International Science and Technology Journal
المجلة الدولية للعلوم والتقنية

Volume 36 العدد
Part 2 المجلد

http://www.doi.org/10.62341/amka8005

Received	2025/03 /20	تم استلام الورقة العلمية في
Accepted	2025/ 04 /15	تم قبول الورقة العلمية في
Published	2025/ 04 /17	تم نشر الورقة العلمية في

Modulate the Tonal Quality of the Human Voice by Employing Frequency Filter Effects.

Adel Masoud Khalifa Asker

Electrical and Electronic Department, Higher Institute of Science and Technology, Nalut, Libya

asker36916@gmail.com

https://orcid.org/0009-0003-2884-3915

Abstract

The technique of changing the tone of a voice can be done by controlling its frequency components after passing through the high, medium, and low frequency control channels to eliminate the power spectrums of unwanted frequencies and maintain the necessary frequencies to achieve the desired sound. This has been done after comparing the spectrum of the voice recorded segment with some spectral frequencies of the same sentence with a more male tone. A commonly used vocal phrase spoken in a female voice, such as a woman saying "As-salamu alaykum," a well-known Islamic greeting, was chosen. An attempt was made to try changing its tone to that of a male voice saying the same sentence; "As-salamu alaykum." The results were compared with the display of the resulted spectral frequencies of the same phrase in original male voices, taking into account the findings of previous studies in this field (Norman D. Cook, 2002). Comparing the results, a similarity rate exceeding 80 percent was obtained, and this makes an FDAtool use as one of the best choices offered by MATLAB for advanced filtering.

Keywords: tone of voice, vocal phrase, frequency components.

http://www.doi.org/10.62341/amka8005

تعديل جودة الصوت البشري باستخدام تأثيرات المرشحات الترددية

عادل مسعود خليفة عسكر
قسم التقنيات الكهربائية و الإلكترونية ،المعهد العالي للعلوم و التقنية، نالوت، ليبيا
asker36916@gmail.com
https://orcid.org/0009-0003-2884-3915

الملخص

يمكن إجراء تغييرات في نغمة أي صوت البشري و ذلك عن طريق التحكم في مكوناته الترددية بعد تمريره عبر قنوات للتحكم بالترددات العالية والمتوسطة والمنخفضة، و يتم ذلك بإزالة أطياف طاقة الترددات الغير المرغوبة والحفاظ على الترددات المرغوبة و اللازمة للحصول على نغمة الصوت المطلوب. وقد تم في هذا العمل مقارنة أطياف ترددية لمقطع صوتي تم تسجيله بأطياف ترددية لنفس الجملة تم تسجيلها أيضا و لكن بنغمة ذكورية أكثر، تم اختيار عبارة صوتية شائعة الاستخدام و هي تحية إسلامية مشهورة تقول فيها امرأة بصوت أنثوي، "السلام عليكم". وجرت محاولة لتغيير نغمتها إلى نغمة أكثر رجولية تكرارا نفس الجملة ؛ "السلام عليكم". و تمت مقارنة النتائج مع مخططات للترددات الطيفية الناتجة عن نفس العبارة بأصوات ذكورية أصلية، مع مراعاة نتائج الدراسات السابقة في هذا المجال (Norman D. Cook, 2002). وبمقارنة النتائج، تم الحصول على معدل تشابه يتجاوز 80 % بين هذه الأطياف، مما يجعل استخدام الـ FDAtool هو أحد أفضل الخيارات المتاحة التي يقدمها MATLAB للفلترة المتقدمة.

الكلمات المفتاحية: نغمة الصوت، المكونات الترددية، عبارة صوتية.

Introduction

A Great attention is now paid to the appropriate tone of voice to address the chosen audience when marketing a product or merchandising service or simply answering some inquiries. Discerning the tone of voice is extremely important because of its impact in many ways, whether emotional, personal, social, or other areas in which convincing engagement is sought. These may include politics, religion, drama, the military, or the arts. It may extend to dealing with animals, in particular pets (Neil Patel, 2024)

The challenge facing those interested in this field is to propose a system that aims to modify human vocal recordings in a known rule-

http://www.doi.org/10.62341/amka8005

based manner, such as the one we have indicated. To obtain a realistic performance, the main parameter, which is the rhythm, must be controlled. A model for other modifications was presented in a previous work (Bastos, Wilson, and Sidney J. Levy, 2012). During the recent developments of that system, it became clear that the pronunciation and dynamics modifications must be applied on a note-by-note basis. For this reason, the analysis and synthesis approach in our work was considered the best option for such operations because it is possible to modify the dynamic level of each note in a realistic way by changing both its amplitude and its spectrum.

