

Multi-spectral Texture Characterisation for Remote Sensing Image Segmentation

Filiberto Pla¹, Gema Gracia¹, Pedro García-Sevilla¹,
Majid Mirmehdi², and Xianghua Xie³

¹Dept. Llenguatges i Sistemes Informàtics, University Jaume I, 12071 Castellón, Spain
{pla, ggracia, pgarcia}@lsi.uji.es

²Dept. of Computer Science, University of Bristol, Bristol BS8 1UB, UK
majid@compsci.bristol.ac.uk

³Dept. of Computer Science, University of Swansea, Swansea SA2 8PP, UK
x.xie@swansea.ac.uk

Abstract. A multi-spectral texture characterisation model is proposed, the Multi-spectral Local Differences Texem – MLDT, as an affordable approach to be used in multi-spectral images that may contain large number of bands. The MLDT is based on the Texem model. Using an inter-scale post-fusion strategy for image segmentation, framed in a multi-resolution approach, we produce unsupervised multi-spectral image segmentations. Preliminary results on several remote sensing multi-spectral images exhibit a promising performance by the MLDT approach, with further improvements possible to model more complex textures and add some other features, like invariance to spectral intensity.

Keywords: Texture analysis, multispectral images, Texems.

1 Introduction

Multi and hyperspectral sensors acquire information in several spectral bands, which generate hyperspectral data in high dimensional spaces. These systems have traditionally been used to perform tasks in remote sensing, and are being introduced and developed in other application fields like medical imaging or product quality assessment. Multispectral image data are used in order to estimate and analyze the presence of types of vegetation, land, water and other man made objects, or to assess the quantity of substances, chemical compounds, or physical parameters, e.g. temperature, providing a qualitative and quantitative evaluation of those features.

Standard multispectral image interpretation techniques barely exploit the spectral-spatial relationships in the image. The multi-spectral image data is basically treated as a set of independent spectral measurements at each pixel location, without taking into account their spatial relations. In order to exploit hyper-spectral imagery in applications requiring high spatial resolution, e.g., urban land-cover mapping, crops and vegetation mapping, tissues structure identification, it is necessary to incorporate spatial [1], contextual [2] and texture information in the multi-spectral image classification and segmentation processes.

To the best of our knowledge, there are no texture characterisation methods for multispectral images with high number of bands. Such methods are unaffordable to be used directly in gray level and colour images due to the increase of dimensionality in texture characterisation. Multi-band images techniques have been traditionally restricted to three-band colour images, by processing each channel independently, taking into account spatial interactions only within each channel [3]. Other approaches decompose the colour image into luminance and chromatic channels, extracting texture features from the luminance channel [4]. There are works that try to combine spatial interaction within each channel and interaction between spectral channels, applying gray level texture techniques to each channel independently [5], or using 3D colour histograms as a way to combine information from all colour channels [6].

Another group of techniques try to extract correlation features between the channels for colour texture analysis, like in [7], where spatial and spectral interactions are simultaneously handled. Such techniques assume the image to be a collection of epitomic primitives, and the neighbourhood of a central pixel to be statistically conditionally independent. A more recent approach based on these premises is the Texem model [8], consisting of a Gaussian mixture model representation for colour images using conditional dependency in neighbouring and cross-channels information. The gray level Texem model assumes spatial conditional dependency within the pixel neighbourhood. The Texem model will be the basis of the work presented in this paper for texture characterisation in multispectral images.

2 The Texem Model

The Texem model [8] is a texture characterisation method that models the image as a generative process where a set of image primitives generate the image by superposition of image patches from a number of texture exemplars, Texems.

This generative model uses a Gaussian mixture to obtain the Texems that have generated an image. The Texems are derived from image patches that may be of any size and shape. In this work, square image patches of size $N=n \times n$ have been considered. The image I is decomposed as a set of $Z = \{Z_i\}_{i=1}^P$ patches, each one belonging to any of K possible Texems, $T = \{t_k\}_{k=1}^K$. A patch vector at a central pixel location i is defined as $Z_i = (g_{i1}, \dots, g_{iN})$, with the gray level values $g_{ij} = I(x_{ij}, y_{ij})$ at pixel locations $ij = i1, \dots, iN$ in the patch grid. Each Texem is modelled as a Gaussian distribution, therefore, given the k th Texem t_k , the likelihood of a patch Z_i is expressed as a Normal distribution

$$p(Z_i | t_k, \theta_k) = G(Z_i; \mu_k, \Sigma_k) \tag{1}$$

where $\theta_k = (\alpha_k, \mu_k, \Sigma_k)$ is the parameter set defining the Gaussian in a mixture, with the prior probability α_k , mean μ_k and covariance Σ_k , constrained to $\sum_{k=1}^K \alpha_k = 1$.

