TEXTURE ANALYSIS USING GABOR FILTERS FOR IMAGE CLASSIFICATION BASED ON SVM *

Yi-Ting Tsai [1] and Jin-Tsong Hwang [2]
[1]Graduate Student, Dept. of Real Estate and built Environment, National Taipei University,
No.,151, University Rd., San Shia, New Taipei City, Taiwan; Tel: +886-2-26748189#67426;
E-mail: tsai_ting@hotmail.com

[2]Associate Professor, Dept. of Real Estate and built Environment, National Taipei University,
No.,151, University Rd., San Shia, New Taipei City, Taiwan; Tel: +886-2-26748189#67426;
E-mail: jthwang@mail.ntpu.edu.tw

KEY WORDS: SVM, Gabor Filter, Classification

ABSTRACT: The texture is an important factor in region-based segmentation of images. Texture can be seen in many images from multispectral remote sensed data to microscopic photography. Despite its importance, there is no unique and precise definition of texture. Each texture analysis method characterizes image texture in terms of the features it extracts from the image. Therefore, it depends not only on studying the images but also on the goal for which the image texture is used and the features that are extracted from the image. Gabor filters provide means for better spatial localization however their usefulness is limited in practice because there is usually no single filter resolution at which one can localize a spatial structure in natural texture. In this paper, we present a methodology based on computing a set of textural measures with Gabor filter, and then, the Gabor texture features combined with original bands of image, PCA, and NDVI, etc. were adopted as the characteristic vector of training samples for SVM, and Decision Tree classification. Finally, traditional classification schemes of Maximum Likelihood were comparatively studied.

1. INTRODUCTION

Texture is a fundamental characteristic in many natural images and also plays an important role in computer vision and pattern recognition. Texture analysis is an essential step for many image processing applications such as industrial inspection, document segmentation, remote sensing of earth resources, and medical imaging. Most approaches can be divided loosely into three categories: statistical, model-based, and signal processing methods. Statistical methods characterize an image in terms of numerical features, which are derived from the Fourier power spectrum, gray level run length, and co-occurrence matrices (He, D. C. and Wang, 1992). Model-based methods such as the Markov random field (Cohen, F. S., et al., 1991) and the simultaneous autoregressive models (Mao, J. and A. K. Jain, 1992) use the model parameters as features in texture classification or segmentation. Gabor filters provide means for better spatial localization however their usefulness is limited in practice because there is usually no single filter resolution at which one can localize a spatial structure in natural texture. There are many of different definitions of texture in (Vyas, V. S. and Priti R., 2006). The time- frequency transformed based method of texture discrimination, which is in turn based on Gabor filters is done.

The availability of images acquired by these very high spatial resolution sensors leads to a new set of possible applications, which require mapping the Earth surface both with great geometrical precision and a high level of thematic detail. In this context, great attention is devoted to the analysis of urban scenes, with applications such as road network extraction and road map updating, transportation infrastructure management, the monitoring of growth in urban areas, and discovering building abuse (Volpe, F. and L. Rossi, 2003). The pixel-based system is aimed at obtaining accurate and reliable maps both by preserving the geometrical details in the images and by properly considering the spatial context information (Bruzzone, L. and Lorenzo C., 2006).

A SVM is a new machine learning technique developed on the basis of statistical learning theory, and it is the most successful realization of statistical learning theory. In this paper, our choice of feature extraction includes Principal Component Analysis and Gabor filter process, respectively. In our work, Matlab was used for Gabor filter process; Weka 3.6.1 was used for decision tree classification; LibSVM 2.9 was used for SVM classification.

2. FEATURE EXTRACTION

2.1 Principal Component Analysis (PCA)

* This work was supported in part by the National Science Council of Taiwan under Grant (NSC-99-2410-H-305-075）

The PCA employs the statistic properties of spectral bands to examine band dependency or correlation. This kind of transformations is based on the same mathematical principle known as Eigen value decomposition of the covariance matrix of the multispectral image bands to be analyzed. The aspect of PCA analysis that can be seen in this illustration pertains to the variability within bands. Once the transformation has taken place, PCA band 1accounts for the maximum amount of variability or contrast possible in the image and PCA band 2 accounts for the second largest amount. This trend is likely to continue in the first few PCA bands, with the remainder containing less and less useful information. In this study, the first PCA bands with 97.4% of contained accumulation entire information is selected, and then this file is prepared for classification. And then, the first PCA file is transformed by Gabor filter for textures extraction

2.2 Gabor Filter

A Gabor filter is obtained by modulating a sinusoid with a Gaussian. For the case of one dimensional signal, a 1D sinusoid is modulated with a Gaussian. This filter will therefore respond to some frequency, but only in a localized part of the signal. For 2D signals such as images, consider the sinusoid. By combining this with a Gaussian, we obtain a Gabor filter.

