

音声を効果的に聞こえにくくする楽曲の提示手法

加藤 健人[†] 岡 誠[‡] 森 博彦[‡]

東京都市大学大学院工学研究科[†] 東京都市大学知識工学部[‡]

1. 背景

役所や医療機関の窓口や受付では個人情報が入り口でやりとりされ、その内容がそれを知る必要のない第三者に漏洩する危険性がある。このような事態を防ぐことをスピーチプライバシーと呼び、近年の個人情報保護に対する意識の高まりから重視されてきている。

しかし、窓口や受付といった場合は、来訪者が自由に出入りできるような空間的に開けていることが必要とされる。ゆえに、壁などを用いて音声を遮蔽することは困難である。したがって、このような場でスピーチプライバシーを実現するためには、他の音を用いて会話内容を聞こえにくくすることが有効であり、その手法の開発が求められている。

2. 関連手法および先行研究

音声を聞こえにくくするための既存手法は、スピーチプライバシーを確保するために使用することが困難である。耳栓や耳覆い型の聴覚保護具は各人が装着する必要がある、不特定多数の利用者に対する適用が難しい。アクティブノイズコントロール(ANC)は、制御部やスピーカが他手法と比較して高価であり、音声のような音圧レベルや周波数変動の激しい音を聞こえにくくすることは技術的に困難である。サウンドマスキングシステムは、空調音に似た音を提示することで音声を聞こえにくくする手法で、上記の手法の問題点を回避できる。しかし、この手法では新たな騒音源が増えることになり、その場に騒音源を増やしたくないと考える利用者にとっては導入に抵抗感がある。

上記の問題点を解決するため、Kato[1]はBGMとして導入に抵抗感の少ないと考えられる楽曲を、不特定多数の利用者に提示することで音声を聞こえにくくすることを提案し、どのような楽曲が音声を効果的に聞こえにくくするかについて

検討した。この研究では、音声を効果的に聞こえにくくする楽曲を選択するための指標として加重平均スペクトル距離(WSPD)[2]を用いた。WSPDは音声と楽曲の周波数特性と音圧レベルを用いて算出される指標である。

実験の結果、WSPDにもとづいて選択された楽曲は、音声を効果的に聞こえにくくすることがわかった。また、音声を効果的に聞こえにくくする効果が得られる楽曲の平均音圧レベルは、音声の音圧レベルよりも5dB以上大きい必要があることがわかった。

次に、上記手法の課題点について述べる。実験では、聞こえにくくする対象となる音声の有無や音声の音圧レベルの変化によらず、楽曲を常に音声を効果的に聞こえにくくすることが期待される平均音圧レベルで提示していた。この場合、音声が存在しない場合でも楽曲が音声を効果的に聞こえにくくすることのできる平均音圧レベルで提示されることになり、音声が存在しない比較的静かな状況で提示されるBGMとしては不適切な可能性がある。さらに、会話中に音声の音圧レベルが高くなった場合、楽曲提示による音声を聞こえにくくする効果が得られない可能性がある。そのため、音声の有無や音声の音圧レベル変化に応じて、楽曲の音圧レベルを作業者が不快に感じない程度に変化させることが課題として挙げられた。

3. 目的

WSPDにもとづいて楽曲を選択し、その楽曲が音声を効果的に聞こえにくく、かつ聴取者に不快感を与えないよう提示する手法を提案する。さらに、実験を通じてその有用性を検証することを目的とする。

4. システム概要

本システムの構成は、音声をマイクロホンで電気信号に変換し、コンピュータを用いた処理部へ入力する入力部と、入力信号にもとづいて楽曲選択アルゴリズムを駆動し、音声を聞こえにくくする楽曲を選択し、かつ音声の音圧レベルに応じて楽曲の音圧レベルを変化させる処理部とで構成

Music playing system to be effectively difficult to hear the speech noise.

