



Queensland University of Technology
Brisbane Australia

This is the author's version of a work that was submitted/accepted for publication in the following source:

Li, Zhengrong, Hayward, Ross F., Walker, Rodney A., & Liu, Yuee (2011)
A biologically inspired object spectral-texture descriptor and its application to vegetation classification in power-line corridors. *IEEE Geoscience and Remote Sensing Letters*, 8(4), pp. 631-635.

This file was downloaded from: <http://eprints.qut.edu.au/41278/>

© Copyright 2011 IEEE

Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Notice: *Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source:*

<http://dx.doi.org/10.1109/LGRS.2010.2098391>

A Biologically Inspired Object Spectral-Texture Descriptor and Its Application to Vegetation Classification in Power Line Corridors

Zhengrong Li, Ross Hayward, Rodney Walker, and Yuee Liu

Abstract—The use of appropriate features to represent an output class or object is critical for all classification problems. In this paper, we propose a biologically inspired object descriptor to represent the spectral-texture patterns of image-objects. The proposed feature descriptor is generated from the pulse spectral frequencies (PSF) of a pulse coupled neural network (PCNN), which is invariant to rotation, translation and small scale changes. The proposed method is first evaluated in a rotation and scale invariant texture classification using USC-SIPI texture database. It is further evaluated in an application of vegetation species classification in power line corridor monitoring using airborne multi-spectral aerial imagery. The results from the two experiments demonstrate that the PSF feature is effective to represent spectral-texture patterns of objects and it shows better results than classic color histogram and texture features.

Index Terms—feature descriptor, pulse spectral frequency, PCNN, rotation and scale invariance, spectral-texture analysis, vegetation species classification

I. INTRODUCTION

Trees, shrubs and other vegetation are of continued importance to the environment and our daily life. However, vegetation touching power lines is a risk to public safety and the environment, and one of the main causes of power supply problems. Aerial remote sensing techniques have great potential in assisting vegetation management in power line corridors [1]. In power line corridors, some tree species are of particular interest and are generally categorized as undesirable species and desirable species. The undesirable species with fast growth rates that also have the potential to reach a certain mature height which can pose risks to the infrastructure should be identified and removed.

The use of appropriate features to characterize an output class or object is fundamental for all classification problems. There is no generically best feature for image classification. The selection of an appropriate feature descriptor must reflect a specific classification task in hand and usually need to be obtained through experimental evaluation. Tree crowns often present different shapes when viewed from different directions

and positions (Fig. 1). Due to the perspective view and the spatial resolution of aerial photographs, the sizes and shapes of tree crowns can look quite different. This motivates the use of appropriate features to represent image structures which are invariant to rotation and scale changes.

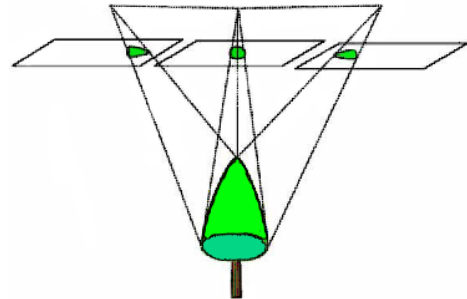


Fig. 1. Tree crown shapes from triple views.

Rotation invariant features have been investigated in image texture classification for a long time, with many of them generated from filtered images or by converting rotation variant features to rotation invariant features using a circular neighbor set [2]. The human eye is remarkable in its ability to interpret color-textured objects and there are a number of models developed for image feature analysis based on biological models of the visual cortex [3]. To the best of our knowledge, there has been little research to validate the capabilities of biologically inspired feature extraction mechanisms in remote sensing image classification problems. In this context, we present a biologically inspired spectral-texture feature descriptor by utilizing pulse spectral frequency (PSF) of a pulse coupled neural network (PCNN) model. This feature is capable of capturing the local structure of image and is invariant to rotation, translation and scale changes. The proposed method is evaluated against several classic color and texture descriptors in standard texture classification using USC-SIPI texture database and also an application of vegetation species classification using airborne multi-spectral images.

