Selection of Training Sets for the Characterisation of Multispectral Imaging Systems

Paolo Pellegri [ab]***, Gianluca Novati*** [ab]***, and Raimondo Schettini*** [a]
[a] ***DISCo, University of Milano – Bicocca***
[b] ***ITC, National Research Council***
Milano, Italy

Abstract

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 correlation method and the training set (if any) on which the method relies. In this paper we propose two novel approaches in selecting a training set to be used for the characterisation of a multispectral acquisition system. In both cases the selected training sets will have low numerosity and broad applicability. We also test both approaches on the data obtained from a real acquisition, comparing the reconstructed reflectances with the measurements obtained using a spectrophotometre.

Introduction

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 correlation method and the training set (if any) on which the method relies. The output of a generic multispectral acquisition system may be denoted as

$$\mathbf{a}(\mathbf{x}) = [\ a_i(\mathbf{x})\ ]_i\,, \tag{1}$$

where i is an index that varies with the filter used (or the spectral band examined), and $\mathbf{x}$ is a two-dimensional vector identifying the point considered within the acquired scene. If M filters are used, then $\mathbf{a}(\mathbf{x})$ is an M-dimensional vector. The reflectance of the object at point $\mathbf{x}$ is a function of the wavelength λ, and can be denoted as $R(\mathbf{x}, \lambda)$; however, 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 then sampled at a discrete number of values of λ, and the reflectance is expressed as

$$\mathbf{r}(\mathbf{x}) = [\ R(\mathbf{x}, \lambda_j)\ ]_j\,, \tag{2}$$

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 correlation between the system output and the corresponding reflectance, the system characterisation function

$$\mathbf{a}(\mathbf{x}) \mapsto \mathbf{r}(\mathbf{x}) \tag{3}$$

must be described or estimated in some way; this is usually done by means of an empirical model based on a chosen training set and correlation method (such as, for instance, linear models[1] or polynomial regression[2]). The quality of the estimation depends on the correlation 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 and a certain correlation 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 correlation 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[3] employs an algebraic method to select a training set of low numerosity; most other authors (such as Herzog et al.[2]) either employ as their training set the set of reflectances to be reconstructed, or simply avoid the issue. One likely cause for this approach is that many authors work on simulations 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 two different approaches in selecting a ‘good’ training set from an initial array of available colours (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 approach, which we call the Camera Output Analysis Method, is 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 characterisation models on the data obtained in a real acquisition.

The Hue Analysis Method

Ideally, we may expect the 'representativeness' of a training set to improve as the number of 'different' colours included in it increases. Since in this context colours 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 colour, 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 colours in the target are known, the corresponding LHC coordinates are computed, and the L coordinate is normalised (or simply ignored) so that the colours are projected onto an HC plane. This plane is then divided into *n* sectors of equal angular width, *n* being the number of colours to be selected. For each sector, the colour 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* colours (see Figure 1). In this way the chosen colours 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 colours chosen with no regard to their spatial disposition.

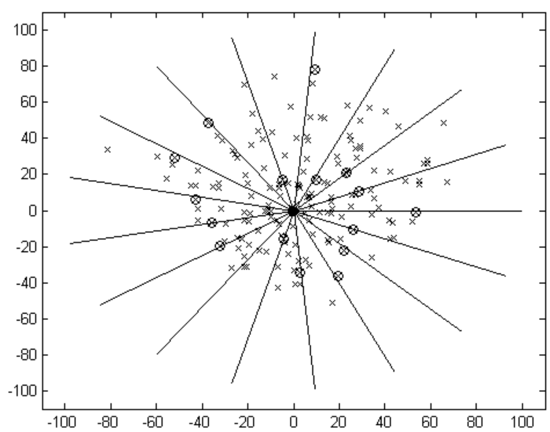

Figure 1. An example of how training set colours are chosen in the HC plane. The central half-lines of all the sectors are drawn, and the chosen colours are circled.

The Hue Analysis Method may pose two geometrical issues, which arise, respectively, when no target colours lie within a sector (see Figure 2), or when a colour is the nearest to both of two neighbouring sectors (in which case the colour lies on the boundary between the two sectors). The Hue Analysis Method cannot be applied in either of these cases, and while the second instance is rare, the first may present more frequently, especially as the number of elements desired in the training set grows larger. However, 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 HC plane; then, if either of the two situations described arises, the initial target must be replaced.

