

## 超解像技術による画像拡大の検討

厚木 七恵<sup>†1</sup> 永田 明徳<sup>†1</sup>

現在, 表示機器の解像度が飛躍的に向上したため, 大画面で表示すると画質が低下してしまうが, 拡大しても画質が粗くならない拡大技術である超解像技術が研究されてきた. 超解像技術とは, 元の画像の高周波成分に関する情報を復元することで, 画質の低下を防ぐ拡大技術である. 本研究では, 一定の速度で文字が移動し時間によって文字の位置や形も変わる映画のクレジットタイトルに絞り, 映像を複数枚の画像ととらえ, 画像から時間ごとに移動した同じ形の文字を複数個集めて画像にある同じ文字を重ね, 超解像処理を行った. 複数枚の超解像画像と従来の超解像画像の PSNR の値を比較し評価した結果, PSNR の値が高くなったため重複文字の超解像には効果があるとわかった.

### Image enlargement by super-resolving technology

NANAE ATSUGI<sup>†1</sup> and AKINORI NAGATA<sup>†1</sup>

In recent years, the resolution of display apparatus is improving by leaps and bounds. However, it is not easy for us to read the letter written there from the picture of the low resolution which has already existed. Then, the super-resolving technology in which image quality does not become coarse even if it expands the picture of low resolution has been studied. Super-resolving technology is expansion technology which restores the information about the high frequency ingredient of the original picture, and prevents deterioration of image quality. In this research, I focus on the letter of the credit title of the movie which moves at a fixed speed in the inside of a screen, and also changes a pattern by time due to resolution. We are collecting two or more same characters that presumed the position of the letter to move and moved, and piling them up, and performed super-resolving processing. I compared and evaluated the value of PSNR of the super-resolving picture of two or more sheets, and the conventional super-resolving picture. As a result, since the value of PSNR became high, it turned out that there is an effect in super-resolving of a duplicate letter.

### 1. はじめに

現在, ディスプレイなどの表示機器の解像度が飛躍的に向上したため, NTSC の字幕やテロップなどの文字を大画面で表示すると拡大されてしまい, 画質が低下してしまう. また, 監視カメラなどの性能や記憶容量に制限のある機器では, 低解像度の動画像しか保存することができないことが多く部分的に拡大すると低下した画質が目立ってしまう. そこで近年, 拡大しても画質が粗くならない拡大技術である超解像技術が研究されてきた.

超解像技術とは, 画像に隠されている折り返しひずみという, 元の画像の高周波成分に関する情報を復元することで, 画質の低下を防ぐ拡大技術である<sup>1)</sup>. 超解像処理は, 高解像度化 (Image Registration) と高画質化 (Deconvolution) という2つのアプローチがある<sup>2)</sup>. 高解像度化とは, 変形, 移動する被写体を捉えた複数枚の画像を重ね合わせることで, 1枚の高解像度画像を作成する技術である. 被写体の変形, 移動を検出するにはオプティカルフローを用いた Lucas-Kanade 法やブロックマッチング法などの手法がある. 時間の変化により, 変形, 移動する被写体を捉えた複数枚の画像を重ね合わせることでピクセルの数が増し, 解像度を高くすることができる.

複数枚の画像を重ね合わせ高解像度の画像を作成するには再構成型超解像度を用いる. 再構成型超解像度では, カメラの撮影モデルにしたがって高解像度の画像を復元する. 撮影モデルとは被写体とカメラのレンズや検出素子などを介して記録された画像の情報量の差を数学的にモデル化したものである. 動画像のフレーム間の動きをサブピクセル精度で検出し, 動きに合わせて前後から参照したフレームの画素値を処理フレームの標本値として追加する. その後高解像度画像の画素値を算出する<sup>3)</sup>. 高解像度画像の画素値の算出は再構成処理と呼ばれ, 主な手法として, Back-Projection 法<sup>4), 5)</sup>, MAP 法<sup>6), 7)</sup>, POCS 法<sup>8)</sup>などが挙げられる.

