Audio Engineering Society

Convention Paper

Presented at the 123rd Convention
2007 October 5–8 New York, NY

Aspects of Reverberation Echo Density

Patty Huang[1] and Jonathan S. Abel[1]

[1] *Center for Computer Research in Music and Acoustics (CCRMA), Stanford Univ., Stanford, CA, 94305, USA*

Correspondence should be addressed to Patty Huang (pph@ccrma.stanford.edu)

ABSTRACT

Echo density, and particularly its time evolution at the reverberation impulse response onset, is thought to be an important factor in the perceived time domain texture of reverberation. In this paper, the psychoacoustics of reverberation echo density is explored using reverberation impulse responses synthesized via a Poisson process to have a variety of static and evolving echo densities. In addition, a recently proposed echo density measure called the normalized echo density, or NED, is explored, and related via a simple expression to echo density specified in echoes per second using echo patterns with static echo densities. A continuum of perceived time-domain texture was noted, from "sputtery" around 100 echoes per second to "smooth" above about 20,000 echoes per second, at which point it was perceptually identical to Gaussian noise. The character of the reverberation impulse response onset was explored for various rates of echo density increase, and ranged from "sputtery" for long mixing times to "instantly smooth" for short mixing times.

1. INTRODUCTION

In a reverberant environment, radiated sound interacts with objects and surfaces to create reflections. Further interactions create additional reflections, and a listener in the environment would experience an increasing arrival density over time. The time evolution of echo density, and particularly its rate of increase, are commonly thought to influence the perceived time-domain timbre of a space, with large, rapidly increasing echo densities being desired.

In the context of developing an artificial reverberator based on pseudorandom sequences of echoes [4, 5], Rubak and Johansen cited literature and presented listening tests suggesting that reverberation impulse responses maintaining echo densities greater than about four thousand echoes per second may be treated statistically and sound similar to one another.

In [6], the variance over time of reverberation impulse response energy (which can be indicative of echo density) was seen to correlate with the subjective texture of late reverberation. In other recent work [7], temporal correlations in reverberation impulse response frequency bands were shown to relate to the "temporal color" of the reverberator. The idea was that noise-like segments of a reverberation impulse response would maintain small autocorrelation sidelobes compared to their peak values, whereas segments with undesirable periodicities would have relatively larger sidelobes.

In [1, 2], a measure of reverberation echo density was presented wherein successive frames of the reverberation impulse response were analyzed to produce a so-called normalized echo density profile, or NEDP. The NEDP examines the percentage of impulse response taps lying outside the local standard deviation, and is normalized to range from near zero, indicating few echoes, to around one, indicating a fully dense reverberation having Gaussian statistics. The echo density measure was seen to be insensitive to level, equalization, reverberation time and sampling rate, and it was seen to correlate well with theoretically expected echo densities as a function of time for a number of measured and artificial reverberation impulse responses. (An indicator of echo density based on impulse response kurtosis was described by Usher in [3, pp. 152–153]. In [1], impulse response tap kurtosis and outlier percentage were seen to have similar performance.)

In this paper, we focus on the psychoacoustics of reverberation echo density, and explore the perception of reverberation impulse responses synthesized to have a range of static and evolving echo densities. We also develop a formula relating echo density measured in echoes per second and echo density on its normalized scale.

Section 3 defines the normalized echo density, and presents an expression relating it to the echo density measured in echoes per second, called here the absolute echo density. Section 4 describes the psychoacoustics of static and evolving echo densities, and Section 5 summarizes the results. We begin by describing the method used to generate echo patterns having a prescribed echo density profile.

2. ECHO PATTERN GENERATION

The analyses in this paper make use of echo patterns having a variety of static and evolving echo densities. These echo patterns are generated using a Poisson process to define the time intervals between echo arrivals, with amplitudes drawn from Gaussian distributions. The echo patterns are generated such that the average energy in a time interval of a given length is constant throughout the pattern. In this way, the perception of various echo densities can be explored, and by imposing the appropriate decay as a function of frequency, these patterns may be turned into reverberation impulse responses with identical energy profiles.

