A Hybrid Fuzzy Logic-Neural Network Approach For Multi-path Separation Of Underwater Acoustic Signals

Abigail Lee-Leon *†, Chau Yuen *, Dorien Herremans *
* Singapore University of Technology and Design (SUTD), 8 Somapah Rd, Singapore 487372
† Thales Solutions Asia Pte Ltd, 21 Changi North Rise, Singapore 498788

***Abstract*—Underwater acoustic channels are generally recognized as one of the most difficult communication media in use today. One of the most important constraints of underwater communications is the chaotic acoustic propagation of the signals. The aim of a multi-path separator is to extract individual correlated signals from their mixtures. We present an algorithm, Tag Receiver, that is similar to the rake receiver, but interpreted from the neural network viewpoint. First, the received signal is split into frames. Next, the frames are tagged with a predicted number of multi-path in the framed segment via fuzzy logic on extracted audio features. Finally, the frame and the number of multi-path are inputted into a neural network, which separates the signals. The analysis of the proposed algorithm shows a 1.2dB improvement when bit error rate (BER) is 10^{-3} over the conventional rake receiver.**

***Keywords*—Multi-paths Separation, Fuzzy logic, Neural Networks, Signal Processing, Feature Extraction**

I. Introduction

Acoustic signals produced in underwater or shallow water environments undergo multi-path propagation due to the sound propagating at a very low speed. This causes the observed signals to consist of many delays that are over tens or even hundreds of milliseconds long, and are summed corrupted copies of the desired source [1]. Multi-path formation in the ocean is governed by two effects —sound reflecting off underwater surfaces like bubbles and seabed, and sound refraction in the water due to density change. These effects will cause an elongation of the path traveled and thus a time delay.

Multi-path formation is a reverberation characteristic that diminishes the perceived quality of the received signals. In extreme cases, this characteristic can even corrupt comprehensibility of the signal. Multi-path separation aims to reduce or eliminate the damaging effects of reverberation [2]. This task is completed by either leveraging on the multiple paths or determining the most prominent single path. Research into this has progressively gained enormous interest from different research fields, such as speech recognition [3], music separation [4] and communications [5].

If we look into the field of human speech, we would notice that human speech is highly subjected to reverberation and multi-path propagation. The idea is to create a system that mimics the human perception and cognition. It has been proven that human ears can differentiate between twenty-four frequencies without considering phase differences [6]. The human ears can distinguish noise and speech patterns, both of which are in fact linked to changes in frequency. For instance, a human is able to identify a person's voice and echoes of the voice.

There are two trends of multi-path separation —uncorrelated signals and correlated signals. The first is considered a classical problem in speech or audio signal processing called "cocktail-party problem" [7] . The individual sources are first mixed and then extracted using unique features. The latter is observed as more difficult and can be separated using a direction of arrival estimation [8]. Another method is the rake receiver, used for Code Division Multiple Access (CDMA) cellular systems. This method combines multi-path components and attempts to detect the time delayed versions of the transmitted signal by creating individual correlation receivers for each of the multi-path signals. However, this can only be done when their multi-path components have a relative time delay that is greater than a chip period thus making the signals uncorrelated.

Fuzzy logic is considered to be an adaptive machine learning technique. There has been much research done into fuzzy logic methods applied to many areas of communications, such as fuzzy logic filter for coherent detection [9] and Adaptive Modulation techniques [10]. Fuzzy logic is a machine learning technique that can categorize the changes in the features extracted from the received signal. It was first proposed by [11] to construct human-like reasoning and problem solving characteristics into a more efficient mathematical process. Due to this feature, fuzzy logic is able to handle imprecise data, non linear functions and can be used with other techniques for different types of problem solving [12].

The proposed algorithm, Tag Receiver, considers the task of multi-path separation in a very similar way as the rake receiver, but the process is interpreted from the viewpoint of neural networks. The proposed algorithm is a hybrid model that integrates fuzzy logic and neural networks. The fuzzy logic tags the segments of received signal with the predicted number of multi-paths, which in turn directs the segment into the correct neural network for path separation.

