Applied Acoustics 117 (2017) 167–172

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Technical note

Effects of simultaneous sound arrivals on direction-of-arrival estimates of the polar energy time curve

Brian T. Thornock [1], Timothy W. Leishman *, Brian E. Anderson, J. James Esplin [2]

Acoustics Research Group, Department of Physics and Astronomy, Brigham Young University, N283 Eyring Science Center, Provo, UT 84602, United States

A R T I C L E I N F O

Article history:
Received 3 March 2016
Received in revised form 16 August 2016
Accepted 7 October 2016
Available online 8 November 2016

Keywords:
Reflection
Simultaneous
Directional
Polar ETC
Impulse response
Room acoustics

A B S T R A C T

Acoustic reflections in rooms are best understood and controlled when their causal surfaces have been accurately localized. Several in situ measurement techniques have been proposed over the years to pinpoint reflections via estimated directions of arrival at chosen field positions. The polar energy time curve (Polar ETC) is one approach requiring six cardioid impulse response measurements along the Cartesian axes about a point in a room. This paper presents a theoretical, numerical, and experimental investigation of its estimation errors due to simultaneous sound arrivals. Theoretical developments provide insights into their causes. Numerical examples based on the method of images explore conditions leading to the errors and likely error magnitudes. Experimental work using discrete planar reflectors in an anechoic chamber produced controlled reflections to further investigate the effects. An altazimuth-mounted laser and mirrors enabled reflection tracing and a means of measuring and comparing arrival directions to those estimated by the Polar ETC. The efforts demonstrate that the Polar ETC provides credible angular estimations for individual arrivals, but spurious results for most simultaneous arrivals. Single arrivals produce absolute errors of only a few degrees, while simultaneous arrivals of similar amplitudes produce errors ranging sporadically from a few degrees to many degrees.

1. Introduction

Many methods have been devised over the years to estimate the directions of arrivals of reflected waves in architectural acoustics measurements. One approach is the polar energy time curve (Polar ETC), which requires six cardioid impulse responses (IRs) along the Cartesian axes about a field point [1,2]. This is of practical interest to acousticians and audio engineers because a single cardioid microphone may be used for the measurements, with sequential orientations in each Cartesian direction. The angles of arrivals are subsequently calculated using ETCs (envelopes) of the six IRs—the magnitudes of their complex time signals including the IRs as real parts and Hilbert transforms as imaginary parts [3–6].

Researchers initially understood the ETC to be proportional to either instantaneous energy density or sound intensity. Becker and others [2,3] used it under these assumptions. However, it has since been shown to be based on an acausal operation and to not correctly represent energy flow [7,8]. Moreover, its components are not quadratic in the linear field variables, i.e., not proportional to squared acoustic pressure or particle velocity, and cannot be considered energy quantities. Other names have subsequently been used to describe the ETC, including “time response” [9] and “envelope time curve” [10]. Regardless, the Polar ETC does allow angular estimations of sound arrivals and has been widely used for that purpose. Recently, Esplin et al. provided insights into the method, including observations about noncardioid microphone directivities and background noise [11]. This paper provides additional insights through an investigation of the underlying assumption that no more than one sound arrival occurs during a given sample of its six digitized IRs [2].

Other direction-of-arrival estimation methods have also arisen over the years. For example, Yamasaki et al. [12] and Sekiguchi et al. [13] developed cross-correlation methods for an array of four microphones. Gover et al. introduced a complex alternative involving a 32-microphone spherical array with a beamforming algorithm [14–16]. Choi et al. employed a five-microphone array with one microphone at the geometric center of a tetrahedral array [17]. Noël et al. added intricacy to the arrangement by using up to 15 microphones and more involved matrix calculations [18]. Abdou et al. used intensity measurements to characterize

* Corresponding author.
E-mail address: tim_leishman@byu.edu (T.W. Leishman).
[1] Current address: Cirrus Logic, 800 West 6th Street, Austin, TX 78701, United States.
[2] Current address: Labcyte, Inc., 1190 Borregas Avenue, Sunnyvale, CA 94089, United States.

http://dx.doi.org/10.1016/j.apacoust.2016.10.004
0003-682X/

directional room response characteristics [19], while Essert [20] and Farina et al. [21,22] employed Ambisonics microphone arrays with intensity-based schemes. Each of these methods has its benefits and drawbacks, as does the Polar ETC. However, since the latter requires only a single microphone and analyzer channel, and has been commercially available for many years, it continues to be a practical, cost-effective tool.

The purpose of this paper is to explain errors that arise with the Polar ETC when multiple reflected waves arrive simultaneously at the microphone position (within one sample). It typically cannot resolve their individual arrival angles. To the authors' knowledge, no other study has reported or clarified this issue. Of the alternative methods described in the previous paragraph, only those presented by Gover et al. [14] and Choi et al. [17] attempted to resolve individual angles of simultaneous arrivals. Their methods employed 32 and 5 microphones, respectively, providing at least 4 microphones that did not detect any two reflections simultaneously. They also described qualitatively why other techniques fail in this regard. This paper uses detailed theoretical developments and several computational and experimental examples to demonstrate how the underlying single-arrival assumption of the Polar ETC is unsatisfactory. The issue is particularly relevant at later IR times when temporal densities of reflection arrivals are high, leading to common angular estimation errors. The paper also explains how the effects manifest themselves according to relative arrival strengths for simple cases.

