

卒業論文概要書

Summary of Bachelor's Thesis

Date of submission: 02/08/2023 (MM/DD/YYYY)

学科名 Department	情報通信	氏名 Name	張 歌	指 導 教 員 Advisor	渡辺 裕 印
研究指導名 Research guidance	オーディオビジュアル 情報処理研究	学籍番号 Student ID number	1w193081-6		
研究題目 Title	フィルムグレインノイズを含んだ画像の超解像に関する研究				

1. 研究背景

近年、4K モニターやテレビなど高解像度のディスプレイを持っていることが珍しいことでもなくなった。その影響で映像コンテンツをより良い画質で楽しみたいという需要が増え、昔販売された映画を高解像度のディスプレイでも楽しめるようリマスターすることが増えた。しかし元の映像となるフィルムが何らかの理由で消失してしまった場合、すでに販売されている映像でリマスターする必要があり、そこで超解像技術が必要となる。



図2 従来手法の超解像の例

本研究ではそのフィルムグレインノイズを残したまま映画を超解像行うことを目的としている。ただし本研究では映画の動画ではなく映画の一フレームを切り取った画像にたいして主眼を置く。

2. 研究目的

古いフィルム映画にはフィルムグレインノイズと呼ばれる独特の丸い形状をしたノイズがある。一部の映画愛好家ではそのノイズが映画における重要な構成要素の一つであると考えている。従来の超解像技術ではそのフィルムグレインノイズをただのノイズとして除去してしまう、または結果として画がのっぺりとしてしまう、またはフィルムグレインノイズ同士が結合しもやもやしたがぞうになるなど映画らしさがなくなってしまう。(図1,2を参照)



図1 元画像[3]

3. 提案手法

フィルムグレインノイズを含む画像を end-to-end で学習するのではなく、元画像をフィルムグレインノイズとフィルムグレインノイズを取り除いた画像を2つに分けそれぞれを超解像し、出力されたその2つの結果を合成することでフィルムグレインノイズを含む画像を超解像したものとみなす。こうすることでフィルムグレインノイズの特徴を最大限に残しながら超解像を行うことができる。

4. 実験

4.1 実験方法

フィルムグレインノイズを含んだ画像を用意するのが最も好ましいが、その画像の多くが映画の一部分の切り出しであり、著作権の問題がある。

フィルムグレインノイズはある一定の法則で発生しているため、そこで本研究ではグレインノイズの発生を再現するプログラムを利用し著作権フリーの画像（図3）にフィルムグレインノイズを合成したものを元画像とする（図4）。



図3 元画像[3]



図4 グレインノイズを合成した画像

4.1 実験結果



図6 提案手法



図7 Real ERSGAN

Real ESARGAN ではフィルムグレインノイズが消え非常にのっぺりしているが（図7）、本研究の手法では元のグレインノイズを残しつつ超解像されていることがわかる。（図6）

さらにフィルムグレインノイズは高周波成分が主であるため、画像をハイパスフィルターに通して画像の高周波成分を比較することにより、画像のフィルムグレインノイズがどれほど残っているのか、またフィルムグレインノイズの分布がどのようになっているかを調べることでできる。本研究の提案手法と従来手法である Real ERSGAN が出力した画像を図8,9に示す。



図8 提案手法



図9 Real ESARGAN

図8,9より本研究の提案手法は従来手法である Real ESARGAN よりも多くのフィルムグレインノイズが残っていることがわかる。

以上の結果より本研究の提案手法はフィルムグレインノイズを含んだ画像の超解像において従来手法より優れていることがわかる。

5. 結論

本研究では Real ESARGAN などのとは違い、フィルムグレインノイズ部分と元画像とで2つに分離しそれぞれ分けて超解像を行うことがグレインノイズを含む画像に対しての超解像に有効であることがわかった。

今回では画像に対して研究であるが、本来の目的であるフィルムグレインノイズを含んだ動画の超解像について1フレームごとに超解像するには効率が悪いいため、本研究手法をもとに動画用に効率よく超改造できる方法がないかを検討する必要がある。

