Testing paradigms for assistive hearing devices in diverse acoustic environments

Ram Charan M C, Hussnain Ali, *John H. L. Hansen*

Cochlear Implant Laboratory – Center for Robust Speech Systems
The University of Texas at Dallas, Richardson TX, USA
{RamCharan.ChandraShekar, hussnain.ali, john.hansen}@utdallas.edu

Abstract

Many individuals worldwide are at risk of hearing loss due to unsafe acoustical exposure and chronic listening experience using personal audio devices. Assistive hearing devices(AHD), such as hearing-aids(HAs) and cochlear-implants(CIs) are a common choice for the restoration and rehabilitation of the auditory function. Audio sound processors in CIs and HAs operate within limits, prescribed by audiologists, not only for acceptable sound perception but also for safety reasons. Signal processing(SP) engineers follow best design practices to ensure reliable performance and incorporate necessary safety checks within the design of SP strategies to ensure safety limits are never exceeded irrespective of acoustic environments. This paper proposes a comprehensive testing and evaluation paradigm to investigate the behavior of audio devices that addresses the safety concerns in diverse acoustic conditions. This is achieved by characterizing the performance of devices with large amounts of acoustic inputs and monitoring the output behavior. The CCi-MOBILE Research-Interface(RI) (used for CI/HA research) is used in this study as the testing paradigm. Factors such as pulse-width(PW), inter-phase gap(IPG) and a number of other parameters are estimated to evaluate the impact of AHDs on hearing comfort, subjective sound quality and characterize audio devices in terms of listening perception and biological safety.

Index Terms: Assistive Hearing Devices, Cochlear Implants, Hearing Aids, Research Interface, Inter-phase gap

1. Introduction

According to an estimate from the World Health Organization (WHO), almost a billion-young people worldwide are at risk of hearing loss due to unsafe listening habits [1]. Nearly 50% of teenagers and young adults aged 12–35 years, in middle- and high-income countries, are exposed to unsafe acoustic conditions, primarily from the use of personal audio devices. Depending on the level of hearing loss, hearing-aids (HAs) and cochlear implants (CIs) can be used to restore auditory function of hearing impaired individuals for those who meet the candidacy criteria [2]. Several concerns regarding experimental safety, stimulation levels (current/charge), perception, and neurophysiology for Assistive Hearing Devices (AHDs) have been addressed in literature [3]. Some of these aspects concern best-practices from a safety and ethics perspective, the prevention of biological or neural damage [4], the prohibition of uncomfortably loud presentation of sounds, and customization of stimuli presentation [5]. For electric stimulation, the guidelines from the FDA set a conservative and safe upper limit of 216 mC/cm^2 for clinical applications [6-7]; 100 mC/cm^2 is the recommended limit by the FDA for RI (investigational devices) [3]. Although, clinically appropriate loudness levels can be provided by the audiologists by measuring maximum acceptable loudness levels where gross adjustments can be performed across all electrodes, it is likely that research will involve generation of stimulation patterns with parametric values that deviate from the clinical parameters. The loudness of an electrical stimulation (i.e. a pulse) is related to its charge, which is a product of amplitude and pulse duration, both with a complex relationship with loudness perception. For lower stimulation rates of 100 pulses per second (pps), loudness is best modeled as a power function of pulse amplitude; and for higher stimulation rates (>300 pps), loudness is best modeled as an exponential function of the pulse amplitude [8]. The balance of charge between the two phases in biphasic and multiphasic pulses is designed to prevent irreversible corrosion of electrodes and the potential deposit of metal oxides at the electrode–tissue interface [9]. The magnitude of loudness increases as a function inter-phase gap (gap between cathodic and anodic pulses of a biphasic pulse) [10-11]. Reliable performance of AHDs can be assured by restricting the values of stimulation parameters to prescribed target-ranges and the necessary checks in design to ensure that the safety limits are never exceeded [12].

