Fuzzy Equalization of Musical Genres

Marie González
Department of
Electrical Engineering
School of Engineering
Pontificia Universidad Católica
de Chile
mrgonzalez@uc.cl

Patricio de la Cuadra
Center for Research in
Audio Technologies
Music Institute
Pontificia Universidad Católica
de Chile
pcuadra@uc.cl

Rodrigo F. Cadiz
Center for Research in
Audio Technologies
Music Institute
Pontificia Universidad Católica
de Chile
rcadiz@uc.cl

ABSTRACT

In this article, we propose an audio equalizer of musical genres based on fuzzy logic. Widely used audio playback software, such as VLC or iTunes, propose genre-specific equalization curves to be applied for the whole duration of the music. These curves are the same for all songs belonging to a specific genre, and they do not take into account the specifics and diversity of each song. We propose a different strategy. Research in music information retrieval has revealed a significant number of audio descriptors that allow for the recognition and description of diverse musical genres. We use some of these descriptors to feed a fuzzy logic inference system, whose outputs are the required equalization levels for each frequency band. The rules of the system are derived from the analysis of a well known music database encompassing ten different musical genres. Our results indicate that our approach works for songs that exhibit multiple genre characteristics, that are difficult to classify into one category, or that mix genres.

1. INTRODUCTION

Audio equalization using digital signal processors has been a subject of research for several decades. An equalizer aims for the correction of the magnitude and the phase response of an audio chain [1]. The equalization of audio is required in a great number of distinct situations such as radio broadcasting, recording studios, or music listening while driving or at home.

Equalization is useful to correct for uneven frequency responses of sound reproduction systems, which essentially means static equalization curves for all genres. Consequently, it appears that equalization based on genre is unnecessary and not even desirable. This argument strengthens considering all the work and effort that mastering engineers have put into balancing the audio spectrum in the studio. However, we would like to emphasize that genre-based equalization is a common reality. Car radios and home audio playback systems contain this kind of equalization. Very popular audio software such as VLC or iTunes base their equalization curves on genre information, as it is shown in figure 1. We simply want to propose a different strategy for genre equalization and we don't have the intention at this moment of arguing against or in favor of it.

These curves, although somewhat similar in different software, appear to be based mostly on experience and customized by each vendor, and we have not found scientific or widely accepted methodologies for their design. Similar curves are also present in the majority of audio reproduction equipments and car radios. Each one of these curves aims for a specific musical genre such as rock, pop, jazz or electronic. The user is supposed to select one of them when listening to a particular piece of music. However, what curves should be used for music that mix several styles, songs that contain a significant amount of musical diversity, or when the genre one is listening to does not appear in the lists of available equalizations? In order to tackle this problem, we propose a fuzzy logic-based equalization scheme that allows for a more meaningful equalization of music, including mixed musical genres or genres outside the realm of traditional music.

This article is structured as follows. First, fuzzy logic is presented as discussed as a valid technique for non-linear mappings of data. Next, we present our fuzzy logic-based audio equalizer including previous analysis of the audio data, determination of fuzzy inference rules, output variables and computational implementation. In the following section, we describe one simple experiment we conducted in order to test our equalization scheme and later the main results are presented. Finally, conclusions and future work are addressed.

2. FUZZY LOGIC

Fuzzy logic [2] [3] [4] [5] [6] is a concept derived from the mathematical branch of fuzzy sets [7] that applies *multi-valued logic* to sets or groups of objects. In a narrow sense, fuzzy logic refers to a logical system than generalizes traditional two-valued logic for reasoning under uncertainty, allowing multiple values of truth. In a broad sense, it refers to all the

