口唇の動き特徴の個人識別法への適用

学生員

根田 雅稔

(秋田大学)

正 員

西田 眞

(秋田大学)

正 員

石井 雅樹

(秋田県工業技術センター)

非会員

佐藤 和人

(秋田県工業技術センター)

Application of the Feature of Lip Motion in Continuous Image to Personal Identification
Masatoshi Konda, Student Member, Makoto Nishida, Member (Akita University), Masaki Ishii, Member, Kazuhito Sato, Non-Member (Akita Prefecture Industrial Technology Center)

The aim of this paper is to apply the feature of lip motion in continuous image to personal identification. At first, four feature points are selected around the lip region. Next, the following are respectively computed: distance between the upper and lower and distance between the right and left. Finally, personal identification is carried out using the features that made by the measurements in continuous image.

By using the Mahalanobis distance, twenty persons are accurately distinguished. The experimental result indicates the effectiveness of the feature of lip motion to personal identification.

キーワード：個人識別，口唇，動き特徴，動画像処理

1. はじめに

近年，個人識別を目的として，第三者には容易に真似することのできない身体的特徴を用いたバイオメトリックス（生体認証）に関する研究が盛んに行われている⁽¹⁾。バイオメトリックスを用いた場合，パスワードを記憶する必要がなく，また，データを盗難されたり，改変されたりする可能性が極めて低いといった利点を有しているため，新たな個人認識手法として期待されている。

筆者らはこれまで，発話時の口唇の動きに着目し，言語認識に有効な特徴量の抽出法について検討を加えてきた⁽²⁾。その中で，発話に伴う口唇の動き特徴には個人差の含まれていることが認められた。このことは被験者の口唇の動き特徴の差異を用いることにより，個人識別が可能であることを示唆するものである。そこで本研究では，発話に伴う口唇の動き変化を個人照合するための特徴量とし，各被験者の特徴データの分散を考慮可能なマハラノビス距離⁽³⁾を用いた個人識別法について検討を加えたので報告する。

2. 使用画像データおよび解析手法

<2. 1> 使用画像データ 本研究では被験者とカメラの位置を固定（距離 1.5m）し，被験者 20 名（A～T）が単語「あいいうえお」を 5 回ずつ発話した時の動画像を CCD カメラを用いて取得した。その後，動画像を 30 フレーム/秒の時系列静止画像データに変換し，解析に用いた。画像サイズは横 320 ピクセル，縦 240 ピクセル，階調は 8 ビット（256

グレースケール）である。なお，発話の開始時および終了時を把握するため，発話前後では口唇を閉じた状態とした。また，同一被験者でも発話時間の相違等により得られる特徴の異なることが予想されるため，各被験者に対し，発話画像データ取得時にはできるだけ均一に発話するように事前指導した。上記の条件下において取得した 5 回の発話画像データのうち，3 回分の発話画像データの平均を各被験者の特徴データ，残り 2 回の発話画像データをテストデータとして用いて個人識別法の有効性の検証を行った。

<2. 2> 解析手法 発話に伴う口唇の主な動きは口唇の上下方向の開閉，および左右方向の伸縮によるものとに大別される⁽⁴⁾。そこで本研究でも，発話時の口唇の動き情報を表わす特徴量を求めるため，口唇の上下および左右の時系列的な距離変化に注目した。具体的には図 1 に示すように口唇の上下左右に位置する口唇輪郭上の 4 点（A～D）を特徴点として設定した。さらに，図 2 に示すように，特徴点 A－B 間の距離を口唇の横幅（diXi），特徴点 C－D 間の距離を口唇の縦幅（diYi）とし，口唇の横幅（diXi）と縦幅（diYi）を要素としたベクトルの大きさ $\left| \overrightarrow{diZ_i} \right|$ を(1)式を用いて求め，注目する静止画像データにおける特徴量とした。

$$\left| \overrightarrow{diZ_i} \right| = \sqrt{(diX_i)^2 + (diY_i)^2} \quad \dots \quad (1)$$

ここで，
i：時系列静止画像データにおけるフレーム No.

