ROBUST WIND NOISE DETECTION

Justin A. Zakis, Christine M. Tan

Wolfson Dynamic Hearing
2 Chapel St, Richmond, VIC 3121, Australia.

ABSTRACT

Wind noise can be a major problem with audio devices such as hearing aids, cochlear implants, phones and headsets. Previous wind-noise detection algorithms generally assume that large level and/or phase differences between two microphones indicate wind noise, while small differences indicate its absence. However, differences may exist without wind noise due to unmatched microphones, acoustic reflections, or the phase shift caused by the microphone spacing. This paper shows that previous algorithms do not always correctly differentiate between wind and non-wind causes of microphone signal differences, which could lead to the inappropriate engagement of wind-noise reduction processing. A novel algorithm is presented, which performs an efficient statistical analysis of the microphone signals that is substantially more robust against non-wind causes differences, and hence false wind-noise detection, in an exemplary hearing-aid application.

Index Terms— Wind noise detection, hearing aids, cochlear implants, phones, consumer audio

1. INTRODUCTION

When a microphone is in an air flow, pressure variations from flow turbulence deflect the microphone diaphragm, which results in wind noise in the microphone output signal. Turbulence can exist in environmental wind [1] and is locally generated where an air flow moves around an object [2]. For the latter, the wind-noise spectral peak frequency is proportional with the flow speed and inversely proportional with the diameter of the object [2]. Therefore, small objects such as microphone ports generate higher-frequency wind noise than larger objects such as the human head [2].

Wind noise is problematic for a variety of audio devices. For instance, surveys report a 58 % satisfaction rate with hearing-aid performance in wind, compared with 61 % for noisy situations and 91 % for one-on-one conversation in quiet [3]. Hearing-aid wind noise levels can exceed low-frequency speech levels at a wind speed of only 3 m/s (11 km/h) and also exceed high-frequency speech levels at 6 m/s (22 km/h) [4]. Speech masking by wind noise may occur at even higher frequencies in the impaired auditory system. Wind noise can also clip hearing-aid electret microphone circuits by 12 m/s (43 km/h), resulting in equivalent input levels of up to 116 dB SPL [4]. Therefore, aggressive wind-noise *reduction* algorithms are needed for reducing the undesirable effects of wind noise on listening comfort and possibly speech understanding, and a reliable means of *detecting* the presence or absence of wind noise is important for appropriately controlling such algorithms.

Previous wind-noise detection (WND) algorithms typically compare samples of two microphones [5-7]. They tend to assume that large level and/or phase differences between microphones indicate the presence of wind noise (i.e. slow-moving turbulence), while small differences indicate its absence. However, large differences may also exist due to unmatched microphones, acoustic reflections, or the phase shift caused by the microphone spacing. It is important that WND algorithms correctly differentiate between wind and non-wind causes of differences between microphones, otherwise non-wind sounds may be falsely detected as wind noise. This could lead to the inappropriate use of wind-noise reduction algorithms, which could have unintended consequences, such as reduced speech intelligibility and sound quality, in the *absence* of wind.

However, despite the importance of reliable WND, the authors are not aware of any peer-reviewed papers that compare the efficacy of two-microphone WND methods, which have typically only been described in patents or books [5-8]. This paper extends the literature by presenting a comparative evaluation of two previous WND methods that have been described in the context of hearing-aid processing [8]. This paper further extends the literature by describing and co-evaluating the novel Chi-squared (χ^2) WND algorithm, which differs in that it performs an efficient statistical analysis of the microphone signals to determine the presence or absence of wind noise. The WND algorithms were evaluated with a hearing aid and a selection of wind and non-wind stimuli. Comparisons of WND algorithm output data for different stimuli were used to establish robustness against false WND. This paper is organized as follows. Section II describes the previous and χ^2 WND algorithms. Section III describes the evaluation methods. Section IV presents the evaluation results. Section V provides a discussion and future research directions. Section VI draws conclusions from the research.

