A SYSTEM-IDENTIFICATION-ERROR-ROBUST METHOD FOR EQUALIZATION OF MULTICHANNEL ACOUSTIC SYSTEMS

Wancheng Zhang, Emanuël A. P. Habets, and Patrick A. Naylor

Imperial College London, Department of Electrical and Electronic Engineering

ABSTRACT

In hands-free communications, speech received by a microphone is distorted by room reverberation that can reduce the intelligibility of speech. An approach to dereverberation is firstly to estimate the impulse responses of the acoustic channels between the speaker and the microphones and secondly to design a multichannel equalization system based on the estimated impulse responses. Traditional equalization techniques are designed without the consideration of estimation errors that are commonly introduced by the system identification process. In this work, a System-Identification-Error-Robust Equalization Method (SIEREM) for the equalization of multichannel room acoustic systems is presented. Experimental results for dereverberation using SIEREM applied to estimates of single-input multiple-output acoustic systems with known level of estimation errors show that the proposed equalization design significantly outperforms existing methods in the presence of both synthetic and real system identification errors.

Index Terms— dereverberation, room acoustics, channel estimation error, multichannel equalization, robustness

1. INTRODUCTION

In hands-free communications, in which a speech signal is acquired by microphones placed at some distance from the speaker in a room, the observed speech signals are distorted by room reverberation. Reverberation is the process of multipath propagation of an acoustic sound in an enclosed space and can reduce the intelligibility of speech. The observed reverberant signal generally therefore consists of a direct sound and reflections. The process of dereverberation is to recover the original (dry) speech signal from the signal(s) acquired by a single microphone or multiple microphones. Hereafter, we refer to the multiple acoustic channels from the source to multiple microphones as a multichannel acoustic system.

Acoustic channels are usually modeled as finite impulse response (FIR) filters. An approach to dereverberation is firstly to estimate the impulse responses of the acoustic channels and secondly to design a multichannel equalization system based on the estimated impulse responses [1]. The estimated impulse responses obtained through system identification (SI) always include errors that can be introduced by limitations and inaccuracies of the SI process.

The estimated multichannel acoustic system can be exactly inverted by a set of FIR filters using the multiple-input/output inverse theorem (MINT) [2], subject to specific conditions as explained in Section 3.1. However, the fact that the designed equalization system is able to equalize the estimated acoustic system does not mean that it is able to equalize the true acoustic system. When the designed equalization system is employed to equalize the true acoustic system, the response from the source to the output of the equalization system will deviate from the desired response due to the aforementioned estimation errors. In this paper, our objective is to find a robust equalization system design method that takes into account these system identification errors (SIEs).

The problem is formulated in Section 2. Section 3 reviews existing room acoustics equalization methods and evaluates their performance when SIEs are present. In Section 4, the proposed System-Identification-Error-Robust Equalization Method (SIEREM) is presented. The performance of the proposed method is evaluated in Section 5.

2. PROBLEM FORMULATION

Consider a speech signal $s(n)$ propagating through an M-channel acoustic system $\mathbf{h} = [\mathbf{h}_1^T \; \ldots \; \mathbf{h}_M^T]^T$ where n denotes the discrete time index. The acoustic channel between the source and the mth microphone is characterized by its impulse response $\mathbf{h}_m = [h_m(0) \; h_m(1) \; \cdots \; h_m(L-1)]^T$, $m = 1, \ldots, M$, where $\{\cdot\}^T$ denotes the transpose operation. In this work we assume that the acoustic channels are time-invariant. Using the reverberant speech signals

$$x_m(n) = s(n) * h_m(n) + v_m(n) \quad \text{for } m = 1, \ldots, M, \quad (1)$$

estimates of the room impulse responses (RIRs) $\mathbf{h}_m$ can be obtained using an SI technique [3], where $*$ denotes linear convolution and $v_m(n)$ denotes additive noise at the mth microphone.

