

リアルタイム遮蔽物透視システムのための 微小な位置ズレのある RGB-D 画像対の位置合わせ

Registration of RGB-D Image Pair with Small Displacement for Real-time See-through System

○学 橘高 達也 (東京大学) 正 藤井 浩光 (東京大学)
正 山下 淳 (東京大学) 正 浅間 一 (東京大学)

Tatsuya KITAKA, The University of Tokyo, kittaka@robot.t.u-tokyo.ac.jp
Hiromitsu FUJII, The University of Tokyo
Atsushi YAMASHITA, The University of Tokyo
Hajime ASAMA, The University of Tokyo

This paper presents a method of real-time registration of RGB-D image pair. For remote operation of robots in dangerous situations such as disaster sites, providing robot operators with appropriate visual information is important. So far, we have proposed the “see-through system”, in which multiple images acquired from RGB-D sensors are integrated to present images of objects as if they are seen through obstacles. However, uncertainty of information of sensor positions and postures may cause small displacement between images, which leads to visually inconsistent output images. In this paper, corresponding points are efficiently detected by temporal and spatial tracking, and sensor positions and postures are robustly corrected by iteratively reweighted least squares (IRLS). The experimental results show the ability to significantly decrease image displacement in real time.

Key Words: Computer vision, Diminished reality, RGB-D sensor, Registration

1 緒言

災害現場などの危険な現場でのロボットの遠隔操作において、オペレータへの適切な視覚情報の提示は重要である [1]。作業対象物が遮蔽物に隠れて見えない場合、オペレータが作業対象物を認識することが困難になり、作業効率の低下につながる。

コンピュータビジョンの分野では、映像上で視覚的に不要な物体を隠蔽・消去、もしくは障害となる物体を透過させる技術が提唱されており、隠消現実感 (Diminished Reality, DR) と呼ばれている [2]。筆者らはこれまでに、図 1 のように 2 台の RGB-D センサを用いたリアルタイム遮蔽物透視システムを提案してきた [3]。RGB-D センサは、通常のカメラと同様の RGB 情報に加えて、画像中の各点からセンサまでの距離情報も取得可能なセンサであり、環境の 3 次元的情報が取得可能である。このシステムでは、RGB-D センサをロボットの正面およびアームに取り付け、アームの順運動学によってセンサの位置・姿勢を常に更新する。そして、2 台の RGB-D センサの映像に座標変換を施した上で統合し、透視映像を生成する。

しかし、DR における研究課題として、「高精度な幾何学的整合の達成」が挙げられている [2]。統合する 2 つの映像に位置のズレが存在する場合、図 2 のように生成される透視映像中で物体の輪郭が不連続になる、物体が二重に見えるなどの現象が生じる。これはセンサの位置・姿勢の推定誤差などが原因で発生し、一般的な DR において解決すべき問題の 1 つとなっている [2]。

筆者らの提案システム [3] においてもこの問題が存在する。このシステムでは、事前に行う RGB-D センサのキャリブレーションおよびアームの手先と RGB-D センサの位置関係のキャリブレーション [4] の結果と、現在のアームの関節角情報を用いて、2 台の RGB-D センサの位置関係を導出し、それぞれの RGB-D センサから取得した映像に座標変換を施すことで 2 つ映像の位置合わせを行う。しかし、事前のキャリブレーションを正確に行っても、アームの関節角情報の誤差などが原因で、アームを運動させるにつれて微小な位置のズレが発生していくという問題があった。これに対して、手動での微調整によって映像を正確に位置合わせするという方法も考えられるが、3 次元的な並進・回転

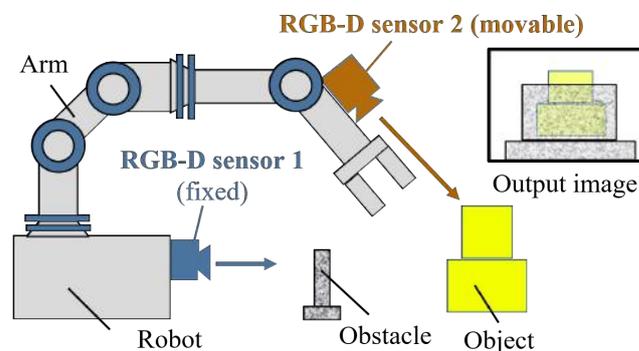


Fig. 1 See-through system [3]

