

複数のハンディカメラの協調利用による遮蔽物体除去映像のオンライン生成

On-line Video Synthesis for Removing Occluding Object via Complementary Use of Multiple Handheld Cameras

榎本 暁人[†], 正会員 斎藤 英雄[†]

Akihito Enomoto[†] and Hideo Saito[†]

Abstract We propose the use of an on-line video synthesis system that makes complementary use of multiple handheld cameras and thus removes occluding objects. We assumed that the same scene is captured with multiple handheld cameras, but an object occludes it. First, we used an ARTag marker to calculate a projection matrix, and then we estimated a homography between two captured images that was based on the projection matrix. We used this homography to transform the planar area into the image seen by the viewer. Finally, we blended the pixel value in accordance with the difference of the pixel values of the captured image and of the warped images at the same position. Doing this means that we remove the occluding object that cannot be approximated as a planar area. Our experimental results prove that this system can be used to remove the occluding object in a dynamic scene.

キーワード：拡張現実感, Augmented Reality, Diminished Reality, 複数カメラ, 遮蔽物体除去

1. まえがき

コンピュータビジョンの分野において、実写映像とCG等の仮想物体の合成は盛んに行われてきた。中沢ら¹⁾はA4用紙を置いたテーブルを動画で撮影し、各フレームにおける用紙の頂点を追跡することにより推定したテーブルに対するカメラの位置姿勢に基づいて、仮想物体のCGを合成する手法を提案した。このような実写映像と仮想物体の合成画像をオンラインで実時間処理し、ユーザに新しい視覚情報を提供する技術は、近年、拡張現実感(Augmented Reality)として様々なアプリケーションで利用されている²⁾。仮想物体をカメラで撮影された実写映像に付加させることでユーザを視覚的に支援することができるので、拡張現実感の技術は、ナビゲーションシステム、医療、設計、教育、ゲームなど、幅広い分野においての利用が期待されている。

仮想物体を実環境に存在するかのように重ね合わせて表示する技術である Augmented Reality に対して、適切な背景画像を遮蔽物体領域に重ね合わせることにより画像中

からあたかも遮蔽物体を取り除いたかのような映像を生成する技術は Diminished Reality と言われる。Diminished Reality の従来研究は大きく分けて二つのアプローチに分けることができる。

一つ目のアプローチは、1台のカメラで撮影された動画像の時系列情報を用いる手法である。Lepetit ら³⁾は、ユーザが指定した境界線を追跡してオクルージョンを検出することでシーン中の遮蔽物体を取り除く手法を提案した。Mann ら⁴⁾は、入力シーンから動画像列に撮影される平面物体を取り除き、別の平面テクスチャに置き換える手法を提案した。また、Wang ら⁵⁾は動画像列を複数のレイヤーに分割し、一つのレイヤーを取り除く場合には同じ動画像をレンダリングする手法を提案した。

二つ目のアプローチは、複数の位置に設置したカメラに撮影された画像を用いる手法である。Zokai ら⁶⁾は、大きく離れた3視点から撮影された静止画像の投影モデルを用いて適切な背景を生成することでユーザが指定した遮蔽物体の除去を行う手法を提案した。三つのカメラの役割はユーザカメラが1台、他の2台のカメラが背景画像生成用のカメラである。Zokai らの手法は背景領域の3次元復元を行うため、背景が非平面の場合でも扱うことができた。実際に、Zokai らの手法は工場の再設計において用いられ、有用性を証明した。しかし、複数のカメラの役割をあらかじめ決めておく必要があり、さらにすべてのカメラが固定で

2007年11月1日受付、2008年3月21日最終受付、2008年4月3日採録

†慶應義塾大学 大学院 理工学研究科

(〒223-8522 神奈川県横浜市港北区日吉3-14-1, TEL 045-563-1141)

†Graduate School of Science and Technology, Keio University

(3-14-1, Hiyoshi, Kouhoku-ku, Yokohama, Kanagawa 223-8522, Japan)

ある必要があった。

本論文では、お互いの位置姿勢関係が未知で常に移動しているフリーハンドの複数カメラを協調利用することにより、それぞれのカメラに撮影された遮蔽物が除去された画像をオンラインで生成することが可能なシステムを提案する。

Augmented Reality や Diminished Reality の分野では、複数カメラは盛んに利用されている。佐藤ら⁷⁾は、頭上に設置された固定カメラを用いて、ユーザの頭部の位置・姿勢を追跡することで、ユーザカメラの外部パラメータを修正する手法を提案した。Pilet ら⁸⁾は、複数の固定カメラを用いて、AR アプリケーションのための幾何学的キャリブレーション・光学的キャリブレーションを行うシステムを提案した。Klein ら⁹⁾は固定カメラを用いて、タブレット PC のに取り付けられた LED の追跡を行った。LED を追跡することで、タブレット PC に設置されたカメラの位置・姿勢を正確に推定できる。市原ら¹⁰⁾はロードサイドカメラを用いて、NaviView と呼ばれる運転支援システムを提案した。上記に述べたように、複数カメラは様々な拡張現実感の研究で用いられてきたが、複数カメラを用いたシステムの大部分が以下の制約を有している。

