

# Psychophysical Evaluation of Gamut Mapping Techniques Using Simple Rendered Images and Artificial Gamut Boundaries

Ethan D. Montag and Mark D. Fairchild

**Abstract**—Using a paired comparison paradigm, various gamut mapping algorithms were evaluated using simple rendered images and artificial gamut boundaries. The test images consisted of simple rendered spheres floating in front of a gray background. Using CIELAB as our device-independent color space, cut-off values for lightness and chroma, based on the statistics of the images, were chosen to reduce the gamuts for the test images. The gamut mapping algorithms consisted of combinations of clipping and mapping the original gamut in linear piecewise segments. Complete color space compression in RGB and CIELAB was also tested. Each of the colored originals (R,G,B,C,M,Y, and Skin) were mapped separately in lightness and chroma. In addition, each algorithm was implemented with saturation ( $C^*/L^*$ ) allowed to vary or retain the same values as in the original image. Pairs of test images with reduced color gamuts were presented to twenty subjects along with the original image. For each pair the subjects chose the test image that better reproduced the original. Rank orders and interval scales of algorithm performance with confidence limits were then derived. Clipping all out-of-gamut colors was the best method for mapping chroma. For lightness mapping at low lightness levels and high lightness levels particular gamut mapping algorithms consistently produced images chosen as most like the original. The choice of device-independent color space may also influence which gamut mapping algorithms are best.

## I. INTRODUCTION

THE RANGE, or *gamut*, of colors produced by different imaging devices varies greatly from one to another. A major problem in the reproduction of an image from an original device, say a cathode ray tube (CRT), to a destination device, a color printer, is how to adjust the colors that are nonreproducible in the destination device in order to create a veridical reproduction. *Gamut mapping* is the term used to describe this process.

Gamut mapping has been described as being perhaps a “printer’s art” that necessarily needs the aesthetic input from a human operator [1]–[3]. To this end, various investigators [1], [2] have developed interactive techniques in which various parameters are adjusted interactively until a visually pleasing result is attained. However, current research has attempted to develop automated gamut mapping algorithms that could be implemented in commercial devices. As has been noted by Fairchild [4], a further complication to this end is that

Manuscript received March 28, 1996; revised January 28, 1997. This research was supported by the NSF-NYS/IUCRC and NYSSTF-CAT Center for Electronic Imaging Systems. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ping Wah Wong.

The authors are with the Munsell Color Science Laboratory, Center for Imaging Science, Rochester Institute of Technology, Rochester, NY 14623 USA (e-mail: edmpci@ritvax.isc.rit.edu).

Publisher Item Identifier S 1057-7149(97)04904-X.

gamut mapping is application dependent. A business graphic, a pictorial image, and a computer generated image may require different mapping techniques for optimal image quality.

The goal of this research is to evaluate a variety of gamut mapping techniques in order to assess whether particular algorithms can be chosen which produce the best gamut-limited images for simple pictorial imagery. As a first step to addressing this problem and making it more tractable, we have simplified it in a number of ways. For these experiments we have used a CRT as both the original and destination device avoiding problems of change of color appearance under differing media and viewing conditions. By creating artificial gamut boundaries based on the statistics of our original images, we avoided the complication of the characterization and calibration of hard copy media while being able to assess the effects of equivalent levels of gamut mapping for different hues. The gamuts were limited only in one color dimension at a time. In three separate experiments the gamut was limited in lightness at the high lightness levels of the gamut, the low lightness levels of the gamut, and in chroma. For simplicity, we refer to the gamut mapping of lightness at high lightness levels as lightness mapping at the top of the gamut. Likewise, gamut mapping the low lightness region of the gamut will be referred to as lightness mapping at the bottom. Our original images consisted of seven computer rendered single-colored spheres floating in front of a gray background. These simple images contained many of the characteristics of real pictorial images while allowing us to selectively test different areas of color space individually.

In the next section of the paper, we present the framework for the implementation of gamut mapping in our experiments. We begin by discussing the choice of device-independent color space used in our experiments. This is followed by an explanation of our device characterization, gamut determination, and choice of stimuli. We then present the algorithms that we tested in our experiments. In the Experimental section of the paper, we describe the methodology used to collect our psychophysical data and report and discuss the results of the experiments. This is followed by our conclusions.

## II. IMPLEMENTATION

### A. Device-Independent Colorimetry

For gamut mapping between different devices, a common color language is needed in order to specify colors

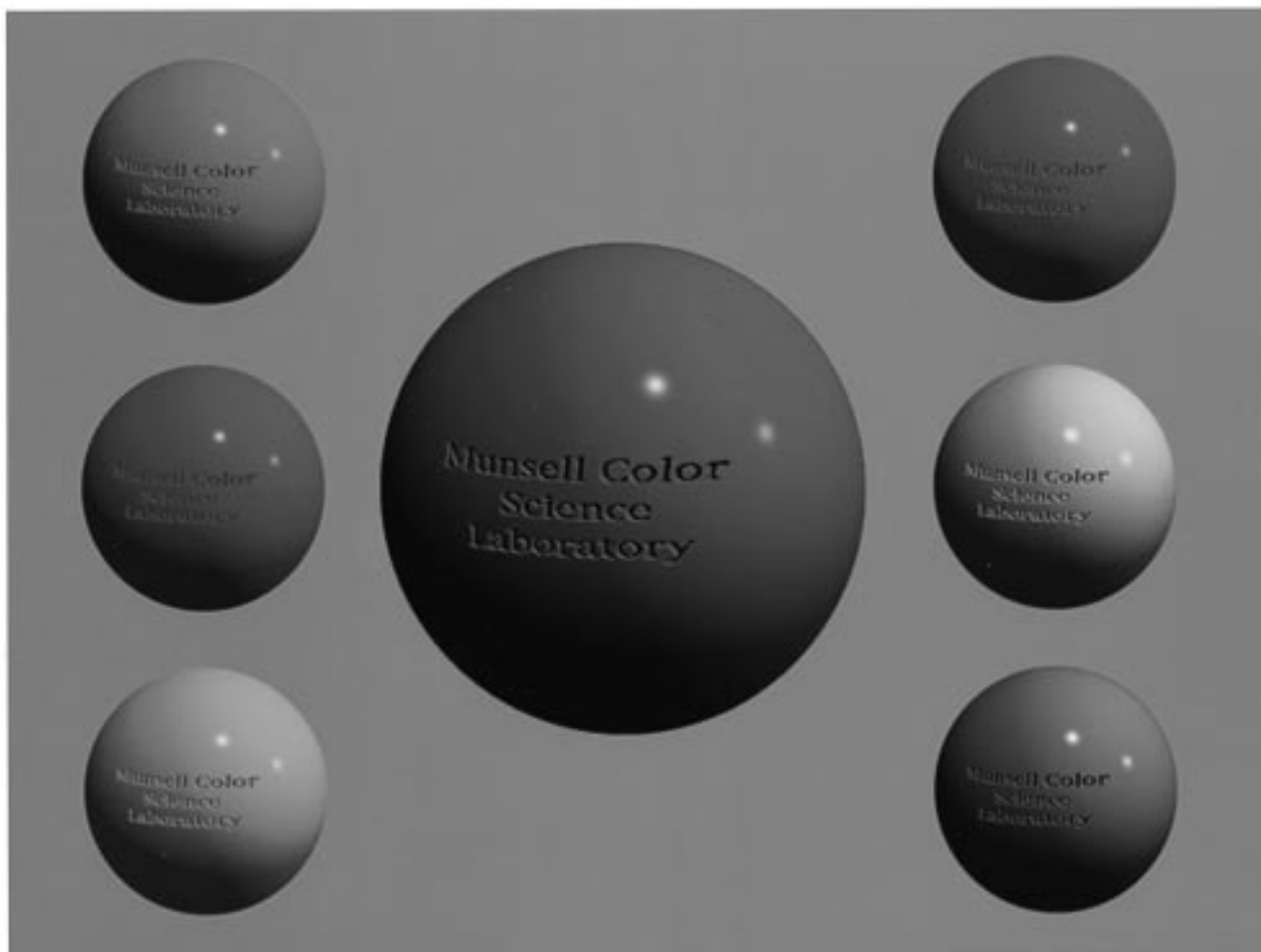


Fig. 1. Seven colored spheres contained in the seven original images.

colorimetrically. For such device-independent color, a color appearance model that can accurately account for differences in viewing conditions between the two devices would be the ideal choice [4]. In this way, perceptual attributes of color such as hue, chroma, and lightness can be specified and altered independently. In addition such models have built-in features that allow specification of certain parameters affecting appearance such as the white point of the illuminant, level of adaptation, surround conditions, etc. A number of models have been proposed for this use and have been the focus of much research [5]–[7]. In order to simplify the problem, the gamut mapping in our experiments is performed on the same device using artificial gamut boundaries. Therefore, the viewing conditions and level of adaptation are held constant.

