Jurnal Penyelidikan dan Pendidikan Kejuruteraan. Jilid 1, 2004

Labview™ for Nutra-Biostrip in Herbal Quality Assessment

MOHD NOOR AHMAD [1*], MAXSIM YAP MEE SIM [2], MOHD KAMAL BIN MOHAMED RAMLY NIL [2], ALI YEON MD SHAKAFF [3], CHANG CHEW CHEEN[2]

[1] School of Material Engineering, KUKUM, 02600, Kangar, Perlis, Malaysia
[2] School of Chemical Sciences, Universiti Sains Malaysia, 11800 Penang, Malaysia
[3]School of Computer & Communication Engineering, KUKUM, 02600, Kangar, Perlis, Malaysia

ABSTRACT

In this work, we introduce the approach on the development of a stand-alone laptop based data acquisition of an array sensor system, namely Nutra-BioStrip coupled with pattern recognition algorithm for herbal quality assessment. The array sensor system control program, developed in LabView 6.1 programming languages allow data acquired from the array sensor to be analyzed by means of Principal Component Analysis (PCA) and displayed in the form of an interactive two-dimensional cluster mapping with detail statistical analysis results for rapid and real-time herbal quality assessment.

Keywords: *Herbal analysis; Array sensor; Pattern recognition; Principal Component Analysis*

INTRODUCTION

The use of computers for data acquisition, data analysis and data presentation in the field of sensor technology has increase rapidly in the last several years. This increased has paralleled the availability of user-friendly graphical programming languages which allow rapid program development and an execution speed comparable with more conventional languages such as PASCAL or C. LabView is a powerful and versatile graphical programming environment that was developed primarily to facilitate data acquisition and analysis since the software becomes the actual "instrument" (the so called "virtual instrument" or VI) [1,2]. So instead of dedicated devices or instruments, the personal computer, equipped with a multi-function I/O card and suitable software is turned into a flexible instrument capable of performing a variety of task, i.e. acquiring data, statistical data analysis, and data storage.

In this paper, a stand-alone laptop based data acquisition of an array sensor system, namely Nutra-BioStrip coupled with pattern recognition algorithm developed in the LabView environment is reported for herbal quality assessment. As a high level graphics-oriented language, LabView that comes together with a library of mathematical subroutines and other utilities, provides programming solutions that is less tedious compared other programming languages. Its programming ease, fast in data acquisition and analysis execution, modularity, flexibility, expandability and the simplicity in handling very large vectors and matrices are some of the reasons its' being incorporated in the system. The report will be based on description of the program development for data acquisition and data analysis namely Principal Component Analysis (PCA) and finally the projection of herbal analysis results in the developed program.

The sensor array system generates data of high densionality, which is hard to handle and visualize. However, as PCA is a linear feature-extraction technique, which finds new directions in the pattern space and explains the maximum amount of variance in the data set as possible, it is a method that have been proved to be effective for discriminating responses in array sensor especially in the application of complex system i.e. herbals. Although a lot of statistical software for PCA analysis are readily available in the market, we developed our own software for the sensor system as the aim of this work is to develop a computerized portable sensor system with stand-alone program that supports multiple tasking capable of performing rapid and real-time herbal analysis which is intended for laboratory and field use. Besides, it is also more economical in such a way that we have flexibility in upgrading or modified the program, which suits our experimental application that changes from time to time without any required charges.

The increasing demands for traditionally used herbal products worldwide have prompted healthcare professionals and researchers to offer new ways of assessing quality, efficacy and safety. However, since herbal medicines with a large number of chemical components are very complex, the most current approach in herbal validation and standardization is herbal analysis of the whole herb (holistic approach) instead of its components. Standardization by conventional analytical methods such as GC, HPLC & TLC refers to standardizing herb to its marker compound (with separation of constituents). Besides time & cost consuming, the problem with this type of standardization is that although it may identify a particular herb to the consumer, it does not mean that the compound to which the herb is standardized is the one responsible for the herb's desired effect. A plant's benefit is a result of the combined activity among the many active components known to be responsible for a particular herbal used (synergistic effects vs. reductionism approaches). Hence, the advent of the Nutra-BioStip sensor system, together with pattern recognition algorithm is highly desired for real-time quality assessment of herbs.

