

## 予測誤差信号の振幅特性に基づく低処理コストな Adaptive Interpolation Filter 設計手法

吉野知伸 内藤整 酒澤茂之

高圧縮符号化において、動き補償予測は極めて重要な符号化技術の一つである。これまでの研究において、小数精度動き補償における補間フィルタ性能の改善手法である、適応型補間フィルタ手法(Adaptive Interpolation Filter; AIF)が提案されている。Wiener Filter をベースとしたこれらの手法は多くの処理コストを要するため、リアルタイム性が求められるコーデックへの適用は困難であった。これに対して、本研究では、少ない処理コストで AIF に準じた性能を有する補間フィルタの設計手法を提案する。提案手法では、動き補償誤差を推定し、同誤差を周波数領域で解析することにより、大きな MC 誤差を抑制するフィルタを設計する。符号化実験により、従来の AIF 手法に対して 30% の処理量で、従来方式と同等の符号化性能が得られることを確認した。

### Low complexity Scheme for Designing Adaptive Interpolation Filter Based on Amplitude Characteristic Analysis of Predicted Error Signal

Tomonobu Yoshino, Sei Naito and Shigeyuki Sakazawa

Motion-compensated (MC) prediction is one of the most effective coding tools for high-compression video coding. In former works, adaptive interpolation filter (AIF) technology, which improves the coding performance of the fractional-pel MC, has been proposed. As for the real-time codec system, the implementation of AIF based on the Wiener filter algorithm is difficult because of its high computational complexity. To overcome this problem, in this paper, we propose a low-complexity AIF scheme, which can maintain the coding performance achieved by the conventional AIF. The key technology is the efficient filter design process based on the amplitude characteristic analysis with the precise estimation of MC error signal. The experimental result confirmed that almost the same coding performance was maintained at only 30% of the complexity compared to the conventional AIF.

#### 1. はじめに

IPTV などに代表されるコンシューマ向け映像サービスでは、昨今、コンテンツの高解像化が進んでいる。現在は HDTV が主流であるが、将来的には更に高解像化し、走査線 4000 本クラスの超高精細映像(Ultra-High Definition TeleVision; UHDTV)[1,2]の普及が期待されている。また、映像サービスにおける高品位化に伴い、番組の多チャンネル化も予想される。映像の圧縮方式については、H.264 が標準方式の中で最も優れた性能を有するが、今後の映像サービスにおける映像情報量の増大を考慮すると、更に高性能な圧縮方式が求められる。

標準の符号化方式における重要な要素技術の一つに、動き補償予測が挙げられる。動き補償予測では、オブジェクトの動きに正確に追従することで、高い符号化性能を得ることができる。このため、H.264 や MPEG-2 に代表される標準符号化方式では、小数画素精度での動き補償予測が採用されている。H.264[3]では、In-Loop Filter の機能を有する Wiener Filter に基づく補間フィルタが用いられている。動き補償における最適な In-Loop Filter 特性は、絵柄に依存することが知られている。これまでの研究で、符号化性能の改善を目的として、適応的な補間フィルタ手法(Adaptive Interpolation Filter; AIF)が提案されている。文献[4-6]では、Weiner Filter に基づいて、フレーム毎に適応的な補間フィルタ係数を与える手法が提案されている。しかしながら、これらの手法では、予め MC 誤差を算出し、更に同誤差を最小化する係数を計算するため、非常に大きな処理コストが必要であった。文献[7]では、AIF における処理コストを抑制する手法が提案されている。同手法では、フィルタ係数算出過程における演算量を抑制する。しかしながら、最小化される MC 誤差を事前に得るための処理コストは依然として残されたままであり、リアルタイムコーデックなどでの実用に向けて、同処理コストの大幅な削減が求められる。

以上の背景から、本研究では、適応的な補間フィルタ設計の低処理コスト化を目的として、In-loop Filter の定性的な特徴に基づいてフィルタ係数を適応的かつ高速に決定する手法を提案する。提案手法は、Wiener Filter に基づく手法[4-7]で前提としている厳密な MC 誤差および、フィルタ係数の計算的な求解を必要としない。提案手法では、処理フレームの原画像と参照先フレームの局所復号画像から、フィルタ係数算出において最小化対象となる MC 誤差を推定する。また、原画像及び推定された MC 誤差の振幅特性に基づき、フィルタ特性を決定する。

