THE ANNALS OF "DUNĂREA DE JOS" UNIVERSITY OF GALATI
FASCICLE III, 2017, VOL. 40, NO. 2, ISSN 2344-4738, ISSN-L 1221-454X
ELECTROTECHNICS, ELECTRONICS, AUTOMATIC CONTROL, INFORMATICS

EVALUATION OF TWO ALGORITHMS FOR CHANGE DETECTION BASED ON VIBRATION SIGNALS PROCESSING

Dorel Aiordachioaie

Dunarea de Jos University of Galati, Electronics and Telecommunication Department, Electronics, Information Technology and Communications Research Center (CCETIC), Domneasca-47, Galati-800008, Romania

Abstract: Change detection and diagnosis are important activities and research directions, in the field of system engineering and conditional maintenance of the equipment and industrial processes. The processed signals are coming from vibration generated by incipient faults in mechanical structures, e.g. bearings. Classical algorithms based on various version of CUSUM do not have enough performances to use intensively in real industrial application. The present work considers two new algorithms for change detection working on real industrial data of radial bearings. One is based on classical CUSUM criterion applied to the Renyi entropy. The second one is based on energy processing distributed over time-frequency region. The algorithms are tested on real recorded data. The results indicate good behavior and performance of the proposed algorithms, and define the rationale to implement them in commercial software product for change detection and diagnosis.

Keywords: signal processing, change detection, fault detection, Renyi entropy, vibration, time-frequency transform.

1. INTRODUCTION

In the field of change detection and process diagnosis, two paradigms are commonly used. The first one is based on the equations of the model. A change in the evolution of the signals of the observed process and those of the model is the first step in the detection of a change. The second approach is based on statistical signal processing of the signals coming or extracted from the studied process. The signals could be processed directly by computing various statistic moments and criteria or are processed to identify their mathematical models. The above approaches are well explained in various books and articles as: (Patton, 1989; Gertler, 1998; Isermann, 1997) for the general problem of change detection and diagnosis; (Chen and Patton, 1999; Sobhani-Tehrani and Khorasani, 2009; Venkatasubramanian, *et al*, 2003) for change detection based on process modeling and fault identification; (Gustafsson, 2001; Basseville, 1997; Mangoubi, 1998) for statistic signal processing.

The highlighted classical methods have some limitations in the case of complex signal and processes. The new methods consider space changes for decision, new entropies measures and advanced signal processing methods, as those based on time-frequency approach.

The present work considers three algorithms for change detection, namely: CA (Classic algorithm based on CUSUM method); AA1 (Advanced algorithm based on CUSUM applied to the Renyi entropy); AA2 (Advanced algorithm based on the energy processing spread in the time-frequency domain and by using the Wigner-Ville transforms). Real data records are considered from two sources, (Case Western Reserve University Bearing Data Center, 2017; VIBROCHANGE, 2017). The results of the algorithms are compared and discussed from accuracy point of view.

The paper is organized in five main sections. Section two introduces the basic of the proposed algorithms, i.e. the CUSUM criterion, the Renyi entropy and Wigner-Ville time-frequency transform. Section three introduces the main features of the data used for experiments. Section four presents the algorithms mainly based on pseudo code. Section five presents the experimental results based on computer simulation and the last section concludes the paper.

2. THE BASIC KNOWLEDGE

This section briefly introduces the basic knowledge used in this work, i.e. CUSUM criterion, the Renyi entropy and Wigner-Ville time-frequency transform.

