Training Set Selection for Multispectral Imaging Systems Characterization

Paolo Pellegri†‡*, Gianluca Novati†‡ and Raimondo Schettini†

†*DISCo, University of Milano – Bicocca, Milano, Italy*
‡ *ITC, National Research Council, Milano, Italy*

Deriving the actual multispectral data from the output of the acquisition system is a key problem in the field of multispectral imaging. Solving it requires a characterization method and the training set (if any) on which the method relies. In this paper we propose three novel approaches in selecting a training set to be used for the characterization of a multispectral acquisition system. The first approach, which we call the Hue Analysis Method, is based on colorimetric considerations; the second and third approaches, which we call the Camera Output Analysis Method and the Linear Distance Maximization Method respectively, are mainly based on algebraic and geometrical facts. In all three cases the selected training sets will have relatively low numerosity and broad applicability. We also test our three approaches, as well as an approach from another author and a random selection method, on the data obtained from a real acquisition. We then compare the reconstructed reflectances with the measurements obtained using a spectrophotometer. Our results indicate that all our methods can be substantial improvements compared to a random selection of the training set, and that the performances of the Linear Distance Maximization Method make it the best choice among all the methods tried for application in a general context.

Journal of Imaging Science and Technology 48: 203–210 (2004)

Introduction

A multispectral imaging system allows the reconstruction of the reflectance in the spectrum of visible light of the objects in a scene.[1,2] Such a system is built around specialized hardware (the acquisition system) that can vary its 'sensitivity' to light with respect to different wavelengths: this is usually accomplished by including in the system a set of traditional optical filters or a tunable filter with variable transmittance, and taking several shots of the scene with a different filter enabled for each shot. The behavior of a generic multispectral acquisition system can be modeled by expressing its output $a_i(\mathbf{x})$ as a function of the relevant parameters of the acquisition process[3]; formally, this can be written as

$$a_i(\mathbf{x}) = \int_{\lambda_1}^{\lambda_2} E(\mathbf{x},\lambda)R(\mathbf{x},\lambda)S_i(\lambda)d\lambda \qquad (1)$$

where i is an index that varies with the filter being used, $\mathbf{x}$ is a two-dimensional coordinate vector identifying a point in the scene, λ is the wavelength, E is the energy that reaches point $\mathbf{x}$ in the scene, R is the reflectance, and S_i is the 'sensitivity' of the system when filter i is enabled. The integration (which is replaced by a summation in actual computations) is performed between λ_1 and λ_2, as the system is assumed to be blind outside this interval.

As Eq. (1) shows, however, the output of the acquisition system is not the actual reflectance data: deriving these data is a key problem in the field of multispectral imaging. To solve this problem, a characterization method must be established and the training set (if any is required) on which the method relies must be chosen. If the system output for point $\mathbf{x}$ is denoted as

$$\mathbf{a}(\mathbf{x}) = [\, a_i(\mathbf{x}) \,]_i, \qquad (2)$$

and M filters are used, then $\mathbf{a}(\mathbf{x})$ is an M-dimensional vector. The corresponding reflectance $R(\mathbf{x}, \lambda)$ is a function of the wavelength λ, but since in practice it is not easy (or even always possible) to give an analytical form to R, a sampling of its value is customarily considered instead. The light spectrum is accordingly sampled at a discrete number of values of λ, and the reflectance is expressed as

$$\mathbf{r}(\mathbf{x}) = [\, R(\mathbf{x}, \lambda_j) \,]_j, \qquad (3)$$

where j is an index that varies with the sample wavelengths. If N sample values of λ are considered, then $\mathbf{r}(\mathbf{x})$ is an N-dimensional vector.

To establish a relationship between the system output and the corresponding reflectance, the system characterization function

$$\mathbf{a}(\mathbf{x}) \mapsto \mathbf{r}(\mathbf{x}) \qquad (4)$$

which links the output $\mathbf{a}(\mathbf{x})$ at a point $\mathbf{x}$ in the scene to the corresponding reflectance $\mathbf{r}(\mathbf{x})$, must be described

Original manuscript received July 1, 2003

* pellegri@itc.cnr.it

Supplemental Material—*Appendix with Tables AI, AII and color version of Fig. 6* can be found on the IS&T website (www.imaging.org) for a period of no less than two years from the date of publication.

or estimated in some way; this is usually done by means of an empirical model based on a chosen training set and characterization method, e.g., linear models[4] or polynomial regression[5]. The quality of the estimation depends on the characterization method and training set selected; it improves if a 'good' training set is available while a 'bad' training set may negatively affect the resulting estimation. Therefore, a certain characterization method may be rejected because it appears to yield poor results, when these are really due to the training set. There are also cases in which the context dictates the use of a specific characterization method; in this event the quality of the approximation is influenced by the training set alone.

Despite all this, the literature to date on multispectral acquisition seems to place little emphasis on the problem of choosing a 'good' training set. Hardeberg[6] employs an algebraic method to select a training set of low numerosity; most other authors either employ as their training set the set of reflectances to be reconstructed,[3,5] or simply avoid the issue. One likely cause for this approach is that many authors work on simulations[7] rather than on real acquisitions. In real acquisitions large training sets are unwieldy, and changing training sets often for different applications is inefficient, so the importance of having a few small training sets of broad applicability is more likely to be stressed.

