# Bottlenose Dolphin Whistle Categorization Using Eigenwhistle Based Approach

Michael MacBride, Lucas Roth

*School of Engineering, Edith Cowan University*
*270 Joondalup Dr, Joondalup WA 6027, Australia*
Michael.MacBride60@gmail.com

*Abstract*— **Dolphins produce a wide variety of whistles that vary their contour shapes in terms of frequency band and spatial length. Characterization of whistles is based on the premise that dolphins share a function-specific repertoire of whistles and each one is used in a particular behavioural context which has not been investigated adequately in the literature. Therefore, the automated categorization of whistles introduced in the paper helps marine biologist to correlate them with animal's behavioural contents. Dolphin whistles visualized in time-frequency representation are passed through a digital bandpass filter to throw off undesired and noisy information. A well-known image processing technique is adapted to process the spectrograms and extract salient features named as eigenwhistles for all training samples of different types. In the evaluation phase, testing whistles are projected on the training eigenwhistle space and based on the Euclidean metric, they are assigned to different classes of whistles. The results demonstrate the capability of eigenwhistles to correctly classify whistles with an accuracy of 90% on a total of 500 calls with 7 call types.**

*Keywords*— **Whistle categorization, Eigenwhistle, Dolphin, Whistle perception.**

## I. INTRODUCTION

Marine mammals has been in the core attention of biologists and organizations in the past years. Since the ocean is vulnerable to harm by human activities and other environmental causes, regular census studies and conservation plans have been developed and implemented for the protection of sea lives. Accumulating many field data and creation of marine labs led to the convergence of a new research subject that studies on perception of natural animal vocalizations and how to associate different calls with behavioural contexts [1].

Bottlenose dolphins are one of the most social and cooperative species and usually live in groups of several individuals helping them to strengthen their bonds and learn feeding and swimming skills rapidly [2]. On the other hand, they have a strong sound production and reception ability hearing across a broad frequency band, up to 150 kHz [3]. Bottlenose dolphins can produce narrowband whistles ranging from 3.5 kHz to 10 kHz with fundamental frequency centred at 7 kHz and traveling up to 25 km away from the caller. Such characteristics make whistles ideal for social and communication purposes. Additionally, they vocalize broadband sounds named as clicks or echolocations used mostly for ocean navigation, foraging and feeding [4].

One of the early studies on dolphin whistles revealed that the frequency of whistle contours changes over time [5]. In another early work, vocalizations of captive bottlenose dolphins captured to identify the whistler showed that each individual produced a stereotyped and distinct whistle named as "signature whistle" that is unique among the other dolphins [6]. These whistles are not only used for individual recognition but also for group cohesion [7]-[8]. Recent studies even correlated different vocalizations with behavioural activities. Panova *et al* determined the relationship between Beluga whale's behavioural activity and underwater sounds that were classified into five major whistle types, four types of pulsed tones, echolocations and noise calls [9]. In [10], it was shown that not only different types of dolphin whistles are used in specific behavioural contexts, but also the number of behavioural contexts correlated with certain whistle type grows with the frequency of the whistle types.

A common task in bioacoustics is to determine a species repertoire of vocalizations [11]. Usually, the spectrograms of various calls are analysed followed by extraction of important features and then grouping them into several types. One of the drawbacks of previous researches is that it required whistle contour tracing that is very tedious and computationally expensive [12]. Gannier *et al* used a custom developed software to extract whistle contours to measure several temporal and frequency parameters [13]. In recent years, there have been some researches to avoid contour extraction and consider whistle spectrogram as an image and apply successful image processing filters to extract distinguishing features. In [14], three different feature extraction methods such as Gabor wavelets, Fourier descriptors, and time-frequency features were applied on dolphin whistles and sparse representation classifier categorized them into four different whistles types with a superior accuracy of 98% with Gabor features. Roch *et al* computed cepstral features of delphinid vocalizations and trained Gaussian mixture models for different species in order to predict the whistle type [15]. Another approach introduced by Esfahanian *et al* used local binary patterns adapted from image processing techniques in order to capture important features to feed into K-Nearest Neighbour (KNN) whistle detectors [16].