Although not intended as a color appearance space, CIE 1976  $L^*a^*b^*$  was chosen as the device-independent color space for these experiments, since it has correlates to perceived lightness, hue, and chroma, is easily invertible, and has been shown to perform adequately as an appearance space [5]. Although CIE 1976  $L^*u^*v^*$  has been chosen by others [2], [8] because it has a perceptual correlate to saturation, CIELAB performs significantly better as a color appearance space [5]. Because CIELAB does not contain a constant luminance chromaticity diagram, a correlate of saturation is not defined.

However, we used the ratio of  $C^*/L^*$  as an analog of saturation in CIELAB. In this way we could manipulate either chroma or saturation to determine which attribute was more salient for the preservation of image fidelity.

#### B. Device Characterization, Gamut Determination, and Imagery

Because different imaging devices have different gamuts, the amount of gamut mapping required is determined by both the device gamuts and the range of colors in the image. For real devices, determining the corresponding color coordinates for colors common to the two gamuts in an independent color space is a difficult procedure involving precise characterization of both devices and interpolating between measured or modeled data. For these experiments, only the computer monitor needed to be calibrated and characterized.

We took the complete color gamut of our monitor to be the gamut of our original device. Artificial gamut boundaries, in chroma and lightness, were then established based on the statistics of our original images. That is, for each image, a cut-off value for chroma or lightness was determined so that a certain percentage of the pixels of the object in the original image was beyond the gamut limit.

TABLE I  
HUE ANGLE AND IMAGE CUT-OFF VALUES FOR EACH IMAGE IN ALL THREE EXPERIMENTS

Image	Hue angle	Lightness at top		Lightness at bottom		Chroma	
		cut-off $L^*$	max $L^*$	cut-off $L^*$	min $L^*$	cut-off $C^*$	max $C^*$
Red Sphere	33°	53.5	87.3	24.4	9.4	76.7	103.9
Green Sphere	142°	68.6	90.6	35.8	16.2	87.4	118.2
Blue Sphere	-61°	40.9	86.2	14.6	3.2	92.6	126.6
Cyan Sphere	-161°	70.7	90.9	37.3	17.2	45.1	60.6
Magenta Sphere	-30°	56.9	87.8	27.3	11.2	77.4	105.4
Yellow Sphere	101°	78.2	95.6	41.1	19.9	66.1	89.4
Skin Sphere	49°	57.6	95.9	28.1	11.3	32.3	45.2

Our original images consisted of seven images of individual colored spheres floating in front of a gray background. The colors of the spheres were red, green, blue, cyan, magenta and a “skin” tone. These images were created in a three-dimensional (3-D) ray-tracing rendering program. Some manipulation of the colors in the images was needed to prevent some of the gamut mapping routines from shifting colors out of the CRT gamut. Fig. 1 shows in one image the seven spheres used in our original images. The nominal hue angle is tabulated in Table I. The objects in our images contain a full range of lightness and chroma. The images have gradual shading, highlights, and a range of contrasts that are altered dramatically depending on the method of gamut mapping used. In this way they are representative of real pictorial images.

As mentioned above, gamut mapping is application dependent. Even for pictorial images, previous investigators [2], [10] have found that the preferred gamut mapping techniques may depend on image content. Because our images contain single hues, we can test whether image dependence may be due to interactions between different hues. Since the same proportion of each colored image is being mapped, preference for different algorithms for the different colored images would indicate that the choice of algorithm depends on the region of color space being mapped. The lightness cut-off for gamut mapping at the top was chosen so that one-quarter of all the pixels in the sphere were out of gamut. One-third of the pixels in the sphere were below the cut-off in the lightness mapping at the bottom experiment. For chroma mapping, the cut-off was chosen so that one-half of the pixels were out of gamut. Table I shows the cut off values for chroma and lightness for each image in the three experiments.

Fig. 2(a)–(c) show the  $L^*$  and  $C^*$  values of the pixels and the gamut cut-off values for the green image in the three experiments. A different original was used for each colored image in the three experiments. For both lightness mapping experiments, a less saturated image was used in order to keep the colors in the image within the real gamut of the CRT. The differences in the pixel values for the two originals in the lightness mapping experiments were the result of changes in the monitor set-up and calibration between these two experiments.

### C. Mapping Algorithms

Three experiments tested limiting the gamut in lightness,  $L^*$ , both at the top (high lightness values) and at the bottom (low lightness values) and limiting the gamut in chroma,  $C^*$ . Following previous research [2], [8]–[12], hue was kept

constant while lightness and chroma were compressed and clipped in various ways. Fig. 3(a) shows conceptually the rationale for our choice of techniques. The abscissa ( $A_{in}$ ) represents the original gamut of the particular color attribute (chroma or lightness) to be mapped and the ordinate is the gamut available in the gamut limited reproduction device ( $A_{out}$ ). The solid horizontal line shows the location of the gamut limit. The upper boundary of the shaded region is the mapping produced by clipping the out-of-gamut colors to the gamut boundary. The lower boundary of the shaded region is the mapping produced by linearly scaling the color attribute to just fit into the limited gamut. The shaded region is therefore the region in which the optimal gamut technique lies. Our mapping techniques tested this region with curves made up of two or three linear segments. Gentile *et al.* [9] also tested piecewise linear compression and reported that these curves gave results that were similar to smoother curves based on nonlinear functions. Fig. 3(b) represents the same region for mapping lightness at the dark end of the gamut. It is identical to Fig. 3(a) turned upside down. Since only lightness is mapped in this manner, the letter  $L$  is substituted for  $A$ .

Fig. 3(a) shows the gamut scaled to the maximum value,  $A_{input}^{MAX}$ , of the attribute attainable on the original device. Gamut mapping to this point will be referred to as device dependent mapping. Alternatively, the mapping algorithms can scale the reproduction in the destination device to the point  $A_{image}^{MAX}$ , which is the maximum value of the attribute in the original image. This mapping, referred to as image dependent mapping, is more expensive to implement since image analysis is needed to find the maximum chroma and lightness values in the image. For mapping lightness at the bottom of the gamut, device dependent mapping is when the image is scaled to  $L^* = 0$  in the reproduction device. For image dependent mapping, the smallest  $L^*$  value in the image,  $L_{image}^{MIN}$ , is the point to which the reproduction is scaled.

For each of the above mapping algorithms drawn from this space, the gamut mapping was accomplished in two ways. For lightness mapping at both the top and bottom of the gamut, one gamut limited image was produced in which the chroma values were the same as in the original image and one image was produced in which the  $C^*$  values were adjusted so that the saturation,  $C^*/L^*$ , in the new image was identical to the original. Likewise for chroma mapping, one image was produced in which the  $L^*$  values were the same as in the original and in a second reproduction, the  $L^*$  values were recalculated in order to preserve the original’s saturation. This is shown schematically in Fig. 4.

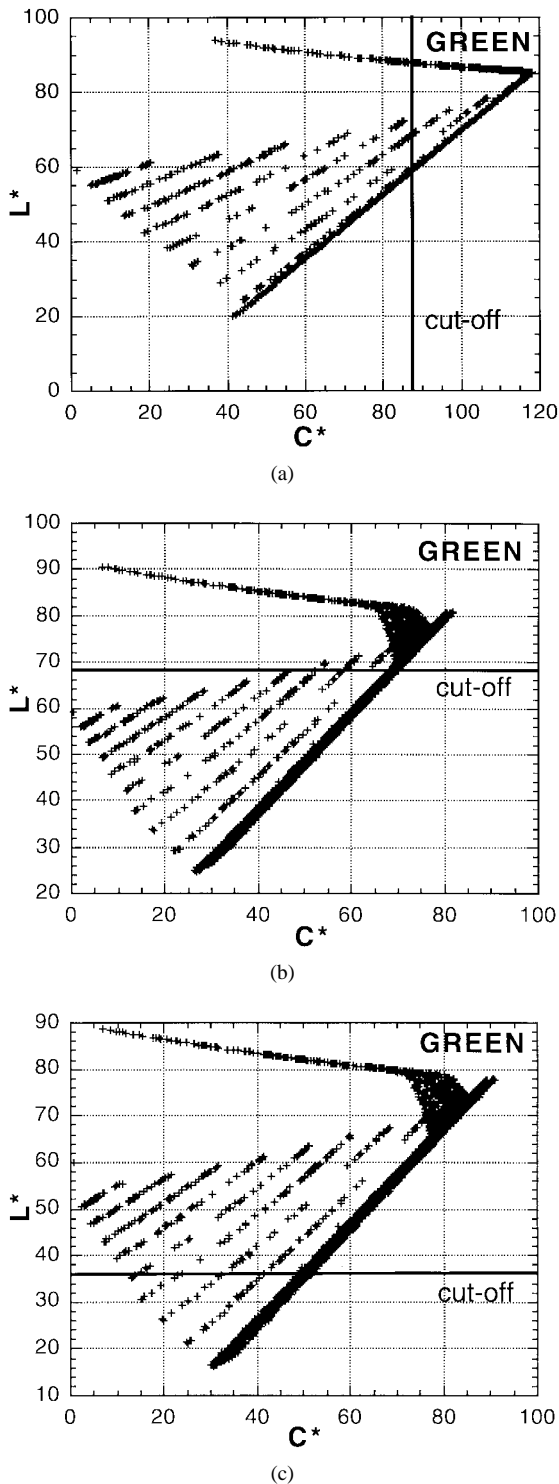


Fig. 2. Plots of the  $L^*$  versus  $C^*$  values for the green original images for the three gamut mapping experiments showing the cut-off values for the reduced gamuts. (a) Chroma mapping. (b) Lightness mapping at high values of  $L^*$ . (c) Lightness mapping at low values of  $L^*$ .

