

修士論文要旨 (2021 年度)

短時間離散フーリエ変換を用いた光線空間の雑音除去法

Light Field Image Denoising

Based on Short-Time Discrete Fourier Transform

電気電子情報通信工学専攻 菅野 友彦

1. まえがき

近年、三次元空間の視覚的な情報を、空間を伝搬する光線の情報として表現する方法である、光線空間法 (Light Field) が注目されている。この方法を用いることで、より高画質の自由視点画像や可変焦点画像の生成、物体との距離の高精度な測定、よりリアルな VR (Virtual Reality) 空間の実現などへの応用が期待されている。しかし、この光線空間のデータを取得するためにはライトフィールドカメラなどの機械や大量の写真が必要で、原理上、目に見えるレベルでの雑音が避けられないという問題がある。

光線空間画像の雑音除去には、動画の雑音除去に用いられる手法を拡張して適用することが少なくない。本研究では、齊藤ら [1] が動画の雑音除去のために用いた三次元平均値分離型短時間離散フーリエ変換 (3-D MS²T-DFT) 法を参考にしている。これは原画像の信号はガウス分布などの簡単なモデルで表すことができるという理論 [2] のもと、雑音付き動画から原画像の信号のパラメータを推定し、動画の復元を行う手法である。

しかし、実際に適用して高精度な雑音除去を行うためにはあらかじめ設定しなくてはならないパラメータ数が多く、それらを動画のサイズや特徴に合わせて適切に設定する必要がある。

本論文では、それらのパラメータのうち、雑音除去精度に大きな役割を果たすと考えられる雑音信号の高周波部分指定用パラメータを、雑音付き動画の雑音信号の分散値から自動的に推定し、

雑音処理を行う手法を提案する。これにより、手当たり次第に一つ一つ行っていたパラメータ設定作業を効率化でき、時間の大幅短縮を実現できる。

2. 光線空間

まず、本論文で扱っている光線空間と雑音の原因についてまとめる。光線空間の情報を取得する原理の多くは図 1 のようにモデル化することができる。平行な 2 平面において、1 平面がカメラ (センサ) の存在する面 ($s-t$ 面)、もう 1 平面が光線の来た方向を特定するためのスクリーン (レンズアレイ) 面 ($u-v$ 面) となる。それぞれの平面において光線が通過した点の座標を取得することで、四次元の光線空間の情報 $L(s, t, u, v)$ を得ることができる。これが、光線空間が四次元情報として扱われる理由である。

光線空間情報取得の際、雑音がのってしまう原因は、センサ存在面の 1 点に入る光を、光が来た方向を表すレンズアレイ面のすべての点で分け合うからである。この結果、一つの光線情報あたりの光の量は小さくなり、解像度も下がる。

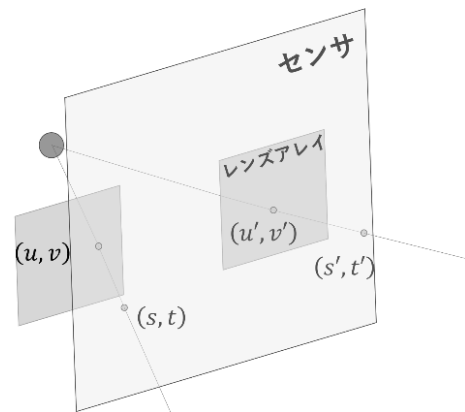


図 1 光線空間のモデル

3. 提案手法

本論文では、雑音除去において設定する必要があるパラメータのうち、雑音除去精度に大きな役割を果たすと考えられる雑音信号の高周波部分指定用パラメータを、雑音付き動画像に付加されている雑音信号の分散から自動的に推定し、効率よく雑音処理を行う手法を提案する。

3.1 高周波部分指定用パラメータ

図2に3-D MS²T-DFT法の手順を示す。高周波部分指定用パラメータは、三次元DFT係数から雑音信号を推定する際に高周波部分を選択するために用いる。DFT係数の高周波部分には原画像信号がほぼ存在せず、雑音信号のみが存在する。そこで、①の部分でパラメータ値を利用し、取得した値から雑音信号を推定して、②の部分で雑音信号部分のDFT係数を小さくすることで雑音信号と原画像信号を分離する。パラメータ値は $[d_1, d_2, d_3]$ の3値で1組となっており、それぞれ縦・横・時間方向の値を表している。

