Apple Internal Quality Classification Based on Electronic Signal Analysis Using Sparse Principal Component Analysis

Mengbo You[1], Cheng Cai[1], Huiling Ma[2]
[1] College of Computer and Information Engineering,
[2] College of Life Science
Northwest A&F University, Yangling, Shaanxi, 712100
E-mail: chengcai@nwsuaf.edu.cn

***Abstract*— Apple internal quality classification is fairly important in fruit sales, packing and storage for producers and sellers. But internal properties like firmness, ascorbic acid content, and Ethylene production rate are difficult to measure without damaging the tested apples. Therefore, we propose a damage-free apple internal quality approach using such electronic signals as impedance, admittance, conductance, etc. With these multidimensional electronic signals, we adopt Principal Component Analysis (PCA), Sparse Principal Component Analysis (SPCA), and Gray Relational Analysis (GRA) to classify apple internal quality. According to three sets of our experiments, it is verified that PCA, SPCA and GRA all work well, and that SPCA outperforms the other two in terms of requirement of electronic signals and classification accuracy.**

I. INTRODUCTION

Apple internal quality classification is fairly important in fruit industries, including packing, distribution and storage for producers and sellers. For example, apples with high quality should be packed more carefully and stored together to guarantee freshness and delay deterioration. There have been some external apple quality detection systems detecting surface defects[1,2]. The internal quality properties like firmness, ascorbic acid content, and Ethylene production rate are difficult to measure without damaging the tested apples. Some researches show that apple dielectric properties are correlated with their internal quality[3]. In this paper, we propose a novel damage-free apple internal quality classification system based on electric signals analysis. For each apple testing sample, we extract 14 electronic signals including impedance, admittance, and conductance, etc. and 7 internal quality features such as firmness, vitamin, etc.

All testing samples are grouped into different quality categories based on their internal quality properties using hierarchical clustering. We apply Principal Component Analysis (PCA), Sparse Principal Component Analysis (SPCA), and Gray Relational Analysis (GRA) using electronic signals to predict the sample's quality categories.

This paper is organized as follows. Both the electronic signals and internal features to be measured are described in Section 2. Section 3 elaborates on the three algorithms of PCA, SPCA, and GRA. Section 4 continues to compare their internal quality classification performances. To conclude, Section 5 analyzes and summarizes the three sets of our experiments.

II. ELECTRONIC SIGNAL AND INTERNAL QUALITY FEATURE DEFINITION

We obtained the apple quality classification experimental data set which consists of 36 observations every 20 days from day 1 to day 220. And each observation includes 14 electronic signals and 7 internal quality features as defined in Table I and Table II respectively. As displayed in Fig. 1, biological device cannot provide damage-free measurements as electronic device do, but its measuring results can give credible classification of apples using clustering methods. Here, we explore techniques to predict the classification results only by electronic signals.

TABLE I. Apple electronic signals

Signal No.	Symbol	Property
1	Z	impedance (Ω)
2	C_s	static capacitance in series equivalent circuit mode (F)
3	L_s	inductance in series equivalent circuit mode (H)
4	R_s	effective resistance in series equivalent circuit mode (Ω)
5	C_p	static capacitance in parallel equivalent circuit mode (F)
6	L_p	inductance in parallel equivalent circuit mode (H)
7	R_p	effective resistance in parallel equivalent circuit mode (Ω)
8	G	conductance (S)
9	Y	admittance (S)
10	θ	impedance phase angle (°)
11	D	loss coefficient = tan(θ)
12	Q	quality factor
13	ε'	relative permittivity
14	ε''	dielectric loss factor = ε' tan(θ)

TABLE II. Apple internal quality features

Signal No.	Symbol	Property
15	H	firmness (kg•cm^{-2})
16	T	total soluble solid admittance (%)
17	A	titirable acid (%)
18	r	ratio = T/A
19	V_c	ascorbic acid (mg•kg^{-1})
20	r_i	respiration intensity (mgCO_2•kg^{-1}•h^{-1})
21	E	ethylene (μL•kg^{-1}•h^{-1})

III. ELECTRONIC SIGNAL EXTRACTION

In apple quality degeneration each feature seems to rise or fall in value, but there is no unified integral variation trend. It will be irrational to review and categorize this batch of apples only by one or some of these features, but there is a great possibility that some of these indicators may perform as redundant or negative factors in classification. Then an efficacious and scientific algorithm is in urgent need to rationally classify these apples and identify their corresponding quality levels.

A. Principal Component Analysis

Principal Component Analysis (PCA)[4] is a classical multivariate method used frequently in many fields for its simplicity. It is usually used to reduce dataset dimension. PCA transforms linearly a number of correlated variables to a smaller number of uncorrelated variables called principal components. The first principal component best describes the variability in the data, and each succeeding component best describes the remaining variability. All these principal components form a projection in which the dataset is most informative[5]. And the results of the analysis largely depend on the scaling of the specified data matrix[6]. Here we will introduce the derivation of PCA by using the covariance method[7].

