Applied Acoustics 70 (2009) 556–562

Contents lists available at ScienceDirect

Applied Acoustics

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

A hybrid MLS technique for room impulse response estimation

Joel Preto Paulo [a,c,*], Carlos Rodrigues Martins [b], J.L. Bento Coelho [c]

[a] Instituto Superior de Engenharia de Lisboa, Lisboa 1950-062, Portugal
[b] Escola Náutica Infante D. Henrique, Paço D'Arcos, Oeiras 2780, Portugal
[c] CAPS, Instituto Superior Técnico, TULisbon, Lisboa 1049-001, Portugal

ARTICLE INFO

Article history:
Received 21 April 2007
Received in revised form 10 July 2008
Accepted 11 July 2008
Available online 18 September 2008

Keywords:
Acoustical measurement techniques
Room impulse response
Low SNR
Room acoustics

ABSTRACT

The measurement of room impulse response (RIR) when there are high background noise levels frequently means one must deal with very low signal-to-noise ratios (SNR). If such is the case, the measurement might yield unreliable results, even when synchronous averaging techniques are used. Furthermore, if there are non-linearities in the apparatus or system time variances, the final SNR can be severely degraded. The test signals used in RIR measurement are often disturbed by non-stationary ambient noise components. A novel approach based on the energy analysis of ambient noise – both in the time and in frequency – was considered. A modified maximum length sequence (MLS) measurement technique, referred to herein as the hybrid MLS technique, was developed for use in room acoustics. The technique consists of reducing the noise energy of the captured sequences before applying the averaging technique in order to improve the overall SNRs and frequency response accuracy. Experiments were conducted under real conditions with different types of underlying ambient noises. Results are shown and discussed. Advantages and disadvantages of the hybrid MLS technique over standard MLS technique are evaluated and discussed. Our findings show that the new technique leads to a significant increase in the overall SNR.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

The maximum length sequence (MLS) technique is often used for measuring room impulse response function (RIR) usually leading to large signal-to-noise ratios (SNR) [1,2]. However, for the cases of high background noise levels, the SNR can be quite low. The estimation of sound insulation indexes, energy decay time/reverberation time (RT60) and the magnitude frequency response, as well as other objective parameters related to the subjective perception of the human auditory system, may then be unreliable [3–7].

In room acoustics, the background noise level and/or distortion of the impulse response are divided primarily into two components. The first component deals with ambient noise, which depends on environmental conditions such as traffic noise, resulting from poor facade insulation and structure-borne vibration, and the HVAC systems in the building; and/or people talking nearby during an event. This noise component is, for all intents and purposes, beyond the control of the acoustician. The second component is related to (i) the electroacoustic apparatus (mostly due to loudspeaker distortion at low frequencies or to the power amplifier overload), (ii) the inappropriate length of the ML sequence, or (iii) the number of frames used in the time average procedure [8]. This second noise component can be controlled (a) during the acoustical tests and/or (b) afterwards, during the post-processing stage.

In most situations, the room being tested can be modeled by means of a linear time invariant system (LTI) corrupted by added noise [14].

The LTI system can be described as follows:

$$\mathrm{Res}(n) = SeqMLS(n) \circledast h(n) + N(n) = y(n) + N(n), \tag{1}$$

where $SeqMLS(n)$ is the system input MLS frame, $\circledast$ stands for the circular convolution, $y(n)$ refers to room response and $N(n)$ is the noise during measurement. Here, the linear convolution is overwritten by the circular convolution operation given the periodicity of the MLS frames.

The ML sequences are termed pseudo-random signals, given their random phase behavior. However, they are deterministic signals. Therefore, the synchronous temporal averaging technique can be applied in order to improve the SNR. This technique performs a summation, where the correlated signals (the ML sequences) are amplified and the uncorrelated disturbance noise components are attenuated. The increment of the SNR resulting by averaging the synchronized M sequences is given as $\Delta\mathrm{SNR} = 10\log 10M$. Therefore, for each doubling of the number of the sequences, an increase of 3 dB for the SNR can be expected. This result is valid for any type of noise assuming low correlation between different

* Corresponding author. Address: Instituto Superior de Engenharia de Lisboa, Lisboa 1950-062, Portugal. Tel.: +351 218317207.
E-mail address: jpaulo@deetc.isel.ipl.pt (J.P. Paulo).

0003-682X/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.apacoust.2008.07.007

frames. For music signals used as noise, where a certain correlation exists between the captured frames, the increment of the SNR rises asymptotically to the theoretical value when the number of averages is increased.

Problems of time variance occur frequently in acoustical measurements. This phenomenon can be attributed mainly to air mass movement and to temperature gradients, which are more relevant in open or semi-open air environments such as arenas and stadiums. In closed spaces this effect is related to (i) the long-term duration of the acoustical measurement essentially due to air or to fast temperature drifts of the loudspeaker coil and (ii) HVAC systems or movement of people or objects (for instance, in large auditoriums or in public areas like transportation stations). Assuming the validity of the LTI model, whenever subtle time variance problems occur, a balance between the number of averages required and the increase of the MLS frame length should be considered [9,10]. In this case, the length of the MLS sequences should be increased (doubling the MLS sequence length leads to an increase of the SNR by 3 dB) when the inter-sequence time variance (between different sequences) is more visible than the intra-sequence time variance (during the same sequence). However, for long ML sequence length, the intra-sequence time variance could prevail against inter-sequence time variance. Therefore, a comprehensive optimization of the acoustical measurement set-up is then recommended and a pre-emphasis procedure should be applied.

