Spectral Imaging Target Development Based on Hierarchical Cluster Analysis

Mahnaz Mohammadi, Mahdi Nezamabadi, Roy S. Berns, and Lawrence A. Taplin
Munsell Color Science Laboratory, Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, New York, USA

Abstract

Agglomerative hierarchical cluster analysis was used to group similar spectra from a large database of samples. Based on angles between reflectance vectors of members of a cluster, a reflectance vector was selected as representative of that cluster. Representative samples were grouped together and stored as new calibration targets. Simulated wide-band imaging with glass filters was performed using these new calibration targets and a transformation matrix from digital signals to reflectance was derived. Different verification targets were reconstructed using the transformation matrix; the spectral and colorimetric accuracy of the reconstruction was evaluated. It was shown that beyond a threshold number of samples in the calibration target, the performance of reconstruction became independent of the number of samples used in the calculation. The average spectral RMS for a calibration target consisting of 24 samples selected based on clustering were found to be less than 3.2% for GretagMacbeth ColorChecker DC, GretagMacbeth ColorChecker Rendition Chart, and Esser Test Chart TE221.

I. Introduction

The spectral estimation accuracy of a multi-channel visible spectrum imaging (MVSI) system is dramatically impacted by the calibration target. The idea of selecting a set of spectral reflectance factor values as calibration targets among a series of samples was the basic idea of this research. Agglomerative hierarchical cluster analysis was employed as a statistical technique to achieve this goal. In previous research, the hierarchical clustering method proposed by Wan and Kuo[1] extracted color features based on different color spaces for image retrieval. The focus of this paper is on spectral cluster analysis. There are several common calibration targets available for this purpose, for example, the GretagMacbeth ColorChecker DC (CCDC), the GretagMacbeth ColorChecker Rendition Chart (CC), and the Esser Test Chart TE221 (Esser). These calibration targets have 240, 24, and 283 samples, respectively. These targets were designed with different criteria than required for spectral estimation. For example, the range of pigmentation of the CCDC is limited, particularly for blues. In much of the research performed previously,[2,3] a separate target of 56 blues using artist acrylic paints and including cobalt blue and ultramarine blue was used as a calibration target along with the CCDC. A similar target was used in this research. Different sets of spectral reflectance values were extracted from a larger set of reflectance values that consisted of CCDC, CC, Blues, and Esser, using agglomerative hierarchical cluster analysis. Vector analysis was employed for selecting the representative samples of each cluster. The extracted samples were used as calibration targets in a simulated wide-band imaging with glass filters.

Compared to a principal components method, cluster analysis has the advantage in that it generates directly a selected set in spectral space that is more intuitive than projection of spectral space into a space of eigenvectors. The eigenvector space usually has fewer dimensions than the original spectral space and variability is not equally distributed among each dimension; thus is less intuitive in this sense.

II. Cluster Analysis

Cluster analysis (CA) is a way of grouping on the basis of similarities or distances (dissimilarities). The inputs required are similarity measures or data from which similarities can be computed using different distance functions.[4] In this research, correlation distance was employed as the distance function. The dissimilarity between two reflectance vectors, $\mathbf{R}'_r = [R_{r1}, R_{r2}, ..., R_{rp}]$ and $\mathbf{R}'_s = [R_{s1}, R_{s2}, ..., R_{sp}]$ of the samples based on their correlation distance was calculated using equation (1)

$$dist(\mathbf{R}_r, \mathbf{R}_s) = 1 - \frac{(\mathbf{R}_s - \overline{\mathbf{R}}_s)'(\mathbf{R}_r - \overline{\mathbf{R}}_r)}{(\mathbf{R}_r - \overline{\mathbf{R}}_r)'(\mathbf{R}_r - \overline{\mathbf{R}}_r)^{1/2} (\mathbf{R}_s - \overline{\mathbf{R}}_s)'(\mathbf{R}_s - \overline{\mathbf{R}}_s)^{1/2}} \quad (1)$$

where

$$\overline{\mathbf{R}}_r = \frac{1}{n} \sum_p \mathbf{R}_{rp} \quad and \quad \overline{\mathbf{R}}_s = \frac{1}{n} \sum_p \mathbf{R}_{sp}$$

The p in equation (1) is for wavelength ranging from 380 (nm) to 750 (nm) with 10 (nm) intervals. The next step is linking the pair of objects that are in close proximity. There are several clustering algorithm; in this research the *Agglomerative Hierarchical Clustering* was employed.

