

# 心理・体調変化の解析を目的とした口唇の動き抽出法の改善

## Improvement of Lip Motion Extraction Method for Analysis of Psychology and Physical Conditions

中村 悦郎<sup>1</sup>                      高橋 毅<sup>1</sup>                      景山 陽一<sup>1</sup>  
Etsuro Nakamura              Tsuyoshi Takahashi              Yoichi Kageyama  
石井 雅樹<sup>2</sup>                      西田 眞<sup>1</sup>  
Masaki Ishii                      Makoto Nishida  
<sup>1</sup>秋田大学                              <sup>2</sup>秋田県立大学  
Akita University                      Akita Prefectural University

**Abstract:** Lip motion features such as changes in the lip width and length provide important information for analyzing psychology and physical conditions. Our previous studies revealed that we can analyze psychology and physical conditions by lip motion features. Therefore, it is important to develop a method for extracting the lip width and length from facial image data with high accuracy. In this paper, we propose a lip motion extraction method on the basis of color and shape information in the oral region.

### 1. はじめに

近年、超高齢社会の到来に伴う諸課題への対策の一つとして、クオリティ・オブ・ライフ（以下、QOLと表記）の重要性和その維持の必要性が注目されている。すなわち、個々のQOLが維持・向上可能であれば、生活習慣病予防やストレス対策につながるため、社会の安定[1][2]、介護や医療費等の社会的コスト削減[2]に寄与できると考える。また、QOLは笑いから喚起される「喜び」やアニマルセラピー等により得られる「穏やかさ」によって向上すると報告されている[3]。すなわち、これら2つの感情を検出することは、超高齢社会におけるQOLの維持・向上に寄与できると考える。

本研究グループではこれまでに、発話に伴う口唇の動きが①発声の有無にかかわらず使用可能、②非接触で取得可能、③表情や音声との併用が可能である等の利点を有していることに着目し、口唇の縦幅および横幅の変化（以下、口唇の動き特徴と表記）から「喜び」の情動検出、および体調変化の検出が可能であることを明らかにしている[4]。したがって、口唇の動き特徴を自動取得する技術を開発することは、QOLの認識に寄与すると考える。しかしながら、従来手法[5]では、口唇周辺の陰影や肌の色の個人差に起因して口唇形状抽出に失敗し、口唇の動き特徴を良好に取得できない場合を認めた。

そこで本研究では、口唇の動き特徴に着目した心理・体調変化の解析を目的とし、その前処理である口唇の動き特徴取得手法を改善する。具体的には、口内領域および上唇、下唇の厚さに着目した補正手法に関して検討を加えた。

### 2. 使用データ

一般的な蛍光灯下（照度約500~1000lx）においてCCDビデオカメラ（Point Grey Research; Grasshopper）を用い、被験者10名（a~j）が「本人の氏名」を6回発話する様子を撮影した（図1参照）。撮影した発話動画データを時系列静止画像（30fps）に分割し、これを使用データとした。なお、本研究は「秋田大学手形地区におけるヒトを対象とした研究に関する倫理規定第6条第2項」に基づき、被験者の同意を得た上でデータを取得し使用している。

### 3. 口内領域の色情報解析

従来手法では、L\*a\*b\*色空間において、赤味を表すa\*および黄味を表すb\*に着目し、口唇形状を抽出している。しかしながら、口唇周辺の陰影や肌の色の個人差に起因し、口唇領域の過剰抽出が生じる場合を認めた。そこで本研究では、口内領域が口唇周辺の陰影や肌の色の個人差に影響されにくいと仮定し、口内領域および上唇・下唇の厚さを用いた補正処理の検討を行った。本章では口内領域を抽出するために、口内領域の色情報解析を行った。はじめに、使用データを以下の①~③の3種類に分類した（図2参照）。

