24

Time-frequency analysis using Bayesian regularized neural network model

Imran Shafi, Jamil Ahmad, Syed Ismail Shah and Ataul Aziz Ikram
Iqra University Islamabad Campus, Sector H-9
Pakistan

1. Introduction

During the last twenty years there has been spectacular growth in the volume of research on studying and processing the signals with time–dependant spectral content. For such signals we need techniques that can show the variation in the frequency of the signal over time. Although some of the methods may not result in a proper distribution, these techniqes are generally known as time–frequency distributions (TFDs) (1, Boashash 2003). The TFDs are two–dimensional (2–D) functions which provide simultaneously, the temporal and spectral information and thus are used to analyze the non–stationary signals. By distributing the signal energy over the time–frequency (TF) plane, the TFDs provide the analyst with information unavailable from the signal's time or frequency domain representation alone. This includes the number of components present in the signal, the time durations and frequency bands over which these components are defined, the components' relative amplitudes, phase information, and the instantaneous frequency (IF) laws that components follow in the TF plane.
There has been a great surge of activity in the past few years in the TF signal processing domain. The pioneering work in this area is performed by (2, Claasen & Mecklenbrauker 1980), (3, Janse & Kaizer 1983), and (4, Boashash 1978). They provided the initial impetus, demonstrated useful methods for implementation and developed ideas uniquely suited to the situation. Also, they innovatively and efficiently made use of the similarities and differences of signal processing fundamentals with quantum mechanics. Claasen and Mecklenbrauker devised many new ideas, procedures and developed a comprehensive approach for the study of joint TFDs. However Boashash is believed to be the first researcher, who used various TFDs for real world problems. He developed a number of new methods and particularly realized that a distribution may not behave properly in all respects or interpretations, but it could still be used if a particular property such as the IF is well defined. The research presented in (6, Flandrin & Escudie 1980) transcribed directly some of the early quantum mechanical results, particularly the work on the general class of distributions, into signal analysis. The work in (3, Janse & Kaizer 1983) developed innovative theoretical and practical techniques for the use of TFDs and introduced new methodologies remarkable in their scope.
Historically the spectrogram has been the most widely used tool for the analysis of time–varying spectra. The spectrogram is expressed mathematically as the magnitude–square of the short–time Fourier transform (STFT) of the signal, given by

$$S\left(t,\omega\right)=\left|\int_{-\infty}^{\infty}s\left(t\right)h(t-\tau)e^{-j\omega\tau}d\tau\right|^{2} \tag{1}$$

where $s(t)$ is the signal and $h(t)$ is a window function. Nevertheless, the spectrogram has severe drawbacks, both theoretically, since it provides biased estimators of the signal IF and group delay, and practically, since the Gabor–Heisenberg inequality makes a tradeoff between temporal and spectral resolutions unavoidable. However STFT and its variation being simple and easy to manipulate, are still the primary methods for analysis of the signals with time varying spectral contents and most commonly used today.

There are other approaches with a motivation to improve upon the spectrogram, with an objective to clarify the physical and mathematical ideas needed to understand time–varying spectrum. These techniques generally aim at devising a joint function of time and frequency, a distribution that will be highly concentrated along the IFs present in a signal and cross terms (CTs) free thus exhibiting good resolution. One form of TFD can be formulated by the multiplicative comparison of a signal with itself, expanded in different directions about each point in time. Such formulations are known as quadratic TFDs (QTFDs) because the representation is quadratic in the signal. This formulation was first described by Eugene Wigner in quantum mechanics (7, Wigner 1932) and introduced in signal analysis by Ville (8, Ville 1946) to form what is now known as the Wigner–Ville distribution (WVD). The WVD is the prototype of distributions that are qualitatively different from the spectrogram, produces the ideal energy concentration along the IF for linear frequency modulated signals, given by

$$W(t,\omega)\triangleq\frac{1}{2\pi}\int_{-\infty}^{\infty}s^{*}(t-\frac{1}{2}\tau)s(t+\frac{1}{2}\tau)e^{-j\omega\tau}d\tau \tag{2}$$

It is found that the spectrogram results in a blurred version (5, Cohen 1995), which can be reduced to some degree by use of an adaptive window or by combination of spectrograms. On the other hand, the use of WVD in practical applications is limited by the presence of non-negligible CTs, resulting from interactions between signal components. These CTs may lead to an erroneous visual interpretation of the signal's TF structure, and are also a hindrance to pattern recognition, since they may overlap with the searched TF pattern. Moreover If the IF variations are non–linear, then the WVD cannot produce the ideal concentration. Such impediments, pose difficulties in the correct analysis of non–stationary signals, are dealt in various ways and historically many techniques are developed to remove them partially or completely. They were partly addressed by the development of the Choi–Williams distribution (9, Choi & Williams 1989), followed by numerous ideas proposed in literature with an aim to improve the TFDs' concentration and resolution for practical analysis (10, Shafi et al. 2009). Few other important non–stationary representations among the Cohen's class of bilinear TF energy distributions include the Margenau–Hill distribution (11, Margenau & Hill 1961), their smoothed versions (12, Hippenstiel & Oliveira 1990), and others with reduced CTs (13, Jeong & Williams 1992) are members of this class. Nearly at the same time, some authors also proposed other time–varying signal analysis tools based on a concept of scale rather than frequency, such as the scalogram (14, Daubechies 1990) (the squared modulus of the wavelet transform), the affine smoothed pseudo WVD (15, Rioul & Flandrin 1992) or the Bertrand distribution (16, Bertrand 1988). The theoretical properties and the application fields of this large variety of these existing methods are now well determined, and wide–spread. Although many other QTFDs have been proposed in the literature, no single QTFD can be effectively used in all

possible applications. This is because different QTFDs suffer from one or more problems (5, Cohen 1995).
An ideal TFD function roughly requires the four properties namely (i) high clarity i.e high concentration along individual components, (ii) CTs' elimination, (iii) good mathematical properties, and (iv) lower computational complexity. These characteristics are necessary for an easy visual interpretation of their outcomes and a good discrimination between known patterns for non-stationary signal classification tasks. To analyze the signals well, choosing an appropriate TFD function is important. Which TFD function should be used depends on what application it applies on. On the other hand, the short comings make specific TFDs suited only for analyzing non–stationary signals with specific types of properties and TF structures. Half way in this decade, there has been an enormous amount of work towards achieving high concentration and good resolution along the individual components and to enhance the ease of identifying the closely spaced components in the TFDs. The aim has been to correctly interpret the fundamental nature of the non–stationary signals in the TF domain.
We shall present a novel Bayesian regularized artificial neural network (ANN) based method for computing highly informative TFDs. The proposed method provides a way to obtain a non-blurred and high resolution version of the TFDs of signals whose frequency components vary with time. The resulting TFDs do not have the CTs that appear in case of multicomponent signals in some distributions such as WVD, thus providing visual way to determine the IF of non-stationary signals. It is proved that Bayesian inference framework and ANN learning capabilities can be successfully applied in the TF field, where they have not been used before.

2. Bayesian regularized Neural Network based Framework for Computing De-blurred TFDs

This section presents the Bayesian regularized ANN model (BRNNM) based correlation vectored taxonomy algorithm to compute the TFDs that are highly concentrated in the TF plane (23, Shafi et al. 2008). The degree of regularization is automatically controlled in the Bayesian inference framework and produces networks with better generalized performance and lower susceptibility to over–fitting. The grayscale spectrograms and pre–processed WVD of known signals are vectored and clustered as per the elbow criterion to constitute the training data for multiple ANNs. The best trained networks are selected and made part of the localized neural networks (LNNs). Test TFDs of unknown signals are then processed through the algorithm and presented to LNNs. Experimental results demonstrate that appropriately vectored and clustered data and the regularization, with input training under Mackay's evidence framework, once processed through LNNs produce high resolution TFDs.
Bayesian regularization involves modifying the usually used objective function, such as the mean sum of squared network errors (24, MacKay 1992).

$$mse = \frac{1}{N}\sum_{k=1}^{K}(e_k)^2 \tag{3}$$

where mse, e_k, and N represent MSE, network error and network errors' taps for averaging respectively. It is possible to improve generalization if the performance function is modified by adding a term that consists of the mean of the sum of squares of the network weights and biases

$$msereg = \gamma mse + (1-\gamma)msw \tag{4}$$

where $\gamma, msereg$, and msw are the performance ratio, performance function and mean of the sum of squares of network weights and biases, respectively. msw is mathematically described as under:

$$msw = \frac{1}{n}\sum_{i=1}^{n}(w_i)^2 \tag{5}$$

using this performance function causes the network to have smaller weights and biases, and this forces the network response to be smoother and less likely to over fit. Moreover it helps in determining the optimal regularization parameters in an automated fashion.

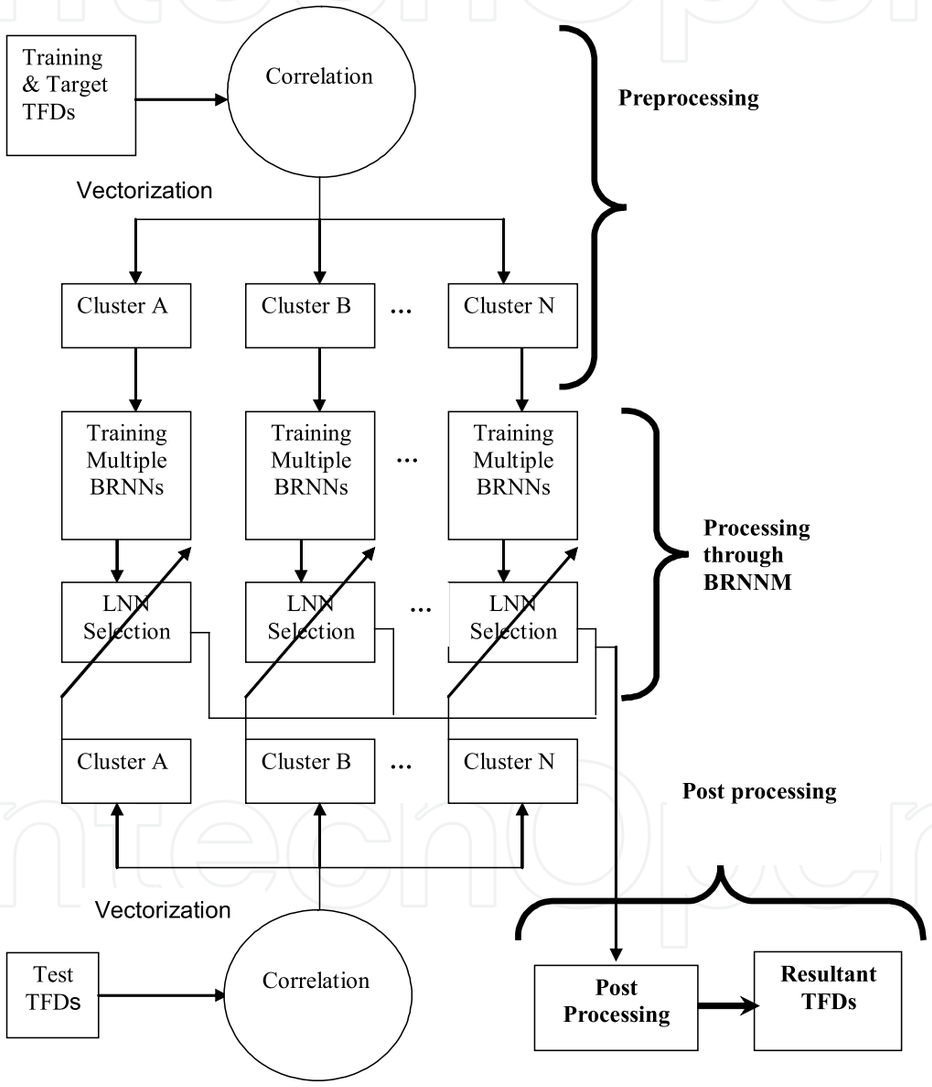


Fig. 1. Flow diagram of the method

Fig. 1 is the overall block representation of the proposed ANN based framework. This block diagram highlights three major modules of the method that include (i) pre–processing of training data, (ii) processing through the BRNNM and (iii) post–processing of output data. These modules and the rationale of the proposed method are described below:

2.1 Pre–processing of Training Data

Fig. 2 depicts the block diagram for this module. It consist of four major steps, namely (i) two–step pre–processing of target TFDs, (ii) vectorization, (iii) subspaces selection and direction vectors, and (iv) correlation and taxonomy. They are described as follows.

