Effect of spatial structure on visual tolerance

Q Chen, TLV Cheung and S Westland, Centre for Colour Design Technology, University of Leeds (UK)

Abstract

It is known that the reflectance spectra of both natural and man-made surfaces may be represented efficiently using linear models. A key question, however, is how many basis functions of a linear model are necessary for a given accuracy of representation. The question is ill-posed, however, since it is understood that the number of basis functions required depends to a great extent on the intended application of the linear model. However, in one study it was shown that more than six basis functions were required to ensure that the largest colour difference in the set of spectra was less than 1.0 CIELAB unit and therefore it is reasonable to assume that, for many applications where relatively large patches of spatially uniform colour are present, six of basis functions will be required since CIELAB colour differences of unity or more in such circumstances are known to be noticeable. However, the magnitude of colour difference that would be visible in a complex or natural image is not so well established. A recent psychophysical study demonstrated that although five basis functions produced on average unit error in CIELAB space, original natural images were psychophysically indistinguishable from their linear-model approximations only if there were at least 8 basis functions. The aim of this study is to psychophysically investigate the effect of spatial structure on the number of basis functions required to colorimetrically reproduce spectral images.

Introduction

It is known that the reflectance spectra of both natural and man-made surfaces may be represented efficiently using linear models. A key question, however, is how many basis functions of a linear model are necessary for a given accuracy of representation. Many studies[1-3] have been carried out to estimate the minimum number of basis functions for spectral reproduction. The question is ill-posed, however, since it is understood that the number of basis functions required depends to a great extent on the intended application of the linear model[4]. However, in one study it was shown that more than six basis functions were required to ensure that the largest colour difference in the set of spectra was less than 1.0 CIELAB unit[5] and therefore it is reasonable to assume that, for many applications where relatively large patches of spatially uniform colour are present, six of basis functions will be required since CIELAB colour differences of unity or more in such circumstances are known to be noticeable. However, the magnitude of colour difference that would be visible in a complex or natural image is not so well established[6]. A recent psychophysical study[7] demonstrated that although five basis functions produced on average unit error in CIELAB space, original natural images were psychophysically indistinguishable from their linear-model approximations only if there were at least 8 basis functions. The aim of this study is to psychophysically investigate the effect of spatial structure on the number of basis functions required to colorimetrically reproduce spectral images.

Experimental

A set of 24 reflectance spectra were chosen from the set of 1269 Munsell reflectance spectra[8] using a colour-selection technique[9] that aimed to ensure that the selected samples were evenly distributed in the (approximately) visually uniform CIELAB colour space. Figure 1 shows the chromaticities of the 24 reflectance spectra when viewed by the CIE 1964 standard observer under D65 illumination in CIE *xy* coordinates. A linear model consisting of basis functions was derived from the full 1269 Munsell data set and subsequently used to represent the 24 reflectance spectra using n basis functions where $n \in \{1,2,...,9\}$. Mondrian-like images were created using n basis functions to represent the spectra of the patches. CIE tristimulus values were computed (illuminant D65 and 1964 standard observer data) for each spectral representation in the linear model and a monitor characterization model was used to transform the *XYZ* values to monitor *RGB* values[9] for display purposes.

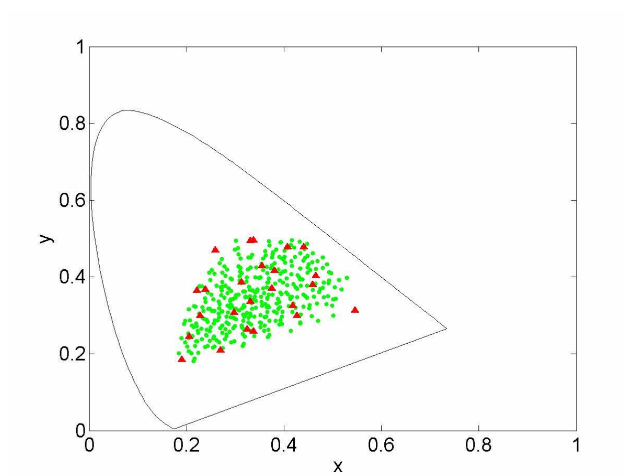
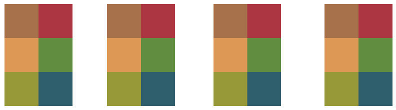
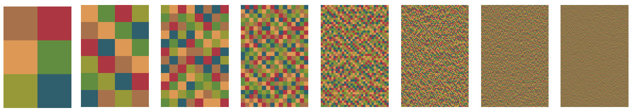


