

TrakMarkのための仮想化現実モデルを用いた ベンチマークスイートの開発とカメラトラッキング手法評価への適用

牧田 孝嗣[†] 大隈 隆史[†] 石川 智也[†] Laurence NIGAY^{††} 蔵田 武志[†]

[†] 産業技術総合研究所 サービス工学研究センター 〒305-8568 茨城県つくば市梅園 1-1-1

^{††} Université J. Fourier, France

E-mail: [†]{k.makita,takashi-okuma,tomoya-ishikawa,t.kurata}@aist.go.jp, ^{††}laurence.nigay@imag.fr

あらまし 3次元モデルを用いて作成された画像データは、カメラ位置姿勢の真値が既知であるため、カメラトラッキング手法の推定精度を評価する際に有用である。そこで我々は TrakMark と呼ばれるプロジェクトの活動の一環として、仮想化現実モデルを用いた画像データの作成、及び公開に取り組んできた。しかし、データ提供者の観点からは、データ作成の人的コストや、データの可搬性に課題があった。一方、ユーザの観点からは、カメラパスなどが想定シーンと合わない場合への対応や、性能評価作業のサポートが望まれている。そこで本研究では、拡張現実 (Augmented Reality: AR) や複合現実 (Mixed Reality: MR) におけるカメラトラッキング手法の精度評価を想定し、仮想化現実モデルを用いた評価用データセットの生成、及びトラッキング手法の性能評価を支援するためのベンチマークスイートを提案する。本稿では、ベンチマークスイートのコンポーネントとして、歩行モーションを含むカメラパス生成を支援するツールを試作した結果、及び試作ツールを用いて生成したデータセットによるカメラトラッキング手法の評価結果について報告する。

キーワード カメラトラッキング, 拡張現実感, 複合現実感, ベンチマーク, 仮想化現実モデル

A benchmark suite with virtualized reality models for TrakMark and its evaluation with a camera-tracking method

Koji MAKITA[†], Takashi OKUMA[†], Tomoya ISHIKAWA[†], Laurence NIGAY^{††}, and Takeshi KURATA[†]

[†] Center for Service Research, National Institute of Advanced Industrial Science and Technology
Umezono 1-1-1, Tsukuba, Ibaraki, 305-8568 Japan

^{††} Université J. Fourier, France

E-mail: [†]{k.makita,takashi-okuma,tomoya-ishikawa,t.kurata}@aist.go.jp, ^{††}laurence.nigay@imag.fr

Abstract Sequences of computer graphics scene image with camera parameters used for rendering the scene are useful for benchmarking camera-tracking methods for augmented reality. In this research, we propose a benchmark suite for camera-tracking methods. The benchmark suite is composed of virtualized reality models, camera paths, and a tool for generating the image sequences with camera parameters and supporting for comparing ground-truth with estimated camera parameters by a camera-tracking method to be evaluated. In this paper, we show a prototype of the tool for generating the sequences using walking motion models. Moreover, benchmarking results of a tracking method with the data sets are described.

Key words Camera tracking, Augmented and Mixed reality, Benchmarking, Virtualized reality models

1. はじめに

画像処理技術を用いたトラッキング手法の精度を評価するには、カメラ位置姿勢や特徴点の位置などの真値データ

(Ground-truth data)、もしくはそれに準ずる参照データ (Reference data) が有用である。そのため、これまでにオプティカルフロー推定を評価するための画像シーケンス [1] や、テンプレート画像を用いたカメラトラッキングを評価するた

