ISSN: 2278 – 7798
International Journal of Science, Engineering and Technology Research (IJSETR)
Volume 2, Issue 8, August 2013

A Review of Different Approaches Applied For The Estimation of Reverberation Time

Pallavee Jaiswal
M.Tech Scholar
Dr. C.V. Raman University

Dr. Moushmi Kar
Department of ECE
Dr. C.V. Raman University

***Abstract*— The reverberation time is one of the most prominent acoustic characteristics of an enclosure. Its value can be used to predict speech intelligibility, and is used by speech enhancement techniques to suppress reverberation. Speech enhancement aims to improve speech quality, intelligibility and degree of listener fatigue. The reverberation time is estimated by the two different ways. (1) With the knowledge of room acoustics. (2) Without the knowledge of room acoustics. Estimating the RT using only the observed reverberant speech signal i.e, blind estimation is required for speech evaluation and enhancement technique. Many blind methods have been developed. This is a review paper and its objective is to provide an overview of the verity of blind estimation of RT that has been propose for the speech enhancement method. Section – 1 give the introduction. Section – 2 give the reverberation. Section – 3 give the various reverberation time estimation methods. Section – 4 give the method for blind T60 estimation. Section – 5 conclusion.**

***Index Terms*— Blind estimation, Reverberation time, Room acoustics,**

I. INTRODUCTION

When speech is produced in an enclosed environment, the acoustic signal follows multiple paths from source to receiver, resulting in reverberations. Reverberant signals sound distant and suffer from perceptual artifacts such as coloration and echoes. With the advances in hands-free telephony, reverberation has become a burden, in particular, to applications such as automatic speech recognition (ASR) and hearing aids. Quantifying reverberation is not easy and, commonly, the so called reverberation time (RT) is used. RT, by definition, is the interval required for the sound energy to decay by 60dB after the sound source is turned off. Larger RT results in speech signals with decreased quality and intelligibility. Traditionally, room impulse responses or room geometry and absorptive properties are used to measure RT. The reverberation time is one of the important features for the speech enhancement method. By speech enhancement, it refers not only to noise reduction but also to dereverberation and separation of independent signals. In this paper, we present historical review of different approaches used for the estimation of reverberation time.

II. REVERBERATIION

Reverberation is the process of multi-path propagation of an acoustic signal from its source to the microphone. The received signal generally consists of a direct sound, reflections that arrive shortly after the direct sound (commonly called early reverberation)

and reflections that arrive after the early reverberation (commonly called late reverberation). The combination of the direct sound and early reverberation is sometimes referred to as the early sound component. The different sound components will now be discussed in more detail.

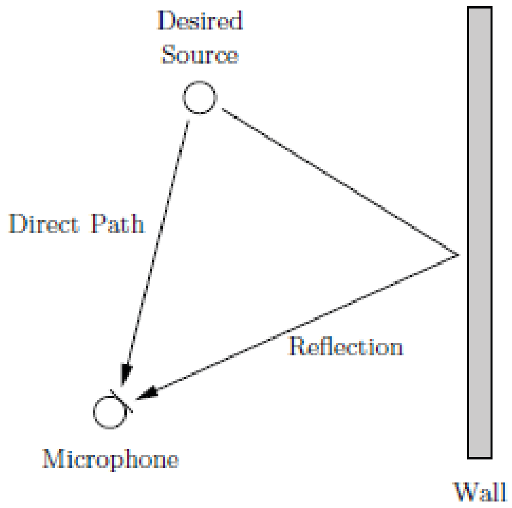

Figure 2.1 Illustration of the direct path and a single reflection from the desired source to the microphone.

Direct Sound: The first sound that is received through free-field, i.e., without reflection, is the direct sound. In case the source is not in line of sight of the observer there is no direct sound. The delay between the initial excitation of the source and its observation is dependent on the distance and the velocity of the sound.