1) 固定カメラが必ず 1 台以上存在している。

2) 固定カメラはユーザの視点に応じて自由に動き回るカメラの位置や姿勢情報を与えたり他の視点からの映像情報を与えるための支援に用いられる。

このように、従来の複数のカメラを利用した AR システムでは、各カメラにはそれぞれ異なる独自の役割が与えられているものであった。

これに対して、本論文では、すべてのカメラが固定されておらず、ユーザによって自由に動き回っているという状況において、各カメラがその他のカメラの支援をお互いに行うことのできる複数カメラシステムを提案する。そして、このような状況における複数カメラシステムを用いて、各カメラ視点において対象物を遮蔽する物体を除去した映像のオンライン生成法を提案する。

本手法では、遮蔽物体を除去するために別の視点で撮影された背景テクスチャを合成しているが、中沢ら¹⁾の手法のように、1 台のカメラにより撮影された異なる視点の画像を利用して同様に、遮蔽物体を除去した画像を生成することも可能である。しかし、本論文の提案手法では、複数カメラにより撮影された異なる視点の画像を利用した処理をフレーム毎に行うことで、対象シーンの時間的変化にも対応可能である。

さて、本論文で提案するシステムを Diminished Reality の従来研究と比較すると、Zokai らのアプローチと同様に、異なる視点から撮影された複数の画像を使用するという意味では同様である。しかし、Zokai らの手法において、1 台のカメラは対象シーン撮影用のユーザカメラ、他の 2 台は背景画像撮影カメラというように、それぞれのカメラの役割は分かれていたのに対して、本論文の提案手法では、す

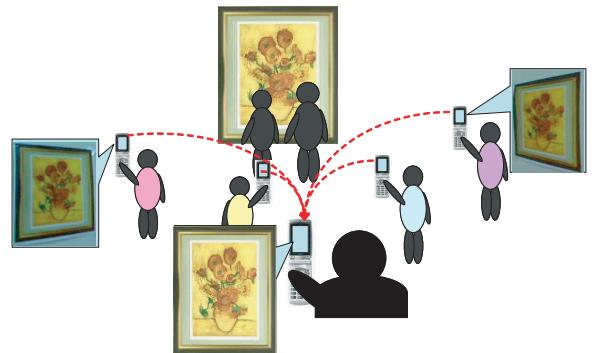


図 1 提案システム
Proposed system.

べてのカメラがユーザカメラになり、同時に背景画像撮影カメラにもなる。つまり、複数のカメラ画像を協調的に利用することにより、すべてのカメラにおいて遮蔽物体の除去が同時に可能となっている。また、Zokai らの手法においては遮蔽物体領域をユーザが指定するなどの手動の入力が必要であった。これに対し、提案手法では手動操作が必要でなく、また ARTag¹¹⁾を用いて簡易的にキャリブレーションを行えるため、オンラインシステムが可能となっている。オンライン処理を行える利点として、ユーザの自由な視点移動、遮蔽物体の移動、対象シーンの変化に対応して処理を行えるということがあげられる。

本稿の構成について述べる。まず 2 章では提案システムの概要について記述する。3 章では、オンラインでの Diminished Reality を実現するための提案手法について述べる。4 章では、本システムを用いて遮蔽物体を除去した結果を示すことで本システムの有効性を示す。そして 5 章で結論を述べる。

2. 提案システム

現在、あらゆる環境中に多くのカメラが存在している。例えば、ほとんどの人がカメラ付携帯電話を持っている。本論文では、図 1 に示されるように、複数の人が持つカメラを利用することで Diminished Reality システムを構築する。すべてのカメラの画像を共有することで、すべてのユーザ視点において遮蔽物体が除去された画像を生成することが出来る。このシステムを応用することで、カメラ付携帯電話を持っているすべての人が遮蔽物体のない対象シーンを見ることができる。例えば、美術館において有名な絵画が前にいる人々によって隠れてしまっても、本システムを用いれば遮蔽物体のない絵画の映像を手に入れることができる。

このコンセプトに基づいて、我々は図 2 に示すようなプロトタイプの実験システムを構築した。このシステムは複数の可動カメラから成り立つ。すべてのカメラは異なるユーザが持ち、異なる視点から同一のシーンを撮影する。実験システムにおいて、我々は異なる PC に接続された複数の Web カメラを用いる。それぞれの PC はネットワークに繋がつ

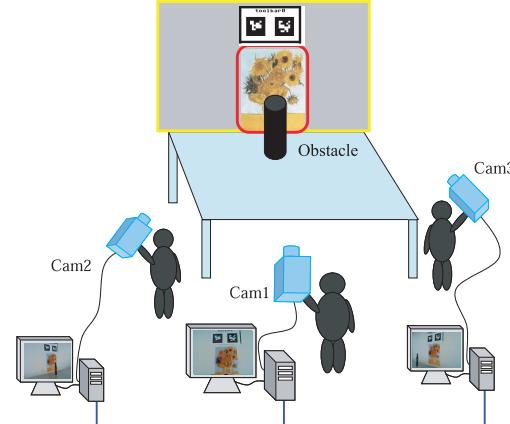


図 2 プロトタイプの実験システム
Prototype experimental system.

