



隠消現実感における両眼視野不整合についての分析と対策

松木 ひとみ¹⁾, 森 尚平¹⁾, 柴田 史久¹⁾, 木村 朝子¹⁾, 田村 秀行²⁾

1) 立命館大学大学院 情報理工学研究科 2) 立命館大学 総合科学技術研究機構

(〒525-8577 滋賀県草津市野路東 1-1-1)

概要: 複合現実感 (MR) が現実世界と仮想世界との融合を目指すのに対し, 隠消現実感 (DR) は現実世界から物体を視覚的に隠蔽・消去することを目指す. DR も MR 同様に Head-Mounted Display (HMD) を用いた実地体験型のシステムが望まれるが, これを実現した例は少なく, 両眼立体視に関する問題把握や考察はされてこなかった. 本研究では, そういった問題を両眼視野不整合として分析し, 実験システムを構築して実証した. また, その対策についても言及し, 同様に実証した.

キーワード: 複合現実感, 隠消現実感, 両眼立体視, Image-Based Rendering

1. はじめに

隠消現実感 (Diminished Reality; DR) は, 現実世界に存在する不要な物体を視覚的に隠蔽・消去・透過させる技術である. DR は 1994 年に S. Mann によって提唱された Mediated Reality[1] という概念的枠組みに属し, 現実世界に仮想物体を重畳し現実感を拡張する拡張現実感 (Augmented Reality; AR) や複合現実感 (Mixed Reality; MR) とは逆の概念である. 現実世界への情報の加算を行う AR/MR だけでは表現力に限界があるため, 減算する DR 技術と一緒に利用する必要がある.

DR の実現方法は種々存在する. 例えば, Selective Photographyのごとく光景から不要な色をなくす, 不要物体を仮想物体で覆い隠す, 背景を推定して不要物体を覆い隠すことで視覚的に物体を除去して見せる等である. 特に, 物体を除去する DR 技術は, 対象領域の周辺から背景を推定する Image Inpainting-Based 法と観測を伴って背景を再構成する Observation-Based 法に大別される. 中でも Observation-Based 法は AR/MR における挑戦的課題として注目されており, 本研究の研究対象に当たる.

AR/MR と同様に没入感の高い DR 体験を実現するためには, 当然 Head-Mounted Display (HMD) を用いた実地体験できる形式が好ましい. にもかかわらず既存の Observation-Based DR (OB-DR) においては, 単眼カメラを用いたビデオシースルー (Video-Seethrough; VST) 型のシステムの提案が多く, VST-HMD を用いて両眼立体視可能なシステムの実現例が見られない. 即ち, OB-DR の両眼立体視に関する問題は事実上未確認である.

本論文では, OB-DR による処理結果を両眼立体視可能な VST-HMD に表示した際に生じる問題を両眼立体視野

不整合と呼称し, 分析及び対処法に関して検討する. 具体的な手順は以下の通りである.

- ・ 既存 OB-DR システムを例にとった, 両眼視野不整合の発生要因の特定と整理
 - ・ VST-HMD を通して両眼立体視が可能な OB-DR システムを構築し, 両眼視野不整合を具体的に例示
 - ・ 両眼視野不整合の対処法の検討と例示
- 尚, 特に断りのない限り, DR は OB-DR を, HMD は VST-HMD を指すものとする.

2. 事前観測ベースの DR での両眼立体視

先述の歴史的背景, 既存の DR システムで用いられる機材や技術といった面から考えて, DR システムは AR/MR システムの拡張形態であると捉えることができる. よって, そのシステム構成には類似点が多く, HMD を利用した DR システムは, 以下の 4 つの手順に則って実装できる. また, 各ステップで発生する誤差が両眼立体視野不整合となって現れる.

背景の観測とデータ化: DR では, 不要物体によって隠される光景 (隠背景) を事前ないし実時間で観測する必要がある. 事前か実時間かは状況によって異なる. 例えば, 除去対象が現時点で存在する場合はその背後を多数の視点から実時間観測する[2, 3], 望ましい光景を邪魔するように不要物体が置かれても後にその光景を取り戻せるよう, カメラで事前観測する[4, 5]といった方法が考えられる.

このデータ収集は DR 体験とは独立しているため, ここでは HMD を用いる必要はない. 隠背景の観測結果は, その再構成方法に応じて, テクスチャ付き 3 次元モデル, ボクセル[6], 簡易な 3 次元幾何モデルと視点依存画像群[4]といった形式で記述される. これを隠背景データと呼ぶ.

