JAIST Repository
https://dspace.jaist.ac.jp/

Title	室の伝達特性の測定を必要としない室内音響指標のブラインド推定の研究
Author(s)	鵜木，祐史
Citation	科学研究費助成事業研究成果報告書: 1-6
Issue Date	2013-06-02
Type	Research Paper
Text version	publisher
URL	http://hdl.handle.net/10119/11373
Rights	
Description	研究種目：挑戦的萌芽研究，研究期間：2011～2012，課題番号：23650086，研究者番号：00343187，研究分野：聴覚・音声信号処理，科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

様式Ｆ－１９

科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成２５年６月２日現在

機関番号：13302
研究種目：挑戦的萌芽研究
研究期間：2011～2012
課題番号：23650086
研究課題名（和文）室の伝達特性の測定を必要としない室内音響指標のブラインド推定の研究
研究課題名（英文）Study on blind method of estimating room acoustic characteristics without measuring room acoustics

研究代表者
鵜木　祐史（UNOKI MASASHI）
北陸先端科学技術大学院大学・情報科学研究科・准教授
研究者番号：00343187

研究成果の概要（和文）：

本研究では，室内の伝達特性を測定せずに，観測された信号（残響音声など）から室内音響指標（音声伝送指標，残響時間，音の明確性）をブラインド推定する方法を提案した．評価シミュレーションの結果，提案法が原理的に正しく室内音響指標をブラインド推定できることがわかった．また，フィールド実験の結果，人が部屋に居るときでも居ないときでも，提案法が，残響音声から室内音響指標を正確に逆推定できることがわかった．

研究成果の概要（英文）：

This study proposed simplified and specified methods of blindly estimating the room acoustic characteristics from observed reverberant speech signals, based on the concept of the modulation transfer function. Results of simulations revealed that the proposed approach could blindly estimate the room acoustic characteristics. Results of experiments also revealed that the proposed approach could be used to effectively estimate the room acoustic characteristics from reverberant speech signals in various room acoustics even if people were in the room.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
交付決定額	2,900,000	870,000	3,770,000

研究分野：聴覚・音声信号処理
科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス
キーワード：室内音響指標，変調伝達関数，音声伝送指標，ブラインド推定，残響特性

１．研究開始当初の背景

音バリアフリーやユニバーサルデザインの志向から，利用者に係わらず円滑な音声コミュニケーションを実現するための要素・統合技術の確立が急務となっている．そのため，音環境設計における質的な検討（音声の明瞭性の向上，誘導音の検知力向上など）は，最重要課題である．

室内音響の設計を行なう際，古典的にみると最適残響時間（中心周波数が 500 Hz のオクターブ帯域内の残響時間と室容積の関係）の利用が主であったが，現在は室内インパルス応答から求められる室内音響指標を利用することが主となっている．代表的なものとして，ISO 3382-1（1997 改訂）では，残響時間（T_{60}）や音の透明性(C 値)，音の明確性（D 値）があり，IEC 60268-16 では，音声伝送指標（STI: Speech Transmission Index）がある．これらは，音声了解度や聴き取り難さといった室内音響の「質」を主観的に評価するのに重要なものであり，音バリアフリーを目

指した室内音響設計では非常に重要な検討項目である．しかし，室内インパルス応答の実測を前提とするため，聴力保護等の目的から人を排除して，人がいない室の特性が比較的安定しているうちに測定するか，あるいは複数回測定した後でその平均値を実測値として利用しなければならない．そのため，人を排除できないような音環境（駅構内や空港内といった公共環境）や，人や物の出入りが激しく，時々刻々変化するような音環境には，現状の手法を適用することができない．

これまでのところ，室内音響指標を逆推定する試み（室内音響特性を統計的に学習する方法を除く）は非常に少ない．提案された手法の多くは，学習ベースの方法であり，大量の測定データを必要とする点が問題である．これらの方法では，人を排除できないような環境の音データを必要とするため，根本的な解決には至っていない．

