

# 卒業研究発表

## 顔と手の動きを組み合わせた

## 個人識別の評価

---

67150040 池田 拓矢

指導教員 椋木 雅之

平成31年2月15日

# 研究背景

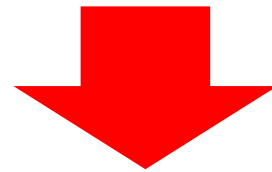
---

近年、生体情報を用いて個人認証を行うことが増加

ex) iphoneの指紋認証

銀行ATMの静脈認証 など

一つの生体情報での個人認証は、偽装や改ざんによるすり抜けの恐れ



複数の生体情報を用いて認証(マルチモーダル認証)を行うことが注目されている

# マルチモーダル識別

---

## 顔と手の動きを用いた個人識別を扱う

顔 → 指紋や虹彩などに比べ心理的抵抗が少ない

手の動き → 個人の癖が出やすく、  
顔の情報が分かりにくい状態でも取得しやすい



これらを組み合わせることで  
個人識別の**精度を上げる**ことができる

# 従来のマルチモーダル識別

---

**スコアレベルでの統合** (従来のマルチモーダル識別)

それぞれの相違度を統計学的に統合



**特徴量レベルでの統合**

統合しにくいいため、あまり行われていない

# 特徴量抽出の技術の進化

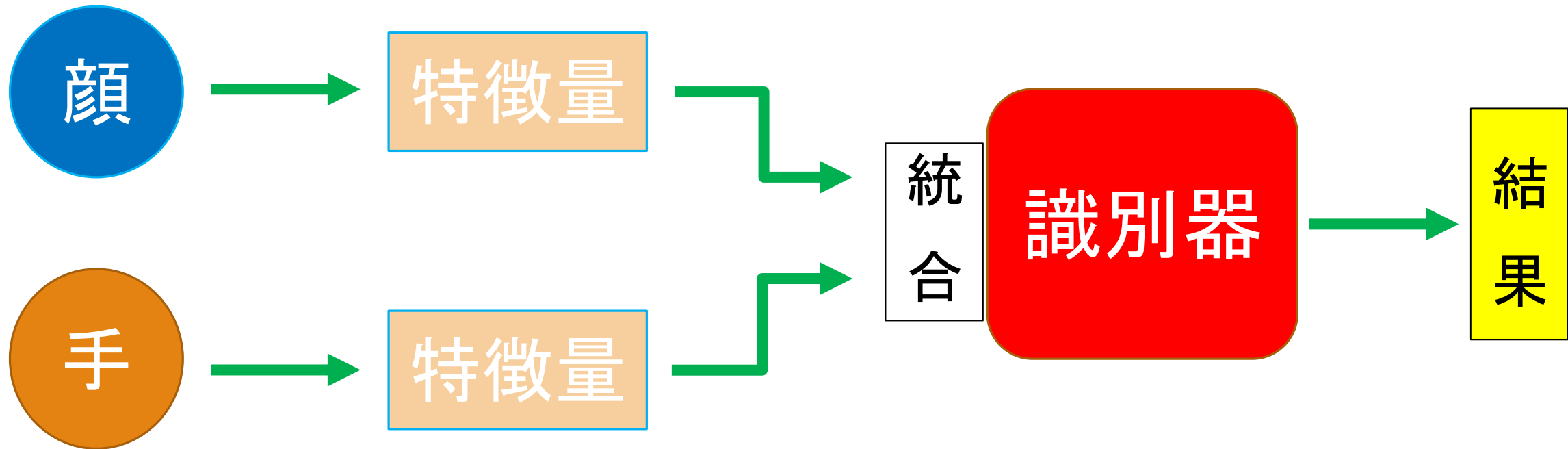
---

近年、深層学習の進化により、特徴量の優劣が重要であることが明らかになってきている



特徴量に適切に優劣をつければ、特徴量レベルで統合しても十分な結果が得られる可能性がある

# 提案手法



顔と手の動きのマルチモーダル情報を  
特徴量レベルで統合して個人識別に利用する手法を評価

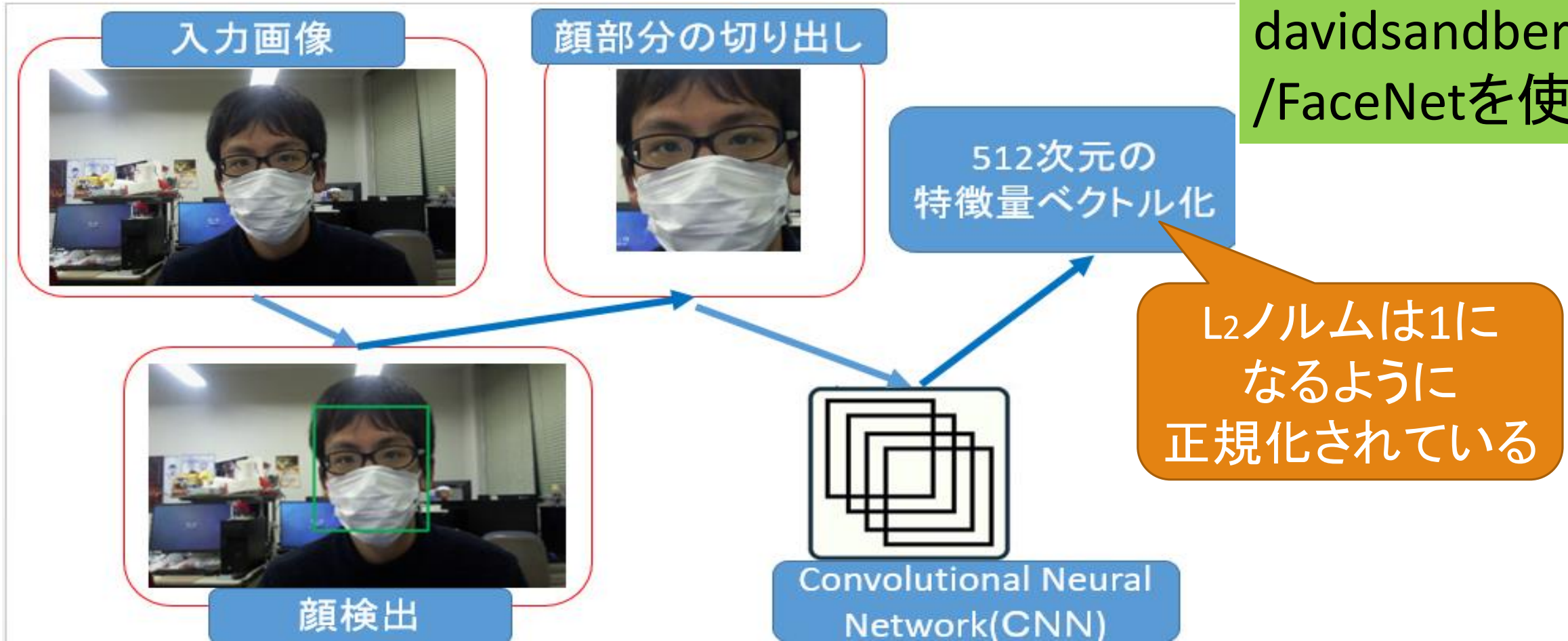
# 特徴量レベルでの統合評価の手順

---

1. 顔の特徴量の取得
2. 手の動きの特徴量の取得
3. 特徴量の統合
4. 識別器による判定

# 顔の特徴量の取得の流れ

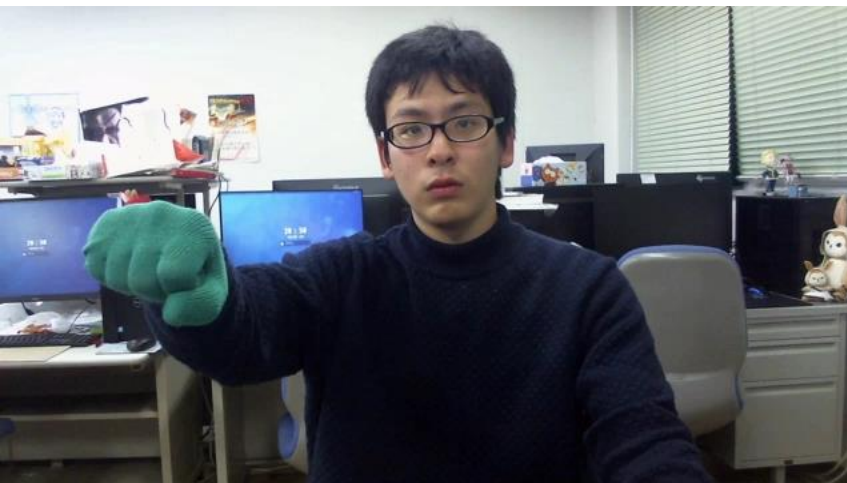
特徴抽出器  
として顔認識用  
ライブラリ  
davidsandberg  
/FaceNetを使用





