

三次元 MS²T-DFT を用いた CFA 動画データの雑音除去 CFA Video Data Denoising using 3-D MS²T-DFT

小松 隆[†] 西沢 豪[†] 中村 聡[†]
Takashi Komatsu Takeshi Nishizawa Akira Nakamura

1. はじめに

最近の動画は Bayer 方式の色フィルタ配列 (CFA) を用いた単板カラー撮像素子により撮像されることが多い。CFA 撮像された RAW データは、デモザイク処理され RGB 信号が生成される。CFA データに白色ガウス性雑音を加わる場合でも、デモザイク処理された RGB データには、デモザイク処理の影響で、空間相関および色間相関を持つ雑音が含まれるようになり、その除去が難しくなる。よって、デモザイク前の CFA データに対して直接雑音除去を行う方式が望ましい。しかし、これまで提案されているカラー動画雑音除去法は、観測された RGB 信号に雑音を加わるモデルを想定したものが多く、CFA 動画データに直接適用しても効率的な雑音除去は実現できない。

筆者らは先に、Bayer CFA 動画を直接 3 次元 DFT することで、色間の相関を活用した雑音除去が可能となることを示した[1]。本稿では、CFA 動画データに、三次元平均値分離型短時間 DFT 雑音除去法 (以下 MS²T-DFT 法) を適用することで CFA 動画データに含まれる雑音をより効率的に除去する雑音除去手法を提案する。

2. CFA 画像の DFT と雑音除去

Allaysson らは、CFA 画像の DFT により luminance 成分と chrominance 成分とが分離されて、DFT の局所領域に画像信号が集中する性質を利用したデモザイク手法を提案した[2]。

図 1 (a) は RGB 画像を示し、(b) はその Green 画像の 2 次元 DFT 画像である。一方、(c) は Bayer CFA 画像を示し、(d) は CFA 画像の 2 次元 DFT 画像である。DFT 画像は複素係数の大きさのみを明るさで示している。(b) では画像信号の電力が中央の直流付近に集中しているが、(d) では水平・垂直周波数の高い、画像の 4 辺、4 隅

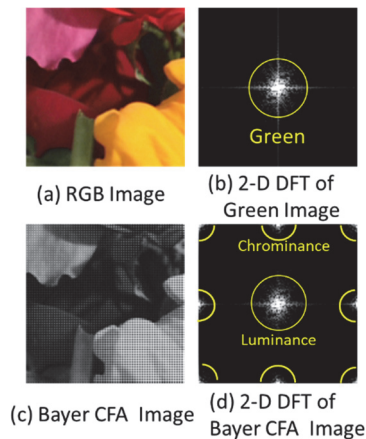


図 1 CFA の DFT

にも集中して信号が表れていることがわかる。

画像の雑音除去法の 1 つに変換領域で雑音除去する手法がある。例えば、雑音を白色ガウス性とモデル化すると、雑音の DFT 係数は統計的には全ての周波数成分に均一に分布する。一方、画像信号は直流成分付近に信号が集中する。よって、雑音を含む画像を DFT し、直流成分付近の変換を残し、DFT 逆変換することで雑音除去が行える。

Bayer CFA 画像の場合、2 次元 DFT により図 1 (d) に示すように、画像信号が複数の局所領域に集中する。この領域

の信号を残し、2 次元 DFT 逆変換することで雑音除去が可能となる。

動画の場合、フレーム間にも強い相関を有している。変換領域での雑音除去法をフレーム方向にも拡張し、3 次元 DFT 領域で雑音除去することで、フレーム間相関も利用した効率の良い雑音除去法を構築できる。