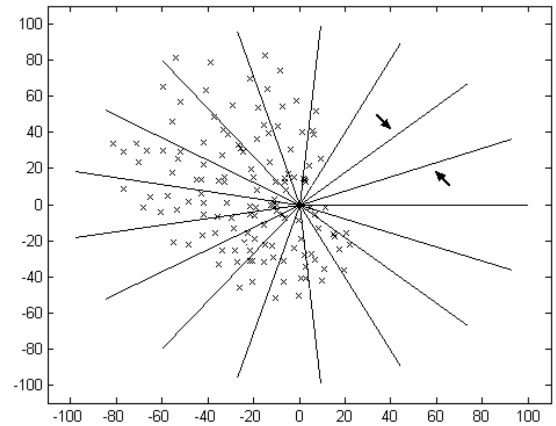

Figure 2. No colours lie within the sectors corresponding to the indicated half-lines: the Hue Analysis Method cannot be applied.

The Camera Output Analysis Method

If the characterisation 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 camera output vectors to the *N*-dimensional vector space of (sampled) reflectances. In this context, for a training set to be representative enough to use to reconstruct the reflectance of any colour, then it must span the whole characterisation function domain. This means that it must include a subset which is a basis for the vector space of the camera output vectors. This requirement can be met by simply including in the training set *M* colours the corresponding camera output vectors of which are linearly independent. However, although theoretically sufficient, this approach may not give satisfactory results. As our experimental results demonstrate, the characterisation models built on 'randomly chosen' training sets of linearly independent colours may be very imprecise. In our study, an analysis of the spatial distribution of the camera output vectors corresponding to the colours 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 camera output vectors.

Although the magnitude of this error may be small, if the colours 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 camera output space, so that the decomposition of an arbitrary colour on the basis of the training set colours 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 colours chosen for the training set well apart from one another. Principal component analysis is applied to all available colours, so that the principal eigenvectors are identified. Then, for each eigenvector in order of relevance, the colour not already chosen 'nearest' to that eigenvector is included in the training set, for a total of *M* vectors. The distance used to measure 'nearness' is the cosine of the angle between the camera output vector of a colour and the eigenvector concerned (see Figure 3).

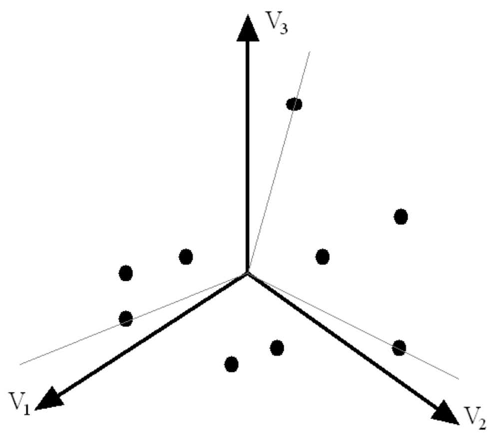

Figure 3. A fictional example of three eigenvectors and nine target colours. The colours nearest to each eigenvector in the sense of angular distance are marked.

This strategy tries to maximise the orthogonality of the colours 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 colours is observed in the camera output space, the relative linear distance of the colours chosen for the training set may prove to be more important than their orthogonality to this end. To provide for this case, we designed a second version of the Camera Output Analysis Method to maximise the relative linear distance of training set colours while still taking into account their anisotropic spatial disposition (as described by the eigenvectors). In this variation all available colours are decomposed on the basis of the eigenvectors, and, for each eigenvector in order of relevance, the colour not already chosen with the greatest absolute coordinate on that eigenvector is included in the training set, for a total of *M* vectors (see Figure 4).

