

カメラ間のシャッタータイミングのずれを利用した高速 3 次元位置推定

清水 彰一[†] 藤吉 弘亘[†]

[†] 中部大学 〒 487-8501 愛知県春日井市松本町 1200

E-mail: †shiyou@vision.cs.chubu.ac.jp, hf@cs.chubu.ac.jp

あらまし 複数のカメラを用いたマルチベースラインステレオは、多様な分野で用いられている。高速に移動している対象物の 3 次元位置をステレオ視により推定する場合、異なる視点の各カメラ画像を同期して取得する必要がある。本稿では、3 台の非同期カメラによる曲面を用いた高速 3 次元位置推定について提案する。3 台のカメラのシャッタータイミングが異なるため、最新フレームの 3 次元位置は前 2 フレームの 3 次元位置を用いて線形予測により求める。さらに、最新フレームの画像中の点からの光線情報により 3 次元位置を修正する。また、前 2 フレームの 3 次元位置を求める際、画像上の運動軌跡から求められる光線群から 3 次元空間における曲面を作成し、その曲面と対象画像から得られる光線が交わる位置を 3 次元位置とする。本提案手法では、各カメラのシャッタータイミング毎に 3 次元位置を出力することにより、90 fps 周期で最新フレームにおける対象物の 3 次元位置を得ることが可能である。これにより、ロボット制御におけるビジュアルフィードバックの高速化が期待できる。

Fast 3D Position Measurement Using Cameras Shutter Timing Adjusted

Shoichi SHIMIZU[†] and Hironobu FUJIYOSHI[†]

[†] Chubu University Matsumoto-cho 1200, Kasugai, Aichi 487-8501 Japan

E-mail: †shiyou@vision.cs.chubu.ac.jp, hf@cs.chubu.ac.jp

Abstract A multiple baseline stereo vision system has been proposed and used in various fields. They require cameras to synchronize with each other for tracking objects accurately measuring depth. In this paper, we present a method for calculating 3D positions of objects using three unsynchronized cameras. The 3D position at the latest frame is estimated by linear prediction using previous two frames. Then, the 3D position is calculated once more using constraint by viewing ray information in 3D space obtained from the last image. In calculation of 3D positions at previous two frames for linear prediction, a curved surface passing through the viewing rays of previous three frames is generated, and the 3D position is calculated as an intersection point between the ray and the curved surface. Since the 3D position is calculated at every shutter timing of each cameras, the 3D position can be obtained at 90 fps. This can be useful for controlling a robot as a fast visual feedback system.

1. はじめに

対象物の 3 次元位置を求める手法として、複数のカメラを用いたマルチベースラインステレオが提案されており、多様な分野で用いられている [1]。一般にステレオ視の際、運動している対象物の 3 次元位置を正確に求めるためには、各カメラの画像を同時に取得する必要がある。一方、複数の非同期カメラを用いた対象物の 3 次元位置や距離画像を求める手法として、森らによる人物追跡法 [3] や Zhou らによる距離画像の生成法 [4] 等が提案されている。複数の非同期カメラを用いた人物追跡法は、複数台のカメラから得られる画像内の人間と床の接触を検出し、床面上の位置を出力する。人の移動を対象としているため、人

物追跡システムのビジュアルフィードバックのサイクルは平均 5 fps と遅く、速度の早い物体運動を正確に復元することは難しい。2 台の非同期カメラを用いた距離画像の生成法では、存在しない画像を非同期カメラによって発生するシャッタータイミングの時間ずれとオプティカルフローを用いて自動生成し、生成した画像と実画像の対応より、距離画像を取得する。時間ずれは、基準画像上に各カメラ 2 点ずつ、過去 4 点の特徴点を求め、その比率により得られる。基準画像に存在しない特徴点の基準画像への投影には、エピポーラ直線と基準画像上の 2 特徴点により構成される直線の交点を求めることにより行う。Zhou らの手法の出力は、時刻 $t-1$ (1 フレーム前) の出力であり、最新フレームでの出力ではないという問題がある。

我々は、複数のカメラのシャッタータイミングのずれを用いて、最新フレームにおける3次元位置を高速に求める手法を提案した[2]。最新フレームにおける3次元位置を求めるために、前2フレームの3次元位置を再計算し、最新フレームの3次元位置を線形予測により求める。線形予測の際に用いる前2フレームの3次元位置はステレオ視により求める。しかし、各カメラは同期していないため、対応点が存在しない。そこで、前後フレームの画像座標から、内挿により擬似的に存在しない画像点を求める。この擬似的に求めた点と実際に検出した画像中の対象点でステレオ視を用いて3次元位置を求める。しかし、この内挿する点は線形補間で求めているため、対象物が円運動している場合などに誤差が大きくなる。また、3次元空間で等速運動するものを観測しても、画像面上では等速運動とはならないため、3次元空間で補間の方が有利であると考えられる。