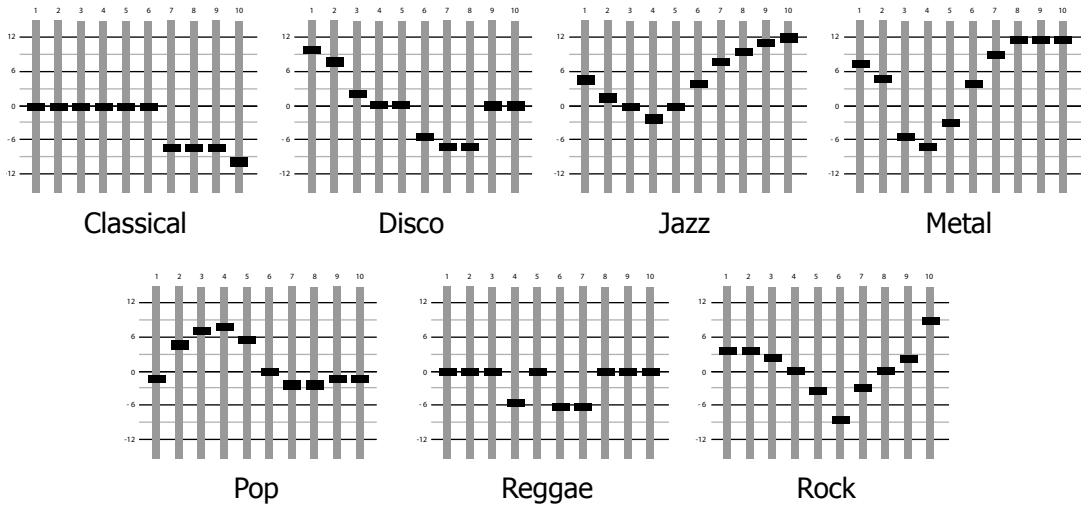

Figure 1. Equalization curves as proposed by the VLC media player for the musical genres Classic, Disco, Jazz, Metal,Pop, Reggae and Rock. The EQ has 10 bands, centered on 60, 170, 310, 600, 1000, 3000, 6000, 12000, 14000 and 16000 Hz. Pictures shows the gains on dB in a range from -12 to 12 dB.

theories and technologies that employ fuzzy sets [8]. In general, when fuzzy logic is applied to computers, it allows them to emulate the human reasoning process, quantify imprecise information and make decisions based on vague and incomplete data [5].

Fuzzy systems are powerful and work in a way that resembles some characteristics of human behavior. Parallel computation of fuzzy rules reduces drastically the computation time compared to a traditional mathematical approach. Fuzzy systems allow approximation of highly nonlinear systems with high accuracy. It is not necessary to know any mathematical model in advance to approximate *any* system [5]. Fuzzy logic allows us to build systems using common sense, and the fuzzy rules can be discussed, tuned, and detuned easily. These facts makes fuzzy logic a very appropriate tool to study and model non-linear dynamical systems and to handle complexity.Fuzzy logic systems have been widely used in engineering and control applications [4] [5] , and also in several artistic and musical applications such as sound synthesis [9], polyphonic note extraction [10], digital restoration [11], musical decisions [12], and audio-visual composition [13].

3. GENRE DESCRIPTORS

Some authors have proposed the usage of audio descriptors to identify different musical genres [14–16]. [16] shows timbre descriptors like a useful way to classify genres, so we use these descriptors to detect when a song belongs to one genre or another. Based on this knowledge, we decided to use the MIR toolbox, a very complete musical information retrieval analysis tool for MATLAB [1] The descriptors we used were zerocross, centroid, low energy, rolloff, flux and MFCC coefficients [17].

We used the GTZAN Genre Collection [2] , a well known database in the Music Information Retrieval community. It contains one hundred audio tracks of thirty seconds long of each of the following genres: Blues, Classical, Country, Disco, Hiphop, Jazz, Metal, Pop, Reggae and Rock. Ninety tracks per genre were select to be analyzed and the other ten were reserved for testing.

We checked the usability of these descriptors for genre classification. We conducted analysis of variance (ANOVA) tests for every descriptor. This statistical test determines significant differences in groups of data [18]. It estimates the variance on a normal distribution for each group and provides a parameter p that represents the probability that two or more sets belong to the same group. A value less than 0.01 determines that groups are independent sets. For each genre, 90 songs were selected and their audio descriptors were computed. The highest p-value obtained was 0,000239. This means that the eighteen chosen descriptors have a small probability of belonging to the same group, and we than deduct that these genres could be indeed characterized by the chosen parameters.

4. AUDIO EQUALIZATION BASED ON FUZZY LOGIC

We propose an equalization system as the one described in figure 2. We start by computing the eighteen audio descriptors of section 3 in a window of 30 seconds, using the MIR toolbox [17]. As previously discussed, these descriptors are supposed to represent specific aspects of different musical genres and styles. We then fuzzyfy these descriptors, and use them as inputs for a fuzzy logic inference engine. Using several decision rules derived from the descriptor analyses (described in section 4.3), the fuzzy system determines

[1] `http://www.mathworks.com`

[2] available from `http://marsyas.info/downloads/datasets.html`

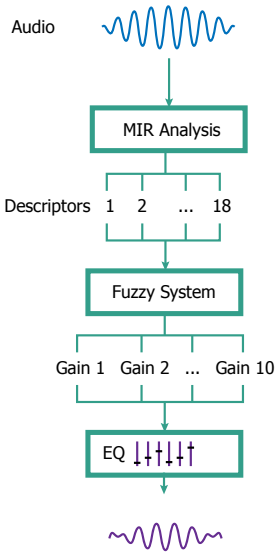

Figure 2. An incoming music signal is analyzed using the MIR toolbox resulting in the calculation of 18 genre descriptors, which are later used as the inputs for the fuzzy system. The fuzzy outputs correspond to the gains of a 10-band audio equalizer.