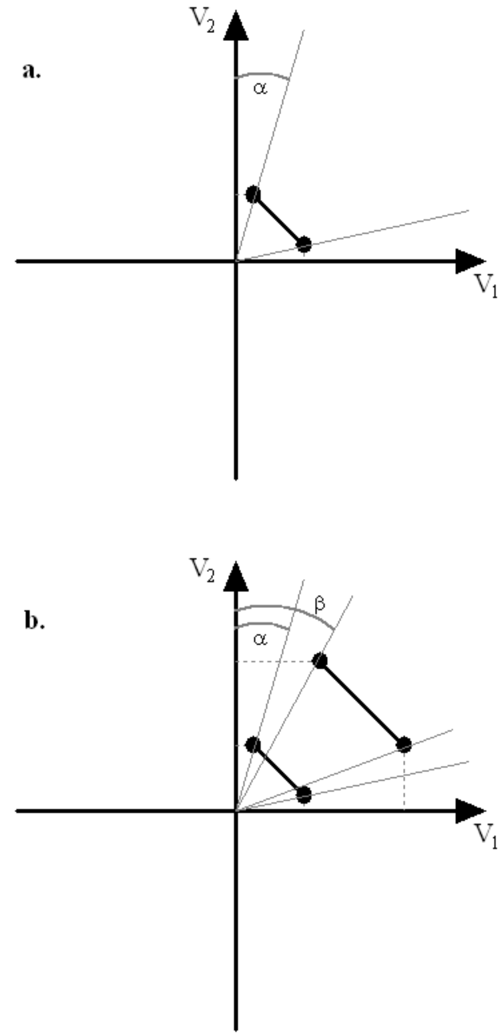

Figure 4. (a) Two colours which are nearest to different eigenvectors in the sense of angular distance may still be close to each other. (b) Colours which have maximum absolute coordinate on those eigenvectors may be more spaced.

There may be cases, however, in which even this approach cannot guarantee that the selected colours will be sufficiently spaced. Particularly, if one considers two eigenvectors the associated singular values of which are small, the colours selected in correspondence to those eigenvectors may still be close to each other. This happens because the variability over the set of the target colours of the coordinate associated with an eigenvector decreases as the associated singular value decreases: if two colours 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 to this problem: for each eigenvector, both the colour with the greatest coordinate and the colour with the smallest coordinate are selected and included in the training set, for a total of 2*M* colours. As figure 5 shows, this approach usually guarantees that for each pair of eigenvectors there will be a corresponding pair of training set colours that are more distant from each other than those colours that would be selected using the second version of the method.

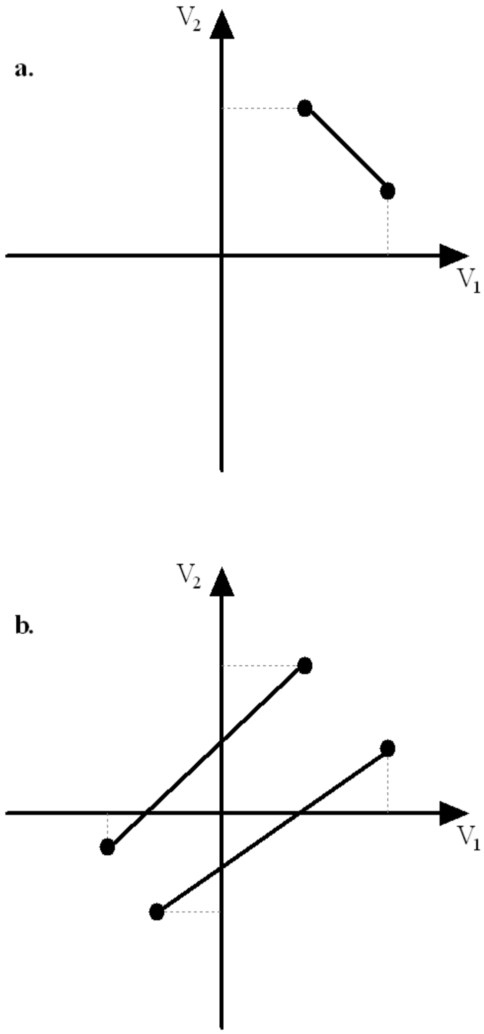

Figure 5. (a) Two chosen colours corresponding to different eigenvectors may still be close using the second version of Camera Output Analysis Method. (b) Pairs of colours chosen with the third version will usually be more spaced.