そこで本稿では、文献[2]の手法における過去の3次元位置を求める際に、画像上の運動軌跡から求められる光線群から3次元空間における曲面を生成し、その曲面と対象画像から得られる光線が交わる位置を3次元位置として出力する手法を提案する。また、提案手法をシミュレーションと実カメラを用いた実験によりその有効性を評価する。

2. シャッタータイミングのずれを利用した3次元位置推定

本稿では、3台のカメラを用いた高速3次元位置推定について検討する。3台の各カメラ間のシャッタータイミングを1/90秒ずつずらし、この時間ずれを考慮して各シャッタータイミング毎に3次元位置推定を行う。以下に3台のカメラを用いた3次元位置推定の手法を示す。

Step1 前2フレームの3次元位置を計算

Step2 線形予測により最新フレームの3次元位置を計算

Step3 最新フレームの画像面上の点を通る光線情報による3次元位置の修正

各カメラのシャッタータイミング毎に、上記のStep1～3により3次元位置を求める。従って、本手法は1秒間に90ポイントの3次元位置を出力することが可能となる。以下に各Stepにおける処理について示す。

2.1 前2フレームの3次元位置計算

図1に前2フレームの3次元位置計算について示す。各カメラ画像より対象物の2次元画像座標 (u, v) を求める。カメラ1によって撮影された物体の画像座標を (u_t^1, v_t^1) 、カメラ2の画像座標を (u_t^2, v_t^2) 、カメラ3の画像座標を (u_t^3, v_t^3) とする。画像座標は、キャプチャ終了時の時刻を基に時系列の順に並べられる。本来ならば、ステレオ視の際、同時刻 t の対応点を求める必要があるが、提案手法における3台のカメラ間は $\delta(\delta = \delta_{12} = \delta_{23} = \delta_{31} = 1/90 \text{ 秒})$ の時間ずれを持つ。そのため、図1に示すように時刻 $t-1$ におけるカメラ3の画像座標 (u_{t-1}^3, v_{t-1}^3) に対応するカメラ1の画像座標は存在しない。そこで、以下に示すアルゴリズムを用いて対応する時刻におけるカメラ1上の画像座標 (u_{t-1}^1, v_{t-1}^1) を求め、3次元位置 P_{t-1} をステレオ視により計算する。

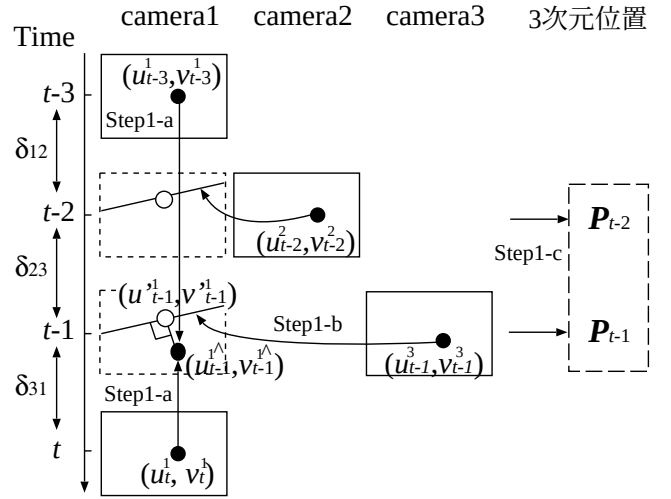


図1 前2フレームの3次元位置計算

Step1-a カメラ1において、前後フレームにおける実際に観測された2点の画像座標より $t-1$ における画像座標 (u_{t-1}^1, v_{t-1}^1) を次式により推定する。

$$\begin{aligned} \hat{u}_{t-1}^1 &= \frac{(\delta_{12} + \delta_{23})u_t^1 + \delta_{31}u_{t-3}^1}{\delta_{12} + \delta_{23} + \delta_{31}}, \\ \hat{v}_{t-1}^1 &= \frac{(\delta_{12} + \delta_{23})v_t^1 + \delta_{31}v_{t-3}^1}{\delta_{12} + \delta_{23} + \delta_{31}} \end{aligned} \quad (1)$$

Step1-b 時刻 $t-1$ におけるカメラ3の観測された画像座標点 (u_{t-1}^3, v_{t-1}^3) から対応するカメラ1上のエピポーラ線 e^3 を求める。Step1-aで計算した $(\hat{u}_{t-1}^1, \hat{v}_{t-1}^1)$ が最も近いエピポーラ線上の点 (u_{t-1}^1, v_{t-1}^1) を $t-1$ におけるカメラ1の画像座標とする。

