

ホモグラフィに着目した画像特徴点に基づく 位置・姿勢のマーカレス計測に関する研究

A Study on Marker-less Position and Orientation Measurements Based on Image Feature Points Focusing on Homography

12D5202 水地 良明

指導教員 崔 龍雲

ABSTRACT

This thesis describes a vision-based marker-less approach for the real-time measurement of the position and orientation relative to a target object. Such marker-less measurement without the installation of transmitter-receivers or artificial markers is an optimal solution to expand the application of position and orientation measurements. The proposed approach is focused on a homography estimated from image feature points for real-time position and orientation measurements of a target which has poor texture or redundant features. As an application for measuring the position and orientation of a target which has poor texture and shape changes, a monocular camera-based measurement of a palm is described. As the other application for measuring the position and orientation of a target which has redundant features, a stereo camera-based measurement for relative vessel positioning without any installation on a target is also described. Results of these applications will demonstrate that the proposal is applicable and useful for the real-time position and orientation measurement of a target which has poor texture or redundant features.

Keywords: Marker-less Measurement, Position and Orientation, Image Feature Point, Homography

1 序論

対象との位置・姿勢の実時間計測を要する応用が多岐の分野にわたって展開されている。中でも、画像センサを用いたマーカレス計測は、計測対象への受信機等の付与やインフラの設置が困難な場面での応用に有用な方法の一つである。画像に基づく位置・姿勢の計測では、事前に用意した三次元モデルと入力画像、もしくは異なる位置で撮影された複数画像間の共通する特徴を検出し、それらの対応から位置・姿勢を推定する方法が一般的である。画像特徴点を用いた位置・姿勢の推定は、照合に多くの繰り返し計算を要する領域やエッジに基づく方法に比べ、計算コストを低減しやすいため、実時間計測に適した方法と言える。一方で、解を得るために位置関係の変化しない複数の対応点が必要であるため、対象物の形状や見かけの変化、冗長的な特徴に依存する誤対応の影響を受けやすいことが特徴として挙げられる。

そこで本研究では、これらの影響を低減する方法の一つとして、対象を平面とみなすよう制約を設け、ホモグラフィ（平面射影変換）に着目することで対象物との位置・姿勢を計測するアプローチを試みる。本研究では、まず、単眼カメラを用いて形状変化を伴う手のひらの位置・姿勢をマーカレスで計測する応用について述べる。明瞭なテクスチャを持たず、形状変化を伴う人間の手は、特徴点に基づく位置・姿勢のマーカレス計測が難しい対象の一つである。シルエットやエッジに基づいて逐次的に手の形状と位置・姿勢を解析する方法^[1]も提案されているが、計算コストが膨大であるため、実時間計測への適用は難しい。また、単眼カメラを用いた位置・姿勢の計測では計測対象の三次元モデルが必要となるため、利用者毎に異なる手の三次元モデルを簡便に獲得する方法の実現も課題となる。そこで、手の形状変化に関わる指の動作を平面的に限定することで、形状変化を伴う手のひらの三次元モデルの獲得および位置・姿勢のマーカレス計測を実現する。これにより、処理能力の乏しいウェアラブルデバイスを利用した AR における、仮想情報とのインタラクションを実現するためのインタフェースとしての応用を可能とする。次に、大型船の位置・姿勢をマーカレス計測するシステムについて述べる。船舶のような大型対象の位置・姿勢の計測を局所領域のみから高精度に計測することは難しいため、複数の観測点を設け、各観測点に対応する計測器間の基線長を十分に確保することが望ましい。また、移動する船体の位置・姿

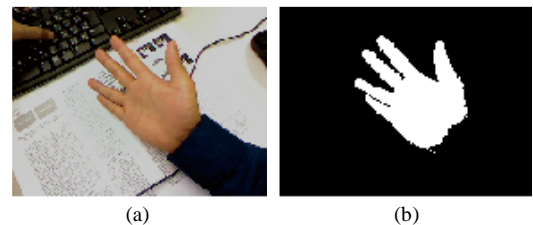


Fig.1 Hand segmentation: (a) a captured RGB image, and (b) the extracted hand region excluding low brightness region.