EXPERIMENTAL

Equipment

The interface card used for data acquisition was a NI PCMCIA-6024E low-cost multifunction DAQ card featuring 16 single-ended or 8 differential gain programmable analog inputs multiplexed to a 12-bit ADC, two analog outputs based on 2 12-bit DACs, 8 digital I/O lines and a 2 24 bit counters/timers. In DMA mode, the analog sampling frequency extends up to 200kHz, and the analog signal output up to 10kHz. The card was interface to a Pentium 4 processor, 256MB Ram running LabView 6.1 under windows XP. LabView is an object-oriented software-programming tool using a graphical programming environment. Each program (so called VI) is composed of two levels: 1) the front panel which is the graphical user interface (GUI) contain controls for input operations, and; ii) the block diagram in which the actual programming code is structured by interconnecting icons representing operators, values and actions [3]. which suits our experimental application that changes from time to time without any required charges.

The BioStrip, which is the array sensor strip (Figure 1) was fabricated using screen-printing technology. The detail fabrication procedure is as reported in the previous studies [3].

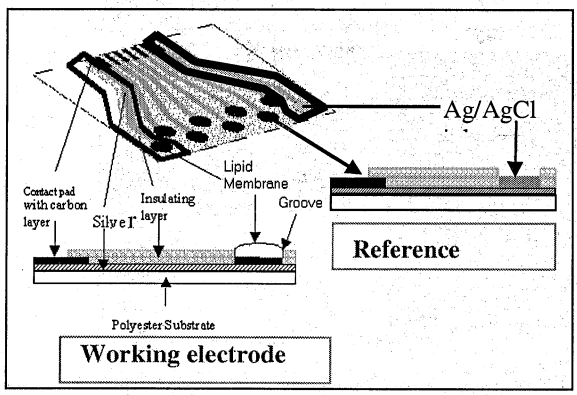

Figure 1: Cross- sectional view of BioStrip

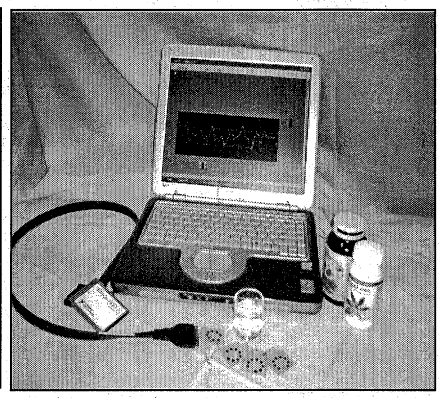

Figure 2: Nutra-BioStrip sensor system.

Software Configuration

Data Acquisition System

The Biostrip sensor is connected to the DAQ card with a 2.54 mm pitch standard connector attached to the 1 meter NI R6868 Ribbon I/O Cable (see Fig.2). The Biostrip sensor array is an integration of 8 working electrodes and a reference electrode together on a single strip based on screen-printing technology. The 8 working electrodes are connected to eight Analog Input channel and the reference node is grounded (single ended analog input). The DAQ card is plugged-in to the laptop's PCMCIA slot. NI PCMCIA-6024 is chosen as it is a low-cost data acquisition card which uses E Series technology to deliver high-performance and with a fast sampling rate, which is up to, 200kS/s. NI Measurement & Automation Explorer (MAX) is used to access and configure the DAQ card, which connects to the system.

Eight channels are used for the acquisition in which each channel acquires 500 samples with 2000 scans/s and a -1v to 1v input signal range. Over sampling and averaging is applied to reduce noise in the analog measurements. This can be easily done with LabView by using the AI Acquire Waveforms VI and the Mean VI. However, readings acquired from the BioStrip are low in voltage. Therefore, it has been changed to the unit of mV. Data collected from real-time data acquisition can then be saved into the computer hard disk as a spreadsheet file together with information of parameters such as time and sample name. The front panel for data acquisition of the Nutra-BioStrip sensor system is illustrated in Figure 3. This particular panel displays functions like time interval setting, sensor response, sensor readings, 8 different sensor channels which can be selected, and basic controls like quit, stop, start and save data function.

