ADDING THE ROOM TO THE MIX: PERCEPTUAL ASPECTS OF MODAL RESONANCE IN LIVE AUDIO

Carlo Bolla and Alessandro Palladini
Music Tribe UK
`carlo.bolla,alessandro.palladini`
`@musictribe.com`

Bruno M. Fazenda
Acoustics Research Centre
University of Salford
Salford, UK
`b.m.fazenda@salford.ac.uk`

ABSTRACT

The problem of room acoustics correction in live sound is still an open one. In particular, the audible artefacts caused by low frequency resonances are still a major factor in determining the perceived quality of a live show.

Despite many years of research into room acoustics correction, very little research has been done into how this influences the sound at large venues and the difficulties this causes to the working practices of live sound engineers.

In this work, we show how perceptual models of modal resonance in rooms can be applied to designing a novel room response analysis tool that can be used to intelligently guide the mix process and to design music production tools that facilitate the work of sound engineers.

1. INTRODUCTION

Live audio is often challenged by the acoustic conditions of modern concert venues, which are typically not designed with optimal sound reproduction in mind. Despite the evolution of technology for live sound, the acoustic conditions of a venue are still a significant (if not the main) cause of poor sound quality.

In the last 40 years, extensive research has been done in the area of room response equalization. Sound system technicians can now employ a wide set of tools to perform the tuning of a reproduction system with the purpose of improving the sound quality [1]. A typical approach is to employ equalization techniques based on one or more measurements of the room impulse response (RIR) [2].

Despite this, very little attention has been given to understanding how the listening conditions of a venue influence and challenge the creative process of live sound mixing. Moreover, most of the measurement and corrective tools often provide poor or no insights about the psychoacoustics of the listening environment and about the proposed corrections, making the setup of system equalization often an empirical and complex process.

In this work, we show how we can take into account perceptual models of modal resonance to design novel RIR analysis tools able to highlight the key perceptual issues in the low frequency range. Such tools, can be used by system technicians to aid the system correction, by sound engineers to guide the mix process, and by intelligent music production tools to suggest or automate acoustically informed mixing decisions. Such tools, in particular when automation is employed, can drastically reduce the effort required from the user, avoiding fatigue and releasing more time for creative delivery.

2. ROOM ACOUSTICS AND PSYCHOACOUSTICS OF MODAL RESONANCE IN LIVE SOUND

The characteristics of a room response are best analysed in the time-frequency domain, where objective assessment of the measured signal can be related to acoustic properties of the environment and how they influence human perception of sound.

An impulse response, obtained from a sound source in a specific position of a real environment, can be divided into three parts [3]: direct sound, early reflections, and late reflections. Direct sound and early reflections contribute to the localisation of the sound source and also to the perception of its timbre [4], while the late reverberation provides cues on the dimensions of the room [5].

However, while recording and mixing studio rooms have relatively simple and desirable acoustic properties that make their impulse response easy to analyse, live venues can exhibit complex time-frequency patterns that can make their inspection and interpretation complex and time consuming. For these reasons, although system equalisation is common practice in live sound, its application is limited to simple amplitude equalisation techniques that aim at making the desired RIR as smooth and neutral as possible. Such procedures are predominantly manually applied and rely often on RIR compensation algorithms that offer little or no insights about the proposed corrections.

Within the range of acoustic artefacts introduced by reproduction systems installed in venues, low frequency room modes constitute a challenging and important class of issues that can significantly affect the quality of live events. Although several modal control techniques have been proposed [6], such methods are often designed by criteria that are perceptually unfounded [7]. Moreover, large venues can exhibit complex patterns of multiple modal resonance that makes their interpretation and mitigation difficult.

Perceptual models can significantly simplify the analysis of a complex RIR by focusing on the key factors that determine the perceived quality of a room response. In particular, previous research on the subject has shown that temporal decay of room modes are a key factor in determining perceived low frequency quality [7].

Further research on the perception of low frequency resonance has highlighted the existence of frequency and level dependent perceptual modal thresholds (PMT) as a function of modal decay [8], which have been modelled in [9] with the following regression model:

$$PMT(\omega) = 0.15 + \frac{755}{\omega^2}, \tag{1}$$

where ω is the modal frequency under investigation expressed in *Hz*. Such works, inspired the creation of an automatic method to transform the impulse response of a room into a psycho-acoustically weighted frequency spectrum which greatly simplifies the identification of problematic room modes.

