METHODS TO CLASSIFY OR GROUP LARGE SETS OF SIMILAR UNDERWATER SIGNALS

L. J. Hamilton
Defence Science & Technology Organisation (DSTO), 13 Garden St, Eveleigh, NSW 2015
Les.hamilton@dsto.defence.gov.au

Three methods of classifying large sets of acoustic signals are briefly discussed. The purpose of the discussion is to broadcast the existence and summary details of the methods to a wider audience. Large implies that hundreds of signals to several tens of thousands of signals may be detected. The signals of interest are broadly Gaussian, Rayleigh, or sinusoidal in shape, and of finite duration, such as seabed echoes and beaked whale chirps. The classification methods are (1) feature analysis, (2) direct statistical clustering of signals treated as single-valued curves, and (3) matched filtering with use of normalisations and kurtosis of the cross-correlation function output of the matched filter. Method (1) has been used for many years in several fields of science. It is suitable for many applications, but classifies on proxies, not the actual signals, which may lead to loss or distortion of information. Method (2) is a post processing operation suitable for signals with well defined signal to noise ratio which can be well aligned in time. Although simple in concept, it is a recent innovation, as it was apparently not previously realised that it could be done. A suitable clustering algorithm can classify signals into groups or sets where each set has a different average shape from the other sets, and can also classify signals forming quasi-continuums, such as those which can be viewed as morphing from one shape to another. Method (3) is suitable for detection and classification of stereotypical signals (those with strongly repeating waveform or signal shape), including weak signals in noisy backgrounds. In the usual application of matched filtering, classification is made solely on the un-normalised amplitude of the cross-correlation function. A novel extension of method (3) is to provide a confidence estimate for the classification through the kurtosis of the normalised autocorrelation function. The kurtosis is observed to be related to the degree of signal distortion or malformation relative to the template signal. When incoming signals of the same type vary in energy and degree of distortion or malformation, this scheme greatly outperforms standard matched filtering.

INTRODUCTION

Underwater acoustics may detect and classify inputs received by hydrophones over time intervals which can yield high numbers of received signals. Detection and processing may be carried out in real-time or executed as a post-processing activity. For example, real-time monitoring of offshore sites may be carried out to determine the presence of marine mammals, and to find to which species they belong.

The first step in signal processing is signal detection, often in the presence of noise or unwanted signals. The signals of interest to the present paper are broadly Gaussian, Rayleigh, or sinusoidal in shape and of finite duration (for examples see Figures 1 and 2). A second step is signal pre-processing or conditioning, where acoustic propagation effects and removal of artefacts caused by the measuring system are allowed for. The present paper does not specifically deal with these topics. It assumes signal pre-conditioning, and is instead concerned with the problem of classifying preprocessed received signals into groups with similar properties, particularly shape. When received signal types are very different from each other in one or more properties this is not necessarily a difficult problem. Problems can arise if, for example, signal shapes morph smoothly from one shape to another. This situation arises in the classification of seabed echoes stimulated by echosounders, or other active sonar types, when seabed properties along a transect change smoothly from one type to another, and consequently so do the echo shapes. Classification methods discussed are (1) feature analysis, (2) unsupervised statistical clustering applied directly to single-valued curves, and (3) matched filtering with normalisations and use of kurtosis as a classification parameter. These are versatile and relatively routine signal processing methods, suitable for a wide range of detection and classification applications. However, the direct clustering method for classification of signals is a recent innovation which is not widely known. Nor perhaps is an appreciation of the need for normalisation in matched filtering, or the fact that there is more useful information from the output of matched filtering than an amplitude.

THREE METHODS FOR CLASSIFICATION OF UNDERWATER SIGNALS

Method 1 – Classification through feature detection (reduction of a signal to a set of proxy parameters) and statistical clustering