カメラ, シーン, 不要物体のトラッキング: 目的に応じて, カメラ, シーン, 不要物体のトラッキングを行う. これま

で AR マーカ[3, 4]や特徴点ベースの手法[5]を利用する DR 手法が提案されているが、物理センサを利用する方法も考えられる。他にも、HMD に装着されたステレオカメラを利用した、ビジョンベースのカメラ位置姿勢推定法を実装することも考えられる。

不要物体の認識とマスク画像の生成: 既存の DR 手法では、除去対象領域を決定するために、除去対象領域とそれ以外とを区別するマスク画像を生成する。これまでに、手動での切り出し[2]、3次元バウンディングボックス[4]、半自動抽出法[5, 7]を用いる手法が提案されている。

隠背景の再構成と現画像への合成: 隠背景データとマスク画像を利用して現画像の除去対象領域に再構成した隠背景を重畳する。隠背景は、ホモグラフィ変換を用いた画像の張り合わせ[5]、Image-Based Rendering (IBR) [4]等で再構成される。体験者はこの結果を HMD を通して、両眼立体視する。

3. OB-DR における両眼視野不整合

本章では、両眼立体視野不整合が発生する具体的な原因について分析する。その原因を明確化するため、図 1 に実画像、再構成画像、マスク画像の関係図を示す。また表 1 に両眼視野不整合と要因、次章にて述べる解決策の一覧を示す。

3.1 問題の明確化

ステレオカメラ較正の誤差: ステレオカメラ較正が正しく行われていない場合、除去対象領域がシーンに対して浮き出る、引っ込むといった不整合が発生する。また、歪み補正值や内部パラメータが正しく得られていない場合、ビジョンベースのカメラ位置姿勢推定がうまく行われなかったり、プロジェクション行列が正しく得られないため隠背景の再構成結果を合成する際に幾何学的不整合が発生したりする。しかしながら、HMD のステレオカメラ較正は事前に十分に時間をかければ解決できる問題である。

ステレオ画像表示の非同期: 両眼用のマスク画像生成、隠背景再構成、実画像表示は全て同期している必要がある。この同期がずれてしまう可能性は以下の 3 つの組み合わせについてである。

- ・ 実画像: 同期していないカメラ、ディスプレイを用いることで発生するが、然るべきハードウェアを利用すれば解決できる問題である
- ・ マスク画像/再構成結果: 任意に非同期にしない限り発生しない。例えば、実時間性を保つために、片方の処理だけで善しとする等が考えられる
- ・ 実画像とマスク画像/再構成結果: 時間的整合性の問題として知られる。カメラトラッキング、除去対象の追跡、マスク画像生成、隠背景の再構成の遅れが原因で発生する

隠背景の再構成精度: 隠背景の再構成結果が不正確な場合、両眼視野不整合を発生させる要因となる。具体的には、下記の要因が隠背景の再構成精度に影響を与える。

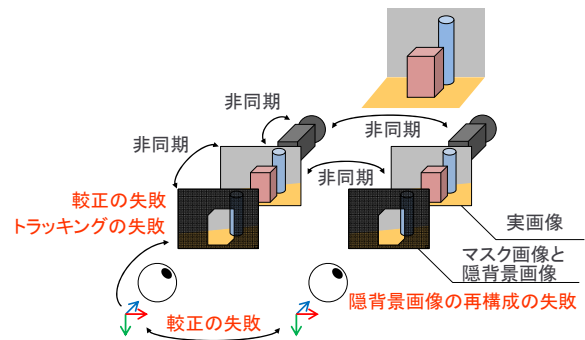


図 1 実画像、再構成画像、マスク画像の関係図

- ・ カメラ位置姿勢推定誤差: DRにおける主要な問題の一つであり、Computer Vision (CV) 及びAR/MR分野にて広く研究されている
- ・ 背景の計測及び再構築の誤差: カメラ位置姿勢推定と同様、DRにおける主要な問題の一つであり、この幾何学的不整合は、後続の隠背景の再構成に影響を与える
- ・ 照明や光学系の違いによる光学的不整合: 照明推定、カメラやシーンの光学的モデリングが不正確な場合、再構成された隠背景と実シーンとの間で光学的不整合が発生する

マスク画像の生成: HMDを用いたDRでは、生成されるマスク画像間で一貫性を保てない場合がある。

- ・ スクリーンスペースでの物体認識: 左右画像上で個別に除去対象の物体認識処理された場合、両者の結果が大きく異なる可能性がある
- ・ 3Dバウンディングボックス: カメラ位置姿勢推定法に起因する位置ずれが発生する。しかし、3次元幾何形状を利用するため、左右で大きく異なる結果とはなり難い