Eigenface is a well-known appearance-based methods used for many applications such as face recognition, object identification, etc [17]-[18]. The main idea is to find optimal basis for in a subspace that has a very low dimension by applying principle component analysis (PCA) [19]. The novelty of this research is to adapt such a powerful technique on whistle type classification of bottlenose dolphin data.

This paper has been organized as followings. Section 2 describes pre-processing methods such as denoising and band-pass filtering applied on the whistle spectrograms. Section 3 outlines the eigenwhistle algorithms and how it contributes to feature extractions. Sections 4 shows the experimental results for whistle type classification followed by conclusive remarks.

## II. PRE-PROCESSING

To effectively visualize marine mammal vocalizations and investigate the time-frequency variations of contours, Fourier transform is implemented on the signal windowed typically with an overlapping factor. The resulting two-dimensional matrix can be displayed as an image called "spectrogram" whereas the x and y axes represent time and frequency scales, respectively. The bottlenose dolphin signals were organized into 100 ms frames using 80% overlap and Hamming window of 1024 samples in length at sampling frequency of 80 kHz.

The corpus utilized in this research consists of dolphin vocalizations with various types of whistles. Therefore, a thorough examination of dataset led to identification of 7 types of signature whistles that have unique contour shapes and characteristics. Fig 1. illustrates spectrograms of each whistle type along with their names used for the purpose of this research.
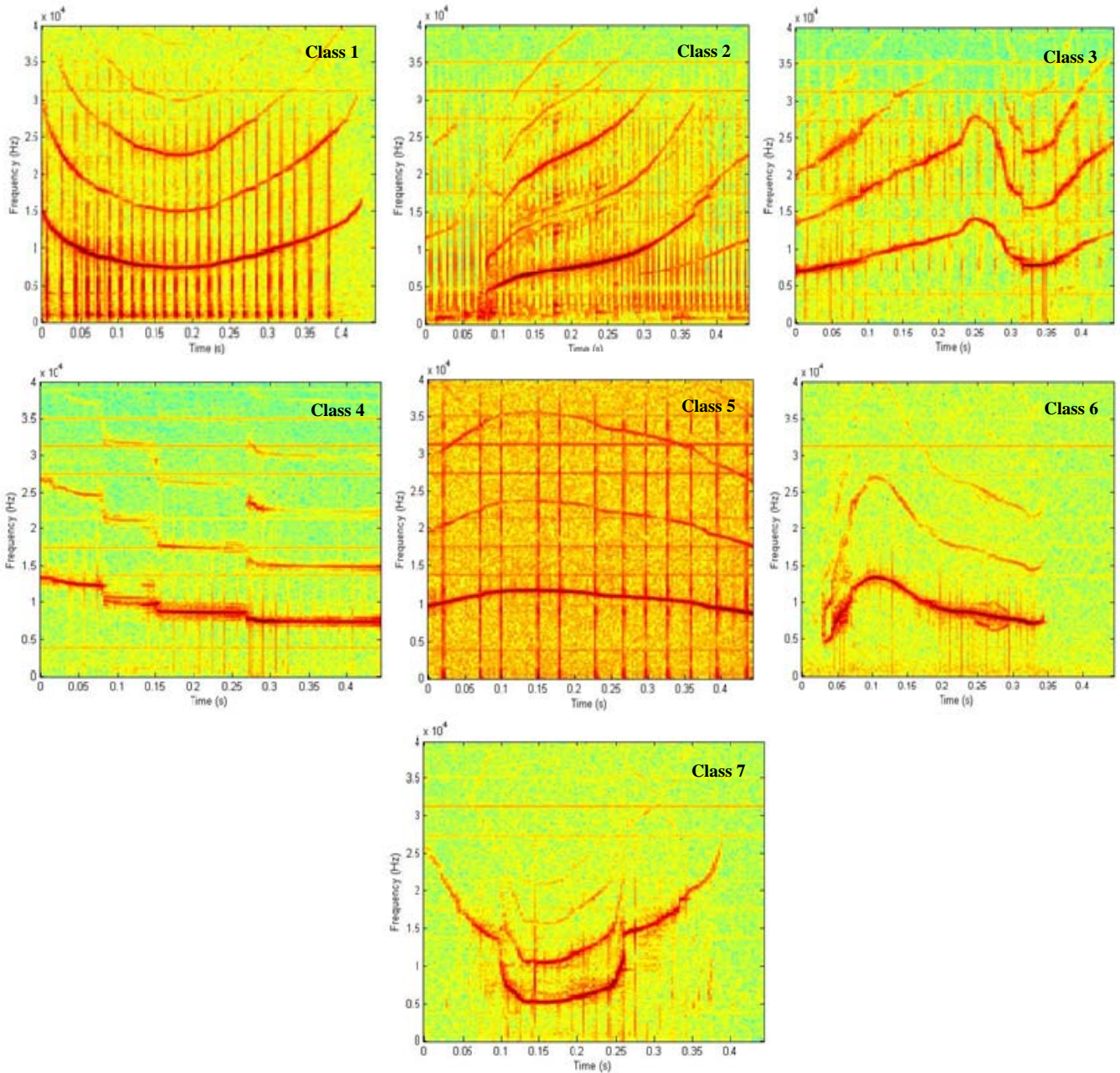


