Original Paper

Audiology & Neuro-Otology

Audiol Neurootol 2005;10:6–21
DOI: 10.1159/000081544

Received: January 13, 2004
Accepted after revision: May 18, 2004
Published online: October 14, 2004

Automated Analysis of the Auditory Brainstem Response Using Derivative Estimation Wavelets

Andrew P. Bradley[a] Wayne J. Wilson[b]

[a]Cooperative Research Centre for Sensor Signal and Information Processing (CSSIP), School of Information Technology and Electrical Engineering, and [b]Division of Audiology, School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane, Australia

Key Words
Auditory brainstem response · Automatic peak detection

Abstract
In this paper, we describe an algorithm that automatically detects and labels peaks I–VII of the normal, suprathreshold auditory brainstem response (ABR). The algorithm proceeds in three stages, with the option of a fourth: (1) all candidate peaks and troughs in the ABR waveform are identified using zero crossings of the first derivative, (2) peaks I–VII are identified from these candidate peaks based on their latency and morphology, (3) if required, peaks II and IV are identified as points of inflection using zero crossings of the second derivative and (4) interpeak troughs are identified before peak latencies and amplitudes are measured. The performance of the algorithm was estimated on a set of 240 normal ABR waveforms recorded using a stimulus intensity of 90 dBnHL. When compared to an expert audiologist, the algorithm correctly identified the major ABR peaks (I, III and V) in 96–98% of the waveforms and the minor ABR peaks (II, IV, VI and VII) in 45–83% of waveforms. Whilst peak II was correctly identified in only 83% and peak IV in 77% of waveforms, it was shown that 5% of the peak II identifications and 31% of the peak IV identifications came as a direct result of allowing these peaks to be found as points of inflection.

Introduction

The auditory brainstem response (ABR) remains the most widely and successfully used auditory evoked potential in clinical practice [Jewett and Williston, 1971; Jacobson, 1985; Hall, 1992; Delgado and Ozdamar, 1994; Neu et al., 1999; Battista et al., 2000; Sininger et al., 2001]. Part of this success stems from the relative ease in which the ABR is analysed, that is, most clinicians simply identify the peaks of interest (primarily peaks I, III and V) and then compare their latencies, and occasionally their amplitudes, to matched normative data.

Whilst the standard ABR analysis process is relatively simple, it is also manual, usually being completed by an audiologist or ABR technician. However, manual peak labelling results in at least two significant limitations. First, there is the potential for failure when too many ABR waveforms must be analysed either in too short a period of time, as in neonatal threshold assessment or intra-operative monitoring, or over too long a period of time, as in intensive care unit neurological monitoring [Jacobson, 1985; Hall, 1992]. Second, there is the potential for mislabelling in the hands of a novice operator [Gabriel et al., 1980; Hall, 1992]. A possible solution to these problems is to automate, or semi-automate, the peak labelling process [Pool and Finitzo, 1989; Delgado and Ozdamar, 1994].

Previous attempts to automate ABR wave labelling have varied depending on whether the aim was simply to

KARGER
Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1420–3030/05/0101–0006$22.00/0

Accessible online at:
www.karger.com/aud

Dr. Wayne Wilson, PhD, MAudSA, CCP
Division of Audiology, School of Health and Rehabilitation Sciences
University of Queensland
Brisbane QLD 4072 (Australia)
Tel. +61 7 3365 1797, Fax +61 7 3365 1877, E-Mail w.wilson@uq.edu.au

detect the presence (or absence) of a reliable ABR waveform at low stimulus levels, as required for hearing threshold estimation [Boston, 1989; Cebulla et al., 2000; Salvi et al., 1987; Sininger et al., 2001] or to measure individual ABR peaks at varying stimulus levels and then interpret them via latency-intensity analysis [Delgado and Ozdamar, 1994; Vannier et al., 2002]. The former task is often posed as a *signal detection* task, while the latter is often posed as a *signal analysis* task. Therefore, a method designed for signal detection may not be suitable for signal analysis and vice versa. In this paper, we concentrate on the signal analysis task: that is, quantifying ABR waveforms obtained from suprathreshold stimuli. Used alone, the algorithm is directly applicable to intra-operative monitoring and intensive care unit neurological monitoring. However, it may be beneficial to combine the algorithm with a preliminary signal detection step in order for it to be applied to hearing threshold estimation (with stimulus levels near threshold). In addition, we note that only at suprathreshold intensity levels are all seven Jewett waves likely to be present [Vannier et al., 2002]. As stimulus intensity is reduced, the signal-to-noise ratio (SNR) of the waveforms will be reduced (as ABR amplitude is reduced against an approximately constant background noise) [Boston, 1989].

Methods previously used for analysis of ABR peaks in suprathreshold waveforms have included: detecting the local maxima in low-pass filtered waveforms [Pratt et al., 1989], designing filters *matched* to the peaks of interest [Vannier et al., 2001; Vannier et al., 2002; Woodworth et al., 1983] and applying artificial neural networks (ANNs) [Alpsan et al., 1994; Popescu et al., 1999; Tian et al., 1997]. However, the most commonly used method has been to band-limit the ABR waveform and then to estimate its gradient (first derivative). The peaks and troughs of interest are then found via the zero crossings in the gradient waveform (with the peaks being distinguished from the troughs by the direction of the zero crossing) [Boston, 1989; Delgado and Ozdamar, 1994; Fridman et al., 1982; Gabriel et al., 1980; Pool and Finitzo, 1989; Popescu et al., 1999]. Once candidate zero crossings are found, individual peaks can be labelled according to expected latency and amplitude criteria [Gabriel et al., 1980; Pool and Finitzo, 1989], estimates of the amount of noise present [Boston, 1989] or previous knowledge of peak latencies from higher stimulus intensity waveforms [Vannier et al., 2002].

In addition to band-limiting the waveform and estimating its gradient, multiresolution detection methods have also been proposed that initially detect the primary ABR peak (V), also known as the pedestal peak, using a narrowband filter centred around 100 Hz. More accurate estimates of the peak latency can then be made using a slightly broader band-pass filter centred on 500 Hz [Pratt et al., 1989; Delgado and Ozdamar, 1994; Popescu et al., 1999]. This *coarse-to-fine* [Mallat, 1999] strategy improves detection at low stimulus intensities as the narrowband filter has an improved SNR, whilst the broader band filter has improved temporal resolution [Marr and Hildreth, 1980]. However, it should be noted that at high stimulus levels, the SNR may be such that a multiresolution approach is not required as peaks can be reliably detected with a single (broadband) filter.

Whilst many previous systems have used various multiresolution detection methods, only one [Popescu et al., 1999] has formalised this process using wavelet analysis [Daubechies, 1992; Mallat, 1999]. Because of this, the majority of the previous systems can be criticised for being potentially ad hoc and therefore suboptimal. For example, the filters used in Delgado and Ozdamar [1994] were single cycles of a truncated sinusoid of various frequencies (200, 500 and 900 Hz). While these filters meet the primary condition for being a wavelet (they integrated to zero), they are discontinuous in gradient and so have poor time-frequency resolution. Similarly, the multiresolution method for detecting peak V described in Pratt et al. [1989] used filters with approximately the same bandwidth and so the low frequency filter had the same SNR and temporal resolution as the medium frequency filter. In addition, it has been common practice to apply the band-limiting and derivative estimation filters in separate operations [Gabriel et al., 1980; Fridman et al., 1982; Delgado and Ozdamar, 1994]. Clearly, as both differentiation and low-pass filtering are linear operations, it would be preferable to combine them into a single filtering operation as is the case in wavelet analysis [Daubechies, 1992; Mallat, 1999].

Another limitation of many of the previous automated wave labelling systems has been that they have only attempted to identify the primary peaks of the ABR (I, III and V) [Pool and Finitzo, 1989; Pratt et al., 1989] or alternatively the primary IV-V complex [Popescu et al., 1999]. None of them attempts to identify waves II, IV, VI or VII. In addition, it is not clear in a number of these previous systems how they can be adapted to differing acquisition parameters (e.g. stimulus frequency or polarity) and to waveform abnormalities (usually increased interwave latencies).

In view of these limitations, this paper describes a peak-trough finding algorithm that extends previous work by:

(1) utilising wavelet analysis to efficiently estimate the first derivative of the ABR waveform to detect *all* seven peaks (I–VII);

(2) utilising the second derivative of the ABR waveform to detect peaks II and IV when they cannot be found via the first derivative, and

(3) providing a small set of easily understood parameters that can be used to adapt the algorithm to specific waveform morphologies.

Materials and Methods

Subjects

Sixty subjects (22 male and 38 female) aged between 18 and 55 years were convenience sampled from the general, adult population of Gauteng, South Africa. All subjects demonstrated pure tone audiometry thresholds of 20 dBHL or better at 250, 500, 1000, 2000, 4000 and 8000 Hz and normal acoustic immittance measurements (as defined by Northern [1980]) bilaterally. They also had no self-reported history (past or present) of auditory dysfunction, neural dysfunction, head injury or current medication that could affect the ABR and their ABR waveforms were considered within normal limits by standard normative databases [Hall, 1992].

ABR Recording

All ABR recordings were conducted in sound-treated and electrically shielded booths at a large university hearing clinic. Subjects initially filled out a pretest questionnaire and underwent pure tone audiometry, using a two-channel Grason-Stadler Incorporated GSI 10© diagnostic audiometer and acoustic immittance testing, using a GSI 33© acoustic immittance meter. If the pre-ABR selection criteria were met, ABR testing was conducted using a Navigator© Evoked Potential unit (Biologic Systems Incorporated) running Evoked Potential© software, version 4.31 for DOS©.

