Energy distribution evaluation using Renyi entropy measures with application in EEG data analysis

THEODOR D. POPESCU
National Instituite for Research and Development in Informatics
8-10 Averescu Avenue, 011455 Bucharest, ROMANIA

Abstract- The "corrected" EEG recordings, after artifact removing, may be the subject of further investigations, for example segmentation and energy distribution, resulting new informa- tion to be used for feature extraction, of great help for medical diagnosis. The paper presents a generally method for energy distribution evalua- tion using measures of R´enyi entropy. The pre- sented approach ensures the possibility of quan- titative analysis of the information contained in time-frequency distribution of EEG signals. The proposed procedure is applied with good results in the analysis of a sample lowpass event-related potentials (ERP) data, collected from 13 scalp and 1 EOG electrodes.

Keywords- Energy distribution, time- frequency analysis, R´enyi entropy, EEG data analysis.

1. Introduction

THE electroencephalogram (EEG) analysis, repre- senting the electrical activity of the brain, makes use of different tools, specific to signal processing. The results are used for diagnosis of diseases which involves the function of cortical neurons. The evolution of EEG signals, in the case of a disease, is presented under the form of abnormal frequencies or anomalies in the ampli- tude of EEG waves.

The signals are stored and processed using different techniques, able to point out some links between par- ticular frequencies or spectra, to specific characteristics of diseases. Some expertise and software are necessary for the interpretation of the analysis results obtained for EEG recorded data.

The problems that occur in the functioning of the brain produce a distinct pattern in EEG data, and this can be estimated making use of pattern analysis. The analysis process refers to the ability of the user to iden- tify different patterns in data, making the object of the analysis. It represented a fundamental tool in diagnosis of brain diseases, in the last years.

Sometimes, the EEG analysis uses different tech- niques for change detection in signals. The problem is to determine if a change produced and determine the time when it occurred. The specific techniques oper- ate in time, frequency or time-frequency fields and use measures of "distance", statistical inference, fuzzy logic, artificial intelligence, among others. These techniques could be applied on the original EEG data, or on the pre-processed data. The purpose is to amplify the the effect of the changes. Sometimes, post-processing tech- niques are applied on the results obtained in the phase of change detection.

The amplitude levels of the EEG signals, obtained in the primary processing stage, can be extracted and clas- sified, as a first step in decision process; unfortunately, these results could be affected by the noise. The time- frequency analysis of signals, used for energy distribution evaluation, even it involves a more computation effort, offers new facilities to determine change detection in- stants, [1]. Also, the algorithms for modeling of signals, using parametric model, prove useful in such analysis, even with some limitations.

The time-frequency analysis [2] provides a simpler way for analysis and evaluation of non-stationary EEG signals; it refers to energy distribution analysis in pre- determined time and frequency areas [1]. The mea- sures based on entropy, make use of Kullback-Leibler and Rényi "distances", as well as Jensen difference [3] and offer new facilities in this field.

The problem of energy distribution of EEG signals make the object of many papers: [4], where the authors use a wavelet-neural network classifier, [5], with applica- tion in diagnosis, based on energies of EEG subbands, making use of discrete wavelet transform and support vector machine, [6] with an application of entropies in automated diagnosis of epilepsy, [7], energy feature ex- traction of EEG signals with a case study and [8], an application of the short-term Rényi entropy to estimate the number of EEG signal components, among others.

The paper presents some techniques, which include time-frequency analysis and energy distribution, as well as the use together, assuring a framework for EEG data analysis, resulting new information to be used for feature extraction. Section 2 refers to time-frequency analysis,

offering an overview of the techniques, with fundamental principles, advantages and disadvantages. In Section 3, we review some existing measures for energy distribution of signals, and insist on Rényi entropy measures, able to assure the possibility of quantitative analysis of of the information, contained in time-frequency distribution of EEG signals. Finally, Section 4 presents a case study having as subject the analysis of a sample lowpass event -related potentials (ERP) data, collected from 13 scalp and 1 EOG electrodes.

2. Time-frequency analysis of EEG signals

Time-frequency analysis proves to be useful in the analysis, processing and parameter estimation of signals, like EEG, where the spectral content of these signals changes in time. The basic objective of TFA is to find a function able to describe energy density of the signal like a function of time, t, and frequency, ω, simultaneously.

The time-frequency representations (TFRs) of a signal are classified depending of the analysis approaches [9]. So, in one approach, the signal, making the object of the analysis, uses time-frequency (TF) functions, resulting after translating, modulating and scaling of a basis function with a definite time and frequency location. So, the TFR, for signal $x(t)$, is given by

$$TF_x(t,\omega) = \int_{-\infty}^{+\infty} x(\tau)\phi_{t,\omega}^*(\tau)d\tau = < x, \phi_t, \omega >, \quad (1)$$

in which $\phi_{t,\omega}$ are the basis functions (the TF atoms) and $*$ is the complex conjugate. The $\phi_{t,\omega}$ functions are square integrable, with finite energy [10]. To this category belong short-time Fourier transform (STFT) [11], wavelets [10], [12], and matching pursuit algorithms [10], [13], among others.