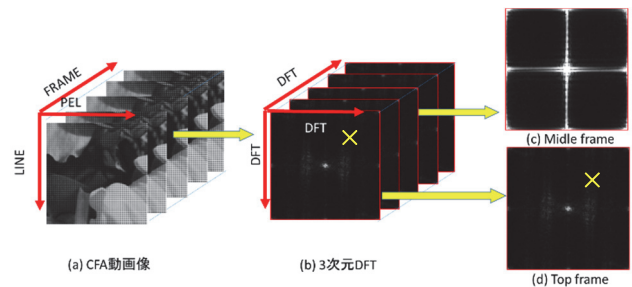


図 2 CFA 動画の 3 次元 DFT

図 2 には CFA 動画を (a) に、その 3 次元 DFT を (b) に、(b) の中央と先頭の複素係数の大きさを (c) と (d) に示した。3 次元 DFT は、空間方向に対する 2 次元 DFT 結果に、さらに時間 (フレーム) 方向に 1 次元 DFT を施すものである。モノクローム動画の 3 次元 DFT は各フレームの luminance 成分に対して時間方向に DFT するものであるが、CFA 動画の 3 次元 DFT は各フレームの luminance 成分と chrominance 成分とに対して時間方向に DFT するものと考えられる。

この結果より、時間周波数の低いフレーム付近の、局所領域に画像信号が集中している様子がわかる。この DFT 係数に位相保存型 Shrinkage を適用し 3 次元 DFT 逆変換することで、CFA 動画の雑音除去が可能なることを筆者らは文献[1]で示した。

3. 3-D MS²T-DFT による CFA 動画雑音除去

画像は時間・空間いずれの方向に対しても非定常となることが多い。動画全体を 3 次元 DFT するよりも、動画から切り出された小ブロックに対して 3 次元 DFT を適用するほうが画像信号の疎表現の観点では有利となる。一方、DFT は有限区間信号を周期性信号として解析する。動画の 3 次元 DFT の場合、周期性によりブロック境界で不連続を生じ、画像信号に由来しない複数の DFT 係数値が生じてしまう。ブロック境界での不連続の影響を回避する目的で ST-DFT が考案されている。ST-DFT は信号に対して境界付近で係数がゼロとなる重み窓をかけてから DFT するものである。ただし、直流信号に重み窓を掛けると、窓関数の形状に由来する交流成分を生じる。そこで ST-DFT を画像に適用する場合、切り出したブロックの平均値を差し引いておくことで、窓関数の影響が軽減され、動画信号のより疎な表現が可能となる。

筆者らは先に、3-D Mean-Separation-type Short Time-DFT (以下 3-D MS²T-DFT) を用いた動画雑音除去法を提案

した[3]。3-D MS²T-DFT 法は動画像から切り出された 3 次元ブロックに平均値分離 3 次元 ST-DFT を適用し、変換係数に位相保存型の Bayes 縮退処理を施すことで動画像の雑音除去を行う方式である。3-D MS²T-DFT 法はフレーム間のマッチング処理を用いていない。例えば 3 次元 DFT を行うと、フレーム間での平行移動は DFT 係数の位相成分のみに反映される。位相保存型の縮退処理は動き情報を保存したまま時間空間相関を利用した雑音が除去可能な方式である。

VBM4D[4]のような代表的な動画雑音除去法は、ブロックマッチングにより対応するブロックをフレーム間で求める処理が必要となっている。CFA 画像をモノクロ画像とみなしてフレーム間でブロックマッチングを行うと、CFA の影響で赤画素と青画素は画像の縦・横とも 2 画素おきに配置されているため、縦・横とも 2 画素精度での動き量しか求められず、CFA 動画雑音除去に適しても、本来の性能を十分に発揮した雑音除去方式とはならない。一方、モノクロ動画用の 3-D MS²T-DFT 雑音除去法はフレーム間でマッチング処理を行わない。モノクロ動画の場合と同様に、CFA 動画データに対してもフレーム間の動き量は DFT 係数の位相成分に反映され、雑音除去は DFT 係数の振幅成分のみに対して行われる。'フレーム間でマッチング処理を行わない'という特徴を持つ 3-D MS²T-DFT 雑音除去法は、CFA 動画像に恰好の雑音除去法といえる。

