

# テクスチャの明度変化と局所性を考慮したパターン類似度を用いたエネルギー最小化による画像修復

河合 紀彦<sup>†</sup>      佐藤 智和<sup>†</sup>      横矢 直和<sup>†</sup>

Image Inpainting by Minimizing Energy Function Based on Pattern Similarity Considering Brightness Change and Spatial Locality

Norihiko KAWAI<sup>†</sup>, Tomokazu SATO<sup>†</sup>, and Naokazu YOKOYA<sup>†</sup>

あらまし 本論文では、写真についた傷など画像内の不要部分を取り除き、その欠損領域を自動的に修復する新たな手法を提案する。従来、欠損領域の修復に関しては、欠損領域と欠損領域以外のデータ領域のパターン類似度を用いて欠損領域全体の尤もらしさを表す目的関数を定義し、それを最適化することで画像修復を行う手法が提案されている。しかし、従来提案されている目的関数の最適化による修復では、以下に挙げる2つの問題により不自然なテクスチャが生じやすい。(1) データ領域におけるテクスチャパターンの種類に限りがあるため、単純なテクスチャのコピーによる修復を前提とした場合には、違和感のない修復に最低限必要なパターンがデータ領域に存在しないことが多い。(2) データ領域と欠損領域間のパターン類似度のみによる指標は自然なテクスチャの再現条件としては不十分である。本研究ではこのような問題に対して、テクスチャの明度変化を許容したパターン類似度を用いることで表現できるテクスチャパターンの拡張を行う。また、パターン類似度以外にテクスチャの局所性を表すコスト関数を用いることで不自然なテクスチャの生成を抑制する。実験では、様々な画像に対して欠損領域の修復を行い、従来手法との比較およびアンケート評価に基づく定性的評価により提案手法の有効性を示す。また、従来行われてきた評価方法である少人数・少数の画像による主観評価および結果画像と原画像の差分に基づく定量的評価の問題点を実験により明らかにする。

キーワード 画像修復, 画像補間, Image Inpainting, エネルギー最小化, パターン類似度

## 1. はじめに

インターネットの普及に伴い、個人が日常的に撮影した写真や映像をホームページやブログに掲載することが一般的に行われている。このような目的で、過去に撮影済みのアナログ写真をスキャナなどで電子化する際、アナログ写真の物理的な損傷(キズ, よごれ等)によりそのままの利用が難しい場合がある。また、デジタル写真や映像においては、利用意図に沿わない物体などが写っているため、そのままでは利用しづらい場合も起こりうる。このような問題に対して、写真についた傷や意図せず写りこんでしまった物体などの画像内の不要部分を取り除き、欠損領域を自動的に違和感なく修復することで画像の利用価値を高める画像

修復に関する研究が盛んに行われている。これら画像修復の手法は、輝度値の連続性を考慮する手法と欠損領域以外の領域(以下、データ領域)のテクスチャを用いる手法に大別できる。前者は、輝度値の連続性を考慮して欠損領域の周りから輝度値を滑らかに補間することで画像修復を行う手法[1]~[8]であり、写真に付いた引っかき傷のような細かい領域に対しては良好な修復画像を得ることができる。しかし、大きな領域を修復した場合には細かいテクスチャが表現できず、不鮮明な画像が生成されるという問題がある。そのため、近年では複雑なテクスチャも表現できる後者の手法が主流となっている。

データ領域のテクスチャを用いる手法は、欠損領域とデータ領域のパターン類似度に基づいてデータ領域のテクスチャを欠損領域に合成することで画像修復を行う。このような手法には、テクスチャの輝度値をパターン類似度の指標として用いる手法[9]~[18]と、

<sup>†</sup> 奈良先端科学技術大学院大学, 奈良県生駒市高山町 8916-5  
Nara Institute of Science and Technology, 8916-5 Takayama,  
Ikoma, Nara, 630-0192 Japan

フーリエ空間，ウェーブレット空間，固有空間等の特徴空間上の距離を類似度の指標として用いる手法 [19]～[22] がある．後者の手法は特徴空間への変換コストが大きく，またテクスチャの輝度値を用いた場合でも画像修復においては十分良好な結果が得られることから，本論文では前者の手法に着目する．

Efros ら [9] は，データ領域から欠損領域の境界部分に最も類似したテクスチャを欠損領域の外側から逐次的に合成する手法を提案している．この手法は複雑なテクスチャを再現することができるが，最終的に生成される画像の品質がテクスチャの合成順に大きく依存し，不連続なテクスチャが生じやすいという問題がある．そのため，決定済みの画素数，エッジの強さ，類似度等を考慮することで，違和感の少ない画像を生成可能なテクスチャの合成順を決定する手法 [10]～[12] が提案されている．しかし，これらの手法においても，やはり欠損領域の外部のテクスチャが複雑な幾何学的構造を持つ画像に対しては，不連続なテクスチャが生じやすいという問題がある．この問題に対して，欠損領域全体の尤もらしさを表す目的関数を定義し，それを最適化するアプローチが提案されている [13]～[15]．これらの手法では，グラフカット，確率伝搬法，EM アルゴリズムを用いて目的関数を最適化することで画像修復を行っている．

全体最適化による手法 [13]～[15] では，結果がテクスチャの合成順に依存せず，欠損領域全体に対して最適な画像を生成できるが，それでもなお以下に挙げる 2 つの問題により不自然な画像が生成される場合がある．(1) データ領域におけるテクスチャパターンの種類に限りがあるため，単純なテクスチャのコピーによる修復を前提とした場合には，違和感のない修復に最低限必要なパターンがデータ領域に存在しないことが多い．(2) データ領域と欠損領域間のパターン類似度のみによる指標は自然なテクスチャの再現条件として不十分である．従って，多くの画像に対して良好な結果を得るためには，これら 2 つの問題を考慮する必要があり，これまでもいくつかの試みがなされてきた．問題 (1) に対して，文献 [16] ではデータ領域におけるテクスチャの幾何学的な変形を許容することで，修復に用いるテクスチャパターンを拡張している．また，テクスチャの明度変化などの光学的な変化についても許容することで，より多くのパターンを修復に利用することが可能になると考えられるが，現在までにそのような研究事例は存在しない．問題 (2) については，

修復の前処理としてテクスチャの境界部分を手動で指定または自動で推定する手法 [17], [18] が提案されている．これらの手法では，テクスチャの境界を保つようにテクスチャを合成することで不連続なテクスチャを生じにくくしているが，複雑なテクスチャを持つ画像を対象とした修復では，修復に有効な境界部分を指定または推定することが難しい．

本論文では，多くの画像に対して違和感の少ない修復結果を得るために，従来とは異なるアプローチを採用する．まず，問題 (1) に対して，従来手法では考慮されてこなかったデータ領域のテクスチャの明度変化を許容したパターン類似度を用いることで，表現できるテクスチャパターンの拡張を行う．また，問題 (2) について，多くの画像において類似したテクスチャ同士が近傍に存在する確率が高い性質（以下，テクスチャの局所性）を考慮し，類似パターン間の画像上での距離に応じたコスト関数を用いることで，不自然なテクスチャの生成を抑制する．本研究では，テクスチャの全体最適化を用いる Wexler らの手法 [15] を基礎とし，これら 2 つの拡張を行うことにより高品位な画像修復を行う手法を提案する．更に，本論文では従来の研究ではあまり議論されてこなかった修復結果の評価方法の妥当性についても考察する．具体的には，従来から行われてきた少人数・少数の画像による主観評価および結果画像と原画像の差分に基づく定量的評価の問題点を実験により明らかにする．

