

複合現実オーサリング環境のための ウェアラブルセンサと環境側センサ統合の検討

大隈隆史† 蔵田武志† 西田佳史‡ 堀俊夫‡ 坂上勝彦†

†産業技術総合研究所 知能システム研究部門 〒305-8568 つくば市梅園 1-1-1 中央第二

‡産業技術総合研究所 デジタルヒューマン研究ラボ 〒135-0064 東京都江東区青海 2-41-6

E-mail: {takashi-okuma, t.kurata, y.nishida, t.hori, k.sakaue}@aist.go.jp

あらまし 本研究では複合現実環境を観察しながらコンテンツのオーサリングが可能なシステムの構築を目指し、その第一段階としてウェアラブルセンサと環境側センサの両者から得られる情報を統合して、作業対象と入力デバイスのユーザの視点に対する位置と姿勢を推定する手法を検討する。ウェアラブルセンサとしては視点からの映像を撮影する小型カメラ、頭部回転量を測定する回転角センサ、環境側センサとして対象の3次元位置を計測する超音波センサを用いる。筆者らの提案したトップダウン・ボトムアップ統合アプローチによる追跡手法で環境側センサを用いることで、データベース中に登録された実環境情報が少ない段階から対象実物体の追跡を安定化する。

キーワード 複合現実感, トラッキング, オーサリング環境, センサフュージョン

A study of multi-sensor data fusion of wearable/environment-embedded sensors for mixed reality authoring environment

Takashi OKUMA† Takeshi KURATA† Yoshifumi NISHIDA‡

Toshio HORI‡ and Katsuhiko SAKAUE†

† Intelligent systems institute, National Institute of Advanced Industrial Science and Technology (AIST)

Tsukuba Central 2, 1-1-1 Umezono, Tsukuba, Ibaraki, 305-8568, JAPAN

‡ Digital Human Laboratory, National Institute of Advanced Industrial Science and Technology (AIST)

2-41-6, Aomi, Koto-ku, Tokyo, 135-0064, JAPAN

E-mail: {takashi-okuma, t.kurata, y.nishida, t.hori, k.sakaue}@aist.go.jp

Abstract In this paper, a tracking method for a mixed reality authoring environment is described. In this method, positions and orientations of viewpoint, an object and an input device are determined using data set from wearable camera, wearable rotation sensor, and ultrasonic position sensor embedded in the environment. A vision-based tracking method that uses integration of a top-down approach and a bottom-up approach is used with a position sensor to track targets stably while entries of database are insufficient.

Keyword Mixed Reality, tracking, authoring environment, sensor fusion.

1. はじめに

現実環境と仮想環境を融合して提示する複合現実感 (Mixed Reality: MR) 技術は潜在的な応用分野の広さから近年盛んに研究されている技術である[1][2]。例えば機械のメンテナンスをするときに作業対象の機械に作業手順をCGで重畳表示する3次元オンラインマニュアルや、現実の建築現場に設計図から起こした仮想の建物を合成して見せる景観シミュレーションなど、現実環境におけるユーザの作業を支援する実用への

期待が高い。MR技術が研究され始めて以来、現実環境と仮想環境の位置合わせがその根幹となる技術として研究されている。これまでに環境中に配置された人工マーカを画像処理によって認識して位置合わせする手法[3][4][5][6]や、複数センサの結果を統合してより正確な位置合わせを実現する手法[7]が登場している。また、Wearable Computingと組み合わせ、屋外環境下における位置合わせ技術の研究や、Wearableな設定では使いづらい従来型の位置センサ(磁気センサや超音

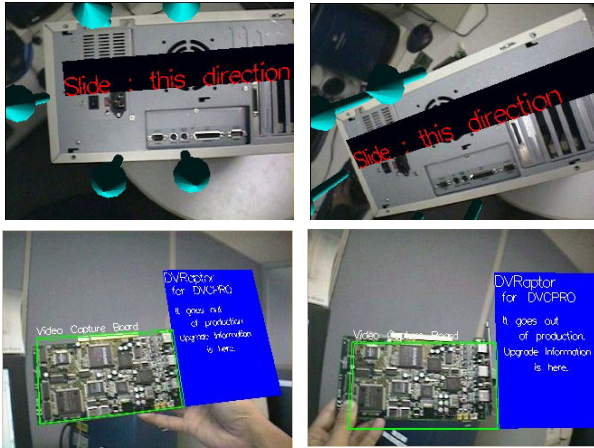


図 1 : 3次元オンラインマニュアル出力例

波センサ)を用いない位置合わせ技術などが研究されている。著者らは特別に環境を調整しない場所でのMR環境を実現するため、観察対象の認識とマーカを用いない位置合わせを実現する研究を進めている[8][9][10][11]。図1はマーカのない実物体を追跡し、物体に対する作業手順が合成提示されている例を示している。

