

遠隔操縦建機のための 屋外環境における遮蔽物透視システム*

長野 樹** 藤井浩光*** 橘高達也† 淵田正隆††
深瀬勇太郎††† 青木 滋††† 鳴海智博††† 山下 淳†† 浅間 一††

See-through System Using LiDAR for Tele-operated Construction Machines

Tatsuki NAGANO, Hiromitsu FUJII, Tatsuya KITTAKA, Masataka FUCHIDA,
Yutaro FUKASE, Shigeru AOKI, Tomohiro NARUMI, Atsushi YAMASHITA and Hajime ASAMA

Demand for teleoperation of construction machines at disaster sites is increasing to prevent the secondary disasters. In teleoperation, operators can see the environment using cameras mounted on the construction machines. However, there is a problem that the camera images have blind spots occluded by obstacles such as the arm of the construction machine itself, which cause a decrease of work efficiency. This paper presents a method to generate images with few blind spots by seeing through foreground obstacles and visualizing background objects. The proposed method uses two RGB cameras and LiDAR to acquire texture data with the three-dimensional information of the environment. And then, it projects the background image acquired by one camera to the other camera by using the three-dimensional information, and integrates the two camera images into a see-through image. In the experiment with an actual construction machine, we succeeded in generating see-through images in real time.

Key words: diminished reality, teleoperation, construction machine, LiDAR, outdoor environment, texture mapping

1. 序 論

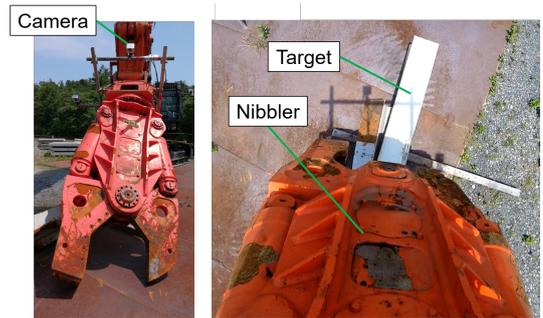
地震や火山噴火などの災害現場では、作業員の安全を確保し、二次災害の発生を避けることを目的として、無人の建機を遠隔操作することによる災害対応に期待が高まっている¹⁾。

遠隔操作では、オペレータは建機に搭載したカメラの映像を確認しながら作業を行うことが一般的である。その際、カメラ映像の視野の狭さや死角の存在による作業効率の低下が問題となっている²⁾。その解決方法として、カメラを複数搭載して、様々な視点からの映像を提示することで広い視野を確保することが一般に行われてきた³⁾。しかし、遠隔作業中のオペレータは、複数のモニタを提示した場合にも特定のモニタのみを注視していることが指摘されている²⁾。その上、複数のモニタを見比べながら作業をすることは、オペレータへの負担につながるという問題もある⁴⁾。したがって、作業効率を向上させるためには、視野が広く死角の少ない1つの映像を提示することが有効であると考えられる。

災害現場における建機の重要な作業の1つとして、本研究では図1のようなアームを有する建機を用いた屋外でのガレキの撤去作業に焦点を当てる。建機の手先には図2(a)に示すようなニブラと呼ばれる作業用のアタッチメントが取り付けられており、ニブラを開閉することでガレキの把持・破断を行う。そのため、遠隔操作で効率よく作業を行うには、ニブラの開閉



Fig. 1 A construction machine with arm



(a) A nibbler (b) The view from the camera mounted on the arm

Fig. 2 The problem in tele-operation of construction machines

状態に加えてその周囲の環境を把握可能な映像提示が必要である。従来の映像提示では、図2(b)で示すようなニブラを見下ろす位置に設置したカメラからの映像が使われてきた。しかしこの映像では、広い範囲でニブラによる死角が生じており、その裏側(以下、背景と呼ぶ)の状況を把握することが不可能である。このような状況では把持対象全体を確認できないため、細

* 原稿受付 平成30年5月17日

掲載決定 平成30年7月13日

** 学生会員 東京大学大学院(東京都文京区本郷7-3-1)

*** 千葉工業大学(千葉県習志野市津田沼2-17-1)

† 安川電機(福岡県北九州市八幡西区黒崎城石2-1)

†† 正会員 東京大学大学院

††† 清水建設(東京都中央区京橋2-16-1)

長いガレキの先端など不安定な部分を掴んでしまうことや背景に存在する障害物に衝突する危険が生じる。したがって、1つの映像上でニブラ本体とその背景を同時に視認可能なことが必要である。

映像中の死角の問題に対して、本研究では遮蔽物を透視したような映像（以下、透視映像と呼ぶ）を生成する手法に着目する。透視映像は遮蔽物がある領域に、その背景の映像を重畳することで生成される。背景の情報を取得する手法として、事前に環境をセンシングしておく手法⁵⁾⁶⁾や周囲の画素情報から推定する手法⁷⁾⁸⁾が提案されている。しかし、遠隔操作で使用するためには、未知環境においても実際の背景の状況をオンラインで取得・提示することが必要不可欠であるため、これらの手法は適用できない。

また、Enomotoらは、背景が2次元平面であると仮定した手法を提案している⁹⁾¹⁰⁾。この手法では、別視点のカメラから得た背景映像を平面射影変換によって視点を変換して合成することで透視映像生成している。しかし、ガレキなどが散在する災害現場を2次元平面に近似することは困難であり、この手法は適用できない。