We present here three different approaches in selecting a 'good' training set from an initial array of available colors (which we call the 'target'). By 'good' we mean that the elements in the chosen training set will be as few as possible and its applicability as broad as possible within the limits suggested by the operational context. The first approach, which we call the Hue Analysis Method, is based on colorimetric considerations; the second and third approaches, which we call the Camera Output Analysis Method and the Linear Distance Maximization Method respectively, are mainly based on algebraic and geometrical facts. We have employed these methods to select different training sets from an initial common target, and tested the corresponding system characterization models on the data obtained in a real acquisition.

Methods for the Selection of Training Sets

The Hue Analysis Method

Ideally, we may expect the 'representativeness' of a training set to improve as the number of 'different' colors included in it increases. Since in this context colors are represented by their reflectances, the feature that most clearly sets them apart from one another is the shape of the reflectance curve. Although in colorimetric terms a reflectance curve subsumes all the characteristics of a color, the property that most directly reflects the shape of the curve is hue.

As the name suggests, with the Hue Analysis Method the selection of the training set is based on hue. Assuming that the reflectances of the colors in the target are known, and a suitable illuminant is chosen, the corresponding $L^*C_{ab}^*h_{ab}$ coordinates are computed[8], and the L^* coordinate is normalized (or simply ignored) so that the colors are projected onto an $C_{ab}^*h_{ab}$ plane. This plane is then divided into n sectors of equal angular width, n being the number of colors to be selected. For each sector, the color inside the sector and nearest (in the sense of angular distance) to the central half-line, i.e., the half-line that cuts the sector in equal halves, is included in the training set, for a total of n colors (see Fig. 1). In this way the chosen colors are as widely spaced as possible and cover the whole plane, i.e., the whole range of hues, avoiding the bias and/or local overfitting that could affect the resulting model were the colors chosen with no regard to their spatial disposition.

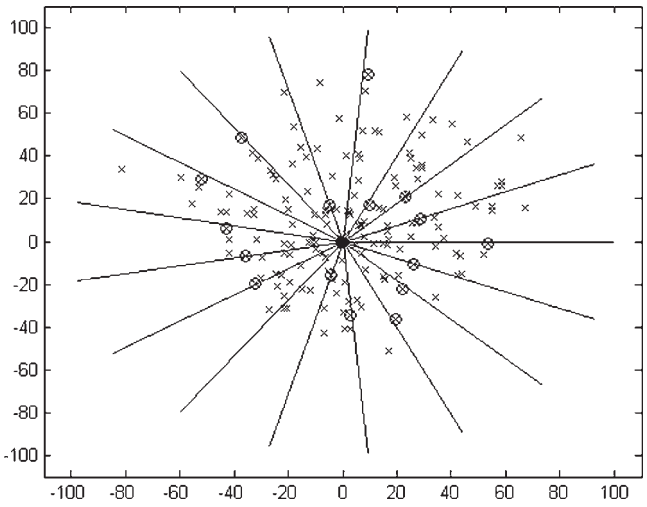

Figure 1. An example of how training set colors are chosen in the $C_{ab}^*h_{ab}$ plane. The central half-lines of all the sectors are drawn, and the chosen colors are circled.

The Hue Analysis Method may pose two geometrical issues, which arise, respectively, when no target colors lie within a sector (see Fig. 2), or when a color is the nearest to both of two neighboring sectors, in which case the color lies on the boundary between the two sectors; while the second case is rare, the first may present more frequently, especially as the number of elements desired in the training set grows larger. Sometimes these obstacles can be overcome by simply rotating the sector boundaries (in any direction), so that they become placed where the sector half-lines were originally placed. This shift preserves the number of sectors, and does not violate the basic ideas of the method, but is not always sufficient; for instance, it does not work if two adjacent sectors are both empty. Another possible workaround is that of lowering the desired numerosity of the training set until all the resulting sectors include at least one target color; however, the numerosity of the selected training set may prove too small for the given context. Anyway, if neither of these two workarounds leads to acceptable results, the Hue Analysis Method cannot be applied. It should also be noted, though, that in view of selecting a training set with broad applicability (if not a 'universal' training set), the initial target should itself include a good choice of hues and range over the whole $C_{ab}^*h_{ab}$ plane; then, if either of the two issues described arises, the possibility of replacing the initial target should be considered.

The Camera Output Analysis Method (COAM)

If the characterization function of the acquisition system can be assumed to be linear, then it can be approximated by an empirical linear model; this model will be a linear function from the M-dimensional vector space of system output vectors to the N-dimensional vector space of (sampled) reflectances. The generic form of the system characterization function given in Eq. (4) then becomes

$$\mathbf{r}(\mathbf{x}) = \mathbf{M}\mathbf{a}(\mathbf{x}), \tag{5}$$

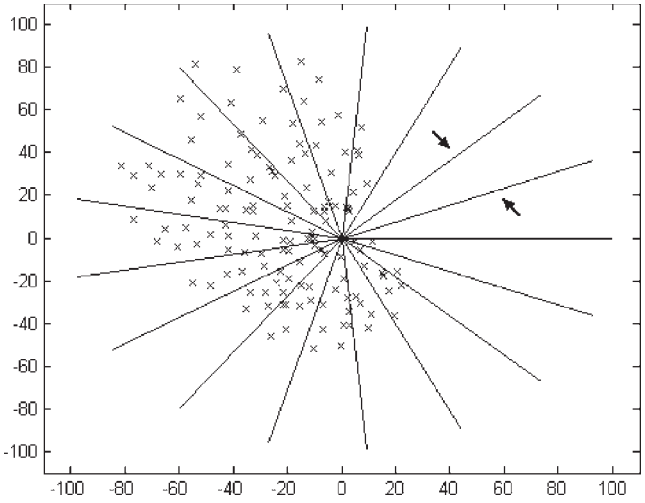

Figure 2. No colors lie within the sectors corresponding to the indicated half-lines: the Hue Analysis Method cannot be applied unless the desired training set numerosity can be lowered.

