I J C T A, 9(2) 2016, pp. 1381-1391

Principal Component Analysis for Harmonic Separation in Electric Signal with Hardware Implementation

G. Malathi *, T. Venkatesan and S. Ravi*****

ABSTRACT

A serious concern in the quality of power delivered to the commercial and industrial networks is the harmonic contents in the electrical system due to substantial penetration of non linear-loads. In order to enhance the power quality in modern power system, harmonic content level needs to be identified and separated. Harmonic separation is critical in power quality aspects due to non-optimal effects in power system and electrical apparatus. In this paper, a fast, simple and effective Principal Component Analysis method to separate the harmonic from the mixed input signal is presented. The present method has an intrinsic ability to detect and separate a number of independent harmonic signals from the random mixture. By separating harmonic signals as individuals, the power quality delivered by the power system connected to various types of non-linear loads is improved. The Principal Component Analysis algorithms are developed and tested in MATLAB/SIMULINK and the results confirms effectiveness of the present approach for power quality improvement. An ex perimental prototype using Raspberry-pi hardware is also presented.

Key words: Harmonic separation, MATLAB, Power quality, Principal Component Analysis, Raspberry-Pi, Total Harmonic Distortion

1. INTRODUCTION

The large scale use of non-linear power electronics loads in the modern power distribution system causes many power quality problems like high current harmonics, low energy efficiency, low power factor, malfunctioning of sensitive devices and so on. Several investigations [1-4] to quantify the problems associated with electric power networks with non-linear loads have been carried out. In general, non-linear loads draw non-sinusoidal currents that contain harmonic currents. These harmonic currents interact with impedance of the power distribution system and create voltage distortion that can affect both the distribution system equipment and the load connected to it. Thus, it is important to develop a suitable compensating device that can avoid negative consequences of harmonics problems. This paper is aimed to apply an efficient mathematical technique known as principal component analysis to deal power quality problems due to harmonic issue.

Principal Component Analysis (PCA) possesses much of the active filter characteristics intrinsically in its harmonic separation process for improving power quality. It is a powerful statistical technique in data analysis to transforms a set of highly correlated zero mean random variables into a small number of de-correlated variables called principal components [5-6]. The components are usually ordered according to decreasing variance. Among the various starting points and optimization criteria for PCA derivation, the most important are minimization of the mean square error in data compression, finding mutually orthogonal directions in the data having maximal variances, and de-correlation of the data using orthogonal

* Research Scholar (PT), Faculty of EEE, Anna University, Chennai-600 025, India, *Email: ursermala@gmail.com*

** Dept. of Electrical and Electronic Engineering, K.S. Rengasamy College of Technology, Tiruchengode-637 215, India.

*** Dept. of Electrical and Communication Engineering, Dr. MGR Educational and Research Institute University, Chennai-600 095, India.

transformations. It can be used for compression of large amount data without loss of information and can extract relevant information by projecting the original data sets into the new orthogonal space **[7]**. Hence, PCA has been recognized as promising statistical tool in recent years for many areas like face recognition, image processing, rock analysis and power quality analysis [8-10]. PCA application to Power quality problems has been successfully proposed by many authors [11-12]. Okan Ozg onenel et al.[13] have developed PCA based hybrid algorithm for power quality monitoring and ensured the capability of this algorithm for detection and the classification of the distorted power quality waveforms. This paper models the problems of harmonic separation using PCA. The result shows that the method presented in this paper has real time implementation feasibility.

2. PCA BASED ALGORITHMS

The PCA based algorithms for the present investigation are given as follows.

2.1. PCA based harmonics separation algorithm

i. **Mixture signal:** The mixing of signals is done in the simulation with basic signals.

ii. **Remove mean:** Initially, the mean value is removed from the sample (i.e., mean = 0). The training data can be amplitude-scaled in basically two ways: so that the values of the patterns lie between -1 and 1, or so that the values of the patterns lie between 0 and q. These two types of amplitude scaling are usually referred to as min/max scaling. The function “premnmx” is available in the MATLAB Neural Network toolbox [ ] to scale in the input data in the range [-1, 1]. The “premnmx” function has an option to scale the input data or both the input and target data. There is another important scaling process called mean centering and variance scaling. Mean centering and variance can be carried out separately or together. In the MATLAB Neural Network toolbox [ ], the function “prestd” will mean-center and variance-scale both for the input data and the target data (or occasionally) just the input data.

iii. **Pre-whitening:** This process normalizes the variance of the observed signals to unity and provides better stability properties and converges faster. The input vector $X(k)$ are whitened by applying the transformation

$$V(k) = VX(k) \tag{1}$$

where $V(k)$ is the k^{th} whitened vector and V is the whitening matrix.

iv. **Separation and estimation:** The signals that were correctly mixed are separated into individual original signals using PCA separation algorithm.

v. **Correlation:** This is used to determine the phase relation between the recovered and the initial signal. A negative sign denotes 180° phase shift.

vi. **180° Phase shifter:** The PCA can occasionally recover an antipodal signal and in such occasions this block is used to recover the original phase.

