

|              |   |
|--------------|---|
| Title        | 音情景解析の概念にもとづいた音声プライバシー保護  |
| Author(s)    | 赤木, 正人; 入江, 佳洋  |
| Citation     | 電子情報通信学会論文誌 A, J97-A(4): 247-255  |
| Issue Date   | 2014-04-01  |
| Type         | Journal Article   |
| Text version | publisher   |
| URL          | <a href="http://hdl.handle.net/10119/11936">http://hdl.handle.net/10119/11936</a>   |
| Rights       | Copyright (C)2014 IEICE. 赤木正人, 入江佳洋, 電子情報通信学会論文誌 A, J97-A(4), 2014, 247-255.<br><a href="http://www.ieice.org/jpn/trans_online/">http://www.ieice.org/jpn/trans_online/</a> |
| Description  |   |

## 音情景解析の概念にもとづいた音声プライバシー保護

赤木 正人<sup>†a)</sup> 入江 佳洋<sup>††</sup>

Privacy Protection for Speech Based on Concepts of Auditory Scene Analysis

Masato AKAGI<sup>†a)</sup> and Yoshihiro IRIE<sup>††</sup>

あらまし 我々が生活している環境にはたくさんの音が存在するため、外界から目的音の情報を得るためには、目的音のみを選択し必要な情報を得ることが重要である。複数の音から目的音を選択的に聴取可能であるような聴覚上の働きは、カクテルパーティ効果と呼ばれる。この効果が生じる主な要因として、Bregman が紹介した音情景解析 (Auditory Scene Analysis: ASA) が知られている。音情景解析が上手く行えることにより、我々は、数メートル先のヒソヒソ声でも注意を向けて聴くことができる。しかし同時に、他人に聞かれない会話、例えば銀行店舗や病院外来、薬局の店頭などでの機密やプライバシーに関する会話を他人に聞かれる可能性もでてくる。本論文では、音情景解析についての知見を逆に利用することで、音声プライバシー保護を行う手法について紹介する。この手法は、会話音声と同時に音声の言語意味情報を担う音韻性を曖昧にする音を再生することによって、音情景解析がなるべく有効に働かない音環境を生成し、発話内容を不明瞭にする。

キーワード 音声プライバシー保護、音情景解析、カクテルパーティ効果

## 1. ま え が き

完全に仕切られていることが少ない銀行店舗や病院外来、薬局の店頭などでは、周りに会話が漏れてしまい、会話の機密やプライバシーが十分に保護されていない場合がある。例えば、預貯金の金額、自らの病気のことなど、他人には知られたくない情報が漏れないか心配になることもあるし、これらが待合スペースに聞こえてくることも、ある意味、迷惑である [1]。特に、薬局でのプライバシー保護の必要性については、服薬指導時の会話漏洩に対する薬剤師へのインタビュー調査等 [2]~[4] で議論されている。

一般に、会話の機密を保護するための最も良い方法は、遮音壁等を用いて発話者を隔離することであるが、銀行店舗や病院外来、薬局の店頭などのオープンな場所でのプライバシーの保護には、遮音壁を用いることは望むべくもない。現在各方面で研究されている手段

は、マスキングの現象を利用して会話を聞こえなくすることである。

従来のマスキング技術として、ピンクノイズ・BGMなどの音を重畳する方法がある。中でも、帯域制限したピンクノイズは音声をマスクするために最も有効であるとされている [5], [6]。しかし、これらは定常的なマスキング音であり、マスキング量は大きい、うるささも感じる。人の声の特性に近いマスキング音も提案されており、特に会話者本人の音声特性をもつマスキング音の効果が高いという報告もある [7] が、マスキング音を常時呈示しておく必要がある。銀行店舗や病院外来、薬局の店頭などで使用する場合には、うるささなどの不快感を覚えない音であることが望ましい。

本論文では、プライバシー保護の新しい手段として、音情景解析の概念にもとづいた、音の「知覚的融合」に関する知見を用いた手法 [8] を紹介する。この方法では、会話音声と同時に、音声の言語意味情報を担う音韻性を曖昧にする音 (以下、防聴音 (sound for hearing protection: SHP) と呼ぶ) を再生する。ある音が聞こえているときに、もう一つの音を聞かせると一方がかき消され聞こえなくなってしまう「パワーマスキング」とは異なり、会話音声と防聴音が知覚的に融合して一つの音に聞こえることにより、会話音声を

<sup>†</sup> 北陸先端科学技術大学院大学情報科学研究科, 能美市 School of Information Science, JAIST, 1-1 Asahidai, Nomi-shi, 923-1292 Japan

