

デプスカメラにより取得される 部分形状テンプレートのマッチングによる 3次元物体のリアルタイムトラッキング

坪井 一菜^{1,a)} 鐘ヶ江 資子¹ 小山田 雄二¹ 杉本 麻樹¹ 斎藤 英雄¹

概要: プロジェクタを用いてコンピュータグラフィックス (CG) 物体を実環境に直接重畳する Spatial Augmented Reality (SAR) の研究では, CG を実環境の物体に重畳する為実環境の物体の 3 次元位置をリアルタイムで求める必要がある. 3 次元物体のトラッキングではしばしば物体の全周データのマッチングが用いられているが, 全周データの作成は困難である. また SAR は光源の変化が起きる環境中であるためにカラー画像を用いたトラッキングは困難である. 本論文ではデプスカメラのみを用いて物体の部分形状テンプレートを作成し, 光源が変化する環境中での 3 次元物体のトラッキング手法を提案しカラー画像を用いずに SAR を実現する. 提案手法ではまず複数の異なる視点からトラッキング対象の部分形状テンプレートを取得しテンプレートの互いの位置関係を算出する. 次に形状の合致するテンプレートとフレームの初期位置合わせを行う. 続いて初期位置合わせを行ったフレームからフレーム間マッチングを行いトラッキングを継続する. 加えて, 物体の角度変化を計測し見え方の変化が起きた場合, テンプレートの位置関係をもとに部分形状テンプレートから現フレームに合致するテンプレートを選択し再初期化を行う. 実験結果より光源変化がが起きている環境中でもトラッキングが行われていることが確認できた. また提案手法を用いて SAR が実現できた.

キーワード: Spatial Augmented Reality, デプスカメラ, リアルタイムトラッキング, 3 次元点群

Realtime 3D Object Tracking Using Partial Shape Templates

KAZUNA TSUBOI^{1,a)} MOTOKO KANEGAE¹ YUJI OYAMADA¹ MAKI SUGIMOTO¹ HIDEO SAITO¹

Abstract: This paper presents a new realtime 3D object tracking method using a single depth camera for Spatial Augmented Reality systems. The method is based on template matching and frame-to-frame tracking of point cloud data captured from depth camera. Point clouds of different views of the object are first trained as partial shape templates. Then the best shape matching template is chosen and initialized to the current frame. Tracking continues by frame-to-frame matching of the point cloud. New template is chosen when the view of the object has changed. Experiment result shows that our method is able to track 3D object in different lighting conditions.

Keywords: Spatial Augmented Reality, Depth Camera, 3D object, tracking, realtime

1. 本研究の背景

コンピュータグラフィックス (CG) 物体を実環境に重

畳することによって実空間に CG 物体が存在するような感覚を実現する技術, 拡張現実感 (AR) の研究が行われており, 現実空間に視覚的に情報を付加することができる事から, 作業支援, ナビゲーションシステム, ゲームなど幅広い分野で利用されている.

¹ 慶應義塾大学大学院理工学研究科
Graduate School of Science and Technology, Keio University
^{a)} kazunat@hvrl.ics.keio.ac.jp

AR を実現するにあたり重要な要素がある．1つは仮想物体を実環境のどこに表示するのかという幾何学的問題の解決，もう1つはCG物体の表示方法である．

CG物体を現実世界に違和感なく重畳するには，実環境を把握しその状況に合わせてCG物体を重ね合わせる必要がある．実環境を把握する方法としてカメラの位置姿勢を推定することが有用であり，コンピュータビジョンの技術を用いたカメラの位置姿勢の推定にはマーカを用いる方法 [4] がある．また近年マイクロソフトが開発した Kinect により高価であったデプスカメラが手軽に手に入るようになった．デプスカメラを用いるとカラーカメラでは取得が困難であったカメラと実環境の3次元位置の把握がカメラ1台できることから，デプスカメラを用いたARの実現が注目を浴びている．

