

鮮明な画像撮影のための高速追従型アクティブカメラ

大池 洋史[†] 呉 海元[†]
加藤 丈和[†] 和田 俊和[†]

本論文では、高速に運動している物体を対象として、それを鮮明に撮影するためのアクティブカメラ制御方法を提案する。知覚モジュール（検出部）では、視線方向を変更しても運動視差が生じない視点固定型カメラで画像を撮影し、単色対象の追跡を安定化させるために、色空間内でのヒストグラムのシフトを行う改良型 mean shift 法によって画像上での対象の重心位置を追跡する。行動モジュール（制御対象）では、パン・チルトユニット (PTU) の速度制御に PID 制御を採用する。これは、対象の速度を P 成分、位置を I 成分、そして対象の予測運動を D 成分に対応付けることにより、対象の運動予測を行いながら、カメラの回転速度を対象の速度に合わせると同時に、対象を画像の中心でとらえるというカメラ制御が実現できるからである。実験結果より、本手法がアクティブカメラ制御法として有効であることを確認した。

A High-Performance Active Camera System for Taking Clear Images

HIROSHI OIKE,[†] HAIYUAN WU,[†] TAKEKAZU KATO[†]
and TOSHIKAZU WADA[†]

This paper presents a new method for tracking an object actively by controlling a pan-tilt camera. While tracking a moving target, we can capture motion-blur-free target images, because this method automatically synchronizes the target motion with the camera motion. For the apparent target tracking on image space, we propose a modified mean shift tracking method that adjusts the color histogram location in the color space dynamically for realizing robust tracking of monochrome object. As for the pan-tilt control, we employed the PID control scheme. This is because if the P component correspond to the object speed, the I and D components can be corresponded to the angular position (direction) and the angular acceleration, respectively. This means the PID control is suitable for controlling the angular speed and position of the pan-tilt unit simultaneously. The extensive experiments demonstrate that our method, i.e., PID based pan-tilt control, is effective for the motion synchronization between the target and camera.

1. はじめに

コンピュータビジョン (CV) の研究では、主にカメラによって撮影された画像の解析問題が取り扱われてきた。しかし、高速に運動する物体を対象とした認識・理解の問題では、画像の解析問題に加え、どのようにアクティブカメラを制御して鮮明な画像を撮影するか、という問題が重要になってくる。

色やエッジなどを正確に解析するためには、対象をぶれなく鮮明に撮影した画像が必要であり、そのためには対象の運動にあわせたカメラ制御が必要となる。すなわち、能動追跡システムにおいて、カメラ制御は画像解析に基づいて行われ、カメラ制御によって画像の鮮明さが変化し、画像解析の精度に影響を与えると

いうように、画像解析とアクティブカメラの制御は密接な相互依存関係を持っている (図 1 参照)。したがって、これまで CV ではあまり研究されてこなかったカメラ制御の問題は、能動追跡システムにおいては画像解析と同等の重要性を持つと言える。

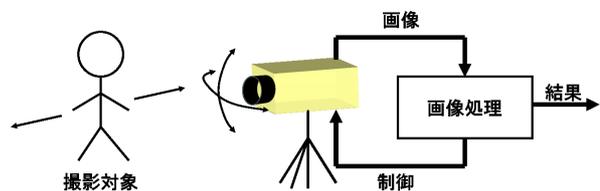


図 1 画像処理とアクティブカメラ制御

本論文では、高速に運動している物体を対象として、その画像を鮮明に撮影するために、

- 対象の運動速度とカメラの回転速度を一致させる

[†] 和歌山大学
Wakayama University

- 対象を画像の中央でとらえる
- 対象の運動予測を行う

という3条件を同時に満足するアクティブカメラ制御方法を提案する。

1.1 従来研究とその問題点

ロボットビジョンの分野で一般的に用いられるビジュアルサーボ⁵⁾では、対象の運動を考えず、「見え方」だけに基づいてカメラを動かす、はじめに与えられた目標の画像が撮影できるまで、「見る」「動く」という動作を繰り返す手法であり、高速に運動している物体の追跡には向いていない。

パン・チルト・ズームカメラを用いて運動物体を実時間で追跡する研究も行われている¹⁾²⁾。これらの手法では、カメラで撮影された画像内の追跡対象の「位置」だけ着目している。すなわち、追跡対象の位置を予測して、「次の撮影時刻において、パン、チルトの値をいくらに設定すれば対象を画像中央部にとらえられるか」という基準でパン・チルトカメラの制御を行っている。この手法では、Stop-and-Goの動作を繰り返すことしか実現できず、対象撮影時にはカメラが常に停止していることになる。このため、高速に運動する物体をぶれなく鮮明に撮影することは本質的に不可能であると言える。(図2左)

