

アフィン運動記述を利用した 明度相関による視覚追跡

足立 佳久 中村 恭之 浅田 稔
y-adachi@cv.ccm.eng.osaka-u.ac.jp
大阪大学 工学部 電子制御機械工学科
〒565 大阪府吹田市山田丘 2-1

単一のウィンドウに基づいた相関法による視覚追跡アルゴリズムでは、目標物体の見え方に変化が起きる場合や、オクルージョンの発生によって目標物体の一部が隠される場合には、追跡を失敗する可能性が大きくなる。

本論文では、アフィン運動記述を利用した相関法による視覚追跡アルゴリズムを提案する。まず、複数のウィンドウを用いて1つの目標物体を追跡し、追跡時に得られる各ウィンドウからの移動ベクトルを用いて目標物体領域の運動を表現する。この運動をアフィン運動モデルにより近似し、アフィン運動パラメータを推定する。この運動パラメータをもとに参照画像を更新することで、目標物体の見え方の変化に対応した視覚追跡を実現する。また、目標物体のオクルージョンの検出は、複数のウィンドウの相関誤差値の分散から求める。

Correlation-Based Visual Tracking enhanced by Affine Motion Description

Yoshihisa Adachi, Takayuki Nakamura, and Minoru Asada
Dept. of Mech. Eng. for Computer-Controlled Machinery
Osaka University, 2-1, Yamadaoka, Suita, Osaka 565, Japan

Tracking a target by conventional correlation-based algorithm with a single reference image often fails on the following situation: 1) a view of the target image changes. 2) a part of the target image is occluded.

In this paper, we propose a method of correlation-based visual tracking enhanced by affine motion description. Multiple tracking windows are utilized for tracking a target. The motion of the tracked region is described by multiple motion vectors each of which obtained from each tracking window, and it is approximated by an affine motion model. Based on the estimated parameters of affine transformation, our method detects the change in the view of the target image, then updates the reference image. And we can also detect the occlusion by calculating variance of SAD (Sum of Absolute Difference).

We give the results for real image sequences to show the validity of our method.

1 はじめに

実環境内でロボットを動かす場合、視覚によってその環境の状態や変化を常に把握していなければならない。そのためには高速な視覚処理機能が必要となる。注視領域を実時間で処理できるシステムとして、局所明度相関を用いた高速演算処理装置トラッキングビジョン [1] [2] がある。しかし、このシステムは、連続して入力される濃淡画像と参照画像から、ウィンドウ内の個々の画素の輝度値の差(相関誤差値)によって相関度を計っているだけで、追跡している物体の形状をもとに追跡を行っているわけではない。故に、単一のウィンドウのみを用いて、初期に取得した参照画像(テンプレート)を更新せずに追跡を行うと、目標物体が回転運動や姿勢変化をおこした場合や、他の物体によって隠された場合には相関誤差値が大きくなり見失う可能性が大きくなる。そのため、追跡対象を見失う前に参照画像の更新や、オクルージョンの判定を行う必要が生じてくる。

森ら [3] は局所明度相関による参照画像の更新に関する研究で、参照画像の記憶を利用する記憶画像主導型の手法を提案している。しかし、追跡過程においてオクルージョンがおこる場合を想定していない。また、参照画像を追加、更新、削除して複数記憶していくが、そのための基準として相関誤差値を用いているために、目標物体の濃度パターンによって閾値を設定し直す必要がある。

また、Darrellら [4] は単一のウィンドウを用い、参照画像と入力された画像の各画素値を比較し得られた相関係数を評価し参照画像の更新を行った。しかし、これもオクルージョンがおこる場合を想定していないため、もしオクルージョンが生じた場合には隠している物体を含んだ新たな参照画像をつくってしまい、対象を見失う可能性が大きくなる。

一方、Nakamuraら [5] は、濃淡画像でのブロック間の明度相関による実時間の追跡において、追跡物体が隠される場合を想定し、複数のウィンドウを組み合わせたものを用い、オクルージョンの判定することにより追跡できるようにしているが、これは、初期に取得した参照画像を更新しないため、目標物体が回転運動や姿勢変化をおこした場合に見失ってしまう。

