Med Biol Eng Comput (2008) 46:263–272
DOI 10.1007/s11517-007-0289-4

ORIGINAL ARTICLE

A novel method for automated classification of epileptiform activity in the human electroencephalogram-based on independent component analysis

Marzia De Lucia · Juan Fritschy · Peter Dayan · David S. Holder

Received: 18 December 2006 / Accepted: 20 November 2007 / Published online: 11 December 2007
© International Federation for Medical and Biological Engineering 2007

Abstract Diagnosis of several neurological disorders is based on the detection of typical pathological patterns in the electroencephalogram (EEG). This is a time-consuming task requiring significant training and experience. Automatic detection of these EEG patterns would greatly assist in quantitative analysis and interpretation. We present a method, which allows automatic detection of epileptiform events and discrimination of them from eye blinks, and is based on features derived using a novel application of independent component analysis. The algorithm was trained and cross validated using seven EEGs with epileptiform activity. For epileptiform events with compensation for eyeblinks, the sensitivity was 65 ± 22% at a specificity of 86 ± 7% (mean ± SD). With feature extraction by PCA or classification of raw data, specificity reduced to 76 and 74%, respectively, for the same sensitivity. On exactly the same data, the commercially available software Reveal had a maximum sensitivity of 30% and concurrent specificity of 77%. Our algorithm performed well at detecting epileptiform events in this preliminary test and offers a flexible tool that is intended to be generalized to the simultaneous classification of many waveforms in the EEG.

M. De Lucia (✉) · J. Fritschy · D. S. Holder
Medical Physics and Clinical Neurophysiology,
University College London, London, UK
e-mail: marzia.delucia@ucl.ac.uk

P. Dayan
Gatsby Unit of Computational Neuroscience,
University College London, London, UK

Keywords Electroencephalogram · Independent component analysis · Automatic classification · Epileptiform events · Eye-blinks artefacts

1 Introduction

1.1 Background

The human Electroencephalography has been employed as a routine clinical test for over 70 years and is widely used for diagnosis of epilepsy, and, to a lesser extent, brain abnormalities such as encephalopathies or dementia [5]. The principal abnormalities in epilepsy involve short transients due to neuronal depolarization and repolarization termed *spikes*, *spike wave complexes* and *sharp waves* (Fig. 1). These waveforms may occur in sequences, which are called *polyspikes* and *spike-and-wave* [18]. Other possible abnormalities include slow activity due to disruption of underlying brain mechanisms or to continuous abnormalities in brain tissue. In practice, all these pathological brain activities are relatively easy to diagnose. The main difficulty in EEG analysis lies in distinguishing these *epileptiform* waveforms from the many patterns which occur due to normal activity, such as alpha rhythm at 8–12 Hz, movement and muscle artefacts, eyeblinks (Fig. 2), or normal physiological activity.

Despite computational advances, current clinical practice is still for expert observers to interpret EEGs by eye. Over the past three decades, there has been considerable interest in finding automated ways to analyse the EEG in order to save time consuming human analysis. This is a special problem for longer recordings. Until now, these have mostly concentrated on detection of seizures and the spikes and sharp waves which indicate epileptic activity; sensitivities of up to

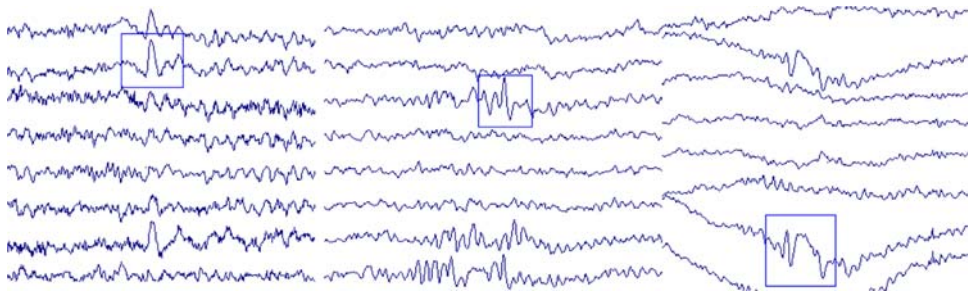

Fig. 1 Three examples of epileptic waveforms. From the *left* to the *right*, a *sharp wave*, a *spike wave*, and a *spike-and-wave*, characterized by a slow wave which follows the spike

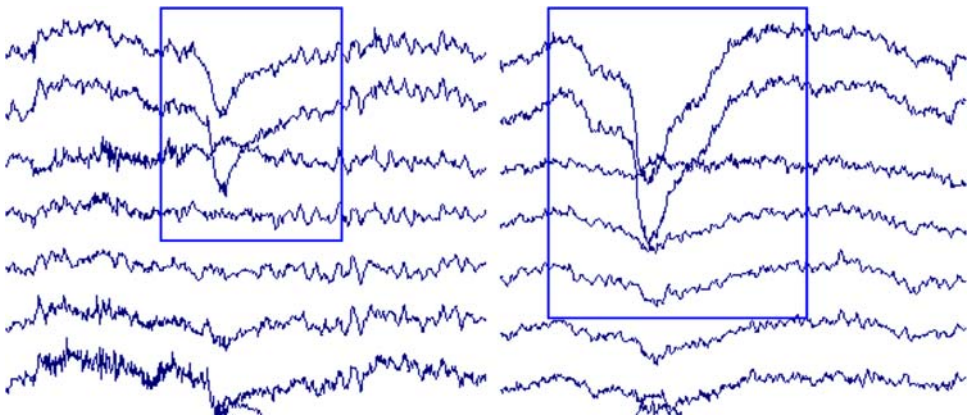

Fig. 2 Examples of *eye-movement* and *eye-blink* artefacts (*left* and *right* boxes)

about 79% for seizures and 89% for spikes and sharp waves have been reported for commercial software [42, 43].