***Figure 1**: Chromaticity in CIE xy coordinate for 24 reflectance spectra (red triangles) used to create the first set of Mondrians and 384 reflectance spectra (green dots) used to create the second Mondrians.*

The 24 reflectance spectra were randomly divided into 4 groups of 6. Each group was then used to define the reflectance characteristics of patches in Mondrian images (384 × 256 pixels). In the simplest case the Mondrian images were composed of 6 patches each of size 128 × 128 pixels; in the most complex case there were 95744 patches each of size 1 × 1 pixels (see Figure 2). Irrespective of the number of patches (6, 24, 96, 384, 1536, 6144, 24576 or 95744) only 6 reflectance spectra were used in the definition of the patches of any Mondrian (see Figure 2b). This created a series of Mondrian images where the spatial complexity of the image varied but the spectral properties were kept constant.

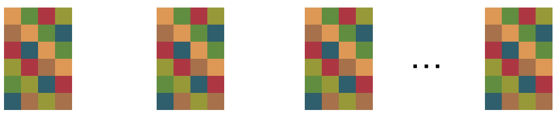


(a)



128×128 *64×64* *32×32* *16×16* *8×8* *4×4* *2×2* *1×1*

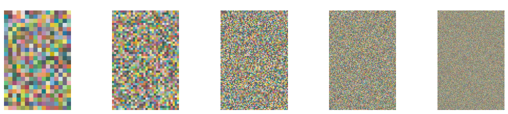
(b)



1 basis function *2 basis functions* *3 basis functions* … *9 basis functions*

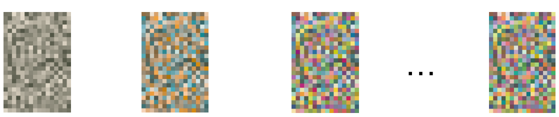
(c)

***Figure 2**: Illustration of the first set of Mondrian images (a) constituted from 24 reflectance spectra (b) with varying patch size in units of pixel, and (c) the produced images by linear modeling of 1 to 9 basis functions for each class of spatial complexity Mondrian image.*



16×16 *8×8* *4×4* *2×2* *1×1*

(a)



1 basis function *2 basis functions* *3 basis functions* … *9 basis functions*

(b)

***Figure 3**: Illustration of the second set of Mondrian images constituted from 384 reflectance spectra with (a) varying patch size in units of pixel, and (b) the produced images by linear modeling of 1 to 9 basis functions for each class of spatial complexity Mondrian image.*

A second set of Mondrian images was created based on a subset of 384 reflectance spectra. The chromaticities of the 384 reflectance spectra are illustrated in Figure 1. Rather than using only 6 reflectance spectra as in the first set of Mondrian images, each image in the second set of Mondrians contains 384 spectrally unique patches. The series of Mondrian images was generated in a similar way as the first set of Mondrian images; however, in order to include all 384 colours in all images of the second series, the largest patch size was 16 × 16 pixels (resulting in 384, 1536, 6144, 24576 or 95744 patches). A typical sequence is illustrated in Figure 3. As in the first series this created a set of Mondrian images where the spatial complexity of the image varied but the spectral properties were kept constant. However, in the second series the variety of colours (both spectrally and colorimetrically) is much larger than in the first set.

The configuration of the psychophysical experiment was such that in each trial three images were displayed in a horizontal row on a computer monitor. The centre image always consisted of original reflectance spectra; the left-hand and right-hand images were randomly selected so that one was identical to the centre image and the other consisted of linear-model reconstructions of the spectra in the centre image. For the reconstructed spectra the number of basis functions n was varied but remained the same in any one image or trial. Observers were informed that the central image was the original and were forced to choose whether the left-hand or right-hand image was an identical match to the original. The experiment was therefore a two-alternate forced-choice (2AFC) paradigm. It was assumed that when the reconstructed image was very different to the original then the observer would correctly select the original image with high probability. Conversely, when the quality of the reconstruction was high it would be difficult for the observer to identify the original and discrimination performance would tend towards 50% correct responses. Note, therefore, that poor visual performance in the visual discrimination task will correspond to an effective linear model.