Gabor filter is a linear filter whose impulse response is defined by a harmonic function multiplied by a Gaussian function. It is optimally localized as per the uncertainty principle in both the spatial and frequency domain. This implies Gabor filters can be highly selective in both position and frequency, thus resulting in sharper texture boundary detection. Gabor filter related segmentation paradigm is based on filter bank model in which several filters are applied simultaneously to an input image .The filters focus on particular range of frequencies .If an input image contains two different texture areas, the local frequency differences between the areas will detect the textures in one or more filter output sub-images .

Gabor filters can be configured to have various shapes, bandwidths, center frequencies and orientations by the adjustment of suitable parameters. By varying these parameters a filter can be made to pass any elliptical region of spatial frequencies.

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{\overline{x^2}}{\sigma_x}+\frac{\overline{y^2}}{\sigma_y}\right)\right]\cos\left[2\pi j\frac{\bar{x}}{\lambda}\right] \tag{1}$$

$$\bar{x} = x\cos\theta + y\sin\theta\,;\quad \bar{y} = -x\sin\theta + y\cos\theta$$

Where σx and σy characterize the spatial extent and bandwidth of the filter and determine the effective size of the neighborhood of a pixel in which the weighted summation takes place. θ specifies the orientation of the Gabor filters. λ is the wavelength of which is the wavelength of the cosine factor $\cos(2\pi\bar{x}/\lambda)$, determines the preferred spatial frequency 1/λ of the receptive field function g(x,y). A filter will respond stronger to an edge with a normal parallel to the orientation θ of the sinusoid. The Gauss window reflects the location of the Gabor filter both in the time and frequency domain, and limits the range of the oscillation function. Gabor filter can tolerate image slight distortion by using the Gauss window. The Fourier transform of the Gabor function in (1) is given by

$$G(u,v) = \exp\left[-\frac{1}{2}\left(\frac{(u-F)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right)\right] \tag{2}$$

Where σu = 1/(2πσx), σv = 1/(2πσy). In most cases, letting σx =σy = σ is a reasonable design choice. The center frequency of the Gabor function is defined by (u,v). The radial center frequency is defined as $F = \sqrt{u^2 + v^2}$ and the orientation as $\theta = \tan^{-1}(v/u)$. The Gabor function is most interesting when studied in the frequency domain. It is then a Gaussian function shifted in frequency to position (u,v) i.e. at a distance F from the origin in the orientation. In 1946, Dennis Gabor proposed the expansion of a wave in terms of Gaussian wave packets. An example of such a wave packet is a sine wave multiplied by a Gaussian function. If a signal is modulated by a Gaussian window of a certain width and central time, then a Fourier expansion of the modulated signal gives a measure of the local spectrum. Clearly such a spectrum is not unique since the width of the Gaussian is arbitrary; but nevertheless, such local spectral are extremely useful. If a collection of local spectral is computed for a suite of window positions, the result is a time-frequency decomposition called a Gabor transformation (Manjunath B. S. and W. Y. Ma, 1996). Given a textured image I(x, y) consisting of known textures A and B, find the Gabor filter that best discriminates A and B in the output O(x, y). A properly designed Gabor filter can produce an output image O(x, y) exhibiting some types of discontinuity at the texture boundaries. When the two textures differ from each other, then one designs a Gabor filter that produces a step change in O(x, y) at the texture boundaries.

