NTNU – Trondheim
Norwegian University of
Science and Technology

# Subjective evaluation of the audio bandwidth extension program MaxxBass; Lab and web based listening tests monitoring the effect on single and mixed music tracks.

Emilia Graban

Master of Science in Electronics
Submission date: July 2015
Supervisor: Jan Tro, IET

Norwegian University of Science and Technology
Department of Electronics and Telecommunications

**CONTENT**

## ABSTRACT (in English)

The following thesis presents a subjective evaluation of the audio bandwidth extension algorithms available in the MaxxBass Program [1]. The thesis explains the principle of low frequency perception and low frequency synthesis algorithms. The listening tests are conducted using a portable devices, under laboratory conditions and through the website. The recordings consist of mixed original and processed tracks. The perception of the processed recordings were evaluated. Tests are conducted on a diverse group of people. The test results indicate factors which affect the differences in the evaluation record. Moreover, this thesis presents a functional and non – functional requirements which are necessary to prepare and carry out the testing process. The analyzed results indicate that original recording are evaluated better, although individual preference depends on the music genre and algorithm settings.

## ABSTRACT (in Polish)

Poniższa praca magisterska przedstawia subiektywna ocenę algorytmów wzmacniania niskich częstotliwości dostępnych w programie MaxxBass. Opisano proces percepcji niskich częstotliwości oraz potrzebę wzmacniania basów i sposób działania algorytmów. Przeprowadzono testy odsłuchowe w warunkach laboratoryjnych oraz za pośrednictwem strony internetowej, korzystając z urządzeń mobilnych. Nagrania przygotowano przez zmiksowanie ścieżek oryginalnych i przetworzonych. Zastosowano modyfikacje algorytmów i zbadano, w jaki sposób wpływają one na odbiór i ocenę przetworzonych nagrań. Praca zawiera opis wymagań funkcjonalnych i poza-funkcjonalnych koniecznych do przygotowania i przeprowadzenia testów. Zaprezentowane rezultaty i analiza testów wskazują, że lepiej oceniono nagrania oryginalne – chociaż indywidualna preferencja zależy w dużej mierze od gatunku muzycznego, do którego należy nagranie oraz parametrów algorytmów.

**ACKNOWLEDGEMENT**

Foremost I would like to express my sincere gratitude to my supervisors: Prof. Bożena Kostek and Prof. Jan Tro.

I would like to thank Prof. Kostek for suggested topic of My Master Thesis., which really inspired my interest in this area. I appreciate her willing to cooperate with another University and encouragement to follow on studying abroad. I would like to thank for every comment and correction. Her guidance was really important for me.

My gratitude also goes to Prof. Tro for opportunity to take part in his course, support of my studying and guidance in all the time of writing my Master Thesis. I would like to thank for every meeting after which I was motivated to further work with new ideas. I thank for his enthusiasm and confidence.

Besides my advisors, I would like to thank the consultant Piotr Hoffmann for his support, answer for every question and proposed solutions.

My sincere thanks goes to Prof. Bogdan Wiszniewski, Prof. Andrzej Czyżewski, Dr. Anna Bobkowska and Katarzyna Jezierska for allowing to study abroad, at reputable University with international students. It was a great experience both in terms of life and scientific interest.

Foremost, I would like to thank my family and my friend who helped me to process the test. I appreciate their time and entanglement.

# 1. INTRODUCTION

The following Chapter presents the aim of the M.Sc. thesis, a concept of subjective tests and some insights on how to build an audio database. The thesis starts with a description of perception of low frequencies in general and explains the need of their enhancement in portable devices. Additional phenomena affecting their perception are shortly reviewed. This Chapter reviews also the content of the M. Sc. thesis.

## *1.1. Perception of Low Frequencies*

Pitch is one of the most important features in sound. It determines whether sound perceptually seems to be low or high in the subjective scale. Sound frequency could be determined physically, however, it often depends on the personal experience. That may depend on various phenomena occurring in brain.

Sound frequency depends on the wavelength. Low frequency waves are long, so they require large-sized speakers. Currently, technology emphasizes the miniaturization of devices that reproduce sounds. In everyday life, mobile devices are used to play music, listen to audiobooks or watch movies. Unfortunately, striving for lighter and smaller devices requires a reduction in the size of the speaker, which causes distortion in bass reproducing, and also in perception. Low frequencies are cut off. In popular application it is used to increase energy in the low frequency range, however, it leads to distortion and speaker overload.

Auditory illusion happens in human brain very often. One of them is the *Missing Fundamental* – we hear the fundamental frequency which is not physically present in the signal, this is because sound has its harmonics. Both in pure and complex sound, it is possible to generate successive harmonics to create the illusion of bass [2], this phenomenon will be described in more detail in Section 2.

This auditory illusion is in the basis of synthesis of low-frequency algorithms called Virtual Bass Synthesis (VBS). They are used in many commercial programs. One of them is MaxxBass [1], which is described in the following Section.

### ***1.2. Aim of the Thesis***

The aim of this study is to determine to what extent it is possible to improve quality of recorded music playback on a mobile device. A key element of the work is to conduct listening tests, then analyze and summarize their results. The result of the thesis is the constructed audio database comprising not only the original recordings but also those improved by low frequency synthesis algorithms. Mixed soundtracks are contained in the database. Recordings differ in genres and their content.

# 2. LOW FREQUENCIES ENHANCEMENT

The following Chapter describes the principle of frequency distinguish in physical and biological way. The audio bandwidth extension is presented, its main principle and mathematical formulas. The additional parameters which affect enhancement are described – for example the characteristics of the speakers. It is answered why linear enhancement is not uses. The factors affect the perception of frequency and volume are described – especially subconscious phenomena. There are information about Missing Fundamental and lists of the systems based on these effect. At the end MaxxBass – the commercial program, which was used to process tests signals is presented.

## 2.1. *Pitch Perception*

The possibility of hearing and differentiating pitch is mainly based on the basilar membrane characteristics [3]. Figure 2.1 shows the division of the frequency ranges where the basilar membrane is responsive to stimulation by sound depending on the frequency. Vibration measurements and differences in the stimulation of the auditory nerves enable to understand the activity of the basilar membrane. It has a specific construction – it is narrower and more stiff at the basal end but wider and more flexible at the apical end. These differences allow distinguishing audio frequency differences depending on which part of the basilar membrane reacts. In response to the high frequency, the base of the basilar membrane reacts, but for the low frequencies the top does [4]. This phenomenon was discovered by Georg von Bekesy, who was awarded by Nobel Prize [5]. For the complex tone composed of several frequencies, perceived pitch does not depend on the place where the basilar membrane reacts the most. It is the value of the fundamental frequency with multiples harmonics in the test signal which has more influence on this effect [6].

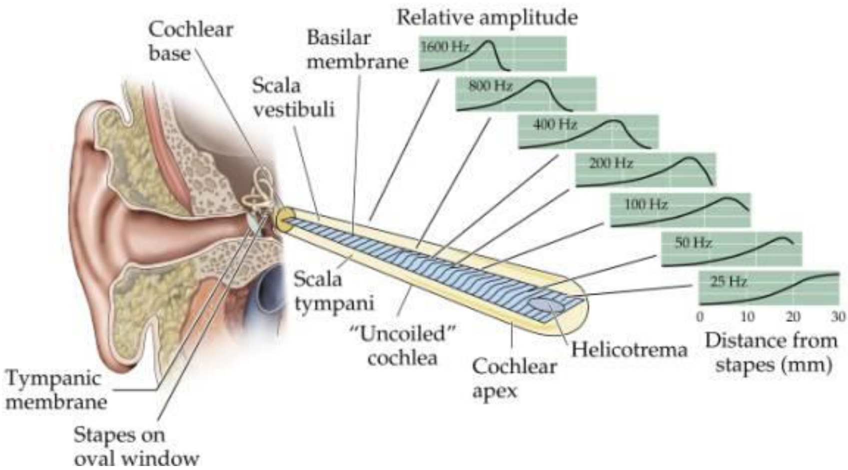


Figure 2.1. Division of the frequency ranges where the basilar membrane is responsive to stimulation by sound depending on the frequency [7]

## 2.2. *Audio Bandwidth Extension*

One of the definition given by the IEEE Standard Dictionary of Electrical and Electronics Terms says that bandwidth of a signal transmission system is: "The range of frequencies within which performance, with respect to some characteristic, falls within specific limits" [8].

Audio Bandwidth Extension is a signal processing leading to increase width of the band. The algorithm takes into account the properties of the auditory system which has a big influence in audio perception [9].

### *2.2.1. Speaker Size and Efficiency*

Bandwidth Extension methods have been developed for portable devices which have limited production charges and small dimensions. Introducing these solutions does not increase costs. The main idea is to implement signal processing algorithms in order to overcome the physical and acoustic limitations.

There are four main method of Bandwidth Extension. One of them is Psychoacoustic Bandwidth Extension for Low Frequencies. It is the psychoacoustic processing in which the low frequency band are shifted above the cutoff frequency of the speaker [9]. Each reproducing sounds device has frequency response limits which is related to a sound energy emission. Sound pressure level is plotted in the frequency as the function of speaker response amplitude. However, usually speaker is inefficient in higher and lower ranges of frequencies and there are lot of distortion. Level of these low frequencies distortion determines quality of the speaker [10].

The efficiency $\eta$ of the speaker increases with the cone area and decreases with increasing mass, what is shown in formulas below (2.1), (2.2) [10].

$$\eta \sim (\frac{s}{m})^2 \quad (2.1)$$

$$f_1 = \frac{1}{2\pi}\sqrt{\frac{k_t}{m}} \tag{2.2}$$

where, *S* is the cone area, *m* is the mass, $f_1$ is the frequency, $k_t$ is the total compliance - suspension and influence cabinet

However, to obtain low frequency $f_1$, the $k_t$ value should be small but mass should be large - which decreased the efficiency. The efficiency of the loudspeaker in low frequencies range cannot be improved by physical methods. It is required to find properties in the psychology of hearing that could improve perceived sound regardless of the above mentioned limitations [10].

### *2.2.2. Bandwidth Extension Algorithm*

The Bandwidth Extension Algorithm are presented in the graph below (see Fig. 2.2).

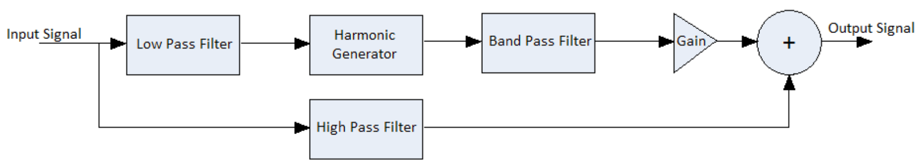


Figure 2.2. Bandwidth Extension Algorithm [11]

The output signal reproduced through the speakers is the sum of two signals. Portion of the first signal is filtered by the low-pass filter, then it is modify by nonlinear transformation or Phase Vocoder techniques where low frequencies are transferred to higher ranges [11]. Signal is filtered by band-pass filter in order to obtain proper spectrum and timbre [9] and specify the number of harmonics which will be generated. At the end signal is passed through an amplifier where is attenuated or amplified [11]. Finally, the first modified signal is mixed with the second one which is delayed [9] or had a limited power which is needed to reproduce lowest frequencies by a loudspeaker.

It is important to take into consideration the properties of the auditory system, which allow receiving mixed signal as a single stream. Three major features are distinguished: pitch, timbre and loudness [12]. The proportion of those characteristics should be preserved – disparity could be noticeable in the output signal. It is one of the limitations of perception among many other restrictions on the algorithm implementation as the memory, computational and signal format limitation [9].

To perform any non-linear signal transformation it is required to determine one dominant frequency. For this purpose, in the complex tone *zero crossing spectral representation* is calculated with formula 2.3 presented below [13].

$$\cos(\pi\gamma) = \frac{\int_0^{\pi} \cos\omega dF(\omega)}{\int_0^{\pi} dF(\omega)} \tag{2.3}$$

where, $\omega$ is the pulsation, $F$ is the continuous function.

Based on it, the dominant frequency is assigned – its zero crossing rate is the most similar to the zero crossing spectral representation for the whole signal. The nonlinear devices algorithm is able to find the dominant frequency and compute its harmonics [9].

### 2.2.3. *Traditional Methods*

There are a lot of traditional - linear methods which can enhance the bass by an amplifier. However, they provide low enhancement and high distortion. These methods require lot of energy supply. Moreover, such simple, linear bandwidth extension is not sufficient for reducing the very low, undesirable frequencies coming from the speaker [10].

One of the psychoacoustic methods which enhance bass is frequency doubling. The frequency range, where the frequency *f* is located, is shifted from $\frac{f_1}{2} \leq f \leq f_1$ to the $f_1 \leq f \leq 2f_1$. This method gives high distortion, change value of the pitch and make low frequencies louder [10].

## 2.3. *Missing Fundamental Effect*

If an signal consists of harmonic frequencies for example: $f_1$ = 300 Hz, $f_2$ = 450 Hz, $f_3$ = 600 Hz and so on, the frequency $f_0$ is equal to 150 Hz and is the fundamental one. This signal is perceived as low sound even if the fundamental frequency is filtered out. It is called the *Missing Fundamental* effect and it is presented in Fig. 2.3 .

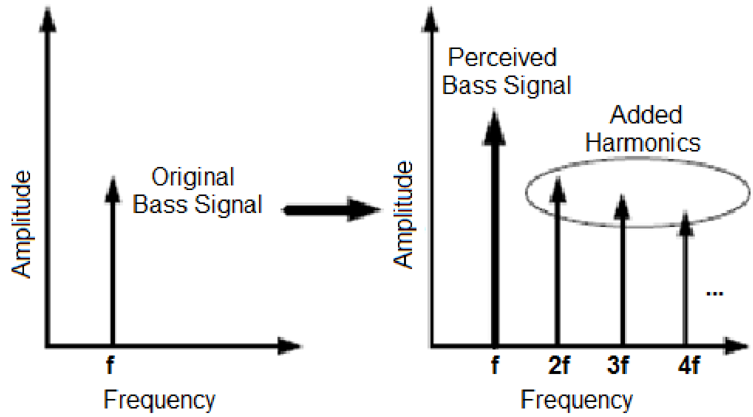


Figure 2.3. Missing Fundamental Effect [14]

The *Missing Fundamental* effect occurs in psychoacoustics and depends on the distortion of the wave propagating in the ear. This is because the human ear is a non-linear device and can cause audible distortion. This phenomenon allows recognizing which instruments are played in the song, even if the low-frequencies are cut out [6].

