

# CONTENT BASED ACOUSTIC DATA RETRIEVAL SYSTEM FOR UNDERWATER TARGET RECORDS USING UNSUPERVISED NEURAL NETWORKS

Shameer K. Mohammed<sup>1</sup>, Suraj Kamal<sup>2</sup>, Supriya M.H.<sup>3</sup>

<sup>1</sup>Research scholar, Research and Development Center, Bharathiar University, Coimbatore, India.

<sup>2</sup>Research Scholar, Department of Electronics, Cochin University of Science and Technology. Kochi, India.

<sup>3</sup>HOD, Department of Electronics, Cochin University of Science and Technology. Kochi, India.

\*\*\*

**Abstract** - Underwater acoustic target recognition systems as well as human experts utilize large set of exemplar records as a means for inductive training that are usually obtained from archived observations. Due to the staggering volume of such archives it is often hard to manage or retrieve particular contents. Content based query systems are often employed in such scenarios in order to retrieve records based on similarity. However finding similarity is not a straight forward task especially with underwater acoustic records. Target feature extraction schemes based on expert domain knowledge in combination with discriminative pattern classifiers that can generalize on complex functions are utilized in order to retrieve similar contents. In the proposed system spectral features are used along with an unsupervised self-organizing feature map classifier to find similar contents. Kohonen Self Organizing Map, a topology preserving unsupervised neural network that projects high dimensional data into lower dimensional space where similar contents distributed adjacently. A global map of the archived contents is created and the mapped terrain is tagged with textual content descriptors. The query item is mapped to the co-ordinates of the global map and a Euclidean minimum distance measure is used to retrieve the best matching content. As the search is carried out in a lower dimensional feature map it is less susceptible to errors due to variations in data and more efficient in computational means.

**Key Words:** Underwater Target Classifier, Feature Vector, Spectral Features Kohonen Self Organizing Map, Unsupervised Neural Networks, Euclidian Minimum Distance, Spectral Roll off

## 1. INTRODUCTION

Automated target recognition and classification is a challenging task in underwater ambience due to the presence of several interfering targets as well as the complex channel behavior. Although the human perceptual ability to identify and categorize acoustic targets is astonishingly high even in cluttered observations, algorithmically engineered target recognition systems often suffer from the lack of ontological knowledge encoded within. Hence, target recognition systems generally need large amount of exemplar training data as a means of prior knowledge, in order to produce acceptable success rates. Such knowledge is generally obtained from the records of long term field recon-naissance as well as the several decades of aggregated data that are often archived in warehouses. However, as the list of known targets increases, it becomes a laborious task to maintain the archive as well as to retrieve particular targets among from the listed ones. Manual categorization of large acoustic target records is inefficient both in time and consistency. Content based query systems are often employed for acoustic surveillance and strategic data warehousing. In this paper, an unsupervised neural network based method is being presented as a means for content indexing and retrieval. The self-organizing map (SOM) is a well-known unsupervised neural network (Kohonen, et al., 2000) [1] which is able to capture the higher order latent structures in the data and visualize it as partitioned maps. The unsupervised technique is rather efficient than the supervised techniques as it can learn from the data directly without any explicit training especially in systems where database updates are frequent.

The semantic gap between the user level descriptions of an acoustic target record such as the perceptual qualities of the record, to the computed ones of the same in the data flow pipeline, needs to be bridged effectively in order to build an efficient indexing system. In the proposed system, an acoustic content descriptor database is being maintained as the search index for the actual target signal recording which in turn holds the extracted signature features that describe the collected signals in a compact and expressive manner. The gap between the low level

content descriptors and the high level target labels is bridged by the self-organizing map, which is efficient in representing the emergent ontological relations. A progressive learning system scheme is employed which keeps track of the content additions into the main database and automatically updates the descriptor table metadata. The SOM learns the specific features of all archived data in the background. When a new query is presented to the system, the extracted feature map is matched against the content map generated by the SOM network. The best matching target signal is then retrieved from the corresponding map coordinates.

