

長岡技術科学大学大学院
工学研究科修士論文

題 目

BTV 正則化に基づく魚眼画像の
マルチフレーム超解像に関する研究

指導教員

准教授 杉田 泰則

著 者

電気電子情報工学専攻
信号処理応用研究室
19503291 丸田 裕輝

令和 3 年 2 月 9 日

ABSTRACT

A Study on Multi-frame Super-Resolution of Fisheye Images Based on BTV Regularization

Author : 19503291 Hiroki MARUTA
Supervisor : Yasunori SUGITA

In recent years, information terminal devices equipped with cameras such as smart-phones have been rapidly developing. Especially, fisheye cameras which have the wide angle of field, have been developed in particular. However, due to the characteristics of fisheye cameras, the image quality is not high. Therefore, there has been active research about super-resolution of fisheye images. There are two main types of super-resolution methods. The single-frame type is super-resolution which an input image is a single image, but the usable information is small, the most parts of high-frequency components have to be estimated. On the other hand, the multi-frame type uses multiple images, so it can use more information and achieve high accurate super-resolution. However, there is a problem that the fisheye image has a large distortion and blur from the center region to the peripheral region, and the image quality is deteriorated by ones. Therefore, super-resolution which a part of the central region of the fisheye image is often used, but super-resolution for the peripheral region or the entire fisheye image has rarely been considered.

In this paper, we propose a method (Method 1) that uses image segmentation to improve the image quality of peripheral regions and the entire fisheye image by using BTV regularization based super-resolution which one of a method of multi-frame super-resolution. Method 1, the input image is divided into multiple regions, super-resolution is performed for each region, and multiple divided images are combined to create a single super-resolution image. We thought that this method was effective super-resolution to different distortions and blurring. In addition to Method 1, we proposed a method to limit the input images by similarity (Method 2). This method calculates the similarity between the reference frame image and other frame images in the input frame image group, and uses input images which only those frame images of exceeded a threshold value for multi-frame super-resolution.

In order to confirm the effectiveness of the proposed method, we compared the super-resolution results of the conventional method and the two proposed methods. Method 1 performed super-resolution by dividing the image into 5 and 9 segments, and the results of two proposed methods were almost the same as the conventional method in static video.

However, it was confirmed that the image quality was lower than that of the conventional method in the video with movement. Method 2 improved the image quality by limiting the input images which was thought to have a bad influence, and the image quality was equal to or little lower than that of the conventional method. But it didn't achieve the purpose which improving the image quality of the entire fisheye image. We couldn't achieve our purpose of improving the entire image quality of the fisheye image. However, we also confirmed that Method 2 requires less computation and less time for super-resolution processing than the conventional method. These results confirm the effectiveness of the proposed method for super-resolution of fisheye images.

目次

| | | |
|-------|-----------------|----|
| 第 1 章 | はじめに | 1 |
| 1.1 | 研究背景 | 1 |
| 1.2 | 研究目的 | 2 |
| 1.3 | 本論文の構成 | 2 |
| 第 2 章 | 超解像の基礎知識 | 3 |
| 2.1 | 超解像 | 3 |
| 2.2 | シングルフレーム超解像 | 4 |
| 2.3 | マルチフレーム超解像 | 4 |
| 2.3.1 | 画像モデル | 4 |
| 2.3.2 | MAP 法に基づく超解像手法 | 5 |
| 第 3 章 | 従来の超解像手法 | 7 |
| 3.1 | 魚眼画像の特徴 | 7 |
| 3.2 | 歪み補正 | 7 |
| 3.3 | 位置合わせ | 8 |
| 3.3.1 | テンプレートマッチング | 9 |
| 3.3.2 | SURF 特徴量による台形補正 | 10 |
| 3.4 | BTV 正則化に基づく超解像 | 11 |
| 3.5 | 問題点 | 11 |
| 第 4 章 | 提案手法 | 12 |
| 4.1 | 画像分割による超解像 | 12 |
| 4.1.1 | 画像分割 (5 分割) | 12 |
| 4.1.2 | 画像分割 (9 分割) | 12 |
| 4.1.3 | 画像合成 | 13 |
| 4.2 | 類似度による入力画像の制限 | 14 |

| | | |
|-------|----------------------------|----|
| 4.2.1 | 類似度が超解像に及ぼす影響の調査 | 14 |
| 第 5 章 | シミュレーション | 16 |
| 5.1 | シミュレーション条件 | 16 |
| 5.2 | 歪み補正 | 17 |
| 5.3 | シミュレーション結果 | 18 |
| 5.3.1 | 提案手法 1 | 18 |
| 5.3.2 | 提案手法 2 | 18 |
| 5.4 | 考察 | 22 |
| 5.4.1 | 提案手法 1 の考察 | 23 |
| 5.4.2 | 提案手法 2 の考察 | 23 |
| 5.4.3 | 処理時間の比較 | 28 |
| 第 6 章 | 結論 | 29 |
| 6.1 | まとめ | 29 |
| 6.2 | 今後の課題 | 30 |
| 謝辞 | | 31 |
| 参考文献 | | 32 |

第 1 章

はじめに

本章では，本論文の研究背景及び目的，本論文の構成について述べる．

1.1 研究背景

画像や映像は目で見た景色を撮影してデータとして保存することで，いつでも何度でも景色を視聴することができるため，日常生活で欠かせないものとなっている．特に画像・映像の圧縮技術やスマートフォンなどのカメラを搭載した情報端末が急速に発展したことから，近年では手持ちの情報端末で撮影して，その場で画像・映像を確認・共有することができるようになった．このような背景から画像処理技術に対する需要は拡大している．

景色などの撮影に使われる機器の一つに魚眼カメラがある．近年では，この魚眼カメラを搭載した監視カメラやドライブレコーダーなどの様々な製品が増加傾向にある．魚眼カメラは魚の目を模したレンズが使われており，通常のカメラと比較して，より広範囲を撮影でき，通常のカメラが複数台で撮影できる範囲を魚眼カメラ 1 台でカバーできる点である．しかし，レンズ歪み，ぼけなどの影響により，通常のカメラよりも画質が低いという問題がある．そのため，画質向上を目的として超解像を応用した研究 [1] が行われている．

画像処理技術の分野を担う技術の 1 つに超解像がある．超解像とは，エッジなどの高周波成分が失われた観測画像の高周波成分を復元する技術である．この技術は，低コストで画素数の少ないカメラで撮影された画像の高画質化，4K や 8K ディスプレイに表示するための高画質化などへの応用として期待されている．この超解像は大まかに 2 種類の手法に分類される．1 つ目が，1 枚の観測画像に対して，超解像を行うシングルフレーム型 [5] と呼ばれるものである．2 つ目が，複数枚の観測画像を用いて，各観測画像で不足している画素を補間する再構成型，またはマルチフレーム型 [3] と

も呼ばれるものである。シングルフレーム型は観測画像の失われた高周波成分を学習やデータベース、非線形関数などから推測するものであるため、正しい高周波成分が復元されるか分からない。それに対して、マルチフレーム型は基準の観測画像から失われた高周波成分を他の観測画像の情報で補って求めるため、シングルフレーム型よりも精度が高く高周波成分を復元することができる。そのため、本論文ではマルチフレーム型に基づく超解像に焦点を当てて行う。

1.2 研究目的

現在までに魚眼画像に対してマルチフレーム超解像を適応した研究は少ない。これは急速に魚眼カメラが発展したこと、魚眼画像特有の歪みによってマルチフレーム型で用いる位置合わせ処理が困難であることが原因である。実際に先行研究 [9] では、魚眼画像の中央領域付近のみに着目した超解像が検討されているが、歪みの大きい周辺領域の超解像については検討されていない。そこで本研究では周辺領域を含めた魚眼画像全体の超解像を提案する。提案法では観測された魚眼画像が中央領域と周辺領域で歪み、ぼけが異なることに着目して、観測画像の各領域に分割する。そして、分割画像ごとに BTV 正則化に基づくマルチフレーム超解像を行い、各分割画像を 1 枚の画像に合成する。最後に従来法とのシミュレーション結果を比較し、提案法の有効性を明らかにする。

1.3 本論文の構成

本論文の構成は以下の通りである。第 1 章では、本論文の研究背景と目的を述べた。第 2 章では、マルチフレーム超解像の基礎技術について述べる。第 3 章では、従来の超解像手法の各処理について紹介し、その手法が抱える問題点について述べる。第 4 章では、提案手法の構成を示し、従来法との違いを述べた。第 5 章では、シミュレーションに必要な条件について述べ、魚眼画像を対象とした超解像の評価を行う。また、従来法と提案手法の比較を行い、提案手法の有効性を示す。第 6 章では、これまで述べた内容をまとめ、結論を述べる。

第 2 章

超解像の基礎知識

2.1 超解像

超解像 (Super Resolution) は、一般的に低解像度画像（以下，LR 画像とする）から高解像度画像（以下，HR 画像とする）に変換する技術だと思われているが，実際には観測画像から失われたエッジなどの高周波成分を復元する技術のことである．例えば，図 2.1 のような高周波成分が少ない画像があったときに，超解像をすることで図 2.2 のように高周波成分が復元され，その結果エッジが明確になる．このような超解像技術は主に撮影環境が悪くノイズが多く発生するような撮影，テレビのようなリアルタイムで高画質の映像を提供する必要があるが通信環境の都合で低画質の映像しか送れない場合などに使用されることが多い．例えば，ブラックホールの観測 [4]，低解像度映像を高解像度映像に変換するテレビやスマホなどの応用例がある．



図 2.1 高周波成分が少ない画像



図 2.2 推定画像

1.1 節に記述しているが，超解像は主に 2 種類の手法に分類することができる．1 つ目が 1 枚の画像を用いて超解像を行うシングルフレーム超解像，2 つ目が複数枚の画像を組み合わせるマルチフレーム超解像がある．各手法の詳細については，2.2，2.3 節で説明する．

2.2 シングルフレーム超解像

シングルフレーム超解像は、1枚の観測画像を用いて1枚の超解像画像を生成する手法の総称である。マルチフレーム超解像に比べて、観測画像が1枚だけであるため、画像、映像のどちらでも適用することができる。しかし、超解像に使用できる情報が観測画像1枚のみであるため、マルチフレーム超解像よりも精度が悪い。このシングルフレーム超解像には様々な種類があり、辞書型 [5]、非線形型 [6]、フラクタル型 [7]、ディープラーニング型 [8] など含まれている。辞書型とは、事前に LR 画像と HR 画像の同じ部分を切り抜いてパッチ（辞書）を作成しておき、LR 画像を探索し辞書に登録されているパターンがあった場合に、その対応関係から HR 画像のパッチへと置き換える手法である。非線形型は LR 画像にハイパスフィルタをかけて高周波成分を取り出し、その結果を非線形関数を用いて強調して LR 画像に加算することで高周波成分を復元する手法である。フラクタル型はフラクタル次元という形状の複雑さを表す次元を用いて、画像を拡大する際のフラクタル次元ルールを推定・設定することでその拡大ルールに従う超解像画像を生成する手法である。ディープラーニング型は SRCNN(Super-Resolution Convolutional Neural Network) を用いて、入力を低解像度画像、出力を高解像度画像を与えて学習させる手法である。

2.3 マルチフレーム超解像

マルチフレーム超解像は、複数枚の観測画像を用いて、1枚の超解像画像を生成する手法の総称であり、再構成型とも呼ばれている。超解像に使用できる情報が複数枚の観測画像であり、シングルフレーム超解像よりも情報量が多いため、より精度が高いことが特徴である。しかし、計算量が多くリアルタイム性に欠けることが問題点となっている。この手法では、同じ対象物・景色をサブピクセルだけずれた位置から撮影した複数枚の観測画像が必要になる。その観測画像のうち1枚を基準画像として、その画像と他の観測画像の位置合わせを行い超解像を行う。

2.3.1 画像モデル

一般的なマルチフレーム超解像の画像モデルを式 (2.1)、図 2.3 に示す。

$$\mathbf{Y}_i = \mathbf{D}\mathbf{B}_i\mathbf{M}_i\mathbf{X} + \mathbf{n}_i \quad \text{for } 1 \leq i \leq p \quad (2.1)$$

\mathbf{Y}_i は観測画像 ($1 \times N_1$ ベクトル)、 \mathbf{D} はダウンサンプリング行列、 \mathbf{B}_i はボケを表す行列、 \mathbf{M}_i は動きを表す行列、 \mathbf{X} は推定画像 ($1 \times N_2$ ベクトル)、 \mathbf{n}_i はノイズ、 N_1 は観測

画像の画素数, N_2 は推定画像の画素数, p は入力画像数である. また, **DBM** をまとめて劣化行列 **A** として表されることもある. 式 (2.1) は, 劣化行列 **A** に推定画像のベクトル **X** をかけることで, 観測画像ベクトル **Y** になるという意味である. また, **Y**, **X** のような画像をベクトルに変換する方法は図 2.4 のように, 画像の 1 行目から順にピクセルを並べて配置している. カラー画像の場合は **RGB** の 3 チャンネルそれぞれに対して, 同様のベクトルを作成する.

