A Hybrid Neural Network and Maximum Likelihood Based Estimation of Chirp Signal Parameters

Samir Shaltaf[1] and Ahmad Mohammad[2]
[1]Department of Electrical Engineering, Aljouf University, Saudia Arabia
[2]Department of Electronics Engineering, Princess Sumaya University for Technonlogy, Jordan

Abstract: *This research introduces the hybrid Multilayer feed forward Neural Network (NN) and the Maximum Likelihood (ML) technique into the problem of estimating a single component chirp signal parameters. The unknown parameters needed to be estimated are the chirp-rate, and the frequency parameters. NN was trained with several thousands noisy chirp signals as the NN inputs, where the chirp-rate and the frequency parameters were embedded into those chirp signals, and those parameters were used as the corresponding NN output. The NN resulted in parameter estimates that were near the global maximum point. ML gradient based technique then used the NN output parameter estimates as its initial starting point in its search of the global point parameters. The ML gradient based search improved the accuracy of the NN parameter estimates and the new estimates were very much near the exact parameter values. Hence it can be said that NN working in corporation with the ML gradient based search results in accurate parameter estimates for the case of large signal to noise ratio.*

Keywords: *Chirp parameter estimation, frequency estimation, NN, ML.*

Received August 28, 2011; accepted December 29, 2011; published online August 5, 2012

1. Introduction

Chirp parameter estimation is a problem of well-known applications, such as radar, sonar, and seismography. Spectral analysis techniques have been used in chirp signals estimation and detection. Those techniques are mostly based on the Maximum Likelihood (ML) principle [1]. The corresponding ML function has many local optimal points beside the global optimal point which renders gradient-based search techniques unsuitable since they may get stuck at any one of these local optimal points. To overcome this problem, a simple but yet effective high resolution grid search is used but at a high computational cost. There are other suggested techniques for solving this problem; such as phase unwrapping [3, 4], which is suitable for estimation of single chirp signal under high signal-to-noise ratio. O'Shea [5] dealt with a cubic phase FM chirp signal where it is preprocessed and converted into quadratic phase chirp signal where then the 3D search problem is reduced to 1D search problem. An improved recursive cyclostationary based algorithm is suggested to obtain estimates of frequency and chirp rate parameters where error propagation effect is reduced resulting in improvement in estimation accuracy [7]. An approximate ML estimator was introduced with low computational complexity where a weighted linear combination of the phases of the received signal samples were used to obtain the chirp signal parameters with the aid of a phase unwrapping algorithm [8].

Our proposed research introduces multilayer feed forward Neural Network (NN) as the main estimation engine. Trained NN results in near global parameter estimates of a ML function. The NN then supplies its near global parameter estimates to a ML gradient based maximization technique. Once the gradient based technique is given the initial estimate it locks onto the region where the global maximum point exists and it steers its self onto that point. This hybrid technique results in accurate chirp signal parameter estimates.

The NN is trained with several thousands of chirp signals used as inputs and the chirp signal parameters as the corresponding outputs. The data window length used as the NN input had a size $N=256$ samples. Consider dealing with a chirp signal $s(n)=sin(\omega_0 n+\beta_0 n^2)$ with additive white Gaussian noise resulting in noisy signal $x(n)$.

$$x(n) = s(n) + w(n), n = 0,1,..., N-1 \quad (1)$$

where $w(n)$ is zero mean white Gaussian noise with variance σ_w^2 and uncorrelated with the chirp signal. The chirp signal parameters (ω_o, β_o) represent the frequency and the chirp-rate parameters respectively.

The NN was trained with large set of noisy chirp signals as its input, and the chirp-rate and frequency as its corresponding output parameters. Large signal to noise ratios were considered in this research since no filtering of the chirp signal was used. The noisy chirp signals must not be contaminated with large noise levels in order to obtain good parameter estimates from the NN. Filtering of chirp signals before applying them to the NN will be best, but classical filtering techniques can not produce clean chirp signals if the chirp parameter is large. It is well known fact that chirp signal with a large chirp-rate parameter value ends up with wide bandwidth. In such case, noisy chirp signals can not be adequately filtered using classical filtering techniques. Noise reduction of noisy holographic signals was considered by [2], where trained NN was used to filter noisy input signals and resulted in noise filtered signal.

