31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Multivariate Analysis of Intracranial Pressure (ICP) Signal Using Principal Component Analysis

N. Al-Zubi, L. Momani, A. Al-kharabsheh and W. Al-Nuaimy

Abstract— **The diagnosis and treatment of hydrocephalus and other neurological disorders often involve the acquisition and analysis of large amount of intracranial pressure (ICP) signal. Although the analysis and subsequent interpretation of this data is an essential part of the clinical management of the disorders, it is typically done manually by a trained clinician, and the difficulty in interpreting some of the features of this complex time series can sometimes lead to issues of subjectivity and reliability.**

This paper presents a method for the quantitative analysis of this data using a multivariate approach based on principal component analysis, with the aim of optimising symptom diagnosis, patient characterisation and treatment simulation and personalisation. In this method, 10 features are extracted from the ICP signal and principal components that represent these features are defined and analysed. Results from ICP traces of 40 patients show that the chosen features have relevant information about the ICP signal and can be represented with a few components of the PCA (approximately 91% of the total variance of the data is represented by the first four components of the PCA) and that these components can be helpful in characterising subgroups in the patient population that would otherwise not have been apparent. The introduction of supplementaty (non-ICP) variables has offered insight into additional groupings and relationships which may prove to be a fruitful avenue for exploration.

I. INTRODUCTION

Nowadays, the diagnosis and management of various neurological and head injuries involove monitoring and recording vast amounts of data relating to the instantaneous pressure of the cerebrospinal fluid within the skull cavity [1]. While current diagnosis procedure depends mainly on the mean value of this intracranial pressure (ICP) signal in addition to the surgeons observation of the clinical symptoms and neuroimages [2],[3], the ICP waveform proves to carry a lot of useful information and the data acquired from waveform recording are valuable to extract relevant information about the state of the patient [4]. However, these data sets are large and complex, and physicians, biomedical engineers, or surgeons are limited in the extent to which they can draw meaningful interpretations from visual inspection of this complex waveform. Other parameters derived from the raw ICP waveform are expected to offer additional insight into the nature of the underlying condition and the patients' response to treatment.

This paper presents a multivariate analysis method based on these additional features extracted from the ICP waveform that allows the representation of time varying data, and a Principal Component Analysis (PCA) that can assist in the interpretation of this data. Although PCA is among the most popular methods in analysis of multivariate signals, to the authors' knowledge, it has not yet been used in the analysis nor interpretation of ICP data. The present work uses PCA to identify the correlation between a number of signal features, and identify the most revealing features, additionally categorising patients into clusters according to their ICP features.

Such analysis can potentially be of great help in providing an enhanced and less subjective approach to the traditional methods for diagnosing neurological disorders and hydrocephalus in particular. By fusing multiple clinical measurements coupled with patients' symptoms, this analysis can be used for following up patients' state and predicting clinical outcomes.

TABLE I
ICP FEATURES EXTRACTED

Feature	Equation
Mean	$\frac{1}{N}\Sigma\, x_i$
Curve Length	$\Sigma\, x_i - x_{i-1}$
Energy	$\Sigma\, x_i^2$
Nonlinear Energy	$\Sigma\, -x_i.x_{i-2}+x_{i-1}^2$
Katz FD	$\Sigma \frac{log(k-1)}{log\frac{max\ (\Sigma_i\sqrt{(x_i-x_1)^2+i^2})}{\Sigma_i\sqrt{(x_{i+1}-x_1)^2+1}}+log(k-1)}$
Hurst	$\ln\left(\frac{range(x_i)}{\sigma_x(x_i)}-\frac{i}{2}\right)$
Shannon Entropy	-Σ f(x).log(f(x))
Peak Power	max(PSD)
Peak Frequency	index(max(PSD))
Spectral Entropy	-Σ PSD.log(PSD)

II. MATERIALS AND METHODS

A. Clinical Data

Digital ICP recordings were taken using an intraparenchymal pressure probe in the form of a Codman MicroSensor™ miniature strain gauge mounted on a nylon catheter and inserted into the frontal lobe parenchyma via small burr holes in the skull. Approximately 1482 hours of anonymous ICP recordings from 40 patients were analysed: 15 patients (37%) were suffering from hydrocephalus with 547 hours of ICP recordings, and others had different conditions, but all were

