

## 多段階ハフ変換による 3-D 認識のための動き検出

Motion Detection for 3-D recognition by Multistage Hough transform

中島 重義、周 明勇、濱 裕光、山下 一美 (大阪市立大学工学部)

Shigeyoshi NAKAJIMA, Mingyong ZHOU, Hiromitsu HAMA and Kazumi YAMASHITA

Faculty of Engineering, Osaka City University

### Abstract

We propose a method for motion detection from moving images using Hough transform. The inputs are feature points separated in the moving images and the outputs are the velocity vectors of the feature points. We can recognize a feature point whose moving is approximated linear as a straight line in the 3-D space whose axes are two of 2-D image and one of time. Generally Hough transform has a disadvantage that it needs lots of computations and memories. But with our method the needed amount of memories is in proportion of the number of pixels of the image and the needed computation is in proportion of the product of the number of pixels and the number of images frames.

### [ 1 ] はじめに

動画像から物体の動きを検出する問題には、多くの研究者が取り組んでいる。特に剛体の動きについては過去に様々な提案がなされた。<sup>1) 2) 3) 4) 5)</sup>我々はこの剛体の動きについて、動画像上で十分な数の特徴点の2次元的位置と速度が複数の瞬間について観測できればハフ変換の手法を用いて剛体の動きが回復できることを示した。<sup>6)</sup>しかし、この手法のためには動画像の特徴点の位置と速度が安定的に観測されることが必要であった。動画像上の特徴点の検出を3次元空間での直線検出とみなせば、この問題は3次元のハフ変換の手法で解ける。しかし、高次元のハフ変換は多量の計算と記憶容量を必要とし、単純なハフ変換では現実的な時間内でこの問題を解くには多くの困難があった。今回、我々は、同等の変換を2次元のハフ変換で実行する事を提案し、その計算が現実的な計算時間と記憶容量で実現できることを示す。

### [ 2 ] 剛体の3次元の動き

剛体の動きは剛体の中心の移動、中心回りの回転で近似的に表現される。実際の物理的な動きでは重力その他の相互作用があるが、ここでは考えず、物体は自由な運動をしているものとする。物体の中心座標が $(x_g, y_g, z_g)$ で、その速度ベクトルが $(\alpha_g, \beta_g, \gamma_g)$ 、各軸回りの角速度が $(\theta, \phi, \psi)$ であるとする。観測される特徴点の3次元空間での位置を $(x, y, z)$ 、回転による成分と平行移動の成分を足し合わせた速度ベクトルを $(\alpha, \beta, \gamma)$ であるとする。次の式が成立する。

$$\begin{aligned}\alpha &= \alpha_g + \phi(z - z_g) - \psi(y - y_g) \\ \beta &= \beta_g + \psi(x - x_g) - \theta(z - z_g) \\ \gamma &= \gamma_g + \theta(y - y_g) - \phi(x - x_g).\end{aligned}\tag{1}$$

観測された特徴点の画像上の2次元的位置を $(X, Y)$ とし、その速度ベクトルを $(A, B)$ とする。次の式が成立する。

$$\begin{aligned}
X &= \frac{fx}{z}, \\
Y &= \frac{fy}{z}, \\
A &= \frac{f\alpha}{z} - \frac{f\gamma x}{z^2}, \\
B &= \frac{f\beta}{z} - \frac{f\gamma y}{z^2}.
\end{aligned} \quad (2)$$

これらより、特徴点の固有の位置座標である  $(x, y, z)$  を消去すると次式が得られる。

$$\begin{aligned}
& \frac{(f\alpha_g - X\gamma_g + Xy_g\theta - fz_g\phi - Xx_g\phi + fy_g\psi)}{(A + \frac{XY\theta}{f} - f\phi - \frac{X^2\phi}{f} + Y\psi)} \\
= & \frac{(f\beta_g - Y\gamma_g + Yy_g\theta + fz_g\phi - Yx_g\phi - fx_g\psi)}{(B - \frac{YX\theta}{f} + f\theta + \frac{Y^2\theta}{f} - X\psi)}.
\end{aligned} \quad (3)$$