The estimate of an acoustic system always includes some errors due to under-modeling of channel order and/or the existence of the additive noise. However, as is common practice in the current literature, we assume the channel orders are known or can be correctly estimated. We also note that the acoustic system $\mathbf{h}$ can only be estimated up to an unknown scaling factor [3]. Without loss of generality and to be consistent with earlier work [3], we express the estimated acoustic system $\hat{\mathbf{h}} = [\hat{\mathbf{h}}_1^T \; \ldots \; \hat{\mathbf{h}}_M^T]^T$ as

$$\hat{\mathbf{h}} = \frac{1}{\gamma}(\mathbf{h} - \mathbf{e}), \quad (2)$$

where γ is defined as

$$\gamma = \frac{\hat{\mathbf{h}}^T \mathbf{h}}{\hat{\mathbf{h}}^T \hat{\mathbf{h}}}, \quad (3)$$

and $\mathbf{e} = [\mathbf{e}_1^T \; \ldots \; \mathbf{e}_M^T]^T$ with $\mathbf{e}_m = [e_m(0) \; e_m(1) \; \ldots \; e_m(L-1)]^T$. Following the definition of γ in (3), the error vector $\mathbf{e}$ in (2) represents the projection misalignment vector [3] that can be geometrically interpreted as the vector from the projection of $\mathbf{h}$ on the linear manifold of $\hat{\mathbf{h}}$ to $\mathbf{h}$ and therefore is independent of the scaling factor introduced by the SI process. In general, an equalization system $\mathbf{g} = [\mathbf{g}_1^T \; \mathbf{g}_2^T \; \ldots \; \mathbf{g}_M^T]^T$ where $\mathbf{g}_m = [g_m(0) \; g_m(1) \; \ldots \; g_m(L_i - 1)]^T$ can be computed that satisfies

$$\sum_{m=1}^{M} \hat{h}_m(i) * g_m(i) = d(i) \quad \text{for } i = 0, \ldots, L + L_i - 2, \quad (4)$$

where $d(i)$ defines the target impulse response. We now define the equalized impulse response (EIR):
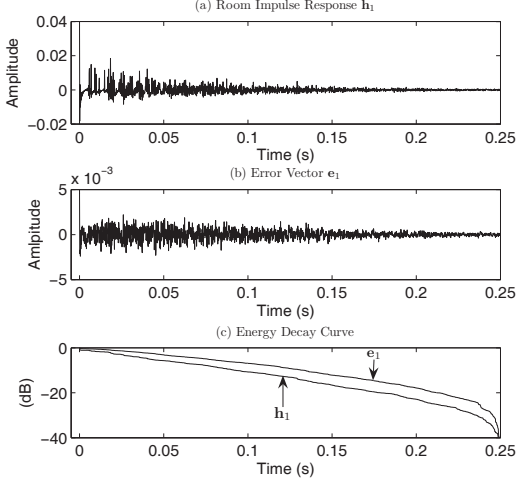
Fig. 1. An example of RIR and SIEs; (a) $\mathbf{h}_1$; (b) $\mathbf{e}_1$, SIEs introduced to $\mathbf{h}_1$; (c) Energy decay curves of $\mathbf{h}_1$ and $\mathbf{e}_1$.

$$b(i) = \sum_{m=1}^{M} h_m(i) * g_m(i) \qquad (5)$$

$$= \gamma d(i) + \sum_{m=1}^{M} e_m(i) * g_m(i). \qquad (6)$$

The EIR, $b(i)$, may still distort the speech signal severely due to the second term in (6). It is this problem that is addressed in this paper.

We now provide an example of the SIEs. A 6-channel acoustic system, the RIRs of which are generated using the image method [4], is identified with the well known NMCFLMS algorithm [3]. The room dimensions are 6.4 m × 5 m × 3.6 m (length × width × height), the positions of the speaker and the central microphone are $(3, 1, 1.6)$ and $(3, 2, 1.6)$ respectively, the intermicrophone distance is 5 cm, the reverberation time T_{60} is 0.6 s and the additive noise is set to signal-to-noise ratio SNR = 30 dB. Without loss of generality, it is assumed that the direct-path propagation time l_1 from the source to the 1st microphone is the shortest. As in [3], the propagation time l_1 was trimmed for all channels in all experiments in this paper[1]. Also, the RIRs used for experiments are truncated to $L = 2000$ corresponding to 0.25 s with a sampling frequency of $f_s = 8$ kHz.