Four families of gamut mapping algorithms were drawn from this space, as follows.

1) *Scaling and Clipping (SC)*: The first set of algorithms are shown in Fig. 5(a)–(c) and are denoted “SC.” Fig. 5(a) and (b) shows the six algorithms used for lightness mapping at the top of the gamut and chroma mapping, respectively.

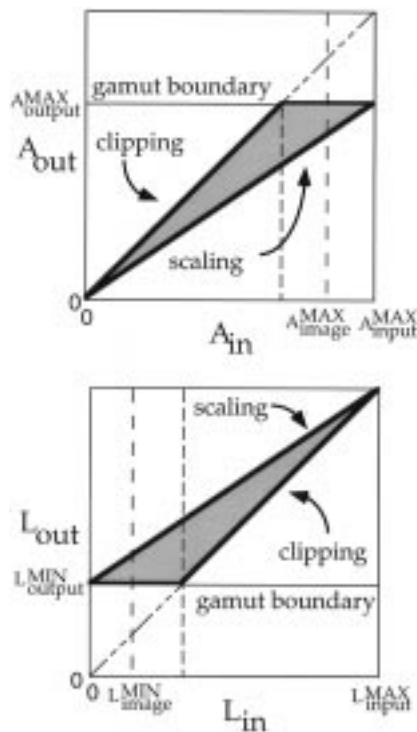


Fig. 3. Shaded region represents the area from which the linear piecewise gamut mapping techniques are drawn. (a)  $A_{in}$  represents the gamut of the color attribute being mapped for the original imaging device and  $A_{out}$  represents the attribute gamut for the destination device. (b) The same region drawn for gamut mapping the low lightness region of the gamut.

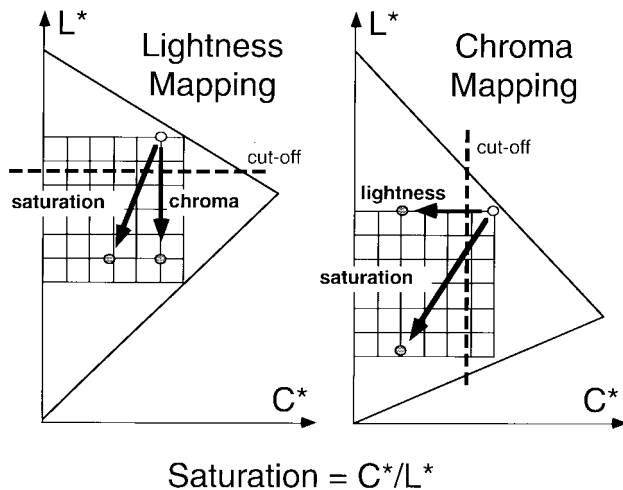


Fig. 4. Lightness mapping and chroma mapping were accomplished in two ways shown here schematically. For lightness mapping, either the chroma or saturation of the original image could be maintained in the reproduction. For chroma mapping either lightness or saturation could be held constant. The outer triangle represents the gamut of the original input device.

The equation of these curves is

$$A_{out} = \begin{cases} A_{in} \frac{assymp}{k}, & 0 \leq A_{in} \leq k, \\ assymp, & k \leq A_{in} \leq A_{input}^{MAX} \end{cases} \quad (1)$$

where “assymp” is the cut-off value for the gamut boundary,  $A_{in}$  is the original lightness or chroma value,  $A_{out}$  is the mapped value, and the value  $k$  sets the location of the inflection point.

TABLE II  
SUMMARY OF THE GAMUT MAPPING ALGORITHMS AND THEIR NOTATIONS

	Mapping Algorithm				
	Scaling & Clipping	Knee Functions	Gentile et al.	Three Segments	Color Space Scaling
Lightness Mapping at High $L^*$	0SC (clipping) .25SC .5SC .75SC ImSC (image dependent) 1SC (device dependent)  + s	.25KF .5KF .75KF  + I or D + s	1/3Gea 2/3Gea  + I or D + s	3S   + I or D + s	CIELAB RGB
Lightness Mapping at Low $L^*$	1SC (image dependent) 1/3SC 2/3SC 0SC (clipping)  no + s	.25KF .5KF .75KF  + I only no + s	1/3Gea 2/3Gea  + I only no + s	3S   + I or D no + s	CIELAB RGB
Chroma Mapping	0SC (clipping) .25SC .5SC .75SC 1SC (image dependent) DSC (device dependent)  + s	.25KF .5KF .75KF  + I or D + s	1/3Gea 2/3Gea  + I or D + s	3S   + I or D + s	CIELAB RGB
Appended notation:		+ s $\rightarrow$ saturation of original maintained + I or D $\rightarrow$ image or device dependent implementation			

For lightness mapping, the notation indicates the location of the inflection by the fractional distance between complete clipping (OSC) and complete device dependent scaling (1SC). “ImSc” represents complete image dependent scaling where  $k$  is equal to the maximum lightness value in the image. For chroma mapping the notation is slightly different. The notation indicates the location of the inflection by the fractional distance between complete clipping (0SC) and complete image dependent scaling (1SC). “DSC” represents complete device dependent scaling where  $k$  is equal to the maximum chroma value attainable by the device.

The equation for lightness mapping at the bottom of the gamut is

$$L_{\text{out}} = \begin{cases} \text{assymp}, & 0 \leq L_{\text{in}} \leq k, \\ \frac{(\text{assymp} - 100)}{(k - 100)}(L_{\text{in}} - 100) + 100, & k \leq L_{\text{in}} \leq 100. \end{cases} \quad (2)$$

Fig. 5(c) shows the notation and curves for the four algorithms used in the experiment. Only image dependent algorithms were used so that the notation indicates the inflection by its fractional distance between complete clipping (OSC) and complete image dependent scaling (1SC).

2) *Knee Functions (KF)*: The second set of algorithms, called *knee functions*, is shown in Fig. 6(a). The first segment has a slope of one and the second segment compresses lightness or chroma to either the maximum value in the image (KFI) or for the device (KFD). The equation for these curves is

$$A_{\text{out}} = \begin{cases} A_{\text{in}}, & 0 \leq A_{\text{in}} \leq k, \\ k + \frac{(A_{\text{in}} - k)(\text{assymp} - k)}{(A^{\text{MAX}} - k)}, & k \leq A_{\text{in}} \leq A^{\text{MAX}} \end{cases} \quad (3)$$

where  $A^{\text{MAX}} = A_{\text{image}}^{\text{MAX}}$  for image-dependent mapping and

$A^{\text{MAX}} = A_{\text{input}}^{\text{MAX}}$  for device-dependent mapping. An  $I$  or a  $D$  is appended to the notation to indicate image or device-dependent mapping.

The inflection point is indicated in the notation by the fractional distance between zero and the cut-off value setting the gamut boundary. For example, for mapping the lightness of the cyan sphere with a cut-off value of  $L^* = 70.7$ , .75KFI and .75KFD both have the same lower segment up to  $L^* = 53$ , but the former compresses the lightness to  $A_{\text{image}}^{\text{MAX}}$ , the maximum value in the image,  $L^* = 90.9$ , and the latter to  $A_{\text{input}}^{\text{MAX}}$ , the maximum value of the device,  $L^* = 100$ .

For lightness mapping at the bottom, an inverted set of these three curves were used (not shown). Only the image dependent form was used. The equation for these is

$$L_{\text{out}} = \begin{cases} \text{assymp} + \frac{(L_{\text{in}} - L_{\text{image}}^{\text{MIN}})(k - \text{assymp})}{(k - L_{\text{image}}^{\text{MIN}})}, & 0 \leq L_{\text{in}} \leq k, \\ L_{\text{in}}, & k \leq L_{\text{in}} \leq 100, \end{cases} \quad (4)$$

3) *Gentile et al. (Gea)*: The next family curves, Gentile *et al.* (Gea), uses a technique described in [9]. Shown in Fig. 6(b), these two-segment curves have an inflection located at the value of the cut-off on the abscissa but the slopes of the linear segments vary according to the fractional distance along the vertical line extending between complete compression and complete scaling. Again, the right-hand segment can be either determined by the image (GeaI) or the device (GeaD) maximum values.