To elicit the ABR waveforms, alternating click stimuli, driven by an electrical square wave of 100 μs duration, were presented at a rate of 21 clicks/s to each subject's test ear via EAR 3A insert phones (giving a 0.8-ms delay for onset of the ABR waveform) at 90 dBnHL. The non-test ear was masked using white noise at 50 dBnHL in order to prevent the non-test ear from contributing to the ABR. All levels were as indicated by the Biologic Evoked Potential software, which had undergone factory calibration together with the insert phones. No attempt was made to measure the actual dB SPL at the eardrum.

To record the ABR waveforms, disposable silver-silver chloride electrodes were placed on each subject's high forehead (non-inverting), test ear earlobe (inverting) and non-test ear earlobe (ground). Impedances were maintained at <5 kΩ as measured by the biologic 'impedance check' function (2 μA at 20 Hz). Other recording parameters were: number of averages = 2048; amplifier gain = 150000 (sensitivity or artefact level ± 16.3 μV); band-pass filter = 30–3000 Hz; notch filter = 50 Hz; analogue to digital conversion accuracy = 8 bit (the maximum allowable on the Biologic Evoked Potential software), and time epoch = 15.0 ms (giving a sampling frequency of 34.133 kHz). To remove possible stimulus artefact, the first 20 sample points were preset ('blocked') to zero.

Two ABR waveforms were recorded from each ear of each subject (giving a total of 240 waveforms) and transferred in ASCII format to a personal computer running Mathworks Matlab© software version 6.5. The proposed peak-finding algorithm was applied as a Matlab M-file.

The Algorithm

For each ABR waveform, the peak-finding algorithm proceeded in three steps, with the option of a fourth:

Step 1: all candidate peaks in the ABR waveform were identified using a first derivative estimation wavelet;

Step 2: peaks I–VII were identified from the candidate peaks using latency and morphology criteria;

Step 3: if required, peaks II and IV were identified as points of inflection using a second derivative estimation wavelet, and

Step 4: interpeak troughs were identified and the final peak latencies and amplitudes calculated.

Flowcharts for this process are shown in figures 1 and 2 and the details of this process are described below.

Step 1: Identification of all Candidate Peaks and Troughs

The latencies and amplitudes of all peaks and troughs (called 'candidate' peaks and troughs) were calculated as follows (fig. 1, steps 1.A–1.C). First, their latencies were identified as those of the zero crossings in the estimate of the first derivative of the ABR waveform. Second, their up-going 'trough-to-peak' (A) and down-going 'peak-to-trough' (B) amplitudes were calculated as the preceding trough-to-peak and peak-to-following-trough amplitudes, respectively. Visually insignificant peaks with amplitude below a specified threshold (0.01 μV in this study) were removed from the list of candidate peaks.

Step 2: Identification of Peaks I–VII

Once all candidate peaks had been identified, the algorithm searched these candidates for peaks I–VII (fig. 1, steps 2.A–2.G). The search was performed in the prespecified order of peaks V, III, I, II, IV, VI and VII, based both on their general order of clinical importance and ease of either manual or automatic detection. As a result, the identification of one peak determined the bounds on the possible locations of the next peak. This served to both reduce the number of candidate peaks to search and to ensure that once a peak was labelled, it could not be a candidate for a subsequent peak. However, this also meant that subsequent peaks could only be found if the prerequisite peaks were also found, a fact which has been argued to be advantageous for algorithm robustness [Pool and Finitzo, 1989].

The labelling of all peaks I–VII utilised the same two-stage process, the parameters of which are summarized in table 1, and the details of which are outlined below.

Step 2, Stage 1: Selection of a Candidate Peak Based on Latency Alone

The first step in labelling each of the peaks I–VII was to identify the expected latency of the peak being searched for. After this, the candidate peak nearest to this latency was identified as the 'most likely' peak (fig. 2, stage 1). For peaks I, III and V, the expected latencies were the average latencies (μ) previously reported from a larger normative database that had used the same ABR equipment and test protocols as our study and the same subjects plus an additional 60 others [Wilson, 2001]. For peaks II and IV, the expect-

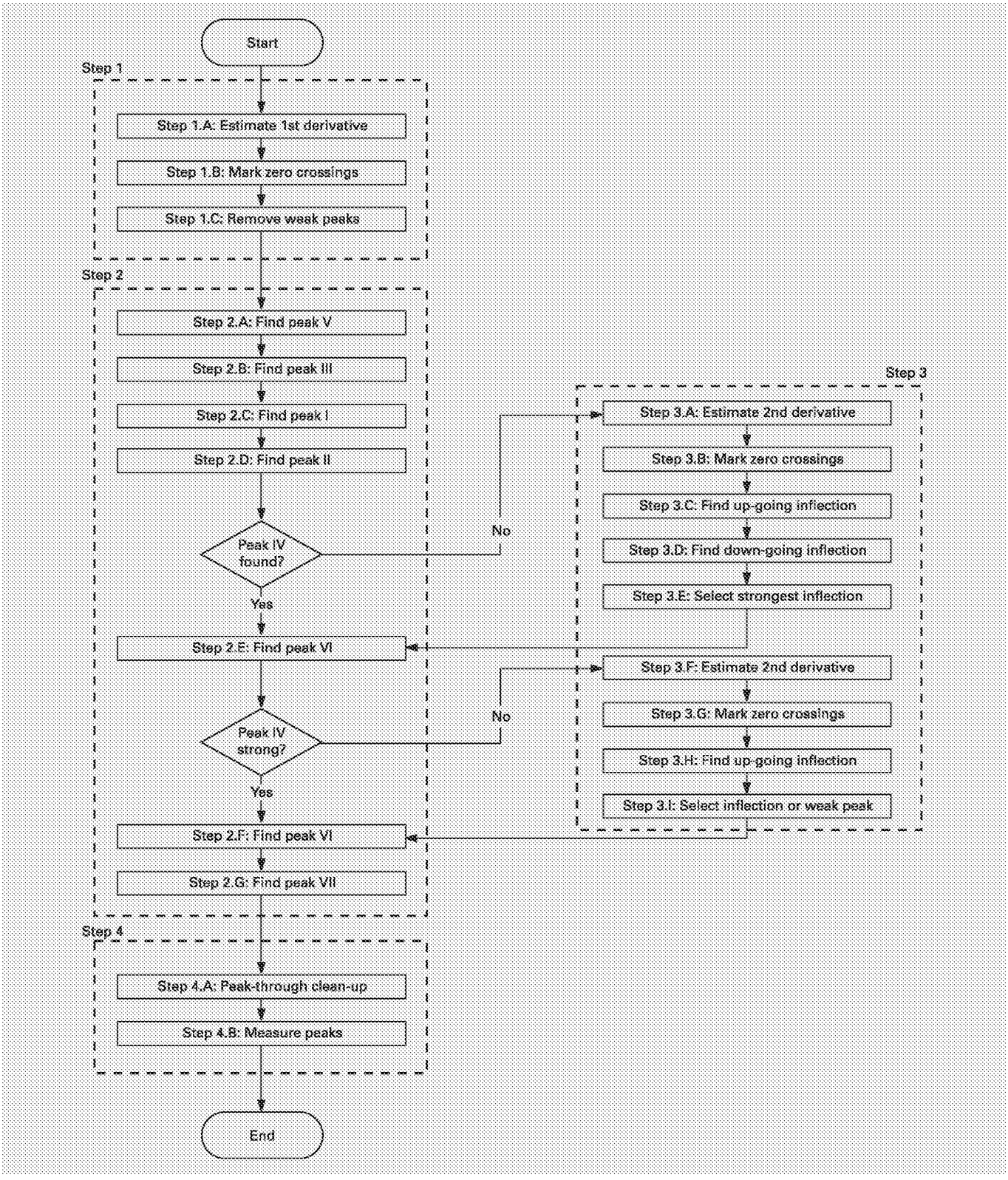

Fig. 1. Flowchart of the proposed algorithm. Note in step 3.E that the strongest inflection is the one closest to being a zero crossing (that is, with the minimum absolute gradient) and peak IV is defined to be 'strong' if it has a trough-to-peak or peak-to-trough height ≥0.05 μV.