There are situations where the energy decay range is not large enough for an adequate estimation of the acoustical room parameters. These situations occur when (i) the SNR is very low, i.e. the maximum acoustical power provided by the measurement equipment is not sufficient to excite the enclosure under test for all frequency bands, and, (ii) the presence of people inside the room is a pre-condition such as in public areas or during performance events. In these cases, the acoustical levels need to be kept below a pre-defined threshold, usually in relation to loudness contours, so that the people will not be disturbed.

In such cases, measurement times may increase substantially if averaging techniques are used. For example, let us assume one is interested in increasing the SNR by 20 dB. Here, 100 sequences need to be sent to the room, which comes to an expend time of about 3.5 min for a frame duration of 2 s. However, if this SNR increment is not sufficient and a supplementary increase of 10 dB is required, the measurement time will expand to 35 min (1000 MLS sequences). Moreover, if time variance phenomena occur during the measurements, the SNR will be upper-bound and, subsequently, the increase expected for the SNR will not be achieved. This upper limit for the SNR gain is a result of the increased noise from adding up several ML sequences that do not exactly match (the sequences are stretched or compressed randomly). In practice, this upper limit is observed by the acoustician when doubling the number of sequences, the SNR does not increase 3 dB.

The noise inside the enclosure being tested can be characterized as (i) a quasi-stationary process with quite lengthy intervals, for instance transportation noise or industrial noise, or (ii) non-stationary noise with an intermittent behavior that changes both in time and in frequency with pronounced energy fluctuations. The research presented here deals with situations of high levels and non-stationary noise and examines the acoustic energy both in time and frequency. By taking these issues into account, a noise-robust, MLS-based technique was investigated.

2. Method

The room being assessed is excited by a set of M ML sequences. An entire set of frames is sent to the room one after another or with delays between them. The purpose of the latter procedure is to keep the ambient noise statistics as different as possible during the presence of each MLS frame. However, each time the test signal is applied to the room, at least two concatenated sequences are sent. This procedure assures that the second frame holds all of the room's acoustical information. The first sequence of each set should be discarded. Obviously, this procedure is of less importance for MLS frame lengths several times greater than the expected RT.

In the case of RIR measurement where there are high levels and non-stationary background noise, the mean square, MS, value of the resulting captured signal (a measure of the acoustic energy of the sequence filtered by the room plus noise) must be kept to a minimum in order to exclude further noise energy components [11–13].

The analysis is described separately with regard to time and frequency to provide a better understanding of this new technique.

2.1. Segmented MLS technique

This method consists of splitting each captured MLS frame into N segments followed by the estimation of the MS value for each windowed segment. From each transmitted sequence, the segments with the lowest energy is extracted and a new sequence is built by using the overlap-add method. The corresponding arrangement is depicted in Fig. 1.

This procedure acts as a noise scrambler, decorrelating the noise added to each new frame due to the contribution of segments originated in different frames, at different times. The resulting sequence corresponds to the lowest energy Seq_{MLS}, thus ensuring the highest SNR value.

The ML sequence is segmented by applying the time windowing method,

$$S_{\text{MLS}}(n) = \sum_{k=1}^{N} \text{Seg}_{\text{MLS}_K}(n) = \sum_{k=1}^{N} S_{\text{MLS}}(n)\ w(n - kL/N + 1), \tag{2}$$

with

$$w(n) = \begin{vmatrix} 1 & 0 \leqslant n \leqslant L/N \\ 0 & n > L/N \end{vmatrix},$$

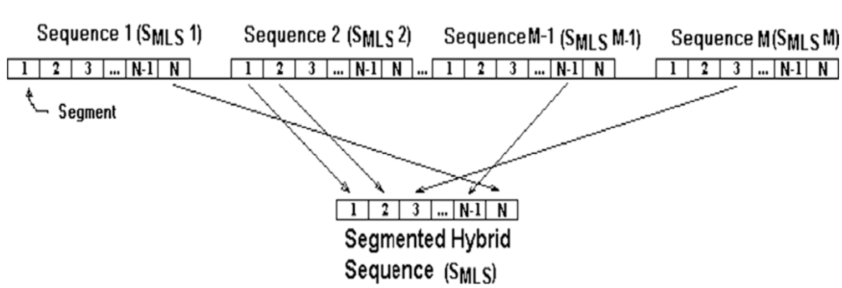

Fig. 1. Implementation of the segmented MLS technique.

