

## ワープ変形を利用した 動画間の照明転送手法

中川西宏友<sup>†</sup> 岡部誠<sup>‡\*</sup> 尾内理紀夫<sup>†‡</sup>

2つの異なる照明条件の人物動画が与えられた際、一方の動画から他方の動画へ照明条件を転送する手法を提案する。各動画では、異なる俳優が異なる照明条件のもとで演技していることを仮定している。本手法ではまず、転送元の動画にワッピングを施し、体の輪郭を転送先の動画に合わせる。次に各フレーム間で低周波成分を転送することにより照明転送を実現する。これは人間の目は低周波な照明変化にのみ敏感である事実を利用している。本手法により、一般的なビデオカメラで撮影された動画間で照明転送が可能であることを示した。

### Illumination Transfer between Videos using Image Warping

Hiroto Nakagawasai<sup>†</sup> Makoto Okabe<sup>‡\*</sup> and Rikio Onai<sup>†‡</sup>

We propose a method for illumination transfer between two videos. We assume a different actor plays in a different illumination condition in each video. First, we warp each frame of the source video so that the contour of the actor's body fits to the contour in the target video. We then synthesize the resulting frame by transferring the low frequency component from the source to the target: this is based on the fact that human visual system is sensitive not high frequency but only low frequency of illumination change. We demonstrate illumination transfer between videos captured using a casual video camera.

### 1. はじめに

近年、テレビ番組や映画の制作で映像合成が一般的になってきているが、リアルな合成映像を作るためには、素材となる画像間における照明条件の不整合を取り除く事が必要不可欠である。そのための手法として、プロジェクタなどによって目的の照明条件となる環境光を再現してから被写体を撮影する手法[1][2][3]や、被写体の3次元形状などを計測しコンピューター・グラフィックスとして再現したものを目的の照明条件でレンダリングする手法[4][5][6]、既に撮影された被写体の画像を何らかの方法で変換するポスト・プロダクション的な手法[7][8][9]などが挙げられる。

ポスト・プロダクション的な方法の1つとして、異なる画像間で照明条件を転送させる技術（イルミネーション・トランスファー、照明転送）がある。この技術は静止画を対象としたものは頻りに研究されている[7][8][9]。Liら[7]は異なる人物の顔をそれぞれ別の照明条件下で撮影した画像を利用した照明転送の手法を提案している。この手法は、まず目や鼻といった人間の顔特有の特徴をキーとして利用して転送元の画像から転送先の画像へワープ変形を行った後、それぞれの顔画像の照明に依存する成分・依存しない成分へ分解し、最後に転送元画像の照明依存成分と転送先画像の照明に依存しない成分を合成するというものである。この方法は2つの入力画像のみから照明の転送を行えるという点でとても優れているが、顔画像に特化した手法であるため、対象範囲が狭く、一般的な画像にも利用できる手法であるとは言えない。

動画に対する照明転送の手法の1つがPeersら[9]が提案した「Light Stage」と呼ばれる装置を利用した手法である。Light Stageでは、球面上に設置されたLEDライトを様々な組み合わせで高速に照射することによって、その内部の被写体上にあらゆる照明条件を再現することが可能である。Peersらの手法では、照明条件を変更したいターゲット画像に対し、Light Stageによって撮影された画像郡の中から「ターゲットと同様の照明条件の画像」と「転送したい照明条件の画像」を選び、画素ごとの画素値の比率を表す中間画像を生成し、その中間画像によってターゲットの画像を変換する。この手法は2次元的な画像処理技術のみで実現されているが、まるで3次元形状を利用しているかのようにあらゆる角度からの光のリアルな再現に成功している。しかし一方で、この手法に必要なLight Stageを構築するためには金銭的・時間的にも高いコストが必要である。またLight Stageを使って撮影可能な空間は装置中央の範囲のみであるため、全身を撮影するのであればそれだけ装置全体は巨大化し、歩く・走るなど