3.2 マスク画像の窓を通して見た隠背景再構成結果

実際問題として、両眼視野不整合は上記の原因が複数組み合わせられて発生する。ただし、初めの 2 つの問題は然るべきハードウェアを利用するか準備段階で時間をかければ解決できる問題である。よって、ここでは、残りの問題について議論する。特に、DR ではマスク画像という窓を通して隠背景再構成結果を見ることになる点に注目する。

両眼視野闘争: 2章で述べた通り、除去対象は除去対象領域、つまり、マスク画像によって切り抜かれた窓内に生成される隠背景の再構成結果で上書きされる。ここに発生する覗き窓（隠背景の再構成結果）とその周辺領域（実画像）との差異は理想的にはないが、実際は幾何学的/光学的不整合によって見た目には差が生じる。仮に、覗き窓から見える世界が両目用の各画像で全く異なる場合、覗き窓内全体で両眼視野闘争が発生するし、程度が低くとも幾何学的/光学的不整合がその領域を明らかにする。

このように画面内に覗き窓を設けることは、Window Violation を引き起こす。つまり、3D 映画と同様に、左右の画像に互いに存在しない部分が含まれるため、両眼視野闘争が発生する[8]。3D 映画の場合、観客の視界の両

表 1 OB-DR における両眼視野不整合

処理手順	両眼視野不整合	原因	解決策
ステレオカメラ較正	除去対象領域における不自然な立体感の知覚	不正確な較正	時間をかけた事前の較正
左眼/右眼用画像の同期	両眼視野闘争	ステレオカメラ/ディスプレイの非同期	適切なハードウェアの利用
隠背景画像の再構成	隠背景の覗き窓における Window Violation	幾何学的不整合	RBAB 処理 (4.3 章)
	3D バウンディングボックスの露出 (MaMo 効果)	光学的不整合	
	立体感の切り換わり	疎な IBR (例: VDTM)	密な IBR (例: LFR)
マスク画像の生成	両眼視野闘争	2D 物体の追跡	3D バウンディングボックス

端でこの問題が発生する一方で、DR の場合、除去対象領域周辺でも発生する。

この問題を軽減するために、除去対象領域とその周辺領域で幾何学的整合性を保つ必要がある。例えば、Light Field Rendering (LFR) といった高精度な再構成法を利用する方法が有効である。他にも、領域間の境界を目立たせなくする処理を実装することも解決策となり得る[5]。

立体的に露呈する窓：仮に幾何学的整合性が達成されたとしても、光学的整合性によって、依然、覗き窓が露呈する。例えば、照明位置が変化して、シーンの陰影が変化することで、隠背景の再構築結果と実画像の切れ目が目立ち、カニツアの錯視のように除去対象領域が認識できる。更に、両眼用のマスク画像にはしばしば視差が存在するため、除去対象領域が立体的に浮き出て見えてしまう。我々は、この DR 特有の現象を MaMo 効果と名付けた。

IBR による奥行知覚の変化：最新の CV 技術を用いても高品質なテクスチャ付き 3 次元モデルの再構成はいまだに困難な課題である。こうした背景から、[4]のように、IBR を基礎とした隠背景の再構成法を用いることが DR に適した方法だと考えられる。特に LFR は、複数枚の視点依存画像をブレンディングして任意視点画像を再構成するため、対象シーンの幾何形状に関係なく高品質な隠背景を再構成できる。

IBR 手法の種類や、視点依存画像の枚数、シーンの幾何形状の近似度合は、隠背景投影画像における観測者の奥行知覚に影響する。中でも、IBR 手法の一種である View-Dependent Texture Mapping (VDTM) のように、少ない枚数の視点依存画像を利用する場合は、観測者の視点位置により、用いる視点依存画像が切り替わるので、隠背景投影画像の視差が変化してしまう場合がある。

4. 実装と実験

分析した両眼視野不整合の内、下記の 2 つを確かめる。

実験 1：IBR による奥行知覚の変化

実験 2：立体的に露呈する窓

4.1 実験用システムの構築

本実験では、輻輳点が適切に設定された COASTER 方式(ディスプレイとカメラの光軸が一致する方式)の HMD (Canon HH-A1) に表示した結果を目視で確認した。即ち、