Let the data set $X=\{x_1, x_2, \ldots x_n\}$ be an $m\times n$ matrix, where m is the number of features, n is the number of observations, and $x_i(i=1,2,\ldots,n)$ is one of the observations. The average observation and the difference after subtraction of the mean can be defined as

$$\bar{x} = \frac{1}{m}\sum_{i=1}^{n} x_i \ , \quad X' = x_i - \bar{x} \tag{1}$$

The goal is summarized as follows. Find a certain orthogonal matrix P so that the covariance matrix is

$$C_Y = \frac{1}{n-1} YY^T \text{ , subject to } \ Y = PX' \tag{2}$$

and the m orthogonal row vectors in P are the principal components If X^TX is a rank r, square, symmetric $m\times m$ matrix, we can use singular value decomposition [8] to solve the following equation:

$$X'^T X' v_i = \lambda_i v_i, i = 1, 2, ..., r \tag{3}$$

The final form of the decomposition is: $X=U\Sigma V^T$ which

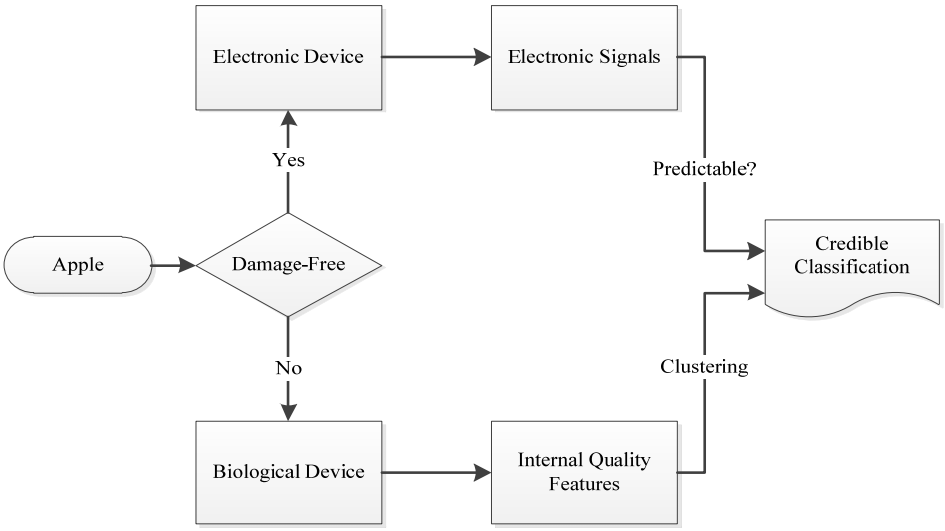

Fig.1 Apple internal quality classification predicted by electronic signals

TABLE III. Eigenvalues and contributions to variance of Principal Components

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Eigenvalue	*1.604*10^11*	*4.648*10^5*	*1.336*10^5*	*2863.0*	*334.8*	*156.3*	*0.3036*	*0.001127*
Percentage (%)	*0.999996*	*0.00000289809*	*8.32872*10^(-7)*	*1.785*10^(-8)*	*2.08768*10^(-9)*	*9.74462*10^(-10)*	*1.89325*10^(-12)*	*7.0264*10^(-15)*

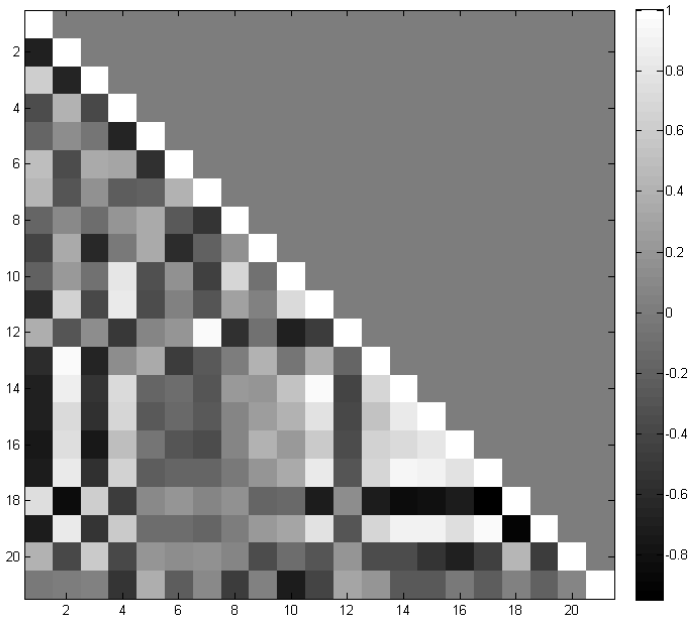

Fig. 2 Correlation coefficient matrix (Each row or column is one feature.)