2. MR オーサリング環境

MR環境を実現するためには、位置合わせに用いられる現実環境の情報と、合成される仮想物体形状情報の2種類のデータが必要になる。実際にMR環境を構築するときには、利用する位置合わせ手法に応じた現実環境情報の収集と、仮想物体のモデリングを事前に別々に行い、後にMR環境に提示してテストを行う。このためMR環境コンテンツの設計者は、実際に合成結果を確認しないことには、収集した実環境情報が十分であったか、モデリングした仮想物体が重畳される実環境で意図したように観察されるかを検討することができず、コンテンツ生成に多大な労力を要する。そこで、ユーザが利用するMR環境に提示するコンテンツを作成するために、よりリッチなMR環境を利用できるMR環境設計者が、精度よく実物体情報を収集しつつ、仮想物体の記述を行えるMRオーサリング環境を構築することが重要となる。

本研究の最終的な目標は、MR環境の設計者が合成結果を確認しながら現実環境情報の収集と仮想物体のオーサリングを対話的に行うことを可能にするMRオーサリングシステムに関する技術確立することにある。例えば、コンテンツの設計中に、ある視点からのCG合成がウェアラブルMRに使用できる情報だけで正しく提示できるか確認したり、正しく提示できない場合には必要な実物体情報を追加したりするといった機能を提供することで、MR環境はより実用的なものになると考えられる。

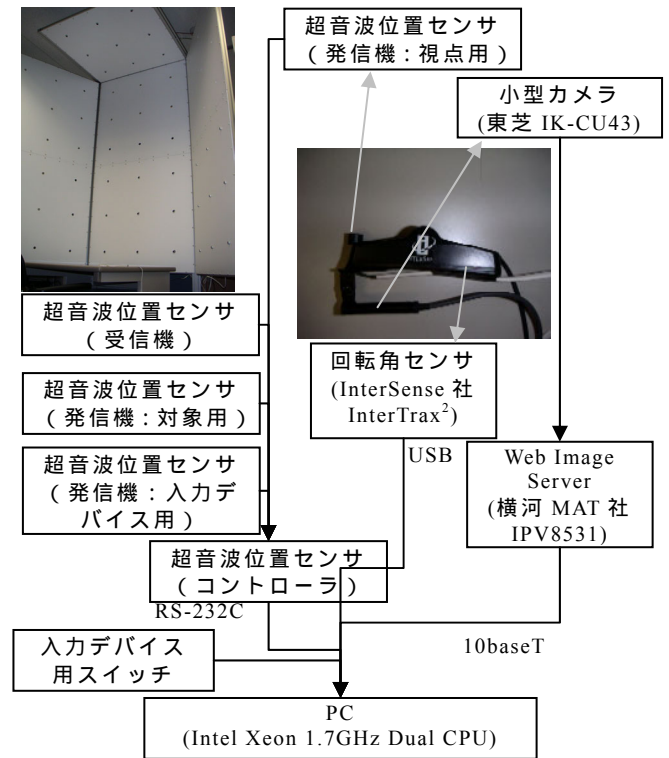


図 2 : システム構成

本研究で開発を目指すオーサリングシステムは、一般ユーザ向けの環境よりも多くのセンサやマーカを利用できるリッチな環境で、MR環境を設計するための機能を提供することを前提とする。また、このオーサリングシステムを用いて設計されるMR環境として、一般ユーザ向けに特殊なセンサや人工的なマーカを使用しないウェアラブルMR環境をターゲットとする。本報告ではその第一報として、一般ユーザが使用すると想定される着用型のセンサとMR環境のデザイナーのみが利用可能な環境側のセンサからのデータを統合してデザイナーの視点位置姿勢を追跡する手法について検討する。