の 6 自由度を考慮する必要があるため、手動で映像を正確に位置合わせするのは困難である。よって、透視映像における位置ズレを自動で補正する手法を構築することが重要である。

森ら [5] は、SURF 特徴量を用いた特徴点マッチングによりホモグラフィ行列を求めることで、透視映像における位置ズレを補正する手法を提案した。しかし、この手法は環境が平面のみで構成されている場合のみ適用可能であり、さらに事前に遮蔽物が存在しない環境の画像を準備しておく必要がある。また、SURF や RANSAC などの計算コストの大きい処理を含むため、リアルタイムでの実行に不向きであるという課題もある。

本研究は災害現場などの未知環境での適用を想定するため、環境の制約や事前情報の有無によらず適用可能な手法を構築する必要がある。また、本研究はロボットの遠隔操作に適用することを想定するため、リアルタイムで補正を行うことが重要であり、計算コストを抑える必要がある。

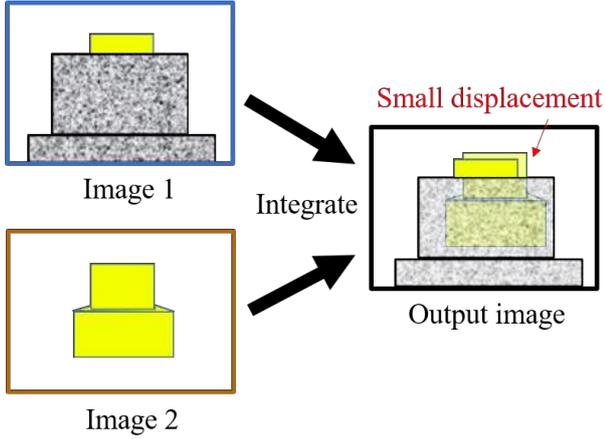


Fig. 2 Problem of displacement in a see-through image

座標変換後の位置のズレに対処するため、本研究では座標変換後の2つの映像から対応点を検出し、その位置の誤差を最小化するようにRGB-Dセンサの外部パラメータをリアルタイムで補正する手法を提案する。以下、使用する2つのRGB-Dセンサをセンサ1およびセンサ2、それぞれから取得された画像に座標変換を施したものを画像1および画像2と呼ぶ。

2 提案手法

提案手法の概要を図3に示す。提案手法は、対応点検出、誤対応点の除去、センサ位置・姿勢補正の3つの処理からなる。

対応点検出では、画像1および画像2からKLT tracker [6]を用いて空間・時間方向に特徴点検出・追跡を行う。誤対応点の除去は、各対応点の位置のズレのばらつきに着目し、マハラノビス距離による外れ値の判定に基づいて行う。センサ位置・姿勢補正は、位置ズレが微小であるという仮定から補正量に関する一次方程式を導出し、IRLS（反復再重みづけ最小二乗法）で解くことを行う。

2.1 対応点検出

センサの3次元的位置・姿勢を補正するためには、3次元空間中での対応点が最低3組得られていれば良い。ここでは、少ない計算コストで少数の対応点を正確に求める手法として、KLT trackerを用いる。KLT trackerは動画における移動体追跡などに多く用いられ、特徴点検出・特徴点追跡の2つの処理からなる。特徴点検出は、ある時刻に撮影された画像から、Harrisオペレータを用いてコーナ点を検出する処理である。特徴点追跡は、前の時刻の画像と次の時刻の画像からオプティカルフローを用いて、特徴点の移動先を求める処理である。

本研究では画像1および画像2の位置ズレは微小であるという仮定から、図4のようにKLT trackerを空間方向の特徴点追跡、すなわち同一時刻における画像1と画像2の間の対応点検出にも用いることが可能である。また時間方向についても、特徴点検出を毎時刻行うことと比べて、特徴点追跡は前の時刻の特徴点座標の周辺のみで計算を行うため、計算コスト・精度の両面において有利な手法である。よって、提案手法では特徴点追跡を積極的に活用する。