N. Al-Zubi, L. Momani, A. Al-kharabsheh and W. Al-Nuaimy are with Electrical Engineering and Electronics Department, University of Liverpool, Brownlow Hill, Liverpool, UK {nael,momani,karabshe,wax}@liv.ac.uk

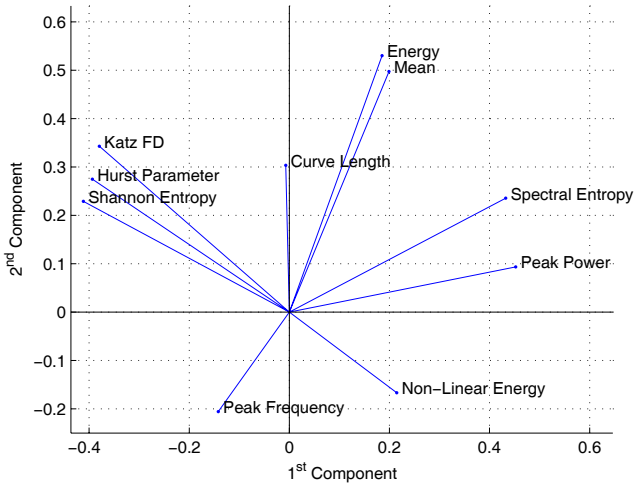

Fig. 1. ICP features represented in the first and second component of the PCA plane

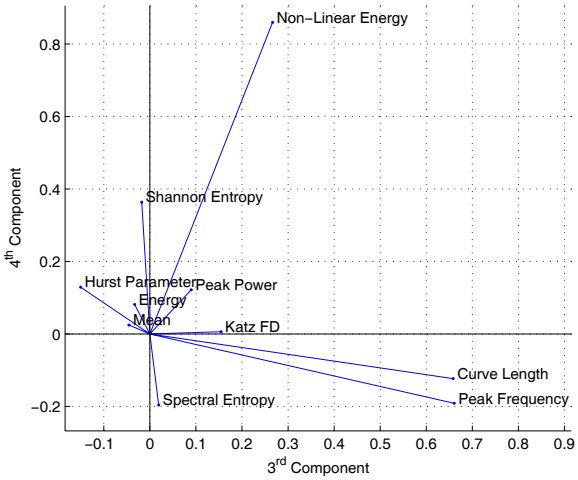

Fig. 2. ICP features represented in the third and fourth component of the PCA plane

exhibiting symptoms. Ages of all patients ranged from 1 to 20 years, and male to female ratio was 1:1 (20 males and 20 females). Data was recorded at different sampling rates, 40Hz, 100Hz and 400Hz, and recording length varied from 5 to 75 hours.

B. ICP Signal Features Extracted

Ten features are extracted from the ICP signal (Table I), based on features suggested by Wiggins *et al.* [5] for ECG signals and D'Alessandro *et al.* [6] for EEG signals. These features are extracted using a moving time sequence window of 6 seconds then averaged over one hour periods, and these features are briefly described as follows. The mean value is already used as a diagnosis parameter by physicians and surgeons, although recent studies report its poor indication of patient status [7]. The curve length is used for observing changes in amplitude and frequency without sensitivity to the measure of self-similarity. Energy of the ICP signal shows the position trend of the signal according to certain level. The nonlinear energy measures the energy of the signal proportional to its amplitude and frequency. Katz fractal dimension and Hurst parameter are used to measure long sequence dependency and self similarity of the signal. They are important in looking for patterns within the data. Shannon entropy is a measure of the randomness of the signal. Peak frequency represents the frequency of the maximum power, where studies show that this frequency represents the heart rate [1]. It is claimed [6] that the spectral entropy measures the regularity of the power spectrum of the signal, and represents a measure of the distribution of the frequency.