また高画質化とは, ピンぼけ, 手ぶれ, 被写体ぶれ, ノイズなどによって生じた画質の劣化を推定し, 元の画像を復元することである. これには Lucy-Richaedson 法 (LR 法) や Iterative Back Projection 法 (IBP 法), Bilateral Total Variation 法 (BTV 法) などがある<sup>2)</sup>. しかし LR 法と IBP 法ではノイズも忠実に再現してしまうため, ガウシアンフィルタを用いてノイズを除去することができる BTV 法が使われる事が多い. BTV 法では元の画像ベクトルを導くために, 画質劣化の原因のピンぼけ, 手ぶれ, 被写体ぶれ, ノイズについ

<sup>†1</sup> 東京工科大学大学院 バイオ・情報メディア研究科 Graduate School of Tokyo University of Technology

て、ガウシアンフィルタを適用し画像の復元を行う。また、エッジ保持平滑化を行うことで、画素間の距離ではなく輝度の差を見て重みを付けることが可能になる。そのため、画像をスムーズにしつつもエッジ部分を量すことなく処理を行うことができる。

しかし、既存手法では人物やモノなどの、一般的なオブジェクトに対しては効果があるが、動画に含まれる文字に対しては、効果があまり見られないという問題がある。人物やモノなどの一般的なオブジェクトに比べ、文字には高周波成分が多く含まれているため、輪郭周りにノイズが発生しやすく、既存手法で超解像度化しても画質が向上したように見えないためである。

本研究では、文字に対しては新しいアプローチである重複文字の超解像処理を行うことを提案する。また、文字以外の部分には既存手法を適用することで、文字に対しても効果のある超解像度化、画像拡大を行っていく。

まず2章で既存研究について3章で提案手法について述べ、4章で検証を行う。最後に、5章で考察を行う。

## 2. 超解像処理

この章では超解像処理について説明する。

### 2.1 高解像度化 (Image Registration)

高解像度化とは、変形、移動する被写体を捉えた複数枚の画像を重ね合わせることで、1枚の高解像度画像を作成することである。図1のように時刻の変化 ( $t \sim t + \Delta t$ ) により、変形、移動する被写体を捉えた複数枚の画像を重ね合わせることで、1枚の高解像度な画像を作成する。複数枚の画像を重ね合わせることで、ピクセルの数が増すので解像度が高くなる。

### 2.2 再構成型超解像度

複数枚の画像を重ね合わせ高解像度の画像を作成するには再構成型超解像度を用いる。再構成型超解像では、カメラの撮影モデルにしたがって高解像度の画像を復元する。撮影モデルとは、被写体とカメラのレンズや検出素子などを介して記録された画像の情報量の差を、数学的にモデル化したものである。撮影モデルとして、入力された低解像度画素の輝度値  $y_j (j = 1, \dots, M)$  を、対応する高解像度画像上の各画素  $x_k (k = 1, \dots, N)$  の加重平均で表す<sup>3)</sup>。このモデルは低解像度とその位置や画像のぼけ方によって決定される重み  $w_{jk}$  を高解像度がその重みとした式1で表せる。

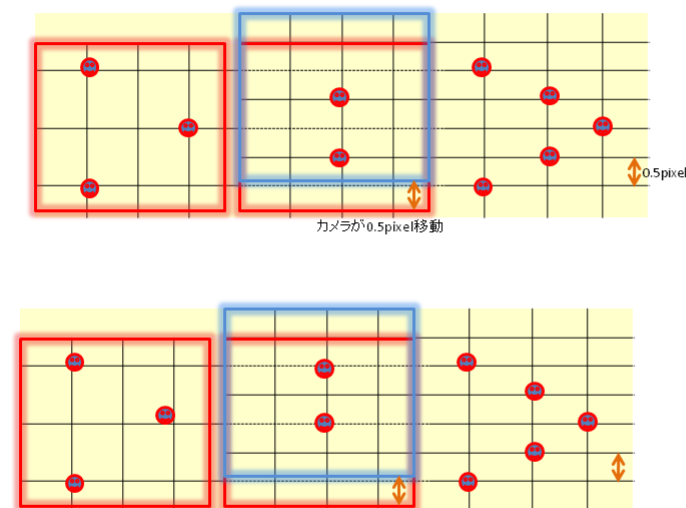


図1 時刻の変化により、変形、移動する被写体  
Fig. 1 Subject Transformation and moves by change in time

