Munsell Reflectance Spectra Represented in Three-Dimensional Euclidean Space

A. Kimball Romney,* Tarow Indow

School of Social Sciences, University of California—Irvine, Irvine, CA 92697-5100

Received 22 December 2001; revised 31 May 2002; accepted 22 July 2002

Abstract: In this article we describe the results of an investigation into the extent to which the reflectance spectra of 1269 matt Munsell color chips are well represented in low dimensional Euclidean space. We find that a three dimensional Euclidean representation accounts for most of the variation in the Euclidean distances among the 1269 Munsell color spectra. We interpret the three dimensional Euclidean representation of the spectral data in terms of the Munsell color space. In addition, we analyzed a data set with a large number of natural objects and found that the spectral profiles required four basis factors for adequate representation in Euclidean space. We conclude that four basis factors are required in general but that in special cases, like the Munsell system, three basis factors are adequate for precise characterization. © 2003 Wiley Periodicals, Inc. Col Res Appl, 28, 182–196, 2003; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/col.10144

Key words: Munsell system; basis factors; color spaces; geometric aspects; and spectrophotometry

INTRODUCTION

The primary aim of this article is to determine how well 1269 Munsell standard color samples are represented in a three-dimensional space on the basis of their spectral reflectance curves. No human assessment is involved in this representation. Two methods of analysis are used. In one method, the dissimilarities between spectral patterns are captured as Euclidean distances and hence the space in which we embed colors is an Euclidean space E^3. The other method is a more direct approach to patterns of spectra in which no such constraint is imposed and the assumed space is a vector space $\mathbf{R}^3$. Both analyses yielded essentially the same results. It will be shown that the set of Munsell colors form a configuration of 1269 points in E^3 according to patterns of their spectral reflectance curves.

Several researchers (reviewed later in the article) have investigated the question of how many basis factors are required to represent spectral reflectance curves adequately. In contrast, we are only aware of two published examples of the scaled relationships among a complete set of Munsell color samples based on their spectral reflectance curves. Such a description based on purely physically derived spectral data would be of potential interest by itself, as well as to determine the relationship between a physical description and a perceptual description. Such a description also has implications for a variety of questions including the number of basis factors necessary to describe the spectral reflectance curves.

Previous researchers have discussed scaling the spectra of colors but usually do not present actual empirical results. For example, Koenderink and van Doorn[1] present an idealized spherical configuration of color space obtained by singular value decomposition (SVD) but do not show the empirical evidence for the structure. Lenz and Meer[2] have carried out one of the two empirically based studies of the structure of color spectra that we have located. They analyzed "a database consisting of reflectance spectra of 2782 color chips, 1269 from the Munsell system and the rest from the NCS system."[2] The Munsell spectra are the same ones we analyze below. They find that the spectral data are described by coordinate vectors which lie in a cone and they therefore define a hyperbolic coordinate system to represent the data. They present a three-dimensional figure of the distribution of the chips (Fig. 1(b) in Ref. 2) that is somewhat like the upper left panel of our Fig. 2. The structure we describe in this article could be characterized as a set of nested cone-like structures, each made up of a single Munsell Chroma with lower Chroma values being inside higher

*Correspondence to: A. Kimball Romney (e-mail: akromney@uci.edu).

Contract grant sponsor: The research was funded in part by NSF Grant number SES-0136115 to A.K.R. and W.H. Batchelder.

values. The narrow tips of each cone-like structure are found at the lowest Value levels.

The second empirically derived structure of color based on measured reflectance spectra was reported by Burnes, Cohen, and Kurznetsov.[3] From a data set of 427 Munsell color chips under C illuminant measured from 380–770 nm in 10 nm steps, they plotted a configuration of color chips and a curve for wavelengths in "Fundamental Color Space." The space was spanned by three orthoganal axes, Luminosity, Red, and Violet. Points of the same Munsell Value level form a plane and each plane is tilted with regard to vectors representing a power level of both equal energy illumination and illuminant C.

Representation of Color Samples with Metric Multidimensional Scaling

The present study is based on the 1269 color chips of the Munsell color book (1976 matte edition). The data were downloaded from the web site: http://www.it.lut.fi/research/color/database/database.html. The data set is: Munsell colors matt (spectrophotometer measured). The spectra were obtained from a Perkin-Elmer lambda 9 UV/VIS/NIR spectrophotometer from 380 nm to 800 nm at 1 nm resolution using a measurement geometry of 0/d (integrative sphere). Because in subsequent research we will want to compare the physically measured spectra with human perceptually derived structures, we limited the data to the approximate limits of human vision, namely, the 430 nm to 660 nm segment. This results in a data matrix $\mathbf{S}_{N\times M} = (s_{i\mu})$ with $N = 1269$ rows of Munsell chips and $M = 231$ columns of 1 nm spectral intervals.

It has been known for some time that classical metric multidimensional scaling (MMDS) recovers exact distances when applied to data known to be Euclidean.[4–7] In order to obtain a Euclidean representation of the Munsell color space as described by reflectance spectra, we applied metric scaling to the matrix of Euclidean distances. The dissimilarity between two spectral patterns, $s_{i\mu}$ and $s_{j\mu}$, $\mu = 1\ 2, \ldots,$ 231, is converted to a Euclidean distance d_{ij} for all combinations of colors. Matrix $\mathbf{D}_{1269\times 1269} = (d_{ij})$ was analyzed by MMDS to obtain a matrix $\mathbf{A} = (a_{i\alpha})$ on orthogonal coordinate axes $\mathbf{f}_\alpha$ embedded in three dimensional Euclidean space $\mathbf{E}^3$. Our notational conventions as well as the mathematical and statistical details of all computations are presented in Appendix I. Appendix II summarizes in a single table the descriptive statistics relevant to all computational procedures carried out on the Munsell data.

To measure how well the data are fit with just three dimensions, we examined the scatter diagram (or Shepard diagram) between the original 804,546 distances d_{ij} and the computed estimates of these distances $\hat{d}_{ij}$. The results are shown in Fig. 1. Since the contributions of higher dimensions are excluded, $\hat{d}_{ij} \leq d_{ij}$. The variance accounted for was estimated by $r^2 = .9994$. We also plotted the distribution (not shown) of the distances among the color samples and found that the square root of these distances is very nearly normally distributed.

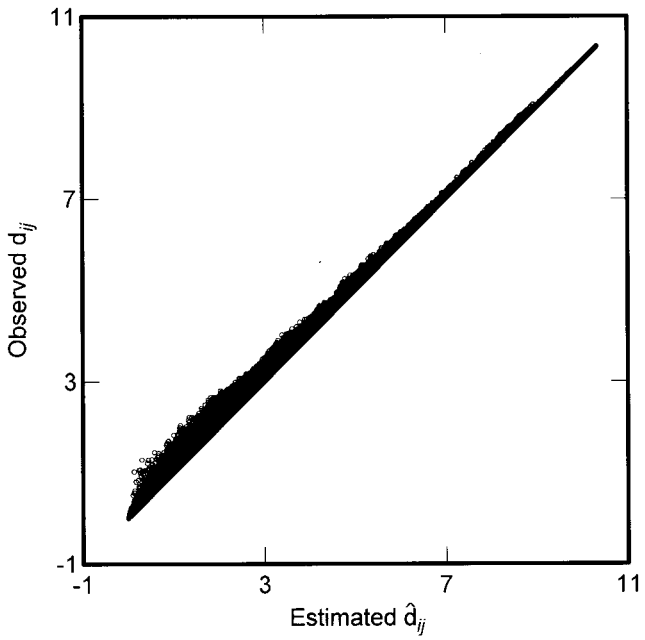

FIG. 1. A scatter plot showing the relationship between the original 804,546 Euclidean distances d_{ij} among the spectral curves of 1269 Munsell color chips and the $\hat{d}_{ij}$ estimated from the three dimensional Euclidean representation.

In order to obtain more detailed insight into the overall structure, we oriented it by trial and error visual rotation. We used two criteria. The first criterion was that when viewed from "above," the achromatic point was centered on the zero point of the second and third dimensions. The second criterion was to minimize the width of the Munsell value levels when viewed from the side perpendicular to the first axis. We used software that allowed us to manipulate a visual image of the rigid figure on the computer screen and iterated by eye until a satisfactory orientation was obtained. The rotation and translation was rigid, insuring that the Euclidean distances among the points were invariant. The axes are labeled A1, A2, and A3 for convenience. These coordinates constitute the matrix $\mathbf{A} = (a_{i\alpha})$ as specified in Appendix I. A plot of the results is shown in Fig. 2.

The units on the axes represent the actual Euclidean distances as calculated from the original spectra from 430 nm to 660 nm using Equation (3) in Appendix I. The left panel in Fig. 2 shows the resulting representation. The colors in the plots show the five Munsell primary hues in their natural color (red, yellow, green, blue, and purple) for the four variations (2.5, 5, 7.5, and 10) for each Hue, with all the intermediate hues labeled as gray. The color coding is for identification purposes only and is not meant to portray actual Munsell colors.

