Electronic Nose For Black Tea Quality Evaluation Using Kernel Based Clustering Approach

Ashis Tripathy ramakrishnam1984@gmail.com
Assistant Professor
Department of Electronics & Instrumentation Engineering
Siksha O' Anusandhan University
Bhubaneswar, Odisha, 751030, India

A. K. Mohanty akmohanty@gmail.com
Assistant Manager, OPTC Ltd.
Balasore, Odisha, 756029, India.

Mihir Narayan Mohanty mihir.n.mohanty@gmail.com
Associate Professor
Department of Electronics & Instrumentation Engineering
Siksha O' Anusandhan University
Bhubaneswar, Odisha, 751030, India

Abstract

Black Tea is conventionally tested by human sensory panel called "Tea Tasters", who assign quality scores to different tea samples. This paper proposed a method of separation using the device named as electronic noise. The various tea samples have been analyzed using the popular method of separation, like PCA and LDA. For better separation among different scores of tea samples, the kernel based PCA as well as kernel based LDA methods have been considered in this case as the clustering algorithm. The method exhibits a better performance than those of traditional methods. Also the separation index has been evaluated and shows its efficacy.

Keywords: Kernel, Feature Space, Nonlinear Mapping, Electronic Nose, Black Tea, PCA, KPCA, LDA, KLDA.

1. INTRODUCTION

The electronic nose technology has been successfully employed for recognition and quality analysis of various food and agro products, viz., wine [1], cola [2], meat [3], fish [4], coffee [5], etc. Instrumental evaluation of black tea quality is quite complex because of the presence of innumerable compounds and their multidimensional contribution in determining the final quality of tea. Present day practice in the tea industry is that experienced tea tasters are employed for this purpose and gradation of tea is done based on their scores. This method is purely subjective and an objective assessment using a low-cost instrument is a dire necessity in the tea industry today. For evaluation of quality of black tea aroma using electronic nose, pioneering work had been done by Dutta et al.[6], where the efficacy of the electronic nose instrument in classifying black tea aroma in different processing stages was established. Co-relation of electronic nose data and tea taster's score has been demonstrated in [7], wherein a stable model was developed and applied on data collected from some gardens of north and north-east India. Electronic nose has also been used successfully for detection of optimum fermentation time during the tea manufacturing process [8].

Different scores of tea samples have been separated using Kernel based Principal Component Analysis (KPCA) and Kernel based Linear Discriminant Analysis (KLDA) clustering techniques. KPCA and KLDA are the excellent statistical learning techniques [9]. Those methods are used to cluster the different groups of data samples. Due to its flexibility and good performance these are widely applied to various learning scenarios. Clustering of black tea aroma using electronic nose has been considered. The applicability of kernel principal component analysis and kernel linear discriminant analysis for data clustering has been demonstrated.

In this paper Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), KPCA, KLDA techniques have been examined successfully and their response are shown in fig 1, fig 2, fig 3, fig 4, respectively. Also the separation index is calculated for all the above techniques to identify the clarity of clustering in between different groups of data samples. Here the kernel based clustering is introduced and designed for the quality evaluation of tea samples and also various kernels have been tested.

The paper is organized as follows. Section-1 has introduced the work, Section-2 is the explanation of experimental setup, Section-3 describes the methodology for data analysis. Section-4 explains the separation index as the measuring parameter for separation, Section-5 discusses the result and section-6 concludes the work.

2. EXPERIMENTAL SETUP

2.1. Customized Electronic Nose Setup for Black Tea

A customized electronic nose setup has been developed for quality evaluation of tea aroma. The electronic nose consists of metal oxide semiconductor sensors. Each sensor is meant to sense specific chemicals for various tea samples. Five gas sensors from Figaro, Japan – TGS-832, TGS-823, TGS-2600, TGS-2610 and TGS-2611 constitute the sensor array for the setup. The experimental conditions of the electronic nose for classification of black tea aroma are given as follows:

- Amount of black tea sample = 50 grams,
- Temperature = $60^0 C \pm 3^0 C$,
- Headspace generation time = 30s,
- Data collection time =100s,
- Purging time = 100s,
- Airflow rate = 5 ml/s.

Dry tea samples have been used during the experiments in order to avoid the effect of humidity. The above experimental conditions have been optimized for black tea quality evaluation on the basis of repeated trials and sustained experimentation.

