Applied & Computational Mathematics Seminar

Real-time wind noise detection and suppression with neural-based signal reconstruction for multi-channel low-power devices

Anthony D. Rhodes

1/22/18

Portland State University

Outline

- Research goals & motives
- Digital signal processing overview
- Wind noise detection: *RTWD* algorithm (*real-time wind noise detection*)
- Wind noise suppression: *ANWS* algorithm (*attentive neural wind suppression*)
- Results/Conclusion

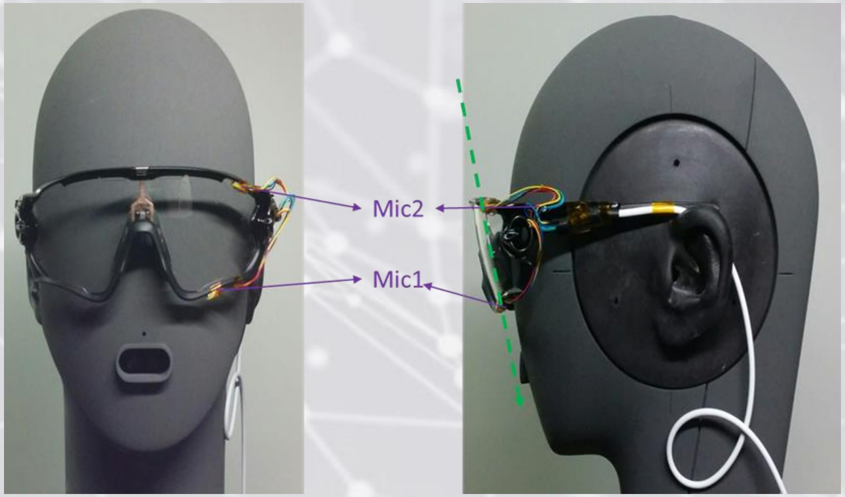
Research Goals & Motives

Goals:

(*) Develop robust, generalizable algorithms capable of real-time deployment of the following tasks:

(i) Accurate wind/noise detection

(ii) Wind noise suppression / accurate signal reconstruction for ASR (automatic speech recognition)



Challenges/Novelties:

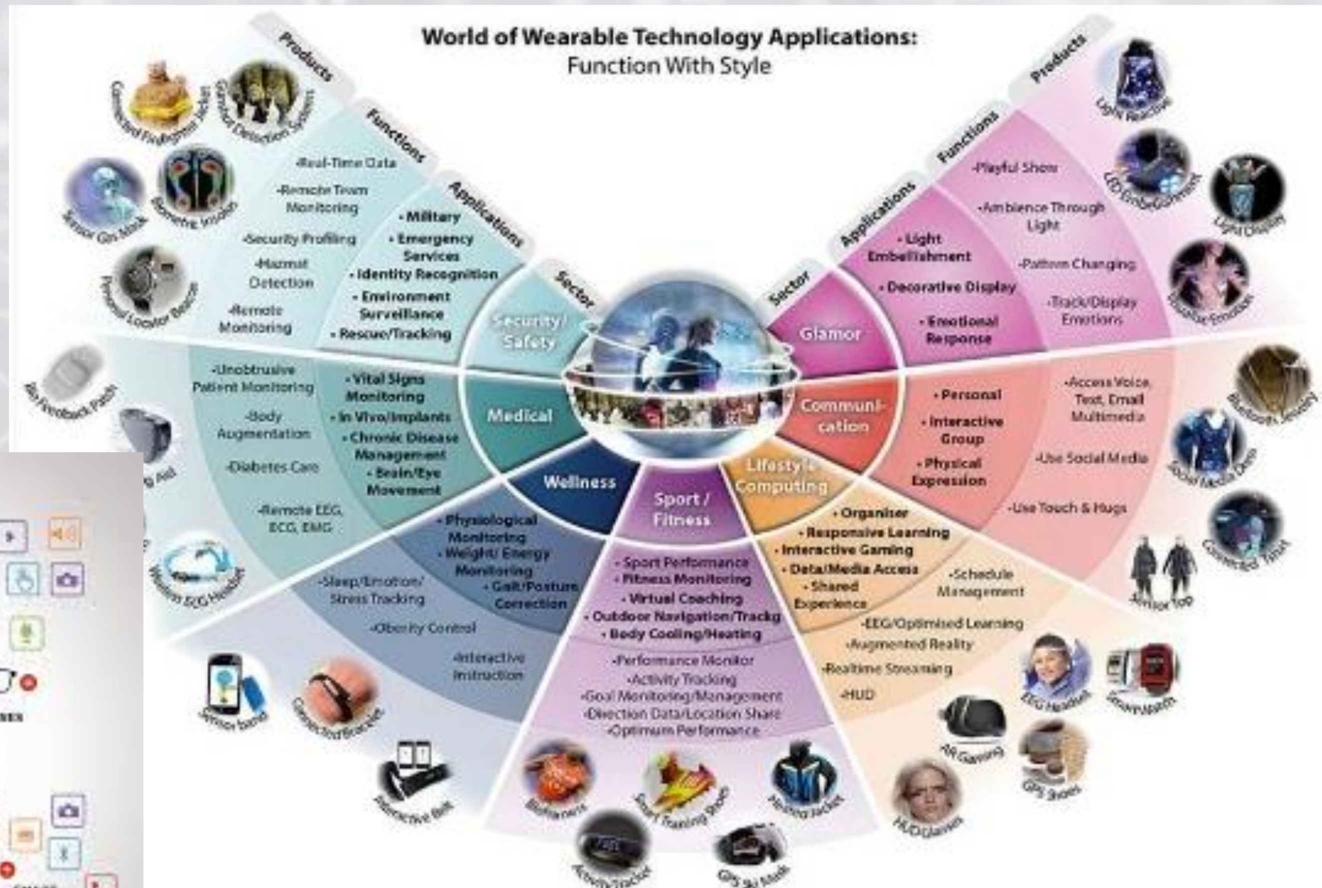
(*) Low-power regime

(*) Multi-microphone device

(*) Solving noise detection/suppression crucial for improving potential IoT (internet of things) connectivity.

