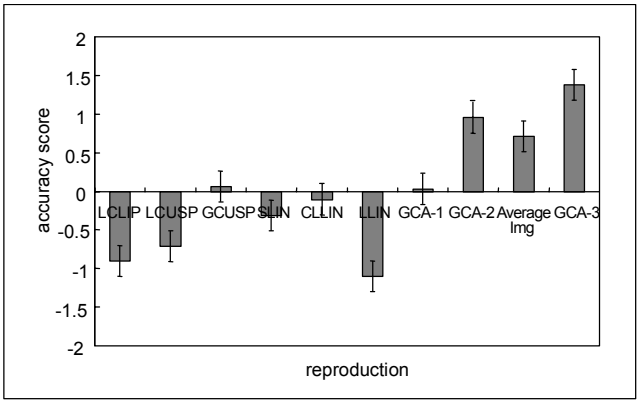

Fig. 5. Overall scores averaged for four images for the second experiment (Error bars represent a 95% confidence interval.)

[STEP 1] Lightness Compression

This step is the same as step 1 of GCA-2.

[STEP 2] Convergent Point Determination

□ Upper level

If $L^* < K_0$

then the upper convergent point is $f(K_0)$.

Otherwise,

$$\textit{the upper point} = f(x) = 45 \times \left(\frac{(x - K_0)}{100} \right)^2 + K_1 .$$

□ Lower level

If $L^* < K_1$

then the upper convergent point is $g(K_1)$.

Otherwise,

$$g(x) = 45 \times \left(\frac{x}{100} \right)^2 - 2.8 \times \left(\frac{x}{100} \right) + K_2$$

with the terms defined as follows:

x: original gamut cusp's lightness of a color,

K_0: lowest lightness values in the cusps' lightness of six primary and secondary reproduction gamuts,

K_1: mean of lightness values of six primary and secondary reproduction gamuts × 1.40, and

K_2: lowest lightness of reproduction gamut.

This step was obtained by optimizing an individual parameter and then fitting it on the experimental data.

[STEP 3] Recompress lightness and chroma simultaneously

$$d_3 = 1.12 \times \left(\frac{d_1 \times d_2}{d_0} \right), \tag{6}$$

with the terms defined as follows:

d_0: distance between the convergent point and the original media's gamut boundary,

d_1: distance between the convergent point and the original color,

d_2: distance between the convergent point and the reproduction media's gamut boundary, and

d_3: distance between the convergent point and the compressed point.

When joining the line between a color (d_1) and a convergent point from Step 2, d_0 and d_2 can be defined using a one-dimensional equation. A coefficient of 1.12 was obtained by fitting the experimental data.

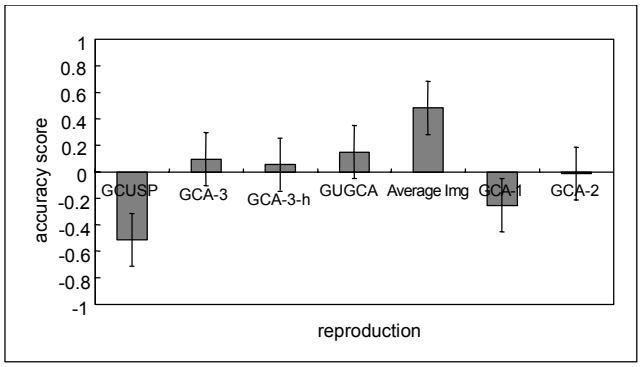

Fig. 6. Overall scores averaged for four images for the third experiment. (Error bars represent 95% confidence interval.)

