

平成 22 年 5 月 31 日現在

研究種目：若手研究（B）

研究期間：2008 年度～2009 年度

課題番号：20700167

研究課題名（和文）

携帯電話に実装可能な口唇動作抽出センサの開発

研究課題名（英文）

Development of Lip Movement Sensor for Mobile Phone

研究代表者

吉田 孝博（東京理科大学・工学部・助教）

研究者番号：10385544

研究成果の概要（和文）：

照射した赤外光の口唇での反射の変化により口唇動作をセンシングする口唇動作抽出センサにおいて、携帯電話での利用時に問題となるセンサ位置の変動に対応させた。本研究では、9つのセンサを携帯電話の文字盤部分に配置し、有効なセンサ信号を発話単位で自動的に選択する手法を開発した。口唇動作のみによる 50 単語認識では、センサ位置の変動があっても平均認識率 74.1%が得られ、従来と比較して 7.2%の改善が実現できた。

研究成果の概要（英文）：

Lip movement sensor senses speaker's lip movement by reflection of infrared light. When the sensor uses on mobile phone, there is a problem that the sensor position is different in each time. In this study, we improved robustness for the sensor position by installing nine lip movement sensors on the mobile phone and selecting the best sensor signal automatically. We achieved 74.1% Japanese fifty words recognition rate only by lip movement under different sensor position by each user. It was 7.2% improvement.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008 年度	800,000	240,000	1,040,000
2009 年度	800,000	240,000	1,040,000
年度			
年度			
年度			
総計	1,600,000	480,000	2,080,000

研究分野：ヒューマンマシンインタフェース、デジタル信号処理

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：携帯電話、口唇動作、センサ、単語認識

## 1. 研究開始当初の背景

発話時の口唇動作には、発話内容と関連のある有用な情報が含まれるため、携帯電話が利用者の口唇動作を得られると、通話中に音声を発しない口唇動作による携帯電話の補助操作や電話自動応答サービス等のユーザビリティ向上、音声を発しない口唇動作によるパスワード認証機能による携帯サイト利用時のセキュリティ向上など、携帯電話の高度化が実現できる。

しかし、口唇動作情報のセンシングには、ビデオカメラで動画像を撮影し各フレーム単位で画像処理を行い口唇形状を抽出する必要があり、撮像デバイスや画像処理用の高速・大規模な回路が必要である。このため、サイズ・コスト・消費電力の面で携帯電話への実装には向かない。

そこで当研究グループの従来研究にて、低コスト、小サイズ、そして口唇動作抽出に要する計算処理量が極めて少なく、携帯電話への実装に適した特徴を持つ口唇動作抽出センサを提案した。しかし、携帯電話利用時に問題となる携帯電話の位置の変動によるセンサ位置ずれの影響などへの対応が課題として残されていた。

## 2. 研究の目的

本研究課題では、口唇動作抽出センサの携帯電話への実装を目的として、携帯電話利用時の位置とセンシング性能の基礎調査、口唇動作センサのマルチチャンネル化、そして携帯電話利用時のセンサ位置ずれへの対応手法の開発を行った。

## 3. 研究の方法

(1) 携帯電話の持ち方の個人差や携帯電話のサイズ・形状等により、携帯電話と口唇の位置関係が変化し、センサ位置がずれる。このセンサ位置ずれによる影響を調査するため、携帯電話通話中の携帯電話と利用者の口唇との位置関係およびセンサ位置による口唇動作抽出精度を調査し、センサ改良のための基礎データを得る。

(2) 口唇動作をセンシングできる範囲を拡大するため、口唇動作抽出センサを携帯電話の文字盤上に3×3のマトリクスで配置し、

マルチチャンネル化した口唇動作抽出センサを試作する。

(3) センサ位置ずれによる影響を回避するため、マルチチャンネル口唇動作抽出センサから得られる9chのセンサ信号から、有効なセンサ信号を発話単位で自動選択する手法を開発する。