点  $P_i$  に対し、(3)式が成立すれば、次の値が0になる。

$$\begin{aligned}
& h_i(\alpha_g, \beta_g, \gamma_g, \theta, \phi, \psi) \\
= & (f\alpha_g - X_i\gamma_g + X_iy_g\theta - fz_g\phi - X_ix_g\phi + fy_g\psi) \\
& \times (B_i - \frac{Y_iX_i\theta}{f} + f\theta + \frac{Y_i^2\theta}{f} - X_i\psi) \\
& - (f\beta_g - Y_i\gamma_g + Y_iy_g\theta + fz_g\phi - Y_ix_g\phi - fx_g\psi) \\
& \times (A_i + \frac{X_iY_i\theta}{f} - f\phi - \frac{X_i^2\phi}{f} + Y_i\psi).
\end{aligned} \quad (4)$$

この値を丘登り法とハフ変換を併用すれば正しい運動パラメータ  $(\alpha_g, \beta_g, \gamma_g, \theta, \phi, \psi)$  が得られ、座標  $(x_g, y_g, z_g)$  と  $(x, y, z)$  が回復できる。

### [ 3 ] 2次元の動きの検出

前節での画像上で移動は短い時間の間においては、時間パラメータ  $t$  にしたがう線形な動きに近似できる。

$$\begin{aligned}
X(t) &= X_0 + At \\
Y(t) &= Y_0 + Bt.
\end{aligned} \quad (5)$$

この動きは  $(X, Y, t)$  の三次元空間での直線に近似できる。この直線に原点から降ろした垂線の足を  $(a, b, c)$  とする。以下の式が成立する。

$$\begin{aligned}
0 &= Aa + Bb + c \\
a &= X_0 + Ac \\
b &= Y_0 + Bc.
\end{aligned} \quad (6)$$

さて、次の式を満足する位置  $(l, m, n)$  を考える。

$$l^2 + m^2 + n^2 = lX + mY + nt. \quad (7)$$

これは、 $(X/2, Y/2, t/2)$  を中心とする球面であり、原点と  $(a, b, c)$  と  $(X, Y, t)$  が球面上にある。さて、原点と  $(a, b, c)$  を結ぶ線分を直径とし、ベクトル  $(A, B, 1)$  に垂直な平面に属する円の上にある点を  $(d, e, f)$  とする。  $(d, e, f)$  と  $(a/2, b/2, c/2)$  と  $(X/2, Y/2, t/2)$  を結んだ三角形は、  $(a, b, c)$  と  $(a/2, b/2, c/2)$  と  $(X/2, Y/2, t/2)$  を結んだ三角形と合同であり、その結果  $(d, e, f)$  と  $(X/2, Y/2, t/2)$  との距離は  $(a, b, c)$  と  $(X/2, Y/2, t/2)$  の距離と同じである。よって  $(d, e, f)$  は常に上記の球面上にあることになる。  $(d, e, f)$  の位置は  $(a, b, c)$  の位置とベクトル  $(A, B, 1)$  により依存し、時間成分  $t$  には依存しない。この事は、異なる  $t$  に対する球面が常に  $(d, e, f)$  が形成する円を含むことを示す。よって、観測された  $(X, Y, t)$  に対応する球面上に蓄積を行えば、蓄積は  $(d, e, f)$  の上に集中する。この集中を検出し、さらにそれらの形成する円を検出できれば、ベクトル  $(A, B, 1)$  と位置  $(a, b, c)$  とが検出できたことになる。