## 2. IMPLEMENTATION

A top level architectural diagram of the proposed system is depicted in Figure 1. As most of the records are obtained in different recording setups and measurement ambiances, the low level signal parameters may vary drastically. The records are normalized in order to alleviate the ambiguities that may arise due to these anomalies. The archive manager process, also responsible for training the SOM network, continuously waits for update requests from the system manager task. The metadata usually contain where and when the record is carried out (typically geographical co-ordinates and time), what the class of target it belongs to and other important parameters describing the content.

The metadata pre-processor attributes these textual parameters with the associated content in the main database. This helps in retrieving contents via text based queries. The query parser discriminates between the textual and content based queries and invoke responsible functions accordingly. The archive manager task handles all requests to the main database where all records are stored, and is invoked through the system manager task. On certain special occasions such as the database holds an unmapped item, the archive manager directly invokes classifier in order to remap the global content map.

### 2.1 Feature extraction

The content retrieval system utilizes a classification model that best maps the query signal to the contents of the database in order to retrieve similar records if there exists any. However, the map-ping is often found to be complex due to many reasons such as; acoustically similar patterns can be produced by very different sources, extraneous dimensionality of the measurements, variations in SNR and measurement ambience etc.. Signal pre-processing and feature extraction schemes based on extensive knowledge on the characteristic features of the noise spectra that are unique among the individual classes are used in order to alleviate the curse of dimensionality as well as to yield better invariant representations in various situations. The process of feature extraction transforms the measurement space into a new lower dimensional subspace, where individual sources have disjoint manifolds by obviating the trivial or irrelevant

information. These feature representations corresponding to the observations are also stored in the database as low level content descriptors along with the original measurement and the metadata.

Several discriminative features are devised for sonar target recognition during the last few decades. Among them, spectral features are computationally less expensive and more descriptive. A feature set is formed based on standard low-level (SLL) spectral shape features [2] [3] such as spectral flux, spectral roll off, and have been extracted from the acoustic signals so as to generate the content descriptors.

#### 2.1.1 Spectral flux

Spectral Flux is a measure of the amount of spectral variation in the frame by taking the Euclidean distance between two consecutive frames of normalised spectra. It is computed as the squared sum of the difference between the corresponding values of the magnitude spectrum of the successive signal frames in a short time analysis window.

$$S_f(k) = \sum_{i=0}^{n-1} s(k, i) - s(k-1, i) \quad (1)$$

$S_f(k)$  is the spectral flux,  $s(k, i)$  is the value of the  $i$ th bin in the  $k$ th frame,  $s(k-1, i)$  is in the previous frame.

#### 2.1.2 Spectral Roll off

The spectral roll off is the point at which the  $n$ th percentile of the power spectrum where  $n$  is 85% to 95%. It is usually considered that, about 95% of the magnitude distribution is concentrated below the roll off point. This is also shows a measure of spectral skewness.

$$\sum_{k=1}^M |x_r[k]| = 0.85 \sum_{k=1}^{N/2} |x_r[k]| \quad (2)$$

#### 2.1.3 Spectral Centroid

The spectral centroid represents the centre of mass or the barycentre of the magnitudes of the spectrum computed as the weighted mean of the spectral components, where  $x(n)$  is the weighted frequency value of  $n$ th bin, and  $f(n)$  is the centre frequency.

$$S_c = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)} \quad (3)$$

## 3. SELF-ORGANIZING FEATURE MAPS

The self-organization ability or the plasticity of the biological nervous system towards variations in the perceived world is fascinated signal processing experts for long time. Based on a common Hebbian mechanism, the

network self-organizes cooperatively and simultaneously to produce respective fields (RFs) or orientation maps. The Self-Organizing Maps, proposed in [4] is a fully connected, single layer, linear, unsupervised as well as competitive artificial neural network model. The network can project higher dimensional data into an l-dimensional lattice of weights while preserving the topological relations in the original feature space. Such topological maps of features can be obtained by abstracting out the elements in common in the feature space i.e. the similarity in the input signals are preserved as much as possible in the representation space.