背景をオンラインで、かつ3次元的に取得する手法として、著者らは複数のRGB-Dセンサを用いた手法を提案している¹¹⁾¹²⁾。RGB-Dセンサは、カメラの映像に加えて画素ごとに深度情報を取得可能なセンサである。環境を色情報付きの3次元点群として取得することで、映像の任意視点への変換が可能となる。それによって、複数視点からの映像を1つの視点に投影・合成することで透視映像を生成している。しかし、一般的に実用化されているRGB-Dセンサの多くは、赤外光を使用して測距を行っている。そのため、太陽光の影響が大きい屋外環境では、赤外光の読み取りが困難で正確な測距ができず、屋内での使用に限定された手法である。

一般に、屋外での測距手段としてLiDAR (Lighting Detection and Ranging) と呼ばれる、レーザ光を用いたTime of Flight方式のセンサが用いられている。本研究では、LiDARで得られる3次元情報を用いて複数台のカメラ映像を統合することによって、未知の屋外環境においてオンラインで透視映像を生成することを目的とする。

2. アプローチ

2.1 要求仕様

第1章での議論から提案手法が満たすべき要求仕様を以下のようにまとめる。

- 1つの映像上で遮蔽物とその背景を同時に視認可能なこと
- 屋外環境で適用可能なこと
- オンラインで映像提示可能なこと
- 背景映像を正確な位置に提示可能なこと

ガレキの撤去作業では、1つの映像上でニブラとその背景を視認できることが重要であり、本研究では、透視映像の生成によってこの問題の解決を図る。また、従来研究では考慮されていなかった、屋外環境への適用を想定している。加えて、遠隔操作での使用を想定しているため映像のオンライン性が不可欠である。本研究においてオンラインであるとは、センサから送られてきた情報を即座に処理して遅延なく映像を提示することである。佐藤ら¹³⁾による建機の遠隔操縦の熟練オペレータに対する調査では、9 fpsでの映像提示で十分使用に耐え得るこ

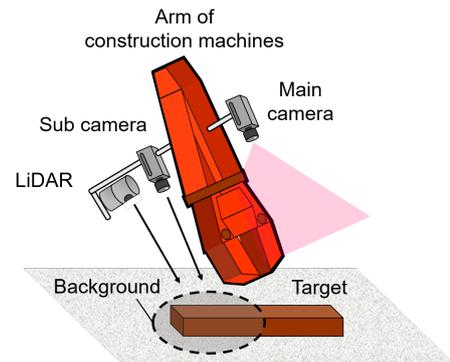


Fig. 3 The schema of the system

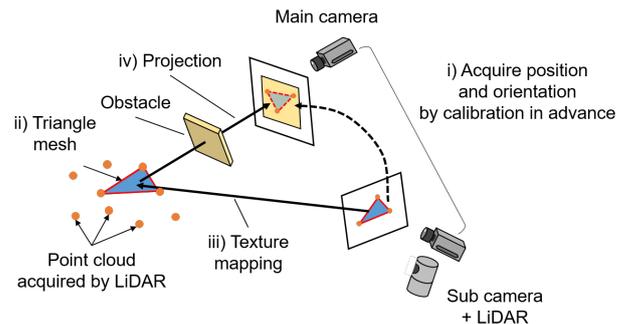


Fig. 4 The flow of the proposed method

とが示されている。そこで本研究では、映像のフレームレートとして9 fpsを要求仕様を設定する。最後に、背景の映像を正しい位置に提示することも重要である。ガレキを正確に把持するためには、提示する背景映像と実際の背景との位置ずれが建機のニブラのサイズと比較して十分小さいことが求められる。また、ニブラが離れたところから把持対象にアプローチする過程において、常に設定した要求精度を満たす必要がある。定量的には、4.1節において提案システムを搭載する建機を明確にしたのちに、その操作特性を踏まえて適切に設定する。

2.2 システムの構成

図3に提案システムの構成を示す。システムは、建機のアームに取り付けた2台のカメラとLiDARからなる。ニブラを見下ろす位置に取り付けられたカメラ（以下、メインカメラと呼ぶ）がオペレータの視点となるカメラであり、この映像中のニブラによる死角を問題として取り上げる。そして、メインカメラからは見えない背景の映像と3次元情報を取得するために、アームを挟んで反対側にカメラ（以下、サブカメラと呼ぶ）とLiDARを設置する。また、各センサは遠隔操縦中に相対的な位置姿勢関係が変化しないようにすべて同じリンクに固定する。

2.3 提案手法の流れ

以下に提案手法の流れを示す（図4）。i)は事前に行われ、ii)～iv)がオンラインでの処理となる。

- 各センサ間の相対的な位置姿勢関係を求める。
- LiDARで得られる点群からメッシュを生成し、背景の3次元復元を行う。
- 生成したメッシュにサブカメラ映像から対応するテクスチャをマッピングする。
- テクスチャ付きメッシュを座標変換によってメインカメラに投影、メインカメラ映像と合成する。

まず、カメラの映像とLiDARの3次元情報を統合するため

に、事前のキャリブレーションによってセンサ間の相対的な位置姿勢関係を求める。これによって、LiDAR で得られた点群の、2台のカメラ画像座標系において対応する座標を知ることが可能となる。

ここで、本研究で使用する LiDAR では、前述の RGB-D センサと比べると取得できる 3 次元点群の密度が疎になるという問題がある。LiDAR は、照射角が異なる複数本のレーザ光を 360 deg 回転させ、その反射時間を計測することで測距を広範囲に行っている。LiDAR による測距はレーザ光の走査線上のみ可能であり、照射されるレーザ光の本数も限られているため、スキャン方向に垂直な方向に対しては環境を粗くしかセンシングできず、疎な 3 次元情報しか得られない。