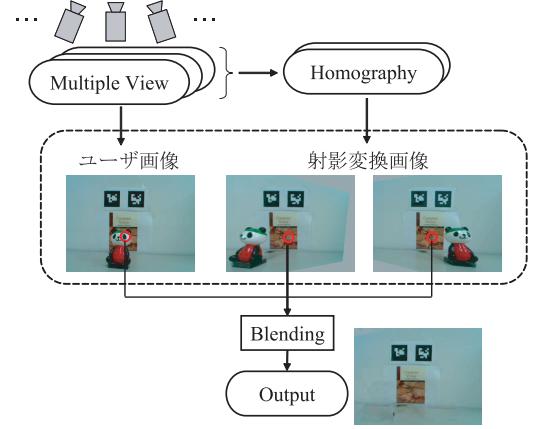


図 3 処理の流れ
Overview of proposed method.

ているので、ネットワークを通じて撮影された画像データの共有が可能である。ユーザは環境中に配置された ARTag マーカが検出される範囲で自由にカメラを動かすことで、オンラインでのキャリブレーションが可能となる。ARTag マーカを対象平面と同一平面上に配置した場合は、オフラインでの手動作業が必要でない。対象シーンと ARTag マーカが配置された平面が同一平面上でない場合、前処理としてオフラインで対象シーンの 3 次元位置を取得することで Diminished Reality を行うことが可能である。我々の提案システムでは、3 台以上のカメラの使用が可能であるが、今回は図 2 に示されるように、最小のシステム構成として 3 台のカメラを用いて提案システムの有用性を検証する。このシステムにおいて、我々は全てのカメラにおいて他の 2 台のカメラの画像を支援情報として遮蔽物体の除去を行う。本論文中で、我々は、遮蔽物体除去が行われるカメラをユーザカメラ、一方ユーザカメラの支援に用いられる他の 2 台のカメラをサポートカメラと呼ぶ。

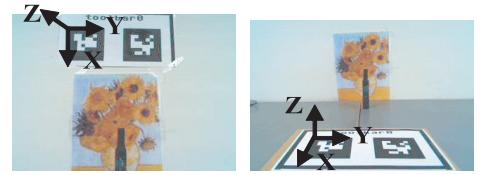
3. 遮蔽物体の除去

ここでは、提案システムを実現するために本研究で提案する手法について述べる。図 3 で示されるように、我々の提案手法は大きく分けて以下の二つの処理から成り立つ。

- (1) 射影変換画像の生成
- (2) ブレンド処理

一つ目の処理である射影変換画像の生成では、対象シーンは平面であると近似することで撮影された画像を他のカメラ視点に変換する。平面射影変換を行うために、ARTag を用いてカメラのオンラインキャリブレーションを行う。

二つ目の処理であるブレンド処理では、撮影された画像と他カメラの射影変換画像を統合する。我々は、ブレンド処理において、対象平面以外の領域すなわち遮蔽物体部分を除去する。以下、それぞれの処理について詳しく述べる。



(a) 対象平面中に配置 (b) 垂直な平面に配置
図 4 世界座標系の定義
Definition of the world coordinate system.

3.1 射影変換画像の生成

(1) ARTag による射影行列算出

我々は、カメラの射影行列 \mathbf{P} を算出するために ARTag と呼ばれるソフトウェアを利用する。ARTag とは画像処理によって、カメラの入力画像中からマーカを検出することで、オンラインでカメラの射影行列 \mathbf{P} を簡易的に算出することのできるソフトウェアである。ARTag により算出される射影行列は図 4 に示すように、XY 平面がマーカと水平、マーカの鉛直上向き方向が Z 軸の正となる世界座標系を持つ。

(2) 対象平面に対する Homography の算出

マーカ検出により得られた射影行列 \mathbf{P} を用いることで対象平面に対する Homography を算出する。対象平面と画像平面間の Homography が算出することができれば、対象平面に対するカメラ間での Homography は(1)式のようにして求まる。図 5 で示される対象平面上でのある点 \mathbf{M} と画像 A、画像 B 上での対応点 $\tilde{\mathbf{m}}_A$, $\tilde{\mathbf{m}}_B$ 間での Homography がそれぞれ \mathbf{H}_A , \mathbf{H}_B で表されたとする。この時、画像 A から画像 B への Homography(\mathbf{H}) は一般に以下の式で表される。

$$\mathbf{H} = \mathbf{H}_B^{-1} \mathbf{H}_A \quad (1)$$

次にカメラ間の Homography を求めるために必要な対象平面と画像平面間の Homography の算出法について述べる。図 4(a) で示されるように、マーカを対象平面上に配置した場合は世界座標系の $Z=0$ 平面が対象平面となる。射影行列 \mathbf{P} より $Z=0$ の時を考えると

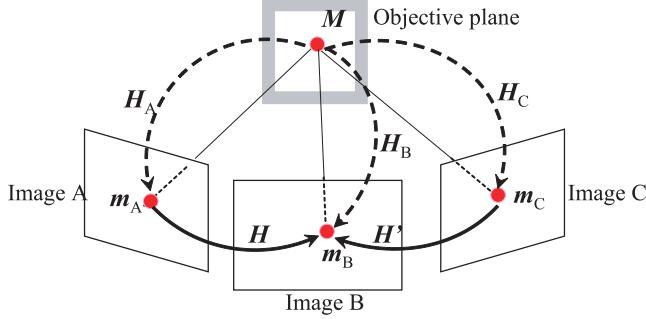


図 5 カメラ間の Homography
Homography of the objective plane between cameras.