<sup>†</sup> 電気通信大学大学院情報理工学研究科  
Graduate School of Informatics and Engineering, The University of Electro-Communications

<sup>‡</sup> 電気通信大学工学情報理工学部  
Faculty of Informatics and Engineering, The University of Electro-Communications

\* 独立行政法人科学技術振興機構さきがけ  
JST PRESTO

の動きを伴う場合はそれに応じて動く床などを設置しなければならず、装置の構築コストはさらに高くなる。

そこで我々は、家庭でも使われる一般的なデジタルビデオカメラで撮影された動画のみを利用した照明転送の手法を提案する。提案手法では、Liら[7]の手法におけるワーピングを顔でなく体全体の輪郭に対して行うことで、対象を顔画像から体全体まで拡張している。画像において必要な入力は、2人の人物を異なる照明条件下で撮影した2種類の動画のみであり、従来手法よりも低コストで照明転送の実現が可能である。

まず入力として、異なる俳優を異なる照明条件のもとで撮影した2種類の動画を用意する。転送先動画の各フレームに対し、もっとも姿勢が近いと思われるフレームを転送元動画から選ぶ。次に選ばれたフレームをワープ変形により俳優の体輪郭を転送先フレームと一致させる。そして変形したフレームから低周波成分を取り出し、転送先の高周波成分と合成する。人間の目は低周波な照明変化に敏感なため、こうすることで照明条件が転送されたように見える画像を得ることができる。

本論文では、2章で提案手法全体の概要について述べ、3章で具体的な処理の内容を述べる。そして4章で提案手法によって得られた結果を示し、5章でまとめと今後の課題について述べる。

## 2. 提案手法の概要

以下に、提案手法の全体構成図を図1によって示しながら、おおまかな処理の流れを解説する。

まず入力として照明を変更したい動画(A)と、目的の照明条件を満たした別人の動画(B)を用意する。動画を与えられたシステムは、ターゲットのあるフレーム(a)に対し、ワーピングによって変形しやすいソースのフレーム(b)を選び、ワーピングを行うことで被写体の輪郭が(a)と一致したフレーム(b')を得る。

次に、(a)と(b')の明度成分を高周波成分・低周波成分へ分解する。この時、高周波には照明の条件によらない画像のエッジなどの成分が、低周波には光の陰影などによって起こる緩やかな明度変化などの成分にあらわれる。

ここで、出力したいフレーム(c)を考える。もしも(c)の高周波成分と低周波成分がそれぞれ分かれば、分解の逆変換によって(c)を合成することが可能である。

(a)と(c)に写っている人物は同一人物なので、(c)の高周波成分は(a)の高周波成分( $a_{high}$ )とほぼ一致する。また、(b')はワーピングによって(a)および(c)の輪郭に一致しているため、(c)の低周波成分は(b')の低周波成分( $b'_{low}$ )とほぼ近似できる。したがって、これら( $a_{high}$ )と( $b'_{low}$ )を合成することにより得られるフレームは(c)に近似できる。以上の処理をターゲットの全フレームに対してかけることにより、照明転送が施された動

