Independent Component Analysis and Evolving Fuzzy Neural Networks for the Classification of Single Trial EEG Data

Carl Stuart Leichter[1] , Andrzej Cichocki[2] , and Nik Kasabov [1]

[1]Department of Information Science, University of Otago, Dunedin, New Zealand.
[2] Laboratory for Advanced Brain Signal Processing, RIKEN, Wako-Shi, Japan
(E-mail: [1]leica666@infoscience.otago.ac.nz, 2:cia@brain.riken.go.jp, 3 : nkasabov@otago.ac.nz)

Abstract

The paper presents a novel model for classification of EEG data based on independent component analysis (ICA) as a feature extraction technique, and on evolving fuzzy neural networks – as a classification modeling technique. One of the problems in such models is that some of EEG channels and model variables are redundant, 'noisy', and have a detrimental effect on the performance of the whole system. Here the problem of removing the 'noisy' independent components is explored through applying an exhaustive a-posteriori search algorithm. In the experiment performed and reported in the paper single trial EEG data from four different stimulus conditions are transformed into independent components via ICA. One-by-one, individual independent components from each possible pairing of the stimuli are suppressed and the residual data are projected, through the estimated mixing matrix, back to the EEG electrodes. Feature vectors are extracted from the resulting data and are used to train and test a two-class EFuNN classifier. The independent components whose suppression leads to the best performance enhancement are considered to be the noise components. After all six possible pairings of the four stimuli have been used, the identified noise components are removed from the data and, after a back-projection and feature extraction, the residual data is used to train and test a four class EFuNN. EFuNN classification performance using this a-posterior ICA denoising technique is compared to an analogous PCA denoising technique as well as compared to performance using raw EEG data.

Keywords: Independent Component Analysis, ICA, EEG, Classification, evolving fuzzy neural networks

1 Introduction

Analysis of single trial EEG is an extremely difficult problem. Wherever possible, cognitive science research tends to avoid the use of single trial data in favour of grand averaged data. This is because single trial EEG are notoriously noisy, as they are contaminated with a variety of noise sources emanating from the EEG subject (e.g.: electromyographic signal sources, electro-ocular signal sources, EKG, etc) [1,2]. Furthermore, in "real-world" environments, EEG can also be contaminated by environmental noise sources such as line and electrical equipment noise.

However, cognitive science research notwithstanding, "real-world" applications of EEG will require analysis of single trial data within uncontrolled environments. Alertness monitoring and brain-computer interfaces are two examples of such applications. Alertness monitoring would be used in occupations that are vigilance critical such as air traffic control or extended submarine sonar operations. These applications will require the on-line, real-time analysis of EEG data [3, 4]. These data are likely to be very non-stationary and the analytical techniques used will have to be capable of adapting to the changing nature of the data. Because of their ability to adapt to novel data, Evolving Fuzzy Neural Networks (EFuNN) should be well suited to such a task [5].

The signal of interest in EEG is generated by large collections of neurons in the brain that act in a coherent and organised fashion [6]. Since they emanate from the various cortical regions of the brain, these signals will henceforth be referred to as the "cortical components". However, the signal power of these components is much lower than the power of other bioelectric events; that is why it is so difficult to extract these cortical component signals from the background bioelectric noise. Independent Component Analysis (ICA) is one method of extracting these cortical signals from the multitude of noise sources that mask them.

ICA is a solution to the classic signal-processing problem known as "the cocktail party problem". In this scenario, an array of microphones records the mixture of voices at a cocktail party. If an observer were to listen to the signal from any one of the microphones, s/he would be unable to understand the unintelligible jumble of conversation recorded. Assuming that there are at least as many microphones as there are voices, ICA can transform the data from all the microphones into separate voice signals. Similarly, in MEG and EEG analysis, ICA can separate the cortical components from the non-cortical bioelectric events as well as from the environmental noise sources [7].

However, once such a separation is achieved, the problem then becomes one of identifying and suppressing the noise components. Again, in real-world applications, this identification and suppression will have to be accomplished on-line and in real-time. A connectionist based information system (CBIS) such as an EFuNN, would also be well suited to this task. The architecture of such a system is illustrated in Figure 1.