CG物体の表示方法としては，ビデオスルーが一般的である．ビデオスルーではコンピュータ，携帯電話などの画面でカメラから撮影した実環境の映像にCG物体を重畳する．またヘッドマウントディスプレイ (HMD) は装着することによって眼前に画面が配置されより没入感を得ることのできるディスプレイである．これらのビデオスルータイプの他にプロジェクタを用いて現実環境に光でCG物体を重畳する手法 [8] も存在する．これらの研究は Spatial Augmented Reality (SAR) と呼ばれ，プロジェクタを用いてARを行うと光で実環境に直接CG物体の重畳を行う為，肉眼で重畳の様子を観測できることから臨場感のあるARの実現が期待でき，実環境の立体物とのインタラクションを行うSARのシステムの構築研究 [5] も行われている．

2. 関連研究

本章では関連研究とその問題点及び本研究の位置づけについて述べる．まずSARの研究について，次に3次元物体のトラッキングに関する研究について，続いてデプスカメラによるトラッキングと3次元点群の位置合わせについて述べ，最後に関連研究と本研究の位置づけについて述べる．

2.1 Spatial Augmented Reality

SARを実現するにはプロジェクタで光を投影する対象物体の3次元位置の把握が必要である．あらかじめ実環境を計測しておくことも有用だが，光の投影対象が動的に変化する場合は投影対象を実時間でトラッキングする必要がある．Bandyopadhyayらは [2] の研究において投影対象に赤外線と移動磁気センサをつけ物体の移動のトラッキングを行い，表面に投影を行った．Manarらは [6] の研究において物体にモーションキャプチャ用のマーカをつけ物体の移動を把握した．しかしながらセンサやマーカを用いる場合使用できる物体は限られる．提案手法ではトラッキング対象物体のテンプレートの登録を行い，トラッキング対象に

物理的なセンサやマーカを用いずにトラッキングを行う．

2.2 3次元物体のトラッキング

3次元物体のトラッキングを行う手法としては，古くから3次元モデルをあらかじめ持ちそのモデルとカラーカメラで写した2次元画像からの点や線によるマッチングを行う方法が用いられている．Harrisらは作成した3次元モデルと2次元の線分マッチングを行うことによってトラッキングを実現した [3]．Azadらはあらかじめ3次元モデルから作り出した見え方の画像と2次元画像のマッチングによってトラッキングを実現した [1]．しかしながらSARではトラッキング対象に光を投影し時間推移の中で物体の色調が変化するため画素値を用いた手法ではトラッキングが非常に困難である．また3次元物体の全周モデルは作成が困難である．また，プロジェクタで投影するCG物体の色情報を用いて投影した様子を撮影した画像から色を分離する研究もおこなわれているが，投影条件によって実際に映射される色調がコンピュータ上の数値と異なる場合があり，計算による色の分離は非常に困難である．

2.3 デプスカメラを用いたトラッキング

デプスカメラを用いると色調ではなくカメラと物体までの距離を画像として得ることができる．ParkらはRGBカメラとデプスカメラを組み合わせることによりテクスチャを用いずに剛体のトラッキングを実現した [7]．しかしながら，手法中ではカラー画像を用いてデプスマップの修正を行っている他，カラー画像を元に作成したテンプレートを用いている．SAR環境は光源が変化する環境であることから，カラー画像を含むトラッキングは困難であり，本研究ではデプスカメラ1台のみを用いてトラッキングを行う．

2.4 Point Cloud と位置合わせ

ロボティクスの分野ではロボットの移動や位置把握のために，複数視点から実環境を撮影した Point Cloud を用いて3次元モデルを作成する研究が盛んに行われている [10]．Point Cloud とは3次元点の集まりであり，デプスカメラなどを用いることで取得することが可能である．3次元モデル作成には異なる視点の Point Cloud を同じ座標系で表すことによって位置合わせを行う．互いの Point Cloud における対応点を用いてターゲットとなる Point Cloud への回転と平行移動の行列を求め，Point Cloud を移動させることによって同じ座標系で表すことが可能である．本研究はこの Point Cloud の位置合わせの手法を部分形状テンプレートとフレーム間の Point Cloud のマッチングに適用することでトラッキングを実現した．

2.5 本研究と関連研究の位置づけ

本研究ではカラー画像を一切用いずにデプスマップのみ

を入力としてトラッキングを行う。またトラッキングを行う 3 次元物体のモデルを Point Cloud で表された部分形状テンプレートとして分割して保持し、Point Cloud のマッチングを行うことによって、高速化を行う。