めのデータセット [2] などが提案されている。

そこで、日本バーチャルリアリティ学会複合現実感研究委員会内に設けられたワーキンググループ：TrakMarkWG では、拡張現実感（AR）、複合現実感（MR）のトラッキング手法の評価方法の標準（以下、プロセス標準）、及び評価に用いるデータセットの形式の標準（以下、フォーマット標準）の策定に向けた活動が行われている [3]。これまでに我々は、TrakMarkWG の活動の一環として、仮想化現実モデルのレンダリングにより作成した画像のシーケンスを、各時刻における画像とレンダリングに用いたカメラパラメータの組の連続データセットとして公開 [4] してきたが、データセット提供者の観点からは、データセット作成の人的コストが大きい場合や、画像群のデータ量が大きく、データセットの可搬性が低い場合があるという課題があった。一方、ユーザの観点からは、評価対象とするトラッキング手法が想定するようなカメラの動きがすでに提供されているデータセット中に含まれていない場合に適切な評価ができないという問題がある。また、トラッキングの性能評価のために、トラッキング成功率や処理速度、仮想物体の投影誤差など複数の項目を評価する必要があるため、ユーザによる性能評価のサポートを行う機構が望まれている。そこで本研究では、仮想化現実モデルを用いたデータセット生成、及びトラッキング手法の性能評価を支援するためのベンチマークスイートを提案する。以降、ユーザによる性能評価をベンチマーキングと呼ぶ。また、仮想化現実モデルのレンダリングにより作成した画像を生成画像と呼ぶ。

2. ベンチマーキングをサポートするフレームワーク

ベンチマーキングをサポートするためのフレームワークの概要を図 1 に示す。本研究では、サーバを用いて、3 種類のベンチマークリソース（A：データセット群、B：ベンチマークスイート、C：支援ツール拡張）の提供、及びベンチマーキング結果の公開を行うことを想定している。以下、それぞれのベンチマークリソースについて述べる。

2.1 データセット群

前述のように、各時刻における生成画像とレンダリングに用いたカメラパラメータの組の連続データセットを、様々な仮想化現実モデル及びカメラパスを用いて生成したデータセット群を適用している。データセットを用いたベンチマーキングには、前述の課題（データセットが想定シーンと合わない場合がある、データ量が大きい、ベンチマーキング実施に手間がかかる）がある。しかし一方では、ダウンロードのみで手軽に利用可能であるといった利点があり、ベンチマークスイート導入前の、小規模な予備実験に適している。

2.2 ベンチマークスイート

ベンチマークスイートは、ベンチマーキングの支援のためのツール（以下、支援ツール）と、その入力サンプルとして用いられる仮想化現実モデルデータとカメラパスデータで構成される。以下では、支援ツールが備える機能について述べる。

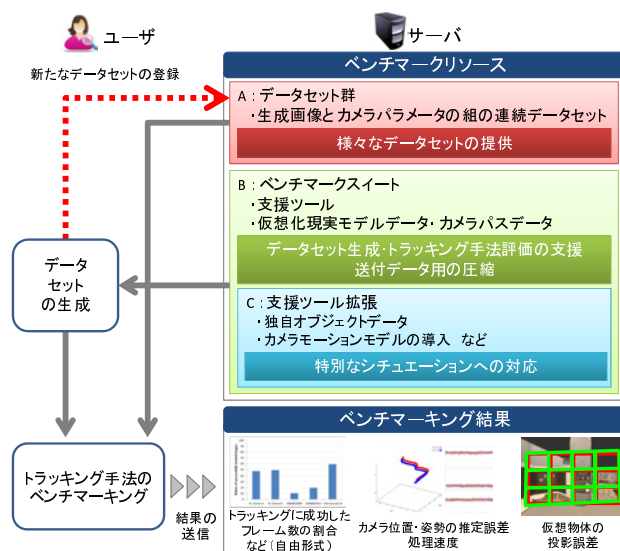


図 1 ベンチマークフレームワークの概要

新しいデータセットの生成：支援ツールを用いて、仮想化現実モデルを見回しながらカメラパスを自作することで、独自のデータセットが作成できる。また、自作したデータセットをサーバに新規登録できる。支援ツールの設計にあたっては、手動で全てのフレームのカメラパラメータを設定してカメラパスを生成する作業の人的コストが大きい点を考慮し、本ツールでは、モデル空間内に離散的に制御ポイントを設定することで、設定した制御ポイント間のデータが自動的に補間される手法 [5] を適用する。

既存のデータセットの生成：ベンチマークスイートには、既存のデータセットの生成に用いられた仮想化現実モデル、及びカメラパスが含まれるため、ユーザによる既存のデータセットの生成も可能である。ベンチマークスイートのデータ量は、仮想化現実モデル画像のデータ量と比べ非常に小さいため、本機能は、送付データ量の圧縮につながる。

ベンチマーキングの支援：データセットを用いて、自身の持つトラッキング手法によって推定カメラパスを作成し、本ツールに入力することで、プロセス標準に即したベンチマーキング結果の自動生成や、サーバへの結果の送信ができる。