### [ 4 ] 全空間にわたる蓄積と円検出の工夫<sup>6)</sup>

画像のサイズを  $N^2$  とし、ある瞬間に検出される特徴点の数の平均を  $M$  とする。本手法では画

像のフレーム上では特徴点は分離している必要がある。一般に  $M \ll N$  である。また、観測された  $t$  のサンプリングが  $T$  回であったとする。一般的な状況では  $N$  と  $T$  では一般に  $N$  の方が大きい。  $T$  が大きすぎると直線への近似ができない。十分な精度でハフ変換を行うためにはハフ空間のサイズは  $N^3$  以上である必要があり、それだけの記憶容量が必要である。この空間に一つの球を蓄積するとほぼ  $N^2$  回の蓄積が必要であり、全時間にわたって検出される特徴点は  $MT$  であるから、  $MTN^2$  回の蓄積が必要である。この蓄積の後、蓄積の集中を検出するためにはこの全空間にわたってフィルタ操作が必要であるから、  $N^3$  に比例する計算量が必要である。特徴点を抽出するクライテリアが一定であるとすると、一枚の静止画像から抽出される特徴量の平均数  $M$  は  $N$  に比例し、動きが直線に近似されるためには  $T$  がある値以下に制限されるはずであるから、結局全計算量は  $TN^3$  に比例することになる。

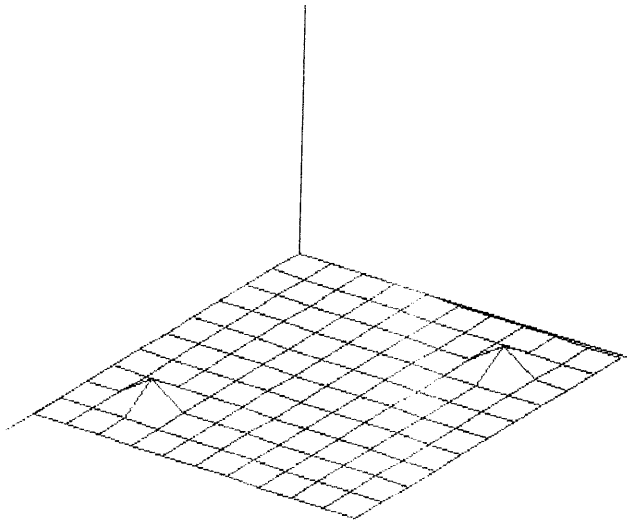


図1 画像の例 (時刻  $t_1$ )

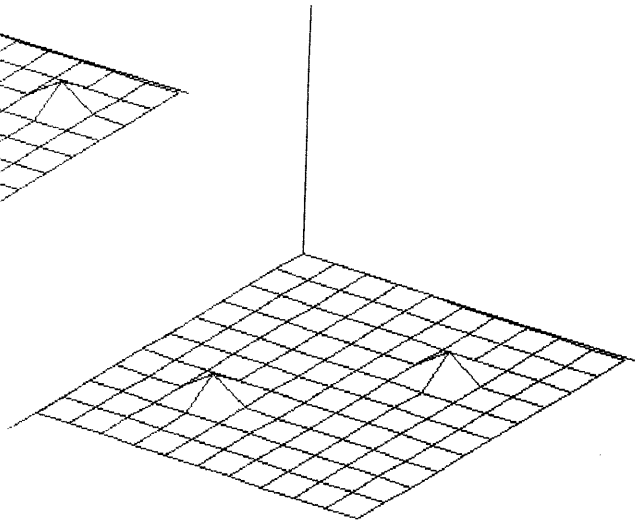


図2 画像の例 (時刻  $t_2$ )

[ 5 ] 数枚の平面上の蓄積と円検出

全空間にわたる蓄積はコストのオーダーが画素数のオーダーを越すが、全空間にわたる蓄積は必ずしも必要が無い。この円が平面と交わる部分では平面上に集中点が出現する。図1から図3までは異なる  $t$  において観測したとして合成した画像である。2点異なる速度で移動している。この速度に関わらず集中点を検出するためには、6枚の平面が必要である。6枚の平面は  $lm$  平面、 $mn$  平面、 $n1$  平面とそれらの二つの間