Early Reverberation: A little time later the sounds which were reflected off one or more surfaces (walls, floor, furniture, etc.) will be received. These reflected sounds are separated in both time and direction from the direct sound. The reflected sounds form a sound component which is usually called early reverberation. Early reverberation will vary as the source or the microphone moves within the space, and gives us information about the size of the space and the position of the source in the space. Early reverberation is not perceived as a separate sound to the direct sound so long as the delay of the reflections does not exceed a limit of approximately 80-100 ms with respect to the arrival time of the direct sound. Early reverberation is actually perceived to reinforce the direct sound and is therefore considered useful with regard to speech intelligibility. This is often referred to as the precedence effect. This reinforcement is what makes it easier to hold conversations in closed rooms compared with outdoors. Early reverberation is mainly important in so-called small-room acoustics since the walls, ceiling and floor are really close. Early reverberation also causes a spectral distortion called coloration.

Late Reverberation: Late reverberation results from reflections which arrive with larger delays after the arrival of the direct sound. They are perceived either as separate echoes, or as reverberation, and impair speech intelligibility.

There are two possible way to estimate the reverberation time. (1) With the knowledge of room acoustics, and (2)Without the knowledge of room acoustics. A number of reverberation compensation approaches require some knowledge about the room reverberation parameters. This becomes an issue with the practical deployment of the algorithm in unknown environments. It required to propose completely blind reverberation compensation algorithms that do not require any knowledge about reverberation parameters. Instead of guiding our dereverberation optimization problems with prior knowledge about reverberation parameters, we need to propose and successfully guide our optimizations using generic speech knowledge in terms of its feature auto-correlation sequences, feature sparsity, and feature probability distributions.

III. REVERBERATION TIME ESTIMATION METHODS

There are various method proposed for the estimation of reverberation time which are required for the speech enhancement technique.(1) with the knowledge of room acoustics, and (2) without the knowledge of room acoustics.

(1) With the knowledge of room acoustics.

The reverberation time play a important role in speech enhancement methods. Earlier all the method is primarily concerned with the modeling the impulse response of a real room, it is instructive to consider room acoustic modeling method which simulate the impulse response of a room. A brief overview of various room acoustic modeling method is presented. Mathematically the sound propagation is described by the wave equation. An impulse response from a source to a microphone can be obtained by solving the wave equation. Since it can seldom be expressed in an analytic form the solution must be approximated. There are three main modeling methods, as illustrated in Fig. 3.1, viz., wave based, ray-based and statistical [19]. The ray-based methods, such as the ray-tracing [20] and the image-source method [10], are the most often used. The wave-based methods, such as the Finite Element Method (FEM), Boundary Element Method (BEM) [21, 22] and Finite-Difference Time-Domain (FDTD) [23] methods, are computational more demanding. In real-time auralization4 the limited computation capacity requires simplifications. A frequently used simplification consists of modeling the direct path and early reflections individually and the late reflections by recursive digital filter structures. The statistical modeling methods, such as the Statistical Energy Analysis (SEA), have been widely used in aerospace, ship and automotive industry for high frequency noise analysis and acoustic designs. They are not suitable for auralization purposes since those methods do not model the temporal behavior of a sound field.

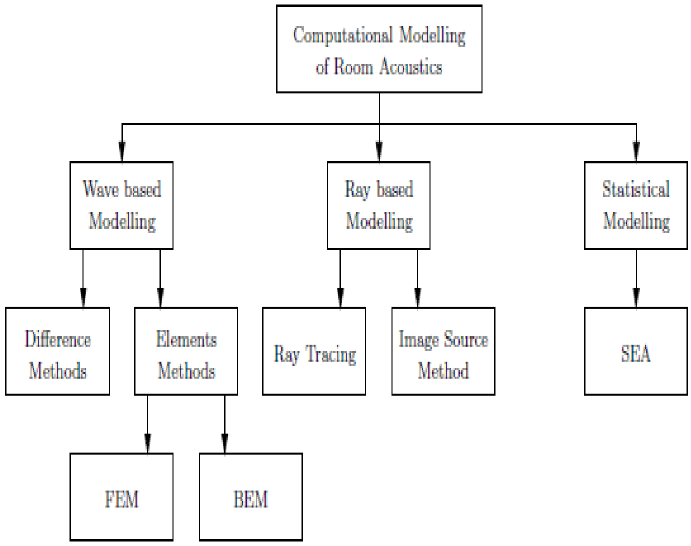

Figure 3.1 Room acoustic models are based on sound rays (ray-based), on solving the wave equation (wave-based) or some statistical method [19]