A second category of time-frequency distributions (TFD) is given by Cohen's shift invariant class distributions, where the TFR is represented by a kernel function [2]. Because such of distributions will be used in the case study in Section 4, having as subject energy distribution evaluation in EEG signals, we give some details in the following.

As we mentioned before, in the case of this approach, the TFR is characterized by a kernel function, which the properties reflected by the constraints on the kernel, assuring desirable properties [2]. These TFRs can be described by

$$TFD_x(t,\omega) = \quad (2)$$

$$= \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x(u+\frac{\tau}{2})x^*(u-\frac{\tau}{2}) \times$$

$$\times \phi(\theta,\tau) \exp^{-j\theta t - j\tau\omega + j\theta u} du d\tau d\theta$$

with $\phi(\theta,\tau)$ a two-dimensional kernel function, ensuring the properties of the representation. In this case, the commonly used distributions for TFDs are Wigner distribution, Choi-Williams distribution, and spectrogram [2].

In selecting of a such distribution, to be used in practice, a great attention must be paid to its properties. The well-known spectrogram (SP), in general a distribution free of cross-terms, raises issues concerning the undesirable trade-off between the time and frequency resolution, see [14], [15] and [16]. The Wigner-Ville distribution (WVD) assures a high time-frequency resolution, but the presence of cross-terms raises issues in practice [14]. The problem of cross-terms removing, but keeping a high time-frequency resolution, is solved by other TFDs. So, the Choi-Williams distribution (CWD) [15], [17] and the binomial distribution (BD) [18] overcomes the limitation of WVD to a large extent, but it is lost the time-frequency resolution. The BD overcomes these problems, but some small synchronization and cross-terms could be present. Taking into account the above remarks, the BD appears to be a better choice in EEG signal analysis, solution used in the case study.

The last two distributions are known as the so-called reduced interference distribution (RID) and they also belong to the Cohen's class; RID is by itself an extension of the WVD [19].

The energy preservation and the marginals are the most desired properties for TFDs and are given as follows; they are satisfied when $\phi(\theta,0) = \phi(0,\tau) = 1 \forall \theta, \tau$ /citeCOH1.

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} TFD_x(t,\omega) dt d\omega = \quad (3)$$

$$= \int_{-\infty}^{+\infty} |x(t)|^2 dt = \int_{-\infty}^{+\infty} |X(\omega)|^2 d\omega$$

$$\int_{-\infty}^{+\infty} TFD_x(t,\omega) d\omega = |x(t)|^2, \quad (4)$$

$$\int_{-\infty}^{+\infty} TFD_x(t,\omega) dt = |X(\omega)|^2$$

Starting from the formulas (3) and (4), an analogy between TFD and the probability density function (pdf) of a two-dimensional random variable can be established [20].

As it is know, all TFDs tend to the same goal, but each has to be interpreted, according to its own properties: important interference terms, only positive terms, perfectly localized on particular signals, etc. Starting from the knowledge of these properties, information extraction has to be done carefully. It is desirable to use a distribution that can reveal the features of the signal as clearly as possible.

3. Energy distribution of EEG signals

One of the simplest procedure for evaluation of energy concentration is via time-frequency analysis. The idea is to evaluate the concentration of energy at different time instant or frequency band, or in some particular time and frequency region. In [9] is given an overview of recent advances for time-frequency representation using energy concentration, while a review of some existing measures

with their comparison makes the object of [1]. Also, some measures for distribution concentration for mono-component signals are presented in [21]. Other measures for time-frequency distribution concentration are defined by Jones and Parks [22], for more complex signals, and by Williams et al. [23], [24] and Flandrin [25], in the form of the distribution norms ratio and Rényi entropy. The distribution energy was also used for optimal kernel distribution design [26], [27], [28]. All these measures are based on the distribution norms, resulted as sums of distribution values at a power greater than one.

Various and efficient modifications of energy distribution have been proposed in [22] and [23]. The distribution norm has been divided by a lower-order norm, while some strict constraints were imposed on the kernel forms in [26]. In some cases, the normalized forms of the measures are not quite appropriate for the cases where there are some components of approximately equal energies, whose concentration are very different. A local concentration measure is proposed in [22], but it increases the computational effort. Some of these difficulties are overcome by the measures for distributions concentration presented in [1]. A briefly review of Rényi entropy measures, used to describe the amount of information contained in time-frequency distribution, as an approach for the EEG signals analysis, is given in the following.

A new entropy measure can be evaluated, to extract the information contained in a given position of $t = n$, after the local frequency content resulted after time-frequency distributions evaluation.

The class of Rényi entropy measures has been introduced in by Williams at al. [23], [24], with some contributions of Flandrin et al. [25], which established the properties for this measure. In [29] can be found a study of the properties and possible application of the Rényi measures, as well as mathematical foundations for quadratic TFRs. Starting from these results have been developed new measures used for signal information estimation and its complexity in time-frequency plane. The Rényi entropy measures are characterized by some interesting and useful properties, that make them natural for time-frequency analysis [29].

The Rényi information measures make the object of many papers. So, in [30] is described an application of Rényi information measures for spectral change detection without any assumption about the input; they are based on the evaluation of information measures in time-frequency plane, in particular the spectrogram. The Rényi entropy measures resulted as an extension of Shannon entropy and are well-suited as a support for analysis, segmentation and classification algorithms. Some entropy functionals resulted from Rényi information divergence are given in [31].