But, in order to use a CBIS in this fashion, it must be trained and tested with known noise and signal components. As yet, no such collection of components exists. Furthermore, most methods of identifying such components are very laborious and require expert visual examination of the independent component data. It would be useful to have a method to automatically identify the components that represent noise with respect to the kinds of applications previously described. This method could be used to generate training examples that will be later used to train a CBIS to identify and suppress the noise components.

2 Methods and Materials

The EEG data was collected in an experiment that used four stimulus conditions. In the auditory stimulus case, a 1Khz tone of 50 mSec in duration was presented to the subject in one-second intervals. The visual stimulus had the same duration and interval of the auditory stimulus and consisted of a white circle on a black background. The mixed auditory and visual stimulus combined the two stimuli already described, with the auditory stimulus presented first and then the visual stimulus presented to 100 mSec later. The fourth case is when no stimulus was presented. The EEG data were acquired using a 64 electrode EEG system that was filtered using a 0.05Hz to 500 Hz band- pass filter and sampled at 2Khz.

For the purposes of this experiment, the 2KHz-sampling rate was reduced via down-sampling to 200Hz. Also, only the 19 channels corresponding the international 10-20 EEG standard were used.

Baseline EFuNN performance was established using this raw EEG data. For both raw and denoised EEG data, when a stimulus mark was encountered in the data, the 180 time samples that immediately followed the stimulus presentation were taken from the data for each of the 19 electrodes. The post-stimulus data selection is illustrated in Figure 2.

This data was then transformed with a 180 point DCT. Then, the RMS band power for each of the five classical EEG bandwidths, Delta (1 - 3.5Hz), Theta (4 – 7Hz), Alpha (8 – 13Hz), Beta (14 – 26Hz) and Gamma (29 – 50Hz) were computed and normalised by using the RMS power from the entire 180-point DCT. The five normalised band powers were computed for the data from each of the 19 EEG electrodes. This resulted in a 95-component feature vector that represented the subject's EEG response to the given stimulus. Feature vector generation is illustrated in Figure 3.

Initially, a four class EFuNN was trained and tested with feature vectors generated from the raw EEG data. Due to computer memory constraints, only 37 of the 50 available stimulus trials for each stimulus were used. The entire gamut of training/testing partitions of the data were used, ranging from only 1 training vector with 36 testing vectors to the other extreme using 36 training vectors with only 1 test vector. The baseline raw EEG performance of the EFuNN is illustrated in Figure 4.

The FASTICA [8] algorithm was then used to perform ICA on the same EEG data. The maximum number of iterations (`'maxNumIterations'`) for FASTICA was set to 5000 with the hyperbolic tangent [tanh (**a1*u**)] used as the non-linear function and **a1** = 0.01. FASTICA calculated the independent components and also generated estimates for the separation and mixing matrices.

The ICA data "noise" components were defined as those components that degraded the ability of the EFuNN to classify between the various stimulus condition EEG. This is a very broad interpretation of the concept of noise, since some of these components might actually represent interesting signals. However in this particular instance, because they interfere with the EFuNN's classification of EEG, these components are considered to be noise.

The noise components were identified in the following fashion. For a given pairing of the four stimuli (e.g. auditory vs. visual), the training partition was fixed to 27 training vectors. This partition sets aside 10 testing vectors. Then, an exhaustive search was made to determine which component from each stimulus should be suppressed to maximize EFuNN test performance. Suppression of a specific component was achieved by zeroing out the time series data for that component and then projecting the residual data back to the electrodes for feature vector extraction and classification. Suppression of noise components is illustrated in Figure 5.

The following section of pseudo code shows a-posterior identification of the noise components with respect to EFuNN classification. For this example, the classification is between auditory and visual stimulus EEG.

Pseudo-code for noise component identification is given below:

```
for aud _IC = 1 to 19
  for vis _IC = 1 to 19
       suppress aud _IC in the auditory data
       suppress vis _IC  in the visual data
       project the residual data back to the electrodes
       extract feature vectors from the resulting data
       use feature vectors to train/test a 2 class EFuNN
       if this is best the performance so far
          save aud _IC number
          save vis _IC number
     end {for vis _IC_ number = 1 to 19}
end {for aud _IC_number = 1 to 19}
```

As can be seen, the algorithm runs through an exhaustive search of 361 (19x19) iterations to identify only two noise components: one from each of the stimulus types. There may exist more noise components, but to improve speed performance, only two were sought for each pair of stimuli. Seeking more noise components would require re-running the algorithm with the remaining 18x18 = 324 possible component combinations. On a PIII computer clocked at 733 MHz, this algorithm took 100 minutes to search through the initial 361 combinations required to find the first two noise components.

