COLOR AND TEXTURE FEATURE FUSION USING KERNEL PCA WITH APPLICATION TO OBJECT-BASED VEGETATION SPECIES CLASSIFICATION

Zhengrong Li [1,2], Yuee Liu [2], Ross Hayward [1], Rodney Walker [2]

[1] Queensland University of Technology, Australia
[2] Northwest A&F University, China

ABSTRACT

A good object representation or object descriptor is one of the key issues in object based image analysis. To effectively fuse color and texture as a unified descriptor at object level, this paper presents a novel method for feature fusion. Color histogram and the uniform local binary patterns are extracted from arbitrary-shaped image-objects, and kernel principal component analysis (kernel PCA) is employed to find nonlinear relationships of the extracted color and texture features. The maximum likelihood approach is used to estimate the intrinsic dimensionality, which is then used as a criterion for automatic selection of optimal feature set from the fused feature. The proposed method is evaluated using SVM as the benchmark classifier and is applied to object-based vegetation species classification using high spatial resolution aerial imagery. Experimental results demonstrate that great improvement can be achieved by using proposed feature fusion method.

Index Terms— geographic object-based image analysis (GEOBIA), color-texture feature fusion, kernel principal component analysis, local binary patters, vegetation classification

1. INTRODUCTION

The use of traditional coarser spatial resolution satellite imagery has proven inadequate for discriminating species-level vegetation in detailed vegetation studies [1]. Airborne high spatial resolution (HSR) imagery provides more information for detailed observation of vegetation. However, traditional classification algorithms based on single pixel analysis are often not capable of extracting the information we desire from HSR images. In recent years, object-based approaches become popular in HSR image classification, which has proven to be an alternative to the pixel-based image analysis and better results can be expected [2]. It is noted that in the remote sensing and GIS community, a new discipline of geographic object-based image analysis (GEOBIA) has gained wide spread interest, although a critical discussion has arisen concerning whether or not geographic space should be included in the name of this concept in order to discriminate from other discipline like computer vision and biomedical imaging, which also conduct object-based image analysis (OBIA) [2].

A good object representation or object descriptor is one of the key issues in object based image analysis. Texture and color are two fundamental features in describing an image, but prior research generally focus on extracting color and texture feature as separate entities rather than a unified image descriptor [3]. The use of color and texture information collectively has strong links with the human perception, and this motivates investigating how to effectively fuse color and texture as a unified descriptor to improve the discrimination over viewing color and texture features independently. Although the motivation of using color and texture information jointly in object-based image classification is clear, how best to combine color and texture in a unified object descriptor is still an open issue. Huang el al. [4] proposed a multiscale spectral and spatial feature fusion method based on wavelet transform and evaluated in very high resolution satellite image classification. Zhang et al. [5] extracted texture features using multi-channel Gabor filters and Markov random fields integrated the two features using a neighbourhood-oscillating tabu search approach for high-resolution image classification. However, these methods extract features from fixed window size and do not consider all pixels within an object as a whole. Moreover, heavy computational burden is induced by combining multiple features, which may cause the 'the curse of dimensionality' problem and decrease the performance of the classifier.

In this paper, we propose an object-based color texture fusion method based on kernel principal component analysis (Kernel PCA). This method has been evaluated in a specific application: vegetation species classification in power line corridors using airborne HSR imagery.

2. OBJECT-BASED COLOR AND TEXTURE FEATURE EXTRACTION

The object-based classification is substantially different from a per-pixel classification as it is conducted in object-feature space. Successful object-based image analysis results largely depend on the performance of image segmentation. Since we are going to classify the species among trees, tree crowns are the only image-objects of interest in our research. The aim of segmentation is, therefore, to detect and delineate all tree crowns from images while eliminating other image regions. We have developed an automatic tree crown detection and delineation algorithm by utilizing spectral features in a pulse coupled neural network followed by post-processing using morphological reconstruction [6]. Although the automatic segmentation is satisfied from visual assessment, decomposition of tree clusters is occasionally poor. Since the main aim of this research is to evaluate the effectiveness of kernel PCA in object-level color and texture feature fusion, manual segmentation is used to minimize the influence of under-segmentation.

The image-objects generated from segmentation is arbitrary-shaped, however, texture measurements are usually extracted based on the texture property of pixels or small blocks within the rectangular shaped region. Therefore, in this paper, the arbitrary-shaped objects are extended to a rectangular area for texture extraction. This can be achieved by padding zero or mean value outside the object boundary, or obtaining the inner rectangle from the object. Zero padding introduces spurious high frequency components leading to degrading the performance of the texture feature, while the inner rectangle cannot usually represent the property of the entire object well. Mean-intensity padding has shown better performance than the other two approaches [7] and thus is adopted in this paper. Firstly, the minimum bounding rectangle is obtained from the image segment, and then the area which is outside of the segment and inside of the minimum bounding rectangle is padded using the mean value of pixels in the region.

