

# Curvelet 変換を用いた網状物体の除去による隠消現実感

松井 琢朗<sup>\*1</sup> 張 起瑞<sup>\*2</sup> 武富 貴史<sup>\*3</sup>

プロプスキ アレクサンダー<sup>\*4</sup> サンドア クリスチャン<sup>\*5</sup> 加藤 博一<sup>\*6</sup>

**Abstract** – 本報告では、映像中に存在するフェンスのような網状物体を自動で検出し視覚的に除去する隠消現実感手法を提案する。提案手法では、カメラで撮影された動画像のフレームに対し、周波数解析の一種である Curvelet 変換を用いた網状物体の検出及び除去を行う。まず、画像を FFT により周波数領域へと変換し、周波数領域での解析を行うことで周期パターンの検出を行う。その後、検出された周期パターンに対して Hough 変換に基づく網状物体の絞り込みを行う。最後に、検出された網状物体領域に対して、Curvelet 係数の補間処理を行うことで網状物体を除去した画像の生成を行う。実験では、シミュレーション画像を用いた定量評価結果と実画像を用いた結果を示す。

**Keywords** : 隠消現実感, サブバンド分解, Curvelet 変換, 画像修復

## 1 はじめに

隠消現実感 (DR: Diminished Reality) とは、リアルタイムで映像に写り込んだ不要な物体を視覚的に除去する技術である。DR は映像に仮想物体を重畳させる拡張現実感 (AR: Augmented Reality) とは対となる技術として位置づけられている [1]。このような DR 技術は、AR 技術と組み合わせることで空間のシミュレーションなどを行うことができる。具体的には、位置合わせ用の AR マーカを DR 技術で除去した後に仮想物体を表示することで、最終的に生成される映像では AR マーカが存在しない合成結果を得ることができる [2]。さらに医療の分野への応用も検討されている [3]。

現在までに、DR を実現するための様々な手法が提案されている。DR による物体の除去手法は大きく、画像修復技術による推定を行う手法 [4][5][6] と、事前に背景情報を計測しておく手法 [7][8][9] に分けられる。これらの手法に共通している点として、除去対象がマーカ以外の場合には除去対象となる物体を手動で指定する必要がある。これは、一般的な映像の場合には、除去対象となる物体はユーザにより異なるため自動で指定することが困難である。そのため、初期フレームなどにおいてユーザが除去対象を指定するアプローチが採られている。一方で、我々の身の回りの環境には野球の観戦や動物園での動物の鑑賞などのように、安全性の面からネットや金網越しに物体を鑑賞するせざるを得ない場合が存在する。本研究では、このよう

な網状物体越しの鑑賞は、臨場感を低下させる一因になっていると考え、このような網状物体を映像中から自動で除去する手法を提案する。具体的には、撮影された動画像に対して Curvelet 変換 [10] という周波数解析の一種を適用することで、動画像中より網状物体の検出および除去を行う。サブバンド分解を行うことで周波数領域での解析が可能となり、網状物体の周期パターンを自動的に検出することができる。また、検出された領域に対して Curvelet 係数の補間処理を施すことで、網状物体の存在しない映像を生成する。

## 2 関連研究

本節では、これまでに提案されている隠消現実感手法について、実計測に基づく手法と実計測に基づかない手法に大別し概観する。

### 2.1 実計測に基づく手法

実計測に基づく手法では、除去対象となる物体の背景情報を事前またはリアルタイムで計測された情報を用いて除去することで画像の生成を行う。Meerits らは、環境中に設置した RGB-D カメラを用いて背景情報の取得を行い、この背景情報をユーザ視点画像へ投影することで物体を除去する手法を提案している [7]。除去対象物体は AR マーカを用いて検出する方法、物体の色と形を指定することで検出する方法の二種類の方法で実現している。この手法では、物体の背景が動的なシーンにおいても幾何的に整合性のとれた画像を生成することができる。しかし、ユーザ視点カメラとは別に環境中に RGB-D カメラを設置することが必要となり、手軽に隠消現実感画像を生成することは難しい。

