Reprinted from
Journal of Perceptual Imaging® 3(2): 020401-1–020401-8, 2020.

Using Gaussian Spectra to Derive a Hue-linear Color Space

Luke Hellwig and Mark D. Fairchild▲

Munsell Color Science Laboratory, Rochester Institute of Technology, Rochester, New York 14623
E-mail: lfh9787@rit.edu

Abstract. *A new color space, $I_GP_GT_G$, was developed. $I_GP_GT_G$ uses the same structure as IPT, an established hue-uniform color space utilized in gamut mapping applications. While IPT was fit to visual data on the perceived hue, $I_GP_GT_G$ was optimized based on evidence linking the peak wavelength of Gaussian-shaped light spectra to their perceived hues. The performance of $I_GP_GT_G$ on perceived hue data was compared to the performance of other established color spaces. Additionally, an experiment was run to directly compare the hue linearity of $I_GP_GT_G$ with those of other color spaces by using Case V of Thurstone's law of comparative judgment to generate hue-linearity scales. $I_GP_GT_G$ performed well in this experiment but poorly on extant visual data. The mixed results indicate that it is possible to derive a moderately hue-linear color space without visual data.*
[DOI: 10.2352/J.Percept.Imaging.2020.3.2.020401]

1. INTRODUCTION

Gamut mapping is an important process by which colors specified within one device's color gamut are mapped to fit within or fill out the color gamut of a second device. Typically, gamut mapping involves adjusting the chroma and lightness of colors while holding hue constant because changes in hue are generally more objectionable [17]. Thus, it is important for the color space where gamut mapping occurs to accurately predict perceived hue. The degree to which a color space maps stimuli of the same perceived hue to the same hue angle represents its hue linearity. In a hue-linear color space, it is simple to adjust the chroma of a color without changing its perceived hue by adjusting the color along the line that passes through the achromatic origin (an iso-hue line). In a color space with poor hue linearity, changing chroma or lightness also changes the perceived hue although this crosstalk can be reduced with the use of look-up tables [7]. Thus, performing gamut mapping in a color space with poor hue linearity could lead to hue differences between input and output devices. Thus, hue linearity is an important feature of color spaces used for gamut mapping.

Although CIELAB is a commonly used uniform color space, significant hue non-linearity has been documented in CIELAB for purple-blue (PB) colors [1, 17]. This non-linearity—and the demand for a color space that better predicts perceived hue—has led to the development of color spaces with improved hue linearity such as IPT in 1998 [2] and $J_za_zb_z$ in 2017 [18]. These color spaces consist of a matrix transform from CIE XYZ coordinates to LMS pseudo-cone space, a non-linear compression in cone space, and a matrix transform from LMS to an opponent coordinate system, where one dimension (I in IPT) corresponds to lightness and the other two dimensions correspond approximately to red versus green (P) and yellow versus blue (T), similar to CIELAB's coordinate system. The hue of stimuli in IPT is quantified by its polar angle in the PT plane relative to the +P axis.

In general, these color spaces are derived by mathematically fitting transformations to experimental visual data. For example, IPT improved on the hue linearity of CIELAB and CIECAM97s by fitting the transformation from CIE XYZ to IPT to constant hue datasets from Hung and Berns [9] and Ebner and Fairchild [3] in addition to measurements of the Munsell color order system [10]. IPT has proven to be successful in subsequent experiments [19], and it has formed the basis for Dolby's IC_TC_P color space [20].

Experimental results from Mizokami et al. [15] opened up a new path to the development of a hue-linear color space. The paper connected spectral properties to perceived hue. Specifically, the researchers found that Gaussian-shaped light spectra of varying bandwidth but the same peak wavelength were perceived by observers to have the same hue [15]. This result would suggest non-linear compensation in the neural coding of color to account for the fact that single-peak-wavelength Gaussian spectra would have different cone excitation ratios at different bandwidths [15]. For neural coding to be connected to Gaussian spectra, though, it would be required that Gaussian spectra serve as an effective representation of the stimuli that we encounter in natural scenes. Further work by this research group explored whether Gaussian spectra "could accurately approximate natural spectra with a small number of parameters" [14]. They found that Gaussian spectra performed similarly to linear models with the same number of parameters [14]. In conjunction, these results establish the plausibility of using features of Gaussian spectra to optimize a color space for hue linearity as opposed to fitting the transformations to visual data.

Work by Mirzaei and Funt [12, 13] related to categorizing object colors added more evidence to the efficacy of

▲ IS&T Member.

Received Feb. 21, 2020; accepted for publication July 29, 2020; published online Sep. 16, 2020. Associate Editor: Bernice E. Rogowitz.
2575-8144/2020/3(2)/020401/8/$00.00

using the peak wavelength of Gaussian spectra as a hue predictor. Their approach involves finding a wraparound Gaussian spectrum that is metameric to a stimulus under a specified illumination and using the peak wavelength of the Gaussian spectrum as a direct hue descriptor [13]. (A wraparound Gaussian spectrum is where a function at 780 nm continues at 380 nm up to the wavelength complementary to the peak wavelength. This is one approach that allows the entire chromaticity diagram to be represented by Gaussian functions, including purple and magenta colors, which requires preferential stimulation of S and L cones versus M cones.) The researchers used the peak wavelength hue descriptor to train a genetic algorithm to optimize hue boundaries for Munsell colors [13]. They found that their Gaussian-based system worked better than CIECAM02 for this task. These results agreed with other qualitative measurements [13].

