Wide-Area Psychoacoustic Correction for Problematic Room-Modes Using Nonlinear Bass Synthesis

ADAM J. HILL, *AES Student Member,* **AND MALCOLM O. J. HAWKSFORD,** *AES Fellow*
(ajhilla@essex.ac.uk) (mjh@essex.ac.uk)

University of Essex, School of Computer Science & Electronic Engineering, Colchester, Essex, CO4 3SQ, UK

Small room acoustics are characterized by a limited number of dominant low-frequency room-modes that result in wide spatio-pressure variations that traditional room correction systems may find elusive to correct over a broad listening area. A psychoacoustic-based methodology is proposed whereby signal components coincident only with problematic modes are filtered and substituted by virtual bass components to forge an illusion of the suppressed frequencies. Although this approach can constitute a standalone correction system, the impetus for development is for use within well-established correction methodologies. A scalable and hierarchical approach is studied using subjective evaluation to confirm uniform wide-area performance. Bass synthesis exploits parallel nonlinear and phase vocoder generators with outputs blended as a function of transient and steady-state signal content.

0 INTRODUCTION

Room-modes in small- to medium-sized closed acoustical spaces often cause wide variations in low-frequency response across a listening area. This spatiotemporal variance will result in largely different impressions of a room and/or sound system, commonly with adjacent listeners experiencing antithetical acoustical conditions.

A wealth of research exists concerning room-mode correction/suppression, including passive, active, and hybrid systems. Many well-established correction systems perform effectively in decreasing the consequences of room-modes but often fall short concerning spatial variation minimization and can require highly complex signal processing. Depending on the system configuration, there are often room-modes that are nearly impossible to fully correct, resulting in an incomplete solution to the problem at hand.

As room-modes are a physical phenomenon, it is proposed that a psychoacoustical method could strengthen a physically-based correction technique by easing the physical requirements of the system, allowing for problematic room-modes to be addressed within the psychoacoustical domain.

A system is presented in this paper whereby the "principle of the missing fundamental" (or virtual bass synthesis) is utilized to create the impression of the presence of certain narrow frequency bands, while in actuality these bands are removed from the audio signal. This substitution procedure effectively eliminates the physical reinforcement of the most problematic room-modes with the aim of reducing spatial variation across a listening area while maintaining consistent signal fidelity. This approach was developed to operate within the structure of the chameleon subwoofer array (CSA) low-frequency room-mode correction algorithm [1].

Common small-room, low-frequency correction procedures will be briefly discussed, highlighting the performance of each system followed by an analysis of the standard procedures used to implement the principle of the missing fundamental including a novel hybrid dual-generator system. The virtual bass procedure will be described in the context of room-mode suppression with results of subjective evaluation included for validation. Finally, the new methodology of targeting virtual bass synthesis to substitute just for problematic room-modes will be discussed in the context of existing low-frequency correction systems as a means of improving wide-area subjective performance. A major claim for this sparsely applied virtual bass approach is that a frequency-selective and signal-dependent application of virtual bass synthesis largely overcomes the synthetic quality that is an intrinsic hallmark of this system.

All occurrences of simulated data in this paper utilize a Finite-Difference Time-Domain (FDTD) acoustical modeling toolbox that has been described in detail in [2] and is freely available online as an open-source project [3].

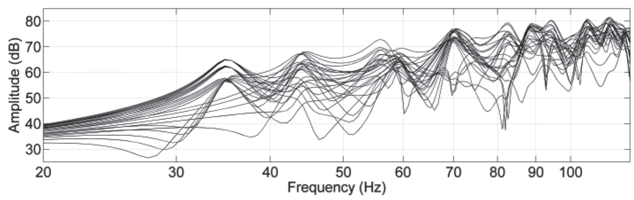

Fig. 1 Simulated frequency responses of 25 listening locations in a 5 m x 4 m x 3 m room with a subwoofer in the room corner.

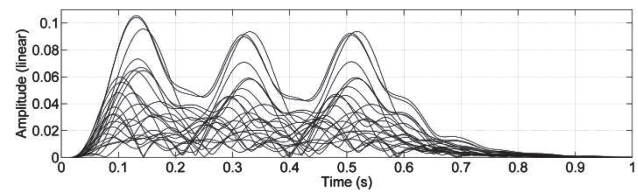

Fig. 2 Simulated 80 Hz tone burst measurements at 25 listening locations in a 5 m x 4 m x 3 m room with a subwoofer in the room corner.

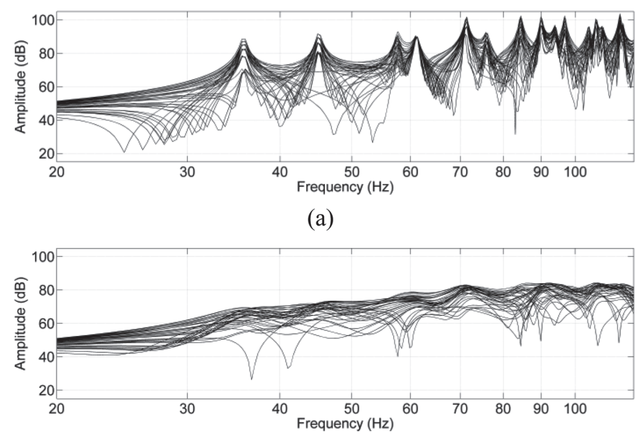

Fig. 3 Simulated frequency responses over a listening area with (a) 2% wall absorption and (b) 20% wall absorption.

1 LOW-FREQUENCY ROOM ACOUSTICS

Room acoustics are largely influenced by room-modes below the Schroeder frequency, as defined by Eq. 1 [4].

$$f_s = 2000\sqrt{\frac{RT_{60}}{V}} \tag{1}$$

where the Schroeder frequency, f_s (Hz), is characterized by the reverb time, RT_{60} (s), and the room volume, V (m^3). This limit for the low-frequency range of a closed space operates on the principle that above the Schroeder frequency room-modes become sufficiently spatially and spectrally dense to not be subjectively distinct, largely due to masking within the human hearing mechanism [4].