以下に本論文の構成を示す。2 節では、AIF に関する従来手法を示す。3 節では、MC 誤差を推定するアプローチと、原画像と同誤差の振幅特性からフィルタ特性を決定するアプローチを提案する。4 節では、符号化実験により、超高精細映像に対する

符号化性能を評価する。最後に、5節で結論を述べる。

## 2. RELATED WORKS

### 2.1 Adaptive Interpolation Filter

MC+DCT 符号化方式において、MC に In-loop Filter を施すことで、予測誤差が抑制され、符号化性能が向上することが知られている。H.264/AVC では、小数精度 MC の補間フィルタに、Wiener Filter に基づく Loop Filter が採用されている。下記に、Wiener Filter に基づく補間フィルタの基本的な考え方を示す。

簡単のため、1次元信号について考える。t フレーム目における i 番目の原画像の画素値を  $I(i,t)$ 、タップ数を  $L$ 、フィルタ内の位置  $k$  の係数を  $w(k)$ 、符号化済み画像の参照により得られる t フレーム目における i 番目の予測値を  $p(i,t)$  とする。このとき、MC 誤差は式(1)で表わされる。

$$I(i,t) - \sum_{k=-L/2}^{L/2} w(k) \times p(i+k,t) \quad (1)$$

ここで、式(2)に示すような、MC 誤差の二乗和を最小化するフィルタ係数を数学的に得ることで、映像特徴に適したフィルタを設計可能である[4]。

$$\min_w \left( \sum_i \left( I(i,t) - \sum_{k=-L/2}^{L/2} w(k) \times p(i+k,t) \right)^2 \right) \quad (2)$$

### 2.2 従来研究

これまでの研究で、補間フィルタの性能を改善する手法が提案されている。文献[4]では、2-D および 3-D の適応的な補間フィルタが提案されている。しかしながら、文献[4]ではシーケンス全体で一つのフィルタ係数セットだけであり、シーン変化に追従できず、効果は限定的であった。文献[5]では、フィルタ係数をフレーム毎に適応的に更新する手法が提案されている。同手法により、時間的な映像特徴変化に追従可能なフィルタを設計可能である。これに対し、文献[6]では、画面内で AIF が有効な箇所の特化してフィルタ係数を算出する手法が提案されている。同手法により、局所的に適応的なフィルタの設計が実現されている。

文献[5,6]では、フレーム毎に、Wiener Filter に基づいて MC 誤差を最小化するフィルタ係数を数学的に求めることで、映像信号に適応的なフィルタを設計可能である。しかしながら、フレーム毎に式(2)で最小化対象となる MC 誤差を得るための処理が必要となる。この処理には、予め一旦符号化処理を行うのと同様な処理コストが必要と

なる。更に、式(2)の最小化問題を解くための処理が必要であり、同処理では特に自己/相互相関を演算する処理量が支配的となる[7]。

これらの処理には多大な処理コストを要するため、ハードウェア実装などの実用的な用途には文献[5,6]の手法は適さない。これに対し、文献[7]では、式(2)の最小化問題を解くための処理コストを抑制する手法が提案されている。同手法では、問題を解く際に有効な情報とならないブロックを予め識別し、不要な演算を抑制している。しかしながら、MC 誤差の算出処理は依然として必須であり、十分な処理コスト削減は達成されていない。

## 3. 提案方式

### 3.1 Adaptive Interpolation Filter の特徴

本節では、AIF の特性について解析的に考察する。AIF におけるフィルタの主な役割の一つとして、MC における In-loop Filter が挙げられる。In-loop Filter では、符号化歪みなどに起因する MC 誤差を予め抑制することで、MC 誤差の符号量を効果的に抑制する。In-loop Filter 処理を周波数領域で考えた場合、MC 誤差に該当する周波数成分を抑制することを意味する。抑制の対象として、原画像で大きな振幅を持つ成分を選択すると符号化性能の低下を招くため、原画像で小さな振幅を有する成分に限定すべきである。