However, there are several issues with using peak wavelength as a direct descriptor of hue. Narrowband Gaussian spectra with peak wavelength at either end of the visible range (e.g., greater than 700 nm) have identical or very similar chromaticities even as the peak wavelength varies. This singularity would lead to ambiguity in the method and would cause visually identical stimuli to be mapped to different hue bins. Additionally, the Mirzaei & Funt method is tailored to being an illuminant-invariant descriptor of object colors and was not developed for use as a color space to perform gamut mapping. As such, the research did not include the standard quantitative assessment of hue linearity that could be directly compared with other color spaces: a measurement of the angular spread of constant hue loci [19].

An alternative approach explored here is to use Gaussian spectra to generate predicted constant hue loci, which can then be used to optimize a color space transform where points on a constant hue locus map to the same hue angle. This approach is similar to the derivation processes for IPT [2] and $J_z a_z b_z$ [18] although where the constant hue loci for those color spaces were the result of visual experiments, these constant hue loci are generated from Gaussian spectra of a single peak wavelength. The color space derived in this article following the above method is referred to as $I_G P_G T_G$ given that the form of the transform was chosen to match IPT's definition. This allowed a direct comparison of the efficacy of the Gaussian-based hue loci for generating a hue-linear color space with the traditional method of hue loci based on visual data.

The use of a color space to test the Gaussian hue hypothesis posited by previous papers [13–15] addresses the shortcomings of previous attempts described above. Because Gaussian spectra are used as parameters in the optimization process as opposed to being used as direct hue descriptors, problematic singularities for long-wavelength narrowband spectra are avoided. Additionally, the structure of the color space allows for direct use in gamut mapping applications. This structure also requires less computation time than that of the Mirzaei & Funt method, which requires a Gaussian metamer to be calculated for each stimulus [13].

Two experiments were conducted to compare the hue linearity of $I_G P_G T_G$ with that of other established color spaces—CIELAB, CAM16-UCS, and IPT—and the Munsell color order system. The experimental results indicated that $I_G P_G T_G$ matched or exceeded the hue linearity of the comparison color spaces.

Existing visual data related to perceived hue were transformed into $I_G P_G T_G$ as an additional validation of the color space's hue linearity. In this case, $I_G P_G T_G$ performed much more poorly than the comparison color spaces. The mixed results of the experiment and the visual data analysis indicate that a color space derived from Gaussian spectra is plausible but is unlikely to supplant established color spaces.

Hue linearity is one aspect of broader hue uniformity. Much research has also gone into measuring uniform hue scaling [4], which relates to the uniformity of differences between hue angles. Hue scaling uniformity is a measure of whether the difference in hue between colors with hue angle 0° and colors with hue angle 10° is the same as the difference in hue between those with hue angles 10° and 20°. Hue-linearity uniformity, as we use it here, is a measure of whether all colors with hue angle 10° have the same perceived hue. For the purpose of this article, only hue linearity was evaluated given its special importance in gamut mapping applications as discussed above.

2. COLOR SPACE DEVELOPMENT

The structure of IPT served as the basis for $I_G P_G T_G$. IPT coordinates are defined by their transformation from 1931 CIE XYZ coordinates with a D65 white point. XYZ values are first transformed to an LMS cone space using a 3-by-3 matrix transform. Non-linear compression is applied to each dimension in the LMS cone space before another 3-by-3 matrix transforms the LMS coordinates into IPT coordinates, where I represents the brightness/lightness dimension and P and T represent chromatic dimensions of perception (roughly, red versus green and yellow versus blue). For more information on the structure of IPT, see [2]. This structure was deemed appropriate for $I_G P_G T_G$ because of its invertibility, simplicity (while still providing the same number of degrees of freedom for optimization as used to derive other hue-linear color spaces), and the success of IPT as a hue-linear color space.

The first step in optimizing the transform to $I_G P_G T_G$ was to choose Gaussian-shaped spectra to be used for optimization. To fully and evenly sample the gamut of possible stimuli, the peak wavelength and the bandwidth of Gaussian spectra were optimized to fit an evenly spaced grid of coordinates in CIELAB (Figure 1). The simulated spectra were then grouped into sets of equal peak wavelength, which would be expected by our model to have the same perceived hue. Each peak-wavelength group contained at

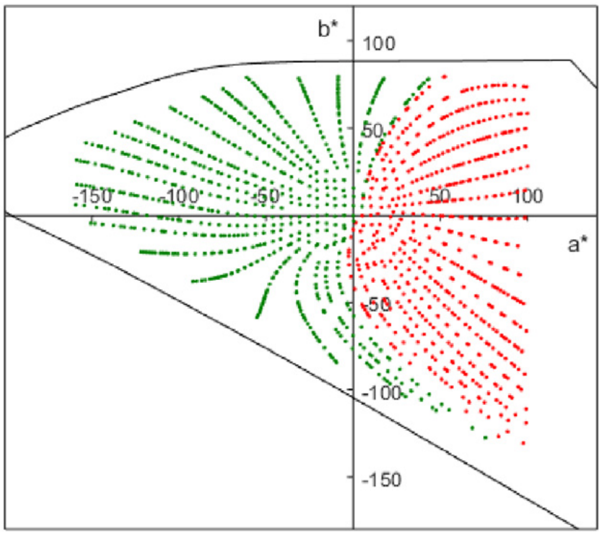

Figure 1. CIELAB a^*b^* coordinates corresponding to the single (green) and double (magenta) Gaussian spectra used for optimization. The spectral locus is shown in black. All points in this figure have $L^* = 50$.

