STATISTICAL MEASURES OF EARLY REFLECTIONS OF ROOM IMPULSE RESPONSES

Rebecca Stewart and Mark Sandler

Centre for Digital Music, Queen Mary, University of London
London, UK
{rebecca.stewart|mark.sandler}@elec.qmul.ac.uk

ABSTRACT

An impulse response of an enclosed reverberant space is composed of three basic components: the direct sound, early reflections and late reverberation. While the direct sound is a single event that can be easily identified, the division between the early reflections and late reverberation is less obvious as there is a gradual transition between the two.

This paper explores two statistical measures that can aid in determining a point in time where the early reflections have transitioned into late reverberation. These metrics exploit the similarities between late reverberation and Gaussian noise that are not commonly found in early reflections. Unlike other measures, these need no prior knowledge about the rooms such as geometry or volume.

1. INTRODUCTION

A room can be assumed to be a linear time-invariant system where the impulse response (IR) of the system can be found by recording a broadband signal within the room. Often the IR is convolved with non-reverberant audio to add artificial reverberation by simulating recording the audio in the room.

An IR of a space consists of direct sound, early reflections, and late reverberation. The early reflections are a set of discrete reflections whose density increases until individual reflections can no longer be discriminated and/or perceived. While the direct sound is a single event that can be easily identified, the early reflections and late reverberation of an IR are more difficult to label. For this paper, the transition time of an IR will be the earliest point in time when the density of the reflections has reached a perceptual threshold in which individual reflections can no longer be distinguished.

Since the seminal publications by Schroeder [1] and Moorer [2], digital artificial reverberators have contained a component intended to create discrete echoes in order to simulate early reflections and a component whose intent is to create a set of reflections as dense as possible. Later developments with feedback delay networks acknowledge that prior to high frequency attenuation, the reverberator should produce white noise [3] and Moorer in [2] first discussed using frequency-shaped Gaussian noise to simulate the energy in late reverberation. The transition from early reflections to late reverberation can then be modeled as a deterministic system that transitions into a stochastic one [4]. The statistics regarding early reflections are of greater concern here because the statistics of late reverberation has been covered in depth, particularly in [3]. Abel and Huang [5] have also recently explored similar statistics as are examined here, looking at measures of reverberation quality, particularly for judging artificial reverberation.

The late reverberation of an IR tends towards a normal distribution, unlike the energy from early reflections. A progression towards a more normal distribution occurs as time increases and the acoustic energy within the space becomes more mixed [4]. A measurement of distribution can then be used to determine whether a point in time is more or less deterministic. That is, whether it is within the early reflections or late reverberation.

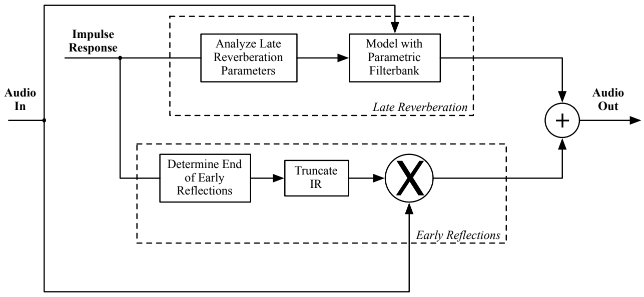


Figure 1: Basic design of a hybrid reverberator.

The ability to determine the mixing time of a space is especially relevant to hybrid reverberation. Hybrid reverberation uses both convolution and recursive filterbank techniques, as can be seen in Fig. 1. An IR is truncated to ideally contain the early reflections. The truncated IR is convolved with the dry audio and then the late reverberation is simulated with a filterbank. A precise measure of when a room is first mixed is important so that all perceptually relevant information is preserved in the early reflections of the truncated IR while the size of the truncated IR is as small as possible to reduce the length of the convolution.

1.1. Definitions of Early Reflections

Early reflections are loosely defined as a set of echoes that have not reached a perceptual threshold, and that can be described by the mathematical relationships that define the dispersion of echoes in a space such as

$$\frac{dN_r}{dt} = 4\pi \frac{c^3 t^2}{V} \tag{1}$$

where N_r is the number of reflections, t is the time from the direct sound, c is the speed of sound, and V is the volume of the room [6].

