

スピーチプライバシー保護機能をもたせた電子ポスター - ポスラサウンドパネル[®] for スピーチプライバシー -

茂出木 敏雄[†]

橋本 聡[‡]

片山 貴信[‡]

庄司 藤男[‡]

スピーチプライバシーを保護する方法として白色雑音のような音響信号を付加する方法が提案されているが、返って不快感を増強させる問題が指摘されており普及していない。筆者らは、音楽・自然音など任意の鑑賞用の音響信号に対して、白色雑音と同程度に会話音声に対して聴覚マスキングが効果的に働くように特殊なフィルタ加工を施す手法について提案する。大手調剤薬局の待合室に、既開発のフラットパネル型スピーカを設置して、本提案手法の効果を検証した結果、フィルタ加工による音質歪みについては殆ど認識されず、マスキング音が所定の音圧で届くエリアには薬剤師と顧客との会話音が聞き取りにくくなるという効果を確認することができた。本実演では会話音を模擬した小型スピーカを用いて、スピーチプライバシー保護効果を提示する。

An Electronic Poster Providing Speech Privacy Protection Functions - Posla Sound Panel[®] for Speech Privacy -

TOSHIO MODEGI[†]

SATOSHI HASHIMOTO[‡]

TAKANOBU KATAYAMA[‡]

FUJIO SHOJI[‡]

In order to protect speech privacy, a method of generating white noises on purpose has been proposed. However, this method will lead us to uncomfortable environments and has been not widely used. We propose a novel filtering method for generally used music or natural sound signals in order to emphasizing their auditory masking functions against speech sounds like white noises. We have installed our previously developed flat-panel loudspeaker in a waiting room of a large-scale prescription pharmacy, and experimented our proposal there. As a result, we could not recognize any play-back sound quality distortion in masker sounds, whereas we could find difficulty to listen to conversation sounds between couples of a pharmacist and a client as far as heard masker sounds had proper level. This demonstration presents speech privacy protection effects using small loud speakers simulating human speech voices.

1. はじめに

各種面談・打ち合わせ場面で交わされる会話音声には個人情報や企業機密情報などが含まれていることが少なくない。個人情報保護法の制定に伴い、個人を特定するデータに対するセキュリティ意識は高まってきたが、口頭から漏れ聞こえてくる個人情報についても注目されるようになってきた。特に、医療機関ではインフォームドコンセント、調剤薬局では薬剤師による服薬指導が義務付けられており⁸⁾、国土の狭い日本では、遮音パネルを敷設し面談環境を整備するなど、ハード的にスピーチプライバシーを保護することは必ずしも容易ではない。

そこで、ヒトの聴覚マスキング特性を活用し、白色雑音に近いエアコンノイズのような音響信号を対象の室内に流すエネルギーマスキング（エネルギーの大き

い音波で会話音を聞こえなくする手法）という手法が提案された¹⁾²⁾³⁾。しかし、音量が大きい白色雑音が流れている環境では返って不快感が増し、会話時の相手の声が聞き取りにくくなるという問題が指摘され、なかなか普及に至っていない。文献 4)のように流す音響信号の煩さを低減したり、BGMを合成して快音化する方法¹⁾も提案されているが、抜本的な解決には至っていない。

これに対して、ヒトの音声信号成分を重ねて会話音声の内容を秘匿化するインフォメーションマスキングという方法が提案されている⁵⁾⁶⁾⁷⁾。前述のエネルギーマスキング手法に比べ、弱いマスキング信号音圧で会話音を聞き取りにくくする効果があるが、マスキング音として使用する音声信号成分は、会話音声の当事者の声にできるだけ近い特性をもつサンプルを用いないと秘匿化効果が十分にでないという問題がある⁶⁾。文献 7)では、会話音声をリアルタイムにマイクロフォンで拾い、音韻のフォルマント特性を変化させるように、特殊な雑音成分をリアルタイムに生成して重畳する方法を提案している。本方法ではシステムが大掛かりに

[†] 大日本印刷（株）情報コミュニケーション研究開発センター
Media Technology Research Center, Dai Nippon Printing Co., Ltd
[‡] （株）DNP東北 P&I 開発センター / 開発営業センター
P&I / Business Development Center, DNP Tohoku Co., Ltd.

