Audio Engineering Society

Convention e-Brief

Presented at the 144th Convention
2018 May 23–26, Milan, Italy

Sound masking on iOS devices. Masking everyday noises and tinnitus

Lorenzo Rizzi[1], Nadir Bertolasi[1], and Gabriele Ghelfi[1]

[1] *Suonoevita Ingegneria, via Cavour 18, Lecco, Italy*

Correspondence should be addressed to Author (rizzi@suonoevita.it)

ABSTRACT

The research started studying the mobile devices abilities and limits to perform an environmental noise analysis using the internal microphone recording. Perceptual masking algorithms have been implemented to specifically modify natural sounds for the analyzed noise. Then sound pleasantness algorithms have been applied to obtain better masking sounds. This set-up is being extended to the masking of tinnitus: the tinnitus frequency selection is being proposed to the user through sine-sweep listening.

1 Introduction

The aim of our research is developing an application to analyze environmental noise and masking annoying noises and generating relaxing sounds, pleasant to the listener's ear.

In order to examine and extend the known studies, an independent experiment has been carried out. Its aim being that of evaluating how two external microphones and a microphone on a 5s iPhone behave in frequency (this is related to electrical noise, reaction in frequency and sensitivity). The results have been compared to a professional sound pressure level meter data.

We observed how iPhone reacted to energy changes provoked by an artificially produced noise. It has been defined the frequency and sound-level domains that allows observation and a more reliable analysis, with functional data.

A method based on a psychoacoustic model has been employed to generate environment masking sounds on the device itself. We researched into the best algorithm performance allowing an analysis of the incoming signal and its processing in order to obtain synthesis metrics in a reasonable period of time. All this has been implemented natively on iOS by exploiting API Core Audio to obtain more direct and finer control of the device hardware.

The result of the process in an equalization. It can be applied to a bank of natural sounds inserted into the application with the aim to reproduce a masking sound with headphones or external loudspeakers.

The quality of the signal produced is improved without distorting it or compromising its masking properties but keeping it pleasant to one's ear.

Two methods are employed for pleasantness of reproduced sounds improvement, where one is based on the tonality of the disturbing signal controlling a filter in the equalization phase, while the other is based on the gain between adjoining levels allowing to keep the sound nature.

To conclude, in order to evaluate the application efficacy, the application was tested on a sample of eleven people.

2 Smartphones and SPL

To study the potentialities offered by smartphone hardware it is important to realize where relevant results may be expected referred to the hardware limits.
There are studies where smartphones applications and professional SPL Meters are compared. Interesting a 2015 study [1] on a sample of smartphones where 65 are iOS and 35 are Android. The results show a high level of inaccuracies when noise levels under 27 dB (A) are measured. They are generally 5.33 dB (A) and also above 90 dB (A) there is an over-estimation where average deviations of -3.57 dB(A) can be observed.
Android smartphones show larger measurement variation: this is due to the high variety of producers assembling different MEMS and software. It is generally observed better measurement precision on iOS devices.
In order to make a thorough analysis of the results, an independent study has been carried out by the authors. The aim was to understand the energetic estimation capacities related to frequency and to investigate potentialities offered by two smartphone external microphones already on the market (IK Multimedia iRig and mic W I 436) compared to one integrated in a 5s iPhone.

The test has been planned so to artificially generate a white noise at the following LAeq 30 dB(A) , 35 dB(A) , 40 dB(A) , 45 dB(A). For each of the above a measurement has been made using a professional class I SPL meter (Larson and Davis 831) and respective smartphone recordings with each external microphone as well as with the integrated one. Each recording lasted 2 minutes with a variable environment noise of about LAeq = 27 dB(A). Each track has been worked out giving its sound pressure spectrum in third octave bands and calculating differences in decibels along the frequency axis between adjoining levels.

According to this experimentation results it is possible to expect relevant differential measurements when LAeq > = 35 dB (A).
Our analysis allowed to define the frequency range where the devices show results that are coherent with the sound level meter (between 300 Hz and 7000 Hz). External microphones allow to both obtain more stationary estimates compared to the integrated model and lower the frequency limit to 150-200 Hz, yet they fail to improve the estimate capacities for low energy levels.

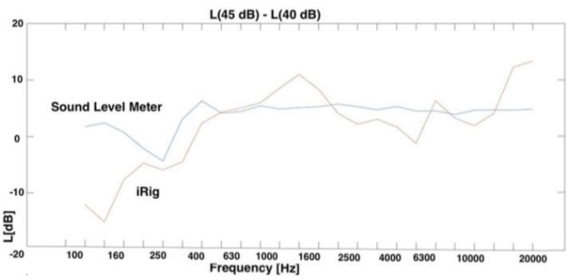

Figure 1. Comparison of Sound level meter and iRig microphone

3 Implementation of the masking algorithm

After the preliminary experiments, the aim of our application has been set to mask annoying noises above LAeq >= 35 dB(A) as well as obtaining privacy for conversation. The purpose will be helping users' studying, helping daytime concentration and obtaining stress reduction.
We chose to adopt an algorithm working on a psychoacoustic model. We proceeded by making improvements to both reduce the computation load and process results within the identified analysis range.
The incoming audio signal is sampled on a buffer with 1024 samples where an FFT is calculated. Then each result is mapped on a Bark Scale in order to carry the incoming signal on a psychoacoustic model divided on 25 bands. Such transformation allows the mapping of frequencies on distances equally perceived by the human ear.