For a time-frequency distribution, $P_x(n, k)$, the Rényi entropy measure has the form:

$$R_\alpha = \frac{1}{1-\alpha} \log_2 \left(\sum_n \sum_k P_x^\alpha(n, k) \right) \quad (5)$$

where n is the time variable and k the frequency variable,

with $\alpha \geq 2$ recommended for distribution time-frequency measures, [25]; n and k are discrete values. For $\alpha = 2$, oscillatory cross-terms would increase the energy, leading to false conclusion. In the case $\alpha = 3$ the detection fails to point out the presence of oscillatory cross-terms. Starting from these remarks have been introduced the normalized Rényi entropy measures. This normalization is done in various ways, resulting different measure definitions. [25] and [1] propose some normalization schemes for Rényi entropy, described below:

- *Normalization with the signal energy* produces

$$RE_\alpha = \frac{1}{1-\alpha} \log_2 \left(\frac{\sum_n \sum_k P_x^\alpha(n, k)}{\sum_n \sum_k P_x(n, k)} \right) \quad \text{with } \alpha \geq 2 \quad (6)$$

This measure is quite similar with the non-normalized measure form, excepting its magnitude. This normalization offers the possibility for comparison of different distributions.

- *Normalization with the distribution volume,*

$$RV_3 = -\frac{1}{2} \log_2 \left(\frac{\sum_n \sum_k P_x^3(n, k)}{\sum_n \sum_k |P_x(n, k)|} \right) \quad (7)$$

If the distribution includes oscillatory terms, after summing them, in absolute value, the large cross-terms will decrease measure RV_α. This highlights smaller concentration, due to cross-terms appearance. The volume-normalized form of measure has been used in [23]. Other ways of normalization are presented in [25].

- *Quantum normalization*

In this case, the time-frequency distribution, at a given instant $t = n$, is assimilated to a wave function and in the general case, $\alpha = 3$ [32], results

$$R_3 = -\frac{1}{2} \log_2 \left(\sum_n \sum_k P_x^3(n, k) \right) \quad (8)$$

The normalizing process affects exclusively to index k, when the operation is restricted to a single position n to satisfy the condition $\sum_k P_x(n, k) = 1$ in such position.

The measure (8), for a given n, is rewritten as follows:

$$R_3(n) = -\frac{1}{2} \log_2 \left(\sum_k P_x^3(n, k) \right) \quad (9)$$

The normalization proposed in [32] resulted to be most suitable in signal analysis. The values of $R_3(n)$ depend upon the size N of the window in (9) and it can be shown that they are within the interval $0 \leq R_3(n) \leq \log_2 N$. Hence, the measure can be normalized by applying $\hat{R}_3(n) = R_3(n)/\log_2 N$.

Concerning the selection of Rényi entropy order α it depends to a great extend on the statistical distribution of the analyzed data. The order $\alpha = 3$ offers a good sensitivity for a great diversity of distributions, experimental proven, without need to remove some data, especially in complex optimization problems. This choice proved good results in the investigated case studies. Also, in [33], it is recommended and used with good results, the order 3 of Rényi entropy. Some authors use and investigate Rényi entropy of order 2, which in this case, is reduced to Shannon entropy.

4. Case study
4.1 EEG data

The data used in this case study represent an EEG time series collected from 13 scalp and 1 EOG electrodes. For artifact removing from data, has been used Independent Component Analysis (ICA) with high-order statistics [34], applied to sample lowpass ERP data for 2 epochs and 312 frames per data epoch, at 312.5 Hz sampling rate. The data are used in many case studies (see [35], among others).

The "corected" EEG data for 2 conditions, after artifact removing, are shown in Fig. 1 and Fig. 2 for the channels 1-7 (Fz, Cz, Pz, Oz, F3, F4, C3) and for the channels 8-14 (C4, T3, T4, P3, P4, Fpz, EOG), respectively.

4.2 Time-frequency analysis

The time-frequency analysis of the "corrected" EEG signals has been performed using the reduced interference distribution (RID) [19]; see Fig. 3 and Fig. 4, for channels 1-7 and channels 8-14, respectively. As we mentioned in Section 2, RID represents an extension of the Wigner-Wille distribution [2], belonging to the Cohen's class of distribution. It was computed with a kernel based on the Hanning window, for the number of frequency bins, Nf = 624, identical with the time instants, time smoothing window, Lg = 204, frequency smoothing window, Lh = 512, and a threshold of 2%.

4.3 Energy distribution

The time-frequency analysis results, presented above, have been used for evaluation of energy concentration in EEG signals and feature extraction, at certain time or frequency values, or more generally, in some time and frequency region of interest. In this case, the Rényi entropy proved to be a measure of information contained in time-frequency distribution [3].

Starting from time-frequency analysis results, we present in Fig. 5 and Fig. 6, the normalized Rényi information, for channels 1-7 and channels 8-14, respectively, as a measure of energy concentration for a window with length $N = 32$ and a constant bias added to signal of 7.5.