Agglomerative Hierarchical Clustering Method

The agglomerative hierarchical method starts with individual objects. The most similar objects are first grouped, and these initial groups are merged according to their similarities. The merging of clusters can be performed under different linkage criteria. These criteria are *single linkage* (minimum distance or nearest neighbor), *complete linkage* (maximum distance or farthest neighbor), and *average linkage* (average distance).[4] In order to select the linkage criteria, the CCDC, CC, Blues, and Esser targets were clustered separately under single, complete, and average linkage method. The performance of clustering of these targets was evaluated using *Cophenetic correlation coefficient*, r_{coph},[5] *Cophenetic correlation coefficient* refers to the correlation between the actual entity correlations in the similarity matrix and the predicted values from the hierarchical tree, called a dendrogram. The range of a positive correlation is $0 \le r_{coph} \le 1$. The results of comparison of the three linkage methods are summarized in Table I. This table shows that the *single linkage method* had poor performance except for the Blues target, which includes only blue pigments. Since the performance of the *average linkage method* was better than the *complete linkage method,* the former was selected as the linkage method.

Table I. Comparison between performances of the three linkage methods based on Cophenetic correlation coefficient (single, complete, and average).

	Cophenetic correlation coefficient			
	CC	Blues	CCDC	Esser
Single	0.47	0.76	0.47	0.35
Complete	0.74	0.78	0.65	0.76
Average	0.74	0.82	0.76	0.80

The distance between two clusters using the average linkage method is the average distance between all pairs of items. The mathematical expression for distance between the two clusters is

$$dist(r,s) = \frac{1}{n_r n_s} \sum_{i=1}^{n_r} \sum_{j=1}^{n_s} dist\left(\mathbf{R}_{ri}, \mathbf{R}_{sj}\right) \qquad (2)$$

where n_r and n_s are the number of reflectance vector (object), $\mathbf{R}_{ri}$ and $\mathbf{R}_{sj}$ are i*th* and j*th* reflectance vector in cluster *r* and *s*. After proper merging of all groups, a hierarchical tree would be formed and clustering can be interrupted at any defined threshold.

The next step was selecting a representative spectral reflectance of each cluster. Vector analysis was employed for this purpose.

III. Vector Analysis

The angle θ between two vectors, $\mathbf{R}'_r = [R_{r1}, R_{r2}, ..., R_{rp}]$ and $\mathbf{R}'_s = [R_{s1}, R_{s2}, ..., R_{sp}]$, in a plane can be represented as the inner product of them divided by the product of the length of the two vectors.[4] The mathematical expression of this definition is

$$\cos(\theta) = \frac{\mathbf{R}'_r \cdot \mathbf{R}'_s}{L_s L_r} \qquad (3)$$

where

$$L_s = \sqrt{\mathbf{R}'_s \cdot \mathbf{R}_s}$$
$$L_r = \sqrt{\mathbf{R}'_r \cdot \mathbf{R}_r}$$

We first define the $i \times i$ angle matrix Θ for each cluster:

$$\underset{(i \times i)}{\Theta} = \begin{bmatrix} 0 & \theta_{1,2} & \theta_{1,3} & . & \theta_{1,i} \\ \theta_{2,1} & 0 & \theta_{2,3} & . & \theta_{2,i} \\ \theta_{3,1} & \theta_{3,2} & 0 & . & \theta_{3,i} \\ . & . & . & . & . \\ \theta_{i,1} & \theta_{i,2} & . & \theta_{i,i-1} & 0 \end{bmatrix}$$

This matrix contains the angles between the reflectance vectors included in a cluster. The mean of angles of each vector with the others was calculated. A vector, which had the minimum average angle with the other vectors, was selected as the representative sample of the issued cluster.