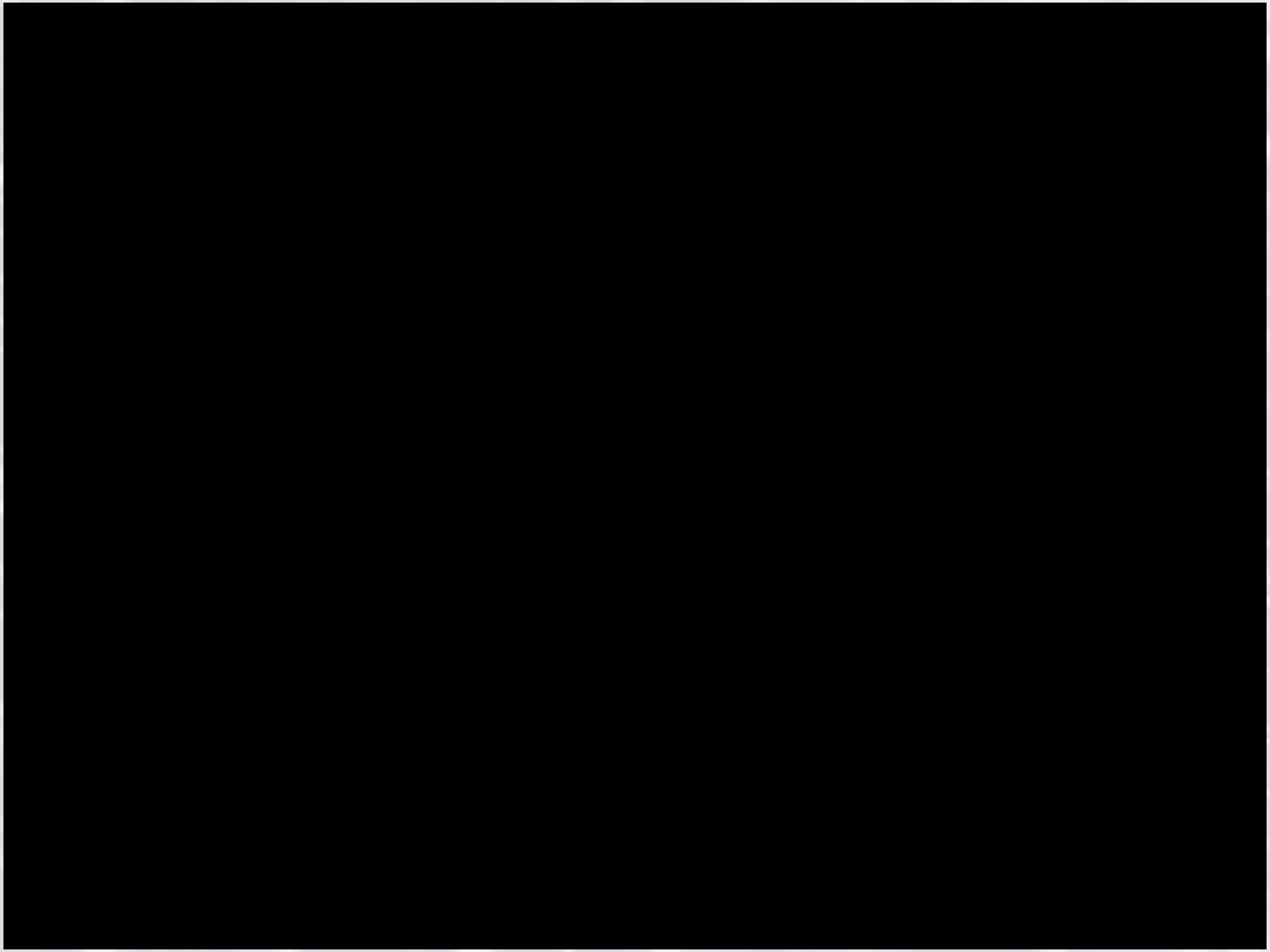
Research Goals & Motives

<u>Wearable Devices & the IOT</u>



Research Goals & Motives

Wearable Devices & the IOT

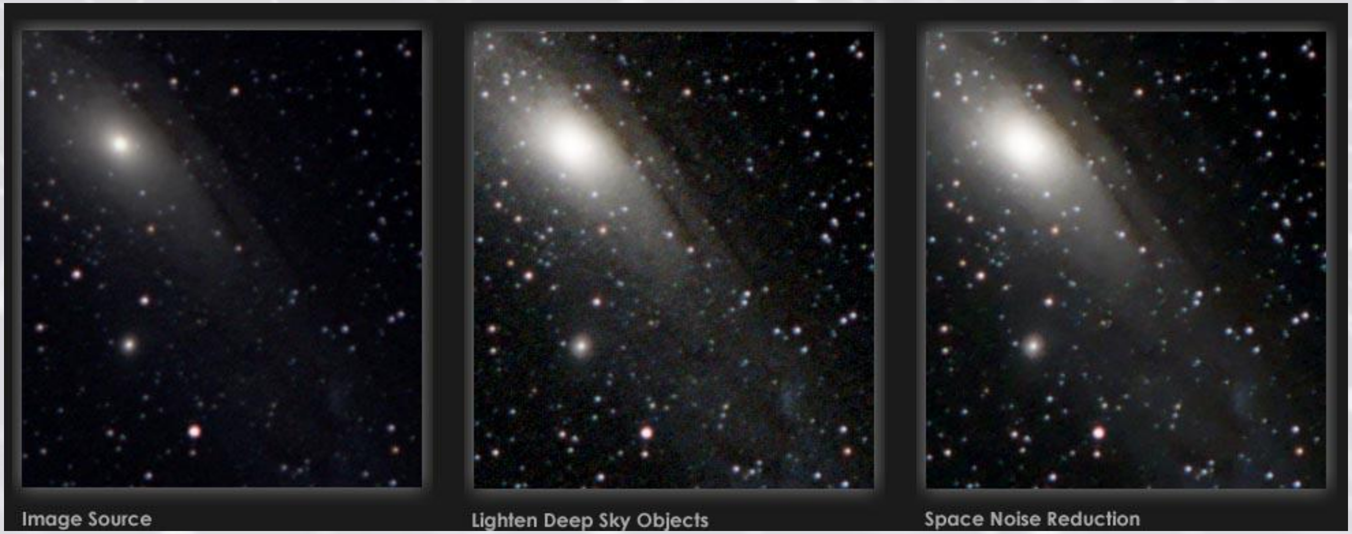


Noise: The Bane of Signal Processing

(*) "Noise" is the unwanted modification of a signal suffered during capture, storage, transmission, processing, or conversion.

(*) Noise reduces (or potentially eliminates) the presence of useful information in a signal.

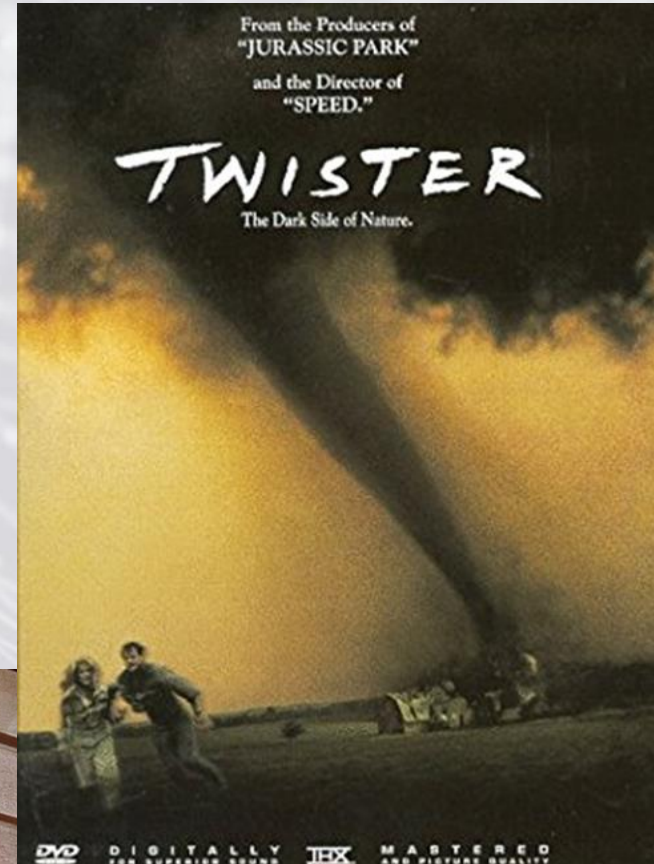
(*) Noise reduction/suppression is the process of recovering the original signal from the noise corrupted one – the most common goal in the design of signal processing systems (e.g. filters).



Wind Noise: The Bane of Audio Signal Processing

(*) Because wind noise is a predominant source of audio interference, it creates a common – albeit challenging – setting for voice-driven applications for wearable devices, including ASR (automatic speech recognition).

The Disastrous Effects of Wind



The Disastrous Effects of Wind

*Spanish Armada Defeated July 29, 1588 (it was a very windy day)

The Disastrous Effects of Wind

Digital Signal Processing

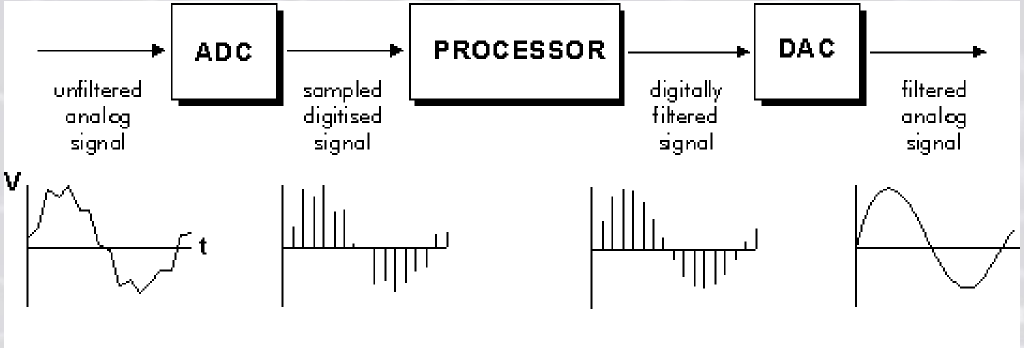
(*) Digital Signal Processing (DSP) is the mathematical manipulation of the numerical representation of a digital signal (for some advantageous end).

(*) A basic DSP system is composed of:

(1) An ADC providing digital samples of an analog input

(2) A digital processing system

(3) A DAC converting processed sample to analog output



What's a signal?

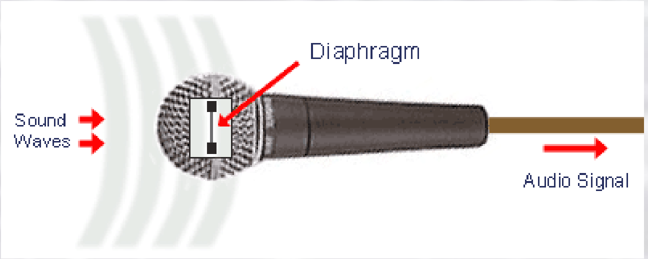
(*) An analog signal is defined as any representation of a physical quantity that varies over time, has a value at all instants and contains information (e.g., temperature, electrical voltage, human voice, light intensity).

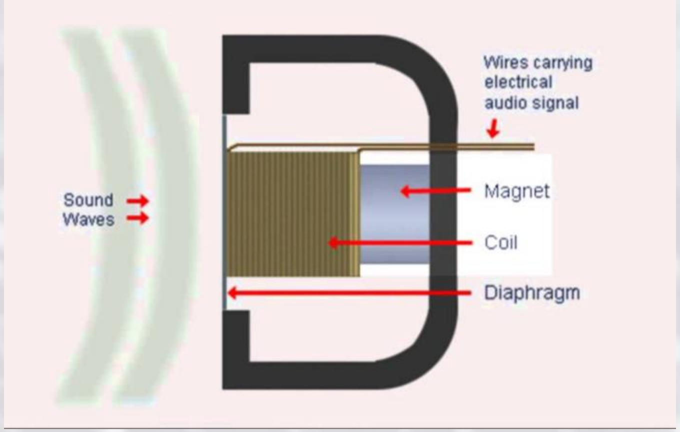
What's the role of the processor?

(*) Typically, to compress the signal and/or reduce presence of noise.

Digital Signal Processing

Consider a microphone as a basic ADC:





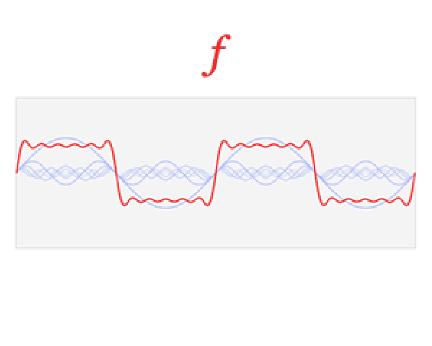
(*) Inside the microphone, the diaphragm vibrates as it receives sound waves.

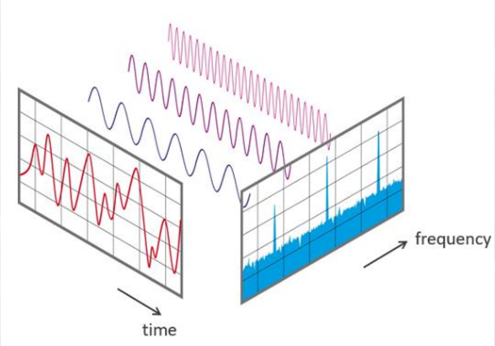
(*) The coil, attached to the diaphragm oscillates back and forth.