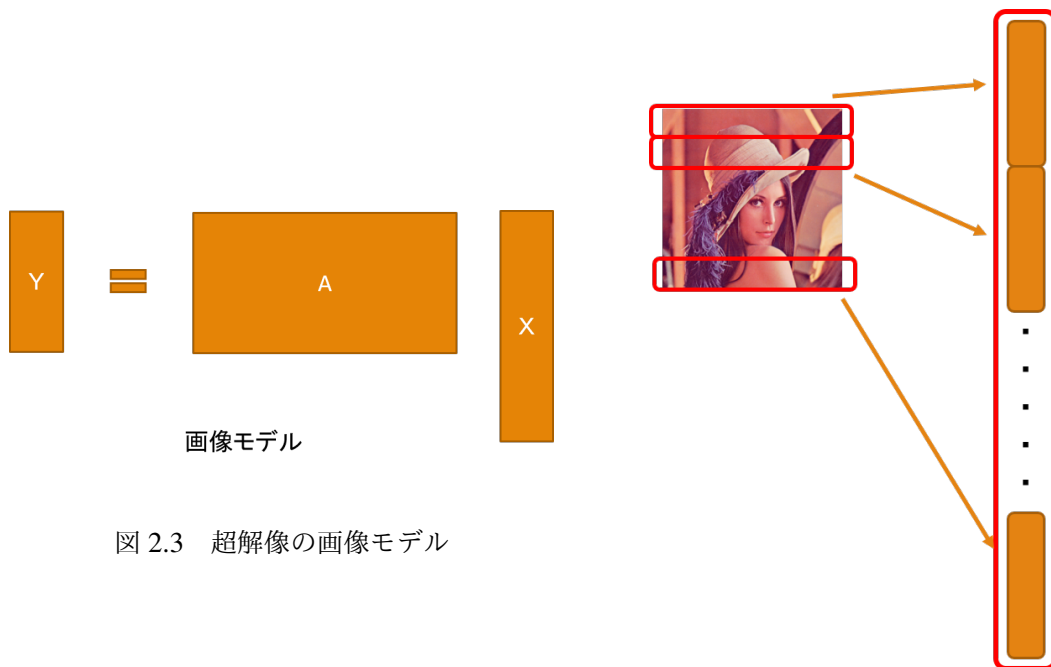


図 2.3 超解像の画像モデル

図 2.4 画像からベクトルへの変換法

2.3.2 MAP 法に基づく超解像手法

マルチフレーム超解像で頻繁に用いられる手法にベイズ推定に基づく事後確率の最大化問題を解いて再構成を行う MAP(Maximum A Posteriori) 法と呼ばれる手法がある. この最大化問題は,

$$\max_{\mathbf{X}} P(\mathbf{X}|\mathbf{Y}_1, \dots, \mathbf{Y}_N) \quad (2.2)$$

と表すことができる. この時, \mathbf{Y}_i は観測画像, \mathbf{X} は推定画像, N は入力画像数, $P(\mathbf{X}|\mathbf{Y}_1, \dots, \mathbf{Y}_N)$ は推定画像 \mathbf{X} に対する事後確率分布を表している. この事後確率はベイズの定理より,

$$P(\mathbf{X}|\mathbf{Y}_1, \dots, \mathbf{Y}_N) = \frac{\sum_{\mathbf{X}}^{i=1} P(\mathbf{Y}_i|\mathbf{X})P(\mathbf{X})}{P(\mathbf{Y})} \quad (2.3)$$

と表すことができ、対数を取ると、

$$\log P(\mathbf{X}|\mathbf{Y}_1, \dots, \mathbf{Y}_N) = \log \sum_{i=1}^N P(\mathbf{Y}_i|\mathbf{X}) + \log P(\mathbf{X}) - \log P(\mathbf{Y}) \quad (2.4)$$

となる。この時、 \mathbf{X} についての最大化問題を解いているため、 \mathbf{X} に依存していない第3項を無視すると、

$$\max_{\mathbf{X}} \left[\log \sum_{i=1}^N P(\mathbf{Y}_i|\mathbf{X}) + \log P(\mathbf{X}) \right] \quad (2.5)$$

と、式 (2.2) を書き換えることができる。このとき、式 (2.5) の $P(\mathbf{Y}_i|\mathbf{X})$ は尤度分布、 $P(\mathbf{X})$ は \mathbf{X} に対する事前確率分布である。また、MAP 法の典型的な尤度分布および事前確率分布を以下に示す。

$$P(\mathbf{Y}_i|\mathbf{X}) \propto \exp \left[-\frac{\|\mathbf{Y}_i - \mathbf{A}_i \mathbf{X}\|^2}{2\sigma^2} \right] \quad (2.6)$$

$$P(\mathbf{X}) \propto \exp \left[-\frac{\|\mathbf{H}\mathbf{X}\|^2}{2\lambda^2} \right] \quad (2.7)$$

\mathbf{A}_i は劣化行列、 σ^2 はノイズの分散、 λ^2 は事前確率分布のパラメータ、 \mathbf{H} はハイパスフィルタを表す行列である。この式 (2.6), (2.7) を用いて、式 (2.5) の最大化問題を最小化問題に変換すると、

$$E(\mathbf{X}) = \min_{\mathbf{X}} \sum_i \|\mathbf{Y}_i - \mathbf{A}_i \mathbf{X}\|_2^2 + \alpha \|\mathbf{H}\mathbf{X}\|_2^2 \quad (2.8)$$

となる。 α は拘束パラメータで $\alpha = \sigma^2/\lambda^2$ の関係であり、 $E(\mathbf{X})$ は MAP 法の評価関数であるため、この評価関数の最小化問題として解くことで推定画像 \mathbf{X} を求めることができる。

第 3 章

従来の超解像手法

魚眼画像は通常レンズとは異なる歪みの大きい魚眼レンズで撮影するため、中央部から周辺部に行くほど歪みが大きくなり、2.3.2 節の手法では超解像を行うことが不可能である。そのため、歪みを考慮した魚眼画像用の超解像を検討する必要がある。本章では、魚眼画像のための BTV 正則化に基づく超解像の従来法 [9] について説明する。また、この従来法が抱える問題点を考察し、改善する手法を検討する。

3.1 魚眼画像の特徴

魚眼画像とは、魚眼レンズを取り付けたカメラで撮影された画像（図 3.2）のことである。魚眼画像を用いる理由は通常画像に比べて広範囲を撮影することができるためである。通常のカメラであれば複数台でカバーする必要のある視野を魚眼カメラなら、より少ない台数でカバーすることができる。これに対して、欠点は広範囲を撮影するためにレンズが歪ませているため、魚眼画像そのままでは画像処理を行うには向いていないことである。そのため、魚眼画像を用いる際は事前に歪みを補正する必要がある。

3.2 歪み補正

通常カメラで撮影された画像や図 3.2 のような特殊なレンズを用いたカメラで撮影された画像は、そのカメラで用いられているレンズ特有の歪みが画像に載っている。超解像を行う際に歪みが残っていると、対象物が同じ位置、形状で画像に映らない場合があり、位置合わせができない可能性がある。そのため、図 3.2 のような歪みのある画像に対して歪み補正処理を行い、図 3.3 の歪みのない画像に変換する。

図 3.1 に示す魚眼レンズを単純化したモデルで歪み補正のパラメータを計算する。

光軸ずれがない場合，撮像面上の点 $p(u, v)$ とレンズ面上の点 $P(x, y, z)$ との間には次式の関係が成り立つ．

$$P = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} u \\ v \\ f(u, v) \end{bmatrix} \quad (3.1)$$

f は撮像面中心からの距離 $\rho = \sqrt{u^2 + v^2}$ であるため， $f(u, v) = f(\rho)$ と書ける．さらに f を多項式で近似すると式 (3.2) になる．

$$f(\rho) = a_0 + a_1\rho + a_2\rho^2 + \cdots + a_n\rho^n \quad (3.2)$$

撮像面と光軸にはずれが発生するため，理想の点 (u, v) ，実際の点 (u', v') より式 (3.3) のモデルを採用した．

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} c & d \\ e & 1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \end{bmatrix} \quad (3.3)$$

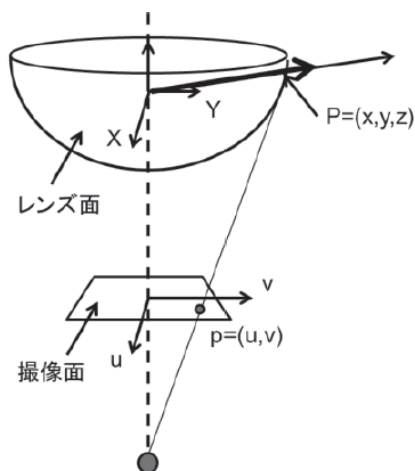


図 3.1 魚眼カメラに近似した球面モデル

以上より， $a_0, a_1, a_2, \dots, a_N, c, d, e, c_x, c_y$ を計算し歪みを補正する．これらのパラメータは図 3.2 のようにチェスボードを複数枚撮影し，格子模様の交点座標から歪みを推定することで算出する．

3.3 位置合わせ

マルチフレーム超解像は再構成処理を行う前に入力画像群（従来法では 16 枚が最適）の位置をサブピクセル精度で合わせる必要がある．そのため従来法では，位置合



図 3.2 魚眼画像



図 3.3 歪み補正結果

わせの手法にテンプレートマッチングと SURF 特徴量 [10] による台形補正が用いられている。

3.3.1 テンプレートマッチング

テンプレートマッチング（正規化相互相関）とは，テンプレート画像に一致する領域を探索する手法であり，今回は平行移動に対する位置合わせを行うために用いる．テンプレートマッチングには以下の式 (3.4) を用いる．

$$R = \frac{\sum_i \sum_j \mathbf{I}(i, j) \mathbf{T}(i, j)}{\sqrt{\sum_i \sum_j \mathbf{I}(i, j)^2} \sqrt{\sum_i \sum_j \mathbf{T}(i, j)^2}} \quad (3.4)$$

このとき， R は類似度， $\mathbf{T}(i, j)$ はテンプレート画像の輝度値， $\mathbf{I}(i, j)$ は対象画像の輝度値である．この類似度 R が 1 に近いほど似ている領域となる．テンプレート画像で図 3.5 の対象画像上を探索し，類似度が最大になった位置を求めて，中心まで平行移動することで位置合わせを行う．例えば，図 3.4 は図 3.3 の中心を切り抜いてテンプレート画像としたとき，類似度が最大となる位置が図 3.3 よりも図 3.5 が上にずれているため，対象画像を上方にずれている分だけ下方に平行移動して図 3.6 のように位置合わせを行う．



