

**Title:** Leveraging Screen-Integrated Speakers for Hand-Pose Recognition in Mobile Interfaces

(画面一体型スピーカーを活用したモバイルデバイスの近接領域でのハンドポーズ認識)

**Authors:** Kaori Ikematsu\*, and Kunihiro Kato\* (\*joint first authors)

(池松 香 (LINE ヤフー株式会社 上席研究員) 、(加藤 邦拓 (東京工科大 助教))

**Journal:** International Journal of Human-Computer Interaction, pp.1–21, (2025).

**掲載年月:** 2025 年 12 月

**研究概要:** 本研究は、スマートフォンのディスプレイ全体を振動させて音を出力する Acoustic Surface (画面一体型スピーカー) 技術を活用し、画面上空 (近接領域) でのハンドポーズ認識を実現するアクティブ音響センシング手法を提案しています。従来のスピーカー構成に比べ、認識可能なエリアを画面全体に広げ、タッチ操作とシームレスに組み合わされる新しいインタラクションを提供します。

**研究背景:** スマートフォンに内蔵されたセンサを用いて、画面上空・デバイス周辺領域でのジェスチャ入力を実現する手法は、これまでに多くの研究者によって模索されてきました。既存手法 (カメラ、静電容量タッチセンシング) では、追加のアタッチメントの実装や、非公開のフレームワークの使用が必要でした。また、音響センシングを用いた手法では、従来のスマートフォン (上部スピーカー・下部マイク) のハードウェア構成では指向性の影響で画面上空の広い範囲をカバーすることが困難であるという課題がありました。

**研究成果:** 本研究では、ディスプレイ全体を振動板として利用する Acoustic Surface 技術を用いることで、安定した信号を広範囲に放射し、追加ハードウェアなしで画面近接 (5cm 以内) のハンドポーズを認識可能にします。提案手法では、人耳には聞こえない高周波音 (16–24 kHz) をスマートフォン画面スピーカーから放射し、手による反射音の変化を内蔵マイクで集音して解析します。反射音の信号に対して FFT (高速フーリエ変換) を用いて、可聴域を除いた 172 個の特徴量を抽出します。このトク著量に対して、ランダムフォレスト (RF) や SVM などの機械学習モデルを用いて、ポーズを分類します。これにより、

・**片手ポーズ (Single-hand poses):** デバイスを保持した手の親指による 5 つの空中ポーズ + ニュートラル (静止) 状態、

・**両手ポーズ (Double-hand poses):** デバイスを保持していない手による 4 つの形状 (人差し指のみ、ピンチ、全指など)

+ ニュートラル状態、がそれぞれ識別可能となりました。

**社会的・学術的なポイント:** Acoustic Surface により追加ハードなしで画面近接領域の静的ハンドポーズ認識を高精度に実現した点、グリップ判定を先行させる階層型モデルの有効性と設計指針を確立した点に大きな意義があります。これにより、スマートフォンの片手操作時のアクセシビリティ向上に加え、既存デバイスをそのまま活用して高度なショートカット入力を可能にする実用的な応用事例を提示しました。

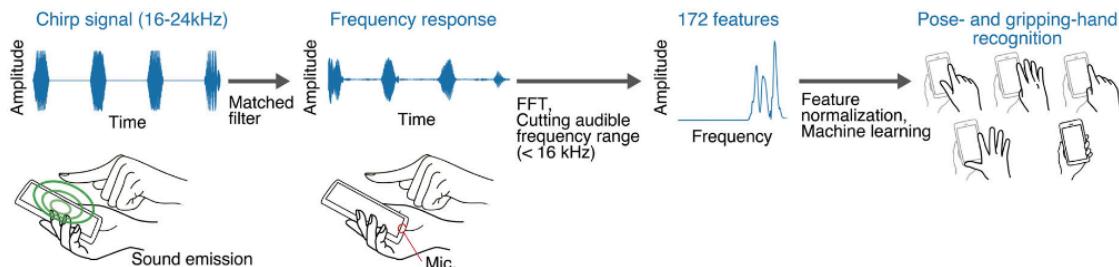


図 1 手のポーズ認識のパイプライン。画面スピーカーからチャーブ信号を出力し、手に反射した音をマイクで取得。反射音を解析し、機械学習によって手のポーズを識別する。

**用語解説:**

Acoustic Surface : スマートフォンの画面全体を振動させて音を発生させる技術。LG や Huawei など一部のスマートフォン製品に採用されている。