II. SPECTRAL-TEXTURE REPRESENTATION USING THE PULSE SPECTRAL FREQUENCY OF A PCNN

A. Pulse Coupled Neural Networks

The pulse coupled neural network (PCNN) is a relatively new biologically inspired spiking neural network model based on the understanding of visual cortical models of small mammals. Most PCNNs are based on the Eckhorn model [4] sharing a common mathematical foundation but with variations

Manuscript received xx xx, xx. This work was conducted within the CRC for Spatial Information, and with the collaboration between the Queensland University of Technology (QUT), the Australian Research Centre for Aerospace Automation (ARCAA) and Ergon Energy Australia.

The authors are with the Queensland University of Technology (QUT), Brisbane, QLD, 4001, Australia (e-mail: zhengrong.li@qut.edu.au; ra.walker@qut.edu.au; yuee.liu@qut.edu.au).

each having their own unique terms [5]. PCNNs are spatial-temporal-coding models which attract much attention from researchers in that they mimic real neurons better and have more powerful computation performance than traditional neural network models due to the use of time.

In this paper, a simplified spiking cortical model called unit-linking PCNN [6] is employed to generate image features. We introduce multi-spectral channels in this PCNN model. Compared with original unit-linking PCNN, the advantage of this model is that it has more external inputs so that both spectral and spatial information are considered in the derived features. Fig. 2 illustrates the structure of this PCNN model. Each neuron corresponds to one pixel in an input image.

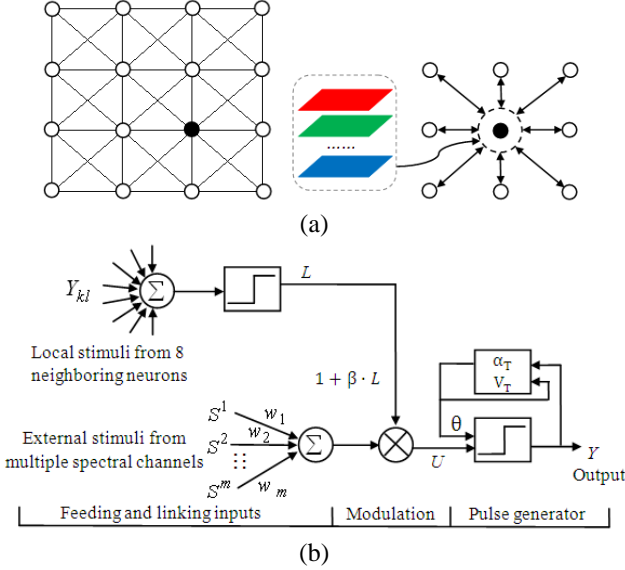


Fig. 2. The structure of the PCNN model. (a) Each neuron is coupled with its 3×3 neighboring neurons, receiving local stimuli (i.e. pulse outputs) from its neighboring neurons and also the external stimuli from the corresponding pixel values. (b) local stimuli and external stimuli are modulated and input to the pulse generator. The neuron pulses if the modulated input is larger than a dynamic threshold.

The model can be mathematically represented as:

$$F_{i,j}^m(t) = S_{i,j}^m(t) \quad (1)$$

$$L_{i,j}(t) = \begin{cases} 1, & \text{if } \sum_{k,l \in N(i,j)} Y_{k,l} > 0 \\ 0, & \text{else} \end{cases} \quad (2)$$

$$U_{i,j}(t) = (1 + \beta L_{i,j}(t)) \sum_{m=1}^M (w_m F_{i,j}^m(t)) \quad (3)$$

$$Y_{i,j}(t) = \begin{cases} 1, & \text{if } U_{i,j}(t) > \theta_{i,j}(t) \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

$$\text{delta}\theta_{i,j}(t) = \frac{d\theta_{i,j}(t)}{dt} = -\alpha_{i,j}^T + V_{i,j}^T Y_{i,j}(t-1) \quad (5)$$

