

# 適切な証明写真をレコメンドするアルゴリズムについて

宮地 美希<sup>1</sup> 北 栄輔<sup>1,2,a)</sup>

**概要:** 本研究では、ユーザーが証明写真として適切と思うものを選択するためのアルゴリズムについて述べる。電子的な顔画像データから顔特徴量を抽出する。顔特徴量を説明変数として、写真が証明写真として良い写真か否かを判別するアルゴリズムをニューラルネットワーク (NN) とサポートベクターマシン (SVM) を用いて定義する。解析例において、1名の被験者の写真をもとにニューラルネットワーク (NN) とサポートベクターマシン (SVM) を学習する。それを他の10人の被験者の写真に適用して良い写真であるかどうかを推定する精度を評価する。

**キーワード:** 写真, ニューラルネットワーク, サポートベクターマシン, Kinect.

## 1. 緒論

自動証明写真機は、証明写真の撮影手段として広く普及している。写真館での撮影に比べて安価で手軽に撮影できる利点はあるが、撮影時に被写体の様子を確認しながら撮影できないため、不自然な表情で写ってしまう場合がある。

近年、画像解析技術の向上により、被写体が笑顔を見せた瞬間に撮影を行うスマイルシャッターという機能が開発されている。しかし、笑顔写真が全て証明写真に適している訳ではないこと、証明写真サイズに合わせた写真でないこと、メガネ着用時等に認識率が低下することなど解決すべき課題もある。そこで本研究では、証明写真機で撮影された複数の写真から証明写真として適切な写真だけを選択するアルゴリズムについて提案する。

ユーザーは証明写真として適切と思われる自分の写真を選択する。顔画像データから算出した顔特徴量を説明変数として、ユーザーが証明写真として適切と判断した写真を判別するアルゴリズムをニューラルネットワークまたはサポートベクターマシンで定義する。これらは、すでに判別されている顔画像の特徴量と証明写真として適切かどうかの判定結果によって学習させる [1]。表情表出時には、眉と目の端点や中心点、口周辺の点などの多くの点が特徴的に動くと考えられている。そこで、表情研究分野では、経時的に撮影した静止画を用いて顔特徴点の変化に着目して表情認識を行うことが多い。本研究では、顔面上に配置した

顔特徴点から、顔特徴量を算出して用いる [2]。なお、画像の撮影には Microsoft 社の Kinect を利用する。

本論文の構成は以下のようになっている。第2節でアルゴリズムについて説明し、第3節では解析結果を示す。第4節は本論文のまとめである。

## 2. アルゴリズム

### 2.1 顔特徴量

顔特徴点 50 個を図 1 に示す。このうち、図における  $N_1$  を基点として定義する。

本研究では、Kinect で撮影した画像上の顔特徴点 50 個から、2 種類の顔特徴量 30 個を算出する。顔特徴量とは、座標値である顔特徴点から算出される、顔画像分析時に利用される特徴と定義する。

この算出精度は、適切な写真判定時の予測精度に大きく影響する。本研究では、1 種類の顔特徴量を、Kinect と被験者との距離に起因する顔特徴量の算出精度低下を防ぐため、基点  $N_1$  と任意の顔特徴点とのユークリッド距離により算出する。左右の瞳孔間距離  $pd$  とすると、次式で計算される。

$$F = \sqrt{\frac{(x_i - x_{N_1})^2 + (y_i - y_{N_1})^2}{pd}} \quad (1)$$

ここで、 $x_{N_1}, y_{N_1}$  と  $x_i, y_i$  は基点と顔特徴点  $i$  の平面座標系における  $x, y$  座標を示す。

また、2 種類目の顔特徴量は、顔のパーツ間の関連を考慮するため、2 個の顔特徴点間における  $x, y$  座標の各差分値を取ることで算出する。本研究では、算出された 1 種類目の顔特徴量のうち、表情表出時に特徴的に動くことされる