以下，2 節では，テクスチャの明度変化と局所性を考慮したエネルギー関数を最小化することで画像修復を行う提案手法について述べる．3 節では，まず修復結果について従来手法との比較およびアンケート評価に基づく定性的評価により提案手法の有効性を示す．次に，従来行われてきた修復結果の評価方法の妥当性について考察する．最後に 4 節で，まとめと今後の課題について述べる．

## 2. テクスチャの明度変化と局所性を考慮したエネルギー最小化による画像修復

提案手法の処理の流れを図 1 に示す．本研究では，まず写真上の傷や不要な物といった修復したい領域を画像上で手動により指定し (a)，何らかの方法を用いて欠損領域に初期値となる画素値を与える (b)．次に，今回提案するエネルギー関数を最小化することで，欠損領域の修復を行う (c)．

本節では，まず従来から用いられてきたパターン類

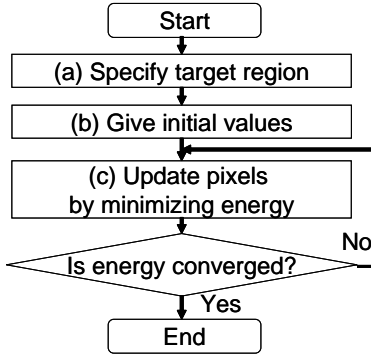


図 1 画像修復の処理の流れ  
Fig.1 Procedure of the proposed image inpainting method.

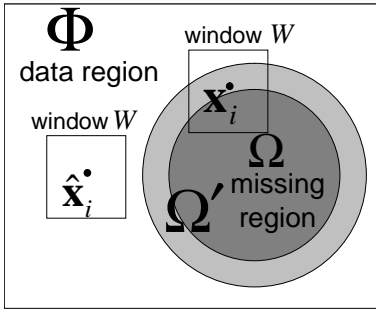


図 2 画像上の各領域  
Fig.2 Illustration of missing and data regions in an image.

似度 SSD (Sum of Squared Differences) によるエネルギー関数の定義について概説し、次に、今回提案するエネルギー関数とその最小化手法について述べる。

### 2.1 パターン類似度 SSD によるエネルギー関数の定義

本節では、Wexler らが提案した、SSD による画像修復のための目的関数 [15] について概説する。ただし、文献 [15] では目的関数を確率密度関数として定義しているが、本研究ではこれを本質的に同等なエネルギー関数として再定義する。

文献 [15] では、図 2 に示すように、画像をユーザが指定した欠損領域  $\Omega$  を含む領域  $\Omega'$  と、画像内の  $\Omega'$  以外のデータ領域  $\Phi$  に分け、領域  $\Omega'$  内の画像の尤もらしさをデータ領域  $\Phi$  内の画像パターンを用いて定義する。ここでは、画像内において一定の大きさの正方形ウィンドウ  $W$  内に一部でも  $\Omega$  が含まれるウィンドウの中心画素の集合を  $\Omega'$  とし、欠損領域の尤もらしさを表すエネルギーを、領域  $\Omega'$  内の画素  $x_i$  とそれに対

応するデータ領域  $\Phi$  内の画素  $\hat{x}_i$  周辺のパターン類似度 SSD の重み付き総和として以下のように定義する。

$$E_{org} = \sum_{i=1}^{N_{\Omega'}} w_{x_i} SSD(x_i, \hat{x}_i) \quad (1)$$

ただし、 $N_{\Omega'}$  は領域  $\Omega'$  内の画素数を表す。また、 $SSD(x_i, \hat{x}_i)$  は以下のように表される。

$$SSD(x_i, \hat{x}_i) = \sum_{p \in W} \{I(x_i + p) - I(\hat{x}_i + p)\}^2 \quad (2)$$

ここで、 $I(x)$  は画素  $x$  の画素値を表す。また、ここでは重み  $w_{x_i}$  として、領域  $\Omega' \cap \bar{\Omega}$  では各画素の値が固定値となるため  $w_{x_i} = 1$  を、領域  $\Omega$  では境界に近いほど画素値の信頼度が高くなるため  $w_{x_i} = c^{-d}$  ( $d$  は  $\Omega$  の境界からの距離、 $c$  は定数) を用いる。従来手法 [15] では、定義したエネルギー関数  $E_{org}$  を最小化する欠損領域内の画素値  $I(x_i)$  とそれに対応する類似パターン位置  $\hat{x}_i$  を算出することで、欠損領域を修復する。

### 2.2 明度変化と局所性を考慮したエネルギー関数の拡張

本研究では、式 (1) による欠損領域の尤もらしさに基づくエネルギー関数  $E_{org}$  を、テクスチャの明度変化と局所性を考慮して拡張する。具体的には、テクスチャの明度変化を許容するために、輝度値の補正係数を導入する。また、テクスチャの局所性を考慮するために、欠損領域とデータ領域の対応する画素間の距離に応じたコスト項をエネルギー関数に加えることで、エネルギーを新たに次のように定義する。

$$E = \sum_{i=1}^{N_{\Omega'}} w_{x_i} [SSD'(x_i, \hat{x}_i) + w_{dis} SD(x_i, \hat{x}_i)] \quad (3)$$

ここで、 $SSD'(x_i, \hat{x}_i)$  は明度の変化を考慮したパターン類似度、 $SD(x_i, \hat{x}_i)$  は画素  $x_i$  と画素  $\hat{x}_i$  の距離に応じたコスト項であり、 $w_{dis}$  は局所性の強さを示す重みである。以下、 $SSD'(x_i, \hat{x}_i)$  と  $SD(x_i, \hat{x}_i)$  について詳述する。

[テクスチャの明度変化を許容したパターン類似度]

一般的な画像中において、照明条件が不均一である場合などに、局所的なパターン構造は同一であっても明度の異なるテクスチャの組が存在する場合がある。

このようなテキストパターンは、利用できるデータ領域が限られる一枚の画像からの画像修復に有用であると考えられるが、従来手法 [15] で用いられる SSD(式 (2)) は、テキストパターンの明度変化を許容していないため、明度の異なるテキストを修復に利用することができない。本研究では、式 (2) に明度の補正係数を加え、テキストの明度変化を許容することで、より多くのパターンを修復に利用し、修復画像の品質を向上させる。具体的には、テキストの明度変化を許容したパターン類似度  $SSD'$  を以下のように定義する。

$$SSD'(x_i, \hat{x}_i) = \sum_{p \in W} \{I(x_i + p) - \alpha_{x_i, \hat{x}_i} I(\hat{x}_i + p)\}^2 \quad (4)$$

