Murdoch UNIVERSITY

MURDOCH RESEARCH REPOSITORY

http://dx.doi.org/10.1109/ASSPCC.2000.882494

Jan, T., Zaknich, A. and Attikiouzel, Y. (2000) Separation of signals with overlapping spectra using signal characterisation and hyperspace filtering. In: The IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium AS-SPCC., 1 - 4 October, Lake Louise, Canada, pp. 327 - 332.

http://researchrepository.murdoch.edu.au/20194/

Copyright © 2000 IEEE

Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Separation of Signals with Overlapping Spectra using Signal Characterisation and Hyperspace Filtering

Tony Jan
The University of Melbourne
Parkville, Victoria, Australia
tonyj@ee.mu.oz.au

Anthony Zaknich
The University of Western Australia, WA, Australia
tonko@ee.uwa.edu.au

Yianni Attikiouzel
The University of Western Australia, WA, Australia
yianni@ee.uwa.edu.au

Abstract

For separation of signals with overlapping spectra, classical linear filters fail to perform effectively. Nonlinear filters such as Volterra filters or Artificial Neural Networks (ANNs) can perform better but their implementations are often impractical due to their computational complexity. In this paper an ANN based hyperspace signal modeling is used to separate signals with overlapping spectra. The computational complexity of the ANN is reduced significantly by a simple feature extraction utilizing the unique temporal characteristics of the signals. The results show that difficult signal separation and filtering can be achieved efficiently by employing an ANN and an effective feature extraction.

1.0 Introduction

The frequency spectrum of human speech ranges between 400 - 4000 Hz for telephone communications. A linear Low Pass Filter (LPF) with cutoff frequency of approximately 4.3 kHz can be used to recover the speech signal and attenuate other undesired signals. However when additive interference between analogue speech and digital pulse signals occurs, classical linear filters can not separate the two signals effectively because of the large spectral overlap.

Two spectra are displayed in figure in order to explain the paragraph above. The dotted spectrum is a noiseless analogue speech signal filtered by a 4.3 kHz linear LPF. The solid line spectrum represents a mixture of analogue speech and digital pulse signals filtered by a 4.3 kHz linear LPF.

The spectrum shows undesired signal components which have been transmitted through the linear LPF. These undesired signals of overlapping spectra degrade the speech signals to unintelligible sounds contaminated by mechanical type noise.

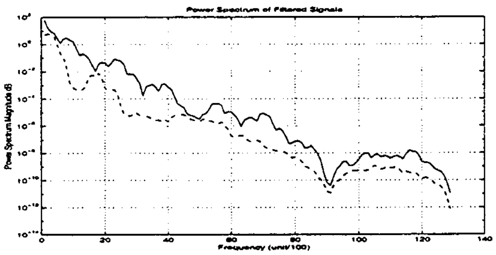

Figure 1: Power Spectra of Filtered Signals

In such a case, a simple and efficient system is required that can separate analogue speech signals from digital pulse signals in order to enhance the quality of voice communications.

Multi-band noise cancellation systems [1] and sub-band filter-banks [2] proved useful in general noise cancellation but did not improve the cancellation of noise with overlapping spectra [2]. Moir and Campbell [3] used a multi-band nonlinear FIR (MBNFIR) based noise canceller for speech enhancement however the noise with overlapping spectra remained an unsolved serious problem.

A differentiator-integrator pair can be used to extract and remove the digital pulse signal. This time domain filtering is simple and effective but its performance is not satisfactory when the amplitude of the digital square signals is small compared to the amplitude of the analogue speech signals.

Le [4] used a Multi-Layer Perceptron (MLP) to remove noise with overlapping spectra such as low bit CELP code noise added to speech signals. The MLP successfully removed the noise, however the size of training data set, the time varying nature of the signals and unavailability of the training pairs for all relevant input-output pairs remain unsolved.

0-7803-5800-7/00/$10.00 ©2000 IEEE

Over the past decade, there has been an increasing interest in the use of Artificial Neural Networks (ANNs) for solving complex real-world problems [5], [6]. ANNs can perform well for many difficult tasks such as the separation of signals with overlapping spectra which seemed almost impossible with classical linear filtering techniques. However, the implementations of ANNs in real applications are often limited because of their massive computational complexity. In order to use ANNs, it is imperative to reduce the complexity of problems by applying an effective feature extraction.