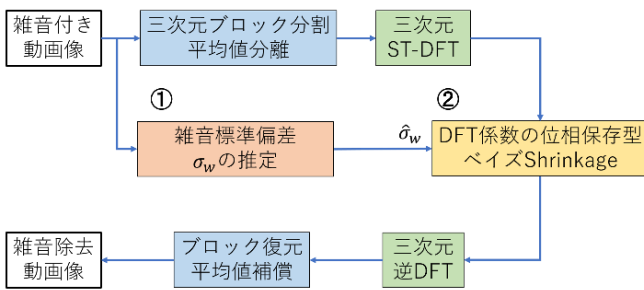


図2 3-D MS²T-DFT法の手順

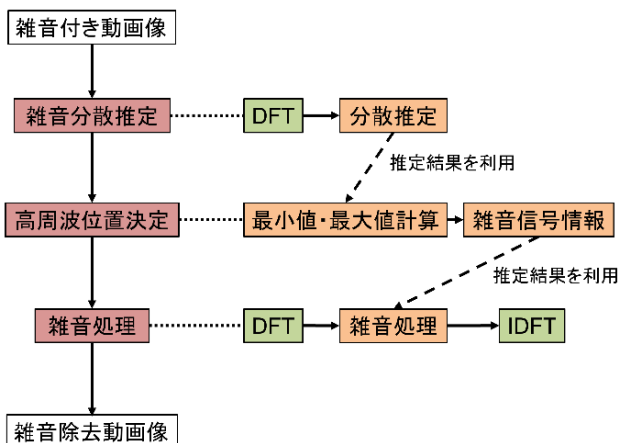


図3 提案手法の流れ

3.2 冗長色変換

本提案手法では、取り出したDFT係数に対して冗長色変換[3]し、4色成分値で処理を行っている。変換は以下の式で行っている。

$$\begin{bmatrix} O_1 \\ O_2 \\ O_3 \\ O_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} r \\ g \\ b \end{bmatrix}$$

3.3 提案手法の適用手順

提案手法の流れを図3に示す。本提案手法では、大きく分けて3つのステップによって雑音を処理していく。以下にその順番を示す。

- ① σ^2 値推定
- ② d 値推定
- ③ 雑音処理

各ステップの手順について示す。

①雑音が付加している動画像に対して処理ブロックサイズを動画サイズと同じにして三次元DFT処理を行う。高周波部分指定用パラメータを用いて画像信号のDFT係数を取り出し、雑音信号を推定することで雑音信号の分散推定値 σ^2 を取り出す。このとき、冗長色変換により1組のパラメータ値 d に対して4つの対応する分散推定値を取得できる。この動作をパラメータの組み合わせを変えて繰り返す。本提案手法では、 d 値を高周波範囲に絞って推定している。例として、 $256 \times 256 \times 64$ の動画サイズの場合、 $[d_1, d_2, d_3] = [97 \sim 129, 97 \sim 129, 32 \sim 33]$ の範囲で変更し、2178回繰り返した範囲で分散値を推定している。

②4つの分散推定値 σ^2 のうち、最小値を取り出す。この操作を各パラメータ値で行い、その中で分散推定値 σ^2 が最大のものを選ぶ。この最大値に対応する d 値1組を自動推定結果とする。これで d 値を適切に推定することができる。

③②で自動推定した値を用い、3-D MS²T-DFT法を適用する。まず、高周波指定用パラメータ値を用いて雑音の標準偏差値 σ_w を再度推定する。次に動画像の全画素値の平均値をとり、各画素値から平均値を差し引く。三次元DFTを行い、冗長色変換したうえで1次モーメント(絶対値)・2次

モーメント・4次モーメントを計算する。 σ_w と 2次モーメント・4次モーメントから位相保存型 Bayes Shrinkage を行い、復元 DFT 係数を計算する。最後に逆 DFT 処理を行い、差し引いた分の平均値を加える。

従来法と異なり、光線空間で適用する際はできるだけ処理ブロックを大きくする方がよい。そのため、処理ブロックサイズを動画像のブロックサイズと同じとし、ブロックをスライドして繰り返し冗長に計算する処理は行わない。