3. PROPOSED METHOD

The aim of the proposed method is to inform the user about potential issues caused by room modes which have perceptually relevant decay time and therefore become noticeable and can degrade the quality of the listening experience.

After the RIR has been measured [10, 11], the identification of the room modes starts from the calculation of the short time Fourier transform (STFT) of the measured impulse response, creating a waterfall plot. Waterfall plots, i.e. three-dimensional plots of the magnitude of a STFT, combine, in a very effective way, temporal and frequency characteristics of a RIR and can be used to visually identify issues caused by the acoustics conditions of the listening environment.

In order to increase the interpretability of the model, a psychoacoustic reference criteria such as the modal thresholds measured in [8] are introduced while performing the STFT analysis. The analysis is focused on the 32 - 250 Hz band, where the PMTs have been defined. An example of PMTs is shown in Figure 1 where they are superimposed on the STFT of a real RIR. Such approach, can constitute an informative visualization tool that can lead to a better understanding of the localisation of modal resonances: modes whose decay exceed the PMTs, are likely to be perceived as unwanted resonances and thus need to be properly treated. Unfortunately, such approach becomes quickly time consuming for complex RIR that exhibit multiple modal frequencies like those shown in Figure 1 which have decay larger than the PMTs (in particular: 46 Hz, 58 Hz, 68 Hz, 93 Hz, 125 Hz, 134 Hz and 145 Hz).

A further level of priority, based again on psychoacoustic criteria, needs to be defined to further refine the analysis of the RIR. As suggested in [6], to achieve optimal modal equalization, it is necessary to take into account the decay profile of each mode, as their unique shape carries important information about how human hearing will perceive resonances above the modal threshold. This considerations led to the creation of a mode detection function (MDF) defined as:

$$MDF(\omega) = \sum_{n=0}^{N} a_{\omega,n} \left\{ \frac{L_{\omega,n}}{PMT_{\omega,n}} \right\}, \tag{2}$$

that is a weighted sum of n ratios between the decay profile and the PMTs at each level $L_{\omega,n}$; the weighting coefficient $a_{\omega,n}$ is set as a function of the frequency and the number of detection levels and it can be empirically tuned by means of subjective listening tests.

The MDF is thus obtained following these steps:

1. STFT is calculated in the frequency range of interest and then normalised to maximum 0 $dBFS$;
2. PMTs are superimposed on the STFT;
3. the detection levels ($L_{\omega,n}$) are set (a minimum of three is advisable) depending on the target accuracy;
4. the MDF is calculated for each frequency bin in the STFT.

An example of the resulting MDF calculated for the RIR of Figure 1 is shown in Figure 2, together with its normalised magnitude spectrum. The *MDF* parameters are the followings ($\forall\omega \leq 250\,Hz$):

- $N = 6$;
- $a_{\omega,n} = 1 \quad \forall n \in [1, ..., N]$;
- $L_\omega = -[10, 20, 30, 40, 50, 60]\,dB$;

It can be noticed that the proposed MDF is a smoothed and perceptually weighted version of the amplitude spectrum of the RIR: while the amplitude spectrum presents several peaks and valleys, typical of complex RIR, the MDF exhibits only a few prominent peaks whose magnitude is proven to be correlated with the perceived modal resonance by means of subjective listening tests (not published).

Figure 2 clearly shows that the 68 Hz mode has the larger magnitude therefore, the user tend to rate it as the most critical one while the MDF gives priority to the 134 Hz mode. It is worth mentioning that both 92 and 134 Hz modes have similar magnitude on the FFT spectrum (hence similar energy content), what makes them ranked differently is the energy vs time distribution: the 134 Hz shows a longer decay therefore a stronger peak in the MDF.

When inspected visually, such diagram can serve as a very intuitive visual aid for sound engineers to guide the equalisation process both during system setup as well as mixing. This algorithm can also be used by an intelligent music production system to devise semi-supervised or unsupervised room equalisation strategies.

