あいまいさを考慮した色彩情報による口唇抽出アルゴリズム

学生員 白澤 洋一*

正員 西田 眞*

A Lip Extraction Algorithm by Using Color Information Considering Obscurity

Yoichi Shirasawa*, Student Member, Makoto Nishida*, Member

This paper proposes a method for extracting lip shape and its location from sequential facial images by using color information. The proposed method has no need of extra information on a position nor a form in advance. It is also carried out without special conditions such as lipstick or lighting. Psychometric quantities of a metric hue angle, a metric hue difference and a rectangular coordinates, which are defined in CIE 1976 $L^*a^*b^*$ color space, are used for the extraction. The method employs fuzzy reasoning in order to consider obscurity in image data such as shade on the face.

The experimental result indicate the effectiveness of the proposed method: 100 percent of facial images data was estimated a lip's position, and about 94 percent of facial images data was extracted its shape.

キーワード: 顔画像, 色彩情報, ファジィ推論, メトリック色相差, 近傍画素

Keywords: facial image, color information, fuzzy reasoning, metric hue difference, vicinity pixel

1. まえがき

近年、音声などの聴覚的情報に加え、顔画像などの視覚的情報に着目した読唇および個人識別の研究が行われている^{(1)~(4)}。視覚的情報である発話に伴う口形の時系列変化は固有の特徴を有し、読唇および個人識別を行う上で有効な情報となり得る。したがって、口形の時系列変化を特徴量として用いるためには、発話に伴い形状の変化する口唇を抽出する手法、ならびに取得される特徴の定量化を行う手法の開発が必要である。

筆者らはこれまでに、一般的な室内環境下で取得された顔画像データを対象とし、読唇および個人識別に有効な特徴量の抽出法について検討を加えてきた⁽¹⁾⁽²⁾。その中で、発話に伴う口の開閉といった“口唇の動き特徴”が読唇および個人識別に利用可能であることを明らかにした。しかし、これまでの手法では初期フレームで口唇位置を示す特徴点を手動により設定する必要があり、背景を含む顔画像から口唇を自動抽出するまでには至らなかった。画像処理による口形の時系列変化に着目した口唇形状の抽出法は報告されているものの、被験者に対して口紅の塗布、ならびに補

助的な照明の利用などといった特殊な条件下で行われている事例^{(3)~(5)}が多く、通常の条件下で取得されたデータを用いた検討はあまり行われていないのが現状である。この理由としては、顔面における肌の色味の分布が様々ではないこと、口唇輪郭も明確であるとは限らないことが挙げられる。そのため、色彩情報に着目した顔面における肌の抽出法も報告されているものの、口唇も肌の一部として誤抽出されている事例が多い^{(6)~(8)}。呉らは、肌とそれ以外の領域を2値的に判定せず、各色がどのくらい肌色に似ているかという度合いで表現している⁽⁹⁾。しかしながら、肌のみの色彩情報に着目しているため、肌と口唇の色差が小さい人物の場合、肌と口唇を明確に分類することは困難であると予想される。本論文では、口唇抽出を目的として肌を抽出するため、肌と口唇を明確に分類する必要がある。そこで、肌と口唇それぞれの色彩情報に着目した肌および口唇抽出について検討を加えた。

口唇位置の抽出法としては、口唇の形状などに関する情報を有するテンプレートを用いる手法が報告されている⁽¹⁰⁾⁽¹¹⁾。しかしながら、これらの手法ではテンプレート、すなわち形状に関する情報を事前に設定する必要がある。

そこで本論文では、通常の室内環境下（日常一般的と考えられる蛍光灯による照明下、照度 150 lx 以上）⁽¹²⁾で取得された発話に伴う時系列顔画像を対象とし、被験者に対して口紅の塗布や補助的な照明を利用することなく口唇の位

* 秋田大学工学資源学部情報工学科
〒010-8502 秋田市手形学園町 1-1
Department of Computer Science and Engineering,
Faculty of Engineering and Resource Science, Akita
University,
1-1 Tegata Gakuen, Akita, 010-8502

置および形状を自動抽出する手法を提案する。

肌および口唇それぞれが有する個人差は大きいこと、口唇と肌の色差が小さい人物が存在すること、ならびに画像中での肌や口唇の色彩情報は照明条件によって大きく変動すること^{(13)~(15)}等を考慮し、提案手法では口唇抽出処理に必要な色彩情報は、被験者を撮影した時の画像データより取得した。また、天候および撮影時間帯などの違いにより、取得した画像全体の色味は変動する。加えて、補助的な照明を用いていないため、色ムラの生じることを確認している。そこで、提案手法ではこれらが色彩情報に与える影響を低減することを目的として、背景に模様のない一様な壁を用いた。また、被験者を撮影する前に、あらかじめ背景である壁（背景画像）の撮影を行い、口唇抽出の際には背景画像と時系列顔画像、それぞれの色彩情報を用いた。