In the lower left panel of Fig. 2, one is looking down through the color space with the achromatic point centered at zero. In the middle left panel, the separate value levels of the Munsell system are viewed from the side. Points for colors of the same value form a plane and all planes are parallel and sloped at an angle of about 37.8 degrees with regard to A1. In the left top panel the perspective of the

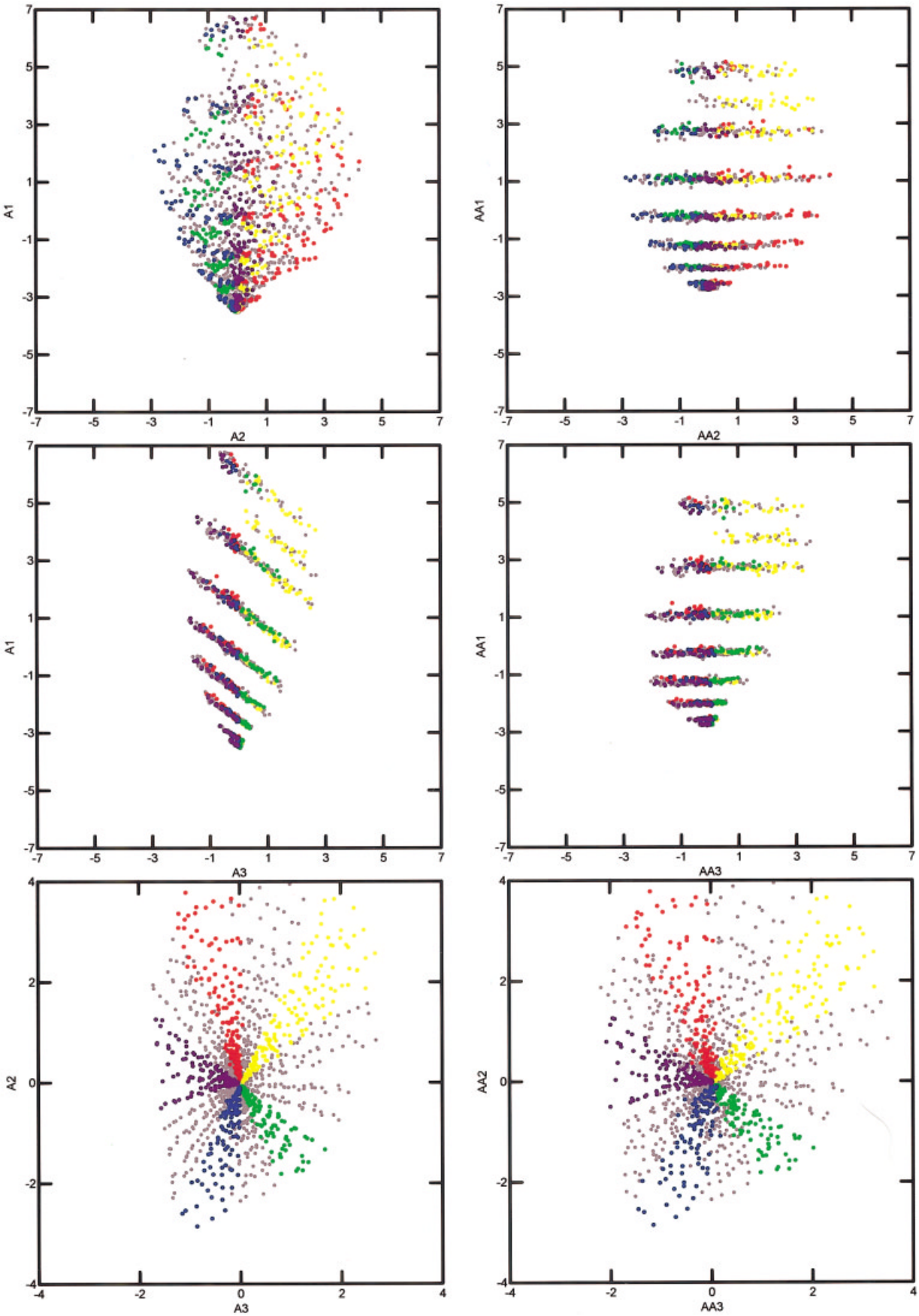

FIG. 2. Scatter plots representing the 1269 Munsell spectra scaled in Euclidean space (left panel) and altered to best represent the Munsell axes (right panel). See text for explanation of color coding and other details.

middle panel is rotated 90° around the vertical axis resulting in an overlap of the slanted planes.

The three orthogonal Euclidean axes may be interpreted in the following way. A1 is the achromatic axis and should relate to the mean energy level of the spectra. Let us denote the mean of the 231 spectral points $s_{i\mu}$ by ME_i for each of the 1269 spectra *i*. The plot of the relation between ME_i and A1 is shown in Fig. 3. There is a close correspondence between the two variables ($r^2 = .99998$) which gives us confidence that the orientation of the figure arrived at by successive visual rotations is satisfactory. The higher the value of ME_i, the larger the Munsell Value of that color *i*. However, ME_i is not constant for colors *i* of the same value, because Munsell Value is the luminous intensity (Y). Axis A1 represents difference of ME_i between Value as well as the variation of (ME_i) within each plane of Value. The luminous intensity Y (Munsell V) is captured as a tilted axis in the left panel in Fig 2.

Axis A2 is oriented along the 10R and 2.5B of the Munsell hue color space. It may be noted that in this physical representation the color chips along this axis would all have roughly the same mean spectra at any given Munsell Value. Axis A3 is oriented along 7.5GY and 5P of the Munsell hue color space. Note that in this case the mean spectral value (ME_i) of the color chips on this axis, at a given Munsell Value level, would increase considerably as one moves from the GY end to the P end of the axis. It is interesting that such a result should emerge from an objective description of the physically measured spectra of the Munsell chips.

In the left panel in Fig. 2, A1 does not correspond to Munsell Value and the bottom plot is a slanted plane of Munsell Hue and Chroma. Hence we have reoriented and modified the perspective to one that corresponds to the Munsell notation. The modified perspective is seen in Fig. 2 in the right panel, and the axes have been labeled AA1, AA2, and AA3. The axis AA1 has been tilted about 37.8 degrees from A1 so that all planes of constant Munsell Value are horizontal as seen in the top and middle plots. The rotation displaced the alignment among planes of constant value and we have horizontally shifted each V-plane on the AA3 axis until its achromatic point is centered at AA3 = 0. The magnitude of the shift is given in Table I. The new vertical axis AA1 passes through these achromatic points and is perpendicular with all planes of constant Value. The bottom plot is the projection of all V-planes along AA1 to a plane spanned by AA2 and AA3. These new coordinates in the right panel are directly related to Munsell value, hue, and chroma. To relate distance $\hat{d}_{ij}$ in the solid to dissimilarity between $s_{i\mu}$ and $s_{j\mu}$, however, we have to use the solid in the left panel. In the remainder of the paper we refer to the representation in the right panel of Fig. 2 as the Munsell Model.

TABLE I. Amount of shift at each Value level necessary to bring the achromatic point to zero on axis AA3 of the Munsell color space representation.

Munsell Value level	Amount of shift
2.5	−2.1
3	−1.9
4	−1.5
5	−1.0
6	−0.3
7	0.8
8	2.1
8.5	2.9
9	3.6

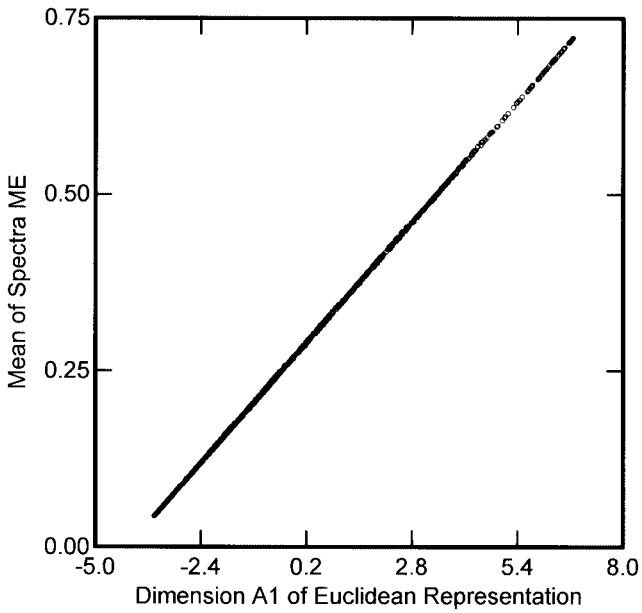

FIG. 3. Scatter plot of the first axis, A1, of the Euclidean representation and the mean Value ME of the 231 spectral points characterizing the corresponding Munsell color sample.

Representation of Color Samples and Spectra into a Common Space

In the preceding analysis, we cannot visualize the relationship of the spectral points $s_{i\mu}$ to the configuration $\{P_i\}$. In order to plot color samples and spectra in a common space, the matrix $\mathbf{S}_{1269\times231}$ was reanalyzed by SVD using procedure (2) of Appendix I to obtain an estimated matrix $\hat{\mathbf{S}}_{1269\times231} = \mathbf{P}_{1269\times m}\,\Delta_{m\times m}\mathbf{Q}'_{m\times231}$ reconstructed from *m* basis factors. The matrices **P** and **Q** respectively define colors *j* and spectral values μ in an *m*-dimensional space, and Δ is a diagonal matrix of $\sqrt{\lambda_\alpha}$, $\alpha = 1, 2, \ldots, m$, where λ_α are eigen values of the matrix $\mathbf{S}'\mathbf{S}$. The first three column vectors in $\mathbf{P}\Delta$ were found to correspond fairly closely to A1 to A3. Elements of the first column, $p_{j\alpha}\sqrt{\lambda_1}$, are correlated with the corresponding elements in A1 with $r^2 = 0.9988$. The configuration of colors *j* in Fig. 4 is the plot of $p_{j2}\sqrt{\lambda_2}$ and $p_{j3}\sqrt{\lambda_3}$ that corresponds to the $\{P_j\}$ in the plot of A2 and A3 in Fig. 2. The space spanned by the first three basis factors is a vector space, but the configuration defined in this space and the configuration $\{P_i\}$ defined in a three-dimensional Euclidean space are found to be essentially the same. The curve in Fig. 4 is the plot of the second and third columns of $\mathbf{W} = \mathbf{Q}\Delta$, $q_{\mu2}\sqrt{\lambda_2}$ and $q_{\mu3}\sqrt{\lambda_3}$, where $\mu = 1$

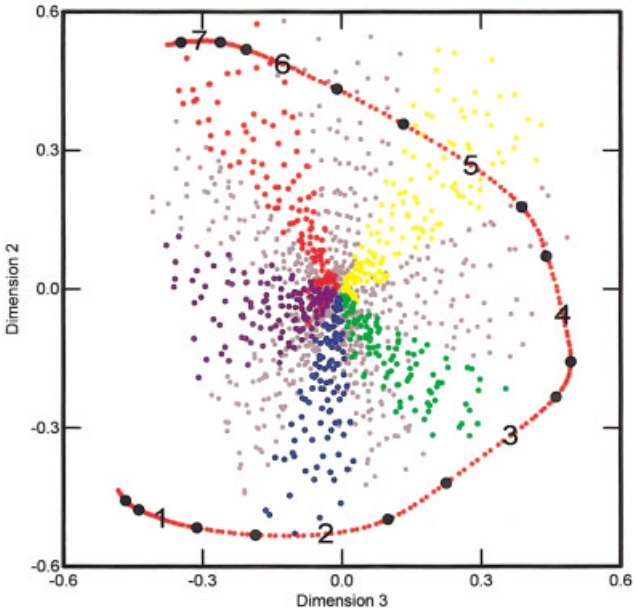

FIG. 4. Plot of 1269 Munsell color chips showing both the color chips and each spectral position from 430 nm to 660 nm obtained from a SVD of original spectra data. The numbers on the spectral line represent interval of 30 nm beginning in the lower left where "1" indicates 460 nm. The black filled circles indicate 10nm intervals. Color coding same as Fig. 2.

