Fast Detection of Wave V in ABRs Using a Smart Single Sweep Analysis System

Daniel J. Strauss[1,3], Wolfgang Delb[2], Peter K. Plinkert[2], Helmut Schmidt[3]
[1] Key Numerics, Saarbruecken, Germany
[2] Saarland University Hospital, Homburg/Saar, Germany
[3] Institute for New Materials, Saarbruecken, Germany

Abstract— **The analysis of auditory brainstem responses (ABRs) is accepted to be the most reliable method for the objective diagnosis and quantification of hearing loss in newborns. However, in currently available setups, a large number of sweeps has to be averaged to obtain a meaningful signal at low stimulation levels due to a poor signal–to–noise ratio.**

In this study, we present a new approach to the detection of wave V in ABRs using a smart single sweep analysis system. A small number of sweeps is decomposed by optimized tight frames and evaluated by a kernel based novelty detection machine. This hybrid supervised learning scheme is combined with an inter–sweep dissimilarity tracing for the final decision making.

At the challenging stimulation level of 30 dB, our system reached a reasonable specificity and sensitivity for the detection of wave V in a fraction of the measurement time of conventional schemes.

Keywords– **Hearing Screening, Auditory Brainstem Responses, Single Sweep Analysis, Adapted Filter Banks, Wavelets, Kernel Machines, Smart Systems**

I. Introduction

The detection of wave V in auditory brainstem response (ABR) measurements is the standard method in the objective diagnosis and quantification of hearing loss in children. Due to a lack of cooperation in young children, often a sedation or a narcosis is necessary to obtain reliable measurements. As the diagnosis and treatment of pediatric hearing loss has to be done as early as possible in order to avoid a serious delay in speech and intellectual development [1], [2], the examination of hearing under narcosis has often to be applied in very young children. As these children are often neonatal intensive care unit patients, there is an increased risk of complication due to the general anesthesia.

A large number of universal newborn hearing screening (NHS) programs has been established so far [3], [4]. The screening methods that are currently available include otoacoustic emissions [4] and ABR measurements [5] at a fixed stimulation level (35 dB HL). Due to methodological (otoacoustic emissions) and cost problems (ABR measurements), only the exclusion of a hearing loss and not the determination of a given hearing threshold is possible using these methods. Also the specificity of these methods is less than 100 %. Therefore NHS programs are mostly organized as 2 or 3 stage procedures which causes the problem that many children get lost to follow up, resulting in a loss of effectiveness of these programs. An automated and objective audiological method able to quantify the hearing threshold within a short measurement time (less than 2 or 3 minutes) without the need of sedation or general anesthesia would be able to drastically reduce the number of audiological examination under anesthesia that have to be performed. Also the determination of the auditory threshold could be done in the first stage of screening programs. Therefore such a method would make a second or third stage unnecessary, reduce the cost of universal hearing screening programs, and would thus help to increase the effectiveness of NHS–programs.

In this paper, we present a new approach to the fast detection of wave V in ABRs employing a smart single sweep analysis system which is based on adapted signal decompositions, machine learning, and single sweep dissimilarity tracing.

II. Methodology

Our scheme consists basically of 4 major stages: (1) data acquisition, (2) feature extraction, (3) novelty detection, and (4) dissimilarity tracing. These steps are summarized below and finally assembled at the end of the this section.

Methodically, stage (2) and (3) can be summarized as hybrid tight frame–kernel machine proposed in [6] as general approach for biosignal recognition. This scheme is just shortly sketched below, adjusted to our problem. For details we refer to [6], [7].

A. Data Segments

The study group consisted of 16 normal hearing probands (threshold <10dB HL between 0.5kHz and 6kHz). This group was decomposed into a learning group of 6 probands and an independent test group of 10 probands. For the final performance assessment of our system, the test group is extended by 10 deaf subjects.