Zv energy is calculated on each band as follows:

$$Z_v = \frac{1}{m}\sum_{k=1}^{m} Z(k) \qquad (1)$$

where v is the Bark index and Z (K) is the partial energy for the frame with m index.
In order to emulate and consider the masking effect between adjoining critical bands a spreading function has been applied calculated with B.M.R. Schroeder's definition.
Then a convolution between the Sij result and the Zv band energy is operated:

$$C_i = S_{ij} * Z_v \qquad (2)$$

We decided to keep memory of the pre-calculated analysis data of masking signals in order to obtain the above- mentioned result in a short period of time and to avoid keeping the user waiting.
A calculation of the masking threshold follows our analysis, meaning the minimum annoying level still perceivable even with a masking sound. The valued found need be remapped in the Bark dominion.
A better method than deconvolution (which proved to be unstable) has been applied: renormalization. Each threshold level for each band is multiplied for the inverse of the Gain applied to the spreading function.
The estimated levels are used to instruct an equalizer which will apply a gain coherent with the range of frequencies in charge of the masking operations.

4 Natural sounds and their pleasantness

Another aspect has been taken in care: that is the production of pleasant sounds inducing a sense of well-being and relaxation in the listener as well as eliminating environment noises.
We noticed, when looking in psychology bibliography, that natural sounds produced by the wind, rain, waves etc. do have beneficial effects on the listener, they can improve one's concentration level, they can also calm those who find themselves in a stressing, busy environment [3, 4].

The choice of sounds offered to the listener includes just tracks of this type plus the hairdryer sound which we found is often used for this purpose.

When working on sound equalization we tried to keep the frequency structure of the sounds used in the masking process so as to improve their pleasant impact on the listener.
We applied two methodologies one of which is based on the Gain Ratio calculus. The factor has been calculated as follows:

$$G_i = {M_i} \Big/ {\frac{1}{24}\sum_{j=1}^{24} M_j} \qquad (3)$$

with Mi being the signal coefficient for the i-esim band. This metrics allows to detect bands where the signal carries a smaller amount of energy: the equalization registers this value in order to avoid applying gains on frequency points that may distort the original sound.

The other methodology is driven by a signal tonality measurement: if this is high we expect to have as a result from the analysis a much higher gain value in one band compared to the others. If this value were applied directly during the equalization phase, the resulting signal would be quite clearly distorted.
This is the reason according to which, if the tonality level is registered as high, a median filter is applied during the equalization phase to smooth the gain level changes and the adjoining bands.

5 App testing

The application was subjected to a first listening test on a sample of eleven people aged between 20 and 50. The listeners were asked to evaluate masking capacities and sound generated pleasantness in two different situations: one with the annoying environment noise of a pellet stove, the other with speech noise. Both tests were conducted with a comparison of a docking station (JBL on stage II) and commercial headphones (Sennheiser HD 598SE) to reproduce masking sounds. In both cases the results show high degrees of satisfaction (4.2 out of 5) referred to the masking efficacy. As reported in *table 1* headphones give a slightly higher effect in masking action, this is due to the fact that sound is applied to ear with no dispersions, focusing on it the action. Indeed, docking station has a directional propagation of sound: masking effects are stronger in these regions, and lower in other areas. This will induce user to be more easily distracted by environment short disturbance events, particularly if he moves around the room.

MEDIUM	**MEAN (1-5)**
Loudspeaker	4.2
Headphones	4.5

Table 1 – Masking evaluation scores by the users

Judgments referred to the generated and listening pleasantness, revealed different opinions showing a high degree of subjectivity although they were never completely negative. According to a prejudice connected with habits and cultural background, listeners tend to consider “odd” a sound that doesn’t come within a musical range. Yet the results are fairly good (3.6 out of 5): they also put into evidence how the methods applied are effective to keep the chosen sounds naturalness. An important aspect that has to be noted is high standard deviation in users’ opinions regarding pleasantness and ‘boredom’ for long hearings, meaning that evaluation of how much a sound is tedious for long hearings is again very subjective.