<sup>††</sup> グローリー株式会社新事業推進統括部, 姫路市 New Business Promotion Division, GLORY LTD., 1-3-1 Shimoteno, Himeji-shi, 670-8567 Japan

a) E-mail: akagi@jaist.ac.jp

別の音に聞き取らせ発話内容を不明瞭にする、言うなれば「情報マスクング」[9]に根差した方法である。以下、本手法の基本コンセプト、構成方法を紹介する。

## 2. 音声のプライバシー保護：消すこと vs 解らなくすること

### 2.1 遮音：音を物理的に消す

音声のプライバシー保護にとって最も良い方法は、会話音声を特定の空間から漏れ出ないようにする、あるいは、その音声を物理的に消すことである。前者を達成するためには、対話している人たちを防音壁あるいは密閉された部屋に閉じ込めれば良いが、店舗の都合上困難な場合が多い。後者を達成するためには、例えばアクティブノイズコントロール (ANC) などを使用して、対話している人たちがいる狭い範囲より外側の全ての場所で対話音声を消せば良いが、対象範囲の広さを考えれば非現実的である。

### 2.2 パワーマスクング：大きな音で知覚的に消す

声を物理的に消すのではなく、図 1 のように別の音を呈示することにより、目的の音を聞こえなくする、あるいは、音は聞こえても何を言っているのか解らなくすることができる。これらの方法でも、音声のプライバシー保護は可能となる。

地下鉄の中など大きな音が存在する環境では、普段話している声の大きさではコミュニケーションができない。これは、二つの音がほぼ同時期に存在するとき、一方の音によって最小可聴値が上昇し、他方の音を聞こえなくするために生じる。この原理を応用して、会話音声をマスクするために会話音声レベル以上の音量をもつ音を呈示すれば、会話音声を知覚的に聞こえなくすることができ、結果として音声のプライバシー保護が行える。マスクする音には、図 2 のように、ホワイトノイズ、ピンクノイズなどが用いられ、帯域制限したピンクノイズが音声をマスクするために有効であ

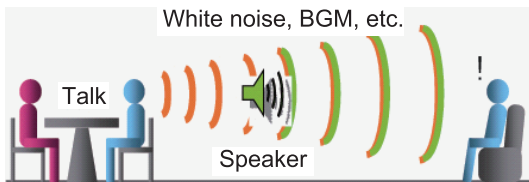


図 1 ホワイトノイズやピンクノイズなどのマスクング音、あるいは BGM による音声プライバシー保護

Fig. 1 Speech privacy protection by power masking with white noise, pink noise, BGM, etc.

ると言われている [5], [6]. また、話者本人の声を重畳したパルプノイズが有効であるという報告もある [10].

しかし、プライバシー保護を確実にするためにはマスク音のレベルを上昇させなければならず、結果として「うるさく」感じられてしまう。これを避けるためにマスク音のレベルを下降させると、マスクングできない周波数領域が現れることとなり、会話音声が聴取可能となる。すなわち、プライバシー保護のためには、「うるさい」と感じられたとしてもマスク音のレベルをある程度高く保つ必要がある。このような状況では、顧客相手の銀行、薬局などでの利用は難しい。

### 2.3 人が勝手に意味を取り間違える

一方、聴取者が勝手に意味を取り間違える、すなわち、「話者があたかも別のことを話しているように聴取者には聞こえる」状況を作ることで、音声プライバシー保護は成り立つ。

われわれが生活している環境にはたくさんの音が存在しており、聞こうとする目的の音は常に別の音に邪魔されている。このため、われわれが外界から目的音の情報を得るためには、聴覚により目的音を選択し、必要な情報を得る必要がある。二つ以上のメッセージが混在していても一つのメッセージを選択的に聴取可能であるような聴覚上の効果を「カクテルパーティ効果」と呼んでいる。カクテルパーティ効果が生じる要因としては、音の始まりの違い、音の到来方向の違い、音の高さの違い、音色の違い、また、音声の場合には言語的知識、経験、話し相手と対面している場合は唇の動きなどの視覚情報などが関係していると見られている [11].

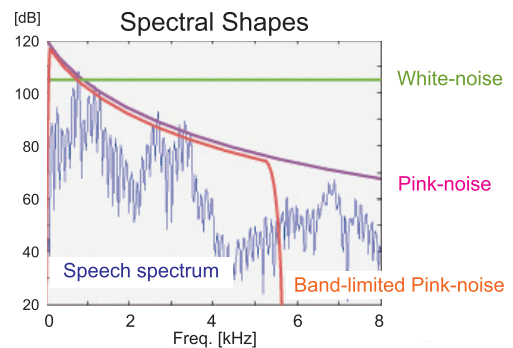


図 2 ホワイトノイズ、ピンクノイズ、帯域制限されたピンクノイズによるマスクングの概念図

Fig. 2 Conceptual graph of power masking with white noise, pink noise, and band-limited pink noise.