1.2 本研究の基本アイデア

本論文では、図2右に示すように鮮明な画像を撮影するために、Stop-and-Goではなく、対象を撮影しながらパン・チルトカメラを制御し、1)対象の運動予測を行いつつ、2)対象とカメラの速度を一致させ、さらに、3)追跡対象を画像中央部にとらえるというパン・チルトカメラの制御方法を提案する。また、このシステムで採用したmean shift法は単色物体の追跡が非常に不安定であるという欠点があるため、単色物体でも安定に追跡できるように改良したmean shift法も提案する。この方法は、画像空間内でmean shiftを行うだけでなく、色空間内でのヒストグラム位置もシフトさせる方法である。これらの提案手法に基づいて、能動追跡システムを試作する。

試作システムは、以下に示す知覚モジュール(検出部)と行動モジュール(制御対象)から構成される。

知覚モジュール(検出部)では、パン・チルトを変更しても運動視差が生じない視点固定型カメラを用いて画像を撮影し、改良型mean shift法を用いて画像上の対象追跡を行う。追跡結果は、画像上の対象の重心位置として行動モジュールに渡される。このように行動モジュールの状態に依存しない対象追跡手法を用いているため、行動モジュールの内部状態(カメラパラ

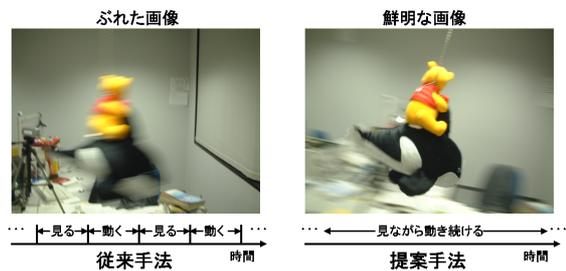


図2 従来手法と提案手法における取得できる画像と行動の流れ

メータの現在値)を参照しなくても、画像解析が行え、画像撮影時にカメラを停止させる必要がない。

行動モジュール(制御対象)では、知覚モジュールから送られてきた対象の重心位置と現在のパン・チルトの値から運動対象の方位と角速度を計算し、それに従ってパン・チルトの回転速度を制御する。この制御には、PID制御を採用する。これは、対象の角速度をP成分、方位をI成分、そして対象の予測運動をD成分に対応付けることにより、前述の3条件を同時に満足するパン・チルト制御が実現できるからである。

2. PID制御による実時間対象追跡

2.1 鮮明な画像の取得におけるPID制御の有効性

本研究の目標である「鮮明な画像の取得」を達成するために、次の2つを満たす必要がある。

- 追跡対象を画像の中心で捉えていること
- アクティブカメラの回転角速度と、カメラを中心とした追跡対象の移動速度が一致していること

そこで、提案システムはa.画像中心、b.追跡対象の移動速度を制御の目標値とする。それにあわせたPID各成分の役割は以下ようになる。

P成分 知覚モジュール(検出部)から得られる、追跡対象の現在の角速度を追従する。

I成分 速度の積分成分、つまり追跡対象の移動距離。この項によって、画像上での追跡対象の位置と画像中心とのずれが補正できる。

D成分 速度の微分成分、つまり追跡対象の加速度。この項によって、対象の未来の角速度が予測できる。

このように、PID制御則を用いることにより、対象の速度予測をしながら、対象とパン・チルトユニットの速度と方位を一致させることが可能である。

2.2 実時間対象追跡問題における制御の特殊性

提案システムは図3に示し、PID調節部、制御対象、検出部1、検出部2の4つのブロックから構成する。提案システムでは、視点固定型パン・チルト・ズーム

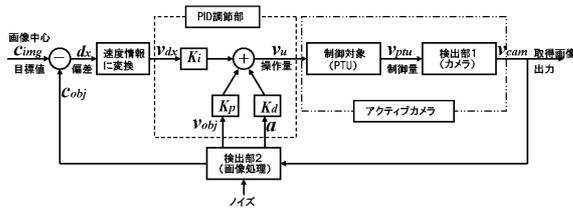


図3 システムのブロック図

(FV-PTZ: Fixed Viewpoint Pan-Tilt-Zoom) カメラ (図4) を利用し、システムの検出部1 (カメラ) が制御対象 (PTU: Pan-Tilt Unit) の上に固定されている (以下、アクティブカメラと呼ぶ)。検出部1 で得られた画像に対して、検出部2 で追跡対象の検出・追跡処理を行い、観測値・予測値を計算し、PID 調節部でこれらの値と目標値を比較し、操作量を決定する。



