

遠隔地における複合現実空間の共有 —対面型卓上作業システムの構築—

南谷 真哉† 北原 格‡ 亀田 能成‡ 大田 友一‡

†筑波大学 大学院システム情報工学研究科 知能機能システム専攻

〒305-8573 茨城県つくば市天王台1-1-1

E-mail: †minatani@image.iit.tsukuba.ac.jp ‡{kitahara, kameda, ohta}@iit.tsukuba.ac.jp

あらまし 遠隔地に存在するユーザで複合現実空間を共有することにより、現実空間では実現不可能な協調作業環境の提供を目的とした“遠隔協調型複合現実感”に関する研究に取り組んでいる。本稿では、対面型卓上作業を題材とし、協調作業に必要な情報、具体的には互いの空間に実在もしくは仮想的に存在する物体の視覚情報を、実時間で獲得・伝送・提示することにより、空間を共有している感覚を引き起こすシステムについて述べる。

キーワード 協調型複合現実感, ビルボード, 臨場感通信, 卓上作業

Remote Shared Mixed Reality System —Remote Collaboration by Using Face-to-Face Tabletop Type—

Shinya MINATANI† Itaru KITAHARA‡ Yoshinari KAMEDA‡ and Yuichi OHTA‡

†University of Tsukuba, Department of Intelligent Interaction Technologies

1-1-1 Tennoudai, Tsukuba Science City, Ibaraki, 305-8573 Japan

E-mail: †minatani@image.iit.tsukuba.ac.jp ‡{kitahara, kameda, ohta}@iit.tsukuba.ac.jp

Abstract We are conducting on “Remote Shared Mixed-Reality” that enables multiple users in distant places to share a same Mixed-Reality space. In this paper, we propose a novel face-to-face tabletop type remote collaboration system to share the same tabletop MR space by two users sitting in distant places. In order to realize the system, the system has to execute the all processes, capturing, sending, and rendering the appearance of users, real/virtual objects in real time.

Keyword Shared Mixed-Reality, Billboard, Tele-Immersion, Table-Top Communication

1. はじめに

複数のユーザが同一の複合現実感空間において協調作業を行う協調型複合現実感（協調 MR）では、視覚情報の共有により、円滑なコミュニケーションが可能となる[1][2]。しかし、互いの姿や空間中に存在する実物体の見え方を共有するために、全ユーザが同一の実空間に存在していることが前提条件となるため、離れた場所に存在するユーザ同士での協調作業に適用することはできない。一方で、ネットワーク技術の発達により、テレビ会議システムなどを介して遠隔地のユーザとコミュニケーション・協調作業をする機会が日常的になりつつある。我々は、ネットワーク技術を活用することにより協調 MR が有する物理的制限を取り除き、遠隔地における複合現実空間の共有を目的とした“遠隔協調型複合現実感（遠隔協調 MR）”の研究に取り組んでいる。

MR 空間中のユーザは、現実空間に仮想物体を重畳

した映像を観察するためにヘッドマウントディスプレイ（HMD）を装着している。特に、現実世界の見え方もビデオデータとして撮影・提示を行うビデオスルー型 HMD を利用する場合、CG で描画される仮想物体だけでなく、周りの人物や背景などの実物体の見え方もビデオ映像として電子的に取り扱うため、ビデオ会議同様、ネットワークを介して遠隔地への伝送や提示が可能となる。我々はこの特性に着目し、遠隔地の人物像の撮影・伝送・提示処理を実時間で実行することにより、図 1 に示すように空間を越えて MR 空間を共有している感覚を与えるシステムの実現を目指して入る。

これまで臨場感通信（Tele-Immersion）に関する様々な研究が行われているが[3][4]、複合現実感と組み合わせることにより臨場感の向上を実現させる例は、それほど見当たらない。MR 技術を用いた遠隔コミュニケーションシステムとして、実空間中に設置したスクリ