IV. Calibration and Verification Targets

The spectral data of all the targets were combined. The combined target has 603 samples and is called Target A. Cluster analysis was performed; the hierarchical tree is shown in Figure 1. A representative sample of each cluster was determined using vector analysis. In order to have different calibration targets for comparing the performance of MVSI, clustering was interrupted at different thresholds. The thresholds used for clustering the dendrogram of 603 samples were determined by visual assessment, with the goal of creating calibration targets with different numbers of samples. A series of targets in the range of 2 to 321 samples was created at the different threshold values. As an independent verification target, some samples were made of typical artist's pigments using the Gamblin Conservation Colors. All the targets were measured using an integrating sphere spectrophotometer, specular component excluded. A target consisting of 20 samples was also made by a random selection from the Esser target (Random) based on a uniform probability distribution. The random sampling procedure was repeated 25 times. Simulated imaging was performed with each of 25 selected targets and the results described herein are the average of the 25.

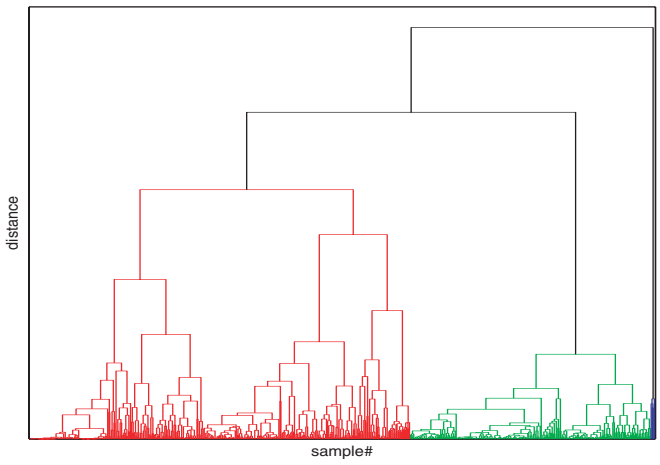

Figure 1. Dendogram of full data set (603 samples).

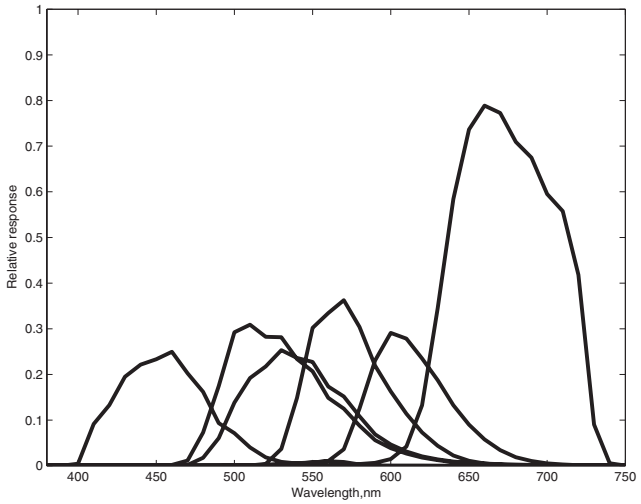

Figure 2. Spectral sensitivity of the designed filters combined with camera spectral sensitivity and IR cut-off filter.

V. Simulated Wide-Band Imaging with Glass Filters

Each of selected targets based on clustering methods were used as a calibration target in a simulated wide-band imaging with glass filters in order to reconstruct the CCDC, CC, Blues, Target A, Esser, and Gamblin targets. Each individual target was also used as a calibration target and the system performance for both cases were compared.

