Technical Report #76 (Institute for Brain and Neural Systems, Brown University)

Feature extraction from acoustic backscattered signals using wavelet dictionaries

Nathan Intrator
Institute for Brain and Neural Systems*
Brown University
Providence, RI 02912

Quyen Q. Huynh
Institute for Brain and Neural Systems and
Coastal Systems Station
Naval Surface Warfare Center
Panama City, FL 32407-7001

Gerald J. Dobeck
Coastal Systems Station
Naval Surface Warfare Center
Panama City, FL 32407-7001

June, 1997

1 Introduction

Discrimination problems differ in nature from reconstruction tasks. While in reconstruction, it is the mean squared error that is often used to measure the quality of the scheme, classification requires a different measure which often is not related to the former. The discrimination power of a certain basis or a set of basis function is not necessarily connected to the quality of reconstruction associated with this set. Furthermore, the degree of relevance of the orthonormality constraint to the quality of the discrimination is questionable. For example, linear discriminant analysis [1] searches for linear projections which maximize the between-class variance divided by the sum of within-class variance. Such projections do not necessarily coincide with the principal components of the data which are the directions that optimize MSE reconstruction.

A successful approach to discrimination is based on an appropriate preprocessing to create an efficient signal representation, which then leads to an efficient dimensionality reduction. The next step is again some combination of feature extraction and classification. In this paper we briefly review several methods for finding data representation via optimal decomposition of wavelet basis functions and discuss their reconstruction properties. We then discuss some signal decomposition methods for the purpose of discrimination. This is followed by a brief discussion on a combination of feature extraction and classification scheme and with discrimination results on two acoustic data-sets.

*On leave, School of Mathematical Sciences, Tel-Aviv University

2 Optimal basis function decomposition for reconstruction

2.1 Entropy based algorithms

Coifman and Wickerhauser [2] presented a simple and fast algorithm for finding the local best basis (BB) in a wavelet packet (WP) library basis functions. The search is very simple and fast due to the orthogonality condition between the basis functions at each level and the inclusion properties of basis functions between different levels. Choosing between different possible bases is done via the entropy of the coefficients, namely the speed of decay in coefficient values, which indicated the degree of compression of the representation.

2.2 Basis pursuit

Unlike the search in orthogonal bases as done in the best basis method, one can search in an overcomplete dictionary of basis functions. This has been proposed by Daubechies and termed the Method of Frames [3]. Among different representations for the same signal, one searches for a representation whose vector of coefficients has the smallest l^2 norm. This approach leads to a quadratic optimization problem that is solved via a system of linear equations. Recently Chen et al. [4] presented a Basis Pursuit method (BP) that is very similar to the methods of frames; It decomposes a signal using dictionary elements so that the coefficients have the smallest l^1 norm among all such decompositions. This optimization can be performed by advanced linear programming techniques [5]. Chen et al. demonstrate that for certain signals, the convergence of a basis pursuit algorithm is faster than that of a best basis representation.

2.3 Matching pursuit

The Matching Pursuit algorithm [6] is an iterative algorithm, which does not explicitly seek any overall goal, but merely applies a simple and local rule repeatedly. It is a forward model selection that adds at each step the single most correlated new atom among all those not included yet in the model. The algorithm is very powerful for orthogonal basis selection, but may fail for non-orthogonal dictionaries.

3 Optimal basis decomposition for discrimination

3.1 Local discriminant bases

The local discriminant base (LDB) [7, 8] creates a time-frequency dictionary such as WP or local trigonometric functions (CP), from which signal energies for each basis coordinates are accumulated for each signal class separately. Then, a complete orthonormal basis is formed using a distance measure between the distributions of those energies from each class.

The original algorithm [7] attempted to extract best basis from the energies (squared values) of the WP, which is the direct approach to finding a best basis for a class of patterns [2]. Unfortunately, when the distance measure is applied to these energy coefficients, or more generally to the distribution of the energies, then the interpretation of the new basis is not clear anymore and the optimality properties are not so apparent. Moreover, noticing that the energies may not be so indicative for discrimination, Saito and Coifman [8] have suggested to use a different non-linear function of the basis function of the coefficients (instead of a just square values) so as to alleviate this problem. However, this approach takes us even further away from interpretation and optimality of the best basis approach.