Denoting by $\rho(t)$ the absolute echo density as a function of time t, the time interval τ between arrivals at time t is drawn from an exponential distribution,

$$\wp(\tau; t) = \exp\{-\tau/\rho(t)\}. \tag{1}$$

In other words, the time of the $(i+1)$st echo t_{i+1} is given by the time of the ith echo t_i plus an exponential random variable scaled according to the echo density at that time.

The amplitude $\alpha(t)$ for the echo at time t is drawn from a zero-mean Gaussian distribution with variance scaled by the inverse echo density at that time,

$$\alpha(t) \sim \mathcal{N}(0, 1/\rho(t)). \tag{2}$$

The idea behind the variance scaling is as follows. Because of the statistical nature of the echo arrival times, the echoes may be considered independent, and the energy in a given time interval is approximately the sum of their square amplitudes. By scaling the echo amplitudes according to the local echo density, the energy will be roughly constant throughout the echo pattern.

After generating the echo arrival times and amplitudes, the echo pattern is generated using sinc interpolation. The examples presented here have a sampling rate of 44.1 kHz and a 17-point sinc interpolation filter was used. The perception of echo density depends on the duration (or bandwidth) of the echo, and a second-order Butterworth lowpass filter was applied to achieve different echo durations.

Figure 1 shows a set of echo patterns generated according to static echo densities ranging from 10 to 1,000,000 echoes per second. Figure 6 shows synthetic reverberation impulse responses generated according to quadratically increasing echo densities. The range of echo densities in evident in these time

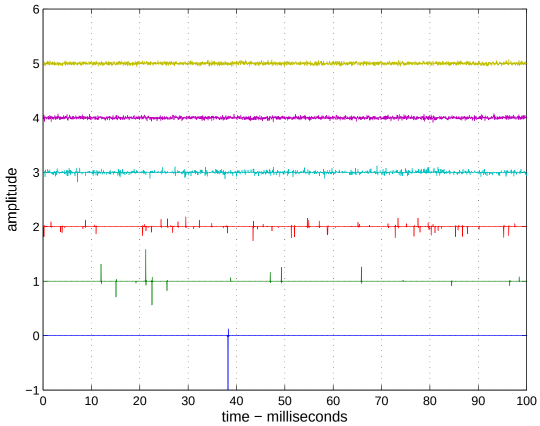

Fig. 1: Synthesized echo patterns maintaining statistically constant echo densities of 10, 100, 1000, 10,000, 100,000, and 1,000,000 echoes per second (offset bottom to top).

series. As expected, the echo amplitude decreases with increasing echo density.

3. NORMALIZED ECHO DENSITY

3.1. Definition

Over a sliding reverberation impulse response window, the normalized echo density profile $\eta(t)$ is the fraction of impulse response taps which lie outside the window standard deviation:

$$\eta(t) = \frac{1/\mathrm{erfc}(1/\sqrt{2})}{2\beta+1} \sum_{\tau=t-\beta}^{t+\beta} \mathbf{1}\{|h(\tau)| > \sigma\}, \qquad (3)$$

where $h(t)$ is the reverberation impulse response (assumed to be zero mean), $2\beta+1$ is the window length in samples, σ is the window standard deviation,

$$\sigma = \left[\frac{1}{2\beta+1} \sum_{\tau=t-\beta}^{t+\beta} h^2(\tau)\right]^{\frac{1}{2}}, \qquad (4)$$

$\mathbf{1}\{\cdot\}$ is the indicator function, returning one when its argument is true and zero otherwise, and $\mathrm{erfc}(1/\sqrt{2}) \doteq 0.3173$ is the expected fraction of samples lying outside a standard deviation from the mean for a Gaussian distribution.

The normalized echo density profile is more generally computed using a positive weighting function $w(t)$ so as to de-emphasize the impulse response taps at the sliding window edges:

$$\eta(t) = \frac{1}{\mathrm{erfc}(1/\sqrt{2})} \sum_{\tau=t-\beta}^{t+\beta} w(\tau)\mathbf{1}\{|h(\tau)| > \sigma\} \qquad (5)$$

with

$$\sigma = \left[\sum_{\tau=t-\beta}^{t+\beta} w(\tau)h^2(\tau)\right]^{\frac{1}{2}} \qquad (6)$$

and where $w(t)$ is normalized to have unit sum $\sum_\tau w(\tau) = 1$.