Methodology

The frequency response in the FDAtool we used in MATLAB shows the balance "volume levels" of the frequencies (tonal balance) in the same way as the equalizer setting turned on for headphones or speakers with "absolute flat frequency response." What we mean is that while tuning the equalizer, and you decided to add middle bass tone, you push up the levers that tune the frequencies you want to hear on more volume levels. The frequency response of the headphones is dependent on the sound pressure level of the frequency of the reproduced harmonic signal at the headphone output. A parametric equalizer (EQ), which is the most common and versatile type of equalizer, is considered to be a multi-band variable equalizer that allows every aspect of the equalization range to be adjusted ideally for both broad changes and precise adjustments. The octaves are for example the frequency ranges: 20 to 40 Hz, 250 to 500 Hz, 3 to 6 kHz as represented in table1 and figure 1.

Table 1: Names of frequency responses with ranges
(Understanding Audio Frequency Range in Audio Design, 2002):

Range	Name
20 - 40 Hz	Low Bass
40 - 80 Hz	Mid Bass
80 - 160 Hz	Upper Bass
160 - 320 Hz	Lower Midrange
320 - 640 Hz	Middle Midrange
640 Hz - 1.28 кHz	Upper Midrange
1.28 - 2.56 кHz	Lower Treble
2.56 - 5.12 кHz	Middle Treble
5.12 - 10.2 кHz	Upper Treble
10.2 - 20.4 кHz	Top Octave

http://www.doi.org/10.62341/amka8005

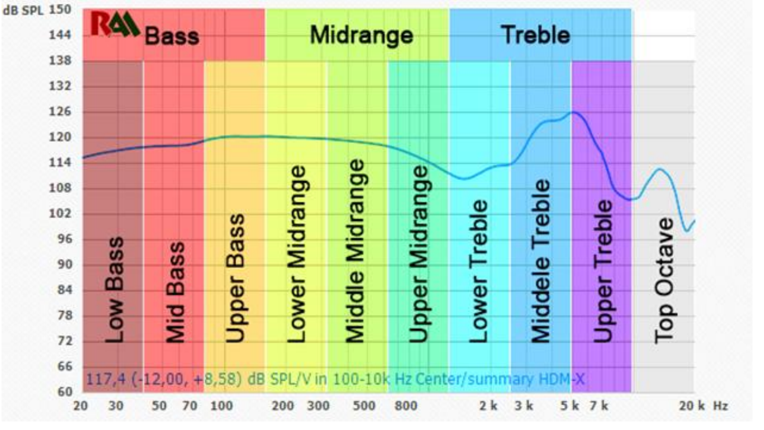

Figure 1. The frequency range is divided into low, medium and high frequencies
(Understanding Audio Frequency Range in Audio Design, 2002)

Filters affection in the equalizers

Equalizers are deliberately designed to create fairly minor changes in the signal. For more drastic effects, such as removing some regions of the signal entirely, a filter is required, where it is just a circuit that sharply reduces the amplitude of signal frequencies outside of specified limits. The unaffected region is called the passband, and the filter type is named after the passband as low-pass, high-pass, or band-pass (Davis, Melissa, 2009). The point where the signal attenuation becomes noticeable (a reduction of 3 dB) is termed the cut-off frequency. A low-pass filter with a cut-off of 500Hz will attenuate any signal of frequency above 500Hz. The attenuation provided by a filter is never absolute. The frequency response graph of a filter in Figure 2 shows that the amplitude of the signal decreases as the frequency moves past its maximum. The actual rate of this decrease is a filter design parameter called a slope. The slope is a general description of the response from the cut-off. To begin with, the curve at which the signal begins to roll off is not a sharp and well-defined frequency. It is a standard practice to take the frequency at which the signal level has fallen by 3 dB as a useful criterion in describing the characteristics of that graph. This frequency is called, the turnover frequency or the 3 dB down point (Ambili Vipin, 2018). Different circuits vary in the shape of the curve near the cut-off point, and the possible shapes have names such as Bessel and Chebychev, after the creators of the mathematical formulas involved. Another design parameter that

http://www.doi.org/10.62341/amka8005

affects the shape of the filter curve is known as "Q." (Parametric Audio Equalizer on mathworks.com).
The derivation of Q is too complex for this discussion but it is useful to know that filters with a high value of Q have an amplitude bump near the cut-off frequency and have a tendency to oscillate at this frequency.