3. 実験システム

3.1. 構成

実験システムの構成を図2に示す。ウェアラブルセンサとして頭部の回転量を計測する回転角センサ、デザイナーの視点からの映像を取得する超小型カメラを用いる。これは前述のように、一般ユーザが利用するMR環境で使用されるセンサを想定している。環境側のセンサとしては、文献[12][13]で開発した低コストの超音波式の位置計測センサを用いる。この超音波センサは環境側に多数の受信機を設置しておき、計測環境内に存在する複数の送信機の位置を計測する。本研究ではデザイナーの視点位置、オーサリングに使用する位置入力デバイスおよびオーサリングの対象となる実物

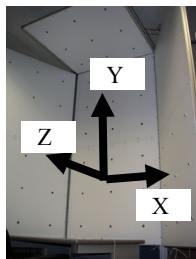


図 3 - 1 : 設定した位置センサ座標系(図中の矢印は方向のみを示す . 原点は床上に設定した .)

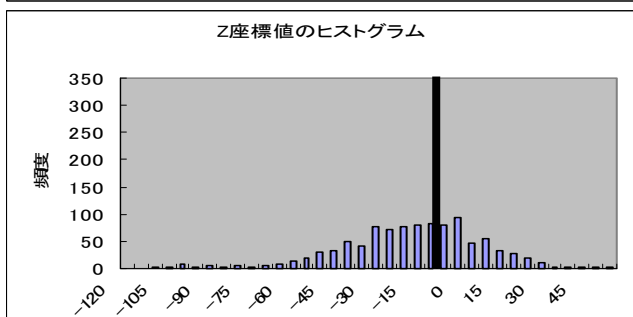
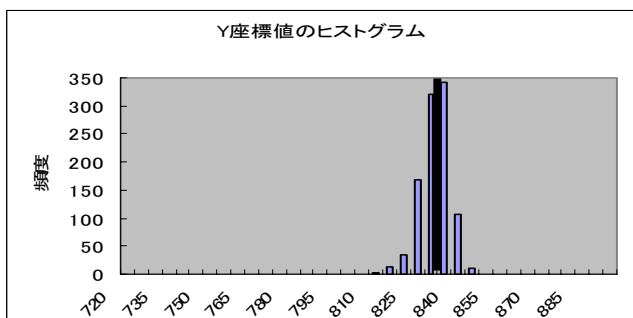
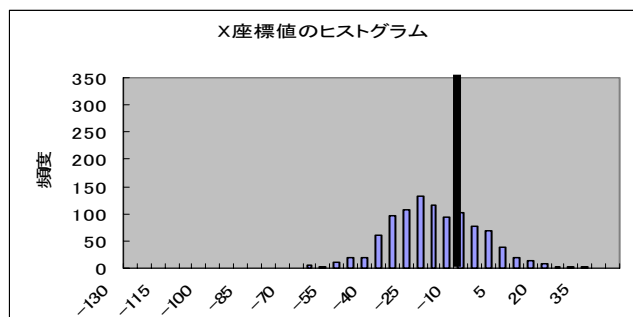


図 3 - 2 : 超音波位置センサの出力分布 (静止時 . 太線は手作業で計測した値を示す .)

体に送信機を取り付けて利用することを想定している . 各センサとの通信および追跡処理アルゴリズムは Intel Xeon 1.7MHz Dual CPU の PC によって実現されている .

3.2. 超音波位置センサ

超音波位置センサを利用することで , 視点 , 追跡対象実物体 , 入力デバイスの各々の位置を知ることができる . このため , ウェアラブルセンサのみを用いた手法に比べて位置姿勢の計算に利用できる情報が増え , 対象の認識処理を簡略化することができる .

図 3 - 1 に示されるような空間で , 27 個の受信機を天井 , 正面及び向かって右側の壁に 30cm 間隔で 9 個

ずつ配置し , 位置計測制度の予備実験を実施した . 図 3 - 2 に予備実験の結果として , 送信機静止時の計測値 (1000 回分) の分布をヒストグラムを用いて示す . X 軸方向に約 $\pm 20\text{mm}$, Y 軸方向 (鉛直方向) に約 $\pm 10\text{mm}$, Z 軸方向に約 $\pm 40\text{mm}$ のブレが観測されている . 文献 [12][13] においては日常生活の記録を目的としてセンサを利用しているので , 移動平均をとることでセンサ出力のブレを平滑化して記録するため , 実用上問題は無い . しかし MR オーサリングシステムの視点位置計測には実時間性と精度が要求される . そこで対象と視点に取り付けられたセンサによって計測された値から計算される相対位置は後のステップで予測値として利用することで計測誤差の影響を押さえる .