When detection of any signal at all is important then the exceedance of a threshold value for a single parameter, such as signal to noise ratio, may be sufficient to decide a signal has been received. In the more general case of classification, rather than detection, two or more parameters may be required to decide that a particular type of signal has been received. Feature detection classifies on time and/or frequency domain properties of signals such as peak height, peak position relative to the start of the signal, duration, kurtosis, skewness, Fourier and wavelet transform coefficients, fractal dimension, and

time domain or spectral moments. In cases when it is not known which proxies are optimal for classification, such as when these may vary in time or space, hundreds of proxies may be formed [1]. Formation of proxies may require curve fitting, application of assumed models of signal shape or signal formation, and specification of user criteria such as what height, width, and separation define peaks. Multimodal signals can cause complications for feature detection. In post processing the features or proxies (m in number) may be subject to Principal Components Analysis (PCA) to reduce them to n independent or orthogonal principal components ($n < m$). The principal components are combinations of the features which best describe a particular data set in terms of the variance. The actual number of components used for classification is typically chosen to account for 95% or more of the variance of the data set [1]. However, three components are often used, and the remainder discarded, so as to enable visualisation in a pseudo three-dimensional space as point data. The relation of the principal components to echo properties or to physical processes may not be known, but a visual assessment of the distribution of the proxy points is used to discover any trends or distribution patterns which can be used to segment the point distribution into classes. Automatic segmentation of the three dimensional point data can be made by statistical clustering, simulated annealing, or simple segmentation into voxels (a voxel is a volume element in three dimensional space, analogous to a pixel or picture element in two dimensions). Examples of feature detection followed by PCA and simulated annealing or clustering may be found in [1] where these techniques are applied to seabed echoes based on proxies calculated in time and frequency domains.

Statistical Clustering

Statistical clustering views the n proxy components (or the n parameters) as coordinates in an n-dimensional space. Each set of coordinates identifies an n-dimensional point in the space. The function of clustering is to automatically find if points form discrete groups (clusters) which can be used as classes. This process begins by finding which points are close together and which are far apart. In a general framework, distances between any two points are calculated by measures such as the Minkowski distance:

$$\left(\sum_{i=1}^{n} | x_i - y_i |^p \right)^{1/p} \qquad (1)$$

where $p >= 1$, and x_i and y_i are vectors with the same number of elements (n in the present notation). The extended or n-dimensional Euclidian metric is given for $p = 2$, and the n-dimensional Manhattan metric for $p = 1$. Other metrics, such as entropy, may be suitable for some data sets. Points are assigned to trial groups or clusters, and in an iteration process are then moved to other clusters if this improves a global cost function. The statistical clustering algorithm employed in the present work is the CLARA (Clustering LARge Applications) algorithm of [2], described in the next section.

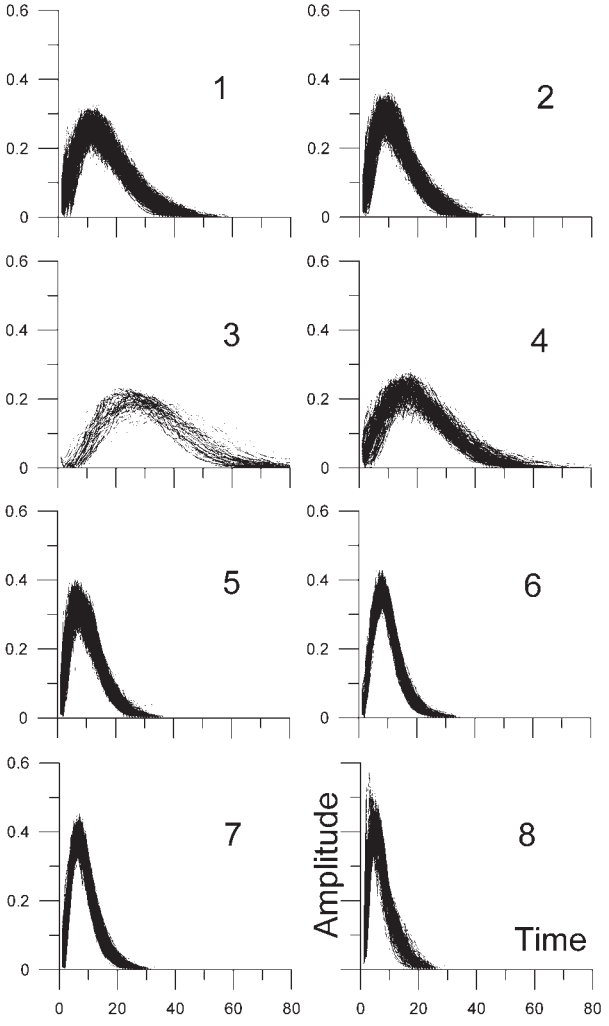

Figure 1. Statistical clustering of seabed echoes (clusters of single-valued curves) from Sydney Harbour into eight groups. Each group or cluster has a different basic shape from other groups, and all groups have relatively narrow dispersion or spread about a well defined central tendency (see Figure 2 for central tendency of the clusters).

