

無人化施工のためのカメラと RGB-D センサを用いた半隠消映像の生成

杉本 和也 † 藤井 浩光 † 山下 淳 † 浅間 一 †

† 東京大学

E-mail: sugimoto@robot.t.u-tokyo.ac.jp

Abstract

災害発生時の対応作業の際には、二次災害軽減のために、バックホウを遠隔操作する技術が導入されている。バックホウを遠隔操作する際、オペレータはバックホウに搭載されたカメラの映像をもとに操作を行う。しかし作業対象物の映像がアームに遮られてしまい、作業効率が著しく低下する問題がある。そこで本研究では、アームによって遮られて見えない作業対象物を透視可能な半隠消映像を生成するシステムの開発を行なった。

1 序論

本研究では、無人化施工における遠隔操作のための半隠消映像の生成手法を提案する。災害発生時の対応作業の際には、二次災害を軽減するために、建設機械を遠隔操作する無人化施工のための技術が導入されている。特に瓦礫の撤去、土砂の採掘といった作業では、図1のような車体にアームがついたバックホウ（油圧ショベル）が用いられている。バックホウを遠隔操作する際、オペレータはバックホウに搭載されたカメラの映像をもとに操作を行う [1]。しかし、カメラ映像を用いた操作の場合、搭乗時の操作と比較して作業効率が著しく低下する [2]。この原因の1つとして、手前にあるアームに遮られて作業対象物の映像が十分に得られないことが挙げられる。

この問題を解決するためには、バックホウに複数のカメラを搭載することが有効である [3]。複数のカメラを搭載することで、死角となる領域が減少するためである。しかし多視点の映像を見比べながら作業を行う必要があり、操作に熟練を要するばかりか、長時間の作業はオペレータへの負担となり作業効率の低下を招く原因になる。そこでアームによって遮られて見えない作業対象物を1つの映像上で見えるようにすることは有効である。

近年、コンピュータビジョンの分野では、現実世界に存在する物体を映像上で隠蔽・除去する技術が注目されており、エンターテインメントの分野において数多く報告されてる [6, 7, 8]。文献 [4] では、映像上で物体を完全に消しきることなく、半透明状態にすることで

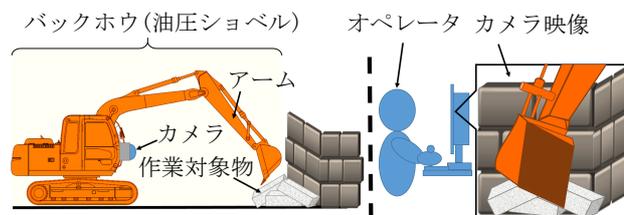


図1 バックホウを用いた遠隔操作

その物体に隠れている背景が透視できるような表現のことを「半隠消現実感」と呼称している。そこでバックホウを用いた作業において「半隠消現実感」を用いることは、アームによって遮られて見えない作業対象物を1つの映像上で確認可能であると考えられる。筆者らはこれまでに複数カメラとロボットアームを用いて、カメラ映像上のロボットアーム部分を半透明にし背景が透視可能な半隠消画像を試みてきた [5]。しかし文献 [5] では、静止画像から動画への応用は出来ておらずリアルタイムに半隠消映像を作成することが不可能であった。そこで本稿ではカメラと RGB-D センサを用いた際に、半隠消映像をリアルタイムに生成する手法について述べる。

2 提案手法

本研究で半隠消映像を生成する際、対象となる環境の3次元計測を行う必要がある。3次元計測を行う手法としては、ステレオ視を用いる手法が存在する。ステレオ視を用いる場合は各カメラ映像の対応点を求める必要がある。しかし土木作業の現場では、土砂や瓦礫といった色の変化に乏しい物体が背景物体となるため、各カメラ映像の対応点を求めることが困難になる。

次に、レーザセンサを用いる手法が存在する。しかしレーザセンサは高価なことが多く、緊急を要する災害対応には不向きである。また、比較的安価な平面照射型のレーザセンサを動かし、3次元環境を測定する手法も存在するが、バックホウの激しい振動に耐久可能なシステムを作るのは困難を極める。

