

時空間フィルタを用いた運動領域の抽出と 物体追跡への適用

A method for extracting a moving object
region with spatiotemporal filters and
an its application to tracking

田中 恭治*

Kyoji Tanaka

* 山口県工業技術センター

* Industrial Technology Institute,
Yamaguchi Prefectural Government

白井 良明**

Yoshiaki Shirai

** 大阪大学

** Osaka Univ.

Abstract This paper describes a method for extracting a specified moving object region, assuming that each object moves in an arbitrary direction, including objects that move in opposite directions. For the extraction of an object which moves at a speed of more than 1 pixels/frame, we use spatio-temporal filters which are lowpass in the spatial frequency dimension.

We succeeded in tracking a fast moving object by using a window which is fit to a high energy region of the filtered image; the window fitting process is based on a PYRAMID METHOD.

We confirmed that this extraction method is useful even under conditions in which each object moves in an arbitrary direction with arbitrary velocity and/or in which the background is not stationary.

1 まえがき

運動する物体のスペクトルは速度に比例して時間周波数軸方向に傾く。この性質を利用し、時空間フィルタリングにより運動物体を認識する方法がいくつか提案されている。Anderson、Burtらは画像をラプラシアン-ガウシアンフィルタに通しピラミッド構造化し、等価的に時空間バンドパスフィルタリングを行い、この出力から運動物体の認識と注目を行った[1]。

LeeとLinはこれを改良し、全周波数バンドの出力から物体の運動エネルギーを再構築し、よりロバストな運動の決定を行った[2]。Burt、Hingoraniらは空間周波数軸方向にバンドを切り換えることで異なる速度の物体を識別している[3]。これらの方法においては時空間フィルタのインパルス応答を時間と空間で独立に扱っている。このため、時空間フィルタは周波数領域において全象限対称になり、正の方向の動きと負の方向の動きが分離できない。

Heegerは時空間フィルタとして楕円型の3次

元GABORフィルタを用い、物体の速度を検出している[4]。このフィルタは全象限対称とはならないため動きの方向により物体を分離可能である。しかし、この研究によれば速度1pixel/frame以上の処理はピラミッド階層化(多分解能化)で対処しており、また加速度運動について考慮されていない。さらに、カメラが移動することにより背景が動く場合についても考慮されていない。

本研究では、時空間周波数領域において各速度に対応した時空間フィルタを設定し、画像をこのフィルタのインパルス応答とコンボリューションする。この時、空間周波数は低域を通過するように制限を設ける。これにより、任意の向きに速度1 pixel/frame以上の運動をする複数物体から注目する運動領域を抽出する。

また、物体の速度が限定できない場合や加速度運動を考慮し、速度に対応してフィルタのチャンネルを複数用意し出力が最大のチャンネルを選択することで注目する物体をロバストに抽出す

る。

第2章では運動物体の時空間周波数領域における性質について記述する。第3章では時空間フィルタについて述べる。第4章ではピラミッド法を用いてフィルタ出力の高エネルギー領域に追跡窓をかけ、運動物体の領域を決定する。

2 時空間周波数領域における運動物体の性質

空間 (x, y) において速度 (u, v) で運動している物体の画像が信号 $h(x, y)$ として得られた場合を考える。この場合、時空間 (x, y, t) 上においては信号 $f(x, y, t)$ は

$$f(x, y, t) = h(x+u \cdot t, y+v \cdot t) \quad (1)$$

と表される。

ここで、 $f(x, y, t)$ の3次元フーリエ変換および $h(x, y)$ の2次元フーリエ変換をそれぞれ

$$\begin{aligned} F[f(x, y, t)] &\equiv F(\omega_x, \omega_y, \omega_t) \\ &= \int \int \int f \cdot \exp(-j\omega_x x) \exp(-j\omega_y y) \\ &\quad \exp(-j\omega_t t) \, dx dy dt \\ F[h(x, y)] &\equiv H(\omega_x, \omega_y) \\ &= \int \int h \cdot \exp(-j\omega_x x) \exp(-j\omega_y y) \, dx dy \end{aligned} \quad (2)$$

とすると、

$$F(\omega_x, \omega_y, \omega_t) = H(\omega_x, \omega_y) \delta(\omega_x u + \omega_y v - \omega_t) \quad (3)$$