(430 nm) to 231 (660 nm). The black dots on the spectral plot are placed at 10 nm intervals with the numbers being at 30 nm intervals beginning with the number "1" at 460 nm.

There are strong potential theoretical implications implicit in the dual plot of color chips and spectral wavelength. One immediate observation is that the projection of a vector from the achromatic point out to the spectral curve provides a prediction about the perceived color of a monochromatic light with a given wavelength. Studies of unique hues provide positive empirical evidence of this correspondence.[8] The fact that the spectral curve does not encircle all the hues is clearly consistent with the well known fact that there is no unique wavelength that produces a perceived color of purple.

We can check on the accuracy of the three dimensional reconstruction of the spectra by reconstructing an estimated $\hat{\mathbf{S}}$ matrix of $m = 3$. The correlation between the 293,139 (1269 $\times$ 231) pairs of values $s_{j\mu}$ in $\mathbf{S}$ and $\hat{s}_{j\mu}$ in $\hat{\mathbf{S}}$ is $r^2 = .988$.

Figure 5 shows a sample of 40 spectra of the Munsell samples all at value level 7 and chroma level 8 in which the original data $s_{j\mu}$ are plotted in blue and the reconstructed values $\hat{s}_{j\mu}$ are plotted in red. Figure 5 begins with the spectra of 2.5R and proceeds through the Munsell color circle ending with 10RP. Note that the progression from one spectrum to another is rather smooth and that the similarity between the first and last spectra appear about the same as between other adjacent pairs throughout the spectra. This illustrates why the colors at the two ends of the spectra merge seamlessly so that there is no physical or perceptual "gap" between so-called low and high wave length colors.

What is shown in Fig. 5 is the similarity between two spectral patterns $s_{i\mu}$ and $\hat{s}_{i\mu}$ for the 40 Munsell colors of 7/8. In order to obtain a visual notion of how close the positions of colors produced by $s_{i\mu}$ and $\hat{s}_{i\mu}$ are, a further analysis was carried out. To a stacked matrix $\mathbf{S}_{80\times231}$ consisting of $s_{i\mu}$ and $\hat{s}_{i\mu}$, we applied procedure (1) to obtain an embedding of the 80 points in a 3-dimensional space $\mathbf{E}^3$. For an extended treatment of scaling stacked matrices, see Romney, Moore, and Brazill.[9] A comparison between points $(a_{i\alpha})$ for $s_{i\mu}$ and points $(a_{i\alpha})$ for $\hat{s}_{i\mu}$, $a = 2, 3$ is shown in Fig. 6. In this plot the symbols for the original spectra are plotted in blue and the symbols for the reconstructed spectra are plotted in red. It may be seen that the pair of points representing the original and the reconstructed are so close that it is difficult to see the relatively small differences.

In theory the figure should be in the form of a smooth closed curve. The deviations from an ideal curve illustrate the presence of error variance from a variety of sources. This variance would include such factors as errors in production of the Munsell color samples, deterioration and soiling of the color samples, possible confusions in labeling, transcription and transmission errors in copying and transferring the original data, and the physical measurement errors in obtaining the spectra, etc.

Relating the Spectral Representation to Psychological Representations

We turn now to the question of how the representation of the color samples shown in the right panel of Fig. 2 relate to previous psychophysical studies of the relations among the colors. In Fig. 7 we display some additional detail of the overall structure by plotting the Munsell Value levels from 3 through 8 separately by level. Each panel is perpendicular to the achromatic point and thus the AA1 axis. The color coding is the same as in Fig. 2. For purposes of comparison, the sample of spectra displayed in Fig. 5 and Fig. 6 is highlighted in black in the lower left panel. The range of views presented in Fig. 7 facilitate the examination of various characteristics of the color space altered to fit the Munsell Model. It is possible to see how the outer limits of chromaticity varies with value. For example, reds have relatively higher Chroma at low Value levels and yellows relatively higher Chroma at higher Value levels.

We ask whether there are changes in hues at different value levels by plotting several possible Chroma levels for the principal hues (5R, 5Y, 5G, 5B, and 5P) at selected Value levels. Figure 8 shows the chromaticity vectors for Munsell Value levels 4, 5, 6, 7, and 8 for the principal hues plotted in blue. The red vectors show the directions of these Munsell principal Hues in the corresponding configuration constructed by multidimensional scaling from human assessment of color differences.[10, 11] The two configurations embedded in E^3, one derived from Euclidean definition of dissimilarities between spectral patterns and one constructed by MDS from color differences assessed by human observers, are close in their structure.

The present results, as shown in Fig. 7, cover a much wider range of the Munsell solid than was possible in the earlier studies. In the MDS studies, the matrix $\mathbf{D} = (d_{ij})$ was incomplete in the sense that only color differences of mod-

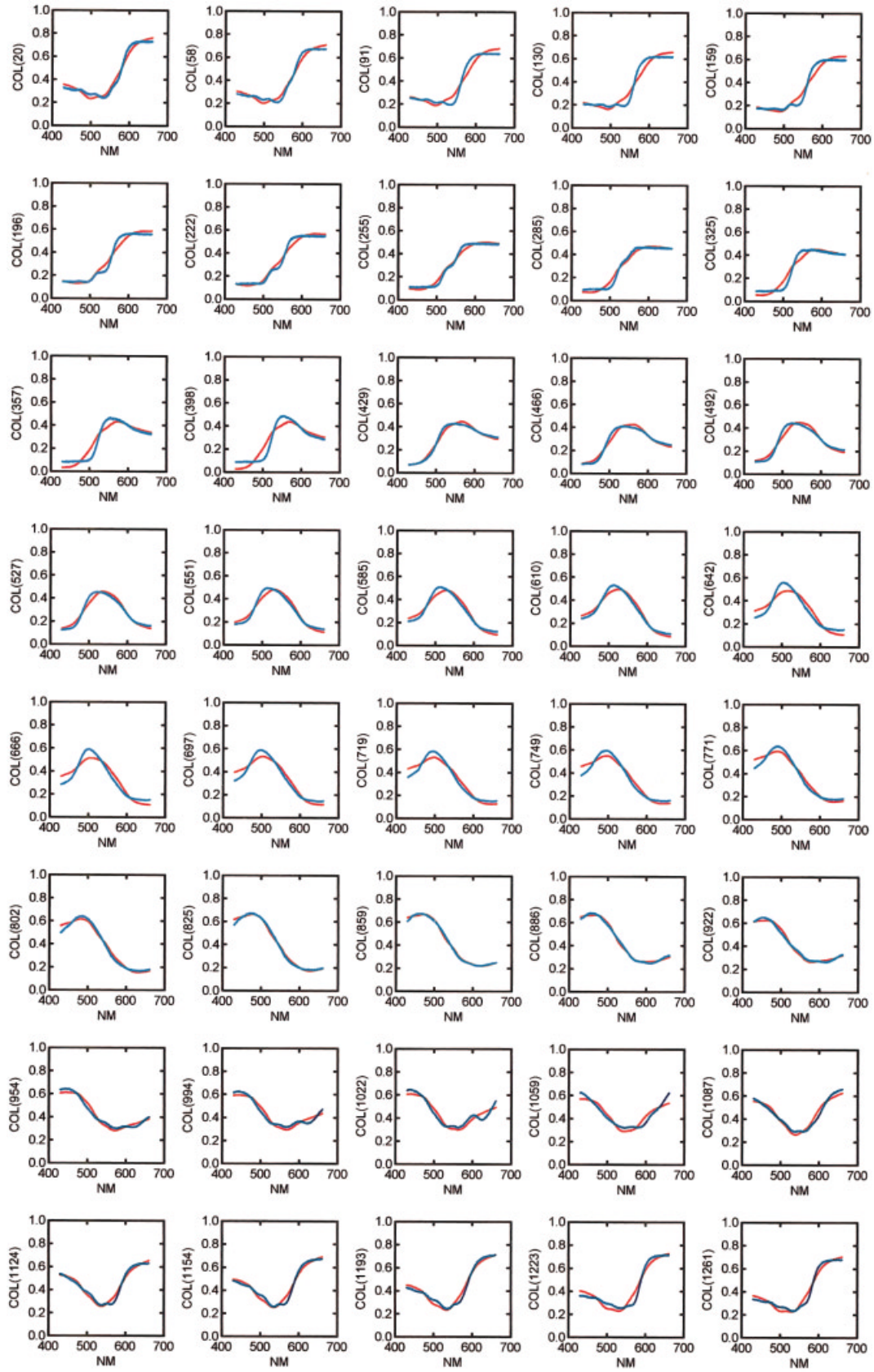

FIG. 5. A sample of 40 spectra (from 430 nm to 660 nm) of Munsell color chips spanning the hue circle at Value 7 and Chroma 8. The blue lines show the actual spectra $s_{i\mu}$ and the red lines show the reconstructed spectra $\hat{s}_{i\mu}$ from three SVD factors. These same color chips are also represented in Figs. 6 and 7.

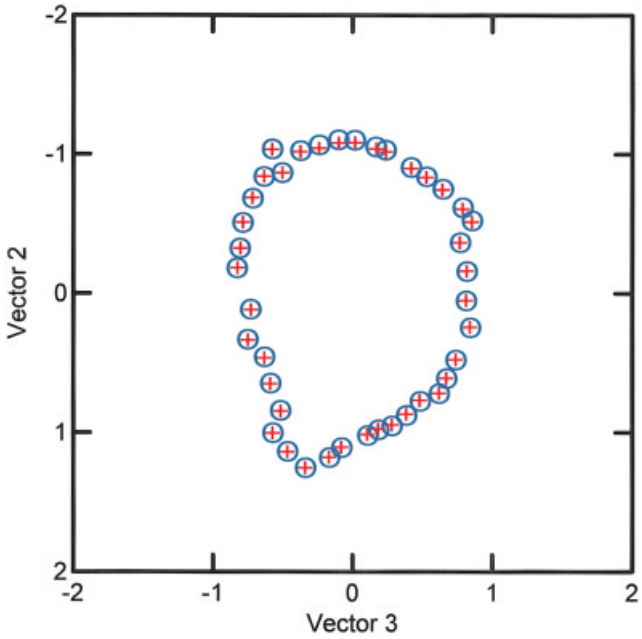

FIG. 6. Plot of 40 pairs of spectra from Fig. 5 representing Munsell color chips at Value 7 and Chroma 8 where the blue circles represent the original spectra and the red crosses represent the reconstructed spectra. The locations of the same chips are highlighted in the lower left panel of Fig. 7.

erate size were assessed because it is not easy to make psychological judgments of color differences beyond a certain range.[12] That is the reason the space embedding the Munsell solid was called a 3-D manifold with locally Euclidean metric.[10] In the present study, **D** is complete in the sense that it includes differences among all Munsell colors. The Munsell color solid as a whole is embedded in E^3. In the MDS experiments, Indow could not show Hue and Chroma configuration in each level of Value as in Fig. 7. We tried to define equi-chroma contours in each plot in Fig. 7. It was found rather complicated and in this article we do not go beyond Fig. 8 in exploring this relationship.