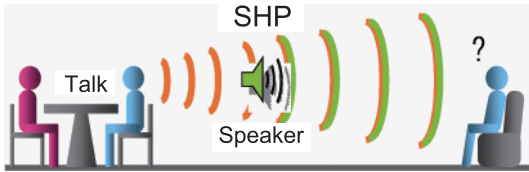


図3 会話音声と知覚的に融合する音（防聴音：SHP）による音声プライバシー保護

Fig. 3 Speech privacy protection by sound to accelerate perceptual fusion (SHP).

カクテルパーティ効果では、選択的分離知覚，すなわち，複数の音の集まりの中から意味ある音をまとまりのある音（音脈）として抜き出して聞いていると考えられている．この知覚現象を分凝（ぶんぎょう，Segregation）と呼ぶ．分凝の中でも，複数の音が一つの音脈として知覚される場合を融合という．ここで仮に，二つの音声が存在するにもかかわらず，これらが知覚的に融合し一つの音しか聞こえない場合を考えてみよう．これを「知覚的融合」と呼ぶこととする．この場合，融合した音は，個々の音声をもつ言語情報とは異なる新たな情報をもつこととなる．

本論文では，プライバシー保護の新しい手段として，音の知覚的融合に関する知見に基づいて，会話音声と同時に会話音声と知覚的に融合する音（防聴音）を再生する方法を提案する．会話音声と防聴音があたかも一つの音に聞こえる状況を作ることにより，会話音声を不明瞭にし，音声プライバシーを保護する（図3参照）．

会話音声と防聴音の融合が不十分である場合，複数の音が聞こえることとなり，会話を聴取できる可能性がある．このため，元の会話音声と防聴音を聴取者が一つの音として聞き取り，違和感を覚えないようにするための処理を加える．融合を促進させることのできる十分な要因はまだ明らかになっていないことから，カクテルパーティ効果の主要な要因である音情景解析の知見を逆に利用し，音源を分凝して聞く聴覚特性自体を欺く条件を積極的に作り出すことによって，聴取者が勝手に意味を取り間違えるように，複数の音を知覚的に融合させる．

次章以降，本手法の詳細について説明する．

### 3. 「知覚的融合」：音情景解析からの知見

#### 3.1 音情景解析

Bregman は，カクテルパーティ効果が有効に働いた

めの主要な要因を説明するために，音情景解析（Auditory Scene Analysis）という題名の書を著した [12]．この本の中で Bregman は，複数の音が存在する環境での音聴取を次のように説明した．

人間が音声あるいは音楽を聞く場合，まず音の始まりの違い，音の到来方向の違い，音の高さの違い，音色の違いなどを用いて音響的な特徴をバラバラにわけ（分離），そして，言語的知識，経験などを用いて似た者同士を寄せ集め（群化），時間方向につながり良く時間軸上のイベントをならべる（音脈の形成）という三つの処理を行った上で聞き取っている．この一連の働きを既にも書いたように分凝と呼ぶ．分凝とは，属性が異なるものを分解するだけではなく，属性の成り立ちまで立ちかえて群化を行い，意味ある情報を形成することである．

群化及び音脈形成の際に適用される規則として，規則 (1)：音の立ち上がり/立ち下りの同期性，規則 (2)：倍音構造，規則 (3)：変化の滑らかさ，そして，規則 (4)：周波数成分の振幅変化の類似性，という四つの規則が提案されている [13]．図4に規則の概要を示す．それぞれの規則について，左図は規則に従っているため音脈が一つ，右図は規則から外れた特徴が存在するため音脈が二つ聞こえる音となる．

目的の音と別の音が混在していたとしても，一般に二つの音は同時には始まらない，同じ基本周波数・倍音構造はもたない，別の音が鳴り始めるとその時点で音圧の急激な変化が現れ滑らかな変化とならない，そして，異なる音は類似の変化をしない．分離・群化・音脈形成の過程を経る段階で上記の規則を適用することにより，複数の音を分解・再構成することが可能となり，結果としてこれらの中から同じ性質をもつ目的の音が取り出しやすくなる．すなわち，一つの音のつながり＝音脈として目的の音を知覚することができる．

#### 3.2 知覚的融合の促進

ここで仮に，物理的に二つの音が存在するにもかかわらず全てが融合して一つの音脈となり，単独の音しか聞こえない場合，すなわち，知覚的融合が生じる場合を考えてみよう．知覚的融合が生じると，音声为正しく分凝できなくなり聞こえてくる音は本来もっている言語情報とは異なる情報をもつ．このため，聴取者は聞こえてくる音から意味ある言語情報を取得できず，結果として，音声プライバシー保護が可能となる．すなわち，会話音声と防聴音に対して積極的に知覚的融合が生じる環境を作ることが，音声プライバシー保護

