

Elaborated Reichardt Detectorによる動きベクトル推定

高橋 陸良[†] 中村 栄治[†] 沢田 克敏[†]

生物の視覚情報処理の数理モデルである ERD (Elaborated Reichardt Detector) により、動画中の動きベクトルを推定するアルゴリズムについて提案する。従来の ERD による動き検出においては、動きの方向成分のみを求めることが主眼であったが、提案手法では動きの速さ成分をも含めた動きベクトルそのものを算出することを可能にした。空間および時間領域においてフィルタリングを行う ERD に含まれるフィルタパラメータを最適化することにより、動きベクトルの速さ成分を算出する手法である。数値計算シミュレーションにより提案するアルゴリズムを検証するとともに、ビデオ画像を用いた検証実験を行い、提案手法の有効性を確認した。

Motion perception using Elaborated Reichardt Detector

CHIKARA TAKAHASHI,[†] ELJI NAKAMURA[†] and KATSUTOSHI SAWADA[†]

This technical report presents a novel algorithm capable of estimating motion vectors based on the so called Elaborated Reichardt Detector or ERD which is a computational model for how visual information being processed in animals. The ERD has mainly been considered as a direction estimator of motion. However, the proposed algorithm has added to the ERD a speed estimator aspect by optimizing filter parameters in the temporal and space domain included in the ERD. The performance of our algorithm has been examined via various experiments such as computer simulations and motion analysis using real world video images in order to demonstrate how the algorithm can estimate motion vectors accurately.

1. はじめに

1.1 研究の背景

Elaborated Reichardt Detector (以下 ERD) は、生物視覚における運動の検出過程を表す数理モデルとして工学者 van Santen と心理学者 Geroge Sperling が提案したものである⁴⁾。これは脳科学者の Werner Reichardt が、視覚刺激に対する応答についてゾウムシを実験対象に研究し提案した Elementary Motion Detector (EMD と呼ばれる)³⁾ を基にしている。両論文は視覚情報処理の運動視を説明する理論として参照され、画像処理工学の面でもオプティカルフロー検出に応用されている。また、さらに幅広くビジョンチップ⁷⁾ や、視覚を持った小型飛行船²⁾ など、多数の工学的応用もなされている。著者らは、2000 年よりこの理論を拡張し動画像の任意方向の移動方向検出⁶⁾ や、速度検出⁸⁾ の手法、フォトトランジスタを入力デバイスとして用いた動画像認識アナログ回路の研究⁹⁾ を行っている。

1.2 研究の目的

従来、ERD は主に動きベクトルの方向成分の推定に利用され、ERD により大きさ成分を推定しようとの積極的な試みはなされて来なかった。本研究では、ERD により 1 次元信号の移動方向のみならず、移動速さの情報を持つ動きベクトルを推定する手法を提案する。特に、ERD を空間フィルタと時間遅延フィルタによる周波数解析の道具として捉え、これを理論的に検証する。対象とする観察信号は、輝度の時間変化なしに並進運動する輝度信号とする。つまり時間の経過とともに輝度分布やコントラストが変化してしまう 2 次運動¹⁰⁾ や、像面中の輝度パターンが小領域ごとに移動するローカルモーションは対象外とする。また、他のオプティカルフロー検出法と比較し、提案手法の持つ特徴や長所短所について考察する。

2. ERD について

2.1 ERD の観察信号

ERD の観察信号 L (Luminance) を定義する。まず、一次元信号 $L(x)$ を以下のようなフーリエ級数展開の形で表す。これは周波数 f_x の周期信号となる。 x は空間方向の座標を表す。 a_n は各周波数の正弦波の振

[†] 愛知工業大学大学院工学研究科
Graduate School of Engineering, Aichi Institute of
Technology

幅であり、 Φ_n は位相を表している。

$$L(x) = \sum_{n=0}^{\infty} L_n(x) = \sum_{n=0}^{\infty} a_n \sin(2\pi n f_x x + \Phi_n) \quad (1)$$

これが速度 dv で動くような $L(x, t)$ を ERD の観察信号と定める。 d (direction) は進行方向を表し、左方向の場合 $d = 1$ 、右方向の場合 $d = -1$ とする。 v は速度の絶対値となる。

$$L(x, t) = L(x + dvt) \quad (2)$$

$$= \sum_{n=0}^{\infty} a_n \sin(2\pi n f_x (x + dvt) + \Phi_n) \quad (3)$$

$$= \sum_{n=0}^{\infty} a_n \sin(2\pi n f_x x + 2\pi n f_x dvt + \Phi_n) \quad (4)$$

ここで時間周波数 $f_t = f_x v$ とおき、

$$L(x, t) = \sum_{n=0}^{\infty} a_n \sin(2\pi n f_x x + 2\pi n d f_t t + \Phi_n) \quad (5)$$