The normalized Rényi information for ERP filtered data without artifacts can make the object of segmentation using parametric models and an optimal procedure [36], [37], resulting new information for medical diagnosis. A such procedure, operating on the new space of
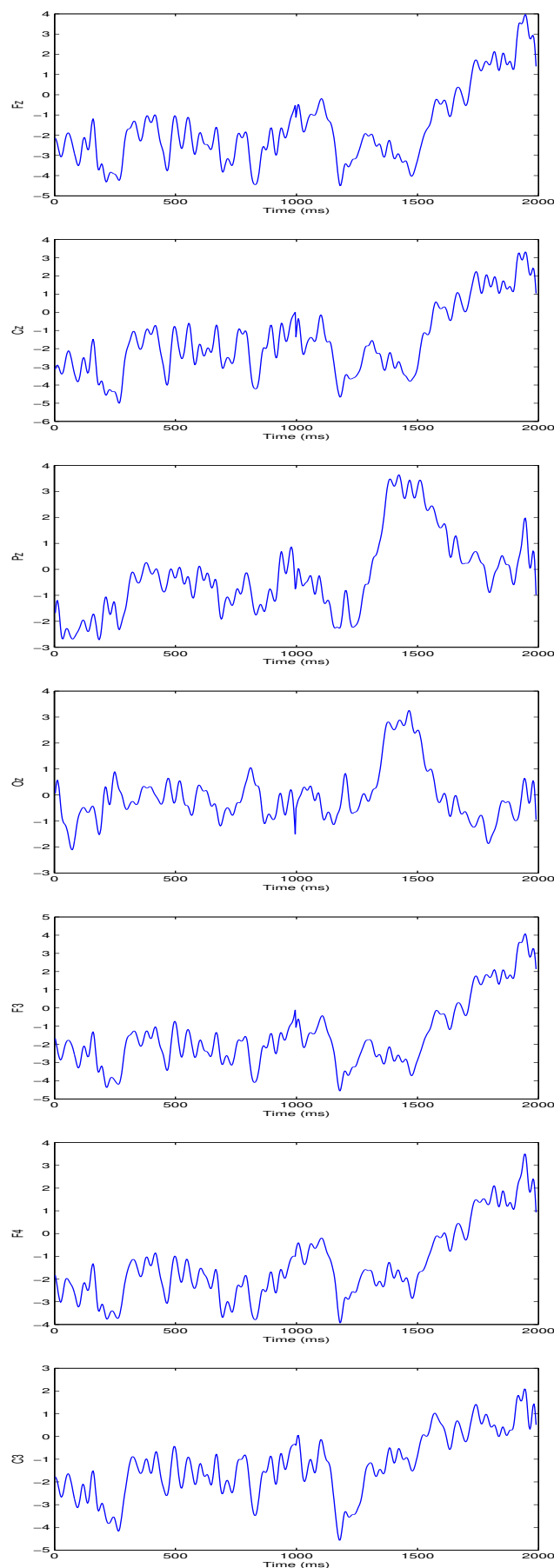

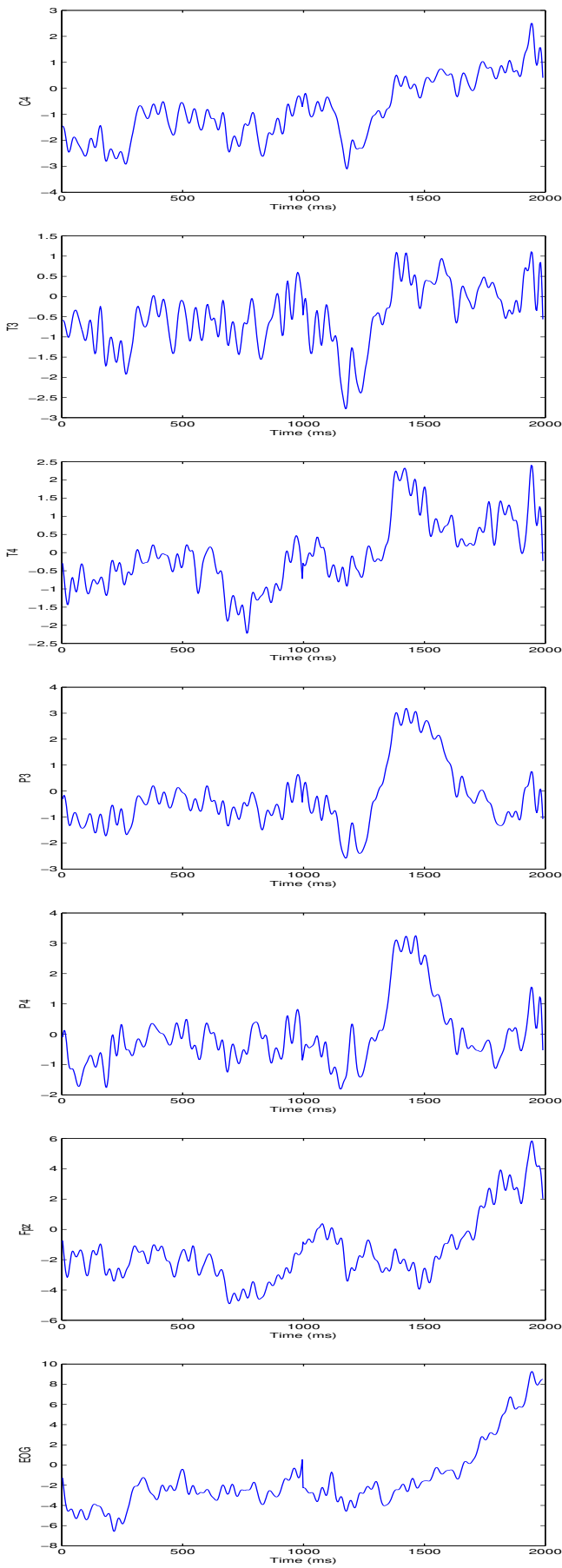Fig. 1: ERP filtered data after artifact removing for 2 epochs, channels 1-7