画(C)を得ることが可能である。

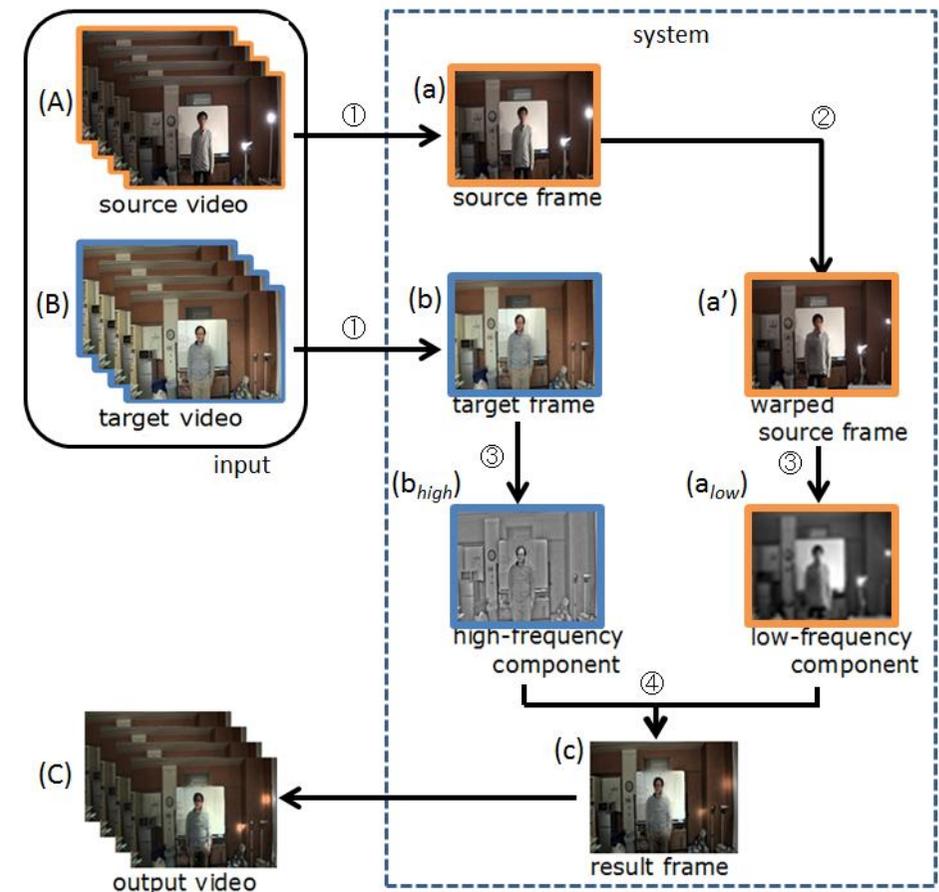


図1 全体構成図 - ①転送先の各フレームに対して似た姿勢のフレームを転送元より選び、②転送元フレームを転送先フレームの体輪郭に合わせ変形し、③転送元の低周波成分と転送先の高周波成分を取り出し、④最後にそれらを合成する

### 3. 各モジュールの機能

#### 3.1 前処理

##### ・ 人物領域の抽出と、輪郭の類似度比較

ソースのフレームがターゲットに似た姿勢をしていれば、特徴点も一致しやすくワーピングの精度も高まると考えられる。提案手法では、人物領域の輪郭を類似度比較することで似た姿勢の判別を行い、ワーピングに適した最も姿勢の近いフレームを取得する。

図2は、人物領域の特定と体輪郭の特徴量を求める様子を表している。まず、各フレームから人物が写っている領域を求める。提案手法では、単純な背景差分によって得られた複数の領域にラベリング処理によってラベルを付与し、面積の最も大きな領域のラベルをもつ画素を人物領域としている(図2で水色に着色してある領域。その他の白い領域は背景の微細な変化によるノイズである)。次に人物領域の重心から輪郭線上の点までの長さを16方向分について測定する(図2における赤線の長さ)。この長さを成分とした16次元のベクトルを特徴ベクトルとし、ベクトルのユークリッド距離がもっとも小さくなる画像であればあるほど類似度が近いものとした。この類似度を利用することで、1つのターゲットフレームに対し、ソース動画全体の中からもっとも変形に適していると思われるフレームを検出することが可能である。

