

時間方向への外挿・内挿に基づく前景領域抽出とフレーム補間

千代 裕介[†] 阪野 貴彦^{††} 大石 岳史^{†††} 池内 克史^{†††}

[†] 東京大学大学院学際情報学府 〒 113-0033 東京都文京区本郷 7-3-1

^{††} 東京大学生産技術研究所 〒 153-8505 東京都目黒区駒場 4-6-1

^{†††} 東京大学大学院情報学環 〒 113-0033 東京都文京区本郷 7-3-1

E-mail: †{chishiro,vanno,oishi,ki}@cvl.iis.u-tokyo.ac.jp

あらまし 本研究では、人物映像による複合現実感環境 (Mixed Reality, 以下 MR) への人物コンテンツ埋め込みを大きな目的とし、課題となるマスク処理のコストと MR の時間的整合性問題の解決を行う。マスク処理では既知のマスク領域の外挿から人物領域を推定し、映像の見え特性のみに依存しない自動抽出手法を提案する。大量の人物映像から効率的にマスクを抽出する事が可能であり、MR における物語性の演出が容易となる。時間的整合性では人物映像と MR のフレームレート非同期による見えの崩れを防ぐため、マスクの内挿によるオクルージョンに安定したフレーム補間手法を提案する。

キーワード 複合現実感, 人物コンテンツ, 映像の領域分割, フレーム補間, 仮想歴史体験

1. はじめに

複合現実感 (Mixed Reality, 以下 MR) は、実世界映像に CG で描かれた仮想物体をリアルタイムに重畳する技術である。作業支援や認知支援、さらには娯楽体験の価値向上などが可能であり、教育、医療をはじめ幅広い分野で応用が進んでいる [1]~[5]。

MR の新たな活用として、歴史的建築物のデジタル復元が注目されている。一部または全てが失われた遺構において、CG 復元された遺跡のようすを重ね合わせて表示することにより、あたかも当時の風景が甦ったかのような体験が出来る [6], [7]。現地の空気に触れながら鑑賞することで、シアターで鑑賞する VR に比べてより高い臨場感もたらされる。

これまでの歴史体験 MR では風景のデジタル復元が主であったが、今後は多数の登場人物、ナレーション、場面転換等を通して物語性を付加していく事が重要となる。歴史を題材とするエンターテインメント作品においても、当時の人々の織りなすドラマが鑑賞者を歴史の世界に引き込む一助となっている。また、歴史的な事由の提示することによって教育的効果も期待できる。

物語性の演出において重要となる人物モデルについては、背景との親和性を考えると、実写ベースである事が望ましい。大蔵ら [8] は屋外環境でステレオ撮影した人物を 3 次元復元し、MR システムへの組み込みを行った。しかしコンテンツの取得コスト、描画コストは共に大きく、多数の登場人物の配置には適さない。そこで、本論文では役者の演じる実写映像 (以下人物映像) を 2D テクスチャとして用いる。素材の敷居を低くする事で、歴史ドラマ、映画等のアーカイブ映像から人物コンテンツを作成する事も可能となり、完成度の向上を期待でき

る。また描画時の負担も小さい事から、現在注目を集めるユーザ移動型 MR [9] で使用されるポータブル PC、スマートフォンにも組み込みやすいと考えられる。

人物映像の MR コンテンツ化には大きく 2 つの課題がある。1 つ目は、MR に不要な背景領域を除去し、前景 (人物) 領域のみを抽出するマスク処理の作業コストである。前景の誤抽出は MR の見えに著しく悪影響を与えるため、人間の目で一枚ずつ確認しながら行わなければならない。人的コストが膨大となるため、何らかの自動化処理が必要となる。

2 つ目は、時間的整合性の崩れである。MR システムは描画シーンの複雑さに応じてフレームレートが可変となるため、人物映像のフレームレートと同期させるのが難しい。これまでは人物映像側のフレームを水増し、中抜きして時間の進みを調整したが、見えに違和感が生じてしまう問題があった。解決するには予め映像中の人物の動きを抽出し、MR 実行時にフレーム補間表示を行う必要がある。

本論文では人物映像の MR コンテンツ化支援のために、マスク処理と時間的整合性の問題を解決する手法を提案する。2 つの問題の解決には、マスク領域の時間方向に対する外挿、内挿という共通の枠組みを用いる。本論文で提案する手法を以下に示す。

- マスク外挿による前景領域抽出
- マスク内挿によるフレーム補間

前景領域抽出では、連続フレームのマスク間で輪郭の DP マッチングを行い、これを次フレームに外挿する事で前景領域の推定を行う。マスク領域の情報に注目する事で、人物領域の色分布やオクルージョンに安定した処理を行う事が出来る。また、オプティカルフローによる推定も同時に行うことで、様々な映像に対して精度の良

い抽出を実現する．

マスク内挿によるフレーム補間では，同じく輪郭の DP マッチングを用いて解空間を限定したオプティカルフローの推定を行い，補間画像の生成を行う．輪郭の対応関係を内挿するため，輪郭付近において精度の良い補間を行う事が出来る．

本論文の構成は以下の通りである．次章では，関連研究について述べる．第 3 章ではマスクの外挿による前景領域の抽出について説明し，第 4 章では，マスクの内挿によるフレーム補間について説明する．第 5 章では実験結果について示し，第 6 章にて本論文のまとめを行う．

2. 関連研究

2.1 映像からの前景領域の抽出

一般にマスク処理と言えば撮影時に特定色のスクリーンを背景とするクロマキー処理などが広く用いられているが，本研究で対象とする映像の背景は未知とする．これによって，商業ドラマなどの既存のアーカイブ映像を MR に組み込めること，また，駆ける馬など屋外での移動撮影が必要なコンテンツの作成も容易となる．

