

# 手の動きに追従する UI における タップジェスチャの認識方法の研究

田村 律起<sup>†</sup> 野田 雄希<sup>‡</sup> 水谷 晃三<sup>†‡</sup>

帝京大学工学部情報電子工学科<sup>†</sup> 帝京大学大学院理工学研究科<sup>‡</sup>

## 1. はじめに

センサによりユーザの手を認識し、その動きに追従するようにプロジェクタからユーザインタフェース (UI) を投影する研究が行われている。例えば Harrison らの OmniTouch[1]では、ユーザの肩にセンサとプロジェクタを装着し、手のひらや腕に UI を投影した。筆者らは、天井にセンサやプロジェクタを設置し、ユーザの手のひらに UI を投影・操作するシステムの実現方法を研究している。天井に機器を設置することで、端末を携帯したり機器を装着したりせずに UI を操作できるコンピューティング環境の実現を目指している。本研究では筆者らの試作システムにおいてタップジェスチャを認識するための方法を検討する。

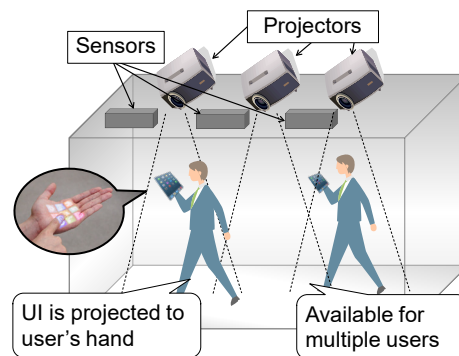


図1 システムの構成

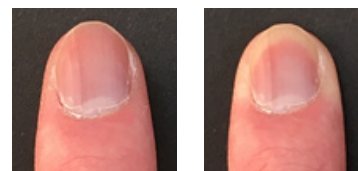


図2 タップ時の爪の色の変化  
(左 非タップ時, 右 タップ時)

## 2. 手の動きに追従する UI

筆者らは、天井に下方へ向けて設置した RGB-D センサを用いて手を認識し、認識した手の位置にプロジェクタを用いて UI を投影するシステムの実現方法を提案している[2]。図1はシステムの配置例である。UI は手の動きに追従しながら投影されるため、筆者らはこのような UI を Followable User Interface (FUI) と呼んでいる。現状、FUI の試作システムでは一方の手に投影された UI を他方の手の指でタップして操作するような仕組みを実現できていない。天井に配置したセンサと手のひらの間には距離があり、指先の数 cm の動きの変化を正確に捉えることが課題となっている。そこで本研究では、タップ時の指先の色変化を用いてタップジェスチャを認識する方法を検討する。

## 3. 方法

### 3.1. タップ認識手法の検討

Sun らは、平面上でのタップ認識方法として、タップ時の爪や指先の色の変化によりタップを

A Study of a Tap Gesture Recognition Method for Followable User Interface

Ritsuki Tamura<sup>†</sup>, Yuki Noda<sup>‡</sup>, Kozo Mizutani<sup>†‡</sup>

<sup>†</sup>Department of Information and Electronic Engineering, Faculty of Science and Engineering, Teikyo University

<sup>‡</sup>Graduate School of Science and Engineering, Teikyo University.

認識する EigenNail 法を提案した[3]。また、小川らは、限定的な条件下でのみの認識を可能としていた EigenNail 法の制限を緩め、より汎用的な場面での認識を可能とするための方法を提案した[4]。これらの研究では、硬さのある面をタップすることを主に想定している。図2はタップ時の指先の色の変化の様子である。色の変化を用いることでタップの瞬間がわかりやすく、カメラとの距離に影響されにくい正確なタップ認識が期待できる。しかしながら、本研究の場合は手のひらという柔らかい面へのタップが想定される。そこで、これら既存研究を参考に、RGB センサによって手のひらと指の接触時の指先の色の変化を検出し、タップを認識することができるかどうかを検証することにした。

### 3.2. 認識方法

本研究では、EigenNail 法において SVM を用いる方法[4]と CNN (Convolutional Neural Network) を用いる方法を検討、実装した。SVM による方法では、まず画像から手のひら上の指を切り出し、照明などの光の影響を抑えるために輝度による正規化処理を行う。次に、これらの処理を行った画像データセットに対し、主成分分析による次元削減を行う。これにより画像の情報量