6. The Third Experiment

In the third experiment, twelve observers conducted psychophysical experiments using pair comparisons to verify the performance of the generalized algorithm, the generally usable gamut compression algorithm (GUGCA). The four images used in the second and third experiments were again used. Each image was processed using six algorithms—the GCUSP, GCA-1, GCA-2, GCA-3, GCA-3-h, and GUGCA—as well as the average observer image. GCA-3-h is the GCA-3 extended by modeling the hue shifting behavior from the first experiment. Figure 6 shows the overall z-score (accuracy) results based on Thurstone's law of comparative judgment [9]. The GCA-3, GCA-3-h, GCA-2, GUGCA, and the average observer images had very similar accuracy scores taking into account the confidence level. The hue-shifting algorithms did not give better results than the non-hue-shifting algorithms. It is

encouraging that the GUGCA performed even slightly better than the GCA-3 in some test images and color region experimental data.

7. Summary

Twelve observers took part in the second experiment to evaluate the three newly-developed GCAs. The images used in the experiment were transformed by the three newly-derived algorithms, the six existing algorithms—the LCLIP, LCUSP, GCUSP, SLIN, CLLIN, and LLIN—and the averaged observer images obtained in the first experiment.

The accuracy scores (z-scores) of the results showed that the newly-derived algorithms together with the averaged observer image generally outperformed the existing algorithms. The GCA-3 performed the best amongst the three new GCAs studied and the averaged observer data. It also gave similar color matching performance across all the test images and color regions. This indicates that the algorithm is more or less image-independent. Hence, the GCA-3 was used as a basis for developing the GUGCA. Among the previous algorithms, the GCUSP performed the best.

The GCA-3 performed the best in the second experiment. However, this algorithm was developed based upon the media gamuts of the original and reproduction devices used in the data acquisition experiment of this study. Hence, it could not be expected to take into account a wide range of the original and reproduction media. To overcome this problem, a generic GCA was developed and was named GUGCA. Three different procedures of the GCA-3 were generalized, i.e., not dependent on any pair of media gamuts.

In the third experiment, twelve observers conducted pair comparison assessments to evaluate the GUGCA together with the other algorithms using the four images used in the two previous experiments. The results showed that the GUGCA, the observer-averaged image, and GCA-3, which was the best algorithm in experiment GC2, generally gave similar matching performances in terms of z-scores. The color matching performance of the GUGCA was generally stable across test images and color regions in comparison to the other algorithms. The results showed that the GUGCA could be used generally (any image and any imaging media) in color image reproduction as a gamut compression algorithm.

V. Extension Algorithm Development

1. Procedure of Experiment GE1

Twelve observers took part in this experiment. Before the main experiment, they were trained for about one hour to verify the experimental procedure and their understanding of lightness, chroma, and hue attributes. On average, each image took between 25 to 50 minutes to complete the adjustment. Four images, IT8, Ski, Orchid, and Smile, were used. This experiment was conducted twice for each image. The results for observer accuracy and repeatability in experiment GE1 are given in Table 5.

Table 5. Observer accuracy and repeatability (experiment GE1) in $\Delta E_{CMC(1:1)}$ Unit.

	Observer Accuracy	Observer Repeatability
Minimum	0.90	0.04
Median	4.85	5.14
Mean	4.96	5.39
Maximum	17.45	12.03
Standard Deviation	2.55	2.37

For accuracy and repeatability, the variances of the extension experiments were larger than those of the compression experiments. Two images, the original and reproduction, were shown in compression experiments, though only one initial image was shown in the extension experiments, and then observers adjusted the initial image according to the degree of pleasantness as judged by the observers. The larger variance was caused by this difference between two experimental methods.

2. Analysis of Experimental Results

Seventy-six plots were generated to study the relationship between the initial and observer reproduction images. The reproduction images were calculated by averaging each observer's data for each image. As in the compression experimental data, we analyzed the figures of lightness and chroma between the initial and reproduction (extended) images, hue variations plotted between the hue angles and hue differences (ΔH^*), and data on the L^*-C^* planes. In this study, some typical examples are presented to describe the analysis and development of the algorithm.

A systematic pattern for lightness extension was found for all four of the images. The results from these were combined and plotted in Fig. 7, which also shows a minimum L^* value of about 25. This is caused by the limitation of the initial images generated within the reproduction (printer) gamut boundary having lightness ranging from 24.67 to 100.0. The solid line indicates a perfect agreement between the initial and

reproduction results if all data points lie on the line. The results clearly show that the reproduction images are darker and slightly higher in contrast than the initial images.