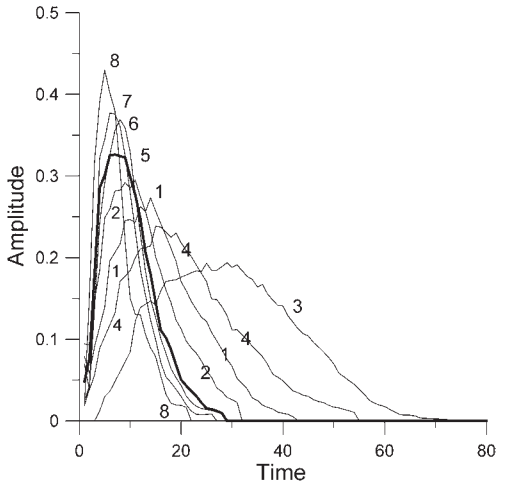

Figure 2. Medoids (central tendencies) of the eight clusters of Figure 1.

Method 2 – Classification through direct statistical clustering of signals treated as geometrical objects

The pseudo-Gaussian (Figures 1, 2) and sinusoidal signals of interest to the present paper can be described as single-valued curves. These have only one ordinate value for each abscissa value, whereas a multi-valued curve such as a circle has two for all but two points of its span. The classification of single-valued curves by description or reduction to features (Method 1) has been the orthodoxy for some decades in many scientific fields. Feature analysis has been used for example in geology to classify cumulative grain size curves, in underwater acoustics to classify seabed echoes, and in oceanography to classify wind-wave spectra. It has been demonstrated that a different approach is possible [3-5]. The signals in each of these cases form single-valued curves which can be regarded as discrete geometrical objects. Single-valued curves can be directly grouped or classed using suitable clustering algorithms. This approach is model free, and requires no curve fitting or selection and calculation of proxies. The actual curves are used, and data are not distorted or lost before analysis begins.

The CLARA algorithm of [2] has been demonstrated to be suitable for clustering of curves [3-5], although it was not designed for this activity. For a description of the CLARA algorithm see [2] and remarks in [5]. The CLARA algorithm is intended to cluster a minimum of 100 objects. A sister algorithm called PAM (Partitioning Around Medoids) can be used for less than 100 objects [2]. When clustering curves, the objective is that each cluster contains curves of similar shape, and that each group has a different basic curve shape than other groups. A distance metric, e.g. the multi-dimensional Euclidian or Manhattan distances referred to earlier, is used to decide whether curves are similar or dissimilar in properties. For CLARA, each curve in the entire data set is assigned to one (and only one) of the groups. Alternatively, fuzzy clustering algorithms assign probabilities of membership of an object to all clusters. The individual curve most closely approximating the central tendency of a cluster is termed a medoid [2]. A combination of non-standardisation of parameters and Manhattan distance metric was found to produce best results for curves. Alternatively, standardisation allows identification of outlier curves and of sets of curves most different from others [3]. A feature of CLARA demonstrated by [3-5] is that it can successfully partition a set of similarly shaped curves forming a quasi-continuum in their space. As an example, seabed echoes may form a quasi-continuum when seabed properties change gradually from one place to another, rather than jumping from one type to a completely different type. Another useful characteristic of CLARA is that clusters appear independent of data numbers. Clusters holding very small numbers of curves can be formed in classification of large data sets if they are geometrically different from other curves, providing a sufficient number of clusters is requested [3,4]. Another advantage of CLARA is that results do not depend on the order that objects are input to it, unlike K-means algorithms [2].

