Fundamenta Informaticae 151 (2017) 325–339
DOI 10.3233/FI-2017-1495
IOS Press

Tea Category Identification using Computer Vision and Generalized Eigenvalue Proximal SVM

Shuihua Wang
School of Electronic Science and Engineering
Nanjing University
Nanjing, Jiangsu 210046, China

Preetha Phillips
Department of Supply Chain Management
W. P. Carey School of Business
Arizona State University
P.O. Box 873406, Tempe, AZ 85287, USA

Aijun Liu
College of Communications Engineering
PLA University of Science and Technology
Nanjing, Jiangsu 210007 - China

Sidan Du*
School of Electronic Science and Engineering
Nanjing University, Nanjing, Jiangsu 210046, China
coff128@nju.edu.cn

Abstract. (Objective) In order to increase classification accuracy of tea-category identification (TCI) system, this paper proposed a novel approach. (Method) The proposed methods first extracted 64 color histogram to obtain color information, and 16 wavelet packet entropy to obtain the texture information. With the aim of reducing the 80 features, principal component analysis was harnessed. The reduced features were used as input to generalized eigenvalue proximal support vector machine (GEPSVM). Winner-takes-all (WTA) was used to handle the multiclass problem. Two kernels were tested, linear kernel and Radial basis function (RBF) kernel. Ten repetitions of 10-fold stratified cross validation technique were used to estimate the out-of-sample errors. We named our method as GEPSVM + RBF + WTA and GEPSVM + WTA. (Result) The results showed that PCA reduced the 80 features to merely five with explaining 99.90% of total variance. The recall rate of GEPSVM + RBF + WTA achieved the highest overall recall rate of 97.9%. (Conclusion) This was higher than the result of GEPSVM + WTA and other five state-of-the-art algorithms: back propagation neural network, RBF support vector machine, genetic neural-network, linear discriminant analysis, and fitness-scaling chaotic artificial bee colony artificial neural network.

*Address for correspondence: School of Electronic Science and Engineering, Nanjing University, Nanjing, Jiangsu 210046, China

Keywords: Tea category identification; computer vision; color histogram; wavelet packet entropy; winner-takes-all; radial basis function; artificial neural network; pattern recognition; support vector machine

1. Introduction

Tea is an aromatic beverage [1]. This paper focused on green, black, and Oolong tea, because they are the most popular categories in China. There were scientific evidences that tea could help prevent some diseases, such as neurodegenerative disease [2], blood pressure and cardiovascular disease [3], Parkinson's disease [4], Alzheimers disease [5], breast cancer [6], etc.

Developing tea-category identification (TCI) system is of essential importance for tea market, since different tea category corresponds to different customer types who are willing to afford their favorite tea categories [7]. Traditional approaches have shortcoming to accurately assess, predict, and control of tea quality. Recently, scholars proposed use new measurement devices and new computer vision approaches. Herrador and Gonzalez [8] selected 8 types of metals. back propagation neural network (BPNN) were used with recall rate of 95%. Chen, Zhao [9] proposed to use near-infrared (NIR) reflectance spectroscopy. Radial basis function support vector machine (RBFSVM) was picked up for classification.

On the other hand, computer vision systems attract more and more popularity due to its cost effectiveness, consistency, superior speed, and high accuracy. Jian, Xianyin [10] employed computer vision approaches to identify the tea on the basis of their color and shape parameters. They employed a genetic neural-network (GNN) as the classifier. Chen, Zhao [11] employed twelve texture features and twelve color features. The 24 features were reduced to 11. linear discriminant analysis (LDA) and PCA helped to generate the recognition system. Borah, Hines [12] described a new feature estimation approach for recognizing 8 different grades over CTC tea. Gill, Kumar [13] pointed out that image analysis was non-destructive for tea identification. They used granule color, size, texture and shape features. Laddi, Sharma [14] analyzed ten graded tea samples. The best result was obtained in the condition of dark-field illumination. Zhang, Wang [15] utilized an artificial neural network (ANN) to classify fruits. They proposed fitness-scaling chaotic artificial bee colony (FSCABC) method to train the ANN.

In this work, we chose the latter school of tea identification, i.e., computer-vision based methods. To realize it, we combined three stages: (I) Feature extraction, (II) feature reduction and (III) classification. In Stage I, we used color histogram and wavelet packet entropy to extract features. In Stage II, PCA was introduced to reduce features. In Stage III, we introduced generalized eigenvalue proximal support vector machine (GEPSVM) with linear and RBF kernels.

The structure of the rest was organized as follows. Next section described the sample preparation and image acquiring procedures. Section Computer Vision-based Feature Processing showed the feature extraction and reduction methods. Section Classifier offered a rather novel SVM variant, dubbed as GEPSVM. Besides, the multiclass technique was introduced. Section Implementation of proposed methods depicted K-fold stratified cross validation, and the implementation of proposed methods. Section Experiments and Results compared our methods with state-of-the-art Approaches. Discussions can be found in Section Discussion. Final section Conclusion and Future Research concluded the paper and gave future research directions. We explained the acronyms in Table 5.

