

モバイル端末におけるドップラー効果を用いたハンドジェスチャ認識手法の提案

村石 映

Hayuru MURAISHI

1 はじめに

近年、人間にとってより直感的かつ自然な操作を可能にするジェスチャ認識に注目が集まっている。また、スマートフォンやタブレット端末といったモバイル端末が近年急速に普及している。近年のモバイル端末は地磁気センサや加速度センサといった様々なセンサを内蔵しており、これらセンサを用いてジェスチャ認識をする研究が多く存在する。例として、モバイル端末内蔵のカメラを用いたものやモバイル端末内蔵の照度センサを用いてハンドジェスチャを認識する研究がある。しかし、例にあげた2つの研究は照度が低く暗い環境においてはジェスチャの認識率が低くなるという課題がある。

本研究では、モバイル端末を用いたジェスチャ認識における課題を解決する手法として、モバイル端末内蔵のマイクとスピーカを用いたハンドジェスチャ認識手法を提案する。提案手法は専用の特殊デバイスを必要とせず、モバイル端末1台のみでジェスチャを認識可能とする。本研究では、モバイル端末に手が近づく場合と離れる場合の2種類のジェスチャのドップラー効果を測定する。測定したデータをもとに2種類のジェスチャの認識精度の検証を行う。この検証をもとにドップラー効果を用いたジェスチャ認識手法の可能性について考察する。

2 関連研究

モバイル端末を用いて、ハンドジェスチャ認識を行う研究が多く存在する。例として、モバイル端末に内蔵している照度センサを用いたジェスチャ認識の研究がある¹⁾。内蔵照度センサを用いたジェスチャ認識では照度センサによってモバイル端末表面の照度値を取得する。ハンドジェスチャを行うとジェスチャを行った際に生じる影によって、照度値に揺らぎが生じる。照度値の揺らぎからハンドジェスチャを複数識別する。

Song²⁾らは、スマートフォン内蔵のRGBカメラを用いてジェスチャの認識を行う。カメラ画像から手の形を認識しジェスチャを複数識別する。しかし、これらの手法は明るさの足りない暗い環境では認識率が低くなる課題がある。提案手法では、モバイル端末内蔵のスピーカとマイクを用いてハンドジェスチャ認識を行う。これにより使用環境の明るさに左右されないジェスチャ認識を可能とする。

3 ドップラー効果を用いたジェスチャ認識

提案手法では、モバイル端末内蔵のスピーカとマイクを用いてドップラー効果を検知し、ジェスチャを認識する。人には聴き取ることのできない非可聴域の音を内蔵のスピーカから出力する。スピーカから音を出力した状態で、内蔵のマイク付近でハンドジェスチャを行う。ジェスチャを行うと、出力した音が手に反射してマイクに入る。手に反射した音は手の移動によってドップラー効果が起きており、スピーカから発した周波数とは異なる周波数となる。そこでスピーカから発した周波数前後の周波数の強さを計測することによりドップラー効果を検知する。ドップラー効果を検知することでジェスチャを識別する。本研究ではモバイル端末に手が近づくジェスチャを APPROACH、遠ざかるジェスチャを AWAY と定義し2種類のジェスチャを識別する。それぞれの手の動きのイメージを Fig. 1 に示す。また2種類のジェスチャによって起こるドップラー効果の発生例をグラフとして Fig. 2 に示す。

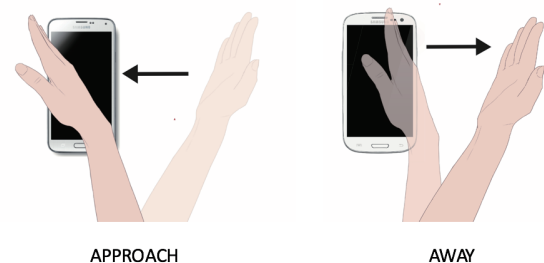


Fig. 1 APPROACH と AWAY のイメージ図

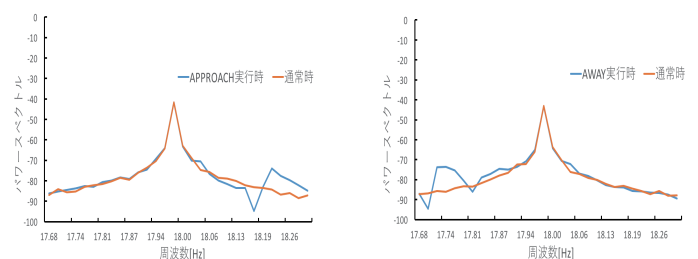


Fig. 2 APPROACH と AWAY のドップラー効果発生図

また以下にジェスチャを識別するために用いる式を示す。スピーカで発した周波数より高い周波数域に用いる

式が d_α , 低い周波数域に用いる式が d_β である. ジェスチャの識別を行う式は N である.