where t refers to time (the number of iterations); (i, j) indicates the index of the current neuron (i.e. pixel (i, j)) and (k, l) indicates the neighboring field of the neuron (i.e. 3×3 window)); m indicates that the external input from m^{th} channel of the image; $Y_{i,j}$ is the pulse output of the neighboring neurons; $U_{i,j}$

is the internal activity; $V_{i,j}^T$ is the threshold magnitude scale (greater than 1); $\theta_{i,j}$ is the dynamic threshold which controls whether the neuron pulse or not. In this paper, the firing threshold is linear-depending threshold. $\alpha_{i,j}^T$ is the maximum value of the input image; The linking strength coefficient β determines the weight of linking input to the internal status of the neuron. The weight factor w_m is the importance of m^{th} spectral channel (M is the total number of channels, $\sum_{m=1}^M w_m = 1$).

B. Properties of Behaviors of PCNN

Unlike most other neural network models, the processing is automatic and there is no training required in a PCNN. The PCNN algorithm consists of iteratively computing until some stopping criterion is reached. Through iterative computation, neurons produce a temporal series of pulse outputs, which indicates the pulse status of each neuron (pixel). At each iteration, different neurons fire sequentially according to the internal status of neurons and the firing threshold. Similarities in the input pixels cause the associated neurons to pulse synchronously, thus indicating similar structures.

The dynamic properties of PCNN are very complex. In this section, the property and behavior of a single neuron is analyzed under the assumption that there are no linking connections. Biologically there is a fatigue period called the refractory period after a neuron fires. A neuron cannot be captured by other neurons if it is in the refractory period. Fig. 3, obtained from equation (3)-(5), illustrates the periodical pulse of a neuron (i, j) , and its capture and refractory period. $U_{i,jmax}$ is the possible maximum value of the internal activities, $T_{Ri,j}$ is the refractory period and $T_{Ci,j}$ is the capture period, $F_{i,j}$ is the combination of the external stimulus $F_{i,j} = \sum_{m=1}^M (w_m \times F_{i,j}^m(t))$.

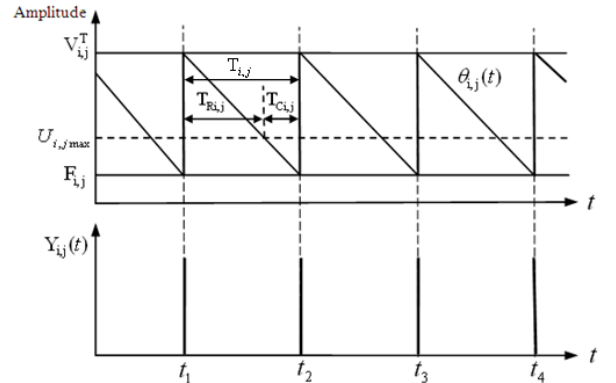


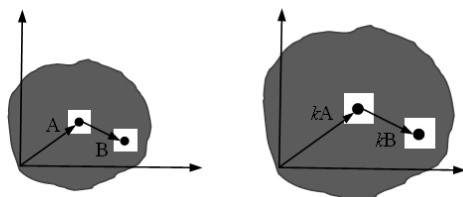
Fig. 3. Periodical pulse of the neuron (i, j)

When a neuron (i, j) fires, its threshold increases to $V_{i,j}^T$, and then linearly decreases to $U_{i,jmax}$ to make the neuron fire again. The pulse process continues periodically with the time $T_{i,j}$. During the refractory period $T_{Ri,j}$, the neuron cannot fire no matter whether its neighbors fire or not because its threshold $\theta_{i,j}$ is larger than the maximum internal activity $U_{i,jmax}$. Only during the capture time $T_{Ci,j}$, the neuron can pulse or be excited by other neurons as the threshold is lower than the maximum internal activity $U_{i,jmax}$. There is a period changing from the refractory time to the capture time. Therefore, PCNN mimics the mechanism of the biological neuron.