For mapping lightness at the top of the gamut and for chroma mapping the equation of these curves takes the form

$$A_{\text{out}} = \begin{cases} A_{\text{in}} \left( (1 - k) \left( \frac{\text{assymp}}{A^{\text{MAX}}} \right) + k \right), & 0 \leq A_{\text{in}} \leq \text{assymp} \\ \text{assymp} \left( A_{\text{in}} \left( \frac{(1 - k)}{A^{\text{MAX}}} \right) + k \right), & \text{assymp} \leq A_{\text{in}} \leq A_{\text{input}}^{\text{MAX}} \end{cases} \quad (5)$$

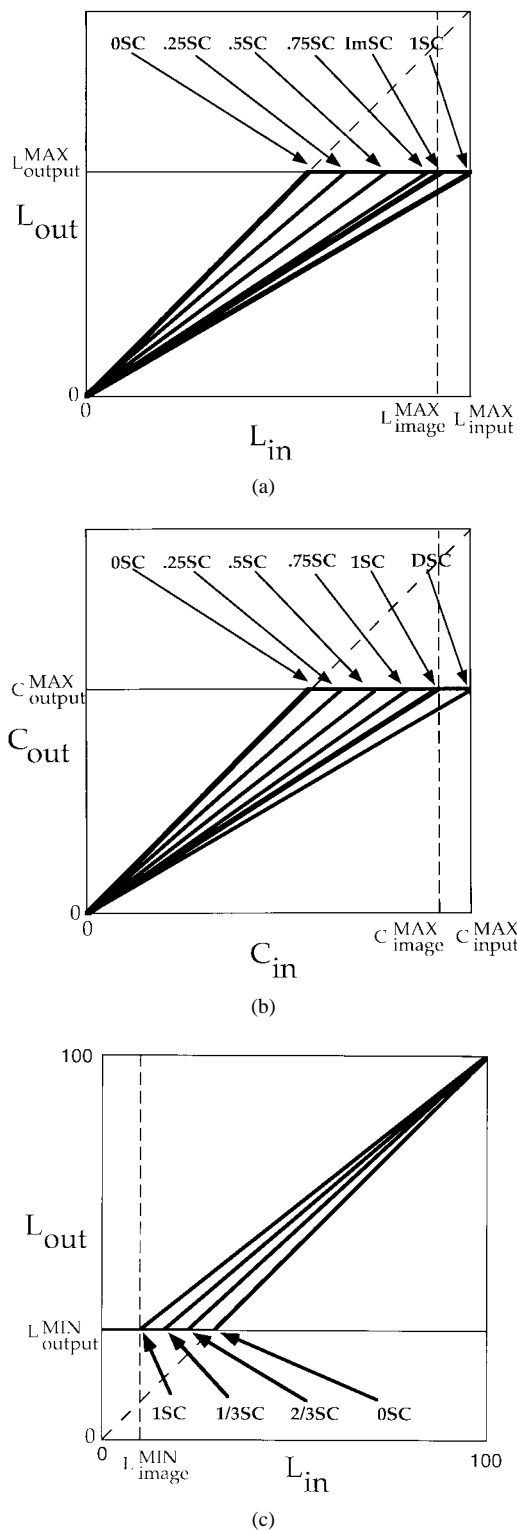


Fig. 5. Scaling and clipping (SC) algorithms. (a) Lightness mapping at the top of the gamut. (b) Chroma mapping. (c) Lightness mapping at the bottom of the gamut.

where  $A^{MAX} = A_{image}^{MAX}$  for image dependent mapping and  $A^{MAX} = A_{input}^{MAX}$  for device-dependent mapping. Here,  $k$  has a value between zero, which produces a curve that linearly scales the complete gamut, and one, which produces a curve that clips the gamut. Since the SC algorithm was used to produce the complete clipping and scaling curves, only the curves produced

by  $k$  values of 1/3 and 2/3 are used.

For mapping lightness at the bottom of the gamut, curves derived by the following equation were used:

$$L_{out} = \begin{cases} \frac{(L_{in} - L_{image}^{MIN})(1-k)(100 - \text{assymp})}{(100 - L_{image}^{MIN})} + \text{assymp}, & 0 \leq L_{in} \leq \text{assymp}, \\ \left( \frac{\text{assymp}(k-1) - kL_{image}^{MIN} + 100}{100 - L_{image}^{MIN}} \right) (L_{in} - 100), & \text{assymp} \leq L_{in} \leq 100. \end{cases} \quad (6)$$

Again, these curves are the identical to Fig. 6(b) inverted. Only the image-dependent algorithms using  $k = 1/3$  and  $2/3$  were used.

4) *Three Segment (3S)*: The last group of curves are curves made up of three segments shown in Fig. 6(c). The equation of these curves is

$$A_{out} = \begin{cases} A_{in}, & 0 \leq A_{in} \leq k1 \\ \frac{(\text{assymp} - k1)(A_{in} - k1)}{(k2 - k1)} + k1, & k1 \leq A_{in} \leq k2, \\ \text{assymp}, & k2 \leq A_{in} \leq A_{input}^{MAX}. \end{cases} \quad (7)$$

The values  $k1$  and  $k2$  are the values of the two inflections on the abscissa. For the device-dependent form of the algorithm, 3SD,  $k1 = \text{assymp}/2$  and  $k2 = (\text{assymp} + A_{input}^{MAX})/2$ . For the image-dependent form of the algorithm, 3SI,  $k1 = (A_{image}^{MIN} + \text{assymp})/2$  and  $k2 = (\text{assymp} + A_{image}^{MAX})/2$ .

For lightness mapping at the bottom of the gamut, the equations of the 3S curves are

$$L_{out} = \begin{cases} \text{assymp}, & 0 \leq L_{in} \leq k1, \\ \frac{(k2 - \text{assymp})(L_{in} - k1)}{(k2 - k1)} + \text{assymp}, & k1 \leq L_{in} \leq k2, \\ L_{in}, & k2 \leq L_{in} \leq 100. \end{cases} \quad (8)$$

For the device-dependent curves, 3SD,  $k1 = \text{assymp}/2$  and  $k2 = (\text{assymp} + 100)/2$ . For the 3SI image-dependent curves,  $k1 = (L_{image}^{MIN} + \text{assymp})/2$  and  $k2 = (\text{assymp} + L_{image}^{MAX})/2$ .

For lightness mapping, at the top and bottom, each algorithm was used twice to create one test image with the same hue and chroma as the original and one test image with the same hue and saturation as the original. Similarly for chroma mapping, each algorithm was used to create two images: one with the same lightness and hue as the original, and one with the same saturation and hue. Therefore, there were a possible total of 36 test images to be tested for each original using the piecewise linear algorithms. A lower-case "s" was appended to the notation to indicate that the gamut mapped image retained the same saturation as the original, the absence of the "s" indicates that lightness (for chroma mapping) or chroma (for lightness mapping) was held constant.

For chroma mapping and lightness mapping at the top of the gamut, the number of algorithms used to create test images was sometimes less than 36, because different algorithms would sometimes lead to functions that were very similar. For lightness mapping at the bottom of the gamut, only image-dependent algorithms for KF and Gea were used based on

TABLE III  
RESULTS FROM THE THREE MAPPING EXPERIMENTS SHOWING FOR EACH IMAGE THE ALGORITHMS USED TO CREATE TEST IMAGES FOR THE 20 SUBJECTS ORDERED BY THEIR RANKED PERFORMANCE. THE BEST ALGORITHM HEADS EACH COLUMN AND IS GROUPED TOGETHER ABOVE THE DOUBLE LINE WITH THOSE FROM WHICH IT IS NOT SIGNIFICANTLY DIFFERENT

Lightness Mapping at the Top						
Red	Green	Blue	Cyan	Magenta	Yellow	Skin
.25SCs	1/3Geals	1/3Geals	1/3Geal	1/3Geals	RGB	.5SCs
2/3Geals	.5SCs	.5SCs	.5SCs	.5SCs	1/3Geals	1/3Geals
1/3Geals	ImSCs	.25KFIs	1/3Geals	1/3Geal	.5SCs	.75SCs
.5SCs	CIELAB	CIELAB	RGB	3SDs	CIELAB	.25KFIs
.25KFIs	.25KFIs	ImSCs	ImSCs	.25KFIs	ImSCs	ImSCs
ImSCs	RGB	1/3Geal	CIELAB	.5SC	.25KFIs	CIELAB
CIELAB	3SDs	.5SC	ImSC	ImSCs	.5KFIs	1/3Geal
1/3Geal	1/3Geal	.25KFI	.25KFI	RGB	1SCs	.75SC
	.5KFIs		.25KFIs	CIELAB		
			.5KFI			
			.5KFIs			
Lightness Mapping at the Bottom						
Red	Green	Blue	Cyan	Magenta	Yellow	Skin
3SI	0SC	0SC	2/3SC	2/3SC	3SI	0SC
0SC	2/3SC	3SI	1/3SC	3SI	1/3Geal	2/3SC
2/3SC	3SI	3SD	3SI	2/3Geal	2/3SC	3SI
2/3Geal	.25KFI	2/3SC	1SC	3SD	1/3SC	2/3Geal
3SD	1/3SC	2/3Geal	1/3Geal	1/3SC	3SD	3SD
1/3SC	1/3Geal	1/3SC	.25KFI	1/3Geal	2/3Geal	1/3SC
1/3Geal	3SD	1/3Geal	3SD	1SC	1SC	1/3Geal
.25KFI	2/3Geal	.25KFI	2/3Geal	RGB		.25KFI
1SC	1SC					1SC
	RGB					
Chroma Mapping						
Red	Green	Blue	Cyan	Magenta	Yellow	Skin
2/3Geals	0SC	3SIs	0SC	0SC	0SC	0SC
1/3Geals	.25SC	.25SC	.25SC	.25SC	.25SC	.25SC
.75SCs	.75KFI	.5SC	3SI	2/3Geal	2/3Geal	2/3Geal
1SCs	2/3Geal	RGB	.75KFI	.5SC	3SI	3SD
RGB	3SI	2/3Geal	2/3Geal	3SI	.75KFI	3SI
.25SC	.5SC	.75SC	.5SC	2/3GealD	.5SC	.75KFI
CIELAB	3SD	.75SCs	1/3Geal	1/3Geal	1/3Geal	.5SC
.25KFIs	1/3Geal	1/3Geal	.75SC	.75SC	.5KFI	.75SCs
2/3Geal		1SC	.5KFI	.25KFI	.75SC	1/3Geal
.5SC		.25KFI	CIELAB	1SC		.75SC
3SDs						
3SI						
2/3GealD						
.75SC						
1/3Geal						