Furthermore, the signal processing (SP) module is central to AHDs, and the choice of algorithm, number of channels/filter-banks, architectural design, programmability, and implementation strategies effect and can influence the desired signal quality and power in HAs or electric stimulation in CIs. Other factors include processing delay, spectral and temporal resolution, signal to noise ratio, signal envelope, attack and release times for automatic-gain control etc., [13]. These attributes of the processed signal/electrical stimulation determine the speech/music quality, speech intelligibility, temporal fine structure, timbre and customizability based on user preference. Furthermore, HAs and CIs typically include device interconnectivity with external devices like smartphone, laptop or TV using Bluetooth, Wi-Fi and other extra add-on features. These devices have a strong focus on minimizing power, cost and overall device size and enhancing the life and usability of the device. AHD engineers are faced with a challenge to optimize the design within the aforementioned constraints to attain a feasible product by achieving a fine balance between irreconcilable features. Ultimately, the responsibility of delivering a safe listening experience lies with the manufacturers and researchers of AHDs.

The acoustic signal processed by the human auditory system can be perceived as speech or non-speech signals: music and environmental sounds. While the speech signals provides necessary phonetic data for the brain to process the audio message, the non-speech sounds such as music and environmental sounds provide key information for patient's daily activities (e.g., fire alarms, car horns), and perceptual

relationship to the surrounding environment which contributes to the patient's overall well-being. The impact of CIs and HAs on speech, language, and communication among the hearing-impaired and the resulting benefits in these fields are well established [14]. Music and environmental sounds remain a challenge for auditory stimuli among hearing aid and cochlear implant users. The individual aspects of music: melody, rhythm, and timbre present challenges and difficulties faced by the users with HA/CIs while listening to music in most settings. Various strategies in electrical-to-cochlear pitch mapping, pitch processing strategies, and the ability to preserve residual hearing in the implanted ears, together with an increased recognition improves the performance of music perception [15]. Research shows environmental sound perception for these patients, can be improved by having a low-cost computer based training and rehabilitation [16]. Even among experienced CI users and hearing-impaired patients with high speech perception scores, environmental sound perception remains poor.

While the evaluation of any proposed speech processing algorithm requires a corpus, test plan, and evaluation criteria, the evaluation of a HA/CI RIs are not as well defined. The proposed work is focused upon developing a robust and rigorous testing and evaluation paradigm, commonly termed as the "burn-in" process, to investigate the behavior of AHDs. Burn-in process is carried out in two phases to identify unstable, unexpected and anomalous stimulus that could inadvertently cause discomfort, pain, permanent tissue damage or threaten life itself of the CI user: (i) Audio-test phase: Test the HA/CI RI in diverse acoustical conditions spread wide across broad range of audio, speech, acoustic signal and address safety concerns by characterizing its performance over experimental observations. (ii) Researcher-test phase – Test the HA/CI RI for all the possible combinations of signal processing parameters which could be selected by a researcher for algorithmic design and characterize the output stimulation pattern for biological safety and listening perception. For the purposes of this study, only the Audio-Test Phase portion is considered.

This paradigm is characterized by: stimulation levels, loudness, customizability, perceptual and neurophysiological impact on comfort and sound quality. A speech battery of over thirteen major databases are used to simulate wide-ranging acoustical conditions. The CCi-MOBILE RI (used for CI and HA research) is used as the AHD to determine the safety and efficacy of simulated acoustic conditions [17]. Biological safety is ensured by providing symmetric biphasic pulses, where charge of the pulse is determined by the product of its amplitude, phase duration and phase reversal prevents ionic imbalance [6]. DIET (CIC4 - Decoder Implant Emulator Tool) box manufactured by Cochlear Corp., is used to record stimulation parameters such as the overall charge per second, pulse width, interphase gap, and timing errors. The correlation between stimulation parameters and recorded error indicative of unsafe conditions is shown in this study to determine the impact on biological safety and loudness levels.

2. Proposed acoustical testing and evaluation paradigm

The proposed comprehensive testing and evaluation paradigm is developed as a protocol to test AHD. This involves considering various available signals in acoustic space to test, analyze, and characterize the performance of the AHDs with respect to the acoustic environment. The acoustic space consists of speech, noise, and music as shown in Fig. 1. This test battery is used to represent acoustic signal classes based on their nature, human auditory perception, and auditory signal processing competence. The corresponding collection of acoustic corpora is shown in Fig. 2.
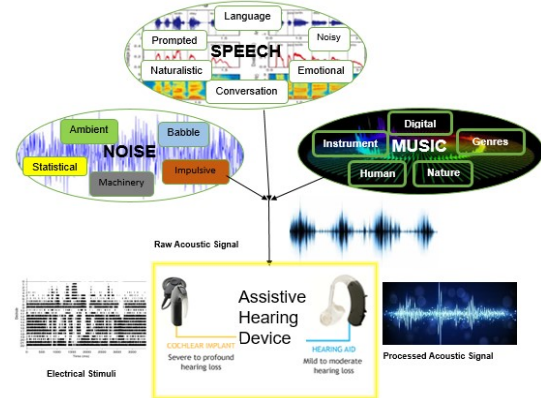

Figure 1: *Block diagram of proposed acoustic testing and evaluation paradigm*

2.1. Acoustic Space Classification

The test battery can be separated into three categories: speech, music and noise.