2.3 Parameter Selection of Gabor Filter

The standard deviation of the Gaussian factor determines the size of the receptive field. λ is the wavelength of which is the wavelength of the cosine factor $\cos(2\pi\bar{x}/\lambda)$, determines the preferred spatial frequency 1/λ of the receptive field function g(x,y). DeValois et al. (DeValois, R. L., et al., 1982) propose that the input to higher processing stages is provided by the more narrowly tuned simple cells with half-response spatial frequency bandwidth of approximately one octave. This value of the half-response spatial frequency bandwidth corresponds to the value 0.56 of the ratio, which is used in the simulations of this study. Since λ and σ are not independent (σ/λ=0.56), only one of them is considered as a free parameter which is used to index a receptive field function. In our work, we set the value of σ/λ as 0.5. For ease of reference to the spatial frequency properties of the images, we choose λ based on the semivariogram estimation on the area of padding field and tree at this study area, shown as Fig. 1(b) and (c). The λ was chooses as the distance in which the difference of the semivariogram from the sill becomes negligible. In this paper, the distance is 6 pixels for padding field, and 9 pixels for tree area, respectively. Fig. 1(d) illustrates the application of the filter bank on an input image which contains texture of tree and padding field, shown as Fig. 1(a). It seems that column of 3 (with λ equals 9) gives better results to distinguish these two classes than the others. Therefore, we choose σ equals 4.5, and λ equals 9 for full image Gabor Filter transformation, respectively.

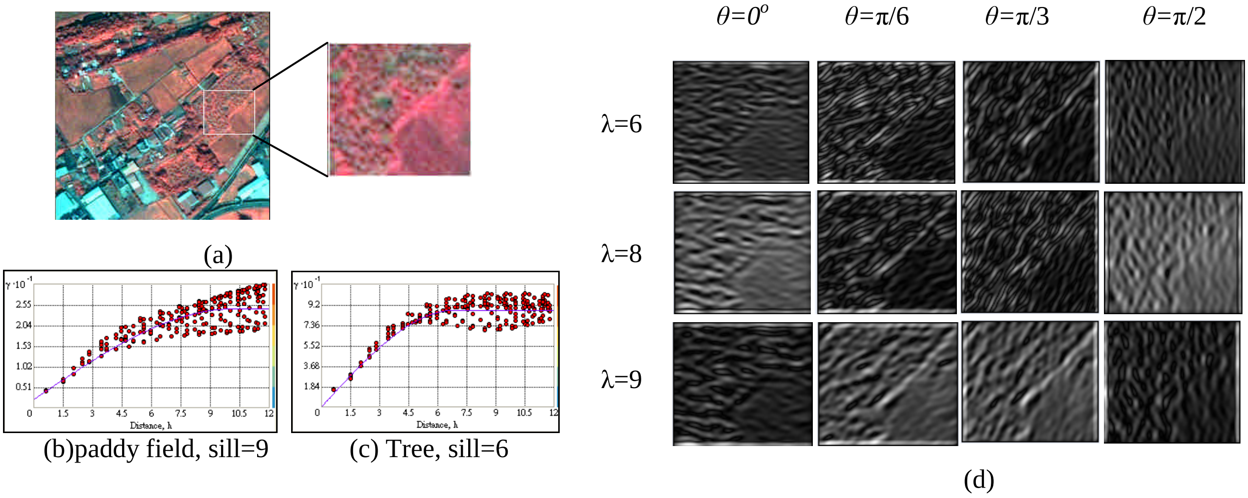

Figure 1 The input image is shown as (a), semivariogram of paddy field and tree shown as (b), and (c), respectively. (d) shows the images arranged in an 3 x4 matrix correspond to the outputs of the different channels of the filterbank. The rows correspond to different preferred wavelengths (6, 8,and 9), and the columns to different preferred orientations.

Gabor filter form a complete but non-orthogonal basis set and any given function f(x, y) can be extended in terms of these basis functions. Each filter is fully determined by choosing the four parameters in θ, f, σx, and σy. For feature extraction, we use 1/9 for f, 4.5 for σx and σy, respectively. The window size of filter is set to 11x11 with at orientations (0, π/6, π/3, π/2, 2π/3, and 5π/6).