The rest of this paper is organized as follows: Section 2 presents the details of the proposed NN technique for the estimation of the chirp signal parameters. Section 3 presents the infamous classical ML estimation technique used for fine tuning the NN estimated parameters. Section 4 presents the simulation results of the proposed technique and the total training and testing time required by the NN. Conclusion is presented in section 5.

2. Neural Network Based Chirp Parameter Estimation

It is well known that maximization of the likelihood function results in accurate reliable parameter estimates. But usually a likelihood function may have several maximum points, where a gradient based search technique may get stuck at any one of them. NN in this research were trained to obtain estimates that are near the global maximum point, and such near global estimates can be considered as good initial starting points for gradient based search techniques. Hence NN works jointly with gradient based technique to produce accurate ML based parameter estimates.

This research proposes an estimation technique that uses NN for a direct estimation of frequency and chirp-rate parameters of noisy chirp signals. Noisy chirp signals must not be contaminated with large noise levels in order to obtain good parameter estimates. Two layers and three layers NN were tested for the estimation of chirp-rate and frequency parameters. It was found through simulation that the three layers NN resulted in more accurate parameter estimates than that of the two layers NN; hence the two layers NN was excluded.

Hyperbolic tangent nonlinearity functions were used for the two hidden layer neurons, while linear transfer function was used for the output neurons. Resilient backpropagation training algorithm was used for training the network. Riedmiller *et al.* [6] had noticed that the nonlinear transfer functions; the hyperbolic tangent and the log sigmoid; of the neurons have very small gradient for large input values. The small gradient values result in slow convergence for the NN in the training phase because back-propagation is a gradient based learning algorithm. In order to overcome this problem, the sign of the gradient is used instead of its small value for updating the NN parameters. This results in a major improvement on the speed of convergence of the NN.

The noisy signal $x(n)$ in equation 1 is converted to a NN input vector $X=[x(0), x(1), \ldots, x(N-1)]$. The NN input size is equal to the input vector size which is N. The first hidden layer size was chosen to be 20 neurons, and the second layer size was 10 neurons. The output layer size was equal to the number of the parameters needed to be estimated.

A total of 2000 noisy chirp signals were used in the training phase as the NN inputs. The NN was trained directly with the input data without any preprocessing or filtering. In another phase, the NN was trained with preprocessed data. The preprocessing step was obtained through windowing the data vector using hanning window. It turned out that using windowed data resulted in more accurate parameter estimates.

In this research the data set is multiplied by the hanning window before applying it as input to the NN. This step is performed in order to minimize the effect of the endpoints of the data set. The endpoints of the input data vector are multiplied with the small values of the end points of the hanning window. Also, the middle part of the data set is enhanced by getting multiplied by the large values of the middle part of the hanning window.

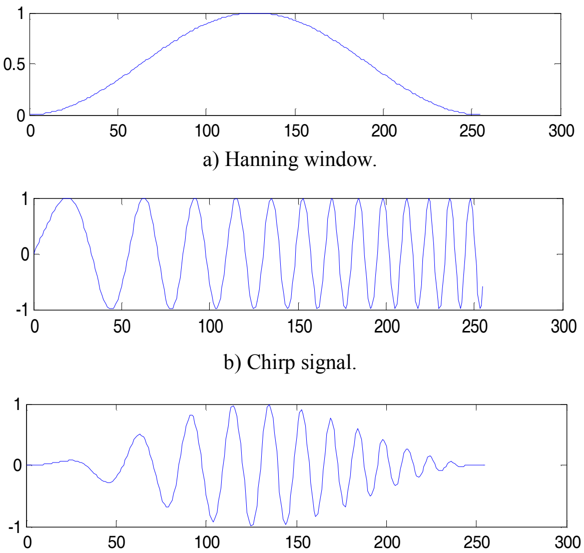
The hanning window $w(n)$ is obtained as follows:

$$w(n) = 0.5(1 - \cos(2\pi \frac{n}{N-1})), n - 0,1,\ldots, N-1 \quad (2)$$

The data $x(n)$ was multiplied with the window $w(n)$ and resulted in the windowed data $x_h(n)$ where

$$x_h(n)=x(n)w(n), \quad n=0, 1, \ldots, N-1 \quad (3)$$

The resulting $x_h(n)$ was used as the NN input signal. Figure 1 below presents the hanning window, the chirp signal, and the windowed chirp signal.



a) Hanning window.

b) Chirp signal.

c) Windowed chirp signal.

Figure 1. Presents the hanning window, the chirp signal, and the windowed chirp signal.