As an example, the true impulse response $\mathbf{h}_1$ and the resulting error vector

$$\mathbf{e}_1 = \mathbf{h}_1 - \gamma\hat{\mathbf{h}}_1, \qquad (7)$$

which is part of $\mathbf{e}$, are shown in Fig. 1(a) and (b) respectively; (c) shows the energy decay curves (EDCs) [5] of $\mathbf{h}_1$ and $\mathbf{e}_1$, where the EDC of an impulse response $\mathbf{a} = [a(0)\ a(1)\ \ldots\ a(L_a - 1)]^T$ is defined as

$$\mathcal{E}(i) = \frac{1}{\|\mathbf{a}\|_2^2} \sum_{j=i}^{L_a - 1} a^2(j), \qquad (8)$$

where $\| \cdot \|_2$ denotes ℓ_2-norm.

It can be seen in Fig. 1(b) that the overall temporal shape of the $\mathbf{e}_1$ can be modeled by a random sequence with an exponential decay. It can also be seen that the EDCs of $\mathbf{e}_1$ and $\mathbf{h}_1$ are similar as shown in Fig. 1(c). The error vectors $\mathbf{e}_m$, for $m \in \{2, \ldots, M\}$, share the above described properties of $\mathbf{e}_1$, which will be used to derive the SIEREM in Section 4.

---

[1]The propagation time l_1 represents only a bulk propagation delay and is not significant to either the SI or the equalization system design.

3. REVIEW OF EXISTING METHODS

3.1. LS and MINT

The equalization system $\mathbf{g}$ can be obtained by solving the system of equations (4), where the target impulse response is given by

$$d(i) = \begin{cases} 0 & \text{if } 0 \le i < \tau; \\ 1 & \text{if } i = \tau; \\ 0 & \text{otherwise,} \end{cases} \qquad (9)$$

with an integer delay τ. In matrix form, (4) can be written as

$$\hat{\mathbf{H}}\mathbf{g} = \mathbf{d}, \qquad (10)$$

where $\hat{\mathbf{H}} = [\hat{\mathbf{H}}_1 \cdots \hat{\mathbf{H}}_M]$ and $\mathbf{d} = [d(0) \cdots d(L + L_i - 2)]^T$ with $\hat{\mathbf{H}}_m$ an $(L + L_i - 1) \times L_i$ convolution matrix of $\hat{\mathbf{h}}_m$.

When two or more channels are employed, exact solution(s) to (10) exist when the following two conditions are both satisfied [2]:

1. $\hat{H}_m(z)$, the z-transforms of the multichannel RIRs $\hat{\mathbf{h}}_m$ do not share any common zeros.
2. $L_i \ge L_c = \lceil \frac{L-1}{M-1} \rceil$ [6][2], where $\lceil \kappa \rceil$ denotes the smallest integer larger than or equal to κ.

If both conditions are satisfied, the solution with minimum ℓ_2-norm can be obtained using

$$\mathbf{g} = \hat{\mathbf{H}}^+ \mathbf{d}, \qquad (11)$$

where $\{\cdot\}^+$ denotes Moore-Penrose pseudo-inverse [8]. If any one or both conditions are violated, (11) gives a multichannel LS solution.

3.2. Weighted LS

The weighted LS method [9] is used for the over-determined cases when the condition 2 is violated. The weighted LS solution is obtained by minimizing the following cost function

$$J = \|\mathbf{W}(\hat{\mathbf{H}}\mathbf{g} - \mathbf{d})\|_2^2, \qquad (12)$$

where $\mathbf{W} = \text{diag}\{\mathbf{w}\}$ denotes the diagonal weighting matrix with $\mathbf{w} = [w(0) \cdots w(L + L_i - 2)]^T$.

The solution is given by

$$\mathbf{g} = (\mathbf{W}\hat{\mathbf{H}})^+ \mathbf{W}\mathbf{d}. \qquad (13)$$

When conditions 1 and 2 are both satisfied and $w(i) \ne 0$, for $i = 0, \cdots, L + L_i - 2$, then the solutions given by (13) and (11) are the same [10].