3.2 Discriminant pursuit

Buckheit and Donoho [9] have introduced the discriminant pursuit (DP) algorithm which follows the approach of basis pursuit, in the sense that it is not constrained by seeking only orthogonal discriminant basis functions, but can search in the overcomplete WP or CP dictionary. The discrimination power of each basis function is measured by:

$$D_i(X,Y) = \frac{|E_X[wp_i(x)] - E_Y[wp_i(y)]|}{\mathrm{STD}_X(wp_i(x)) + \mathrm{STD}_Y(wp_i(y))}, \tag{1}$$

which is a 1-dimensional form of Fisher discriminant analysis criterion [1]. It is our experience that often, the additional flexibility leads to inferior results. This happens when the dimensionality is high and the number of training patterns is relatively small. If the WP representation is sparse, then for every basis function there are very few patterns which contribute to its value, thus the variability is large and outliers are more likely to cause trouble. There is another problem associated with this approach; Since the wavelet packet transformation is linear, it follows that $E_X[wp_i(x)] = wp_i(E_X[x])$. Thus, if the mean of each signal set is zero, there is no discrimination power in the means. A simple example is the discrimination between two signals of the form: $\sin(\omega t + u)$ and $\sin(2\omega t + u)$, where $u \sim U[0, 2\pi]$.

4 Hybrid dimensionality reduction/discrimination scheme

Because of the potential benefits of bringing all possible kinds of information to bear on the problem of dimensionality reduction, numerous attempts have been made to combine unsupervised with supervised learning for that purpose. Typically, these approaches use a hybrid learning rule to train a network, which then develops a reduced-dimensionality representation of the data at its hidden layer. We have presented a general framework based on projection pursuit regression, in which a penalty term may be added to the cost function minimized by error back propagation, for the purpose of measuring directly the goodness of the projections [10] (see Figure 1). This

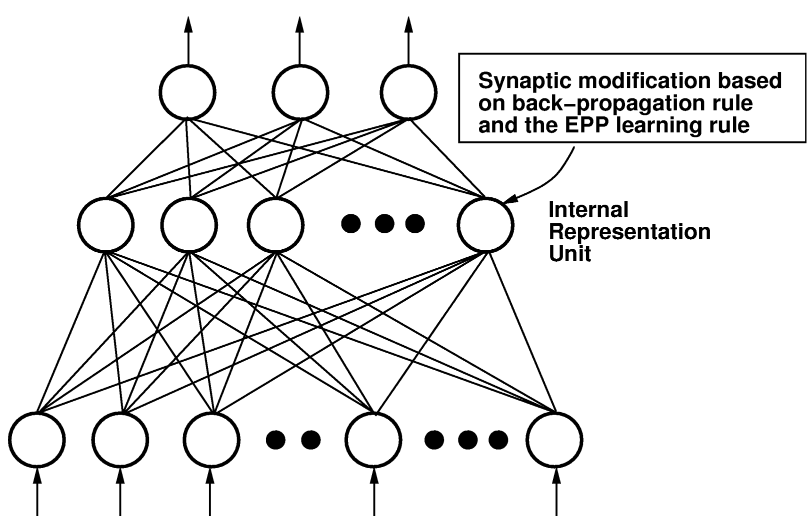

Figure 1: A hybrid EPP/PPR neural network (EPPNN).

emphasizes the choice of the "right" prior, as a means to improve the bias/variance tradeoff [11] which is much more crucial in overparametrized problems. Penalty terms derived from projection

Whale/Porpoise Discrimination

	Coiflet-3/3
BB	10.5(0.3)
LDB	52.4(1.3)
Wavelet	12.6(0.5)
Quadratic	30.4(0.7)
LDA	13.9(0.7)
BB 9	11.9(0.4)
BB Daub-4	14.8(0.8)

Table 1: Percent classification error rate of whale porpoise discrimination using various discriminant basis representations from a Coiflet-3 mother wavelet. The discrimination is performed from most discriminative 3 coefficients extracted by each method.

pursuit constraints tend to be more biased towards the specific problem at hand, and therefore are more likely to better control the variance while not inflating the bias contribution to the error.