The principle of the CUSUM method was introduced by (Page, 1954), in the context of products quality, but can be extended to other signals as well to detect changes in the mean of the signal, with independent samples, identically distributed before and after the change time. If a signal vector **x** is considered, the most used and general expression of the CUSUM criterion is

$$CS(i) = \sum_{j=1}^{i} (x(j) - k), \quad i = 1,2,..., N \tag{1}$$

where k is the reference value which corresponds to the difference between the average values from the two states or regimes, normal and abnormal. The reference value, k, is computed by using knowledge from the model of the signal $x(t)$, via plausibility ratios. A change occurs if

$$CS(i) > Th \tag{2}$$

where Th is a trigger parameter, bigger than the nominal value. The evolution of the variable CUSUM is linear with the slope, changing when the mean of the signal changes. This is particularly important because some versions of the CUSUM methods reset the origin of the samples, when evaluation of the criterion (2) is true. The second step or stage in the change detection method is the estimation of the change point. In the case of the batch processing the estimation is

$$M = \arg\max\{CS\} \tag{3}$$

The point M is the last point before the change starts and the point (M+1) is the first after the change. The recursive equation which corresponds to the on-line detection procedure is

$$CS(i) = CS(i-1) + x(i) - \bar{x}, \quad i = 1,2,..., N; CS(0) = 0 \tag{4}$$

More consideration and examples are available e.g. in (Woodall and Montgomery, 2014; Pastell and Madsen, 2008; Siettos and Russo, 2013; Xin, *et al*, 2013; Tam, 2009).

The Renyi entropy is intensively used in the field of statistical signal processing, especially in non-stationaries conditions, being able to estimate the number of the components of complex signals and the degree of randomness in various signal representation framework, in time or frequency domains. In this work, the 2nd order Renyi entropy is used by considering

$$H_2(\mathbf{P}) = -\log\left(\sum_{j=1}^{N} P(j)^2\right) \tag{5}$$

where $P(j), j=1,\ldots,N$ are the probabilities of a finite set of independent events from a discrete information source.

The computation of the entropy needs the probabilities set. If the data set has N samples and the observation vectors of size mx1, the Renyi entropy estimator is based on (Parzen, 1962), window with Gaussian kernel, (Erdogmus, *et al*, 2002):

$$\hat{H}_2(\mathbf{x}_m, \sigma) = -\log \frac{1}{N} \sum_{n=1}^{N}\left(\frac{1}{N}\sum_{k=1}^{N}\left(G(x_m(n) - x_m(k), 2\sigma^2\right)\right) \tag{6}$$

The kernel size, σ should be small (relative to the standard deviation of the data), (Hild, *et al*, 2006) suggests values between 0.1 and 2 for unit-variance signals.

An important category of time–frequency transforms are based on the signal energy distribution in the time–frequency domain. They are characterized by a kernel function. The properties of the representation are reflected by simple constraints on the kernel that produces the time-frequency representation with prescribed, desirable properties, (Cohen, 1989). A mathematical description of these transforms can be given by

$$TFD_x(t,\omega) = \frac{1}{4\pi^2}\iiint x\left(u+\frac{\tau}{2}\right)\cdot x^*\left(u-\frac{\tau}{2}\right)\cdot \phi(\theta,\tau)\cdot e^{-j\theta\cdot t - j\tau\omega + j\theta\cdot\tau}du d\tau d\theta \quad (7)$$

where $\phi(\theta, \tau)$ is a two-dimensional kernel function, determining the specific representation in this category, and hence, the properties of the representation. The star (*) represents complex conjugation operator. The basic distribution in this approach is the Wigner Distribution. If *x(t)* is a continuous (possible complex) signal, the Wigner distribution of the signal *x(t)* is defined (in time domain) as, (McFadden and Wang, 1990),

$$W_x(t,f) = \int_{-\infty}^{\infty} x\left(t+\frac{\tau}{2}\right)\cdot x^*\left(t-\frac{\tau}{2}\right)\cdot e^{-j2\pi f\tau}d\tau \quad (8)$$

For the cases where *x(t)* is an analytic signal, the Wigner distribution is called Wigner-Ville distribution or Winger-Ville transform. This distribution satisfies a large number of desirable mathematical properties, as it is described in the specialized literature, e.g. (Auger, *et a*l, 1996; Hlawatsch, and Boudreaux-Bartels, 1992). It has also some drawbacks, as the apparition of the cross-terms. The coefficients of the time-frequency transform define a time-frequency image.