In order to demonstrate this phenomenon, two signals were generated in the Audacity Program [15]. The first one consists of five tones: *f1* = 100Hz 0.8; *f2* = 200Hz 0.6; *f3* = 300Hz 0.4; *f4* = 400Hz 0.2; *f5* = 500Hz 0.1; the second one consist of four tones: *f1* = 200Hz 0.6; *f2* = 300Hz 0.4; *f3* = 400Hz 0.2; *f4* = 500Hz 0.1. The comparison between these two signals makes impression of similarity. It is important to consider that the most important value in recognition pitch by a human brain is not the lowest frequency but the whole structures of harmonics. The difference between harmonics allows our brain to hear “missing fundamental”.

The M*issing Fundamental* effect is the basis of the bass enhancement techniques which will be described further.

## 2.4. *Virtual Bass Synthesis*

Virtual Bass Enhancement is a technique that is based on the phenomenon of *Missing Fundamental*. It is used in mobile devices, because it is characterized by low cost, possibility of being used in devices with limited dimensions and it provides maximum bass performance [2]. There are two main methods namely: Phase Vocoder (PV) and Nonlinear Device (NLD) [16]. Both algorithms convert an input signal by adding harmonics to achieve the illusion of low-frequency perceiving. In addition, there are other algorithms which may improve quality of signal. For example, they allow adjusting the volume of the sound which is heard subconsciously, they may decrease distortion heard, etc.

### *2.4.1. Phase Vocoder*

The first approach to virtual bass creating is the phase vocoder. It involves signal processing in the frequency domain. The signal is divided into short intervals (50-250 ms) and the Fast Fourier Transform is computed. STFT in time instant $n = s \cdot R_a$ for the s-th frame is given by following formulas (2.4) [17] :

$$X(sR_a, k) = |X(sR_a, k)| e^{j\phi(sR_a, k)} \qquad (2.4)$$

where $R_a$ determines the shift of the STFT first stage. To obtain improved spectrum $|Y(sR_s, k)|$ is required to modify the amplitude $|X(sR_a, k)|$ and phase $\phi(sR_a, k)$.

If short frames are windowed and overlapped, the new signal is created with formula presented below (see 2.5) [17]:

$$y(n) = \sum_{s=-\infty}^{\infty} f(n - sR_s) y_s(n - nR_s) \tag{2.5}$$

One of the phase vocoder advantages is to maintain constant phase during the signal processing, whereby the output signal has a better quality [16]. The disadvantage of this method is the trade-off in time and frequency distribution.

A Phase Vocoder allows full control of harmonics that are contained in the output signal. The amplitude of each harmonic is determined by equal-loudness contour (see Fig. 2.4) [18].

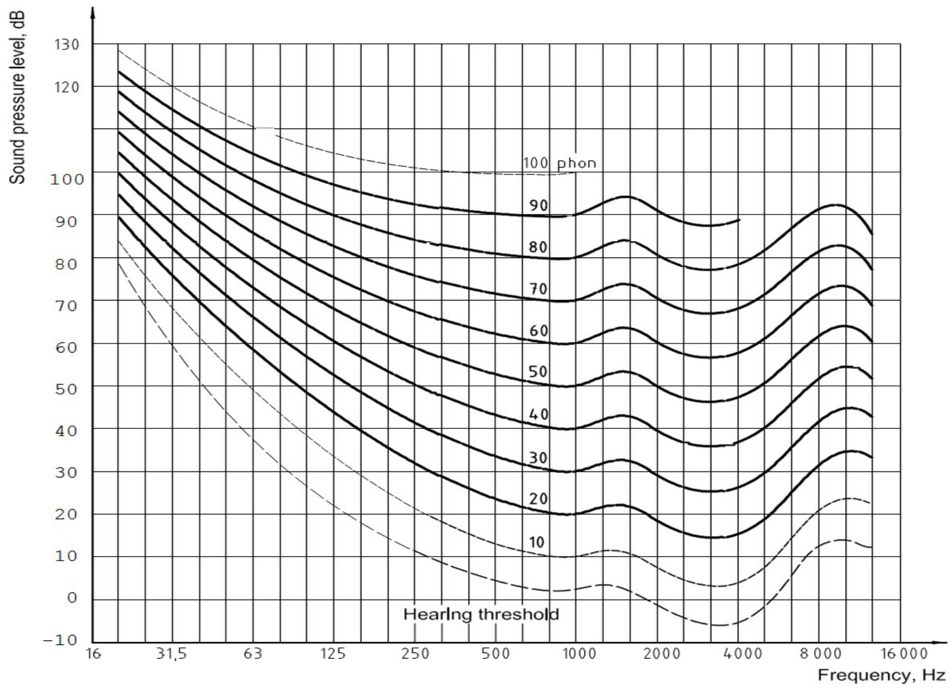


Figure 2.4. Equal-loudness contour [18]

Equal-loudness contour shows the relation between sound pressure level in the objective scale expressed in dB and the subjective scale perceived perceptually expressed in phones. Wider range of SPLs occurs in higher frequencies. In the low frequency range more energy is needed to perceive a louder sound with a higher value of sound pressure level. In Virtual Bass Enhancement techniques, amplitudes are adjusted basing on the polynomial model which is calculated with reference to the equal-loudness contour. The cost of these calculations is not high because it concerns only the low-frequency ranges. However, such computations limit the use of Phase Vocoder in the real time applications. Moreover, this technique causes a smearing effect and the signal has artifacts. Phase Vocoder is not used in applications to boost the virtual bass [16]. It is most useful in audio processing, for example pitch shifting or time stretching [19].

### *2.4.2. Nonlinear Device*

The second approach to Virtual Bass Enhancement generating is the Nonlinear Device (NLD). This technique is used for signal processing in the time domain by adding a set of harmonics to the input signal. Sinusoidal signal is transformed by the algorithm by calculating Discrete Fourier Transform. Nonlinear device decides which harmonics to generate, and how high the amplitudes will be [20]. Generally, the algorithm chooses the suitable coefficients for the selected function [16].

However, the distortion appears very often, especially for complex sound. Most virtual bass algorithms are a combination of the signal passed through the low pass filter and the signal from the NLD passed through the band-pass filter [20]. There is one more reason of distortion which raises an important issue. One of the techniques uses a full-wave rectifier (FWR), which generates only even-order harmonics. To determines harmonics spectrum a pure-tone of frequency $f_0$ are added to input signal and Fourier series $b_k$ of the full-wave-rectified signal are computed by the following formula (3.1) [10].

$$b_k = f_0 \int_0^{\frac{1}{f_0}} |\sin(2\pi f_0 t)| e^{-2\pi i k f_0 t} dt \ = \begin{cases} \frac{2}{\pi(1-k^2} & for\ even\ k \\ 0 & for\ odd\ k \end{cases} \tag{2.6}$$

For sound with the frequency of 200 Hz, the following harmonics will be introduced: 400 Hz, 800 Hz, 1200 Hz and so on, which means that the fundamental frequency perceived by brain has the value of 400 Hz instead of 200 Hz.

It is important to preserve the homogeneity in virtual bass enhancement. The ratio between amplitude of the input and output signals should be proportional. If the nonlinear device is used for automatic gain control there is a possibility to multiply - scale the signal to the expected value, and after passing through the NLD rescale to the original value [21]. Generated harmonics are multiplied by the weight of each other with $g_i$ factor. For the pure tone signal this is computed according to the formula (2.7) [10]:

$$x(t) = \sin(2\pi f_0 t) \tag{2.7}$$

Harmonics are multiplied by each other. For example, generating four harmonics is presented in Figure 2.5 [10].

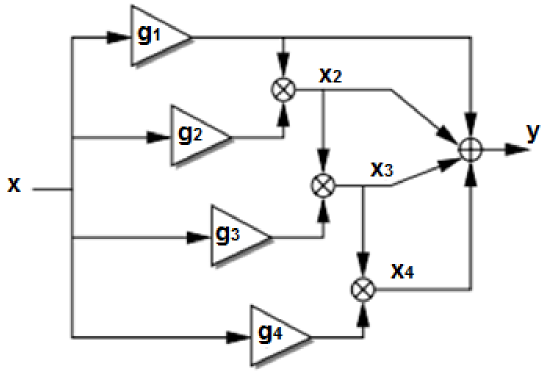


Figure 2.5. Example of generating four harmonics by multiplication [10]

The output signal is the sum of *x*, $x_2$, $x_3$ and $x_4$. For example the $x_2$ is given by (2.8)

$$x_2(t) = \frac{g_1}{g_2}[1 - \cos(2 \cdot 2\pi f_0 t)] \qquad (2.8)$$

The output signal *y(t)* is calculating according to the following formula (2.9):

$$y(t) = h_0 + \sum_{i=1}^{2}[h_{2i-1}\sin(2i-1) \times 2\pi f_0 t) + h_{2i}\cos(2i \times 2\pi f_0 t)] \qquad (2.9)$$

$h_{0\text{-}4}$ are the scale factors given by:

(2.10)

$$h_0 = \frac{g_1 g_2}{2}\left[1 + \frac{1}{4} g_3 g_4\right]$$

$$h_1 = g_1\left[1 + \frac{3}{4} g_2 g_3\right]$$

$$h_2 = -\frac{g_1 g_2}{2}[1 + g_3 g_4]$$

$$h_3 = -\frac{g_1 g_2 g_3}{4}$$

$$h_4 = -\frac{g_1 g_2 g_3 g_4}{8}$$

It is required to specify the factors $g_{1\text{-}4}$:

(2.11)

$$g_1 = h_1 + 3h_3$$

$$g_2 = -2\frac{h_2 + 4h_4}{h_1 + 3h_3}, \qquad h_1 \neq -3h_3$$

$$g_3 = \frac{2h_3}{h_2 + 4h_4}, \qquad h_2 \neq -4h_4$$

$$g_4 = -2\frac{h_4}{h_3}, \qquad h_3 \neq 0$$

The number of harmonics and their amplitudes are based on the $g_i$ values and the number of multiplications [10].

NLD is a generator which can be used in real time applications [16]. The most popular commercial software that uses the NLD is MaxxBass, which in addition, is able to adjust the amplitude.

### *2.5. MaxxBass*

MaxxBass is a commercial program to process sounds, which is mainly used for the synthesis of virtual bass. The majority of operations are based on psychoacoustic properties. MaxxBass allows increasing low frequencies without a signal attenuation. Moreover, it does not require any additional supply or a large-sized loudspeaker [2]. The input signal is processed according to the scheme from Figure 2.6.

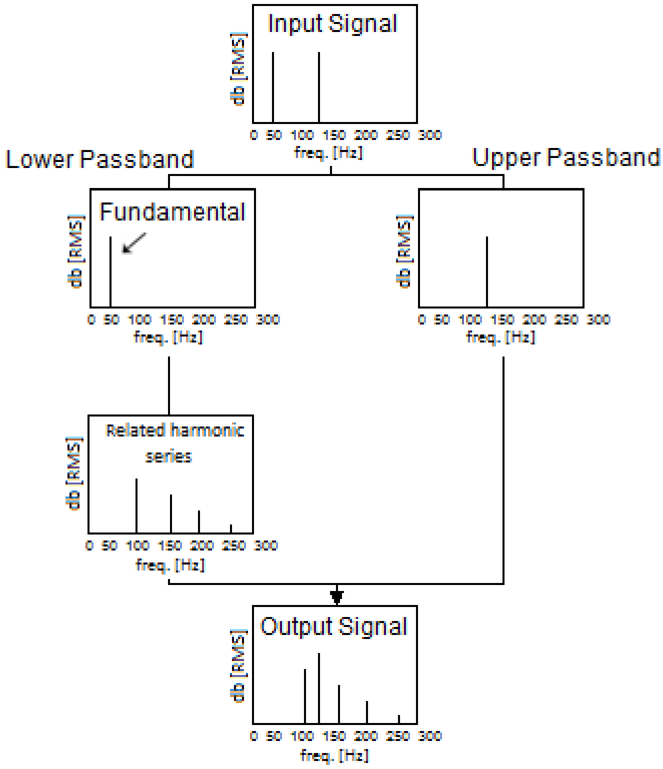


Figure 2.6. Input Signal Processing in the MaxxBass System [2]

Frequencies with the value less than specified in the low-pass filter (depending on the physical characteristics of the speaker) generate their harmonics according to a predetermined algorithm, and then they are converted to the signal with the signal components at frequencies higher than the limit-pass filter. Thus, psychoacoustic information is based on the whole spectrum of the output signal.

In order to achieve a better quality of signal, the amplitude of the harmonics is calculated based on the special formulas and loudness principles [17]. The dynamics of the resulting signal is adjusted according to the equal-loudness contour (described in Section 2.4.1).

Decreasing the sound pressure level of the input signal for *x* dB causes a decrease of energy in the output signal for $R \cdot x$ phones, where SPL to phone ratio is expressed by the formula (2.11) [17].

$$R(f) = \frac{1.0}{\ln(f) \cdot 0.241 - 0.579} \tag{2.11}$$

The remainder of the ratio SPL and phones for *n*-th harmonic are expressed by the following formula (2.12) [17].

$$RR(f, n) = 1 - 1\ln(n) \cdot 0.241 \cdot R(f) \tag{2.12}$$

Energy for the *n*-th harmonic is given by (2.13) [17]:

$$E_h(f, n) = RR(f, n) \cdot E_f + \rho \quad [dB] \tag{2.13}$$

where $E_f$ is energy of the fundamental frequency, $E_h$ energy of *n*-th harmonic and $\rho$ is an equalization constant.

The sound pressure level is enhanced so that low frequencies preserve the desired value in subjective scale expressed in phones. Processing based only on equal-loudness contour may cause sound appearing unnatural thus calculations are based on a logarithmic scale. For the lower frequencies the sound pressure level a have narrower range.

The MaxxBass system allows increasing low frequencies up to one and a half octave [22]. All the frequencies are higher than the limit of the filter and can be heard from the speakers. Therefore, the system is considered to be very effective.

