



# 顔画像を用いた感情認識システムの開発

## ～Emotion recognition system using facial images～

中村拓 上田純也 岡嶋克典 横浜国立大学

nakamura-taku-dh@ynu.jp, ueda-junya-pc@ynu.jp, okajima@ynu.ac.jp

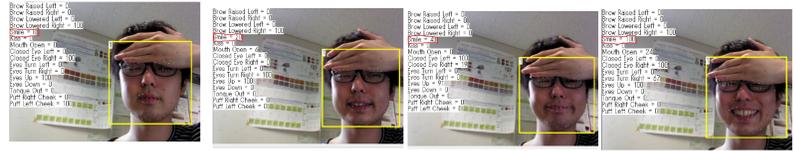
### 背景と目的

表情検出から感情推定を行う手法として、機械学習とAction Unitを用いたパターンマッチングがある。

しかし、これまでの方法では、表情変化が小さい日本人には対応出来ていないのが現状である。そこで今回、顔の静止画像からリアルセンスと機械学習を用いた日本人の表情解析システムの開発について報告する。また、ヒトの自発的な表情を解析し、その特徴を抽出することでヒトが他者の表情を認知するメカニズムについても検討した。

### 既存の顔認識システム

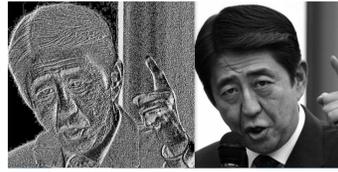
海外の顔認識サービスの結果



既存の表情検出システム(リアルセンスカメラ, Intel社製)では、パターンマッチング形式で“口の形でのみsmileと判定している”

と考えられる。上の写真では手で眉の情報を消し、全てで目をつぶることで口の形以外の情報を使わないようにしても「smile値」は0から100まで変化している。

### 顔の特徴抽出フィルタの作成



元画像出典: squallchannel.doorblog.jp

LBP特徴量で見る顔画像

→LBP特徴量は計算が早く、照明条件に強い  
ため、表情検出でも一般的に使われる。

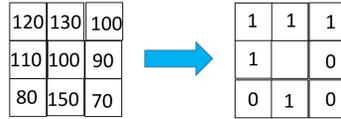
表情を認識する上で、特徴点の幾何学的関係や変位の他に使える情報はないのか？

→「しわ」

左上左の画像にメディアンフィルタを一度かけるとしわが強調される。これを鼻や頬に局所的に用いると有効なフィルターになる可能性がある。

眉間

鼻下



### 機械学習の利用

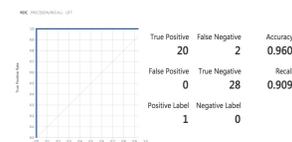
右図のように、目の中心からの8ベクトルと口の中心からの4ベクトルを両目間距離により正規化した特徴ベクトルを用いる。



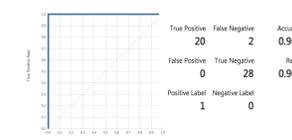
訓練画像150枚 テスト画像50枚

2クラス分類(笑顔)

Two-Class Boosted Decision Tree



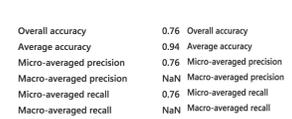
Two-Class Logistic Regression



8クラス分類

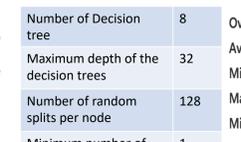
Multiclass Decision Forest(左)

Multiclass Logistic Regression(右)

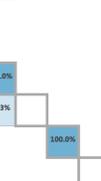
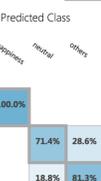
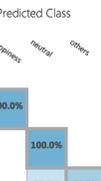
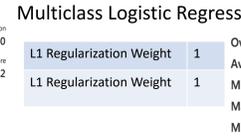


3クラス分類(笑顔・ニュートラル・それ以外)