さらに、蛍光灯の周期的な明るさ変化であるフリッカ、ならびに下唇の下方に生じる陰影の影響を低減するため、色空間として明るさ変化に対するロバスト性が高く、人間の視覚系に沿った均等色空間である CIE 1976 L*a*b*色空間⁽¹⁶⁾（以下、L*a*b*色空間と略記する）に着目した。なお、特徴量としてはL*a*b*色空間で定義された知覚色度指数および心理メトリック量を用いた。また、フリッカおよび陰影の影響による色味の変動をあいまいさと仮定し、マムダニのファジィ推論法⁽¹⁷⁾を用いた口唇の位置および形状抽出を行った。提案手法の有用性を検討するため、これまで時系列顔画像を用いた読唇を目的とする研究で多く用いられてきた特徴量、すなわち明度情報に着目した口唇抽出法^{(3)~(6)}により得られた結果との比較を行ったので報告する。

2. 口唇抽出法

本論文では、色彩情報を特徴量とする処理を施すことにより口唇位置の抽出と形状の推定を行った。照明の周期的な明度変化に伴う色味の変動や口唇輪郭の境界が明確でない等、顔画像を構成する画素の有する色彩情報はあいまいさを有している。そこで、色彩情報としては明度変化に対してロバスト性を有するL*a*b*色空間で定義された特徴量を用いた。また、画像情報の持つあいまいさ、発話に伴う肌と口唇における色味の時系列的な変動の相違、ならびに単独画素と近傍画素との関連情報などを考慮することにより、ファジィ推論による口唇の自動抽出を行った。

提案手法における口唇抽出処理の流れを図1に示す。

〈2・1〉 L*a*b*色空間 顔面の肌と口唇の相違を示す特徴を有し、照明の影響に対してロバスト性を有することに着目し、L*a*b*色空間の知覚色度指数である a^* 、 b^* 、ならびに心理メトリック量である h_{ab} ⁽¹⁶⁾を特徴量として用いた。

RGB色空間からL*a*b*色空間への変換はXYZ色空間を経由して行う。RGB色空間からXYZ色空間⁽¹⁶⁾への変換式を(1)式に示す。

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.607 & 0.174 & 0.201 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.117 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \dots\dots\dots (1)$$

ここで、XYZ色空間への変換式である(1)式はNTSC方式に基づいた式である。また、XYZ色空間を(2)式~(4)式によりL*a*b*色空間に変換する。

図1 提案手法における口唇抽出処理の流れ

Fig. 1. Flowchart of the lip extraction procedure.

$$L^* = \begin{cases} 116 \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} - 16 & \text{if } (Y/Y_n) > 0.008856 \\ 903.29 \left(\frac{Y}{Y_n} \right) & \text{if } (Y/Y_n) \leq 0.008856 \end{cases} \quad (2)$$

$$a^* = 504.3 \left[\left(\frac{X}{X_n} \right)^{\frac{1}{3}} - \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} \right] \quad (3)$$

$$b^* = 201.7 \left[\left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} - \left(\frac{Z}{Z_n} \right)^{\frac{1}{3}} \right] \quad (4)$$

ここで、 L^* は明度指数である。また、 a^* および b^* は色相と彩度に関係する量であり、 a^* は正に大きい程赤に、負に大きい程緑であることを表す。

X_n, Y_n, Z_n は標準の光 $C^{(16)}$ における完全拡散面の三刺激値である。 X_n, Y_n, Z_n の値は次のように表される⁽¹⁶⁾。

$$(X_n, Y_n, Z_n) = (98.072, 100.000, 118.225) \quad (5)$$

提案手法で着目する色相に相関のあるメトリック色相角 h_{ab} は(5)式で定義される。ここで、 $a^* > 0, b^* > 0$ の場合、 h_{ab} は $0^\circ \sim 90^\circ$ の範囲の値となる⁽¹⁶⁾。

〈2.2〉 前処理 データ取得時の照明条件の変化から生じる色むらを除去するため、被験者毎に顔画像から背景画像の a^*, b^* それぞれの差である $\Delta a^*, \Delta b^*$ を算出した。なお、 $\Delta a^*, \Delta b^*$ は(6)式で示される。また、明度情報を表す L^* は各フレームの顔画像の値を用いた。顔画像および背景画像の画像例を図1(a)、図1(b)にそれぞれ示す。

$$\left. \begin{aligned} \Delta a^* &= a_f^* - a_w^* \\ \Delta b^* &= b_f^* - b_w^* \end{aligned} \right\} \quad (6)$$