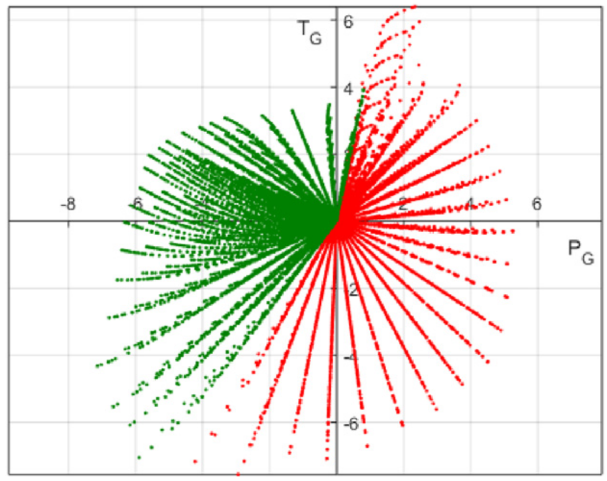

Figure 2. P_GT_G coordinates of the single (green) and double (magenta) Gaussian spectra used for optimization. Hue angle spread is not evidence of a lack of hue linearity but rather represents the limit of optimization.

least 32 spectra to ensure full chroma sampling for each group.

Gaussian spectra can be used to generate metamers to any color except colors that correspond to the stimulation of L and S cones without M cone stimulation. (Such colors are typically purple or magenta.) To represent these colors for optimization, spectra consisting of two Gaussian functions were simulated, one with a central wavelength of 380 nm and the other with a central wavelength of 700 nm. The hue of such combined Gaussian spectra can be controlled by the ratio between the bandwidths of the two Gaussian functions. A spectrum's chroma can be controlled by the overall bandwidth of the function. Since the two Gaussian functions overlap, the greater of the two functions at each wavelength was taken as the radiance of the spectrum at that wavelength. Double Gaussian spectra were simulated to fit an evenly spaced grid of coordinates in CIELAB (Fig. 1). The stimuli were then grouped based on the ratio between the 380 nm and 700 nm functions' bandwidths with at least 32 spectra in each group. (The bandwidth ratios within each group were set equal.)

For each unique spectrum, the total radiance was adjusted to match nine luminance levels corresponding to Munsell values 1–9 for a D65 white point at 100 cd/m^2. CIE XYZ coordinates were calculated for each spectrum using the 1931 Standard Observer with 1-nm sampling.

The transform from XYZ to $I_GP_GT_G$ was then non-linearly optimized in MATLAB to minimize the root-mean-square (RMS) hue angle difference from the mean hue angle for stimuli with the same peak wavelength (or the same bandwidth ratio for double Gaussian spectra) by using IPT as the starting point for optimization. A second optimization objective was to minimize the difference between the I_G and I (from IPT) values for each stimulus. Using multiple luminance levels ensured that the optimized formula could handle a variety of luminances. However, this method also involved the assumption that Gaussian spectra of the same peak wavelength but different luminance levels have the same hue. The Bezold–Brücke hue shift effect suggests that this assumption may not hold true for nearly monochromatic spectra [6]. Additionally, to avoid the trivial solution to the optimization problem where all peak wavelengths are mapped to the same hue angle, a third optimization objective was added. The objective was for the transform to evenly space the hue angle of 40 Munsell colors with value 5 and chroma 5. The weight of this objective was minimized as much as possible without causing optimization to revert to the trivial solution. The optimized transform (Eqs. (1)–(5)) had mean and maximum hue angle standard deviations of 2.3° and 14.1°, respectively, for the 31 sets of Gaussian spectra used in optimization (Figure 2). There was a 8.5% root-mean-square difference between I_G and I values for the Gaussian optimization spectra. It should be noted that the hue angle spread for these points does not directly represent a failure of hue linearity. Rather, the spread present for certain hues in Fig. 2 merely represents the limits of the optimization process to map Gaussian spectra of equal peak wavelength to the same hue angle. Although it is plausible that including additional parameters in the XYZ to $I_GP_GT_G$ transform would improve the optimization, there was circumstantial evidence that the improvement would have been minimal as the optimization function did not use all of the degrees of freedom provided to it. Additionally, changing the form of the transform would reduce our ability to directly compare this method of fitting to the fitting process used for IPT.

Like IPT, $I_GP_GT_G$ assumes a D65 white point. For stimuli with a different white point, it is recommended to transform the stimuli's XYZ values using CAT16, the chromatic adaptation transform embedded in CAM16, before converting to $I_GP_GT_G$ [11]. This is done by converting the XYZ values to the LMS cone space and then applying a von Kries type adaptation, where the LMS values of the stimuli are scaled to the LMS values of the test white point. The scaled LMS coordinates are then transformed back into XYZ, where they will have a D65 white point [11].