3. 目的

前述の通り、動的な SAR の実現における従来の 3 次元物体のトラッキングはセンサを必要とし使用できる物体は限られた。またカラーカメラでの 3 次元物体のトラッキングはあらかじめ 3 次元モデルを用意することによって実現されたが、モデルの作成は困難であり、さらに明度値を用いてマッチングを行う為に光の変化に弱く、光源の変化する環境である SAR での利用は適さない。したがって本研究ではデブスカメラによるデブスマップからトラッキングする物体の部分形状テンプレートを登録しテンプレートと現フレームのマッチングを行い、光現環境の変化にも頑健なトラッキングを実現する手法を提案する。

4. 提案手法

本研究では固定したデブスカメラから得たデブスマップを用いて、動的な剛体のトラッキングを行う場合を想定する。デブスマップはカメラと被写体までの距離を表した画像であることから、3 次元点の集まり Point Cloud として扱うことが可能である。このことにより点群の位置関係を用いて点の特徴を表す特徴量 Fast Persistent Feature Histogram (FPFH) [9] を用いることが可能である。この FPFH を用いて、追跡する物体のテンプレートと現フレームの位置合わせ及びフレーム間における物体の移動推定を組み合わせることにより、3 次元物体の位置姿勢の推定を行いトラッキングを実現する。

次項でまず Point Cloud の位置合わせについて述べ、その後提案手法の各処理について詳しく述べる。

4.1 Point Cloud の位置合わせ

4.1.1 Fast Persistent Feature Histogram

ある Point Cloud をもう一方の Point Cloud にあわせるためには、双方の Point Cloud における対応点のペアを求め、回転と平行移動の行列を求める必要がある。精度よく位置合わせを行う為には、より正しい対応点のペアを求めることが重要である。本手法においては Point Cloud の対応点を求める為に Rusu らによる特徴量 FPFH を用いる。FPFH は Point Cloud 中の点とその周囲の点の関係を 4 つのスカラー値で表した値をヒストグラムで記述した特徴量である。この特徴が似た点の対を位置合わせに用いる Point Cloud から求めることにより位置合わせに用いる対応点を取得することが可能である。

4.1.2 Point Cloud 位置合わせアルゴリズム

本手法では Point Cloud の位置 Sample Consensus Align-

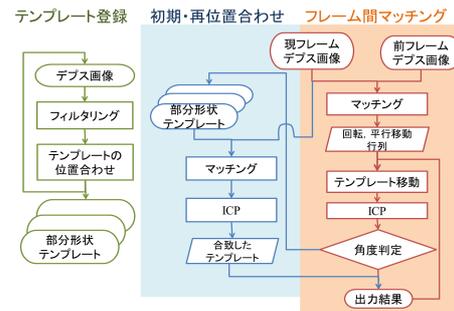


図 1 提案手法の概要

Fig. 1 flow chart

ment[10] アルゴリズム (SAC-IA) と Iterative Closest Point(ICP) の 2 種類のアルゴリズムを組み合わせる。

SAC-IA はターゲットの Point Cloud のうちテンプレートからランダムに選んだ点との特徴量が最も近い点を対応点として選択する。求めた対応点から求めた回転行列と平行移動行列を用いて移動したテンプレートとターゲットの Point Cloud 間のエラーを算出する。対応点の組み合わせ選択を複数回行い、エラーの値が最も少ない行列を求める。特徴量として FPFH を用いるにより精度の高い Point Cloud の位置合わせが可能である。

ICP アルゴリズムはターゲットの Point Cloud のうちテンプレートからランダムに選んだ点ともっとも近い点を対応点として選び、点の選択、移動を再帰的に用いることによってテンプレートを徐々にターゲットへ近づける方法である。前提として、用いる 2 点の位置合わせは既に大まかに行われているという条件があるため、形や位置が大きく異なると正しく位置合わせを行うことはできない。この 2 つのアルゴリズムを組み合わせることによってより正確な位置合わせを実現することが可能である。