参考文献

- [1] Realistic Film Grain Rendering
<http://www.ipol.im/pub/art/2017/192/>
- [2] Image Super-Resolution Using Deep Convolutional Networks
<https://arxiv.org/pdf/1501.00092.pdf>
- [3] FDDB - Face Detection Data Set and Benchmark
<http://vis-www.cs.umass.edu/fddb/>

2023 年度 卒業論文

フィルムグレインノイズを含んだ画像の超解像に関する研
究

Research on super-resolution of images containing film grain
noise

指導教員 渡辺裕 教授

提出日：2023 年 8 月 2 日

早稲田大学 基幹理工学部 情報通信学科

1W193081-6

張歌

目次

第1章 序論	1
1.1 研究背景	1
1.2 研究目的	1
1.3 本論文の構成	1
第2章 関連研究	2
2.1 まえがき	2
2.2 Realistic Film Grain Rendering	2
2.3 超解像における従来手法	2
2.3.1 ESRGAN	2
2.3.2 SRCNN	3
2.3.3 Texture Hallucination for Large-Factor Painting Super-Resolution	3
2.5 むすび	3
第3章 提案手法	4
3.1 まえがき	4
3.2 提案手法	4
3.3 むすび	5
第4章 実験結果と考察	6
4.1 まえがき	6
4.2 データセット	6
4.3 評価指標	6
4.4 実験結果	7
4.5 考察	11
4.6 むすび	11
第5章 結論と今後の課題	12
5.1 結論	12
5.2 課題	12
謝辞	13
参考文献	14
図表一覧	15

第1章 序論

1.1 研究背景

近年、4K モニターやテレビなど高解像度のディスプレイを持っていることが珍しいことでもなくなった。その影響で映像コンテンツをより良い画質で楽しみたいという需要が増えたが、昔販売された映画を高解像度のディスプレイで視聴すると映像がボケ、せつかくの高解像度も意味をなさなくなってしまう。よって高解像度の環境においても映画をきれいに楽しめるようにと昔の映画を高解像度の環境でも耐えうるようにリマスターすることが増えた。しかし元の映像となるフィルムが何らかの理由で消失してしまった場合、すでに販売されている映像でリマスターする必要がある、そこで超解像技術の出番が出てくる。

1.2 研究目的

古いフィルム映画にはフィルムグレインノイズと呼ばれる独特の丸い形状をしたノイズがある。一部の映画愛好家ではそのノイズが映画における重要な構成要素の一つであると考えており、映画が映画たらしめる要素の一つであると考えている。だが従来の超解像技術ではそのフィルムグレインノイズをただのノイズとして除去してしまい、結果として画がのっぺりとし、映画らしさがなくなってしまう。

本研究では映画を超解像する際にのっぺり感が出てしまう要因の一つであるそのフィルムグレインノイズの消失を防ぎ、フィルムグレインノイズを残したまま映画を超解像行うことを目的としている。ただし本研究では映画の動画ではなく映画の一フレームを切り取った画像にたいして主眼を置く。

1.3 本論文の構成

本論文の構成を以下に示す。

第1章は本章であり、本論文の研究背景、研究目的について述べる。

第2章では本研究で用いる関連技術について述べる。

第3章では本研究での提案手法について述べる。

第4章では実験の結果及び考察について述べる。

第5章では本論文の結論及び今後の課題について述べる。

第2章 関連研究

2.1 まえがき

本章では、フィルムグレインノイズ発生原理である Realistic Film Grain Rendering, 従来手法である ESRGAN、SRCNN などについて説明する。

2.2 Realistic Film Grain Rendering

フィルムグレインノイズは映像を銀フィルムに現像する際に起こる化学反応により発生する。[1]その発生するフィルムグレインノイズの分布は正規分布によるものである。この発生原理をプログラムで再現した。このプログラムを利用しフィルムグレインノイズを合成した画像を図 2.1,2.2 に示す。



