International Journal of IT-based Public Health Management
Vol. 3, No. 2 (2016), pp. 7-14
http://dx.doi.org/10.21742/ijiphm.2016.3.2.02

Respiratory Signal Analysis using PCA, FFT and ARTFA (A Generalized Comment)

Varun Gupta and Monika Mitta[1]

NIT, Kurukshetra

Abstract

Sinus patient are increasing day by day in the world, may be human as well as animals. That's why today signal analysis has been the need to know the diseases in the patient. Biomedical signal processing has great importance in the life of every human and animals. In absence of biomedical signal processing (BSP) signals cannot be analysed resulting in failure of disease acknowledgment. In this paper we have analysed the respiratory signals of Sinus and Normal Person using principal component analysis (PCA), Fast Fourier Transform (FFT) and Autoregressive Time-Frequency Analysis (ARTFA). PCA is used to derive a relatively small number of decor related linear combinations (principal components) of a set of random zero-mean variables while also retaining the signal information as much as possible. Autoregressive time-frequency analysis allows us to follow the changes in frequencies involved in the signal through time. For this we require to see the frequency changes in time. FFT examine the signal in frequency domain that calculates the spectral function. This paper describes the application of principal component analysis (PCA), a technique allowing the reduction of the data set dimensionality. In this paper we have calculated the variance of First Principal Component and Second Principal Component in Sinus and Normal Person and these are 86.94%, 13.05% and 92.733%, 7.266% respectively.

Keywords: *Biomedical signal processing, Principal Component Analysis (PCA), Autoregressive Time-Frequency Analysis (ARTFA), Fast Fourier Transform (FFT), variance.*

1. Introduction

Principal component analysis (PCA) is a technique that is generally used for reducing the dimensionality of multivariate datasets. Considering a vector of n random variables x for which the covariance matrix is £, the principal components (PCs) can be defined by [14][15].

$$z = Ax \quad \textbf{(1)}$$

Where z is the vector of n PCs and A is the n x n orthogonal matrix with rows that are the eigenvectors of £. The eigenvalues of £ are proportional to the fraction of the total variance accounted for by the corresponding eigenvectors.

FFT is the fast way to compute Discrete Fourier Transform (DFT).The amplitudes andphases of the sinusoidal components can be estimated using the DFT and is represented mathematically as

* ZltllK
x (K) = - 2 JflJ x (n) e ““ST- **(2)**

For a given biomedical signal x(n)(may be airflow and respiration) whose sampling period is T with N number of total samples (NT is therefore the total duration of the signal segment).

Article history:
Received (March 21, 2016),Review Result (May 07, 2016), Accepted (June 28, 2016)

ISSN: 2205-8508 (Print), eISSN: 2207-3965 (Online) IJIPHM

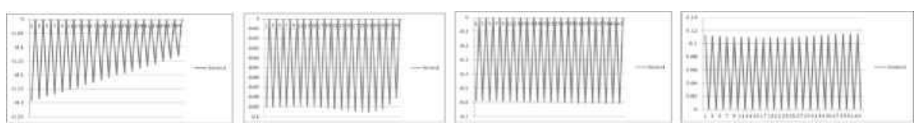
Thespectrum X(k) is estimated at multiples of -, where fs is the sampling frequency [22].It is well known, that an analysis based on an AR model offers the advantage of automatic spectral decomposition, based on a residual integration algorithm (Mainardi L.T. et al., 1995).The AR model is part of the ARMA system family together with the Moving Average (MA) model, and the Autoregressive Moving Average (ARMA) model. Other parametric method is the Prony method applicable for the analysis of transient signals [20]. The AR model can also be represented as to predict the signal sample based on q past values of the signal weighted by constant coefficients. We estimate the best model by trying to minimize the mean squared error (MSE) between the signal sample predicted by the model and the actual measured signal sample. In an AR model, the signal x(n) is represented in terms of its prior samples as follows [22]:

$$x[n] = e[n] - a1\, x[n-1] — a2\, x[n-2] — \bullet \bullet \bullet \ldots aqx[n-q] \quad \textbf{(3)}$$

Where assumed to be zero mean white Gaussian noise with a variance of .q = order of the AR model, x [n — i] = signal sample i time periods prior to the current sample at n, at coefficients or parameters of the AR model

2. Acquisition of fata

In this paper we have done the analysis of respiratory signal through Principal component analysis .Analysis through Empirical Mode decomposition is still running. First of all we are acquiring the data through MP36 BIOPAC machinery and after that we are analyzing the signal through BIOPAC acknowledge software.