4.2 提案手法の各処理について

本手法の流れは大きく部分形状テンプレート登録とトラッキングに分かれる。図 1 に提案手法の概要を示す。

テンプレートの登録においてはまず追跡する対象のテンプレートの登録を行う。3 次元の物体は見え方によって形状が異なることより、面に対応した部分形状テンプレートを用意する。続いてトラッキングではまず初期位置合わせでテンプレートと現フレームの位置合わせを直接行う。次にフレーム間マッチングでは現フレームと一つ前のフレームの位置合わせを行い、回転と平行移動の行列を算出する。フレーム間マッチングで求めた行列を用いてテンプレートの移動を行い現在のフレームにおける物体の位置を検出する。またこの行列から物体の回転が求まり、回転角に応じて最適テンプレートの再選出を行い再初期化を行う。

4.2.1 部分形状テンプレートの登録

カメラの奥行き方向の軸を Z 軸、横方向を X 軸、縦方向を Y 軸とする。トラッキングを行う 3 次元物体は体積があ



(a) 左から見た物 (b) 正面から見た (c) 右から見た物
体 物体 体



(d) 左からのテ (e) 正面のテン (f) 右からのテ
ンプレート プレート ンプレート

図 2 見え方の違いと登録したテンプレート
Fig. 2 view changes and trained template

るため、物体が X 軸または Y 軸方向に回転すると物体の見え方も変化する。図 2(a) から図 2(c) はトラッキングに用いたモデルの一つである。この 3 つの図からわかるように同じ物体でも視点が異なると物体の見え方も異なる。本手法ではトラッキング対象の一部分から見た Point Cloud をテンプレートとして登録する。このテンプレートを部分形状テンプレートと呼び、部分形状テンプレートを複数用いることによって見え方の変化に対応する。複数の部分形状テンプレートの 1 つに残りのテンプレートの位置合わせを行い互いの位置関係を求めることにより、複数のテンプレートを 1 つの同じ物体として扱うことができる。登録時には物体を時計回りに回転させ順番に応じて連続したラベルをつける。このラベルは後述するトラッキング再初期化の処理に用いる。図 2(d) から図 2(f) は実際に登録したテンプレートの例である。

4.2.2 初期位置合わせ

初期位置合わせでは部分形状テンプレートの Point Cloud と現フレームの Point Cloud の位置合わせを直接行う。SAC アルゴリズムを用いて、部分形状テンプレートの Point Cloud 内のランダムに選ばれた点の FPFH に最も近い点を現フレームにおける Point Cloud の点から選び、対応点とする。その対応点をもとに部分形状テンプレートと現フレーム間の回転と平行移動行列を得る。部分形状テンプレートは複数存在するが、行列で移動した部分形状テンプレートの点と最もユークリッド距離に近い現フレームの点の間の距離の平均値を求め、最も値が小さいテンプレートが現フレームの形状に最も合致したテンプレートとして得ることができる。本手法においては処理速度をリアルタイムに近づけるために SAC の再帰数を少なくしている。SAC の再帰数が少ない為、現フレームと移動したテンプレートの位置には多少の誤差が生じている。フレームの位

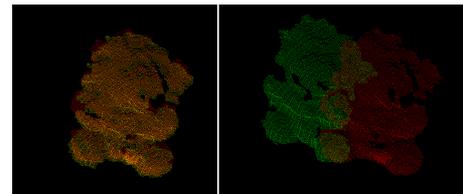


図 3 フレーム間隔の違いによるポイントクラウドの違い

Fig. 3 difference between frames

置合わせをより確実にするために icp アルゴリズムを用いて移動したフレームと現フレームの誤差を縮める。SAC によって結果の良い初期位置が与えられていることから icp アルゴリズムの再帰数を少なくすることが可能であり、処理速度向上に貢献している。

4.2.3 フレーム間マッチング

フレーム間マッチングではフレーム間の回転、平行移動行列を用いてトラッキング対象の位置推定を行う。単純な位置合わせの問題と比べて、剛体のトラッキングでは時系列の変化の様子を追うことが可能であり、これは大きな利点である。図 3 の左はトラッキングする物体が移動しているシーケンスの中で連続したフレームの様子、右は 5 フレームあいた場合の様子を表している。フレーム間隔がいている場合に比べて連続したフレームでは変化が少なくなることがわかる。このように長い時間で見ると一見大きな変化も、ごく短い時間のわずかな変化の積み重ねであると仮定することが可能である。