(*) The coil moves back and forth through the magnetic field produced by the magnet; as a result an electric current flows through it.

Digital Signal Processing

(*) DSP typically involves a Fourier Transform (e.g. DFT).





(*) The Fourier transform measures whether a frequency is present in a particular function.

(*) Computing the Fourier transform for a function f, resolves f into a Fourier series: a linear combination of sines and cosines.

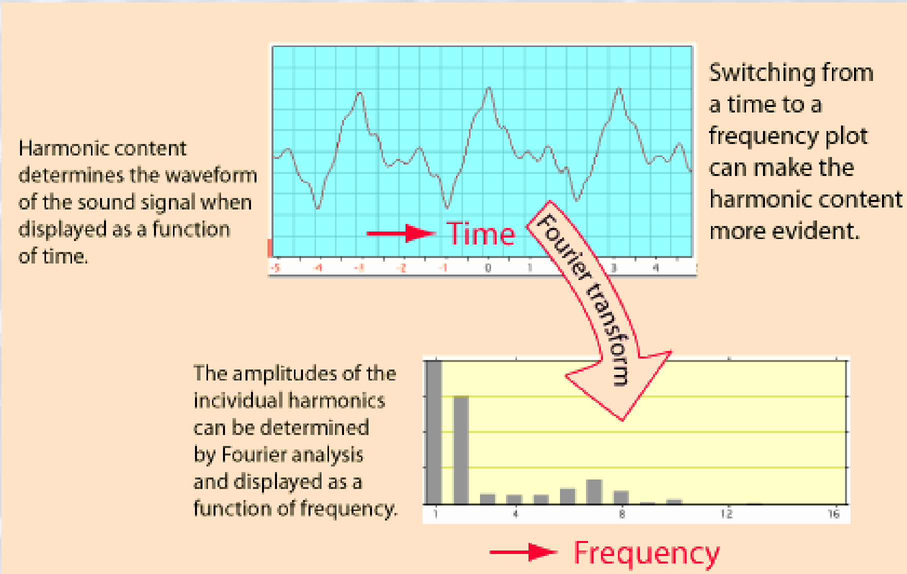
(*) The component frequencies of these sines and cosines spread across the frequency spectrum, are represented as Dirac delta functions in the frequency domain.

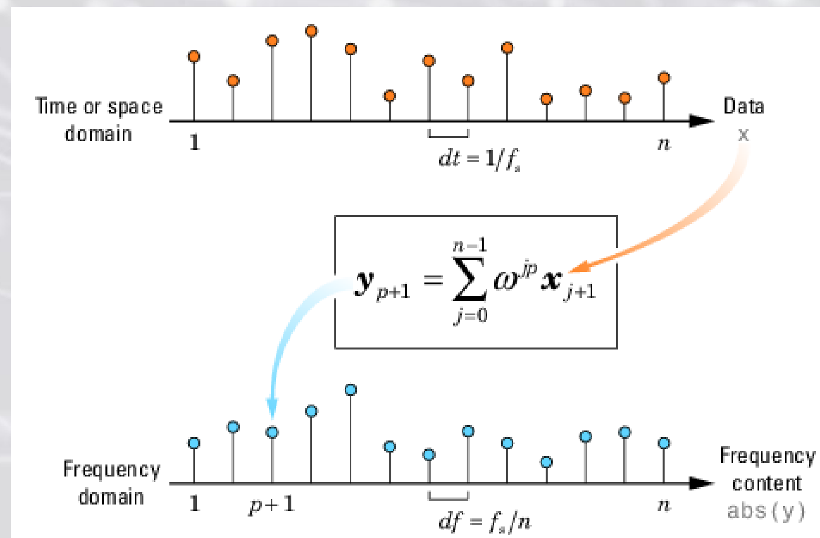
(*) The frequency domain representation $\hat{f}$, is the collection of these peaks at the frequencies that appear in the resolution of the function.

Famously, Tukey & Cooley published a general version of the FFT (fast Fourier Transform) requiring O(n log n) run-time vs. O(n^2).

Gilbert Strang on the FFT: "the most important numerical algorithm of our lifetime."

Digital Signal Processing





Real-Time Wind Detection

(*) Many commercial devices in use today rely heavily on passive solutions to mitigating wind noise (e.g. physical dampening devices).

(*) While these techniques can provide simple, approximate solutions to wind noise reduction, their effectiveness can nevertheless be limited even in moderate wind conditions.

(*) We believe instead that more active (i.e. software-driven) approaches can also be leveraged, in addition, to achieve state-of-the-art wind noise suppression for wearable devices.

Real-Time Wind Detection

(*)To this end we develop robust, software-driven wind noise detection and suppression algorithms operational in low-energy, multiple microphone regimes.

(*) Limitations in computational and memory resources provide a significant challenge for noise detection and signal reconstruction tasks with wearable and smart devices.

(*)Because ASR systems are commonly highly sensitive to the presence of interfering noise, we also require our noise suppression system to be both reliable in moderate and even low wind noise regimes and to furthermore minimize the introduction of ectopic, reconstructed signal distortions.

Related Research/Previous Work

(*) Previous research in active noise detection and related tasks in audio signal processing has chiefly relied on identifying *a priori* (or conversely: by learning) discriminative features that indicate the presence of interfering noise.

- Nelke et al. [19], for instance, use short-term mean (STM) features in the time-domain as the basis for a low-dimensional wind indicator.
- Relying on the assumption that the magnitude of the spectrum of wind noise can be roughly approximated by a linear decay over the frequency, [23] proposes learning a negative slope fit (NSF) model for wind classification.
- Freeman et al. [7] train a neural network to build a general noise classification system; see also: [32], [21], [25], [38], [28], [23].

(*) In each case, these various approaches violate either the low computational limitations or desired ASR sensitivity threshold for our consumer applications, and/or failed to make genuine use of multi-channel signals.

Related Research

(*) In general, signal reconstruction and noise reduction tasks typically necessitate even more computational and memory resources than detection and classification tasks.

- Popular examples include full spectrum neural "denoising" approaches [16], [2], non-negative sparse coding (NNSC) [30], [26], and subspace-based methods [17], [4], [3], [8].

$$X = Xs + Xn \approx [D_s\ Dn]\begin{bmatrix} H_s \\ H_n \end{bmatrix} = DH$$

$$A = \sum_{k=1}^{m} \sigma_k u_k v_k^T$$

- Attempts to "sparsify" signal reconstruction systems to reduce their computational and memory requirements often come at a significant performance cost. While effective against point-wise interference sources, we found, for example, adaptive beamforming (particularly the MVDR and GSC algorithms, see: [11], [29], [35], [15]) approaches to be largely unsuccessful for clean signal reconstruction in the case of diffuse wind, or when the interference signal vector strongly aligned with the source signal.

$$min_w\ w^H R_n w \ \text{ s.t. } \ w^H \mathbf{a}(\theta)=1$$

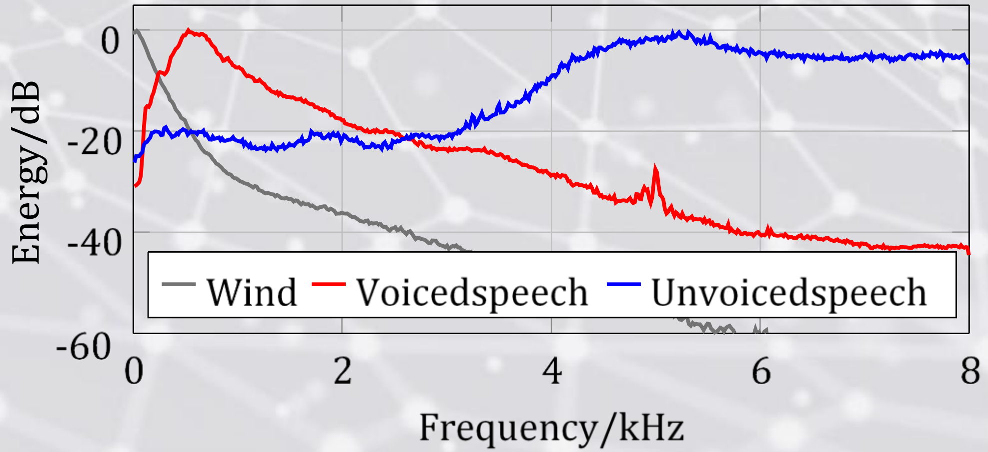
$$\hat{X}_k(t,f) = Y_k(t,f) - \hat{N}_k(t,f)$$

- Similarly, spectral subtraction ([33], [34]) and various filtration procedures ([6], [25]) commonly fail for ASR-based signal reconstruction tasks due to the non-stationarity of wind noise.

Real-Time Wind Detection

- A sufficiently precise detection of wind noise is the first step towards suppression of noise in captured signals [1].
- We seek discriminative (preferably low-dimensional) features that can be used to accurately determine the presence of wind.
- Features for wind detection commonly rely on short-term statistics.

- In particular, the spectral energy distribution for very low frequencies (< 10 Hz) for wind is discernable from that of speech.



Real-Time Wind Detection (SSC)

- We first consider sub-band centroids (SSC) features for wind classification [32].
- Samples are captured from wearer voice and segmented into several frames and frequency analysis is performed via FFT.