2.2. PCA based whitening algorithm

Here, the standard PCA based approach is used for whitening. This process can simultaneously compress information optimally in the mean-square error sense and filter possible noise. The following steps are used in the PCA based pre-whitening process.

i. Remove the mean value from the composite readings $s[k]$

ii. Form the covariance matrix of the observed signal $E\{s_k s_k^T\}$.

iii. Extract the matrices D and E, Where, $D = \text{diag}[\lambda(1), \ldots\ldots, \lambda(M)]$ is a $M \times M$ diagonal matrix, $E = [c(1), \ldots\ldots., c(M)]$ is a $L \times M$ matrix, $\lambda(i)$ is the i_{th} largest eigenvalue of the data covariance matrix $E\{s_k s_k^T\}$, and $c(i)$ is the respective i_{th} principal eigenvector.

iv. Form whitening matrix V using $V = D^{-1/2} * E^T$ and pre-whiten the inputs using $v[k] = V*s[k]$

3. METHODOLOGY

The methodology of implementation involved in this paper is discussed as follows.

3.1. Mixing of signal

i. Include built in library modules for user space linking in runtime and initialize the variables.

ii. The mixed signal (i.e. fundamental + harmonics) is stored (by acquiring the samples online) in a file. The separation algorithm reads this data along with the mixing matrix (A) using "fgetc" function.

In step (ii), the mixing signal includes

a. the file containing only the signal (to be separated)

b. the file containing constant (and is random) values represented as matrix "A" and

c. file containing possible frequencies that have been mixed and need to be separated

iii. Read the mixed signal and convert the mxn matrix using user defined function

$$X(k,:) = X1(1,((k-1)*\text{no. samp}) + 1: (k*\text{no.samp})).$$

iv. Declare this variable as persistent *nu*.

3.2. Processing the mixed signal

i. Using the "A" mxn matrix, mix the signals with A and remove the mean. This is required so that the correlation and covariance of the mixture is same.

ii. Form a Kron matrix by taking all the possible products between the elements of x and y.

iii. The userspace program when hardcored for running directly in hardware without the help of developer platform, requires certain functions to be used for cross compiling and compatibility. This includes

a. str_to_ double(s, i) to convert string into double

b. double, use str_to_num(s, i) to convert string into number

3.3. Pre-whitening

i. Calculate average of the co-variance matrix and find its eigenvalues and eigenvectors.

ii. Sort the eigenvalues in ascending order and extract the diagonal elements of the matrix.

iii. Rotate the matrix by 180° in anticlockwise direction in order to arrange it in column wise. Restore it in matrix form using "diag ( )" function.

iv. Arrange the eigenvector matrix in indexed matrix form.

v. Use "fliplr ( )" function to preserve row and flip column in left/right direction.

vi. Calculate the whitening matrix.

vii. Calculate the whitened input vector and scale it by dividing by sqrt(200).

3.4. Generation of weight matrix

i. The trained or initial value of the weights are read (using fgetc function) and converted into a form containing only rows.

ii. The transformation of this weight matrix into a mxn matrix is done again using the user defined function, $W(k, :) = W1(1, ((k-1)*p) + 1:(k*p));$. This variable is again declared as persistent *nu*.

3.5. Largest Eigen value based dimensional reduction

In this paper, the feature vectors in the N-dimensional space are projected to a m-dimensional subspace such that the spanned space is formed by the eigenvectors of the covariance matrix corresponding to the m largest eigenvalues.

This method of dimensional reduction ensures

(i) Maximizing the sum of the variances of the components

(ii) Minimizing the Mean Square Error (MSE) between an *N*-dimensional vector and the "*m*" dimensional projection of it.

