独立成分分析を応用したロボット聴覚による残響下におけるバージイン発話認識

○武田龍† 中臺一博‡ 高橋徹† 駒谷和範† 尾形哲也† 奥乃博†
†京都大学大学院情報学研究科 ‡(株) ホンダ・リサーチ・インスティチュート・ジャパン

ICA-based Robot Audition for recognizing barge-in speech under reverberation

*Ryu Takeda†, Kazuhiro Nakadai‡, Toru Takahashi†, Kazunori Komatani†, Tetsuya Ogata†, and Hiroshi G. Okuno†
†Graduate School of Informatics, Kyoto University
‡Honda Research Institute Japan, Co., Ltd.

Abstract— This paper describes a new method based on independent component analysis (ICA) for enhancing a target source and suppressing other interference sources supposed that the latters are known. The method can provide barge-in capable robot audition system in a reverberant environment. We can separate late-reverberations of user's speech and robot's speech by ICA with a multi-input model. To reduce computational complexity to the linear order of the reverberation time, we applied 1) spatial sphering and 2) partial independent component estimation to ICA. Experimental results show that our method outperforms our previous ICA-based known-noise cancelation method.

Key Words: Robot Audition, ICA, Barge-In, late-reverberation, multi-input system

1. はじめに

人とロボットの自然な対話を実現する上で，ロボットの発話中にユーザの発話（**バージイン**）を許容することは不可欠である．ロボットに装着されたマイクにはロボット自身の発話が入り込むため，自分自身の発話はユーザ発話の認識性能を低下させる原因となる．また，環境特性への対処（残響）も必須である．

我々は，バージインを許容するロボット音声対話を実現するため，独立成分分析 (ICA) に基づく適応フィルタ (以下，Semi-Blind ICA と呼ぶ) を用いてロボットの自発話の抑圧を行ってきた [1]．Semi-Blind ICA は，ロボット発話の残響を短時間周波数領域でモデル化することで，計算量の削減と分離性能の両立を実現している．一方，ユーザ発話の残響に関しては何も対策をしていない．そのため，残響環境下での，特に後部残響音が音声認識率低下の原因となっていた（図 1）．

本稿では，Semi-Blind ICA の混合過程を多入力系に拡張することにより，瞬時混合 ICA でユーザ発話の残響とロボット発話の同時抑圧を実現する．ICA を用いる理由は，他の手法とは異なり，ブラインド残響抑圧やブラインド分離との統合が容易かつ自然に行えるからである．また，音響モデル適応などによる対処は，環境に関する事前情報が必要であるため，今回は扱わない．残響抑圧は精度の点でフレーム単位の処理が困難なため，本研究では ICA のリアルタイム実装で行われているブロック処理を想定して設計した [2] （図 2）．

2. Semi-Blind ICA

音の混合モデルは，処理効率などの観点から，すべて短時間フーリエ変換（STFT）を適用後の信号（スペクトル）で記述される．基本的な考え方は，“別フレームの音源は別音源として ICA で分離する”，ことである．

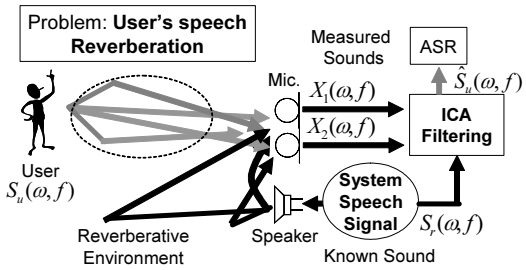

Fig.1 System overview and our problem

2·1 混合モデル

ここでは，ロボット発話の残響に対する混合過程を示す．周波数 ω，フレーム f における，ロボット発話（既知信号）のスペクトルを $s_r(\omega, f)$，マイク 1 での観測スペクトルを $x_1(\omega, f)$ とする時，