Hasegawa らは 1 台のカメラを用いて過去映像を利用した実計測に基づく手法を提案している [8]。この手

\*1 奈良先端科学技術大学院大学

\*2 奈良先端科学技術大学院大学

\*3 奈良先端科学技術大学院大学

\*4 奈良先端科学技術大学院大学

\*5 奈良先端科学技術大学院大学

\*6 奈良先端科学技術大学院大学

法では、除去対象を歩行者としており、映像中の歩行者は HOG 特徴量を用いて検出している。検出された歩行者領域に対して、歩行者が存在しない過去フレームの背景情報を利用することによって歩行者の除去を行っている。1 台のカメラを用いて実背景の計測に基づく隠消現実感映像の生成が可能であるが、画像の合成を Homography 変換を用いて行っており、対象までの距離が近い場合には視差が生じ違和感のない合成が難しいという問題がある。

Mori らは対象となるシーンの三次元構造を事前に撮影した画像群から復元を行い、これらの画像群と三次元構造を用いて View-dependent Texture Mapping によって除去対象領域の画像を生成する手法を提案している [11]。この手法では、シーンの三次元構造を考慮した除去対象領域の画像生成を行うことにより、事前に撮影された画像群と異なる視点で撮影された映像に対しても違和感のない合成が可能となっている。

一方で、三原らはライトフィールドカメラを用いて撮影された映像中から不要な物体を除去する手法を提案している [12]。この手法では、ライトフィールドカメラによって取得された光線情報より、不要物体に対応する光線を除去することによって映像の生成を行っている。除去対象となる領域については、特定のフレームで指定した後他フレームに伝播することで半自動で生成を行っている。ライトフィールドカメラを用いることで、実背景の情報を同時に取得することが可能であるが除去対象領域に対応する光線が取得されない場合には画像修復手法を併用する必要がある。

## 2.2 画像修復に基づく手法

一方、実計測に基づかない手法は、入力画像中のテクスチャを解析することによって除去対象領域を補完する画像修復技術に基づくアプローチが採られる。

Herling らは画像修復技術を応用し、除去対象領域外のテクスチャ情報を用いて除去対象領域のテクスチャを生成するアプローチを提案している [5, 6]。これらのアプローチでは、除去対象領域が平面で近似できる場合においては違和感のない映像の生成を実現することができるが、除去対象領域が平面で近似できない場合においては視点位置を移動した場合に除去対象領域の内外でテクスチャに不連続性が生じてしまう。

Kawai らは、Visual SLAM (Simultaneous Localization and Mapping) を用いてカメラの動きとシーンの三次元構造を取得し、取得された画像群と三次元構造を用いて物体の除去を行う手法を提案している [4]。この手法では、ユーザが指定した除去対象領域をいくつかの平面で近似した後に、各平面に対応するテクスチャを画像修復技術を用いて生成している。その後、生成されたテクスチャを映像中へ投影することによって、

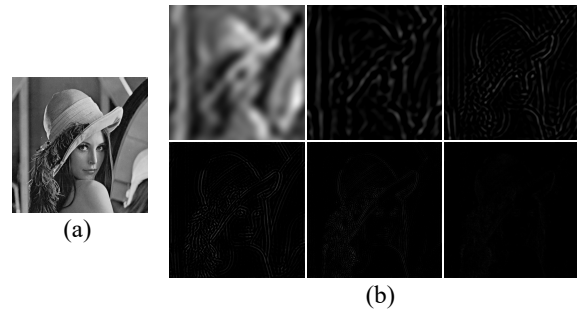


図1 Curvelet 変換を用いたサブバンド分解例、  
(a) 入力画像、(b) サブバンド群

隠消現実感映像の生成を行っている。さらに、Kawai らは、マーカを用いた拡張現実感を対象として、取得された映像中からマーカを除去する手法を提案している。この手法では、マーカ周辺の領域が平面で近似できない場合においても、マーカ周辺に存在する特徴点の動きを解析することでテクスチャの変形を許容した画像修復を実現している。