Given a set of sample patches extracted from an image, the generative Gaussian Mixture model of the K Texems that generated that image can be estimated by the Expectation Maximisation (EM) algorithm [9]. Thus, the probability density function of any image patch Z_i will be given by the Gaussian mixture model,

$$p(Z_i | \alpha) = \sum_{m=1}^K \alpha_m p(Z_i | t_m, \theta_m) \tag{2}$$

A straightforward way to extend the gray level Texem model to colour images would be consider instead of the image patch $Z_i = (g_{i1}, \dots, g_{iN})$ of N pixel values in a monochrome image, an image patch at pixel i in a three band colour image, e.g. an RGB image, as $Z_i = (g_{i1}^R, \dots, g_{iN}^R, g_{i1}^G, \dots, g_{iN}^G, g_{i1}^B, \dots, g_{iN}^B)$. This increases the feature patch dimensionality in a proportional way with respect the number of bands. In order to avoid the increase in dimensionality of the generative Texem model for colour images, the pixels $i=1, \dots, N$ within the patch are assumed to be statistically independent within the Texem, with each pixel value following a Gaussian distribution in the colour space [8]. Thus, now the likelihood of a patch Z_i given the k th Texem t_k is expressed as a joint likelihood of the N pixels belonging to the patch, that is,

$$p(Z_i | t_k, \theta_k) = \prod_{j=1}^N G(Z_{j,i}; \mu_{j,k}, \Sigma_{j,k}) \tag{3}$$

where now the k th Texem parameters $\theta_k = (\mu_{1,k}, \Sigma_{1,k}, \dots, \mu_{N,k}, \Sigma_{N,k})$ are the mean $\mu_{j,k}$ and covariance $\Sigma_{j,k}$ of the $j=1, \dots, N$ pixels in the Texem grid. The mean $\mu_{j,k}$ and the covariance $\Sigma_{j,k}$ of each pixel are now defined in the colour space.

3 Segmentation with Inter-scale Post-fusion

In general, to model the texture features of an image appropriately, several Texem sizes are needed. Alternatively, instead of using different Texem sizes to characterise the set of patches that may generate an image, the same Texem size can be used in a multiresolution scheme, assuming each resolution level is generated from a Texem set independently [10]. However, applying multi-resolution to image segmentation needs a fusion process in order to integrate the information across the different image resolution levels, from coarser to finer levels. We follow this approach in this paper.

4 Multi-spectral Local Difference Texems - MLDT

The colour Texem model described in section 2.2 can be easily extended from colour images, usually represented by 3 bands, to any number of bands B . However, multi and hyper-spectral images may contain order of hundred bands to represent each pixel location. This will lead to the estimation of the Gaussian mixtures in

very high dimensional spaces, involving more computational complexity issues and to the so-called curse of dimensionality, when having a limited number of samples to estimate the Gaussian mixtures. In order to cope with such a high dimensionality problem, each image patch Z_i at a pixel location i in a multi-spectral image I with B bands will be defined as follows,

$$Z_i = (\bar{g}_{i1}, \dots, \bar{g}_{iB}, d(\mathbf{g}_{i1}, \bar{\mathbf{g}}), \dots, d(\mathbf{g}_{iN}, \bar{\mathbf{g}})) \tag{4}$$

where \bar{g}_{ib} is defined as

$$\bar{g}_{ib} = \frac{1}{N} \sum_{j=1}^N g_{ijb}; \quad b = 1, \dots, B$$

denoting the mean value of the N pixels in the image patch grid for each band $b=1, \dots, B$; and

$$d(\mathbf{g}_{ij}, \bar{\mathbf{g}}_i) = \frac{1}{B} \sum_{b=1}^B |g_{ijb} - \bar{g}_{ib}|; \quad j = 1, \dots, N$$

is the $L1$ norm of the spectral differences between pixels $j=1, \dots, N$ in the image patch Z_i at pixel location i and the mean spectrum $\bar{\mathbf{g}}_i = (\bar{g}_{i1}, \dots, \bar{g}_{iB})$ in the image patch Z_i . In the present work, $L1$ norm has been used, which would represent in a continuous spectral representation the area between two spectral power spectra, although other norms or spectral difference measures could be used. Analogously, instead of the patch mean $\bar{\mathbf{g}}_i$ as the patch spectral reference, other possible spectral image patch representatives could be used, e.g. the median spectral pixel or the spectrum of the central pixel in the patch.

