

5E-4

表情コミュニケーションのための 事例に基づく表現モデルの作成

近藤崇† 角所考‡ 美濃導彦‡

† 京都大学工学部情報学科

‡ 京都大学総合情報メディアセンター

1 はじめに

ネットワークコミュニケーションなどにおいて表情情報を送信しようとする際、匿名性を保存したい場合や、サイズの大きな実画像の表示が困難な多人数でのコミュニケーションなどの場合にはユーザの表情を合成顔画像で置き換えて送信する必要が生じる。

顔の造作や各部位の変化には個人差があるため、ユーザの実画像（客観表情）の変化に忠実に合成顔（合成表情）を合成しても、それがユーザの本当に意図したものと一致するとは限らない。よってユーザの客観表情と合成表情との対応関係を定める表情表現モデルが必要である。

本研究では、いくつかの合成表情に対して、ユーザがそれに対応すると思うユーザ自身の客観表情を事例として蓄積し、これらの事例をユーザが望む合成表情と客観表情との間の対応関係を定める表情表現モデルとすることを目指す。

2 事例に基づく表情表現モデル

2.1 表情のパラメータ表現

ユーザのとっている表情をカメラで観測し、その顔画像（客観表情）を適当なパラメータを用いて n 次元ベクトル $E = [x_1, \dots, x_n]$ として表現する。一方、合成表情を表情変形のパラメータを用いて m 次元ベクトル $D = [y_1, \dots, y_m]$ として表現する (n, m はパラメータの次元数)。このとき、客観表情と合成表

情はそれぞれ n 次元、 m 次元のパラメータ空間中の一点として表せる。また、 m は n 以下であるとする。

2.2 事例の対話的獲得

客観表情と合成表情の対応の事例を以下のようなユーザとのインタラクションを通じて対話的に獲得する。

1. システムが合成表情のパラメータ空間中の点のサンプリングによって合成表情 D_i を生成し、ユーザにそれに対応すると思う表情 E_i をとってもらう ($i = 1, \dots, s$)。この E_i と D_i のペアを事例として蓄積する。(2の未知客観表情に対応する合成表情の推定のために $s \geq m + 1$ が必要)
2. ユーザがシステムに対し表情 E を示す。 E に対しユーザが望む真の合成表情 D を 2.3 節の手順に従って推定しこれを \hat{D} とする。ユーザが \hat{D} を自分が意図した表情として許容できるとすれば、 D はすでにある事例から推定可能であるので新たな事例は追加しない。
3. ユーザが \hat{D} を自分が意図した表情ではないとした場合はユーザが直接合成顔を編集して意図した表情 D を作成する。 E を D と対応付け、新しい事例として蓄積する。
4. 2 と 3 を繰り返す。

2.3 事例間の補間による合成表情の推定

$m+1$ 個以上の合成表情に対して、対応する客観表情が事例として得られているものとする。客観表情のパラメータ空間中で未知の客観表情 E との距離が最も小さい $m+1$ 個の事例を E_1, \dots, E_{m+1} とし、これに対応する合成表情をそれぞれ D_1, \dots, D_{m+1} とする。

“Case-Based Model to Synthesize Facial Expressions for Network Communication.”

KONDO Takashi, KAKUSHO Koh, MINOH Michihiko

† Department of Information Science, Kyoto university

‡ Center for information and multimedia studies, Kyoto university

Email kondo@kuis.kyoto-u.ac.jp

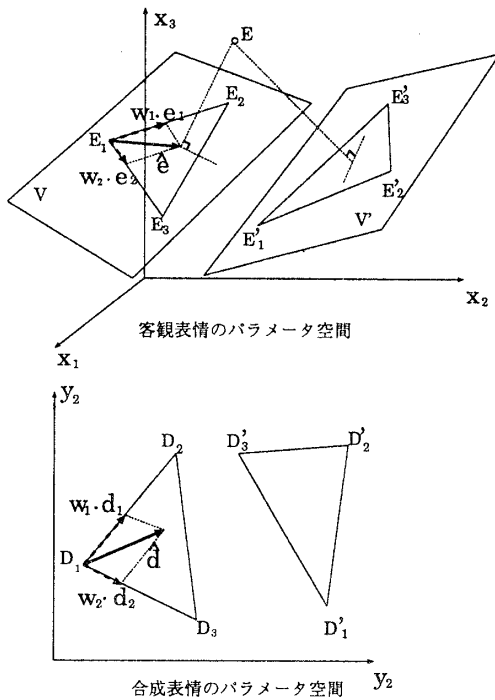


図 1: 補間による合成表情の推定 ($n = 3, m = 2$)

さらにベクトル $e_1, \dots, e_m, d_1, \dots, d_m, e, d$ を次式のように定義する。ただし、以下で定義されるベクトル e_1, \dots, e_m が一次独立であるとする。

$$\begin{aligned} e_i &= E_{i+1} - E_1, & e &= E - E_1 \\ d_i &= D_{i+1} - D_1, & d &= D - D_1 \end{aligned} \quad (1)$$

次に e を e_1, \dots, e_m の張る R^n の部分空間 $V = \langle e_1, \dots, e_m \rangle$ へ直交射影し、この射影ベクトル \hat{e} を次のように表す。

$$\begin{aligned} \hat{e} &= [e_1, \dots, e_m] w \\ w &= [w_1, \dots, w_m]^T \end{aligned} \quad (2)$$

ここで得られるベクトル $w = [w_1, \dots, w_m]^T$ を補間の重みとして用い、 E に対応する D の推定ベクトル \hat{D} を次式で計算する。

$$\begin{aligned} \hat{d} &= [d_1, \dots, d_m] \underbrace{w}_{\text{補間の重み}} \\ \hat{D} &= \hat{d} + D_1 \end{aligned} \quad (3)$$

3 実験

実験では、顔の各部位に付けた8つのマーカーと瞳の中心の画像中での位置座標（鼻の頭につけたマーカーを原点とする。計20個）を客観表情のパラメータとする。一方、合成表情のパラメータは瞳の座標、左右の眉と唇の曲がり具合（計7個）とした。10種類の表情を事例として与え（図2は内4種類）、図3の左側の客観表情を与えたときの推定合成表情を右側に示す。主観的には、唇や右眉の曲がり少し不十分であるが、ある程度意図したものとなっている。今後は複数の被験者で同様の実験を行い有効性を確認した上で、事例を増やして実験を行う必要がある。

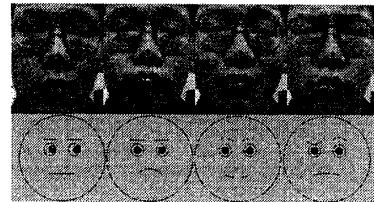


図 2: 実験で与えた事例 (4 種類)

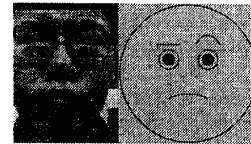


図 3: 実験結果

4 おわりに

本稿では、事例の対話的獲得と観測事例に対応する合成表情間の補間によりユーザの意図に合う表情表現モデルを作成するための手法を示した。現在は客観表情のパラメータとして顔につけたマーカーの位置を用いているが、今後はユーザに負担のないものとして画像の画素値や表情認識の研究で提案されている各種の特徴 [1] の利用についても検討する予定である。また、効率的な事例蓄積のためのプランニング、より適切な部分空間の構成法についても検討していく予定である。

参考文献

- [1] 八木、森島他 “顔情報処理のための共通プラットフォームの構築”, 情処研報 CVIM110-9, 1998.