- Define the spectral centroid for time frame λ with respect to the bin range [μ1, μ2]:

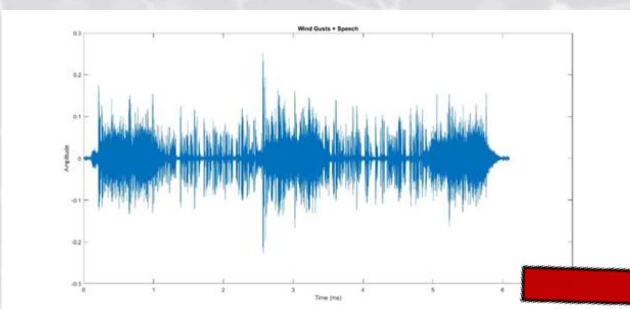
$$\Xi_{\mu 1,\ \mu 2}(\lambda) = \frac{\sum_{\mu=\mu 1}^{\mu 2} |X(\lambda,\mu)|^2 \cdot \mu}{\sum_{\mu=\mu 1}^{\mu 2} |X(\lambda,\mu)|^2}$$

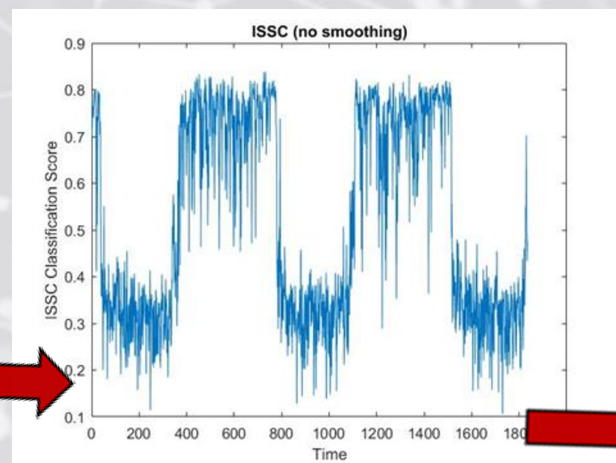
- In order to detect wind, we consider the sub-band range: [0,10], as in [3].
- Define the SSC-based wind indicator:

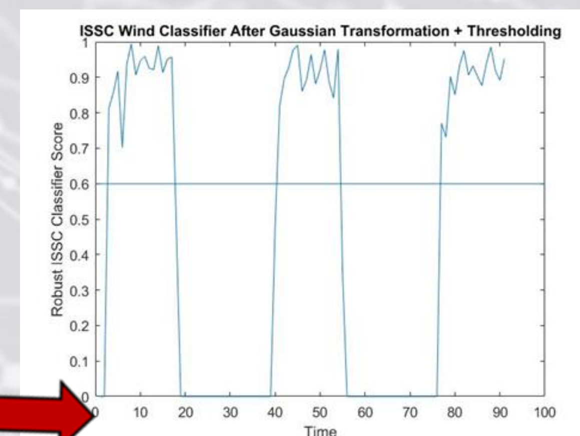
$$I_{SSC}(\lambda) = \frac{\mu 2 - \Xi_{\mu 1,\ \mu 2}(\lambda)}{\mu 2} \in [0,1]$$

Real-Time Wind Detection (SSC)

- Because of the low-dimensional spectral representation used for the SSC method, the wind indicator function tends to be very noisy and frequently unstable.
- To generate a more robust model, we apply a smoothing procedure (500ms windows), followed by an inverse Gaussian transformation of the ISSC function with graceful thresholding for robust wind classification.







Speech + Wind Signal

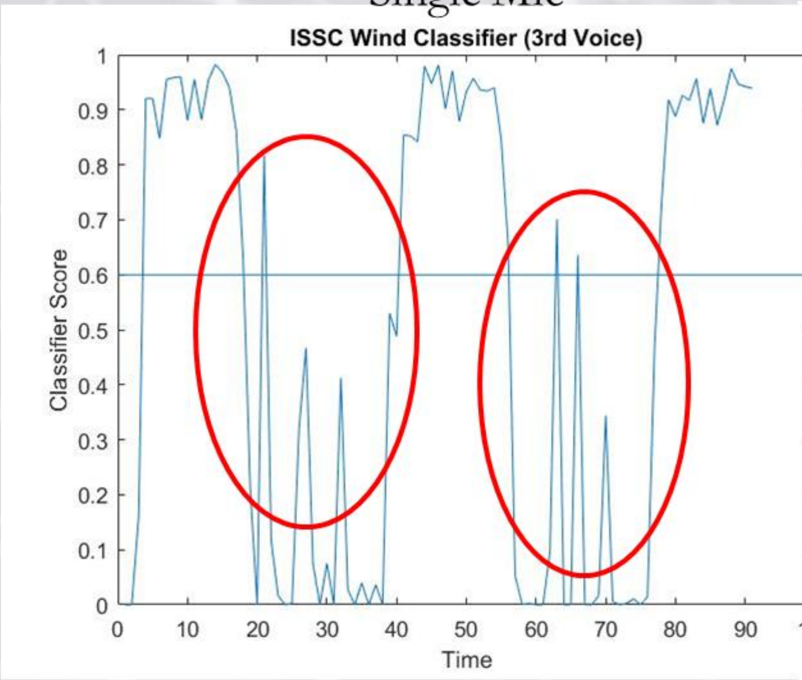
ISSC (no smoothing)

ISSC Wind Classifier

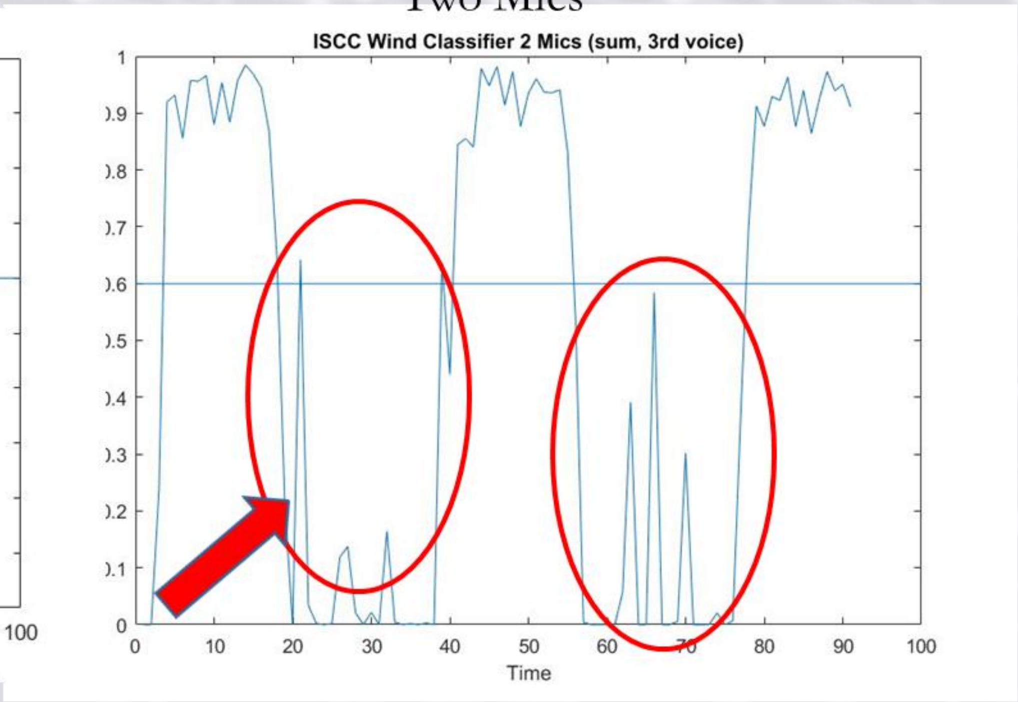
Real-Time Wind Detection (SSC)

- To improve SSC-based wind classification for multi-channel audio, we additionally apply a max operation to promote robustness in the case of the non-stationarity of wind noise and to safeguard against microphone occlusion in head-worn devices.

Single Mic



Two Mics



Classification
error reduction

Real-Time Wind Detection (SSC)

- By themselves, we found that transformed SSC features can be used to accurately detect the presence of wind for wearable devices in the case of moderate to strong wind (15 mph+).

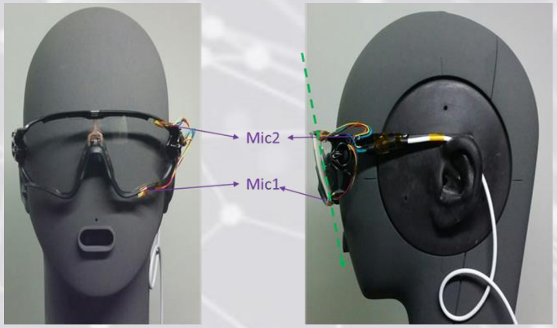
- However, this method alone renders a large quantity of false-positive results for low wind speed regimes (<= 10mph), which can be a critical range for ASR applications.

- To reduce this sensitivity and thereby improve classification in low wind intensity scenarios by <u>decreasing instances of false-positive readings</u>, we additionally incorporate coherence-based features into our algorithm.

Real-Time Wind Detection (MC)

- More specifically, we average the magnitude of the coherence (MC) for the current frame of captured audio.

- Values close to one indicate the presence of a strong power "transfer" between the two channels, whereas values close to zero show a weak power transfer.
- For example, wind alone should yield a small MC value, whereas speech alone produces a large MC value.



- Finally, we tune the classification algorithm so that when both wind and speech are present simultaneously, wind detection "overwhelms" the presence of speech. Together, we gracefully threshold the SSC and coherence features to achieve high accuracy for wind detection across a broad spectrum of wind intensities.