The remainder of this paper is organized as follows. Section II provides an overview of the system model. Section III describes the components of the Tag receiver. In Section IV,

Preprint accepted for publication: Lee-Leon A., Yuen C., Herremans D.. 2019. A Hybrid Fuzzy Logic-Neural Network Approach For Multi-path Separation Of Underwater Acoustic Signals. 89th IEEE Vehicular Technology Conference. Kuala Lumpur, Malaysia.

978-1-7281-1217-6/19/$31.00 ©2019 IEEE

the simulations used to test the design are presented together with their results. Section V discusses the conclusion and future work.

II. System Model and Channel

In Underwater Acoustic Communication (UWAC), the channel is complex to understand and model. Many tractable communications channel models are derived from a simplified version of Maxwell's equation. However, there are many examples where this does not hold, such as when the signal is not electromagnetic or when the propagation of the signal is either complex or poorly understood [13]. Therefore, in this section, we describe the channel and system model used for the testing of the Tag receiver.

First, we consider a single transmitter-receiver system employing 16-Quadrature Amplitude Modulation (QAM) in a multi-path channel with Rayleigh fading and Addictive White Gaussian Noise (AWGN). Fig.1 shows a block diagram of the system model.

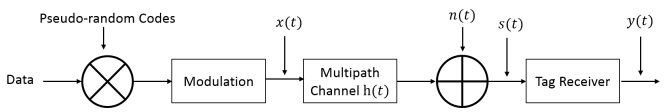

Fig. 1: Block Diagram of the system

To simplify the description of the channel, the transmitted signal $x(t)$ is relayed through a simple multi-path channel modeled by:

$$s(t) = \sum_{i=0}^{N} a_i b(\theta_k(t)) \circ x(t - \tau_i) + n(t) \tag{1}$$

where the symbol $\circ$ represents the Hadamard product, $s(t)$ is the received signal, $n(t)$ is awgn, N is the number of multi-paths, a_i is the amplitude variation caused by the scattering and fading of the channel, and τ_i is the time delay. The phase variation due to the refraction and reflection on water surfaces and objects is modeled as:

$$b(\theta_k(t)) = [1 \quad e^{-j\theta_1(t)} \ldots \quad e^{-j\theta_k(t)}] \tag{2}$$

where k is the length of the signal and θ_k is the phase shift corresponding to the change in angle due to scattering.

In the next section, we will describe the proposed "Tag receiver', which will include trying to estimate the weight vectors, which provides estimates of the desired signal at the neural network output.

III. Proposed Method for Multi-path Separation

The proposed method, Tag receiver, is described in ths section. With the purpose of acquiring signals in reverberant environments, a receiver that is able to take advantage of obtaining and combining the individual multi-paths would be beneficial. The multi-paths can serve as reinforcement of the direct path signal and improve the signal component.

A. Tag Receiver

The conventional rake receiver uses correlators called "fingers" to coherently combine the multi-path signals using the channel knowledge of each multi-path signal. This knowledge can be obtained by transmitting training sequences at a specific time interval.

The proposed algorithm, Tag receiver, uses two machine learning models to label the number of multi-path signals and then separate them into their individual signals. The fuzzy logic section tags the time-windowed received signal with the predicted number of multi-paths by distinguishing the previously identified features that are more sensitive to the change in number of multi-path signals. After this, the correct neural networks is selected, which in turn can separate the signal into their paths.

An overview of the system is shown in Fig.2. The detailed description of the components are shown in section III-B and section III-C.