3.3. ウェアラブル回転角センサ

実験システムでは頭部に着用する回転角センサとして InterSense 社の InterTrax² を用いた . InterTrax² は静止時と運動時で異なるセンシングアルゴリズムを用いるため , 静止時には安定した計測値を出力する . しかしながら微小な運動を加え続けると計測値のドリフトが観測されるため , 出力値をそのまま利用せずに InterTrax² からの出力のフレーム間の回転差分量のみを用いることでドリフトの影響を避けるように後段の処理を設計した .

4. 相対位置姿勢の推定

図 4 に本実験システムで用いる追跡対象物体の視点に対する相対位置姿勢の追跡アルゴリズムの概略を示す . 図 4 では処理の流れを太矢印 , データの流れを細矢印で示している . この追跡アルゴリズムは基本的に画像処理に基づく手法 [11] であり , その他のセンサ情報は後段の画像処理をより効率よく実行するために利用される . MR オーサリング環境ではオーサリング対象となる実物体 , 位置入力デバイスの視点に対する相対位置姿勢を求める必要がある . 本システムでは超音波位置センサで視点と対象の位置を計測し , 相対位置の予測値として後段の処理で利用する . また , ウェアラブル回転角センサは画像が入力される時点での視点の姿勢を計測し , 後段の処理ではフレーム間の視点の回転差分量のみを用いる .

文献 [11] ではウェアラブルセンサのみを用いて相対位置姿勢を推定するので , 物体が動いたのか視点が動いたのかを判定することはできない . このため , 追跡に用いる予測値の計算において , 物体が動いた場合や視点が動いた場合など複数の状況を想定していた . しかし , 環境側位置センサを導入することで環境に対する視点の位置 , 環境に対する物体の位置をそれぞれ計測することができるので , 物体が動いたのか視点が動いたのかを判定することができるため , 予測処理を軽