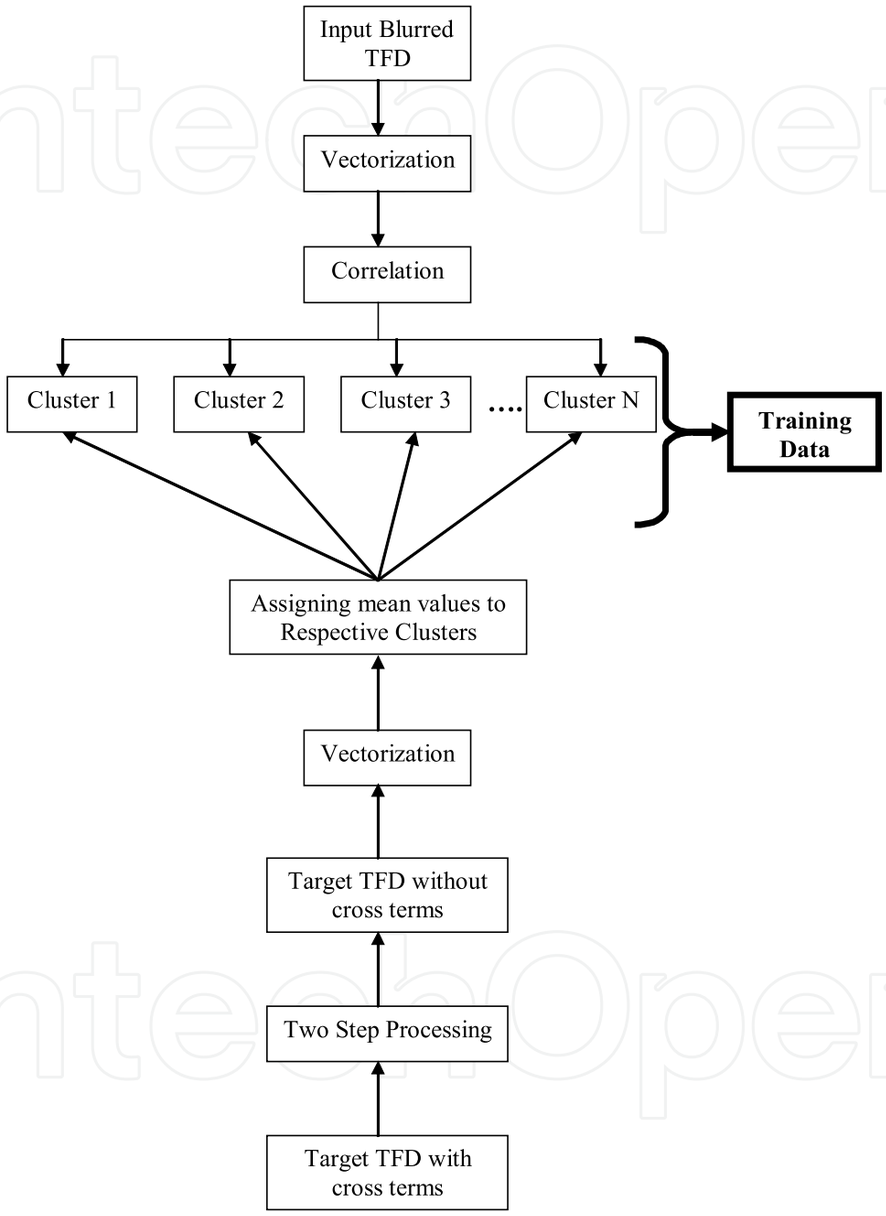


Fig. 2. Pre-processing of training data

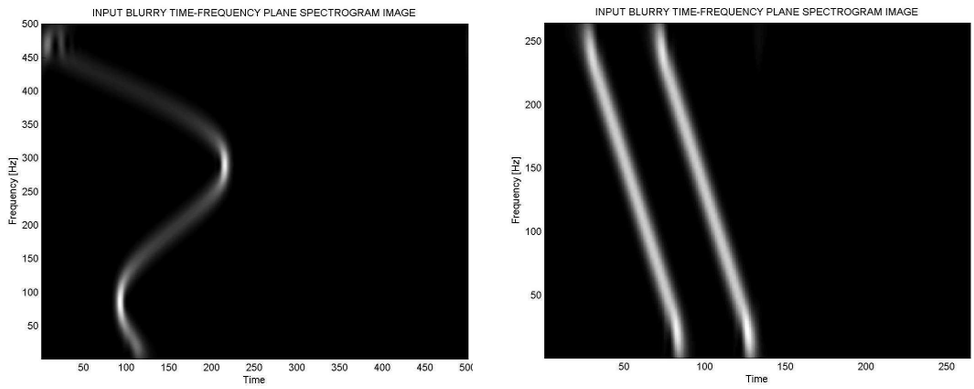
2.1.1 Two–step pre–processing of target TFDs

The highly concentrated WVD of various known signals are used as the target TFDs. As will be shown in Fig. 4, the WVD suffers from CTs which make them unsuitable to be presented

as targets to the ANNs (17, Hagan, Demuth & Beale 1996). The CTs are therefore eliminated before the WVD is fed to the ANN. This is achieved in two steps:

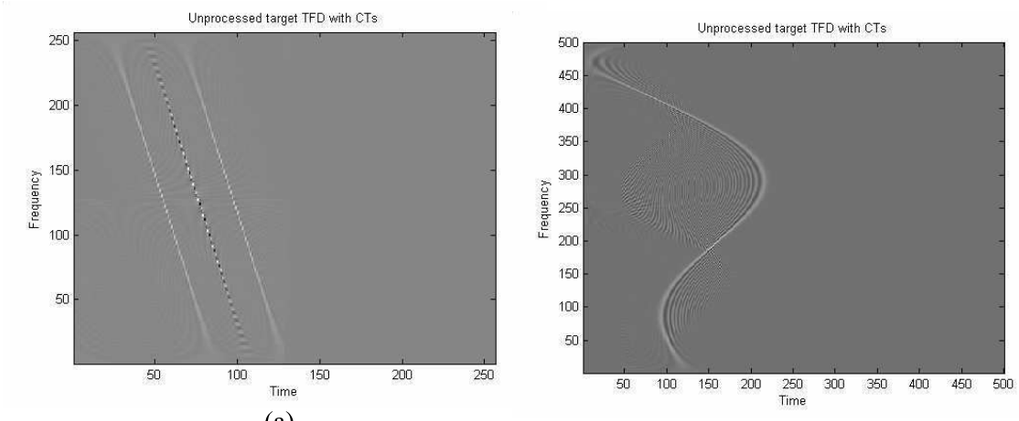
1. The WVD is multiplied point by point with spectrogram of the same signal obtained with a hamming window of reasonable size.
2. All values below a certain threshold are set to zero.

The resultant target TFDs are shown in Fig. 5, which are fed to the ANN after vectorization described as follows.



(a) (b)

Fig. 3. The spectrograms used as input training images of the (a) sinusoidal FM, and (b) parallel chirp signals.

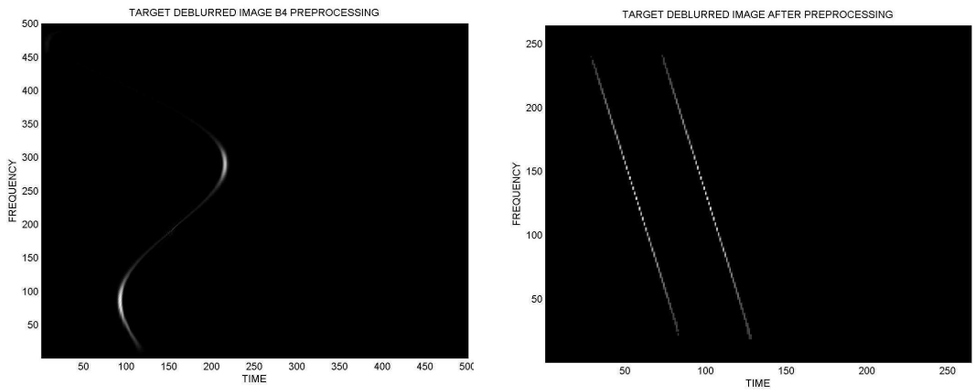


(a) (b)

Fig. 4. Target TFDs with CTs unsuitable for training ANN taking WVD of the, (a) parallel chirps' signal, and (b) sinusoidal FM signal.

2.1.2 Vectorization

(1) Input TFDs. Fig. 3 depicts input spectrograms. A TFD is considered as 2–D image which can be mathematically described as a matrix of pixels depicting grayscale values e.g.,



(a) (b)

Fig. 5. Target TFDs without CTs suitable for training ANN after pre-processing WVD of the, (a) parallel chirps' signal, and (b) sinusoidal FM signal.

$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$.These pixel values can be used to generate vectors, for example, a vector of length three will contain three pixel values of a row/column of TFD image. The suitable vector length is decided after experimenting with various vector lengths (3, 5, 7 and 9). The decision is made based on visual results. Each input TFD image is thus converted to vectors. These vectors are paired with the vectors obtained from target TFDs, to be subsequently used for training.

(2) Target TFDs. After CTs' removal from the target WVD, they are vectored. Next the mean values of these vectors are computed. For example, if $\langle a_{11}, a_{12}, a_{13} \rangle$ is a pixel vector of the input TFD and $\langle b_{11}, b_{12}, b_{13} \rangle$ is the vector representing the same region of the target TFD, then $\frac{(b_{11}+b_{12}+b_{13})}{3}$ will become the target numerical value for the input vector. Mean values are taken as targets with a view that the IF can be computed by averaging frequencies at each time instant, a definition suggested by many researchers (5, Cohen 1995).

2.1.3 Subspaces selection and direction vectors

1. *Elbow Criterion*. The elbow criterion is a common rule of thumb to determine what number of clusters should be chosen. It states that number of clusters be chosen so that adding another cluster does not add sufficient information (18). More precisely, if the percentage of variance explained by the clusters is plotted against the number of clusters, the first clusters will add much information (explain a lot of variance), but at some point the marginal gain will drop, giving an angle in the graph (the elbow). On the following graph (Fig. 6) which is drawn for the problem in hand, the elbow is indicated by the "goose egg". The number of clusters chosen is therefore three.
2. The number of subspaces N_s into which vectors are distributed is decided based on elbow criterion in relation to underlying image features like edges present in the data. The edge is considered because it is one of the important image underlying features and characteristics. Moreover it is well established fact that blurring mostly causes loss of edge information (19, Gonzalez & Wintz 1987). An edge could be ascending $(1, 2, 3)$,

descending $(3,2,1)$, wedge $(1,3,2)$, flat $(1,1,1)$, triangular $(1,3,1)$ etc. Empirically it is found that going from three to four clusters does not add sufficient information, as the end result has no significant change in entropy values as indicated in Table 1 and evident from Fig. 6. The impact of clustering is noted for six different test images (TIs), described in section 3. As a result of this study, $N_s = 3$ is chosen considering the first three most general types of edges.

3. The sub space direction vectors v_n $(n = 1,2 \ldots N_s)$ are selected that will best represent the subspaces. As these subspaces are defined on the basis of edges, so three directional vectors v_h, v_c, v_l are computed in the following manner:
   (a) v_h is obtained by rearranging (any) 3 integers in descending order.
   (b) $v_c$is obtained by rearranging (any) 3 integers in a wedge shape where the highest value occurs in the middle and values on either side are in descending order.
   (c) v_l is obtained by rearranging (any) 3 integers in ascending order.
4. All the direction vectors v_h, v_c, v_l are normalized.