図 2.1 合成前の画像[5]



図 2.2 グレインノイズ合成後の画像

2.3 超解像における従来手法

2.3.1 ESRGAN

SRGAN は敵対的生成ネットワークを利用した低解像度の画像から高解像度の画像を出力するためのニューラルネットワークの一つである。[2]生成器と識別器の2つのニューラルネットワークを軸としており、生成機が生成した超解像画像と本物の画像とを識別機で比較し学習することでより自然な超解像を行うことができるようになった。

2.3.2 SRCNN

SRCNN は画像の一部をピックアップし、非線形マッピングを通じて低解像度の画像と高解像度の画像との間でそれぞれのピクセルがどのように変換されているかを学習したニューラルネットワークである。[3]学習する際、評価関数として PSNR を用い、この値が大きくなるように学習を行う。構造が単純であるため学習の速度が速くかつ従来の超解像技術と比べ品質が比較的高い。

2.3.3 Texture Hallucination for Large-Factor Painting Super-Resolution

この関連研究は超解像の際にテクスチャが失われてしまう、また不快なアーティストが生じてしまうのを防ぎながら超解像する研究である。[4]

この従来手法では高い周波数成分のテクスチャの再構築に重点を置き、新しい損失関数を定義し超解像した画像と元画像との間の損失を最小化になるようにネットワークを学習する。

2.5 むすび

本章では本研究で参考にしたフィルムグレイノイズと超解像に関する関連研究について述べた。

第3章 提案手法

3.1 まえがき

本章では、従来の超解像とは異なるアプローチでフィルムグレイノイズを含んだ画像を超解像することについて述べる。

3.2 提案手法

画像を **end-to-end** で学習し新たなニューラルネットワークを構築することが考えられるが、フィルムグレイノイズの成分が大量に存在するため、一つ一つのフィルムグレイノイズの特徴を残しながら超解像するのは非常に難しい。パラメータの調整に失敗すると図 3.1 のようにフィルムグレイノイズ同士が結合しもやもやした画像になってしまう、またはフィルムグレイノイズが消えてしまう。

そこで関連研究で述べた元画像と元画像を縮小した画像とでマッピングを行い、ネットワークを構築することでテクスチャを保持しながら超解像を行う **Texture Hallucination for Large-Factor Painting Super-Resolution** が考えられたが、高品質なデータセットを用意できない都合上、低品質なデータセットでも学習可能なネットワークを考案する必要があった。

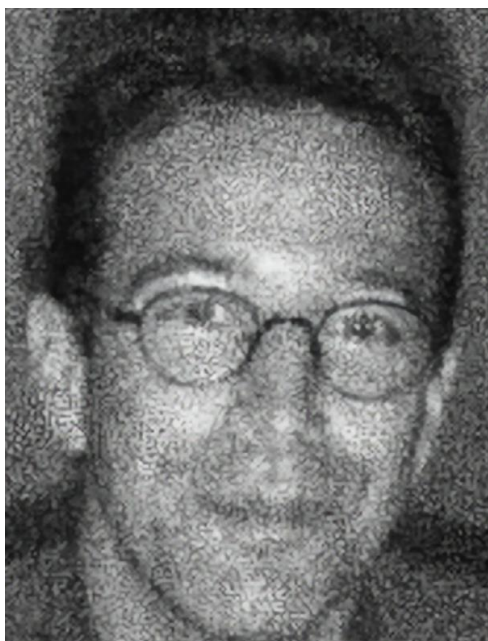


図 3.1 超解像の失敗例[5]

そこで本研究ではフィルムグレインノイズは高周波成分であるというところに着目した。高周波成分であるため低解像度においてもフィルムグレインノイズを取り出しやすい。この点を活かし、元画像と縮小した画像とでフィルムグレインノイズの部分を取り出し、画像を超解像した際にフィルムグレインノイズがどのように分布しているかを学習しフィルムグレインノイズの分布専用のニューラルネットワークを構築する。

