

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 (2)

松木 ひとみ*1 森 尚平*1 池田 聖*1 柴田 史久*1 木村 朝子*1 田村 秀行*2

Abstract --- 複合現実感 (MR) が現実世界と仮想世界との融合を目指す技術であるのに対し、隠消現実感 (DR) は現実世界から物体を視覚的に隠蔽・消去・透過することを目指す技術である。DR においても MR 同様にステレオカメラ及びディスプレイを備えた Head-Mounted Display (HMD) を用いる実地体験型のシステムであることが望まれるが、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両眼視野不整合をシミュレートすることで実証した。また、その対策についても言及し、同様に実証した。

Keywords: 隠消現実感, 複合現実感, 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) とは逆の概念であり、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ている。現実世界への情報の加算を行う AR/MR だけでは現実感の表現力に限界があるため、減算する DR 技術と併用する必要があるとされる。

DR の実現方法は種々存在する。特に物体を除去する DR 技術のうち、対象領域の周辺から背景を推定する Inpainting-Based DR (IB-DR) や観測を伴って背景を再構成する Observation-Based DR (OB-DR) 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、当然ステレオカメラ及びディスプレイを備えた Head-Mounted Display (HMD) を用いる実地体験の形式を取ることが好ましい。にもかかわらず、既存の DR システムにおいては、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あまり見られない。即ち、DR の両眼立体視に関する問題は事実上未確認であると言える。

本論文では、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、これまで我々が分析及び対処法に関して検討してきた内容[2]を踏まえてさらに検討を進める。過去の研究[2]では、以下の手順で検討を進めた。本研究もそれに倣うものとする。

- ・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 OB-DR における両眼立体視と両眼視野不整合

[2]の文献で述べたように、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、その各手順で発生する誤差が両眼立体視野不整合となって現れる (図 1)。

- ・隠背景の観測とデータ化
- ・カメラ、シーン、不要物体のトラッキング

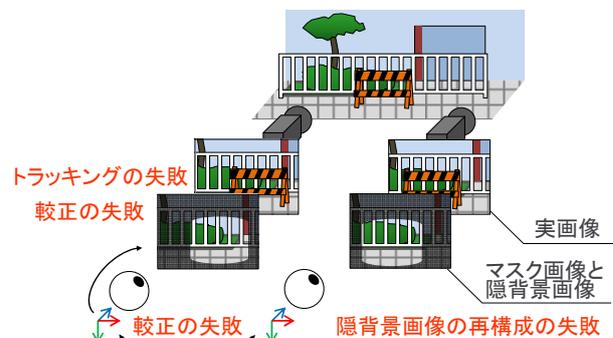


図 1 実画像、再構成画像、マスク画像の関係図
Fig. 1 Relationships between a stereo camera, video streams from the camera, and masked hidden view recovery

*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

表 1 OB-DR における両眼視野不整合 ([2]より, *を本論文で拡張)

Table 1 Binocular mismatching in OB-DR (*The underlined content is the new contribution in this paper from [2])

処理手順	両眼視野不整合	原因	解決策
ステレオカメラ校正	除去対象領域における不自然な立体感の知覚	不正確な校正	時間をかけた事前の校正
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
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
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 (3.3 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	
	立体感の変化	疎な IBR (例: VDTM)	<u>Cyclops Blending Field</u> を用いた <u>VDTM*</u> 密な IBR (例: LFR)

・除去対象領域を示すマスク画像の生成
・データ化した隠背景再構成と体験時画像への合成
さらに、具体的に両眼視野不整合が発生する原因を以下の様に考察した(表 1)。

- ・ステレオカメラ校正の誤差
- ・ステレオ画像の取得及び表示の非同期
- ・不適切なマスク画像の生成
- ・隠背景再構成に関する誤差

このうち、初めの 2 つの原因は、現実的には、然るべきハードウェアを利用し、かつ準備段階で時間をかければ解決できる問題であると考え、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と、以下が問題となる。

問題 1: 幾何学的不整合により発生する除去対象領域内での両眼視野闘争及び Window Violation [3]

問題 2: 光学的不整合により立体的に知覚される除去対象領域

問題 3: Image-Based Rendering (IBR) により除去対象領域内で知覚される奥行きの変化

3 実装と実験

分析した両眼視野不整合に対して、下記の 3 つの実験を行い、各問題の発生の確認とその対処法について議論する。

実験 1: 光学的不整合に起因して除去対象領域が立体的に見えること(問題 2)を確認し、Region Blending Across the Border (RBAB) 処理によりこれが軽減されることを確認する。

実験 2: 密な Light Field 及び疎な Light Field を構成し DR 処理した結果を比較し、後者において、除去対象領域内で Windows Violation (問題 1) 及び、知覚される奥行きがフレーム間で異なることを確認



図 2 提示されるステレオ画像の例

Fig. 2 An example of stereo images in the experiments

する(問題 3)。

実験 3: 疎な IBR を想定し実装した View-Dependent Texture Mapping (VDTM) [4] と Cyclops Eye Blending Field を利用した VDTM を比較し、後者において、除去対象領域内で知覚される奥行きの変化(問題 3)が軽減されることを確認する。

3.1 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 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、フルカラーのサイドバイサイド方式のステレオ映像で確認した(図 2)。尚、紙面の都合上、以降は Cyan-Red のアナグリフ形式の結果を掲載する。

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した。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 (Flea3 GigE)、照明 (太陽光として ARRI 社製 HMI575、環境光として 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 3 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。除去対象は工事の柵とした。尚、カメラ及び HMD は事前に校正されており、それぞれの内部パラメータと歪み補正值、左



(a) システムの全景

(a) A complete view of the system



(b) 実験シーンのクローズアップ

(b) Close up view of the experimental scene

図3 実験システム

Fig. 3 Experimental system

右カメラへの座標系変換行列は既知である。尚、撮影の最中はホワイトバランスの変更は行わず、除去対象から落ちる影は手で設定したバウンディングボックスによって決定される除去対象領域内に収まるものとした。

3.2 実験用データの作成

隠背景データとして、極座標系で記述した球面 Light Field ($r: 400 \text{ mm}$, $-2.5^\circ \leq \varphi \leq 2.5^\circ$, $0^\circ \leq \theta \leq 60^\circ$, 200×20 のカメラアレイ) である LF_D 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_D から視点依存画像を 5 枚毎に 1 枚と間引いて LF_S を作成した。続いて、除去対象をシーン中に設置し、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] の 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] の 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_S と LF_D 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] の 2 章の「隠背景の再構成と現画像への合成」)。それぞれの結果を DR_{S-1} と DR_{D-1} とする。

3.3 実験 1

本実験では、3.2 節に示したデータに加え、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 SI_2 を作成した。次に、 LF_D を使用して SI_2 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_{D-2} をそれぞれ得た。加えて、 SI_2 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。

SI_2 が与えられたとき、 LF_D に α マップを適用することにより、DR 結果 DR_{D-2-M} を得た。最後に、HMD を用いて DR 処理結果を提示及び観察した。

図 4 に DR_{D-2} , DR_{D-2-M} の 10 および 100 フレーム目を示す。 DR_{D-2} と DR_{D-2-M} を比較すると「光学的不整合による立体的に知覚される除去対象領域 (問題 1)」が DR_{D-2} で見られるが、マスク画像のエッジが RBAB 処理により目立たなくなることで DR_{D-2-M} ではそれが軽減されている。これらの現象は除去対象領域に 3D バウンディングボックスを使用している Cosco ら [5] や、Zokai ら [6] の場合にも発生するものと考えられる。

3.4 実験 2

本実験では、DR 処理結果 DR_{S-1} と DR_{D-1} を HMD に表示して、観察を行い、比較した。図 5 に 10 および 100 フレーム目の DR_{S-1} と DR_{D-1} を示す。 DR_{S-1} 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変化が発生していることが分かる。一方で、 DR_{D-1} で同じ現象は確認されなかった。以上より、2 章で述べた「IBR による奥行知覚の変化 (問題 3)」を確認した。加えて、隠背景の Window Violation が見られることを確認した (問題 1)。

3.5 実験 3

Cosco ら [5] の手法のように VDTM を利用する際には、ある体験時視点における視点依存画像の合成割合を表す Blending Field (BF) [7] を計算する。本実験では、通常の BF の計算方法に加え、2 章で示した問題 3 の発生を軽減する新たな BF の計算方法を導入し、これにより問題 3 が実際に軽減されることを確認した。

実験の準備として、まず、 LF_D から 134 枚の画像を取り出し、Agisoft 社の PhotoScan を用いてそのシーンの 3D モデルを獲得した。次いで、その 3D

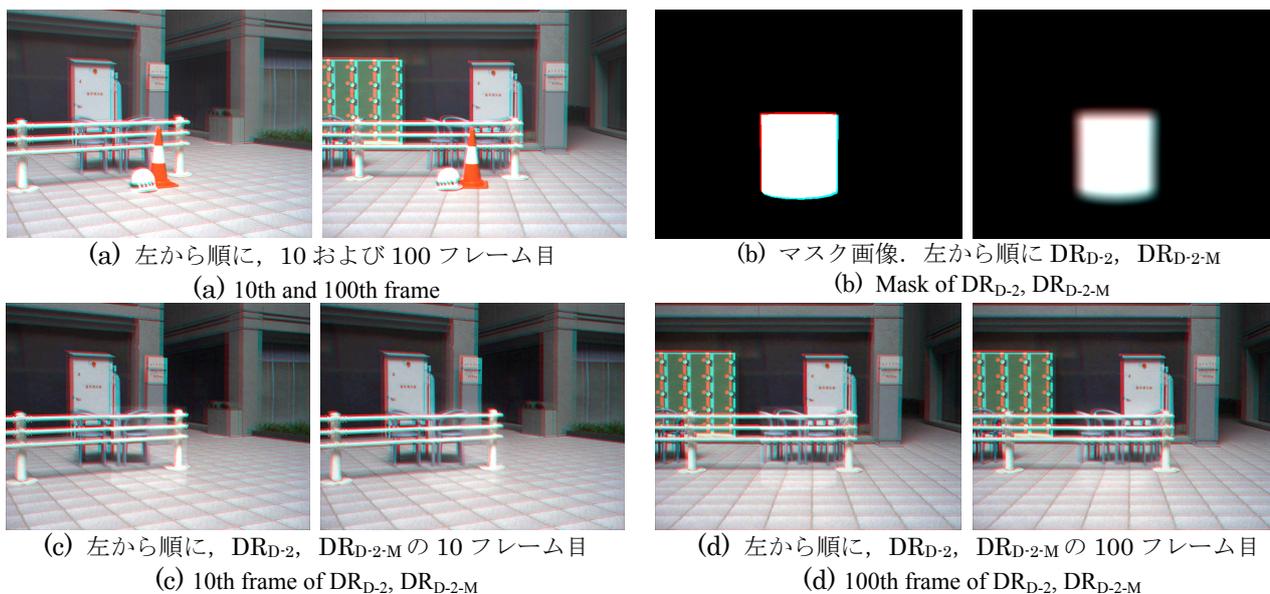


図 4 実験 1 の結果
Fig.4 Results of Experiment 1

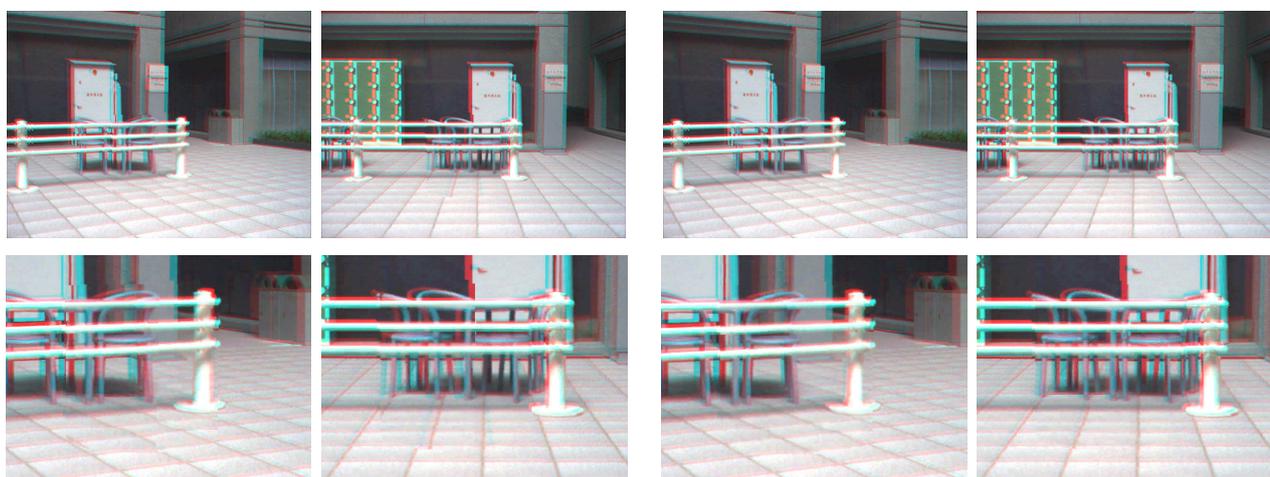


図 5 実験 2 の結果. 左から DR_{S-1} の 10 および 100 フレーム目, DR_{D-1} の 10 および 100 フレーム目. 下段は各拡大図
Fig.5 Results of Experiment 2. Left to right, 10th and 100th frame of DR_{S-1} , and 10th and 100th frame of DR_{D-1} .

モデルを基にして, 手動で新たに 3D モデルを作成し, VDTM のための Geometry Proxy [7] として用いた. この Geometry Proxy を用いて, 2 通りの BF を計算する. 一方は, 左右のカメラ位置姿勢でそれぞれの BF を計算し, ステレオ画像を描画した. これを「Stereo VDTM」と呼ぶ. もう一方は, 両目の中間点のカメラ位置姿勢を計算し, その位置姿勢を基に左右共通の BF を計算して, ステレオ画像を描画した. これを「Cyclops VDTM」と呼ぶ. 図 6 に示すのは, その概念図である.

図 6 に示す通り, Stereo VDTM では事前撮影された視点依存画像の枚数が少ない場合, 視点位置によっては, 左右に表示される画像の視差が本来よりも大きくなってしまふ場合がある. 一方で, Cyclops VDTM では, 両目で共通の BF を利用するため, こうした問題が発生しない. しかし, この方法によっ

て, VDTM による一視点における両眼立体視の効果はなくなり, 運動視差のみが得られることになる. 換言すれば, 両眼視差による奥行きは Geometry Proxy [7] において感じられるようになる.

図 7 は, Stereo VDTM と Cyclops VDTM によって隠背景を再構成した結果である. この結果から, 先に示した効果が得られていることが確認できる.

4 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. 実施した 3 つの実験では, 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, 定めた両眼視野不整合が実際に発生することを確認した. 特に重要な DR に

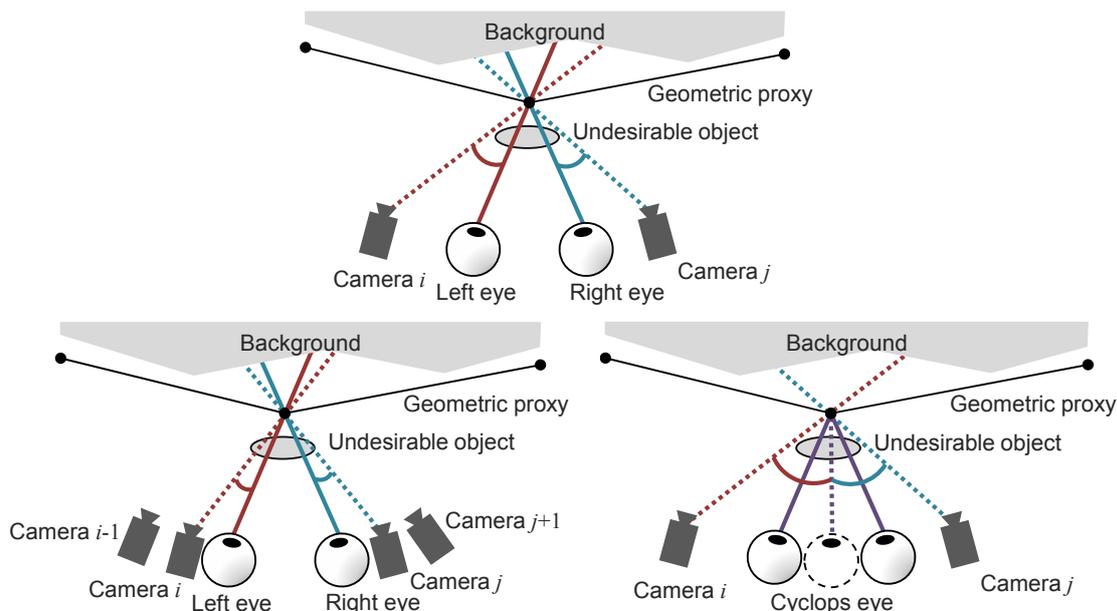


図 6 IBR による両眼視野不整合. 疎に覆われた視点依存画像は除去対象領域の窓において両眼視野闘争を引き起こす (上). しかし密な場合は基本的にはこの問題を解決し (左), cyclops eye blending field を基にした VDTM はこの影響を軽減する (右)

fig.6 Binocular mismatching due to IBR. Sparsely covered view-dependent images cause binocular rivalry in the hidden view window (top). Although densely covered view-dependent images essentially solve this problem (left-down), VDTM based on cyclops eye blending field can reduce the effect (right-down)

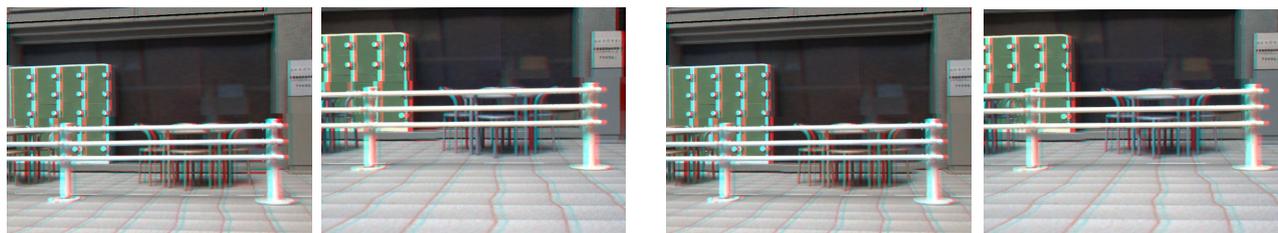


図 7 2 視点における Stereo VDTM (左) と Cyclops VDTM (右)

Fig.7 Stereo VDTM (left) and Cyclops VDTM (right) at two view points

における両眼視野不整合に解決策を提案した。

具体的には、2 値のマスク画像を用いた際に除去対象領域浮き出て見える問題に RBAB 処理が効果的であること、DR において VDTM を利用する際に、Cyclops VDTM によって奥行知覚に関する問題を部分的に解決できることを示した。

今後は、OB-DR 及び IB-DR における両眼視野不整合をさらに調査してゆく。

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] 松木ひとみ, 森尚平, 柴田史久, 木村朝子, 田村秀行: 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, 第 20 回日本バーチャルリアリティ学会大会論文集, 21B-2, pp. 230 - 233 (2015.9).

[3] B. Mendiburu: "3D movie making: Stereoscopic digital cinema from script to screen," Focal Press, 2009.

[4] P. E. Debevec, C. J. Taylor, and J. Malik: "Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach," Proc. SIGGRAPH, pp. 11-20, 1996.

[5] F. I. Cosco, C. Garre, F. Bruno, M. Muzzupappa, and M. A. Otaduy: "Augmented touch without visual obtrusion," Proc. ISMAR, pp. 99 - 102, 2009.

[6] S. Zokai, J. Esteve, Y. Genc, and N. Navab: "Multiview paraperspective projection model for diminished reality," Proc. ISMAR, pp. 217 - 226, 2003.

[7] C. Buehler, M. Bosse, L. McMillan, S. Gortler, and M. Cohen: "Unstructured lumigraph rendering," Proc. SIGGRAPH, pp. 425 - 432, 2001.