この問題を解決するために、本研究では点群のメッシュ化によってレーザ間の 3 次元情報を補間する。そして、生成したメッシュにサブカメラから得られるテクスチャ情報を付与し、それをメインカメラ映像に投影、合成することで遮蔽物を透視したような映像を生成する。上記の ii)~iv) を高速に処理することによってオンラインでの映像生成を可能にする。

3. 提案手法

3.1 座標系の定義

本章で使用する座標系の定義について述べる。ある点 P を、環境中の任意の場所に設置したワールド座標系 Σ_W において ${}^W\mathbf{p} = [x_w, y_w, z_w]^T$ と表す。同様に、カメラ座標系 Σ_C 、LiDAR 座標系 Σ_F における表現を ${}^C\mathbf{p} = [x_c, y_c, z_c]^T$ 、 ${}^F\mathbf{p} = [x_f, y_f, z_f]^T$ とする。そして、 Σ_W から Σ_C への座標系の変換を 3 行 3 列の回転行列 ${}^C\mathbf{R}_W$ と 3 次元の並進ベクトル ${}^C\mathbf{t}_W$ を用いて以下の式のように定義する。

$${}^C\mathbf{p} = {}^C\mathbf{R}_W {}^W\mathbf{p} + {}^C\mathbf{t}_W. \quad (1)$$

同様に、 Σ_C から Σ_F への変換は ${}^F\mathbf{R}_C$ と ${}^F\mathbf{t}_C$ を用いて表される。さらに、 ${}^C\mathbf{p}$ を画像上に投影したときの同次座標をカメラ画像座標系 Σ_M で ${}^M\mathbf{p} = [u, v, 1]^T$ と表し、カメラの内部パラメータ \mathbf{A} によって次の式で表現できる。

$${}^M\mathbf{p} \simeq \mathbf{A} {}^C\mathbf{p}, \quad (2)$$

ここでは、 \simeq は定数倍の不定性を表すものとする。

3.2 カメラと LiDAR のキャリブレーション

本節では、Zhang ら¹⁴⁾¹⁵⁾ が提案したカメラと LiDAR のキャリブレーション手法について述べ、両者の位置姿勢関係を表す回転行列 ${}^F\mathbf{R}_C$ と並進ベクトル ${}^F\mathbf{t}_C$ を求めることを目的とする。この手法では、図 5 のように大きさが既知の 2 色の四角形のパターンが交互に並んだチェッカーボードを利用し、カメラと LiDAR が同時にチェッカーボードを観測している必要がある。まず、Zhang ら¹⁶⁾ のキャリブレーション手法を用いてカメラの内部パラメータ \mathbf{A} と Σ_W から Σ_C への変換 ${}^C\mathbf{R}_W$ 、 ${}^C\mathbf{t}_W$ を求める。ここでは、ワールド座標系 Σ_W をチェッカーボードの平面が $z_w = 0$ となるように設定する。この時、カメラの光学中心からチェッカーボードに下した垂線のベクトル \mathbf{n} をカメラ座標系 Σ_C で次の式 (3) を用いて求めることができる。

$$\mathbf{n} = -\mathbf{r}_3(\mathbf{r}_3 \cdot {}^C\mathbf{t}_W), \quad (3)$$

\mathbf{r}_3 は回転行列 ${}^C\mathbf{R}_W$ の第 3 列目を表す。 \mathbf{n} の大きさ $\|\mathbf{n}\|$ はチェッカーボードからカメラまでの距離に等しい。

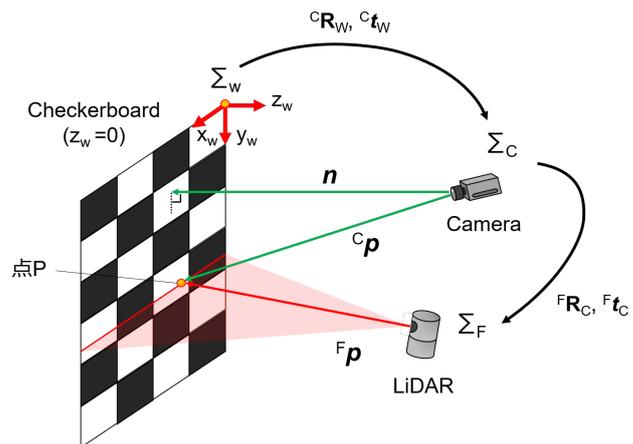


Fig. 5 Camera-LiDAR calibration

また、LiDAR によって得られるチェッカーボード上の点 P のカメラ座標系 Σ_C 、LiDAR 座標系 Σ_F での座標 ${}^C\mathbf{p}$ と ${}^F\mathbf{p}$ には、式 (4) のような関係が成り立つ。

$${}^F\mathbf{p} = {}^F\mathbf{R}_C {}^C\mathbf{p} + {}^F\mathbf{t}_C, \quad (4)$$

${}^F\mathbf{p}$ は LiDAR がチェッカーボード上に照射した複数本のレーザ光のうち、回転軸に垂直なレーザ光で取得した点である。 ${}^F\mathbf{p}$ は Kassir ら¹⁵⁾ による手法によって全点群の中から検出する。また、点 P はチェッカーボード上の点であることから ${}^C\mathbf{p}$ と \mathbf{n} について次の式のように表せる。

$$\mathbf{n} \cdot {}^C\mathbf{p} = \|\mathbf{n}\|^2. \quad (5)$$

さらに、式 (5) に式 (4) を代入すると、次の式 (6) を導くことができる。

$$\mathbf{n} \cdot ({}^F\mathbf{R}_C^{-1} ({}^F\mathbf{p} - {}^F\mathbf{t}_C)) = \|\mathbf{n}\|^2. \quad (6)$$

