The properties of multiple CMF determinations using alternative primary sets Part II: A data unification methodology

Document Version
Accepted author manuscript

Link to publication record in Manchester Research Explorer

Citation for published version (APA):

Published in:
Color Research and Application

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.

Download date:27. Jan. 2020
The Properties of Multiple CMF Determinations Using Alternative Primary Sets Part II: A Data Unification Methodology

D. P. Oulton

UMIST, Manchester M60 1QD, United Kingdom

Received 7 March 2003; revised 17 December 2003; accepted 4 March 2004

Abstract: The spectral power values representing Thornton’s “alternative primary” PC, NP, and AP colour matching functions (CMF) are compared with the power values representing the 49-observer Stiles–Burch average definition. The Thornton measurements are first converted by matrix transformation into a data set expressed in terms of spectral power at the Stiles–Burch primary wavelengths. Graphs and power ratios are used to compare the definitions for two alternative matches to the same visual stimulus. A triplet of n:n spectral-power ratios (one in each dimension, R, G, and B) is used to quantify the differences between the alternative matches. The relationship between the Thornton PC and Stiles–Burch match-definitions is then found to deviate from the expected power-ratio of 1:1 after matrix transformation. The revealed relationship is an internally consistent and smooth function of matched wavelength, which has a different nonlinear characteristic in each R, G, and B dimension relative to the Stiles–Burch reference model. The “Thornton bow-tie” phenomenon is also demonstrated between a pair of maximum saturation CMF definitions made with alternative primaries. The implicit differences in neutral axis definition represented by the bow-tie diagram are linked to differences in trichromatic unit (T-unit) definition. In this case, the conventional CMF normalization process is postulated to be inaccurate at the wavelengths concerned, resulting in incompatibility between the T-unit definitions of the two primary sets being compared. The conventional N→3 T-unit definition of visual neutrality equating Illuminant S_e to a single R:G:B power ratio is extended, by adding an extra N→N mapping to the definition. The resulting N→N →3 mapping is in principle a fully determined redefinition of three-dimensional T-unit equivalence, in which many R:G:B ratios for a comprehensive set of visually neutral metamers can be mapped by N→N transformation onto the conventional single ratio. The effect of N→N mapping is to transform spectral power distributions (SPDs) into spectral effect distributions (SEDs) expressed in T-units. The SPD/SED transform, thus defined, is proposed as a method for unifying CMF determinations made with alternative primaries. The expected outcome is that after transforming SPDs by N→N mapping into SEDs the definitions for all visually matching metamers will be demonstrably interconvertible by matrix product. © 2004 Wiley Periodicals, Inc. Col Res Appl. 29, 438–450, 2004; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/col.20069

Key words: colorimetry; trichromatic units; colour matching functions; data-set unification; colorimetric vectors

INTRODUCTION

Part I of this article, subtitled “Evidence and Modeling,” describes a novel data-transformation method using a generic linearization and normalization construct. The method allows data sets to be unified that are not linearly transformable by matrix multiplication of spectral-power match definitions. Part II is intended to demonstrate in detail why the proposed functionality is necessary to unify the Thornton and Stiles–Burch data sets. A methodology rather than a numeric solution to the current Thornton/Stiles–Burch example is presented. The construction of a definitive and testable numeric model using the proposed method is regarded as premature, until additional experimental data has been accumulated, and the apparent inconsistencies between the Thornton and Stiles–Burch data sets have been investigated further and confirmed by independent experi-
The point of creating and presenting the method first (without experimental proof of correctness) is that it establishes more clearly the nature and potential meaning of those inconsistencies and aids the design of new experiments.

A two stage data-unification method is proposed in which a common three-dimensional definition of T-unit value is first established between the data sets. The method involves redefining the three-dimensional T-unit of spectral effect, as a fully determined $N \rightarrow N \rightarrow 3$ mapping from spectral power to spectral effect. The mapping is from many monochromatic $R:G:B$ spectral-power ratios established by Maxwell method matches to a visually neutral stimulus. The product of the mapping is a single three-dimensional T-unit definition that is a nominal constant of all members of the compared data-set definitions. Each alternative primary set then has a unique subdefinition of T-unit balanced equivalence calculated by matrix transformation from the CMFs of a master “reference definition” of the visual response. The reference definition (in the current exercise, that of the Stiles–Burch average observer) consists of a set of defining primaries and CMFs together with a T-unit definition in the form of a single $R:G:B$ power ratio defining a neutral balance of visual effect for the chosen primaries.

Part I analyzed the theoretical challenges represented by W. A. Thornton’s new colour matching functions (or CMFs), and sought to establish Thornton’s results within the context of conventional colorimetry. Thornton’s measured single observer definitions used in the analysis were quoted from his articles entitled “Towards a More Accurate and Extensible Colorimetry” in which important deviations from uniform additivity under the trichromatic generalization and the laws of additive colour mixing were demonstrated.

A transformed CMF match definition $P_{\lambda}$, expressed in terms of stimulus power values, is derived as the matrix product $P_{\lambda} = M_{\lambda} \times I_{\lambda}$, where $I_{\lambda}$ and $P_{\lambda}$ are three-element visual-match definitions of a common spectral stimulus at wavelength $\lambda$ and $M_{\lambda}$ is the appropriate unit-value transformation matrix formed by reference to the CMF matches at the six defining primary wavelengths of the two CMF sets. The Stiles–Burch reference definition and the transformed definition $P_{\lambda}$ can then be compared directly, because they both specify the same spectral stimulus in terms of power at the Stiles–Burch primary wavelengths.

Three sets of Thornton CMF definitions (each made using a different alternative primary set) are first converted by matrix transformation into predictions of the equivalent CMF matches using the Stiles–Burch primaries. The Thornton definitions are then compared with the Stiles–Burch experimental CMF definitions, working throughout in terms of stimulus power values. The adopted measure of difference is an $n:n$ power ratio in each of three dimensions ($R$, $G$, and $B$).