勢を得るためには、各計測器を用いて計測対象となる計測点を常に追従しながら計測点との距離と方向を計測する必要がある。このような計測点の追従や距離・方向の計測を実現するためには、複数のカメラから得られる画像間の対応点を検出し、視差を得る必要がある。しかしながら、埠頭や浮体プラント、船舶など、実環境で想定される対象は冗長的な特徴を多く含むため、多数の特徴の誤対応により、船位保持に要求される精度で距離・方向の計測を実現するための正確な視差を得ることは難しい。そこで、計測対象を平面的に捉え、画像間のホモグラフィから視差を推定することにより、任意に指定した計測点の追従と平行等位ステレオの原理に基づく距離・方向の算出を可能とする。これにより、マーカレスでの船位計測を実現する。

本論文では、上記 2 つの応用とそれらの評価実験の結果から、ホモグラフィに着目した画像特徴点に基づく位置・姿勢のマーカレス計測の応用における有用性について明らかにする。

2 手のひらの位置・姿勢のマーカレス計測

ハードウェア技術の発展に伴い、小型計算機とカメラ、ヘッドマウントディスプレイを組み合わせたウェアラブルデバイスをを用いた AR^[2]が実現されている。さらに、カード状の人工マーカ^[3]を利用することで、表示された仮想情報とのインタラクションを可能とする方法^[4]も提案されている。しかしながら、インタラクションに利用する場合、人工マーカの持ち運びや装着はウェアラブルデバイスの利便性を低下させてしまう。この解決方法として、マーカレス計測した手の位置・姿勢を用いたインタフェースを実現する。

2.1 手のひらの位置・姿勢計測の概要

提案する手のひらの位置・姿勢の計測手法では、手の姿

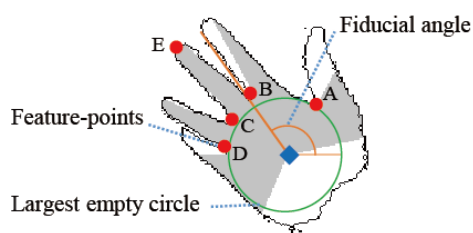


Fig.2 Overview of object-points detection.

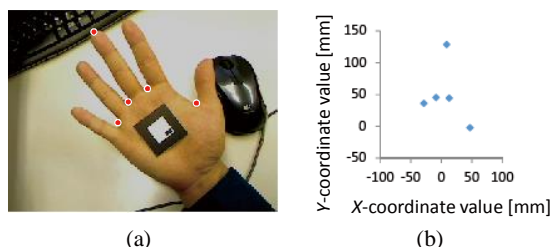


Fig.3 Measurement of a hand model: (a) an input image with detected object-points and a pasted AR marker, and (b) an obtained hand-model.

勢やスケールに依存せず手領域を抽出するため、色情報を利用して Fig. 1(a)の入力画像から Fig. 1(b)のように手領域を抽出する。ここでは、手領域内の相対的に輝度の低い領域を除外することで、各指が分離された手領域を抽出している。

得られた手領域から輪郭を抽出し、Fig. 2 のようにに影響を受けにくい特徴点 A~E を検出する。ここでは、手領域内における最大空円を検出し、その中心を基準とする相対的な輪郭点の分布から基準角を決定する。この基準角を用いて、Fig.2 の手領域内の影付き部分のように、最大空円の中心から再近傍の手領域輪郭画素までの距離を得る。この距離の変化に基づいて、指の間の凹部を特徴点 A~D として検出する。さらに、点 B, C 間に存在する輪郭画素において、最も距離の大きい画素を特徴点 E として検出する。