where N is the number of segments in each sequence, L is the sequence length and $L_{seg} = (L+1)/N$ is the length of each segment, except for the last one that is equal to $[(L+1)/N]$-1, due to the odd number of samples of each ML sequence. Although the rectangular window is expressed in Eq. (2) for simplicity's sake, other types of window shapes are usually employed with 50% overlap. The response of the room to an MLS signal is given as

$$y(n) = \sum_{k=1}^{N} y_k(n) \tag{3}$$

where

$$y_k(n) = y(n)w(n - kL/N + 1) = [S_{\mathrm{MLS}}(n) * h(n)]w(n - kL/N + 1)$$

Considering the influence of the disturbing ambient noise, it follows that

$$\mathrm{Res}(n) = \sum_{k=1}^{N} \widehat{\mathrm{R}}\mathrm{es}_k(n)$$

where

$$\begin{aligned}\widehat{\mathrm{R}}\mathrm{es}_k(n) &= \mathrm{Res}(n)w(n - kL/N + 1)\\ &= [S_{\mathrm{MLS}}(n) * h(n) + N(n)]w(n - kL/N + 1)\end{aligned} \tag{4}$$

The SNR is calculated after the estimation of the MS value for each segment

$$\begin{aligned}\mathrm{SNR}_{\mathrm{Seg}} &= \frac{\mathrm{MS}(y_1(n)) + \mathrm{MS}(y_2(n)) + \cdots + \mathrm{MS}(y_N(n))}{\mathrm{MS}(N_1(n)) + \mathrm{MS}(N_2(n)) + \cdots + \mathrm{MS}(N_N(n))}\\ &= \frac{\mathrm{MS}((n))}{\sum_{k=1}^{N}\mathrm{MS}(N_k(n))}.\end{aligned} \tag{5}$$

The overall SNRSeg can be maximized by minimizing the denominator of Eq. (5). The MS value of the $y(n)$ is proportional to the acoustical power of the test signal used to excite the room. This value remains the same for the whole set of sequences. For this reason, this value was kept constant in order to compare and evaluate the performance of the segmented MLS technique against the standard MLS technique.

Hence, in this technique the segments which have the lowest MS value among the whole set of captured sequences are selected. Therefore,

$$N_k(n) = \min\{\mathrm{MS}(N_i(n))\}, \quad i \in [1, M]. \tag{6}$$

The improvement of this arrangement compared with the standard technique, ΔSNR, can be expressed by

$$\Delta\mathrm{SNR} = \frac{\mathrm{SNR}_{\mathrm{Seg}}}{\mathrm{SNR}_{\mathrm{Stand}}} = \frac{\mathrm{MS}(N(n))}{\sum_{k=1}^{N}\mathrm{MS}(N_k(n))} \geqslant 1 \tag{7}$$

where $N_k\,(n)$ is defined in Eq. (6).

2.2. Spectral MLS technique

At the receiver point, a filter bank is applied to each sequence by splitting the audio spectrum into multiple complementary sub-bands. The energy within each frequency sub-band is then computed. The sub-bands with the lowest spectral energy will be chosen to compose a new sequence. This resulting sequence will have the highest SNR, assuming the energy of the sequences sent is kept constant. An overview of this procedure is illustrated in Fig. 2.

The MS value of the system response $y(n)$ is constant for the whole set of sequences (the acoustical energy of the signal excitation is kept constant for all the ML sequences). Then, Eq. (3) yields

$$\begin{aligned}\mathrm{SNR}_{\mathrm{Spect}} &= \frac{\mathrm{MS}(y_1(n)) + \mathrm{MS}(y_2(n)) + \cdots + \mathrm{MS}(y_B(n))}{\mathrm{MS}(N_1(n)) + \mathrm{MS}(N_2(n)) + \cdots + \mathrm{MS}(N_B(n))}\\ &= \frac{\mathrm{MS}(y(n))}{\sum_{j=1}^{B}\mathrm{MS}(N_j(n))}\end{aligned} \tag{8}$$

The maximum value of SNRSpect in Eq. (8) is obtained by minimizing the MS value of the noise $N_k(n)$ within each sub-band,

$$N_j(n) = \min\{\mathrm{MS}(N_i(n))\}, \quad i \in [1, M]. \tag{9}$$

This new method leads to an increase in the SNR expressed as

$$\Delta\mathrm{SNR} = \frac{\mathrm{SNR}_{\mathrm{Spect}}}{\mathrm{SNR}_{\mathrm{Stand}}} = \frac{\mathrm{MS}(N(n))}{\sum_{j=1}^{B}\mathrm{MS}(N_j(n))} \geqslant 1 \tag{10}$$

where $N_j(n)$ is defined by Eq. (9), and SNRSpect and SNRStand are the SNR of the Spectral MLS technique and the standard MLS technique, respectively.