Auditory evoked potentials were obtained using a commercially available device (Evostar, Pilot–Blankenfelde, Berlin, Germany) in a sound–proof chamber. In each measurement, 15 clicks per second were presented monaurally at an intensity of 30, 40, 50, and 70 dB (HL) with an inter–stimulus interval of 60ms. Artifacts were excluded from the analysis by the internal artifact filter of the system.

Single sweeps, i.e., the responses to individual clicks, were recorded using electrodes placed at the neck, the vertex and the upper forehead, respectively. Electrode impedances were below 5 kΩ in all measurements (filter: 0.15kHz–3kHz, sampling frequency: 20kHz).

B. Feature Extraction in Single Sweeps

A hybrid wavelet–support vector classification has been introduced in [7] which employs lattice structure based wavelet and frame decompositions for feature extraction tasks in waveforms which are tailored for support vector classifiers with radial kernels. In particular, it provides a feature extraction which allows for an inclusion of a priori knowledge and leads to a maximal margin of the scheme and is thus conform with the maximal margin theorem [8] of statistical learning theory.

Our objective in this paper is novelty detection instead of binary classification as it is better suited for abnormality detection [9]. Nevertheless, the feature extraction stage is closely related to classification.

The original wavelet–support vector classifier as proposed in [7] relies on *multilevel concentrations*. $\xi(\cdot) = \|\cdot\|_{\ell^p}^p$ $(1 \leq p < \infty)$ of coefficient vectors of adapted wavelet or frame decompositions as feature vectors, i.e, scale features. These feature vectors incorporate the information about local instabilities in time as a priori information. For the classification of ABRs, we also include the morphological information of the waveforms as features as the discriminant information which separates the physiological and pathological sweeps is also reflected in the transient evolution of ABRs.

Since we are interested in a shift–invariant classification scheme, we may only evaluate the morphology of ABRs as a whole and not the exact latency of transient features. A possible way to realize this is by the use of entropy which is already employed to evaluate the subbands of wavelet and wavelet packet decompositions for the purpose of signal compression, see [10] and [11]. When using an appropriate entropy in connection with the tight frame decompositions, it is invariant to shifts of the sweeps. We define the entropy of a sequence $\mathbf{x} \in \ell^2$ by

$$E(\mathbf{x}) = -\sum_{n\in\mathbb{Z}} \frac{|x[n]|^2}{\|\mathbf{x}\|_{\ell^2}^2} \ln \frac{|x[n]|^2}{\|\mathbf{x}\|_{\ell^2}^2}. \tag{1}$$

Note that $E(\cdot)$ is also the well known *Shannon entropy* [12] but one where the probabilistic events are replaced by normalized energies of the samples, i.e., we do not deal with the probabilistic concept of the entropy here.

For a fixed ABR sweep $\mathbf{x}$, we define the function

$$\begin{aligned}\zeta^{\mathbf{x}}(\vartheta) &= (\zeta_1^{\mathbf{x}}(\vartheta), \ldots, \zeta_{2J}^{\mathbf{x}}(\vartheta)) \\ &= \left(\|\mathrm{d}_1^\vartheta\|_{\ell^1}, \ldots, \|\mathrm{d}_J^\vartheta\|_{\ell^1}, E(\mathrm{d}_1^\vartheta), \ldots, E(\mathrm{d}_J^\vartheta)\right),\end{aligned}$$

and set $\zeta^i(\vartheta) := \zeta^{\mathbf{x}_i}(\vartheta)$ $(i = 1, \ldots, M)$. Here d_j^ϑ denotes the coefficients of a shift–invariant lattice structure based octave–band tight frame decomposition, parameterized by the angle vector ϑ, see [7]. The number J is the decomposition depth. The first J elements of this feature vector carry multilevel concentration of the subbands in ℓ^1, i.e., a scale information. The second J elements carry the morphological information reflected in the entropy as defined in (1). Note that $\zeta_i(\vartheta)$ is totally invariant against shifts of the individual sweeps. We used decomposition level 2 to 5 in this study as these levels carried the substantial signal information.