## 2.3 本研究の位置づけ

本研究では、映像中に写りこむ網状物体を画像修復のアプローチを用いて除去を行う。これまでに提案されている手法とは異なり、Curvelet 変換を用いた周波数領域での画像修復を行うことで視覚的な違和感の少ない映像の生成を目指す。また、本手法では、除去対象となる網状物体は周波数変換後に高周波成分を取り出すことで自動的に抽出を行う。

## 3 Curvelet 変換を用いた網状物体の除去

本研究では、映像中に写りこんだ網状物体を自動的に除去する。提案手法では、まず、入力フレームに対して Curvelet 変換を行う。その後、画像の高周波成分を取り出すことで、網状物体の領域を検出し網状物体除去のためのマスク画像を生成する。最後に、生成されたマスク画像を用いて欠損部分に対応するテクスチャを Curvelet 変換に基づく画像修復によって推定する。以下では、各処理について詳述する。

### 3.1 Curvelet 変換

Curvelet 変換は、物体のエッジに対する最適なスパース表現として知られている [10]。また、Curvelet 変換は画像ノイズの除去などの様々な不良設定問題において良い性能を示している。

Curvelet 変換は以下のように表される [13]。

$$c(j, l, k) = \sum_{\mathbf{u} \in \Omega} f(\mathbf{u}) \overline{\phi(j, l, k)} \quad (1)$$

ここで、 $f$  は画像のような二次元信号、 $c(j, l, k)$  は Curvelet 係数を表し、 $j, l, k$  は異なるスケールのサブバンド、方向、空間周波数窓の位置を表す。また、 $\phi(j, l, k)$

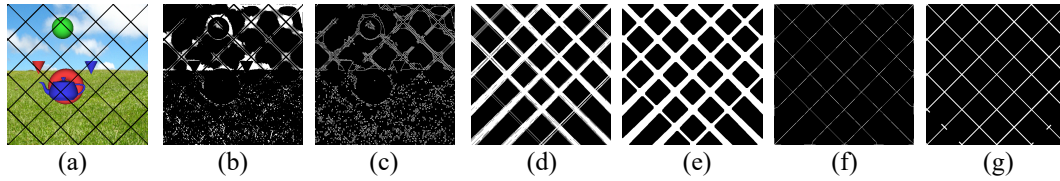


図2 (a) 入力画像, (b) 高周波成分のみを抽出した結果, (c) 輪郭成分を抽出した結果, (d) Hough 変換に絞り込みの結果, (e) クロージング処理を行った結果, (f) 細線化処理を行った結果, (g) 生成されたマスク画像

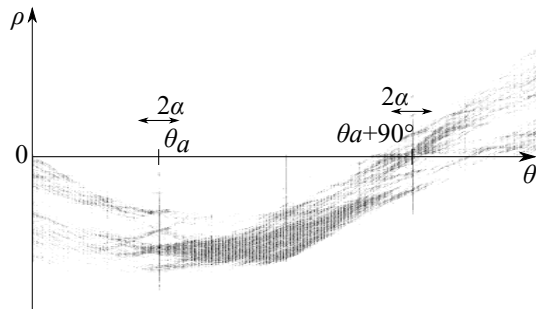


図3 網状物体を構成する直線の検出

は Curvelet 変換の母関数である. 図1に入力画像を6つのサブバンドに分解した例を示す. 本研究では, CurveletLab の非等間隔フーリエ変換 (USFFT) を使用して Curvelet 変換の実装を行い実験に用いた<sup>1</sup>.

### 3.2 周波数特徴に基づく網状物体の検出

ここでは網状物体の検出から, 画像修復に用いるマスクの生成までについて述べる.

#### サブバンド分解による高周波成分の抽出

FFT(Fast Fourier Transform) を行った結果から高周波成分に対応するサブバンドを抽出することによって, 網状物体の候補領域を抽出する. 網状物体の候補領域の抽出例を図2に示す. 網状物体が含まれている入力画像(図2(a))に対して Curvelet 変換を行い, 高周波成分に対応するサブバンドを抽出し逆 FFT 変換によって網状物体の候補領域を抽出した結果が図2(b)となっている. この画像に対して, 輪郭情報を抽出した結果が図2(c)となっている. このようにして取得された網状物体の候補領域について, 次項で述べる Hough 変換を用いた絞り込みを行い画像修復のための最終的なマスク画像を生成する.