In this paper an ANN is used to separate signals with overlapping spectra using the temporal characteristics of the signals. However, computational efficiency is attained by employing a feature extraction and a simple hyperspace signal modeling.

The proposed approach for the separation of the analogue speech and digital pulse signals is as follows:

1. The signal space representation of the mixed signal is analyzed to select the most distinct features of the signals that can simplify the signal separation. (Refer section 2.1)
2. Using the selected temporal characteristics of the signals, the ANN performs hyperspace signal modeling. (Refer section 2.2)
3. The output of the modeling ANN system (estimate of digital signal) is used to extract the pulse signal components from the mixture of pulse and speech signals.

The efficiency and effectiveness of this approach is compared with that of MLP and other signal separation methods such as the differentiator-integrator pair and the linear filters. Refer to section 3.4.

Though a simple example is used in this contribution, the results reveal a very important underlying feature extraction and hyperspace processing concept useful for many signal processing applications including blind source identification/separation.

2.0 Signal Characterization and Hyperspace Filtering

The ANN is used to model the digital pulse signals using hyperspace processing. Hyperspace processing involves deliberately mapping inputs to the most effective feature space in which the desired separation is easily achieved and then mapping back to the required response space [9]. (see figure 4)

In order to achieve effective hyperspace processing, it is very important to find the most distinguishable features of the signals. In this section, the time delayed input signals are observed for feature extraction.

2.1 Signal Space Representation

When the tap delayed mixed signals of noiseless speech and digital pulse signals are shown in two dimensional signal vector space, the signal constellations of speech and digital signals can be easily distinguished and this can be used for signal separation. For more complex problems, higher dimensions or other data transforms [9] are required for effective signal separation.

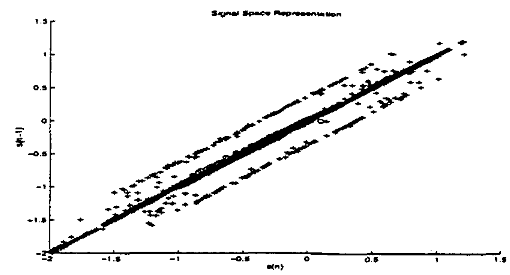

Figure 2: Signal Space Representation

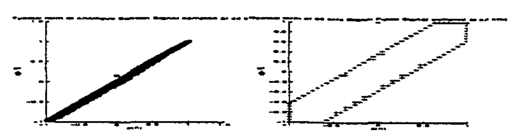

Figure 3 a, b: Typical Formations of the Signals

The two dimensional signal vector space of a noiseless mixture of analogue and digital signals is shown in figure 2. The signal constellations show clear separation. The digital pulse signal components form rectangular constellations (fig. 3b) because of their discontinuities whereas the analogue speech signals components form sharp elliptical signal constellations (fig, 3a) within the middle part of the two dimensional signal vector space. The most distinct feature of the signals proved to be the discontinuities of the pulse signals. This feature is used to simplify the hyperspace signal modeling. (see figure 4)

2.2 Hyperspace Processing using an ANN

The signal modeling of the pulse signals requires a particular nonlinear mapping between the input space and the desired response space. (see fig. 4)

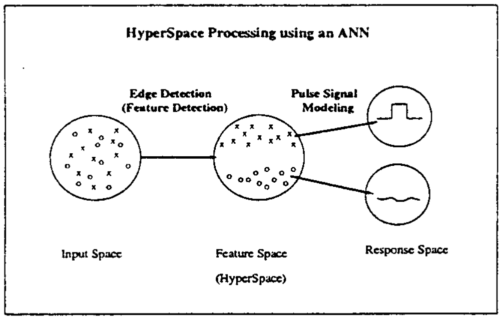

Figure 4: Hyperspace Processing

The mapping between the input space and the response space without a feature space can be extremely complicated because of the complex nature of human speech in the input space. In order to simplify the nonlinear mapping, the feature space is devised that reduces the dimensionality of problems by removing redundant information. Utilizing the edge feature of the pulse signals on the feature space, the nonlinear mapping process reduces to a simple edge detection and pulse signal modeling based on an ANN. (see fig. 4)