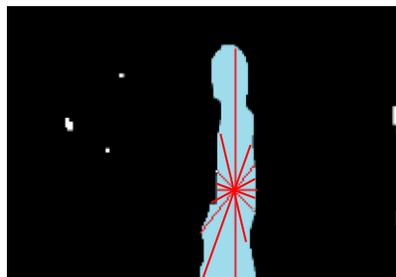


図2 人物領域の決定と輪郭の特徴量 - 最も大きい領域(水色部分)の重心から輪郭までの長さ(赤線)を成分としたベクトルを特徴量とする。

#### 3.2 2段階のワーピング

提案手法ではSIFT-Flow[10]を利用したワープ変形を行なっている。しかし、仮に似た姿勢のフレームを選んでワープ変形をしていたとしても、動画の全フレームに対し被写体の輪郭が崩れないように変形させるのは難しい課題である。特に光源の向きが大きく変化する場合などは、特徴点のマッチングが失敗しやすくなってしまい、それ

によって意図しない変形によって輪郭が崩れてしまうという不具合が起きやすくなってしまふ。図3におけるフレームAからフレームWの変形が特徴点のマッチングに失敗した例である。マッチングがうまくいっていない部分は、崩れた部分(collapsed area)として出力結果に現れているのが確認できる。

そこで、あらかじめ各フレームを手作業でワーピングさせ、ある程度輪郭が一致した画像を用意しておき、その画像に対してSIFT-Flowにかけるという2段階のワーピングを行う。手作業でのワーピングではMizuiら[11]の手法を利用しfeature-based image warping[12]によって変形を行う。このようにし先に輪郭を合わせておくと、SIFT-Flowによってピクセルが大きく移動させる必要がなくなる。なので、必要以上の変形を行わないようにSIFT-Flowのパラメータを調整することができるようになる。これにより、直接変形すると変形に失敗してしまうフレームでも俳優の体輪郭が大きく崩れないように変形することができる。図3のフレームAは直接変形してもうまく変形出来なかったが、手動で変形させたフレームA'を経由することによって得られたフレームW'には画像の崩れが無くなっている。

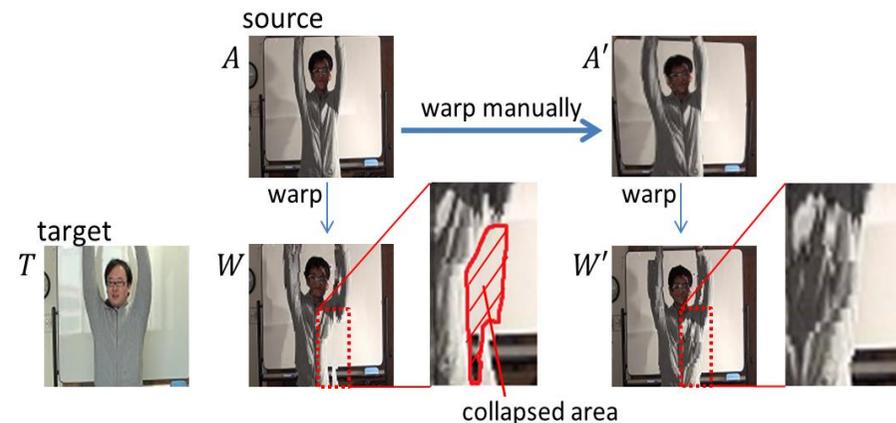


図3 2段階のワーピング - source(A)からtarget(T)へ直接SIFT-Flowにかけて得た画像(W)には画像の崩れ(collapsed area)が現れているが、手動によって変形したもの(A')にSIFT-Flowをかけると(W'), その崩れが改善されている

### 3.3 照明転送

#### ・ ラプラシアンピラミッド

周波数の分解にはラプラシアンピラミッド[13]を利用する。ラプラシアンピラミッドによる変換は可逆変換であるので、分解された各周波数成分は逆変換によって元の画像に戻すことができる。

提案手法では、ターゲットとソースの各フレームをラプラシアンピラミッドで低周波成分と高周波成分に分解した後、ソースの低周波成分とターゲットの高周波成分を使って逆変換していくことで照明条件の異なる新しいターゲットのフレームを合成できる。

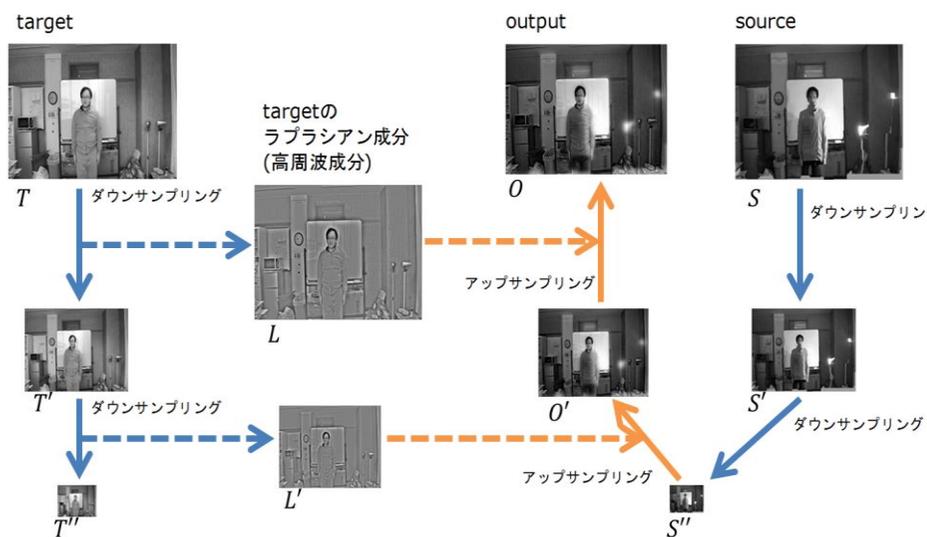


図 4 ラプラシアンピラミッドを用いた照明転送 - target, source それぞれの画像(T,S)からダウンスAMPLING( $T', T'', S', S''$ )し、その際に失われた高周波成分によりtargetのラプラシアンピラミッド(L, L')を形成する。そしてsourceの低周波成分( $S''$ )にtargetの高周波成分を加えてアップサMPLING( $O', O$ )する事によって、照明転送された画像 output を得る。

#### ・ 時間軸方向の平滑化

実際に照明条件の転送を行なってみると、動画がちらついて見えることがある。このちらつきは明度成分が短時間で不自然に増減することにより起こる。

そこで、低周波成分に時間軸方向の平滑化をすることで、低周波成分における明度変化を穏やかにする。提案手法では $t$ 軸方向に対する1次元メディアンフィルタをかけることでこの平滑化を行なっている。これによって、問題となっていた動画のちらつきの大部分を取り除くことができる。

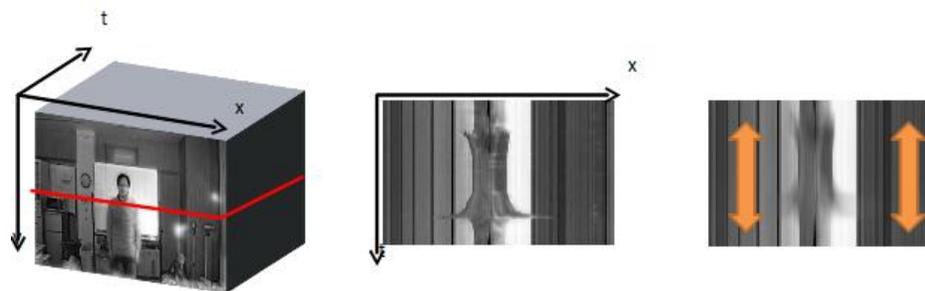


図 5 時間軸平滑化 - 動画を縦・横・時間の3軸で表し(左側)、時間軸に平行な面で切り取ったもの(中央)を時間軸方面に平滑化する(右側)