2.2. Sample Collection

One of the major problems is to collect the tea samples, as the tea industries are spread over dispersed locations in India. Also the quality of tea varies considerably on agro-climatic condition, type of plantation, season of flush and method of manufacturing. Experiments were carried out for approximately one-month duration each at the tea gardens of the following industries:

i) Khongea Tea Estate
ii) Mateli Tea Estate
iii) Glenburn Tea Estate
iv) Fulbari Tea Estate

The industries have multiple tea gardens spread across north and north-east India and the teas produced in their gardens are sent everyday to the tea tasting centers for quality assessment. All the companies had expert tea tasters and for our experiments, one expert tea taster was deputed by the respective industries to provide taster's score to each of the samples, which were subsequently considered for the discrimination study with the computational model.

3. Data analysis METHODOLOGY

3.1. Principal Component Analysis (PCA)

Principle component analysis has been widely used in modeling the statistics of a set of multi-dimensional data [10]. Based on PCA, KPCA method provides a technique for nonlinear feature extraction in the sample data. The nonlinearity is introduced by first mapping the data from the original input space into a higher dimensional feature space F using a nonlinear map $\Phi : R^N \rightarrow F$,where R is the set of real numbers and N is the dimension of the original input space and linear PCA is then performed in F using the mapped samples $\Phi(\mathrm{X}_k)$, X_k being the sample data.

By using PCA data may be expressed and presented in such a way as to highlight their similarities and differences. Since patterns in data can be difficult to observe in data of high dimension, PCA is a powerful tool for analyzing data. In the vector space, PCA identifies the major directions, and the corresponding strengths, of variation in the data. PCA achieves this by computing the eigenvectors and eigenvalues of the covariance matrix of the dataset. Keeping only a few eigenvectors corresponding to the largest eigenvalues, PCA can be also used as a tool to reduce the dimensions of the dataset while retaining the major variation of the data.PCA Plot for 174 data samples are shown in fig.1.

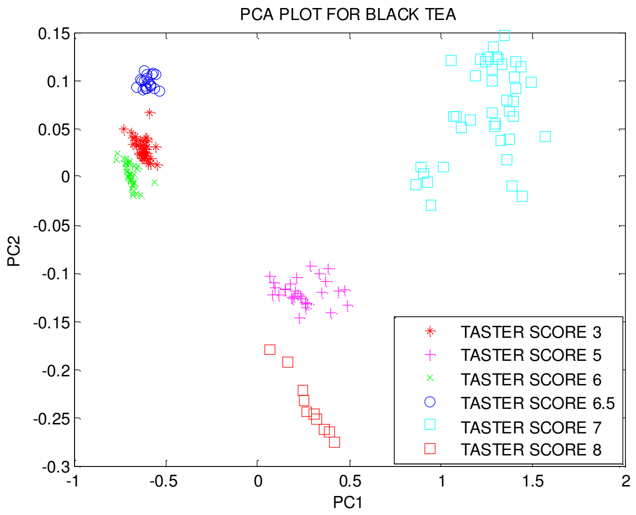

FIGURE 1: PCA plot for black tea samples

3.2. Kernel Principal Component Analysis (KPCA)

KPCA is an extension of principal component analysis using techniques of kernel methods [11]. In general, linear operation of PCA is done in a reproducing kernel Hilbert space with a non-linear mapping. In kernel PCA [12], an arbitrary function Φ is chosen, which is non-trivial. Generally the dimension of arbitrary function Φ is very high. So we generally try to avoid working in the Φ-space, which is known as 'feature space '.By using function Φ we can create the *N*-by-*N* kernel *K*=k(**x**,**y**)=(Φ(*X*),Φ(*Y*)) which represents the inner product of feature space. The kernel function is meant for the high dimensional features representation and is represented by

$$y_k = \sum_{i=1}^{M} \alpha_i^k k(\mathbf{x_i}, \mathbf{x}) \qquad (1)$$

where ***M***, is the number of samples , α_i^k is the i^{th} value of κth eigenvector of kernel ***K***, y_k is the κth (*q*...........*M*) value of sample after transforming, $\mathbf{x_i}$ is the i^{th} original sample, **x** is the original sample which is to be transformed, *q* is the sequence number of the first non zero eigen values in an ascending order. The dot product of the samples in the feature space is defined as