To perform the edge detection and pulse signal modeling, the training data set requires examples of input-output pairs that can train the ANN to detect discontinuities in the incoming input signals and build a digital signal model when the discontinuities are detected. In other words, the training of the ANN should produce zero output when the input signal vectors fall into the signal constellation in figure 3a. But when the testing input signal vectors fall into other signal constellation as in figure 3b, the ANN should generate the digital pulse signal model. The size and complexity of the training data is small and efficient since the training data set carries only the training samples that represent the discontinuities in the input signals and their corresponding outputs. Refer section 3.1 for further detail.

In order to model the digital pulse signals using the training data set, the Modified Probabilistic Neural Network (MPNN) [13] is selected for the following reasons.

1. Nonlinear function approximation ability.
2. Good performance in varying noise conditions [13].
3. Improved performance over the differentiator-integrator pair (Refer to section 3.4).
4. More structural information (radial basis function) available about the ANN [10].

The modified probabilistic neural network was initially introduced by Zaknich et al in 1991 [14]. It is closely related to Specht's general regression neural network [11] and both are related to Specht's Probabilistic Neural Network (PNN) [12] classifier.

If it can be assumed that there is a corresponding scalar output y_i for each local region of the input space which represented by a centre vector $\mathbf{c}_i$, then the general algorithm of MPNN given in equation 1 can be used for nonlinear function approximation within any acceptable accuracy.

The general algorithm for the MPNN/GRNN is:

$$\widehat{y}(\mathbf{x}) = \frac{\sum_{i=1}^{M} Z_i y_i f_i(\mathbf{x})}{\sum_{i=1}^{M} Z_i f_i(\mathbf{x})} \tag{1}$$

where

$$f_i(\mathbf{x}) = exp\frac{-(\mathbf{x}-\mathbf{c}_i)^T(\mathbf{x}-\mathbf{c}_i)}{2\sigma^2} \tag{2}$$

$\mathbf{x}$ = Input Vector.
$\mathbf{c}_i$= centre vector for class i in the input space.
σ = Learning parameter.
y_i= output related to $\mathbf{c}_i$
Z_i = no. of vectors $\mathbf{x}_j$ associated with each $\mathbf{c}_i$.
M = Number of unique centres $\mathbf{c}_i$.

A Gaussian radial basis function is often used for $f_i(x)$ as defined in equation 2. However, many other suitable radial basis functions can be used. Training simply involves finding the optimal σ giving the minimum mean squared error (mse) of the network output minus the desired output for a representative testing set of known sample vector pairs by a convergent optimization algorithm [10]. In summary, the MPNN is capable of any nonlinear mapping which makes them the perfect candidate for nonlinear signal modeling.

The MLP can be trained in a similar way but shows less flexibility for varying noise conditions. This is discussed in section 3.4.

2.3 Recovery of Signals

The final estimate of digital pulse signal is used to extract the digital signal components from the mixture of analogue speech and digital pulse signals. This results in the recovery of both speech and pulse signals. Refer to section 3.2 and 3.3 for the experimental results.

3.0 Experimental Results and Analysis

The signal separation method described above was applied to real speech and digital pulse signals.

3.1 Experiments

Figure 5a shows the 34150 points of human speech signal sampled at 44100 Hz. In order to recover this signal from the contaminated speech signals (figure 5b), the training data set was developed that satisfies the conditions described in section 2.2. The whole training data set consisted of 3500 samples. Figure 6 shows part of the training data set.

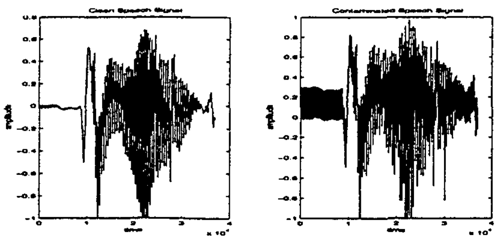

Figure 5 a, b: Real Speech Signals.