Fig. 2: ERP filtered data after artifact removing for 2 epochs, channels 8-14
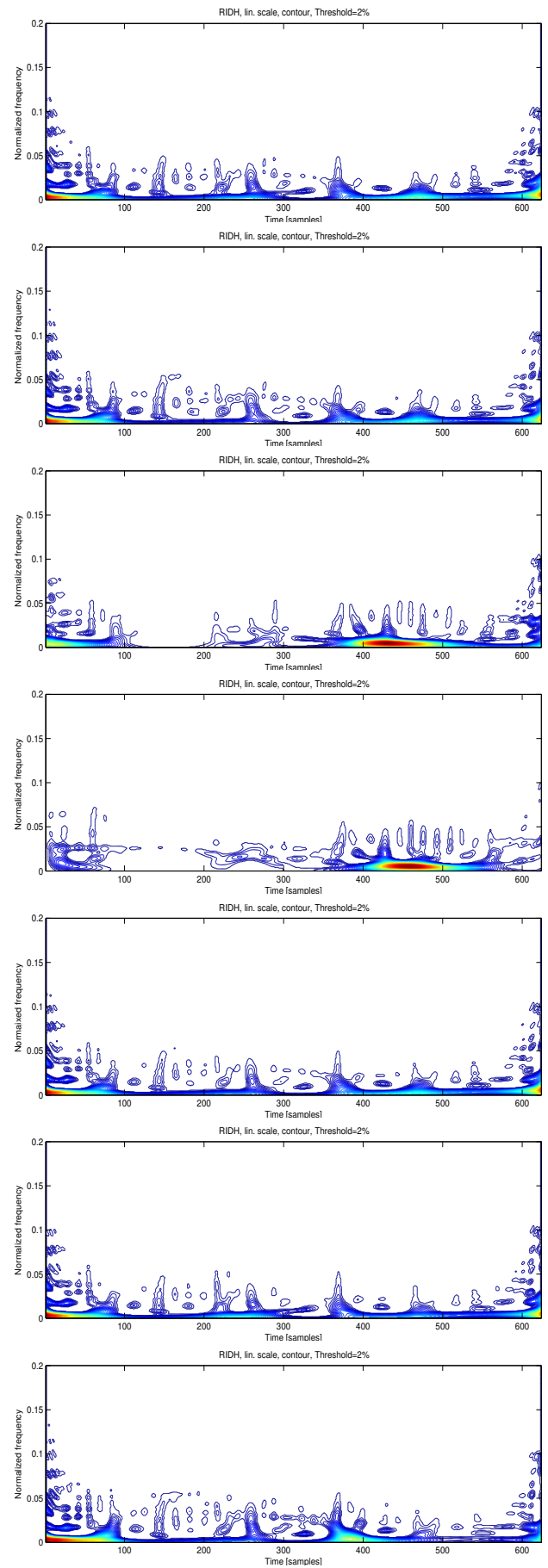
Fig. 3: Reduced interference distribution for ERP filtered data without artifacts: channels 1-7 (2 epochs)
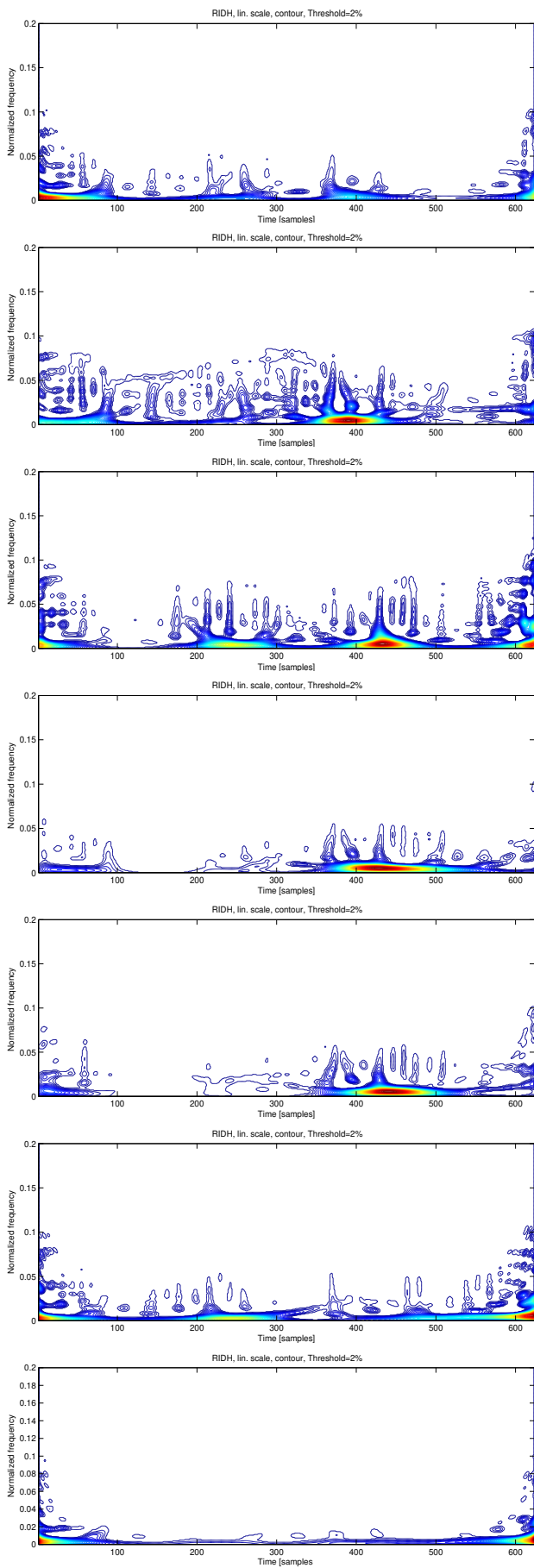
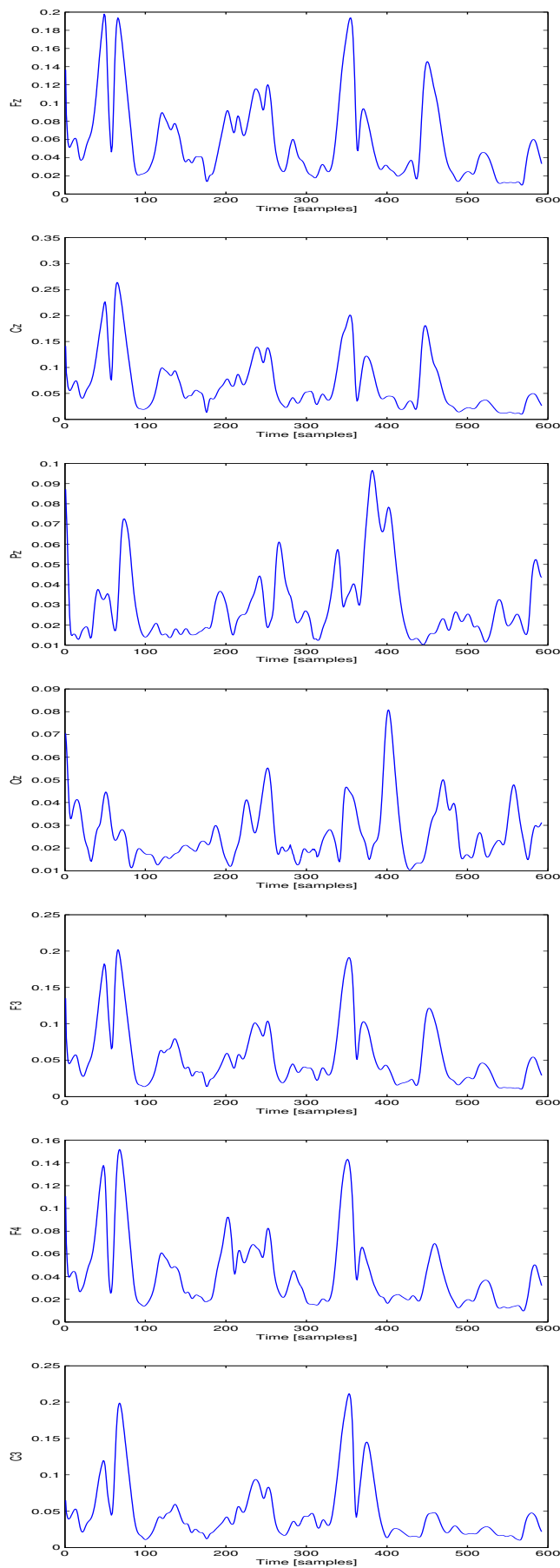Fig. 4: Reduced interference distribution for ERP filtered data without artifacts: channels 8-14 (2 epochs)