Image patch Z_i is then a $B+N$ dimensional vector, with B the number of bands and N the number of pixels in the image patch grid. Note that given a patch size N , for any spectral number of bands B , the dimensionality of the texture feature vector has always a fixed part size of N difference terms, and only the mean spectral pixel increases linearly as the number of bands B increases. This is a desirable property of the texture characterisation in the multi-spectral domain, since the complexity of the Texem model is controlled, keeping dimensionality to an affordable way. In addition, if a band reduction technique is used as a previous step [11], the Texem dimensionality can even be kept at a more reduced and affordable level.

The MLDT characterisation captures in a compact way the difference patterns within an image patch in a multi-spectral image, and is thus able to represent the integrated spatial and spectral information in a single representation. Using the image patch representation expressed in (8) will enable the use of the gray level Texem model in section 2.1 directly, keeping spectral and spatial dependencies in the generative model at an affordable feature vector dimensionality.

5 Experimental Data

In order to test the validity of the proposed MLDT characterisation, it has been applied to a set of three hyper-spectral remote sensing images captured with different sensors:

- DAISEX099 project provides useful aerial images about the study of the variability in the reflectance of different natural surfaces. This source of data, referred as HyMap in figures, corresponds to a spectral image of 700×670 pixels and 7 classes of identified crops and other unknown land use class, acquired with the 128-bands HyMap spectrometer during the DAISEX-99 campaign (<http://io.uv.es/projects/daisex/>). In this case, 126 bands were used, discarding the lower SNR bands (0, 64). Figure 1 (left) represents a pseudo-colour image composed from three of the 126 bands.
- Satellite PROBA has a positional spectra-radiometric system (CHRIS) that measures the spectral radiance. The images used in this study come from the CHRIS-PROBA mode that operates on an area of 15×15 km, with a spatial resolution of 34m obtaining a set of 62 spectral bands that range from 400 to 1050 nm, 641×617 pixels and 9 classes of identified crop types and other unknown land use classes. In this case, 52 bands were used, discarding the lower SNR bands (25, 33, 36-37, 41-43, 47, 50, 53). Figure 2 (left) represents a pseudo-colour image composed from three of the 52 bands.
- A third multi-spectral image of seven spectral bands and 512×512 pixels, obtained from LandSat-7 of an area around the Kilauea Volcano, in Hawaii. This image will be referred as LandSat-7. Figure 3 (left) represents a pseudo-colour image composed from three of the 7 bands.

Considering multi-spectral images can contain a huge amount of information with a high number of bands, and taking into account that most of these bands are very correlated [11], it seems logical that the dimensionality reduction problem in multi-spectral images has to be linked with the texture characterisation problem, since trying to combine correlations simultaneously in the spectral and spatial domain can be computationally expensive.

In order to exploit inter-band correlation to reduce the multi-spectral band representation, the unsupervised band reduction technique by [11] has been used to reduce the bands of HyMap and CHRIS to the seven most relevant bands. This band reduction technique exploits inter-band correlation to reduce the multi-spectral band representation by means of theoretic information concepts. The seven bands selected by the band selection algorithm for HyMap images have been bands (1, 28, 41, 52, 79, 107, 122), and for CHRIS image, bands (0, 9, 20, 30, 40, 46, 59).

With the selected bands for every image of this dataset, an unsupervised image segmentation algorithm has been applied, based on an EM algorithm for a Gaussian mixture model, fixing the number of Texems as input parameter, and the inter-scale post-fusion strategy pointed out in section 3. The results are discussed in the next section.

