Segmentation of natural scenes: Clustering in colour space vs. spectral estimation and clustering of spectral data

Jia Song, Eva M. Valero, Juan L. Nieves

Optics Department, Faculty of Science, University of Granada, Spain
Email: songjia815@gmail.com

In this paper, two approaches are implemented and compared in order to determine which one offers the better segmentation quality for natural scenes: one is finding the best colour space for colour-based segmentation, and the other is segmenting by using the spectral data obtained by estimation from sensor responses. Eight colour spaces including perceptual and non-perceptual spaces are evaluated, and pseudo-inverse spectral estimation method is used for obtaining the estimated spectral reflectances from three simulated sensor responses. K-means and spectral clustering using Nyström approximation are used for segmenting an image after an adaptive quantisation step by mean-shift method is applied on the image. The segmentation results are evaluated by measuring the degree of matching with the segmentation benchmark using a similarity metric based on Jaccard index, and the segmentation benchmark is created by manual labelling. Results show that using estimated spectral data for colour image segmentation of natural scenes can achieve better or equally good results as the best colour space among the tested eight colour spaces.

Received 31 October 2013; revised 19 May 2014; accepted 20 May 2014

Published online: 15 July 2014

Introduction

One of the most important and challenging tasks in image analysis and computer vision is image segmentation, which involves the process of determining the regions corresponding to the meaningful objects in the observed scene [1]. Segmentation of natural scenes is particularly challenging due to the variety of objects they contain and the complexity of the spatial structures. Different image features can be analysed in their corresponding feature spaces separately or jointly for the segmentation process, and various segmentation methods have been proposed (for a review on segmentation algorithms, see [2]). With the development of spectral imaging techniques, recent studies pay more and more attention to using spectral information for segmentation [3,4].

It is plausible to assume that using spectral information can help achieve better segmentation results than using only colour information due to the fact that spectral data offers a much more complete way of characterising the physical properties of the object surfaces than colour coordinates [4]. However it is not always feasible to obtain full spectral information unless some estimation algorithms are used to obtain it from a set of sensor responses [5,6,7]. Whether using estimated spectral information for image segmentation has the same advantages as real spectral information over colour spaces is investigated in this study.

In this paper, we perform clustering-based colour image segmentation in colour space and estimated spectral space. Two approaches are implemented and compared in order to determine which one offers the better segmentation quality for natural scenes: one is finding the best colour space among a representative selection of standard colour representation systems, and the other is segmenting by using the spectral data obtained by estimation from sensor responses. In practice, segmentation using the best colour space usually involves the process of colour space selection, which is a process very challenging and data dependent, while using estimated spectral data avoids the selection scheme and requires a training process for spectral estimation.

The paper is organised as follows. Section 2 introduces the data set and detailed methods applied in this work. Section 3 illustrates the experiment results and discussions. In section 4, several relevant conclusions and future work are discussed.

Method

Data Set, Segmentation Benchmark and Segmentation Evaluation

The set of colour images used in this work are from the Foster and Nascimento's database [8,9,10]. The spectrum for each pixel has been sampled at 10 nm intervals over 400-700 nm, resulting in 31 wavelength bands. As in this data set the segmentation benchmark is not available, it is created by manual labelling in this work. A pre-processing step using the mean-shift algorithm [11] is applied to perform an adaptive quantisation on the image before segmentation to reduce insignificant details in the natural scenes while preserving the edge information [12]. Figure 1 shows a colour rendering of the six 31-band spectral images tested in this paper, the manually labelled benchmark and the quantised images.