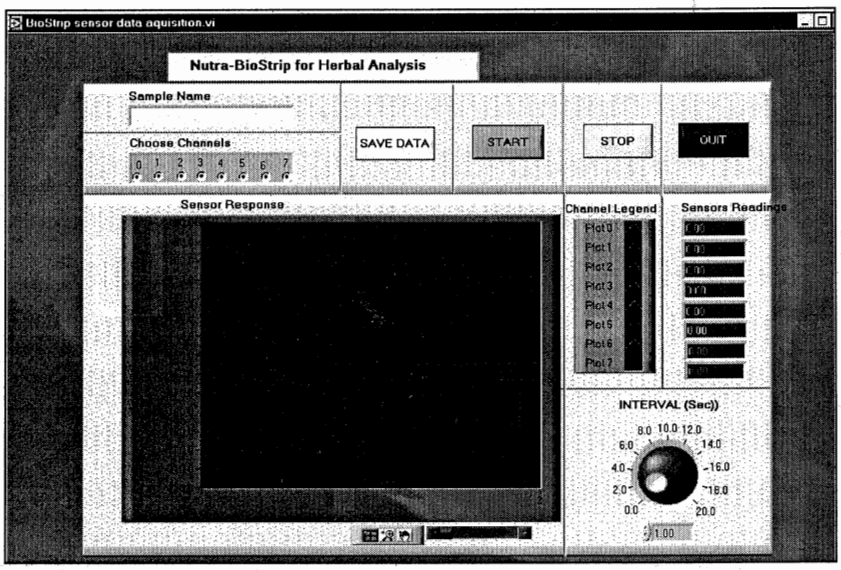

Figure 3 : BioStrip sensor data acquisition VI front panel.

Principal Components Analysis (PCA)

The programming task for this part appeared to be much more challenging as the application deals with large matrices and vectors. However, LabView Linear Algebra Mathematics functions SubVIs somehow makes the calculation easier to execute.

Principal Components Analysis (PCA) is a multivariate projection method, which extracts and highlights systematic variation in a data matrix to give us an overview of dominant pattern and major trends in the data. PCA transforms the original set of variables into a smaller set of linear combinations that account for most of the variance of the original set. The purpose is to determine factors (i.e. principal components) in order to explain as much of the total variation in the data as possible with as few of these factors as possible [4,5].

Besides, PCA is also an unsupervised method and is used to decide whether a set of pattern classify naturally into groups or not. Therefore it is able to highlight some clusters without having any prior knowledge of the classes to be expected and it proves the performance of the sensor system. The result is expressed in score plots, where trends, groups and outliers among the measurements can be observed, and loading plots, where relevance and similarity between variables can be seen [6].

First of all, the spreadsheet data, which has been saved from data acquisition, will then be processed in the Microsoft Excel environment. Data set which needed to be analyzed using PCA should be arranged in the manner of 8 Columns (8 different channel of sensor output) x Rows (Depending on the number of samples feed-in). Then, the data will be saved as text file (Tab delimited).

This text (Tab delimited) file data will then be imported into the Program for analysis. In most instances we are dealing with measurements collected at one point of time on n individuals for p>1 variables or characteristics. We will use the notation X_{ij} to indicate the reading for the *i*th object on the *j*th variable. Alternatively, it can be written as

$$\mathbf{X}= (X_{ij}),\ i=1, 2,\ldots, n,\ j=1, 2,\ldots, p.$$

The analysis continues to the next level where the notation will be used. **Figure 4** shows the block diagram where the calculation is executed and continued to the **Figure 5.**

After that, centroid is calculated from the data. Centroid is the row vector of the means of **X**, denoted by $\overline{\mathbf{X}}$ and computed by the following matrix operation:

$$\overline{\mathbf{X}}' = \frac{1}{n}\mathbf{1}'\mathbf{X}$$

Where **1**′ denotes a 1 × n unit vector.