The hybrid MLS technique corresponds to the interconnection of the segmented MLS and spectral MLS methods. This technique is viewed as an acoustic energy time/frequency map split by cells corresponding to the windowed segments analyzed in frequency bands. The assembly of new MLS frames by choosing the cells belonging to deeper valleys (time/frequency segments with lower energy) yields to better results than the standard MLS technique

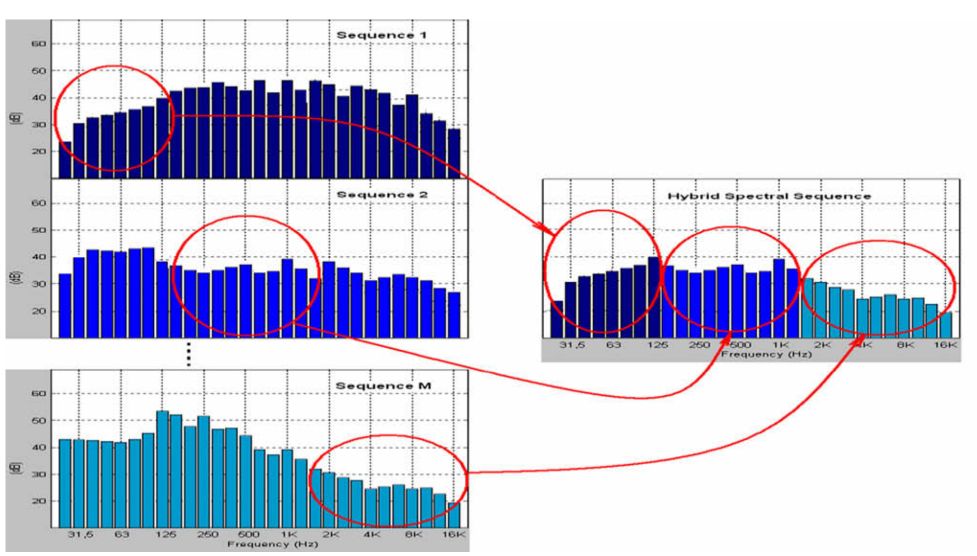

Fig. 2. Spectral MLS technique procedure for composing the sequence with the lowest sub-band energy.

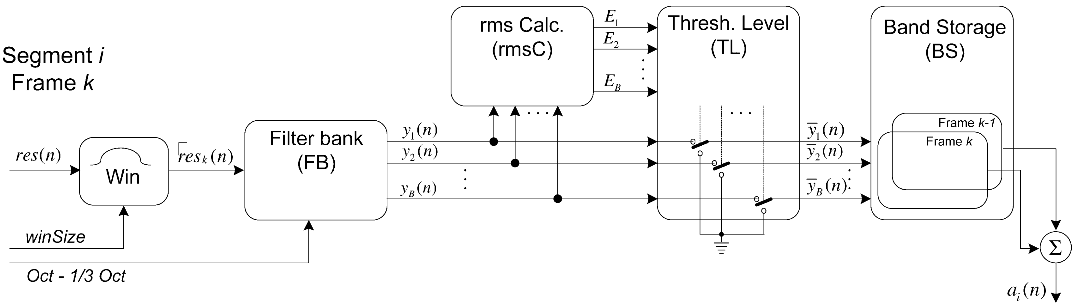

Fig. 3. Complete block diagram for the hybrid MLS technique for the analysis/processing of a segment and accumulation, $a_i(n)$ with the homologous segment from the previous frame.

for high levels and non-stationary noise profiles. This result is closely related to the noise profile and time variance occurrence, though for achieving a pre-defined SNR value or a magnitude frequency response error, fewer MLS frames are required for the averaging process.

In this technique, from each set of M sequences a new one is built, providing it has the lowest energy. However, this procedure may lead to worse results than those exhibited by the standard averaging technique (by summing up all the sequences) in situations where several frames or homologous segments (segments corresponding to the same position in different frames) share low energy noise. To improve the performance of the hybrid MLS technique, a threshold level is applied to the energy cells as a decision-making rule for rejecting segments. All the segments showing energy lower than this threshold value are used to build new hybrid MLS frames, thus, increasing the overall SNR with the application of the averaging technique. Replication of the segments in the case of incomplete frames is possible. Although the replicated segments do not contribute to increasing the SNR, there are others with different noise profiles. The increase in the SNR depends on the ratio between the different segments and the replicated segments. Thus, the final SNR strictly depends on the particular situation. The complete block diagram for the hybrid MLS technique is depicted in Fig. 3.

3. Experiments

The acoustical tests were performed in a small room with dimensions of approximately $6\,\text{m} \times 4\,\text{m} \times 3\,\text{m}$ $(L \times W \times H) = 72\,\text{m}^3$ with a mid-frequency RT60 of about 0.6 s. The measurement setup consisted of a dedicated software platform composed of modules to generate the MLS signal and implement the hybrid MLS technique; a PC with a stereo sound card; a sound source; and an omni-directional 1/2-in. (12.6 mm) type 1 condenser microphone, in accordance with the recommendations described in the ISO 3382 standard.

The assessment of the hybrid MLS technique was carried out by using eight non-consecutive frames, M. Two MLS frame duration, was tested for comparison purposes. All the frame durations had to be longer than the RT60 to avoid the time aliasing effect. Thus, a frame duration of about the RT60 and several times the RT60 was applied. The sequence length, L, of 32,767 and 131,071 samples (corresponding to a sequence order, $K = \log 2(L+1)$, of 15 and 17, respectively), with the sampling frequency, f_s, of 44,100 Hz met this requirement, since the sequence duration, L/f_s, is about 0.74 s and 2.97 s. All the tests were evaluated with 16 bits/samples.

Different types of disturbance signals and different SNR values were used. Optimal segment length was also investigated. The best results were produced by the sequence length of 1024 samples for speech and 2048 samples for music signals.