ここで、 $\alpha_{x_i, \hat{x}_i}$  はテキストの明度変化を許容するための輝度値の補正係数であり、 $x_i, \hat{x}_i$  それぞれの画素の周辺の平均輝度値を用いることで、データ領域のテキストの明度を欠損領域のテキストの明度に合わせる。ただし、実際の画像において比較的大きな明度変化を式 (4) のような定数倍の変化として近似すると違和感のある画像が生成されやすいため、ここでは  $\alpha_{x_i, \hat{x}_i}$  の値を式 (5) に示す一定範囲内 ( $1 - D \leq \alpha_{x_i, \hat{x}_i} \leq 1 + D$ , ただし  $D$  は  $0 < D < 1$  の定数) に限定する。

$$\alpha_{x_i, \hat{x}_i} = \begin{cases} 1 - D & (\beta_{x_i, \hat{x}_i} < 1 - D) \\ \beta_{x_i, \hat{x}_i} & (1 - D \leq \beta_{x_i, \hat{x}_i} \leq 1 + D) \\ 1 + D & (\beta_{x_i, \hat{x}_i} > 1 + D) \end{cases}, \quad (5)$$

$$\beta_{x_i, \hat{x}_i} = \frac{\sqrt{\sum_{q \in W} I(x_i + q)^2}}{\sqrt{\sum_{q \in W} I(\hat{x}_i + q)^2}} \quad (6)$$

輝度値補正係数  $\alpha_{x_i, \hat{x}_i}$  を用いることによって、画像中でテキストの明度が連続的に変化している箇所を欠損領域がまたぐ場合においても、中間的な明度のテキストを生成することで、式 (4) のコストを下げることができ、結果的に急激な明度変化を抑制することができる。

[テキストの局所性]

多くの画像では類似したテキスト同士が近傍に存在する確率が高い。ここでは、このようなテキストの局所性を考慮したコスト関数  $SD$  を、シグモイド関

数を用いて次のように定義する。

$$SD(x_i, \hat{x}_i) = \frac{\|W\|}{1 + e^{\{-K(\|x_i - \hat{x}_i\| - X_0)\}}} \quad (7)$$

ただし、 $K, X_0$  は定数、 $\|W\|$  はウインドウ内の画素数を表す。テキストの局所性に関して、注目画素が存在する物体領域の範囲内では類似したテキストが存在する確率が一樣に高いと仮定する。同様に、物体領域の範囲外では類似したテキストが存在する確率が一樣に低いと仮定する。以上の仮定から、一定範囲の内外で値がほぼ一定となるシグモイド関数を用いて式 (7) に示す局所性に関するコスト関数を定義する。なお、欠損領域内における実際の物体の大きさは未知であるため、式 (7) において物体領域を一定の大きさとして仮定している。局所性を考慮することにより、注目画素の近傍に存在する適切なテキストが優先的に選択される確率が高くなる。

### 2.3 エネルギー最小化による画素値と類似パターン位置の更新

本研究では、Greedy Algorithm の枠組みを用いて式 (3) で定義したエネルギー  $E$  を最小化する。ここでは、類似パターンの組  $(x_i, \hat{x}_i)$  を固定し、かつ輝度値の補正係数  $\alpha_{x_i, \hat{x}_i}$  の変化が欠損領域内の画素値の変化に対して微小であると仮定することで、エネルギー  $E$  を欠損領域  $\Omega$  内の各画素で独立に扱えることに着目し、

- (i) 各画素  $x_i$  に対する類似パターン位置  $\hat{x}_i$  の更新
- (ii) 欠損領域内の画素値の並列的な更新

の2つのプロセスをエネルギーが収束するまで繰り返すことで、画像全体のエネルギーを最小化する。

プロセス (i) では、欠損領域内の画素値を全て固定することで、対応する類似パターン位置を更新する。具体的には、データ領域  $\Phi$  内の画素に対して  $SSD'$  と  $SD$  を算出し、以下の式を満たすパターン位置  $\hat{x}_i$  を決定することで類似パターン位置を更新する。

$$f(x_i) = \hat{x}_i = \operatorname{argmin}_{x' \in \Phi} (SSD'(x_i, x') + w_{dis} SD(x_i, x')) \quad (8)$$

プロセス (ii) では、類似パターンの組を固定し、式 (3) で定義したエネルギー  $E$  を最小化する欠損領域内の画素値  $I(x_i)$  を画素並列に更新する。以下では、パターンの組を固定した場合の画素値  $I(x_i)$  の算出手法について詳述する。ここではまず、エネルギー  $E$  を、

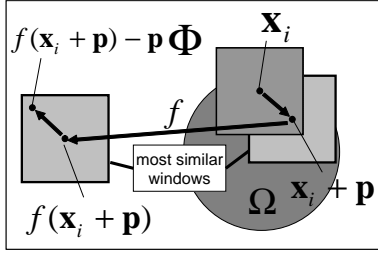


図3 エネルギー算出における画素の関係  
Fig.3 Relationship between pixels in energy calculation.

欠損領域内の各画素の要素エネルギー  $E(x_i)$  に分解する．図3に示すように，更新対象となる画素の位置を  $x_i$ ， $x_i$  を中心とするウインドウ  $W$  内の任意の点を  $x_i + p$  ( $p \in W$ ) とする．このとき，画素  $x_i + p$  を中心とするパターンに対して式(8)で求まる類似パターンの位置は  $f(x_i + p)$  であり，この類似パターンにおいて  $x_i$  と対応する画素の位置は  $f(x_i + p) - p$  となる．ここで，注目画素  $x_i$  に関する  $E$  の要素エネルギー  $E(x_i)$  は， $x_i$  と  $f(x_i + p) - p$  の画素値の関係， $x_i + p$  と  $f(x_i + p)$  の画素の周りの平均輝度値， $x_i$  と  $f(x_i)$  の位置の関係から算出でき，以下のように表すことができる．

$$E(x_i) = \sum_{p \in W} w_{(x_i+p)} \{I(x_i) - \alpha_{(x_i+p)(f(x_i+p))} I(f(x_i+p) - p)\}^2 + w_{dis} \frac{\|W\|}{1 + e^{-K(\|x_i - f(x_i) - X_0\|)}} \quad (9)$$

このとき，欠損領域全体のエネルギー  $E$  と各画素での要素エネルギー  $E(x_i)$  の関係は，以下のように表せる．

$$E = \sum_{i=1}^{N_\Omega} E(x_i) + C \quad (10)$$

$C$  は，領域  $\bar{\Omega} \cap \Omega'$  内にある画素に関するエネルギーであり，ここでは類似パターン位置が固定されているため定数として扱える．ここで，要素エネルギー  $E(x_i)$  の変数は  $I(x_i)$  であるため， $E$  を欠損領域内のある画素の画素値  $I(x_i)$  で偏微分すれば，エネルギー  $E$  を最小化する  $I(x_i)$  の必要条件は次式で表せる．

$$\frac{\partial E}{\partial I(x_i)} = \sum_{j=1}^{N_\Omega} \frac{\partial E(x_j)}{\partial I(x_i)} = 0 \quad (11)$$

このとき，式(9)内の輝度補正係数  $\alpha$  も変数  $I(x_i)$  を含んでいるが，画素値  $I(x_i)$  の変化に対する  $\alpha$  の変化は微小であると仮定し，

$$\frac{\partial \alpha_{x_j f(x_j)}}{\partial I(x_i)} = 0 \quad (\forall x_j \in \Omega') \quad (12)$$

とおけば， $\partial E(x_j)/\partial I(x_i) = 0$  ( $j \neq i$ ) となるため，式(11)は次のように表せる．

$$\frac{\partial E}{\partial I(x_i)} = \frac{\partial E(x_i)}{\partial I(x_i)} = 0 \quad (13)$$