The mean corrected scores can be obtained once $\overline{\mathbf{X}}$ has been found. Denoting by $\mathbf{X}_d$ the n×p matrix of mean corrected scores, we have

$$\mathbf{X}_d = \mathbf{X} - \mathbf{1}'\overline{\mathbf{X}}$$

It is now an easy task to compute the variance s^2, an estimator of the population variance σ^2, in matrix terms.

$$S^2 = \frac{1}{n-1}\mathbf{X}_d'\mathbf{X}_d$$

Once the variances have been found, we can place in a diagonal matrix, **D**. Standardized data can be easily obtained by

$$\mathbf{X}_S = \mathbf{X}_d\mathbf{D}^{-1/2}$$

Where $\mathbf{D}^{-1/2}$ is the inverse of the square root of **D**. Or,

$$\mathbf{X}_S = \mathbf{X}_d\sigma^{-1}$$

Where σ is standard deviation that square root of variance.

It is easy to obtain the standard deviation and variance by using LabView as the function of Standard Deviation and Variance are readily exist in sub VI. Then, the mean corrected sums-of-squares and cross-products matrix, which is denoted by **S** is obtained. The **S** is obtained by working with the mean corrected data matrix $\mathbf{X}_d$;

$$\mathbf{S} = \mathbf{X}_d'\mathbf{X}_d$$

Once the mean corrected sums-of-squares and cross products matrix **S** has been found, it is easy to obtain the variance matrix, denoted here by **C**. Where

$$\mathbf{C} = \frac{1}{n-1}\mathbf{S}$$

Mean while, the correlation matrix **R** is obtained from pre- and post multiplying **S** by $\mathbf{D}^{-1/2}$;

$$\mathbf{R} = \frac{1}{n-1}(\mathbf{D}^{-1/2}\mathbf{S}\ \mathbf{D}^{-1/2})$$

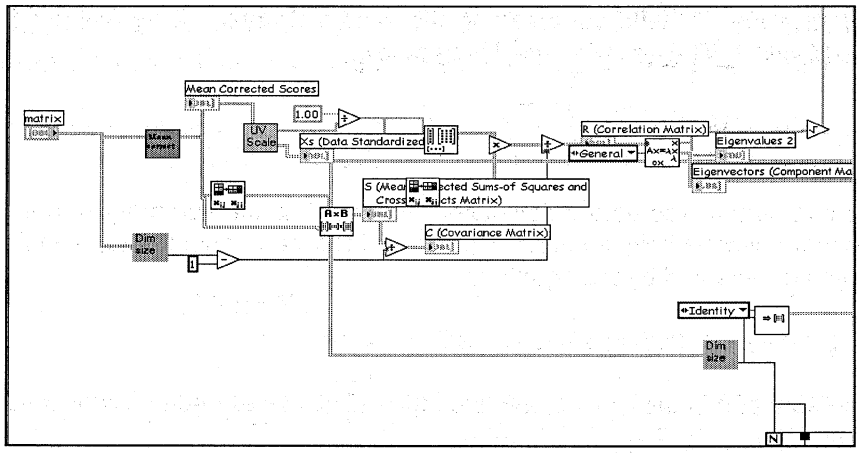

Figure 4. The block diagram of the Principal Components Analysis (PCA) VI. (continue with **Figure 5**)

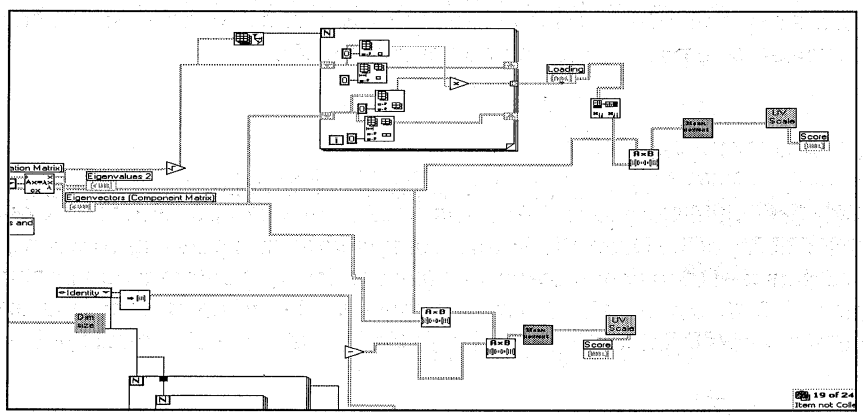

Figure 5. The block diagram of the Principal Components Analysis (PCA) (Continuation from **Figure 4**)