Once the stimuli have been transformed to a D65 white point, the optimized $I_GP_GT_G$ transform is

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix}_{\text{lin}} = \begin{bmatrix} 2.968 & 2.741 & -0.649 \\ 1.237 & 5.969 & -0.173 \\ -0.318 & 0.387 & 2.311 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{1}$$

$$L = \left(\frac{L_{\text{lin}}}{18.36}\right)^{0.427} \tag{2}$$

$$M = \left(\frac{M_{\text{lin}}}{21.46}\right)^{0.427} \tag{3}$$

$$S = \left(\frac{S_{\text{lin}}}{19435}\right)^{0.427} \tag{4}$$

$$\begin{bmatrix} I_G \\ P_G \\ T_G \end{bmatrix} = \begin{bmatrix} 0.117 & 1.464 & 0.130 \\ 8.285 & -8.361 & 21.40 \\ -1.208 & 2.412 & -36.53 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix}. \tag{5}$$

3. COLOR SPACE VERIFICATION

3.1 *Visual Data on Hue Uniformity*

One method to compare color spaces' ability to predict perceived hue is to use existing datasets. The Munsell color order system [1, 10], the NCS system [1, 8], and the Hung and Berns dataset [9] all specify colors that have the same perceived hue according to visual experiments. Analyzing how color spaces handle these datasets thus provides insight into their hue linearity.

Colorimetric coordinates of 40 Munsell constant hue loci, 24 NCS constant hue loci, and 12 hue loci from the Hung and Berns dataset were transformed into $I_GP_GT_G$, IPT [2], CIELAB (using a D65 white point) [1], and CAM16-UCS (using a 100 cd/m^2 D65 white point with a dim surround) [11]. In an ideal hue-linear color space, all of the colorimetric coordinates for a single Munsell/NCS/Hung and Berns hue locus would map to a single hue angle. Thus, the RMS difference from the mean hue angle for colors within one hue locus is a measure of how well these four color spaces perform in their prediction of hue. This is a standard metric used to quantify the hue linearity of color spaces [19]. The results of this analysis are shown in Figures 3–5.

$I_GP_GT_G$ performs more poorly than the three comparison color spaces for both the Munsell and NCS datasets with especially poor performance in Munsell. However, $I_GP_GT_G$ was able to match the performance of CIELAB for the Hung and Berns data. Nonetheless, the results indicate that $I_GP_GT_G$ has worse hue linearity than the comparison color spaces.

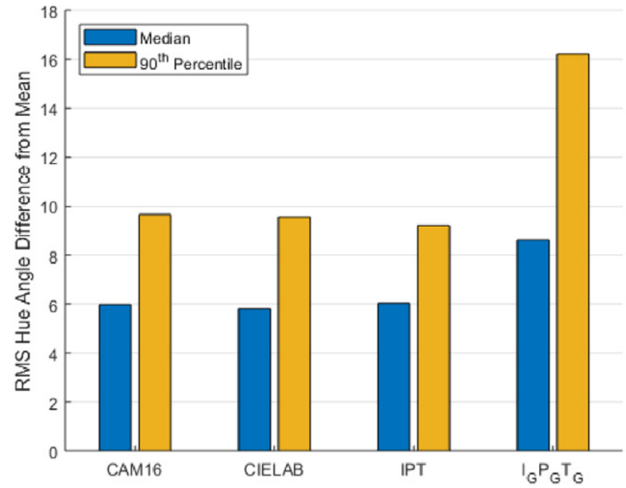

Figure 3. Median and 90th percentile root-mean-square hue angle difference from mean hue angle over 40 Munsell hues for four color spaces. Lower values indicate greater hue linearity according to the Munsell data.

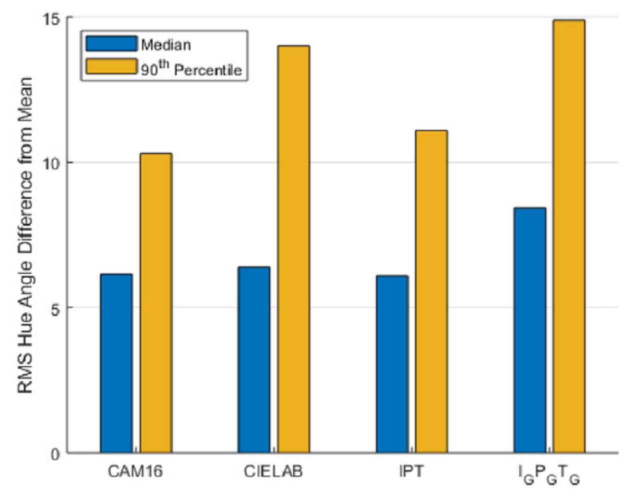

Figure 4. Median and 90th percentile root-mean-square hue angle difference from mean hue angle over 24 NCS hues for four color spaces.

3.2 *Experiment 1*

3.2.1 *Motivation*

A psychophysical experiment was designed and implemented to directly compare the hue linearity of $I_GP_GT_G$ to three color spaces and a color order system all known for hue linearity: IPT, CIELAB, CAM16, and Munsell.

3.2.2 *Methods*

A paired comparison stimulus presentation was used to compare color space predictions for 11 hues. Each stimulus consisted of a two-by-three grid of color patches with the same hue angle in the test color space (or the same Munsell hue for the Munsell system) (similar to Figure 6). Stimuli were generated by transforming six Munsell colors with the

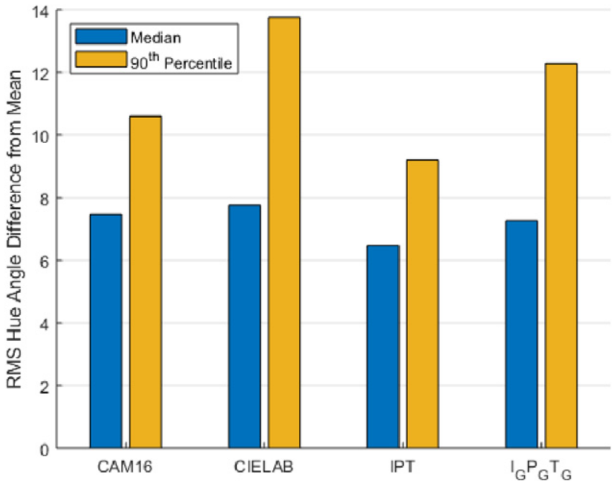

Figure 5. Median and 90th percentile root-mean-square hue angle difference from mean hue angle over 12 Hung and Berns hue loci for four color spaces.