As stated before, the axis AA1 corresponds to Munsell value in the sense that colors of the same Value have almost the same values of AA1. However, this does not imply that AA1 and Value are linearly related. The box plot in Fig. 9 shows the mean (red dots) and scatter of AA1 values in each level of Munsell V. In analogy to the frequently used functional form between value and Y, Value is taken as the dependent variable in the curve[a]

$$V = 5.282(AA1 + 3.203)^{0.335} - 1.6$$

that is plotted in red in Fig. 9.

[a]In accordance with the 1976 CIE lightness function[13] $V = (L^*/10) = 11.6\ (Y/Y_n)^{0.333} - 1.6$ *where* Y_n is the value of Y of the nominally white object, the expression for Value was obtained by fitting the expression $V = a(AA1 + c)^b - 1.6$. The parameter c is necessary because AA1 is defined from an arbitrary origin. It is of interest that the exponent b has turned out close to 0.333 in spite of the fact that AA1 corresponds to the mean spectral energy ME, not to the luminous intensity Y. The parameter a depends upon the unit of scale of the abscissa, V = 10 and 0 when Y/Y_n is 1.0 and 0.003 whereas V = 10 and 0 when AA1 = 8.357 and −3.175.

The scatter of AA1 values at each level of Value is more clearly shown in Fig. 10. Here all color chips at each of the 10 Munsell Hues designated 5 are shown on the x-axis while the AA1 axis is plotted on the y-axis. The horizontal lines of points each represent a Munsell Value level, namely, 2.5, 3, 4, 5, 6, 7, 8, 8.5, and 9. If the relationship between Munsell value and AA1 were perfect, the measures perfect, and the calculations perfect, then each row and column would form a straight line and each cluster of points in the figure would be a single dot. Note that at the lowest value levels there is more variability on the hue axis than the AA1 axis that correlates with Munsell Value, i.e., there is more spreading on the horizontal axis than on the vertical axis at the lower part of the figure. Figure 10 illustrates both the degree of accuracy of the overall picture as well as the correspondence between the structures derived from the Munsell Model and human perceptions. Taken together, these examples are sufficient to show that there is a remarkable fit between the structure of the spectral data and the perceptual structure. We are preparing a detailed analysis of the relationship between the physical and perceptual structures for future publication.

Implications for the Number of Basis Factors Required for Color Description

In this section, we briefly review some previous findings on basis factors after which we present our results and compare them to the prior research. Cohen was the first to reduce spectral reflectance curves $s(\lambda)$ to a small number of basis factors.[14] His data was drawn from $s(\lambda)$ of 433 Munsell chips published by Nickerson[15] in which reflectances are given at 10 nm intervals from 380 nm to 770 nm (40 values). Due to computer limitations, he randomly selected data from 150 Munsell chips and analyzed $\mathbf{S} = (s_{i\mu})$, 150 × 40 with singular value decomposition (SVD) "almost identical" to that discussed by Simons.[16] The only difference seems to be that Simons recommended a "correction by column means $\bar{s}_\mu$" in advance and showed the analysis of $(s_{i\mu} - \bar{s}_\mu)$, whereas Cohen analyzed **S** directly. He extracted four eigenvectors from the minor product moment matrix **S′S,** 40 × 40, and showed that 99.18% of the data was accounted for by the first three. He gave two examples of how well data values $s_{i\mu}$, i = 5R 6/6 and N 4/, were reproduced by loadings of spectra of these colors on the three eigenvectors (as in Fig. 5 of this article).

Maloney[17] repeated Cohen's analysis on s(λ) of 462 Munsell master standard chips obtained from the Macbeth Corporation (400 nm to 700 nm with 10 nm intervals, 31 values) and on $s(\lambda)$ of 337 natural objects collected by Krinov[18] (400 nm to 650 nm with 10 nm intervals, 26 values). His criteria were not only the percentage of data accounted for but also the closeness between the data spectra $s(\lambda)$ and the reconstructed spectra $\hat{s}(\lambda)$. We showed the closeness graphically in Fig. 5 whereas he used R^2 as a measure of fit. He also tried discrete Fourier transformations (DFT) of $s(\lambda)$. On the basis of these two analyses, he

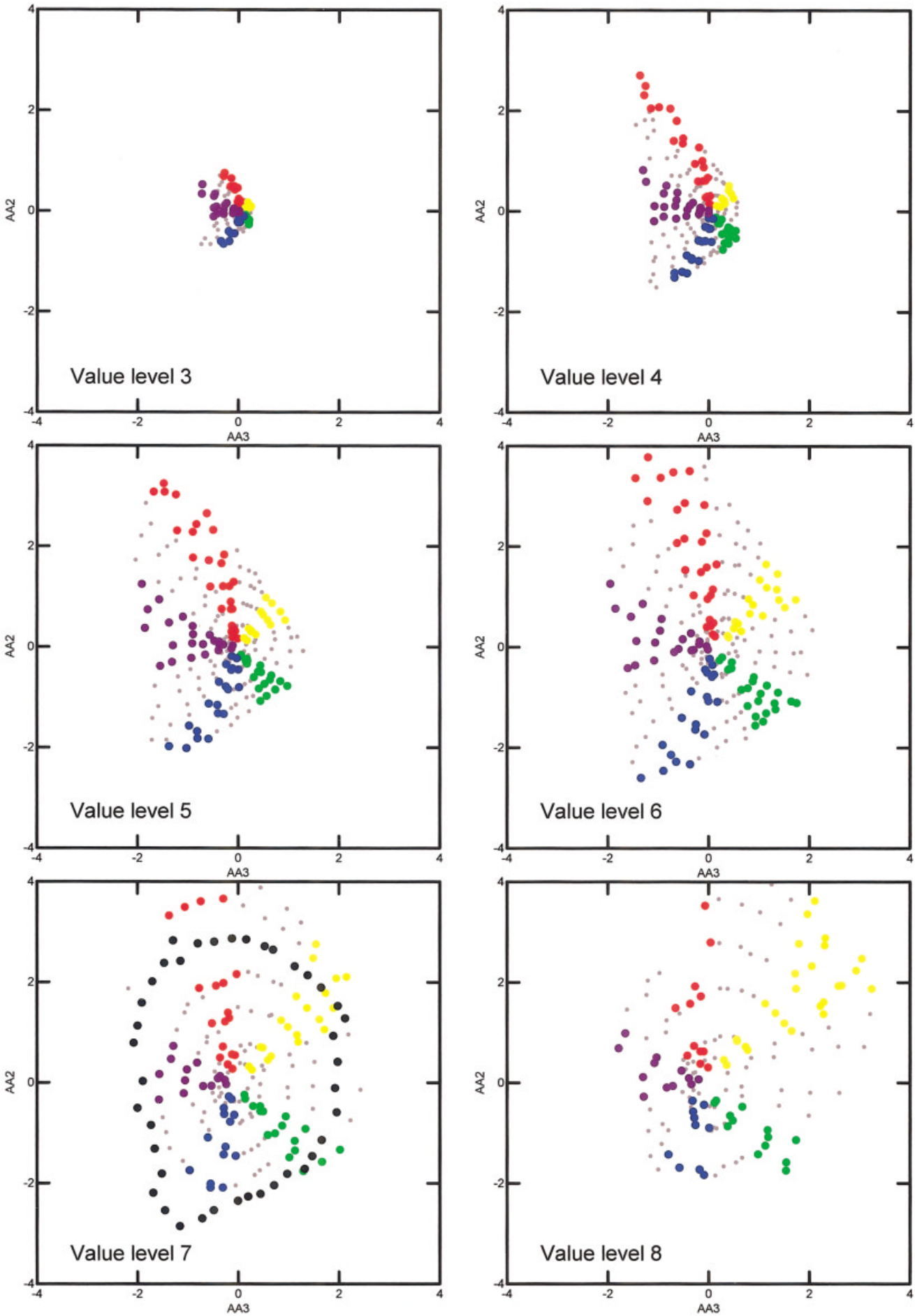

FIG. 7. Plot of the Munsell Model axes AA2 and AA3 at each of 6 Value levels. The sample of 40 spectra at Value 7 and Chroma 8 are highlighted in black in the lower left panel.

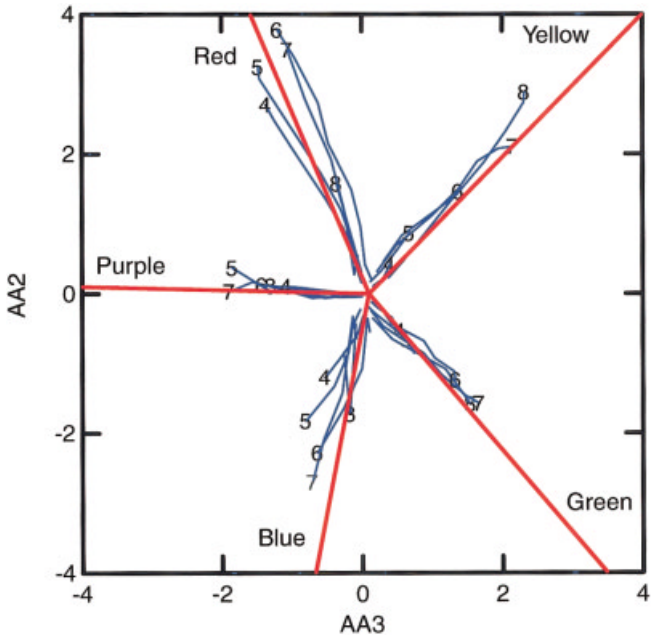

FIG. 8. Plot showing the loci of constant hue in blue at value levels 4, 5, 6, 7, and 8 for the Munsell hues 5R, 5Y, 5G, 5B, and 5P obtained from the plots of Fig. 7. The corresponding vectors of the psychophysical perceptual space obtained by Indow and Aoki[11] and Indow[10] are shown in red.

concluded that five to seven, not three, basis factors were necessary to account for the data. The conclusion is retained in his more recent article.[19]