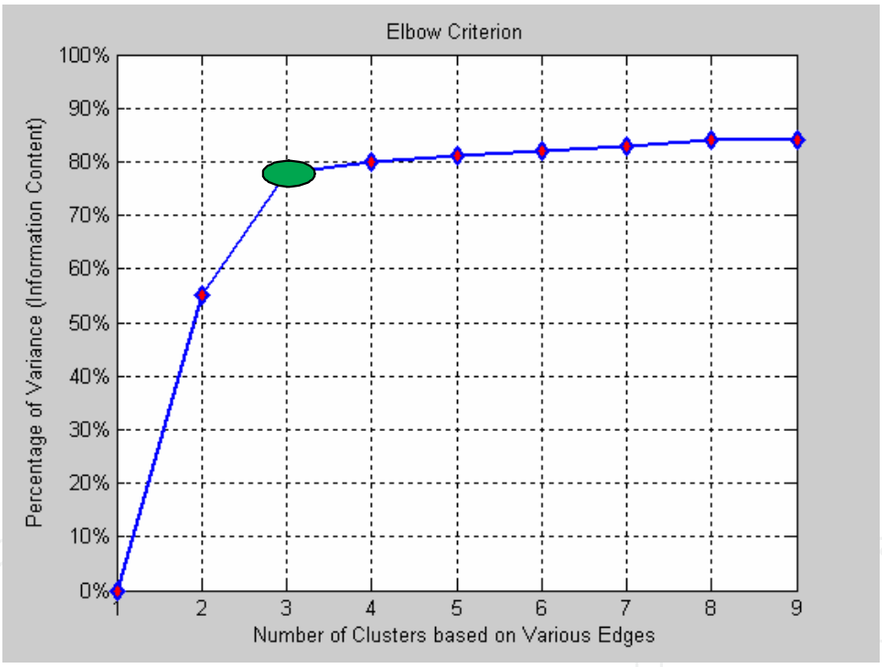


Fig. 6. Elbow criterion

2.1.4 Correlation & Taxonomy

1. An input vector x_i is chosen from input spectrogram. The correlation between each input vector x_i from input TFD and each direction vector v_h, v_c, v_l is calculated, i.e. $t_{ij} = x_i^T v_j$ where $j = h, c, l$.
2. There will be N_s product values obtained as a result of last step for each input vector x_i. To find the best match, if t_{ic} has the largest value then this indicates that the input x_i is most similar to the directional vector v_c, which implies that the vector is wedge type.

Description	E_Q (bits) for test TFDs					
	TI 1	*TI 2*	*TI 3*	*TI 4*	*TI 5*	*TI 6*
No cluster	20.539	18.113	18.323	19.975	21.548	17.910
2 clusters	13.523	12.294	12.421	11.131	14.049	11.940
3 clusters	**8.623**	**6.629**	**7.228**	**5.672**	**8.175**	**6.948**
4 clusters	8.101	6.300	7.202	5.193	8.025	6.733
5 clusters	7.998	6.187	7.111	5.012	7.939	6.678
6 clusters	7.877	6.015	7.019	5.995	7.883	6.661

Table 1. Entropy values vs clusters

3 Step (2) is repeated for all input vectors. Consequently all the vectors are classified based on the type of edge they represent and N_s clusters are obtained. The statistical data revealing various vector types in the two TFD images depicted in Fig. 3 is shown in Table 2.

4 The pairs of vectors from training and target TFDs are formed. These pairs are divided into training set and validation set for training phase and by observing error on these two sets, the aspect of overfitting is avoided.

These steps of vectorization, correlation and taxonomy are further elaborated in graphical form by Fig. 7.

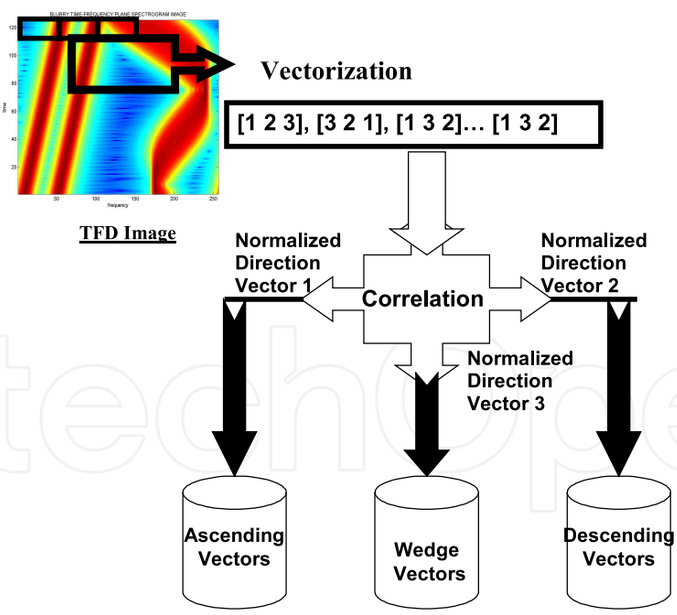


Fig. 7. Vectorization, correlation and taxonomy of TFD image.

2.2 Processing through Bayesian Regularized Neural Network Model

Fig. 8 represents this module. There are three steps in this module, namely (i) training of BRNNM, (ii) selecting the LNNs, and (iii) testing the LNNs. They are discussed as under.

Various parameters (input training TFDs)	*Cluster* 1 *ascending edge type vectors*	*Cluster* 2 *descending edge type vectors*	*Cluster* 3 *wedge edge type vectors*
Sinusoidal FM signal	19157	18531	112
Parallel chirps' signal	4817	4959	52
The best ANN	$ANN-3$	$ANN-2$	$ANN-1$
Time consumed for training	308 *seconds*	114 *seconds*	55 *seconds*
MSE converged	2.54×10^{-4}	3.56×10^{-4}	1.38×10^{-2}

Table 2. Cluster parameters

2.2.1 ANN Training

1. Since the ANN is being used in a data–rich environment to provide high resolution TFDs, it is important that it does well on data it has not seen before, i.e. that it can generalize. To make sure that the network does not become over trained. the error is monitored on a subset of the data that does not actually take part in the training. This subset is called the validation set other than the training set. If the error of the validation sets increases the training stops. For this purpose, alternate pairs of vectors from input and target TFDs are included in training and validation sets.
2. The input vectors represented by x_i and the mean values y_i of the pixel values, of the corresponding window from the target TFD image are used to train the multiple ANNs under Bayesian framework. There are three ANNs trained for each cluster, being the smallest numerical value to check the advantage of training multiple ANNs. This selection has no relation with the number of subspaces or direction vectors.
3. Step (2) is repeated until all pairs of input and corresponding target vectors are used for training.

2.2.1.1 Topology, Architecture and Training TFDs

To address the stated problem, Bayesian Regularized LMB training algorithm is used with feed forward back propagation ANN architecture and 40 neurons in the single hidden layer. This architecture is chosen after an empirical study (20; 21, Shafi et at. 2006, Ahmad et al. 2006). We experiment with various training algorithms using different parameters such as different activation functions between layers, number of hidden layers and number of neurons. Also the positive impact of localised processing by selecting the best trained ANN out of many is ascertained (22, Shah et al. 2007). The 'tansig' and 'poslin' transfer functions are used respectively representing the hidden layer of sigmoid neurons followed by an output layer of positive linear neurons. Multiple layers of neurons with nonlinear transfer functions allow the network to learn linear and nonlinear relationships between input and output vectors. The linear output layer lets the network produce values outside the range $[-1.+1]$.
The spectrograms and WVD of the two signals are used as input and target TFDs respectively for training the ANNs. The first signal is a sinusoidal FM signal, given by:

$$x(n)=e^{-j\pi[\frac{5}{2}+\{0.1\sin(\frac{2\pi n}{N})\}]n} \tag{6}$$

where N refers to the number of sampling points. The spectrogram of this signal is depicted in Fig. 3(a). The respective target TFD, obtained through WVD, is depicted in Fig. 5(a).

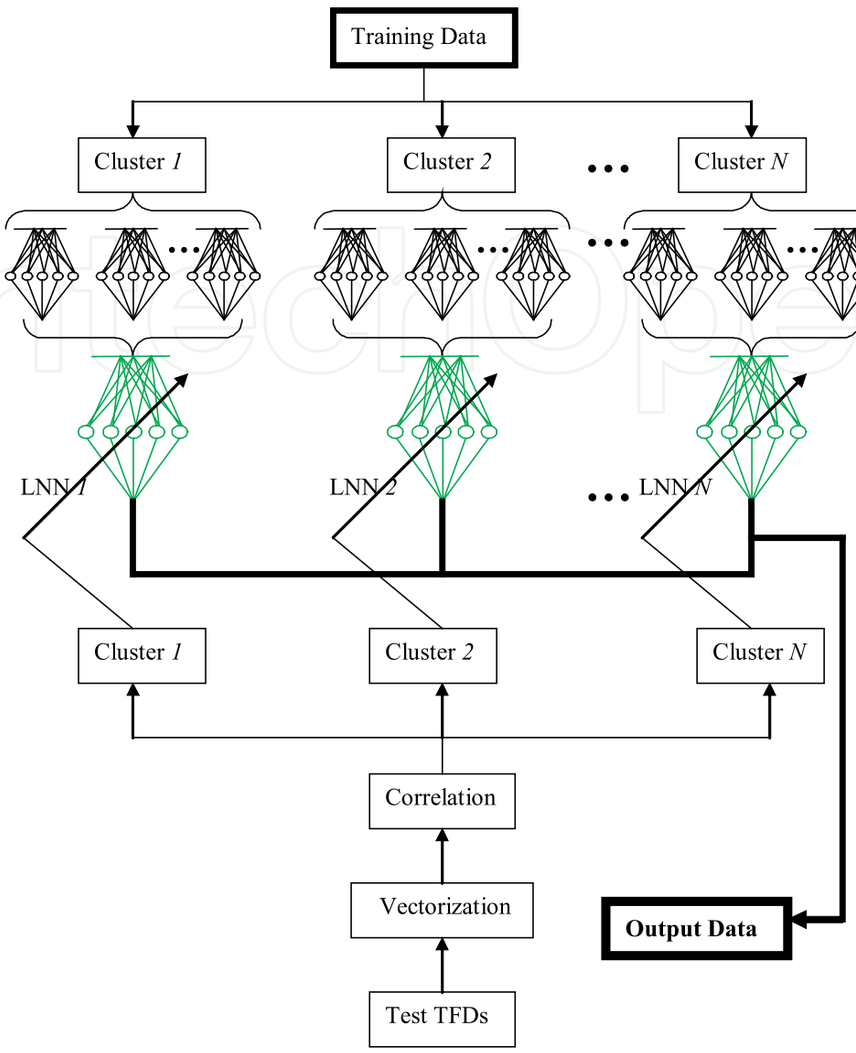


Fig. 8. Bayesian regularised neural network model

The second signal is with two parallel chirps given by:

$$y(n) = e^{j(\frac{\pi n}{4n})n} + e^{j(\frac{\pi}{3}+\frac{\pi n}{4n})n} \tag{7}$$

The spectrogram of this signal is depicted in Fig. 3(b). The respective target TFD, obtained through the WVD, is depicted in Fig. 5(a).

2.2.2 LNNs' selection

1. As mentioned above, three ANNs are being trained for each cluster and the best for each cluster is required to be selected. The "training record" is a programmed structure in which the training algorithm saves the performance of the training–set, test–set, and the validation–set, as well as the epoch number and the learning rate. By keeping track of the network error or performance, accessible via the training record, the best network is selected for each cluster. These best networks selected for various clusters are called the LNNs.

2. Using multiple networks for each cluster is found to be advantageous because the weights are initialized to random values, and when the network begins to over–fit the data, the error in the validation set typically begins to rise. If this happens for a specified number of iterations, the training is stopped, and the weights and biases at the minimum of the validation error are obtained. As a result, various networks will have different MSEs in the last training epoch. The ANN with minimum MSE is the winner and is included in the LNNs. There are three ANN trained for each of three clusters, and as recorded in Table 2 it is found that $ANN-3$ and $ANN-2$ are the best for the first and second clusters respectively, and the $ANN-1$ is found to be the best for the third cluster only. It is assumed that these selected ANN are optimally trained and will posses better generalization abilities.

2.2.3 ANN Testing

1. Test TFDs are converted to vectors (z_i) and clustered after correlating with the direction vectors, as done for the training TFDs.
2. Each test vector z_i is fed to the LNN trained for the particular type and the results are recorded.

2.3 Post–processing of the Output Data

This module is illustrated in Fig. 9. After testing phase, the resultant data is post–processed to get the resultant TFD. As we obtain one value for each vector of length three from test TFD after processing through the LNNs. There are two possibilities to fill the rest of two pixels, either (i) replicate the same value for other two places, or (ii) use zero padding around this single value to complete the number of pixels. Zero padding is optimal because it is found to reduce the blur in TF plane. Next the resultant vectors of correct length are placed at their original places from where they were correlated and clustered. These vectors are placed according to the initially stored grid positions.