3. DATA RECORDS

Data were considered for the case of faults in bearings, available from (Case Western Reserve University Bearing Data Center, 2017), and briefly described in Table 1. Three faults are considered, as *F*1 (Inner race), *F*2 (Ball) and *F*3 (Outer race). The case *F*0 means no faults. Four sizes of the faults were considered, by introducing the cases 1 to 4. The file data0 contains the first 5,000 elements of *fd*. All names beginning with "*d*" indicates a file with 5,000 samples from normal conditions (no faults) and 6,000 samples for the cases with faults. In total, each data file has 11,000 samples with sampling rate of 12,000 sample/s. In the case of the fault *F*3, there are four cases, depending on the transducer positions and orientations. The following test cases could be considered:

(9) T0: **D**_0 = *d*0 // *fault free*

(10) T1: **D**_1 = *d*0+*d*1+*d*6+*d*9+*d*14 // *F*1

(11) T2: **D**_2=*d*0+*d*2+*d*7+*d*10+*d*15 // *F*2

(12) T3: **D**_3= *d*0+*d*3+*d*8+*d*11 // *F*3

Table 1. Data Test Set

Case	Fault size	Faults (with files)					
		*F*0	*F*1	*F*2	*F*3		
		Free	Inner Race	Ball	06HH	03HH	12HH
0	0.000 "	*d*0	-	-	-	-	-
1	0.007"	-	*d*1	*d*2	*d*3	*d*4	*d*5
2	0.014"	-	*d*6	*d*7	*d*8	-	-
3	0.021"	-	*d*9	*d*10	*d*11	*d*12	*d*13
4	0.028"	-	*d*14	*d*15		*	

The previous data has the advantage of physic compliance but the disadvantage of the variable length of the data vector. Thus, a more simple data structure were considered by considering a matrix composed of 15 columns of N =11,000 elements as

$$\mathbf{D} = [\mathbf{D}_1 \quad \mathbf{D}_2 \quad \mathbf{D}_3 \quad \mathbf{D}_4] \quad (13)$$

$$\mathbf{D}_1 = [d1 \quad d2 \quad d3 \quad d4 \quad d5] \quad (14)$$

$$\mathbf{D}_2 = [d6 \quad d7 \quad d8] \quad (15)$$

$$\mathbf{D}_3 = [d9 \quad d10 \quad d11 \quad d12 \quad d13] \quad (16)$$

$$\mathbf{D}_4 = [d14 \quad d15] \quad (17)$$

The first 5,000 elements of each column are from the fault free vector, i.e. *d*0.

4. DESCRIPTION OF THE ALGORITHMS

The basic structure of data processing and implementation of the algorithm are presented in Fig. 1. Computer simulation were conducted with a sliding rectangular window of length *n*=1,000 samples over the set of *N* samples. Data was low pass filtered by using a cut-off frequency of 5,000 kHz. An algorithm includes a computation block for calculus of various change detection criteria, including parameter estimations. The values of the criteria are compared with one or more reference values, in order to decide which hypothesis is true. Depending on the selected value an action is triggered, e.g. activating an alarm.

The threshold values are suggested by the end user of the equipment, which known better the features and the working regimes of the process. These values could be for various degrees of the faults, as incipient (small), important (medium) or very important or major (high). Three algorithms are tested: CA – an algorithm based on CUSUM criterion; AA1 – an algorithm which uses the Renyi entropy for the input of the CUSUM criterion; AA2 – an algorithm based on the energy distribution through the time-frequency domain.