Estimation of the number of classes present in a data set is discussed in [3,5]. Companion algorithms to the clustering provide a quasi-independent estimate of the number of clusters. For CLARA this estimate is termed the Silhouette Coefficient. However, [3] recommends that users form from 2 clusters upwards (for example, 2, 3, 5, 8, 10, 20, …) until no more useful results are obtained for the particular data set being examined. Data exploration is an essential part of any examination of large data sets, and it is usually better to request many clusters, rather than a few. If discrete clusters do exist in a data set then CLARA and the Silhouette Coefficient will find them, but further information may be revealed by forming more clusters than recommended [3,5]. The effectiveness of the clustering can be checked by (1) examination of overplots of cluster medoids (the central tendencies of the clusters), and (2) examination of overplots of the curves forming each cluster for uniformity of properties (shape, location, central tendency, and spread). Figures 1 and 2 provide examples of medoids and their parent clusters.

Many clustering algorithms require too much processing power, computer memory, or processing time to be tenable for analysis of large data sets. Software program CLARA overcomes data size and processing time limitations by coupling statistical sampling and clustering techniques. The algorithm first clusters several sets of randomly chosen subsamples (for example, 5 sets with 200 objects in each for a data set with a total of 900 objects), then uses the particular subsampling returning best results to cluster the entire data set. This provides a fast algorithm suitable for processing of large data sets, at the possible expense of accuracy. However, the scheme has proved robust. The present author has used CLARA to cluster about 45,000 objects, and divide and conquer schemes can be used to increase this figure. It is also noted [3] that CLARA can be used to very quickly examine large data sets when the number of randomly chosen samples placed into the data sets is initially made very small. Running CLARA in this fashion provides a quick-look facility for data examinations, which can also be used to quickly estimate the number of clusters necessary to discover the structure in a data set.

Example of clustering of curves

An example is provided by classification of seabed echoes received by an echosounder [5]. Echoes are first compensated for acoustic propagation losses, corrected for artefacts caused by sampling effects, and are then normalised to unit energy. Normalisation removes relative amplitude information between signals, but this can be retained if necessary. The artefacts are caused by sampling the echo at fixed times instead of times corresponding to a set of particular incident angles on the seabed [6]. More samples are received for some particular angular range as depth increases, even if seabed type remains the same. Echoes received at different depths from the same seabed type are dilated or compressed relative to some reference depth, and this effect must be removed for comparisons of echo shapes to be meaningful (see [1] for more details). Figure 1 shows clusters and Figure 2 shows cluster medoids for a data set of seabed echoes from Sydney Harbour. The medoids generally morph from high amplitude, short duration signals to low amplitude, long duration signals. The longer duration signals are received from rougher or softer surfaces (muds) than the shorter duration, higher peaked echoes (which are from sands).

Discrete clusters do not exist for this data set, but the CLARA algorithm is able to sensibly partition the curves. For data sets such as this the clusters can be mapped to geographical space to provide a segmentation of seabed properties according to the seabed acoustic response stimulated by the echosounder. If this provides spatially coherent patterns then in principle only a few seabed samples need be taken to label or classify the seabed types indicated by the acoustics. A similar procedure can be carried out for multibeam sonar backscatter response curves [7], as an alternative to feature extraction and image processing methods. In both cases, because the actual backscatter response curve is used, geographical mappings of seabed areas with similar and dissimilar responses indicated by the clustering are standalone mappings which in themselves do not require validation. The same cannot be said when proxies are used. Some seabed samples or video must be taken if labels describing the physical or biological properties of the seabed classes are required.

Underwater signals are also used for purposes other than detection and classification of targets (including the seabed). For example, see [8] for statistical clustering of profiles of water current speeds with depth obtained from an Acoustic Doppler Current Profiler in the Changjiang estuary, Shanghai.

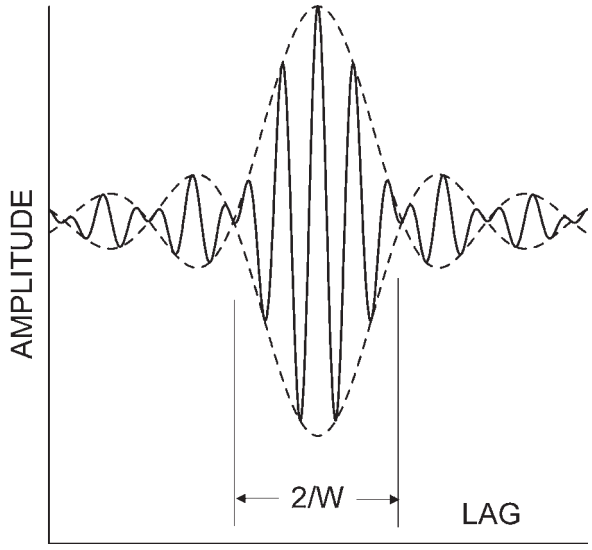

Figure 3. The autocorrelation function for a chirp signal, approximated as a sinc function. W=bandwidth (Hz) of the chirp.

