

最近の技術から

2段階処理による局所運動検出モデル*

緒形昌美

ソニー(株)総合研究所 AV 第1部 〒141 東京都品川区北品川 6-7-35

1. まえがき

視覚系から得られる動きの情報は、生体が周囲の状況を把握する上で非常に有効な手掛かりとなる。このため、生体における運動視知覚メカニズム、特にその初期レベルにあたる局所運動検出のメカニズムについては、心理学や生理学の分野で古くから研究がなされ、優れた数理モデルも幾つか提案してきた。また工学の分野でも、柔軟な視覚情報処理能力を有する機械を実現すべく多くの運動検出法が検討されてきた。

これらの従来法は、大きく二つのカテゴリー、すなわちフィルタリング的アプローチ、およびマッチング的アプローチに分類できる。前者は生体モデルに多く採用されており、後者が抱える対応問題を解く必要はないが、フィルタの出力値が正しく速度に対応するためには、フィルタの時空間的広がりに対応する範囲にわたって速度が一定である必要がある。もしその範囲で速度が変化してしまえば大きな誤差が生じる^{1,2)}。一方マッチング法は、基本的には点ごとの処理であり、速度変化に強いが、これまでの研究からもわかるように特徴点間の対応付けを正しく行うことは容易ではない。本稿ではこれら二つのアプローチを効果的に組み合わせた新しい局所運動の検出法について述べる³⁾。

2. フィルタリング処理

視覚から得られる入力情報は、図1(a)に示すような2次元画像の時系列と考えられる。このような時空間データに対し、スカラセンサー¹⁾とよばれる狭帯域通過形の時空間フィルタを施す。スカラセンサーは空間位置(x, y)、および空間周波数(u_s, θ_s)に選択性を持っている。したがって、入力パターンの空間周波数成分が周波数領域においてほぼ一様に分布していると仮定すれば、センサーからの出力パターンはセンサー自体の形状同

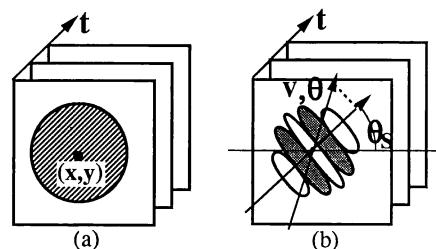


図1 入力画像(a), 出力画像(b)

様、 θ_s 方向に周波数 u_s で変化する正弦波状の格子縞となる(図1(b))。このパターンが速さ v で θ の向きに運動すると、位置 (x, y) には正弦波状の時間的輝度変化が生じ、その時間周波数 f_t は、

$$f_t = u_s \cdot v \cdot \cos(\theta - \theta_s) \quad (1)$$

となる。 u_s, θ_s は用いたセンサーによって決まるため、センサー出力の時間波形の周波数 f_t を検出することにより、(1)式は運動速度(v, θ)に関する拘束条件を与えることになる。しかし、実際に検出される時間周波数成分は、空間フィルタのバンド幅や、フィルタの積分範囲における速度変化等の影響で、通常は周波数軸上で広がりを持っている。したがって、速度に関する拘束条件は

$$f_{t\min} \leq u_s \cdot v \cdot \cos(\theta - \theta_s) \leq f_{t\max} \quad (2)$$

となる。ここで、 $f_{t\min}, f_{t\max}$ は、それぞれ検出された時間周波数の最小値、最大値である。スカラセンサーのパラメータ u_s, θ_s を変えることにより、異なる成分に関する拘束条件を得ることができるが、ここでは空間周波数 u_s が同じ成分は同じ速度で運動していると仮定し、 θ_s の異なるセンサーから得られた拘束条件の連立不等式の解として速度候補を検出し、それらすべてを次のマッチング処理に送る。

