

An Evaluation of Otoacoustic Emissions as a Biometric

Neil J. Grabham, Matthew A. Swabey, Paul Chambers, Mark E. Lutman, Neil M. White *Senior Member, IEEE*, John E. Chad and Stephen P. Beeby *Senior Member, IEEE*

Abstract— This paper presents a comprehensive overview of an investigation into the use of Otoacoustic Emissions (OAE) as an identification biometric. OAE could be important as a biometric identifier in applications where users wear headsets since it is discrete and difficult to spoof. OAE are very low level (~17 dB Sound Pressure Level (SPL)) sounds emitted from the human ear as part of the normal hearing process. They can occur spontaneously or be invoked by a suitable stimulus, these being known as Transient Evoked Otoacoustic Emissions (TEOAE) and Distortion Product Otoacoustic Emission (DPOAE). An initial visual comparison shows that otoacoustic emissions are clearly distinctive and are stable over a six month period. A biometric analysis based on the Euclidean distance measurement of TEOAE recordings in the temporal domain was performed on pre-recorded datasets captured for medical purposes and data were collected specifically for this study. For a database of 23 subjects, the predicted Equal Error Rate (EER) was 1.24% for a 90% confidence interval. DPOAEs also demonstrated biometric potential but the level of discrimination is inferior to TEOAE. The combination of DPOAE and TEOAE into a multimodal analysis was demonstrated to be feasible although the potential improvement in performance is yet to be quantified. Finally the use of Maximum Length Sequencing (MLS) was investigated to reduce capture time without decreasing performance. This demonstrated a reduction in capture time for a TEOAE from 1 minute to 5 seconds with a visual analysis of a 4th order MLS showing good stability and reproducibility. OAEs can potentially be used as a biometric and benefit from their small template size (512 data points in our TEOAE biometric) and simple analysis. The level of background noise is the most significant practical factor that affects biometric performance.

Manuscript received December 15, 2011. This work was supported by the UK Engineering and Physical Sciences Research Council for their support for this work under grant EP/E015522/1.

Neil J. Grabham, Neil M. White, and Stephen P. Beeby are with ECS – Electronics and Computer Science at the University of Southampton, Southampton SO17 1BJ, UK (email spb@ecs.soton.ac.uk).

Mark E. Lutman, is with the Hearing and Balance Centre at the Institute of Sound and Vibration Research, University of Southampton, Southampton SO17 1BJ

John E. Chad is in Biological Sciences at the University of Southampton, Southampton SO17 1BJ, UK

Matthew A. Swabey was with the University of Southampton. He is now manager of the Mixed-Signal Labs in the School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907-2035 USA.

Paul Chambers was with the University of Southampton from July 2008 to July 2009.

Index Terms— Biometric Identification, Otoacoustic Emissions, Security, Identification, Verification

I. INTRODUCTION

THE hearing process in humans is generally well known. Sound waves enter the ear and are ultimately sensed by hair cells located in the cochlea (inner ear). However, it is not widely appreciated that the outer hairs cells also function as actuators that add energy to the incoming sound wave. This active process is believed to overcome resistive losses within the ear. The active process enables humans to detect low levels of sound which would otherwise be lost in the ear. Otoacoustic emissions (OAE) are emitted sounds from the ear which occur as a by-product of this active process [1]. During the active process, energy invariably escapes from the ear causing sound to be re-emitted. OAEs were originally theorized in 1948 by Gold [2], but they were not experimentally detected until 1978 by Kemp [3], who measured the sound signal emitted from the ear cavity in response to a click stimulus.

The signal that Kemp measured in response to a click is now known as a transient evoked otoacoustic emission (TEOAE). There are, in fact, a range of OAEs that can be recorded and these are summarized as follows [4]:

1) Transient Evoked Otoacoustic Emission (TEOAE): the complex sound emitted from the ear canal several milliseconds after a very short click stimulus has been applied. TEOAEs are sometimes termed ‘cochlear echoes’.

2) Spontaneous Otoacoustic Emission (SOAE): continuously emitted tones that can be detected in at least 30% of young healthy ears without any stimulus [4].

3) Stimulus Frequency Otoacoustic Emission (SFOAE): emissions resulting from the application of a single frequency stimulus. They are the result of reflections of travelling waves within the cochlea escaping via the middle ear and adding to the sound energy in the ear canal.

4) Distortion Product Otoacoustic Emission (DPOAE): sound emitted from the ear in response to a stimulus comprising two tones at different frequencies. Because the active process is nonlinear, intermodulation distortion is created at sums and

differences of integer multiples of the input frequencies, forming the distortion products.

Otoacoustic emissions are recognized clinically as being indicative of the status of cochlear function [5, 6, 7]. The TEOAE is the most frequently used method of clinically evaluating OAEs due to its comparatively robust response; there are several commercially available products that measure TEOAE (for example the Otodynamics Echoport series).

Biometrics refer to the methods used for recognizing individual human subjects based upon one or more intrinsic physiological characteristic or behavioural trait. Common example methods include fingerprints, retina scan, facial recognition and gait analysis. The shape of the ear has been used as a biometric [8] and previous studies have indicated that OAEs contain information unique to an individual, suggesting they are a physiological biometric. The first observation was differences in the frequency and number of spontaneous otoacoustic emissions between the genders [9]. Following this, Whitehead *et al* observed the difference in emissions from individuals from different ethnic origins [10]. This paper presents an overview of research carried out at the University of Southampton, exploring the use of OAEs as a biometric. The motivation for the research is that OAEs offer potential to perform an unobtrusive biometric analysis during any process where the individual is using headphones, telephone handsets or mobile phones.

This paper summarizes the early work of Swabey *et al* [11, 12] who devised a biometric analysis based upon TEOAE data in the time domain, which is presented with new data, and includes the latest results from an expanded cohort of test subjects. In addition to the expanded early work, this paper presents an evaluation of DPOAEs for use as a biometric along with a combination of DPOAE and TEOAE results to form a multimodal analysis. The new dataset contains information from both ears, compared to the analysis on one ear per individual used to date, which provides improved biometric performance. Finally, the paper presents the initial results from measures employed to reduce the recording time (or improve detail) using maximum length sequencing.

II. OTOACOUSTIC EMISSIONS

TEOAEs are extensively used to screen neonates for hearing difficulties [13]. TEOAEs are stimulated with a short acoustic pulse, ideally tending towards an impulse with a flat frequency spectrum. In practice the stimulus tends to be approximately 1 ms in duration and have lower and upper 3 dB cut-off frequencies of around 0.5 and 5 kHz. The stimulus typically possesses a peak magnitude of around 80 dB sound pressure level (SPL). The TEOAE emerges approximately 2–3 ms after the stimulus and finishes 10–20 ms after the

stimulus. They are collected using a probe positioned in the opening of the ear containing an earphone and microphone [5] (see figure 1). The probe should be suitably sealed into the ear canal in order to reliably detect the emission. To capture a typical set of data, the stimulus is applied and the resulting TEOAE recorded 512 times which is then averaged to improve the signal-to-noise ratio by $\sqrt{512}$. Recorded signals commenced 3 ms after the stimulus onset, in order to avoid capturing artefact caused by ringing of the stimulus [14].