The toughest part in Principal Components Analysis is to extract the components. We can generate Principal Components from two types of data input: (1) a variance-covariance matrix and (2) a correlation matrix. In this LabVIEW program, we use the correlation as the data input ($\mathbf{R}_{ij}$). First, the eigenvalues and eigenvectors are obtained from the correlation matrix. The equation used here is the p simultaneous linear equations.

$$(\mathbf{R}-\lambda_{(1)}\mathbf{I})\gamma_{(1)} = 0$$

where $\lambda_{(1)}$ is the largest eigenvalue of $\mathbf{R}$, and $\gamma_{(1)}$ is the corresponding eigenvector. The problem now become: maximize $\gamma'\mathbf{R}\gamma$, with respect to γ, subject to $\gamma'\gamma = 1$. The γ coefficients must satisfy the p simultaneous linear equations.With LabView, this equation can be easily executed by using the Complex EigenValues & Vectors (Advanced Only) function. After obtaining the eigenvalues and eigenvectors, Loadings, **P** is calculated. Loadings are the

weights (influence of the X-variables on the scores Complex EigenValues & Vectors (Advanced Only) **T**. The equation used is as follow:

$$\mathbf{P} = \gamma_{ij}\sqrt{\lambda_j}$$

Then the component scores are calculated. Scores, **T** is the summary of the original X variables that describe how the different rows in **X** (observations) relate to each other. The scores are obtained by the equation:

$$(\mathbf{X}s \times \mathbf{P}')$$

T = mean corrected and standardized of the

This is the equation being used in this program. Besides, Scores can also be calculated this equation:

$$\mathrm{Y} = \left(\mathrm{I} - \frac{1}{\mathrm{n}}\mathrm{E}\right)\mathrm{X}\,\mathrm{A}$$

where **X** is the n×p data matrix, **E** is the n×n matrix with one in every position, and **A** is the p×r matrix whose columns are the first r eigenvectors of **S.**

RESULTS AND DISCUSSION

In this study, Principal Component Analysis (PCA) has been performed to examine the discriminative ability of the Nutra-BioStrip sensor system in distinguishing between i) different types of extracts for *Orthosiphon Stamineus* (cat's whisker) ii) different geographical origin for samples of *Orthosiphon stamineus* and finally iii) the different parts and extracts of *Labisia pumila* (kacip fatimah). Figure 6 illustrates the data set being displayed in the Data Table Vi.

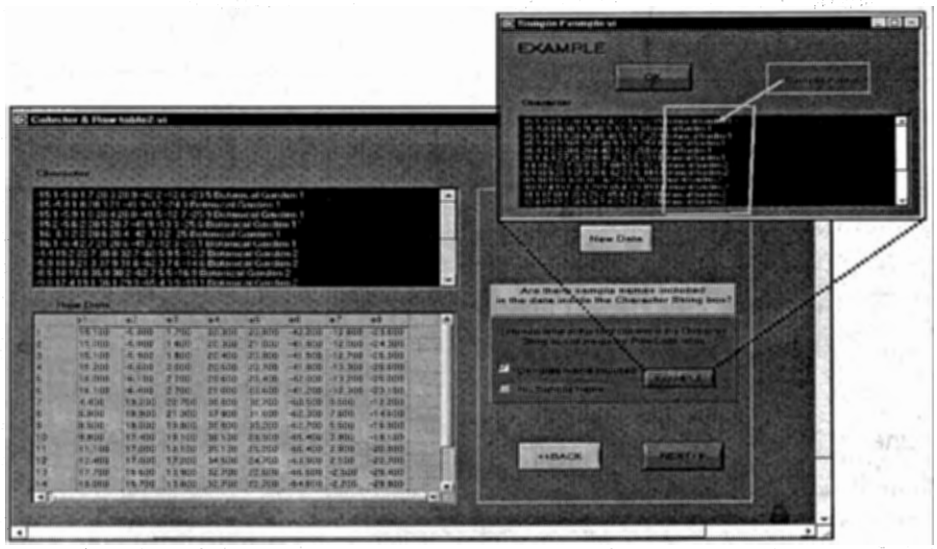

Figure 6. The Front Panel of the Character & Data Table Vi.