$$y_j = \sum_k w_{jk} x_k (j = 1, \dots, M) \quad (1)$$

式1は  $x_k$  を求めるための線形連立方程式を与える。  $m < n$  となるためこの線形連立方程式の解は1つに定まらない。線形連立方程式を求めるのは下記の方法を用いる。式2を使い  $y_i$  の式に対する誤差  $e_j$  を求める。

$$e_j = y_j - \sum_k w_{jk} x_k \quad (2)$$

誤差  $e_j$  の値が出来るだけ小さくなるように、式 3 を用いて  $x_k$  を更新する。

$$x_k \leftarrow x_k - \beta w_{jk} e_j \quad (3)$$

式 2, 3 を数回から数十回繰り返す。

この方法で得られた高解像度の画像は、期待する高解像度と比べぼやけた画像になってしまうことが多い。これは線形連立方程式の数が足りない ( $m < n$ ) 為である。線形連立方程式を追加する ( $m$  の数を増やす) には、別の位置の輝度値が必要になる。別の位置の輝度値を調べるには他の方法を用いる。輝度値を調べるには、まず被写体のオプティカルフローを捉える必要がある。被写体のベクトルを捉えることの出来る LK 法を用い、被写体のオプティカルフローを求める。LK 法は同一物体の局所領域内ではオプティカルフローが一定になると仮定する空間的局所最適化法の 1 つであり、画面座標  $(x, y)$  及び時刻  $t$  における注目画素の濃度を  $I(x, y)$  とすると、局所領域  $w$  におけるオプティカルフロー  $(u, v)$  は式 4, 5 で求められる。しかし、LK 法で求めたオプティカルフローはあまり正確ではない。そのため、対応する輝度値の平均二乗誤差を放物線の関数で近似し、その関数値が最小となる位置に修正する。

$$u = \frac{\sum_w \frac{\partial I}{\partial x} \cdot [(J(P) - I(P))]}{\sum_w (\frac{\partial I}{\partial x})^2}, v = \frac{\sum_w \frac{\partial I}{\partial y} \cdot [(J(P) - I(P))]}{\sum_w (\frac{\partial I}{\partial y})^2} \quad (4)$$

$$I(P) = I(x, y, t), J(P) = I(x, y, t + \delta t) \quad (5)$$

オプティカルフローを求めたら、画像のフレーム間の対応付けを行い対象フレームに参照フレームの画素をコピーする。これにより違う領域の画素を利用出来るようになり (M が N に近づく) 解像度画像がより正確に復元される。

### 2.2.1 高画質化 (Deconvolution)

高画質化とは、図 2 のように元の画像  $X$  にピンぼけ、手ぶれ、被写体ぶれ、ノイズなどの何らかの劣化フィルタ  $H$  がかったために、観測される画像  $Y$  が劣化してしまう。劣化した画像  $Y$  から、元の画像  $X$  を推測し、ピンぼけ、手ぶれ、被写体ぶれ、ノイズなどによって生じた画質の劣化を復元することである<sup>2)</sup>。

元の画像  $X$  は式 6 から求めることが出来る<sup>2)</sup>。  $H$  が既知の場合、  $H$  と観測値  $Y$  から  $X$  を推定し、  $H$  が未知の場合、観測値  $Y$  のみから  $H$  と  $X$  を推定する。画像がどのように劣

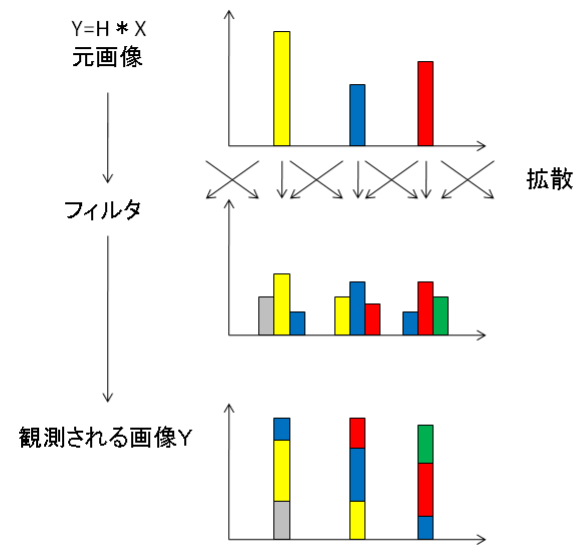


図 2 画像の劣化  
Fig.2 Degration of image

化していたかを推測できる場合は元画像が失われているため殆どない。その為、  $H$  は未知の場合がほとんどである。

$$Y = H * X \quad (6)$$

$Y$ : 劣化した画像ベクトル (既知)