2.3 支援ツール拡張

拡張現実の多様な応用シナリオを想定し、現実的な前提条件や制約条件のもとでの評価を可能にすることで、評価対象とするトラッキング手法の得意・不得意な応用分野などの特徴を分析することができる。そこで本フレームワークでは、分野の発展に伴って想定シナリオが増える場合等に柔軟に対応するために、支援ツールを拡張する枠組みを準備している。以下に、想定する支援ツール拡張の例を挙げる。

例 1：独自オブジェクトの追加

トラッキング手法の中には、マーカやランドマーク、ポスター等、独自のオブジェクトを利用する手法や、動物体の存在を想定している手法などがある。そこで、独自オブジェクトの読み込み機能を支援ツールに追加する。

例 2：カメラモーションモデルの導入

AR や MR では、ハンドヘルド、ヘッドマウント、車載型、などの、さまざまなカメラが想定される。そこで、カメラモーションモデルを支援ツールに導入し、想定するカメラ特有の動きを効率的に生成できるように支援する。

3. 実 験

提案するフレームワークの有効性を確認するとともに、今後、支援ツールで提供する予定の誤差評価機能に用いる指標を検討するために、現在実現済みのフレームワークを実際に用いてベンチマーキング実験を行った。以下に、本実験に用いたベンチマークスイート、カメラトラッキング手法、及びベンチマークの指標についてそれぞれ述べる。

3.1 ベンチマークスイート

我々が従来作成したデータ生成ツール [5] を支援ツールのベースとして用い、支援ツール拡張のコンポーネントとしてカメラモーションモデルを導入する機能を試作した。

本試作では、ヘッドマウントカメラを想定し、文献 [6] に基づいた歩行モーションモデルを導入した。ユーザは図 2 に示す 6 つのパラメータを設定し、続いてモデル内に制御ポイントを複数個設置することで、制御ポイント間のカメラパラメータが自動生成され、歩行モーションの揺れを再現したカメラパスを生成できる。作成したデータの例を図 3 に示す。また、作成したデータセットの仕様を表 1 に示す。本実験で作成した画像は、歪みの無い画像であり、カメラ内部パラメータとしては、投影中心 C_x, C_y 、及び焦点距離 F_x, F_y を使用した。なお、支援ツールを用いて複数回のデータ作成実験を行った結果、移動と見回しを含むシーケンスで、1 分以内程度の物であれば、概ね 10 分程度の作業で作成が可能であった。

3.2 カメラトラッキング手法

ベンチマーキングの対象となるカメラトラッキング手法として、キーフレームマッチングと特徴点追跡を併用したカメラトラッキング手法 [7] を基にした手法を用意した。以下に、用意した手法に関して、事前準備、キーフレームマッチング、及び特徴点追跡についてそれぞれ述べる。

3.2.1 事前準備

事前準備として、はじめに、キーフレームとしてカメラ位置姿勢が既知の画像を用意する。次に、各画像について SURF [8] による特徴点検出を行い、結果を保存しておく。最後に、検出された特徴点の 3 次元座標を付与する。本実験では、3.1 節において作成された生成画像（フレーム番号：0～348）から 4 枚（フレーム番号：0, 90, 180, 270）を選び、SURF による特徴点検出を行った。その後、仮想化現実モデ