Existing methods for automatic detection of epileptiform events are conventionally based on one or both of two principles: (1) *spike morphology*, in which a set of typical features that characterize spikes is utilized, and (2) *localization of transient events*, in which spikes are represented as transients that can be distinguished from ongoing background activity. Algorithms in the first class attempt to represent spikes by a set of features that make it possible to discriminate spikes from other events. These are constructed to capture and represent explicitly all the relevant information in the input, and eliminate redundancies that would corrupt the subsequent discrimination. Some features are readily related to spike morphology, such as duration, amplitude and kurtosis [10, 15, 16, 43]; others use more general signal processing methods such as waveform decomposition into Fourier or wavelet components [14, 25, 29]. Algorithms in the second class aim to detect epileptiform events by measuring their deviation from stationarity. Autoregressive analysis has been used both in single channel or multi channel EEG for the extraction of relevant coefficients for characterizing the non-stationarity of any epileptic events [1, 11]. The same principle has been implemented using nonlinear measures based on information theory that are able to capture the transitions between complex and less complex dynamics [8].

In a typical detection algorithm design, these principles are implemented in a tree-like structure algorithm or a multistage classifier. One common idea is to use different features of the EEG as inputs to one or more artificial neural networks, whose outputs are then passed through a set of rules based on fuzzy logic [35, 43]. Another is to use a highly structured network with several input and hidden layers [33, 40]. Such architectures can behave in complex ways, which often makes it difficult to diagnose and rectify flaws in classification.

1.2 Purpose

The long-term purpose of this work is to develop an algorithm, which can automatically detect all the clinically relevant waveforms in the EEG. Here, we report a novel, generalizable, method, implemented in this case for the special cases of major classes of epileptiform activity, and the discrimination of these from eyeblinks. These two features were selected because of the clinical significance of the former, and the false positives caused by the latter. We present preliminary results on a sample database of seven EEGs and compare these results to the commercial software Reveal (http://www.eeg-persyst.com/index.html), as well as other possible approaches for feature extraction prior to classification.

2 Methods

2.1 Design

Our algorithm falls into the class of spike morphology algorithms described above. First, the raw EEG was projected onto selected informative feature dimensions; second, these multidimensional features were used as the basis of a classifier. Note that, for the moment, we only considered single epileptic waveforms in the scalp EEG, thus excluding seizures and also polyspike sequences of these single waveforms.

The goal of feature extraction is to discover dimensions in the input that collectively enable simple, readily generalizable, separation of the various classes. This task is subject to various trade-offs, most notably between the power of the classifier and the power of the features, with the latter being subject to severe constraints based on the limited amount of information available in the training set. We therefore started by restricting our consideration to simple linear and quadratic classifiers, and sought small numbers of feature dimensions. We considered spectral, principal component analysis (PCA) and independent component analysis (ICA) approaches to determine appropriate features. The algorithm that performed most competently was based on ICA and a quadratic classifier. We, therefore, report the results achieved by this method.

In previous work, the most popular choice of features involves spectral [9], or wavelet-based [14, 25, 29] decompositions, and methods driven more directly by the data such as PCA and ICA have attracted much less study. However, the empirical bases are well suited to our goal, since they are automatically sensitive to the characteristics of the patterns involved. PCA and ICA are ways of extracting empirical bases that differ in subtle, but important, respects [17]. Indeed, ICA has previously been shown to be useful for analyzing epileptiform discharges, but only given a degree of hand-tuning [26–28, 32]. We employed a novel variant of ICA in a fully automatic procedure, and showed that it out-performs PCA and is at least competitive with Reveal on the same input dataset.

2.2 Independent component analysis (ICA) and its novel use in this study

The ICA is a general tool for analyzing multivariate data. ICA aims to represent input data as a linear mixture of an ensemble of (initially unknown) *sources*, which are assumed to be statistically independent [19, 22]. This model can be conveniently expressed using vector-matrix notation. We denote by **X** the matrix whose rows are the U (n-dimensional) input variables

$$X_i = (x_{i1}, x_{i2}, \ldots, x_{in}) \quad i = 1,..,U \tag{1}$$

and by **S,** the matrix of the V underlying sources

$$S_j = (s_{j1}, s_{j2}, \ldots, s_{jn}) \quad j = 1,\ldots,V \tag{2}$$

If **A** is the matrix of mixing coefficient a_{ij} ($i = 1,\ldots,U$, $j = 1,\ldots,V$), then ICA can be expressed by the following linear model:

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{S} \tag{3}$$

Without any prior knowledge on the properties of the sources or of the mixing process, the goal of ICA is to recover a version of the mixing matrix **A** and of the independent sources (*components*) **S**, given only the observation variables **X.**

Technically, a collection of random variables **S** is said to be independent if their joint distribution factorizes as a product of their marginal distributions:

$$P(S) = P(S_1, S_2, \ldots S_V) = P(S_1) \cdot P(S_2), \ldots, P(S_V) \tag{4}$$

There are many different ways or principles for quantifying the degree of independence of a collection of variables, and concomitant algorithms that optimize these quantities. Several possible implementations of ICA have, therefore, been proposed [3, 20, 21, 34]. In the following we will consider a 'fixed point' iteration algorithm (FastICA), that aims to maximize the *negentropy* of the variables, which is an information-based measure of independence [20, 21]. The problem is well defined if the sources have non-Gaussian marginal statistics; it is also most straightforward to extract fewer sources than input data $U \geq V$.

Not only are there different principles and algorithms, but also there are various ways to apply ICA to physiological data [24]. One application of ICA, commonly used for fMRI data, is called *spatial ICA* (sICA). In this, the signals at each point in space are decomposed into V *spatially* independent sources whose projections are fixed across time. sICA is conventionally applied to functional MRI [6, 30, 31], for which there are often many more spatial than temporal dimensions. Here, the number of sources, V, is bounded by the number of time steps of the data that were acquired.

By contrast, the most common approach for EEG signals is called *temporal ICA* (tICA) (Fig. 3a). Signals are decomposed at each time into *temporally* independent sources whose spatial projections (i.e. mappings onto the channels) are fixed across time. The maximum number of independent components is the number of EEG channels. Temporal ICA makes good sense for signals such as eye blinks, eye movements and muscle activity [2, 37–39], since the sources underlying these signals are spatially compact and independent from the rest of the activity picked up by the EEG.

