Comparative study of shift-invariant symmetric wavelets and cosine local discriminant basis in noisy transients classification

Eric Delory(1) and John R. Potter(2)
Acoustic Research Laboratory
EE Dept, 27 Engineering Drive 3
National University of Singapore
(1)eric@arl.nus.edu.sg
(2)johnp@arl.nus.edu.sg

Abstract- Linear transforms in the time-frequency domain have proven very effective for non-stationary signal analysis. As such, wavelets and local trigonometric transforms have attracted substantial attention during the last decade and both methods have undergone extensive refinement. One recent development is the availability of algorithms for wavelet design and shift-invariant (SI) algorithms for both the wavelet and cosine transforms. Discriminant analysis of features extracted from a projection into a transformed space can be sensitive to shifts in the time-domain. The application of a shift-invariant transform is therefore attractive. Here, a comparative study is performed between two SI-wavelet transforms, the conventional transforms and the local cosine transform, on four classes of signals corrupted with underwater background noise. Training sets are reduced to 50 samples per class, and SNR is low in order to reflect real applications where samples are rare and noisy. Features are extracted from local discriminant bases to maximise the class separability. A Fischer discriminant analysis is performed on bases components to sort the coefficients according to their discriminant power. In our application, a maximum of 99% classification success rate is achieved and results show that performance differences between transforms can reach more than 20%.

I. INTRODUCTION

Time-frequency (TF) transforms have proven very useful in non-stationary signal analysis, in as various as numerous fields such as biomedical engineering, seismology, telecommunications, multimedia, satellite imaging, and underwater acoustics. Consequently, this last decade has been very dense in widening this family of signal processing tools and related research is living a real revolution. Increasing computing power coupled with the optimisation of newly introduced transforms such as the wavelet transform have generated robust and fast methods for signal decomposition and identification. Although a wide variety of TF-transforms are available and have been applied with relatively high efficiency in many fields, the concern is to find the appropriate TF-transform for a given problem.

In this paper, one of the many issues of signal processing is treated: the discriminant analysis of non-stationary signals with linear TF-transforms.