# 3. AUDIO DATABASE

The following chapter contains information about the database with audio recordings created to conducting a subjective testing of low-frequency synthesis algorithms applied to portable devices. There is a description of the processed music genres, instruments and songs source. The way of amplifying the basses and parameters used In MaxxBass are presented. There are description of the final. At the end there are the information about the technical parameters of database

## *3.1. Original Sound Sources*

The created database is a collection of test signals processed and structured in specific way. There are five songs which differ in music genres. All of them were taken from the *'Mixing Secrets'* Free Multitrack Download Library [23]. This website provides both mixed and separate audio tracks. Recorded pieces of music are available in two options: Full Multitrack and Edited Excerpt. They are sorted according to music genres. Additionally, there is the discussion zone which allows users to comment and share the samples of their processed recordings.

## *3.2. Music Genres*

Each track differs in music genre. The following ones were selected: Heavy Metal, Indie, Pop, Reggae and Rock. All of them have different content of low frequencies. It allows recognizing a difference in the results obtained for Virtual Bass Synthesis algorithms.

a) Heavy metal is characterized by overridden electric guitars, intensive rhythm, loud vocal and clear bass [24]. The most important element is the amplified and distorted electric guitar, which interferes with the singer. Heavy metal produces a lot of energy in specific frequency ranges in discontinuous time [24]. There are also other instruments determining the heaviness of the music - bass and drums add a lot of power but the whole percussion enhances timbre and volume. In the created database there is *Dark Ride – “Dead Enemies”* song including original tracks mixed with processed, separated audio tracks of four instruments: Electric Guitar, Bass Guitar, Drums and Kick.

b) Pop is music genre characterized by eclectic line, with elements from other styles. Originally called popular. In this work the *Patrick Talbot - "Blue"* is used along with processed the following instruments: Drums, Bass Guitar, Electric Guitar and Vibraphone.

c) Indie is a genre similar to Indie rock and Pop focusing on the music melodic line. The study is based on *Ben Carrigan – “Hey Carrie Anne”*. The Song is mixed with processed tracks of Timpani, Bass Guitar and Strings.

d) Reggae music is created from a combination of drums in the high tones of electric instruments and bass guitars. It has a distinctive rhythmic pattern connecting the bass line and drum bit. The leisurely tempo, steady and pulsating rhythm and extended bass line make reggae songs relaxing and hypnotic [25]. The song entitled *Arise – “Run Run Run”* is used to create a mix with original tracks and processed vocal track, Bass, Hi Hat and Organs.

e) Rock is music genre characterized by various types of guitars and drums, with clearly delineates rhythm and singing. There is a lot of improvisation visible within a song. The most important feature is the collectivity during the music making. In the database *Angles in Amplifiers – “I'm all right”* is used with processed separated tracks: vocal, piano, bass and acoustic guitar.

### ***3.3. Separated Audio Tracks***

In the created database Virtual Bass Enhancement algorithms were applied to the to the separated tracks with instruments and vocals. Selected paths are different from each other in the low frequency content. It is often possible to determine the amount of bass included in the audio by listening to the recording, but for more objective analysis there is a possibility to plot the time-frequency spectrogram. It shows the sound pressure level distribution as a function of frequency changes in time. Such a plot shows which frequencies are dominant. Below, there are two spectrograms plotted using the Audacity Program [15]. The first one, shown in Figure 3.1, is of piano audio track and the second one, presented in Figure 3.2, of the bass guitar from the same song.

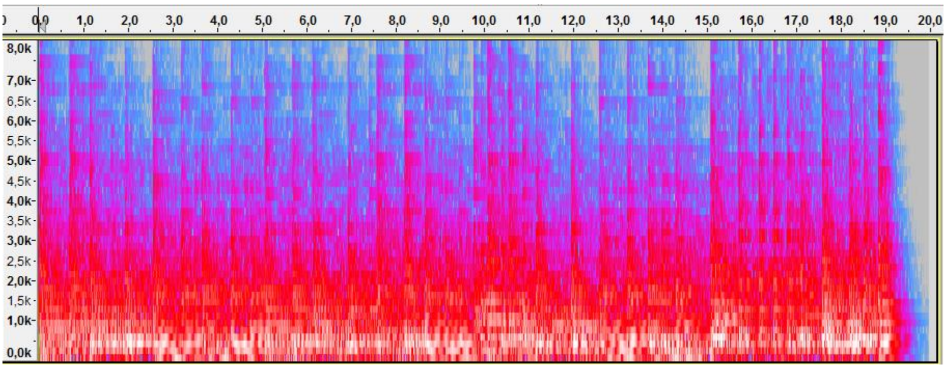

Figure 3.1. Spectrogram for audio track with the piano

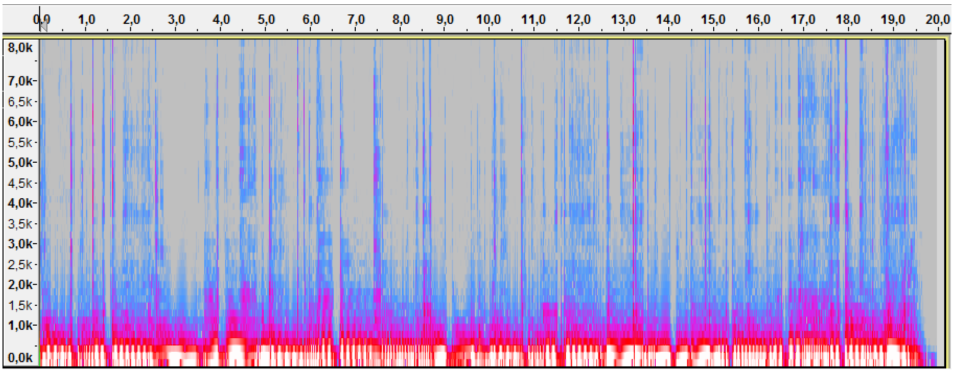

Figure 3.2. Spectrogram for audio track with the bass guitar

The vertical axis shows frequency expressed in Hertz and on the horizontal one time expressed in seconds. The color indicates the intensity. These graphs allow noticing that piano signal intensity has higher values in a wider frequency range, but intensity for the bass signal focuses on a very narrow range of a low value. It is possible to hear the difference by comparing these two instruments.

In the created database audio tracks with the following instruments were used:

a) The drum is an instrument constructed of shells and a stretched membrane belongs to a drums family. The drums differ significantly in frequency range. The drums differ from each other in the membrane tension, shape and the size – dimension is from 18 to 24 inches [26]. The extracted sound depends on the shape and a skin type.

b) The timpani is a percussion instrument from the group of impacted membranophones [27], the family of timpani. It consists of a membrane stretched on a bowl-shaped base. The sounds are generated after hitting the top of the instrument. Bases are equipped with a pedal which controls the tension of the membrane and adjusts the frequency. It is the first solution which allows changing the pitch when player's hands are busy with hitting on the membrane [28].

c) The kick - bass drum is the largest drum in the whole drum kit, characterized by the strong construction and low sound. It is from the family of drum. For the high frequency attenuation and the low frequency emphasis there is a rolled material or a small pillow inside the drum. It generates powerful low pitches of an indefinite frequency [29].

d) The hi-hat is a part of a drum kit belonging to a group of instruments with an indefinite pitch [27] from the plate family. It is composed of two plates mounted on a stand. The movement of the plates is generated by pressing on the pedal. The pitch changing is possible by adjusting the distance between the plates.

e) The vibraphone is an instrument belonging to the group of idiophones from the family of metal [30]. The bars are connected by tubes which act as resonators and generate vibrating sounds. The vibration is added due to rotating blades inside the tube, on and off by an electric monitor. The most common vibraphone operates in three octaves, starting from the F. Less popular are four octave instruments which scale starts from *C* below the middle *C*.

f) The piano is included in a family of stringed impacted instruments [31]. The sound frequency of the piano is in the range from 30 Hz to 4000 Hz. A vibrating string has one fundamental frequency and partial series. If the complex sound contains the fundamental frequency and frequency with its doubled value, it is expected to get the most pure tone [29].

g) The organ is a set of many pipes which sounds one note with a fixed pitch, intensity and timbre. It belongs to family of brass. The pipes are supplied from the common air reservoir [32]. The range of the large scale organ is ten octaves from 16 Hz to 20 kHz, which is comparable to the range of sounds perceived by a human ear.

h) The guitar is a chordophone string instrument [27] which has from 4 to 18 strings. The cover is made of wood, strings are made from of the gut, nylon or steel. It has a characteristic structure. The strings differ in mass, length and tension whereby it is possible to generate the specific frequencies and their harmonics. The following types of guitar are distinguished:

- Acoustic - strings are activating resonances in the top plate, bottom plate and the cavity, causing sound radiation.
- Electric guitar has an microphone which reacts to movements of the metal strings and generate an electrical signal, which is amplified or send to the set of the speakers [33].
- Bass is the largest instrument that produces the lowest frequency from the viol family [20]. Basses have four strings E1, A1, D, G [34] which make sounds respectively with the frequency:  41.2 Hz, 55 Hz, 73.4 Hz and 98 Hz.

i) Voice is the instrument generated in speech mechanism. For comparison the frequency of the adult male voice is in the range from 85 Hz to 180 Hz, woman's from165 Hz to 255 Hz.

## *3.4. Music Signal Processing*

Experiments consisted in several stages. The first one was performed to pre-process separate audio tracks. Every test signal has been adjusted to the duration of 10 seconds - for this purpose the free version of Adobe Audition [35] was used. In addition, this program allows using more plug-ins, for example free version of the MaxxBass algorithm [1], of which the user's interface is presented in Figure 3.3.


Figure 3.3. The user's interface of MaxxBass

### 3.4.1. *The MaxxBass Features*

The MaxxBass technical specification has been described in Section 2.5. The prepared user's interface is very easy to use and allows changing the number of parameters which affect the output signal.

On the screen presented in Figure 3.3 there is chart showing the content of the original and enhanced bass. The vertical axis represents the sound pressure level expressed in dB, and the horizontal axis the frequency expressed in Hz.

Sliders on the right side allow increasing or decreasing the content of the original and enhanced bass. For test signal played through small speakers, the MaxxBass documentation [22] recommends to mute the original bass and replace it with reinforced one.

The slider below the graph allows shifting the frequency in which the signal is amplified. For this value, the original bass is removed but enhanced bass is imposed. This change affects the smoothness and the impression of full bass. If a test signal contains lots of low frequencies, the slider should be set to a lower value.

To increase the amplification of the output signal it is possible to modify the value of the decay - it changes the generated harmonics. For the higher value of harmonics the probability that the sound will be weakened is increased. Too low value can cause the sound not to be received for small speakers. With the varying decay, the right edge of the enhanced harmonic is moving. A suitable adjustment of this value depends strongly on the type of sound. To prevent clipping, it is recommended to use a low value of decay for the input signals which have high frequency.

It is possible to adjust the slope of imposed harmonics. The high-pass filter is applied on the signal. There are three possibilities: the first one is 6 dB per octave - it removes very low frequencies, the second one is 12 dB per octave - used in most situations, and the third one is 24 dB per octave used for the sounds reproducing through small speakers because it decreases the possibility of pushing the transducer too far [22]. The results of applying the sequential filters are shown in Figures 3.4 a-c.

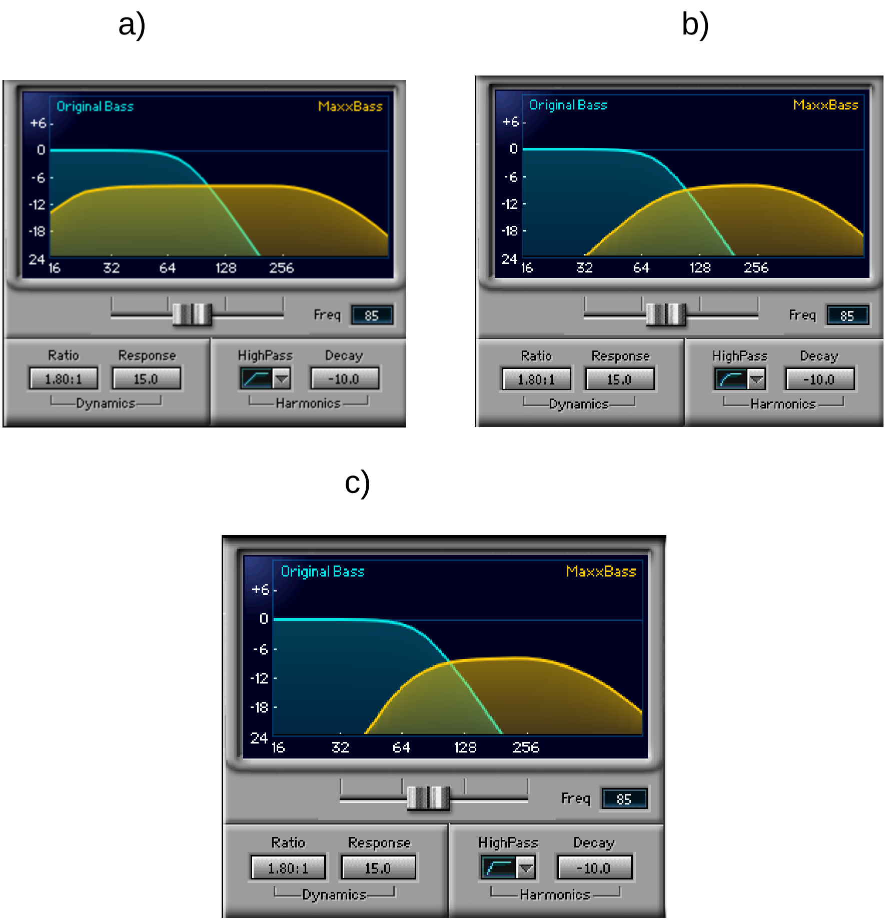


Figure 3.4. The differences in the overlap the of original and enhanced signals depends on the high-pass filter slope:

a) 6 dB, b) 12 dB, c) 24 dB

In addition, MaxxBass allows changing the dynamics of the sound by setting the Ratio and Response parameters. It controls the compression of generated harmonics. More frequent sampling increases the harmonic level but decreases the dynamic.

The Response value is expressed in milliseconds and determines the response time of harmonics. It is recommended to set a higher value of the response for short sounds such as kick but a lower value for longer one for example jazz bass [22].