(a) (b) (c) (d)

Figure 1. Selected air flow and respiration signals of sinus patient(left side a, b) and

Figure 2. Selected air flow and respiration signals of normal person (right side c ,d)

3. Analysis of Acquired Data

The analysis of acquired data is done through simply the mathematical expressions such as eigen values and eigen vectors [4]. Let, A = n x n matrix. The scalar 1 is an eigenvalue of A if there exists a non-zero vector v such that,

$$Av = 1v \quad \textbf{(4)}$$

Where Vector v is called an eigenvector of A corresponding to eigenvalue 1. For each eigenvalue 1, the set of all vectors v satisfying Av = 1v is called eigen space of A corresponding to eigenvalue 1.We can rewrite the condition Av = 1v as,

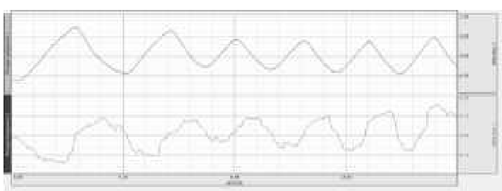
$$(A- 1I)\, v = 0 \quad \textbf{(5)}$$

Frequency analysis (FFT) permits us to interpret which frequencies are involved in our biomedical signal/signal. The problem is the great length of the data/signal. If we transform the whole signal then we may interpret approximately which frequencies the signal contains. Peaks are visible if we use short time period. FFT require the samples of the data being analyzed/interpreted to be evenly spaced in time. The result is divided by the number of samples in the window. Because the FFT spectrum of the rectangular window is sin c function having sinusoidal decaying ripples, FFT X (f can be a distorted version of the original spectrum of biomedical signal x(n). Spectral contents from one frequency component tend to leak into neighbouring frequency components due to the convolution summation operation. That's why it is recommended to window the signals. Many windows, such as Hamming, Hanning, Kaiser Bessel, Blackman, Blackman (-61 db), etc., are available in biomedical digital signal-processing toolbox.

4. Results and discussion

The main objectives of biomedical digital signal processing are noise/artefact removal, precise quantification of signal model and its components through analysis that means system identification for modelling and control purposes, feature extraction for resulting function, and prediction of future pathological. ARTFA model calculates a new set of AR parameters each time a new sample is available.

Principal Components Analysis has the applications of dimensionality reduction, determination of linear combinations of variables, feature selection, multidimensional data visualization, and identification of underlying variables [15].Components with the low % variances are known as minor components (MCs) are regarded as unimportant or associated with noise, whereas largest variances are regarded as important. However, in some applications, the



:

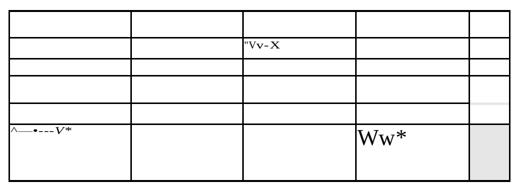


Figure 3. Showing principal components of sinus patient (left side) and Figure 3. showing principal component of normal person (right side)

MCs are of the same importance as the PCs, which is noteworthy here [14][15]. In the proposed algorithm the MCs reveal meaningful information. In the case of feature extraction and dimension reduction, PCA proposes a method based on the eigen structure of data covariance matrix. If signals are zero-mean, the covariance and correlation matrices are identical.