same hue—and a range of chromas and values—into the test color space. Each color's hue angle was then set equal to the mean hue angle in that color space for the set. This provided similar variations in lightness and chroma between sets of stimuli for different color spaces. Eleven hues were tested, corresponding to ten Munsell hues spread evenly throughout the hue circle plus 10PB, since this is a known hue problem for hue linearity in CIELAB [9].

Stimuli were presented on a calibrated Eizo ColorEdge CG279X LCD monitor with DCI primaries and a D65 white point at 400 cd/m^2. The background was achromatic and had a CIELAB lightness of 50. For each stimulus presentation, observers were shown two sets of patches for the same hue. They were asked to determine "which set of patches is most uniform in hue." (Uniformity of perceived hue among patches of the same hue angle is a direct measure of the hue linearity of the color space.) With 11 hues and 5 color spaces, this led to 110 paired comparisons per observer. Paired comparisons were presented in random order, and the arrangement of each stimulus was randomized. Fifteen observers (7 females and 8 males) with normal color vision participated in the experiment.

3.2.3 *Results*

Observer responses were analyzed using Case V of Thurstone's law of comparative judgment [5]. Individual responses were compiled into matrices for each hue, where each matrix entry was the number of times that the color space corresponding to the column was chosen over the color space corresponding to the row. Converting these values to z-scores and then averaging along each column lead to the color space's scale value on an interval scale of hue linearity for each hue [5]. These scale values are relative and are z-scores scaled in units of standard deviations along the psychological dimension being tested. A scale value of zero indicates that the color space has average hue linearity for that hue. A color space with a value of one is one standard deviation better than a color space with scale value zero for that hue. It is assumed that this assessment of hue linearity is unidimensional and that the uncertainty in psychometric judgments is normally distributed with equal dispersion. The assumption of equal variances required by Case V was confirmed using a chi-square test following the method described by Engeldrum [5].

Ninety-five percent confidence intervals for the scale values were estimated using a formula developed by Montag [16]. The Montag formula is based on a Monte Carlo simulation, and it uses the number of stimuli and the number of observations to estimate the uncertainty in scale values calculated using Case V of Thurstone's law of comparative judgment. The results are shown in Figure 7. For seven of the hues, there was no significant difference in hue uniformity for the five color spaces. For 5P, CIELAB and CAM16 performed slightly better than IPT and Munsell. For 5YR, $I_GP_GT_G$ performed significantly worse than the other four color spaces. For 5PB and 10PB, $I_GP_GT_G$ and IPT performed significantly better than the other three color spaces, with CIELAB performing worst for 10PB.

3.3 *Experiment 2*

3.3.1 *Motivation*

A second experiment was performed to further investigate the poor performance of $I_GP_GT_G$ for the 5YR hue. This was also a confirmation that an improvement to the transform used to adjust the white point of the optimization spectra did not significantly change $I_GP_GT_G$'s performance.

3.3.2 *Methods*

As Gaussian spectra increase in bandwidth, their chromaticity approaches that of Illuminant E, the equal energy spectrum. Thus, the Gaussian spectra used to optimize the matrix transform would naturally have a white point of Illuminant E had we not transformed the white point to D65. Prior to performing the second experiment, we switched from using a von Kries type transform in XYZ space to using CAT16, the chromatic adaptation transform embedded in CAM16 [11]. The $I_GP_GT_G$ definition was then re-optimized, resulting in Eqs. (1)–(5).

The second experiment repeated the methodology of the original experiment. However, only the three hues with significant results in Experiment 1—5YR, 5PB, and 10PB—were used. The stimuli for 5YR in Experiment 1 contained both orange and brown patches. Observers reported that it was difficult to judge hue across this color name boundary. Therefore, for Experiment 2, 5YR was split into two sets of stimuli: orange and brown. Additionally, the process by which stimuli were generated for this hue was modified to reduce the overall color difference between different sets of patches. The hue angle in the test color space was still constant within each set of patches.

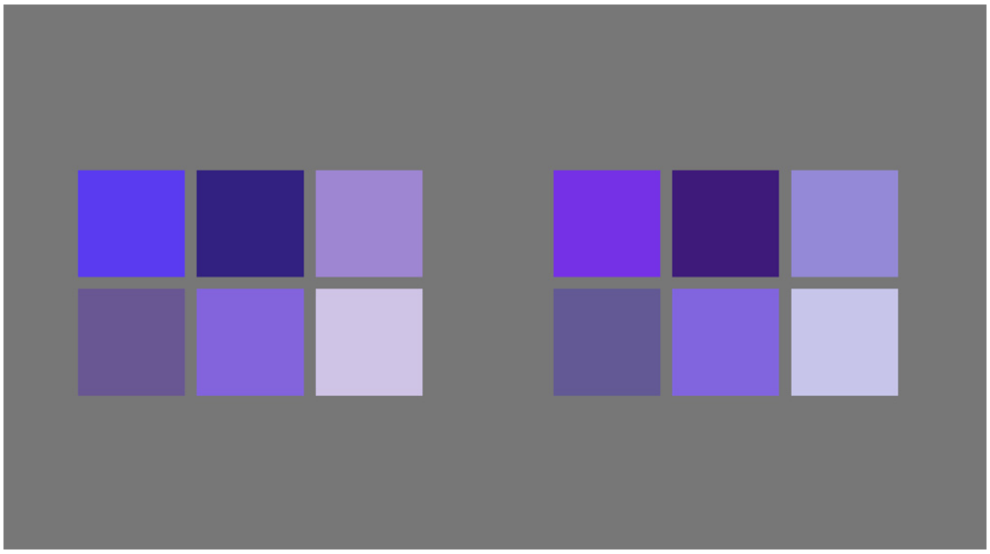

Figure 6. GUI design for Experiment 2.