トラッキングする対象が 3 次元の剛体であっても、フレーム間の幾何学的フレーム間の幾何学的位置の変化は小さく、トラッキングする物体の見え方の変化も小さいと考えられることから、位置合わせに用いる Point Cloud の形の差異も小さい。このことを利用し、フレーム間マッチングではと現在のフレームと 1 つ前のフレームの位置合わせを行い、算出した回転と平行移動行列をテンプレートにかけることにより現フレームにおけるトラッキング対象の位置を推定する。しかしながら、時間を積み重ねていくと蓄積誤差が生じる。蓄積誤差の解消法としてフレーム間の移動を考慮したテンプレートにさらに icp アルゴリズムを用いる。icp アルゴリズムの位置合わせを加えることによってテンプレートをできるだけ現フレームにあわせることにより持ち越す誤差をできるだけ少なくする。

4.2.4 角度判定による再初期化と最適テンプレート選定

トラッキングを継続していく中で 3 次元物体が x 軸、y 軸方向に回転した場合、現在用いているテンプレートの形状が現フレームにおける物体の見え方と合致しなくなる。そこで再初期化の処理を行うことによって現在の見え方に対応するテンプレートの再選出を行う。

再初期化処理への判別方法にはテンプレートが移動した角度を測定することによって行う。フレーム間マッチングで算出された回転行列からはその移動が x 軸、y 軸、z 軸



図 4 セットアップ

Fig. 4 Pro-Cam Snapshot



図 5 トラッキングに使用した物体

Fig. 5 3D objects

についてどれほど回転したのかという角度を算出することが可能である．フレーム間マッチングは時系列に変化を追っていることから，初期化を行った直後のテンプレートの位置から移動角度を加算していくことにより物体が初期化後からどれだけの角度を回転したのかということが分かる．トラッキング対象物体の見え方は z 軸に関する回転による変化は少ないことから， x 軸， y 軸に関する回転のみを考慮に入れる．

現在の見え方に対応するテンプレートの選出はテンプレートの相互関係を利用行う．テンプレートの登録において各テンプレートの相互関係が求められていることから物体のが回転により変化した見え方に合致するテンプレートを予測することが可能である．使用していたテンプレートのラベルに対して最も近いラベルのテンプレートを新たなテンプレートとして用いる．テンプレートを選出した後 SAC アルゴリズムと ICP アルゴリズムを用いて現フレームとの位置合わせを行う．またトラッキングにおいて連続したフレームを追っていることからフレーム間の変化は少ないことが予想され，もし回転の数値が大幅に変化した場合はトラッキングが失敗したとみなすことができることから 4.2.2 で用いた初期化と同様に全テンプレートの中から現在に合致するテンプレートを選出し位置合わせを行う．

5. 実験と結果

本章では提案手法を用いた 3 次元物体のトラッキング実験と，SAR 環境下におけるトラッキング実験を行った．カメラとトラッキング対象物体の位置関係は以下のとおりである．固定されたカメラとプロジェクタに対してトラッキング対象の物体を移動させた．また使用するプロジェクタとカメラはキャリブレーション済みである．図 4 はカメラとプロジェクタの位置関係を表す．

使用環境は以下の通りである．

- デブスカメラ：Microsoft Kinect
- 使用 OS：Windows 7 32bit
- CPU: Intel(R)Core(TM)i7-2600 3.40GHz
- メモリ：3.49GB
- プロジェクタ：Epson EB-W8