${}^F\mathbf{p}$ 、 \mathbf{n} は既知であるので、式 (6) は ${}^F\mathbf{R}_C$ 、 ${}^F\mathbf{t}_C$ に関する 1 つの方程式になる。求めたい未知数 ${}^F\mathbf{R}_C$ 、 ${}^F\mathbf{t}_C$ は位置姿勢の 6 自由度であり、1 つの ${}^F\mathbf{p}$ と \mathbf{n} の組み合わせで 1 つの拘束条件を与える。したがって、最低 6 つの組があればすべての未知数を求めることが可能である。チェッカーボードを様々な位置に動かし、多数の ${}^F\mathbf{p}$ と \mathbf{n} のデータの組を得ることで、Levenberg-Marquardt 法¹⁷⁾ による最適化計算により ${}^F\mathbf{R}_C$ と ${}^F\mathbf{t}_C$ を推定する。

3.3 メッシュの生成

2.3 節で示したように LiDAR で得られる点群の密度は疎であり、環境の 3 次元情報を十分に取得できない。そのため本研究では、点群から近傍の 3 点を選び出し、三角メッシュを生成することによる補間手法を提案する。メッシュの生成手法としては、全点群の中から近傍点を探索する手法¹⁸⁾ などが代表的である。しかし点群の数が多い場合、計算量が膨大となるため、オンラインでの処理は困難になる。そこで、LiDAR で得られる点群がレーザ光の走査線上に分かれて並んでいることを利用し、隣り合う 2 本のレーザからそれぞれ近傍点を選ぶことで高速にメッシュを生成する手法を提案する。

具体的な手法について説明する。まず、図 6 に示すように物体に照射されたレーザ光の走査線を端から順番にレーザ $L_1, L_2, \dots, L_j, \dots$ と呼ぶこととする。また、レーザ L_j 上で取得した点群の適当な点を基準として $P_{(1,j)}, P_{(2,j)}, \dots, P_{(i,j)}, \dots$ のように定める。 $P_{(1,j)}, P_{(2,j)}, \dots, P_{(i,j)}, \dots$ は同じレーザ光によ

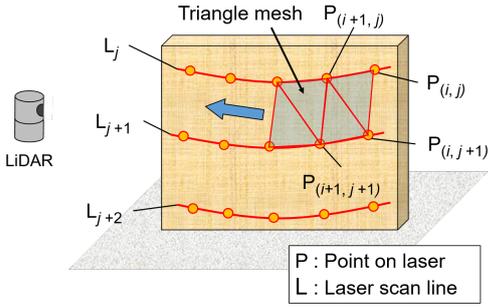


Fig. 6 Proposed mesh generation method

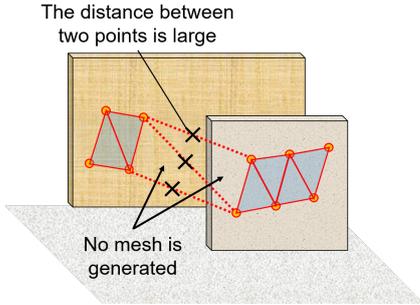


Fig. 7 An example when mesh is not generated

て取得されたため、LiDAR 座標系の原点から見た仰角が等しくお互いに隣り合う点である。さらに $P_{(i,1)}, P_{(i,2)}, \dots, P_{(i,j)}, \dots$ の方位角が等しくなるように定義すると、 $P_{(i,j)}$ と $P_{(i,j+1)}$ は同じレーザー光で取得した点以外では最も近傍にあると仮定できる。提案手法では、以下の処理を 2 本のレーザー L_j と L_{j+1} で行うことで、そのレーザー間を埋め尽くすようにメッシュを生成する。また、その処理の様子を図 6 に示す。

- i) 2 本のレーザー光からそれぞれ $P_{(i,j)}, P_{(i,j+1)}$ を選び、これらを基準点とする。
- ii) 基準点と隣合う点である $P_{(i+1,j)}, P_{(i+1,j+1)}$ を追加し、2 つの基準点との距離をそれぞれ調べる。
- iii) 距離が設定した閾値の範囲内である場合、この 4 点で 3 角メッシュを 2 つ生成する。
- iv) 基準点を $P_{(i+1,j)}, P_{(i+1,j+1)}$ に更新し、2) からの処理を繰り返す。

ここで、図 7 に示すように隣り合う点同士であっても、レーザー光の照射位置が奥行き方向に離れていれば必ずしも近距離にあるとは限らない。点同士が離れていると大きく引き伸ばされたメッシュが生成され、3.4 節で説明するテクスチャマッピングの際にメッシュに歪んだ映像が付与されてしまう。そこで、iii) で基準点との距離が設定した閾値を超えていた場合はメッシュの生成は行わず、基準点の更新のみを行う。

本研究では、LiDAR で取得した背景の点群情報から、すべてのレーザー間でこの処理を繰り返し、3 次元のポリゴンとして背景の形状を復元する。

3.4 テクスチャマッピング

3.3 節では、メッシュの生成により背景の 3 次元形状の復元を行った。この 3 次元情報を用いて、サブカメラで取得した背景映像をメインカメラ視点の映像に変換する。

3.2 節で各カメラの内部パラメータ、LiDAR とカメラの位置姿勢関係がすでに求められているため、式 (2), (4) から導かれる次の式 (7) によって生成したメッシュを各カメラの画像上に

投影することが可能である。

$${}^M \mathbf{p} \simeq \mathbf{A}^F \mathbf{R}_C^{-1} ({}^F \mathbf{p} - {}^F \mathbf{t}_C). \quad (7)$$

