Interspeech 2018
2-6 September 2018, Hyderabad

Classification of nonverbal human produced audio events: a pilot study

Rachel E. Bouserhal[1,4], *Philippe Chabot*[1], *Milton Sarria-Paja*[3], *Patrick Cardinal*[2], *Jérémie Voix*[1,4]

École de technologie supérieure
Departments of [1]Mechanical Engineering and [2]Software and Information Technology Engineering
1100 Notre-Dame St W, Montréal, Québec, Canada
[3]Universidad Santiago de Cali
Calle 5 No 62-00 Pampalinda, Cali, Colombia
[4]Centre for Interdisciplinary Research in Music Media and Technology
527 Rue Sherbrooke O, Montréal, Québec, Canada

rachel.bou.serhal@etsmtl.ca, philippe.chabot.1@ens.etsmtl.ca, milton.sarria00@usc.edu.co, patrick.cardinal@etsmtl.ca, jeremie.voix@etsmtl.ca

Abstract

The accurate classification of nonverbal human produced audio events opens the door to numerous applications beyond health monitoring. Voluntary events, such as tongue clicking and teeth chattering, may lead to a novel way of silent interface command. Involuntary events, such as coughing and clearing the throat, may advance the current state-of-the-art in hearing health research. The challenge of such applications is the balance between the processing capabilities of a small intra-aural device and the accuracy of classification. In this pilot study, 10 nonverbal audio events are captured inside the ear canal blocked by an intra-aural device. The performance of three classifiers is investigated: Gaussian Mixture Model (GMM), Support Vector Machine and Multi-Layer Perceptron. Each classifier is trained using three different feature vector structures constructed using the mel-frequency cepstral (MFCC) coefficients and their derivatives. Fusion of the MFCCs with the auditory-inspired amplitude modulation features (AAMF) is also investigated. Classification is compared between binaural and monaural training sets as well as for noisy and clean conditions. The highest accuracy is achieved at 75.45% using the GMM classifier with the binaural MFCC+AAMF clean training set. Accuracy of 73.47% is achieved by training and testing the classifier with the binaural clean and noisy dataset.

Index Terms: nonverbal, classification, hearing protection, biosignals

1. Introduction

The human body communicates countless nonverbal signals (heartbeat, blinking, coughing, etc...) that can be revealing of one's health and emotional state. Capturing and classifying such nonverbal events has gained much interest over the years [1, 2]. Namely, the combination of well performing machine learning algorithms and the boom in wearable devices has opened up a door for continuous health monitoring [3, 4]. Often, these nonverbal signals are either inaudible (blinking), too faint (clearing the throat), or overlooked (teeth clicking) when considering the communication signals sent by the human body. Fortunately, the ear canal can act as an efficient medium to these nonverbal signals.

When the ear canal is blocked at the entry, there is a buildup of energy from soft tissue and bone conduction causing an amplification in the bone-conducted sounds in the ear canal. This phenomenon is called the *occlusion effect* [5]. Thus, by way of the occlusion effect, intra-aural devices that create an acoustical seal in the ear canal have access to an extensive variety of human produced verbal and nonverbal audio events. Martin and Voix [3], have shown that the occluded ear is a reliable place to capture breathing and heartbeat signals for health monitoring. However, other relevant signals such as blinking, coughing and clicking of the teeth can also be recorded from inside the occluded ear. In this paper, two new applications for the classification of nonverbal human produced audio events are introduced.

The accurate classification of such signals allows for diverse applications beyond continuous health monitoring. One such application, is addressing a significant hurdle with in-ear dosimetry. To ensure safety and avoid noise-induced hearing loss, workers in noisy environments are usually equipped with hearing protection devices and dosimeters calculating the individual's noise dose over the work day. Recently, in-ear dosimetry has gained much interest as it offers a more accurate representation of the worker's true noise dose [6, 7]. An obstacle with in-ear dosimetry is the effect of physiological signals on the accumulated noise dose [8, 9]. To an untrained dosimeter the levels inside the ear caused, for example, by clearing the throat or coughing are mistakenly added to the dose calculation leading to inaccurate dosimetry readings at the end of the workday. Therefore, the ability to classify and reject physiological noise from the calculation of the daily noise dose will lead to a more accurate representation of the worker's noise exposure.

Another application, is to use subtle voluntary actions such as explicit tongue and teeth clicking to replace audible verbal command when necessary in human-machine interactions. Silent interfaces as such have gained much attention in recent years, as they provide an inconspicuous way of communication that is robust to ambient noise and accessible to people with speech impairment [10, 11]. Tongue movements have been extensively explored as a silent interface for people with limited mobility [12, 13]. Although the ear has also been explored as a means for a silent interface [14], to the best of our knowledge, none have explored the opportunity of using nonverbal acoustic signals captured inside the occluded ear for human-machine interaction.

Nonverbal acoustic events such as gunshots, sirens and screams [15, 16] as well as people walking and closing doors [17] have been classified for security purposes. Only a few have classified human produced acoustic nonverbal events such as different types of cough [18] and blinking [19]. Often, Support

10.21437/Interspeech.2018-2299

vector machine (SVM) are used for non-verbal acoustic event recognition [16, 17]. While neural network algorithms such as Multi Layer Perceptron (MLP) and Convolutional Neural Networks (CNN) have also been used successfully [17], they require larger datasets. Gaussian Mixture Models (GMM) are also used and have good results in audio classification problems [2].

This paper investigates the classification of non-verbal human produced audio events captured inside an occluded ear. Using Mel-Frequency Cepstral Coefficients (MFCCs), their derivatives, zero-crossing rate (ZCR), and auditory-inspired amplitude modulation features (AAMF) as features, the performance of an SVM, GMM and MLP classifier is validated using cross validation and compared. The data collection and classifier description are presented in Section 2. The results are shown in Section 3, followed by the conclusions in Section 4.

2. Methodology

2.1. Data Collection

Non-speech audio signals are recorded as described by [3]. In each ear, participants were equipped with an intra-aural device developed by EERS Global Technologies Inc. (Montreal, Canada), shown in Fig. 1. Each earplug contains an in-ear microphone to capture audio signals occurring in the occluded ear. Audio data was recorded using a multichannel digital audio recorder (H4n, Zoom Corporation, Tokyo, Japan) at a sampling rate of 48 kHz and 24 bit resolution. Data was collected from 25 participants, consisting of 19 males and 6 females, aged between 21 and 53, with an average age of 28. Participants were instructed to perform the actions associated with each nonverbal audio event (teeth chattering, tongue clicking, etc...) for at least ten seconds. In certain situations, participants were asked to repeat until a clear signal could be recorded. Audio signals were then labeled by hand post hoc. For this study talking (t) and the following 10 nonverbal audio events were used for classification: clicking of teeth softly (cts), clicking of teeth loudly (ctl), tongue clicking (cl), blinking forcefully (bf), closing the eyes (ce), closing the eyes forcefully (cef), grinding the teeth (gt), clearing the throat (clt), saliva noise (sn), and coughing (c). Since these signals are captured by way of the occlusion effect, their bandwidth is limited to < 2 kHz [5, 20]. For this reason, the signals were downsampled to 8 kHz to limit the bandwidth to informative data.

Envisioning that this type of application is pertinent in noisy settings, it is of interest to investigate the classification of signals degraded by noise. Therefore, a noisy dataset was created post-hoc by adding factory noise from the NOISEX-92 database [21] at 10 dB signal-to-noise ratio (SNR). In situations where the SNR is much lower, denoising algorithms would be used to clean the captured in-ear signals [3, 22]. For the scope of this paper, we will look at signals degraded by noise but not to the extent where denoising algorithms are engaged.

2.2. Feature Vectors

Samples of 400 ms are extracted for each of the aforementioned audio events captured from both the left and right ear. Table 1 shows the amount of monaural samples extracted for each class, i.e. the total number of samples is doubled binaurally. Typically, MFCCs have been used as features [15, 16, 17] to train classifiers for both verbal and non verbal events. Some have added wavelets [15], Perceptual Linear Prediction (PLP), or ZCR [16] to the feature vector. In this work, events are classified using MFCCs as features extracted using a window

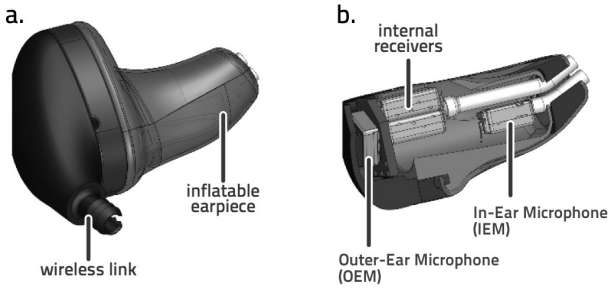

Figure 1: *The intra-aural device used to record the nonverbal audio signals captured inside the ear.*

Event	Number of samples
Clicking of teeth softly (cts)	246
Clicking of teeth loudly (ctl)	304
Tongue clicking (cl)	364
Blinking forcefully (bf)	207
Closing the eyes (ce)	286
Closing the eyes forcefully (cef)	329
Grinding the teeth (gt)	170
Clearing the throat (clt)	163
Saliva noise (sn)	213
Coughing (c)	219
Talking (t)	526

Table 1: *The total of 400 ms samples for each class.*

length of 50 ms with a 25 ms overlap. Parameters such as window length and overlap were selected after a pilot experiment, where 50 ms and 25 ms, respectively, were shown to provide the best results for the task at hand. Each MFCC vector consists of 13 MFCCs, delta, and delta-delta coefficients resulting in a 39 dimensional vector for each frame. In addition, due to the atypical nature of the nonverbal signals captured from inside the occluded ear, it is of interest to investigate other features as a complement or replacement of the MFCCs. In this work, we investigate the use of the auditory-inspired amplitude modulation features (AAMF), which were first proposed by Sarria et al. (2017) [23] to enhance speaker verification with whispered speech. When compared to standard MFCCs, AAMFs are said to have high discriminative capabilities and could, thus, be useful for the purposes of this work. AAMFs were proposed from blocks of spectrograms consisting of multiple consecutive short-time frames. As this leads to high-dimensional feature representations, principal component analysis (PCA) was adopted to reduce the number of variables in the feature space [23]. In this approach, it is assumed that an observed time-domain signal is the result of multiplying a low-frequency modulator (temporal envelope) by a high-frequency carrier, and the analysis is carried out by using acoustic subbands. In the case of speaker recognition applications, the modulation frequency (modulation domain) represents the frequency content of the subband amplitude envelopes and it potentially conveys information about speaking rate and other speaker specific attributes [24]. Hence, it is expected that AAMF features would be more informative than standard MFCC when describing changes in slow varying amplitude characteristics. As suggested in [23], time contexts of 200ms (a matrix with 216 elements - 27 acoustic bands $\times$ 8 modulation bands $=$ 216 dimensions) can be col-

lapsed into a vector and used as standard features after applying PCA to keep the 40 first components retaining 98.7% of cumulative variance.

Three different structures for the feature vector are used and compared. The first consists of the extracted MFCCs, the delta and the delta-delta from each frame concatenated into a one-dimensional vector appended with the ZCR of the entire sample. The second is framed (-F) and consists of 15 one-dimensional MFCC vectors appended with the ZCR of each frame for each sample. The third follows a similar structure as the second but includes context (-C). Each of the 15 feature vectors contains the 4 preceding frames, the target frame, and the 4 following frames. The choice of including 4 frames before and after was done ad hoc by investigating the change in average accuracy over all classifiers. An example of the change in accuracies as a function of the number of contextual frames used is shown in Section 3. For the remainder of the paper, for convenience, the feature vectors based on MFCCs, deltas, delta-deltas, and ZCR are referred to simply as the MFCC feature vectors.

Finally, we investigate the effects of training with samples captured from only one of the two ears versus including binaural samples. This is of interest because several factors contribute to the shape of the captured signal in an occluded ear: the shape of the ear canal and the way the intra-aural device fits in the ear [25, 26]. Therefore, even though the signals were captured simultaneously from both ears, using binaural data could serve as a type of data augmentation caused by slight differences in the left and right signals.

2.3. Classification

For classification, SVM, GMM, and an MLP neural network are used and compared. The hyper-parameters of the models are chosen to optimize the overall accuracy of the model over all classes.

A one-vs.-all classifier is used for the SVM classifier with a linear kernel to compute the 11 hyperplanes needed for the classification. For the GMM, a diagonal covariance matrix is used with 2 Gaussians components per class for the one-dimensional structure and 15 Gaussians components for the framed and contextual structure. Two hidden layers with a rectified linear activation function and a linear activation function for the output layer are used for the MLP. The network is trained using the cross entropy loss function and the Adam method of optimization [27].

When using the framed and contextual structure, one feature vector does not represent an entire sample. Therefore, the 15 feature vectors are classified independently and the final classification decision for the sample is made based on the most occurring class from each frame.

2.4. Validation

For the validation, a 10 fold cross validation is used where the different participants are separated in different batches to avoid testing on a trained test subject. The accuracy is then calculated for each fold by averaging the accuracy over each class to remove any weighting caused by the varying number of samples for each class.

3. Results

To start we use the clean binaural dataset to train and test all three classifiers using the three different feature vector structures. Fig. 2 shows the average accuracies over all classes for each classifier. As can be seen from Fig. 2, the GMM classifier using the contextual feature vector structure (GMM-C) results in the highest average accuracy of 75.47%, followed by the GMM with the framed feature vector (GMM-F) at 73.88%. The worst performing classifier is the MLP with the framed feature vector (MLP-F) at 34.03% accuracy. The degraded performance of the MLP could be attributed to the size of the dataset. In addition, the superior performance of the GMM classifier is advantageous, considering the applications of interest, as it was the fastest classifier to execute.

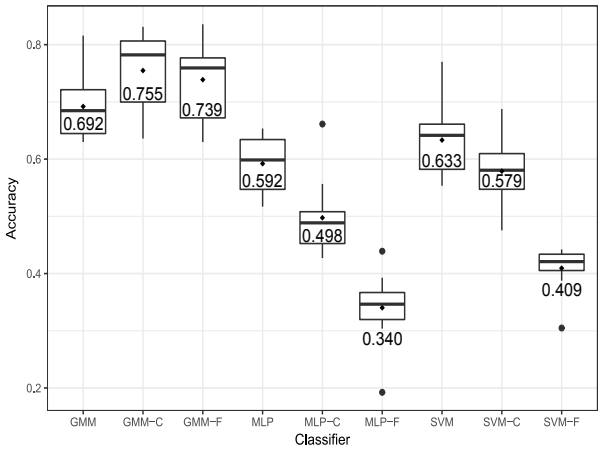

Figure 2: *Accuracy for each classifier using the binaural clean dataset to train and test, averaged over the 10-folds, where (-F) denotes the framed feature vector, (-C) denotes the contextual feature vector, and no affix denotes the one-dimensional feature vector.*

Using the GMM-C classifier, we compare the results between using a monaural and a binaural dataset. The mean accuracy over all classes for the GMM-C classifier are shown in Table 2. There is a benefit in using the binaural dataset versus monaurally. As previously discussed, variations in fit and ear canal shape contribute to differences in the captured in-ear signal. Due to this type of data augmentation caused by using the binaural dataset and the desire to implement this on a low complexity embedded system, only the performance of the GMM classifier with the contextual feature vector structure using the binaural dataset is presented for the remainder of the paper.

To better understand these nonverbal signals, as well as the performance of the classifier, the confusion matrix of the GMM-C classifier is shown in Fig. 3. It can be seen that besides speech, the most accurately classified event is coughing (c) at 85.2% accuracy, followed by closing the eyes (ce) and clearing the throat (clt) at 82.5% accuracy each. The most difficult event to classify is clicking of the teeth loudly (ctl) at 63.3%.

The general performance of the classifier could be attributed to the use of MFCCs, which are designed to characterize speech. In order to add complementary information extracted with AAMF features, we investigate the use of a fusion scheme, as it has shown to be effective at improving performance in many applications by combining the strengths of different feature representations [28, 29]. Particularly, we look at the effects of using fusion at the feature level of the AAMFs concatenated with the MFCCs. Figure 4 shows the confusion matrix when using the AAMFs in addition to MFCCs to train the GMM classifier. The overall accuracy as a consequence of

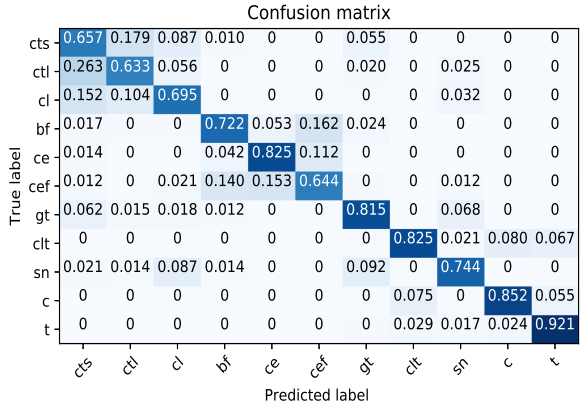

Figure 3: *Confusion matrix of the GMM classifier with MFCC contextual features trained and tested with the clean dataset.*

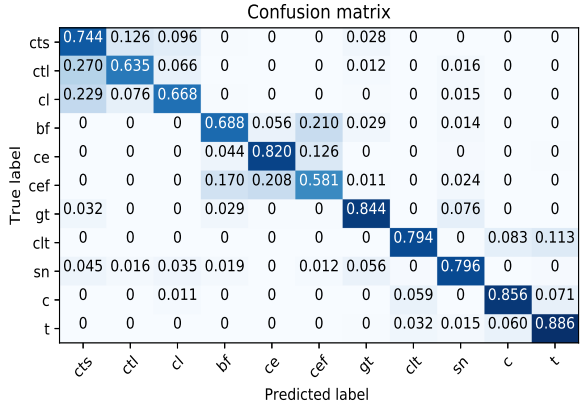

Figure 4: *Confusion matrix of the GMM classifier with MFCC + AAMF contextual features trained and tested with the clean dataset.*

complementing the MFCCs with AAMFs is 75.54%. Although this may seem an unremarkable difference from simply using MFCCs, it is important to study the confusion matrices in both conditions. The classification accuracy of 4 nonverbal classes increases when complementing with AAMFs. Namely, the classification accuracy of clicking the teeth softly (cts) is increased by 8.7%, that of saliva noise (sn) is increased by %5.2, and of grinding the teeth (gt) by 2.9%. The increase in classification accuracy for certain classes is indicative of the nature of those classes. Specifically, we can infer that these classes are better categorized by changes in their slow-varying amplitude envelope.

Next, we investigate the effect of a noisy dataset on the GMM-C classifier using MFCCs alone as well as their fusion

Dataset	Left	Right	Binaural
Overall Mean	0.750	0.733	0.755

Table 2: *Overall mean for the GMM classifier with contextual features (GMM-C) using monaural versus binaural datasets.*

		Test Dataset	
Features	Train Dataset	Clean	Noisy
MFCC	Clean	0.754	0.243
MFCC	Clean&Noisy	0.731	0.705
MFCC+AAMF	Clean	0.755	0.329
MFCC+AAMF	Clean & Noisy	0.735	0.728

Table 3: *Overall mean for the GMM classifier with MFCC contextual features (GMM-C) using Clean, Noisy and Clean & Noisy datasets for training and testing.*

with AAMFs for training. When trained with only the clean MFCC features, the performance of the classifier drastically decreases to 24.3%. Training with both noisy and clean samples increases the robustness of the classifier to noise but degrades the overall classification of clean signals as can be seen from Table 3. The classification accuracy of noisy events increases to 70.45% when using a combined (noisy and clean) training dataset, at a cost of a 2.37% decrease when classifying clean signals. Introducing the AAMFs to complement the MFCCs and using both clean and noisy samples for training increases the overall accuracy of the classifier when testing noisy and clean signals to 72.79% and 73.47% respectively, as presented in Table 3. In general, adding the fusion scheme of MFCCs with AAMFs increases the robustness of the classifier toñoise.

4. Conclusions

Classifying nonverbal human produced audio events can open up the door to many applications beyond health monitoring. In this paper, classification of nonverbal events was achieved with up to 75.54% average accuracy across all classes for clean events and 73.47% for noisy events with the GMM classifier and the contextual feature vectors using a fusion of MFCCs and AAMFs. Complementing the MFCC features with AAMFs increased the classification accuracy for some classes but not others. This suggests that ensemble learning could be beneficial in increasing the overall classification accuracy.

The results of the addition of AAMFs further suggest that the nature of theses nonverbal events should be more closely investigated and understood. One potential avenue to explore, is using features that are more adapted to low frequencies to accommodate for the limited bandwidth of bone conduction. The major challenges for this type of classification are the choice of features as well as the desire to limit computational complexity.

5. References

[1] B. M. Asl, S. K. Setarehdan, and M. Mohebbi, "Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal," vol. 44, no. 1, pp. 51–64, 2008. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0933365708000559

[2] T. Drugman, J. Urbain, and T. Dutoit, "Assessment of audio features for automatic cough detection," in *19th European Signal Processing Conference*. IEEE, 2011, pp. 1289–1293. [Online]. Available: http://ieeexplore.ieee.org/abstract/document/7074201/

[3] A. Martin and J. Voix, "In-ear audio wearable: Measurement of heart and breathing rates for health and safety monitoring," pp. 1–1, 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7959201/

[4] A. Pantelopoulos and N. G. Bourbakis, "A survey on wearable sensor-based systems for health monitoring and prognosis," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 1, pp. 1–12, 2010.

[5] R. E. Bouserhal, T. H. Falk, and J. Voix, "In-ear microphone speech quality enhancement via adaptive filtering and artificial bandwidth extension," *The Journal of the Acoustical Society of America*, vol. 141, no. 3, pp. 1321–1331, 2017.

[6] K. Mazur and J. Voix, "A case-study on the continuous use of an in-ear dosimetric device," *The Journal of the Acoustical Society of America*, vol. 133, no. 5, pp. 3274–3274, 2013.

[7] C. Smalt, S. K. Davis, P. Calamia, J. Lacirignola, O. Townsend, C. Weston, and P. Collins, "On-body and in-ear noise exposure monitoring," *The Journal of the Acoustical Society of America*, vol. 141, no. 5, pp. 3732–3732, 2017.

[8] F. Bonnet, J. Voix, and H. Nélisse, "The opportunities and challenges of in-ear noise dosimetry," *Canadian Acoustics*, vol. 43, no. 3, 2015.

[9] H. Nélisse, M.-A. Gaudreau, J. Boutin, J. Voix, and F. Laville, "Measurement of hearing protection devices performance in the workplace during full-shift working operations," *Annals of occupational hygiene*, vol. 56, no. 2, pp. 221–232, 2011.

[10] J. Freitas, A. Teixeira, M. S. Dias, and S. Silva, *An Introduction to Silent Speech Interfaces*. Springer, 2017.

[11] A. Teixeira, N. Vitor, J. Freitas, and S. Silva, "Silent speech interaction for ambient assisted living scenarios," in *International Conference on Human Aspects of IT for the Aged Population*. Springer, 2017, pp. 369–387.

