

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 9 日現在

機関番号：13301

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23500207

研究課題名(和文) 講演音声の高品質収録システム技術の研究

研究課題名(英文) On making quality recordings of distant speech by using a 2D microphone array

研究代表者

三好 正人 (MIYOSHI, MASATO)

金沢大学・電子情報学系・教授

研究者番号：40396205

交付決定額(研究期間全体)：(直接経費) 2,800,000円、(間接経費) 840,000円

研究成果の概要(和文)：講演音声等から生成される音声アーカイブの高品質化に資することを目的として、遠隔音声収録システム技術を検討した。操舵可能なスポットライト形状の指向性收音ビームを用いて低雑音な音声收音を実現するマイクロホンアレイ技術、収録音声の明瞭性の向上に寄与する音響特徴3種の特定とこれらの調整方法に関する知見や、ヒトの聴知覚の特徴を利用してアレイ開口長の短縮に起因する低音感の不足を補い、アレイの小型化(可搬性や設置性の向上)を果たす方法等の成果を上げた。これら成果は電子情報通信学会主催の研究会等で口頭発表した。

研究成果の概要(英文)：In this research, we obtained the following major achievements through a study about distant speech recording. (1) 2 dimensional microphone array producing an easily steerable spotlight-shaped directional beam. (2) Enhancement of the intelligibility of captured speech by controlling 3 acoustic features such as mean pitch frequency, pitch frequency range and spectral intensity around 3 kHz. (3) Bass augmentation of speech captured through a portable microphone-array of an insufficient aperture length by using psychoacoustic phenomenon MISSING FUNDAMENTAL. These achievements are expected to be useful for making quality recordings of lecture speech.

研究分野：総合領域

科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス

キーワード：遠隔音声収録 指向性收音 雑音除去 明瞭性 開口長

1. 研究開始当初の背景

講演会や授業等から作成される音声アーカイブの高品質化(低雑音化・高明瞭化)は、e ラーニングサービス等の拡充に重要な役割を果たすものと期待される。

収録音声の低雑音化については、会場雑音等による SN 比(收音音声と雑音のパワー比)低下や、(話者の)ハンドマイクやタイピンマイクの使用に伴う煩雑さ、収録音声音量変動やポップノイズ(息雑音)混入等を防ぐことが重要である。また、収録音声の高明瞭化については、同音声に残留する雑音等を取り除くことや、(話者の)話し方等によって低下した明瞭性を改善することが重要である。

2. 研究の目的

本研究の目的は、上記3つの課題夫々の解決方法を開発すると共に、それら方法を統合・組み合わせたシステム技術を開発し、前記高品質音声アーカイブ生成に資することと定めた。

3. 研究の方法

具体的な検討項目は次の様に定めた。最終年度(H24年度)には、各項目で検討した要素技術の統合方法も検討した。

(1)スポットライト收音

話者(講演者や質問者)の遠隔収録に適したマイクロホンアレイ技術に着目し、まるでスポットライトの様に、話者方向に対する收音感度を上げ、同音声を高 SN 比収録する方法(指向性收音ビーム生成方法)を検討した。

(アレイから見た)話者位置(方向)の変化や移動に対応するため、簡易に、上記スポットライト型收音ビームの方向変更(操舵)が出来ることを要件とした。また(音声による)話者方向同定の可能性を鑑み、既存の音源方向推定技術との組合せ方法についても調査した。

(2)音声信号回復・強調処理

①雑音重畳音声を周波数分析し、周波数毎に、予め推定した雑音を差し引くことで音声品質の回復を図る周波数減算法(SS法)に着目した。具体的には、SS法と雑音減算を変調周波数領域^注で行う方法(MSS法)を比較した。