Wave- based methods

The most accurate results can be achieved by the wave-based methods. An analytical solution for the wave equation can be found only in extremely simple cases such as a rectangular room with rigid walls. Therefore, numerical methods such as FEM and BEM [22, 21] are often used. The main difference between these two element methods is in the element structure. In FEM, the space is divided into volume elements, while in BEM only the boundaries of the space are divided into surface elements. The elements interact with each other according to the basics of wave propagation. The size of these elements has to be much smaller than the size of the wavelength for every particular frequency. At high frequencies, the required number of elements becomes very high, resulting in a large computational complexity. Therefore, these methods are suitable only for low frequencies and small enclosures.

Another method for room acoustics simulation is provided by the FDTD method [23, 24]. The main principle of this method is that derivatives in the wave equation are replaced by corresponding finite differences. The FDTD method produces impulse responses that are better suited to auralization than FEM and BEM. The main benefit of the element methods over FDTD methods is that one can create a denser mesh structure where required, such as locations near corners or other acoustically challenging places.

In all the wave-based methods, the most difficult part is the definition of the boundary conditions. Typically a complex impedance is required, but it is hard to find that data in existing literature.

Ray-based methods

The ray-based methods are based on geometrical room acoustics [18].The most commonly used ray-based methods are the ray-tracing [20] and the image method [10]. The main difference between these methods is the way the reflection paths are calculated [19]. To model an ideal impulse response from a source to a receiver all possible sound reflection paths, commonly called rays, should be discovered. In ray-tracing methods the sound power emitted by a sound source is described by a finite number of rays. These rays propagate through space and are reflected after every collision with the room boundaries. During that time, their energy decreases as a consequence of the sound absorption of the air and of the walls involved in the propagation path. When the rays reach the receiver, an energy calculation process is performed. When all rays are processed the impulse response is obtained. Rays can be selected at random, based on a fixed interval or restricted to a given range of angles. Due to this the ray-tracing methods are by no means exhaustive, whereas the image method finds all the rays. However, while the image method is limited to geometries that are formed by planer surfaces the ray-tracing method can be applied to geometries that are formed by arbitrary surfaces.

It should be mentioned that all ray-based methods are based on energy propagations. This means that all effects involving phase differences such as refraction or interference are neglected. This is admissible if the sound signals of interest are not sinusoids or other signals with small frequency bandwidth but are composed of many spectral components covering a wide frequency range. Then it can be assumed that constructive and destructive phase effects cancel each other when two or more sound field components superimpose at a point, and the total energy in the considered point is simply obtained by adding their energies. Components with this property are often referred to as mutually incoherent [25]. The image method, which was developed by Allen and Berkley in 1979, is probably one of the methods most commonly used in the acoustic signal processing community.

(2) Without the knowledge of room acoustics.

The reverberation time, T60, is one of the key parameters used to quantify room acoustics. It can provide information about the quality and intelligibility of speech recorded in a reverberant environment, and it can be used to increase robustness to reverberation of speech processing algorithms. T60 can be determined directly from a measurement of the acoustic impulse response, but in situations where this is unavailable it must be estimated blindly from reverberant speech.