ここで，

$$\frac{\partial E(x_i)}{\partial I(x_i)} = \sum_{p \in W} w_{(x_i+p)} \{2\{I(x_i) - \alpha_{(x_i+p)(f(x_i+p))} I(f(x_i+p) - p)\} \{1 - \frac{\partial \alpha_{(x_i+p)(f(x_i+p))}}{\partial I(x_i)} I(f(x_i+p) - p)\}\} \quad (14)$$

となることから，式(12)，(13)，(14)より， $E$  を最小化する画素値  $I(x_i)$  は以下のように算出される．

$$I(x_i) = \frac{\sum_{p \in W} w_{(x_i+p)} \alpha_{(x_i+p)(f(x_i+p))} I(f(x_i+p) - p)}{\sum_{p \in W} w_{(x_i+p)}} \quad (15)$$

なお，式(15)は式(12)を前提とした近似解であるが， $I(x_i)$  が収束するに従って輝度値の補正係数  $\alpha$  の値も収束するため，エネルギーが収束するにつれて良い近似解となる．

また，処理全体を通して効率的に大域最適解に近い解を得るために，粗密法による処理も行う．具体的には，あらかじめ画像ピラミッドを作成しておき，一定サイズのウインドウを用いて上位層から下位層へ順次エネルギー最小化処理を行う．その際，上位層の結果を下位層の初期値として投影することで良い初期値を与え，処理の高速化と局所解の回避を行う．また，元のサイズの画像に対しては，ウインドウサイズを順次小さくしながらエネルギー最小化処理を行うことで，細部までテクスチャを再現する．

### 3. 実験

本節では，まず提案手法と従来手法を含む5種類の手法を実装し，様々な特徴を持つ100枚の画像(200

表 1 本実験での各種パラメータの設定  
Table 1 Parameters in experiment.

Window size	$W$	max $9 \times 9$ min $3 \times 3$
Weight for sigmoid function	$w_{dis}$	120
Parameter in sigmoid function	$K$	0.4
	$X_0$	20
Range of coefficient $\alpha$	$D$	0.1

× 200 画素) を対象とした画像修復実験を行い、修復結果の比較およびアンケート評価に基づく定性的評価により提案手法の有効性を示す。次に、従来から画像修復に関する多くの研究で用いられてきた、少人数・少数の画像による主観評価および結果画像と原画像の差分に基づく定量的評価について、アンケートの評価結果を用いて考察を行い、問題点を明らかにする。

### 3.1 修復結果の比較およびアンケート評価による提案手法の評価

本節ではまず、実験で用いた 100 枚の画像のうち、画像内の同一の物体領域上においてテクスチャに対する明度変化やパターンの幾何学的な構造 (以下、パターン構造) の変化が特徴的な 4 枚の画像について、本手法が基礎としている Wexler らの手法 [15] による修復結果と提案手法による修復結果を比較することで、提案手法の有効性を示す。次に、テクスチャの逐次的な合成により修復を行う手法の中で代表的な Criminisi らの手法 [11]、Wexler らの手法 [15]、提案手法によるそれぞれ 100 枚全ての修復結果に対するアンケート評価に基づく定性的評価を行うことで、提案手法の有効性を客観的に判定する。

なお、本実験では、PC(CPU:Xeon 3.2GHz メモリ:8GB) を用いて修復処理を行い、エネルギー関数の各種パラメータは、事前に数枚の画像に対して行った予備実験の結果から、経験的に表 1 に示すように設定した。また、欠損領域は原画像の一部を塗りつぶすことにより手動で与え、欠損領域の画素値の初期値として欠損領域の境界画素の輝度値の平均値を与えた。

#### 3.1.1 様々な特徴を持つ画像を対象とした修復結果の比較

本項では、実験に用いた 100 枚の画像のうち図 4(a) に示す、同一の物体領域上においてテクスチャに対する明度変化やパターン構造の変化が特徴的な 4 枚の画像について提案手法と従来手法の修復結果を比較し、考察する。図 4(b) に手動で与えた欠損領域を、同図

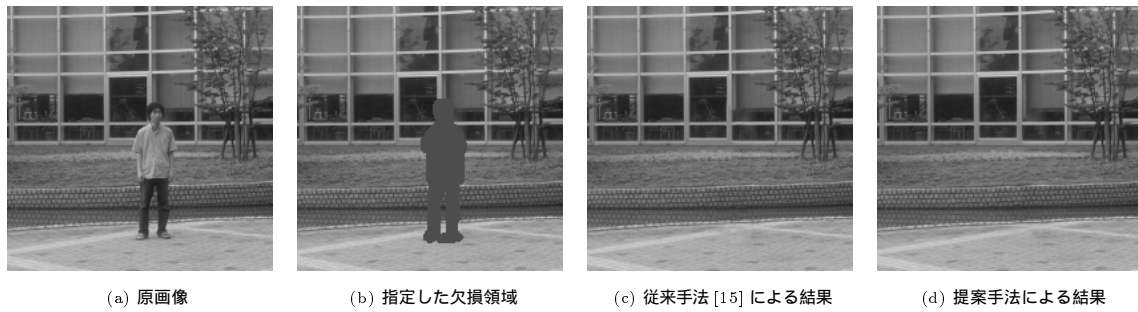
中 (c) に 2.1 節で述べたパターン類似度 SSD をエネルギーとした従来手法 [15] による修復画像を、(d) に提案手法を適用した修復画像を示す。以下、画像 (I) ~ (IV) に関して考察する。

画像 (I) は欠損領域周辺の同一の物体領域に対する照明条件が一樣でテクスチャの明度変化が小さく、また、パターン構造の変化も小さい画像である。このような画像に対する修復結果においては、従来手法、提案手法ともに違和感の少ない画像が生成されており、主観的には違いが感じられない。

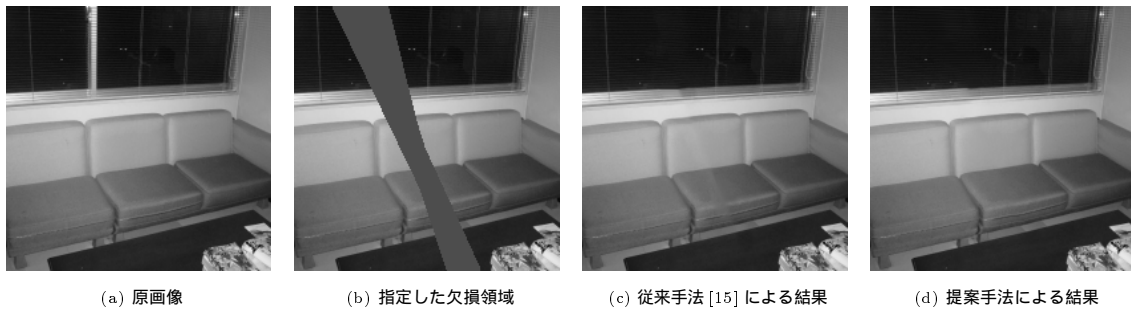
画像 (II) は、欠損領域周辺の同一の物体領域においてテクスチャが平坦でパターン構造は類似しているが、照明条件が一樣でなく明度変化が大きい画像である。従来手法では、座席部と背もたれ部の修復領域において、輝度値が滑らかに続いていないため偽エッジが発生し違和感が生じている。これに対して、提案手法では明度変化を考慮しているため輝度値が滑らかにつながり、違和感の少ない画像が生成されている。