研究代表者は，これまでに，上記の問題を解決するため，系の測定をせずに観測した信号から残響時間 T_{60} を逆推定する方法を開発した（特許第 5077847）．本技術は非常に独創性の高い研究として認められ，応用展開が期待されている．この方法は，変調伝達関数（MTF）の概念に基づいているが， STI の導出には MTF も利用されるため，本方法を拡張することで，T_{60} だけでなく STI も同時予測が可能である．

２．研究の目的

本研究の目的は，室内の伝達特性を測定せずに，観測された音信号（音声や音楽といった身の回りの音）そのものから室内音響指標（残響時間 T_{60}，音声伝送指標 STI，音の明確性 D 値）を逆推定する方略を明らかにすることである．これらの客観的な指標は，一般的な室内音響の特徴付けだけでなく，音環境設計に関する主観的な「質」（例えば，音声の明瞭性や聞き取り難さ）を問うときに必要不可欠である．これらは国際規格（ISO 3382-1）に従い，室内インパルス応答を測定した上で算出しなければならないが，人を排除できない音環境や室の特性が時々刻々変化する音環境には適用できない．提案法を実現できれば，例えば，駅や空港といった音の明瞭性・了解性の悪い音環境でも，簡便で即時的な室内音響指標を推定でき，安心安全な音環境改善策を提言できる．

３．研究の方法

本研究では，図１に示す変調伝達関数に基づき，室内音響指標（残響時間，STI，D 値）をブラインド推定する方法を確立するために，申請者によって既に提案された「MTF に基づいた残響時間のブラインド推定法」を足がかりとして，次の３つの課題に取り組んだ．

(1) 室の統計的な音響特性を調査して，変調伝達関数のモデル化を行う．
(2) 様々な室の音響特性を網羅して，MTF の関数モデルの精緻化を行う．
(3) MTF の逆推定法を加えた室内音響指標のブラインド推定法を構築する．

課題(3)の洗練化の作業には，課題(1)ならびに課題(2)との間で主観・客観評価を通じてインタラクティブに行った．最後に，国際規格（ISO 3382-1 や IEC 60268-16）に準拠した測定システムと提案法のフィールドテストを行ない，高精度に室内音響指標を逆推定可能であることと，既存の測定システムを適用できない音環境において正確に室内音響指標を推定可能であることを実証した．

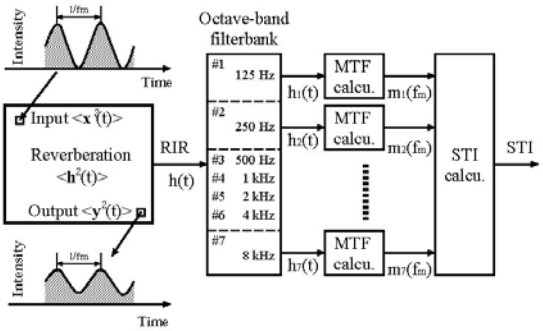

図１　変調伝達関数の概念に基づいた入出力信号の関係と MTT/STI の導出

４．研究成果

(1) 課題(1)と課題(2)に関する結果

課題(1)では，SMILE2004 データベース（建築と環境のサウンドライブラリ）を利用して，多種多様な室内インパルス応答を分析し，MTF のモデル化を行なった．先に検討した室内インパルス応答（RIR）は Schroeder の統計的インパルス応答であり，

$$h(t) = a\exp(-6.9t/T_R)\mathrm{c}(t) \qquad (1)$$

と表された．ここで a は振幅項のパラメータであり，T_R は残響時間である．これに対し，本研究では一般化 RIR モデルとしてこのモデルを拡張した．

$$h(t) = at^{-(n-1)}\exp(-6.9t/T_R)\mathrm{c}(t) \qquad (2)$$

ここで，n は RIR の時間包絡線の過渡部の上昇度合いを表すパラメータである．従来の Schroeder の室内インパルス応答の近似特性（指数型の減衰特性）にパラメータ n を一つ

加えることで初期の過渡特性も表現できるようにモデルを拡張した．

それぞれの MTF はこれらの正規化したパワースペクトルを求めることで得られる（a=1）．

①Schroeder の RIR モデルの MTF：

$$m(f_m, T_R) = \left[1 + \left(2\pi f_m \frac{T_R}{13.8}\right)^2\right]^{-(1/2)} \quad (3)$$