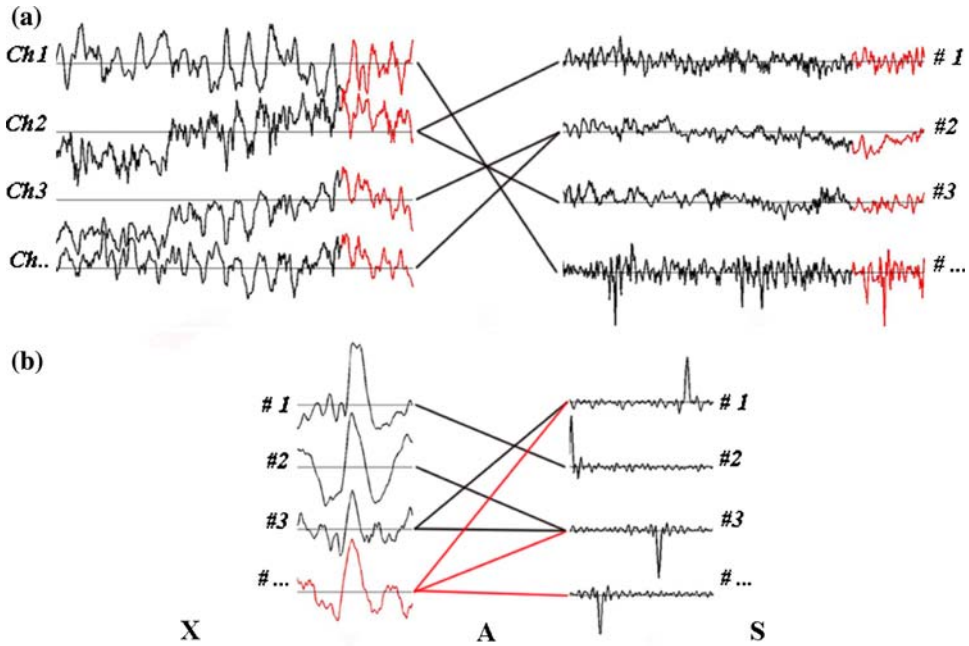

Fig 3 **a** Temporal ICA: EEG signals (**X**) picked up at some spatial points (Ch1, Ch2, …), can be represented as a linear mixture of independent components (**S**); when new data is measured, the independent components are updated correspondently in time. (b) Our proposed use of ICA: A collection of spikes (**X**) were decomposed into a mixture of independent components (**S**) through the mixing matrix **A;** when a new spike was added to the dataset of spikes, **S** was kept fixed and the mixing matrix **A** was updated. The four independent components represented here were a subset of the whole set of 105 independent components obtained from one of the cross validation split of our dataset

For our purposes, tICA was problematic, since we expected that more than one independent component accounts for the epileptic activity, and that different patients would have different foci in different locations, therefore having quite different underlying spatial maps on different electrodes.

We, therefore, designed what we believe to be a novel variant of ICA [7]. We expected the form of spikes to be relatively independent of the location of the electrode on which they appear, therefore allowing all the spikes to be lumped together. Then, like sICA, except with instances of spikes replacing voxels (Fig. 3b), we sought a representation in which the activations of the components are sparse and independent across the time steps in the spike window (Fig. 4). This method of representation is able to characterize abrupt changes in the amplitude of the signal, especially if they arise from changes in phase rather than changes in frequency. In this respect, Fourier-like or PCA-based representations perform less well, as they are less well temporally localized [4] (Fig. 5).

The novel aspect of our ICA variant arises from the fact that during testing, any potential spike was treated as an extra input into the model (Fig. 3b). The features for each spike (red links in Fig. 3b) were evaluated by projecting the independent components onto the spike itself (red waveform in Fig. 3b). This projection is in fact the set of mixing (weights) coefficients that, in the underlying ICA generative model, would allow the spike to be created from the ensemble of independent sources learned from the training set.

It is worth emphasizing that the independent sources were extracted just from the training database of spikes. This was done since we expected that there will be a bigger difference between the projection of spikes and non-spikes onto a representation that is precisely tailored to the vagaries of spikes themselves. The same argument applies potentially to other signals of interest such as eyeblinks, to which we also applied our method.

2.3 EEG training and validating data

Training was based on seven EEGs chosen from clinical recordings in subjects with confirmed epilepsy (Table 1). Data were collected using a 21-channel EEG system at a sample rate of 256 Hz, following the 10–20 system for the electrode positions [23] and a Micromed (Micromed, Italy) Brainquick EEG system utilizing the SystemPlus software suite. All the recordings showed some form of epileptic activity, either spikes or sharp waves (Fig. 1). Seven recordings (124 min total duration) were labeled by an experienced EEG technician and were used for both training and testing. The labels included the time of the centre of the epileptiform activity and the channels in which these occurred. The EEGs were filtered between 0.5 and 70Hz, and a notch filter at 50 Hz was applied. The EEG was displayed in common average montage.

Training and validation proceeded according to a leave-one-out cross-validation procedure on the seven recordings. In each case, testing was performed on one full recording and training was undertaken on a subset of the remaining six. This was repeated seven times. The training set comprised all the sharp and spike waves collected from the six recordings together with a separate, hand-picked,

