RESEARCH ARTICLE OPEN ACCESS

Snorm–A Prototype for Increasing Audio File Stepwise Normalization

[1]Sachin M. Narangale, [2]Prof. Dr. G. N. Shinde
[1]School of Media Studies, Swami Ramanand Teerth Marathwada University, Nanded-431606, India
[2]Pro-Vice Chancellor, Swami Ramanand Teerth Marathwada University, Nanded-431606, India

ABSTRACT
This paper introduces a novel concept SNORM (Step NORMalization) for increasing normalization. It is a prototype algorithm for increasing normalization based on loudness factor of the audio. The function effect normalization plays a vital role in loudness control. The proposed experiment carried out for increasing normalization based on step wise increase yield the variations in peaks of the audio file. The experimental results are shown in the form of graphical analysis of the plot spectrum values of frequency analysis. From the results, it is clearly apparent that, the normalization values are increased at different levels. The function SNORM can set a new benchmark in the field of audio industry for the processes of increasing normalization.
SNORM can be substantial in the audio broadcast systems for applications in live audio streaming, news broadcast, sports coverage, live programming where the loudness control mechanism is essential. For the selective or predictive loudness control systems SNORM can be effectively applied.
Keywords: *Normalization, Loudness Control, Track normalization, Audacity, Audio Retrieval, Text to Speech application, Frequency Analysis*

I. INTRODUCTION

The retrieval systems for audio are major game players in the music industry. The content-based audio information retrieval system (AIRS) for different electronic equipments is creating lot of attention. AIRS defined properly will definitely perform various retrieval operations including recognition, referencing and recommendation. Content-based music information retrieval (MIR) systems such as Shazam, SoundHound, and Gracenote have already been developed for the iPhone, iPad, and other similar Smartphone devices [1]. Audio retrieval is a mechanism that searches music that is being played. The audio identification and retrieval system has to identify the different background noises. For building AIRS, audio tokenization or unique identification is necessary. These unique identification measures act as the identity of the audio to be retrieved from the database. This audio unique identification measure contains brief details of the audio file or sometimes a frame of audio too. The growth of the audio industry has invariably led to the demand for quick and correct audio data retrieval. The audio retrieval system resulting lot of information none of the use is more. The audio retrieval rates need to be improved. To improve the retrieval rate, before the query is to be fired for the audio, first, it is essential to create a audio lookup reference table (ALURT). For audio to be matched the audio unique identification is maintained as the private key of retrieval. Due to this the audio retrieval probability of intended audio increases quantitatively. In this research, a mechanism in the form of prototype pre-processing method for retrieving the intended audio from ALURT is presented. The novel proposed mechanism SNORM works on the principle of normalization. A study has been evaluated for increasing normalization stepwise to counter the normalization values of the audio file.

Normalization is a process of reducing peak amplitude of the audio signal to a defined intended level or to the corresponding average of the frame of an audio signal. Normalization is the process of modification of the amount of gain or amplitude of the audio signal. The peak amplitude is reduced to a target level in normalization. Audio Normalization is a pre-process of the system used in audio compression [2]. Another effective use of normalization can be seen in the process of loudness control. Normalization works out to smooth out the variations in the loudness. Peak control and loudness control can be achieved using normalization [3].

Normalization is an effect provided in the Audacity to guarantee the complete representation of every audio feature element. The process involves the methodology of subtracting the mean of audio feature and the resultant value is divided by its standard deviation [4].

Various application areas of the normalization process are:

- Process of loudness control
- Smooth out the variations in the loudness

- To guarantee the complete representation of every audio feature element
- To check audio compatibility with the reference signal
- For every audio input track, the speaker may not speak at an expected level
- In the research of text to speech mechanism, audio has to be normalized
- Different channel surfing over television network if doesn't have audio normalization done, the audience may get disturbed by the sudden increase in loudness.