3. Discussion on Experimental Results

The discussion on experimental results by the proposed approach and performance evaluation of various bilinear distributions is presented in this section. It uses objective methods of assessment to evaluate the performance of de–blurred TFDs estimated through BRNNM (henceforth the NTFDs). The objective methods allow quantifying the quality of TFDs instead of relying solely on visual inspection of their plots. Performance comparison with various other quadratic TFDs is provided too. This section is organized in two subsections. The subsection 3.1 discusses the NTFDs' performance basing on the visual results and carrying out their information quantification by measuring the entropy values only. In subsection 3.2, the concept and importance of TFDs' objective assessment is described using both real life and synthetic signals.

3.1 Visual Interpretation and Entropy Analysis

In the first phase, five synthetic signals are tested to evaluate the effectiveness of the proposed algorithm basing on visual results and their entropy analysis. They include (i) a two sets of parallel chirps signal intersecting at four places, (ii) a mono–component linear chirp signal, (iii) combined quadratic swept–frequency signals whose spectrograms are concave and convex parabolic chirps respectively, (iv) a combined crossing chirps and sinusoidal FM signal and (v) a quadratic chirp signal. The spectrograms of these signals are shown in Figs.

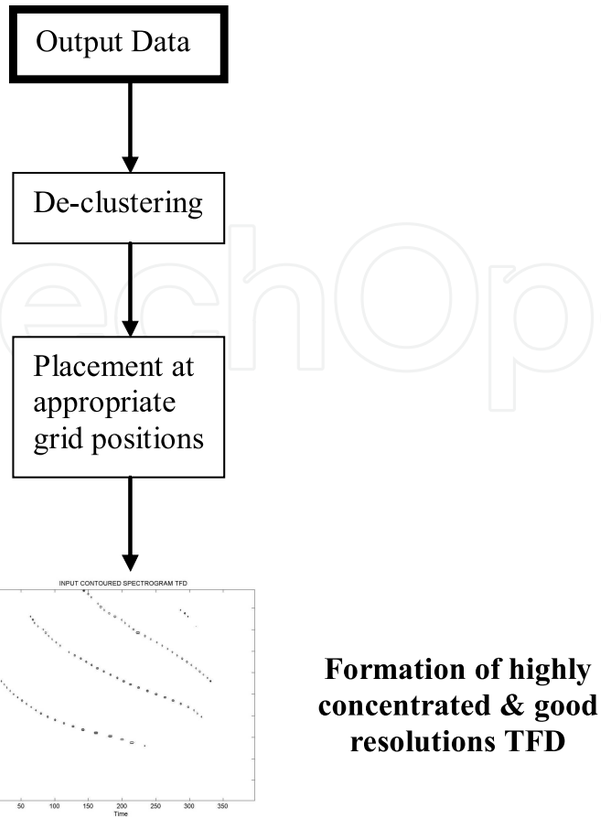


Fig. 9. Post-processing of the output data

10(a) to 10(e) respectively. Keeping in mind that estimation of the IF is rather difficult at the intersections of chirps, the first and fifth test cases are considered to check the performance of proposed algorithm at the intersection of the IFs of individual components present in the signals.

The spectrogram of the two sets of parallel chirps signals crossing each other at four points is fed as the first test signal, depicted in Fig. 10(a), is obtained by:

$$TS_1(n) = e^{j\left[\pi - \frac{\pi n}{6N}\right]n} + e^{j\left[\frac{\pi}{3} - \frac{\pi n}{6N}\right]n} + e^{j\left[\frac{\pi n}{N}\right]n} + e^{j\left[\pi + \frac{\pi n}{N}\right]n} \tag{8}$$

The second test signal is a mono–component chirp signal given by:

$$TS_2(n) = e^{j\left[\pi + \frac{\pi n}{N}\right]n} \tag{9}$$

The spectrogram of the resultant signal is depicted in Fig. 10(b).

The third test signal is obtained by point–by–point addition of two quadratic swept–frequency signals whose spectrograms are concave and convex parabolic chirps respectively. Mathematically both the signals can be obtained by manipulating different parameters of following equation:

$$TS_3(n) = \cos\left[2\pi\left(\frac{\partial}{1+\beta}\right)\left(n^{(1+\beta)}\right) + f_0 + \frac{\theta}{360}\right], \tag{10}$$

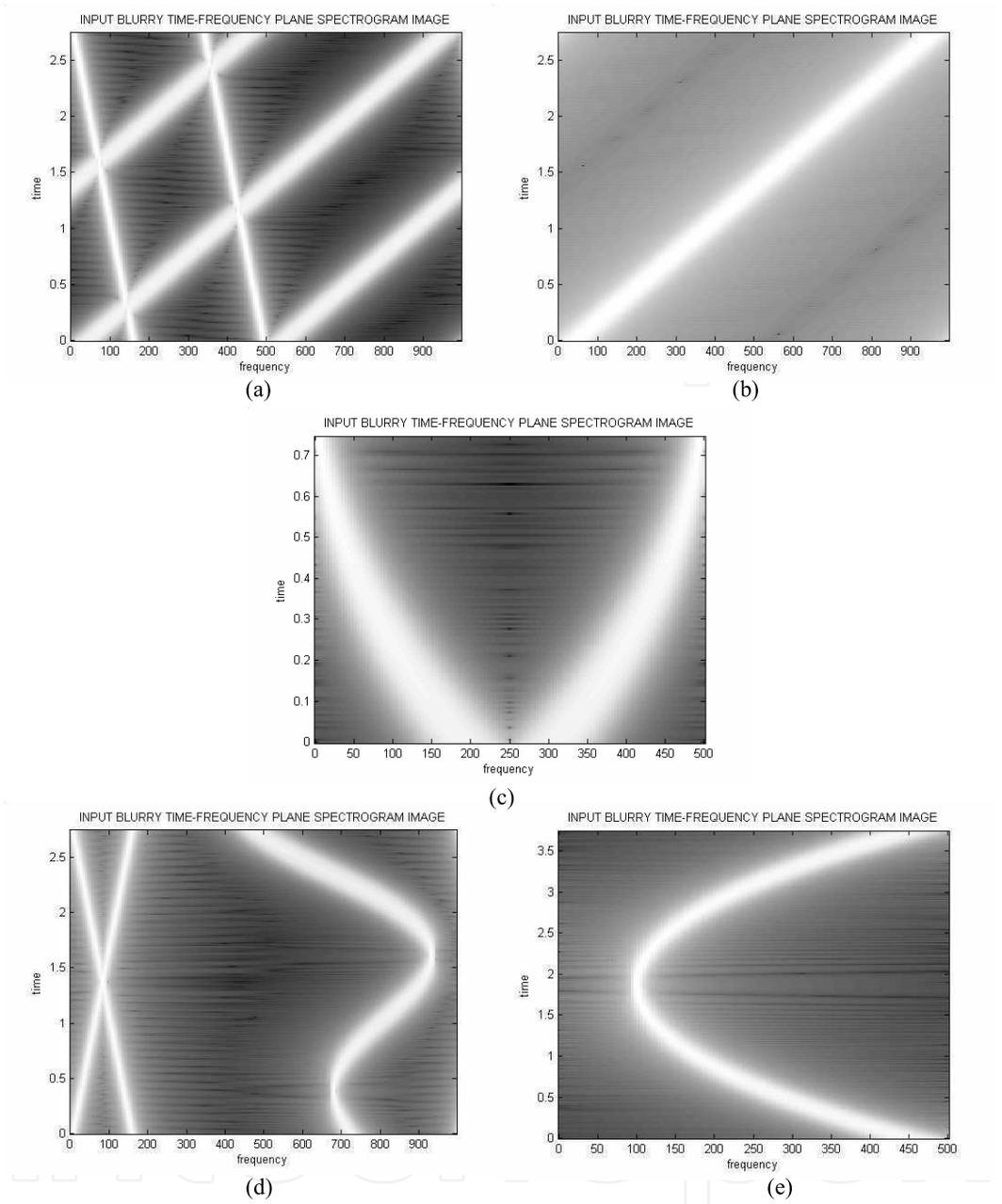


Fig. 10. Test TFDs (a) Crossing chirps (TI 1), (b) mono-component linear chirp (TI 2), (c) combined quadratic swept-frequency signals whose spectrograms are concave and convex parabolic chirps respectively (TI 3), (d) combined sinusoidal FM and crossing chirps (TI 4), and (e) quadratic chirp (TI 5)

where

$$\partial = (f_1 - f_0)\,\rho^{(-\beta)}$$

The method	Resultant E_Q (bits) for test TFDs				
	TI 1	*TI 2*	*TI 3*	*TI 4*	*TI 5*
NTFD	**8.623**	**6.629**	**5.672**	**8.175**	**6.948**
WVD	21.562	10.334	18.511	20.637	18.134
Spectrogram	28.231	18.987	27.743	28.785	23.774

Table 3. Entropy values for various techniques

here $\beta, f_{0,} f_1, \theta$ and ρ are defined as the matching string constant, start frequency, frequency after one second, initial phase of signal and sample rate respectively. The spectrogram of the first quadratic swept–frequency signal is concave parabolic chirp which starts at 250 Hz and go down to 0 Hz at a 1 kHz sample rate; whereas spectrogram of the second quadratic swept–frequency signal is a convex parabolic chirp starting at 250 Hz and going up to 500 Hz at a 1 kHz sample rate. These aspects are evident in the combined spectrogram depicted in Fig. 10(c).
Another test signal is obtained by combining crossing chirps and sinusoidal FM signal as:

$$TS_4(n) = e^{j\left[\frac{\pi n}{N}\right]n} + e^{j\left[\pi+\frac{\pi n}{N}\right]n} + e^{j\pi\left[\frac{1}{2}-(0.1\sin\left(\frac{2\pi n}{N}\right))n\right]} \tag{11}$$

The spectrogram of the signal is depicted in Fig. 10(d).
Yet another test signal is a quadratic chirp which starts at 100 Hz and cross es 200 Hz at 1 *second* with a 1 kHz sample rate. It is obtained from Eqn. (10) after necessary adjustment of different parameters. The spectrogram of this signal is depicted in Fig. 10(e).

3.1.1 Resultant NTFDs – Experimental Results

The five synthetic test signals are: a combined parallel chirps signal crossing at four points, a mono–component linear chirp signal, combined quadratic swept–frequency signals whose spectrograms are concave and convex parabolic chirps respectively, combined crossing chirps and sinusoidal FM signals without any intersection and a quadratic chirp signal. The spectrograms of these signals constitute test image 1 (TI 1), test image 2 (TI 2), test image 3 (TI 3), test image 4 (TI 4), and test image 5 (TI 5). They are depicted in Figs. 10(a–e) respectively. The entropy expression given by $E_Q = -\sum_{n=0}^{N-1} Q\left(n,\omega\right)\log_2 Q\left(n,\omega\right)d\omega \geq 0$ is used to quantify the TFDs' information, which has an inverse relation with the information (25, Gray 1990).
The entropy values for different TFDs have been recorded in Table 3, which are the lowest for the NTFDs than other technique like WVD and the spectrogram. TI 1 and TI 5 are taken into account to check the performance of the proposed algorithm for estimation of the IFs at the intersections along the individual components in the signals. Even though estimation of IF is considered rather difficult at intersections, the algorithm performs well as depicted in Figs. 11(a) and (d). The test images including TI 2, TI 3 and TI 5 present the ideal cases to check the performance of the proposed algorithm trained with signals of different nature. The resultant TFD images are highly concentrated along the IF of individual components present in the signal as shown in Figs. 11(b), (c) and (e).