$$d_\alpha = \sum_{i=1}^{n/2} \{|P[f_{0+i}] - P_L[f_{0+i}]\}| \quad (1)$$

$$d_\beta = \sum_{i=-n/2}^{-1} \{|P[f_{0+i}] - P_L[f_{0+i}]\}| \quad (2)$$

$$N = d_\alpha - d_\beta \quad (3)$$

f_0 : スピーカから発する周波数 [Hz]
 P : 今回取得したパワースペクトルの値
 P_L : 前回取得したパワースペクトルの値
 n : 使用する標本数

4 ジェスチャ認識精度検証

4.1 実験概要

本実験では, モバイル端末を用いて APPROACH と AWAY の 2 種類のジェスチャの計測を行い, 計測結果からドップラー効果によるハンドジェスチャ認識手法の精度検証を行う. 出力する音の周波数は非可聴域の 18kHz である. マイクが取得するサンプリングレートは 44.1kHz であり, サンプルした音に対して標本点数 4096 で高速フーリエ変換を行う. 高速フーリエ変換によって得られる, パワースペクトルの波形を約 0.3 秒間隔で取得する. この際, スピーカで出力する周波数 18kHz の前後 17.68kHz から 18.3kHz までの周波数を 21.53Hz 間隔で取得する. また取得した周波数それぞれに対応するパワースペクトルを取得する. 18kHz 付近におけるパワースペクトルの変化によってジェスチャを認識可能か検証を行う. 被験者は 20 代の大学生 5 人であり, 使用するモバイル端末は Nexus 7 である.

4.2 実験結果と考察

全被験者の 18kHz のデータを用いて 1 個抜き交差検証を行った. 被験者ごとのデータを用いた 1 個抜き交差検証は, テストデータおよび訓練データの両方を 1 人の被験者のデータのみで構成し, 被験者ごとの分類精度を評価する. 1 個抜き交差検証の結果を Table 1 に示す. 結果より APPROACH と AWAY の認識率は低い, APPROACH を AWAY と認識することや AWAY を APPROACH と認識する誤認識は少なかった. 一方, ジェスチャを行っても NONE と認識する割合が平均 29.0% であり, ジェスチャを行っても何もしていないとシステムが認識する割合が高かった. これらのことから, ジェスチャを認識するまでジェスチャを繰り返すことで APPROACH と AWAY の認識は可能であると考え. また APPROACH と AWAY の認識率の低さは, NONE

の分散の値に対して APPROACH と AWAY の分散の値が大きいことが挙げられる. この問題は訓練データの個数を増やすことで分散の値が小さくなるため改善すると考える.

Table 1 全被験者データに対する 1 個抜き交差検証結果

	APPROACH	AWAY	NONE
APPROACH	68.0 %	0 %	32.0 %
AWAY	2.0 %	72.0 %	26.0 %
NONE	0 %	0 %	100.0 %

また, 被験者ごとにデータを分割し, ある被験者 1 人のデータをテストデータ, その他の被験者のデータを訓練データとするユーザ抜き交差検証を行った. ユーザ抜き交差検証では未知のユーザに対する汎用性を評価する. ユーザ抜き交差検証の結果を Table 2 に示す. 結果から, 全体の平均認識率は 88.7% と高い正確性を示し, 未知のユーザに対して高い汎用性があると考え. もっとも結果の悪かった被験者の認識率は 73.3% であり, 理由としてジェスチャの速度が遅いことが挙げられた.

Table 2 全被験者分のユーザ抜き交差検証平均値

	APPROACH	AWAY	NONE
APPROACH	78.0 %	0 %	22.0 %
AWAY	2.0 %	98.0 %	2.0 %
NONE	6.0 %	0 %	90.0 %

5 結論

既存のモバイル端末を用いたジェスチャ認識手法における, 暗い環境での認識率低下の課題に対して, 課題を解決するハンドジェスチャ認識手法を提案した. 提案手法はモバイル端末内蔵のスピーカとマイクを用い, ドップラー効果によってジェスチャを識別する. ジェスチャの認識率は比較的高く, 提案手法を用いたジェスチャ認識が実現可能であることを示した.

参考文献

- 1) Kento Matsui and Hiroto Aida and Hikaru Ichikawa and Mitsunori Miki: Hand Gesture Interface using Light-dark changes in an Illuminance meter built in mobile devices, WORLD-COMP2015
- 2) Song, Jie and Sörös, Gábor and Pece: In-air Gestures Around Unmodified Mobile Devices, UIST,978-1-4503-3069-5(2014)