# 手の動きの特徴量の取得

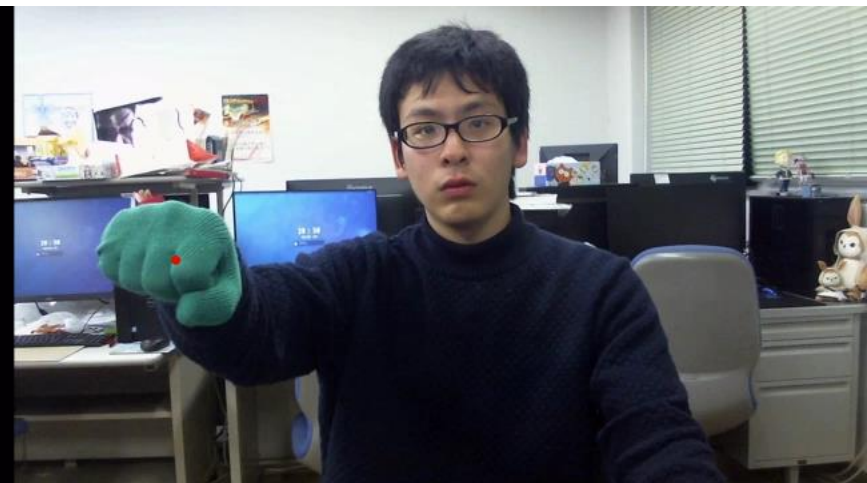
---



三角形を描く



緑色の軍手の領域を  
二値化



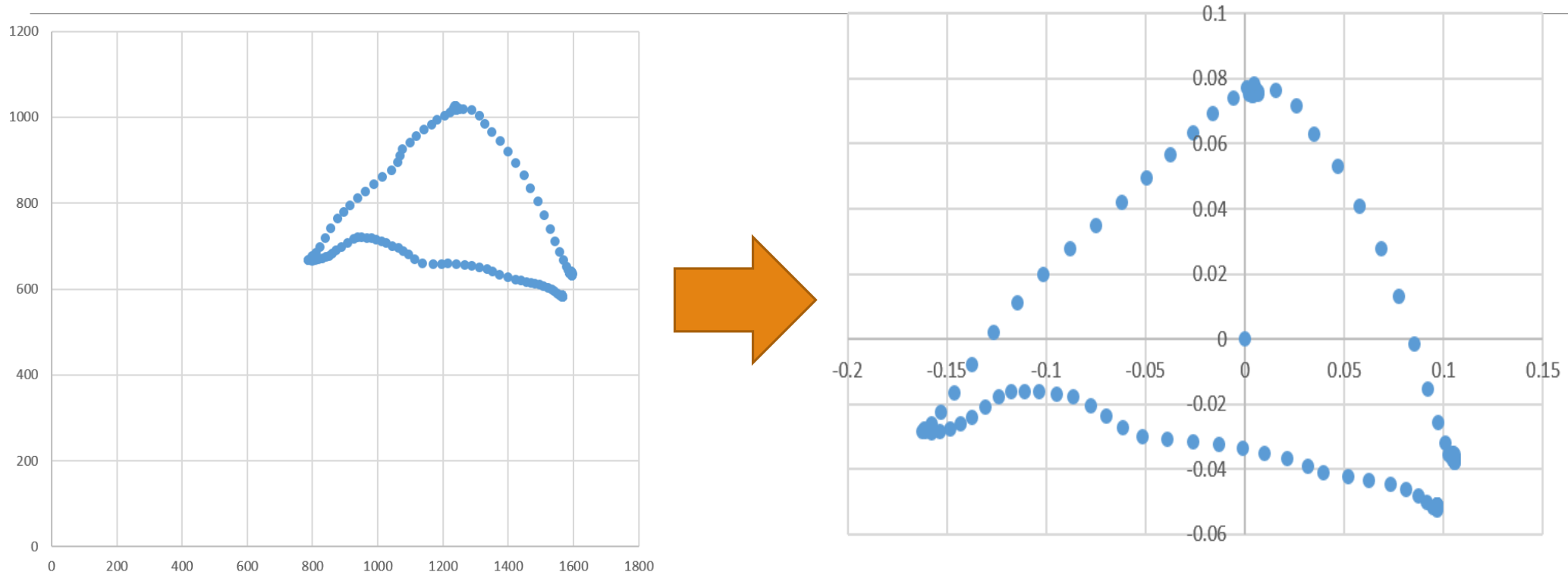
重心座標を取得

# 特徴量の正規化

---

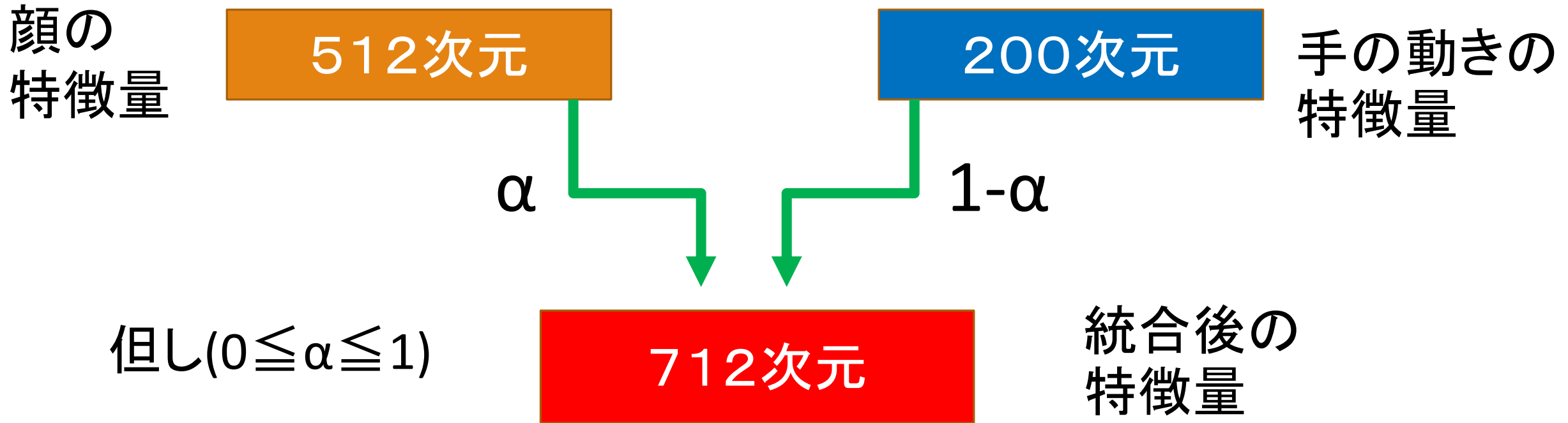
- a. 手の動きの座標の時間合わせ
- b. 手の動きの座標の位置合わせ
- c. ベクトルの $L_2$  ノルムの正規化

# 特徴量の正規化



こうしてできた100点の座標列を  
200次元の手の動きの特徴量とする

# 特徴量の統合



※ $\alpha=0$ のときは手の動きの特徴量ベクトルのみ、

$\alpha=1$ のときは顔の特徴量ベクトルのみでの識別を意味する

# 識別器による判定

---

使用する識別器: Collaborative Mean Attraction(CMA)法

CMA法とは:

未知のテストデータを既知のカテゴリのどれかに分類

どうやって?

各カテゴリの評価値を求め、最小となったカテゴリに分類

# 実験

---

- 顔と手の動きの特徴量の抽出
- 特徴量に重みづけて統合
- 統合した特徴量で識別
- 重みを変えたらどのように変化するか実験

# 実験データ

## 顔画像と手の動きの動画を使用

被験者数: 4人

### 【顔画像】

学習用: なにも装着していない画像

テスト用: サングラス画像、マスク画像

### 【手の動きの動画】

三角形を描く

種類	学習用	テスト用
顔画像	40(枚) × 3(異なる場所で撮影) × 4(人) = 480枚 (1920 × 1080ピクセル)	50(枚) × 3(異なる場所で撮影) × 4(人) = 600枚 (サングラス・マスクそれぞれについて用意)
手の動きの動画	2本(一人当たり) × 4(人) (30fps, 1920 × 1080ピクセル)	8本(一人当たり) × 4(人)