TABLE IV. The eigenvectors corresponding to eigenvalues in descending order (a column is an eigenvector)

Features	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Z	-0.02647	0.05813	-0.05424	-0.8156	0.5231	-0.2321	-0.01297	0.0003457
Cs	-4.432*10^(-13)	-1.706*10^(-13)	-2.614*10^(-14)	-2.145*10^(-14)	4.897*10^(-13)	7.362*10^(-13)	8.044*10^(-12)	-4.467*10^(-11)
Ls	-9.917*10^(-10)	2.013*10^(-9)	-1.731*10^(-9)	-3.88*10^(-8)	-1.579*10^(-8)	-5.775*10^(-8)	1.166*10^(-6)	-4.218*10^(-6)
Rs	-0.008616	-0.02013	-0.01928	-0.2285	0.07227	0.9692	-0.0486	0.002356
Cp	-3.934*10^(-13)	6.179*10^(-13)	2.499*10^(-13)	-3.106*10^(-12)	-7.512*10^(-12)	-1.336*10^(-11)	-7.385*10^(-11)	-7.217*10^(-9)
Lp	-1.096*10^(-9)	1.046*10^(-9)	-2.916*10^(-9)	-2.94*10^(-8)	-3.01*10^(-9)	2.129*10^(-8)	3.046*10^(-7)	2.861*10^(-6)
Rp	-0.08398	0.2938	-0.9477	0.09038	0.009492	0.007716	0.01223	-0.00014
G	-3.629*10^(-6)	8.567*10^(-7)	2.215*10^(-6)	-9.084*10^(-5)	-2.555*10^(-7)	-1.486*10^(-5)	-0.002206	-0.1191
Y	-1.079*10^(-5)	9.44*10^(-6)	2.93*10^(-7)	-8.927*10^(-5)	-0.0002458	0.0002494	-0.006015	-0.07007
θ	0.03676	-0.05345	0.03839	0.5204	0.8481	0.05796	-0.02957	-0.00368
D	-0.0001729	-0.0003884	-0.000287	-0.002754	-0.002931	0.001788	-0.003341	-0.9904
Q	-0.001558	0.005769	-0.01097	0.007453	-0.0353	-0.04579	-0.9982	0.003988
ε′	-0.9445	0.2738	0.1735	0.04837	0.01994	0.01068	0.0003162	-0.0001647
ε"	-0.3142	-0.9122	-0.2585	-0.03375	-0.009133	-0.03417	-0.0003002	0.0005372

signifies that the matrix X can be converted to an orthogonal matrix, a diagonal matrix and another orthogonal matrix[9]. And the orthogonal matrix P can be easily obtained from this decomposition.

Principal Component Analysis requires the calculation of the covariance matrix of a data set. First the linear relation between the feature measurements presented by the correlation coefficient matrix (Shown by Fig. 2) is analyzed. Let $\mathrm{Coe}(X)_{ij}$ denote the value in the i-th row and the j-th column which means this number measures the correlation between the two feature variables. But a great number of values in the correlation coefficient matrix make it difficult to make comparisons unless we already have a pair of target features.

Table III presents the eigenvalues and contributions to principal components variance. The first column explains that the eigenvalue of PC2 is 0.999996, which verifies that some of the principal components can construct fairly good mapping space. The second row displays the energy percentage of each principle component. Table IV shows the eigenvectors corresponding to eigenvalues in descending order (each column represents one eigenvector). Obviously, value distribution in all the features appears to be random.

B. SPCA

PCA finds the linear combinations of the original variables so that the derived variables are able to capture maximal variance. PCA can be estimated via the singular value decomposition (SVD) of the data matrix. PCA makes itself successful mainly by this optimal property. The first several principal components can sequentially maximize the variability of the data matrix and remains uncorrelated at the same time[10]. However, the weakness is that principal components are linear combinations of all the original variables, or rather their loadings are mostly nonzero, which makes it difficult to interpret the derived variables. It should be considered to change the weighting coefficients of insignificant principal components to be zero.

SPCA[10] is one of such combination algorithms. In fact, the L1-norm-based penalization technique of lasso (proposed by Tibshirani in 1996[11]) can produce accurate and sparse models via variable selection.

In 2003, Jolliffe & Uddin introduced SCoTLASS [12] to get modified principal components with possible zero loadings. And in the same year, a generalization of lasso, the elastic net, is presented by Zou & Hastie[13] to improve the lasso. In 2006, Zou et al. [10] put forward the SPCA algorithm which is based on the fact that PCA can be treated as a regression-type optimization problem whose regression criterion can be integrated by the elastic net directly. Next, we will review SPCA briefly.