そこで本研究では、RGB-D センサを用いる。RGB-D センサとは、センサ1台で RGB の色情報に加えてセンサから対象物までの距離情報が取得可能なセンサであ

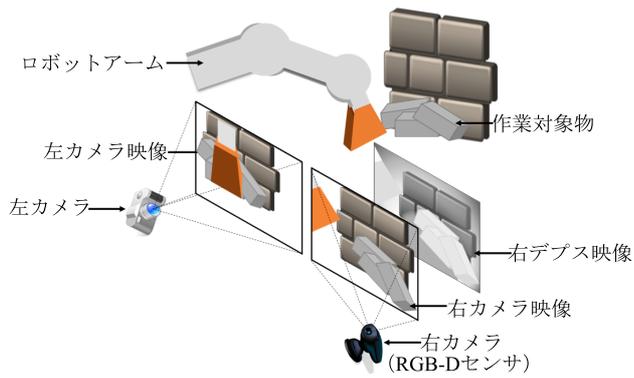


図2 実験環境の概念図

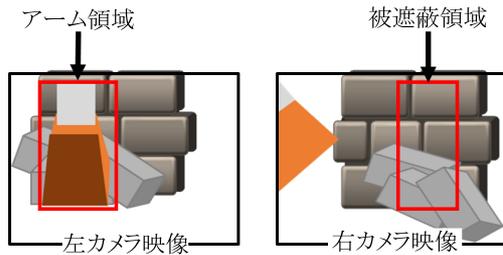


図3 各領域の名称

る。ステレオカメラ、レーザセンサと比較してより安価で簡便に背景物体の3次元位置計測が可能である。

実験環境の概念図、およびカメラとRGB-Dセンサの位置関係を図2に示す。カメラおよびRGB-Dセンサは必ず視差が得られるように設置する。本研究では簡単のために、カメラを左カメラ、RGB-Dセンサを右カメラと呼ぶことにする。この際、左右のカメラから得られるカラー映像をそれぞれ左カメラ映像、右カメラ映像と呼ぶことにする。さらに右カメラ (RGB-Dセンサ) から得られる距離情報を含む映像を右デプス映像と呼ぶことにする。また左カメラ映像上に映し出されているロボットアーム部分をアーム領域とし、右カメラ映像中でアーム領域の背景に対応する部分を被遮蔽領域とする (図3)。

以上の環境において、半隠消映像を生成するためには以下の3つの処理を行う必要がある。

- (1) アーム領域の取得
- (2) 被遮蔽領域の取得
- (3) 半隠消映像の合成

2.1 アーム領域の取得

左カメラ映像中のアーム領域は、アームの各関節角度を取得することで追跡が可能になる。ロボットアームの各関節角度を随時取得することで、あらかじめ用意したロボットアームの形状データに基づく順運動学計算により左カメラ映像中のアーム領域の取得を行う。

図4のように設定したワールド座標系において、ロボットアームの3次元形状でデータお

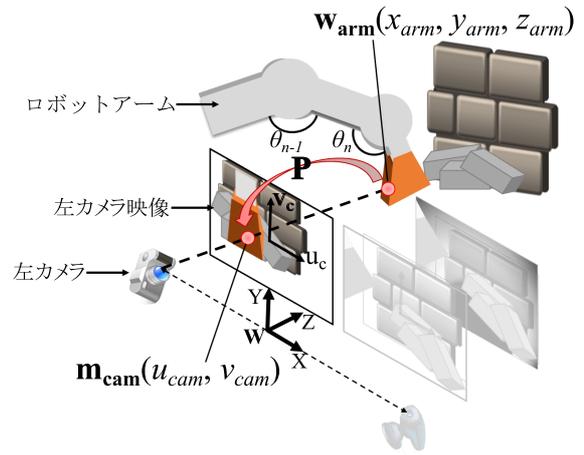


図4 アーム領域を取得する際の概念図

よび各関節角度から計算可能なアーム位置座標 $\mathbf{W}_{arm} = [x_{arm}, y_{arm}, z_{arm}, 1]^T$ が、左カメラ映像の画像座標系において $\mathbf{m}_{cam} = [u_{cam}, v_{cam}, f_{cam}, 1]^T$ に像を結んだとき、行列 \mathbf{P} を用いると以下のような関係がある。また、 \mathbf{P} は左カメラの内部パラメータおよびセンサの位置・姿勢を表す行列である。

$$\mathbf{m}_{cam} \simeq \mathbf{P}\mathbf{W}_{arm} \quad (1)$$

ここで \simeq は、左辺は右辺の0でない定数倍と等しいことを意味する。上式を用いて \mathbf{m}_{cam} を求めることにより、左カメラ映像中のアーム領域を取得する。