なり、会話音声を取得して別の場所に加工して流すという行為自体がプライバシー侵害につながる可能性があるという懸念もあり、実用化が難しい。

そこで筆者らは、鑑賞用の音楽や自然音に対して白色雑音と同等にエネルギーマスキング効果を発揮できるようにフィルタ加工を行い、室内に快適なBGMを流すだけでスピーチプライバシーを保護できる新規な手法を提案する。また、文献 9)10)の技術を用いて開発したフラットパネル型スピーカでフィルタ加工したBGMを流すと、本来の会話音声を邪魔しない音量で、室内の広範な領域に聴覚マスキング効果を働かせることができることを調剤薬局で実証したので¹¹⁾、本稿ではその結果を報告する。

2. 提案するスピーチプライバシー保護手法

文献 1)のような不快な白色雑音を使用しなくても、一般的な鑑賞用の音楽や自然音などのBGMをある程度の音量を上げて再生すると、会話音が聞こえにくくなることが知られている。しかし、聴覚マスキングの原理上、図1右上に示されるように、BGMの周波数特性により、いくら音量を上げて音声に対するマスキングが働かない場合が生じる。そこで、同図右下のように秘匿化対象音声に豊富に含んでいて、BGMに欠如している周波数成分の近傍成分に対して増強するようなフィルタ加工処理を施すことを提案する。この際、電子透かし埋め込みと同様に、再生されるBGMの品質に対してできる限り影響を与えないように、フィルタ加工を行うことがポイントになる。

この方法でフィルタ加工されたBGM音楽信号をそのまま会話音に重ねても、ある程度音量を上げないと秘匿化効果が発揮されない。そこで、筆者らは図2に示されるような秘匿化対象音声の平均スペクトルとBGMのスペクトルに共通するスペクトル成分を先にフィルタ加工したBGM信号より削除する方法を併せて提案する。このように減弱されたBGM信号を秘匿化音声に重ねると、秘匿化対象音声信号の一部の成分を奪い取ってBGM音楽信号の一部として機能するようになる。即ち、比較的弱い音量でもマスキング効果を発揮できるようになる。ただし、音声・音楽に共通するスペクトル成分を厳密に算出することは、秘匿化対象音声を特定できない以上、困難である。そこで、図2の上段に示されているように、秘匿化対象音声の平均スペクトルに適当なスケールを施して減弱した平均スペクトルを共通スペクトルとして使用するようになった。