During the adjusting features it is possible to listen to the mixed output signal and separated paths – the enhanced and original bass in the Output Section. It enables controlling the overloaded recording and noticing whether the signal is processed well enough to hear the improvement.

MaxxBass offers 16 designed settings for different types of applications. According to the MaxxBass documentation [22] four of them are suitable for processing music that will be played through small speakers.

### 3.4.2. *The Used Parameters Settings*

For the purpose of conducting subjective testing of low-frequency synthesis algorithms three settings have been used. One of them is called "Multimedia" and it is originally designed and available in MaxxBass. For systematization purpose it is called "Setup A" in the rest of the text. Its characteristic is shown in Figure 3.5.


Figure 3.5. The user's interface of the MaxxBass with „Setup A"

The graph above shows that the original bass is completely filtered and has an infinite value. The Maxx Bass is boosted to the value of 0 dB. The frequency is set to an average value of 120 Hz. The slope of the filter is 24 dB per octave and as is mentioned in Section 3.4.1 it is recommended for sounds that will be reproduced through the small speakers. The value of decay is – 15 dB on average, if the range of the decay is within -9 and -24. The input sound is heard and  it is not attenuated.

Properties of "Setup B" are shown in Figure 3.6.


Figure 3.6. The user’s interface of the MaxxBass with „Setup B”

In this case, the starting setup was "Lo-fi system enhancer" and the variables were subsequently adjusted, based on the subjective impression and purpose of each features. The frequency value was set very high – to 220 Hz, the original bass below these thresholds was removed and the enhanced bass was imposed. There is also a very high value for the MaxxBass harmonics, it is 10 dB. The filter slope value is 12 dB per octave which is the most common setting. The response of 20 results that time for generation harmonics is respectively high so it is expected to archive a better result for short sounds. The differences in the output signal are very distinctive.

The last setting, i. e. “Setup C" is shown in Figure 3.7


Figure 3.7. The user’s interface of the MaxxBass with „Setup C”

Figure 3.7 shows that the original bass signal was replaced completely by the enhanced one. Generated harmonics in the output signal are set to 5.3 dB. The frequency has the average value of 120 Hz, but the decay -12 is a little bit higher than in the previous cases. Changes have been applied to the designed settings of "BoomBaxx" dedicated to small speakers.

The differences between the settings are collected in the following Table 3.1. The meaning and the average values for each feature are thoroughly described in the Section 3.4.1.

Table 3.1. The values of parameters affecting the algorithms set in the MaxxBass

| Setup | A | B | C |
|---|---|---|---|
| Original bass [dB] | Filtered | 0 | Filtered |
| Maxx bass [dB] | 0 | 10 | 5.3 |
| Frequency [Hz] | 120 | 220 | 120 |
| Slope of the filter [dB/octave] | 24 | 12 | 6 |
| Decay | -15 | -15 | -12 |
| Response | 15 | 20 | 20 |

## 3.5. *Audio Track Mixing*

The reference recording has been created by mixing all of the original tracks. In addition processed tracks have been mixed with the rest of the raw (unprocessed) tracks. For example, the song Angles in Amplifiers – “I'm all right” consists of 12 tracks: Kick, Snare, Overhead, Toms, Percussion, Bass, Piano, Electric Guitar, Acoustic Guitar and three recordings of Voices. The low frequencies have been amplified (in three different ways described in Section 3.4) in the track with Piano, Bass, Acoustic Guitar and one of the voices. Finally, they have been mixed with the rest of unprocessed recordings. As a result, four excerpts have been created for each of five songs.

0.5 seconds of fed-in and fed-out have been added to every recording. It provides a linear increase of volume in sound at the beginning and at the end  of the except of the song in order to improve quality of listening.

Volume level has been adjusted to -16 LUFS for every song. LUFS stands for Loudness Units Full Scale and it is used by the European Broadcast Union (EBU). 1 LUFS is equal to 1 dB. This is an European equivalent unit for LKFS (K-weighted Loudness Full Scale). K-weighted curve shown in Figure 3.8 determines the loudness level [36].

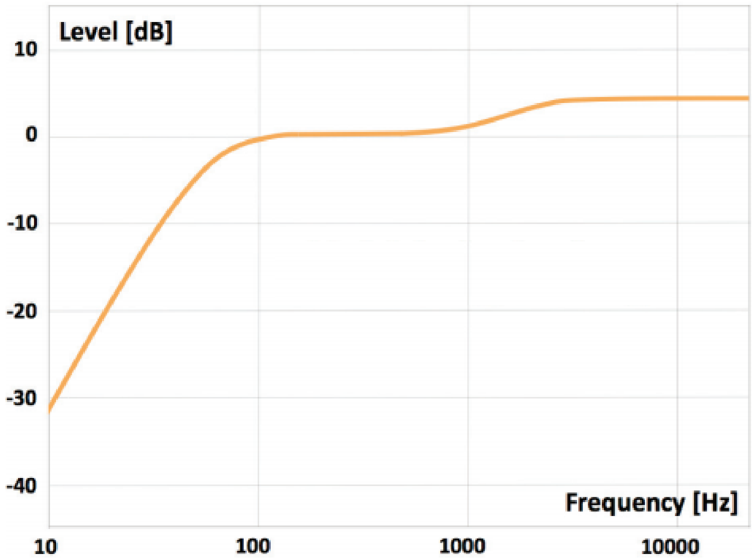


Figure 3.8. K-weighted curve [36]

The K-weighted curve was designed based on the results of listening tests. For lower frequencies, the sound pressure level expressed on an objective scale of decibels is perceived as more quiet. To achieve an equal loudness, low frequencies have higher value of sound pressure level. Designed filters have been standardized by the International Telecommunication Union.

## 3.6. *Technical Parameters of Sounds*

All of these test signals are saved as a mono-channel, in the Audio waveform format 24-bit Integer with a sampling rate of 44100 Hz. As mentioned before, they are of 10 seconds duration.

## 3.7. *Structure and Size of Database*

For the purpose of listening tests five full songs and 17 separated audio tracks have been processed. Resulted from that there are 22 original recordings and 66 processed recordings which makes a total of 88 tracks. The whole database has 253 MB and the sum of the recording duration time is 29 minutes and 38 seconds.

## 3.8. *Metadata and Legal Aspects*

In the metadata layer all the recordings are marked by the information about music genre and the number of the setting. For example Reggae 1.wav. All test signals have legal situation which allowed processing and presenting them to the public.

# 4. TEST PREPARATION AND PROCEDURE

The following Section presents the reason and aim of tests proposed. The conditions and tests scenario, both in the laboratory and via the Internet, are described. It summarizes test conditions, number of people taking part in listening tests, also the listening environment and loudspeaker types are presented.

## *4.1. Test Hypothesis*

The reason for performing subjective tests is the ability to collect subjective opinion of people whose impressions depend on subconscious feelings. The collected and analyzed results provide some useful information for the new systems, devices or algorithm designers. The developers take care of customers' satisfaction. Listening tests which answer the question "which sounds do you prefer?" allow improving created systems and avoiding disappointment after work.

The main aim of the tests was to check whether the Virtual Bass Enhancement algorithms presented in various combinations described in Section 3.4 change the recordings and whether they improve them. In addition, the age, gender and type of speakers were taken into consideration. It was analyzed if the test results depend on music genre and test environmental.

## 4.2. Methodology

### 4.2.1. Types of methodology

There are many types of listening tests. Each of them is designed for a different purpose. The right choice depends not only on the evaluating test signal features but also on the time spent on testing. The results form is taken into consideration. It is requisite to analyze the requirements and the purpose of the test. [37] The most common tests are: ranking, category judgement, semantic differential exploration of associated imagination on sound perception and pair comparison.

### 4.2.2. Choosen Methodology

Taking into consideration a set of recordings and the expected format of the results the test of VBS algorithms was performed based on the Pair Comparison test.

The Pair Comparison test consists of a pair excerpts where one is the reference. Tracks are selected based on preset criteria, for example which one is louder or better on a subjective scale. The test offer a choice of 2 or 3 - point scale: better / worse and better / the same / worse. The disadvantage of a three-evaluation approach is that it may be difficult to analyze and draw a conclusion if most of the answers are "the same". On the other hand, if a tester has to choose only between better / worse, and does not perceive the difference, he must choose even if he is not convinced. This test is best suited to assess very similar recordings, because every time a reference recording is listened to. Short-term memory allows memorizing the sound level and distinguished between sound excerpts. The disadvantage of the test is the fact that it is getting more fatiguing with each pair of listening excerpts.

During this kind of test, it is important to repeat recordings to investigate the stability of the listener's decisions and increase the objectivity of the results. Another important factor is the random order in setting recordings in pairs. The reference recording should not always be presented as the first or the second one . The length of each recording should be between 8 and 15 seconds and the interval between them 2 seconds. The break after one pair should last 5 seconds. If the test consists of more than 15 pairs, there should be 3 - 4 minute break after them [38].

The purpose of the study is to determine whether the recording processed by VBS algorithms is evaluated to be better than the original one and on what scale it was decisive. The preferred excerpt (including the level of convenience on the scale from 1 to 3) was chosen from each pair. If two recordings seemed similar the "+" has been chosen and if one fragment was definitely evaluated higher the "+++" has been clicked. This scale allows examining the listener's preferences in detail - thorough analyzing and making statistics results. The test often contains different pairs which are very similar – this kind of test enables immediate evaluation - when the recording is still in short-term memory - the decision is more objective.

## 4.3. *Test Procedure*

The test was divided into six parts. The first one includes some general questions: Name, Age, Gender and the Type of speakers. The answers to these questions allow drawing a further conclusion. Particularly, in the evaluation of low frequency synthesis algorithms the age, gender, type and size of the speaker could have an impact on preferences. Subsequent part of the test (2-6) allows comparing a song of a different kind of music. There are four pairs on every page - three of them contain a reference recording along with processed one in various ways. The last pair is a repetition of one from the above pairs - to enhance the objectivity of the test and evaluate stability of the listener’s answers. For each pair the reference recording is in a different position – to avoid the tester’s biased opinion.

The test has been carried out for 30 persons. 9 of them have done it in the laboratory conditions - it is not a large group, but sufficient to draw conclusions. The rest of the tests have been conducted using a website described in Chapter 5.

### *4.3.1. The Test in the Laboratory*

#### *a) The Laboratory*

In order to provide the best acoustic environment the test was performed in a laboratory available in the Multimedia Systems Department at the Gdansk University of Technology.

b) The Listeners Characteristic

Nine people took part in the test. Some of them deal with the processing of sounds as a job, for the rest of them listening to music is only the way of entertainment.

c) The Device Parameters

In order to provide the best quality of sound Toshiba NB550 PCs were used. There are built-in Harman Kardon speakers with the size of 25.7 mm (10.1 '). It ensured sound quality at the highest level combining Dolby @ Advanced Audio technology. The power in every one is 2 Watts. The maximum volume is approximately 90vdB. Additionally, the computer is equipped with processor AMD C-50.

d) Test scenario

- Meeting in Laboratory
- Brief information about VBS algorithms and their modifications
- Manual test – the tests length, intervals between excerpts in pairs and pages, making the right choice
- Presenting the Website on the Computer
- Testing under my supervision - the tester uses the website himself.
- Transfer the data

Each test has been carried out individually. All the people heard the processed songs for the first time. The best conditions for the environment and equipment have been provided.

### ***4.3.2.*** *Website*

In the test - via the website - 27 people took part. They were mostly family members, friends and their friends. Most of the results were complete, but some of them contain answers only to the main form and had to be discarded.

# 5. WEBSITE

The following Section contains information about the created website containing the tests prepared. There are the advantages and disadvantages of such a form. Functional requirements and the scenario are described. The subpages and transitions between them are presented. The description of functional requirements related to the design of the website is provided,

## 5.1. *Chosen methodology*

In order to carry out the tests a website [39] has been created. The choice of such a form is due to the ability to collecting number of results - anyone who has an access to the Internet and uses the computer with speakers or headphones is able to perform the test. The second advantage of this form is the simplicity of collecting the results - all the answers are automatically sent to an e-mail, there is no chance to lose them. Another advantage is the ability to prepare the logical structure and user-friendly interface using HTML, as well as sending e-mails with the results using PHP. The disadvantage is that there is no supervision during the tests. Moreover, the environmental conditions might be insufficient. Recordings which are played on different devices sound differently so the assessment might be unequal for all the results. In addition, without any control during the test, there is no certainty that it has been carried out conscientiously and all the answers are subjective. To avoid this problem, several tests have been performed in the laboratory conditions under supervision of the author of this thesis. The investigation of the stability in the testers decision by double rating  the same pair might improve the objectivity of the tests.

## 5.2. *Requirement Specification*

Requirements for creating a website can be divided according to the conditions of the test into two groups:

a) In the supervised laboratory conditions

- Intuitive interface,
- Verification of transfer empty answers,
- The opportunity to return to the home page,
- Ergonomic.

b) Shared as a link to the website:

Satisfy all of the above and the following:

- Provide manual,
- Provide brief information about VBS algorithms,
- Preliminary questionnaire – tester doesn't feel anonymous,
- Questions about the speaker and their characteristics,
- Friendly interface - an incentive to solve the test.

## 5.3. Requirements Description

### 5.3.1. Functional Requirements

The website consists of two parts. The first one with the general information is divided into two web pages: The Home page with information about the test and the manual and the page with the description of VBS algorithms. The second part of the website includes a questionnaire and a listening test form. Both of them have been created taking into consideration their purpose. Figure 5.1 presented below shows all pages and available transitions between them. From the home page, which opens as the first one, users are able to go to page with VBS algorithms description or to the test form. From each subsequent page there is possibility to return to the main page or follow-up the test.