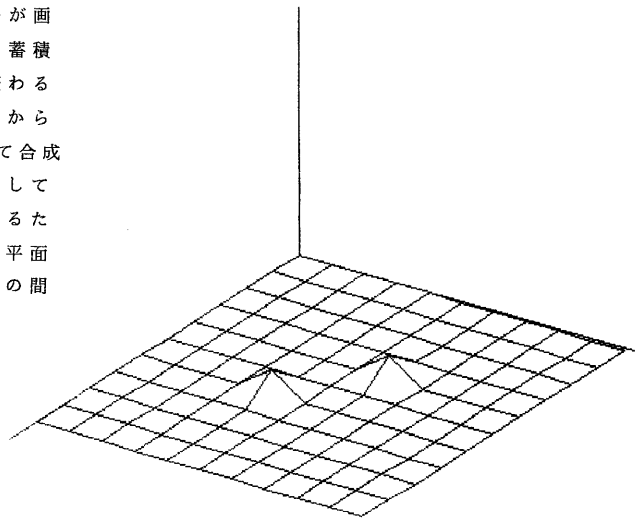


図3 画像の例 (時刻  $t_3$ )

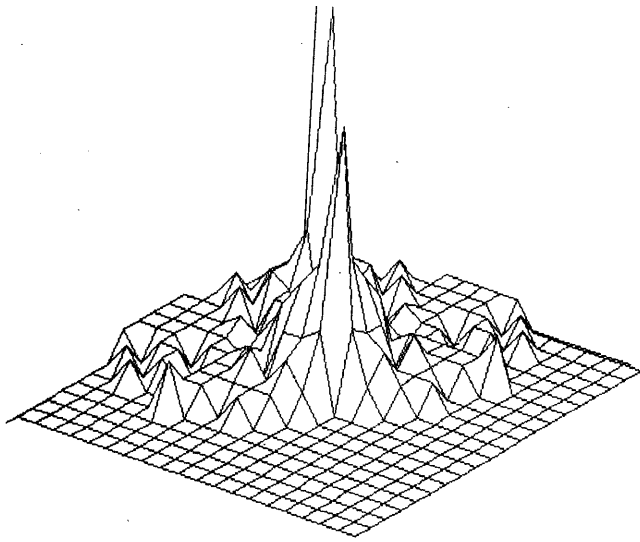


図4 1m平面上的蓄積

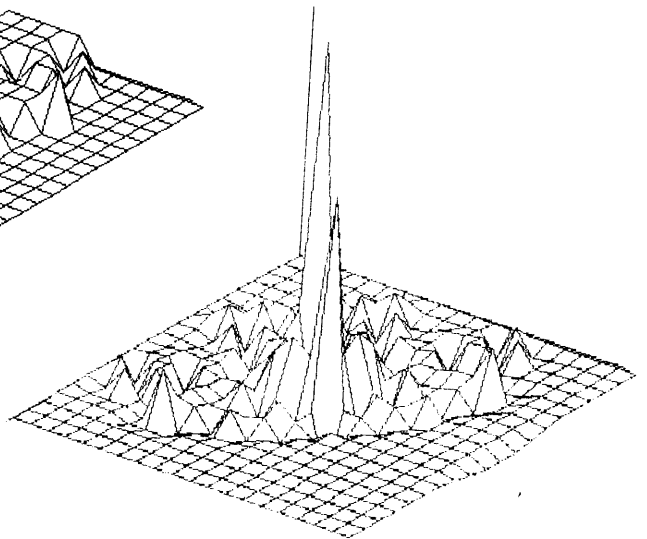


図5 ラプラシアンフィルタの結果

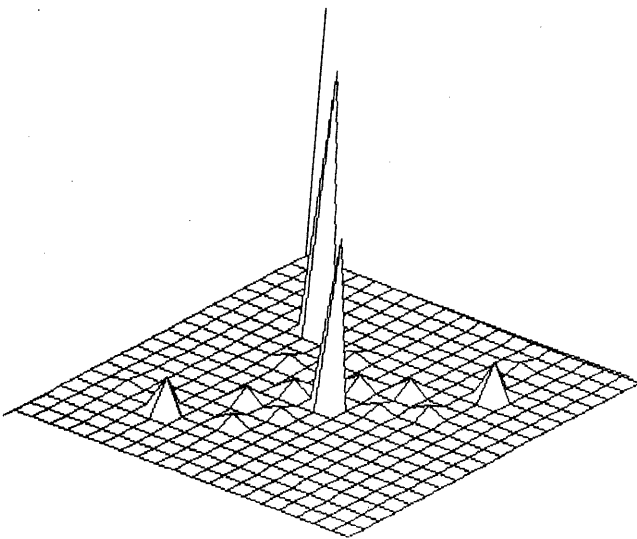


図6 8近傍より高い部分

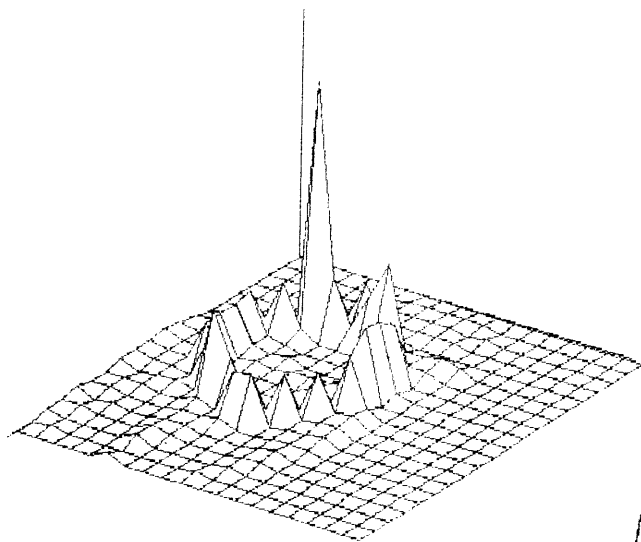


図7 正しい面上の蓄積

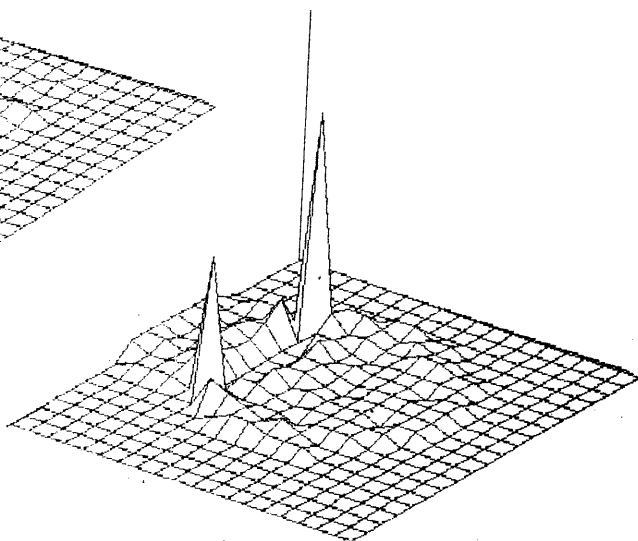


図8 誤った面上の蓄積

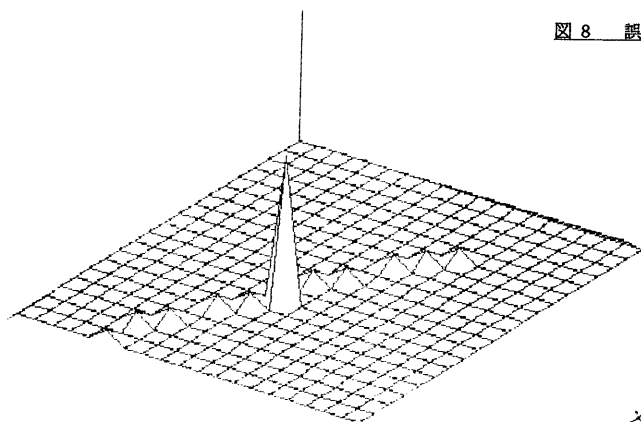


図9 正しい面上の正射影

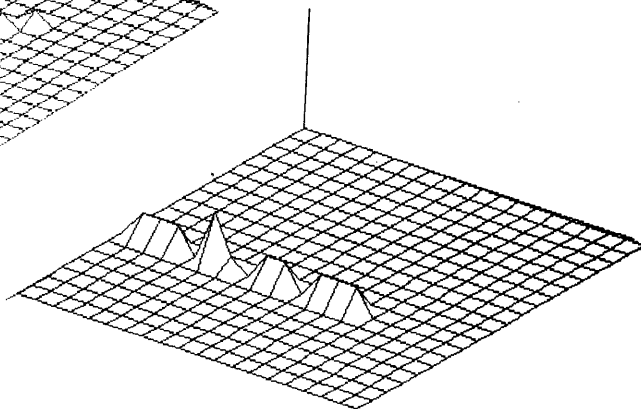


図10 誤った面上の正射影

にある平面である。よって、必要な記憶容量は  $6N^2$  に比例する。一つの特徴点に対して、一枚の平面の上に一つの円を描き、その計算量はほぼ  $N$  に比例する。よって、その計算量は、ほぼ  $MNTN$  に比例する。前節の議論に従えば、これはほぼ  $TN^2$  に比例する。

図4は図1などで示された動画像に対して  $1m$  平面の上に蓄積された結果である。前述の円が平面と交わり、矢印で示した2点の集中点が出ている。これに対してこの集中点を強調するためにラプシアンフィルタをかけた結果が図5である。さらに、このように強調された集中点を検出するためにするために八近傍より高い部分だけを残すと図6のようになる。このように残った集中点の値は、その特徴点がどれだけ多くの瞬間において観測されたかによる。図中サンプリング雑音による誤ったピークも残っているが、真の集中点との差は歴然としている。