2.1.1. Speech

The speech test battery can be further subcategorized based on language (english, french etc.), background noise (machinery, babble etc.), type of speech production: prompted (news reporter, orator etc.) or naturalistic (candid talk, public spaces), conversational speech (telephone, video/audio chatting etc.), and emotional content (happiness, excitement, etc.). The following speech databases were considered:

a) AzBio: 1000 sentences recorded from 4 talkers, 2 male (ages 32 and 56) and 2 female (ages 28 and 30) [18].

b) IEEE: 72 lists of ten phrases each.

c) Consonant Nucleus Consonant (CNC) Test: 10 lists of monosyllabic words with equal phonemic distribution across lists with each list containing 50 words (500 words total).

d) NOIZEUS: Noisy speech corpus - 30 IEEE sentences (produced by 3 male and 3 female speakers) corrupted by 8 different real-world noises at different SNRs. Noise: AURORA database - suburban train noise, babble, car, exhibition hall, restaurant, street, airport, and train-station noise.

e) Language Database (LRE): NIST Language Recognition Evaluation Test Set - Amharic, Haitian, English, French, Hindi, Spanish, Urdu, Bosnian, Croatian, Georgian, Korean, Portuguese, Turkish, Vietnamese, Yue Chinese, Dari, Persian, Hausa, Mandarin Chinese, Russian, Ukrainian, Pushto.

f) TIMIT: 630 speakers of 8 major dialects of American English, each reading 10 phonetically rich sentences: time-aligned, orthographic, phonetic and has word transcriptions as well as a 16-bit, 16kHz speech waveform file for each utterance.

g) DARPA RATS: The Robust Automatic Transcription of Speech (RATS): All audio files are presented as single-channel, 16-bit PCM, 16000 samples per second.

2.1.2. Music

The music test battery can be subcategorized based on genres (jazz, pop etc.), instruments (violin, veena, etc.), production (digital synthesizer, gramophone, etc.), human singing (chorus, chanting, etc.), or natural musical pattern (humming bees, birds chirping, etc.). The following music databases were considered:

a) MARSYAS GTZAN Music: Genre classification [19]. 1000 audio tracks each 30 seconds long and contains 10 genres, each represented by 100 tracks.

b) MARSYAS GTZAN Music Speech: The dataset consists of 120 tracks, each 30 seconds long. Each class (music/speech) has 60 examples. The tracks are all 22050Hz Mono 16-bit audio files in .wav format

2.1.3. Noise

The noise test battery can be further subcategorized based on its statistics (pink noise, white noise, etc.), speech shape (babble, crowd noise, etc.), temporal characteristics (gunshot, explosives, etc.), machinery (air-conditioner, washing machine, etc.), and naturalistic ambience (tornado, ocean waves, etc.). The following noise databases were considered:

a) ESC: 2000 short clips comprising 50 classes of various common sound events.

b) Freesound Project: an abundant unified compilation of 250 000 unlabeled auditory excerpts extracted from recordings

c) UrbanSounds: 27 hours with 18.5 hours of annotated sound event occurrences across 10 sound classes.

d) Gunshots-Airborne: Free Firearm Sound Library – More than 1,100 files which has 7.48 gigabytes of memory, 1,106 sound effects, 192 kHz/24-bit WAV files (with some 96 kHz tracks).

Figure 2: *Collection of acoustic corpora*

2.2. Testing and evaluation scheme

The selected test battery is used to analyze the behavior of the AHD (the CCi-MOBILE RI) [3]. Acoustic signals are provided to the AHD for processing; the out of a CI is an electric signal/pulse and output from a HA is an acoustic signal. Each output is further examined for its functionality, operational accuracy, impact on loudness and establishing measures for safer listening experience.

