パン ・チルト角速度及びズーム制御による動物体の追跡撮影手法

A Shooting Method for Tracking a Moving Object by Pan/Tilt Speed and Zoom Control

東 和秀† 亀田 能成‡ 美濃 導彦‡
† 京都大学大学院情報学研究科
‡ 京都大学総合情報メディアセンター
Kazuhide HIGASHI, Yoshinari KAMEDA, and Michihiko MINOH
† Graduate School of Informatics, Kyoto University
‡ Center for Information and Multimedia Studies, Kyoto University

概要: 首振りカメラのパン ・チルト角速度、ズーム値を制御し、動物体を追跡撮影する手法を提案する。動物体の追跡撮影では、映像上での動物体の見かけの位置の変化が小さく、かつ見かけの大きさが大きくなることがよいとされる。本手法では、まず、ポジションセンサによって動物体の位置を測定し、測定時刻以後の動物体の速度を推定、位置を予測する。次に、予測結果から透視投影カメラモデルに基づいて追跡撮影として適当なパン ・チルト角速度及びズーム値を算出し、カメラを制御する。実験によって、本手法は上記の条件を満たすような映像を撮影できることを確認した。

Abstract: We propose a shooting method for tracking a moving object by pan/tilt speed and zoom control of a video camera. It is said that the video camera should shoot the object so as to make the location change of the object in the video image small and its size as large as possible. To satisfy these two conditions, first the spatial location of the object is measured by a position sensor and then the spatial location and velocity of the object are predicted in the next step. After that, these two predicted values are transformed into camera parameters (pan/tilt speed and zoom value) in order to control the camera. Experimental results showed that the video taken with our method satisfies the conditions described above.

1 はじめに

講演会、舞台、スポーツなどのイベントを撮影した映像は、テレビニュースの中で使用したり、近年のネットワーク技術を使って世界中に発信したりと、様々な用途に使用される。イベントには一般に多くの撮影対象が存在するので、例えばスポーツの場合、試合の流れを撮影した映像を見たい人やある特定の選手の映像を見たい人などが存在し、映像への要望は様々である。多種多様な映像を複数のカメラを制御して撮影するシステムに対する需要は高くなりつつある。

イベントを様々な方法で撮影する試みが行われている。石塚ら [1] は、カメラによって講義の状況を判断し、講義の撮影を行っている。しかし、映像から講師が外れないように、常に広角で撮影している。和田ら [2] は、動物体の位置が未知であるという条件下で、動物体を撮影した画像を処理することによって動物体の位置を測定し、カメラのパン ・チルト角度を一定の角速度で変化させながら追跡撮影を行う手法を提案している。しかし、追跡撮影に用いるカメラのパン ・チルト角速度は一定で、動物体の速度が急に変化するといった、パン ・チルト角速度を変化させなければならない状況には対応することができない。

動物体の撮影では、映像を見て動物体の様子が分かるように撮影しなければならない。そのためには、映像上での動物体の見かけの位置の変化が小さく、かつ見かけの大きさが大きくなければならないと考えられる。本稿では、講義中の講師の撮影を対象に、この条件を満たすような映像を撮影する手法を提案する。したがって、動物体の大きさは既知、位置 ・速度は未知、個数は 1 つと仮定する。

映像上での動物体の見かけの位置の変化を小さくするにはカメラのパン ・チルト角速度を、見かけの大きさを大きくするにはカメラのズーム値を制御することになる。ゆえに本研究では、映像上での動物体の見かけの位置の変化とカメラのパン ・チルト角速度の関係、

† 〒 606-8501 京都市左京区二本松町
Email higashi@kuis.kyoto-u.ac.jp
URL http://www.imel1.kuis.kyoto-u.ac.jp