Fig. 1. Spectrograms of bottlenose dolphin whistles categorized into seven distinct types

It is observed that there is a fundamental whistle with the lowest centre frequency followed by several harmonics at higher frequencies with less intense contours. The fundamental whistles of bottlenose dolphins among these classes have a frequency varying between 3 kHz and 15 kHz. Additionally, the vertical red line in the spectrograms are associated with echolocation clicks produced simultaneously with the whistles but they are considered as undesired sounds as the focus of this research is on fundamental whistles.

The raw data collected from recording devices in the sea are contaminated by ocean ultra low frequency sounds and various noise sources generated by ships and boats engines and other species such as shrimps in shallow waters. On the other hand, depending on which animal and what type of call are under study, a specific frequency range may be of more importance. Therefore, the pre-processing phase is an inseparable component of such underwater studies.

To limit the frequency content of dolphin signals in which fundamental whistles are mostly seen, a 6th order Butterworth band-pass filter with low and high cutoff frequencies of 3 kHz and 15 kHz, respectively, are implemented on dolphin signals. Fig 2. shows the spectrogram of a band-passed whistle from class 6. To normalize the spectrogram within the desired frequency band ($f_i$), the mean value of each frequency band ($Sm(f_i)$) is subtracted from the frequency bin and then divided by that bin's standard variation ($Sstd(f_i)$) as below:

$$S_N(f_i) = \frac{S(f_i) - Sm(f_i)}{Sstd(f_i)} \quad for \quad i = 1,...,L \qquad (1)$$

where $S$ and $S_N$ is the original and normalized spectrograms, respectively, and L is number of frequency bins. The spectrogram values are thresholded in order to isolate the fundamental whistle for further processing as shown in Fig 3.
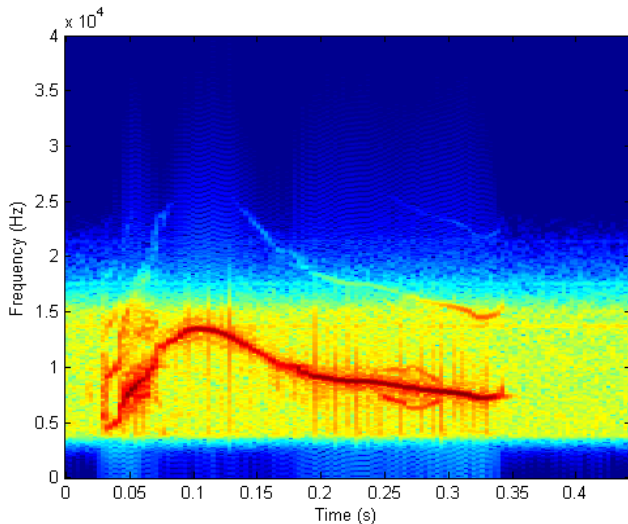