In this paper, an extension of the classic Local Binary Pattern (LBP), the uniform LBP (ULBP) is used as texture feature. ULBP contains at most two bitwise (0 to 1 or 1 to 0) transitions The occurrence histograms of the ULBP are computed using $P = 8, 16, 24$, with $R = 1, 2, 3$ respectively, which is claimed to have the best performance of the local binary patterns in the experiments conducted by Ojala *et al.* [8]. The features are obtained by combining the three sets of features together with a dimension of 857.

Color histograms are the most widely used statistical features in computer vision. They are often used for the illumination independent characterization of the color distribution of the pattern. In this paper, color histograms extracted each spectral channel are used to represent the color feature of each image-objects (tree crowns). Another advantage of using color histogram is that it can be easily integrated with ULBP texture features as they are both statistical features and can be combined in one histogram.

3. KERNEL PCA FOR FEATURE FUSION

It is often difficult to classify objects using single feature descriptor. Therefore, feature-level fusion plays an important role when multiple features are used in the process of object classification. The advantages of feature fusion are: 1) the most discriminatory information from original multiple feature sets can be derived by the fusion process; 2) the noisy information can be eliminated from the correlation between different feature sets. In other words, feature fusion is capable of deriving and gaining the most effective and least-dimensional feature vectors that benefit the final classification [9].

The proposed object-level feature fusion is described as follows. After object segmentation, color and texture features are extracted from image-objects. Different feature vectors are normalized and then serially integrated. After that, kernel PCA is used to globally extract the nonlinear features from the integrated feature sets as well as to reduce the dimensionality. The intrinsic dimensionality of the serial fused features is estimated using a maximum likelihood method to select the target dimensionality from the kernel PCA fused feature. Finally, features selected from Kernel PCA are used as the input to classifiers for further analysis.

We first use a serial fusion strategy [9] by simply combining different feature vectors into one set of feature union-vector. Different features vectors are combined into one set of feature union-vector. As the features are different on the value scope, they are initialized into range [-1,1] by Gaussian criterion. Suppose α and β are two feature vectors which are extracted from the same image-object. The integrated feature union-vector is defined by $\gamma = \binom{\alpha}{\beta}$. Obviously, if feature vector α is m-dimensional and β is n-dimensional, then the serial integrated feature vector γ is (m+n)-dimensional. The integration of color histogram and ULBP has a high dimensionality of 1878, which cause the 'curse of dimensionality' problem. We employ kernel PCA as a fusion method to define the best subspace for nonlinear feature extraction and reduce the dimensionality of the serial fused feature.

Kernel PCA is the reformulation of traditional linear PCA in a high-dimensional space using a kernel function. Kernel PCA firstly map the original input vectors x_i into a high-dimensional feature space $K(x_i)$ and then calculate the linear PCA in $K(x_i)$. Performing PCA in the high-dimensional feature space can obtain high-order statistics of the input variables, which is also the initial motivation of kernel PCA [10]. However, it is difficult to directly compute both the covariance matrix and its corresponding eigenvectors and eigenvalues in the high-dimensional feature space. Therefore, kernel tricks are employed to avoid this difficulty and the principal eigenvectors are computed from the kernel matrix, rather than the covariance matrix.

Assuming that the kernel matrix is centered, i.e. $\sum_{i=1}^{M} K(x_i) = 0$, where M is the number of input variables, kernel PCA computes the principal d eigenvectors v_i of the centered kernel matrix. The corresponding eigenvectors of the covariance matrix β_i can be derived from the eigenvectors of the kernel matrix α_i as follows:

$$\beta_i = \frac{1}{\sqrt{\lambda}} Q \alpha_i \tag{2}$$

where $Q = [K(x_1),...,K(x_M)$ is the mapped data matrix on the high-dimensional feature space.