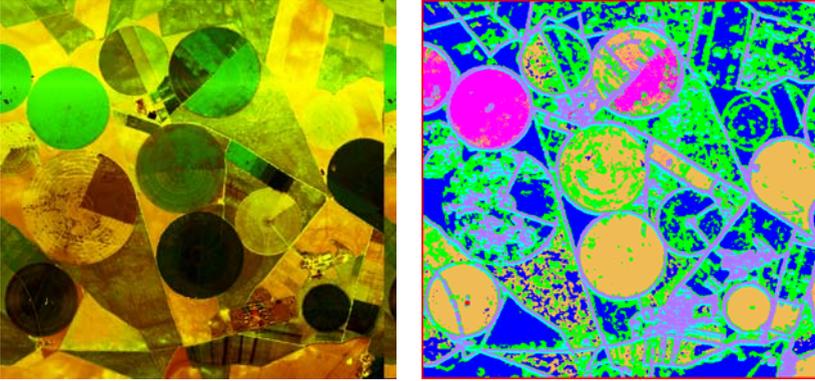


Fig. 1. HyMap pseudo-colour image (left) and its MLDT-based segmentation (right)

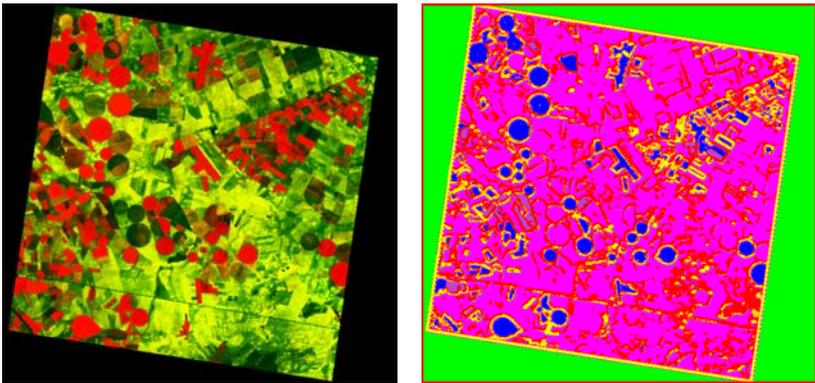


Fig. 2. CHRIS pseudo-colour image (left) and its MLDT-based segmentation (right)

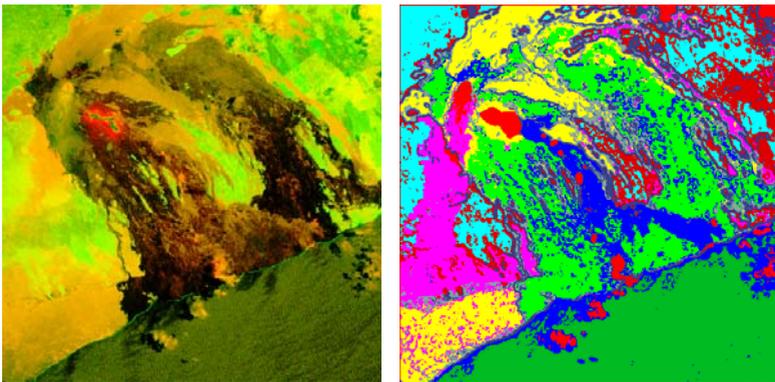


Fig. 3. LandSat-7 pseudo-colour image (left) and its MLDT-based segmentation (right)

6 Results

Figure 1 (right) shows the result of the proposed method to the HyMap multispectral image, using the selected 7 image bands, with $L=3$ multi-resolution levels, $K=12$ Texems and $N=7\times 7=49$ image patch size. The Texem model was trained with 3000 random image patches in each level. Note how the MLDT multi-spectral characterisation and inter-scale post-fusion segmentation has been able to identify the most important texture types in the image, grouping them in a satisfactory way. Note how the image Texems found model most of the structures of the image, finding the main regions corresponding to the different crop and land uses in the image.

Analogously, Figure 2 (right) shows the result of the algorithm for the CHRIS multispectral image, using the selected 7 image bands, with $L=3$ multi-resolution levels, $K=8$ Texems and $N=7\times 7=49$ image patch size. The Texem model was trained with 4000 random image patches in each level. In this case, the Texems found correspond to the main three types of land uses. Note how it is also modelled the different types of crop/land borders, where we can distinguish fairly well at least two different border structures modelled by their corresponding Texems.