が成り立つ。ここで、 $\delta(*)$ はディラックのデルタ関数である。

ゆえに、速度 (u, v) で運動する場合 $f(x, y, t)$ の周波数成分は

$$\omega_t = \omega_x u + \omega_y v \quad (4)$$

の平面上に現れる。すなわち、スペクトルは (ω_x, ω_t) 領域で u 、 (ω_y, ω_t) 領域で v の傾きを示す(図1)。

また、速度 $(-u, -v)$ で運動する場合

$$F(\omega_x, \omega_y, \omega_t) = H(\omega_x, \omega_y) \delta(\omega_x u + \omega_y v + \omega_t)$$

であり、スペクトルは (ω_x, ω_t) 領域で $-u$ 、 (ω_y, ω_t) 領域で $-v$ の傾きを示す。

ここで、 u と v を独立と見なすと (ω_x, ω_t) 領域で傾き u を中心とした近傍領域を通過させるフィルタを用いることで速度 u で運動する信号 $f(x, t)$ の成分が抽出できる。同様に (ω_y, ω_t) 領

域で傾き v を中心とした近傍領域を通過させるフィルタを用い速度 v で運動する信号 $f(y, t)$ の成分が抽出できる。

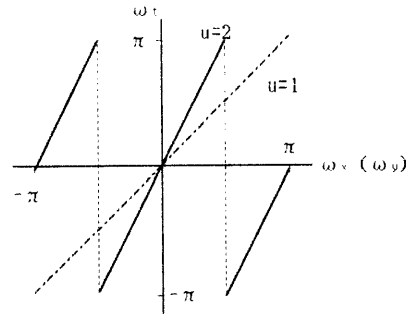


図1 周波数領域における運動物体の性質

ところでサンプリング定理により、 $u > 1$ の場合および $u < -1$ の場合、空間周波数 ω_x の高域成分は折り返しを生じる。このことより、通過域が空間高周波まで及ぶような時空間フィルタを用いた場合、抽出すべき運動とは逆方向に高速で運動する物体の折り返し成分をも抽出してしまう。

本研究においては、許容最高速度を u_{max} とした時、空間周波数 $\omega_x \leq \pi / u_{max}$ の範囲では正常に抽出できることに注目し、この範囲でバンドパスするようにフィルタを設定することにより $u \leq u_{max}$ ($u_{max} > 1$)の運動を抽出する。

3 運動の抽出

3.1 時空間フィルタリング

前述したように u, v を独立とみなし領域 (ω_x, ω_t) 、領域 (ω_y, ω_t) において独立に時空間フィルタを用いる。

本研究では、フィルタ $G(\omega_x, \omega_t)$ を第1象限または第2象限において、傾き u を中心とした近傍領域を通過するフィルタとして設定し、フィルタのインパルス応答 $g(x, t)$ と信号 $f(x, t)$ のコンボリューション結果の実数項を取り出すことで速度 u の物体の信号成分を抽出する。

実数フィルタの場合、そのインパルス応答 $g(x, t)$ は

$$g(x, t) = g_R(x, t) + jg_I(x, t) \quad (5)$$

である。但し、 $g_R(x, t)$ 、 $g_I(x, t)$ は $g(x, t)$ の実

数部、虚数部を示す。

この時、時空間フィルタリングは $f(x, t)$ と $g(x, t)$ とのコンボリューション

$$f * g = f * g_R + j f * g_I \quad (6)$$

で表される。ここで、実数項のみ取り出すと、そのフーリエ変換は、

$$\begin{aligned} F[f * g_R] \\ = F(\omega_x, \omega_t) \cdot \frac{G(\omega_x, \omega_t) + G(-\omega_x, -\omega_t)}{2} \end{aligned} \quad (7)$$

となる。但し、 $\frac{G(\omega_x, \omega_t) + G(-\omega_x, -\omega_t)}{2}$ は複素共役を示す。

また、 $F(\omega_x, \omega_t) = F(-\omega_x, -\omega_t)$ を考慮すると、 $f * g_R$ により領域 (ω_x, ω_t) の第1象限と第3象限、あるいは第2象限と第4象限の成分が抽出できる(図2)。