As an example, consider the impulse response shown in Figure 2, measured in the lobby of the Knoll, a building at Stanford University. The direct path is evident, as are a number of early reflections giving way to a noise-like late field response. Also shown is the echo density profile, $\eta(t)$, computed according to (3) using a 20 ms long sliding boxcar window, and tap histograms for several of the windows used in computing the echo density profile.

The echo density profile starts near zero and increases over time to around one, where it remains for the duration of the impulse response. This is typical of echo density profiles of measured impulse responses and high-quality synthetic reverberators.

As described in [1], what sets one profile apart from another is the rate of increase and the time at which a value near one is first attained, indicating the start of the late field. This behavior is reflected in window histograms, calculated at various times along the impulse response. At the beginning of the impulse response, the histograms are peaked at small tap values compared to the Gaussian, indicating the concentration of the window energy in a relatively small number of taps. After the echo density profile nears a value of one, the histograms take on a Gaussian form, consistent with the notion that the late field has started.

3.2. Relationship to Absolute Echo Density

In [1], normalized echo density was seen to be an indicator of the perceived texture of reverberation impulse responses. In exploring the psychoacoustics of reverberation echo density, it is therefore useful

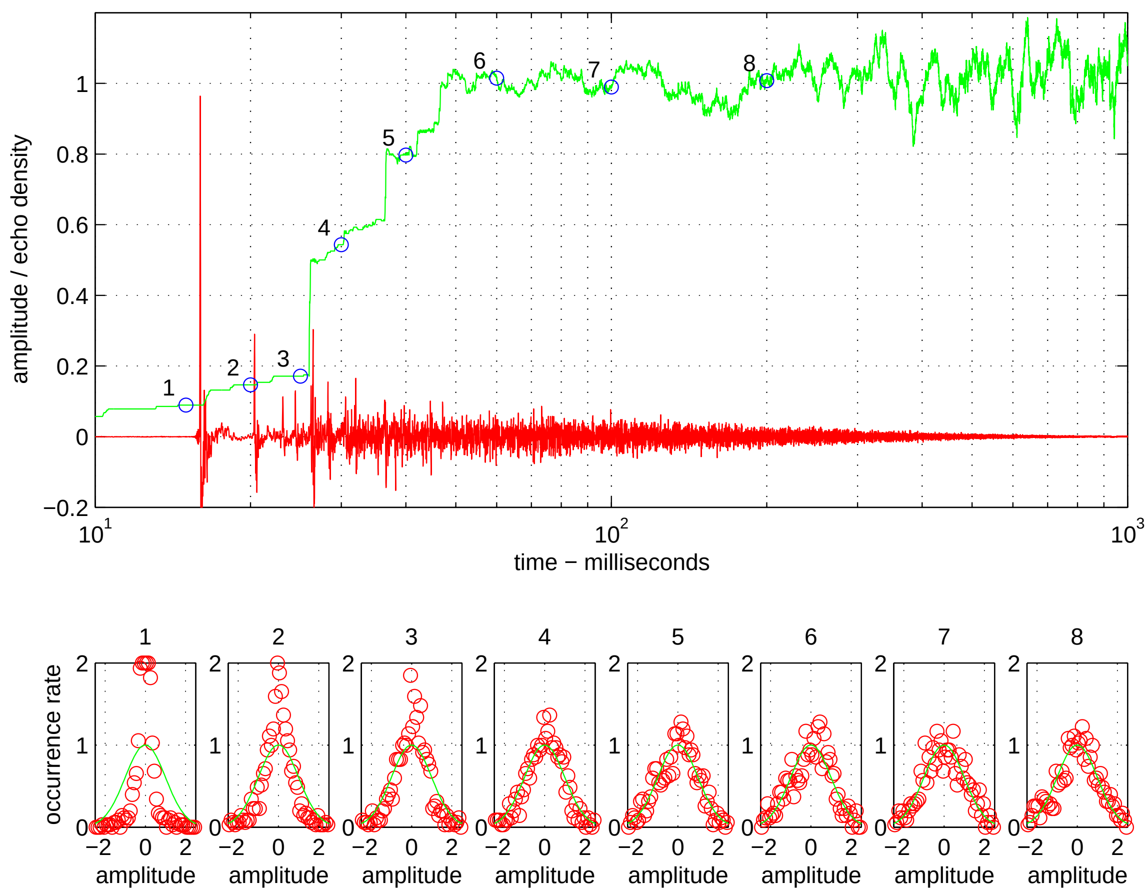

Fig. 2: Top: Measured room impulse response (red, near bottom, 44.1 kHz samples) and its echo density profile (green, near top). Note that the time axis is on a logarithmic scale. Bottom: Window histograms centered at times indicated by the blue circles in top plot. The Gaussian distribution is indicated by a green line. Normalized occurence rates greater than 2.0 are plotted at this value.

to have an expression relating echo density in echoes per second to normalized echo density.

Consider the case of a very sparse echo pattern, with echoes of duration δ and density $\rho << 1/\delta$. In a time window of duration β, the expected number of echoes is then $\rho\beta$, and the number of outlying samples in the window will be $\beta\rho\delta$, as almost the entirety of each echo is expected to lie outside the window standard deviation. The normalized echo density in the window of length β is then roughly $\rho\delta$. For small echo densities, the normalized echo density is expected to be proportional to the absolute echo density.

