Audio Engineering Society

Convention Paper

Presented at the 132nd Convention
2012 April 26–29 Budapest, Hungary

The Hand Clap as an Impulse Source for Measuring Room Acoustics

Prem Seetharaman[1], Stephen P. Tarzia[1*]

[1] *Northwestern University, EECS Department, Evanston, IL, USA*

Correspondence should be addressed to Prem Seetharaman (prem@u.northwestern.edu)

ABSTRACT

We test the suitability of hand clap recordings for measuring several acoustic features of musical performance and recording rooms. Our goal is to make acoustic measurement possible for amateur musicians and hobbyists through the use of a smartphone or web app. Hand claps are an attractive acoustic stimulus because they can be produced easily and without special equipment. Hand claps lack the high energy and consistency of other impulse sources, such as pistol shots, but we introduce some signal processing steps which mitigate these problems to produce reliable acoustical measurements. Our signal processing toolchain is fully-automated, which allows both amateurs and technicians to perform measurements in just a few seconds. Using our technique, measuring a room's reverberation times and frequency response is as easy as starting a smartphone app and clapping several times.

1. INTRODUCTION

Modern digital recording technology has made it possible for amateur, student, and semi-professional musicians to easily produce high-quality recordings in inexpensive, makeshift "studios". While recording novices may have equipment that provides excellent sound fidelity their recordings will likely be flawed by problems in the recording room's acoustics. Recording novices almost universally lack the experience and measurement tools needed to diagnose and correct the room's acoustical flaws. We believe that easy-to-use, inexpensive acoustic measurement tools can be developed—for example a smartphone app which provides an acoustics assessment in a few simple steps.

However, standard acoustic measurement methods are too cumbersome to be widely adopted. The

*This work was partially supported by the National Science Foundatation (NSF) via grant CNS-0720691 and by a Dr. John N. Nicholson Fellowship.

impulse response is a complete characterization of a room's acoustics for a given pair of output and listening positions. As such, impulse responses are commonly measured in concert halls and recording studios. These measurements serve as an important diagnostic tool for acoustical engineers seeking to improve the "sound" of the space. However, accurate impulse response measurements are difficult to make. By definition, an impulse is very short, lasting perhaps several milliseconds. Yet, it must contain sufficient energy to overwhelm the background noise even after the impulse's energy is spread over several seconds of reverberation. So, a very high output power is needed from the impulse source; shots from starter pistols and popped balloons are typically used by acoustical engineers.

This work seeks to use recordings of hand claps as the basis for a room's acoustic characterization. As impulses, hand claps are inferior to balloon pops and pistol shots. They have lower energy, longer duration, and more inconsistency. The methods outlined in this paper mitigate these factors, allowing for consistent acoustic measurements. Thus, we open the door to truly easy acoustic measurement: just start recording, clap a few times, and our fully-automated toolchain gives you a room characterization.

1.1. Related work

The use of approximate acoustical impulses for room measurement has been well-studied in the acoustical engineering community, and acousticians have adopted impulse-generation techniques which suit the professional engineer. To measure reverberation time, the International Organization for Standardization (ISO) requires that the source have output power 45 dB above the background level [1]. Typically, a special-purpose, high-power, dodecahedral loudspeaker is used. Beranek suggests using pistol shots or electrical discharges as impulse sources [2]. A more sophisticated approach is the integrated impulse response method [3]. By spreading the stimulus over a long time period, this method eliminates the need for a source with high power output: any omnidirectional loudspeaker can be used. For our purposes, however, this method is not easy enough; a loudspeaker must be brought into the room and a long white noise sound must be played. We would like the user to be able to simply walk into a room carrying a handheld recorder or smartphone, clap several times and then move on to the next room.

Griesinger gives a lengthy discussion of the disadvantages of each typical impulse sound source [4]. Regarding hand claps he concludes that their easy availability and non-threatening nature (compared to pistol shots) make them a useful option. His measurements also show that claps are more uniformly omnidirectional than are balloons. Pätynen et al. measured balloon impulses in more detail [5]. Horvat et al. study balloons and firecrackers [6]. Sumarac-Pavlovic et al. use a wooden clapping device as an impulse [7].

This work is part of a larger effort at Northwestern University to explore acoustical measurement on smartphones. One of the authors has released an iOS app called "Batphone" which uses measurement of the ambient sound spectrum to recognize previously-visited places [8].