Real-Time Wind Detection (MC)

- Define 2-channel coherence as the ratio of the cross power spectral density (CPSD) and auto power spectral densities (APSDs):

$$\Gamma(\lambda, \mu) = \frac{\phi_{x_1 x_2}(\lambda, \mu)}{\sqrt{\phi_{x_1 x_1}(\lambda, \mu)\phi_{x_2 x_2}(\lambda, \mu)}}$$

- Where the power spectral densities (PSDs) are estimated by the recursive smoothed periodgram [9]:

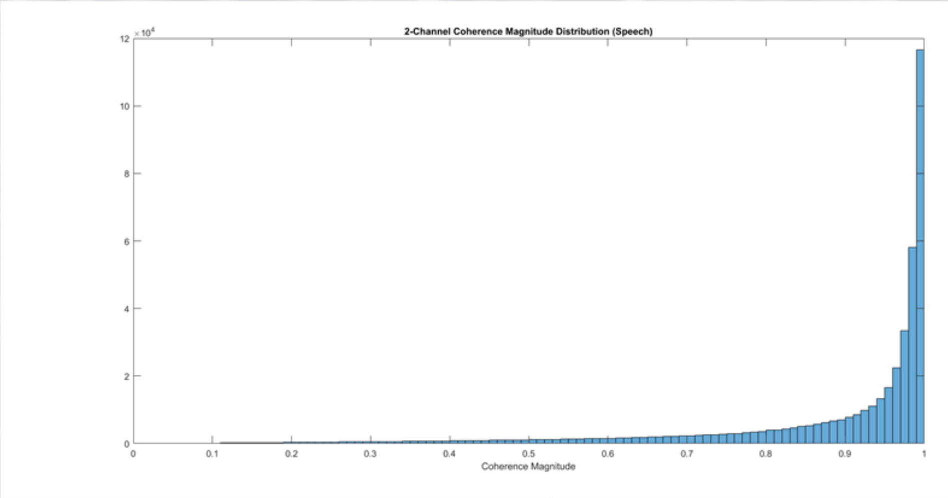
$$\phi_{x_i x_j}(\lambda, \mu) = \alpha_s \phi_{x_i x_j}(\lambda - 1, \mu) + (1 - \alpha_s) X_i(\lambda, \mu)\, X_j{}^H(\lambda, \mu)$$

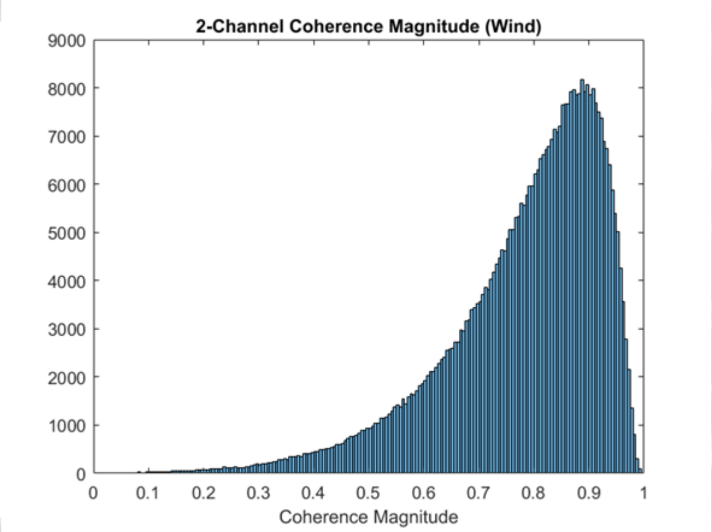
- Here α is a smoothing constant set heuristically (α=0.8) and *X* represents the short time spectrum of the signal.

Real-Time Wind Detection (MC)

- [25] Showed that the magnitude of coherence can be used to discriminate between speech and noise. To this end, from the 2-channel coherence, define the magnitude of coherence:

$$MC(\lambda, \mu) = |\Gamma(\lambda, \mu)|$$



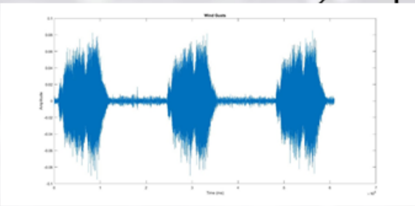


2-Channel Coherence Magnitude Histogram (Speech)

2-Channel Coherence Magnitude Histogram (Wind)

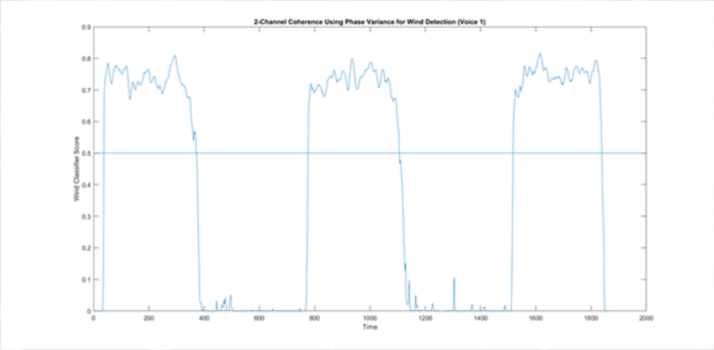
Real-Time Wind Detection (MC)

- We test 2-channel coherence features (i.e. MC, phase) for wind detection using 4 distinct voices; experiments yielded (near) perfect classification.

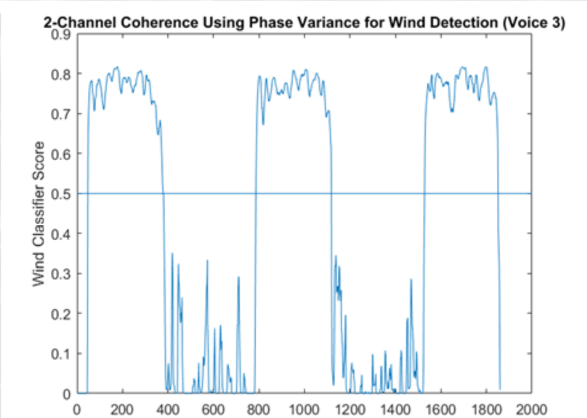


Wind (ground-truth)

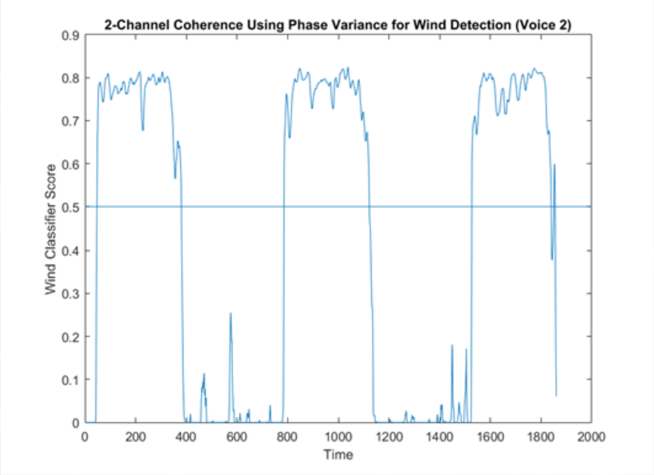
Voice 1



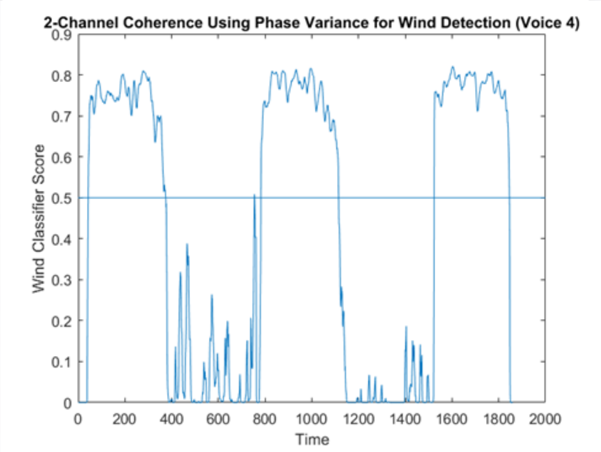
Voice 3



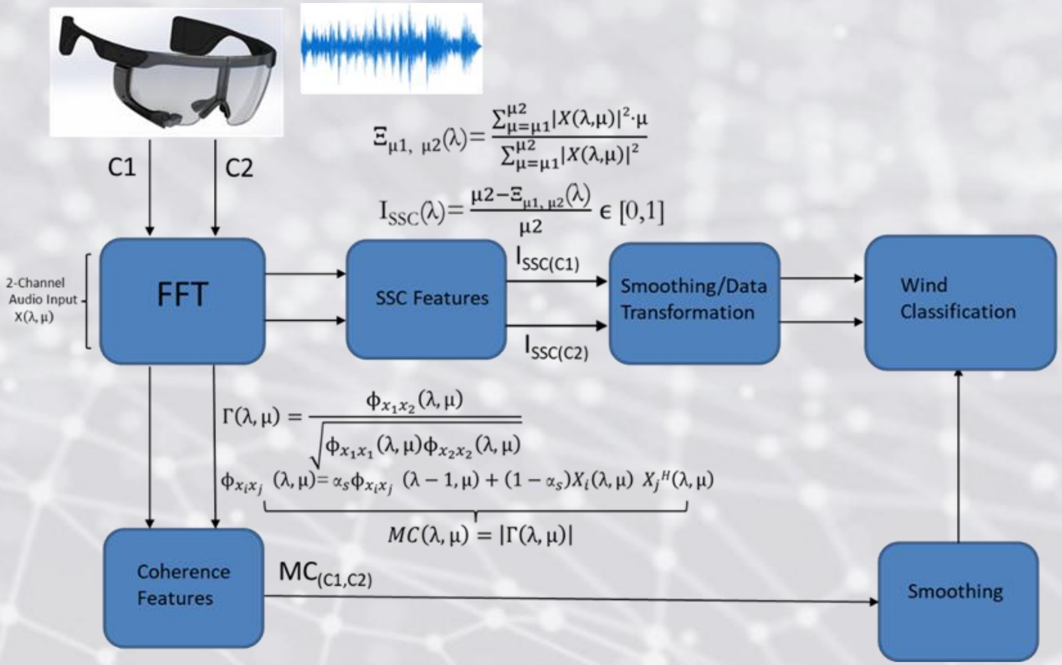
Voice 2



Voice 4



Real-Time Wind Detection



(*) Schematic of the **RTWD algorithm** in full.