The *N*-dimensional random vector "*x*" is approximated as

$$\hat{x} = \sum_{i=0}^{m-1} y_i e_i + \sum_{i=m}^{N-1} c_i e_i \tag{2}$$

where c_i are nonrandom constants and e_i, $i = 0, 1, 2, N\text{-}1$, constitute an orthonormal basis. Based on (i) and (ii) above, the minimum MSE is $E = \|x - \hat{x}\|^2$ and is guaranteed with the eigenvectors of $\sum_x$ forming the orthonormal basis and e_i for $i = m, \ldots, N\text{-}1$, for correspond to the *N-M* smallest eigenvalues. Let

$$x = \sum_{i=0}^{N-1} x_i e_i, x_i = \langle x_i, e_i \rangle = e_i^T x$$

The projection in the m-dimensional subspace spanned by, say, $e_{i,}$, $i = 1, 2, \ldots, m$, is

$$\hat{x} = \sum_{i=0}^{m-1} x_i e_i \text{ and } x - \hat{x} = \sum_{i=m}^{N-1} x_i e_i \text{ (or)}$$

$$E = \|x - \hat{x}\|^2 = E\left[\sum_{i=m}^{N-1} \sum_{j=m}^{N-1} x_i x_j e_i^T e_j\right] = E\left[\sum_{i=m}^{N-1} x_i^2\right]$$

$$E = \|x - \hat{x}\|^2 = \sum_{i=m}^{N-1} E[e_i^T x x^T e_i] = \sum_{i=m}^{N-1} e_i^T \operatorname{Re}_i \tag{3}$$

Thus, the problem is now to minimize the above, subject to the constraint

$$e_i^T e_i = 1, \; i = m, \ldots, N-1$$

Hence, using Lagrange multipliers, this is equivalent with minimizing

$$Q = \sum_{i=m}^{N-1} e_i^T \operatorname{Re}_i + \sum_{j=m}^{N-1} \lambda(1 - e_j^T e_j)$$

$$\frac{\partial Q}{\partial e_i} = 2(\mathrm{R}e_i - \lambda_i e_i) = 0 \Rightarrow \mathrm{R}e_i - \lambda_i e_i,\ i = m,\ ...,\ N-1$$

Hence, e_i must be eigenvectors of R. Substituting the above in the cost function we obtain

$$E[\|x-\hat{x}\|^2] = \sum_{i=m}^{N-1} \lambda_i \tag{4}$$

Since $\lambda_i > 0$ then the minimum value is achieved if $\lambda_m, \ldots\ \lambda_{N-1}$ are the smallest eigenvalues. The sum variance of its components, assuming zero mean values, is

$$\sum_{i=0}^{m-1} E\left[x_i^2\right] = \sum_{i=0}^{m-1} e_i E\left[xx^T\right] e_i = \sum_{i=0}^{m-1} \lambda_i \tag{5}$$

which is obviously maximized for the above choice.

In this case, the error is

$$x - \hat{x} = \sum_{i=m}^{N-1} \left(x_i - c_i\right) e_i \text{ (or)}$$

$$E[\|x-\hat{x}\|^2] = \sum_{i=m}^{N-1} E\left[\left(x_i - c_i\right)\right]^2 \tag{6}$$

The gradient of the above with respect to c_i is

$$2E\left[(x_i - c_i)\right] = 0 \Rightarrow c_i = E\left[x_i\right]$$

Hence,

$$E\left[\|x-\hat{x}\|^2\right] = \sum_{i=m}^{N-1} E\left[(x_i - E\left[x_i\right]^2\right]$$

$$E\left[\|x-\hat{x}\|^2\right] = \sum_{i=m}^{N-1} e_i^T E\left[\left(x_i - E\left[x_i\right]\right)\left(x_i - E\left[x_i\right]\right)^T e_i\right]$$

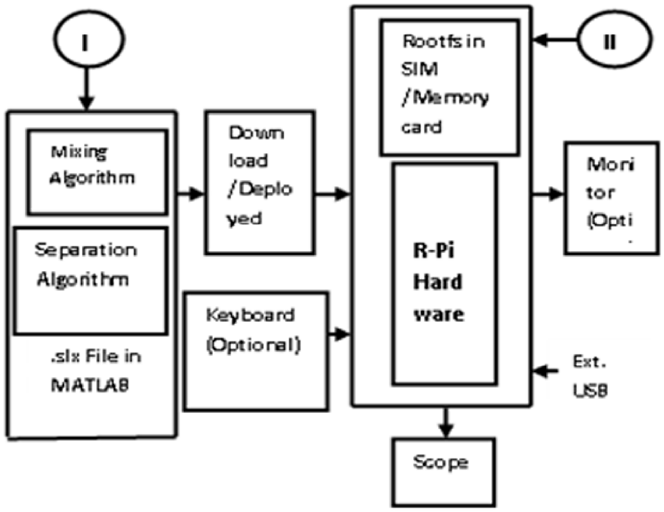
i.e.,

$$E\left[\|x-\hat{x}\|^2\right] = \sum_{i=m}^{N-1} e_i^T \sum e_i \tag{7}$$

4. HARDWARE IMPLEMENTATION

The Raspberry-Pi (R-Pi) primarily intended to run Linux kernel-based operating systems, offers a capable affordable prototyping platform for a wide range of products, with onboard Ethernet, plus a bunch of available General Purpose Input/output (GPIO) pins. The blocks used to build the hardware is shown in Figure. 1(a). In Module-1, a model for mixing and separation of signals is done separately using MATLAB/ SIMULINK. This model is deployed in to R-Pi hardware in Module-II using R-Pi package which support in MATLAB 2013b.