C. Novelty Detection

Suppose we are given a set of M samples and a description is required. We try to find sphere with a minimum volume, containing all data in the hard case (no outliers in learning set) and most of the data in the soft case (the learning set may contain outliers). Instead of constructing this sphere in the original space, we construct it in a high dimensional feature space which is induced by a kernel of a reproducing kernel Hilbert space [13]. All patters which lay outside the sphere are detected as novel instances which do not correspond to the learned class [9], [14]. The minimal sphere can be obtained by the following optimization problem:

$$\min_{\mathbf{a}\in\mathcal{F}_K, R\in\mathbb{R}, \mathbf{u}\in\mathbb{R}^M} R^2 + \lambda \sum_{i=1}^{M} u_j \tag{2}$$

subject to

$$\begin{aligned}\|\Phi(\zeta_i(\vartheta)) - \mathbf{a}\|^2 &\leq R^2 + u_i \quad (i = 1, \ldots, M), \\ u_i &\geq 0 \quad (i = 1, \ldots, M).\end{aligned} \tag{3}$$

where the $\Phi : \mathcal{X} \subset \mathbb{R}^J \longrightarrow \mathcal{F} \subset \ell^2$ denotes the feature map from the pattern space to kernel feature space, $\mathbf{a}$ is the center of sphere [9].

For the embedding of the feature extraction in the minimal sphere approach above, the objective is now to find optimal lattice angels $\hat{\vartheta}$ such that a learning set of M sweeps $\mathcal{A}(\vartheta) = \{\zeta^i(\vartheta)) \in \mathcal{X} : i = 1, \ldots, M\}$ is as compact as possible in the feature space, i.e.,

$$\hat{\vartheta} = \arg\min_{\vartheta\in\mathcal{P}^2} \left\{ \sum_{i=1}^{M} \|\Phi(\zeta^i(\vartheta)) - \Xi\|_{\mathcal{F}_k}^2 \right\},$$

where $\mathcal{P}^2$ denotes the lattice parameter space for filters of order 5 (see [7] for details of the lattice parameterization) and Ξ the feature center. For radial kernels of the SVM, problems of this type can be transformed from the feature to the original space and solved by genetic algorithms, see [7].

D. Single Sweep Dissimilarity Tracing

Our single sweep dissimilarity tracing in based on the fact that ABRs are due to neural synchronizations of the spontaneous brain activity [15]. This can be seen as a transition from a disordered or high entropy state (the spontaneous activity) to an ordered or low entropy state (the neural response upon the auditory stimulation), a fact that was exploited in [16] for the final analysis of averaged responses in view of binaural interaction.

Such entropy considerations imply that we have a high decorrelation of the individual sweeps in the pathological case. On the other hand, a sequence of sweeps exhibit an inner correlation due to the neural synchronization as response to the auditory stimulation in the physiological case.

To take this into account, we determine the degree of dissimilarity between the feature vectors of consecutive sweeps. The dissimilarity analysis of the feature vectors has the following advantage: additionally to the high–frequency noise filtering,

latency jitters of the responses can be excluded from the analysis due the shift–invariance of the feature vectors. Both factors corrupt also physiological responses.

As dissimilarity measure, we apply the *J–Divergence* [17] of consecutive sweeps which is the symmetric version of the cross–entropy. Mathematically more precisely, for a fixed ϑ two feature vectors $\zeta^i(\vartheta)$ and $\zeta^{i+1}(\vartheta)$ of consecutive sweeps $\mathbf{x}_i$ and $\mathbf{x}_{i+1}$ are normalized in ℓ^1, and the degree of dissimilarity is determined by

$$\mathcal{D}_{\text{JDIV}}\left(\zeta^i(\vartheta), \zeta^{i+1}(\vartheta)\right) = \sum_{n=1}^{2J} \zeta_n^i(\vartheta) \log \frac{\zeta_n^i(\vartheta)}{\zeta_n^{i+1}(\vartheta)} + \sum_{n=1}^{2J} \zeta_n^{i+1}(\vartheta) \log \frac{\zeta_n^{i+1}(\vartheta)}{\zeta_n^i(\vartheta)}; \quad (4)$$

with the convention $\log 0 = -\infty$, $\log(\gamma/0) = \infty$ for $\gamma > 0$, $0 \cdot (\pm\infty) = 0$.

