

顔動画における頑強な特徴点追跡の一手法

野村 有加 (指導教員：伊藤 貴之)

1. 概要

近年、コンピュータによる人間の顔表情の認識に関するさまざまな研究が発表されている。その研究の多くは、コンピュータが人の顔領域、顔器官、顔特徴点を認識し、それらを動画像において追跡しその変化を観察することを目標としている。この技術によって、対象人物の表情、心理、動作などを理解することが可能になると考えられる。

従来手法の一例として、顔動画像から抽出した複数の特徴点を動画像において追跡し、対象人物の表情を解釈する手法について議論する。この手法において問題となる点の一つに、顔器官の急激な動き、顔の向きの変化、特徴点の隠蔽、などに対する精度や頑強さがあげられる。近年では、この点を考慮して制度頑強性を重視した様々な手法が提案されているが、未だ議論の余地があると考えられる。そこで本報告では、既存研究での顔特徴点追跡の手法を応用し、ある一定時間にわたって撮影した顔動画像を用いて、従来研究での顔特徴点追跡の手法を応用し、顔動画像における特徴点追跡を頑強にするための一手法について提案する。

2. 関連研究

顔特徴点や特定部位の抽出や追跡に関する研究は、旧来から活発に発表されている[1-4]。これらの抽出や追跡を頑強にするためには、顔領域や顔特徴点の隠蔽への対策を議論する必要がある。これをうけて近年では、部分的な遮蔽に対応した顔領域の検出手法[5]が発表されている。また、顔特徴点の隠蔽に対する頑強さを実証した手法の一例として、特徴点を覆い隠しても追跡を続行できる手法[6]が発表されている。

本報告の提案手法は文献[6]の手法と同様に、顔特徴点抽出にテンプレートマッチングを適用し、動画像において直前時刻の抽出結果を参照し、その近傍領域において顔特徴点を追跡する。それに加えて本報告では、一部の顔特徴点検出結果の信頼性が低い場合に、残りの信頼できる特徴点を追跡し続け、その結果から信頼性の低い特徴点の位置を推定する、という方針で特徴点を追跡する手法を提案する。

3. 提案内容

本手法では、1台のデジタルビデオカメラで撮影された顔動画像内の注目すべき特徴点を、以下の手順によって追跡する。

1. まず動画像の最初のフレームにおいて、テンプレートマッチングで類似部分を検索し、特徴部位の位置を決定する。
2. 以降のフレームでは、1.の処理において得られた特徴点をもとに、特徴点追跡手法 Lucas-Kanade 法により追跡を行う。それに加え、以下の手順により、特徴点が発見できなかった場合の処理を行う。
 1. 前フレームでの特徴点位置の周辺領域において再度テンプレートマッチングを行う。

2. テンプレートと結果領域との類似度が低い場合は、特徴点が隠蔽状態にあると判断し、その適切な位置を推定する。

3.1 特徴点追跡の基本手法

表情変化を分析するために重要な顔特徴点を抽出し、動画像に対してそれらの追跡を行う。まず特徴の抽出には、あらかじめ用意したテンプレート画像を用いて、テンプレートマッチングで類似部分画像を検索する、これによって、顔特徴点の位置を決定する。ここで $M \times N$ 画素のテンプレート画像を、動画像から抽出された $M \times N$ 画素の部分領域に重ね合わせるとする。この重ね合わせる領域の左から i 番目、上から j 番目の画素値を $I(i, j)$ とし、 $T(i, j)$ をテンプレート画像の左から i 番目、上から j 番目の画素値とする。このときテンプレートマッチングには、以下の式(1)に示される類似度 R の算出式を用いる。

$$R = \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} |I(i, j) - T(i, j)| \dots(1)$$

この R 値が最小となるとき部分領域を類似部分画像とみなし、提案手法における顔特徴点の抽出結果とする。

一例として、文献[7]のデータベース上にある顔動画像において、テンプレートマッチングによって顔特徴点を抽出した結果を示す。ここでは図1(左)に示す被験者Aの顔動画像における最初のフレームから図1(右)で示すように両目尻と両口端を抽出してテンプレート画像とし、テンプレートマッチングと追跡点の初期設定を試みた。その結果を図2に示す。

そして次のフレーム以降は、特徴点追跡手法の一つである Lucas-Kanade 法により追跡を行う。Lucas-Kanade 法は、動画像における運動物体のオプティカルフローを局所勾配法によって推定する手法である[8]。

3.2 異常検出された場合の対応

提案手法では、前節で述べた顔特徴点の追跡において、検出された特徴点の信頼性を判定しながら処理を行う。提案手法では、以下の状況のどちらかが発生した場合に、その顔特徴点検出結果を異常検出であるとみなす。

- ・特徴点が検出されなかった場合。
- ・特徴点の初期状態と、移動後の点との周辺画素値の差分が、ある定数を越えた場合。

この場合において提案手法では、前のフレームでの特徴点位置を中心とする正方形小領域に限定したテンプレートマッチングを行う。この結果、式(1)の類似度 R の最小値が閾値を越えた場合は、検出結果の誤差が大きいとみなし、より広い領域での類似部分を検索する。この結果、式(1)の類似度 R の最小値が閾値よりも小さくなれば、類似度の高い結果が得られたとみなし、この検出結果を適切な検出結果であると判断する。この処理でもテンプレート画像に類似する位置が検出されない場合は、その顔特徴点は、手前の物体に隠蔽されているか、顔全体が横や

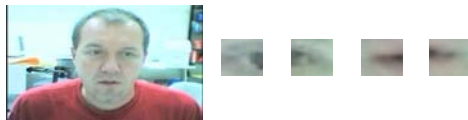


図1: 被験者Aの顔動画1フレーム目(左)と
テンプレート画像(両目尻・両口端)(右)