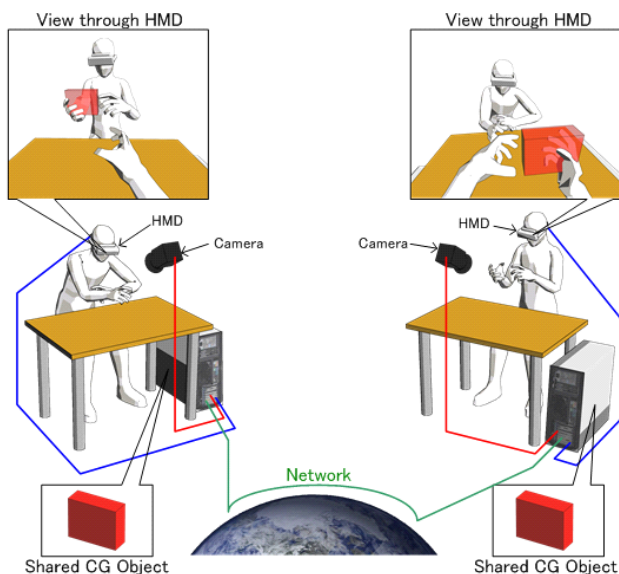


図 1：遠隔協調型複合現実感システム

ーンや壁面にプロジェクタを用いて遠隔地のユーザの見え方情報を投影し、協調作業を行うシステムが提案されている[5][6]。これらのシステムでは、遠隔地の通信相手が現実の空間中に存在するような映像を提示することにより、空間を“共有しているような”インタラクションの実現は可能であるが、実際には、現実空間と遠隔地は映像提示スクリーンによって分断されているため、相手側の空間に入り込んでのインタラクションを行うことはできない。協調作業の観点から遠隔コミュニケーションシステムを考えると、ジェスチャーや顔の向きなどの非言語情報と共に、共同で作業を行う空間を、互いに入りこめる状態で共有し、その中のインタラクションが重要である[7]。

本研究では、協調作業の一例として、図 2 に示すような対面型卓上作業を取り上げ、実空間を分断することなく複合現実感の提示が可能な HMD を用いて、卓上空間にユーザが入り込んでのインタラクションを実現する遠隔協調型複合現実感システムを構築する。

2. 遠隔協調型複合現実感システム

実空間と仮想空間を違和感無く合成するためには、幾何的整合性、光学的整合性、時間的整合性（時間遅れ）の三つの整合性が重要とされている。我々は、これらの中でも、円滑なコミュニケーションの実現に欠かせないものとして、時間的整合性に重点を置いている。また、ユーザの表情の変化などの細かい非言語情報の伝達に影響することから、人物像や物体像を撮影した映像の解像度も重要であると考えている。このような考えに基づき、提案する遠隔協調型 MR システムでは、人物像や実物体像の映像を高解像度で撮影し、

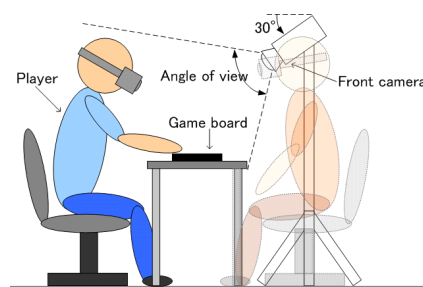


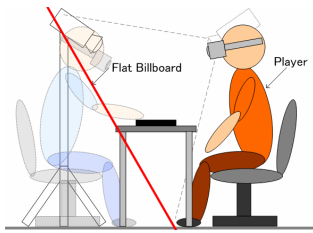
図 2：提案システムのレイアウト

かつ、それらをリアルタイムで伝送・提示することを目標とする。

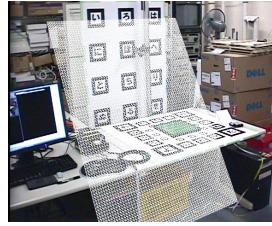
遠隔地のユーザや実物体の 3 次元的な見え方の変化を再現する手法として、3 次元モデルを生成する手法[8][9]や、イメージベースレンダリングによる手法[10][11]などがある。しかしこれらの手法では、3 次元モデル生成や自由視点映像生成に多くの計算コストが必要となるためリアルタイム処理が困難であり、また、膨大な量の生成された 3 次元モデルデータをリアルタイムで伝送することは、現代の広帯域ネットワークをもってしても現実的ではない。これらの計算コスト・データ量は、撮影映像の解像度が増えれば増えるほど増大する傾向にあるため、従来のシステムでは、リアルタイム処理と、人物や実物体の映像を高解像度で撮影することは、相反する関係にある。

我々の提案するシステムでは、対象物体の 3 次元形状を忠実に表現する代わりに、1 枚の平面で形状を表現するビルボードを採用している。この手法では、物体の 3 次元形状を復元するのではなく、対象物体の位置・姿勢と観察者の視点位置からビルボードの 3 次元空間中での位置・姿勢を算出するもので、対象となるシーンを撮影するために必要なカメラは 1 台で済み、モデリングやレンダリングの計算コストは格段に少ない。高解像度のカメラで撮影することにより、ビルボードに投影される画質を容易に向上させることが可能な上に、伝送するデータはカメラ 1 台分の映像データのみであるため、現行の広帯域ネットワークでも十分にリアルタイム伝送が可能であり、リアルタイム処理と、人物や実物体の映像の高解像度の両方の実現が可能という特長を有している。

一般的なビルボードの形状（地面に垂直に立てられた平面）は、モデリング対象となる物体（人物像）の実際の形状とは異なるため、ユーザの視点が撮影時にカメラが設置された位置以外に移動する場合、提示される映像に歪みが生じる。我々は、できるだけ対象物体（人物像）の表面とビルボード面の距離が近くなるような形状にビルボードを変型させることにより、この問題を解決する。この新しいビルボードの表現手法

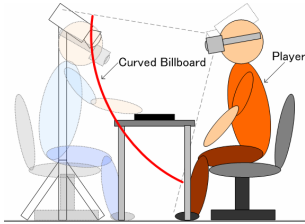


(a) 横から見た形状と位置

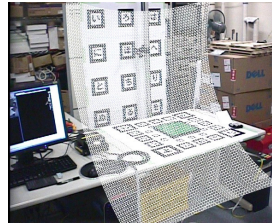


(b) 実空間に投影した図

図 3：傾倒平面ビルボード

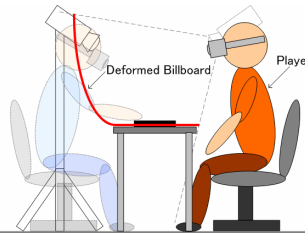


(a) 横から見た形状と位置

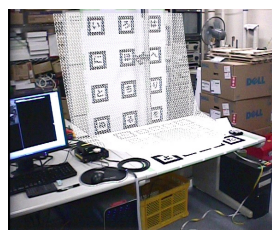


(b) 実空間に投影した図

図 4：曲面ビルボード



(a) 横から見た形状と位置



(b) 実空間に投影した図

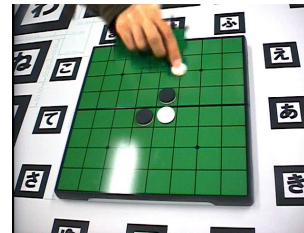
図 5：変型ビルボード

を、変型ビルボードと呼ぶ。

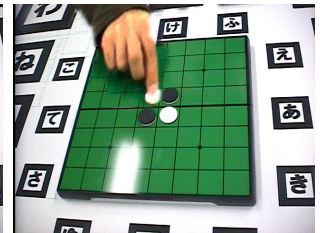
我々が想定する作業環境の条件を満たすような遠隔協調 MR システムのプロトタイプとして、オセロゲームを題材として取り上げ、遠隔協調作業システムを構築した。対戦相手の映像と盤上の状態をリアルタイムで取得・伝送・提示することにより、遠隔地のユーザ同士でも、まるで目の前にいる相手と対戦しているかのような自然な協調作業を可能にする。

3. ビルボードによる人物像のレンダリング

図 2 は、我々の提案する遠隔協調 MR システムのレイアウトを示しており、二人のユーザが“仮想的に”向かい合って座った状態を想定している。対戦相手に伝送する人物像や盤面状況の獲得のために、超広角レンズを取り付けた一台のカメラ（以降、正面カメラと呼ぶ）が双方に配置されている。正面カメラは、正面から見える人物像全体と卓上面が画角に収まる位置で、水平から 30 度ほど下に傾けて固定している。これは、遠隔地の対戦相手の視点の位置姿勢と画角とほぼ同等になるように設定されている。正面カメラで撮影され



(a) 曲面ビルボードの場合



(b) 変型ビルボードの場合

図 6：曲面／変型ビルボードの見え方の比較

た人物像は、ネットワーク経由で遠隔地に伝送された後、ビルボードに投影される。その際、正面カメラのパラメータとビルボードの位置・姿勢・形状に合わせて 2 次元射影変換を施すことにより、ユーザの視線方向の変化に対応した自由視点映像が生成・提示される。ユーザの視点の移動に応じてビルボードの位置姿勢を変化させることにより、実空間中の決まった位置に人物像が提示される。

傾倒平面ビルボード

まず我々は、ビルボードの形状に関する検討を行った。図 3 は、一般的に用いられるビルボード形状である平面ビルボードを、正面カメラの俯角に合わせて傾けた“傾倒平面ビルボード”のレイアウトと、実空間中におけるビルボードの配置を示している。正面カメラには、広範囲を見回すような映像生成に対応するため超広角レンズを取り付けているが、このレンズの歪曲収差により、提示される人物像が歪んでしまう。内部カメラパラメータを用いて収差を除去することは可能であるが、計算コストが増大してしまう。

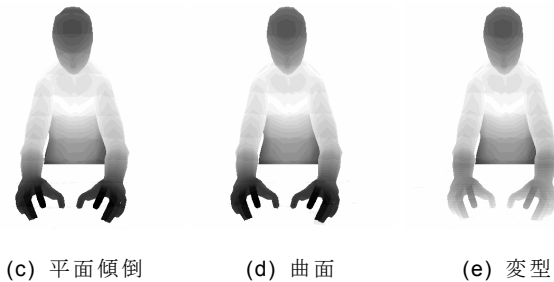
曲面ビルボード

そこで、図 4 に示すように、超広角レンズの歪曲収差に合わせて湾曲している 2 次曲面形状のビルボード“曲面ビルボード”を考案した。ビルボードにマッピングされる座標は予めレンズの内部パラメータを基に算出されているため、歪み補正の計算はグラフィックアクセラレータの幾何計算機能を用いて行うため、非常に高速である。また、曲面ビルボードは傾倒平面ビルボードに比べて人物表面との距離が若干近くなる、つまり、椅子に座った人物表面は平面よりも二次曲面に近いので、視点移動時の誤差も軽減される。

この曲面ビルボードを用いることにより、卓上面よりも奥にあるユーザの上半身像は、ビルボードの表面と人物表面との距離が近いため、違和感なく提示することができる。しかし、卓上面付近まで伸ばしているユーザの腕や手の像は、ビルボード表面からの距離が極端に離れてしまう。そのため、図 6(a) に示すように、ユーザが視点位置を移動すると、卓上面付近に提示される人物像は実際の位置から大きくずれてマッピングされてしまう。



(a) 3次元人物モデル(前面) (b) 3次元人物モデル(側面)



(c) 平面傾倒 (d) 曲面 (e) 変型

変型ビルボード

この問題を解決するため、我々は新たな形状のビルボード“変型ビルボード”を考案した。図5に示すように、ユーザの上半身がマッピングされるビルボードの上半分は曲面ビルボードと同様にレンズの歪曲収差に合わせて変型させ、ユーザの腕や手がマッピングされる下半分は机の面に合わせた平面ビルボードとなっている。この二つの領域は、ビルボード全体が連続面となるように滑らかに繋げられている。テクスチャがマッピングされる座標は、カメラの位置とレンズの歪曲収差に合わせて予め算出されているため、レンズによる歪みは実空間全体で補正されている。その結果、図6(b)に示すように、ユーザが視点を移動させても、相手の腕や手を常に正しい位置に違和感なく提示させることができる。図6の(a)と(b)は、どちらも同じ視点から見た、同じマス目に駒を置いている様子を示しているが、図6の(a)では、正しい場所に置かれているように見えない。

変型ビルボードの妥当性を評価するため、図7(a)(b)に示すような3次元人物モデルを用いて、人物表面とビルボード表面の間の距離を測定した。図7(c)~(d)にはその結果がグレースケール画像として示され、色の濃い部分ほど距離が遠いことを意味する。この結果からも、変型ビルボードを用いることにより、モデリング対象である人物形状により近い形状表現が可能であることがわかる。

4. システム構成

図8に、今回提案する遠隔協調型MRのプロトタイプシステムの概観を示す。プレイヤーは互いにゲーム盤の置かれた机の前に座っている。机の向う正面には、垂直画角 89.08 度の超広角レンズが取り付けられたカ

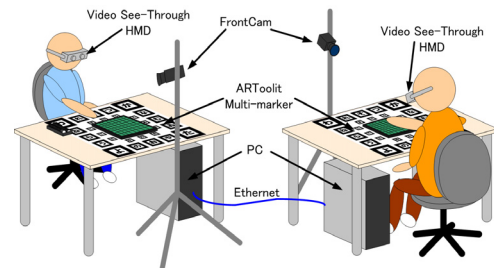


図8：プロトタイプシステムの概観

メラ（正面カメラ）が一台設置されており、プレイヤーの人物像全体と卓上面全体を撮影する。その光軸は水平から 30 度下に傾けてられた状態で固定されている。卓上面に置かれる実物体（本システムの場合はオセロの駒）の状態は、正面カメラの映像から抽出される。双方のプレイヤーは、両眼に映像を提示する二つの LCD と、外界の映像を取得する二つのカメラを内蔵したビデオシースルー型 HMD を装着している。HMD の位置と姿勢は、ARToolkit のマルチマーカ機能を利用し、内蔵カメラから得られたマーカ映像を基に算出する[12]。これらのカメラと HMD は互いのプレイヤー側にある一台の PC に接続され、その PC 同士がネットワークで接続されている。このネットワークを通して、プレイヤーの人物像や卓上の状態がリアルタイムで相手に送られる。

本システムの処理の流れ図を図9に示す。これは大きく“キャプチャステージ”と“レンダリングステージ”の二つに分かれている。これら全ての処理は、互いのプレイヤー側の1台のPCによってビデオレート処理される。

4.1. 人物像の取得

図10(a)には、IEEE1394b インタフェースを通してPCに送られた正面カメラからの映像の例を示している。この映像から、予め撮影しておいた背景画像を用いて背景差分処理を行い人物像マスクを生成する。マスクされた人物像は JPEG 方式で圧縮される。この JPEG データと共に、ランレングス圧縮されたマスクデータ、その他の付加情報（マスクのサイズなど）などが格納されたデータストリームが1フレーム毎に生成される。

4.2. 卓上面の実物体認識

提案するオセロゲームシステムでは、正面カメラからの映像を基に、ゲーム盤上の全ての駒の位置と色を自動的に認識している。オセロの駒の色は白か黒と定められており、また盤面は緑色と決まっている。そこで、今回は単純に $L^*a^*b^*$ 表色系を用いた色ベースの識別手法を導入した。ゲーム盤の全てのマス目の中心座標を予めキャリブレーションしておき、以下に示すア

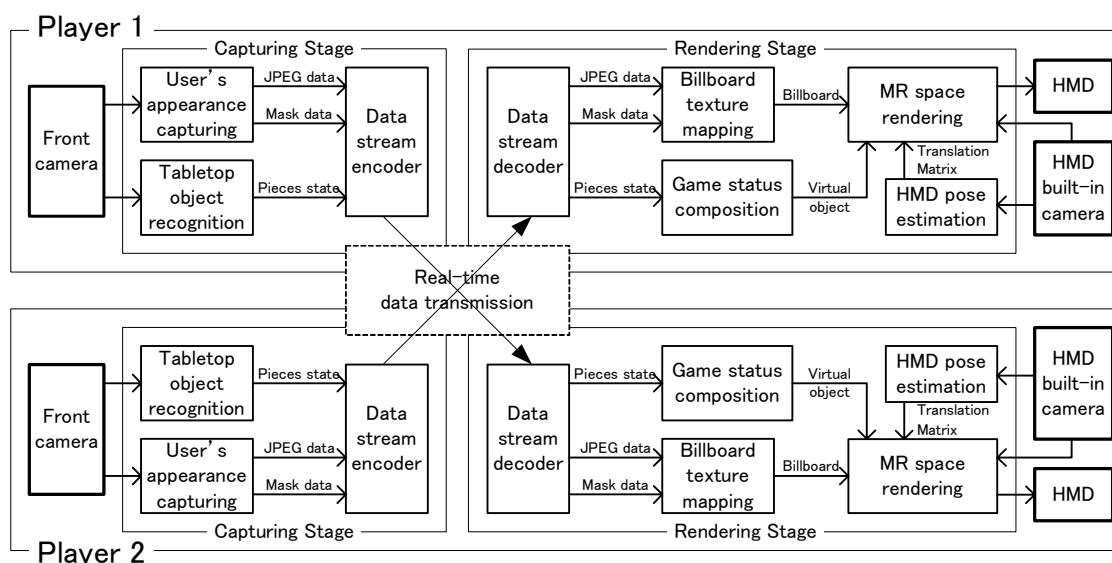


図9：提案システムの処理の流れ図

ルゴリズムによって各マス目の状態を認識する。

- 1) a^* と b^* の値が共に閾値以下の場合には駒があり、そうでない場合は、駒は置かれていないと認識。
- 2) L^* の値が閾値以下の場合には黒駒、そうでない場合は、白駒が置かれていると認識。

駒や盤面以外の物体による誤認識を防ぐため、認識されるマス目に人物像マスクの領域が重なっている場合には、そのマス目はプレイヤーの手で隠されていると判断し、前の状態が維持されていると認識する。この手法では、プレイヤーが新しい駒を置いた場合にのみ認識結果が更新されるため、新しい駒を置いた時点を容易に検出できる。新しい駒が置かれると、その時点の盤面状況がデータストリームの付加情報として相手に送られる。

4.3. リアルタイムデータ転送

本システムでは時間遅れの少ないリアルタイム映像伝送を目指して、データストリームの同期にはTCP方式を、データストリームの連続伝送にはUDP方式を用いる複合伝送方式を実装した。データストリームは1フレーム毎に、UDP方式の最長フレームサイズに区切られて伝送され、その同期や遅延の制御をTCP方式の通信で行っている。

4.4. 変型ビルボードへのテクスチャマッピング

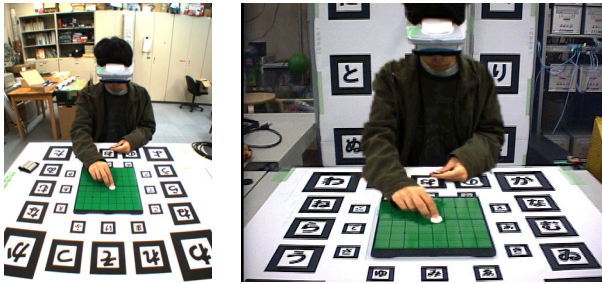
相手から送られてきたデータストリームを受信すると、まず人物像マスクに平滑化フィルタがかけられ、そのマスクの画素値をアルファ要素として用いた人物像のテクスチャを生成する。生成したテクスチャはアルファブレンドを用いて変型ビルボードにマッピングされる。平滑化フィルタによりエッジ付近がぼかされているため、背景である実空間映像と相手の人物像が自然に合成される。

4.5. HMDの位置・姿勢の獲得

HMDには正面を撮影するカメラが内蔵されていることを利用し、HMDの位置と姿勢の獲得にARToolkitのマルチマーカ機能を用いることにした。ビルボードが重畳される範囲全てをカバーするために、卓上面や正面カメラの後ろに立てられたボードに多数のマーカを配置した。HMD内蔵カメラがこれらのマーカの一部を捉えると、実空間に固定されたMR空間の基準点と内蔵カメラとの間の剛体変換行列を算出することができる。

4.6. MR空間のレンダリング

MR空間のレンダリングはHMDの左右の画面共にビデオレートで行われる必要がある。まず、HMD内蔵カメラで撮影された映像を、実世界背景としてそのまま画面いっぱいにレンダリングする。次に、人物像がテクスチャとしてマッピングされた変型ビルボードオブジェクトを、予めキャリブレーションしておいた定位置にレンダリングする。さらに、対戦相手から送られてきた盤面状況を基に、実空間には存在しない相手が置いた駒をCGオブジェクトとしてMR空間中にレンダリングする。オセロのルール上、相手の駒によってひっくり返された自分の駒も、CG物体として実物体の駒を覆い隠す形でレンダリングされる。最後に、HMD装着者の頭の位置と姿勢にあわせるため、4.5節で述べたARToolkitで得られた剛体変換行列をMR空間全体の座標に掛け合わせる。これらの処理をビデオレートで、かつできるだけ時間遅れが生じないように処理することにより、遠隔地にいるはずの対戦相手が、まるで実世界の目の前に座って対戦しているかのように感じさせることができる。



(a) 正面カメラ映像 (b) HMD を通してみた映像

図 10：遠隔協調 MR システムの実行例

5. 実験結果

現在の我々の実験環境では、プレイヤーPC に Dell XPS710 (Core2Extreme 2.67 [GHz])を、正面カメラとして PGR Flea2 (IEEE1394b, 768x1024 [pixels], 30 [FPS])を、ビデオスルー型 HMD として LCD 解像度が片目で 640x480 [pixels]のステレオ型ものを使用している。この環境下で、15~30 [FPS]でのレンダリングを実現できており、ネットワークの品質や帯域などに依存するが、概ねビデオレートでのレンダリングを実現できている。ビットレートは概算値で 130 [KB/frame]以下であり、ネットワーク以外の要因による遅延は 1 秒以下に抑えることができています。

図 10(b)に示すように、HMD を通してみた映像には、実際には目の前にはいない遠隔地の対戦相手の人物像が実世界に重畳され、あたかも目の前にいるかのような自然な対戦風景を提示できている。また、対戦相手が遠隔地の盤上に置いた駒を CG として盤上に提示することで、オセロゲームのルールに忠実な自然な対戦を実現できている。

6. まとめ

我々は、遠隔協調型 MR 技術を用いた対面型卓上遠隔協調作業システムを提案した。遠隔協調 MR のプロトタイプシステムとして、遠隔オセロゲームを実装し、その実現可能性を検証した。変型ビルボードというレンダリング手法を用いることで、人物像のリアルタイム伝送、レンダリングを容易にし、人物像の自然な提示と、作業領域となる卓上面付近において常に実世界に対して正しい位置にレンダリングできることが確認された。また、対戦相手が遠隔地の盤上に置いた駒や、挟まれてひっくり返された自身の駒を CG 物体として盤上に置かれているかのように提示することにより、実世界で駒を打つ感覚はそのままに、自然なオセロの対戦を実現することができた。

卓上状況の認識や遠隔地での再現手法を適応させることで、オセロ以外にも様々な卓上型遠隔協調 MR システムを構築できると思われる。

しかし、自然な映像の提示のためには、少なくとも以下のような問題の解決が必要であると考えている。それは、ビルボードに投影される人物像、CG 物体の駒、実物体である作業者自身の腕や手、盤上に実際に置かれている実物体の駒、それぞれのオクルージョンの問題である。特に作業者自身の腕や手が CG 物体に隠されてしまうことは本来ありえないため、HMD 内蔵カメラの映像を基に実物体の領域を切り出して CG よりも手前にレンダリングしなおす手法を検討する必要がある。

文 献

- [1] Y. Ohta and H. Tamura. "Mixed Reality - Merging Real and Virtual Worlds -," Springer-Verlag, 1999.
- [2] K. Kiyokawa, H. Iwasa, H. Takemura and N. Yokoya. "Collaborative Immersive Workspace through a Shared Augmented Environment," In Proc. of SPIE '98, Vol.3517, pp.2-13, 1998.
- [3] H. Takemura, Y. Kitamura, J. Ohya and F. Kishino. "Distributed Processing Architecture for Virtual Space Teleconferencing," In Proc. of ICAT'93, pp.27-32, 1993.
- [4] S. Sugawara, G. Suzuki, Y. Nagashima, M. Matsuura, H. Tanigawa, M. Moriuchi. "InterSpace: Networked Virtual World for Visual Communication," In IEICE transactions on information and systems, E77-D(12), pp.1344-1349, 1994.
- [5] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs. "The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays," In Proc. of SIGGRAPH'98, pp.179-188, July 1998.
- [6] J. Mulligan, K. Daniilidis. "View-independent scene acquisition for tele-presence," In Proc. of ISAR2000, pp.105-108, 2000.
- [7] K. Kiyokawa, M. Billinghurst, S. E. Hayes, A. Gupta, Y. Sannohe, and H. Kato. "Communication Behaviors of Co-located Users in Collaborative AR Interfaces," In Proc. of ISMAR2002, 2002.
- [8] H. Towles, Wei-Chao Chen, R. Yang, Sang-Uok Kum, H. Fuchs, N. Kelshikar, J. Mulligan, K. Daniilidis, L. Holden, B. Zeleznik, A. Sadagic, and J. Lanier. "3D Tele-Collaboration Over Internet2," In Proc. of ITP2002, 2002.
- [9] K. Daniilidis, J. Mulligan, R. McKendall, G. Kamberova, D. Schmid, and R. Bajcsy. "Real-Time 3D Tele-immersion," In the Confluence of Vision and Graphics, A. Leonardis et al. (Eds.), Kluwer Academic Publishers, pp. 253-266, 2000.
- [10] E. Cooke, P. Kauff, and O. Schreer. "Image-Based Rendering for Tele-Conference Systems," In Proc. of WSCG2002, 2002.
- [11] P.J. Narayanan, P. Rander, and T. Kanade. "Constructing Virtual Worlds Using Dense Stereo," In Proc. of ICCV'98, pp. 3 - 10, 1998.
- [12] H. Kato and M. Billinghurst. "Marker tracking and HMD calibration for a video-based augmented reality conferencing system," In Proc. of IWAR99, ACM, pp.85-94. 1999.