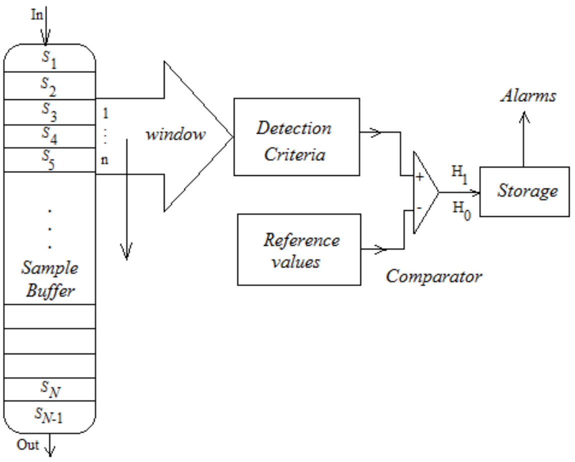

Fig.1. The basic structure of the data processing and algorithm implementation

In the case of the CUSUM, the change detection point is estimated by the coordinate (time moment) of the minimum value of the CUSUM variable. The algorithm AA2 is based on Wigner-Ville transform. For each record of N sample a sliding non-overlapping window of $n<<N$ samples is considered. The algorithm estimate an interval of change, i.e. the length of the sliding window must be smaller as possible. Each window of observation generates a time-frequency frame/image. There are two aspects of decisions: (i) quantitative, based on the change of the energy in frequency domain, from one frame to another one; (ii) qualitative, based on the content of each time-frequency image and also on the evolution of the content when the frames are changing. The change criterion is based on the maximum variation of the energy between frames (in number of *nf*). The normalized energy is computed for each frame by

$$E(j) = \frac{1}{n}\sum_{k=1}^{n} S(k)^2,\ j = 1:nf\ = (N/n) \quad (18)$$

The decision criterion could be complex as in (Aiordachioaie, 2013) or a simple one which selects the maximum difference between two consecutive frames

$$j^* = \arg\max\left\{\left|E(j) - E(j-1)\right|,\ j = 1:nf\right\} \quad (19)$$

5. EXPERIMENTAL RESULTS

In the previous data set, the true change value is 5001. The CUSUM based algorithm (CA) has relatively good results. The estimated values are around 4992, excepting the case #12 when a value of 5130 was obtained. The maximum absolute error is 129 and the mean square error is 1185. In the next two figures, two cases are presented: case #1 in Fig. 2 and case #12 in Fig. 3.

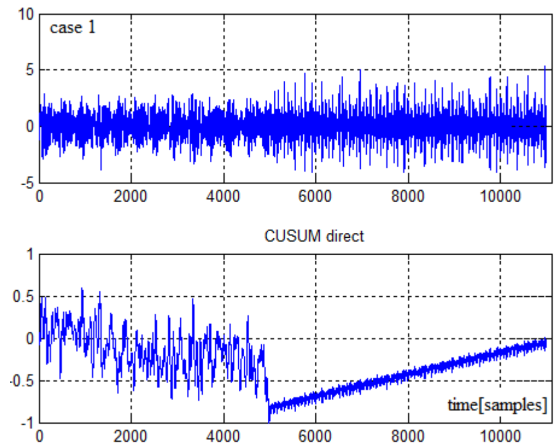

Fig.2. Results of the CA in the case #1.

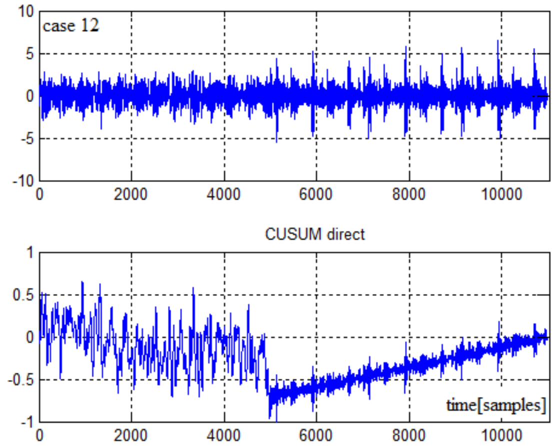

Fig.3. The results of the CA in the case #12

The results of the AA1 algorithm are much better. The maximum absolute error is 43 and the mean square error is 315.6. The results obtained for the set of two test cases, i.e. #1 and #12, are presented in Fig. 4 and 5. Similar results are obtained with the second data source, i.e. (VIBROCHANGE, 2017).