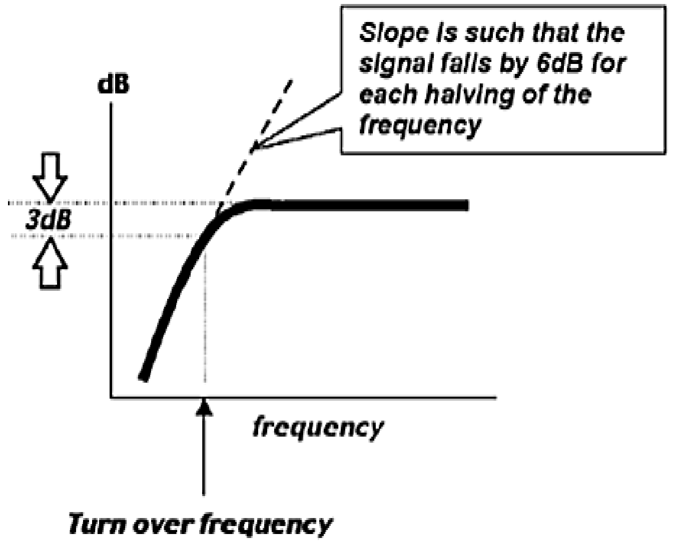

Figure 2. The slope in the filter design
(Ambili Vipin, 2018).

Applying the Parametric Equalizer

The parametric equalizer is the most powerful and flexible type of equalizers, where the mid-range bands in a parametric equalizer have three adjustments: gain, centre frequency, and quality factor (Q) or bandwidth, so the parametric equalizer allows the operator to add a peak or notch at an arbitrary location into the audio spectrum. For other frequencies, away from the peak or notch, the parametric equalizer does not modify the spectral content, as its response to volume where it is one (0 dB). Adding a peak can be useful to help an instrument to be heard in a complex mix, or to intentionally add colouration to an instrument's sound by boosting or reducing a particular frequency range. Notches could have mitigated unwanted sounds, including removing power line hum (500 Hz or 600 Hz and sometimes their harmonics) and reducing feedback (Parametric Audio Equalizer on mathworks.com). To remove artefacts without affecting the rest of the audio, a narrow bandwidth "female voice" for the phrase "Peace be upon you" will be used. One section of the

http://www.doi.org/10.62341/amka8005

parametric equalizer is created using MATLAB code designed with an "FDAtool" add-on MATLAB tool to create the required filters for the first-order peak/notch filter, or in certain cases, the first-order shelving filter for the lowest and highest frequency bands. The transfer functions for the first-order shelving filter and the first-order peak/notch filter are derived. The multiband parametric equalizer performs multiband parametric equalization independently across each input channel using specified centre frequencies, gains, and quality factors. It can be configured with up to 10 bands "while just 3 bands are used here", and can be done by just adding low-shelf and high-shelf filters, as well as high-pass (high-cut) and low-pass (low-cut) filters. To perform multiband equalization on each channel of the input signal according to its original characteristics in the MATLAB code, the input must be a real-valued, double-precision, or single-precision matrix, because the equalizer treats each column of the input as an independent channel. Figures 3 and 4 show both female voice signals (original and equalized) in real-time and frequency domains.