背景が未知の映像から前景抽出を行う手法として，グラフカット [10], [11] による領域分割が提案されている．隣接画素の輝度差を考慮する事で，領域の連続性を拘束したセグメンテーションが可能である．Boykov らによるグラフカット [10], [11] は静止画が対象であったが，映像へ適用する研究も多く報告されている．Wang ら [12] は映像中の全ピクセルをノードとする三次元グラフを構築し，グラフカットを行った．前処理として簡易なセグメンテーションを行い，処理速度を向上させている．Li ら [13] は局所学習された色分布から領域尤度を求め，フレーム毎にグラフカットを行った．ユーザは数フレーム毎に正しい領域情報を与える必要がある．Bai ら [14] はオプティカルフローによってマスク形状の伝搬を行う事で，前景と背景の色分布が類似した映像でも安定した抽出を実現した．フレームバイフレームで処理が進み，操作性にも優れている．

これまでに提案された手法の問題として，抽出精度が映像の見えの特性のみに依存する事が挙げられる．例えば色分布の違いを手がかりに分割を行う手法 [12] では，輪郭付近の色分布が類似している場合に誤抽出が生じる．オプティカルフローを用いてこれを補う手法 [14] も提案されているが，輪郭付近ではオクルージョンが発生し易いため，オプティカルフローの精度自体が信用できない．背景の見えによっては誤抽出領域が拡大する恐れもある．本論文では既知のマスク領域の外挿によって前景領域の推定を行う手法を提案する．

2.2 フレーム補間

フレーム補間では，まず前処理としてフレーム間の動き情報の抽出を行い，得られた動き情報を基に中間フ

レームにおける補間画像の生成を行う．

動き情報を抽出する手法として，オプティカルフローが有名である．Horn ら [15] はエネルギー式の最小化によって，隣接画素のフローを平滑化する手法を提案し，現在広く用いられている．ただし，オプティカルフローは映像によって精度が不安定という問題がある．特に (1) オクルージョンや，(2) 大きく推移する局所領域において精度が低下する可能性が大きい．人物領域の輪郭の補間精度は MR の見えに強く影響を与えられられるため，(1)(2) の場合においても安定した精度で推定を行える事が望ましい．

(1) の問題に対し，Kolmogorov ら [16] はマルチラベルのグラフカットを用いて連続フレーム内の全ての画素を 1 対 1 に対応付ける事により，オクルージョンの検出精度を向上させた．また Xiao ら [17] はフレーム間の推移を双方向から求め，両者の比較からオクルージョン関係を求めた．しかしどちらの手法も映像の見えのみに依存する問題点は残る．

(2) の問題に対し，Brox ら [18] は特徴量によるマッチングを行うことで，大きく推移する局所領域に対しても精度の良いフロー計算を行った．しかしテクスチャの少ない領域では適用できず，汎用的な効果は期待できない．

本論文では前処理として連続フレームの人物領域のマッチングを行い，解空間を限定したオプティカルフローの推定手法を提案する．提案手法によって (1)(2) の問題を同時に解決する事が出来る．

3. マスクの外挿による前景領域の抽出

本章では，既知のマスク領域の外挿によって前景領域の抽出を行う手法を提案する．マスク外挿の概要図を図 1 に示す．まず抽出対象フレーム (t とする) の前 2 フレームのマスクから輪郭を抽出し，マッチングを行う．マッチングより得られた輪郭画素の推移を外挿する事で，フレーム t における推定マスクを得る事が出来る．これを色分布による推定結果と足し合わせ，グラフカットによって前景抽出を行う．本手法により，既存のグラフカット手法の長所を活かしつつ，色分布やオクルージョンの特性に安定した抽出が可能である．

本手法ではフレーム $t-2, t-1$ の情報を用いてフレーム t における前景の抽出を行う．抽出処理は以下の 3 ステップから構成される．

STEP1: 前景領域の推定

STEP2: 領域尤度の算出

STEP3: グラフカットによる前景抽出の高精度化

STEP1 では DP マッチングを用いてマスクの外挿を行い，前景領域の推定を行う．ただし，外挿だけでは回転運動などの複雑な動きに対処できない．そこでオプティカルフローの推定も同時に行い，信頼性の比較から両者の選択を行う．STEP2 では推定された前景領域と色分布に基づき，フレーム t の全画素に対して前景・背景の

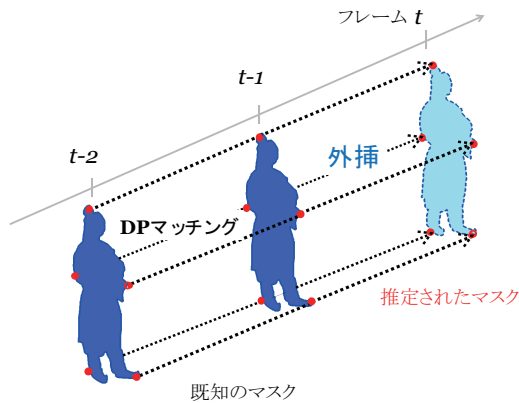


図 1 マスクの外挿の概要図

領域尤度を算出する．色分布の学習は Bai ら [14] の手法を用いた．STEP3 では得られた領域尤度と隣接画素の輝度差からエネルギー式を定義し，グラフカットを用いて前景抽出の精度を高める．

フレーム t での画像及びマスクを I_t, L^t とすると，以上 3 つのステップに必要な入力 $I_{t-1}, I_t, L^{t-2}, L^{t-1}$ であり，求める出力は L^t である．この処理を全フレームに逐次適用する事で，映像中の全ての前景領域が得られる．ただしユーザによる初期入力として，映像の第 1, 2 フレームのマスクを与えるものとする．