It is well known that the training phase of the NN requires long training time, especially when faced with large set of input vectors where the inputs are of large dimensions. But once the NN is trained, the total run time in the testing phase is minimal.

3. Maximum Likelihood Based Estimation

ML parameter estimates of chirp signal in equation 1 is obtained by maximizing the probability density function $f(X, \Phi)$ with respect to the parameters vector $\Phi=[\beta, \omega]$. The symbol X represents the input data vector $[x(0), x(1), \ldots, x(N-1)]$.

$$f(X;\Phi) = \frac{1}{(2\pi\sigma_w^2)^{N/2}} exp\left[-\frac{1}{2\sigma_w^2} \sum_{n=0}^{N-1} \left(x(n) - sin(\omega n + \beta n^2) \right)^2 \right] \quad (4)$$

Maximization of the density function can be obtained through minimization of $J(\Phi)$ where:

$$J(\Phi) = \sum_{n=0}^{N-1} \left(x(n) - sin(\omega n + \beta n^2) \right)^2 \quad (5)$$

Steepest decent approach is used to minimize $J(\Phi)$ in terms of ω and β as follows:

$$\hat{\omega}_k = \hat{\omega}_{k-1} - \mu \left. \frac{\partial J(\Phi)}{\partial \omega} \right|_{\hat{\omega}_{k-1}} \quad (6)$$

$$\hat{\beta}_k = \hat{\beta}_{k-1} - \mu \left. \frac{\partial J(\Phi)}{\partial \beta} \right|_{\hat{\beta}_{k-1}} \quad (7)$$

where $\hat{\omega}_k$ and $\hat{\beta}_k$ are the k_{th} estimates of the unknown parameters ω_o and β_o:

$$\frac{\partial J(\Phi)}{\partial \omega} = -2\sum_{n=0}^{N-1} n \left(x(n) - \sin(\omega n + \beta n^2) \right) \cos(\omega n + \beta n^2) \quad (8)$$

$$\frac{\partial J(\Phi)}{\partial \beta} = -2\sum_{n=0}^{N-1} n^2 \left(x(n) - \sin(\omega n + \beta n^2) \right) \cos(\omega n + \beta n^2) \quad (9)$$

Steepest descent gradient based search technique produces accurate parameter estimates if the initial parameters point is near the global minimum point of $J(\Phi)$. Usually they end up getting stuck at any of the local minimum points of $J(\Phi)$ if the initial point was far from the global minimum point and near a local one. For illustration purposes, consider dealing with the case of known frequency and unknown chirp parameter. The plot of the cost function $J(\beta)$ shown in Figure 2 corresponds to the case of noise free chirp signal with known $\omega_o=0.06$, and β_o assumed unknown where it was assigned a value of 0.0004 in the signal $s(n)$. The plot shows several local minimum points beside the unique global minimum point at the exact value of $\beta_o=0.0004$.

When dealing with noisy data, the NN produces near global chirp parameter estimates of the function $J(\Phi)$. Accurate parameter estimates can be obtained by the NN when low noise cases are considered. Higher noise levels results in higher estimation error variances. To minimize the NN estimation error variances, the gradient-based technique is used to fine tune the NN parameter estimates. Hence, NN is used as the first stage estimator where its estimates are not far from the true global optimal point. Then gradient based technique can be used to fine tune the NN estimates.

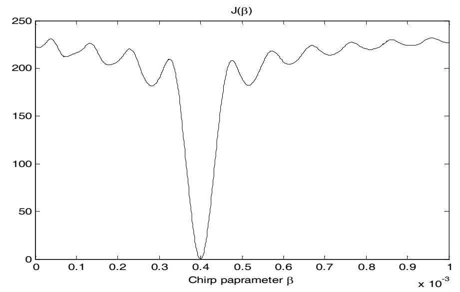


Figure 2. Function J(β) with minimum at β=0.0004.

4. Simulation

In this section several simulation scenarios were performed using NN. The NN was trained and tested for a single unknown chirp-rate parameter in subsection 4.1. Subsection 4.2 deals with the case of unknown frequency and chirp-rate parameters.