上記の考えを適用した本研究の提案手法を図 3.2 に示す。フィルムグレインノイズを含んだ画像を超解像する際には、まず画像からフィルムグレインノイズを抽出し、フィルムグレインノイズがない画像とフィルムグレインノイズのみの画像に分け、そのフィルムグレインノイズをフィルムグレインノイズの分布専用のニューラルネットワークに通して新たなフィルムグレインノイズの分布を生成、そしてフィルムグレインノイズがない画像を超解像し、その2つを合成することで超解像した画像が出来上がる。

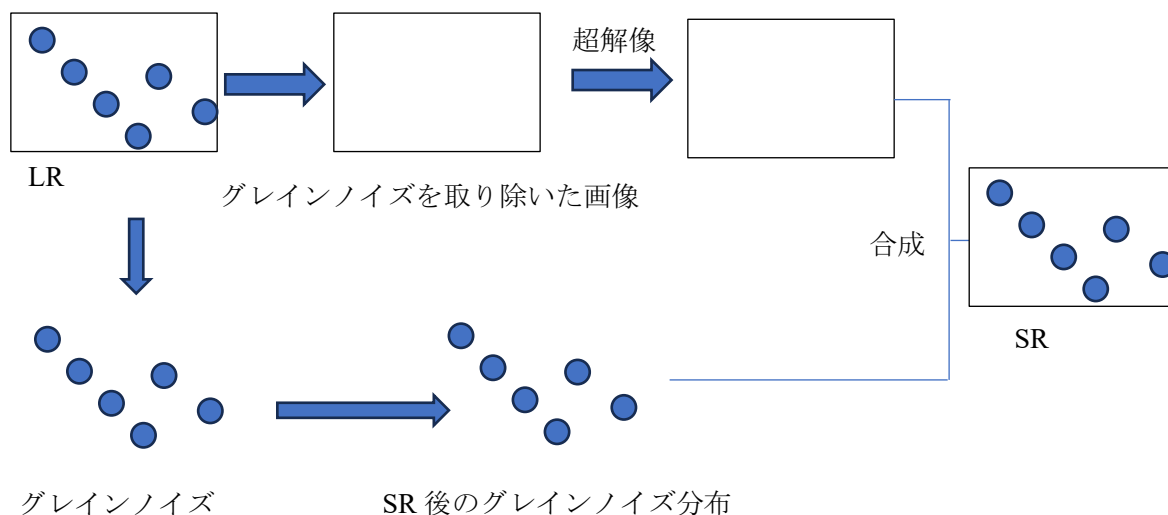


図 3.2 提案手法の概要図

超解像した後のフィルムグレインノイズの分布の学習ネットワークとして SRCNN の技術を応用した。SRCNN のピクセルがどのように変換されているかを学習したニューラルネットワークの構造がフィルムグレインノイズの分布の学習に似ており、適していると考えたためである。

3.3 むすび

本章では超解像に関する新たなアプローチについて述べた。

第4章 実験結果と考察

4.1 まえがき

本章では、提案手法に基づく実験の概要、結果及び考察について述べる。実験では画像の周波数成分を分析することで評価を行った。

4.2 データセット

フィルムグレインノイズを含んだ画像を用意するのが最も好ましいが、その画像の多くが映画の一部分の切り出しであり、著作権の問題がある。フィルムグレインノイズはある一定の法則で発生しているため、そこで本研究ではフィルムグレインノイズの発生を再現するプログラムを利用し著作権フリーの画像にフィルムグレインノイズを合成したものを元画像とする。学習用かつ実験用の元画像には Face Detection Data Set and Benchmark[5]の画像データセットを用いた。