- ① 口が開いている、かつ歯が存在する画像
- ② 口が開いている、かつ歯が存在しない画像
- ③ 口が閉じている画像

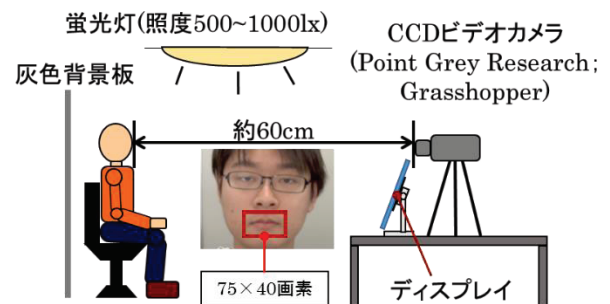


図1 データ取得環境

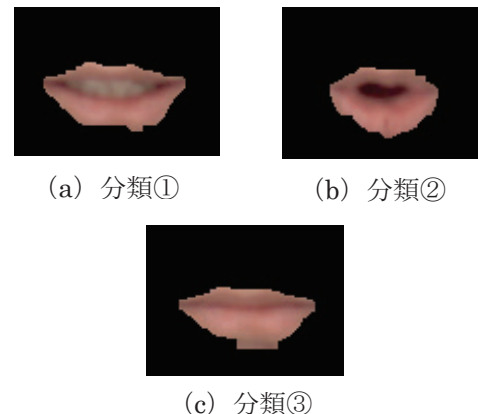


図2 口唇画像の分類

次に、口唇横幅を2分する垂線と、そこから横方向(x座標)に±10画素離れた垂線上における各画素の明度値(L\*)および赤味(a\*)を算出した。その結果、口内領域と口唇部の境界には下記に示す3つの特徴が存在した。

- 特徴1 分類①と③の場合、明度分布の谷が存在する(図3参照)。
- 特徴2 分類②の場合、明度分布における正の傾きおよび負の傾きが最大となる(図4参照)。
- 特徴3 分類①と②の場合、赤味の分布における正の傾きおよび負の傾きが最大となる(図5参照)。

## 4. 提案手法

### 4.1 前処理

口唇抽出処理を実施した画像において、口内領域が欠損し、口内領域と口唇部の境界における特徴1~3の抽出が困難となる場合がある(図6(a)参照)。このため、前処理として欠損の補正処理を施した(図6(b)参照)。

### 4.2 口内領域抽出処理

#### 4.2.1 明度分布の谷が存在する画素の抽出処理

垂線上の任意の画素Aとその上下の画素B(上)、C(下)の明度値を比較し、条件Iを満足した場合は谷、条件IIを満足した場合は山と判定した。

$$(B > A = C) \vee (A < B \wedge A < C) \vee (B = A < C) \quad (\text{条件I})$$

$$(B < A = C) \vee (A > B \wedge A > C) \vee (B = A > C) \quad (\text{条件II})$$

#### 4.2.2 明度値分布の谷におけるノイズの除去

谷と判定された画素とその前後に存在する山との明度値の差を  $t_1$  (前),  $t_2$  (後), 最も深い谷とその前後に存在する山との明度値の差を  $T_1$  (前),  $T_2$  (後)

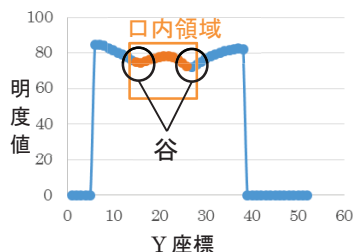


図3 明度分布における谷

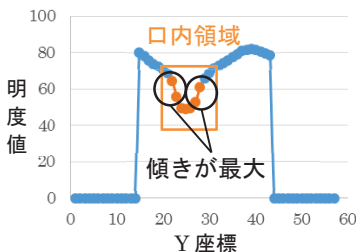


図4 明度分布の傾きが最大となる画素

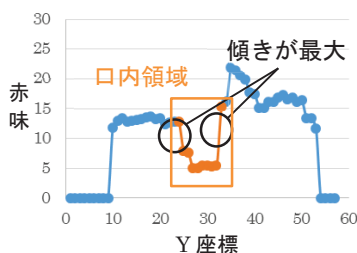


図5 赤味の分布の傾きが最大となる画素

とした場合に、以下の条件III, IVが成立した谷をノイズと判定した。

$$t_1 < 6.0 \wedge t_2 < 6.0 \quad (\text{条件III})$$

$$\min(t_1, t_2) / \min(T_1, T_2) < 2.0 \quad (\text{条件IV})$$

### 4.2.3 傾きの最大値算出処理

垂線上に隣接した画素A, 画素C間における  $a^*$  および  $L^*$ の傾きを算出し、正および負の傾きの最大値をそれぞれ算出した。なお、算出した最大値の絶対値が4.0以下の点をノイズと見なし除外した。