図2: テンプレートマッチング結果と追跡点決定(黄点)

下を向いている, などの理由で検出困難であると判断する. その場合には, 他の特徴点の移動距離などを参考に, 検出困難な特徴点の推定位置を算出する. この処理には, 文献[6]内で顔情報の計測で用いられている処理を用いる. 信頼性に基づく重み付けを考慮した処理を, 以下の評価式に従って行う.

$$E = \sum_{i=0}^{N-1} w_i | \mathbf{P} \mathbf{x}_i + \mathbf{t} - \mathbf{y}_i |^2 \dots(2)$$

ここで N は特徴点の数, \mathbf{x}_i は顔特徴点の位置ベクトル, \mathbf{y}_i は対応する特徴点が観測された位置ベクトル, w_i は観測の信頼性, \mathbf{t} および \mathbf{P} は並進ベクトルおよび回転行列, そして E はエラー値を表している. まず本手法では文献[6]と同様に, 式(2)における E を最小化する \mathbf{t}, \mathbf{P} を求める. 続いてこの \mathbf{t}, \mathbf{P} 値を用いて, 検出困難と判断された顔特徴点の位置を推定する.

以上の処理により, 顔特徴点がどのような理由で検出困難であるか判断できるようになると考えられる. また, 推定した顔特徴点の近傍領域にテンプレートマッチングを適用することで, 検出可能な状態に戻った顔特徴点を適切かつ迅速に検出できるようになると考えられる.

4. 実行結果

ここでは, 被験者 A の顔動画を用いて, 既存手法と本手法とを適用し, それぞれの結果の違いを示す. 図3は Lucas-Kanade

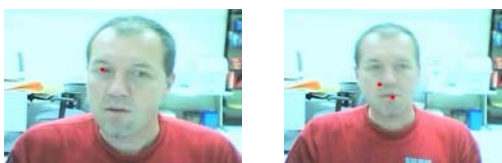


図3: Lucas-Kanade法のみ適用した追跡

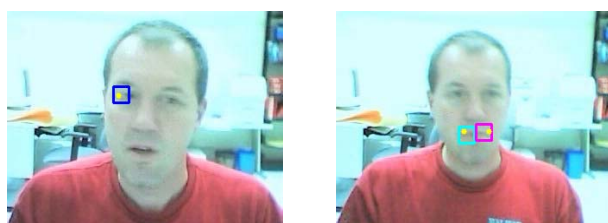


図4: 本手法を適用した追跡

法のみ適用した結果であり, 図4は本手法による結果画像である. 両図ともに同フレームでの結果を示した. 図3では赤点で示すように, フレーム15では右目尻, フレーム53では両口端が, それぞれ誤検出されている. それに対し, 本手法の結果である図4では, 再テンプレートマッチングにより, 特徴点が正しい位置(黄点)に検出されていることがわかる.

5. まとめと今後の課題

本報告は, 隠蔽などによる異常検出を考慮した顔特徴点追跡の一手法を提案した. 従来手法で見られる, テンプレートマッチングとオプティカルフロー推定による特徴抽出に加えて, 抽出された特徴点の信頼性に応じた処理を行うことで, 自然な状態の人物を対象にしても頑強な特徴点追跡が可能となる. この手法により, より自然な状態の人物を対象として, 顔特徴点を追跡できるようになり, より自然な状態から繊細な感情や心理状態を認識できると考えられる. この技術を利用することで例えば, 人間のコンピュータ使用時における疲労やストレスなどを認識し, ユーザの状態に合わせたコンテンツを提供するサービスや, またユーザ同士のコミュニケーションを支援するアバターなどに有効活用できると考えられる.

今後の課題として, 3.2節で述べた異常検出処理の処理時間や, 追跡結果の精度などについて, 検証を行う予定である.

謝辞

東京大学情報学環学際情報学府の山口泰教授には研究の全般的なアドバイスを頂きました. ここに感謝の意を表します.

参考文献

- [1] 福井, 山口, 「形状抽出とパターン照合の組合せによる顔特徴点抽出」, 電子情報通信学会論文誌 D-II, Vol. J80-D-II, No. 8, pp. 2170-2177, 1997. 8.
- [2] 林, 橋本, 鷺見, 「頑健性と精緻性を備えた顔特徴点追跡による顔方向推定」, 電子情報通信学会論文誌 D-II, Vol. J84-D-II, No. 8, pp.1762-1771, 2001. 8.
- [3] 岡, 佐藤 「頭部変形モデルの自動構築を伴う実時間頭部姿勢推定」画像の認識・理解シンポジウム(MIRU2005), 167-173, 2005.7
- [4] David Cristinacce and Tim Cootes, "Feature Detection and Tracking with Constrained Local Models", Proc. British Machine Vision Conference, Vol. 3, pp. 929-938, 2006.
- [5] K, Hotta et al. , "A Robust Face Detector under Partial Occlusion", IEEE International Conference on Image Processing, pp. 597-600, 2004.10.
- [6] 松本, 怡土, 竹村, 小笠原, 「リアルタイム顔・視線計測システムの開発と知的インタフェースへの応用」, 情報処理学会論文誌コンピュータビジョンとイメージメディア, Vol. 47, No. SIG15(CVIM16), pp. 10-21, 2006.10.
- [7] Institute for Information Technology, "NRC-IIT Facial Video Database", <http://synapse.vit.iit.nrc.ca/db/video/faces/cvglab/>
- [8] B. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision" Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI '81), April, 1981, pp. 674-679.