As each principal component is a linear combination of the n variables, its loadings can be recovered by regressing the target principal component on the n variables. Let U_i denote the i-th principal component. $\forall\lambda>0$, the ridge estimates and the solution are given as follows:

$$\widehat{\beta}_{ridge} = \arg\min_{\beta} \left|Y_i - X\beta\right| + \lambda\left|\beta\right|^2$$

$$V_i = \widehat{v} = \frac{\widehat{\beta}_{ridge}}{\left|\widehat{\beta}_{ridge}\right|} \tag{4}$$

It is obvious that if $n>p$ and X is a full rank matrix, the theorem does not need λ to be positive. And if $p>n$ and $\lambda=0$, V_i will not be the unique solution of ordinary multiple regression. The same situation appears when $n>p$ and X is not a full rank matrix. But PCA can achieve a unique solution under any circumstances. After the addition of the L1 penalty, the optimization problem will be:

$$\widehat{\beta} = \arg\min_{\beta} \left|Y_i - X\beta\right|^2 + \lambda\left|\beta\right|^2 + \lambda_1\left|\beta\right|_1,$$

$$V_i = \frac{\widehat{\beta}}{\left|\widehat{\beta}\right|} \tag{5}$$

Then will be the estimate of V_i, and XV_i will be the estimate of the i-th principal component. After the addition of the lasso penalty into the criterion, the optimization problem will be:

Considering the first k principal components, let α and β

TABLE V. Sparse principal components and their variance contributions

Features	SPC1	SPC2	SPC3	SPC4	SPC5	SPC6	SPC7	SPC8
Z	0	0	0	0	0.6627	0.5158	0	0
Cs	0	0.07289	0	0.03758	0.115	0	0.02191	0.4809
Ls	0	0.07289	0	0.03758	0.115	0	0.02196	0.153
Rs	0	0	0	0.6229	0	0.4836	0	0
Cp	0	0.07289	0	0.03758	0.115	0	0.02191	0.4809
Lp	0	0.07289	0	0.03758	0.115	0	0.02196	0.1185
Rp	0	0	0	0	0	0	0	0
G	0	0.07283	0	0.03823	0.08497	0	0.1816	0
Y	0	0.07272	0	0.0395	0.02568	0	0.4969	0
θ	0	0.679	0.707	0	0	2.542*10^(-5)	0	0
D	0	0.07005	0	0.06821	0	0	0.4671	0
Q	0	0.04718	0	0.3142	0	0	0	0
ε′	0.684	0	0.01235	0	0	0	0	0
ε"	0.1794	0	0	0	0	0	0	0
Percentage (×100%)	0.8781	0.04617	0.05509	0.009619	0.005347	6.361*10^(-5)	0.0001049	8.693*10^(-6)
Cumulative percentage (×100%)	0.8781	0.9243	0.9794	0.989	0.9944	0.9944	0.9945	0.9945

denote a $p \times k$ matrix respectively. And X_i is the i-th row vector of X. Then $\forall \lambda > 0$,

$$\left(\hat{\alpha}, \hat{\beta}\right) = \arg\min_{\alpha,\beta} \sum_{i=1}^{n} \left| X_i - \alpha\beta^T X_i \right|^2 + \lambda \sum_{j=1}^{k} \left|\beta_j\right|^2 + \sum_{j=1}^{k} \lambda_{1,j} \left|\beta_j\right|_1$$

$$\alpha\alpha^T = I_k, \quad \hat{\beta}_i \propto V_i, i = 1, 2, ..., k \qquad (6)$$

If $\alpha = \beta$, the minimization under the orthogonal constraint on α is exactly the first k loading vectors of ordinary PCA.

In apple electronic signal data analysis, we use SPCA to select a set of base principal components with sparse loading to construct the new mapping space, which avoids mid-identifying the important variables and guarantees minimal information loss at the same time. Table V shows the sparse principal components and their variance contributions. It is obvious that these basis vectors are restricted to sparse loadings. In the next step, projections on these components will be computed and used as the new representation of apple quality.

C. *Grey Relational Analysis*

Grey Relational Analysis (GRA)[14] is developed as part of the grey system theory[15], which has been widely adopted for data analysis in various fields. As a method in grey system theory, GRA is introduced here specially for analyzing discrete data series. Its procedure involves the following steps.