と書いてもよい。 $L(x, t)$ のうち、 $n = 0, 1$ の項である $L_{0,1}(x, t) = a_0 \sin(\Phi_0) + a_1 \sin(2\pi f_x x + 2\pi d f_t t + \Phi_1)$ (6)

を特に S (SinWave) とおき、 $\Phi_0 = \pi/2$ として変数名を変更し

$$S(x, t) = I_0 + a \sin(2\pi f_x x + 2\pi d f_t t + \Phi) \quad (7)$$

と定義する。このような信号を正弦波運動縞という。図1に例を示す。この図は空間周波数 $f_x = 0.05$ 、速度 $dv = 1$ 、幅 = 100 pixel、高さ = 100 pixel にて生成した左方向に並進運動する正弦波である。

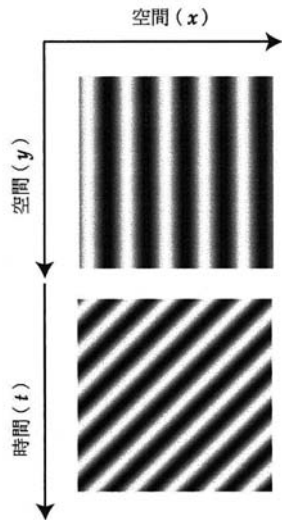


図1 正弦波運動縞 (上) と、その時空間プロット (下)

2.2 ERD の各ステップ

ERD 法による信号処理の各ステップについて述べる。各ステップは空間フィルタ FSA(Filter for Spatial Average)、時間遅延フィルタ FTD(Filter for Time Delay)、乗算、時間平均 FTA(Filter for Time Average)、減算の5ステップに分かれている。図2に、ERD の構造を示す。 \otimes は乗算を、 \oplus は加算処理を表している。 $y_{l,0}(t)$ 、 $y_{r,0}(t)$ は FSA からの出力信号、 $y_{l,1}(t)$ 、 $y_{r,1}(t)$ は FTD からの出力信号、 $y_{l,2}(t)$ 、 $y_{r,2}(t)$ は乗算処理後の信号、 $y_{l,3}(t)$ 、 $y_{r,3}(t)$ は FTA からの出力信号である。 $y_4(t)$ は加減算の結果であり、ERD の最終出力である。

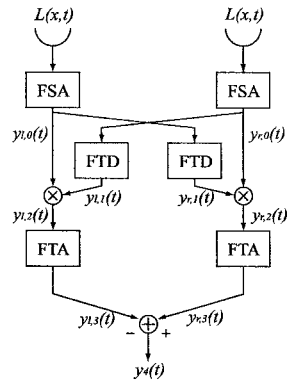


図2 ERD の構造

2.2.1 空間フィルタ

まず、観察信号 $L(x, t)$ に対して窓関数 $\Gamma_h(x)$ により空間フィルタをかける。これを0ステップ目の信号 $y_{h,0}(t)$ として、下式で表す。 h は hand の略で、 r (右側) もしくは l (左側) が入る。

$$y_{h,0}(t) = \int_{-\infty}^{\infty} \Gamma_h(x) L(x, t) dx \quad (8)$$

2.2.2 時間遅延フィルタ

次に、空間フィルタを経た $y_{h,0}(t)$ に対し、インパルス応答 $h(t)$ の時間遅延フィルタをかける。 \tilde{h} は h の反対側からの入力を表す。時間遅延フィルタの結果である $y_{h,1}(t)$ は、 $h(t)$ と $y_{h,0}(t)$ とのたたみ込み積分となる。

$$y_{h,1}(t) = \int_{-\infty}^{\infty} h(t - \tau) y_{h,0}(\tau) d\tau \quad (9)$$

2.2.3 乗算

$y_{h,0}(t)$ と $y_{h,1}(t)$ との乗算を行う。この後の時間平均フィルタにて時間積分をとることで、その出力 $y_{h,3}(t)$

は $y_{h,0}(t)$ と $y_{h,1}(t)$ の内積に比例した値となる。

$$y_{h,2}(t) = y_{h,0}(t)y_{h,1}(t) \quad (10)$$

2.2.4 時間平均

$y_{h,2}(t)$ の時間平均を取る。無限の過去から現在までの平均を取ることで、 $y_{h,3}(t)$ は時間の関数でなく観察信号に規定される $y_{h,3}(L)$ となる。

$$y_{h,3}(t) = y_{h,3}(L) = \lim_{s \rightarrow \infty} \frac{1}{t+s} \int_{-s}^t y_{h,2}(\tau) d\tau \quad (11)$$