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \simeq \mathbf{P} \begin{pmatrix} X \\ Y \\ 0 \\ 1 \end{pmatrix} \simeq \begin{pmatrix} p_{11} & p_{12} & p_{14} \\ p_{21} & p_{22} & p_{24} \\ p_{31} & p_{32} & p_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} \quad (2)$$

$$\simeq \mathbf{H} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}$$

となり、 3×4 の射影行列 \mathbf{P} を 3×3 の行列 \mathbf{H} で表すことができる。すなわち、この式は 2 次元座標から 2 次元座標への変換となる。よって、世界座標系の $Z=0$ 平面から画像座標系への Homography が算出できる。

3.2 ブレンド処理

ここでは、遮蔽物体の除去をするために行われるブレンド処理について詳述する。ブレンド処理とは、ユーザ視点画像と 2 枚の射影変換画像間で対応する画素の明度値の差に応じて、明度値を合成する処理である。ブレンド処理を行うことで実際に遮蔽物体の除去を行う。

(1) 明度値の補正

明度値を合成する前に、まず明度値の補正を行う。複数のカメラを用いるため、光源の方向等によりそれぞれのカメラで得られる画像の全体的な明度が異なる事がある。そのため、明度値の補正を行い各画像の全体的な明度を合わせる。明度値の補正には ARTag マーカ中の白い部分を利用する。図 6 は実際に画像の全体的な明度値の補正を行った結果である。図 6 の画像 A におけるマーカ中の白い部分の明度の平均値を \mathbf{C}_A 、同じく画像 B におけるマーカ中の白い部分の明度の平均値を \mathbf{C}_B とするとガンマ変換より以下の式が成り立つ。

$$\mathbf{C}_A = 255 \left(\frac{\mathbf{C}_B}{255} \right)^{\frac{1}{\gamma}} \quad (3)$$

(3) 式において未知数である γ の値を求め、画像 B をガンマ変換することにより、全体的な明度を画像 A に合わせることができる。図 6(c) は、画像 B の明度を画像 A に合わせた結果である。この明度値の補正処理を射影変換画像に対して行うことで、各画像の全体的な明度をユーザ画像に合わせることができる。ただし、強い鏡面反射光が、いずれかのカメラに撮影されてしまい、カメラ間の明度が極端

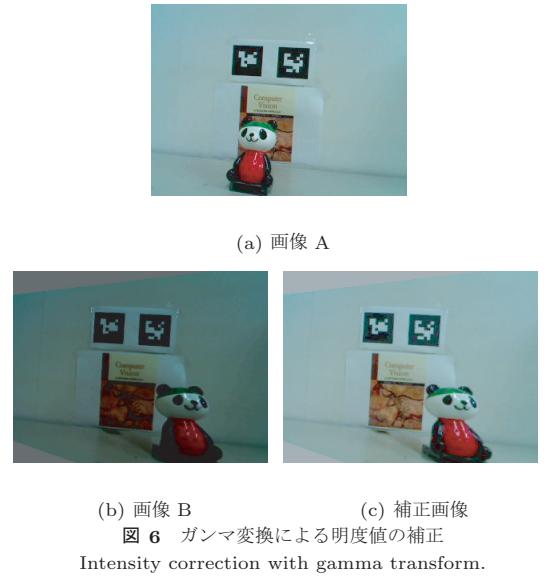


図 6 ガンマ変換による明度値の補正
Intensity correction with gamma transform.

に異なってしまうような場合には完全に補正できるわけではない。ここで補正は、あくまでもユーザ画像に他のカメラに簡易な手法で近づけることを狙ったものに過ぎない。

(2) ブレンド関数

次に、対応する三つの画素の明度値をどのように合成することで遮蔽物体の除去を行うのかについて述べる。遮蔽物体領域においては、図 7 のように、ユーザ視点画像の注目画素が遮蔽物体の表面であり、裏の対象平面が見れない。しかしながら、2 台のサポートカメラでは遮蔽物体の裏の対象平面の情報を得ることができる。よって、この 2 台のサポートカメラから得られる対象平面の情報を利用することで遮蔽物体の除去を実現する。

まず、ユーザ視点画像の明度値が 2 枚の射影変換画像の内、少なくとも一方の射影変換画像の明度値とほぼ等しい場合、この場合はユーザ視点画像においては対象平面領域であると考えられるのでユーザ視点画像の明度値を選択する。次に、ユーザ視点画像の明度値のみが他の二つの明度値と大きく異なる場合、この場合はユーザ視点画像においては遮蔽物体領域であると考えられるので、他の二つの明度値を選択する。