In [4], Blesser describes the mixing time as "how long it takes for there to be no memory of the initial state of the system. There is statistically equal energy in all regions of the space after the mixing time." He estimates this to be approximately three times the mean free path and directly a property of the geometry of the room. Here the mixing time is accepted to be after the transition from early reflections to late reverberation is complete. The upper

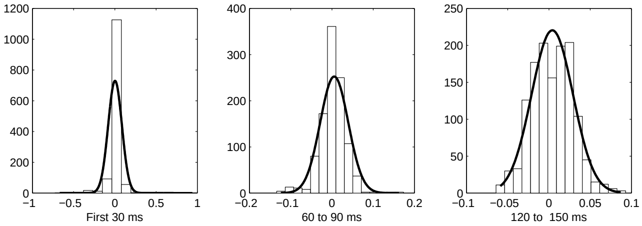


Figure 2: Progression in time of histograms and fit normal functions of 30 ms windows of an IR. Note the differing vertical and horizontal scales. The absolute values of the vertical scale are not important, but rather the relative values between the histogram bars. The range of the horizontal scale decreases in subsequent windows because the signal is decreasing in amplitude. It is important to note the distribution of the samples, not the absolute values.

limit of mixing time has been discussed in [7] and further in [3] to be

$$t_{mixing} = \sqrt{V} \tag{2}$$

where V is the volume in m^3.

Standardized measurements of room acoustics divide the IR of a room into an early and late portion in order to calculate early lateral energy, clarity and definition. When measuring for musical material, the early portion is defined as the first 80 ms [8]. In most literature, it is accepted that the early reflections are contained within the first 80 ms [9].

The transition from early reflections from late reverberation is either defined by a point in time regardless of the room properties, usually 80 ms, or is calculated based on physical properties of a room, most commonly volume. Using a single point in time regardless of the space is inaccurate, but access to the dimensions of a space may be impractical or impossible. A blind method that can determine the transition point between early reflections and late reverberation without knowledge of the measurements of the space is needed. Two possible methods are described here.

2. MEASURES OF DISPERSION

The transition from early reflections to late reverberation can be observed in several domains. This paper only addresses the transition with regard to a Gaussian distribution in the time domain, but other factors can be considered. As discussed in [10] and [11], the frequency distribution tends towards a Rayleigh distribution. As only monophonic impulse responses are studied here, the spatial properties of the room are not being considered.

2.1. Standard Deviation

The standard deviation of a group of samples is a measure of the spread of the samples and is defined as

$$\sigma = \sqrt{\mathbf{E}(x^2) - (\mathbf{E}(x))^2} \tag{3}$$

where $\mathbf{E}(x)$ is the expected value of x.

In a normal distribution, approximately one third of the samples lie outside one standard deviation of the mean and approximately two thirds of the samples are within one standard deviation of the mean. Early reflections, however, have more samples within one standard deviation and fewer outside (see Fig. 2). The progression from early reflections to late reverberation can then be observed through the ratio of samples outside one standard deviation versus inside.

As an IR progresses in time the ratio of samples gradually approaches approximately one third. Fig. 3 shows the ratio outside one standard deviation versus inside for a 30 ms window. The ratio is divided by $efrc(1/\sqrt{2})$, the expected value of samples outside one standard deviation, to normalize for Gaussian distribution. This is similar to what is done in [5].

Examples for three different spaces can be seen in Fig. 3. Fig. 3(a) is a measurement from a smaller space, a 350 seat concert hall, than those in Fig. 3(b) and 3(c) which are both large, reverberant churches. The measuring of the IRs from the churches is described in [12]; the concert hall was measured by the authors using the same technique and equipment.

The normalized ratio outside versus inside one standard deviation approaches one as time progesses, but it makes for a poor measure. It is difficult to identify the point in time when it can first be assumed that the room is mixed. While the curve approaches a threshold, that threshold is not constant amongst IRs from different spaces and is difficult to extrapolate from the curve. However, it is clear that the IR does gradually transition from less diffuse early reflections to more diffuse late reverberation. Using the ratio of samples outside to inside one standard deviation does not allow an easy selection of a transition point.