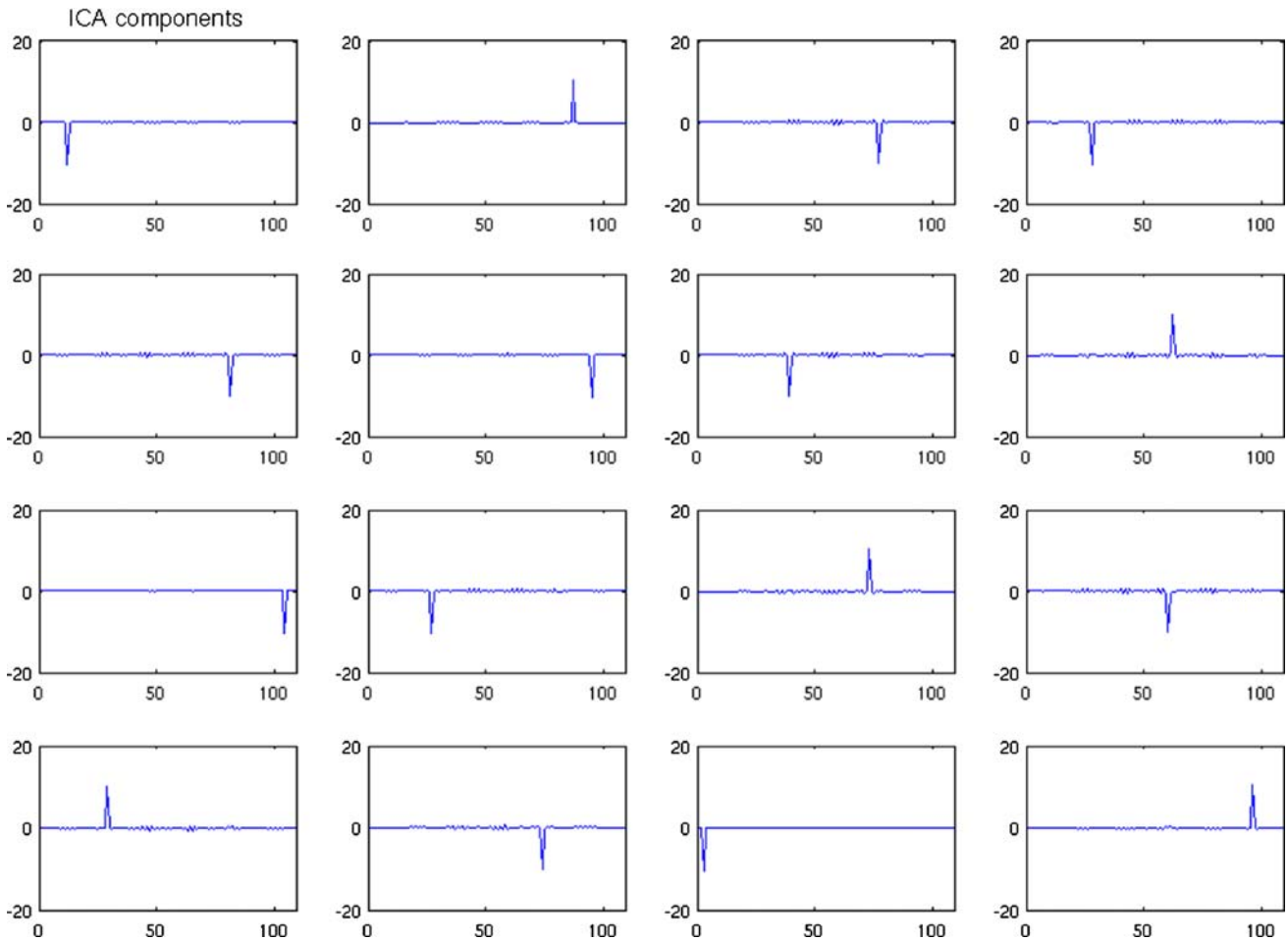

Fig. 4 Example of 16 independent components that provided the temporal basis used for feature extraction. These components are a subsample of the 105 independent components obtained in the training stage using six out of the seven EEG recordings. The components are all well localized in time as well as in frequency, which provides a good basis for representing the abrupt amplitude change of epileptic events

23 min subset of the six recordings. This subset was chosen to capture the range of variability in the data, whilst avoiding artefacts. The total number of sharp waves and spikes in each training set was between 499 and 611.

2.4 Description of the method of classification

In the training set, each spike was positioned in the centre of a window of fixed length. The window was $L = 109$ samples long, corresponding to about half a second. This length was chosen in order to include the longest sharp wave and also some background activity. The epileptiform activity was stored in a matrix $\mathbf{X}^s$ (of dimension $n^s \times L$), where n^s was the total number of epileptiform waves in the training set. $\mathbf{X}^s$ represents the first class in the problem. The background EEG waveforms were stored in a matrix $\mathbf{X}^b$ (of dimension $n^b \times L$), where n^b was the number of background EEG waveforms in the training set:

$$\mathbf{X}^k = \begin{pmatrix} X_1^k \\ X_2^k \\ \dots \\ X_n^k \end{pmatrix} \quad k = s,b \tag{5}$$

where the ith row of $\mathbf{X}^k$ is the generic waveform $x_i^k = (x_{i1}^k\ x_{i2}^k,\dots,x_{iL}^k)$, and where k can assume as many indexes as many classes to be discriminate by the algorithm.

For each waveform, we computed the maximum amplitude occurring within a sub-interval of the original temporal window. This subinterval consisted of 40 time frames starting at the 40th time point of each waveform. Without loss of generality we disregarded a priori all the waveforms that did not have this maximum value in the middle of the 109 time-frame window. This choice was made in order to reduce the computational time required by the algorithm, and because those criteria did not exclude any epileptic waveforms. The same principle applies to both training and testing of the algorithm.