ここで、

- a_f^* : 顔画像の各画素の a^* 値、
- a_w^* : 背景画像の各画素の a^* 値、
- b_f^* : 顔画像の各画素の b^* 値、
- b_w^* : 背景画像の各画素の b^* 値

である。また、 Δa^* および Δb^* のメトリック色相角 h_{ab}' は(7)式で求めた。

$$h_{ab}' = \arctan\left(\frac{\Delta a^*}{\Delta b^*}\right) \times \frac{180}{\pi} \quad (7)$$

次に、口唇の輪郭を損なうことなく雑音の除去および平滑化を行うため、 3×3 サイズのメディアンフィルタ⁽¹⁸⁾を用いて平滑化処理を施した。前処理済の画像例を図1(c)に示す。

〈2.3〉 ファジィ推論法によるクラス分類 ファジィ推論の知識ベースは、推論規則、前件部メンバーシップ関数および後件部メンバーシップ関数より構成され、推論の合成則としてはマムダニの方法⁽¹⁷⁾を採用した。フレームNo.1では「肌クラス」、「口唇クラス」および「壁クラス」の3クラス、フレームNo.2以降のフレームでは口唇周辺に

着目して処理するため、「肌クラス」、「口唇クラス」の2クラスをそれぞれ設定した。各フレームの対象データ（前処理済画像の各画素）に対し、各クラスの帰属度 y'_i ($i=1 \sim 3$) をそれぞれ算出し、帰属度の最も高いクラスを分類結果として出力した。 $i=1 \sim 3$ はそれぞれ「肌クラス」、「口唇クラス」および「壁クラス」に対応する。ただし、各クラスの帰属度が全てしきい値 T_1 ($T_1=10.0$) 以下の場合には着目画素がどのクラスにも帰属しないと仮定し、「その他クラス」として分類した。提案手法におけるファジィ推論処理の概要を以下に示す。

(1) 前件部メンバーシップ関数の生成 推定のための特徴量として、フレームNo.1では各被験者における前処理済の顔画像から得られた $\Delta a^*, \Delta b^*$ を用いた。また、フレームNo.2以降のフレームでは $\Delta a^*, \Delta b^*$ および h_{ab}' を用いた。このため、特徴空間は各特徴量を座標軸とする多次元空間となり、特徴ベクトルは各クラスの画素のRGB値から算出された知覚色度指数および心理メトリック量である。

① 教師データ取得： 前処理済画像のフレームNo.1において、オペレータが被験者毎のデータから「肌」、「口唇」、ならびに背景である「壁」と判断した領域より、それぞれ30画素ずつ手動で抽出した。

さらに、 $\Delta a^*, \Delta b^*$ および h_{ab}' を算出し、各領域におけるそれぞれの平均および標準偏差、すなわち Δa^* の $\bar{C}_i$ と σ_i 、 Δb^* の $\bar{C}_i$ と σ_i 、 h_{ab}' の $\bar{C}_i$ と σ_i を教師データとした。なお、特定の部位に偏ることがないように教師データを取得している。

② 前件部メンバーシップ関数の生成： 前件部のメンバーシップ関数として、各特徴データの平均値 $\bar{C}_i$ を中心とし標準偏差 σ_i の α_i 倍および $\alpha_i\beta_i$ 倍をそれぞれの底辺とする三角形の関数を定義した（図2参照）。なお、 α_i の値は各教師データの標準偏差の大きさと各特徴データ間における平均値の距離を考慮し設定した。さらに、各クラスの特徴を考慮して重み付け β_i を行った。予備実験の結果、 $\alpha_1=\alpha_2=3.0$ 、 $\alpha_3=1.5$ 、 $\beta_1=\beta_3=1.0$ 、 $\beta_2=2.0$ とした。また、壁クラス α_3 は1.5に固定した。

(2) 後件部メンバーシップ関数の生成 後件部には2種類の関数 F_1 および F_2 を用いた。 F_1 および F_2 は(8)式、(9)式でそれぞれ定義される。用いた後件部メンバーシップ関数を図3に示す。

$$F_1 = \mu(x) \quad (8)$$

$$F_2 = \sqrt{\mu(x)} \quad (9)$$

(3) ファジィ推論規則 着目画素の各クラスの帰属度を推定するファジィ推論規則を(10)式に示す。

$$\begin{aligned} \text{rule1: } & A_1 \rightarrow F_2 \\ \text{rule2: } & A_2 \rightarrow F_1 \\ \text{rule3: } & A_3 \rightarrow F_1 \\ \text{rule4: } & A_1 \text{ and } A_2 \text{ and } A_3 \rightarrow F_2 \\ \text{input: } & x_1 \text{ and } y_0 \text{ and } z_0 \\ \hline \text{output: } & y' \end{aligned} \quad (10)$$

図2 前件部メンバーシップ関数

Fig. 2. Membership function of condition part.