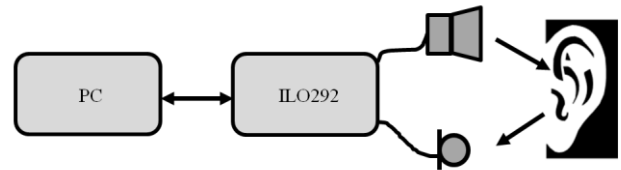


Fig. 1. Block diagram of the OAE capture system

DPOAEs are typically recorded by presenting a stimulus that consists of two pure tones, at a frequency ratio of $f_2/f_1=1.2$. The response from the ear measured by the ILO292 system will be at a frequency of $2f_1 - f_2$. To perform the biometric analysis, a series of tests are carried out using a range of frequencies. The resulting plot of $2f_1 - f_2$ SPL emitted by the ear against stimulus frequency is used to form the identity vector for the individual. Sample identity vectors from four individuals are given in figure 2.

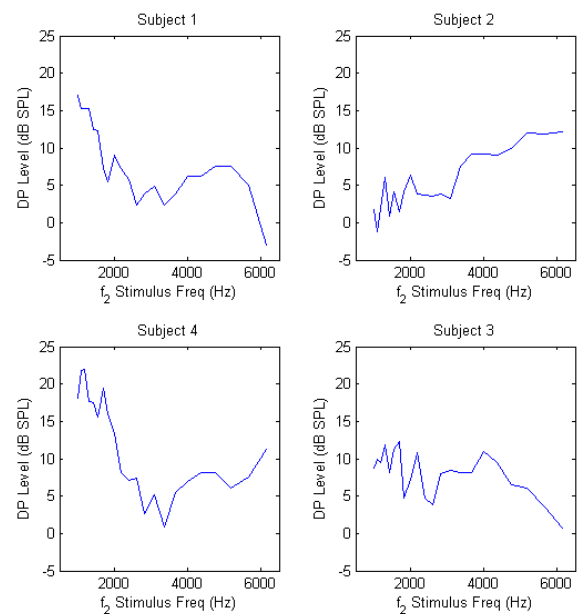


Fig. 2. DP vectors from four subjects

To be a useful biometric the OAEs should demonstrate stability over long periods. A simple visual comparison of four TEOAEs recorded from four different subjects is shown in figure 3. There is a clear difference in the appearance of the emissions from the individuals. These emissions were recorded using an Otodynamics ILO292 system with a UGD TE+DPOAE probe and ILO V6 clinical software. Owing to

subject confidentiality this data set is not available to the public.

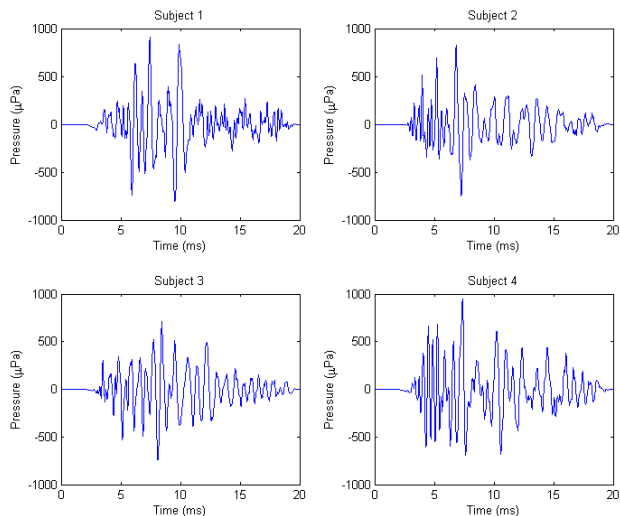


Fig. 3. TEOAE in the time domain showing the responses from four individual subjects.

A. Visual Analysis

The repeatability and stability of OAEs in the short and longer term can also be quickly demonstrated by a simple visual comparison. Figure 4 shows 10 sample TEOAE recordings from a subject overlaid in the same plot. Each recording made using the Otodynamic ILO292 system consists of a number of individual transient recordings which have been individually validated for signal to noise ratio before being averaged together to reduce overall noise, this is an internal function of the instrument. These were recorded at the same sitting. Figure 5 shows 10 TEOAE recordings from the same individual overlaid in the same plot comprising two sets of five recordings taken 6 months apart.

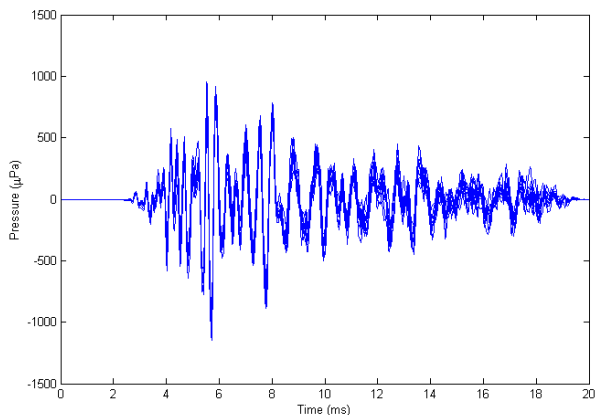


Fig. 4. Ten TEOAE recordings from the same individual recorded at the same sitting.

TABLE I
EXPERIMENTAL DATASET PROPERTIES

Parameter	Dataset 1	Dataset 2	Dataset 3	Dataset 4
No. of Subjects, n	23	760	561	33
No. of TEOAE per subject, m	10	2	2	4L + 4R
Stimulus type	DNLR, 80 dB SPL	Impulse, 70 dB SPL	DNLR, 70 dB SPL	DNLR, 80dB SPL
Sampling Frequency	25 kHz	12.5 kHz	12.5 kHz	25 kHz
No. of recordings per TEOAE avg.	512	256	256	512
Length of Recording	20.48 ms	20.48 ms	20.48 ms	20.48 ms
Elapsed time between recordings	6 month	Single sitting	Single sitting	Single sitting
Age range of subjects	Adults 20-60 years	Neonates (newborn)	Neonates (newborn)	Adults 20-40 years

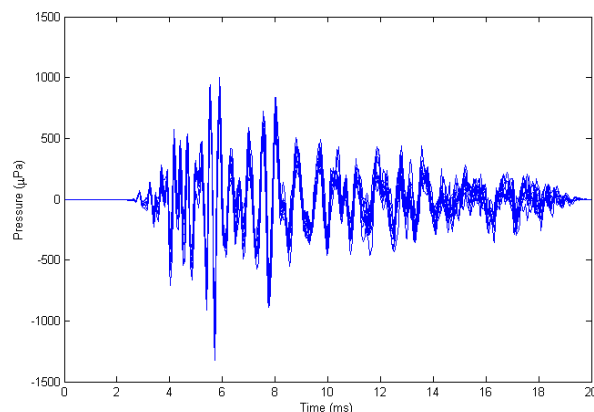


Fig. 5. Ten TEOAE recordings from the same individual comprising two sets of five recordings taken 6 months apart.