これらに対し、本研究では、複数のウィンドウを用いて1つの目標物体を追跡することによりオクルー

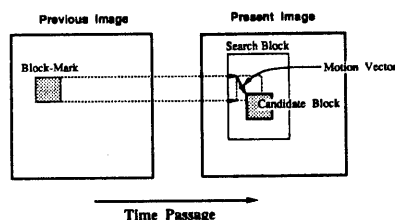


図 1: テンプレート・マッチング

ジョンの判定を可能にし、また、各ウィンドウの移動ベクトルを用いて、全体の運動をアフィン運動近似し、得られたアフィン運動パラメータを評価して目標物体の見え方の変化を検知し、新たな参照画像を取得することにより対象を見失うことなく追跡する手法を提案する。

2 アフィン運動記述を用いた視覚追跡

2.1 各ウィンドウの追跡

視覚追跡は、注視領域での局所明度相関演算に基づき行う。図1のように注視領域近傍で探索を行い、参照画像と相関度が最も高い領域に対象が移動したとしてその移動ベクトルを得る。相関値は、計算量を減らすため SAD(Sum of Absolute Difference) 法を用い、

$$Dist(i, j) = \sum_n \sum_m^{ysize \times xsize} |f_r(m, n) - f_c(m + i, n + j)|$$

と定義する。ここで、 $Dist(i, j)$ は点 (i, j) における相関誤差値を、関数 $f_r(x, y), f_c(x, y)$ はそれぞれ参照画像と候補ウィンドウ内における点 (x, y) の輝度値を、 $xsize \times ysize$ は参照画像のサイズを示す。

2.2 アフィン運動近似

M 個のウィンドウを一組として用いて、追跡を行うとする。

ウィンドウ群の探索領域の中心を原点としたときの、各ウィンドウの中心の初期位置を $x_k = [x_k \ y_k]^T$ とする。そして、各ウィンドウが $x'_k = [x'_k \ y'_k]^T$ に

移動したとき、その移動ベクトルは、 $\delta_k = [\delta_x \ \delta_y]^T = [x'_k - x_k \ y'_k - y_k]^T$ で表される。ただし、 $k=1,2,\dots,M$ とする。

このとき、 x_k と x'_k の関係は、変換行列 A と並進ベクトル d で

$$x'_k = Ax_k + d$$

と表すことができる。例えば、目標物体が並進したとき、つまり各ウィンドウが同一方向に移動したとき、変換行列 A と並進ベクトル d は、 $A = E$ 、 $d = \delta_k$ となる。ただし、 E は単位行列を表す。

各ウィンドウの動きをアフィン運動近似するため最小自乗法により、誤差

$$\epsilon = (x'_k - (Ax_k + d))^2$$

を最小にする A と d を推定する。

$$Y = \begin{bmatrix} x'_1 & x'_2 & \dots & x'_M \end{bmatrix}^T$$

$$X = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_M \end{bmatrix}^T$$

$$B = \begin{bmatrix} d & A \end{bmatrix}^T$$

とおくと、

ここでもし $X^T X$ の行列式が0でなければ、その逆行列 $(X^T X)^{-1}$ が存在して、

$$\hat{B} = (X^T X)^{-1} X^T Y = \begin{bmatrix} \hat{d} & \hat{A} \end{bmatrix}^T$$

で、 B つまり A と d を推定することができる。

2.3 参照画像の更新時期の判定

図2のような配置の9個のウィンドウを用いて一つの目標物体の追跡を行うことにする。1個のウィンドウのサイズは、16[pixel]×16[pixel]とする。

この9個のウィンドウの移動ベクトルにより得られたアフィン運動パラメータ

$$\hat{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

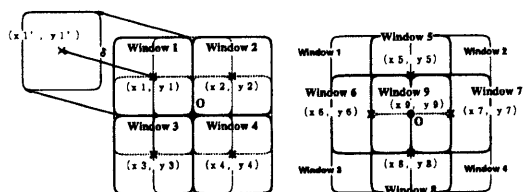


図2: ウィンドウの初期配置

の各成分を評価し、参照画像の更新時期を判定する。

目標物体が静止しているときであっても、その濃度パターンによっては、ウィンドウが x, y 方向それぞれに+2[pixel]から-2[pixel]動く場合があることが経験的に得られた。