図 3.4 テンプレート画像



図 3.5 対象画像

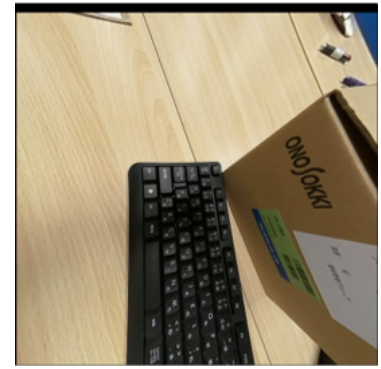


図 3.6 テンプレートマッチング後

3.3.2 SURF 特徴量による台形補正

拡大縮小や回転に対してロバストな特徴量である SURF 特徴量を検出し，基準画像（図 3.7 左）と対象画像（図 3.7 右）のマッチングを行い，ホモグラフィ行列を求めて台形補正を行う．例えば，図 3.7 は SURF 特徴量を検出し，マッチングを行った画像である．このとき，対象画像の緑線内の領域は基準画像全体に相当しており，マッチングしている点の中から RANSAC アルゴリズム [11] で 4 点選び，ホモグラフィ行列を求めてホモグラフィ変換を行うことで擬似的に対象画像を引き延ばすことができ，同じ解像度，領域の画像を生成することができる．これはテンプレートマッチングでは対応できない回転・拡大縮小に対しての位置合わせとして用いられている．



図 3.7 SURF 特徴量によるマッチングと台形補正

3.4 BTV 正則化に基づく超解像

BTV 正則化に基づくマルチフレーム超解像のモデルを式 (3.5) に示す.

$$\hat{\mathbf{X}} = \underset{\mathbf{X}}{\operatorname{argmin}} \left[\sum_{k=1}^N \|\mathbf{A}_k \mathbf{X} - \mathbf{Y}_k\| + \lambda \sum_{l=-P}^P \sum_{m=-P}^P \alpha^{|m|+|l|} \|\mathbf{X} - \mathbf{S}_x^l \mathbf{S}_y^m \mathbf{X}\| \right] \quad (3.5)$$

この時, \mathbf{X} は推定画像, \mathbf{Y} は観測画像, \mathbf{A} は劣化行列, N は画像数, λ は画像の滑らかさの拘束強度, P は最大の平行移動量, \mathbf{S}_x^l は x 方向に l 画素だけ平行移動させる行列, \mathbf{S}_y^m は y 方向に m 画素だけ平行移動させる行列とする. 式 (3.5) の第2項は正則化項と呼ばれ, 画像全体の変動量に制限を与えることで再構成処理による画像の崩れを防止し, 安定して超解像画像を生成することができる. この問題を最急降下法を用いて展開したものを式 (3.6) に示す.

$$\mathbf{X}_{n+1} = \mathbf{X}_n - \beta \left[\sum_{k=1}^N \mathbf{A}_k^T \operatorname{sign}(\mathbf{A}_k \mathbf{X}_n - \mathbf{Y}_k) + \lambda \sum_{l=-P}^P \sum_{m=-P}^P \alpha^{|m|+|l|} (\mathbf{I} - \mathbf{S}_x^{-l} \mathbf{S}_y^{-m}) \operatorname{sign}(\mathbf{X}_n - \mathbf{S}_x^l \mathbf{S}_y^m \mathbf{X}_n) \right] \quad (3.6)$$

この時, β は収束のステップ幅, n は繰返し回数, \mathbf{I} は単位行列とする. 例として, 式 (3.9) を用いて, サブピクセル精度で異なる 16 枚の原画像 (図 3.8) を入力画像とした場合, 解像度を縦横 2 倍に超解像した結果を図 3.9, 線形補間結果を図 3.10 に示す.

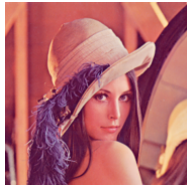


図 3.8 原画像

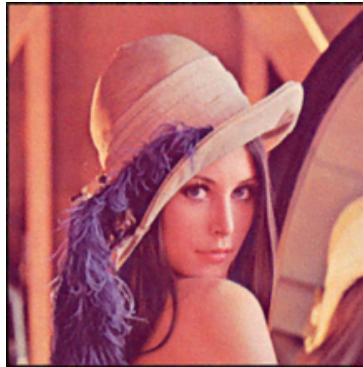


図 3.9 BTV 正則化



図 3.10 線形補間

3.5 問題点

図 3.3 より, 歪み補正結果の画像では周辺部がぼけており, 画像情報の一部欠落がみられる. 入力画像にこのような低画質な領域を含む画像を用いるため, 精度が悪く, 超解像を行っても結果を悪化させることがある. そのため, 本研究では魚眼レンズの周辺部を含めた魚眼画像全体の超解像手法の検討を行う必要がある.

第 4 章

提案手法

従来法では，魚眼画像の周辺部における超解像の精度が悪くなるという問題点があった．そのため，周辺部での精度向上することで全体の精度向上にもつながると考えて次の手法を提案する．

4.1 画像分割による超解像

魚眼画像は中央部と周辺部で歪みやぼけが異なるため，画像全体に対して超解像しても精度が向上しないと考えた．そのため，入力画像を複数領域に分割して，それぞれの分割領域ごとに SURF 特徴量を用いた台形補正，超解像を行い，最後に各超解像画像を合成して 1 枚の超解像画像を生成する．この手法を用いれば，分割した画像ごとに最適な超解像を行うことができると考えた．

4.1.1 画像分割（5 分割）

入力画像を中央部を 1 領域，周辺部を 4 領域の合計 5 領域に分けて分割した．図 4.1 の原画像を 5 分割した例を図 4.2 に示す．これは中央部に比べて周辺部の方が歪みとぼけが大きいため，周辺部は周辺部だけで処理を行うことで超解像の精度向上につながると考えた．

4.1.2 画像分割（9 分割）

入力画像を等間隔に 9 枚の画像に分割した．図 4.1 の原画像を 9 分割した例を図 4.3 に示す．これは 4.1.1 節の 5 分割よりも細かく分割することでさらに精度向上につながると考えられるため，分割数を増やして，9 分割として超解像を行う．

また，4.1.1，4.1.2 節では領域を分割する際にそれぞれの領域がが少しだけ重なるよ

うに大きめに分割した．これは提案手法で用いる BTV 正則化を用いたマルチフレーム超解像（3.4 節）を行うと，画像端の画素が欠落し黒くなり，画像合成時に図 4.4 のような黒線が超解像画像に入ってしまうためである．

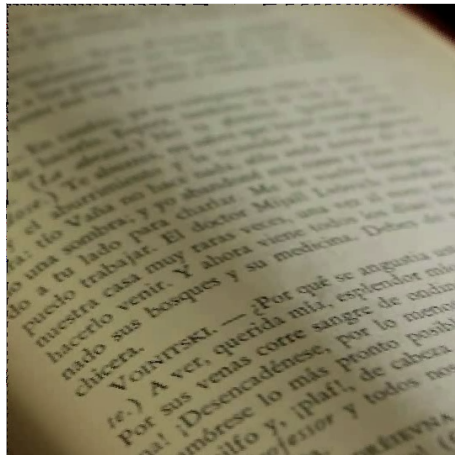


図 4.1 原画像

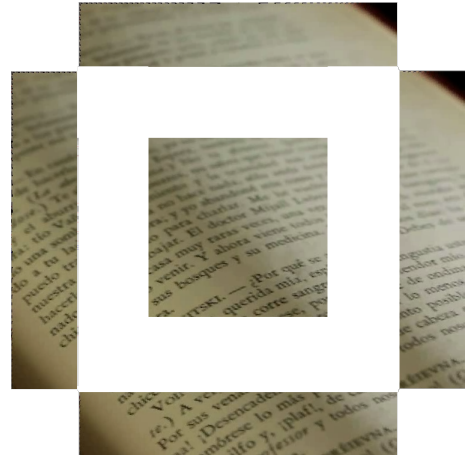


図 4.2 5 分割例

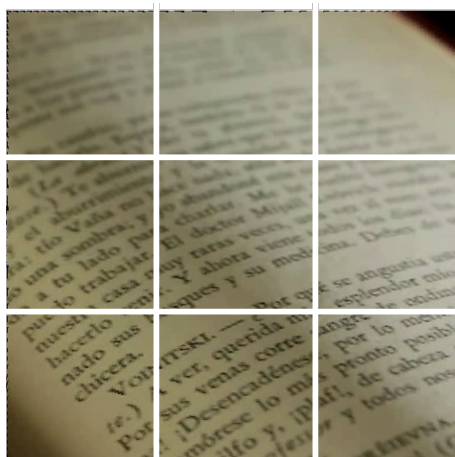


図 4.3 9 分割例



図 4.4 画像合成の失敗例

4.1.3 画像合成

各領域で超解像を行った後は，4.1.2 節で説明した通り，各領域が重なるように分割したため，重ならない内側の領域だけを切り抜く．そして，対応する位置に領域ごとの超解像画像を張り付けて，1 枚の超解像画像を生成する．

4.2 類似度による入力画像の制限

位置合わせを行う際に 16 枚のフレーム画像から基準のフレーム画像を選択し、基準フレーム画像から離れたフレーム画像ほど平行移動や台形補正によってデータの欠損が発生する。従来法では、このデータの欠損がある画像も超解像の入力画像として用いていた。しかし、BTV 正則化ではサブピクセル精度でズレている画像を用いることで超解像を行うことが前提となっているため、このデータ欠損が悪影響を与える可能性があると考えられる。そのため、テンプレートマッチングでも用いた類似度 (式 (3.4)) が一定以上でなければ、入力画像から除外することを考えた。

4.2.1 類似度が超解像に及ぼす影響の調査

類似度が超解像に及ぼす影響を調査するために類似度の上位 4 枚で超解像した結果、下位 4 枚で超解像した結果を次に示す。

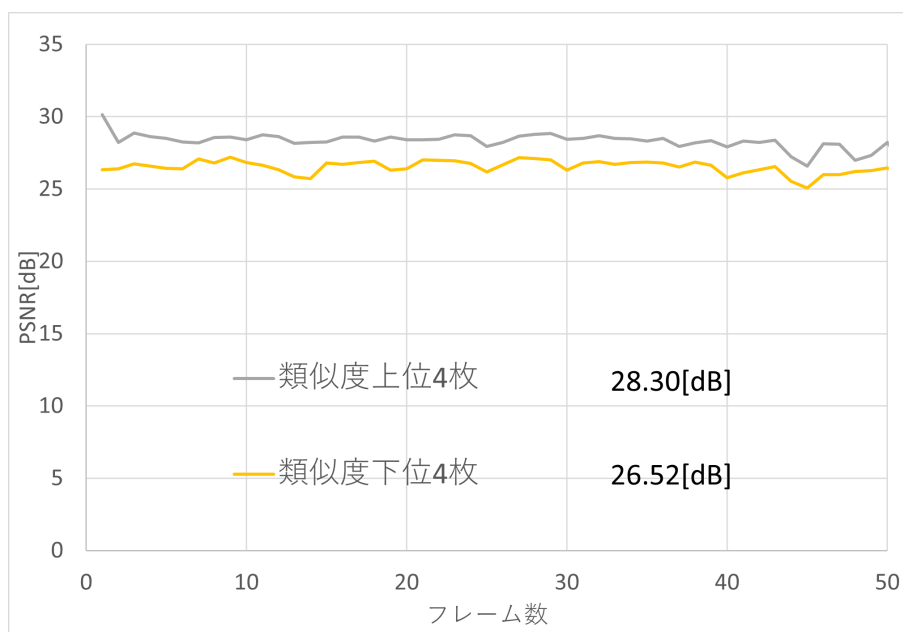


図 4.5 類似度上位 4 枚と下位 4 枚比較 (PSNR)

図 4.5, 4.6 より, PSNR, SSIM とともに類似度上位 4 枚で超解像した結果の方が下位 4 枚よりも上を推移しており, 平均 PSNR, SSIM においても上位 4 枚の方が高いことが確認できる。これにより, 類似度が高いフレーム画像を用いて超解像することで画質が向上すると考えられる。

例として, 実験的に決定した類似度の閾値が 0.99 のときの画像比較を図 4.7 に示

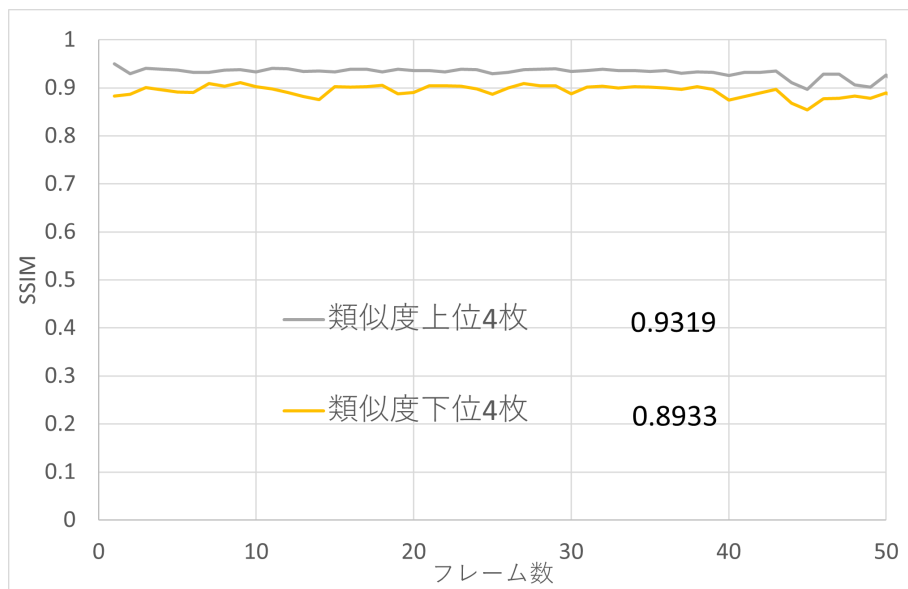


図 4.6 類似度上位 4 枚と下位 4 枚比較 (SSIM)

す. 3.3 節で位置合わせを行っているため, 大半のフレーム画像で類似度が 0.99 以上の値となっているが, 一部で特徴点を検出できずに SURF 特徴量の正しいマッチングができないため, ホモグラフィ変換によって画像がズレてしまうことがある. そのため, 類似度が 0.99 に満たないフレーム画像を入力画像から除外することを考えて, 閾値を 0.99 とした.