The deployment of model provides executable .o files. The executable files (mixing and separation) is called using application program (C-Program) which runs on R-Pi board separately. The user input data are stored in “in.txt” file and mixed and separation of signals model use this file for further processing. The



(a) **(b)**

Figure 1 (a): Blocks used in hardware implementation **Figure 1 (b): Experimental prototype**

mixed and separated signals spectrum is displayed in CRO using audio-port of the R-Pi. The experimental prototype shown in Figure. 1(b) includes R-Pi hardware.

5. RESULTS AND DISCUSSION

The observed three signals are generated from $x(k) = As(k)$, where A is the mixing matrix for $k = .02$ to 20. This represents a set of three instantaneous mixtures. The observations are pre-whitened using batch whitening process. To perform separation, non linear PCA subspace learning rule is used. A set of random initial weights was selected from a Gaussian distribution (with zero mean and unit variance) and then the columns of the weight matrix were ortho-normalized. The correlation coefficient is computed for each of the separated signal with respect to the known source signal. The correlation of

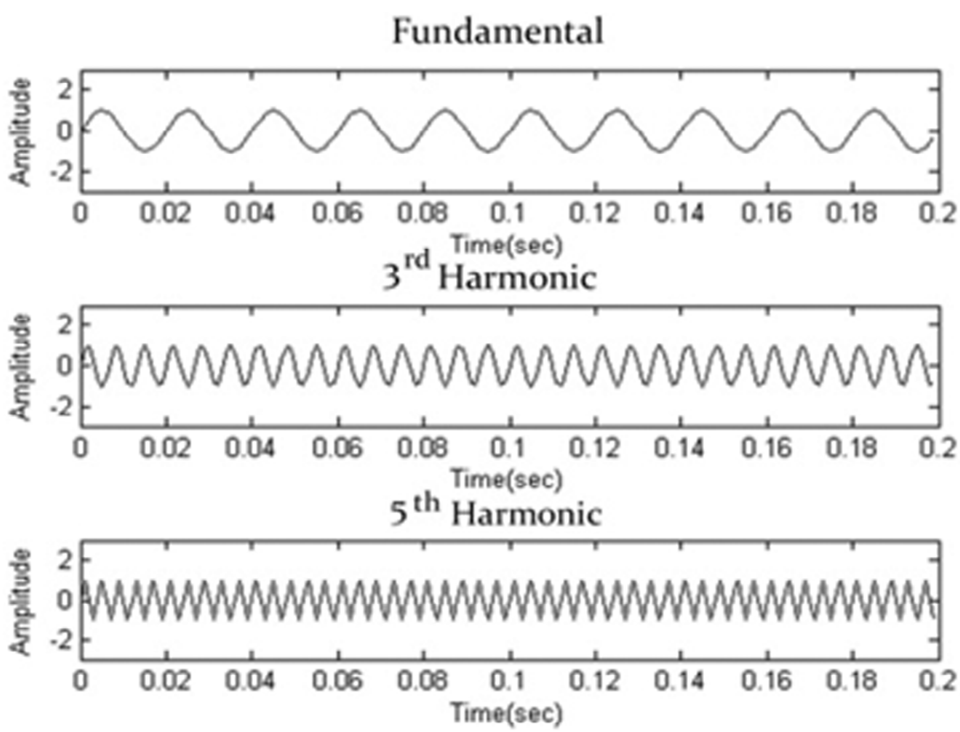


Figure 2: Input signal (Fundamental + Harmonics)

the separated signal with respect to actual (known) signal is perfect. A negative correlation coefficient infers a 180° phase shift occurred in the separated output. Figure.2 represents the three original sinusoidal source signals of fundamental 50 Hz, 3rd harmonic 150 Hz and 5th harmonic 250 Hz that are sampled at f_s = 500 Hz.

The separated output (using the PCA separation algorithm) is shown in Figure 3 with the FFT output of input signal in Figure 4.