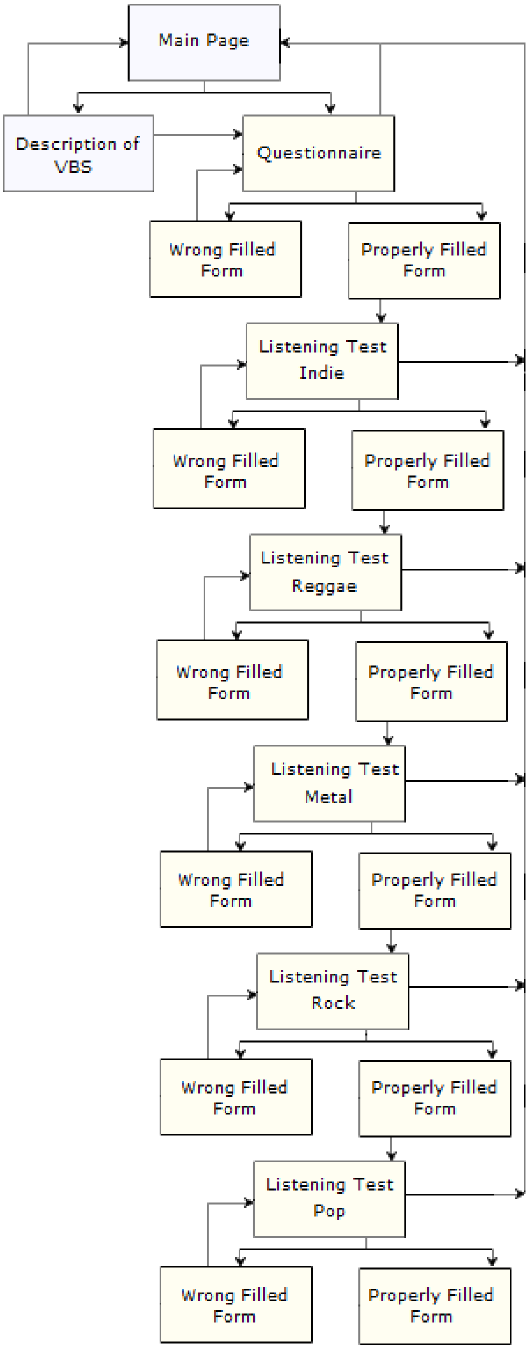


Figure 5.1. Web pages and transitions between complete or incorrectly filled form

a) The Information Section

The information part contains instructions to perform the test. There are answers to the following questions: how to prepare for the test, how long it takes, how to select answers properly and continue the test and so on. In addition, there is some information about test pages contents and their division. In the information section there is description of the Virtual Bass Enhancement algorithm, its purpose and use. There is brief information about the missing fundamental effect which is based on low frequencies enhancement. In addition there are two pictures which help to understand the ideas.

b) The Tests Section

At the beginning there is a general questionnaire, which contains questions about the name, gender, age, range and a type of speakers. The user is not anonymous which increases the truthfulness in the test. This information allows drawing additional conclusions. For example, whether age or gender affect noticing major differences between the recordings. There is *Continue* button and after press on it, the answers are checked - whether all fields are filled in or not. If there are any missing gaps, the information is shown and a user is asked to return and complete the previous page. A correctly filled in form allows moving to the main part - listening tests. They consist of five pages. Visually all of them are very similar, the main difference is only in the title and icon characteristic for a particular kind of music. However, they contain different recordings. Each of five web pages presents a different kind of music. On one web page there are four pairs of recordings. Each pair contains an original and processed recording presented in random order. Below every recording there are three buttons labeled with the amount and color of "+". It represents the certainty with which one recording of the pair is appreciated more. A tester, by choosing a recording that he like more, must identify his convenience on the scale. If the difference is small he indicates green "+"; if it is clear he clicks on red "+++". After clicking on the "*Continue" button it is* checked whether all pairs were assessed - if not, tester is asked to return and complete the form. Afterwards a tester is redirected to the next website. After the last response transfer, an information on completion the test and gratitude. appear The data contained in the questionnaire and during listening tests are sent to an e-mail. There is a possibility to return to the main page and start the test again. It is not necessary to fill in the form again after returning to the previous page.

### *5.3.2. Non functional requirements*

Communication and testers understanding are extremely important especially during the test without supervision. To achieve an objective test results it is important to present the test in a simple and intuitive way. The user's interface design is presented in Section 6.1. A manual at the beginning and an explanation during the test also in case of errors or incorrectly filled in form are provided. The interface is created to avoid the tester distraction. There are simple colors and clear letters. The website contains only the necessary information. The elements which improve the attractiveness of websites are changing graphics depending on the type of music on every page with a listening test. To fulfill ergonomics requirements the recordings are in the cumulative point and the button *Continue* is located directly under them. The size of the buttons and the subtitles are in a special size adapting to the screen size, to increase the simplicity of pressing.

### *5.3.3. System Requirements*

The created website works on Google Chrome browser with the Internet connection. It is required to have headphones or speakers (built-in or external). The Website is compatible with every screen size. There is only English version available.

### *5.3.4. Stability of answers*

In order to study the stability of answers on every page the fourth pair of sounds to be evaluated was added. It is a repetition of one of the previously pair (original /processing by Setup A; original /processing by Setup B; original /processing by Setup C). In order to collect an equal number of repetitions for a pair processed by each setting, after receiving 10 results the pair was changed. After collecting 30 results, changes followed after receiving 2 answers.

# 6. WEBSITE DESIGN AND IMPLEMENTATION

This Chapter presents the website design and implementation. It is explain how the graphic affects the testing. The functionality of each button is described. The most important source code is presented. There is a list of technologies and tools used to create the web pages. It is shown how the data have been sent to the server and how answers were received by an e-mail.

## *6.1. User interface design*

One of the most important aspects when creating the website is its appearance. The first impression is very important and should not discourage or cause embarrassment. For listening tests it is particularly important to stay focused on the music. All further stimuli from the world should be eliminated. The website should allow easy movements and returns between screens. Below the screenshots of each page are presented.

a) Main Page

Below, in Figure 6.1. there is the Home Page screen. The header contains the name of the test – it is possible to move to the website with a description of Virtual Bass Enhancement algorithms. There is some information about the test and the instructions. At the bottom there is an icon with an arrow and the words "let's try" - it is the link to the next page - questionnaire.

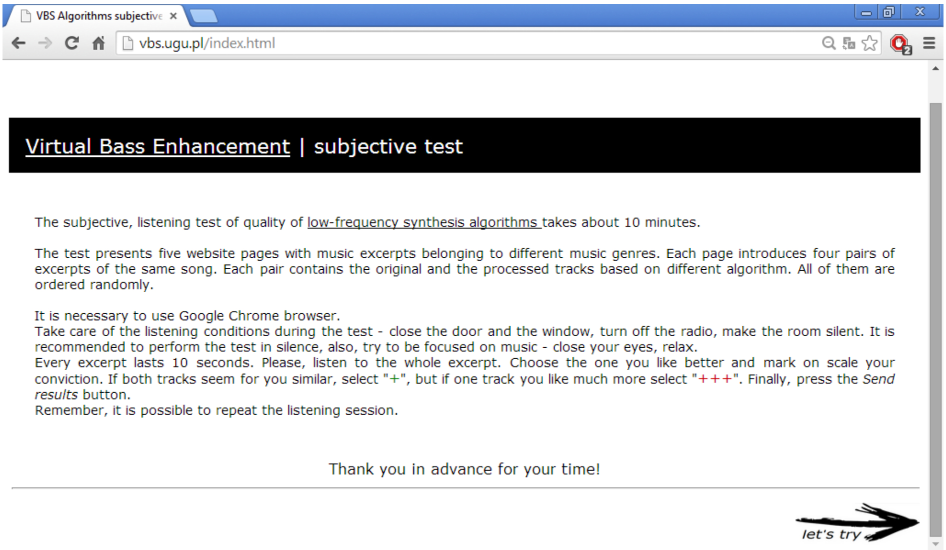


Figure 6.1. The Main Page Interface

b) VBS algorithms description

In Figure 6.2 there is a screenshot with the VBS algorithms description. There are information and images which allow a better understanding of the principle. It is possible to return to the home page using the icon in the upper right corner or go to the questionnaire by clicking on the button with an arrow.

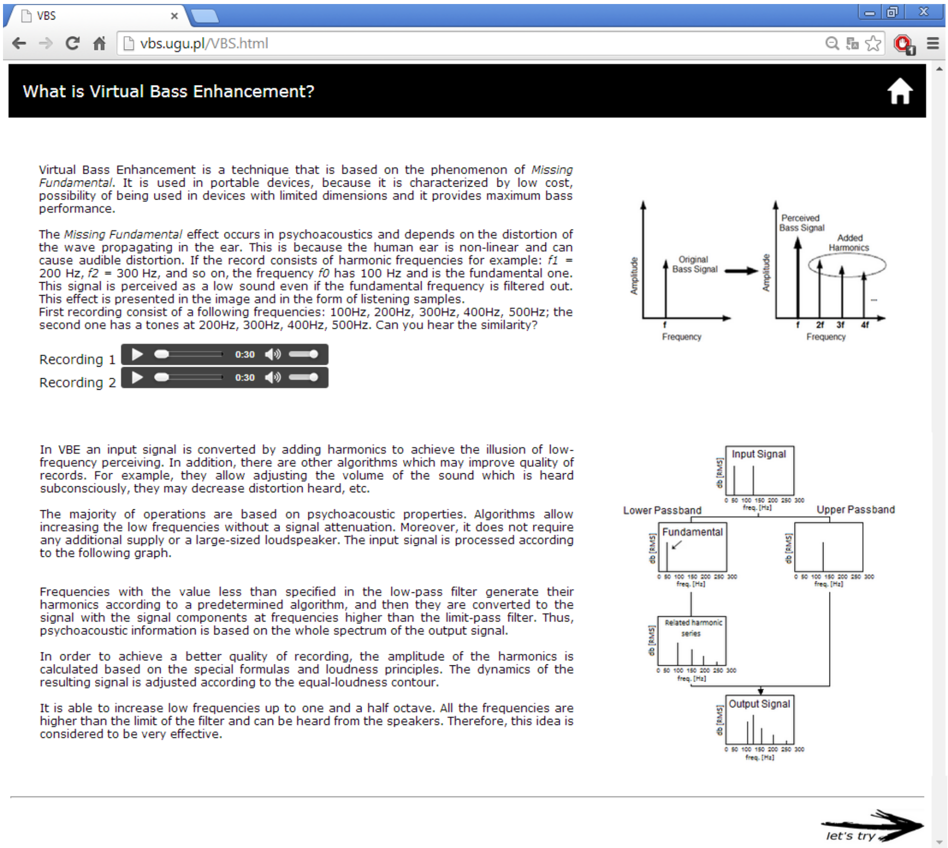


Figure 6.2. The VBS algorithms description

c) Questionnaire

Below, in Figure 6.3 there is a screenshot with the Questionnaire. In the upper right corner there is an icon with an image of a home which allows returning to the main page and start the test one more time. Below there are four fields to fill in and just below them - in order to maintain the ergonomics - there is a *Continue* button which allows executing the test.

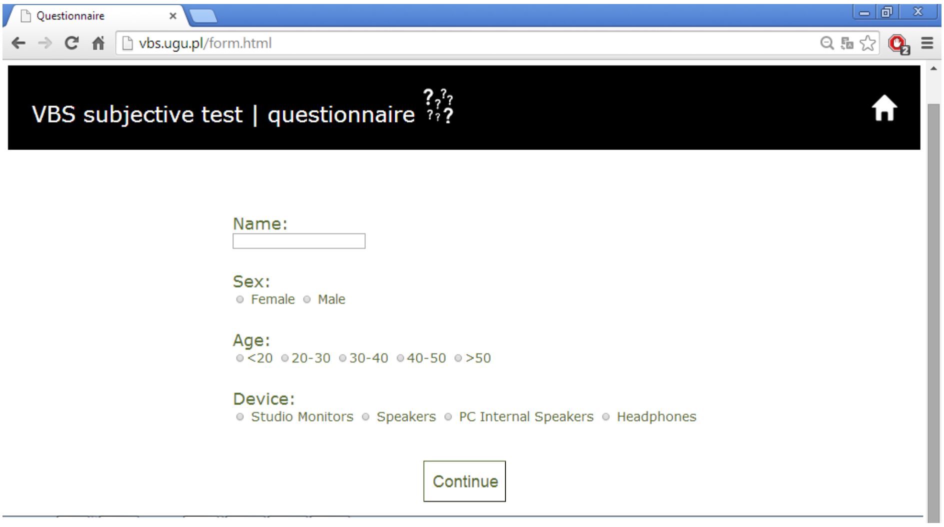


Figure 6.3. The Questionnaire Interface

d) Data transfer

After clicking on the *Continue* button the page informing about the correct or incorrect filling in form is presented. If every field has been marked, there is information about data transfer and it is possible to go to the listening test. The arrows are labeled with a type of music that will be presented at the next page. The screen of this page is shown in Figure 6.4.

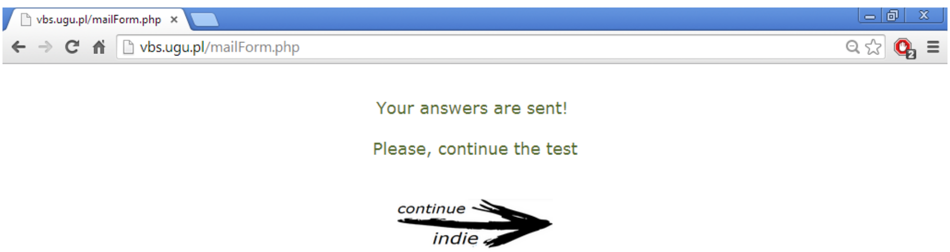


Figure 6.4. Confirmation of sent answers

In the case of incorrectly completed form, the information about the error appears and a listener is requested to fill in the missing gaps. Figure 6.5 presents a screenshot from this page.

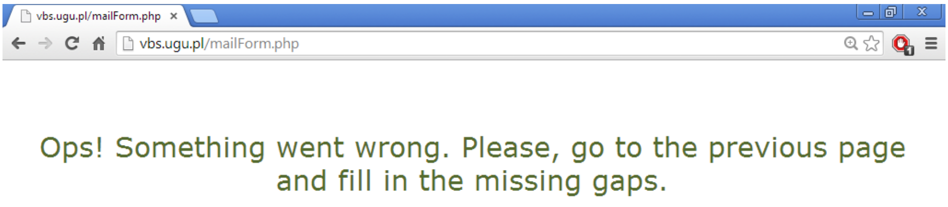


Figure 6.5.Incorrectly filled in form interface

e) Listening tests

Figure 6.6. shows the screenshot of one the 5 web pages of listening tests. All pages have a different name in the header and various icons correspondent to the type of music. Icons used for all the web pages are presented below in Figure 6.7. On every page there are four pairs of recordings. Clicking on the *play* button allows playing the sound. It is possible to reproduce a recording several times. The buttons located respectively below the recordings express the testers opinion. In order to increase the intuitiveness, the colors were used - red is responsible for clearly audible difference, and green for similar recordings. In addition there is a various amount of pluses. It is possible to select one of the six buttons prepared for each pair. There is also a home icon which allows accessing the home page and a button *Continue* redirecting to the page described in Section 6.1.d.