注:周波数分析のためフーリエ変換した音声を更にフーリエ変換すると変調周波数分析出来る信号に変換される。

②音声の基本周波数(F0)分析や周波数分析を用いて、明瞭性に関係の深い音響特徴を特定し、聴覚試験を用いて、これらの操作による明瞭性改善効果を評価した。

(3)要素技術の統合

上記要素技術の統合方法を検討した。

4. 研究成果

(1)スポットライト收音

マイクロホン素子を等間隔直線状に配列し、隣接素子出力信号間の振幅差を調整することにより、同配列(アレイ)の正面方向の收音感度を最大化出来る。また、この最大收音感度方向(收音ビーム方向)は、上記隣接素子出力信号間の時間差調整により、任意方向(φ_0 [rad])へ向けること(操舵)も出来る(図1)。

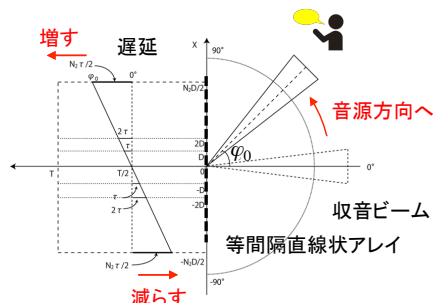


図1. 指向性收音ビームの操舵

この原理を元に、マイクロホン素子を縦(上下)方向に等間隔直線配置したアレイを用いて、同アレイに直交する水平面内の收音感度を最大化し、更に、同アレイを1組として、これを横(左右)方向に等間隔配列し、各組間の時間差を調整することで、簡易な水平面内操舵が可能なスポットライト型指向性收音ビームの生成方法を考案した(図2)。

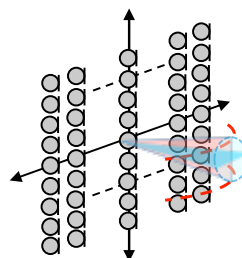


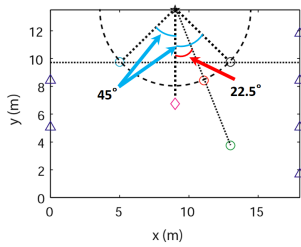
図2. スポットライト型指向性收音ビーム

提案方法による收音効果を確認するために行ったシミュレーション実験の結果を図3に示す。同図(a)は音源(話者)の配置を、(b)は次式を用いて評価した提案方法の雑音低減効果(NRR: Noise Reduction Rate)を夫々示す。

$$NRR = 10 \log_{10}(P_{M1}/P_{M2}) \text{ [dB]}$$

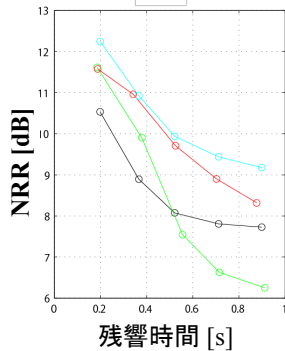
ここで、 P_{M1} はアレイ中央の全指向性マイクロホン素子で收音した雑音音量を、 P_{M2} はスポットライト型收音ビームで收音した雑音音量を夫々表す。アレイ(図2)は、縦横夫々4.25 cmの等間隔配置した19個のマイクロホン素子で構成した。

図3(b)より、残響時間が0.2秒程度の場合には良好なNRR値(10~12 dB)の得られたことが分かる。残響時間の増加につれてNRR



(a) 音源配置図

○, ●, ○, ○: 音源(話者)
△, ◇: 雑音源



(b) NRRと残響時間の関係

図 3. 実験

値が減少する理由のひとつは、雑音の拡散性が増すためだと考えられた。この問題に対処する方策のひとつは、指向性收音ビームの幅を（広い周波数範囲に渡って）狭めることである。收音ビームの幅はマイクロホンアレイの開口長（縦横の寸法）による。例えば、300 Hzの音に対して20°幅の收音ビームを生成するための開口長は4.25 mである。

しかし、開口長の伸長はマイクロホンアレイの設置性や可搬性等の利便性を損なう。そこで、この問題への対処方法（開口長の短縮方法）は音声信号回復処理との統合を検討する際の課題とした。検討内容は下記(3)にまとめた。