3. METHODLOGY

3.1 Support Vector Machines (SVM)

SVM is a statistic classification method proposed by (Cortes, C. and V. Vapnik, 1995). It is originally designed for binary classification. Based on structural minimization risk principle from computational learning theory, it tries to find the separating hyper-plane with maximum margin to separate positive and negative samples from the training set. The simplest SVM is a binary classifier, which was mapping to a class and just can identify an input image data belongs to the class or not. To produce a SVM and the corresponding class c, the SVM must be given a set of training samples including positive and negative samples. Positive samples belong to c and negative samples do not. After image preprocessing, all samples can be translated to n-dimensional vectors. SVM tries to find a separating hyper-plane with maximum margin to separate the positive and negative examples from the training samples. The standard SVM is a linear inductive learning classifier where data in input space are separated by the hyperplane:

$$f(x) = w^T x + b \quad (3)$$

with maximal geometric margin $2/\|w\|^2$, where w is a vector, normal to the hyperplane and $|b|/\|w\|^2$ is the perpendicular distance from the hyperplane to the origin. The objective of the learning phase of standard SVM is to maximize the geometrical margins between classes in the feature space.

There are two kinds of multi-class SVM system, one-against-all and one-against-one. The one-against-all SVM must train k binary SVM where k is the number of classes. The ith SVM is trained with all samples of ith class as positive samples, and takes all other examples to be negative samples. After setting up all SVM with positive and negative samples, it trains all k SVM. Then it can get k decision functions. For a testing data, all the decision values are computed by all decision functions and choose the maximum value and the corresponding class to be its resulting class. The one-against-one SVM is that for every combination of two classes i and j, it must train a corresponding SVMij. Therefore, it will train $k(k - 1)/2$ SVM and get $k(k - 1)/2$ decision functions. For an input data, all the decision values are computed and use a voting strategy to decide which class it belongs to. If sign ($w_{ij} \cdot x+b_{ij}$) shown x belongs to ith class, then the vote for the ith class is added by one. Otherwise, the jth class is added by one. Finally, x is predicted to be the class with the largest vote. In this paper, a nice and efficient LIBSVM developed in C++ was used for SVM classification (Chang, C. C. and C. J. Lin, 2009).

3.2 Decision Tree (DT)

Unlike conventional statistical and neural/connectionist classifiers, which use all available features simultaneously and make a single membership decision for each pixel, the DT uses a multi-stage or sequential approach to the problem of label assignment. The labelling process is considered to be a chain of simple decisions based on the results of sequential tests rather than a single, complex decision. Sets of decision sequences form the branches of the DT, with tests being applied at the nodes. DT construction involves the recursive partitioning of a set of training data, which is split into increasingly homogeneous subsets on the basis of tests applied to one or more of the feature values. These tests are represented by nodes. The univariate DT applies a test to a single feature at a time, whereas the multivariate DT uses one or more features simultaneously. Labels are assigned to terminal (leaf) nodes by means of an allocation strategy, such as majority voting. At one time, DTs were designed manually, using spectral plots. The decision tree Classification algorithm provides an easy to understand description of the underlying distribution of the data. The objective of Classification is to build a model of the class label based on the other attributes. After a model is built, it can be used to determine the class label of unclassified dataset. In the past decade, automatic methods of decision tree design have been developed. Decision trees can be constructed relatively quickly, compared to other methods. Another advantage is that decision tree models are simple and easy to understand. In this study, DT algorithms, J48 in WEKA version 3.6.1 was used for decision tree classification.

4. EXPERIMENT

4.1 Study Area

Quick Bird images are high-resolution satellite images. The images of Earth's surface is recorded in four spectral bands:0.45-0.52μm(blue), 0.52-0.60μm(green), 0.63-0.69μm(red), and 0.76-0.90μm(near-IR), by a multispectral scanner with the spatial resolution of 2.8m. Simultaneously, a panchromatic image within the range of 0.45-0.9μm and with the spatial resolution of 0.7m, is recorded. Here, the classification was done based on the example of fragment of an image, which was recorded on March 2, 2009. The test image was taken over the San Shia, New Taipei City, Taiwan. The image size being tested is approximately 512 by 512 pixels at the resolution of 2.8m. Each pixel was digitized to 11-bit precision. There are five species comprise the canopy of this study area. They are: tree, paddy field, road, building, and water. The most of canopy are paddy field which is collected early in the growing season and field of tree. Fig. 2 shows a false color composition of this scene with training and test samples area. We generated 8,471 training samples and 11,076 testing samples referring to the ground truth.

Figure 2 Training and testing area location show on data set. The number of training and testing samples for each class are shown in parenthesis at the legend.

