

表情シーケンス画像からの個別表情抽出に関する検討¹

4 V-7

— ある表情の表出過程における別の表情への変化が発生する場合 —

大塚 尚宏 大谷 淳

(株) ATR 知能映像通信研究所

1. はじめに

表情は内部的な感情の外部表現であると考えられ、外界からの刺激等によって感情が変化するとともに変化する。さらに、社会的なコンテキストにおいては人間の感情の変化は複雑かつ多様となり、このような場面において複雑な表情変化が発生するものと推測される。したがって、表情シーケンス動画からそれぞれ異なった表情カテゴリーの区間を抽出し認識すること(本稿では表情抽出と呼ぶ)は、顔動画像の知的符号化および計算機との表情を媒介とした対話を実現するためには重要な技術である。筆者らは隠れマルコフモデル(HMM)を用いた表情抽出手法を提案し[1,2]、6種類の基本表情を実時間で抽出できることを示した。そこでは、5状態からなるleft-to-right型のHMMにより、表情変化にとまって検出される特徴ベクトルの時間変化をモデル化した。HMMの各状態は表情筋の異なった収縮状態に対応させ、緩和(無表情)→収縮→表情表出→伸長の順に遷移して初期状態に戻るように構成した。したがって、無表情から始まりある単一の表情が表出されて無表情に戻るといったシーケンスしか抽出することができず、ある表情が表出されている状態から無表情を介さずに別の表情に変化するようなシーケンスを抽出することはできなかった。

本稿では、従来法により表情毎に独立に構成されたHMMをもとに、任意の2つの表情の直接変化に対応する状態およびそれらの状態と元のHMMの表情表出状態と間の遷移を付加することにより、ある表情の表出過程における別の表情への変化が発生する場合にも表情を正しく抽出できる手法を提案する。

2. 表情抽出のためのHMMの構成法

HMMを用いた表情抽出では、前処理により得られる特徴ベクトルの時系列から時刻毎に状態の分布確率を求め、分布確率が最大となる状態をその時刻に占有されている状態と推定する。さらに、占有されている状態が

¹A Study of Extracting Regions Displaying Facial Expressions from Image Sequences — Case of direct transition from one facial expression to another one —

Takahiro Otsuka, and Jun Ohya

ATR Media Integration & Communications Research Labs.

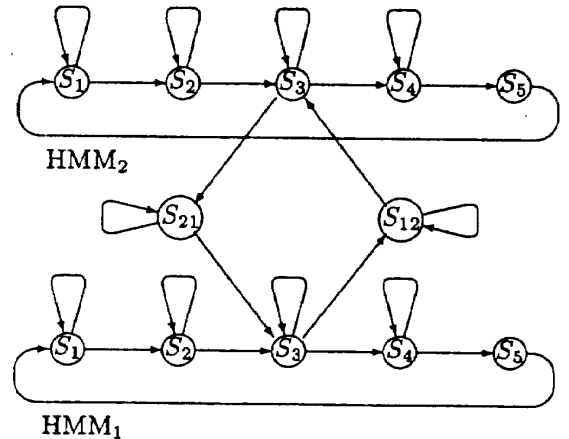


図 1: HMM の構成 ($N = 2$ の場合)

ある表情の表出状態である場合には、対応する表情が表出されているものと抽出する。時刻 t における状態 S_i の分布確率 $P_i(t)$ は以下のフォワードアルゴリズムにより求められる。

$$P_i(t) = b_i(O_t) \sum_{j=1}^N P_j(t-1) a_{ji} \quad (1)$$

ここで、 a_{ji} は状態 S_j から S_i への遷移確率、 $b_i(O)$ は状態 S_i にあるときに特徴ベクトル O が出力される出力確率分布、 O_t は時刻 t において検出される特徴ベクトル、 N はカテゴリー数である。また、出力確率分布 $b_i(O)$ は次式のように正規分布で近似できるものと仮定する。

$$b_i(O) = \mathcal{N}(O, \mu_i, \sigma_i) \quad (2)$$

ここで、 μ_i, σ_i は正規分布の平均、標準偏差である。以下では、 $P_i(t)$ をフォワード確率と呼ぶ。

フォワード確率はすべての状態からの寄与の総和として定義されるが、各状態に意味を付与することにより特定の状態間の遷移のみが可能となるようにすることができる。

無表情を介さない表情変化を抽出するためには、図 1 に示す HMM を用いる。ここでは、 N が 2 の場合として、1 番目および 2 番目のカテゴリーの表情が単独に表出されるシーケンスを抽出するループ HMM₁, HMM₂

と、1(2)番目のカテゴリーから2(1)番目のカテゴリーに直接変化する際に占有する状態 S_{12} (S_{21}) および HMM₁, HMM₂ の (表情表出) 状態 S_3 へのリンクとから構成される HMM を示す。