サブカメラ画像上に投影することで生成したメッシュの画像上における対応位置を知ることが可能となる。対応する画像上のテクスチャをメッシュにマッピングすることで、背景のテクスチャ付きの 3 次元形状を生成することが可能となる。さらに、そのメッシュのメインカメラ画像上での対応関係も知ることができるため、メインカメラから背景を見たような映像を生成することができる。

3.5 映像の合成

本節では、3.4 節で生成されたメインカメラ視点の背景映像とメインカメラ映像の合成について述べる。本研究では、背景映像が取得できた領域に式 (8) に示すようなアルファブレンディングを適用することで 2 つの映像を合成する。

$$\mathbf{I}_{out} = \alpha \mathbf{I}_s + (1 - \alpha) \mathbf{I}_m \quad (0 \leq \alpha \leq 1). \quad (8)$$

$\mathbf{I}_m, \mathbf{I}_s$ がそれぞれメインカメラとサブカメラ (背景) の映像の RGB 値を表し、式 (8) により合成映像 \mathbf{I}_{out} が出力される。 α は、メインカメラ映像に重畳するサブカメラ映像の不透明度を表すパラメータである。 α の値を変化させることで遮蔽物の透視の度合いを変化させることができる。 $\alpha = 0$ のときメインカメラの映像がそのまま出力され、逆に $\alpha = 1$ のとき映像中の該当領域が完全な透明となり背景のみが出力される。背景映像を取得できなかった範囲は $\alpha = 0$ とし、 $\mathbf{I}_{out} = \mathbf{I}_m$ となる。

4. 実験

4.1 実験環境

提案手法の有効性を検証するために、図 8 のような環境で実験を行った。建機は日立建機 ZX350B を使用し、ガレキの撤去作業時にアームの先端に取り付けるニブラには日本ニューマチック工業社 S-36XCR を使用した。

4.1.1 使用するセンサ

提案システムでは、2 台のカメラと 1 台の LiDAR を使用する。カメラは、解像度が $2,048 \times 2,048$ pixel の Point Grey Research 社の Grasshopper3 GS3-U3-41C6C-C を用い、入力フレームレートを 15 fps に設定した。カメラのレンズには広い視野を確保するために、焦点距離 8 mm、画角が 79.7 deg の 1 インチ対応の固定焦点レンズを用いた。さらに、LiDAR は、2.1 節で挙げた映像のオンライン提示に関する要求仕様を満たすために

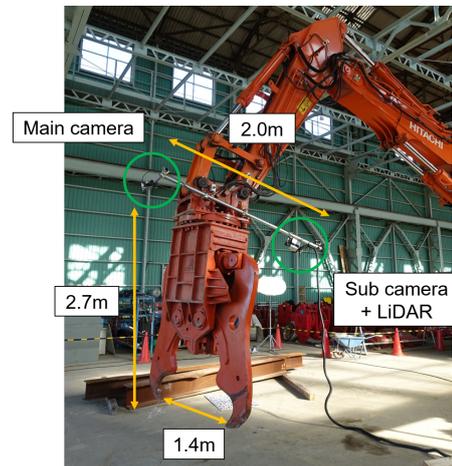


Fig. 8 Experimental setup

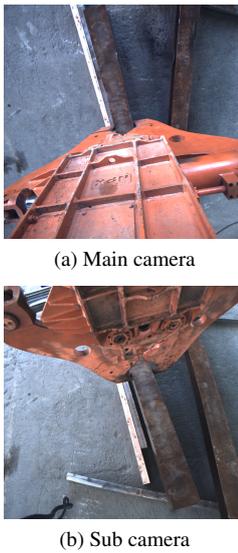
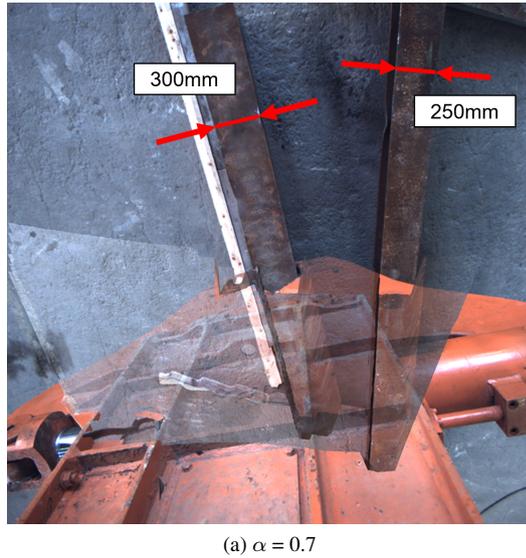


Fig. 9 Original image



(a) $\alpha = 0.7$



(b) $\alpha = 1.0$

Fig. 10 See-through image by the proposed method

10 fps の速度で測距可能な Velodyne LiDAR 社の HDL-32e を使用した。この LiDAR の測定視野は水平方向 360 deg, 鉛直方向 41.3 deg (+10.67 deg ~ -30.67 deg), 測定距離は 1 ~ 100 m である。

また、センサの設置位置に関しては、把持作業の様子を大きく写すためにニブラの近くに設置することが求められるが、同時に把持対象物の破片がセンサに衝突して破損する危険も高まる。センサの建機からの張り出しも同様の危険があることを考慮して、ニブラの回転軸上方である先端から 2.7 m の位置に長さ 2.0 m の単管フレームを取り付け、その両端にカメラと LiDAR を配置する。これらはすべて同じリンクに固定されているため、アームを動作させてもお互いの位置姿勢関係は変化しない。