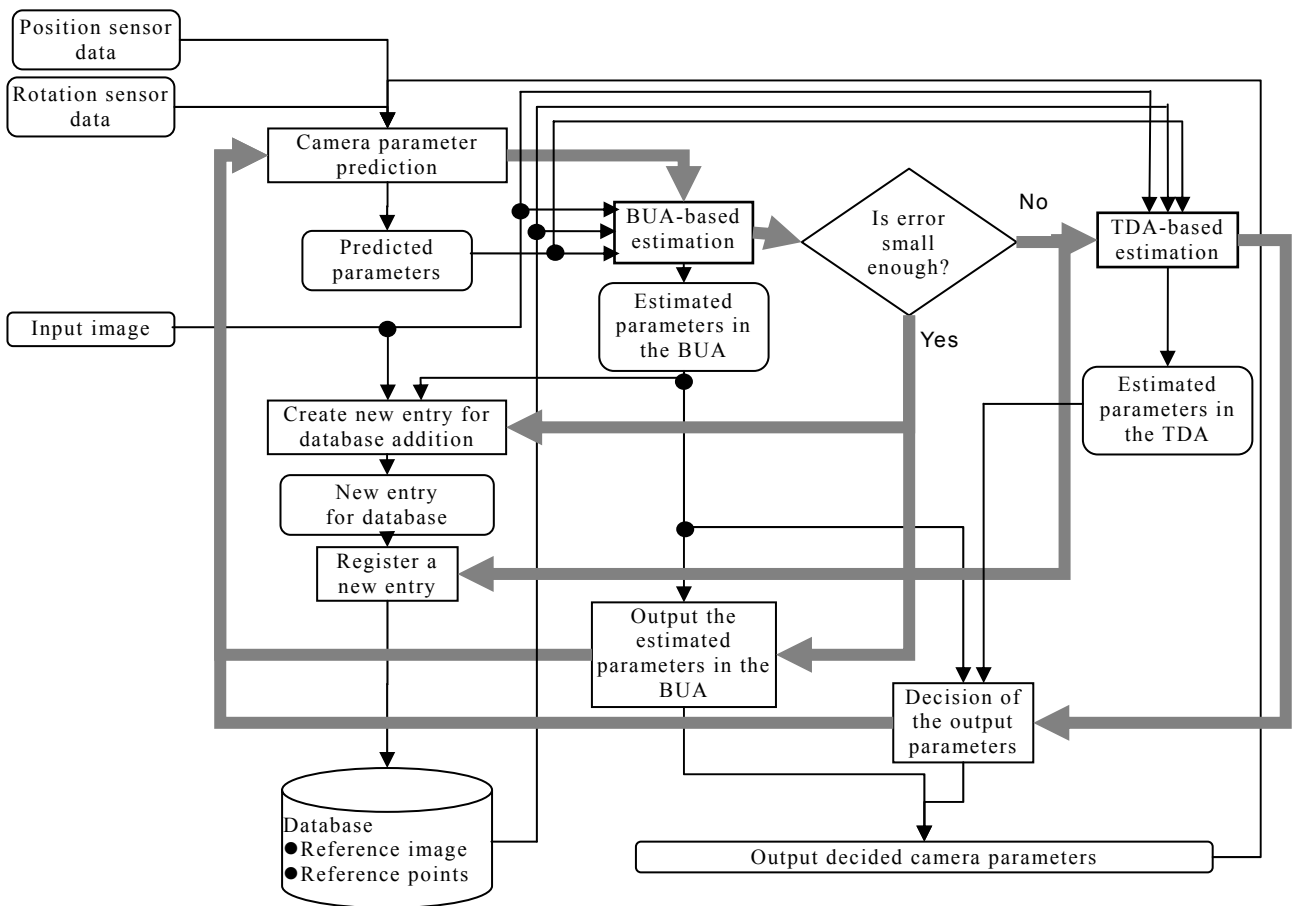


図 4：相対位置姿勢追跡手法の流れ

減することができる。

この追跡手法では多数の解候補をボトムアップアプローチにより解を決定する。求めた解の誤差(後述)が小さい場合はこれを出力とし、誤差が大きい場合にはトップダウンアプローチからも解を求めて、より誤差の小さい解を出力する。また、解の誤差が小さい状態での入力画像と対象の位置姿勢を必要に応じてデータベースに自動登録することで、基準点のアピランス変化に対応し、追跡範囲を拡大する。追跡対象とする3次元物体について、以下の項目を事前知識として利用する。

- カメラ内部パラメータ。
- 基準画像：追跡対象を撮影した画像。
- 基準点：位置合わせの基準となる追跡対象上の点。Lucas-Kanade 法[15]で基準画像から選択した追跡対象上の自然特徴点を基準点とし、実環境での3次元座標を計測しておく。

4.1.1. ボトムアップアプローチ (BUA)

提案手法で用いる BUA の位置姿勢推定ステップでは複数の基準画像の利用と基準画像の光軸周りの回転処理によって基準点検出が可能な領域を拡大する。また、ロバスト統計量の導入によって位置姿勢計算処理

の特徴点の誤追跡に対するロバスト化を図る。以下に提案手法で導入したそれぞれの工夫について説明する。

複数基準画像の利用と基準画像の回転

提案手法では画像局所領域の勾配情報を用いて繰り返し演算で基準点を Lucas-Kanade 法で追跡する。繰り返し演算の初期値となる基準点の2次元座標は前フレームの位置姿勢の推定結果、超音波センサデータから計算された相対位置、回転角センサによって得られたフレーム間の姿勢変化量を用いて予測される位置姿勢によって既知特徴点の3次元座標を画像平面上に投影した位置とする。Lucas-Kanade 法は特徴点の平行移動のみを仮定しているため、アピランス変化に強いとはいえない。このため提案手法では登録された複数の基準画像を用いる。登録基準画像の増加に伴う計算量の増加を最小限にするため、前フレームの推定結果の計算に使用された基準画像と、予測位置姿勢と一番近い位置姿勢が登録されている基準画像のふたつ(一致する場合はひとつ)を選択して利用する。また、対象の形状の影響を受けない画像の光軸まわりの回転を考慮し、前フレームの推定結果と回転角センサによって得られた前フレームから現在までのカメラの回転量を用いて選択された基準画像を回転し、特徴点追跡に