Comparison

The two methods we proposed operate on different premises and follow different approaches, but both allow the selection of training sets that are widely applicable. The Hue Analysis Method is independent from the operating conditions of real acquisitions, and makes it possible to build a training set of any numerosity: it could potentially be used to select a truly 'universal' training set. However, since the spatial disposition of the colours in the LHC space depends on the illuminant chosen for the computation of the LHC coordinates, this illuminant must be the same as or very similar to the illuminant used in the acquisitions to which the resulting characterisation model will be applied. The selected training set will be adequate as long as the environmental illuminant does not change. The same holds true for the Camera Output Analysis Method, which operates on the camera output data obtained from a real acquisition of the target: in this case, the characterisation model obtained depends on the acquisition conditions, but may be used to reconstruct the reflectances of any surface acquired under the same conditions. The numerosity of the selected training set will also be low, since it depends on the number of different filters used / bands considered in the acquisitions.

Experiments

We have used the Hue Analysis Method and the three versions of the Camera Output Analysis Method to select different training sets from the colours included in the Macbeth Digital Camera target (MDC). Each of the selected training sets was employed to approximate the characterisation function of our multispectral acquisition system, and the resulting characterisation models were then taken to reconstruct the reflectances of all the colours in the MDC target from the output data obtained by a real acquisition (see Figure 6).

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, a high-quality Rodagon lens, a VariSpec Tunable Filter, and a cut-off optical filter for infrared and ultraviolet radiations. We used 33 different configurations of the Tunable Filter for band selection. Since previous experiments had shown that our system response function was linear, we adopted a linear model to approximate the system characterisation function; consequently, each selected training set was used to compute a corresponding linear model, employing singular value decomposition for system inversion.

To evaluate the quality of the approximations, the reconstructed reflectances were compared with measurements of the same reflectances obtained using a Spectrolino spectrophotometre. 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 $\mathbf{r'} = [\, \mathbf{r'}_j \,]_{j = 1, \ldots, M}$ is the reconstructed reflectance, then our distance d is defined as

$$d(\mathbf{r}, \mathbf{r'}) = \max_{j=1,\ldots,M} \left| r_j - r'_j \right| . \tag{4}$$

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 was small then the corresponding RMS was also small (while the converse was not necessarily true).

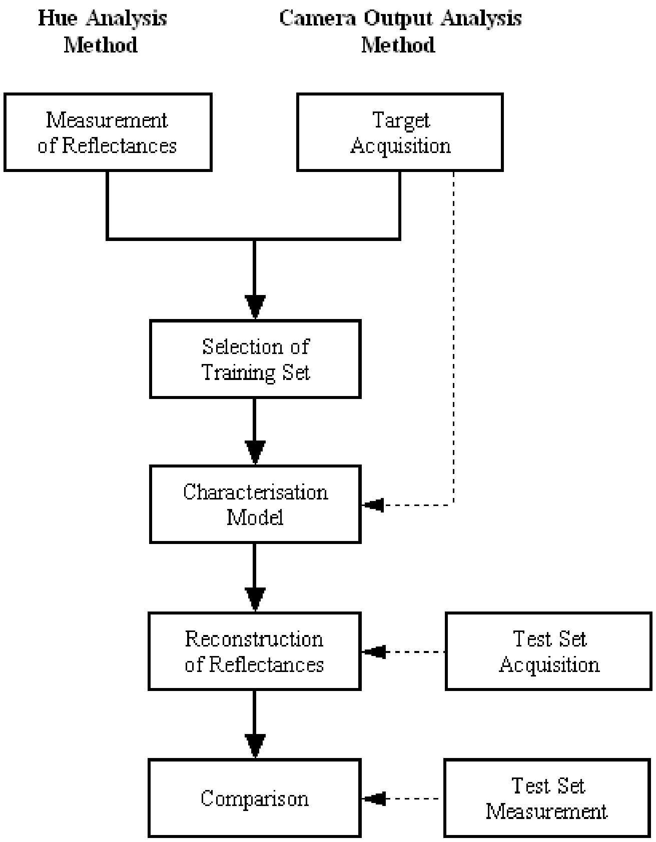

Figure 6. The whole procedure for selecting and evaluating training sets.

In the case of the Hue Analysis Method, the XYZ coordinates of the target colours were computed, using the D65 illuminant and the standard 2° CIE observer, from the corresponding measured reflectances, and the LHC coordinates were then obtained by means of standard colorimetric formulas. The HC plane was divided into 33 and 66 sectors, with the centre of the first sector at H = 0, and two training sets of corresponding numerosity were selected. These values were chosen to allow comparison with the results obtained using the Camera Output Analysis Method. Finally, for each model/training set we reconstructed the reflectances of the colours in the test set (the MDC target), and compared them with the measured reflec-tances we had obtained: the results are shown in Table 1.