As will be shown later, additional noise components from each stimulus were eventually identified.

Even if the algorithm is restricted to searching for only one noise component from each stimulus, a four-way search using a 4-class classifier requires unacceptably long computational time. The number of possible combinations required to exhaustively search for noise components grows exponentially with the number of stimuli: 19^4 = 130321! This is 361 squared, so the algorithm would have to run for 36,100 minutes (25 days) to identify the noise components.

Instead, a more efficient pair-wise search method was used. With the four stimulus conditions, there are 12 possible pair-wise combinations, but only 6 of those are unique:

1) Auditory versus Visual
2) Auditory versus Mixed Auditory and Visual
3) Auditory versus No Stimulus
4) Visual versus Mixed Auditory and Visual
5) Visual versus No Stimulus
6) Mixed Auditory versus _Visual and No Stimulus

The other six combinations merely represent a reordering of the stimuli within the pair (e.g. from Auditory versus Visual to Visual versus Auditory) Since the EFuNN classification performance is invariant with respect to class order, it is not necessary to use these other pairings. The previously presented pseudo-code algorithm is now generalised and extended for this four-class search.

```
Class1 = Auditory_ Stimulus _EEG
Class2 = Visual_ Stimulus _EEG
Class3 = Mixed_Auditory _Visual_ Stimulus _EEG
Class4 = No_Stimulus_EEG

for ClassA = Class1 to Class3
   for ClassB = ClassA to Class4
      for A _IC = 1 to 19
        for B_IC= 1 to 19
          suppress A _IC_ in the ClassA ICA data
          suppress B _IC_ in the ClassB ICA data
          project the residual data back to the electrodes
          extract feature vectors from the resulting data
          train/test the 2 class EFuNN with this data
          if this is the best performance so far
             save A _IC  number
             save B _IC number
        end {for B _IC = 1 to 19}
      end {for A_IC = 1 to 19}
      store the best performance for this stimulus pair
    end {for ClassB = ClassA to Class4)
end {for ClassA = Class1 to Class3}
```

Once the noise components for each stimulus pairing have been identified, they are all suppressed and the residual data from all four of the stimulus conditions are used to train/test a four class EFuNN classifier.

3. Results

The pair-wise search identified the components shown on Table 1 (next page) as noise within their respective EFuNN classification tasks.

Table 1: Pair-wise Classification Noise Components

Stimulus Pairing		Noise Component	
Stimulus Pair Condition 1	Stimulus Pair Condition 2	Stimulus 1 Noise component	Stimulus 2 Noise component
Auditory	Visual	3	15
Auditory	Mixed Auditory/ Visual	6	13
Auditory	No Stimulus	11	2
Visual	Mixed Auditory/ Visual	1	4
Visual	No Stimulus	1	4
Mixed Auditory/ Visual	No Stimulus	1	8

On the basis of this pair-wise search, the independent components suppressed for the 4-class classifier are shown in Table 2.

Table 2: Accumulated Noise Components

Stimulus Condition	Components Suppressed
Auditory	3, 6, 11
Visual	1, 15
Mixed Auditory/Visual	1, 4, 13
No Stimulus	2, 4, 8

With these components suppressed, for all possible training/testing partitions, the EFuNN exhibited a marked performance improvement against the baseline EEG. This enhanced performance is illustrated in Figure 6.

The same search and suppression algorithm was also applied using PCA instead of ICA. In this instance, “noisy” *principle* components were sought for suppression. For the unrealistically low training/testing partitions, PCA denoising outperforms ICA denoising. However, in the more realistic cases, ICA denoising outperforms PCA denoising. Furthermore, unlike PCA, the ICA components represent specific cortical processes. These results are shown in Figure 7.