映像上での動物体の見かけの大きさとカメラのズーム値の関係から、動物体を最適に撮影するためのパン・チルト角速度、ズーム値を求め、実際にカメラを制御して動物体を撮影する。動物体の移動速度は一定であるとは限らないので、その位置・速度に応じたカメラ制御が必要になる。我々は、パン・チルト角速度やズーム値を様々な値に変化させながら、動物体の追跡撮影を行う。

本稿では、次の 2 つの過程を経て動物体の追跡撮影を行う。

- 速度推定・位置予測... 空間内での動物体の位置を測定する。その測定と前回の測定をもとに、測定時刻以後の動物体の速度を推定し、位置を予測する。

- 撮影... 透視投影カメラモデルを用いて、予測された空間内の動物体の位置・速度に応じたカメラ制御方法を決定し、動物体の追跡撮影を行う。

本稿の構成は次の通りである。まず 2章で、システム全体の実行スケジュールについて述べる。3章では、動物体の速度推定・位置予測について述べる。4章では、動物体の追跡撮影手法について述べ、5章では、カメラの制御について述べる。実験結果を 6章に示し、7章で結論を述べる。

2 システム全体の実行スケジュール

撮影を行う空間に世界座標系 (空間) を導入する。世界座標系での動物体のある時刻 T^iでの位置を$\boldsymbol{M}_{T^i}$とする。

撮影を行うカメラに画面座標系 (平面) を導入する。画面の中心を原点とし、ある時刻 T^iでの動物体の画面座標系での位置を$\boldsymbol{m}_{T^i}$、速度を$\boldsymbol{m}'_{T^i}$とする。

システム全体の実行スケジュールは図 1、及び図 2のようになる。

計算機はポジションセンサに世界座標系での動物体の位置を要求する (図 1矢印※ 1)。ポジションセンサは常に動物体の位置を測定・保存しており、計算機からの要求があれば、最新の測定結果を計算機に送る (図 1矢印※ 2)。i 回目の位置要求の時刻を T^iとし、測定は計算機がポジションセンサに位置要求を行った時刻に行われたものとする。これは、ポジションセンサの位置測定間隔が十分短ければ有効な仮定である。計算機は、i 回目の位置要求での返信結果と $i-1$ 回目の返信結果をもとに測定時刻以後の動物体の速度を推定し、推定速度をもとに測定時刻以後の動物体の位置を予測する。位置予測結果をカメラの画面座標系、パン・チルト角度に変換し、変換した値をもとに、カメラのパン・チルト角速度とズーム値に関する制御命令を決定し、カメラに送る (図 1矢印※ 3)。カメラは、命令を受け取ると、命令を受け取ったことを計算機に知らせ (図 1矢印※ 4、ACK メッセージ)、その後、カメラを動かし始める。しかし、カメラはすぐに動き始めず、応答待機時間が生じる。計算機は、カメラが命令を受け取ったことを ACK メッセージによって確認すると、次の制御命令を決定する手続きに入る。次の制御命令の決定と同時に、カメラではすでに決定した制御命令で運動を行っている。i 回目の位置要求からカメラ運動までの一連の手続きを手続き i とする。

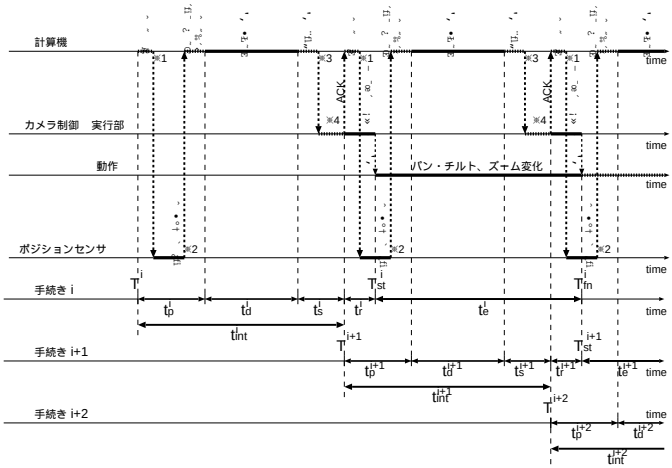


図 1: システムの流れ

手続き i に関して、位置要求から速度推定、位置予測にかかる時間を t_p^i、制御命令の決定にかかる時間を t_d^i、計算機が制御命令を発行してからカメラが ACK メッセージを返すまでの時間を t_s^i、カメラの応答待機時間を t_r^i、カメラが動いている時間を t_e^iとする。t_{int}^iを