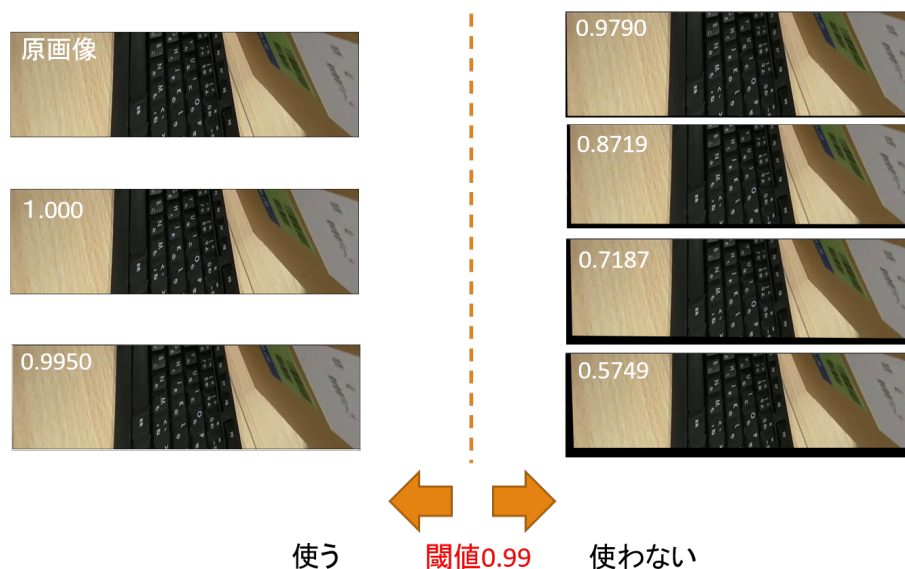


図 4.7 類似度比較

第 5 章

シミュレーション

本章では BTV 正則化に基づく魚眼画像の超解像シミュレーションを行い、従来法と提案手法の結果を比較し、4 章で述べた提案手法の有用性の検討を行う。

5.1 シミュレーション条件

入力動画のフレーム画像を図 5.1, 5.2 に示す。また、従来法と提案手法で用いる入力動画は同一なものを使用する。

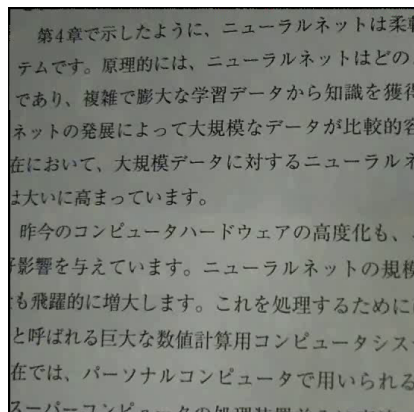


図 5.1 動画 1 のフレーム画像

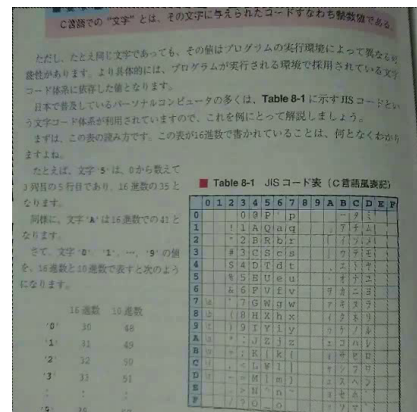


図 5.2 動画 2 のフレーム画像

超解像が高周波成分を復元する技術であるため、文字などの高周波成分を多く含まれている図 5.1, 5.2 の動画を選定した。動画 1 は途中でカメラが右に移動するシーンがあり、動画 2 ではカメラが静止している。今回のシミュレーションでは動画に対して、縦横 2 倍ずつ、ピクセル数では 4 倍の超解像処理を行っている。例えば、本節のシミュレーションでは解像度が 270×270 の動画に対して超解像処理を行い、 540×540 の超解像動画を生成している。また、本研究で使用している全ての手法では複数枚の

フレーム画像を入力して1枚の超解像画像を生成しているため、入力フレーム画像として基準となるフレームの前後16フレームを用いる。

5.2 歪み補正

また本研究では、魚眼レンズを用いて撮影した魚眼画像を用いるのではなく、通常のレンズで撮影した画像に対して魚眼化フィルタをかけて魚眼画像に変換した画像を用いる。魚眼化フィルタと逆の処理を行うことで魚眼画像の歪み補正を行う。これは、魚眼画像の歪み補正が困難であり、確実に歪みが補正されるとは限らなく、魚眼化フィルタを用いれば理想的な歪み補正を行うことができると考えたためである。魚眼化フィルタには多くの魚眼レンズで採用されている等距離射影方式を用いる。光軸からの角度を θ 、焦点距離を f 、画像中心からの距離を r とする場合、等距離射影方式の魚眼レンズモデルを以下の式 (5.1) に示す。

$$r = f\theta \quad (5.1)$$

等距離射影方式とは、画像中心から離れるほど歪みが大きくなり、距離が同じであれば放射状で同じ歪み方をする方式である。以下に通常画像から魚眼画像への変換式を示す。

$$u = \frac{2x - w}{w}, v = \frac{2y - h}{h} \quad (5.2)$$

$$r = \sqrt{u^2 + v^2} \quad (5.3)$$

$$u' = \frac{\left(\frac{u}{1-fr} + 1\right)w}{2}, v' = \frac{\left(\frac{v}{1-fr} + 1\right)h}{2} \quad (5.4)$$

x, y は通常画像の座標、 w, h は通常画像の解像度、 u, v は通常画像の中心を原点として正規化（-1 から 1）した座標、 u', v' は変換後の魚眼画像の座標、 r は原点からの距離、 f は焦点距離を表している。

また、歪み補正の変換式を以下に示す。

$$u = \left(\frac{2u'}{w} - 1\right)(1 - fr), v = \left(\frac{2v'}{h} - 1\right)(1 - fr) \quad (5.5)$$

$$x = \frac{(u + 1)w}{2}, y = \frac{(v + 1)h}{2} \quad (5.6)$$

原画像、魚眼画像変換、歪み補正の結果を図 5.3, 5.4, 5.5 に示す。図 5.3 の原画像に魚眼フィルタをかけて、図 5.4 の魚眼化した画像に変換し、最後に魚眼フィルタとは逆の変換を行うことで図 5.5 の歪み補正を行った画像を生成する。

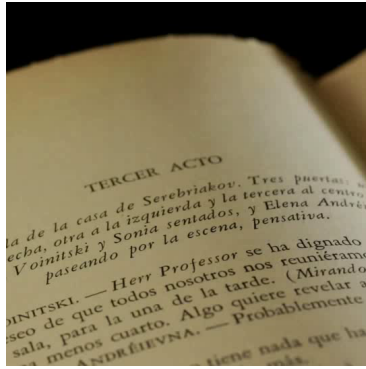


図 5.3 原画像



図 5.4 魚眼化した画像

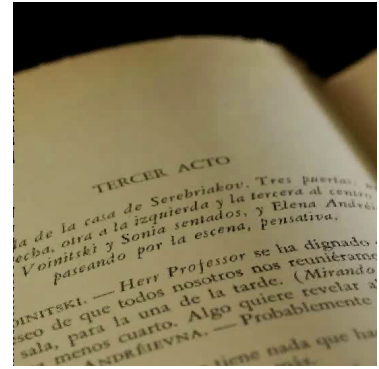


図 5.5 歪み補正後

5.3 シミュレーション結果

5.3.1 提案手法 1

提案手法 1 では入力フレーム画像を複数領域に分割し、それぞれの領域に対して超解像した後、画像合成で 1 枚の超解像画像を生成した。まず、動画 1, 2 の従来法の結果を図 5.6, 5.9 に示す。左上が原画像、右上が評価画像、左下が超解像画像、右下が線形補間画像となっている。そして、5 分割と 9 分割で超解像を行った結果を図 5.7, 5.8, 5.10, 5.11 に示す。

図 5.6, 5.7, 5.8 より、超解像画像の方が線形補間画像よりも文字が鮮明に表示されているため、高周波成分が復元できていることが確認できた。また、図 5.6 の超解像の PSNR(Peak Signal to Noise Ratio) と SSIM(Structural similarity) [13] を図 5.12, 5.13 に示す。

図 5.12, 5.13 より、ほぼ全てのフレームで従来法と提案手法 1, 2 の PSNR, SSIM にほとんど違いがないことが分かった。また、平均 PSNR, SSIM でも従来法と従来手法 1, 2 に大きな違いは発生しなかった。このことから各領域に分割して超解像した場合では同程度の品質の動画を生成することができる。

5.3.2 提案手法 2

基準となるフレームとの類似度が低いフレーム画像を使用しない処理を追加し、入力フレーム画像の制限を行った。5 分割超解像より、分割領域が小さく SURF 特徴量による位置合わせが困難であると考えられる 9 分割超解像の結果の画像を図 5.16, 5.17 に示す。5 分割, 9 分割超解像を行った時の画像全体の PSNR と SSIM を図 5.18, 5.19, 5.20, 5.21 に示す。