4.1.2 位置ずれの要求精度

本項では、映像の位置ずれの定量的な要求精度について述べる。2.1 節で述べたように正確な把持を行うためには、位置ずれがニブラのサイズ、特に刃と刃の間の幅と比較して小さいことが必要である。本実験で使用するサイズは、ニブラを最大まで開いた時の間隙の幅が約 1.4 m である。そこで、その 20 分の 1 に当たる 0.07 m を要求精度として設定する。本実験では、センサから把持対象までの距離が 7.0 m である位置からニブラを近づけていくことを想定し、このとき常に位置ずれが 0.07 m 以内となることを要求する。

4.1.3 閾値の設定

ここでは、3.3 節で述べたメッシュの生成における距離の閾値の設定を行う。閾値が大きいと映像の欠けは少ないが、一方で歪んだ映像が生成される。本実験では、映像の視認性の観点から出力映像の歪みよりも欠けを問題視し、LiDAR と地面との距離も考慮に入れて、映像の欠けが少なくなるように試行錯誤的に 500 mm に設定した。選ばれた点がすべて 500 mm 以内の距離に存在すればメッシュを生成する。

以上の実験環境のもと、オペレータが実際に建機を操縦し、把持対象である鋼材にアプローチを行い、そのときの映像をオンラインで処理し透視映像を生成した。鋼材には、幅 300 mm と 250 mm の 2 種類の H 形鋼を用いた。なお、映像処理用の PC にはクロック数 2.70 GHz の Intel Core i7-6820HQ CPU を

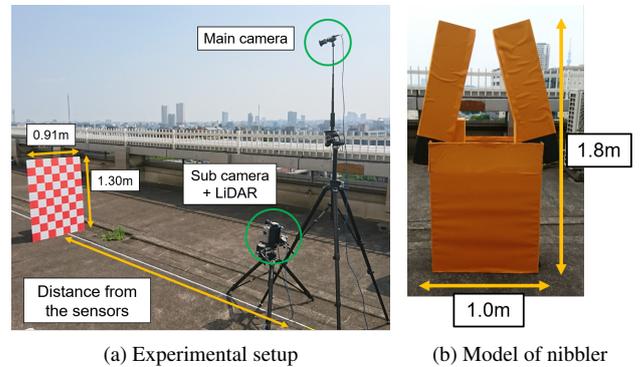


Fig. 11 Experiment for evaluation

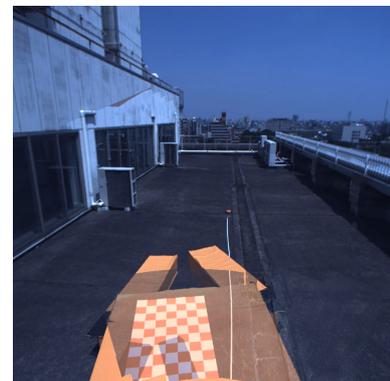


Fig. 12 See-through image in outdoor environment

使用した。

4.2 実験結果

4.1 節で述べた実験環境において得られた結果を図 9, 図 10 に示す。図 9(a), 図 9(b) はそれぞれ合成前のメインカメラ, サブカメラの映像である。どちらの映像でもアームによる死角が存在し、把持対象であるガレキを一部しか把握できない。図 10 が提案手法によって得られた透視映像である。アルファブレンディングによる映像合成時の α の値を $\alpha = 0.7$, $\alpha = 1.0$ とした結果を図 10(a), 図 10(b) にそれぞれ示す。提案手法によりアームを半透明化し、サブカメラで得られた背景の映像を重畳することによって、隠されていたガレキ全体を認識できること

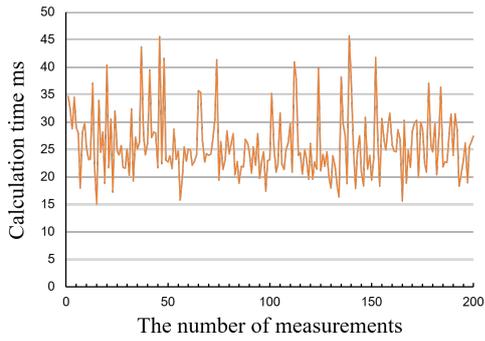


Fig. 13 Calculation time of mesh generation

を確認した。それによって、複数の映像を見比べたり、アームを動かして視点を変えたりしなくとも、ガレキの掴むべき位置や障害物に衝突する危険性を判断することが可能となった。

4.3 評価

本節では、2.1 節で要求仕様として挙げた、屋外環境で適用可能な手法であるか、または映像のオンライン性と背景映像の提示位置の正確さの3点について評価を行う。

4.3.1 屋外環境での透視映像生成

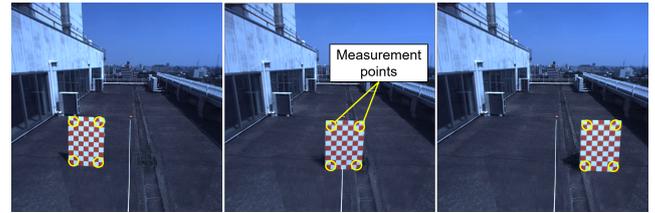
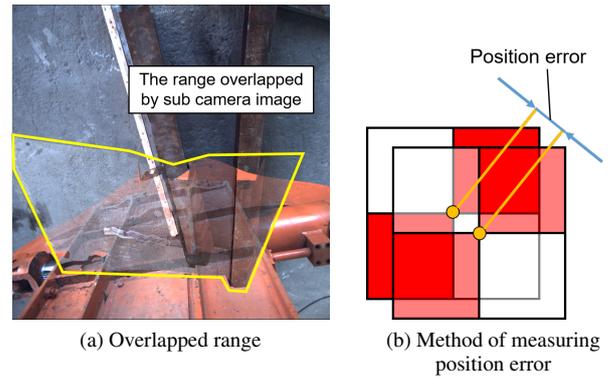
2.1 節で示したように、提案手法は屋内・屋外の環境を問わずに適用可能であることが求められる。そこで、図 11(a) で示すような屋外環境においても透視映像が生成可能であるか検証した。図 11(b) に示すニブラの模型を2台のカメラの間に挿入し、この模型を半透明にして表示した。出力映像を図 12 に示す。すると、ニブラの模型によって遮蔽された背景に置かれているチェッカーボードを透視によって視認することができた。したがって、太陽光の影響を受ける日中の屋外環境であっても、提案手法によって遮蔽物を透視する映像が生成可能なことを確認した。

