

# 手の動きに追従する UI における 手の認識とタグ付けによる判別の研究

岩崎真大<sup>†</sup> 松岡雄介<sup>†</sup> 土屋博雅<sup>†</sup> 水谷晃三<sup>†</sup> 荒井正之<sup>†</sup>  
帝京大学工学部ヒューマン情報システム学科<sup>†</sup>

## 1. はじめに

センサにより利用者の手の動きを認識し、その動きに追従するようにプロジェクタから UI (User Interface) を投影する研究が行われている。例えば奥村らの Lumipen ではセンサによって手の動きを捉え、プロジェクタの投影軸をミラーによって制御することで手の動きに追従する UI が実現されている[1]。しかし、複数人の手の動きに同時に対応する方法は議論されていない。

複数人に対して手の位置を認識して UI を投影するために、筆者らは手の形を機械学習させた物体検出器を用いる方法を検討した[2]。複数人の手に対して位置を認識することができたが、それは手の形をしたものが認識されているだけであり、深度センサから取得したフレームの前後で同じ手として認識できていない。それぞれの手に UI を継続的に投影するためには、この問題を解決する必要がある。本研究では、認識された手の情報を利用して前後のフレーム間で同じ手を連続的に判別する方法を検討する。

## 2. システム概要

図 1 に先行研究のシステムの配置を示す。上方にプロジェクタと深度センサを設置する。プロジェクタの投影範囲をユーザの移動可能範囲に限定する。深度センサを用いて投影範囲内の手のひらの位置を認識し、UI はプロジェクタ出力の一部分を使用して手のひらに投影する。

先行研究では、深度センサから取得したデータに対して手の高さとなる領域の閾値を設け、その範囲内をグレースケールで描画する。これを深度画像と呼び、解像度は 512 pixel×424pixel である。深度画像に対して手の形を学習させた OpenCV[3]のカスケード分類器を用いて画像処理を行うことにより手の領域を検出し、これをもとに投影する UI の位置を決定する。

### A Study of Hands Recognition and Distinction by Attachment of Tags in Followable User Interface.

<sup>†</sup>Iwasaki Masahiro, Matsuoka Yusuke, Tsuchiya Hiromasa, Mizutani Kozo, Arai Masayuki : Department of Human Information Systems, Faculty of Science and Engineering, Teikyo University.

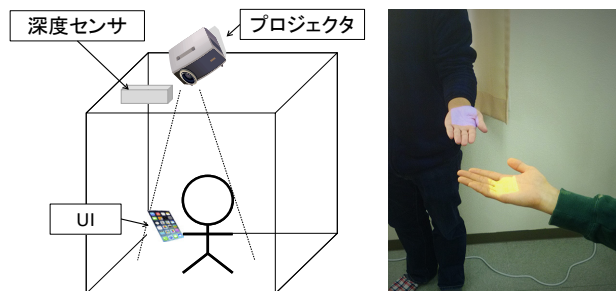


図 1 先行研究のシステムの配置 (左) と簡易的な UI の投影例 (右) [2]

## 3. フレーム間での連続的な手の判別方法

### 3.1 概要

認識された手をフレームの前後で同じユーザの手だと連続的に判別するために、カスケード分類器で検出された手の領域ごとに仮のタグをつける。タグには、タグの名前、物体検出により得られた手の領域の位置と大きさ、手の高さとなる領域内の手の面積、手の検出率、深度センサのデータなどを属性値に含める。

タグを管理するためのリストを設け、リストが空であるときは仮のタグをリストに追加する。すでにリストにタグが格納されている場合は仮のタグとリストに格納されているタグを比較していき、類似するものがあればリスト内のタグの属性値を仮のタグのものに更新する。この際、タグ名など一部の属性値は更新せずに残す。同じ手の検出領域に同じ名前のタグが与え続けられれば、フレームの前後で連続的に手を判別できていることになる。

また、生成されたタグが使われなくなった、もしくは手のない領域の誤検出により不適切なタグが生成された場合に対しては、タグを消去する機構も設けることで判別精度を高める。

### 3.2 タグの類似判定

仮のタグとリストで管理されているタグの類似判定は以下の 2 方式を組み合わせる。

#### ① 位置情報による判定

タグに含まれる位置情報を比較し、ある一定範囲内で位置が同じであることを類似するタグであると判別する。

## ② 深度値を活用した手形状による判定

①の手法では手が素早く移動した際に、同じ手として扱うべきタグを適切に判別できない。そこで、深度値を活用した手形状による判定方法を導入する。

図2に示すように、検出された手の領域の重心Cから同心円状に深度値をサンプリングする。図中では3つのサンプリングを行う例を示しており、1周目は半径を $r$ とし、2周目以降は間隔 $s$ となるようにして複数のサンプルを取得する。各サンプルは360度分の深度値からなるが、一番外側の円の深度分布を用いて手首がある角度を取得し、この角度を起点としたサンプルとなるように調整する。さらに、これらのサンプルのデータを20度ごとのブロックに分けて平均値を算出し、そのブロックごとの値の差の平方和が閾値 $t$ 以下のものを類似するタグであると判定する。