The accuracy of the separated result can be identified with help of spectrum analysis as shown in Figure 5 (a-c)

Figure 5(a-c)**.** FFT of separated signals

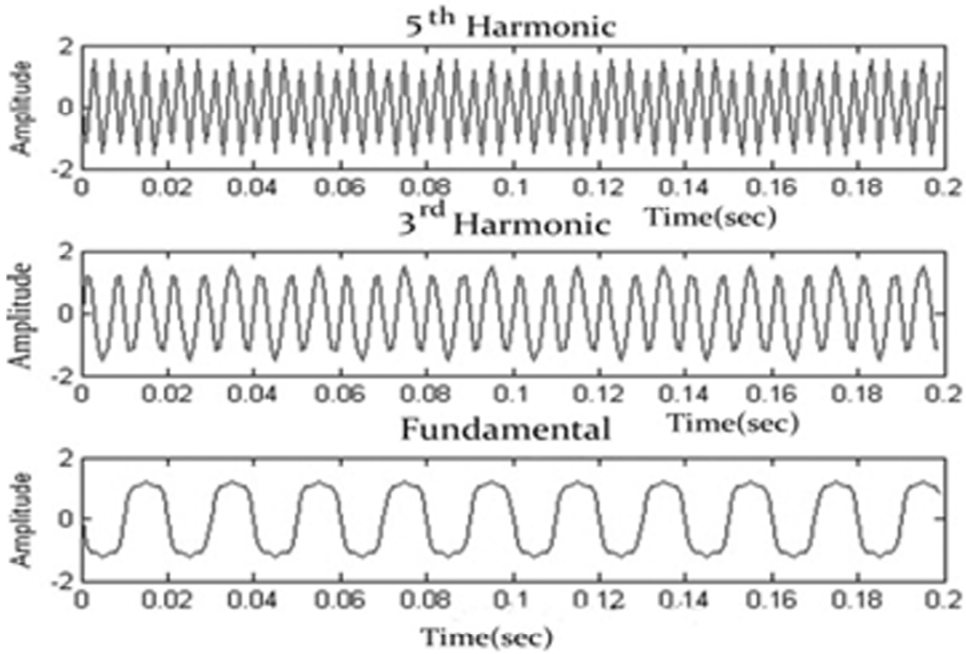


Figure 3: Separated signals

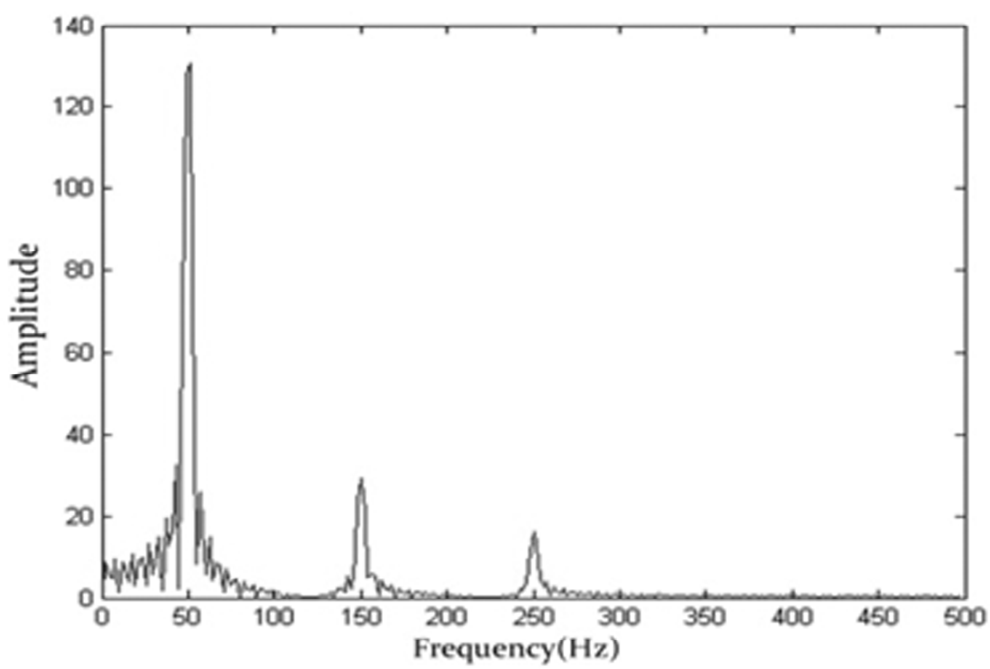


Figure 4: FFT of input signal (Fundamental more prominent but harmonics present)