<sup>1</sup> 名古屋大学  
Nagoya University, Nagoya 464-8601, Japan

<sup>2</sup> 神戸大学  
Kobe University, Kobe 657-0013, Japan

a) kita@is.nagoya-u.ac.jp

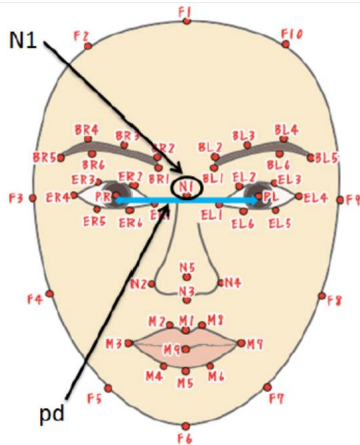


図 1 顔面に配置した 50 個の顔特徴点

Fig. 1 Arranged facial feature 50 points on face

目, 眉, 口, および顔の向きを判断するために顔の輪郭点を示す 10 個を利用する. また, 顔特徴点 2 個を組み合わせる, 2 種類目の顔特徴量 20 個を加えた, 合計 30 個の顔特徴量を写真選択の説明変数として用いる.

## 2.2 Kinect

本研究では, 被写体を撮影するデバイスとして Microsoft 社の Kinect を用いる. Kinect は RGB カメラ, 深度センサ, マルチアレイマイクロフォン等を内蔵しており, 人間の三次元位置情報や動き, 声等を認識することが可能である. 本研究では RGB カメラ, 深度センサを利用して Kinect から電子的な画像を取得し, 被写体の三次元位置情報及び骨格座標を取得する.

## 2.3 実験データの取得

実験データは以下のようにして取得する.

- (1) Kinect を用いて被験者の写真を指定枚数だけ撮影する.
- (2) 撮影した顔画像データから, WebAPI を用いて顔特徴点の座標値を求める.
- (3) 顔特徴点から式 (1) によって顔特徴量を計算する.
- (4) 被験者自身が, 撮影した写真について, 証明写真として良い写真 (Good), または悪い写真 (Bad) を判定する.

## 2.4 判別式

アルゴリズムでは, 説明変数として 30 個の顔特徴量を, 目的変数として証明写真として良い写真 (Good), 悪い写真 (Bad) をとる. 関係式はニューラルネットワークまたはサポートベクターマシンを用いて定義する.

説明変数ベクトルを  $z$ , 目的変数を  $f$ , これらの非線形関係式を  $F$  とすると, 以下のように定義できる.

$$f=F(z) \quad (2)$$

$$z=\{z_1, z_2, \dots, z_{30}\} \quad (3)$$

$$f=\{\text{Good, Bad}\} \quad (4)$$

表 1 写真判定における予測精度

Table 1 Estimate accuracy rate regarding photo judgement

被験者	ANN(%)	SVM (%)
A	70	75
B	55	65
C	60	70
D	35	30
E	70	80
F	55	50
G	50	60
H	80	85
I	70	70
J	75	75
平均	62	66

## 3. 解析例

Kinect を用いて撮影した被験者 1 名の 90 枚の写真を用いて, ニューラルネットワークまたはサポートベクターマシンで学習する. このデータを元に, 20 代の男性 5 名と女性 5 名の合計 10 名の被験者の顔画像について判定する. 各被験者に対してそれぞれ 20 枚の写真を撮影し, 証明写真として使用したいか否かを, 各被験者に判定してもらう. 被験者の判定と各アルゴリズムに基づく判定の一致率を正答率として, アルゴリズムの精度評価を行う. 2 つのアルゴリズムに対する各被験者の予測精度を表 1 に示す. ニューラルネットワークの場合は予測精度は 35% から 80% であり, 平均予測精度は 62% である. サポートベクターマシンの場合は予測精度は 30% から 85% であり, 平均予測精度は 66% である.

## 4. 結論

本研究では, 証明写真として適切な写真を判定するためのアルゴリズムについて述べた. ニューラルネットワークとサポートベクターマシンを用いて, 電子的な顔画像データから算出した顔特徴量を説明変数として, 写真が証明写真として良い写真か否かを判別するアルゴリズムを定義した. 解析例において, 1 名の被験者の写真をもとに 10 名の被験者の写真の判定を行った. ニューラルネットワークにおける平均予測精度は 62%, サポートベクターマシンにおける平均予測精度は 66% であった. また, 予測精度に個人差が大きいことがわかる. 今後は, 一層の予測精度の改善が求められる.

## 参考文献

- [1] 石川博, 新見礼彦, 白石陽, 横山昌平『データマイニングと集合知』, 共立出版, 2012 年
- [2] 野宮浩揮, 宝珍輝尚. 顔特徴量の有用性推定に基づく特徴抽出による表情認識, 知能と情報 (2011).