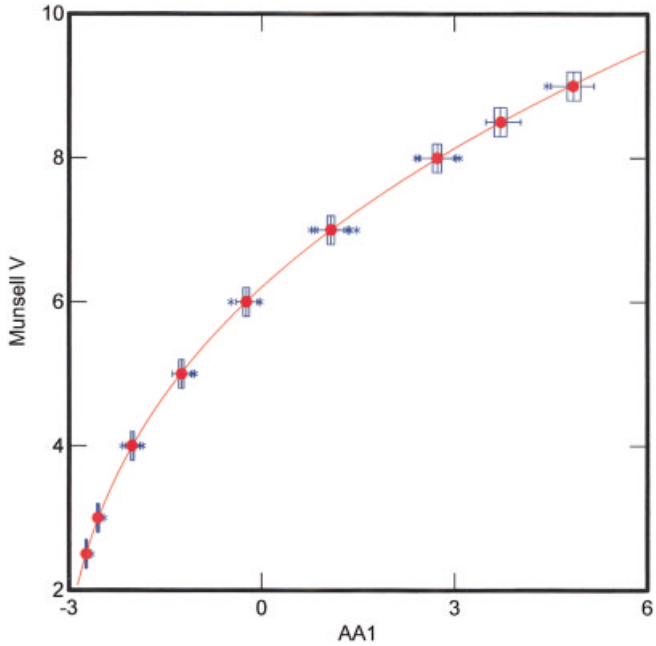

FIG. 9. A standard box plot showing the relationship between Munsell Value levels (the y-axis) and the axis AA1 as in the right panel of Fig. 2 (the x-axis). Values of AA1 of colors of constant Value are not exactly constant. Red dots represent means, boxes contain 50% of cases and median line, whiskers contain the range of values that falls within 1.5 H spread of the hinges, and asterisks show outliers. The fitted red curved line is described in the text.

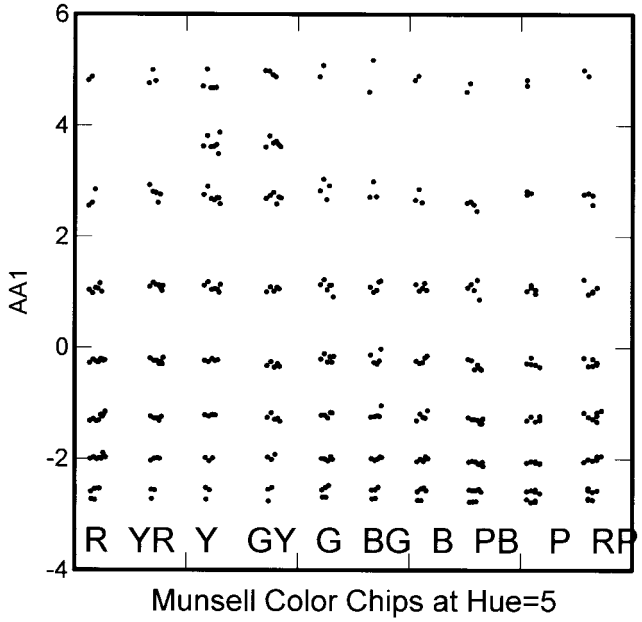

FIG. 10. A plot showing the first axis AA1 of the Munsell Model on the y-axis and the 10 categories of H = 5 of the Munsell color system on the x-axis.

Parkkinens, Hallikainen, and Jaaskelainen[20] measured $s(\lambda)$ of 1257 Munsell matte color samples. Jaaskelainen, Parkkinenn, and Toyooka[21] used these data and a newly added set of data, $s(\lambda)$ of 218 plants. In both sets of data, spectra covered 400 nm to 700 nm with 5 nm intervals (61 values). The analysis of $\mathbf{S} = (s_{i\mu})$ was carried out with SVD without mean correction. They introduced a new criterion, the agreement between $s(\lambda)$ and $\hat{s}(\lambda)$ in terms of CIE(x,y). According to these studies, the number of vectors needed for accurate reconstruction of spectra ranges from three to eight.

Vrhel, Gershon, and Iwan[22] measured 354 spectra $s(\lambda)$ (64 Munsell chips, 120 Dupont paint chips, and 170 natural and man-made objects), where λ covers from 390 nm to 730 nm with 2 nm intervals. Then they reduced these data to 31 values of 10 nm intervals. By the use of the SVD analysis procedure of Trussell[23] with mean correction, they extracted 7 basis factors. They compared data $s(\lambda)$ and $\hat{s}_m(\lambda)$ reconstructed from m basis factors and mean values in two ways. One method is the differences between the two by the CIE $L^*a^*b^*$ (Table II of their article) and the other is plots of $s(\lambda)$ and $\hat{s}_m(\lambda)$ in the Chromaticity diagram, m = 3 to 7.

Following the line advanced by Stiles, Wyszecki, and Ohta,[24] Barlow,[25] and Buchsbaum and Gottschalk[26] pointed out that $s(\lambda)$ can be represented by a finite number n of significant samples where n is related to the range T of λ and the limiting frequency B of Fourier transforms of $s(\lambda)$ and that if T is 300 nm and B is 0.005 cycle/nm, then n = 3. They showed that the gamut spanned by three samples can cover the range of maximum chroma Munsell colors in the CIE chromaticity diagram. However, if B for Fourier transforms of CIE color matching functions is 0.02 to 0.01 cycle/nm, then n becomes 6 to 12. This paradoxical situation was discussed by Maloney.[17,19] In contrast to the studies referred to above, the three-dimensional representation

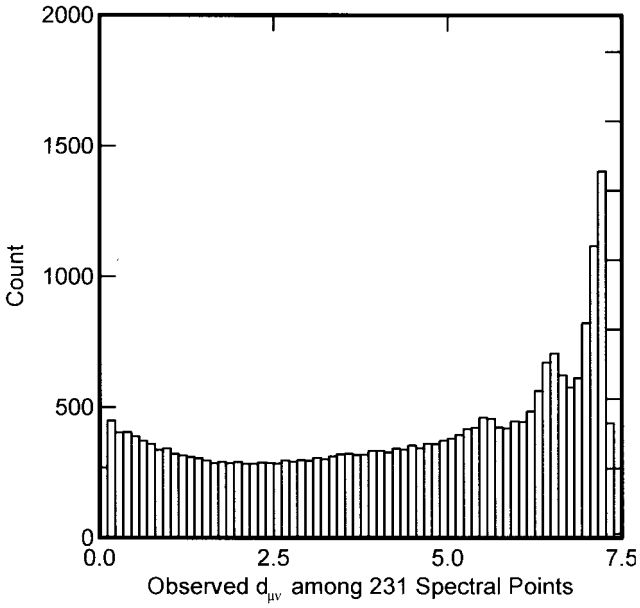

FIG. 11. A histogram showing the distribution of the 26,565 interpoint distances, $d_{\mu\nu}$ among the spectral points as represented in three dimensional Euclidean space.

gave satisfactory results in the present study, which may be due to some differences in data and procedures between the studies. One difference is that our spectral values $s(\lambda)$ are more fine-grained. Spectra were defined with 1-nm interval and the number of $s_{i\mu}$, M, is 231 in each color i. Another difference is that we limited range of λ to between 430 nm to 660 nm. Large discrepancies between $s(\lambda)$ and $\hat{s}(\lambda)$ tend to occur outside this region (e.g., Fig. 4 in Ref. 20). A final difference in procedure that might have an effect is embedding the structure in Euclidean space.

We turn now to the results of the analysis of the spectral information in our original data matrix $\mathbf{S}_{1269\times231}$. The curve from 430 to 660 nm in Fig. 4 shows $W_\alpha = (w_{\mu\alpha})$, each representing the overall weights of 231 wavelength values μ in the respective basis factors α ($\alpha = 2$ and 3). The configuration of 1269 colors in this plot is based on P2 and P3 of SVD analysis obtained by procedure (2) in Appendix I. For comparison we also calculated the Euclidean distances among all wavelengths μ and ν, $\mathbf{D}_{231\times231} = (d_{\mu\nu})$, and used metric scaling to embed them in E^3 by the procedure (3) in Appendix I. The results are given in Appendix II. Fig. 11 shows that histogram of $d_{\mu\nu}$ is U-shaped. The data are well accounted for by the first three factors (97.82%), and $\hat{d}_{\mu\nu}$ based on $m = 3$ are correlated with $d_{\mu\nu}$ with $r^2 = 0.9999$. Cohen[14] extracted four basis factors by SVD directly applied to $(s_{i\mu})$. His first variable I consisting of all positive components represents the general contributions of variables. He reports that the variables II to IV "plots as a near-perfect helix." In our embedding points in E^3 by the procedures (1) and (3), the first three variables exhibit helix patterns. In other words, the curve from 430 to 660 nm in Fig. 4 is a spiral in E^3. The U-shaped histogram of $d_{\mu\nu}$ in Fig. 11 may be understood from the curve in Fig. 4. The excess number of large $d_{\mu\nu}$ represents distances between λ's in the opponent directions in this plot (e.g., λ's in the intervals around "1" and "5") as well as distances between $s_{i\mu}$ and $s_{i\nu}$ at opposite extremes of dimension 1. The large number of small $d_{\mu\nu}$ represents distances between λ's within short intervals at both ends in the plot (e.g., λs below "1" and above "6").