②一般化 RIR モデルの MTF:

$$m(f_m, T_R, n) = \left[1 + \left(2\pi f_m \frac{T_R}{13.8}\right)^2\right]^{-(2n-1/2)} \quad (4)$$

MTF の逆推定で音源特性（音声の特徴）を利用するため，室内音響測定と音源特性の再分析ならびにこれらの対応関係の調査も行った．その代表的な結果を図２に示す．

図２は，一般化 RIR が実測の RIR を正確に近似できているかどうかを確認するために，二つの RIR モデルの時間包絡線と MTF を実測の RIR の時間包絡と MTF に適合したものである．図２(a)から，一般化 RIR モデルが Schroeder の RIR モデルよりも実測の RIR の時間包絡線を正確に近似できていることがわかった．また，図２(b)から，一般化 RIR モデルが Schroeder の RIR モデルよりも実測の RIR の MTF を正確に表現できていることがわかった．

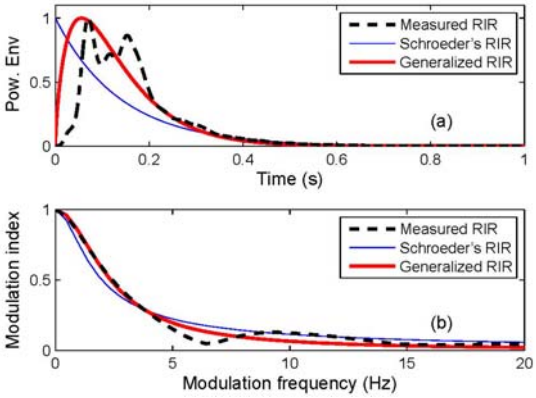

図２　実測の RIR を二つの RIR モデルで適合したときの結果：(a) RIR のパワーエンベロープと(b) 変調度の変化（MTF）．

(2) 課題(3)に関する結果

課題(3)として，モデル化された MTF から室内音響指標を算出する方法を検討した．残響時間の推定法をベースに残りの指標（STI，D 値）の推定法を実装し，室内インパルス応答の統計的モデル化と室内インパルス応答と MTF の対応づけを評価した．図３は，提案法もブロックダイアグラムである．

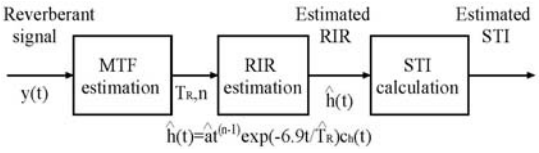

図３　提案法のブロックダイアグラム

図３の MTF 推定（1 段目）では，式(3)を利用した場合に簡便法（従来法），式(4)を利用した場合に提案法となる．図 3 の RIR 推定（2 段目）では，式(1)を利用した場合に簡便法（従来法），式(2)を利用した場合に提案法となる．いずれも，観測された信号の変調スペクトルと式(3)ないし式(4)の MTF を RMS 最小化の意味で適合させることでパラメータ推定を行う．得られたパラメータ（簡便法では T_R のみ，提案法では T_R と n）を利用して，IEC 60268-16 の計算法に基づき STI を算出する．D 値に関しては，推定された RIR から ISO 3382-1 の計算に基づき算出する．

図４に残響音声を利用したパラメータ推定時の変調伝達関数の推定結果の一例を示す．図中の上段は，RIR の時間包絡線に周期性がみられる信号を利用したときの結果，下段は残響音声を利用したときの結果である．また，左段はクリーンな音声，右段は残響環境で観測されたときの結果である．いずれも MTF の推定がうまくいっていることがわかる．