The comparison is between alternative-primary CMF definitions of visual matches to a common reference stimulus. The compared match definitions are expressed in units of stimulus power and nominally equated by matrix transformation. After transformation, the alternative-primary Thornton PC, NP, and AP definitions are redefined in terms of power at the Stiles–Burch primary wavelengths. The data transformation method is described in Part I and summarized below.

**Data Transformation and Spectral Power Value Analysis**

The hypothesis to be tested is as follows: “Each of the nine additional PC, NP, and AP wavelengths used by Thornton has an intrinsic and potentially distinct pattern of spectral power values relative to that used in the Stiles–Burch experiments in matches to the same reference stimulus”.

A transformed CMF match definition $P_{\lambda}$, expressed in terms of stimulus power values, is derived as the matrix product $P_{\lambda} = M_{\lambda} \times I_{\lambda}$, where $I_{\lambda}$ and $P_{\lambda}$ are three-element visual-match definitions of a common spectral stimulus at wavelength $\lambda$ and $M_{\lambda}$ is the appropriate unit-value transformation matrix formed by reference to the CMF matches at the six defining primary wavelengths of the two CMF sets. The Stiles–Burch reference definition and the transformed definition $P_{\lambda}$ can then be compared directly, because they both specify the same spectral stimulus in terms of power at the Stiles–Burch primary wavelengths.

Three definitions of the matrix $M_{\lambda}$ are required, expressed respectively by reference to unit value PC, NP, and AP stimuli. These source matrices are defined (at 1-nm interval precision) by linear interpolation between the 10-nm CMFs values supplied by Thornton.
The source matrices can be used to convert milliwatts of power at the Stiles–Burch wavelengths into a prediction of Thornton’s measurements using primaries at PC, NP, and AP wavelengths. The inverse transformation predicts the Stiles–Burch experimental average by expressing the AP wavelengths. The inverse transformation predicts the Thornton’s measurements using primaries at PC, NP, and AP wavelengths into a prediction of Stiles–Burch primary wavelengths. The predictions are made by tristimulus sum, scaled according to the power present in the Thornton matches at a given wavelength. A worked example of the relevant calculation is given in Part I.

Using the full set of CMF definitions, this calculation can be repeated 31 times for each of the PC, NP, and AP 10-nm interval CMFs. This yields 93 transformed spectral-triplet definitions involving 279 spectral power values. All 279 “converted-value” predictions can be plotted against their measured-value equivalents in the Stiles–Burch data published by Wyszecki and Stiles. The resulting T versus S (i.e., Thornton versus Stiles–Burch) graphs can be used to reveal the relevant relationships.

The ratio \( \partial T : \partial S \) is used extensively to analyze the differences between the predicted and measured values, where \( \partial T \) is a stimulus power value forming an element of a transformed Thornton match and \( \partial T \) is a prediction of \( \partial S \), where \( \partial S \) is the equivalent stimulus power value established experimentally by Stiles and Burch for the same match. Partial differential symbols are used, because each spectral triplet in the Thornton CMF matches, yields three simultaneous \( \partial T : \partial S \) ratios [one in each dimension (\( R, G, \) and \( B \)) of the Stiles–Burch definition]. They are, however, simultaneous products rather than true partial differentials of the overall relationship.

The relationship between the data sets is revealed by spectral power ratios \( \partial T : \partial S \) in T versus S graphs. For example, if systematic nonlinearity is present that differs characteristically from primary to primary, it will then become apparent as \( \partial T : \partial S \) variations that are a function of Stiles–Burch stimulus values in T versus S plots.

- If the spectral-power ratio used to normalize the Stiles–Burch CMFs is a constant that is common to all four CMF determinations, all the derived \( \partial T : \partial S \) ratios will also have a constant value of 1:1 to the level of accuracy represented by the measured data. A constant 1:1 \( \partial T : \partial S \) ratio indicates that the Thornton and Stiles–Burch data sets will exhibit both forward and inverse matrix transformability of spectral power values.
- If Thornton’s measurements deviate erratically from the Stiles–Burch values, an equivalently erratic pattern of transformability failure will be revealed by the \( \partial T : \partial S \) ratios.
- If the spectral power values in the source data sets are internally consistent and functionally related, then the \( \partial T : \partial S \) ratios should mirror any intrinsic relationship between them as smooth and consistent deviations from a ratio of 1:1.

### Tristimulus Sum and Chromaticity Coordinate Analysis

The cumulative scalar value of a stimulus is modeled by tristimulus sum, and the implicit vector direction in colour space is modeled by chromaticity, calculated as the relative three-dimensional weighting provided by the normalized CMFs. An implicit and potentially important distinction is three-dimensional weighting provided by the normalized CMFs. An implicit and potentially important distinction is the cumulative scalar value of a stimulus is modeled by tristimulus sum, and the implicit vector direction in colour space is modeled by chromaticity, calculated as the relative three-dimensional weighting provided by the normalized CMFs. An implicit and potentially important distinction is thus made between vector direction and vector magnitude when chromaticity coordinates are calculated.

However, Thornton points out that the conventional chromaticity calculation produces anomalies. The chromaticity divisor \( R + G + B \) tends toward or passes through zero at some wavelengths in some NP and AP CMF matches when large negative tristimulus values are present. This causes the chro-
maticity locus to approach or pass through infinity in such cases.

Chromaticity definitions are recalculated in the following analysis using absolute tristimulus values as the divisor when calculating signed chromaticity definitions. The chromaticity definition then becomes a pure $R:G:B$ ratio that is independent of the sign of the component values, and the observed distortions in the chromaticity locus are thus removed. There is no effect in all-positive systems such as the calculation of $x,y,Y$ coordinates in the CIE model.

The alternative chromaticity calculation is used in two ways as follows:

1. *In spectral stimulus matches*. In the first analysis the chromaticity definitions of the CMFs are calculated using the adopted convention, and the loci of the transformed Thornton PC CMF matches are compared with the equivalent Stiles–Burch definitions. The intent is to reveal variations in the $R:G:B$ tristimulus ratios, particularly at the "crossover" points representing $B/G$ and $G/R$ numeric equivalence. The variation in $R:G:B$ ratios at other wavelengths is also regarded as important in relation to the spectral definition of T-unit value.