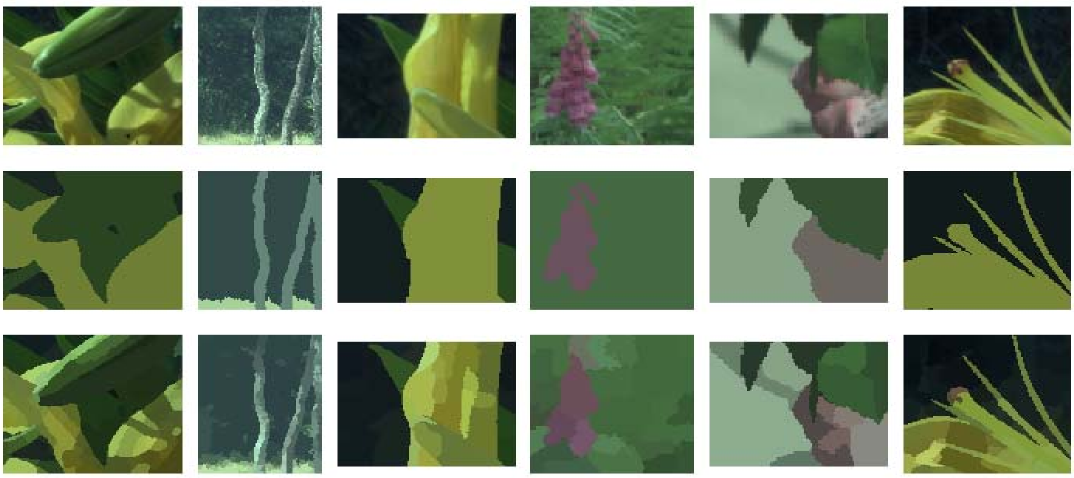

Figure 1: Test images (first row), segmentation benchmarks (second row) and quantised images by mean-shift (third row).

The segmentation results are evaluated quantitatively by computing a similarity metric that characterises the degree of matching with the benchmark. The similarity metric is defined as Jaccard Index [13] weighted by the size of each cluster, and is shown in Equation (1).

$$SIM = \sum_{k=1}^{c} \frac{n_k}{N} J(S_k, C_k) \qquad (1)$$

where c is the number of clusters in the segmentation result, n_k is the number of pixels belonging to the k-th cluster in the segmentation result, N is the total number of pixels in the image, S_k is the set of pixels in the k-th cluster in the segmentation result, C_k is the corresponding cluster in the benchmark which has the largest common area with S_k, $J(S_k,C_k)$ is the Jaccard Index which quantifies the similarity between S_k and C_k, and is defined in Equation (2).

$$J(S_k, C_k) = \frac{S_k \cap C_k}{S_k \cup C_k} \qquad (2)$$

The range of the evaluation metric value is from 0 to 1. A larger similarity metric value indicates a better segmentation result, and 1 indicates a perfect segmentation.

Colour spaces

The huge variety of colour spaces that have been developed over time can be classified into a four categories according to their definitions and properties [14].

1. The primary spaces are based on the trichromatic theory, which assumes that it is possible to match any colour by mixing an appropriate amount of the three primary colours. We selected (R,G,B), (X,Y,Z), (r,g,b) and (x,y,z) from this category in our work. (r,g,b) and (x,y,z) are normalised primary spaces of (R,G,B) and (X,Y,Z) by dividing each colour component value by the sum of the three tristimulus values.
2. The luminance-chrominance spaces are compounded of one colour component which represents the luminance and two colour components which represent the chrominance. We selected (L*,a*,b*) and (L*,u*,v*) in this work. We also tested one component (L*) and two-component combination (a*,b*) and (u*,v*) in segmentation in order to evaluated the luminance and chrominance components separately.
3. The statistical independent component spaces are resulted from different statistical methods like PCA. We used (I1,I2,I3) in this work, which is based on the Karhunen-Loeve transformation proposed by Ohta in 1980 [15].
4. The perceptual spaces try to quantify the subjective human colour perception by means of three measures: intensity, hue and saturation.

The colour spaces are calculated by linear and non-linear transformations from (R,G,B) sensor responses, which are simulated from spectral reflectance data using a set of spectral responsivities corresponding to a commercial CCD camera and the scene illumination is always assumed to be the standard illuminant D65.

Spectral estimation

In this work we use pseudo-inverse (PI) [16] algorithm to estimate the spectral reflectance of each pixel from simulated camera responses in an image scene. PI is the simplest linear regression estimation method based on the pseudo-inverse of the training camera responses matrix.

The segmentation performances of original spectral data (SP), normalised spectral data (NSP), estimated spectral data (ESP) and normalised estimated spectral data (NESP) are compared with the eight colour spaces evaluated in this paper. For NSP and NESP, the reflectance vector is normalised by its Euclidean norm so that the reflectance vector has unit Euclidean length. The normalisation reduces the effect of brightness change in the same way as the normalised primary colour spaces ((r,g,b) and (x,y,z)) do, while at the same time preserves the spectral signature of the surface and has the same advantage as using spectral information over colour spaces.