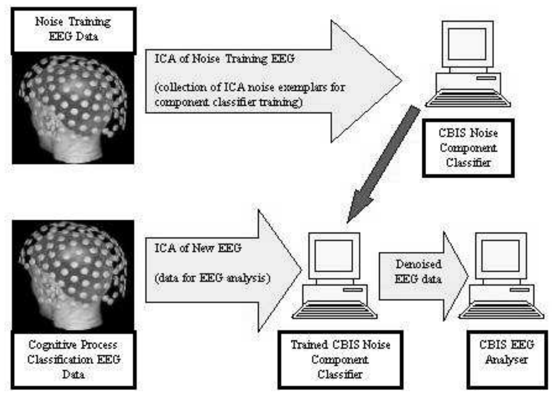

Figure 1: Multiple CBIS Architecture for Denoising and Classifying EEG

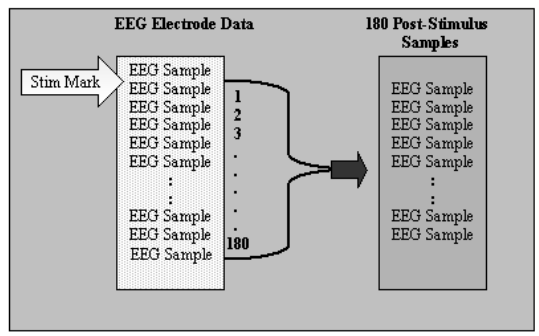

Figure 2: Selection of 180 Post-stimulus Samples

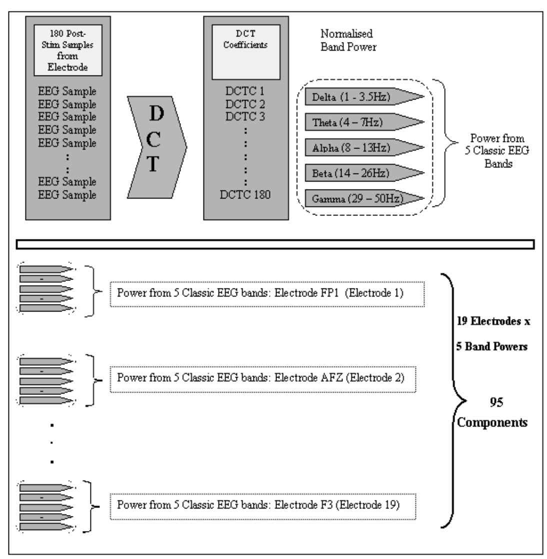

Figure 3: Generation of Feature Vectors

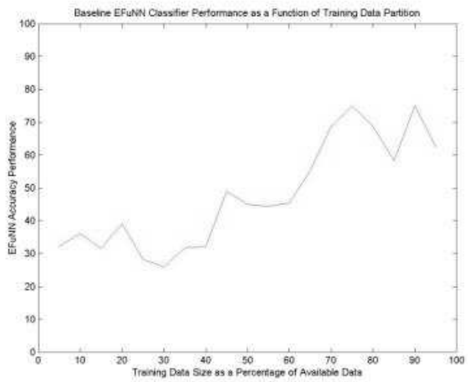

Figure 4: Baseline EFuNN Performance

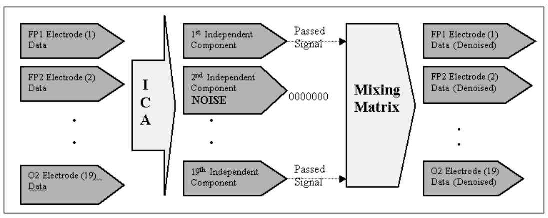

Figure 5: ICA suppression of EEG noise

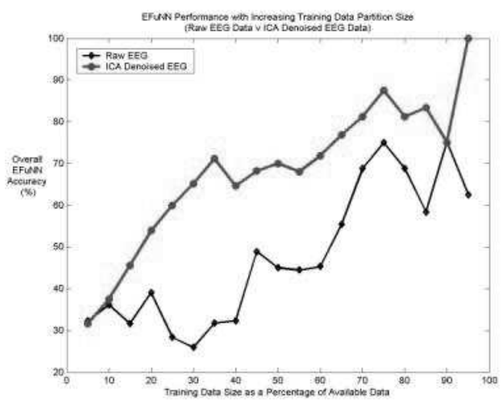

Figure 6: EFuNN Performance (Raw EEG v Denoised EEG)