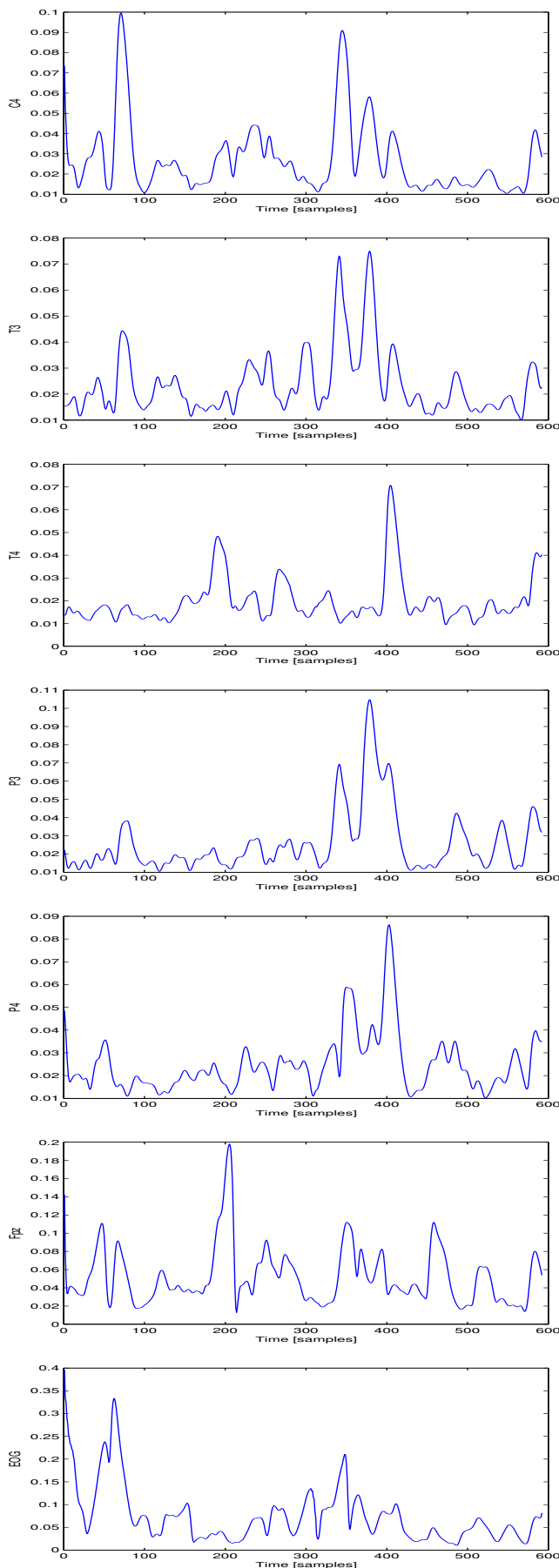Fig. 5: Normalized Rényi information for ERP filtered data without artifacts: channels 1-7 (2 epochs)