図4 アクティブカメラ

このようなシステムを構築するためには、以下の2つの制御の特殊性を考慮しなければならない。

- a. 検出部と制御対象の時間的特性の異なる構造をもったモジュールをうまく統合しなければならない。検出部 離散時間サンプリングされた画像情報に基づいて対象の運動速度、加速度を推定する。制御対象 PTU は次に新しいコマンドを受けるまで、最後に受けたコマンドに依存して時間的に連続に運動する。カメラは PTU の上に固定されているので、同じ運動特性を持っている。
- b. 画像情報 (色やエッジなど) を用いて安定に対象を追跡するためには、追跡対象が画像中に鮮明に写っていないといけない。本研究では、検出部1 (カメラ) によって得られる画像の質は、追跡対象に対するアクティブカメラの速度の正確さに依存する。また、アクティブカメラの動く速度の正確さは、検出部2 における画像上での対象追跡処理の正確さに依存する。検出部2 における処理の正確さは、検出部1 によって得られる画像の質に依存する。このように、検出部と制御対象の間に相互依存関係が生じる。したがって、操作量を決定する際、これらの相互関係を考慮しなければならない。

3. mean shift 法の改良

本論文では、アクティブカメラで得られた画像から運動対象を実時間で追跡するために mean shift 法³⁾ を改良して用いる。

3.1 mean shift 法

基本的な mean shift の計算手順は以下のように表すことができる。

1. 初期化: 追跡対象の位置を手動で楕円を指定し、その楕円内部の色ヒストグラムをモデルヒストグラムとしてあらかじめ求めておく。
2. 追跡: 新しく入力された画像に対して前フレームでの楕円位置からスタートして、モデルと入力画像の楕円中の色ヒストグラムの類似度が大きくなる方向に楕円を移動し、類似度が最大となる位置を追跡対象の位置と推定する。この処理を反復する。上記の計算を行う際、楕円の中心部で高い値を持ち周辺部で低い値を持つように定義された滑らかな kernel 関数を用いて各画素に重みづけをする。これは、ヒストグラム間の重なりによって評価される類似度関数を空間的に滑らかにする効果を持つ。この滑らかな類似度関数の特性から、山登り計算によって効率的に類似度が最大となる楕円配置を計算することができ、類似した物体が近接した場合でも誤って追跡することはほとんど起きない。

3.2 単色対象追跡のための mean shift 法の改良

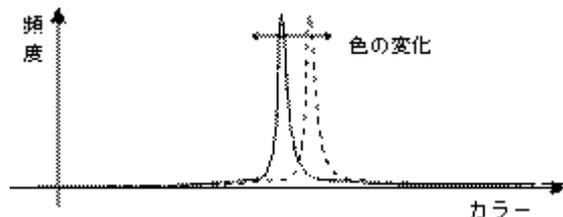


図5 色ヒストグラム間類似度の問題

上述のように mean shift 法は、高速であり類似物体の近接に強いという利点を持っているが、単色対象の追跡に弱い。これは、対象が単色の物体であるため鋭い単峰性のカラーヒストグラムが得られ、ヒストグラムの重なりを評価して類似度を求めているため、照明などの影響で対象の色が変化すると、図5に示すようにその変化がわずかでもヒストグラム間の重なり部分が急激に減少するためである。本論文では、この問題を解決するために、対象を追跡する際、以下に示す方法によって、現フレームのヒストグラムと、モデルヒストグラム同士が重なるようにシフトする方法を提案

する．

画素値を i , モデルヒストグラムを $h_m(i)$, 入力画像から得られたヒストグラムを $h_c(i)$ とし, 類似度 sim を以下のように定義する．

$$sim = \sum_{i=0}^n \min\{h_m(i), h_c(i)\} \quad (1)$$

これは, モデルヒストグラムが最大となる画素値 p_m と, 入力画像から得られたヒストグラムの値が最大となる画素値 p_c をそれぞれ求め, それらに対応させて類似度を計算する方法である． $\Delta_{mc} = p_m - p_c$ とし, 以下の式でモデルヒストグラムをシフトし, 正規化を行う．

$$H_m(i) = h_m(i + \Delta_{mc}) \quad (2)$$

正規化されたヒストグラムを用いて 類似度 sim を計算する．

$$sim = \sum_{i=0}^n \min\{H_m(i), h_c(i)\} \quad (3)$$

以上の方法を採用することにより, シャープな単峰性ヒストグラムを持つ単色物体の追跡は安定化するが, 逆に多峰性ヒストグラムを持つ物体の追跡が不安定になる可能性がある．この問題に対処するためには, 追跡対象のモデルヒストグラムを獲得した際に, その分散の大小に応じて通常の mean shift 法, 改良型 mean shift 法のいずれを用いるかを判断する必要がある．