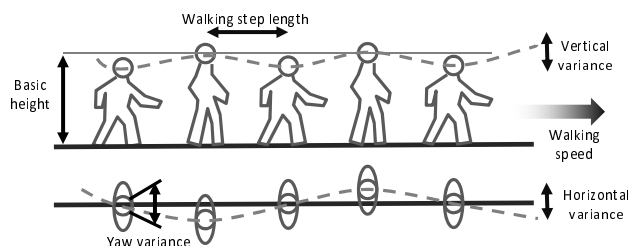


図 2 歩行モーションのモデル化

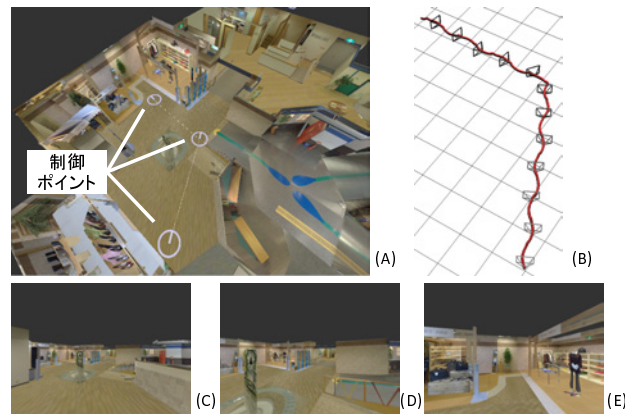


図 3 カメラパスとモデル画像の生成結果 (A：制御ポイント設定時の様子、B：生成されたカメラパス、C～E：生成されたモデル画像の例)

ルとキーフレームのカメラ位置姿勢を用いて、検出された特徴点の 3 次元座標を算出した。図 4 に、キーフレーム画像、図 5 に、特徴点の 3 次元位置を可視化した様子をそれぞれ示す。

3.2.2 キーフレームマッチング

カメラ位置姿勢が未知の状態では、入力画像に対して SURF を検出した上で、各キーフレームの SURF を順に検索して、特徴点の対応付けを行う。十分な数の特徴点に対応付けられた場合には、マッチングが成功したと判断し、対応付けられた点の 2 次元座標と 3 次元座標の組の集合を用いて、RANSAC 法 [9] により入力画像のカメラ位置姿勢の推定値を算出する。マッチングが成功した場合には、次フレームより特徴点追跡に移行する。マッチングが失敗した場合には、次フレームにおいて再びキーフレームマッチングの処理を行う。

3.2.3 特徴点追跡

本フェーズでは、前フレームのカメラ位置姿勢が既知であることを前提として、KLT Tracker [10] [11] を用いた特徴点追跡により現在のフレームにおける特徴点の二次元座標を求め、マッチング時と同様に RANSAC 法を用いて位置姿勢の推定値を算出する。特徴点追跡が成功した場合には、次フレームにおいて再び特徴点追跡を行う。特徴点追跡が失敗した場合には、次フレームよりキーフレームマッチングに移行する。

3.3 ベンチマークの指標

本実験では、ベンチマークの指標として、カメラ位置、カメラ姿勢、仮想物体の投影誤差、の 3 つを取り上げた。以下

表 1 データセットの仕様

フレーム数	349 [枚]
時間長	約10.5 [秒]
画像解像度	640×480 [画素]
カメラ内部パラメータ	$C_x = 320, C_y = 240, F_x = 240, F_y = 240$