2.3. Proposed acoustic test platform

The CCi-MOBILE RI, developed by Ali *et al.* 2016 (UT-Dallas) is used to determine the safety and efficacy of acoustic listening condition. Behind-the-ear (BTE) microphones and radio-frequency (RF) transmission coils are used to deliver the electrical signal to CI users through an interface board as described in [17]. The CCi-MOBILE RI is connected to a personal computer (PC) with access to an enormous collection of acoustical databases. The computation needed to process the digitized acoustic signal is drawn from the PC to generate corresponding electrical stimulations from the board to the coil and thus the implant user. First, the acoustic data is encoded using the transmission protocols of the CI device in the FPGA before streaming. The RF coil is connected to the DIET (CIC4 Decoder Implant Emulator) box (manufactured by Cochlear Corp), and this DIET box is connected to the PC using the USB and the output RF signal (comprising of electric stimuli data for all electrodes) can be recorded, decoded, and digitized and accessed on a PC. Python and C++ libraries enable data logging of output stimuli for long durations of pulses with precise timing and recording of current level of pulses. A program was generated to capture and record the various characteristics of the electric cues generated from the CCi-MOBILE RI. These characteristics are discussed in section 3
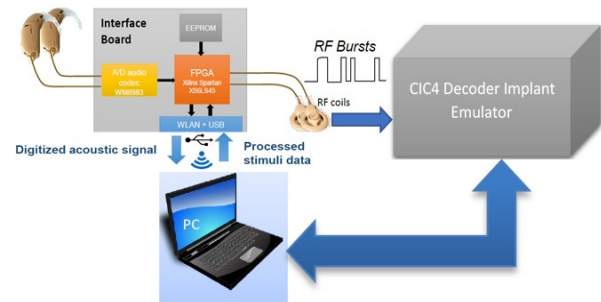

Figure 3: *Schematic functional block diagram of acoustic testing and evaluation platform*

3. Experimental results

The analysis and evaluation of the AHD (CCi-MOBILE RI used in this study) under diverse acoustical conditions, also termed as the Audio-test phase of the Burin-in process is presented in this section. A standard biphasic stimulation with a stimulation rate of 1000pps, pulse width of 25µs, and IPG of 8µs was used. The analysis of the performance of the CCi-MOBILE RI is carried out by considering these electrical stimulation factors: simulated intracochlear current (charge/sec – simulated intracochlear current experienced by the implant; pulse width errors (ΔPW – difference between designed pulse width and experimentally observed); pulse width balance error (PWBal – difference between experimentally observed anodic and cathodic pulse widths in a stimulation cycle); inter-phase gap errors (ΔIPG – difference between designed IPG and experimentally observed); timing error (ΔT – difference between designed stimulation cycle and experimentally observed);
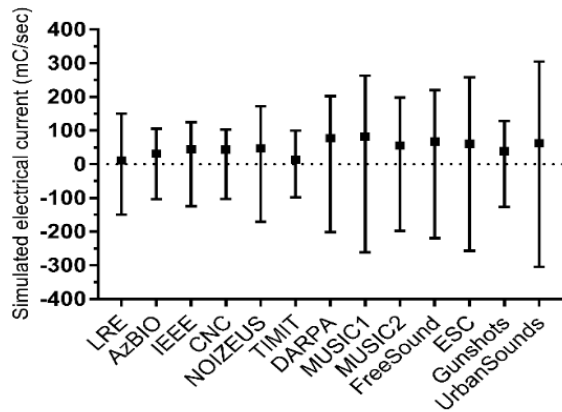
Figure 4: *Simulated intracochlear current*

3.1. Analysis of the behavior of the CCi-MOBILE RI in diverse acoustical environment

A set of 13 databases are used to simulate 260hrs of speech, 46hrs of music and 76hrs of noise. The dynamic behavior of the simulated intracochlear current (charge/sec) is captured, by continuously streaming the audio files contained in every database, and the corresponding simulated intracochlear current (charge/sec) is recorded every second. The simulated intracochlear current (charge/sec) is used to characterize the behavior of the CCi-MOBILE RI in various acoustic environments. The Fig. 4 describes the maximal variations of charge and the average of the charge across the electrodes, observed during the anodic and the cathodic phases of stimulation.