4. 速度と位置を考慮した制御法

4.1 検出部・制御対象の統合

追跡対象の3次元運動速度は, 画像上での対象の見掛けの角速度と, アクティブカメラの運動角速度の和で表される．画像上での追跡対象の見掛けの角速度は, 連続する撮影時刻間での画像上での対象位置によって求められる．

一方, アクティブカメラの角速度は, 時間的に連続に変化するものである．実時間で対象を追跡し, フレームレートを落とさないようにするには, カメラのサンプリングに同期して, アクティブカメラの角速度も同じ周期で変化する速度信号として計算しなければならない．

これを実現するためには以下のことが必要である．

- 制御対象 (PTU) の電子的, 機械的運動特性を調べておくこと
- 制御対象に命令を送るタイミングの決定
- 画像取得と, 制御対象に命令を送るタイミングとの時間差を考慮すること

本論文では, 検出部・制御対象の統合に伴って, 以下の内容を考慮して速度指令を送信するタイミングを決定する．

- すべての処理におけるフレームレートを落とさない
- 画像処理に要する処理時間
- 速度指令を送る際に必要となる時間
- 速度指令を出してから等速運動になるまでの時間
- パンとチルトの両方を制御するための所要時間
- 速度指令を出す周期はできるだけ短くする
 1. を考慮し, 制御対象に速度指令を送信するタイミングはカメラのサンプリングに同期して行う．

2. ~ 6. はシステムを構成する環境に大きく依存するが, フレームレートの m 倍 (m は整数) が速度指令を送信するサンプリングレートとなる．

本論文では, 制御対象の連続的な動きを離散化して, 検出部の離散的な時間特性に合わせている．こうすることにより, フレームレートが安定し, 速度予測の精度が高まる．

4.2 追跡対象の角速度推定

フレーム k が撮影された時刻を t_k とする． k における追跡対象の角速度 $v_{obj}(k)$ を次の式で表す．

$$v_{obj}(k) = v_{Robj}(k) + v_{cam}(k) \quad (4)$$

ここで, $v_{Robj}(k)$, $v_{cam}(k)$ はそれぞれ k における追跡対象の見掛けの角速度, アクティブカメラの角速度を表す．以下, $v_{Robj}(k)$, $v_{cam}(k)$ を求める方法について述べる．

4.2.1 追跡対象の見かけの角速度推定

k における追跡対象の画像上での見掛けの速度 $v_{iobj}(k)$ を次の式で求める．

$$v_{iobj}(k) = \frac{c_{obj}(k) - c_{obj}(k-1)}{\Delta t} \quad (5)$$

ここで, $c_{obj}(i)$ は i における追跡対象の画像内の重心座標, Δt はフレーム間の時間である．次の式より, $v_{iobj}(k)$ を実空間での見かけの角速度 $v_{Robj}(k)$ に変換する．

$$v_{Robj}(k) = f \cdot v_{iobj}(k) \quad (6)$$

ここで, f は焦点距離である．

4.2.2 制御対象の角速度推定

あらかじめ求めておいた制御対象の運動特性から, カメラのフレームレートと同期して, 離散的速度信号として計算する．つまり k と $(k-1)$ のフレーム間の制御対象の平均角速度を求める．4.1 の c. で述べた, 考慮すべき点として, $k = m \cdot n + 1$ の時, k と $(k-1)$ フレーム間において, 撮影時刻と制御対象が新しい命令で動きはじめるまでのディレイを考える必要がある．図6に撮影時刻と制御対象が動き始めるまでの時間の

関係を示す。ここで Δt_p は検出部が画像を取得してから $v_u(i)$ を決定し、命令を制御対象に送信し終わるまでの時間を表している。 $v_u(i)$ は後に述べる制御則にしたがって決定する。 Δt_p は計算機の内部クロックによって求めることができる。これらを考慮すると、あるフレーム k における PTU の角速度 $v_{ptu}(k)$ は次式のように求められる。

$$v_{ptu}(k) = \begin{cases} \frac{1}{\Delta t} \left(\int_{t_{k-1}}^{t_{k-1} + \Delta t_p + \tau} g_{v_u(n \cdot m - 1)}(t) dt \right. \\ \quad \left. + \int_{t_{k-1} + \Delta t_p + \tau}^{t_k} g_{v_u(n \cdot m)}(t) dt \right) & (k = m \cdot n + 1) \\ \frac{1}{\Delta t} \int_{t_{k-1}}^{t_k} g_{v_u(n \cdot m)}(t) dt & (\text{上記以外}) \end{cases} \quad (7)$$