In summary: Following the FFT step,

(1) SSC-based wind indicator values are computed for each channel, a windowed smoothing procedure (500ms) followed by an inverse Gaussian transformation is performed and subsequently a max operation is applied across the 2-channel signal.

(2) Concurrently, we compute the 2-channel coherence features and then determine the average MC value for the given time frame; we apply smoothing for robustness.

(3) Binary wind classification is finally determined based on a <u>tunable, conjunctive thresholding using the transformed SSC and coherence-based features together</u> (e.g., when both feature values meet specific thresholding criteria, the signal is classified as containing wind).

Attentive Neural Wind Suppression (ANWS)

- We devise a novel wind suppression algorithm, ANWS, for use with low-computation, multiple-microphone devices.

- Recently, [33], [28] have demonstrated the promise of applying deep neural networks (DNNs) to the task of clean audio signal reconstruction. However, due to their computational demands and extensive training data requirements, these approaches have heretofore rarely been applied successfully to low-power devices.

- To circumvent these issues, we train a relatively low-dimensional, shallow neural network to reconstruct the wearer speech signal from wind-corrupted audio specifically in the spectral regions that are most adversely affected by wind noise; see: [22].

- In this way, we say that the neural-based signal reconstruction is a parsimonious process that *attends* to the regions of greatest need for signal reconstruction.

Attentive Neural Wind Suppression (ANWS)

- This *attentive spectral region* identification can feasibly be accomplished in one of two ways:

(1) We apply prior knowledge about the spectrum of the noise class that has corrupted our signal.

(2) We use an *a posteriori* learning approach, where a noise approximation is first made (in combination with a classification/detection algorithm), and then relevant corrupted frequency bins are identified according to a separate feature/spectral analysis.

Attentive Neural Wind Suppression (ANWS)

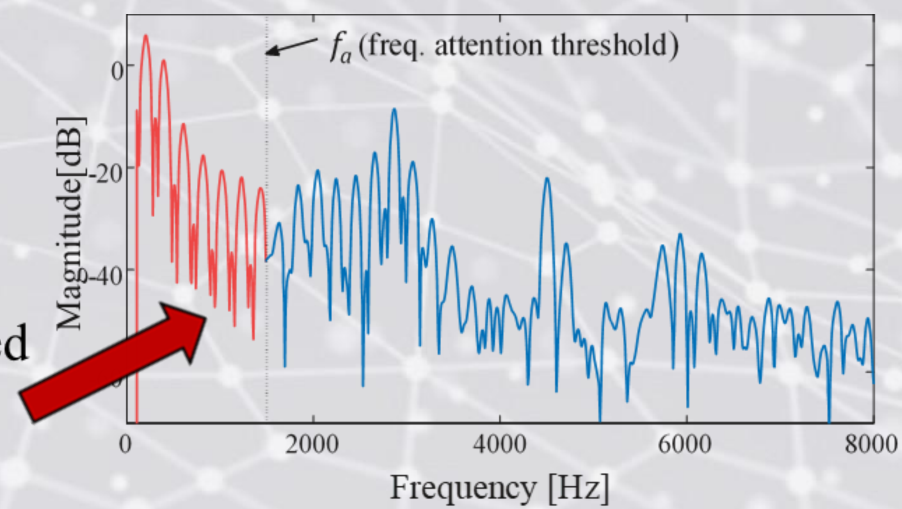
- This approach bears <u>several distinct advantages </u>for the noise reduction task:

(1) The model can be learned with a relatively small amount of data.

(2) The data representation is low-dimensional.

(3) Generally, the speech signal remains largely undistorted by the reconstruction process.

Reconstructed Signal



Attentive Neural Wind Suppression (ANWS)

- We develop a shallow, low-dimensional, feed-forward NN for wind noise suppression.
- The input to the network consists of context-expanded frames (see below) of the noisy signal. As in [12], [38], we use the log-power spectra features of a noisy utterance n^u for the short-time Fourier transform.

$$N(t, f) = log|STFT(n^u)^2|$$

- Let n_t be the t^{th} frame of $N(t, f)$. We express the multi-channel, context-expanded input vector to the NN as:

$$y_t = \left[n_{t-r}^{(1)}, \dots n_t^{(1)}, \dots, n_{t+r}^{(1)}, n_{t-r}^{(2)}, \dots, n_{t+r}^{(2)}\right]$$

- Where the parameter r represents the "context-horizon" and the superscripts here indicate the channel identification.

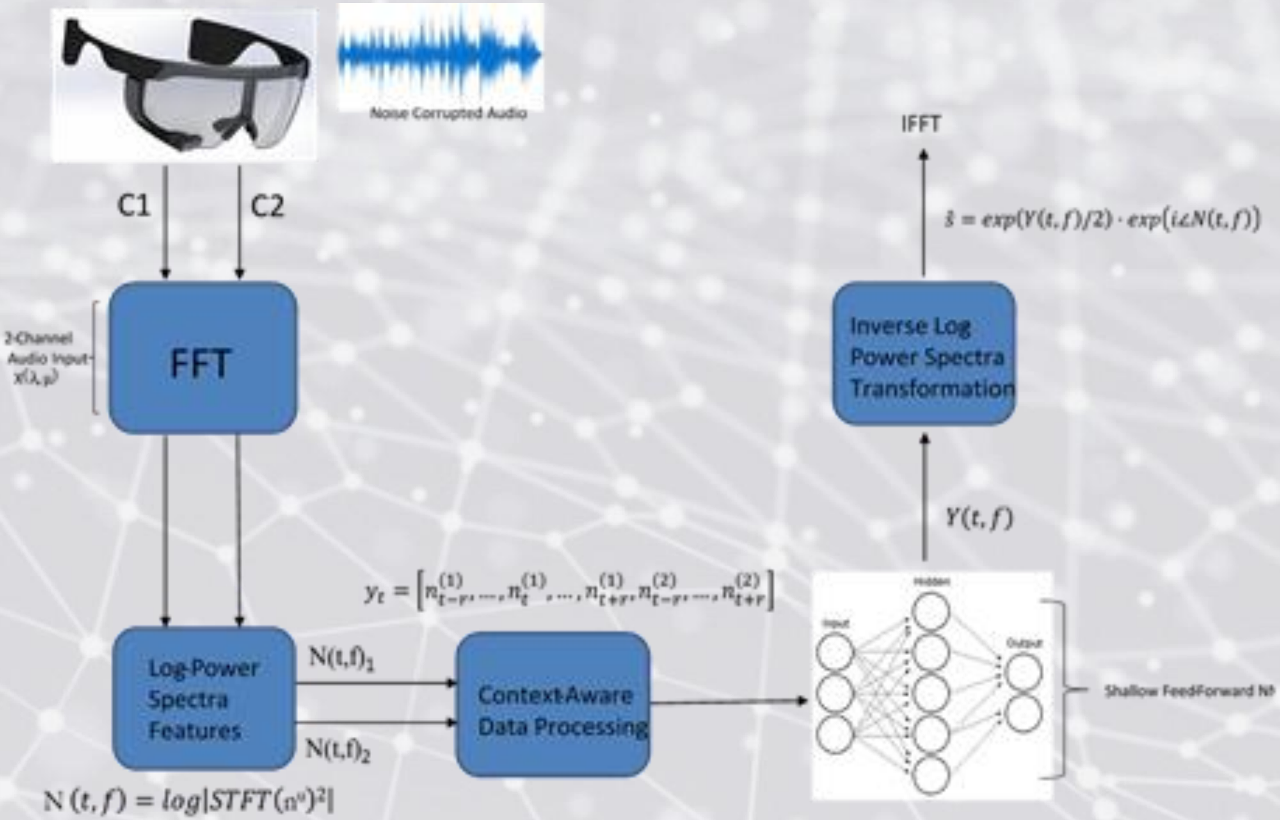
Attentive Neural Wind Suppression (ANWS)

$$y_t = \left[n_{t-r}^{(1)}, \ldots n_t^{(1)}, \ldots, n_{t+r}^{(1)}, n_{t-r}^{(2)}, \ldots, n_{t+r}^{(2)}\right]$$

- Using $r = 3$, we train a shallow NN with 150 hidden nodes, using conjugate gradient backpropagation on only 5 minutes of noisy speech and clean audio sample pairings for training.
- Note that noise-aware NN training [28] and larger microphone vector configurations are straightforwardly accommodated by the ANWS algorithm.
- The reconstructed signal $\hat{s}$ is obtained by applying the following "inverse" operation sequence to the output of the NN, represented by $Y(t,f)$:

$$\hat{s} = exp\big(Y(t,f)\big) \cdot exp\big(i\angle N(t,f)\big)$$

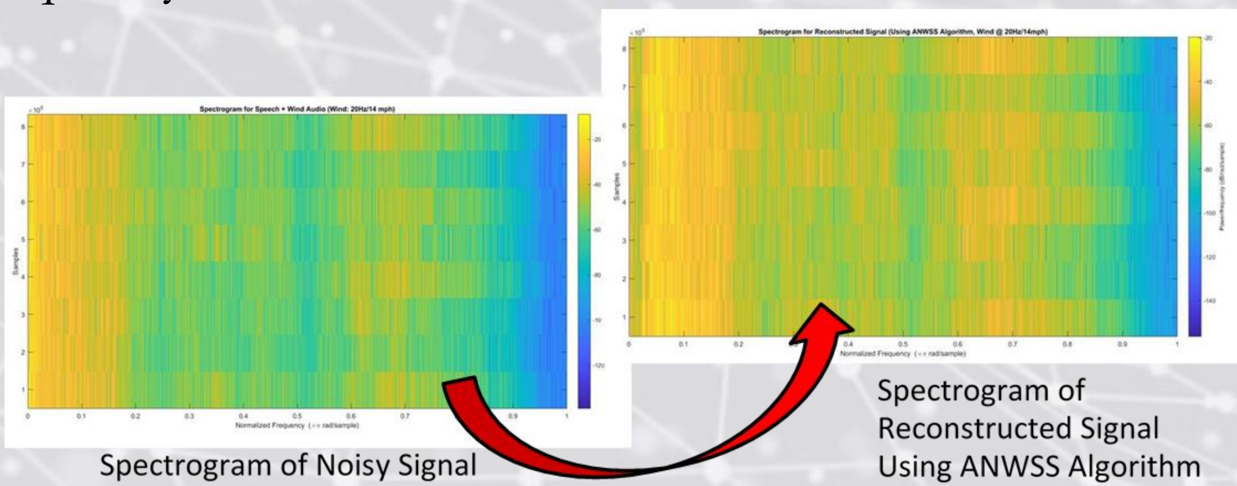
Attentive Neural Wind Suppression (ANWS)



(*) Schematic of the **ANWS algorithm** in full.