3.1 前景領域の推定

本節ではフレーム t における前景領域を推定する手法について説明する．提案する手法ではまず L^{t-2}, L^{t-1} から輪郭 C^{t-2}, C^{t-1} を抽出し，輪郭間の対応を DP マッチングによって求める．得られた輪郭上の推移を t に外挿する事で， L^t の推定を行う．輪郭の情報を用いる事で，映像の見えにのみ依存するという先行研究の課題を解決する事が出来る．

ただし輪郭 DP マッチングの課題として，複雑な動きをする領域において形状推定にエラーが生じてしまう事が挙げられる．これはフレーム $t-2, t-1$ 間の推移をそのままフレーム t へ外挿したことによるもので，回転運動などに対応する事が出来ない．そこで， I_{t-1}, I_t 間のオプティカルフロー推定も同時に行い，マスクの外挿による推定結果と相補的に用いる事で精度の向上を図る．なお，どちらの推定結果を用いるかは前景内の局所領域ごとに選択するものとする．選択の基準として推定の信頼性を算出し，より信頼できる推定手法を適応的に選択する．

本節で提案する前景領域の推定の流れを以下にまとめる．

- 輪郭 DP マッチングによる前景領域の推定
- オプティカルフローによる前景領域の推定
- 両推定手法の適応的な選択

それぞれの詳細について次に述べる．

3.1.1 輪郭 DP マッチングによる前景領域の推定

L^{t-2}, L^{t-1} の輪郭マッチングを行い，外挿処理からフレーム t における前景領域 L^t の推定を行う．輪郭マッチングには Fujiyoshi ら [19] の DP マッチングを用いた．マッチング処理の流れは以下の通りである．

- (1) L^{t-2}, L^{t-1} から輪郭 C^{t-2}, C^{t-1} を抽出
- (2) 輪郭画素の重心 x_G^{t-2}, x_G^{t-1} を算出
- (3) C 上の全ての画素に対して x_G との距離を DP マッチングのコストとして設定
- (4) 始点と終点が一致するように DP マッチング
- (5) 最小コスト時のバックトレースにより C^{t-2}, C^{t-1} の対応関係を取得

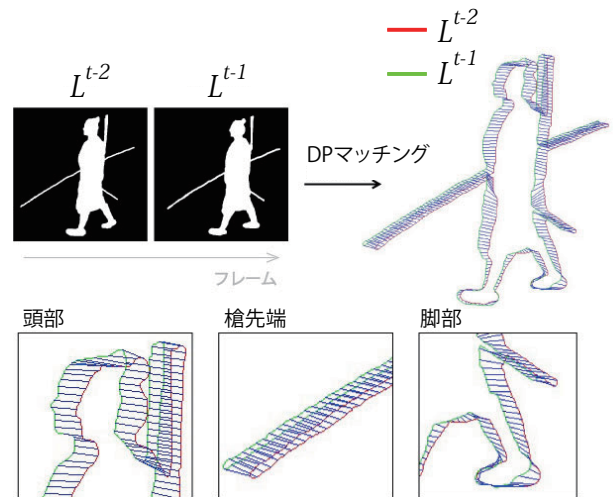


図 2 DP マッチングによる輪郭の対応結果

DP マッチングの結果を図 2 に示す．この対応結果より， C^{t-1} 上の各画素の推移量を得る事が出来る． C^{t-1} 上の画素を $c_k (k = 0, 1, \dots, N_c - 1)$ ，得られる推移量を $u_{DP}(c_k)$ とすると，フレーム t における輪郭画素 \tilde{c}_k は，

$$\tilde{c}_k = c_k + u_{DP}(c_k) \quad (k = 0, 1, \dots, N_c^{t-1}) \quad (1)$$

と計算でき， \tilde{c}_k を $k = 1, 2, \dots$ と順に結ぶ事で求める前景領域の推定 \tilde{L}_{DP}^t を得る．

3.1.2 オプティカルフローによる前景領域の推定

一般的にオプティカルフローの計算コストは大きい．しかし本抽出システムでは，人の目で抽出エラーが無いが確認しながらフレーム処理を進めるため，処理のレスポンスは早い方が好ましい．そこで本手法では高速かつ精度も良い Wedel ら [20] のアルゴリズムを用いた． I_{t-1}, I_t 間の推移を推定した結果を図 3 に示す．槍などの細長い領域や，色分布の類似したオクルージョン付近で推定にエラーが生じている事がわかる．

輪郭上の推移を $u_{OF}(c_k)$ とすると，オプティカルフローによる前景領域の推定結果 \tilde{L}_{OF}^t は式 (1) と同様に求める事が出来る．

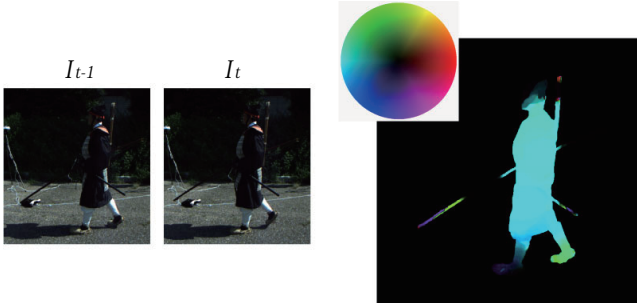


図3 オプティカルフローの推定結果

3.1.3 両推定手法の適応的な選択

両推定の結果を比較するため、 $\tilde{L}_{DP}^t, \tilde{L}_{OF}^t$ を I_t に重ねたようすを図4に示す．頭部、槍先端ではマスク外挿が良い推定結果を示している．まず頭部では、頭と背景の色が似ているため、オプティカルフローによる推定はうまく行えなかった．また槍の先端では、槍の幅に比べて大きく推移しており、オプティカルフローではやはりうまく計算できていない．一方で、脚部ではオプティカルフローによる推定の方が良い結果を示した．これは足が歩行によって回転運動をしており、マスク外挿では対応できなかったためと思われる．

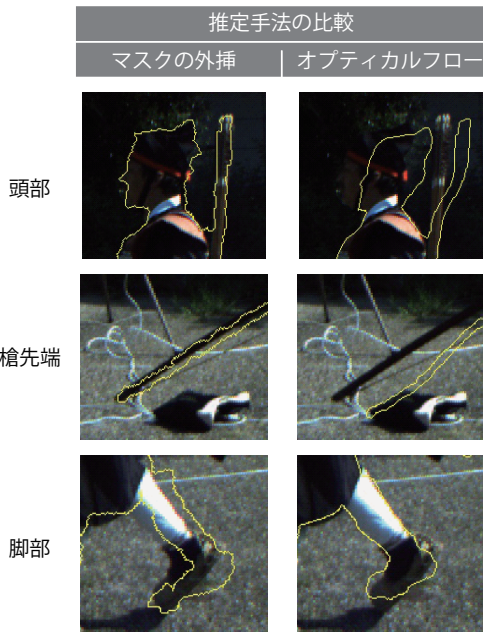


図4 両手法の推定結果の比較 (黄線は推定された前景領域の輪郭)

このように両手法はそれぞれ異なる長所を持つので、局所領域ごとに最適な推定手法を選択していくのが望ましい．そこで、本論文では両推定の信頼性を算出し、この値に基づいて推定手法を適応的に選択する手法を提案する．信頼性の算出は L^{t-1} の輪郭 C^{t-1} 上の画素ごとに行い、 $c_k (k = 0, 1, \dots, N_c - 1)$ は信頼性の指標として f^{DP}, f^{OF} の値を持つものとする．

信頼性の算出は推移前後の見えの変化をブロックマッ

チングによって評価した．探索ブロックを W として、以下の式より求める．

$$f^X = \frac{\sum_{x \in W} \{I_t(x + u_X(c_k)) - I_{t-1}(x)\}^2 \cdot L^{t-1}(x)}{\sum_{x \in W} L^{t-1}(x)} \quad (2)$$

なお、 $L^t(x)$ はある画素 x が前景 \mathcal{F} に属する場合 1 を、背景 \mathcal{B} に属する場合 0 を返す．式 (2) より求まる信頼性 f^{DP}, f^{OF} を用いて、 c_k における推移 $u(c_k)$ を以下の式より選択する．

$$u(c_k) = \begin{cases} u_{DP}(c_k) & \text{if } f^{DP} < f^{OF} \\ u_{OF}(c_k) & \text{otherwise.} \end{cases} \quad (3)$$

選択された推移を用いて前景領域の推定を行う．推移の選択結果および、推定された \tilde{L}^t を図5に示す．

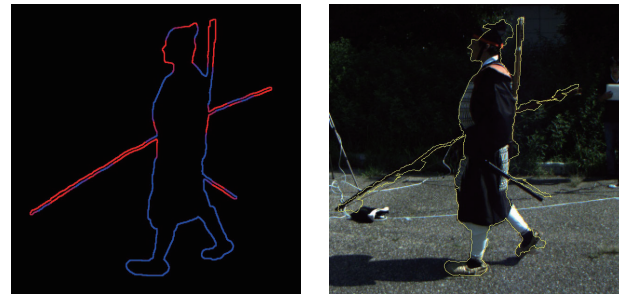


図5 (左) 推定手法の選択結果：赤線はマスク外挿、青線はオプティカルフローによる推定を選択した画素、(右) 推定された前景領域

3.2 領域尤度の算出

本小節ではグラフカットに用いる領域尤度の算出手法について説明する．既に前景領域の推定形状は得られたが、図5を見ると輪郭が凹凸になっていたり、また一部でエラーが生じてしまっている．そこで、色の情報も考慮して領域尤度を求め、次小節で述べるグラフカットを行う事で抽出精度の向上を行う．

領域尤度の算出は Bai ら [14] の手法を用いた．まず輪郭 C^{t-1} 上にお互い重複するように正方形の窓を設定する．本研究では窓の一边を 50 ピクセルとする．次に、窓ごとに混合正規分布による I_{t-1} の色モデルの学習および、前述の推移の代表量の算出を行う．そして得られた推移を用いて窓を I_t に伝搬させ、ある画素 x の色モデルによる領域尤度 $p_c(x)$ を求める．最後に推定された前景領域を領域尤度に加え、求める領域尤度 $p_{\mathcal{F}}(x)$ を

$$p_{\mathcal{F}}(x) = f_s(x)\tilde{L}^t(x) + (1 - f_s(x))p_c(x) \quad (4)$$

より得る．なお、 $f_s(x)$ は x から輪郭までの距離に応じて $\tilde{L}^t(x)$ の減衰する速さを決める変数で、それぞれの窓ごとに求めておく． $f_s(x)$ によって、前景と背景の色分布が類似しているときは $\tilde{L}^t(x)$ の影響が増すように、逆の場合は $p_c(x)$ 、つまり色モデルの影響が増すように、両者を重み付けする事が出来る． $p_{\mathcal{F}}(x)$ が求まったら同様に背景の領域尤度 $p_{\mathcal{B}}(x)$ の計算も行う．

3.3 グラフカットによる前景抽出の高精度化

最終的な前景抽出は、以下の式で定義されるエネルギー関数 $E(X)$ を最小化する事で得られる [10] .

