

オプティカルフローとスネークの協調処理による口唇運動の解析

○大久保 雅史 渡辺 富夫

岡山県立大学情報工学部

1. 緒言

機械読唇においては、口紅等の人工的な処理を施すことなく、口唇の輪郭を高精度に抽出する動画処理手法の開発が実用化の第一歩である¹⁾。

本研究の目的は、動的輪郭抽出法 (SNAKES)²⁾ と Optical flow³⁾ との協調的処理 (以下、Optical SNAKESという) により動画から口唇の輪郭を順次抽出し、時系列変化を解析することで読唇を試みるものである。本報告では、開発した Optical SNAKES の概略と読唇の基本である単母音の認識に対して Optical SNAKES を適用し、単母音発声時の口唇輪郭の時系列変化の解析結果について述べ、本手法の有効性を示す。

2. システムの概要および実験方法

本システムでは、SNAKES の初期輪郭を直前の画像に対する SNAKES の結果得られた各点の座標値と、2つの画像から求められる Optical flow とにより決定する。本システムを利用した口唇の輪郭抽出実験の概念図を Fig. 1 に示す。

ビデオカメラまたはビデオから顔画像 (1秒間に30フレーム) を取り込む。すべての画像に対して1フレーム前の画像との Optical flow を求める (①, ②)。オペレータが1枚目の画像に対してのみ口唇を取り巻くように SNAKES の初期輪郭をマウスを使ってディスプレイ上で与える。

本システムは、与えられた初期輪郭を基に1枚目の口唇の輪郭を抽出する。2枚目以降は、直前の画像の口唇輪郭 (③) と Optical flow (④) からその画像での初期輪郭を求め SNAKES により口唇の

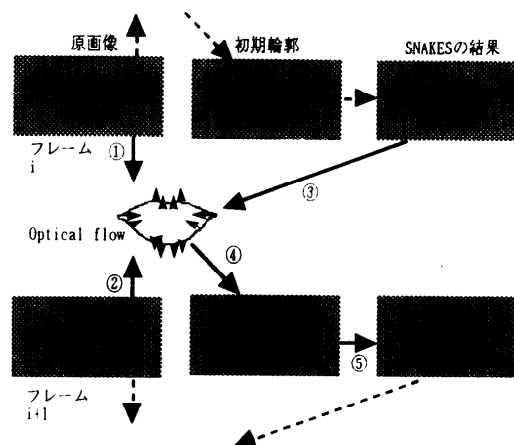


Fig. 1 SNAKES と Optical flow の協調的処理の流れ

輪郭を順次抽出する (⑤)。以上の結果として各フレームでの口唇の輪郭を構成する点群の座標が得られる。

3. 口唇輪郭の時系列変化の解析

読唇の基本の一つとして単母音の認識が挙げられる⁴⁾。そこで、上記システムを用いて5母音の発声による口唇輪郭の時系列変化を調べた。

5母音を連続的に発声した場合の口唇輪郭の時系列変化を Fig. 2 に示す。この図から5母音の発声がどのあたりのフレームでなされたかが、推測可能である。つぎに、得られたデータから静止画の読唇の研究でよく用いられるパラメータ: 口唇の横幅 (Width)・高さ (Height)・周囲長 (Length) 面積 (Area) について解析を行った。結果を Fig. 3 (a) ~ (c) に示す。このように、動画処理から得られるデータを用いることにより、静止画では得ることが非常に煩雑となる母音の発声からつぎの母音の発声への口唇の変形が簡便に求められる。

Analysis of Lip Movement by Concurrent Processing between Optical Flow and Snakes
Masashi Okubo, Tomio Watanabe
Okayama Prefectural University
111 Kuboki, Soja, Okayama, 719-11, Japan

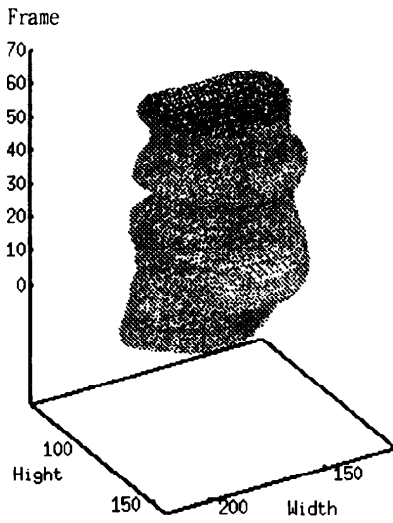


Fig. 2 5母音の発声による
口唇の時系列変化

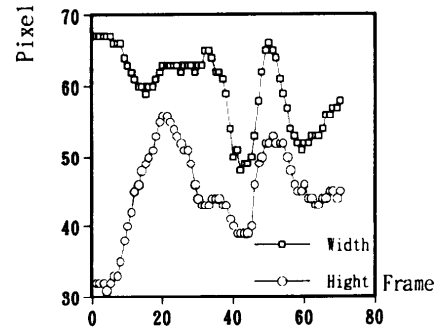
4. 結言

本研究の結果、Optical SNAKESは比較的正確に口唇輪郭を連続的に得ることが可能であること、口唇輪郭の時系列変化は、読唇の研究に対して新たなパラメータを供与し、認識率を上げる可能性が高いことを示した。

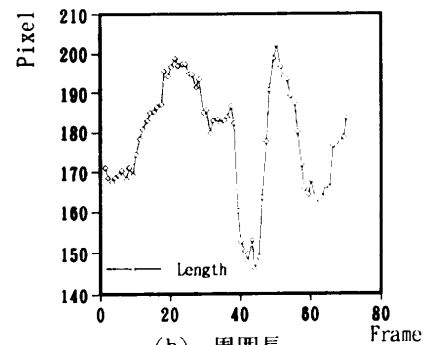
なお本研究でのシステム開発は、著者らの参加する科学技術庁省際基礎研究「霊長類行動のマルチメディア情報処理とニューロコンピュータによる解析システムの開発」プロジェクトによる。

参考文献

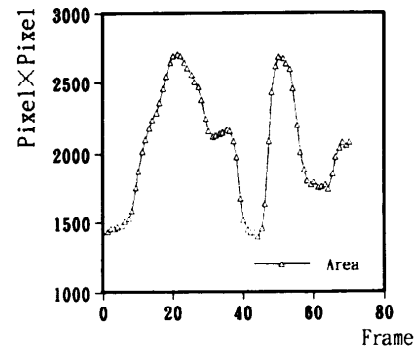
1) 中野政身・渡辺富夫：ステレオ視による母音口形の識別，日本機械学会論文集（C編），60巻570号，pp. 525/530（1994）．
 2) Michael KASS, Andrew WITKIN and Demetri TERZOPOULOS：Snakes：Active Contour Models, International Journal of Computer Vision, pp. 321/331（1988）．



(a) 横幅と高さ



(b) 周囲長



(c) 面積

Fig. 3 口唇の時系列変化の解析結果

3) G. ADIV：Determining tree dimensional motion and structure from optical flow generated by several moving objects, IEEE Trans. Vol. PAMI 7, No. 4, pp. 384/401（1985）．
 4) 渡辺富夫：読唇による母音の自動認識，日本機械学会論文集（C編），53巻496号，pp. 2613 -2616（1987）．