$$k(\mathbf{x_i},\mathbf{x_j}) = \Phi(X_i)^T g\Phi(X_j) \tag{2}$$

where X_i ,X_j (i,j = 1,2,……………M) are random samples of data sets. We use polynomial kernel function, which is defined as

$$k(\mathbf{x_i},\mathbf{x_j}) = (\Phi(X_i)g\Phi(X_j)+1)^d \tag{3}$$

After getting the value of κ, centralization of data is required to perform an effective principal component analysis, we centralize K to become $K^{'}$.

$$K^{'} = K - 1_N K - K1_N + 1_N K1_N \tag{4}$$

where 1_N denotes N X N matrix having each value 1.We use $K^{'}$to perform kernel PCA. KPCA result for 174 data samples are shown in fig.2.

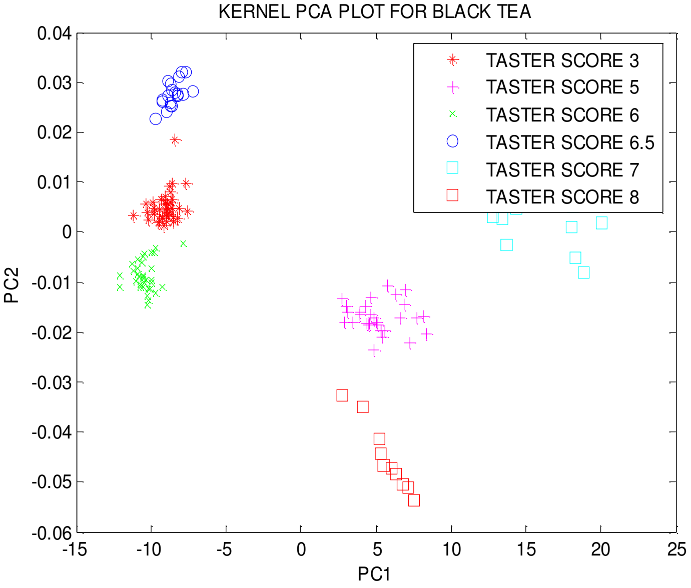

FIGURE 2: KPCA plot for black tea samples

3.3.Linear Discriminant Analysis(LDA)

Dimensionality reduction and classification project high-dimensional data onto a low dimensional space. The data achieves maximum class of separation and is done using Linear Discriminant Analysis (LDA) method [15]. The derived features are linear combinations of the original features. These coefficients are considered from the transformation matrix. In classical LDA, by minimizing the within-class distance and maximizing the between-class distance simultaneously optimal transformation is obtained, thus achieving maximum class discrimination. It has been applied successfully in many applications. Specifically, LDA seeks a transformation matrix W that maximizes the ratio of the 'between class cluster' to the 'within class cluster'.

The within class scatter matrix 'S_W' and between class scatter matrix 'S_b' are defined as

$$S_w = \sum_{i=1}^{C}\sum_{x\in c_i}(x-\mu_i)(x-\mu_i)^T \tag{5}$$

where

C - the number of classes ,

C_i - a set of data belongs to the i^{th} classes and

μ_i - the mean of i^{th} class.

The within class scatter matrix is the degree of scatter within classes as a summation of covariance matrices of all classes.

The between class scatter matrix S_b is defined as

$$S_b=\sum_{i=1}^{c} n_i(\mu_i-\mu)(\mu_i-\mu)^T \tag{6}$$

The between class scatter matrix represents the degree of scatter between classes as a covariance matrix of means of all classes. We seek a transformation matrix W that is some sense maximizes the ratio of the between-class scatter and the within class scatter.
The criterion function *J(W)* is defined as

$$J(W)=\frac{W^T S_b^{\Phi}\mathbf{w}}{W^T S_w^{\Phi}\mathbf{w}} \tag{7}$$

The transformation matrix *W* as one that maximizes the criterion function *J(W)* can be obtained. The column of optimal *W* are the generalized eigen vectors $\mathbf{w_i}$ that corresponds to the largest eigen values as

$$S_b\mathbf{W_i}=\lambda_i S_w\mathbf{W_i} \tag{8}$$

The result using LDA for 174 data samples are shown in fig.3.

