NEURAL NETWORK FOR SONOGRAM GAP FILLING

Henrik Klebæk, Jørgen Arendt Jensen, and Lars Kai Hansen

Electronics Institute, Build. 349, Technical University of Denmark, DK-2800 Lyngby, Denmark

Abstract

In duplex imaging both an anatomical B-mode image and a sonogram are acquired, and the time for data acquisition is divided between the two images. This gives problems when rapid B-mode image display is needed, since there is not time for measuring the velocity data. Gaps then appear in the sonogram and in the audio signal, rendering the audio signal useless, thus making diagnosis difficult.

The current goal for ultrasound scanners is to maintain a high refresh rate for the B-mode image and at the same time attain a high maximum velocity in the sonogram display. This precludes the intermixing of the B-mode and sonogram pulses, and time must be shared between the two. Gaps will appear frequently in the sonogram since, *e.g.*, half the time is spent on B-mode acquisition. The information in the gaps can be filled from the available information through interpolation. One possibility is to use a neural network for predicting mean frequency of the velocity signal and its variance. The neural network then predicts the evolution of the mean and variance in the gaps, and the sonogram and audio signal are reconstructed from these.

The technique is applied on *in-vivo* data from the carotid artery. The neural network is trained on part of the data and the network is pruned by the optimal brain damage procedure in order to reduce the number of parameters in the network, and thereby reduce the risk of overfitting. The neural predictor is compared to using a linear filter for the mean and variance time series, and is shown to yield better results, *i.e.*, the variances of the predictions are lower. The ability of the neural predictor to reconstruct both the sonogram and the audio signal, when only 50% of the time is used for velocity data acquisition, is demonstrated for the *in-vivo* data.

1 Introduction

Pulsed ultrasound systems are widely used for probing the human circulatory system for finding the distribution of blood velocities in arteries, veins, and heart chambers. The velocity distribution is found by emitting a series of pulses into the human body, and acquiring a single sample at a fixed depth for each pulse emitted. The velocity of the blood scatterers generates a signal with a frequency proportional to the velocity v_z of the blood along the ultrasound beam direction [1]. The frequency is given by

$$f_p = \frac{2v_z}{c} f_0 \tag{1}$$

where c is the speed of sound and f_0 the center frequency of the emitted pulse. A spread in velocities gives a corresponding spread in frequencies and the power density spectrum of the received, stochastic signal thus reveals the velocity distribution. A typical display is to show the power density spectrum as a function of time, where the x-axis is time, the y-axis is frequency, and the gray scale value corresponds to power density. Such a display is called a sonogram or spectrogram. A Fourier transform is done on small segments of data to calculate the power density. Often 64 to 128 pulse-echo lines are used, and this necessitates, *e.g.* 12 to 25 ms of data at a pulse repetition frequency of 5 kHz.

To assist in placing the interrogation volume or range gate, a dynamic B-mode image is shown alongside the sonogram display. Time for data acquisition is then split between the two images. The highest velocity that can be detected is dependent on the pulse repetition frequency f_{prf} as

$$v_{max} = \frac{c}{2} \frac{f_{prf}}{2f_0} \tag{2}$$

The interspacing of the B-mode pulses with the sonogram pulses then conflicts with the demand of a high v_{max}, since f_{prf} is reduced by the fraction of time used on the B-mode image. One solution is to alternate between acquiring a number of B-mode lines and sonogram lines, as shown in Fig. 1. Gaps then appear in the sonogram, when the B-mode image is acquired, and this severely distorts the audio signal. This is disruptive to the physician, and the gaps should be filled.

A gap filling approach has been suggested by Kristoffersen and Angelsen (1988). They used a FIR filter determined by the power density spectrum of the signal

Figure 1: Alternate duplex mode imaging.

prior to the gap, and passed noise through it to synthesize the audio signal. A second approach is to use the time reversed signal prior to the gap to fill the hole [1].

In this paper the use of a neural network is suggested. The network is trained on the spectral signal and then used for predicting the missing data to properly tack the time evolution of the non-stationary power density spectrum. Neural networks are used extensively for time series prediction (see, e.g., Weigend and Gershenfeld [3]).