2. Theoretical developments

While the single-arrival assumption has been used in many direction-of-arrival estimation methods, an illustrative example demonstrates its shortcomings. Consider an elongated rectangular room with dimensions 30.0 m × 8.0 m × 4.5 m and wall, floor, and ceiling absorption coefficients of $\alpha_w = 0.28$, $\alpha_f = 0.51$, and $\alpha_c = 0.37$, respectively. Its IR was simulated using the Allen-Berkley method of images [23], which calculates the IR between two points in a rectangular room as a sum of contributions from numerous image sources. It characterizes geometric spreading and absorption of six room walls, each of which may have a distinct absorption coefficient. For this work, the method was adapted from Fortran to Matlab [24]. The omnidirectional source and receiver positions were located at coordinates (25.0 m, 1.5 m, 1.25 m) and (5.0 m, 7.0 m, 2.5 m), respectively. For the given room geometry, these positions showed strong late reflections, reflection clustering, and other interesting IR properties. The sampling frequency was 192 kHz, corresponding to a sampling interval of 5.2 µs.

Figure 1 shows the IR along with a histogram revealing the number of sound arrivals per sample. The direct arrival at 60 ms closely precedes a stronger peak comprising two simultaneous arrivals. Single arrivals occur for several milliseconds thereafter. At about 90 ms, reflections from the far end of the room begin to arrive. Double arrivals become more prominent and, in one case, four simultaneous arrivals produce a conspicuous IR peak. Thus, while the single-arrival assumption is reasonable for the first 90 ms, it breaks down later as the average number of arrivals per sample clearly exceeds one and three or more simultaneous arrivals become increasingly common. Since simultaneous arrivals in this and other room IRs are very likely, their effects on direction-of-arrival estimates should not be overlooked.

If the Polar ETC employs a true cardioid microphone, as is typically assumed, its estimated elevation and azimuthal angles of incidence at the field point are ordinarily calculated as

$$\theta = \sin^{-1}\left(\frac{ETC_{+z} - ETC_{-z}}{2ETC_0}\right) \tag{1}$$

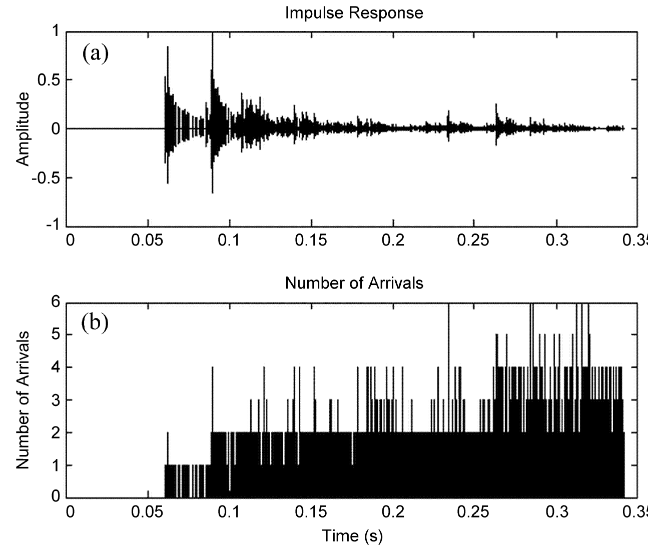

Fig. 1. The (a) impulse response and (b) arrival histogram of a numerically simulated room with dimensions 30.0 m × 8.0 m × 4.5 m, an omnidirectional source at (25.0 m, 1.5 m, 1.25 m), an omnidirectional receiver at (5.0 m, 7.0 m, 2.5 m), and nonuniform wall absorption.

and

$$\phi = \tan^{-1}\left(\frac{ETC_{+y} - ETC_{-y}}{ETC_{+x} - ETC_{-x}}\right), \tag{2}$$

respectively, where ETC_i denotes the cardioid ETC oriented towards the ith Cartesian direction (i.e., $i = \pm x, \pm y$, or $\pm z$) and ETC_0 denotes the omnidirectional ETC [2,11]. [The 2 in the denominator of Eq. (1) should be dropped when using normalized cardioid directivity functions.] The ETC itself is defined as

$$ETC[t] = \sqrt{h^2[t] + \hat{h}^2[t]}, \tag{3}$$

where $h[t]$ is the IR sampled at discrete time t and $\hat{h}[t]$ is its associated Hilbert transform.

An IR $h[t]$ with N samples and sampling frequency f_s may be defined as

$$h[t] = h[n/f_s] = \sum_{n=0}^{N-1} a_n \mathbf{e}_n, \tag{4}$$

where a_n is the amplitude of the nth sample and $\mathbf{e}_n$ is a unit impulse sequence comprising unity at the nth sample and zero at all other samples (i.e., $\mathbf{e}_n$ is a basis vector facilitating construction of the discrete-time IR).

If one assumes for simplicity that the directional IR $h_i[t]$ has a single nonzero sample consisting of M simultaneous arrivals, it may be represented as

$$h_i[t] = a_n \mathbf{e}_n = \mathbf{e}_n \sum_{m=1}^{M} A_m D_i(\theta_m, \phi_m), \tag{5}$$

where $\mathbf{e}_n$ is the unit impulse sequence for the sample at time $t = n/f_s$, A_m is the amplitude of arrival m, and $D_i(\theta_m, \phi_m)$ is the directivity function value for the arrival incident upon the microphone from angle (θ_m, ϕ_m). For an omnidirectional microphone, the directivity function is constant for all angles, while for a true cardioid microphone oriented in the $+x$ direction (as one example), the unnormalized directivity function is $D_{+x}(\theta_m, \phi_m) = (1 + \cos\theta_m \cos\phi_m)$.

The ETC for the ith direction is

$$ETC_i[t] = \sqrt{\left(\mathbf{e}_n \sum_{m=1}^{M} D_i(\theta_m, \phi_m) A_m\right)^2 + \left(\hat{\mathbf{e}}_n \sum_{m=1}^{M} D_i(\theta_m, \phi_m) A_m\right)^2}, \tag{6}$$