2. *In neutral stimulus matches*. An analysis based on predicted neutral-stimulus matches is used to establish if Thornton’s neutral-metamer bow-tie diagrams are also an implicit property of his PC spectral stimulus matches. Stiles–Burch spectral-triplet definitions are first calculated for a set of metameric neutral stimuli that match Illuminant $S_c$. Each match definition is derived from the maximum saturation CMF definitions as a matrix product, using one variable and two primary wavelength stimulus definitions in each case. The calculation is the inverse of that used to extract the CMFs from Maxwell method neutral matches.3

The transformed Thornton PC CMFs are next used to re-calculate the chromaticity of each matching primary triplet. If Thornton’s bow-tie diagrams are also an implicit property of his maximum-saturation CMF matches, the recalculated chromaticities will yield a systematic bow-tie diagram.

**THE EVIDENCE FOR EXPERIMENTAL CONSISTENCY AND SYSTEMATIC RELATIONSHIPS**

Evidence is sought for significant differences between the data sets after they have been nominally unified by matrix transformation of power values. After transformation, the Thornton data set is reweighted by reference to the Stiles–Burch primary wavelengths of 444, 526, and 645 nm. Any revealed differences are then relative to the exact 1:1 equality between the data sets (at those wavelengths) that is generated by matrix transformation.

**Data Consistency**

Experimental consistency is examined by comparing the transformed Thornton CMFs and derivatives of them, with their Stiles–Burch equivalents. In principle, the PC CMFs are likely to be the most consistent, because they involve the minimum of negative tristimulus values and this makes them intrinsically more robust. The power represented by the reference stimulus being matched never drops far below 50% of the total power present across all matches. In NP matches the amount of modifying stimulus power added to the reference stimulus is greatly increased, and in many AP matches the amount of modifying stimulus required can exceed 80% of the reference stimulus.

Figures 1a, 1b, and 1c illustrate Thornton’s PC, NP, and AP CMFs after matrix transformation. With one exception they appear to relate consistently at all wavelengths to the Stiles–Burch measurements. In Fig. 1c inconsistency is regarded as significant only in the AP relationship below...
520 nm. The large differences in peak height evident at 600 nm in all three graphs turn out to represent a significant and systematic departure from matrix transformability that is revealed more clearly in Fig. 2.1 The trends are again relative to the exact 1:1 ratio at the primary wavelengths, produced by the matrix transformation. The fact is emphasized that the deviations are all from the constant T-unit power-ratio that is equated to Illuminant SF and used for CMF normalization, in the Stiles–Burch reference model.

The graphs of Fig. 2 are regarded as representing two distinct instances of unit additive-value difference. A wavelength-dependent locus of varying spectral effect appears to be superimposed on distinct nonlinearities whose characteristics are dependent on the stimulus power of the primaries used in each match. By virtue of CMF normalization, matrix transformability can be demonstrated if the transformation produces a precise 1:1 $\frac{dT}{dS}$ relationship at all wavelengths and stimulus power levels. The graphs of Fig. 2 each include a graph of the required 1:1 $\frac{dT}{dS}$ relationships between the alternative CMF definitions.

Each individual point on the spectral loci of Figs. 2a, 2b, and 2c represents a solution by matrix transformation. Virtually all of these solutions are incorrectly predicted using a single matrix. The difference is relative to the established weighting provided by the Stiles–Burch CMFs and its associated constant T-unit definition. Clear systematic nonlinearity is evident in the graphs for the transformed data at 444 nm. For example, in Fig. 2c the prediction by matrix transformation at 490 nm differs by a factor in excess of 2 from the linear reference value calculated using a single matrix. By contrast, the data in the graphs for the 526- and 645-nm dimensions are not as strongly nonlinear relative to the Stiles–Burch reference data. It is, however, evident from Figs. 2a and 2b that a constant matrix relationship would still produce significant transformation errors, mainly because of the superimposed wavelength-dependent effects. In Fig. 2b the calculated prediction by a linear 1:1 $\frac{dT}{dS}$ relationship is incorrect by a factor close to 1.2 at 600 nm and in Fig. 2a it is incorrect by a factor close to 0.7 at 590 nm. Further, in Fig. 2 and the match at 490 nm, the transformed PC prediction specifies more watts than those used in the Stiles–Burch 444-nm stimulus and fewer watts are predicted than those in the Stiles–Burch stimuli at 526 and 645 nm (see also the explanation via chromaticity ratios under recalculated chromaticity definitions).

Superficially, it seems that the conventional model in which the tristimulus sum is calculated using a constant weighting of power values must be abandoned before the alternative match definitions can be unified by matrix transformation. It is, however, the implementation of CMF constancy that is challenged, not the principle itself. A relatively simple redefinition of the relationship between power value and T-unit additive value (using the proposed SED model elaborated under data unification methodology) is
sufficient to restore the tristimulus sum weightings to their correct status as a linear colorimetric constant.

Relationships Common to All Three Primary Sets

The differences in the 444-, 526-, and 645-nm relationships of Fig. 2 may be essentially a property of the stimulus wavelength being matched and would then be common to all three CMF comparisons. Alternatively, they may be a property of the primary stimuli used in the match that may vary from one primary set to another. The NP and AP data would then be expected to produce potentially different graphs from the PC versions in Fig. 2. Comparisons of the source matrices of relationship1 indicate that very different wavelength combinations predominate in the PC as opposed to the NP and AP CMF matches before transformation into a 444-nm Stiles–Burch definition. Figure 3 demonstrates that there is very little difference in graphed effect, whether it is the PC 452-nm stimulus that predominate in the transformed match or a combination of all three NP or AP primaries with opposing positive and negative values.

Figure 3 supports Thornton’s finding that there is a visual effect (hypothetically unquantified in conventional colorimetric calculations) that varies as a function of the wavelength being matched and is common to measurements using all three primary sets. A simple linear modification of the CMFs might correct this effect at all wavelengths and stimulus intensities were it not for the differences of non-linearity represented in Figs. 2a, 2b, and 2c.