Five observers (QC, VC, WL, CF, SC), with normal colour vision and normal or corrected-to-normal visual acuity, were recruited to take part in the psychophysical experiment. All five observers participated in the experiment using the second set of Mondrian images. The viewing distance of this experiment was fixed at 160cm so that the series of images generated patch sizes in the range 2.4 to 38.4 cyc/deg. Each observer undertook at least 24 repeats for the 5 (spatial complexity) × 9 (linear model n) conditions leading to 5400 trials.

Three of the observers (QC, VC, WL) participated in the experiment using the first set of Mondrian images. Two of the observers (QC and VC) viewed the stimuli at 160cm so that the series of images generated patch sizes in the range 0.30 to 38.4 cyc/deg. Each observer undertook 16 repeats for the 8 (spatial complexity) × 9 (linear model n) × 4 (selections of 6 spectra) conditions leading to a total of 4608 trials. All three observers also repeated the experiment at half the viewing distance (80 cm). Each observer undertook 15 repeats for the 8 (spatial complexity) × 9 (linear model n) × 4 (selections of 6 spectra) conditions leading to a total of 12960 trials.

Results and discussion

In order to quantify the reproduced reflectance colorimetrically, CIELAB colour differences were calculated between each original reflectance spectra and its representation in the linear model with $n \in \{1,2,...,9\}$. Table 1 shows the average colour difference between the original images and their linear-model representations for different values of n. The data in Table 1 are consistent with data reported in similar studies[1,2,5].

Psychophysical data were fitted by psychometric functions fitted using *psignifit* (vers. 2.5.6), a software package which implements the maximum-likelihood method[11]. Figure 4 shows an example of one observer's results for a set of stimuli of a particular spatial complexity. For each psychometric function the threshold of discrimination performance was set to 75% correct responses.

Table 1: Average and maximum CIELAB colour differences between original and approximated images generated using different numbers of basis functions.

number of basis functions	24 reflectances		384 reflectances	
	average	max	average	max
1	37.49	70.74	32.11	71.93
2	30.48	109.47	22.83	109.47
3	6.14	19.92	3.79	28.65
4	3.71	12.44	2.34	19.16
5	1.48	6.04	1.08	6.04
6	1.39	4.22	1.00	5.31
7	0.40	1.61	0.31	1.78
8	0.38	1.38	0.25	1.94
9	0.27	0.77	0.19	1.18

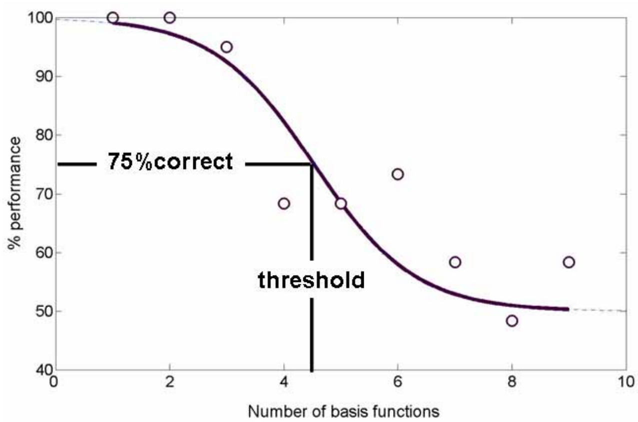


***Figure 4**: Observer's percent-correct result for one sequence with a particular spatial structure. The data is fitted by logistic regression. The threshold is chosen corresponding to 75% correct.*

