

Motion JPEG 2000 を利用したマルチ解像度変換システム

石川 孝明[†] 渡辺 裕[†]

[†] 早稲田大学 大学院 国際情報通信研究科
〒 367-0035 埼玉県本庄市西富田大久保山 1-0-1-1
E-mail: †takaxp@fuji.waseda.jp

あらまし 近年、ビデオ符号化におけるスケーラビリティの重要性が高まっている。主なスケーラビリティとして、空間、品質、時間スケーラビリティがある。2002年に国際標準化した Motion JPEG 2000 は、優れた空間スケーラビリティと品質スケーラビリティを有し、超高解像度映像の符号化方式として期待されている。しかし、Motion JPEG 2000 で採用されている周波数帯域のオクターブ分割によるサブバンド符号化では、デコード後の解像度が $1/2^n$ に制限される事が明らかとなっている。この問題に対し、従来手法ではデコードした低域画像とデシメーションフィルタを組み合わせる事によりマルチ解像度変換を実現している。しかし、自然画像を入力とした場合の変換方法には検討の余地がある。本研究では、デコードした低域画像と組み合わせるデシメーションフィルタの処理量を定式化し、フルデコード画像に対するフィルタリングと同等の処理量で、より高次のフィルタリングが可能であることを示し、Motion JPEG 2000 とデシメーションフィルタを組み合わせたマルチ解像度変換システムの有効性を示す。

キーワード Motion JPEG 2000, スケーラビリティ, 解像度変換, マルチレート変換, ローパスフィルタ

A Multi-resolution Conversion System using Motion JPEG 2000

Takaaki ISHIKAWA[†] and Hiroshi WATANABE[†]

[†] Graduate School of Electronics, Information and Communication Engineering Studies, Waseda University
1-0-1-1 Okuboyama, Nishi-Tomida, Honjo, Saitama, 367-0035 Japan
E-mail: †takaxp@fuji.waseda.jp

Abstract Recently, the importance of scalability functions in video codecs such as SNR, temporal and spatial scalability has been increased. Motion JPEG 2000, which is the Part 3 of the JPEG 2000 standard series, has the advanced SNR and spatial scalability. It is achieved by subband division using wavelet transform and Embedded Block Coding with Optimal Truncation (EBCOT). However, the spatial scalability is limited to the cases where decimated resolution is $1/2^n$ of the source image. When the source image has high resolution like Digital Cinema or HDTV, then we cannot get an SDTV or QVGA size image by decoding Motion JPEG 2000 betstream directly. We developed a software “down-conversion” system, which can generate sequences of different resolutions, such as HDTV, SDTV, CIF, QVGA and QCIF, from Digital Cinema-size video sequences. The system combines LL sub-band of source and the Lanczos3 filter.

Key words Motion JPEG 2000, Scalability, Resolution Conversion, Multi-rate Conversion, Low-pass filter

1. ま え が き

近年、静止画像符号化および動画符号化におけるスケーラビリティの重要性が高まっている。符号化データのスケラビリティには複数の種類があるが、2000年に標準化された JPEG 2000 は優れた品質スケーラビリティと空間スケーラビリティ

を有する [1]。また、JPEG 2000 は、 4096×2160 [pel] にも及ぶ超高解像度画像の符号化が必要とされるデジタルシネマの符号化方式として有力視されている。今後は、デジタルシネマのような高解像度かつ高品質なソースを、Motion JPEG 2000 [2] などのスケーラビリティを有する符号化方式で符号化することにより、単一のソースからテレビや携帯向けの複数の

低解像度および低品質のコンテンツを、デコード処理のみで取得することが可能となる。このように、ワンソース・マルチユースを実現するアプリケーションが現実的となっている。

しかし、JPEG 2000 および Motion JPEG 2000 で利用されている周波数帯域のオクターブ分割によるサブバンド符号化では、デコード後の画像解像度が $1/2^n$ に制限される事が明らかとなっている [3]。この問題に対し、我々の提案するマルチ解像度変換法では、符号化データのデコード処理とデシメーションフィルタを組み合わせる事により任意の解像度の出力画像の生成を実現しており、出力画像の品質選択基準として Effective Tap Length を提案している [3]。