Method 3 – Matched Filtering

Matched filtering cross-correlates one or more signal templates with time series to detect particular signals with known waveforms. This yields a sinc type cross-correlation function (Figure 3). The template is used as a sliding time window, and the exceeding of some threshold in peak amplitude of the resulting cross-correlation functions is used to indicate a detection. This operation can be performed in near real-time or in post processing. A major advantage of matched filtering is that it is very efficient at detecting signals buried in noise (e.g. [9]). It can also be used to detect signals overlapping in time.

If signals originally identical in waveform and source strength are received from very different distances, their received amplitudes and energies may vary greatly. If the propagation distances and paths are unknown, then compensation for spreading, scattering, and absorption losses can not be made, and genuine signals may be erroneously rejected by matched filtering. To overcome this [10] normalised the template signals and detected signals to unit energy. After normalisation of the detected signal, the cross-correlations were recalculated. It was observed that incoming signals, including spikes, which were partially correlated with a template could have a cross-correlation function of high peak amplitude, giving an erroneous classification. This spurious effect was reduced by searching for spikes and very short signals, and also by normalising the cross-correlation peak amplitude and kurtosis by the corresponding template parameters of the autocorrelation function.

Two examples of amplitude normalised autocorrelation functions for marine mammal chirp templates are shown in Figure 4. Their shapes are similar, but the peak of one function is noticeably broader than the other. The kurtosis parameter was found to be able to quantify this difference in shape very well, so much so that it could be used as the primary measure of the reliability of the detection. Kurtosis was observed to be related to the degree of distortion or malformation of the detected waveform compared to its template. Through the normalisations and use of kurtosis a low-amplitude well-formed signal scores higher than a distorted high amplitude signal. In this scheme the degree of distortion of the waveform is most important to the classification, not the signal to noise ratio of the detected signal.

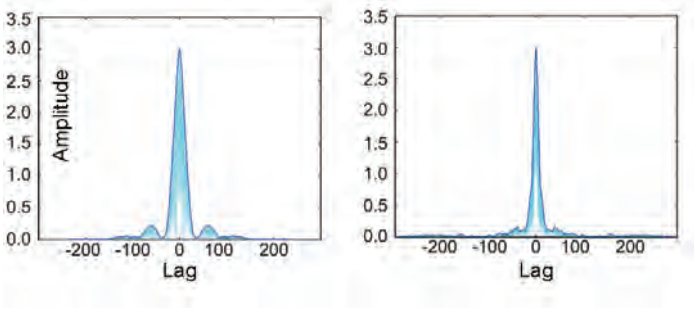

Figure 4. Two examples of amplitude normalised autocorrelation functions for marine mammal vocalisations. Note the different peak widths, which can be differentiated by the kurtosis parameter.

Example of matched filtering including allowance for ambiguity

The matched filtering scheme with normalisations and use of kurtosis as the primary detection parameter was applied by [10] to classification of the chirp vocalisations of beaked whales (Ziphiidae), a family of toothed whale or odontocete. Chirps are oscillating signals for which the frequency increases or decreases with time (Figures 5 and 6).