(4) 開発したマルチチャンネル口唇動作抽出センサと、有効なセンサの自動選択手法を評価するため、複数の被験者に単語を発話してもらい、実験用の単語発話データベースを構築し、50単語認識実験を行う。

## 4. 研究成果

### (1) センサ位置ずれに関する基礎調査

携帯電話利用時の位置ずれに対応できるセンサの実現を目的とし、センサの位置ずれによる口唇動作抽出精度への影響について調査を行った。また、携帯電話通話中の携帯電話の実際の下端位置を、30名の被験者から測定して、当センサの改良に必要な口唇動作抽出可能位置の範囲の特定を行った。

図1に、各センサ位置に対する単語認識率と携帯電話下端位置の関係を示す。この調査は、口唇周辺の120箇所における単語発声時の口唇動作を抽出し、単語音声認識エンジンによる単語認識率にて評価を行った結果である。センサ位置の座標系は、x軸を口唇の端点同士をつなぐ線、y軸を鼻梁と口唇の中心とし、顔表面から5mm離れた位置とした。×印は携帯電話下端位置である。図中の色は、赤いほど認識率が高いことを意味する。

図1より、赤色で示される口唇中央付近と口唇側面(水平方向)の2領域では90%前後の単語認識率が得られ、高い精度で口唇動作を抽出できるが、それらの領域から離れると抽出精度が大きく低下することが分かる。

また、携帯電話下端の位置は、図2中の×印の位置で分布しており、携帯電話の形状や利用者の持ち方に左右されるため、これらの位置ずれに対応するためには、センシング可能な範囲をこれらの分布をもとに広げる必要があるといえる。

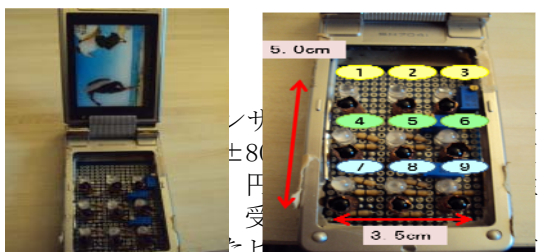
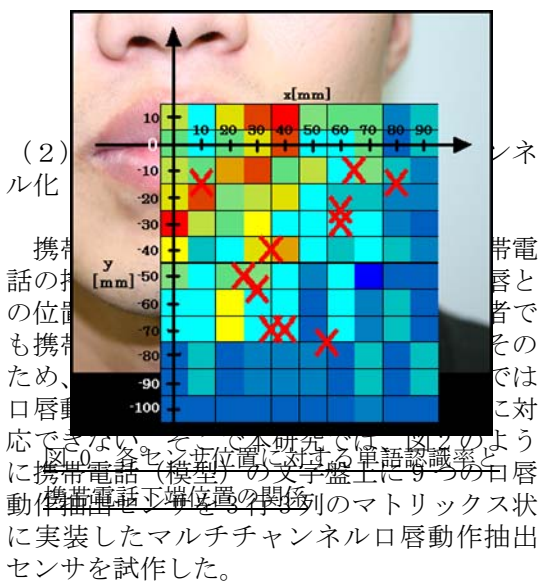


図0 各センサ位置に対する単語認識率と携帯電話下端位置の関係  
図1 携帯電話型マルチチャンネル口唇動作抽出センサ

光波長とフォトトランジスタのピーク波長は、近い組み合わせを選ぶことによりセンサの感度向上を図っている。また、フォトトランジスタの可視光カットフィルタにより、蛍光灯などの外来光の影響を低減させている。