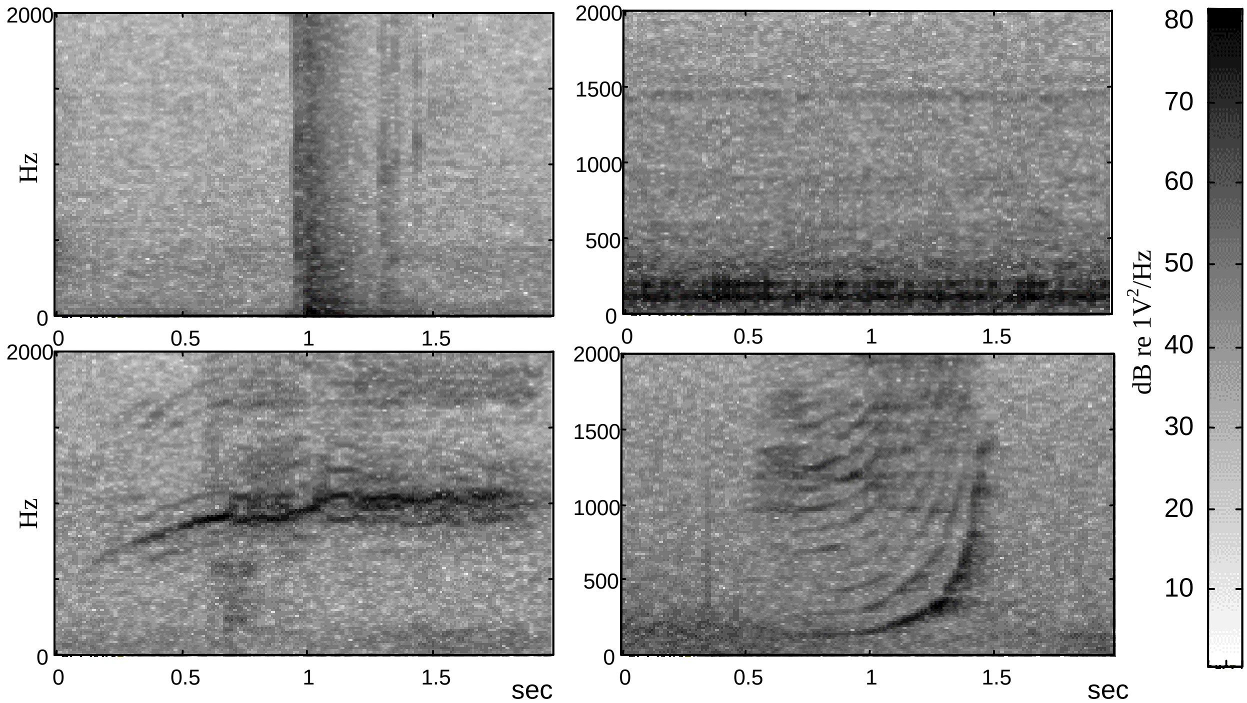

Fig.1: Spectrograms of four signal types, sampled at 4kHz: Top-left is a slamming door sound, top-right is a mechanical rumble sound, bottom-left is a whale vocalisation from a humpback whale (known as 'scream'), bottom-right another whale vocalisation. Classification was performed on very noisy versions of these signals (SNR~0dB).

A comparative study is presented, focused on two linear transform types, the wavelet packet transform (WPT) and the cosine packet transform (CPT). Shift-invariance is also implemented for the wavelet packet transform (SIWPT) whereas, although an SI algorithm exists for CPT, the SICPT algorithm [1] will only be mentioned for the reader's information but not applied, for some reasons explained later in the paper. Shift-invariance is often introduced as a criterion for classification robustness and one objective of this work is to compare the SI transforms with their non-SI relatives. WPT was chosen for this study for the availability of SI algorithms, recently introduced in [2]. One drawback of SI algorithms is a higher computational complexity, which must be taken into account when performances are compared. Two wavelet-scalet pairs of different lengths but from the same family are examined and compared with the cosine function. Here we present the results of a local discriminant basis (LDB) classification where four classes of noisy signals are pre-processed by the five transforms (2 wave-packet transforms, their SI versions, and the cosine-packet transform). For each LDB a Fischer linear discriminant analysis (LDA) is performed to order the transform vectors according to their discriminant power. Classification training and validation results are given for different coefficient numbers.