Step1-c カメラ1上の推定した画像座標 (u_{t-1}^1, v_{t-1}^1) から世界座標空間の直線 l_{t-1}^1 とカメラ3上の観測された (u_{t-1}^3, v_{t-1}^3) より直線 l_{t-1}^3 を求め、その交点を $t-1$ における対象の3次元位置 P_{t-1} とする。

以上の処理を $t-2$ においても行い、前2フレームの3次元位置を求める。

2.2 線形予測による最新フレームの3次元位置推定

既に計算された前2フレームの結果 P_{t-1} と P_{t-2} を用いて、図2に示すように、最新フレーム t における予測位置 $\hat{P}_t = [x_w, y_w, z_w]^T$ を次式より求める。

$$\hat{P}_t = P_{t-1} + v_{t-1}\delta, \quad v_{t-1} = \frac{P_{t-1} - P_{t-2}}{\delta} \quad (2)$$

式(2)は、前2フレームより速度を線形に求めたものである。このとき、最新フレームはカメラ1の結果とする。

2.3 光線情報による3次元位置の修正

最新フレーム t におけるカメラの画像座標から求める3次元空間の直線を求める。最新フレーム t における結果がカメラ1の場合、世界座標における原点 O からカメラ1への平行移動ベクトルを $T^1 = [Tx, Ty, Tz]^T$ 、物体の画像座標 (u_t^1, v_t^1) から求められた世界座標における直線 l_t^1 の傾きを表すベクトルを $r_t^1 = [x_v, y_v, z_v]^T$ とすると、図2に示す光線は次式に示す直線 l_t^1 と表現できる。 k は実数である。

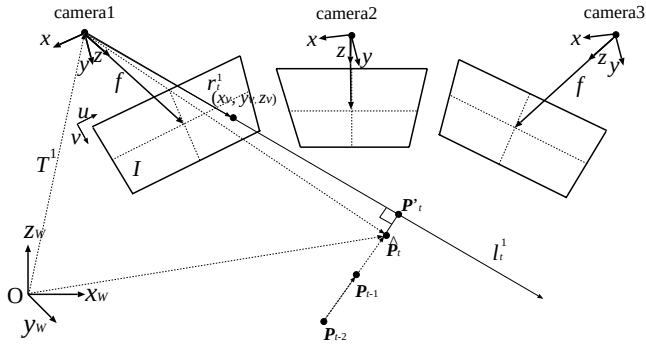


図2 光線情報による3次元位置推定

$$l_t^1 = k r_t^1 + T^1 \quad (3)$$

本来ならば、式(2)により求めた予測位置 $\hat{P}_t$ はこの直線 l_t^1 上に存在する。しかし、予測位置 $\hat{P}_t$ は過去に計算した位置から推定したものであり予測誤差を含んでいる。そのため、予測位置 $\hat{P}_t$ は図2に示すように必ずしも最新フレームの結果より得られる直線 l_t^1 上に存在するとは限らない。

そこで、予測位置 $\hat{P}_t$ が垂直に交わる直線上の点 P'_t を求める。 P'_t は予測位置 $\hat{P}_t$ の直線 l_t^1 の傾きを表すベクトル r_t^1 方向への正射影ベクトルであり、次式で求めることができる。

$$P'_t = \frac{(\hat{P}_t - T^1) \cdot r_t^1}{|r_t^1|^2} r_t^1 + T^1 \quad (4)$$

これにより、線形予測で求めた $\hat{P}_t$ に最も近いカメラ1上の画像座標点 (u_t^1, v_t^1) を通る直線上の点 P'_t を対象の最新フレーム t による3次元位置とする。カメラ2、カメラ3が最新フレームの場合も画像座標 (u_t^2, v_t^2) 、 (u_t^3, v_t^3) から各直線 l_t^2 、 l_t^3 を求め、上記の処理を同様に行う。

3. 曲面による3次元位置推定

2.章で述べた手法では、前2フレームの3次元位置を求める際に、カメラが非同期のため同時刻の対応する画像点が存在しないため、画像平面上の線形補間による内挿により仮想の対応点を求め、ステレオ視により3次元位置を計算している。しかし、内挿により求めた画像平面上の点は線形補間の結果であり、円運動する物体などの際には誤差が生じる。また、画像平面上の内挿は、3次元空間中の運動を内挿することにはならない。そこで、本稿では前2フレームの3次元位置を求める際に、画像平面上の運動軌跡から求められる光線群から3次元空間における曲面を作成し、その曲面と求めたい時刻に得られたカメラ画像から求められる対象物体の光線が交わる位置を3次元位置として出力する手法を提案する。