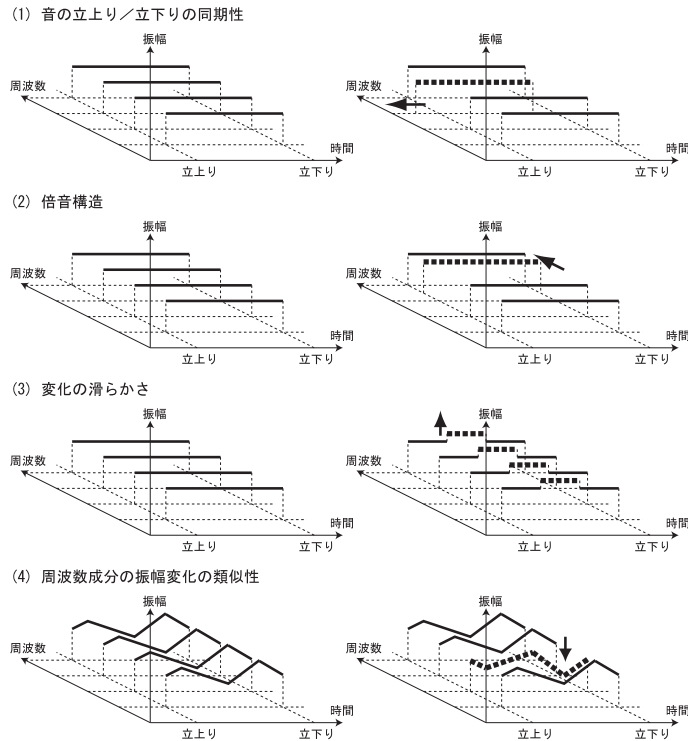


図 4 各項目の左パネルが一つの音、右パネルが二つの音に聞こえる。(1) 高調波の 1 本が他のものより早く立ち上がっており同期がとれていないため分離して聞こえる。(2) 高調波の 1 本が倍音構造からずれているため分離して聞こえる。(3) 振幅が急激に変化しているため滑らかさが保たれず、突出した時間の間の音はその前後の音とは分離して聞こえる。(4) 高調波の 1 本の振幅変調が他の高調波と異なるため分離して聞こえる [14].

Fig. 4 Bregman's four psychoacoustically heuristic regularities [14]. Sounds in left column are perceived as one sound and sounds in right column are perceived as two sounds. Rule (1): Common onset/offset, Rule (2): Harmonicity, Rule (3): Gradualness of change (smoothness), and Rule (4): Common changes occurring in the acoustic events.

につながる。

そこで、防聴音に次の条件を課す。

<条件> 防聴音は会話音声とできる限り融合しなければいけない

すなわち、会話音声と防聴音、それぞれから分離された音響特徴が、似た者同士として寄せ集められ、一緒に群化されなければいけない。このためには、図 4 に示した Bregman の四つの規則について、一つの音として聞こえるための条件をなるべく満たすように、防聴音を作成する必要がある。

本手法では、防聴音作成のために、Bregman の四つの法則のうち、規則 (1)：音の立ち上がり/立ち下りの周波数間での同期性、規則 (2)：倍音構造（共通の基本周波数変化）、及び、規則 (3)：変化の滑らかさ、

をまず考慮することとする。更に、話者本人の声を重畳したバブルノイズが有効であるという知見 [10] を考慮して、防聴音には話者の個人性 [15], [16] を残すことを考える。この操作は、高周波数帯域において、規則 (4)：周波数成分の振幅変化の類似性を満たす。

## 4. 構成例

2., 3. で示した提案手法の基本コンセプトにもとづいたシステムの構成例を紹介する。図 5 及び図 6 に、アルゴリズムの概要と各処理の出力を示す。

### 4.1 アルゴリズム

(1) マイクロホンで取り込まれた会話音声を、FFT によりフレームごとに振幅スペクトル (図 6(a)) と位相スペクトルに変換する。

- (2) 振幅スペクトルは、対数変換後に IFFT し、振幅ケプストラムを得る。
- (3) 振幅ケプストラムは、リフタリングにより、

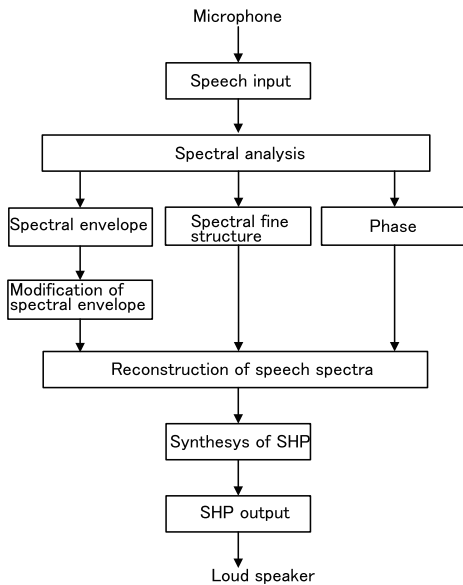


図5 アルゴリズムの概要  
Fig.5 Dataflow.