# 実験データ(学習用・顔)なにも装着していない



各人120枚 × 4

合計480枚



# 実験データ(テスト用・顔)サングラス装着



150枚 × 4

合計600枚

# 実験データ(テスト用・顔)マスク装着



各人150×4

合計600枚

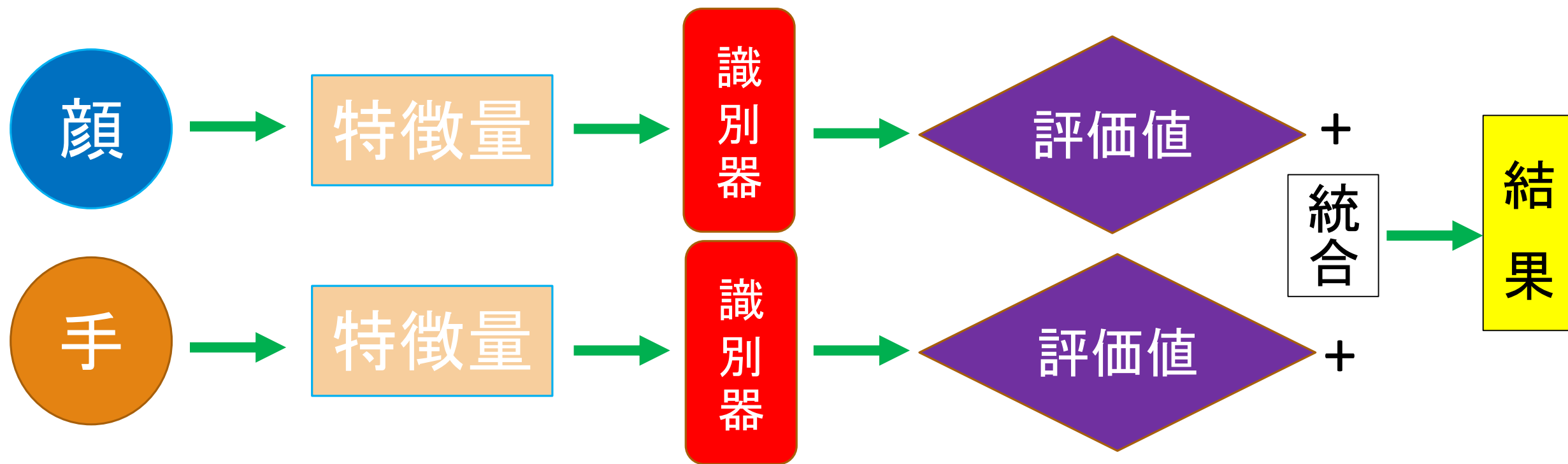
# 実験データ(テスト用・学習用)・手の動き



学習用:  
 $2 \times 4 = 8$ 本、

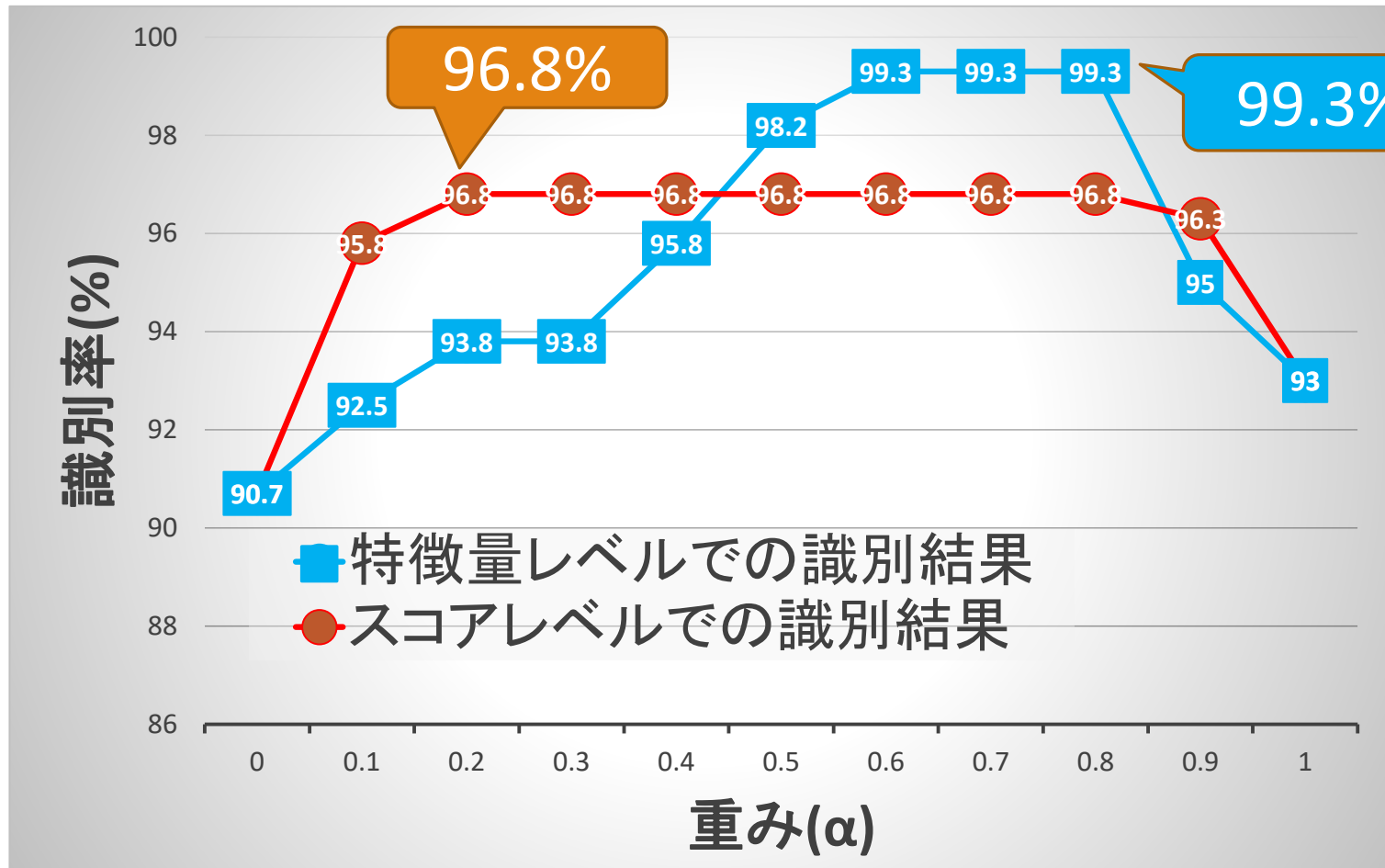
テスト用:  
 $8 \times 4 = 32$ 本

# 比較実験



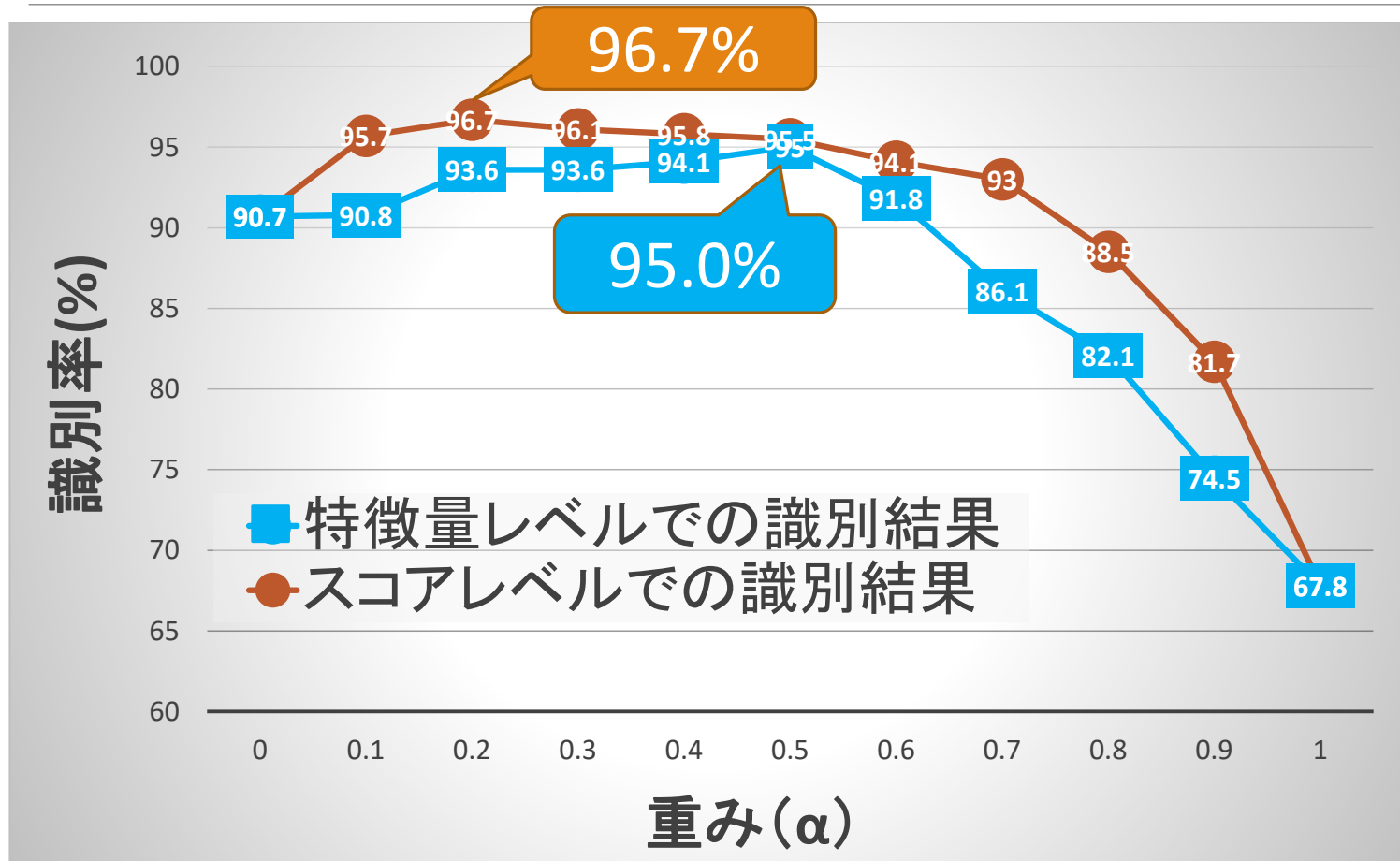
$$d_{int}^i = \alpha d_{face}^i + (1 - \alpha) d_{hand}^i (0 \leq \alpha \leq 1)$$

# 実験結果(サングラス画像)



特徴量レベルで  
統合したほうが  
2.5ポイント良い

# 実験結果(マスク画像)



スコアレベルで  
統合したほうが  
1.7ポイント良い

# 実験結果

---

- いずれの場合についても、  
単独の場合より識別率の向上がみられる
- 必ずしも特徴量レベルで統合した場合の方が  
高い識別率を得られるとは限らない

# まとめと今後の課題

---

- 顔と手の動きを特徴量レベルで統合し、識別を行い、スコアレベルでの統合と比較を行った
- 特徴量レベルで統合した方が良い場合と、スコアレベルで統合した方が良い場合があった