便宜上、画像座標 (i, j) の位置でのユーザ視点画像の明度値を \mathbf{p}_V 、射影変換画像 n の明度値を \mathbf{p}_n とする。ここで、画像座標 (i, j) の位置での出力画像の明度値を \mathbf{q} とすると、この明度値 \mathbf{q} を決定するためのブレンド関数を我々は以下のように \mathbf{p}_V と \mathbf{p}_n の大小関係により幾つかの場合に分けて定義するものとした。

$$Case1 \quad if(\mathbf{p}_V \approx \mathbf{p}_1 \approx \mathbf{p}_2) \quad \mathbf{q} = \mathbf{p}_V \quad (4)$$

$$Case2 \quad if(\mathbf{p}_V \approx \mathbf{p}_1 \neq \mathbf{p}_2) \quad \mathbf{q} = \mathbf{p}_V \quad (5)$$

$$Case3 \quad if(\mathbf{p}_V \approx \mathbf{p}_2 \neq \mathbf{p}_1) \quad \mathbf{q} = \mathbf{p}_V \quad (6)$$

$$Case4 \quad if(\mathbf{p}_V \neq \mathbf{p}_1 \approx \mathbf{p}_2) \quad \mathbf{q} = (\mathbf{p}_1 + \mathbf{p}_2)/2 \quad (7)$$

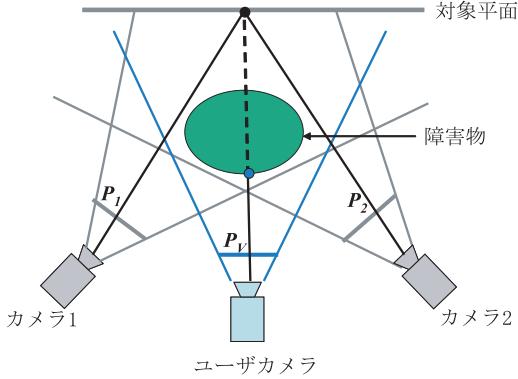


図 7 障害物に遮蔽された対象平面と三つのカメラの配置関係
Geometrical relationship between three cameras and objective plane behind obstacle.



図 8 ブレンド処理
Blending process.

Case1 は注目画素が全ての画像において、対象平面領域の場合である。この場合は、出力画像の明度値にはユーザ画像の明度値を用いる。*Case2* は注目画素が射影変換画像 2 の遮蔽物体領域、*Case3* は注目画素が射影変換画像 1 の遮蔽物体領域に存在する場合である。このような場合には、*Case1* と同様に出力画像の明度値にはユーザ画像の明度値を用いる。*Case4* は注目画素がユーザ画像の遮蔽物体領域に含まれる場合である。この場合、ユーザ視点画像の明度値は他の 2 枚の射影変換画像の明度値とは大きく異なる。この時には、図 8 のように出力画像の明度値には 2 枚の射影変換画像の明度値を用いる。*Case4* のみユーザ画像の明度値を用いないが、ブレンドを行う前に、明度値の補正を行い全体的な明度をユーザ画像に合わせているため、出力において明度値が不連続になることを防いでいる。*Case1*～*Case4* の式を用いることで、ユーザ視点画像の遮蔽物体領域では、別の視点から射影変換された画像の明度値を用いることとなるので、遮蔽物体が出力画像において除去されるのである。このブレンド処理は明度値の差のみを用いて遮蔽物体の除去を行っているため処理に時間がかかりず、また遮蔽物体領域指定の手動の処理が必要でないのでオンラインシステムに適している。

3.3 対象平面以外への ARTag マーカの配置

前項で述べたように、ARTag マーカを対象平面上に配置した場合は、対象平面に対するカメラ間の Homography がオンラインで算出できる。しかしながら、たとえ ARTag マーカを対象平面上に配置することができなくても、オフライン処理によって ARTag マーカと対象平面の位置関係を取得することで遮蔽物体除去の処理を行うことができる。

ARTag マーカを対象平面以外の位置に設置した場合は、オフライン段階において対象平面の 3 次元位置を指定する必要がある。

ここでは、マーカを対象平面に対して垂直な平面に配置した場合を考える。マーカを垂直な平面に配置した場合は、世界座標系の $X=d$ の平面が対象平面となる。この場合、対象平面に対する Homography は以下の式によりもとまる。

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \simeq \mathbf{P} \begin{pmatrix} d \\ Y \\ Z \\ 1 \end{pmatrix} \simeq \begin{pmatrix} p_{12} & p_{13} & p_{14} + dp_{11} \\ p_{22} & p_{23} & p_{24} + dp_{21} \\ p_{32} & p_{33} & p_{34} + dp_{31} \end{pmatrix} \begin{pmatrix} Y \\ Z \\ 1 \end{pmatrix} \quad (8)$$

$$\simeq \mathbf{H} \begin{pmatrix} Y \\ Z \\ 1 \end{pmatrix}$$

4. 実験・結果

ここでは、実験を通して提案システムの有用性を証明する。マーカを対象平面と同一平面上に配置した場合、マーカを垂直な平面に配置した場合の二つのマーカ配置について、提案手法を用いて遮蔽物体の除去が実現できているかを調べた。本実験は、最小のシステム構成としてカメラ 3 台を用いて行った。Web カメラには、Logicool の QCam Fusion を用いた。なお、各カメラの入力画像サイズは 320 × 240 画素である。本実験で用いた PC のスペックは以下の通りである。