2.2.5 加減算

最後に右側出力と左側出力の差分をとることで、2つの相関を比較する。この正負により観察信号の移動方向を決定する。

$$y_4(L) = y_{r,3}(L) - y_{l,3}(L) \quad (12)$$

2.3 ERD の出力値

前節で述べた ERD $\rightarrow L(x, t)$ を入力した場合、その出力 $y_4(L)$ は以下のように、 $n = 1 \sim \infty$ の線形結合の形で表される。入力信号は $n = 0$ から始まっているが、その出力には含まれない。

$$y_4(L) = -d \sum_{n=1}^{\infty} a_n^2 y_s(L_n) y_t(L_n) \quad (13)$$

$$y_s(L_n) = |H_l(2\pi n f_x)| |H_r(2\pi n f_x)| \sin\{\phi_r(2\pi n f_x) - \phi_l(2\pi n f_x)\} \quad (14)$$

$$y_t(L_n) = -|H_t(2\pi n f_t)| \sin\{\phi_t(2\pi n f_t)\} \quad (15)$$

$y_s(L_n)$ は ERD 出力のうち空間フィルタに関連する部分、 $y_t(L_n)$ は同じく時間平均フィルタに関連する部分である。 $|H_l(2\pi n f_x)|$ 、 $|H_r(2\pi n f_x)|$ は空間フィルタの窓関数 $\Gamma_l(x)$ 、 $\Gamma_r(x)$ の利得である。また $\phi_l(2\pi n f_x)$ 、 $\phi_r(2\pi n f_x)$ は同じく窓関数の位相遅れである。 $|H_t(2\pi n f_t)|$ は時間遅延フィルタのインパルス応答 $h(t)$ の利得、 $\phi_t(2\pi n f_t)$ は位相遅れとなっている。

式 (13) より、ERD の入出力間には時間的・空間的線形性があることがわかる。

2.3.1 空間エイリアシング

$y_s(L_n)$ のうち、 $\sin\{\phi_r(2\pi n f_x) - \phi_l(2\pi n f_x)\}$ は、窓関数 $\Gamma_l(x)$ と $\Gamma_r(x)$ の位相遅れの差により、正負が決定する。つまり窓関数の選び方や入力信号の空間周波数 f_x によっては $y_4(L)$ の正負が反転してしまい、移動方向を正しく求められなくなる。これを空間エイリアシングという。このため、常に $0 < \phi_r(2\pi n f_x) - \phi_l(2\pi n f_x) < \pi$ もしくは $-\pi < \phi_r(2\pi n f_x) - \phi_l(2\pi n f_x) < 0$ となるような窓関数を選ぶ必要がある。

2.3.2 時間エイリアシング

前項と同様に $y_t(L_n)$ のうち、時間遅延フィルタの位相遅れである $\sin\{\phi_t(2\pi n f_t)\}$ は、時間遅延フィルタのインパルス応答 $h(t)$ の位相遅れにより正負が決

定する。よって時間遅延フィルタの選び方や入力信号の時間周波数 f_t によっては $y_4(L)$ の正負が反転してしまい、移動方向を正しく求められなくなる。これを時間エイリアシングという。 $\phi_t(2\pi n f_t)$ が一定の範囲 $0 < \phi_t(2\pi n f_t) < \pi$ もしくは $-\pi < \phi_t(2\pi n f_t) < 0$ に限定されるような時間遅延フィルタを選ぶ必要がある。

2.4 採用した空間フィルタ、時間遅延フィルタ

ここでは空間フィルタと時間遅延フィルタを提案し、そのフィルタが時間・空間エイリアシングを防止できることを示す。

2.4.1 空間フィルタ

今回提案する空間フィルタの窓関数を、下式に示す。

$$W(x) = e^{-\frac{(x-x_c)^2}{4\sigma^2}} \quad (16)$$

$$\Gamma_l(x) = \frac{1}{\sigma\sqrt{\pi}} W(x) \cos\{f_0(x-x_c)\} \quad (17)$$

$$\Gamma_r(x) = \frac{1}{\sigma\sqrt{\pi}} W(x) \sin\{f_0(x-x_c)\} \quad (18)$$

これは van Santen らがその論文の中で提案した窓関数⁴⁾の係数を、筆者が調整したものである。これ以降、特に限定しない場合この空間フィルタを用いることにする。このような関数をガボール (Gabor) 関数と呼ぶ。 f_0 は窓関数の空間角周波数 (周波数でないことに注意) であり、 x_c は中心位置となる。 σ は窓の大きさを表し、 $\sigma = \sqrt{2}/f_0$ のとき正弦波または余弦波の約 2 周期分が窓の中に入る。図 3 は、 $f_0 = 0.02\pi$ 、 $\sigma = \sqrt{2}/f_0$ として $W(x)$ 、 $\Gamma_l(x)$ 、 $\Gamma_r(x)$ のグラフを描いたもので、横軸は空間座標 x を表している。式