5.1 異なる形状の物体に対するトラッキング実験

トラッキング実験では図 5 にある大きさ，形，色の異

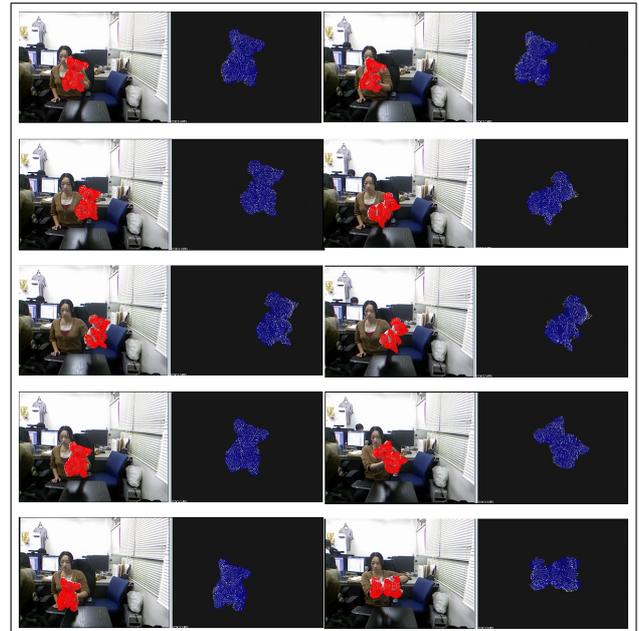


図 6 物体 1 のトラッキング結果

Fig. 6 Object 1 result

なる物体のトラッキングを行った．右から物体 1，物体 2 とする．図 6 から図 7 はそれぞれの物体のトラッキングを行った結果を示している．図の右はテンプレートマッチング結果であり，青の点は現在のフレームの Point Cloud，白の点はテンプレートの Point Cloud を表す．図の左における赤い点は RGB カメラで写した映像にテンプレートの 3 次元位置を投影した結果である．フレームは左上から右下へ進む．

色や形大きさにかかわらずどの物体もカメラに対する平行移動，回転を行ってもトラッキングが持続していることが分かる．複数種類のトラッキング対象物体を用いてもトラッキングが行えていることから物体の形状をあらかじめ持たずに任意の物体のトラッキングが行うことが可能である．また，表 1 は処理速度を表す．物体 1 の 1 フレームの点の数は約 1500 点，物体 2 の 1 フレームの点は 2200 点である．フレームレートは約 4fps で緩やかに動いた物体のトラッキングが実現できている．

表 1 処理速度

Table 1 Speed

| | 初期化 [ms] | フレーム間マッチング [ms] | 全行程 [ms] |
|------|----------|-----------------|----------|
| 物体 1 | 249.2 | 151.2 | 280.6 |
| 物体 2 | 331.0 | 289.0 | 390.9 |

5.2 プロジェクタ照射環境下における SAR

本手法を用いて得た物体のトラッキング結果を用いて，3 次元物体の表面の色を変化させる SAR を行った．トラッキングの結果より得たテンプレートの位置に色を照射する