Improved prediction by a plausible spectral level modification of the CMFs is demonstrated in his more recent articles by Thornton. However, this linear solution can be only partial and data specific. For example, in Fig. 2(c) no linear solution is available that simultaneously resolves the graphed \( \partial T / \partial S \) ratios at 490 and 444 nm.

Recalculated Chromaticity Definitions

Nominally equivalent chromaticity definitions for maximum-saturation matches are next compared, using values derived from the Stiles–Burch CMFs and the transformed Thornton PC CMFs. Matrix transformation is again responsible for equating the compared definitions at the primary wavelengths.

Figure 4 indicates that the Thornton definitions of chromaticity differ both systematically and smoothly from the Stiles–Burch definitions. It also demonstrates that (after matrix transformation) the proportions in which the \( R, G, \) and \( B \) stimuli are mixed by the Thornton average observer, are significantly different from those by the Stiles–Burch average observer. In particular, the blue–green and red–green balance points (i.e., points of equivalent T-unit value) are relocated spectrally by a full 10 nm in the Thornton definition as opposed to the Stiles–Burch definition. By implication, the spectral power ratio defining T-units of the Stiles–Burch primaries is not a matrix function of the T-unit ratio represented in the CMFs measured by Thornton using PC-wavelength primaries.

Similar graphs have been generated using the NP and AP data. The graphs thus derived were generally similar to Fig. 4 with similar blue–green and green–red balance-point displacements. They are, however, significantly less smooth, and the \( R:G:B \) ratios derived from the AP CMFs differ radically from the Stiles–Burch ratios below 520 nm.

Neutral-Metamer Power Ratios as a Definition of Spectral T-Unit Value

The T-unit defining \( R:G:B \) ratio equated to Illuminant \( S_b \) is characteristic of a given set of defining primaries and is also used to normalize the CMFs. A different T-unit ratio is characteristic of each alternative primary set. For any given pair of CMF sets, the relevant \( R:G:B \) ratios are then in principle derivable by matrix product from either set of CMFs. The appropriate (forward or inverse) matrix for either CMF set is established by the weightings at the six primary wavelengths. Thus, in principle, all possible primary tristimulus ratios are equivalent by matrix product and derivable from a single master CMF set, which has been

**FIG. 4** Recalculated chromaticity definitions for each primary are plotted as a function of the matched wavelength. The alternative convention of using the absolute-value sum as the divisor \( R+G+B \) is reflected in the shape of the graphs.
The hypothesis is next tested that this relationship is true for the Thornton PC and the Stiles–Burch data sets. The primary tristimulus ratio is also characteristic of the spectral triplet metamers that lie on the neutral axis. Thus the hypothesis is tested by calculating a set of spectral-triplet metamers that lie on the Stiles–Burch neutral axis and then plotting their tristimulus ratios relative to the Thornton PC neutral axis. The hypothesis is untrue in this case, and the result is a bow-tie diagram of deviant tristimulus ratios.

Power values at 10-nm intervals were generated by iteration to equivalence with Illuminant SE, using the Stiles–Burch CMFs. The wavelengths and values are tabulated at alternate 20-nm intervals in Table I, omitting the intervening 10-nm interval steps. The calculated tristimulus ratio definitions for the Stile–Burch metamers, and their redefined values according to the transformed Thornton CMFs, are tabulated alongside the power values for comparison. The implication of Fig. 5 and Table I is that the tristimulus ratios of the normalized Thornton PC and Stiles–Burch CMFs are not mutual matrix products of either CMF set.

The discussion in Part I, of reference 1 related the bow-tie phenomenon to the differences in visual neutrality definition when comparing Maxwell method and maximum saturation method CMF determinations. Here the comparison is between two maximum saturation CMF determinations made with alternative primary sets, and a clear bow-tie relationship is again evident between the relevant neutral axis definitions.

**DISCUSSION**

**The Bow-Tie Diagrams**

The preferred interpretation of Fig. 5 is that the failure of matrix transformability is caused by an incorrect unit mapping from watts to psychophysical T-units. Any constant three-dimensional difference in the T-unit normalization ratio is then revealed as propagated effects modified by CMF weighting. Thus the multilobe bow-tie effect increases and decreases and changes direction in the chromaticity plane over wavelength in parallel with CMF weighting. The diagrams are therefore a direct product of the conventional link between watts and T-units, when the relationship is established by a single R:G:B ratio relative to Illuminant SE.

<table>
<thead>
<tr>
<th>Matching triplets</th>
<th>Power value mW</th>
<th>Stiles-Burch tristimulus ratios</th>
<th>Thornton tristimulus ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>G</td>
<td>B</td>
</tr>
<tr>
<td>(645) (526) (400)</td>
<td>2164.62</td>
<td>1117.29</td>
<td>16132.67</td>
</tr>
<tr>
<td>(645) (526) (444)</td>
<td>2306.39</td>
<td>1075.61</td>
<td>654.93</td>
</tr>
<tr>
<td>(645) (526) (444)</td>
<td>2306.39</td>
<td>1075.61</td>
<td>654.93</td>
</tr>
<tr>
<td>(580) (526) (480)</td>
<td>737.57</td>
<td>886.19</td>
<td>657.76</td>
</tr>
<tr>
<td>(580) (526) (480)</td>
<td>737.57</td>
<td>886.19</td>
<td>657.76</td>
</tr>
<tr>
<td>(600) (526) (480)</td>
<td>952.34</td>
<td>1020.28</td>
<td>656.19</td>
</tr>
<tr>
<td>(600) (526) (480)</td>
<td>952.34</td>
<td>1020.28</td>
<td>656.19</td>
</tr>
<tr>
<td>(620) (526) (480)</td>
<td>2306.39</td>
<td>1075.61</td>
<td>654.93</td>
</tr>
<tr>
<td>(620) (526) (480)</td>
<td>2306.39</td>
<td>1075.61</td>
<td>654.93</td>
</tr>
<tr>
<td>(640) (526) (480)</td>
<td>2306.39</td>
<td>1075.61</td>
<td>654.93</td>
</tr>
<tr>
<td>(640) (526) (480)</td>
<td>2306.39</td>
<td>1075.61</td>
<td>654.93</td>
</tr>
<tr>
<td>(660) (526) (480)</td>
<td>19128.62</td>
<td>1096.35</td>
<td>654.96</td>
</tr>
</tbody>
</table>
Correlated Explanations