単眼カメラを用いて位置・姿勢を推定するためには、特徴点の三次元的な位置関係が既知でなくてはならない。本手法では、カード状のマーカを用いた簡易的な方法により、オフラインで一度のみ特徴点の位置関係を手のモデルとして計測する。この処理では、まず、Fig.3(a)のように手のひらの中央にカード状の AR マーカを載せ、手全体が画面内に収まるように画像を取得する。AR マーカは手のひら平面上に乗せられているため、特徴点は AR マーカの三次元平面上に存在するとみなすことができる。これにより、カメラの焦点・手の特徴点間を結ぶ直線とマーカ平面の交点を算出することで、Fig.3(b)のような手のモデルを得る。

事前に取得した手のモデルと入力画像中の特徴点の対応から手のひらの位置・姿勢を推定する。特徴点は手のひら平面上に存在するため、手のひら平面と画像平面間の位置・姿勢を表すホモグラフィ行列は自由度が 8 となり、4 組以上の対応点からカメラの内部パラメータに基づいて計算することが可能となる。手のモデルの特徴点と入力画像中の特徴点の 5 点の対応から投影誤差が最小となるホモグラフィ行列を算出することで、手のひらの位置・姿勢計測を得る。

2.2 手のひらの位置・姿勢計測の評価実験

手の位置・姿勢および指の間隔の変化を含む状況下における、提案手法により計測された位置・姿勢に基づくオブジェクト描画の精度を評価する実験を行った。実験では、まず、手の位置・姿勢および指の間隔の変化を含む手の入

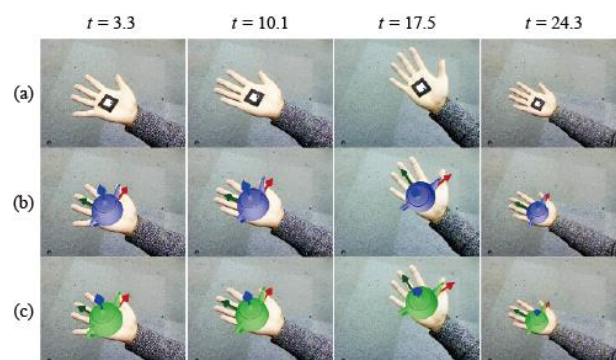


Fig.4 Qualitative comparison between displayed objects based on the postures of the AR marker and the palm: (a) captured images, (b) displayed objects based on the posture of the AR marker, and (c) displayed objects based on the posture of the palm

力画像列を準備し、提案手法による手のひらの位置・姿勢の推定を行った。次に、ARToolKit^[3]を利用して手のひらに貼付された AR マーカの位置・姿勢を推定した。最後に、両手法により推定された位置・姿勢に基づいて三次元オブジェクトを描画し、その見かけの差から、オブジェクトの描画における提案手法の精度を評価した。

Fig. 4 に両手法によって推定された位置・姿勢に基づいて三次元オブジェクトを描画した実験の結果の一部を示す。Fig. 4(a) が実際に実験で使用した入力画像列の一部である。これらの入力画像は、USB カメラ(Webcam C500, Logitech 社製)を用いて、320×240[pixel]の解像度で、 $t=0.0$ から $t=30.0$ まで 30 秒間取得されたものである。このとき、手全体が常にカメラの視野内に収まる範囲で、隣接する指の間隔を広げたり、狭めたりしながら、手の位置・姿勢を変化させた。人の手のひら上に貼付した一辺 40mm の AR マーカから推定された位置・姿勢に基づいて、三次元オブジェクトを表示したものが、Fig. 4(b) である。同図では、オブジェクトとして位置・姿勢を判別しやすい座標軸とティーポットを描画している。また、提案手法により推定された手のひらの位置・姿勢に基づいて三次元オブジェクトを表示した結果が、Fig. 4(c) である。 $t=3.3$ と $t=17.5$ の結果を比較すると、ほぼ同じ指の動きを伴う手に対して、手の姿勢のみが変化している。また、 $t=10.1$ と $t=24.3$ の結果を比較すると、手の位置が大きく変化しているが、いずれの結果においても手のひらに基づいて AR マーカとほぼ同じ位置・姿勢でオブジェクトを描画できている。同様に、 $t=3.3$ と $t=10.1$ の結果では、指の間隔が異なっているが、AR マーカと手のひらに基づいて描画されたオブジェクトの位置・姿勢に大きな差は見られない。これにより、指の動きに依存することなく手のひらの位置・姿勢を推定し、オブジェクトを重畳表示可能であることが示された。