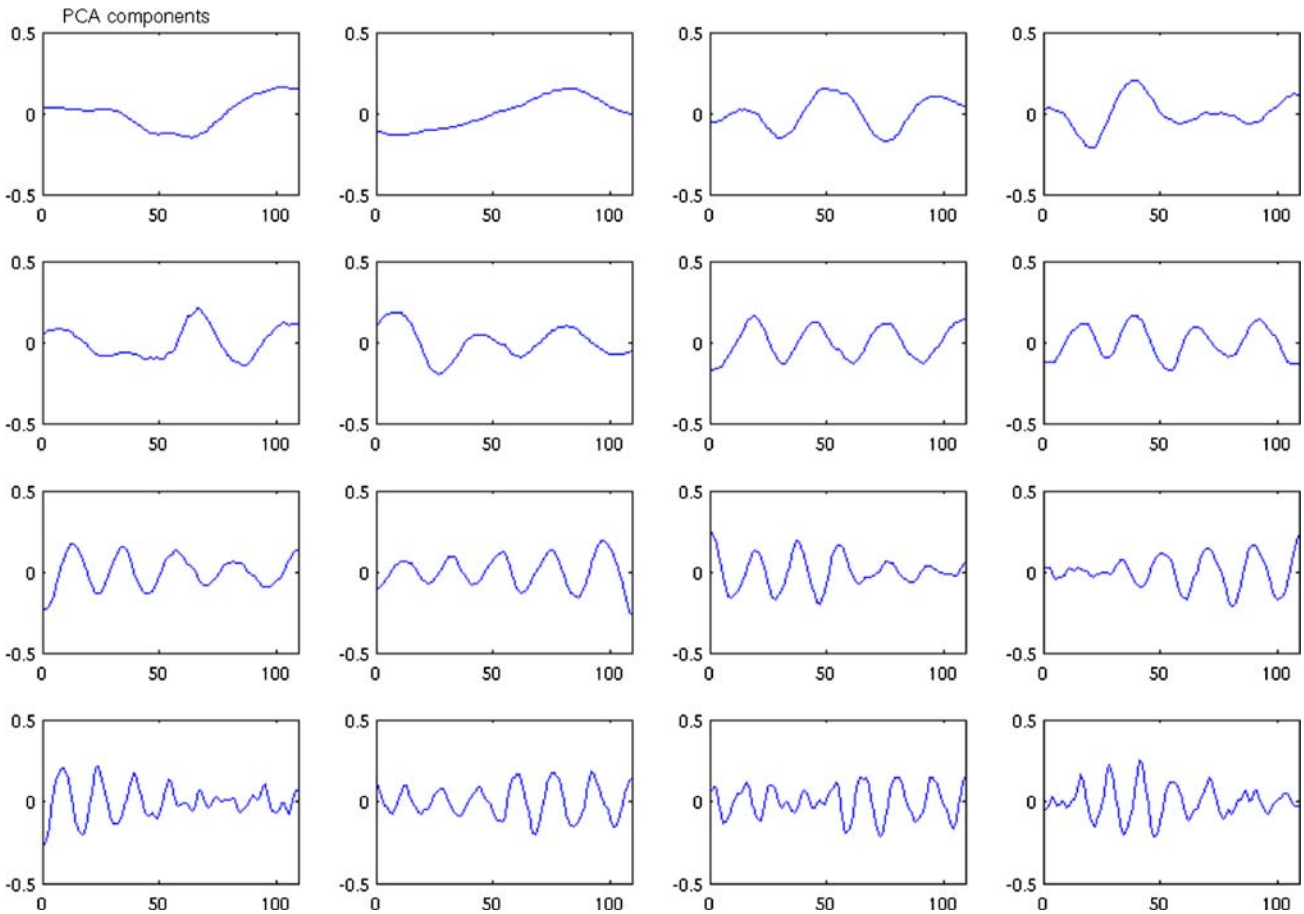

Fig. 5 Example of principal components extracted in the training stage using six out of the seven EEG recordings, ordered by descending variances. These components can be considered in the same terms as Fourier analysis, since each of them can be seen as representing a different frequency component spread across the whole temporal window. This set of representational coordinates was used to provide a point of comparison for the ICA-based coordinates

In the first step of the algorithm, independent components **S** were extracted from $\mathbf{X}^s$ according to the model expressed by (3). The input data for ICA were preprocessed by centering and whitening. The ICA decomposition was performed using Fastica [21], [20] and activation function $g(u) = u^3$. The independent components were fixed for both classes, and their projections onto $\mathbf{X}^k$ were treated as lists of features for each class. In the second step of the algorithm, we trained a quadratic classifier based on these projections [36].

In the test stage, a sliding window of length L was shifted sample by sample along each channel, and projections onto the independent components were computed for each EEG segment that met the above criteria.

For any of the cross-validation splits in the dataset, the maximum number of independent components was 105. Since the basis is undercomplete, computation of the pseudoinverse of the matrix **S** was required.

The algorithm was then generalized to detect a further class which comprised eye movements and eye blinks (eye artifacts; Fig. 2), as these artefacts appeared to provide the main source of classification errors in the two-class case. In the feature extraction stage, the three classes—epileptiform

Table 1 Dataset description

Diagnosis	# Epileptiform events	Duration (min)	Birth year
Idiopathic generalized epilepsy	10	5.76	1983
Left temporal partial lobe epilepsy	55	23.23	1973
Focal epilepsy	40	20.0	1956
Focal spike wave activity	17	4.17	1976
Temporal lobe epilepsy	122	20.01	1983
Generalized spike and wave and partial temporal epilepsy	265	24.36	1981
Generalized spikes and spikes and wave, burst of polyspikes and wave	112	23.48	1991

activity, eye artefacts and rest of the EEG—were all projected on the same single basis obtained from ICA using only the epileptiform activity. We generalized the quadratic discriminator to allow a multiclass classification, fitting a mixture of Gaussians model to the features of the different classes, and computing the posterior probabilities of each class [36].

The model was optimized by extracting the number of independent components that showed the best performance of the quadratic classifier across the seven splits. To this end, we reduced the input dimension n^s of $\mathbf{X}$ to between 1 and 105 dimensions using Principal Component Analysis before extracting the same number of independent components. In the following we will always refer to this algorithm as ICA-algorithm, either when it is based on PCA followed by ICA, or on ICA alone when no dimension reduction is needed.

We evaluated the performance based on the average ROC curve across the seven cross-validation splits. Given the multiple classes, we evaluated the ROC curve in the testing stage of each of the seven splits by fixing the prior probability for the eyeblink class in order to get the maximum sensitivity possible for detecting epileptiform activity, and tuning the prior probability between epileptiform and the background EEG classes. In the ROC curve evaluation, we collapsed all the marks across channels along one single channel and considered the marks within the L-window as one single event. A mark is labeled as true positive when at least one of the spikes at a certain time point is detected. This criterion makes it possible to make a fair comparison with the Reveal software (see below).

Two aspects of the ROC curve are of particular interest: the *area*, which measures the overall performance of the algorithm, and the *sensitivity* reached at the rate of six false positives (FPs) per minute. Since EEGs are usually displayed with 10 s per page, a higher rate of FPs per minute would require a revision of classifier marks for every page, and so would not save the time of the technician or electrophysiologist. In order to emphasize the difference between the three representations, we also report the accuracy obtained at a mid-range value of 65% sensitivity.

3 Comparison with other feature-based classifiers and Reveal

The performance of the ICA-based representation was compared with other feature-based classifiers, and the commercially available software Reveal.

We considered two possible choices of features, using each time the same classifier as described above. The first important control was done using the raw, temporal, representation of the input data themselves. The second comparison was performed using PCA.