the results of the other two mapping experiments. Also for lightness mapping at the bottom, we only used the algorithms in the form that retained the original images' chroma and allowed the saturation to vary. Adjusting chroma in order to retain the originals' saturation led to problems of colors moving out of the CRT gamut for many of the images. While producing these images, it was apparent that they were not producing images that were as faithful to the originals as those that held chroma constant and allowed saturation to vary.

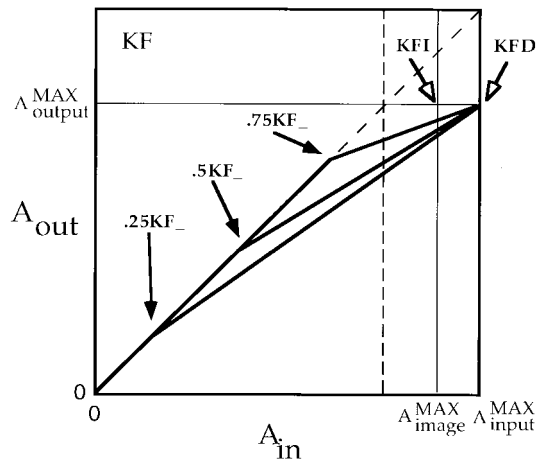
In addition to the above piecewise linear algorithms, we created test images by scaling the color gamut in CIELAB space and in the monitor RGB space, denoted "CIELAB" and "RGB," respectively, bringing the maximum number of algorithms to 38. For lightness at the top and chroma, all three dimensions of each space were scaled by the same constant toward black so that the image just fit into the new gamut. The scaling of the image in CIELAB space toward black was equivalent to the complete image dependent scaling of the image while preserving saturation. Although we did not realize this until after completion of the experiments, it provided a

useful check on our results. For lightness scaling at the bottom, the color spaces were scaled toward white by the same constant in all three dimensions until the image was in gamut. This is shown schematically in Fig. 7. A summary of the algorithms and their notations is presented in Table II.

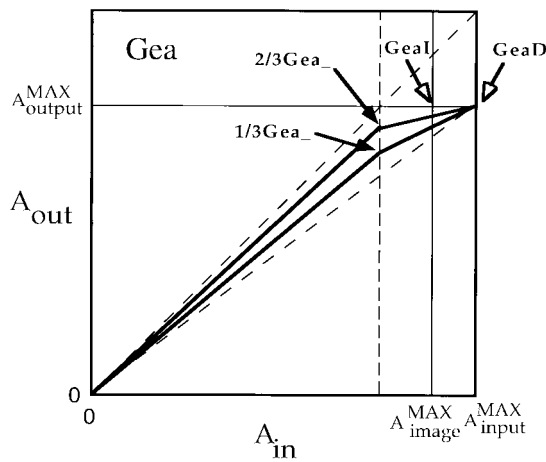
### III. EXPERIMENT

#### A. Procedure

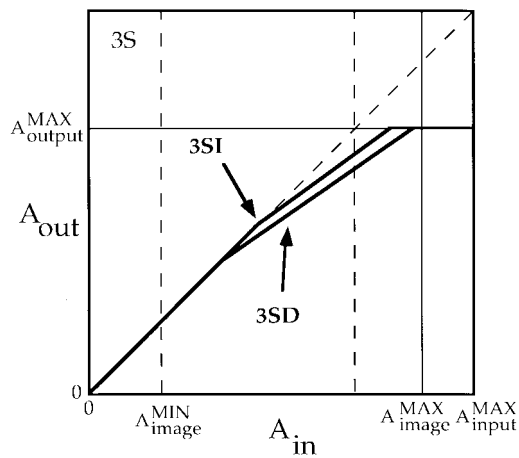
Each image was 422 × 344 pixels presented on a Sony 20-in diagonal Trinitron™ CRT subtending approximately 5.5° at a viewing distance of approximately 0.75 m. The monitor was calibrated and characterized following the procedures found in Berns *et al.* [13]. The original and two gamut mapped reproductions (the test images) were presented simultaneously in each trial on a neutral gray background ( $x = 0.324, y = 0.343, 38.8 \text{ cd/m}^2$ ) The subjects task was to choose which of the reproductions more faithfully reproduced the object represented in the original. More specifically the instructions stated, "... For example, you sell the object pictured in the



(a)



(b)



(c)

Fig. 6. (a) "Knee function" (KF) algorithms. (b) "Gentile *et al.*" (Gea) algorithms. (c) "Three segment" (3S) algorithms.

original and you would like the original picture reproduced in your brochure. Your two printers give you what is shown in the reproductions. Which of the two reproductions would you choose for your brochure?" Subjects freely viewed the monitor and proceeded at their own pace.

With up to  $n = 38$  test images, a paired comparison experiment would need  $n(n - 1)/2 = 703$  trials for one subject to

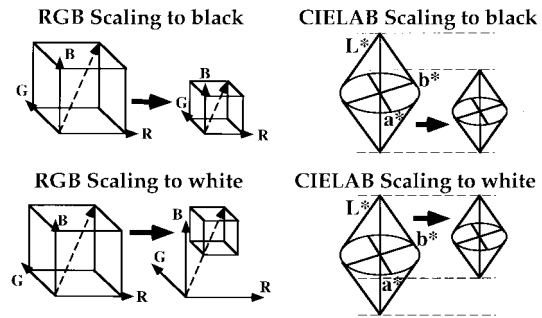


Fig. 7. RGB and CIELAB color space scaling.

view each pair only once. For seven original images and three mapping experiments, this would be prohibitive. To reduce the number of trials to a more practical amount, one subject, author EM, performed the complete paired-comparison experiment for all the test images created using all the algorithms. Each original colored image and one gamut mapping region were tested per session, making 21 sessions. Every trial pair was presented twice in random order to counterbalance screen position. Based on the rank order of the test subject's results, the top six to 15 algorithms for each original image were then used in a paired-comparison experiment with other subjects. An example of one of the ordered histograms showing the frequency at which images produced by each algorithm were chosen as the member of a pair more like the original is given in Fig. 8. Nine test images to the left of the vertical line (excluding 1/3GeaDs, due to its similarity to 1/3GeaIs) were presented to the rest of the subjects as test images for the magenta original in the lightness mapping at the top experiment. Table III shows the algorithms used to create test images for each image in all three experiments. The best algorithms selected by the rank order from EM's data for use in the experiment were predominantly image dependent.

Twenty color-normal subjects participated in each of the experiments in this order: lightness mapping at the top of the gamut, chroma mapping and lightness mapping at the bottom of the gamut. Subjects typically took less than one hour to complete one experiment. Again, every pair was presented twice in a random order to counterbalance screen position.

### B. Results and Discussion

Based on Thurstone's Law of Comparative Judgment [14], the following procedures were used to determine which algorithms performed the best. The data for all twenty subjects were combined. For each of the seven originals in each experiment, the proportion of times each image was chosen over all the other images was converted into a normalized Z-score. These were then averaged over image to produce an interval scale value for that image. Higher scale values indicate a higher level of similarity to the original. The scale values were adjusted so that the least similar image for each original had a value of zero.