4.3 評価指標

画像の評価指数として SSIM や PSNR が有名であるが、フィルムグレインノイズの有無が SSIM や PSNR の数値に大きな影響を与えなかったため本研究の評価指数として不適切であると考えた。そこで 3.2 提案手法で述べた通り、フィルムグレインノイズは高周波成分であるため、まず画像を離散フーリエ変換し周波数成分に分解する。そしてハイパスフィルターを通して低周波成分をカットし高周波成分のみを比較することで、フィルムグレインノイズがどれほど残っているのかを定量評価することができる。画像をハイパスフィルターに通した例を図 4.1,4.2 にしめす。



図 4.1 グレインノイズを含んだ元画像[5]



図 4.2 ハイパスフィルターに通した結果

4.4 実験結果

元画像、本研究の提案手法で超解像した画像,従来手法である RealESRGAN で超解像した画像の3セットを図4.3~図4.12に示す



Original/PSNR



提案手法/ 24.91



RealESRGAN/ 26.56

図 4.3[5]



Original/PSNR



提案手法/ 26.51



RealESRGAN/ 29.08

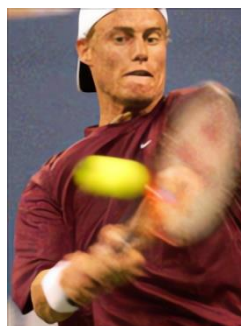
図 4.4[5]



Original/PSNR



提案手法/ 26.77

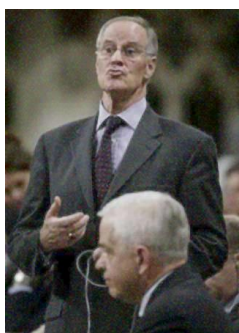


RealESRGAN/ 29.87

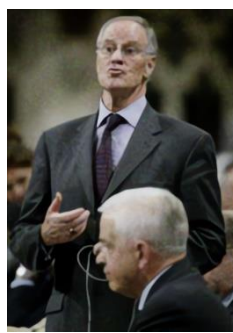
図 4.5[5]



Original/PSNR



提案手法/ 25.38



Real ESRGAN/ 30.82

图 4.6[5]



Original/PSNR



提案手法/ 29.63



Real ESRGAN/ 29.40

图 4.7[5]



Original/PSNR



提案手法/ 25.10



Real ESRGAN/ 29.19

图 4.8[5]



Original/PSNR



提案手法/ 25.11



Real ESRGAN/ 29.87

图 4.9[5]



Original/PSNR

図 4.10[5]



提案手法/ 26.87



Real ESRGAN/ 29.21



Original/PSNR

図 4.11[5]



提案手法/ 24.30



Real ESRGAN/ 31.08



Original/PSNR

図 4.12[5]



提案手法/ 27.50



Real ESRGAN/ 29.76

さらにフィルムグレイノイズは高周波成分に値するため、画像をフーリエ変換しハイパスフィルターに通し、高周波成分がどのように分布しているのかを可視化することでフィルムグレイノイズがどの程度残っているかを客観的に知ることができる。本研究と従来手法である Real ESRGAN が出力された画像をハイパスフィルターに通した画像を図 4.13 に示す。



Original/PSNR

提案手法/ 21.18

Real ESRGAN/ 20.20

図 4.13[5]

また画像を2次元フーリエ変換にかけ、高い周波数帯域において輝度の合計を計算することにより、フィルムグレインノイズがどれほど残っているか数値化することができる。この結果を図 4.14～図 4.16,表 1 に示す。

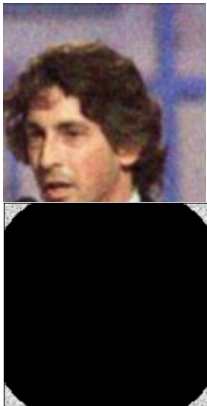


図 4.14 元画像[5]



図 4.15 提案手法



図 4.16 Real ESRGAN

	輝度の合計
元画像	50597
提案手法	44071
Real ESRGAN	43577

表 1 各手法における高周波成分の輝度の合計

4.5 考察

図 4.3～図 4.12 より Real ESRGAN ではフィルムグレインノイズが消え非常にのっぺりしているが、本研究の提案手法では元のフィルムグレインノイズを残しつつ超解像されていることがわかる。しかし多くの画像で PSNR の値について、提案手法は Real ESRGAN と比べ無視できないほど低かった。これは提案手法の最終段階である出力されたフィルムグレインノイズと画像との合成でフィルムグレインノイズの色が再現しきれていないことが考えられる。