3.3. Discussion

The approaches reviewed above are generally designed without the consideration of SIEs. The performance of these methods when SIEs are present is now studied for a 2-channel system. The $\hat{\mathbf{h}}_1$ and $\hat{\mathbf{h}}_2$ obtained in Section 2 are used in this study. The normalized projection misalignment (NPM) of $\hat{\mathbf{h}} = [\hat{\mathbf{h}}_1^T\ \hat{\mathbf{h}}_2^T]^T$ with respect to $\mathbf{h} = [\mathbf{h}_1^T\ \mathbf{h}_2^T]^T$ is -12 dB, where NPM is defined as [11]

$$\text{NPM} = \frac{\|\mathbf{h} - \gamma\hat{\mathbf{h}}\|_2^2}{\|\mathbf{h}\|_2^2} \qquad (14)$$

with γ defined in (3). Parameters are set to $L_i = L_c, \tau = 0$.

Fig. 2 shows the EDCs of the EIRs. It can be seen that when SIEs are present, MINT completely fails to equalize the acoustic system. In fact, the EDC of the EIR obtained using MINT is worse than the EDC of $\mathbf{h}_1$. Weighted LS, with $w(i) \ne 0$ for $i = 0, \cdots, L + L_i - 2$, gives the same EIR as MINT.

---

[2]It should be noted that the exact solution(s) to (10) always exist when $L_i \ge L - 1$ [7]. Unfortunately, this cannot be guaranteed when $L_i \ge L_c$. However, it has been proved in [6] that exact solution(s) exist for almost all cases.
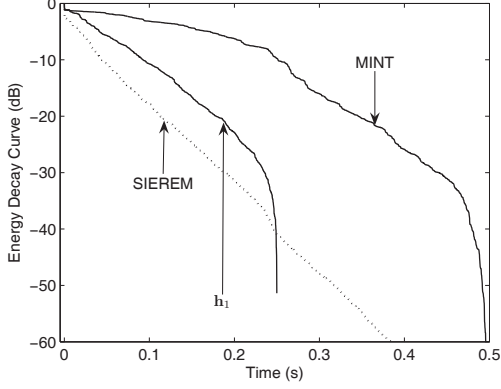
Fig. 2. EDCs of $\mathbf{h}_1$ and EIRs obtained with MINT and the proposed SIEREM.

4. PROPOSED METHOD

In this Section, a System-Identification-Error-Robust Equalization Method (SIEREM) is proposed that takes the robustness of equalization systems to SIEs into consideration.

Since the SI process introduces an unknown scaling factor γ, our equalization system $\mathbf{g}$ is designed to equalize $(1/\gamma)\mathbf{h}$, rather than $\mathbf{h}$. However, assuming that equalizing $(1/\gamma)\mathbf{h}$ gives

$$b'(i) = \sum_{m=1}^{M} \frac{1}{\gamma} h_m(i) * g_m(i), \qquad (15)$$

using (5) we see that the resulting EIR $b(i) = \gamma b'(i)$, i.e., the acoustic system is equalized only up to a scaling factor γ. Therefore, we aim to obtain

$$\mathbf{g}^{\circ} = \arg\min_{\mathbf{g}} \left\| \mathbf{W} \left[(\hat{\mathbf{H}} + \frac{1}{\gamma}\mathbf{E})\mathbf{g} - \mathbf{d} \right] \right\|_2^2 \qquad (16)$$

where $\mathbf{E}$ is formed by $\mathbf{e}$ and has the same form as $\hat{\mathbf{H}}$. Since $\mathbf{e}$ is unknown, $\mathbf{E}$ is also unknown. In order to find $\mathbf{g}$ that minimizes (16), we replace $(1/\gamma)\mathbf{E}$ by an estimate $\hat{\mathbf{E}}$.