Such an architecture which combines feature extraction and classification was found to be superior to sequential methods of first extracting features and then classifying in several real-world applications [10, 12, 13]

4.1 Non-linear feature extraction from wavelet representation

This sections briefly discusses an unsupervised learning rule that is combined with the supervised learning rule (as described above). This unsupervised rule searches for multi-modality in the projection space. Exploratory projection pursuit theory [14, 15] tells us that search for structure in input space can be approached by a search for deviation from normal distribution of the projected space. Furthermore, when input space is clustered, a search for deviation from normality can take the form of search for multi-modality, since when clustered data is projected in a direction that separates at least two clusters, it generates multi-modal projected distributions.

It has been recently shown that a variant of the Bienenstock, Cooper and Munro neuron (BCM) [16] performs exploratory projection pursuit using a projection index that measures multi-modality [17]. This neuron allows modeling and theoretical analysis of various visual deprivation experiments and is in agreement with the vast experimental results on visual cortical plasticity. A network implementation which can find several projections in parallel while retaining its computational efficiency, was found to be applicable for extracting features from very high dimensional vector spaces [18, 17]. This method is applied to feature extraction in the problems discussed in the next section.

5 Mammal acoustic signal discrimination

The types of signals explored in this study are the marine mammal sounds of porpoise and sperm whale which were recorded at a sampling rate of 25 kHz at various locations such as the Gulf of Maine, the Mediterranean and the Caribbean sea. We consider large data files where the signal consist intermittently of mammal sounds and background noise. Each of these files contains whale or porpoise sounds, but not both. Several data sets of length 32768 samples corresponding approximately to 1.3 seconds, were extracted from these large files. These data sets which contained mammal sounds mixed with background noise, were used for training and testing.

Full discussion of the results appears in [13]. In this paper we present results from different choices of basis functions using a discrimination measure. These were not the best results obtained on this data (about 5%) which used a (nonlinear) discrimination based on linear combinations of basis functions from a fixed (wavelet) basis [13]. What is of interest here, is that although the signal is acoustic, better results were obtained with a Coiflet mother wavelet and the large variability in results based on different basis-search approaches. These results should then be contrasted with the active sonar results presented in the next section.