II. RELATED WORK

Normalization has been carried out on attacked audio samples before the watermark extraction to be compatible with the reference signal [5]. Reference signals are the audio reference signals which can be considered while ALURT preparation. For the listening tests, the normalization of the gains of the audio is necessary. The auralized samples undergo normalization process on the equivalent level for the listening test by altering gains. The increase in graphical computing power and the system capacity is demanding increased aspiration to achieve higher visual reliability in non-real world environments and the video games. As the visual fidelity perception is getting more attention, the role of audio in multimedia is getting inattentiveness. But the question is, whether the visual fidelity is sufficient for higher reliability? No, the audio perception is equally important in the reliability. To achieve the attention in the visual fidelity, audio needs to be normalized at particular levels [6]. For live broadcast systems, for example news broadcast, every time the speaker may not be a perfect speaker with monotonous maintained audio levels. Sometimes, the environmental factors may generate some noise that increases the audio signals all of a sudden, disturbing the continuity of the viewer. Similarly, in case of different channel viewing over television network, some channel may not control the loudness of the audio signals. The audio loudness control mechanism through normalization is essential.

In the case of text-to-speech paradigm, the text normalization based on textual grammar and pronunciation is needed. The audio generation based on the text reading has to have the audio normalization pre-processing methodology for loudness control [7].

The classification systems to become robust to loudness and channel changes like CD channel to telephone channel use normalization [8].

III. EXPERIMENTAL SETUP

This research includes various experiments executed on open source freeware Ubuntu, a Linux desktop operating system. For audio recording and editing an open source, free and cross-platform software Audacity was used. Audacity gets perfectly installed and configured with the Ubuntu. The features of Audacity in this research are used for prototype designing related to normalization. For GUI between Audacity and Ubuntu a C++ library wxWidgets was used. wxWidgets facilitates developers with a single code base for creating cross-platform applications for Windows, Mac OS X, Linux and other platforms.

A desktop system with hardware configuration 2GB RAM, i3 processor, 3.08GHz and 350GB HDD was used for experiments. Experiments started with recording a sample audio file "hello123.mp3" using Audacity recording feature. Normalization experiments were conducted on "hello123.mp3". Different stepwise normalization parameters were incorporated in the code file "Normalize.cpp". The frequency analysis spectrum plots were captured using screen capturing tools and presented in following figures. Plot spectrum analysis values were analyzed against the frequency values in squared form. The comparative analysis chart has been presented in this paper below.

Snorm Algorithm

Step 1: Start
Step 2: ProcessFirst – take a track
Step 3: Transform track into bunch of buffer blocks
Step 4: Get the offset value and ratio parameters set.
Step 5: SNORM function
Step 6: SNORM Step 2,4
Step 7: Adjust frames based on step parameters and offset values.
Step 8: Analyze Track
Step 9: Analyze Data
Step 10: Return

IV. RESULTS AND DISCUSSION

This research presents the results of executing SNORM stepwise function for increasing gain of audio file by molding the effect normalization in Audacity. For analyzing the results the graphical chart analysis has been used. The representation of increased gain with SNORM has been depicted in the figure 5.

Figure (1) is the graphical Plot spectrum of the audio file "hello123.mp3" without any editions to it. The frequency analysis plot spectrum

ISSN : 2248-9622, Vol. 6, Issue 3, (Part -4) March 2016, pp.171-174

has been shown in this figure. The audio file “hello123.mp3” is a sample audio file created using recording feature of Audacity.

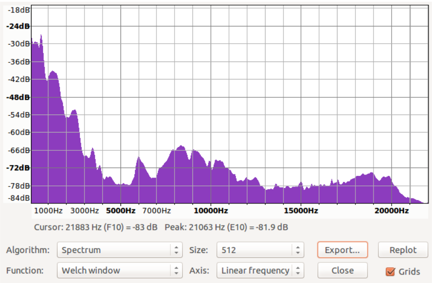


Figure 1: Plot spectrum of the Original audio file “hello123.mp3” – frequency analysis

Figure (2) is the graphical Plot spectrum of the audio file “hello123.mp3” after the normalization process has altered the gain of the audio file. Normalization as stated earlier, is a process of altering peak values.