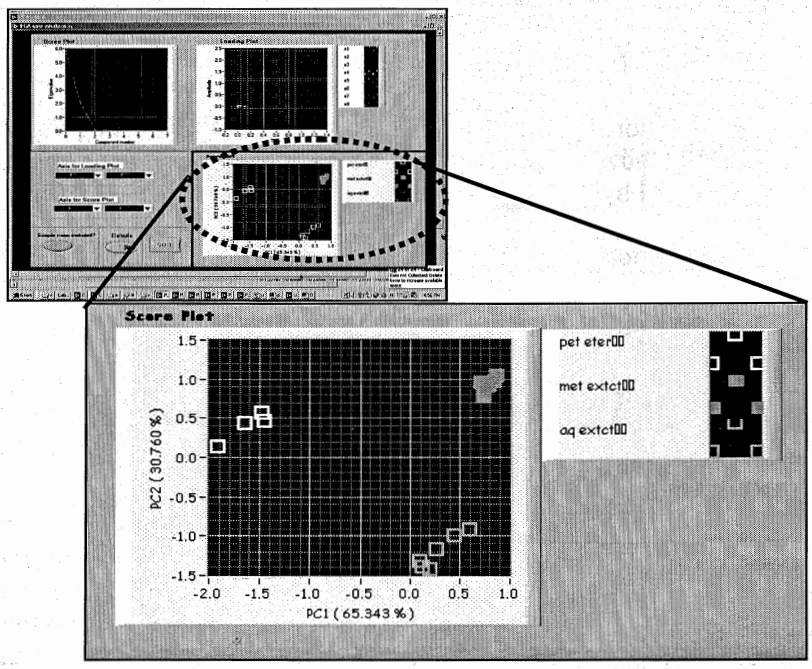

Figure 7. The front panel of the PCA .user interface VI. It shows the scree plot for three different types of *Orthosiphon stamineus* extracts, loading plot and the score plot.

The PCA score plot in **Figure 7** shows the two-dimensional PCA analysis of different extracts of *Orthosiphon stamineus* using the first and second principle component. The scores are divided moderately into three groups with the first two components PC1 (65.34%), PC2 (30.76%) explaining 96.10% of the total system variance. This result show that the BioStrip effectively discriminate the different types of extraction i.e. water extract, methanol extract and pet-ether extract.

In conducting quality control of herbal products, one of the main important tasks is to establish the true botanical identity of the raw herb material. The source and quality of raw materials play a pivotal role in guaranteeing the quality and stability of herbal preparations. Even though herbals come from the same species, the quality and efficacy are somewhat different according to the growing conditions based on geographical origin. However, it is not easy to determine the geographical origin based on the analytical tools as well as through visible inspection. There are at least a number of major components, which are slightly different according to growing conditions, and we cannot select only several specific components as essential criteria [8]. Therefore, Nutra-BioStrip sensor system can be an excellent candidate for the particular purposes.

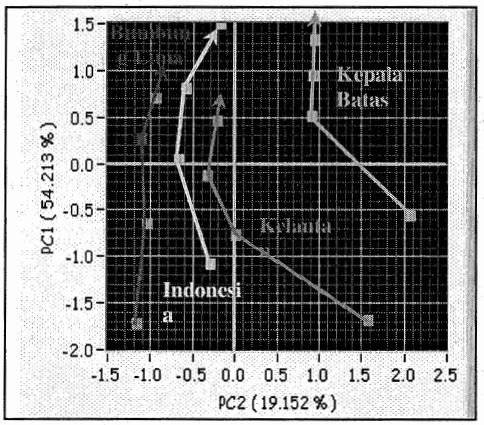

Figure 8. PCA score plot for the different geographical origin of *Orthosiphon staminues* aqueous extrac