E. The Assembled Detection Scheme

The assembled recognition scheme in shown in Fig. 1. There are two additional averaging stages. The partial averaging stage after the adapted tight frame feature extraction improves the robustness of the scheme by averaging a small number of feature vectors to one. The averaging stage after the dissimilarity evaluation computes the mean dissimilarity of the sweep sequence which is used for the decision making in this stage. If the mean dissimilarity is below a predefined threshold $\tau_{\mathcal{D}}$ we have a physiological detection here and a pathological output otherwise. This threshold can be defined from the learning set by setting $\tau_{\mathcal{D}}$ equal to the maximal dissimilarity in the learning set plus its standard deviation.

The output of the novelty detector is physiological if 90% of the partially averaged feature vectors are within the sphere, i.e., detected as normal instances.

The outputs of the dissimilarity evaluation and the novelty detector have both to be physiological in order to get a final physiological decision. This optimizes the system with respect to the sensitivity which is indispensable for its use in screening programs.

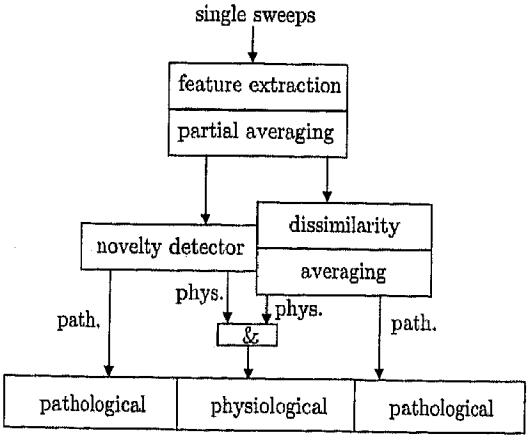

Fig. 1. The assembled detection scheme.

III. Results

In Fig. 2 we have shown a single sweep sequence of 1000 sweeps, partially averaged over 10 sweeps to illustrate how traces of wave V look at an stimulation level of 30 dB. In Fig. 2 (top) we have an physiological example and in Fig. 2 (bottom) an pathological example. The trace of wave V for the physiological example is clearly noticeable in this figure as well as the decorrelation of the sweeps in the pathological case.