を削減するとともに特徴を抽出し、認識を行いやすくする。その後、得られた特徴量を用いてクラスタリングを行い、データセットを2つのクラスタに分類する。それぞれのクラスタに対しSVMを用いて識別機を生成し、これを用いてタップを判定する。CNNを用いる手法では、SVMを用いる手法と同様に、指先の切り出し、正規化処理を行った後、認識モデル学習を行う。認識モデルにはEfficientNet[5]を用いた。

#### 4. 実験および結果

前述の方法によるタップ認識の評価実験を行う。実験に用いるデータセットの画像は、試作システムで用いているRGB-Dセンサと画角がほぼ同じであるiPhone XRのRGBカメラを用いて、天井に設置した場合を想定して約1.2mの距離から3840×2160画素で動画撮影したものをを用いた。図3(左)に示すように、被験者はカメラの真下で左手を手のひらを上にして構え、右手の人差し指を用いて手のひら上をタップする。その様子を撮影し、タップ時、非タップ時の画像をそれぞれ切り出す(図3右)。撮影は3人の被験者(20代および40代男性)に対して行い、被験者ごとにタップ時、非タップ時の画像をそれぞれ30枚ずつ、計180枚の画像を用いてデータセットを作成した。2名分のデータを用いて学習を行い、残りの1名分のデータを使って評価した。

表1にSVMを用いた方法およびCNNを用いた方法による結果を示す。それぞれの値は、3人分を入れ替えながら学習と認識を行った結果の平均値である。例えばタップ状態の画像をタップと認識した画像数はSVMでは23.3、CNNでは28.7となった。この結果、SVMの正解率は86.1%、適合率は93.3%となった。CNNの正解率は94.5%、適合率は93.5%となった。

#### 5. 考察

実験の結果から、どちらの方法でも、手のひら上である程度の精度でタップ状態を認識できることがわかった。CNNの方が高精度な結果となったが、認識時の計算負荷が大きいためリアルタイムでの処理を踏まえるとどちらが最適であるかは検討の必要がある。また、学習に用いた被験者のデータの組み合わせによっては精度が低くなることもあった。これは、爪の色やタップ時の変化に個人差があるためだと考えられる。本実験では被験者が3人と少なかったため、実用的な認識精度を得るには、より多くのデータを集めた認識モデルの作成が必要である。

また、今回の実験では指の向きやカメラとの



図3 タップ時の様子(左)と指先画像(右)の例

表1 モデル予測結果(SVM/CNN)

		認識結果	
		非タップ	タップ
正解	非タップ	28.3 / 28.0	1.7 / 2.0
	タップ	6.7 / 1.3	23.3 / 28.7

距離は一定としたが、FUIではUIを投影する手の高さや向きはさまざまである。また、UIを投影している中での認識になるため、その影響についても考慮する必要がある。これらを踏まえたタップ判定方法の検討が必要である。

#### 6. おわりに

本研究では、手に追従するUIにおけるタップジェスチャの新たな認識方法として、指の色の変化を用いた方法について検討し、評価を行った。今後は課題の解決や精度の向上を行い、実用的なシステムの実現を目指す。

#### 謝辞

本研究の一部はJSPS科研費21K12163の助成を受けた。

#### 参考文献

- [1] Harrison, Chris and Benko, Hrvoje and Wilson, Andrew D.. OmniTouch: Wearable Multitouch Interaction Everywhere, Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, pp.441-450, 2011.
- [2] 山崎雄太, 大塚友章, 高笠綾華, 水谷晃三, 荒井正之. 複数人の利用者の手の動きに追従するユーザインタフェースの研究, 情報処理学会第78回全国大会, 6Y-03, 2016.
- [3] Y. Sun, J. M. Hollerbach and S. A. Mascaró. EigenNail for Finger Force Direction Recognition, Proceedings 2007 IEEE International Conference on Robotics and Automation, 2007, pp.3251-3256.
- [4] 小川修平, 杉田尚基, 岩井大輔, 佐藤宏介: EigenNail法における複数識別器による指先押下認識の改善とそのメニュー選択インタフェース応用に向けた基礎検討, 映像情報メディア学会誌, Vol.67, No.12, pp.J425-J433, 2013.
- [5] Tan, Mingxing and Le, Quoc. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, Proceedings of the 36th International Conference on Machine Learning, Vol.97, pp.6105-6114, 2019.