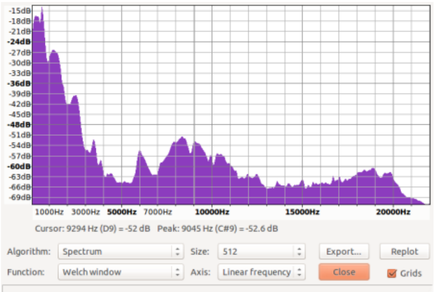


Figure 2: Plot spectrum of the Normalized audio file “hello123.mp3” – frequency analysis

Figure (3) is the graphical Plot spectrum of the audio file “hello123.mp3” after the normalization process has altered the gain of the audio file up to step 2. SNORM is executed on the audio file. This figure, clearly it is seen that, the gain values has been altered. The amplitude values are changed. The SNORM has increased the normalization of audio files.

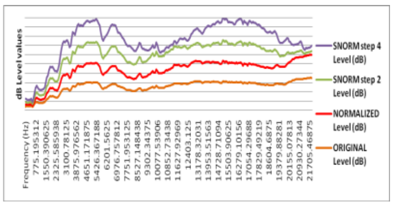


Figure 5 Graphical Analysis of Original audio file and SNORM applied files

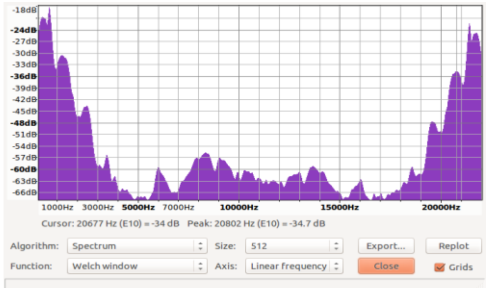


Figure 3: Plot spectrum of the SNORM step 2 Normalized audio file “hello123.mp3” – frequency analysis

Figure (4) is the graphical Plot spectrum of the audio file “hello123.mp3” after the normalization process has altered the gain of the audio file up to step 4. SNORM is executed on the audio file. This figure, clearly it is seen that, the gain values has been altered. The amplitude values are changed. The SNORM has increased the normalization of audio files. The increase in the peaks or amplitude is clearly seen. The objective of step 4 normalization is achieved. The frequency analysis can be identified as the symmetrical representation in this figure.

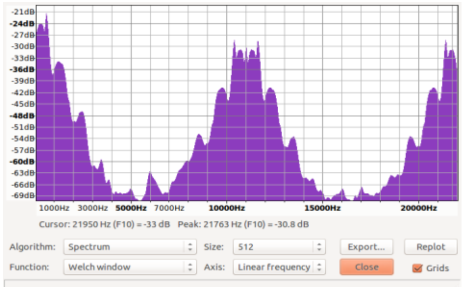
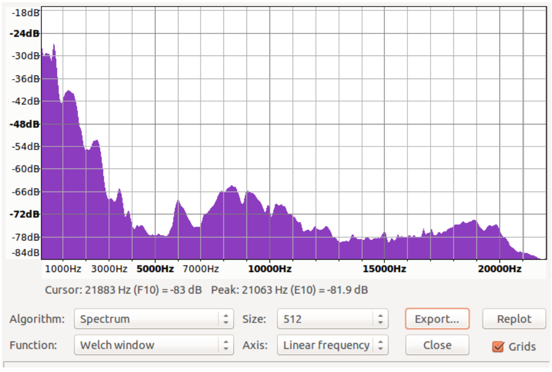


Figure 4: Plot spectrum of the SNORM step 4 Normalized audio file “hello123.mp3” – frequency analysis