From visual inspection we can see that TEOAEs are highly repeatable in the short term and even over a period of six months. Furthermore, the appearance of the TEOAEs suggests a high degree of idiosyncrasy. A more comprehensive analysis to support this argument is provided by Swabey *et al* [12].

III. BIOMETRIC PERFORMANCE

Whilst visual inspection provides an indication of the potential for using OAEs as a biometric, it needs to be quantified in order to determine performance. This has been achieved by determining the False Accept Rate (FAR) and False Reject Rate (FRR). The FAR is construed as the proportion of fraudulent users incorrectly passed as valid and the FRR is the proportion of valid users incorrectly rejected. Another metric used to compare the performance of biometrics is the Equal Error Rate (EER) which is the percentage error at which the FAR and FRR are equal. In all these metrics a lower percentage value indicates better performance. Within a particular application, a higher FAR might be accepted to ensure a lower FRR or vice versa.

A. Experimental Datasets

For the initial investigation of biometric performance, four TEOAE datasets have been examined. Three of these were collected by Lutman *et al* [7] originally for the purpose of assessing screening. A further set has been collected specifically for the purposes of biometric analyses and has continually been updated and expanded. The properties of the datasets are summarised in Table 1 and described below.

Dataset 1: The adult DNLR set

These emissions were collected using the derived nonlinear method (DNLR) technique which is used to reduce artefacts and increase the signal to noise ratio (SNR) for TEOAEs [6]. It exploits the invertible nature of TEOAE, where if an equal and opposite stimulus is applied an equal and opposite TEOAE is produced. It also relies on the nonlinear relationship between emission and stimulus amplitude. For this dataset, four stimuli were applied, three lower level (84 dB peak equivalent SPL) whilst the fourth was negative with a magnitude three times greater than the lower level. The four resulting TEOAEs were then summed together, thereby eliminating all linear components of the TEOAE. This dataset was recorded from 23 adult volunteers from students and staff at the University of Southampton aged between 18–50 years. The TEOAEs were acquired in one sitting using an Otodynamics ILO292 instrument in a quiet room environment. Five separate recordings of two TEOAE pairs were obtained. The probe was removed and reinserted between recordings, with care taken to maintain position. After each fitment, individual TEOAEs were scaled to have a magnitude of -1 to 1 which is done to reduce the effects of differing signal coupling which is in turn due to changes in position of the probe in the ear and amount of damping.

Datasets 2 and 3: Neonates

The data in sets 2 and 3 were collected from neonates as part of a large study to test whether TEOAEs were a sensitive test of hearing impairment. The tests were carried out at various hospitals using the Programmable Otoacoustic Emission Measurement System (POEMS), which was custom-built for this experiment due to the lack of any commercial units at that time [5]. Two separate datasets have been recovered from these data, after all recordings from individuals with hearing difficulties and duplicates were removed. Dataset 2 contains 760 raw TEOAE, comprising the pressure recording of the sound at the entrance of the ear canal without signal processing. Dataset 3 which is taken from the same cohort consists of 561 DNLR TEOAE recordings. The reduced number of recordings is due to some of the subjects in Dataset 2 not having a valid secondary recording with an inverted stimulus which is required for use with the DNLR technique.

Dataset 4: Biometric set

For each subject both transient-evoked (TEOAE) and distortion-product (DPOAE) emissions were recorded. To

TABLE II
EXPERIMENTAL DATASET PROPERTIES

Parameter	
Predicted EER	EERs predicted by the crossovers of the: Lower 95% confidence interval CDF, Optimally fitted CDF, Upper 95% confidence interval CDF
Actual False Acceptances	Actual number of false acceptances at each of the three MLE thresholds
Actual False Rejections	Actual number of false rejections at each of the three MLE thresholds
FARs	FARs at the each of the three MLE thresholds, with confidences from the 'Rule of 30' or 'Rule of 3' as appropriate
FRRs	FRRs at the each of the three MLE thresholds, with confidences from the 'Rule of 30' or 'Rule of 3' as appropriate
d'	The d' scores for the three CDFs at the Lower 95%, Optimal and Upper 95% intervals

allow for dual-ear verification techniques to be assessed, four recordings of both emissions were made per ear, per subject which gives a total of 16 separate emission recordings per subject. All recordings were performed using the ILO292 system described previously, in a typical office environment, with the ear probe removed and refitted between recordings and all recordings completed in one session. The TEOAE were recorded using a DNLR technique to improve the SNR (as used with Dataset 1) and the DPOAE recordings used a fixed frequency ratio of 1.22 with stimulus levels of 65 and 55 dB SPL for stimulus frequencies f_1 and f_2 respectively. Data were collected from 33 healthy adults, none of whom featured in any of the other data sets.

B. Data Analysis

Each TEOAE waveform was analysed by treating the waveform samples as a vector. The Euclidean distance between vectors was used to measure separation between individuals. The sets of genuine and impostor distances generated are then quantified by fitting parametric statistical distributions. The protocol used is described in detail by Swabey *et al* [12] and is based upon the work of Mansfield and Wayman who address the use of statistical tools to evaluate the performance of a biometric where there is a lack of data [15]. The protocol consists of standard statistical estimation (described below) along with numerical checks to ensure their validity:

1. A set number of TEOAE recordings from one subject are averaged to form a template. These recordings are then excluded from the remainder of the analysis.

2. Compare individual TEOAE recordings against the templates from the same subject and from the other subjects and calculate the intra- and inter-subject distances.