図 3 に提案法の概要を示す。CFA 動画像からサイズ $M \times M \times N$ ブロックを切り出し、3-D MS²T-DFT を適用する。DFT 係数に位相保存型 Bayes Shrinkage を適用する。3次元短時間 DFT の逆変換を行い、さらに分離した平均値の補填を行う。これらの雑音除去結果を CFA データに戻す。なお上記のブロックの切り出しは空間・時間方向とも冗長な切り出しとしている。CFA 動画像に対しては、次章で述べる雑音の標準偏差の推定手順を除き、モノクローム動画用の 3-D MS²T-DFT 法と同一処理手順で CFA 動画像に適用できる。

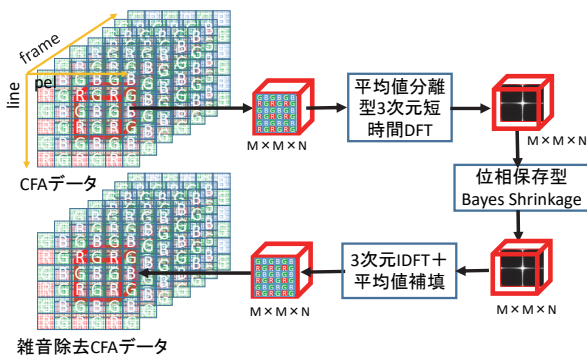


図 3 3-D MS²T-DFT による CFA 動画像雑音除去法の概念図

4. Bayer CFA 動画像中の雑音標準偏差の推定

モノクローム動画用の 3-D MS²T-DFT 法では、画像に含まれる白色ガウス性雑音の標準偏差 σ を、与えられた動画像の DFT 係数値から推定している。時空間周波数が共に高い 3 次元 DFT 係数値には画像信号成分はほとんど含まれず、主に雑音成分に起因すると考えられる。そこで動画像中から切り出された多数のブロックより得られる時空間周波数が共に高い 3 次元 DFT 係数値に、ロバスト推定を適用し雑音の標準偏差 σ を推定している。

CFA 動画像に 3 次元 DFT を適用した場合、雑音成分が大部分と見なしうる変換係数は、①空間周波数が中間で、②時間周波数が最も高い、DFT 係数である。この 3 次元 DFT 係数の位置を図 2 の(d)に黄色の×印で示した。CFA 動画像中から切り出された多数のブロックより得られる×印の位置の DFT 係数値にロバスト推定を適用し雑音の標準偏差 σ を推定する。具体的にロバスト推定は、多数のブロックからえられた×印の位置の DFT 係数値の大きさのメディアン値に 0.6745 を掛けた値を標準偏差の推定値としている。

5. 雑音除去性能の評価

提案法の雑音除去性能評価を行った。評価手順を示す。

[性能評価手順]

- ① カラー動画像の各画素を Bayer CFA 順序で間引き、CFA 動画像とした。
- ② CFA 動画像データに白色ガウス性雑音を加えた。
- ③ 雑音付加 CFA 動画像に提案法の雑音除去手順を適用し、雑音除去 CFA 動画像を生成した。
- ④ 雑音除去 CFA 動画像をフレーム毎にデモザイクし、カラー原動画像との PSNR を求めた。

シミュレーションには表 1 に示した ITE の標準カラー動画像を用いた。評価の諸元は表 2 とした。

表 1 シミュレーションに使用した動画像

画像名	使用フレーム
Walk through the Square	#600~#663
Flamingoes	#600~#663
Whale Show	#0~#63
Japanese Room	#600~#663
Woman with the Bird Cage	#600~#663
Intersection	#600~#663

表 2 評価シミュレーションの諸元

諸元	設定値
3-D ST-DFT のブロックサイズ	32画素×32ライン×16フレーム
ST-DFT の重み窓	三角窓 (32タップ, 16タップ)
ブロック切り出しストライド	水平・垂直方向4、フレーム方向2
付加雑音の標準偏差	24, 36, 48 (画素値0~255の場合)
デモザイク法	文献[5]