Clustering based colour image segmentation

In this paper, we perform colour image segmentation using clustering based methods. K-means (KM) [17] and spectral clustering (SC) using Nyström approximation [18] are used for performing the segmentation on the quantised image. KM follows an iterative process of minimising the sum of squared Euclidean distance of all the data points to their assigned cluster centres. By using Euclidean distance metric, KM tends to find hyper-spherical clusters with the same shape and orientation. SC is based on graph theory and finds clusters that have minimum disassociation with each other by minimising the normalised cut, and it takes into account both the intra- and inter- cluster similarity. Nyström approximation is applied in the SC process in order to improve the time and memory efficiency. The number of desired clusters is specified in advance as a prior knowledge.

Results

The average and standard deviation of the segmentation results (evaluated by weighted Jaccard Index) for all the feature spaces over the six test images using KM and SC are shown in Table 1 and Table 2, and partial segmentation results using KM are shown in Figure 2. Higher average and lower standard deviation indicates a better and more stable performance of a feature space.

Based on the experimental results we make the following observations:

1. K-means vs. spectral clustering
   SC performs better than KM in segmenting the six tested images by having higher average and lower standard deviation in the performance for most feature spaces.
2. Comparison between colour spaces
   Different colour spaces have different performances, and there is no colour space that has the best performance in segmenting all the images. The performance of the same colour space varies quite much in segmenting different images. For example in Figure 2, (H,S,V) is the best colour space for segmenting image 2, but it has the worst performance in segmenting image 1. It is the same with (R,G,B) in image 3 and image 4. Based on the average and standard deviation of the results in Table 1 and 2, (r,g,b) is the best colour space for segmenting the six tested images among the eight colour spaces evaluated in this work.
3. Normalised feature space vs. non-normalised feature space
   The normalised spectral data NSP and NESP and normalised colour space (r,g,b) and (x,y,z) perform better than the non-normalised spectral data and the colour spaces who are sensitive to brightness change. For example in image 1 and image 3 in Figure 2, because of shadows or highlights, the objects are inhomogeneous in colour appearance. NESP and (r,g,b) can successfully segment the objects and outperform the other feature spaces. (L*,a*,b*), (L*,u*,v*) and (H,S,V) are robust to lightness change to some degree but not as good as the normalised feature spaces.

4. Estimated spectral data vs. spectral data
   NESP performs similarly but not as well as the NSP, and we assume that its performance could be improved by increasing the spectral estimation accuracy.
5. Estimated spectral data vs. the best colour space
   NESP performs better than the best colour space (r,g,b) by having higher average and lower standard deviation in the performance over the six tested images.

Feature	(R,G,B)	(r,g,b)	(X,Y,Z)	(x,y,z)	(L*,a*,b*)	(L*,u*,v*)
KM	0.7398	0.9220	0.6934	0.8981	0.8635	0.8771
SC	0.8386	0.9497	0.7731	0.8740	0.9322	0.9490
Feature	(H,S,V)	(I1,I2,I3)	SP	NSP	ESP	NESP
KM	0.7839	0.7441	0.7638	0.9544	0.7381	0.9379
SC	0.8349	0.8047	0.8226	0.9676	0.8195	0.9665

Table 1: Average segmentation results of all the feature spaces using KM and SC.

Feature	(R,G,B)	(r,g,b)	(X,Y,Z)	(x,y,z)	(L*,a*,b*)	(L*,u*,v*)
KM	0.1523	0.0674	0.1017	0.0828	0.1713	0.1457
SC	0.1463	0.0481	0.1481	0.1175	0.0610	0.0497
Feature	(H,S,V)	(I1,I2,I3)	SP	NSP	ESP	NESP
KM	0.2309	0.1480	0.1518	0.0299	0.1510	0.0651
SC	0.1545	0.1357	0.1838	0.0173	0.1865	0.0158

Table 2: Standard deviation of the segmentation results of all the feature spaces.