C. Principal Component Analysis (PCA)

PCA [8] is one of the most used multivariate methods, and aims to optimally representing data sets and a large number of variables to identify patterns in the data and highlight their distribution. This method generates a new set of variables (principal components) each of which is a linear combination of the original variables. All the principal components form an orthogonal basis for the space of the data. PCA takes as input a matrix of 10 columns to the features listed above, and rows corresponding to an hour's worth of observations. With standardised data, the outputted matrix "coeff" represents the correlation between principal components and variables and "scores" represents individual readings in the principal component space. Biplot, as they have good geometrical representation, can be used to represent variables, where the cosine of the angle between variables vectors reflect their correlation, and the length of the variable vector represents the amount of variance explained on the principal space. Additional supplementary variables and readings can be represented in the original calculated principal components where they did not contribute to the original analysis, which enables the interpretation of some variables or some readings which excluded from analysis because they may have been calculated or collected in different situations, and also enables the projection of new data [9].

III. RESULTS AND DISCUSSION

A. Interpretation of Principal Components

In total, 91% of the variance of the data was found to be represented by the first four components of the PCA. The first component represents 40.5% of the variance of the data, its positive part is represented by the spectral entropy of the ICP signal (i.e. regularity of its spectrum) and the peak power, while the negative part is represented by the Katz fractal dimension and Hurst parameter which measure the long dependance of the signal, and the entropy which measures the randomness of the amplitude of the signal, as shown in Fig. 1. The first component is thus mainly defined by the regularity (randomness) of the ICP signal. The second component represents 26% of the variance of the data, and is represented by the energy and the mean value of the ICP signal, as in Fig. 1. The third component

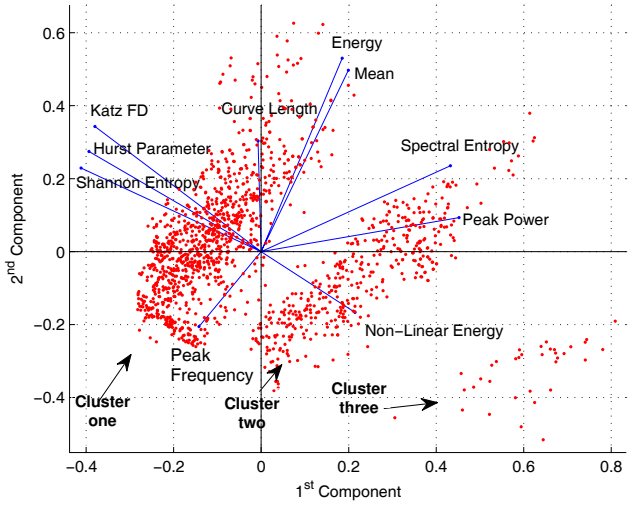

Fig. 3. Analysis of the patients represented in the first and second components of the PCA plane, showing three clusters of patient types

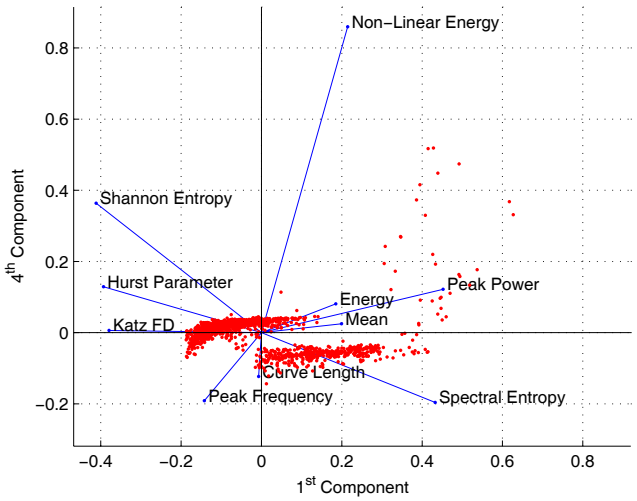

Fig. 4. The three clusters as shown from the first and fourth components of the PCA

represents 15.8% of the variance, and represented by the curve length and the peak frequency as shown in Fig. 2. The fourth component represents 8.3% of the variance and represented by the average non-linear energy as in Fig. 2.