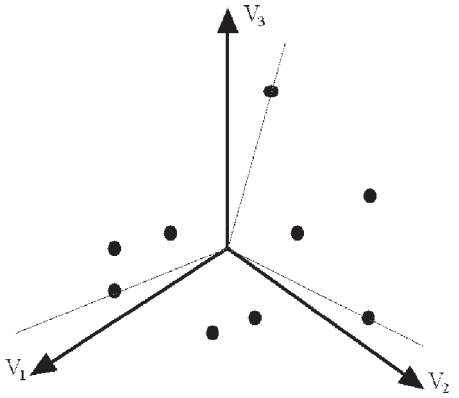

Figure 3. A fictional example of three eigenvectors and nine target colors. The colors nearest to each eigenvector in the sense of angular distance are marked.

where **M** is the matrix associated with the linear model employed to approximate the function.

In this context, for a training set to be representative enough to use to reconstruct the reflectance of any color, then it must span the whole characterization function domain. This means that it must include a subset which is a basis for the vector space of the system output vectors. This requirement can be met by simply including in the training set M colors whose corresponding system output vectors are linearly independent. However, although theoretically sufficient, this approach may not give satisfactory results. As our experimental results show, the characterization models built on 'randomly chosen' training sets of linearly independent colors may be rather imprecise. In our study, an analysis of the spatial distribution of the system output vectors corresponding to the colors in the target showed that, despite the great variety in their colorimetric characteristics, the vectors tended to cluster together, due probably to the contribution of the illuminant used for the acquisition. Therefore, the likely cause of the bad results of randomly chosen training sets is the unavoidable error in measurement which affects the system output vectors. Although the magnitude of this error may be small, if the colors are close to each other, their distances and relative positions may be very sensitive to even small differences. Such differences may cause severe warping in the geometry of the system output space, so that the decomposition of an arbitrary color on the basis of the training set colors is conspicuously incorrect, and this may cause unsatisfactory results when reconstructing the corresponding reflectance.

With the Camera Output Analysis Method we attempt to solve these problems by implementing a strategy to space the colors chosen for the training set well apart from one another. Principal component analysis[9] is applied to all available colors, so that the M principal eigenvectors $\mathbf{e}_i$, $i = 1, \ldots, M$, are identified. Then, for each eigenvector in order of relevance, the color not already chosen 'nearest' to that eigenvector is included in the training set, for a total of M vectors. The distance d used to measure 'nearness' is the cosine of the angle between the system output vector **a** of a color and the eigenvector concerned (see Fig. 3).

$$\mathbf{d}(\mathbf{a}, \mathbf{e}_i) = \frac{\langle \mathbf{a}, \mathbf{e}_i \rangle}{\|\mathbf{a}\| \|\mathbf{e}_i\|}. \tag{6}$$

This strategy tries to maximize the orthogonality of the colors chosen for the training set, so that the impact of measurement errors on the geometry of the space is kept to a minimum. However, if a very tight clustering of the available colors is observed in the system output space, the relative linear distance of the colors chosen for the training set may prove to be more important than their orthogonality.

To provide for this case, we designed a second version of the Camera Output Analysis Method to maximize the relative linear distance of training set colors while still taking into account their anisotropic spatial disposition (as described by the eigenvectors). In this variation all available colors are projected on the basis of the eigenvectors to retrieve their coordinates in the eigenvectors system; for a color with output vector **a**, the corresponding coordinate c_i on eigenvector $\mathbf{e}_i$ will be

$$c_i = \mathbf{a} \cdot \mathbf{e}_i . \tag{7}$$

Then, for each eigenvector in order of relevance, the color not already chosen with the greatest absolute coordinate on that eigenvector is included in the training set, for a total of M vectors (see Fig. 4); this means that the color $\mathbf{a}_0$ chosen for eigenvector $\mathbf{e}_i$, must satisfy the condition

$$\left| \mathbf{a}_0 \cdot \mathbf{e}_i \right| \geq \left| \mathbf{a} \cdot \mathbf{e}_i \right| \tag{8}$$

for every color **a** not already chosen.

There may be cases, however, in which even this approach cannot guarantee that the selected colors will be sufficiently spaced. Particularly, if one considers two eigenvectors whose associated singular values are small, the colors selected in correspondence to those eigenvectors may still be close to each other. This happens because the variability over the set of the target colors of the coordinate associated with an eigenvector decreases as the associated singular value decreases: if two colors lie within the same 'quadrant' with respect to the corresponding eigenvectors, they are likely to be very near to each other.