$$\begin{pmatrix} x_1(\omega,f) \\ \boldsymbol{S}_r(\omega,f) \end{pmatrix} = \begin{pmatrix} a(\omega) & \boldsymbol{h}^T(\omega) \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} N(\omega,f) \\ \boldsymbol{S}_r(\omega,f) \end{pmatrix}, \quad (1)$$

$$\boldsymbol{S}_r(\omega,f) = [s_r(\omega,f),\ldots,s_r(\omega,f-M)]^T, \quad (2)$$

$$\boldsymbol{h}(\omega) = [h(\omega,0),\ldots,h(\omega,M)]^T, \quad (3)$$

と表現する．ここで，$\boldsymbol{I}$ は $(M+1)\times(M+1)$ の単位行列，$h(\omega, m)$ は遅延フレームが m である周波数成分 $s_r(\omega, f-m)$ の伝達係数，$A(\omega)$ は変数，$N(\omega, f)$ はユーザ発話（雑音）のスペクトルである．

2·2 ICA を用いた既知信号の分離

混合過程が瞬時混合として扱えるため，ICA を適用することでロボット発話を分離する．

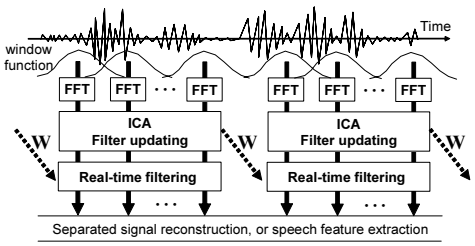

Fig.2 Signal flow in real-time implementation

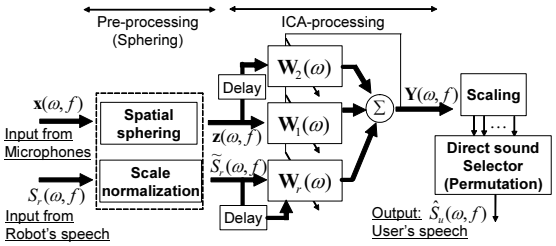

Fig.3 Signal flow in ICA processing

分離過程は次式のようになる．以降，表記の簡単化のため，周波数 ω を省略する．

$$\begin{pmatrix} \hat{N}(f) \\ \boldsymbol{S}_r(f) \end{pmatrix} = \begin{pmatrix} \hat{a} & \boldsymbol{w}^T \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} X_1(f) \\ \boldsymbol{S}_r(f) \end{pmatrix}, \quad (4)$$

$$\boldsymbol{w} = [w(0), \ldots, w(M)]^T. \quad (5)$$

ここで，$\boldsymbol{w}$ は $M+1$ 次の分離フィルタである．

$\hat{N}, \boldsymbol{S}_r$ の結合確率密度と周辺確率密度の積との距離である Kullback-Leibler Divergence（KLD）を最小化することで，分離フィルタを推定する．非ホロノミック拘束と自然勾配法により，以下の学習則が得られる．

$$\boldsymbol{w}^{[j+1]} = \boldsymbol{w}^{[j]} - \mu E[\phi(\hat{N}(f))\boldsymbol{S}_r^H(f)]. \quad (6)$$

ここで，E は期待値演算，^{H} はエルミート転置を表し，非線形関数 $\phi(x)$ は $\tanh(100|x|)e^{\theta(x)}$ を用いる [3].

2·3 問題点と課題

Semi-Blind ICA のモデルでは，ユーザ発話の残響を雑音 $N(f)$ として扱っている．このため，STFT の窓長を超える残響を分離することができず，残響時間の長い環境では音声認識率が低下してしまう．窓長を長くすると逆に分離性能自体が低下してしまう [4].