II. TRANSFORMS

A. Wavelet Packet Transform

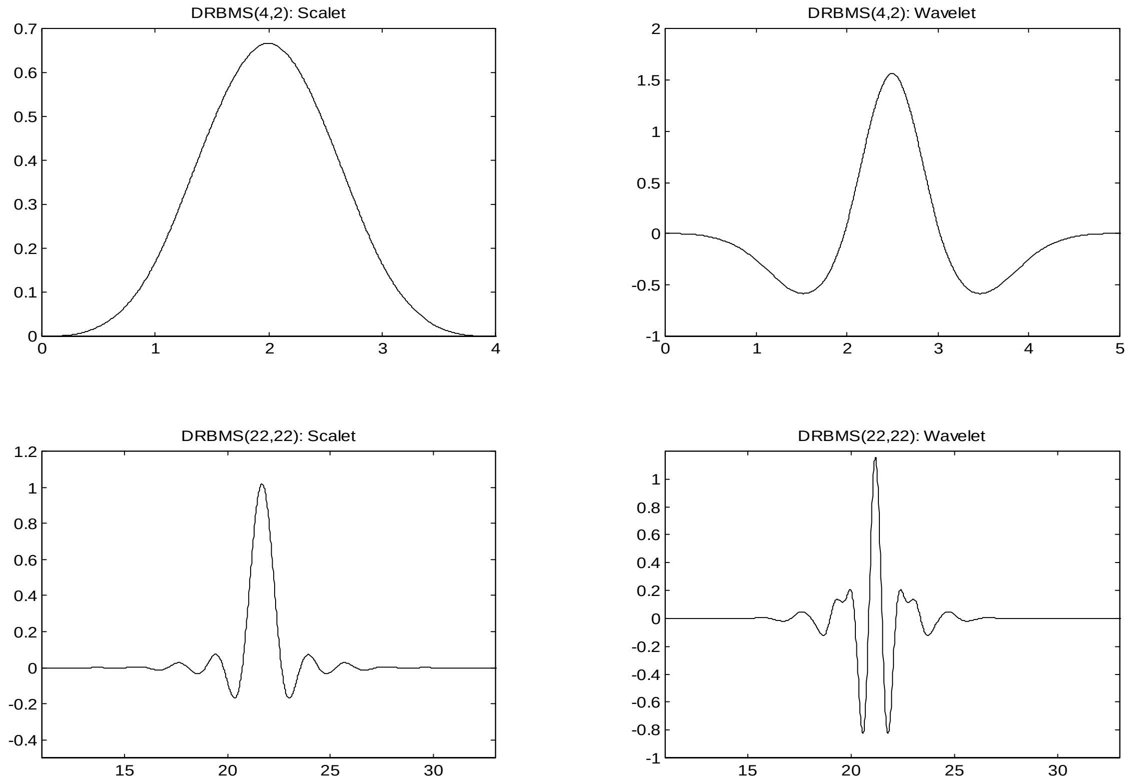

Fig. 2: Continuous scalet and wavelet from the (22,22)-and (4,2) parameters Daubechies real biorthogonal most selective family, resulting in respectively (43,45) and (5,7) filter bank lengths. All functions are symmetric, scalet and wavelet are respectively used for low and high-pass filtering.

WPT of a signal of length *N* performs a dyadic division of the frequency axis, using a fast filter bank algorithm that requires $O(2KN\log_2 N)$ operations, where K is the filter bank length. The resulting representation is an array of $(J+1)$ rows and *N* columns, where *J* is the number of dyadic decompositions. The filter bank low and high-pass filters used in this paper are the Daubechies real biorthogonal most selective (DRBMS) scalet and wavelet with respectively, 43 and 45 coefficients (long filters), and 5 and 7 coefficients (short filters). The coefficients were calculated with the WavBox from Carl Taswell [3]. Details about the wavelet choice can be found in [5] as the only purpose here is to compare filter lengths, not the types. Wavelet and scal-ets continuous functions are displayed in Figure 2.

B. Cosine Packet Transform

Also known as dyadic Local Cosine Transform, CPT of a signal of length *N* performs a dyadic division of the time axis, applying to each time bin a discrete cosine transform. Although the decomposition does not have to be dyadic, it has the advantage of creating a tree structure like in the wavelet packet case, allowing the implementation of fast best-basis search algorithms [6,7]. The decomposition complexity for *J* levels is $O(JN\log_2 N)$ which is comparable to the WPT algorithm complexity.

C. Shift-Invariant WPT

The fast dyadic WPT algorithm involves down-sampling by 2 at each decomposition level, leaving the “choice” between the even and odd sub-sampled transform coefficients at each level. Consequently, any even and odd shifted versions of the same signal will be transformed in a different set of coefficients. As signal classification relies on the positions and values of these coefficients, a shift-invariant algorithm is needed. A solution to keep WPT shift-invariant is to return to its original non sub-sampled transform. A far more elegant method is described in [6], which does not increase memory allocation and only raises the algorithm complexity by 2 in our case, as our decomposition is dyadic.

D. Shift-invariant CPT

The non shift-invariance of CPT is more obvious than in the WPT case, as any shift in the signal, which is not a multiple of the shortest window size, will locally result in a different spectral representation. A trivial method, which would consist in calculating all shifted transforms and choose the least costly one (for a given information cost), would be shift invariant, but the memory and computation time for such a transform would be far too high. A fast and low-memory SI-CPT algorithm was recently introduced, as part of a wider family of SI adaptive local trigonometric decompositions [1]. The algorithm complexity is $O(N(J+2^{logN-J+1})log_2N))$, which is generally higher than SIWPT but still allows for real-time decomposition. The SICPT algorithm results in selecting the time-shift, which produces the “cheapest” discrete packet transform, according to a given information cost function. The transform is therefore not only shift-invariant, but also provides a high level of energy compaction. However, in transient classification, signal centring in time is often performed before off-line processing in the time-frequency domain; SICPT shifts the energy centre by an arbitrary value, so that the resulting Fourier bases are not anchored in the time domain. As our classifier relies on such centring, it was decided not to use SICPT in the current study.

III. DISCRIMINANT ANALYSIS

The coefficient table resulting from a packet transform gives a redundant signal decomposition, where a basis can be selected out of the 2^J available. “Best-basis” algorithms [7] are fast and widely used in signal compression but do not enhance discrimination between signal classes. The Local Discriminant Bases (LDB) algorithm [8] performs a “most discriminant” basis search. First, a rapid algorithm, which uses a cross-entropy function, is applied to all classes packet tables. This step is important as it pre-processes the data for the extraction of better features. As the most discriminant basis is a set of N vectors, a Linear Discriminant Analysis (LDA) [9] orders them according to their discriminant power, for the extraction of the k most discriminant vectors. The other vectors are discarded.