図3 後件部メンバーシップ関数

Fig. 3. Membership function of operation part.

ここで, rule1は Δa^* , rule2は Δb^* , rule3は h_{ω}' に着目したファジィ規則であり, rule4は Δa^* , Δb^* および h_{ω}' のAND演算を行うファジィ規則である。また, $A_1 \sim A_5$ は各特徴データの条件部メンバーシップ関数, F_1 , F_2 は各特徴データの後件部メンバーシップ関数である。 x_0 , y_0 および z_0 は入力変数であり, それぞれ着目画素における Δa^* , Δb^* および h_{ω}' に対応する。なお, y は各クラスの推論結果である。

(4) デファジィ化処理 最終出力値である帰属度 y_j は荷重平均法により(11)式を用いて求めた。

$$y_j = \frac{y_1 \cdot w_1 + y_2 \cdot w_2 + y_3 \cdot w_3 + y_4 \cdot w_4}{w_1 + w_2 + w_3 + w_4} \dots\dots\dots(11)$$

ここで, y_j ($j=1 \sim 4$)は(10)式のファジィ規則rule1～rule4それぞれの帰属度である。また, w_j は重み係数である。なお, 肌および口唇の推定においては赤みの度合いに対応する Δa^* が重要と考え, w_1 の値を大きく設定した。予備実験を行ったところ, $w_1=1.5$, $w_2=w_3=w_4=1.0$ とした時の結果が良好であったため, 本論文ではこれらの値を使用した。

(2.4) 口唇の抽出 肌は口唇と比較し, 発話に伴う時系列的な色味の変動が小さい。そこで提案手法では, 直接「口唇クラス」に分類された画素を抽出せず, 以下の処理を施すことにより口唇を抽出した。なお, フレームNo.1では(2.3)節の処理で, またフレームNo.2以降のフレームでは後述する(2.6)節の処理によりクラス分類された画素を対象とした。

- (1) 2値化処理⁽¹⁸⁾: 「肌クラス」の画素を輝度値255(白), 「肌クラス」以外の画素を輝度値0(黒)とした。2値化処理を施した画像例を図1(d)に示す。
- (2) ラベリング処理⁽¹⁸⁾: 上記処理(1)で2値化された画素を対象として膨張・収縮処理⁽¹⁸⁾およびラベリング処理を

施し, 口唇位置の抽出を行った。各ラベルの特徴パラメータとして面積, ならびに面積に対する「口唇クラス」に分類された画素数の割合(以下, 口唇クラス帰属率と表記する)を設定した。ここで, 口唇クラス帰属率を設定したのは, 口唇に関する特徴も考慮して処理することによって下記処理(3)で他の部位が誤って抽出される可能性を低減するためである。なお, 提案手法では, 口唇の位置や形状に関する情報を事前に設定する必要はない。

- (3) 口唇の抽出: 顔面に存在する各部位の中で一番大きい面積を有するのは口である場合が多い。しかしながら, まゆげなどの面積が口唇と比較して大きい場合も有り得る。そこで, 提案手法では誤抽出の防止を目的として上記処理(2)で特徴パラメータである口唇クラス帰属率がしきい値 T_2 ($T_2=0.3$)以上のラベルの中で, 特徴パラメータの面積が最大のものを口唇に対応するラベルと仮定した。
- (4) 膨張・収縮処理: 口唇に対応するラベルにおいて, 上記処理(1)で誤識別された画素がごま塩雑音となことを確認している。そこで, 口唇に対応する画素に対して膨張・収縮処理を施し, 雑音を除去した。また, 得られた画像と原画像のマッチングにより口唇を抽出した。提案手法における口唇抽出結果例を図1(e), 図1(j)に示す。

(2.5) 口唇周辺領域の設定 口唇の形状抽出の精度向上, ならびに計算処理の軽減を目的とし, フレームNo.2以降のフレームでは口唇周辺に注目した領域(以下, 口唇周辺領域と略記する)で口唇の抽出処理を施した。口唇周辺領域における設定処理の概要は以下になる。

- (1) フレームNo.n ($n=1 \sim$ 最終フレーム)において(2.4)節の処理によって抽出された口唇の形状から縦, 横方向のヒストグラムをそれぞれ作成し, 口唇の縦幅(Xwd)および横幅(Ywd)を取得した。提案手法で作成したヒストグラムの例を図1(f)に示す。次に, 得られたXwdおよびYwdから口唇位置の中心座標 $C(Xc, Yc)$ を算出した(図4参照)。
- (2) 発話に伴う口唇の変化に対応した口唇を抽出するため, 座標 $C(Xc, Yc)$ を中心として横幅 $2Xwd$, 縦幅 $2Ywd$ と設定した四角形をフレームNo.n+1の口唇周辺領域とした。口唇周辺領域の設定例を図1(h)に示す。