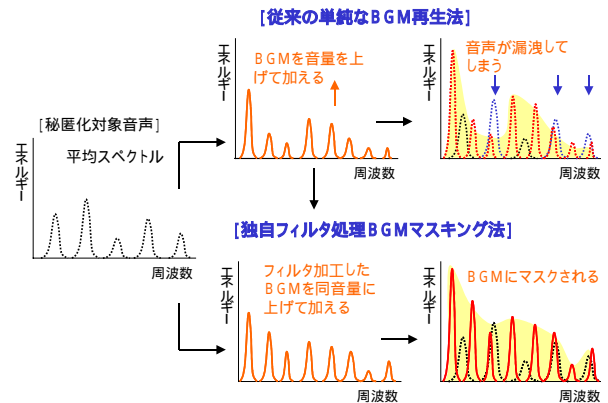


図1 BGMを用いたスピーチプライバシー保護手法の基本概念

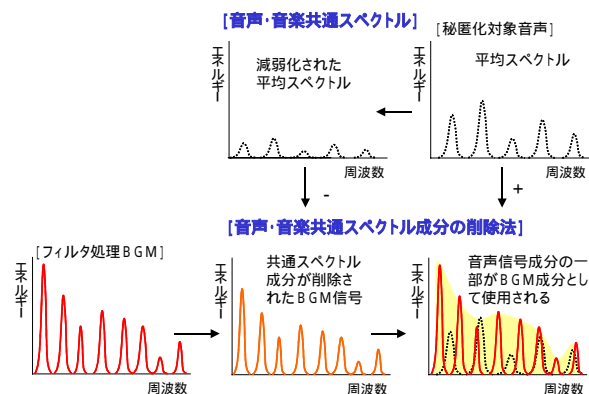


図2 秘匿化対象音声とBGM音楽に共通するスペクトル成分の削除手法の概念

スピーチプライバシー保護の目的は会話音を聞こえなくすることではなく、会話音を理解できないようにすることにある。従って、会話音の全ての周波数成分をマスキングする必要は無く、母音を構成する一部のフォルマント成分、子音を構成する主要な周波数成分を部分的にマスキングできるレベルに、フィルタ加工強度、共通スペクトル成分の削除割合及びBGM再生音量を調節するようにする。また、スピーチプライバシー保護を導入することにより、主となる会話音を邪魔することがないように配慮する必要がある。これについては、会話を行っているエリアにはBGMの再生音圧が低くなるよう、スピーカにある程度の指向性をもたせることで解決できる。筆者らは、既開発のフラットパネル型スピーカを採用することで本問題を解決した。

スピーチプライバシー保護を実現する上で技術的に障害となるのは、カクテルパーティ効果で知られるヒ

トの音脈に対する補間能力である。ヒトは特定の音源・会話音に注目して聴取すると、聴覚マスキングにより聞き取れない信号成分を前後の音脈を参考にして自動的に推定補間してしまう能力がある。この場合は、会話音より BGM 音源の方に注意をそらせるように誘導する必要がある。対策としては、BGM の音量を上げるか、聴取しているヒトが興味をもちやすい音楽コンテンツを選択する手段が挙げられる。

次節では、会話音声に対するマスキングが効果的に働くように再生対象の BGM 信号をフィルタ加工する手法、音声と音楽の共通スペクトル成分を削除する手法について詳細を述べる。

3. 音響信号に対するスピーチプライバシー保護機能の埋め込み方法

地点 A において交わされる会話音声信号に対して、所定の距離だけ離れた地点 B において聴取できないように、地点 C より流す BGM 音楽信号を生成する方法として、図 3 に示されるような信号処理手法を提案する。まず、日本人の平均的な男声・女声を含むモノラル音声信号のサンプルをテンプレートとして準備し所定のフレーム間隔で周波数解析を行い、全フレームに対する平均値スペクトルと最大値スペクトルを算出する。前者については、図 3 最上段に示されるように所定のスケール値をかけて減弱化し、これを共通スペクトルとする。次に、再生対象とする BGM 音楽信号に対しても同一のフレーム間隔で周波数解析を行い、各フレームごとに得られる瞬時スペクトルに対して前後の近傍フレームの瞬時スペクトルを用いて平滑化する。各フレームごとに平滑化された瞬時スペクトルの値で先に算出したテンプレート音声の最大値スペクトルの値を周波数別に除算することにより、図 3 中央に示されるフィルタ関数を算出する。続いて、図 4 に示されるように、BGM 音楽信号に対して前記と同一のフレーム間隔でフーリエ変換を行い、周波数次元で各フレームごとに算出されたフィルタ関数を乗算するとともに、減弱化した平均音声スペクトルの減算し、フーリエ逆変換を行って時間次元のフィルタ加工されたマスキング音楽信号を取得する。この信号を会話音を聞かせたくないエリアに向けてスピーカで再生するようにする。