4.3.2 映像のオンライン性

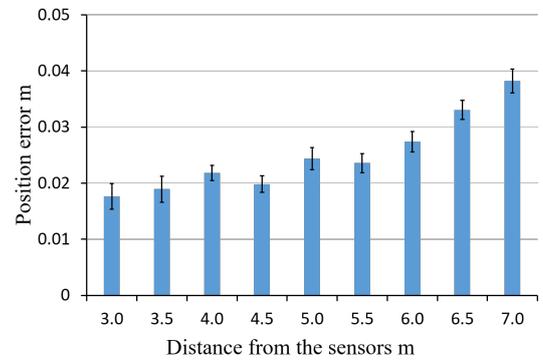
映像のオンライン性は遠隔操作のための映像提示において不可欠な要素である。そこで、出力された透視映像のフレームレートの計測を行った。任意のタイミングで10秒間計測すると、平均で10.1 fps という結果となった。また、メッシュの生成処理にかかる計算時間を計測すると、任意のタイミングの200回の計測で平均25.8 ms であった。その計測結果を図 13 に示す。メッシュの生成は約40 fps で処理可能であるが、出力される映像はLiDARの入力フレームレートがボトルネックとなり10 fps が上限となっている。提案手法によりメッシュの生成処理が高速に行われたことによって、2.1 節で示した要求仕様を満たすことに成功した。

4.3.3 映像合成時の位置ずれ

ここでは、2台のカメラの映像を合成したときの背景映像の位置ずれについて定量的な評価を行う。評価は、センサ間の遮蔽物を取り除いた環境で行う。遮蔽物が存在しないため、位置ずれが生じていなければメインカメラの映像とそこに重畳するサブカメラの映像は一致するはずである。定量的な評価のために図 11(a) のように環境中に大きさが既知のチェッカーボードを地面に垂直に設置して、映像合成時のパターンのずれを計測する。本実験では、1辺が0.13 m の正方形が10×7個並んだチェッカーボードを使用し、図 14(b) に示すように白と赤の四角形で囲まれた格子点の位置ずれを評価した。図 14(a) は実機実験においてサブカメラ映像を重畳した範囲である。位置ずれ



(c) 12 measurement points



(d) Evaluation results of position errors

Fig. 14 Evaluation of position error

は、この範囲で要求精度を満たす必要があるため、図 14(c) に示すようにチェッカーボードを3か所に移動させ、それぞれで格子点の4つの端点のずれを計測し、合計で12点の平均をとって評価した。また、図 8 のようにシステムを建機に搭載したときのセンサからニブラの先端までの距離はおおよそ3.0 m であり、ガレキを把持するためにはこの距離でのずれが最も重要である。さらに、7.0 m の高さからガレキにアプローチしていくことを想定し、チェッカーボードとメインカメラとの水平方向の距離を3.0 m~7.0 m まで0.5 m おきに变化させて、それぞれ計測を行う。

計測結果を図 14(d) に示す。エラーバーは12個の計測箇所の標準誤差を表している。図 14(d) によると、センサから3.0 m 離れた地点での位置ずれは0.018 m であった。0.018 m という位置ずれは、4.1.2 項で示した0.07 m 以内という要求精度を満たしている。位置ずれはセンサから離れるほど大きくなる傾向があり、7.0 m の地点では最大0.038 m のずれが生じた。したがって、本項の評価実験で計測したいずれの地点においても4.1.2 項で設定した要求仕様を満たしていることが示され、背景映像のずれが十分小さい正確な映像提示を実現した。

4.4 考察

本節では、4.3.3 項で示した位置ずれについて考察を行う。位置ずれの要因としては以下の4点が考えられる。

- i) 環境をポリゴンメッシュに近似していることによる歪み
- ii) 操縦時の振動や衝撃によるセンサのずれ
- iii) 単管フレームのたわみ
- iv) センサのキャリブレーション誤差

i) について、図 10 に示す透視映像においても鋼材のテクスチャが引き伸ばされていることがはっきりと確認できるため、これが最も大きな要因として考えられる。提案手法では、ある閾値を設定し、その距離の範囲内に存在する点群で三角メッシュを生成することで、LiDAR だけでは十分に取得できない環境の 3 次元情報の補間を行っている。実機実験においては、把持対象である鋼材と地面との距離が近く、この間でメッシュが生成されたことが映像の位置ずれを引き起こしている。閾値を小さく設定することで歪みを解消することも可能だが、一方で 4.1.3 項に述べたように映像が欠けることで視認性に難が生じることも考えられる。また、ii) センサの荷重による単管フレームにたわみや iii) 衝撃や振動によって固定したセンサが動いてしまうことも考えられる。たわみの影響を最も強く受けるのは、荷重の大きさと張り出しの距離から LiDAR とカメラのセンサ系である。単管フレームの太さや材質も考慮に入れてたわみの計算を行うと、自重の影響も含めて 0.65 mm、たわみ角は 0.054 deg という計算結果となった。これは他の要因と比べて小さく、また iii) についてもセンサの取り付けを慎重に行うことで位置ずれを軽減することが可能である。iv) ではカメラの内部パラメータとカメラと LiDAR の位置姿勢関係を表す外部パラメータの 2 種類の誤差がそれぞれ生じている。これらのパラメータはチェッカーボードの位置姿勢を様々に変え、より多くの映像や点群情報を用いて最適化計算を行うことで誤差を小さくすることが可能である。また、ニブラによる把持作業では、ニブラが写る映像中央部の位置ずれが作業効率に大きく影響を与えたと考えられる。そのため、キャリブレーションでの最適化計算において重みを加えることで、映像中央部の重要度を増して各パラメータを求めることが改善案として挙げられる。