In order to see whether the method of calculation to obtain $\mathbf{A} = (a_{\mu\alpha})$ made much difference in the shape of the basic variables, two more analyses of $\mathbf{S} = (s_{i\mu})$ were added according to the procedures in (4) of Appendix I. Thus, we have four sets of curves representing the overall effects of λ on the basic variables: 1) ($w_{\mu2}$, $w_{\mu3}$, $w_{\mu4}$) of W in the SVD procedure (2). As in the analysis of Cohen,[14] the first vector W1 in this analysis represents the general contributions of variables and all $w_{\mu1} > 0.2$) ($a_{\mu1}$, $a_{\mu2}$, $a_{\mu3}$) in E^3 by the procedure of (3). 3) ($a_{\mu2}$ $a_{\mu3}$ $a_{\mu4}$) by PCA based on $\mathbf{C}$ of the procedure (4). 4) ($a_{\mu2}$, $a_{\mu3}$ $a_{\mu4}$) by PCA based on $\mathbf{R}$ of the procedure (4). In 3 and 4, the first vector a_1 is the direct current component (ME or general intensity) and all $a_{\mu1} > 0$. Since these analyses result in different magnitudes, we standardized the scores for comparison. Results are illustrated in the three panels of Fig. 12. All four curves are

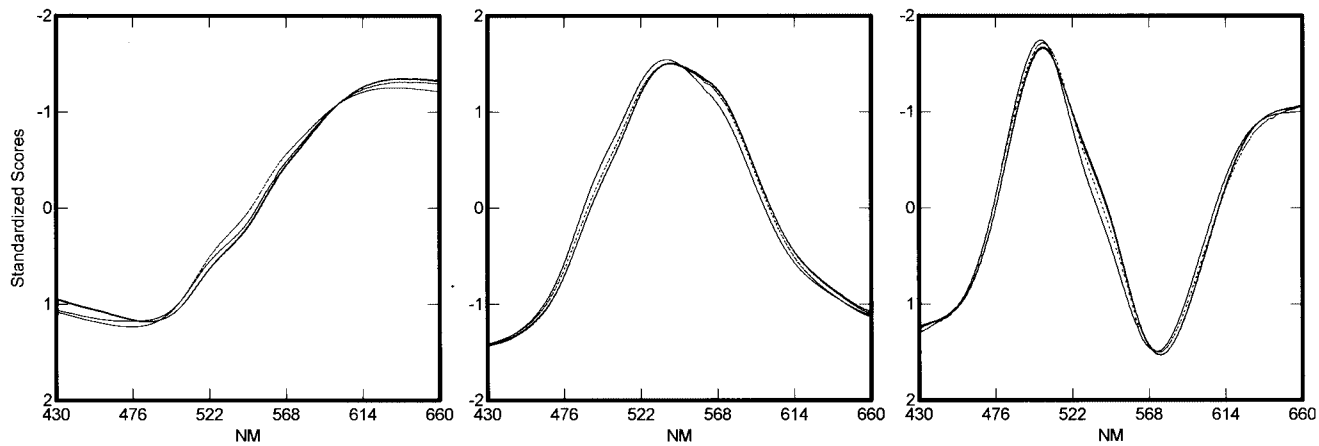

FIG. 12. A plot showing a comparison of four separate methods for computing basis factors using standardized scores of the first three basis factors.

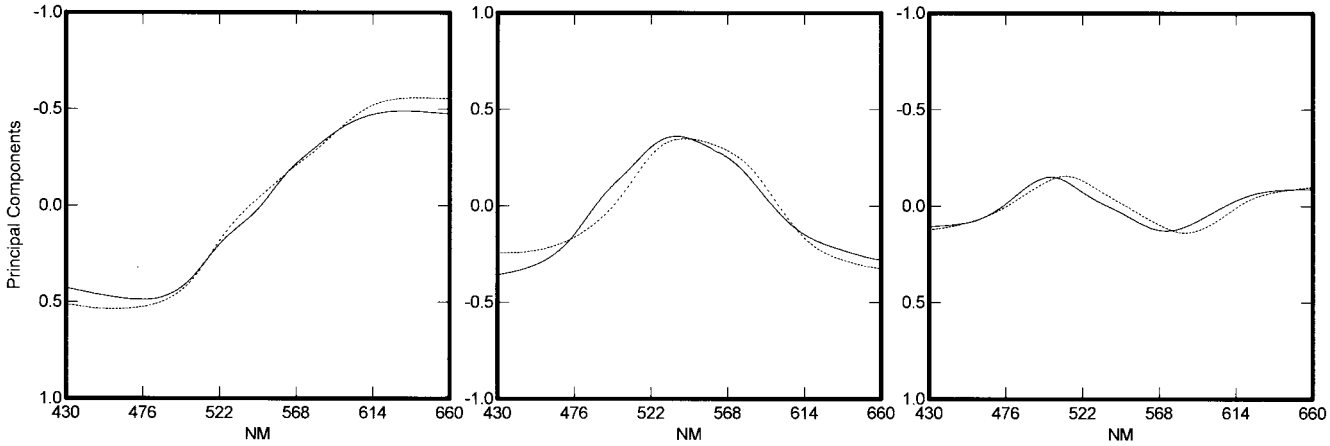

FIG. 13. A plot showing a comparison of the first three basis factors for two separate data sets. The present data set is shown in thick lines, while the Vrhel, Gershon, and Iwan data set is shown in thin lines. Both were obtained with a principal component analysis of the correlations.

correlated above .99. The figure shows that there is no practical effect of choosing one method over any other as they all are virtually identical.

The question of how similar basis factors are when computed on different data sets seems relevant to the question of how many to retain. To obtain information on this question, we downloaded from the web http://ftp.cns.nyu.edu/pub/ltm/SSR/ the data $s(\lambda)$ of Vrhel, Gershon, and Iwan[22] from 390 to 730 nm with 2 nm intervals. These data were also analyzed by Maloney.[19] We censored the set limiting it to the range of 430 nm to 660 nm and analyzed $\mathbf{S}_{354\times116} = (s_{i\mu})$ by PCA based on the correlation matrix $\mathbf{R}$ of the procedure (4) in Appendix I. We then compared the results with those from the same procedures on the Munsell data set. The results are shown in Fig. 13 where the thick line is the Munsell data and the thin line is the Vrhel, Gershon, and Iwan[22] data. The results show fundamental similarities but more variation between these samples of data than when computed on the same data by different methods of analysis as in Fig. 12.

We applied the same test used in Fig. 6 to $\mathbf{S}_{354\times116} = (s_{i\mu})$ of Vrhel, Gershon and Iwan[22] with 2 nm intervals. Namely, we first obtained $\hat{\mathbf{S}}_{354\times116} = (\hat{s}_{i\mu})_{\mathrm{m}}$ by the SVD procedure (2) in Appendix I, based on m basis factors. We then stacked $\mathbf{S}$ and $\hat{\mathbf{S}}_{\mathrm{m}}$ into a single matrix of 708×116 and computed Euclidean distances d_{ij} among all 708 rows to obtain (a_{i1}, a_{i2}, a_{i3}), coordinates of 708 points in E^3, by the procedure (1) of Appendix I. Fig. 14 shows the configuration in the plane of the second and third dimensions where a vertical slash represents (a_{i2}, a_{i3}) for 354 observed spectra and a horizontal slash represents $(a_{i2}, a_{i3})_{\mathrm{m}}$ for 354 reconstructed spectra. Figure 14a is the case that $m = 3$ and Fig. 14b is the case that $m = 4$. These are analogs of Fig. 7 and Fig. 8, respectively, in Vrhel, Gershon and Iwan.[22] They converted row vectors $\mathbf{s}_{\mathrm{i}}$ and $\hat{\mathbf{s}}_i$ to points in the CIE chromaticity diagram while we show the points in the plane of the second and third dimensions in which Euclidean distances between corresponding points represent similarity of $\mathbf{s}_i$ and $\hat{\mathbf{s}}_i$. Squared correlations r^2 between $\mathbf{s}_i$ and $\hat{\mathbf{s}}_i$ based on 4 basis factors approach 1 beyond six decimal places. When $\hat{\mathbf{s}}_i$ is defined by 3 basis factors (Fig. 14a), some discrepancies are visible. In this view the $\hat{\mathbf{s}}_i$, represented by horizontal slashes, are nearer the center of the plot than their actual counterparts $\mathbf{s}_{\mathrm{I}}$, indicting that more basis factors are necessary to provide a closer fit. When $\hat{\mathbf{s}}_i$ is defined by 4 basis factors (Fig. 14 b), points are virtually on top of each other. This is the situation we observed with 40 Munsell colors in

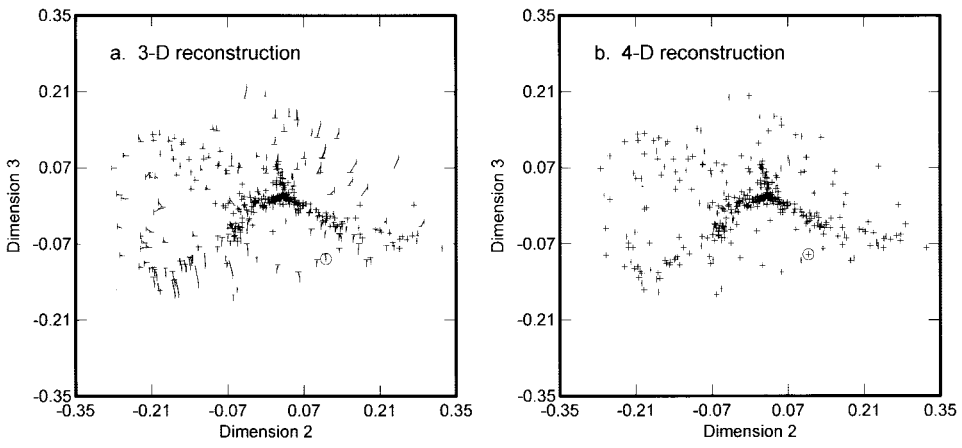

FIG. 14. Plots showing the relation between the location of the actual Euclidean location of the 354 samples from Vrhel, Gershon, and Iwan and the location based on SVD reconstructions of spectra (14a is based on three SVD factors and 14b is based on four SVD factors), plotted as a vector from the actual to the reconstructed position. A vertical slash represents the actual location and the horizontal slash represents the reconstructed location. Sample number 77 is shown in a circle below and to the right of the center of the plot.

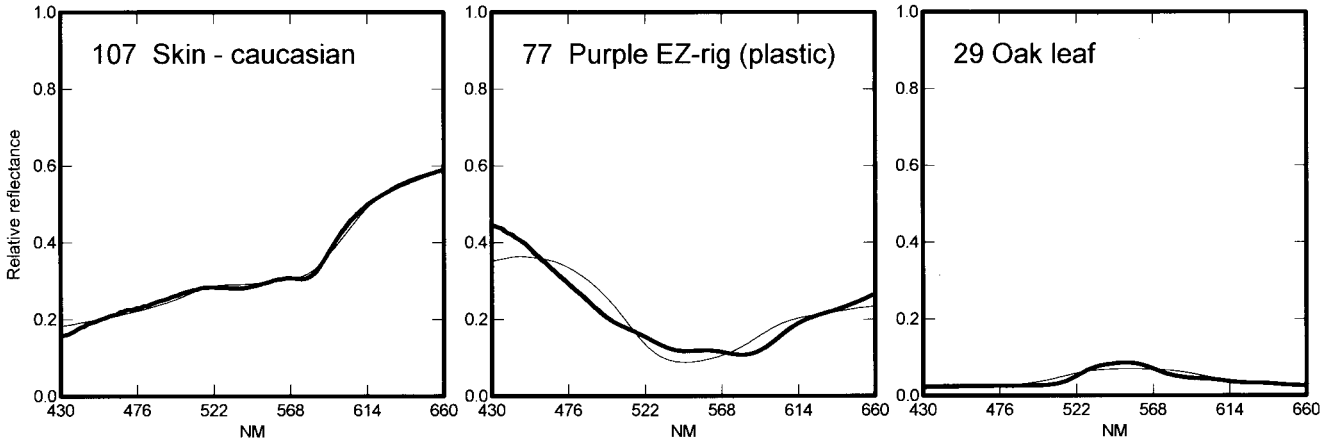

FIG. 15. A plot showing a comparison between the actual spectra (thick lines) and the four dimensional reconstructed spectra (thin lines) of three samples analyzed by Maloney. Sample 77 is highlighted with a circle in Fig. 14.