2.2 被遮蔽領域の取得

被遮蔽領域を右カメラ映像より求める。概念図を図5に示す。左右のカメラの光軸方向をそろえ、視点の高さも等しく設置することで、平行ステレオ [9] を用いた三角測量の原理による位置計測が可能となる。そのため2.1節で取得したアーム領域の画像座標 $\mathbf{m}_{cam} = [u_{cam}, v_{cam}, f_{cam}, 1]$ に隠れている背景物体の位置座標 $\mathbf{W}_{back} = [x_{backi}, y_{backi}, z_{backi}, 1]$ が、右カメラ映像上で $\mathbf{m}_{snsr} = [u_{snsri}, v_{snsri}, f_{snsr}, 1]$ に像を結ぶとき以下の関係式が成立する。

$$x_{backi} - \frac{b(u_{cam} + u_{snsri})}{2(u_{cam} - u_{snsri})} = 0 \quad (2)$$

$$y_{backi} - \frac{bv}{(u_{cam} - u_{snsri})} = 0 \quad (3)$$

$$z_{backi} - \frac{bf}{(u_{cam} - u_{snsri})} = 0 \quad (4)$$

f は左右のカメラの焦点距離であり $f = f_{cam} = f_{snsr}$ である。また、 b は左右のカメラ間の距離であり $v = v_{cam} = v_{snsr}$ である。添え字の i は右映像上の画像座標において $v_{snsr} = u_{cam}$ のライン上における u 軸の値でありピクセル単位である。右カメラ映像の各ピクセルは、RGBの色情報と任意のワールド座標系における位置座標情報を兼ね備えている (図5)。そのため

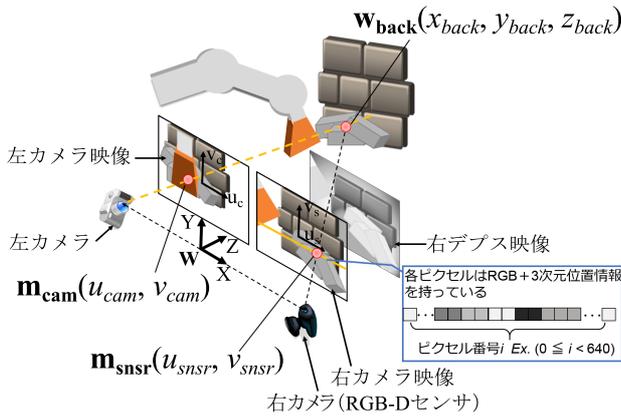


図5 被遮蔽領域を取得する際の概念図

このラインに沿って、式(2)~(4)の左辺の絶対値を最小にする i の探索を行うことでアーム領域に隠れている被遮蔽領域の取得が可能となる。

2.3 半隠消映像の生成

2.1節で取得したアーム領域を半透明にし、2.2節で得られた被遮蔽領域と共に左カメラ映像に投影および合成処理を行う。投影位置は2.1節で求めた左カメラ映像上のアーム領域とする。本研究では、2.1節で取得したアーム領域を半透明にし2.2節で取得した被遮蔽領域と合成処理を行う際に以下の式を用いる。 I は各映像の各ピクセルにおけるRGBの色成分である。また α は不透明度を表す係数である。 $\alpha = 0$ で完全な透明となり、 $\alpha = 1$ で不透明な状態を示す。

$$I_{output} = \alpha I_{arm} + (1 - \alpha) I_{back} \quad (5)$$

3 検証実験

3.1 実験概要

前章で述べた処理を用いて、視差のある2つ映像からアーム領域の半隠消映像を生成する。バックホウを遠隔操作する際に行う作業としては、前方の瓦礫の撤去などが存在する。そこで本実験では、バックホウを用いた瓦礫の撤去作業を想定し、図6のように2台のRGB-Dセンサ (ASUS: Xtion Pro Live) およびロボットアーム (YASKAWA: MOTOMAN-HP3J) を用いて、前方の瓦礫にアクセスする際にロボットアームに隠れて見えない瓦礫が透視可能な映像の生成を行う。同じRGB-Dセンサを2台用いているが、提案手法と同様、一方のRGB-Dセンサを左カメラ、もう一方のRGB-Dセンサを右カメラと呼ぶことにする。またこの際、左カメラを通常のカメラとして、右カメラをRGB-Dセンサとして用いた。