Fig. 2 The spectrogram of a band-passed whistle from 6th class

## III. EIGENWHISTLE ALGORITHM

There are various solutions introduced in the literature for the challenging problem of face recognition. One of the most common methods based on the appearance is eigenface [33f]. The basic idea behind the eigenface is to reduce the information by Principle Component Analysis (PCA). Due to its successful implementation and high accuracy rate, this technique has been adapted accordingly and named "eigenwhistle" for whistle type identification. There are training and testing phases involved in this approach that are explained in details at following.
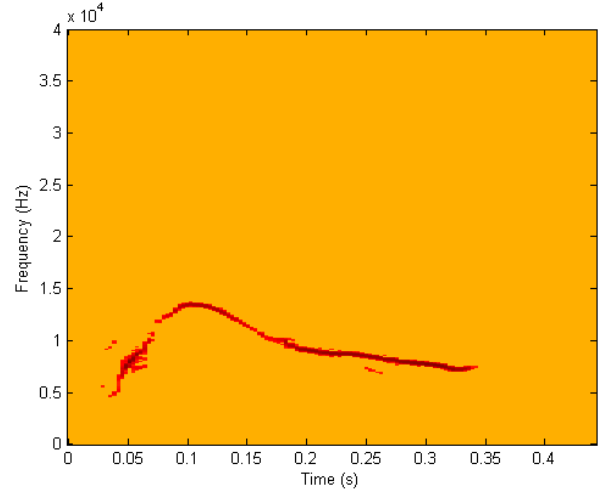
### A. Training phase



Fig. 3 The isolated 6th-class fundamental whistle after denoising spectrogram

Suppose, each whistle spectrogram is of size $n_r$ and $n_c$ denoting the number of rows and columns, respectively. Then, each spectrogram is vectorized column-wisely as a 1-D vector ($\Gamma_i \in \mathbb{R}^{(n_r . n_c) \times 1}$) as depicted in Fig 4.
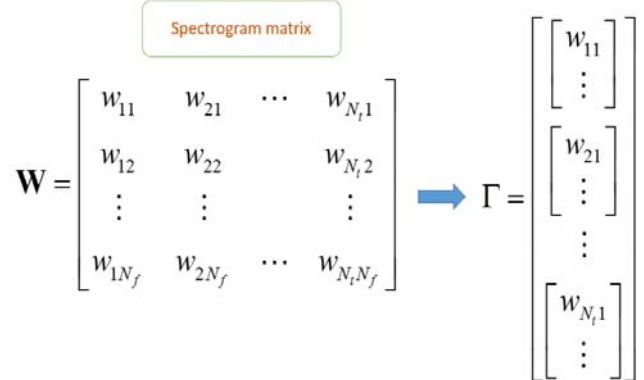


Fig. 4 Illustration of vectorizing spectrogram matrix

If training database consists of M whistles:

$$\Gamma = \{\Gamma_1, \Gamma_2, ..., \Gamma_M\} \qquad (2)$$

The average whistle spectrogram can be computed as:

$$\Psi = \frac{1}{M}\sum_{i=1}^{M}\Gamma_i \qquad (3)$$

The difference whistle spectrogram is calculated by subtracting mean vector from each vectorized spectrogram:

$$x_i = \Gamma_i - \Psi \qquad (4)$$

The concatenation of all 1-D spectrograms constructs a large matrix of size $(n_r \times n_c) \times M$

$$X = [x_1, x_2, ..., x_M] \qquad (5)$$

The covariance matrix $\Im$ is established as:

$$\Im = \frac{1}{M}XX^T = \frac{1}{M}\sum_{i=1}^{M}x_i x_i^T \qquad (6)$$

To make the inner product less computationally expensive, the following R covariance matrix is obtained

$$R = \frac{1}{M} X^T X \qquad (7)$$

Now, an eigenvector matrix $Q$ of size $M{\times}M$ can be founds followed by calculation of eigenwhistle matrix as

$$Y = X.Q \qquad (8)$$

At the end, whistle matrix is projected on the eigenspace:

$$Z = Y^T.X \qquad (9)$$

where each column of $Z$ represent the feature vector of each whistle.