6 Tinnitus Treatment

Tinnitus is the medical term that defines the perception of a disturbance that is not produced by a source in the environment and it does not exist in the acoustic space where the subject lives. In the literature many different causes are discussed, starting from overexposure to loud music that is a problem very diffused between musicians [5]. Psychologic condition of the person can affect tinnitus disease too: excessive stress or traumatic condition have an impact on its accentuation. The intensity of Tinnitus may vary from individual to individual and can be centered in different frequencies. This phenomenon is caused by the damage of outer air cells and inner air cells that cause generation of problematic areas in the cochlea that generate high intensity electrical pulses to the brain for certain pitches: the human brain reacts causing an excessive neural activity that brings a perceived noise to the subject. In 1953 Heller and Bergam’s study reported that individuals suffering of tinnitus constantly perceive annoyance, but if they are moved to environment with high background noise, the tinnitus is masked by ambient noise which hides the disturbance and brings relief to the subject. This study highlights the personal and psychologic nature of Tinnitus and indicates that frequency masking techniques can influence its perception. Article [7] compares different therapies for tinnitus treatment, starting from Masking (Vernon, 1977). The sense of relief is introduced by the use of energetic masking of disturbance frequencies that inhibits neural activity and acts totally removing disturbance or only partially hiding it (Partial Masking). This technique produces an important outcome which is Residual Inhibition: clinical test demonstrates that broadband noise can have an influence on tinnitus hiding for minutes.

Sound therapy is one of the most recent techniques: it uses sounds to alter the perception of Tinnitus having real clinical effects on individuals. In this article the use of natural sounds is proposed as sound therapy in order to produce relief and relax to the subject. An important phase to reach this aim is to execute pitch match in order to understand which is the frequency where individuals feel pain. As highlighted in [5] one of most recent technique is to use three consecutive pitch matches: the user is asked to indicate the frequency where most of the pain is concentrated. The application reproduces three times a slow sweep (chirp) signal between 500Hz and 1500Hz. The user is asked to press stop button each time he feels tinnitus reaching higher intensity. Then a resultant value is obtained by averaging these results. It is important to report that although this method is one of the most precise reported in literature, often it has a limitation due to insufficient ear training of the subject. The Article

[5] suggests that higher precision can be reached by use of a training program of the individual in order to coach him how to recognize pitch and be more precise in individuation of the disturbance.

The pitch individuated with this technique is then used to instruct masking algorithm proposed in this article: it is simulated the presence of a tonal disturbance bypassing environment analysis phase. The frequency set by the user through pitch matching is reported in bark scale, in order to evaluate the ear reason in the cochlea where neural activity provides most of the disturbance felt by the patient.
Masking algorithm hence propose a natural sound with a resulting equalization which acts in masking the tinnitus and bringing relief to the individual. It is also important to consider psychological influence of this method: as reported in this section, one of the causes of Tinnitus is stress and disease suffered by the subject, which many times is influent in Tinnitus emphasizing. Natural sounds and pleasantness techniques described in this article address this issues with the objective to produce the best sense of disturbance relief and relax the user.

7 Conclusions

Our study of iOS devices allowed to single out the limits connected with the hardware. Such limits made us reconsider the frequency domain and the chances to detect noises with LAeq inferior to 35 dB(A). As a consequence, the areas of both analysis and use where the application can obtain maximum efficaciousness, have been chosen. Furthermore, areas of interest with disturbances centered in range of frequency below 150 Hz can be excluded. In order to have access to fine grained and deep control of hardware, iOS has been chosen as the development platform.
According to the algorithm the above-mentioned hardware limits have been considered to find a good ratio between calculation times and masking efficacy.
When tested the application, the eleven listeners reported generally positive evaluations, but a high grade of subjectivity in evaluation of sound pleasantness is to be taken in consideration. The study investigates also Tinnitus relief treatment with partial masking techniques: the same algorithm for annoyance hiding can be used bypassing the analysis phase to produce an equalization that is based on user-given pitch frequency.

Future improvement may be obtained by making the algorithm adaptable and by allowing very long listening to produce more pleasant sounds.

References

[1] Enda Murphy Eoin A. King. *Testing the accuracy of smartphones and sound level meter applications for measuring environmental noise.* Applied Acoustics, 106(16), 2015.

[2] B M. R. Schroeder B. Atal and J. L. Hall. *Optimizing digital speech coders by exploiting masking properties of the human ear.* J. Acoust. Soc. Am., 66(6):1647–1652, Dec 1979.

[3] Elliott Salamon Minsun Kim John Beaulieu George B Stefano. *Sound therapy induced relaxation: down regulating stress processes and pathologies.* Med Sci Monit :, 9(5)(RA96-0), 2003

[4] Laura J. Loewen Peter Suedfeld. *Cognitive and arousal effects of masking office noise*. Sage, 24(3), 1992. Laboratory Number: 602

[5] Jaime Serquera, Winfried Schlee, 2,Rüdiger Pryss, Patrick Neff And Berthold Langguth *Music Technology For Tinnitus Treatment Within Tinnet* . AES 58th International Conference, Aalborg, Denmark, 2015 June 28–3

[6] Stamatia Staikoudi, *Contemporary theories of tinnitus generation, diagnosis and management practices*. AES 128th Convention, London, UK, 2010 May 22–25

[7] Derek J. Hoare, Grant D. Searchfield, Amr El Rafaie, James A. Henry. *Sound Therapy for Tinnitus Management: Practicable Options*. Journal of the American Academy of Audiology/Volume 25, Number 1, 2014