用いる。

LMedS 基準を用いた位置姿勢計算処理

n 個の追跡された基準点の 2 次元座標と 3 次元座標から PnP 問題の解法[16]によりカメラ位置姿勢の解析解を求めることができる。しかしながら n 個全てが正確に追跡されることはまれであり、これを前提として解を求めると正確な解を得られない。そこで提案手法ではロバスト統計量として代表的な LMedS 基準を用いて以下の手順で解を求める。

手順1 n 個の特徴点からランダムに 3 点を選択し、P3P 問題の解法を用いて解候補を計算する。

手順2 次式で定義される解候補の誤差値を計算する。

$$err^2 = \text{med} \left((x_i - \tilde{x}_i)^2 + (y_i - \tilde{y}_i)^2 \right) \quad (1)$$

ただし、 (x_i, y_i) は i 番目の基準点の画像上での観測座標値、 $(\tilde{x}_i, \tilde{y}_i)$ は i 番目の基準点を解候補で画像平面上へ投影したときの座標値、 $\text{med}(f(i))$ は全ての i における $f(i)$ の中央値を表すものとする。

手順3 手順 1. と 2. を m 回繰り返す。割合 r の特徴点 が正確に追跡された場合に確率 p より高い確率で正しい解を求めるための繰り返し回数 m は次式により計算される。

$$p < 1 - (1 - r^3)^m \quad (2)$$

手順4 誤差値最小となる解候補を解として出力する。

4.1.2. トップダウンアプローチ (TDA)

提案手法では BUA によって信頼できる解が得られなかった場合、TDA による位置姿勢推定を行う。本研究では TDA として ConDensation アルゴリズム[14]を用い、以下の手順を毎フレーム繰り返すことで位置姿勢を推定する。

手順1 サンプル生成処理：前フレームの状態に応じてサンプル集合を生成する処理。詳細は後述。

手順2 観測処理：入力画像を観測して各サンプルの評価値を計算する。各サンプルの評価値の計算にも err を用いるためには基準点 3 次元座標を各サンプルで画像平面上に投影した座標 $(\tilde{x}_i, \tilde{y}_i)$ を初期値とする特徴点追跡により対応する基準点の入力画像中の座標 (x_i, y_i) を求めなければならない。計算量が膨大になる。そこで、提案手法では事前知識の準備で用いた Lucas-Kanade 法で入力画像中の自然特徴点を検出し、各々の $(\tilde{x}_i, \tilde{y}_i)$ に最も近い検出特徴点 (x'_i, y'_i) で (x_i, y_i) を代用して式(1)の誤差値を計算し評価する。

手順3 代表値決定処理：推定された解の確率密度からの代表値として誤差値最小のサンプルを出力する。

サンプル生成処理

効果的なサンプルを重点的に生成するために、BUA の計算結果、超音波位置センサから得られた視点と対

象の相対位置、及び回転角センサから得られた姿勢変化量を利用する。実際には以下の手順で時刻 t のサンプル集合 $\{s_t^{(1)}, s_t^{(2)}, \dots, s_t^{(N)}\}$ を生成する。ただし、 N は全サンプル数、 s_t は時刻 t のサンプルで位置姿勢を表す。

まず、全体の 1/2 のサンプル $\{s_t^{(1)}, s_t^{(2)}, \dots, s_t^{(N/2)}\}$ を、時刻 t の BUA の計算結果を分布中心とするランダムサンプリングによって生成する。残りの 1/2 のサンプルについては時刻 $t-1$ の出力が BUA、TDA のどちらにより計算されたかで生成方法を変える。

時刻 $t-1$ の出力が BUA で計算された場合：

超音波センサデータから計算された相対位置、時刻 $t-1$ の出力と回転角センサデータから得られる姿勢変化量をもとに予測した姿勢を分布中心とするランダムサンプリングによって $\{s_t^{(N/2+1)}, s_t^{(N/2+2)}, \dots, s_t^{(N)}\}$ を生成する。

時刻 $t-1$ の出力が TDA で計算された場合：

まず、ConDensation アルゴリズムに従ったサンプル選択処理を行う。すなわち、時刻 $t-1$ のサンプル集合 $\{s_{t-1}^{(1)}, s_{t-1}^{(2)}, \dots, s_{t-1}^{(N)}\}$ から各サンプルの評価値の比に従って全サンプル数の 1/2 を選択する。選択されたサンプル集合を $\{s_t^{(1)}, s_t^{(2)}, \dots, s_t^{(N/2)}\}$ とする。