(1). Generate the reference vector and the comparison vector selected from observations:

$$x_0 = \{a_{01}, a_{02}, ..., a_{0m}\}^T,$$

$$x_i = \{a_{i1}, a_{i2}, ..., a_{im}\}^T, i = 1, 2, ..., n \qquad (7)$$

(2). Compute the difference between x_0 and x_i:

$$\begin{aligned} \Delta_i &= \left| x_0 - x_i \right| \\ &= \{\left|a_{01} - a_{i1}\right|, \left|a_{02} - a_{i2}\right| \\ &\quad , ..., \left|a_{0m} - a_{im}\right|\}^T \end{aligned} \qquad (8)$$

(3) Get the global maximum difference value and minimum value in the whole difference vector:

$$\Delta_{\max} = \max_{\forall i}(\max \Delta_i),$$

$$\Delta_{\min} = \min_{\forall i}(\min \Delta_i) \qquad (9)$$

(4) Transform each data point in each difference vector to grey relational coefficient. The grey relational coefficient of the j-th data point in the i-th difference vector:

$$\delta_i(j) = \frac{\Delta_{\min} + \mu\Delta_{\max}}{\Delta_i(j) + \mu\Delta_{\max}} \qquad (10)$$

$\Delta_i(j)$ is the i-th value in the difference vector Δ_i. And μ is a coefficient between 0 and 1, used to compensate the effect of $\Delta_{\min}$ in case $\Delta_{\max}$ is an extreme value in the vector. In general, the value of μ is set to 0.5.

(5) Compute grey relational grade for each difference vector. Assume that data points in the series are of the same weights, and the grey relational grade for the *i-th* scale item will be:

$$\Gamma_i = \frac{1}{n} \sum_{i=1}^{n} \delta_i(m) \qquad (11)$$

In general, a scale item with a high value of Γ indicates that the respondents, as a whole, have a high degree of favored consensus on the particular item[16].

Fig. 3 (data for detail shown in Table VII in Appendix) displays the grey relational grades among features. For example, the value in the first row and the second column represents the grey relational grade between feature Z and feature Cs. And the value of the number shows the overall degree of standardized deviance of feature Cs from feature Z. An important characteristic of this table is that when feature ε" and feature ε′ serve as the reference data series, the grey relational grades of all the rest features are below 0.60. By contrast, when other features are taken as the reference data series, the grey relational grades of the rest features are around 0.90. Besides, the grey relational grades of the rest features are around 0.50 when feature Rp is the reference indicator. All these facts prove that the 3 electric signals (ε", ε′ and Rp) do not fit into the general tendency that other features show.

Generally internal quality features describe the trend of quality decline. But the 3 electric signals (ε", ε′ and Rp) are negative factors in terms of grey relational grade. Then it will be reasonable to remove them and evaluate the apple internal quality by other signals. Here we choose a mapping scheme to complete the experiment, and elaborate on it in detail in the next step.

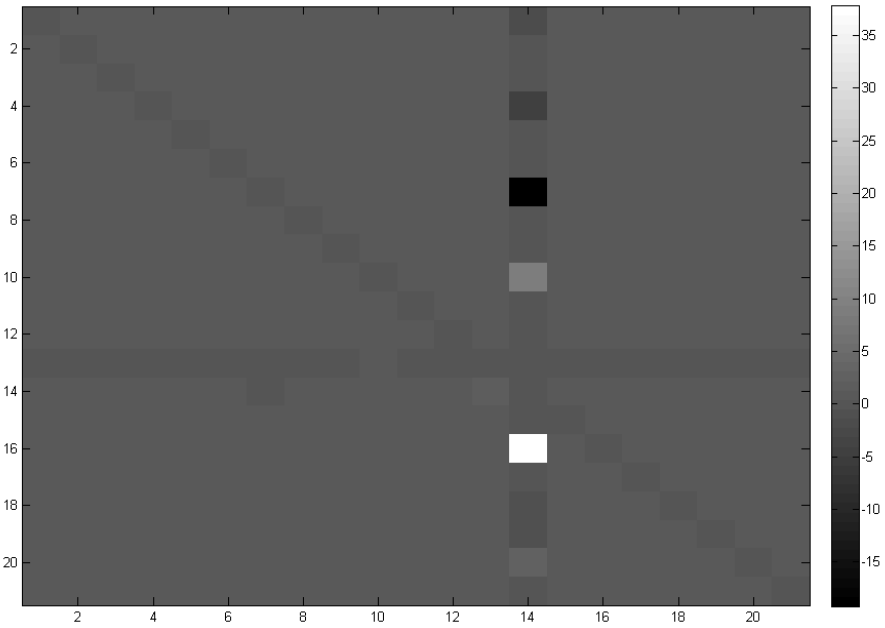

Fig. 3 The grey relational grades between each internal quality feature and others

TABLE VI. The classification performance in experiments of GRA

Number of basis vectors	1	2	3	4	5	6	7
Recognition Rates	0.705	0.72	0.77667	0.86833	0.905	0.89167	0.88833