Using the generated components, it would appear from Fig. 3 that three clusters of patients can be identified. If the data used had been descriptive, it might have been possible to project the status of the patients on these clusters, describing clusters in terms of other physiological parameters. These clusters are best identified by the first and the fourth components, Fig. 3 shows clusters in the first and second component, while Fig. 4 shows clusters from the first and fourth components view, where each dot represents one hour average of calculated parameters from ICP signal. Referring to the interpretation of components, the first cluster has higher irregularity compared to the other two (referring to the first component) and less non-linear energy (referring to the fourth component), which the third cluster has the highest non-linear energy and the least irregularity.

B. Additional Supplementary Variables

Additional supplementary variables can be projected on the generated component space. Fig. 5 shows the projection of the *Gender* and *Age* variables on the first and second components. The correlation between the first component and the *Gender* variable is approximately 0.3, which agrees with the manipulated data, where cluster 1 has male to female ratio of 1:2 and clusters 2 and 3 have male to female ratio of 2:1. This means that the males in this study show higher peak power and spectral entropy compared to the females, while females have higher Katz FD and Hurst parameter. Also there is a correlation of 0.3 between the *Gender* variable and the third and the fourth components, which means that males appear to exhibit higher non-linear energy and peak frequency compared to the females. With a correlation of 0.13, between gender variable and the peak frequency which corresponds to the heart rate, as shown in Fig. 6, this is expected as heart rate is slightly higher for males than females.

Fig. 5 represents the projection of the *Age* supplementary variable on the third and fourth components. With the small range of ages among patients, a good correlation of approximately 0.5 is shown between the *Age* variable and the second, third and fourth component, consistent with visual interpretation the data. This suggests that the older the patients are, the higher the non-linear energy is.

IV. CONCLUSIONS AND FUTURE WORK

This work has demonstrated the suitability of using PCA for the analysis of extracted features from ICP signals of patients suffering from hydrocephalus and other neological disorders. Although limited in scope and utilising limited data, this study shows the appropriateness of PCA for analysing features extracted from ICP signal and highlights the most revealing ones. The first four components represent 91% of the total variance of the data, the first and the second together represent more than 66% of the variance. Supplementary variables were added to find the relationship between the gender and age variables and the established principal components. Moreover, clustering patients according to the principal components has been facilitated, allowing the characterisation and comparison of subsets of the patient population that might otherwise not have been immediately clear.

With the ability of this method to cluster different groups of patients, a more elaborate and meaningful classification process can be applied to stratify patients into different categories according to perceived risk and improvement. Moreover, according to each group, this method can potentially provide an effective way of determining suitable parameter settings for programmable shunts used in treat-

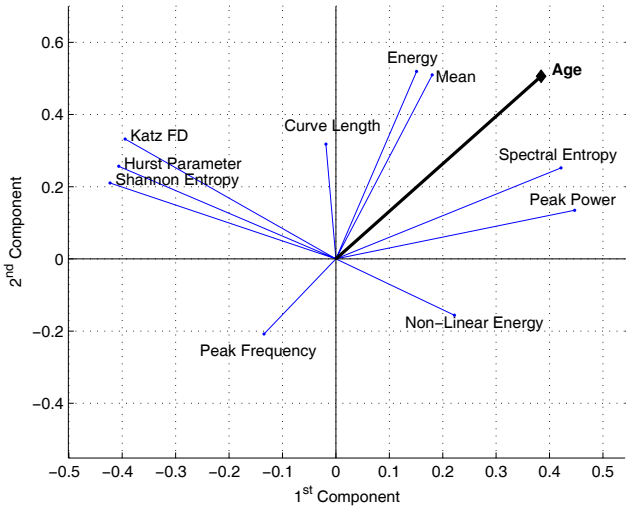

Fig. 5. Age as a supplementary variable is projected on the first and second components of the PCA plane