$H$ : 劣化フィルタ (既知もしくは未知)

$X$ : 元の画像ベクトル (未知)

### 2.3 PSNR(Peak Signal-to-Noise Ratio)

PSNR とは Peak Signal-to-Noise Ratio(ピーク信号対雑音比) の略で、単位はデジベル(dB) で表せられる。PSNR は信号の理論ピーク値と誤差の 2 乗平均を用いて評価しており、値 255(最大濃淡値) を誤差の標準偏差で割った値である。dB 値が高いほど拡大した画像が元画像に近いことを表す。一般的には 40dB 以上あれば綺麗と言われている。PSNR は式 7 で求められる。

$$PSNR = 10 \log_{10} \frac{255^2 \cdot w \cdot h}{\sum_{x=0}^{w-1} \sum_{y=0}^{h-1} (p_1(x, y) - p_2(x, y))^2} \quad (7)$$

$w$  : 画像の幅

$h$  : 画像の高さ

$p_1$  : 元画像

$p_2$  : 拡大画像

### 3. 提案手法

重複文字の超解像では、動画などで流れる映像を複数枚の画像にとらえ、画像から時間ごとに移動した同じ形の文字を複数個集めていき、画像にある同じ文字を重ね、超解像処理を行う。画像はところどころ劣化しており、複数枚の画像にある文字も例外なく、別の場所が劣化していることが多い。特に先に論じたように文字は高周波成分が多いため、輪郭にノイズが発生し、文字がぼやけてしまう。ここでいう劣化は高周波成分が失われていることにより、グレースケールの部分がぼやけてしまっていることである。今回の手法を行うことでぼやけた文字が結果的に、重ねられた文字は多ければ多いほど強調され、流麗な文字となると思われる。

今回は映画のクレジットタイトルに限定する。また、クレジットタイトルは黒の背景に白文字であるものであり、キャストやスタッフ、企業、団体などの文字が下から上に流れていくものを利用する。クレジットタイトルは画面の下から上へ一定の速度で文字が移動するため、時間によって文字の位置やノイズの入り方が変化し、低解像であるほどパターンの変化が大きくなる。そこで、移動した文字をキャプチャし、そのキャプチャした複数枚の画像の文字を重ね超解像処理を行う。複数枚の超解像画像と従来の超解像画像の PSNR の値を比較し、複数枚の超解像画像の PSNR の値が高くなったら効果があるといえる。

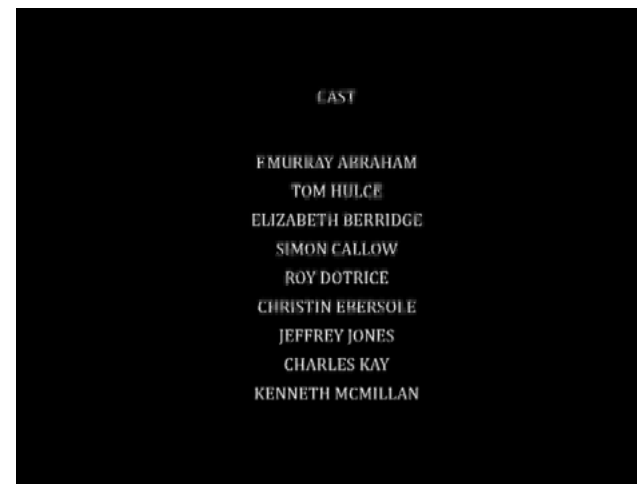


図 3 480 × 360 のクレジット動画  
Fig. 3 480 × 360 Credit movie

### 4. 検証

提案手法の有用性の検証として、予備実験を行った。提案手法では画像を重ねる際、輝度値や画素と画素との対応付けを行う位置合わせ処理等は計算して決めるが、今回は簡易的に手動で位置を合わせ、画像は加重平均として重ねる。