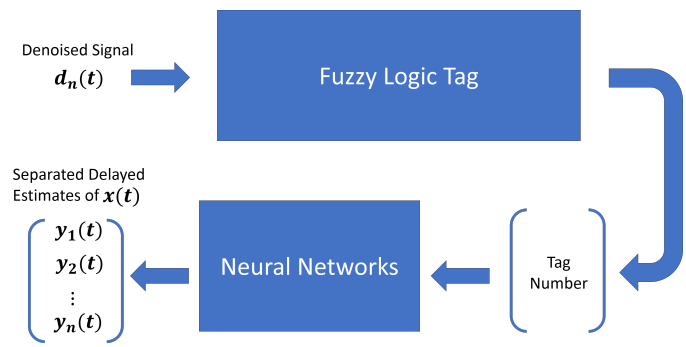

Fig. 2: Multi-path Separator Process

B. Fuzzy Logic Tag (FLT)

For the first part of the Tag receiver, the basic implementation of the Fuzzy Logic Tag (FLT) is to estimate the number of paths in the segment of received signal $s(t)$. Fig. 3 depicts the process of the FLT.

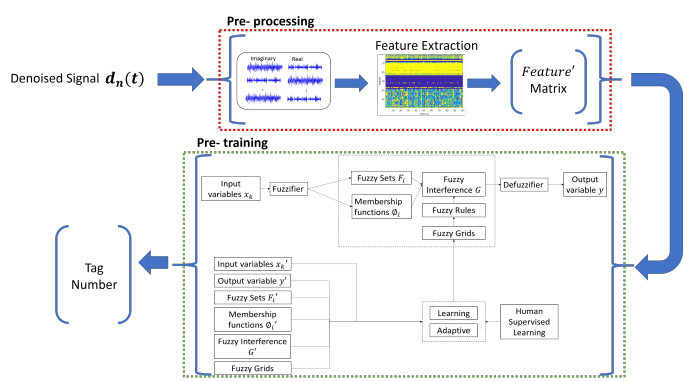

Fig. 3: Fuzzy Logic Tag Process

The input for the FLT is the received signal $s(t)$. For the pre-processing stage, the signal is broken down into frames and converted into the corresponding feature matrix, $Feature$. The features then undergo the process of optimization, thus

Preprint accepted for publication: Lee-Leon A., Yuen C., Herremans D.. 2019. A Hybrid Fuzzy Logic-Neural Network Approach For Multi-path Separation Of Underwater Acoustic Signals. 89th IEEE Vehicular Technology Conference. Kuala Lumpur, Malaysia.

deciding which sets of features (listed in Table I) are more strongly impacted by the number of multi-paths.

1) Pre-processing: The feature extraction process involves a temporal analysis of features, whereby individual features are calculated from 10ms half-overlapping frames of $s(t)$. This feature set, shown in Table I, was based on work done by D. Li and I.K. Sethi [14].

TABLE I: Features corresponding to rows in $Feature'$ Matrix

Feature number	Feature
1	Zero-Crossing Rate
2	Energy
3	Entropy
4	Root-mean-square
5	Integrated
6	Cosine
7	Spectral Centroid
8	Spectral Entropy
9	Spectral Flux
10	Spectral Roll-off
11-23	MFCC
24	Harmonic Ratio
25	Fundamental Frequency
26-38	Chroma Vector
39-59	FFT
60-79	Gradient FFT

In feature selection, we consider two aspects of the individual features—the value and the variance, to the number of multi-paths. For the evaluation, the Fisher's discriminant ratio modeled in [15] was used:

$$f(p) = \frac{(\mu_p - \mu_{p+1})}{\sigma_p - \sigma_{p+1}} \tag{3}$$

where μ_p, μ_{p+1}, σ_p and σ_{p+1} are the mean values and the within-class scatter of the pth class.

We observed three main features with high sensitivity to a change in the mutli-path signal. The integrated signal over time (Feature number 5) and a section of the FFT (Feature number 40) had a normalized weightage of 0.051 and 0.057 respectively. The highest normalized weightage of 0.7771 was found in sections of the FFT gradient (Feature number 59), along with 0.047 (Feature number 61) and 0.204 (Feature number 78).