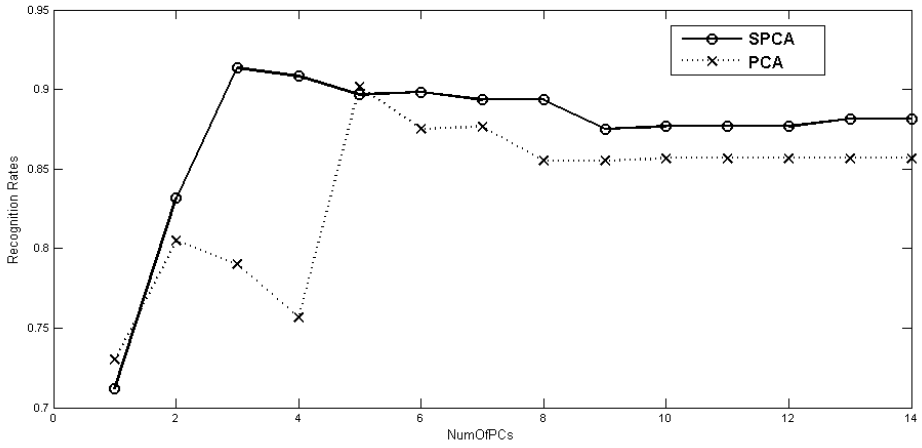

Fig.4 Comparison of the classification performance between experiments of PCA and SPCA

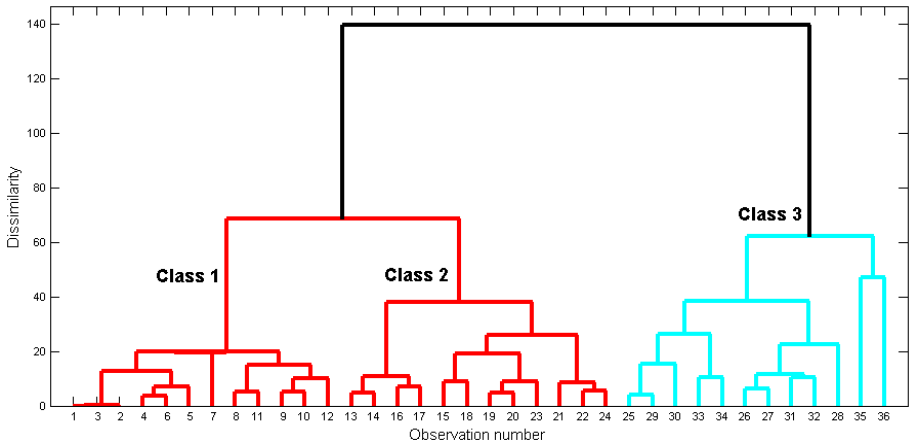

Fig.5 Classification made by Hierarchical Clustering when the distance measure of Hierarchical Clustering is Standardized Euclidean distance and the linkage criterion is the complete link

IV. Classification and Recognition

It is easy to classify apple samples by time of storage. But as time goes on, the quality of apples after being culled from trees is lowered day by day. Besides, different preservation techniques result in great difference of quality variation. And it is also possible that apples, either of different species or of the same one, possess diverse qualities and properties. For these reasons, clustering apples by quality instead of time would be a better way out.

It is widely accepted that internal quality features such as firmness and contents of Ethylene production rate can reveal apple quality. But an obvious disadvantage of clustering apples by quality is that internal quality features are difficult to measure without damaging the tested apples. So it will be convenient and efficient to measure apple quality just through electronic signals if we can use these signals to identify different classes. Then we introduce apple clustering based on internal quality features using hierarchical clustering.

Widely used in statistics, hierarchical clustering is aimed to build a hierarchy of clusters. There are generally two types of approaches to hierarchical clustering:

(1). Agglomerative: It is a "bottom up" strategy: each observation starts with its own cluster, and pairs of clusters are merged as one moves up the hierarchy.

(2). Divisive: It is a "top down" strategy: all observations start with a single cluster, and splits are performed recursively as one moves down the hierarchy.

As a rule, these merges and splits are determined in a greedy manner. The results of hierarchical clustering are usually presented in a dendrogram. Here we use the Agglomerative strategy to build the hierarchy of sample clusters.

For the purpose of deciding which clusters should be combined, Euclidean distance is used as the pair-wise similarity measure while Complete-link is used as the measure between two sets of samples. The apple quality categories defined by clustering are shown in Fig. 5.

To verify the variable selection made by all the techniques mentioned above, we use the Linear Discriminant Classifier (LDC) to identify the belonging of the cluster samples. We use a series of basis vectors extracted from the training set as the mapping space. For example, in PCA, we choose a number of principal components to compute the projections, which will reduce the electronic signal dimension. In SPCA, we also have the same basis vectors (Sparse Principal Component, SPC) to perform the transformation. In GRA, the gray relational grades of the 7 internal quality features express the same variable interpretation as principal components (See Table VII in Appendix), when the 14 electrical signals serve as reference. And they can be treated as basis principal components computed with special loadings. Then we introduce the classification methods and algorithms to perform signal selection or projection, and classifiers also given to categorize testing samples into the known clusters of the training set.