the ratios in which the inputs represent each one of seven predetermined musical styles, We only use seven genres, which are the ones that VLC provides information for; these are: Classical, Disco, Jazz, Metal, Pop, Reggae and Rock. The outputs of the system are obtained by defuzzycation and represent the gains of a 10-band equalizer, proposing a customized equalization curve.

4.1 Fuzzy Inputs

The eighteen descriptors are the inputs of the fuzzy system. All of them were described as fuzzy sets with the following five membership functions: Very Low (VL), Low (L), Medium (M), High (H) and Very High (VH). VL sets were described by a zmf function, VH with a smf and other sets with gaussmf [3].

4.2 Fuzzy Outputs

The outputs of the fuzzy system represent the intensity levels of an 10-band equalization curve, such as the one used in VLC. We chose this software for comparison purposes. The bands are centered around 60, 170, 310, 600, 1000, 3000, 6000, 12000, 14000 and 16000 Hz. The gains are calculated in dBVU scale, ranging from -12dB to 12dB. VLC does not provide information about the bandwidth of each band. However, this does not affect our study, as we just use the gains provided by VLC as scalars, thus we do not use their equalizer. We use their gains as the reference, and we calculate different ones with our fuzzy system, and we apply both sets of gain with the same equalizer, described in section 4.4, to the audio database. We can then compare under similar conditions the effect of both approaches.

Every output was defined as a fuzzy set with nine triangular membership functions, as it is customary in the design of fuzzy systems. We treated each frequency band of the equalizer as equals, as we have no access to VLC's internal equalizer equations in order to determine if there are significant changes for each band.

4.3 Fuzzy Rules

In order to determine appropriate rules for out system, the value of every audio descriptor was computed for each of the 900 training songs. For each genre and descriptor the first three percentiles were calculated and translated into their corresponding fuzzy memberships. We then obtain three rules to be associated to each genre: the inputs are fuzzy forms of the three percentiles for each descriptor.

At the same time, for every detected genre, we associated a corresponding equalization curve. As discussed in section 4.2, we simply used the gains provided by VLC, as shown in figure 1. It is important to mention that these curves are very similar to the ones available on others media players such as iTunes or Windows Media Player, and, consequently, we think that it constitutes a good standard for comparison.

In order to give a higher relevance to the average representation of the values, rules calculated with the second percentile have a weight of 0.5 and the other ones have a weight of 0.25.

4.4 Implementation

The MIR toolbox for Matlab was used to obtain the parameters discussed in section 4.1 from an audio signal. As described in figure 2, these correspond to the inputs for the fuzzy system, which calculates gains for each band. We used the audioplayer equalizer [4], a set of very customizable MATLAB routines, to apply the proposed equalization curve. In the final stage of our process, the final audio output is calculated and saved as an audiofile.

5. EXPERIMENTS AND RESULTS

To test the proposed system, we considered the song *Bicycle Race* from the group *Queen*, that was tagged as belonging to different genres (Rock, Pop) in the website *last.fm* [5].

Figure 3 shows the proposed fuzzy curve for this song. The first four frequency bands appear to be a mixture of rock and pop, while in the middle and high bands the proposed curve is more similar to the pop curve than to the rock one.

[3] zmf,smf anf gaussmf are typical functions described on Matlab and used for fuzzy systems

[4] Copyright (c) 2012, Maxim Vedenyov. Downloaded for http://www.mathworks.com/matlabcentral/fileexchange/34739-equalizer–audioplayer–gui

[5] `http://www.last.fm/music/Queen/_/Bicycle+Race/+tags`

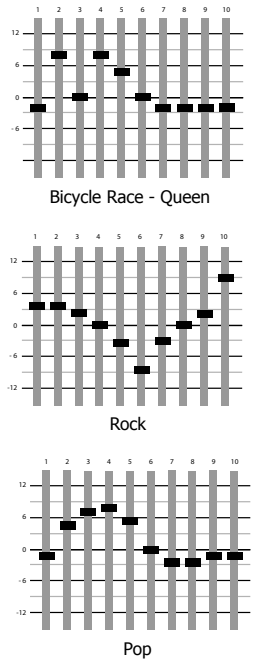

Figure 3. Fuzzy equalization curve proposed by our system for the song *Bicycle Race* by *Queen*. As this song has been tagged both as rock and pop, it should not be equalized by a curve designed only for one specific genre, but with a curve that captures characteristics of both genres. The proposed curve, at the top, is a mixture of the specific genre curves for rock and pop.