3.1 光線群から曲面の作成

カメラのシャッタータイミングを1/90秒ずらしているため、図3のように各カメラ画像は交互に取得される。ここで、最新フレームをカメラ3とする。各時刻 t 、 $t-3$ 、 $t-6$ におけるカメラ3上の対象物の画像座標 (u_t^3, v_t^3) 、 (u_{t-3}^3, v_{t-3}^3) 、 (u_{t-6}^3, v_{t-6}^3) より、3次元空間中の直線 l_t^3 、 l_{t-3}^3 、 l_{t-6}^3 を式(3)により求める。次に、求めた直線のパラメータである k に任意の値を代入

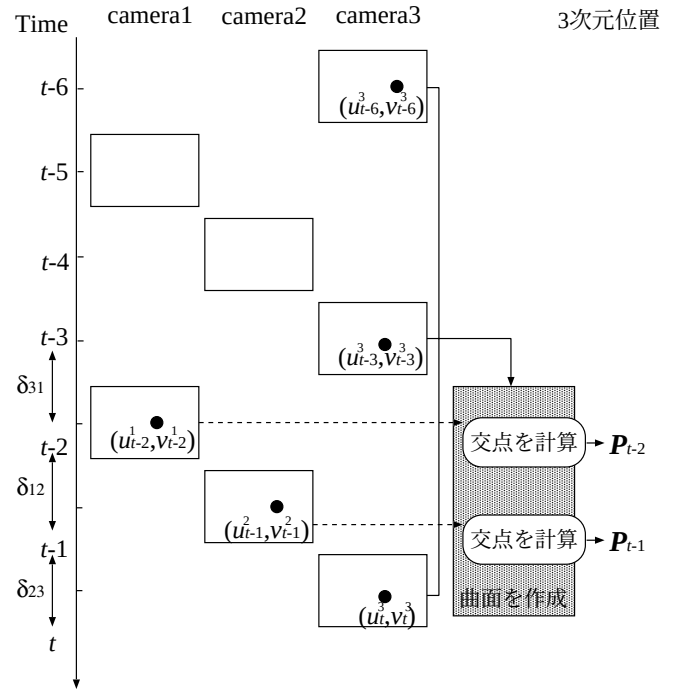


図3 前2フレームの3次元位置計算

することで3次元空間中の点を求めることができる。本手法では、対象空間をカバーするように各直線に対して5つ k の値を代入し、直線につき5点ずつ計15点のデータを得る。求めた15点のデータを入力値としB-スプライン[5]を用いて曲面を求める(図4参照)。B-スプラインの空間座標 (x, y, z) はそれぞれ媒介変数 (s, t) を用いて以下の式で表す。

$$x = x(s, t), \quad y = y(s, t), \quad z = z(s, t) \quad (5)$$

また、これらの座標は、 s と t をそれぞれ変数とするB-スプラインの積によって求めることが可能である。

$$\begin{aligned} x(s, t) &= \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \alpha B_{i,K}(s) B_{j,L}(t), \\ y(s, t) &= \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \beta B_{i,K}(s) B_{j,L}(t), \\ z(s, t) &= \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \gamma B_{i,K}(s) B_{j,L}(t) \end{aligned} \quad (6)$$

式(6)により光線上の15点から、図4(a)に示すように曲面を補間することが可能となる。このとき、媒介変数 (s, t) を任意に設定することで、対応する曲面上の点 (x, y, z) を得ることができる。

3.2 直線と曲面の交点の探索

3次元位置 P_{t-1} は、時刻 $t-1$ のカメラ2の画像座標 (u_{t-1}^2, v_{t-1}^2) から得られる直線が曲面と交わる点 P_{t-1} として求める(図4参照)。しかし、曲面はB-スプラインにより求められているため、直接交点を求めることは容易ではない。そこで本手法では、曲面を細分化し平面の集合体として処理することにより、交点を求める。以下にその手順を示す。