In the first case, the number of features equals the temporal dimension of the input data, which is fixed to match the number of time frames defining each waveform. Therefore, this comparison was made only with complete ICA representation, in order to minimize the impact of the number of feature dimensions on the performance and allow a fair contrast (Fig. 6). The second comparison was performed using as features the PCA components rather than the ICA components. This second case allowed a comparison for any undercomplete choice of representation of the waveforms (Fig. 7). In all cases, we measured relative performance by computing and comparing the areas under the ROC curves for the various cases (Figs. 6, 7).

Reveal was also run on the same dataset. In this program, spikes, sharp waves, and spikes and waves are marked when they occur as single waveforms or in a run (focal or generalized or partially generalized). All these events are defined as spikes in the 'mark detector'. The mark is accompanied by a probability index, termed a perception level p ($0.1 < P < 1$), and the channel of the spike focus.

The perception level quantifies the degree of uncertainty about the marked waveform; the lower the perception value, the lesser the confidence in the detection.

To treat Reveal in the same manner as our algorithm, we considered its marks to be correct if they matched true spikes within a window of length L. Since, unlike Reveal, we did not include polyspikes in our classification, we disregarded any marks occurring within a polyspike. The perception levels were used to generate the ROC curves, where the maximum of the perception value corresponds to the minimum sensitivity of Reveal and vice-versa. That the

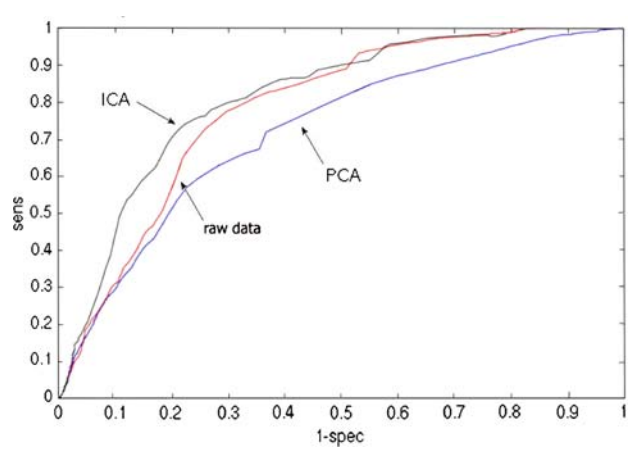

Fig. 6 Average ROC curve for the three methods: ICA and PCA with 105 components extracted, and raw data

perception levels are discrete makes for a rather rough estimate of the ROC curve (Fig. 8).

4 Results

When the ICA-algorithm was optimized for epileptiform detection, the best performance for all seven splits was for the case that just three components were extracted (Figs. 7, 8, 9).

The corresponding ROC curve (Fig. 8) gave a sensitivity of 65% (±22%) at a specificity of 86% (±7%). With a PCA algorithm based on three components, the specificity was 76%(±1%) at the same sensitivity (Fig. 8). Allowing any possible number of extracted components, the PCA-based classifier gave an underlying ROC curve area that was always lower than the ICA-algorithm (Fig. 7; $p < 0.05$). The three ROC curves obtained for our algorithm at 105 independent components extracted, 105 principal components and raw data gave respectively a specificity of 80% ± 12%, 64% ± 26% and 74% ± 7% all at a sensitivity of 65% (Fig. 6).

The ICA-algorithm when three components were extracted provided 17% ± 17% in sensitivity and 96% in specificity, at a rate of six false positives per minute (Fig. 8).

For Reveal, the highest average sensitivity was 30% and a specificity of 77%. In one recording, Reveal did not detect any spikes. Our algorithm showed a higher accuracy than Reveal only for sensitivities higher than 17% (Fig. 8). ROC curves areas for Reveal compared to our algorithm restricted to this highest sensitivity reached by Reveal were not significantly different (Student's t test).

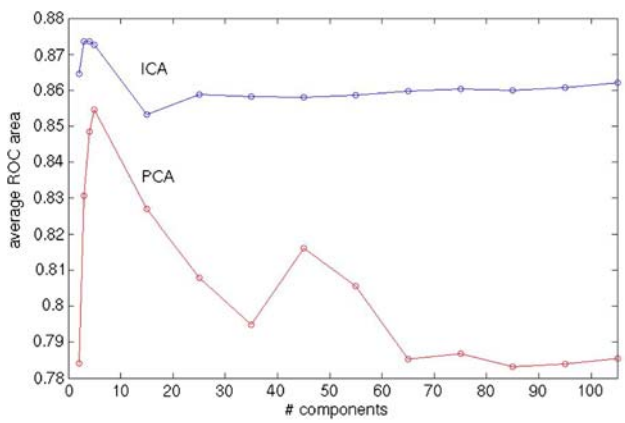

Fig. 7 Average ROC areas as a function of the number of independent components extracted for ICA and PCA-based algorithms; in both cases all the seven curves showed a maximum at three independent components

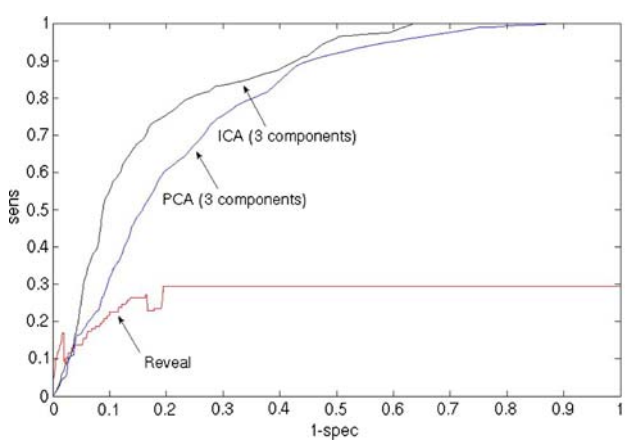

Fig. 8 Average ROC curve at three independent components extracted (best parameter choice) for ICA and PCA and average ROC curve obtained running reveal on the same dataset