音声認識を行う場合，窓長内の残響（伝達特性）は乗法性の雑音として扱われ，Cepstrum Mean Normalization（CMN）等で対処できるため，あまり問題にならない [5]．解決すべき課題は次のようになる．

- ユーザ発話の解析窓長に収まらない残響の抑圧

3. ユーザ発話残響への対応

方針 多入力系（複数マイクロホン使用）へ拡張することで，瞬時混合モデルで別フレームに入るユーザ発話の残響を含めることが可能なことを示す．次に，瞬時混合 ICA を用いて残響を抑圧する手順を示す．

3·1 多入力系への拡張

マイク $1, ..., L$ の入力に対し，観測スペクトルを $x_1(f), ..., x_L(f)$ と表す．この時，ベクトル $\boldsymbol{x}(f), \boldsymbol{X}(f)$ およびユーザ発話スペクトル $\boldsymbol{S}_u(f)$ を次式で定義する．

$$\boldsymbol{x}(f) = [x_1(f), \ldots, x_L(f)]^T, \quad (7)$$

$$\boldsymbol{X}(f) = [\boldsymbol{x}(f), \ldots, \boldsymbol{x}(f-N)]^T, \quad (8)$$

$$\boldsymbol{S}_u(f) = [s_u(f), \ldots, s_u(f-K-J)]^T. \quad (9)$$

ここで，ユーザ発話に関する，$L(N+1) \times (K+J+1)$ の伝達特性行列 $\boldsymbol{H}$ を考える．

$$\boldsymbol{h}(i) = [h_1(i), h_2(i), \ldots, h_L(i)]^T, \quad (10)$$

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{h}(0) & \cdots & \cdots & \boldsymbol{h}(K) & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{h}(0) & \cdots & \cdots & \boldsymbol{h}(K) \end{pmatrix}. \quad (11)$$

$L(N+1) = K+J+1$ を満たすとき，$\boldsymbol{H}$ は $L(N+1) \times L(N+1)$ の正方行列となる．つまり，瞬時混合系で記述することができる．ロボット発話（既知信号）を含んだ全体の過程は以下のように表現できる．

$$\begin{pmatrix} \boldsymbol{X}(f) \\ \boldsymbol{S}_r(f) \end{pmatrix} = \begin{pmatrix} \boldsymbol{H} & \boldsymbol{H}_r \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{S}_u(f) \\ \boldsymbol{S}_r(f) \end{pmatrix}. \quad (12)$$

$\boldsymbol{I}$ は $(M+1) \times (M+1)$ 単位行列，$\boldsymbol{H}_r$ は既知信号の $L(N+1) \times (M+1)$ 伝達特性行列である．

$$\boldsymbol{H}_r = \begin{pmatrix} \boldsymbol{h}_r(0) & \cdots & \boldsymbol{h}_r(M) \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{h}_r(0) \end{pmatrix}, \boldsymbol{h}_r(i) = \begin{pmatrix} h_{1r}(i) \\ \vdots \\ h_{Lr}(i) \end{pmatrix} \quad (13)$$

3·2 瞬時混合 ICA による分離フィルタ推定

$\boldsymbol{S}_u(f), \boldsymbol{S}_r(f)$ の各要素が統計的に独立であれば，ICA を用いて残響成分を分離することができる．本節では，1) 強制時間相関除去による近似球面化変換，2) 部分独立成分推定による直接音抽出，により，高収束かつ計算量を抑えた残響・既知信号抑圧手法を説明する．ICA の処理の概要を図 3 に示しておく．

強制時間相関除去による近似球面化 ICA における収束を高速化するため，前処理として球面化変換を行う．球面化は，時空間相関行列 $\boldsymbol{R}$ の固有値 $\boldsymbol{\Lambda}$，および固有ベクトル $\boldsymbol{E}$ を以下のように用いて行われる．

$$\boldsymbol{R} = \begin{pmatrix} E[\boldsymbol{X}(f)\boldsymbol{X}^H(f)] & E[\boldsymbol{X}(f)\boldsymbol{S}_r^H(f)] \\ E[\boldsymbol{S}_r(f)\boldsymbol{X}^H(f)] & E[\boldsymbol{S}_r(f)\boldsymbol{S}_r^H(f)] \end{pmatrix}, \quad (14)$$

$$\boldsymbol{Z}(f) = \boldsymbol{E}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{E}^H \begin{pmatrix} \boldsymbol{X}(f) \\ \boldsymbol{S}_r(f) \end{pmatrix}. \quad (15)$$