画像 (III) は、欠損領域周辺の同一の物体領域に対する照明条件が一樣でテクスチャの明度変化は小さいが、物体の大きさと形が一樣でないため、画像上におけるテクスチャパターンの構造も一樣でない画像である。この画像では、欠損領域周辺が一見類似したテクスチャに見えるが、小石の大きさが均一でなく、単純な SSD による類似度を用いた場合、欠損領域とデータ領域のテクスチャの適切な対応付けが難しい。これにより、従来手法では、修復領域の一部で適切な対応付けが行われず、結果として修復領域が白っぽく浮き出ている。これに対して、提案手法ではテクスチャの局所性を考慮することで、パターン構造の変化にも関わらず欠損領域周辺のパターンが欠損領域内のテクスチャに対応付けられ、周辺領域と類似した色のテクスチャで欠損領域が修復されている。

画像 (IV) は、欠損領域周辺の同一の物体領域に対する照明条件が一樣でないため明度変化が大きい。また、中央に写る物体の三次元構造は規則的であるが、透視投影効果により画像上でのテクスチャパターンの構造が連続的に変化している。従来手法では、修復領域内部の壁や床の明度の急激な変化により違和感が生じている。また、黒い四角の模様も再現されず完全にぼけてしまっている。これに対して提案手法では、輝度値が滑らかにつながり、またパターンも連続的なものが生成されているため、比較的違和感の少ない画像が生成できている。



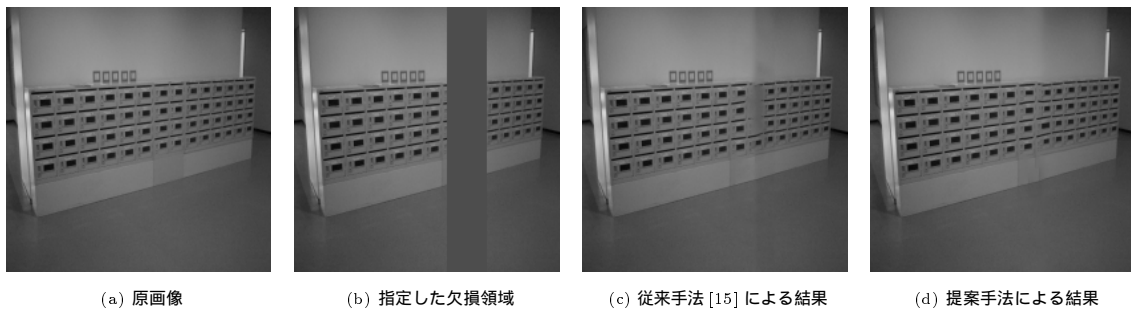
画像 (I) : テクスチャの明度変化やパターン構造の変化が小さい画像



画像 (II) : テクスチャの明度の大きな変化を伴う画像



画像 (III) : パターン構造の変化を伴う画像



画像 (IV) : テクスチャの明度の大きな変化とパターン構造の変化を伴う画像

図 4 様々な特徴を持つ画像に対する画像修復  
Fig. 4 Image inpainting for four representative images.

表 2 従来手法 [15] と提案手法による処理時間の比較  
Table 2 Processing time.

	Conventional method	Proposed method
Image (I)	2'17"	8'45"
Image (II)	3'17"	12'15"
Image (III)	3'28"	18'59"
Image (IV)	5'25"	18'32"

次に、従来手法と提案手法による処理時間の比較を行った。表 2 に各画像に対する各手法の処理時間を示す。同表から提案手法では従来手法に対して 3~5 倍程度の処理時間が必要であることが分かる。これは、輝度値補正係数と局所性を考慮した項の計算コストが増えたことによる。

### 3.1.2 アンケート評価に基づく定性的評価実験と考察

本項では、提案手法と従来手法を含む 5 種類の手法による各々 100 枚の修復画像を用いた 37 名の被験者によるアンケート評価実験について述べる。なお、本実験では以下の 5 手法で画像修復を行った。

手法 A テクスチャの逐次的合成による Criminisi らの手法 [11]

手法 B テクスチャの全体最適化による Wexler らの手法 [15]

手法 C 提案手法においてテクスチャの明度変化のみ考慮したエネルギー関数を用いた場合

手法 D テクスチャの局所性のみ考慮したエネルギー関数を用いた場合

手法 E テクスチャの明度変化と局所性の双方を考慮したエネルギー関数を用いた場合

また、被験者は 20 代前半の男女であり、全員コンピュータを日常的に扱っている。本実験は、被験者により修復画像を主観的に評価してもらうことで、提案手法の有効性を客観的に示すことを目的とする。

#### ・評価方法

被験者には、アンケート評価のためのウェブページにアクセスしてもらい、ウェブページ上に並べられた 100 枚の入力画像に対する 500 枚の修復画像に対して 5 段階の点数評価を行ってもらった。修復結果の採点ページ上では、5 種類の手法により修復した結果画像を入力画像ごとにランダムな順序で並べ、評価者には各画像がどの手法による修復結果であるのかを知らせなかった。また、本実験では、修復画像を個人のホームページや書籍・雑誌等の写真として利用することを前提として、全く使えない画像を 1 点、十分使える画

表 3 100 枚の画像に対する点数の平均点と最高点を取った回数

Table 3 Average scores and number of best-scoring images from 100 images.

	Average of score	Number of highest-scored images
Method A [11]	2.21	7
Method B [15]	3.24	7
Method C	3.39	21
Method D	3.42	20
Method E	3.60	45

表 4 有意水準 5% と設定した時の t 検定による各手法間を比較した場合の検定統計量 (有意差がある場合は数値に下線)

Table 4 T-value produced by the t-test between a pair of methods with a 5% significant level. (If significant difference is observed, the t-value is underlined.)

Method	B [15]	C	D	E
A [11]	<u>9.053</u>	<u>9.138</u>	<u>10.870</u>	<u>11.046</u>
B [15]	-	<u>2.225</u>	<u>4.575</u>	<u>5.919</u>
C	-	-	0.403	<u>4.397</u>
D	-	-	-	<u>3.351</u>

像を 5 点という回答基準で採点してもらった。

#### ・アンケートの評価結果と考察

手法 A~E により出力された各々 100 枚の修復結果につけられた点数の平均値と各手法が最高点を取った回数を表 3 に示す<sup>(注1)</sup>。表 3 から、平均点で比較すると、手法 E、手法 D、手法 C、手法 B、手法 A の順で高く、最高点の回数で比較すると、手法 E、手法 C、手法 D、手法 A・B の順となった。また、各画像への評価点を基に有意水準を 5% を設定した t 検定により各手法間を比較することで統計的な裏付けを行った。表 4 に各手法間の検定統計量と有意差の有無を示す。同表に示すように、手法 C と手法 D の間には有意な差が見られず、他は有意な差が見られた。従って、手法 E、手法 C・D、手法 B、手法 A の順に多くの画像に対して良好な結果を得られる手法であると言える。また、テクスチャの明度変化のみを考慮した手法 (手法 C) とテクスチャの局所性のみ考慮した手法 (手法 D) は、どちらも従来手法 (手法 A・B) よりも良い評価を得ており、各々の要素をエネルギー関数に導入するこ