シミュレーションでは比較のため Bayer CFA 動画像をまずデモザイクし、得られた RGB 動画像に対して VBM4D 雑音除去法[4]を適用する方式の性能評価を行った。また、Bayer CFA 動画像をモノクローム動画像と見なしモノクロ用 VBM4D 雑音除去法を適用し、得られた雑音除去 CFA 動画像をデモザイクする方式との比較も行った。提案法と 2 種類の比較雑音除去法の流れを図 4、図 5、図 6 に示す。

VBM4D 法の適用には公開されている MATLAB のプログラムを用いた。VBM4D 法は与えられた動画像データに対

して付加雑音の標準偏差を推定し、雑音除去の強度を自動的に決定するモードを有している。しかし、デモザイキング後の動画像にこのモードを使用しても良好な雑音除去結果は得られなかった。そこで、原動画像を参照し PSNR が最も高くなる雑音除去強度を個々の動画像シーケンス毎に求め、この雑音除去強度値を性能評価に使用した。

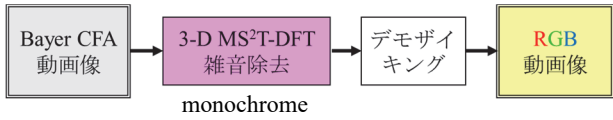


図 4 提案法による雑音除去処理の流れ①

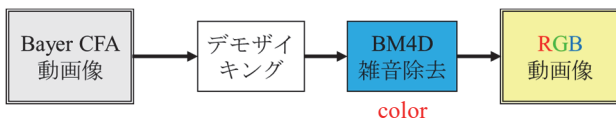


図 5 比較法 1 による雑音除去処理の流れ②

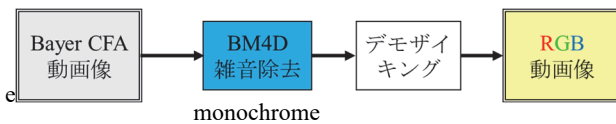
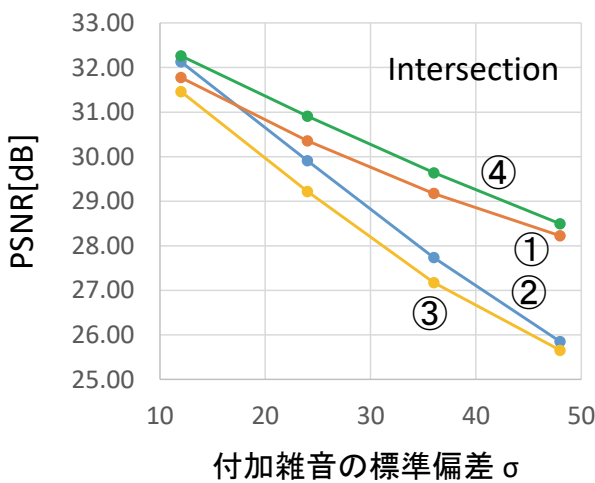


図 6 比較法 2 による雑音除去処理の流れ③

図 8 に評価結果を示した。茶色①は提案法、青色②は比較法 1、オレンジ③は比較法 2 である。この結果より、Bayer CFA 動画像に含まれる雑音が多くなるほど、デモザイキング前に雑音除去する提案法の有効性が確認できる。また、VBM4D を直接 Bayer 動画像に適用する比較法 2 では十分な雑音除去性能が得られていないことがわかる。

6. Bayer CFA 動画像の 2 段階雑音除去法

図 8 の結果は、CFA 動画像に含まれる雑音が多い場合、具体的には雑音標準偏差がおおむね 30 以上において提案法による雑音除去結果は比較法 1 の雑音除去結果を上回る



(a)

PSNR が得られている。一方、雑音がさほど多くない場合、デモザイキング後に動画雑音除去する比較法 1 のほうが有効な場合があることを示している。通常デモザイキングはフレームごとの処理であり、動画像としてのフレーム間での画素値の対応関係や相関関係などは考慮されていない。付加雑音が少ない場合、雑音のデモザイキングに及ぼす悪影響が少なく、デモザイキング後に動画雑音除去を行うことで、観測値に含まれる雑音とデモザイキングにより生じる雑音を一括して除去でき、比較法 1 が有利となっていると考えられる。