The third version of the Camera Output Analysis Method provides a simple although partial workaround

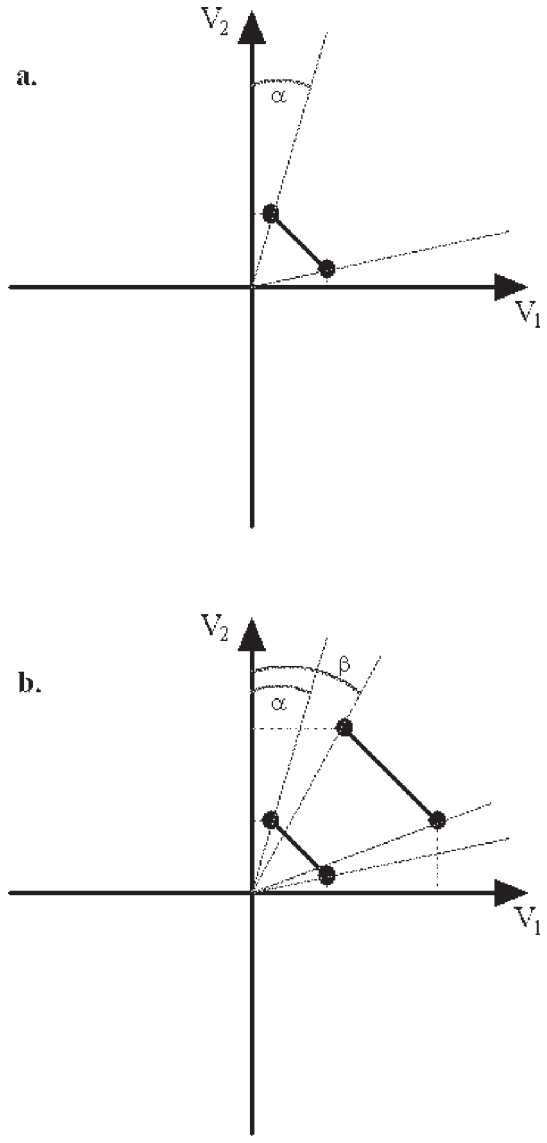

Figure 4. (a) Two colors which are nearest to different eigenvectors in the sense of angular distance may still be close to each other. (b) Colors which have maximum absolute coordinate on those eigenvectors may be more spaced.

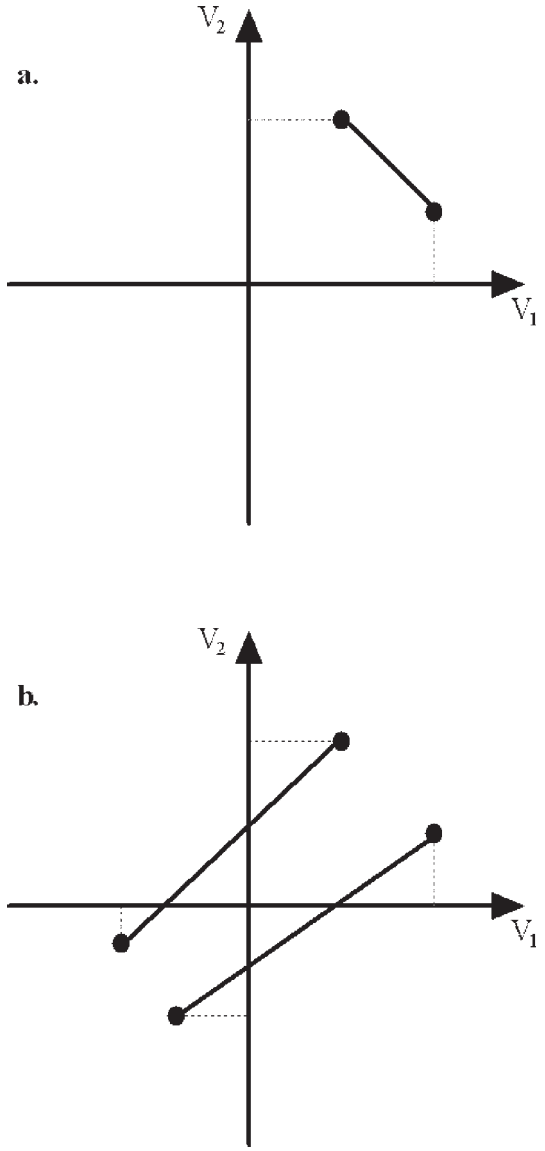

Figure 5. (a) Two chosen colors corresponding to different eigenvectors may still be close using the second version of the Camera Output Analysis Method. (b) Pairs of colors chosen with the third version will usually be more spaced.

to this problem: for each eigenvector, both the color with the greatest coordinate and the color with the smallest coordinate are selected and included in the training set, for a total of 2*M* colors. As Fig. 5 shows, this approach usually guarantees that for each pair of eigenvectors there will be a corresponding pair of training set colors that are more distant from each other than those colors that would be selected using the second version of the method. As a partial drawback, the number of colors included in the training set doubles.

The Linear Distance Maximization Method (LDMM)

As our experimental results will show, the second strategy of the COAM performed better than the first one in our trials. We therefore decided to go further ahead in that direction and investigate the use of the linear distance between system output vectors as the only criterion to choose the training set, with no regard to the principal components. To do so, we devised a third approach, which we call the Linear Distance Maximization Method, in which the colors for the training set are chosen iteratively based on their distance from those already chosen. As the first color $\mathbf{t}_1$ of the training set, which cannot be chosen by comparison with previously chosen colors, we select the color whose associated system output vector has maximum norm among all the target colors, i.e., the brightest color in the target for the chosen acquisition conditions. Then, we select the second color $\mathbf{t}_2$ as that color not already chosen that has maximum distance from the first color. The distance d used is the standard Euclidean distance in $\mathbf{R}^M$ (M is the dimension of the system output vector space). To select the other colors, we proceed iteratively: at step k, for each color $\mathbf{a}$ not already chosen we compute the minimum distance $\delta(\mathbf{a})$ from it to the colors already included in the training set:

$$\delta(\mathbf{a}) = \min_{l=1,\ldots,k-1} \mathrm{d}(\mathbf{a}, \mathbf{t}_l). \quad (9)$$

Then, we select as the kth training set color $\mathbf{t}_k$ the color $\mathbf{a}_0$ that has maximum δ; formally,

$$\mathbf{t}_k = \mathbf{a}_0 \mid \delta(\mathbf{a}_0) \geq \delta(\mathbf{a}) \quad (10)$$

for every color $\mathbf{a}$ not already chosen.