2) Pre-training: Since we aim at including multiple delays in multi-paths, we used a normally distributed random generator for creating our datasets. The set of randomly delayed multi-paths is inputted as training data and the number of multi-paths is used as a label. We used the MATLAB anfis library [1] to model complex system behaviors using simple logic rules, and then implement these rules in a fuzzy inference system.

The concepts of fuzzy logic revolve around the idea of "linguistic variable" (x_i), "linguistic quantity" (fuzzy sets) and "membership function" (ϕ_{ij}). The function ϕ_{ij} is known as the membership function. It is used to determine the degree of which x_i belongs to a fuzzy set, which is usually stored in a numerical value between 0 and 1.

Consider the following system, where the input variable is a set of the framed detected signal $x(t) = \{x_1, x_2, ..., x_k\}$ and the output is the estimated number of mutli-path signals in that segment $y = Tag\_Number$, where each frame consists of the same pre-determined number of sampled points (time windowing). $x(t)$ is mapped into the fuzzy interference G, which is described by i number of fuzzy sets $F_i(x) = \{F_i^1, F_i^2, ..., F_i^k\}$ and membership functions $\phi_i(x) = \{\phi_i^1, \phi_i^2, ..., \phi_i^k\}$. These sets and membership functions are determined through the training of the FLT algorithm.

As an illustration of the implementation, consider the situation where the number of fuzzy sets is set at $i = 3$. Each set represents one of the following features —f_1, f_2 and f_3.

If the first set f_1 represents FFT (feature number 40), then $F_1(x)$ can be characterized into term sets, such as {No peak, low peak , moderate peak, large peak, max peak} and mapped onto G. An example of the membership function is depicted in Fig. 4.

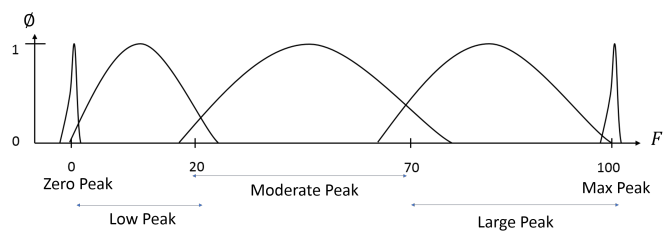

Fig. 4: Membership Function

To evaluate the output of a fuzzy interference system, the degree of which each set of $x(t)$ belongs to each $F_i(x)$ should be determined. The fuzzifier first maps the each framed $x(t)$ to members in G, given by:

$$F_i(x_k) = g(x_k), \quad F_i(x) \in G$$

During training, the FLT input variable set x_k is first mapped into a fuzzy set $F_1(x_k)$ with a degree $\phi_1(x_k)$, another set $F_2(x_k)$ with another degree $\phi_2(x_k)$, and so on. A set of fuzzy rules are created to build a base $R = \{R_1, R_2, ..., R_i k\}$, where the rules assume the form of "$if...then...$"; "if x_k is in $F_1(x_k)$ with a $\phi_1(x_k)$ degree of belonging, $F_2(x_k)$ with a $\phi_2(x_k)$ degree of belonging, and..., then y_k is $F'(y_k)$."

Once the training is completed, the output y is inputted into the Neural Network section of the algorithm.

C. Neural Network

Depending on the Tag Number, the framed signal is put through a different neural network. This approach considers the task of multi-path separation in a very similar way as the rake receiver, but the process is interpreted from the viewpoint of neural networks.

In Fig. 5, the model for the channel is presented: the instantaneous mixture of multi-paths is achieved by

[1] https://www.mathworks.com/matlabcentral/fileexchange/36098-adaptive-neuro-fuzzy-inference-systems-anfis-library-for-simulink

summation of the overlapping sound waves. Next, the signal separation is accomplished through a neural network trained on a dataset of desired separated signals.