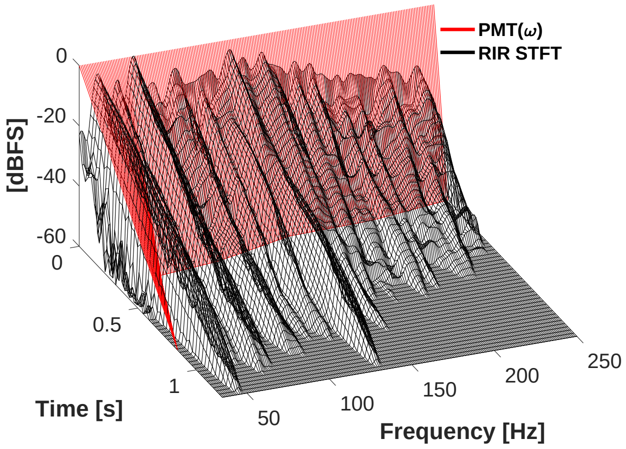

Figure 1: *Superposition of perceptual modal thresholds (red lines) as defined by [8] and STFT of a RIR exhibiting multiple modal resonances.*

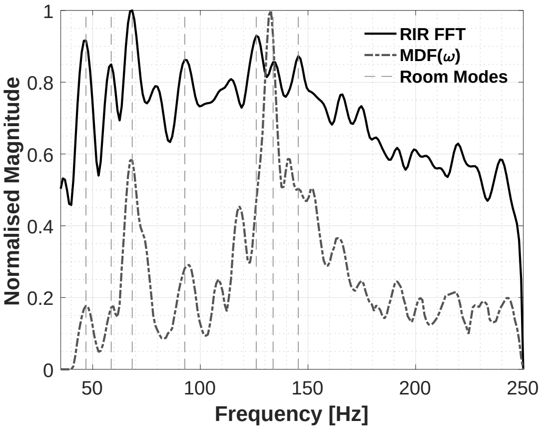

Figure 2: *Comparison between RIR amplitude spectrum and the proposed modal detection function.*

4. CONCLUSIONS

In this paper we introduced a novel, perceptually driven method, to identify audible artefacts due to the modal behaviour of a room. The algorithm simplifies and speeds up the complex task of room equalisation in a live context, focusing on modal resonances. Such tool can be used as a visual aid by sound engineers to manually set up modal equalisers, or by automated, intelligent music production tools to suggest modal equalisation presets as well as automatic correction of room impulse responses in a completely unsupervised manner.

Future developments of this work include the extension of the frequency range of the analysis as well as the development of more effective perceptually motivated modal correction strategies.

5. REFERENCES

[1] B. McCarthy, *Sound Systems: Design and Optimization : Modern Techniques and Tools for Sound System Design and Alignment*. Focal Press/Elsevier, 2010.

[2] S. Cecchi, A. Carini, and S. Spors, “Room response equalization — a review,” *Applied Sciences*, vol. 8, no. 1, 2018.

[3] H. Kuttruff, *Room Acoustics, Fifth Edition*. Taylor & Francis, 2009.

[4] H. Haas, “The influence of a single echo on the audibility of speech,” *J. Audio Eng. Soc.*, vol. 20, no. 2, pp. 146–159, 1972.

[5] F. E. Toole and S. E. Olive, “The modification of timbre by resonances: Perception and measurement,” *J. Audio Eng. Soc*, vol. 36, no. 3, pp. 122–142, 1988.

[6] A. Mäkivirta, P. Antsalo, M. Karjalainen, and V. Välimäki, “Modal equalization of loudspeaker - room responses at low frequencies,” *J. Audio Eng. Soc*, vol. 51, no. 5, pp. 324–343, 2003.

[7] B. M. Fazenda, M. R. Avis, and W. J. Davies, “Perception of modal distribution metrics in critical listening spaces—dependence on room aspect ratios,” *J. Audio Eng. Soc*, vol. 53, no. 12, pp. 1128–1141, 2005.

[8] B. M. Fazenda, M. Stephenson, and A. Goldberg, “Perceptual thresholds for the effects of room modes as a function of modal decay,” *J. Acoustical Society of America*, vol. 137, no. 3, pp. 1088–1098, 2015.

[9] J. Heddle, “Room acoustics for listening,” in *Proceedings of Acoustics2016*, (Brisbane, Australia), November 2016.

[10] A. Farina, “Advancements in impulse response measurements by sine sweeps,” in *Audio Engineering Society Convention 122*, May 2007.

[11] D. D. Rife and J. Vanderkooy, “Transfer-function measurement with maximum-length sequences,” *J. Audio Eng. Soc*, vol. 37, no. 6, pp. 419–444, 1989.