実際の DR 体験を模倣して、体験者の視界を HMD で覆い、フルカラーの映像で確認した。

実験用システムでは HMD の代わりに単眼カメラを取り付けたロボットアームを制御することによりステレオ画像を取得する。実験用システムは PC、ロボットアーム (Denso Wave 社製 VS-087)、単眼カメラ (Flea 3 GigE)、照明 (太陽光として ARRI 社製 HMI575、環境光として 2 台の ARRI 社製 650Plus) で構成した。HH-A1 の構成に合わせて、カメラの解像度は 640×480 画素であり、画角 40.5° (H)、焦点距離 6.5mm のレンズを使用した。

図 2 に、撮影シーンとして構築した 1/12 スケールのミニチュアセットを示す。除去対象は三角コーン及びヘルメットとした。尚、カメラ及び HMD は事前に較正されており、それぞれの内部パラメータと歪み補正值、左右カメラへの座標変換行列は既知である。撮影の最中はホワイトバランスの変更は行わない。

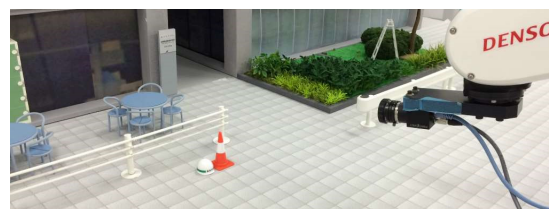


図 2 1/12 スケールのミニチュアセットを用いた実験システム

4.2 実験 1

隠背景データとして、極座標系で記述した球面 Light Field ($r: 400 \text{ mm}$, $-2.5^\circ \leq \phi \leq 2.5^\circ$, $0^\circ \leq \theta \leq 60^\circ$, 200×20 のカメラアレイ) である LF_A を構成した。これは 2 章で述べた「背景の観測とデータ化」の段階に相当する。続いて LF_A から視点依存画像を 5 枚毎に 1 枚に間引いて LF_B を作成した。続いて、除去対象をシーン中に置き、200 枚のステレオ画像列 SI_1 を記録した。各フレームに対応しカメラ位置姿勢はロボットアームから取得した (2 章の「カメラ、シーン、不要物体のトラッキング」に該当)。

そして、不要物体を囲う長方形のバウンディングボックスを 3 次元空間内に配置した (2 章の「不要物体の認識とマスク画像の生成」)。これらの情報から、 LF_A と LF_B を使用して SI_1 のステレオ画像すべての隠背景を再構成した (2 章の「隠背景の再構成と現画像への合成」)。それぞれの結

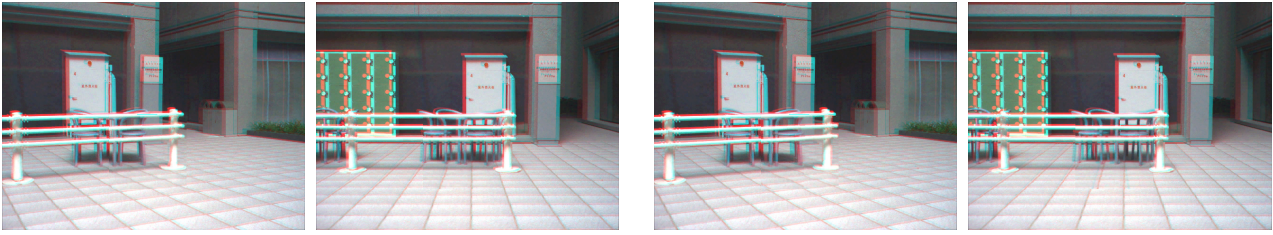


図3 実験1の結果*. 左から順に, DR_{1-A}の10および100フレーム目, DR_{1-B}の10および100フレーム目

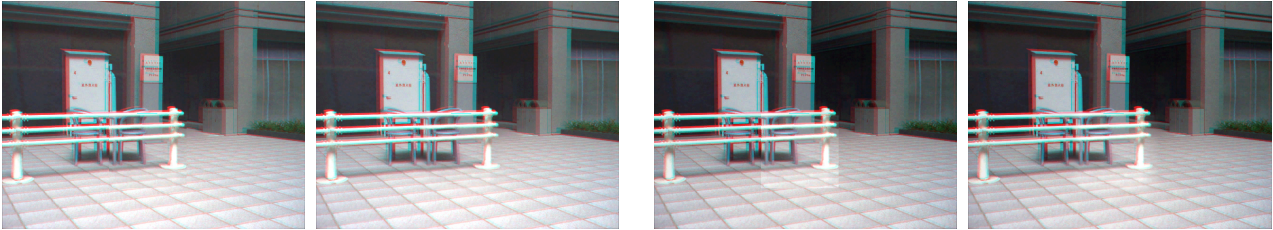


図4 実験2の結果*. 左から順に, DR_{1-A}, DR_{1-A-M}, DR_{2-B}, そしてDR_{2-B-M}の10フレーム目

*本来フルカラーのサイドバイサイド形式で提示されるが, 紙面の都合上, Cyan-Redのアナグリフ形式の結果を掲載する

果を DR_{1-A} と DR_{1-B} とする. 最後に, DR 処理結果 DR_{1-A} と DR_{1-B} を HMD に表示して, 観察を行い, 比較した.

図3に10および100フレーム目の DR_{1-A} と DR_{1-B} を示す. DR_{1-B} の2フレームを比較すると, IBR を間引いたことによる奥行の切り替わりが発生していることが分かる. 一方で, DR_{1-A} で同じ現象は確認されなかった. 以上より, 3.2章で述べた「IBRによる奥行知覚の変化」を確認した. さらには, 隠背景の Window Violation が見られる.

4.3 実験2

本実験では, 故意に光学的不整合を発生させるために, 画素値を20増加させて SI₂ を作成した. 次に, 実験1と同様に LF_A と LF_B を使用して SI₂ すべてのステレオ画像の隠背景を再構成し, DR 処理結果 DR_{2-A} と DR_{2-B} をそれぞれ得た. 加えて, SI₁ や SI₂ に隠背景の再構成結果を α ブレンドし, 画像を滑らかに合成するため, マスク画像にぼかし処理を施して α マップを作成した. 以降, この合成処理を Region Blending Across the Border (RBAB) 処理と呼ぶ. SI₁ と SI₂ が与えられたとき, LF_A と LF_B に α マップを適用することにより, DR 結果 DR_{1-A-M}, DR_{1-B-M}, DR_{2-A-M}, そして DR_{2-B-M} をそれぞれ得る. 最後に, HMD を用いて DR 処理結果を提示及び観察し, 実験1と2の全ての結果を比較した.

図4に DR_{1-A}, DR_{1-A-M}, DR_{2-B}, そして DR_{2-B-M} の10フレーム目を示す. DR_{1-A} と DR_{1-A-M} を比較すると, 隠背景の Window Violation が RBAB 処理により軽減された. また「幾何学的不整合による立体的に露呈する窓」が DR_{2-B} で見られるが, DR_{2-B-M} ではそれが軽減されている.

5. むすび

本稿では, ビデオシースルー型 HMD を用いた事前観測ベースの DR における両眼視野不整合の問題を明らかにした. また, 両眼視野不整合を実証するために実験用システムを構築した. 実験では, 既存の DR 手法を両眼立

体視できるように拡張した手法を利用した. 最後に, 特に重要な DR における両眼視野不整合に関して解決策を例示した. その1つとして, RBAB 処理を利用することで, マスク画像により形成される覗き窓が立体的に浮き出る問題を軽減できることを示した.

謝辞 本研究の一部は, 科研費・基盤研究 (S) 「複合現実型情報空間の表現力基盤強化と体系化」と学振・特別研究員奨励費 (課題番号 25・9193) 「複合現実空間の表現力向上のための2つのアプローチ」による.

参考文献

- [1] S. Mann: “Mediated reality,” TR 260, M.I.T. Media Lab Perceptual Computing Section, Cambridge, Ma, 1994.
- [2] S. Zokai, J. Esteve, Y. Genc, and N. Navab: “Multiview paraperspective projection model for diminished reality,” Proc. ISMAR, pp. 217 - 226, 2003.
- [3] A. Enomoto and H. Saito: “Diminished reality using multiple handheld cameras,” Proc. ACCV, pp. 130 - 150, 2007.
- [4] F. I. Cosco, C. Garre, F. Bruno, M. Muzzupappa, and M. A. Otaduy: “Augmented touch without visual obtrusion,” Proc. ISMAR, pp. 99 - 102, 2009.
- [5] Z. Li, Y. Wang, J. Guo, L.-F. Cheong, and S. Z. Zhou, “Diminished reality using appearance and 3D geometry of internet photo collections,” Proc. ISMAR, pp. 11 - 19, 2013.
- [6] S. Jarusirisawad and H. Saito: “Diminished reality via multiple hand-held cameras,” Proc. ICDCS, pp. 251 - 258, 2008.
- [7] V. Lepetit and M.-O. Berger: “An intuitive tool for outlining objects in video sequences: Applications to augmented and diminished reality,” Proc. ISMR, pp. 159 - 160, 2001.
- [8] B. Mendiburu: “3D movie making: Stereoscopic digital cinema from script to screen,” Focal Press, 2009.