2 Neural network solution

Our simulation tool for design of adaptive neural network architectures is described in detail in [4]. In particular, the networks are *pruned* to obtain the optimal connectivity. The advantage of the pruned networks is that they perform better on test data, *i.e.*, they do not simply memorize the training data but are able to *generalize* better than fully connected architectures. They are also able to select out the relevant inputs since inputs that have too low signal to noise ratios are pruned away. Based on statistical theories of generalization, the best generalization is expected from the least complex networks that can "solve the task".

In our simulator we search for the best network with a pruning scheme based on the Optimal Brain Damage technique of [5]. We first train a large network that can easily implement the training set. Subsequently we compute the *saliency* of the weights of the network and delete a fraction of the lowest saliency weights. The network is re-trained and the procedure repeated as long as the pruned network can implement the training set. For noisy problems, of course, one would tolerate a certain amount of error on the training set in order not to *overfit*. To quantify the tradeoff between misfit and overfit, hence providing a pruning stop criterion, we use a variant of Akaike's Final Prediction Error (FPE) [6], in particular our modified FPE takes into account that the network parameters are trained under a regularization constraint (weight decay), for more details on the training and regularization procedures consult [4].

The objective of the present system is to simulate the duplex mode audio signal, which, as seen in Fig. 1, can be characterized as a stochastic bandlimited signal with slowly varying spectral parameters (basicly driven by the heartbeat cycle). In order to simulate this signal we propose a simple model of the "instantaneous" power density spectrum at time t. Our model takes the form of a Gaussian peak centered at $\mu(t)$ and having a width $\sigma(t)$. Given a predicted time course of the parameters we can easily simulate the stochastic signal by multiplying the envelope on the spectrum of a white-noise random signal. It should be noted that the Gaussian assumption is by no means crucial for the approach, any parameterized spectral ansatz might be used, however, the Gaussian ansatz do provide a simple two-parameter approximation that we have found sufficient for our purposes [7].

The initial architecture is an ordinary feed-forward network with one hidden layer. Two networks are used to estimate the parameters of the model spectrum, one network estimates the center of mass of the power density spectrum $\mu(t)$ and another network estimates width $\sigma(t)$ of the power density spectrum. The networks are trained independently and both are equipped with a linear output units to accommodate an unlimited dynamic range. The activation function of the hidden neurons is the hyperbolic tangent.

The database for training of the network (estimation of network weights) is generated by estimating center of mass and width for an ultrasound signal (described in the next section). The target parameters were obtained by use of "trimmed" means and standard deviation calculation. The training error is the usual sum of squared errors and is calculated over a subset of the available data base, the remaining part been held out for testing of the system. The learning algorithm adjusts the weights of the network in order to minimize the error function. For the hidden layer, a second order modification of the back-propagation algorithm with batch learning is used [4]. The output neuron weights are trained using the matrix inversion algorithm described in [8].

The algorithm used to adapt the neural network architecture can then be summarized:

1. Train the fully connected neural network.

2. Use the Optimal Brain Damage method to calculate the saliencies of all the weights in the network.
3. Remove the 5% lowest ranking weights in the network according to saliency. If a hidden neuron is disconnected from the output neuron the hidden neuron is removed.
4. Retrain the network using the learning algorithm.
5. Stop when no more weights remain.

For each network configuration we have computed, the training error, the estimated test error (using FPE), and the true test error. As an estimate of the best generalizing network we chose the FPE-optimal network (*i.e.*, the network having the lowest estimated test error).

3 Experimental setup

The neural network algorithm was trained and tested on *in-vivo* data from the carotid artery of a healthy male. A B&K Medical 3535 ultrasound scanner with a 5 MHz type 8545 convex array transducer was used for the acquisition. The scanner was connected to our dedicated sampling system [9], that used a sampling frequency of 15 MHz, and acquired 50 rf samples for each pulse emission. The pulse repetition frequency was 5 kHz and a total of 19000 lines or 3.8 seconds of data was acquired.

The rf data was preprocessed by the matched filter/Hilbert transform approach described by Jensen (1996). A filter matched to the emitted pulse is first applied on the data and then a Hilbert transform is used to calculate the imaginary part of the signal. One complex set of samples in then extracted for each line, and this is the complex signal used for spectral estimation.