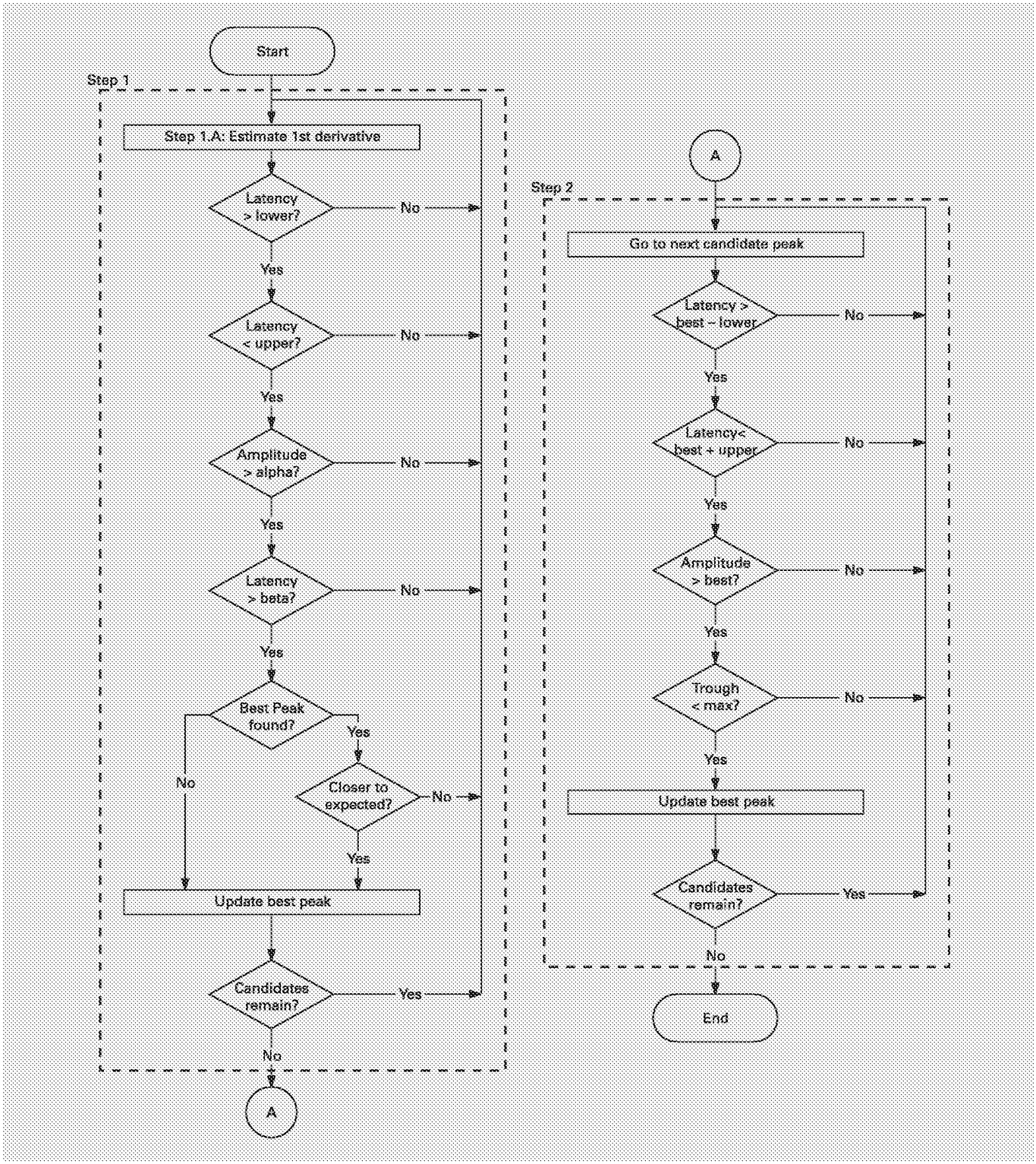

Fig. 2. Flowchart for the two-stage process of step 2, which labels peaks I–VII from the list of 'candidate peaks' (the zero crossings in the first derivative waveform). Here, 'upper' and 'lower' relate to the upper and lower search limits of each stage as per table 1; the 'best' peak refers to the 'most likely' peak in stage 1 and the 'strongest' peak in stage 2; 'max' is a constant defining the maximum trough depth that can separate the current 'best' and a potential 'stronger' peak in stage 2 (fixed at 0.05 μV).

Table 1. Parameters used in the two-stage identification of the peaks I–VII as part of step 2 of the algorithm

Peak No.	Stage 1					Stage 2	
	Expected peak latency	Upper search limit	Lower search limit	Minimum up-going amplitude	Minimum down-going amplitude	Upper search limit[1]	Lower search limit[1]
V	μ_5	$\mu_5 + 10\sigma_5$	$\mu_5 - 10\sigma_5$	α_5	β_5	2δ	$\delta/10$
III	μ_3	V – δ	$\mu_3 - 5\sigma_3$	α_3	β_3	$\delta/2$	$\delta/2$
I	μ_1	III – δ	$\mu_1 - 5\sigma_1$	α_1	β_1	$\delta/2$	$\delta/2$
II	(III + I)/2	III – δ	I + δ	α_2	β_2	$\delta/2$	$\delta/2$
IV	(2V + III)/3	V – $2\delta/3$	III + δ	α_4	β_4	$\delta/2$	$\delta/2$
VI	V + 1.6	V + 6δ	V + δ	α_6	β_6	4δ	2δ
VII	VI + 1.6	VI + 6δ	VI + δ	α_7	β_7	4δ	2δ

Latencies are expressed in milliseconds and amplitudes in microvolts. Mean ($\mu_{1,3,5}$) and standard deviation ($\sigma_{1,3,5}$) values were based on data previously published in Wilson [2001], as shown in table 2. The predefined minimum latency separating any two peaks was $\delta = 0.45$ ms and the minimum up-going (α) and down-going (β) peak amplitudes were $\alpha_{1-7} = 0.01$ μV, $\beta_{1-4,\,6-7} = 0.01$ μV, and $\beta_5 = 0.1$ μV, respectively.

[1] From the latency of the candidate peak selected in stage 1.

Table 2. Mean ($\mu_{1,3,5}$) and one Standard Deviation ($\sigma_{1,3,5}$) values for the latencies (ms) of ABR peaks I, III, and V, as derived from Wilson [2001]

Peak No.	Males			Females		
	18–30 years	31–45 years	46–60 years	18–30 years	31–45 years	46–60 years
I	2.40 ± 0.12	2.30 ± 0.15	2.44 ± 0.21	2.27 ± 0.09	2.34 ± 0.11	2.36 ± 0.15
III	4.63 ± 0.16	4.59 ± 0.19	4.64 ± 0.22	4.47 ± 0.11	4.68 ± 0.20	4.68 ± 0.17
V	6.44 ± 0.19	6.39 ± 0.20	6.50 ± 0.22	6.23 ± 0.13	6.45 ± 0.22	6.52 ± 0.25

The stimulus was a 90-dBnHL click presented through insert earphones. N = 20 for each age/gender group. Right and left ear values have been averaged.

ed latencies were based on those of peaks I and III, and III and V, respectively. For peaks VI and VII, the expected latencies were based on those of peaks V and VI, respectively. The expected latencies of peaks I, III and V are shown in table 2.

To be selected as the 'most likely' peak, the candidate peaks also had to:

(1) occur within a range around the expected latency, referred to here as the 'search limit'. For peaks I, III and V, these search limits were based on a prespecified number of standard deviations (σ), also taken from the previously reported database [Wilson, 2001] and the latencies already found for peaks V or III as appropriate; for the remaining peaks, the search limits were based on the latencies of already found peaks;

(2) be a minimum latency, δ (≥ 0.45 ms), away from an already labelled peak;

(3) have a minimum up-going (α) and down-going (β) amplitude based on the expected peak amplitude and the estimated noise floor. For example, the minimum down-going amplitude for peak V was $\alpha_5 = 0.1$ μV.

Step 2, Stage 2: Local Search for a 'Stronger' Peak

Once the 'most likely' peak had been selected (fig. 2, stage 2), all candidate peaks were again searched in an attempt to locate a 'stronger' peak that:

(1) had an amplitude greater than the 'most likely' peak found in stage 1 (based on latency alone);

(2) lay within a range of acceptable latencies, again referred to here as the 'search limit' [for peaks I–V, this search limit was quite small (fractions of δ), while for peaks VI and VII, it was quite large (multiples of δ)], and

(3) was separated from the 'most likely' peak by a local trough of a depth less than a prespecified maximum of 0.05 μV for all peaks.

If a 'stronger' peak was found, then it was labelled as the peak found. If not, the original 'most likely' peak (based on latency alone) was labelled as the peak found.

Step 3 (Optional): Detecting Peaks II and IV as Points of Inflection

As already mentioned, peaks II and IV often do not present themselves as individual peaks in the ABR waveform, that is, as a local maximum with a gradient that passes through zero. Instead, they often appear as points of inflection, that is, as a change from convexity (decreasing gradient) to concavity (increasing gradient), or vice versa, on the shoulder of an adjacent peak. Therefore, a point of inflection may not have a gradient that passes through zero. However, a point of inflection will be a turning point in the gradient, either a local minimum or a local maximum and therefore can be detected as a zero crossing in the second derivative (for example fig. 5). Therefore, if peaks II or IV could not be found as zero crossings in the first derivative, an attempt was made to find them as zero crossings in the second derivative (fig. 1, steps 3.A–3.I).

In searching for points of inflection, it was assumed that peak IV could only appear as a point of inflection on the up-going slope of peak V (fig. 1, step 3.F–3.I), while peak II could appear as a point of inflection on either the down-going slope of peak I or the up-going slope of peak III (fig. 1, step 3.A–3.E). Therefore, the algorithm had to be capable of finding both concave-to-convex inflections on down-going slopes and convex-to-concave inflections on up-going slopes. These were simply identified as negative- and positive-going zero crossings in the second derivative waveform (relating to troughs and peaks in the first derivative waveform), respectively (for example fig. 5).