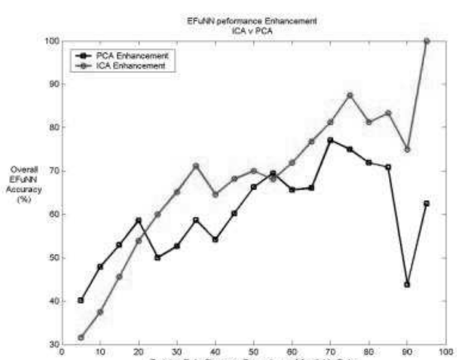

Figure 7: ICA Denoising versus PCA Denoising

4.Summary and Conclusions

This research demonstrates that ICA denoising can dramatically improve the EFuNN classification of EEG from various stimuli. Furthermore, in this respect, ICA denoising is superior to the analogous PCA denoising technique. It must be noted that such performance gains were only found in an a-posteriori fashion; so this research has limited applicability at this time. However, it should be recalled that the primary purpose of this technique was the identification and collection of noise independent components; the presented algorithm clearly demonstrates this capability. An accumulation of EEG noise components will eventually be used to train a noise-suppression EFuNN, whose purpose will be to recognise and suppress noise components *prior* to feature vector generation for eventual classification by another EFuNN.

Another possible application of this technology is within the field of medical diagnosis and neuro-surgery. Future research will use epileptic EEG data in an attempt to identify the most dominant epileptogenic independent components. In this application, rather than search for and suppress components in order to *enhance* EFuNN recognition, the algorithm will search for and suppress the components that lead to maximal *degradation* of data recognition. In this case, the suppressed component should represent the most dominant epileptic signal component. When this epileptic signal component is projected back to the EEG electrodes, source localisation techniques may then be used to locate the epileptogenic foci for eventual surgical excision.

Evolving connectionist systems, and EFuNN in particular, allow for on-line learning and model modification based on a continuous stream of input data [9]. This feature can be used for on-line creation and on-line modification of models of cognitive functions based on continuous EEG streams of data. The models would allow also for a fusion of different sources of information (e.g. EEG and fMRI). These are challenging tasks for a future development.

Acknowledgements

This research was undertaken during participation in the RIKEN Brain Science Institute 2001 Summer Internship Program. I would like to thank Hovagim Bakardjean, the EEG Engineer at the Advanced Brain Signal Processing Laboratory, for his support and encouragement.

References

[1] Oja, E., Karhunen, J., Hyvärinen, A., Vigário, R. & Hurri, J. Neural Independent Component Analysis - Approaches and Applications. In: Kasabov, N. & Amari, S. (eds.), Brain-like Computing and Intelligent Information Systems. Singapore 1998, Springer, pp. 167-188.

[2] Hori, G. Amari, S. Chichocki, A. Mizuno, Y. Okuma, Y. Aihara, K. Using ICA of EEG for Judgment of Brain Death. Proc. ICONIP- 2000 pp1216-1220, 2000

[3] Makeig, S. Elliot, F. S. Postal, M. First Demonstration of an Alertness Monitoring Management System. Naval Health Research Center Rpt No 93-36, San Diego, CA. March 1994

[4] Jung, PJ. Makeig, S. Stensmo, M. Sejnowski, TJ. Estimating Alertness form the EEG Power Spectrum IEEE Transactions on Biomedical Engineering, 44(1), pp 60 –69. Jan 1997

[5] Kasabov, N. The ECOS Framework and the ECO Leaning Method for Evolving Connectionist Systems. Journal of Advanced Computational Intelligence. 2(6) pp 195 – 202. 1998

[6] Steyn-Ross, M. Steyn-Ross, DA. Sleigh, J. Liley, D. Theoretical Electroencephalogram Stationary Spectrum for a White-Noise-Driven Cortex: Evidence for a General Anesthetic Induced Phase Transition. Physical Review E 60 (6) pp 7299 – 7311, Dec 1999

[7] Hyvärinen, A., and Oja, E., Independent Component Analysis: Algorithms and Applications. Neural Networks, 13(4-5):411-430, 2000

[8] FASTICA is available for download from http://www.cis.hut.fi/projects/ica/fastica/

[9] Kasabov, N. Adaptive learning system and method, Patent New Zealand, PCT/NZ01/00059, April 2000