図2のウィンドウ1~9の x_k が、 x, y 方向に+2[pixel]から-2[pixel]動くことを考えると、変換行列 \hat{A} の各成分 $a_{11}, a_{12}, a_{21}, a_{22}$ は、

$$0.75 \leq a_{11} \leq 1.25 \quad (1)$$

$$-0.25 \leq a_{12} \leq 0.25$$

$$-0.25 \leq a_{21} \leq 0.25$$

$$0.75 \leq a_{22} \leq 1.25$$

の範囲をとる。

よって式(1)の範囲を満たしているとき、目標物体の見え方はほとんど変化していない、つまり $\hat{A} \approx E$ であるとし、ウィンドウ群の探索領域の中心を並進ベクトル d で移動させ、そのままの参照画像を用いて追跡を続ける。

逆に $a_{11}, a_{12}, a_{21}, a_{22}$ の4つのうち1つでも式(1)の範囲を満たしていないものがあつたとき、つまり $\hat{A} \neq E$ とみなされるときは、探索領域の中心を \hat{d} で移動させた後、各ウィンドウを図2のような初期配置の位置関係に戻し、その位置の画像を新たな参照画像として取得し追跡を続ける。

2.4 アフィン運動近似の問題点とその対応

アフィン運動近似は、2次元平面内の動きを近似したものである。そのため、鉛直軸まわりに回転している物体の追跡に用いると、図3のように参照画像に背景を含んでしまうことがあり、誤動作が生じる。

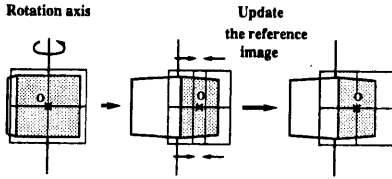


図 3: 鉛直軸まわりに回転している物体の追跡

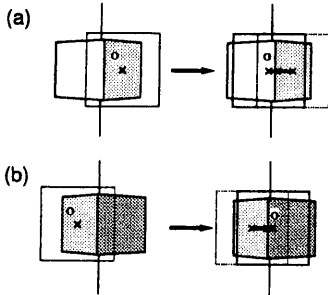


図 4: 探索領域の移動

この問題に対して、回転物体の探索領域を図 4 のように移動させれば、背景を含まない参照画像を取得することができる。追跡物体が鉛直軸まわりにのみ回転している場合は、推定された変換行列 \hat{A} が、

$$\hat{A} = \begin{bmatrix} a_{11} & 0 \\ 0 & 1 \end{bmatrix} \\ a_{11} < 1$$

となる。しかし、 \hat{A} からの情報のみでは回転方向を決定できない。

そこで、 \hat{A} より物体が鉛直軸まわりに回転したことがわかった場合、図 2 のように配置した 9 個のウィンドウ群の左右に図 5 のように 3 個ずつウィンドウを配置し、その後の数フレームの間、中央の 9 個のウィンドウの x, y 方向の動きと左側の 3 個のウィンドウの動きのマハラノビス距離 $D_l^2(i) (i=1,2,3)$ を求める。同様にして、中央の 9 個と右側のマハラノビス距離 $D_r^2(i) (i=1,2,3)$ を求める。

$\sum_{i=1}^3 D_l^2(i) < \sum_{i=1}^3 D_r^2(i)$ の場合は、図 4(a) と同様の場合とみなし、図 5 のように 9 個のウィンドウの中心と左側のウィンドウの中心を結んだ直線の中心に探索領域の中心 O を移動する。

逆に、 $\sum_{i=1}^3 D_l^2(i) > \sum_{i=1}^3 D_r^2(i)$ の場合は、図 4(b)

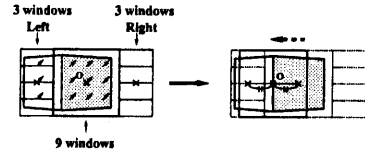


図 5: 左右のウィンドウの配置

と同様の場合とみなし、探索領域の中心 O を右に移動させる。

同様にして、9 個のウィンドウ群の上下に 3 個ずつウィンドウを配置すれば、水平軸まわりに回転している物体にも対応できる。

3 オクルージョンの判定

アフィン運動パラメータが式 (1) の範囲を満たしている ($\hat{A} \approx E$ である) という仮定のもとで、目標物体の一部が他の物体によって遮蔽された場合、遮蔽領域のウィンドウの相関誤差値は、非遮蔽領域のウィンドウのそれに比べて大きくなる。