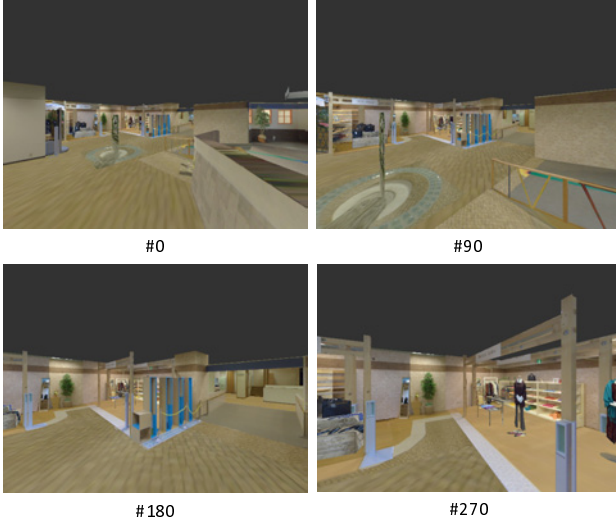


図 4 キーフレーム画像

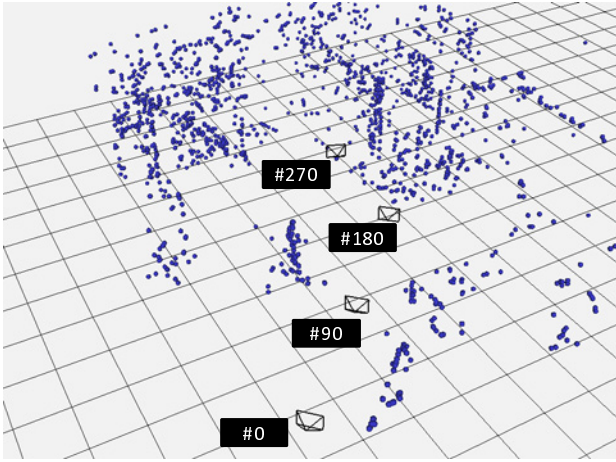


図 5 SURF によって検出された特徴点の 3 次元位置を可視化した様子（錐台は、キーフレームのカメラ位置姿勢を示す）

では、各々の指標についてそれぞれ述べる。

3.3.1 カメラ位置

カメラ位置を 3 次元ベクトル T で定義し、カメラ位置の真値を T_g 、推定値を T_e として、誤差 E_T を 2 つの位置の間のユークリッド距離として、以下のように定義した。

$$E_T = ||T_g - T_e|| \quad (1)$$

3.3.2 カメラ姿勢

カメラ姿勢を、世界座標系からカメラ座標系への変換行列として 3×3 の回転行列 R で定義し、カメラ姿勢の真値を R_g 、推定値を R_e として、2 種類 (A, B) の評価指標を定義した。以下に、それぞれの評価指標について述べる。

A: オイラー角を用いた誤差

文献 [12] と同様に、回転行列 R から Z-X-Y 系のオイラー角を 3 つ算出し、それぞれの角度の差の絶対値を誤差として定義した。図 6 に、オイラー角の算出に用いた式を示す。ただし、各軸周りのオイラー角を、 r_x, r_y, r_z 、回転行列 R の i 行 j 列の成分を、 m_{ij} とする。

$m_{32} = 1$ のとき	$m_{32} = -1$ のとき	$m_{32} \neq \pm 1$ のとき
$r_x = 90$	$r_x = -90$	$r_x = \arctan(m_{32})$
$r_y = 0$	$r_y = 0$	$r_y = \arctan(-\frac{m_{31}}{m_{33}})$
$r_z = \arctan(\frac{m_{21}}{m_{11}})$	$r_z = \arctan(\frac{m_{21}}{m_{11}})$	$r_z = \arctan(-\frac{m_{12}}{m_{22}})$

図 6 オイラー角の算出式

B: Difference Matrix を用いた誤差

文献 [13] と同様に、Difference Matrix : R_d を用いて算出される角度 θ_{R_d} を誤差として定義した。以下に、 R_d 、及び θ_{R_d} の式を示す。

$$R_d = R_g R_e^T \quad (2)$$

$$\theta_{R_d} = \arccos((\text{tr}(R_d) - 1)/2) \quad (3)$$

ただし、 R_e^T は R_e の転置行列、 $\text{tr}(R_d)$ は R_d の対角成分の和である。

3.3.3 点の投影誤差

推定したカメラ位置姿勢を AR や MR に適用し、仮想物体を描画する場合には、仮想物体を描画した際の位置ずれの大きさが重要な指標となる。そこで、3 次元空間中に点を設定し、真値、及び推定値を用いて点を投影した際の 2 次元位置のずれの大きさを誤差として定義した。本実験では、図 7 に示すように、カメラ位置姿勢の真値によって出来る投影面の前方に、投影面に平行な仮想平面を考え、仮想平面上に 9 つの点 (A, B, C, ..., I) を設定して、投影誤差の評価を行った。図 8 に、真値のカメラ位置姿勢を基準とした座標系における、仮想平面上の点の座標を示す。ただし、 a は、カメラ位置から仮想平面までの距離であり、 C_x, C_y, F_x, F_y は、表 1 に示したカメラの内部パラメータである。

4. トラッキング手法のベンチマーキング

ここでは、3 節で述べた内容で行ったベンチマーキングの結果を示し、考察を述べる。

4.1 カメラ位置の誤差

図 9 に、カメラ位置の誤差を示す。本結果から、第 200 フレームを超えた辺りからの誤差が発散していることが見て取れる。一方で、誤差の値のみでは、大局的な誤差の特徴は掴みにくい。実際に、図 10 に示したように、カメラ位置の真値と推定値を可視化してみた所、主に高度と奥行き方向のずれが増加していく様子が観察された。