The underlying differences in neutral-metamer ratio between the Thornton and Stiles–Burch data sets are revealed by the bow-tie diagram in Fig. 5. Inspection of Figs. 1 to 4 reveals that the two major lobes of Fig. 5 correlate at least visually in both size and direction with the graphed deviations in all four figures. For example, there is a clear correlation between Figs. 4 and 5. The displacement in magnitude and direction of the chromaticity loci in Fig. 4, correlates with the magnitude and direction of neutral stimulus-ratio loci of Fig. 5. Further, the different chromaticity ratios at 490 nm in Fig. 4 correlate with the deviations of $\frac{\delta T}{\delta S}$ ratio at 490 nm in Fig. 2.

Additionally, the direction and magnitude of nonlinear difference in Fig. 2 correlate with the lobes of Fig. 5 and may also represent propagated effects of ratio difference distorted by CMF weighting. That is to say, differences in the T-unit $R:G:B$ ratios may distort otherwise linear relationships. The distinct nonlinearities of Fig. 2 would only represent a confirmed visual response characteristic, if the $R:G:B$ ratios as a function of stimulus power, are significantly different for a set of neutral metamer with constant spectral composition when measured over a range of power values.

Finally, an interpretation of the wavelength dependent effects in Figs. 1 and 3 based on $R:G:B$ normalization ratios, would explain the insensitivity to the choice of alternative primaries (but see also Unanswered Questions). The predictions in Part I are supported in each of these correlations, that “a spectrally distinct mapping of watts onto three-dimensional T-units is both necessary and sufficient to represent the differences between the Thornton and Stiles–Burch data sets.”

In general, the bow-tie phenomenon is predicted to arise in two cases. First, it will occur in a conventional colorimetric comparison between a Maxwell method and a maximum saturation method CMF determination by a single observer under a single primary set. Second, it is likely to occur between alternative primary maximum saturation CMF definitions made by a single observer. In either comparison, the effect is predicted to disappear if the single $R:G:B$ ratio link between watts and T-units is replaced by the proposed SPD/SED relationship elaborated later under The Data Unification Methodology.

Spectral Power Ratios

It is clear from the spectral power ratios $\frac{\delta T}{\delta S}$ in the graphs of Fig. 2 that the transformed Thornton primary-power data are a smoothly nonlinear function of the Stiles–Burch primary power used in each CMF match. Critically in the current context, the nonlinearity has a different characteristic in each 444-, 526-, and 645-nm dimension of the Stiles–Burch model.

The differences in nonlinear characteristic are sufficient to violate the requirement for uniform additive properties under the laws of additive colour mixing. The validity of the laws is, however, not in question; the laws are assumed to be correct. The intent of the proposed data unification method is therefore to render the data into a uniform T-unit definition of additive properties such that both data sets relate to those laws in a mutually transformable and therefore consistent manner.

The proposed precursor transform seeks to map the distinct $\frac{\delta T}{\delta S}$ ratios of Fig. 2 onto a uniform and constant three-dimensional definition of T-units to establish transformability. The presence of a wavelength-dependent effect requires that the transform must be N-dimensional. If, in addition, the nonlinear characteristic differs from wavelength to wavelength as in Fig. 2, then the transform must include an equivalent set of $N$ nonlinear mapping functions as described later under The Data Unification Methodology. Figure 2 also clearly demonstrates that the smoothly changing nonlinearities, proposed as a modeling construct, are evident in the relationships between Thornton’s measurements and the Stiles–Burch reference data set.

There is insufficient evidence to determine if the wavelength-dependent loci of $\frac{\delta T}{\delta S}$ ratios in Fig. 2 would change shape as a function of the power value in the matched reference stimulus (wavelength constant). The proposed precursor transform is therefore a generic normalization and linearization, which can model any possible smooth and continuous relationship in $N$ dimensions revealed by further experiment.

A set of Maxwell method neutral matches at a range of wavelengths and a range of photopic intensity levels is needed to establish the true nature of these wavelength-dependent relationships. Because there are insufficient data, it is regarded as premature to attempt a definitive numeric unification of the Thornton and Stiles–Burch data sets.

UNANSWERED QUESTIONS

Several important features of Thornton’s adopted primary stimuli and the measurements made with them raise questions that cannot be answered without further experimental work. The NP and AP CMF determinations require substantial additions of modifying stimuli to the matched reference stimulus at each CMF wavelength. This is reflected in the strongly cross-dependent nature of the source matrices in relationship (1). A significantly different set of matrices is derived if the Stiles–Burch data rather than the Thornton data are used to generate these data. The relevant matrices are derived by matrix inversion of the Stiles–Burch definition of the Thornton primaries rather than by the alternative route, where the matrix is compiled directly from the Thornton definition of the Stiles–Burch primaries. The inverted matrices are given below in relationship (3) for comparison with those in relationship (1). Interpolated 1-nm interval power values were used throughout both derivations.
The differences between relationship (3) and relationship (1) are highlighted by referring to the AP definition of the Stiles–Burch 444-nm primary. Anyone seeking to repeat Thornton’s AP determination is predicted by these matrices to require nearly 34 watts of monochromatic radiant power at 653 nm when attempting to match 1 watt at 444 nm. By contrast, Thornton only used about 11 watts in this match.

The alternative of working throughout using 10-nm interval definitions has also been calculated. This approach produces yet another set of significantly different matrices compared with those in relationships (1) and (3). In particular, the predicted power needed for the AP match at 444 nm rises to 41.8 watts based on 1-nm interval interpolated Stiles–Burch definitions and 10.7 watts by interpolation to a 1-nm interval within Thornton’s measured data. Although the AP definition for a 444-nm stimulus has been highlighted, the majority of the 27 matrix elements in relationship (3) are significantly different in percentage terms from those in relationship (1).