Simulated image acquisition was composed of a high performance Roper Scientific, Inc. Photometric Quantix 6303E that uses a Kodak blue enhanced KAF6303E CCD. The spectral sensitivity of the camera was measured previously.[2,3] A set of six glass filters was optimized for the best colorimetric and spectral performance.[3] The combined response of the camera sensitivity including filter transmittance and IR cut-off filter is shown in Figure 2. Illuminant D65 was used in all computations. The camera system delivers12-bit images, that is digital counts between 0 and 4095. The simulated exposure time for each channel was computed to prevent clipping. For a perfect diffuser digital counts were adjusted to 3800. Hence a security margin below the maximum digital count was reserved for highlights. The digital counts for a pixel with known reflectance factor can be computed using equation (4)

$$\mathbf{D}_i = \sum_{\lambda=380}^{750} \mathbf{R}_\lambda \cdot \mathbf{L}_\lambda \cdot \mathbf{S}_{\lambda,i} \qquad (4)$$

where λ is wavelength, $\mathbf{D}_i$ is the digital count of the i*th* channel, $\mathbf{R}_\lambda$ is spectral reflectance factor of a pixel, and $\mathbf{L}_\lambda$ is relative spectral power distribution of illuminant D65. The $\mathbf{S}_{\lambda,}$ is spectral sensitivity of the camera for the i*th* channel.

The digital counts for simulated image of the CCDC, CC, Blues, Gamblin, Random, Target A and Esser calibration targets were computed using equation (4). A noise matrix consisted of 50 values with normal distribution and zero mean value and standard deviation of 2.5% of the digital count of each sample of each calibration target was generated and added to that sample. In this way each sample has an associated Gaussian noise of 5% of its original digital count values. This amount of noise was more than real measured noise for typical imaging.[2,3] For any calibration target and using a generalized pseudo inverse based on Singular Value Decomposition (applied by PINV function in Matlab), a transformation matrix, **T**, for converting digital counts to reflectance values, was found. Knowing the transformation matrix and a set of digital counts, one can calculate estimated reflectance factor values according to equation (5) for that set of digital counts

$$\hat{\mathbf{R}} = \mathbf{T} \cdot \mathbf{D} \qquad (5)$$

where, $\hat{\mathbf{R}}$ is estimated reflectance value matrix, **D** is the digital count matrix, and **T** is the transformation matrix. Since reflectance values were measured from 380 to 750 (nm) with 10 (nm) intervals, 38 values, and there were six digital counts associated with each reflectance matrix, the transformation matrix was a (38×6) matrix. Therefore at least six simultaneous equations were required at each wavelength to calculate elements of the transformation matrix. Since there are 50 noisy versions of each reflectance sample, the linear equation for determination of transformation matrix was over determined.

The estimated reflectance factor values for calibration targets and independent verification target were calculated using equation (5). The spectral accuracy of the estimated reflectance was analyzed by root mean square error (RMS) of spectral reflectance presented in percentage format. Color

differences between estimated and original samples were also calculated using the ΔE_{00} color difference formula for illuminant D65 and 1931, 2 degree standard observer. As described above, 50 samples with noise added to them were generated for each sample of each target; hence the mean value of ΔE_{00} and spectral RMS were used in performance comparison.

VI. Comparison of Results

The average spectral accuracy is shown in Table II. As expected, the best spectral estimations were obtained when the same target was used for calibration and verification. In the most practical situations calibration and verification targets are not the same but the performance obtained by the same calibration and verification can be interpreted as the upper limit of expected performance. The Gamblin target was reserved as an independent target and was not used in either the clustering process or the calculation of the transformation matrix, **T**.

Table II. Comparison of average spectral RMS%, for different calibration targets in estimation of verification targets. N indicates the number of samples in each target.