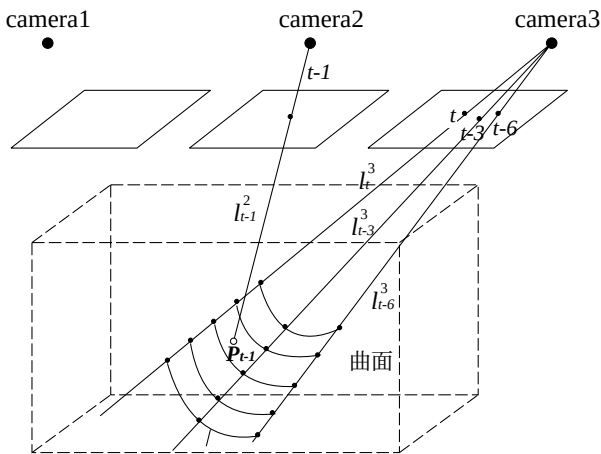


図4 曲面の作成

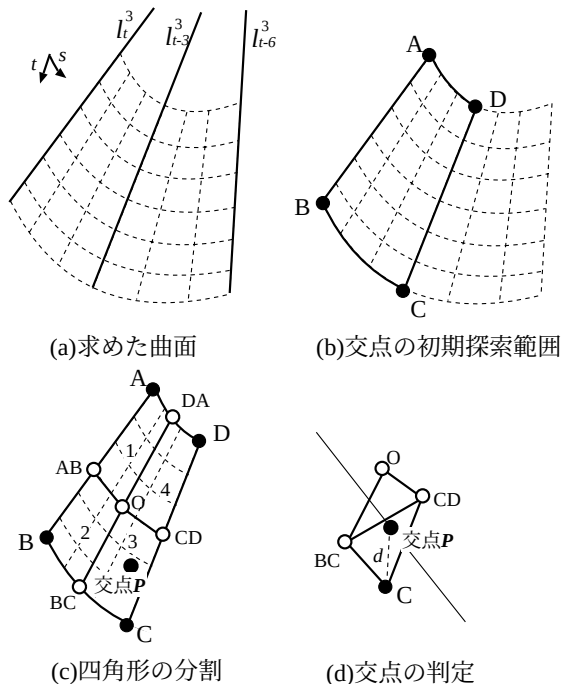


図5 直線と曲面の交点計算

(1) 対象物の運動軌跡である光線群から求めた曲面において、時刻 $t-1$, $t-2$ における対象物体は、時刻 t から $t-3$ までの範囲内に存在するため図5に示す頂点 A, B, C, D とした曲面領域を初期探索範囲とする。

(2) 曲面上の頂点 A, B, C, D により囲まれた領域に対して、媒介変数 s, t を制御して、曲面上の中点 (AB, BC, CD, DA) を求め、4 分割する (図5(c) 参照)。

(3) 4 分割された各平面において、平面区間内にカメラ 2 から得られる光線 l_{t-1}^2 との交点 P が存在するか判定を行う。このとき、各平面の頂点から 3 頂点を選び、図5(d)に示すように 2 つの三角形が存在する平面を求める。その平面とカメラからの直線が交差する点を求め、その点が三角形内に存在するか内外判定を行う (図5(d) 参照)。

(4) 内外判定の結果、内側と判定された平面領域を初期領域として、(2), (3) を繰り返す。最終的に、分割された平面の

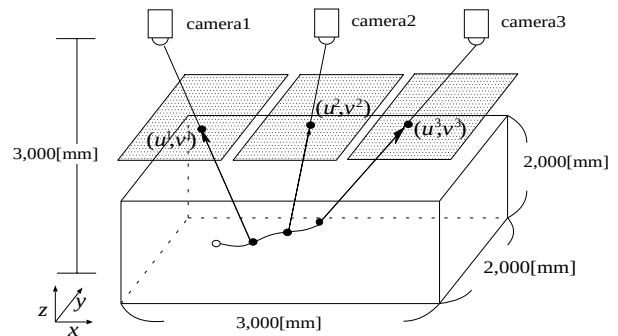


図6 仮想カメラ配置

頂点と交点 P との距離 d がしきい値以内であれば、繰り返し処理を終了する。

予め任意の距離しきい値を設定することで、必要となる精度に対応した 3 次元位置を求めることが可能である、

3.3 直線と平面のなす角

画像中の対象物がカメラの視点間を結ぶように直線運動をしている場合、作成した曲面と光線が平行となり、交点を求めることが不可能となる場合がある。そのため、最新フレームと時刻 $t-3$ において求められた光線から平面を作成し、その平面と前 2 フレームの光線と交わる角度を求め、選択処理を行う。求められた角度が平行に近い場合は、2.1 節で述べた前後フレームを用いた画像平面上の内挿による手法を用いて前 2 フレームの 3 次元位置を求める。

4. シミュレーション実験

本提案手法による 3 次元空間中の対象物の運動を復元するシミュレーション実験により、評価を行う。

4.1 対象物の運動復元

対象物が仮想世界座標空間 (3,000 × 2,000 × 2,000 mm) を移動していると仮定し、その運動復元を行う。3 台のカメラは、高さ 3,000 mm に平行となるように設置してあるとする (図6 参照)。このとき、3 次元空間内の対象物の運動として、以下に示す 3 種類の等速・非等速運動を対象とする。