Multiclass Decision Forest



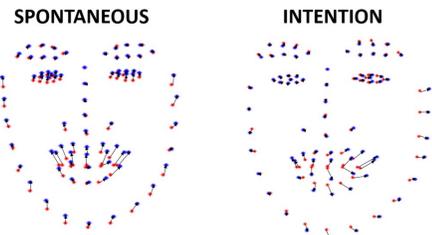
Multiclass Logistic Regression



### 自然な笑顔と故意的な笑顔

自然な笑顔と故意な笑顔の違い  
→様々な手段で被験者を自発的に笑わせ、故意に笑った顔との特徴差を考察した。

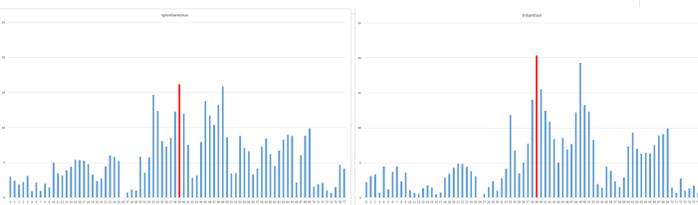
赤点: 真顔を願った顔の特徴点  
青点: 各笑顔を観測した時の特徴点



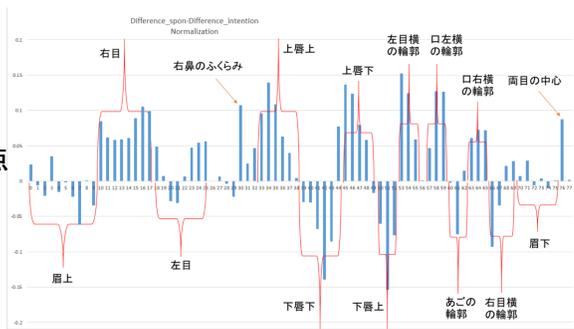
(結果)

故意的な笑顔(右グラフ)は口辺りの特徴点の位置変位量が大きく、他のパーツは小さい。

それに対して自発的な笑顔(左グラフ)は強度は右に劣るが全体的に大きく、特に目の辺りの位置変位量が大きい。



上の左右のグラフをそれぞれの特徴点番号でNo.10とNo.18の距離(目の間の距離)で割り、差(自発的-故意的)を取って分析した。



(結果と考察)

自発的な笑顔と意図的な笑顔の間で目に大きな違いが表れることが分かった。これは人の表情の認知メカニズムにおいて目から多くの情報を得るという説を支える結果となった。

この結果をふまえ、今後我々の表情検出器には口の形のみならず、目の変形の特徴を考慮した判定アルゴリズムを検討していくことで精度の向上を目指す。

### 表情の平均

ニュートラル顔とは？

→表情判定においては基準顔を設定し、各表情の中立(表情空間の原点)とすることが一般的である。

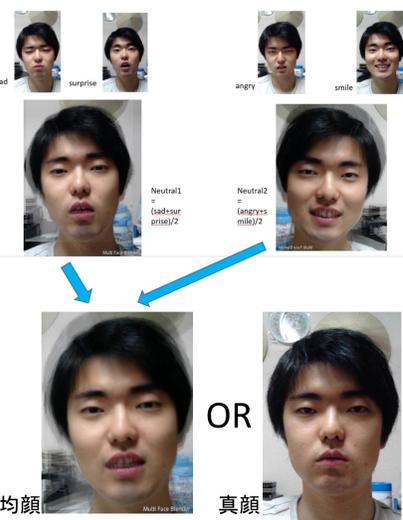
(基準顔案)

- 1, いわゆる真顔を基準顔とする
- 2, 各表情の平均した表情を基準顔とする

(考察)

2の平均顔を研究室の学生ら10人ほどにどの表情をしていると思うか尋ねた結果、1つの表情が際立つことはなかった。

平均化したことで各表情に存在する特徴が混ざり合い、どの表情と判別しにくくなったと考えられる。どの表情を基準とするか、また平均顔ならば万人の平均顔か個人のシステム実行時数フレームの平均顔にするのか検討の必要がある。



- 笑顔とニュートラルの認識率は高い
- 3クラス分類までは精度が良い
- ネガティブな微表情の識別精度は低い
- 1個人のみ画像を利用した結果である
- →どこまで汎用能力があるか
- →個人ごとの表情識別器として活用するか

### 今後の方針

- デプス画像やIR画像・ヨーピッチローの傾き情報など新しい情報を追加
- 有効な特徴量を増やし認識精度を高めると同時に、顔が正面でない場合のロバスト性向上を期待
- 深層学習の検証
- 効果的な特徴量の探索と抽出フィルタの作成
- 人間の表情認知メカニズムの解析



深度画像

IR画像