- CPU : Intel Pentium4 Processor 3.2GHz
- メモリー: 2048MB
- OS : Windows XP

4.1 マーカを対象平面上に配置

マーカを対象平面上に配置した場合について、対象平面の前に置かれた遮蔽物体を本手法を用いて除去した。

まず、3 台のカメラの内、中央のカメラをユーザカメラとして処理を行った時の結果を図 9 に示す。図 9(a) が入力画像、図 9(b) が出力画像である。図 9(a) においてユーザカメラは自由に視点を移動しているが、このように視点が移動した場合でも提案システムを用いることで遮蔽物体が除去され、見たい対象平面の情報が得られることがわかる。次に、明度値の補正の有用性について示す。図 10 は補正の有無による結果の比較である。図 10(c),(d) はそれぞれ補正なし・ありの時の出力画像である。図 10(c) は補正をしていない場合の除去結果であるが、射影変換画像中の遮蔽物体がノイズとして結果に影響を及ぼしてしまっている。しかしながら、図 10(d) に示されるように、明度値の補正を施すことでノイズの影響を軽減し、出力画像の明度値が不連続になること防いでいる。明度値の補正を行うことでより違和感のない除去を行うことができる。明度値補正の有効性について、定量的に評価した結果を表 1 に示す。評価には、実際に遮蔽物体を除いて撮影した正解画像(図 10(b))と、補正なし・ありのそれぞれの出力画像との

表 1 明度値補正の定量的評価
The quantitative evaluation of intensity correction.

出力画像	PSNR(dB)
補正無	26.7
補正有	32.7



(a) 入力 (b) 出力
図 9 マーカを対象平面上に配置
Marker on the object plane.



(a) 入力 (b) 出力
図 11 すべてのカメラにおける除去結果
Removig result in all cameras.



(a) 入力画像 (b) 正解画像
(c) 出力画像 (補正無) (d) 出力画像 (補正有)
図 10 明度値補正の有無による結果の比較
Comparison of result by using of intensity correction or not.

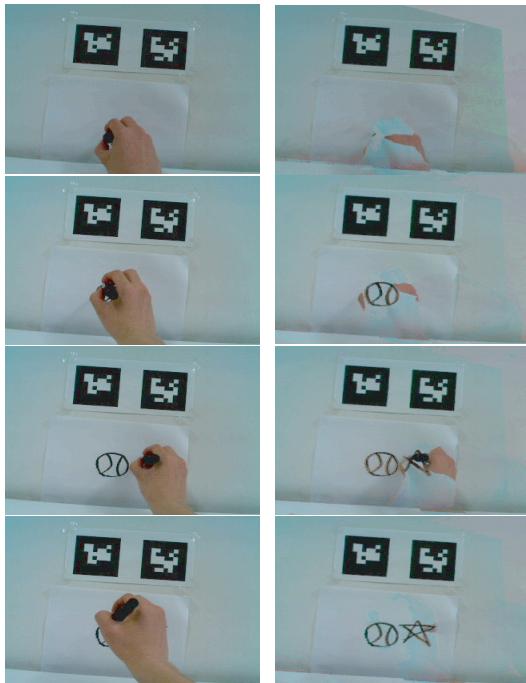
PSNR(Peak Signal-to-Noise Ratio) を用いた。PSNR は値が大きいほど、正解画像との差が少ないということを示している。定量評価の結果より、補正ありの出力画像の方が PSNR の値が大きく、より正解画像との差が少ないことがわかる。

また、3 台のカメラ全てをユーザカメラとし、それぞれのカメラで遮蔽物体を除去した結果を図 11 に示す。図 11 は上から順に、中央・左・右カメラをユーザカメラとした時の入力画像と出力画像であるが、どの位置からでも遮蔽物

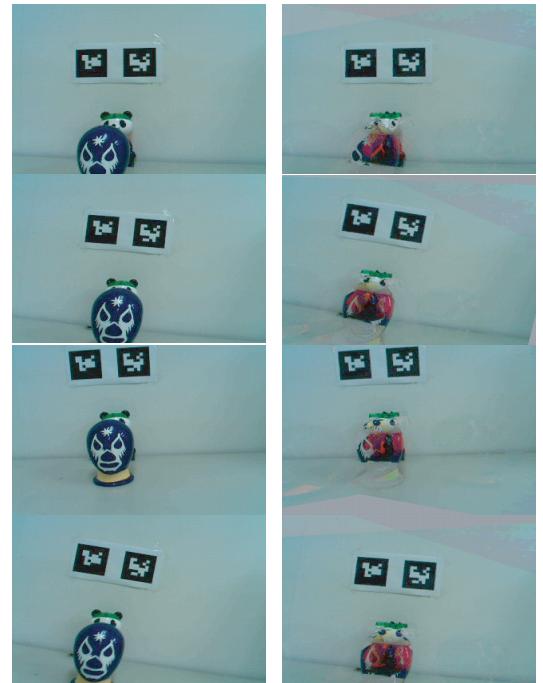
体が除去されている。提案システムにおいては、個々のカメラが固有の役割を持っているわけではなく、すべてのカメラがユーザカメラにもサポートカメラにもなり得る。そのため、各カメラ画像を協調的に利用することにより、すべてのカメラ中から遮蔽物体を除去することができる。

