

Title	音声伝送指標を基準としたスピーチプライバシー保護の研究
Author(s)	柏原, 佑太
Citation	
Issue Date	2017-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/14139
Rights	
Description	Supervisor: 鶴木 祐史, 情報科学研究科, 修士

音声伝送指標を基準としたスピーチプライバシー保護の研究

柏原 佑太 (1510010)

北陸先端科学技術大学院大学 情報科学研究科

2016年2月10日

キーワード: 音声伝送指標 (STI), 変調伝達関数 (MTF), 室内インパルス応答 (RIR), スピーチプライバシー保護, 音声の聴き取り.

音声伝送指標 (STI) は, 室の音声伝送品質の評価に利用される客観評価尺度であり, 主観評価尺度である聴き取りにくさと高い相関を持つ指標である. したがって, STI は通常, 部屋の聴取者による音声の明瞭度を予測するために使用される. STI は, 音声伝送性能を表すもので, bad から excellent までの 5 段階評価がされている. STI は, 変調伝達関数 (MTF) の概念に基づき, 室内インパルス応答 (RIR) から算出される. したがって, STI および RIR を使用して, 室内における音声の聴き取りを予測することができる. 本論文では, この STI を基準として, STI を操作することで, 漏えい音声を聴き取りにくくすることによって, スピーチプライバシー保護を行うことを目的とする.

実環境では, 室の伝送経路を直接操作して STI を制御することは非常に難しい. しかし, 直接音に対して, 後部残響に対応する音刺激を遅延和として加算呈示することで見かけ上, RIR を畳み込むことと等価な音響処理が可能である. さらにはこの方法によって STI を制御することで音声の聴き取りを制御することも可能かもしれない. 本論文では, この仕組みに基づき, 直接音と後部残響で構成した RIR を用いて, そのパラメータを操作することで STI を制御することで, 音声の聴き取りを制御し, スピーチプライバシーが保護されたかどうか検討する.

ある室における音声の聴き取りの状況を予測するためには, その室の STI, つまり, その室の RIR を事前に知っておかなければならない. RIR は, 一般的に, 直接音, 初期反射, 後部残響の三つの成分で構成される. これらの成分のうち後部残響がもっとも音声の聴き取りに影響を与えることは知られている. 直接音と後部残響からなる統計的 RIR モデルを利用して, そのモデルパラメータの変化に伴う STI の変化を系統的に予測できることを示した. これらの結果に基づけば, 後部残響を構成するモデルパラメータを操作することで STI を制御できるだけでなく, さらには音声の聴き取りも制御できるし, スピーチプライバシー保護を行うことができる. そのために, まず, 拡張型 RIR を後部残響とした後部残響モデルを定義した. このモデルから, MTF を算出した. すると, 低域通過特性が見られた. この MTF から STI を求めた結果, T_h と T_t によって, STI が 0.2 から

1.0まで制御できることがわかった。拡張型 RIR を後部残響とした後部残響モデルは、立ち上がりを制御するパラメータ T_h や立ち下がり制御するパラメータ T_t , 振幅項 a を容易に制御する事ができる。この方法は提案法と呼ばれる。

提案法の性能評価を行うために、STI を変化させたときに単語了解度試験、聞き取りにくさ、わずらわしさがどのように変化するかを検討した。実験刺激には、親密度別単語了解度試験用データベース (FW07) を用いた。これらの音声は、4 モーラで構成されている。音声には、単語親密度が 1.0 から 7.0 の男性発話者 (mya) の音声を用いた。提案法を用いて、STI が 0.875, 0.675, 0.525, 0.375, and 0.230 となるような RIR を作成した。実験刺激は FW07 と作成した RIR を畳み込んで作られた。刺激の総数は、STI 5 条件 × 単語親密度 4 条件 × 20 単語 × 実験数 3 の 1200 単語であった。サンプリング周波数は、48 kHz だった。被験者は、単語了解度試験と聞き取りにくさにおいて男性 6 名、わずらわしさにおいて男性 3 名とした。被験者全員は、正常聴力を持ち、日本語を母語としている。単語了解度試験では、了解度は、被験者のこたえた単語の正解率とした。聞き取りにくさの聴取実験では、実験参加者に、聴取音声の印象を、表??のように「聞き取りにくくはない」、「やや聞き取りにくい」、「かなり聞き取りにくい」、「非常に聞き取りにくい」の 4 つの評価から 1 つ強制選択させた。「聞き取りにくくはない」の数から聞き取りにくさを算出した。その集計の結果、次の式のように、「聞き取りにくくはない」以外の選択数を刺激の総数で除算した値を集計した。聞き取りにくさの聴取実験では、実験参加者に、聴取音声の印象を、表??のように「わずらわしくはない」、「ややわずらわしい」、「かなりわずらわしい」、「非常にわずらわしい」の 4 つの評価から 1 つ強制選択させた。「わずらわしくはない」の数から聞き取りにくさを算出した。その結果、単語了解度試験においては、単語親密度に注意すれば、STI の操作によって、単語了解度を操作できることがわかった。また、聞き取りにくさとわずらわしさは、単語親密度にかかわらず STI の操作によって、制御されることがわかった。これにより、STI を操作することによって、主観評価尺度である、単語了解度、聞き取りにくさ、わずらわしさを制御することができたといえる。この結果から、スピーチプライバシー保護ができることがわかった。

最後に、提案法が他手法よりも優れているかどうかを検討するために比較実験を行った。STI が 0.23 のときに、提案法と残響音声とピンク雑音について比較するため、単語了解度、聞き取りにくさ、わずらわしさについて聴取実験を行った。刺激の総数は、3 手法 × 単語親密度 2 条件 × 20 単語 × 実験数 3 の 360 単語であった。被験者は、23 歳から 31 歳の男性 8 名とした。その結果、提案法は、他の 2 手法とほとんど同等の性能を発揮することができた。さらに、提案法は、他の 2 手法よりも信号体雑音比 (SNR) が低くなったため、効率よく STI が制御できたといえる。これにより、他の 2 つのスピーチプライバシー保護の技術よりもスピーチプライバシー保護が容易にできることがわかった。

本論文では、次の 3 点を明らかにした。(1) 本研究における原理を確立するために後部残響モデルを定義し、どのような後部残響が、STI に影響があるか検討した。拡張型 RIR の立ち上がり立ち下がり制御するパラメータ T_h と T_t が STI に大きく影響を与えると

いうことである。(2) 提案法の性能評価を行うために、STIを変化させたときに単語理解度試験、聴き取りにくさ、わずらわしさがどのように変化するか検討した。STIを操作することによって、スピーチプライバシー保護ができることを明らかにした。(3) 他の方法と比較し、提案法が他の方法（残響音声とピンク雑音）よりもどのように良いか検討した。従来のスピーチプライバシー保護の技術よりもスピーチプライバシー保護が容易にできることを明らかにした。