4.2 Data Processing

In this section, we provide experimental results using the image data set described above. The classification system described in section 3 is trained with different types of input features in supervised fashion. There 6 types of input features considered in the classification experiments conducted in this work can be summarized as follows:

(1) Original: In this case, we use the full of original spectral signatures available in the Quick Bird data as input to the proposed classification system.

(2) Original+PCA1: Here, we apply layer stacked processing with original image to the first PCA component of original image.

(3) Original+NDVI: In this case, we use NDVI combined with original spectral image.

(4) Original+Gabor Filter: We use one frequency and 8 orientations Gabor Filter for textures extraction. And then, that were combined with original spectral image.

(5) Original+NDVI+Gabor Filter: We use NDVI, and Gabor Filter for texture extraction. And then, that were combined with original spectral image.

(6) Original+PCA1+NDVI+GaborFilter: Adding PCA1 band to the 5) of input features.

5. RESULTS AND ANALYSES

In this study, the first PCA component band with 97.4% of contained accumulation entire information was selected, and then this file was prepared for classification. The first PCA file was transformed by Gabor filter for textures extraction. This transform provides a methodology for texture analysis in different orientations. In this paper, there are six orientations with Gabor filter transform.

The training samples are classified first by SVM, decision tree, and maximum likelihood. And then, testing samples are involved for accuracy assessment pixel by pixel with correspondent area on classified images. These combined feature vectors are fed into the SVM classifier initially for training from the known examples and for predicting the labels of unknown samples once the training is complete. As described above, SVM is principally a binary classifier. We used a one-against-one decomposition scheme for approaching into a multiclass classifier. It is well known that SVM generalization performance (estimation accuracy) depends on a good setting of hyperparameters C; γ and the kernel parameters. The problem of optimal parameter selection is further complicated by the fact that SVM model complexity (and hence its generalization performance) depends on all three parameters. The radial basis function (RBF) is used due to its superiority over other kernels for most of the applications. It always pays off when using optimal parameter values for the respective kernel. For handling the problem, a grid search found the parameter values $C = 2048$, $\gamma = 2$ for the RBF. These are used in all of the following experiments to perform the training.

Experimental results of dataset are given in Table 1 which is focused on features combination among original image, PCA, NDVI, and texture features extracted by Gabor filter. There are six types of combination of features in Table 1. The classification results of full image shown as Fig. 3. It seems that both of features of PCA and NDVI do not improve the classification correct rate. According to the error matrix, the class of padding field and tree received lower producer's accuracy of 85% and 77%, respectively. It is clear that these two of classes are similar spectral. So, it is difficult to distinguish between each other just based on spectral information. The time-frequency transformed based method of texture discrimination, which is in turn based on Gabor filters, is done in this paper. It shows that in the case of combination with Gabor filter features improved correct rate significantly according to type 4 to type 6 in Table 1. In the most of case, the SVM method gave the highest correct classification rate within these three methodologies. Decision tree and SVM have their superiority respectively.

Table 1. Summary of Features Combination and It Correct Rate.

Features	**Classification Results**					
	Max. Likelihood(%)		***Decision Tree(%)***		***SVM (%)***	
	Overall accuracy	***Kappa***	***Overall accuracy***	***Kappa***	***Overall accuracy***	***Kappa***
Type 1	82.22	0.71	84.73	0.76	85.50	0.79
Type 2	82.15	0.71	84.60	0.76	85.99	0.78
Type 3	82.32	0.72	84.68	0.76	85.50	0.77
Type 4	89.75	0.84	91.07	0.86	91.68	0.87
Type 5	90.04	0.84	91.62	0.87	91.68	0.87
Type 6	89.84	0.84	91.76	0.87	91.76	0.87

6. CONCLUSION

This paper has proposed a method of features extraction by NDVI, PCA, and Gabor filter transform operator, and then, combine these features for Quick Bird remote sensing images classification. The schemas of classification include SVM and Decision Tree. The proposed algorithms are evaluated and compared with the Maximum Likelihood ones. The result shows that PCA and NDVI features for classification do not increased the correct rate in this study. Image textures extracted by Gabor filter are significant improving the classification result. SVM classifier gives the best classification result in the most of case.

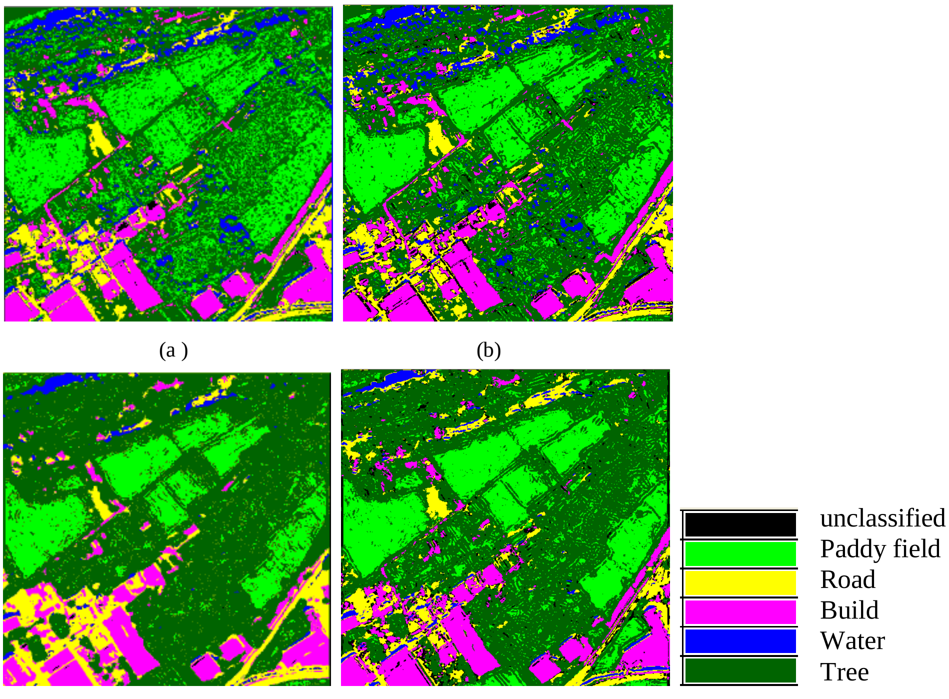

Figure 3. Results of full image classification: (a) Type 1 features; (b) Type 6 features(by Max. Likelihood); (c) Type 6 features(by Decision Tree); (d) Type 6 features(by SVM).

References

Cohen, F.S., Z. Fan, and M.A. Patel, 1991. Classification of rotated and scaled textured images using Gaussian Markov random field models, IEEE Trans. Pattern Anal. Mach. Intell. 13, pp. 192−202.

DeValois, R. L., D. G. Albrecht, and L. G. Thorell, 1982. Spatial frequency selectivity of cells in macaque visual cortex, Vis. Res., 22, pp. 545–559.

Franz, E., M. R. Gebhardt, and K. B. Unklesbay, 1991. The use of local spectral properties of leaves as an aid for identifying weed seedlings in digital images, Trans, ASAE, 32(2), pp. 682–687.

Guyer, D. E., G. E. Miles, M. M. Shreiber, O. R. Mitchell, and V. C. Vanderbilt, 1986. Machine vision and image processing for plant identification, Trans. ASAE., 29(6), pp. 1500–1507.

He, D.C. and L. Wang, Unsupervised textural classification of images using the texture spectrum, Pattern Recognition 25 (1992) 247−255.

Manjunath, B. S., and W. Y. Ma, 1996. Texture features for browsing and retrieval of data, IEEE Trans. Pattern Analysis Machine Intelligence, 18(8), pp. 837-842.

Mao, J. and A.K. Jain, Texture classification and segmentation using multiresolution simultaneous autoregressive models, Pattern Recognition 25 (1992) 173−188.

Soille, P., 2005. Beyond self-duality in morphological image analysis, Image and Vision Computing, 23(2), pp. 249–257.

Volpe, F., and L. Rossi, 2003. Quickbird high resolution satellite data for urban application, in Proc. 2nd GRSS/ISPRS Joint Workshop Data Fusion and Remote Sens. Over Urban Areas, May, pp. 1–3.

Vyas, Vibha S., and Priti Rege, 2006. Automated Texture Analysis with Gabor filter, GVIP Journal, 6(1), July, pp.35-41.

Yonekawa, S., N. Sakai, and O. Kitani, 1996. Identification of idealized leaf types using simple dimensionless shape factors by image analysis," Trans. ASAE, 39(4), pp. 1525–1533.