提案手法では事前にパラメータ N を定めておき、 N 組以上の対応点が前の時刻で得られている場合はそれらの特徴点を時間方向で追跡し、現在の画像1、画像2における特徴点とする。 N 組以上の対応点が得られていない場合は、現在の画像1および画像2に対してそれぞれ特徴点検出を行う。そしていずれの場合においても、画像1から画像2へ、および画像2から画像1への空間方向の特徴点追跡により、対応点を得る。これらの対応点は画像上の2次元座標 $(u, v)^T$ として得られるが、RGB-Dセンサより得られる奥行き z の情報を加えることで、3次元座標 $(u, v, z)^T$ として得ることが可能である。

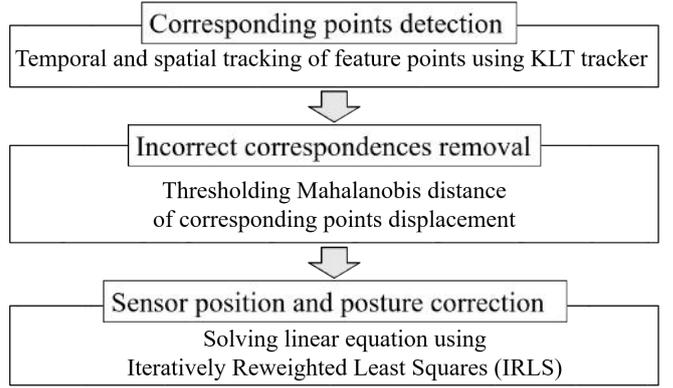


Fig. 3 The flow of the proposed method

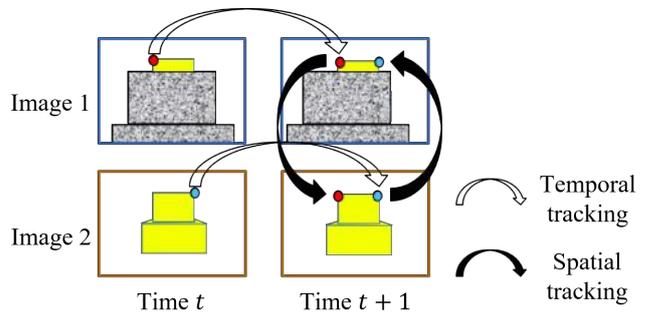


Fig. 4 Temporal and spatial tracking of feature points using KLT tracker

2.2 誤対応点の除去

前節の処理で得た対応点には誤対応点が含まれている場合がある。以下、2.1節の処理で得た対応点の組数を n 、画像1および画像2における対応点の座標を $\mathbf{p}_{1i} = (u_{1i}, v_{1i}, z_{1i})^T$, $\mathbf{p}_{2i} = (u_{2i}, v_{2i}, z_{2i})^T$ ($i = 1, 2, \dots, n$) とする。

画像1と画像2の位置のズレは微小であるという仮定より、対応点の位置のズレ $\mathbf{d}_i = \mathbf{p}_{1i} - \mathbf{p}_{2i}$ は微小である。このとき、センサ1とセンサ2の姿勢の違いが微小であると仮定すると、どの対応点も同様の大きさ・向きで位置のズレが生じる、すなわち \mathbf{d}_i のばらつきは小さいと考えられる。ただし、対応点 $\mathbf{p}_{1i}, \mathbf{p}_{2i}$ が誤対応点であった場合 \mathbf{d}_i は正しい対応点の位置のズレとは異なる値をとる。よって、 \mathbf{d}_i が外れ値をとるような対応点 $\mathbf{p}_{1i}, \mathbf{p}_{2i}$ は誤対応点である可能性が高いと考えられる。

