Graduate Research Day 2013
Florida Atlantic University

College of Engineering and Computer Science

Sparse Representation Classification of Dolphin Whistles Using Local Binary Patterns

Mahdi Esfahanian, Dr. Hanqi Zhuang, Dr. Nurgun Erdol

Electrical Engineering; Florida Atlantic University

Compressive sensing has been recently emerged in the area of pattern recognition and signal processing. In this research, a Sparse Representation Classifier (SRC) is adapted and applied to the classification of bottlenose dolphin whistles. This technique relies on near completeness of the training features, rendering their choice no longer crucial as long as certain criteria are met. Classes of dolphin calls are determined by a linear basis pursuit algorithm that minimizes the l1-norm of the error vector. Also Signal sparsity is ensured via the employment of a robust, effective, and computationally simple Local Binary Pattern (LBP) operator that eliminates the need for costly denoising and contour tracing operations. The performance of the SRC is evaluated and then compared with outputs obtained from K-Nearest Neighbor (KNN) and two feature extraction methods: Time-Frequency Parameters (TFP) and raw pixels. The experimental results demonstrate superior capability and accuracy of the proposed method on classifying dolphin whistles into distinct call types. The method can be generalized to all narrowband signals with time varying spectra.

Sparse Representation Classification of Dolphin Whistles Using Local Binary Patterns

Mahdi Esfahanian, Hanqi Zhuang, and Nurgun Erdol

Department of Computer and Electrical Engineering and Computer Science

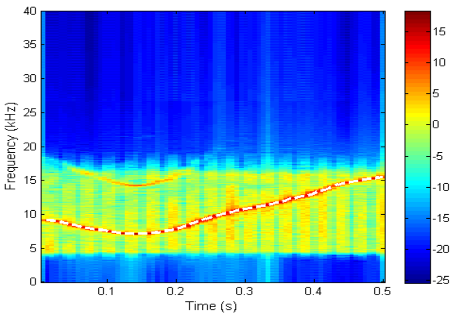
Southeast National Marine Renewable Energy Center

MOTIVATION

- A major task in bioacoustic research is analyzing recorded repertoires of various species to either be associated with a behavior or determine animal identity.
- Bottlenose dolphins can produce whistles that are narrow-band long duration sounds from 1 kHz up to 24 kHz.
- Marine mammal vocalizations are usually corrupted by different acoustic sources making the pre-processing task necessary and costly.
- Most feature extraction techniques require tracing the contour of whistles in order to obtain distinguishing information for recognition purposes.

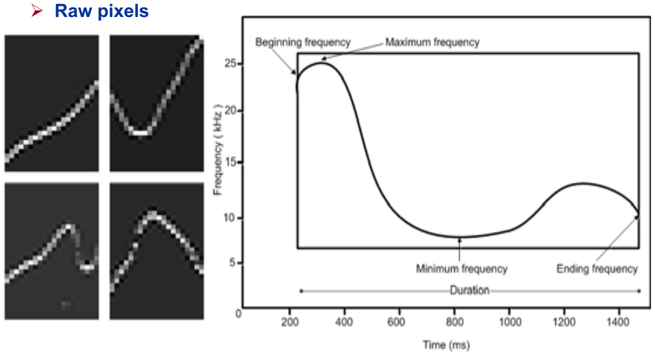
METHODOLOGY

- Pre-processing stage may implement band-pass filtering with cut-off frequencies at 4 kHz and 16 kHz as well as spectral denoising.
- If needed to contour tracing, the highest frequency peak of isolated fundamental whistle is picked at each time instance.



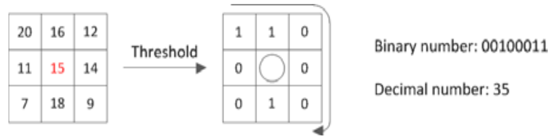
FEATURE EXTRACTION

- Conventional:
  - Time-Frequency Parameters (TFPs)
  - Raw pixels

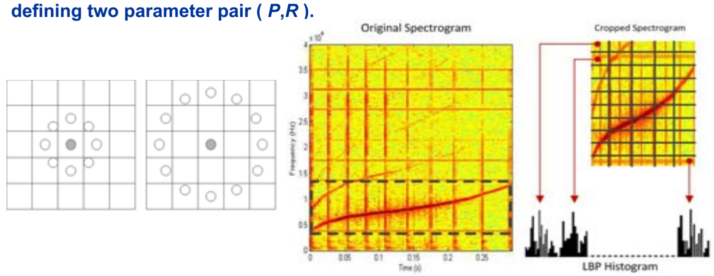


- Local Binary Patterns (LBP)

Neighboring pixels are assigned the binary digit (bit) 1 if larger than the center pixel and 0 otherwise.



The LBP operator can be extended to a circularly symmetric neighborhood by defining two parameter pair (*P*,*R*).