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

2. WIND-NOISE DETECTION ALGORITHMS

2.1. Cross/Auto Correlation

This previous WND algorithm calculated [6, 8]:

$$\Gamma = \frac{\sum_{n=-k+1}^{k} x(n)y(n-l)}{\sum_{n=-k+1}^{k} x^2(n-l)} \tag{1}$$

where *x(n)* and *y(n)* are samples of microphones *x* and *y*, respectively, *l* is the correlation lag (l=0 for this evaluation), and k>0 for correlation over a vector (block) of 2*k* samples. The value Γ should approach 1.0 when *x(n)* and *y(n)* have similar amplitude and phase, such as for non-wind sounds, and tend toward 0.0 as *x(n)* and *y(n)* become less similar, such as for wind noise [8].

2.2. Difference/Sum

This previous WND algorithm calculated [7, 8]:

$$\Delta = \frac{\sum_{n=-k+1}^{k} \left| x(n) - y(n) \right|^2}{\sum_{n=-k+1}^{k} \left| x(n) + y(n) \right|^2} \tag{2}$$

where *x(n)* and *y(n)* are samples of microphones *x* and *y*, respectively, over a block of 2*k* samples. The value Δ should approach 0.0 when *x(n)* and *y(n)* have similar amplitude and phase, such as for non-wind sounds, and tend towards 1.0 as *x(n)* and *y(n)* become less similar, such as for wind noise [8].

2.3. Chi-Squared

The novel χ^2 WND algorithm adapts the χ^2 test of independence to estimate the dissimilarity between two microphone signals. The χ^2 test is traditionally used in social sciences and quality control to compare two or more sets of observed *categorical* data. Thus, adapting the χ^2 test to WND requires converting the microphone samples to categorical data. In devices that represent samples with the twos-compliment number system (and zero DC bias), it is simple to categorize samples as having either a positive (≥0) or negative (<0) value based on the sign bit. This process is used to construct the observed data matrix, **O**:

$$\mathbf{O} = \begin{Bmatrix} \sum_{n=-k+1}^{k} \mathrm{POS}(x(n)) & \sum_{n=-k+1}^{k} \mathrm{NEG}(x(n)) \\ \sum_{n=-k+1}^{k} \mathrm{POS}(y(n)) & \sum_{n=-k+1}^{k} \mathrm{NEG}(y(n)) \end{Bmatrix} \tag{3}$$

where *x(n)* and *y(n)* are samples of microphones *x* and *y*, respectively, and **POS** and **NEG** are functions that return the number of positive and negative samples, respectively, over a block of 2*k* samples. The expected data matrix, **E**, is then calculated as in [9]:

$$e_{ij} = \frac{\sum_{m=1}^{c} o_{im} \cdot \sum_{m=1}^{r} o_{mj}}{N} \tag{4}$$

where *e* and *o* are elements of **E** and **O**, respectively, *r* and *c* are the number of rows and columns, respectively, of **O**, and *N* is the sum of elements in **O**. Each e_{ij} is the product of the row *i* and column *j* sums of **O**. The χ^2 statistic is then [9]:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\left(o_{ij} - e_{ij}\right)^2}{e_{ij}} \tag{5}$$

The χ^2 value is zero when the ratio of positive to negative sample counts is the same for both microphones (i.e. o_{11}:o_{12}=o_{21}:o_{22}), as approximated with non-wind sounds, and increases as the ratios become less similar, as for wind noise. Compared with previous algorithms, the χ^2 WND algorithm ignores level differences, and may be less sensitive to phase differences since it does not compare the microphone signals on a sample-by-sample basis. For this application to WND, given known constants (**O** row sums = 2*k*, *N*=4*k*,), Eq. 4-5 can be replaced with a single equation that only uses *N* and the negative sample counts:

$$\chi^2 = \left(o_{12} - o_{22}\right)^2 \times \left(\frac{1}{\left(o_{12} + o_{22}\right)} + \frac{1}{\left(N - o_{12} - o_{22}\right)} \right) \tag{6}$$

Divide operations can be efficiently implemented in real time with a look-up table of length *N*. In order to avoid division by zero, if o_{12} or o_{22} equals zero it is set to 1. This rule and Eq. 3 and 6 were used to evaluate the χ^2 WND.

3. EVALUATION METHODS

A medium-sized, behind-the-ear hearing aid ("BTE2" from [4]), with a typical spacing of 13 mm between its unmatched microphones, was modified so a thin, shielded, multi-core cable entered near its base (well clear of the microphone ports) and was connected to each microphone's power, ground, and signal terminals. All other hearing-aid circuits were disconnected from the microphones and unpowered. The other end of the cable was connected to a 1.5-V cell to power the microphones, and the stereo input of a 32-bit, external sound card (108 dB dynamic range). The sound card's maximum input level was above the microphone clipping level. Stereo digital recordings of both microphone outputs were made while the hearing aid was presented with wind and non-wind stimuli.