領域 (ω_x, ω_t) については、上述の x を y に置き換えることで全く同様に議論できる。

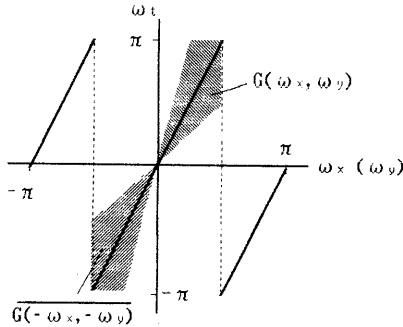


図2 時空間フィルタ

結局、 $G(\omega_x, \omega_t)$ は次の条件を満たすものとする。(1) 各速度に対応する周波数成分の傾きを中心とした近傍領域を通過帯域とする。

(2) 第1象限または第2象限を通過させる実数フィルタとする。(3) 許容最高速度 u_{max} のとき、 $\omega_x \leq \pi / u_{max}$ を通過帯域とする(図3)。

3.2 フィルタの選択

物体が加速度運動する場合や物体の速度が限定できない場合、あるいはカメラの移動により観測される速度が変動する場合を考慮すると、速度に応じてフィルタ $G(\omega_x, \omega_t)$ の通過領域を変更する必要がある。このため、各速度に対応させたフィルタ $G_1(\omega_x, \omega_t) \sim G_M(\omega_x, \omega_t)$ を予

め用意する。ここで、 m ($1 \leq m \leq M$) はフィルタのチャンネルを表す。

あるフレームにおける物体の瞬間速度 (u, v) が次のフレームにおいて $(u + \Delta u, v + \Delta v)$ に変化した

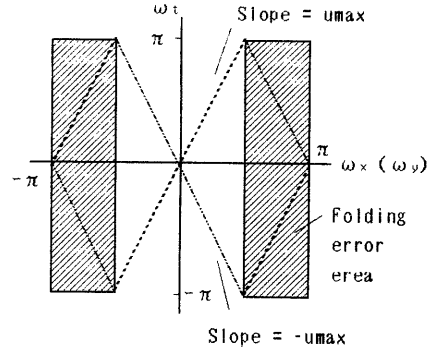


図3 フィルタの禁止領域

場合、これに対応してチャンネルを切り換える。 Δu 、 Δv はある程度小さく、1フレームにおけるチャンネル m の変化は ± 1 以内とする。この仮定のもとで u (または v) に対応するチャンネル i とその隣接チャンネル $i-1$ 、 $i+1$ ($1 \leq i-1, i, i+1 \leq M$) を用いて並列にフィルタリングし、最大値を出力するチャンネルを次のフレームにおける中心チャンネル i とする。

ところで、ある速度の物体とそれにほぼ等しい速度の物体とが共存する場合、フィルタの更新は不安定になる。このため、各チャンネルからの出力の比較は4章に述べる追跡窓内において行う。

すなわち、各辺を x_{min} 、 x_{max} 、 y_{min} 、 y_{max} とする追跡窓内において、速度 u についてのチャンネル m における関数 $E_u(m)$ と速度 v についてのチャンネル n における関数 $E_v(n)$ を

$$\begin{aligned} E_u(m) &= \iint | \operatorname{Re} [f(x, y, t) * g_m(x, t)] | dx dy \\ E_v(n) &= \iint | \operatorname{Re} [f(x, y, t) * g_n(y, t)] | dx dy \\ &\text{for } x_{min} \leq x \leq x_{max}, y_{min} \leq y \leq y_{max} \end{aligned}$$

とし、

$$p = \arg \max_{m=1, \dots, M} E_u(m), \quad q = \arg \max_{n=1, \dots, M} E_v(n)$$

なるチャンネル p 、 q からの出力を u 、 v に最も適応したチャンネルからの出力

$$\begin{aligned} Ou(x, y, t) &= | \operatorname{Re} [f(x, y, t) * g_p(x, t)] | \\ Ov(x, y, t) &= | \operatorname{Re} [f(x, y, t) * g_q(y, t)] | \end{aligned} \quad (8)$$

とする。

3. 3 水平移動成分と垂直移動成分の合成

複数の運動物体から特定の速度ベクトル (u,v) のものだけを抽出する。このためには時空間 (x,t) におけるフィルタ出力と時空間 (y,t) におけるフィルタ出力を合成する必要がある。