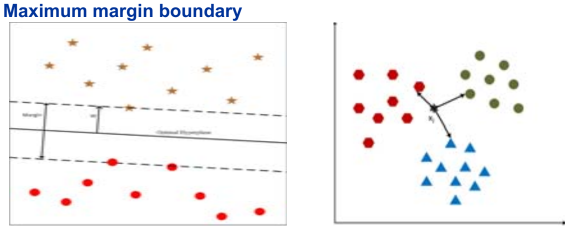


The feature vector assignment using LBP is comprised of following steps:

a) Label all the pixels through LBP coder
b) Divide the spectrogram into M small rectangular regions
c) Obtain histogram of each region, column-wisely
d) Concatenate all the histograms into one vector

CLASSIFICATION METHODS

- Conventional:
  - *K*-Nearest Neighbor (KNN)
    - Based on the closet distance to the training data
  - Support Vector Machine (SVM)
    - Maximum margin boundary



- Sparse Representation Classifier (SRC)

The application of compressive sensing to classification has been recently emerged where sparse representation of the test sample is obtained using a training data set.

- Consider a dictionary $U=[U_1,U_2,...,U_d]$ where each class consists of n_i training vectors $U_i=[u^1,u^2,...,u^{ni}]$ for $i=1,2,...,d$. A test sample y from i-th class can be represented as a linear combination of all training samples from d distinct classes as:

$$y=\sum_{i=1}^{n_1}\beta_1^i u_1^i+\sum_{i=1}^{n_2}\beta_2^i u_2^i+\ldots+\sum_{i=1}^{n_d}\beta_d^i u_d^i \quad \text{or} \quad y=U\alpha_0$$

In this case, all the elements of weighting vector α_0 should be zero but those associated with the same class: $[0,\ldots,0,\beta_i^1,\ldots,\beta_i^{n_i},0,\ldots,0]$

- Now the signal y is deduced to be sparse in domain U and the convex optimization problem can be solved:

$$\bar{\alpha}=\operatorname{argmin}\|\alpha_0\|_1 \quad \textit{subject to} \quad \|y-U\alpha_0\|_2\le\varepsilon$$

- Once the sparse solution is obtained, the classification procedure of SRC is summarized as follows:
  1) For each class, find the residual between the reconstructed sample and the test sample: $r_i=\|y-U_i\bar{\alpha}_i\|_2 \quad \textit{for} \quad i=1,2,...,d$
  2) The class of test sample is determined by: $class(y)=\arg\min_i(r_i)$

RESULTS

- The best result is achieved by combination of SRC and LBP, as expected.

Comparison of SRC performance using different feature extraction techniques

Feature method	RAW	LBP	TFPs
Accuracy (%)	92	100	92

Confusion matrix of classifiers' outputs in the following order: [SRC+RAW, KNN+LBP, SVM+LBP, SRC+LBP]

	First class (%)	Second class (%)	Third class (%)	Fourth class (%)
1st class	[92, 92, 100, 100]	[8, 8, 0, 0]	0	0
2nd class	[12, 0, 6, 0]	[88, 100, 94, 100]	0	0
3rd class	0	0	[83, 100, 100, 100]	[17, 0, 0, 0]
4th class	0	[0, 14, 0, 0]	0	[88, 100, 94, 100]

CONCLUSIONS

- A new algorithm that combines a compressive-sensing based technique called Sparse Representation Classifier and a feature vector set derived from Local Binary Patterns was developed and applied to the classification of dolphin whistle types.
- The marriage of the SRC classifier and LBP feature set achieves superior classification accuracy rates and eliminates the need to remove echo location clicks, denoise, and trace whistles trajectories, all of which are costly and tedious preprocessing tasks.
- The proposed algorithm was tested against other classifier-feature combinations in terms of overall classification accuracy. Particular attention was paid to the individual contributions of SRC and LBP. It was observed that LBP improved other classifiers, and SRC with raw data was better than KNN and SVM classifiers with other features.

REFERENCES

Ahohen, T., Hadid, A., and Pietikäinien, M. "Face recognition with local binary patterns," European Conference on Computer Vision, pp. 469-481, 2004.

Baraniuk, R. G. "Compresseive sensing," IEEE Signal Processing Magazine, pp. 118-121, 2007.

Candès, E. J., and Wakin, M. B. "An introduction to to compressive sensing," IEEE Signal Processing Magazine, pp. 21-30., 2008.

Ojala, T., Pietikäinen, M., and Harwood, D. "A comparative study of texture measures with classification based on featured distribution," Pattern Recognition, vol. 29, pp. 51-59, 1996.

Ojala, T., Pietikäinen, M., and Mäenpää, T. "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, pp. 971-987, 2002.

Oswald, J. N., Barlow, J., and Norris, T. F. "Acoustic identification of nine delphinid species in the eastern tropical pacific ocean 19, 20-37.," Marine Mammal Science, vol. 19, pp. 20-37, 2003.

ACKNOWLEDGMENT

This project was partially supported by SNMREC.