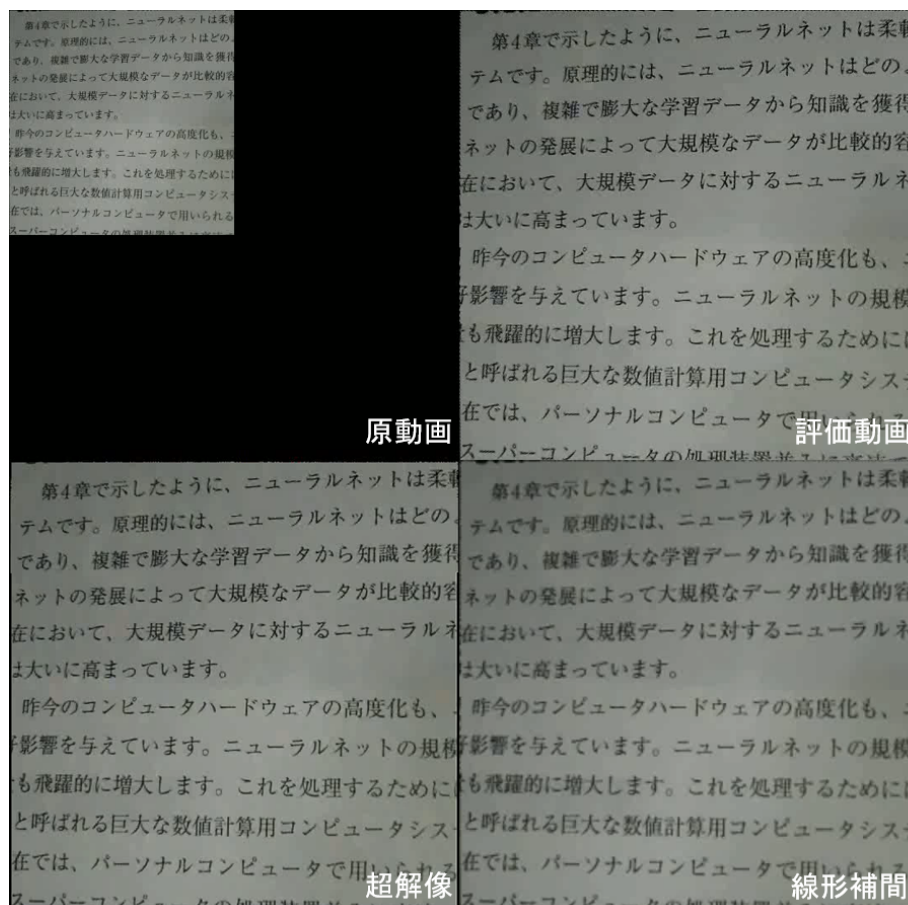


図 5.6 従来法の超解像結果（動画 1）

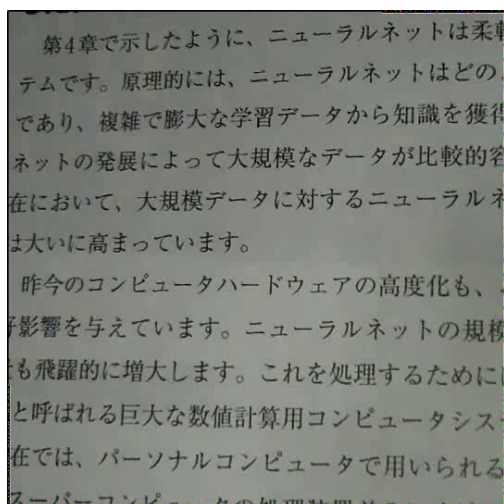


図 5.7 提案手法 1 の 5 分割超解像結果（動画 1）

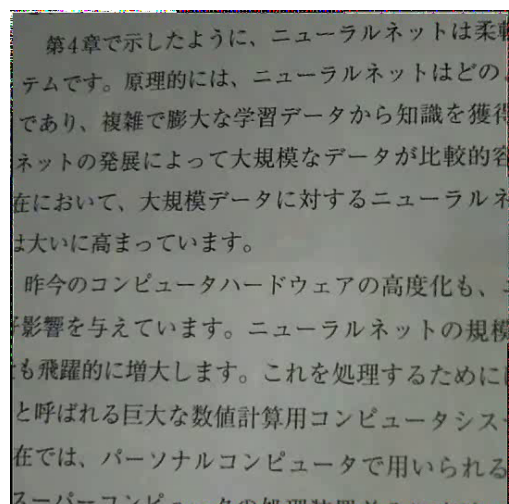


図 5.8 提案手法 1 の 9 分割超解像結果（動画 1）

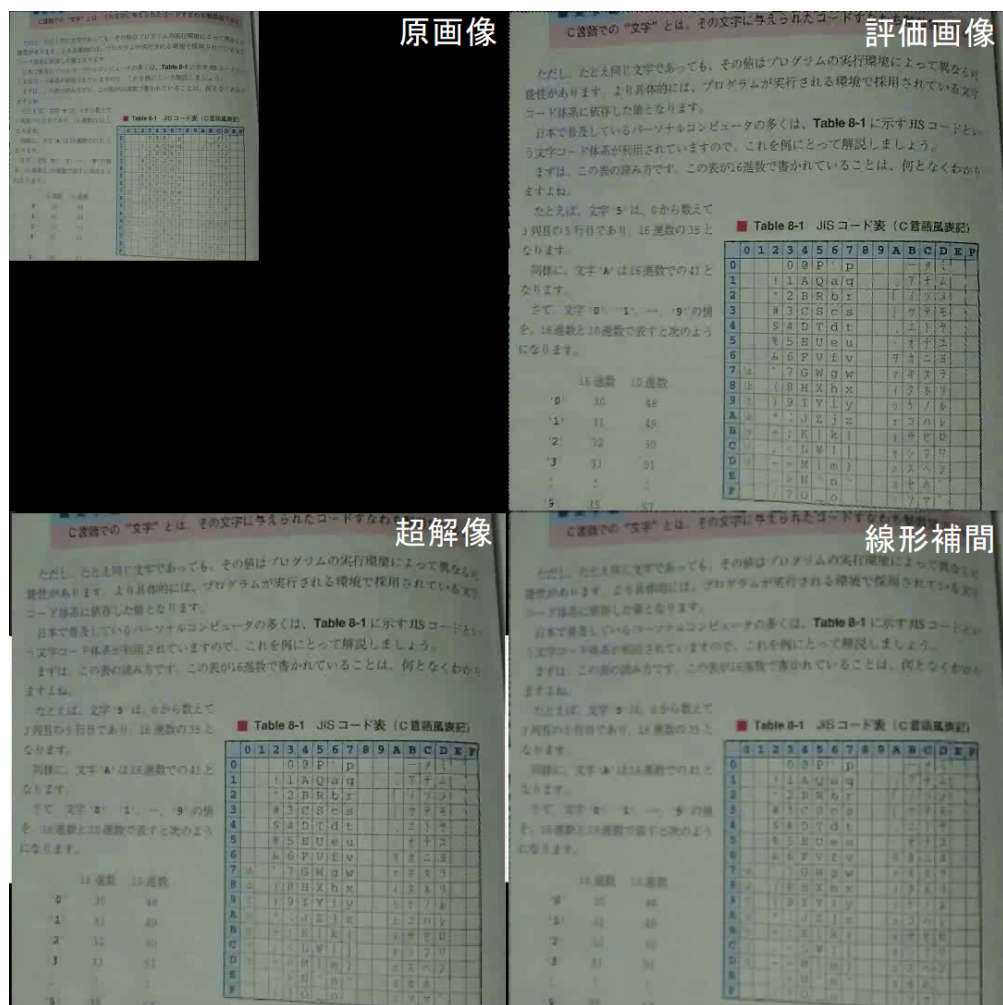


図 5.9 従来法の超解像結果 (動画 2)

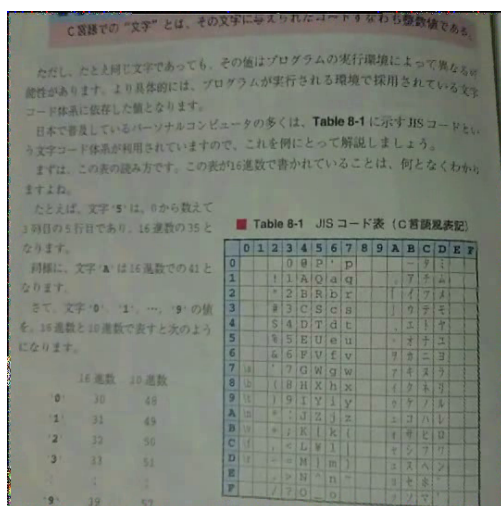


図 5.10 提案手法 1 の 5 分割超解像結果 (動画 2)

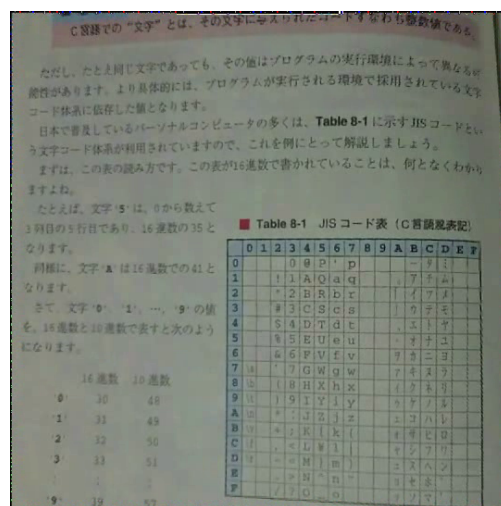


図 5.11 提案手法 1 の 9 分割超解像結果 (動画 2)

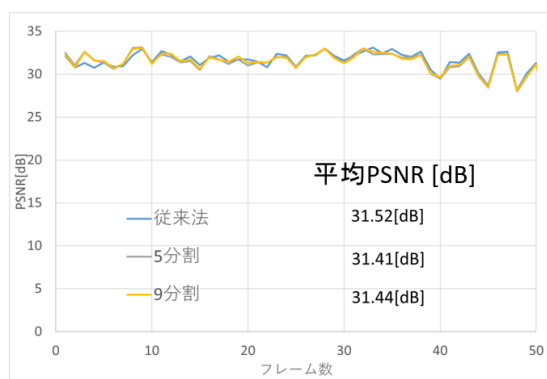


図 5.12 PSNR 比較 (動画 1)

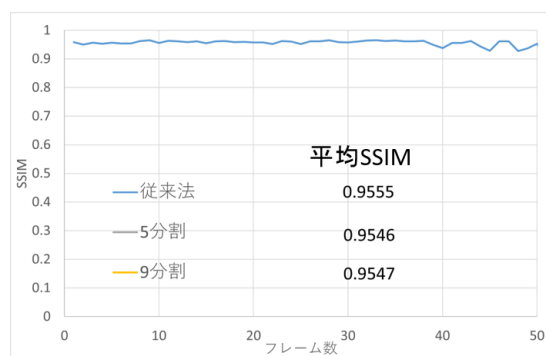


図 5.13 SSIM 比較 (動画 1)

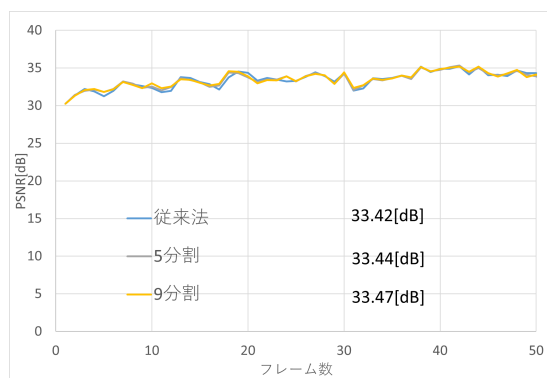


図 5.14 PSNR 比較 (動画 2)

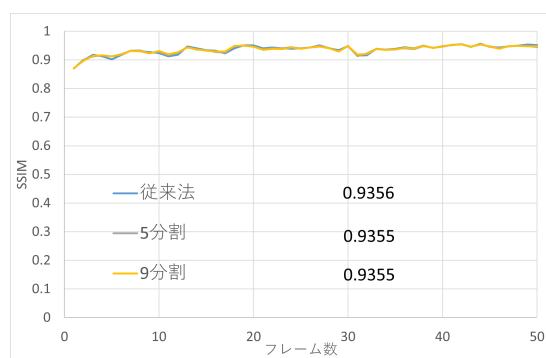


図 5.15 SSIM 比較 (動画 2)

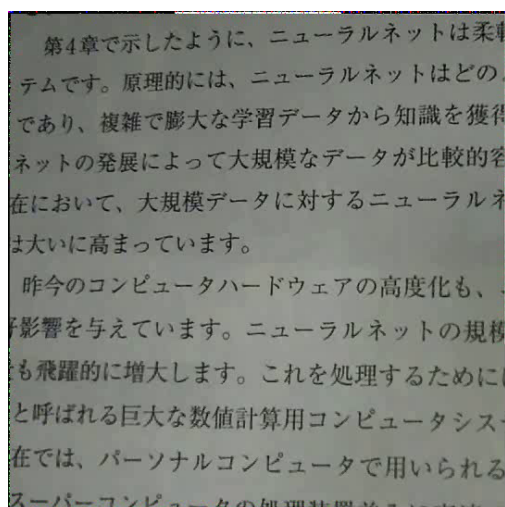


図 5.16 提案手法 2 の 9 分割超解像結果 (動画 1)