図 1 に、ある素材における原画像の振幅  $|F(\omega)|$  と MC 誤差  $|D(\omega)|$  の比を示している。同結果は、4節で示す素材 A および符号化条件におけるものである。簡単のため、ある水平ラインにおける 1/2 画素位置の結果について議論する。図 2 に、代表的な Wiener Filter ベースの補間フィルタである、文献[6]のフィルタ特性を示す。図 1 より、中域から高域において、原画像の振幅に対して MC 誤差の振幅が大きくなる傾向が分かる。一方、図 2 より、AIF では当該帯域を抑制する特性を持つフィルタが設計されることが分かる。特に、振幅比が 1 を超える頻度が高くなる  $\pi/3$  付近以降が、阻止域もしくは遷移域となっている。図中では、閾値を点線で示している。同帯域以降では MC 誤差の振幅が原画像よりも大きく、参照先である符号化済み画像においても、当該周波数帯では誤差成分が支配的である。従って、当該周波数はフィルタにより抑制されることが求められる。

### 3.2 提案手法

#### (a) 基本アプローチ

3.1 節の解析に基づき、補間フィルタを設計する手法を提案する。本手法では、MC 誤差と原画像から抑制されるべき周波数成分を予測し、同予測に基づいてフィルタの希望特性を決定する。以下に、提案手法を構成する処理ステップを示す。

- Step 1. 処理フレームの原画像と MC 参照先フレームの符号化済み画像の差分を得る。
- Step 2. 同差分の振幅特性 $|D(\omega)|$ と処理フレームにおける原画像の振幅特性 $|F(\omega)|$ をそれぞれ求め、両者の比(振幅特性比)  $|D(\omega)|/|F(\omega)|$ を得る。振幅特性は当該信号をフーリエ変換することにより得る。
- Step 3. 振幅特性比の連続する  $M$  個の移動平均値によりフィルタ特性  $W(\omega)$ を決定する。移動平均値について、閾値  $Th$  を初めて超える周波数  $\omega_0$  と、最大となる周波数  $\omega_1$  を求める。このとき、 $(\omega_0 + \omega_1)/2$  までを通過域、それ以降を阻止域となるフィルタを設計する。
- Step 4.  $W(\omega)$ を逆フーリエ変換して、6-tap の対称なフィルタ係数  $w(k)$ を得る。

#### (b) MC 誤差推定

前節の手順では、MC 誤差を簡易的に推定するために、原画像と参照先画像の単純フレーム間差分を MC 誤差とみなしている。しかし、カメラワークなどの動きを伴う映像に対しては、同差分により MC 誤差を推定すると、推定精度が低いために動オブジェクト付近に 2 本のオブジェクト境界(エッジ)が現れる。したがって、推定された MC 誤差の振幅特性には、実際の MC 誤差には出現しない、同エッジに起因する特性も含まれてしまう。結果的に、誤ったフィルタ特性が設計される懸念がある。