IV. METHODS FOR BLIND T60 ESTIMATION

1. Method 1: Spectral Decay Distributions (SDD)

The method proposed by Wen et al. [8] is based on the spectral decay distributions of the observed speech and assumes a statistical model for the AIR. Frequency dependent decay rates are estimated by applying a least squares linear fit to the log-energy envelope in each frequency band in the Discrete Fourier Transform (DFT) domain. The negative-side variance of the distribution of the decay rates is demonstrated to correlate with the room decay rate and is, thus, used to predict reverberation time. This approach requires training in order to map the values from the negative-side variance to T60. In the training phase, a 2nd-order polynomial mapping function is calculated using reverberated speech with known T60.

2. Method 2: Modulation Energy Ratios (MER)

Falk and Chan [17] proposed a non-intrusive quality measure for dereverberated speech based on the Speech-to- Reverberation Modulation energy Ratio (SRMR). The method considers the energy in eight modulation frequency bands, varying logarithmically between 4 and 128 Hz and calculated from 23 acoustic frequency bands obtained from a gamma tone filter bank. It is observed that the low modulation frequency energy (4 -18 Hz) is relatively insensitive to reverberation while the energy at high modulation frequencies (29 - 128 Hz) increases almost linearly with T60. This leads to the SRMR measure which is the ratio of the average energy in the low modulation frequencies to the high modulation frequencies. It is also shown that the inverse of the SRMR is highly correlated with T60. Obtaining the values for T60, requires some form of training and, similarly to Method 1, a 2nd-order polynomial mapping function is calculated from reverberant speech with known T60.

3. Method 3: Maximum Likelihood (ML)

The method proposed by L¨ollmann et al. [7] is inspired by the method from Ratnam et al. [6]. It uses a statistical model of the sound decay of reverberant speech, following a reverberation model similar to that of Method 1. This is then used to develop a Maximum Likelihood (ML) approach for the T60 estimation. In order to improve the computational efficiency, the speech signals are down sampled before the estimation and there is a pre-selection approach to detect plausible decays before these are used in the ML estimation procedure. Furthermore, the estimated T60 for each frame is used in a histogram and smoothing procedure in order to increase the robustness of the estimates. This algorithm has also an option for a fast tracking of the T60. However, tracking is not considered in this evaluation. Unlike the previous two methods this method does not require training.

V. CONCLUSION

This is a review paper and various reverberation time estimation method were described which required for implementing the speech enhancement technique. There are two types of method for estimating the reverberation time.1) with the use of room acoustics 2) without the use of room acoustics. Evaluation of method which uses room acoustics revealed that this method is very complicated and require more computation. The methodology involved in estimating the reverberation time using room acoustics is quite complex. The second method i.e. blind estimation of reverberation time is very easily evaluate as it only required the reverberant signal, it reduces the calculation portion. Blind estimation are more suitable than the method require knowledge of room acoustics.

VI. REFERENCE

[1] W. C. Sabine, Collected Papers on acoustics (Originally 1921), Peninsula Publishing, 1993.

[2] Carl F. Eyring, “Reverberation time in “dead” rooms,” J. Acoust. Soc. Am., vol. 1, no. 2, pp. 217–241, Jan. 1930.

[3] E. A. P. Habets, Single- and Multi-Microphone Speech Dereverberation using Spectral Enhancement, Ph.D. thesis, Technische Universiteit Eindhoven, 2007.

[4] ISO3382, “Acoustics – measurement of room acoustic parameters,” 2009.

[5] M. R. Schroeder, “New method of measuring reverberation time,” J. Acoust. Soc. Am., vol. 37, pp. 409–412, 1965.

[6] R. Ratnam, D. L. Jones, B. C. Wheeler, W. D. O’Brien, Jr., C. R. Lansing, and A. S. Feng, “Blind estimation of reverberation time,” J. Acoust. Soc. Am., vol. 114, no. 5, pp. 2877–2892, Nov. 2003.

[7] H. W. L¨ollmann, E. Yilmaz, M. Jeub, and P. Vary, "An improved algorithm for blind reverberation time estimation," in Proc. Intl. Workshop Acoust. Echo Noise Control (IWAENC), Tel-Aviv, Israel, Aug. 2010.