図 12 は対象平面上の白紙に人が絵を書いているという状況である。すなわち、これは遮蔽物体が移動し、かつ対象平面の情報が時間的に常に変化している場合である。この時ユーザカメラにおいては、人の手が遮蔽物体となってしまい描いている絵(ボールと星)が隠れてしまっている。しかし、提案手法を用いることで人の手を消し、描いている絵を見ることができた。提案手法では、前フレームの情報を用いずに現フレームの情報のみで遮蔽物体の除去を行っているため、フレーム間でシーンが変化するような場合でも除去を行うことができる。出力画像には一部手が消えず残っている部分があるが、これはユーザ視点画像中の遮蔽物体領域と射影変換画像中の遮蔽物体領域が重なってしまったため、ブレンド処理がうまくいかず遮蔽物体が残ってしまったものと考えられる。

上記の実験結果では、見たい対象が完全な平面のみであったが、対象が平面ではなく立体の場合に対して、本手法を用いて処理を行った結果を図 13 に示す。図 13 より、対象が平面の時と比べて出力画像における見たい対象物体はボケてしまっている。このようにボケてしまった理由として、



(a) 入力
図 12 対象シーンが変化した場合の結果
Result of dynamic scene.



(a) 入力
図 13 対象が非平面の場合の結果
Result in nonplanar scene.

マーカを対象平面上に配置した場合は世界座標系 $Z=0$ の位置に平面が存在していると仮定して処理を行っている。しかしながら、実際は立体なため平面同士の対応付けである Homography を用いて処理を行った際に誤差が生じてしまうためである。

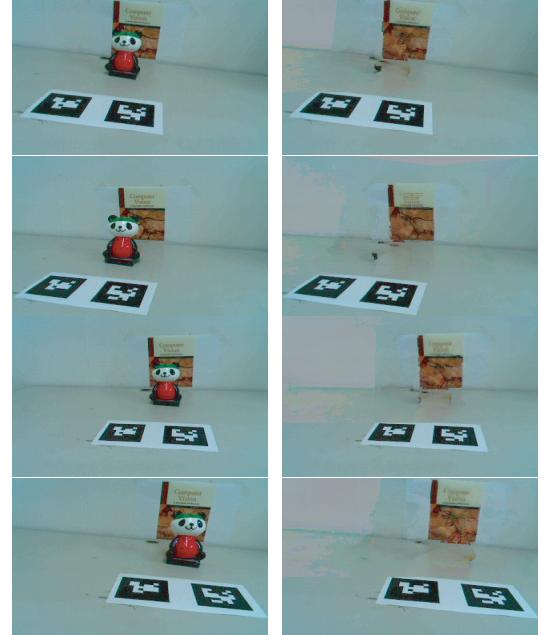
なお、処理速度はネットワークを通じて処理を行うと約 1fps であり、処理速度の観点からはまだ充分なものとは言えない。しかしながら、画像処理部分(射影変換画像の生成+ブレンド処理)のみの処理速度では約 7fps であった。また、本システムは、遮蔽物体除去にはブレンド処理を用いたので手動の操作が必要ない。よって、現段階ではネットワークの通信速度が処理速度に大きな影響を及ぼしているが、今後通信速度の向上が図られれば充分オンラインでの遮蔽物体除去システムであるといえる。

4.2 マーカを対象平面と垂直な平面上に配置

マーカを対象平面と垂直な平面上に配置した場合について本手法の処理を行った。マーカを対象平面と垂直な平面上に配置した場合は、対象平面の 3 次元位置情報が必要となる。今回はオフライン処理を行い、前もって対象平面の 3 次元位置を算出しておいた。結果を図 14 に示す。図 14 より、マーカを対象平面と垂直な平面上に配置した場合でも入力画像における遮蔽物体が除去されていることがわかる。このように、対象平面のマーカからの 3 次元位置さえ分かれば、処理を行うことができる。

4.3 異なるスペックのカメラを用いた結果

ここでは、異なるスペックのカメラを用いた場合の



(a) 入力
図 14 マーカを対象平面と垂直な平面上に配置
Marker on the perpendicular plane to the object plane.

結果を示す。上記の実験結果では、カメラ 3 台とも Logicool の Qcam Fusion を用いていたが、異なる 3 機種のカメラ、UCAM-E1W30(ELECOM), Qcam Fusion(Logicoool), Qcam Pro(Logicoool) を 1 台ずつ同時に用いた結果を図 15 に示す。図 15 は上から順に、UCAM-E1W30, Fusion, Pro をユーザカメラとして除去をした結

表 2 カメラの名称とスペック
A name and the specifications of the cameras.

カメラ名	画素数	撮像デバイス
ELECOM UCAM-E1W30	30 万画素	CCD
Logicool Qcam Fusion	130 万画素	C-MOS
Logicool Qcam Pro	200 万画素	C-MOS

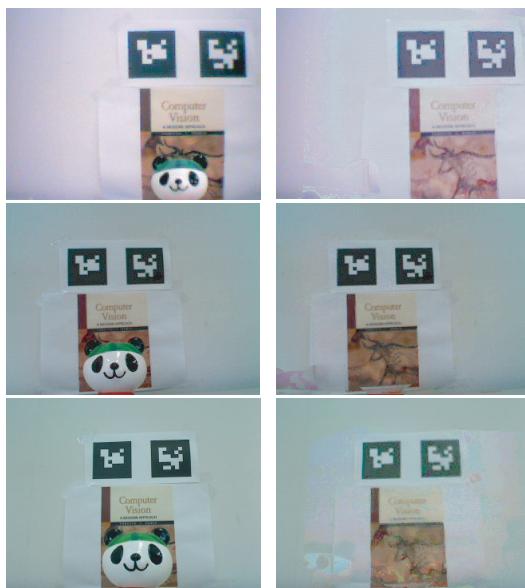


図 15 異なるスペックのカメラを用いた場合の結果
Results with different camera specifications.