Conclusions

The performance of natural scene colour image segmentation, when considering representation of the images in different colour spaces, is dependent on the scene content. We cannot find a colour space that performs well for all types of image content. Using estimated spectral data for segmentation can achieve better or equally good results as the best colour space among the tested eight colour spaces. Based on the findings of this paper, we can perform spectral estimation on the RGB image data and use normalised estimated spectral data for segmentation to achieve a better result in average without the need of performing colour space selection.

As using the full spectrum for clustering based colour image segmentation may include redundant information and is less efficient, we can try to apply some pre-processing steps in our future work, such as spectrum reduction based on statistical analysis [19]. We can also try to improve the spectral estimation quality by applying some training selection schemes [20] or using alternative spectral estimation approaches that outperform pseudo-inverse, such as kernel-based methods [7], and thereby improve the segmentation accuracy of estimated spectral data.

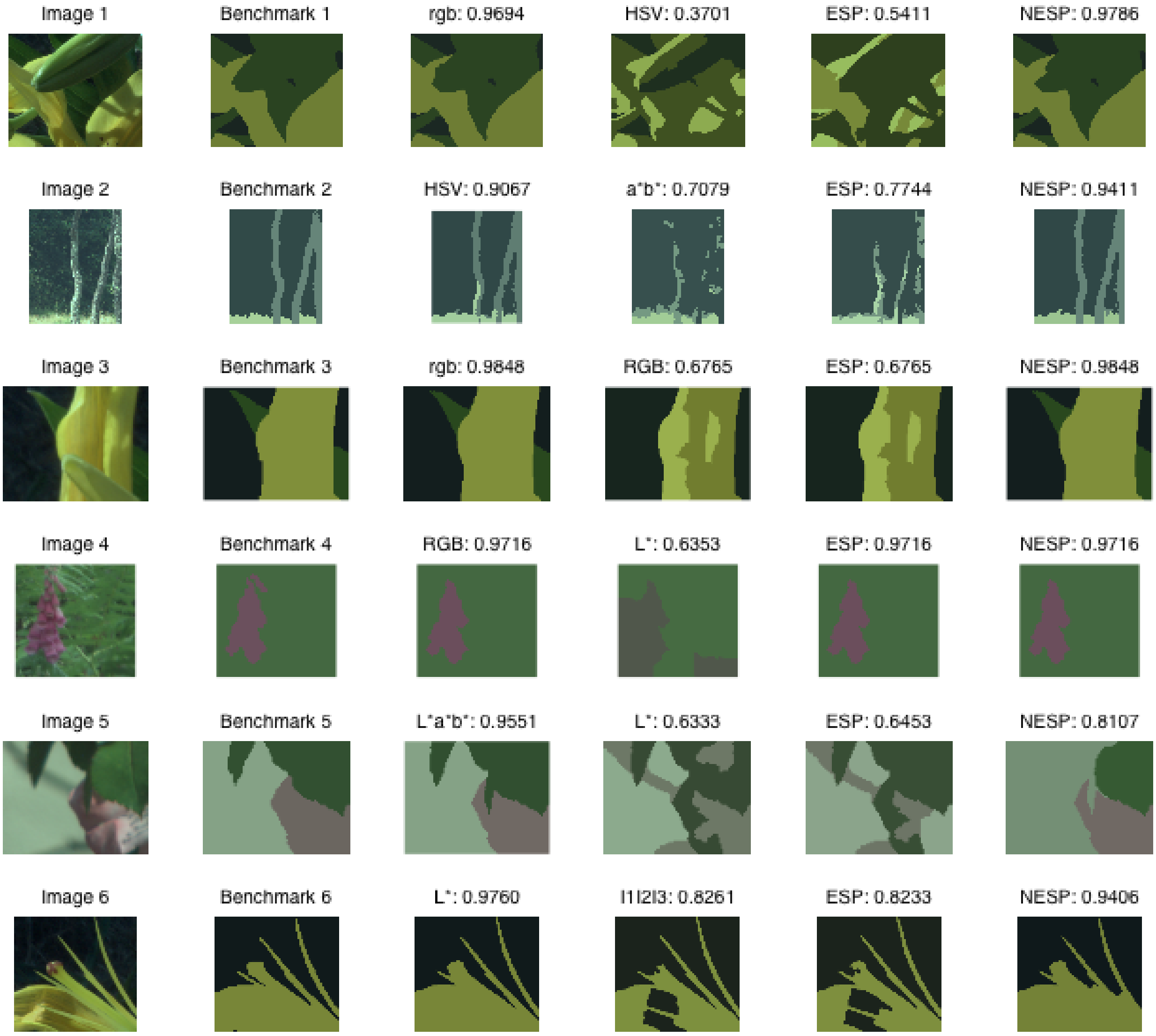

Figure 5: Anthraquinone – Laccaic acid [lac dye].