Fig. 6: Normalized Rényi information for ERP filtered data without artifacts: channels 8-14 (2 epochs)

decision, provided by Rényi entropy, assures more robust change detection in EEG signal analysis, than in the case of change detection in original EEG signals, after artifact removing.

5. Conclusions

A generally method for energy distribution evaluation using Rényi entropy measures, able to quantitative analysis of the information contained in time-frequency distribution of EEG signals, is presented. The normalized Rényi information for ERP filtered data without artifacts can make the object of further investigation, offering new information for medical diagnosis. The proposed procedure is applied with good results in the analysis of a sample lowpass event -related potentials (ERP) data, collected from 13 scalp and 1 EOG electrodes. The approach offers a simpler analysis and interpretation of the EEG signals behavior, providing new solution in investigation of cortical processes.

Acknowledgment

References

[1] L. Stankovic, "A measure of some time-frequency distributions concentration", Signal Processing, pp. 621-631, 2001.

[2] L. Cohen, Time-Frequency Distribution, Prentice Hall, New York, 1995.

[3] S. Aviyente, "Information processing on the time-frequency plane", Proc. IEEE International Conference Acoustics, Speech, and Signal Processing (ICASSP '04), pp. 617-620, 2004.

[4] I. Omerhodzic, A. Avdakovic, A. Nuhanovic, K. Dizdarevic, "Energy distribution of EEG Signals: EEG signal wavelet-neural network classifier", World Academy of Science, Engineering and Technology, vol. 61, pp. 1190-1195, 2010.

[5] J. E. Jacob, G. K. Nair, T. Iype, A. Cherian, "Diagnosis of encephalopathy based on energies of EEG subbands using discrete wavelet transform and support vector machine", Neurol Res Int., Art. 1613456, 2018.

[6] U. R. Acharya, H. Fujita, V. K. Sudarshan, S. Bhat, J. E. W. Koh, "Application of entropies for automated diagnosis of epilepsy using EEG signals: A review", Knowledge-Based Systems, vol. 88, pp. 85-96, 2015.

[7] J. Li, S. Sun, "Energy feature extraction of EEG signals and a case study", 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hong Kong, China, pp. 2366-2370, 2008.

[8] J. Lerga, N. Saulig, V. Mozetic, R. Lerga, "Number of EEG signal components estimated using the

short-term Renyi entropy", 2016 International Multidisciplinary Conference on Computer and Energy Science (SpliTech), Split, Croatia pp. 1-6, 2016.

[9] E. Sejdić, I. Djurović, J. Jiang, "Time-frequency feature representation using energy concentration: An overview of recent advances", Digital Signal Processing, pp. 153-183, 2009.

[10] S. G. Mallat, A Wavetet Tour of Signal Processing, Second Ed., Academic Press, 1999.

[11] K. Gröchenig, Foundations of Time-Frequency Analysis, Birkhäuser, 2001.

[12] I. Daubechies, Ten Lectures on Wavelets, SIAM, 1992.

[13] S. G. Mallat, Z. Zhang, "Matching pursuits with time-frequency dictionaries", IEEE Trans. Signal Process., pp. 3397-3415, 1993.

[14] B. Boashash, Time frequency signal analysis and processing: A compressive reference, Elsevier, 2003.