In a manner similar to the L^* plots, the C^* values between the initial and reproduction colors were also plotted for each image. We also analyzed the hue shift. First, the ΔH^* (=2 $(C^*_o C^*_r)^{1/2}$ × sin $(h_o – h_r/2.0)$) values were calculated and plotted against the hue angles for each image. The trend showed that colors were adjusted mainly in three hue areas—red to be bluer, blue to be greener, and magenta to be bluer—but almost no change for the other colors, such as yellow, green and cyan. However, the magnitudes of adjustment were very small, between +2 and –2 in ΔH^* units. This implies that the hue shift was not essential for developing GEAs in this experimental domain.

Figure 8 shows the L^*-C^* plots for the color shift before and after the adjustment of the primary and secondary hues for the Ski image. In each plot, the initial and reproduction colors are shown using circle and cross symbols (a, β in Fig. 8), respectively, and the dashed line represents the image gamut boundary (χ in Fig. 8). Hue areas were divided according to the primary and secondary colors of the CRT gamuts. The patterns of color shifts for the Ski image in each of the primary and secondary hues were very similar to those in the IT8 and Orchid images, but not the Smile image. All the colors in the Smile image were adjusted in the lightness direction to become darker, but not in the chroma direction. This indicates that the image content has an impact on gamut extension. The Smile image includes the face of a girl and has different image characteristics than the other images. All the plots were used to develop the new GEA. Most people have a notional skin color and it seems observers produced a stimulus of this color in the experiment. 'Memory color' refers to this phenomenon that recognizable objects often have a prototypical color that is associated with them [10].

To investigate hue-dependency in this extension experiment, we calculated three different values: the lightness of the cusp of the original medium ($L^*_{(cusp,\ medium)}$ in Table 6), lightness of the cusp of the image gamut ($L^*_{(cusp,\ image)}$), and the slope of the extension shift ($(L^*_r – L^*_i)/(C^*_r – C^*_i)$) (Σ in Table 6; the subscript i denotes the initial image). This analysis investigated how much information of the gamut boundary is statistically related to the observer data of the experiment GE1 data. A higher correlation means that the lightness-chroma extension function should be a hue variant. Table 6 shows correlation results.

Table 6 shows the correlation results of individual test images and the case of 'all images' that represents correlation between all observer data against the gamut information of four test images. As the table shows, the poor correlation results we obtained demonstrated that gamut extension is not related to gamut shape: hue-invariant in this experimental domain (test images and medium gamut). This verifies the hue-invariant chroma extension function of GEA-1.

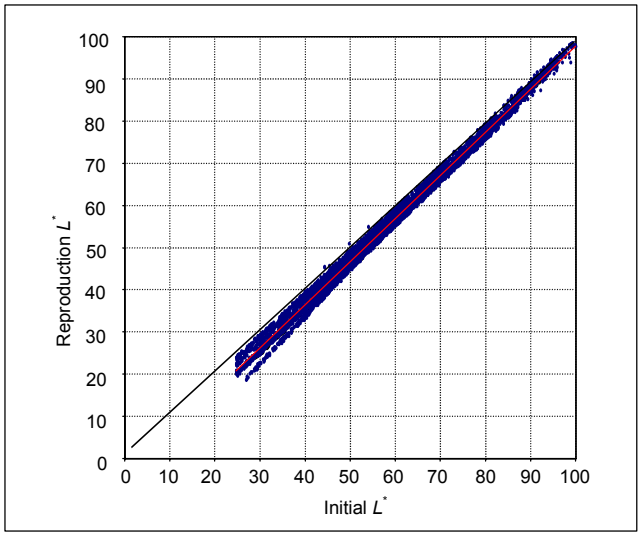

Fig. 7. Initial L^* vs. reproduction L^*, combination of four images.