また画像の高周波成分を比較すると図 4.13 より本研究が出力された画像が Real ESRGAN より多くのグレインノイズが残されていることがわかる。さらに PSNR 比でも提案手法のほうが従来手法と比べて高い。これはフィルターに通すことで色に関係なく純粋にグレインノイズの分布の比較ができたからと考えられる。また表 1 より高周波成分における輝度の合計が提案手法のほうが従来手法よりも元画像により近いことから、定量的に提案手法が従来手法より優れていることがわかる。

以上の結果より画像をフィルムグレインノイズとフィルムグレインノイズを取り除いた画像とで分けて超解像することがフィルムグレインノイズを残したまま超解像する点において一定の有効性があることがわかった。

4.6 むすび

本章では、従来手法と本研究の手法との超解像した画像の比較、そして従来手法では残すことができなかったフィルムグレインノイズを本研究の手法で残すことができたことについて述べた。

第5章 結論と今後の課題

5.1 結論

本研究では画像をフィルムグレインノイズとフィルムグレインノイズを取り除いた画像とで分けて超解像することがフィルムグレインノイズを残したまま超解像する新たな手法を提案した。その結果、従来手法と比べ多くのフィルムグレインノイズを残すことに成功し、映画らしさを失うことなく超解像することができた。超解像技術は画像の種類によって得意不得意が分かれるため、それぞれの得意分野で超解像を行う有効性があるということがわかった。

5.2 課題

本研究では主に画像に対しての研究を行ったが本来の目的である映画、つまり動画に対しての超解像にたいして有効な手法を提示することができなかった。動画ではフィルムグレインノイズが常に動いており、また本研究の手法では動画の1フレームごとに超解像するしか方法はなく非常に効率が悪い。さらに色の再現性に難があった。よってフィルムグレインノイズの動きや色を考慮し、本研究の手法を応用した動画に対して有効な手法を考える必要がある。

謝辞

本研究にあたり，熱心かつ丁寧にご指導をいただいた渡辺裕教授に感謝いたします。

また，常日頃から相談に乗ってくださり，研究に関する助言やご指摘をくださった研究室の皆様にお礼申し上げます。

最後に，私をここまで育ててくださり，精神面や金銭面で多くの力添えをしていただいた家族に心より感謝致します。

参考文献

- [1] Image Super-Resolution Using Deep Convolutional Networks
<https://arxiv.org/pdf/1501.00092.pdf>

- [2] ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks
<https://arxiv.org/pdf/1809.00219.pdf>

- [3] Realistic Film Grain Rendering
<http://www.ipol.im/pub/art/2017/192/>

- [4] Realistic Film Grain Rendering
<https://arxiv.org/pdf/1912.00515>

- [5] FDDB - Face Detection Data Set and Benchmark
<http://vis-www.cs.umass.edu/fddb/>

図表一覧

図 2.1	合成前の画像.....	2
図 2.2	フィルムグレインノイズ合成後の画像.....	2
図 3.1	超解像の失敗例.....	4
図 3.2	提案手法の概要図.....	5
図 4.1	フィルムグレインノイズを含んだ元画像.....	6
図 4.2	ハイパスフィルターに通した結果.....	6
図 4.3~図 4.12	実験結果.....	7
図 4.13	ハイパスフィルターを通した結果.....	10
図 4.14	2次元フーリエ変換かけた後の元画像における高周波成分.....	10
図 4.15	2次元フーリエ変換かけた後の提案手法における高周波成分.....	10
図 4.16	2次元フーリエ変換かけた後の Real ESRGAN における高周波成分.....	10
表 1	各手法における高周波成分の輝度の合計.....	10