			Verification Targets					
			Esser	CCDC	CC	Blues	Gamblin	Target A
		N	283	240	24	56	63	603
Calibration Targets	Esser	283	2.1	2.4	2.6	4.1	3.6	2.5
	CCDC	240	2.5	1.9	2.6	4.8	3.9	2.5
	CC	24	2.5	2.3	2.4	3.6	3.6	2.5
	Blues	56	5.2	4.4	4.8	2.1	6.0	4.5
	Gamblin	63	2.5	2.5	2.8	3.6	3.3	2.6
	Target A	603	2.3	2.2	2.5	3.4	3.5	2.4
	Random	20	2.5	2.7	3.0	4.5	3.9	2.8
	Clusters	2	8.6	8.3	9.1	13.0	10.7	8.9
		4	4.7	4.9	4.9	6.7	6.8	5.0
		6	4.1	3.2	3.6	5.3	5.3	3.8
		9	4.5	3.3	4.0	5.0	5.7	4.0
		12	3.7	3.2	3.4	3.4	4.5	3.5
		14	3.7	3.1	3.3	3.5	4.4	3.4
		16	3.4	2.8	3.2	3.0	4.1	3.1
		18	3.4	2.8	3.2	3.0	4.1	3.1
		24	3.1	2.6	2.7	3.2	3.9	2.9
		28	3.3	2.7	2.9	2.9	4.0	3.0
		48	3.5	2.9	3.0	2.8	4.1	3.2
		71	2.8	2.6	2.8	2.8	3.7	2.7
		127	2.5	2.4	2.7	3.2	3.6	2.5
		168	2.4	2.4	2.6	3.3	3.6	2.5
		203	2.4	2.3	2.6	3.3	3.6	2.5
		281	2.3	2.3	2.6	3.3	3.6	2.4
		321	2.3	2.3	2.6	3.4	3.5	2.4

As can be seen from Table II, the average spectral RMS is lower than 4% for most targets, except the Blue target and those targets containing less than 12 samples. This suggests that the performance of the system becomes independent of the number of samples used in the calibration targets beyond some threshold. In the case of calibration targets selected by the clustering method, this effect is also seen in Figure 3 where the average color difference between each verification target and its corresponding estimate falls rapidly and become constant as the number of samples in the calibration target increases.

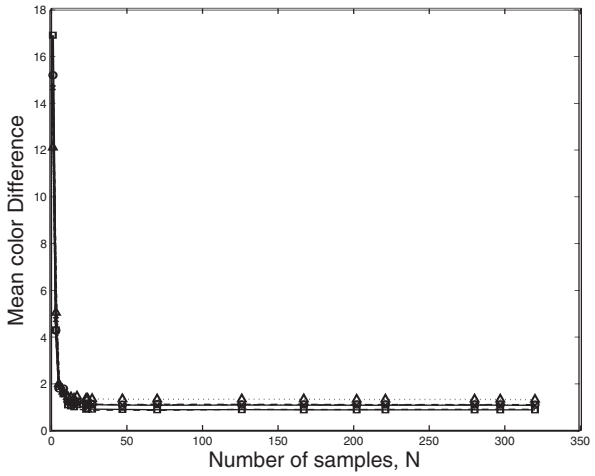

Figure 3. Mean color difference versus number of samples in calibration targets selected based on clustering method. Dashed line with triangle is CCDC target; solid line with square is Esser target; line and circle is CC target; solid line is Blue target; dashed and dot is Gamblin target; solid line and x-mark is Target A.

From Table II and Figure 3 one may suggest that the threshold is equal to the 24 samples in the calibration target. The CC also has 24 samples and can be compared to the selected cluster-based target. An interesting case is the Blue verification target in which all calibration targets based on clustering method and with more than 24 samples have lower average spectral RMS than other calibration targets. For example, a calibration target based on clustering method and consisting of 24 samples had better performance in terms of average spectral RMS than the CCDC with 240 samples, that is, an average spectral RMS of 3.2% and 4.8%, respectively. For the independent verification target, Gamblin, both the CCDC and the cluster-based calibration targets have the same average spectral RMS of 3.9%. Since the Blue target has been used in clustering process with other available targets and the final selected samples have been used in calculation of transformation matrix, it is reasonable to assume that the presence of the Blue target has added some extra piece of information to the transformation matrix. This added piece of the information is not included in the

transformation matrix in the Random target. The average spectral RMS of 4.5% in the estimation of the Blues verification target was obtained for the Random calibration target.