本稿では、同手法を実際の自然画像に対して適用する場合には必ずしも正確な推定結果を得ることができないことを示し、自然画像に対するマルチ解像度変換を行う場合のデコード画像の性質を明らかとし、その画像品質が、デシメーションフィルタの性能に依存する事を示す。また、デコードした低域画像と組み合わせるデシメーションフィルタの処理量を定式化し、フルデコード画像に対するフィルタリングと同等の処理量で、より高次のフィルタリングが可能であることを示し、提案するマルチ解像度システムの有効性について述べる。

2. 従来システム

2.1 システムフロー

従来システムは、図1のように Motion JPEG 2000 のデコダとデシメーションフィルタを組み合わせることで任意の解像度を取得する事ができる。システムは、まず、Motion JPEG 2000 のビットストリームから、空間スケーラビリティを利用して縮小画像を得る。この時、取得可能な縮小画像は、 $\frac{1}{2^n}$ に制限されるため、デシメーションフィルタにより画素を間引き、任意の解像度でデコード画像の生成を実現している。実際のシステムフローでは、さらにピクセルアスペクト比やフレームレートの違いも考慮して、SDTV 用コンテンツや QVGA サイズの携帯向けコンテンツを生成する。

2.2 Effective Tap Length

図1のシステムフローにおいて、空間スケーラビリティにより取得可能な縮小画像の種類は、最終的にシステムが出力する画像の解像度が低ければ低いほど、多様性を増す。例えば、デジタルシネマサイズから QVGA サイズの縮小画像を得るためには、図2のように3通りの低域画像が取得可能である(以降、

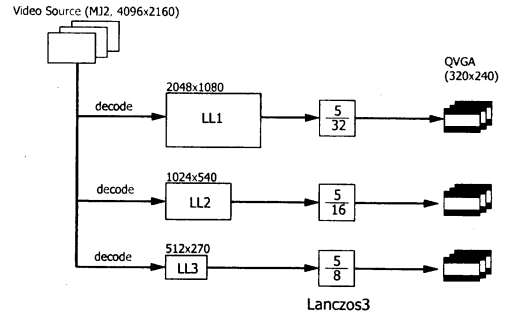


図2 Decoding path from 4096x2160 to 320x240

選択可能な処理フローを、デコードパスと表現する)。これに対し、最高品質の出力画像を得るために、従来手法では、式(1)から式(3)に示す定義式(“Effective Tap Length”)を用いて、どのデコードパスを選択することで最高品質の出力画像を取得できるかについての指標を定義している。これは、Wavelet フィルタの有効タップ数と Lanczos3 フィルタ [4] のタップ数を加算した値で定義される。本システムで用いる Wavelet フィルタは、JPEG 2000 で定義されている 5/3-Integer フィルタと 9/7-Daubechies フィルタである [1]。

式(1)は、原点を除く正の領域におけるタップ数を示している。奇数タップの FIR フィルタとして定義される Wavelet フィルタは、対称性を利用して式(2)のように、Wavelet フィルタによる分解数 n と、そのタップ数 t (5/3-Integer フィルタは 5tap, 9/7-Daubechies フィルタは 9tap) で表される。

$$A(n, t) = \left(\frac{t-1}{2}\right) [2^{n-1} + 2^{n-2} + \dots + 2^1 + 2^0] \quad (1)$$

$$\begin{aligned} TAPS_{DWT}(n, t) &= 2A(n, t) + 1 \\ &= 2 \sum_{k=0}^{n-1} \left(\frac{t-1}{2}\right) 2^k + 1 \\ &= (t-1)(2^n - 1) + 1 \quad (n > 0) \end{aligned} \quad (2)$$

t : DWT フィルタのタップ数, n : 分解数

また、Lanczos3 フィルタのタップ数は、縮小率 r に対し、式(3)で定義される。このように、Lanczos3 フィルタの特性は、フィルタのタップ数が縮小率により決定される点にある。

$$TAPS_{LANC3}(r) = \lfloor \frac{1}{r} \cdot 3 \rfloor \cdot 2 + 1 \quad (3)$$

以上より、最終的にシステムがフィルタリング処理で利用するフィルタのタップ数は、式(2)および式(3)を用いて、式(4)のように定義される。

$$\begin{aligned} TAPS_{(r, n, t)} &= TAPS_{DWT} + TAPS_{LANC3} \\ &= \begin{cases} \lfloor \frac{1}{r} \cdot 3 \rfloor \cdot 2 + 1 & (n = 0) \\ \lfloor \frac{1}{r} \cdot 3 \rfloor \cdot 2 + (t-1)(2^n - 1) + 2 & (n \geq 1) \end{cases} \end{aligned} \quad (4)$$