In order to obtain the low-dimensional feature representation, the data is projected onto the eigenvectors of the covariance matrix β_i. The result of low-dimensional data representation Y is:

$$y_i = \left\{ \sum_{j=1}^{n} \beta_1^j K(x_j, x_i), \ldots, \beta_d^j K(x_j, x_i) \right\} \tag{3}$$

where β_1^j indicates the j^{th} value in the vector β_1, and K is the kernel function that was also used in the computation of kernel matrix. The mapping performed by kernel PCA relies on the choice of the kernel function. In this paper, a Gaussian kernel is employed which is widely used in many applications.

$$K(x_i, x_j) = exp\left(-\frac{\left\| x_i - x_j \right\|^2}{2\delta^2} \right) \tag{4}$$

In order to automatically obtain the optimal dimension number of the fused color-texture feature by kernel PCA, a maximum likelihood estimator (MLE) [11] is employed to estimate the intrinsic dimensionality. MLE consider the data points in the hypersphere as a Poisson process, in which the estimated intrinsic dimensionality d around data point x_i in given k nearest neighbors is given by

$$\hat{d}_k(x_i) = \left(\frac{1}{k-1} \sum_{j=1}^{k-1} log \frac{T_k(x_i)}{T_j(x_i)} \right)^{-1} \tag{5}$$

where $T_k(x_i)$ represents the radius of the smallest hypersphere with center x_i that covers k neighboring data points.

4. EXPERIMENT AND RESULT

A. Data Collection

There are few public databases available for vegetation species classification by means of remotely sensed imagery. The data used in this study were collected in rural Queensland Australia in November 2008 for research into vegetation management in power line corridors. The reason why we need species information of individual trees is that vegetation management in power line corridors is based on their potential risks to power lines. For example, species with fast growing rates and that also have the potential to reach a mature height of more than four meters are defined as undesirable species. These undesirable species often pose high risks to power lines and therefore should be identified and removed. The image data collected multi-spectral imagery with 15 cm spatial resolution.

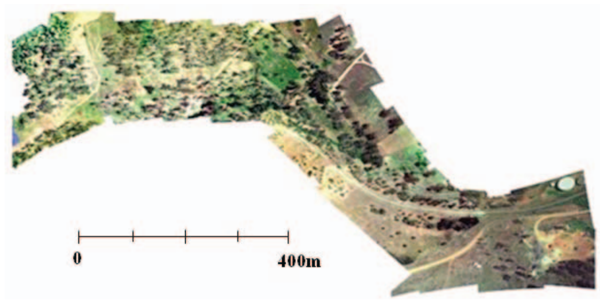

Figure 1 Experiment Test Site

The experiment data include high resolution multi-spectral imagery as well as ground truth data. The ground truth data of vegetation species were obtained from a field survey with domain experts' participation. Figure 1 shows a mosaic of the test area generated from aerial images acquired from the trial. It should be noted that classifying all types of species in power line corridors requires significantly more resources than are currently available, however, classifying species in a given test area as a proof of concept is possible. In this research, we focus on three dominant species in our test field: Eucalyptus tereticornis, Eucalyptus melanophloia, and Corymbia tesselaris. We abbreviate the species names to Euc_Ter, Euc_Mel and Cor_Tes. Through field survey with botanist's participation, 121 trees were selected and labeled for the experiment with 64 Euc_Ter, 30 Euc_Mel and 27 Cor_Tes. The criterion is that tree crowns are big enough so that they can be visually identified from the aerial images.

B. Experiment Setup and Results

To evaluate the performance of the proposed feature fusion method, the overall classification accuracy of the fused color-texture feature are compared with single feature vectors and serial integrated feature vector through the same classifier. A support vector machine (SVM), which is well developed to handle classification problems, is employed in this research as the benchmark classifier. Radial basis function (RBF) is used in the kernel function of the SVM algorithm which is suggested by many previous research. V-fold cross validation technique is employed in the experiment, and 10 folders were selected for the cross validation. The dataset is partitioned into 10 groups, which is done using stratification methods so that the distributions of categories of the target variable are approximately the same in the partitioned groups. 9 of the 10 partitions are collected into a pseudo-learning dataset and A SVM model is built using this pseudo-learning dataset. The rest 10% (1 out of 10 partitions) of the data that was held back and used for testing the built model and the classification error for that data is computed. After that, a different set of 9 partitions is collected for training and the rest 10% is used for testing. This process is repeated 10 times, so that every row has been used for both training and testing. The classification accuracies of the 10 testing datasets are averaged to obtain the overall classification accuracy.

From the experiment, the overall classification accuracies of color histogram and ULBP texture features are 80.17% and 71.07% respectively. The serial integration of these two features shows better performance over single feature with an overall accuracy of 83.47%. To evaluate of performance of fused feature using kernel PCA, we use a step-by-step model justification method. We justify the dimensionality from 2 to 8 with step 2, and from 10 to 100 with step 5 for the fused feature vectors. Figure 2 shows the classification accuracy curve at different dimensions. As we can see from the figure, the kernel PCA fused feature performs much better than single feature and serial integrated feature. However, it is still based on the assumption that user can specify a good target dimensionality. The estimation of intrinsic dimensionality using MLE is employed as the automatic selection of optimal number of dimensions. From our experiment, the intrinsic dimensionality of the integrated ULBP and RGB histogram feature is 39.3016, which conforms to the result from Figure 2 where the best accuracy (95.04%) is obtained in dimension 40.