To ensure that the algorithm found the strongest point of inflection, that is, the one closest to being a zero crossing in the first derivative, zero crossings in the second derivative waveform were only selected if they:

(1) were either positive going for down-going slope inflections or negative going for up-going slope inflections;

(2) had a first derivative magnitude between zero and a predetermined maximum of 0.05 μV/ms, and

(3) had a latency that was between a predetermined lower and upper bound. For down-going slope inflections, the lower bound was the previously labelled peak plus δ/2 ms (within which separate peaks were not expected to occur), while the upper bound was the trough associated with the next labelled peak minus δ/2 ms. For up-going slope inflections, the bounds were the trough of the previously labelled peak and the next labelled peak (again, plus or minus δ/2 ms as appropriate). For example, the algorithm would search for peak IV as the strongest point of inflection between the trough after peak III and peak V (for example fig. 5).

Again, this process illustrates the philosophy behind the algorithm of finding the stronger peaks first (V, III and I) and then using this information to constrain the search for the weaker peaks (II, IV, VI and VII).

Finally, if peaks II or IV were found as a point of inflection, then an additional (pseudo) trough was inserted into the list of 'candidate troughs'. Note that this pseudo trough does not change the waveform in any way, but is required both to preserve the implicit peak-trough-peak sequence of zero crossings and as a point of reference for measuring peak amplitudes. Currently an arbitrary decision was made to mark the sample after the point of inflection as the pseudo trough to ensure that step 4 works correctly.

Step 4: Peak Trough Cleanup

Once as many as possible of the peaks I–VII had been found and labelled, it was necessary to mark all of the interpeak troughs so that peak amplitudes could be reliably measured (fig. 1, steps 4.A–4.B). Initially, the trough preceding the first peak and the deepest trough following the last peak were kept. If either of these troughs did not exist, then the first/last sample (respectively) was marked as a trough. Next, all unlabelled peaks either before, after or in between the labelled peaks were removed from the list of 'candidate peaks' and the deepest troughs between the labelled peaks kept, while all other troughs were removed from the list of 'candidate troughs'. Finally, each labelled peak's latency, up-going amplitude (A) and down-going amplitude was measured.

Derivative Estimation

In this paper, the derivative estimation filters used were members of the family of Gaussian wavelets, the second derivative (Laplacian) of which is more commonly known as a Mexican hat wavelet [Daubechies, 1992; Mallat, 1999]. Derivatives of Gaussian filters combine both smoothing and derivative estimation in a single operation (and so are band-pass filters, as shown in figure 4). In fact, because of the linearity of the convolution operation, they are equivalent to applying Gaussian smoothing to the derivative of the signal (or vice versa) [Marr and Hildreth, 1980]. Favourable properties of the Gaussian wavelets include: their minimisation of time-frequency uncertainty (they minimise the product of time and frequency resolution) and their ready implementation as a linear phase finite impulse response filter with minimal error [Marr and Hildreth, 1980]. Phase linearity is a direct result of their symmetrical filter coefficients, as shown in their impulse response in figure 3. The previous use of quadratic spline wavelets (that are effectively the first-order derivative of a cubic spline) [Popescu et al., 1999], which have both linear phase and compact support, was not considered here as they do not have optimal time-frequency properties when estimating the first and second derivatives used in this paper [Unser et al., 1992].

As the ABR waveforms being analysed were all generated with 90 dBnHL stimuli, the derivative estimation was applied using a (20-tap) finite impulse response filter with an upper cut-off frequency of approximately 7 kHz. This filter clearly passes all of the frequency components contained in the ABR waveform [Boston, 1981; Fridman et al., 1982], whilst removing some high-frequency noise.

Algorithm Parameterisation

It should be noted that the parameter values quoted in table 1 directly relate to the expected latency and amplitude values of the ABR peaks I–VII and can be modified to suit specific waveform acquisition conditions or waveform morphologies. For example, the search parameters for the primary ABR peaks (I, III and V) are:

(1) the expected (average) latency of each peak (μ_1, μ_3 and μ_5 in table 1);

(2) the expected deviation around the average latency (σ_1, σ_3 and σ_5 in table 1);

(3) the minimum up-going and down-going amplitude of each peak (α and β in table 1), and

(4) the minimum latency separating any two consecutive peaks (two constants derived from δ in table 1).

Therefore, a maximum of 6 parameters, which all relate to peak latency and amplitude, are required to adapt the algorithm's search for each ABR peak. However, in reality, the values of these parameters (α, β and δ) are codependent and can be made common across a number of the peaks. This significantly reduces the total number

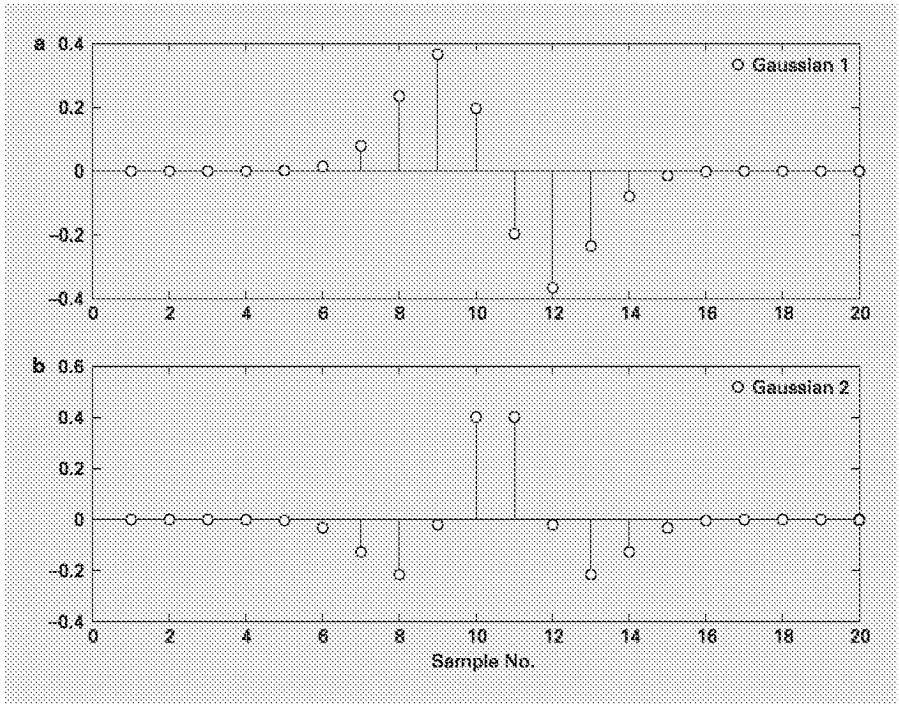

Fig. 3. Impulse responses of the first (**a**) and second (**b**) derivative of Gaussian wavelets.

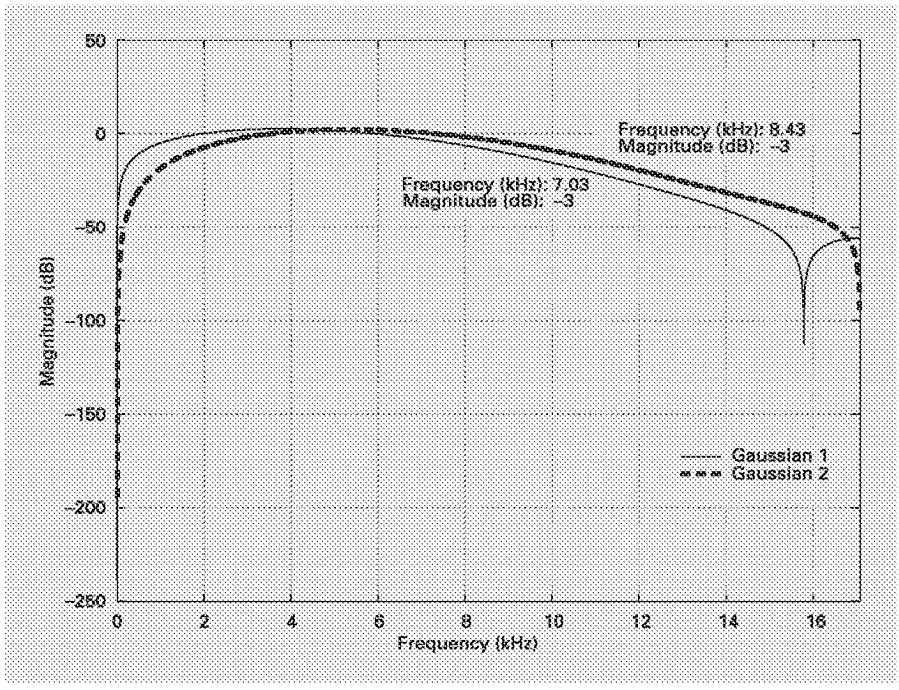

Fig. 4. Magnitude of the frequency response of the first (solid line) and second derivative (dashed line) of Gaussian wavelets.

of parameters that have to be specified in table 1. In addition, the latency parameters defining the minor ABR peaks (II, IV, VI and VII) are defined in relation to the primary peaks (I, III and V) and so may not need to be modified at all.

As this study attempted to demonstrate the efficacy of the algorithm on normative data, we have specified the expected latency and deviation of peaks I, III and V from published data matched for stimulus level and subject age, gender and test ear [Wilson, 2001]. The expected latency and deviation of peaks II, IV, VI and VII are then defined in relation to the algorithm's estimated latencies of peaks I, III and V. Finally, we selected all other parameters (α, β and δ) experimentally and have presented our results for these where appropriate.