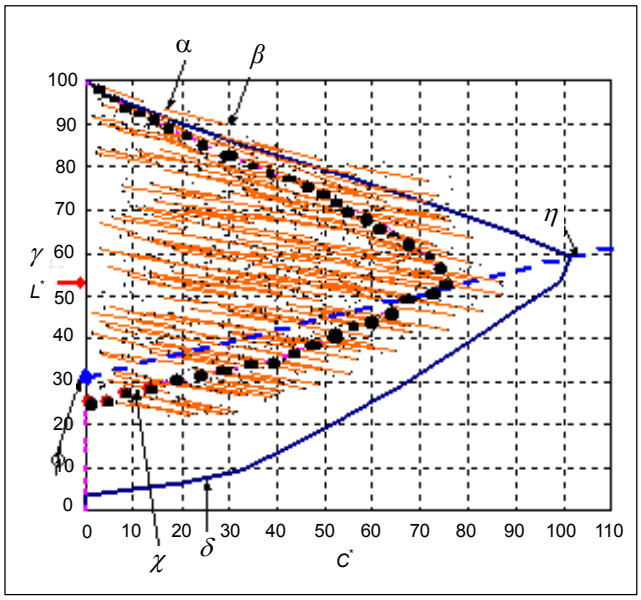

Fig. 8. Results of experiment GE1, analysis of L^*-C^*, Ski, red and yellow (α: initial color; β: reproduction color; χ: image gamut boundary; δ: original medium gamut boundary; φ: intercept point of η; γ: lightness of cusp of reproduction medium; η: connecting cusps between original and reproduction media).

As Fig. 7 shows, all initial images were adjusted to become darker and to extend their lightness ranges. A linear function was devised to fit this trend. The results are given in Table 7 for each individual image and the combined data. In addition, the modeling accuracy is reported in terms of the correlation of

determination (R^2). The results clearly show that the model developed for each image was very similar and each model fitted the experimental data well. The lightness extension equation, L_r=1.02L_o – 4.44, is derived by experimental data.

The analysis of chroma data shows that there is a distinct linear gamut extension for three of the test images, i.e., the majority of colors are linearly located above the 45° line in the IT8, Ski, and Orchid images. However, in the Smile image, the chroma values of the initial and reproduction images are divided into two groups: one group lies on the 45° line and the other is located above the 45° line. The group around the 45° line includes red, yellow, and green in the Smile image. This is a typical example illustrating that the image content has a large impact on observers when performing the gamut extension. The lightness dependent chroma extension suggested by Hoshino [11], [12] was not observed in the data from experiment GE1.

For chroma extension modeling, the GEA-1 can be used as a basis for further model evaluation. The Smile image was not included in modeling the GEA-1 due to its anomalous extension pattern, particularly in the lower range. Prediction accuracy and the linearity of chroma extension in the GEA-1 are again acceptable from the statistical analysis point of view (R^2=88%). A coefficient was obtained using the least square fitting technique of statistics analysis software (SPSS version 6.0).

Based upon the analysis above, we developed the gamut extension algorithm GEA-1 as follows.

[STEP 1]

$$L_r = 1.02L_o - 4.44. \quad (7)$$

[STEP 2] A Linear Chroma Mapping

$$C_r = \theta\ C_o, \quad (8)$$

where θ =1.33, which was obtained using the least square fitting technique of statistics analysis software.

As mentioned, the other models (GEA-2 to GEA-6) were devised i) to test the GEA-1 and ii) to encompass the possible perception of pleasantness in higher extension rates. Experiment GE1 data shows that the observers did not fully adjust chroma onto the CRT gamut boundary. Particularly, developing the group-2 models (Table 8) was an attempt to cover the different observer behavior in the lower chroma levels in the Smile test image. This behavior was analyzed as the skin tone colors considered as memory color. This was also evaluated in experiment GE2.