1.2. Acoustic features

The acoustic features that we extract in this paper are reverberation time, frequency response, and frequency decay. These three features are correlated with the perceived sense of space, warmth, and clarity. An accurate measurement of each of these would allow the amateur recording engineer to evaluate the suitability of a room for recording and to alter the room as needed. For example, if the room has a short reverberation time, it is suited to studio work. Frequency response allows the engineer to identify how warm or bright the sound is, and correct accordingly. For studio work, a flat frequency response is ideal. Finally, frequency decay tells the engineer how clear the sound is. A substantially longer decay time in lower frequencies results in a "muddier" sound that is far from ideal. Rooms can be fixed in a variety of ways. Denser material lining the walls can result in shorter reverberation times, and other materials can alter which frequencies are absorbed and which are emphasized.

2. METHOD

In this section we assume that, as input, we have a single recording of several hand claps recorded in an otherwise silent room. We assume that the claps are well-spaced so that their reverberations do not overlap but we otherwise have no knowledge of the times at which the claps occurred. In this section, we

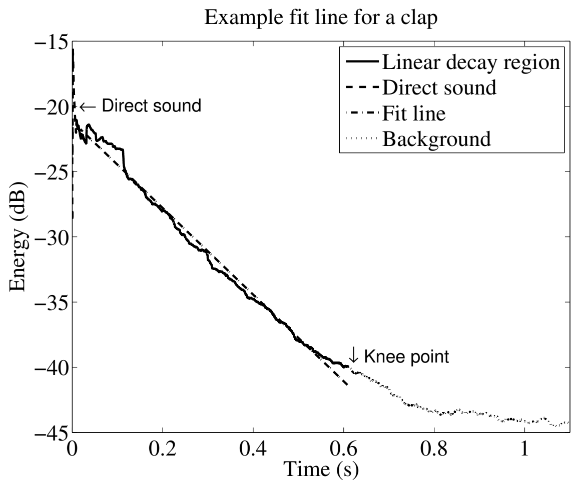

Fig. 1: Reverberation time calculation using the line-fitting and extrapolation method. The plots shows the clap's energy envelope, divided into three regions.

describe how we automatically extract three acoustic features from such a recording.

2.1. Segmentation

Our first task is to identify the onset time of each clap and to split the input recording into a series of individual clap recordings. We accomplish this by simply using an energy threshold. We split the recording into 10 ms windows and calculate the energy in each. Scanning through these windows, we recognize a new clap onset when the energy level exceeds ten times the background level and we recognize the end of a that clap when the energy level returns below twice the background level.

The background energy level mentioned above can be calculated explicitly or implicitly. In the explicit approach, a recording of the quiet room is needed. For example, in the smartphone implementation, the app would prompt the user to record a sample of the quiet room. In the implicit approach, we use the minimum or 5th-percentile value from the distribution of the energy levels observed in all 10 ms windows. In our experiments we simply assume that the first 100 ms of the recording represents the background level; so the background level is measured explicitly.

2.2. Reverberation time

In architectural acoustics, reverberation time is defined as the time it takes for an impulse to decay 60 dB, called the RT_{60} [2]. It is impractical to produce a sound 60 dB above the background level, so RT_{60} is calculated by finding the decay slope and extrapolating to 60 dB, as shown in Figure 1; the assumption here is that the sound level decays with an approximately constant rate.

We adopt a line-fitting method to calculate RT_{60}, as follows. We first calculate the signal energy envelope by squaring the samples, smoothing with a window of 10 samples, then converting to decibels; this produces the signal shown in Figure 1. The next task is to fit a decay line to the energy envelope, as shown again in Figure 1. When fitting the line, we omit the first 10 ms of the envelope because this early region contains what is known as the "direct sound", which has a higher rate of decay. Thus, the fit line starts at 10 ms and ends at some unknown "knee" point where the linear decay region ends.

The fundamental problem for calculating RT_{60} is determining which time to label as the knee. A choice of knee too early will result in little data being used for the linear fit, and thus causing the slope calculation to be innacurate. Choosing the knee too late will cause the decay slope to be too shallow. In our approach, We try all possible knee points starting 100 ms after the direct sound and choose the point which minimizes the RMS-error-per-second of the fit line. We found that fitting reliability improved after adding a constraint requiring that the knee be some reasonable minimum distance from the beginning (one third of the expected decay region length).