3.2 Objective Assessment

In this subsection, the objective measures are used to analyze the NTFDs' performance in comparison to other TFDs. The aim has been to find, based on these measures, the highly

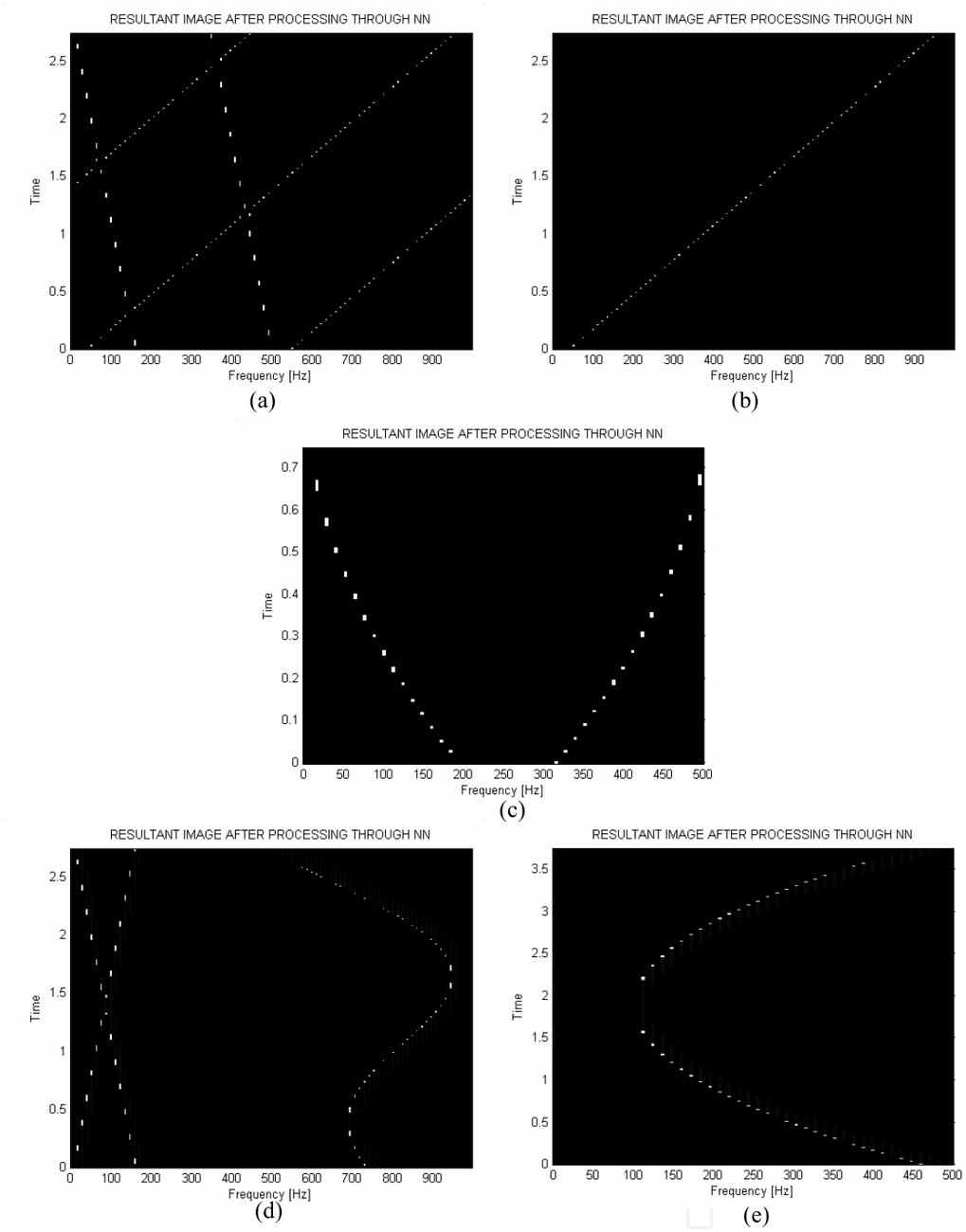
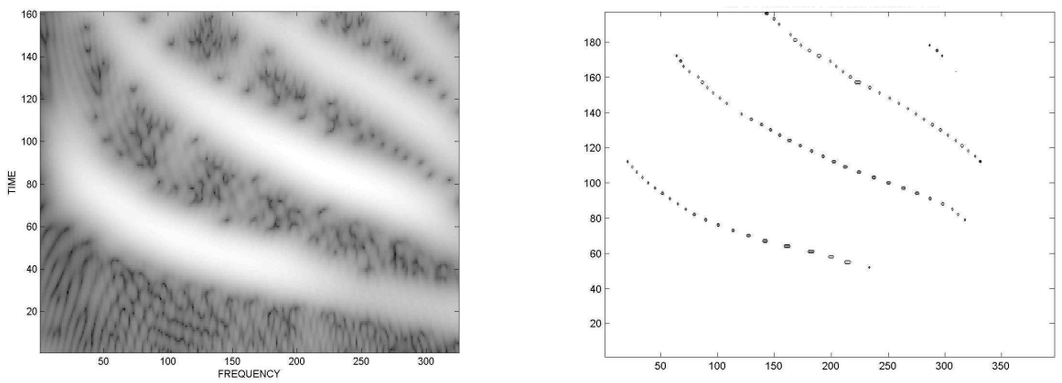


Fig. 11. Resultant TFDs after processing through correlation vectored taxonomy algorithm with LNNs for (a) Crossing chirps (TI 1), (b) mono-component linear chirp (TI 2), (c) combined quadratic swept-frequency signals whose spectrograms are concave and convex parabolic chirps respectively (TI 3), (d) combined sinusoidal FM and crossing chirps (TI 4), and (e) quadratic chirp (TI 5)

informative TFDs having the best concentration and the highest resolution. Five new examples, including both real life and synthetic multicomponent signals, are being considered. The

signals include (i) a multicomponent bat echolocation chirp signal, (ii) a two–component intersecting sinusoidal FM signal, (iii) a two sets of nonparallel, nonintersecting chirps' signal, and (iv) a closely spaced three–component signal containing a sinusoidal FM component intersecting the crossing chirps. The respective spectrograms, termed as test image A (TI A), test image B (TI B), test image C (TI C), and test image D (TI D), are shown in Figs. 12(a), 14(a)–16(a) respectively. As an illustration of the evaluation of the NTFDs' performance through Boashash concentration and resolution measures in (26, Boashash & Sucic 2003), we have further considered a closely spaced multicomponent signal containing two significantly close parallel chirps. The spectrogram of this signal, termed as test image E (TI E), is depicted in Fig. 17(a). The resultant NTFDs for the test signals are shown in Fig. 12(b) & Figs. 14(b)– 17(b) respectively. The visual results are indicative of NTFDs' high resolution and concentration along the IF of the individual component present in the signals.



(a) (b)

Fig. 12. Test TFDs for bat chirps signal, (a) the spectrogram TFD, and (b) the resultant TFD after processing through proposed framework.

Fig. 13. Resultant TFD obtained by the method of (28, Baraniuk & Jones 1993).

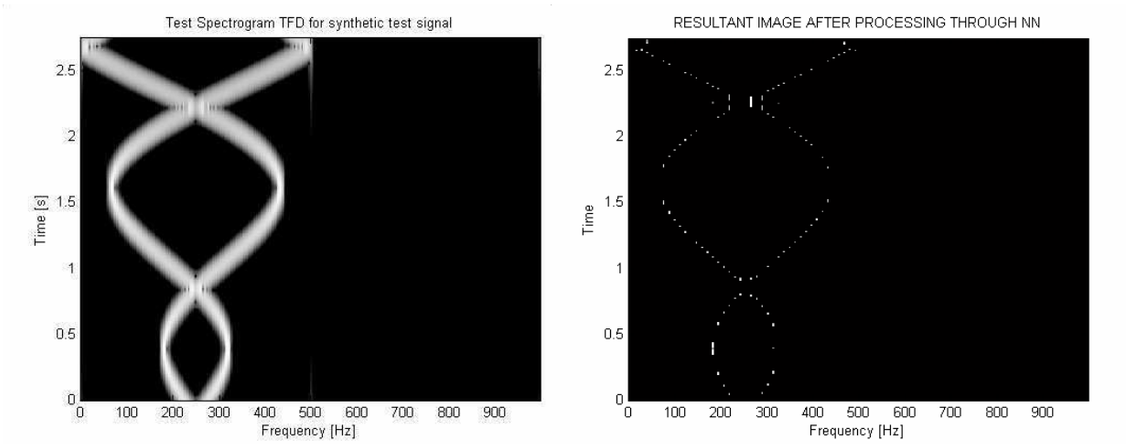


Fig. 14. (a) The test spectrogram (TI 2) $[Hamm, L = 90]$, and (b) The NTFD of a synthetic signal consisting of two sinusoidal FM components intersecting each other.

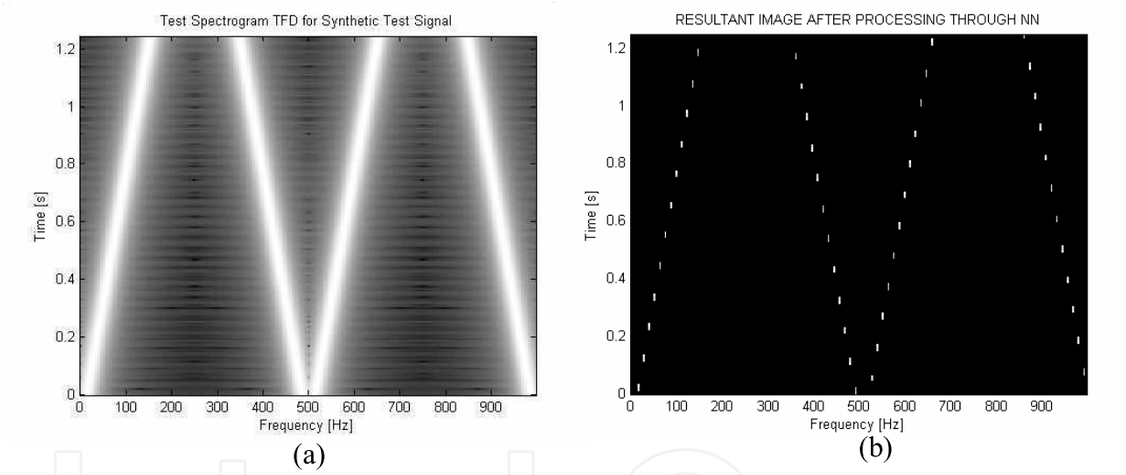
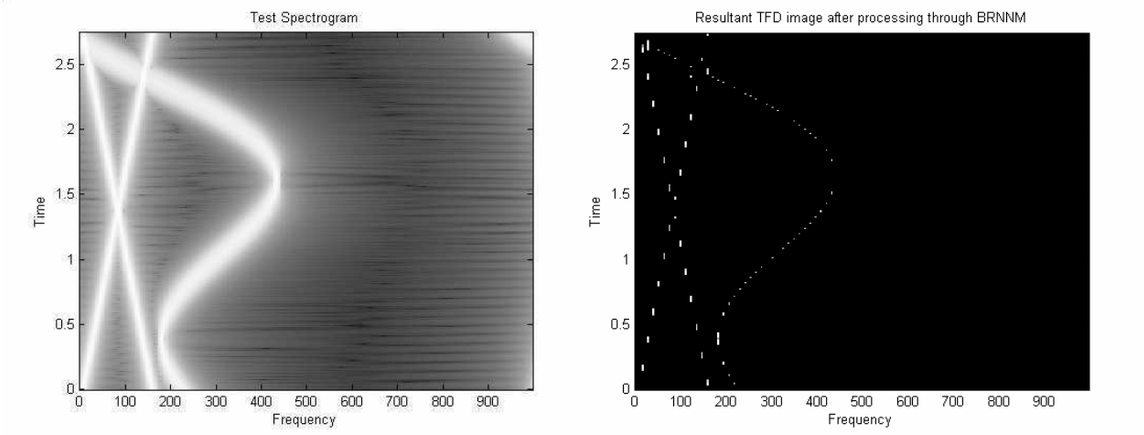


Fig. 15. (a) The test spectrogram (TI 3) $[Hamm, L = 90]$, and (b) The NTFD of a synthetic signal consisting of two-sets of non-parallel, non-intersecting chirps.