The hanning window type with 50% overlap was chosen since it is a suitable compromise between passband frequency response and side lobe rejection.

The tests were conducted on the full audio band, for the frequency spectrum split into four complementary frequency bands or into octave and 1/3 octave bands.

A FIR filter type of 550th order was used for the implementation of the complementary band-pass filters of the filter banks due to its linear phase. With the full audio band, care was taken in setting up the filter bank for the narrow bandwidth of the band-pass filters for the octave and 1/3 octave bands. Significant discrepancies among the filter coefficients and their very small values (close to zero) are noted, implying appreciable degradation of the filtered signals, mainly in low frequency bands. These obstacles were avoided by using multi-rate filtering techniques with two different sampling frequencies: 4410 samples/s for the first four octave bands (or for the first twelve on the 1/3 octave bands) and 44,100 samples/s for the remaining filter bands.

In order to guarantee reliable results, the input and the output signals were synchronized. Only the first sequence needs to be synchronized, since all the others refer back to the first. Thus, we took into account (i) the output/input latency of the audio cards channels in simultaneous playing/recording mode and (ii) the direct sound propagation delay inside the room. Since autocorrelation methods lead to poor results owing to the low SNR used in the experiments an alternative procedure was applied. The first problem was solved by connecting the audio card line out R (right output) with the MLS test signal, directly to the audio card Line In R (right input), with an appropriate gain level adjustment. This arrangement avoids the problem of having to know with precision the starting point of each sequence due to the operating system and sound board latency, which depend on the particular computer configuration. The sound propagation time was estimated by measuring the distance between the sound source and the microphone devices. Although error occurs in measuring the exact distance, especially in large rooms, this error can be overlooked in light of the MLS frame length and if concatenated sequences are applied, as explained in Section 2. For instance, an error of 1 meter corresponds to a slice between frames of about 130 samples (for the sampling rate of 44,100 Hz)at normal ambient temperature. At worst, this delay represents a minor degradation on the cross-correlation peak (the peak value is reduced by the same amount of the delay, in samples, with a time shift of the same amount).

In order to keep the magnitude frequency response as linear as possible, and to guarantee suitable acoustical power to avoid loudspeaker distortion, an adequate equalization of the whole measurement set-up was applied [15–19]. However, since this study is focused on comparing the performance of two different techniques, the test conditions were kept the same, thus reducing the importance of this issue.

In order to inspect the SNR dependence on the MLS signal parameters, several tests were carried out. Only speech and music tracks were used as noise signals, given their long-term non-stationary statistic properties. Also we were interested in studying the minimization of disturbance to audiences by using signal tests during live performances. Fig. 4 shows the SNR gain of the hybrid MLS technique against the standard MLS technique, or equivalently, the standard MLS to the hybrid MLS background noise ratios (the test signal acoustic power is the same for both techniques), for different types of added ambient noise.

We shall now discuss our conclusions regarding MLS frame length. For the sequence order of $K = 17$ (bottom figure) there is an increase in the SNR gain of about 9 dB for speech signal and 3 dB for music compared to the sequence order of $K = 15$ (top figure), at least for the values of SNR centered on -6 dB. This means that if the sequence length is increased, there is better chance of finding segments with lower energy; therefore, more noise components are excluded. Apparently, for low input SNR values (below -9 dB) and for $K = 17$ better results are achieved for all kind of signals (this can easily be seen by the shape of each set of curves related to the different MLS order).

The results in Fig. 4 show that the improvement of SNR yielded by the hybrid MLS technique, when compared to the standard MLS technique, is significant even in the case of moderate SNR, up to 0 dB. For high SNR values above 0 dB, the noise energy is low; thus, both techniques can lead to good results. For very low SNR, below -12 dB, the noise severely contaminates the test signal, thus invalidating the measurements for both techniques in most cases. Overall, the hybrid MLS technique always led to better results when compared to the standard MLS technique.

As can be expected, the results depend closely on the type of noise. The best results were obtained with speech. This is explained by the strong non-stationary behavior of speech signals and their short bandwidth when compared to the broad band of musical signals. The time envelope of the speech signals allows the segmented MLS algorithm time analysis to find lower energy segments, and the short-term stationary voiced part, made up essentially of tonal components, is analyzed by the filter bank using the Spectral MLS algorithm.

An additional experiment was conducted using a John F. Coltrane jazz track for disturbance with several input SNR values. Fig. 5 shows the results for the estimation of the RIR magnitude frequency response (left side) and RIR energy decay (right side). Fig. 5a1 and a2 correspond to the segmented MLS technique – full audio band; (b1), (b2): octave bands; (c1), (c2): 1/3 octave bands. The input SNR was (i) -9 dB – in green and (ii) -15 dB – in red.

Fig. 5, right side, also depicts the linear regression of the energy decay curve for the estimation of the RT60 and the background noise energy, only the initial part of the RIR (up to about 0.17 s) is shown for simplicity. The SNR is estimated by the gap between the maximum energy and the noise floor values given by these straight lines.