5. 結 論

本研究では、建機を遠隔操作するために、カメラ映像中の遮蔽物による死角の問題に対して透視映像を生成することで死角を低減する手法を提案した。従来研究では未解決であった屋外環境での透視映像の生成に成功し、遠隔操作に必要な映像のオンライン性と正確な映像提示を実現した。透視映像によって隠れていた背景を同時に視認可能となったことで、複数の映像を見比べることなく 1 つの映像のみで周囲の状況把握が可能となった。

今後の課題として、4.4 節で議論したように位置ずれの精度を改善することが挙げられる。また、提案した透視映像だけでは、環境によっては奥行き感覚がつかみづらい傾向がみられた。そのため、深度マップや環境の 3D モデルを提示することで 3 次元情報を明示的に示し、遠隔操作の操作性を高めることが考えられる。加えて、従来の映像提示と提案手法による透視映像提示で遠隔操作における作業効率の比較実験を行い、提案

手法が操作性の向上に有効であるか検証を行う必要がある。

謝 辞

本研究の一部は、総合科学技術・イノベーション会議が主導する革新的研究開発推進プログラム (ImPACT) の一環として実施したものである。

参 考 文 献

- 1) S. Tadokoro: Rescue Robotics Challenge, Proceedings of the 2010 IEEE Workshop on Advanced Robotics and its Social Impacts, (2010) 92.
- 2) 関川 健一, 本間 政幸: 無人化施工機械の操作性向上の検討, http://www.hrr.mlit.go.jp/library/happyoukai/h20/pdf/b_b_10hokugi.pdf (2006).
- 3) 古屋 弘, 栗生 暢雄, 清水 千春: 3D 画像と体感型操縦を用いた「次世代無人化施工システム」, 大林組技術研究所報, **76**, (2012), 1.
- 4) A. Nishiyama, M. Moteki, K. Fujino and T. Hashimoto: Research on the Comparison of Operator Viewpoints between Manned and Remote Control Operation in Unmanned Construction Systems, Proceedings of the 30th International Symposium on Automation and Robotics in Construction, (2013), 772.
- 5) F. Cosco, C. Garre, F. Bruno, M. Muzzupappa and M. A. Otaduy: Augmented Touch without Visual Occlusion, Proceedings of the IEEE International Symposium on Mixed and Augmented Reality 2009, (2009) 99.
- 6) S. Mori, F. Shibata, A. Kimura and H. Tamura: Efficient Use of Textured 3D Model for Pre-observation-based Diminished Reality, Proceedings of the 2015 IEEE International Symposium on Mixed and Augmented Reality Workshops, (2015) 32.
- 7) J. Herling and W. Broll: Pixmix: A Real-time Approach to High-quality Diminished Reality, Proceedings of the IEEE International Symposium on Mixed and Augmented Reality 2012, (2012) 141.
- 8) N. Kawai, T. Sato, Y. Nakashima and N. Yokoya: Augmented Reality Marker Hiding with Texture Deformation, IEEE Transactions on Visualization and Computer Graphics, **23**, 10, (2017) 2288.
- 9) A. Enomoto and H. Saito: Diminished Reality Using Multiple Hand-held Cameras: Proceedings of the 8th Asian Conference on Computer Vision, **7**, (2007) 130.
- 10) P. Barnum, Y. Sheikh, A. Datta and T. Kanede: Dynamic Seethroughs: Synthesizing Hidden Views of Moving Objects, Proceedings of the IEEE International Symposium on Mixed and Augmented Reality 2009, (2009) 111.
- 11) 藤井 浩光, 杉本 和也, 山下 淳, 浅間 一: 遠隔操作ロボットのための複数の RGB-D センサを用いた半隠消映像のオンライン生成, 精密工学会誌, **81**, 12, (2015), 1185.
- 12) 橘高 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, **83**, 3, (2017) 235.
- 13) 佐藤 貴亮, 藤井 浩光, A. Moro, 杉本 和也, 野末 晃, 三村 洋一, 小幡 克実, 山下 淳, 浅間 一: 無人化施工用俯瞰映像提示システムの開発, 日本機械学会論文集, **81**, 823, (2015), 14-00031.
- 14) Q. Zhang and R. Pless.: Extrinsic Calibration of a Camera and Laser Range Finder (Improves Camera Calibration), Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, **3**, (2004), 2301.
- 15) A. Kassir and T. Peynot: Reliable Automatic Camera-laser Calibration, Proceedings of the 2010 Australasian Conference on Robotics and Automation, (2010).
- 16) Z. Zhang: A Flexible New Technique for Camera Calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence, **22**, 11, (2000), 1330.
- 17) J. J. More: The Levenberg-Marquardt Algorithm: Implementation and Theory, Numerical Analysis Lecture Notes in Mathematics, **630**, (1978), 105.
- 18) Z. C. Marton, R. B. Rusu and M. Beetz: On Fast Surface Reconstruction Methods for Large and Noisy Point Clouds, Proceedings of the 2009 IEEE International Conference on Robotics and Automation, (2009), 3218.