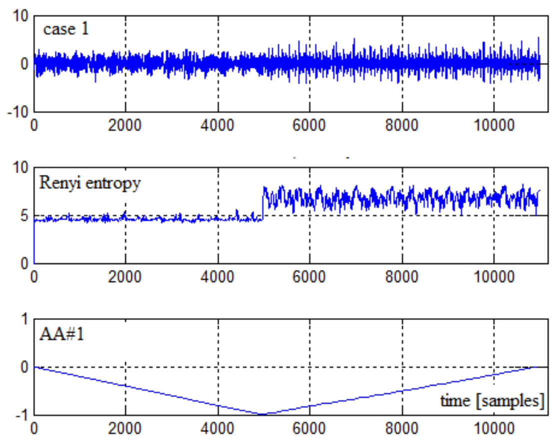

Fig.4. The results of the AA1 for the case #1

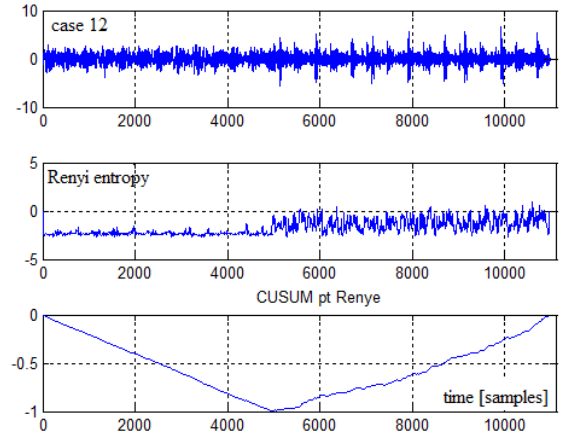

Fig.5. The results of the AA1 for the case #12

The next algorithm is AA2, and it is based on Wigner-Ville transform. For each record of N = 11,000 samples a non-overlapping sliding window of n = 1000 samples is considered. In Fig. 6 and 7 the results of AA2 for the test cases #1 and #12 are presented. Each figure has four sub-figures. On the left side, on the vertical direction, the power spectrum is presented. On the bottom left corner, the evolution of the energy over intervals (frames) is presented. On the bottom side, the evolution of the signal during the frame is indicated. Finally, the Wigner-Ville image is presented on the right side. The information content of these images is not explored here.

There are some parameters which should be clarified. The most important is the length of the observation window. Computer based simulation show an optimum length between 500 and 600 samples. Smaller lengths do not catch the change and longer lengths do not provide enough precision. In fig. 8, the evolution of the energy over 22 frames is presented. At the bottom of the figure, the evolution of the gradient is presented. The biggest value (represented by a circle) indicates a change in the frame no. 11, i.e. in the interval (5000, 5500).