Room-modes are a consequence of standing waves between one or more set of parallel reflecting surfaces and arise at frequencies with integer multiples of their half-wavelengths fitting perfectly within the standing wave pattern (Eq. 2) [5]. Listeners experience largely different steady-state and transient responses at these frequencies, depending on their location within the complex standing wave pattern (Fig. 1 & 2).

$$f_m = \frac{c}{2}\sqrt{\left(\frac{\eta_x}{x}\right)^2 + \left(\frac{\eta_y}{y}\right)^2 + \left(\frac{\eta_x}{z}\right)^2} \tag{2}$$

where the room-mode frequencies, f_m, are calculated for η_x, η_y and η_z from zero upwards with x, y, and z representing the dimensions (in meters) of a rectangular space and c, the speed of sound in air (m/s).

The widely varying frequency response over a listening area, as shown in Fig. 1, demonstrates how greatly the low-frequency steady-state acoustical response differs between closely spaced listening locations. At some locations certain frequencies are overpowering while at other locations the same frequencies are virtually non-existent. The related transient responses highlight the spatiotemporal variation among listeners, often causing difficulty in perceiving detailed time-domain nuances (i.e., following the bass line) within a signal, as illustrated in Fig. 2 with an 80 Hz tone burst [6].

2 COMMON ROOM-MODE CORRECTION PROCEDURES

Low-frequency room-mode correction can be approached using a number of well-known passive and active procedures. Each of these techniques addresses the modal problem from a different perspective, resulting in varying advantages and disadvantages between the methodologies. The positive and negative aspects of each approach must be considered when selecting a correction strategy that meets system requirements. In this section commonplace methods will be highlighted and discussed in terms of how they could benefit from a supplemental psychoacoustically-based procedure.

2.1 Passive Correction–Absorption

Increasing a room's surface absorption is a simple technique to reduce modal problems within a space by decreasing wall reflections, thus limiting the buildup of standing waves. This causes low-Q room-modes, resulting in less pronounced resonances due to the increased modal overlap (Fig. 3).

Fig. 3 highlights a noticeable decrease in sharp resonances as absorption is increased tenfold, although spatial variance only decreases by a marginal amount of 5.0%. Even though the space exhibits fewer sharp resonances with added wall absorption, a strong variance still exists between listeners.

Bass-traps are commonly implemented to provide additional absorption at key room-modes. A variety of bass-traps are available in practice, including porous absorbers and resonating absorbers. These approaches each have their advantages and disadvantages, as discussed in [5, 7, 8].

Even if it were possible to increase a room's absorption level tenfold (regardless of the absorption technique utilized), such as in Fig. 3, spatial variance remains large. In this example, three narrow bands (centered at 40, 57 and 85 Hz, respectively) could be removed from the physical

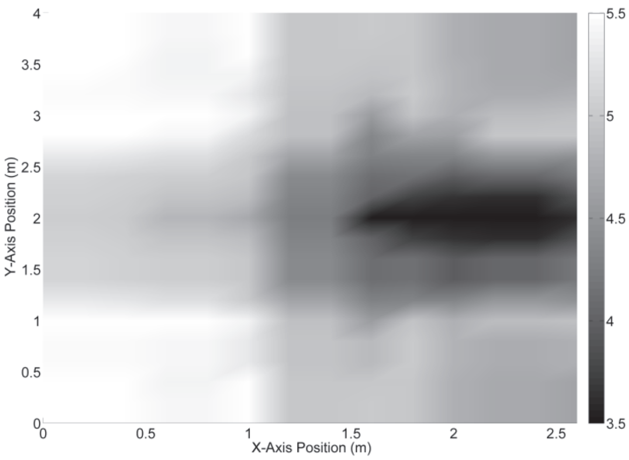

Fig. 4 Spatial variance (dB) at various single subwoofer placements over the first half of a 5 m x 4 m x 3 m room.

reproduction signal and replaced psychoacoustically with virtual bass to further reduce spatial variance.

2.2 Passive Correction–Source Placement

In situations where additional absorption is not practical and there are minimal available signal processing options within a system, intelligent subwoofer placement can provide a significant reduction in spatial variance.

Often the goal is to achieve maximum low-frequency output without excessive amplification requirements. This is achieved by keeping subwoofer to room-mode coupling in mind [9]. When a subwoofer is placed at an antinode of a room-mode, coupling will be maximized. When the subwoofer is placed at a node, coupling will be minimized (theoretically zero) due to placement at the standing wave's zero crossing [5]. Placing the subwoofer in a corner has the added benefit of the Waterhouse effect, whereby each adjacent boundary contributes an additional 6 dB to the sound pressure level, giving an 18-dB boost at a corner location [10].

While this simple expedient projects greater low-frequency energy into the room, it does not provide an equal response at all listening points. This is due to the strong dependence on source-to-listening location coupling. A single subwoofer should be placed close to as many nodes as possible to provide more uniform coverage at low-frequencies. Although the center of a room generally contains the most frequency nodes common to a single point, it is normally an impractical location for a subwoofer. In addition, center placement does not benefit from the Waterhouse effect, resulting in lower system output.

Due to the drawbacks of central subwoofer placement, a compromise can be made by placing the single subwoofer at a ground level wall midpoint (Fig. 4). This placement provides a (simulated) 16% spatial variance reduction compared to corner placement (while central placement results in a 38% reduction).

As with passive absorption, single subwoofer placement cannot provide sufficient spatial variance reduction to guarantee equal listening experiences for all listeners and the

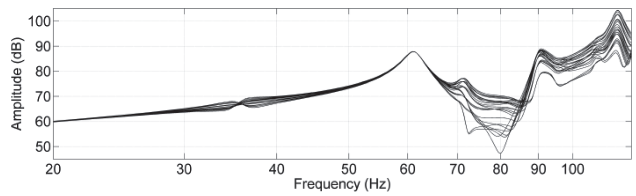

Fig. 5 Simulated frequency responses over a listening area with a four subwoofer system consisting of one omnidirectional unit at each wall midpoint.