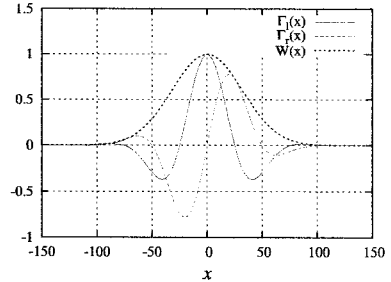


図 3 ガボール関数

(18) で表される空間フィルタを採用した場合、その利得と位相遅れは以下のようになる。

$$|H_l(2\pi n f_x)| = e^{-\sigma^2(f_0 - 2\pi n f_x)^2} + e^{-\sigma^2(f_0 + 2\pi n f_x)^2} \quad (19)$$

$$\phi_l(2\pi n f_x) = 2\pi n f_x x_c \quad (20)$$

$$|H_r(2\pi n f_x)| = e^{-\sigma^2(f_0 - 2\pi n f_x)^2} - e^{-\sigma^2(f_0 + 2\pi n f_x)^2} \quad (21)$$

$$\phi_r(2\pi n f_x) = 2\pi n f_x x_c + \frac{\pi}{2} \quad (22)$$

よって、 $y_s(L_n)$ の正負を決定する部分は

$$\sin\{\phi_r(2\pi n f_x) - \phi_l(2\pi n f_x)\} = \sin\left(\frac{\pi}{2}\right) = 1 \quad (23)$$

となり、空間エイリアシングを防止できることが示された。このとき、ERD の出力 $y_4(L_n)$ のうち空間フィルタに関連する部分 $y_s(L_n)$ は式 (14) に式 (19)~(22) を代入して、下のように表される。

$$y_s(L_n) = e^{-2\sigma^2(f_0 - 2\pi n f_x)^2} - e^{-2\sigma^2(f_0 + 2\pi n f_x)^2} \quad (24)$$

2.4.2 時間遅延フィルタ

今回提案する時間遅延フィルタのインパルス応答を、下式に示す。

$$h(t) = \begin{cases} \frac{2}{\tau} e^{-t/\tau} & \text{if } t \geq 0, \\ 0 & \text{otherwise.} \end{cases} \quad (25)$$

これは、W.Reichardt がその論文の中で提案した時間遅延フィルタ³⁾ の係数を調整したものである。これ以降、特に限定しない場合この時間遅延フィルタを用いることにする。 τ はフィルタの時定数である。 $h(t)$ はシンプルな RC フィルタと同じ特性を持っている。この振幅と位相遅れは下式で表される。

$$|H_t(2\pi n f_t)| = \frac{2}{\sqrt{1 + (2\pi n f_x v \tau)^2}} \quad (26)$$

$$\phi_t(2\pi n f_t) = -\tan^{-1}(2\pi n f_x v \tau) \quad (27)$$

となる。よって

$$\sin\{\phi_t(2\pi n f_t)\} = \sin\{-\tan^{-1}(2\pi n f_x v \tau)\} \in \left[-\frac{\pi}{2}, 0\right] \quad (28)$$

より、時間エイリアシングを防止できることが示された。ここで、ERD の出力 $y_4(L_n)$ のうち時間遅延フィルタに関連する部分 $y_t(L_n)$ は式 (15) に式 (26), (27) を代入して、下のように表される。

$$y_t(L_n) = \frac{4\pi n f_t \tau}{1 + (2\pi n f_x v \tau)^2} \quad (29)$$

2.4.3 観察信号の進行方向

式 (13) と $y_t(L_n) \geq 0$, $y_s(L_n) \geq 0$ より、ERD の出力 $y_4(L)$ の符号は d により一意に決まることがわかる。つまり、信号が左方向に進んでいるとき ($d=1$) は $y_4(L) < 0$, 右方向に進んでいるとき ($d=-1$) は $y_4(L) > 0$ となる。よって観察信号の進行方向が不明である場合も、ERD の出力の符号より進行方向の推定値 \hat{d} は以下のように求められる。

$$\hat{d} = \frac{y_4(L)}{|y_4(L)|} \quad (30)$$

また $v=0$ で信号が静止しているときは、 $y_t(L_n) = 0$ より $y_4(L) = 0$ となる。

3. ERD による一次元信号の移動速度推定

3.1 原理

ERD の観察信号が式 (7) の $S(x, t)$ で示される正弦波運動縞ならば、空間フィルタと時間遅延フィルタ

の最適パラメータ f_0^* , τ^* を求めることで移動速度が推定できる。その方法を示す。具体的には、観察信号 $S(x, t)$ の空間周波数 f_x , 時間周波数 f_t が未知のとき、ERD を応用してその推定値 \hat{f}_x , \hat{f}_t を求め、下式より速度の推定値 \hat{v} を算出する。

$$\hat{v} = \frac{\hat{f}_t}{\hat{f}_x} \quad (31)$$

3.1.1 空間フィルタの最適化

観察信号 $S(x, t)$ に対する ERD 出力は、

$$y_4(S) = -da^2 y_s(S) y_t(S) \quad (32)$$

となる。ここで、空間フィルタに関連した項 $y_s(S)$ は以下ようになる。

$$y_s(S) = e^{-2\sigma^2(f_0 - 2\pi f_x)^2} - e^{-2\sigma^2(f_0 + 2\pi f_x)^2} \quad (33)$$