As shown in Section 2, the overall temporal shape of $\mathbf{e}_m$ can be modeled to a first approximation by a random sequence with an exponential decay rate similar to that of $\mathbf{h}_m$. Assuming the power of SIEs is uniformly distributed along the channels, we can use $\hat{\mathbf{e}}_m = [\hat{e}_m(0) \ \dots \ \hat{e}_m(L-1)]$, where

$$\hat{e}_m(i) = \beta \cdot \varepsilon_m(i) \cdot e^{-\alpha i}, \qquad (17)$$

to form $\hat{\mathbf{E}}$, where $\varepsilon_m(i)$ is an uncorrelated Gaussian random sequence with zero mean and unit variance, β a multiplicative factor related to the power of $\hat{\mathbf{e}} = [\hat{\mathbf{e}}_1^T \ \dots \ \hat{\mathbf{e}}_M^T]^T$ and the decay rate α is related to the reverberation time T_{60} by [5]

$$\alpha = \frac{3\ln(10)}{\mathrm{T}_{60} \cdot f_s}. \qquad (18)$$

It should be noted that the decay rate α can be estimated from $\hat{\mathbf{h}}_m$.

With different realizations of the sequence $\varepsilon_m(i)$, (17) provides good and bad estimates of $(1/\gamma)\mathbf{E}$, and with $(1/\gamma)\mathbf{E} = \hat{\mathbf{E}}$, the performance of $\mathbf{g}$ obtained from (16) with different realizations of the sequence $\varepsilon_m(i)$ vary much. However, we desire a $\mathbf{g}$ which performs well on average. Therefore, the $\mathbf{g}$ that minimizes

$$J = \mathrm{E}\left\{ \left\| \mathbf{W} \left[(\hat{\mathbf{H}} + \hat{\mathbf{E}})\mathbf{g} - \mathbf{d} \right] \right\|_2^2 \right\} \qquad (19)$$

is preferred, where $\mathrm{E}\{\cdot\}$ represents expectation operation.

The $\mathbf{g}$ that minimizes (19) can be obtained by computing the derivative of J with respect to $\mathbf{g}$ and subsequently solving

$$\frac{\partial J}{\partial \mathbf{g}} = \mathbf{0}_{[ML_i \times 1]}. \qquad (20)$$

Using (19) and (20), we can obtain

$$\mathbf{g} = (\hat{\mathbf{H}}^T \mathbf{W}^T \mathbf{W}\hat{\mathbf{H}} + \mathrm{E}\{\hat{\mathbf{E}}^T \mathbf{W}^T \mathbf{W}\hat{\mathbf{E}}\})^{-1} \hat{\mathbf{H}}^T \mathbf{W}^T \mathbf{W}\mathbf{d}. \qquad (21)$$

The matrix $\mathbf{R} = \mathrm{E}\{\hat{\mathbf{E}}^T \mathbf{W}^T \mathbf{W}\hat{\mathbf{E}}\}$ is a diagonal matrix with $r(j)$, for $j = 1, \dots, ML_i$, on its diagonal, where

$$r((m-1) \cdot L_i + j) = \beta^2 \sum_{i=0}^{L-1} w^2(i+j-1)e^{-2\alpha i} \qquad (22)$$

for $m = 1, \dots, M$ and $j = 1, \dots, L_i$.

Now we derive the relationship between the multiplicative factor β and the NPM of the estimate $\hat{\mathbf{h}}$. Since $\varepsilon_m(i)$ is uncorrelated with $\hat{h}_m(i)$, it can be assumed that $\mathrm{E}\{\hat{\mathbf{h}}^T\hat{\mathbf{e}}\} = 0$, i.e., the vector $\hat{\mathbf{e}}$ is on average orthogonal to $\hat{\mathbf{h}}$.