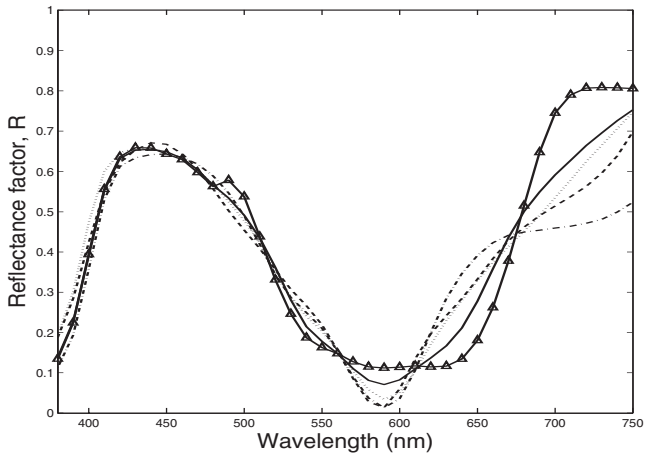

Figure 4. Spectral reflectance of cobalt blue (PB 28) and reconstructed reflectance factor based on different calibration targets. Solid line with triangle is cobalt blue; dashed line is CC target estimate; dash-dot line is CCDC target estimate; dot line is Esser target estimate; solid line is estimate of cluster-based target with 24 samples.

The reconstructed reflectance factors of cobalt blue (PB 28) and cobalt violet (PV 14) pigments are shown in Figures 4 and 5 for different calibration targets, respectively. As can be seen from Figure 4, all the targets have poor performance in the long wavelengths and cannot match the tail pattern of cobalt blue. However, the calibration target based on clustering method had slightly better performance in estimating the tail shape. In the short wavelengths, less than 470 (nm), the target based on clustering has the best matching to the shape of the reflectance curve. Although none of the calibration targets could reconstruct the reflectance curve of cobalt blue perfectly in the area of 500 (nm) to 620 (nm), the cluster-based target had the best fit. As shown in Figure 5, in the case of cobalt violet (PV14) all of four calibration targets were not perfect in reconstructing the reflectance curve in the short wavelength though the general pattern was estimated. All four reconstructed curves have a local maximum at 420(nm) as the original cobalt violet has a peak at the same position. The cluster-based estimation has been successful to catch the minimum peak at the 580(nm). None of calibration targets were able to match long tail of cobalt violet. The limitation in the long wavelength spectrum may be attributed to the camera spectral sensitivities as well as the appreciable noise introduced in the simulated imaging system.

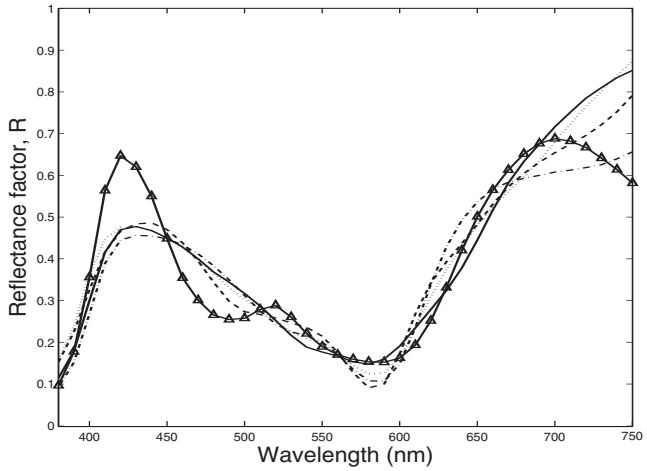

Figure 5. Spectral reflectance of cobalt violet (PV14) and reconstructed reflectance factor based on different calibration targets. Solid line with triangle is cobalt violet; dashed line is CC target estimate; dash-dot line is CCDC target estimate; dot line is Esser target estimate; solid line is estimate of cluster-based target with 24 samples.

The spectral and colorimetric performance of the reconstructed cobalt blue and cobalt violet are tabulated in Table III. The better performance of cluster-based calibration target can be seen in the reconstruction of cobalt blue. That is because the calibration target had a piece of information during the clustering process by using the Blues target for making target A. The performance of the reconstruction of cobalt violet was poor. The results show that having a calibration target containing more samples does not guarantee better performance.