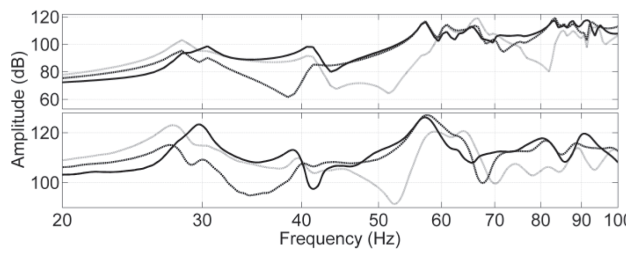

Fig. 6 Simulated (top) and measured (bottom) frequency responses at three listening locations due to one subwoofer at the front left corner of the room.

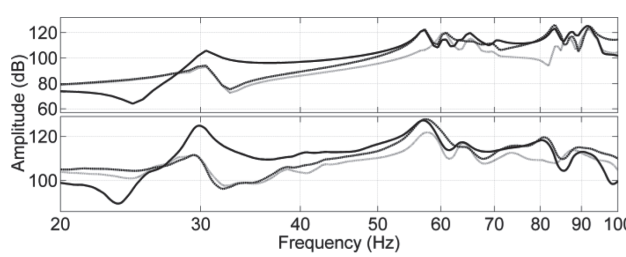

Fig. 7 Simulated (top) and measured (bottom) frequency responses at three listening locations due to one subwoofer in each front corner of the room.

problem therefore requires further consideration for scenarios demanding high-accuracy.

Research has been conducted using multiple omnidirectional subwoofers to provide further spatial variance reduction, concluding that four subwoofers located at wall midpoints is the most practical solution (Fig. 5) [5, 11].

This configuration provides a more uniform listening experience throughout a space, but with lower efficiency due to the destructive interference used to limit the buildup of standing waves and also the nodal placement of the sources (low source-to-room coupling).

To provide additional illustration of this point, measurements were taken in the University of Essex Audio Research Laboratory listening room (dimensions 6.05 m x 5.79 m x 2.80 m) with a square nine-point measurement grid centered at (3.50 m, 3.00 m, 1.50 m) and point spacing of one meter. Three listening points were analyzed for both a single subwoofer system with front-left corner placement (Fig. 6) and a dual-subwoofer system with one subwoofer in each front room corner (Fig. 7).

The simulated and measured responses illustrate precisely how destructive interference between multiple subwoofers can suppress lower-order axial room-modes, thus reducing spatial variance between listening points. There is a clear spatial variance reduction between 30 – 60 Hz, when comparing the single and dual-subwoofer systems shown in Fig. 6 and 7, respectively.

Subwoofer configuration has also been explored in the context of 5.1 surround sound [12] where it is stressed that subwoofer position(s) in relation to the other loudspeakers must be considered to avoid any subjectively unpleasant interference. Without any additional signal processing it is unlikely to find a practical position where the subwoofer integrates perfectly with the main and surround loudspeakers, therefore significant spatial variance is likely to occur in the crossover region. Subjectively, however, it has been argued that subwoofer placement in a 5.1 audiovisual configuration is noncritical, possibly due to distraction by the visual components [12].

The various approaches to spatial variance reduction via subwoofer placement could benefit from supplemental psychoacoustical reinforcement. In the case of the four-subwoofer system with wall midpoint placement (Fig. 5) the bulk of spatial variance has been suppressed, except at a narrow band centered around 80 Hz. Physical reinforcement of this band could be replaced psychoacoustically to further equalize all listening locations. The issue presented in [12] concerning significant spatial variance around the crossover region of a 5.1 surround sound system could be addressed in this manner, subjectively replacing the crossover region.

2.3 Active Correction–Parametric Equalization

A simple to implement room-mode suppression technique involves parametric equalization. This method usually targets three to five of the most problematic room-modes by applying notch filters centered at these frequencies to limit their reinforcement within the sound system. This strategy limits the buildup of standing waves at these frequencies and can help to reduce spatial variance.

While this low-frequency correction method does not address all modal problems within a space, it can be used as a quick fix for the worst acoustical problems in a room. Problematic modes can be identified and handled automatically with room measurements or addressed manually by ear.

The drawback to this technique is that it eliminates information in the targeted frequency bands; therefore listeners may miss key elements of the audio signal in exchange for modal suppression. Again, supporting psychoacoustical reinforcement could reintroduce the missing data into the audio signal, thus delivering all auditory information to the listeners.

2.4 Active Correction–Single-Point Equalization

A common correction technique similar to automatic parametric equalization is single-point equalization. This operates by taking a measurement at the primary listening point in a room, generally with a maximum length sequence (MLS) test signal. Once the frequency response is calculated from the measurement, an inverse filter can be generated. It is noted in [13] that an electroacoustic system is likely to result in an unstable filter due to the mixed-phase nature of the system. There are two primary options around this: First, a filter can be generated based on the magnitude response. This will ensure stability, but does not provide a complete equalization solution, as phase is ignored. Alternatively, an optimization approach can be implemented with error minimization in mind, resulting in a best-case equalization solution without risking instability [13]. These methods can be applied either as static filters [13] or as filters which vary in real-time, depending on the response of the system [14].

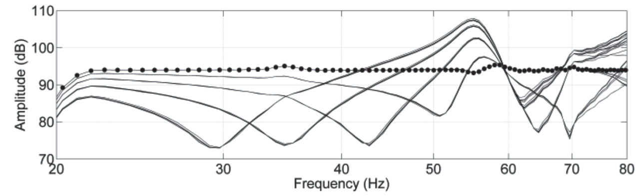

Fig. 8 Simulated frequency responses over a listening area after single point equalization (dotted line = target EQ point).

Single-point equalization can be applied over the entire audible frequency range (20 Hz – 20 kHz) and typically performs well at the target point. Unfortunately, this is not the case for the non-target points, where source-to-listener coupling is different from the target location. This causes low correlation between the correction benefits at the target point and non-target points. Spatial variance will usually not reduce with single-point equalization (Fig. 8); therefore this technique is only effective for scenarios where there is only one listening location/listener using the system at a time or the uncorrected listening area naturally exhibits very low spatial variance.