The net effect of such ill-conditioned matrices is to make error propagation extremely uneven. The matrices also indicate that any differences are typically either exaggerated or minimized differently in each dimension. This makes the transformed AP and NP data a potentially unsafe basis for diagnostic interpretation. It was therefore decided to base the data unification method, and the conclusions concerning matrix transformability, mainly on the PC measurement data.

### The Data Unification Methodology

The Thornton and Stiles–Burch data sets both describe the same group of visual sensations using stimulus-match definitions in watts. In consequence it should be possible to unify the two alternative data sets by matrix transformation of stimulus power definitions. However, systematic failures of matrix transformability between such data sets are evident in the sections from The Basis for Data-Set Comparison to Discussion. The following data unification method resolves these transformability failures by a redefinition of the conventional N→3 colorimetric mapping of watts onto T-units.

The proposed solution is to redefine T-units using a comprehensive set of neutral metamers representing all possible sets of defining primaries and introduce an additional N→N stage of mapping to model their R:G:B ratios. The result is an overall N→N→3 T-unit definition, which is in principle a general definition of equivalent spectral effect in three dimensions, common to all the wavelengths present in the evaluated metamers. The introduced N→N mapping can therefore be considered to convert the SPD definition of any given visual stimulus into a spectral effect distribution (SED). The set of all possible SEDs can now be classified more accurately into visually equivalent subsets by conventional tristimulus sum, because each spectral dimension of an SED shares a common unit definition.

In principle, there is no reason to alter either the definition of visual neutrality that equates Illuminant S_E with a single T-unit R:G:B ratio or the definition of visual equivalence provided by the standard observer tristimulus sum. However, by virtue of the proposed N→N mapping, the definition of visual neutrality becomes more complete. In effect, the mapping establishes a distinct and unambiguous definition of the spectral response, for each individual added milliwatt of stimulus power in each spectral component of an SED.

The set of N distinct functions equate Illuminant S_E with a matrix of many individual R:G:B ratios ordered by stimulus power and wavelength. Each of the N functions transforms the data for a single spectral component of the input SPD. The only proposed change in visual equivalence calculations is to use SEDs rather than SPDs as the input definition of a visual stimulus. Conventional colorimetric calculations otherwise remain intact and unaltered.

#### Additive Equivalence Mapping Using Visually Neutral Metamers

The first step is to establish a definitive N→N mapping using a comprehensive set of metamers, each consisting of a triplet of monochromatic primaries. In principle, the data set used should represent a full range of both defining stimulus wavelengths and photopic stimulus power values. Each data-set member then represents a specific combination of primary wavelengths and supplies an R:G:B power ratio that is visually neutral at a specific visual intensity level. In effect, the T-unit defining relationship between Illuminant S_E and a single R:G:B ratio is thus evaluated...
The Generalization from Neutral to Nonneutral Metameric Matches

The developed $N \rightarrow N \rightarrow 3$ transform first establishes and then maps a fully evaluated set of neutral-metamer SEDs onto a single T-unit definition. The trichromatic generalization then validates the mapping for the set of all SEDs that have the same T-unit definition. The only additional information required to predict their visual equivalence is provided by the master CMFs, once they have been normalized using the master T-unit $R:G:B$ ratio.

After an appropriate SPD$\rightarrow$SED transform has been established, the next nominal step is to normalize a candidate master CMF set derived by independent experiment using the master T-unit definition and test its predictions. Forward and inverse matrix transformability of data sets can be used to validate these CMFs (and also their primary definitions) as the remaining components of an overall master reference standard. In practical terms it would make more sense to adopt an existing standard observer as the reference definition (of both T-units and CMFs) and establish the $N \rightarrow N$ mapping relative to that definition.

The proposed data unification method is regarded as analogous to that adopted by W. D. Wright and his colleagues when they used the spectrally distinct $V_\lambda$ definition of visual response magnitude and T-unit $R:G:B$ normalization of the CMFs for the 1931 standard observer definition. The only difference is the prior conversion of SPDs into SEDs. The proposed SPD/SED mapping is conceived as a generic method for establishing data transformability. As defined, the mapping has sufficient dimensions and functionality to represent any possible smooth and continuous variation in the matrix relationship between experimentally established stimulus power and visually established spectral effect.

FUTURE EXPERIMENTAL WORK

Some preliminary guidelines for a set of appropriate confirmatory experiments are now suggested. The key primary task is to confirm by verified independent experiment, that the differences between the Thornton and Stiles–Burch data sets represent significant properties of the visual response to stimulus power.

In view of the diagnostic inadequacies of NP and AP maximum saturation matches, they are regarded as unsuitable as a basis for fundamental investigation. I agree wholeheartedly with Thornton, that locating the PC primary wavelengths at 450, 530, and 610 nm is somewhere near to the optimal choice. Any experiment should capture the data using these primary wavelengths, and then establish a consistent T-unit definition of visual neutrality at the spectral level.

Attention is directed to the discussion by Wyszecki and Stiles of the theoretical differences between maximum saturation CMF determinations and Maxwell method neutral-axis matches using spectral triplets. Several articles have been published, notably by Crawford and Wyszecki, that indicate that stimulus intensity-dependent deviation (from uniform and proportional relationships) is evident in Maxwell method measurements. It is plausible to speculate that the measurements concerned reflect data-set relationships of the type now found to be implicit in Thornton’s maximum saturation PC measurements. Maxwell method neutral-stimulus matches are regarded as the correct basis for the design of the fundamental experiments required to define a consistent T-unit definition of visual neutrality. They offer direct measurement of the $R:G:B$ ratios in visually neutral metamers.