このような集中点の組み合わせにより、前述の円の位置を予測する。予測された円の位置が正しければその平面の上での蓄積は図7のように集中点が円の上に並ぶが誤っている場合は図8のように集中点が孤立する。正しい予測の場合の正射影は図9のように高い集中点があり、誤った予測に従った場合は図10のように高い集中点はない。図9における射影の集中位置は図6の原点でない集中点二つと原点がなす三角形の外心の位置から予測できる。この部分だけの射影をとればその計算量は  $T$  に比例する。集中点の数はほぼ  $M$  に比例するから、その二つの組の数は  $M^2$  に比例する。よって、全計算量のオーダーは  $M^2T$  であるが、前述の議論にしたがえば、これは  $TN^2$  に比例する。よって、全計算量は、 $TN^2$  に比例する。

## [6] 結論

動画像の全画素数に比例するオーダーの計算量で、ハフ変換により画像上の特徴点の対応をとる手法を提案した。今後は種々の実際の動画像に対して、特徴点の検出手法および、平面上の蓄積へのフィルタの改良などを計りたい。

## 参考文献

- 1) T. J. Brodia and R. Chellappa, "Estimation of Object Motion Parameters from Noisy Images," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-8, No. 1, Jan. 1986.
- 2) W. Burger and B. Bhanu, "Estimating 3-D Egomotion from Perspective Image Sequences," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-12, No. 11, Nov. 1990.
- 3) G. Adiv, "Determining Three-Dimensional Motion and Structure from Optical Flow Generated by Several Moving Objects," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-7, No. 4, Jul. 1985.
- 4) K. Arber, W. E. Snyder, H. Burkhardt, and G. Hirzinger, "Application of Affine-Invariant Fourier Descriptors to Recognition of 3-D Objects," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-12, No. 7, Jul. 1990.
- 5) J. Aisbett, "An Iterative Estimation of the Motion Parameters of a Rigid Body from Noisy Displacement Vector," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-12, No. 11, Nov. 1990.
- 6) S. Nakajima, M. Zhou, H. Hama, and K. Yamashita, "Three-dimensional motion analysis and structure recovering by multistage Hough transform," Proc. of VCIP'91, Nov. 1991.