果である。表 2 は、それぞれのカメラのスペックである。なお、最大フレームレートは 3 台とも 30fps である。

提案手法において、特定のカメラに依存する処理はしていない。そのため、異なるカメラを用いた場合でも原理的には除去を行うことができる。実際に図 15 に示されるように、3 台とも異なるスペックのカメラを用いた場合でも遮蔽物体の除去を行うことができた。

5. む す び

本論文では、複数のハンディカメラの協調利用による遮蔽物体除去映像のオンライン生成システムを提案した。提案システムにおいて、すべてのカメラは固定である必要がなく、各カメラの画像を協調利用することによりすべてのカメラにおいて遮蔽物体の除去を実現した。また、提案システムでは、前フレームの時系列情報を用いずに現フレームにおける複数視点画像を利用することにより、時間的にシンクが変化するような状況でも処理が行えることを示した。

本論文では、簡易的にキャリブレーションを行うために ARTag マーカを用いたが、最近は、SIFT 特徴量¹²⁾などの自然特徴点の検出に基づいたキャリブレーション手法が多数提案されており、このような手法を利用すれば将来的には、マーカレスでのシステム実現は充分可能であると思われる。また、今回の処理では領域を平面近似して処理を行っているため、対象が立体の時は対象にボケが生じてしま

ったが、今後カメラの台数を増やす、オンラインでの奥行き算出などをを利用して立体にも対応できるように処理を改良する予定である。

【文 献】

- 中沢祐二、中野誠巳、小松隆、斎藤隆弘：“画像中の特徴点に基づく実写画像と CG 画像との動画像合成システム”，映情学誌，51, 7, pp.1086-1095(1997)
- R.Azuma, Y.Baillot, R.Behringer, S.Feiner, S.Julier and B.MacIntyre: “Recent advances in augmented reality”, IEEE Computer Graphics and Applications, 21, 6, pp.34-47 (Nov. 2001)
- V.Lepetit and M.-O.Berger: “A semi-automatic method for resolving occlusion in augmented reality”, In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2, pp.225-230 (Jun. 2000)
- S.Mann and J.Fung: “VideoOrbits on Eye Tap Devices for Deliberately Diminished Reality or Altering the Visual Perception of Rigid Planar Patches of a Real World Scene”, Proceedings of the Second International Symposium on Mixed Reality, pp.48-55 (Mar. 2001)
- John Y. A. Wang and Edward H. Adelson: “Representing moving images with layers”, IEEE Transactions on Image Processing Special Issue: Image Sequence Compression, 3, 5, pp.625-638(Sept. 1994)
- S.Zokai, J.Esteve, Y.Genc and N.Navab: “Multiview Paraperspective Projection Model for Diminished Reality”, Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality, pp.217-226(Oct.2003)
- K.Satoh, S.Uchiyama and H.Yamamoto: “A Head Tracking Method Using Bird's-Eye View Camera and Gyroscope”, Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality, pp.202-211(Nov.2004)
- J.Pilet, A.Geiger, P.Lagger, V.Lepetit and P.Fua: “An All-In-One Solution to Geometric and Radiometric Calibration”, Proceedings of the Fifth IEEE and ACM International Symposium on Mixed and Augmented Reality, pp.69-78(Oct.2006)
- G.Klein and T.Drummond: “Sensor Fusion and Occlusion Refinement for Tablet-based AR”, Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality, pp.38-47 (Nov.2004)
- E.Ichihara and Y.Ohta: “NaviView: Visual Assistance Using Roadside Cameras- Evaluation of Virtual Views”, 3rd IEEE Conference on Intelligent Transportation Systems, pp.322-327 (Oct.2000)
- M.Fiala: “ARTag, a fiducial marker system using digital techniques”, Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition , 2, pp.590-596(Jun.2005)
- I.Skrupnyk and D.G.Low: “Scene Modelling, Recognition and Tracking with Invariant Image Features”, Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality, pp.110-119 (Nov.2004)



えのもと あきひと
複本 晃人 2007 年、慶應義塾大学理工学部情報工学科卒業。現在、慶應義塾大学大学院理工学研究科前期博士課程に在学中。コンピュータビジョン、特に拡張現実感に関する研究に従事。



ときとう ひでお
常藤 英雄 1992 年、慶應義塾大学院博士課程修了。同年同大学助手、1995 年、専任講師、2001 年助教授、2006 年、情報工学科教授。コンピュータビジョン、画像処理に関する研究に従事。1997～1999 年、日本学術振興会海外特別研究員として米国カーネギーメロン大学ロボティクス研究所へ留学。2000～2003 年科学技術振興事業団さきがけ 21『情報と知』研究員。博士（工学）正会員。