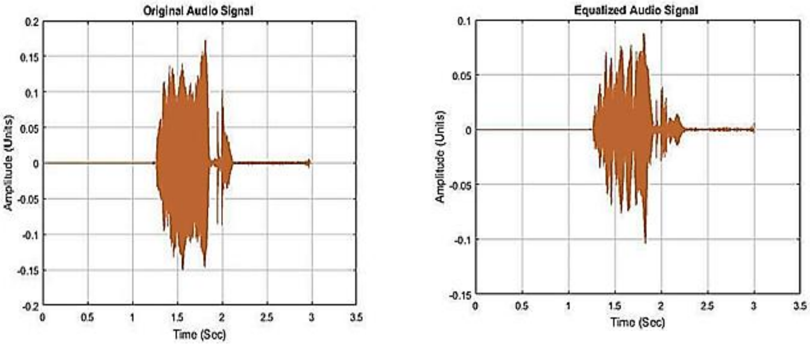

Figure 3. The original and equalized signals in the real-time domain

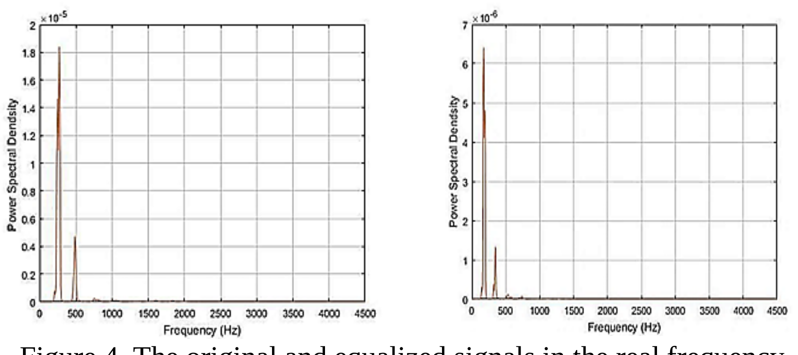

Figure 4. The original and equalized signals in the real frequency domain

http://www.doi.org/10.62341/amka8005

The Cumulative Frequency Boundaries

The cumulative frequency graph is a line graph displaying the cumulative frequency of each class at its upper-class boundary. The upper boundaries are marked on the horizontal axis, and the cumulative frequencies are marked on the vertical axis. The graph should start at (or just before) the lower boundary of the first class (where the cumulative frequency is zero), and end at the upper boundary of the last class . (RECOMMENDATION ITU-R SF.675-4).

The graph should increase from left to right, with the points evenly spaced along the horizontal axis. It is clear from Fig.4 that many of the frequency components have been removed after the signal has been filtered, leaving only the desired frequencies. The frequency spectrum density calculated using the FDAtool is approximately 0.6404 W/Hz.

Filters Effect

Equalizers are intentionally designed to create fairly subtle changes in the filtered signal. After the filter reduces the amplitude of frequency signals outside of specified limits, the unaffected area is called the passband (Parametric Audio Equalizer on mathworks.com).

Figure 5 shows parametric equalizer filters designed for the three bands using the FDAtool in MATLAB:

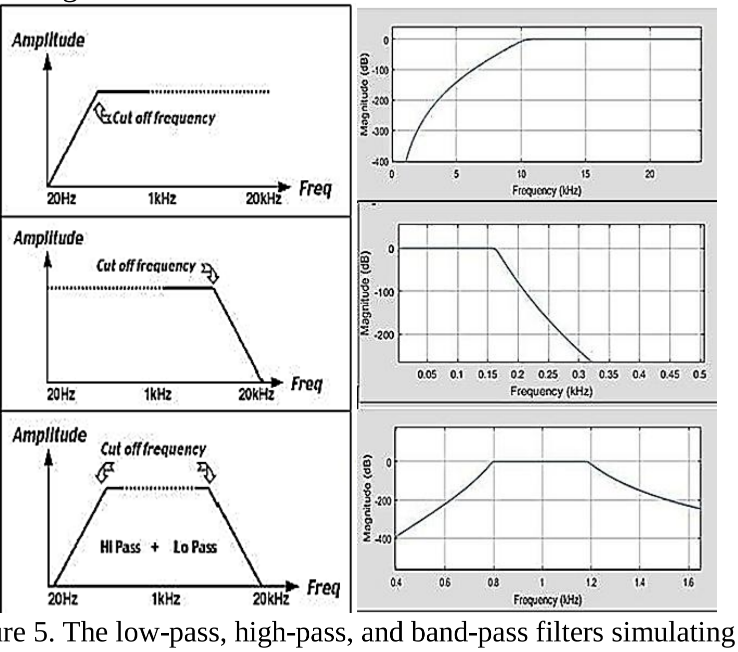

Figure 5. The low-pass, high-pass, and band-pass filters simulating in designs using the FDAtool.

http://www.doi.org/10.62341/amka8005

The Desired Tone Estimation

The power spectral density function can be estimated via the MATLAB pwelch command, which is a part of the Signal Processing Toolbox. The pwelch command, however, is not available in every MATLAB software configuration. Thus, a more fundamental approach is needed. This paper gives the results of source code for calculating the power spectral density using MATLAB based on the Fast Fourier transform (FFT).