カメラは制御対象である PTU に固定されているので、カメラの角速度 $v_{cam}(k)$ は $v_{ptu}(k)$ と一致する。

$$v_{cam}(k) = v_{ptu}(k) \quad (8)$$

以上式 (4)~(8) を解くことによって k フレームにおける追跡対象の角速度を推定することができる。

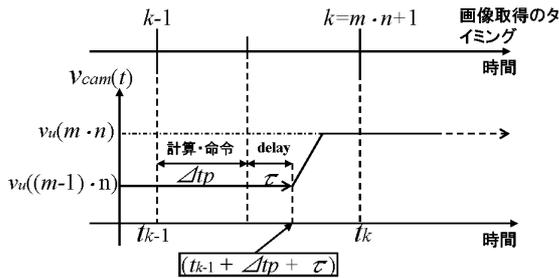


図 6 撮影時刻と、制御対象が動き始めるまでの時間のずれ

4.3 PID 制御則に基づく操作量の決定

推定された追跡対象の角速度をもとに操作量を決定するとき、P, I, D それぞれの成分を以下のように計算する。

P 成分 式 (4) で求めた追跡対象の実際の角速度。

I 成分 画像中心 c_{img} と、追跡対象の重心 $c_{obj}(k)$ のずれを補正する項である。この項は位置成分であるが、制御対象に与える命令は角速度形式なので、角速度形式に変換し、I 成分 $v_{dx}(k)$ として加える。

$$v_{dx}(k) = \frac{f \cdot (c_{obj}(k) - c_{img})}{\Delta t} \quad (9)$$

D 成分 追跡対象は、操作部に命令を送るサンプリング時間内に等角加速度運動をしていると仮定して、追跡対象の運動を予測する。D 成分 $a(k)$ は以下の式で表される。

$$a(k) = \frac{v_{obj}(k) - v_{obj}(k-1)}{\Delta t} \quad (10)$$

以上の式 (4)(9)(10) を用いて、制御則が決定できる。

しかし、これらの式から求められる値は、すべて検出部 2 の画像処理結果に依存するため、ノイズとして間違った追跡結果が混ざっている可能性がある。もし間違った速度情報を用いて操作量を決定してしまうと、システムが不安定になる。そこで本研究では過去 z フレーム間の速度情報と、 w 個の加速度情報のそれぞれの平均値を制御則に用いることによってシステムの安定化を計っている。以下に本研究の制御則を示す。

$$v_u(k) = K_p \cdot \frac{1}{z} \sum_{i=0}^z v_{obj}(k-i) + K_i \cdot v_{dx}(k) + K_d \cdot \frac{1}{w} \sum_{i=0}^w a(k-i) \Delta t \quad (11)$$

但し、 K_p, K_i, K_d はそれぞれ PID 成分の係数である。

5. プロトタイプシステムの作成

以上に述べた制御法を用いてプロトタイプシステムを作成した。

このシステムでは PTU として Directed Perception, Inc. model PTU-C46, カメラとして SONY DFW-VL500 を用い、PTU にカメラを固定し FV-PTZ カメラとして用いている。この PTU のパンとチルトは同時に動かすことができ、ハーモニックギアが使用されているためギアの間隙によるバックラッシュはない。撮影できる範囲はパン： $-180^\circ \sim 180^\circ$ 、チルト： $-80^\circ \sim 30^\circ$ 、静止以外の最低角速度はパン・チルトともに約 $2.9315^\circ/\text{秒}$ 、最大角速度はカタログスペックではパン・チルトともに約 $205.8251^\circ/\text{秒}$ となっているが、実際にこの速度で動作させるとパルスモータが脱調を起こす危険性が高い。

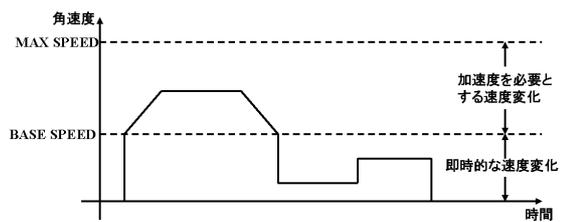


図 7 制御対象の速度変化の波形

本プロトタイプで使用する FV-PTZ カメラの速度変化特性は、図 7 に示した通り台形である。あらかじめ設定されている BASE SPEED までの加速および BASE SPEED からの減速は、即時的に行われる。本研究ではこれらの所要時間を 0 として考える。