An analysis of variance [15], [16] based on the Z-scores gave significant results for all seven images in all three experiments. Confidence intervals were calculated by correcting the overall error term in the analysis of variance (ANOVA)

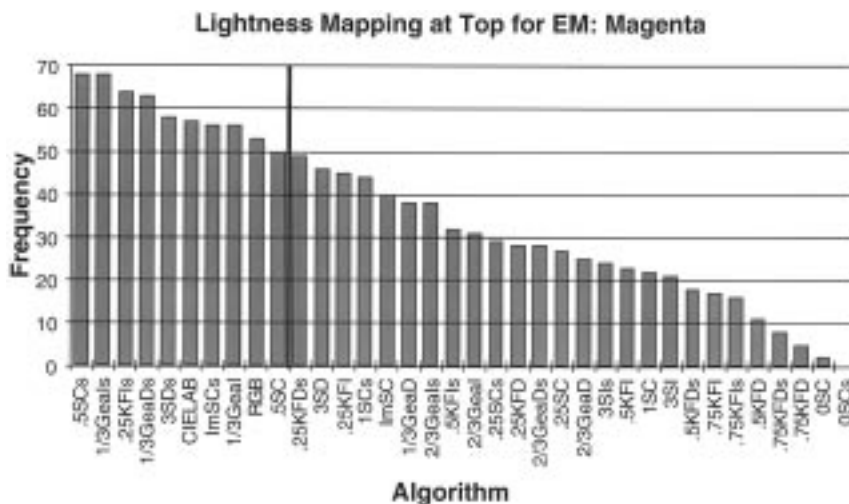


Fig. 8. Ordered histogram of the frequency of preferences for subject “EM” for the magenta image in the lightness mapping at the top experiment.

for multiple comparisons by the Tukey test [17]. One example for each experiment is shown in Fig. 9(a)–(c). A difference between two images is significant when the error bar of one does not extend past the mean of the other.

Fig. 9(a), for example, shows that the best test image for lightness mapping at the top was produced by the algorithm 1/3GeaIs. That is the Gea scaled so that the maximum  $L^*$  value in the image is just within the new gamut (I) and the chroma is adjusted to preserve the saturation in the original ( $s$ ). The 1/3GeaIs test image is not significantly different from the .5SCs image but is statistically better than the rest of the images tested, since the lower extent of its error bar is above the means of those images as shown by the horizontal line. Table III summarizes the results from all three experiments. The best algorithm leads each column followed by the others in rank order of performance. The algorithms that were not statistically different from the best algorithm are grouped together above the double line in each column. Using the above example for the blue image for lightness mapping at the top, Table III shows the 1/3GeaIs and .5SCs algorithms separated from the others by the double line.

As mentioned previously, no device-dependent algorithms were tested in the lightness mapping at the bottom experiment except for 3SD because the image dependent algorithms were generally better than the device dependent algorithms in the chroma mapping and lightness mapping at the top experiments.

For lightness mapping at the top, the best test images were the ones that retained the saturation in the original (as indicated by the small “s” in the notation). The one exception to this was the cyan image where chroma was held constant in the best algorithm (1/3GeaI), although it was not statistically different from .5SCs which did hold saturation constant. For the yellow image, RGB scaling had a similar effect of preserving the saturation of the original image, since the whole gamut was scaled in three dimensions. RGB scaling for the yellow image was not statistically better than four other algorithms, which kept saturation constant. Fig. 10(a) shows the original magenta image (left) and the images created by the 1/3GeaIs (center) and 1/3GeaI (right) algorithms to illustrate the best test image

and the effect of maintaining the chroma versus the saturation of the original image.

In order for the reproduction of the sphere to look compatible to the original when the lighter parts of it are mapped to lower levels, the chroma of the mapped region must be decreased in proportion so that the reproduced sphere has the identical saturation as the original. Although for lightness mapping at the bottom of the gamut, only algorithms that preserved the chroma of the original were tested, it was apparent while creating the stimuli for the experiment that boosting the chroma in order to maintain the original’s saturation does not produce images that are adequate reproductions of the original. That is, chroma needs to be maintained for the best image reproduction. This points out a curious apparent anisotropy in the relationship of lightness to colorfulness along its scale.

In general, the best chroma mapping test images retained the same lightness as the original images allowing saturation to vary. There were two exceptions: red and blue. More importantly, these two images were the only ones that did not have OSC as the statistically best algorithm. That is, for all the images, except red and blue, the best chroma mapping algorithm was the one where all out of gamut colors were clipped in chroma at the gamut boundary without changing lightness. This result agrees with the findings of Gentile *et al.* [9], who also found that clipping chroma while keeping lightness and hue constant was the best chroma mapping technique using CIELUV as the device-independent color space. Fig. 10(b) shows the images produced by complete clipping (the best algorithm, OSC, in the center) and complete image-dependent scaling (algorithm 1SC on the right) for the original green image (left).

That the red and blue images were exceptions in the chroma mapping experiment, both for choice of mapping algorithm and for which color attributes to hold constant, may be a result of the choice of device-independent color space. Hung and Berns [18] had subjects determine constant hue loci on a CRT for various hues and compared these results to the predicted constant hue loci in various color appearance spaces. When constant hue loci for hues of equal lightness

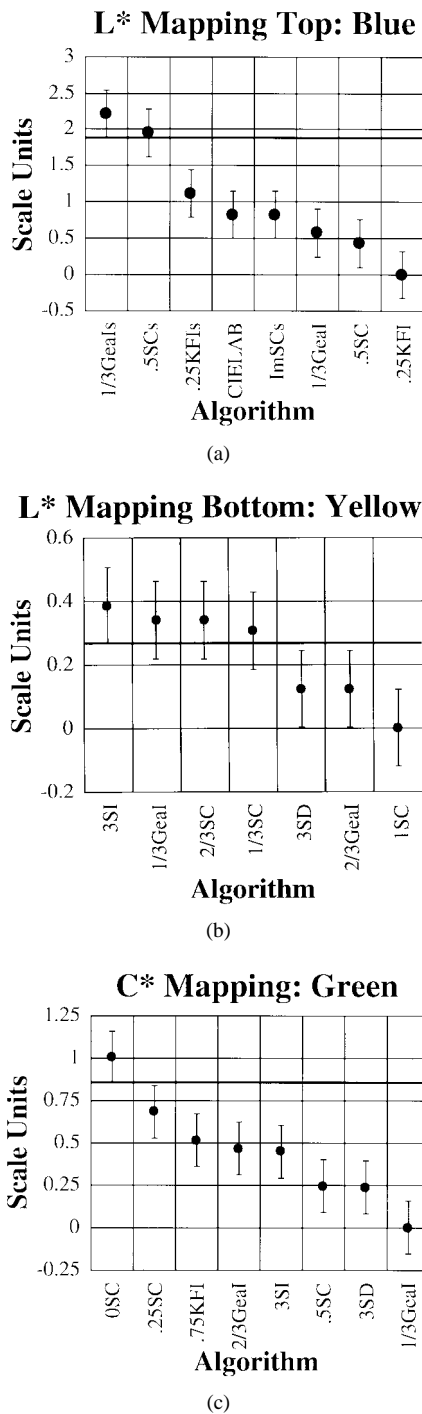


Fig. 9. Results for one original in each mapping experiment. (a) Lightness mapping at the top. (b) Lightness mapping at the bottom. (c) Chroma mapping.

were plotted in CIELAB space, the greatest deviations from linearity were in the red and blue regions of the color space. This confirmed their qualitative observations that when gamut mapping colors in CIELAB space, the blues became reddish and the reds turned yellowish. In our experiment the same was true. Subjects had to contend with changes in hue as well as lightness and chroma when comparing images.

Similarly, that the cyan image in lightness mapping at the top, was the exception to the general finding that holding

saturation rather than chroma constant leads to better mapping may also be a result of the choice of device independent color space. Hung and Burns [18] also compared constant hue loci for colors that varied in lightness. They found that in the cyan region of CIELAB space, colors on the subjects' constant hue loci did not have the same hue angle at different lightness levels. Therefore in our experiment, compressing lightness also produced hue shifts for the cyan image.

Taking a look at lightness mapping at the top again, the 1/3Geals algorithm and the .5SCs algorithm were predominantly better performers. Although not easily apparent in Figs. 3(a) and 4(b), these two algorithms are very similar for the lower part of the curve up to the inflection in the 1/3Geals algorithm when plotted together with the same cut off level. Fig. 11 shows these two algorithms plotted for the magenta image. The upper part of the curve past this bend affects the lightest areas of the image, which contain the highlights. The rate of compression, as indicated by the slopes of the curves, for all but the highest lightness values in the image, given by these two algorithms appears to produce the best mapping. The 1/3Geals algorithm produces a smoother transition at the higher lightness values, which may account for its predominance as the best algorithm. An algorithm, similar to Wallace and Stone's "knee function" [11] that is a linear function that approaches the cut-off at a tangent, would perhaps produce optimal mapping. The yellow image did not lead to the same level of discrimination as the other images since the best algorithm was not significantly different from four others, although 1/3Geals and .5SCs were both in the top group.

The red image is, however, an exception with algorithm .25SCs performing statistically better than the rest. Although this may be an indication of one area where the best gamut mapping method may be dependent on color, the 2/3Geals, 1/3Geals, and .5SCs algorithms had scale values very close to that of the .25SCs algorithm. Although we have used the Tukey test [17] in order to control the rate of false positives in our analysis, as the number of comparisons increases so does the possibility of rejecting the null hypothesis when it is actually true.