Attentive Neural Wind Suppression (ANWS)

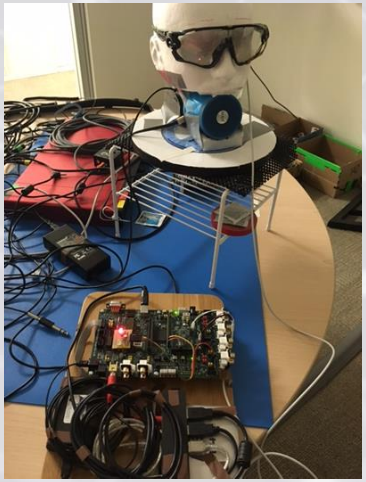
- The effectiveness of the ANWS is further illustrated by a spectrogram analysis for both noisy and subsequently reconstructed signals.
- In the figure below, the spectrogram of a wind corrupted signal is shown to be strongly dominated by extreme low frequencies (i.e. wind noise), whereas the corresponding reconstructed signal displays a more uniform frequency distribution.



- Spectrogram analysis for noisy signal versus reconstructed signal. Here the horizontal axis represents normalized frequency and the vertical axis represents time (equivalently: "samples"). <u>Yellow colors indicate frequency content with higher power; blue indicates low power</u>.

Experimental Results

- We tested our wind noise suppression system, including RTWD and ANWS algorithms, in real-time, under difficult, low-power conditions using a high-end wind simulator.
- We ported our algorithms to a Cirrus DSP (5.5 MIPS); for FFT we used 200ms audio "chunks", with 25 frames per chunk, comprising 16ms frames and 8ms overlap. Our smart glass device was affixed with a light, windscreen foam, so that our test conditions reflected the capabilities of a commercial-ready device.
- We used a competitive, proprietary ASR algorithm for measuring WER (word error rate) as an evaluative metric for wind noise suppression.



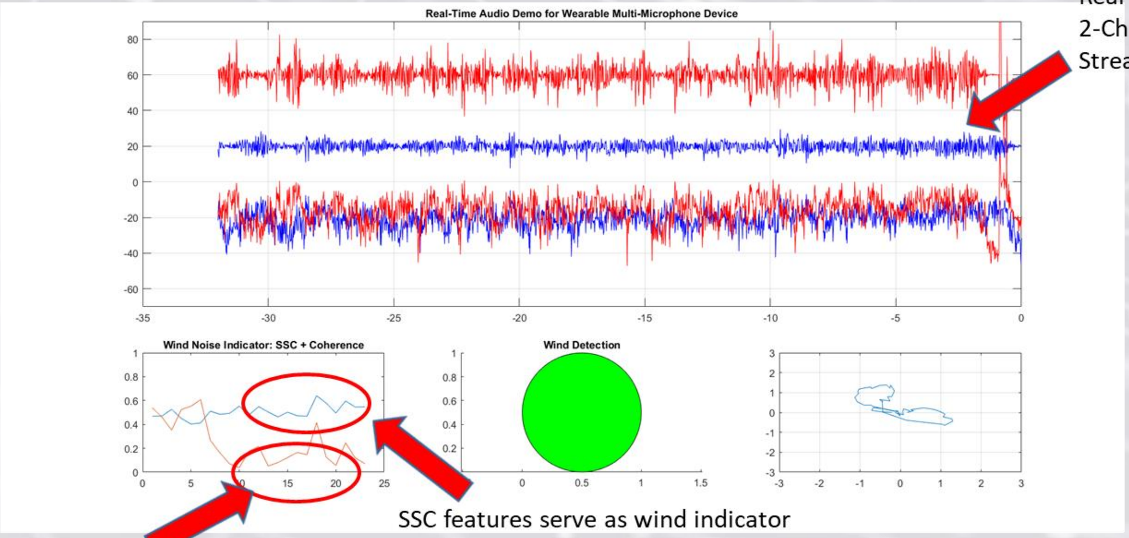
Experimental Results

- The RTWD algorithm experiments yielded a very strong detection accuracy (approximately 90%) in challenging, low wind intensity scenarios (~6 mph).

- These results are comparable with state-of-the-art active approaches used in wearable devices such as hearing aids.

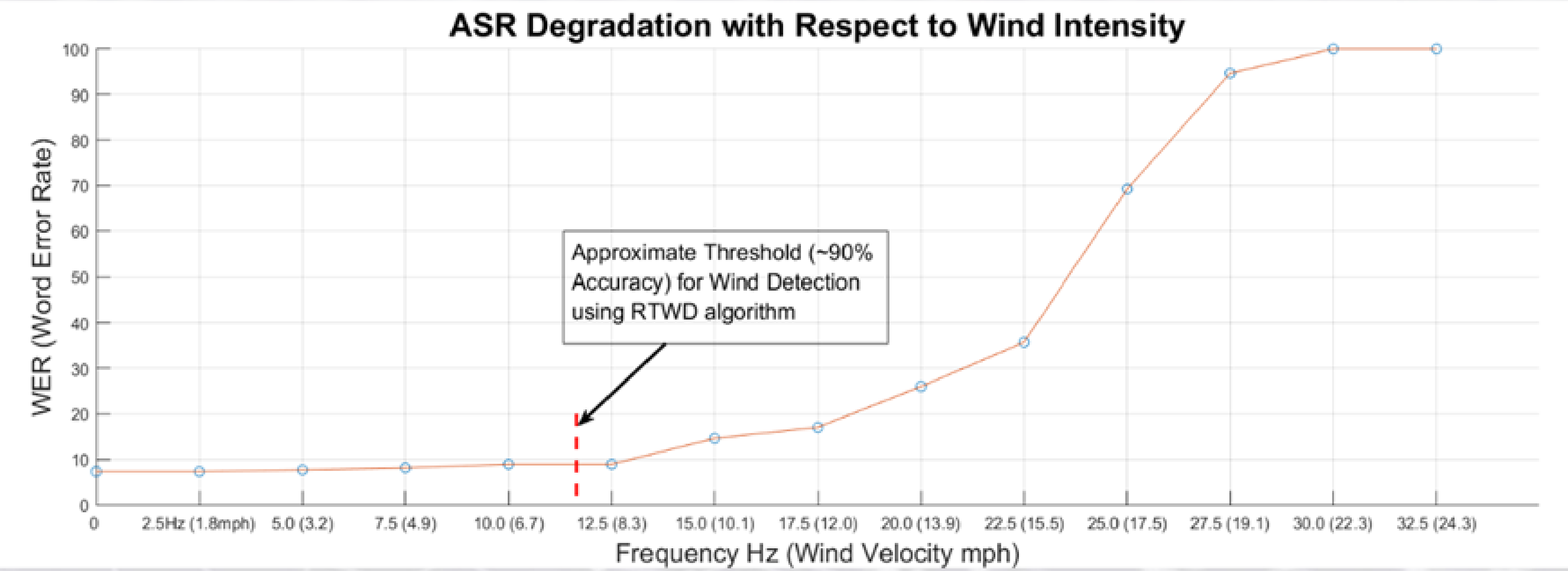
- In the case of medium and strong wind (10 mph+) the detection accuracy was nearly perfect; the algorithm furthermore performed very well even in the case of partial or full microphone occlusion (viz., for one channel), as well as in both cases of directed and diffuse wind.

Real-Time
2-Channel Audio
Stream



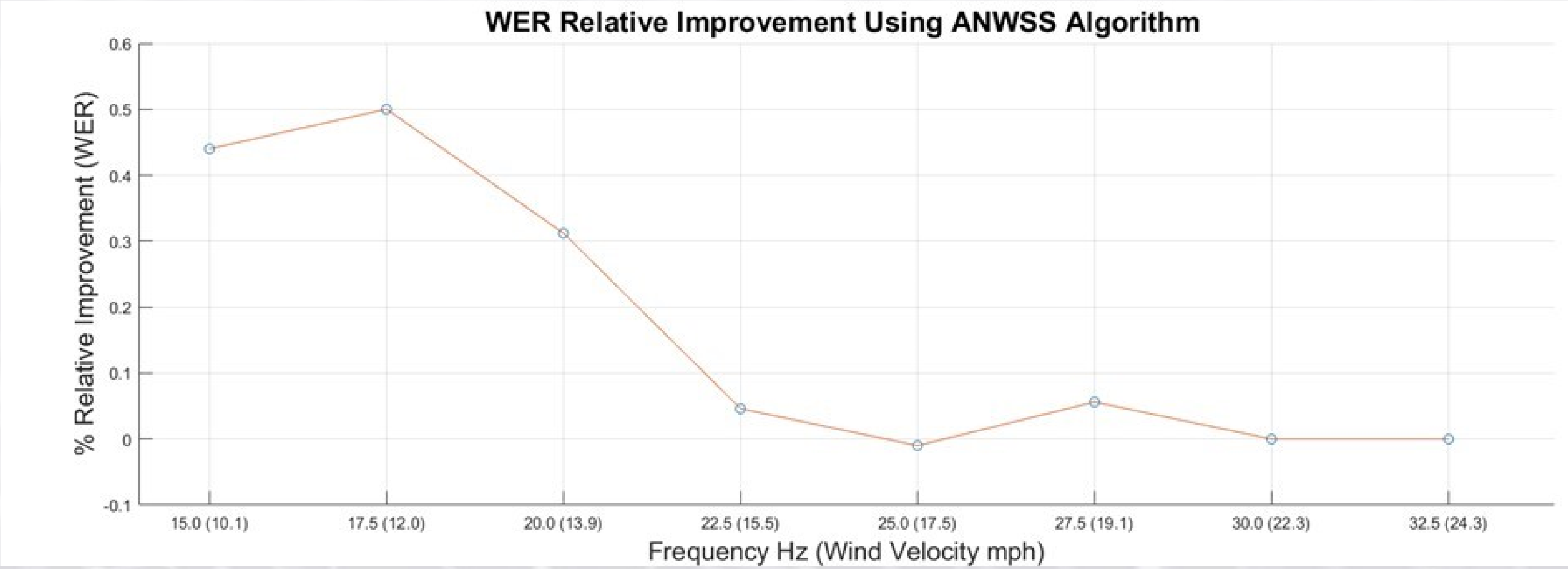
Coherence features indicate degree to
which audio stream is "Speech-like"