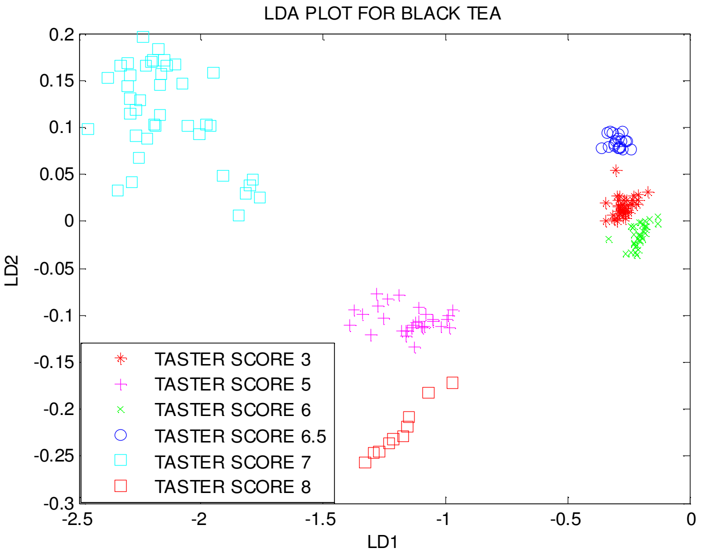

FIGURE 1: LDA plot for black tea samples

3.4. Kernel Linear Discriminant Analysis (KLDA)

The reduction of feature vector components in the feature extraction is done using the standard Linear Discriminant Analysis(LDA). The Kernel Linear Discriminant Analysis (KLDA) is a non-linear expansion of standard LDA to project the feature vectors onto the best discriminant features, while the non-linear projection is implicitly performed by the so called kernel trick. This is a way to represent the scalar-product of non linearly transformed feature vectors without performing the transformation itself. The resulting formulation is expressed as an eigenvalue problem, similar to the linear one. The size of the eigen value problem is equal to the number of input training vectors is a major setback. To get the largest eigen values and the corresponding eigen vectors we use the efficient Kernel Discriminant Analysis (KDA) algorithm.

Since nonlinear extension of LDA, kernel LDA [14] essentially performs LDA in feature space, Φ.
Using nonlinear mapping: Φ: C → f || X → Φ(x).
where C is a compact subset of R^N, linearly non-separable configuration becomes separable in Φ.
We can rewrite the objective function as

$$W_{opt}=\arg\max_{w}\left[\left(w^T s_b^{\Phi} w\right)/\left(w^T s_w^{\Phi} w\right)\right] \tag{9}$$

where,

$$S_b^\Phi = \sum_{i=1}^{j} n_i(\mu_j^\Phi - \mu^\Phi)(\mu_j^\Phi - \mu^\Phi)^T \tag{10}$$

$$S_w^\Phi = \sum_i \left(\Phi(X_i) - \mu_{k_i}^\Phi\right)\left(\Phi(X_i) - \mu_{k_i}^\Phi\right)^T \tag{11}$$

where Φ is the nonlinear mapping function,
μ_i - the mean of i^{th} class.
X_i – input data set
KLDA result for 174 data samples are shown in fig.4.
Similarly, as the nonlinear extension of LDA, kernel LDA [14] essentially performs LDA in feature space, Φ. Using nonlinear mapping: Φ: $C \rightarrow$ f || X $\rightarrow$ Φ(x); where C is a compact subset of R^N linearly non-separable configuration becomes separable in Φ.