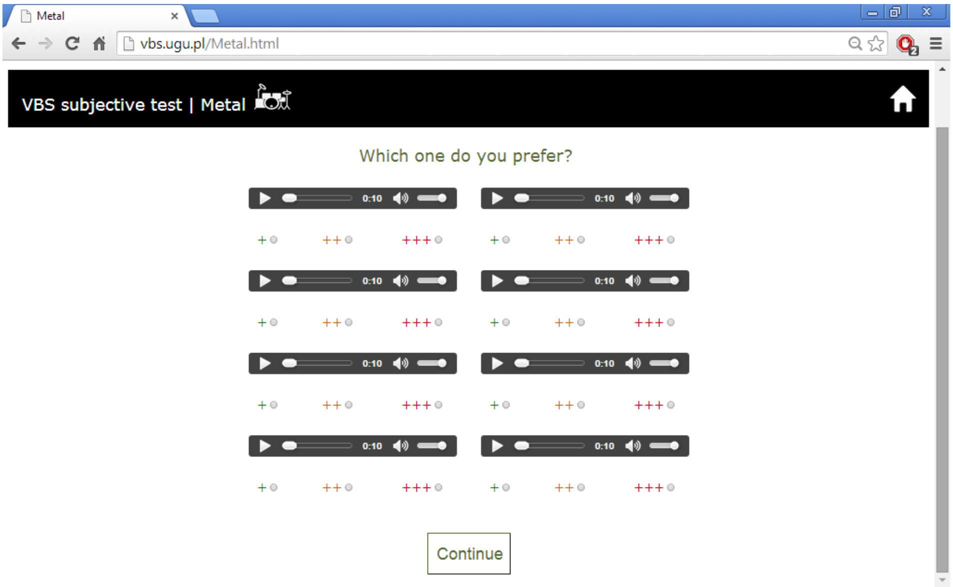


Figure 6.6. The Listening Test Interface

a) b) c) d) e)

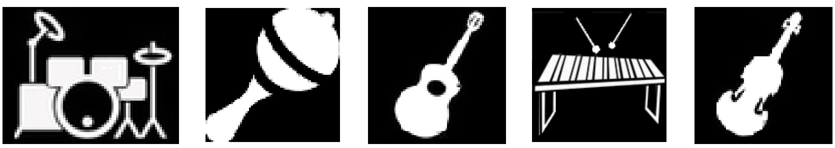

Figure 6.7. Icons used in the Listening Test Form depending on the type of music:

a) Metal, b) Reggae, c) Rock, d) Pop, e) Indie.

f) The Last Page

After completing the whole test and clicking on the *Continue* button, a page presented in Figure 6.8 is displayed. There is information that the test has been completed.

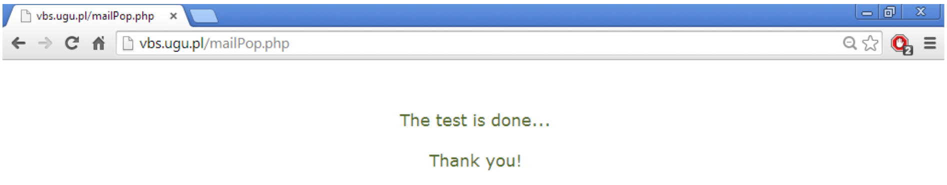


Figure 6.8 The Last Page

## 6.2. *Implementation*

### *6.2.1. Software - HTML*

The source code of the website has been written in HTML using Notepad. This is dedicated to create websites. It allows managing a design and adding such actions as playing music or moving between the pages. Below the excerpts of the code are presented.

1. add a header and a link to the Home Page as an icon

```
<tr>
    <td colspan="2" bgcolor="black">
        <table width="100%" cellpadding="15">
        <tr>
            <td align="left" valign="middle" bgcolor="black"> <font size="5" color = "white" face="Verdana">
            VBS subjective test | Metal </font><img src="metalIcon.png" width="60" height="40" /></td>
            <td align="right" valign="middle" bgcolor="black"><a href="index.html"><img src="home.png" width="40"
            height="40" /></a></td>
        </tr>
        </table>
    </td>
</tr>
```

2. play a sound

```
<tr>
    <td colspan="3"><audio src="recordings\metal3.wav" controls="controls"></audio></td>
    <td colspan="3"><audio src="recordings\metalOrg.wav" controls="controls"></audio></td>
</tr>
```

3. Display buttons - changing their appearance and adding the action

```
<tr>
    <td align="center"><font size="4" color = "green" face="Verdana">+</font><input type="radio" name="metal3" value
    ="3+"/></td>
    <td align="center"><font size="4" color = "#CC6600" face="Verdana">++</font><input type="radio" name="metal3"
    value="3++"/></td>
    <td align="center"><font size="4" color = "#CC0000" face="Verdana">+++</font><input type="radio" name="metal3"
    value="3+++"/></td>

    <td align="center"><font size="4" color = "green" face="Verdana">+</font><input type="radio" name="metal3" value
    ="org+"/></td>
    <td align="center"><font size="4" color = "#CC6600" face="Verdana">++</font><input type="radio" name="metal3"
    value="org++"/></td>
    <td align="center"><font size="4" color = "#CC0000" face="Verdana">+++</font><input type="radio" name="metal3"
    value="org+++"/></td>
</tr>
```

4. Send an e-mail with the answers

```
<form action="mailMetal.php" method="post">
```

```
            <td align="center" >
            </br>
            <input style="background-color:white; height:60px; width:120px; border-color:#556B2F; border-size:6px;
            color:#556B2F; font-size:1.5em"  name="send" type="SUBMIT" value="Continue">
            </td>
```

### *6.2.2. Using PHP to data transfer*

PHP is an object-oriented language for generating web pages in real time. Clicking on the *Continue* button redirects to a page created in PHP. At the beginning the properly filled in form is checked:

```
    if(!empty($_POST['metal1'])){
    if(!empty($_POST['metal2'])){
    if(!empty($_POST['metal3'])){
    if(!empty($_POST['metal2a'])){
```

If the form is completed the data is sent to the specified email:

```
if(isset($_POST['send'])) {
    $wiadomosc = "'Results : \n '". $_POST['mailMetal']
    . " \n\n metal1: ". $_POST['metal1']
    . " \n\n metal2: ". $_POST['metal2']
    . " \n\n metal3: ". $_POST['metal3']
    . " \n\n metal2a: ". $_POST['metal2a'];
    $temat = 'Results';
```

```
    if(mail("emiliagraban@gmail.com", $temat, $wiadomosc, "From: vbs@vbs.ugu.pl"))
    {
```

Additionally, the information about data transferring is displayed and transit to the next page is available.

```
    echo '
      <table width="100%" cellspacing="0" cellpadding="15">
      <tr>
      <td align="center" valign="middle" bgcolor="white"> <font size="5" color = "556B2F" face="Verdana">Your
      choices are sent!
    </br></br>Please, continue the test</br></br></br><a href="Rock.html"><img src="goRock.png" width="225"
    height="75" /></font></td>
      </tr>
      </table>
    ';}
```

## 6.2.3. *Data Collecting*

The results are sent instantly to the specified e-mail address. Below, in Figure 6.9 there is a screenshot from the received message

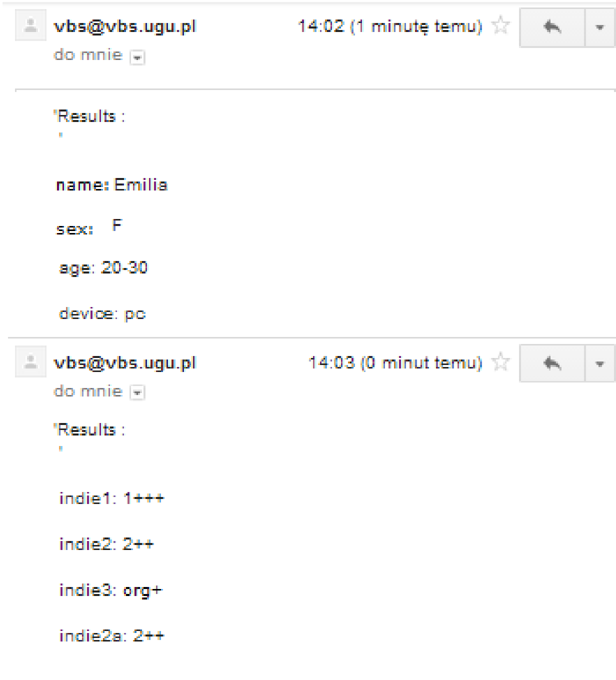


Figure 6.9. The results received by an e-mail

In the second mail the first result is presented as “indie1: 1+++” which means that the recording processed using algorithm called Setup A as described in Section 3.4.2 was assessed as better than the original recording, with grade "+++". However in the third pair the original recordings won but rated with the "+", which means the difference was not very clear.

### *6.2.4. Putting a Website on a Server*

To place the website on the Internet, it is required to use a server and appropriate network bandwidth. In order to carry the work out the free hosting [40] was used. Supplier allows selecting a domain and provides the information needed to send data from the local computer. Figure 6.10 shows the process of transferring resource from the local computer to the network using Total Commander.

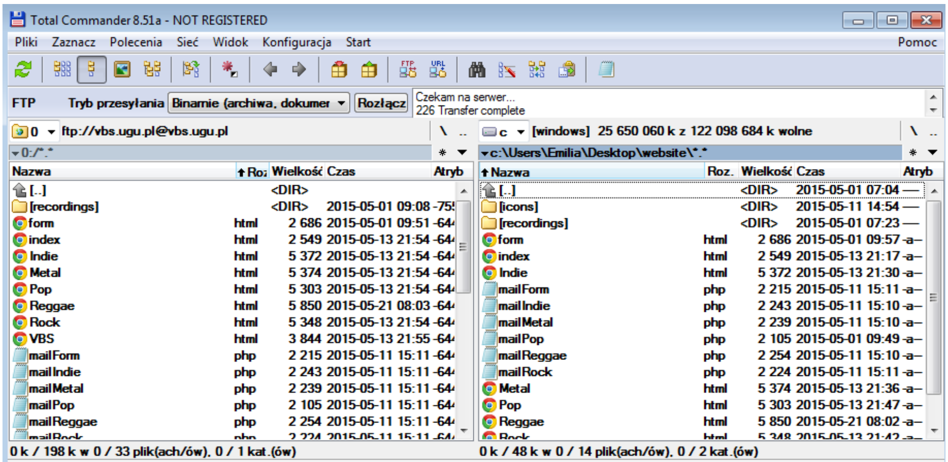


Figure 6.10 Resources transfer from the local computer to the server.

# 7. RESULTS AND DISCUSSION

The following Section provides the results of the listening tests. The most important relations are presented with appropriate comments. There is also some information about the stability of the listeners responses.

## 7.1. *Results Presentation*

The Virtual Bass Enhancement Algorithms tests were conducted on a group of 36 people. All the results were automatically sent to an email and entered in a table. In order to present significant correlations appropriate calculations and charts were created. All the results are available as an attachment to this work. The contents of the work contained only the calculations results, more important graphs and comments.

The subjective tests were designed to examine whether the audio processing with algorithms VBS bring expected results and make the recording preferable. Recordings differed in genres. The bass was enhanced in tracks with various instruments for every recording. Three different settings of VBS algorithms were used. The testers were of different age and gender. According to many dependencies, a lot of conclusions could be drawn.

Table 7.1 presents the portion of the collected data. The mark "2 P" means that the processed recording get 2 pluses "++" . “3 O” means "+++" for the original one. “Indie A” is a name of recording processed by algorithms called "Setup A" which is widely described in Section 3.4.

Table 7.1. The tests results from 6 listeners

| Name | Alicja | | Magdalena | | Rafał (Studio) | | Julia | | Johannes | | Paweł (Studio) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Gender | F | | F | | M | | F | | M | | M | |
| Age | 40-50 | | >50 | | 20-30 | | 20-30 | | 30-40 | | 20-30 | |
| Device | H | | S | | PC | | H | | PC | | PC | |
| Double | | x | | X | | x | | x | | x | | X |
| Indie A | 2 P | 2 P | 1 P | 1 O | 1 P | | 2 P | | 2 O | | 2 O | |
| Indie B | 1 O | | 1 P | | 1 P | 2 P | 2 P | 3 O | 3 O | | 2 O | |
| Indie C | 1 P | | 1 P | | 1 P | | 1 P | | 3 O | 1O | 1 P | 3 O |
| Reggae A | 3 O | 3 O | 1 P | 2 P | 1 P | | 3 O | | 1 P | | 2 P | |
| Reggae B | 3 O | | 2 O | | 1 O | 1 P | 3 O | 2 O | 1 O | | 3 O | |
| Reggae C | 3 O | | 1 P | | 3 O | | 3 O | | 1O | 2 O | 2 O | 2 O |
| Metal A | 3 P | 1 P | 1 O | 2 O | 3 O | | 3 O | | 3 P | | 1 P | |
| Metal B | 2 O | | 1 P | | 2 O | 3 O | 1 P | 2 O | 2 P | | 3 O | |
| Metal C | 1 O | | 2 O | | 3 P | | 3 P | | 1 P | 3 P | 1 O | 1 O |
| Rock A | 3 P | 1 P | 1 P | 2 P | 2 O | | 2 P | | 3 P | | 2 P | |
| Rock B | 3 P | | 1 O | | 2 P | 2 P | 2 O | 2 O | 3 O | | 2 O | |
| Rock C | 2 P | | 1 P | | 2 P | | 2 O | | 2 O | 1 O | 3 O | 1 P |
| Pop A | 3 P | 2 P | 2 O | 1 O | 3 O | | 3 O | | 1 O | | 1 P | |
| Pop B | 1 P | | 1 O | | 2 O | 1 P | 2 O | 3 P | 2 O | | 2 O | |
| Pop C | 1 P | | 1 O | | 3 P | | 2 P | | 2 O | 3 P | 1 P | 2 P |
| Suma P | 19 | | 8 | | 14 | | 13 | | 10 | | 8 | |
| Suma O | 13 | | 10 | | 16 | | 21 | | 20 | | 20 | |

### *7.1.1. General Results*

The graph in Figure 7.1 shows the responses from all of the 36 responders. They compared two recordings: the original and the processed one. Firstly they decided which one they prefered and marked their convintion by the quantity of "+". It was possible to select from 1 to 3 "+".

