

# 家電との「アイコンタクト」による操作手法の提案とそれを用いたスマートライトの開発

佐々木 雄司<sup>1,a)</sup> 中屋 悠資<sup>2</sup> 尾崎 正和<sup>1</sup> 増井 俊之<sup>1</sup>

**概要:** 多くの電子機器を単一の方法で操作する必要性が指摘されており、多くのデバイスを音声コマンド等で操作する手法の開発が求められている。そのためには操作対象の識別が必要である。ラベリングなどの手法がとられてきたが、名前を憶えるのが難しい、名前を知らない人が使えないなどの問題がある。これらを解決するために視線を用いて操作対象を指定する研究が行われているものの、正確性の低さから実用化に至っていない。本研究では、視線で選択されようとしている家電がインジケータによってフィードバックすることで、人間との「アイコンタクト」を実現し、ストレスのない操作を可能にする手法を提案する。

## 1. はじめに

多数の家電や電子機器が家庭に存在することでリモコンやスイッチが増えており、単一の方法でこれら多くのデバイスを操作する必要性が指摘されている [1]。最近では、スマートフォンやスマートスピーカーの普及により、ネットワークに接続した家電、いわゆるスマート家電が広く利用されるようになってきた。多くのスマート家電では、スマートスピーカーやスマートフォンのアプリから遠隔操作できることが特徴のひとつとなっている。特に、音声を用いた操作は、利便性だけでなく移動することが難しい高齢者や障がい者の生活をサポートする面でも有効だと言われている [2]。

多くのデバイスを1つの方法で制御するには操作対象を識別する必要がある。操作対象の指定には家電に名前を登録することが考えられるが、事前に命名スキームを知っておく必要があり、訪問者はそれを推測できないという問題がある [3]。「ただいま」といった状況を示すキーワードを用いて事前に登録した複数の家電の操作を行えるようになるものもあるが、正確な操作に向かないうえ、キーワードを覚えている必要がある点では家電に名前を登録しておくことと変わりはない。

この問題を解決するために操作対象の指定に視線を用いる研究が行われており、視線による選択と音声に限らない様々な命令方法を組み合わせることで家電を操作しようとしている。Brumitt らの研究によると、ライトを操作す

るときに対象物を全く見ない人は9%であり、それ以外は66%が常に見ており、25%は必要に応じて見ていた [3]。このことから操作対象の指定に視線を用いるのは自然であるといえる。

Kim らは室内に設置した視線推定モジュールを用いてユーザが見つめているデバイスを特定、さらに同じく設置しているハンドジェスチャ認識モジュールでユーザが意図する操作を特定することで家電を操作するというシステムを提案している [4]。しかし、視線の推定は完全ではなく誤検出される場合があり、正確な操作対象指定は難しい。Smith らは室内に1つの視線検出デバイスを置いてユーザが何を見ているのか推定するのではなく、カメラをユーザが見ているかどうか判別する手法を提案している [5]。このカメラを家電に取り付けることで、ユーザが家電を見ているかの判別に応用できる。しかし、機械学習モデルのトレーニングが環境やユーザに依存するため、一般的に高い精度を実現することは難しいと指摘されている [6]。

このように、視線検出により操作対象を特定する手法は広く研究されており、特定されたデバイスに対してハンドジェスチャや音声コマンドなどによる操作を行えるシステムが開発されている。視線検出により操作対象を特定するには (1) 室内に設置された1つのカメラで何を見ているか特定する方法、(2) 操作対象の家電それぞれにカメラを取り付けて見つめられているか判断する方法などがあるが、いずれも精度が十分とはいえず、実用に至っていない。

このような問題を解決するため、操作対象がユーザの注視に反応してインジケータを点灯させる、人間同士のアイコンタクトのような仕組みを考案した。視線の誤検知による誤作動を抑制し、同時に家電との対話感をユーザに提供