## 【課題】

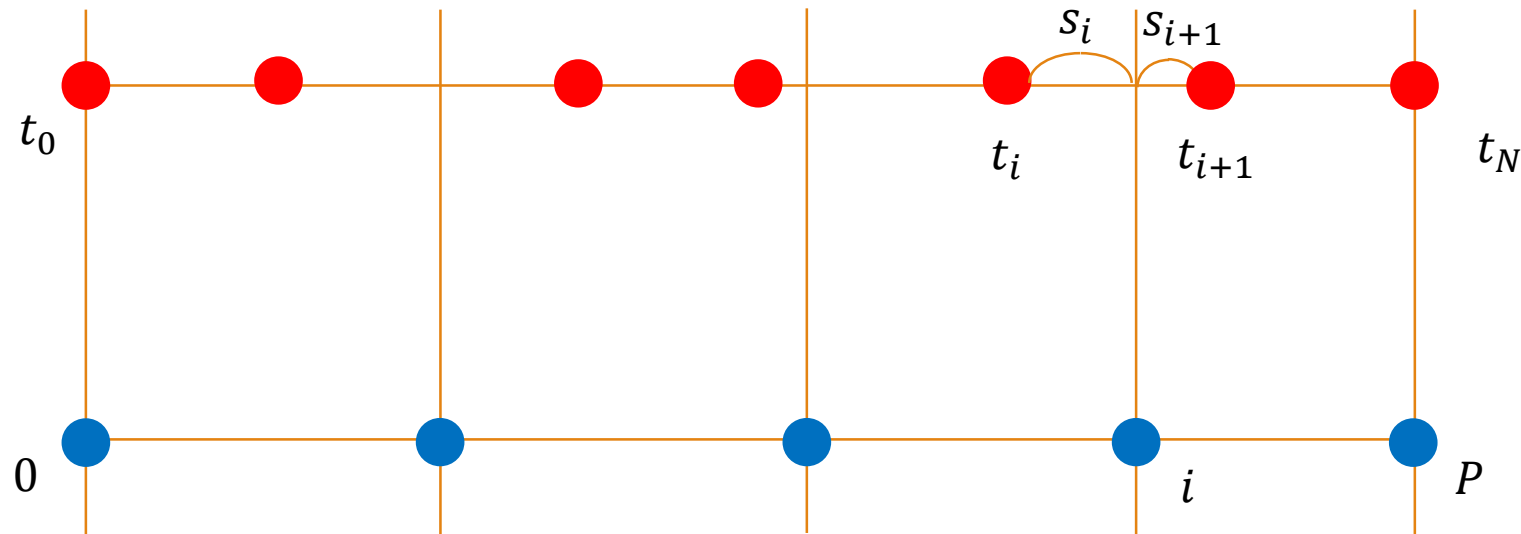
- 被験者が多い場合
- 正面画像以外のものを使った場合
- 信頼性の高い特徴量同士で組み合わせた場合



# ここから補足

---

# 手の位置の座標の時間合わせ



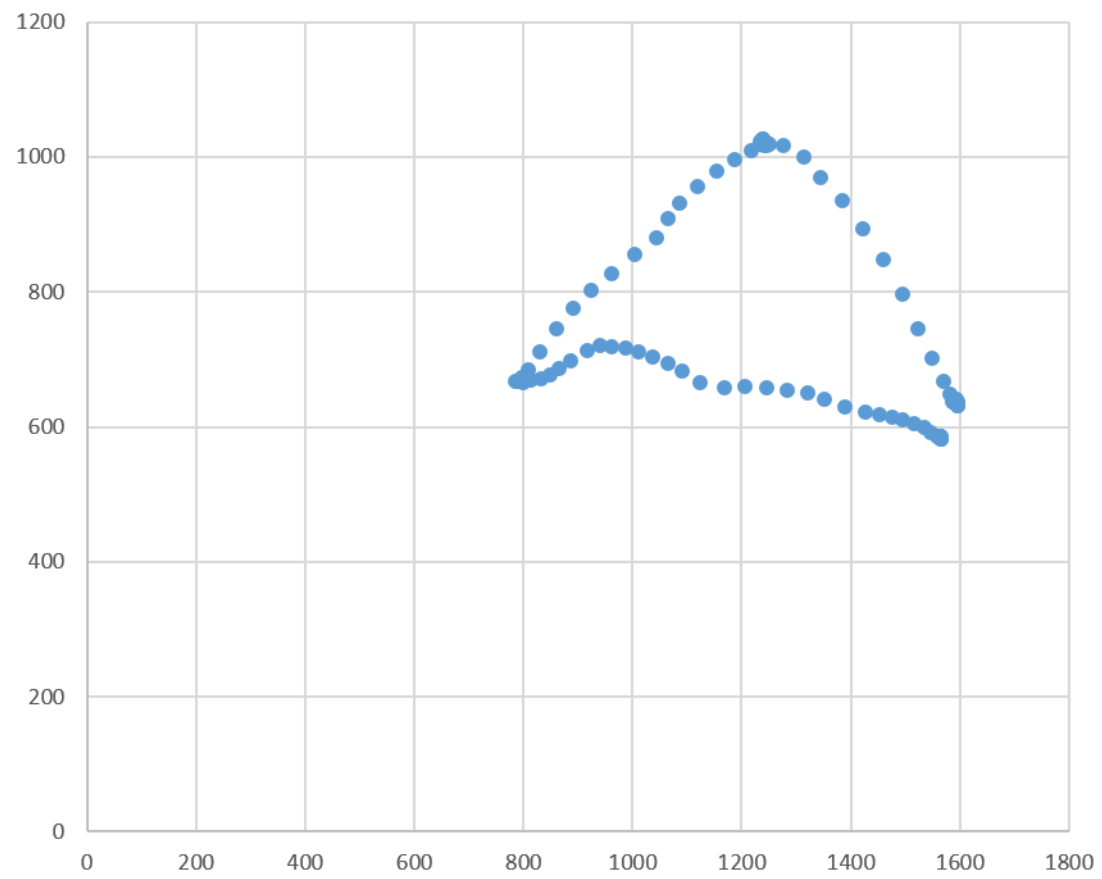