Table III. Comparison between the spectral and colorimetric performance of predicting cobalt violet (PV14) and cobalt blue (PB28) using different calibration target.

Calibration target	Cobalt violet		Cobalt blue	
	$\Delta E_{00}(D_{65})$	RMS%	$\Delta E_{00}(D_{65})$	RMS%
Esser	0.8	7.4	1.0	9.6
CCDC	1.0	7.6	1.2	13.3
CC	0.7	8.8	0.8	8.3
Cluster (24)	0.2	9.0	0.4	5.4

It may be intuitive to know what the simulation program selects, if the program was forced to select just four samples out of 603 available options. Based on clustering logic the four selected samples should differ as much as possible. Figure 6 presents those four samples, which are red, green, blue, and black. This was reasonable since red, green, and

blue are the three most different colors among the available colors; they have the least possible reflectance overlap. The black is a special case that can also be representative of all dark colors.

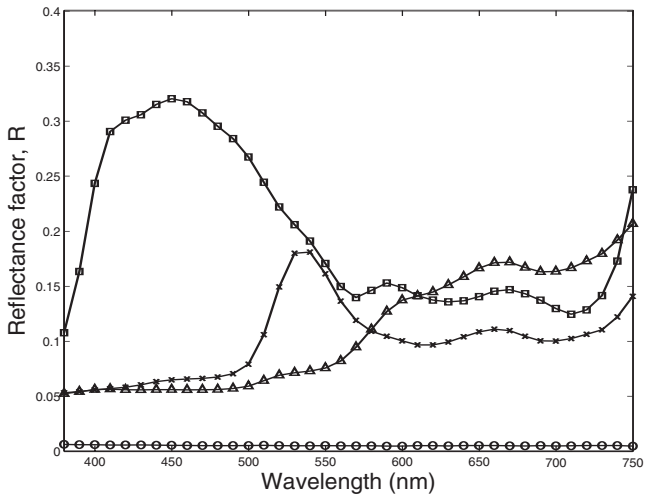

Figure 6. Spectral reflectance factor of selected four representative samples out of 603 available sample options.

VII. Conclusions

The key calculation in spectral imaging is to find a proper transformation matrix from digital signals to spectral reflectance. The transformation depends on the camera design, calibration target samples, and the method of determining the matrix coefficients. It was shown that beyond a threshold number of samples the performance of reconstruction of spectral reflectance became independent of the number of samples. The clustering method was able to reduce the number of samples used for calibration and helped fine-tune the system. In the case of the Blue target, which is strongly biased toward a specific hue, the cluster-based calibration targets had better performance than common calibration targets since the cluster-based calibration targets had additional information during the clustering process. Furthermore, the clustering method can be more intuitive than methods using principal component analysis.

References

1. X.Wan, C. Jay Kuo, A New Approach to Image Retrieval with Hierarchical Color Clustering, *IEEE, **8**, 5, September (1998).*
2. See http://www.art-si.org/ for the listings of relevant publications.
3. S. Quan, Evaluation and Optimal design of Spectral Sensitivities for Digital Color Imaging, PhD thesis, Rochester Institute of Technology, 2002.
4. R.A. Johnson, D.W. Wichern, *Applied Multivariate Statistical Analysis*, Fifth edition, Prentice Hall,(2002).
5. H. Charles Romesburg, *Cluster Analysis For Researchers*, Lifetime Learning Publications (1984).

Acknowledgments

This research was supported by the Andrew W. Mellon Foundation, the National Gallery of Art, Washington, DC, the Museum of Modern Art, New York, and Rochester Institute of Technology,

Biography

Mahnaz Mohammadi received her B.S. degree in Textile engineering from AmirKabir University at Tehran in 1993 and a M.S. in Polymer Engineering with an emphasis in color science from AmirKabir University in 1996. She worked for two years for Iran Color Research Center. Since 2002, she is a Ph.D. candidate in Imaging Science at the Munsell Color Science Laboratory of Rochester Institute of Technology.