2.3. Frequency decay

A common acoustic measurement is the reverberation time in various frequency bands, which we call the frequency decay. To compute frequency decay, we use the same method for reverberation time described above; the only difference is that we now supply the energy envelopes for the various frequency bins. We get these energy envelopes from the rows of the clap's power spectrogram matrix. In particular, we calculate the spectrogram using 10 ms Hamming windows restricted to the half-octave frequencies of interest.

2.4. Frequency response

We define room frequency response to be the ratio of the reverberant sound power spectrum to the direct sound power spectrum. Recall that the direct sound is the sound observed in the early part of the clap. This is the sound which travelled directly from the source to the microphone without any reflections. Thus, if the source sound is very short, that first 50 ms window should contain the unaltered, "dry" source sound. The reverberant sound is that in the decay region of Figure 1. We compute power spectra using Welch's method using 50%-overlapping, 10 ms rectangular windows and Goertzel's algorithm (rather than an FFT) for restricting it to the log-spaced frequencies of interest. After computing the power spectra in the direct sound and decay regions, we convert each into energy spectra by dividing by the region's length. Finally, divide every element of the reverberant sound energy spectrum vector by the corresponding element in the direct sound energy vector to obtain the frequency response vector.

2.5. Availability

To simplify the presentation, many details have been omitted from the method descriptions above. All the details can be found in the toolchain's MATLAB code, which is available on the author's website `http://claps.seeth.org`.

We are also developing a smartphone app called "ClapIR" for Apple's iOS platform (compatible with iPhone, iPad, and iPod Touch hardware). The app will implement the entire measurement toolchain described in this paper and our hope is that it will bring acoustic measurement capabilities to a broad audience.

3. EXPERIMENTS

To validate our acoustic measurement methods, we recorded a series of 19 claps in Pick-Staiger concert hall at Northwestern University. We set up a Zoom H4n handheld recorder in the center of row E on the main floor, and clapped 19 times from center stage. The frequency response of the recorder's microphone is nearly flat, so we ignore its effect in this paper. The clap power observed at the microphone averaged 26.4 dB above the background level with a standard deviation of 4.4 dB. Pick-Staiger is a 989-seat hall, ideal for classical and symphonic works. The hall can be put into two configurations, chamber and symphonic. The former consists of acoustic panelling to make the hall smaller and more focused. The latter configuration is the one our dataset is drawn from. We also recorded an additional 12 claps in another concert hall at Northwestern University, and achieved similar results.

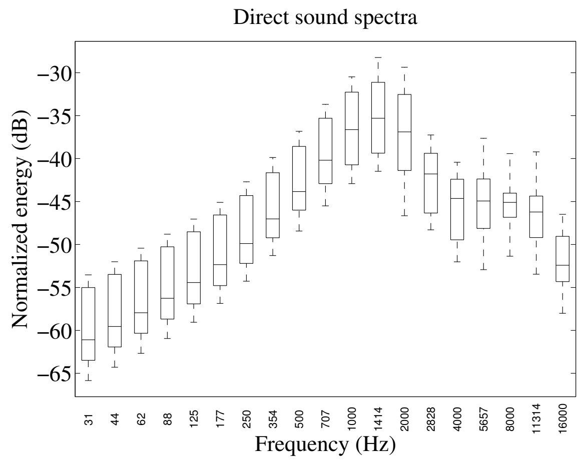

Fig. 2: Direct sound spectra for hand claps in the experiment. The box plot shows the quartiles of our 19 clap samples.

The building staff commissioned an acoustical survey of Pick-Staiger in 1999. In this survey balloon pops were used as the impulse source. We obtained the report and used their reported measurements as a reference point. We evaluate our measurements by comparing to the balloon survey where possible and by simply looking for consistency among our samples in other cases.

3.1. Clap spectra

Figure 2 shows the spectra for the claps in the experiment. We observe a peak in energy at around 1 kHz with a 5 dB/octave drop when moving to higher or lower frequencies. Our results match those of Griesinger [4, Fig. 6]. An interesting feature of the clap spectra is that its shape does not vary much with changes in clap volume; notice that the minimum and maximum values have roughly equal spacing at all of the frequencies.