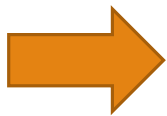
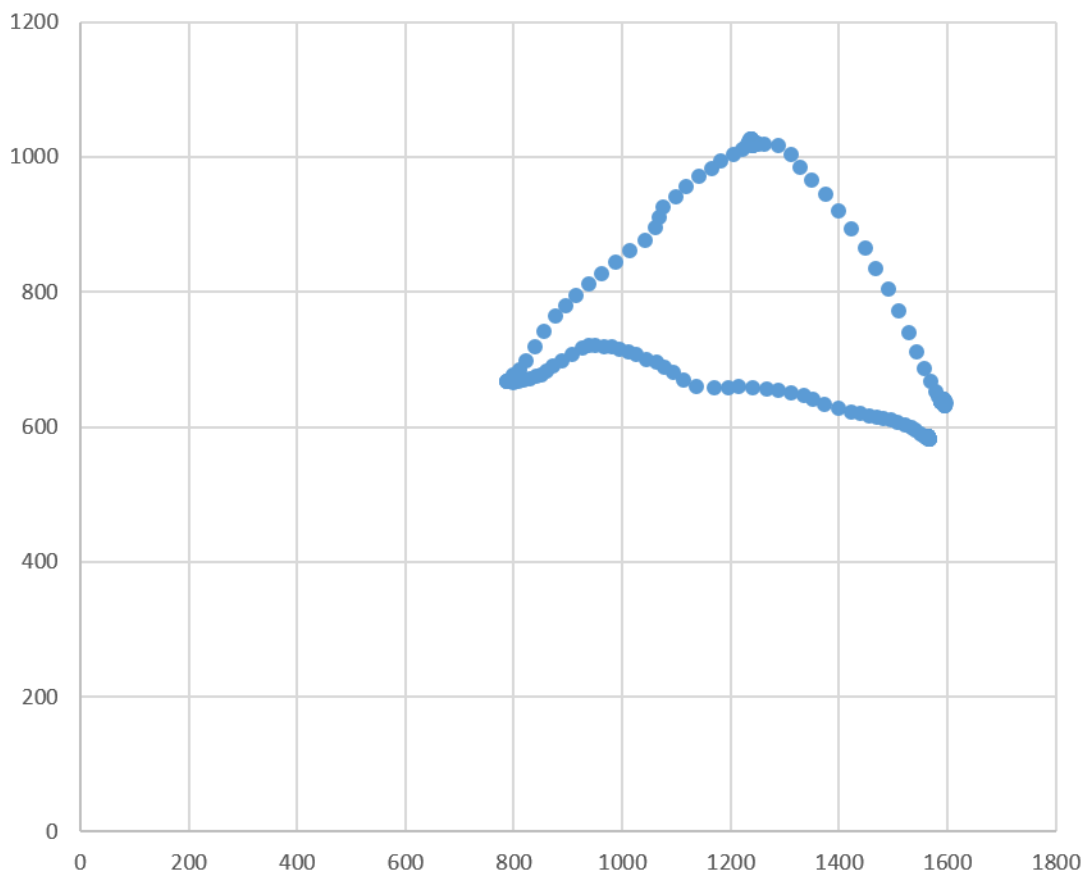
$$t_i = \left[ \frac{N}{P_i} \right] (\lceil \rceil : \text{ガウス記号})$$