3. Fit a probability density function (PDF) to the intra- and inter-subject distances using maximum likelihood estimation (MLE) [16] from the family of lognormal and normal

TABLE III
EXPERIMENTAL RESULTS

Parameter	Dataset 1	Dataset 2	Dataset 3	Dataset 4
Predicted EERs (%)	0.47/1.17/2.59	1.77/2.29/2.94	3.39/4.58/6.04	N/A
Actual FARs (%)	0.592/1.27/1.95	1.66/2.18/2.86	3.5/4.6/5.94	0.47/2.64/12.31
No. of actual FARs	21/45/69	9560/12593/16472	10997/14453/18654	318/1790/8322
Rule of 30 FAR confidence (%)	$\pm 50/\pm 30/\pm 30$	$\pm 10/\pm 10/\pm 10$	$\pm 10/\pm 10/\pm 10$	$\pm 10/\pm 10/\pm 10$
Actual FRRs (%)	0.0/0.0/0.0	2.5/2.11/1.45	0.89/0.89/0.54	4.83/2.65/1.61
No. of actual FRRs	0/0/0	19/16/11	5/5/3	102/56/34
Rule of 3 FRR lower bound (%)	1.86	N/A	N/A	N/A
Rule of 30 FRR confidence (%)	N/A	$\pm 50/\pm 50/\pm 50$	N/A	$\pm 30/\pm 30/\pm 30$
d'	8.9/8.55/8.19	5.17/5.05/4.92	6.56/6.43/6.29	N/A

continuous PDF functions respectively. The MLE process generates the maximally likely distribution parameters and the 95% confidence interval relating to those parameters.

4. Using the parameters from the MLE process, calculate to a desired accuracy the crossover point of the distribution cumulative density functions (CDF) using a gradient descent algorithm. This value is the decision threshold that equalizes the FAR and FRR, giving the EER, and is the maximally likely threshold (MLT).

5. Multiple distributions corresponding to the 95% confidence interval per parameter and optimal parameter are calculated. We obtain four distributions for both the intra- and inter-subject distances; the maximum and minimum crossover values from these distributions are then used as the upper and lower thresholds that give a conservative estimate of the upper and lower 95% confidence bounds on the MLE calculated EER.

6. Other statistical and numerical results, as defined in Table II, are derived from the dataset at the three values of the decision threshold as obtained from the MLE process to guarantee its credibility. By taking the observed FAR and FRR error rates, with bounds determined using the ‘Rule of 30’ or ‘Rule of 3’ as appropriate [17], we can assess the quality of fit in the tail of the inter- and intra-user distance PDFs.

C. Results

The numerical results from the data analysis are shown in Table III for datasets 1, 2 and 3. The values given for each property correspond to the lower 95% confidence interval, the MLT, and at the upper 95% confidence interval. For dataset 4 the values are for lower 95% confidence interval for EER, measured EER and upper 95% confidence interval for EER. Predicted EER and d' are not available as an MLT analysis was not done for this dataset. The FAR and FRR error rate confidence interval is estimated from the ‘Rule of 30’ or one of its variants at 90% confidence where the figure is presented as a mid-point and the \pm bound. Where no errors have occurred (i.e. FRR for dataset 1), the ‘Rule of 3’ at 95% confidence has been applied. The number of errors for the

FRR on dataset 3 ($0 < \text{FRRs} < 10$) means that neither the Rule of 30 nor the Rule of 3 can be applied.

The tabulated results show that dataset 1 demonstrates very promising performance from such a simple analysis. Using the rule of 30 on the FAR result we can see that the predicted EER is within the $\pm 30\%$. For the FRR, the rule of 3 only supports a lower bound of 1.86% at 95% confidence interval. If we take a 90% confidence interval then the predicted EER becomes 1.24%. This dataset was captured in a quiet room environment with low levels of background noise (~ 30 dB SPL).

For dataset 2, the FARs and FRRs closely correspond to the predicted EER, which is 2.29%. Due to the large number of comparisons for the FAR analysis the confidence interval from the rule of 30 is good being ± 10 . However, due to the small number of FRR errors the confidence interval from the ‘Rule of 30’ has moved from ± 30 to $\pm 50\%$.

The results for dataset 3 show significant deviation from the fitted distribution, and this is reflected in Table III by the deviation in the FRR from the predicted level. In this case, however, the actual FRR is much smaller (i.e. better) than the predicted EER. However, there are not sufficient FRR errors to establish a confidence for the rule of 30, nor are there zero errors which would allow the rule of 3 to be used. Overall, the results from this dataset are worse than the other two sets.

To illustrate the biometric performance of TEOAEs a Receiver Operating Characteristic (ROC) plot is shown in figure 6 for the TEOAE data from dataset 1, compared using the Euclidean distance technique. From the plot an EER of 2.3% can be measured, which is within the calculated 95% confidence interval reported in Table III

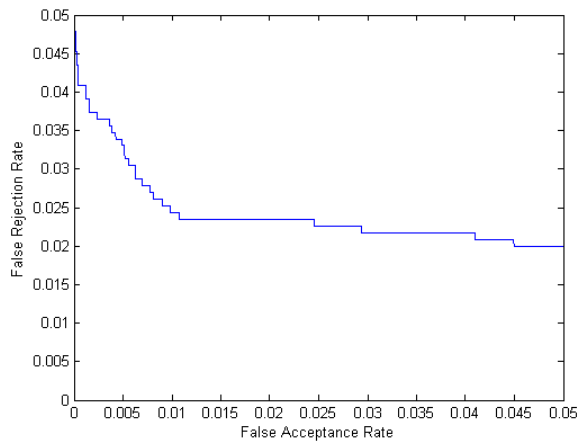


Fig. 6. ROC plot for TEOAE data from dataset 1 compared using the Euclidean distance technique.

IV. ANALYSIS DEVELOPMENT

The early work exploring the potential for OAE-based biometrics has most recently been expanded upon by investigating other factors and parameters, which may improve biometric performance. Previous attempts to improve biometric performance by exploring a Cepstral analysis has been presented elsewhere [18]. Experimentally the Euclidean distance was found to perform better in the temporal domain than in the cepstral domain. The cepstrum transformation was not found to improve the performance of the biometric system and only served to obscure the biometric information.

The optimisation presented in this paper addresses the biometric performance obtained when using DPOAEs, combining DPOAEs and TEOAEs to form a multimodal analysis and using both ears in the biometric analysis. Furthermore, the length of time taken to record suitable data has also been investigated by using maximum length sequencing to speed up data capture.

A. DPOAE Based Biometrics

DPOAE were recorded using the ILO292 system in the DP test mode and this forms part of dataset 4. Tests were performed for 22 frequency pairs with resulting emission frequencies in the range of 660 to 4110 Hz. The SPL of the $2f_1-f_2$ distortion product (DP) at each of these 22 frequencies is used to form an identity vector for that user. DP data have been recorded for the subjects in data set 4, with four recordings taken from each ear. The DP values for four subjects from data set 4 are plotted in figure 2 for comparison. Figure 7 shows four separately recorded sets of DP values for the same subject.