Eigenvalues=0.00636069 & 0.000955443(sinus patient) and = 0.108343 & 0.00848956(Normal)

0.1 *Eigenvector matrix* = [0.993407 —0.1146371 1463 7 -0.9934071 (sinus patient) and 0.985834 -0.1677261 6772 6 -0.985834] (Normal person)

-0. 1

Variance of First Principal Component-

$\%PC_1 = \frac{0.00636069}{0.007316133} = 14.92$ (sinus patient) and $= \frac{0.108343}{0.11683256} = 92.733$ (Normal)

Variance of Second Principal Component-

$\%PC_2 = \frac{0.000955443}{0.007316133} = 13.06$ (sinus patient) and $= \frac{0.00848956}{0.11683256} = 7.244$ (Normal)

4.1. Respiration signal analysis Results of sinus patient-

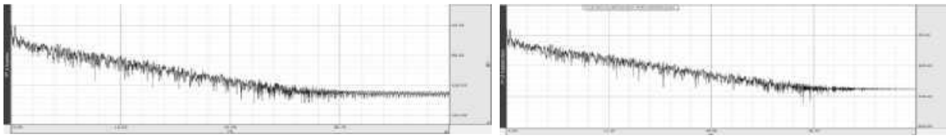


Figure 4. (a) FFT analysis of sinus patient with 4096 point of FFT with zeros padding and hamming window(left side) and (b) FFT analysis of sinus patient with 4096 point of FFT with zeros padding and Hanning window (right side)

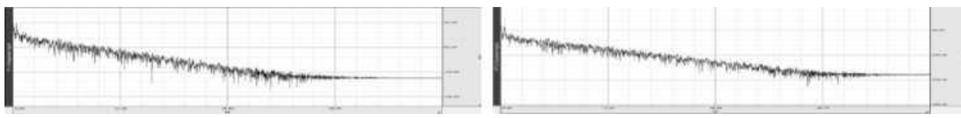


Figure 5. (a) FFT analysis of sinus patient with 4096 point of FFT with zero spading and bartlet twindow (leftside) and (b) FFT analysis of sinus patient with 4096 point of FFT with zero spading and blackman (-61db) window (rightside)

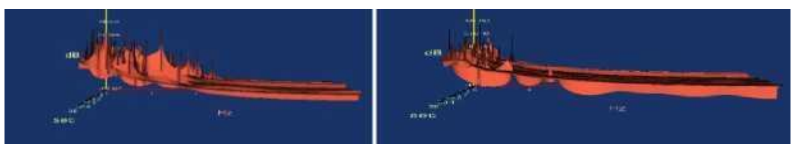


Figure 6. (a) Respiration signal time-frequency analysis of sinus patient and (b) airflow signal time-frequency analysis of sinus patient with time interval 1sec, model order 15, frequency resolution 1024 points and amplitude scaling is in decibels

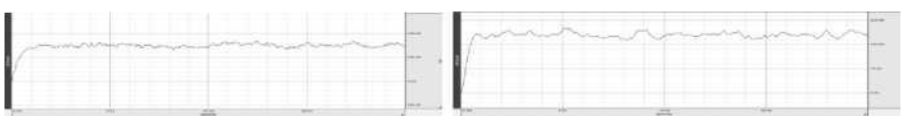


Figure 7. (a) Respiration signal autoregressive modelling of sinus patient with model order 10 and samples length 3806 and (b) airflow signal autoregressive modelling of sinus patient with model order 10 and samples length 3806

4.2. Respiration signal analysis results of normal person-

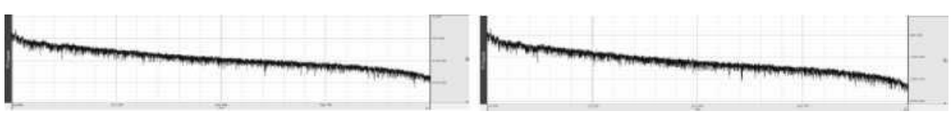


Figure 8. (a) FFT analysis of normal person with 4096 point of FFT with zeros padding and bartlett window and Fig. 8(b) FFT analysis of normal person with 4096 point of FFT with zeros padding and hanning window

In our study purpose Hamming and Blackman (-61db) window are the right choice because it suppresses the edges much more efficient, than the other window function.