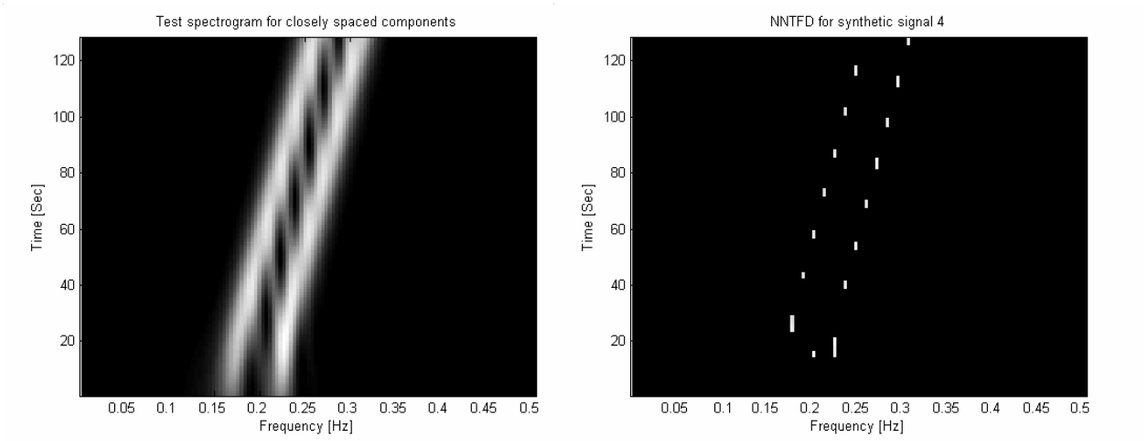
3.2.1 Real Life Test Case

Real life data for bat echolocation chirp sound (adopted from (27)) provides an excellent multicomponent test case. The nonstationary nature of the signal is only obvious from its TFD. The spectrogram of this signal is shown in Fig. 12(a), and the resultant NTFD is depicted in Fig. 12(b). The result for the same test case TFD is computed using an existing optimal kernel method (OKM) (28, Baraniuk & Jones 1993) and is plotted in Fig. 13. The OKM proposes a signal–dependent kernel that changes shape for each signal to offer improved TF representation for a large class of signals based on quantitative optimization criteria. On close monitoring the OKM's output depicted in Fig. 13, it is revealed that this TFD does not fully recover all the components, thus losing some useful information about the signal. Whereas the NTFD is



(a) (b)

Fig. 16. (a) The test spectrogram (TI 4) $[Hamm, L = 90]$, and (b) The NTFD of a synthetic signal consisting of crossing chirps and a sinusoidal FM component.



(a) (b)

Fig. 17. (a) The test spectrogram (TI 5), and (b) the NTFD of test case E.

not only highly concentrated along the IF of the individual components present in the signal but also more informative showing all the components.
For further analysis, slices of the test and resultant NTFDs are taken at the time instants $n = 150$ and $n = 310$ (recall that $n = 1, 2, \ldots, 400$) and the normalized amplitudes of these slices are plotted in Fig. 18. These instants are chosen because three chirps are visible (see Fig. 12(b)) at these time instants. Fig. 18 confirm the peaky appearance of three different frequencies at these time instants. It is worth mentioning that the NTFD not only recovers the fourth component (the weakest) but it has the best resolution i.e. narrower main lobe and no side lobes.

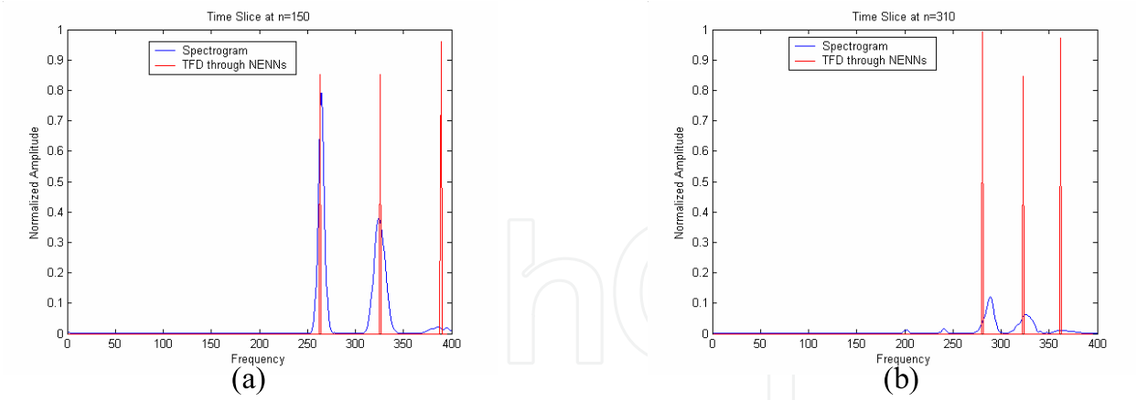


Fig. 18. The time slices for the spectrogram (blue) and the NTFD (red) for the bat echolocation chirps' signal, at n=150 (left) and n=310 (right)

3.2.2 Synthetic Test Cases

Further four specially synthesized signals of different nature are fed to the model to check its performance at the intersection of the IFs and closely spaced components, keeping in mind that estimation of the IF is rather difficult in these situations. The test cases are described as under:

3.2.2.1 Test case 1.

The first one is the synthetic signal consisting of two intersecting sinusoidal FM components, given as:

$$SynTS_1(n) = e^{-i\pi\left(\frac{5}{2}-0.1\sin(2\pi n/N)\right)n} + e^{i\pi\left(\frac{5}{2}-0.1\sin(2\pi n/N)\right)n} \tag{12}$$

The spectrogram of the signal is shown in Fig. 14(a).

3.2.2.2 Test case 2.

The second synthetic signal contains two sets of nonparallel, nonintersecting chirps once plotted on the TF plane. Mathematically it can be written as:

$$SynTS_2(n) = e^{i\pi\left(\frac{n}{6N}\right)n} + e^{i\pi\left(1+\frac{n}{6N}\right)n} + e^{-i\pi\left(\frac{n}{6N}\right)n} + e^{-i\pi\left(1+\frac{n}{6N}\right)n} \tag{13}$$

The spectrogram of the signal is shown in Fig. 15(a).

3.2.2.3 Test case 3.

It is a three–component signal containing a sinusoidal FM component intersecting two crossing chirps. It is expressed as:

$$SynTS_3(n) = e^{i\pi\left(\frac{5}{2}-0.1\sin(2\pi n/N)\right)n} + e^{i\pi\left(\frac{n}{6N}\right)n} + e^{i\pi\left(\frac{1}{3}-\frac{n}{6N}\right)n} \tag{14}$$

The spectrogram of the signal is shown in Fig. 16(a). The two components (sinusoidal FM and chirp components) are very close in between $150 - 200\ Hz$ near $0.5\ sec$. This is to confirm the model's effectiveness in de–blurring closely spaced components.

3.2.2.4 Test case 4.

This particular test case is adopted from Boashash (26, Boashash & Sucic 2003) to compare the TFDs' concentration and resolution performance at the middle of the signal duration interval by Boashash performance measures. The signal consists of two linear frequency modulated signals whose frequencies increase from 0.15 to 0.25 Hz and from 0.2 to 0.3 Hz, respectively, over the time interval $t\forall[1,128]$. The sampling frequency is $f_s = 1\ Hz$.The authors in (26, Boashash & Sucic 2003) have found the modified B distribution $(\beta = 0.01)$ as the best performing TFD for this particular signal at the middle after measuring the signal components' parameters needed in Boashash resolution measure (see Table 5). The signal is defined as;

$$SynTS_4(n) = \cos\left(2\pi\left(0.15t + 0.0004t^2\right)\right) + \cos\left(2\pi\left(0.2t + 0.0004t^2\right)\right) \tag{15}$$

The spectrogram of the signal is shown in Fig. 17(a).
The above mentioned test cases are processed through the BRNNM and the resultant NTFDs are shown in Figs. 14(b)–17(b). High resolution and concentration along the IF of individual components is obvious once inspecting these plots visually.

3.2.3 Performance Evaluation

To evaluate the performance, numerical computations by the methods like the *ratio of norms based measures, Shannon & Rényi entropy measures, normalized Rényi entropy measure* and *Stankovic measure* are recorded in Table 4. The entropy measures including Shannon & Rényi entropies with or without normalization make excellent measures of the information extraction performance of TFDs. By the probabilistic analogy, minimizing the complexity or information in a particular TFD is equivalent to maximizing its concentration, peakiness, and, therefore, resolution (29, Jones & Parks 1992). To obtain the optimum distribution for a given signal, the value of ratio of norms based and Boashash resolution measures should be the maximum (30, Jones & Parks 1990), whereas TFDs' yielding the smallest values for Stankovic and Boashash concentration measures are considered as the best performing TFD in terms of concentration and resolution (26; 31, Boashash & Sucic 2003, Stankovic 2001)
The values in Table 4 refer to the NTFDs as the best TFDs by various criteria. This can be better observed by plotting these measures separately for various TI's (i.e. TI A–TI D) shown in Fig. 19. Few singularities are mainly attributable to inherent shortcomings and derivations' assumptions, e.g. simple Rényi entropies, being unable to detect zero mean CTs, indicate ZAMD as the best concentrated TFD. However the more often used volume normalized Rényi entropies are the minimum for the NTFDs[1].

[1] Here the abbreviations for different methods include the spectrogram (spec), Wigner–Ville distribution (WVD), Choi–Williams distribution (CWD), Zhao–Atlas–Marks distribution (ZAMD), neural network based TFD (NTFD), Margenau–Hill distribution (MHD), Born–Jordan distribution (BJD), Simple Neural network based method (SNN) without clustering the data and the optimal radially Gaussian kernel TFD method (OKM).

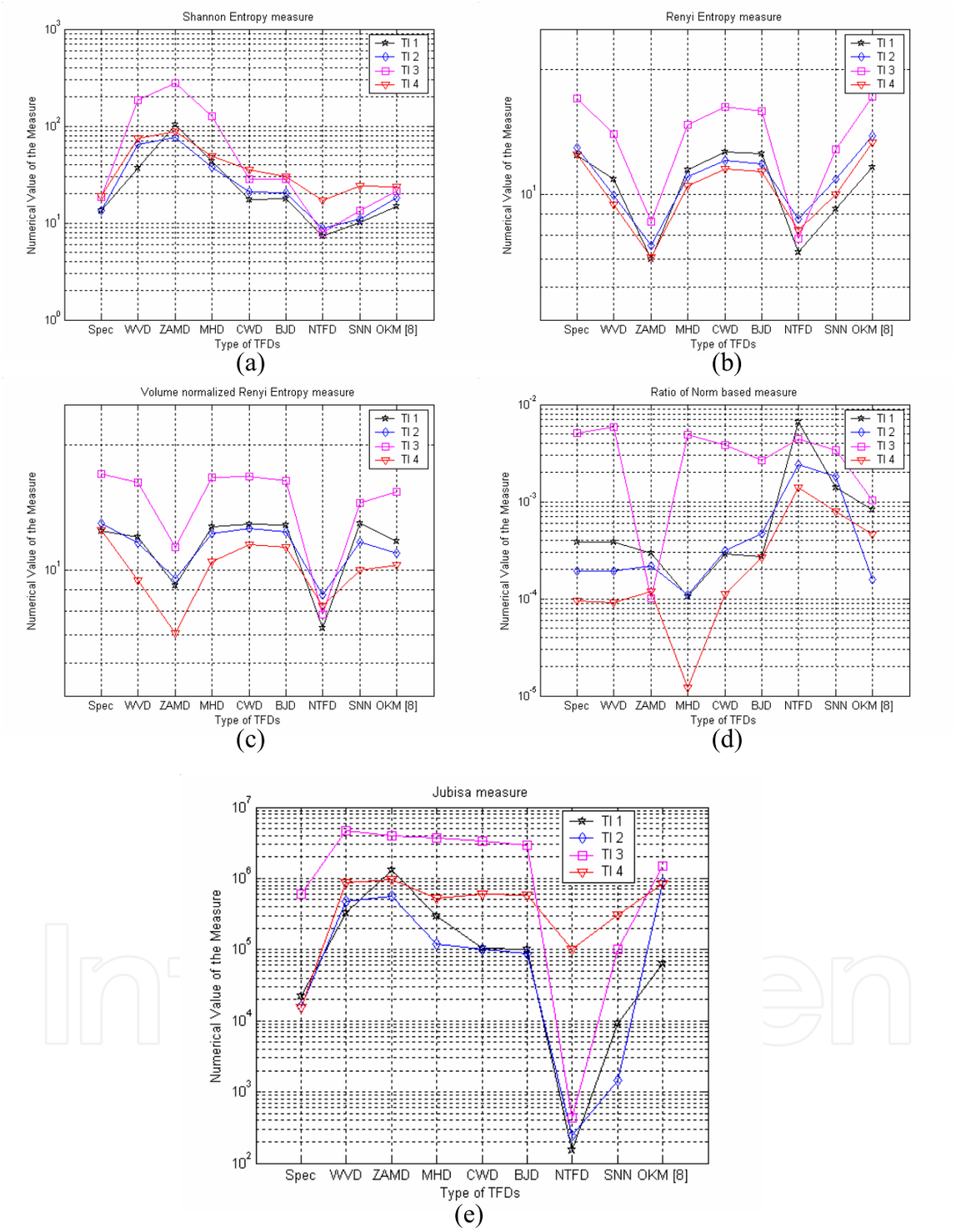


Fig. 19. Comparison plots, criteria vs TFDs, for the test images A–D, (a) The Shannon entropy measure, (b) Rényi entropy measure, (c) Volume normalized Rényi entropy measure,(d) Ratio of norm based measure, and (e) Stankovic measure.