と表される。この関数 $y_s(S)$ を最大とするようなパラメータ f_0^* は

$$f_0^* = 2\pi f_x \quad (34)$$

と表される。よって観察信号 $S(x, t)$ の空間周波数 f_x が未知のとき、 $y_s(S)$ を最大とするようなパラメータ f_0^* を求めれば、推定値 \hat{f}_x は下式により求められる。

$$\hat{f}_x = \frac{f_0^*}{2\pi} \quad (35)$$

また、 f_0^* を選んだときの $y_s(S)$ の値は

$$y_s(S) = 1 - e^{-16} \simeq 1 \quad (36)$$

となる。式 (33) の $y_s(S)$ に適当な値を代入してグラフを描いて確かめてみる。図 4 に、 $f_0 = 1, \sigma = 10$ として描いたグラフを示す。 $y_s(f_0^*)$ で極大となることが分かる。

3.1.2 時間遅延フィルタの最適化

上述した $y_4(S)$ のうち、時間遅延フィルタに関連した部分は式 (29) より

$$y_t(S) = \frac{4\pi f_t \tau}{1 + (2\pi f_x v \tau)^2} \quad (37)$$

となる。この関数 $y_t(S)$ を最大とするようなパラメータ τ^* は、下式で表される。

$$\tau^* = \frac{1}{2\pi f_t} \quad (38)$$

よって観察信号 $S(x, t)$ の時間周波数 f_t が未知のとき、 $y_t(S)$ を最大とするようなパラメータ τ^* を求めれば、推定値 \hat{f}_t は下式により求められる。

$$\hat{f}_t = \frac{1}{2\pi \tau^*} \quad (39)$$

また、 τ^* を選んだときの $y_t(S)$ の値は

$$y_t(S) = 1 \quad (40)$$

となる。空間フィルタと時間遅延フィルタの係数調整はこのように入出力を正規化するために行われている。図 5 に $\tau = 1/2\pi$ としたときの $y_t(S)$ のグラフを示す。 $f_x = 1/2\pi\tau = 1$ となる点で最大となっていることが分かる。

3.1.3 推定速度 \hat{v} の算出

ここまでで求めた \hat{f}_x , \hat{f}_t より、式 (31) を用いて

推定速度 \hat{v} を算出する。

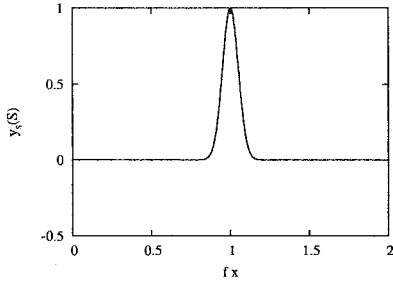


図4 $y_s(u)$ のグラフ

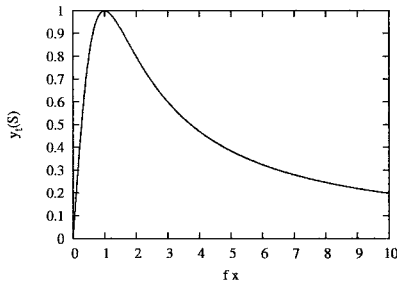


図5 $y_t(S)$ のグラフ

3.2 ERD による時空間周波数解析

3.1 では単一周波数の正弦波運動縞の移動速度が推定できることを述べた。これを応用して、複数の周波数を持つ自然画像の移動速度を推定することも可能となる。複数の ERD にて信号を計測することで、動画の時間・空間周波数を解析することが出来る。ERD による速度推定はこれを利用している。以下、その詳細を述べる。

3.2.1 空間周波数の測定

空間フィルタのパラメータ f_0 を変えながら観察信号 $L(x, t)$ に対して ERD をかけ、その応答特性を測定することは、窓関数 $\Gamma_l(x)$ 、 $\Gamma_r(x)$ にて周波数解析を行うことに等しい。これは観察信号が複数の空間周波数を持つ場合にも応用できる。

図6に示す自然画像の $y = 50$ での輝度断面が図7である。一次元信号の速度推定においては、この輝度断面を ERD に観測させることになる。図8に空間フィルタの応答特性 $y_A(f_0/2\pi)$ と、フーリエスペクトル $|F(f_x)|$ を示す。一般に自然画像は DC 成分を最も多く含み、高周波成分になるにつれて急激に減少する傾向がある。図6の画像はエッジを多く含むため、フーリエスペクトルを見ると DC 成分の次は