On the graph below the blue lines mean the sum of the pluses given for the recordings with enhanced bass during the whole test. Red color is for sum of pluses received by the original recordings.

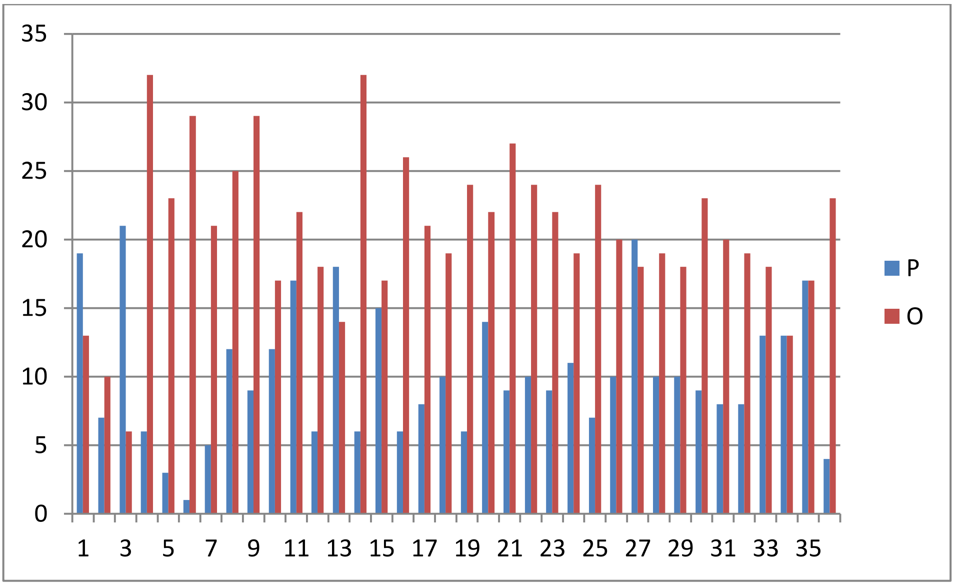


Figure 7.1. Distribution of pluses given for original and processed sound from 36 testers

More often the original recordings were evaluated as better ones. They received a twice more pluses (2,016). The original recordings got 744 pluses, and the processed ones 369. The graph shows that one of the listeners (No. 6) during the entire test admitted only one plus for the processed recording and 29 for original one. The average of results gives dominance for original recordings, but there were also testers which liked the processed recordings most. For example tester No. 3 gave 21 pluses for the processed song and only 6 for the original one.

*7.1.2. Music Genre*

The results were analyzed in more detail. The chart below shows the distribution of the results according to music genre.

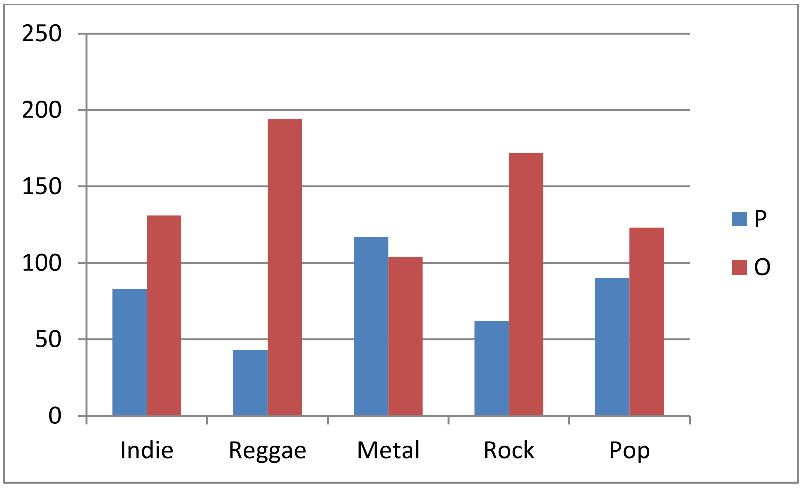


Figure 7.3. Distribution of pluses given for recording depending on the music genre

The music genre affects the assessment of the tracks. For each of them different tracks were processed. Table 7.2 presents which instruments were changed depending on music genre.

Table 7.2. The processed instrument depending on music genre

| Music genre | Instruments |
|---|---|
| Indie | Timpani, Bass Guitar, Strings |
| Reggae | Bass, Hi Hat and Organs |

| | |
|---|---|
| Metal | Electric Guitar, Bass Guitar, Drums, Kick |
| Rock | Vocal, Piano, Bass, Acoustic Guitar |
| Pop | Drums, Bass Guitar, Electric Guitar, Vibraphone |

Type of the instrument and its participation in a particular song have a big influence on final algorithms result. Various instruments have a different amount of low frequencies which are changed by VBS algorithms.

Song belonging to Metal were evaluated better with the enhanced bass. This excerpt is composed of 20 tracks but only 4 of them have been processed. A Tester while assessing the songs of Reggae and Rock definitely preferred the original ones. In these recordings the distortion was heard. The recording belonging to Rock is composed of 12 tracks, and 4 of them have been processed including vocals and piano. In Reggae song there were Hi Hat and Organs. These instruments have a lot of high frequencies.

### *7.1.3.MaxxBass Settings*

The evaluation of recordings depends on parameters set in the MaxxBass. They are thoroughly described in Section 3.4.2. Figures 7.2 shows the interfaces sequentially for Setup A, B and C.



A B


C

Figure 7.2. Interfaces for Setup A, B and C set in MaxxBass.

The chart in Figure 7.3 shows the distribution of the amount of "+" for recording depending on the parameters set in the MaxxBass.

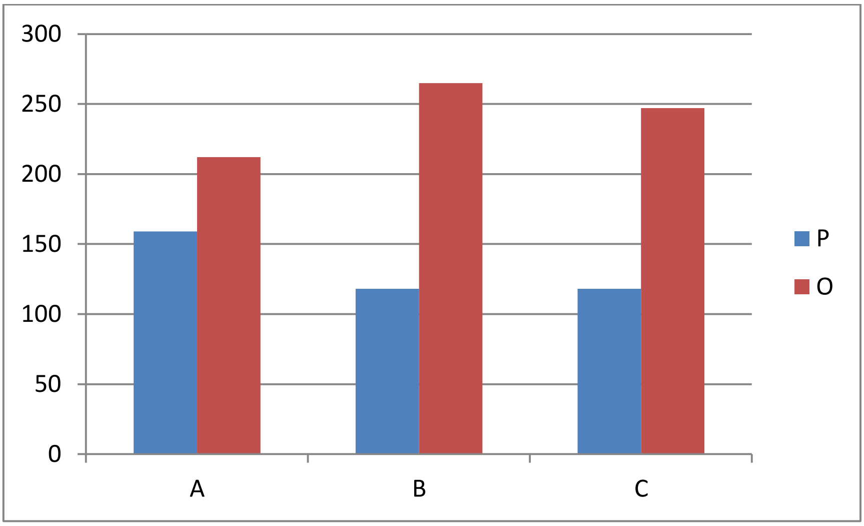


Figure 7.3. Distribution of pluses given for recording depending on the parameters set in the MaxxBass.

None of the MaxxBass settings improved the recordings. For the setup B original tracks obtained 265 pluses and with enhanced bass only 118. Processing in such a way causes a lot of distortion and tracks are heard unnatural. The parameters in the setup B were chosen based on a subjective assessment in the individual tracks – not in the mix. Below 220 Hz original frequencies have been replaced by generated harmonic – it is relatively high. The value of the harmonic is also too large - 10 dB. These values deviate from the recommendation and cause distortions.

Recordings processed with Setup A provided much better effect. Cut-off frequency was set to 120Hz. The slope of the filter was set at 24dB per octave which is dedicated to the sounds reproduced by speakers with limited size. These tracks got 159 pluses. On average the original recordings were evaluated better but received less pluses than with Setup B and C.

### *7.1.4. Music Genre and Settings*

The graph below (Figure 7.4) shows the results depending on the music genre and MaxxBass Settings

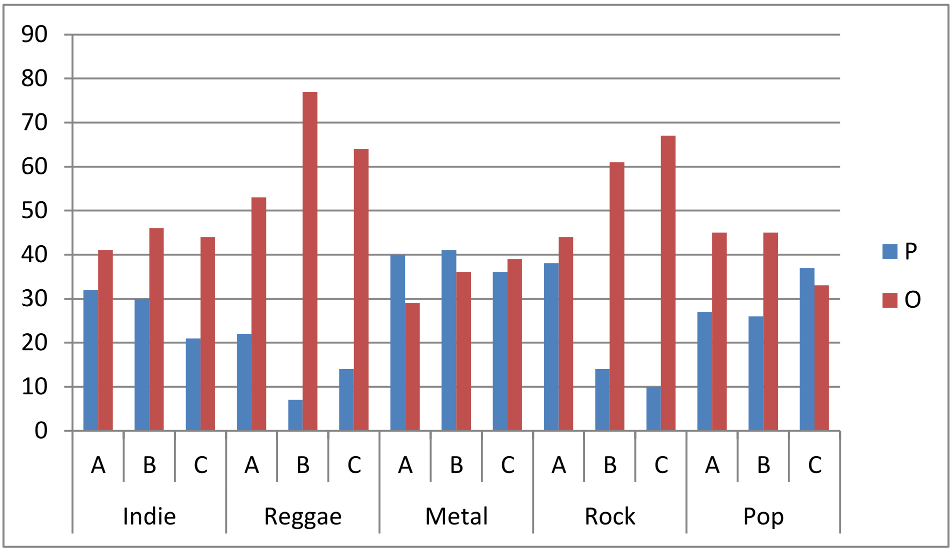


Figure 7.4. Distribution of pluses given for original and processed recording depending on the music genre and MaxxBass Settings

Analysis of this chart provides additional conclusions. On average, processing with Setup C worsen the track but the song designated as Pop was improved. In Reggae music, track processed in the same way has been evaluated worse and the original recordings obtained almost 5 times more pluses.

For the Rock music changes set as C provides a higher assessment than by A and B, while in Metal the recordings processed by Setup B was better than by Setup A. It may be related to the amount of high-frequency in Rock and high value at which original frequencies were increased (Setup B - 220Hz).

### *7.1.5. Testers' Age*

The test results were also analyzed depending on age. Below the graph in Figure 7.5 indicated the sum of pluses for the processed and original recording depending on age. Most testers belong to the group of 20-30 years old (14 people), four people under 20 years took part in the test, six people aged 30-40, seven between 40 and 50 and five people over 50 years. The distribution shows on the chart indicates that young people preferred the original recording. Older people chose equally the original and processed tracks. With increasing age the hearing is worse and some distortions could not have been heard. Both of the tracks were evaluated better.

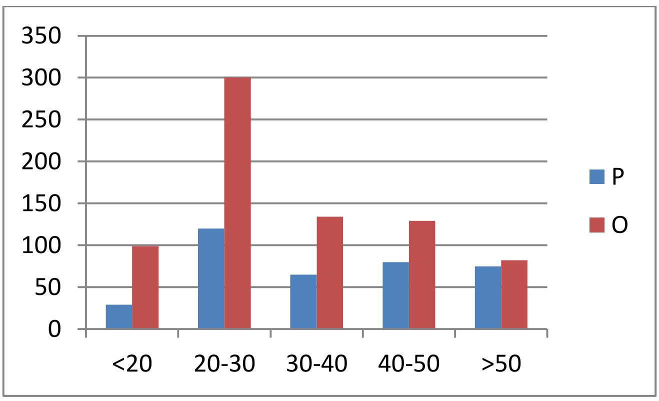


Figure 7.5. Distribution of pluses given for original and processed recordings depending on testers' age

### *7.1.6. Testers' Gender*

In the charts from Figure 7.6. the distribution of pluses for recording processed and original in relation to gender is shown. 19 women and 17 men took part in the test. The ratio of pluses given for original recording to processed for women is 2.18, but for men 1,86.It means that men like recordings with enhanced bass more.

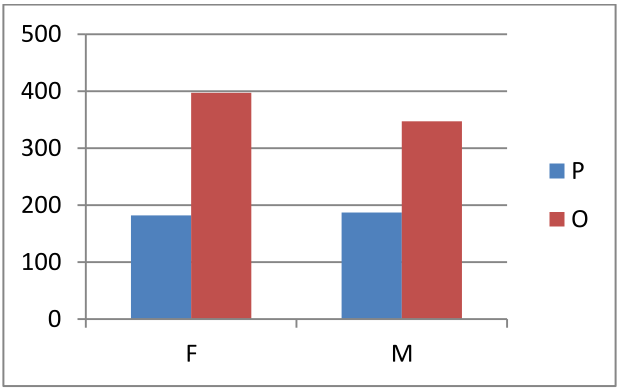


Figure 7.6. Distribution of pluses given for original and processed recordings depending on testers' gender

## *7.2. Differences In Responses*

In order to study the stability of the audience responses in the test the one of pairs was doubled. For every music genre only one pair was repeated. In every test it was a different pair so the results for Setup A, B and C were collected double from 12 students. Responses were compared for the same pairs and the absolute difference was calculated. The following scoring scale was assumed:

3 P (+++ for processing track) = 6

2 P = 5 1 P = 4 1 O = 3 2 O = 2 3 O = 1

If the tester comparing the original and processed record for the first time assigned “++” to the original track (2 O = 2 points), and the second time," ++ " to the processed recording (2 P = 5 points) made a mistake by 3 points. Table 7.3 shows the sum of the mistakes from all of the testers in every pair.