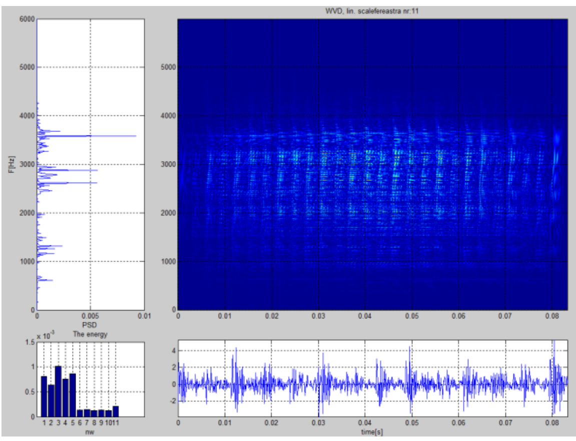

Fig.6. The results of the AA#2 for the case 1

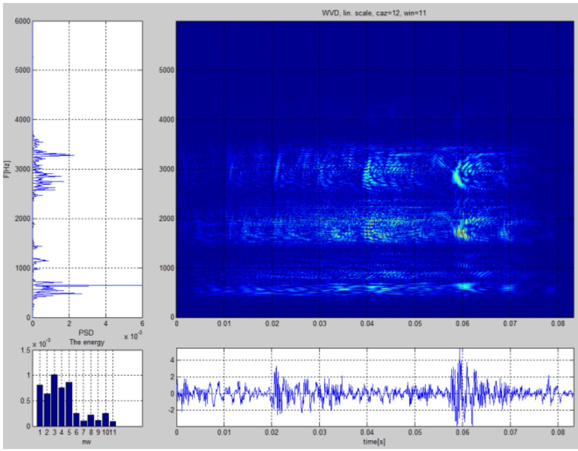

Fig.7. The results of the AA2 for the case #12

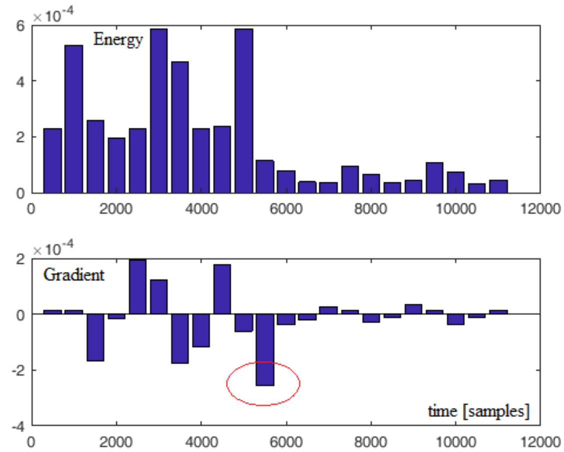

Fig.8. Details of the energy evolution for AA2, case #12

6. CONCLUSIONS

The main objective of the paper was to present some results concerning change detection in vibrational processes (processes which generates vibrations), by using two advanced methods, based on Renyi entropy and time-frequency transforms.

The reference method was based on CUSUM, a well-known detection method used in statistical signal processing. Data from two sources was considered, both from experimental processes, but the results from only one source were presented and discussed.

The simulation results show better performances for the algorithm based on Renyi entropy. The uses of the time-frequency paradigm, here by the Wigner-Ville transform, reveals new sources of information based on the time-frequency images associated to the time-frequency transforms. Attention must be paid to the optimization of some parameters, as the length of the observation window. In the case of the AA1, the window is sliding in small steps and the computation time is high. In the case of AA2, based on the change in the energy, the length must be optimized in the sense to have quick response (small width of the observation window) and robustness (high width of the window).

Taking into account the fact that the time scale is at the level of sampling period and the sampling rate is very high, the length of the window is not a major drawback for mechanical vibration of ordinary electrical machines, i.e. with angular speed less of 5,000 rpm.

ACKNOWLEDGMENT

The work was partly supported by the Romanian Council for Research (UEFISCDI) under PN-II-PT-PCCA-2013-4-0044 and Financial Contract 224/2014, "Experimental model for change detection and diagnosis of vibrational processes using advanced measuring and analysis techniques model-based (VIBROCHANGE)".

7. REFERENCES

Patton, R.J., Frank, P., and Clark, R. Eds. (1989). *Fault Diagnosis in Dynamic Systems – Theory and Application*. Prentice Hall.

Gertler, J., (1998*). Fault Detection and Diagnosis in Engineering Systems*. Marcel Dekker.

Isermann, R., (1997). Supervision, fault-detection and fault-diagnosis methods – An introduction, *Control Engineering Practice*, **5** (5).

Chen, J. and Patton R.J., (1999). *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer Academic Publishers.