Fig. 6 with 3 basis factors. The methods of embedding spectral data in low-dimensional Euclidean space seem to provide rather precise representation for spectra in general (the data includes 170 natural and man-made objects) and is not limited to just the Munsell system.

It is interesting to note that the entire Munsell sample of 1269 spectra are fit about as well with three SVD factors as the Vrhel, Gershon, and Iwan sample is with four SVD factors. For the Munsell sample, the vectors corresponding to the observed and reconstructed spectra all have r^2 approaching 1 beyond five decimal places and the points plot virtually on top of each other (plot not shown). Apparently the variability within a heterogeneous sample that includes a variety of natural objects requires an additional dimension for full characterization.

For the set of 1269 Munsell samples, three basis factors are sufficient. For the set of Vrhel, Gershon and Iwan,[22] one more basis factor seems to be necessary if we want to have agreement as shown in Fig. 14b. This difference may be due to the following fact. As can be seen in Fig. 5, spectral patterns $s(\lambda)$ of the Munsell chips have only one hump, upward or downward, in the spectral range of 430 nm to 660 nm. The set of Vrhel, Gershon and Iwan consists of 64 Munsell chips, 120 Dupont paint chips, and 170 natural and man-made objects and there may be spectral patterns $s(\lambda)$ of more complex form. Thus, it is an open question whether three basis factors are sufficient for Munsell colors having $s(\lambda)$ of more complex form (e.g., displays on a monitor) when analyzed in the same way.

Maloney[19] shows patterns of $W_{\mu\alpha}$, $\alpha = 1$ to 8, that he obtained in his analysis of $\mathbf{S}_{170\times171} = (s_{i\mu})$ of Vrhel, Gershon and Iwan[22] in order to have perfect fit between $s(\lambda)$ and reconstructed $\hat{s}(\lambda)$ in the range from 390 nm to 730 nm. For $\alpha > 4$, $W_{\mu\alpha}$ have many humps in the range from 430 nm to 660 nm. If $\mathbf{D} = (d_{ij})$ of Equation (3) in Appendix I for the Munsell colors is analyzed directly, we could have extra-factors having similar form of $W_{\mu\alpha}$. However, we could have almost perfect reproduction of 1269 Munsell samples in E^3 disregarding these extra basis factors. The most important criterion to see how many basis factors are necessary is to see coincidence between the locations of colors and reconstructed colors. Maloney[19] presents comparisons of reconstructing spectra with three and eight basis factors (see Figs. 20.2 and 20.3 of Ref. 19) by plotting observed spectra and reconstructed spectra for each comparison. Our Fig. 15 shows our reconstruction based on four basis factors for three of his examples. The fit in sample 77 (purple EZ-rig) appears to be rather poor. The position of this sample in Fig. 14 is shown by a circle (below and to the right of the center). The discrepancy between $s(\lambda)$ and $\hat{s}(\lambda)$ is not reflected in the plots in Fig. 14. No discrepancy is visible in either Fig. 14a or Fig. 14b. The implication of this is that, even though the spectral profiles may differ in detail, they may code similar locations of the colors represented.

DISCUSSION

Cohen concluded his well cited (over 100) two-page classic with the sentence, "It is odd that the Munsell reflection data should collapse to rank three, and perhaps odder still that the vectors plot in a structure so elegant."[14] We find an equally elegant representation of the Munsell color chips in 3-dimensional Euclidean space. In the Euclidean space, the first dimension is virtually identical to the mean of the spectra characterizing the chip. When the first axis of the Euclidean representation is rotated (37.8 degrees) and shifted to represent the Munsell Model, the structure corresponds almost perfectly to the Munsell color solid. The first axis corresponds to Munsell value and the second and third axes map the Munsell hue circle with high accuracy while Munsell chroma is represented as distance from the achromatic point. Preliminary investigations show a high degree of similarity between this spectrally derived structure and previous studies of perceptual similarities among the colors.[10–12,27]

The overall structures that we have described may be analyzed with linear methods that are simple and well known. The findings would appear to have implications for further research relating physical and perceptual structures. The findings also place constraints on a variety of practices and theories. For example, our demonstration that color

space can be adequately represented in low dimensional Euclidean space sugests that it is unnecessary to use more complicated spatial models in higher dimensions. Our demonstration that four basis factors are sufficient in general, and that three are sufficient for the Munsell colors, has implications for practices in this regard.

ACKNOWLEDGMENTS

We thank Robert Newcomb for writing programs for rotating the Euclidean structures, and we thank Ti-Lien Hsia and Joel Schwarzbart for writing the scaling programs. Markku Hauta-Kasari kindly provided additional information about the spectrophotometer measurements. Roy D'Andrade and Carmella Moore provided very helpful advice at many stages of the research.

APPENDIX I. CALCULATION PROCEDURES FOR SINGULAR VALUE DECOMPOSITION AND EMBEDDING COLORS IN EUCLIDEAN SPACE

Define a data matrix $\mathbf{S} = (s_{i\mu})$ with N colors, $i = 1, 2, \ldots, N$ and M values on the spectrum, $\mu = 1, 2, \ldots, M$. A matrix of this form can be reproduced by the triple product of a singular value decomposition (SVD).

$$\hat{\mathbf{S}}_{N\times M} = \mathbf{P}_{N\times m}\Delta_{m\times m}\mathbf{Q}'_{m\times M} \tag{1}$$

where Δ is diagonal matrix with $\sqrt{\lambda_\alpha}$ on the diagonal, 0 elsewhere, $\alpha = 1$ to m, $m = N$ or M at most, and $\mathbf{P}'\mathbf{P} = \mathbf{I}$ and $\mathbf{Q}'\mathbf{Q} = \mathbf{I}$ with $\mathbf{I}$ equal to an identity matrix and the prime indicates the transpose of a matrix. The data matrix $\mathbf{S}$ was analyzed in two ways, both based on SVD.

(1) Embedding of Colors in an m-Dimensional Euclidean Space E^m (Metric Multidimensional Scaling)

Logic. Configuration of N points $\{P_i\}$ in E^m is given by the N by m coordinates $\mathbf{A} = (a_{i\alpha})$ on orthogonal coordinate axes $\mathbf{f}_\alpha$. Suppose N vectors from the origin of $\mathbf{f}_\alpha$ to P_i, $i = 1, 2, \ldots, N$, then scalar products of vectors, b_{ij}, form a matrix $\mathbf{B}$ that may be written as

$$\mathbf{B}_{N\times N} = (b_{ij}) = \mathbf{A}\mathbf{A}' \tag{2}$$

Define eigenvalues of $\mathbf{B}$ as λ_μ, $\mu = 1, 2, \ldots, m$, in the order of magnitude. Then, m columns of $\mathbf{A}$ are given by m eigenvectors of $\mathbf{B}$ if the unit of N elements in each column is defined to satisfy $\sum_i a_{i\mu}^2 = \lambda_\mu$, $\mu = 1, 2, \ldots, N$. If we take $m = N$ and $\mathbf{A}$ is $N \times N$ and Eq. (1) holds strictly, and if $\lambda_1, \lambda_2, \ldots, \lambda_m > 0$ and λ_{m+1} on, are close to 0, then $a_{i\mu} \approx 0$ for $\mu > m$. Now $\mathbf{B}$ can be reproduced by $\hat{\mathbf{B}}$ defined by $\mathbf{A} = (a_{i\alpha})$, $\alpha = 1, 2, \ldots, m$, $m < N$, which means that $\{P_i\}$ is in a m-dimensional subspace of E^m. The total sum of squared coordinates equals the trace of $\mathbf{B} = \sum_\mu \lambda_\mu$ and the coefficient $\sum_{\alpha=1}^{m} \lambda_\alpha$/trace ($\mathbf{B}$) corresponds to the portion of $\{P_i\}$ embedded in E^m. The axes $\mathbf{f}_\alpha$ are defined one by one to include the largest portion of $\{P_i\}$, and hence $\mathbf{f}_\alpha$ thus obtained are defined only for this purpose. Once $\mathbf{A}$ is obtained and the structure of $\{P_i\}$ becomes apparent, $\{P_i\}$ can be represented by any coordinate axes that are defined by an orthogonal rotation of $\mathbf{f}_\alpha$ and translation. The new set of axes and coordinates can be denoted as $\mathbf{f}_\alpha$ and $\mathbf{A}$ again because Equation (2) still holds with these.

Steps of Analysis. 1. Compute matrix of Euclidean distances using the formula:

$$d_{ij} = \sqrt{\sum_{\mu=1}^{M} (s_{i\mu} - s_{j\mu})^2} \tag{3}$$

2. $\mathbf{B}_{N\times N} = (b_{ij})$ may be obtained from $\mathbf{D}_{N\times N}$ by

$$b_{ij} = \frac{1}{2}\left(\frac{1}{N}\sum_i d_{ij}^2 + \frac{1}{N}\sum_j d_{ij}^2 - \frac{1}{N^2}\sum_i\sum_j d_{ij}^2 - d_{ij}^2\right),$$

where the origin of $\mathbf{f}_\alpha$ is at the centroid of $\{P_j\}$. In this analysis $\mathbf{B}$ corresponds to $\mathbf{S}\mathbf{S}'$ of Eq. (1).

3. Calculate $\mathbf{A} = (a_{i\alpha})$, m eigenvectors of $\mathbf{B}_{N\times N}$, $\alpha = 1$ to m, where

$$\sum_i a_{i\alpha}^2 = \lambda_\alpha, \sum_i a_{i\alpha}a_{i\beta} = 0 \text{ and } \sum_i a_{i\alpha} = 0.$$

In this analysis $\mathbf{A}$ corresponds to $\mathbf{P}\Delta$ of Eq. (1).