4. 実験

提案手法の有効性を示すために実験を行った。実験は、数値計算ソフトウェア MATLAB を使用し行った。動画は「Mandrill 16」, 「Mandrill 64」, 「Mandrill 128」, 「Chess16」を用意し、実験に用いた。「Mandrill」は画像処理分野で一般的に使用される標準画像データベース SIDBA(Standard Image Data-Base)のうちの一つであり、実験では「Mandrill」の1枚の画像を時間ごとに水平に動かすことで動画像を作成した。末尾の数字はそれぞれの画像の時間方向のフレーム数を表す。また、「Chess」はスタンフォード大学が提供する「The (New) Light Field Archive」のうちの一つである。また、雑音は用意したカラー画像に対して白色ガウス雑音を付加した。その設定値である σ 値は 24 を採用した。動画像復元の評価については、実際に復元した動画像と雑音付き動画像を比較した際の目視による見た目と PSNR 値を用いて性能を評価した。

4.1 σ^2 - PSNR の関係

まず、雑音信号分散推定値 σ^2 と復元性能値 PSNR の関係をグラフ化して確認した。雑音が載った「Mandrill 64」の動画像に対し、ブロックサイズが $16 \times 16 \times 16$ の場合と動画サイズの場合で実験し、縦軸を PSNR 値、横軸を σ^2 値にした際のグラフを図4, 5に示す。なお、紙面の都合上、色成分値1のグラフのみ示す。

ブロックサイズが 16 のときは σ^2 値が小さくなるほど PSNR 値が大きくなったが、ブロックサイ

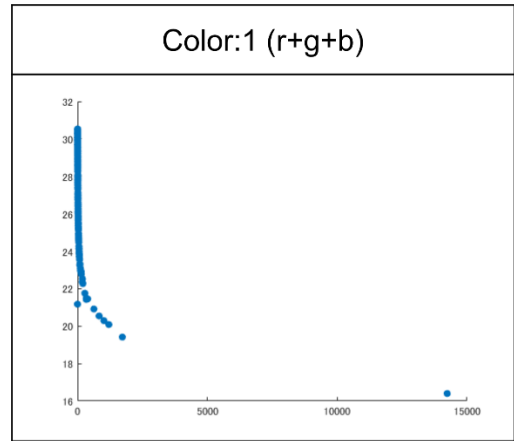


図4 「Mandrill 64」 ブロックサイズ 16

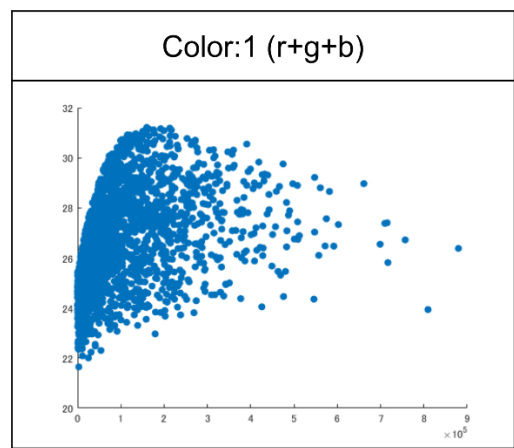


図5 「Mandrill 64」 ブロックサイズ=動画

ズを大きくすると σ^2 値が 0 より大きいところで PSNR 値が最大となり、全体的に値が散らばった。この結果より、 σ^2 値と PSNR 値には関係があり、ブロックサイズ=動画サイズのときは本来存在する σ^2 値に近いところで PSNR 値が高くなっていると考えられる。

4.2 σ^2 による推定法検証

次に、ブロックサイズ=動画サイズのときに適切にパラメータ値 d を設定するための最適な σ^2 値の取り方を検証した。 d 値1組あたり4つの σ^2 値の平均値・最小値をそれぞれ計算し、縦軸を PSNR 値、横軸を σ^2 にした際のグラフを図6に示す。

4つの σ^2 値のうち平均値を取ると羽根状の散らばった分布になったが、最小値を取ると σ^2 値が大きいほど PSNR 値が高い値に集まった。この結果より、4つの σ^2 値のうち最小値を取ってそのなか