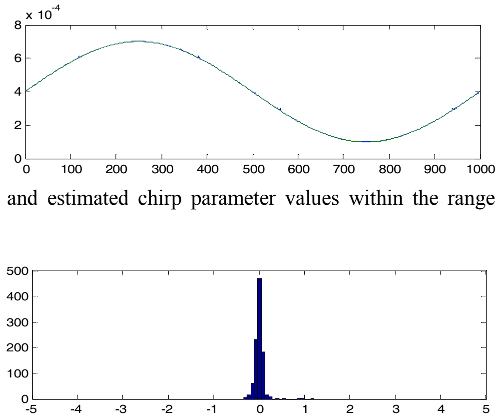
4.1. Chirp-Rate Parameter Estimation

In this simulation the chirp signal had a known frequency $\omega_o=0.06$ and unknown chirp β_o. In the training phase of the NN, 2000 chirp parameter values of the chirp signal $x(n)$ were randomly selected using a uniform random generator with a value in the range [0.0001 , 0.0007]. In the testing phase, 1000 chirp-rate parameter values where chosen from a curve representing the shape of a raised sine function. The raised sine function that generates the j_{th} chirp parameter value is:

$$\beta_0(j) = 0.0004 + 0.0003\, sin(0.002\pi j) \quad (10)$$

Each of these chirp-rate parameter values is embedded into the signal $s(n)$, where a Gaussian noise is added to $s(n)$ resulting in signal $x(n)$. The resulting 1000 signals are used in the testing phase of the NN. The trained NN then gives a chirp parameter estimate to each of these signals. Figure 3 below shows the exact and the estimated chirp-rate curves for the case of noiseless unwindowed data. Since noiseless data were used, the estimated parameters by the NN where very accurate to the point that the true parameters curve and the estimated parameters curve merged and looked as a single curve. To further illustrate the difference between the true and the estimated parameters, the estimation error histogram is presented in the lower part of Figure 3. The j_{th} estimation error between the exact j_{th} chirp parameter ($\beta_0(j)$) and its estimate ($\hat{\beta}_o(j)$) is presented in equation 11:

$$e(j) = \beta_o(j) - \hat{\beta}_o(j) \quad (11)$$



a) Exact and estimated chirp parameter values within the range [0.0001, 0.0007].

b) Histogram of the estimation error.

Figure 3. NN trained and tested with noiseless unwindowed data.

Table 1 Presents simulation results of NN trained with unwindowed (unprocessed) data. The first row is the NN test results corresponding to the case of NN trained with noiseless data. The second row is the NN test results corresponding to the case of NN trained with noisy data. In both cases, the NN was tested under three different Standard Deviation (STD) noise levels of 0.0, 0.01, and 0.1. Both rows present the Standard Deviation Estimation Error (SDEE) of the chirp-rate parameter obtained by the NN. Table 2 presents the SDEE of a NN trained and tested using windowed data. It is clear by comparing the results of Table 1 and Table 1 that NN trained and tested with windowed data resulted in smaller SDEE values. It is also clear that NN trained with noisy data produce smaller SDEE than those trained with noiseless data when tested with noisy data. Table 3 presents the estimation error of the steepest decent algorithm where it had its initial parameters from the NN. It is clear that the steepest descent SDEE is much better than that of the NN. That was the main idea behind using the steepest descent technique in corporation with the NN.

Table 1. SDEE of the chirp-rate parameter using unwindowed data.

	Testing Noise Levels		
	STD=0.0	**STD=0.01**	**STD=0.1**
NN Trained with Noiseless Data	6.2855e-4	0.0078	0.0963
NN Trained with Noisy Data	0.0076	0.0067	0.0491

Table 2. SDEE of the chirp-rate using windowed data.

	Testing Noise Levels		
	STD=0.0	**STD=0.01**	**STD=0.1**
NN Trained with Noiseless Data	3.2094e-4	0.0037	0.0423
NN Trained with Noisy Data	0.0016	0.0021	0.0153

Table 3. SDEE of the chirp-rate obtained with steepest descent gradient based method.

Testing Noise Levels		
STD=0.0	**STD=0.01**	**STD=0.1**
0.0000	2.6110e-4	0.0037

4.2. Chirp-Rate and Frequency Parameters Estimation

In this part, windowed data was used in the training and testing phases. The frequency parameter ω was chosen from a uniform random generator within the range [0.0, 0.08]. The chirp-rate parameter was also chosen from a uniform random generator within the range [0.0001 0.0008]. By comparing results of Table 4, it is clear that NN trained with noisy data results in smaller estimation error than NN trained with noiseless data when noisy data was used in the testing phase. Again, steepest descent algorithm results in improvement of the SDEE for the estimated parameters.

Table 4. SDEE of chirp-rate and frequency parameters.