低ケフレンシ部（スペクトル包絡：音韻情報，図 6(b)）と高ケフレンシ部（スペクトル微細構造：基本周波数及び倍音，図 6(c)）に分離する。FFT により再び対数振幅スペクトルに変換した包絡と微細構造のうち、スペクトル包絡を変形する（図 6(d)）。変形は、例えば、対数スペクトルの平均を保存する軸を境に上下反転する操作で行う。これにより、規則 (1)：音の立ち上がり/立ち下がりの周波数間での同期性と規則 (3)：変化の滑らかさを保証する。

(4) スペクトル微細構造と変形されたスペクトル包絡を加えることにより、対数振幅スペクトル（図 6(e)）を得る。これにより、規則 (2)：倍音構造（共通の基本周波数変化）を保証する。変形された対数振幅スペクトルの高域部分を、元の会話音声の対数振幅スペクトルと入れ替える（図 6(f)）。これにより、話者の個人性、及び、高周波数帯域において規則 (4)：周波数成分の振幅変化の類似性を保証する。

(5) 保存しておいた位相スペクトルと変形作成した振幅スペクトルを掛け合わせ、IFFT することにより防聴音の波形を得る。

防聴音は、図 7 に示すように、A の位置で收音され

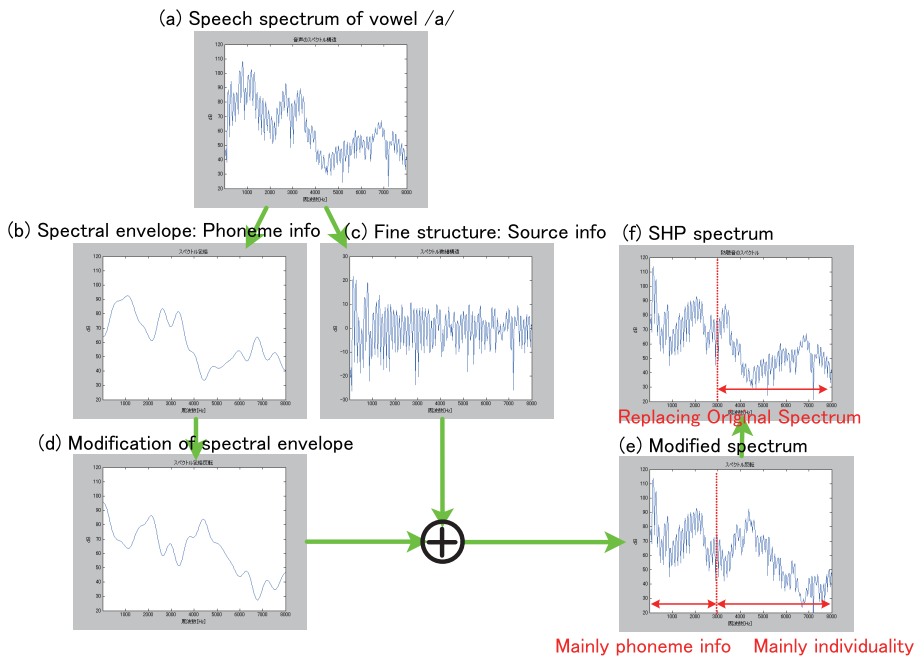


図 6 処理により出力されたスペクトルの例。各スペクトルは、横軸：0-8 kHz，縦軸：20-120 dB で描かれている。

Fig.6 Example of resultant spectra of processes. Horizontal axis: 0-8 kHz, Vertical axis: 20-120 dB.

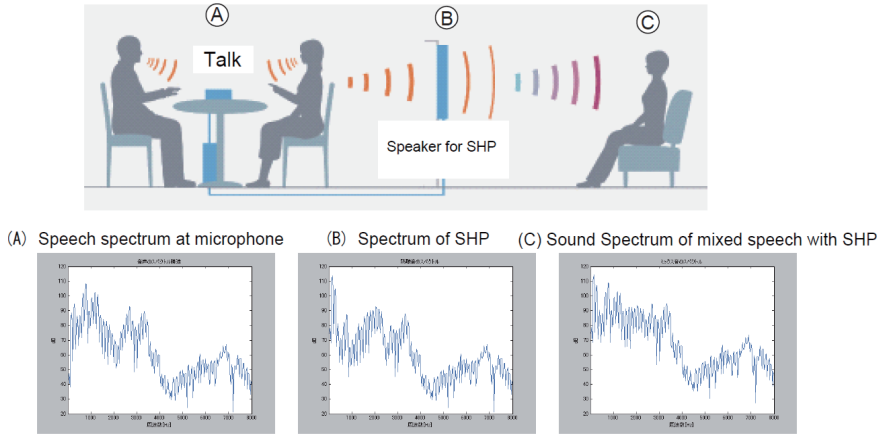


図7 音声プライバシー保護システム. 各スペクトルは、横軸：0-8 kHz、縦軸：20-120 dBで描かれている。

Fig. 7 Proposed system for speech privacy protection. Horizontal axis: 0-8 kHz, Vertical axis: 20-120 dB.