この相関行列のサイズは $(L(N+1)+M+1) \times (L(N+1)+M+1)$ であり，固有値分解の計算量は行列サイズの 3 乗オーダである．特に，遅延フレーム N に対して，計算量が $O(N^3)$ となるので，実用的な観点からしても演算量を抑える工夫が必要である．

そこで，次のように強制的に時間相関除去と既知信号・観測信号相関除去を行う．

$$E[\boldsymbol{X}(f)\boldsymbol{X}^H(f)] = \begin{pmatrix} \boldsymbol{R}(0) & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{R}(0) \end{pmatrix}, \quad (16)$$

$$E[\boldsymbol{S}_r(f)\boldsymbol{S}_r^H(f)] = \begin{pmatrix} \lambda_r & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_r \end{pmatrix}, \quad (17)$$

$$E[\boldsymbol{X}(f)\boldsymbol{S}_r^H(f)] = \boldsymbol{0}, \;\; E[\boldsymbol{S}_r(f)\boldsymbol{X}^H(f)] = \boldsymbol{0}. \quad (18)$$

ここで，空間相関行列 $\boldsymbol{R}(0) = E[\boldsymbol{x}(f)\boldsymbol{x}^H(f)]$，分散 $\lambda_r = E[s_r(f)s_r^H(f)]$ である．これらの要素を用いた式（15）の実行は，観測信号には空間的球面化，既知信号にはスケールの正規化を行うことを意味する．

結局，この近似した球面化では，観測信号 $\boldsymbol{X}(f)$ および既知信号 $\boldsymbol{S}_r(f)$ は次の変換を受ける．

$$\boldsymbol{z}(f) = \boldsymbol{V}_0\boldsymbol{x}(f), \;\; \boldsymbol{V}_0 = \boldsymbol{E}_0\boldsymbol{\Lambda}_0^{-1/2}\boldsymbol{E}_0^H \quad (19)$$

$$\tilde{s}_r(f) = \lambda_r^{-1/2}s_r(f). \quad (20)$$

$\boldsymbol{E}_0, \boldsymbol{\Lambda}_0$ は $\boldsymbol{R}(0)$ の固有ベクトルおよび固有値である．

部分独立成分推定　ここでは，L 個の独立成分 $\boldsymbol{y}(f) = [y_1(f), \cdots, y_L(f)]^T$ を抽出するような，次の分離過程を設定する．というのは，通常の ICA の分離過程では，分離行列の次元が $L(N+M+2) \times L(N+M+2)$ となってしまい，計算量が膨大となるためである．

$$\begin{pmatrix} \boldsymbol{y}(f) \\ \boldsymbol{Z}_2(f) \\ \tilde{\boldsymbol{S}}_r(f) \end{pmatrix} = \begin{pmatrix} \boldsymbol{W}_1 & \boldsymbol{W}_2 & \boldsymbol{W}_r \\ \boldsymbol{0} & \boldsymbol{I}_2 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_r \end{pmatrix} \begin{pmatrix} \boldsymbol{z}(f) \\ \boldsymbol{Z}_2(f) \\ \tilde{\boldsymbol{S}}_r(f) \end{pmatrix} \quad (21)$$

$$\boldsymbol{Z}_2(f) = [\boldsymbol{z}(f-1), \cdots, \boldsymbol{z}(f-N)], \quad (22)$$

$$\tilde{\boldsymbol{S}}_r(f) = [\tilde{s}_r(f), \cdots, \tilde{s}_r(f-M)]. \quad (23)$$

ここで，$\boldsymbol{W}_1$ は $L \times L$ 分離行列，$\boldsymbol{W}_2$ は $L \times LN$ 分離行列，$\boldsymbol{W}_r$ は $L \times (M+1)$ 分離行列，$\boldsymbol{I}_2, \boldsymbol{I}_r$ はそれぞれ対応する単位行列である．