(注1): 実験に用いた 100 枚の画像と修復結果およびその評価を <http://yokoya.naist.jp/research2/inpainting/> に示す。



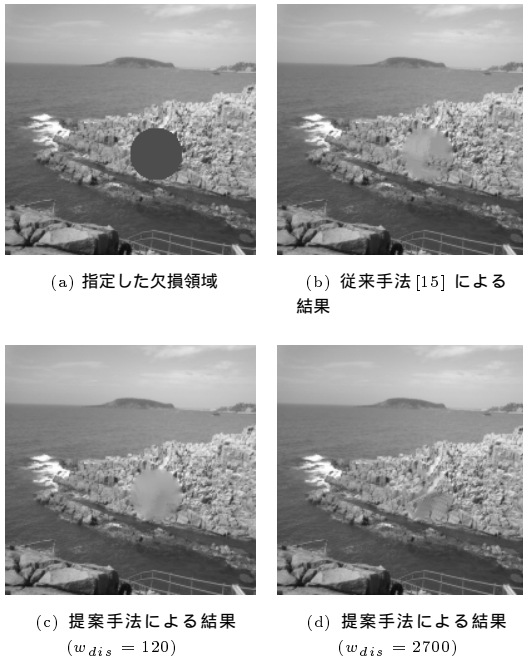


図 5 従来手法・提案手法共に評価が低い画像  
Fig. 5 Example images for which both the proposed and conventional methods have a problem.

との有効性が確認できる。

ただし、100 枚の画像のうち 7 枚の画像は、従来手法 [15] (手法 B)・提案手法 (手法 E) とともに修復結果に問題があり、どちらの結果に対する被験者間の平均点も 1 点台となった。しかし、これらの画像のうち、図 5(a) に示す画像では、欠損領域の周りでテクスチャパターンの幾何学的な構造に周期性がなく様々なパターンが混在しているため、局所性を考慮したにも関わらず、データ領域と欠損領域のテクスチャの対応付けがうまくいかずにぼけてしまい、違和感のある画像が生成されている。これに対して、局所性に関する重みを高め、 $w_{dis} = 2700$  とした上で修復を実行したところ、図 5(d) に示す違和感の少ない画像が生成された。このように、単一の重み係数による修復では必ずしも最適な画像が得られず、更に多くの画像に対して良好な結果を得るためには、画像の特徴を考慮してパラメータを動的に決定する必要がある。

### 3.2 画像修復手法における評価方法の考察

従来の画像修復に関する文献では、主に以下に述べる 2 種類の方法により結果画像の評価を行っている。すなわち、一方は、小人数・少数の画像による主観評価法であり、もう一方は、原画像と結果画像の画素値

の差分に基づく定量的評価法である。しかし、このような評価方法の妥当性についてはあまり議論されてこなかった。本節では、従来の画像修復の研究で行われてきた評価方法を 3.1.2 項で得られたアンケート評価結果を用いて考察し、問題点を明らかにする。

#### 3.2.1 アンケート評価における被験者数と画像枚数の考察

従来行われてきた多くの研究では、少人数の被験者によって少数の画像を主観的に評価することが一般的に行われている。しかし、一般的には画像や人による評価のばらつきが存在するため、手法の優劣が正しく評価されていない可能性がある。ここでは、前述の入力画像 100 枚、被験者数 37 名によるアンケート評価結果における平均点を基準とした手法の優劣を正解とし、実験に用いる画像枚数および被験者数を減少させた場合の結果の信頼性を検証する。

本論文では、結果の信頼度を示す指標として、順位誤差  $RE$  (Ranking Error) を以下のように定義する。

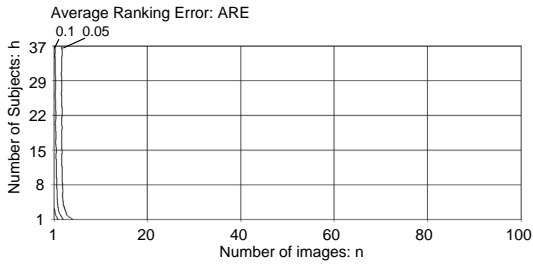
$$RE(n, h) = \begin{cases} 0 & (\forall i, g(i, 100, 37) - g(i, n, h) = 0) \\ 1 & (\text{otherwise}) \end{cases}$$

(16)

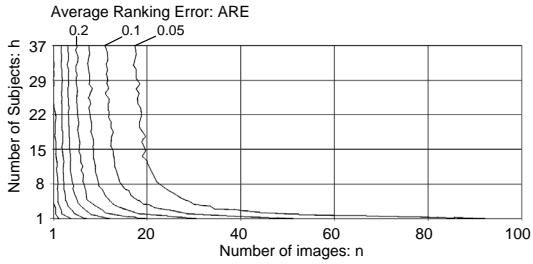
ただし、 $g(i, n, h)$  は画像枚数  $n$  および被験者数  $h$  でアンケート評価を行った際の手法  $i$  ( $i \in \{A, B, C, D, E\}$ ) の順位を表す。本実験では画像枚数 100 枚、被験者数 37 名の中からランダムに  $n$  枚の画像と  $h$  名の被験者を選び  $RE(n, h)$  を算出する。これを 10000 回施行し、 $RE(n, h)$  の平均値として平均順位誤差  $ARE(n, h)$  を求めることで結果の信頼性を検証する。

図 6 に画像枚数  $n$ 、被験者数  $h$ 、平均順位誤差  $ARE$  の関係について、主観評価の平均点に大きな差があった 2 手法 (手法 A と手法 E) を用いた場合 (a)、平均点に有意差はあるものの差が比較的小さい 2 手法 (手法 D と手法 E) を用いた場合 (b)、5 手法全て (手法 A ~ 手法 E) を用いた場合 (c) の結果を示す。同図中には  $ARE$  の値が 0.05 変化するごとに等値線が引かれている。また、各図では  $g(i, 100, 37)$  として、3.1.2 項で得られた結果から、表 5 に示す値を用いる。なお、手法 C と手法 D の間には有意な差がないため、図 6(c) ではそれぞれ 3 か 2 を当てはめて、少なくとも一方の  $RE$  が 0 の場合は、 $RE$  を 0 として  $ARE$  を算出している。

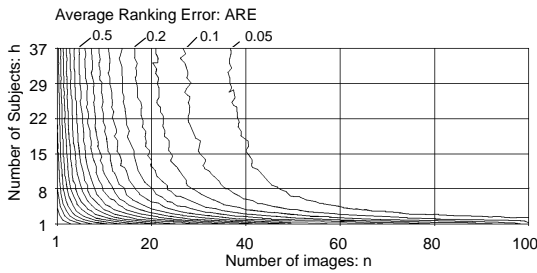
手法の順位が正解の順位と異なった場合、 $RE$  が 1



(a) 手法 A と手法 E に関する ARE



(b) 手法 D と手法 E に関する ARE



(c) 手法 A ~ E に関する ARE

図 6 画像枚数と被験者数と平均順位誤差の関係

Fig.6 Relationships among the number of images and subjectives and ARE.

表 5 図 6(a) ~ (c) で用いた  $g(i, 100, 37)$  の値  
Table 5 Value of  $g(i, 100, 37)$  in Fig6(a)-(c).