そこで、デモザイキングの前と後との 2 段階で雑音除去する方式とすることで、雑音除去性能が向上すると期待できる。Bayer CFA 動画像の 2 段階雑音除去法の処理の流れを図 7 に示す。

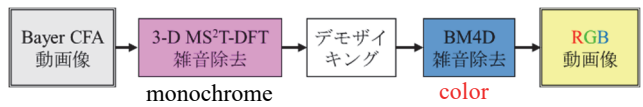


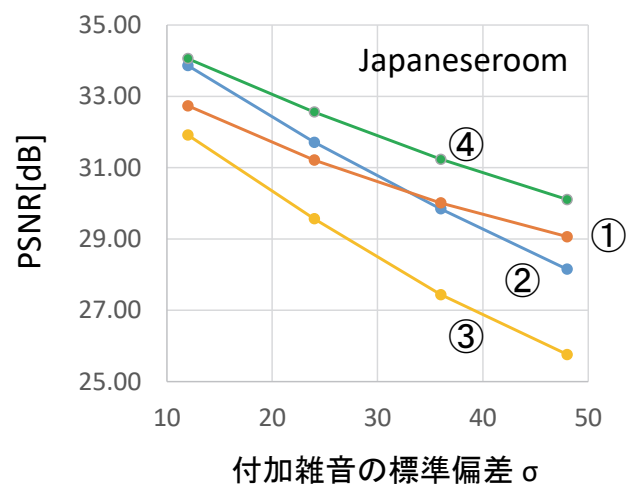
図 7 2 段階雑音除去法の処理の流れ④

まず、Bayer CFA 動画像を 3-D MS²T-DFT 法で 1 段階目の雑音を除去する。次にデモザイキング後に BM4D 法により 2 段階目の雑音除去を行う。ただし、1 段階目での雑音除去の程度は 8 割程度、すなわち推定した雑音の標準偏差の推定値を 0.8 倍して 1 段階目の雑音除去を行った。

2 段階方式の雑音除去性能を図 7 に緑色④で示した。使用したいずれの動画像においても、デモザイキングの前に 3-D MS²T-DFT を単独で用いる方式①や、デモザイキングの後に VBM4D を単独で用いる方式②よりも高い雑音除去性能を示していることがわかる。

7. おわりに

モノクローム動画像用の 3-D MS²T-DFT 雑音除去法を用いることで、デモザイキング処理前の Bayer CFA 動画像に対して直接雑音除去することが可能であることを示した。付加



(b)