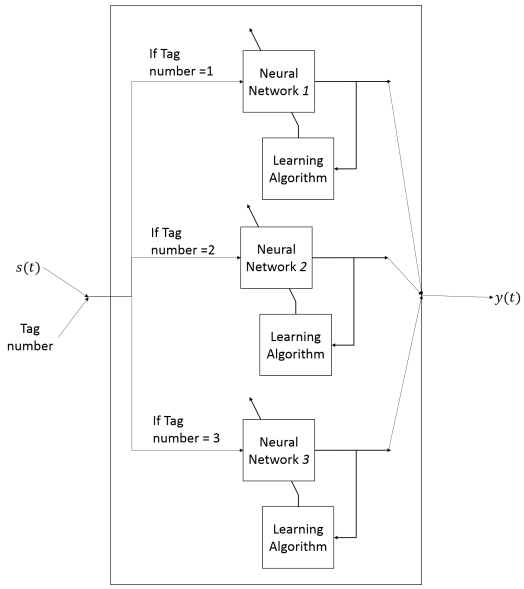

Fig. 5: Block Diagram for Neural Network Separator

The feed-forward neural network works to separate the signals via:

$$y(t) = \mathbf{W} \cdot s(t) \tag{4}$$

where $y(t)$ is the estimation of the signal and $\mathbf{W}$ is the separation matrix corresponding to the neutral network weights.

A multi-layer feed-forward neural network consists of neuron that are constructed into layers. The neural network is designed in such a way that the first layer, which the input layer, takes in the Tag Number from the FLT and $s(t)$. The size of the last layer, called the output layer, is determined by the Tag Number. If Tag Number is 3, the input signal of $k \times 1$ will be outputted as a signal matrix of $3k \times 1$. For this reason, the Tag Number "directs" the signal segment to a different neural network for the separation.

The connection between the ith and jth neurons—N_i and N_j—are characterized by weight and bias coefficients, w_{ij} and b_{ij}, which reflects the importance of the neurons' relationship and is modeled by the equation:

$$N_i = f(\ell_i) \tag{5}$$

$$\ell_i = b_{ij} + \sum w_{ij} N_j \tag{6}$$

where $f(.)$ is the activation function.

In Table II, the activation and number of nodes used in the experiments are shown.

TABLE II: Tag Number Neural Networks

Tag Number	Number of layers	No. of connecting nodes	Type of activation
1	0	[-]	[-]
2	3	[500,50,50]	[Sig, Sig, Linear]
3	3	[500,100,50]	[Sig, Sig, Linear]

IV. Simulation Results

In this section, we evaluate and compare the above proposed architectures and training strategies. As a threshold for comparison to the above proposed method, we used the maximal ratio combiner (MCR) described in [16], [17].

For the Tag receiver proposed in Section III-B, we calculated the average error rate since the accuracy can be misleading at times when dealing with a variation in delays and noise levels. This was done with a delay that is at most 0.2s (where the transmission rate is 50 bits/s). The average results of 100 experiments is shown in Fig. 6 and Fig. 7.

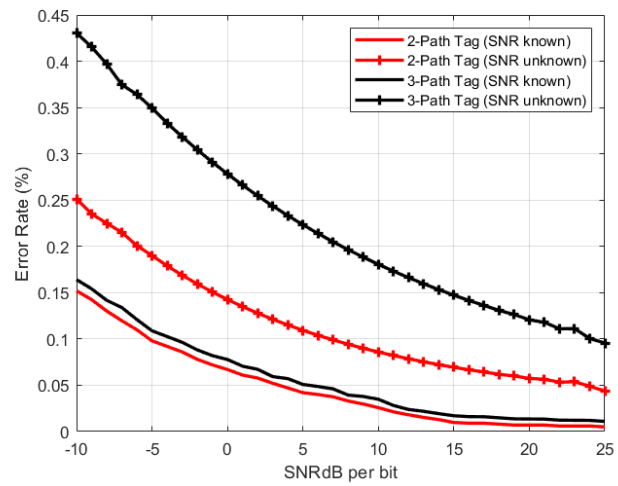

Fig. 6: Error Rate for the Tag receiver after traning and testing 2-paths and 3-paths multi-paths tagging (dataset size of 10^4)