3. マッチング処理

フィルタリング処理によって得られた速度候補の中から、正しい速度を特徴点マッチングによって決定する。そのため、各時刻における入力画像から特徴点を抽出

* 本研究は筆者が(株)ATR 視聴覚機構研究所出向中に行なったものである。

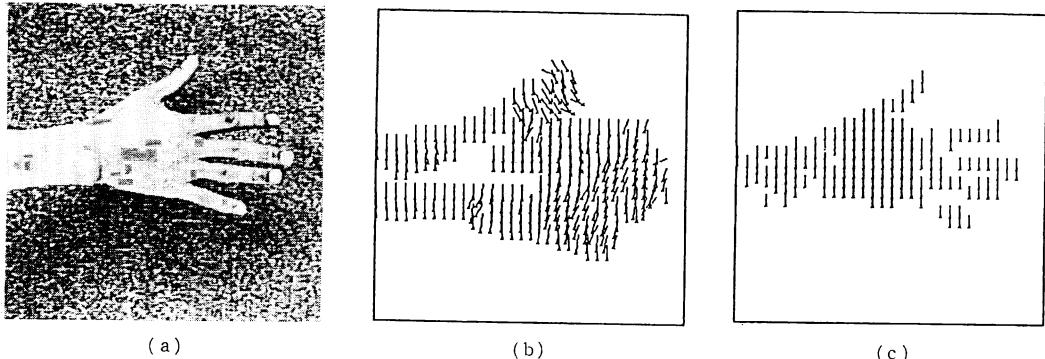


図 2 入力画像(a), フィルタリング処理(b), 2段階処理(c)

し、特徴画像 $F_n(i, j, k)$ を生成する。 n は特徴の種類を表し、複数の種類の特徴画像を用意しておく。用いる特徴としては、輝度の極大値や極小値、エッジに相当するゼロ交差点などが考えられる。特徴画像は、特徴 n が時刻 k における位置 (i, j) に存在する場合には 1、存在しない場合には 0 の値を持つ。この特徴画像を用い、時刻 k 、位置 (i, j) における速度が、速度候補中のある速度 $\mathbf{v} = (v_x, v_y)$ である可能性を次式により評価する。

$$S(\mathbf{v}) = \sum_{n=0}^{N-1} \text{NXOR}(F_n(i, j, k), F_n(i + v_x \Delta t, j + v_y \Delta t, k+1)) \quad (3)$$

ここで、 Δt はフレーム間の時間間隔であり、NXOR は排他論理和の否定を計算する。速度候補中のすべてに対してこの値 $S(\mathbf{v})$ を計算し、最大値を与える速度を時刻 k の位置 (i, j) に割り当てる。ただし、 $S(\mathbf{v})$ がある閾値以下の場合はいかなる速度も割り当てない。

4. 計算機実験

図 2(a) に入力時系列画像の第一フレームを示す。時間の経過とともに人間の手が真下に向かって移動していく。速さは約 0.9 pixels/frame である。周囲のテクスチャは静止している。図 2(b) にフィルタリング処理だけで求めた速度場を、図 2(c) に本方法によって求めた速度場を示す（推定された速度は下向き 0.75 pixels/frame）。手の輪郭にあたる部分でとくに精度が向上している。

5. むすび

従来独立に用いられてきたフィルタリング的アプローチとマッチング的アプローチを効果的に組み合わせ、より正確に局所運動を検出する方法について述べた。

最近の心理実験によると、生体の運動視知覚メカニズムにおいては線形処理と非線形処理が併用されていることが示唆されているが⁴⁾、このような方略が生体の優れた運動検出能力を支えているとも考えられる。本稿で述べた方法はこのフレームワークに沿ったものとなっており、その有効性は計算機実験によっても明らかであるが、最適なインプリメンテーションという点ではさらに検討が必要であろう。この意味で生体においてどのように二つの処理が併用されているかは興味深い。

文 献

- 1) A. B. Watson and A. J. Ahumada : "Model of human visual motion sensing," J. Opt. Soc. Am. A, 2 (1985) 322-341.
- 2) 緒形昌美、佐藤隆夫：“空間周波数チャンネル間の相互作用を考慮した運動視知覚モデル”，信学論(D-II), J73-D-II (1990) 207-214.
- 3) M. Ogata and T. Sato: "Motion-detection model with two stages: spatiotemporal filtering and feature matching," J. Opt. Soc. Am. A, 9 (1992) 377-387.
- 4) T. Sato: "Reversed apparent motion with random dot patterns," Vision Res., 29 (1989) 1749-1758.

(1992年6月1日受理)