The results show that for the standard MLS technique, the error of the estimated RIR is considerably higher – up to 15 dB – over a wide frequency range. This technique is more vulnerable to the low frequency bands, with the first two octave bands showing a magnitude deviation of about 20 dB. For the segmented MLS technique

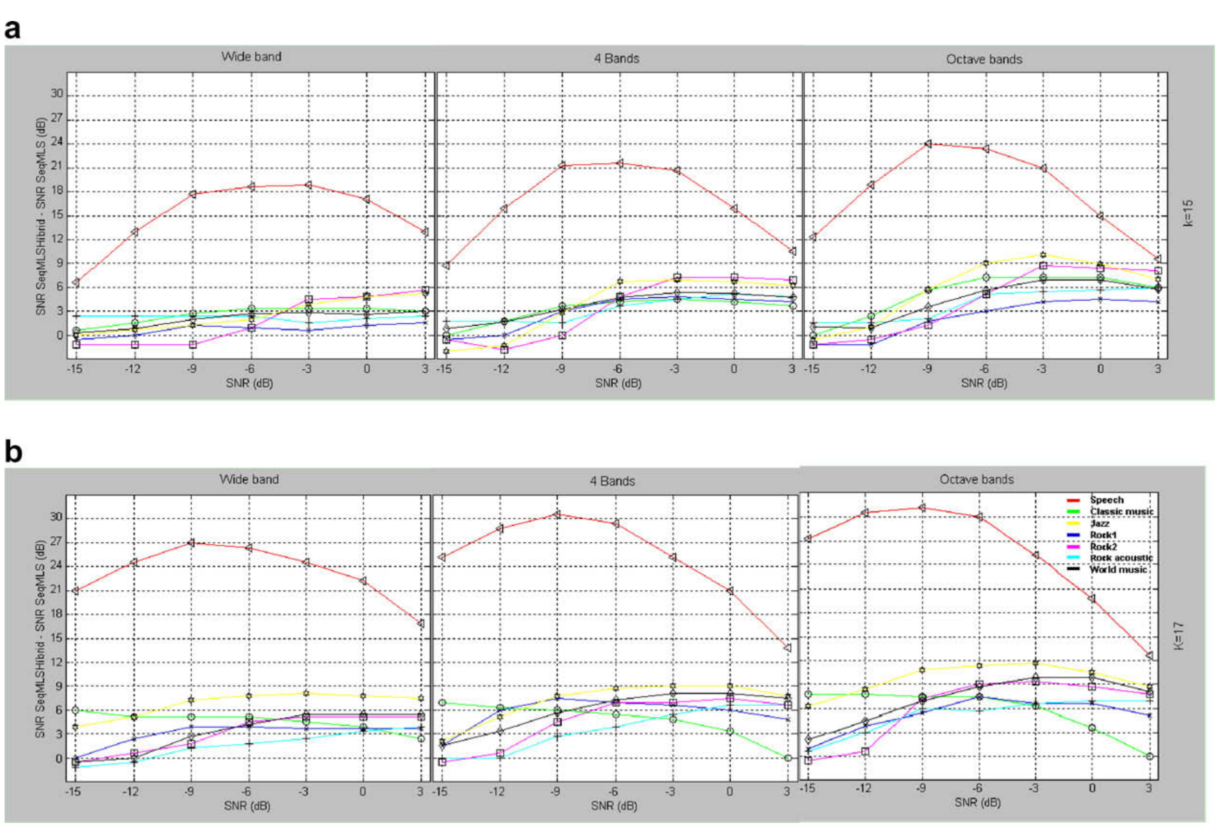

Fig. 4. SNR gain of the hybrid MLS technique against the standard MLS technique vs. input SNR. The curves correspond to different types of ambient noise (i) speech (Δ), (ii) classical music ($*$), (iii) jazz (♦) and, (iv) rock music (□). The topmost figure correspond to a sequence order of $K = 15$ and the bottom figure to $K = 17$. The first column refers to the segmented ML sequence and the 2nd and the 3rd columns to the spectral ML sequence.