We also devised a second version of the LDMM in which we employed a different distance, the maximum absolute difference of components; for two colors $\mathbf{a}_1$ and $\mathbf{a}_2$, this distance is computed as follows:

$$\mathrm{d}(\mathbf{a}_1, \mathbf{a}_2) = \max_{i=1,\ldots,M} |a_{1i} - a_{2i}|. \quad (11)$$

We employed this distance because we wanted to test an approach in which one large difference in a component would prove more important than several small ones (these would 'add together' if the Euclidean distance were used). In doing so, we attempted to 'weigh' hue more than luminance.

Other Methods

To allow a comparison with our methods, we also considered two other different approaches: a method introduced by Hardeberg,[6] and a biased random selection. The method by Hardeberg selects the colors to be included in the training set based on their reflectances using an iterative process similar to the one we employed in our LDMM method. The color whose reflectance vector has maximum norm among all the target colors is chosen as the first color of the training set. This is the 'most reflective' color available, but it is not necessarily the brightest one for a given environmental setup. All subsequent training set colors are then chosen iteratively: if the colors already selected are indicated as $\mathbf{r}_1$, …, $\mathbf{r}_{k-1}$, then at step k the color $\mathbf{r}$ not already chosen that maximizes the ratio of the smallest to the largest singular value of the matrix $[\mathbf{r}_1 \ldots \mathbf{r}_{k-1}\ \mathbf{r}]$ becomes the kth color in the training set. Compared to our LDMM approach, Hardeberg's method then shows two main differences: first, instead of system output vectors it considers reflectances; second, it uses the ratio of the smallest to the largest singular value and not distance as its selection criterion.

The other selection method we employed is a random method biased to avoid including several similar colors in the training set, as this could lead to local overfitting. Each available color was assigned a probability of being selected based on its similarity to other colors. Similarity was assessed by visual inspection of the target, using hue and then lightness as similarity criteria. Similar colors were grouped together, and the colors within each group were assigned decreasing probabilities starting from a value common to all groups and sufficiently high to allow positive probabilities for all colors in the target. After that, random selection was performed using the assigned probabilities.

Experiments

We have used our three proposed methods to select different training sets from the colors included in the Macbeth ColorChecker DC target (MDC),[10] which contains 177 different color patches (see Fig. 6). Each of the selected training sets was employed to approximate the characterization function of our multispectral acquisition system, and the resulting characterization

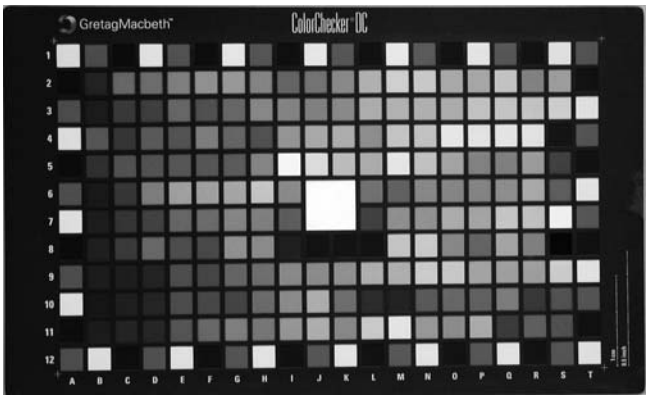

Figure 6. The Macbeth ColorChecker DC target (grayscale).*Supplemental Material—Color version of Fig. 6 can be found on the IS&T website (www.imaging.org) for a period of no less than two years from the date of publication.*

models were then used to reconstruct the reflectances of all the colors in the MDC target from the output data obtained by a real acquisition (see Fig. 7). To allow a comparison with our methods, we also applied the same procedure with some training sets that were selected using the other two approaches we considered above: these are, respectively, the method introduced by Hardeberg, and biased random selection.

For completeness, we added a further trial performed using the entire MDC target as training set. The results obtained in this case could be seen as a sort of 'benchmark', especially in view of assessing whether reduced training sets can decrease the impact of any data noise. However, good results in this trial could simply indicate that the target is indeed 'representative' of the whole color spectrum, as we had assumed, so that considering all colors does not lead to any noticeable local overfitting. Also, the computations involved in this case are much more costly, raising time needed by roughly two orders of magnitude, from tens of seconds to a few hours on an average computer; therefore, using the whole MDC target as training set in actual acquisitions would prove rather unwieldy.

To ensure that all the trials were conducted under the same operating conditions, we performed one acquisition of the whole MDC target and used the results with all the selected training sets. Our multispectral acquisition system consisted of a Photometrics CoolSnap digital camera with a resolution of 1392 by 1040 pixels and a dynamic range of 12 bits (from Roper Scientific), a high quality 28 mm *f*/4 Rodagon lens (from Rodenstock), a VIS2 VariSpec Tunable Filter[11,12] with a nominal bandwidth of 40 nm at 500 nm (from Cambridge Research and Instrumentation), and a cut off optical filter for infrared and ultraviolet radiations (from Andover Corporation). We used 31 different configurations of the Tunable Filter for band selection; these configurations were chosen so that the peak in the filter transmittance varied between 400 nm and 700 nm with steps of 10 nm. Since previous experiments had shown that our system response function was linear, we adopted a linear model to approximate the system characterization function; consequently, each selected training set was used to compute a corresponding linear model, employing singular value decomposition[13] for system inversion. We considered 15 singular values (and the associated basis vectors) when working with reflectances and 23 values when working with system output vectors; these numbers were chosen after an exhaustive search because