5.1 FV-PTZ カメラの運動特性

FV-PTZ カメラの運動特性を実験的に調べた。

5.1.1 電子的運動特性

- 1つの命令を送るために、必要な処理時間は約0.01秒である。
- 実際の角速度 x (度/秒) から制御対象特有のパラメータ X (pos/秒) への変換式は

$$X = 19.4436943x \quad (12)$$

になる。実際の処理における X は整数なので小数点以下は切り捨てる。

5.1.2 機械的運動特性

- BS を BASE SPEED, a_c を制御対象の角加速度、命令 $v_u(i)$ が入力された時刻を t_{in} , t_{in} における初速を v_o , 制御対象が角加速度運動をしている時間を Δta , $td = t_{in} + \tau$, $ts = td + \Delta ta$ とし、 $g_{v_u(i)}(t)$ を以下の式で表す。

$if((v_u(i) \leq BS) \cap (v_o \leq BS))$ (BASE SPEED 以下)

$$g_{v_u(i)}(t) = \begin{cases} v_o & (t_{in} \leq t < td) \\ v_u(i) & (t \geq td) \end{cases} \quad (13)$$

$else if(v_u(i) \geq v_o)$ (加速時)

$$\Delta ta = \frac{v_u(i) - \max\{BS, v_o\}}{a_c} \quad (14)$$

$$g_{v_u(i)}(t) = \begin{cases} v_o & (t_{in} \leq t < td) \\ a_c t + \max\{BS, v_o\} & (td \leq t < ts) \\ v_u(i) & (t \geq ts) \end{cases} \quad (15)$$

$else$ (減速時)

$$\Delta ta = \frac{v_o - \max\{BS, v_o\}}{a_c} \quad (16)$$

$$g_{v_u(i)}(t) = \begin{cases} v_o & (t_{in} \leq t < td) \\ -a_c \cdot t + v_o & (td \leq t < ts) \\ v_u(i) & (t \geq ts) \end{cases} \quad (17)$$

- 命令が送信してから制御対象が動き始めるまでのディレイは $\tau = \frac{1}{60}$ 秒である。

5.2 FV-PTZ カメラの初期設定

FV-PTZ カメラの初期設定として、BASE SPEED, 加速度を決めておく必要がある。

式(13)(15)(17)を比較してみると、式(13)だけがステップ関数になっている。本システムは、フレームレートに同期した離散的な信号を処理することによって制御対象を制御する。ステップ関数というのは一種の離散的な信号なので、検出部・制御対象を統合させる際都合がよい。よって、できるだけ式(13)を多用するためにBASE SPEEDを高く設定したい。また、できるだけ早く目標速度に達してほしいので加速度も高く

設定したい。

しかし、これらのパラメータを高く設定しすぎると機械的な問題が生じる。図7に示すように、制御対象の運動速度はBASE SPEEDまで一気に上がる。この時、PTU内部に大きな負荷がかかり、急激な速度変化が起こるとパルスモータが脱調を起こしてしまう。加速度も高く設定しすぎると、同様の事態が生じる。プロトタイプシステムでは、このような事態が生じないように、できるだけパラメータを低く設定した：

$$\text{BASESPEED } BS = 100^\circ/[\text{sec}] \quad (18)$$

$$\text{acceleration } a_c = 1590^\circ/[\text{sec}]^2 \quad (19)$$

5.3 命令送信のタイミング

カメラのサンプリングレートは毎秒30フレームなので、すべての処理をあわせて約0.0333秒以内に完成する必要がある。あらかじめ調べた制御対象の電子的運動特性より、一つ命令を送るために約0.01秒を要することが分かった。もし1フレームの間にパン・チルトそれぞれについて命令すると、残りの処理時間は約0.0133秒となる。これだけの処理時間では、対象追跡の画像処理、操作量の決定は行えない(提案システムの計算機のCPUはIntel Pentium4 3GHzである)

したがって、パン・チルトの速度制御指令は交互に行い、それぞれ2フレームに一回命令を送ることにした。式(18)(19)と、今回使用するカメラの最高運動角速度から計算できるように、2フレーム後には、制御対象であるPTUの角速度が操作量に必ず達しており、4.1で述べた条件4も満たしている。