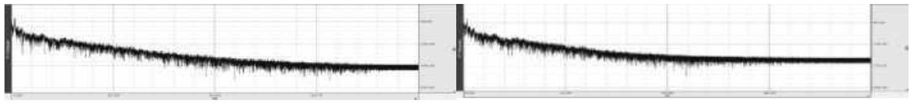


Figure 9. (a) FFT analysis of normal person with 4096 point of FFT with zeros padding and hamming window and (b) FFT analysis of normal person with 4096 point of FFT with zeros padding and blackman (-61db) window

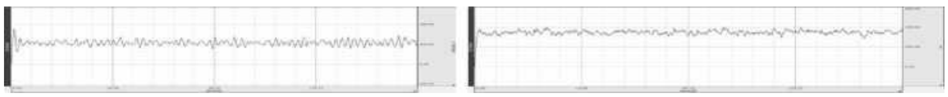


Figure 10. (a) Airflow signal autoregressive modelling (left) and (b) Respiration signal Autoregressive modelling with model order 10 and samples length 17832(right)

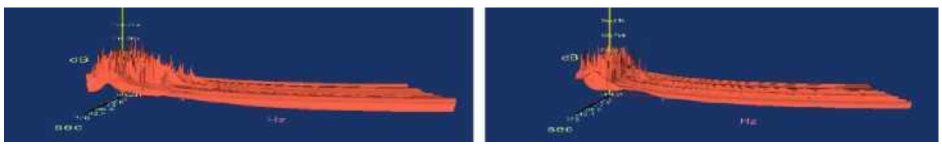


Figure 11. (a) Respiration time-frequency analysis of normal person with time interval 1 sec, model order 15, frequency resolution 1024 points and amplitude scaling is in decibels(left side) and (b) Airflow time-frequency analysis of normal person with time interval 1 sec, model order 15, frequency resolution 1024 points and amplitude scaling is in decibels (right side)

For Sinus patient analysis Kaiser Bessel window is the right choice because we can vary the side lobe of the window as well. FFT based analyses has errors known as leakage effect. It effect occurs when the FFT is calculated from a block of data which is not periodic. To nullify this appropriate windowing functions must be applied. In the time- frequency transform methods we have to use short time Fourier transform. It uses the fast Fourier transform with a sliding window of the data that is to be analysed. To determine the STFT, we are to need the specification of the window length, the window overlap and the type of the window. In order to select how long our window should be then by hit and trial method we have to examine the FFT on various numbers of samples but in this paper we have given the best sample results that we have performed on BiopacAcqKnowledge 4.0 software.

In AR Modelling it is very challenging to estimate the correct model order because this finds the number of poles in the model transfer function. If we have model order too small, then the power spectral estimation is to be biased more in the power spectrum toward the dominant peaks. If we have large model order than required, it enhance to spurious peaks in the power spectral estimation of the signal. As the AR model ARTFA has the merit of giving smooth spectral components and accurate estimation of the power spectrum.

Autoregressive (AR) technique estimates the power spectrum density of the signal. Therefore AR methods are not facing the problem of spectral leakage and thus show better frequency resolution. PSD estimation is done by calculating the coefficients.

5. Conclusion

We have shown the results that shows the respiratory analysis of Sinus Patient & Normal person. The principal component coefficients obtained from the application of principal component analysis to beats provide a surrogate for the respiratory signal from which respiratory rate can be accurately estimated. In the acquired signal of sinus patient there are more instability in the signal i.e. having very high & very low frequency components. A very high & very low frequency component represents the sinus problem (fast inspiration & expiration). In the results it is also clear that peaks in sinus patient is smaller than normal patient. If we talk about the certainty in the signal then you can also satisfied with the result.

The reason behind of we have chosen different window function is to correct the effect of leakage effect and we got the truth that the Hamming window and Blackman (-61db) window are the right choice to our purpose, for it suppresses the edges much more efficient, than the other window function. The use of such a shape of the window together with the overlap of the large data windows is a correct method to avoid the leakage problems.

Fast Fourier transform (FFT) is the fast way to compute DFT but scholars/researchers need to know the resolution limitations and the impact of signal windowing on the . Chaotic signals are neither periodic nor stochastic, which makes very difficult to predict beyond a short time. The problem in prediction is due to their high sensitivity to initial conditions. FFT does not best suited for short-duration signals/data. The spectral resolution is directly proportional to the ratio of sampling frequency to the total number of points. So if we want a resolution of 10 Hz in the spectrum, then we want to use at least 10-s duration of the signal.

The final Auto-Regressive time-frequency analysis provides high time and frequency resolution without any interference/distortion. We have got truth that ARTFA may be used for short segments of signal and in high noise/artefact environments.