Experimental Results



- These results augur favorably for wind noise suppression when we consider the nature of ASR degradation with respect to wind intensity (see Figure).
- From our experiments, we observed a negligible decline in WER for wind intensities less than 8 mph. In the range 9-15 mph, WER was moderate (indicating that quality clean speech reconstruction is still achievable); beyond wind speeds of 15 mph, however, WER grows sharply.

Experimental Results



- WER for ASR was significantly reduced using the ANWS algorithm, showing the considerable potential of this method. In particular, the algorithm performs very well in moderate to strong wind regimes for which ASR degradation is most precipitous;
- At 12 mph, for example, ANWS reduced WER by 50% – see Figure above.
- Although accurate ASR in severe wind conditions (25 mph+) may be generally unfeasible, the ANWS-based reconstructed audio under these extreme conditions is nonetheless still commonly comprehensible to a human listener, indicating the potential further utility of ANWS as a noise suppression method for human-to-human audio communications.

Conclusion

- We successfully developed a novel, robust and strongly competitive, low-energy wind noise suppression system portable to wearable and smart devices endowed with multi-channel capacities.
- Future iterations of this system would likely yield improved results by utilizing a data-driven process to dynamically learn *attentive spectral regions* for signal reconstruction, in addition to incorporating *noise-aware* training [28].
- More generally, the method we advance, which is built around the idea that different noise classes possess characteristic, learnable spectral energy distributions, could potentially be applied across a broad range of noise sources.
- In this way, we imagine that a future noise classification-suppression system grounded in this approach could provide an indispensable tool (e.g., through "context-awareness" and object-class localization capabilities) in the development of a <u>fully-realized, "intelligent" audio system and the incipient IoT</u>.

12mph wind noise + voice

12mph wind + voice (ANWS reconstruction)

18mph wind noise + voice

18mph wind + voice (ANWS reconstruction)

- Thanks for listening. Questions and comments are welcome.

<u>Patents based on work presented</u>:

Rhodes, A. D., Kar, S., *Efficient Wind Detection using Multiple Microphones for Headworn Devices.*

Rhodes, A. D., Kar, S., *Neural-Based Signal Reconstruction using Multiple Microphone for Headworn Devices*. (Pending)

<u>Conference paper</u>:

Rhodes, A. D., (2017). *Real-Time Wind Noise Detection and Suppression with Neural-Based Signal Reconstruction for Multi-Channel, Low-Power Devices*. (Submitted)

References

[1] Alexandre, E., Lucas, C., Álvarez, Utrilla, M., "Exploring the Feasibility of a Two-Layer NN-Based Sound Classifier for Hearing Aids," *EUSIPCO*, 2007.
[2] Bagchi, D., Mandel, M., Wang, Z., He, Y., Plummer, A., Fosler-Lussier, E., "Combining Spectral Feature Mapping and Multi-Channel Model-Based Source Separation for Noise-Robust Automatic Speech Recognition," *ASRU*, 2015.
[3] Chen, Wenjun et al. "SVD-Based Technique for Interference Cancellation and Noise Reduction in NMR Measurement of Time-Dependent Magnetic Fields." Ed. Andreas Hütten. Sensors (Basel, Switzerland) 16.3 (2016): 323. PMC. Web. 19 Sept. 2017.
[4] Doclo, S., Moonen, M., "Robustness of SVD-Based Optimal Filtering For Noise Reduction in Multi-Microphone Speech Signals," Proc. of the 1999 *IEEE International Workshop on Acoustic Echo and Noise Control* (IWAENC'99), Pocono Manor, Pennsylvania, USA, Sep. 27-30 1999, pp. 80-83.
[5] Doclo S., Dologlou, I., Moonen, M., "A Novel Iterative Signal Enhancement Algorithm for Noise Reduction in Speech," *ICSLP* 1998.
[6] Fischer, D., Gerkmann, T., "Single-Microphone Speech Enhancement Using MVDR Filtering and Wiener Post-Filtering," *ICASSP*, 2016.
[7] Freeman, C., Dony, R.D., Areibi. S.M., "Audio Environment Classification for Hearing Aids Using Artificial Neural Networks with Windowed Inpu,". *CIISP*, 2007.
[8] Ghasemi, J., Karami Mollaei, M.R., "A New Approach Based on SVD for Speech Enhancement," *CSPA*, 2011.
[9] Gerkmann, Timo, et al., "Phase Processing for Single-Channel Speech Enhancement," *IEEE Signal Processing Magazine*, March 2015.
[10] Hansen, P., Jensen, S., "Subspace-Based Noise Reduction for Speech Signals via Diagonal Triangular Matrix Decompositions," *EURASIP*, 2007.
[11] Heymann, J., Drude, L., Haeb-Umbach, R., "'Neural Network Based Spectral Mask Estimation for Acoustic Beamforming," *ICASSP* 2016.
[12] Kumar, A., Florencio, D., "Speech Enhancement in Multiple-Noise Conditions Using Deep Neural Networks," *INTERSPEECH*, 2016.
[13] Leng, S., Ser, W., "Adaptive Null Steering Beamformer Implementation for Flexible Broad Null Control. *Signal Processing*, Vol. 91, Issue 5, pp. 1229-1239, 2011.
[14] Lilly, B.T., Paliwal, K.K., "Robust Speech Recognition Using Singular Value Decomposition Based Speech Enhancement *TENCON*, 1997.
[15] Loizou, P., *Speech Enhancement: Theory and Practice*, CRC Press, 2013.
[16] Lu, X., Taso, Y., Matsuda, S., Hori, C., "Speech Enhancement Based on Deep Denoising Autoencoder," *INTERSPEECH*, 2013.
[17] Maj, JB., Moonen, M. & Wouters, J. *EURASIP J. Adv. Signal Process.* (2002) 2002: 852365.
[18] Murphy, Kevin. *Machine Learning: A Probabilistic Perspective*, MIT Press, 2012.
[19] Nelke C., Jax, P., Vary, P., "Wind Noise Detection: Signal Processing Concepts for Speech Communication," *DAGA*, 2016
[20] Nelke, C., "Wind noise short term power spectrum estimation using pitch adaptive inverse binary masks", *ICASSP*, 2015.
[21] Nelke C., Vary, P., "Dual Microphone Wind Noise Reduction by Exploiting the Complex Coherence," *Speech Communication*, 2014.
[22] Nelke, C., Chatlani, N., Beaugeant, C., Vary, P., "Single Microphone Wind Noise PSD Estimation Using Signal Centroids," *ICASSP*, 2014.
[23] Nemer, E., et al., "Single-microphone wind noise suppression," Patent 2010/00209, 2010.
[24] Ochiai, T., Watanabe, S., Hori, T., Hershey, J., "Multichannel End-to-End Speech Recognition," *ICML*, 2017.
[25] Park, J., Park, J., Lee, S., Hahn, M., "Coherence-Based Dual Microphone Wind Noise Reduction by Wiener Filtering," *ICSPS* 2016.
[26] Schmidt, M., Larsen, J., Hsiao, Fu-Tien, "Wind Noise Reduction Using Non-Negative Sparse Coding," *Machine Learning for Signal Processing*, 2007.
[27] Schmidt, M., Larsen, J., "Reduction of Non-Stationary Noise Using a Non-Negative Latent Variable Decomposition," *Machine Learning for Signal Processing*, 2008.
[28] Seltzer, M., Yu, D., Wang, Y., "An Investigation of Deep Neural Networks for Noise Robust Speech Recognition," *ICASSP*, 2013.
[29] Shao, W., Wang, Wei-cheng, "A New GSC based MVDR Beamformer with CS-LMS Algorithm for Adaptive Weights Optimization," *CISP*, 2011.
[30] Sun, M., Li Y., Gemmeke, J., Zhang, X., "Speech Enhancement Under Low SNR Condition via Noise Estimation Using Sparse and Low-Rank NMF with Kullback-Leibler Divergence," *IEEE Transactions of Audio, Speech and Language Processing*, Vol. 23, Issue 7, 2015.
[31] Thomas, M., Ahrens, J., Tashev. I., "Optimal 3D Beamforming Using Measured Microphone Directivity Patterns," *IWAENC*, 2012.
[32] Vary, P., Martin, R., *Digital Speech Transmission. Enchancement, Coding and Error Concealment*, Wiley Verlag, 2006.
[33] Vaseghi, S., *Advanced Signal Processing and Noise Reduction*, "Spectral Subtraction." Wily & Sons, 2000.
[34] Verteletskaya, E., Simak. B., "Noise Reduction Based on Modified Spectral Subtraction Method." *IAENG*, 2011.