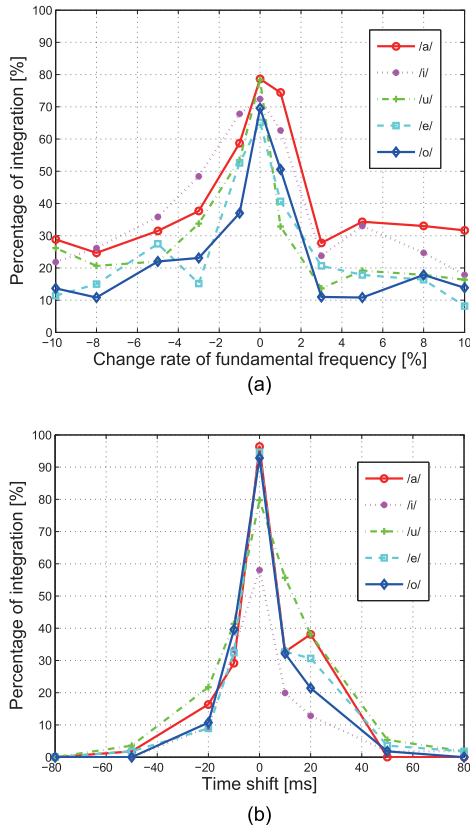


図8 防聴音の基本周波数をシフトした場合 (a) と、防聴音呈示の開始時間をシフトした場合 (b) の知覚的融合率. 会話音声として単独母音を用いている [17].

Fig. 8 Percentage of perceptual fusion as (a) a function of F0-difference of the two sounds and (b) a function of time-shift of the two sounds [17].

た音から上記のアルゴリズムにより作られ、会話者と聴取者の間 (B の位置) に置かれたスピーカから呈示される。これにより、聴取者の位置 (C) では、会話音声と防聴音が一緒に到来し、知覚的融合が図られることとなる。

#### 4.2 知覚的融合が成り立つ条件

Minowa らの研究 [17] によれば、このアルゴリズムによって作られた防聴音が知覚的融合を生じるためには、ヘッドホン受聴の場合においては次の条件を満たす必要がある。

図 8 は、防聴音の基本周波数を上下にシフトした場合 (a) と、防聴音呈示の開始時間を前後にシフトした場合 (b) の知覚的融合が生じる割合である。(a) の場合には規則 (2) : 倍音構造 (共通の基本周波数変化) が保証されず、(b) の場合には規則 (1) : 音の立ち上がり/立ち下りの周波数間での同期性が保証されない。図を見ると、基本周波数が 1% ずれるだけで、また、立ち上がりが 10 ms ずれるだけで、知覚的融合が成り立たなくなっている。このため、基本周波数は誤差が 1% 以内、防聴音を呈示するまでの遅延時間は 10 ms 以内を達成することが望ましい。なお、実環境では残響が存在するため、遅延時間がもう少し長くても知覚的融合は生じる。

### 5. 性能評価

#### 5.1 単語理解度

本手法の有効性を評価するために、先行研究にて音声をマスクするために最も有効であるとされている定

常的なピンクノイズと本研究で提案する防聴音を用いた場合との比較実験結果を示す。実験では、防聴音を呈示するスピーカを被験者から 1.5 m の位置へ、会話音声呈示するスピーカを被験者から 3 m の位置へ設置し、単語理解度試験を行う。音声データとして用いたのは、NTT 親密度別単語理解度試験用音声データベース [18] の単語音声である。単語は、男性話者 mya の 5 種類の親密度別に同じ数ずつランダムに選んだ。呈示音圧は、実験前にそれぞれのスピーカからピンクノイズを呈示し、被験者の耳の位置で騒音計 (A 特性) にて 50 dB になるように調整した上で、会話音声を 50 dB、ピンクノイズも 50 dB、そして、防聴音を 44~56 dB の範囲で呈示した。被験者には、再生時は必ずスピーカの方を見るように指示し、再生された音声は何と言っているのかを聞こえたとおりに紙に書き込ませた。実験環境を図 9 に示す。被験者は聴力正常な大学院生 3 名 (A~C) である。

結果を図 10 に示す。図 10 中の横軸において、44, 50, 56 dB が提案手法で作成した防聴音の場合、Pink が 50 dB のピンクノイズを呈示した場合である。縦軸は単語の完全正答率である。値が小さい方がより

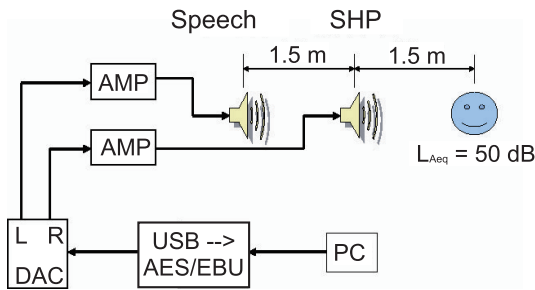


図 9 実験概要  
Fig.9 Experimental setup.