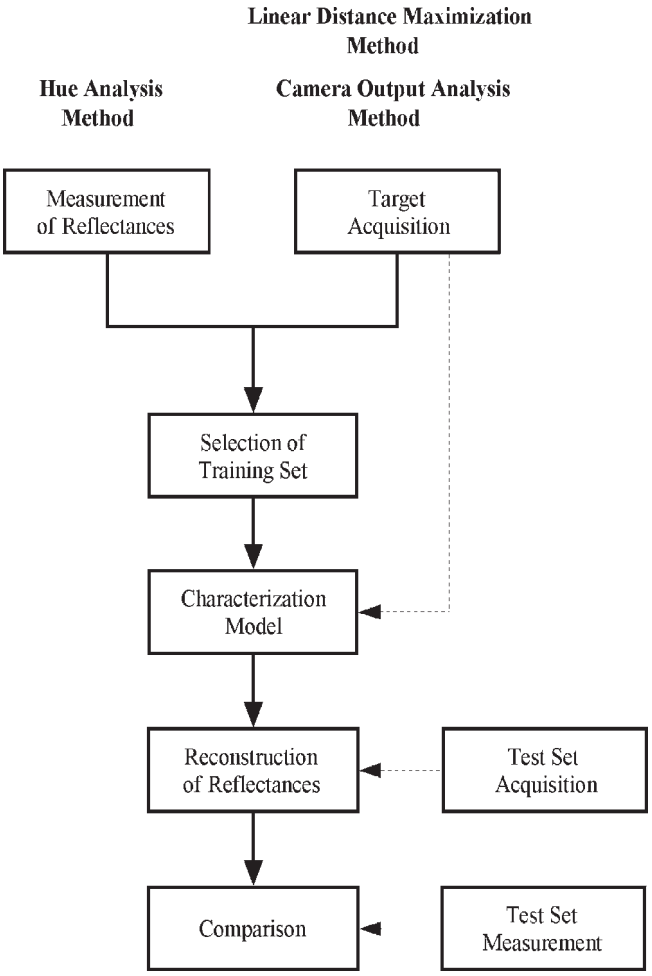

Figure 7. The complete procedure for selecting and evaluating training sets.

they yielded the best results. Anyway, the three methods we proposed can be used with different acquisition systems as well; in particular, their applicability is independent of the number of filters employed.

To evaluate the quality of the approximations, the reconstructed reflectances were then compared with measurements of the same reflectances obtained using a Minolta CM-2002 spectrophotometer; these measurements were taken between 400 nm and 700 nm, at intervals of 10 nm. As a measure of the precision of the reconstruction, we considered the maximum absolute difference computed on all the components of the reflectance vector: formally, if $\mathbf{r} = [r_j]_{j = 1, \ldots, M}$ is the measured reflectance and $\mathbf{r}' = [r'_j]_{j = 1, \ldots, M}$ is the reconstructed reflectance, then our distance d is defined as

$$d(\mathbf{r}, \mathbf{r}') = \max_{j=1,\ldots,M} |r_j - r'_j| \qquad (12)$$

We chose this distance instead of the customarily employed RMS for two reasons: we were more interested in the maximum error than in the mean error; and if our distance were small then the corresponding RMS was also small, while the converse is not necessarily true.

In general, we selected training sets with 16, 31, or 62 colors included. The value 31 was chosen because it was the dimension of both the system output vector space and the (measured) reflectances vector space; the values 62 and 16 (respectively, the double and the rounded half of 31) were added to investigate if any stable and consistent dependency of the quality of the results from the numerosity of the training set could be guessed or established, as well as to allow a comparison with the third version of the COAM method. For every training set, the errors obtained comparing the reconstructed reflectances to the corresponding measures were computed according to Equation 12. In each case, the mean and maximum errors observed across all the colors in the test set are reported.

TABLE I. Mean and Maximum Errors for the Hue Analysis Method on the MDC Target

Hue Analysis Method	Mean Error	Maximum Error
16 colors	0.0103	0.1101
31 colors	0.0074	0.0688
62 colors	0.0056	0.0668

TABLE II. Mean and Maximum Errors for the Three Versions of the Camera Output Analysis Method on the MDC Target

Camera Output Analysis Method	Mean Error	Maximum Error
31 colors – nearest	0.0099	0.1077
31 colors – absolute largest coordinates	0.0062	0.0550
62 colors – largest / smallest coordinates	0.0052	0.0395

TABLE III. Mean and Maximum Errors for the Two Versions of the Linear Distance Maximization Method on the MDC Target

Linear Distance Maximization Method	Mean Error	Maximum Error
16 colors – Euclidean distance	0.0088	0.1000
16 colors – maximum absolute difference of coordinates	0.0087	0.1199
31 colors – Euclidean distance	0.0061	0.0705
31 colors – maximum absolute difference of coordinates	0.0057	0.0687
62 colors – Euclidean distance	0.0049	0.0457
62 colors – maximum absolute difference of coordinates	0.0046	0.0378

In the case of the Hue Analysis Method (HAM), the XYZ coordinates of the target colors were computed, using the D65 illuminant and the standard 2° CIE observer, from the corresponding measured reflectances, and the $L^*C_{ab}^*h_{ab}$ coordinates were then obtained by means of standard colorimetric formulae.[8] In all cases, the center of the first sector in the $C_{ab}^*h_{ab}$ plane was set at $h_{ab} = 0$. The results for the HAM method are shown in Table I.

