EC 601
Mini project 2 Report
Zhou Shen
“Speech Processing” - Neural Network use in DSP

Introduction to the topic

First, below is the definition of speech processing from Wikipedia: “Speech processing is the study of speech signals and the processing methods of signals. The signals are usually processed in a digital representation, so speech processing can be regarded as a special case of digital signal processing, applied to speech signals.” [1]

Second, definition of Digital Signal Processing is below: “Digital signal processing (DSP) is the use of digital processing, such as by computers or more specialized digital signal processors, to perform a wide variety of signal processing operations. The signals processed in this manner are a sequence of numbers that represent samples of a continuous variable in a domain such as time, space, or frequency. Digital signal processing and analog signal processing are subfields of signal processing. DSP applications include audio and speech processing, sonar, radar and other sensor array processing, spectral density estimation, statistical signal processing, digital image processing, signal processing for telecommunications, control systems, biomedical engineering, seismology, among others. DSP can involve linear or nonlinear operations. Nonlinear signal processing is closely related to nonlinear system identification[1] and can be implemented in the time, frequency, and spatio-temporal domains.” [2]

Third, definition of Kalman Filter is below: “In statistics and control theory, Kalman filtering, also known as linear quadratic estimation (LQE), is an algorithm that uses a series of measurements observed over time, containing statistical noise and other inaccuracies, and produces estimates of unknown variables that tend to be more accurate than those based on a single measurement alone, by estimating a joint probability distribution over the variables for each timeframe.” [3]

The state-of-art paper is about a neural network algorithm for real-time signal processing. This paper consists of several parts: a general overview of the network, performance in worst test cases, discussion for several capabilities.

The author of the paper introduces us a neural network algorithm called STOCHASM, which was developed for the purpose of real-time signal detection and classification, basically is the most important processes of digital processing. Its main purpose is dealing with transient signals having low signal-to-noise ratios (SNR). This algorithm was first developed in 1986 for real-time fault detection and diagnosis of malfunctions in ship gas turbine propulsion systems (Malkoff, 1987). It subsequently was adapted for passive sonar signal detection and classification. Currently, its newest version for information fusion and radar classification have been developed. The performance of the algorithm is impressive. In the worst possible case of three sets of totally overlapping double chirp signals, it recognized target correctly within 95% correct. [4]

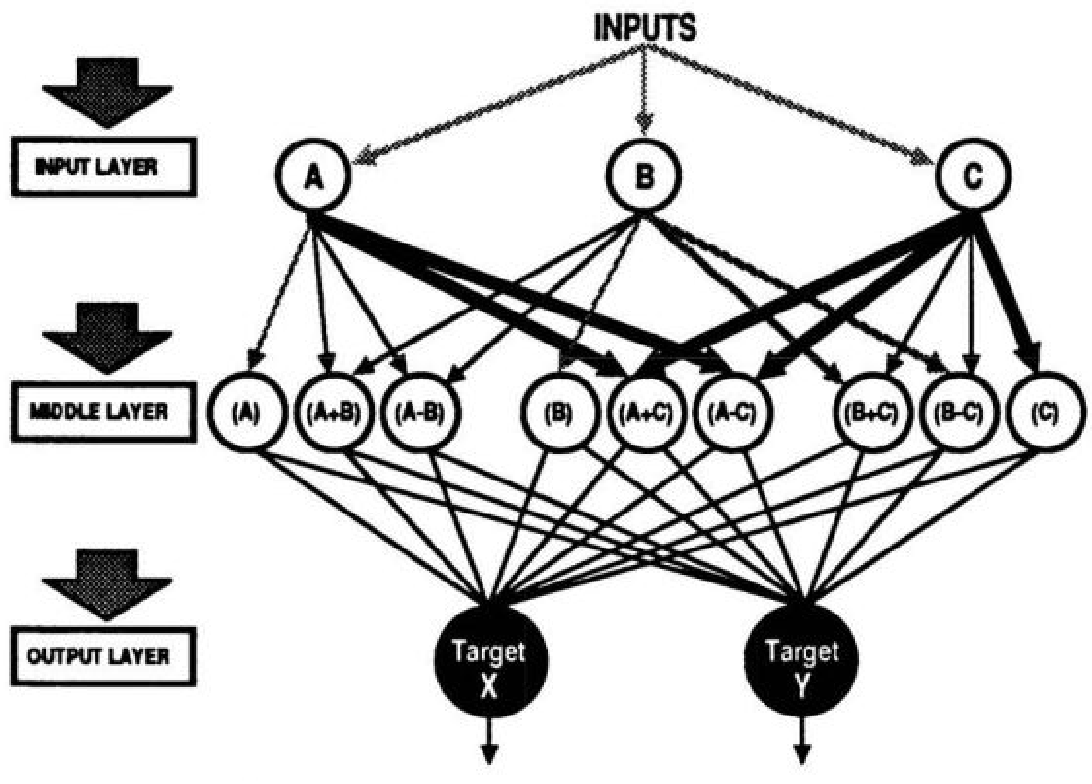

Figure 1. STOCHASM network architecture

Summary of references

1. “Speech Processing.” Wikipedia. Wikimedia Foundation, February 13, 2019. https://en.wikipedia.org/wiki/Speech_processing.
2. “Digital Signal Processing.” Wikipedia. Wikimedia Foundation, October 13, 2019. https://en.wikipedia.org/wiki/Digital_signal_processing.
3. “Kalman Filter.” Wikipedia. Wikimedia Foundation, September 28, 2019. https://en.wikipedia.org/wiki/Kalman_filter.

4. Malkoff , Donald B. “A Neural Network for Real-Time Signal Processing .” https://papers.nips.cc/paper/284-a-neural-network-for-real-time-signal-processing.pdf. Moorestown Corporate Center , n.d. Accessed October 14, 2019.

Analysis of results

According to the content of the paper we just discussed, we can obviously see the huge advantage of using this neural network algorithm during speech digital signal processing. Below are the advantages and disadvantages I can think of.

Pros:

1. This neural network algorithm has the ability to deal with transient signals having low signal-to-noise ratios (SNR), even the noise can be classified.
2. It performs well in the presence of either Gaussian or non-Gaussian noise,

even where the noise characteristics are changing.

3. Improved classifications result from temporal pattern matching in real-time,

and by taking advantage of input data context dependencies.

4. The network is trained on-line. Single exposures of target data require one

pass through the network. Target templates, once formed, can be updated

On-line.

5. The algorithm is implemented in parallel code on a Connection Machine. Simulated signals embedded in noise and subject to considerable uncertainty, are classified within 500 ms of onset.”

Cons:

1. Current worst case is only 3 digital signal resources, may try more resources for detecting.
2. May improve the temporal pattern matching capabilities.

Recommendations for a person who wants to develop or use such systems

Person who needs this algorithm to detect the correct signal from multiple signals like background noises.

Group who would like to develop this algorithm to find target voice signal from chaos environment.

Conclusions

Speech processing is currently using in more frequently in daily life as technology grows quickly. Speaking in noisy places becomes popular, therefore, handling quickly for multiple digital signals with stable correct rates is the key for developers. Within this state-of-art technology introduced in paper, we can find target signal in a short time with high correct rate.