実験環境の座標系および機器の位置・姿勢を図6に示す。Z軸方向に150mm、Y軸方向に250mmの位置に左カメラを設置し、さらに左カメラからX軸方向に

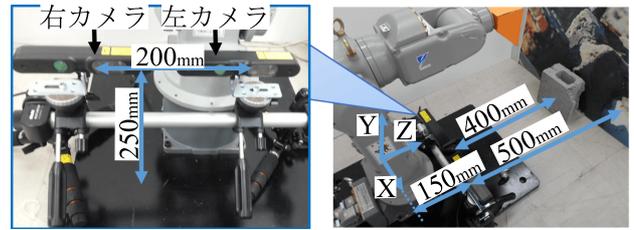
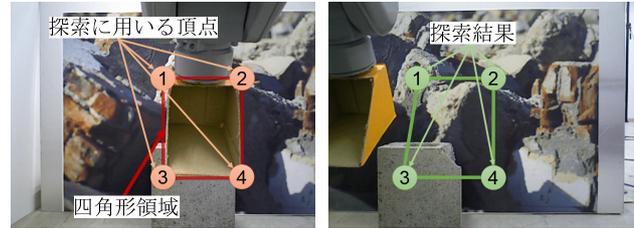


図6 実験環境図



(a):左カメラ映像

(b):右カメラ映像

図7 入力画像

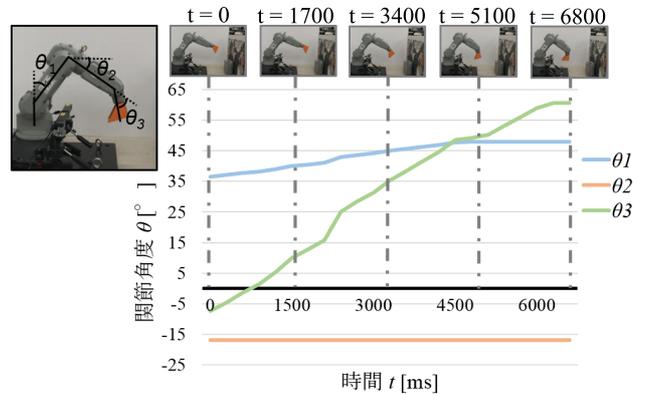


図8 各関節角度の変動図

200mm 離れている位置に右カメラを設置した。取得映像サイズは 640×480 pixel とした。

実際の作業現場で散在した瓦礫の撤去作業を行うことを想定し、瓦礫のモデルとして擬似的に再現した異なる2つの位置・形状の物体を用いた。1つ目の物体としては画像が貼り付けられている幅1000mm、高さ800mmの板状のボードとした(図6)。2つ目の物体としてはボードの手前に置かれた幅200mm、高さ200mm、奥行き100mmの直方体形状の瓦礫とした(図6)。

使用したロボットアームモデルに対し、図8のように各関節角 $\theta_1 \sim \theta_3$ を設定した。ロボットアームが前方の瓦礫にアクセスする動作を想定し、 $\theta_1 \sim \theta_3$ を図8のように変動させた際の各カメラ映像を入力映像とした。

各カメラからの入力映像を図7に示す。図7(a)が左カメラから得られる映像であり、図7(b)が右カメラから得られる映像である。本実験ではこれらの映像の内、左カメラ映像上のアーム領域から成る四角形領域(図7(a))を半透明にすることで半隠消映像を生成した。