$f_x = 0.02$ の成分が最大となっている。 $y_A(f_0/2\pi)$ も同様に $f_x = 0.02$ がピークを示しているため、最適な空間フィルタのパラメータは $f_0^* = 0.02$ となる。

以上の議論と式13より、 $y_A(f_0/2\pi)$ は観察信号のフーリエスペクトルによく似た傾向となることがわかる。そして観察信号の空間周波数成分のうち、DC成分を除く最も支配的な周波数成分に対して最大の値を示し、これが f_0^* となることがわかる。

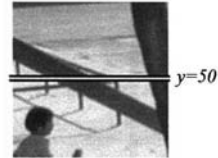


図6 自然画像

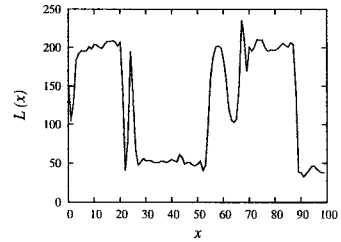


図7 自然画像の輝度断面

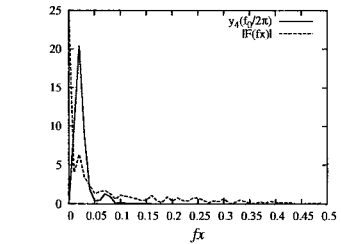


図8 自然画像に対する空間フィルタの応答

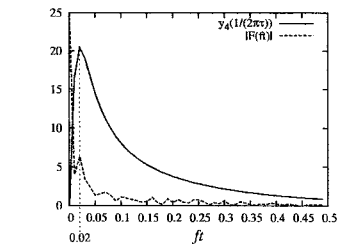


図9 自然画像に対する時間遅延フィルタの応答

3.2.2 時間周波数の測定

図12の自然画像を $dv = 1$ となるように左方向に移動させた動画を用いて同様の測定を行ったのが図

9 である。空間周波数の測定と同じく、フーリエスペクトルは時間周波数 $f_t = 0.02$ の点でピークを持ち、 $1/(2\pi\tau^*) = 0.02$ となることがわかる。

3.2.3 移動速度の推定

以上より、複数のパラメータを持つ ERD は、動画の周波数解析を行っていることがわかる。 f_0^* , τ^* から動画の持つ空間・時間周波数のうち最も支配的な成分を算出できるため、複数の周波数を持つ自然画像においても、ERD による速度推定が可能となる。

4. 速度推定実験

4.1 概要

前章で説明した ERD による速度推定の実験を行う。実験方法と、3 種類行った実験について述べる。

4.1.1 実験方法

複数の ERD を用いて、空間フィルタと時間遅延フィルタの最適パラメータ f_0^* , τ^* を求め、動画の移動速度を推定する方法について具体的に述べる。

4.1.2 複数の ERD

複数の ERD が動画を観察する仕組みは図 10 のような模式図で表される。以下、各ステップごとにその処理について説明する。

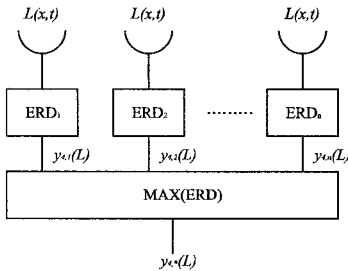


図 10 並列処理を行う複数の ERD

Step 1.異なるパラメータ f_0 , τ を持った n 個の $ERD_1 \sim ERD_n$ に観察信号 $L(x, t)$ を入力し、各々の出力を求める。

Step 2. $ERD_1 \sim ERD_n$ の出力値 $y_{1,1}(L) \sim y_{4,n}(L)$ のうち最大値 $y_{4,*}(L)$ を出力した ERD^* を探す。その ERD^* の持っているパラメータが最適な f_0^* , τ^* となる。

Step 3. f_0^* , τ^* より \hat{f}_x , \hat{f}_t を求め、 $\hat{v} = \hat{f}_t / \hat{f}_x$ より推定速度を算出する。

なお、動画は $L(x, y, t)$ のように空間の次元が 2 次元となっているが、 $y = \text{一定}$ で $L(x, t)$ とした輝度断面を入力して 1 次元の移動方向を求めることとする。また、 $L(x, t)$ が RGB 表色系よりなるカラー画像の場合、下式を用いてグレースケールの輝度値に変換して

観察信号とする。 Y の値が輝度値 $L(x, t)$ となる。

$$Y = 0.299R + 0.587G + 0.114B \quad (41)$$

4.1.3 実装方法

Microsoft の DirectShow を用い、PC 上で動画処理システムを作成した。実験 I の自然画像を並進運動させた画像はオープンソースの画像処理ライブラリである OpenCV にてプログラムを作成し、生成した。

4.2 実験 I: 自然画像を並進運動させた動画 (公園)