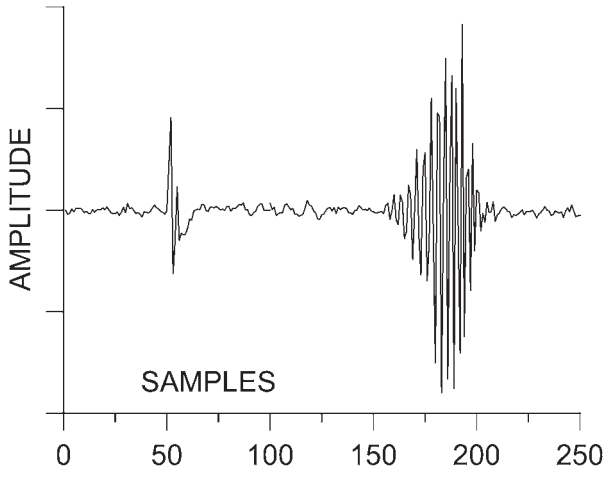

Figure 5. Examples of two stereotypical whale vocalisations from the digital library of the Marine Biological Library, Woods Hole Oceanographic Institute (www.mblwhoilibrary.org). The right hand waveform is a probable Blainville's beaked whale (*Mesoplodon densirostris*) chirp. Note the relative differences in duration of the two waveforms.
(Filename Set3_A1_042705_CH11_H11_A0300_0330.WAV, Time 11:14.0500 to 11:14.0534).

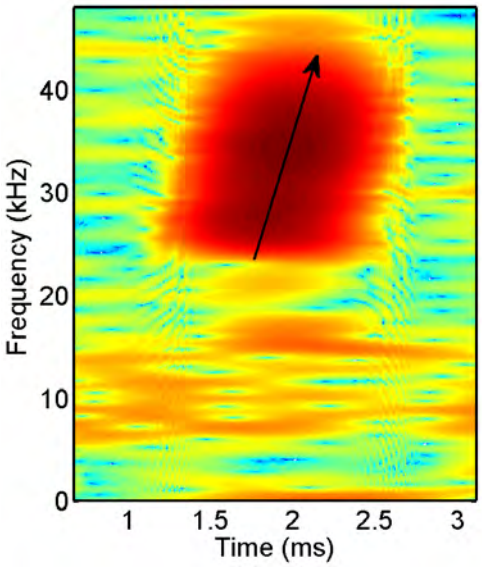

Figure 6. Spectrogram of a chirp with frequency upsweep. For the waveform (a Blainville's beaked whale chirp) see Figure 5.

A further processing step was necessary to obtain reliable detections, as a property of chirp signals can result in detection ambiguity, even though chirp waveforms may have very obvious time domain differences. The autocorrelation function of a chirp is approximately a cosine modulated by a sinc function (Figure 3). The width between the primary nulls of the sinc function is 2/W, where W is the chirp bandwidth (Hz). The frequency of the cosine is the median frequency of the chirp [9,11]. Chirps with similar bandwidth will produce sinc envelopes with similar widths between the primary nulls, regardless of chirp duration or centre frequency. Chirps with similar median frequencies and similar bandwidth have similar autocorrelation functions, including phase, regardless of marked differences in durations or number of cycles of waveform. This matched filtering ambiguity must be resolved by other information on the signal. Simple time or frequency domain rules suffice for some cases. For example, beaked whale chirps have much longer signal durations than the click sounds and chirps of other odontocetes such as the false killer whale (*Pseudorca crassidens*) vocalisations which caused ambiguity. Simple duration criteria were used to separate possible Ziphiids and non-Ziphiids before the final classification step.

The scheme was largely self-verifying in that the number of detections of a particular beaked whale species for some particular confidence level over some particular length of time plateaued as the initial matched filter amplitude criterion specifying a detection was decreased. In contrast, the number of detections for standard matched filtering methods, when used without additional rules, kept increasing without apparent limit (Figure 7). The normalised scheme also produced very few false detections for higher confidence values.

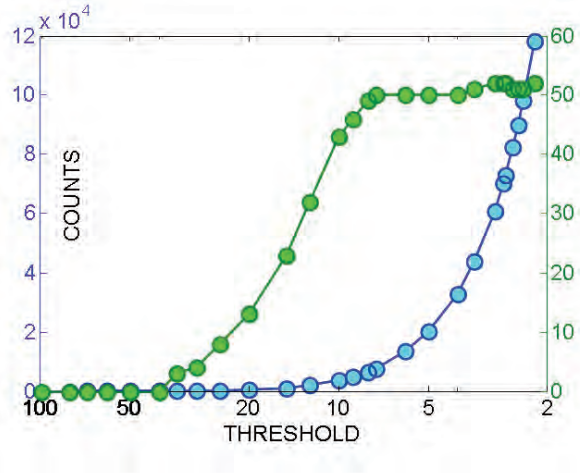

Figure 7. Number of detections as a function of detection threshold for the normalised two parameter matched filtering scheme (green circles) compared to standard matched filtering (blue circles). Note the different ordinate scales. The normalised scheme plateaus at 50 detections as the standard scheme rises to 120,000 detections.