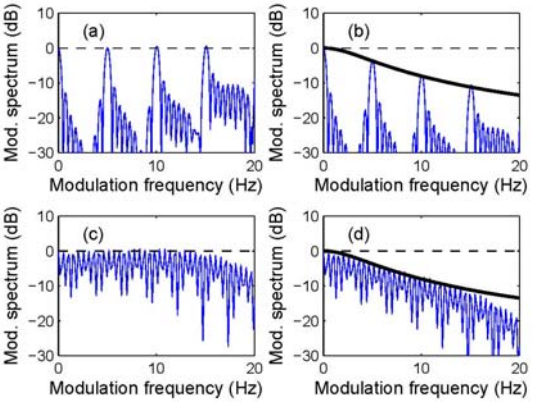

図４　残響音声からの変調伝達関数の推定

(3) モデル原理の評価

課題(3)の成果を検証するために提案法ならびに簡便法の推定評価を行った．ここでは課題(1)で利用した SMILE データベース中の 43 個の RIR を利用した．まず，図５に人工的な AM 信号（変調周波数 5 Hz，雑音キャリア）を利用した場合の音声伝送指標（STI）の推定結果を示す．図中の・が簡便法（従来法）の推定値を，○が提案法の推定値を示し，番号は SMILE データベースの RIR 番号である（１～43 番）．図中の RMS は実効誤差を示す．

推定値は図中の直線上にある場合，最適推定であるといえる．

結果から，提案法ならびに簡便法によりSTI をブラインド推定可能であることがわかった．また，RMS 誤差から，提案法が簡便法よりも正確に推定できることもわかった．ここでは，STI のみの検討であり，残響時間ならびにD 値の推定結果は割愛した．なお，評価シミュレーションで利用したRIR の情報は表１に要約した．

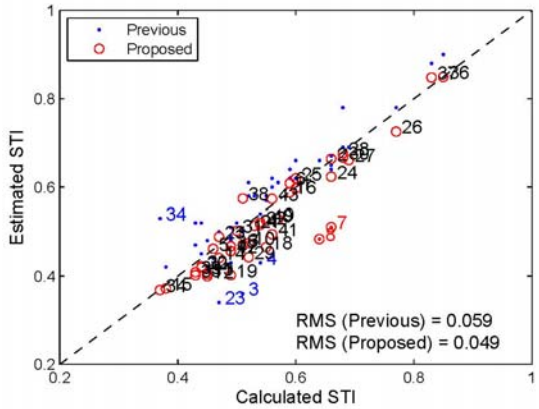

図５　残響 AM 信号からの STI 推定の結果

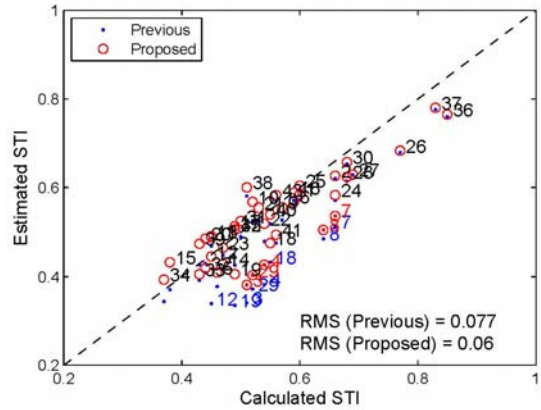

図６　残響音声からの STI 推定の結果

次に図６に残響音声信号からの推定結果を示す．図中のフォーマットは図５と同じである．図６にて推定値が理想直線上の近くにあることから，提案法ならびに簡便法とも残響音声からSTI を推定可能であることがわかった．RMS 誤差の値から，提案法が簡便法よりも正確に推定可能であることもわかった．

このときの残響時間ならびにD 値の推定結果を図７ならびに図８に示す．図中の縦棒は推定結果の標準偏差を，その中心は平均値を示す．図中の曲線は推定値の回帰曲線（2 次関数）である．これらの結果から，いずれも残響音声から推定可能であるが，理想値（直線上）に完全に一致しているわけではない．RMS 誤差からみると，両者とも改善の余地が残されているがこれは今後の課題とする．