次に、選択されたサンプルのうち $\{s_t^{(1)}, s_t^{(2)}, \dots, s_t^{(N/4)}\}$ をランダムウォークで更新して $\{s_t^{(N/2+1)}, s_t^{(N/2+2)}, \dots, s_t^{(3N/4)}\}$ を作成する。

最後に、選択されたサンプルのうち $\{s_t^{(N/4+1)}, s_t^{(N/4+2)}, \dots, s_t^{(N/2)}\}$ の各々の姿勢パラメータと姿勢変化量から計算される姿勢と超音波センサデータから計算された相対位置を中心にランダムウォークで更新して $\{s_t^{(3N/4+1)}, s_t^{(3N/4+2)}, \dots, s_t^{(N)}\}$ を作成する。

4.1.3. アピランスの自動追加登録

本追跡手法では視点移動に伴う基準点のアピランス変化のに対応するため、複数視点から撮影した基準画像と基準点を登録しているが、安定して追跡するために必要な画像の撮影位置は追跡対象の形状やテクスチャに依存する。そこで、追跡処理中に必要に応じてアピランスを追加する方針をとる。具体的には、ボトムアップアプローチの誤差が大きくなり、トップダウンアプローチが開始されるときに新しいアピランスが必要であると判断して、直前のフレームの入力画像、位置姿勢、及び追跡に成功した基準点の 2 次元座標を新しいアピランスとしてデータベースに追加登録する。

5. 実験と考察

本システムの動作を確認するために、テーブル上に固定した物体を視点を動かしながら観察する動作について、視点に対する物体の相対位置姿勢の追跡実験を行った。図 5 はこのとき推定された位置姿勢を示す。途中、画像上の動きの大きい場所で追跡に失敗するこ