#### Hough 変換を用いた網状物体領域の絞り込み

FFT 変換によって抽出した網状物体の候補領域(図2(c))に対して Hough 変換を行い直線成分を抽出する. 図3に Hough 変換によって得られる投票の結果を示す. ここで, 本研究では網状物体の検出を目的としていることから,  $\rho - \theta$  空間において  $\theta_a, \theta_a + 90^\circ$

の位置に投票が集中することが考えられる. そこで,  $\theta_a, \theta_a + 90^\circ$  での投票値が最大となる位置での  $\rho, \theta$  の組み合わせを網状物体を形成する直線の組として抽出する. ただし, 直線の検出誤差の影響を考慮し, 投票値は  $\pm\alpha$  のマージンを設定した上で算出を行う. 図2(d)に Hough 変換による絞り込みの結果を示す. さらに, 検出された直線の内部についてもマスク領域を設定するために図2(d)に対してクロージング処理を施す(図2(e)). 最後に, 図2(e)に対して細線化処理(図2(f))を行い, 線上に存在するピクセルの色情報を用いて網状領域の色情報を取得し, 図2(e)よりこの色情報を用いた絞り込みを行うことで最終的なマスク画像を生成する(図2(g)).

### 3.3 Curvelet 変換を用いた画像修復

生成されたマスク画像を用いて網状物体領域に対して Curvelet 変換に基づく画像修復処理を行う. この補間に関する問題は ROF モデルとして考えることができ[14], 誤差関数  $E_{total}$  以下のように表すことができる.

$$E_{total} = E_{prior} + \lambda E_{noise} \quad (2)$$

ここで, 以下に示すように  $E_{prior}$  は Curvelet 変換係数のスパース性を表す.

$$E_{prior} = \sum_{c \in W} \|c(j, l, k)\|_1 = \sum_{u \in \Omega} \|\Phi f(\mathbf{u})\|_1 \quad (3)$$

また,  $E_{noise}$  は画像全体  $\Omega$  からマスク領域  $I$  を除いた領域における観測された画像  $\tilde{f}(\mathbf{u})$  と推定された画像  $f(\mathbf{u})$  の差を表す.

$$E_{noise} = \sum_{u \in \Omega \setminus I} \|f(\mathbf{u}) - \tilde{f}(\mathbf{u})\|_2 \quad (4)$$

Curvelet 変換は物体のエッジ情報をスパースな表現に変換したものであるため,  $E_{prior}$  項によってこのスパース性に関する制約を導入する. ただし,  $L_0$  ノルムによるスパース性の表現を用いた場合には最適化が困難であるため,  $L_1$  ノルムを用いた近似を行っている. また,  $\Phi$  は Curvelet 変換を表し, この変換を推定画像  $f(\mathbf{u})$  に対して適用することで Curvelet 変換係数を得る.

<sup>1</sup><http://www.curvelet.org>

最終的に誤差関数は以下の形で表される。

$$E_{total} = \sum_{u \in \Omega} \|\Phi f(\mathbf{u})\|_1 + \lambda \sum_{u \in \Omega \setminus I} \|f(\mathbf{u}) - \tilde{f}(\mathbf{u})\|_2 \quad (5)$$

$\lambda$  は  $E_{noise}$  の影響をコントロールするための定数であり、本研究では網状物体領域外の画素値は修復の前後で同一の値をとるように非常に大きな値を設定する。最終的に、式 (5) を文献 [15] で紹介されているアルゴリズムを用いて最小化することで網状物体を除去した画像を得る。