[15] L. Cohen, Time frequency distribution: A review, Proc. of IEEE, pp. 941-981, 1989.

[16] P. Flandrin, Temps-fréquence, Hermes, 1998.

[17] F. Hlwatsch, G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations", IEEE Signal Processing Magazine, pp. 21-67, 1992.

[18] M. Hussain, B. Bouashash, "Multi-component IF estimation", Proc. of the 10th IEEE Workshop on statistical signal and array processing - SSAP, pp. 559-563, 2000.

[19] J. Jeong, W. J. Williams, "Kernel design for reduced interference distributions", IEEE Transactions on Signal Processing, pp. 402-412, 1992.

[20] S. Aviyente, "Toward a theory of information processing on the time-frequency plane", Proc. SPIE 5559, Advanced Signal Processing Algorithms, Architectures, and Implementations XIV, 2004.

[21] L. Cohen, "Distributions concentrated along the instantaneous frequency", SPIE, Adv. Signal Process. Algebra Arch. Imp., pp. 149-157, 1990.

[22] D. L. Jones, T. W. Parks, A high resolution data-adaptive time-frequency representation", IEEE Trans. Signal Process, pp. 2127-2135, 1990.

[23] T. H. Sang, W. J. Williams, "Rényi information and signal dependent optimal kernel design", Proc. of the ICASSP, pp. 997-1000, 1995.

[24] W. J. Williams, M. L. Brown, A. O. Hero, "Uncertainity, information and time-frequency distributions", SPIE Adv. Signal Process. Algebra Arch. Imp., pp. 144-156, 1991.

[25] P. Flandrin, R. G. Baraniuk, O. Michel, "Time-frequency complexity and information", Proc. of the ICASSP, pp. 329-332, 1994.

[26] R. G. Baraniuk, D. L. Jones, "A signal dependent time-frequency representation: Optimal kernel design", IEEE Trans. Signal Process., pp. 1589-1602, 1993.

[27] R. G. Baraniuk, D. L. Jones, "Signal-dependent time-frequency analysis using radially Gaussian kernel", IEEE Trans. Signal Process., pp. 263-284, 1993.

[28] D. L. Jones, R. G. Baraniuk, "An adaptive optimal-kernel time-frequency representation", IEEE Trans. Signal Process, pp. 2361-2372, 1995.

[29] R. G. Baraniuk, P. Flandrin, A. J. E. M. Janssen, O. J . J. Michel, "Measuring time-frequency information content using the Rényi entropies", IEEE Transactions on Information Theory, pp. 1391-1409, 2001.

[30] M. Liuni, A Robel, M. Romito and X. Rodet, "Rényi information measures for spectral change detection", CASSP, Prague, Czech Republic hal-01161299, 2011.

[31] J. F. Bercher, "On some entropy functionals derived from Renyi information divergence", Information Sciences, Elsevier 178 pp. 2489-2506 hal-00276749, 2008.

[32] R. Eisberg, R. Resnick, Quantum Physics, Wiley, 1974.

[33] G. Yu, "A concentraed time-frequency analysis tool for bearing fault diagnosis", IEEE Transactions on Instrumentation and Measurement, vol. 69, pp. 371-381, 2020.

[34] Th. D. Popescu, "Artifact removing from EEG recordings using independent component analysis with high-order statistics", International Journal of Mathematical Models and Methods in Applied Sciences, vol. 15, pp. 76-85, 2021.

[35] A. Delorme, S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis", Journal of Neuroscience Methods, vol. 134, pp. 9-21, 2004.

[36] Th. D. Popescu, "Signal segmentation using changing regression models with application in seismic engineering", Digital Signal Processing, pp. 14-26, 2014.

[37] Th. D. Popescu, "Signal segmentation using maximum a posteriori probability estimator with application in artifact "corrected" EEG data", International Journal of Circuits, Systems and Signal Processing, vol. 15, pp. 1336-1345, 2021.

Theodor D. Popescu was born in Rosiori de Vede, Romania, on July 4, 1949. He received his M.Sc. degree and Eng.Sc.D. (PhD) degree, both in Automatic Control, from "Politehnica" University of Bucharest in 1972 and 1983, respectively. He has been a senior research scientist with the National Institute for Research and Development in Informatics, ICI Bucharest, since 1972. His research activities are focused on the fields of system identification and parameter estimation, time series analysis, signal processing, change detection and diagno-

sis, independent component analysis, blind source separation, time-frequency analysis, biomedical signal analysis. He is author or co-author of more than 150 papers and of five books. He was awarded with "Tudor Tanasescu" Prize of the Romanian Academy in 1995, for his works in the field of detection of changes in dynamics of signals and systems and General Association of Engineers in Romania Prize in 2000 in the field of Information Technology. He was granted by The Royal Society (1992), DAAD (1992, 1995, 1999), JSPS (1996, 1999), Australian Research Council (2002), CERGE-EI (2002). He held conferences and seminars in the Institute of Systems Science, Beijing, Institute of Technology, Beijing, University of Lancaster, Wuppertal University, Kyushu Institute of Technology, Oita University, Saga University, Tottori University, Kyoto Institute of Technology, CERGE-EI Prague and Spatial Co. Kuwait. He has been a Senior Member of IEEE since 2002.