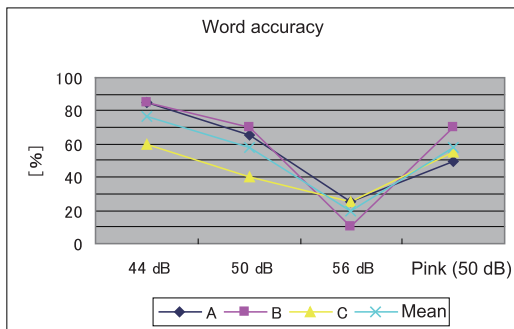


図 10 単語正答率  
Fig. 10 Correct answer rate of the words.

プライバシー保護度合いが高いことを示す。実験の結果、ピンクノイズを呈示した場合と本研究の手法 (50 dB の場合) を比較すると、単語正答率においてはほぼ互角の結果となっている。ピンクノイズは現有で最も良いと言われている手法であるが、本手法はピンクノイズと同程度の保護度合をもつことがわかる。また、防聴音を用いた場合には、音圧を少し大きくする (56 dB の場合) だけで正答率が低くなっている。被験者の内観報告によれば、防聴音は耳障りではなく、56 dB に大きくした場合でも違和感はない、一方、ピンクノイズはうるささが大きく音声が完全に分離して聞こえ、違和感も大きく感じる、ことが明らかとなっている。

### 5.2 不快感

防聴音を用いた場合の不快感に関して、岸らが次の研究を行っている [19]。この論文では、防聴音のみならず、音声逆再生音、ピンクノイズをマスキング音として用いた場合の不快感の許容レベルを聴取実験により明らかにしている。不快感の許容レベルとは、マスキングノイズの不快感についての許容上限となる提示音圧レベルである。参考文献 [19] 中の結果の抜粋を図 11 に示す。図中の F が防聴音、N がホワイトノイズに -5 dB/oct. のイコライジングをかけた音である。

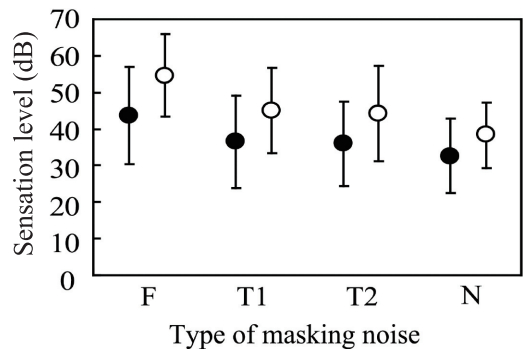


図 11 マスキングノイズの不快感についての許容感覚レベル。参考文献 [19] から抜粋。この図中の F が提案手法、N がホワイトノイズに -5 dB/oct. のイコライジングをかけた音、T1/T2 は、音声素片をランダムに並べ替えた音。○は待合スペースで何もせず待っている状況を想定、●はオフィスで働いている状況を想定。許容感覚レベルは F 型が最も高く N 型が最も低い。

Fig. 11 Averaged acceptable levels of unpleasantness for each type of masking noise. F: SHP, N: pink noise, and T1/T2: time-randomized/time-reversed reproduction speech. The black and white circles indicate long-time and short-time situations, respectively. For the acceptable level, type F is the highest and type N is the lowest.



不快感の許容レベルは F 型が最も高く N 型が最も低い。すなわち、防聴音は、音圧が高くても最も不快感が小さい手法であることがわかる。

## 6. む す び

本論文では、音情景解析についての知見を逆に利用した新たな音声プライバシー保護法を紹介した。この方法は、会話音声と同時に音声の音韻性を曖昧にする防聴音を再生することによって、音情景解析がなるべく有効に働かない音環境を生成し、発話内容を不明瞭にする。本手法を利用した製品 [20] が既に発売され、銀行、病院、調剤薬局などに導入が始まっている。

本研究のプライバシー保護手段の実用化は、防音個室以外の半個室等でもプライベートな会話を保護できるようになるばかりでなく、心理学的興味として音源分離知覚による心理的要因も発見できると考えられる。

