3台のカメラを用いた擬似ステレオによる3次元位置推定

Pseudo Stereo Method for Measuring 3D Position Using Three Cameras

清水彰一† 藤吉弘亘† 長坂保典† 高橋友一††

Shoichi Shimizu, Hironobu Fujiyoshi, Yasunori Nagasaka, and Tomoichi Takahashi †中部大学、 †† 名城大学

Chubu University, Meijo University

shiyou@vision.cs.chubu.ac.jp

Abstract

Multiple cameras have been used to get a view of a large area. In some cases, the cameras are layouted so that their views are overlapped to get a better view as a whole than non_overlapping layout. 3D information of the overlapping areas that are covered with two or three cameras can be get by stereo vision methods. Shifting shutter timings of cameras and using our pseudo stereo vision method, we can output 3D information faster than 30 fps.

In this paper, we propose a pseudo stereo vision method using cameras with different shutter timings. Using three cameras, two types of shutter timings are discussed. In three different shutter timings, 90 times of 3D position for a sec are obtained because proposed method can output 3D positions at every shutter timings of three cameras. In two different shutter timings, it is possible to calculate the 3D position at 60 fps with better accuracy.

1 はじめに

ロボットが行動するためには,センサによる環境認識・理 解技術から得られる情報を基に行動プランニングを行い, ロボットを制御するフィードバック機構が必要である.視 覚センサでロボットを迅速にかつ正確に制御するために は,ビジュアルフィードバックのサイクルを高速に行う必 要があり,特に,カメラから得られる画像から有用な情報 を得るための高速なビジョンシステムが不可欠となる.そ の際には,ロボット等の対象とする物体の3次元位置を より正確に検出する必要がある.ロボカップにおいても, ループシュートを打つサッカーロボット [Muratec, 2003] が登場しており,その際には,ボールの軌跡をサッカー フィールド上の2次元位置ではなく,3次元位置を高速に 求める必要がある.

ビジョンシステムの高速化として,ロボカップ小型リーグ では,60 fpsの高速カメラが用いられている [R.D'Andrea, 1999,加藤,2002].これは,60 fpsのカメラを用いたダブ ルバッファリングによる高速処理を行うものである.し かし,このシステムでは,1台のカメラを使用しているた め,対象物の3次元位置を求めることができない.

対象物の3次元位置を推定する手法としてマルチベー スラインステレオが既に提案されており,多様な分野で 用いられている[Okutomi, 1993]. 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距離画像を求める手 法として, Zhou らによる距離画像の生成法[Zhou, 2003] が提案されている.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,2台の非同期カメラを用いた3次元位置推定法 を提案している[Shimizu,2004].2台のカメラ間のシャッ タータイミングのずれを利用し,60 fps で対象物の3次 元位置を得ることが可能である.本報告では,3台のカメ ラを用いた際に考えられる2種類のシャッタータイミング における3次元位置推定法について提案する.



Figure 1: 同期・非同期のシャッタータイミング





2章では、2種類のシャッタータイミングとその手法に ついて、3章では、シミュレーションによる移動物体の運 動復元について実験を行い、提案手法の有効性を示す.4 章では、実際の実験において、高速に円運動する対象物の 3次元位置推定の実験を行い、提案手法の有効性を示す.

2 複数のカメラを用いた3次元位置推定

ステレオ視では対象物の 3 次元位置を求めるために, 2 枚 の画像を同時刻で取得する必要がある.一般的なステレ オ視では,通常のカメラ (30 fps)を用いた場合, 3 次元位 置の出力は最大 30 fps となる.2 台のカメラによるステ レオ視におけるカメラのシャッタータイミングを図 1(a) に示す.カメラ間が同期していない場合,各カメラ間に 時間ずれが生じる(図 1(b)).最新画像がカメラ 2 の場合, (u_t^1, v_t^1) と時刻 $t' = t + \delta$ の対応点 $(u_{t'}^2, v_{t'}^2)$ (時間ずれを 含んだ対応点)を用いてステレオ視により高速に移動する 対象物の 3 次元位置を求めると,図 2 に示すように誤差 が生じる.

これに対し, 我々は, 非同期カメラを用いた3次元位 置推定法を提案している [Shimizu, 2004].シャッタータ イミング毎に3次元位置を求めることにより, 早いサイ クルでの3次元位置の出力を得ることができる.以下に, 3台のカメラを用いた2種類のシャッタータイミングにお ける3次元位置推定法について示す.