物体 i の速度ベクトル $v_i=(u_i,v_i)$ と別の物体の速度ベクトル v_j が $v_i \neq v_j$ の関係においても $u_i=u_j$ あるいは $v_i=v_j$ となることがある。このことより $0u(x,y,t)$ または $0v(x,y,t)$ が物体 i と j の成分を共に含むことがありうる。

従って、両出力を合成し

$$Y(x,y,t)=(0u(x,y,t)+0v(x,y,t))^{1/2} \quad (9)$$

とすることで特定の速度ベクトルのものを抽出する。

4 運動領域の決定

複数の運動物体の中の1つに注目し、その物体を抽出するため追跡窓を用いる。対象の大きさが限定できない場合や、1つの物体が接近して見かけの大きさが変化する場合は考慮すると、追跡窓はフレーム毎に位置と大きさを更新する必要がある。

本研究においてはピラミッド法を利用し、追跡窓の更新をフレーム毎に行う。

あるフレームにおいてフィルタを通過した画像 $Y(x,y,t)$ を適当なしきい値 TH で2値画像 $I(i,j)$ とした後、階層 N までピラミッド $Pl_k(i,j) \sim Pl_N(i',j')$ を構築し、coarse to fineの方向に窓を決定する。

この手順は以下のようなになる。ここで、窓の各辺 $xmin, xmax, ymin, ymax$ の初期値は追跡開始時に注目する対象を含むように荒く設定する。

- ```

step1 $Pl_0(i,j) \leftarrow I(i,j)$
 $Pl_k(i,j) \leftarrow \max_{m,n=0,1} \{ Pl_{k-1}(2i+m, 2j+n) \}$
for $1 \leq k \leq N$
step2 $xmin \leftarrow xmin/2^N, xmax \leftarrow xmax/2^N$
 $ymin \leftarrow ymin/2^N, ymax \leftarrow ymax/2^N$
 $k=N$ とし、 $Pl_k(i,j)=1$ の領域の輪郭と窓の各辺 $xmin, xmax, ymin, ymax$ が接するように各辺を更新する。
step3 $xmin \leftarrow xmin \cdot 2, xmax \leftarrow xmax \cdot 2$

```

$$ymin \leftarrow ymin \cdot 2, ymax \leftarrow ymax \cdot 2$$

$k=k-1$  とし、 $Pl_k(i,j)=1$  の領域の輪郭と窓の各辺  $xmin, xmax, ymin, ymax$  が接するように各辺を  $\pm 1$  する。

step4  $k=0$  までstep3を繰り返す。

step5  $xmin \leftarrow xmin - OFST, xmax \leftarrow xmax + OFST$

$$ymin \leftarrow ymin - OFST, ymax \leftarrow ymax + OFST$$

step6 次のフレームにおいてstep1に戻る。

ここで、窓サイズのオフセット  $OFST$  は許容最高速度を  $umax$ 、時空間フィルタの時間方向タップ数を  $T$  とした時、 $OFST \geq umax \cdot T / 2$  に選ぶ。

更新後の追跡窓は図4のようなになる。

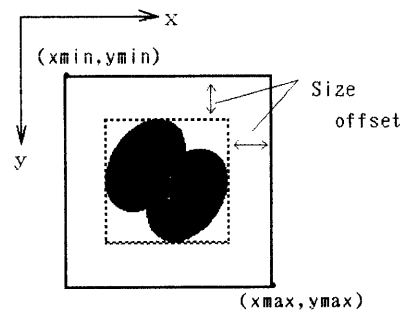


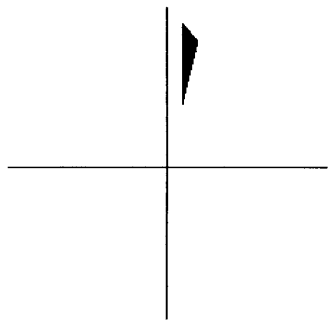
図4 抽出された領域にかかる追跡窓

### 5 実験

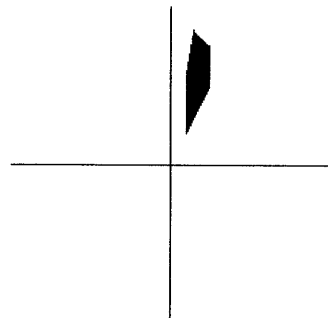
#### 5. 1 バッチ処理によるフィルタの設計

フィルタのインパルス応答を各チャネル毎にテーブルとして予め用意する。時間方向7タップ、空間方向15タップで構成するフィルタとし、 $-umax \sim umax$  までについてチャネル数7とした。フィルタは周波数領域において図5(a)のように設計した。同図において黒はゲイン1、白はゲイン0である。但し、同図はチャネル3~6であり、チャネル0~2は空間周波数軸を中心としてチャネル6~4に対称である。なお、同図が示すように設計したフィルタにおいては、実際のフィルタがギブス現象により負の空間周波数領域にまで広がることを防ぐため、予め空間低周波をカットした。