Finally, Figure 3 (right) shows the result of the algorithm for the LandSat-7 multispectral image, using the 7 image bands, with $L=2$ multi-resolution levels, $K=10$ Texems and $N=3\times 3=9$ image patch size. The Texem model was trained with 3000 random image patches in each level. The results on this image show how well the different land and water types have been extracted, being able to discriminate even distinct water sediments levels. Another important detail is how the spectral information has also been able to detect the area covered by the smoke from the volcano, which cannot be visually appreciated very well from the pseudo-colour composition, only near the volcano, but the MLDT characterisation has been able to represent.

It is worth noting that when using 7 image bands and a 49 pixel patch size, the MLDT vector has $7+49=56$ dimensions, and the Texem model in this case is defined by a single Gaussian of 56 dimensions in a Gaussian mixture model. In the case of a 9 pixel patch size, Texem dimensionality reduces to $7+9=16$ dimensions. If the multi-spectral image had 50 dimensions, the Texem dimensionality for a $3\times 3=9$ patch size would be $50+9=59$, which is still affordable.

7 Conclusions

As a final remark, the basis for a multi-spectral texture characterisation technique has been built, with promising results and affordable complexity to deal with the huge amount of data a multi-spectral image may contain, capturing the essential properties of the spatial and spectral relationships.

Possible improvements and variations of the proposed MLDT characterisation can be done in several areas, for instance, if the mean spectra of image patches are dropped, the MLDT vector then reduces considerably and for a given image patch size, the MLDT vector dimension is constant (the image patch size) for any number of bands. This is a particular interesting property to be studied in further work.

Eventually, more tests should also be directed to model Texems as several Gaussian components, as pointed out in [10], exploring some hierarchical clustering structure as in [12], to merge Gaussian modes to form clusters.

Acknowledgments. This work has been partially supported by grant PR2008-0126 and projects ESP2005-00724-C05-05, AYA2008-05965-C04-04/ESP, CSD2007-00018 and PET2005-0643 from the Spanish Ministry of Science and Innovation, and by project P1 1B2007-48 from Fundació Caixa-Castelló.

References

1. Plaza, A., Martínez, P., Plaza, J., Pérez, R.: Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing* 37(6), 1097–1116 (2005)
2. Camps-Valls, G., Gomez-Chova, L., Munoz-Mari, J., Vila-Frances, J., Calpe-Maravilla, J.: Composite kernels for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters* (3), 93–97 (2006)
3. Haindl, M., Havlicek, V.: A Simple multispectral multiresolution Markov texture model. In: *International Workshop on Texture Analysis and Synthesis*, pp. 63–66 (2002)
4. Dubuisson-Jolly, M., Gupta, A.: Color and texture fusion: Application to aerial image segmentation and GIS updating. *Image and Vision Computing* 18, 823–832 (2000)
5. Palm, C.: Color texture classification by integrative co-occurrence matrices. *Pattern Recognition* 37(5), 965–976 (2004)
6. Mirmehdi, M., Petrou, M.: Segmentation of color textures. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(2), 142–159 (2000)
7. Jovic, N., Frey, B., Kannan, A.: Epitomic analysis of appearance and shape. In: *IEEE International Conference on Computer Vision*, pp. 34–42 (2003)
8. Xie, X., Mirmehdi, M.: TEXEMS: Texture exemplars for defect detection on random textured surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29(8), 1454–1464 (2007)
9. Bouman, C.A.: Cluster: An unsupervised algorithm for modelling Gaussian mixtures (April 1997), <http://www.ece.purdue.edu/~bouman>
10. Xie, X., Mirmehdi, M.: Colour image segmentation using texems. *Annals of the BMVA* 2007(6), 1–10 (2007)
11. Martinez-Uso, A., Pla, F., Sotoca, J.M., Garcia-Sevilla, P.: Clustering-based Hyperspectral Band Selection using Information Measures. *IEEE Transactions on Geoscience & Remote Sensing* 45(12), 4158–4171 (2007)
12. Pascual, D., Pla, F., Sánchez, J.S.: Non Parametric Local Density-based Clustering for Multimodal Overlapping Distributions. In: Corchado, E., Yin, H., Botti, V., Fyfe, C. (eds.) *IDEAL 2006*. LNCS, vol. 4224, pp. 671–678. Springer, Heidelberg (2006)