V. Results and Summaries

We selected the testing set from the original data set randomly, but kept the structure so that each class has at least one observation. We repeated the process of random selection and recognition 100 times, and computed the average recognition rate as the final result. The comparison of these three variable processing algorithms is shown in Table VI and Fig. 4. Here, 'recognition rates' means the ratio of the right-matched number to the total number of testing samples. In Fig. 4, the horizontal axis denotes the number of principal components or sparse principal components.

In detail, the highest recognition rate 0.9133 is reached by SPCA when the number of remaining SPCs is 3, the distance measure of Hierarchical Clustering is Standardized Euclidean distance, and the linkage criterion is the complete link. Fig. 5 shows the classification of Hierarchical Clustering in the same argument setting. In the three sets of experiments, we verify that these 3 methods are all workable in extracting features and SPCA is better than the other two in classification. According to our experimental results, when PCA and GRA are chosen as feature selection methods, their performances are acceptable. And the best scheme is to select features by means of SPCA.

VI. ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (60975007), the Natural Science Foundation of Shaanxi Province, China (2010JQ8019, 2010K06-15), the Special Foundation for Young Scientists of Northwest A&F University, China (Z11120822), the Special Foundation for Basic Scientific Research of Central Universities, China (QN2009093) and the Natural Research Foundation of Shaanxi Tobacco Company, China (K332021105).

APPENDIX:

TABLE VII. The grey relational grades between each internal quality feature and others

Comparison / Reference	Z	Cs	Ls	Rs	Cp	Lp	G	Rp	Y	θ	tanθ	Q	ε′	ε"	Firmness	TSS	Titirable acid	TSS/acid ratio	Ascorbic acid content	Respiration intensity	Ethylene production rate
Z	-	0.888	0.888	0.885	0.888	0.888	0.481	0.888	0.888	0.954	0.888	0.884	1.207	-1.644	0.887	0.885	0.888	0.818	0.874	0.872	0.885
Cs	0.909	-	0.930	0.928	0.930	0.930	0.511	0.930	0.930	0.995	0.930	0.926	1.179	-0.280	0.929	0.928	0.930	0.861	0.918	0.914	0.928
Ls	0.909	0.930	-	0.928	0.930	0.930	0.511	0.930	0.930	0.995	0.930	0.926	1.179	-0.280	0.929	0.928	0.930	0.861	0.918	0.914	0.928
Rs	0.894	0.916	0.916	-	0.916	0.916	0.501	0.916	0.916	0.981	0.916	0.912	1.188	-4.689	0.915	0.914	0.916	0.847	0.903	0.899	0.913
Cp	0.909	0.930	0.930	0.928	-	0.930	0.511	0.930	0.930	0.995	0.930	0.926	1.179	-0.280	0.929	0.928	0.930	0.861	0.918	0.914	0.928
Lp	0.909	0.930	0.930	0.928	0.930	-	0.511	0.930	0.930	0.995	0.930	0.926	1.179	-0.280	0.929	0.928	0.930	0.861	0.918	0.914	0.928
Rp	0.783	0.809	0.809	0.806	0.809	0.809	-	0.809	0.809	0.877	0.809	0.805	1.273	-19.195	0.808	0.806	0.809	0.738	0.793	0.794	0.807
G	0.909	0.930	0.930	0.928	0.930	0.930	0.511	-	0.930	0.995	0.930	0.926	1.179	-0.280	0.929	0.928	0.930	0.861	0.918	0.914	0.928
Y	0.909	0.930	0.930	0.928	0.930	0.930	0.511	0.930	-	0.995	0.930	0.926	1.179	-0.279	0.929	0.928	0.930	0.861	0.918	0.914	0.928
θ	0.979	0.996	0.996	0.996	0.996	0.996	0.560	0.996	0.996	-	0.996	0.992	1.144	8.772	0.996	0.995	0.996	0.930	0.986	0.979	0.993
tanθ	0.909	0.930	0.930	0.928	0.930	0.930	0.511	0.930	0.930	0.995	-	0.925	1.180	-0.251	0.929	0.928	0.930	0.861	0.917	0.913	0.927
Q	0.906	0.927	0.927	0.926	0.927	0.927	0.509	0.927	0.927	0.993	0.927	-	1.181	-0.088	0.927	0.925	0.927	0.859	0.915	0.911	0.925
ε′	0.323	0.346	0.346	0.341	0.346	0.346	0.161	0.346	0.346	0.393	0.346	0.344	-	-0.103	0.344	0.341	0.346	0.299	0.330	0.338	0.344
ε"	0.570	0.599	0.599	0.593	0.599	0.599	0.300	0.599	0.599	0.662	0.599	0.595	1.598	-	0.596	0.593	0.599	0.534	0.579	0.586	0.596
Firmness	0.903	0.924	0.924	0.922	0.924	0.924	0.507	0.924	0.924	0.989	0.924	0.920	1.183	0.369	-	0.922	0.924	0.855	0.911	0.908	0.922
TSS	0.896	0.917	0.917	0.915	0.917	0.917	0.502	0.917	0.917	0.983	0.917	0.913	1.187	37.770	0.917	-	0.917	0.848	0.904	0.901	0.915
Titirable acid	0.909	0.930	0.930	0.928	0.930	0.930	0.511	0.930	0.930	0.995	0.930	0.926	1.179	-0.259	0.929	0.928	-	0.861	0.917	0.913	0.927
TSS/acid ratio	0.845	0.869	0.869	0.866	0.869	0.869	0.468	0.869	0.869	0.935	0.869	0.865	1.222	-1.143	0.868	0.866	0.869	-	0.855	0.853	0.866
Ascorbic acid content	0.875	0.898	0.898	0.896	0.898	0.898	0.489	0.898	0.898	0.964	0.898	0.894	1.199	-1.023	0.897	0.895	0.898	0.828	-	0.882	0.895
Respiration intensity	0.901	0.922	0.922	0.921	0.922	0.922	0.506	0.922	0.922	0.988	0.922	0.918	1.184	2.182	0.922	0.920	0.922	0.853	0.910	-	0.920
Ethylene production rate	0.908	0.928	0.928	0.927	0.928	0.928	0.510	0.928	0.928	0.994	0.928	0.924	1.180	-0.214	0.928	0.927	0.928	0.860	0.916	0.912	-