Figures 5 and 6 show the results of the first and second experiment respectively. For the experiment with only six distinct colours, the thresholds (in terms of number of basis functions n corresponding to 75% correct) were pooled over the four sets of 6 spectra and shown separately for each observer. In Figure 5 the thresholds are plotted against spatial frequency. We note that in each stimulus many spatial frequencies are present. However, for each stimulus the first harmonic frequency has been calculated and is used to represent the scale properties of the stimulus. As the size of the patches was reduced (left-to-right in the middle row of Figure 2) the spatial frequency of the first harmonic increased. Recall again that when the threshold number of basis functions is small, this implies that observers' ability to discriminate between originals and reproductions for that condition is poor and that the linear model is effective. Figure 5 indicates that there is an effect of spatial frequency (or image complexity) on the number of basis functions required that is large compared with the standard error bars. It might be thought that when the patch size is small (corresponding to high spatial frequencies in Figure 5) observers would be unable to discriminate between original and linear-model reproductions even when the number of basis functions used in the linear model is small. However, the opposite trend is observed.

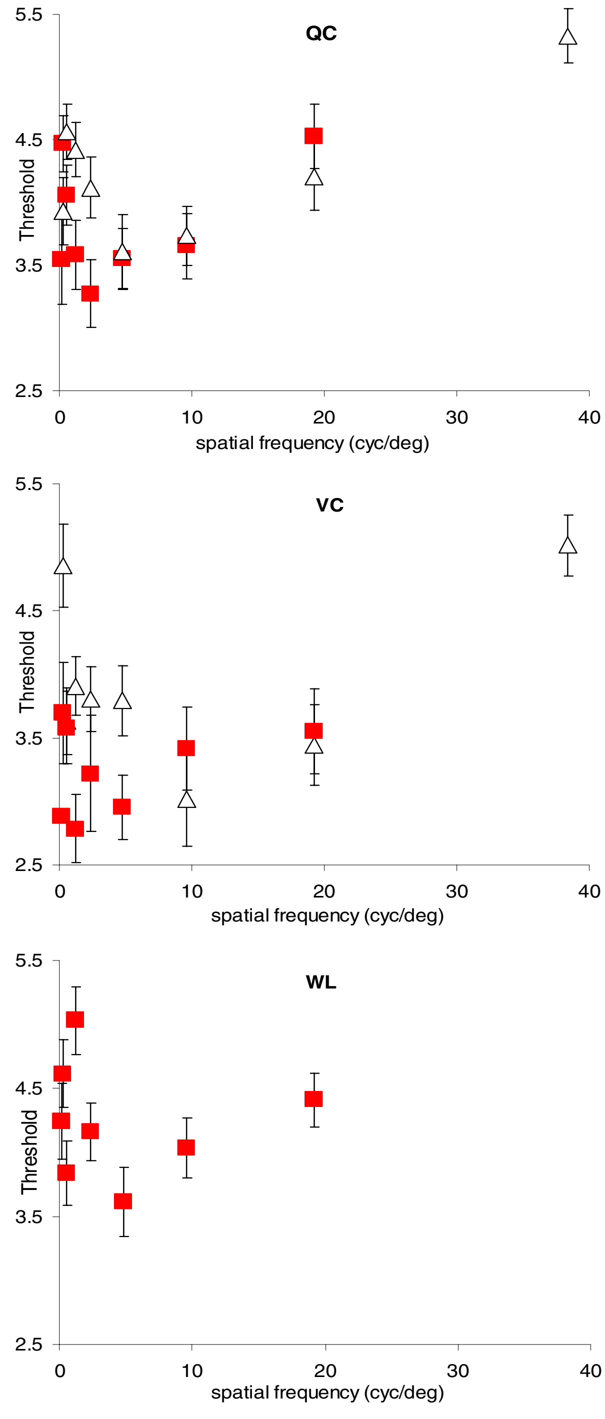


***Figure 5**: Discrimination thresholds for three observers. Threshold is plotted as number of basis functions against spatial frequency for all 4 image groups (±1 standard error is shown) under 2 viewing distances (squares: 80cm; triangles: 160cm).*