2.5 Active Correction–Other Techniques

There exist a number of additional low-frequency correction techniques that have been addressed in previously published literature. Many of these routines involve multiple-point equalization whereby measurements are taken throughout the listening area and are grouped based on similarity and/or weighted based on location importance to give significant spatial variance reduction [15-20].

Some of these methods employ fixed equalization (one-time measurements) while others utilize adaptive systems whereby measurements are continuously taken as the system operates, leading to a problem in having measurement microphones throughout the listening area at all times. This problem has been eliminated with the system presented in [19] since measurements are taken in close proximity to the subwoofer.

An additional room-mode correction method that has been the topic of investigations is active absorption [21-24]. Active absorption combines the principles of passive correction methods and single/multiple point equalization methods. This method generally operates with a set of one or two primary sources placed at one end of a room. An array of secondary sources is placed at the opposite end of the room, with each drive unit usually containing a measurement microphone to monitor the signals received from the primary sources. These secondary sources will reproduce their measurements with reverse polarity in an attempt to suppress wall reflections, giving a traveling wave in the room as if it were anechoic.

Active absorption can require a large number of secondary drive-units to effectively suppress high-order modes, making it difficult and costly to implement. When properly calibrated, though, these systems can create a virtual anechoic environment where all points (at a sufficient distance from the secondary units) will experience the same response both in the time and frequency domains.

Source/listener placement in either the multi-point equalization or active absorption systems is likely to cause troublesome correction at a select few frequencies due to nodal placement requiring unrealistic energy levels. Replacing these frequencies with virtual bass will limit the requirements on the physical correction system, increasing efficiency while decreasing spatial variance.

3 VIRTUAL BASS SYNTHESIS

This section describes a low-complexity room-mode correction process that builds upon the concept of parametric equalization by incorporating a psychoacoustically-motivated procedure known as *virtual bass*. The process is compatible with a wide range of sound reinforcement systems, as highlighted in the previous section. The core problem with conventional parametric equalization is that in order to reduce spatial variance it suppresses the frequencies that maximally excite room-modes, thus filtering potentially important audible information. The enhancement presented in this paper seeks to compensate for this loss of information by substituting a subjectively equivalent signal based upon virtual bass synthesis.

This is not an entirely novel concept. A system has been patented [25] whereby virtual bass is utilized to avoid excessive "noise pollution" from a home theater system into neighboring residences. This methodology targets a frequency band spanning 50 – 120 Hz. The input signal is passed through a band-stop filter corresponding to the target frequency band and then, in parallel, the input signal has virtual bass applied to boost the impression of low-frequency energy. The two signals are finally summed to form the output signal. Unlike the system in [25], the concept proposed in this work targets multiple narrow frequency bands in order to diminish spatial variance due to room-modes, as opposed to addressing "noise pollution" and/or band limited loudspeakers.

Virtual bass synthesis operates on the doctrine of the missing fundamental. The missing fundamental, or the residue pitch, is a result of the complex pitch-extraction mechanism within the inner ear and the brain. When presented with a spectrally-complex sound, the pitch extraction mechanism attempts to make sense of the received signal by relating various spectral components to one another [26]. Eqully spaced spectral components result in a perceived pitch corresponding to the greatest common factor of the frequency values (in Hz) that falls within the audible range of 20 Hz – 20 kHz. For instance, if the source contains spectral components at 200, 300, 400, and 500 Hz the overall perception will correspond to a harmonically-rich tone at 100 Hz.

This effect can operate using only two higher harmonic components (e.g., second and third) of the fundamental. Adding additional harmonics will increase the sharpness of the signal timbre (sound quality) as the average frequency of the components increases [26].

When applying the missing fundamental for low-frequency applications, it is important to keep the average frequency of all spectral components to a minimum so that the perceived pitch is as close in timbre to the fundamental as possible. Minimizing the amount of harmonic components introduced also preserves the fidelity of the source signal since these virtual bass components are a form of distortion, which should ideally be kept to a minimum.

There are two primary implementations of the virtual bass effect where both offer unique advantages and disadvantages. These two techniques are presented in the following sections.

3.1 Nonlinear Device Virtual Bass

A nonlinear device (NLD) is the most common harmonic generator implemented within virtual bass systems (such as in [25]) for a number of reasons. First, the NLD is memoryless, allowing for real-time applications. NLDs generally operate using a polynomial approximation of a chosen function. The calculated coefficients are then applied to the input signal as defined in Eq. 3.

$$y = \sum_{i=0}^{N} h_i x^i \tag{3}$$

where, h is a vector containing the N polynomial coefficients with x and y representing the signal input and output, respectively [27, 28].

The NLD virtual bass technique operates in the time domain, applying the effect over all spectral components of the signal. However, this process normally introduces intermodulation distortion to the signal if there are two closely-spaced spectral components in the input signal. While it has been argued that these components cause minimal auditory artifacts due to psychoacoustical masking at the Basilar membrane in the inner ear [27], intermodulation distortion is an unwanted peripheral to the NLD virtual bass system, which must be handled with care.

Early virtual bass research utilized a full-wave rectifier (FWR) for the NLD [29]. The FWR is simple to implement, but suffers from the fact that it generates only even-order harmonics. A FWR applied to a 100 Hz pure tone would result in harmonic distortion introduced at 200, 400, 600 Hz, and so on. Following the principle of the missing fundamental, this harmonic series should result in a perceived pitch of 200 Hz rather than 100 Hz. The perceived pitch is a full octave higher than the target pitch perception that results in an inaccurate virtual bass effect.

This problem has led to a significant body of research to develop the ideal NLD for virtual bass applications. A

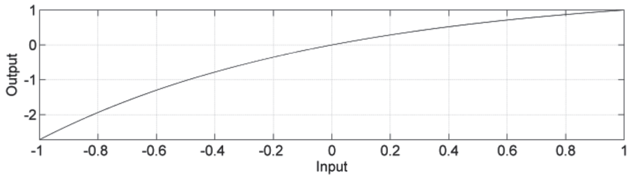

Fig. 9 Input-output relationship for exponential NLD virtual bass.