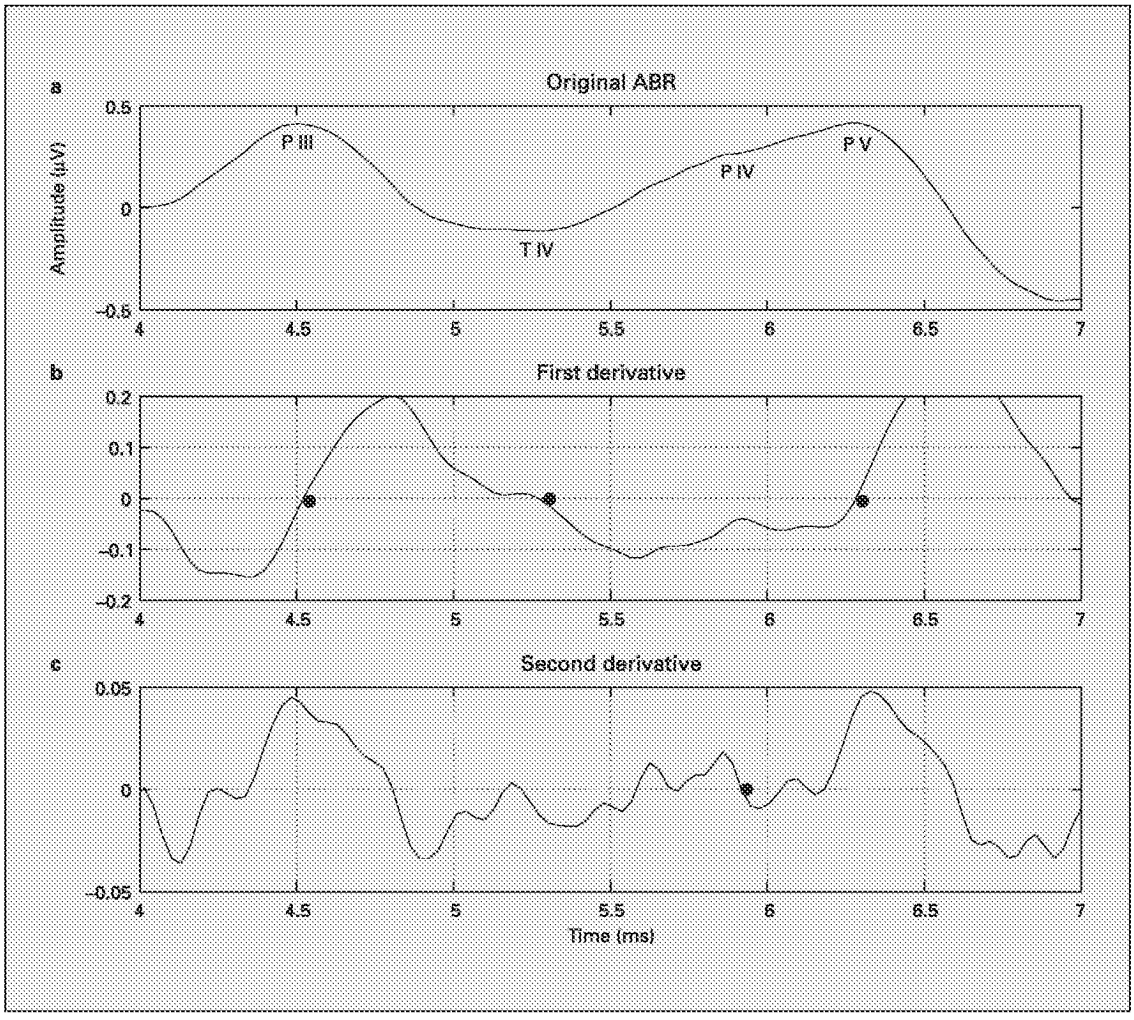

Fig. 5. An example III/IV/V complex (**a**) showing peaks III and V as zero crossings of the first derivative (**b**) and IV as a point of inflection (**c**). That is, peak IV is a local maximum in the first derivative waveform (**b**) and hence a zero crossing in the second derivative waveform (**c**).

Data Analysis

An expert audiologist (Wayne Wilson, this study's second author) with over 10 years of experience in ABR analysis labelled all seven peaks in each ABR waveform and calculated their latency. These latencies were measured taking into account not only the individual ABR waveform being analysed, but also its repeat trace in the same ear and the two waveforms from the other ear of the same subject. These peak latencies were then taken to be the 'gold standard', against which the peak latencies determined by the proposed algorithm were compared.

The peak latencies determined by the algorithm were considered correct if they were:

(1) within 0.2 ms of the gold standard (true positive); the threshold of 0.2 ms was chosen as it is a commonly used clinical standard for peak latency reproducibility in ABR [Vannier et al., 2002], and

(2) determined not to be present by both the algorithm and the audiologist (true negative).

The latencies determined by the algorithm were considered incorrect if they were:

(1) beyond 0.2 ms of the gold standard (false negative; note that this type of error could have been considered to be both false positive and false negative);

(2) determined to be present by the algorithm, but not present by the audiologist (false positive), and

(3) determined not to be present by the algorithm, but present by the audiologist (false negative).

A receiver operating characteristic curve could not be constructed as none of the primary peaks (I, III and V) were absent in the data set and so labelled as negative examples by the expert. However, the mean absolute latency error for each peak was also calculated, as this has been previously reported in the literature [Pool and Finitzo, 1989; Tian et al., 1997; Vannier et al., 2001].

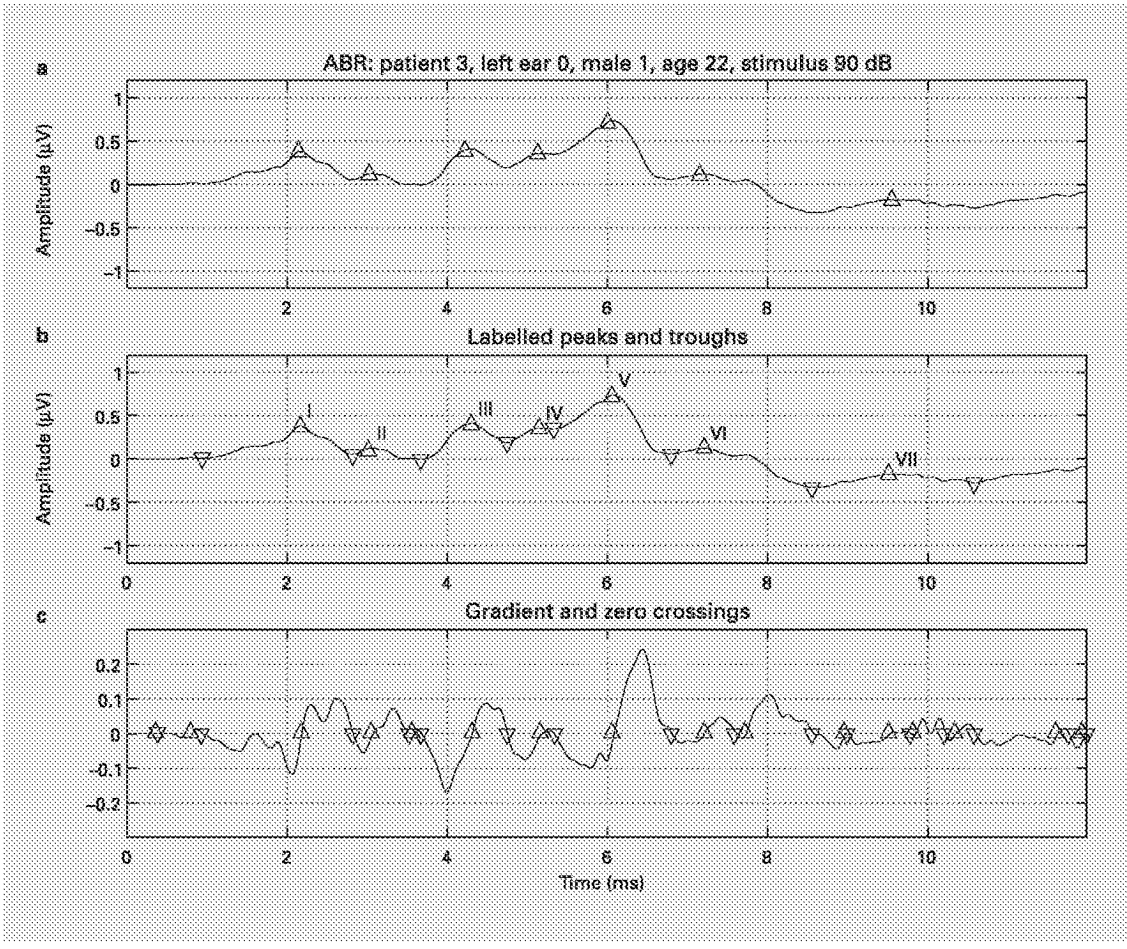

Fig. 6. An ABR labelled by an expert (**a**) and by the proposed algorithm (**b**). In addition, the first derivative of the ABR waveform and all 'candidate' peak and trough zero crossings are also shown (**c**).

Table 3. Accuracy, sensitivity, specificity, and mean absolute latency error (MAE) of the proposed algorithm

	Peak No.						
	I	II	III	IV	V	VI	VII
Accuracy	0.96	0.83	0.98	0.77	0.98	0.75	0.46
Sensitivity	0.96	0.84	0.98	0.79	0.98	0.79	0.66
Specificity	N/A	0.00	N/A	0.36	N/A	0.00	0.00
MAE, ms	0.03	0.12	0.05	0.12	0.06	0.20	0.37

In this case, a correctly labelled peak must be labelled to within 0.2 ms of an expert audiologist. Note that for peaks I, III, and V, all waveforms were labelled by the expert and so specificity was not calculated.