$$t_{int}^i = t_p^i + t_d^i + t_s^i \qquad T^{i+1} = T^i + t_{int}^i$$

とする。$t_p^i, t_d^i, t_s^i, t_r^i, t_e^i$は一定とは限らないので、時間間隔 t_{int}^iも一定ではない。

手続き i に関して、カメラが実際に動いている時刻は $(i+1)$ 回目の位置要求後の $T^{i+1} + t_r^i (= T^i + t_{int}^i + t_r^i)$ から $(i+2)$ 回目の位置要求後で、手続き $(i+1)$ でのカメラ運動が始まる $T^{i+2} + t_r^{i+1} (= T^{i+1} + t_{int}^{i+1} + t_r^{i+1})$ となる。しかし手続き i の中で制御命令を決定する時点では、これらの時刻は分からない。そこで、$t_{int}^i \approx t_{int}^{i+1} \approx t_{int}^{i-1}, t_r^i \approx t_r^{i+1} \approx t_r^{i-1}$と仮定し、

$$T^{i+1} = T^i + t_{int}^i \approx T^i + t_{int}^{i-1} = \overline{T^{i+1}}$$

$$T^{i+2} = T^{i+1} + t_{int}^{i+1} \approx \overline{T^{i+1}} + t_{int}^{i-1} = \overline{T^{i+2}}$$

とすることにより、時刻$\overline{T^{i+1}} + t_r^{i-1}(T_{st}^i$とする$)$ から$\overline{T^{i+2}} + t_r^{i-1}(T_{fn}^i$とする$)$ までカメラが動くと仮定する。

カメラは、異なる種類の制御は同時に行うことができるとし、パン・チルト運動とズーム値変更運動を同時に行う。しかし、同種の運動を複数同時に行うことはできないとし、カメラは同種の新しい命令を受け取ると、現在実行中の命令に代えて、新しい制御命令を実行するものとする (図 2中の矢印)。そのため、カメラはt_{int}時間、一定のパン・チルト角速度運動を行うことになる。図 2は、位置要求からカメラへの制御命令発行までの計算段階 (Calculation Phase) とカメラのインターフェース応答待機段階 (Waiting Phase)、カメラがパン・チルト・ズームの運動を行っている運動段階 (Rotation Phase) が、パイプライン化されていることを示している。t_{int}間隔で新しい制御命令が決定し、カメラ制御が変化していく。次の手続きの計算段階及び応答待機段階では、前の手続きの運動段階が同時に行われる。

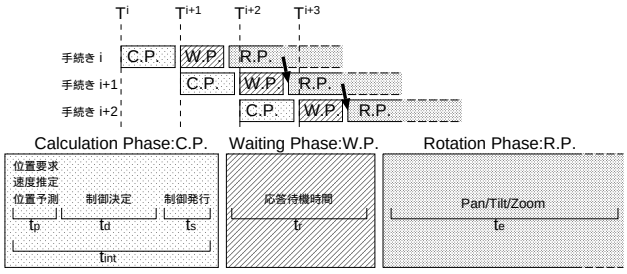


図 2: パイプライン化

3 動物体の速度推定・位置予測

世界座標系での動物体のある時刻 T^iでの位置$\boldsymbol{M}_{T^i}$をポジションセンサを使用して測定する。

次に、ある測定時刻 T^i以後の動物体の速度を推定する。動物体は位置測定時刻以後、等速直線運動を行うと仮定する。この仮定は、時間間隔が十分短ければ、十分有効な仮定である。

$$\boldsymbol{M}'_{T^i} = \frac{\boldsymbol{M}_{T^i} - \boldsymbol{M}_{T^{i-1}}}{T^i - T^{i-1}} \tag{3.1}$$