DISCUSSION

Three methods have been briefly discussed for classification of underwater signals. These are (1) feature analysis, (2) direct statistical clustering of signals treated as single-valued curves, and (3) matched filtering with normalisations. Method (1) is simple and is suitable for many applications, but classifies on proxies, not the actual signal, and may lose or distort information. Method (2) is a post processing operation suitable for signals with well defined signal to noise ratio which can be well aligned in time. A suitable clustering algorithm can classify sets of signals where each set has a different average shape from the other sets, and can also classify signals forming quasi-continuums, such as signals which can be viewed as morphing from one shape to another. Method (3) is suitable for stereotypical signals with strongly repeating shape, and is very

good at detecting such signals in noisy backgrounds.

In order to apply clustering directly to signals of different durations the signals must have well defined start points or be able to be well aligned in time. For signals such as seabed echoes alignment should not be made by echo peak, because the peak position is determined largely by the interaction of the acoustic wavelength and seabed roughness, and does not occur at some fixed interval after initial contact of the output echosounder pulse with the seabed (see Figure 2). For seabed echoes it is usually possible to automatically determine robust start points by simple amplitude criteria [5]. However, the criteria are a function of echosounder make and model, since they depend on the shape of the output pulse.

Cross-correlation is generally sufficient to align and detect highly stereotyped signals. It is not necessary to be able to find well defined start points for them, and weak and noisy signals may be detected and aligned. However, identical signals of the same source strength received from different and unknown ranges will produce different cross-correlation amplitudes. Unless normalisation of some kind is used this negates the use of matched filtering for signal classification. Incoming signals of unwanted type which are partially correlated with the desired signal in the time domain may also cause erroneous classification. Matched filters are very efficient at detection of signals, including signals buried in noise, but not necessarily at discrimination. If matched filter banks are to be effective, then users must check and allow for ambiguity between templates, and must also be aware of the other factors which may lead to ambiguity in classification.

REFERENCES

[1] J.M. Preston, "Acoustic classification of seaweed and sediment with depth-compensated vertical echoes", *Proceedings of OCEANS 2006*, Boston, USA, 18-21 September 2006

[2] L. Kaufman and P.J. Rousseeuw, *Finding groups in data: An introduction to cluster analysis*, John Wiley, New York, 1990

[3] L.J. Hamilton, "Clustering of cumulative grain size distribution curves for shallow-marine samples with software program CLARA", *Australian Journal of Earth Sciences* **54**, 503-519 (2007)

[4] L.J. Hamilton, "Characterising spectral sea wave conditions with statistical clustering of actual spectra", *Applied Ocean Research* **32**(3), 332-342 (2010)

[5] L.J. Hamilton, "Acoustic seabed classification for echosounders through direct statistical clustering of seabed echoes", *Continental Shelf Research* **31**, 2000-2011 (2011)

[6] D.A. Caughey and R.L. Kirlin, "Blind deconvolution of echosounder envelopes", *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing Conference* (ICASSP'96), 6, 3149-3152 (1996)

[7] L.J. Hamilton and I. Parnum, "Seabed segmentation from unsupervised statistical clustering of entire multibeam sonar backscatter curves", *Continental Shelf Research* **31**(2), 138-148 (2011)

[8] Z. Cao, X.H. Wang, W. Guan, L.J. Hamilton, Q. Chen, D. Zhu, "Observations of nepheloid layers in the Yangtze estuary, China, through phase corrupted Acoustic Doppler Current Profiler speeds", *Marine Technology Society Journal* **46**(4), 60-70 (2012)

[9] A. Hein, *Processing of SAR data: fundamentals, signal processing, interferometry*, Springer-Verlag, Berlin, New York, 2004

[10] L.J. Hamilton and J. Cleary, "Automatic detection of beaked whale calls in long acoustic time series from the Coral Sea", *Proceedings of OCEANS 2010*, Sydney, Australia, 24-27 May 2010

[11] C. de Moustier, *Fourth Asia-Pacific coastal multibeam sonar training course*, Cairns, Australia, 14-19 August 2000