## 4 実験

提案手法の有効性を示すため、シミュレーション画像と実画像に対して、網状物体の検出および除去を行った。

### 4.1 シミュレーション画像を用いた実験

シミュレーション実験では、まず、Curvelet 変換を用いた画像修復の性能を評価するために、マスク画像を与えた場合の画像修復結果について評価を行った。ここでは、比較のため従来手法として OpenCV<sup>2</sup> に実装されている偏微分方程式を用いた画像修復手法と比較を行った [16]。図 4 に修復結果を示す。また、表 1 に修復結果の PSNR の値を示す。提案手法、従来手法ともに PSNR の値は同等の値となっている。しかし、図 5 に修復結果の一部を拡大した図を示す。図より、修復された画像のエッジ部分においては、提案手法により復元された結果は滑らかな補間が行われていることが確認できる。これは、Curvelet 変換を用いることで、物体のエッジが滑らかに接続されるように復元された結果だと考えられる。このように、Curvelet 変換を用いることで、視覚的に自然な画像の生成が可能である。

次に、網状物体の検出精度について検証を行った。提案手法によって生成されたマスク画像を図 6 に示す。また、網状物体を構成する画素の検出率（検出された画素数/網状物体領域の全画素数）を表 2 に示す。結果より、全体の 90% 以上の網状物体領域を検出できていることが確認できる。ただし、画像端に存在している短い線分に関しては、Hough 変換に基づく直

表 1 完全なマスクを用いた実験結果 PSNR[dB]

画像	提案手法	従来手法 [16]
シーン A	25.817	25.558
シーン B	25.761	25.599
シーン C	27.627	27.065

線検出の際に投票数が少なくなり検出が難しい。このような問題に対して、今後、短い線分に対しても全体の網状物体のパターンを考慮することで検出可能なアルゴリズムの開発が必要である。

### 4.2 実画像を用いた実験

実画像に対して提案手法を適用した結果を図 7 に示す。実験では、画像中に存在する金網を提案手法を用いて除去を行った。実験結果より、自動でマスク画像の生成および画像の修復が行えていることが確認できる。ただし、修復対象領域が大きな場合にはばげが生じてしまうという問題がある。

## 5 まとめ

Curvelet 変換を用いた網状物体を自動的に検出、除去する手法を提案した。実験では、サブバンド分解と Hough 変換を用いることで、網状物体を自動的に検出できることを確認した。また、シミュレーション画像および実画像を用いた画像修復実験を行い、提案手法を用いることで画像中より網状物体を自動的に検出し除去することが可能であることを確認した。

今後の課題として、網状物体の検出精度の改善および時系列情報を用いた高精度化があげられる。

## 謝辞

本研究の一部は科学研究費補助金（若手研究（B）, No.26730091）によって行った。

## 参考文献

- [1] 森尚平, 一刈良介, 柴田史久, 木村朝子, 田村秀行. 隠消現実感の技術的枠組と諸問題: 現実世界に実在する物体を視覚的に隠蔽・消去・透視する技術について. 日本バーチャルリアリティ学会論文誌, Vol. 16, No. 2, pp. 239–250, Jun. 2011.
- [2] Norihiko Kawai, Masayoshi Yamasaki, Tomokazu Sato, and Naokazu Yokoya. Diminished reality for ar marker hiding based on image inpainting with reflection of luminance changes. *Proc. ITE Trans. Media Technology and Applications*, Vol. 1, No. 4, pp. 343–353, Oct. 2013.
- [3] Fabien Mourgues, Frederic Devernay, and Eve Coste-Maniere. 3d reconstruction of the operating field for image overlay in 3d-endoscopic surgery. *Proc. IEEE and ACM International Symposium on Augmented Reality*, pp. 191–192, Oct. 2001.
- [4] Norihiko Kawai, Tomokazu Sato, and Naokazu Yokoya. Diminished reality considering back-