なお, 本論文では使用する全画像において, 更新された口唇周辺領域はフレーム毎に口唇を包含することを目視により確認している。

(2.6) 口唇周辺領域における口唇抽出処理 一般的に室内環境下で取得されたデータを用いる場合, 照明などの影響により口唇の下部に陰影の生じることが多い。そこで提案手法では, 口唇形状における抽出精度の向上を目的とし, 単独画素の情報だけではなく, 近傍画素との関連情報も有効な特徴量になると考え, 以下の処理を施した。

- (1) 図5に示すように口唇周辺領域を座標 $C(Xc, Yc)$ の中

図4 Xwd,Ywdおよび座標C

Fig. 4. Feature width Xwd, and Ywd, and point C.

図5 口唇周辺領域

Fig. 5. Region around the lip.

心として領域A~Dの4つに分割した。

- (2) 近傍画素との関連情報としては、明度変化に対してロバスト性を有することに着目し、 $L^*a^*b^*$ 空間のメトリック色相差 ΔH_{ab}^* (16)を用いた。 ΔH_{ab}^* は(12)式で定義される。

$$\Delta H_{ab}^* = \left[(\Delta E_{ab}^*)^2 - (\Delta L^*)^2 - (\Delta C_{ab}^*)^2 \right]^{1/2} \quad (12)$$

ここで、 ΔE_{ab}^* は色差、 ΔL^* は明度指数 L^* の差、ならびに ΔC_{ab}^* はメトリッククロマ C_{ab}^* の差であり、(13)式~(15)式でそれぞれ定義される(16)。

$$\Delta E_{ab}^* = \left[(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2 \right]^{1/2} \quad (13)$$

$$\Delta L^* = L_2^* - L_1^* \quad (14)$$

$$\Delta C_{ab}^* = C_{ab2}^* - C_{ab1}^* \quad (15)$$

ここで、

L_1^* : 着目画素のメトリック明度、

L_2^* : 近傍画素のメトリック明度、

C_{ab1}^* : 着目画素のメトリッククロマ、

C_{ab2}^* : 近傍画素のメトリッククロマ

である。

口唇周辺領域に存在する各画素(以下、着目画素と略記する)と、その近傍8画素の内の2画素(以下、それぞれ近傍画素a、近傍画素bと表記する)それぞれの ΔH_{ab}^* を算出するため、図6に示す3x3のマスクA~Dを設定した。以下、着目画素と近傍画素aのメトリック色相差を ΔH_{aba}^* 、着目画素と近傍画素bのメトリック色相差を ΔH_{abb}^* と表記する。

- (3) 上記処理(2)で設定したマスクA~Dを図7に示す様に、領域AではマスクAを口唇周辺領域の左上、領域BではマスクBを右上、領域CではマスクCを左下、ならびに領域DではマスクDを右下からそれぞれ横方向に

図6 マスクA~D

Fig. 6. Mask A to D.

ラスト走査した。領域A~Dにおいて、各マスクの着目画素、近傍画素aおよび近傍画素bに対応する口唇周辺領域の画素を対象として図8に示す流れ図に従い、クラスを推定した。

- (4) 上記処理(1)~(3)によりクラス分類された画素を対象とし、(2・4)節の処理手順に従い口唇を抽出した。

口唇周辺領域における処理の流れは図1(g)~図1(k)のようになる。ここで、図1(i)は(2・4)節の処理により2値化処理を施した例、図1(j)は口唇抽出結果例、図1(k)は(2・5)節の処理によりヒストグラムを作成し口唇の横幅と縦幅を設定した例をそれぞれ示している。

3. 使用データおよび比較手法

(3・1) 使用画像データ 被験者15名(a~o)が単語「秋田(あきた)」と1回ずつ発話した画像データをCCDビデオカメラ(SONY製:DCR-TRV900)により取得した。1回の発話に伴うフレーム数は被験者により異なるが約50~80フレームであった。なお、2.3節のファジィ推論に用いる教師データは、被験者毎に画像データのフレームNo.1より取得した。

本論文で用いた画像データの取得条件を以下に示す。また、取得画像データの概要を表1に示す。

- ① 模様のない様な壁の前で被験者が椅子に座った状態で撮影
- ② 被験者および壁とビデオカメラの位置をそれぞれ距離1.2mおよび1.8mに固定
- ③ 日常一般的と考えられる蛍光灯による照明下(照度150lx以上の室内)
- ④ 被験者は眼鏡をはずし、口ひげはない