そこで提案手法では、外れ値の判定に基づく誤対応点の除去を行う。 \mathbf{d}_i のばらつきには方向があると考えられるため、ばらつきの方向を考慮した外れ値の判定として、ここではマハラノビス距離による閾値処理を行う。まず、 $i = 1, 2, \dots, n$ における \mathbf{d}_i の平均 $\bar{\mathbf{d}}$ と分散共分散行列 Σ を求める。次に、 $i = 1, 2, \dots, n$ について $\bar{\mathbf{d}}$ と \mathbf{d}_i のマハラノビス距離 $D_i = \sqrt{(\mathbf{d}_i - \bar{\mathbf{d}})^T \Sigma^{-1} (\mathbf{d}_i - \bar{\mathbf{d}})}$ を計算する。これがある閾値 T を超えている点を外れ値、すなわち誤対応点と見なし除去する。以降、除去されなかった対応点の組数を n' とし、それぞれの座標を $\mathbf{p}'_{1i} = (u'_{1i}, v'_{1i}, z'_{1i})^T$, $\mathbf{p}'_{2i} = (u'_{2i}, v'_{2i}, z'_{2i})^T$ ($i = 1, 2, \dots, n'$) とする。

2.3 センサ位置・姿勢補正

2.3.1 補正量を求める方程式の導出

透視映像の生成においては、各センサの絶対的な位置・姿勢ではなく、センサ1とセンサ2の相対位置が重要である。よって、ここではセンサ1を固定し、センサ2の外部パラメータを補正する問題を考える。センサ2の外部パラメータは、セ

ンサ 1 からセンサ 2 への回転行列 $\mathbf{R}_{3 \times 3}$ 、並進ベクトル $\mathbf{t} = (t_x, t_y, t_z)^T$ で表される。ただし、回転行列はロドリゲスの回転公式により回転ベクトル $\mathbf{r} = (r_x, r_y, r_z)^T$ と 1 対 1 で対応するため、以降は回転行列ではなく回転ベクトルを考える。このとき、外部パラメータの補正は $r_x, r_y, r_z, t_x, t_y, t_z$ の補正量 $\Delta r_x, \Delta r_y, \Delta r_z, \Delta t_x, \Delta t_y, \Delta t_z$ を求める問題に帰着する。よって、 $\mathbf{x} = (\Delta r_x, \Delta r_y, \Delta r_z, \Delta t_x, \Delta t_y, \Delta t_z)^T$ とおき、 \mathbf{x} が満たすべき方程式を導出する。

外部パラメータの補正により、画像 2 における対応点の座標 \mathbf{p}'_{2i} は変化し、外部パラメータの補正が理想的に行われれば対応点の位置ズレ $\mathbf{d}'_i = \mathbf{p}'_{1i} - \mathbf{p}'_{2i}$ は全て 0 になると考えられる。また、画像 1 と画像 2 の位置ズレは微小であるという仮定から、補正量も微小であると考えてよい。よって、補正量に対する $\mathbf{p}'_{2i} = (u'_{2i}, v'_{2i}, z'_{2i})^T$ の変化が $\mathbf{d}'_i = (\Delta u'_i, \Delta v'_i, \Delta z'_i)^T$ に等しいという拘束を、テイラー展開の 1 次の項まで考えることにより、補正量についての以下の 1 次方程式 (1) を得る。

$$\mathbf{A}_{3n' \times 6} \mathbf{x} = \mathbf{b}_{3n' \times 1} \quad (1)$$

$$\Leftrightarrow \begin{bmatrix} \frac{\partial(u'_{21}, v'_{21}, z'_{21})}{\partial(r_x, r_y, r_z, t_x, t_y, t_z)} \\ \vdots \\ \frac{\partial(u'_{2n'}, v'_{2n'}, z'_{2n'})}{\partial(r_x, r_y, r_z, t_x, t_y, t_z)} \end{bmatrix} \begin{bmatrix} \Delta r_x \\ \Delta r_y \\ \Delta r_z \\ \Delta t_x \\ \Delta t_y \\ \Delta t_z \end{bmatrix} = \begin{bmatrix} \Delta u'_1 \\ \Delta v'_1 \\ \Delta z'_1 \\ \vdots \\ \Delta u'_{n'} \\ \Delta v'_{n'} \\ \Delta z'_{n'} \end{bmatrix}$$