As mentioned previously, the CIELAB scaling of the image toward black was equivalent to complete image-dependent scaling of the image while preserving saturation. Although we tried to eliminate redundancy, one of the anonymous reviewers pointed this out. However, since these two methods produced identical images, we can check our methodology by examining whether these two algorithms were ever significantly different from each other. Looking at the results for lightness mapping at high values of  $L^*$ , we see that the ImSCs and CIELAB algorithms are next to each other in the ranking all but once. These two algorithms were never found to be significantly different statistically for lightness mapping, thus validating our methodology. For chroma mapping, these two algorithms (1SCs and CIELAB) were paired only for one image in the experiment and they were not significantly different.

The lightness mapping at the bottom experiment did not show the same level of discrimination as the other two experiments. More algorithms were not significantly different

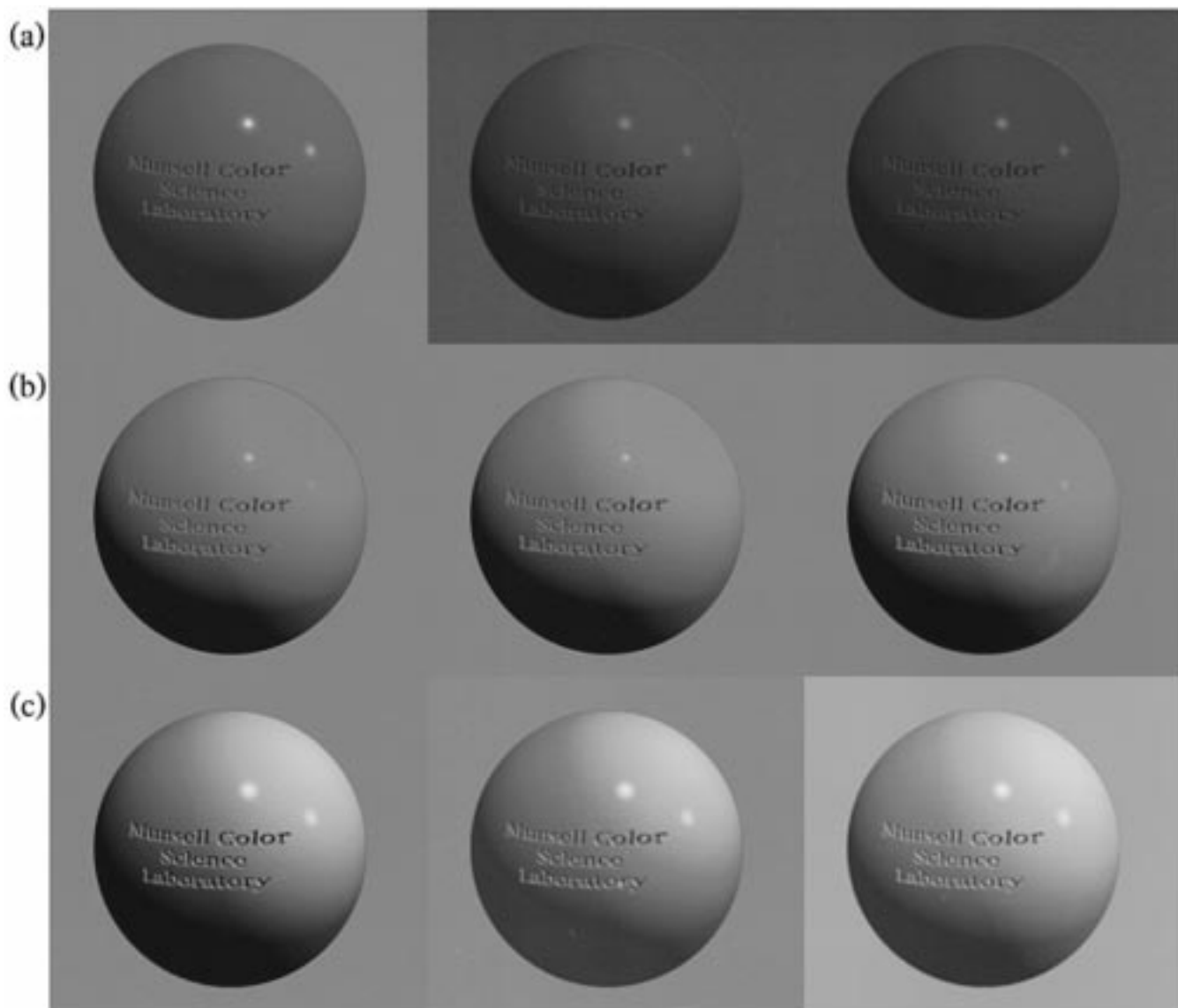


Fig. 10. (a) Original magenta image used in the lightness mapping experiment is reproduced on the left. In the center is the image created by the 1/3Geal algorithm, which was determined to be most like the original. The image produced using the 1/3Geal algorithms is shown on the right. (b) The green original in the chroma mappin experiment is on the left. The image produced by complete clipping, 0SC, the best algorithm is in the center. Complete image dependent scaling, 1SC, is shown on the right. (c) The original yellow image for lightness mapping at the bottom is on the left followed by the test image produced by 3SI, the best algorithm, and the image created by complete image-dependent scaling (1SC).

from the best one. There was also less consistency than the other two experiments as demonstrated by the lack of predominance of one algorithm over the rest. This would indicate that at least for the images and the levels chosen in this experiment, the choice of mapping algorithm was not as critical for good performance. However, there are certain characteristics shared by the aforementioned algorithms that are noteworthy.

The test image produced by the 3SI algorithm was in the top group for all the images. The next predominant algorithm was 2/3SC which was used to create test images for all seven originals and was in the top group five out of these seven times. 0SC was the next predominant, making the top group four times—three times as the best algorithm. Interestingly, 0SC was only tested with those four original images because it was not in the top ranking in EM's original experiment. These

three algorithms are shown in Fig. 12 for the green image. The 2/3SC and 3SI algorithms produce a similar amount of clipping at the bottom of the gamut. The 3SI and 0SC algorithms both leave most of the color gamut unchanged at higher lightnesses. This would indicate that optimal gamut mapping for lightness at the bottom might include some small amount of clipping in the darkest region followed by a small transitional region of compression followed by a large region of no gamut mapping at the higher lightnesses. This would be similar to the knee function inverted for gamut mapping at the bottom of the gamut with a linear section with a slope of one. These results are consistent with those of Hoshino and Burns [10]. As an example of lightness mapping at the bottom, Fig. 10(c) shows the original yellow image, the test image produced by 3SI, the best algorithm, and the image created by complete image-dependent scaling (1SC), respectively.

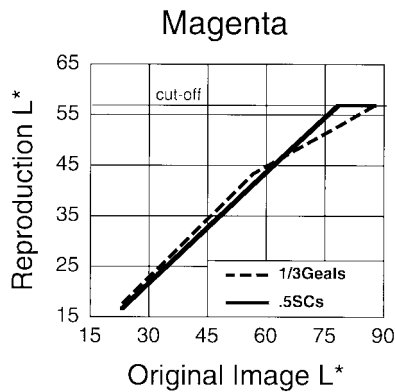


Fig. 11. 1/3Geals and .5SC's algorithm produced the images that were chosen as most like the magenta original for mapping lightness at high values of  $L^*$ .

#### IV. CONCLUSIONS

In general, for each mapping experiment, there was a fair amount of consistency in the choice of best mapping algorithm across color. This finding is encouraging in that it indicates that gamut mapping may be automatable. In these experiments the cut-off values were determined so that the same area of the spheres were out of gamut in all seven images. In actual gamut mapping, due to differences in lightness in the location of various hues in color space, the amount of mapping necessary for different hues will vary. Further experiments are needed in order to see if the consistency seen in these experiments holds when the cut-off values are constant across hue and when multiple hues are included in one image.

The linear piecewise algorithms used in this experiment all kept hue constant in CIELAB. This was an *a priori* decision based on the previous results of other investigators and principles derived from experience in graphic arts [1], [2], [8–12]. This would apply to the case of pictorial imagery. For computer-generated imagery and business graphics, gamut mapping techniques that alter hue in order to maximize chroma, select particular hues, or perform some other noncolorimetric color-enhancement technique may be more desirable [19].

In our case, in which we simulated pictorial graphic imagery, the choice of CIELAB as our device-independent color space may have affected the performance of the various gamut mapping techniques. For those regions of CIELAB in which the perceived constant hue loci deviated the most from those in the color space, the algorithms chosen by our observers were different than those in the more linear regions of the color space. The choice of color space may also have influenced which appearance attributes needed to be held constant (saturation or chroma). Ideally, a color appearance space with linear constant hue loci would better allow mapping of particular color attributes without altering appearance on other attributes. Of those tested by Hung and Berns [18], Hunt's model demonstrated the least hue error for constant hues of the same lightness and CIELAB was closest to linear for constant hues varying in lightness. The evaluation and development of color appearance spaces is an active area of research that, in the future, may provide even better tools for gamut mapping.