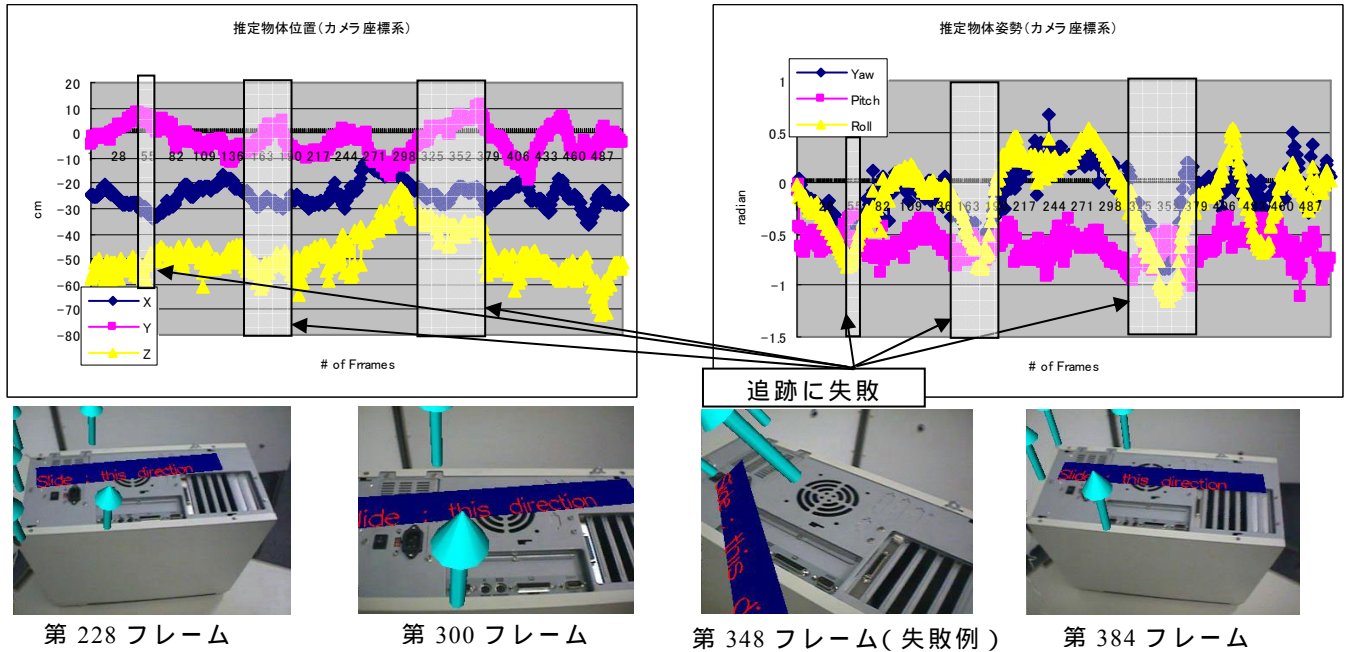


図 5 : 実験結果

とがあったが、その後復帰することができている。今後は、超音波センサと回転角センサからより精度の良い予測値を計算する手法について検討することで位置姿勢の追跡の頑健性を向上し、複合現実オーサリング環境に必要な実物体上に観測された新規特徴点の3次元位置計測の機能の付加や、環境側センサの入力値を利用しないで出力を確認する機能など、実際にオーサリング作業を行う環境を構築していく予定である。

6. 謝辞

本研究の一部は、文科省の科学技術振興調整費の支援による。

文 献

- [1] Azuma, R.T., "A survey of augmented reality," Presence, vol.6, No.4, pp.355-385, 1997.
- [2] Azuma, R., Bailiot, Y., Behringer, R., Feiner, S., Julier, S., and MacIntyre, B., "Recent Advances in Augmented Reality", IEEE Computer Graphics and Applications, Vol. 21, No. 6, pp. 34-47, 2001.
- [3] Neumann, U., and Cho, Y., "A self-tracking augmented reality system," In Proc. VRST 96, pp. 109-115, 1996.
- [4] Neumann, U., You, S., Cho, Y., Lee, J. and Park, J., "Augmented Reality Tracking in Natural Environments," Mixed Reality - Merging Real and Virtual Worlds, Ohmsha & Springer-Verlag, pp. 101-130, 1999
- [5] Rekimoto, J., "Matrix: A Realtime Object Identification and Registration Method for Augmented Reality," APCHI'98, 1998.
- [6] Kato, H., and Billinghurst, M., "Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System," In Proc. the 2nd IEEE and ACM International Workshop on Augmented Reality '99, pp.85-94, 1999.
- [7] You, S. and Neumann, U., "Fusion of Vision and Gyro Tracking for Robust Augmented Reality Registration," In Proc. VR2001, pp. 71-78, 2001.
- [8] Weavy: <http://www.is.aist.go.jp/weavy/>
- [9] Okuma, T., Kurata, T., and Sakaue, K., "Real-Time Camera Parameter Estimation for 3-D Annotation on a Wearable Vision System," IEICE Trans. Inf. Syst., Vol.E84-E, No. 12, pp.1668-1675, 2001.
- [10] Okuma, T., Kurata, T., and Sakaue, K., "VizWear-3D: A Wearable 3-D Annotation System Based on 3-D Object Tracking using a Condensation Algorithm," In Proc. IEEE Virtual Reality 2002, pp.295-296, 2002.
- [11] 大隈隆史, 蔵田武志, 坂上勝彦, "ボトムアップ・トップダウン統合アプローチとアピアランス自動登録によるマーカを用いない3次元物体追跡手法," 日本VR学会第7回大会, pp.99-102, 2002.
- [12] 西田佳史, 相澤洋志, 堀俊夫, 柿倉正義, "超音波センサを用いた対象物のセンサ化に基づく人の日常活動の認識," 日本機械学会ロボティクス・メカトロニクス講演会'02 講演論文集, 1A1-J07, 2002.
- [13] 西田佳史, 相澤洋志, 堀俊夫, 柿倉正義, "超音波3次元タグを用いた人の日常活動の頑健な計測~冗長なセンサ情報に基づくロバスト位置推定~, " 第20回日本ロボット学会創立20周年記念学術講演会予稿集, 3C18(1)-(4), 2002.
- [14] Isard, M. A., "Visual Motion Analysis by Probabilistic Propagation of Conditional Density," Ph.D. thesis, Department of Engineering Science, University of Oxford, 1998.
- [15] Lucas, B.D. and Kanade, T., "An Iterative Image Registration Technique with an Application to Stereo Vision," Proc. DARPA Image Understanding Workshop, pp.121-130, 1981
- [16] Haralick, R.M., Lee, C.-N., and Ottenberg, K., "Analysis and solutions of the three point perspective pose estimation problem," Proc. CVPR '91, pp.592-598, 1991.