(3・2) 比較に用いた明度情報による抽出法 本論文では蛍光灯下で取得された画像を対象としている。このため、鼻および下唇の下部にはそれぞれ陰影が生じる。また、提案手法では明度変化に対するロバスト性を高めるため、明度情報を除去した色彩情報に着目するとともに、照明変化などにより生じる各特徴量の変動をあいまいさと仮定している。

そこで、提案手法の有用性を検討するため、時系列顔画像を用いた読唇の研究において、特徴量として用いられている明度情報に着目した口唇抽出法(3)~(5)により得られた結

果との比較を行った。

なお、本論文における明度情報に着目した手法とは、顔画像をグレースケールに変換し、次に、提案手法で設定した口唇周辺領域内において、オペレータが最適であると判

表 1 画像データの概要

Table 1. The details of image data for analysis.

frame size	320×240pixel
composition of data	RGB 8bit
total number of frames	974 frames

図 7 マスクの走査方向

Fig. 7. Direction in each mask for scanning.

図 8 口唇周辺領域におけるクラス分類処理の流れ

Fig. 8. Flowchart of the classification procedure in the region around the lip.

断したしきい値（輝度値）を試行錯誤的に手動で設定し 2 値化処理を施したものである。

4. 実験結果および検討

発話に伴う時系列的な口唇抽出を目的とする場合、口唇に関する情報として位置および形状を把握することは重要である。そこで本論文では、口唇の位置および形状を評価指標として着目し検討を加えた。

提案手法、ならびに明度情報に着目した手法により得られた口唇抽出結果の例を図 9 に示す。なお、各抽出結果から口唇周辺領域（2・5 節参照）を切り出したものを示している。明度情報に着目した手法（図 9(c)参照）では、陰影部分が口唇として誤抽出されているのに加え、下唇の中央部分が抽出されていないなど、口唇形状の正確な抽出は困難であることが分かる。これに対し、提案手法（図 9(b)参照）では良好に口唇の抽出されている様子が認められる。

次に、被験者 a~o が/a/, /き/, /た/と発話した時の提案手法により得られた口唇抽出結果の一部を図 10(a)~(o)に示す。口唇位置に着目した場合、口唇は原画像から良好に抽出されていることが分かる。また、対象とした全ての画像（合計 974 枚）についても口唇位置が正しく抽出されていることを確認している。これは、口唇の位置や形状に関する情報を事前に対象画像に付加することなく、時系列顔画像から口唇の位置抽出が可能であることを示している。

口唇形状に着目した場合、提案手法により得られた結果では 974 例中 920 例（約 94%）のデータにおいて背景（白色の壁）を含む顔画像から良好に口唇形状を推定することができた。

ここで、本論文では口唇形状に着目した場合、下記の 3 つの条件を満たすものを抽出成功とみなした。

- ① 口唇外側輪郭が滑らかに抽出されている。
- ② 口唇における左右の両端広角部が輪郭に沿って抽出されている。
- ③ 上唇と下唇が分離することなく抽出されている。

また、残り約 6% のデータについても口唇輪郭の一部が欠損する事例などが存在するものの、口唇形状の把握は容易である。

図 9 口唇抽出結果の比較

Fig. 9. Comparison of the extraction results.

したがって、一般的な室内環境下で取得された発話に伴う時系列顔画像において、口唇以外の顔の部位や背景などを認識することなく、提案手法により口唇を良好に抽出可能であることが明らかとなった。

5. まとめ

本論文では、一般的な室内環境下で口紅の塗布など特別な処理を施すことなく口唇を抽出することを目的とし、フuzzy推論を用いた口唇抽出法を提案しその有用性について検討を加えた。本論文で得られた成果をまとめると次のようになる。

- (1) 特徴量として、知覚色度指数 a^* 、 b^* および心理メトリック量 h_{ab} を用いることにより、提案手法は顔面の肌および口唇をそれぞれ抽出可能であることを明らかにした。
- (2) 肌は口唇と比較して発話に伴う時系列的な色味の変動が小さいことに着目した。さらに、抽出した肌領域の差

分として得られる画像にラベリング処理を施し口唇抽出を行った。照明の影響による色味の変化をあいまいさと仮定するとともに、着目画素と近傍画素の関連を考慮して処理を施したところ、提案手法は口唇を良好に抽出可能であることを明らかにした。

- (3) 提案手法は、口紅の塗布や照明の工夫といった特別な処理を必要とせず、通常の室内環境下で取得した時系列顔画像を対象として、口唇抽出が可能であることを明らかにした。

なお、本アルゴリズムを用いることにより、フレーム No.1 の画像から教師データを取得することで、口唇位置および形状の自動抽出が可能となった。ただし、本論文では教師データの自動取得について検討するまでには至らなかった。したがって、教師データの自動取得手法について、今後検討を加える予定である。また、被験者が同一画像中に複数存在する場合について検討するまでには至らなかったため、

図 10(1/2) 提案手法における口唇の抽出結果例

Fig. 10(1/2). Examples of extraction results of lip shape and its location with the proposed method.