3.2. Reverberation time

Our hand clap recordings gave extremely consistent reverberation time measurements. The average RT_{60} was 1.74 seconds with a standard deviation of

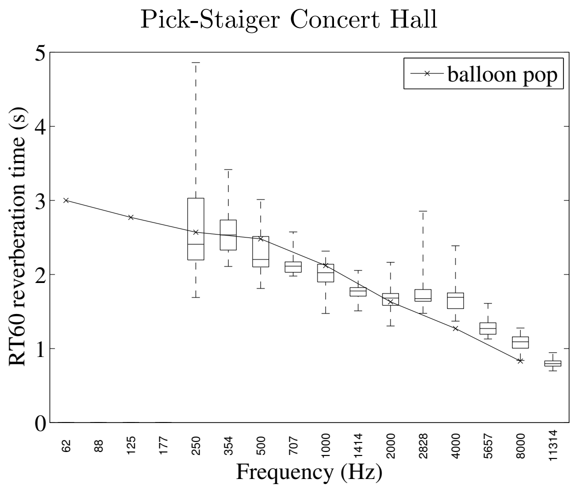

Fig. 3: Frequency-dependent reverberation time for the Pick-Staiger Hall, Northwestern University. Results were similar for hand claps and balloon pops. The box plot shows the quartiles of our 19 clap samples. The solid line plots the average results reported by the prior acoustical survey using balloons.

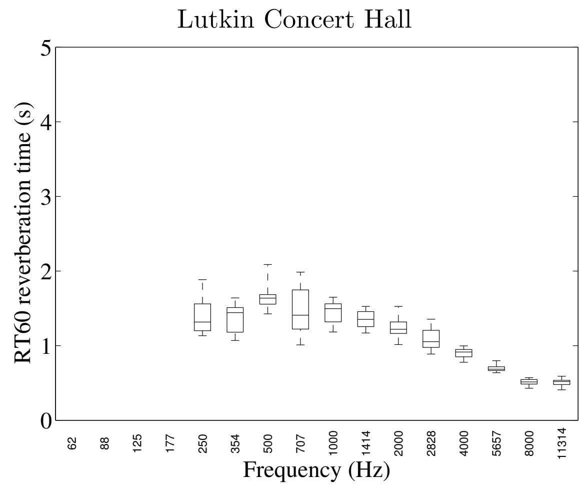

Fig. 4: Frequency-dependent reverberation time for Lutkin Hall, Northwestern University. The box plot shows the quartiles of our 12 clap samples from this hall.

just 0.07 seconds. The high consistency of reverberation times, despite the variation in individual claps, supports our proposal to use them as impulse sources.

3.3. Frequency decay

Figure 3 plots the RT_{60} reverberation times across frequency half-octaves. Frequencies above 250 Hz gave very consistent results, while lower frequencies were unreliable. Those below 250 Hz did not exhibit sufficient energy above the background level to meaningfully measure the decay slope. This is validated by Figure 2 which shows that claps lack low-frequency energy.

Note that balloons and other impulse sources also lack low-frequency content. However, claps seems to provide about an octave less low-frequency bandwidth than do balloons. The solid line in Figure 3 shows the results obtained by the balloon acoustic survey. Our clap results were very consistent with the prior balloon results, except at low frequencies where the clap lacked sufficient energy to compute a reasonable reverberation time.

Other rooms To demonstrate that hand clap measurements are useful in a variety of rooms, we made recordings two additional rooms: Northwestern University's Lutkin Hall (a 400-seat concert venue) and a makeshift home recording studio. Figures 4 and 5 show very consistent frequency decay measurements within these rooms. Of course, the studio exhibited much shorter reverberation times than the concert halls, as expected. We conjecture that clap-based measurements are more reliable in small rooms because relatively little energy is needed to excite the room.

3.4. Frequency Response

Figure 6 plots the frequency response distribution observed for our claps. As in the previous section, the measurements below 250 Hz have high variability and thus provides little confidence. However, the middle frequencies, which are most important for music, were consistent.