### 4.3 全画像に対するノイズ除去処理

口内領域を抽出した結果、口内領域と口唇部の境界が白い線で出力された。しかしながら、図7に示すように、ノイズを完全に除去することが困難な画像が複数認められた。そこで、「マスクを用いたノイズ除去処理」および「孤立点の除去処理」を実施し、ノイズの除去を行った。

#### 4.3.1 マスク画像を用いたノイズ除去処理

マスク画像を用いたノイズ除去処理は以下の(1)~(3)の手順で行う。

- (1) 図8(a)に示す口内領域と口唇部の境界に対して、膨張処理を施す。
- (2) 図8(b)に示す膨張処理を実施した口内領域と口唇部の境界に対して、面積が最大となる領域のみを残し、マスク画像とする(図8(c)参照)。
- (3) マスク画像と図8(a)に示す画像の論理積を求め、ノイズを除去する(図8(d)参照)。その結果、図9のようにノイズが除去された画像を取得する。



(a) 処理前 (b) 処理後

図6 前処理

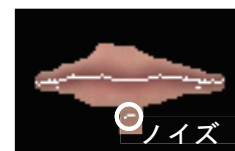
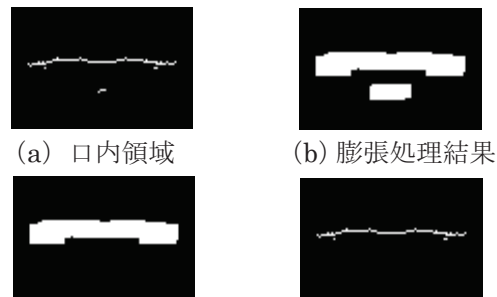


図7 口内領域抽出処理結果



(a) 口内領域 (b) 膨張処理結果 (c) マスク画像 (d) マスク処理結果

図8 マスクを用いたノイズ除去処理



図9 マスクを用いたノイズ除去処理結果

### 4.3.2 孤立点の除去処理

対象となる孤立点は、不連続であることに着目し、ノイズとして除去を行った。具体的には、口内領域と口唇部の境界である白い点の画素に着目し、左右5画素内の白い点が2つ以下の場合にはノイズと判定した。

### 4.4 境界の補間処理

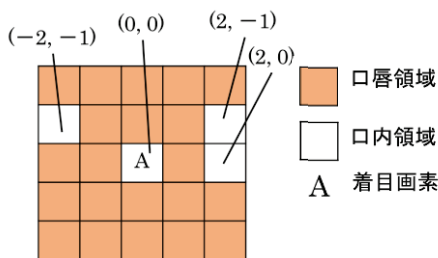
口内領域と口唇部の境界が不連続となる場合を認めたため、境界線を補間する処理を(1)～(5)のように実施した。また、補間処理例を図10に示す。

- (1) 口内領域と口唇部の境界である白い画素の任意の点を着目画素(A)とし、それを中心として5×5画素の処理領域を設定する。
- (2) 着目画素(A)を原点(0,0)とし、処理領域内に存在する白い画素の座標を取得する。(図10(a)参照)
- (3) x座標が正となる画素群に存在する白い画素のy座標の平均値mを算出した。また、x座標が負となる画素群に対しても同様に平均値nを求める。なお、平均値は整数とし、小数点以下を切り捨てる(図10(b)参照)。
- (4) (1, m), (-1, n)を白で塗りつぶす(図10(c)参照)。
- (5) すべての口内領域に属する画素に対して(1)～(4)を繰り返す。

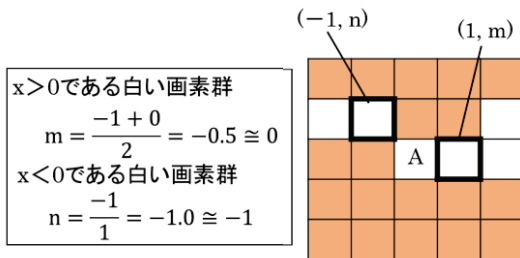
### 4.5 補正処理

提案する補正処理では、上唇の厚さの推定値と実際に取得した上唇の厚さの比較を行い、実際の上唇の厚さが厚い場合は過剰抽出と判定し、補正処理を行う。具体的には、発話1回分に対してi～viiの処理を施した。