Results

Example Waveforms

Figure 5 illustrates a peak III-IV-V complex where peaks III and V were correctly found as zero crossings in the first derivative and peak IV was correctly found as a point of inflection. It should be noted that peak IV (approximately 5.9 ms) occurs at a local maximum in the first derivative waveform (that is, as the gradient changes from increasing to decreasing) and so appears as a zero crossing in the second derivative.

Figure 6 illustrates an example ABR where all seven peaks were correctly labelled by the proposed algorithm (to within 0.2 ms of the expert). In addition, figure 6 also illustrates the eight preserved troughs that were then used to measure the up-going and down-going amplitudes for each peak (the values of which are not reported here). The

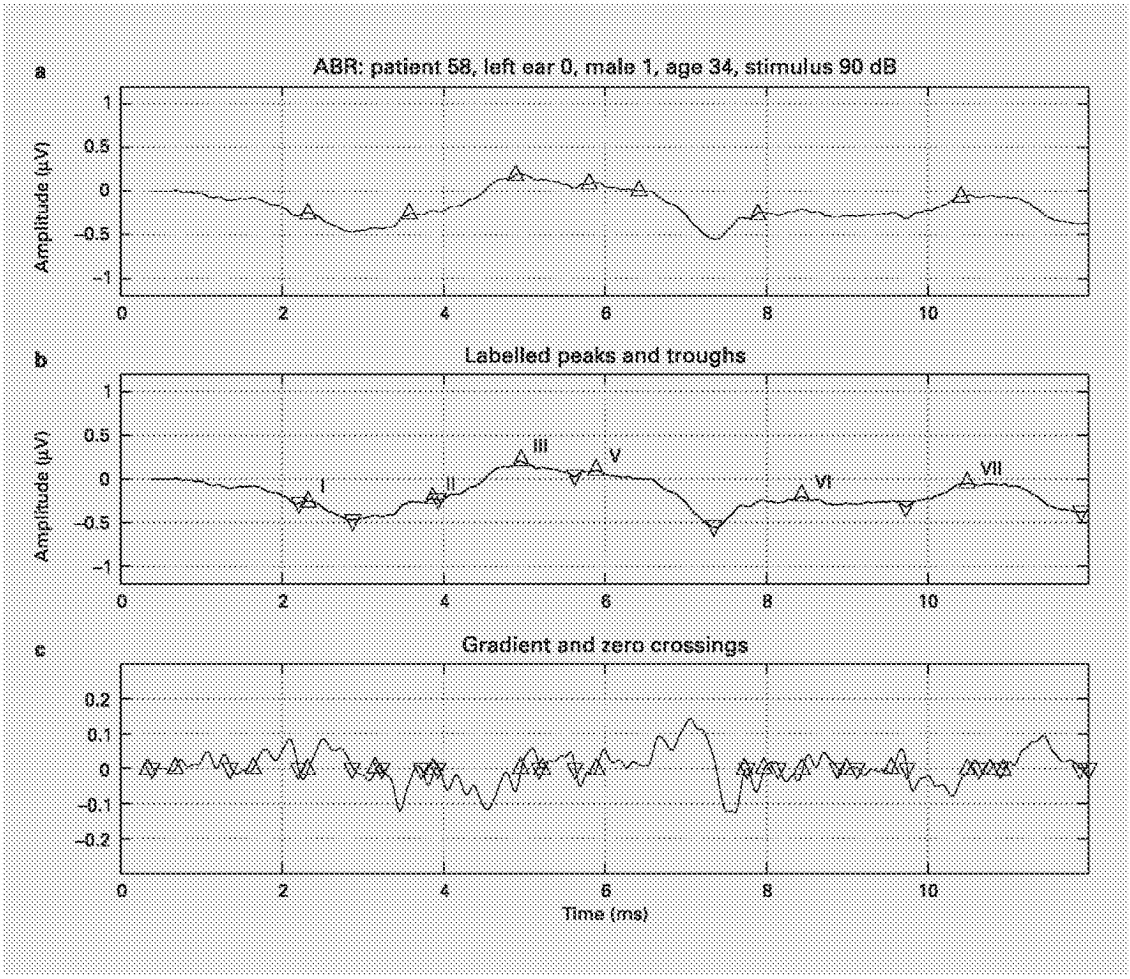

Fig. 7. An example ABR labelled by an expert (**a**) and by the proposed algorithm (**b**), with the gradient and 'candidate' peak and trough zero crossings shown for reference (**c**). The proposed algorithm incorrectly labels peaks II, IV, V and VI; the greatest number of mislabelled peaks in the whole data set of 240 waveforms.

accuracy of the trough labelling was not quantitatively evaluated in the present study as peak amplitudes are less frequently used in clinical practice and were therefore considered of lesser importance. However, qualitatively, the correct troughs appeared to be preserved in all the cases investigated.

Figure 7 illustrates the ABR with the greatest number of mislabelled peaks from the whole data set of 240 ABR waveforms. Here the proposed algorithm has incorrectly labelled peaks II, IV, V and VI, while I, III and VII are correct to within 0.2 ms of the expert.

Algorithm Performance

Table 3 shows the overall algorithm accuracy (probability of a correct response), sensitivity (probability of a true positive) and specificity (probability of a true negative) to within 0.2 ms of the expert. Table 3 also shows the mean absolute latency error for each peak.

If the algorithm is constrained to detect peaks II and IV using the first derivative only, accuracy on these peaks falls from 83 to 79% and from 77 to 53%, respectively. In addition, if the algorithm is constrained to label peaks using stage 1 of step 2 only, that is, always selecting the 'most likely' peak based on latency alone, then accuracies on peaks I and III remain constant at 96 and 98%, respectively, while accuracy on peak V falls marginally from 98 to 97%. Accuracy on peak II increases from 83 to 84%, but falls more significantly on peaks IV, VI and VII from 77 to 58%, 75 to 60% and 46 to 35%, respectively.

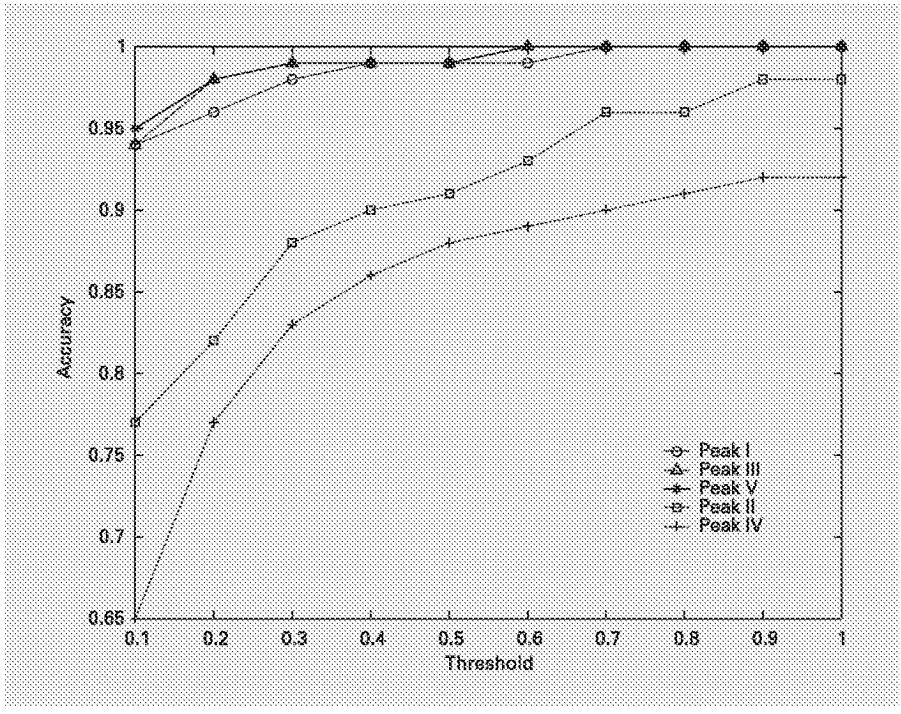

Fig. 8. Increase in detection accuracy of peaks I–V as the latency tolerance threshold is varied from 0.1 to 1 ms.