6 Backscattered sonar data discrimination

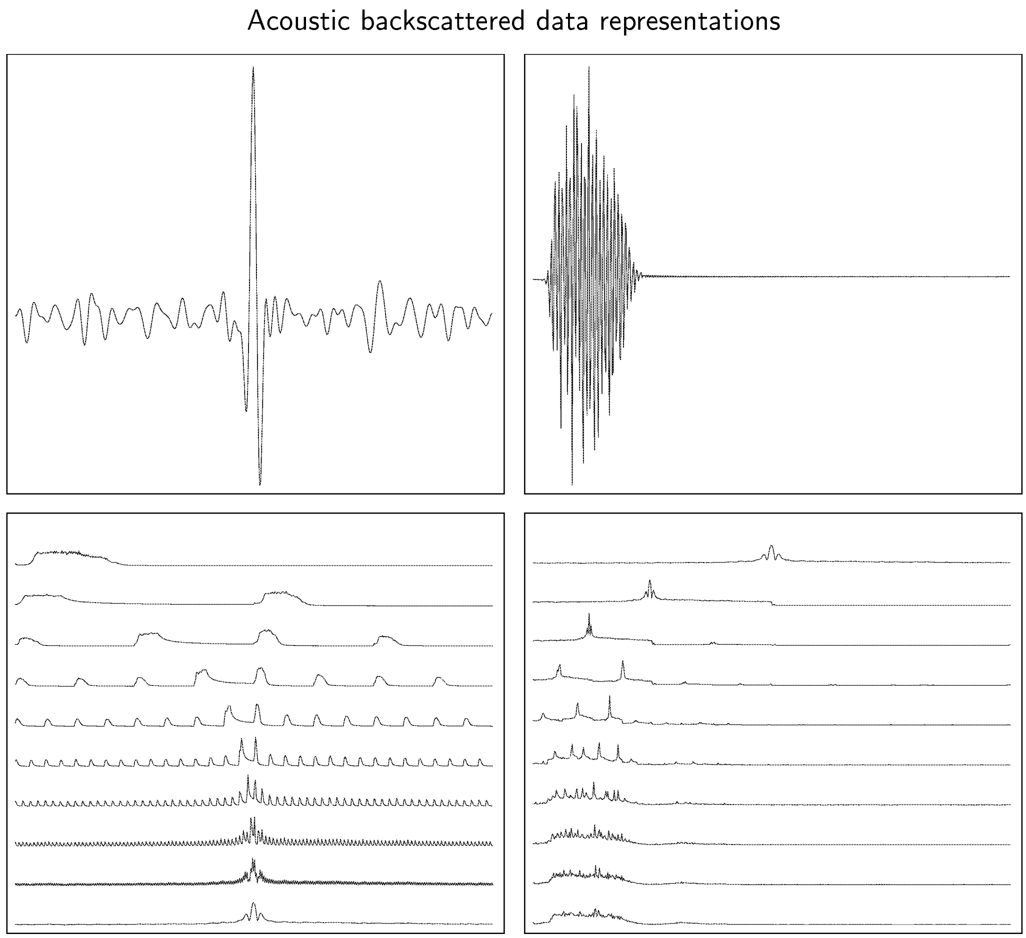

Figure 2: A typical signal from the data-set. Top left: Raw signal, right: Best Basis from Local Cosine. Bottom left: Local Cosine packet, right Wavelet Packet (Symmlet-6).

This application involved an active backscatter data set of mine-like objects. The data was collected at the Naval Surface Warfare Center (NSWC) by Gerald Dobeck. The task was to distinguish between man-made and non-man-made objects. There were six objects in the data: metal cylinder, cone-shaped plastic object, water-filled barrel, limestone rock, granite rock, and a water-logged wooden log. The data-set contained seven different frequency bands, however in this preliminary study, only one frequency band, an FM sweep between 20 to 60 kHz was used. The targets were suspended in a large water tank, while cylindrical objects were suspended horizontally. Measurements were collected in 5 degree increments on a rotating target around a vertical axis. Every second measurement was used for testing, thus the train and test data were interleaved and both included measurements at 10 degree increments. The signal/noise ratio was between 12-20db, where the

Backscattered data classification

	Local Cosine	Symmlet-6	Coiflet-3
BB	18.6(1.2)	33.1(1.4)	41.1(1.4)
LDB	18.7(0.8)	40.0(1.4)	39.9(1.6)
Wavelet	28.9(1.3)	42.7(1.8)	41.8(1.0)
Quadratic	18.7(1.5)	29.9(1.3)	30.3(0.8)

Table 2: Percent classification error rate of mine-like object discrimination using various discriminant basis representations from a Coiflet-3 and Symmlet-6 mother wavelets and a local cosine wavelet packet. The discrimination is performed from most discriminative 15 coefficients extracted by each method.

noise was an accurate modeling of the bottom reverberation. It appears that there is very little sensitivity to the noise level between these levels. Further details about the data as well as other classification results are given elsewhere in this volume [19]. Figure 2 depicts various signal representations of the backscattered data. Next to the raw signal at the top, a best-basis representation based on Coifman's algorithm [2] is shown. This basis was extracted from a local cosine packet that is presented at the bottom left panel. A wavelet packet representation using a Symmlet-6 mother wavelet is shown on the bottom right.

The results of the backscattered data are in Table 2. It appears that a local cosine representation is more appropriate for this acoustic signal and several discriminant basis selection yield similar results. These results reflect an improvement of around 3% which is due to BCM constraints and ensemble averaging over noise injected inputs [20, 21, for implementation details]. The results are generated assuming that false positive and false negative errors have the same weight. Surprisingly, linear discrimination from wavelet packet representation as suggested by [9] did not yield good results in any of the wavelet representations. This may be due to the small size of the training set which may lead to a very sparse representation in which a small number of training patterns cause a nonzero value for some of the wavelet basis and appear to have a high discrimination value.