4.2 カメラ姿勢の誤差

図 11 に、オイラー角を用いて算出した姿勢の誤差、図 12 に、Difference Matrix を用いて算出した姿勢の誤差をそれぞれ示す。オイラー角を用いて算出した誤差の場合、誤差が 3 つの値で分離して表現されるため、各々の誤差を観察しても、実際の誤差傾向を把握することが困難である。一方で、Difference Matrix を用いて算出した誤差の場合、カメラの向きの差を単一の値で観察できるため、オイラー角の場合と比較して、誤差の直観的な把握がしやすいと言える。

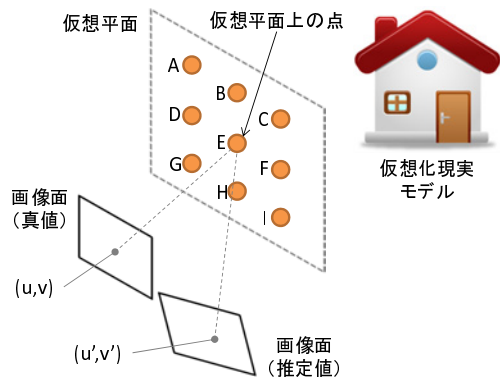


図 7 仮想平面上の点の投影誤差

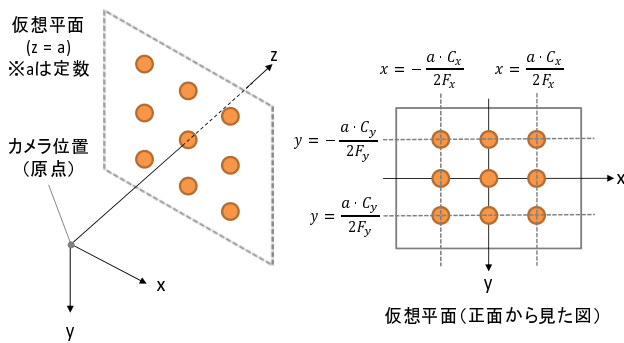


図 8 仮想平面上に設定した点の座標

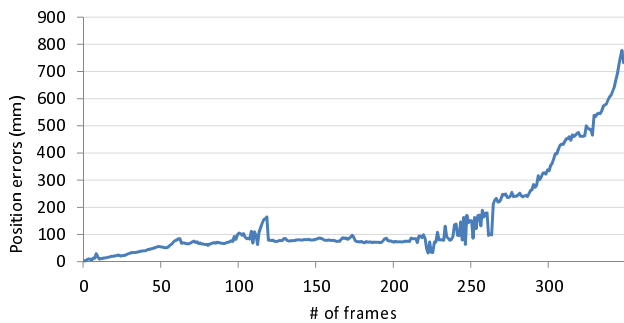


図 9 カメラ位置の誤差

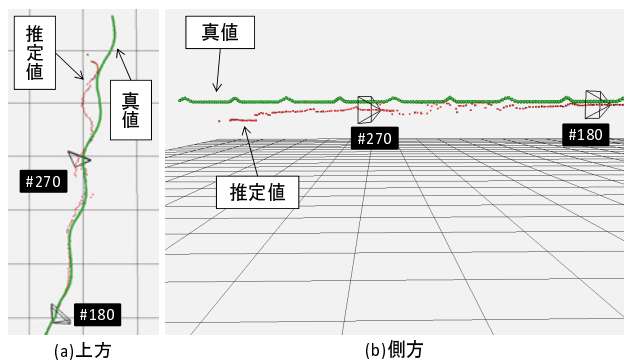


図 10 カメラ位置の真値と推定値の可視化結果

4.3 点の投影誤差

3.3.3 節で述べた手法により、点の投影誤差を算出した。本実験では、カメラ位置から仮想平面までの距離 a を、以下の 4 パターン (1000mm, 3000mm, 5000mm, 7000mm) で設定した。図 13 に各パターンでの投影誤差を示す。また、