\mathbf{A} は RGB-D センサの内部パラメータおよび外部パラメータ、特徴点の座標を用いて解析的に求めることが可能である。また、 \mathbf{b} は各対応点の位置ズレの値を並べたベクトルであり、特徴点の座標を用いて容易に計算可能である。この方程式を \mathbf{x} についての最小二乗法で解くことで補正量を求められる。

ただし、 u, v は画像上の座標であり pixel 単位であるのに対し、 z は一般的に mm 単位で得られるため、単位の異なるこれらの量を最小二乗法で同列に扱うのは不適切である。また、実際にはこの方法では補正の精度が良くない場合が多い。その原因として、RGB-D センサの性質上、奥行き z の不確かさが大きいという点が挙げられる。特に、画像上で物体の輪郭線上の点などの z の値を参照する場合、画像上の座標が少しずれただけで z の値が大きく異なってしまう場合がある。そこで、 z の成分が最小二乗法に寄与する重みを調整するため、適当な定数 s をかけた 1 次方程式 (2) を、次節の方法で解くことで補正量を求める。

$$\mathbf{A}'_{3n' \times 6} \mathbf{x} = \mathbf{b}'_{3n' \times 1} \quad (2)$$

$$\Leftrightarrow \begin{bmatrix} \frac{\partial(u'_{21}, v'_{21}, s z'_{21})}{\partial(r_x, r_y, r_z, t_x, t_y, t_z)} \\ \vdots \\ \frac{\partial(u'_{2n'}, v'_{2n'}, s z'_{2n'})}{\partial(r_x, r_y, r_z, t_x, t_y, t_z)} \end{bmatrix} \begin{bmatrix} \Delta r_x \\ \Delta r_y \\ \Delta r_z \\ \Delta t_x \\ \Delta t_y \\ \Delta t_z \end{bmatrix} = \begin{bmatrix} \Delta u'_1 \\ \Delta v'_1 \\ s \Delta z'_1 \\ \vdots \\ \Delta u'_{n'} \\ \Delta v'_{n'} \\ s \Delta z'_{n'} \end{bmatrix}$$

2.3.2 補正量の計算

式 (2) の単純な解法として最小二乗法があるが、最小二乗法は外れ値の影響を受けやすいという問題がある。提案手法では 2.2 節で述べた誤対応点の除去処理を事前に行うが、全ての誤対応点を除去できるとは限らず、除去できなかった少数の誤対応点の影響で補正の精度が低下する場合がある。

そこで、少数の外れ値にロバストであり、計算コストも小さい解法として、反復再重みづけ最小二乗法 (Iteratively Reweighted Least Squares, IRLS) を用いる。IRLS は、最小二乗法において各行の重みを反復的に調整していくことで、明示的に外れ値を検出することなく外れ値の影響を小さくすることが可能な解法である。ここでの重みは式 (2) で導入した s とは別に、残差 $\mathbf{b}' - \mathbf{A}'\mathbf{x}$ を用いて $3n'$ 行のそれぞれに対し計算・更新される。重みの更新方法には様々なものが提案されているが、ここでは初期値によらず大局解に必ず収束する手法として、文献 [7] を参考に Huber のウェイト関数を使用する。

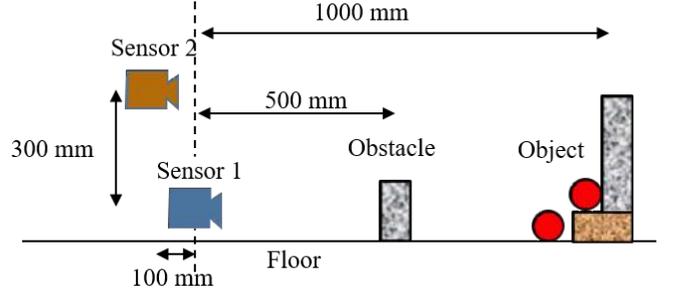


Fig. 5 Experimental environment



Fig. 6 The input image of sensor 1 (left) and 2 (right)