4. CONCLUSION

We found that the three acoustic features that we sought to measure (reverberation time, frequency decay, and frequency response) could all be reliably measured using hand claps, with some caveats. Reverberation times and frequency responses below 300 Hz were not reliable, due to the lack of low-

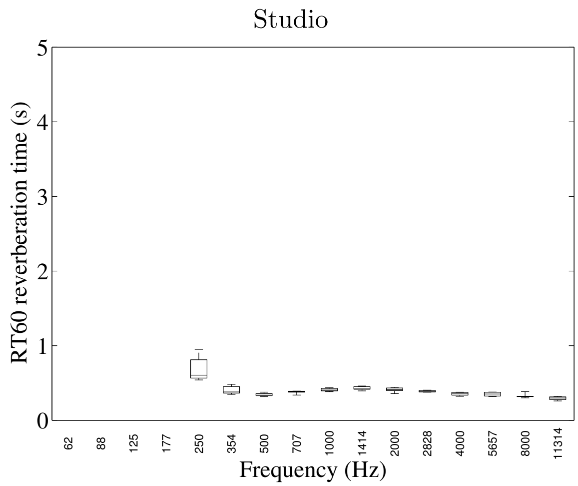

Fig. 5: Frequency-dependent reverberation time for a small studio. The box plot shows the quartiles of our 6 clap samples. In this small room, the hand-clap results are very consistent.

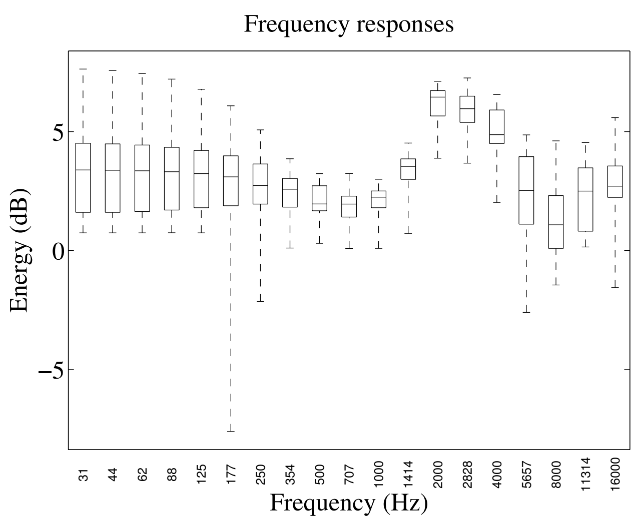

Fig. 6: Computed frequency response for Pick-Staiger Hall. The box plot shows the quartiles of our 19 clap samples.

frequency energy in hand claps. However, this is a problem that also exists for other impulse sources such as balloons, so we do not see this as a barrier to adopting hand claps.

Our current task is to extend this work by building useful acoustic measurement tools which can be used even by recording novices. Our results show that such tools can be build on the basis of hand-clap acoustic impulses. The end product will be implemented via a smartphone app which prints plots such as those in this paper after being supplied with a few hand-clap examples. We believe that such tools would be incredibly valuable to musicians and we hope that the automated signal processing steps that we have implemented will enable this and other novel tools.

5. REFERENCES

[1] Intl. Organization for Standardization, Geneva, Switzerland, *ISO 3382:1997 Acoustics – Measurement of the reverberation time of rooms with reference to other acoustical parameters*, 1997.

[2] Leo L. Beranek, "Measurement of the acoustic properties of rooms, studios, and auditoriums," in *Acoustical Measurements*, pp. 781–791. American Institute of Physics, revised edition, 1988.

[3] M. R. Schroeder, "Integrated-impulse method measuring sound decay without using impulses," *J. Acoustical Society of America*, vol. 66, no. 2, pp. 497–500, 1979.

[4] David Griesinger, "Beyond MLS – occupied hall measurement with FFT techniques," in *In 101st Convention of the Audio Engineering Society*, Nov. 1996.

[5] Jukka Pätynen, Brian F.G. Katz, and Tapio Lokki, "Investigations on the balloon as an impulse source," *J. Acoustical Society of America*, vol. 129, no. 1, pp. EL27–EL33, 2011.

[6] Marko Horvat, Kristian Jambrosic, and Hrvoje Domitrovic, "A comparison of impulse-like sources to be used in reverberation time measurements," *J. Acoustical Society of America*, vol. 123, no. 5, pp. 3501, 2008.

[7] Dragana Sumarac-Pavlovic, Miomir Mijic, and Husnija Kurtovic, "A simple impulse sound source for measurements in room acoustics," *Applied Acoustics*, vol. 69, no. 4, pp. 378–383, Apr. 2008.

[8] Stephen P. Tarzia, Peter A. Dinda, Robert P. Dick, and Gokhan Memik, "Indoor localization without infrastructure using the acoustic background spectrum," in *Proc. Intl. Conf. on Mobile Systems, Applications, and Services (MobiSys)*, June 2011, pp. 155–168.