References

[1] A. Hyvarinen, J. Karhunen and E. Oja, “Independent Component Analysis”, John Wiley & Sons Inc., New York, **(2001).**

[2] R. Vigario, V. Jousmaki, M. Hamalainen, R. Hari, and E. Oja, “Independent component analysis for identification of artifacts in magneto encephalographic recordings”, In Advances in Neural Information Processing Systems, MIT Press, Vol. 10, pp. 229-235, **(1998).**

[3] S. Makeig, A. J. Bell, T. P. Jung, and T. Sejnowski, “Independent component analysis of electro encephalographic data” Advances in Neural Information Processing Systems, MIT Press, Vol. 8, 145-151, **(1996)**

[4] G. D. Brown, S. Yamada and T. J Sejnowski, “Independent components analysis (ica) at the neural cocktail party”, Trends in neuroscience, Vol. 24, pp. 54-63, **(2001).**

[5] P.O. Hoyer and A. Hyvarinen, “Independent component analysis applied to feature extraction fromcolour and stereo images”, Network: Computation in Neural Systems, Vol. 11, pp. 191-210, **(2000).**

[6] P. Langley, E.J. Bowers, and A. Murray, “Principal Component Analysis as a Tool for Analyzing Beat-to-Beat Changes in ECG Features: Application to ECG-Derived Respiration”, Vol. 57, No. 4, pp. 821-829, **(2010)**.

[7] M Emdin, A Taddei and M Varanini, “Compact Representation of Autonomic Stimulation on Cardiorespiratory Signals by Principal Component Analysis”, **(1993).**

[8] X. Zhaoyan, X. Shiman and S. Qiuyan. “The Empirical Mode Decomposition Process of nonstationary signals”.

[9] R. Yan and R.X. Gao “A Tour of the Hilbert-Huang Transform: An Empirical Tool for SignalAnalysis”, Vol. 10, No. 5, pp. 40-45, **(2007).**

[10] V. Gupta, M. Mittal, G. Singh and S. Kumar Pahuja, “Fourier Transform of Untransformable signals using Pattern Recognition Technique”, Indonesia IEEE, **(2010).**

[11] Z. Huang “A Novel spectral Analysis Method of Atrial Fibrillation Signal Based on Hilbert-HuangTransform”, Vol. 1, pp. 825-828, **(2005).**

[12] E. Gottlieb and S.M. Armour, “Mitochondrial respiratory control is lost during growth factor deprivation”, Vol. 99, No. 20, pp. 12801-12806, **(2002).**

[13] A. Otero, “Palliative Performance Status, Heart Rate and Respiratory Rate as Predictive Factors ofSurvival Time in Terminally Ill Cancer Patients”, Vol. 31, pp. 485-492, **(2006).**

[14] M.P.S. Chawla, “A comparative analysis of principal component and independent component techniques for electrocardiograms”, Neural Computing & Applications, **(2008).**

[15] M. Taheri, B. Ahmadi, R. Amirfattahi and M. Mansouri, “Assessment of depth of anaesthesia using Principal Component Analysis”, Journal of Biomedical Science and Engineering, Vol. 2, pp. 9-15, **(2009)**.

[16] http://news.mit.edu.

[17] Neural Networks. Theory and Applications, Adaptive and Learning Systems for Signal Processing, Communications and Control, John Wiley & Sons Inc., New York.

[18] M. Turk and A. Pentland, “Eigenfaces for recognition”, Journal of Cognitive Neuroscience, pp.71-86, **(1991).**

[19] S. Y. Kung, K. Diamantaras, and J. Taur, “Neural networks for extracting pure/constrained/oriented principal components. In J. R. Vaccaro, editor”, SVD and Signal Processing Elsevier Science, Amsterdam, pp. 57-81, **(1991).**

[20] I. Feigler, “Time frequency analysis of ECG signals,” Computational Biology, **(2013).**

[21] J.L.A. Carvalho1, A.F. Rocha1, L.F. Junqueira, Jr, J. Souza Neto, I. Santos and F.A.O. Nascimento, “A Tool for Time-Frequency Analysis of Heart Rate Variability”, EMBC, pp. 2574-2577**, (2003).**

[22] J. Muthuswamy, “BIOMEDICAL SIGNAL ANALYSIS”, The McGraw-Hill Companies, **(2004).**