$$E(X) = \lambda \cdot R(X) + B(X) \quad (5)$$

X は全画素のラベル (前景または背景) の集合であり、 $R(X)$ はデータ項、 $B(X)$ は平滑化項、 λ は重み係数である。 $R(X)$ には領域尤度 $p_F(x)$ 、 $p_B(x)$ に反比例するような値を設定する。 $B(X)$ には隣接画素の輝度差に反比例するような値を設定する。 $B(X)$ と適切な λ の設定によって、隣同士の画素のラベルが連続するように拘束する事が出来る。

式 (5) の $E(X)$ をグラフカットによって最小化し、求める前景領域を得る。

4. マスク領域の内挿によるフレーム補間

本章ではマスク領域の内挿によるフレーム補間手法を提案する。フレーム補間処理は次の2つのステップから構成される。

STEP1: フレーム間の動き情報の抽出

STEP2: 中間フレームにおける補間画像の生成

STEP1 では連続するフレーム間において、人物領域がどのように推移したかを推定する。本手法では前処理としてマスク間のマッチングを行い、解空間を限定したオプティカルフローの推定を行った。これにより、特に輪郭付近の推定を安定して行う事が出来る。初期入力としてマスク情報を与える必要があるが、これは3章で提案した抽出システムにより効率的に求める事が出来る。STEP2 では抽出された動き情報を用いて中間フレームにおける人物領域の位置を補間、描画を行う。補間画像の描画手法は様々あるが、本手法ではMRシステム実行時にかかる負荷を最低限に抑えるため、パッチワーピングの手法を用いた。

4.1 フレーム間の動き情報の抽出

4.1.1 マスク間のマッチング

マスク間のマッチングは3章同様、輪郭のDPマッチングを用いて行う [19], [21]。輪郭付近の見えやオクルージョンに依らない推移の推定が可能である。ただしDPマッチングは輪郭同士で行われるため、人物領域内の全ての画素についてもフローを求める必要がある。輪郭上の画素を $c_k (k = 0, 1, \dots, N_c - 1)$ 、DPマッチングより得られた c_k の推移を $u_{DP}(c_k)$ とし、求める人物領域内のある画素 x のフローを $u_{INI}(x)$ とすると、

$$u_{INI}(x) = \frac{\sum_k u_{DP}(c_k) (|x - c_k| + \epsilon)^{-2}}{\sum_k (|x - c_k| + \epsilon)^{-2}} \quad (6)$$

x から各輪郭画素までの距離の2乗に反比例した重みで推移を足し合わせている。 ϵ は近傍の輪郭上のフローの重み付けに影響する指標であり、本論文では $\epsilon = 0.5$ とした。求めた $u_{INI}(x)$ のマップを図6(中央)に示す。

4.1.2 解空間を限定したオプティカルフロー

得られた $u_{INI}(x)$ のマップは、人物領域の内側において正確な推定が出来ていない。そこでオプティカルフローによる精度の向上を行う。本論文ではSandら [22] の手法を用いた。精度が良い分処理コストが大きいが、3章とは違ってシステムのレスポンスに注意する必要はない。各画素のフローの初期値に $u_{INI}(x)$ をセットし、輪郭画素のフローを $u_{DP}(c_k)$ に固定して、オプティカルフローの反復計算を行った。輪郭画素のみフローを固定するのは、人物領域外にマッチングを取らないようにするためである。また領域外とのフロー平滑化も行わないように平滑化項を設定する。得られたフロー $u(x)$ のマップを図6(右)に示す。

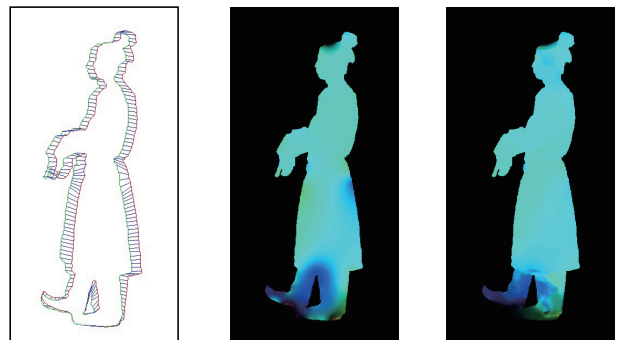


図6 (左) 輪郭 DP マッチング結果, (中央) $u_{INI}(x)$ のマップ, (右) 得られた動き情報 $u(x)$ のマップ

4.2 中間フレームにおける補間画像の生成

補間画像の生成には、描画時のコストが小さいパッチワーピングを用いた。まずフレーム上にパッチを配置し、次に動き情報から中間フレームにおけるパッチ位置を線形補間により求める。最後にパッチのテクスチャを変形して描画する。より高速な描画の為に、テクスチャが疎な領域ではパッチの配置密度も小さくした。

5. 実験

5.1 マスク外挿による前景領域抽出

複数の人物映像に対して提案手法を適用し、前景領域の抽出を行った。結果を図7に示す。

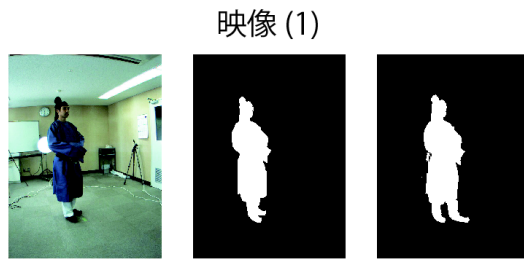
提案手法の有効性を確かめるため、3つの手法に分けて比較実験を行った。

- マスクの外挿
- オプティカルフロー (Baiら [14])
- 両手法の適応的選択 (提案手法)