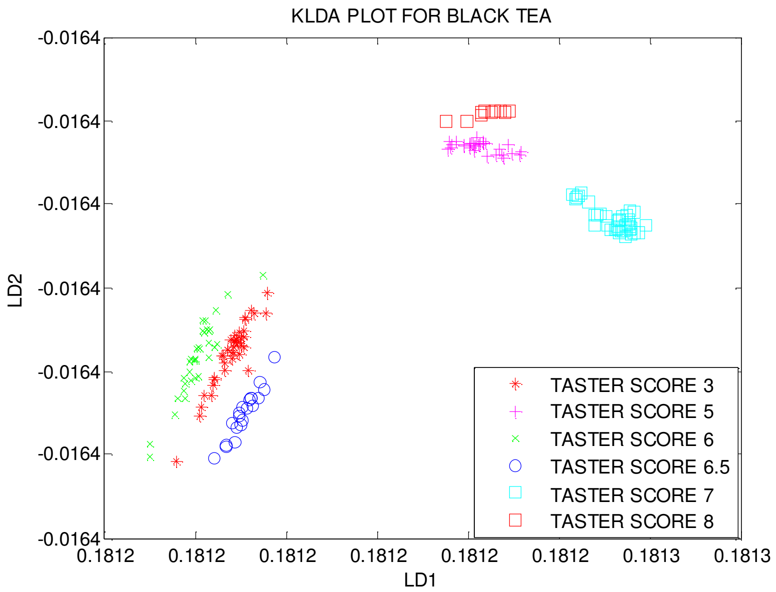

FIGURE 4: KLDA plot for black tea samples

4. SEPARABILITY INDEX

In general, decision on discrimination among data samples is not in proper way. So to extract the clarity of discrimination among the different groups of data samples we need a separability criterion. The separability measure is defined by the ratio of the trace of the 'between class scatter matrix' (S_B) to that of the 'within class scatter matrix' (S_W), and the expressions are given below:

$$S_B = \sum_{i=1}^{c} n_i (m_i - m)(m_i - m)^T \tag{12}$$

$$S_W = \sum_{i=1}^{c}\left(\sum_{j=1}^{n_i}(x_{i,j} - m_i)(x_{i,j} - m_i)^T\right) \tag{13}$$

where C is the number of classes, n_i denotes the number of samples in the i^{th} class, and $x_{i,\,j}$ denotes the j^{th} sample in the i^{th} class. m_i mean vector of the samples in the i^{th} class and m denotes the mean vector of the samples. Here the separation index values are obtained by using different clustering techniques for the black tea samples.

5. RESULTS AND DISCUSSION

Experimentations with electronic nose have been performed with 174 finished tea samples and sensor output signatures are logged in the computer. Total data size becomes 174×5.i.e.174 samples and each sample having 5 dimensions, as we have five sensors in our electronic nose system. PCA, KPCA, LDA, KLDA techniques have been applied on the 174 data samples .The PCA, KPCA, LDA, KLDA results are shown in Figure1,2,3,and 4 respectively. Due to collection of tea samples from different gardens and difference in their plugging time, samples may have different scores, still it is observed that they are overlapped to each other. Due to agro-climatic condition of a particular location, specific season of flush and clonal variation for the tea plant, the taster score 3 and 6 are overlapped to each other.

In this paper we applied PCA and LDA techniques, but we observed the overlapping nature of data samples, so to avoid this overlapping we applied the nonlinear approach called, KPCA and KLDA techniques, which have given better clarity in clustering as compare to PCA and LDA. Also sometimes it is not possible to properly visualize the clarity of clustering in data samples, so to identify the percentage of clarity of separation we calculated the separation index [18]. We have calculated the separation index for PCA which gives 56.7777, for LDA 56.7964, for KPCA 78.0039, and for KLDA, it is 84.9883. From the above separation index value definitely it has been observed that KPCA has a better clarity than linear PCA, and KLDA has better clarity than linear LDA. Table-1 shows different clustering techniques and their separation index values.

TABLE 1: separation index value in various techniques

Technique	Separation index value
PCA	56.7777
KPCA	78.0039
LDA	56.7964
KLDA	84.9883

6. CONCLUSION

The objective of the work was the clustering of multi sensor array data with tasters' scores in a discriminant fashion using electronic nose. As quality of black tea differ due to seasonal variations and different parameters of tea manufacturing process, it is most important to obtain the performance in discrimination on the different groups of black data samples. The KPCA and KLDA may be helpful as it shows the capability to reduce nonlinearity, and conversion of data samples to high dimensional feature space. Also it is very clear that Kernel approach has a great potential for efficient clustering in different groups of data samples. Further optimization may be applied to reduce the features. Also the accuracy is to be increased. It is left for the future scope.

ACKNOWLEDMENTS

The authors are deeply indebted to the Tea Industries, where experiments were performed during the study. The work has been sponsored by the National Tea Research Foundation, Tea Board; Department of Science & Technology and University Grants Commission, Government of India.