Seventeen color normal observers (8 females and 9 males) completed the 40 paired comparisons for Experiment 2 using the same physical parameters as those in Experiment 1. The only other modification was a thin gray border added around each patch to reduce any effect of simultaneous contrast (Fig. 6).

3.3.3 *Results*

Observer responses were again analyzed using Case V of Thurstone's law of comparative judgment. The results for the four scales are shown in Figure 8. Interestingly, $I_GP_GT_G$ was now judged to have similar hue uniformity as the other color spaces for both the orange and brown stimuli from 5YR, which is a significant improvement compared with the results from Experiment 1. This result suggests that the poor performance of $I_GP_GT_G$ on 5YR in Experiment 1 was due to an uncontrolled overall color difference for the 5YR $I_GP_GT_G$ patches from the patches for other color spaces in Experiment 1 and the difficulty in judging hue across the orange/brown color name boundary. $I_GP_GT_G$ performed similarly well for the 5PB and 10PB hues as it did in Experiment 1, indicating that using the new white point correction before optimization did not affect the color space's hue prediction. Compared with its performance in Experiment 1, CAM16 performed better in 5PB and worse in 10PB. Munsell was judged to have higher hue uniformity for 10PB in Experiment 2 than it had in Experiment 1.

4. DISCUSSION

Although extant visual data suggested that $I_GP_GT_G$ has worse hue linearity than CAM16, CIELAB, and IPT, our visual experiment indicated that $I_GP_GT_G$ performs equally well as these color spaces on this metric. It should be noted that slightly different versions of the $I_GP_GT_G$ formula were used for the first psychophysical experiment and the comparison with extant visual data (which is the formula presented in Eqs. (1)–(5)). Originally, $I_GP_GT_G$ was optimized using lower-chroma Gaussian spectra to facilitate the optimization process. However, this led to irregularities in how high-chroma points were mapped at the interface between the single and double Gaussian spectra at high chromas. This issue was addressed through the process described in Section 2. Because the main improvements to $I_GP_GT_G$ were made beyond the gamut of colors used in the experiment, we would expect the final $I_GP_GT_G$ formula to present similar performance in the experiments. This expectation was supported by post hoc statistical analysis.

The conflict between experimental results and statistical analysis of extant visual data suggests that deriving a color space based on spectral properties alone is plausible. However, this color space fails to match the hue linearity of existing color spaces. Typically, uniform color spaces are fit to visual data alone. Although this approach has been successful, developing processes by which we can derive color spaces using first principles is attractive because such spaces would not be directly dependent on the accuracy of experimental data and limited to the gamut covered by such data. However, from a performance perspective, this work did not find compelling data to support the use of $I_GP_GT_G$ over the well-established IPT color space. Furthermore, although a direct correspondence between spectral properties of Gaussian stimuli and the perceived hue has not yet been proven, the results of this article add to the growing body of literature exploring potential connections [13–15].