そこで、各ウィンドウから得られる相関誤差値 $Dist_k (k = 1 \sim 9)$ の分散 σ_D を計算し、それが閾値 σ_{DT} より小さくなったときには、ウィンドウの相関誤差値をクラスター分析し、9 個のウィンドウを 2 つのグループに分割する。このとき、相関誤差値の大きいグループに含まれるウィンドウは遮蔽領域のウィンドウであると特定し、それらの動きを無視する。また参照画像の更新はせずに、現在使用している参照画像で追跡を続ける。

分散 σ_D が閾値 σ_{DT} より小さくなったとき、遮蔽領域であると見なしていたウィンドウの動きを再び考慮し、9 個のウィンドウで追跡を行う。

4 実験結果

2, 3 章で述べた手法を用い、以下の実験を行った。

4.1 視軸まわりに回転する物体の追跡

床面を移動する物体を上からカメラでとらえ追跡を行った。

追跡中の \hat{A} の各成分の値の時間変化を図 6 に示す。また、追跡中に取得した参照画像を図 7 に示す。

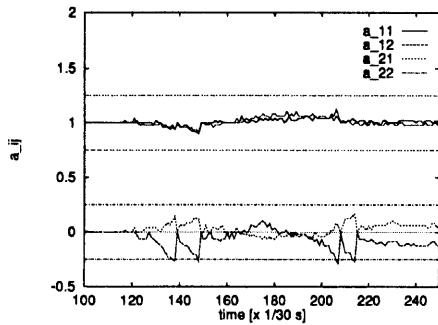


図 6: 変換行列 \hat{A} の各成分 (視軸まわりの回転)

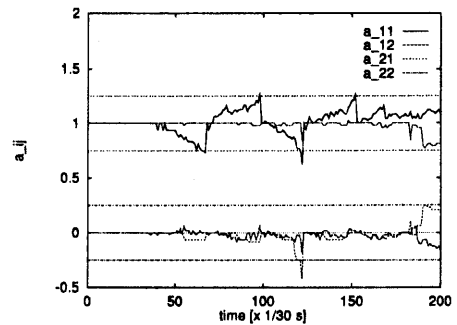


図 8: 変換行列 \hat{A} の各成分 (鉛直軸まわりの回転)

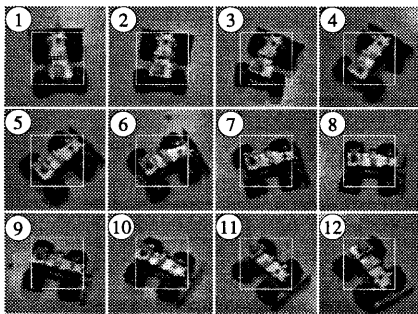


図 7: 取得した参照画像 (視軸まわりの回転)

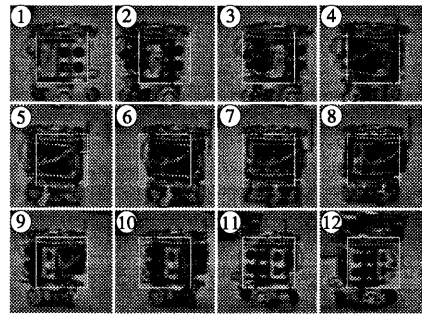


図 9: 取得した参照画像 (鉛直軸まわりの回転)

図 6 から, $\text{time}=138, 148, 207, 214[\times 1/30\text{s}]$ のときに a_{12} が式 (1) の範囲からはずれたため参照画像を更新していることがわかる. この更新時に取得した参照画像は図 7 の 2, 3, 4, 5 である.

本論文で提案するアフィン運動記述を用いた視覚追跡アルゴリズムにより, 視軸まわりに回転する目標物体を追跡することができた.

4.2 鉛直軸まわりに回転する物体の追跡

目標物体が右方向へ並進し, 途中で U ターンをして, 左方向に並進した場合の追跡を行った. このとき, 目標物体は U ターンをする際, 鉛直軸まわりの回転運動を行っている.

追跡中の \hat{A} の各成分の値の時間変化を図 8 に示す. また, 追跡中に取得した参照画像を図 9 に示す.

図 8 では, $\text{time}=67, 69, 98, 122[\times 1/30(\text{s})]$ のときに参照画像を更新したことを示している.