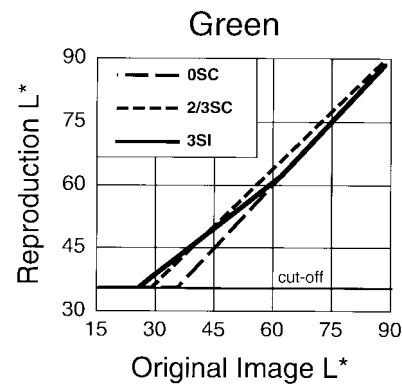


Fig. 12. 3SI, 2/3SC, and 0SC algorithms produced the images that were chosen as most like the green original for mapping lightness at low values of  $L^*$ .

The best gamut mapping algorithms are summarized below. The best performing algorithms from these experiments give indications of those features that the optimal gamut algorithm should possess, since the algorithms used did not exhaustively test every possible technique but sampled the solution space, Fig. 3(a) and (b), using a number of algorithms. Algorithms with smoother transitions between segments and that asymptote the gamut boundary gradually may, no doubt, provide better images.

- For lightness mapping at the top, retaining the saturation of the original with the 1/3Geals algorithm produced the best images. Most of the gamut, up to the cut-off value, was linearly compressed with a slope 1/3 of the way between those produced by complete linear scaling of the lightness values and completely clipping lightness at the cut-off value. This was followed by a smooth transition that scaled the remainder of the lightness to be just within the new gamut.
- For lightness mapping at the bottom, a curve with a slope of zero at the lowest lightness values followed by a moderate transition to linear compression with a slope of one for most of the gamut at the lighter end would perhaps perform well as indicated by the results. Keeping chroma, not saturation, constant produced better images.
- For chroma mapping, simply clipping the out of gamut chroma to the chroma cut-off level produced the most chosen mapping. For chroma mapping, preserving the original's lightness, not saturation, produced better images.

Since for chroma mapping, clipping was the best technique, no image analysis is necessary. However, for the lightness mapping at the top, it is clear from the results that image-dependent algorithms produced better results. For lightness mapping at the bottom only the image-dependent versions of the algorithms were tested based on the results from EM's data, which showed that the image-dependent algorithms were the ones chosen the most for chroma mapping and lightness mapping at the top. For lightness mapping, automated gamut mapping procedures need to analyze the image at least to determine the minimum and maximum lightness values.

It is common practice to set the white point of the reproduction device to  $L^* = 100$  and allow this renormalization to

take care of differences in the luminance capabilities between the original and the reproduction for high  $L^*$ . However, in our experiments, we found that this type of gamut mapping, corresponding to the ISC algorithm, did not produce the best images. Hence, although limitations in high chroma and dark regions of the gamut are stressed in the search for the ultimate gamut mapping strategy, this result indicates that gamut mapping near "white" is also a significant problem, and that the common method of adjusting lightness changes may not be optimal.

When the illumination on a related color is decreased (within the range of photopic vision), neither its chroma nor saturation would change since colorfulness and brightness would both decrease proportionally. In our experiments, when mapping lightness at the high  $L^*$  region of the gamut, a decrease in lightness of a color without an adjustment of its chroma has the effect of increasing the saturation of the color. When saturation is preserved, a decrease in lightness is accompanied by a proportional decrease in chroma. This latter situation is less discordant with natural phenomena and might explain its better performance.

The mapping algorithms that performed the best for lightness mapping were those that scaled the image, thereby preserving variations in lightness. It is well known that variations in lightness are important cues for the perception of shape and scene geometry. In comparison, the clipping algorithm performed best for chroma mapping, indicating that pictorial variations in chroma are not salient cues for the perception of shape. Our experiments would suggest the following order of implementation of the best performing algorithms: i) lightness mapping at the top of the gamut; ii) lightness mapping at the bottom of the gamut; iii) chroma clipping.

The research presented here is the first step in developing automated gamut mapping algorithms. We plan to apply the results of this study to more realistic gamuts in which chroma and lightness need to be mapped simultaneously as well as with the use of multicolored images. Also of interest is the question of gamut expansion when the reproduction device has a gamut larger than the original device.

#### ACKNOWLEDGMENT

The authors thank R. Balasubramanian of Xerox Corp. for his valuable suggestions and helpful discussions, and the subjects for volunteering their time and patience.

#### REFERENCES

- [1] M. C. Stone, W. B. Cowan, and J. C. Beatty, "Color gamut mapping and the printing of digital image," *ACM Trans. Graphics*, vol. 7, pp. 249–292, Oct. 1988.
- [2] M. Wolski, J. P. Allebach, and C. A. Bouman, "Gamut mapping: Squeezing the most out of your color system," in *Proc. IS&T/SID's 2nd Color Imaging Conf.: Color Science, Systems, and Applications*, Nov. 1994, pp. 89–92.

- [3] A. S. Glassner, *Principles of Digital Image Synthesis*. San Francisco, CA: Morgan Kaufmann, 1995.
- [4] M. D. Fairchild, "Some hidden requirements for device-independent color imaging," *SID 94 Digest*, pp. 865–868, 1994.
- [5] T. K. Kim, R. S. Berns, and M. D. Fairchild, "Comparing appearance models using pictorial images," in *Proc. IS&T/SID's Color Imaging Conf.: Transforms and Transportability of Color*, Nov. 1993, pp. 72–77.
- [6] M. D. Fairchild, "Visual evaluation and evolution of the RLAB color space," in *Proc. IS&T/SID's 2nd Color Imaging Conf.: Color Science, Systems and Applications*, Nov. 1994, pp. 9–13.
- [7] K. M. Braun and M. D. Fairchild, "Evaluation of five color-appearance transforms across changes in viewing conditions and media," in *Proc. IS&T/SID's 3rd Color Imaging Conf.: Color Science, Systems and Applications*, Nov. 1995, pp. 93–96.
- [8] L. W. MacDonald, "Gamut mapping in perceptual color space," in *Proc. IS&T/SID's Color Imaging Conf.: Transforms and Transportability of Color*, Nov. 1993, pp. 193–196.
- [9] R. S. Gentile, E. Walowitz and J. P. Allebach, "A comparison of techniques for color gamut mismatch compensation," *J. Imag. Tech.*, vol. 16, pp. 176–181, Oct. 1990.
- [10] T. Hoshino and R. S. Berns, "Color gamut mapping techniques for color hard copy images," in *Proc. SPIE*, vol. 1909, pp. 152–165, 1993.
- [11] W. E. Wallace and M. C. Stone, "Gamut mapping computer generated imagery," in *Proc. SPIE*, vol. 1460, pp. 20–28, 1991.
- [12] E. G. Pariser, "An investigation of color gamut reduction techniques," in *Proc. IS&T Symp. Electronic Prepress Technology—Color Printing*, 1991, pp. 105–107.
- [13] R. S. Berns, R. J. Motta, and M. E. Gorzynski, "CRT colorimetry. Part I: Theory and practice," *Color Res. Appl.*, vol. 18, pp. 299–314, Oct. 1993.
- [14] L. L. Thurstone, "A law of comparative judgment," *Psycholog. Rev.*, vol. 34, pp. 273–286.
- [15] H. A. David, *The Method of Paired Comparison*. New York: Hafner, 1963.
- [16] A. E. Maxwell, "The logistic transformation in the analysis of paired-comparison data," *Br. J. Math. Stat. Psychol.*, vol. 27, pp. 62–71, 1974.
- [17] G. Keppel, *Design and Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1973, pp. 144–166.
- [18] P. Hung and R. S. Berns, "Determination of constant hue loci for a CRT gamut and their predictions using color appearance spaces," *Color Res. Appl.*, vol. 20, pp. 285–295, Oct. 1995.
- [19] K. Spaulding, R. N. Ellson, and J. R. Sullivan, "UltraColor: A new gamut-mapping strategy," *Proc. SPIE*, vol. 2414, pp. 61–68, 1995.



**Ethan D. Montag** received the B.A. in psychology from the University of Pennsylvania, Philadelphia, in 1985, and the Ph.D. in experimental psychology in 1991 from the University of California, San Diego, working in color vision.

He was a post-doctoral fellow at the Center for Visual Science, University of Rochester, Rochester, NY, from 1991 to 1994, where he worked on the interactions between form and color under a National Eye Institute-NIH National Research Service Award Fellowship. In 1994, he started a two-year post-doctoral position at the Munsell Color Science Laboratory, Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, NY, where he worked on color gamut mapping. In 1996, he was appointed Research Assistant Professor at the Munsell Color Science Laboratory. His current interests include color gamut mapping, color vision, color tolerance measurement, and the use of color in information display.

Dr. Montag is a member of the Association for Research in Vision and Ophthalmology, the Optical Society of America, the International Research Group on Color Vision Deficiencies, and the Society for Imaging Science and Technology.

**Mark D. Fairchild**, for photograph and biography, see this issue, p. 900.