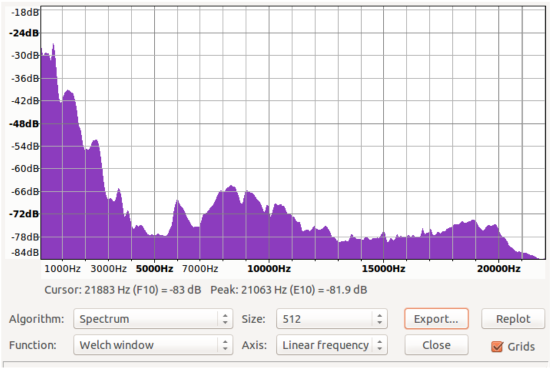


Figure (5) is the graphical waveform presentation of the original audio file "hello123.mp3" and after the normalization process has altered the gain of the audio file up to step 4. The proposed prototype SNORM with step 2 and step 4 are resulting increase in the normalization values of the audio at an average of 101 and 98 percentages respectively.

V. CONCLUSION

This research emphasizes the concept of increasing normalization for smoothing out the variations in gain or loudness. The function effect normalization plays a vital role in loudness control. The proposed experiments carried out for increasing normalization based on step wise increase yield the variations in peaks of the audio file. SNORM is a novel prototype algorithm for increased normalization based on loudness factor of the audio. The proposed prototype SNORM with step 2 and step 4 are resulting increase in the normalization values of the audio at an average of 101 and 98 percentages respectively. SNORM can be substantial in the audio broadcast systems for applications in live audio streaming, news broadcast, sports coverage, live programming where the loudness control mechanism is essential. For the selective or predictive loudness control systems SNORM can be effectively applied.

REFERENCES

[1]. Dae-Jin Kim, Ddeo-Ol-Ra Koo, "Analysis of Pre-Processing Methods for Music Information Retrieval in Noisy Environments using Mobile Devices", International Journal of Contents, Vol.8, No.2, Jun 2012

[2]. Zainab T. Drweesh, Loay E. George, "Audio Compression Based on Discrete Cosine Transform, Run Length and High Order Shift Encoding", International Journal of Engineering and Innovative Technology (IJEIT), ISSN: 2277-3754, ISO 9001:2008 Certified, Volume 4, Issue 1, July 2014

[3]. Manik Gupta, Jozsef Pinter, "Loudness Measurement and Control", Proceedings of NAB Broadcast Engineering Conference, 2012, pp.195-200

[4]. M. Liu, C. Wan, L. Wang, "Content-based audio classification and retrieval using a fuzzy logic system: towards multimedia search engines", Soft Computing 6 (2002) 357 – 364 Springer-Verlag, 2002, 10.1007/ s00500-002-0189-3, pp.357-364

[5]. Janusz Cichowski, Andrzej Czyżewski, Bożena Kostek, "Analysis of impact of audio modifications on the robustness of watermark for non-blind architecture, Multimed Tools Appl (2015) 74:4415–4435 Springerlink.com

[6]. David Rojas, Bill Kapralos, Andrew Hogue, Karen Collins, Lennart Nacke, Sayra Cristancho, Cristina Conati, Adam Dubrowski, "The Effect of Sound on Visual Fidelity Perception in Stereoscopic 3-D", IEEE Transactions On Cybernetics, 2013, pp.1-12

[7]. Arun Soman, Sachin Kumar S., Hemanth V. K., M. Sabarimalai Manikandan, K. P. Soman, "Corpus Driven Malayalam Text-to-Speech Synthesis for Interactive Voice Response System", International Journal of Computer Applications (0975 – 8887), Volume 29– No.4, September 2011, pp.41-46

[8]. Hadi Harb, Liming Chen, "A general audio classifier based on human perception motivated model", Multimed Tools Appl (2007) 34:375–395 DOI 10.1007/s11042-007-0108-9 Springer Science