It can be expected that combining information from several frequency bands will improve these results.

Acknowledgments

This work, including the collection of the data, was supported by the Office of Naval Research. Many discussions with various members of the Institute of Brain and Neural Systems were very helpful.

References

[1] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, pp. 179–188, 1936.

[2] R. R. Coifman and M. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. Info. Theory*, vol. 38, no. 2, pp. 713–719, 1992.

[3] I. Daubechies, "Time-frequency localization operator: a geometric phase space approach," *IEEE Transactions on Information Theory*, vol. 34, pp. 605–612, 1988.

[4] S. S. Chen, D. L. Donoho, and M. Saundres, "Atomic decomposition by basis pursuit," Technical Report, Stanford University, February 1996.

[5] P. E. Gill, W. Murrary, and M. H. Wright, *Numerical Linear Algebra and Optimization*. Addison Wesley, Redwood City, CA, 1991.

[6] S. Mallat and Z. Zhang, "Matching pursuit in a time-frequency dictionary," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3397–3415, 1993.

[7] N. Saito and R. R. Coifman, "Local discriminant bases," in *Proc. SPIE 2303* (A. F. Laine and M. A. Unser, eds.), pp. 2–14, 1994.

[8] N. Saito and R. R. Coifman, "Improved local discriminant bases using empirical probability density estimatin," in *Amer. Stat. Assoc. Proceeding on Statistical Computing (To appear)*, 1996.

[9] J. Buckheit and D. L. Donoho, "Improved linear discrimination using time-frequency dictionaries," Technical Report, Stanford University, 1995.

[10] N. Intrator, "Combining exploratory projection pursuit and projection pursuit regression with application to neural networks," *Neural Computation*, vol. 5, no. 3, pp. 443–455, 1993.

[11] S. Geman, E. Bienenstock, and R. Doursat, "Neural networks and the bias-variance dilemma," *Neural Computation*, vol. 4, pp. 1–58, 1992.

[12] N. Intrator, D. Reisfeld, and Y. Yeshurun, "Face recognition using a hybrid supervised/unsupervised neural network," *Pattern Recognition Letters*, vol. 17, pp. 67–76, 1996.

[13] Q. Huynh, L. N. Cooper, N. Intrator, and H. Shouval, "Classification of underwater mammals using feature extraction based on time-frequency analysis and bcm theory," November 1997. To appear IEEE-SP, Special Issue on NN Applications.

[14] P. J. Huber, "Projection pursuit. (with discussion)," *The Annals of Statistics*, vol. 13, pp. 435–475, 1985.

[15] J. H. Friedman, "Exploratory projection pursuit," *Journal of the American Statistical Association*, vol. 82, pp. 249–266, 1987.

[16] E. L. Bienenstock, L. N. Cooper, and P. W. Munro, "Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex," *Journal Neuroscience*, vol. 2, pp. 32–48, 1982.

[17] N. Intrator and L. N. Cooper, "Objective function formulation of the BCM theory of visual cortical plasticity: Statistical connections, stability conditions," *Neural Networks*, vol. 5, pp. 3–17, 1992.

[18] N. Intrator and J. I. Gold, "Three-dimensional object recognition of gray level images: The usefulness of distinguishing features," *Neural Computation*, vol. 5, pp. 61–74, 1993.

[19] L. L. Burton and H. Lai, "Active sonar target imaging and classification system," in *Proceedings SPIE*, 1997. To appear.

[20] N. Intrator, “Feature extraction using an unsupervised neural network,” *Neural Computation*, vol. 4, pp. 98–107, 1992.

[21] Y. Raviv and N. Intrator, “Bootstrapping with noise: An effective regularization technique,” *Connection Science, Special issue on Combining Estimators*, vol. 8, pp. 356–372, 1996.