4 Example

To illustrate the neural net approach a subject was scanned and two time series of $\mu(t), \sigma(t)$ of length 940 were generated. The first 600 samples of each series were allocated for training while the remaining 340 samples were used for two contiguous test sets each of 170 samples. (We used two test sets to illustrate the variability in estimating the test error.) The networks were configured with an input field of 200 units using a simple lag-space representation as in [4]. Having 200 inputs, hence, leaves 400 samples for training. The networks were trained and pruned as described earlier and the FPE-optimal networks for each of the two time series problems were found. The selection procedure for the $\sigma(t)$ network is shown in Fig. 2. Note that the minimum of the FPE-estimate roughly coincides with the minimum of the first test set while the second test set apparently is predicted best with a somewhat smaller network. This is in line with the results of [4], where it was found that the prediction problem gets increasingly more difficult for longer extrapolation horizons. In table 1 we detail the results for both time series. For comparison we also provide the results for linear networks (corresponding to a simple linear regression model). We note that the pruning procedure in both cases significantly reduces the number of parameters and more importantly the test error (the results reported are mean based on 8 repetitions of the procedure). All errors are normalized by the total variances of the respective time series $(\mu(t), \sigma(t))$ computed from 940 samples.

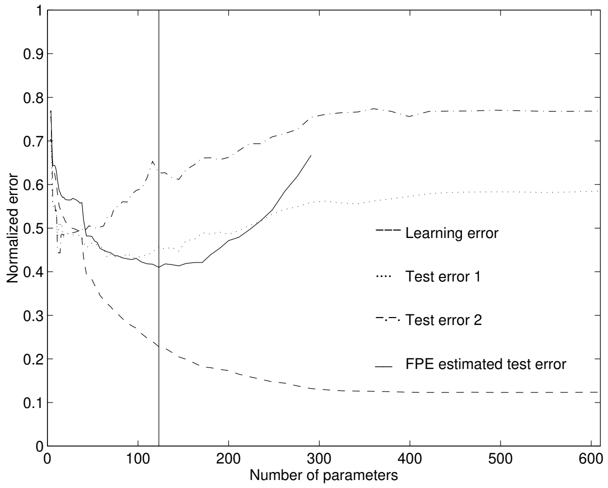

Figure 2: Final prediction error estimate and learning and test errors. The vertical line demarcates the FPE-optimal network.

The resulting FPE-optimal networks are then used to fill sonogram gaps as illustrated in Fig. 3. Both networks are initialized using 200 samples prior to the gap and predicted values are feed back as inputs while the input field slides across the gap. This procedure realises two time series of $\mu(t), \sigma(t)$ for the gaps. To simulate the noisy sonogram signal this spectral envelope is multiplied on the power spectrum of a white noise signal of the appropriate duration. A satisfactory resemblance to the original sonogram is seen.

Acknowledgment

This research is supported by the Danish Research Councils for the Natural and Technical Sciences through the Computational Neural Network Center (CONNECT).