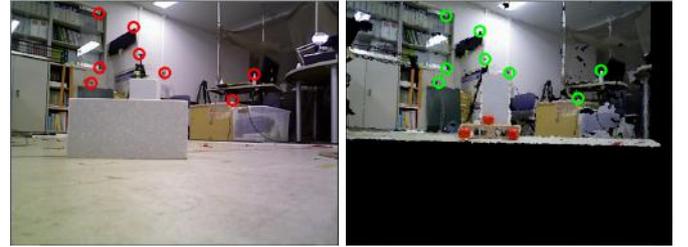


Fig. 7 Coordinate-transformed Fig. 6 and the position of detected corresponding points

3 2つの RGB-D センサを用いた実験

3.1 実験条件

提案手法による位置ズレ補正の有効性を検証するために行った実験について述べる。本実験では、RGB-D センサとして ASUS Xtion Pro Live を 2 台使用し、それぞれ 30 fps のフレームレートで取得した 640×480 pixel の映像を用いて透視映像を生成した。実験環境の様子を図 5 に示す。2 台の RGB-D センサのうち一方をセンサ 1 としたとき、センサ 2 はセンサ 1 から見て上方に約 300 mm、後方に約 100 mm の位置に設置した。また、センサ 1 の約 500 mm 前方に遮蔽物を設置し、約 1000 mm 前方に透視対象を設置した。

実験は以下の手順で行った。まず、事前のキャリブレーションおよび手動での微調整により、筆者らの従来のシステム [3] で位置ズレのない透視映像を生成した。透視映像はセンサ 1 の視点から見た映像とした。その後、センサ 2 を鉛直軸まわりに約 3° 時計回りに回転させ、透視映像に位置ズレを発生させた。この位置ズレを提案手法によって補正した。

提案手法の各種パラメータは試行錯誤的に選択し、 $N = 4$, $T = 2$, $s = 0.01$ とした。また、IRLS の各行の重みが十分に収束するための反復回数を事前に検証した結果から、反復回数は 1 回の補正処理あたり 10 回で固定した。

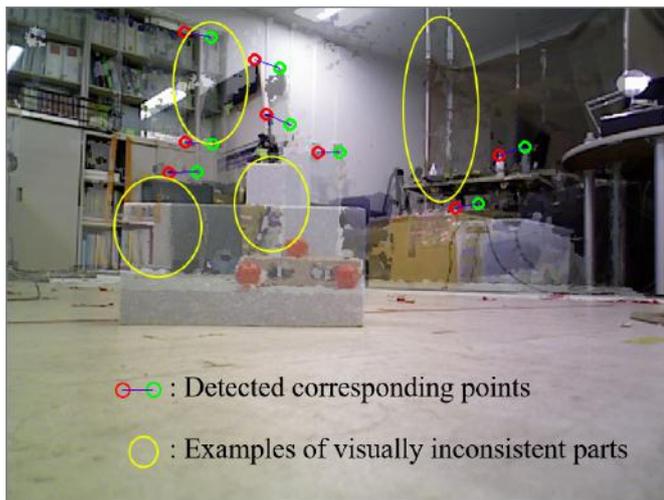


Fig. 8 Output see-through image before correction

3.2 実験結果

センサ 2 を鉛直軸まわりに約 3° 回転させた後の各センサからの入力映像の例を図 6 に示す。センサ 1 からは、遮蔽物に隠れて透視対象は一部しか見えない。

各センサからの入力映像に座標変換を施した結果を図 7 に示す。センサ 1 の視点から見た透視映像を生成するため、センサ 1 からの入力映像は座標変換後も変化しない。一方、センサ 2 からの入力映像はセンサ 1 の視点から見た映像へと座標変換されている。また、それぞれの映像から検出された、誤対応点の除去処理後の対応点の位置を図 7 中に赤および緑で囲って示した。この瞬間の映像からは 8 組の対応点を検出されており、目視で確認できるようにいずれも正しい対応点である。

図 7 の 2 つの映像を統合して生成した透視映像を図 8 に示す。遮蔽物の奥にある透視対象は透視できているが、位置ズレが存在するため、図中に黄色で囲って示した部分などで透視対象の輪郭が不連続である、物体が二重に見えるなどの現象が生じている。なお、図 7 で示した対応点を図 8 中にも示した。このとき、透視映像上では 30 pixel 程度の位置ズレが存在している。