Description	*Test TFD*	*Spec*	*WVD*	*ZAMD*	*MHD*	*CWD*	*BJD*	*NTFD*	*SNN*	*OKM*
Shannon entropy measure	TI A	13.46	36.81	102.23	42.98	17.27	17.73	7.27	10.18	14.68
	TI B	13.45	64.33	76.81	37.74	20.82	20.43	8.75	10.88	18.08
	TI C	18.66	185.49	274.73	126.02	28.08	28.05	7.87	13.45	21.42
	TI D	18.94	74.82	87.30	49.24	35.31	29.92	17.25	24.23	23.57
Ratio of Norm based measure ($\times 10^{-4}$)	TI A	3.81	3.84	2.94	1.05	2.89	2.73	66	13.88	8.32
	TI B	1.94	1.91	2.18	1.10	3.10	4.67	24	18.12	1.59
	TI C	51.23	58.0	1.02	48.71	38.53	26.37	44	33.90	10.26
	TI D	0.95	0.92	1.19	0.12	1.11	2.68	14	8	4.60
Rényi entropy measure	TI A	12.45	10.90	7	11.47	12.67	12.54	7.26	9.25	11.65
	TI B	12.98	9.95	7.56	11.03	12.06	11.85	8.74	10.89	13.82
	TI C	17.07	14.01	8.62	14.74	16.24	15.84	7.85	12.82	17.22
	TI D	12.47	9.48	7.06	10.50	11.54	11.34	8.23	10.03	13.31
Energy Normalized Rényi entropy measure	TI A	12.45	10.90	7	11.47	12.67	12.54	7.26	9.25	11.65
	TI B	12.98	9.95	7.56	11.03	12.06	11.85	8.74	10.89	13.82
	TI C	17.07	14.01	8.62	14.74	16.24	15.84	7.85	12.82	17.22
	TI D	12.47	9.48	7.06	10.50	11.54	11.34	8.23	10.03	13.31
Volume Normalized Rényi entropy measure	TI A	12.45	12.02	9.18	12.75	12.93	12.85	7.26	12.97	11.77
	TI B	12.98	11.62	9.54	12.26	12.60	12.38	8.74	11.68	10.98
	TI C	17.07	16.28	11.35	16.70	16.77	16.41	7.85	14.49	15.43
	TI D	12.47	9.48	7.06	10.50	11.54	11.34	8.23	10.03	10.31
Stankovic measure ($\times 10^5$)	TI A	0.2219	3.30	13.14	2.9200	1.06	1.01	0.0015	0.0912	0.6300
	TI B	0.1600	4.68	5.6266	1.1861	1.0123	0.8946	0.0024	0.0145	8.6564
	TI C	6.03	47.05	39.64	36.47	33.08	29.39	0.0043	0.9973	14.73
	TI D	0.1553	8.67	9.6253	5.1848	6.0110	5.8933	1.0030	3.0223	8.5551

Table 4. Performance Measures Comparison for Various TFDs

Boashash performance measures for concentration and resolution are computationaly expensive because they require calculations at various time instants. To limit the scope, these measures are computed at the middle of the synthetic signal defined in Eqn. (15) and the results are compared with the one reported in (26, Boashash & Sucic 2003). A slice is taken at $t = 64$ and the signal components' parameters $A_{M_1}(64)$, $A_{M_2}(64)$, $A_M(64)$, $A_{S_1}(64)$, $A_{S_2}(64)$, $A_S(64)$, $V_{i_1}(64)$, $V_{i_2}(64)$, $V_i(64)$, $f_{i_1}(64)$, $f_{i_2}(64)$and $\Delta f_i(64)$, as well as the CTs' magnitude $A_X(64)$ are measured. These are then used to calculate the TFDs' normalized instantaneous resolution and modified concentration performance measures $\mathbb{R}_i(t)$ and $\mathbb{C}_n(t)$. The measurement results are recorded in Table 5 and Table 6 seperately for $\mathbb{R}_i(64)$ and $\mathbb{C}_n(64)$. The slice of the signal's NTFD at $t = 64$ is shown in Fig. 20(f).

A TFD that, at a given time instant, has the largest positive value (close to 1) of the measure $\mathbb{R}_i$ is the TFD with the best resolution performance at that time instant for the signal under consideration. From Table 5, the NTFD of synthetic signal given by Eqn. (15) gives the largest value of $\mathbb{R}_i$ at time $t = 64$ and hence is selected as the best performing TFD of this signal at $t = 64$. On similar lines, the TFDs' concentration performance is compared at the middle of signal duration interval. A TFD is considered to have the best energy concentration for a given multicomponent signal if for each signal component, it yields the smallest

TFD (optimal parameter)	$A_M(64)$	$A_S(64)$	$A_X(64)$	$V_i(64)$	$\triangle f_i(64)$	D(64)	$\mathbb{R}(64)$
Spectrogram $(Hann, L = 35)$	0.9119	0.0087	0.5527	0.0266	0.0501	0.4691	0.7188
WVD	0.9153	0.3365	1	0.0130	0.0574	0.7735	0.6199
ZAMD $(a = 2)$	0.9146	0.4847	0.4796	0.0214	0.0420	0.4905	0.5661
CWD $(\sigma = 2)$	0.9355	0.0178	0.4415	0.0238	0.0493	0.5172	0.7541
BJD	0.9320	0.1222	0.3798	0.0219	0.0488	0.5512	0.7388
Modified B $(\beta = 0.01)$	0.9676	0.0099	0.0983	0.0185	0.0526	0.5957	0.8449
NTFD	0.9013	0	0	0.0110	0.0550	0.800	0.9333

Table 5. Parameters and the Normalized Instantaneous Resolution Performance Measure of TFDs for the Time Instant t=64

TFD (optimal parameters)	$A_{S_1}(64)$	$A_{S_2}(64)$	$A_{M_1}(64)$	$A_{M_2}(64)$	$V_{i_1}(64)$	$V_{i_2}(64)$	$f_{i_1}(64)$	$f_{i_2}(64)$	$\mathbb{C}_1(64)$	$\mathbb{C}_2(64)$
Spectrogram $(Hann, L = 35)$	0.0087	0.0087	1	0.8238	0.03200	0.0200	0.1990	0.2500	0.1695	0.0905
WVD	0.3365	0.3365	0.9153	0.9153	0.0130	0.013	0.1980	0.2554	0.4333	0.4185
ZAMD$(a = 2)$	0.4848	0.4900	1	0.8292	0.0224	0.0204	0.2075	0.2495	0.5927	0.6727
CWD$(\sigma = 2)$	0.0176	0.0179	1	0.8710	0.0300	0.0176	0.205	0.2543	0.1639	0.0898
BJD	0.1240	0.1204	1	0.8640	0.0270	0.0168	0.2042	0.2530	0.2562	0.2058
Modified B $(\beta = 0.01)$	0.0100	0.0098	1	0.9352	0.0190	0.0180	0.200	0.2526	0.1050	0.0817
NTFD	0	0	0.8846	0.9180	0.0110	0.0110	0.2035	0.2585	0.0541	0.0425

Table 6. Parameters and the Modified Instantaneous Concentration Performance Measure of TFDs for the Time Instant t=64

1. Instantaneous bandwidth relative to component IF $(V_i(t)/f_i(t))$ and,
2. Sidelobe magnitude relative to mainlobe magnitude $(A_S(t)/A_M(t))$.

The measured results are recorded in Table 6, which indicate that the NTFD of signal given by Eqn. (15) yield the smallest values of $\mathbb{C}_{1,2}(t)$ at $t = 64$ and hence is selected as the best concentrated TFD at $t = 64$. To draw a better comparison, the values of $\mathbb{R}_i$ and $\mathbb{C}_{1,2}$ computed for different TFDs are plotted in Fig. 21.

4. Conclusions

The attempt to clearly understand what a time–varying spectrum is, and to represent the properties of a signal simultaneously in time and frequency without any ambiguity, is one of the most fundamental and challenging aspects of signal analysis. A large pubished scientific literature highlights the significance of TF processing with regard to improved concentration and resolution. However as this task is achieved by many different types of TF techniques, it is important to search for the one that is most pertinent to the application. Although the WVD and the spectrogram QTFDs are often the easiest to use, they do not always provide an accurate characterization of the real data. The spectrogram results in a blurred version and the use of the WVD in practical applications has been limited by the presence of CTs and inability to produce ideal concentration for non–linear IF variations. The spectrogram, for example,

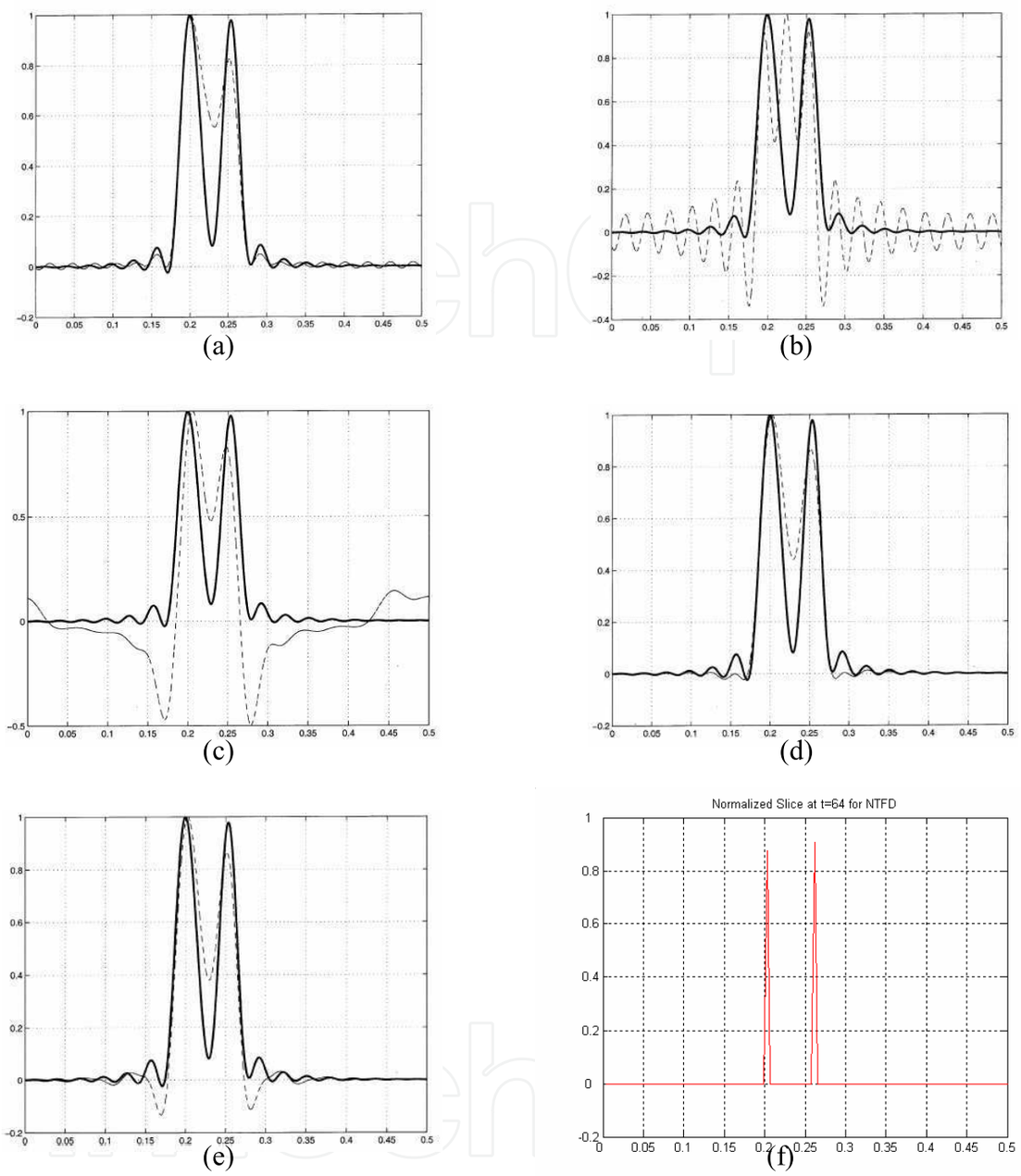


Fig. 20. The normalized slices at $t = 64$ of TFDs, (a) The spectrogram, (b) WVD, (c) ZAMD, (d) CWD, (e) BJD, (f) NTFD. First five TFDs (dashed) are compared against the modified B distribution (solid), adopted from Boashash (26, Boashash & Sucic 2003).

could be used to obtain an overall characterization of the non–stationary signals' structure, and then the information could be used to invest in another QTFD that is well matched to the data for further processing that requires information that is not provided by the spectrogram, the idea conceived and implemented in (32, Shafi et al. 2007).