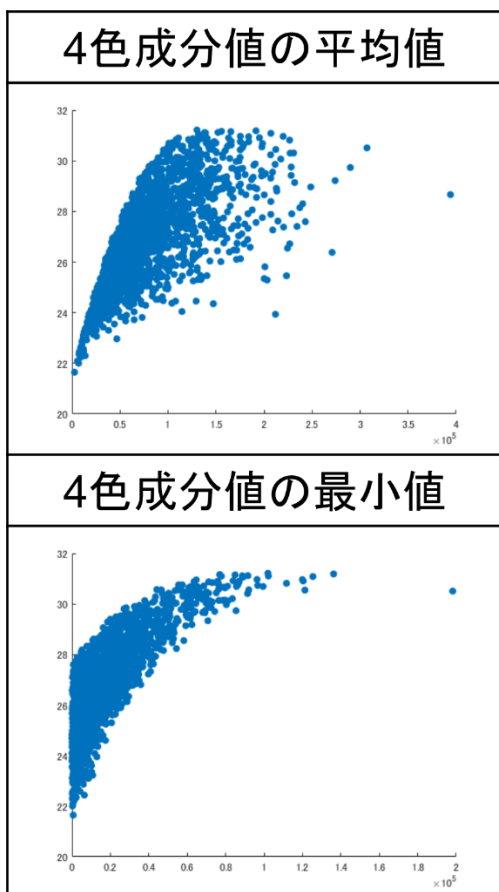


図6 「Mandrill 64」 ブロックサイズ=動画

ら最大値を取り、対応する d 値を採用する手法が適切であることがわかった。

4.3 復元精度検証

最後に、4つの σ^2 値のうち最小値を取って d 値を自動推定して雑音処理を行う提案手法と、一つ一つ d 値を変えて行う従来法の雑音除去性能を比較した。結果を表1に示し、従来法の復元画像と提案手法の復元画像（どちらも「Mandrill 64」）を図7に示す。

結果の値を比較すると、提案手法では大幅に時間短縮ができていたことがわかった。また、復元精度の差は1 dB程度であった。実際の画像で比較

表1 提案手法と従来法の結果比較

動画	自動推定結果			実際(想定最大値)		
	パラメータ	PSNR[dB]	時間	パラメータ	PSNR[dB]	時間(1回あたり)
Mandrill16	[103,98,8]	27.87	5分	[91,129,8]	28.87	12時間(5秒)
Mandrill64	[112,120,32]	30.51	6分	[106,122,33]	31.21	8時間(13秒)
Mandrill128	[99,102,65]	32.34	12分	[118,123,64]	32.65	16時間(27秒)
chess	[171,329,8]	32.52	12分	[171,329,8]	32.52	9時間(16秒)

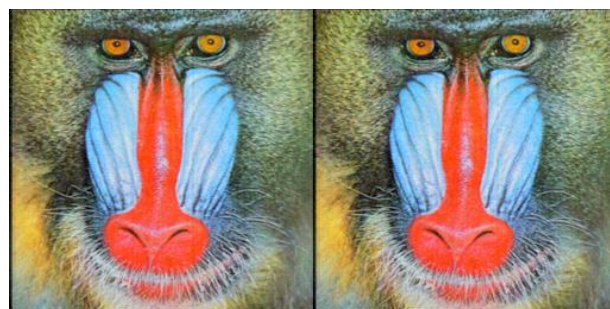


図7 復元性能比較(左:提案手法 右:従来法)

しても、提案手法と従来法では見た目の差はほとんど感じられない。この結果から、提案手法は従来法よりも大幅に時間短縮が可能であるといえる。また、実写画像「Chess」では最大値を取得できたことにより、実際の光線空間ではさらに高い精度で雑音除去が可能であるといえる。

5. おわりに

本論文では、雑音除去精度に大きな役割を果たすと考えられる雑音信号の高周波部分指定用パラメータを、雑音付き動画像に付加されている雑音信号の分散から自動的に推定し、効率よく雑音処理を行う手法を提案した。また、提案手法の検証を行った。その結果、提案手法により、手当たり次第に一つ一つ行っていたパラメータ設定作業を短縮でき、時間の効率化を実現できることが確認できた。さらに、提案手法を光線空間に適用した場合はさらに高い精度で雑音処理と画像復元が可能であろうという結果を得ることができた。

参考文献

- [1] 小松隆, 張鍵, 齊藤隆弘, "平均値分離型三次元 ST-DFT を用いた動画像復元法," 電子情報通信学会論文誌 D, Vol. J100-D, No. 9, pp. 842-845, 2017年9月.
- [2] 齊藤隆弘, 小松隆, "ランダム信号の DFT 係数のための多次元二成分混合型球対称ガウス分布モデルとそのパラメータ推定," 電子情報通信学会論文誌 D, Vol. J101-D, No. 9, pp. 1224-1236, 2018年9月.
- [3] 小松隆, 上田康隆, 齊藤隆弘, "冗長色変換を用いた Color Shrinkage 法," 電子情報通信学会論文誌 D, vol. J93-D, No.9, pp.1693-1696, Sep. 2010.