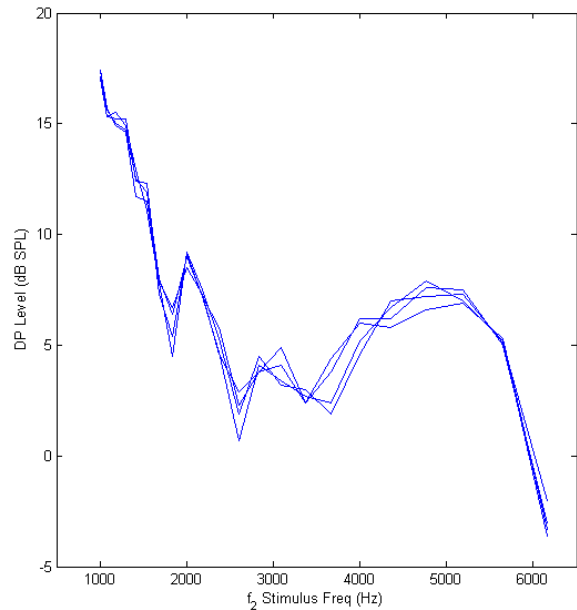


Fig. 7. Four DP vectors from the same subject

Again, the visual inspection indicates that individuals respond differently to the DP stimulus and that biometric information is contained within the DPOAE. Figure 7 indicates there is a reasonable degree of repeatability. As with the TEOAE analysis, the data need to be analysed quantitatively to assess biometric performance.

The initial quantitative test carried out compared a specific ear to all other ears in the dataset, including the other ear from the same subject. Threshold values were found for zero false accept rate, zero false reject rate and equal error rate and these values are reported in Table IV.

Through investigation of the error patterns, it was found that in many cases the closest false acceptance was due to a match with the other ear on the same subject. For instance, when testing using a subject's left ear response this may well be falsely matched to the subject's right ear. This is an understandable observation since the ears from a particular individual are not independent. Therefore, the same tests were again performed but verifying by subject, rather than by ear, whereby returning a match for a given subject's right ear when testing against the emission from their left ear would not be classified as a false acceptance. These values are also shown in Table IV. By adopting this approach the EER reduces from 4.17% to 2.64% for this data set when using the validation criterion defined above

As can be seen from the data in the table, the equal error rate has been reduced by adopting a by-subject verification rule instead of a by-ear rule. As we are primarily concerned with identifying a particular person, using the by-person schema provides improved performance over the by-ear schema without reducing the security of the system.

B. Multimodal TEOAE and DPOAE Analysis

In order to provide a richer biometric with the potential to further improve subject discrimination, tests have been performed on using both the transient evoked and distortion produce otoacoustic emissions within the verification process to give a multimodal biometric [19]. In the verification system utilized in these tests, independent verification is used for the two emissions. The results from these tests are then combined to give an overall accept/reject decision. When combining the two results, both tests must have been passed to give an overall acceptance, with a failure on either test giving rise to an overall rejection. The effect of this is to reduce the rate of false acceptances. By adjusting the threshold levels independently for the TEOAE and DPOAE tests, the overall response of the system can be tailored to the requirements.

The verifications have again been performed using the Euclidian distance technique and applied to data set 4. The analysis has been performed for testing from the full data set of left and right ears and also for left and right ears independently. For such comparisons, the overall error rate (false acceptance rate and false rejection rate added) has been minimized through selection of appropriate threshold values for both the TEOAE and DPOAE tests. An equal error rate was not calculated for this approach because of the additional complexity of the multimodal analysis. The recordings used to form dataset 4 were made using the ILO920 system in both TE and DP modes, with recorded TEOAE emission following

TABLE IV
INITIAL DISTORTION PRODUCT TESTS

Parameter	By Ear	By Subject
Threshold	35 / 77 / 150	39 / 73 / 145
FAR (%)	0 / 4.17 / 48.96	0 / 2.64 / 48.42
No. of Actual FAs	0 / 2866 / 33612	0 / 1790 / 32728
Rule of 30 FAR Confidence (%)	N/A / ± 30 / ± 30	N/A / ± 30 / ± 30
Rule of 3 FAR lower bound (%)	0 / N/A / N/A	0 / N/A / N/A
FRR (%)	28.97 / 4.16 / 0	12.31 / 2.65 / 0
No. of Actual FRs	303 / 44 / 0	260 / 56 / 0
Rule of 30 FRR Confidence (%)	± 30 / ± 30 / N/A	± 30 / ± 30 / N/A
Rule of 3 FRR lower bound (%)	N/A / N/A / 0.28	N/A / N/A / 0.14

Values are for: Zero FAR / EER / Zero FRR

the same form as those shown in figure 3, and the DPOAE vectors the same form as shown in figure 2.

TABLE V
COMBINED TEOAE AND DPOAE

Parameter	All Ears	Left Ears Only	Right Ears Only
Total Error	7.11%	7.10%	4.81%
Overall FAs	852	432	174
Overall FAR	1.24% $\pm 30\%$	2.55% $\pm 30\%$	1.03% $\pm 30\%$
Overall FRs	62	24	20
Overall FRR	5.87% $\pm 30\%$	4.54% $\pm 50\%$	3.78% $\pm 50\%$
TEOAE FAs	13730	12766	2120
TEOAE FAR	20.00% $\pm 30\%$	75.55% $\pm 30\%$	12.54% $\pm 30\%$
TEOAE FRs	26	0	0
TEOAE FRR	2.46% $\pm 50\%$	0% LB=0.57%	0% LB=0.57%
DPOAE FAs	2866	548	596
DPOAE FAR	4.17% $\pm 30\%$	3.24% $\pm 30\%$	3.52% $\pm 30\%$
DPOAE FRs	44	24	20
DPOAE FRR	4.16% $\pm 30\%$	4.54% $\pm 50\%$	3.78% $\pm 50\%$

NB: LB is lower bound from rule of 3

The results for the minimized total error values are shown in Table V. As can be seen from the table, combining the two modalities provides an overall biometric with the false acceptance rates that are reduced compared to the individual TEOAE or DPOAE tests. It should be noted, however, that the performance figures given for the individual TEOAE and DPOAE are based upon the threshold values set during the operation to combine the analysis and are therefore not optimised for each individual analysis. This is most clearly apparent when analysing the TEOAE results where the threshold used has skewed the results giving high FAR and low FRR percentages. The discrepancy between the performance of the left and right ears may be attributable to one or more subjects who had lower quality responses from their left ears due to hearing damage or a build-up of matter within the ear canal. Further investigation with a larger cohort could enable the effects of this to be reduced.

C. Both Ears DPOAE

To further improve the security provided by the system, the same technique used to combine the TEOAE and DPOAE responses can be used to provide verification based upon recordings from both ears. This technique has been assessed for use with the DPOAE recordings in dataset 4. To perform the tests, four separate DPOAE recordings are taken from each of a subject's ears, these identity vectors are then combined in pairs of left and right vectors to give a total of sixteen double length vectors consisting of the data from the left ear followed by the right ear. The sixteen double length vectors created contain all the valid combinations of left and right vectors for that subject.