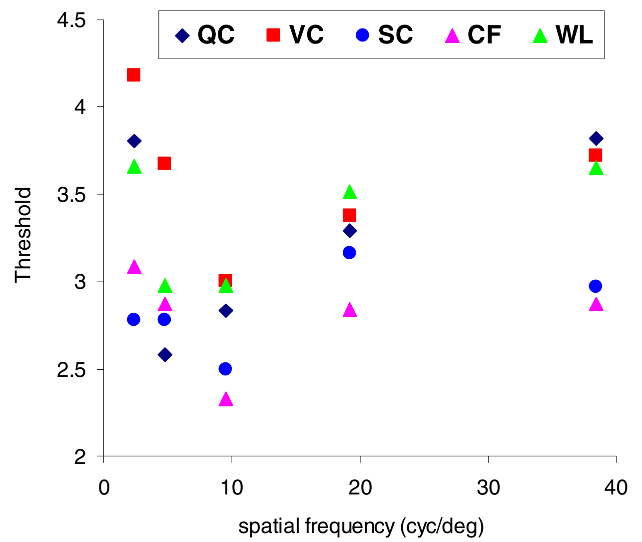
That is, as the patch size is reduced (in terms of degrees of visual angle) observers become better able to discriminate and this is best illustrated by the right-most points in Figure 5.

The right-most square symbol and right-most triangle symbol in each plot corresponds to the same stimulus but viewed from a different distance. Generally, the results show a u-shape; indicating that visual discrimination for this task is poorest for spatial frequencies in the region 10 cyc/deg. This was an unexpected result because in some other visual tasks (such as detection of luminance contrast) performance peaks around 10 cyc/deg. One possible explanation is that the number of distinct colours in each patch is too small. Performance when the patch size is tiny (see right-most image in Figure 2b – where the image is tending towards a spatially uniform patch) might be similar to when the patch size is very large (see left-most image in Figure 2b) when the number of distinct colours is small. It is for this reason that related experiments were conducted using 384 distinct colours rather than 24 colours.

Figure 6 illustrates the results of the experiment using images with 384 distinct colours. The thresholds (in terms of number of basis functions *n*) were pooled over the Mondrian-images set and are shown separately in Figure 6 for each of 5 observers. Although the absolute individual thresholds vary between observers, there is still evidence of the u-shaped plot.

For many applications concerning spatially uniform colours it is considered appropriate to use linear models with only six basis functions[5]. One might imagine that for images of natural scenes the tolerance for discrimination between original images and those represented by basis functions would be greater. There is an analogy with colour difference: The threshold colour difference between two spatially uniform patches is often considered to be somewhat less than 1 CIELAB unit but for images the threshold colour difference is 3-5 CIELAB units[12]. However, in the case of linear-model representations Nascimento *et al.* reported that for natural scenes observers could make discriminations between originals and linear-model representations even when 7 or 8 basis functions were used[7]. Based on their work Nascimento *et al.* suggested that the number of basis functions needed to represent images might need to be revised upwards. This study supports Nascimento *et al.*'s finding in that discrimination performance in our task is surprisingly poor even when relatively large numbers of basis functions are used.

In this study a criterion of 75% correct was used to determine the thresholds and the number of basis functions corresponding to this threshold performance was 3-6 depending upon image complexity. Our data show, however, that as the image complexity increases the number of required basis functions also increases. The disparity between our 3-6 basis functions and Nascimento *et al*.'s 7-8 basis functions may be attributed to the definition of the threshold. Nascimento *et al.* used two criterion levels of performance of 75% and 55% correct and argued that the 55% level was a compromise between the stability of the threshold estimate and the closeness of the criterion to the true chance level of 50%. If in this study a threshold criterion of 55% had been used instead of 75% then it is likely that the number of basis functions required at threshold would have been greater than 3-6. Indeed, in the Nascimento *et al.* study the number of basis functions corresponding to 75% correct (the same criterion threshold as in our study) was 5-6. An alternative approach to the threshold criterion was carried out in Oxtoby and Foster's study[13] where a d-prime analysis suggested that at least 5 basis functions are required for discrimination performance at threshold (in this case defined by d´ = 0) with Mondrian-like images. The Oxtoby and Foster work set a lower limit of 5 basis functions for discrimination performance. However, the novel feature of our work is to vary the image complexity (as defined by spatial frequency of the first harmonic in our graphs). We show that the number of basis functions required for a specified performance varies with the spatial properties of the image.