次に KLD を自然勾配に基づいて最小化し，独立成分を同時に抽出する．学習則は以下のようになる．

$$\boldsymbol{D} = \text{off-diag}\left(\boldsymbol{I} - E[\boldsymbol{\phi}(\boldsymbol{y}(f))\boldsymbol{y}^H(f)]\right), \quad (24)$$

$$\boldsymbol{W}_1^{[j+1]} = \boldsymbol{W}_1^{[j]} + \mu\boldsymbol{D}\boldsymbol{W}_1^{[j]}, \quad (25)$$

$$\boldsymbol{W}_2^{[j+1]} = \boldsymbol{W}_2^{[j]} + \mu\left[\boldsymbol{D}\boldsymbol{W}_2^{[j]} - E[\boldsymbol{\phi}(\boldsymbol{y}(f))\boldsymbol{Z}_2^H(f)]\right], \quad (26)$$

$$\boldsymbol{W}_r^{[j+1]} = \boldsymbol{W}_r^{[j]} + \mu\left[\boldsymbol{D}\boldsymbol{W}_r^{[j]} - E[\boldsymbol{\phi}(\boldsymbol{y}(f))\tilde{\boldsymbol{S}}_r^H(f)]\right]. \quad (27)$$

μ は学習係数，$\boldsymbol{\phi}(\boldsymbol{x}) = [\phi(x_1), \cdots, \phi(x_L)]^T$ は非線形関数ベクトル，off-diag は対角要素を 0 に置き換える作用素である．Fast-ICA などは，正確な球面化を前提としているため，適用は不適切である．

この過程によって抽出される成分も，$\boldsymbol{Z}_2, \tilde{\boldsymbol{S}}_r$ とは独立な成分，つまり，直接音に相当するものだと期待される．ここで，独立性の観点から，隣接フレームの要素である $\boldsymbol{z}(f-1)$ に対応する分離行列の要素を $\boldsymbol{0}$ に拘束する．これは，直接音と隣接フレーム音との独立性があまり高くないことを考慮するためである．

フィルタ初期値・学習係数　周波数 ω の分離行列の初期値は，周波数 $\omega+1$ で推定された $\boldsymbol{W}_1(\omega+1)$ を用いる．一番最初の分離行列は単位行列とする．学習係数には，焼き鈍し法と指数重み付きステップサイズ [7] を併用する．これは，近似球面化で時間相関を無視した影響を抑えるためである．j 回目の反復における，遅延フレーム数 k に当たる，分離フィルタの学習係数 μ_k を次式で定める．

$$\mu_k^{[j]} = \frac{\alpha}{j}\lambda^k + \beta. \quad (28)$$

ここで，α, β, λ は定数である．

スケーリング・パーミュテーション　ICA には出力信号の順番（パーミュテーション）と振幅（スケーリング）を決定できない特徴がある．各周波数毎に ICA を適用するため，信号の再合成時にこれらが問題となる．

スケーリングは Projection Back に基づく方法を利用する [6]．つまり，分離行列の逆行列の対角要素をそれぞれ独立成分に乗ずる．今回の分離行列は式（21）であるので，$\boldsymbol{W}_1\boldsymbol{V}_0$ の逆行列の対角要素を用いる．

パーミュテーションはスケーリングを行った独立成分の平均パワーを利用して解決する．独立成分に直接音が含まれているならば，そのパワーは一番強いはずである．従って，平均パワーが一番大きいものを直接音として選択する．

3.3　計算量の評価

計算量は，近似球面化部で $O(L^3)$，ICA 部で $O(L^2(N+M))$ に落ち着いている．特に，全体の処理時間が遅延フレーム数に対して線形演算量である．