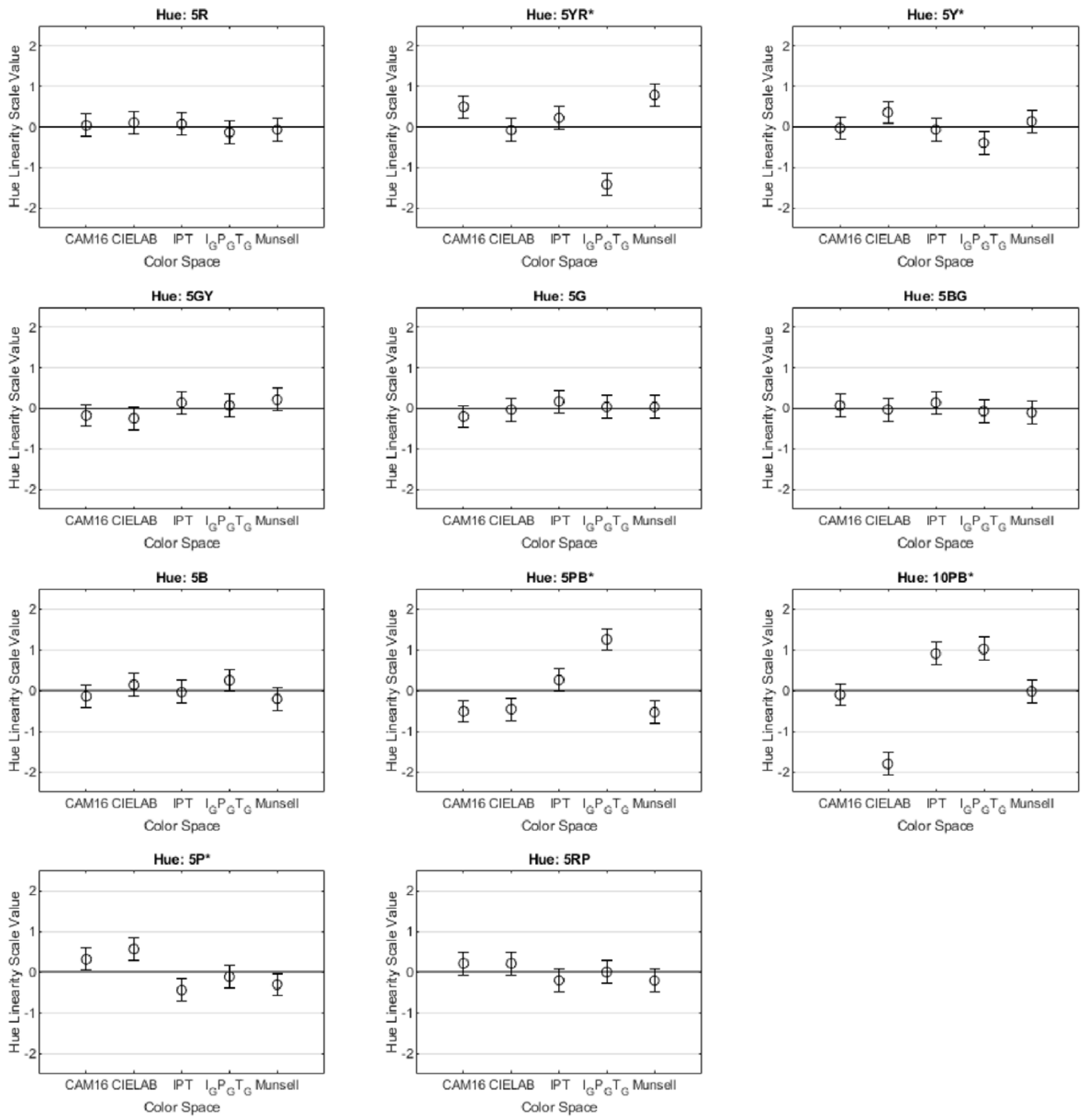

Figure 7. Hue-linearity scales for the 11 hues tested in Experiment 1, which were generated using Case V of Thurstone's law of comparative judgment. A scale value of zero indicates that the color space has average hue linearity for that hue. A color space with a value of one is one standard deviation better than a color space with scale value zero for that hue. 95% confidence intervals were calculated using a method developed by Montag [16]. Statistically significant scale values had confidence intervals that did not include zero (indicated by *). The results indicate that $I_GP_GT_G$ was perceived to have good hue linearity except for hues 5Y and 5YR.

5. CONCLUSION

A new color space, $I_GP_GT_G$, was developed using the premise that Gaussian-shaped light spectra of the same peak wavelength have the same perceived hue regardless of bandwidth. $I_GP_GT_G$ was defined by a transform from CIE XYZ coordinates using the same structure as the well-established hue-linear IPT color space. The transform was non-linearly optimized in MATLAB to minimize the deviation in hue angle for Gaussian spectra of a single peak wavelength. The hue linearity of $I_GP_GT_G$ was then assessed using two methods. First, extant visual data on the perceived hue from the Munsell and NCS color order systems and the Hung and Berns dataset were transformed to $I_GP_GT_G$. This data analysis indicated that $I_GP_GT_G$ has worse hue linearity than CAM16, CIELAB, and IPT. Second, two visual experiments were performed in which observers were asked to directly assess the hue linearity of these color spaces. In this case, $I_GP_GT_G$ matched or exceeded the performance of CAM16, CIELAB, IPT, and the Munsell system. These ambiguous results show that it is at least plausible to derive a hue-linear color space from first principles without the use of visual data in the derivation process.