図 8 様々な雑音除去法の雑音除去性能評

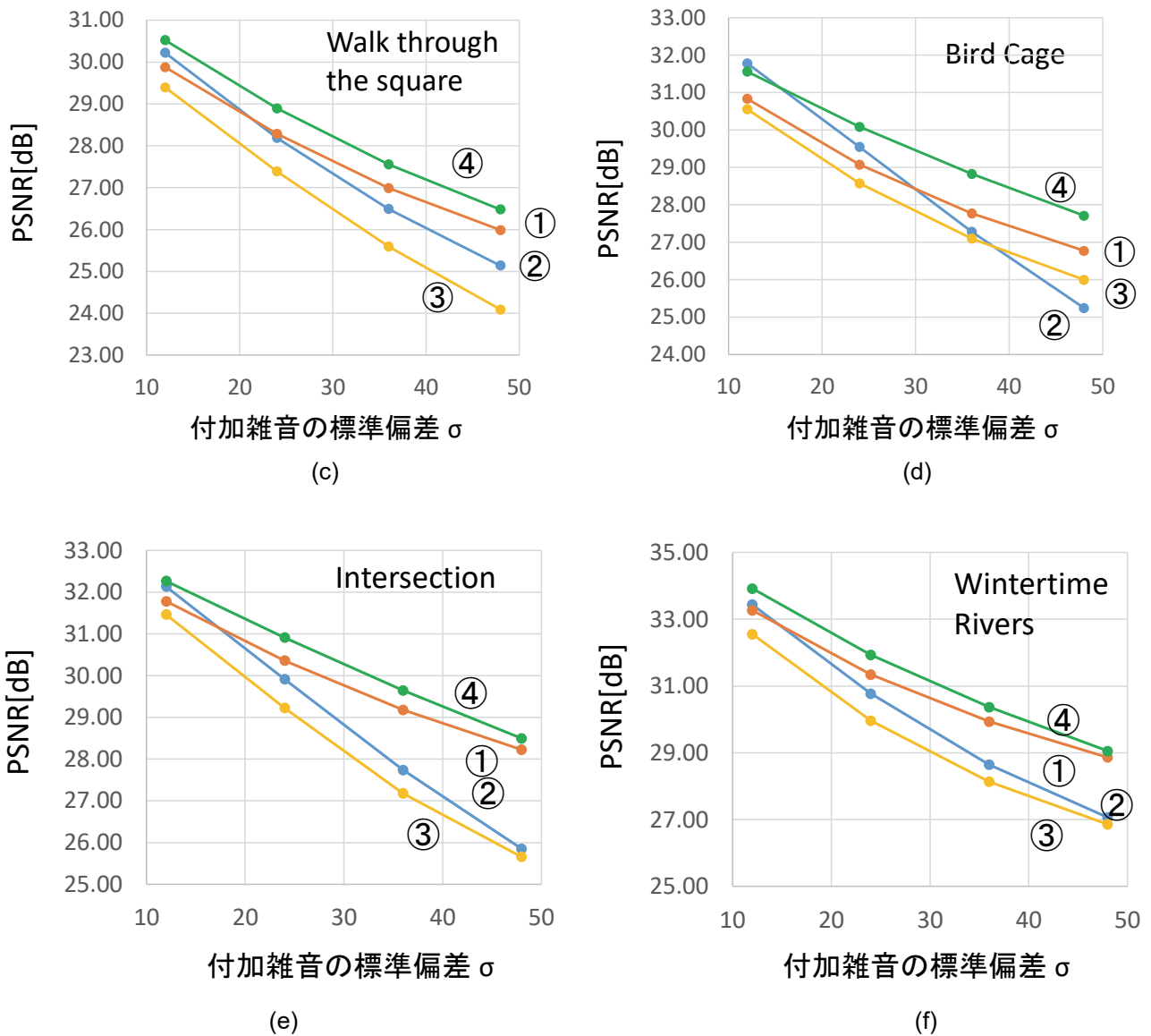


図8 様々な雑音除去法の雑音除去性能評価

雑音が大きい場合、デモザイク処理後の雑音除去よりも、デモザイク処理前の雑音除去がより有効であることをシミュレーションで確認した。デモザイク処理の前後 2 段階の雑音除去法とすることで、各々単独で雑音除去を適用する方式よりも雑音除去性能が大幅に改善することを示した。

参考文献

- [1] 小松, 齊藤, CFA 動画データの雑音除去法の検討, 電子情報通信学会ソサイエティ大会, A-8-11, Sept. 2019.
- [2] David Alleysson, et al, Linear Demosaicing inspired by the Human Visual System, IEEE Trans. on Image Process., vol. 14, no. 4, April 2005.

- [3] 小松, 齊藤, 平均値分離型三次元 ST-DFT 係数の確率分布モデルパラメータの推定と動画復元への応用”, FIT2017, I-002, Sept. 2008.
- [4] M. Maggioni, G. Boracchi, A. Foi, and K. Egiazarian, Video denoising, deblocking and enhancement through separable 4-D nonlocal spatiotemporal transforms, IEEE Trans. on Image Proc., vol. 21, no. 9, pp. 3952-3966, Sept. 2012.
- [5] Malvar, H. S., L. He, and R. Cutler, High quality linear interpolation for demosaicing of Bayer-patterned color images, ICASPP, Volume 34, Issue 11, pp. 2274-2282, May 2004.

† 神奈川大学・工学部
Faculty of Engineering, Kanagawa University