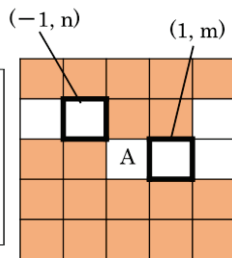
- i. 口唇抽出に成功している画像中から“口の閉じている”画像Pおよび“口の横幅が最も短い”画像Q、発話1回中の“任意の”画像Rをそれぞれ1枚ずつ選定した(図11参照)。なお、画像Pおよび画像Qは、上唇の厚さの推定式を求め際に使用する基準画像である。



(a) 座標の取得



(b) y 座標平均値



(c) 補間処理結果

図10 補間処理例

- ii. 画像P, Q, Rに対して口内領域の横幅 $Wp, Wq, Wr$ を取得し、口唇左端から右端に向かって上唇の厚さ $p(i), q(j), r(k)$ を取得した。
- iii. 式(1)を用いて、 $p(i)$ に対応する $q(j)$ を算出した。

$$Q(i) = q\left(\frac{Wq}{Wp}i\right) = q(j) \quad (1)$$

- iv. 図12に示すように、横軸を口内領域横幅、縦軸を上唇の厚さとした平面上の直交座標系において、2点 $(Wp, p(i)), (Wq, Q(i))$ 間を結ぶ直線の式 $f_i$ (式(2)参照)を算出し、上唇の厚さを口内領域横幅 $Wr$ によって推定する式を算出した(以下、推定式と表記)。

$$f_i(Wr) = \frac{p(i)-Q(i)}{Wp-Wq}Wr - \frac{p(i)-Q(i)}{Wp-Wq}Wp + p(i) \quad (2)$$

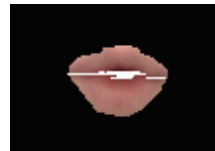
- v. 式(3)を用いて、 $f_i$ に対応する $r(k)$ を算出した。

$$R(i) = r\left(\frac{Wr}{Wp}i\right) = r(k) \quad (3)$$

- vi. 取得した上唇の厚さ $R(i)$ と上唇の厚さ $f_i$ の値を比較し、 $f_i$ の値が $R(i)$ よりも厚い場合は、 $f_i$ の値を厚さとした(図13参照)。これを発話1回のすべてのフレームに対して実施した。
- vii. i～viの処理を下唇の厚さに対しても実施した。



(a) 画像P



(b) 画像Q



(c) 画像R

図11 選定画像例

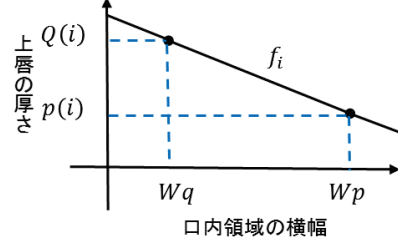


図12 推定式 $f_i$

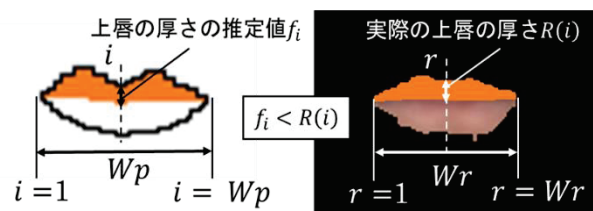


図13 過剰抽出判定



## 5. 実験結果および考察

はじめに、目視、提案手法、ならびに従来手法による口唇抽出を実施し、口唇の縦幅および横幅を矩形として出力した(図14参照)。次に、提案手法および従来手法により抽出した口唇縦幅および横幅と、目視により抽出した口唇との画素の差をそれぞれ算出した(図15参照)。最後に、口唇の縦幅および横幅の差を許容する範囲(以下、許容範囲と表記)を0から±5画素まで1画素刻みで設定し、各許容範囲における累積成功率を算出した。縦幅における累積成功率の結果例を表1に、横幅における累積成功率の結果例を表2にそれぞれ示す。

実験の結果、口唇縦幅に着目すると、被験者10名中9名(a~e, g~j)はすべての許容範囲において抽出成功率が向上した。次に、口唇横幅に着目すると被験者10名中4名はすべての許容範囲において抽出成功率が向上した。これらの結果は、提案手法が縦幅の補正に有用であることを示唆している。縦幅の抽出成功率が向上した被験者の場合、発話に伴う口内領域横幅の変化と口唇の厚さの関係が提案手法で想定した推定式 $f_i$ に適合しているため、抽出成功率が向上したと考える。一方、口唇縦幅の抽出成功率が低下した被験者1名(被験者f)の場合、提案手法で想定している①口唇が中央に圧縮される、もしくは②横方向に引き伸ばされることによって生じる口唇の厚さの変動ではなく、③口を開いた際に口唇の内側が表出し、口唇領域として抽出された場合に生じる口唇の厚さの変動が生じたと考える。このため、推定式 $f_i$ に適合せず、抽出成功率が低下したと考える。