In this work, a Kohonen Self-Organizing Map, a special kind of unsupervised neural network, inspired by the retina-cortex mapping, has been used in order to generate the abstract topological map, where all the known records are mapped. Isolated sectors in the map can be thought as the index of the content database. The query item is mapped against the pre-learned map coordinates for finding the nearest neighbor with a minimum distance measure. The interesting feature of the SOM is that it can map similar data points in closer space in the topology map. The SOM network reorganizes itself into an elastic net of points that are fitted to the feature space during the training process. The map is again much lower in dimension than the original feature space obtained in the feature extraction phase. Hence, the content search is much faster as it is done on the lower dimensional invariant representation. The network can be thought of as an analogue to the tonotopy map found in the auditory regions of mammalian cortex, where local neurons react to identical tonals with progressive frequency sensitivity [5]. An abstract representation of the SOM network with a typical feature map is shown in Figure 3.

### 3. 1 LEARNING THE ARCHIVE MAP

The entire contents of the archived records are used for initial training of the network. While training, the features extracted from the training samples are exposed to the network and the network adjusts its weights to the output nodes in such a way that the structure of the input space is preserved in the output, neurons. The learning rule can be described as

$$w_i = \alpha(t)(x - w_i)U(y_i) \tag{4}$$

where  $w_i$  is the weight vector and  $x$  is the feature vector. The function  $U(y_i)$  is zero unless  $y_i > 0$  so those nodes having some activity are allowed to be represented.  $\alpha(t)$  is the learning rate. Every node can be addressed by its coordinates  $(u, v)$  on the two dimensional space. Similar to other competitive learners a winner is chosen for each input units according to the similarity between the input and the weights. For the input  $x$ , the winning node can be estimated by,

$$\|x - w_c\| = \min_i \{\|x - w_i\|\} \tag{5}$$

where  $c$  is the winning node. The weights of the winning node is updated together with the weights of its physical neighbours in a diminishing gradient along the radial distance. During the course of learning this radius is reduced until it reaches a single node. The learning rule can be rewritten as

$$w_i(t+1) = \begin{cases} w_i + \alpha(t)(x - w_i)U(y_i) & i \in N_c \\ 0 & \text{Otherwise} \end{cases} \tag{6}$$

Where  $N_c$  is the member nodes in the locality. The weights are adjusted until the convergence, where the value represents the physical coordinates of the node.

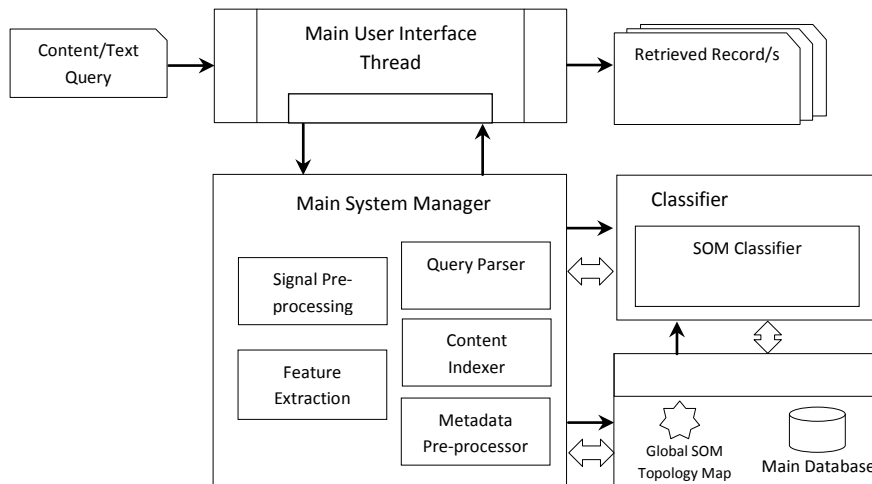
After learning the abstract feature map, the metadata is associated with the map coordinates in order to semantically describe the contents. While each content query is presented to the system it will retrieve the nearest match from the mapped coordinates. However, if the query belongs to an absolute new class, it will be added to the topology by retraining the network. The process is carried out in a separate background thread (i.e. archive manager directly invokes classifier) in order to improve the responsiveness during retraining. Thus the system will learn incrementally from presented queries.