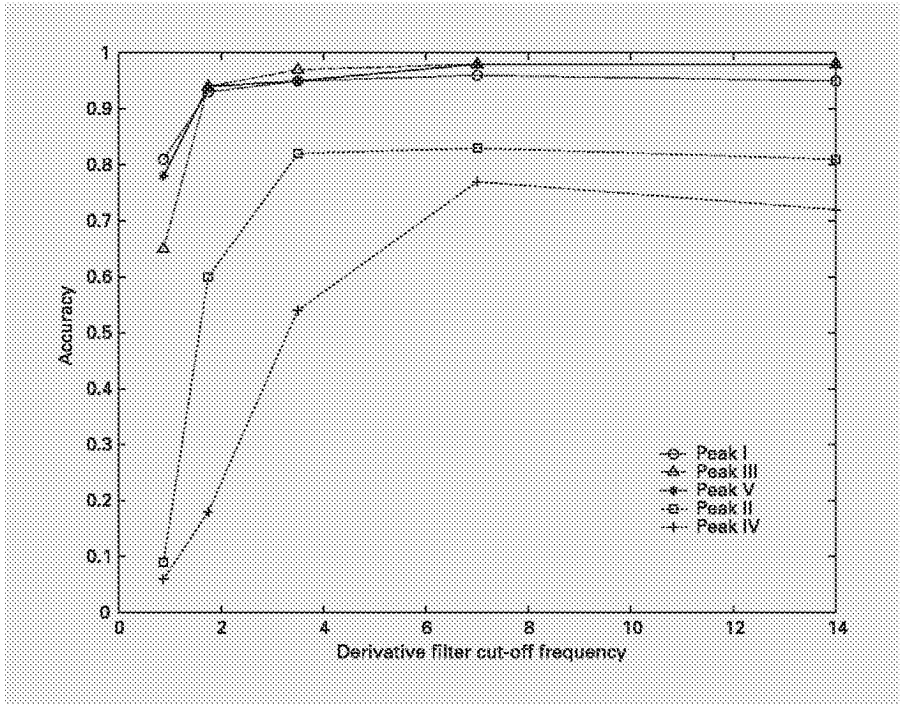

Fig. 9. Change in detection accuracy of peaks I–V as the upper cut-off frequency of the first order Gaussian derivative filter is varied from 875 Hz to 14 kHz. The peak in detection accuracy for all waves is observed at cut-off frequency of 7 kHz.

Figure 8 shows the change in overall algorithm accuracy for peaks I–V as the latency tolerance of a correct detection was varied from 0.1 to 1 ms. It can be seen that all peaks I, III and V in the data set were accurately identified to within 0.7 ms of the expert.

Figures 9 and 10 show the change in overall algorithm accuracy for peaks I–V as the upper cut-off frequency of the derivative estimation filters were varied between 875 Hz and 14 kHz (fig. 9) and the search window width, δ, was varied between 0.15 and 0.75 ms (fig. 10).

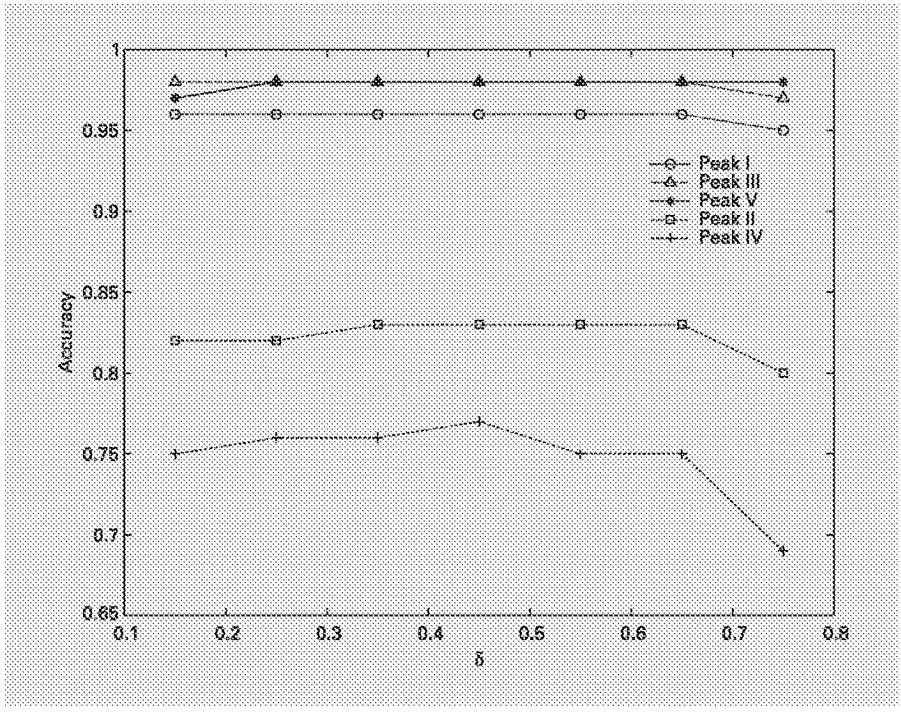

Fig. 10. Change in detection accuracy of peaks I–V as the search window parameter, δ, is varied from 0.15 to 0.75 ms. The peak in detection accuracy for all waves is observed at δ = 0.45 ms.

Discussion

Algorithm Implementation and Parameterisation

It should be noted that for both algorithmic simplicity, and to maximise peak detection accuracy in the good SNRs resulting from the 90-dBnHL stimuli, it was considered desirable to attempt all peak detection at a single wavelet scale (filter cut-off frequency). This decision is supported by figure 9, which shows maximum detection accuracy for peaks I–V at a filter cut-off frequency of 7 kHz. However, figure 9 also shows that the detection accuracy of peaks I, II and V is stable over a wide range of cut-off frequencies, only decreasing significantly as the cut-off frequency falls below about 1500 Hz, which is the approximate bandwidth of the ABR signal. In addition, because peaks II and IV are generally either of a lower amplitude or are not true peaks at all (that is, they are points of inflection), they are more sensitive to the choice of filter cut-off frequency.

Figure 10 illustrates the detection accuracy of peaks I–V is reasonably robust to the choice of the 'search limit' parameter, δ. Only peak IV shows significant variation in detection accuracy as δ is varied between 0.15 and 0.75 ms. This is also supported by the changes in detection accuracy of peaks I, II, III and V of either 0 or 1% when step 2 of the algorithm was constrained to select peaks based on latency alone, that is, using stage 1 of step 2 only. This shows that the more variable peaks (IV, VI and VII) cannot be reliably detected based on latency alone and are only reliably detected when a local search for the strongest peak, as per stage 2 of step 2, is performed.

Algorithm Performance

The proposed algorithm proved to be very accurate (96–98%) in correctly labelling the more clinically reliable ABR peaks (I, III and V) to within a latency of 0.2 ms of the expert. This meant that of the grand total of 720 peaks I, III and V examined (240 waveforms times 3 peaks), the proposed algorithm correctly labelled 698 (97%) of these.

Although direct comparisons with prior systems are difficult, these results are either equivalent to, or better than, recently reported systems whose accuracies vary from 90 to 100% [Gabriel et al., 1980; Delgado and Ozdamar, 1994; Popescu et al., 1999; Vannier et al., 2001, 2002]. In addition, the results were clearly superior to older systems whose peak V accuracies were below 90% [Mason, 1984; Madhavan et al., 1986]. Similarly, the mean latency error of 0.06 ms on peak V was also equivalent to previously reported system errors [Pool and Finitzo, 1989; Tian et al., 1997]. These results are espe-

cially notable considering that both the expert and some of the previously reported systems have used ABR information beyond the single trace analysis used here, for example, comparisons between repeat traces and with traces obtained at higher and lower stimulus levels [Delgado and Ozdamar, 1994; Vannier et al., 2001, 2002]. However, it should be noted that the current algorithm was only evaluated using suprathreshold ABR waveforms.

The proposed algorithm proved to be less accurate (45–83%) in correctly detecting the less clinically reliable ABR peaks (II, IV, VI and VII) to within a latency of 0.2 ms of the expert. The reasons for this reduction in accuracy were thought to be threefold. Firstly, these peaks presented with increased variation in morphology. Secondly, they often presented with reduced amplitude on an approximately constant background noise level, and therefore had a reduced SNR. Finally, the latencies of peaks IV, VI and VII were extremely variable, making selection of the most likely zero crossing less effective. This last point was supported by the finding that, when the algorithm was constrained to label peaks using stage 1 of step 2 only, that is, always selecting the 'most likely' peak based on latency alone, then accuracy on peaks IV, VI and VII fell from 77 to 58%, 75 to 60% and 46 to 35%, respectively. It is also of interest to note that when peaks II and IV were detected using the first derivative of the ABR alone, that is, removing step 3 of the algorithm, then accuracy on these peaks fell by 5 and 31%, respectively. Therefore, the decision to allow the algorithm to use second derivatives to identify peaks II and IV as points of inflection when they cannot be found as true peaks was considered successful, particularly for peak IV. This finding also highlights the question of 'what criteria do audiologists use when marking peaks II and IV?' In the cases where the audiologist is marking a point of inflection, you would expect increased interobserver variation in the measured latencies and a subsequent reduction in the confidence level attached of those measurements. The use of the proposed algorithm could improve these confidence levels, particularly in the case of a novice operator.

It has been proposed that a system must be at least 95% accurate to be clinically effective [Boston, 1989]. By this definition, the proposed algorithm was capable of being clinically applied to the detection of peaks I, III and V on this set of ABR waveforms and was close to being adequate for the detection of peaks II and IV. As well as the level of accuracy demonstrated by the proposed algorithm, the use of derivative estimation would also seem to be the most suitable general-purpose methodology for ABR peak-finding applications. Unlike matched filters [Woodworth et al., 1983; Vannier et al., 2001, 2002], derivative estimation makes no assumptions about the morphology of the peaks, only that the peaks are actually turning points (local maxima) in the waveform. In addition, this condition can be further relaxed if the second derivative is utilised allowing peaks to be found as points of inflection. The proposed algorithm is also conceptually simple and can be directly and easily related to current practice, unlike some of the previous algorithms that, for example, utilise ANN approaches that provide only a 'black-box' solution [Alpsan et al., 1994; Popescu et al., 1999; Tian et al., 1997]. More structured, feature-based approaches, such as the one used in this study, have also been shown to offer improved performance when compared to ANN approaches that often directly accept the ABR waveform as the input vector [Sanchez et al., 1995].