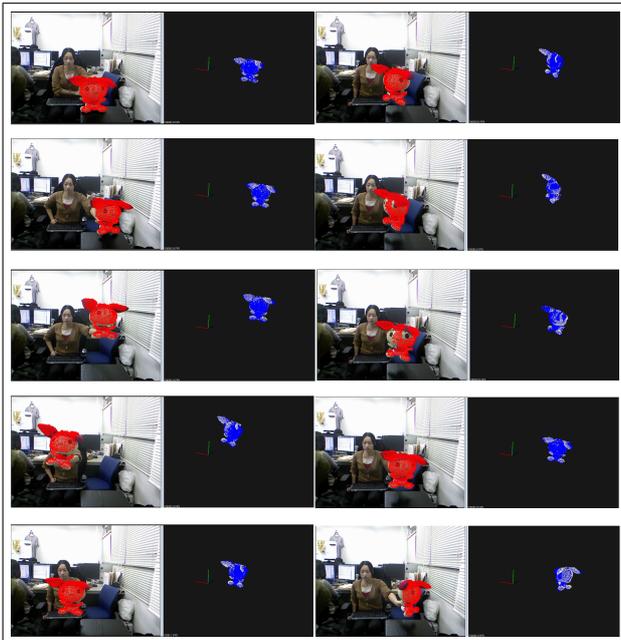


図 7 物体 2 のトラッキング結果
Fig. 7 Object 2 result

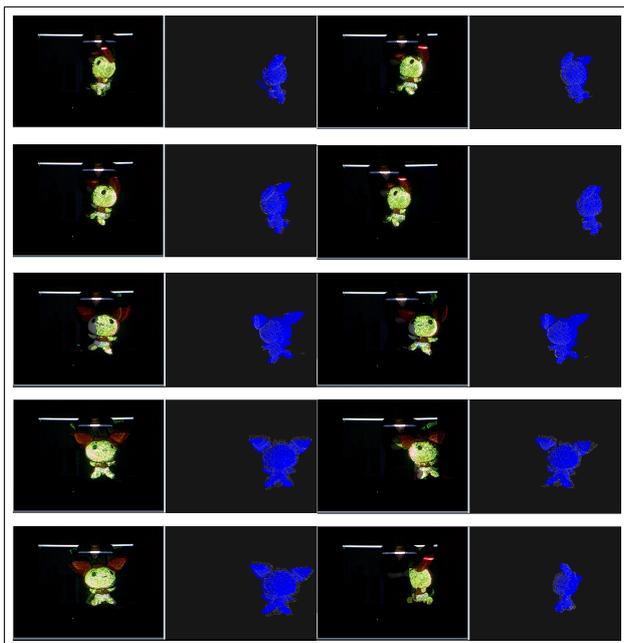


図 8 本手法を用いた SAR
Fig. 8 SAR result

ことで物体全体のテクスチャを変化させることが可能である。図 8 はその実験結果である。3次元物体の位置に光が照射されていることが分かる。物体に対して照射した光がずれているのはトラッキング結果からプロジェクタ照射までの遅延のためである。今後、並列化処理による高速化を図ることで遅延の改善を行い、SAR システムとしての実用性を向上させていきたい。

6. 結論

本研究では光源環境が変化する SAR 環境下で 3次元物

体の位置を把握することを目的とし、デプスカメラを用いて取得した物体の形状を分割した部分形状テンプレートのマッチングによる 3次元物体のトラッキング手法を提案した。実験では光源環境が変化しない環境中でのトラッキングとトラッキング結果を用いて 3次元物体にテクスチャを投影する SAR を行った。結果からも示されているように、部分形状テンプレートをトラッキングを開始する際に作成することによってあらかじめ 3次元物体のフルモデルを用意しなくとも剛体の見え方の違いに対応するトラッキングを実現することができた。また、SAR 実験により光源環境が変化する環境内でもトラッキングが行われることを示した。

参考文献

- [1] P. Azad, D. Munch, T. Asfour, and R. Dillmann. 6-dof model-based tracking of arbitrarily shaped 3d objects. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pp. 5204–5209, may 2011.
- [2] D. Bandyopadhyay, R. Raskar, and H. Fuchs. Dynamic shader lamps : painting on movable objects. In *Augmented Reality, 2001. Proceedings. IEEE and ACM International Symposium on*, pp. 207–216, 2001.
- [3] C. G. Harris and C. Stennett. 3d object tracking at video rate - rapid. In *British Machine Vision Conference (BMVC)*, 1990.
- [4] H. Kato, M. Billingham, I. Poupyrev, K. Imamoto, and K. Tachibana. Virtual object manipulation on a table-top ar environment. In *Augmented Reality, 2000. (ISAR 2000). Proceedings. IEEE and ACM International Symposium on*, pp. 111–119, 2000.
- [5] Jakob Leitner, Michael Haller, Kyungdahm Yun, Woontack Woo, Maki Sugimoto, and Masahiko Inami. Incredible a mixed reality tabletop game experience. In *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology, ACE '08*, pp. 9–16, New York, NY, USA, 2008. ACM.
- [6] M.R. Marner, B.H. Thomas, and C. Sandor. Physical-virtual tools for spatial augmented reality user interfaces. In *Mixed and Augmented Reality, 2009. ISMAR 2009. 8th IEEE International Symposium on*, pp. 205–206, oct. 2009.
- [7] Youngmin Park, V. Lepetit, and Woontack Woo. Texture-less object tracking with online training using an rgb-d camera. In *Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on*, pp. 121–126, oct. 2011.
- [8] Ramesh Raskar, Greg Welch, Kok-Lim Low, and Deepak Bandyopadhyay. Shader lamps: Animating real objects with image-based illumination. In *Proceedings of the 12th Eurographics Workshop on Rendering Techniques*, pp. 89–102, London, UK, 2001. Springer-Verlag.
- [9] Radu Bogdan Rusu, Nico Blodow, and Michael Beetz. Fast point feature histograms (fpfh) for 3d registration. In *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pp. 3212–3217, may 2009.
- [10] R.B. Rusu, N. Blodow, Z.C. Marton, and M. Beetz. Aligning point cloud views using persistent feature histograms. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pp. 3384–3391, sept. 2008.