The Euclidian distance technique was again used to compare the vectors and generate false accept and reject rates by varying the threshold values used. Values have been found for data set 4 for zero false acceptances, equal error rates and for zero false acceptances and are shown in Table VI. The dual ear approach gives an EER of 2.70%.

By comparing the data in Tables VI with that in Table IV, which shows data for tests performed on identities comprising

a single ear, the equal error for the system using both ears rate can be seen to be comparable to that achieved for subject level identification but superior to that achieved for an ear level identification scheme. The false rejection rate at zero false accepts has also reduced. Some improvement over the by-subject identification through single ear DPOAE is seen overall, though this must be offset against the increased measuring requirements as recordings are required from both ears.

If it is desired to operate the identification with either a zero FAR or FRR the use of the dual ear DPOAE approach will provide a lower non-zero acceptance rate than a system using a single ear.

D. Maximum Length Sequences

When recording the TEOAE, it is necessary to make multiple records of the subject's ear response to the stimulus, averaging these recordings to improve the signal to noise level. In standard testing, recordings do not overlap and therefore the required capture time is the duration of a single recording multiplied by the number of recordings. The conventional approach typically took around 60 seconds to average the 512 recordings. This could prove an obstruction to the practical deployment of a biometric system. The use of maximum length sequences (MLS) has been investigated for its suitability to reduce the required capture time for a TEOAE [20].

As with the standard approach, the MLS technique uses a series of pulses to evoke the OAE but the time interval between pulses can be shorter, allowing the responses to overlap. The averaged OAE response is obtained by deconvolution, based on the assumptions that responses sum linearly and the sequence of pulses conforms to the requirements of a MLS. One drawback of this approach is the physiological reduction in the amplitude of the TEOAE with increasing rate of stimulation. This behaviour results in a trade-off between SNR of the averaged response and test time [21].

In this work, MLS of orders 1, 2, 3 and 4 were assessed in conjunction with a single click stimulus; details of the MLS used are shown in Table VII. The number of bits represents the number of time frames in which a stimulus click might be present; the number of clicks entry shows how many of these time frames contain a click; the sequence entry represents the distribution of stimulus clicks within the time frames with a 1 representing a click present in the corresponding time frame, and a 0 representing no stimulus present in that time frame. Finally the clicks per 5 s entry represents the number of click stimuli that are presented and averaged to give a 25-ms emission recording during a 5 second recording process. A MLS sequence of order 1 is equivalent to normal TEOAE recording as there is only one stimulus click within the sampling period and therefore no overlapping of responses.

A TEOAE capture system was implemented using National Instruments data acquisition hardware and a Labview software environment, with the signal recovery performed using functionality coded in MATLAB. For the first test, the system was configured for each MLS order in turn to capture emissions for a total of 5 seconds and to average the results to give a single 25-ms duration emission recording. Within a 5-second recording window, the first order MLS can involve 200 clicks; in the same time the second order realises 400 clicks, the third order realises 792 clicks and the fourth order realises 1568 clicks.

TABLE VII
MLS PARAMETERS

Parameter	Order 1	Order 2	Order 3	Order 4
No. of bits	1	3	7	15
No. of clicks	1	2	4	8
Sequence	1	101	1001110	000100110101111
Clicks per 5 s	200	400	792	1568

Recovered TEOAE responses, captured using the MLS technique and a 5-second recording time span, are shown for one subject in figure 8. In the figure, the initial 3 ms of the recording is not plotted as this contains the stimulus artefact. The response recorded can be seen to be in good agreement up to 15 ms, at which point the emission level from this test subject has decayed. It is expected that further improvement in recording stability can be achieved by increasing the recording duration. For comparison, the ILO292 clinical system typically utilises a recording duration of at least 60 seconds due to the sample processing, noise filtering, level checking and stimulus presentation rate, though this duration may be increased if samples are rejected due to background noise levels.

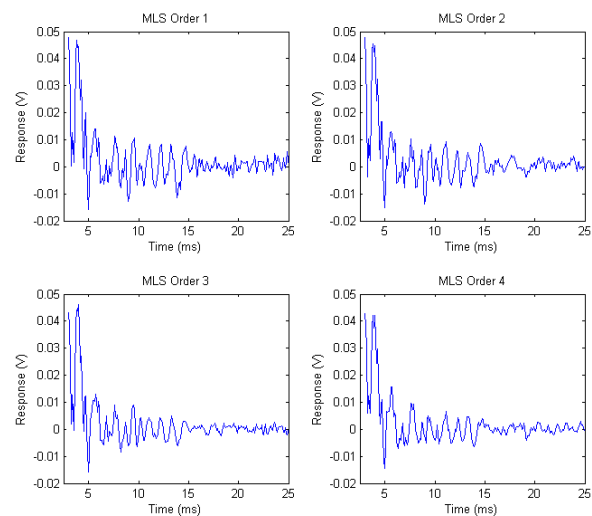


Fig 8. TEOAE recordings captured from one subject using MLS technique with orders 1, 2, 3 and 4

Using the MLS of order 4, with a 5-second recording time frame provides repeatable results. Figure 9 shows five separately recorded emissions from the same subject with the probe being removed from, and then replaced in, the subject's ear between each of the five separate recordings. These recordings can be seen to be in good agreement.

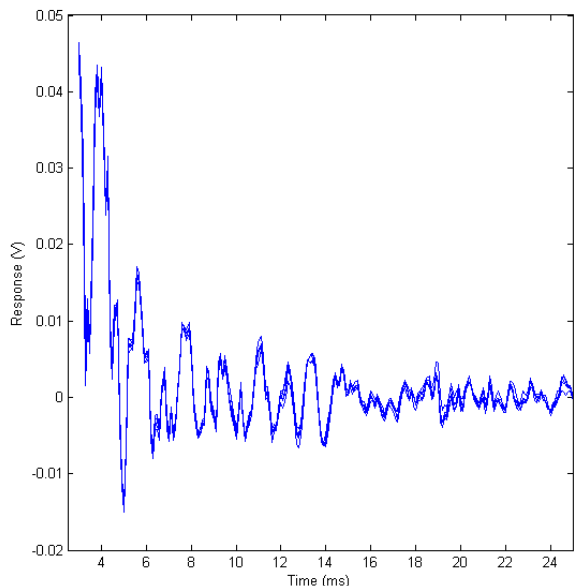


Fig. 9 Five TEOAE recordings from the same subject using a 4th order MLS.