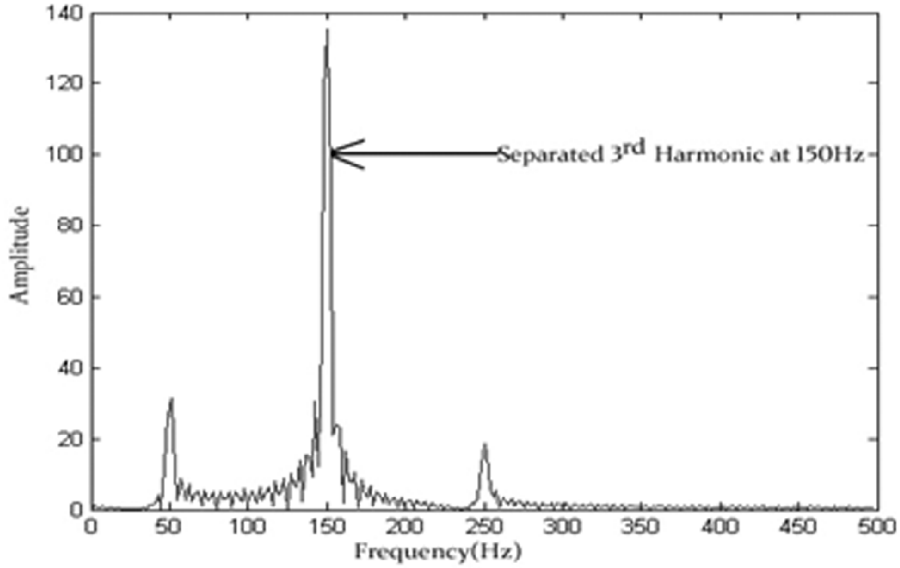


Figure 5 (a): Separation of 3rd Harmonics at 150Hz

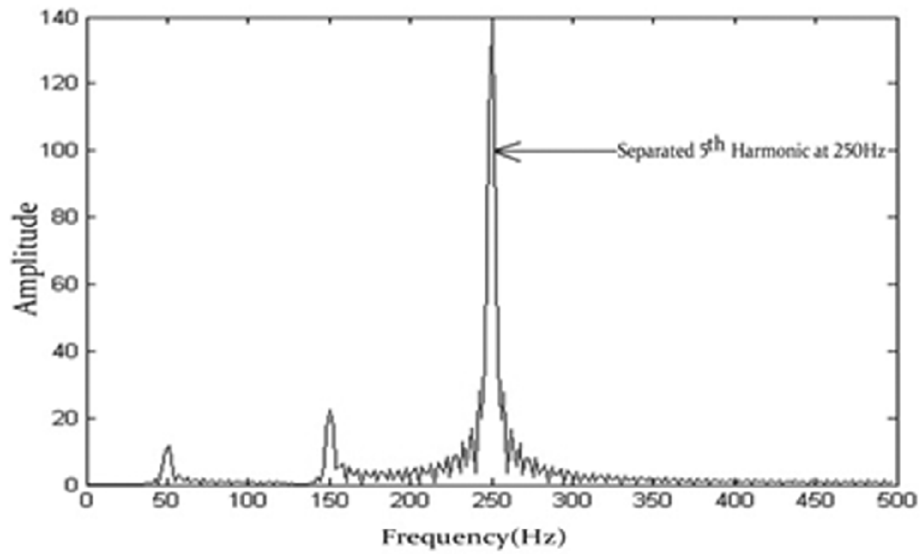


Figure 5 (b): Separation of 5th harmonics at 250Hz

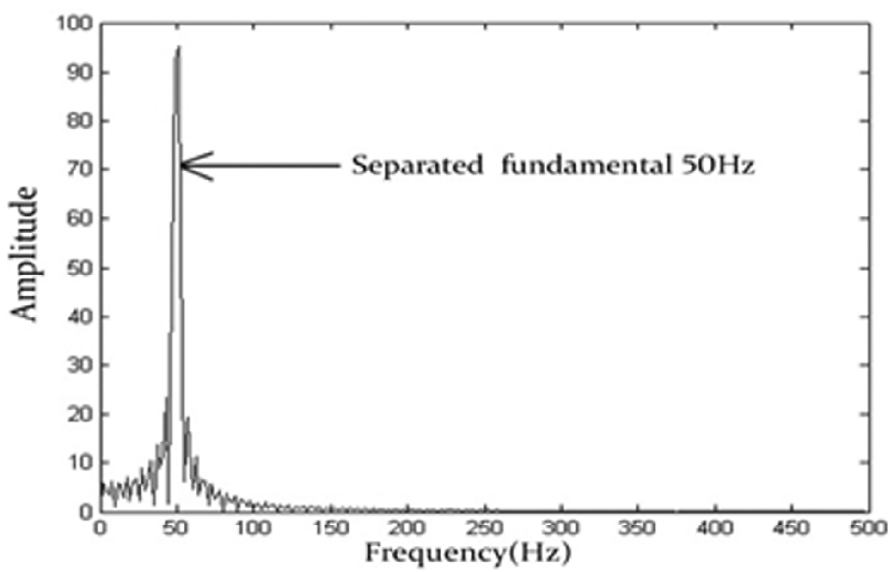


Figure 5 (c): Separation of fundamental 50Hz

5.1. Screen captured result

The screen captured results are displayed in Figure 6(a-e).