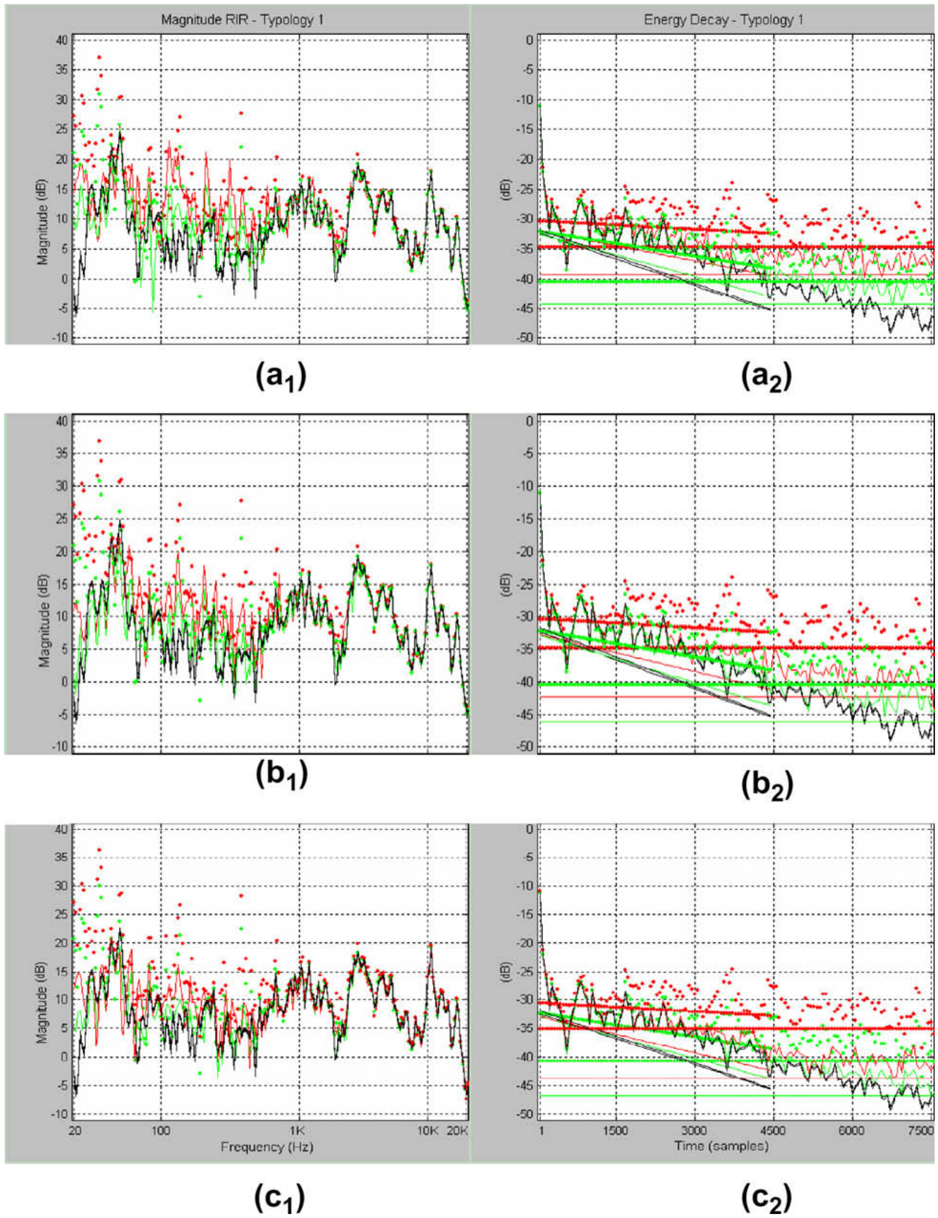

Fig. 5. Estimated RIR magnitude response and RIR energy decay using the coltrane jazz track as noise. The exact response corresponds to the black curve. hybrid MLS technique – filled curves, standard MLS technique – dotted curves. Horizontal straight lines – background noise energy and sloped lines – linear regression of the energy decay curve from the early part of the RIR.

(full audio band), the error is reduced to about 6 dB and for the Spectral MLS technique to 1 dB for a narrower frequency range.

The improvement on the estimated RIR Energy Decay provided by the hybrid MLS technique is significant. A figure of about 9 dB was achieved for the SNR gain. Fig. 5 further shows that with the standard MLS technique, the slope deviation from the exact curve indicates unreliable results for the RIR Energy Decay, thus invalidating the acoustical measurements. In contrast, with the hybrid MLS technique, the decay energy lines are quite coincident and produce minimal error. In situations of very low SNR, the criteria for defining the end point time for the estimation of the energy decay slope is usually a difficult task. In this study, this point was defined by inspection of the energy decay curves. The estimated RT60 and background noise energy for the different techniques are summarized in Table 1. Although these parameters are dependent on the noise profiles, the estimated results are representative enough to evidence the benefits of the new technique.

Table 1
Estimated RT60 and background noise

		RT60 (s)		Backg noise (dB)	
Exact		0, 51		–	
Hybrid MLS	Broadband	0.87	0.61	−44.2	−39.2
	Octave bands	0.68	0.58	−46.3	−42.8
	1/3 Octave	0.61	0.58	−46.9	−43.8
Standard MLS		2.45	0.87	−40.5	−34.9
		−15 (dB) SNR	−9	−15	−9

4. Discussion

4.1. Time domain analysis

The results show that with the hybrid MLS technique there was an increase in the SNR for all situations tested. With all the types of signals used, while maintaining the same SNR, the background noise level was lower than for the standard MLS. This is borne out by the horizontal straight lines (linear regression of background noise energy of the RIR) for the estimated RIR Energy Decay curve in Fig. 5.

A SNR gain of about 18 dB was achieved for speech situations. By increasing the noise bandwidth and using signals with a more stationary time history, such as music material, the SNR gain dropped to less than 5 dB.

The RT60 can be obtained from the oblique straight lines evaluated from the early part of the estimated RIR Energy Decay curve. For an input SNR = −15 dB, the standard MLS technique cannot estimate the RT accurately. On the other hand, with the new technique minimal error is obtained.

4.2. Frequency domain analysis

The results clearly show the influence of the spectrum of the noise component. For the standard MLS technique, the noise spectrum appears almost in the same original frequency range. However,, the hybrid MLS technique spreads the noise components over a large range of the audio spectrum. This indicates a better intra-sequence noise decorrelation owing to a rearrangement of the segments among the sequences.