Table 7.3. The sum of the mistakes

| Music | Setup | R | Sum |
|---|---|---|---|
| Indie | A | 15 | |
| | B | 30 | |
| | C | 21 | 66 |
| Reggae | A | 9 | |
| | B | 7 | |
| | C | 9 | 25 |
| Metal | A | 17 | |
| | B | 14 | |
| | C | 14 | 45 |
| Rock | A | 7 | |
| | B | 11 | |
| | C | 10 | 28 |
| Pop | A | 24 | |
| | B | 25 | |
| | C | 26 | 75 |

The chart from Figure 7.7 shows the sum of mistakes from all users depending on the type of music.

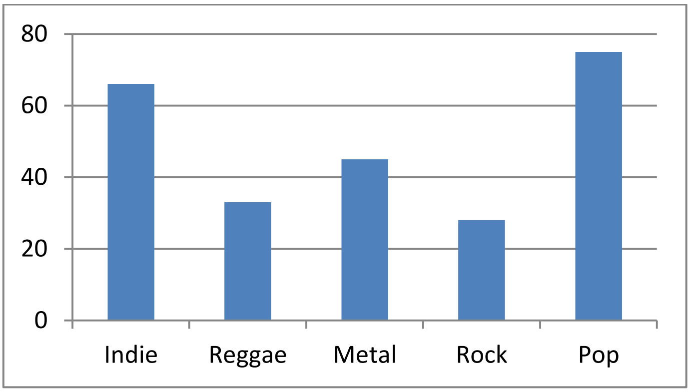


Figure 7.7. Distribution of mistakes depending on types of music

In Pop and Indie music the differences in the responses are the highest, it means that the listeners were not confident in their decisions. In the case of Reggae and Rock the differences are lower and these recordings are actually evaluated better.

It was also analyzed whether the test conditions - in the laboratory or on the website affect the stability of the response. In laboratory 9 tests were carried out and via the website 27. In order to maintain proportion in the graph (see Figure 7.8), the results obtained in laboratory were multiplied by 3. Table 7.4 shows the results

Table 7.4. The sum of the difference responses depending on the tests condition

| Music | Laboratory | L x3 | Website |
|---|---|---|---|
| Indie | 8 | 24 | 58 |
| Reggae | 6 | 18 | 27 |
| Metal | 7 | 21 | 38 |
| Rock | 3 | 9 | 25 |
| Pop | 10 | 30 | 65 |

The chart (Figure 7.8) shows the distribution of all mistakes occurring in the laboratory and on the website.

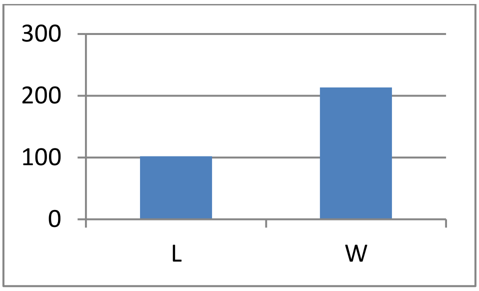


Figure 7.8. Distribution of mistakes depending of tests' environmental.

In the laboratory the testers were more stable in their responses. There were the best conditions provided, the instruction were better explained and the testers were more focused on the test. People were listening more carefully and perceived more differences in recordings. Their answers are more objective.

# 8. CONCLUSION

This chapter contains summary of the study performed within this M.Sc. work, pointing out the most important conclusion. Moreover some ideas for future work and improvement are also presented.

## *8.1. Main Conclusions*

The most general analysis indicates that original tracks were evaluated better than the processed ones. Six main conclusions are listed.

The evaluation of algorithms depends on the factors included in three groups:

- the MaxxBass algorithms parameters: the cut off frequency, the value of original and enhanced bass after processing
- the mixed track parameters: the quantity of separate tracks, the quantity of processed tracks, types of music, content of low and high frequencies in the processed track.
- testers age and gender, tests environmental condition

<u>Conclusion 1:</u> The more low frequencies in processed tracks, the better evaluated output songs.

The song belonging to Metal which is characterized by high content of low frequency as the only one was rated better after processing. The Reggae and Rock songs were evaluated better in original version – they consist of lots of high frequencies.

The conclusion can be drawn that contents of low frequencies in a processing track have the influence on final results.

Conclusion 2: The smaller proportion between the processed and original tracks, the better results.

The proportion between the quantity of processed and original tracks in one song has some influence on the quality of output recording. In Metal song the 4 processed tracks were mixed with 16 original ones – these recordings were evaluated better. The Rock song where 4 enhanced tracks were mixed with 8 original output song was the worst.

Conclusion 3 – The lower cut-off frequency the better results.

The MaxxBass setting called „B", characterized by high cut-off frequency brings about the worst effect. In this case the basses below 220 Hz were modified by MaxxBass algorithms. Original Bass was set to 0 dB but the MaxxBass was enhanced to 10 dB. Meanwhile, the tracks processed by setting „A", where cut off frequency was set to 120 Hz were evaluated better. According to this fact it is assumed that modifications in wide bandwidth lead to distortion and unnatural effect.

Conclusion 4 – The more low frequencies in record the wider bandwidth of modified frequencies is desirable.

It was concluded that lower cut-off frequency brings better effects. However in Metal song, which has a lot of low frequencies the modifications in wider range are desirable. In songs belonging to Reggae and Rock which have processed tracks with higher frequencies, the modified range should be smaller.

Conclusion 5 – Older people and men evaluated the processed and original tracks with the similar value.

It was proven that the confidence in responses and perception of differences between evaluated tracks decrease with age. The evaluation of the processed and original tracks was similar. Younger people and women assessed tracks variously and gave more pluses more often.

Conclusion 6 – The better conditions during the test, the more objective responses.

In the laboratory testers were more steady in their responses. The answers in the first and the second assessment of the same song were almost the same.

## *8.2. Future Work*

Despite the fact that the study gathered quite many results, there is still a need for further research. It is required to conduct more advanced tests. For example, to draw more objective conclusions, the number of listeners representing a particular group of people should be equal. There is also a need to check whether the quantity and types of instruments in the separated audio track improve the recording. In order to increase objectivity of the results, testing conditions and devices should be the same. Tests determining whether there is no hearing loss within the group tested should also be checked. An audiometric test should be taken into consideration for that purpose.

Based on the author's experience gained from the results analysis, if this work starts now more tests under laboratory conditions should be performed. It was confirmed that these results are more reliable – people have assigned the double recordings with the same grade in almost every case.

According to the presented analysis the same excerpt could be processed in two ways. First, separate tracks which contain more high frequencies (for example piano, violin, female voice) could be processed, then, in the same excerpt bass in track characterized by low frequencies (contrabass, bass guitar, drum) could be enhanced. Moreover tracks belonging to music genre significantly different from each other, for example: heavy metal and opera could also be processed. The results could give an answer whether the algorithms parameters (frequencies below which the original signal is amplified; filters slope; response time etc.) would improve music depending on quantity of low or high frequency content. There is a possibility to design a software analyzing the music genre or frequency band content and change algorithms parameters in real time.

Moreover, values of MaxxBass parameter could be changed. They should differ more significantly. Original bass might be filtered out below the high frequency or reduced a little below a relatively low frequency. In the current version the setups called "A" and "C" are quite similar and provide only few conclusions. Also decay and response value could be more differentiated – now they are too similar and that didn't help to find any relationship between them.

The designated website could also be improved. It should enable an access from multiple browsers. Moreover, in the instruction examples that are given provide too small differences between tracks in pairs which sometimes are hard to notice - in particular, on the first subpage presenting Indie music – many testers were confused. The website should also send only one email from one person but containing all the listening test results. In the current version six mails from one person were sent.

Finally, there are a lot of possibilities to extend the tests of the virtual bass enhancement algorithms. New technologies are developing rapidly and all the conclusions and ideas for the research should be taken into consideration.

## REFERENCES


1. Bass Enhancer Plugin, *MaxxBass*, http://www.waves.com/plugins/maxxbass, (Retrieved: 10.03.2015).
2. Ben-Tzur D., Colloms M., T*he Effect of the MaxxBass Psychoacoustic Bass Enhancement System on Loudspeaker Design, 106* Audio Eng. Soc., Convention, Paper 4892, 1999.

3. Shouten J. F., *Acoustics.- The residue and the mechanism of hearing,* Natuurkundig Laboratorium der N. V. Philips' Gloeilampenfabriken, Eindhoven – Holland, Communicated at the meeting of September 28,1940.

4. Purves et.al.: *Neuroscience*, Chapter 12: The auditory system, 223-243.

5. von Bekesy G., Experiments in Hearing, Acoustic Society of America, 1989.

6. Rossing D. T., *Springer Handbook of Acoustics*, ISBN: 978-0-387-30446-5, Springer, 2007.

7. Physiology of the Cochlea, Structural properties of the basilar membrane determine the way it responds to sound, http://www.rci.rutgers.edu/~uzwiak/AnatPhys/Sensory_Systems.html.

8. Jay F., editor., *IEEE Standard Dictionary of Electrical and Electronics Terms*, 2nd edition, 1978.

9. Larsen E. and Aarts R. M., *Audio Bandwidth Extension: Application of Psychoacoustics, Signal Processing and Loudspeaker Design*, *Introduction*, 2004.

10. Larsen E. and Aarts R. M., *Audio Bandwidth Extension: Application of Psychoacoustics, Signal Processing and Loudspeaker Design*, *Chapter 2: Psychoacoustic Bandwidth Extension for Low Frequencies*, 2004.

11. Hoffmann P., Sanner T., Kostek B., *An Approach to Bass Enhancement in Portable Computers Employing Smart Virtual Bass Synthesis Algorithm,* Audio Eng. Soc., Convention, e-Brief 144, Berlin, 2014.
12. Bregman A.S., *Auditory Scene Analysis*, MIT Press, 1990.
13. Kedem B., *Time Series Analysis by Higher Order Crossings*, IEEE Press, New York, 1994.
14. Han-gil Moon, Manish Arora, Chiho Chung, Seong-cheol Jang: *Enhanced Bass Reinforcement Algorithm for Small-sized Transducer*, Audio Eng. Soc., Convention, Paper 7024, Vienna, 2007.
15. Audio Editor, *Audacity*, http://download.cnet.com/Audacity/3000-2170_4-10058117.html, (Retrieved: 10.03.2015).
16. Hill J. A., O. J. Hawksford M.: *A hybrid virtual bass system for optimized steady state and transient performance,* Audio Research Laboratory, University of Essex Colchester, United Kingdom.
17. Mingsian R. Bai, Wan-Chi Lin, *Synthesis and Implementation of Virtual Bass System with a phase-Vocoder Approach*, J. Audio Eng. Soc., Vol. 54, No. 11, pp. 1077-1091, Nov. 2006.
18. ISO/DIS 226, *Acoustic – Normal equal-loudness contours*, 2001.
19. Laroche J., Dolson M., *New phase-vocoder techniques for real-time pitch shifting, chorusing, harmonizing, and other exotic audio modifications,* J. Audio Eng. Soc., Vol. 47, No. 11, pp. 928-936, Nov. 1999.
20. Gerstle B. T., *Tunable Virtual Bass Enhancement,* May, 2009.
21. Larsen E. and Aarts R. M., *Audio Bandwidth Extension: Application of Psychoacoustics, Signal Processing and Loudspeaker Design*, *Chapter 1: From Physics to Psychophysics*, 2004.

22. Maxx Bass Documentation, http://www.waves.com/1lib/pdf/plugins/maxxbass.pdf, (Retrieved: 28.02.2015).
23. *The 'Mixing Secrets' Free Multitrack Download Library* http://www.cambridge-mt.com/ms-mtk.htm#Acoustic (Retrieved: 10.03.2015).
24. Arnett J., *Heavy Metal Music and Reckless Behavior*, Journal of Youth and Adolescence, Vol 20, No 6, 1991.
25. Kosina K., *Music Genre Recognition*, Fachhochschul-Studiengang, Hagenberg, 2002.
26. *How To Play Reggae: Characteristics of Reggae*, www.how-to-play-reggae.com, (Retrieved: 10.03.2015).
27. Drobner M., *Instrumentoznawstwo i akustyka (in Polish)*, Polskie Wydawnictwo Muzyczne Kraków, 1985.
28. Sweeney P., *Drums for the Absolute Beginner*, Alfred Music Publishing, 2004.
29. Bridge R., *Timpani Constructor Paper*, Percussion 1, 1993.
30. Lembessis V., *Physics… in action,* Europhysics News, The University of Southeastern Europe, Athens, Vol. 32 No. 4, 2001.
31. Kostek B., Akustyka Muzyczna, Instrumentoznawstwo (*in Polish*), https://sound.eti.pg.gda.pl/student/akmuz/05-Instrumentoznastwo.pdf (Retrieved: 03.2015).
32. Brown J.C., Vaughnb K. V., *Pitch center of stringed instrument vibrato tones*, Institute of Technology, Cambridge, Massachusetts, 1996.
33. Bugg D., *The role of turkish percussion in the history and development of the orchestral percussion section,* A Monograph, The College of Music and Dramatic Arts, Missisipi, 2003.
34. Reblitz A, *Piano Servicing, Tuning, and Rebuilding. For the Professional, the Student, and the Hobbyist*. Vestal Press, Lanham Maryland, 1993.
35. Audio Editor, *Adobe audition*, trial version, https://creative.adobe.com/pl/products/audition (Retrieved: 10.03.2015).
36. EBU-TECH 3343, *Practical guidelines for Production and Implementation in accordance with EBU R 128,* Geneva, 2011.

37. Head Acoustic, Application Note, *Listening Tests*, 06/11, https://www.head-acoustics.de/downloads/eng/application_notes/ListeningTests_06_11e..pdf (Retrieved: 05.2015).
38. Kostek B., *Technologia Nagrań - Testy Subiektywne*, Politechnika Gdańska, https://sound.eti.pg.gda.pl/student/tn/testy_subiektywne.pdf (Retrieved: 05.2015).
39. VBS algorithms Tests, www.vbs.ugu.pl, (Retrieved: 05.2015).
40. Free hosting for websites, www.ugu.pl,( Retrieved: 05.2015).

## LIST OF FIGURES

## LIST OF TABLES

## LIST OF APPENDIX

1. Source code (*HTML, *PHP)
2. Icons: *png
3. Recordings: *wav
4. Test results *xlsx
    a) General Results
    b) Age and Gender
    c) Differences in responses