## 文 献

- [1] 佐藤 洋, 清水 寧, “スピーチプライバシー研究の歴史と近年の動向,” 日本音響学会誌, vol.64, no.8, pp.475–480, 2008.
- [2] 小山由美, 星 和磨, 羽生敏樹, “調剤薬局におけるスピーチプライバシー保護 —その 1 服薬指導時の会話漏洩の実態調査—,” 日本建築学会大会, 40133, pp.293–296, 2008.
- [3] M. Fujiwara, M. Hata, T. Yamakawa, and Y. Shimizu, “Experimental study of speech privacy with a sound-masking system in pharmacies,” Proc. inter-noise2011, Osaka, Japan, 2011.
- [4] T. Yamakawa, M. Hata, M. Fujiwara, and Y. Shimizu, “The solution of speech privacy secured to a waiting space for a confidential conversation at a pharmacy counter,” Proc. inter-noise2011, Osaka, Japan, 2011.
- [5] 佐伯徹郎, 藤井健生, 山口静馬, 老松健成, “音声をマスクするための無意味定常雑音の選定,” 信学論 (A), vol.J86-A, no.2, pp.187–191, Feb. 2003.
- [6] 佐伯徹郎, 山口静馬, 為末隆弘, “マスキングノイズによるスピーチプライバシー保護に関する一考察,” 日本音響学会誌, vol.61, no.10, pp.571–575, 2005.
- [7] A. Ito, et al., “Oral information masking considering room environmental condition, Part 1: Synthesis of maskers and examination on their masking efficiency,” Proc. inter-noise2007, Istanbul, Turkey, 2007.
- [8] 赤木正人, 入江佳洋, “音情景理解を応用した音声プライバシー保護,” 信学技報, EMM2011-59, Dec. 2011.
- [9] N.I. Durlach, et al., “Note on information masking,” J. Acoust. Soc. Am, vol.113, no.6, pp.2984–2987, 2003.
- [10] T. Komiyama and K. Kondo, “An effective speech privacy system using speaker-dependent babble noise as maskers,” Proc. of inter-noise2011, Osaka, Japan, 2011.
- [11] 赤木正人, “カクテルパーティ効果とそのモデル化,” 信学誌, vol.78, no.5, pp.450–453, May 1995.
- [12] A.S. Bregman, Auditory scene analysis: the perceptual organization of sound, MIT Press, Cambridge, MA, 1990.
- [13] A.S. Bregman, “Auditory scene analysis: hearing in complex environments,” in Thinking in sound: The cognitive psychology of human audition, ed. S. McAdams and E. Bigand, Chapter 2, Oxford Science Pub., pp.10–36, 1993.
- [14] 赤木正人他 (共著), 音響学入門, 日本音響学会 (編), 2.4 節, pp.42–43, コロナ社, 2011.
- [15] T. Kitamura and M. Akagi, “Speaker individualities in speech spectral envelopes,” J. Acoust. Soc. Jpn. (E), vol.16, no.5, pp.283–289, 1995.
- [16] 北村達也, 赤木正人, “単母音の話者識別に寄与するスペクトル包絡成分,” 日本音響学会, vol.53, no.3, pp.185–191, 1997.
- [17] A. Minowa, M. Unoki, and M. Akagi, “A study on physical conditions for auditory segregation/integration of speech signals based on auditory scene analysis,” Proc. NCSF2007, pp.313–316, Honolulu, USA, 2007.
- [18] S. Sakamoto et al., “Complementary relationship between familiarity and SNR in word intelligibility test,” Acoust. Sci. Tech., vol.25, no.4, pp.290–292, 2004.
- [19] 岸 征宏, 森本政之, 佐藤逸人, 黒田直樹, 入江佳洋, “サウンドマスキングシステムのマスキングノイズの不快感,” 平成 23 年秋季音響学会講演論文, 1-5-18, 2011.
- [20] グローリー, “会話保護システム QG-11,” <http://www.glory.co.jp/ir/pdf/k.110823.pdf>, <http://www.glory.co.jp/product/security/gvips.html> (平成 25 年 8 月 23 日受付)

## 赤木 正人 (正員)



昭 54 名工大・工・電子卒。昭 59 東工大大学院博士課程情報工学専攻了。工博。同年電電公社 (現 NTT) 研究所入社。以来, ATR 視聴覚機構研究所, NTT 基礎研究所を経て, 現在, 北陸先端科学技術大学院大学情報科学研究科教授。この間, 米国 MIT, オランダアムステルダム大学, 英国ケンブリッジ大学滞在研究員。音声信号処理, 聴覚機構のモデル化の研究に従事。昭 62 電子情報通信学会論文賞, 平 10, 17, 22, 及び 23 日本音響学会佐藤論文賞, 平 21 信号処理学会 Best Paper Award を受賞。日本音響学会, 電子情報通信学会, 信号処理学会, IEEE, ASA, ISCA 各会員。



入江 佳洋

昭 62 姫路工大・工・金属材料卒，平元同大大学院工学研究科修士課程修了．同年グローリー工業（株）会社（現グローリー（株））入社．以来，文字認識，音声処理システムの技術開発，製品開発に従事，日本音響学会会員．