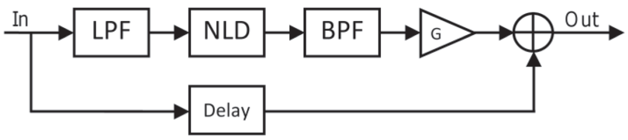

Fig. 10 NLD virtual bass procedure.

wide range of NLDs are presented in [27], where they are each objectively and subjectively evaluated to best judge performance. The second exponential-type NLD in [27] was rated highly in both objective and subjective tests and was therefore chosen as the NLD for this work. The input-output relationship is shown in Fig. 9.

NLD virtual bass systems are implemented with a series of filters to give approximate control of the spectral components of the effect. The input signal is first processed by a low-pass filter (LPF) with a cutoff frequency set to the upper limit of the required low-frequency extension. This low-pass filtered signal is then processed by the NLD, generating the harmonic components.

Next, the NLD output is sent through a bandpass filter (BPF) to remove the fundamental spectral components and to roughly shape the harmonic components. If only a low-frequency boost is required (as opposed to a bandwidth extension), the BPF can be replaced by an LPF. After the BPF, gain is applied to the signal and then combined with a delayed version of the original signal. The overall NLD virtual bass process is shown in Fig. 10.

A widely-utilized commercial NLD-based virtual bass system is called *MaxxBass* [30]. In addition to the system architecture in Fig. 10, *MaxxBass* uses equal-loudness processing to provide a virtual bass effect subjectively equal in level to the unprocessed signal.

3.2 Phase vocoder virtual bass

An alternative to the NLD virtual bass approach has emerged in recent years utilizing a phase vocoder (PV) as the harmonic generator [31]. The PV virtual bass approach provides superior harmonic control, allowing for selective harmonic inclusion. Since this approach operates in the frequency domain, the intermodulation distortion can be effectively avoided, unlike with NLDs.

PVs operate by splitting an input signal into short time-domain windows (generally between 50 – 250 ms). The PV takes the fast Fourier transform (FFT) of each time window, applies the required processing while maintaining phase coherence and then generates the output signal either by sum-of-sinusoids or inverse Fourier transforms where each window is overlap-added to minimize amplitude-modulation effects. This work utilizes the sum-of-sinusoids method.

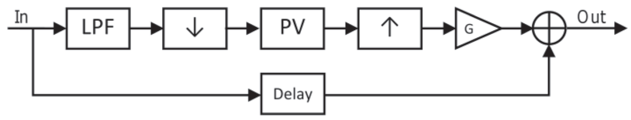

Fig. 11 Phase vocoder virtual bass procedure.

A disadvantage to the PV arises due to the trade-off between time and frequency resolution. Virtual bass systems require adequate frequency resolution to allow for accurate harmonic generation in addition to avoiding intermodulation distortion. Frequency resolution can be determined by Eq. 4.

$$f_{res} = {}^{1}\!/_{t_w} \tag{4}$$

where, f_{res} is the frequency resolution (Hz) and t_{w} is the window length (s). For example, a 125 ms window gives 8 Hz resolution while a 500 ms window gives 2 Hz. This issue leads to smeared transient performance that is clearly evident when applied to audio signals such as drum beats.

Previous solutions to this problem have involved reinitializing the phase within the algorithm when a transient is encountered [32] and also removing any transients from the input signal and then reinserting them, unprocessed, at the PV output [33]. The phase re-initialization solution can prove difficult as it relies on precise transient detection; otherwise, phase re-initialization will occur in excess, causing poor phase coherency for the steady-state signal components. The transient removal method has had low ranking in subjective tests since transient signal components are not addressed within the effect [33].

Even though the PV cannot handle transients perfectly it does perform well on pitched signal components. Unlike the NLD system, PV virtual bass does not require a LPF on the input stage, as the algorithm can selectively apply the effect to frequency bins. Within the PV the selected frequencies are pitch shifted to the desired harmonic frequencies and amplitude adjusted to match any equal-loudness requirements; therefore no BPF or HPF is necessary on the output stage.

Since the PV virtual bass system is more computationally demanding, it is necessary to down-sample the input signal for real-time applications. This requires a LPF before the down-sampling process to avoid any spectral aliasing. Once the signal has been processed, it can be up-sampled to the original sampling rate and recombined with the delayed original signal. The overall PV virtual bass process is shown in Fig. 11.

While PVs are commonly used for audio effects such as pitch shifting and time stretching [34], there are no known commercial applications of PV virtual bass.

3.3 Hybrid Virtual Bass

A virtual bass system that exploits the respective strengths of the NLD and PV systems but circumvents their weaknesses should provide bass synthesis less sensitive to changes in input signal content. When the input signal has

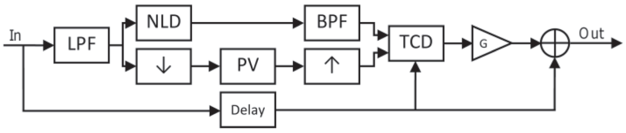

Fig. 12 Hybrid virtual bass procedure.

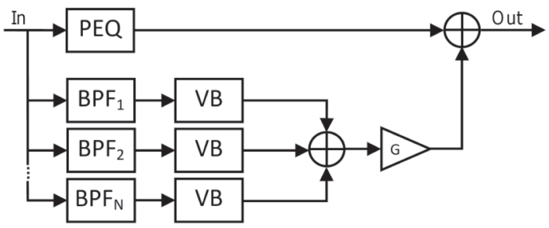

Fig. 13 Virtual bass room correction procedure.

high transient content, the system favors the NLD output and conversely, when the signal is more pitched the PV effect is utilized.

This hybrid approach utilizes a transient content detector (TCD), which analyzes successive time domain windows of the input signal and appropriately weights the respective virtual bass algorithms that are running in parallel (Fig. 12).