The group-1 model was devised from the averaged data of three test images, IT8, Ski, and Orchid, having the chroma extension pattern from the origin (neutral point). The group-2 models were developed using all four averaged data sets. This model divides chroma extension into two parts separated at a C^* value of 23.0 (of chroma) obtained qualitatively. It assigns C^*_r to be the same as C^*_o for C^*_o less than 23, otherwise C^*_r =$\theta\ C^*_o$, where θ is the slope obtained from various plots. Each group of models is divided into two or three in terms of the degree of gamut extension determined by the slope of C^*_r/C^*_o and are given in Table 8. The degree of chroma extension is gradually increased from GEA-2 to GEA-3.

Table 6. Correlation results between the gamut information and observer experimental data.

	Correlation between $L^*_{(cusp,\ medium)}$ and Σ	Correlation between $L^*_{(cusp,\ image)}$ and Σ
IT8	0.46	0.60
Ski	–0.29	–0.16
Orchid	–0.18	0.17
Smile	–0.43	–0.91
All Images	–0.10	–0.39

Table 7. Regression Results of L^*–Gamut Extension.

Image	Model	Accuracy of fitting (R^2)
IT8	L_r=1.04L_o – 5.93	99.99%
Ski	L_r =0.99L_o – 1.38	99.81%
Orchid	L_r =1.03L_o – 5.34	99.99%
Smile	L_r =1.06L_o – 8.14	99.90%
All	L_r =1.02L_o – 4.44	99.48%

Table 8. Five different variations of chroma extension.

Group	GEA	Starting point	Image data used to fit each model	Sub-group	θ	R^2
1	2	(0, 0)	IT8, Ski, Orchid	2+3	1.57	0.87
	3	(0, 0)	IT8, Ski, Orchid	3	1.93	0.85
2	4	(23, 23)	IT8, Ski, Orchid, Smile	1+2+3	1.48	0.54
	5	(23, 23)	IT8, Ski, Orchid, Smile	2+3	2.05	0.29
	6	(23, 23)	IT8, Ski, Orchid, Smile	3	3.53	0.14

3. Experiment GE2

Experiment GE2 was designed to evaluate the performance of the GEA-1 and its five variations as well as the averaged observer data obtained in experiment GE1. We did this experiment to verify that in the GEA-1, the amounts of extension observers applied were indeed sufficient. Twelve observers (six females and six males) took part in this experiment using the pair comparison technique after having received a training session. In the main experiment, a pair of images was shown to each observer and their task was to select which of the two was more pleasant.

Even though Hoshino suggested a wide range of extension methods in his patent [11], [12], his methods could not be included in the experiments of this study because i) detailed extension procedures were not specified, ii) performance evaluation between the methods was not specified, and iii) the data of the user experiments of this study did not agree with those of his methods.

We divided the group into three sub-groups (1, 2, 3) in terms of the extent of gamut extension, i.e., $a = C^*/C^*_o$.

The higher the z-score, the higher the degree of pleasantness as judged by the observers (Fig. 9). The 95% confidence interval for each z-score was calculated. GEA-1 and GEA-2 generally performed better than group 2 algorithms. As expected, the performance of the averaged observer data for each image was also quite satisfactory. This implies that the overall experimental design is reliable. A closer inspection of the z-score results reveals that the GEA-1 performed the best for the IT8 and Smile images, but the second worst for the ORCHID image. This implies that once again the model performance depends upon the characteristics of the image investigated.

During the experiment, each observer was also asked to compare the degree of pleasantness in different color regions of an image. This was normally specified in terms of color objects in each image. The data was also analyzed and the results agree well with individual image data. The difference of z-scores between individual GEAs was not large, although skin colors showed dissimilar results. This is caused by the fact that observers are sensitive to a familiar (memory) color such as skin tone, in carrying out the color extension experiment.

Using the modified computer-controlled interactive software, twelve observers participated in experiments for extending color gamuts to achieve pleasing reproductions. Based on the experimental data, a GEA was developed. After an initial linear lightness extension, five different models were devised according to the degrees of chroma extension.