$$s_i = \frac{N}{P_i} - t_i$$

$$x_i = x_{t_i}(1 - s_i) + x_{t_{i+1}}s_i$$

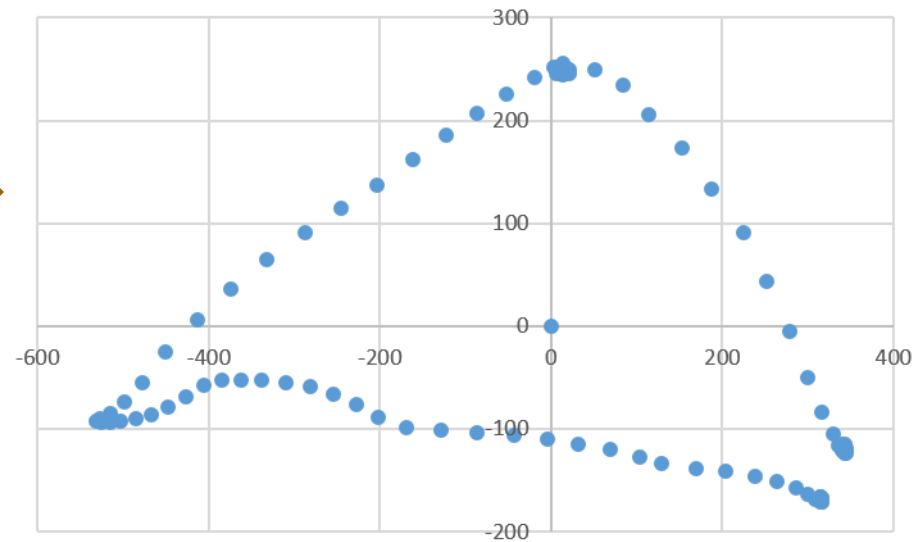
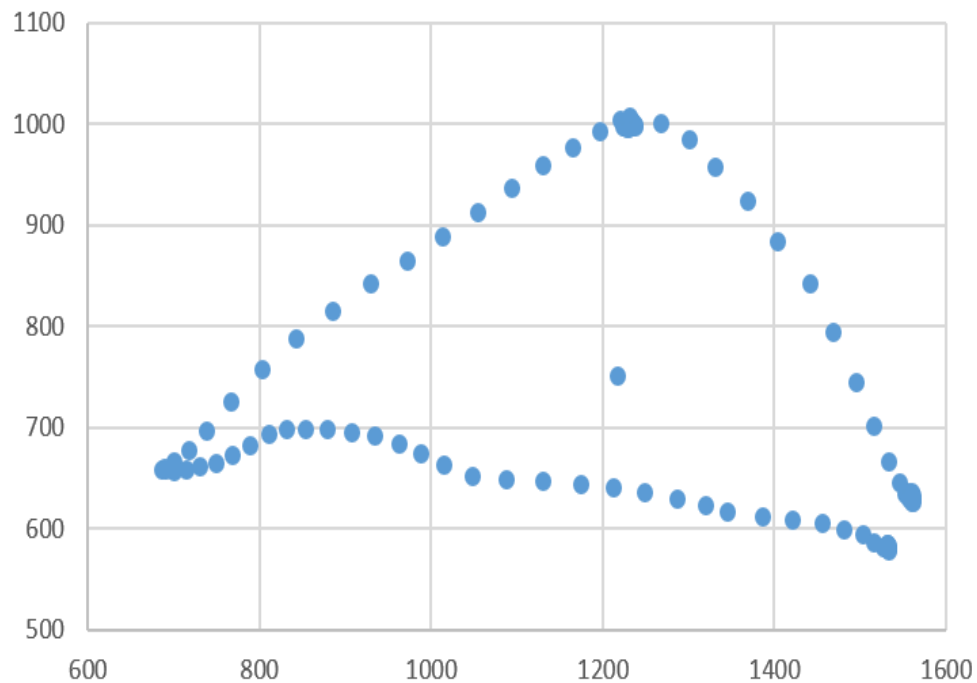
$$y_i = y_{t_i}(1 - s_i) + y_{t_{i+1}}s_i$$