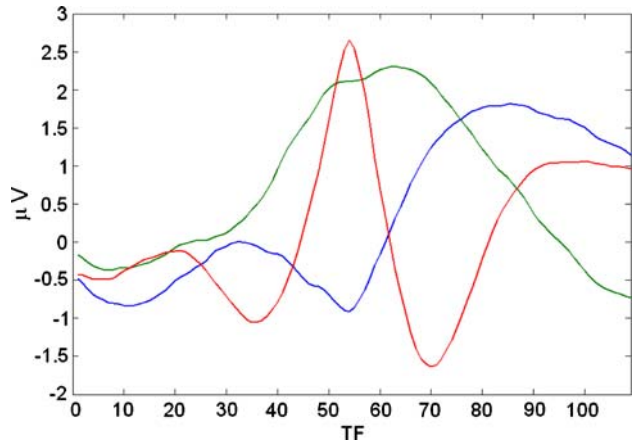

Fig. 9 Example of the three independent components (TF = time frames) extracted in the training stage after reducing data dimension to three through principal component analysis. This low number of independent components is actually that which maximizes the accuracy in the classification of epileptiform waveforms

5 Discussion

We have presented a two-stage algorithm for automatic classification of EEG signals. The first stage consisted of a novel application of an independent components analysis algorithm. This resulted in a set of new representational coordinates that were appropriate for discriminating spikes from other features of the EEG. The second stage consisted of a mixture of Gaussians classifier, based on these coordinates. The classifier was optimized to detect epileptiform activity, eye blinks and eye movement artefacts. A ROC curve analysis for epileptiform events showed that we can expect an average sensitivity of 65% at a specificity of 86%.

Model optimization led to the selection of only three independent components after dimensionality reduction by PCA. The three selected components are clearly related to the salient features that human experts report using to detect a spike or a sharp wave (Fig. 9). One component is clearly responsible for the peakedness of the waveform (red line), the green component reflects the duration of the epileptiform event, and the last blue component represents the rebound, which specially characterizes the spikes and wave. This quite straightforward interpretability of the selected features can be seen as indirect support for the validity of our data driven approach to feature selection based on independent component analysis. More independent components are expected to be selected when more patterns must be detected.

The algorithm has not been fully tested on a naive dataset, but our preliminary results, which are based on a cross validation procedure, appear promising for routine clinical use. The algorithm did not impose high computational demands. A recording of 20 min, which comprises 17 Mb of data, can be processed in about 10 min on a 2.8 GHz dual processor PC. In clinical practice, we expect this time to be practicable for routine daily use. Should the computational time become unfeasible when more classes need to be detected, more powerful computational resources could be considered. One avenue is use of parallel processing by the GRID. Using the Condor GRID resource available at our university with 920 PCs, processing time was reduced by a factor of 25 [12, 13].

It is very hard to compare our algorithm directly against most other suggested methods, since, unlike ours, the datasets on which they have been tested and validated are not in the public domain. We certainly observed even in our own datasets, that there can be quite large differences depending on patient and human expert. In contrast, some other reported performances are better than this, especially in the range of 6 false positives per minute. For instance the 'Gotman spike detector' [16] achieved a sensitivity of 76% with 5.2 FPs/min [40] and another study reported a sensitivity of 73% at 6.1 FPs/min, but our algorithm showed only 17% sensitivity in this range. More reliable are comparisons based on the same dataset. In this respect, we were able to run the commercial algorithm Reveal on our own data. Although this is reported as having a sensitivity and specificity of 89.9 and 99.6%, respectively (based on a large dataset including 40 subjects and 10 controls, which was independently labeled by five experts), when applied to our data, Reveal only achieved a sensitivity of 30% even at the lowest perception level. This prevented us from comparing our results at the level of sensitivity (>50%) usually reported in the literature [41]. Nevertheless, the specificity of Reveal at sensitivity lower than 17% was slightly better than that of our algorithm. We are confident that our data have been appropriately labeled, and have duly put it into the public domain. That our method allows multiple classes suggests that its performance could be further improved if other explicit sources of misclassification are included as extra classes in the classifier. This is a key direction for future work.

Other planned work includes validation with a larger naïve training and cross-validation set, particularly because of the apparent variability in the number of spikes in each dataset. This variability led to problems in setting the threshold consistently between the separate cross-validation splits of the data, an issue with which extra data should greatly help. The next task is to improve the classifier. Since rather few representational components turn out to be optimal, it is possible to visualize the representation of the spikes and non-spikes in the reduced-dimensionality ICA space. This visualization suggests that we might beneficially replace the quadratic classifier, which was chosen mainly for its simplicity, with a nearest neighbor classifier.

In the longer run, our intent is to generalize automatic classification to the other salient aspects of EEG signals. If this can be accomplished, then this could provide a novel tool, which could be used routinely for a quantitative fast analysis suitable for everyday EEG recording.

In conclusion, our novel ICA-based approach is quite accurate in detecting epileptiform events, and is a promising approach for generalizing automatic methods to the detection of other salient aspects of EEG signals.

References

1. Acir N, Guzelis C (2004) Automatic spike detection in EEG by two-stage procedure based on support vector machine. Comput Biol Med 34:561–575
2. Barbati G, Porcaro C, Zappasodi F, Rossini PM, Tecchio F (2004) Optimization of an independent component analysis approach for artefact identification and removal in magnetoencephalographic signals. Clin Neurophysiol 115:1220–1232
3. Bell A, Sejnowski T (1995) An information-maximization approach to blind separaton and blind deconvolution. Neural Comput 7:1129–1159
4. Bell AJ, Sejnowski TJ (1996) Learning the higher-order structure of a natural sound. Netw Comput Neural Syst 7:261–266
5. Binnie CD, Holder DS (1999) Electroencephalography. In: Vinken PJ, Bruyn GW (eds) Handbook of clinical neurology. Amsterdam
6. Calhoun VD, Adali T, Pearlson GD, Pekar JJ (2001) Spatial and temporal independent component analysis of functional MRI containing a pair of task-related waveforms. Hum Brain Map 13:43–53
7. De Lucia M, Fritschy J, Dayan P, Holder DS (2006) The classification of spikes in EEG recordings using features derived from ICA. In: IET 3rd international conference MEDSIP 2006. Advances in Medical, Signal and Information Processing, issue 520, p. 39
8. Diambra L, Malta P (1999) Nonlinear models for detecting epileptic spikes. Phys Rev E 59(1):929–937
9. Dumermuth G (1977) Fundamentals of spectral analysis in electroencephalography. In: A.Rèmond (ed), EEG informatics. A didactic review of methods and applications of EEG data processing. Elsevier, Amsterdam, pp 83–105