The relative signal to noise ratio (SNR) of the MLS acquired TEOAE recordings, referenced to the conventionally acquired response, has been calculated using the equation provided by Bell *et al.* [22]. In all cases the total recording period was 5 s, with a 25-ms emission recording duration. With the four MLS orders tested, order 1 recordings are used as the baseline. The resulting relative SNR for 2nd, 3rd and 4th order MLS sequences are 0.942, 1.137, and 1.4 respectively. From this it can be seen that the use of 2nd order MLS sequence does not improve the signal to noise ratio, but that the 3rd and 4th order MLS do provide an improved signal to noise ratio, when compared with the conventional technique recording for the same time. Further improvement of the SNR may be possible by using higher order MLS. However, as the interval between clicks reduces, the level of the recorded

TABLE VI
DPOAE USING BOTH EARS

Parameter	Zero False Accepts	Equal Error Rate	Zero False Rejects
FAs	0	7308	106526
FAR	0% LB=0.00%	2.70% ±30%	39.40% ±30%
FRs	768	228	0
FRR	9.09% ±30%	2.69% ±30%	0% LB=0.04%

NB: LB is lower bound from rule of 3

TEOAE may also reduce due to physiological effects of the repeated stimulation, therefore reducing the response to each of the individual stimuli in the recording.

V. CONCLUSIONS

TEOAE and DPOAE are clearly distinctive and they can be used to distinguish between individuals. Individual templates are small in size and the basic Euclidean distance biometric analysis is quick to perform. Of the two types of OAE, TEOAE appear to offer the better biometric performance, but this may be due to the reduced degrees of freedom in the DPOAE analysis (22 frequencies) which is limited by the ILO292 system.

There is a variation in the biometric performance results obtained from the different datasets, which highlights the importance of controlling the method of capture to ensure repeatability and consistency. Important parameters appear to be OAE type, recording method, equipment type and background noise level. Further research is required to explore these parameters more fully.

ACKNOWLEDGMENT

The authors wish to gratefully acknowledge funding from the EPSRC under grant EP/E015522/1 and the assistance provided by the test subjects from whom the otoacoustic emission recordings were made.

REFERENCES

- [1] S.A. Gelfand, *Hearing (An Introduction to Psychological and Physiological Acoustics)*, New York: Marcel Dekker, 4th Ed, 2004, pp.165-169.
- [2] T. Gold, "Hearing. II. The Physical Basis of the Action of the Cochlea," *Proceedings of the Royal Society of London. Series B, Biological Sciences*, vol. 135, no. 881, pp. 492-498, 1948.
- [3] D. T. Kemp, "Stimulated Acoustic Emissions from Within The Human Auditory-System," *Journal of the Acoustical Society of America*, vol. 64, no. 5, pp. 1386-1391, 1978.
- [4] M. S. Robinette and T. J. Glattke, *Otoacoustic Emissions: Clinical Applications*, New York: Thieme, 3rd ed, 1997.
- [5] Y. Cope and M. E. Lutman, "Otoacoustic Emissions," *In Paediatric Audiology 0-5 years*, B. McCormick, ed., London: Taylor & Francis, pp. 221-245, 1988.
- [6] F. Grandori and P. Ravazzani P, "Non-linearities of click-evoked otoacoustic emissions and the derived non-linear technique," *British Journal of Audiology*, vol. 27, no. 2, pp. 97-102, 1993.
- [7] M. E. Lutman, A. C. Davis, H. M. Fortnum and S. Wood, S. "Field sensitivity of targeted neonatal hearing screening by transient-evoked otoacoustic emissions," *Ear and Hearing*, vol. 18, no. 4, pp. 265-276, 1997.
- [8] M. J. Burge and W. Burger. "Using ear biometrics for passive identification," *In 14th International Information Security Conference*, Vienna, Austria, 1998, pp. 139-148.
- [9] R. C. Bilger, M. L. Matthies, D. R. Hammel and M. E. Demorest, "Genetic-Implications of Gender Differences in the Prevalence of Spontaneous Otoacoustic Emissions," *Journal of Speech and Hearing Research*, vol. 33, no. 3, pp. 418-432, 1990.
- [10] M. L. Whitehead, N. Kamal, B. L. Lonsburymartin and G. K. Martin, "Spontaneous Otoacoustic Emissions in Different Racial Groups," *Scandinavian Audiology*, vol. 22, no. 1, pp. 3-10, 1993.
- [11] M. Swabey, S. Beeby, A. Brown and J. Chad, (2004) "Using Otoacoustic Emissions as a Biometric," *in Proc. First International Conference on Biometric Authentication (ICBA 2004)*, Hong Kong, 15-17th July 2004, pp. 600-606.
- [12] M. A. Swabey, P. Chambers, M. E. Lutman, N. M. White, J. E. Chad, A. D. Brown and S. P. Beeby, "The Biometric Potential of Transient Otoacoustic Emissions," *International Journal of Biometrics*, vol.1, no. 3, pp. 349-364, 2009.
- [13] B. R. Vohr, L. M. Carty, P. E. Moore and K. Letourneau, K, "The Rhode Island Hearing Assessment Program: Experience with statewide hearing

screening (1993-1996),” *Journal of Pediatrics*, vol. 133, no. 3, pp. 353-357, 1998.

- [14] M. S. Robinette and T. J. Glattke, “*Otoacoustic Emissions, Clinical Applications*,” New York: Thieme, Chapter 2, pp.22, 2002
- [15] A. J. Mansfield and J. L. Wayman, “*Best Practices in Testing and Reporting Performance of Biometric Devices*,” National Physical Laboratory Queens Road, Teddington, Middlesex, TW11 0LW, CMSC 14/02M, 2002,
- [16] T. A. Severini, *Likelihood Methods in Statistics*, Oxford, UK: Oxford University Press, 2000.
- [17] 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, 2000.
- [18] P. Chambers, N. J. Grabham, M. A. Swabey, M. E. Lutman, N. M. White, J. E. Chad, S. P. Beeby, “A comparison of verification in the temporal and cepstrum-transformed domains of Transient Evoked Otoacoustic Emissions for biometric identification.” *International Journal of Biometrics*, Vol. 3, No.3, pp. 246 – 264, 2011.
- [19] A. Ross, A. K. Jain, “Multimodal Biometrics: An Overview”, Proc. of 12th European Signal Processing Conference (EUSIPCO), (Vienna, Austria), pp. 1221-1224, September 2004.
- [20] A. R. D. Thornton, “Click-evoked otoacoustic emissions: new techniques and applications.” *British Journal of Audiology*, Vol. 27, pp.109-115, 1993.
- [21] B. Lineton, A.R.D. Thornton and V.J. Baker, “An investigation into the relationship between input-output nonlinearities and rate-induced nonlinearities of click-evoked otoacoustic emissions recorded using maximum length sequences.” *Hearing Research*, Vol. 219, Issues 1-2, pp. 24-35, 2006.
- [22] S. L. Bell, R. Allen, and M. E. Lutman, “The feasibility of maximum length sequences to reduce acquisition time of the middle latency response.” *J. Acoust. Soc. Am.*, Vol. 109, Issue 3, pp. 1073-1081 (2001)



Neil J. Grabham was awarded a MEng (Merit) degree in Information Engineering from the University of Southampton, UK, in 1998. He obtained his PhD from the University of Southampton, UK, in 2002 for a thesis titled ‘Development of a Thick-Film Magnetostrictive Material: Formulation and Characterisation’.