- 等速運動 (直線): $(x, y, z) = (3,000, 1,200, 0)$ から $(x, y, z) = (0, 1,200, 2,000)$ に向けて直線上を速度 3,000 mm/sec で移動
- 等速運動 (螺旋): $(x, y) = (1,000, 1,000)$ を中心に半径 620 mm, 角速度 4.7 rad/s で螺旋上を移動
- 非等速運動: 高さ 2,000 mm の位置からボールを落下させたときの放物運動 (重力加速度 $g=9.8 \text{ m/s}^2$)

これらの対象物の運動軌跡を、次に推定するカメラ間の時間ずれ毎に、各カメラの仮想画像平面へ投影する。シミュレーション実験では、各カメラの仮想画像平面上の点 (u, v) を用いて、提案手法により 3 次元位置を推定する。また、本実験において画像平面上での観測誤差はない。

4.2 実験結果

図7(a), (b), (c) に各運動を復元した結果を示す。また、表1に前 2 フレームの 3 次元位置計算結果と真値との平均誤差を

表 1 前 2 フレームの 3 次元位置と真値との平均誤差 [mm]

方法	等速運動		非等速運動
	直線	螺旋	
画像平面	0.5	1.0	1.18(0.6)
曲面	0.0	0.2	1.16(0.2)

表 2 3 次元位置の推定値と真値との平均誤差 [mm]

方法		等速運動		非等速運動
		直線	螺旋	
ステレオ視		23.2	21.4	19.6
提案手法	画像平面	1.1	2.0	2.2(1.1)
	曲面	0.0	0.3	2.4(0.5)

示す。表 1 における画像平面は、前 2 フレームの 3 次元位置を内挿を用いることにより求めた結果を示す。曲面は、本稿にて提案した曲面を作成し光線と交わる位置を 3 次元位置として求める手法である。非等速運動の括弧内の値は、対象物体が床に衝突した時に生じる特異点を考慮しない場合の結果である。表 1 より、本手法で求めた過去の 3 次元位置精度は向上していることが分かる。特に、螺旋運動において高い効果が得られている。

表 2 は表 1 で求めた前 2 フレームの 3 次元位置を用いて最新フレームにおける 3 次元位置を文献 [2] の手法で推定した際の誤差である。ステレオ視は時間ずれを含んだ対応点により 3 次元位置を推定した結果である。本手法の最新フレームにおける 3 次元位置精度は、時間ずれを含むステレオ視の結果より精度が良いことがわかる。最新フレームの 3 次元位置は、前 2 フレームの 3 次元位置の線形予測結果を用いるため、提案した曲面による手法が最新フレームの 3 次元位置においても良い精度を得ていることがわかる。

5. 実 験

本提案手法を、1 台の PC と 3 台のカメラを用いて実装し、等速円運動する対象物の 3 次元位置を求める実験を行う。

5.1 本システムの概略

3 台のカメラを用いたビジョンシステムの概略を図 8 に示す。カメラは高さ 2,800 mm の位置に床面上の 2,000 × 3,000 mm の領域が視野に入るように設置した。それぞれのカメラは世界座標 (x_w, y_w, z_w) と画像座標 (u, v) により、キャリブレーション済みである。カメラからの映像信号は、1 台の PC にインストールされた 3 枚のフレームグラブに入力され、それぞれの画像が取り込まれる。また、TV 信号発生装置を用いて、各カメラに NTSC 信号を 1/90 秒ずらして入力する。

それぞれのカメラからの画像は、独立したプロセス process-1, process-2, process-3 で 1/90 秒毎に処理される。各プロセスは、画像処理により得られたカメラ i における対象物の画像座標 (u^i, v^i) と、その処理した画像のキャプチャ終了時刻を UDP 通信により 3 次元位置を計算する process-4 に送信する。

実際にボールを投げたとき 1.5 秒間の復元結果を図 9 に示す。図 9 では、135 ポイントのボールの 3 次元位置を出力しており、90 fps のカメラと同等の出力であることが分かる。本手法の位