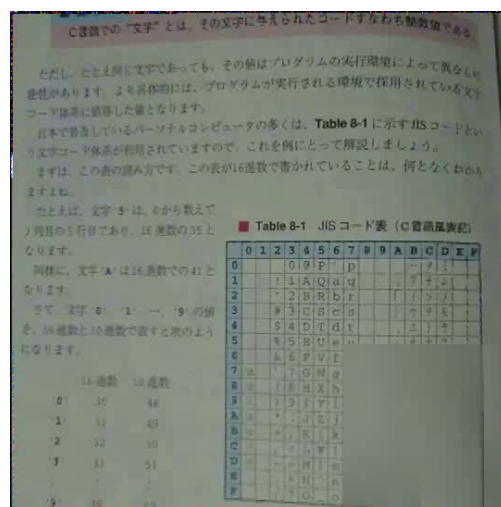


図 5.17 提案手法 2 の 9 分割超解像結果 (動画 2)

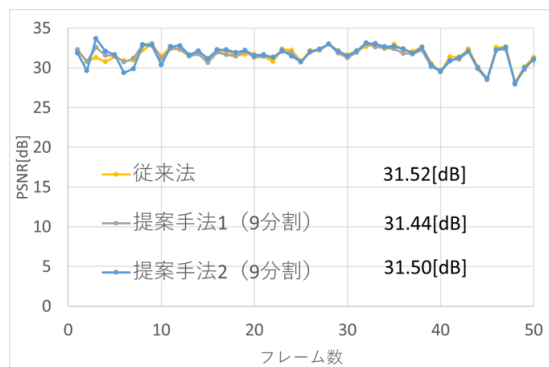


図 5.18 提案手法 2 の 9 分割 PSNR 比較 (動画 1)

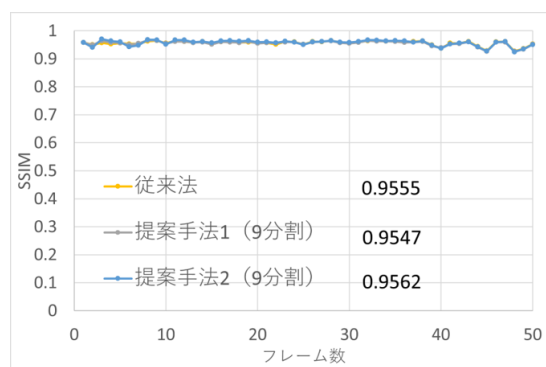


図 5.19 提案手法 2 の 9 分割 SSIM 比較 (動画 1)

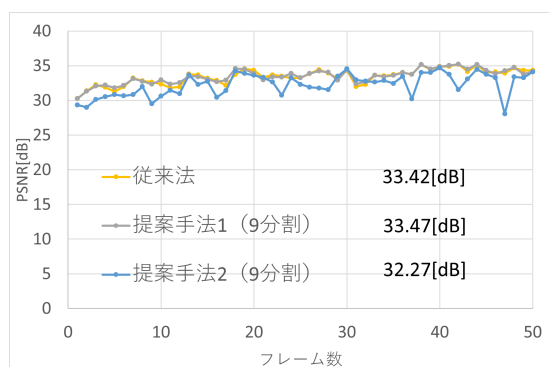


図 5.20 提案手法 2 の 9 分割 PSNR 比較 (動画 2)

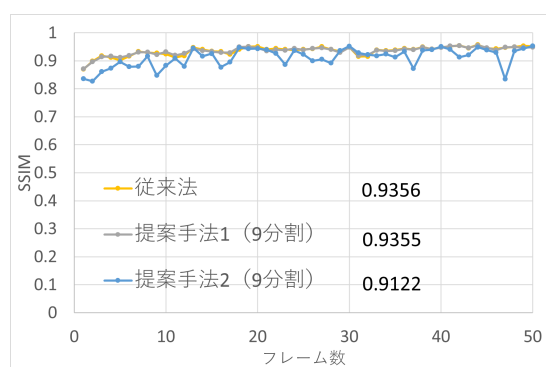


図 5.21 提案手法 2 の 9 分割 SSIM 比較 (動画 2)

図 5.18, 5.19 より, 従来法, 提案手法 1 に比べて提案手法 2 は PSNR, SSIM ともにほぼ変化がないことが確認できる. 図 5.20, 5.21 では従来法と提案手法 1 が同程度の結果となったが, 提案手法 2 は従来法, 提案手法 1 に比べて, 同程度, もしくは少し下を推移していることが分かる. これにより, 提案手法 2 は提案手法 1 に対して優位であるとは言えない.

5.4 考察

前節では, 従来法と分割処理を追加した超解像の結果を比較し, 従来法の方が PSNR, SSIM ともに高くなる, もしくは同程度であることを示した. この節では各提案手法について考察をする.

5.4.1 提案手法1の考察

従来法と提案手法で生成した動画を5分割したときの各分割領域でのPSNRを算出し、分割したときの画像位置と同じ配置でグラフを図5.22, 5.23に並べた。

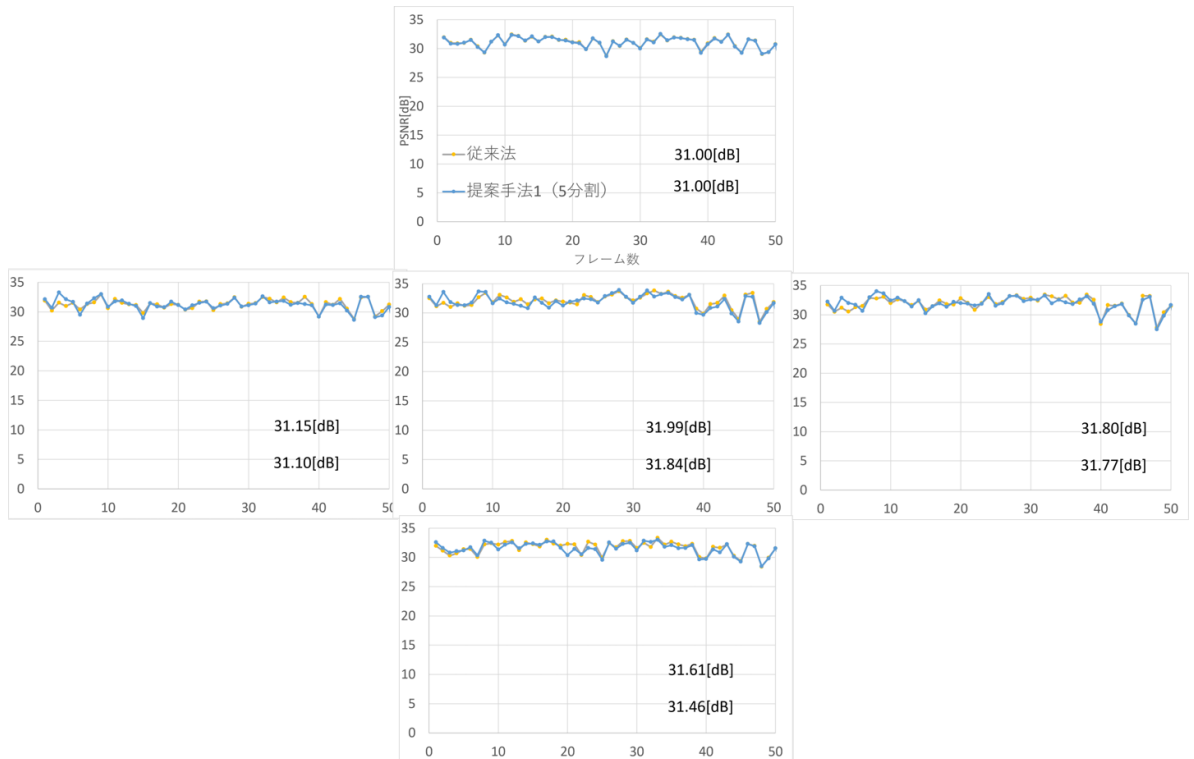


図 5.22 提案手法1の5分割PSNR比較（動画1）

図5.22より、動画1の5分割では上下左右中央の5領域全てにおいて、従来法と提案手法1のPSNRがほぼ同じになった。図5.23より、動画2の5分割でも同様に5領域全てで従来法と提案手法1のPSNRが同程度となることがわかった。

従来法と提案手法1で生成した動画を9分割したときの各分割領域でのPSNRを算出し、分割したときの画像位置と同じ配置でグラフを図5.24, 5.25に並べた。動画1, 2の9分割に関しても5分割と同様の結果となることが図5.24, 5.25より読み取れる。これにより、提案手法1は従来法と同程度の結果となることが分かった。

5.4.2 提案手法2の考察

動画1, 2の9分割超解像したときのPSNRを画像と同じ配置で図5.26, 5.27に示す。

図5.26, 5.27では、第2縦軸を入力フレーム数として黄線で表している。図5.26で

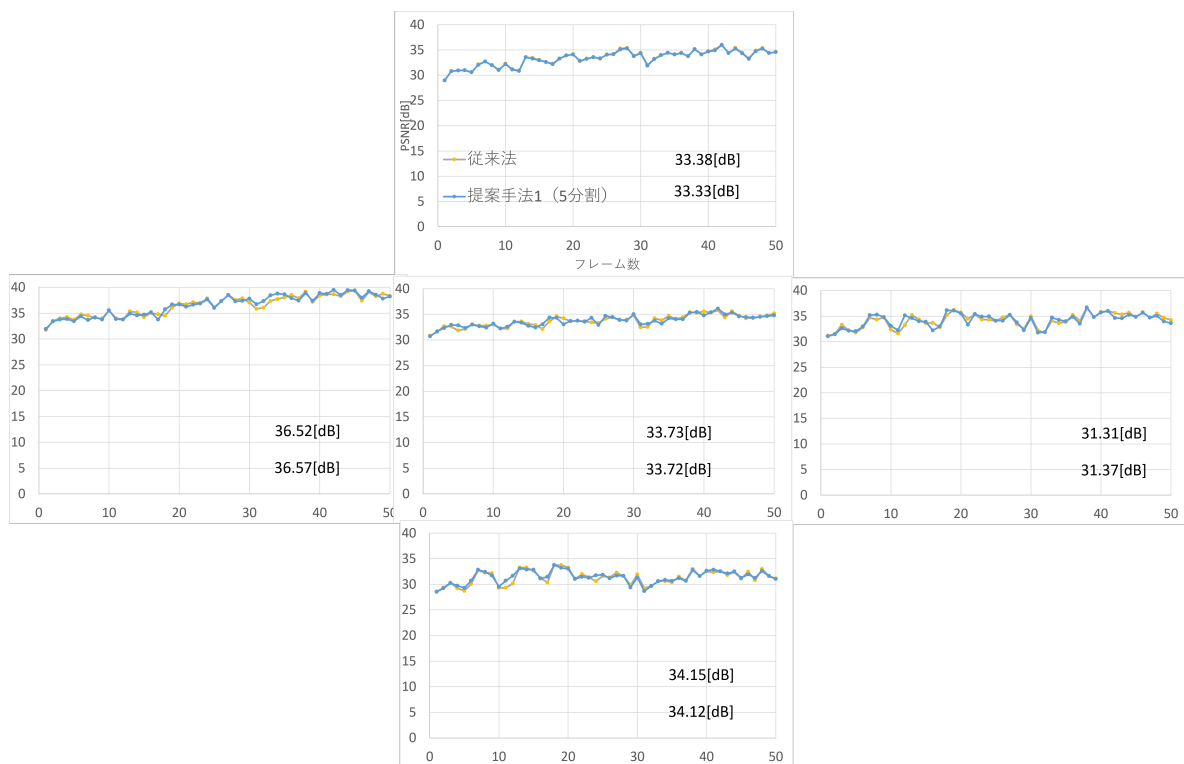


図 5.23 提案手法 1 の 5 分割 PSNR 比較 (動画 2)