The CCi-MOBILE RI exhibited no stimulation configuration over the safety limits for all electrical components measured in this study, and therefore should be regarded as an interface that will function indefinitely within clinical safety limits. The mean and maximum value of simulated intracochlear current, is as shown in Fig. 5, and it is found to be significantly higher for music when compared to speech and these values are significantly higher for noise as compared to speech and music. The 2 speech databases: DARPA RATS and NOIZEUS, containing distorted and noisy speech data respectively, had significantly higher simulated intracochlear currents (charge/sec). A higher simulated intracochlear current affects the perceptual loudness however it does not breach safety limits.
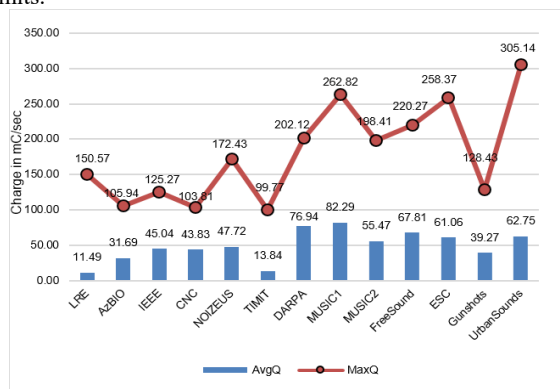

Figure 5: *Performance of charge/sec against database*

3.2. Analysis of the performance of CCi-MOBILE RI

The discrepancies observed in the electrical stimulation parameters: ΔPW, ΔIPG, PWBal, and ΔT, which remained relatively consistent across all the acoustical databases, are considered for the evaluation of the CCi-MOBILE RI. Fig. 6 shows the overall performance of the electrical stimulation parameters observed across all the databases. PWBal and ΔT have a mean value of 0.1μs whereas ΔPW and ΔIPG a mean value of 0.24μs which is relatively more. The experimentally observed PWBal, ΔPW, ΔIPG & ΔT parameters influence and contribute towards the exponential increment in perceptual loudness and can be a cause of concern for acoustical safety and perceptual sound quality. Residual non-zero PWBal contributes towards the irreversible corrosion of electrodes and the potential deposit of metal oxides at the electrode–tissue interface and hence a non-zero PWBal worsens the quality of CI stimulation and reduces the life of CI.
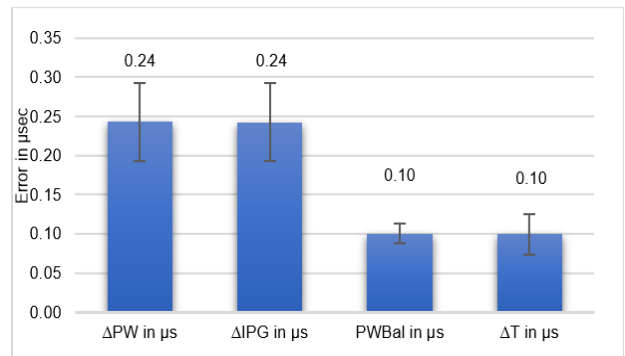

Figure 6: *Error analysis of ΔPW, ΔIPG, PWBal, and ΔT.*

4. Conclusions

A successful demonstration of testing and evaluation paradigm of the AHD under diverse acoustic conditions, also called as the Audio-test phase, was carried out by analyzing the behavior of CCi-MOBILE RI using a total of 380+ hours of acoustical data comprised of speech, music and noise. The CCi-MOBILE RI exhibited no stimulation configuration over the safety limits for all electrical components measured in this study, and therefore should be regarded as an interface that will function indefinitely within clinical safety limits. All of the stimulation parameters: simulated intracochlear current (charge/sec), ΔPW, ΔIPG, PWBal, and ΔT, were shown to effect the perceptual loudness. The behavior of the simulated intracochlear current (charge/sec) is highly dependent on the acoustical conditions and a higher simulated intracochlear current affects the perceptual loudness however it does not breach safety limits. Residual non-zero PWBal worsens the quality of CI stimulation and reduces the life of CI. This study can be used to provide customization, facilitate highly satisfying user preference and establish the conditions suitable for acoustically safer listening experience. The stimulation parameters investigated can be applied on any AHD and thus may promote the standardization, development of a safety-compliance task before the commercialization, establish a bench mark for the use and implementation of a AHD for research purposes.