C. Rotational and Scale Invariant Feature Extraction Using Pulse Spectral Frequency

The spectral time sequence generated from PCNN iterations reflects the image value distribution pattern and thus it can be used to represent image features. Lindblad and Kinser found that PCNN and wavelet transforms have many similarities, however PCNNs are unique in that they generate rotation, translation and scale invariant time signals [7]. This invariance property is identified by observing the same period and number of peaks in the output time signals of a PCNN. The invariance property is not hard to understand if we consider PCNN as an image representation that summarizes the firing neurons in the whole image (region). Since an isotropic neighborhood is adopted, it does not matter which neuron is used to provide the local stimulus. The output of PCNN is not inherently invariant to scale changes because the number neurons affected by the rescaled patch is changed. However, the output pulse frequency of the rescaled patch is stable and the scale changes will only be reflected in the outputs of PCNN by a scale-factor. Here we borrow the analysis method from literature [8] to explain the scale invariance properties of our PCNN model.

For simplicity, we consider image patches rather than single pixels. Assuming one image consists of a certain number of patches and the patch number is independent to scale changes. Each image patch is considered as a whole with its own intensity. The number of pixels in each patch depends on the scale factor. As shown in Fig. 4, after an image is rescaled, the distances of image patches change but the intensity per image patch is constant. When a neuron at patch A receives a linking contribution from a neuron at patch B , the image patch at A goes to kA and B goes to kB after the image is rescaled. However, the feeding input of the neuron does not change (e.g. $F(A) = F(kA)$). Moreover, in the unit-linking PCNN model, the linking input is unified to 0 or 1 and thus it does not depend on the scale factor dependence k . Therefore, the internal activity of the rescaled patch remains the same as that of the original patch and thus the pulse dynamics of each image patch will not change. The only change is the actual number of pixels in each image patch. As a result, if we can normalize this difference PCNN will be invariant to scale changes.



(a) original image

(b) rescaled image

Fig. 4. Geometry for scale invariance

In this paper, we use the pulse spectral frequency (PSF) for rotation and scale invariant object feature extraction. PSF is defined as a normalized histogram which indicates the number of firing pixels in a specified time period (i.e. equation (6)).

$$PSF(t) = N(t)/\max(N) \quad (6)$$

where $N(t)$ is the number of firing pixels at time t and $\max(N)$ is the maximum number of firing pixels in a time period. In order to achieve scale invariance we normalize the number of firing pixels to $[0, 1]$ in the discrete time steps. The dimension of the histogram is equal to the total number of iterations.

However, as stated by Johnson in [8], this scale invariance property may not hold for very small image and large scale changes because the local group around a neuron also changes in scale which cause the internal activity change as well.

The calculation of PSF feature consists of iteratively computing until the user decides to stop. There is currently no automated stop mechanism built. Theoretically, the more the iteration time, the richer information can be derived to characterize the image or object. However, at the meantime, it will cause the high dimensionality of the feature vectors. In our experiments, we use 21 iterations for all PSF features.

III. EXPERIMENTS AND RESULTS

To evaluate the proposed object feature descriptor discussed above, two experiments were conducted. The first experiment assesses PSF features in rotation and scale invariant texture classification. The second experiment evaluates the PSF feature in a real-world application of vegetation species classification in power line corridors. In this section, the details of the two experiments and results are presented.

The benchmark classifiers we used in the experiments include: multilayer perceptron neural networks (MLP), decision tree forest (DTF), and support vector machines (SVM). In this paper, we use the implementation of DTREG for the three classifiers [9]. The same default parameter settings are used in all classification tests. 10-fold cross validation technique is employed in the classification experiment. The classification accuracies of the 10 testing datasets are averaged to obtain the overall classification accuracy.

A. Rotation and Scale Invariant Texture Classification

A dataset of rotated texture images from the USC-SIPI texture database¹ is used in the first experiment. The dataset consists of the 13 Brodatz textures digitized at seven different rotation angles: 0, 30, 60, 90, 120, 150, and 200 degrees (91 images). To test the scale invariance properties, we resize each image to 4 scales (0.25, 0.5, 0.75, and 1). A total of 364 images are used in the experiment with 13 textures (i.e. wool, bark, brick, bubbles, grass, leather, pigskin, raffia, sand, straw, water, weave, and wood).