The one-sided power spectral density function ($X_{\text{PSD}}(\text{f})$) is calculated from the discrete Fourier transform ($X(f)$) as:

$$X_{\text{PSD}}(\text{f}) = \lim_{\Delta f \to 0} \left[\frac{X(f) * f}{2\Delta f}\right] \qquad (1)$$

Where the frequency step is finite in -this practice, and is the inverse of the total measured duration, and $\Delta f = {}^{1}/_{T}$.

A MATLAB code is created to set each constructor or property name to the specified value, and the constructor arguments are given:

- The number of equalizer bands is specified as an integer in the range 1 to 10 (chosen to be 3 bands).
- The order of the odd equalizer bands is specified as an even integer. All equalizer bands have the same order.
- The Low-pass-Filter must be set during construction and cannot be modified after construction and the low-pass filter cut-off frequency is in Hz, specified as true scale equal to 0.
- The High-pass-Filter must be set during construction and cannot be modified after construction and the high-pass filter cut-off frequency is in Hz, specified as true scale equal to 0.
- The half-band input performs oversampling before equalization where the half-band reducer reduces the sampling rate back to the input sampling rate after equalization. The bandpass filter cut-off frequency is in Hz, specified as a true scale equal to 0.
- The centre frequencies of the equalizer bands are in Hz, specified as a row vector with a length equal to the number of bands, this vector consists of real scales in the range from 0 to sampling rate/2.

http://www.doi.org/10.62341/amka8005

- Equalizer range quality factors, specified as a row vector with length equal to the number of bands. The vector consists of real scales in the range from -1.5 to +1.5. Any values outside the range are applicable.
- Maximum or low filter gains in dB, specified as a row vector with a length equal to the number of bands, this vector consists of real scales in the range from 1 to 40. The values are set to be samples.
- The input sampling rate is in Hz, specified as a positive scale, this property can be changed according to the filename object.
- The filename sampling rate is used as the equalizer object sampling rate and then called the setup to reduce the computational overhead of initialization in the audio stream loop.

To achieve a higher resolution estimation of the actual frequency of a discrete frequency component than that provided by the FFT, a weighted average of the frequencies surrounding a detected peak in the power spectrum is performed. This process involves a range that begins with 13 factors from -1.5 to +1.5, increasing by 0.15 in a loop for 40 iterations and 3 integer values for each sample, thereby determining the estimated frequency.

$$f_x = \frac{\sum_{i=1}^{i=40} X(f(i))*i*\Delta f}{\sum_{j=1}^{j=3} X(f(i))} \tag{2}$$

In this context, where (j) denotes the array index corresponding to the apparent peak of the interested frequency, a span of ($j = 3$) is deemed appropriate as it encompasses a range broader than the main lobes of the audio file samples.

The estimation of power in a specific peak discrete frequency component can be achieved by summing the power in the bins surrounding the peak, effectively computing the area under the peak, and then dividing by the noise power bandwidth

$$.P_x = \frac{\sum_{i=1}^{i=3} P_x(i)}{P_N} \tag{3}$$

Where the noise power bandwidth (P_N) is given by:

$$P_N = 10\ \log_{10}\left(\frac{k.B.T}{1mW}\right) \tag{4}$$

Where;

http://www.doi.org/10.62341/amka8005

$$T = Temperature\ in\ Kelvin/°C$$

$$B = Bandwidth\ in\ Hz\ or\ KHz$$

$$K = 1.38064852 \times 10^{-23}\ Joule/Kelvin$$

The calculations in this work were performed using MATLAB code, and part of the code is shown in Figure 6.

```
for j=1:3;
for i=1:13;
x=(0:12.5:150)/100;
z=(150:-12.5:0)/100;
 GLp(i) = x(i);
 GMp(i) = x(i);
 GHp(i) = x(i);
y_Filtered = GLp(i)*y_Lp + GMp(i)*y_Mp + GHp(i) * y_Hp;
% Choose FFT size and calculate spectrum
Nfft = 4096;
[Pxx1,f1] = pwelch(y,gausswin(Nfft),Nfft/2,Nfft,fs);
[Pxx2,f2] = pwelch(y_Filtered,gausswin(Nfft),Nfft/2,Nfft,fs);
% % Get frequency estimate (spectral peak)
Original_FREQ_ESTIMATED = max(f1);
Original_Power_ESTIMATE = max(Pxx1);
Filtered_FREQ_ESTIMATE = max(f2);
```

Figure 6. A part of MATLAB code for the equalizer.