	$i=A$	$i=B$	$i=C$	$i=D$	$i=E$
Fig6(a)	2	-	-	-	1
Fig6(b)	-	-	-	2	1
Fig6(c)	5	4	3 or 2	3 or 2	1

になることから, ARE が 0.05 の場合, 95%の確率で優劣の順位が正解と一致する. 図 6(a) から主観評価による平均点の差が大きい手法 A と手法 E の優劣の比較において, 被験者数  $h = 2$  の時, 画像枚数  $n > 3$  で  $ARE < 0.05$  を満たすことが分かる. よって, 比較的差の大きい 2 者の比較においては, 少人数・少数の画像を用いても, 多人数・多数の画像による主観的評価

表 6 29 枚の画像に対する RMSE とアンケート評価値の平均値

Table 6 Average of RMSE and average score by subjects for 29 images.

	RMSE	Score by subjects
Method A [11]	42.95	1.99
Method B [15]	28.40	3.11
Method C	27.83	3.44
Method D	28.36	3.34
Method E	27.44	3.66

結果に対して, 優劣が入れ替わることは少ない. これに対して, 図 6(b) から, 平均点の差が小さい手法 D と手法 E の比較においては,  $ARE < 0.05$  を満たすためには, 被験者数  $h = 2$  の時, 画像枚数  $n > 49$  が必要であり, 被験者数  $h = 37$  の時, 画像枚数  $n > 18$  の条件が必要であることが分かる. よって, 手法間の評価の差が比較的小さい 2 手法の優劣を評価する場合には, 数人による数枚の画像に対する評価結果が, 多くの被験者および多くの画像を用いた場合の評価結果と異なる可能性が高い. 一方, 図 6(c) から, 5 手法の優劣を検証する場合, 95%以上の確率で正解と一致する ( $ARE < 0.05$ ) ためには, 同図 (b) の場合よりも更に多くの人数, 多くの画像による検証が必要となることが分かる.

### 3.2.2 定性的評価と定量的評価の関係の考察

従来の画像修復に関する研究では, 修復結果の定量的な評価指標として, 修復画像と原画像の画素値の差分に基づいて計算される MSE (Mean Squared Error), RMSE (Root Mean Squared Error), PSNR (Peak Signal-to-Noise Ratio) 等が用いられてきた. 本実験では RMSE を用いて修復画像に対する定量的評価を行い, アンケート評価に基づく定性的評価との関係を考察する. RMSE は以下のように定義される.

$$RMSE = \sqrt{\frac{\sum_{N_{\Omega}} (I_{res}(\mathbf{x}) - I_{org}(\mathbf{x}))^2}{N_{\Omega}}} \quad (17)$$

ただし,  $I_{res}$  は結果画像の画素値を表し,  $I_{org}$  は原画像の画素値を表す. また,  $N_{\Omega}$  は欠損領域の画素数を表す. なお, 本実験で用いた 100 枚の画像のうち 71 枚の画像においては, 特定の物体を取り除くために欠損領域の指定を行っており, 原画像と結果画像を比較することは無意味である. したがって, 物体の位置に関わらず欠損領域の指定を行った残りの 29 枚に対して RMSE の評価を行った.

表 6 に各手法による 29 枚の結果画像に対する RMSE の平均値とアンケート評価の平均値を示す. また 29

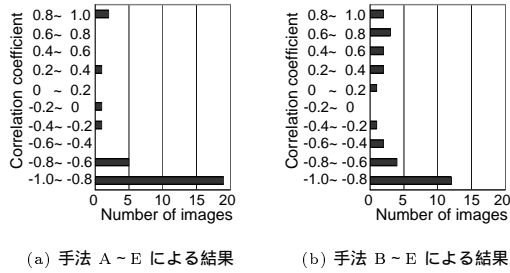


図 7 RMSE とアンケート評価の間の相関係数と画像の枚数の関係

Fig. 7 Distribution of correlation coefficient between RMSE and subjective score.



図 8 RMSE とアンケート評価の間に正相関が得られた画像

Fig. 8 Images with inverse correlation between RMSE and subjective evaluations.

枚の画像に対する RMSE とアンケート評価の間の相関係数の分布を図 7 に示す．同図中 (a) は手法 A ~ E の 5 手法を用いた場合，(b) は手法 A を除いた 4 手法を用いた場合の分布を示している．なお，ここでは相関係数  $Cc$  を以下の式で算出した．

$$Cc = \frac{N_m \sum_i R_i S_i - \sum_i R_i \sum_i S_i}{\sqrt{N_m \sum_i R_i^2 - (\sum_i R_i)^2} \sqrt{N_m \sum_i S_i^2 - (\sum_i S_i)^2}} \quad (18)$$

ただし， $N_m$  は比較に用いた手法の数であり， $R_i$ ， $S_i$  はそれぞれ RMSE の値，被験者による評価値の平均値である．

表 6 と図 7(a) から RMSE とアンケート評価の間には強い負の相関関係が確認できるため，RMSE とアンケート評価の評価結果はおおよそ一致すると考えられる．しかし，図 8 に示す 3 枚の画像においては，正の相関関係が見られた．これら 3 枚の画像は他の画像より比較的周波数が高いテクスチャを持つ画像である．文献 [21] でも指摘されているが，現在定量的評価に用いられている MSE，RMSE，PSNR は 2 つの画像間の対応する画素値の差分を用いて計算されるため，生成される画像の見た目の違和感に関係なく，高周波テ

クスチャの位相のずれに対して敏感に値が変化する傾向がある．従って，RMSE による評価は高周波成分を多く含む画像の評価には適していないと言える．図 7(b) はアンケート評価の評価値が顕著に低い手法 A を除いた 4 手法による相関係数の分布であるが，約 3 分の 2 の画像については負の相関関係，約 3 分の 1 の画像については正の相関関係を示している．これは，RMSE の微小な差が必ずしも画質の差を表していないことを示している．以上のことから，RMSE は人間の主観による画質の評価と一致することも多いが，画像修復結果の絶対的な判定指標として用いることは難しい．

#### 4. むすび

本論文では，従来から提案されているパターン類似度による目的関数を拡張することで画像修復を行う新しい手法を提案した．多くの画像に対して違和感の少ない修復結果を得るために，以下に挙げる 2 点において拡張を行った．

- (1) テクスチャの明度変化の許容によるテクスチャパターンの拡張
  - (2) テクスチャの局所性を考慮したコスト関数の導入
- これら 2 つの要素を考慮したエネルギー関数を新たに定義し，これを最小化することで，不連続・不明瞭なテクスチャの少ない画像修復を行う手法を提案した．

実験では，100 枚の画像に対して従来手法と提案手法により修復を行った結果を比較し，37 名の被験者による定性的評価を行うことで，様々な特徴を持つ多くの画像に対して，提案手法が従来手法よりも良好な修復画像を生成できることを確認した．また，アンケート評価における被験者数や画像枚数を変化させた場合の結果の信頼性についても検証を行い，数人・数枚の画像による主観評価では，手法の優劣を正しく評価できない可能性を示した．最後に，結果画像と原画像の差分に基づく定量的評価と定性的評価との関係について考察し，現在行われている定量的評価についても手法の優劣を正しく評価できていない可能性があることを明らかにした．今後，更に多くの画像に対して良好な結果を得るために，画像の特徴から最適なパラメータを決定する手法を確立する必要がある．また，提案手法の処理時間のうち類似パターンの探索に要する時間が多くを占めているため，データ領域のサンプルデータから分類木を予め作成し，探索回数を削減することで，処理の高速化を実現する．