### B. Testing phase

The whole purpose of this phase is to classify a whistle into one or more of the training whistles.

Assume the test whistle spectrogram is of size $n_r \times n_c$. After it is vectorized as a 1-D vector $P$, the difference test spectrogram is computed as

$$v = P - \Psi \qquad (10)$$

It is then projected into the eigenspace to find the test feature vector

$$u = Y^T v \qquad (11)$$

At the last step, the test whistle spectrogram is classified by comparing $u$ with every column of Z using the Euclidian metric. Therefore, the class corresponding to the smallest distance is assigned to the test whistle.

## IV. EXPERIMENTAL RESULTS

The performance of eigenwhistle algorithm proposed in this research was evaluated on a dataset of bottlenose dolphin whistles that were categorized into seven classes based on their unique temporal and frequency shapes. To remove the bias from the classification results, training whistles were selected in a random process and remaining ones are used as testing samples. It is worthwhile to mention that this approach does not required contour tracing or identifying whistle contour in the spectrogram. Also, it is more robust to variations in the image caused by some pieces of harmonics and echolocation clicks produced simultaneously.

The database used for this study contained 420 whistles of 7 types with different variations and roughly the same number of samples was assigned for each whistle class. Also, half of the samples for each class was randomly selected to construct the training set and the remaining ones established the testing whistles. All the programs and algorithms were written in MATLAB environment.

The confusion matrix of classifications results is shown in Table I. The overall accuracy of eigenwhistle approach with Euclidean similarity metric is 92% while there are only 17 misclassifications occurred with this approach proving the capability of eigenwhistles in identifying various whistle types. Eigenwhistle results is comparable with the accuracy reported in [20] by KNN and different features and even better than SVM with Fourier Descriptor features. Also, the dataset size of this study and number of whistle types is less than this study.

TABLE I. CONFUSION MATRIX FOR CLASSIFICATION OF DIFFERENT TYPES OF WHISTLES

| Confusion Matrix | | Classification | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Labels | 1 | 27 | 1 | 0 | 0 | 0 | 0 | 2 |
| | 2 | 2 | 26 | 2 | 0 | 0 | 0 | 0 |
| | 3 | 0 | 0 | 28 | 0 | 0 | 2 | 0 |
| | 4 | 0 | 0 | 1 | 28 | 1 | 0 | 0 |
| | 5 | 0 | 0 | 0 | 0 | 28 | 2 | 0 |
| | 6 | 0 | 0 | 2 | 0 | 1 | 27 | 0 |
| | 7 | 1 | 0 | 0 | 0 | 0 | 0 | 29 |

It is also observed that the second whistle type with 4 misclassifications (2 first class and 2 third class) had the highest number. One reason is that some segments are the second class is similar to the first and second class. Another reason could be the incapability of Euclidean method to measure their similarities accurately. Also the seventh whistle type has the least number of misclassifications since its range of time and frequency variations don't resemble the other whistle types.

## V. CONCLUSION

In this research, the classification of various types of bottlenose dolphin whistles was implemented and evaluated on a fairly large dataset. A famous image processing technique used for face detection was adapted for identification of different whistle and named as "eigenwhistle". This approach uses the concept of principle component analysis to reduce the dimensionality of data and Euclidean metric to measure the similarity. This approach eliminates one of the main problems in whistle detection and identification called contour tracing. The classification accuracy obtained by eigenwhistle is 92% considering the fact that there exist seven whistle types equally populated in the database. The eigenwhistle approach proved its superior accuracy and robustness against time and frequency variations and other noise effects seen in the whistle spectrograms.

## REFERENCES

[1] J.H. Poole, K. Payne Jr, W.R. Langbauer, C. J. Moss, "The social context of some very low frequency calls of African elephants," Behav Ecol Sociobiol, vol. 22, pp. 385–392, 1988

[2] R.S. Shane, S.H. Wells and B. Wursig, "Ecology, behavior and social organization of the bottlenose dolphin: A review," Marine Mammal Science, vol. 2, no.1, pp.34-63, 1986.