表１　STI ブラインド推定法の評価実験で利用した室内インパルス応答データベース

ID	Room condition	RIR No.	T_{60} [s]
1	MPH 1 (with RB)	301	1.09
2	MPH 1 (without RB)	302	0.80
3	MPH 2 (with RB)	303	1.44
4	MPH 2 (without RB)	304	1.04
5	MPH 3 (with RB)	305	1.93
6	MPH 3 (without RB)	306	1.35
7	MPH 4 (with AB)	307	1.42
8	MPH 4 (without AB)	308	1.54
9	MPH5(14, 000 m3)	319	1.47
10	MPH 6 (19, 000 m3)	321	2.16
11	CCH 1 (5, 600 m3)	309	2.35
12	CCH 1 (d = 6 m)	310	2.34
13	CCH 1 (d = 11 m)	311	2.35
14	CCH 1 (d = 15 m)	312	2.39
15	CCH 1 (d = 19 m)	313	2.38
16	CCH 2 (6, 100 m3)	314	1.14
17	CCH 3 (20, 000 m3)	315	1.96
18	CCH 4 (with AC)	316	1.92
19	CCH 4 (without AC)	317	2.55
20	CCH 5 (17, 000 m3)	323	2.32
21	CCH 6 (1F front)	324	1.77
22	CCH 6 (2F side)	325	1.74
23	CCH 6 (3F)	326	1.69
24	Lecture room	201	1.36
25	Theater hall (3,900 m^3)	318	0.85
26	Meeting room (130 m^3)	401	0.62
27	Lecture room (400 m^3)	402	1.12
28	Lecture room (2,400 m^3)	403	1.09
29	GSH (11,000 m^3)	404	1.54
30	Church 1 (1,200 m^3)	405	0.71
31	Church 2 (3,200 m^3)	406	1.30
32	Event hall 1 (28,000 m^3)	407	3.03
33	Event hall 2 (41,000 m^3)	408	3.62
34	Gym 1 (12,000 m^3)	409	2.82
35	Gym 2 (29,000 m^3)	410	1.70
36	Living room (110 m^3)	411	0.36
37	Movie theater (560 m^3)	412	0.38
38	Atrium (4,000 m^3)	413	1.57
39	Tunnel (5,900 m^3)	414	2.72
40	Concourse in station	415	1.95
41	GSH 2 (1F front)	416	1.53
42	GSH 2 (1F center)	417	1.49
43	GSH 2 (1F balcony)	418	1.40
44	Seminar Room (I-95)	-	0.45 (0.55)
45	AV Laboratory (I-94)	-	0.54 (0.38)
46	IS Lecture Hall	-	0.53 (0.57)
47	IS Lecture Room (I3-4)	-	0.63 (0.47)

（備考）MPH: Multi-purpose hall, CCH: Classic concert hall, GSH: General speech hall, RB: Reflex board, AB: Absorption board, AC: Absorption curtain
ID 44-47 は研究機関の教室等．残響時間の括弧内の数値は人が居る時の結果．