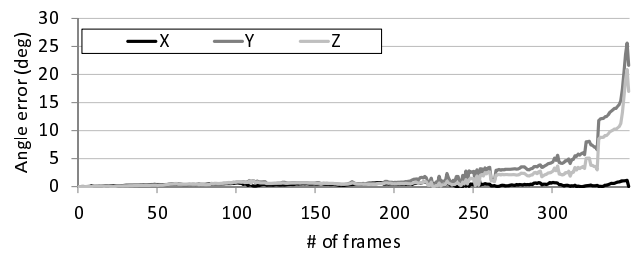


図 11 オイラー角を用いて算出したカメラ姿勢の誤差

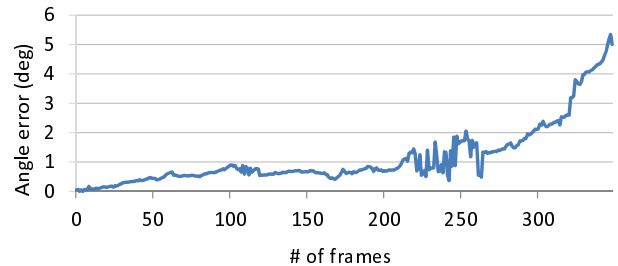


図 12 Difference Matrix を用いて算出したカメラ姿勢の誤差

各パターンでの投影誤差の最大値を表 2 に示す。点の投影誤差の全体的な特徴としては、第 250 フレーム近辺より投影誤差が大きく増加していく傾向があった。各パターンでの投影誤差の最大値については、 $a = 1000$ の場合には、投影誤差の最大値は 132.7[pix] であった。また、 a を大きくした場合、表 2 に示すように、投影誤差の最大値は減少傾向にあった。このような、仮想平面上の点の投影誤差の評価を行うことで、AR や MR を想定したカメラトラッキング手法のベンチマーキングは直観的、かつ定量的に実施できる。一般に、仮想物体の描画に対する要求精度はアプリケーションの内容に依存するが、例えば、“仮想物体の存在位置がカメラから 3000mm 離れた面上、投影誤差の要求精度が 20[pix] 以内”、といった条件を考えた場合、本実験で用いたトラッキング手法では、第 271 フレームまで条件を満たしている、といった評価が出来る。

また、同一仮想平面上において、点の位置の違いによる投影誤差のばらつきが見られた。表 3 に、9 点の投影誤差の最大値と最小値の差を 100 フレーム毎に調査した結果を示す。このようなばらつきが観測されたことで、同一のカメラ位置姿勢の推定値であっても、仮想物体の位置によってその評価にはばらつきがあることが分かった。したがって、仮想物体の位置を事前設定できるアプリケーションを対象とした場合、仮想物体の位置に基づいてベンチマーキングを行うことで、最適なトラッキング手法の選択、トラッキング手法のパラメータ等の調整、などが可能であると考えられる。

5. ま と め

本稿では、カメラトラッキング手法のベンチマーキングをサポートするフレームワーク構築のための、仮想化現実モデルを用いたベンチマークスイートを提案した。支援ツール拡張として歩行モーションモデルを導入する機能を試作した結果、長さが 1 分以内程度のデータセットが 10 分程度の作業