対応点を用いて提案手法によって位置ズレを補正した結果を図 9 に示す。図 8 と比較して、位置ズレが大幅に小さくなったことが分かる。また、補正処理を毎フレームごとに実行した場合のフレームレートを約 30 秒間にわたって計測した。Intel Core i7-3770 (クロック数 3.40 GHz) の CPU を使用して透視映像の生成および位置ズレの補正を行った場合、平均 9.9 fps、標準偏差 3.0 fps となった。

これらの結果から、リアルタイムで位置ズレを補正しながら透視映像を生成することが可能であることが示された。

4 結言

本研究では、透視映像における画像の位置ズレをリアルタイムで補正する手法を提案した。提案手法では、KLT tracker を用いた時間・空間方向の特徴点追跡により少ない計算コストで対応点を求める。そして各対応点の位置のズレのばらつきに着目して誤対応点の除去を行う。さらに、位置ズレが微小であるという仮定より導出した線形方程式を IRLS で解くことで、少数の誤対応点にロバストな補正を実現する。2 つの RGB-D センサを用いた実験によって、提案手法による補正の有効性およびリアルタイム性が示された。

今後の課題として、より正確な対応点の求め方、あるいは対応点によらない画像の位置ズレ検出・補正手法の構築などが挙げら



Fig. 9 Output see-through image after correction

れる。提案手法のように各対応点の位置のズレのばらつきを用いた誤対応点の除去では、誤対応点明らかに正しい対応点の位置からずれている場合にしか対処できない。対応点の周囲のテキスト情報などを用いることで、より正確な対応点を得られる可能性がある。また、テキストが乏しい場合や遮蔽物の存在により 2 つの映像の見た目が大きく異なる場合などは対応点を求めることが困難なため、3 次元形状などの別の情報も活用することが考えられる。

謝辞

本研究の一部は、総合科学技術・イノベーション会議により制度設計された革新的研究開発促進プログラム (ImPACT) 「タフ・ロボティクス・チャレンジ」の援助を受けた。

参考文献

- [1] Masaharu Moteki, Kenichi Fujino, Takashi Ohtsuki, and Tsuyoshi Hashimoto, "Research on Visual Point of Operator in Remote Control of Construction Machinery", Proceedings of the 28th International Symposium on Automation and Robotics in Construction, pp. 532–537, 2010.
- [2] 森 尚平, 一刈 良介, 柴田 史久, 木村 朝子, 田村 秀行, "隠消現実感の技術的枠組と諸問題: 現実世界に実在する物体を視覚的に隠蔽・消去・透視する技術について", 日本バーチャルリアリティ学会論文誌, vol. 16, no. 2, pp. 239–250, 2011.
- [3] 橋高 達也, 藤井 浩光, 山下 淳, 浅間 一, "移動可能な RGB-D センサを用いた任意視点からの遮蔽物透視システム", 精密工学会誌, vol. 83, no. 3, pp. 235–244, 2017.
- [4] Mili Shah, Roger D. Eastman, and Tsai Hong, "An overview of robot-sensor calibration methods for evaluation of perception systems", Proceedings of the 11th Workshop on Performance Metrics for Intelligent Systems, pp. 15–20, 2012.
- [5] 森 尚平, 小向 啓文, 柴田 史久, 木村 朝子, 田村 秀行, "隠消現実感における隠背景平面拘束と周辺参照領域の効果的利用", 日本バーチャルリアリティ学会論文誌, vol. 19, no. 2, pp. 131–140, 2014.
- [6] Carlo Tomasi and Takeo Kanade, "Shape and motion from image streams under orthography: a factorization method", International Journal of Computer Vision, vol. 9, no. 2, pp. 137–154, 1992.
- [7] 和田 かず美, "多変量外れ値の検出～繰返し加重最小二乗 (IRLS) 法による欠測値の補定方法～", 統計研究彙報第 69 号, pp. 23–52, 2012.