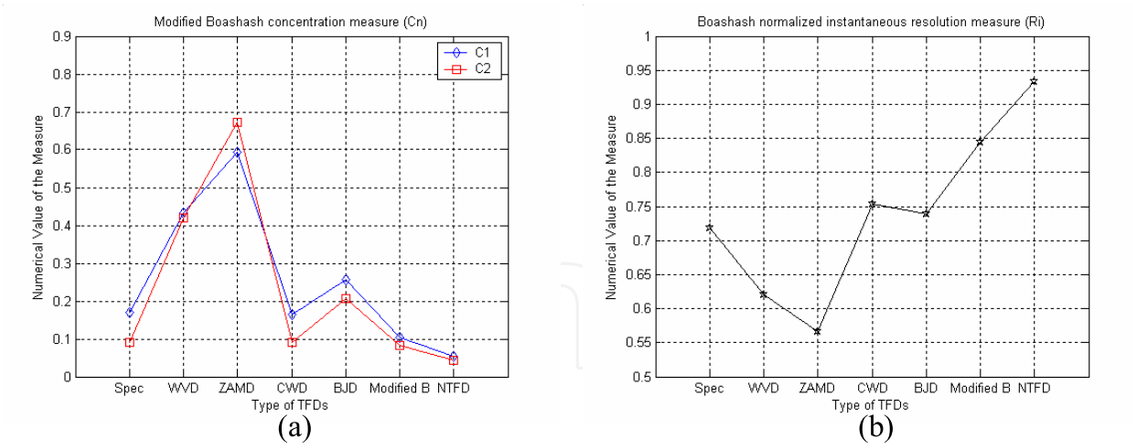


Fig. 21. Comparasion plots for Boashash TFDs' performance measures vs TFDs, (a) The modified concentration measure, and (b) Boashash normalized instantaneous resolution measure

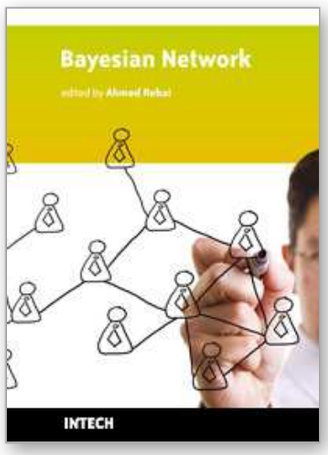
A novel ANN based approach incorporating Bayesian regularization is implemented and evaluated of computing informative, non–blurred and high resolution TFDs. The resulting TFDs do not have the CTs that appear in case of multicomponent signals in some distributions such as WVDs, thus providing visual way to determine the IF of non–stationary signals. The technique explores that the mixture of localized neural networks focused on a specific task deliver a TFD that is highly concentrated along the IF with no CTs as compared to training the ANN which does not receive the selected input. Experimental results presented in section 3 demonstrate the effectiveness of the approach.
For the completeness of proposed framework, the NTFD's performance is further assessed by the information theoretic criteria. These quantitative measures of goodness are used instead of relying solely on the visual measure of goodness of TFDs' plots. The mathematical framework to quantify the TFDs' information is found effective in ascertaining the superiority of the results obtained by the ANN based multiprocesses technique, using both synthetic and real life examples. The NTFD is compared to some popular distributions known for their CTs' suppression and high energy concentration in the TF domain. It is shown that the NTFD exhibits high resolution, no interference terms between the signal components and is highly concentrated. Also it is found to be better at detecting the number of components in a given signal compared to the conventional distributions.

5. References

[1] Boashash, B. (2003). *Time–Frequency Signal Analysis and Processing*, B. Boashash, Ed. Englewood Cliffs, NJ: Prentice–Hall.
[2] Claasen, T. A. C. M. & Mecklenbrauker, W. F. G. (1980). The Wigner distribution–a tool for time–frequency signal analysis; part I: continuous–time signals; part II: discrete time signals; part III: relations with other time–frequency signal transformations. *Philips Journal of Research*, Vol. 35, pp. 217–250, 276–300 and 372–389.
[3] Janse,C. P. and Kaizer,J. M. (1983). Time–frequency distributions of loudspeakers: the application of the Wigner distribution. *Journal of Audio Engg. Soc.*, Vol. 31, pp.198–223.

[4] Boashash, B. (1978). Representation temps–frequence. *Soc. Nat. ELF Aquitaine*, Pau, France, Publ. Recherches, no. 373–378.
[5] Cohen, L. (1995). *Time Frequency Analysis*, Prentice–Hall, NJ.
[6] Flandrin, P. and Escudie, B. (1980). Time and frequency representation of finite energy signals: a physical property as a result of a Hilbertian condition. *Signal Processing*, Vol. 2, pp. 93–100.
[7] Wigner, E. P. (1932). On the quantum correction for thermodynamic equilibrium. *PHYS. Rev.*, Vol. 40, pp. 749–759.
[8] Ville, J. (1946). Theorie et applications de la notion de signal analytique. *cables et Transmission*, Vol. 2, No. 1, pp. 61–74.
[9] Choi, H. and Williams, W.J. (1989). Improved time–frequency representation of multicomponent signals using exponential kernels. *IEEE Trans. Acoust., Speech, Signal Process.*, Vol. 37, No. 6, pp. 862–871.
[10] Shafi, I., Ahmad, J., Shah, S.I., Kashif, F.M. (2009). Techniques to obtain good resolution and concentrated time–frequency distributions–a review. *EURASIP Journal on Advances in Signal Processing*, Vol. 2009 (2009), Article ID 673539, 43 pages.
[11] Margenau, H. and Hill, R. N. (1961). Correlation between measurements in quantum theory. *Prog. Theor. Phys.*, Vol. 26. pp. 772–738.
[12] Hippenstiel, R. D. and Oliveira, P. M. de. (1990). Time varying spectral estimation using the instantaneous power spectrum (IPS). *IEEE Trans. Acoust., Speech, Signal Process.*, Vol. 38, pp. 1752–1759.
[13] Jeong, J. and William, W.J. (1992). Alias–free generalized discrete–time time–frequency distributions. *IEEE Trans. Signal Process.*, Vol. 40, pp. 2757–2765.
[14] Daubechies, I. (1990). The wavelet transform, time–frequency localization, and signal analysis. *IEEE Trans. Inform. Theory*, Vol. 36, pp. 961–1005.
[15] Rioul, O. and Flandrin, P. (1992), Time–scale energy distributions: A general class extending wavelet transforms. *IEEE Trans. Signal Process.*, Vol. 40, pp. 1746–1757.
[16] Bertrand, J. and Bertrand, P. (1988). Time–frequency representations of broadband signals. *Proc. IEEE Intl. Conf on Acoustics, Speech, and Signal Processing (IEEE ICASSP)*, pp. 2196–2199.
[17] Hagan, M.T. Demuth, H.B. & Beale, M. (1996). *Neural Network Design*, Thomson Learning USA.
[18] http://en.wikipedia.org/wiki/Data_clustering.
[19] Gonzalez, R.C. &Wintz, P. (1987). *Digital Image Processing*, 2nd Ed., Addison–Wesley.
[20] Shafi, I., Ahmad, J., Shah, S.I., & Kashif, F.M. (2006). Impact of varying Neurons and Hidden layers in Neural Network Architecture for a Time Frequency Application. *Proc. 10th IEEE Intl. Multi topic Conf., INMIC 2006*, pp. 188-193, Pakistan.
[21] Ahmad, J., Shafi, I., Shah, S.I., & Kashif, F.M. (2006). Analysis and Comparison of Neural Network Training Algorithms for the Joint Time–Frequency Analysis. *Proc. IASTED Intl. Conf on Artificial Intelligence and application*, pp. 193–198, Austria.
[22] Shah, S.I., Shafi, I., Ahmad, J., Kashif, F.M. (2007). Multiple Neural Networks over Clustered Data (MNCD) to Obtain Instantaneous Frequencies (IFs). *Proc. IEEE Intl. Conf. on Information and Emerging Technologies*, pp. 1–6, Pakistan.
[23] Shafi, I., Ahmad, J., Shah, S.I., Kashif, F.M. (2008). Computing De–blurred Time Frequency Distributions using Artificial Neural Networks. *Circuits, Systems, and Signal Processing, Birkhäuser Boston, Springer Verlag*, Vol. 27, No. 3, pp. 277–294.

[24] MacKay, D.J.C. (1992). A Practical Bayesian Framework for Back propagation Network. *Neural Computation*, Vol. 4, No. 3, pp. 448–472.
[25] Gray, R.M. (1990). *Entropy and Information Theory*, New York Springer–Verlag.
[26] Boashash, B. and Sucic, V. (2003). Resolution Measure Criteria for the Objective Assessment of the Performance of Quadratic Time–Frequency Distributions. *IEEE Trans. Signal Process.*, Vol. 51, No. 5, pp. 1253–1263.
[27] http://www–dsp.rice.edu.
[28] Baraniuk, R. G. and Jones, D. L. (1993). Signal–Dependent Time–Frequency Analysis Using a Radially Gaussian Kernel. *Signal Processing*, Vol. 32, No. 3, pp. 263–284.
[29] Jones, D. L., Parks, T. W. (1992). A Resolution Comparison of Several Time–Frequency Representations. *IEEE Trans. Signal Process.*, Vol. 40, No. 2.
[30] Jones, D. L. and Parks, T. W. (1990). A high resolution data–adaptive time–frequency representation. *IEEE Trans. Acoust., Speech, Signal Process.*, Vol. 38, pp. 2127–2135.
[31] Stankovic, LJ. (2001). A Measure of Some Time–Frequency Distributions Concentration. *Signal Processing*, Vol. 81, No. 3, pp. 212–223.
[32] Shafi, I., Ahmad, J., Shah, S.I., Kashif, F.M. (2007). Evolutionary time–frequency distributions using Bayesian regularised neural network model. *IET Signal Process.*, Vol. 1, No. 2, pp. 97–106



Bayesian Network

Edited by Ahmed Rebai

ISBN 978-953-307-124-4
Hard cover, 432 pages
Publisher Sciyo
Published online 18, August, 2010
Published in print edition August, 2010

Bayesian networks are a very general and powerful tool that can be used for a large number of problems involving uncertainty: reasoning, learning, planning and perception. They provide a language that supports efficient algorithms for the automatic construction of expert systems in several different contexts. The range of applications of Bayesian networks currently extends over almost all fields including engineering, biology and medicine, information and communication technologies and finance. This book is a collection of original contributions to the methodology and applications of Bayesian networks. It contains recent developments in the field and illustrates, on a sample of applications, the power of Bayesian networks in dealing the modeling of complex systems. Readers that are not familiar with this tool, but have some technical background, will find in this book all necessary theoretical and practical information on how to use and implement Bayesian networks in their own work. There is no doubt that this book constitutes a valuable resource for engineers, researchers, students and all those who are interested in discovering and experiencing the potential of this major tool of the century.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen