COCHLEAR HEARING LOSS DETECTION SYSTEM BASED ON TRANSIENT EVOKED OTOACOUSTIC EMISSIONS

Hubert Dietl[1], Stephan Weiss[1]

[1] Dept. Electronics & Computer Science, University of Southampton, UK
{hwd,s.weiss}@ecs.soton.ac.uk

Abstract - The aim of this paper is to evaluate the application of the discrete wavelet transform (DWT) and support vector machines (SVM) to transient evoked otoacoustic emissions (TEOAE) in order to achieve a detection of frequency-specific hearing loss (HL). We introduce a system to determine detection rates between groups of persons with normal hearing, high frequency hearing loss, and pantonal hearing loss. The validity and use of our approach is verified on a different patient group.

I. INTRODUCTION

Transient evoked otoacoustic emissions (TEOAE) are used as a clinical standard procedure to detect cochlear hearing loss [2], and measurement equipment [1] is widely available in hospitals. The analysis of TEOAE is usually performed by an human expert. Recently, signal processing detection systems aiming at an automated detection of cochlear hearing loss have been motivated to assist or replace the human expert. These studies aiming at detection of TEOAE apply discrete wavelet transform (DWT) and neural networks [3],[4]. Here, we introduce a system applying a DWT for feature extraction, an energy reduction for feature selection and support vector machines for classification.

Fig. 1 gives an overview of our system. For the feature extraction, a DWT is applied. To select the features of the data, an energy reduction by 5% to 10% is applied to the transformed data resulting in a reduction of coefficients to be used for classification and aiming at a reduction of noisy coefficients. This approach will be outlined in more detail in Sec. 3, following a description of TEOAE data in Sec. 2. The classification of the data is conducted by a support vector machine (SVM) classifier explained in Sec. 4 more explicitly. In Sec. 5, based on the training data, a support vector classification network is found and applied to the test data group yielding detection rates which describe the performance of the system and can be compared with other studies. Finally, Sec. 6 draws the conclusions.

II. TEOAE

TEOAE are low-level sounds produced by the inner ear as a response to an external acoustic stimulus, which are measured in the outer ear canal by sensitive microphones. This broadband click-stimulus contains frequencies between 0.5 and 5 kHz; these frequencies are reflected in the TEOAE and are generally believed to correspond to frequencies that are perceived by the ear [2].

The TEOAE spectrum is latency-dependent: low frequency components possess a prolonged latency. As the TEOAE data is generally very noisy, it requires averaging which is performed by the measurement equipment [1] by 520 stimulus-synchronously recorded responses per ear. Two studies with each approximately 200 ears from the Universities of Homburg and Heidelberg are used for our study whereby the Homburg data represents the training data and the Heidelberg data the test data group. Each study contains three classes of hearing ability, namely persons with normal hearing (NH), high frequency hearing loss (HF), and pantonal hearing loss (PT).

In the next section, the parameterisation and feature selection of the data is described.

III. FEATURE EXTRACTION AND SELECTION

Due to the transient nature of the signals, previous work on the qualitative analysis of TEOAE has focused on time-frequency (TF) methods, such as filter banks [5], matching pursuit [6], or discrete wavelet transforms [4], whereby a quantitative study w.r.t. the achievable distinction of frequency-specific HL has been performed in [4], based on the DWT. The DWT is a fixed transform based on a "mother wavelet" from which the transformation coefficients are derived by scaling, translation and sampling. Here, we have chosen the Mallat wavelet for which good results have been reported in [4] and which can be used for comparison.

Based on a parameterisation of the data by the DWT, representing the feature extraction of the data, an energy reduction by 5% to 10% is conducted which can be described as feature selection. In more detail, the smallest DWT coefficients whose sum-of-squares makes up for 5% to 10% of the total signal power are discarded. Given the spares nature of the DWT, the proportion of such coefficients relative to the total number of coefficients is much greater than the energy reduction proportion, hence reducing computational effort at subsequent stages. The reduced DWT coefficients are the input for the SVM classifier shown next.

IV. SVM CLASSIFICATION

In the following, we briefly explain SVM, [7],[8]. A support vector learning machine calculates a classification network for

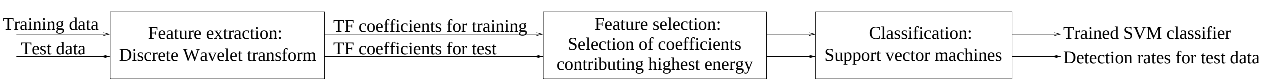

Fig. 1. Overview of the detection system for cochlear hearing loss.

a given training error and minimises the capacity of the learning machine to be able to yield a good classification for test data which is in contrast to neural networks that attempt to minimise the training error for a constant capacity determined by a certain structure choice. We consider a two class classification problem, mutually distinguishing between the three classes of hearing ability groups NH, HF and PT. The training data originates from the Homburg data, while the test data comprises the Heidelberg measurements.

The training data is described as a set of training vectors $\{\mathbf{p}_i\}_{i=1\ \ldots\ M}$ with corresponding binary labels $S_i = 1$ for the one class, e.g. NH, and $S_i = -1$ for the second class, e.g. HF. The SVM conducts a classification of a test vector $\mathbf{t}$ by assigning a label $\hat{S}$ by calculating

$$\hat{S} = \text{sign}(f(\mathbf{t})) \qquad \text{with} \qquad f(\mathbf{t}) = \sum_i \alpha_i S_i K(\mathbf{t}, \mathbf{p}_i) + b. \tag{1}$$

The α_i are called weights and b is the bias, which are SVM parameters and adopted during training by maximising

$$L_D = \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j S_i S_j K(\mathbf{p}_i, \mathbf{p}_j) \tag{2}$$

under the constraints

$$0 \leq \alpha_i \leq C \qquad \text{and} \qquad \sum_i \alpha_i S_i = 0 \tag{3}$$

with C being a positive constant which weighs the influence of training errors. $K(\cdot, \cdot)$ is called kernel of the SVM. If there is a solution for α_i, a value for b is determined. There are several commonly used kernels for SVM, which give some flexibility for the underlying application. Many implementations of kernels can be found in literature, whereby two popular ones are Gaussian and polynomial kernels.

If $K(\cdot, \cdot)$ is positive definite, (2) and (3) is a convex quadratic optimisation problem, which converges towards the global optimum assuringly. This optimisation can be quite demanding in terms of computation time for real-world problems, and therefore, sophisticated algorithms like sequential minimal optimisation (SMO) [7] are used for the solution.

Usually $\alpha_i = 0$ for the majority of i and thus the summation in (1) is limited to a subnet of the $\mathbf{p}_i$, which therefore is called the set of support vectors. For our application, the polynomial kernel of order 2 was found to be the best compromise between computational time, generalisability and yielding a good detection rate for the training data.

To find a significant value for the training error C, a leave-one-out (l-o-o) estimation of the error rate is applied as follows: From the training samples, remove the first example. Train the SVM on the remaining samples. Then test the removed example. If the example is classified incorrectly, it is said to produce a leave-one-out error. In [7], an approach to estimate the maximum l-o-o error is shown avoiding training the SVM more than once, which is also used for our study. By changing the value for C stepwise, the minimum for the l-o-o error is found determining the SVM classification network with a quadratic polynomial kernel. The second parameter for our application for which the l-o-o error is minimised is the energy reduction by 5% to 10%.

V. RESULTS AND DISCUSSION

Having described the detection methods and the data used for our system, we present the results in the following. We use the term sensitivity and specificity to address a correct allocation for a person to a hearing group. E.g. when distinguishing between NH and HF, sensitivity refers to the percentage of persons correctly allocated to the normal hearing group and specificity describes the patients correctly allocated to the group with high frequency HL. For the training data, the values for sensitivity and specificity are all close to 100% for the three possible distinction cases NH vs HF, NH vs PT and HF vs PT. The results for the test data are illustrated in Tab. 1. Also, Tab. 1 indicates the specificity results in [4] for the sensitivity values obtained with our study for comparison. As mentioned in the previous section, the results are optimised w.r.t. the retained energy.

	Test data			results in [4]	
group distinction	sensitivity	specificity	energy reduction	sensitivity	specificity
NH — HF	54%	81%	9%	54%	89%
NH — PT	76%	99%	7%	76%	98%
HF — PT	66%	76%	5%	66%	91%

Tab. 1. Detection rates for the test data; comparison with [4].

The table shows that our approach yields slightly better results for the NH vs PT case than in [4]. The case HF vs PT yields less significant results, which is the most difficult case to separate according to [4]. Moreover, the specificity for the case NH vs HF for our system is not as significant as for the study in [4]. Recapitulating, it can be said that our system separates data with pronounced differences better than e.g. [4]. On the other hand, for less pronounced differences, the performance of our system degrades.

VI. CONCLUSIONS AND ACKNOWLEDGEMENTS

We have introduced a system for the detection of cochlear hearing loss based on signal processing techniques such as wavelet transform and support vector machines. The achieved results were compared with a similar study. The comparison showed some drawbacks of the system. However, when keeping these findings in mind, the system is competitive for the detection of cochlear hearing loss based on otoacoustic emissions.

We would like to thank Profs. Ulrich Hoppe and Sebastian Hoth of the Universities of Homburg and Heidelberg for kindly providing the data.

REFERENCES

[1] Otodynamics Ltd., Hatfield, Herfortshire, UK, *ILO OAE Instrument User Manual*, 5a ed., October 1997.

[2] M. Robinette and T. Glattke, *Otoacoustic Emissions: Clinical Applications*. Thieme Medical Pub, 2. ed., 2001.

[3] G. Buller and M. Lutman, "Automatic classification of transiently evoked otoacoustic emissions using an artificial neural network," *British Journal of Audiology*, vol. 32, pp. 235–247, 1998.

[4] S. Weiss, U. Hoppe, M. Schabert, and U. Eysholdt, "Wavelet Analysis of Transient Evoked Otoacoustic Emissions for Differential Diagnosis of Cochlear Hearing Loss," in *Asilomar Conference on Signals, Systems, and Computers*, (Monterey, CA), November 2001.

[5] P. Ravazzani, G. Tognola, F. Grandori, and J. Ruohonen, "Two-Dimensional Filter to Facilitate Detection of Transient-Evoked Otoacoustic Emissions," *IEEE Transactions on Biomedical Engineering*, vol. 45, pp. 1089–1096, 1998.

[6] K. J. Blinowska, P. J. Durka, A. Skierski, F. Grandori, and G. Tognola, "High Resolution Time-Frequency Analysis of Otoacoustic Emissions," *Technology and Health Care*, vol. 5, pp. 407–418, 1997.

[7] V. Vapnik, *Statistical Learning Theory*. New York: Wiley: Cambridge University Press, 1998.

[8] C. Bahlmann, B. Haasdonk, and H. Burkhardt, "On-line Handwriting Recognition with Support Vector Machines – A Kernel Approach," in *Proceedings of the 8th Int. Workshop on Frontiers in Handwriting Recognition (IWFHR)*, pp. 49–54, 2002.