図 10(2/2) 提案手法における口唇の抽出結果例

Fig. 10(2/2). Examples of extraction results of lip shape and its location with the proposed method.

複数の人が存在する時系列画像の解析手法についても、今後さらに検討を加える予定である。

最後に、日頃有益なご助言を賜っております本学情報工学科景山陽一講師、ならびに石沢千佳子助手に謝意を表します。

本研究は平成14年度共同研究(区分A)経費の支援の下に行ったことを付記する。

(平成14年6月10日受付, 平成14年11月25日再受付)

文 献

- (1) M.Ishi, K.Sato, M.Nishida, and Y.Kageyama: "Feature Extraction and Analysis for Lip Reading in Continuous Image", *TIEE Japan*, Vol.119-D, No.4, pp.465-472 (1999-4) (in Japanese)
石井雅樹・佐藤和人・西田 眞・景山陽一:「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 D, 119, 4, pp.465-472 (1999-4)
- (2) M.Konda, M.Nishida, M.Ishi, and K.Sato: "Application of the Feature of Lip Motion in Continuous Image to Personal Identification", *TIEE Japan*, Vol.120-C, No.5, pp.765-766 (2000-5) (in Japanese)
根田雅穂・西田 眞・石井雅樹・佐藤和人:「口唇の動き特徴の個人識別法への適用」, 電学論 C, 120, 5, pp.765-766 (2000-5)
- (3) K.Matsuoka, T.Furuya, and K.Kurosu: "Speech Recognition by Image Processing of Lip Movements", *SICE Japan*, Vol.22, No.2, pp.191-198 (1986-2) (in Japanese)
松岡清利・古谷忠義・黒須頭二:「顔画像処理による読唇の試み」, 計測自動制御学会論文集, 22, 2, pp.191-198 (1986-2)
- (4) K.Uchimura, J.Michida, M.Tokou, and T.Aida: "Discrimination of Japanese Vowels by Image Analysis", *IEICE Japan*, Vol.J71-D, No.12, pp.2700-2702 (1988-12) (in Japanese)
内村圭一・道田純治・都甲昌美・相田貞蔵:「画像解析による日本語母音の識別」, 信学論 D-II, J71-D, 12, pp.2700-2702 (1988-12)
- (5) H.morimoto, Y.Lee, K.Kamei, and K.Inoue: "Extraction of LipRegion in Consideration of Different Subjects and Experimental Environment", *IEICE Japan*, Vol.J79-D-II, No.1, pp.139-141 (1996-1) (in Japanese)
森本広頼・李 亜兵・亀井旦有・井上和夫:「個人差や撮影環境を考慮した濃淡画像からの口唇領域の抽出」, 信学論 D-II, J79, pp.139-141 (1996-1)
- (6) Y.Adachi, A.Imai, M.Ozaki, and N.Ishii: "Study on Extraction of Flesh-Colored Region", *TIEE Japan*, Vol.120-C, No.12, pp.1927-1932 (2000-12) (in Japanese)
足達義則・今井明宏・尾崎正弘・石井直弘:「肌色領域抽出手法の検討」, 電学論 C, 120, 12, pp.1927-1932 (2000-12)
- (7) T.Yokoyama, K.Tanaka, K.Hisatomi, Y.Yagi, M.Yachida, F.Hara, and S.Hashimoto: "Extracting Contours and Features from Frontal Face Images", *ITE Japan*, Vol.53, No.11, pp.1605-1614 (1999-11) (in Japanese)
横山太郎・田中克明・久富健介・八木康史・谷内田正彦・原 文雄・橋本周司:「正面顔画像からの輪郭特徴抽出」, 映像学誌, 53, 11, pp.1605-1614 (1999-11)
- (8) Y.Abe and M.Hagiwara: "Analysis and Recognition of a Human Head's Movement from an Image Sequence", *IEICE Japan*, Vol.J83-D-II, No.2, pp.601-609 (2000-2) (in Japanese)
安部友一・萩原将丈:「単眼視動画像からの人物頭部動作の解析と認識」, 信学論 D-II, J83-D, 2, pp.601-609 (2000-2)
- (9) H.Wu, Q.Chen, and M.Yachida: "Face Detection from Color Images by Fuzzy Pattern Matching", *IEICE Japan*, Vol.J80-D-II, No.7, pp.1774-1785 (1997-7) (in Japanese)
呉 海元・陳 謙・谷内田正彦:「ファジィパターン照合を用いた色彩画像からの顔検出システム」, 信学論 D-II, J80-D, 7, pp.1774-1785 (1997-7)
- (10) X.Song, C.Lee, C.Xu, and S.Tsui: "Extraction of Facial Organ Features Using Partial Feature Template and Global