最後に、実空間における時刻 T_{st}^i, T_{fn}^iでの動物体の位置$\boldsymbol{M}_{T_{st}^i}, \boldsymbol{M}_{T_{fn}^i}$ を、次のように予測する。

$$\boldsymbol{M}_{T_{st}^i} = \boldsymbol{M}_{T^i} + \boldsymbol{M}'_{T^i}(T_{st}^i - T^i) \tag{3.2}$$

$$\boldsymbol{M}_{T_{fn}^i} = \boldsymbol{M}_{T^i} + \boldsymbol{M}'_{T^i}(T_{fn}^i - T^i) \tag{3.3}$$

4 動物体の追跡撮影手法

ここでは、次の 2 つの課題を達成する撮影手法について考察する。

- 課題 A... 映像上での動物体の見かけの速度を小さくし、見かけの位置が指定した位置を維持するようにする
- 課題 B... 映像上での動物体の見かけの大きさが大きくなるようにする

カメラ制御では、パン角度、パン角速度、チルト角度、チルト角速度、ズーム値を制御変数として想定しているが、それらの値は離散値で、量子化の単位が大きい場合には、常に最適な値を指定できるとは限らない。これを制約αと呼ぶ。

問題は、制約αを考慮しながら課題 A・B を達成することである。

4.1 映像上での動物体の見かけの位置・速度を制御する追跡撮影手法

課題 A は、映像上で指定する位置を$\boldsymbol{a}$とすると、

$$\boldsymbol{m}_T = \boldsymbol{a} \tag{4.1}$$

$$\boldsymbol{m}'_T = \mathbf{o} \tag{4.2}$$

と表すことができる。本研究では、指定する位置を画面の中心、つまり$\boldsymbol{a} = (0, 0)$ とし、カメラの光軸上に動物体が位置するようにする。

時刻 Tにおけるカメラのパン・チルト角度を$\boldsymbol{C}_T$、パン・チルト角速度を$\boldsymbol{C}'_T$、世界座標系での動物体の位置M_Tにカメラの光軸を合わせたときの理想パン・チルト角度を$\overline{\boldsymbol{C}}_T$、動物体にカメラ光軸を追従させたときの理想パン・チルト角速度を$\overline{\boldsymbol{C}'}_T$、ズーム値を Z_Tとする。また、カメラの焦点距離を f_Tとすると、f_Tは Z_Tの関数であるから $f_T = g(Z_T)$、Rを三角関数に関する行列とすると、$\boldsymbol{m}_T$は

$$\boldsymbol{m}_T = g(Z_T) \cdot R(\boldsymbol{C}_T - \overline{\boldsymbol{C}}_T) \tag{4.3}$$

と表せる。よって、映像上での動物体の見かけの位置を制御するのにカメラのパン・チルト角度を利用できる。また、$\boldsymbol{m}'_T$は

$$\boldsymbol{m}'_T = g(Z_T) \cdot R(\boldsymbol{C}'_T - \overline{\boldsymbol{C}'}_T) \tag{4.4}$$

と表せ、映像上での動物体の見かけの速度を制御するのにカメラのパン・チルト角速度を利用できる。

カメラを制御して動物体を撮影する場合、式 (4.1)、(4.2) を満たすためには、$g(Z_T) \neq 0$ であるので、式 (4.3)、(4.4) より、

$$\boldsymbol{C}_T - \overline{\boldsymbol{C}}_T = \mathbf{o} \qquad \boldsymbol{C}'_T - \overline{\boldsymbol{C}'}_T = \mathbf{o}$$