After completing his PhD he was employed as a Research working on a range of projects covering thick-film materials, biometric identification, energy harvesting, sensor devices and wireless sensor networks.

He is currently employed as a Senior Research Fellow with Electronics and Computer Science at the University of Southampton and is currently working on an European Union funded project, TIBUCON, concerning the use of energy harvesting for building management systems. He has authored chapters in the books ‘Electroceramic-Based MEMS’ (New York: Springer, 2005) and ‘Energy Harvesting for Autonomous Systems’ (Norwood, Mass.: Artech House, 2010). His current areas of research cover sensor technology, biometrics, energy harvesting and wireless sensor networks.

Dr Grabham is a Member of the Institution of Engineering and Technology and a Member of the Institute of Physics.



Matthew Swabey obtained a BEng (Honours) degree in Electronic Engineering in 2001. He was awarded his PhD in 2006 for his thesis “The Human Auditory System as a Biometric” from the University of Southampton.

He was appointed as a Research Fellow in 2005 continuing his work in OAE based biometric systems. He became a Teaching Fellow in 2007 at the University of Southampton. In 2011 he joined Purdue University as the Manager of the Mixed Signal Laboratories. He has acted as a consultant to Spirax Sarco, Roke Manor, IBM, Synopsys and D4 Technology Ltd. He has 7 publications in the field of biometrics and IC design.

Dr Swabey won the best paper prize at the 2004 Int. Conf. on Biometric Authentication and a Vice-Chancellors teaching award in 2007.

Paul Chambers obtained his PhD in 2005 from the University of Southampton.

Following his PhD, he was employed as a Research Fellow/part-time lecturer at the University of Limerick until 2007 and at City University in the same role.

Fromm 2008 to 2009 he was a Research Fellow at the University of Southampton and is currently working at City University, London.



Neil M. White (M’01, SM ‘02) is Head of Electronics and Computer Science (ECS) at the University of Southampton. He obtained a PhD from the University of Southampton in 1988 for a thesis describing the piezoresistive effect in thick-film resistors.

He was appointed as a Lecturer within ECS in 1990 and promoted to Senior Lecturer in 1999, Reader in 2000 and was awarded a

personal Chair in 2002. He is co-author of the book *Intelligent Sensor Systems*, which was first published by the Institute of Physics Publishing in 1994. He is also co-author of the books “MEMS: Mechanical Sensors” and “Energy Harvesting for Autonomous Systems”, published by Artech House. His research interests include thick-film sensors, intelligent instrumentation, MEMS, self-powered microsensors and sensor networks. He lectures on digital electronics, electronic measurement techniques and advanced instrumentation and sensors. He has published over 200 scientific papers in the area of sensors and instrumentation systems and holds 10 patents.

Prof. White is a Chartered Engineer, Fellow of the IET, Senior Member of the IEEE, Fellow of the IoP and a Chartered Physicist. He was Chairman of the Instrument Science and Technology (ISAT) group of the IoP from 1997 to 1999. He is a former Director and co-founder of the University spin-out company Perpetuum Ltd., which specialises in vibration energy harvesting. He was the recipient of the 2009 Callendar silver Medal, awarded by the Institute of Measurement and Control for his ‘outstanding contribution to the art of instruments and measurement’.



Mark E Lutman obtained a BSc in Engineering Science in 1971, MSc in Sound and Vibration in 1972 and PhD in Audiology in 1976, all at the University of Southampton.

Following a Research Fellow position from 1975 at the University of Southampton, he was appointed as a Medical Research Council Scientist at the Institute of Hearing Research in Nottingham from 1979 until he took up his present position as Professor of Audiology and Head of the Hearing and

Balance Centre at the Institute of Sound and Vibration Research of the University of Southampton. His research interests include: signal processing for hearing aids and cochlear implants, measurement of cochlear function from otoacoustic emissions, neonatal and adult hearing screening, cochlear implantation, epidemiology of hearing impairment and balance disorder, evaluation of benefit from hearing instruments and noise-induced hearing loss.

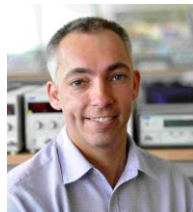
Prof. Lutman was editor of the *British Journal of Audiology* from 1991 to 1995 and President of the *British Academy of Audiology* from 2007 to 2008. He is a Fellow of the *International Collegium of Rehabilitative Audiology*.



John Chad obtained a BSc Physiology & Biochemistry, and a PhD Neurophysiology from the University of Southampton.

His background includes research as a Faculty Research Biologist at UCLA (1978-85) studying neuronal calcium signalling, and research at the Sandoz Institute for Medical Research (1985-1986). He is currently a Senior Lecturer in Biological Sciences at the University of Southampton, UK, and a member of the Southampton Neurosciences Group. In

2002 he co-founded Capsant Neurotechnologies Ltd. which is conducting research into neurodegenerative diseases. His research interests include the application of optics to neuroscience, control of neuronal excitability, neurodegeneration, neuronal self-organisation and neurocomputation. He has over 60 publications in a variety of aspects of neuroscience.



Stephen P. Beeby (M’03, SM’11) obtained a BEng (Honours) degree in Mechanical Engineering from the University of Portsmouth, UK, in 1992. He obtained his PhD from the University of Southampton, UK, in 1998 on the subject of micromechanical resonators.

Following his PhD, he became a Research Fellow in the School of Electronics and Computer Science (ECS) at the University of Southampton. He was

awarded a prestigious EPSRC Advanced Research Fellowship in 2001 to investigate the combination of screen printed piezoelectric materials with micromachined structures. Following the Fellowship, he became a lecturer in ECS, was appointed a Reader in 2008 and was awarded a personal Chair in 2011. His research interests include energy harvesting, MEMS, active printed materials development and biometrics. He is the co-ordinator of an EU Framework Integrated Project 'MICROFLEX' and is principal or co-investigator on a further 6 projects. He is a co-founder of Perpetuum Ltd, a University spin-out based upon vibration energy harvesting formed in 2004. He has co-authored one book, 'MEMS Mechanical Sensors' (Artec House, Inc., Boston, London, 2004) and co-edited 'Energy Harvesting for Autonomous Systems' (Artec House, Inc., Boston, London, 2010). He has over 150 publications in the field and 8 patents.

Prof. Beeby is a member of the Institution of Engineering and Technology and the Institute of Physics.