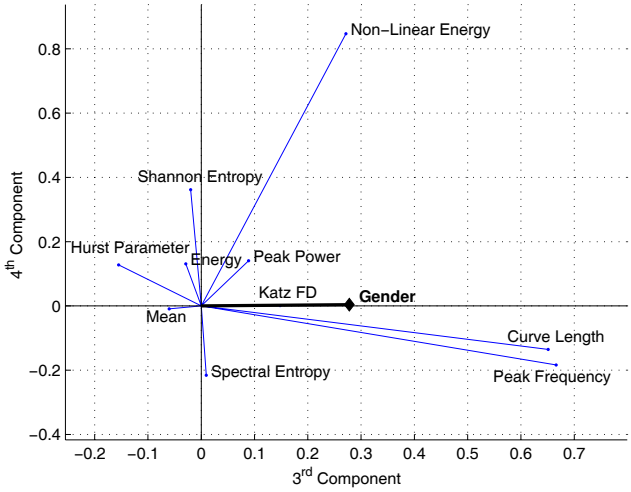

Fig. 6. Gender as a supplementary variable is projected on the third and fourth components of the PCA plane

ment of hydrocephalus, a process which is currently highly subjective. Further studies can be done to investigate the follow up of patients, by acquiring sequent data for patients and represented as a trajectory in the PCA components, these data should include, in addition to the ICP signal, indication about the patients' state such as clinical symptoms and other measured values such as blood pressure and heart rate. For hydrocephalus patients where valves are implanted as part of shunting systems, other data such as valves settings, can be analysed and configured, thus helping in improving these shunts and finding appropriate settings for each cluster of patients.

V. ACKNOWLEDGMENTS

The authors are grateful to Dr M. Al-Jumaily and Mr C. Mallucci from the Walton Neurological Centre, UK, for providing the clinical data and expertise.

REFERENCES

[1] M. Czosnyka, P. Smielewski, I. Timofeev, A. Lavinio, E. Guazzo, P. Hutchinson and J. Pickard, Intracranial pressure: More than a number, *Neurosurg Focus*, vol. 22, 2007, pp 265-271.

[2] W. Pfisterer, F. Aboul-Enein, E. Gebhart, M. Graf, M. Aichholzer and M. Mhlbauer, Continuous intraventricular pressure monitoring for diagnosis of normal-pressure hydrocephalus, *Acta Neurochirurgica Journal*, vol. 149, 2007, pp 983-990.

[3] A. Marmarou, H. Young and G. Aygok, Diagnosis and management of idiopathic normal-pressure hydrocephalus: a prospective study in 151 patients, *J Neurosurg* , vol. 102, 2005, pp 987-997.

[4] W. Pfisterer, F. Aboul-Enein, E. Gebhart, M. Graf, M. Aichholzer and M. Muhlbauer, Continuous intraventricular pressure monitoring for diagnosis of normal-pressure hydrocephalus, *Acta Neurochirurgica*, vol. 149, 2007, pp 983-990.

[5] M. Wiggins, A. Saad, B. Litt and G. Vachtsevanos, Evolving a Bayesian classifier for ECG-based age classification in medical applications, *Applied Soft Computing Journal*, vol. 8, 2008, pp 599-608.

[6] M. DAlessandro, R. Esteller, G. Vachtsevanos, A. Hinson, J. Echauz and B. Litt, Epileptic seizure prediction using hybrid feature selection over multiple intracranial EEG electrode contacts: a report of four patients, *IEEE Trans. Biomed. Eng.*, vol. 50, 2003, pp 603-615.

[7] R. Hutchinson E. Lin, W. Poon and T. Oh, Systems analysis applied to intracranial pressure waveforms and correlation with clinical status in head injured patients, *British Journal of Anaesthesia*, vol. 66, 1991, pp 476-482.

[8] I. Jolliffe, *Principal Component Analysis*, Springer Series in Statistics; 2004.

[9] M. Wiggins, A. Saad, B. Litt and G. Vachtsevanos, Optimal representation of supplementary variables in Biplots from principal component analysis and correspondence analysis, *Biometrical Journal*, vol. 45, 2008, pp 491-509.