本研究ではP=100



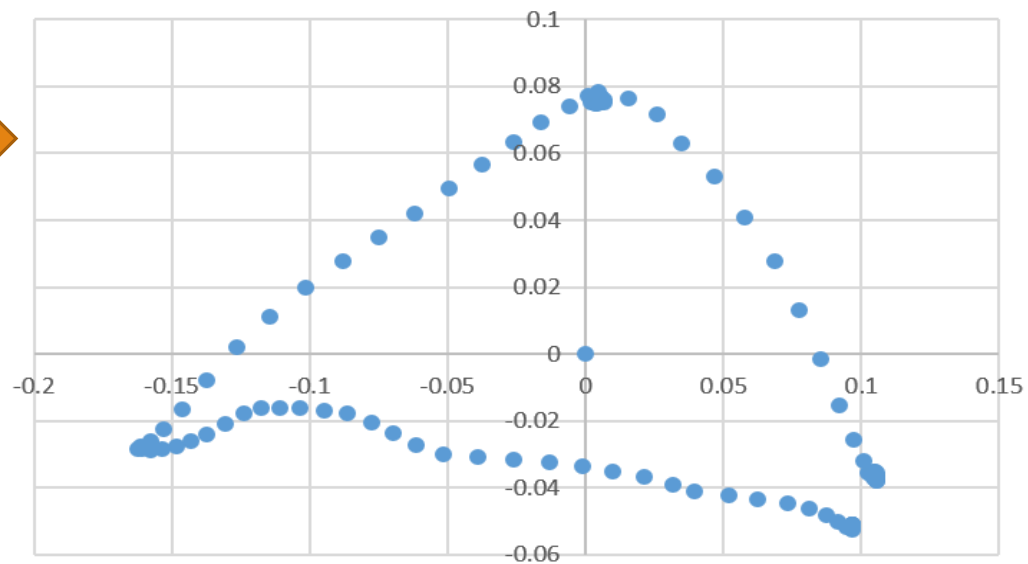
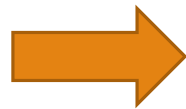
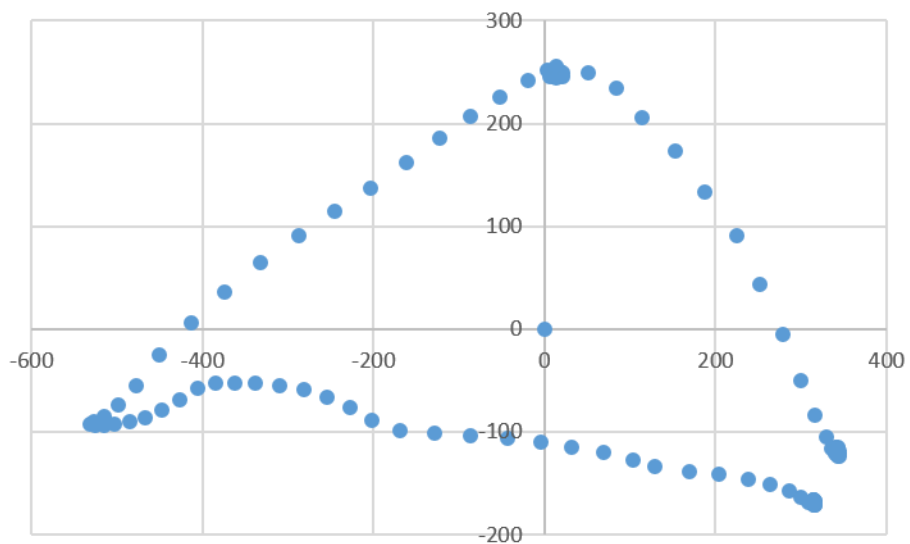
# 手の位置の座標の位置合わせ

$$\bar{x} = \frac{1}{N} \sum_{i=0}^{N-1} x_i \quad x'_i = x_i - \bar{x}$$
$$\bar{y} = \frac{1}{N} \sum_{i=0}^{N-1} y_i \quad y'_i = y_i - \bar{y}$$



# ベクトルのL<sub>2</sub>ノルムの正規化

$$f = \frac{x}{\|x\|}$$
$$\|x\| = \sqrt{\sum_{i=0}^{N-1} x_i^2 + \sum_{i=0}^{N-1} y_i^2}$$



# 比較実験

---

顔と手の動きをスコアレベルで統合した手法と比較を行う

顔の特徴量、手の動きの特徴量をそれぞれ別々に識別器に与え、それぞれの評価値の算出を行い、同カテゴリ同士で足し合わせ、最小となったカテゴリに分類する。

足し合わせの際には、どちらの評価値をどの程度重視するか、重みづけを行う

$$d_{int}^i = \alpha d_{face}^i + (1 - \alpha) d_{hand}^i (0 \leq \alpha \leq 1)$$