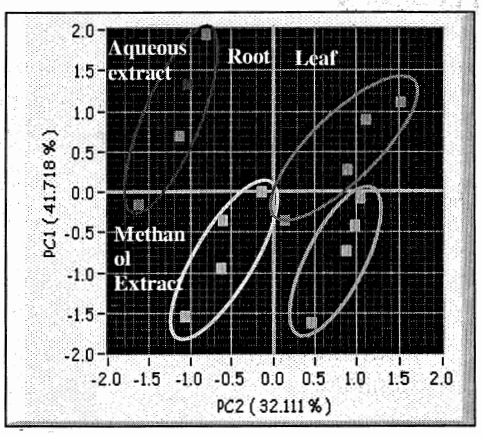

Figure 9. PCA plot for *Labisia pumila* parts and extraction identification from the Nutra-BioStrip sensor system

Figure 8 presents the PCA plot discriminating between four different geographical origin of the *Orthosiphon stamineus* sample in aqueous extract, i.e. Kepala Batas, Kelantan, Indonesia and Bumbung Lima, with the first and second principle component explaining 73.37% of the total variance. As can be observed, there is a clear distinction between the four different samples where PC2 distinguishing the four respective sample origins into four well define space. The PC1 axis is the axis representing changes in concentration with the arrows showing the transition shift from lower to higher concentration of the aqueous samples prepared. The results shown demonstrated that the subtle differences in chemical composition between plant varieties, which are difficult to interpret by physical observation, could easily be determined using array sensor incorporating pattern recognition principle.

As shown in **Figure 9**, the PCA plot discriminating different parts & extracts of the *Labisia pumila* samples. PC1 and PC2 together containing 73.83% of the variance of all data set. A good separation between all kinds of sample category is observed with PC1 discriminating the two different aqueous and methanol extracts while PC2 distinguishing the different parts category with the "root" category being on the negative region of PC2 and 'leaf" on the positive region.

CONCLUSION

Quality control and standardization begin with a well-defined plant material through a regulated extraction procedure. Inadequate quality control allows adulteration and often toxic products to reach the market, which can directly jeopardize public safety. The methodology presented here provides a useful starting point for those interested in developing a computerized based data acquisition array sensor system, incorporating pattern recognition algorithm in the LabView environment for rapid and real-time sample analysis specifically for application in herbal quality control. Future plans involved incorporating additional chemometric algorithm in the currently developed software system using Linear Discriminant Analysis (LDA), Principal Component Regression (PCR) and Partial Least Square (PLS) for a more detailed qualitative and quantitative analysis.

REFERENCE

1. J.H.Moore. (1995). Artificial Intelligence Programming With Labview: Genetic Algorithms for Instrumentation Control and Optimization, *Computer Methods and Programs in Biomedicine,* 47 : 73-79.

2. A.Economou, S.D.Bolis, C.E. Efstathiou, G.J.Volikakis. (2002). A "Virtual" Electroanalytical Instrument for Square Wave Voltammetry. *Analytica Chimica Acta*, 467 :179-188.

3. A.S.A.Rahman, M.M.S.Yap, A.Y.Md Shakaff, M.N.Ahmad, Z.Dahari, Z.Ismail, M.S.Hitam, (2004). A Microcontroller-Based Taste Sensing System for the Verification of Eurycoma longifolia, *Sensors & Actuators B*, 101:191-198.

4. William R. Dillon and Mathew Goldstein. (1984). *Multivariate Analysis Methods and Applications*. NY: John Wiley and Sons.

5. Edmund R. Malinowski. (2002). *Factor Analysis in Chemistry,* Third Edition. WILEY-INTERSCIENCE.

6. G.Polder, G.W.A.M. Van der Heijden, I.T.Young. July (2002). Hyperspectral Image Analysis for Measuring Ripeness of Tomatoes, *An ASAE Meeting Presentation:1-6.*

7. *LabVIEW Basics 1 Course Manual*. National Instruments Corporation.

8. Y.A.Woo, H.J.Kim, J.H.Cho, H.Chung. (1999). Discrimination of Herbal Medicines according to Geographical Origin With Near Infrared Reflectance Spectroscopy and Pattern Recognition Techniques. *Journal of Pharmaceutical and Biomedical Analysis*, 21 :407-413.