とならなければならない。しかし、制約αのため、映像上での見かけの速度が$\mathbf{o}$になるようなパン・チルト角速

度$\overline{\boldsymbol{C}'}_T$が求められても、そのパン ・チルト角速度をカメラが実現できるとは限らない。つまり、制約αのもとでは、一般に$\boldsymbol{m}'_T \neq \boldsymbol{o}$となるので、$\boldsymbol{m}_T$は一定の値をとることができない。したがって、課題 A の目標である式 (4.1)、(4.2) の両方を同時に満足させることは困難である。式 (4.1) と式 (4.2) のどちらを優先させるのかは、カメラの制御機能に依存すると考えられる。

我々が想定するカメラは、量子化の単位が、パン ・チルト角度に比べてパン ・チルト角速度の方が大きく、粗い制御となる。つまり、角速度の離散値の選択を誤ると、その誤差は大きいものとなる。そこで、本稿では、パン ・チルト角速度を優先して決定する。

4.2 映像上での動物体の見かけの大きさを制御する追跡撮影手法

課題 B は、映像上に動物体のある部分が入るような最大ズーム値 Z_T^{max} にズーム値 Z_T を近づけることを意味し、次のように表すことができる。

$$Z_T = Z_T^{max} \tag{4.5}$$

動物体の大きさは既知であると仮定したので、世界座標系でのカメラの焦点位置と動物体の位置から Z_T^{max} を求め、式 (4.5) が満足されるようにする。

5 カメラの制御

3 章で得られた動物体の位置予測は、世界座標系で表現されている。しかし、撮影において満足させられるべき式 (4.1)、(4.2)、(4.5) は、すべてカメラの画面座標系において表現されている。そこで本章では、世界座標系での予測結果を、画面座標系、カメラのパン ・チルト角度に変換し、パン ・チルト角度を用いてカメラ制御を行う手法について述べる。

5.1 パン ・チルト角速度の制御

本節では、課題 A を達成するためのパン ・チルト角速度制御手法について述べる。

映像上での動物体の見かけの速度$\boldsymbol{m}'_T$は、パン ・チルト角速度$\boldsymbol{C}'_T, \overline{\boldsymbol{C}'}_T$で制御することができる。よって、ここから、$\boldsymbol{m}'_T$の議論は$\boldsymbol{C}'_T$、$\overline{\boldsymbol{C}'}_T$で行う。

カメラの理想パン ・チルト角速度$\overline{\boldsymbol{C}'}_T$を求めるために、まず、世界座標系での動物体の位置測定結果を、カメラのパン ・チルト角度に変換する。世界座標系でのある座標を画面の中心に位置させるカメラのパン ・チルト角度を求める関数を h とし、位置測定時刻 T^i での動物体の推定位置$\boldsymbol{M}_{T^i}$から、カメラでの理想パン ・チルト角度$\overline{\boldsymbol{C}}_{T^i}$を求める。

$$\overline{\boldsymbol{C}}_{T^i} = h(\boldsymbol{M}_{T^i}) \tag{5.1}$$

同様に、時刻 $T^i + t$ でのカメラでの動物体の理想パン チルト角度$\overline{\boldsymbol{C}}_{T^i+t}$を求めることができるので、$\overline{\boldsymbol{C}'}_{T^i}$は、

$$\overline{\boldsymbol{C}'}_{T^i} = \lim_{t \to 0} \frac{\overline{\boldsymbol{C}}_{T^i+t} - \overline{\boldsymbol{C}}_{T^i}}{t}$$

で与えられる。

しかし我々は、T^i_{st}から T^i_{fn}までの間、1 つの制御しか行わず、その間$\boldsymbol{C}'_{T^i}$を$\overline{\boldsymbol{C}'}_{T^i}$と常に一致させ続けることはできない。そこで、どのような固定値$\boldsymbol{C}'_{T^i}$を与えるかが問題となる。本稿では

$$\boldsymbol{C}'_{T^i} = \frac{\overline{\boldsymbol{C}}_{T^i_{fn}} - \overline{\boldsymbol{C}}_{T^i_{st}}}{T^i_{fn} - T^i_{st}} \tag{5.2}$$

とする。ゆえに、制御決定手法は、以下のようになる。

1. 式 (5.1) を用いて、時刻 T^i_{st}, T^i_{fn} での動物体の予測位置$\boldsymbol{M}_{T^i_{st}}, \boldsymbol{M}_{T^i_{fn}}$ から、動物体にカメラ光軸を追従させたときの理想パン ・チルト角度$\overline{\boldsymbol{C}}_{T^i_{st}}, \overline{\boldsymbol{C}}_{T^i_{fn}}$を求める。

2. 式 (5.2) を用いて、動物体の平均速度を計算し、$\boldsymbol{C}'_{T^i}$のパン角速度、チルト角速度に一番近い、カメラが実現可能なパン角速度、チルト角速度を求める。

3. カメラにパン ・チルト角速度制御命令を送る。

このように決定したパン ・チルト角速度でカメラを動かすと、パン角度は図 3のように推移する。チルト角度も同様である。

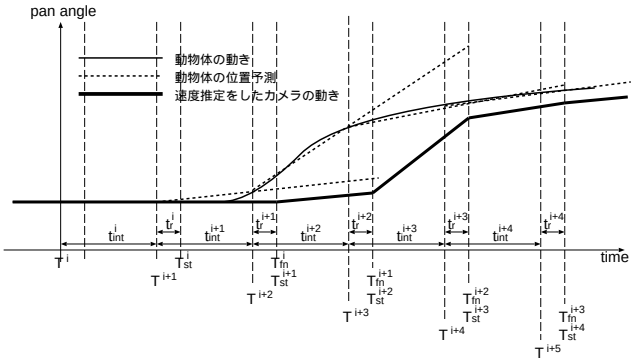


図 3: 動物体の位置予測とカメラの動作

5.2 ズーム値の制御

本節では、ズーム値 Z^i_T の制御手法を考察する。ズームの速度は一定であると仮定する。

ズーム値の制御が行われるのは、課題 B において、式 (4.5) を満足させるときであるから、Z_Tは式 (4.5) の $Z^{max}_{T^i}$を求めることによって得られる。

6 実験

6.1 実験環境

本手法に基づくカメラ制御システムを実装し、京都大学情報工学教室第二講義室において位置・速度が未知である人間を追跡撮影する実験を行った。世界座標系の X、Y、Z 軸を図 4のように設定した。

実験で使用するハードウェア環境を表 1に示す。

表 1: ハードウエア環境

計算機	Sun Ultra10-Elite3D 333MHz	× 1 台
	PC/AT Celeron 500MHz	× 4 台
ポジションセンサ	INTERSENSE IS-600 Mark2	
	本体	× 1 台
	X-Bar	× 2 台
	ビーコン	× 1 個
カメラ	SONY EVI-D30	× 3 台

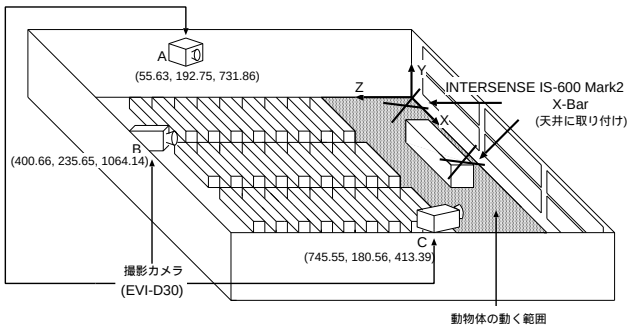


図 4: 実験環境 (単位は cm)

人間は講義室の前方で動き回るものとする。

データの通信を制御するために Sun Ultra10-Elite3D を 1 台、ポジションセンサを使用して動物体の位置測定を行うために INTERSENSE IS-600 Mark2 X-Bar を天井に取りつけ、PC/AT Celeron 500MHz を 1 台用意した。また、各カメラの制御決定・命令の発行のために PC/AT Celeron 500MHz を各 1 台用意した。計算機は 100BaseT で接続され、RS232C を通してカメラを制御する。