REFERENCES

[1] J. Lozano, J. P. Santos, M. Aleixandre, I. Sayago, J. Gutierrez, and M. C. Horrillo. "Identification of typical wine aromas by means of an electronic nose." IEEE Sensor J., 6(1) pp. 173-178, 2006.

[2] Kermani, B. G., Schiffman, S. S., and Nagle, H. T, "Performance of the Levenberg-Marquardt neural network training method in electronic nose applications." Sens. Actuators B, 110(1) pp. 13-22, 2005.

[3] Boothe, D. D. H., and Arnold, J. W.“Electronic nose analysis of volatile compounds from poultry meat samples, fresh and after refrigerated storage.” J. Sci. Food Agric., 82(3) pp. 315-322, 2002.

[4] O’Connell, M., Valdora, G., Peltzer, G. and Martin Negri, R., “A practical approach for fish freshness determinations using a portable electronic nose.” Sens. Actuators B, **80(2)** pp. 149-154, 2001.

[5] Pardo, M., and Sberveglieri, G., “Coffee analysis with an electronic nose.” IEEE Trans. Instrum. Meas., 51(6) pp. 1334-1339, 2002.

[6] Dutta, R., Hines, E. L., Gardner, J. W., Kashwan, K. R., and Bhuyan, M., “Tea quality prediction using a tin oxide-based electronic nose: An artificial intelligence approach.” Sens. Actuators B, 94, pp. 228-237, 2003.

[7] Bhattacharyya, N., Bandyopadhyay, R., Bhuyan, M., Tudu, B., Ghosh, D., and Jana, A., “Electronic nose for black tea classification and correlation of measurements with "Tea Taster" marks.” IEEE Trans. Instrum. Meas., 57(7), pp. 1313-1321, 2008.

[8] Bhattacharyya, N., Seth, S., Tudu, B., Tamuly, P., Jana, A., Ghosh, D., Bandyopadhyay, R., Bhuyan, M., and Sabhapandit, S. “Detection of optimum fermentation time for black tea manufacturing using electronic nose.” Sens. Actuators B,122(2), pp. 627-634, 2007.

[9] Scholkopf, B., Smola, A., and Muller, K-R., “Nonlinear component analysis as a kernel eigenvalue problem.” Technical Report No.44, Max-Planck Institute, Germany, 1996.

[10] Wall, M. E., Rechtsteiner, A., Rocha, L. M., “Singular value decomposition and principal component analysis.” in A Practical Approach to Microarray Data Analysis, Berrar, D. P., Dubitzky, W., and Granzow, M. Eds., Norwell, MA: Kluwer, (Chapter 5), pp.91–109, 2003.

[11] Hoffmann, H., “Kernel PCA for novelty detection.” Pattern Recognition, 40,pp. 863 – 874, 2007.

[12] Kim, K.I., Park, S.H., and Kim, H.J “Kernel principal component analysis for texture classification.” IEEE Signal Processing Letters, 8(2), pp.39-41., 2001.

[13] Dy, J. G., and Brodley, C. E., “Feature Selection for Unsupervised Learning.” J. Machine Learning Res., 5, pp.845-889, 2004.

[14] Scholkopf, B., Smola, A., and Muller, K-R.,”Nonlinear Component Analysis as a Kernel Eigenvalue Problem.” Neural Computation, vol.10,no.5,pp.1299-1319,1998.

[15] Duda D.S.R, Hart P. Pattern Classification. Wiley, New York, 2001.

[16] Belhumeur, N., Hespanha J., and Kriegman, D.,” Eigen faces vs. Fisher faces: Recognition Using Class Specific Linear Projection.” Proc.ECCV,pp.45-58,1996.

[17] Liu, C.J, and Wechsler, H.,”A Shape-and Texture-Based Enhanced Fisher Classifier for Face Recognition.” IEEE Trans.Image Processing. Vol.10, no.4. pp.598-608, 2001 .

[18] R. O. Duda, D. G. Stork, P. E. Hart, Pattern classification, 2nd edition, John Wiley and Sons, (pp. 115), (2001).

[19] Sangita D. Bharkad and Maneshkokare, “Performance evaluation of distance matrices: application to fingerprint recognition.” International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI), volume: 25, pp.777-806, 2011.