4.　評価実験

ここでは，1) 従来の Semi-Blind 適用時のユーザ発話単語正解率, 2) 本手法適用時のユーザ発話単語正解率，の比較評価を行う．

4.1　実験設定

録音条件と評価用データ　インパルス応答は 2 種類測定した．4.2 m × 7 m および 7.55 m × 9.55 m の広さの部屋で，残響時間 (RT20) はそれぞれ，240msec, 670msec である．サンプリングレート 16kHz で録音した．マイクは Honda ASIMO の頭部に設置されている 2 本のマイクロホンを用いた．ユーザ発話に対応するスピーカは, ASIMO から見て正面から 0 °，45 °，90 °，270 °，315 °の 5 方向に設置した．ASIMO との距離は 1.5 m である．これらを表 1 にまとめる．

評価用データは録音したインパルス応答を畳み込んだ ASJ-JNAS の評価用データセット 200 文（男女各 100 文）を用いた．この 200 文を相手発話とし，ロボット発話には異なる男性話者の発話 200 文を用いた．ユーザ発話とロボット発話の長さはだいたい同じに設定し，文章の長さは 1 秒から〜10 数秒である．

音声認識と分離パラメータ　音声認識エンジンは Julius[8] を使用した．音響モデルは，クリーン音声 200 話者 (男性 100 人, 女性 100 人) 分の ASJ-JNAS 新聞記事読み上げ及び音素バランス文 計 150 文で学習したトライフォン (3 状態 8 混合の HMM) である．音声認識特徴量は，MFCC（$12+\Delta 12+\Delta Pow$）25 次元を用いた．認識に用いた話者の音声は学習データに含まれていない．これを表 2 にまとめる．

音源分離における短時間フーリエ解析の窓長は 64msec, シフト長は 24msec とした．観測スペクトル $\boldsymbol{x}$ の遅延フレーム数 N とロボット発話の遅延フレーム

Table 1 Configuration of data

インパルス応答	16kHz sampling
残響時間 (RT20)	240msec, 670 msec
距離	1.5 m
方向	0 °,45 °,90 °,315 °,270 °
STFT 解析	hanning:64 msec., shift: 24 msec.
入力	2ch (mic.) + 1ch, [-1.0 1.0] 正規化

Table 2 Settings of speech recognition

データセット	男女不特定 200 文
学習セット	クリーン男女各 100 名 (各 150 文)
音響モデル	PTM-Triphone: 3-state, HMM
言語モデル	新聞記事, 語彙サイズ 20k
音響分析	窓長 32 msec., シフト長 10 msec
特徴量	MFCC 25 dim.(12+Δ12+ΔPow)

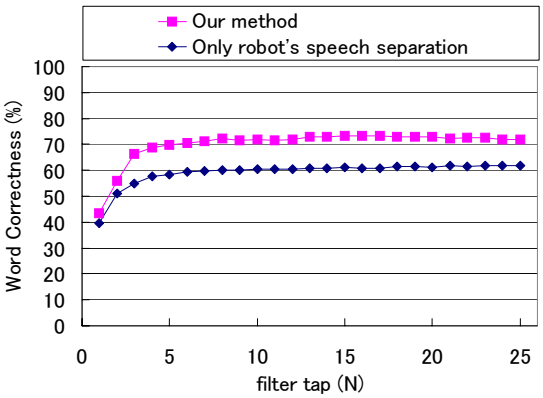

Fig.4 Result1: $RT_{20} = 240$msec

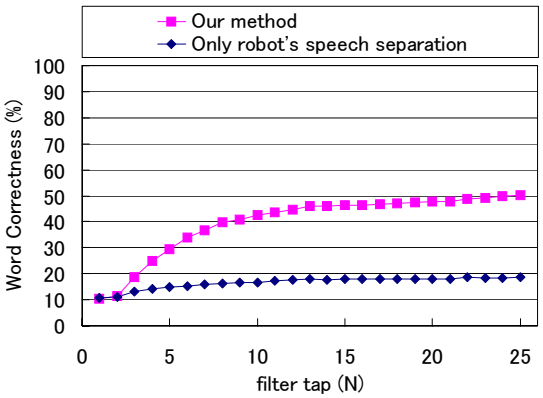

Fig.5 Result2: $RT_{20} = 670$msec

Table 3 Average word correctness (%)