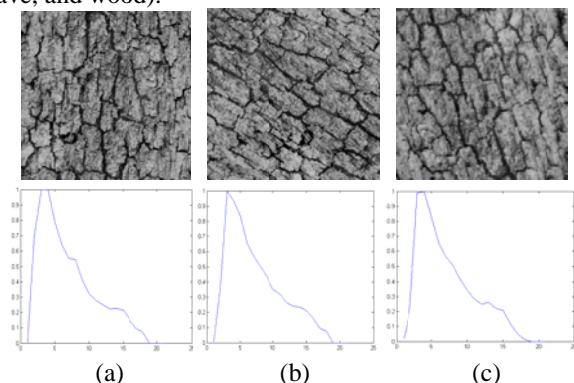


Fig. 5. Examples of bark texture images and their PSF features: (a) scale=1, rotation=0° (b) scale=1, rotation=60° (c) scale=0.5, rotation=200°

PSF features are extracted from the gray-scale texture images and then used as input to build the classifier. Fig. 5

¹ USC-SIPI Database: <http://sipi.usc.edu/database/database.cgi?volume=textures>

shows some examples of bark texture images and their corresponding PSF features at different scales and rotation angles. As we can see, the PSF histograms are not exact the same for the three texture images, but the changing trends of PSF histogram in a time period is approximately the same.

A well-known texture descriptor, local binary patterns (LBP), is also evaluated for comparison purposes. In this experiment, we use the rotation invariant LBP proposed by Ojala et al. [2] and a SVM classifier is employed in the classification test. Table 2 compares the performance of LBP and PSF when textures are rotated only and with both rotation and scale changes. From the results we can see LBP performs slightly better than PSF when images are with rotation only. However, PSF generated much higher classification accuracy when the images have both rotation and scale changes. Table 2 compares the average computational costs of PSF and LBP per image. The experiment is conducted under a desktop PC configuration of core duo 2.66GHz CUP and 2GB memory. Since PSF involves iterative computation of the neural network, the computational cost is much higher than LBP.

Table1 Classification accuracies of PSF and LBP (%)

	Rotation Only	Rotation and Scale
PSF	98.9	99.18
LBP	100	94.51

Table2 Average computational costs of PSF and LBP (s)

	PSF	LBP
computational cost	0.937	0.136

B. Vegetation Species Classification in Power Line Corridor

The experiment data used in the second experiment were captured in a 1.5 kilometres corridor in rural Queensland Australia by a high resolution digital 4-band multi-spectral camera (DuncanTech MS-4100) with a DGPS/INS mounted in the cargo area of a Piper Cub [1]. The 4 spectral bands of the camera are: NIR (800-966nm), red (670-840nm), green (540-640nm), blue (460-545nm). The spatial resolution of the captured images is about 15 cm. It should be noted that classifying all types of species in power line corridors requires significantly more resources than are currently available. In this research, we focus on three dominant species in our test field. We abbreviate the species names to *Euc-Ter*, *Euc-Mel* and *Cor-Tes*. Through a field survey with a botanist's participation, 121 trees were selected and labeled for the experiment with 64 *Euc-Ter*, 30 *Euc-Mel* and 27 *Cor-Tes* trees.

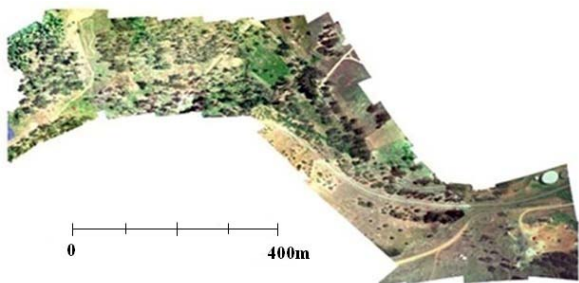


Fig. 6. Experiment test site

Object-based classification is used in this paper. Individual tree crowns are firstly segmented from image and local features are extracted from the crown regions and after that the classification is conducted in object-feature space. Since the