6. CONCLUSIONS AND FUTURE WORK

Many media players, such as VLC, offer the option to equalize the audio stream, and the user must choose from a pre determined set of equalization curves associated to the music genre of the song being listened to. This is acceptable for a song that belongs very clearly to a particular genre, but what happens when a song has features identified with more than one style?

Our fuzzy equalizer comes as a way to decide how to automatically build an EQ curve for music with different styles, based on several decision rules. In traditional equalization, genre-specific equalization curves are used for particular piece of music. The usage of a genre-derived, fixed equalization curve does not work well for music that mix several styles or when the genre being listened does not appear in the lists of available equalizations. In order to tackle this problem, we propose a fuzzy logic-based equalization scheme that allows for a more meaningful equalization for mixed musical genres or for genres outside the realm of traditional music.

Our results indicate that the derivation of cross-genre equalization curves is not a straightforward process, and that nonlinear operations on the typical genre curves, as shown in figure 3, must be applied in order to find a genre-based equalization. These curves are very interesting and unique to each audio track.

Acknowledgments

This project was funded by a undergraduate research grant (Winter 2014) of Vicerrectoría de Investigación, Pontificia Universidad Católica de Chile

7. REFERENCES

[1] B. Bank, "Perceptually Motivated Audio Equalization Using Fixed-Pole Parallel Second-Order Filters," *Signal Processing Letters, IEEE*, vol. 15, pp. 477–480, 2008.

[2] H. Bandemer and S. Gottwald, *Fuzzy sets, fuzzy logic, fuzzy methods with applications*. Chichester ; New York: J. Wiley, 1995.

[3] E. Cox, *The fuzzy systems handbook : a practitioner's guide to building and maintaining fuzzy systems*. Boston: AP Professional, 1994.

[4] G. J. Klir and B. Yuan, *Fuzzy sets and fuzzy logic : theory and applications*. Upper Saddle River N J: Prentice Hall PTR, 1995.

[5] B. Kosko, *Fuzzy Thinking. The new science of fuzzy logic*. New York: Hyperion, 1993.

[6] D. McNeill and P. Freiberger, *Fuzzy logic*. New York: Simon & Schuster, 1993.

[7] L. A. Zadeh, "Fuzzy Sets," *Information and Control*, vol. 8, pp. 338–353, 1965.

[8] J. Yen and R. Langari, *Fuzzy logic : intelligence, control, and information*. Upper Saddle River, N.J.: Prentice Hall, 1999.

[9] R. Cádiz and G. S. Kendall, "A particle-based fuzzy logic approach to sound synthesis," in *Proceedings of the Conference on Interdisciplinary Musicology*, Montreal, Canada, 2005.

[10] G. Monti and M. Sandler, "Automatic Polyphonic Piano Note Extraction Using Fuzzy Logic In A Blackboard System," in *5th International Conference on Digital Audio Effects*, 2002.

[11] M. R. Civanlar, "Digital Restoration Using Fuzzy Sets," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSF-34, no. 4, August 1986.

[12] P. Elsea, "Fuzzy logic and Musical Decisions," 1995, http://arts.ucsc.edu/EMS/Music/research/FuzzyLogicTutor/FuzzyTut.html.

[13] R. Cádiz, "A Fuzzy-Logic Mapper for Audiovisual Media," *Comput. Music J.*, vol. 30, no. 1, pp. 67–82, 2006.

[14] B. L. Sturm, "Two Systems for Automatic Music Genre Recognition: What Are They Really Recognizing?" in *Proceedings of the Second International ACM Workshop on Music Information Retrieval with User-centered and Multimodal Strategies*, ser. MIRUM '12. New York, NY, USA: ACM, 2012, pp. 69–74. [Online]. Available: http://doi.acm.org/10.1145/2390848.2390866

[15] J. Bergstra and B. Kgl, "Aggregate features and AdaBoost for music classification," in *Machine Learning*, 2006, pp. 2–3.

[16] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 5, pp. 293–302, 2002.

[17] Matlab, *MIR*. Prentice Hall PTR, 1995.

[18] R. G. Miller and Jr., *Beyond ANOVA: Basics of Applied Statistics*. CRC Press, 1997. [Online]. Available: https://books.google.com/books?id=nz241IjmSGgC&pgis=1