2.1 カメラ3台におけるシャッタータイミング

3台のカメラを用いた場合,考えられる3種類のシャッター タイミングの組み合わせを図3に示す.図3(a)は,マル チベースラインステレオシステムにおけるシャッタータイ ミングである.本報告では図 3(b) を typeA, (c) を typeB とし, この2種類のシャッタータイミングにおける3次 元位置推定について検討する.typeAの場合,各カメラ 間の時間ずれ △ を 1/90 秒と設定する.この場合,シャッ タータイミング毎に3次元位置を求めるため90 fpsの出 力を得ることが可能となる.一方, typeB では, カメラ1 とカメラ3を同期するように設定し,ステレオ視により 3次元位置を求める.カメラ2が最新フレームの際は,前 2フレームにおいてステレオ視により求めた結果から,線 形予測を用いて3次元位置を求め,光線情報により3次 元位置を修正し出力を得る.また,カメラ2のシャッター タイミングは,カメラ1,3より1/60秒ずれているため, 3次元位置を 60fps で得ることが可能となる.以下に2種 類のシャッタータイミングにおける3次元位置推定法につ いて述べる.

2.2 TYPE-A: 3 つのシャッタータイミング (90 fps) typeA では,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.

2.2.1 前2フレームの3次元位置計算

図4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,*v*^t),



Figure 4: 前 2 フレームの 3 次元位置計算

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 $\delta(\delta = \delta_{12} = \delta_{23} = \delta_{31} = 1/90$ 秒)の時間ずれを持つ.そのため,図 4 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 $(u_{t-1}^{\hat{1}}, v_{t-1}^{\hat{1}})$ を求め,3次元位置 P_{t-1} をステ レオ視により計算する.

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 $(u_{t-1}^{\hat{1}}, v_{t-1}^{\hat{1}})$ を次式により推定する.

$$u_{t-1}^{\hat{1}} = \frac{(\delta_{12} + \delta_{23})u_t^1 + \delta_{31}u_{t-3}^1}{\delta_{12} + \delta_{23} + \delta_{31}},$$

$$v_{t-1}^{\hat{1}} = \frac{(\delta_{12} + \delta_{23})v_t^1 + \delta_{31}v_{t-3}^1}{\delta_{12} + \delta_{23} + \delta_{31}}$$
(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 $(u_{t-1}^{\hat{1}}, v_{t-1}^{\hat{1}})$ が 最も近いエピポーラ線上の点 $(u_{t-1}^{\prime 1}, u_{t-1}^{\prime 1})$ 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 $(u_{t-1}^{\prime 1}, v_{t-1}^{\prime 1})$ 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.

2.2.2 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5に示すように,最新フレームtにおける予測位



Figure 5: 光線情報による 3 次元位置推定

置 $\hat{\boldsymbol{P}}_t = [x_w, y_w, z_w]^T$ を次式より求める.

$$\hat{P}_{t} = P_{t-1} + v_{t-1}\delta, \quad v_{t-1} = \frac{P_{t-1} - P_{t-2}}{\delta}$$
 (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. また,過去の計算結果を用いた位置予測には拡張カルマ ンフィルタを用いたもの[Browning, 2002,金子, 1996]や 3次元曲線が考えられるが,本報告では観察する時間間隔 が1/90秒と短いことから線形に求める.

2.2.3 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まる 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 5 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.

$$\boldsymbol{l}_t^1 = k\boldsymbol{r}_t^1 + \boldsymbol{T}^1 \tag{3}$$

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5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^1_t の傾きを表すベクトル r^1_t 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.

$$\boldsymbol{P'}_{t} = \frac{(\hat{\boldsymbol{P}}_{t} - \boldsymbol{T}^{1}) \cdot \boldsymbol{r}_{t}^{1}}{|\boldsymbol{r}_{t}^{1}|^{2}} \boldsymbol{r}_{t}^{1} + \boldsymbol{T}^{1}$$
(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



Figure 6: 3次元位置推定 (typeB)

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.

2.3 TYPE-B: 2 つのシャッタータイミング (60fps) 線形予測を用いて,より精度の良い3次元位置を求める ためには,前2フレームにおける3次元位置を正確に求 める必要がある.図3(c)の場合,2台のカメラ(カメラ1 とカメラ3)のシャッタータイミングは同期しているため, ステレオ視により3次元位置を求めることができる.図6 に示すようにステレオ視で得られた前2フレームの3次 元位置 $P_{t-1} \ge P_{t-3}$ を用いて,最新フレームの3次元位 置 P'_t を線形予測により推定する.その後,typeA と同様 に,光線情報の制約を用いて3次元位置を修正する.

3次元位置は,2台の同期したカメラによるステレオ視の結果と,1台のカメラから得られる光線情報による修正を加えた結果が交互に出力される.この場合,1秒間に typeA(90ポイント)より少ない60ポイントの出力となる.

3 シミュレーション実験

3次元空間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,提案手法の評価を行う.

3.1 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7参照).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



Figure 7: 仮想カメラ配置

 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 m/s²)

これらの対象物の運動軌跡を,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,本実験おいて画像平面上での観測 誤差はない.

3.2 実験結果

表1に各運動を復元した結果と真値との推定誤差を示す. 表1において, typeAのシャッタータイミングにおける非 同期は,図2に示す時間ずれを含んだ対応点を用いたス テレオ視の結果である.typeAでは,本手法(typeA)の 結果が非同期より良いことは明らかである.

typeBのシャッタータイミングにおける線形予測は,同 期した2台のカメラからステレオ視により前2フレーム の3次元位置を求め,その2点から線形予測した結果で ある.提案手法と線形予測を比較すると,提案手法の精 度が良いことが分かる.これは,線形予測した結果から, 光線情報を用いた修正を行うことにより位置推定精度が 向上しているためである.