以上より、推定される MC 誤差と実際の MC 誤差との乖離を抑制するため、Step1 において、単純フレーム間差分の代わりにグローバル動き補償を導入する。

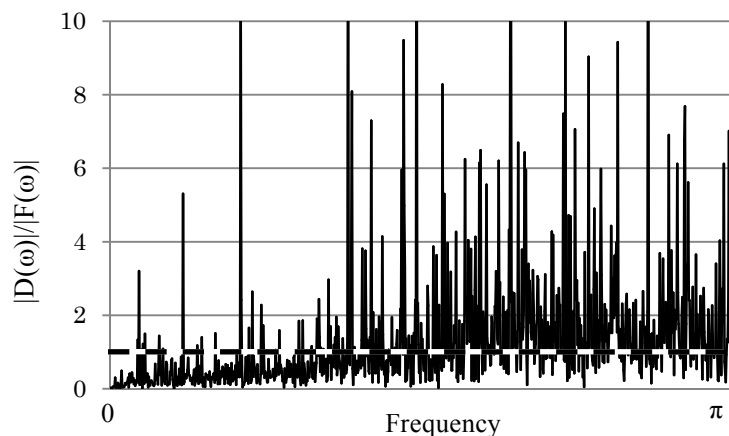


図 1 MC 誤差と原画像の振幅比

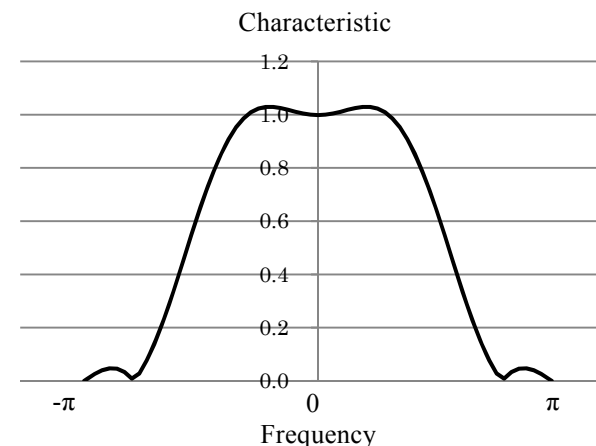


図 2 AIF フィルタ特性

#### 4. 性能評価

符号化実験により、提案手法の符号化性能を評価した。比較手法として、2.1 節に準じたアプローチであり KTA[8]に実装されている Enhanced AIF(EAIF)[9]を用いた。提案手法も KTA に実装した。評価素材として、本研究で対象とする解像度である、超高精細映像を用いた。しかしながら、実験では、KTA における入力映像の制約のため、超高精細映像(7680 画素×4320 ライン)の中央を切り出した素材を用いた。切り出した素材のフレームサイズは、超高精細映像の縦横それぞれ 1/4 に相当する 1920 画素×1080 ラインとした。表 1 に評価素材の特徴を示す。

提案方式として、以下の 3 通りのアプローチによる符号化性能を評価した。

- (a) 3.2 節(a)で示した基本アプローチ
- (b) 3.2 節(b)で示した GMC を導入したアプローチ
- (c) GMC を導入したアプローチに加え振幅特性の演算処理に高速フーリエ変換(FFT)を導入したアプローチ

手法(a)は、提案手法の基本的な性能を評価するためのものである。手法(b)と(c)は、提案方式の更なる改善として、それぞれ符号化性能改善および処理コスト削減の制御を導入したときの性能を評価するためのものである。GMC については、 $\pm 16$  画素/フレームの範囲で、整数画素精度で動きを探索した。表 2 に符号化条件を示す。なお、提案手法におけるパラメータについては、実験的に  $Th=1.0$ ,  $M=32$  とした。また、表 3 では、H.264 リファレンスソフトウェア JM[10]に対する符号化性能改善効果を示している。同評価として、BD-PSNR[11]を用いた。更に表 4 では、EAIF に対する提案手

法による AIF 処理量の削減率を示している。AIF 処理量は、全ての素材および QP における AIF 処理に要した時間の平均値で示している。なお、実験に用いた計算機の性能は、CPU が Intel Xeon X5482 3.2GHz、メモリ 64GB である。

表 1 テスト素材の特徴

素材	特徴
A	向日葵畑を映した映像。緩やかなパンワーク。風で花や葉が揺れる。
B	数匹の蝶を大きく写した映像。蝶が羽ばたいている。静止カメラワーク。
C	日本の祭りの風景。非常に大勢の人が映っている。ほぼ静止カメラワーク。

表 2 符号化条件

GOP 構造	IPPP... (I-pic は先頭のみ)
QP	27,32,37,42
フレーム数	30 枚

表 3 符号化性能(JM に対する BD-PSNR[dB])

素材	EAlF	(a)	(b)	(c)
A	0.45	0.25	0.33	0.28
B	0.35	0.27	0.29	0.26
C	0.36	0.22	0.25	0.24
平均	0.39	0.25	0.29	0.26

表 4 AIF 処理量

	EAlF	(a)	(b)	(c)
AIF 処理時間[msec]	1207.8	262.5	387.2	359.7
削減率[%]	--	78.3	67.9	70.2

表 3 および表 4 の結果より、符号化性能低下を抑制しつつ処理量削減が実現されており、手法(a)では、EAlF に対して、全素材平均で 0.14dB 程度の PSNR 低下で 78.3% の AIF 処理量が削減されている。また、手法(a)に GMC を導入した場合(手法(b))、全素材平均で PSNR 低下が 0.10dB に抑制されている。特に、素材 A のように画面全体の動きを伴う素材では、GMC により動きを容易に推定できるため、他の素材よりも手法(a)に対する PSNR 改善効果が大きい。最も高い BD-PSNR が得られた素材 A について、図 3 に R-D 曲線を示す。図 3 より、提案手法は、JM に対して EAlF とほぼ同等の符号化性能改善効果を有しており、EAlF と同様に高ビットレートで特に有効であることが分かる。ただし、GMC 処理のために AIF 処理量が増大した結果、AIF 処理量削減