The RIR estimation is almost identical to the exact RIR curve, when speech is used. For rock music situations the results were not as good. This is a consequence of the spectral noise contents. Nevertheless, in the high frequency bands, the error was found to be negligible. These conclusions can be derived by the RIR magnitude frequency response and by the magnitude response error (not presented in this paper).

By splitting the audio spectrum into several bands, the Spectral MLS technique significantly reduces the noise influence, thereby increasing the gain of the SNR by a 6 dB figure in some cases.

5. Conclusions and final remarks

A modified MLS measurement method, termed as the hybrid MLS technique, comprising approaches both in the time and in the frequency domains, was developed. This technique is applied to situations where the acoustical room parameters have to be measured under conditions of non-stationary ambient noise, for low or very low SNR, and subtle time variance.

In our experiments, the results obtained under real conditions using the hybrid MLS technique fared better than those achieved using the standard MLS technique, since they resulted in a significant increase in the SNR and lower magnitude frequency response error.

Experiments were conducted for the full audio band, for octave and 1/3 octave bands. Different types of noise (speech, classical music, jazz and, rock music) with different SNR values were used. The sequence order was also investigated for comparison purposes. The results of our experiments showed that SNR improvement is highly dependent on the type of the real performance signal. For example, for speech, a SNR gain of the hybrid MLS technique over the standard MLS technique of about 24 dB can be obtained. In the case of jazz, the SNR gain approaches 10 dB. For rock music, a figure of 6 dB is achieved. These results are linked to the time fluctuations of the signals and their bandwidth when compared to the broadband musical signals. For speech signals, the proposed algorithm finds segments with lower energy and the vowel part, essentially made up of tonal components, which are analyzed more efficiently by the filter bank.

The estimated magnitude frequency response and the energy decay of the RIR were investigated, and the error associated with the hybrid MLS technique was shown to be lower than with the standard MLS technique.

For a pre-defined overall SNR, with the measurement setup adjusted accordingly, a lower number of MLS frames need to be sent to the room being tested. This fact constitutes a noteworthy advantage over the standard MLS technique in situations of slight time variance.

References

[1] Rife DD, Vanderkooy J. Transfer-function measurement with maximum-length sequences. J Audio Eng Soc 1989;37(6).
[2] Vorlander, Michael, Mommertz, E., Guidelines for the application of the MLS technique in buildings acoustics and in outdoor measurements. In: Proceedings of the inter-noise 1997, Budapest, Hungary; August, 1997.
[3] ISO 140. Acoustics – measurement of sound insulation in buildings and of building elements, part 1–12, Draft; 1996.
[4] ISO/DIS 3382. Measurement of reverberation time of rooms with reference to other room acoustical parameters, Draft; 1996.
[5] ISO 3382-1975, Measurement of the acoustical characteristics of rooms for speech and music, International Standards Organization, Geneva, Switzerland; 1975.
[6] ISO 3741, Acoustics – determination of sound power levels of noise sources – precision methods for broad-band sources in reverberation rooms; 1988.
[7] Schroeder MR. New method of measuring reverberation time. J Acoust Soc Am 1965;38:329–61;
Schroeder MR. New method of measuring reverberation time. J Acoust Soc Am 1966;40:549–51.
[8] Bradley JS. Optimizing the decay range in room acoustics measurements using maximum-length-sequence techniques. J Audio Eng Soc 1996;44:266–73.
[9] Vorlander M, Kob M. Practical aspects of MLS measurements in building acoustics. Acust – Acta Acust 1996;83(Supplement):S130.
[10] Borish J, Angell JB. An efficient algorithm for measuring the impulse response using pseudorandom noise. J Audio Eng Soc 1983;31:478–88.
[11] Paulo J, Martins CR, Bento Coelho JL. Room impulse response using segmented maximum length sequences (MLS) for annoyance minimization. In: Proceedings of the tecni-acustica – EAA symposium on architectural acoustics, Madrid; 2000.
[12] Paulo J, Martins CR, Bento Coelho JL. Hybrid M sequences for room impulse response estimation. In: Proceedings of the audio engineering society, 115th convention, New York; October 10–13, 2002.
[13] Paulo J, Martins CR, Bento Coelho JL. Spectral hybrid M sequences for room impulse response estimation. In: Proceedings of the ICSV10, Stocholm; 2003.
[14] Paulo J. Msc. thesis. "Método para Estimação da Resposta ao Impulso de Salas baseado na Técnica de Sequências de Comprimento Máximo, MLS, Lisboa; 2004 [in Portuguese].
[15] Chu WT. Comparison of reverberation measurements using Schroeder's impulse method and decay-curve averaging method. J Acoust Soc Am 1978;63:1444–50.
[16] Faiget L, Legros C. Optimization of the impulse response length: application to noisy and highly reverberant rooms. J Audio Eng Soc 1998;46(9):741–9.
[17] Beranek L. Acoustical measurements. Acoustical Society of America; 1988.
[18] Kuttruff H. Room acoustics. 3rd ed. London, New York: Elsevier Applied Science; 1991.
[19] Vorlander M, Bietz H. Comparison of methods for measuring reverberation time. Acustica 1994;80:207–17.