main aim of this research is to evaluate the effectiveness of object feature descriptors, we assume that the segmentation is perfect and thus individual tree crowns are manually segmented from the images during the field survey. For comparison purposes, we also evaluated some classic color and texture feature descriptors, which include GLCM, Gabor filters, LBP and color histogram features extracted from 4 spectral bands and also HSV color space. It is also worth mentioning that plants have distinctive spectral signature which is often modeled by combinations of reflectance measured in two or more spectral bands [10]. This motivates us to investigate whether spectral-texture features extracted from spectral vegetation indices could help in vegetation species classification. In the experiment, three widely used vegetation index maps are employed: the Normalized Difference Vegetation Index (NDVI), the Soil Adjusted Vegetation Index (SAVI), and the 2-band Enhanced Vegetation Index (EVI2). PSF histogram features are generated from both the original spectral bands and the three vegetation index maps.

Table 3 compares the overall classification accuracies of the PSF feature and three classic texture descriptors using different machine classifiers. The results clearly show that the selection of both feature descriptors and classifiers will strongly influence the classification accuracies. Nevertheless, the PSF feature obtains the best overall classification accuracy on all three benchmark classifiers, which confirm its use as an effective feature descriptor for this data. We also evaluated the performance of PSF features extracted from multiple spectral bands. Table 3 summarizes the overall classification accuracies of these features. *Hist_RGBNIR* and *Hist_HSV* refer to the color histograms extracted from four spectral bands (R, G, B and NIR) and HSV color space; similar names are used for PSF features extracted from four spectral bands and also HSV color space; *PSF_HSV_VI* represents the PSF feature extracted from both HSV color space and three vegetation index maps. From the results, we can see that PSF features show significant improvement over color histograms. While the color histograms characterize the color distribution of the pattern, they do not exploit the spatial layout of the colors. It is also noted that PSF-HSV outperforms the PSF feature calculated from original spectral bands. Another interesting result is that when we incorporate the spectral vegetation index into the PSF-HSV feature, a significant improvement is achieved.

It is worth noting that overall accuracy does not distinguish between the types of errors the classifier makes (i.e. False Positive versus False Negative). Receiver Operating Characteristic (ROC) analysis is a more comprehensive performance measure and provides more detail of the classification result. ROC analysis plots the False Positive Rate (FPR) on the x-axis of a graph and True Positive Rate (TPR) on the y-axis. A ROC graph depicts relative trade-offs between true positive (benefits) and false positive (costs), and the goal in ROC space is to be in the upper-left-hand corner [11]. Fig. 7 presents the analysis results of different feature descriptors using a SVM classifier in ROC space. As we can see, generally most features get better performance for class 'Cor-Tes' than the other two classes. PSF-HSV-VI performs the best for classes 'Euc-Ter' and 'Cor-Tes' and PSF_HSV performs the best for 'Euc_Mol'. Overall the PSF features outperform other color and texture descriptors for all three classes.

We attribute the success of PSF feature to its capability of capturing the local structure of image and its unique property of rotation and scale invariance. These properties make PSF especially useful in object classification from aerial images because the same object type has different shapes when viewed from different heights and directions. Moreover, PSF can be easily extended to represent the spectral-texture patterns by integrating the PSF histograms extracted from the pulse images of multiple spectral bands. However, a critical issue for the use of PSF feature is the adaptive parameter setting of the PCNN

model, which is also a hard problem for all PCNN related research. Optimal feature dimension selection is also an interesting future work.