For the non-wind recordings, the aid was on the right ear of a head-and-torso simulator, which was in the center of a speaker array (0, 90, 180 and 270˚ azimuth, 2-m diameter) in a sound booth. The following stimuli were presented at 65 dB SPL: 1) White noise filtered to match the International Long-Term Average Speech Spectrum (ILTASS); 2) Eight-talker babble; 3) Tone sweep linearly increasing from 0.1 to 8 kHz over 120 s; and 4) Quiet. The ILTASS noise and babble were presented from all four speakers for 30 s. The tone sweep was presented from the front speaker (0˚) to maximize microphone phase (time delay) differences.

The wind recordings were made with the National Acoustic Laboratories' (Chatswood, NSW, Australia) low-noise wind generator [10]. The fan and motor noise at the wind outlet was well below typical wind-noise levels for hearing aids [4]. The aid was located on a head-and-torso simulator's right ear, and the head was placed on top of a vertical length of pipe that was attached to a turntable. With the head in front of the 61×61 cm wind outlet, 31-s recordings were made for 72 combinations of wind speed (3 and 6 m/s) and wind azimuth (0-350˚ at 10˚ intervals). Such wind conditions cover a wide range of microphone level differences and wind-noise bandwidths with this aid [4].

The WND algorithms described in Section II were implemented in MATLAB Simulink. For each wideband (i.e. not pre-filtered) recording, the WND algorithm under test calculated an output value (Γ, Δ or χ^2) for every non-overlapping block of 16 samples at a 20-kHz sampling rate (this replicated hearing-aid applications with minimal computational power for WND). The output values were passed through an IIR filter (b = [0.0032]; a = [1 -0.9968]) to smooth jitter between blocks. Thirty seconds worth of smoothed block values (120 s for the tone sweep) were stored in a data file for analysis.

In addition, the MATLAB script described in [4] was used to calculate the long-term-average input level of each wind-noise recording. As a calibration reference, the script used a recording of white noise at 78.5 dB SPL with the hearing aid placed in a B&K Type 4232 anechoic box.

4. RESULTS

Figures 1 to 3 show the smoothed outputs of the Cross/Auto Correlation, Difference/Sum and χ^2 WND algorithms, respectively. The circle, squares and triangles show the long-term-average WND output for the quiet and 3 & 6 m/s wind recordings, respectively, with the corresponding input levels shown on the black (middle) horizontal axis. The solid and dotted blue curves show the WND output for the ILTASS and babble noise, respectively, over time shown on the blue (bottom) horizontal axis. The red dashed curve shows the WND output for the 120-s tone sweep across frequency shown on the red (top) horizontal axis. All curves start at the origin due to the initial conditions of the smoothing filter.

An ideal WND algorithm should output different sets of values for wind and non-wind sounds with no overlap. The Cross/Auto Correlation WND algorithm should output approximately 1.0 for non-wind sounds and lower values for wind. Fig. 1 showed its output was correct for quiet, ILTASS noise, babble, and low-frequency tones. However, high-frequency tones resulted in wind-like outputs.

The Difference/Sum WND algorithm should output approximately 0.0 for non-wind sounds and higher values for wind. Fig. 2 showed its output was correct for quiet, ILTASS noise, babble, and tones at some frequencies but not others.

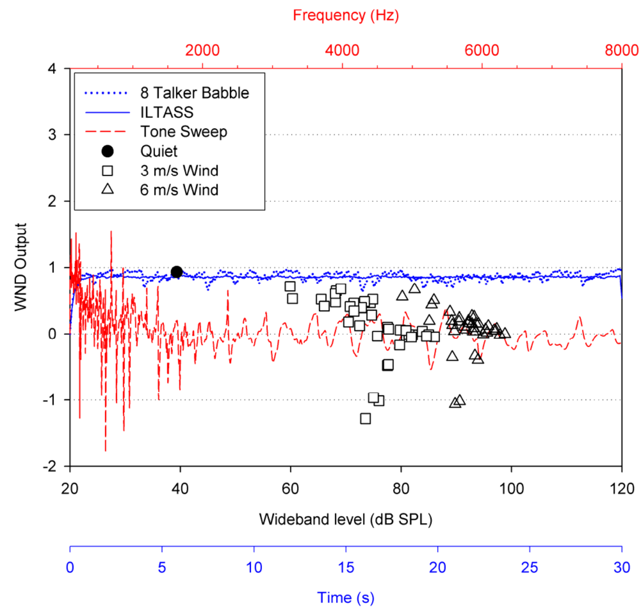

Fig. 1 Smoothed Cross/Auto Correlation WND output values.