この口唇動作抽出センサから得られる信号は、OPアンプを用いた増幅回路で増幅し、9つのセンサの各チャンネルに対してA/D変換してPCに取り込み、 $f_s=90\text{Hz}$ の口唇動作信号を得る。なお、マルチチャンネルA/Dコンバータには、入力チャンネル数16ch、量子化ビット数16bit、最高サンプリング周波数が1MHzであるInterface社製のCSI-360116を用いた。その後、QMFによる帯域分割で口唇動作特徴量に変換する。本研究では、8次FIR型QMF (Quadrature Mirror Filter) を、4段のツリー構成にすることで16分割の周波数帯域を持つフィルタバンクを構成し、センサ信号を帯域分割した。

(3) マルチチャンネルセンサ信号の自動選択による携帯電話利用時のセンサ位置ずれへの対応

携帯電話利用時のセンサ位置ずれにロバストにするためには、9つの口唇動作センサ信号から、有効な口唇動作が得られるセンサを逐次選択する必要がある。

そこで本研究では、得られたセンサ信号のダイナミックレンジ (DR) もしくは振幅により、発話毎に1つのセンサ信号を自動選択する手法を開発し、位置ずれによる影響の低減を図った。

- ・方法1 (DR最大) …1発話毎に、センサ信号の最小電圧と最大電圧の比が最大となるセンサを選択。

- ・方法2 (振幅最大) …1発話毎に、センサ信号の最小電圧と最大電圧との差が最大となるセンサを選択。

(4) センサ信号のみを用いた50単語認識実験による性能評価

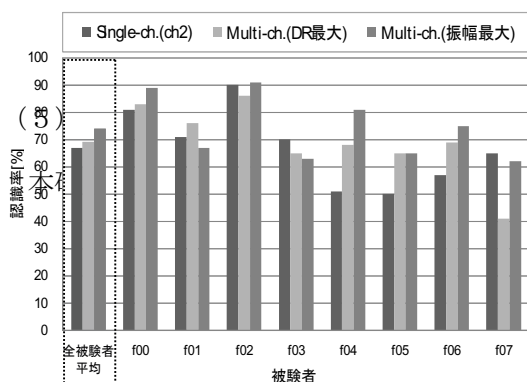
開発したマルチチャンネル口唇動作抽出センサと、有効なセンサの自動選択手法を評価するため、複数の被験者による単語発話データベースを構築し、音声を発しない50単語認識実験を行った。

携帯電話型マルチチャンネル口唇動作抽出センサを、各被験者に通常の話通時の位置に持ってもらい、暗室において50単語の発話データを採取した。この実験条件を表1に示す。

ま  
こ  
こ  
こ  
こ  
こ  
ま  
い

セアルの字首を行なっている。

図2より、被験者によりセンサ信号選択効果が大きく異なるものの、マルチチャンネル口唇動作抽出センサと自動選択（振幅最大）手法の組み合わせで、平均単語認識率 74.1% が得られ、従来の単一センサと比較して 7.2% の改善を実現できた。センサ信号の選択方法は、DR 最大よりも振幅最大のほうが精度は高かった。



チャンネル化し、有効なセンサを発話毎に振幅で自動選択することで、平均認識率 74.1% が得られ、携帯電話利用時のセンサ位置ずれに耐性を持たせることができた。

本研究成果により、低コスト、小サイズ、そして小計算処理量の利点を持つ口唇動作抽出センサの実用化に、より一歩近づくことができた。今後は、外来光による影響の低減などに関して更に実環境適応性を高めることにより、携帯電話への搭載、そして携帯電話の高度化が期待される。

## 5. 主な発表論文等

[学会発表] (計 1 件)

(1) 天野崇、吉田孝博、和田直哉、半谷精一郎、“携帯電話利用時の位置ずれにロバストなマルチチャンネル口唇動作抽出センサ”、電子情報通信学会総合大会、2010/3/18、東北大学

## 6. 研究組織

(1) 研究代表者

吉田 孝博 (東京理科大学・工学部・助教)  
研究者番号：10385544

(2) 研究分担者

(3) 連携研究者

(4) 研究者協力者

天野 崇 (東京理科大学大学院・工学研究科・修士課程 1 年生)