i. Interactive Menu Display using the application (Written in C language) (Figure 6(a))

ii. Input Signal Sample Size Selection (Figure 6(b))

iii. For Illustration, 3 signals of frequency 50Hz,150Hz and 250Hz (fundamental and harmonics) are added in Hardware (Figure 6(c))

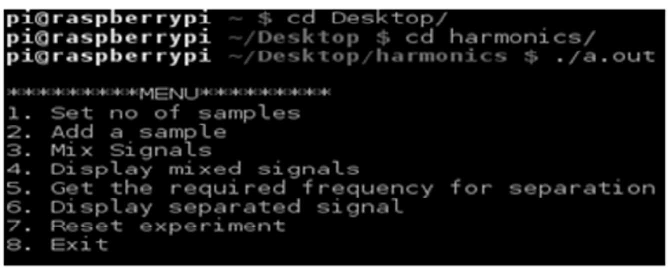
iv. Figure. 6(c)

v. Input signals are mixed successfully in Hardware (Figure 6(d))

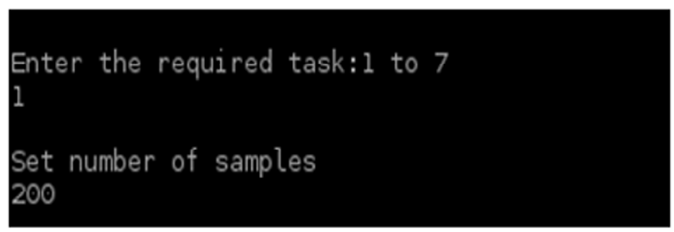
vi. Performing Blind Separation in Hardware for the specified frequency and signal separation is done successfully (Figure 6(e))

5.2. Calculation of Positive Predictive values

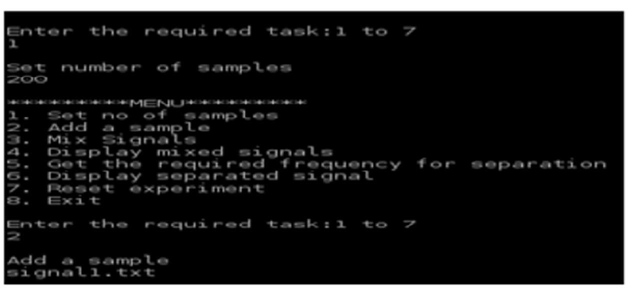
Positive Predictive Value (PPV) evaluates the feasibility or the success of a PCA application for harmonic separation of a signal. It is defined as



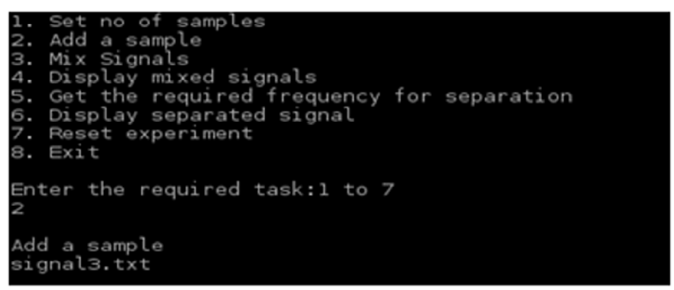
(a)



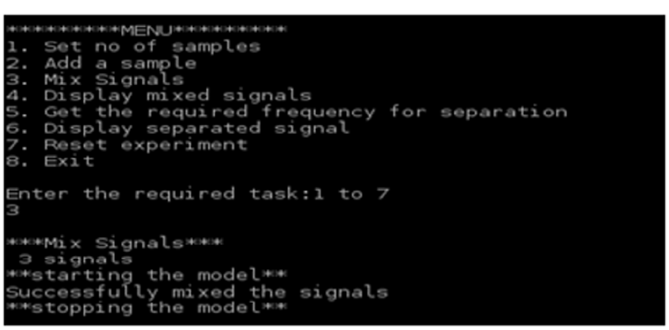
(b)



(c)(i)



(c)(ii)



(d)

(e)

Figure 6 (a-e): Screen captured results

$$PPV = \frac{True\ Positive}{True\ Postive + False\ Postive}$$

Where a “true positive” is the event that the PCA algorithm makes a positive prediction and the signal has a positive result under the gold standard, and a “false positive” is the event that the algorithm makes a positive prediction, and the signal has a negative result under the gold standard.

Table 1 shows the PPV value of a signal separated from a harmonic content mixed signal using PCA technique under different conditions.