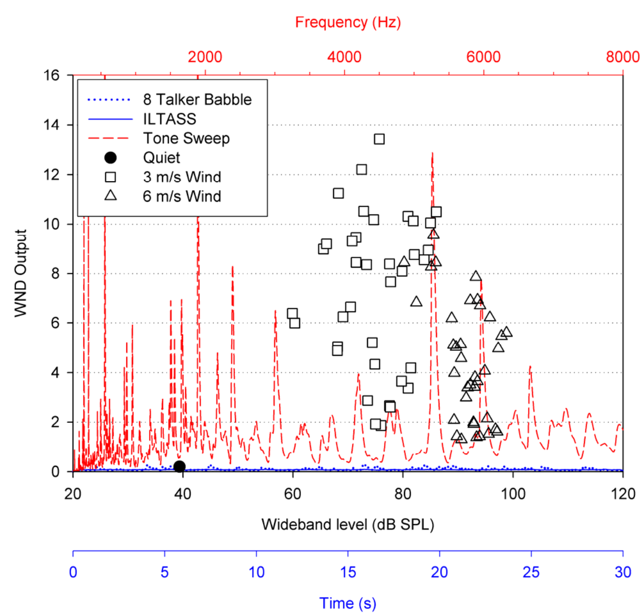

Fig. 2 Smoothed Difference/Sum WND algorithm output values.

The χ^2 WND algorithm should output approximately 0.0 for non-wind sounds and higher values for wind. Fig. 3 shows that its output values for all non-wind sounds were clearly separated from those for wind noise.

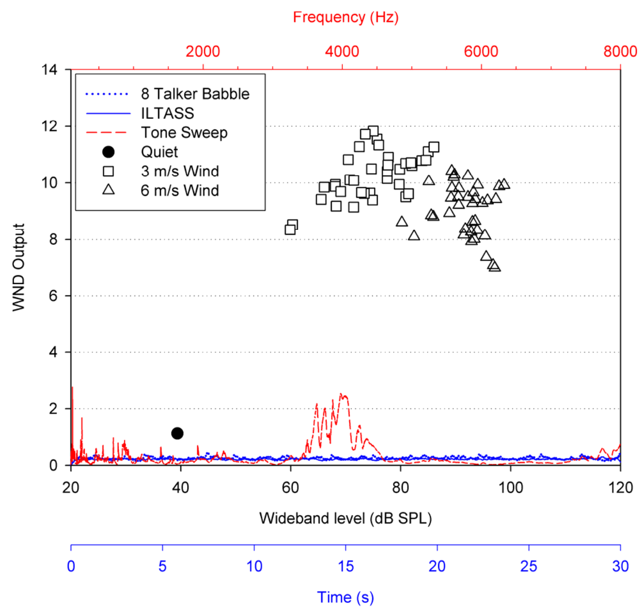

Fig. 3 Smoothed χ^2 WND algorithm output values.

5. DISCUSSION

The two previous and the novel χ^2 WND algorithm used different methods (correlation ratio, power ratio or waveform statistics) to determine the presence or absence of wind noise. As a result, the algorithms differed in their sensitivity to amplitude and phase differences between microphone signals, and hence how well non-wind sounds were differentiated from wind noise. Reliable WND is important to ensure that wind-noise reduction processing is appropriately applied to improve satisfaction with hearing aids in wind. Otherwise, the inappropriate application of such processing in the *absence* of wind noise may negatively affect sound quality and/or speech intelligibility, and hence satisfaction with hearing aids. This would also generally apply to other audio devices, such as mobile phones, telephony headsets, and video recorders.