3. HIGHER ORDER STATISTICS

Moments describe deterministic signals as they “are numerical measures of the degree of similarity between a signal and a product of delayed or advanced versions of itself" [13]. An nth order cumulant is a function of its joint moment of orders up to n [14]. Higher order cumulants contain amplitude and phase information unlike second order statistics (correlation) which are phase-blind. The second order cumulant is the variance, the third order is skewness and fourth order is kurtosis [13].

While moments describe deterministic signals, cumulants are measures for stochastic signals. If a set of random variables are jointly Gaussian, then all information about their distribution is in the moments of an order less than or equal to two. It can then be interpreted that cumulants of order greater than two measure the non-Gaussian nature of a time series [13]. Further, if a non-Gaussian signal is mixed with a Gaussian signal, higher-order cumulants will ignore the Gaussian noise portion of the signal [13].

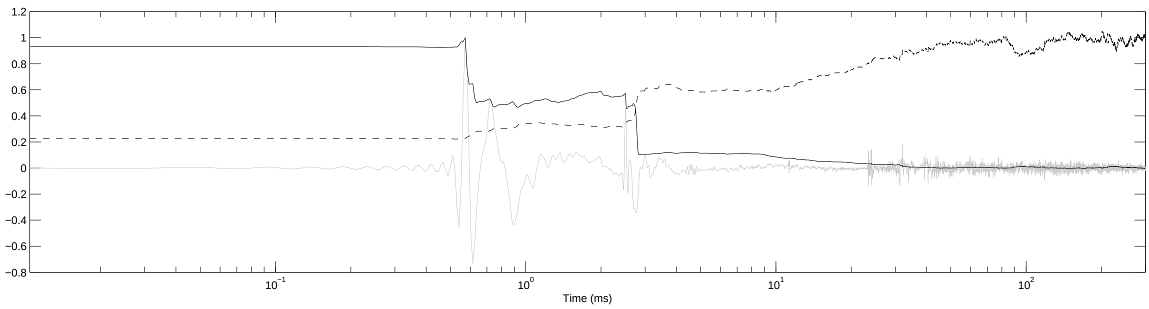
3.1. Kurtosis

The fourth order zero-lag cumulant of a zero-mean process is often referred to as kurtosis and can either be normalized or unnormalized [15]. Here kurtosis will refer to the normalized definition.

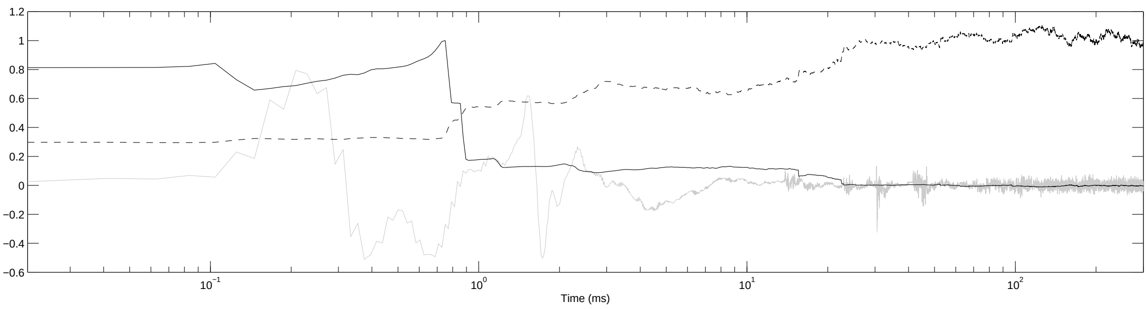
$$\gamma_4 = \frac{\mathbf{E}(x-\mu)^4}{\sigma^4} - 3 \tag{4}$$

where $\mathbf{E}()$ is the expectation operator, μ is the mean, and σ^2 is the standard deviation.