In the case of the Camera Output Analysis Method (COAM), we limited the numerosity of the chosen training sets to the values of 31 for the first and second versions and 62 for the third version; this is consistent with the nature of this method, as it chooses one sample color (two in the case of the third version) for each eigenvector in the space of system output vectors, and, in our experiments, the dimension of this space is 31. On the other hand, since the Linear Distance Maximization Method (LDMM) is not bound by the same constraints, we employed both its versions to select training sets of 16, 31, and 62 colors. Table II shows the results for all three versions of the COAM, while Table III reports the results obtained with the two versions of the LDMM.

TABLE IV. Mean and Maximum Errors for the Method Proposed by Hardeberg[6] on the MDC Target

Hardeberg	Mean Error	Maximum Error
16 colors	0.0106	0.0684
31 colors	0.0071	0.0624
62 colors	0.0055	0.0599

TABLE V. Mean and Maximum Errors for 10 Randomly Chosen Training Sets on the MDC Target

Randomly chosen training sets	Mean Error	Maximum Error
31 colors – average values	0.0092	0.1373
31 colors – best values	0.0075	0.0781

TABLE VI. Mean and Maximum Errors Obtained Using the Entire MDC Target as Both Test Set and Training Set

Whole MDC target as training set	Mean Error	Maximum Error
177 colors	0.0047	0.0356

As stated above, to allow a comparison with our methods, we report here the results obtained with training sets selected using the method of Hardeberg and using a biased random criterion. The results from these trials are shown in Table IV and Table V respectively. For the random selection method, we report the average and best results obtained over ten different training sets of numerosity 31. We also report in Table VI the results obtained using the whole test set as training set.

Last, Table VII summarizes the best results we achieved across all our trials: performances were compared with one another according to whether a method works on measured reflectances or on system output vectors, and distinguishing between mean and maximum errors for each of the three sizes of the training sets. The results obtained using the whole MDC target as training set were not considered here: in fact, since the whole target contains much more colors than the other chosen training sets, a direct comparison is not feasible. However, it should be noted that the LDMM method achieves almost the same results when using the maximum absolute difference of coordinates with only 62 colors in the training set.

As a further reference, the composition of all the training sets that were selected throughout our trials is included in an Apprendix available as Supplemental Material on the IS&T website (http://www.imaging.org/pubs/jist/index.cfm) and with the on-line edition of this Journal.

Discussion

Several remarks can be made based on our results. First, when training set selection was based on measured reflectances, Hardeberg's method generally performed better than our Hue Analysis Method. Of all the methods tried, it is also the one that achieved more consistent results as to the maximum error, which is not very high with 16 training set colors but neither decreases greatly as the numerosity grows larger; this may indicate that the method achieves a 'good' variety in the first selected colors, but adding more does not improve the representativeness of the training set so much. On the other hand, our HAM method showed a marked decrease in the mean error, while the maximum error did not improve much in passing from 31 to 62 colors.

The methods that base their selections on the system output data generally performed better than both Hardeberg's and the HAM method. In particular, the second version of the LDMM achieved the best results in all cases with respect to the mean error, and also showed the best maximum error when using 62 colors; also, the first version of the method performed better than the COAM with respect to the maximum error using 16 colors (though Hardeberg's method did much better in this case). While the second version of the COAM generally achieved good results, the performance of the first version did not entirely meet our expectations. However, an analysis of the spatial disposition of the target colors in system output space revealed that the colors were rather tightly clustered, as we noted above, in such a situation the first version of the COAM is expected to perform more poorly compared to the second version, and the experimental results are consistent with this expectation. Also, this may explain the good performances of the LDMM method.

A few considerations concerning the numerosity of the selected training sets can also be made. First, as the numerosity grew, all methods showed a marked decrease in the mean error; however, this was expected, since the training set was in all cases part of the test set. In general, no method showed any degradation in the maximum error when using 62 colors, which is double the dimension of the characterization function domain; this suggests that in no case did a large training set lead to local overfitting. Only the LDMM and the COAM, though, achieved a marked decrease in the maximum error when passing from 31 to 62 colors; this probably happened because they both work on system output data, and are therefore able to use the 'excess' training set colors to better correct the unavoidable measurement error in the data. Assessing whether all these excess colors are equally useful or only some of them would be sufficient is still an open research issue.

Based on our trials, it could be stated that the methods working on system output data should generally be preferred. The characterization models obtained with such methods depend on the acquisition conditions, but may be used to reconstruct the reflectances of any surface acquired under the same conditions, making them reusable. However, methods working on reflectances are independent from the operating conditions of real acquisitions, and can therefore be used to build training

TABLE VII. Best Performances

	Methods working on measured reflectances		Methods working on system output data	
Numerosity	Mean Error	Maximum Error	Mean Error	Maximum Error
16	HAM	Hardeberg	LDMM max. abs. diff.	LDMM Eucl. Dist.
31	Hardeberg	Hardeberg	LDMM max. abs. diff.	COAM 2nd version
62	Hardeberg	Hardeberg	LDMM max. abs. diff.	LDMM max. abs. diff.

sets with a broader applicability. The only caution, which applies to the HAM method only, is that since the spatial disposition of the colors in the $L^*C_{ab}^*h_{ab}$ space depends on the illuminant chosen for the computation of the $L^*C_{ab}^*h_{ab}$ coordinates, this illuminant should be consistent with the illuminant used in the acquisitions to which the resulting characterization model will be applied. Anyway, as long as cross-illuminant applicability is not a concern, this should reasonably lead to training sets that are better suited to the specific situations considered compared to the more 'generic' training sets selected with Hardeberg's method, although our experimental results do not support this expectation.