	ユーザ単独発話	観測音	ロボット発話分離	本手法 ($N = 12$)
$RT_{20} = 240$msec	74.3	28.2	61.9	74.8
$RT_{20} = 670$msec	26.1	11.0	18.8	44.5

数 M は同一の値を使用している．また，学習係数関係では $\alpha = 0.5, \beta = 0.005, \lambda = 0.9$ とした．ICA の反復計算の上限は 15 回とした．今回の実験は手法の性能自体を評価するため，分離行列の推定にはデータのすべての区間を用いた．

4·2 実験結果・考察

図 4, 5 に，5 つの音源位置に対する単語正解率の平均と遅延フレーム数 N との関係を示す．表 3 に本実験の結果をまとめて示す．

表 3 における観測音は，無対策の場合の認識率であり，ユーザ単独発話はロボット発話がない場合の認識率である．後者は環境特性（残響）が畳み込まれている．クリーン音声の認識率は約 90% 程度であったため，部屋の環境による影響のみで約 10 %〜 60% 認識率が低下している．

図 4, 5 から，全体的にロボット発話の分離のみよりも，残響も同時に抑圧した方が結果が良いことがわかる．RT20=240msec 環境で平均 12.9% ，RT20=670msec 環境で平均 25.7% 単語正解率が改善している．これから残響抑圧の効果が確認できる．特に，RT20=670msec 環境下では，ロボット発話分離のみでは全く認識できていないことがわかる．

5. おわりに

本稿では，バージインを許容するロボット音声対話を目指し，従来の Semi-Blind ICA を拡張し，ユーザ発話の残響への対応を行った．特に，1) 近似球面化，2) 部分独立成分推定により，高収束性と低計算量を実現している．複数環境下での音声認識実験の結果，本手法により約 12〜25% の認識率の改善を確認した．

本手法をベースに音声認識を行うには，残響抑圧処理はまだ十分ではなく，スペクトルサブトラクションやミッシングフィーチャ理論に基づく音声認識，音響モデル適応などとの統合が必要である．特に，今回の抑圧処理で，分離フィルタに残響特性が反映されている可能性がある．今後は，分離フィルタの特性の解析，多音源への拡張や上記処理との統合を行い，リアルタイム実装を行っていく．

謝辞 本研究の一部は科研費基礎研究 (S)，グローバル COE の支援を受けた．

[1] R. Takeda *et al.*: "Barge-in-able Robot Audition Based on ICA and Missing Feature Theory under Semi-Blind Situation", *Proc. IROS2008*, (to appear).
[2] H. Saruwatari *et al.*: "Two-Stage Blind Soure Separation Based on ICA and Binary Masking for Real-Time Robot Audition System", *IROS2005*, pp.209–214.
[3] H. Sawada *et al.*: "Polar Coordinate based Nonlinear Function for Frequency-Domain Blind Source Separation", *IEICE Trans. Fundamentals*, vol.E86-A, No.3, pp.505–510, 2003.
[4] S. Araki *et al.*: "The Fundamental Limitation of Frequency Domain Blind Source Separation for Convolutive Mixtures of Speech", *IEEE Trans. On Speech and Audio Proc.*, vol. 11, no.2, pp.109–116, 2003
[5] R. Gomez *et al.*: "Fast Dereverberation for Hands-Free Speech Recognition", *HSCMA08*, pp.140–143.
[6] N. Murata *et al.*: "An Approach to Blind Source Separation Based on Temporal Structure of Speech Signals", *Neurocomputing*, 41, pp.1–24, 2001.
[7] S. Makino *et al.*: "Exponentially Weighted Stepsize NLMS Adaptive Filter Based on the Statistics of a Room Impuse Response", *IEEE Trans. on Speech And Audio Proc.*, vol.1, no.1, pp.101–108, 1993.
[8] Julius: http://julius.sourceforge.jp/