In the case of a very dense echo pattern, where $\rho >> 1/\delta$, many echoes overlap and the echo pattern samples are expected to take on a Gaussian distribution. In this case, the normalized echo density will fluctuate around 1.

This small and large echo density behavior leads to an expression relating normalized echo density η to the absolute echo density ρ measured in echoes per second,

$$\eta = \frac{\delta\rho}{\delta\rho + 1}, \tag{7}$$

where δ is the echo duration in seconds, or alternatively the inverse echo bandwidth in 1/Hz. By

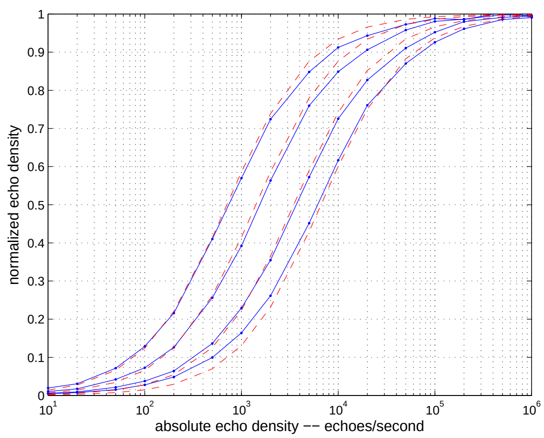

Fig. 3: Measured (solid) and modeled (dashed) normalized echo densities for echo patterns having static echo densities and bandwidths of 1, 2, 5, and 10 kHz (left to right).

rearranging terms, absolute echo density may be expressed in terms of normalized echo density,

$$\rho = \frac{\eta/\delta}{1-\eta}. \tag{8}$$

Figure 3 shows normalized echo density as a function of absolute echo density measured from a set of echo patterns with static echo densities ranging from 10 to 1,000,000 echoes per second and echo bandwidths of 1, 2, 5, and 10 kHz. Also plotted in Figure 3 is the normalized echo density given by (7).[1] Note that (7) closely tracks the normalized echo density over a wide range of absolute echo densities and echo bandwidths.

4. ECHO DENSITY PSYCHOACOUSTICS

4.1. Static Echo Density

Sets of two-second-long echo patterns were generated with constant echo densities ranging from 100 echoes per second to 1,000,000 echoes per second and with different echo bandwidths (1, 2, 5, and 10 kHz). Informal listening tests revealed significant differences in temporal timbre due to both absolute echo density and the echo bandwidth.

[1] The echo durations δ needed in (7) were computed as the inverse warped cutoff frequency of the second-order lowpass Butterworth filter used to generate the echo pattern.

Consider the set of 5 kHz bandwidth echo patterns. Patterns with an echo density of a few hundred echoes per second were "sputtery," with individual echoes audible. As echo density increased, the timbre became "crackly," resembling the sound of a fire. Further increase in echo density then fused the pattern into a more cohesive "coarse" or "rough" texture, with the degree of coarseness and sense of aural friction decreasing with increasing absolute echo density. At densities above 20,000 echoes per second, the roughness became fine-grained indistinguishable from Gaussian noise.