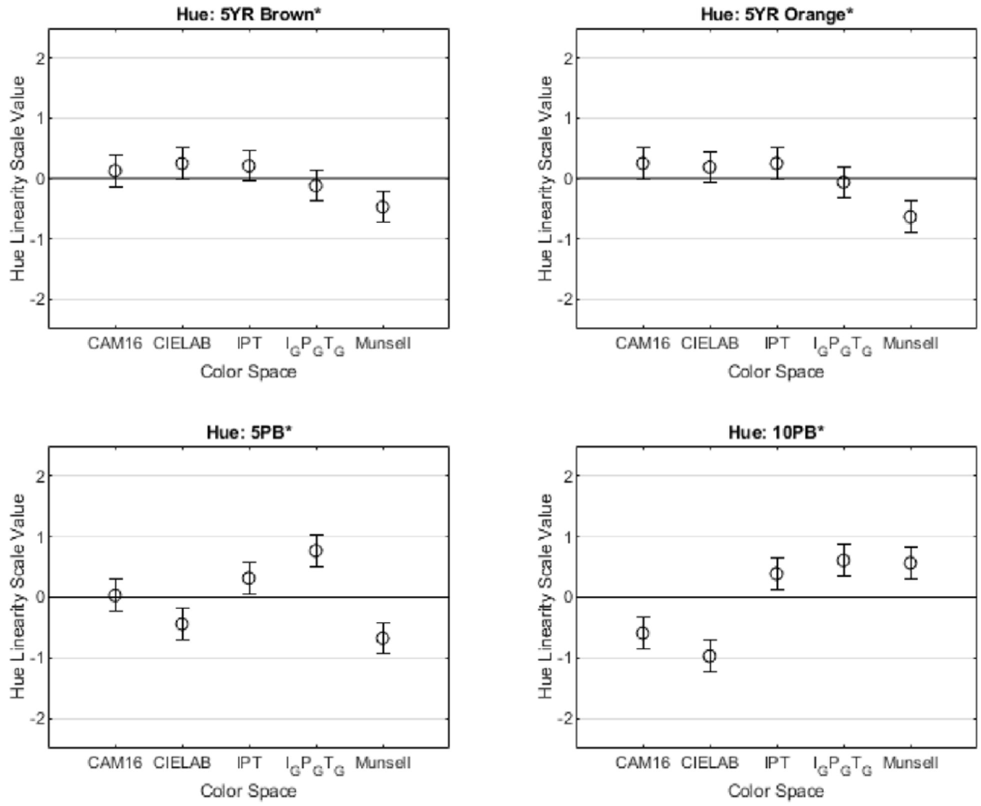

Figure 8. Hue-linearity scales for the four stimuli sets tested in Experiment 2, which were generated using Case V of Thurstone's law of comparative judgment. A scale value of zero indicates that the color space has average hue linearity for that hue. A color space with a value of one is one standard deviation better than a color space with scale value zero for that hue. 95% confidence intervals were calculated using a method developed by Montag [16]. Statistically significant scale values had confidence intervals that did not include zero (indicated by *). The results confirm $I_GP_GT_G$'s good hue linearity for 5PB and 10PB and suggest that the poor result for 5YR and 5Y in Experiment 1 could have been due to experimental design issues that were corrected for Experiment 2.

REFERENCES

1 R. Berns, *Billmeyer and Saltzman's Principles of Color Technology*, 4th ed. (Wiley, Hoboken, NJ, 2019).

2 F. Ebner and M. D. Fairchild, "Development and testing of a color space (IPT) with improved hue uniformity," *Proc. IS&T/SID CIC6: Sixth Color Imaging Conf.: Color Science, Systems and Applications* (IS&T, Springfield, VA, 1998), pp. 8–13.

3 F. Ebner and M. D. Fairchild, "Finding constant hue surfaces in color space," Proc. SPIE **3300** (1998).

4 K. J. Emery, V. J. Volbrect, D. H. Peterzell, and M. A. Webster, "Variations in normal color vision. VII. Relationships between color naming and hue scaling," Vision Res. **141**, 66 (2017).

5 P. G. Engeldrum, "Indirect interval scaling—case V and paired comparison," *Psychometric Scaling* (Imcotek Press, Winchester, MA, 2000), p. 93.

6 M. D. Fairchild, "Color appearance phenomena," *Color Appearance Models*, 2nd ed. (Wiley, Hoboken, NJ, 2005), p. 111.

7 G. Graun and M. D. Fairchild, "Color gamut mapping in a hue-linearized CIELAB color space," *Proc. IS&T/SID CIC6: Sixth Color Imaging Conf.: Color Science, Systems and Applications* (IS&T, Springfield, VA, 1998), pp. 163–168.

8 A. Hård, L. Sivik, and G. Tonnquist, "NCSm natural color system—from concept to research and applications. Part I," Color Res. Appl. **21**, 180 (1996).

9 P.-C. Hung and R. S. Berns, "Determination of constant Hue Loci for a CRT gamut and their predictions using color appearance spaces," Color Res. Appl. **20**, 285 (1995).

10 K. L. Kelly, K. S. Gibson, and D. Nickerson, "Tristimulus specification of the Munsell book of color from spectrophotometric measurements," J. Opt. Soc. Amer. **33**, 355 (1943).

11 C. Li, Z. Li, Z. Wang, Y. Xu, M. R. Luo, G. Cui, M. Melgosa, M. Brill, and M. Pointer, "Comprehensive color solutions: CAM16, CAT16, and CAM16UCS," Color Res. Appl. **42**, 703 (2017).

12 H. Mirzaei and B. Funt, "A robust hue descriptor," *Proc. IS&T CIC21: Twenty-first Color and Imaging Conf.* (IS&T, Springfield, VA, 2013), pp. 75–78.

13 H. Mirzaei and B. Funt, "Gaussian-based hue descriptors," IEEE Trans. Pattern Anal. Mach. Intell. **37**, 2441 (2015).

14 Y. Mizokami and M. A. Webster, "Are Gaussian spectra a viable perceptual assumption in color appearance?," J. Opt. Soc. Amer. A **29**, A10 (2012).

15 Y. Mizokami, J. S. Werner, M. A. Crognale, and M. A. Webster, "Nonlinearities in color coding: compensating color appearance for the eye's spectral sensitivity," J. Vision **6**, 996 (2006).

16 E. D. Montag, "Empirical formula for creating error bars for the method of paired comparison," JEI Lett. **15**, 010502-1 (2006).

17 E. Reinhard, E. A. Khan, A. O. Akyüz, and G. Johnson, "Image properties and image display," *Color Imaging: Fundamentals and Applications* (A K Peters, Ltd., Wellesley, MA, 2008), p. 805.

18 M. Safdar, G. Cui, Y. J. Kim, and M. R. Luo, "Perceptually uniform color space for image signals including high dynamic range and wide gamut," Opt. Express **25**, 15131 (2017).

19 B. Zhao and M. R. Luo, "Hue linearity of color spaces for wide color gamut and high dynamic range media," J. Opt. Soc. Amer. A **37**, 865 (2020).

20 "Report ITU-R BT.2390-0" (ITU-R, Geneva 2016).