HMM のパラメータとしては、ループ HMM_i に関するものは単一の表情シーケンスから学習されたパラメータを用いる。ただし、状態 S_3 からは状態 S_{ij} へのリンクが追加されているので、もとの遷移確率 a_{34} を2つの状態への遷移確率 a_{34} と $a_{3,ij}$ に2等分する。状態 S_{ij} に関するパラメータは、単一の表情シーケンスから学習されたパラメータから算出する。これにより、すべての組み合わせの表情変化のシーケンスを用いてパラメータを学習する必要はなくなる。出力確率分布はカテゴリー i の筋肉伸長とカテゴリー j の筋肉収縮とが同時に起こるものと仮定して次式により求める。

$$b_{ij}(O) = \mathcal{N}(O, \mu_4^{(i)} + \mu_2^{(j)}, \sigma_4^{(i)} + \sigma_2^{(j)}) \quad (3)$$

ここで、 $\mu_j^{(i)}$, $\sigma_j^{(i)}$ はカテゴリー j の状態 S_j の出力分布密度の平均、標準偏差である。一方、遷移確率は次式に示すように、筋肉伸長と筋肉収縮状態のうち継続時間が長い方の遷移確率を選択する。

$$a_{ij,ij} = \max(a_{44}^{(i)}, a_{22}^{(j)}) \quad (4)$$

$$a_{ij,3} = 1 - a_{ij,ij} \quad (5)$$

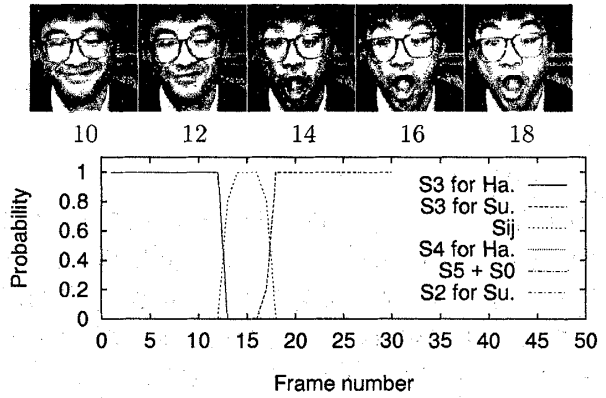
3. 実験結果と考察

本稿で提案した手法を評価するために Indy ワークステーション (SDI 社製, 180MHz) を用いて実験を行った。図2に‘幸福’から‘驚き’に変化するシーケンスの抽出結果例を示す。表情変化が遅い場合にはループに沿って状態遷移しているのに対して、表情変化が速い場合には筋肉の収縮と伸長が同時に起こる状態を通っている。

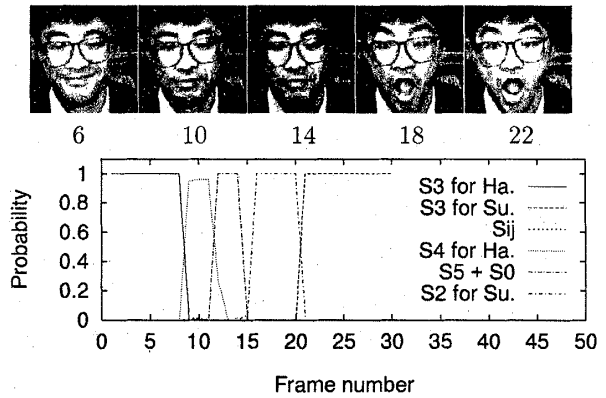
図3に抽出率の表情変化速度への依存性を示す。‘幸福’と‘驚き’とを交互に表出し、2つの表情の表情変化のピーク間の時間差をパラメータとし、抽出率(2つの表情に正しく抽出される割合)を求めた。速い表情変化に対しても追従できることが分かる。

4. むすび

表情シーケンス画像からの個別表情抽出手法をある表情の表出過程における別の表情への変化が発生する場合に拡張した。第一の表情の筋肉伸長と第二の表情の筋肉収縮とが同時に起きるものと仮定し、各表情を抽出する HMM に新たな状態と元の HMM の状態とのリンクを付加した。付加された状態およびリンクに関するパラメータは、上記の仮定に基づいて元の HMM のパラメータから算出することができ、すべての組み合わせのデータを用意してパラメータを学習する必要はない。評価実験を行い速い表情変化に対しても正しく抽出できることが分かった。



(a) 表情変化が速い場合



(b) 表情変化が遅い場合

図2: ‘幸福’から‘驚き’に変化するシーケンスの抽出結果例

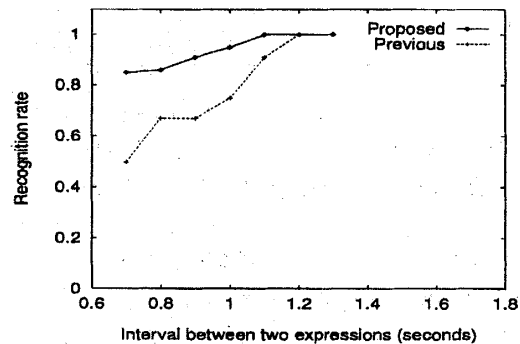


図3: 抽出率の表情変化速度への依存性

[参考文献]

- [1] 大塚、大谷：HMMを用いた動画シーケンスからの個別表情抽出に関する検討，第55回情全大，Vol. 2, pp. 335-336 (1997-9).
- [2] 大塚、大谷：連続した表情シーケンス画像からのHMMを用いた個別表情抽出に関する検討，信学技報，PRMU-97-154 (1997-11).