(1 ≤ i ≤ n)

n : 時系列静止画像データのフレーム数
である。

同様の処理を連続する n 枚の時系列静止画像データにも
施し、各画像データの特徴量をそれぞれ求め、連続する画
像データ間における口唇の動き特徴の変化を追跡した。本
研究では、各画像より得られた特徴量を要素とする判別特
徴ベクトル $\vec{V}_m$ を(2)式を用いて求め、これを話者固有の特性
を表す情報であると仮定した。

$$\vec{V}_m = (diZ_{1m}, diZ_{2m}, diZ_{3m}, \dots, diZ_{nm}) \dots (2)$$

ここで、

$\vec{V}_m$: 被験者 m の判別特徴ベクトル (1 ≤ m ≤ 20)

diZ_{nm} : 被験者 m の n フレーム目の特徴量

である。

さらに、各被験者の判別特徴ベクトル $\vec{V}_m$ とテストベク
トル $\vec{U}$ との距離を最短距離法を用いて求め、最小の値が得
られた被験者にテストデータを帰属させ、個人照合を行っ
た。なお、オペレータが発話指示を行ってから被験者が実
際に発話するまでの時間には、各被験者によりばらつき
の生じることを確認している。そこで本研究では、各被験
者が口唇を開いた瞬間のフレームをスタートフレームと見
なして処理した。

3. 実験結果

マハラノビス距離⁽³⁾を用いて、判別特徴ベクトル $\vec{V}_m$ とテ
ストベクトル $\vec{U}$ の距離を算出した結果の例 (被験者 A) を
表 1 に示す。なお、判別特徴ベクトルとテストベクトルの
距離の値が小さい程、両者は類似していることを意味し
ている。テストベクトルは他の被験者と比較し、被験者 A と
の値が最も小さくなり、良好に識別されている様子が認め
られる。さらに、他の被験者についても識別を行ったところ、
表 2 に示すように 20 人全員(100%)の識別が正しく行わ
れている。以上の結果は発話時の口唇の動きを特徴量とし、
マハラノビス距離を用いた個人識別が可能であることを示
唆している。

4. まとめ

本研究では、顔の中でも著しく動きの変化する口唇に着
目し、発話に伴う口唇の上下左右の動きを特徴量とした個
人識別について検討を加えた。その結果、口唇の時系列的
な動きは個人識別を行う上で有効な特徴量となり得るこ
を示すとともに、本研究で提案するマハラノビス距離を用
いた手法は個人識別を行う上で有用であることを明らかに
した。

提案手法では被験者以外のデータが入力された場合にも、
被験者(特徴データ)のいずれかに分類してしまうため、未登
録者データの棄却法について今後検討する予定である。

図 1 口唇の特徴点

Fig. 1. Four feature points around the lip region.

図2 口唇距離を用いた特徴量 $|diZ_i|$

Fig. 2. Feature used for the personal identification.

表 1 ベクトル間距離の算出結果例(被験者 A)

Table 1. An example of the calculation result (Sample A).

被験者名	被験者A	被験者B	被験者C	被験者D	被験者E
距離	1033.8	70045.0	10709.6	18646.8	7311.2
被験者名	被験者F	被験者G	被験者H	被験者I	被験者J
距離	96585.7	85495.6	68850.5	155279.8	15204.6
被験者名	被験者K	被験者L	被験者M	被験者N	被験者O
距離	15626.4	16971.3	6404.4	18887.2	44463.6
被験者名	被験者P	被験者Q	被験者R	被験者S	被験者T
距離	34771.8	17278.3	25698.7	10035.3	23553.0

表 2 個人識別結果

Table 2. Personal identification result.

	被験者数
識別した人数	20名/20名
誤識別した人数	0名/20名

(平成 11 年 12 月 15 日受付)

文 献

- (1) 小杉:「モザイクとニューラルネットワークを用いた顔画像の認識」,信学論,Vol.J76-D-II,No6,pp.1132-1139 (1993)
- (2) 石井,佐藤,西田,景山:「時系列口唇画像を用いた読唇のための特徴抽出と唇の動き解析」,電学論,Vol.119-D,No4,pp.465-472 (1999)
- (3) 高木,下田:「画像解析ハンドブック」,東京大学出版会,p654 (1991)
- (4) 上條:「口腔解剖学.第 2 巻」,アナトーム社 (1974)