4. $\hat{\mathbf{D}}_{N\times N} = (\hat{d}_{ij})$ reproduces $\mathbf{D}$ where $\hat{d}_{ij} = \sqrt{\sum_\alpha (\alpha_{i\alpha} - \alpha_{j\alpha})^2}$

(2) SVD Analysis of S

Logic. N points are directly represented by coordinates $\mathbf{P} = (p_{i\alpha})$ and spectral values $\mathbf{W} = (w_{\mu\alpha})$ in m-dimensional vector space R^m with

$$\hat{\mathbf{S}}_{N\times M} = \mathbf{P}_{N\times m}\,\mathbf{W}'_{m\times M}.$$

The analysis starts with minor product moment matrix $\mathbf{S}'\mathbf{S}_{M\times M}$ and take its eigenvalues λ_α. If we take $m = M$ then $\hat{\mathbf{S}}_M = \mathbf{S}$. If $\lambda_1, \ldots, \lambda_m > 0$ and λ_{m+1} on, are close to 0, then $\mathbf{S}'\mathbf{S}_{M\times M}$, and the coefficient $\sum_{\alpha=1}^{m} \lambda_\alpha$/trace of $\mathbf{S}'\mathbf{S}$ gives the degree of reproducibility of $\mathbf{S}$ by $\hat{\mathbf{S}}_m$.

Steps of Analysis. 1. Obtain $(\mathbf{S}'\mathbf{S})_{M\times M}$ and its eigenvalues of significantly positive values, λ_a, and corresponding eigenvectors, $q_{\alpha\mu}$. Then $\mathbf{Q}$ in Eq. (1) is given as $\mathbf{Q}_{M\times m} = (q_{\mu\alpha})$ with $\alpha = 1$ to m and $\mathbf{Q}'\mathbf{Q} = \mathbf{I}$.

2. Compute $\mathbf{W}_{M\times m} = \mathbf{Q}\Delta$, the spectral forms of m basic variables, where Δ is a diagonal matrix with $\sqrt{\lambda_\alpha}$ on diagonal, 0 elsewhere,

$$\sum_\mu w_{\mu\alpha}^2 = \lambda_\alpha \text{ and } \sum_\mu w_{\mu\alpha}w_{\mu\beta} = 0$$

In this analysis $\mathbf{W}$ corresponds to $\mathbf{Q}\Delta$ of Eq. (1)

3. Compute $\mathbf{P}_{N\times m}$ in Eq. (1), $\mathbf{P}_{N\times m} = \mathbf{S}\mathbf{Q}\Delta^{-1}$ and $\mathbf{P}'\mathbf{P} = \mathbf{I}$.

4. Compute $\hat{\mathbf{S}}_m = \mathbf{P}\Delta\mathbf{Q}'$ or $\mathbf{P}\mathbf{W}'$ where $p_{i\alpha}$ defines the positions of colors i in this m-dimensional vector space.

(3) Embedding of Wavelength Values in Space E^3

Define a matrix of Euclidean distances $\mathbf{D}_{M\times M}$ using the formula

$$d_{\mu\nu} = \sqrt{\sum_{i=1}^{N} (s_{i\mu} - s_{iv})^2} \qquad (4)$$

and analyze $\mathbf{D}_{M\times M}$ in the same way as in (1) to obtain $A = (a_{\mu\alpha})$, $\alpha = 1, 2, \ldots, m$. Then, the reconstructed distances are given by $\hat{d}_{\mu\nu} = \sqrt{\sum_{\alpha} (a_{\mu\alpha} - a_{\nu\alpha})^2}$.

(4) Principal Component Analyses (PCA) of S

Define by **C** and **R** the covariance matrix and correlation matrix of **S** over *N* colors. Both are $M \times M$, and each can be analyzed by the standard procedure of principal component analysis to yield $\mathbf{A} = (a_{\mu\alpha})$, $\alpha = 1, 2, \ldots, m$ (1's on the principal diagonal in **R**). In both results, $m = 4$ and the first vector $\mathbf{a}_1$ represents the general contribution of λ (direct current component) and all $a_{\mu 1}$ are positive.

APPENDIX II. SUMMARY STATISTICS FROM REPORTED CALCULATION PROCEDURES

Euclidean Scaling and Singular Value Decomposition Analyses

	1269 Munsell Chips (1)*		231 Spectral Points (3)		1269 by 231 Matrix (2)	
	Singular Value	Cum. %	Singular Value	Cum. %	Singular Value	Cum. %
1.	8913.29	77.75	1905.71	70.02	177.08	64.05
2.	1730.24	92.84	667.08	94.53	41.65	79.15
3.	637.66	98.40	89.67	97.82	25.52	88.35
4.	82.22	99.12	39.04	99.25	9.16	91.66

Correlations Between Input Matrix and Reconstructed Matrix by Dimension

	1269 Munsell Chips (1)	231 Spectral Points (3)	1269 by 231 Matrix (2)
1. Dimension	0.93945	0.93322	0.88374
2. Dimensions	0.99121	0.99829	0.96506
3. Dimensions	0.99971	0.99956	0.99389
4. Dimensions	0.99994	0.99992	0.99754

*Note: Parentheses indicate procedure from Appendix I used in calculations.

Cumulative Percent of Singular Values Accounted for by Dimension for Basis Factors

Study or Method	1. Dim.	2. Dim.	3. Dim.	4. Dim.
PCA Correlations (4)	0.776	0.930	0.988	0.995
PCA Covariance (4)	0.832	0.995	0.999	1.000
Cohen (1964) SVD	0.927	0.972	0.992	0.997
Maloney (1986) SVD	n.a.	0.958	0.992	0.996

Note: None of the SVD calculations applied mean corrections to data matrix.

Descriptive Statistics for Computed d_{ij} and $d_{\mu\nu}$, and 3-D Differences between d and $\hat{d}$

Observed Distances	Number	Mean	S.D.	Min	Max
Munsell Chip d_{ij} s (1)	804,546	3.690	2.096	0.000	10.337
Munsell ($d_{ij} - \hat{d}_{ij}$)	804,546	0.035	0.052	−0.000	0.985
231 Spectra $d_{\mu\nu}$ s (3)	26,565	4.293	2.287	0.048	7.409
231 Spectra ($d_{\mu\nu} - \hat{d}_{\mu\nu}$)	26,565	0.085	0.082	0.003	0.461

NOTE ADDED IN PROOF

While in press two items have appeared that are highly relevant: Romney, AK, Indow, T. Estimating physical reflectance spectra from human color-matching experiments. P Natl Acad Sci 2002; 99:14607–14610 and Romney, AK, Indow, T. A model for the simultaneous analysis of reflectance spectra and basis factors of Munsell color samples under D65 illumination in three-dimensional Euclidean space. P Natl Acad Sci 2002; 99:11543–11546.

1. Koenderink JJ, van Doorn AJ. The structure of colorimetry. In: Sommer G, Zeevi YY, editors. Algebraic frames for perception-action cycle. New York: Springer; 2000. p 69–77.
2. Lenz R, Meer P. Non-Euclidean structure of the spectral color space. In: Marszalec EA, Trucco E, editors. EUROPTO: Conference on Polarization and Color Techniques in Industrial Inspection. Munich: SPIE Proceedings, vol. 3826, 1999. p 101–112.
3. Burns, SA, Cohen, JB, Kuznetsov EN. The Munsell color system in fundamental color space. Color Res Appl 1990;15:29–51.
4. Carroll JD, Kumbasar E, Romney AK. An equivalence relation between correspondence analysis and classical metric multidimensional scaling for the recovery of Euclidean distances. Brit J Math Stat Psych 1997;50:81–92.
5. Gower JC. Some distance properties of latent-root and vector methods in multivariate analysis. Biometrika 1966;53:325–338.
6. Torgerson WS. Multidimensional scaling: I. theory and method. Psychometrika 1952;17:401–419.
7. Torgerson WS. Theory and Methods of Scaling. New York: Wiley; 1958.
8. Kuehni RG. Determination of unique hues using Munsell color chips. Color Res Appl 2001;26:61–66.
9. Romney AK, Moore CC, Brazill TJ. In: Blasius J, Greenacre M, editors. Visualization of categorical data. San Diego: Academic Press; 1998. p 329–345.
10. Indow T. Multidimensional studies of Munsell color solid. Psychol Rev 1988;95:456–470.
11. Indow T, Aoki N. Multidimensional mapping of 178 Munsell colors. Color Res Appl 1983;8:145–152.
12. Indow T. Preditions based on Munsell notation. I. Perceptual color differences. Color Res Appl 1999;24:10–18.
13. Wyszecki G, Stiles WS. Color science: concepts and methods, quantitative data and formulae, 2nd Edition. New York:Wiley; 1982.
14. Cohen J. Dependency of the spectral reflectance curves of the Munsell color chips. Psychonomic Sci 1964;1:369–370.
15. Nickerson D. Spectrophotometric data for a collection of Munsell samples. Washington D.C.: U.S.Department of Agriculture, 1957.
16. Simons JL. Application of characteristic vector analysis to photographic and optical data. J Opt Soc Am 1963;53:968–974.
17. Maloney LT. Evaluation of linear models of surface spectral reflectance with small numbers of parameters. J Opt Soc Am A 1986;3:1673–1683.
18. Krinov EL. Spectral'naye otrazhatel'naya sposobnost'prirodnykh obrazovanii. Izd Akad Nauk USSR 1947; translated by G. Beldov. Spectral reflectance properties of natural formations, 1953; Ottawa, Canada: National Research Council of Canada, Technical translation: TT-439.

19. Maloney LT. Physics-based approaches to modeling surface color perception. In: Gegenfurtner KR, Sharpe LT, editors. Color vision: from genes to perception. Cambridge: Cambridge University Press, 1999. p 387–416.
20. Parkkinen JPS, Hallikainen J, Jaaskelainen T. Characteristic spectra of Munsell colors. J Opt Soc Am A 1989;6:318–322.
21. Jaaskelainen T, Parkkinen J, Toyooka S. Vector-subspace model for color representation. J Opt Soc Am A 1990;7:725–730.
22. Vrhel MJ, Gershon R, Iwan LS. Measurement and analysis of object reflectance spectra. Color Res Appl 1994;19:4–9.
23. Trusell HJ. Applications of set theoretic methods to color systems. Color Res Appl 1991;16:31–41.
24. Stiles WS, Wyszecki G, Ohta N. Counting metameric object-color stimuli using frequency-limited spectral reflectance functions. J Opt Soc Am A 1977;67:779–784.
25. Barlow HB What causes trichromacy? A theoretical analysis using comb-filtered spectra. Vision Res 1982; 22:635–643.
26. Buchsbaum G, Gottschalk A. Chromaticity coordinates of frequency-limited functions. J Opt Soc Am A 1984;8:885–887.
27. Indow T. Global color metrics and color-appearance systems. Color Res Appl 1980;5:5–12.