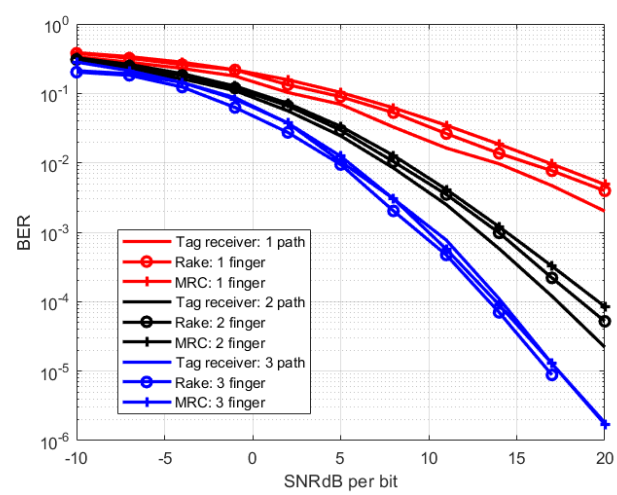

Fig. 7: Comparison of BER of the different multi-path architectures under the conditions that both number of multipath and SNR are unknown

To evaluate the performance of the whole algorithm,Tag receiver, the model was trained using two different training sets, which were synthetically generated by MATLAB. When the signal-to-noise ratio (SNR) is known but the number of paths are unknown, each of the training sets is labeled with the

Preprint accepted for publication: Lee-Leon A., Yuen C., Herremans D.. 2019. A Hybrid Fuzzy Logic-Neural Network Approach For Multi-path Separation Of Underwater Acoustic Signals. 89th IEEE Vehicular Technology Conference. Kuala Lumpur, Malaysia.

SNR and trains a different fuzzy logic system corresponding to the SNR. As seen in Fig 6, the proposed Tag receiver achieves a minimum error rate of 0.02 when trained to tag up to 3 paths and 0.027 for 2 paths. The second dataset contained signals with unknown SNR. On this set, our proposed algorithm performs significantly better than the state-of-the-art.

Fig. 7 presents the results of the Tag receiver against the MCR method and the conventional rake receiver. The average results of 1000 experiments were used to generate the results for the 16-QAM modulation. It demonstrates a slight improvement from the conventional rake receiver.

However, as the number of paths increase, the improvement decreases from 2.5dB to 0.6dB. For 3-paths, the results show that we are able to match the performance of the RAKE receiver. It is also important to understand that a lack of training data, training time, computational capacity, and different number of node and layers may influence the generalizability of these results.

V. Conclusion and Discussion

In this paper, we propose a hybrid technique that uses fuzzy logic and neural networks, called Tag Receiver. The Tag receiver achieves a significant error reduction in comparison to the conventional rake receiver. It must be noted, however, that our hybrid approach causes the fuzzy logic and neural network to be trained separately, resulting in a higher computational complexity.

Although the proposed Tag receiver has an improved BER, the effectiveness of receiver should be further explored in terms of flexibility and computation. The Tag receiver already performs extremely well based on this limited test set, with continuous real life data collection, we can assume its performance will further increase. As these methods are machine learning techniques, the algorithm is extremely dependent on the training set. As such, the neural networks and the fuzzy logic can be further trained on different modulation types.

One possible direction of future work is to investigate the effects of different modulated signals and using other types of neural networks. It may also be interesting to explore the use of back propagation in our proposed neural network.

Another possible future work is to integrate the FLT and the neural network processes into one system. Although individually optimized processing blocks is able to achieve significant results, it cannot be proven to achieve the optimal end-to-end performance. However, an integrated system is unlikely to have such a rigid modular structure as it would be trained for optimized end-to-end performance.

Acknowledgement

This work was supported in part by Thales Solutions Asia Pte. Ltd. and in part by the National Science Foundation of China under the grant no. NSFC 61750110529.