評価の指標として、以下の式によって定義されるエラー率を用いた。

$$\text{抽出エラー率} = \frac{\text{誤抽出された画素数}}{\text{正しいマスク領域の画素数}} \quad (7)$$

なお、正しいマスク領域は手作業により求めたマスクを



	映像 (1)	映像 (2)	映像 (3)
マスクの外挿	5.61 %	21.1 %	6.04 %
Bai ら [14]	1.23 %	16.9 %	3.26 %
提案手法	1.20 %	16.3 %	3.26 %

表 1 各シーケンスでの抽出エラー率

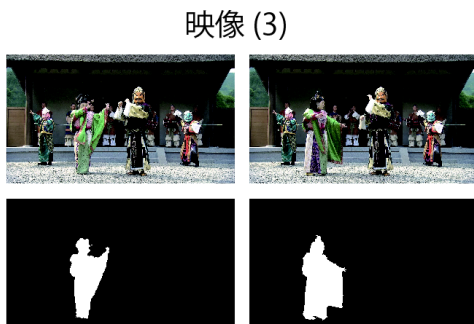
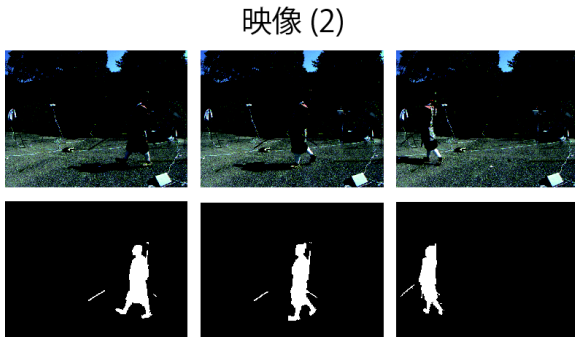


図 7 前景領域の抽出結果

用いる。各映像の抽出エラー率の平均値を表 1 に示す。また、映像 (2) の全フレームのエラー率の推移を図 8 に示す。

全ての映像で提案手法が最も低い平均エラー率を示した。また映像 (2) では、図 8 の示す通りフレームによって抽出率が大きく変動するものの、提案手法は安定して低いエラー率を示した。

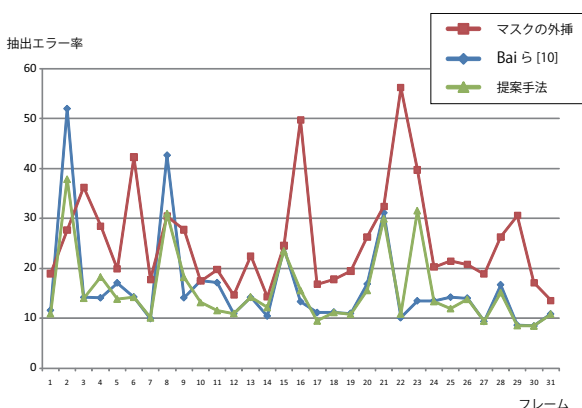


図 8 映像 (2) のエラー率のフレーム推移

提案手法はマスク外挿とオプティカルフロー [14] 双方の長所を活かし、最も安定した抽出を実現した。映像 (1)

では足に着物の影が落ちる事があり、明度が変化する領域でオプティカルフローによる推定精度が悪かった。一方、見えの変化に影響されないマスク外挿では抽出精度が向上し、提案手法もこれを反映して最も良いエラー率を示した。映像 (2) は色分布の類似した頭部、推移の大きい槍などにおいてマスク外挿の推定精度が良く、胴から脚部ではオプティカルフローの推定精度が良かった。提案手法では信頼性に基づく選択が効果的に機能している。映像 (3) ではフレームレートが高いため人物の推移が小さく、オプティカルフローの推定精度が安定していた。提案手法においてもオプティカルフローを選択する領域が多く、エラー率は同じであった。

ただし、本手法で提案する信頼性の定義は色情報のみに依存しているという問題もあるため、この点は改善を行っていく必要がある。

5.2 マスク内挿によるフレーム補間

提案手法を 3 つの人物映像に適用した結果を図 9 に示す。なお、映像 1, 2, 3 のフレームレートは 10, 10, 30 であり、図 9 の示す補間画像はフレームの中間時刻におけるものである。

実験より、本手法の有効性を確認した。映像 (1) では役者が持つ細長い槍の領域、また色分布が類似した頭部において提案手法の有効性が示された。補間画像の拡大を図 10 に示す。Sand ら [22] の手法では槍の先端や頭部において推定エラーが多く、見えに悪影響を与えているのがわかる。映像 (2) では映像のフレームレートを 30fps から 10fps に落として実験を行った。フレームの中抜きをしたものの、前処理の輪郭マッチングがうまく働いたため全体的に良い見栄えとなっている。テキストに比べて動き情報は 0.1% の情報量で済むため、MR 実行時に消費するメモリ量を大幅に削減出来る事が確認された。より高い見栄えを維持するためには、動き情報の抽出はフレームの中抜きをせずに行い、MR 実行時に保持するテキストはキーフレームに限る手法が考えられる。

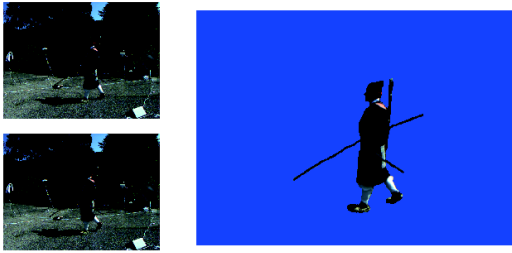
映像 (3) は (2) と同じシーケンスの別フレームだが、人物の領域内でオクルージョンが発生しており、輪郭マッチングによる初期解が一部に悪影響を与えた。第 2 節の手法をマルチレベルに拡張し、複数のマスク領域を用いて前処理を行う手法などが考えられる。