Constraints", *IEICE Japan*, Vol.J77-D-II, No.8, pp.1601-1609 (1994-8) (in Japanese)

宋 欣光・李 七雨・徐 剛・辻 三郎:「部分特徴テンプレートとグローバル制約による顔部品特徴の抽出」, 信学論 D-II, J77-D, 8, pp.1601-1609 (1994-8)

- (11) H.Ilongo and K.yamamoto: "Face and Feature Detection Using Skin Color and Motion", *ITE Japan*, Vol.52, No.12, pp.1804-1847 (1998-12) (in Japanese)
本郷仁志・山本和彦:「動領域内の肌色推定による顔領域および顔部品抽出」, 映像学誌, 52, 12, pp.1804-1847 (1998-12)
- (12) The Illuminating Engineering Institute of Japan: "Lighting Handbook", Ohmsha (1987)
照明学会編:「ライティングハンドブック」, オーム社(1987)
- (13) K.Asanuma, M.Onishi, A.Kojima, and K.Fukunaga: "Extracting Regions of Color and Region Tracking", *TIEE Japan*, Vol.119-C, No.11, pp.1351-1358 (1999-11) (in Japanese)
浅沼克紀・大西正輝・小島篤博・福永邦雄:「色情報と領域追跡情報を用いた人物の顔と手の領域の抽出」, 電学論 C, 119, 11, pp.1351-1358 (1999-11)
- (14) Y.Yokoo, and M.Hagiwara: "Fast Human Faces Detection Method From a Color Image", *TIEE Japan*, Vol.119-C, No.10, pp.1255-1261 (1999-11) (in Japanese)
横尾裕規・萩原将丈:「カラー画像からの高速顔領域抽出法」, 電学論 C, 119, 10, pp.1255-1261 (1999-11)
- (15) K.Takebe, N.Ito, S.Nakauchi, and S.Utsui: "A Digital Color Imaging with Color Constancy", *IEICE Japan*, Vol.J83-D-II, No.8, pp.1753-1762 (2000-8) (in Japanese)
竹部啓輔・伊藤直己・中内茂樹・臼井支朗:「色恒常性を有するディジタルカラーイメージング」, 信学論 D-II, J83-D, 8, pp.1753-1762 (2000-8)
- (16) The Color Science Association of Japan: "Handbook of Color Science [Second Edition]", University of Tokyo Press (1998)
日本色彩学会編:「新編 色彩科学ハンドブック (第2版)」, 東大大学出版 (1998)
- (17) M.Mizumoto: "Fuzzy Reasoning (1)", *SOFT-J*, Vol.4, No.2, pp.256-264 (1992-4) (in Japanese)
水本雅晴:「ファジィ推論(1)」, 日本ファジィ学会誌, 4, 2, pp.256-264 (1992-4)
- (18) M.Takagi and H.Shimoda: "Handbook of image analysis", University of Tokyo Press (1991)
高木幹雄・下田陽久:「画像解析ハンドブック」, 東大大学出版(1991)

白 澤 洋 一

(学生員) 1978年4月6日生。2001年3月秋田大学鉱山学部情報工学科卒業。同4月秋田大学大学院鉱山学研究科博士前期課程情報工学専攻入学。2003年3月秋田大学大学院鉱山学研究科博士前期課程情報工学専攻修了。同4月秋田大学大学院工学資源学部研究科博士後期課程電気電子情報システム工学専攻入学。現在に至る。画像処理による発話認識および個人識別の研究に従事。2001年情報処理学会東北支部学生奨励賞受賞。情報処理学会会員。

西 田 眞

(正員) 1974年3月秋田大学鉱山学部電気工学科卒業。同年4月トヨタ自動車工業(株)入社。75年4月秋田大学鉱山学部助手。講師, 助教授を経て, 96年1月同教授。現在, 秋田大学工学資源学部情報工学科教授。リモートセンシングデータの解析技術とアルゴリズム, 画像情報応用および知識情報システムの研究に従事。88年~89年クラークソン大学客員研究員。84年電気学会論文賞, 88年静電気学会論文賞, 98年国際素材物性学会議優秀論文賞受賞。ファジィ学会, リモートセンシング学会, 写真測量学会, 応用物理学会, 日本素材物性学会, 情報処理学会, 照明学会, IEEE 会員。工博。