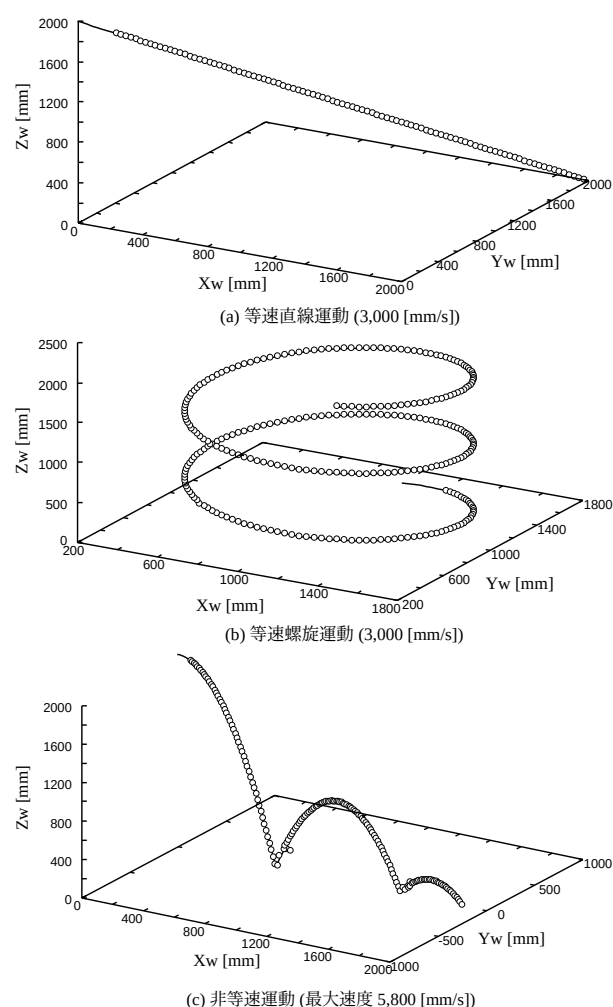


図 7 運動の復元結果

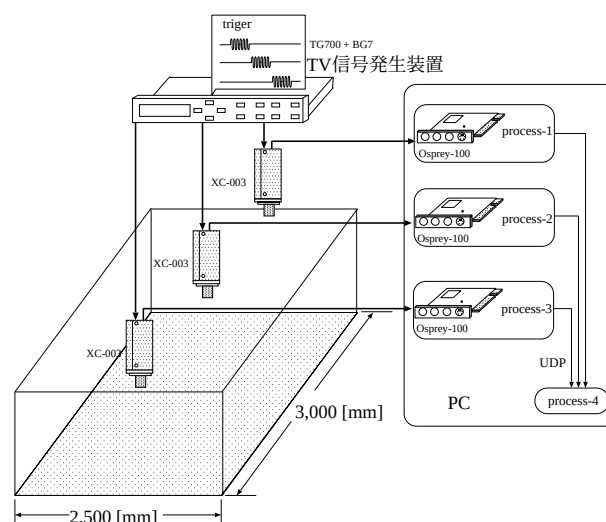


図 8 ビジョンシステムの概略

置推定精度を定量的に評価するために、ボールとターンテーブルによる等速円運動の 3 次元復元による評価を行う。以下にその実験概要と結果を示す。

5.2 実験概要

等速円運動として、図 10 に示すようにターンテーブルとゴルフボールを用いる。長さ 1,000 mm の物差しにボールを

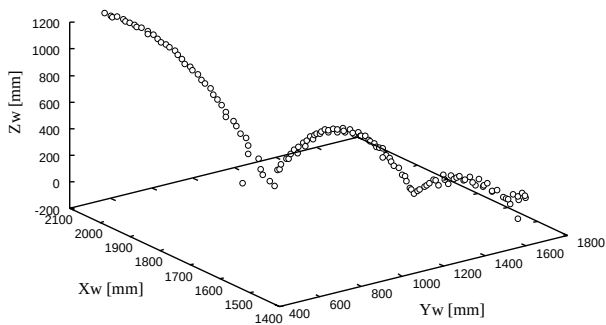


図 9 ボールの運動復元結果

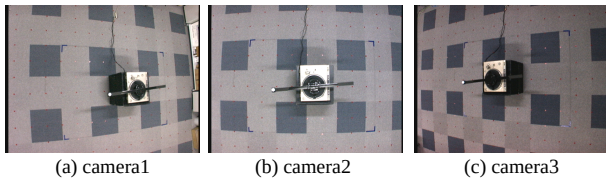


図 10 実験の様子

表 3 半径の平均と分散

方法	平均 [mm]	分散
ステレオ視	504.6	608.9
提案手法	画像平面	512.8
	曲面	509.2

表 4 z 軸の平均と分散

方法	平均 [mm]	分散
ステレオ視	657.8	10502.5
提案手法	画像平面	665.2
	曲面	660.0

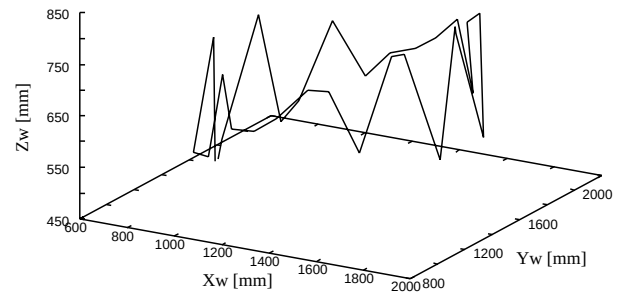
固定し、ターンテーブルを回転させることで半径 500 mm の等速円運動となる。ターンテーブルは高さ 500 mm の箱に置かれており、ボールの床からの高さは 660 mm となる。

5.3 実験結果

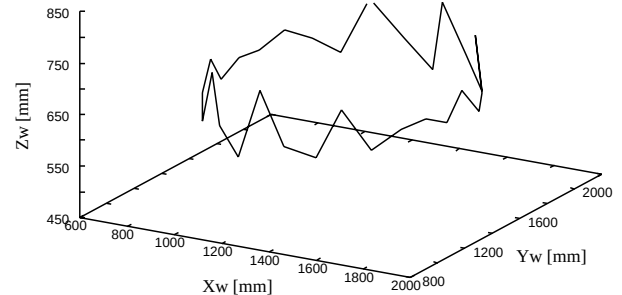
高さ 660[mm] の位置で床面と平行に半径 500[mm] の等速円運動するボールの最新フレームにおける 3 次元位置を前 2 フレームを求める。前 2 フレームの 3 次元位置を各手法で求めたものを用いて最新フレームの 3 次元位置を推定した結果を図 11 に示す。推定された 3 次元位置を $x_w - y_w$ 平面へ投影した際に求まる円の半径 r の平均と分散を表 3 に示す。ボールの円運動における実際の半径は 500[mm] であり、各手法において平均誤差に有意な差はない。しかし、提案手法(曲面)の分散値は小さく、各フレームにおいて推定した位置精度にばらつきが小さいことがわかる。次に、 z_w 軸方向における 3 手法の平均と分散を表 4 に示す。各手法全て、 z_w 軸の平均(ボールの高さ)は、実際の高さ 660[mm] から 5 [mm] 以内であった。半径の場合と同様に、曲面の分散値は最も小さく、実機を用いた実験においても提案手法の精度が良いことがわかる。