5. References

[1] http://www.who.int/mediacentre/news/releases/2015/ear-care/en/

[2] Philipos Loizou, "Signal-Processing Techniques for Cochlear Implants," *IEEE Signal Processing Magazine*, vol. 15, no. 5, pp. 101–130, 1999.

[3] Ali, H., Lobo, A. P., & Loizou, P. C. (2013). "Design and evaluation of a personal digital assistant-based research platform for cochlear implants." *IEEE Transactions on Biomedical Engineering, 60(11), 3060–3073. doi:10.1109/TBME.2013. 2262712*

[4] Douglas B. Mccreery, William F. Agnew, Ted G. H. Yuen, and Leo Bullara "Charge density and charge per phase as cofactors in neural injury induced by electrical stimulation." *IEEE Transactions on Biomedical Engineering. Vol. 37. No. 10. October 1990*

[5] Ruth Y. Litovsky, Matthew J. Goupell, Alan Kan, and David M. Landsberger "Use of research interfaces for psychophysical studies with cochlear-implant users" *Trends in Hearing, Volume 21: 1–15*

[6] D.B Mcreery, W.F.Agnew, T.G.H. Yuen, L.A.Bullara "Damage in peripheral nerve from continuous electrical stimulation:Comparison of two stimulus waveforms". *Medical & Biomedical Engineering & Computing. January-1992.*

[7] Robert V. Shannon, "A Model of Safe Levels for Electrical Stimulation", *IEEE Transactions on Biomedical Engineering, Vol. 39, No. 4. April 1992.*

[8] Zeng FG1, Shannon RV., "Loudness-coding mechanisms inferred from electric stimulation of the human auditory system.", *Science 22 Apr 1994: Vol. 264, Issue 5158, pp. 564-566*

[9] Qian-Jie Fu, and Robert V. Shannon, "Effects of phase duration and electrode separation on loudness growth in cochlear implant listeners Monita Chatterjee", *The Journal of the Acoustical Society of America 107, 1637 (2000)*

[10] Pavel Prado-Guitierrez, Leonie M. Fewster, John M. Heasman, Colette M. McKay, Robert K. Shepherd, "Effect of interphase gap and pulse duration on electrically evoked potentials is correlated with auditory nerve survival", *Hearing Research 215 (2006) 47–55*

[11] Dyan Ramekers, Huib Versnel, Stefan B. Strahl, Emma M. Smeets, Sjaak F.L. Klis, And Wilko Grolman "Auditory-Nerve Responses to Varied Inter-Phase Gap and Phase Duration of the Electric Pulse Stimulus as Predictors for Neuronal Degeneration", *Journal for the Association in Research in Otolaryngology 15: 187–202 (2014).*

[12] D. B. McCreery W.F. Agnew T.G.H. Yuen L.A. Bullara, "Relationship between stimulus amplitude, stimulus frequency and neural damage during electrical stimulation of sciatic nerve of cat". *Medical & Biomedical Engineering & Computing-1995, 33*, 426-429

[13] Krzysztof Kąkol, Bożena Kostek, "A study on signal processing methods applied to hearing aids", *Signal Processing Algorithms, Architectures, Arrangements, and Applications, SPA 2016, September 21-23rd, 2016, Poznań, Poland.*

[14] Allen MC, Nikolopoulos TP, Donoghue GM "Speech intelligibility in children after cochlear implantation." *Am J Otol 19:742–746, (1998)*

[15] Hugh J. McDermott, "Music Perception with Cochlear Implants: A Review", *Trends In Amplification, Vol 8, No 2, 200*4

[16] Valeriy Shafiro,a Stanley Sheft,a Sejal Kuvadia,a and Brian Gygib, "Environmental Sound Training in Cochlear Implant Users", *Journal of Speech, Language, and Hearing Research, Vol. 58 • 509–519 • April 2015*

[17] Hussnain Ali, Ammula Sandeep, Juliana Saba, John H.L Hansen, "CCi Mobile platform for Cochlear Implant and Hearing Aid Research" – *Proc on 1st Conference on Challenges in Hearing Aid Assistive Technology, Stockholm, Sweden. August 19,2017*

[18] G. Tzanetakis and P. Cook "Musical genre classification of audio signals " in *IEEE Transactions on Audio and Speech Processing 2002.*

[19] G. Tzanetakis and P. Cook "Musical genre classification of audio signals " in *IEEE Transactions on Audio and Speech Processing 2002.*