[12] H. Park, M. Kiani, H.-M. Lee, J. Kim, J. Block, B. Gosselin, and M. Ghovanloo, "A wireless magnetoresistive sensing system for an intraoral tongue-computer interface," *IEEE transactions on biomedical circuits and systems*, vol. 6, no. 6, pp. 571–585, 2012.

[13] J. Kim, H. Park, J. Bruce, E. Sutton, D. Rowles, D. Pucci, J. Holbrook, J. Minocha, B. Nardone, D. West *et al.*, "The tongue enables computer and wheelchair control for people with spinal cord injury," *Science translational medicine*, vol. 5, no. 213, pp. 213ra166–213ra166, 2013.

[14] A. Bedri, D. Byrd, P. Presti, H. Sahni, Z. Gue, and T. Starner, "Stick it in your ear: building an in-ear jaw movement sensor," in *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*. ACM, 2015, pp. 1333–1338.

[15] A. Rabaoui, M. Davy, S. Rossignol, and N. Ellouze, "Using one-class svms and wavelets for audio surveillance," *IEEE Transactions on information forensics and security*, vol. 3, no. 4, pp. 763–775, 2008.

[16] J. Portelo, M. Bugalho, I. Trancoso, J. Neto, A. Abad, and A. Serralheiro, "Non-speech audio event detection," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*. IEEE, 2009, pp. 1973–1976.

[17] H. Phan, L. Hertel, M. Maass, and A. Mertins, "Robust audio event recognition with 1-max pooling convolutional neural networks," *arXiv preprint arXiv:1604.06338*, 2016.

[18] J. Schröder, J. Anemiiller, and S. Goetze, "Classification of human cough signals using spectro-temporal gabor filterbank features," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 6455–6459.

[19] H. Sato, K. Abe, S. Ohi, and M. Ohyama, "An automatic classification method for involuntary and two types of voluntary blinks," *Electronics and Communications in Japan*, vol. 100, no. 10, pp. 48–58, 2017.

[20] H. S. Shin, H.-G. Kang, and T. Fingscheidt, "Survey of speech enhancement supported by a bone conduction microphone," in *Proceedings of Speech Communication; 10. ITG Symposium*. VDE, 2012, pp. 1–4.

[21] A. Varga and H. J. Steeneken, "Assessment for automatic speech recognition: Ii. noisex-92: A database and an experiment to study the effect of additive noise on speech recognition systems," *Speech communication*, vol. 12, no. 3, pp. 247–251, 1993.

[22] R. E. Bouserhal, T. H. Falk, and J. Voix, "In-ear microphone speech quality enhancement via adaptive filtering and artificial bandwidth extension," *The Journal of the Acoustical Society of America*, vol. 141, no. 3, pp. 1321–1331, 2017.

[23] M. Sarria-Paja and T. H. Falk, "Fusion of auditory inspired amplitude modulation spectrum and cepstral features for whispered and normal speech speaker verification," *Computer Speech & Language*, vol. 45, pp. 437–456, 2017.

[24] T. Kinnunen and H. Li, "An overview of text-independent speaker recognition: From features to supervectors," *Speech Communication*, vol. 52, no. 1, pp. 12–40, January 2010.

[25] R. Sweetow and C. Pirzanski, "The Occlusion Effect and Ampclusion Effect," *Seminars in Hearing*, vol. 24, no. 4, pp. 333–343, 2003.

[26] M. Dean and F. Martin, "Insert Earphone Depth and the Occlusion Effect," *Am. J. Audiol.*, vol. 9, pp. 131–134, 2000.

[27] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," Dec. 2014, arXiv: 1412.6980. [Online]. Available: http://arxiv.org/abs/1412.6980

[28] E. Khoury, L. El Shafey, and S. Marcel, "Spear: An open source toolbox for speaker recognition based on Bob," in *Proc. ICASSP*, 2014, pp. 1655–1659.

[29] G. R. Doddington, M. A. Przybocki, A. F. Martin, and D. A. Reynolds, "The NIST speaker recognition evaluation - overview, methodology, systems, results, perspective," *Speech Communication*, vol. 31, no. 2-3, pp. 225–254, June 2000.