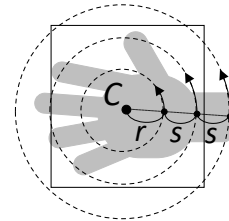


図2 手形状取得のイメージ

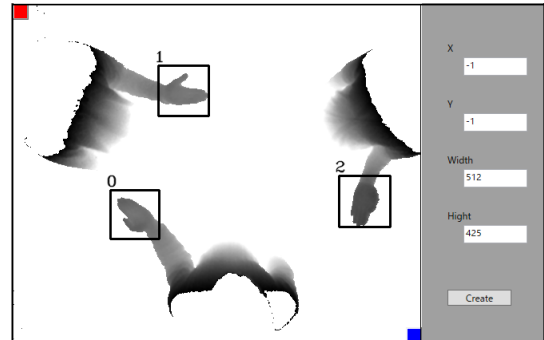


図3 評価実験の試作システム画面例 (一部)

## 4. システムの実装と類似判定の評価

先行研究のシステムに3で述べた機構を実装した。図3にシステムの実行画面の例を示す。3人が同時に手を左右に移動させた場合のようすを表しており、深度画像内の手には物体検出で得られた手の領域の矩形が描画されている。矩形の左上にはタグ名が表示されており、同じタグ名が継続的に付与されることを確認した。

タグの類似判定について、手形状による判定の効果を実験により評価する。ユーザが手を移動させたとき、タグの類似判定が適切に行われていれば同じ手に同じタグ名が与えられ続ける。類似判定が不適切な場合は誤判定となり、同じ手であっても新たなタグ名が与えられてしまう。

実験では、被験者が手を左右または前後方向に2往復分移動させる。その際の動作の速さは、8秒かけて行う場合をL、4秒かけて行う場合をHとする。3人の被験者により、1人ずつ行う場合と3人同時に行う場合で実験を行い、3.2で述べた判定方法の①、①と②の組み合わせについて誤判定回数（および全フレーム中の発生率）を計測した。なお、①における位置の範囲はx軸またはy軸方向に $\pm 30$ pixel、②においてはサンプリング数=3、 $r=15$ pixel、 $s=10$ pixel、 $t=200$ とした。

結果を表1に示す。人数が1人のときの結果については3人分の実験結果の平均値を示している。

## 5. 考察

実験の結果、①による方法のみのときは6.5%以下の誤判定の発生率になった。②の方法の組み合わせにより類似判定の精度の改善効果も確認できた。しかしながら、表中☆で示したケースのように改善効果が得られない場合もあった。その原因としてカスケード分類器により手が検

表1 タグの類似判定実験の結果

人数	移動方向	速さ	誤判定回数 (発生率)		改善度 ((A-B)/A)
			①のみ(A)	①+②(B)	
1	左右	H	2.7 (2.2%)	1.0 (0.9%)	63.0%
1	左右	L	2.0 (0.9%)	1.7 (0.8%)	15.0%
1	前後	H	0.0 (0.0%)	0.0 (0.0%)	0.0%
1	前後	L	0.0 (0.0%)	0.0 (0.0%)	0.0%
3	左右	H	9.0 (6.5%)	8.0 (5.8%)	11.1%
3	左右	L	8.0 (3.3%)	7.0 (2.9%)	12.5%
3	前後	H	2.0 (1.5%)	2.0 (1.5%)	0.0% <sup>☆</sup>
3	前後	L	1.0 (0.4%)	1.0 (0.4%)	0.0%

出されていないフレームが長く続いてしまい、比較するタグの属性値の差が大きくなってしまふことが考えられる。分類器の検出精度を向上させることで、この問題を解決することができると考えられる。

## 6. おわりに

本研究では手の動きに追従するUIにおいてタグ付けによる手の判別方法を検討した。フレーム間で同じ手を連続的に判別するためにタグに含まれる位置情報と手形状の比較による方法を検討し、その効果について述べた。

## 参考文献

- [1] 奥村光平, 奥寛雅, 石川正俊: 高速光軸制御を用いた動的物体への投影型拡張現実感, 東京大学 大学院情報理工学系研究科, Vol.67, No.7, J204-J211, 2013.
- [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-3, 2016.
- [3] OpenCV, <http://opencv.org/>