撮影の実験は 3 ヶ所のカメラで同時に行った。カメラは SONY EVI-D30 を用いた。SONY EVI-D30 の制御命令では、パン角速度は 19 段階、チルト角速度は 15 段階の角速度の中から設定することができる。

6.2 実験・結果

実験により、本手法が動物体の追跡撮影には有効であることを確かめた。ポジションセンサのビーコンを頭の上に置き、人間の顔が画面の中心に来るようにカメラを制御する。t_r は、実験より事前に測定を行い 150ms となることが分かった。

図 5に各カメラが撮影した映像のうち、1.0 秒毎のスナップショットを示す。上からカメラ A、カメラ B、カメラ C(各カメラの位置は図 4参照) である。位置予測の誤差は、時刻 T_{st} で $11.36cm$、T_{fn} で $16.75cm$ となった。

各パラメータの平均時間は $t_p = 13ms$、$t_d = 162ms$、$t_s = 40ms$ となり、t_{int} は $215ms$、つまり $4.65Hz$ となった。イベントを多くのカメラで撮影したいという要望に対しても、本手法では t_{int} が大きくならないという利点がある。

ポジションセンサの位置測定平均間隔は $10ms$ となった。

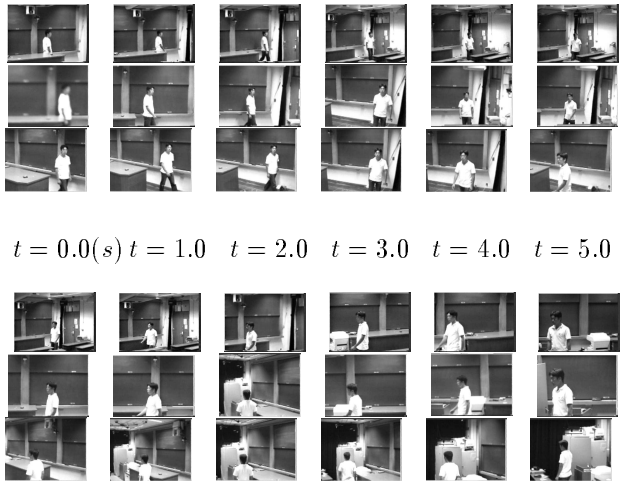
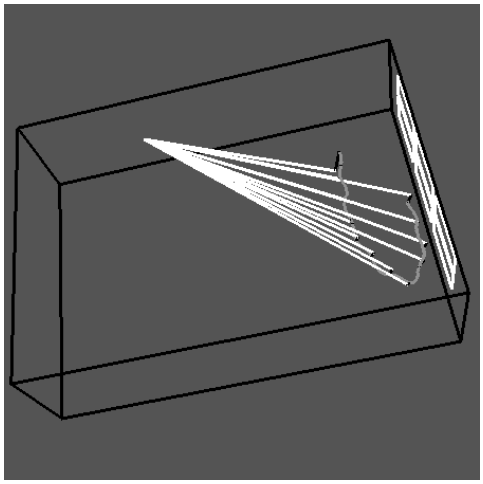


図 5: 各カメラが撮影した映像のスナップショット

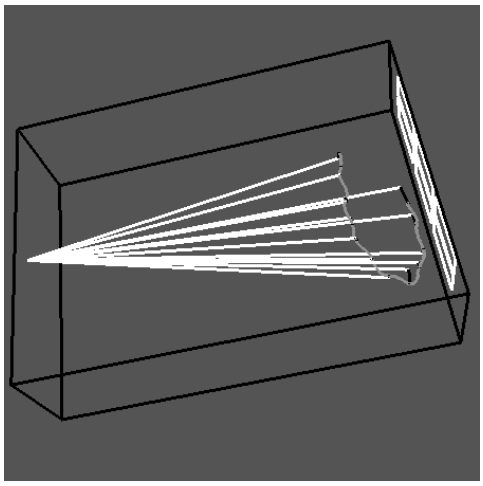
6.3 考察

図 6に動物体の軌跡 (灰) と、動物体の 1.0 秒毎の位置から各カメラの光軸 (白) に下ろした垂線 (黒) を示す。垂線は撮影時の位置ずれを意味する。また、表 2に 1.0 秒毎の各カメラの光軸と動物体の距離を示す。