[8] J. Y. C. Wen, E. A. P. Habets, and P. A. Naylor, "Blind estimation of reverberation time based on the distribution of signal decay rates," in Proc. IEEE Intl. Conf. on Acoust., Speech, Signal Process. (ICASSP), Las Vegas, USA, Apr. 2008.

[9] T. de M. Prego, A. A. de Lima, S. L. Netto, B. Lee, R. W. Schafer, and T. Kalker, "A blind algorithm for reverberation time estimation using subband decomposition of speech signals," J. Acoust. Soc. Am., vol. 131, no. 4, pp. 2811–2816, Apr. 2012.

[10] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," J. Acoust. Soc. Am., vol. 65 no. 4, pp. 9 3–950, Apr. 1979.

[11] T. H. Falk and W.- Y. Chan, "Temporal dynamics for blind measurement of room acoustical parameters," IEEE Trans. Instrum. Meas., vol. 59, no. 4, pp. 978–989, Apr. 2010.

[12] Lebart K, Boucher J M, "A New method based on spectral subtraction for speech enhancement" Acta acustica, Acustica vol. 87 pp359-366. 2001.

[13] M. R. Schroeder, "New Method of Measuring Reverberation Time," *Journal of the Acoustical Society of America*, vol. 37, pp. 409–412, 1965.

[14] T. J. Cox, F. Li, and P. Dalington, "Extracting Room Reverberation Time From Speech Using Artificial Neural Networks," *Journal of the Acoustical Society of America*, vol. 49, no. 4, pp. 219–230, 2001.

[15] M. Wu and D. Wang, "A Pitch-Based Method for the Estimation of Short Reverberation Time," *Acta Acustica United With Acustica*, vol. 92, pp. 337–339, 2006.]

[16] H. W. L¨ollmann and P. Vary, "Estimation of the Reverberation Time in Noisy Environments," in *Proc. of Intl. Workshop on Acoustic Echo and Noise Control (IWAENC)*, Seattle (Washington), USA, Sept. 2008.

[17] T. Falk, C. Zheng, and W.-Y.Chan, "A non-intrusive quality and intelligibility measure of reverberant and dereverberated speech," IEEE Trans. Audio, Speech, Lang. Process., vol. 18, no. 7, pp. 1766–1774, Sept. 2010.

[18] H. Kuttruff, Room Acoustics, Spon Press, London, fourth edition, 2000.

[19] L. Savioja, J. Huopaniemi, T. Lokki, and R. V¨a¨an¨anen, "Creating interactive virtual acoustic environments," Journal of the Audio Engineering Society, vol. 47, no. 9, pp. 675–705, 1999.

[20] A. Kulowski, "Algorithmic representation of the ray tracing technique," Applied Acoustics, vol. 18, no. 6, pp. 449–469, 1985.

[21] M. Kleiner, B.I. Dalenb¨ack, and P. Svensson, "Auralization - an overview," Journal of the Audio Engineering Society, vol. 41, no. 11, pp. 861–875, Nov. 1993.

[22] A. Pietrzyk, "Computer modeling of the sound field in small rooms," in Proc. of the 15th AES Int. Conf. on Audio, Acoustics & Small Spaces, Copenhagen, Denmark, Oct. 1998, vol. 2, pp. 24–31.

[23] D. Botteldoore, "Finite-difference time-domain simulation of low-frequency room acoustic problems," Journal of the Acoustical Society of America, vol. 98, no. 6, pp. 3302– 3308, 1995.

[24] L. Savioja, J. Backman, A. J¨arvinen, and T. Takala, "Waveguide mesh method for low-frequency simulation of room acoustics," in Proc. of the 15th Int. Congr. Acoust. (ICA'95), Trondheim, Norway, June 1995, vol. 2, pp. 1–4.

[25] M.J. Crocker, Handbook of Acoustics, Wiley-Interscience, 1998.