シャッタータイミング			fps	等速		非笙词
		直線		螺旋	中守座	
+	typeA	非同期	90	24.6	20.8	14.8
		提案手法	90	1.1	2.0	1.7
t	typeB	線形予測	60	0.2	1.4	4.4
ſ		提案手法	60	0.2	0.5	1.5

Table 1: 3 次元位置の推定値と真値との平均誤差 [mm]

4 複カメラによる評価実験

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により 評価を行う.



Figure 8: ビジョンシステムの概略



Figure 9: 機器構成

4.1 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, キャリプレーション済みである [Tsai, 1987].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 レームグラバに入力され,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 信号を各タイプ毎にシャッタータイミングを発生させ入力 する.本システムの機器構成を図 9 に示す.

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 で各シャッタータイミン グ毎に処理される.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(*uⁱ*,*vⁱ*)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.

4.2 実験

実際にボールを投げたとき 1.5 秒間の復元結果を図 10 に 示す.図10(a)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.図10(b)では,90 ポイントのボールの3次 元位置を出力しており,60 fps のカメラと同等の出力で あることが分かる.

本手法の評価を行うため,図11に示すようにターン テーブルとゴルフボールを用いて実験を行う.長さ1,000



Figure 10: 3次元位置推定結果



Figure 11: 実験の様子

mm の物差しの端にボールを固定し,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.ターンテーブルは 45 rpm のスピードで回転しており,角速度は $(45 \times 2\pi)/60 = 0.478$ rad となる.

4.3 実験結果

typeA と typeB の 2 種類のシャッタータイミングによる 3 次元位置推定結果の z_w 軸方向における平均と分散を表 2 に示す.共に, 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.typeB の分散値 は typeA より低く,シミュレーション結果と同様に typeB の精度が高いことが分かる.

Table 2: z 軸の平均と分散						
	平均 [mm]	分散				
typeA	664.5	2143.7				
typeB	662.3	112.0				

4.4 まとめ

本報告では,2種類のシャッタータイミングに応じた3次 元位置推定法を提案し,その評価を行った.提案手法は, 3台のカメラ全てのシャッタータイミングが1/90秒ずれて いる場合(typeA),各シャッタータイミング毎に3次元位 置を求めることにより,1秒間に90ポイントの出力を得 ることが可能である.一方,同期した2台のカメラと1/60 秒ずれたシャッタータイミングのカメラの場合(typeB), 60 fps で3次元位置を得ることができ,より正確に求め ることが可能である.

ロボカップ小型リーグでは,2004年から,フィールド サイズが大きくなり,1台のカメラではフィールド全体を 映すことは不可能であると考えられる.本手法は,カメラ 間で重なる位置が在れば,より早く3次元位置を求める ことが可能である.

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た.

参考文献

[Muratec, 2003] Muratec FC: http://ww.muratec.net/robot/

- [R.D'Andrea, 1999] R. D'Andrea, et al: Detailed vision documentation, http://robocup.mae.cornell.edu/
- [加藤, 2002] 加藤恭佑, 日比野晋也, 児玉幸司, 村上和人, 成瀬正: ロボカップのための高速な小型ロボット検出 法について, 情報処理学会研究報告, Vol.CVIM-136, No.16, pp.115-122, 2003.
- [Okutomi, 1993] M. Okutomi, and T. Kanade; A Multiple Baseline Stereo, In Proc. IEEE Trans, PAMI, Vol. 15, No. 4, pp. 353-363, 1993.
- [Shimizu, 2004] S. Shimizu, H. Fujiyoshi, Y. Nagasaka and T. Takahashi: A Pseudo Stereo Vision Method for Unsynchronized Cameras, ACCV2004, Vol.1, pp.575-580, 2004.
- [Zhou, 2003] C. Zhou and H. Tao: Dynamic Depth Recovery from Unsynchronized Video Streams, In Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.351-358, 2003.
- [Bruce, 2000] J. Bruce, T. Balch and M. Veloso: Fast and Inexpensive Color Image Segmentation for Interactive Robots, IROS-2000, Vol.3, pp.2061-2066, 2000.

- [Browning, 2002] B. Browning, M. Bowling and M. Veloso: Improbability Filtering for Rejecting False Positives, In Proc. IEEE International Conference on Robotics and Automation, pp.120-200, 2002.
- [金子, 1996] 金子俊一, 堀内一仁,本多庸悟: 多重カルマ ンフィルタによる3次元運動推定,信学論, Vol.J79-D-II, No.5, pp.840-850, 1996.
- [Tsai, 1987] R. Y. Tsai: A versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses, In IEEE Journal of Robotics and Automation, Vol.RA-3, Num.4, pp.323-344, 1987.