尚、話者方向同定については、縦アレイの指向性感度の高い水平面を話者（壇上の講演者や起立する質問者）の“口の高さ”に合わせることで、着席する聴講者等高さの異なる音源の影響を軽減出来ると云う提案方法の特徴と、音量の大きい音源に対する方向推定精度の高い従来技術（Minimum Norm法）との組み合わせが良好に機能する可能性を簡単なシミュレーション実験で確認した。

(2) 音声信号回復・強調処理

①定常白色雑音を重畳した音声を用いて、周波数減算法(SS法)と変調周波数減算法(MSS法)の雑音低減効果を比較した。音声と重畳雑音のパワー比(SN比)は10dB, 0dB, -10dBとした。10及び0dBの場合はSS法が、-10dBの場合はMSS法が夫々他方を上回った。また、MSS法では、ミュージカル雑音と呼ばれる人工的な雑音の発生は抑えられた。

②まず、音声の明瞭性に関係の深い音響特徴を特定した。話し言葉の発声練習(ボイストレーニング)前後の音声や、静音・雑音環境下の音声を比較し、明瞭性を意識して発声された音声の特徴を分析した。その結果、音声の基本周波数(F0)の平均値は高く、F0の変動幅は大きく、3 kHz付近の周波数の音は強い(スペクトル強度は増す)ことが分かった。

次に、録音音声について、上記3つの音響特徴を操作し、同音声の明瞭性(聞き取りやすさ)の改善程度を評価した。平均F0の調整幅は $\pm 1/2$ オクターブの範囲、F0変動の調整幅は(元音声の)1/2~2倍の範囲、3 kHz付近のスペクトル強度は0~16 dBの範囲で夫々調整した。試験音声は、(外部雑音の遮音されない)開放型ヘッドホンを用い、52 dBの音圧レベルで提示した。外部雑音は、静音(27 dB), 52 dB, 73 dBの3条件とした。被験者は5名とした。次の実験結果を得た。

・3つの音響特徴夫々の明瞭性改善効果を調査した実験では、平均F0とF0変動幅の調整が全ての条件で良好に作用した。これらに比べ、3 kHz付近のスペクトル強度の作用は小さかった。また、雑音の音量や種類によって、音響特徴夫々の最適値は異なることも分かった。

・3つの特徴の組合せによる明瞭性改善効果を調査した実験では、平均F0とF0変動幅の組合せが全ての条件で良好に作用した。この組合せに3 kHz付近のスペクトル強度を加えた場合は、(空調雑音の様な)定常白色雑音下では作用しなかったが、レストランやゲームセンターにおける背景雑音の様に、低周波数域のパワーの大きい雑音下では有効に作用した。

尚、雑音の音量や種類によって異なる各特徴の最適調整値については、今後の検討課題としたい。

(3) 要素技術の統合

前記マイクロホンアレイ開口長の短縮を課題として、スポットライト收音と音声回復処理の統合方法を検討した。開口長4.25 mのアレイで収録した0.3~4 kHzの音声(基準音声)と開口長を短縮し、指向性收音可能な下限周波数を上げて収録した音声(比較音声)の音質を比較した。

図4に指向性收音ビームの周波数特性を示す。指向性收音可能な下限周波数(f_L [kHz])の上昇(開口長短縮)につれて、0.3~ f_L [kHz]間のビーム幅は拡がり、收音音声に重畳する拡散性雑音の量は増加する(青線)。下限周波数 f_L [kHz]によらず、雑音量を一定にするため、ビームの特性と逆特性となる高域通過フィルタ(HPF; 赤線)を用いる。すると、比較音声の低周波数帯域は(基準音声に比較して)抑圧され、音声の品質は劣化する。

比較音声の品質劣化を改善する方法とし

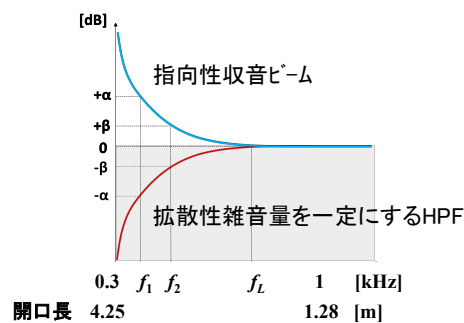


図 4. 指向性收音ビームの周波数特性

て、VB 処理 (Virtual Bass 処理) に着目した。Missing Fundamental と呼ばれるヒトの聴覚現象を用いて、低音感を補う処理である。Missing Fundamental とは、音声の基本周波数 (F0) の連続倍音 (例えば、2~5 倍音等) のみを再生することで (再生されない) F0 が知覚される現象である。