Examining all of the echo patterns, the perceived noise textures appear to fall into three descriptive categories: 1) sputtery, crackly, or sizzly (dependent on the echo bandwidth) 2) rough or coarse, and 3) smooth or windy. In the first group, individual echoes are audible and the texture is not homogeneous. Echo patterns with wider bandwidths of 5 or 10 kHz were best described as sounding "sizzly," similar to oil frying in a pan. In contrast, patterns with less high frequency content (e.g., bandwidths of 1 or 2 kHz) sounded "sputtery." In the second group, echoes fuse into a homogenoeus texture of varying energy, described as "rough," "coarse," or "ragged." The third group is categorized by a smooth, noise-like texture. Depending on bandwidth, it could be described as "sandy" or "windy" and is very similar to Gaussian noise of the same bandwidth.

Figure 4 shows the texture categories for all the static echo density patterns explored. It is important to note that the breakpoints between the three groups occurred at the same normnalized echo density, independent of echo bandwidth, whereas they occurred at much different absolute echo densities for the different echo bandwidths. We therefore suggest using normalized echo density as a predictor of perceived temporal texture.

4.2. Evolving Echo Density

Artificial and measured reverberation impulse responses maintain echo density profiles which grow from low echo densities commensurate with the dimensions of the room to high echo densities indistinguishable from Gaussian noise. As described in

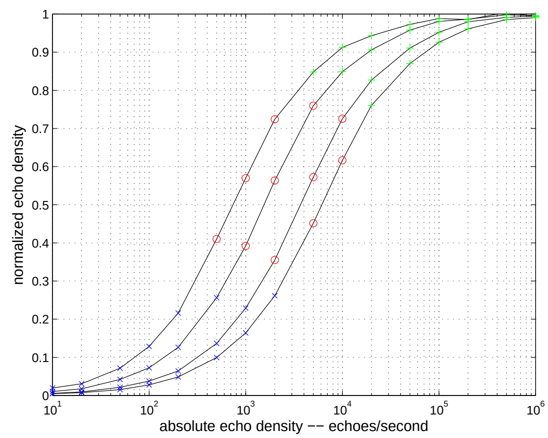

Fig. 4: Echo patterns with static echo density, corresponding to those shown in Figure 3. Listening tests suggested classification into three groups: sputtery (x), rough (o), and smooth (+).

[1], the main feature distinguishing different reverberation echo density profiles is the mixing time—the time it takes to acheive a Gaussian late field. In three-dimensional spaces, echo density increases quadratically with time and proportional to the inverse room volume [8].

To explore the perception of reverberation mixing time, sets of echo patterns were generated having quadratically increasing echo densities and characterized by the time at which the echo density reached a specified threshold. Echo patterns reaching a density of 20,000 echoes per second at times ranging from 10 ms to 1 second were generated. The echo densities were capped at 500,000 echoes per second to reduce computational expense. At this echo density, Gaussian statistics have been achieved for all bandwidths studied.

To generate the reverberation impulse responses, a single set of echo time intervals and amplitudes was generated. The time intervals were then stretched according to the echo densities shown in Figure 5. In addition, the amplitudes were normalized according to the inverse square root echo density so as to achieve constant energy density throughout the pattern. The echo patterns were then converted into impulse responses by applying an exponential win-

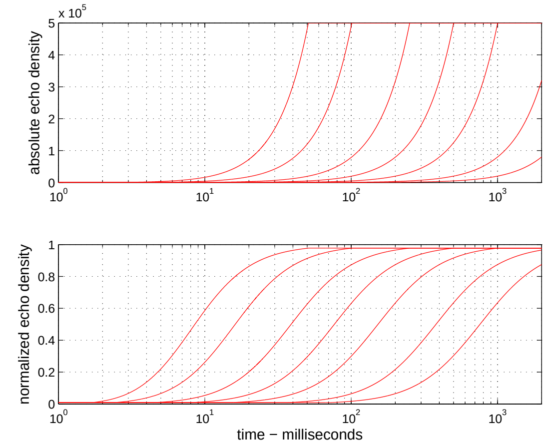

Fig. 5: Normalized echo density profiles (bottom) and absolute echo profiles (top) for a set of evolving reverberation impulse response echo densities.