### 3.4 後処理部

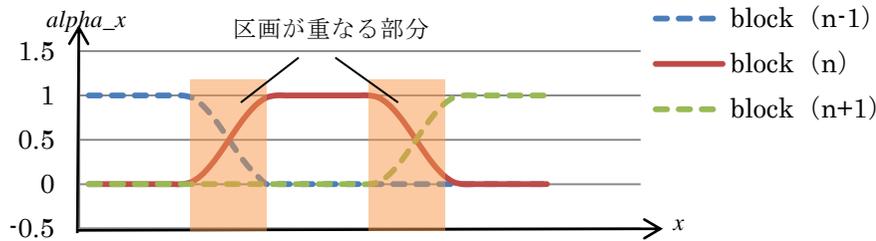
#### ・ 区画ごとのヒストグラムマッチング

ここまでで、ターゲット動画の低周波成分を別動画のものと差し替えることによりある程度照明条件を変化させることに成功しているが、出力結果にはまだターゲット元来の照明条件の情報が取り除き切れておらず、全体的に少しぼけたような印象を受けてしまう。そこで、コントラストや色調などをよりソースの画像の印象に近づけるためヒストグラムマッチングによる見た目の補正を行う。

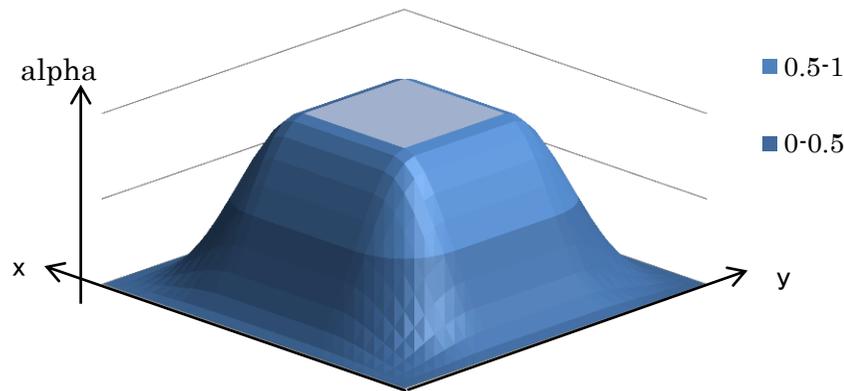
提案手法では、マッチング後の画像にも元画像におけるヒストグラムの局所的な特徴を反映させるため、Okabeら[14]の手法を参考にして画像空間を一樣な矩形の区画に分解し、区画ごとのマッチングを行う。区画の境界線を目立たなくするため、あえて隅が重なるように区画をとり、重なった部分はアルファ・ブレンディングによる合成を行う

まず $x$ 軸上の座標に対し、以下のグラフ1に示すような $alpha_x$ という値を考える。このグラフでは区画が重なる部分は正弦曲線によって、それ以外の部分は直線によって3区画分の $alpha_x$ が描画されていて、任意の $x$ 座標に関してそれらの3つの $alpha_x$

を合計した総和は必ず 1 となっている。同様に、 $y$  軸に対しても  $y$  座標に対応する  $alpha_y$  を求める。 $alpha_x$  と  $alpha_y$  が求められたら両者をかけ合わせてアルファ値  $alpha$  とする。このようにして求められた  $alpha$  もまた、任意の座標について隣接する区画すべての  $alpha$  との総和が必ず 1 になっている。



グラフ 1 x 軸と  $alpha_x$  の関係グラフ。曲線の部分はいずれも正弦曲線になっている



グラフ 2 x 軸・y 軸とある区画における  $alpha$  値の関係を示すグラフ

#### 4. 実験結果

提案手法によって照明転送を施した結果を図 6 および図 7 に示す。どちらもの図に関しても、左列の **target** に対し、中央の **source** の照明条件を転送した物が右列の **result** となっている。