4.2.1 実験に使用した動画

まず、公園の風景をデジタルカメラにて撮影し、JPEG 形式の 2272×1704 pixel の画像を得た。その後、フォトタッチソフトにてニアレストネイバー法で 50% に縮小し、中央部分のトリミングを行った。すると図 11 に示すような静止画像 park.jpg が得られた。この解像度は 1136×100 pixel である。この静止画像の上を、 100×100 pixel という解像度の仮想的なカメラを動かすようにして動画を生成した。生成には OpenCV を用いた。移動速度 $dv = 0.5, 1.0, 2.0$ pixel/frame の 3 段階の移動速度にて 3 種類の動画を作り、それぞれのファイル名を park05.avi, park10.avi, park20.avi とした。図 12, 13 に park10.avi のスナップショット画像を示す。

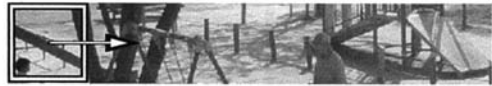


図 11 生成元の静止画像 park.jpg (一部)

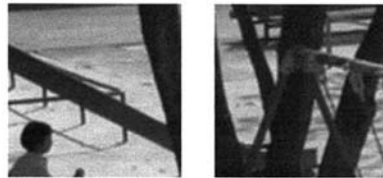


図 12 動画像 park10.avi 0 フレーム目 図 13 動画像 park10.avi 100 フレーム目

4.2.2 ERD のパラメータ

自然画像の速度推定には、 f_x , f_t の範囲を $f_x \in [0.01, 0.05]$, $f_t \in [0.001, 0.05]$ とし、0.001 刻みで候補列を設定した。それぞれの最大値を 0.05 と定めたのは 3.1 より、自然画像の時間・空間周波数は低周波の領域に集まることがわかったためである。この値域をもとに f_0 と τ のパラメータ候補を算出し、図 14 のように定めた。

4.2.3 実験結果と考察

実験結果を図 15 に示す。上から移動速度 $dv =$

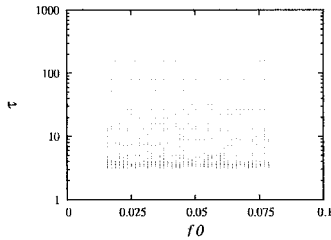


図 14 用意したパラメータ

2.0, 1.0, 0.5 の時の速度推定結果を表している。横軸はフレーム数、縦軸は推定された移動速度である。 $dv = 1.0$ の場合を例にとると、はじめのうち 3.0 pixel/frame 程度が推定されているが、徐々に減少して 50 フレーム目あたりから実際の値である 1.0 付近を前後している。移動速度は全フレームにわたり一定の 1.0 pixel/frame であるが、算出開始直後は ERD に含まれる時間遅延フィルタの過渡現象の影響で速度が安定して出力されるまでに時間 (5τ 程度) がかかる結果となっている。入力される動画像の速度が速ければ速いほど、 τ^* が小さくなるため短時間で定常状態となる。 $dv = 2.0$ のときに 20 フレーム目くらいで正しい速度が推定されているのはそういう理由からである。過渡現象の影響は多少あるものの、

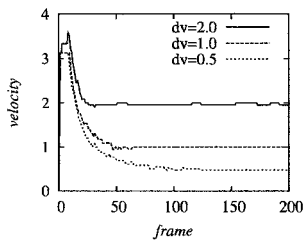


図 15 実験結果

4.3 実験 II: カメラを回転させて撮影した動画像 (研究室)

4.3.1 実験に使用した動画像

上の実験では静止画像の上を仮想的なカメラが動くという特殊な動画像を用いて実験を行っている。続いて、実際に DV カムコーダを回転させ、得られた動画像から速度推定を行う。DV カムコーダを図 16* に示す自動回転台 (楽々ツイビー***) の上に載せ、一定

速度で回転させながら撮影することで像面全体が一様な速度場を持つような動画像 lab.avi を撮影した。回転速度は 17.04 deg/sec である。この速度は 1 周回転する秒数 (21.13 sec) で 360 deg を割ることにより求めた。実験に用いた動画像の一部を図 17, 18 に示す。動画像は解像度 720*480 pixel で撮影したものを 360*240 pixel へ縮小して使用した。水平視野角 37.9 deg で撮影したため、9.5 pixel/deg の解像度となった。よって、回転速度は 5.40 pixel/frame と求められる。フレームレートは 30 fps である。上の移動速度は 1 周分の回転時間からこれらの値を用いて算出した。エンコード形式はノイズの発生を極力抑えるため AVI 形式とした。