6. ま と め

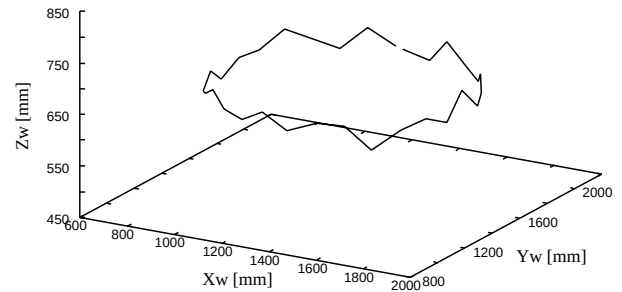
本稿では、前 2 フレームの 3 次元位置は曲面を作成し、時刻 $t-1$ と $t-2$ の画像中の点からの光線と交わる位置を 3 次元位



(a) ステレオ視



(b) 画像平面



(c) 曲面

図 11 3 次元位置推定結果

置として求める手法を提案した。シミュレーション実験により、曲面による推定で前 2 フレームの 3 次元位置がより精度良く求めることが可能であることを示した。また、実機を使用した実験においてもシミュレーションと同様の結果が得られ、本手法の有効性を確認した。

謝辞

本研究は、栢森情報科学振興財団の助成を受けて遂行されています。

文 献

- [1] M. Okutomi, and T. Kanade, "A Multiple Baseline Stereo", Proc. IEEE Trans. PAMI, Vol. 15, No. 4, pp. 353-363, 1993.
- [2] S. Shimizu, H. Fujiyoshi, Y. Nagasaka and T. Takahashi, "A Pseudo Stereo Vision Method for Unsynchronized Cameras", Proc. Asian Conference on Computer Vision, vol.1, pp.575-580, 2004.
- [3] 森 大樹, 内海 章, 大谷 淳, 谷内田 雅彦, 中津 良平, "非同期多視点画像による人物追跡システムの構築". 信学論, J84-D-II, 1, pp. 102-110, 2001.
- [4] C. Zhou and H. Tao, "Dynamic Depth Recovery from Unsynchronized Video Streams". IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.351-358, 2003.
- [5] de Boor, C., "A Practical Guide to Splines", Springer-Verlag, New York, 1978.