REFERENCES

[1] Bennedsen B., Peterson D., Tabb A., "*Identifying defects in images of rotating apples*", *J. Computers and Electronics in Agriculture*, 2005, 2nd ed., vol. 48, pp. 92-102.

[2] Zou X., Zhao J., Li Y., et al. *"In-line detection of apple defects using three color cameras system"*, *J. Computers and Electronics in Agriculture*, 2010, 1st ed., vol. 70, pp. 129-134.

[3] Guo W, Nelson SO, Trabelsi S, et al. *"10–1800-MHz dielectric properties of fresh apples during storage"*, *J. Journal of Food Engineering*, 2007, vol. 83,pp. 562-569.

[4] Smith LI. "*A tutorial on Principal Components Analysis*". URL *http://rangerutaedu/~guerra/Smith02tutorialpdf* (URL accessed on May 29, 2011), 2002, pp. 1-26.

[5] Belhumeur P., Hespanha J., Jman D. *"Eigenfaces vs. Fisherfaces: recognition using class specific linear projection*", *J. IEEE Transaction on Pattern Analysis and Machine Intelligence*, 1997, 7th ed., vol. 19, pp. 711-720.

[6] Wold S, Esbensen K, Geladi P. *"Principal component analysis" J. Chemometrics and Intelligent Laboratory Systems*, 1987, vol. 2, pp. 37-52.

[7] You M, Cai C. *"Wood classification based on PCA, 2DPCA, (2D)^{2}PCA and LDA", J. 2009 2nd International Symposium on Knowledge Acquisition and Modeling*, KAM 2009, vol. 1, pp. 371-374.

[8] Golub GH, Reinsch C. *"Singular value decomposition and least squares solutions" J. Singular value decomposition and least squares solutions*, 1970, 5th ed., vol. 14, pp. 403-420.

[9] Hotelling H. *"Analysis of a complex of statistical variables into principal components", J. Journal of Educational Psychology*, 1933, vol. 24, pp. 498-520.

[10] Zou H, Hastiey T, Tibshirani R. "*Sparse Principal Component AnalysisJ. Journal of Computational and Graphical Statistics"*, 2006, vol.15, pp. 265-286.

[11] Tibshirani R. *"Regression shrinkage and selection via the lasso"*, *J. Journal of the Royal Statistical Society*, 1996, 1st ed., vol. 58, pp. 267-288.

[12] Jolliffe T, Trendafilov NT, Uddin M. *"A Modified Principal Component Technique Based on the LASSO", J. Journal of Computational and Graphical Statistics*, 2003, 3rd ed., vol. 12, pp. 531-547.

[13] Zou H, Hastiey T, Tibshirani R. *"Regularization and variable selection via the elastic net", J. Journal of the Royal Statistical Society*, 2005, 2nd ed., vol. 67, pp. 301-320.

[14] Deng J. "*The Introduction to Grey System Theory*", *J. The Journal of Grey System*, 1989, 1st ed., vol. 1, pp. 1-24.

[15] Tsai C., Chang C., Chen L. *"Applying Grey Relational Analysis to the Vendor Evaluation Model"*, *J. International Journal of The Computer, The Internet and Management*, 2003, 3rd ed., vol. 11, pp. 45-53.

[16] Lin J., Lin C. "*The use of the orthogonal array with grey relational analysis to optimize the electrical discharge machining process with multiple performance characteristics" J. International Journal of Machine Tools and Manufacture*, 2002, 2nd ed., vol. 42, pp. 237-244.