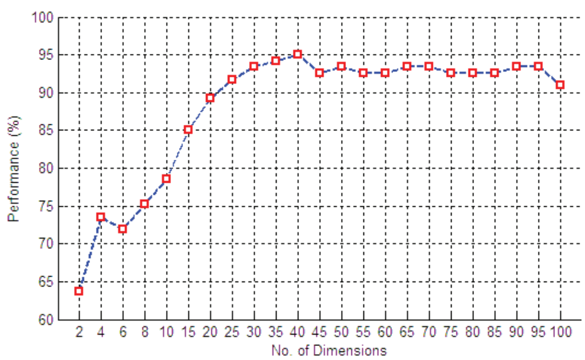

Figure 2 Classification accuracy at different dimensions

In the experiment, the computational costs of the classifiers using different feature vectors are also compared. The analysis time is recorded under a desktop PC configuration of core duo 2.66GHz CUP and 2GB memory. Table 1 summarizes the overall accuracies and analysis time using different feature set, in which KPCA_40 is the optimized fused feature at dimension 40. From the results, we can see that high dimensionality of the original color and texture features and the serial fused feature cause high computational costs, while the using the nonlinear fusion method by kernel PCA can not only improve the classification accuracy but also significantly reduce the computational costs.

Table 1 The overall accuracies and analysis time of single and fused features

Features	Histogram	ULBP	Serial Fusion	KPCA_40
Overall Accuracy	80.17%	71.07%	83.47%	95.04%
Analysis Time	70.32 s	246.29 s	305.83 s	13.63 s

5. CONCLUSION

This paper presents a novel object-based color and texture feature fusion method based on kernel PCA. The method has been evaluated in an application of vegetation classification using aerial imagery. From the experimental results, fusing color and texture features provide improved discrimination over using them independently. Moreover, the proposed nonlinear feature fusion strategy has shown great improvement over the serial fusion strategy, not only on reducing the dimensionality and computational cost, but also on removing noisy information and improving the discriminative power. The proposed object-based feature fusion strategy can be extended for fusing other feature descriptors and some other nonlinear feature selection method such as Local Linear Embedding (LLE) may also be used to substitute the kernel PCA method.

6. REFERENCES

[1] Q. Yu, P. Gong, N. Clinton, G. B. M. Kelly, and D. Schirokauer, "Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery," *Photogrammetric Engineering & Remote Sensing,* vol. 72, pp. 799-811, 2006.

[2] T. Blaschke, "Object-based image analysis for remote sensing," *ISPRS Journal of Photogrammetry & Remote Sensing,* vol. 65, pp. 2-16, 2010.

[3] P. F. Whelan and O. Ghita, "Color texture analysis," in *Handbook of texture analysis*, M. Mirmehdi, X. Xie, and J. Suri, Eds.: Imperial College Press, 2009, pp. 129-164.

[4] X. Huang, L. Zhang, and P. Li, "A multiscale feature fusion approach for classification of very high resolution satellite imagery based on wavelet transform," *International Journal of Remote Sensing,* vol. 29, pp. 5923-5941, 2008.

[5] L. Zhang, Y. Zhao, B. Huang, and P. Li, "Texture feature fusion with neighborhood oscillating tabu search for high resolution image," *Photogrammetric Engineering & Remote Sensing,* vol. 74, pp. 1585-1596, 2008.

[6] Z. Li, R. Hayward, J. Zhang, Y. Liu, and R. Walker, "Towards automatic tree crown detection and delineation in spectral feature space using PCNN and morphological reconstruction," in *IEEE International Conference on Image Processing*, 2009.

[7] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, "Study on texture feature extraction in region-based image retrieval system," in *International Conference on Multimedia Modelling*, 2006.

[8] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution grey-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 24, pp. 971-987, 2002.

[9] J. Yang, J.-y. Yang, D. Zhang, and J.-f. Lu, "Feature fusion: parallel strategy vs. serial strategy," *Pattern Recognition,* vol. 36, pp. 1369-1381, 2003.

[10] X. Xie and K.-M. Lam, "Gabor-based kernel PCA with doubly nonlinear mapping for face recognition with a single face image," *IEEE Transactions on Image Processing,* vol. 15, 2006.

[11] E. Levina and P. J. Bickel, "Maximum likelihood estimation of intrinsic dimension," in *Advances in Neural Information Processing Systems*. vol. 17: The MIT Press, 2004.