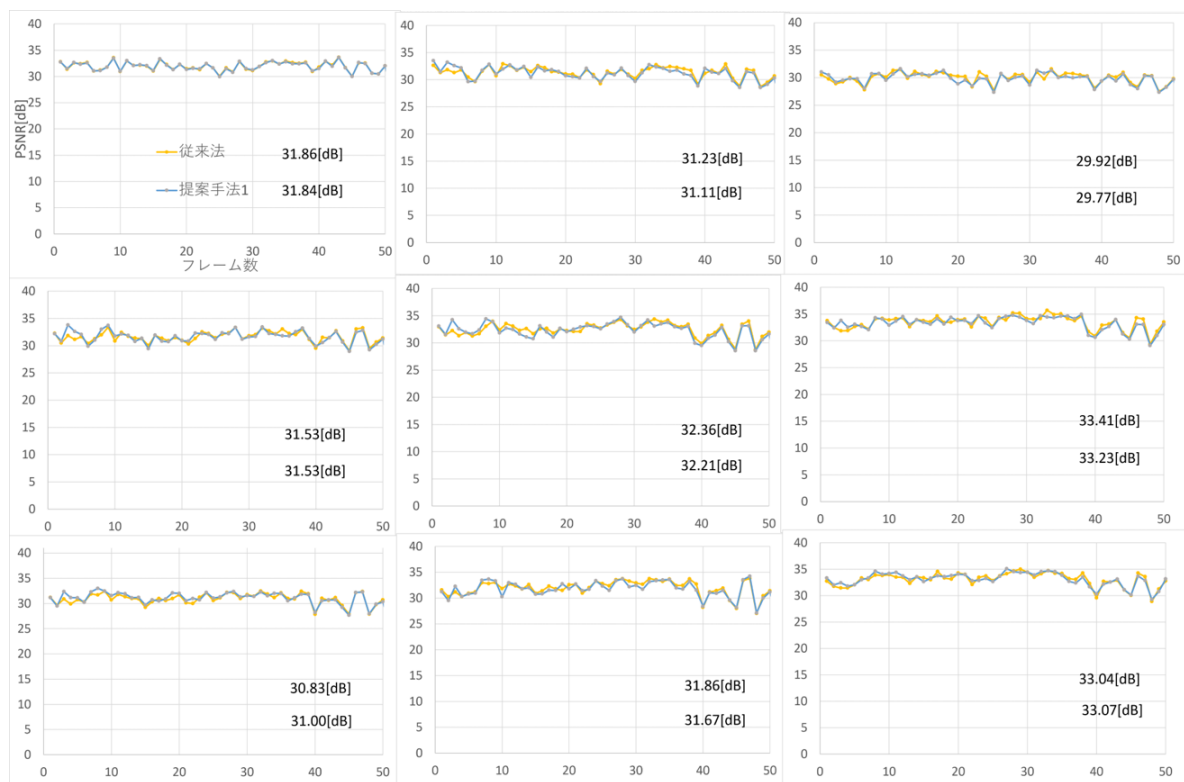


図 5.24 提案手法 1 の 9 分割 PSNR 値比較 (動画 1)

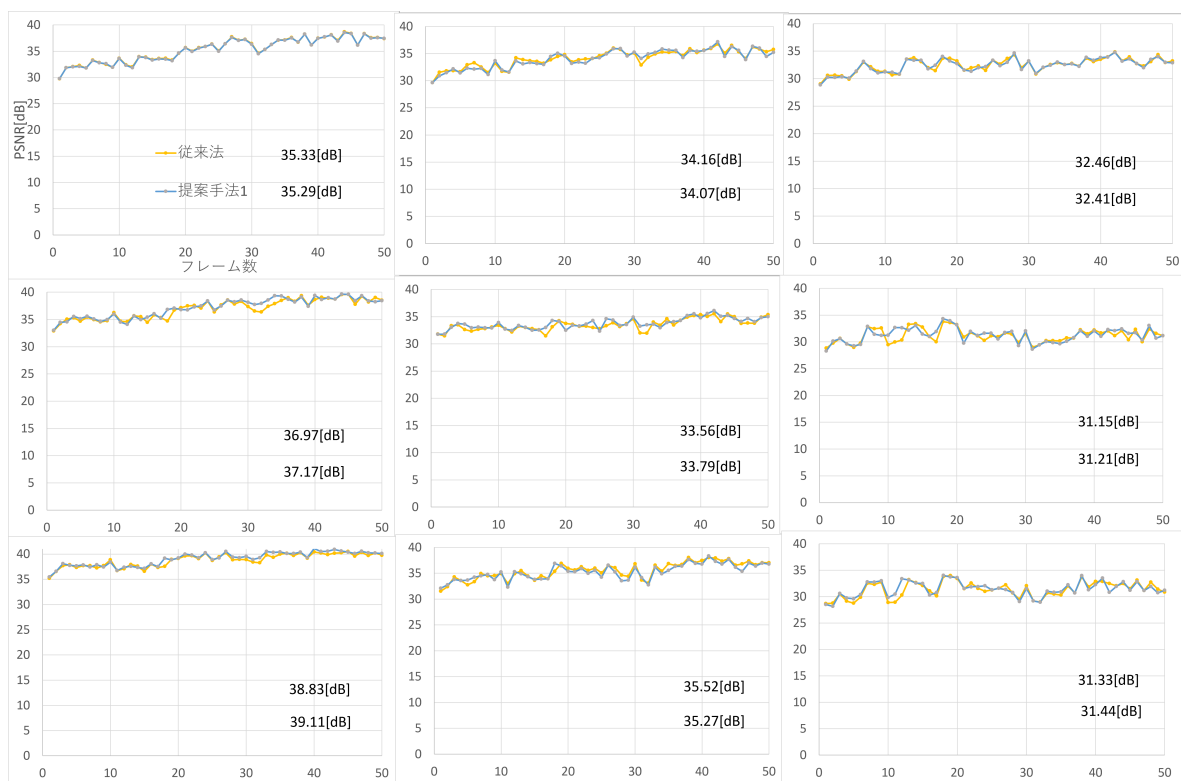


図 5.25 提案手法 1 の 9 分割 PSNR 値比較 (動画 2)

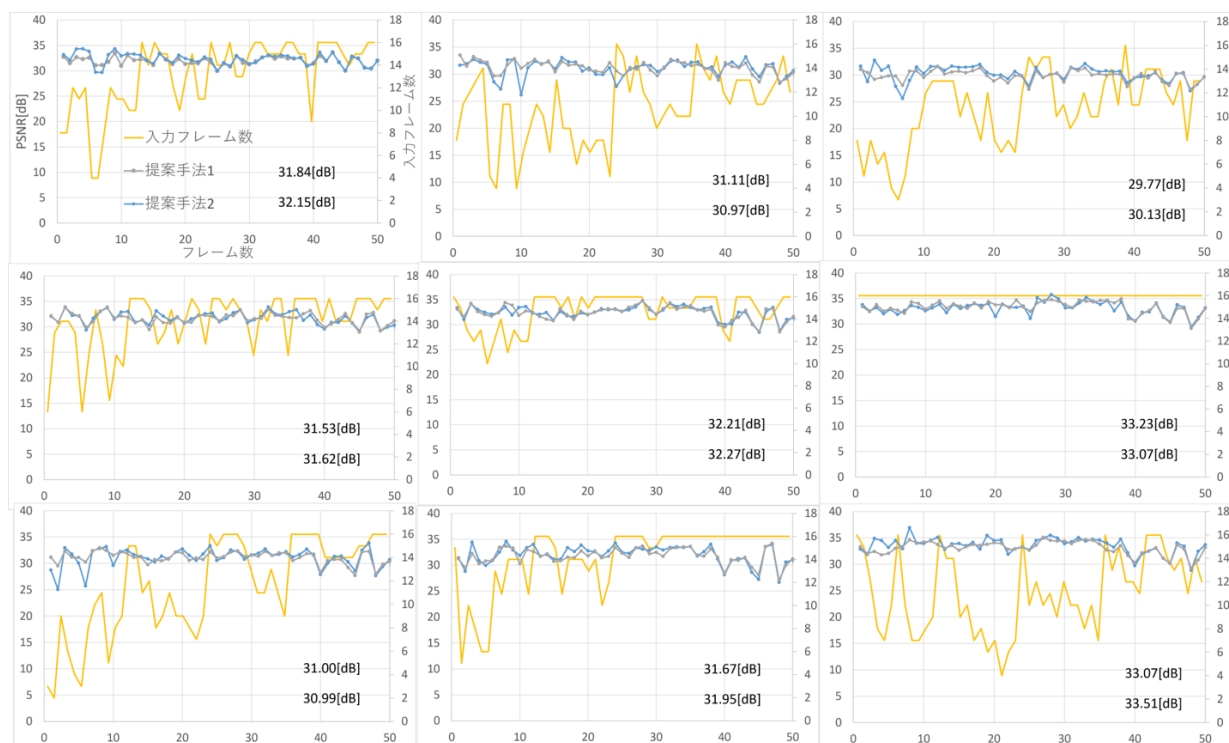


図 5.26 提案手法の 9 分割領域ごとの PSNR 比較 (動画 1)



図 5.27 提案手法の 9 分割領域ごとの PSNR 比較 (動画 2)

は、提案手法 2 で入力フレーム数の制限によってフレーム数が減少した、ほぼすべてのフレームで PSNR が向上していることが読み取れる。これはテンプレートマッチングと SURF 特徴量を用いた台形補正において、位置合わせが不十分でサブピクセル以上に位置がずれてしまった画像を取り除いたため、精度が向上したと考えられる。図 5.27 では、右下以外の領域では提案手法 2 の PSNR の方が高くなっていることが分かる。左下の領域は入力フレームが常に最大の 16 フレームで提案手法 1 と同じ結果になった。図 5.27 の左上、右、右下の領域では、入力フレーム数が極端に少ないときに提案手法 1 よりも PSNR が低くなっていることが確認できる。そのため、入力フレームが少なすぎると画質が低下することが分かった。

さらに従来法と提案手法 2 の結果を 9 分割領域ごとに PSNR で比較したグラフを図 5.28, 5.29 に示す。

図 5.28, 5.29 より、従来法と提案手法 2 を比較すると、PSNR が上下に変動しているが、全体的に見ると提案手法 2 は同程度または若干下を推移していることが確認できる。そのため、従来法で問題となっていた周辺領域を含む画像全体の超解像は提案手法 2 で改善することができない。

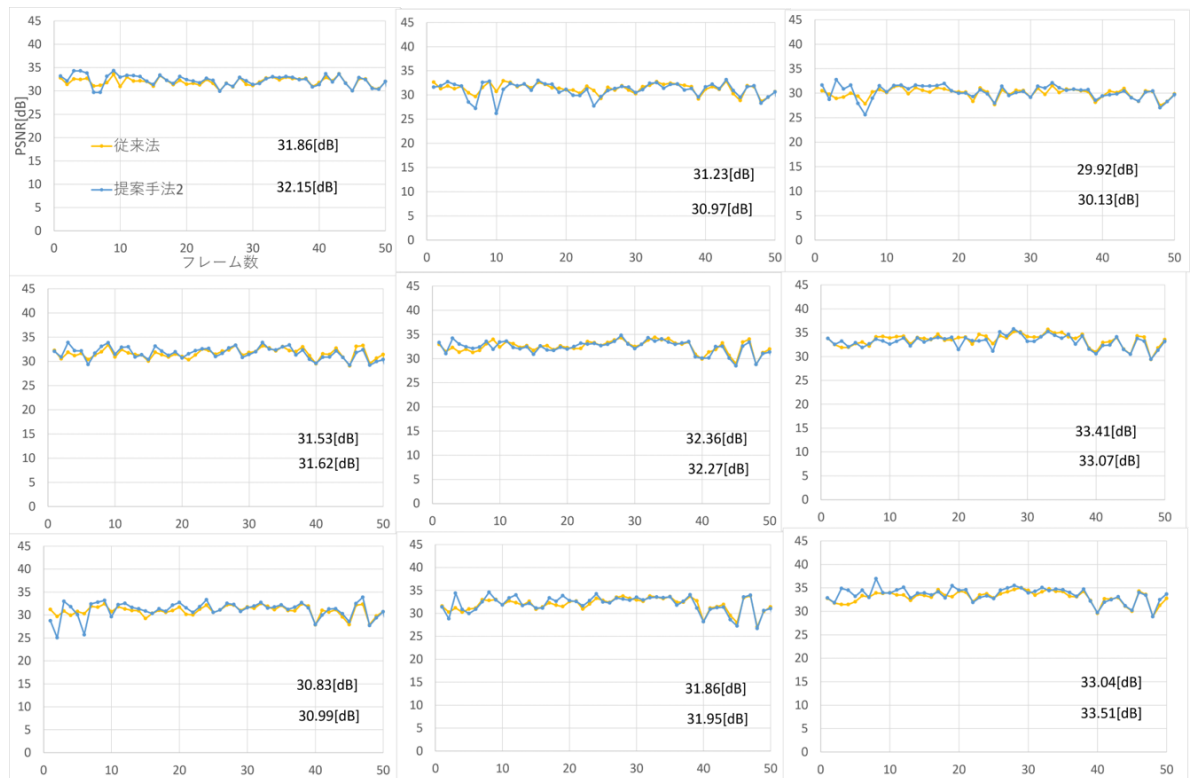


図 5.28 従来法と提案手法 2 の 9 分割領域ごとの PSNR 比較 (動画 1)