The implementation of the hybrid virtual bass procedure is discussed in detail in [35] including results from subjective evaluations that rated the new procedure alongside NLD and PV approaches over a wide range of musical genres. The hybrid approach shows less sensitivity to input content and was therefore chosen as the virtual bass procedure for this work.

4 VIRTUAL BASS ROOM-MODE CORRECTION

Virtual bass can be used as a supplemental component within the parametric equalization structure to help suppress the most problematic room-modes but without losing crucial audio information, as alluded to in section 2.3. When used as a standalone application, virtual bass can often produce an artificial sounding effect which can detract from a natural listening experience. These applications are often targeted toward bandwidth extension of restricted loudspeakers where there are few alternatives to achieve strong low-frequency perception.

However, if the virtual bass effect was limited to a narrow-band application, many of the artifacts may be masked by the surrounding frequencies of physically reproduced energy. The narrow band(s) removed from the signal through parametric equalization could then be reinforced psychoacoustically with the narrow-band virtual bass procedure to maintain any information present within these frequency bins. This room-mode correction approach is illustrated in Fig. 13.

In Fig. 13 the unprocessed signal is sent via the parametric equalization (PEQ) routine with notch filters centered at the most problematic room-modes. At the same time, the unprocessed signal travels in parallel through N bandpass filters (BPF_x) centered at each target frequency and is then run through the hybrid virtual bass procedure (VB), as detailed in section 3.3 of this paper and in [35]. All virtual bass outputs are summed with appropriate gain (G) applied to the resulting signal. The final virtual bass signal is recombined with the parametric equalization output to give the fully processed signal to be sent through the remainder of the signal chain.

Table 1. Musical selections by genre

Genre	Artist	Song
Classical	Frank Zappa	Dog Breath Variations
Jazz	The Bad Plus	Big Eater
Blues	Bernard Allison	Mean Town Blues
Rock	Jeff Beck	There's No Other Me
Pop	Robert Randolph	Diane
Vocals	Blind Boys of Alabama	These Bones Gwine Rise Again
Reggae	Bob Marley	Get Up Stand Up
Country	Drive-By Truckers	Bob
Folk	Alison Breitman	Tenafly
Hard Rock	Audioslave	The Worm

4.1 Subjective Evaluation Procedure

Since the virtual bass effect occurs within the human hearing mechanism and the brain, it is necessary to subjectively evaluate the proposed virtual bass correction procedure. Since the parametric equalization routine removes various narrow bands from the physically reproduced signal, the resulting signal is expected to have slightly reduced low-frequency impact, but with the aspiration of maintaining high fidelity with minimal obvious artifacts due to the harmonic distortion from the virtual bass effect.

Tests were carried out in the University of Essex Audio Research Laboratory listening room. The sound reproduction system consisted of two sealed-box subwoofers placed on the ground at wall midpoints to the left and right of the listening positions along with left and right main stereo loudspeakers. Two adjacent listening locations were chosen where the right location naturally received strong low-frequency energy due to the close proximity of many mode antinodes while the left location received little energy due to nodal placement.

The four strongest room-modes (41, 58, 67, and 84 Hz) were chosen as targets based on FDTD simulations using proprietary software [2, 3] and confirmed with previous room measurements.

Ten high-fidelity musical recordings were chosen for the tests, each from a distinct musical genre as detailed in Table 1.

Subjects were first presented with the unprocessed musical sample and instructed to move between the two seats to judge both the overall sound quality and the low-frequency level variance between the two locations. Sound quality was rated on a one-hundred point scale with one-hundred being the best possible score. Low-frequency variance was also rated on a one-hundred point scale with one-hundred representing significant spatial variance and zero representing no noticeable spatial variance.

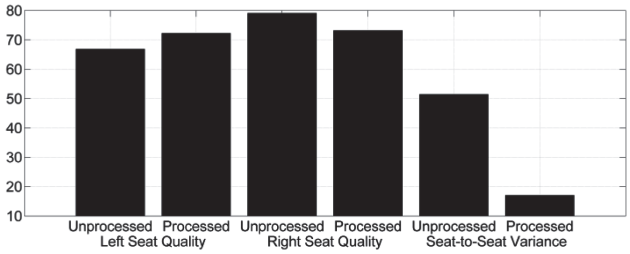

Fig. 14 Virtual bass low-frequency room-mode correction subjective evaluation results.

The test subjects were given a list detailing the musical selections and asked to choose three based on their musical preferences. Next, a sample track was played, allowing subjects to become accustomed to the test procedure. Each unprocessed/processed clip pair was played until the subject had assigned ratings. The entire test generally required fifteen to twenty minutes to complete. The signal processing procedure was not revealed to the listeners to avoid any possible biasing.

4.2 Subjective Evaluation Results

The test subjects were composed of fifteen males and six females ranging in age from twenty-three to sixty-three years old. Each subject completed the test during independent sessions. The subjective evaluation results are presented in Fig. 14.

The subjective evaluation results indicate that the right seat generally received quality ratings in the "good" range which can be largely attributed to strong low-frequency presence. The left seat, on the other hand, received quality ratings in the "fair" range with subjects commenting that they sensed the left seat lacked certain musical information. The differences in quality ratings are reflected in the unprocessed seat-to-seat low-frequency spatial variance ratings falling in the "moderate" range.

After virtual bass processing, though, the subjective ratings show a noticeable shift. The right seat, while rated "good" unprocessed, has decreased to the lower bound of the "good" range. The left seat received "fair" quality ratings unprocessed but has increased into the "good" range. The left and right seats' processed ratings are within 2% of each other, which is strongly exhibited in the processed seat-to-seat low-frequency spatial variance ratings falling in the upper portion of the "not noticeable" range.

The subjective evaluations have shown that virtual bass room correction can provide a reasonable amount of spatial variance reduction between seats. The compromise is that seats with naturally superb responses tend to experience slight decreases in fidelity in order to increase the fidelity of naturally poor sounding seats. The virtual bass ensures that all musical information present in the unprocessed signals is perceptually maintained in the processed signals.