5.3 MR イベントへの適用

5.3.1 MR システムの概要

提案手法を適用し、MR 体験の公開イベントを行った。

映像(1)



映像(2)



映像(3)

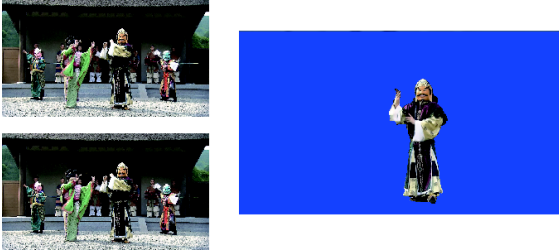


図 9 フレーム補間結果

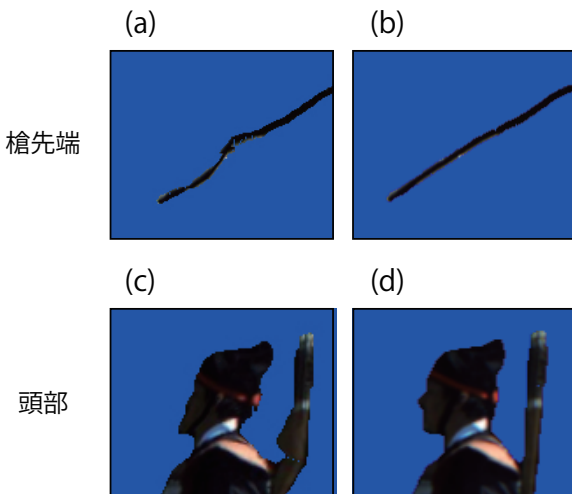


図 10 映像(2)補間結果, 拡大図((a)(c):Sandら[22], (b)(d):提案手法)

8世紀頃に日本の首都であった平城宮の風景を様々な視点から鑑賞する事が出来る。イベントでは、ユーザは会場内を周遊するトラムカーに乗りながら、HMDを通して3DCG復元された平城宮の風景を体験してもらった。なお、建物コンテンツの3Dデータは凸版印刷(株)制作

である。

5.3.2 ビルボードによる表示とテクスチャの切り替え

人物映像より作成するMRコンテンツは2Dモデルであるため、MR実行時にはビルボード表示を行う。しかし本MRシステムは空間内の移動が可能であるため、人物ビルボードの裏側に回り込んだ際に大きな違和感が生じてしまう。

そこで、予め役者を全方向から撮影しておき、MR実行時にはユーザの視点に応じて適宜テクスチャを切り替えて表示した。

5.3.3 シナリオおよび人物コンテンツの作成

当時の役職や儀式の様子を人物コンテンツを通して再現し、物語性のあるMR空間の演出を行った。平城京はかつて朝賀の儀式など国家的儀式を執り行う場所であった事などを参考にし、貴族の参列する儀式の様子、働く役人、警備をする衛士の様子などをシナリオとして設定した。

シナリオに従って当時の役人、衛士、貴族の映像素材の取得を行った。役人同士が話し合う様子や較べ馬を観戦する演技の指導を行い、また、前に述べたとおり役者の撮影は全方向から行われた。次に取得した人物映像に対して本手法を適用し、MRコンテンツとした。

人物コンテンツの作成では、膨大な人物映像のマスク処理を第3章で提案した抽出システムを用いて効率的に行う事が出来た。フレームあたりの処理時間は5~15秒ほど、エラーが生じた場合の補正作業も同程度であった。手作業で行った場合は毎フレーム1分以上かかるため、大幅に作業時間を短縮する事が出来たと言える。

5.3.4 一般公開イベント

作成したMRコンテンツを一般公開イベントにおいて体験者に提示した。体験者はCG復元された平城宮と当時の人々の様子を体験する事が出来る。先述した各シナリオの実機画像を図11に示す。

人物コンテンツがMR体験に与える影響を評価するため、体験者(10代以下~90代の1314名)にアンケート調査を行ったところ、8割を超える体験者が人物コンテンツの重畳を好意的に評価した。本研究がMR体験の価値向上に貢献している事が示された。

一方で“人物が少ない”、“表情などを(もっと近くで)見たい”、“話し声が聞こえると良い”と言った意見も聞かれた。人物が少ないのはMRシーンが非常に広大であった事もあるが、ビデオメモリの容量の制約から用意した人物コンテンツを全て重畳出来なかった事も原因の一つである。今後は第4章で提案したフレーム補間手法を用いてデータ量の圧縮を行い、より多くの人物コンテンツが登場出来る工夫を行う。表情や話し声の追加は今後の課題である。