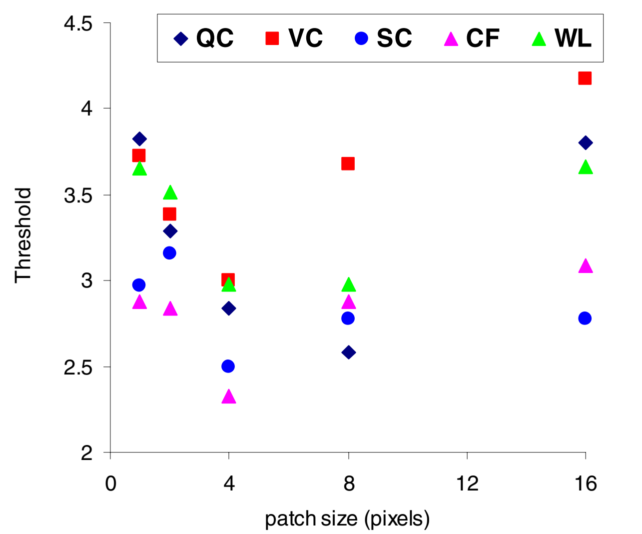


***Figure 6**: Discrimination thresholds for five observers. Threshold is plotted as number of basis functions against spatial frequency (top) and patch size (bottom) under viewing distance of 160cm.*

References

[1] L.T. Maloney, Evaluation of linear models of surface spectral reflectance with small numbers of parameters, *Journal of the Optical Society of America A*, **3** (10), 1673-1683, (1986).

[2] S. Westland, A.J. Shaw and H.C. Owens, Colour statistics of natural and man-made surfaces, *Sensor Review*, **20** (1), 50-55, (2000).

[3] A. Garcia-Beltran, J.L. Nieves, J. Hermandez-Andrez and J Romero, Linear bases for spectral reflectance functions of acrylic paints, *Color Research and Application*, **23** (1), 39-45, (1988).

[4] J.Y. Hardeberg, On the spectral dimensionality of object colours, *Proceedings of the First European Conference on Colour in Graphics, Imaging, and Vision*, Poitiers, France, 480-485, (2002).

[5] H.C. Owens, Colour and spatiochromatic processing in the human visual system, *PhD Thesis*, University of Derby, UK, (2002).

[6] R.S. Berns, *Billmeyer and Saltzman's principles of color technology*, 3rd ed., John Wiley & Sons, New York, (2002).

[7] S.M.C. Nascimento, D.H. Foster and K. Amano, Psychophysical estimates of the number of spectral-reflectance basis functions needed to reproduce natural scenes, *Journal of the Optical Society of America A*, **22** (6), 1017-1021, (2005).

[8] J. Parkkinen, J. Hallikainen and T. Jaaskelainen, Characteristic spectra of Munsell colors, *Journal of the Optical Society of America A*, **6** (10), 318-322, (1989).

[9] V. Cheung and S. Westland, Methods for optimal color selection, in press for *Journal of Imaging Science and Technology*, **50** (5), xx-xx, (2006).

[10] R.S. Berns and N. Katoh, Methods for characterizing displays, in *Colour Engineering*, P. Green and L.W. MacDonald (Eds.), John Wiley & Sons, (2002).

[11] F.A. Wichmann and N.J. Hill, The psychometric function: I. Fitting, sampling and goodness-of-fit, *Perception and Psychophysics*, **63** (8), 1293-1313, (2001).

[12] M. Stokes, M.D. Fairchild and R.S. Berns, Colorimetrically quantified visual tolerances for pictorial images, *ISCC/TAGA comparison of color images presented in digital media*, vol. 2, 757-777, (1992).

[13] E.K. Oxtoby and D.H. Foster, Perceptual limits on low-dimensional models of Munsell reflectance spectra, *Perception*, 34, 961-966, (2005).

Author Biography

Qiao Chen received his BSc degree in Computer Science and Technology from Southwest Jiaotong University (China) in 2000 and his MSc degree in Colour Imaging from the Colour Imaging Institute at the University of Derby (UK) in 2001. He is currently a PhD student at the University of Leeds (UK). His research interests include spectral reflectance modelling, multispectral imaging and human spatial vision.