<sup>1</sup> 慶應義塾大学環境情報学部

<sup>2</sup> 東北大学工学部

<sup>a)</sup> y.sasaki@keio.jp

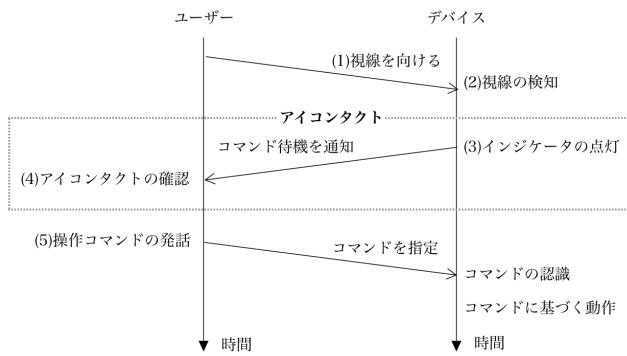


図 1 アイコンタクトの流れ

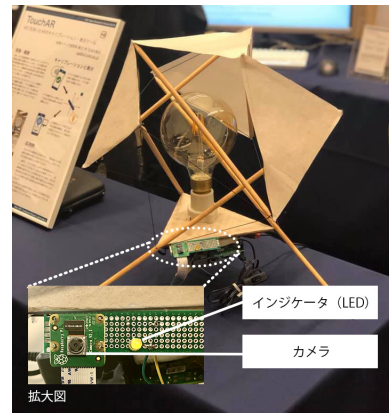


図 2 スマートライト「LuuX」

するスマートライト「LuuX」の開発について述べる。

## 2. アイコンタクトによる家電操作手法

人間同士のアイコンタクトを参考に、人間と家電がお互いに認識し合っていることをインジケータの点灯により示し、視線による操作対象指定を安定的に行えるようにする手法を提案する。ここではこの手法を「アイコンタクト」と称する。

視線による操作対象の検出方法として、ディープラーニングを用いて顔の向きを推定する方法が提案されているが、認識精度が十分でない上、ユーザが意図的に見つめているのか意図せず顔を向けてしまっただけなのか判定することは難しい。これまでのシステムでは、システムがユーザの視線を認識しているかどうかユーザが確認する手段がなかったため、ユーザが認識していないのにコマンドを実行してしまうことがあった。さらに、複数のシステムが近くに並んでいる場合、どちらのシステムがユーザの視線を認識しているか、つまりコマンド待機状態であるかを知ることができなかった。

提案する手法(図 1)では、(1)ユーザが家電を見ると、(2)カメラが視線を検知し、(3)コマンド待機状態であることを示すインジケータが点灯する。(4)ユーザはインジケータの点灯を見て「アイコンタクト」の成立を確認、(5)音声コマンドを発話する。このように「アイコンタクト」を家電選択に取り入れることで、ユーザはシステムに認識されているかどうかを確認した上でコマンド操作することが可能になる。また、操作したい家電を見るという動作は、スマートスピーカーへの指示や通常のリモコンでの操作でも最初に行われることが多いため、ユーザの負担は比較的少なく、自然である。そのため、「家電の側からインジケータを通じてユーザに操作を促す」という UI 上のイディオム [7] を構築できるのではないかと考える。

## 3. スマートライト「LuuX」の開発

前項で説明した「アイコンタクト」を用いて、ライトの

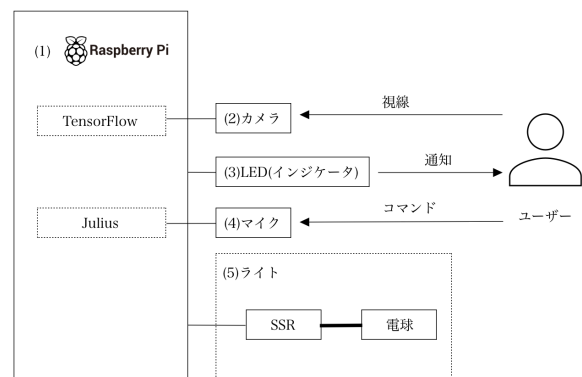


図 3 システム構成図

点灯と消灯を制御できるスマートライト「LuuX」(図 2)を開発した。このデバイスは視線を向けるとインジケータの点灯によりユーザに操作を促し、「点灯」「消灯」などのコマンドを発話するとライトの状態を制御できるというものである。

### 3.1 デバイスの構成

図 3 に示すとおり、(1)Raspberry Pi、(2)顔の向きを判断するためのカメラ、(3)インジケータの LED、(4)音声コマンドを入力するマイク、(5)AC 電源を制御する機構とライトによって構成される。(1)の Raspberry Pi には、ディープラーニングによる顔の向き推定を行えるだけの性能が必要になるため Raspberry Pi 3 B+ を使用している。(2)カメラには Raspberry Pi Camera Module V2、(4)マイクには一般的な USB 入力のマイクをそれぞれ使用した。(5)電源の制御にはソリッドステート・リレー (SSR) を使用し、Raspberry Pi の汎用入出力 (GPIO) ピンの出力を High にすると電源が On になる。

視線の向き判定には TensorFlow Lite [8] を使用し、顔が正面を向いているかどうかを判断している。顔の写真を 5 クラス分 (上下左右正面)、それぞれ 100 枚ずつ集めデータセットを作成した。実験では 84.7% の精度で判定できた。画像の分類ではなく、より正確に顔の向きを推定する技術

も存在するが、処理を行うコンピュータの性能の問題からこの手法を採用した。

音声の認識にはオープンソースのエンジンである Julius[9]を使用した。これは音響モデルや言語モデル、辞書を使用することで、入力された音声を文字に起こしたりコマンドを抜き出したりといった処理をオフラインで行えるというものである。このエンジンに対して、動作のコマンドとして考えられる「点灯」「消灯」「ついて」「きえて」などの言葉を登録した辞書を与え、辞書中のいずれかの言葉である可能性が高い場合に、コマンドが発話されたと判定した。