In a typical LDA, the eigensystem must be solved and feature interpretation is difficult. Here, the basis search algorithm is $O(Nlog_2N)$ and features are well defined in the time-frequency plane.

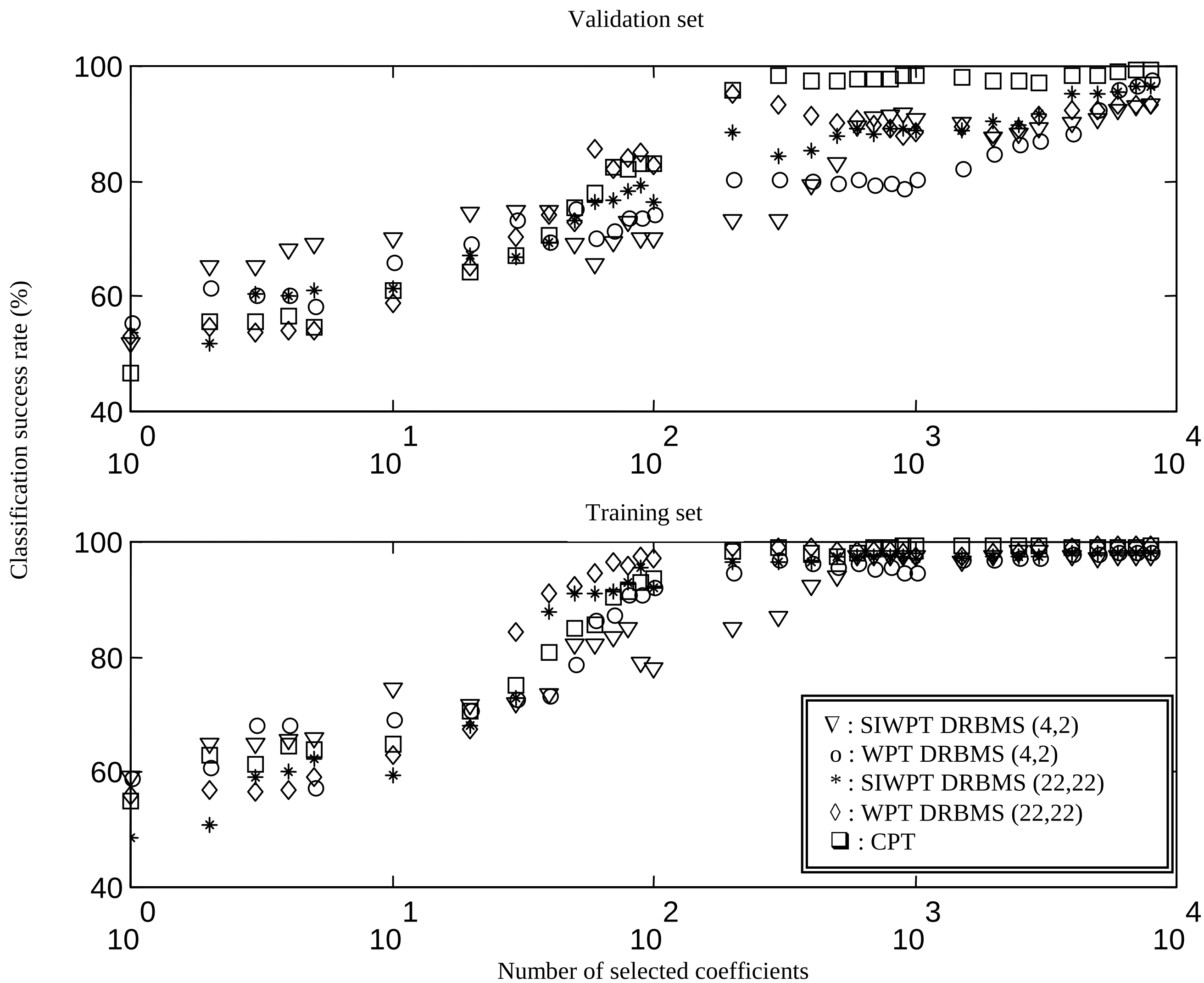

Fig. 3: Classification success rate as a function of the number of coefficient selected in the best discriminant basis, for all transforms. All classes (C1, C2, C3, C4) are considered.

IV. RESULTS

Four classes (C1, C2, C3, C4) of signals were used, all corrupted with underwater background noise (SNR ~ 0dB for all signals). Sample sets were reduced to only 100 signals per class, to reflect many practical studies where samples are lacking. Sample sets were split in two to create the training and validation sets.