Table 1
PPV of separated signals

Separated signals	*True Positive*	*False Positive*	*PPV*
Fundamental	95	8	92.3
3^{rd} Harmonics at 150Hz	140	50	73.9
5^{th} Harmonics at 250Hz	140	35	80.0

This results shows that the PCA technique based filter is highly efficient to separate fundamental frequency (50Hz) and to improve the power quality of the signal.

5.3. Calculation of Total Harmonic Distortion

The Total Harmonic Distortion (THD) is an indicator of the distortion of a current signal. Table 2 shows the current amplitude of a signal separated from a harmonic content mixed signal using PCA analysis.

Table 2
Amplitudes of fundamental and harmonics current measurement data

Order	1	3	5
Amplitude	136	6	3

The harmonic current distortion rate can be calculated by

$$THD = \frac{\sqrt{I_3^2 + I_5^2}}{I_1} X\ 100\% = 4.93\%$$

Thus, it validates that harmonics are effectively separated by using PCA algorithm proposed in this paper.

6. CONCLUSION

The harmonics present in the power system is best removed using PCA algorithm. This algorithm is performed in standalone hardware (R-Pi) and of active type and can effectively remove the different harmonics present in the signal. The achieved THD is acceptable with respect to IEEE norms and allows the hardware unit to be connected inline to effectively separate the harmonics. Since our PCA based solution is expected to take corrective measure for maintaining better power quality from utility and end user, the financial impacts on the country’s economy due to poor power quality power is overcome.

REFERENCES

[1] H. Akagi, “Active Harmonic Filters,” Proceedings of the IEEE 2006,Vol. 93, No. 12, pp. 2128-2141, 2005.

[2] M.J.H. Rawa, D.W.P. Thomas, M. Sumner, “Simulation of Non-linear Loads for Harmonic Studies,” 11th International Conference on Electrical Power Quality and Utilisation (EPQU) 2011, IEEE, pp. 1-6, 2011.

[3] Panafiotis Dimitriadis, Mohamed Darwish, “Harmonic Correction in Power Supplies Feeding Non-Linear loads,” 49th International conference on Power Engineering Conference (UPEC) 2014, IEEE, pp.1-6, 2014.

[4] Wenbo Su, Zhaoxiang Li, Zhenguo Shao, “The Harmonic Calculation Model of Non-linear Load,” China International Conference on Electricity Distribution (CICED) 2014, pp. 866-872, 2014.

[5] Jonathon Shlens, “A Tutorial on Principal Component Analysis,” Institute for Nonlinear Science, University of California, San Diego, 2005.

[6] D. Stathis, D. Myronidis,” Principal Component Analysis of Precipitation in Thessaly Region (Central Greece),” Global NEST Journal, Vol. 11, No. 4, pp. 467-476, 2009.

[7] C.Y. Chang, H.R. Hsu, “Application of Principal Component Analysis To a Radial-Basis Function Committee Machine for Face Recognition,” International Journal of Innovative Computing, Information and Control, Vol. 5, No. 11(B), pp. 4145-4154, 2009.

[8] Jonglak Pahasa and Issarachai Ngamroo, “PSO Based Kernel Principal Component Analysis and Multi-Class Support Vector Machine for Power Quality Problem Classification,” International Journal of Innovative Computing, Information and Control; Vol. 8, No. 3(A), pp.1523-1539, 2012.

[9] Iis Hamsir Ayub Wahab, Adhi Susanto, P. Insap Santosa, Maesadji Tjokronegoro, “Principal Component Analysis Combined With Second Order Statistical Feature Method for Malaria Parasites Classification,” Journal of Theoretical and Applied Information Technology; Vol.63, No. 1, pp. 219-225, 2014.

[10] J. L. Scealy, Patrice de Caritat, Eric C. Grunsky, Michail T. Tsagris & A. H. Welsh “Robust Principal Component Analysis for Power Transformed Compositional Data,” Journal of the American Statistical Association, Vol. 110, Issue 509, pp.136-148, 2015.

[11] Y.Y Hong, B.Y Chen, “Locating Switched Capacitor Using Wavelet Transform and Hybrid Principal Component Analysis Network,” *IEEE* Trans. on Power Delivery; Vol. 22, No. 2, pp. 1145-1152, 2007.

[12] Chuan He, and Qin Shu, Separation and Analyzing of Harmonics and Inter-Harmonics Based on Single channel independent component analysis, Int. Trans. Electr. Energ. Syst. Vol. 25, 169–179, 2015.

[13] Okan Ozgonenel, Umit Kemalettin Terzi, Abdesh Khan, “A hybrid approach for power quality monitoring.” Turk J Elec Eng & Comp Sci., Vol. 20, No. 6, pp. 854-869, 2012.