3 船舶の位置・姿勢のマーカレス計測

船位保持における人的要因による事故の軽減や作業効率の向上を目的とした船位計測では、一般的に、GPS^[5]やレーダ^[6]、レーザ^[7]が用いられる。しかしながら、これらの計測ではいずれも、対象側に受信機や反射板が設置されていることを前提としており、それらが設置されていない対象物との相対的な位置・姿勢を計測することはできない。船舶を対象とした画像に基づくマーカレス計測の従来研究では、自動航行のための障害物検出や衝突回避^[8]は実現されているものの、船位保持を目的とした位置・姿勢計測については言及されていない。そこで、複数のカメラと電動雲台により構成された 2 つのカメラユニットを用い、対象物上の任意計測点に対する計測対象と船体の位置・姿勢

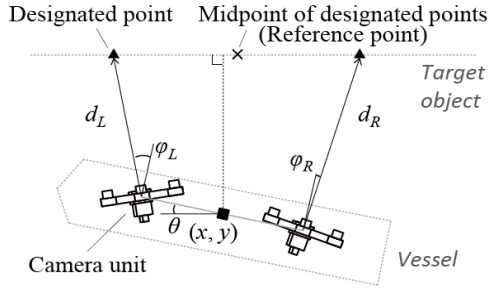


Fig. 5 Overview of the proposed measurement system for vessel positioning.

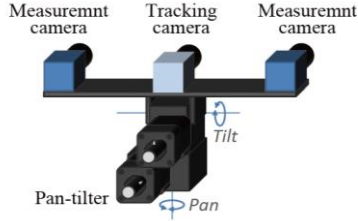


Fig. 6 Configuration of the camera unit.

をマーカレス計測するシステムを実現する。

3.1 計測システムの構成と位置・姿勢計測の概要

Fig. 5 に提案する船位計測システムの概要を示す。本システムは、船体上に設置された2つのカメラユニットから構成される。計測の開始時に、各カメラユニットで任意の計測点を指定し、規範画像を取得する。この規範画像の中心を指定点として、カメラユニットが追跡しながら距離 d_L , d_R と角度 ϕ_L , ϕ_R を計測する。これらの距離と角度から、2つの計測点の midpoint を基準とする船位 (x, y) と姿勢 θ を得る。

Fig. 6 にカメラユニットの構成を示す。カメラユニットは1つの追跡用カメラ、2つの計測用カメラ、電動雲台で構成されている。追跡用カメラは電動雲台のパン軸上に設置されており、広い画角を持っている。計測用カメラは、光軸が平行になるように電動雲台パン軸から等間隔で左右に設置されている。追跡用のカメラにより検出された指定点に基づいて、電動雲台を制御することにより、指定点を補足しながら、連続した計測を可能としている。

指定点を追跡し続けるために、規範画像と現在画像の位置ずれを画像特徴点に基づいて算出する。近年では、多くの特徴量の検出・記述手法が提案されているが、本システムでは、照合における頑健性と計算速度の観点から GPU を用いた SURF^[9] による特徴点照合を採用する。Fig. 7(a) が規範画像から検出された特徴点、Fig. 7(b) が少し異なる視点で取得された画像から検出された特徴点である。また、図中の三角が指定点である。得られた特徴点を照合することで、Fig. 7(c) のような対応点が見られる。ここでは、各線分が2画像間の特徴点を対応を表している。しかしながら、得られた対応点には、冗長的な特徴の存在などにより、多くの誤対応が含まれている。そこで、画像上における対応点同士のユークリッド距離を投票処理することで、Fig. 7(d) のように誤対応を除去する。さらに、得られた対応点を用いて2画像間のホモグラフィを算出する。この時、RANSAC^[10] を適用することで、外れ値を除外してホモグラフィを算出し、Fig. 7(d) 中の矢印のような画像のずれを指定点の並進量として得る。画像内で検出された指定点の並進量 (u, v) に基づいて、カメラの光軸と指定点のなす水平角 ψ_{cx} を次式により計算する。