5 DISCUSSION

The impetus behind the development of this virtual bass room-mode correction procedure is to ease system requirements of existing methodologies where control at select narrow frequency bands may prove elusive. Although subjective evaluations of the standalone virtual bass correction strongly indicate that spatial variance has nearly been eliminated between listening points, test subjects expressed that the corrected system lacks a certain physical impact, which is due to the absence of physical reinforcement at the targeted frequency bands.

In addition to reduced physical impact, overuse of virtual bass correction can result in perceptually unrealistic sound reproduction due to the harmonic distortion introduced. These two issues support the proposal that this form of correction be utilized as a supplement, rather than a replacement, to a well-established room-mode correction system.

As highlighted in section 2, conventional correction often struggles at a few spectrally narrow regions. If the virtual bass procedure was implemented to only handle correction over these frequencies, there would be minimal loss in physical impact while introducing nominal harmonic distortion from the virtual bass algorithm. The limited use of virtual bass is crucial as excessive harmonic distortion leads to a perceived "harshness" or "roughness" in musical signals [36], which is undesirable in high-fidelity sound reproduction.

6 CONCLUSIONS

A wide-area low-frequency room-mode correction method based on a combination of parametric equalization and the virtual bass effect has been presented as a solution for reducing the negative effects of room-modes (primarily high spatial variance) while maintaining all input signal information and reasonable fidelity.

Correction method complexity is reduced by not relying on a purely physical correction system, but using a hybrid physical/psychoacoustical system which can be applied to nearly any sound reproduction system. This technique was primarily developed out of the need to address the few narrow frequency bands that prove difficult to correct for using a chameleon subwoofer array (CSA) [1].

The virtual bass room-mode correction algorithm was subjectively tested by comparing overall sound quality between unprocessed and processed signals as well as seat-to-seat variance between the two versions. Results clearly show a sharp drop in spatial variance due to the virtual bass procedure with a decreased contrast between sound quality ratings at each seat.

This correction procedure was developed as an add-on to a specific low-frequency correction method, but it can be implemented as a supplemental tool within any correction system to provide the best possible correction over the entire low-frequency spectrum.

REFERENCES

[1] A. J. Hill, M. O. J. Hawksford, "Chameleon Subwoofer Arrays—Generalized Theory of Vectored Sources in a Closed Acoustic Space," presented at the 128th

Convention of the Audio Engineering Society, convention paper 8074, 2010 May.

[2] A. J Hill, M. O. J. Hawksford, "Visualization and Analysis Tools for Low Frequency Propagation in a Generalized 3D Acoustic Space." *J. Audio Eng. Soc.*, vol. 59, no. 5, pp. 312–337 (2011 May).

[3] FDTD Toolbox v8.0. www.adamjhill.com/fdtd.

[4] M. R Schroeder, K. H. Kuttruff, "On Frequency Response Curves in Rooms. Comparison of Experimental, Theoretical, and Monte Carlo Results for the Average Frequency Spacing between Maxima," *J. Acoust. Soc. Am.*, vol. 34, no. 1, pp. 76–80 (1962 January).

[5] F. E. Toole, *Sound Reproduction: Loudspeakers and Rooms* (Focal Press, New York, 2008).

[6] S. Linkwitz, "Investigation of Sound Quality Differences between Monopolar and Dipolar Woofers in Small Rooms," presented at the 105th Convention of the Audio Engineering Society, convention paper 4786, 1998 September.

[7] J. A. Fletcher,"The Design of a Modular Sound Absorber for Very Low Frequencies," BBC Research Department, report number RD1992/10. 1992.

[8] E. Reiley, A. Grimani, "Room Mode Bass Absorption through Combined Diaphragmatic & Helmholtz Resonance Techniques: 'The Springzorber,'" presented at the 114th Convention of the Audio Engineering Society, convention paper 5760, 2003 March.

[9] T. Salava, "Subwoofers in Small Listening Rooms," presented at theÂ 106th Convention of the Audio Engineering Society, convention paper 4940,Â 1999 May.

[10] R. V. Waterhouse, "Output of a Sound Source in a Reverberation Chamber and Other Reflecting Environments," *J. Acoust. Soc. Am.*, vol. 30 no. 1, pp. 4–13 (1958 January).

[11] T. Welti, A. Devantier, "Low-Frequency Optimization Using Multiple Subwoofers," *J. Audio Eng. Soc.*, vol. 54, no. 5, pp. 347–364 (2006 May).

[12] N. Zacharov, S. Bech, D. Meares, "The Use of Subwoofers in the Context of Surround Sound Program Reproduction," *J. Audio Eng. Soc.*, vol. 46, no. 4, pp. 276–287 (1998 April).

[13] J. Mourjopoulos, "Digital Equalization Methods for Audio Systems," presented at the 84th Convention of the Audio Engineering Society, preprint 2598, 1988 March.

[14] J. N. Mourjopoulos, E. D. Kyriaskis-Bitzaros, C. E. Goutis, "Theory and Real-Time Implementation of Time-Varying Digital Audio Filters," *J. Audio Eng. Soc.*, vol. 38, no. 7/8, pp. 523–536 (1990 July/August).

[15] S. Bharitkar, C. Kyriakakis, "Objective Function and Automatic Multi-Position Equalization and Bass Management Filter Selection," presented at theÂ 119th Convention of the Audio Engineering Society, convention paper 6608, 2005 October.

[16] A. O. Santillan, "Experimental Low-Frequency Sound Equalization in an Extended Region of an Enclosure Using Adaptive Filters," presented at the 104th Convention of the Audio Engineering Society, convention paper 4755, 1998 May.

[17] A. O. Santillan, "Spatially Extended Sound Equalization in Rectangular Rooms," *J. Acoust. Soc. Am.*, vol. 110, no. 4, pp. 1989–1997 (2001 October).

[18] S. J. Elliot, L. P. Bhatia, D. W. Deghan, A. H. Fu, M. S. Stewart, D. W. Wilson, "Practical Implementation of Low-Frequency Equalization Using Adaptive Digital Filters," *J. Audio Eng. Soc.*, vol. 42, no. 12, pp. 988–998 (1994 December).