Table 2 Overall classification accuracies of PSF and texture measures (%)

	GLCM	Gabor	LBP	PSF
MLP	69.42	71.9	66.94	70.25
DTF	56.2	71.07	71.07	77.69
SVM	69.42	69.42	77.69	77.69

Table 3 Overall Classification accuracies of color histogram features and PSF features in multiple spectral bands (%)

	Hist-RGBNIR	Hist-HSV	PSF-RGBNIR	PSF-HSV	PSF-HSV-VI
MLP	71.9	75.21	75.21	80.17	85.12
DTF	71.9	78.51	80.17	80.99	78.51
SVM	76.03	69.42	79.34	81.82	85.95

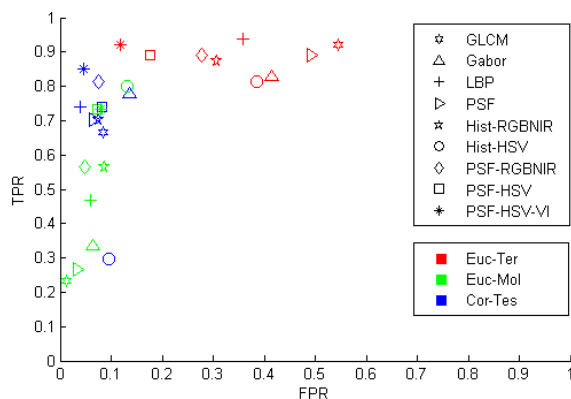


Fig. 7. Analysis of different feature descriptors in ROC space

IV. CONCLUSION

This paper presents a new object spectral-texture feature descriptor which is obtained by calculating the pulse spectral frequency of the outputs of a biological inspired PCNN model. The proposed method has been evaluated against classic color and texture descriptors in two experiments. Experimental results revealed that the PSF feature is able to capture the structure of images and is invariant to rotation and small scale changes. The quantitative evaluation results shows that the proposed feature outperforms color histograms and three classic texture measures. In addition, an interesting result is that incorporating the spectral vegetation index into the PSF features greatly improves vegetation species classification. However, the selection of optimal parameter and feature dimension is still an open problem. In addition, to achieve practical application in vegetation management, precise automatic tree crown segmentation is needed prior to the generation of PSF features.

ACKNOWLEDGMENTS

The authors would like to thank Bred Jeffers from Greening Australia and David Wood from Ergon Energy for their support and assistance in conducting the ground survey.

REFERENCES

- [1] S. J. Mills, M. P. Gerardo, Z. Li, J. Cai, R. Hayward, L. Mejias, and R. Walker, "Evaluation of aerial remote sensing techniques for vegetation management in power line corridors," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, pp. 3379-3390, 2010.
- [2] 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.
- [3] K. Zhan, H. Zhang, and Y. Ma, "New Spiking Cortical Model for Invariant Texture Retrieval and Image Processing," *IEEE Transactions on Neural Networks*, vol. 20, pp. 1980-1986, 2009.
- [4] R. Eckhorn, H. J. Reiboek, M. Arndt, and P. W. Dicke, "A neural network for feature linking via synchronous activity: Results from cat visual cortex and from simulations," in *Models of Brain Function*, R. M. J. Cotterill, Ed. Cambridge: Cambridge University Press, 1989, pp. 255-272.
- [5] Z. Wang, Y. Ma, F. Cheng, and L. Yang, "Review of pulse-coupled neural networks," *Image and Vision Computing*, 2009.
- [6] X. Gu, "Feature Extraction using Unit-linking Pulse Coupled Neural Network and its Applications," *Neural Process Letters*, vol. 27, pp. 25-41, 2008.
- [7] T. Lindblad and J. M. Kinser, "Inherent features of wavelets and pulse coupled networks," *IEEE Transactions on Neural Networks*, vol. 10, pp. 607-614, 1999.
- [8] J. L. Johnson, "Pulse-coupled neural nets: translation, rotation, scale, distortion and intensity signal invariance for images," *Applied Optics*, vol. 33, pp. 6239-6253, 1994.
- [9] P. H. Sherrod, "DTREG predictive modeling software. Users Manual.," www.dtreg.com/dtreg.pdf, 2009.
- [10] M. Rautiainen, "The spectral signature of coniferous forests: the role of stand structure and leaf area index," in *Faculty of Agriculture and Forestry, Faculty of Agriculture and Forestry Helsinki*: University of Helsinki, Finland, 2005.
- [11] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curve," in *the 23rd International Conference on Machine Learning*, Pittsburgh, 2006, pp. 233-240.