As shown in [12], we can express the NPM as follows,

$$\mathrm{NPM} = \frac{\|\mathbf{e}\|_2^2}{\|\mathbf{h}\|_2^2} = \sin^2(\theta), \qquad (23)$$

where θ is the angle between the vectors $\mathbf{h}$ and $\hat{\mathbf{h}}$. Alternatively, we have

$$\frac{\|\mathbf{e}\|_2^2}{\|\gamma\hat{\mathbf{h}}\|_2^2} = \tan^2(\theta). \qquad (24)$$

In order to make on average the NPM caused by $\hat{\mathbf{e}}$ equal to the true NPM, we require that

$$\mathrm{E}\{\|\hat{\mathbf{e}}\|_2^2\} = \left\| \frac{1}{\gamma}\mathbf{e} \right\|_2^2. \qquad (25)$$

Using (17), (23), (24) and (25), we can express β as

$$\beta = \frac{\tan\left[\arcsin(\sqrt{\mathrm{NPM}})\right]}{\sqrt{M \cdot \frac{e^{-2\alpha L}-1}{e^{-2\alpha}-1}}} \cdot \|\hat{\mathbf{h}}\|_2. \qquad (26)$$

Following the above derivation, we conclude that given the decay rate α and the NPM of the SI process we are able to design an equalization system that takes into account the SIEs.

5. PERFORMANCE EVALUATION

In this Section, the performance of the proposed SIEREM is compared with existing methods.

For the intelligibility of the equalized speech signal the suppression of late reflections is more important than the suppression of early reflections [13]. Therefore, it is natural to use some weighting function $w(i)$ for which the amplitudes related to the late reflections are larger than those of the early reflections. We have adopted the following intuitively conceived weighting function that has been verified by informal listening tests to provide good perceptual quality in the equalized speech signal:

$$w(i) = \begin{cases} 1 & \text{if } 0 \leq i \leq \tau; \\ e^{\alpha(i-\tau)} - 1 & \text{if } i > \tau. \end{cases} \qquad (27)$$

In order to evaluate the performance of the equalization system the following measures are used in this work. The first performance measure is the EDC as defined in (8). The second performance measure is the early-to-late reverberation ratio (ELR), also known as the Clarity Index, which is defined as [5]

111

Measures	$\mathbf{h}_1$	MINT	SIRERM
C_{50} (dB)	4.31	-2.86	10.33
T_{30} (s)	0.24	0.44	0.19

Table 1. C_{50} and T_{30} derived from the EDCs shown in Fig. 2.
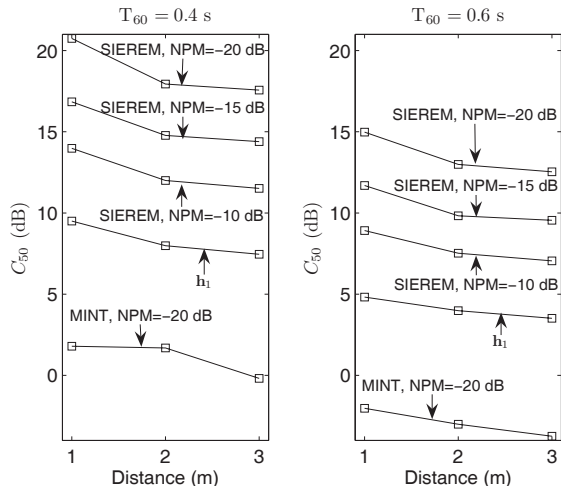
Fig. 3. C_{50} of $\mathbf{h}_1$ and the EIRs obtained with MINT and SIEREM.

$$C_{50} = \frac{\mathscr{E}(0) - \mathscr{E}(n_e)}{\mathscr{E}(n_e)}, \qquad (28)$$

where $n_e = 50$ ms $\cdot f_s$. The third performance measure is T_{30} which is defined as the time interval it takes for the EDC to drop by 30 dB. For the RIRs used in this work it is not always possible to determine properly the time at which the energy decay curve has decreased 60 dB. Therefore, we study the T_{30} rather than the more conventional T_{60} measure.

The EDC of the EIR obtained using SIEREM for the same acoustic system used in Section 3.3 with known NPM and T_{60} is shown by the dashed line in Fig. 2. The C_{50} and T_{30} corresponding to Fig. 2 are given in Table 1. It can be seen that the SIEREM outperforms the MINT. Informal listening clearly concludes that the acoustic system is equalized by SIEREM without introducing in any audible artifacts.