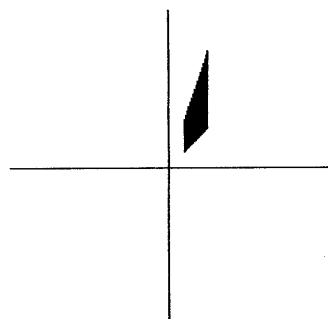
図5(a)のフィルタを上記タップ数で有限とし逆フーリエ変換し、重み関数  $g_R, g_I$  を求めた。この  $g_R, g_I$  をフーリエ変換したものを、すなわち実際のフィルタを図5(b)に示す。同図は黒いほどゲインは大きく、最大(真黒)はゲイン1である。



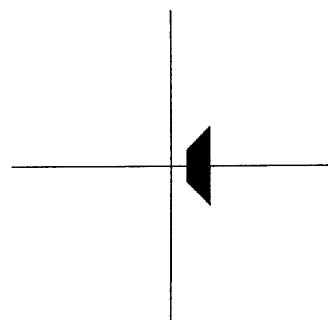
チャンネル 6



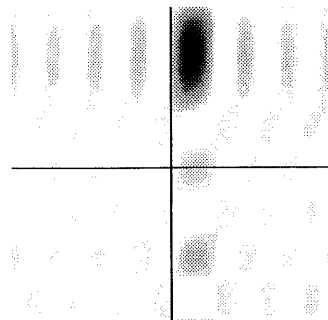
チャンネル 5



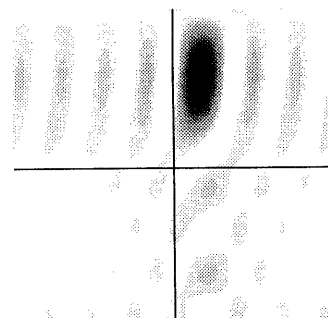
チャンネル 4



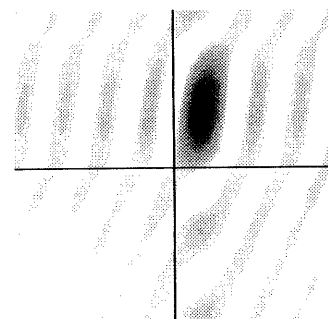
チャンネル 3  
(a)理想フィルタ



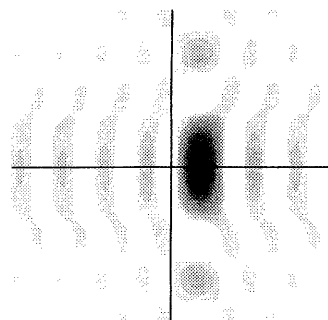
チャンネル 6



チャンネル 5



チャンネル 4



チャンネル 3  
(b)実際のフィルタ

図 5 時空間フィルタ

なお、より先鋭で狭帯域のフィルタはタップ数を増せば実現できるが、同時に処理遅れも増加するため上記のタップ数とした。

### 5. 2 速度フィルタリング

原画像を256×240のサイズとし、許容最高速度を2倍にするため、原画像をガウシアンフィルタに通し128×120のサイズに圧縮したものを入力画像として用いた。図6に原画像を示す。

原画像の背景 $f_{back}(k,l)$ は

$$f_{back}(k,l)=127.5 \cdot \left( 1 + \sum_{m=2}^{20} \sum_{n=2}^{20} \cos(nk/20\pi) \cos(ml/20\pi) \right)$$