6. プロトタイプシステムの性能評価実験

6.1 プロトタイプシステムの性能比較実験

提案した制御則の有効性を定量的に示すために、以下に示す、他の2つの手法と比較した。

手法1: 提案手法。式(11)中の $Kd = 1.0$, $Ki = 1.0/30.0$, $Kd = 1.0$, $z = 3$, $w = 3$ とする。

手法2: 対象の速度予測をしない方法。すなわちPI制御である。他の条件は手法1と同じ。

手法3: 速度固定で位置だけ制御する方法。条件は、BASE SPEED, 命令を送るサンプリングは手法1と同じで、速度は最大速度の205.8251°/秒に設定する。

これらの手法を比較する際に同じ動き、同じ環境で実験しないと比較できないので、繰り返しほぼ同じ動作ができるような振り子を用い、ぶら下がっているボールを追跡対象とした。このボールを同じ位置から自然落下させ、落下しはじめてから300フレーム(約10秒)間追跡した。カメラからボールまでの距離は約

1.6m, ボールの直径は約 20cm である。また,ズームは固定で,校正係数 $f = 1.0/14.0$ (°/pixel) となっている。

提案システムでは検出部と制御対象が一体となっており,実空間での追跡対象の速度と,制御対象の運動速度を比較することが困難である。そこで,性能評価の基準を,前フレームと,現フレームの追跡対象の重心の差とする。もし速度が完全に一致していればこの値は 0 になる。よって,この値が小さいほど速度の差が小さいことを示しており,提案手法の速度追従性の尺度と見なすことができる。

各手法におけるフレーム間の重心のずれにおける絶対値の平均,標準偏差を表 1 にまとめ,表から提案手法である手法 1 の誤差が最も小さく,提案手法の速度追従性が一番優れていることがわかる。

表 1 前・現フレーム間の追跡対象の推定重心位置の平均角度差

$1^\circ = 14pixel$	PAN		TILT	
	平均 (°)	標準偏差	平均 (°)	標準偏差
手法 1	0.150502	0.754835	0.147157	0.695240
手法 2	0.230769	0.840937	0.160535	0.723666
手法 3	0.438127	1.186173	0.190635	0.779959

表 2 画像中心と追跡対象の推定重心位置の平均角度差

$1^\circ = 14pixel$	PAN		TILT	
	平均 (°)	標準偏差	平均 (°)	標準偏差
手法 1	3.180602	5.852119	1.842809	3.313933
手法 2	7.284281	7.177534	2.521739	4.427010
手法 3	5.197324	10.182353	1.685619	3.560023

また画像中心と,追跡対象の重心の差における絶対値の平均,標準偏差を表 2 にまとめる。手法 1~3 を比べると,パン方向では手法 1 が最も誤差が小さい。一方,チルト方向では,手法 1 は手法 3 よりわずかに誤差が大きい。この原因は実験の際,追跡対象の運動に伴う照明条件の変化により,追跡対象の上部の色情報が照明の反射によって損なわれる入力画像系列があり,それによって追跡対象の推定位置が,上下に激しく振動していたためである。手法 3 では予測を行わず,位置情報しか扱っていないので振動による影響は瞬間的に消えるが,手法 1,2 では,対象の運動速度をもとに制御しているため,過去の情報の影響がある程度残る。それによって推定誤差が大きくなってしまったと考えられる。

また文献¹⁾で紹介されている方法では,床面上を円周運動するラジコン車を追跡する実験において,画像中心と追跡対象の推定重心位置の平均角度差は 6.571°

となっている。運動対象が異なるため直接の比較は行えないが,本手法の場合は対象の最高角速度が 150° 以上であるにもかかわらず,平均角度差は 3.180602° となっている。

以上の結果から,今回提案する手法 1 は位置追従性の点でも最も優れていると言える。

比較実験で得られた画像系列の一部を図 8 に示す。frame.1 は初期位置である。frame.15, 89 では追跡対象は画像の右から左に向けて移動している。frame.144 は,振り子運動の折り返し点であるため,追跡対象は静止している。それぞれの手法で得られた画像系列を比較してみると,手法 1 が,最も安定して画像の中心付近で追跡対象を捉えることに成功している。また,追跡対象のぶれも最も少ない。