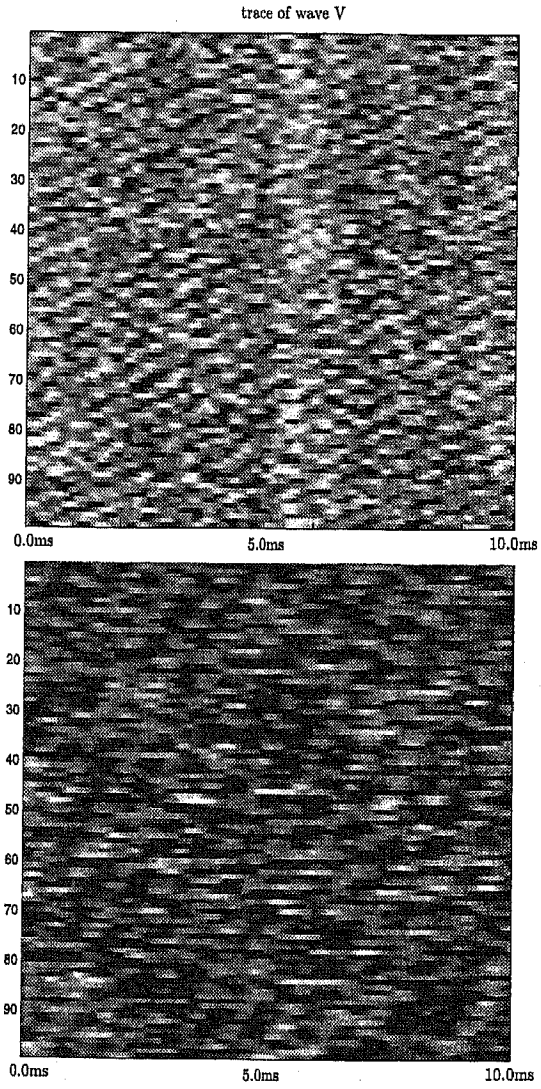

Fig. 2. The single sweep traces for a physiological (top) and a pathological (bottom) example. The normalized amplitude is coded in colors ranging from black (0) to white (1).

In Fig. 3 we have shown how the clustering in the feature space depending on the lattice angles. Only the black regions correspond to an effective frame decomposition which results in a compact representation in the feature space.

For the performance evaluation, our system was initialized, i.e., adapted and learned, with 250 sweeps from each of the 6 probands of the learning group using a Gaussian kernel with a standard deviation of 1.0 (determined by validation) for the construction of the novelty detector. For the testing, our initialized system was applied to a total of 200 sweeps

of each of the 20 data segments in the independent test set (10 physiological and 10 pathological data sets). A number of 10 feature vectors were partially averaged after the feature extraction such that 20 patterns of each data segment were supplied to the novelty detector and the dissimilarity evaluation. For the discrimination of the physiological and the pathological sweep sequences, our system achieved 100 % sensitivity and 90 % specificity based on just 200 sweeps from each data segment at the most challenging simulation level of 30 dB. In conventional systems, a number of 2000 sweeps is usually averaged to obtain a meaningful signal used for the further analysis [15]. Consequently, with the proposed system, just 10 % of the conventional measurement time is necessary to obtain a final machine based decision making which is complete automated and truly objective. It is worth to emphasize that an automated detection of wave V in ABRs at a stimulation level of 30 dB is very challenging, even with 2000 sweeps for conventional schemes.

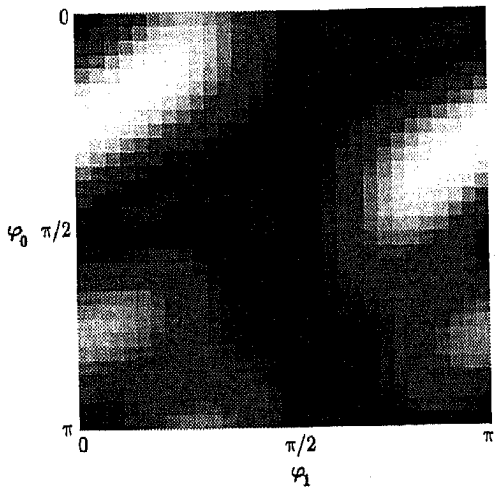

Fig. 3. The clustering the feature space depending on the lattice angles $\vartheta = (\varphi_0, \varphi_1)$. The normalized clustering is coded in colors ranging from black (high clustering) to white (low clustering).

IV. Discussion and Conclusions

We have presented a new approach to the detection of wave V in ABRs using a smart single analysis system. Our scheme is based on an adapted feature extraction in single sweeps by tight frames, a kernel based novelty detector, and a dissimilarity tracing of sweep sequences, taking the larger entropy in the pathological case into account.

Using this system, a discrimination of physiological from pathological probands by the detection of wave V was possible with 100 % sensitivity and 90 % specificity. Compared to conventional schemes, just 10 % of the measurement time is needed to get this result in a completely automated manner.

