

非整列多視点カメラを用いた画像ベース隠背景可視化

森 尚平[†] 前澤 桃子[†] 家永 直人[†] 斎藤 英雄[†]

Shohei MORI[†], Momoko MAEZAWA[†], Naoto IENAGA[†], and Hideo SAITO[†]

[†] 慶應義塾大学 〒 223-8522 横浜市港北区日吉 3-14-1

E-mail: †{mori,momoko_maezawa,ienaga,saito}@hvrl.ics.keio.ac.jp

あらまし 本稿では、非整列多視点カメラを用いて隠された作業領域を可視化する隠消現実感 (Diminished Reality; DR) システムについて述べる。本システムは、環境に設置した非整列多視点カメラを用いてユーザ視点を再構成する際、RGB-D カメラから得られた対象領域内において、非整列多視点カメラからの画素を参照するかを自動計算する。

キーワード 隠消現実感, イメージベースレンダリング, 多視点カメラ

1. はじめに

作業中に用具や機材や自身の手や把持するツールによって視界が遮られることで、作業対象へのアプローチが難しい場合がある。例えば、電子工具に作業対象が遮られる、プラグを把持した自身の手で差込口が隠れる、といった状況である。こうした問題は、AR 表示 [1] や別視点映像の提示によって解決できる可能性があるが、前者は、付加情報を提示するだけであるため、実際の作業対象が見えるようになるわけではなく、後者は、鏡を用いるのと同じで脳内での視点変換が必要となる。これを助けるために、別視点映像を変換する方法もあり、提案システムはこの方法に最も関連している。

別視点映像を変換する方法では、平面の射影変換には歪みが発生し、空間認識が困難になるという限界があるため [2], 多視点カメラを用いる方法が有効である。合成開口法を利用した任意視点画像生成法 [3] では、整列させた数十台のカメラを用いて撮影することで、撮影後の合焦点画像生成や背景の可視化を可能にする。ただし、同期システムや画像の転送の装置が大掛かりになり、撮影システム自体の構築・移動・設置が難しい、といった実用上の欠点が存在する。

そこで、本稿では、数台のカメラのみで構成できる、作業領域可視化用隠消現実感システムを提案する。本システムでは、多視点カメラは概ね作業対象の方を向いていればよく、そのカメラが必要な背景情報を捉えているかどうかはシステム実行時に自動計算されるため、ユーザには事前のカメラの設置と校正以外の負荷を与えない。

2. 提案システム

2.1 多視点カメラの配置と校正

提案システムでは、RGB の多視点カメラに加え、1 台の RGB-D カメラを用いる。よって、ユーザはまず、この多視点カメラ D_m を環境中に配置して作業対象の撮影を開始する。配置した多視点カメラから画像を取得し、カメラの校正を行うことで、各カメラの内部パラメータと外部パラメータを計算しておく。尚、この事前処理には、Agisoft 社製 PhotoScan [4] 等のバンドル調整ツールを利用すればよい。

2.2 前背景の分離

逐次処理では、まず、奥行き画像 I^D から前景と背景を分離する。前景は、除去対象領域の決定 (2.3 節)、背景は、隠背景の再構成 (2.4 節) に利用する。システムの実行を開始してから数十フレームには背景のみが写っている、として、その重み付き平均画像を、初期の背景画像とする。

$$I^D(x, y) = \frac{\alpha I_{curr}^D(x, y) + I_{prev}^D(x, y)}{\alpha + 1} \quad (1)$$

ただし、 α はユーザが与える任意の重み、 I_{curr}^D は取得した現フレームにおける奥行き画像、 I_{prev}^D は 1 フレーム前に得られた奥行き画像の重み付き平均の結果である。よって、初期フレームは $I^D = I_{curr}^D$ とし、次フレームでは、 $I_{prev}^D = I^D$ として、同様の処理を行う。

その後のフレームでは、以下に示す I_{sub}^D によって、画素ごとに背景と前景とを分離する。

$$I_{sub}^D = |I_{curr}^D - I^D| - tolerance \quad (2)$$

$$I^M(x, y) = \begin{cases} 1 & (I_{sub}^D(x, y) > 0) \\ 0 & (othersize) \end{cases} \quad (3)$$

ただし、 $tolerance$ は、ユーザが設けるフレーム間に許容する変化量、 I^M は前背景をそれぞれ 0, 1 で表すマスク画像である。

尚、背景領域と判別された奥行き画像の画素 $I_{curr}^D(x, y)$ は、式 1 によって、逐次更新され、それ以外の画素は I^D から得る。

2.3 除去対象領域の決定

隠背景可視化法には、先行研究にて開発した手法 [X] を用いるが、先行研究では、ユーザ視点画像全領域を再構成していたため、全領域に渡って任意視点画像生成によるアーティファクトが生じることがあった。これを防ぐため、本システムでは、除去対象領域と背景領域を明示的に分けて処理する。

2.2 節で分離した前景画像 $I^M(x, y)$ を、そのまま除去対象領域のマスク画像として用いる場合、奥行き画像中の欠けによって、除去対象物体を完全には切り出せない場合が多い。そこで、前景画像 $I^M(x, y)$ の内、画素値 1 の 2 次元点の集合に対して楕円フィッティングを行い、検出された楕円を 0 それ以外を 1 として、除去対象領域を決定する。