図 16 ツイビー



図 17 lab.avi
0 フレーム目



図 18 lab.avi
100 フレーム目

4.3.2 関連アルゴリズムとの比較

本実験では提案手法と勾配法¹⁾、ブロックマッチング法⁵⁾ との比較も同時に行う。これらの手法は 2 次元の速度推定を可能としているが、条件を揃えるため ERD 法と同じく $y = 120$ の行のみを取り出して 1 次元の移動速度を推定した。勾配法、ブロックマッチング法ともに動画像中の小領域の速度 (= ローカルモーション) ベクトルを推定する方法である。よって画面全体の動き (= グローバルモーション) はこれらのベクトルの平均により算出した。なお、実験パラメータは表 1, 2 のようにした。

表 1 勾配法のパラメータ
誤差係数 α 4.0

4.3.3 実験結果

結果のグラフを図 19 に示す。また、表 3 に 250 フ

* 写真は旭フォトマイクロウェア (株) のウェブサイトより転用
** 楽々ツイビーは旭フォトマイクロウェア (株) の登録商標です

ブロックサイズ	10 pixel
サーチエリアサイズ	20 pixel
誤差の評価	差の絶対値

フレーム目の結果を示す。グラフを見ると、ERD はほぼ 4~6 pixel/frame の間を推移しているが、勾配法とマッチング法は 1~6 pixel/frame の間で大きく変化する結果となった。使用した動画は、PC の液晶ディスプレイや柱などのエッジ以外にも、壁などの一様な領域を多く含んでいる。一様な明るさを持つ領域が移動してもその速度を推定することは困難であり、その部分だけ速度は小さな値となってしまう。両アルゴリズムとも、ローカルモーションベクトルを平均して画面全体の動きを算出する方法を採用しているため、特徴的な画像（エッジなど）を含まない領域で速度が小さな値として推定された分、画面全体の動きも小さな速度として推定される結果となった。

これに対して ERD は構造上、画面全体を入力領域とするため小領域のエッジの有無の影響が小さくなる。よって、より正確な速度推定が可能になったと考えられる。

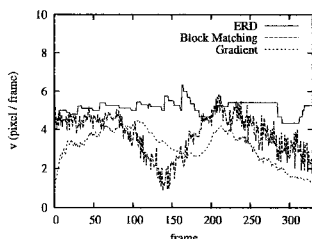


図 19 勾配法、マッチング法との比較

アルゴリズム	速度 pixel/frame
(実測値)	(5.40)
ERD	5.40
勾配法	3.32
マッチング法	4.71

5. おわりに

5.1 まとめ

本論文では、ERD を用いて 1 次元の速度推定を行う手法について述べた。ERD 単体で移動方向を推定する手法、複数の ERD を用いて移動速度を推定する方法について詳説した。また、人工生成画像、カメラで撮影した動画画像によって実験を行い、提案手法の正しさを示した。最後の実験では勾配法、マッチング法

との比較を行い、ERD による速度推定が像面全体の速度場検出に特に優れていることを示した。

5.2 今後の課題

今後の課題を以下に列挙する。

- ERD パラメータ f_0^* , τ^* の推定にあたり、パラメータ候補を多数用意して総当たりしているため、これを最適化して計算コストを減らす必要がある。
- 1次元のみではなく 2次元方向への移動へも対応する必要がある。
- 並進運動以外の運動（回転、剪断、拡大、縮小など）も推定可能にする必要がある。
- 像面全体の動きだけではなく、小領域の動きも高精度に検出する必要がある。

謝 辞

本研究は、平成 18 年度科学研究費、および財団法人日東学術振興財団からの助成金を使い行われました。著者一同、これらの援助に深く感謝いたします。

参 考 文 献

- 1) Berthold K.P. Horn and Brian G. Schunk. Determining optical flow. *ARTIFICIAL INTELLIGENCE*, pp. 185-203, 1981.
- 2) F. Iida and D. Lambrinos. Navigation in an autonomous flying robot by using a biologically inspired visual odometer, 2000.
- 3) Werner Reichardt. Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. *Principles of sensory communication*, pp. 303-317, 1961.
- 4) J.P. Van Santen and G. Sperling. Temporal covariance model of human motion perception. *Journal of the Optical Society of America*, pp. 451-473, 1984.
- 5) 奥富正敏 (編). デジタル画像処理. CG-ARTS 協会, 2006.
- 6) 市村正広. Emd によるリアルタイム動き検出アルゴリズムの基礎的検討. 修士論文, 愛知工業大学 大学院, 2001.
- 7) 浅井哲也. 生体の視覚システムに学んだ動き検出および視覚対象追従運動機能の集積回路化に関する研究. 博士論文, 豊橋技術科学大学, 1999.
- 8) 中村健人. 昆虫の視覚情報処理に基づいたリアルタイム動き検出アルゴリズムの基礎的検討. 修士論文, 愛知工業大学 大学院, 2003.
- 9) 服部巧. Elementary motion detector のアナログ回路による実装. 修士論文, 愛知工業大学 大学院, 2007.
- 10) 日本視覚学会編. 視覚情報処理ハンドブック. 朝倉書店, 2000.