The key experiment is to evaluate the detailed functions necessary in a correct $N \rightarrow N$ mapping. The evaluation is by Maxwell method matches using a comprehensive set of neutral spectral-triplet stimuli. The revealed $N \rightarrow N$ relationship could be nonlinear (relative to the Stiles–Burch stimulus power definitions) to a different extent in each of three or more spectral dimensions. In consequence it is necessary to repeat the full set of spectral determinations at sufficient photopic stimulus intensity levels to characterize these nonlinearities. It is not clear, using the available data, whether the set of $R:G:B$ ratios generated at alternative visual intensity levels will indeed be significantly different. It is therefore recommended to proceed in two stages, first to confirm the origin of the nonlinearities and the effect of visual intensity and then second (if they originate even partly in the visual response) to characterize their nonlinear nature.

The proposed data unification method by T-unit redefinition is expected to be applicable to the unification of CMFs derived by Maxwell method matches with those derived by maximum saturation matches. However, in the proposed experimental plan, Maxwell method neutral-axis matches are given the limited function of trichromatic unit definition, across the full range of visible wavelengths and photopic stimulus intensities. CMF quantification should then be provided by an existing maximum saturation definition. The de facto reference status of the existing standard observer definitions established by this method is an important consideration.

CONCLUSIONS

The first intent of this article is to demonstrate that Thornton’s six-observer average colour matching functions ex-
pressed in watts stand in a detailed nonlinear relationship to the power values in watts determined by Stiles and Burch for matches to the same reference stimulus. The second intent is to establish a hypothetical model of the observed differences. To achieve these objectives, the Thornton data set was compared with the Stiles–Burch 49-observer averaged differences. To achieve these objectives, the Thornton data intent is to establish a hypothetical model of the observed for matches to the same reference stimulus. The second power values in watts determined by Stiles and Burch were converted by matrix transformation into a redefined data set, in which all matches are expressed in terms of stimulus power at the Stiles–Burch wavelengths of 444, 526, and 645 nm.

After transformation Thornton’s measurement data were demonstrated in Fig. 1 to be internally consistent across all three CMF determinations, with the possible exception of the AP CMFs below 520 nm. As shown in Figs. 2 to 5, they were also found to deviate in a smooth and consistent manner from the Stiles–Burch definitions. However, the matrices necessary to transform the NP and AP measurements into Stiles–Burch equivalent values cause error propagation to be very uneven, and the diagnostic value of NP and AP primary matches is probably limited. The evidence from the NP and AP CMFs is therefore given less weight, and the following conclusions are based mainly on the evidence from Thornton’s six-observer average PC CMF determination.

The following overall conclusions are made concerning Thornton’s measurements:

1. The full PC, NP, and AP CMFs provide clear confirmatory evidence of the smoothly varying super- and subadditive differences that are predicted in Part I. The differences occur within the three Thornton data sets and also in the comparisons of all three sets with the Stiles–Burch data set.

2. Differences in the spectral power ratios \( \frac{\partial T}{\partial S} \) are clearly evident between the 444-nm relationship and the 526- and 645-nm relationships of Fig. 2. This is regarded as the critical cause of systematic matrix transformation failure. Matrix transformation of definitions in watts would in principle be a successful (but arbitrary) method of data unification, if the same nonlinear characteristics were present in all three dimensions. The term arbitrary is used here for the reasons given in Part I, concerning the transformation of data sets that stand in a uniformly nonlinear relationship.

3. There is significant evidence (in the spectral loci in Fig. 2) that a distinct and perhaps dominant wavelength-dependent relationship is present. Any proposed data unification method must therefore be spectrally defined.

4. The bow-tie diagrams demonstrated in neutral-stimulus matching experiments by Thornton are also found to be an implicit property of his independent maximum-saturation PC matches. Critically, they reveal a wavelength-dependent variation in the neutral \( R:G:B \) ratio, which is not predicted by either the PC or the Stiles–Burch CMFs.

5. The convention is recommended that the divisor \((R+G+B)\) used to specify chromaticity should be calculated using absolute \( R,G,B \) values. The recommendation has no effect in all-positive coordinate systems such as CIE \([X],[Y],[Z]\). However the alternative convention establishes a pure \( R:G:B \) ratio for the total stimulus power used in the match. It also removes serious distortions in the \( R,G,B \) chromaticity locus for maximum-saturation match definitions, when the stimulus being matched includes significant additions of modifying primary stimulus.

6. A numeric unification of the Thornton and Stiles–Burch data sets using the proposed method has not been attempted in this article, because there is insufficient fundamental data. The intent is therefore only to establish a methodology that characterizes the differences between the data sets. It is recommended that a definitive demonstration of data unification for alternative primary CMFs should be attempted only once the observed nonlinear relationships have been confirmed and fully characterized by the proposed independent experiment with neutral metamers.

7. Theoretically, all the differences between the compared data sets could in principle be the product of systematic differences in either experimental method or measurement. This is regarded as unlikely, but the presence of spectrally distinct nonlinear relationships in the human visual response must be regarded as unproven by Thornton’s experimental results. This needs to be confirmed by independent experiment.

The following conclusions concern the resolution of the observed failures of data-set transformability:

8. The apparent similarity of nonlinear effects revealed by the PC, NP, and AP CMF determinations suggests that they are a function of the matched wavelength rather than of the primaries used to match it. This is regarded as the predictable but not necessarily the consequential result of incorrect T-unit definition. A direct causal link can only be established by future experiment.

9. A novel transformation method using a generic data linearization and normalization construct is described in Part I and elaborated on under The Data Unification Methodology. It is designed to render data sets onto a common comparison basis by converting stimulus-power values into visually equivalent spectral effect values. The intent is to make any pair of alternative data sets mutually transformable by the matrix product of the normalized CMFs. The wider intent is to provide a data unification method capable of characterising super- and subadditive visual effects in general.

10. The proposed data unification method is by T-unit redefinition. The intent is to establish a common definition of unit-additive value for all spectral wavelengths and all photopic stimulus intensities at each wavelength.

11. The appropriate T-unit redefinition is established by Maxwell method matches and quantified by an \( N \rightarrow N \)
The expected outcome is that a single T-unit definition using many neutral metamer definitions. In effect, the three-dimensional T-unit concept is maintained, but its conventional implementation using a single $R:G:B$ ratio (reflecting the use of a single set of defining primaries) is replaced by many definitions (reflecting the properties of many alternative primary combinations).