6. まとめ

本論文では、人物映像を用いて物語性のあるMR空間

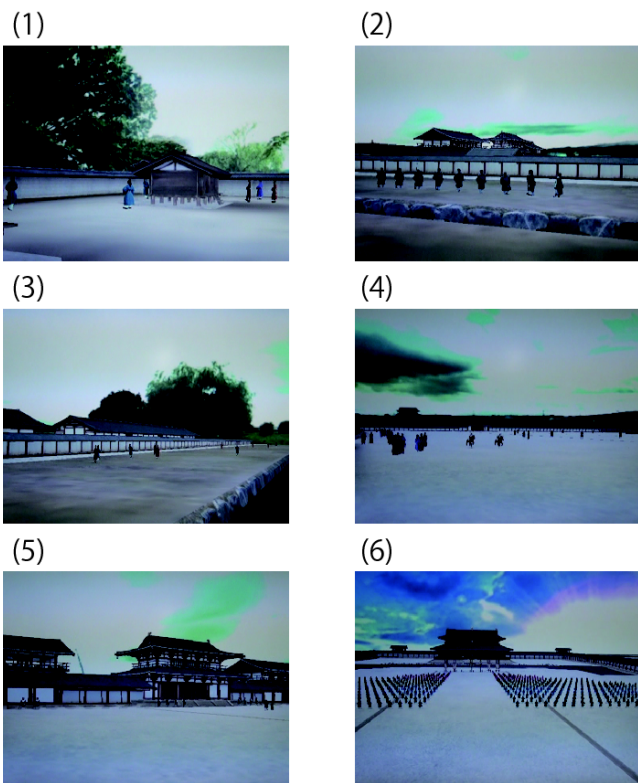


図 11 平城京 MR の実機画像 ((1) 官庁街, (2) 警備する衛士, (3) 行き交う役人, (4) 較べ馬の観戦, (5) 朝堂院に集まる役人, (6) 朝賀の儀式 VR)

を演出することを目的とし, その MR コンテンツ化の支援技術として前景領域の抽出及びフレーム補間手法の提案を行った. 具体的には連続するフレーム間でマスク領域のマッチングを行い, これの外挿により次フレームにおける前景抽出, 内挿により中間フレームにおける見えの補間を行った. また, 作成した人物モデルを実際に MR イベントに適用し, 本手法の有効性を示した. 今後は音声情報などの追加により, 人物映像の MR コンテンツ力を高めたい.

文 献

[1] R. Azuma, "A survey of augmented reality," *Teleoperators and Virtual Environments*, vol. 6, no. 4, pp.355-385, Aug. 1997.

[2] R. Azuma, Y. Baillet, R. Behringer, S. Feiner, S. Julier and B. MacIntyre, "Recent advance in Augmented Reality," *IEEE Computer Graphics and Applications*, vol. 21, no. 6, pp. 34-47, Nov. 2001.

[3] Y. Uematsu and H. Saito, "Multiple Planes Based Registration Using 3D Projective Space for Augmented Reality," *Journal of Image and Vision Computing*, Vol.27, pp.1484-1496, 2009.

[4] 大槻麻衣, 杉原賢次, 中嶋友美, 木村朝子, 柴田史久, 田村秀行, "絵筆の描き味を活かした複合現実型描画システムと筆型対話デバイス," *日本バーチャルリアリティ学会論文誌*, Vol.15, no.3, pp.357-367, Sep.2010.

[5] 南谷真哉, 北原格, 亀田能成, 大田友一, "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," *電子情報通信学会技術研究報告*, pp.91-96, Sep.2007.

[6] 角田哲也, 大石岳史, 池内克史, "高速陰影表現手法を用

いた飛鳥京 MR コンテンツの開発とその評価," *映像情報メディア学会誌*, vol. 62, No. 9, pp. 1466-1473, 2008.

[7] T. Taketomi, T. Sato and N. Yokoya, "AR cultural heritage reconstruction based on feature landmark database constructed by using omnidirectional range sensor," *ACCV 2010 Workshop on e-Heritage*, Nov. 2010.

[8] 大石岳史, 大藏苑子, 川上玲, 阪野貴彦, 池内克史, "全方位カメラを用いた光源環境と対象物の同時撮影法に基づく人物モデルの MR システムへの重畳," *画像の認識・理解シンポジウム*, Aug. 2009.

[9] 大石岳史他"車両を利用したマルチユーザ体験型屋外 MR システム~バーチャル平城宮復元プロジェクト~" *情報処理学会*, vol.175, No.52, pp.1-8, 2011.

[10] Y. Boykov and M. Jolley, "Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, no. 9, pp.1124-1137, 2004.

[11] Y. Boykov and G.Funka-Lea, "Graph cuts and efficient n-d image segmentation," *International Journal of Computer Vision*, Vol.70, no.2, pp.109-131, 2006.

[12] J. Wang, P. Bhat, A. Colburn, M. Agrawala, M. Cohen, "Interactive Video Cutout," *ACM SIGGRAPH 2005 papers*, pp.585-594, 2005.

[13] Y. Li, J. Sun and H.Y.Shum, "Video object cut and paste," *ACM Transactcion on Graphics*, Vol.24, no.3, 2005.

[14] X. Bai, J. Wang, D. Simons and G. Sapiro, "Video Snapcut: robust video obouect cutout using localized classifiers," *ACM SIGGRAPH 2009 papers*, page.1-11, 2009.

[15] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial intelligence*, pp.185-203, 1981.

[16] V.Kolmogorov and R.Zabih, "Computing visual correspondence with occlusions using graph cuts," *IEEE Computer Society*, 2001.

[17] J. Xiao, H. Cheng, H.Sawhney, C. Rao and M. Isnardi, "Bilateral filtering-based optical flow estimation with occlusion detection," *Computer Vision-ECCV*, pp211-224, 2006.

[18] T. Brox and J.Malik, "Large displacement optical flow: descriptor matching in variational motion estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010.

[19] H. Fujiyoshi and A.J.Lipton, "Real-time human motion analysis by image skeletonization," *Applications of Computer Vision*, Fourth IEEE Workshop, pp.15-21, 2002.

[20] A. Wedel, T.Pock, C.Zac, H.Bischof, and D. Cremers, "An improved algorithm for TV-L1 optical flow," *Statistical and Geometrical Approaches to Visual Motion Analysis*. pp.23-45, 2009.

[21] 高畠綾子, 野村太一, 藤吉弘亘, "輪郭形状の対応計算によるシェイプトゥーン," *画像の認識・理解シンポジウム*.2005.

[22] P. Sand and S. Teller, "Particle Video: Long-range motion estimation using point trajectories," *International journal of computer vision*, vol.60,no.2, pp.91-110, 2004.