この際、対象となる四角形の4頂点が右カメラ映像上でどの位置に対応するかを2.2節の手法を用いて探索を行い、対応する点から成る四角形を被遮蔽領域と

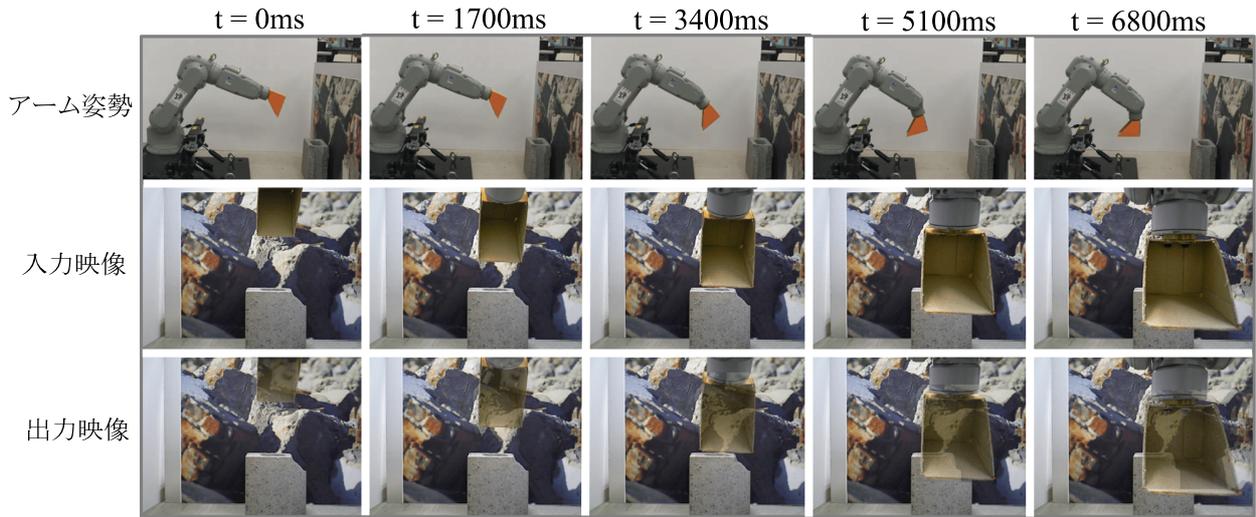


図9 アームの姿勢および入力画像，出力画像群

した(図7(b)). その後, 得られた被遮蔽領域を左カメラ映像中の四角形領域と一致するようにリサイズし合成することで半隠消映像を取得した. また(5)式における透過値を表す係数 α は, 半透明にする全ての領域で0.5とした.

3.2 結果と考察

ロボットアームが前方の瓦礫にアクセスする際の, 時間 t [ms]におけるロボットアームの姿勢, 入力映像(センサ1映像), 出力映像を図9に示す. これら画像群は, 初期位置を $t = 0$ msとし $t = 6800$ msに向けて, 各関節角度を図8のように変化させた様子である. 各フレームの入力画像がら出力画像間を生成する際の処理時間は平均は22msであった. この結果, アーム領域から成る四角形領域に隠れて見えない被遮蔽領域を透視可能な半隠消映像がリアルタイムに生成可能となった.

一方, 例えば図9の $t = 3400$ msおよび $t = 6800$ msの出力映像の比較より, 背景の煉瓦が歪んでいることが確認できる. この原因としては, 図7に示したように, アーム領域から成る四角形領域の頂点情報のみを用いて被遮蔽領域を取得しているためである. この問題を解決するためには, 被遮蔽領域を取得する際に用いるアーム領域上の点を増加させることで, より自然な半隠消映像が取得可能だと考えられる.

4 結論

本稿では, カメラとRGB-Dセンサを使用し, カメラ映像中のアーム領域を半透明にすることでアームに隠れている背景が透視可能な半隠消映像のリアルタイム生成手法を提案した. また本手法を用いることで, 目的とする半隠消映像がリアルタイムに生成可能なことを確認した. 今後は, 生成した半隠消映像を用いた際の遠隔操作性の評価を行う.

参考文献

- [1] 藤本 昭, 松岡 雅博, 茶山 和博, 藤岡 晃: “遠隔操縦ロボット(ロボQ)の開発”, 土木学会第58回年次学術講演会概要集, Vol. 58, No. 6, pp. 61–62, 2003.
- [2] 山口 崇, 石松 豊, 山元 弘: “無人化施工のマンマシンインターフェイスに関する調査”, 平成16年度建設施工と建設機械シンポジウム論文集, Vol. 18, pp. 145–148, 2005.
- [3] 古屋 弘, 栗山 楊雄, 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, No. 76, 2012.
- [4] 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みと諸問題”, 日本バーチャリアリティ学会誌, Vol. 16, No. 2, pp. 239–250, 2011.
- [5] 杉本 和也, 藤井 浩光, 佐藤 貴亮, 山下 淳, 浅間 一: “無人化施工のためのカメラとRGB-Dセンサを用いた半隠消映像の生成”, 2013年度精密工学会春季大会学術講演会講演論文集, pp. 991–992, 2014.
- [6] A. Enomoto and H. Saito: “Diminished Reality Using Multiple Handheld Cameras”, Proceedings of the 8th Asian Conference on Computer Vision 2007, pp. 130–150, 2007.
- [7] B. Avery, C. Sandor and B. H. Thomas: “Improving Spatial Perception for Augmented Reality X-Ray Vision”, Proceedings of the IEEE Virtual Reality Conference, pp. 79–82, 2009.
- [8] 清水 直樹, 橋本 章宗, 植松 裕子, 斎藤 英雄: “デブスカメラを用いたリアルタイム領域抽出による隠消現実感映像生成”, 映像情報メディア学会誌, Vol. 66, No. 12, pp. J549–J552, 2012.
- [9] 出口 光一郎: ロボットビジョンの基礎, コロナ社, 2000.