All evaluated WND algorithms handled the ILTASS noise, babble, and quiet conditions well, while the tone sweep was problematic for the previous WND algorithms. The long-term-average spectrum of the tone sweep recording was calculated with Adobe Audition (Hanning window and 4096-point FFT) and revealed front-minus-rear microphone level differences that varied from -15 to 30 dB across frequency. In contrast, the long-term-average spectrum of the calibration recording suggested that microphone sensitivity differences varied from -1 to 3 dB (front minus rear) over the same bandwidth (caveat: calibration stimulus level differences were specified as no greater than 1 dB between microphones in the B&K Type 4232 anechoic box). Therefore, acoustic reflections and standing waves could dominate microphone sensitivity differences for the tone sweep, so as a result the tone sweep had a combination of level *and* phase differences between microphones. Analysis of the sweep recordings suggests that the standing wave null frequencies differed between microphones, and around these null frequencies room reflections with different phase and/or low-level, non-wind noise could dominate one microphone but not the other. This may explain the spikes in WND output that exceeded the theoretical limits of Eqs. 1 and 2 for clean, pure tones *without* reflections. The correct WND outputs for the ILTASS noise and babble suggests that with these wideband stimuli, the effect of acoustic reflections evened out across frequency and the unmatched microphones had little effect. The latter is advantageous for avoiding the expense of matched microphones, or the use of scarce hearing-aid processing resources for matching filters.

It is likely that the χ^2 WND algorithm better handled the tone sweep because it ignored level differences between microphones. Another possible reason is that it may be less sensitive to small phase differences relative to the block length, since it does not compare the microphone signals on a sample-by-sample basis. However, these algorithmic effects could not be separated with the available data.

Future research could investigate the effect of microphone spacing, block size, use in different types of device (e.g. mobile phones, telephony headsets, video recorders, etc.), and different evaluation stimuli on WND algorithm performance. In applications where more computational power is available for WND, sub-band implementations would be advantageous for controlling wind noise suppression algorithms, since wind noise can exceed speech levels up to at least 8 kHz [4]. The χ^2 WND algorithm could also be extended to compare three or more microphones in applications where this is advantageous.

6. CONCLUSIONS

The novel χ^2 wind noise detection (WND) algorithm was introduced in this paper. This algorithm adapts the χ^2 statistical test of independence, which is commonly used in the social sciences to compare sets of categorical data, to compare microphone signals in order to determine the presence or absence of wind noise. In an exemplary hearing-aid application, the χ^2 WND algorithm was substantially more robust against incorrectly detecting pure tones as wind noise when compared with two previous WND algorithms, which used correlation or level differences to compare microphone signals.

7. REFERENCES

[1] S. Morgan and R. Raspet, “Investigation of the mechanisms of low-frequency wind noise generation outdoors,” *Journal of the Acoustical Society of America*, vol. 92, pp. 1180–1183, 1992.

[2] H. Dillon, I. Roe, and R. Katsch, “The sources of wind noise in hearing aids,” in *The International Hearing Aid Research Conference*, Lake Tahoe USA, August 2000.

[3] S. Kochkin, “MarkeTrak VIII: Consumer satisfaction with hearing aids is slowly increasing,” *Hearing Journal*, vol. 63, no. 1, pp. 19-32, 2010.

[4] J. A. Zakis, “Wind noise at microphones within and across hearing aids at wind speeds below and above microphone saturation,” *Journal of the Acoustical Society of America,* vol. 129, no. 6, pp. 3897-3907, 2011.

[5] T. Dickel and B. Knapp, “Method for operating a hearing aid or hearing aid system, and a hearing aid and hearing aid system,” U.S. Patent 6 882 736, 2005.

[6] K. S. Petersen, G. Bogason, U. Kjems, and T. B. Elmedyb, “Device and method for detecting wind noise,” U.S. Patent 7 340 068, 2008.

[7] G. W. Elko, “Reducing noise in audio systems,” U.S. Patent 7 171 008, 2007.

[8] Kates, J.M., *Digital Hearing Aids*, Plural Publishing, San Diego, 2008.

[9] NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/, 3 Nov 2013.

[10] M. Fishburn, G. Colin-Thome, and P. Always, “Generating a ‘silent’ flow of air to measure the aerodynamic and acoustical performance of a louvre under acoustical laboratory conditions,” *14th International Congress on Sound and Vibration*, Cairns, Australia, 9-12 July 2007.