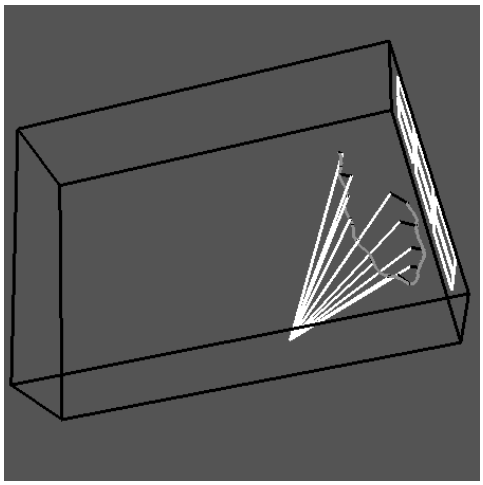
実験では、カメラ B の t=0.0 など、映像上での動物体の見かけの位置が画面の中心から大きく外れてしまうことがあった。原因として、本手法では動物体が等速直線運動をすると仮定していることや位置測定の誤差、時間パラメータの分散が考えられる。しかし、表 2より光軸と動物体の距離は最悪 72cm であり、映像から動物体が外れてしまうことはなかった。また、映像を見て常に人間の表情が分かるように動物体を撮影することができた。



(a) カメラ A



(b) カメラ B



(c) カメラ C

図 6: 動物体の軌跡 (灰) と 1.0 秒毎のカメラの光軸 (白) への垂線 (黒)

表 2: カメラの光軸と動物体の距離 (単位 cm)

時間 (秒)	カメラ A	カメラ B	カメラ C
0.0	70.73	71.71	51.49
1.0	11.35	50.99	57.73
2.0	21.90	25.38	55.58
3.0	27.51	27.74	47.55
4.0	8.39	41.24	48.60
5.0	15.34	6.38	20.65
6.0	12.11	4.29	22.96
7.0	9.90	9.90	14.30
8.0	12.32	12.08	10.62
9.0	22.07	12.86	12.65
10.0	23.68	2.09	33.07
11.0	54.03	24.83	9.30

動物体を映像のほぼ中央に位置させることができたが、右 (左) に移動しているときに動物体の見かけの位置を映像の左 (右) に寄せると、何に向かって移動しているのかが理解しやすくなる。動物体の状態に合わせて映像上での位置を変化させるようにすべきであろう。

7 おわりに

本稿では、パン ・チルト角速度、ズーム値を制御し、動物体を追跡撮影する手法について提案した。提案手法によって、映像上での動物体の見かけの位置の変化が小さく、動物体の見かけの大きさが大きい映像を追跡撮影することができることを示した。

今後の発展としては、本稿のようにカメラの角速度を制御した場合と、カメラの角度を制御した場合との比較や、動物体の速度推定 ・位置予測の精度をあげるためにカルマンフィルタ等を用いること、カメラの角速度に応じてズーム値を変化させることが考えられる。

参考文献

[1] K.Ishizuka, Y.Kameda, M.Minoh, “A Study for Distance Learning Service - TIDE Project -” IEEE International Conference on Multimedia and Expo, Vol.3, pp.1237-1240, 2000.

[2] 和田俊和，浮田宗伯，松山隆司 :視点固定型パン ・チルト ・ズームカメラとその応用，電子情報通信学会論文誌，Vol. J81-D-II, No.6, pp. 1182-1193, 1998