$$\psi_{cx} = \tan^{-1} \frac{u s_u}{f} \quad (1)$$

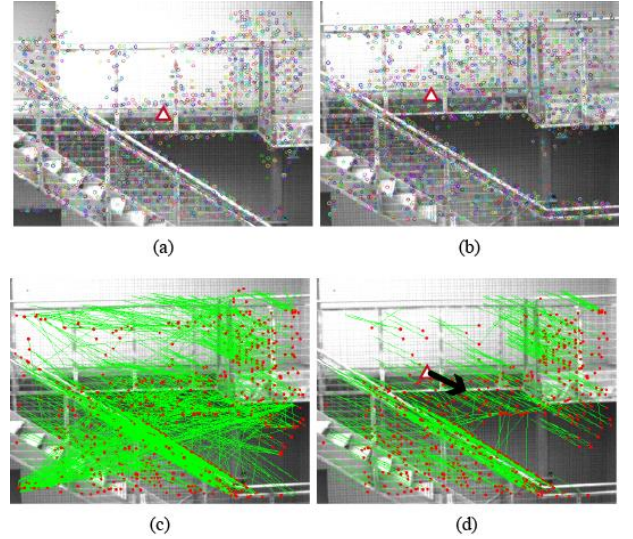


Fig. 7 Estimation of displacement between two image frames: (a) reference image and detected feature points using SURF, (b) detected feature points from an image captured at a slightly different position, (c) matching results of feature points, and (d) matching results without some outliers and estimated translation.

ここで、 s_u は画像ピクセルの横方向のサイズ、 f はレンズの焦点距離である。垂直角 ψ_{cy} についても同様の式から算出する。これらの角 ψ_{cx} , ψ_{cy} が最小となるように電動雲台を制御することで指定点を追跡する。

同様の手順により、左右の計測用カメラ間のホモグラフィを算出することで、各計測用カメラの光軸に対する水平角 ψ_{Lx} , ψ_{Rx} を得る。計測用カメラ間の距離 L_C と置くと、カメラユニットの中心から指定点までの距離 d は次のように得られる。

$$d = \frac{1}{2} \left(\frac{L_C}{\tan \psi_{Lx}} + \frac{L_C}{\tan \psi_{Rx}} \right) \quad (2)$$

また、追跡用カメラから得られた水平角 ψ_{cx} と電動雲台のパン角 ξ_p から指定点の方向角 ϕ を次のように得る。

$$\phi = \psi_{cx} + \xi_p. \quad (3)$$

Fig. 8 に船体の位置・姿勢の導出の計測における幾何学的配置を示す。指定点の midpoint (x_0, y_0) を基準とする船体の位置 (x, y) および姿勢 θ をカメラユニットから得られた距離 d_L , d_R と水平角 ϕ_L , ϕ_R から計算する。これらを用いて、指定点の位置 (x_{TL}, y_{TL}) , (x_{TR}, y_{TR}) を次式により決定する。

$$x_{TL} = -\frac{L_U}{2} + d_L \sin \phi_L, \quad y_{TL} = d_L \cos \phi_L \quad (4)$$

$$x_{TR} = \frac{L_U}{2} + d_R \sin \phi_R, \quad y_{TR} = d_R \cos \phi_R \quad (5)$$

ここで、 L_U はカメラユニット間の距離である。また、姿勢 θ を次式により決定する。

$$\theta = -\tan^{-1} \left(\frac{y_{TR} - y_{TL}}{x_{TR} - x_{TL}} \right) \quad (6)$$

基準点 (x_0, y_0) は次のように得られる。

$$x_0 = \frac{x_{TL} + x_{TR}}{2}, \quad y_0 = \frac{y_{TL} + y_{TR}}{2} \quad (7)$$