[†]Taketo Kato, System Information Engineering, Tokyo City University

[‡]Maloto Oka, Hirohiko Mori, Department of Industrial and Management System Engineering, Tokyo City University

される。また、楽曲はスピーカから提示される。

提示される楽曲の候補としては、BGM としてふさわしく、かつ音声(歌詞)が含まれていないものをあらかじめ選択する。音声(歌詞)が含まれない楽曲を選択する理由は、これらが含まれている楽曲は、楽曲提示と同時に起こられる文章を扱う作業の正確性を低下させる[3] ため、本システムの使用が想定される場に不適だからである。

次に、システムの動作概要を述べる。

システムの実行開始後、マイクロホンから 1 分間に入力された信号を用いて背景雑音の平均音圧レベルを測定する。ただし測定中は、発話が開始されていないものとする。

次に、発話の開始を検知する。音声以外の音源による音圧レベルの上昇は短時間であると仮定し、測定された音圧レベルが背景雑音の平均音圧レベルより 5dB 以上高い状態が 3 秒間継続したときを発話開始と定義する。また、この 3 秒間の平均音圧レベルを音声の音圧レベルとする。

次に、発話開始後のマイクロホンからの入力信号にもとづいて、音声を効果的に聞こえにくくする楽曲を選択する。楽曲選択アルゴリズムは、Kato らと同様に、加重平均スペクトル距離(WSPD)が最小となる楽曲を選択する方法を用いる。

選択された楽曲は、音声を効果的に聞こえにくくする適切な音圧レベルで提示する。Kato らの研究から、楽曲は音声の音圧レベルより 5dB 以上高い場合に、音声を効果的に聞こえにくくことがわかっている。本システムが使用される場では、楽曲は BGM として用いられるため、楽曲の音圧レベルはなるべく小さい方がよいと考えられる。したがって、楽曲の音圧レベルは音声の音圧レベルより 5dB だけ高くなるように提示する。

楽曲の音圧レベルは、楽曲の波形データを調整することで変化させる。また、楽曲の音圧レベルは楽曲の波形データから推測する。

音声の音圧レベルは、背景雑音、音声、提示されている楽曲を含んだ音圧レベルから楽曲の音圧レベル分を減算することで算出し、(音声の音圧レベル+5dB)を楽曲提示の音圧レベルとする。

楽曲の音圧レベルは、一定間隔で更新する。その際には、その時間間隔における音声の音圧レベルの平均値を新たな音声の音圧レベルとして用い、それより 5dB だけ高くなるように楽曲の音圧レベルを再設定する。更新の際には、楽曲の音圧レベル変化によって違和感が生じることが予想されるため、これを軽減するために一定時間で波形データを線形的に増減するフェード処理を適用する。楽曲の音圧レベルの更新間隔とフェード処理時間は、聴取者の違和感が少ない設定を、予備実験

を通じて検討する。

楽曲提示の開始後は、発話が終了するまで楽曲を提示する。発話の終了時刻は、音声の音圧レベルが(背景雑音 + 5dB)を 3 秒間下回った時点とする。発話の終了を検知した後、楽曲が提示される音圧レベルを発話開始前の値へ戻す。

5. 1. 予備実験

違和感の少ない楽曲の音圧レベル変化の方法を検討するために実験を実施した。楽曲の音圧レベルの更新間隔は 1 秒, 5 秒, 10 秒, 15 秒, 20 秒とした。また、更新時のフェード処理時間は、更新間隔の 1/4 秒, 2/4 秒, 3/4 秒, 4/4 秒とした。

音声は、あらかじめ録音した男女の会話音声を用いた。また楽曲は、用意したものの中から WSPD が小さいものを 4 曲選択し、使用した。

音声と楽曲は、ともに着座した被験者の正面 2m 前方の 1 組のスピーカから提示した。マイクロホンは被験者からスピーカまでの距離が同じになるように、被験者の近傍に設置した。

被験者は、システムが提示した楽曲を聴取して、楽曲に違和感があったかを回答した。

5. 2. 実験結果

楽曲の更新間隔が 1 秒では、フェード間隔によらず違和感が報告された。一方、5 秒以上の場合には違和感は報告されなかった。

6. 今後の方針

予備実験の結果を踏まえ、楽曲の音圧レベル更新間隔を 5 秒、フェード時間を 1.25 秒としたシステムを用いて、システムが提示した楽曲が音声を聞き取りにくくする程度を調査するための実験をおこなう。実験では、被験者が音声と楽曲を同時に聴取し、音声の内容を聞き取る作業を予定している。

参考文献

- [1] Kato Taketo, Oka Makoto, Mori Hirohiko. "Music recommendation system to be effectively difficult to hear the speech noise" SICE Annual Conference (2013), 2353 – 2359
- [2] 為末ら 「無意味定常雑音存在下での音声聴取時における心理的印象と聴取成績」 『電子情報通信学会論文誌』 J85-A(2), 2002 年, 258-265 頁
- [3] 門間政亮, 本多薫 「音楽に含まれる言語情報が文章課題に与える影響に関する検討」 『人間工学』 Vol45, No.3 2009 年