4. 提案手法の評価結果

提案手法の評価をするにあたり、調剤薬局の BGM に相応しい明るく静かなクラシック楽曲を音楽素材 10 曲（各平均 3 分程度にカット）を用い、著作権の関

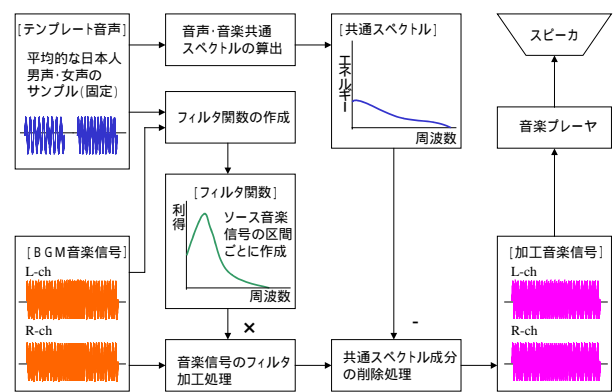


図3 提案するスピーチプライバシー保護機能を BGM 音楽信号に埋め込む信号処理方法

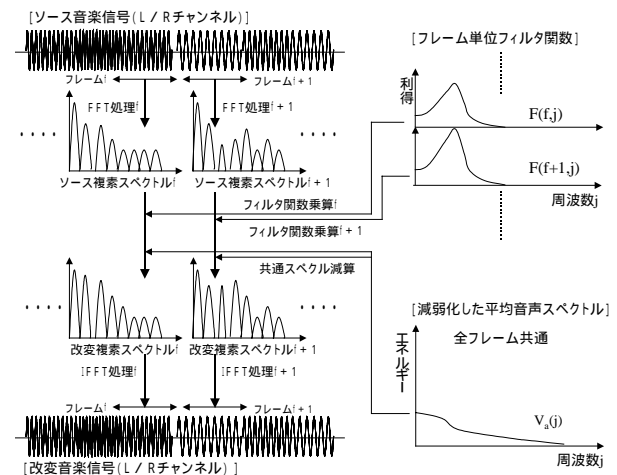


図4 提案する BGM 音楽信号に対する信号処理方法の詳細

係で、全て MIDI 打ち込み素材を用いるようにした。これに対して、既開発の MIDI 符号化ツール「オート符」[13]を用いて客観評価した結果について以下述べる。調剤薬局での実験結果については文献 11)を参照。

図 5 は本提案手法の評価に用いたサンプルの一例で、サンサンス「白鳥」の冒頭 30 秒で、同図(a)は MIDI 打ち込みデータで、これに対し YAMAHA MU500 を用いてオーディオ変換した原音 WAV データに対して、本稿で提案した加工を行った WAV データに対して、同様に MIDI 符号化ツールを用いて MIDI データに変換した結果が同図(b)である。図中の着色された小さな矩形は音符を示し、横軸は時間で、横幅はノートオンからノートオフ区間を示す。縦軸は音高（ノートナンバー）で縦方向の幅でベロシティも示している。従って、矩形の面積は音成分のエネルギーを反映している。図示は省略するが、原音 WAV

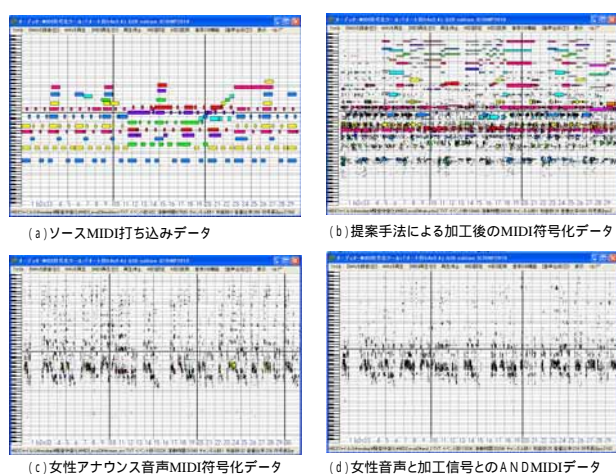


図5 MIDI 楽曲データに対するマスキング効果強調信号処理例
に対する MIDI 符号による評価結果

データに対して、同様に MIDI データに変換した結果と差分をとると、加工処理による信号劣化成分を可視化することができる。この程度の変化であれば、原音との品質上の相違は慎重に聞き比べない限り知覚されないことを確認できた。