基準点に対するカメラユニットの位置 (x_{TL}, y_{TL}) , (x_{TR}, y_{TR}) を次式により決定する。

$$\begin{aligned} x_{CL} &= \left(-\frac{L_U}{2} - x_0 \right) \cos \theta + y_0 \sin \theta \\ y_{CL} &= \left(-\frac{L_U}{2} - x_0 \right) \sin \theta + y_0 \cos \theta \end{aligned} \quad (8)$$

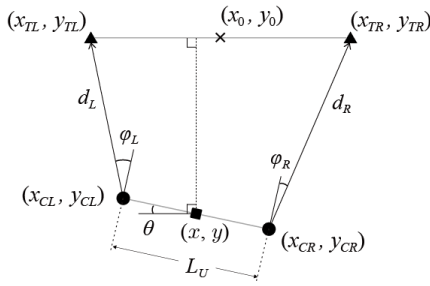


Fig. 8 Geometry for the position and heading.

$$\begin{aligned} x_{CR} &= \left(\frac{L_U}{2} - x_0 \right) \cos \theta + y_0 \sin \theta \\ y_{CR} &= \left(\frac{L_U}{2} - x_0 \right) \sin \theta + y_0 \cos \theta \end{aligned} \quad (9)$$

これにより、船体の位置 (x, y) を次式により決定する。

$$x = \frac{x_{CL} + x_{CR}}{2}, \quad y = \frac{y_{CL} + y_{CR}}{2} \quad (10)$$

3.2 位置・姿勢計測における本システムの性能評価

本システムが船位保持に応用可能な精度を持つか検証するため、モデル船とプロトタイプとなる計測システムを構築し、計測精度を評価した。Fig. 9に評価実験の条件を示す。実験では、全長 2m、幅 1m のモデル船上にカメラユニットを 1316mm 間隔で設置し、計測を行った。カメラユニットには、7.5 μ m のピクセルサイズ、640 \times 480 の解像度を持つカメラを利用した。また、計測用カメラには 16mm、追跡用カメラには 8mm の焦点距離を持つレンズを利用した。計測用カメラの間隔は 300mm とした。計測の開始時には、約 7m の距離にある壁面上の点 A、B を指定した。さらに、船体の実際の動作を得るため、チェスボードパターンと慣性センサ(Xsens MTi-10)をモデル船上に設置した。天井に設置したカメラを用いてチェスボードパターンから 6 自由度の動作を計測^[12]し、評価の基準とした。実験では、全長方向を軸とするロール回転を船体に与えながら、Fig. 9 中に示す x 方向、 y 方向、回転 θ にそれぞれ動作させた。

慣性センサにより得られた実験中の船体のロール回転は、 x 方向への移動時が $\pm 2.7^\circ$ 、 y 方向への移動時が $\pm 2.5^\circ$ 、回転時が $\pm 2.2^\circ$ であった。また、それぞれの周期は順に、1.50-1.72 秒、1.38-1.67 秒、1.43-1.53 秒であった。この時に実験結果として得られた平均誤差と標準偏差を Table 1 に示す。 x 方向および y 方向の平均誤差はいずれも、最も精度の要求される接岸作業における船位計測に必要とされる 0.1m を十分に下回っている。また、各標準偏差についても、0.05m 未満となっており安定した計測結果が得られていることがわかる。姿勢 θ については、平均誤差 0.58° 、標準偏差 0.39° であった。接岸作業等では 0.1° の計測精度が望ましい。しかしながら、本実験におけるカメラユニット間の基線長は約 1.3m となっており、実際の船舶に比べ極端に小さい。回転の計測精度はカメラユニットの基線長に依存して向上するため、実際の船舶上ではこの誤差は数十分の一程度になることが予想される。これらの結果から、提案する計測システムは、船体の位置・姿勢計測において、船位保持に応用しうる十分な精度を持っていることが示された。