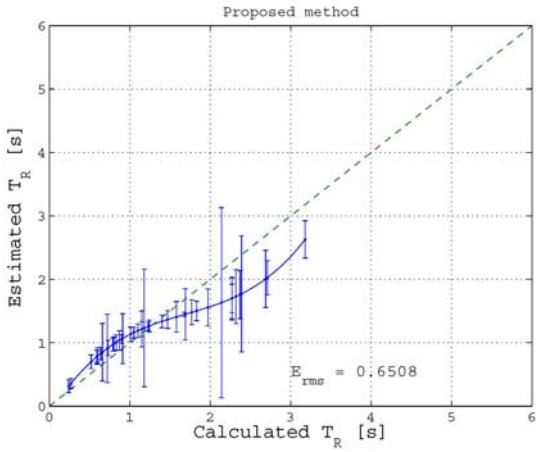

図７　残響音声からの残響時間の推定結果

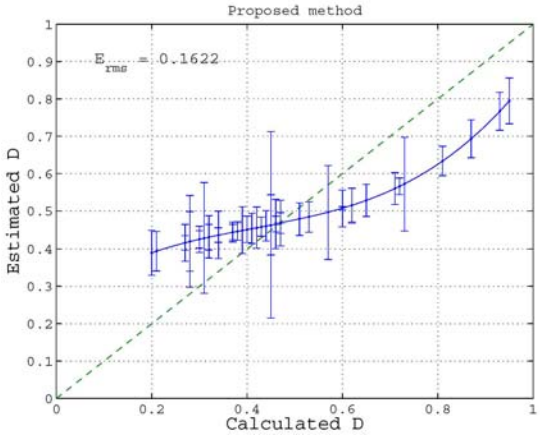

図８　残響音声からのＤ値の推定結果

(4) 総合評価

最後に，平成 23 年度に購入した室内音響測定システム（図９）を利用して，実環境における室内音響指標のブラインド推定法を評価した．この結果，実測の室内インパルス応答の諸特性と同等の特性を推定可能であり，また人がいる環境でも観測信号から正確に MTF/STI をブラインド推定できることを確認した（詳細は表１の 44～47 番を参照）．

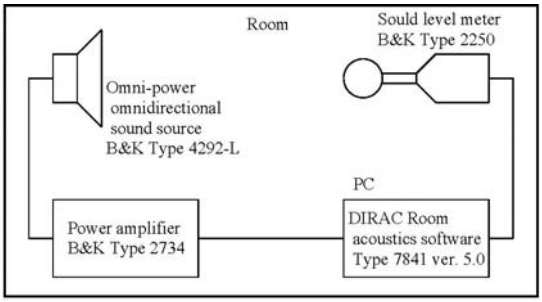

図９　室内音響特性の測定装置

このときの STI の推定結果を図に示す．図中の番号は表１の RIR の番号に対応する．図中の・は人が居ない場合の結果を，×は人が居る場合の結果を示す．推定値はやや理想直線より下にあるが，人が居る，居ないにかかわらず，実環境で残響音声信号を利用して STI を推定できることが確認できた．ここでは残響時間とＤ値の推定結果を割愛するが，図７と図８に掲載したように，若干推定精度が低下する傾向にあった．

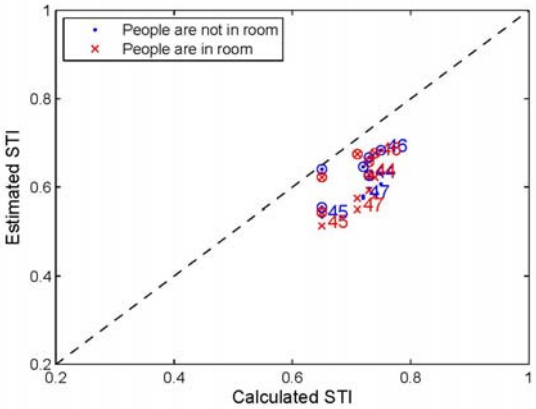

図 10　実環境で推定された STI の結果

(5) まとめ

以上より，MTF の概念に基づいた残響音声からの室内音響指標のブラインド推定法を提案することができた．特に，音声伝送指標 STI の推定に関しては ISO 3382-1 に従った室内インパルス応答の測定ならびに IEC 60268-16ISO に従った STI 計算法から得られた値を，観測された残響音声のみからブラインド推定可能であることを評価シミュレーションならびにフィールド実験により確認することができた．本研究成果より，ブラインド推定に関する骨子は定まった．

今後は本研究で明らかになったいくつかの残された諸問題を解決することで，提案法の精度を高めていく．

５．主な発表論文等

〔雑誌論文〕（計５件）

1. Masashi Unoki, Tomohiro Ikeda, Kyohei Sasaki, Ryota Miyauchi, Masato Akagi, and Nam Soo Kim, “BLIND METHOD OF ESTIMATING SPEECH TRANSMISSION INDEX IN ROOM ACOUSTICS BASED ON CONCEPT OF MODULATION TRANSFER FUNCTION,” Proc. ChinaSIP2013, 査読有，採録済，発表予定(2013/07).
2. Kyohei Sasaki and Masashi Unoki, “Study on Method of Estimating MTF/STI Using Generalized Room

Impulse Response," Proc. 2013 RISP International workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP13), 査読有, pp. 249-252 (2013/03).