On the other hand, methods working on reflectances obviously require that reflectances be known or be estimated in some way; this means that either a good measurement instrument (like a spectrophotometer) must be available, or reflectances must be reconstructed from the system output data. The second scenario is not feasible, since it requires that a characterization model has already been established; this means either that a training set has already been selected, or that none is required, and in both cases there is no need to select a new training set. One could also reconstruct reflectances using a 'temporary' characterization model, but this would likely lead to noisy results. The need for a measurement instrument can instead become a limitation, especially in an industrial environment, where the associated investment (in terms of money, training, and time) could be seen as unnecessary. This could either force the manufacturer to sell pre-characterized acquisition systems, potentially limiting their flexibility, or force the user to have the system characterized repeatedly for different situations or configurations, increasing the operating cost. Also, in both cases, the performance advantages of the methods working on the system output data would not be exploited.

A last note concerns the system characterization. Since we used a linear characterization model and employed standard numerical analysis techniques, it could be argued that the applicability of the proposed methods is restricted to this scenario. However, only our COAM method makes an assumption about the linearity of the system characterization function, while the other methods are based on different premises. This is true for the LDMM method too, since even if it was developed as an 'evolution' of the COAM method, it is not tied to any specific geometric feature of the characterization function. Also, any assumptions about the characterization model do not necessarily imply assumptions about the modeling technique; if different modeling techniques, such as neural networks, were employed, the theoretical validity of the proposed methods would not be affected. However, as other specific issues may arise when a different modeling technique is adopted, further tests with some of these techniques would help assess the general quality of the proposed methods; this will be one of our research priorities in the near future. It would also be interesting to investigate whether these methods would perform well with non-linear systems.

In the end, if measuring reflectances is not an issue, and a training set with very broad applicability (potentially, even a 'universal' training set) is wanted, we would suggest trying Hardeberg's method, especially if the numerosity of the training set is to be very small. However, we generally prefer the second version of our Linear Distance Maximization Method because of its superior performances with mid-sized training sets.

Conclusions

We have proposed three methods, which we have called respectively Hue Analysis Method, Camera Output Analysis Method, and Linear Distance Maximization Method, for selecting a good training set that can be employed in the characterization of a multispectral acquisition system. The selected training sets have relatively low numerosity and broad applicability, and are particularly suitable to be employed in real acquisitions. All three methods have been tested on data obtained from a real acquisition, and the results of the corresponding approximations show a substantial improvement over those obtained using randomly chosen training sets. Our methods have also performed comparably to the method of Hardeberg,[6] with the Linear Distance Maximization Method emerging as the best choice among all the methods tried for application in a general context. People who are interested in using the source code that implements the methods we have introduced are encouraged to contact us.[14]

References

1. F. H. Imai, M. Rosen and R. S. Berns, Comparison of spectrally narrow-band capture versus wide-band with a priori sample analysis for spectral reflectance estimation, *Proc. IS&T/SID 8th Color Imaging Conference*, IS&T, Springfield, VA, 2000, p. 234.
2. F. Schmitt, H. Brettel and J. Hardeberg, Multispectral Imaging Development at ENST, *International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives*, Chiba University, Chiba, Japan, 1999, p. 50.
3. S. Tominaga, Spectral imaging by a multichannel camera, *J. Electronic Imaging*, **8,** 332 (1999).
4. A. Abrardo, L. Alparone, V. Cappellini, and A. Prosperi, Color Constancy from Multispectral Images, *Proc. of 6th IEEE International Conference on Image Processing ICIP '99*, IEEE Computer Society, Kobe, Japan, 1999, p. 570.
5. P. G. Herzog, D. Knipp, H. Stiebig, and F. König, Characterization of novel three and six channel color moiré free sensors, *Proc. SPIE* **3648,** 48 (1999).
6. J. Y. Hardeberg, H. Brettel and F. Schmitt, Spectral characterization of electronic cameras, *Proc. SPIE* **3409,** 100 (1998).
7. P. L. Vora, J. E. Farrell, J. D. Tietz, and D. H. Brainard, Image Capture: Simulation of Sensor Responses from Hyperspectral Images, *IEEE Transactions on Image Processing*, **10(2)** p. 307 (2001).
8. R. W. G. Hunt, *Measuring Colour,* 3rd ed., Fountain Press, London, UK, 1998.
9. I. T. Jolliffe, *Principal Component Analysis*, Springer-Verlag, New York, 1986.
10. http://www.gretagmacbeth.com/files/products/ColorCheckerDC_new_IT.pdf.
11. S. Tominaga and R. Okajima, A Spectral Imaging System and Algorithms for Recovering Spectral Functions, *Proc. 4th IEEE Southwest Symp. on Image Analysis and Interpretation*, IEEE Computer Society, Austin, Texas, 2000, p. 278.
12. S. Poger and E. Angelopoulou, Selecting Components for Building Multispectral Sensors, *IEEE CVPR Technical Sketches (CVPR Tech Sketches) 2001*, IEEE Computer Society Press, Los Alamitos, CA, 2001.
13. J. H. Wilkinson and C. Reinsch, Eds., *Linear Algebra Handbook for Automatic Computation*, Springer-Verlag, New York, 1971, p. 134.
14. http://www.ivl.disco.unimib.it.