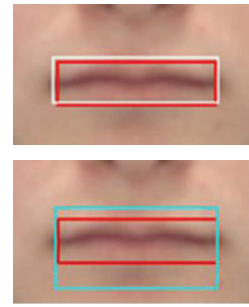
したがって、口唇横幅においては従来手法を用い、口唇縦幅においては提案手法を用いることで、精度良く口唇の動き特徴を取得可能であることが明らかになった。このことは、提案手法を用いることで、従来手法と比較して発話に伴う縦幅のばらつきを増減率を精度良く取得可能となるため、従来研究[4]における喜びの情動判別率向上に寄与すると考える。

## 6. おわりに

本研究では、心理・体調変化の解析を目的とした口唇の動き抽出法[5]の改善を目標とし、口内領域および上唇、下唇の厚さに着目した補正手法に関して検討を加えた。その結果、発話に伴う口内領域の横幅の変化と、上唇および下唇の厚さの変化には比例関係が存在し、それに基づいて口唇の縦幅の補正が可能であることを明らかにした。また、口唇横幅の抽出においては従来手法を用い、縦幅の抽出においては提案手法を用いることで、従来手法のみを用いた場合と比較して精度良く口唇の動き特徴を取得可能であることを明らかにした。

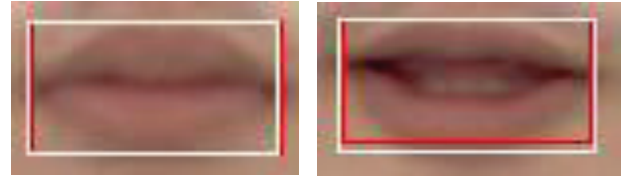
今後は、上唇および下唇の厚さの推定式 $f_i$ に関してさらなる検討を行っていく予定である。

本研究は、JSPS 科研費(No.15K00222)の助成を受け、遂行されたことを付記する。



赤い矩形：目視による矩形  
白い矩形：提案手法による矩形  
青い矩形：従来手法による矩形

図14 過剰抽出判定



(a) 縦幅：同値 横幅：同値  
(b) 縦幅：3画素の差 横幅：2画素の差

図15 縦幅および横幅の差

表1 口唇縦幅における累積成功率(被験者e)

許容範囲(pixel)	0	±1	±2	±3	±4	±5
従来手法(%)	2	6	11	19	24	29
提案手法(%)	11	36	56	69	83	90

表2 口唇横幅における累積成功率(被験者g)

許容範囲(pixel)	0	±1	±2	±3	±4	±5
従来手法(%)	5	15	31	49	67	79
提案手法(%)	9	29	47	64	75	84

## 参考文献

- [1] 加藤, 財務省財務総合政策研究所編: 超高齢社会の介護制度—持続可能な制度構築と地域づくり—, 中央経済社, 2015
- [2] 厚生労働白書: 持続可能な社会保障制度と支え合いの循環〜「地域」への参加と「働き方」の見直し〜, 平成18年度版, <http://www.mhlw.go.jp/wp/hakusyo/kousei/06/> (Accessed 2016/11/20)
- [3] 太湯, 小林, 永瀬, 生長豊: 認知症高齢者に対するイヌによる動物介在療法の有効性, 川崎医療福祉学会誌, Vol.17, No.2, pp.353-361, 2008
- [4] Y. Kageyama, A. Momose, T. Takahashi, M. Ishii, M. Nishida, A. Mohemmed and N. Kasabov: Analysis of Lip Motion Change Arising as a Result of Amusement Feeling, IEEJ Transactions on Electrical and Electronic Engineering, Vol.8, No.5, pp.538-539, 2013
- [5] 佐藤, 成田, 景山, 西田: 口唇形状情報を用いた口唇領域自動抽出処理の改善, 電気学会論文誌C, Vol.130, No.5, pp.873-881, 2010

## 連絡先

秋田大学

景山 陽一

(Tel:018-889-2786, E-mail: kageyama@ie.akita-u.ac.jp)