Sobhani-Tehrani, E. and Khorasani, K., (2009). Fault Diagnosis of Nonlinear Systems Using a Hybrid Approach, *Lecture Notes in Control and Information Sciences*, 383.

Venkatasubramanian, V., Rengaswamy, R., Yin, K, and Kavuri, S.N. (2003). A review of process fault detection and diagnosis: Part I: Quantitative model-based methods, *In Computers & Chemical Engineering*, **27** (3), pp. 293-311.

Gustafsson, F. (2001). *Adaptive Filtering and Change Detection*, Wiley.

Basseville, M. (1997). Statistical approaches to industrial monitoring problems – Fault detection and isolation, *Proceedings of the 11th IFAC/IFORS Symposium on Identification and System Parameter Estimation* – SYSID'97, Kitakyushu, Japan, July 8-11.

Mangoubi, R.S., (1998). *Robust Estimation and Failure Detection*. Springer-Verlag.

Case Western Reserve University Bearing Data Center, (2017). Available: http:// csegroups.case.edu/bearingdatacenter/home.

SKF, *Technical data*, (2017) Available: http://www.skf.com/group/index.html.

VIBROCHANGE, A research contract. (2017). http://www.etc.ugal.ro/VIBROCHANGE/index.php?Lang=En.

Page, E.S., (1954). Continuous Inspection Scheme,. *Biometrika*, **41**(1/2), pp. 100–115.

Woodall, W.H. and Montgomery, D.C. (2014). Some Current Directions in the Theory and Application of Statistical Process Monitoring, *Journal of Quality Technology*, 46(1), pp.78-94.

Pastell, M. and Madsen, H. (2008). Application of CUSUM charts to detect lameness in a milking robot, Elsevier, *Expert Systems with Applications*, **35**, pp. 2032-2040.

Siettos, C.I. and Russo, L. (2013). Mathematical modeling of infectious disease dynamics, *Virulence*, **4**(4), pp. 295-306.

Xin, L., Yu, PL.H. and Lam, K., (2013). An Application of CUSUM Chart on Financial Trading, *in CIS, 9th Int. Conf. on*, pp.178-181.

Tam, D. (2009). A Theoretical Analysis of Cumulative Sum Slope Statistic for Detecting Signal Onset and Offset Trends from Background Noise Level, *The Open Statistics and Probability Journa*l, **1**, pp. 43-51.

Parzen, E. (1962). On estimation of a probability function and mode, *Annals of Mathematical Statistics*, **33**(3), pp. 1065–1076.

Erdogmus, D. et al., (2002). Blind source separation using Renyi's alpha-marginal entropies, *Neurocomputing*, 49(1), pp. 25–38.

Hild II, K.E., Erdogmus, D. and Principe, J.C. (2006). An analysis of entropy estimators for blind source separation, *Signal Processing*, **86**(1), pp.182-194.

Cohen, L. (1989). Time-Frequency Distributions - A Review, *Proceedings of the IEEE*, 77(7), pp. 941–980.

McFadden, P.D. and Wang, W. (1990). Time-Frequency Domain Analysis of Vibration Signals for Machinery Diagnostics. (I) Introduction to the Wigner-Ville Distribution, University of Oxford, *Report OUEL* 1859/92.

Auger, F., Flandrin, P., Gonçalvès, P. and Lemoine, O., (1996). *Time-frequency Toolbox*, CNRS France - Rice University.

Hlawatsch, F. and Boudreaux-Bartels, F. (1992). Linear and Quadratic Time-Frequency Signal Representations, *IEEE Signal Processing Magazine*, pp. 21–67.

Aiordachioaie, D. (2013). Signal Segmentation Based on Direct Use of Statistical Moments and Renyi Entropy, *10th International Conference on Electronics, Computer and Computation* (ICECCO'13), Istanbul, Turkey, pp. 359-362.

Basseville, M. (2013). Divergence measures for statistical data processing - An annotated bibliography, *Signal Processing*, **93**, pp. 621-633.