12. It is strongly recommended that the T-unit redefinition should in practice be used with, and referenced to, an existing standard observer definition. It should also be calculated numerically relative to the conventional definition of visual neutrality provided by Illuminant $S_E$.

13. A necessary corollary of predicting visual equivalence more accurately by redefining T-units is that all compared SPDs must first be converted into spectral effect distributions (SEDs) using the new T-unit definition. This is, however, the only change necessary to otherwise unaltered standard colorimetric calculations.

14. The expected outcome is that a single T-unit definition and set of CMFs can be established as a colorimetric master standard. The reference CMF definition (and candidate master standard) used in the presented analysis is the Stiles–Burch 49-observer average $R,G,B$ definition.

15. A “candidate master CMF-set” can be validated by the accurate matrix-product prediction of a definitive set of metamer matches. The test set could, for example, be sets of alternative–primary CMF definitions, or a similar set of related matching metamers, such as the set of neutral stimuli used by Thornton to generate bow-tie diagrams. As conventionally defined, the Stiles–Burch model fails this test on both counts using the Thornton data sets, according to Figs. 1 to 5.

In principle the Stiles–Burch CMFs would not fail this test if the data were first redefined using the described data unification method. However, the proposed $N\to N\to 3$ unification method only “calibrates” hypothetical visual-response effects. These supposed effects must first be demonstrated to exist using independent verified experimental results before the pass-certificate can be issued. Until then, the data unification method could be dismissed as a pure data-fitting exercise for the Thornton data, without confirmed meaning.

16. A practical confirmation experiment using Maxwell method neutral-metamer matches is described under Future Experimental Work. The experiment is designed to establish that the apparent visual effects represented by the Thornton data are indeed real visual response characteristics and the data unification method is both necessary and sufficient to model them.

17. It is not clear if the $R:G:B$ ratios thus derived will vary as a function of neutral-stimulus intensity when the stimulus wavelengths are held constant (i.e., it is not yet possible to distinguish whether the nonlinearities result from incorrect CMF normalization or are at least in part due to a true visual response characteristic). It is important to establish the origin of these nonlinearities as a first priority, before further experiment.

18. An informal connection is postulated in the appendix between the introduced redefinition of T-unit value and the implicit use of vector-algebraic constructs in the standard colorimetric transform.

**APPENDIX: A VECTOR ALGEBRAIC INTERPRETATION**

An informal vector algebraic interpretation is now presented for the proposed redefinition of trichromatic units. The following interpretation concerns fully quantified sets of perceptual equivalence definitions. They are specified in terms of additive units, either of (physical) stimulus power or of (psychophysical) visual response. In the suggested model an additional $N\to N\to 3$ stage of spectral power rescaling is interposed between the conventional SPD definition of a stimulus and the conventional $N\to 3$ mapping onto equivalence definitions by tristimulus sum matrix product.

T-unit definitions of visual equality in watts have the potential status of basis definitions for vector magnitude when they are referenced to visual neutrality and used for CMF normalization. The relevant definitions must, however, be uniformly equivalent by matrix transformation over the set of all possible triplets of defining primary stimuli that could be used for visual match definition. The use of an $N\to N\to 3$ as opposed to the conventional $N\to 3$ definition of T-units is important. It removes the intrinsic spectral-definition ambiguity of an $N\to 3$ equivalence expressed in watts and therefore enables the theoretically correct vector-sum prediction of all possible instances of perceptual equivalence.

In concept, but without rigorous definition, equivalence of vector magnitude must be established between all the distinct spectral components of the response to light stimuli. Conventionally, equivalence of vector magnitude is established relative to Illuminant $S_E$ by a single $R:G:B$ ratio. This ratio represents the specific set of defining primaries used to establish a definitive set of maximum-saturation CMFs and is used to normalize the CMFs.

In the extended model, a comprehensive set of neutral-metamer ratios (rather than a single $R:G:B$ ratio) is equated to Illuminant $S_E$. Conceptually, this set establishes definitions in watts for all spectral components of the visual response and establishes a spectrally defined unit vector for the set of all possible stimulus combinations or SPDs. This unit vector then becomes the origin of the first $N\to N$ stage of perceptual equivalence mapping, whose notional result is the spectrally defined unit vector (expressed in normalized and rescaled watts) for the set of all possible spectral effect combinations or SEDs.

The proposed $N\to N$ mapping of vector magnitude between SPDs and SEDs is regarded as a clearly permissible mapping of vector definitions, because the redefinition of vector magnitude is scalar, and it is established by $N$ distinct single-dimension operations that equate the notional basis unit of vector magnitude across all SED dimensions.
The concept of unit-vector mapping can in principle be validated by the demonstrable matrix transformability of data sets. The use of vector algebraic manipulation, and the $N \rightarrow 3$ mapping represented by the conventional and well-established colorimetric transform model then also become valid vector-algebraic constructs, and by extension, the CMFs become a demonstrable three-dimensional constant of the model.

Perceptual equivalence matches only define vectors in terms of dimensionality and ordinal magnitude. Therefore the vector sum is only an ordinal classification into equivalence sets. By definition a vector sum of match definitions establishes only point identity, and it implies no spatial or scalar-difference relationship between the tristimulus definitions of colour sensations. It is therefore incorrect to construct a vector space only using definitions of perceptual equivalence.

Before a valid vector space can be established, the three-dimensional vector definitions for the set of all colour sensations must first be extended to include a valid “D-unit of perceptual-difference” (perhaps analogous to the three-dimensional T-unit of perceptual equivalence). Only then can a valid vector space be constructed representing, for example, threshold perceptibility or visual match acceptability, depending on the data used to define the units of difference.

ACKNOWLEDGMENTS

I again acknowledge and thank Dr. Thornton for the use of his much-quoted copyright CMF data. I am also indebted to him for the initial route map “Toward a More Accurate and Extensible Colorimetry,” provided by the title and contents of his articles. These articles have been a potent catalyst, without which none of the presented concepts and methods would have been generated. In acknowledgement, I therefore offer my findings as an additional page to that route map.