We did a subsequent experiment to evaluate the newly-developed GEA-1, its five variations, and the averaged observer data. The results show that group 1 algorithms generally performed better than group 2 algorithms in terms of color pleasantness. It was also found that there were larger observer variations in this kind of experiment as there was no reference image for comparison, unlike in the earlier gamut compression experiments.

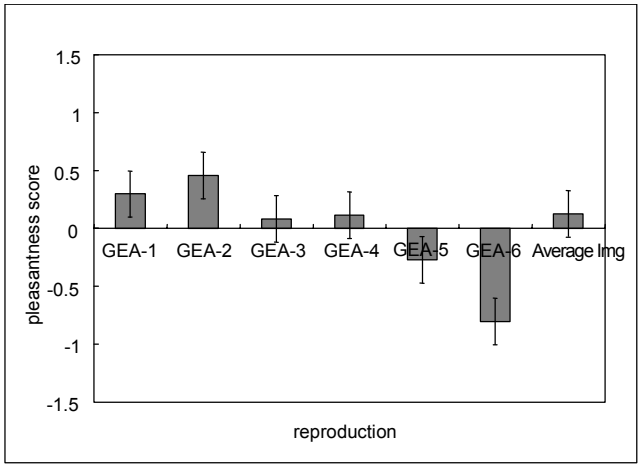

Fig. 9. Overall z-score results of experiment GE2.

VI. Conclusions

Numerous gamut mapping approaches have been proposed and examined in the past, and in most of the studies, algorithms were first defined and then evaluated by observers who judged the algorithms' suitability for a given reproduction intent. An alternative to this approach is to give the observers the opportunity to adjust image colors to better represent the original. We developed an interactive tool for selecting a particular color region of a pictorial image in the CIE $L^*a^*b^*$ color space and for modifying this image to match the original image using color controls based on CIE $L^*a^*b^*$ attributes. The original and reproduction images were limited by the color gamuts of the CRT and printer. The results were used for evaluating and developing GCAs. We represented the experimental results by averaging each individual observer's image pixel by pixel and analyzed them by generating various plots to reveal the color shifts between the averaged and original images. We developed three different algorithms and conducted psychophysical experiments to verify them. We then generalized the best algorithm and further evaluated it. The results showed that the color matching performance of the generalized algorithm was similar to that of the best algorithm.

There has been very little published data on gamut extension. The results of the extension experiments in this study not only give worthwhile data but also suggest a methodology for using the observer experiment tool for further gamut extension research. Characteristics of test images used for the experiments should be taken into consideration in designing the experiments. To develop a gamut compression/extension algorithm, the test images should cover a variety of possibilities

in the color space. We chose four test images because the first experiment took such a long time, an average of 50 minutes per image. To develop a more generally usable algorithm for any medium and any image gamut, more images should be tested in future studies.

Prevalent approaches and research trends in the field of gamut mapping are i) developing a generally applicable algorithm, ii) color imaging manipulation, iii) choice of color space, iv) regional mapping, v) image dependency, and vi) clipping. Clipping, which only changes colors lying outside the reproduction gamut, is the preferred mapping method. Some research results have shown that the minimum ΔE method gives a reasonably good performance for most test images.

We cannot include all methods of previous studies showing the preferred matching performance in the evaluation experiments of this study. First, this study mainly focused on developing an interactive tool and a compression algorithm for using the tool. Second, we needed to analyze the original and reproduction gamuts before we could evaluate the experiments. As the media gamut difference in this study was too large, clipping style methods could not give results as good as those of previous studies. The effect of experimental conditions, e.g., media and images, are of great practical importance in reproducing colors. Future work for enhancing the performance of algorithm development could focus on i) using a different color space, ii) using different algorithms for evaluation, iii) developing a more powerful interactive tool, iv) using image gamuts rather than media gamuts, and v) investigating the influence of image contents.

References

[1] J. Morovic, "Gamut Mapping and ICC Rendering Intents," *Proc. of Int'l Color Management Forum*, 1999, pp. 107-116.

