

Title	音声明瞭度回復を目的とした雑音・残響除去に関する調査研究 [課題研究報告書]
Author(s)	森田, 翔太
Citation	
Issue Date	2010-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/8951
Rights	
Description	Supervisor: 鶴木 祐史, 情報科学研究科, 修士

音声明瞭度回復を目的とした雑音・残響除去に関する調査研究

森田 翔太 (0810062)

北陸先端科学技術大学院大学 情報科学研究科

2010年2月9日

キーワード: 音声明瞭度, 聴き取りにくさ, 雑音・残響除去.

音声コミュニケーションは, 人にとって欠くことのできない情報伝達方法である. しかし, 雑音や残響などの環境によって音声コミュニケーションが阻害されることがある. 例えば, 音声アプリケーションのハンズフリー機能を使う時に, マイクロホンから離れていると会話が阻害され, 話者の発話内容がうまく相手に伝わらないことがある. このような問題を解決するために, 音声の明瞭性を回復する雑音残響除去法が必要となるが, これまでにそのような手法は提案されていない.

本研究では, 雑音残響環境での円滑な音声コミュニケーションの実現を最終ゴールとして, 音声明瞭度の回復を目的とした雑音・残響除去法に関する調査を行った. まず, 円滑な音声コミュニケーションを評価するのに最適な伝達性能の評価方法を調査を行った. 次に, これまでに提案されてきた雑音・残響除去法の調査を行った. 次に, これまでに提案されてきた雑音・残響除去法の調査をした. 最後に, どのようなアプローチにより音声明瞭度を回復する雑音残響除去法の実現ができるかを調査を行った.

音声伝達性能の評価方法を調査した結果, 音声コミュニケーションを評価するためには, 単語の親密度の統制を取りながら単語理解度による評価を行う必要があることがわかった. 単語理解度試験では, 親密度が高い時には, 評価結果に差が現れないにも関わらず, 聴感上では差があることから「聴き取りにくさ」という評価尺度が提案された. 従って, 音声コミュニケーションを考えると音声明瞭度と聴き取りにくさの両方で評価を行う必要があることがわかった. 音声明瞭度の客観評価尺度に Speech Transmission Index (STI) があり, Houtgast と Steeneken によって相関が高いことが示された. しかし, 戸井田らの報告によれば音声明瞭度と STI の相関は低い時もあるとの指摘がなされているが, STI は「聴き取りにくさ」とは非常に相関が高いことが分かっている. 従って, STI は音声明瞭度をうまく表現しきれていないが, 音声コミュニケーションの時に重要である「聴き取りにくさ」をうまく表現していることから, STI は音声コミュニケーションをうまく評価できる客観評価尺度であると考えられる.

続いて、従来の雑音除去法として Spectral Subtraction (SS) 法や Active Noise Cancelling (ANC) 法, Minimum Mean Square Error-Short Time Spectral Amplitude (MMSE-STSA), Winner filtering, 最大尤度法を用いた手法, Real Active Spectral (RASTA) の調査を行った。これらの手法は、雑音のみならば精度良く除去できている。しかし、残響の特性と雑音の特性は全く異なることから、これらの手法で残響を取り除くのは難しい。また、音声明瞭度や聴き取りにくさを回復するような物理指標を使っていないことから、的確に音声明瞭度や聴き取りにくさを回復できないと考える。一方、残響除去法としては、最小位相逆フィルタ処理, Multiple-input/output inverse theorem (MINT) 法, 帯域分割フィルタ処理, Harmonic-based dEReverberation (HERB) についての調査を行った。しかしながら、最小位相逆フィルタ処理や MINT 法, 帯域分割フィルタ処理は、事前に室内インパルス応答を測定する必要があるため実用的でない。また、HERB は単一マイクロホンでブラインド残響処理ができることから他に比べ優れているが、音声明瞭度や聴き取りやすさを回復するような物理指標を用いていないことから、的確に音声明瞭度や聴き取りにくさを回復できない。これらの手法を用いて雑音残響除去を実現するには、雑音除去法と残響除去法を組み合わせる手法でしか実現できない。このような手法は、SS 法と Linear Prediction (LP) を用いた手法や Winner filtering と残響除去に線形フィルタを用いた手法などがある。どちらも雑音除去法と残響除去法の逐次的な処理であり、雑音の成分を減算し、残響成分を線形フィルタリングする手法である。これらの手法は雑音と残響除去手法であり、音声強調を行う手法ではないことから音声明瞭度回復には限界があると考えられる。

一方、音声伝達性能の評価方法である STI と相互関係にある Modulation Transfer Function (MTF) に基づく雑音・残響除去法が提案されている。この手法は、信号の時間包絡線 (エンベロープ) に音声明瞭度に重要な特徴があり、雑音と残響の影響に振幅と位相が影響を受けるため、MTF に基づきエンベロープを回復するという手法である。主観評価である音声明瞭度は、室内の伝達特性 (雑音とインパルス応答) の統計的近似の MTF から計算される客観評価尺度である STI より推定することができる。この STI は、音声明瞭度の試験を雑音残響環境で行える。これらのことより、MTF に基づくと雑音残響環境での音声明瞭度及び聴き取りにくさの回復処理を最善に行えると考える。

音声伝達の評価方法及び従来の雑音・残響除去法の調査を行った結果、MTF に基づいて雑音残響除去を行い音声明瞭度及び聴き取りにくさを回復するというアプローチが最善であると考え、どのような課題があるのかを調査した。その結果、MTF に基づいた雑音残響除去法の実現には、音声のエンベロープ回復だけでは、音声の明瞭度の向上をさせることができないが、キャリアの再生成を行う方法があり、雑音残響環境に頑健な基本周波数推定や音声区間推定などが必要である。今後、これらの推定法の提案を行い、最終的に雑音残響環境において MTF に基づいた音声明瞭度及び聴き取りにくさの回復を実現する。