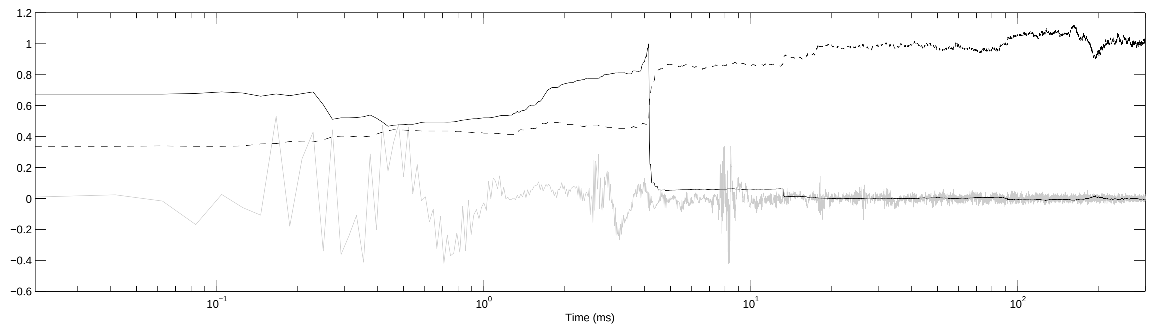
The same sliding window of 30 ms is used for the kurtosis analysis as for the standard deviation measurements. The values have been normalized to one within each IR merely for the sake of the figure. The values nearing zero, denoting a mixed room, and



(a) Sir Jack Lyons Concert Hall, York, UK, a 350 seat concert hall. Reverberation time T_{30} = 4.9 s



(b) York Minster in York, UK, a 330,000 m^3 church. Reverberation time T_{30} = 8.19 s.



(c) St. Andrew's in Lyddington, UK, a 2600 m^3 church. Reverberation time T_{30} = 1.5 s.

Figure 3: The solid gray lines are the IRs, the dashed lines are the normalized ratios of samples outside one standard deviation and the solid black lines are the normalized kurtosis values.

the rate of change of the values are more important than the absolute values, especially those of the peaks in the early reflections.

As can be seen in Fig. 3(a) to 3(c), the plots of the kurtosis show a distinct difference between the early and late energy unlike the plots of the standard deviation which only demonstrated the gradual transition. Two significant points can be seen on the plots: a rapid decrease in the kurtosis value and when the kurtosis value is first zero. All three spaces depicted in Fig. 3(a) to 3(c) have a large decrease in the kurtosis value at approximately the same time, ranging from 2 to 5 ms. Differences occur in the later point, when the kurtosis value is approximately zero. The time is much later in Fig. 3(a) at 45 ms; the larger, more reverberant spaces have earlier kurtosis values nears zero at 26 ms for Fig. 3(b) and 19 ms for Fig. 3(c).

4. FURTHER WORK

The ability to determine the transition time of an IR is useful for determining at what point in an IR an artificial reverberator can be used such as in [3]. This is particularly useful for hybrid reverberators such as in Fig. 1 which use a combination of convolution and filterbank reverberators.

While the kurtosis of an IR shows a significant difference in the early and late portions of a signal, the perceptual implications of this transition have not been explored. Listening tests need to be conducted to determine if sufficient perceptual information is contained in the signal before the transition point. An IR truncated at the transition point needs to hold the same localization information as the complete IR. Only the reverberation tail which contains cues to the size of the space, not details of the localization of the sound source, should be removed.

There are a number of acoustical parameters such as reverberation time, clarity and early decay time that are dependent upon the absorption and diffusion of space. These measurements help characterize a space and quantitatively describe its suitability for music or speech. A study surveying these standard measurements and the transition point could be carried out to determine whether the transition time is useful measure of a space.

5. DISCUSSION

A measure that can determine the point in time when the early reflections have fully transitioned to late reverberation is described. Previous measures either need specific information about the space such as volume, or completely disregard the space and define the transition point to be 80 ms. A method that does not disregard the room properties but still does not need specific dimensions is implemented by analyzing the statistics of the IR of the space.