2. Materials

2.1. Sample preparation

Currently, there are no public dataset for tea leaves. Most researchers used their private data for the experiment. We also obtain the sample by ourselves. The characteristics of tea leave samples are listed in Table 1. All 150 tea sample images of 3 categories were obtained from different brands from different Chinese provinces, with the aim of increasing the generalization ability.

Table 1. Characteristics of tea leave

Category	Number	Origins
Green tea	50	Henan (7); Guizhou (9); Jiangxi (9); Anhui (7); Zhejiang (10); Jiangsu (8)
Black tea	50	Yunnan (12) ; Hunan (15); Hubei (11); Fujian (12)
Oolong tea	50	Fujian (20); Guangdong (30)

(All teas were on stock within four months period)

2.2. Image acquiring

The TCI system used in this study was shown in Figure 1, which consisted of five basic parts: (1) illumination system, (2) a 3-CCD camera, (3) a capture board, (4) computer hardware, and (5) computer software. The lighting arrangements were grouped either front or back lighting [16]. Front lighting is used for surface features, whereas the back lighting is for generating silhouette image. In this study, front lighting was used. The preprocessing process removes the background and selects a region of 256×256 pixels.

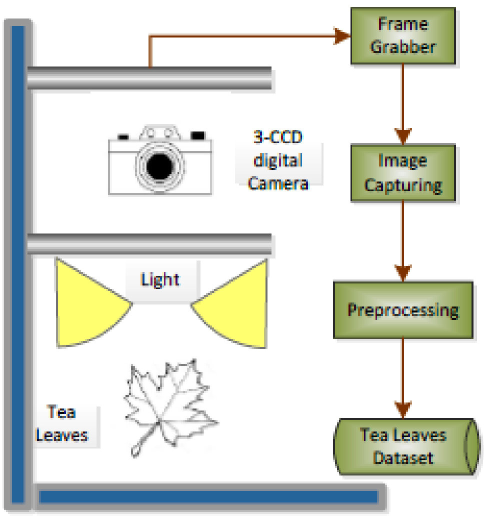
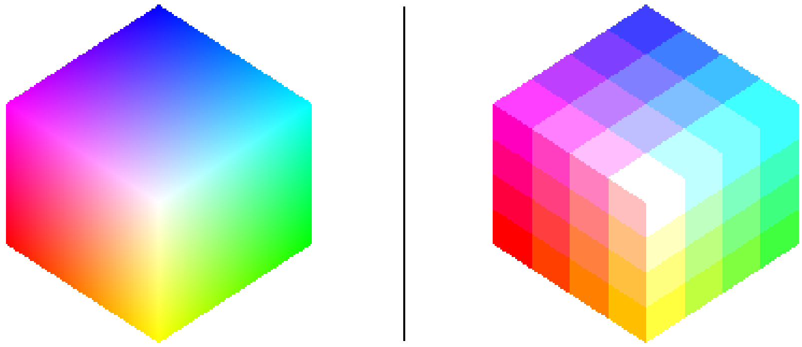


Figure 1. Our TCI system

3. Computer vision-based feature processing

3.1. Color histogram

The color histogram (CH) was widely harnessed to denote color distribution in an image [17]. Figure 2a shows the intensity value of each channel with value from 0 to 255. Then, we use four bins [15] to discretize RGB channels. The four bins are associated with intensities 0-63, 64-127, 128-191, and 192-255, respectively. Hence, 64 new bins representing different colors are produced (See Figure 2b).



(a) Full RGB Space ($256 \times 256 \times 256$) (b) CH Result ($4 \times 4 \times 4$)

Figure 2. Our TCI system

3.2. Wavelet packet entropy

Traditionally, there are many methods to extract texture features. Local binary pattern (LBP) was one traditional method. Nevertheless, wavelet entropy as an emerging texture descriptor, have proven better performance than LBP due to its multi-scale description over order/disorder degree. In this paper, We further enhanced its ability by using wavelet packet entropy (WPE).

Discrete Wavelet packet transform (DWPT) can be treated as an extension of standard discrete wavelet transform (DWT). Figure 3 shows that all nodes in DWPT split further at each decomposition level, which is not allowed for DWT. The continuous WPT of a particular signal $y(t)$ is:

$$C_{u,s} = \int_{-\infty}^{\infty} y(t) w(2^{-s}t - a) dt \tag{1}$$

where s represents the current decomposition level, u the channel number, a the position, w the wavelet function Z the integer and k the index. The next decomposition is written as:

$$C_{2u,\,s+1} = \sum_{a \in Z} h(a - 2k) C_{u,s} \tag{2}$$

$$C_{2u+1,\,s+1} = \sum_{a \in Z} l(a - 2k) C_{u,s} \tag{3}$$

Note that for s-level decomposition, DWT generates $(3s + 1)$ different coefficient sets, while DWPT generates 2^s sets. Hence, the latter provides more information. Nevertheless, their overall coefficient numbers are equivalent due to the downsampling process.