[3] C.S. Johnson, "Masked tonal thresholds in the bottlenosed porpoise," J. Acoust. Soc. Am. 44, 965–967, 1968.

[4] W.W.L. Au, The Sonar of Dolphins. Springer-Verlag, New York, 277 pp. Batteau, D.W., Markey, P.R., 1968. Man/dolphin communication. Final Report for U.S. Naval Ordnance Test Station. China Lake, CA (contract number N00123-67-C-1103), 1993.

[5] J. J. Dreher, "Linguistic considerations of porpoise sounds," J. Acoust. Soc. Am. vol. 33, pp. 1799–1800, 1961.

[6] M.C. Caldwell, D.K. Caldwell, "Individualized whistle contours in bottlenosed dolphins (Tursiops truncatus)," Nature, vol. 207, pp. 434–435, 1965.

[7] M.C. Caldwell, and D.K. Caldwell, "Vocalization of naive captive dolphins in small groups," Science, vol. 159, pp.1121–1123, 1968.

[8]   V. M. Janik, and P. J. B. Slater, "Context-specific use suggests that bottlenose dolphin signature whistles are cohesion calls," Animal Behavior, vol. 56, pp. 829-838, 1998.

[9]   E.M. Panova, R.A. Belikov, A.V. Agafonov, and V.M. Bel'Kovich, "The relationship between the behavioral activity and the underwater vocalization of the beluga whale (Delphinapterus leucas)," Oceanology, vol. 52, pp.79–87, 2012.

[10]  R. Ferrer-i-Cancho, B. McCowan, "A law of word meaning in dolphin whistle types," Entropy, vol. 11, pp. 688–701, 1995.

[11]  D. Wang, B. Wursig, and W. Evans, "Comparisons of whistles among seven odontocete species," in Sensory Systems of Aquatic Mammals, edited by R. A. Kastelein, J. A. Thomas, and P. E. Nachtigall (De Spil, Woerden), pp. 299–323, 1995.

[12]  PJ. Van Otterloo "A Contour-Oriented Approach to Shape Analysis," Englewood Cliffs, New Jersy: Prentice-Hall; 1991.

[13]  A. Gannier, S. Fuchs, P. Quèbre, J.N. Oswald, "Performance of a contour-based classification method for whistles of Mediterranean delphinids," Applied Acoustics, vol. 71, pp. 1063-1069, 2010.

[14]  M. Esfahanian, H. Zhuang and N. Erdol, "A new approach for classification of dolphin whistles," 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, pp. 6038-6042, 2014.

[15]  M.A. Roch, M.S. Soldevilla, J.C. Burtenshaw, E.E. Henderson, and J.A. Hildebrand, "Gaussian mixture model classification of odontocetes in the Southern California Bight and the Gulf of California," J. Acoust. Soc. Am. vol. 121, pp. 1737–1748, 2007.

[16]  M. Esfahanian, H. Zhuang, and N. Erdol, "Using local binary patterns as features for classification of dolphin calls," The Journal of the Acoustical Society of America, vol. 134, no.1, pp. 105-111, 2013.

[17]  W.S. Yambor, B.A. Draper and J.R. Beveridge, "Analyzing PCA-based Face Recognition Algorithms: Eigenvector Selection and Distance Measures," Second Workshop on Empirical Evaluation in Computer Vision, pp. 39-60, 2000.

[18]  M.A. Turk, and A.P. Pentland, "Eigenfaces for Recognition", J. Cognitive Neuroscience, vol. 3, pp. 71-86, 1991.

[19]  K. Pearson, "On Lines and Planes of Closest Fit to Systems of Points in Space", Philosophical Magazine, vol. 2, no. 11, pp. 559-572, 1901.

[20]  M. Esfahanian, H. Zhuang and N. Erdol, "On contour-based classification of dolphin whistles by type," Applied Acoustics, vol. 76, pp. 274-279, 2014.