Results and Discussions

This method, which is based on using a parametric equalizer, applies exclusively to spectra composed of discrete frequency components and is not suitable for continuous spectra. Furthermore, if two or more frequency peaks are found within six separated lines, this can lead to overestimation of power and distortion of the actual frequencies, therefore, it cannot be relied upon to obtain a high degree of similarity between the original spectra and the modified spectra.

Comparing the power spectral density of the resulting female modified signal with the frequency spectrum of a signal emitted by an original male voice, the similarity between the two spectra exceeds 80%, which is considered somewhat acceptable. This extent of similarity between both frequency spectra (original and modified) can be observed from Figure 7.

http://www.doi.org/10.62341/amka8005

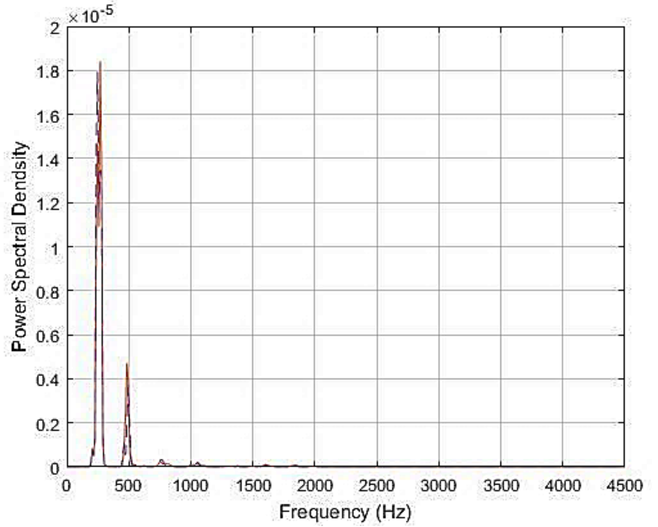

Figure 7. Comparing the similarity between both frequency spectra (original and modified).

Conclusions

This study demonstrated the fundamentals of human vocal intonation modification according to need and usage requirements, where it hypothesized that words spoken with a more emotional tone, as opposed to a neutral tone, attract greater attention and elicit physiological arousal, enhancing memory for concurrent verbal information. Although the current study did not examine any effect of prosody on verbal memory, it indicated the possibility of sustained control over listeners' attitudes toward words, and thus, the words themselves. Words heard by male listeners with a feminine tone were associated more positively than others, while words heard by female listeners with a masculine tone were associated more positively.

This process mitigates the prosodic context, influencing whether an individual will approach or avoid a speech referent in the future. Notably, this occurs independently of the individual's ability to recall the speaker's intonation, suggesting that prosodic moderation of speech value precedes retrieval. Furthermore, since the prosody in the experiments was not task-relevant, this appears to be an implicit process.

http://www.doi.org/10.62341/amka8005

References

[1]Bastos, Wilson, and Sidney J. Levy. 2012. “A History of the Concept of Branding: Practice and Theory,” *Journal of* Historical Research in Marketing, vol. 4, no, 3, pp. 347–68. Howard House: Emerald Group Publishing. Accessed December 11, 2014, DOI:10.1108/17557501211252934

[2] Davis, Melissa. 2009. *The fundamentals of branding*. Switzerland: AVA Publishing Delin, Judy; Abi Searle - Jones and Rob Waller. 2006. https://www.academia.edu/16805438/Tone_of_Voice_in_Information_Design.

[3] Neil Patel, . “*Develop Your Brand’s Tone of Voice*”, Company Founder of NP Digital & Owner of Ubersuggest Published November 4, 2024. https://doi.org/10.1075/ddcs.7

[4] Norman D. Cook. Tone of Voice and Mind. the connections between intonation, emotion, cognition and consciousness, . John Benjamins, 2002. DOI:10.1525/mp.2005.23.1.87

[5] Parametric Audio Equalizer on mathworks.com https://www.mathworks.com/help/ecoder/beaglebone/ref/parametric-audio-equalizer.html?s_tid=srchtitle_site_search_3_Audio%20Equalizer.

[6] RECOMMENDATION ITU-R SF.675-4, “ Calculation of the maximum power density averaged over 4 kHz or 1 MHz(of frequency-modulated and digital carriers)”, 2012.