表 2 網状物体の検出率

画像	検出率 [%]
シーン A	94.4
シーン B	96.7
シーン C	96.9

<sup>2</sup><http://opencv.org/>

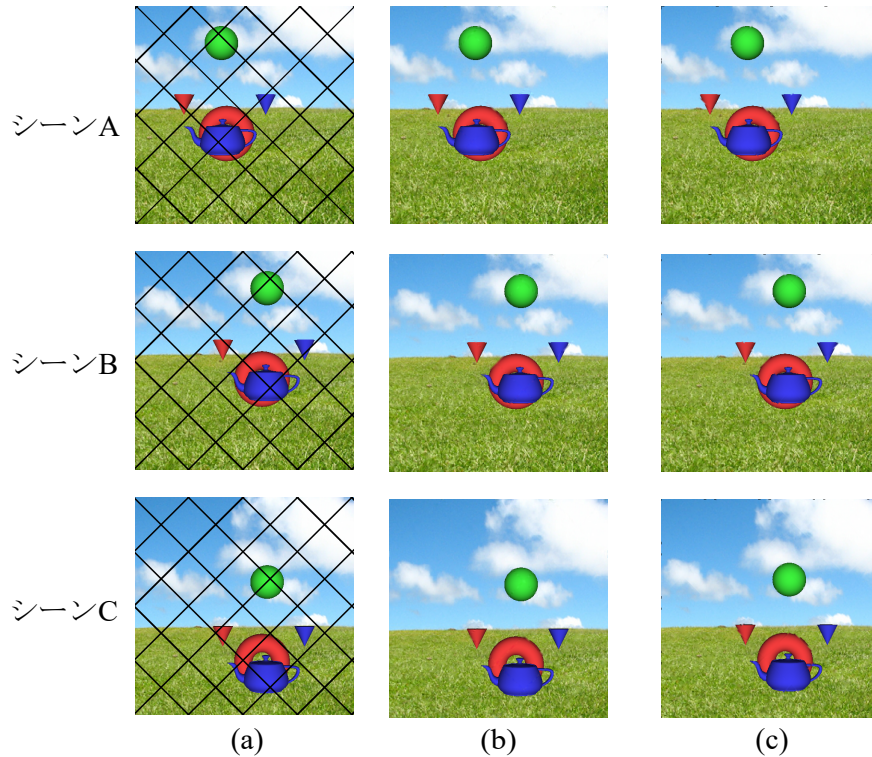


図4 マスク画像を与えた場合の結果, (a) 入力画像, (b) 提案手法, (c) 従来手法 [16]

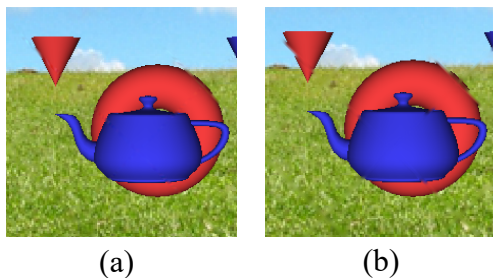


図5 修復結果の拡大図, (a) 提案手法, (b) 従来手法 [16]

ground structures. *Proc. IEEE International Symposium on Mixed and Augmented Reality*, pp. 259–260, Oct. 2013.

- [5] Jan Herling and Wolfgang Broll. Pixmix: A real-time approach to high-quality diminished reality. *Proc. IEEE International Symposium on Mixed and Augmented Reality*, pp. 141–150, Nov. 2012.
- [6] Jan Herling and Wolfgang Broll. Advanced self-contained object removal for realizing real-time diminished reality in unconstrained environments. *Proc. IEEE International Symposium on Mixed and Augmented Reality*, pp. 207–212, Oct. 2010.
- [7] Siim Meerits and Hideo Saito. Real-time diminished reality for dynamic scenes. *Proc. IEEE International Symposium on Mixed and Augmented Reality Workshops*, pp. 53–59, Oct. 2015.
- [8] Kunihiro Hasegawa and Hideo Saito. Diminished reality for hiding a pedestrian using hand-held camera. *Proc. IEEE International Symposium on Mixed and Augmented Reality Workshops*, pp. 47–

52, Sep. 2015.