$\text{time}=67[\times 1/30\text{s}]$ では図 9 の 1 から 2 へ参照画像を更新している. このとき, a_{11} の値が小さくなっていくことから, 目標物体が鉛直軸まわりに回転していることがわかる. そこで, 左右 3 個ずつのウィンドウより得られるマハラノビス距離を比較することにより, $\text{time}=69$ において探索領域を左に移動させ, 3 の参照画像を取得している.

本論文で提案するアフィン運動記述を用いた視覚追跡アルゴリズムにより, 鉛直軸まわりに回転する目標物体を追跡することができた.

4.3 オクルージョンが発生している場合の追跡

目標物体が左から右へ並進し, その一部が静止している他の物体に隠された場合の追跡を行った.

追跡中の各ウィンドウの相関誤差値の分散を図 10

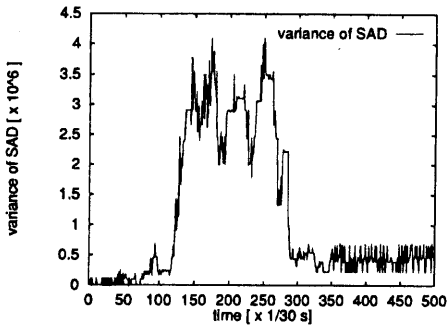


図 10: 相関誤差値の分散

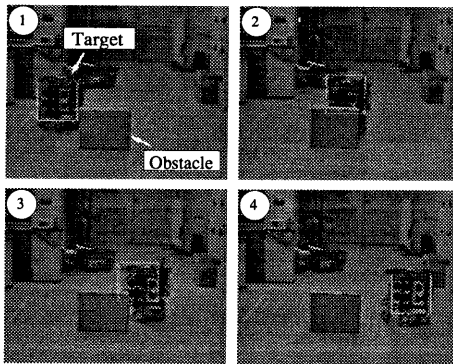


図 11: オクルージョンが発生した場合

に示す。また、追跡中の画像を図 11 に示す。

目標物体は、 $\text{time}=126\sim 285[\times 1/30\text{s}]$ の間、静止した他の物体によって隠されており、図 10 からその間の相関誤差値の分散が大きくなっていることがわかる。

3章で述べた手法により、オクルージョンが発生している場合でも目標物体を追跡することができた。

また、このときの閾値 σ_{DT} は 10^6 である。

5 まとめ

本論文では、アフィン運動記述を利用した相関法による視覚追跡アルゴリズムと、複数のウィンドウによるオクルージョンに対応した視覚追跡アルゴリズムを提案した。これにより、目標物体の見え方が変化した場合とオクルージョンが発生した場合の視

覚追跡を実現した。実環境での実験により、本アルゴリズムの有効性を検証した。

今回は視軸まわり、鉛直軸まわり、水平軸まわりのいずれかがおきた場合に対応できるがこれらが同時におきた場合を想定していない。しかし、アフィン運動パラメータから回転軸を推定することにより、同様のアルゴリズムを用いることができる。

また、オクルージョンが発生している場合の追跡アルゴリズムとアフィン運動記述を用いた追跡アルゴリズムを別々に考慮したが、オクルージョンの判定に、相関誤差値の分散ではなくアフィン運動パラメータを推定する際の最小自乗誤差を用いることにより、この2つのアルゴリズムを融合することができる。

参考文献

- [1] 井上, 稲葉, 森, 立川. 局所相関演算に基づく実時間ビジョンシステムの開発. 日本ロボット学会誌, 第 13 巻, pp. 134-140, 1995.
- [2] 内川, 沢崎, 青木, 森田, 稲本, 佐藤, 稲葉, 井上. ビデオレイトトラッキングビジョンの実用化. 第 12 回ロボット学会学術講演会, pp. 345-346. 日本ロボット学会, 1994.
- [3] 森, 稲葉, 井上. 相関法による視覚追跡における画像の記憶と合成. ロボティクス・メカトロニクス講演会'95 講演論文集 (B), pp. 1080-1083. 日本機械学会, 1995.
- [4] Trevor J. Darrell and Alex P. Pentland. Recognition of Space-Time Gestures using a Distributed Representation. M.I.T. Media Laboratory Vision and Modeling Group Technical Report No.197.
- [5] Takayuki Nakamura and Minoru Asada. Motion Sketch: Acquisition of Visual Motion Guided Behaviors. In *IJCAI'95*, Vol. 1, pp. 126-132, 1995.