マイクロホンアレイの開口長短縮によって劣化した收音音声の音質を VB 処理と雑音低減処理の組合せで回復する方法の概要を図 5 に示す。VB 処理では、推定した F0 の 2~5 倍音を中心周波数とする 1/3 オクターブの周波数範囲夫々を 5 dB 程度増幅した。尚、予備実験の結果、雑音低減処理には周波数減算法 (SS 法) のパラメータ調整を詳細化した一般化 SS 法を適用した。

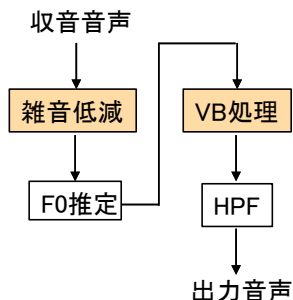


図 5. 收音音声の音質回復方法

男女夫々 2 名の音声について、指向性收音可能な下限周波数 f_L [kHz] を $0.4 \leq f_L \leq 1$ の間で変化させ、(図 5 の方法で) 比較音声を作成した。拡散性雑音を模した白色雑音を音声に重畳した。 f_L [kHz] 以上の指向性收音可能な周波数帯域での SN 比は 10 dB とした。3 名の被験者に、基準音声と比較音声の音質を 7 段階で対比較評価させた (CMOS 評価)。その結果、指向性收音可能な下限周波数を 0.7 kHz としても、即ち、開口長を (4.25 m から) 1.83 m へ短縮しても、音声品質に大きな劣化は生じないことが分かった。

尚、前記音声強調処理方法との統合、変調周波数減算法 (MSS 法) の適用、及び実測した講演音声等を用いる技術検証については、今後の課題としたい。

5. 主な発表論文等

[学会発表] (計 7 件)

- ①阿部悠輝, 丹生賢, 齋藤毅, 三好正人, 指向性收音音声の強調方法について, 信学技法, EA2013-122 (東京, 2014. 3).
- ②高山貴史, 齋藤毅, 三好正人, 雑音下における発話で変化した音響特徴が音声の聞き取り易さに与える影響, 信学技法, EA2013-100 (金沢, 2013. 12).
- ③丹生賢, 阿部悠輝, 齋藤毅, 三好正人, マイクロホンアレイの開口長に起因する收音音声劣化の改善について, 信学技法, EA2013-91 (金沢, 2013. 12).
- ④高山貴史, 齋藤毅, 三好正人, 雑音下で音声明瞭性に寄与する音響特徴について, 平成 25 年度電気関係学会北陸支部連合大会, G-5 (金沢, 2013. 9).
- ⑤丹生賢, 齋藤毅, 三好正人, マイクロホンアレイの開口長收音音声品質の関係について, 平成 25 年度電気関係学会北陸支部連合大会, G-2 (金沢, 2013. 9).
- ⑥丹生賢, 堀内亮輔, 齋藤毅, 三好正人, 遠隔音声のスポットライト收音について, AES ジャパンコンファレンス・仙台 (仙台, 2012. 12).
- ⑦岩崎崇幸, 齋藤毅, 三好正人, 音声の変調スペクトルを用いる雑音低減方法について, 信学技法, EA2011-70 (長野, 2011. 10).

6. 研究組織

(1) 研究代表者

三好 正人 (MIYOSHI Masato)
金沢大学・理工研究域電子情報学系・教授
研究者番号: 40396205

(2) 研究分担者 無

(3) 連携研究者

齋藤 毅 (SAITOU Takeshi)
金沢大学・理工研究域電子情報学系・助教
研究者番号: 70446962

西川 清 (NISHIKAWA Kiyoshi)

金沢大学・名誉教授
研究者番号: 40019774