次に、図 5 (c) は女性アナウンス音声に対して同様に MIDI データに変換した結果である。これに対して、先の図 5 (b) と MIDI データどうしで、同様に既開発の MIDI 符号化ツール[13]を用いて、ノートナンバーに ± 1 半音の許容値をもたせて AND 演算を行った結果が図 5 (d) で、マスキング対象となる音声信号成分を反映している。図示は省略するが、原音 WAV データから変換された MIDI データと同図 5 (c) とを MIDI データどうしで、同様に MIDI 符号化ツールを用いて、ノートナンバーに ± 1 半音の許容値をもたせて AND 演算を行った結果と比較すると、本稿で提案した加工を施すことによりマスキング対象成分が客観的にも増加していることを確認できた。

5. おわりに

本稿では、与えられた任意のスペクトル特性をもつ楽曲に対して、スピーチプライバシーを保護する目的でマスキング効果を強調する信号処理手法を提案し、調剤薬局に持ち込んで提案手法の有効性を確認することができた。今後は、調剤薬局に限らずオフィス一般のスピーチプライバシー保護の目的に本提案手法を適用するにあたり、どのような楽曲をどの程度の音量で提示するべきか、使用するスピーカのパネルの大きさや設置個数などについて最適に見積もれるシステムを

構築する予定である。また、各種不快な機械騒音を快音化するために、より音圧の低い BGM でも同等なマスキング効果が得られるような信号処理手法について検討を進める予定である。

参 考 文 献

- 1) コクヨ(株)「サウンドマスキングパーティション」, <http://www.kokuyo.co.jp/press/2008/12/854.html>. N S K (株)「会議室の環境改善, サウンドマスキングシステム」, <http://www.ipros.jp/product/detail/316231013/>.
- 2) (株)セイ「サウンドマスキングシステム Calma」, <http://www.say-international.com/index.htm>.
- 3) (株)オカムラ「室内音響調整システム, Sound Conditioning System」, http://www.okamura.co.jp/company/topics/kenzai/2009/sound_conditioning_system.php.
- 4) ヤマハ(株)「スピーチプライバシー・ソリューション技術」, <http://www.yamaha.co.jp/news/2010/10012701.html>.
- 5) A. Ito, A. Miki, Y. Shimizu, K. Ueno, HJ. Lee, and S. Sakamoto, "Oral information masking considering room environmental condition, Part 1: Synthesis of Maskers and examination on their masking efficiency," Proc. of inter-noise 2007 (2007).
- 6) グローリー(株) 北陸先端科学技術大学院大学「会話プライバシー保護技術, GVIPS (Glory Voice Intelligent Protection System)」, <http://www.jaist.ac.jp/~akagi/Release.pdf>.
- 7) 小山由美, 星和磨, 羽入敏樹:「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」日本建築学会大会学術講演梗概集 (September 2008) .
- 8) 京セラケミカル「アラミドハニカムフラットパネルスピーカー」, <http://www.kyocera-chemi.jp/news/2007/20071107.html>.
- 9) New Transducers Limited "NXT surface speaker", <http://www.nxtsound.com/index.html>.
- 10) D N P 「ボスラサウンドパネル for スピーチプライバシー」, http://www.dnp.co.jp/news/1221689_2482.html.
- 11) E. Zwicker and E. Terhardt, "Analytical expressions for critical-band rate and critical bandwidth as a function of frequency", Journal of Acoustical Society of America, Vol.68, no.5, pp.1523-1525, November 1980.
- 12) Toshio Modegi: "Evaluation Method for Quality Losses Generated by Miscellaneous Audio Signal Processings Using MIDI Encoder Tool 'Auto-F'," Proc. of IEEE TENCON2010, pp.2066-2071, Fukuoka Japan, November 2010.