[2] F. Ebner and M.D. Fairchild, "Gamut Mapping from Below: Finding the Minimum Perceptual Distances for Color Outside the Gamut Volume," *Color Research and Application*, 1997, pp. 402-413.

[3] P. Green and R.M. Luo, "Developing the CARISMA Gamut Mapping Model," *Proc. of Colour Image Science*, 2000, pp. 244-256.

[4] J.J. Sara, *The Automated Reproduction of Pictures with Non-Reproducible Colors*, Ph.D. Thesis, Massachusetts Institute of Technology, 1992.

[5] J. Morovic and M.R. Luo, "Gamut Mapping Algorithms Based on Psychophysical Experiment," *Proc. of 5th Color Imaging Conf.*, 1997, pp. 44-49.

[6] R.W.G. Hunt, *The Reproduction of Color in Photography, Printing and Television*, Fountain Press, 1995.

[7] R.S. Berns, "Methods for Characterizing CRT Displays," *Displays*, vol. 16, no. 4, 1996, pp. 173-182.

[8] S. Ishihara, *Ishihara's Tests for Color Blindness, 24 Plates Edition*, Kanehara, Tokyo, Japan, 1984.

[9] L.L. Thurstone, "A Law of Comparative Judgment," *Psychophysical Review*, vol. 34, 1927, pp. 273-286.

[10] M.D. Fairchild, "A Preferred Color Reproduction Method for the HDTV Digital Still Image System," *Color Appearance Models*, Addison Wesley, 1998.

[11] T. Hoshino, "A Preferred Color Reproduction Method for the HDTV Digital Still Image System," *IS&T Proc.: Symp. on Electronic Photography*, 1991, pp. 27-32.

[12] T. Hoshino, *Color Estimation Method for Expanding a Color Image for Reproduction in a Different Color Gamut*, US patent 5,317,426, 1994.

Byoung-Ho Kang is a Principal Research Scientist in the Computer Graphics Research Team of ETRI. He received his MS degree in artificial intelligence from the University of Georgia in the USA in 1993 and PhD in Colour Science from Colour and Imaging Institute of University of Derby (United Kingdom) in 2001. He is a member of CIE Technical Committee 8-03 and editorial board of KSIST (Korean Society for Imaging Science and Technology). His research interests include imaging technologies, color science and computer graphics.

Ján Morovic received a first class BA honours degree in print management from the London College of Printing in 1995. He received his PhD in Colour Science at the Colour & Imaging Institute (CII) of the University of Derby in 1998, where the title of his thesis was "To Develop a Universal Gamut Mapping Algorithm." At present he is working as a Lecturer in digital colour reproduction at the CII and he is also the Chairman of the CIE

Technical Committee 8-03 on Gamut Mapping. His research interests include cross-media colour image reproduction with a particular focus on gamut mapping, medium characterization, and the characteristics of complex images.

M. Ronnier Luo is the Director and a Professor of Colour Science at the Colour & Imaging Institute, University of Derby in the UK. He received his BSc in fibre technology from the National Taiwan Institute of Technology in 1981 and PhD in colour physics from the University of Bradford in 1986. He has over 150 publications in the areas of colour measurement, colour difference, colour appearance, and colour reproduction. He chairs the Colour Measurement Committee (CMC) of the UK and two International Commission on Illumination (CIE) Technical Committees: TC 1-52 Chromatic Adaptation Transforms and CIE TC 8-2 Colour Difference Evaluation in Images. He is also a Fellow of the Society for Imaging Science and Technology (FSIS&I) and Fellow of Society of Dyers and Colourists (FSDC).

Maeng-Sub Cho received the BE degree in statistics from Korea University, Korea, in 1976, the MSc degree in computer science from University of Teesside in the UK in 1989 and the PhD degree in computer science from University of Loughborough in the UK in 1993. In 1976, he joined Korea Institute of Science & Technology (KIST) as a Research Scientist. Since 1998, he has worked as a Principal Researcher at ETRI in Korea. His research interests include imaging science, color image processing, and pattern recognition.