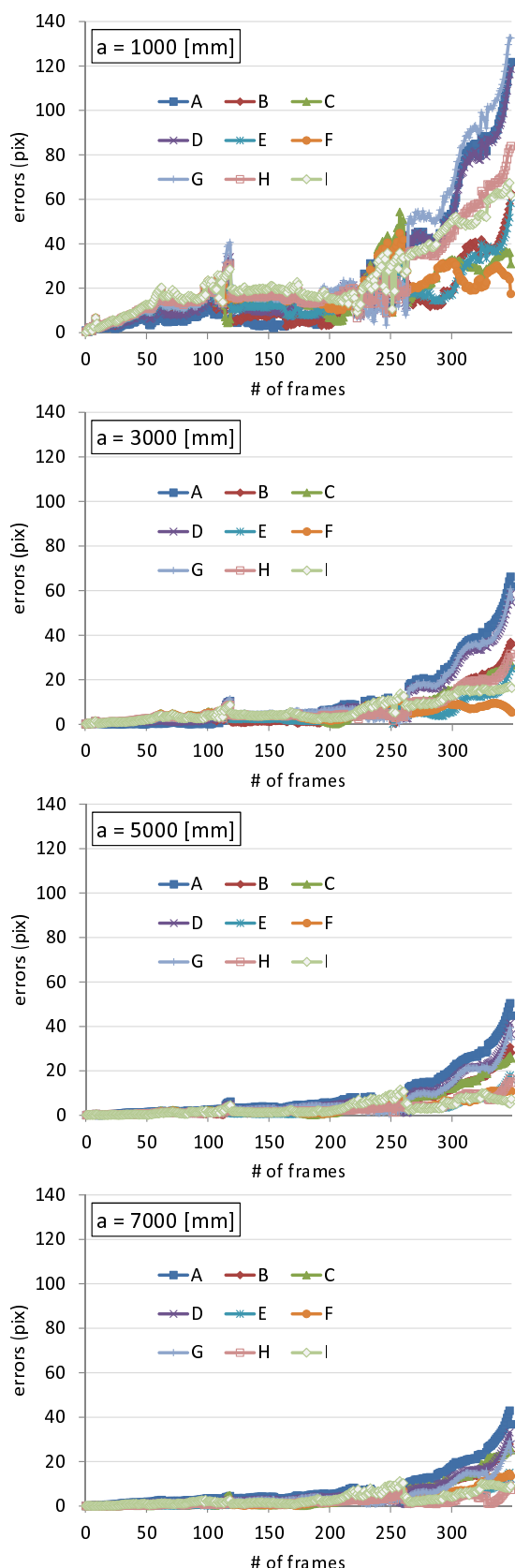


図 13 仮想平面上の点の投影誤差

時間で生成できた。また、位置誤差、姿勢誤差、点の投影誤差の 3 種類の指標を用いてトラッキング手法のベンチマーキング結果を示した。今後は、位置誤差、Difference Matrix を用いた姿勢誤差、点の投影誤差の評価機能や、推定結果の

可視化機能を支援ツールに組み込むための実装を進める予定である。また、ベンチマーキング結果や、新たに登録されたデータセットを共有する枠組みの整備を行う。

謝辞

本研究は、JST 戦略的国際科学技術協力推進事業（研究交流型）「日本 - フランス（ANR）研究交流」として実施されました。

文 献

- [1] I. Austvoll: "A Study of the Yosemite Sequence Used as a Test Sequence for Estimation of Optical Flow", In SCIA2005, pp.659-668, 2005.
- [2] S. Lieberknecht, et al: "A Dataset and Evaluation Methodology for Template-based Tracking Algorithms", In ISMAR2009, pp.145-151, 2009.
- [3] H. Tamura, et al: "Proposal of International Voluntary Activities on Establishing Benchmark Test Schemes for AR/MR Geometric Registration and Tracking Methods", In ISMAR2009, pp.233-239, 2009.
- [4] TrakMark WEB
<http://trakmark.net/>
- [5] K. Makita, et al: "A prototype tool in TrakMark to generate benchmark data sets for visual tracking methods using virtualized reality models", In Proc. of TrakMark2011, 2011.
- [6] E. Hirasaki, et al: "Effects of walking velocity on vertical head and body movements during locomotion", In Exp. Brain Res., vol.127, pp.117-130, 1999.
- [7] 大隈ら: "シームレスな館内ナビゲーションと AR 展示を実現するユーザ測位システム", 電子情報通信学会技術研究報告: 112(25), pp.23-28, 2012.
- [8] H. Bay, et al: "Surf: Speeded up robust features", In Proc. of ECCV2006, pp.404-417, 2006.
- [9] M. A. Fischler, et al: "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography", Comm. of the ACM 24 (6), pp.381-395, 1981.
- [10] C. Tomashi, et al: "Detection and tracking of point features", Technical Report CMU-CS-91-132, 1991.
- [11] B. D. Lucas, et al: "An iterative image registration technique with an application to stereo vision", In Proc. of the 7th Int. Joint Conf. on Artificial Intelligence, pp.674-679, 1981.
- [12] M. Hayashi, et al: "A Study of Camera Tracking Evaluation on TrakMark Data-Set", In Proc. of TrakMark2011, 2011.
- [13] A. Petit, et al: "Evaluation of Model based Tracking with TrakMark Dataset", In Proc. of TrakMark2011, 2011.

表 2 投影誤差の最大値 [pixel]

a	1000	3000	5000	7000
Error(Max)	132.7	66.2	50.4	41.4

表 3 投影誤差の最大値と最小値の差 [pixel]

# of frames a	0	100	200	300
1000	0.86	11.4	14.7	47.2
3000	0.19	5.39	5.81	21.1
5000	0.08	1.22	4.40	16.1
7000	0.11	3.19	3.58	15.6