なる、広帯域の空間周波数成分を有するものを用いた。

運動物体は画像の左上から右下と右下から左上に同じ速さでそれぞれ逆向きに動く設定とした。

この実験画像における運動物体の抽出結果を図7に示す。左上の物体を特定対象とした場合の結果が図7(a)、(b)であり、図7(a)は速度2 pixel/frameでの運動の抽出結果、図7(b)は6 pixel/frameでの運動の抽出結果である。いずれの場合も右下から左上に運動する物体は速度が負、また背景は速度0であり、注目している物体とは速度が異なるため消滅し、特定した運動物体の成分のみ抽出されている。速度 2pixel/frameで動く物体(図7(a))に対してはチャンネルは#4に安定する。チャンネル#4は速度0の成分を若干通過してしまうため、背景の残留成分は図7(a)の方が図7(b)より大きい。図7(c)、(d)は右下の物体を特定対象とした場合の抽出結果である。

### 5. 3 フィルタチャンネルの更新

原画像(図6)において運動物体の速度を変化させ、左上の物体を抽出した時の物体速度とチャンネルの状態の関係を確認した。表1はチャンネルが最適状態に安定した後の結果である。

### 5. 4 追跡窓の更新

運動が図8のように変化(加減速)する場合について追跡窓の更新状況を確認した。この時、パラメータとしては

表1 速度に対するチャンネルの更新結果

|                       |   |   |   |   |   |
|-----------------------|---|---|---|---|---|
| velocity<br>u,v       | 2 | 4 | 6 | 7 | 8 |
| stabilized<br>channel | 4 | 5 | 5 | 6 | 6 |

2値化しきい値 TH = peak\_value\*0.6

追跡窓オフセット OFST = 15

を用いた。ここで、peak\_valueは式(9)により得られた信号 $Y(x,y,t)$ のピーク値である。この追跡結果を図9に示す。但し、本実験の場合フィルタの時間方向タップ数は7であり処理遅れが3フレームある。このため、同図は追跡窓を3フレーム遅れた画像にスーパーインポーズしたものである。

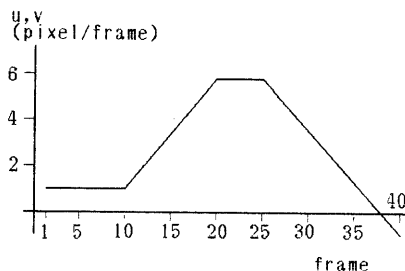


図8 実験に用いた加減速運動

### 5. 5 カメラの運動

カメラが移動する(背景が動く)環境で運動物体を抽出する場合について実験を行った。背景が右下から左上に速度3pixel/frameで動くもとの、左上から右下に速度6で接近する物体と右下から左上へ速度6で遠ざかる物体がある場合において、左上から右下に運動する物体を追跡した結果を図10に示す。図中の黒点は背景の一部である。また、運動物体の表面輝度は球状分布するテクスチャであり、中心が最も明るい。

同図から、背景が動く場合における物体の追跡においても本方式は有効である。また、物体の見かけの大きさが変化する場合においても有効であることが確認される。

### 6 あとがき

時系列2次元画像を時空間フィルタのインパルス応答とコンボリューションし特定の運動物

体の抽出を行った。また、速度1pixel/frame以上の物体の抽出は、空間低周波成分を通過させる時空間フィルタにすることで可能であることを示した。また、各速度に対応するチャンネルを設けフレーム毎にチャンネルを更新することで物体の速度変化に追従し、この出力に対してピラミッド法を利用して追跡窓を形成し、運動物体の追跡を行った。

本方式によれば、従来は難しかったカメラが移動する環境下においてもロバストな抽出、追跡が可能であることを示した。

なお、本方式の構成はフィルタとピラミッドが主であり、両者ともハードウェアで実現できる。このため、実時間処理が可能であり比較的 low cost に実用化が図れる。

#### 文献

- (1) C.H.Anderson,P.J.Burt,G.S.van der Wal:  
"Change Detection and Tracking Using  
Pyramid Transform Techniques",SPIE  
Vol.579 IRCV,pp.72-78,1985
- (2) James S.J.Lee,C.Lin:"A Novel Approach  
to Real-Time Motion Detection",Proc.  
CVPR,pp.730-735,1988
- (3) P.J.Burt,Rajesh Hingorani and Raymond  
J.Kolczynski:"Mechanisms for Isolating  
Component Patterns in the Sequential  
Analysis of Multiple Motion",proc.IEEE  
Workshop on Visual Motion,pp.187-193,  
1991
- (4) David J.Heeger:"Optical Flow from  
Spatiotemporal Filters",proc.ICCV,pp.  
181-190,1987