- [9] 吉岡奨悟, 森尚平, 柴田史久, 木村朝子, 田村秀行. 隠消現実感における半隠消表示法の改良. 日本バーチャルリアリティ学会大会論文集, Vol. 19, pp. 533–536, Sep. 2014.
- [10] Jianwei Ma and Gerlind Plonka. The curvelet transform. *IEEE Signal Processing Magazine*, Vol. 27, No. 2, pp. 118–133, Mar. 2010.
- [11] Shohei Mori, Fumihisa Shibata, Asako Kimura, and Hideyuki Tamura. Efficient use of textured 3d model for pre-observation-based diminished reality. *Proc. IEEE International Symposium on Mixed and Augmented Reality Workshops*, pp. 32–39, Sep. 2015.
- [12] 三原基, 岩口堯史, 田中賢一郎, 久保尋之, 長原一, 向川康博. 光線空間を用いた不要物体除去リフォーカス画像の生成. 情報処理学会研究報告, Vol. 2015, No. 195-52, pp. 1–4, Jan. 2015.
- [13] Emmanuel Candes, Laurent Demanet, David Donoho, and Lexing Ying. Fast discrete curvelet transforms. *Proc. Multiscale Modeling & Simulation*, Vol. 5, No. 3, pp. 861–899, Dec. 2006.
- [14] Leonid I Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, Vol. 60, No. 1, pp. 259–268, Nov. 1992.
- [15] Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, Vol. 40, No. 1, pp. 120–145, May 2011.
- [16] M. Bertalmio, A. L. Bertozzi, and G. Sapiro. Navier-stokes, fluid dynamics, and image and video inpainting. *Proc. IEEE Computer Vision and Pattern Recognition*, pp. 355–362, Dec. 2001.

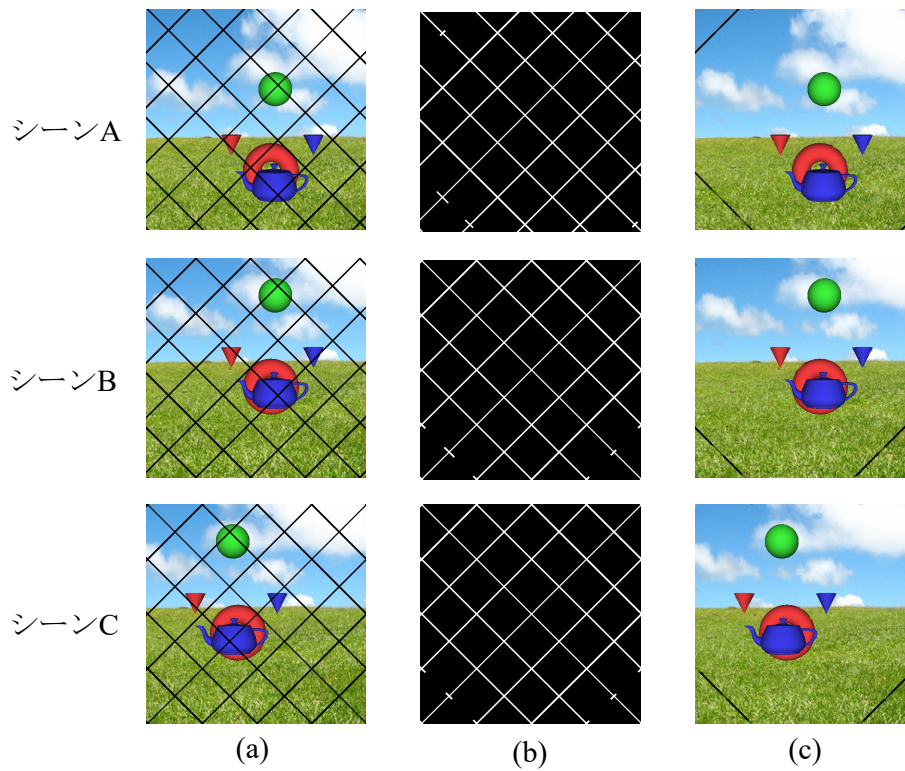


図6 自動で生成したマスク画像を利用した結果, (a) 入力画像, (b) マスク画像, (c) 画像修復結果



図7 実画像を用いた実験結果, (a) 入力画像, (b) マスク画像, (c) 画像修復結果