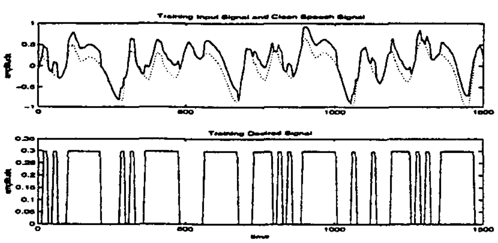

Figure 6: Training Data Set.

The training data set was simple and efficient because of the effective feature extraction employed. The training data set included only the sample input-output pairs that represented the discontinuities in the input signals.

In order to achieve signal separation in noisy (9.8 dB SNR) conditions, a few different input vector dimensions of the training data set were used and the input vector dimension of 8 showed the best performance. The most effective signal separation in noisy conditions was achieved in 8 dimensional signal vector space. The output of MPNN was an estimate of the pulse signal which was later used to extract the pulse signal components from the mixed signals by subtracting it from the mixed signal. The results are shown in the section 3.2 and 3.3.

In order to compare the network performance, the same training data set was used to train both the MLP and the MPNN. The MLP had a 8-3-1 architecture (a single hidden layer with three neurons and the output layer with a single neuron) and 10,000,000 training iteration with Back Propagation Learning algorithm [6]. The training of the MPNN was described in section 2.3. The performances of the MLP and MPNN are compared in section 3.4. The comparisons with other signal separation methods such as a linear FIR filter with 8 taps are also discussed in section 3.4.

The next two sections show the results of the MPNN with hyperspace processing.

3.2 Time Domain Results

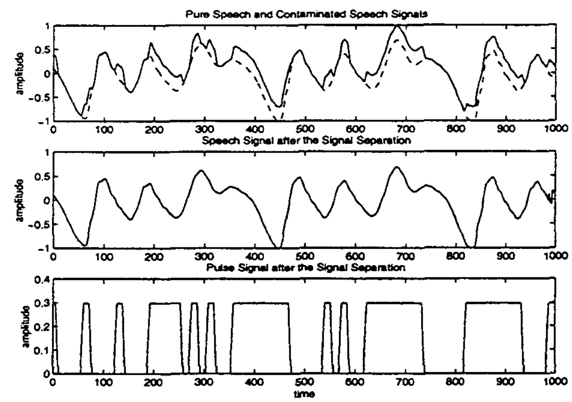

Figure 7: Time Domain Results

The results of the MPNN show complete recovery of the digital pulse signal as well the analogue speech signal. The Signal to Noise Ratio of the speech signal improved from 11.7238 dB to 36.1531 dB after the hyperspace processing.

3.3 Frequency Domain Results

The spectra of the filtered output and the input are shown in figure 8. The results show that the overlapping spectrum of undesired signal was completely removed.

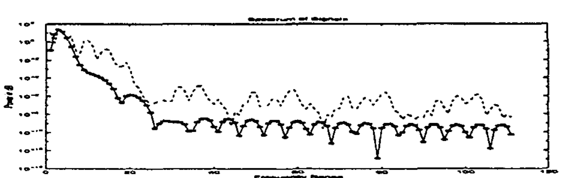

Figure 8: Frequency Domain Results
dotted line: before filtering
star line: after filtering

3.4 Performance Comparison

In order to evaluate the effectiveness and efficiency of the proposed approach, a number of comparisons were made as described in the following subsections.

3.4.1 Linear Filters

The Signal to Noise Ratio (SNR) at the filtered speech output was used to compare the effectiveness of the filters. The hyperspace processing resulted in SNR of 36.1531 dB whereas the linear filters resulted in SNR of 11.3442 dB at the speech output. As expected, the performance of hyperspace filtering was superior.

3.4.2 Differentiator and Integrator pairs

When the amplitude of the pulse signal was small or the noise level was too high, the differentiator and integrator could not perform effectively. This approach may cause trouble in real communication applications where the noise level can not be guaranteed to be small. Also, this method is potentially prone to error for impulse noises in the input signals. The application of the differentiator- integrator pair is limited.