The clinical use of this method, able to quantify the hearing threshold within a short measurement time, without the need of sedation or general anesthesia would drastically by reduce the number of audiological examination under anesthesia that have to be performed in newborns.

Nevertheless, improvements seem still to be possible. For instance, the realization of a patient adapted system seems also to be implementable with formalism below. In this case, the system is initialized with the spontaneous brain activity, i.e., without clicks, and every response to a click is detected as novel instance. Such a system seems to be more robust as it has not to overcome the inter–patient variability. The authors also follow this way.

It is finally concluded, that a very fast detection of wave V in ABRs at low stimulation levels can be implemented by the proposed system and might be a further step to improve the effectiveness of NHS–programs.

References

[1] C. Yoshinaga-Itano, "Benefits of early intervention for children with hearing loss," *Otolaryngol. Clin. North Am.*, vol. 32, pp. 1089–1102, 1999.

[2] C. Yoshinaga-Itano and J. S. Gravel, "The evidence fr universal newborn hearing screening," *American Journal of Audiology*, vol. 10, pp. 62–4, 2001.

[3] B. R. Vohr and A. B. Maxon, "Screening infants or hearing impairment," *J. Pediatr.*, vol. 128, pp. 710–713, 1996.

[4] W. Delb, D. Merkel, K. Pilorget, J. Schnitt, and P. K. Plinkert, "Effectiveness of a teoae based screening program. can a patient tracking system effectively be organized using modern information technology and central data management ?," *Eur. Arch. Otolaryngol.*, 2003, DOI: 10.1007/s00405-003-0662-3.

[5] D. L. Steward, A. Mehl, Thomson V. Hall, J. W. III, and J. Carrol, M. Hamlett, "Infant hearing impairment and universal hearing screening," *J. Perinatol.*, vol. 20, pp. 128–131, 2000.

[6] D. J. Strauss, W. Delb, P. K. Plinkert, and J. Jung, "Hybrid wavelet–kernel based classifiers and novelty detectors in biosignal processing," in *Proceedings of the 25th Silver Anniversary International Conference of the IEEE Engineering in Medicine and Biology Society*, Cancun, Mexico, 2003, pp. 2765–2769.

[7] D. J. Strauss and G. Steidl, "Hybrid wavelet–support vector classification of waveforms," *Journal of Computational and Applied Mathematics*, vol. 148, pp. 375–400, 2002.

[8] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, NY, 1995.

[9] D. M. J. Tax and R. P. W. Duin, "Data domain description using support vectors," in *Proceedings of European Symposium on Artificial Neural Networks '99*, Brugge, 1999.

[10] R. R. Coifman and M. V. Wickerhauser, "Entropy based algorithms for best basis selection," *IEEE. Trans. on Information Theory*, vol. 32, pp. 712–718, 1992.

[11] M. V. Wickerhauser, *Adapted Wavelet Analysis form Theory to Software*, A. K. Peters, Ltd., Wellesley, MA, 1994.

[12] M. T. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, NY, 1991.

[13] G. Wahba, "Support vector machines, reproducing kernel hilbert spaces and the randomized GACV," in *Advanced in Kernel Methods – Support Vector Learning*, B. Schölkopf, C. Burges, and A. J. Smola, Eds., Cambridge, MA, 1999, pp. 293–306.

[14] Asa Ben-Hur, David Horn, Hava T. Siegelman, and Vladimir Vapnik, "Support vector clustering," *Journal of Machine Learning Research*, pp. 125–137, 2001.

[15] J. W. Hall, *Handbook of Auditory Evoked Responses*, Allyn and Bacon, Needham Heights, MA, 1992.

[16] D. J. Strauss, W. Delb, and P. K. Plinkert, "Analysis and detection of binaural interaction in auditory brainstem responses by time–scale representations," *Computers in Biology and Medicine*, 2004, In press.

[17] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Statist.*, vol. 22, pp. 79–86, 1951.