4 結論

本論文では、画像特徴点に基づく位置・姿勢のマーカレス計測のアプローチとしてホモグラフィに着目し、2つの応用について述べた。まず、単眼カメラを用いて形状変化を伴う手のひらの位置・姿勢をマーカレス計測する手法に

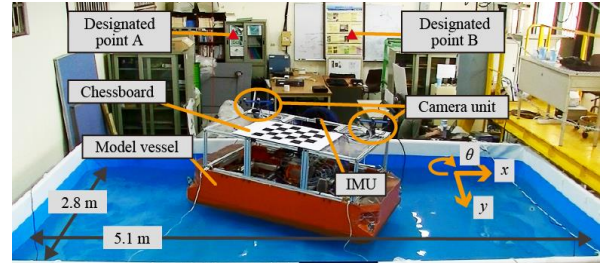


Fig. 9 Condition of the experiment in water.

Table 1 Average errors and standard deviation of the measured positions and headings under the condition with roll motion of the model vessel.

Parameter	Average error	Standard deviation
x	0.053 m	0.037 m
y	0.027 m	0.024 m
θ	0.58°	0.39°

ついて述べ、提案手法を用いて仮想情報を重量表示する評価実験により、その有用性を示した。次に、平行等位ステレオの原理に基づくマーカレスでの船位計測システムについて述べ、モデル船に実装したプロトタイプシステムを用いた評価実験により、船位保持を目的とした位置・姿勢の計測における提案システムの有用性を示した。これらの応用では、計測対象を平面的にのみなすよう制約を設け、ホモグラフィを推定することで、画像特徴点に基づく位置・姿勢計測における形状変化を伴う対象の効率的な計測や誤対応を多く含む対象に対する高精度な計測の課題解決を実現しており、画像特徴点に基づくホモグラフィに着目したアプローチの有用性を明らかにしている。また、これらの成果は、単眼カメラを用いたモデルベースの位置・姿勢計測およびステレオビジョンに基づく位置・姿勢計測における本アプローチに基づく位置・姿勢のマーカレス計測の応用の可能性をそれぞれ示している。

参考文献

- [1] M. de La Gorce, D. J. Fleet, and N. Paragios: "Model-based 3D Hand Pose Estimation from Monocular Video," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 33, No. 9, pp. 1793-1805, 2011.
- [2] S. Henderson and S. Feiner, "Exploring the Benefits of Augmented Reality Documentation for Maintenance and Repair," IEEE Trans. on Visualization and Computer Graphics, Vol. 17, No. 10, pp. 1355-1368, 2011.
- [3] H. Kato and M. Billinghurst, "Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System," Proc. of the International Workshop on Augmented Reality, pp. 85-94, 1999.
- [4] B. H. Thomas, and W. Piekarski, "Glove Based User Interaction Techniques for Augmented Reality in an Outdoor Environment," Virtual Reality, Vol. 6, No. 3, pp. 167-180, 2002.
- [5] M. Ueno, "A GPS-based System for Precise Shipping Guidance and Control," J. marine science and technology, Vol. 5, No. 1, pp. 9-15, 2000.
- [6] Kongsberg, "RADius," <http://www.km.kongsberg.com/>, 2015.
- [7] Marine Technologies, LLC, "CyScan," <http://www.marine-technologies.com/>, 2015.
- [8] T. Huntsberger, H. Aghazarian, A. Howard, D. C. Trotz, "Stereo Vision-based Navigation for Autonomous Surface Vessels," J. Field Robotics, Vol. 28, No. 1, pp. 3-18, 2011.
- [9] H. Bay, T. Tuytelaars, L. Van Gool, "SURF: Speeded-up robust features," Computer Vision and Image Understanding, Vol. 110, No. 3, pp. 346-359, 2008.
- [10] M. A. Fischler, R. C. Bolles, "Random Sample Consensus: a Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," Communications of the ACM, Vol. 24, No. 6, pp. 381-395, 1981.
- [11] Z. Zhang, "A Flexible New Technique for Camera Calibration," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 22, No. 11, pp. 1330-1334, 2000.