References

1. Busin L, Vandenbroucke N and Macaire L (2008), Colour spaces and image segmentation, *Advances in Imaging and Electron Physics,* **151**, 165-168.
2. Fu, King-Sun and Mui J K (1981), A survey on image segmentation, *Pattern Recognition* **13** (1), 3-16.
3. Altman D (1999), Efficient fuzzy clustering of multi-spectral images, *Proceedings of IEEE 1999 International in Geoscience and Remote Sensing Symposium*, **2**, 1594-1596.
4. Li H, Bochko V, Jaaskelainen T, Parkkinen J and Shen I (2008), Kernel-based spectral colour image segmentation, *JOSA A*, **25** (11), 2805-2816.
5. Imai F and Berns R (1999), Spectral estimation using trichromatic digital cameras, *Proceedings of the International Symposium on Multispectral Imaging and Colour Reproduction for Digital Archives*, 44-49.
6. Zhao Y and Berns R (2007), Image-based spectral reflectance reconstruction using the matrix R method, *Colour Research & Application*, **32** (5), 343-351.
7. Heikkinen V, Lenz R, Jetsu T, Parkkinen J, Hauta-Kasari M and Jääskeläinen T (2008), Evaluation and unification of some methods for estimating reflectance spectra from RGB images, *JOSA A*, **25** (10), 2444-2458.

8. Nascimento S, Ferreira F and Foster D (2002), Statistics of spatial cone-excitation ratios in natural scenes, *JOSA A*, **19** (8), 1484-1490.
9. Foster D, Nascimento S and Amano K (2004), Information limits on neural identification of coloured surfaces in natural scenes, *Visual Neuroscience*, **21** (3), 331-336.
10. Foster D, Amano K and Nascimento S (2006), Colour consistency in natural scenes explained by global image statistics, *Visual Neuroscience*, **23** (3/4), 341.
11. Comaniciu D and Meer P (2002), Mean shift: A robust approach toward feature space analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24** (5), 603-619.
12. Xing M, Li H, Jia J and Parkkinen J (2009), Fast spectral colour image segmentation based on filtering and clustering, *Sixth International Symposium on Multispectral Image Processing and Pattern Recognition, International Society for Optics and Photonics*, 74942Q-74942Q.
13. Polak M, Zhang H and Pi M (2009), An evaluation metric for image segmentation of multiple objects, *Image and Vision Computing*, **27** (8), 1223-1227.
14. Vandenbroucke N, Macaire L and Postaire J-G, Colour image segmentation by pixel classification in an adapted hybrid colour space, Application to soccer image analysis, *Computer Vision and Image Understanding*, **90** (2), 190-216.
15. Ohta Y, Kanade T and Sakai T (1980), Colour information for region segmentation, *Computer Graphics and Image Processing*, **13** (3), 222-241.
16. Nieves J, Valero E, Nascimento S, Hernández-Andrés J and Romero J (2005), Multispectral synthesis of daylight using a commercial digital CCD camera, *Applied Optics*, **44** (27), 5696-5703.
17. Steinhaus H (1956), Sur la division des corp materiels en parties. *Bull. Acad. Polon. Sci*, **1**, 801-804.
18. Fowlkes C, Belongie S, Chung F and Malik J (2004), Spectral grouping using the Nystrom method, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **26** (2), 214-225.
19. Rodarmel C and Shan J (2002), Principal component analysis for hyperspectral image classification, *Surveying and Land Information Science*, **62** (2), 115-122.
20. Shen H, Zhang H, Xin J and Shao S (2008), Optimal selection of representative colours for spectral reflectance reconstruction in a multispectral imaging system, *Applied Optics*, **47** (13), 2494-2502.