3.4.3 MLP

The performance of MLP with a single hidden layer was compared with the hyperspace filtering based on the MPNN. The results confirmed the superior performance of hyperspace filtering based on the MPNN over the MLP in both the noiseless and varying noise conditions. The SNR of filtered speech outputs are listed in the table below.

Input Noise Level	MLP	MPNN
12.5182	21.3568	36.1531
10.7317	21.6874	26.0143
8.9823	20.6385	22.9071
5.4774	16.1955	18.1916

Table 1: Comparison of SNR at each filter output (All units are in dB).

The performance of the MLP can improve as the size of the training data set increases, but it also increases the training time and the complexity of the MLP.

4.0 Conclusion and Future Works

The use of an ANN allows almost complete recovery of digital pulse signals as well as recovery of analogue speech signals. This was not possible with a classical linear filter. The method proved to be effective and also efficient because of the reduced computational requirements and training data set by employing hyperspace feature processing. The method presented in this paper applies specifically to the separation of two signals, utilizing temporal signal characteristics. However, this approach of feature extraction and hyperspace filtering is clearly more general than the specific example presented. Research will continue on development of an automated feature extraction and hyperspace filtering algorithm. Such an algorithm can benefit many signal processing applications. For example, effective unsupervised hyperspace filtering algorithms can be used for blind signal separation or identification approaches.

References

[1] E. Toner and D.R. Campbell, *Speech Enhancement using sub-band intermittent adaption*, International Journal of Speech Communication, Vol. 12, pp. 253-259, 1993.

[2] D. J. Darlington and D. R. Campbell, *Sub-band adaptive filtering applied to speech enhancement*, 4th International Conference on Spoken Languages, ICSLP 1996, vol.2, pp. 921-924.

[3] A. Hussain, D. R. Campbell and T. J. Moir, *New Adaptive Non-linear FIR filters for sub-band processing in speech enhancement systems*, IEE Colloquium (digest), 1996, no. 238, IEE, Stevenage, Engl. p 6/1-6/6.

[4] T. T. Le and J. S. Mason, *Artificial neural networks for nonlinear time-domain filtering of speech*, IEEE Proceedings on Vision, Image and Signal Processing, 1996, vol. 143, pp. 149-154.

[5] D. R. Hush and B. G. Horne, *Progress in Supervised Neural Networks: What's new since Lippmann*, IEEE Signal Processing Magazine, pp. 9-39, Jan 1993.

[6] S. Haykin, *Neural Networks expand Signal Processing's Horizons*, IEEE Signal Processing Magazine, pp. 25-49, March 1996.

[7] P. J. Rayner, M. R. Lynch, *A new connection model based on a non-linear adaptive filter*, IEEE ICASSP, pp. 1191-1194, Glasgow, 1989.

[8] ;Moody, J. and Darken, C., "Learning localized receptive fields", Proceedings of the 1988 Connectionist Models Summer School, Editors Touretzky, Hinton and Sejnowski, Publishers Morgan-Kaufmann, 1988, pp. 133-143.

[9] Haykin, S., "Neural Networks, A Comprehensive Foundation", Macmillan College Publishing Co. Inc., 1994.

[10] Zaknich, Anthony and Attikiouzel, Yianni, "Automatic optimization of the modified probabilistic neural network for pattern recognition and time series analysis", Proceedings of the First Australian and New Zealand Conference on Intelligent Information Systems, Perth, Western Australia, December, 1993, pp. 152-156.

[11] Specht, D. F., "A general regression neural network", IEEE Transactions on Neural Networks, Vol. 2, No. 6, November 1991, pp. 568-576.

[12] Specht, D. F., "Probabilistic neural networks", International Neural Network Society, Neural Networks, Vol. 3, 1990, pp. 109-118.

[13] Zaknich, A. and Attikiouzel, Y., "A modified probabilistic neural network signal processor for nonlinear signals", Proceedings of the 13th International Conference on Digital Signal Processing, Santorini, Greece, 1997, July 2-4, pp. 291-294.

[14] Zaknich, Anthony and Attikiouzel, Yianni, "Time series characterization schemes for the modified probabilistic neural network", Australian Journal of Intelligent Information Processing Systems (AJIIPS), Vol. 2, No. 2, Winter 1995, pp. 1-11.