率が 10.4 ポイント減少しており、高速なグローバル動き探索アプローチの必要性を示唆している。

更に、手法(c)にて、振幅特性を演算する際に FFT を導入した場合、AIF 処理量削減率が 2.3 ポイント向上しており、提案手法の更なる高速化が図られていることが分かる。手法(b)と(c)で性能に差が生じている原因は、FFT 導入に起因するものである。FFT に信号を入力するために、入力信号の長さを 2 のべき乗に拡張しており、これにより各周波数成分の振幅について手法間で差が生じる。結果的に、この違いが希望振幅特性や最終的に得られるフィルタ係数に影響する。

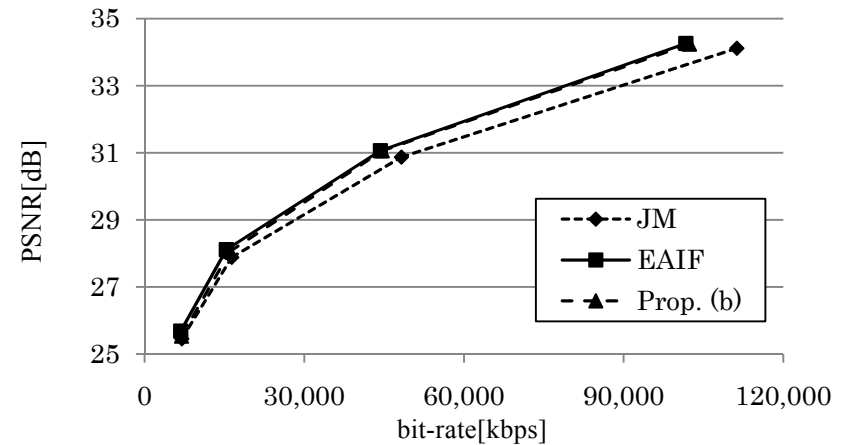


図 3 R-D 曲線(素材 A)

最後に、提案方式のフィルタ特性について考察する。提案方式では、一般的な映像における低域成分の重要性を考慮して、低域における通過域の設定に重点を置いている。一方、高域成分については、基本的にカットする特性の設定となる。

図 4 に、素材 A におけるフィルタ特性を示している。同図では、提案手法および比較のために EAlF のフィルタ特性を示している。なお、同図における提案手法の特性は、3.2 節の方式により決定されたフィルタ係数をフーリエ変換することにより得られるものである。また、図 5 に、当該フレームの原画像の振幅特性を示す。図 4 より、低域の通過特性について、提案手法は EAlF と同等な特性を設計できていることが分かる。一方で、高域では特性の違いが見られる。しかしながら、図 5 から分かる通り、同素材では高域には大きな振幅を持つ成分は存在しない。このため、MC 誤差および符号化性能への影響は小さい。他素材および他符号化条件における提案手法のフィルタ特性についても、図 4 と同様に、低域では EAlF とほぼ同等の特性が得られることを確認した。また、高域で EAlF と異なる場合があるものの、符号化性能への影

響は小さいことを確認している。

本実験では、図5のような空間周波数特性を持つ、自然画像を実験素材として用いており、良好な結果が得られている。しかしながら、提案手法は高域のみを通過するフィルタを設計することが困難であるため、高域に重要な成分が集中する映像に対しては、符号化性能の低下を招く恐れがある。映像特徴への更なる適応性については、今後の課題である。

## 5. まとめ

本研究では、少ない処理コストでAIFに準じた性能を有する補間フィルタを設計する手法を提案した。符号化実験により、超高精細映像において、JMに対する符号化性能改善は最大0.33dBであり、EAIFに対する性能低下は0.06dBであることを確認した。更に、このときのAIF処理量はEAIFの30%程度であることを確認した。