In the second experiment, 2-channel acoustic systems, the RIRs of which are generated using image method [4] with synthetic SIEs of NPM $= -10$ dB, -15 dB and -20 dB, are equalized using MINT and SIEREM. Synthetic error vectors $\mathbf{e}$ are generated such that the overall temporal shape of $\mathbf{e}_m$ resemble random sequences with exponential decay rates corresponding to the T_{60}, and the NPMs of $\mathbf{e}$ meet the desired levels. A systematic way to adjust a generated error vector to the desired NPM can be found in [12] and is used in this work. The distance between the speaker and the nearest microphone is set to 1 m, 2 m and 3 m; the reverberation time T_{60} is set to 0.4 s and 0.6 s. Other parameters are set to $L_i = L_c, \tau = 0$. In each run of the simulation, the relative geometry of the speaker and microphones is kept constant, but the speaker is relocated at a random position in the room. The averaged C_{50} of the resulting EIRs which are obtained using 100 randomly chosen speaker positions and synthetic vectors $\mathbf{e}$, are plotted in Fig. 3. It can be seen that the SIEREM can always equalize the acoustic systems to good effect. However, even with NPM $= -20$ dB SIEs, the resulting C_{50} obtained with MINT are worse than those of the RIRs.

As stated in Section 4, the design parameters for the SIEREM

are the decay rate (or reverberation time) and the NPM level. These parameters are not *a priori* known and therefore need to be estimated. Due to space limitations, elaborate testing results of the robustness of the SIEREM with respect to these design parameters cannot be provided here. In our tests, we have found that the SIEREM is not significantly sensitive to the T_{60} and NPM; estimation accuracy of 20% in T_{60} and ± 10 dB in NPM is sufficient for typical operation.

6. CONCLUSION

In this paper we discussed the equalization of multichannel room acoustic systems, when the multichannel room impulse responses include estimation errors introduced by the system identification process. It was shown that existing equalization techniques fail to equalize the acoustic systems in the presence of system identification errors. Therefore, a new method which takes into account system identification errors was proposed. The evaluation results indicate that the proposed method is more robust to system identification errors than MINT.

ACKNOWLEDGMENT

7. REFERENCES

[1] Y. Huang, J. Benesty, and J. Chen, "A blind channel identification-based two-stage approach to separation and dereverberation of speech signals in a reverberant environment," *IEEE Trans. Speech Audio Processing*, vol. 13, pp. 882–895, 2005.

[2] M. Miyoshi and K. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 145–152, 1988.

[3] Y. Huang and J. Benesty, "A class of frequency-domain adaptive approaches to blind multichannel identification," *IEEE Trans. Signal Processing*, vol. 51, pp. 11–24, 2003.

[4] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.*, vol. 65, no.4, pp. 943–950, 1979.

[5] H. Kuttruff, *Room acoustics*, Taylor & Frances, 4th edition, 2000.

[6] G. Harikumar and Y. Bresler, "FIR perfect signal reconstruction from multiple convolutions: minimum deconvolver orders," *IEEE Trans. Signal Processing*, vol. 46, pp. 215–218, 1998.

[7] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516–525, 1995.

[8] R. Rado, "Note on generalized inverse of matrices," *Proc. Cambridge Philos. Soc.*, vol. 52, pp. 600–601, 1956.

[9] M. Hofbauer, *Optimal linear separation and deconvolution of acoustical convolutive mixtures*, Ph.D. thesis, Swiss Federal Institute of Technology, 2005.

[10] T. N. E. Greville, "Note on the generalized inverse of a matrix product," *SIAM Review*, vol. 8, pp. 518–521, 1966.

[11] D. R. Morgan, J. Benesty, and M. M. Sondhi, "On the evaluation of estimated impulse responses," *IEEE Signal Processing Lett.*, vol. 5, pp. 174–176, 1998.

[12] W. Zhang and P. A. Naylor, "An algorithm to generate representations of system identification errors," *Research letters in signal processing*, vol. 2008, 2008.

[13] E. A. P. Habets, *Single- and multi-microphone speech dereverberation using spectral enhancement*, Ph.D. thesis, Technische Universiteit Eindhoven, 2007.