References

[1] K. C. H. Blom, H. S. Dol, A. B. J. Kokkeler, and G. J. M. Smit, "Blind equalization of underwater acoustic channels using implicit higher-order statistics," in *2016 IEEE Third Underwater Communications and Networking Conference (UComms)*, Aug 2016, pp. 1–5.

[2] P. A. Ranta, A. Hottinen, and Z. . Honkasalo, "Co-channel interference cancelling receiver for tdma mobile systems," in *Proceedings IEEE International Conference on Communications ICC '95*, vol. 1, June 1995, pp. 17–21 vol.1.

[3] D. Wang and J. Chen, "Supervised speech separation based on deep learning: An overview," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 10, pp. 1702–1726, Oct 2018.

[4] C. Demir, M. Saraçlar, and A. T. Cemgil, "Single-channel speech-music separation for robust asr with mixture models," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 4, pp. 725–736, April 2013.

[5] T. Huovinen and T. Ristaniemi, "Effect of channel estimation and multipath on interference cancellation employing blind source separation in the ds-cdma uplink," in *2004 IEEE 59th Vehicular Technology Conference. VTC 2004-Spring (IEEE Cat. No.04CH37514)*, vol. 3, May 2004, pp. 1494–1498 Vol.3.

[6] S. G. Surampudi and R. Pal, "Speech signal processing using neural networks mapping the phonology of sanskrit language using neural networks," in *2015 IEEE International Advance Computing Conference (IACC)*, June 2015.

[7] B. . Z. A. Asari, Hiroki & Pearlmutter, "Sparse representations for the cocktail party problem," *The Journal of neuroscience : the official journal of the Society for Neuroscience. 26. 7477-90. 10.1523/JNEUROSCI.1563-06.2006.*, 2006.

[8] A. A. Kumar, D. Zhang, Z. Lin, S. G. Razul, and C. M. S. See, "Direction finding and multipath mitigation using single antenna," in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2018, pp. 1–5.

[9] A. Pérez-Neira, M. A. Lagunas, A. Jové, and A. Artés, "A fuzzy logic filter for coherent detection in mobile communication receivers," in *9th European Signal Processing Conference (EUSIPCO 1998)*, Sept 1998.

[10] M. F. L. Abdullah and S. W. Bong, "Adaptive differential amplitude pulse-position modulation technique for optical wireless communication channels based on fuzzy logic," *IET Communications*, vol. 8, no. 4, pp. 427–432, March 2014.

[11] L. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338 – 353, 1965.

[12] M. G. Voskoglou, *Stochastic and Fuzzy Models in Mathematics Education, Artificial Intelligence and Management*. Lambert Academic Publishing, 2011.

[13] N. Farsad and A. J. Goldsmith, "Detection algorithms for communication systems using deep learning," *CoRR*, vol. abs/1705.08044, 2017. [Online]. Available: http://arxiv.org/abs/1705.08044

[14] D. Li, I. K. Sethi, N. Dimitrova, and T. McGee, "Classification of general audio data for content-based retrieval," *Pattern Recognition Letters*, vol. 22, no. 5, pp. 533 – 544, 2001, image/Video Indexing and Retrieval.

[15] S. Wang, D. Li, X. Song, Y. Wei, and H. Li, "A feature selection method based on improved fisher's discriminant ratio for text sentiment classification," *Expert Systems with Applications*, vol. 38, no. 7, pp. 8696 – 8702, 2011.

[16] K. Cho and D. Yoon, "On the general ber expression of one- and two-dimensional amplitude modulations," *IEEE Transactions on Communications*, vol. 50, no. 7, pp. 1074–1080, July 2002.

[17] M. K. Simon and M. Alouini, "A unified approach to the performance analysis of digital communication over generalized fading channels," *Proceedings of the IEEE*, vol. 86, no. 9, pp. 1860–1877, Sept 1998.