Windowed	**Parameter Estimates**	**STD=0.0**	**STD=0.01**	**STD=0.1**
NN Trained with Noiseless Data	Chirp Freq	0.0056 0.0041	0.0213 0.0208	0.1910 0.1957
NN Trained with Noisy Data	Chirp Freq	0.0184 0.0176	0.0203 0.0197	0.0775 0.0954
Steepest Descent	Chirp Freq	0.0000 1.0738e-16	0.0011 0.0022	0.0105 0.0207

It can be seen that NN working in corporation with ML gradient based search technique resulted in more accurate parameter estimates than that of a NN working on its own. Gradient based search technique of the ML function would not have been able to obtain the global point parameters estimate if the NN was not used to supply its near global parameters estimate.

5. Simulation Time

The NN simulation was performed on a four year old Acer PC computer with Intel Centrino microprocessor of 2.0GHz processor speed. The three layers NN was trained with 2000 noisy chirp signals, and the total number of training epochs was chosen to be 5000 epochs. The total training time was $T_{training}$=5.02 minutes. The NN was tested with 1000 chirp signals and the total testing time was T_{test}=0.031 seconds, which implies that the testing time per a single chirp signal is 31×10^{-6} second. This shows how fast the processing speed of NN when running in the testing phase.

6. Conclusions

It can be said that NN in corporation with gradient based search technique results in reliable accurate global parameter estimates of a ML function for chirp signal parameters. Gradient search technique by itself can not obtain the global point parameters since the likelihood function have many local maximum points beside the global maximum point. The NN was capable of obtaining the near global point parameter estimates, then the gradient based search technique used the NN estimates to fine tune those estimates. Future work will consider working on finding a non classical filtering technique to filter wide band noisy chirp signals in order to obtain more accurate NN chirp parameter estimates.

References

[1] Abatzoglou T., "Fast Maximum Likelihood Joint Estimation of Frequency and Frequency Rate," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 22, no. 6, pp. 708-715, 1986.

[2] Badri L., "Development of Neural Networks for Noise Reduction," *The International Arab Journal of Information Technology*, vol. 7, no. 3, pp. 289-294, 2010.

[3] Djurie P. and Kay S., "Parameter Estimation of Chirp Signals," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, no. 12, pp. 2118-2126, 1990.

[4] Gershman A. and Pesavento M., "Estimating Parameters of Multiple Wideband Polynomial-Phase Sources in Sensor Arrays," *IEEE Transactions on Signal Processing*, vol. 49, no. 12, pp. 2924-2934, 2001.

[5] O'Shea P., "A Fast Algorithm for Estimating the Parameters of a Quadratic FM Signal," *IEEE Transactions Signal Process*, vol. 52, no. 2, pp. 385-393, 2004.

[6] Riedmiller M. and Braun H., "A Direct Adaptive Method for Faster Backpropagation Learning: the RPROP Algorithm," *in Proceedings of IEEE International Conference on Neural Networks*, San Francisco, vol. 1, pp. 586-591, 1993.

[7] Yan J. and Hongbing J., "Iterative Parameter Estimation of Chirp Signals Based on Cyclostationarity," *in Proceedings of IEEE Conference on Image and Signal Processing*, China, vol. 5, pp. 460-464, 2008.

[8] Yan L., Hua F., and Pooi Y., "Improved Approximate Time-Domain ML Estimators of Chirp Signal Parameters and Their Performance Analysis," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1260-1272, 2009.

Samir Shaltaf received his BSc and MSc degrees in electrical engineering in 1986 and 1988 respectively, MSc degree in Applied Mathematic in 1988, and PhD in sensing and signal processing in 1992, all from Michigan Technological University, USA. He worked at Amman University form 1992-1997, and at Princess Sumaya University since 1997 in Jordan. He is currently doing his Sabatical year at Aljouf University in Saudia Arabia. His research interests inclued in time delay estimation, neural networks, neuro fuzzy logic, adaptive filtering.

Ahmad Mohammad received his BSc in electrical engineering from Ein-Shams University in Egypt in 1981, MSc and PhD degrees in electrical engineering, from Akron University, USA, in 1989, and 1992, respectively. He was an assistant professor in the Electrical Engineering Department, at The Applied Science University, Jordan, from 1993 until 2000. He has been an assistant professor at the Department of Computer Engineering, Princess Sumaya University, Jordan, since 2000. His research interests include control systems, signal processing, and information security.