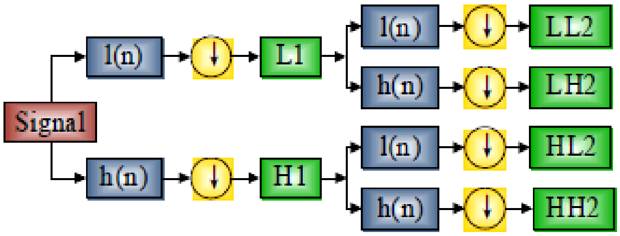


Figure 3. Flowchart of 2-level 1D-DWPT

To calculate WPE, each coefficient sets obtained by DWPT were regarded as an image, and Shannon entropy (SE) were implemented on all such images. SE was defined as

$$S = -\sum_{i=1}^{Z} k_i \log_2(k_i) \tag{4}$$

where i represents the greylevel, Z the total number of greylevels, and k_i the probability of greylevel of i. The entropies over all subbands of discrete wavelet packet transform is termed wavelet packet entropy (WPE), which can capture the wavelet features efficiently and applied to various fields successfully [18-20].

Finally, principal component analysis (PCA) is used to reduce the 80 features, viz., 64 CH features and 6 WPE features. The reduced feature is termed of principal components (PC). The threshold of PCA in this work is set to 99.90%.

4. Classifier

Traditional support vector machine (SVM) is the most popular tool for supervised learning. In this work. We employed its variant, the generalized eigenvalue proximal support vector machine (GEPSVM), with two kernels: linear kernel and RBF kernel.

4.1. Generalized eigenvalue proximal SVM

Mangasarian and Wild [21] proposed the generalized eigenvalue proximal support vector machine (GEPSVM), which discarded the parallelism restrain [22, 23]. Suppose X_1 and X_2 denotes the data of class 1 and 2, respectively. GEPSVM produces following two nonparallel hyperplanes

$$\mathbf{w}_1^T x - b_1 = 0 \quad \text{and} \quad \mathbf{w}_2^T x - b_2 = 0 \tag{5}$$

where w represents the weights and b the biases. Above equation can be transformed to an optimization problem:

$$(\mathbf{w}_1 b_1) = \underset{(w,b\neq 0)}{\arg\min} \frac{||\mathbf{w}^T X_1 - e^T b||^2/||r||^2}{||\mathbf{w}^T X_2 - e^T b||^2/||r||^2} \tag{6}$$

$$r \overset{def}{=} \begin{bmatrix} \mathbf{w} \\ b \end{bmatrix} \tag{7}$$

Here, e represents a vector of any number of ones

Simplifying equation (6) gives

$$\min_{(w,b\neq 0)} \frac{||\mathbf{w}^T X_1 - e^T b||}{||\mathbf{w}^T X_2 - e^T b||} \tag{8}$$

Scholars introduce Tikhonov regularization term to reduces $||r||$

$$\min_{(w,b\neq 0)} \frac{||\mathbf{w}^T X_1 - e^T b|| + \delta ||r||^2}{||\mathbf{w}^T X_2 - e^T b||} \tag{9}$$

here $||$ means the norm and δ is a nonnegative factor. Equation 18 turns to the so-called *Rayleigh quotient* as

$$r_1 = \underset{(r\neq 0)}{\arg\min} \frac{r^T M_1 r}{r^T M_2 r} \tag{10}$$

here M_1 and M_2 are symmetric defined as

$$M_1 \overset{def}{=} [X_1 - e]^T + \delta I \tag{11}$$

$$M_2 \overset{def}{=} [X_2 - e]^T \tag{12}$$

Solutions of Equation 20 can be yielded by solving a generalized eigenvalue problem [23] of

$$M_1 r = \lambda M_2 r, \quad r \neq 0 \tag{13}$$

4.2. Kernel strategy and multiclass technique

Standard SVMs cannot handle linear classification problem, i.e., different categories of data distributed at different sides of a hypersurface; therefore, scholars proposed kernel technique to solve it. We used radial basis function (RBF) kernel because it has been reported to behavior better than any other kernels. The equation of RBF is written as:

$$J(x_i, x_j) = \exp\left(-\frac{|x_i - x_j||^2}{2\sigma^2}\right) \tag{14}$$

where J represents the RBF kernel and σ the scaling factor. The optimal value of parameter σ was obtained by grid-searching method (See Section Optimal Scaling Factor).

Traditionally, SVMs were for binary classification problem (BCP), but three classes exist in this work. The most successful multiclass technique was to transform a multiclass problem into multiple BCPs. Winner-Takes-All (WTA) method was used. Suppose there are totally $D(> 2)$ different classes, in the training phase, D different binary classifiers (either SVM or its variants) are trained. In the test phase, all the D classifiers are run, and the classifier with the largest output is chosen.

5. Implementation of proposed methods

In this study, we used K-fold stratified cross validation (SCV) to evaluate the out-of-sample classification performance of our TCI system. Original samples were randomly partitioned into K mutually exclusively subsets of nearly equal size, in which K-1 subsets were used for training and the last one for validation. The abovementioned process iterated K times until each subset was used exactly once for validation. The K results over validation set from the K runs were then combined in order to yield a single estimation over the whole dataset. Then, The K-fold SCV repeated 10 times in order to reduce the variance of out-of-sample evaluation. K was determined as 10 for fair comparison.

As stated in the introduction, the proposed system consisted of three stages (Figure 4): feature extraction, feature reduction, and classification. Besides, our implementation was two-fold: offline learning and online prediction.

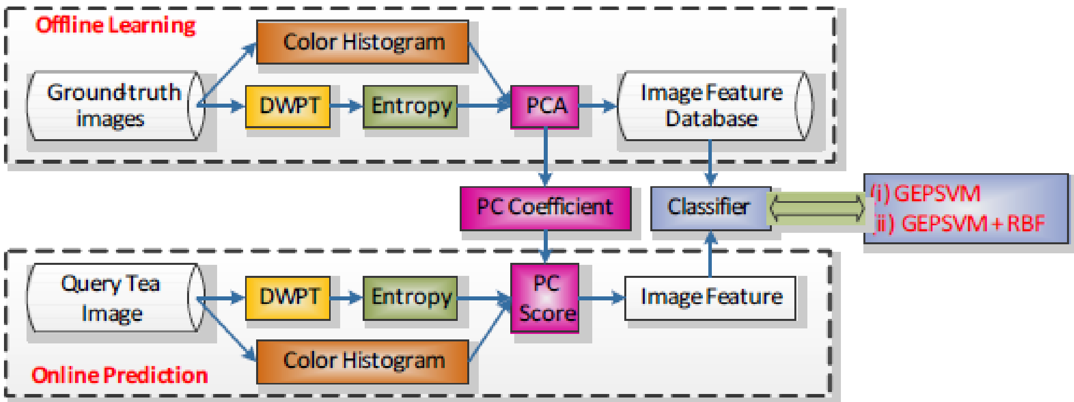


Figure 4. Flowchart of the proposed system

6. Experiments and results

6.1. Experiment tasks

In all, we proposed to use color histogram and WPE as features, PCA as feature reduction, and two classifiers: (i) GEPSVM + WTA; and (ii) GEPSVM + RBF + WTA. Five experiment tasks were designed:

1. We gave the feature extraction results of each category of tea.
2. We showed the relationship between numbers of PCs against explained variances.
3. We tested the two proposed identification methods, and compared them with state-of-the-art TCI approaches.
4. We showed how to find the optimal scaling factor by grid-searching approached.
5. We calculated the computation time of every step in the MCI system.

6.2. CH, DWT, and DWPT

Figure 5 shows the feature extraction results over three categories of tealeaves. Here G, O, and B represent green tea, Oolong tea, and black tea, respectively. The second row shows the samples of tea leave images. The 3-rd row showed the histograms. The 4-th and 5-th rows compared DWT with DWPT. 2-level Haar wavelet was harnessed.

There were three channels (RGB) of each image, so we performed decomposition for each channel and combined them to output the final decomposition result. Entropy was extracted on the 16 subbands of DWPT.

6.3. Principal component analysis

For each tea leave image, its 80 features were formed to a row vector. Afterwards, features of samples were aligned to form a two-dimensional matrix, in which row represents sample and column represents feature. Figure 6 shows the curve of variance explained versus PC number. 1 PC obtains 93.91% of total variances, 2 PCs obtains 99.08%, 3 PCs obtain 99.49%, 4 PCs obtain 99.78%, and finally 5 PCs obtained more than 99.90% of total variance. Therefore, we used 5 PCs due to the threshold of 99.90%.

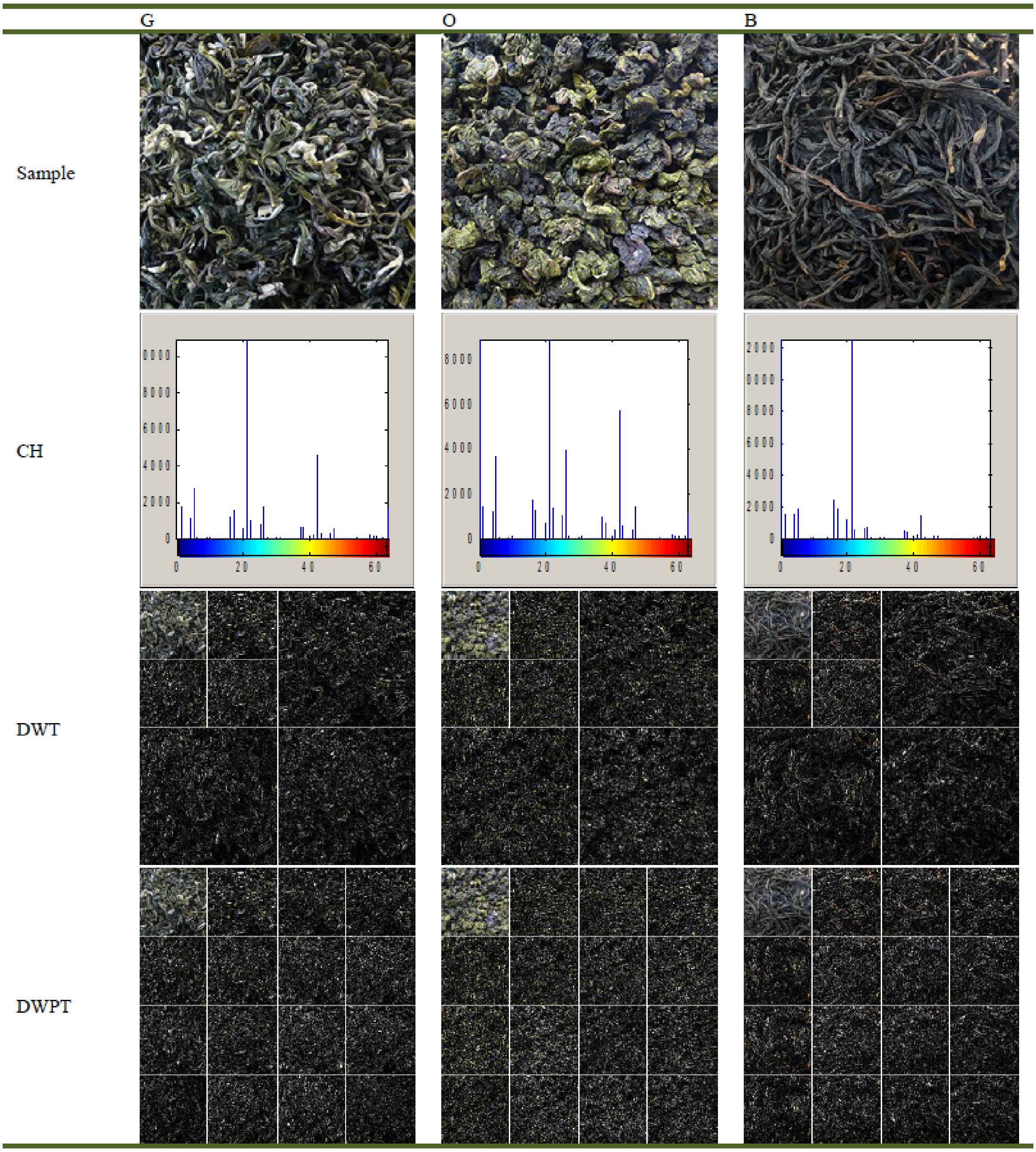
6.4. Kernel comparison

Our two classifiers, (i) GEPSVM + WTA; (ii) GEPSVM + RBF + WTA, were compared to find which kernel is better, linear or RBF. 10 $\times$ 10-fold SCV was used for statistical analysis. The recall results were listed in Table 2. Obviously, the RBF kernel is better.

Table 2. Kernel Comparison

Kernel	**G**	**O**	**B**
Linear	96.1 $\pm$ 1.4 %	98.2 $\pm$ 0.5 %	97.8 $\pm$ 1.1%
RBF	96.3 $\pm$ 1.2 %	99.4 $\pm$ 0.3 %	98.0 $\pm$ 0.8 %

(G = Green tea, O = Oolong tea, B = Black tea)



(G = Green tea, O = Oolong tea, B = Black tea)

Figure 5. Feature extraction of tea leave images

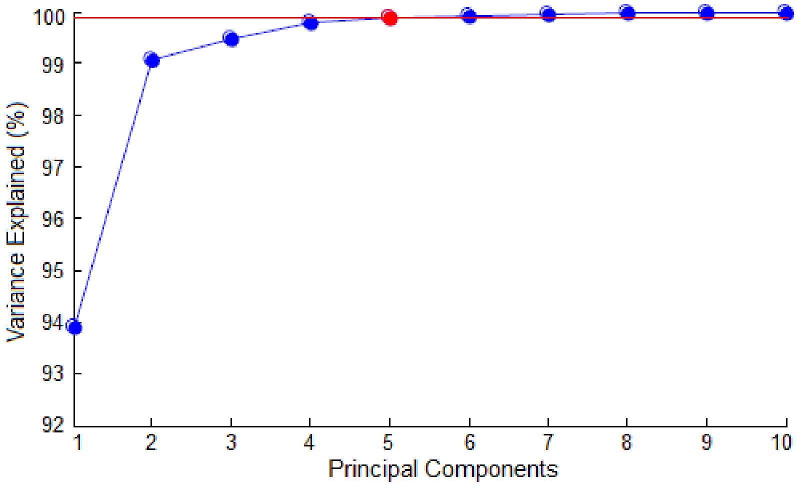


Figure 6. PCA result

6.5. Comparison with state-of-the-art methods

Based on a 10 × 10-fold SCV, we compared our GEPSVM + RBF + WTA with 5 state-of-the-art methods, including BPNN [8], RBFSVM [9], GNN [10], LDA [11], and FSC-ABC-ANN [15]. The results were listed in Table 3. Experiments in literature [8, 9] were also carried out to differentiate the same categories (G, O, and B)

Table 3. Sensitivity rate over a 10x10-fold stratified cross validation (* Data was obtained from the literatures)

Original Feature	(Reduced) Feature #	Classifier	G	O	B	Overall
8 metal	8	BPNN [8]	N/A	N/A	N/A	95%*
3735 spectrum	5	RBFSVM [9]	90%*	100%*	95%*	95%
2 color + 6 shape	8	GNN [10]	95.8%	94.4%	97.9%	96.0%
12 color + 12 texture	11	LDA [11]	96.7%	92.3%	98.5%	95.8%
64 CH + 8 shape + 7 texture	14	FSC-ABC-ANN [15]	98.1%	97.7%	96.4%	97.4%
64 CH + 16 WPE (Our)	5	GEPSVM + RBF + WTA	96.3 ± 1.2%	99.4 ± 0.3%	98.0 ± 0.8%	97.9%

(G = Green tea, O = Oolong tea, B = Black tea)

6.6. Optimal scaling factor

The scaling factor in the proposed GEPSVM + RBF + WTA was assigned with a value of 1.1. To obtain this optimal value, we employed grid searching approach, i.e., let the scaling factor σ change from 0.6 to 1.5 with increment of 0.1. Afterwards, we recorded the corresponding overall sensitivity rate. The results are pictured in Figure 7.

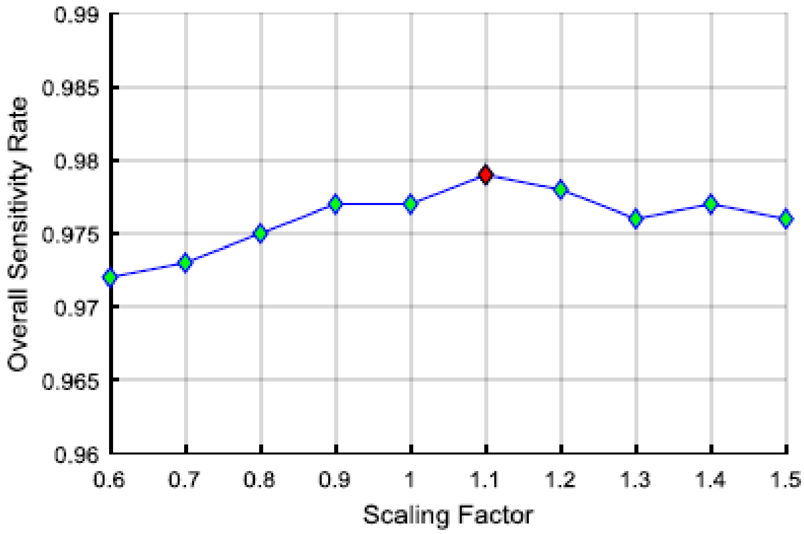


Figure 7. Determining the Optimal Scaling Factor (The red diamond frame represents the best value)

6.7. Time analysis

In the final experiment, Table 4 recorded the computation time. For the offline-learning phase, color histogram, WPE, PCA, and classifier training cost 3.844, 5.215, 0.849 and 0.930s, respectively. For the online-prediction phase, color histogram, WPE, PC score, and tea identification cost 0.059s, 0.073s, 0.004s, and 0.003s, respectively.

Table 4. Time Analysis of GEPSVM + RBF + WTA

Offline learning	**Time (s)**
Color Histogram	3.844
WPE	5.215
PCA	0.849
Classifier training	0.930
Online Prediction	**Time (s)**
Color Histogram	0.059
WPE	0.073
PC Score	0.004
Tea Identification	0.003

7. Discussion

The third row in Figure 5 clearly indicate the color histograms by three categories of tealeaves diff from each other, therefore, the color histogram was an efficient measure. The 4-th and 5-th rows clearly show that DWPT provides 16 subbands while DWT provides 7 subbands for a 2-level decomposition, hence the former provides more features than the latter.

Table 2 showed that the proposed the linear kernel, i.e., the GEPSVM + WTA approach obtains 96.1% 1.4, 98.2 0.5%, and 97.8 1.1% over green tea, Oolong tea, and black tea, respectively. The RBF kernel, i.e., the GEPSVM + RBF + WTA approach yields 96.3 1.2%, 99.4 0.3%, and 98.0 0.8% over the three types of teas. Averaging the recall rates of the three types of teas, we know that the overall recall rate of GEPSVM + WTA is 97.3%, while GEPSVM + RBF + WTA yields 97.9%. In all, we demonstrate the RBF kernel was more effective than linear kernel. RBF augments the overall recall rate slightly (0.6%).

Table 3 shows that the better of our methods,, viz., GEPSVM + RBF + WTA, obtains the average recall rate of 97.9%, better than 5 state-of-the-art approaches including BPNN [8] of 95%, RBFSVM [9] of 95%, GNN [10] of 96.0%, LDA [11] of 95.8%, and FSC-ABC-ANN [15] of 97.4%. Besides, GEPSVM + RBF + WTA used the least features of 5, less than or equal to BPNN [8] of 8, RBFSVM [9] of 5, GNN [10] of 8, LDA [11] of 11, and FSC-ABC-ANN [15] of 14. The excellent recall rate proved the effectiveness of those 5 features used in this study.

There are three reasons. First, the WPE exists better ability to analyze transient features of non-stationary signals than traditional features. Second, the GEPSVM is a rather novel classification method, which drops the parallelism restrain required by canonical support vector machine, so obtaining a more flexible hyperplane. Third, the RBF kernel can map the data to a non-linear space.

The curve in Figure 7 shows that the overall sensitivity rate of this proposed algorithm reaches the highest value as the scaling factor is equal to 1.1. This explains why we determine 1.1 as the optimal scaling factor. In addition, too small or too large scaling factors will deteriorate the classification performance.

Table 4 shows that the whole offline-learning procedure consumes 3.844 + 5.215 + 0.849 + 0.930 = 10.838s, whereas the whole online prediction cost merely 0.059 + 0.073 + 0.004 + 0.003 = 0.139s. The offline learning takes longer time than the online prediction, because the former needed to perform over all 150 images, while the latter just cope with one individual tea leave sample. Besides, the PCA and classifier training in the offline learning had already obtained necessary data (the PC coefficient matrix and the weights/biases of classifier), so the prediction directed used these data to obtain PC scores and identification result.

RBF kernel is a popular kernel that is commonly used in SVM-based classification. Some other kernels, including triangular kernel, quadratic kernel, logistic kernel, polynomial kernel, etc., may produce superior performance to RBF kernel. They will be tested in the future research.

Generally speaking, tea-leaves have random but persistent patterns, and do not contain any detectable quasi-periodic structure. Rossatto, Casanova [24] suggested that fractal theory is better than statistical, spectral, and structural approaches for describing natural textures . Therefore, we will test the fractal analysis for the tea-leaf classification.

8. Conclusion and future research

In this study, we proposed to use color histogram to describe color information, and to use WPE to describe wavelet features. Then, we proposed two classifiers (GEPSVM, and GEPSVM + RBF). Afterwards, WTA technique was introduced to deal with three-class problem. The results showed that our GEPSVM + RBF + WTA method performed the best. In addition, the proposed methods used only 5 features, no more than the feature numbers of existing methods.

In closing, we successfully develop a TCI system of tea category with high classification accuracy. Our contributions are: (i) we introduced a relatively new wavelet packet entropy (WPE) feature for tea category identification; (ii) We introduced GEPSVM and kernel technique that had better performance than standard SVM; and (iii) Our method used the least feature numbers and yielded higher identification rate than other methods.

Future works consist of following aspects: (i) we will test other novel classifiers and other kernel techniques; (ii) some advanced entropy methods may be used, such as Tsallis entropy, Hartley entropy, collision entropy, and Renyi entropy. (iii) new optimization methods, such as clonal selection algorithm, will be used to train the classifier. (v) fractal analysis will be tested.

Conflict of interest

We have no conflicts of interest to disclose with regard to the subject matter of this paper.

Acknowledgment

This paper was supported by the National Natural Science Foundation of China (No.61271231).

References

[1] Yang CS, Landau JM. Effects of tea consumption on nutrition and health. Journal of Nutrition. 2000; 130(10):2409–2412.

[2] Sironi E (et al.). Natural Compounds against Neurodegenerative Diseases: Molecular Characterization of the Interaction of Catechins from Green Tea with A beta 1-42, PrP106-126, and Ataxin-3 Oligomers. Chemistry-a European Journal. 2014;20(42):13793–13800. doi:10.1002/chem.201403188.

[3] Bohn SK (et al.). Effects of black tea on body composition and metabolic outcomes related to cardiovascular disease risk: a randomized controlled trial. Food & Function. 2014;5(7):1613–1620. doi:10.1039/c4fo00209a.

[4] Qi H and Li SX. Dose-response meta-analysis on coffee, tea and caffeine consumption with risk of Parkinson's disease. Geriatrics & Gerontology International. 2014;14(2):430–439. doi:10.1111/ggi.12123.

[5] Lim HJ, (et al.). Green tea catechin leads to global improvement among Alzheimer's disease-related phenotypes in NSE/hAPP-C105 Tg mice. Journal of Nutritional Biochemistry. 2013:24(7):1302–1313. doi:10.1016/j.jnutbio.2012.10.005.

[6] Yiannakopoulou EC. Green Tea Catechins: Proposed Mechanisms of Action in Breast Cancer Focusing on the Interplay Between Survival and Apoptosis. Anti-Cancer Agents in Medicinal Chemistry. 2014;14(2):290–295. doi:10.2174/18715206113136660339.

[7] Shi WM, (et al.). Tea Classification by Near Infrared Spectroscopy with Projection Discriminant Analysis and Gene Expression Programming. Analytical Letters. 2015;48(18):2833–2842. doi:10.1080/00032719.2015.1055574.

[8] Herrador MA and Gonzalez AG. Pattern recognition procedures for differentiation of Green, Black and Oolong teas according to their metal content from inductively coupled plasma atomic emission spectrometry. Talanta. 2001;53(6):1249–1257.

[9] Chen QS, (et al.). Feasibility study on identification of green, black and Oolong teas using near-infrared reflectance spectroscopy based on support vector machine (SVM). Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy. 2007;66(3):568–574. doi:10.1016/j.saa.2006.03.038.

[10] Jian W, Xianyin Z, and ShiPing D. Identification and grading of tea using computer vision. Applied Engineering in Agriculture. 2010;26(4):639–645. doi:10.13031/2013.32051.

[11] Chen Q, Zhao J, and Cai J. Identification of tea varieties using computer vision. Transactions of the Asabe. 2008;51(2):623–628. doi:10.13031/2013.24363.

[12] Borah S, Hines EL, and Bhuyan M. Wavelet transform based image texture analysis for size estimation applied to the sorting of tea granules. Journal of Food Engineering. 2007;79(2):629–639. Available from: `http://dx.doi.org/10.1016/j.jfoodeng.2006.02.022`.

[13] Gill GS, Kumar A, and Agarwal R. Monitoring and grading of tea by computer vision - A review. Journal of Food Engineering. 2011;106(1):13–19. doi:10.1016/j.jfoodeng.2011.04.013.

[14] Laddi A, (et al.). Classification of tea grains based upon image texture feature analysis under different illumination conditions. Journal of Food Engineering. 2013;115(2):226–231. doi:10.1016/j.jfoodeng.2012.10.018.

[15] Zhang Y, (et al.). Fruit classification using computer vision and feedforward neural network. Journal of Food Engineering. 2014;143(0):167–177. doi:10.1016/j.jfoodeng.2014.07.001.

[16] Tai YH, Chou LS, and Chiu HL. Gap-Type a-Si TFTs for Front Light Sensing Application. Journal of Display Technology. 2011;7(12):679–683. doi:10.1109/JDT.2011.2164054.

[17] de Almeida VE, (et al.). Using color histograms and SPA-LDA to classify bacteria. Analytical and Bioanalytical Chemistry. 2014;406(24):5989–5995. doi:10.1007/s00216-014-8015-1.

[18] Li YB, (et al.). Radar emitter signal recognition based on multi-scale wavelet entropy and feature weighting. Journal of Central South University. 2014;21(11):4254–4260. doi:10.1007/s11771-014-2422-5.

[19] Moshrefi R, Mahjani MG, and Jafarian M. Application of wavelet entropy in analysis of electrochemical noise for corrosion type identification. Electrochemistry Communications. 2014;48:49–51. doi:10.1016/j.elecom.2014.08.005.

[20] Yang YH, (et al.). Wavelet kernel entropy component analysis with application to industrial process monitoring. Neurocomputing. 2015;147:395–402. doi:10.1016/j.neucom.2014.06.045.

[21] Mangasarian OL, and Wild EW. Multisurface proximal support vector machine classification via generalized eigenvalues. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2006;28(1):69–74. doi:10.1109/TPAMI.2006.17.

[22] Khemchandani R, Karpatne A, and Chandra S. Generalized eigenvalue proximal support vector regressor. Expert Systems with Applications. 2011;38(10):13136–13142. doi:10.1016/j.eswa.2011.04.121.

[23] Shao YH, (et al.). Improved Generalized Eigenvalue Proximal Support Vector Machine. IEEE Signal Processing Letters. 2013;20(3):213–216. doi:10.1109/LSP.2012.2216874.

[24] Rossatto DR, (et al.). Fractal analysis of leaf-texture properties as a tool for taxonomic and identification purposes: a case study with species from Neotropical Melastomataceae (Miconieae tribe). Plant Systematics & Evolution. 2011;291(291):103–116. doi:10.1007/s00606-010-0366-2.

Appendix

Table 5 contain the mathematical symbols and abbreviations used in this study.

Table 5. Abbreviation List

Abbreviation	Definition
ABC	Artificial bee colony
BCP	Binary classification problem
BP	Back propagation
CCD	Charge-coupled device
CH	Color histogram
DWPT	Discrete wavelet packet transform
FSC	Fitness-scaling chaotic
GEPSVM	Generalized eigenvalue proximal SVM
GNN	Genetic neural-network
LDA	Linear discriminant analysis
NIR	Near-infrared
RBF	Radial basis function
SE	Shannon entropy
SVM	Support vector machine
TCI	Tea-category identification
WPE	Wavelet packet entropy
WTA	Winner-Takes-All