[19] J. A. Pedersen, "Adjusting a Loudspeaker to its Acoustic Environment—The ABC System," presented at theÂ 115th Convention of the Audio Engineering Society, convention paper 5880, 2003 October.

[20] M. Shin, F. M. Fazi, J. Seo, P. A. Nelson, "Efficient 3D Sound Field Reproduction," presented at the 130th Convention of the Audio Engineering Society, convention paper 8404, 2011 May.

[21] P. Darlington, M. R. Avis, "Time/Frequency Response of a Room with Active Acoustic Absorption," presented at the 100th Convention of the Audio Engineering Society, preprint 4192, 1996 May.

[22] A. Celestinos, S. Birkedal Nielsen, "Low Frequency Sound Field Enhancement System for Rectangular Rooms Using Multiple Low Frequency Loudspeakers," presented at the 120th Convention of the Audio Engineering Society, convention paper 6688, 2006 May.

[23] J. Vanderkooy, "Multi-Source Room Equalization: Reducing Room Resonances," presented at the 123rd Convention of the Audio Engineering Society, convention paper 7262, 2007 October.

[24] A. Makivirta, P. Antsalo, M. Karjalainen, V. Valimaki, "Modal Equalization of Loudspeaker-Room Responses at Low Frequencies," *J. Audio Eng. Soc.*, vol. 51, no. 5, pp. 324–434 (2003 May).

[25] R. M. Aarts, "Band Stop Filter," International Patent Number WO 00/57673, 2000 September 28.

[26] J. F. Schouten, R. J. Ritsma, B. Lopes Cardozo, "Pitch of the Residue," *J. Acoust. Soc. Am.*, vol. 34, no. 8, pp. 1418–1424 (1962 September).

[27] N. Oo, W. S. Gan, "Analytical and Perceptual Evaluation of Nonlinear Devices for Virtual Bass System," presented at the 128th Convention of the Audio Engineering Society, convention paper 8108, 2010 May.

[28] N. Oo, W. S. Gan, and M. O. J. Hawksford, "Perceptually-Motivated Objective Grading of Nonlinear Processing in Virtual-Bass Systems," *J. Audio Eng. Soc.*, vol. 59, no. 11, pp. 804–824 (2011 November).

[29] E Larsen, R. M. Aarts, "Reproducing Low-Pitched Signals through Small Loudspeakers," *J. Audio Eng. Soc.*, vol. 50, no. 3, pp. 147–164 (2002 March).

[30] B. T. Daniel, C. Martin, "The Effect of the MaxxBass Psychoacoustic Bass Enhancement System on Loudspeaker Design," presented at the 106th Convention of the Audio Engineering Society, convention paper 4892,1999 May.

[31] M. R. Bai, W. C. Lin, "Synthesis and Implementation of Virtual Bass System with a Phase-Vocoder Approach," *J. Audio Eng. Soc.*, vol. 54, no. 11, pp. 1077–1091 (2006 November).

[32] A. Robel, “A New Approach to Transient Processing in the Phase Vocoder,” presented at theÂ *Proc. 6th Int. Conference on Digital Audio Effects. London, UK*, 2003 September.

[33] F. Nagel, S. Disch, N. Rettelback, “A Phase Vocoder Driven Bandwidth Extension Method with Novel Transient Handling for Audio Codecs,” presented at the *126th Convention of the Audio Engineering Society, convention paper 7711,*2009 May.

[34] J. Laroche, M. Dolson, “New Phase-Vocoder Techniques for Real-Time Pitch Shifting, Chorusing, Harmonizing, and Other Exotic Audio Modifications,” *J. Audio Eng. Soc.*, vol. 47, no. 11, pp. 928–936(1999 November).

[35] A. J. Hill, M. O. J. Hawksford, “A Hybrid Virtual Bass System for Optimized Steady-State and Transient Performance,” *Proc. 2nd Computer Science and Electronic Engineering Conference (CEEC), Colchester, UK. IEEE Xplore*, 2010 September.

[36] C. T. Tan, B. C. J. Moore, N. Zacharov, “The Effect of Nonlinear Distortion on the Perceived Quality of Music and Speech Signals,” *J. Audio Eng. Soc.*, vol. 51, no. 11, pp. 1012–1031 (2003 November).

THE AUTHORS

Adam J. Hill

Malcolm O. J. Hawksford

Adam J. Hill received a B.S. degree in electrical engineering from Miami University in 2007 and an M.Sc. degree with distinction in acoustics and music technology from the University of Edinburgh, Scotland, in 2008. He is currently in the completion stages of a Ph.D. degree within the Audio Research Laboratory at the School of Computer Science and Electronic Engineering, University of Essex, Colchester, UK. His research focuses on the use of networked subwoofer arrays to control low-frequency room responses. Past research has included voice-coil temperature effects in loudspeakers and real-time equestrian horse tracking via GPS/IMU integration. Outside of school he works professionally as a live sound engineer for Gand Concert Sound.

Mr. Hill is a student member of the Audio Engineering Society, IEEE, IOA, and IET.

●

Malcolm O. J. Hawksford received a B.Sc. degree with first-class honors in 1968, a Ph.D. degree in 1972, and a D.Sc. degree in 2008, all from Aston University, Birmingham, UK. He is currently an Emeritus Professor within the School of Computer Science and Electronic Engineering at Essex University, Colchester, UK. Early research embraced delta and sigma–delta modulation (SDM) applied to color TV coding, which under the award of a BBC Research Scholarship led to a method of luminance and chrominance multiplexing exploiting digital time compression, a forerunner of MAC/DMAC. Current interests include audio engineering, electronic circuit design, and signal processing focused on loudspeakers, SDM, PWM linearization, spatial audio, and telepresence.

Dr. Hawksford is a recipient of the AES Publications Award for the best contribution by an author of any age to the *Journal*, volumes 45 and 46, and holds the AES Silver Medal for major contributions to engineering research in the advancement of audio reproduction. He is a chartered engineer and Fellow of the Audio Engineering Society, IET, and IOA. He has been chair of the AES Technical Committee on High Resolution Audio and was a founding member of Acoustic Renaissance for Audio (ARA).