なお、ランプシェードはテンセグリティ [10] を応用した構造を用いてデザインし、後述のデモ展示で実際の利用シーンを想定してもらうために実用性が高いものにした。

### 3.2 処理の流れ

LuuX の電源を入れるとカメラがユーザの顔が正面を向いているかどうか判定し続ける。TensorFlow でカメラからの画像を解析し、正面を向いていると判定された場合にはインジケータの LED を点灯させてユーザを認識していることを示す。インジケータの LED が点灯している間は音声コマンドの入力を有効化し、ユーザからの操作を受け付ける。音声処理エンジンの Julius がコマンドを検出し点灯を意味するものであれば SSR に繋がった GPIO ピンの出力を High に、消灯を意味するものであれば Low に変化させる。SSR は電球に電力を供給する線を制御し、コンピュータから High の入力があるときだけ電球を点灯する。以上のように視線検知とアイコンタクト、さらに音声コマンドによる操作とライトの制御を実現している。

### 3.3 デモ展示の実施

2019 年 11 月 22 日から 23 日にかけて開催された慶應義塾大学 SFC Open Research Forum 2019(ORF)において LuuX のデモ展示を行った。ORF には幅広い分野の研究者、主に高校から大学の学生、企業関係者などが来場し、筆者らが開発した LuuX も実際に使用してもらった。使用した来場者からは以下のようなコメントがあった。

- 見て話すだけでいいのは便利そう
- 顔を向けると反応(インジケータが光る)してくれるのがかわいい
- これまでのスマート家電は使いこなせていなかった  
これまでのスマート家電の使用感と比べ、便利そうだという印象を受けた来場者が多くいた。また、顔を向けると反応する点に親しみを覚えた来場者もあり、人から機械への一方向的な制御よりも心理的な抵抗が下がると考えられる。

## 4. 展望

本研究では、家電の音声コマンド操作における視線を用いた機器選択にアイコンタクトを導入することで対話的

でストレスのない操作を実現した。音声コマンド操作時には操作対象を識別する必要があるが、名前での識別には限界がある。そこで視線による選択の研究が進められているが、どの機器が選択されているか、あるいは選択できているかどうかの確認ができないことから、誤作動が多く実用に至っていない。そこで、ユーザの視線を認識すると機器のインジケータが点灯し、音声コマンド制御が可能であるということをユーザに伝える「アイコンタクト」を提案した。この仕組みは「家電の側からユーザに操作を促す」という新たな UI のイディオムを生む可能性がある。

この手法を用いてライトの点灯と消灯を制御できるようにしたスマートライト LuuX を開発した。このデバイスを ORF にてデモ展示し、来場者に使ってもらったところ、利便性や親しみやすさの点で好評価を得た。

デバイスが近くに複数ある場合にアイコンタクトを用いると、どのデバイスが選択されているか把握するために特に有効であると考えられる。これをより効果的にするために、デバイス同士で通信し同時に複数のシステムが選択されないように調整する仕組みを新たに開発する必要がある。

**謝辞** 本研究の一部は孫正義育英財団の助成により実施されました。

### 参考文献

- [1] Wilson, A. and Shafer, S.: *XWand: UI for intelligent spaces*, Conference on Human Factors in Computing Systems (2003).
- [2] Obaid, T., Rashed, H., Abou-ElNour, A., et al.: *Zigbee Based Voice Controlled Wireless Smart Home System*, International Journal of Wireless & Mobile Networks (2014).
- [3] Brumitt B. and Cadiz, J.: “*Let There Be Light*”: *Examining Interfaces for Homes of the Future*, INTERACT (2001).
- [4] Kim, J., Choi, S. and Jeong, J.: *Watch & Do: A Smart IoT Interaction System with Object Detection and Gaze Estimation*, IEEE Transactions on Consumer Electronics (2019).
- [5] Smith, B. A., Yin, Q., Feiner, S. K., Nayar, S. K.: *Gaze Locking: Passive Eye Contact Detection for Human-Object Interaction*, ACM User Interface Software and Technology (2013).
- [6] Zhang, X., Sugano, Y. and Bulling, A.: *Everyday Eye Contact Detection Using Unsupervised Gaze Target Discovery*, ACM User Interface Software and Technology (2017).
- [7] Cooper, A., Reimann, R., Cronin, D., et al.: *About Face: The Essentials of Interaction Design*, John Wiley & Sons (1995).
- [8] Google: TensorFlow (online), 入手先 (<https://www.tensorflow.org>), (2019.12.20)
- [9] Julius development team: 大語彙連続音声認識エンジン Julius (online), 入手先 (<https://julius.osdn.jp>), (2019.12.20)
- [10] Ingber, D. E.: *Tensegrity: the architectural basis of cellular mechanotransduction*, Annual review of physiology (1997).