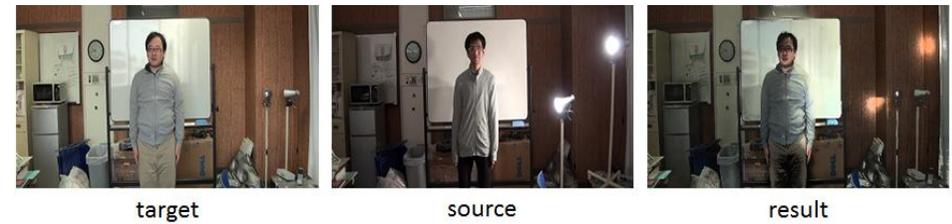


図 6 照明条件の転送 1

図 6 は、同じ室内で撮影された、俳優の動きの小さな動画間における照明転送の結果である。全体的に明るい室内で撮影された **target** に対し、**source** の暗い部屋における照明条件が転送され、**result** は暗い室内で撮影されたような画像となっている。**target** と **source** の画像における被写体の俳優はそれぞれ異なる体型をしているが、ワーピングによって体輪郭に合わせているため、**result** の画像でも被写体の体輪郭がと明瞭に確認できる他、服のしわなどの高周波成分も **target** の画像の物が **result** に反映されているのが確認できる。また、背景を見てみると、**source** 内に写り込んだ照明の残像が **result** にも現れてしまっているが、ホワイトボードを始めとしたその他に写り込んだものに関しては、**source** の照明条件が **result** にも反映されているのが確認できる。変化の少ない背景の画像を用いれば、本手法のワーピングによってある程度対応し、背景に関しても俳優と同様の照明条件を再現することが可能である。

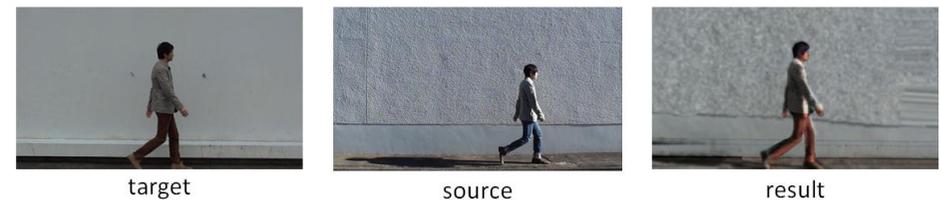


図 7 照明条件の転送 2

図 7 は天候の異なる屋外において撮影した動画間で照明転送を行った結果である。**target** および **source** において、どちらの俳優も左から右へ画面を横断するように大きく移動している。曇りの状態で撮影された **target** に対し、**source** における晴天時の照明条件が転送され、**result** では晴天時に正面から光を受けながら歩いているような画像となっている。**target** と **source** において俳優の位置が異なるが、これは 3.1 にて述べた人物領域の輪郭の類似度によってフレームを選んだ結果である。このように、連続し

た多くのフレームの中から最もワーピングに適した画像を選択できるのは、動画を対象とした本手法ならではの利点でもある。また、例6と同様に地面の影などの背景がこの例でも再現されている。しかし target にはなかった壁の様子が現れている上に、その壁の様様に俳優をワーピングした時の歪みが現れてしまっている。この例のように背景の差分が大きすぎる場合、周波数分解による除去できるノイズにも限界があるため、フィルタの改善など別の工夫によってノイズを取り除く工夫が必要である。

どちらの例にも共通して言える問題点として、エッジ周辺を白く縁取りしたようなアーティファクトが現れるという問題がある。これは解像度が高くなればなるほど顕著になる。この現象はハロー効果と呼ばれるもので、デジタル画像の高解像度化・高精細化でしばしば問題になる現象である。ハロー効果を軽減するには、ダウンサンプリング時にエッジ保存型のローパスフィルタを利用する方法[15]が知られている。



図 8 ハロー効果

## 5. おわりに

本論文では、ワーブ変形を利用した照明転送の手法を提案し、動画を用いた実験と検証を行った。実験の結果、提案手法を利用すれば少ないコストで人間の体全体に対する照明が転送できることを確認した。また、背景に対しても同様の方法である程度の照明転送が可能であることを確認した。