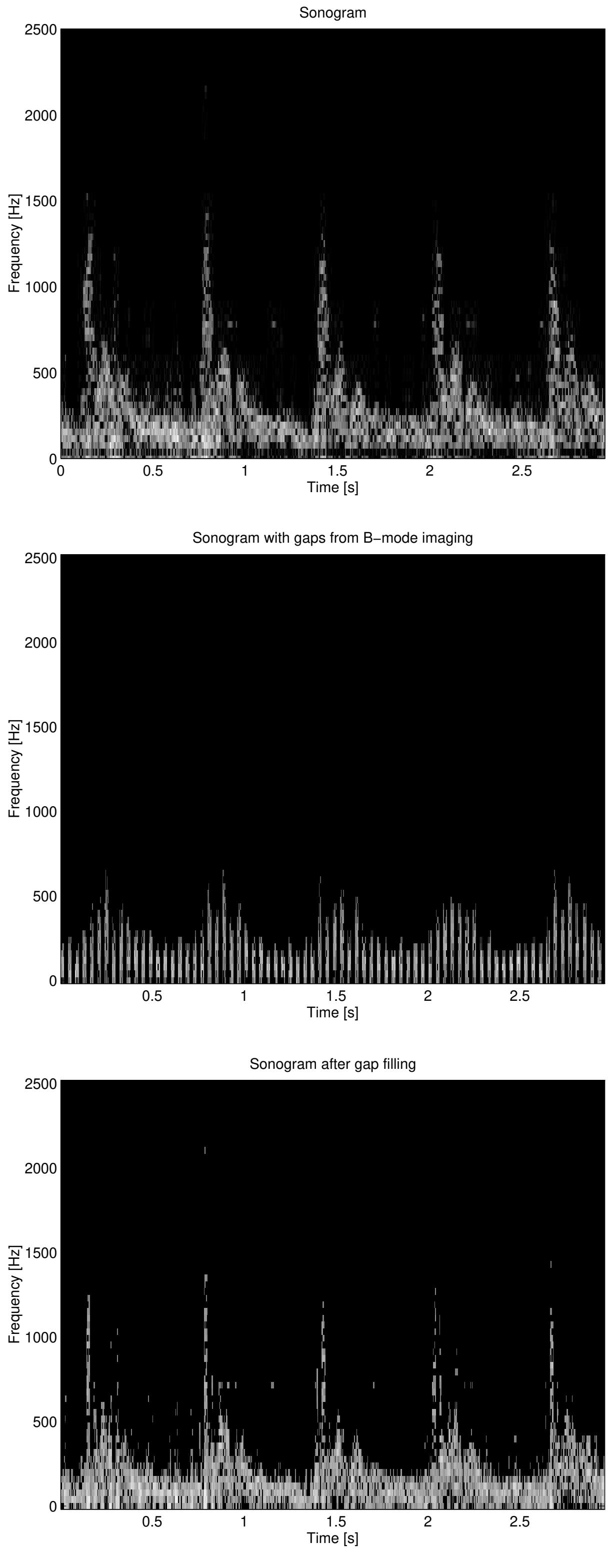

Figure 3: Sonograms from the carotid artery. The top sonogram is the original, the second graph shows it with gaps, and the last with artificially generated segments inserted.

Model	Learning	Test 1	Test 2	Para.
Linear model	0.1381	0.6360	0.6336	201
Full network	0.1557	0.3998	0.4839	607
Pruned network	0.1903	0.3734	0.4600	35-38

Model	Learning	Test 1	Test 2	Para.
Linear model	0.4122	0.6815	0.7336	201
Full network	0.1257	0.5623	0.7670	607
Pruned network	0.2261	0.4495	0.6387	116-123

Table 1: Normalized learning error, test error for two test sets, and number of parameters in the networks for the mean value time series (top table) and for the standard deviation time series (bottom table). The standard deviation on all test error estimates are all less than 0.5% of the estimate.

References

[1] J. A. Jensen. *Estimation of Blood Velocities using Ultrasound: A Signal Processing Approach.* Cambridge University Press, New York, 1996.

[2] K. Kristoffersen and B. A. J. Angelsen. A time-shared ultrasound Doppler measurement and 2-D imaging system. *IEEE Trans. Biomed. Eng.*, BME-35:285–295, 1988.

[3] A. S. Weigend and N. A. Gershenfeld, editors. *Times Series Prediction: Forecasting the Future and Understanding the Past.* Addison-Wesley, Redwood City, CA, 1994.

[4] C. Svarer, L.K. Hansen, and J. Larsen. On design and evaluation of tapped-delay neural network architectures. In E.H. Ruspini et al., editor, *IEEE Int. Conf. Neural Networks*, pages 45–51, 1993.

[5] Y. Le Cun, J. S. Denker, and S. A. Solla. Optimal brain damage. In D. S. Touretzky, editor, *Advances in Neural Information Processing Systems II*, pages 396–404. Morgan Kaufman, San Mateo, 1989.

[6] H. Akaike. Fitting autoregressive models for prediction. *Ann. Inst. Stat. Mat.*, 21:243–247, 1969.

[7] H. Klebæk. Neural network for estimation of ultrasound signals. Master's thesis, Tech Univ. Denmark, Electronics Institute, 1995. (in Danish).

[8] S. A. Barton. A matrix method for optimizing a neural network. *Neural Computation*, 3:450–459, 1991.

[9] J. A. Jensen and J. Mathorne. Sampling system for in vivo ultrasound images. In *Med. Imag. V Symp.*, volume SPIE Vol. 1444, pages 221–231, 1991.