10. Faure C (1985) Attributed strings for recognition of epileptic transients in EEG. Int J Biomed Comput 16:217–229
11. Franaszczuk PJ, Bergey GK (1999) An autoregressive method for the measurement of synchronization of interictal and ictal EEG signals. Biol Cybern 81:3–9
12. Fritschy J, Horesh L, Holder DS, Bayford RH (2005) Using the GRID to improve the computation speed of electrical impedance tomography (EIT) reconstruction algorithms. Physiol Meas 26(2):209–215
13. Fritschy J, Horesh L, Holder DS, Bayford RH (2005) Applications of GRID in clinical neurophysiology and electrical impedance tomography of brain function. Stud Health Technol Inform 112:138–145
14. Goelz H, Jones RD, Bones PJ (2000) Wavelet analysis of transient biomedical signals and its application to detection of epileptiform activity in the EEG. Clin Electroencephalogr 31:181–191
15. Gotman J, Gloor P (1976) Automatic recognition and quantification of interictal epileptic activity in the human scalp EEG. Electroenceph clin Neurophysiol 41:513–529
16. Gotman J, Ives JR, Gloor P (1979) Automatic recognition of interictal epileptic activity in prolonged EEG recordings. Electroenceph clin Neurophysiol 46:510–520
17. Hastie T, Tibshirani R, Friedman J (2001) The elements of Statistical Learning: Data Mining, Inference and Prediction. Springer Series in Statistics
18. Hughes JR (1994) EEG in clinical practice, 2nd edn. Butterworth-Heinemann, Newton, MAJR. Hughes EEG in Clinical Practice
19. Hyvarinen A (1999) Survey on independent component analysis. Neural Comput Surv 2:94–128
20. Hyvarinen A (1999) Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans Neural Netw 10 (3):626–634
21. Hyvarinen A, Oja E (1997) A fast fixed-point algorithm for independent component analysis. Neural Comput 9(7):1483–1492
22. Hyvarinen A, Oja E (2000) Independent component analysis: algorithms and applications. Neural Netw 13:411–430
23. Jasper H (1958) The ten–twenty electrode system of the international federation. Electroencephalogr Clin Neurophysiol 10:371–375
24. Jung T, Makeig S, Mckeown MJ, Bell AJ, Lee T, Sejnowski TJ (2001) Imaging brain dynamics using independent component analysis. Proc IEEE 89(7):1107–1122
25. Kalayci T, Ozdamar O (1995) Wavelet preprocessing for automated neural network detection of EEG spikes. IEEE Eng Med Biol Mag 14(2):160–166
26. Kobayashi K, Akiyama T, Nakahori T (2002a) Systematic source estimation of spikes by a combination of independent component analysis and RAP-MUSIC. I: principles and simulation study. Clin Neurophysiol 113:713–724
27. Kobayashi K, Akiyama T, Nakahori T (2002b) Systematic source estimation of spikes by a combination of independent component analysis and RAP-MUSIC. II: preliminary clinical application. Clin Neurophysiol 113:725–734
28. Kobayashi K, James C, Nakahori T, Akiyama T, Gotman J (1999) Isolation of epileptiform discharges from unaveraged EEG by independent component analysis. Clin Neurophysiol 110(10):1755–1763
29. Latka M, Was Z, Kozik A, West B (2003) Wavelet Analysis of epileptic spikes. Phys Rev E 67:052902
30. McKeown MJ, Brown GG, Jung TP, Kindermann SS, Bell AJ, Sejnowski TJ (1998) Analysis of fMRI data by blind separation into independent spatial components. Hum Brain Mapp 6:160–188
31. McKeown MJ, Jung TP, Makeig S, Brown G, Kindermann SS, Lee TW, Sejnowski TJ (1998) Spatially independent activity pattern in functional MRI data during the stroop color-naming task. Proc Natl Acad Sci USA 95:803–810
32. Ossadtchi A, Baillet S, Moscher JC, Thyerlei D, Sutherling W, Leahy RM (2004) Automated interictal spike detection and source localization in magnetoencephalography using independent components analysis and spatio-temporal clustering. Clin Neurophysiol 115(3):508–522
33. Ozdamar O, Kalayci T (1998) Detection of spikes with artificial neural networks using raw EEG. Comput Biomed Res 31:122–142
34. Pham D-T, Garrat P, Jutten C (1992) Separation of a mixture of independent sources through a maximum likelihood approach. In: Proceedings of EUSIPCO, pp 771–774
35. Ramabhadran B, Frost JD Jr, Glover JR, Ktonas PY (1999) An automated system for epileptogenic focus localization in the electroencephalogram. J Clin Neurophysiol 16:59–68
36. Schlesinger MI, Hlavàc V (2002) Ten lectures on statistical and structural pattern recognition. Kluwer, Dordrecht
37. Shoker L, Sanei S, Wang W, Chambers JA (2005) Removal of eyey blinking artifact from the electro-encephalogram, incorporating a new constrained blind source separation algorithm. Med Biol Eng Comput 43:290–295
38. Vigario RN (1997) Extraction of ocular artefacts from EEG using independent component analysis. Electroencephal Clin Neurophysiol 103:395–404
39. Vigario R, Jousmaki V, Hamalainen M, Hari R, Oja E (1998) Independent component analysis for identification of artefacts in magnetoencephalographic recordings. Adv Neural Inf Process Syst 10:229–235
40. Webber WR, Litt B, Wilson K, Lesser RP (1994) Practical detection of epileptiform discharge (EDs) in the EEG using an artificial neural network: a comparison of raw and parameterized EEG data. Electroencephalogr Clin Neurophysiol 91(3):194–204
41. Wilson SB, Emerson R (2002) Spike detection: a review and comparison of algorithms. Clin Neurophysiol 113:1873–1881
42. Wilson SB, Scheuer ML, Emerson RG, Gabor AJ (2004) Seizure detection: evaluation of the Reveal algorithm. Clin Neurophysiol 115(10):2280–2291
43. Wilson SB, Turner CA, Emerson RG, Scheuer ML (1999) Spike detection II: automatic, perception-based detection and clustering. Clin Neurophysiol 110:404–411