### 3. 2 TAGGING AND METADATA EMBEDDING

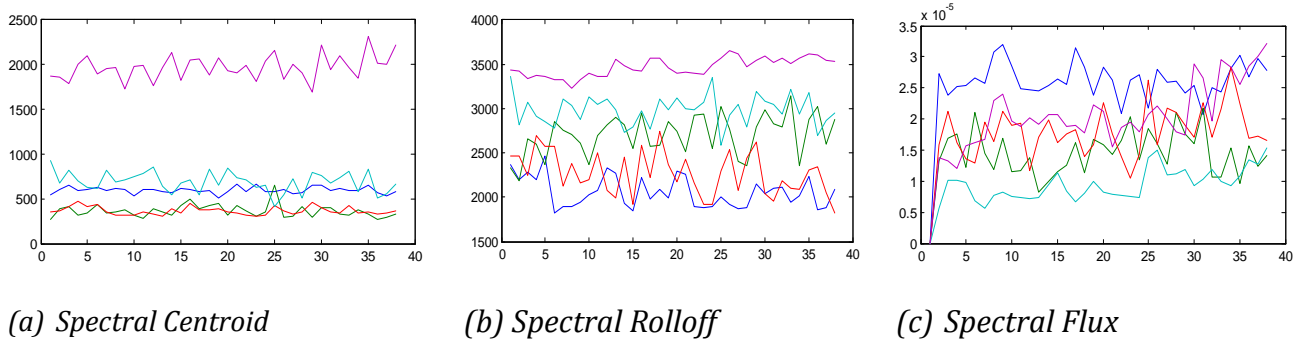
All the records in the database are manually tagged at loading time so as to facilitate textual queries. The metadata is embedded along with the raw observations and extracted features. On all content based queries, the user will be prompted to add it into the database and if the content is absolutely unknown, the user can annotate it with metadata. In order to annotate novel observations of a target that is already archived in the database; content aliasing can be carried out for making further queries easier. Typical contents of the main database are shown in Table 1.

### 4. CONCLUSIONS

Underwater acoustic target recognition systems utilize pattern classifiers for automating the process. However, the inductive learning classifiers require lot of training exemplars for achieving acceptable recognition rate and an extensive set of target records are usually maintained with the system. But, it is often difficult to maintain such large datasets manually which demands for sophisticated content management systems. In this paper an unsupervised neural network based content management/retrieval system has been proposed. Unsupervised networks are recently getting more attention due to their ability to find subtle features from