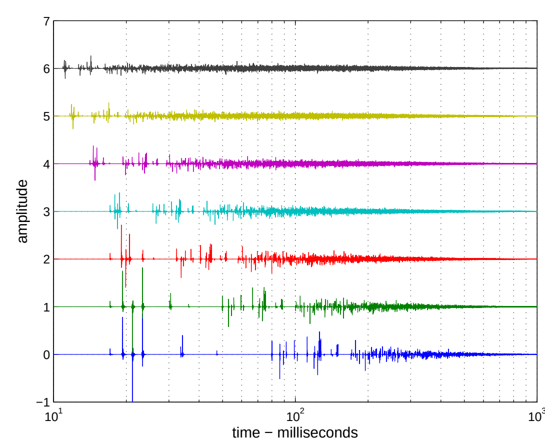

Fig. 6: Reverberation impulse responses having the echo density profiles shown in Figure 5.

dow with a 60 dB decay time of 2 seconds. The impulse responses were also equalized using second-order Butterworth lowpass filters of different bandwidths.

A set of resulting reverberation impulse responses are shown in Figure 6. Note that the impulse responses having long mixing times have larger onset amplitudes and exhibit echo patterns that appear to be time-stretched versions of those appearing at the beginning of the impulse responses with short mixing times.

A number of sets of impulse responses were generated in this manner. Informal listening tests suggested that the various reverberation impulse responses could be classified into three groups. At long mixing times of about 200 ms or more, the onset was "sputtery", with a small number of individual echoes being audible. At medium mixing times, between about 50 and 200 ms, individual echoes were not apparent, but the onset could be described as rough or ragged. At short mixing times, less than about 50 ms, the onset was instantly smooth, perceptually similar to a stepped Gaussian noise burst.

5. SUMMARY

The perception of echo density in reverberation impulse responses was explored. Echo patterns having constant echo density were classified into three groups: "sputtery," "rough," and "smooth." The normalized echo density [2, 1], which indicates closeness to Gaussian statistics, was seen to delineate the groups, with approximate breakpoints at $\eta = [0.3, 0.75]$. By contrast, breakpoints given in terms of absolute echo density (specified in echoes per second) ρ were found to be highly dependent on echo bandwidth. Reverberation impulse responses having quadratically increasing echo densities were generated to explore the effect on mixing time on reverberation impulse response perception. Again, the impulse responses fell into three categories: "sputtery," with a few individual echoes audible, for long mixing times; "ragged," for intermediate mixing times; and immediately smooth and similar to stepped Gaussian noise, for short mixing times.

6. REFERENCES

[1] J. Abel and P. Huang, "A simple, robust measure of reverberation echo density," in Proceedings of the 121st AES Convention, preprint 6985, San Francisco, October 2006.

[2] J. S. Abel and D. P. Berners, *Reverberation acoustics, analysis, and synthesis.*, Tutorial session T21, 117th Convention of the Aes, San Francisco, October 2004.

[3] John S. Usher, *Subjective evaluation and electroacoustic theoretical validation of a new approach to audio upmixing*, Ph.D. thesis, McGill University, Montreal, Quebec, Canada, 2006.

[4] Rubak, P. and Johansen, L., "Artificial reverberation based on a pseudo-random impulse response," in *Proceedings of the 104th AES Convention*, preprint 4875, Amsterdam, Netherlands, May 16–19, 1998.

[5] Rubak, P. and Johansen, L., "Artificial reverberation based on a pseudo-random impulse response II," in *Proceedings of the 106th AES Convention*, preprint 4875, Munich, Germany, May 8–11, 1999.

[6] Huang, P. and Smith, J. O., "Loudness-based display and analysis applied to artificial reverberation," in *Proceedings of the 2004 International Computer Music Conference*, pp. 37–40, Miami, Florida, November 1–6, 2004.

[7] Rubak, P., "Evaluation of artificial reverberation decay quality," in *Proceedings of the 118th AES Convention*, preprint 6395, Barcelona, Spain, May 28–31, 2005.

[8] Pierce, A., *Acoustics*, McGraw-Hill, 1981.