#### 4.1 入力画像の用意

黒が背景で文字が白で画面の下から上へ一定の速度で文字が移動する 480 × 360 のクレジット動画を用意する。この動画を 200 × 150 に低解像化させた動画も用意する。動画の画像を図 3、図 4 に示す。

#### 4.2 文字部分の抽出

文字部分の抽出は、画面の一番下に白い部分が表示され、文字が一定の距離を流れたらキャプチャする。キャプチャした画像の例を図 5 に示す。

#### 4.3 複数枚の画像の超解像

複数枚の画像を用いて超解像を行った.. それぞれ画像の超解像処理結果を図 6 に示す。



図 4 200 × 150 のクレジット動画  
Fig.4 200 × 150 Credit movie



図 5 キャプチャ画像  
Fig.5 Capture image



図 6 複数枚重ねた画像  
Fig.6 Picture of two or more sheets

#### 4.4 評 価

PSNR による評価結果を表 1 に示す。従来の超解像処理の手法はバイキュービック法を用いた。劣化画像、従来の手法、提案手法の中で PSNR の値が一番高いのは提案手法で、PSNR の値は 17.35dB であった。

表 1 検証結果

Table 1 Verification outcome

劣化画像	バイキュービック法	提案手法
16.31	16.40	17.35

## 5. 考 察

検証結果より複数枚の超解像画像と従来の超解像画像の PSNR の値を比較し PSNR の値が高くなったため、重複文字の超解像には効果があるとわかった。バイキュービック法では画像全体を滑らかにするため、文字のような高周波成分が多く含まれた物の超解像ではあまり効果が見られないと考えられる。PSNR の値の差は従来の手法と比べて約 1.0dB 向上しているが、枚数を増やしたり閾値を決めて白はより強く、グレーの部分はより黒くさせることが出来ればさらに PSNR の値が向上すると考えられる。また、文字の位置が合っていないと解像度を向上させることが出来ないため高精度な位置合わせを行うべきだといえる。少なくとも生成されるべき高解像度画像の画素単位の精度は必要である。次の課題として、

## 参 考 文 献

- 1) S. C. Park, M. K. Park, and M. G. King, "Super-Resolution. Image Reconstruction: A Technical Overview," IEEE Signal processing Magazine, pp.21-36, May 2003
- 2) 篠原 修二 : 入門超解像処理アルゴリズム,  
[http://www.aas-ri.co.jp/news/Seminar\\_superResolution.pdf](http://www.aas-ri.co.jp/news/Seminar_superResolution.pdf)
- 3) 松本信幸, 井田孝: 画像のエッジ部の自己合同性を利用した再構成型超解像, 電子情報通信学会 D Vol. J93-D No.2 pp.118-126(2010)
- 4) T. Komatsu, T. Igarashi, K. Aizawa, and T. Saito, "Very high resolution imaging scheme with multiple different-aperture cameras," Signal Process., Image Commun., vol.5, pp.511-526, 1993.
- 5) M. Irani and S. Peleg, "Improving resolution by image registration," CVGIP: Graphical Models and Image Processing, vol.53, no.3, pp.231-239, 1991.
- 6) R.R. Schulz and R.L. Stevenson, "Extraction of high-resolution frames from video sequences," IEEE Trans. Image Process., vol.5, no.6, pp.996-1011, 1996.
- 7) 田中正行, 奥富正敏, "再構成型超解像処理の高速化アルゴリズムとその精度評価," 信学論 (D-II), vol. J88-D-II, no.11, pp.2200-2209, Nov. 2005.
- 8) A.J. Patti, M.I. Sezan, and A.M. Tekalp, "Super-resolution video reconstruction with arbitrary sampling lattices and nonzero aperture time," IEEE Trans. Image

Process., vol.6 no.8, pp.10664-1076, 1997.

- 9) 竹島秀則, 加藤宣弘, 金子敏充: サブピクセルシフト画像を用いた複数フレーム超解像技術, 東芝レビュー vol.64, No.6(2009)