**Fig -1:** The top level architectural diagram of the system

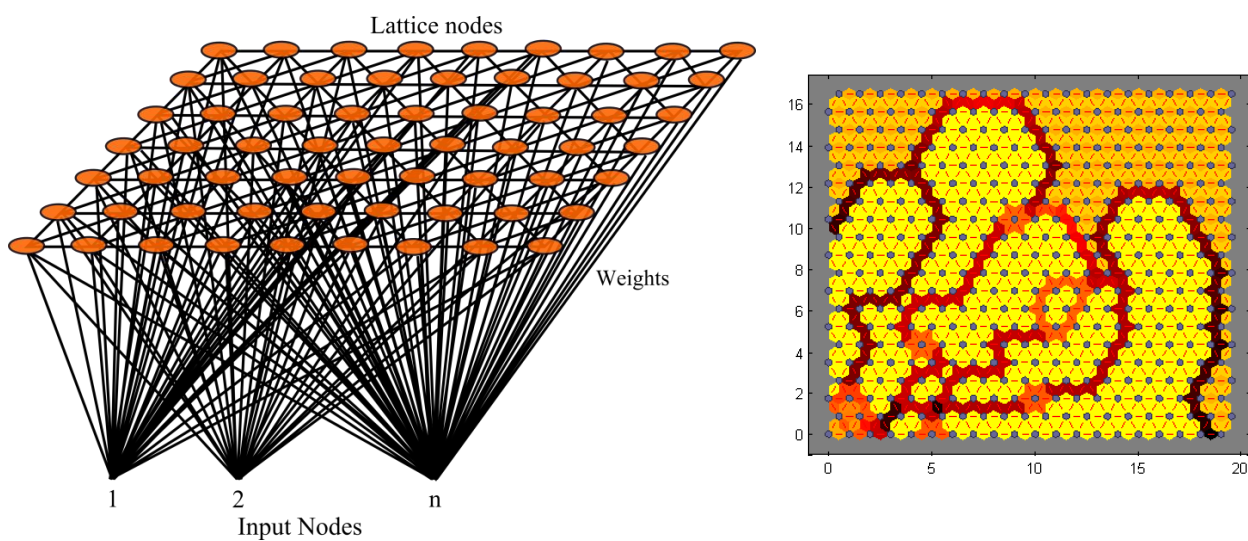


(a) Spectral Centroid

(b) Spectral Rolloff

(c) Spectral Flux

**Fig -2:** Spectral Features of different targets



**Fig -3:** The SOM network and a typical map of features



Table -1: Typical contents of the main database

Record No.	Record Type	Sampling Rate	Length (S)	Features		Captured on		Location
						Date	Time	
01	Ship Engine	44 Ks/S	10	SF	SC	20-04-2013	10:30	9°51'25.779"76°3' 17.6976"
02	Marine Mammal	44 Ks/S	10	SF	SC	20-04-2013	20:45	-

the signal. As one of the best topology pre-serving feature map network, the Kohonen SOM has been used in the system as the discriminative classifier. The SOM provide options for progressive learning, as they are able to detect unknown targets. The other best side of the SOM is its ability to visualize the records as a discriminative map, which helps in yielding better perceivable target space.

A complete acoustic record management system has been implemented with content and textual search facility and the system can be incorporated in oceanographic data warehousing facilities in order to alleviate manual processing hassles. The system yielded more than 80% success rate in test cases with acceptable computational efficiency.

**ACKNOWLEDGEMENTS**

The authors gratefully acknowledge Naval Research Board, New Delhi for the financial assistance and the Department of Electronics, Cochin University of Science and Technology for extending all the facilities for carrying out this work.

**REFERENCES**

- [1] T. Kohonen, H. Jussi and J. K. & J. Laaksonen., "SOM pak: The self-organizing map program package," Helsinki University of Technology, Laboratory of Computer and Information Science, 1996.
- [2] G. Peeters, "A large set of audio features for sound description (similarity and classification)," CUIDADO project, 2004.
- [3] O. Lartillot and P. Toivainen, "A matlab toolbox for musical feature extraction from audio," in In International Conference on Digital Audio Effects, pp. 237-244., 2007.
- [4] T. Kohonen, "Self-organized formation of topologically correct feature maps," Biological cybernetics , vol. 43.1, pp. 59-69, 1982 .
- [5] T. M. Talavage, M. I. Sereno, J. R. Melcher, P. J. Ledden, B. R. Rosen and A. M. Dale, "Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity," Journal of neurophysiology , Vols. 91, no. 3, pp. 1282-1296 , 2004.

**BIOGRAPHIES**



**Shameer K. Mohammed** is working as Project Scientist in the Department of Electronics, Cochin University of Science and Technology 2012 onwards. His areas of interest include, Underwater Acoustics, Artificial Intelligence, pattern recognition, Signal Processing, Automatic Target Recognition, Hidden Markov Model etc.



**Suraj Kamal**, Research fellow at Cochin University of Science and Technology since 2011, is currently Senior Research fellow of the Department of Electronics. His research interests are in Underwater Classifiers, Pattern recognition, Feature Extraction, Deep Learning, Probabilistic Neural Networks etc.



**Dr. Supriya M. H** joined as a faculty in the Department of Electronics, Cochin University of Science & Technology in 1999. Her fields of interests are Target Identification, Signal Processing, Bioinformatics, Steganography and Computer Technology. She has presented papers in several International Conferences in Europe, USA, and France. She is actively involved in Research and Development activities in Ocean Electronics and related areas and has a patent and about 105 research publications to her credit. She is a Life member of IETE and ISTE.