Due to the extremely variable morphology of the ABR, even at suprathreshold stimulus levels, no fully automated ABR analysis algorithm can reasonably expect 100% accuracy. Therefore, algorithms, such as the one proposed in this paper, are probably better suited to semi-automated applications where it can act as an aid to the audiologist, rather than a fully automated replacement. By providing a semi-automated ABR peak labelling system, the proposed algorithm could label what it believes are the peaks and the audiologist could then move or delete the labels as required. Because of this, estimates of the degree of certainty of the algorithm's peak labels were not calculated as they were deemed unnecessary in a semi-automated system. The use of the algorithm as a semi-automated system is further supported by figure 8, which shows that the proposed algorithm correctly labelled all peaks I, III and V in the data set to within an accuracy of 0.7 ms. As a result, an audiologist using the system would not have had to correct any of these labels by more than 0.7 ms.

Conclusions

We have presented an algorithm that automatically detects and labels all seven peaks in normal, suprathreshold ABR waveforms. This algorithm finds peaks via zero crossings in the first derivative of the waveform and then uses a table of expected peak latencies and amplitudes to label peaks I–VII. In addition, when peaks II and IV cannot be found via the first derivative, an attempt is made

to find them as a point of inflection via the second derivative waveform. Two continuous Gaussian wavelets were used to estimate these derivatives, each of which provides optimal time-frequency properties in a single resolution.

On a large data set of normative ABR waveforms, the primary peaks of clinical interest (I, III and V) were labelled with an accuracy of between 96 and 98% to within 0.2 ms, and 100% to within 0.7 ms, of an expert audiologist. Of particular interest was the fact that both peaks II and IV could be detected with an accuracy of 83 and 77%, respectively, when allowed to be detected as a point of inflection when a true peak (local maximum) could not be found. This clearly shows that the expert audiologist must also (perhaps subconsciously) have labelled peaks II and IV as points of inflection when a true peak could not be found.

Limitations and Further Research

The current study was limited by its use of a single suprathreshold stimulus level (90 dBnHL) when recording the ABR waveforms from normal subjects. Part of the reason for this decision was the initial desire to apply the algorithm to the detection of all seven peaks of the ABR in applications such as intra-operative and intensive care unit neurological monitoring. Further research is therefore warranted into the performance of the algorithm on ABR waveforms recorded at lower stimulus levels, especially if the algorithm is to be applied to hearing threshold estimation. For stimulus intensities just above threshold (say, around 30 dBnHL), the algorithm will probably have to be extended to include a course-to-fine multiresolution analysis in order to improve detection rates at these lower SNR levels. In principle, this could proceed as follows. First, the algorithm is applied at a course scale, say with a derivative cut-off frequency of 875 Hz. This increases peak detection reliability, due to the improved SNR of the narrowband filter, but has a reduced temporal accuracy. Therefore, the algorithm is applied again at a finer scale, using the previously found latencies as estimates of the peak latencies at this new scale. This process is repeated a number of times, doubling the bandwidth of the derivative estimation filters each time (in the sequence: 875, 1750, 3500 and 7000 Hz) until the desired level of temporal accuracy has been achieved.

Further research is also warranted into the performance of the algorithm on ABR waveforms recorded from abnormal subjects, especially if the algorithm is to be widely applied to clinical assessments. For this application, initial manual identification of the abnormal peaks may be required so that the expected latencies in table 1 can be updated and the algorithm can continue monitoring from this point. In addition, the algorithm proposed in this paper analysed each ABR waveform in isolation, that is, without recourse to other waveforms acquired from the subject at higher stimulus intensities [Vannier et al., 2001, 2002], the opposing ear or a repeat trace [Pool and Finitzo, 1989]. Rules for integrating this sort of higher-level information should be applied to the proposed algorithm as a postprocessing operation, with the aim of reducing the number of false-positive peak detections.

Finally, we would like to determine if the proposed algorithm is capable of measuring peak latencies to the same level of accuracy as would be expected from a group of expert or novice audiologists. In this way, we would endeavour to repeat the result reported in Gabriel et al. [1980] where their system was reported to be more reliable than an inexperienced audiologist.

Acknowledgements

The authors would like to thank the editor and the anonymous referees for their constructive comments on earlier drafts of this paper.

References

Alpsan D, Towsey M, Ozdamar O, Tsoi A, Ghista DN: Determining hearing threshold from brain stem evoked potentials. Optimising a neural network to improve classification performance. IEEE Eng Med Biol Mag 1994;4:465–471.

Battista RA, Wiet RJ, Paauwe L: Evaluation of three intraoperative auditory monitoring techniques in acoustic neuroma surgery. Am J Otol 2000;21:244–248.

Boston RJ: Spectra of auditory brainstem response and spontaneous EEG. IEEE Trans Biomed Eng 1981;28:334–341.

Boston RJ: Automated interpretation of brainstem auditory evoked potentials: A prototype system. IEEE Trans Biomed Eng 1989;36:528–532.

Cebulla M, Sturzebecher E, Werecke KD: Objective detection of auditory brainstem potentials. Scand Audiol 2000;29:44–51.

Daubechies I: Ten Lectures on Wavelets. Philadelphia, Pennsylvania, Society for Industrial and Applied Mathematics, 1992.

Delgado RE, Ozdamar O: Automated auditory brainstem response interpretation. IEEE Eng Med Biol Mag 1994;2:227–237.

Fridman J, John ER, Bergelson M, Kaiser JB, Baird HW: Application of digital filtering and automatic peak detection to brain stem auditory evoked potential. Electroencephalogr Clin Neurophysiol 1982;53:405–416.

Gabriel S, Durrant JD, Dickter AE, Kephart JE: Computer identification of waves in the auditory brain stem evoked potentials. Electroencephalogr Clin Neurophysiol 1980;49:421–423.

Hall J: Handbook of Auditory Evoked Responses. Needham Heights, Massachusetts, Allyn and Bacon, 1992.

Jacobson JT: The Auditory Brainstem Response. San Diego, College-Hill Press, 1985.

Jewett DL, Williston JS: Auditory evoked far fields averaged from the scalp of humans. Brain 1971;4:681–696.

Madhavan GP, De Bruin H, Upton ARM, Jernigan ME: Classification of brain-stem auditory evoked potentials by syntactic methods. Electroencephogr Clin Neurophysiol 1986;65:289–296.

Mallat SG: A Wavelet Tour of Signal Processing, ed 2. San Diego, Academic Press, 1999.

Marr D, Hildreth E: Theory of edge detection. Proc R Soc Lond B Biol Sci 1980;207:187–217.

Mason SM: On-line computer scoring of the auditory brainstem response for estimation of hearing threshold. Audiology 1984;23:277–296.

Neu M, Strauss C, Romstock J, Bischoff B, Fahlbusch R: The prognostic value of intraoperative BAEP patterns in acoustic neurinoma surgery. Clin Neurophysiol 1999;110:1935–1941.

Northern JL: Clinical measurement procedures in impedance audiometry; in Jerger J, Northern JL (eds): Clinical Impedance Audiometry, ed 2. Acron, American Electromedics, 1980, pp 19–39.

Pool KD, Finitzo T: Evaluation of a computer-automated program for clinical assessment of the auditory brain stem response. Ear Hear 1989;10:304–310.

Popescu M, Papadimitriou S, Karamitsos D, Bezerianos A: Adaptive denoising and multiscale detection of the V wave in brainstem auditory evoked potentials. Audiol Neuro Otol 1999;4:38–50.

Pratt H, Urbach D, Bleich, N: Auditory brainstem evoked potentials peak identification by finite impulse response digital filters. Audiology 1989;28:272–283.

Sanchez R, Riquenes A, Perez-Abalo M: Automatic detection of auditory brainstem responses using feature vectors. Int J Biomed Comput 1995;39:287–297.

Salvi RJ, Ahroon W, Saunders SS, Arnold SA: Evoked potentials: Computer-automated threshold-tracking procedure using and objective detection criterion. Ear Hear 1987;8:151–156.

Sininger YS, Hyde M, Don M: Method for detection of auditory evoked potentials using point optimized variance ratio. US Patent 6,196,977, March 2001.

Tian J, Juhola M, Gronfors T: Latency estimation of auditory brainstem response by neural networks. Artif Intell Med 1997;10:115–128.

Unser M, Aldroubi A, Eden M: On the asymptotic convergence of B-spline wavelets to Gabor functions. IEEE Trans Inf Theory 1992;38:864–872.

Vannier E, Adam O, Karasinski P, Ohresser M, Motsch J: Computer-assisted ABR interpretation using automatic construction of the latency-intensity curve. Audiology 2001;40:191–201.

Vannier E, Adam O, Motsch J: Objective detection of brainstem auditory evoked potentials with a priori information from higher presentation levels. Artif Intell Med 2002;25:283–301.

Wilson WJ: Improving prediction of outcome in the severe acute closed head injured patient – Signal processing and multivariant analysis of the normal auditory brainstem response; PhD thesis presented to the Department of Speech Pathology and Audiology, University of the Witwatersrand, Johannesburg, 2001.

Woodworth W, Reisman S, Fontaine BA: The detection of auditory evoked responses using a matched filter. IEEE Trans Biomed Eng 1983;30:369–376.