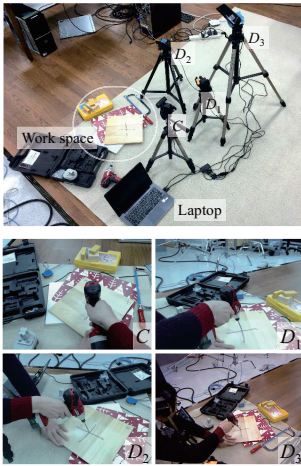


図1 セットアップの例



図2 入力フレーム（上段）と隠消現実感表示（下段）

2.4 多視点画像を用いた隠背景再構成

先述の通り、本システムでは、隠背景可視化法として、先行研究にて開発した手法 [5] を用いる。具体的には、文献 [6] で提案された任意視点画像生成法におけるジオメトリックプロキシ (GP) を、RGBD カメラを用いて実時間で取得する。これに加えて、カメラブレンディングフィールド (CBF) に、除去対象物体を代表する 3 次元点を与え、負の重みとして加え、除去対象物体を含む画素が隠背景再構成に利用されないようにする。

加えて、GP に相当する 3 次元点をユーザ視点に投影し、除去対象領域内であれば、CBF 計算時に 0 の重みを与えることで、この GP が隠背景再構成に利用されないようにする。また、2.3 節で検出された楕円の中心位置に相当する 3 次元位置を奥行き画像から取得し、これを除去対象物体の代表点とする。

3. 実装

3.1 システム構成

以降、ノート PC (CPU: 3.30GHz, RAM: 16GB, GPU: Intel Iris Graphics 550) 1 台、三脚に固定した 3 台の Logicool 社製 USB カメラ 920t を 3 台と Microsoft 社製 Kinect (RGB-D カメラ) 1 台を用いた結果について述べる。それぞれのカメラを D_i ($0 \leq i \leq 3$) (内、 D_0 はユーザ視点 C) とし、画像サイズは全て 640×480 画素とした。図 1 に、本実装で用いた機器構成と得られる入力フレームの例を、提案システムのセットアップ例として示す。開発には、Visual C++ と OpenGL Shading Language 3.3 を用いた。隠背景の再構成には、4 台全てのカメラからの画像を毎フレーム合成した。

3.2 想定シナリオ

図 1 に示す環境下で、ユーザは電動ドリルを把持し、作業空間に置かれた木版に穴を開ける、というシナリオを用意した。穴を開ける箇所にはバツ印を描いた。木版のこの印に対して垂直になるように電動ドリルを当てると、ユーザ視点 C から印が隠されてしまうため、これを可視化する。

3.3 実行結果

図 2 に入力フレームと隠消現実感表示を行った結果を示す。再構成した隠背景の画像をユーザ視点 C の画像に α 合成して、

手を半透明表示した。背景の板や左手が動く動的な空間において、右手と把持した電動ドリルの領域に、作業対象の木版とそこに描かれたバツ印が半透明に表示されている。この結果から、4 台という既存システムと比較して少ない台数かつ非整列のカメラを用いた提案システムが、機能していることが確認できた。

奥行き画像を用いることで、動的かつ非平面な隠背景を再構成できる本システムであるが、一方で、奥行き画像の精度に起因する限界が顕著に現れた。例えば、システム実行中に、一度も奥行き値が取得できなかった画素が存在する場合、その場所に想定する画素が復元できない。こうした画素は、単に RGBD カメラの画角外であるものだけでなく、RGBD カメラからの死角や計測物体の材質にも起因しているため、画像全体に渡って隠消現実感表示できない画素が存在した。

4. むすび

本稿では、手や把持したデバイスによって遮られる作業空間を非整列カメラを用いて可視化する隠消現実感技術型システムについて述べた。背景、除去対象共に同物体を含むシナリオにおいて、本システムが機能することを実データを用いて示した。

謝辞 研究の一部は学振・特別研究員奨励費 (課題番号 16J05114) 及び科学研究費基盤研究 (S) 24220004 の補助により行われた。

文献

- [1] R. T. Azuma: Recent advances in augmented reality, IEEE Computer Graphics and Application, Vol. 21, pp. 34 – 47, 2001.
- [2] M. Goto, Y. Uematsu, H. Saito, S. Senda, and A. Iketani: Task support system by displaying instructional video onto AR workspace, Proc. ISMAR, pp. 83 – 90, 2010.
- [3] B. Wilburn, N. Joshi, V. Vaish, E.-V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy: High performance imaging using large camera arrays, ACM Trans. on Graphics, Vol. 24, No. 3, 2005.
- [4] Agisoft PhotoScan, <http://www.agisoft.com/>
- [5] 前澤桃子, 家永直人, 森尚平, 齋藤英雄: 多視点カメラを用いたライトフィールドレンダリングによる非平面隠背景可視化, 日本 VR 学会複合現実感研究会, MR2017-1, 2017.
- [6] C. Buehler, M. Bosse, L. McMillan, S. Gortler, and M. Cohen: Unstructured lumigraph rendering, Proc. SIGGRAPH, pp. 425 – 432, 2001.