## 文 献

- [1] 前田浩幸, 高橋健一, 太田正光: “欠損画像の修復処理の方式”, 電子情報通信学会論文誌 D, Vol. J69-D, No. 1, pp. 91-97, 1986.
- [2] S. Masnou and J.M. Morel: “Level Lines Based Disocclusion,” Proc. IEEE Int. Conf. on Image Processing, Vol. 3, pp. 259-263, 1998.
- [3] M. Bertalmio, G. Sapiro, V. Caselles and C. Ballester: “Image Inpainting,” Proc. ACM SIGGRAPH2000, pp. 417-424, 2000.
- [4] C. Ballester, M. Bertalmio, V. Sapiro and J. Verdera: “Filling-In by Joint Interpolation of Vector Fields and Gray Levels,” IEEE Trans. on Image Processing, Vol. 10, No. 8, pp. 1200-1211, 2001.
- [5] T. Chan and J. Shen: “Non-texture Inpainting by Curvature-Driven Diffusions (CDD),” Journal of Visual Communication and Image Representation, Vol. 12, No. 4, pp. 436-449, 2001.
- [6] A. Levin, A. Zomet and Y. Weiss: “Learning How to Inpaint from Global Image Statistics,” Proc. IEEE Int. Conf. on Computer Vision, Vol. 1, pp. 305-312, 2003.
- [7] 小川貴弘, 長谷山美紀, 北島秀夫: “オブティカルフローを用いた静止画像における失われた輝度値の復元”, 電子情報通信学会論文誌D-II, Vol. J87-D-II, No. 9, pp. 1786-1795, 2004.
- [8] D. Tschumperlé: “Curvature-Preserving Regularization of Multi-valued Images Using PDE's,” Proc. European Conf. on Computer Vision, Vol. 2, pp. 295-307, 2006.
- [9] A.A. Efros and T.K. Leung: “Texture Synthesis by Non-parametric Sampling,” Proc. IEEE Int. Conf. on Computer Vision, pp. 1033-1038, 1999.
- [10] R. Bornard, E. Lecan, L. Laborelli and J. Chenot: “Missing Data Correction in Still Images and Image Sequences,” Proc. ACM Int. Conf. on Multimedia, pp. 355-361, 2002.
- [11] A. Criminisi, P. Pérez and K. Toyama: “Region Filling and Object Removal by Exemplar-Based Image Inpainting,” IEEE Trans. on Image Processing, Vol. 13, No. 9, pp. 1200-1212, 2004.
- [12] B. Li, Y. Qi and X. Shen: “An Image Inpainting Method,” Proc. IEEE Int. Conf. on Computer Aided Design and Computer Graphics, pp. 531-536, 2005.
- [13] C. Allène and N. Paragios: “Image Renaissance Using Discrete Optimization,” Proc. IEEE Int. Conf. on Pattern Recognition, pp. 631-634, 2006.
- [14] N. Komodakis and G. Tziritas: “Image Completion Using Global Optimization,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 442-452, 2006.
- [15] Y. Wexler, E. Shechtman and M. Irani: “Space-Time Completion of Video,” Trans. on Pattern Analysis and Machine Intelligence, Vol. 29, No. 3, pp. 463-476, 2007.
- [16] I. Drori, D. Cohen-Or and H. Yeshurun: “Fragment-Based Image Completion,” Proc. ACM SIGGRAPH2003, pp. 303-312, 2003.
- [17] J. Sun, L. Yuan, J. Jia and H. Shum: “Image Completion with Structure Propagation,” Proc. ACM SIGGRAPH2005, pp. 861-868, 2005.
- [18] J. Jia and C. Tang: “Image Repairing: Robust Image Synthesis by Adaptive ND Tensor Voting,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 643-650, 2003.
- [19] 東海林健二: “テクスチャ画像における欠損部修復の手法”, 電子情報通信学会論文誌 D, Vol. J71-D, No.9, pp. 1701-1708, 1988.
- [20] 天野敏之, 佐藤幸男: “固有空間法を用いた BPLP による画像補間”, 電子情報通信学会誌D-II, Vol. J85-D-II, No. 3, pp. 457-465, 2002.
- [21] 天野敏之, 佐藤幸男: “kBPLP 法を用いた高次元非線形射影による画像補間”, 電子情報通信学会誌 D-II, Vol. J86-D-II, No. 4, pp. 525-534, 2003.
- [22] 井添慎太郎, 剣持雪子, 小谷一孔: “画像のフラクタル性に基づいたカーネル主成分分析による遮へい領域の原画像推定”, 電子情報通信学会 技術研究報告, PRMU2001-106, pp. 55-62, 2001.

(平成 xx 年 xx 月 xx 日受付)

河合 紀彦 (学生員)

2005 京大・工・情報卒 . 2007 奈良先端科学技術大学院大学情報科学研究科博士前期課程修了 . 現在 , 同大博士後期課程に在学中 . 情報処理学会 , IEEE 各会員 .

佐藤 智和 (正員)

1999 阪府大・工・情報工卒 . 2003 奈良先端科学技術大学院大学情報科学研究科博士後期課程修了 . 現在 , 同大情報科学研究科助教 . コンピュータビジョンの研究に従事 . 2001 本会学術奨励賞受賞 . 博士 (工学) . 情報処理学会 , IEEE 各会員 .

横矢 直和 (正員:フェロー)

1974 阪大・基礎工・情報工卒 . 1979 同大大学院博士後期課程修了 . 同年電子技術総合技術研究所入所 . 以来 , 画像処理ソフトウェア , 画像データベース , コンピュータビジョンの研究に従事 . 1986 ~ 87 マツギル大・知能機械研究センター客員教授 . 1992 奈良先端科学技術大学院大学・情報科学センター教授 . 現在 , 同大情報科学研究科教授 . 1990 , 2007 情報処理学会論文賞受賞 . 2005 情報処理学会フェロー . 工博 . 情報処理学会 , 日本バーチャルリアリティ学会 , 人工知能学会 , 日本認知科学会 , 映像情報メディア学会 , IEEE 各会員 .

**Abstract** Image inpainting techniques have been widely investigated to remove undesired visual objects in images such as damaged portions of photographs and people who have accidentally entered into pictures. Conventionally, the missing parts of an image are completed by optimizing the objective function which is defined based on pattern similarity between the missing region and the rest of the image (data region). However, unnatural textures are easily generated due to two factors: (1) available samples in the data region is quite limited, (2) pattern similarity is one of the required conditions but is not sufficient for reproducing natural textures. In this paper, in order to improve the image quality of completed texture, the objective function is extended by allowing brightness changes of sample textures (for (1)) and introducing spatial locality as an additional constraint (for (2)). The effectiveness of these extensions is successfully demonstrated by applying the proposed method to one hundred images and comparing the results with those obtained by the conventional method. In addition, the relationship between quantitative and qualitative evaluation for these resultant images is discussed.

**Key words** Image Inpainting, Image Completion, Energy Minimization, Pattern Similarity