しかし、背景が大きく異なる時には必要以上に背景の照明を転送してしまうことや、ハロー効果などの問題点も残った。これらは高解像度画像に本手法を応用する際に避けては通れない問題である。これらの問題解決のため、今後は適切にノイズを除去できるフィルタを開発する必要がある。また、低コストで実現できるという提案手法の利点から素人の動画制作における利用場面を想定した場合、ワーピングやフィルタリングなどの調節をユーザーと対話的に行えるようなインターフェースがあれば、ユーザーのイメージした動画により近いものを出力できる有用な編集ツールとして利用できると考えられる。

## 参考文献

- 1) P. Debevec, A. Wender, C. Tchou, A. Gardner, J. Waese, T. Hawlins. "A Lightning Reproduction Approach to Live-Action Compositing", SIGGRAPH 2002 Proceedings, 21, 3, pp.547-556, 2002
- 2) 三ッ峰 秀樹, 深谷 崇史, 山内 結子. "全方位証明による映像合成システム", 映像情報メディア学会誌, 映像情報メディア 59(7), pp.1059-1066, 2005
- 3) M. Okabe, K. Takayama, T. Ijiri, T. Igarashi. "Light Shower: A Poorman's Light Stage Built with an Off-the shelf Umbrella and Projector", SIGGRAPH 2007 Sketches, 2007
- 4) E. Sali, S. Ullman. "Recognizing novel 3-D objects under ner illumination and viewing position using a small number of example views or even a single view", Sixth International Conference on Computer Vision, pp.153-161, 1998
- 5) A. S. Georghiadis, P. N. Belhumeur, D. J. Kriegman. "Illumination-Based Image Synthesis: Creating Novel Images of Human Faces Under Differing Pose and Lighting", IEEE Workshop on Multi-View Modeling and Analysis of Visual Scenes Proceedings, pp.47-54, 1999
- 6) Z. Wen, Z. Liu, T. S. Huang. "Face Relighting with Radiance Environment Maps", IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2003 Proceedings, pp.158-165, 2003
- 7) Q. Li, W. Yin, Z. Deng. "Image-based face illumination transferring using logarithmic total variation models", The Visual Computer, Vol.26, No.1, pp.41-49, 2010
- 8) H. Han, S. Shan, X. Chen, W. Gao. "Illumination Transfer Using Homomorphic Wavelet Filtering and Its Application to Light-Insensitive Face Recognition", IEEE International Conference on Automatic Face & Gesture Recognition, 2008
- 9) P. Peers, N. Tamura, W. Matusik, P. Debevec. "Post-production Facial Performance Relighting using Reflectance Transfer", SIGGRAPH 2007 Proceedings, 52, 2007
- 10) C. Liu, J. Yuen, A. Torralba. "SIFT Flow: Dense Correspondence across Scenes and its Applications", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.33, No.5, 2011
- 11) K. Mizui, M. Okabe, R. Onai. "Choreographing Amateur Performers using Video Examples", SIGGRAPH Asia 2011 Posters, 2011
- 12) T. BEIER, S. NEELY. "Feature-based image metamorphosis", SIGGRAPH 1992 Proceedings, pp.35-42, 1992
- 13) P. Burt, E. Adelson. "The Laplacian Pyramid as a Compact Image Code", IEEE Transactions on Communication, Vol.31, Issue 4, pp.532-540, 1983
- 14) M. Okabe, K. Anjyo, R. Onai. "Creating Fluid Animation from a Single Image using Video Database", Computer Graphics Forum, Vol.30, Issue 7, pp.1973-1982, 2011
- 15) F. Durand, J. Dorsey. "Fast Bilateral Filtering for the Display of High-Dynamic-Range Images", SIGGRAPH 2002 Proceedings, Vol.21, Issue 3, 2002