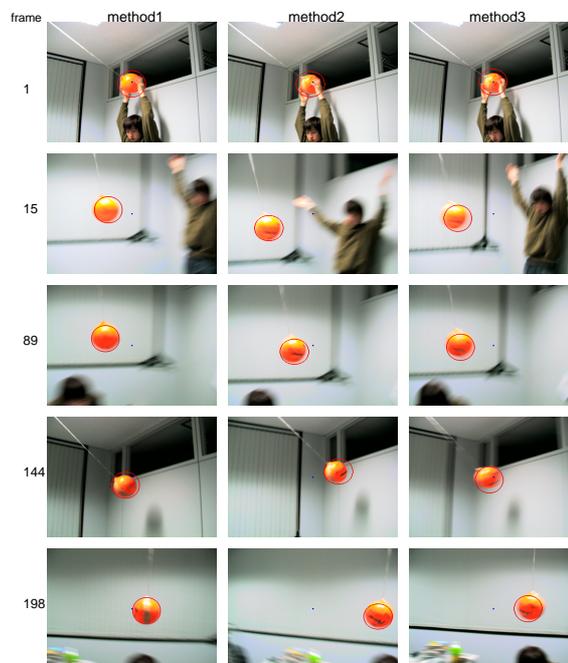


図 8 性能比較実験の画像例

比較実験における対象の運動角速度は最高 150° /秒以上に達しており,提案する手法 1 は高速に移動する物体でも追跡できることがわかる。このように,提案手法は「鮮明な画像を取得し続けるために,追跡対象を画像中心付近で捉え,かつ追跡対象の運動速度とアクティブカメラの運動速度を対応させる」というタスクの達成度が高いと言える。

6.2 人の頭部追跡実験

プロトタイプシステムを用いて頭部追跡実験を行った。追跡によって得られた一部の画像を図 9 に示す。



図 9 人の頭部追跡の様子

frame.1 の時点では静止している．frame.21～100 では、人はほぼ等速運動をしており、頭部を画像の中心付近で捉え、かつアクティブカメラの運動速度が頭部の運動速度に対応しているため鮮明な画像を取得することができる．frame.115 で再び静止し、frame.121～158 は、しゃがんで急激に立つという動作を行っている．frame.158 のときは静止しているが、急激な加速と停止のため、画像中心からずれている．このような急激な加速度の変化に対しては、追従が遅れるが、トラッキングに失敗することは少ない．frame.178～207 では、追跡対象は円を描くように運動している．frame.197 の時、顔の肌色部分の大半が隠れてしまっている画像があるが、追跡に失敗していない．これはモデルのヒストグラムに髪の色情報も入っているためである．frame.260～290 体を右側に振っている．このときも追跡対象である人の頭部を鮮明に捉えることができる．実際の人の運動には、慣性が強く働くので、画像を 30 枚/秒で取得し続ければ追跡に失敗することはほとんど起きない．

表 3 に、この実験における (A) 頭部の画像上での推定位置と画像中心との差 (B) 頭部の画像上での推定位置と前フレームの頭部の推定位置との差、平均、標準偏差をそれぞれ示す．この時の追跡対象の最高運動速度は PAN=64.1°/秒、TILT=81.4°/秒で、6.1 で行った実験における表 1, 2 と比較してわかるように、運動速度が低い対象に対して、より高い精度で追跡を行うことができる．

表 3 (A)・(B) それぞれの平均、標準偏差

1° = 14pixel	PAN		TILT	
	平均 (°)	標準偏差	平均 (°)	標準偏差
(A)	1.070234	2.343131	1.498328	2.924105
(B)	0.026756	0.322277	0.093645	0.571009

7. ま と め

本論文では、視点固定型パン・チルト・ズームカメラを用いて、高速に運動している対象を能動的に追跡し、追跡対象とカメラの運動速度および方位を可能な限り一致させることによって、実時間でぶれのない鮮明な画像を撮影する手法について述べた．知覚モジュール (検出部) では、mean shift 法を改良し、安定かつ高速な対象検出・追跡に成功した．行動モジュール (制御対象) では、PID 制御則を用いることにより、ビデオレート (30 枚/秒) で追跡対象の鮮明な画像を実時間で取得することができ、「見ながら動ける」システムを構築することに成功した．本手法により、高品質の画像が撮影できることから、後に続く認識・理解には欠かせない色情報、エッジ情報などをより高い精度で求めることができることは、CV 研究における画像撮影法として役立つものと思われる．

参 考 文 献

- 1) 松山隆司, 和田俊和, 物部祐亮: 視点固定型パン・チルト・ズームカメラを用いた実時間対象検出・追跡, 情報処理学会論文誌, Vol.40, No.8, pp.3169-3178 (1999).
- 2) 和田俊和, 松山隆司: 能動カメラによる実時間対象追跡, 日本ロボット学会誌, Vol.19, No.4, pp.433-438 (2001).
- 3) Dorin Comaniciu., Visvanathan Ramesh., Peter Meer.: Kernel-Based Object Tracking, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, Vol.25, No.5, pp.564-575 (2003).
- 4) 須田信英: PID 制御, 朝倉書店.
- 5) 出口光一郎: ロボットビジョンの基礎, コロナ社.