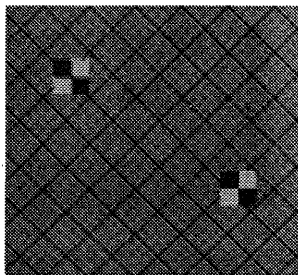
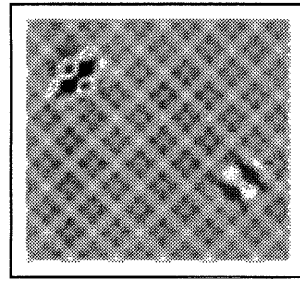
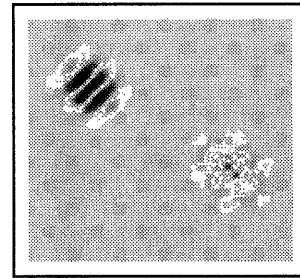


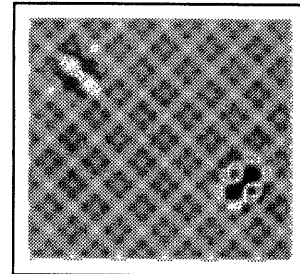
図 6 原画像



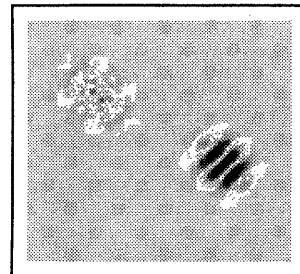
(a) 速度:2pixel/frame



(b) 速度:6pixel/frame

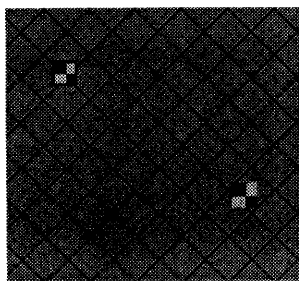


(c) 速度:-2pixel/frame

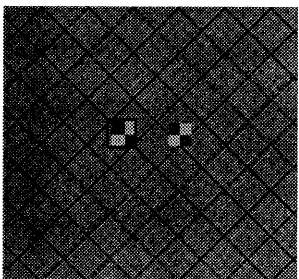


(d) 速度:-6pixel/frame

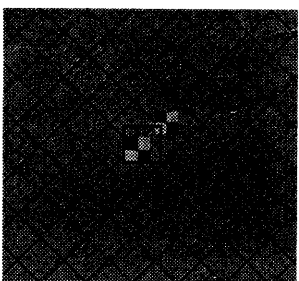
図 7 運動領域抽出結果



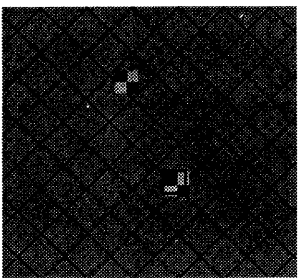
FRAME #10



FRAME #25

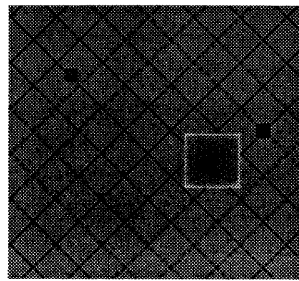


FRAME #27

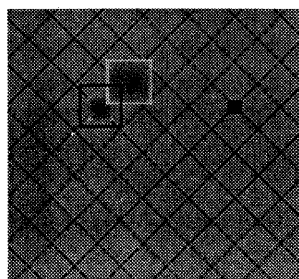


FRAME #35

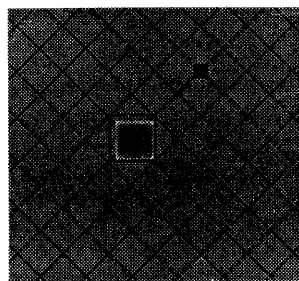
図 9 加減速運動物体の追跡結果



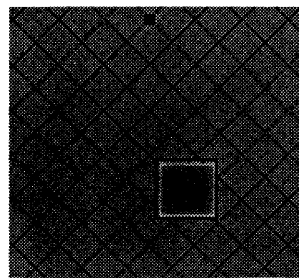
FRAME #7



FRAME #15



FRAME #25



FRAME #40

図 10 背景移動環境下における接近／後退運動物体の追跡結果