3. Masashi Unoki and Xugang Lu, "Unified denoising and dereverberation method used in restoration of MTF-based power envelope," Proc. ISCSLP2012, 査読有, pp. 215-219, Hong Kong (2012/12). DOI: 10.1109/ISCSLP.2012.6423499
4. Yasuaki Kanai and Masashi Unoki, "Study on robust voice activity detection using empirical mode decomposition and modulation spectrum analysis," Proc. ISCSLP2012, 査読有, pp. 400-404, Hong Kong (2012/12). DOI: 10.1109/ISCSLP.2012.6423519
5. Masashi Unoki, Xugang Lu, Rico Petrick, Shota Morita, Masato Akagi, and Ruediger Hoffmann, "Voice activity detection in MTF-based power envelope restoration," Proc. Interspeech2011, 査読有, pp. 2609-2612 (2011/08).

〔学会発表〕(計 9 件)

1. 佐々木 恭平, 鵜木 祐史, "室の音響測定を必要としない音声伝達指標と室内音響特性の推定法の研究," 日本音響学会 2013 年度春季研究発表会講演論文, 1-6-15, pp. 1193-1196, 2013/03/13 (東京工科大, 東京都).
2. Liu Yang and Masashi Unoki, "Improvement of MTF-based power envelope restoration in noisy reverberant Environments," 第 27 回信号処理シンポジウム, pp. 466-471, 2012/11/30 (ANA インターコンチネンタル石垣リゾート, 沖縄県).
3. 鵜木祐史, 森田翔太, ル シュガン, "MTF に基づいたパワーエンベロープ回復処理における統合的な雑音残響除去法," 電子情報通信学会技術報告, EA2012-86, pp. 29-34, 2012/11/16. (大分大, 大分)
4. 佐々木恭平, 鵜木祐史, "室内インパルス応答のモデル化と MTF/STI 推定法の検討," 電子情報通信学会技術報告, EA2012-78, pp. 79-84, 2012/10/28 (牛岳温泉リゾート, 富山県).
5. Masashi Unoki, "Speech Signal Processing Based on the Concept of Modulation Transfer Function," Proc. 2012 IEEE Workshop o Nonlinear Signal Processing (NSP2012), pp. 23, Community Plaza, Shikoku Univ. Tokushima, 25-27, 2012/09/26 (四国大学, 徳島県)
6. 佐々木恭平, 鵜木祐史, "室内インパルス応答のモデル化とその MTF の推定," H24 電気関係学会北陸支部連合大会, 富山県立大学, G7, 2012/09/01 (富山県立大, 富山県).
7. 秋山大知, 鵜木祐史, 金井秀明, "変調スペクトルのリアルタイム分析手法の検討," H24 電気関係学会北陸支部連合大会, G8, 2012/09/01 (富山県立大), 富山県.
8. 鵜木祐史, 池田友洋, 宮内良太, 赤木正人, "変調伝達関数の概念に基づいた音声伝送指標のブラインド推定法の検討," 電子情報通信学会技術報告, EA2011-47, pp. 27-32, 2011/07/08 (キャンパスボート大阪, 大阪府).
9. Masashi Unoki, Tomohiro Ikeda, and Masato Akagi, "Blind Estimation Method of Speech Transmission Index in Room Acoustics," Proc. Forum Acusticum 2011, pp. 1973-1978, 2011/06/29 (Aalborg, Denmark).

〔図書〕(計 0 件)

〔産業財産権〕
○出願状況 (計 0 件)

○取得状況 (計 0 件)

〔その他〕
ホームページ等

６．研究組織
(1)研究代表者
鵜木　祐史 (UNOKI MASASHI)
北陸先端科学技術大学院大学・情報科学研究科・准教授
研究者番号：00343187

(2)研究分担者
赤木　正人 (AKAGI MASATO)
北陸先端科学技術大学院大学・情報科学研究科・教授
研究者番号：20242571

宮内　良太 (MIYAUCHI RYOTA)
北陸先端科学技術大学院大学・情報科学研究科・助教
研究者番号：30455852

(3)連携研究者

なし