Table 1. Maximum and mean errors for the Hue Analysis Method on the MDC target

Hue Analysis Method	Max Error	Mean Error
33 colours	0.0856	0.0076
66 colours	0.0787	0.0061

Each version of the Camera Output Analysis Method was used to select a training set of the proper numerosity. We then reconstructed the reflectances and compared them to the measured reflectances. Table 2 shows the results for all three versions. The results obtained using three randomly chosen training sets of equal numerosity are also reported to allow a comparison.

Table 2. Maximum and mean errors for the three versions of the Camera Output Analysis Method and for randomly chosen training sets on the MDC target

Camera Output Analysis Method	Max Error	Mean Error
33 colours – nearest	0.1382	0.0112
33 colours – largest absolute coordinates	0.0701	0.0077
66 colours – largest / smallest coordinates	0.0405	0.0053
Randomly chosen training sets	**Max Error**	**Mean Error**
33 colours – choice 1	0.1894	0.0184
33 colours – choice 2	0.1623	0.0127
33 colours – choice 3	0.1845	0.0132

Since both the data and the test set used were the same for all the trials, a direct comparison is possible. As the tables show, both of our proposed methods produced results significantly better than those obtained with randomly chosen training sets. Both of the models built using training sets with 66 colours also yielded good approximations of the reflectance vectors, showing no degradation due to local overfitting. On the other hand, the results obtained using the first version of the Camera Output Analysis Method did not entirely meet our expectations. However, an analysis of the spatial disposition of the target colours in camera output space revealed that in our case the colours were rather tightly clustered: as we noticed above, in such a situation the first version of the Camera Output Analysis Method is expected to perform more poorly compared to the second version, and the experimental results are consistent with this expectation.

Conclusions

We have proposed two methods, which we have called the Hue Analysis Method and the Camera Output Analysis Method, for selecting a good training set that can employed in the characterisation of a multispectral acquisition system. The selected training sets have low numerosity and broad applicability, and are particularly suitable to be employed in real acquisitions. Both methods have been tested on data obtained from a real acquisition, and the results of the corresponding approximations are significantly better that those obtained using randomly chosen training sets of comparable numerosity.

References

1. A. Abrardo, L. Alparone, V. Cappellini and Andrea Prosperi, Color Constancy from Multispectral Images, Proc. of 6th IEEE International Conference on Image Processing ICIP'99 pp 570-574 (1999).
2. P. G. Herzog, D. Knipp, H. Stiebig and F. König, Characterization of novel three and six channel color moiré free sensors, Electronic Imaging Conference, San Jose´ (1999) and Proc. SPIE, 3648, pp. 48-59 (1999).
3. J. Y. Hardeberg, H. Brettel and F. Schmitt, Spectral characterisation of electronic cameras, Electronic Imaging: Processing, Printing and Publishing in Color, Zurich, Switzerland, SPIE Proceedings 3409, pp 100-109, (1998).

Biography

Paolo Pellegri received his degree in Mathematics from the University of Milano (Italy) in 2001, and is currently a Ph.D. student at the University of Milano - Bicocca. Since his degree thesis, he has been associated with Italian National Research Council (CNR), where he has been working on multimedia technologies. His current research activities focuses on the developement of efficient algorithms an techniques for colour device characterisation, multispectral imaging and 3D imaging.

Gianluca Novati received his degree in Computer Science in 2001 from the University of Milano (Italy), and is currently a Ph.D. student at the University of Milano - Bicocca. Since 2002 he has been working in the field of multimedia technologies at the Italian National Research Council (CNR). His current research focuses on the acquisition, processing, storage, and rendering of multispectral images.

Raimondo Schettini is an associate professor at DISCo, University of Milano - Bicocca. Since 1987, he has been associated with the Italian National Research Council (CNR), and in 1994 he moved to the Institute of Multimedia Information Technologies, where he is currently in charge of the Image and Color Analysis Lab. He is a member of the CIE TC on multispectral imaging, and he has been General Co-Chairman of the First European Conference on Color in Graphics, Imaging and Vision (2002).