本研究は独立行政法人情報通信研究機構による委託研究「超高精細映像符号化技術に関する研究開発」として実施したものである。

## 参考文献

- [1] ITU-R BT-1769 "Parameter values for an expanded hierarchy of LSDI image formats for production and international programme exchange".
- [2] SMPTE 2036-1-2007, "Ultra High Definition Television - Image Parameter Values for Program Production".
- [3] ITU-T and ISO/IEC JTC1, Rec. H.262 and ISO/IEC 13818-2, Generic Coding of Moving Picture and Associated Audio Information - Part 2: Video, Nov. 1994.
- [4] T. Wedi and H. G. Musmann, "Motion- and Aliasing-Compensated Prediction for Hybrid Video Coding," IEEE trans. on Circuits and Systems for Video Technology, vol.13, no.7, July 2003.
- [5] T. Wedi, "Adaptive Interpolation Filters and High-Resolution Displacements for Video Coding," IEEE trans. on Circuits and Systems for Video Technology, vol. 16, no.4, April 2006.
- [6] Y. Vatis and J. Ostermann, "Adaptive Interpolation Filter for H.264/AVC," IEEE trans. on Circuits and Systems for Video Technology, vol.19, no.2, February 2009.
- [7] K. Ugur, D. Rusanovskyy and M. Gabbouj, "Efficient Calculation of Adaptive Interpolation Filter with Distortion Modelling," ICASSP 2008, pp.673 - 676, April 2008.
- [8] KTA Software, <http://iphome.hhi.de/suehring/tml/>
- [9] Y. Ye, M. Karczewicz, "Enhanced Adaptive Interpolation Filter," ITU-T SG16/Q.6 Doc. T05-SG16-C-0464, Geneva, Switzerland, April 2008.
- [10] H.264 reference software, <http://iphome.hhi.de/suehring/tml/>

[11] G. Bjontegaard, "Improvements of the BD-PSNR model", ITU-T 16/Q6, 35<sup>th</sup> VCEG Meeting, July 2008, VCEG-AI11.

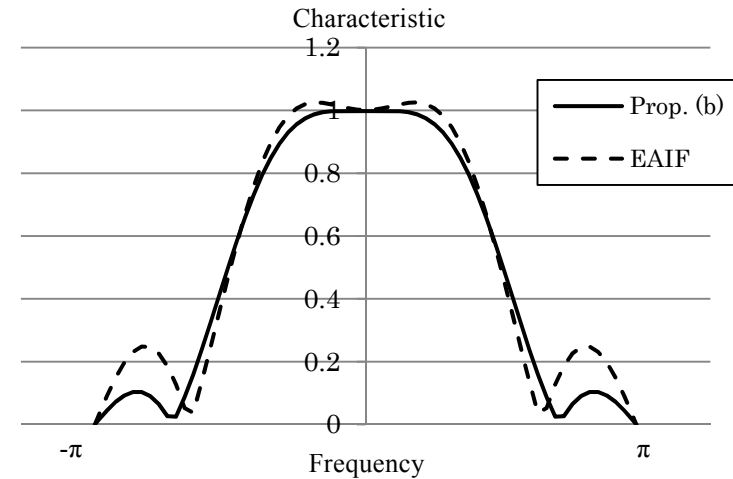


図4 フィルタ特性(素材A, QP=37, 9フレーム目)

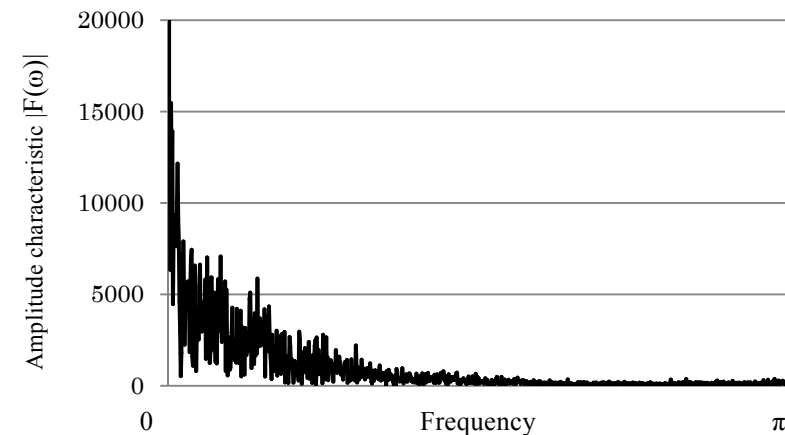


図5 振幅特性(素材A, 9フレーム目)