The sounds used in the experiment were chosen to roughly represent both underwater mechanical and biological sources.
C1: “rumble sound” recorded from an air conditioning system
C2: “slamming door”
C3 and C4 represent recordings of two types of short tonal humpback whale vocalisations.

Figure 1 displays spectrograms of noise-free samples for each class. For each transform, the most discriminant basis was searched and discrimination power was compared for different numbers of features, from 1 to 8000. Classification results are plotted in figure 3.

The validation graph clearly shows that, in average, classification results are very reasonable, but vary substantially with the transform type. Obviously, good to excellent results are obtained with more than 200 coefficients, where CPT outperforms all other transforms without being really affected by the number of selected coefficients. 98% success rate is achieved in average.

WPT(22,22) reaches 95% success rate with 200 coefficients as well but is heavily affected by the coefficient number, with a 5% drop for 500 coefficients.

WPT(4,2) and SIWPT(4,2) (short wavelet) show a poor performance despite reaching 92% for 8000 coefficients. For very few coefficients (less than 20), however, both transforms outperform the others, with the best result obtained with the shift invariant version (78% with 20 coefficients).

In general, shift-invariance did not achieve the expected improvement for both wavelet-scalet pairs, often producing lower performance. However, in this transient analysis, all signals were centred in time and the need for shift-invariance was evidently not as high as for applications where no energy centring is either possible or useful. This result opens the issue of the necessity of SI algorithms for TF-transforms in transient classification, especially when CPU time is costly.

Regarding the classifier, this study shows that local discriminant bases coupled with a Fischer LDA converges towards excellent classification success rates in adverse conditions, i.e. when only a few noisy samples are available. The best and most stable results were obtained when signals were pre-processed by the dyadic local cosine transform (CPT), with already 98,4% success for only 300 coefficients.

ACKNOWLEDGEMENT

This work was funded by Defense Science Organisation, under the grant GR6432.

Eric Delory obtained his MSc in Electrical Engineering from the ENSEA and the MSc in Biomedical Engineering from Paris XII University (France). In 1994 he joined as a research scholar the Institute of Biomechanics of Valencia (Spain). In 1997 he comes to the Acoustic Research Laboratory (NUS, Singapore) to work on the application of time-frequency transforms and pattern recognition methods to marine mammal vocalisations.

Dr. John Potter founded and currently heads the Acoustic Research Laboratory at NUS. Originally trained as an Astrophysicist, he obtained his PhD in Cambridge in Glaciology and Oceanography, on work performed in the Antarctic for the British Antarctic Survey. His migration into Underwater Acoustics came in 1986, when he joined a NATO institute working on ASW. He comes to NUS from Scripps Institution of Oceanography, where he and Prof.

Michael Buckingham developed ADONIS, the first Acoustic Daylight system, producing real-time moving images using only ambient noise.

REFERENCES

[1] I.Cohen, S. Raz and D. Malah, “Orthonormal shift-invariant adaptive local trigonometric decomposition,” *Signal Processing,*vol. 57, pp. 43-64, 1997

[2] J.Weiss , “Translation invariance and the wavelet transform,” *Tech. Rep. Aware, Inc*

[3] C.Taswell, “The Systematized Collection of Daubechies Wavelets”. Technical Report CT-1998-06, Computational Toolsmiths

[4] S.Mallat , *A Wavelet Tour of Signal Processing.* Academic Press, 1998

[5] J.R. Potter and E. Delory, “A wavelet-packet transform as a preprocessor to an SOM-LVQ neural network for humpback whale song classification” *submitted to JASA*

[6] R.R.Coifman and Y. Meyer, “Remarques sur l'analyse de Fourier a fenetres” *C.R. Acad, Sci.,* pp 259-261, 1991.

[7] R.R. Coifman and M.W. Wickerhauser , “Entropy based algorithms for best basis selection,” *IEEE Trans. on Info. Theory*, vol. 38(2), pp. 713-718, Mar, 1992

[8] N. Saito and R.R. Coifman, “Local discriminant bases,” *Mathematical Imaging: Wavelet Appl. in Sig. and Image Proc. II, Proc. SPIE,* vol.2303, 1994

[9] R.A. Fischer, “The use of multiple measurements in taxonomic problems,” *Ann. Eugenics,* vol. 7, pp. 179-188, 1936