Higher order cumulants such as kurtosis are a more convenient measure for the transition time than lower order descriptive statistics. Kurtosis is essentially blind to symmetric probability distributions, unlike standard deviation. This gives a more definite threshold to determine a mixed room since the threshold when using standard deviation is dependent on the room.

Both standard deviation and kurtosis support the model of an IR being a deterministic system transitioning into a stochastic one, however they display this in two different ways. The standard deviation of the IR shows a gradual transition from the early reflections to late reverberation, but the earliest point in which the room is mixed is difficult to find. The kurtosis takes into account phase information while possessing blindness to symmetric distributions. A much sharper transition is then shown allowing a specific transition point to be easily selected.

The estimator in Eq.2 finds the York Minster to have a mixing time of 570 ms, while St. Andrew's is estimated to be mixed by 60 ms. Both are much later times than the kurtosis measurements of 26 and 19 ms respectively. The volume of the Sir Jack Lyon Concert Hall is not available, so the mixing time cannot be estimated from Eq.2.

The mixing time is relative to the beginning of the file containing the IR. In order for the calculated mixing time to be relevant amongst a group of IRs, the beginning of the files need to be standardized, preferably starting at the time of the impulse. Not all IRs contain the time between the impulse and the direct sound and some have the direct sound removed for use in convolution reverberators. This needs to be noted before multiple IRs are compared to each other.

Further work needs to be done to explore the perceptual relevance of this point in time and of the robustness of a truncated IR to still contain localization information. If listening tests show that an IR can be truncated to its transition point and still retain all sufficient localization information, then this method of truncation can be used in applications such as hybrid reverberators. The relationship between the transition time, other acoustical measures, and the properties of the room can also be studied further.

6. ACKNOWLEDGEMENTS

This research was supported by the Audio Engineering Society Educational Foundation and Queen Mary, University of London.

7. REFERENCES

[1] M.R. Schroeder, "Natural sounding artificial reverberation," *JAES*, vol. 10, no. 3, pp. 219–223, October 1962.

[2] James A. Moorer, "About this reverberation business," *Computer Music Journal*, vol. 2, Summer, 1979.

[3] Jean-Marc Jot, Laurent Cerveau, and Olivier Warusfel, "Analysis and synthesis of room reverberation based on a statistical time-frequency model," in *103rd AES Convention*, New York, NY, September 1997.

[4] Barry Blesser, "An interdisciplinary synthesis of reverberation viewpoints," *JAES*, vol. 49, no. 10, pp. 867–903, October 2001.

[5] Jonathan S. Abel and Patty Huang, "A simple, robust measure of reverberation echo density," in *121st AES Convention*, San Francisco, October 2006.

[6] H. Kuttruff, *Room acoustics*, Spon Press, London, 4th edition, 2000.

[7] Jean-Dominique Polack, "Playing billiards in the concert hall: the mathematical foundations of geometrical room acoustics," *Applied Acoustics*, vol. 38, pp. 235–244, 1993.

[8] ISO 3382, "Acoustics-measurements of the reverberation time of rooms with reference to other acoustical parameters," 1997.

[9] Durand R. Begault, *3-D sound for virtual reality and multimedia*, NASA, 2000.

[10] M.R. Schroeder, "Statistical parameters of the frequency response curves of large rooms," *JAES*, vol. 35, no. 5, pp. 307–316, May 1987.

[11] M.R. Schroeder, "Normal frequency and excitation statistics in rooms: model experiments with electric waves," *JAES*, vol. 35, no. 5, pp. 307–316, May 1987.

[12] Damian T. Murphy, "Archaeological acoustic space measurement for convolution reverberation and auralization applications," in *DAFX '06*, September 2006.

[13] Chrysostomos L. Nikias and Athina P. Petropulu, *Higher-order spectra analysis: a nonlinear signal processing framework*, Prentice Hall, 1993.

[14] Jerry M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proc. of the IEEE*, vol. 79, no. 3, pp. 278–305, March 1991.

[15] Ananthram Swami, Jerry M. Mendel, and Chrysostomos L. Nikias, *Higher-order spectral analysis toolbox*, MathWorks, 3rd edition, 1998.