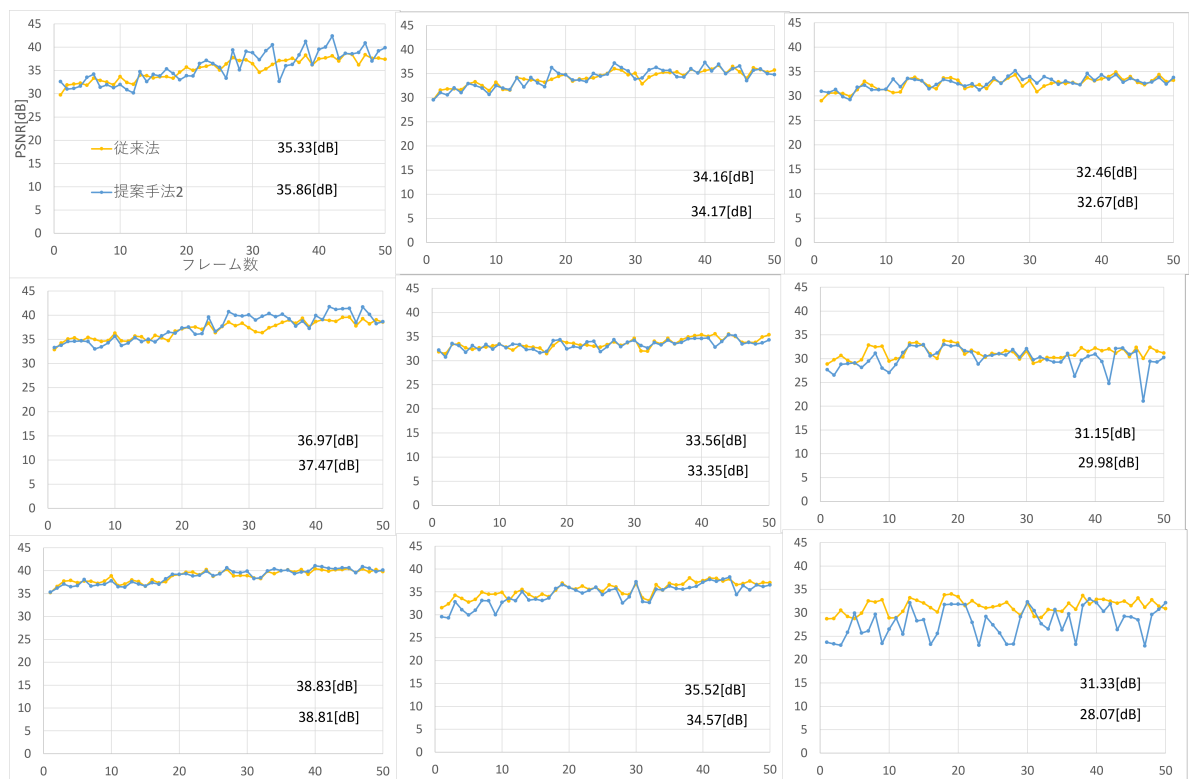


図 5.29 従来法と提案手法 2 の 9 分割領域ごとの PSNR 比較 (動画 2)

5.4.3 処理時間の比較

従来法と提案手法 1, 2 の 5 分割, 9 分割超解像の処理時間を比較すると表 5.1 となる。従来法より提案手法 1, 2 の 9 分割超解像の方が計算量が減るため処理時間が減少

表 5.1 処理時間比較

| | 動画 1[s] | 動画 2[s] |
|---------------|---------|---------|
| 従来法 | 16239 | 6025 |
| 提案手法 1 (5 分割) | 21501 | 6940 |
| 提案手法 1 (9 分割) | 16832 | 5880 |
| 提案手法 2 (5 分割) | 20689 | 5879 |
| 提案手法 2 (9 分割) | 14882 | 4986 |

した。しかし, 5 分割超解像に関しては従来法よりも処理時間が増加しているが, これは図 5.30 のように分割した領域が画像全体の四隅で被ってしまっており, 四隅だけ 2 回ずつ超解像しているため余計に処理時間がかかっていると考えられる。

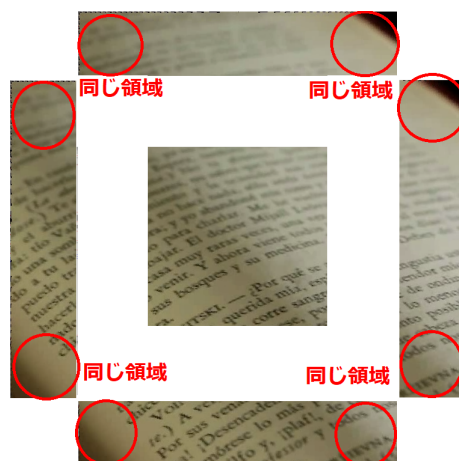


図 5.30 5 分割超解像の分割領域の重複箇所

そして, 提案手法 1, 2 では PSNR や SSIM の結果が同程度, もしくは若干悪化するが, 画像全体に対して超解像する場合に比べて, 領域ごとに並列処理で超解像できるため, 処理時間を短縮できる。そのため, 提案手法 1 は画像を分割して並列処理で超解像することで処理時間の短縮, 提案手法 2 は入力画像数を減らすことでさらなる処理時間の短縮ができるため, 有効であると考えられる。

第 6 章

結論

6.1 まとめ

本論文では、魚眼画像の超解像において、中央領域付近の超解像しか検討されていなかったため、周辺領域を含む魚眼画像全体の超解像手法を提案した。第 1 章では、魚眼カメラ及びそれを応用した製品の需要が高まっている背景に触れ、魚眼カメラの高画質化の必要性とその手法である超解像について簡単に述べた。また、本論文における研究目的についても述べた。第 2 章では、超解像の種類や画像モデル、代表的な手法などの基礎的な知識・技術について述べた。第 3 章では、魚眼画像の超解像に関する既存手法について紹介した。この既存手法では、魚眼画像の比較的画質の良い中央領域付近にのみに対象を絞って超解像を行っており、周辺領域の超解像は精度が悪く画質が向上しないという問題点があった。この問題点を改善するために、第 4 章では入力画像を複数領域に分割して、各領域ごとに超解像を行い、画像合成で 1 枚の超解像画像を生成する手法を提案した。また、複数領域に分割することでサブピクセル精度での位置合わせが困難になる問題があったため、類似度と呼ばれる 2 枚の画像がどの程度似ているかを表す指標を用いて、一定値以下の画像は入力画像から排除する手法も提案した。第 5 章では、2 つの提案手法のシミュレーションを行い、画像の評価指標である PSNR と SSIM を用いて画質の比較を行った。分割して超解像する手法は既存手法に比べて精度が落ちるが、分割領域ごとに並列処理を行うことで超解像処理にかかる処理時間を短くすることができたことが分かった。また、類似度による入力画像の制限の処理を追加することで既存手法とほぼ同等もしくは少し低い画質になるが、超解像に用いる入力画像が減るため、さらなる処理時間の短縮ができることが分かった。したがって、提案手法 1, 2 は従来法と同程度、もしくは少し画質が低下し、画質としては魚眼画像全体の超解像に有用性を見出せなかったが、並列処理が可能となり、処理時間とメモリ使用量が少なくなったことで実際の製品に搭載することが容

易なっと思われ.

6.2 今後の課題

高周波成分の少ない領域では線形補間の方が画質が良くなることが分かったため, 高周波成分が多い領域では提案手法 2, 少ない領域では線形補間を用いることでさらに画質が向上すると考えられるため, 調査, 検討を行う. また, 動きのあるフレーム画像を超解像する場合はモーションブラーにより高周波成分がぼけてしまい, 位置合わせ精度が低下し, 超解像に悪影響を与えてしまうため, 改善方法の調査・検討を行う.

謝辞

本研究を修士論文として形にすることができたのは、担当していただいた杉田泰則准教授の熱心なご指導や、論文の審査において的確なご指示を頂きました、本学電気系岩橋政宏教授ならびに圓道知博准教授、そして多くのご指摘をくださいました信号処理応用研究室の同期・後輩の皆様のおかげです。協力していただいた皆様へ心から感謝の気持ちと御礼を申し上げたく、謝辞にかえさせていただきます。最後に、勉学に励む機会を与え様々な面において支えていただいた両親に深く感謝いたします。

令和3年3月

参考文献

- [1] T. Takano, S. Ono, Y. Matsushita, H. Kawasaki, K. Ikeuchi, "Super resolution of fish-eye images captured by on-vehicle camera for visibility support", IEEE International Conference on Vehicular Electronics and Safety, Volume 67, Issue 2, pp. 99-104, November, 2015
- [2] J. Yang, J. Wright, T. S. Huang, Y. Ma, "Image Super-Resolution Via Sparse Representation", IEEE Transactions on Image Processing, Volume 19, No. 11, pp. 2861-2873, November, 2010
- [3] S. C. Park, M. K. Park, M. G. Kang, "Super-resolution image reconstruction: a technical overview", IEEE Signal Processing Magazine, Volume 20, Issue 3, pp. 21-36, May, 2003
- [4] 本間希樹, 笹田真人, 田崎文得, "スパースモデリング天文学—ブラックホール撮像から時間変動現象まで", 科学研究費補助金 新学術領域研究「スパースモデリングの深化と高次元データ駆動科学の創成」最終成果報告会, December 18-20th, 2017
- [5] William T. Freeman, Egon C. Pasztor, Owen T. Carmichael, MERL, Mitsubishi Electric Research Labs, "Learning Low-Level Vision", International Journal of Computer Vision, Volume 40, pp. 25-47, 2000
- [6] S. Gohsh, "A New Signal Processing Method for Video Image—Reproduce the Frequency Spectrum Exceeding the Nyquist Frequency Using a Single Frame of the Video Image". The Era of Interactive Media, pp. 593-604, August, 2012
- [7] 斎藤翠, 長尾智晴, "フラクタル次元に基づく無制限超解像処理", The 30th Annual Conference of the Japanese Society for Artificial Intelligence, 2016
- [8] Chao Dong, Chen Change Loy, Kaiming He, Xiaoou Tang, "Image Super-Resolution Using Deep Convolutional Networks", Conference on Computer Vision and Pattern Recognition, Dec, 2014
- [9] 高野照久, 松下侑輝, 小野晋太郎, 川崎洋, 池内克史, "視界支援用の車載カメラとして使用される魚眼カメラ画像列を用いた超解像の提案", 生産研究, 67 巻, 2 号,

- pp.99-104, 2015
- [10] H. Bay, T. Tuytelaars, and L. Van Gool, "Speeded Up Robust Features", 9th European Conference on Computer Vision, pp.404-417, 2006
- [11] M.A Fischler, R.C Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography", Communications of the ACM, Volume 24, No. 6, pp. 381-395, 1981
- [12] S. Farsiu, D. Robinson, M. Elad, P. Milanfar, "Fast and robust multiframe super resolution," IEEE Transactions on Image Processing, Volume 13, pp.1327-1344, 2004
- [13] Zhou Wang, Member, Alan Conrad Bovik, Hamid Rahim Sheikh, Eero P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity", IEEE Transactions on Image Processing, Volume. 13, No. 4, pp.600-612, April, 2004