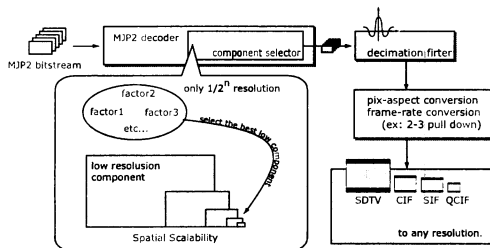


図1 Traditional multi-decoding system

3. 予備実験

式(4)を用いた従来システムのマルチ解像度変換法では、テスト画像をゾーンプレートとして性能評価を行っている。この場合、フィルタの有効タップ数と出力画像の画質に相関性が見られた[5]。しかし、マルチ解像度変換法を実際に適用するコンテンツは、デジタルビデオカメラで撮影された、もしくはフィルムをスキヤニングする事で得られる高解像度の自然画像が主である。そこで、複数の高解像度標準画像 (harbor, wool, cafeteria, bicycle, bride) を用いて再実験を行った。入力画像は、画像サイズを 1920×1080 [pel] の HD サイズにトリミングして用いた。また、最終的な出力画像サイズは、HD サイズを基準として、 768×432 [pel](2/5), 384×216 [pel](1/5), 192×108 [pel](1/10), 96×54 [pel](1/20) の画像サイズ、および縮小率とした。

例として、標準画像 bicycle における実験結果を、表1と表2に示す。表1は、ソースが9/7-Daubechies フィルタにより解析され、EBCOTで圧縮された画像を用いた場合、表2は5/3-Integer フィルタにより解析されEBCOTで圧縮された画像を用いた場合であり、共に理想解像度変換[5]後の画像と比較したPSNRの値である。また、表3、表4に、9/7-Daubechies フィルタの場合と5/3-Integer フィルタの場合の、式(4)より求められる有効タップ数を示す。

システムは、表3および表4において太字に表示されたデコードパスを選択することで、最高品質の出力画像を得ることができる。しかし、有効タップ数と自然画像を入力画像とした場合の出力画像の品質を比較すると、必ずしも高い相関性を得ているとは言えない。これは、従来システムの評価に用いていたゾーンプレートが、全周波数帯に周波数成分を持つことに対し、自然画像は低域成分を中心に周波数成分が分布するため、有効タップ数との相関性が低くなったと考えられる。また、これらの結果は、その他の標準画像においても同じであった。

表1 PSNR (bicycle, Daubechies)

	768x432	384x216	192x108	96x54
Directly Lanczos3	33.807	32.340	33.505	32.010
use LL1(960x540)	32.401	31.948	33.410	31.978
use LL2(480x270)		30.701	33.070	31.866
use LL3(240x135)			31.636	31.498
use LL4(120x68)				26.674

[dB]

表2 PSNR (bicycle, Integer)

	768x432	384x216	192x108	96x54
Directly Lanczos3	33.818	32.345	33.516	32.021
use LL1(960x540)	31.727	32.320	33.315	31.853
use LL2(480x270)		29.545	33.052	31.779
use LL3(240x135)			29.174	31.814
use LL4(120x68)				25.268

[dB]

表3 Effective tap length (Daubechies)

	768x432	384x216	192x108	96x54
Directly Lanczos3	15	31	61	121
use LL1(960x540)	16	24	40	70
use LL2(480x270)		32	40	56
use LL3(240x135)			64	72
use LL4(120x68)				128

[dB]

表4 Effective tap length (Integer)

	768x432	384x216	192x108	96x54
Directly Lanczos3	15	31	61	121
use LL1(960x540)	12	20	36	66
use LL2(480x270)		20	28	44
use LL3(240x135)			36	44
use LL4(120x68)				68

[dB]

このように、従来システムに自然画像を入力する場合には、有効タップ数による出力画像の画質推定は、必ずしも有効であるとは限らないことが分かった。

4. 自然画像を対象としたシステム構成

前章に示したとおり、自然画像では、必ずしも式(4)で与えられる有効タップ数が有効では無いことが分かった。これに対し、予備実験で得られたPSNRの値による出力画像の品質に着目すると、次のような性質があると言える。

(1) ソースのスケラビリティを利用し、低域画像を用いるデコードパスでは、分解レベルの低いLL成分とLanczos3フィルタとの組み合わせが高いPSNRを示している。

(2) 分解レベルの低いLL成分を使用するデコードパスでは、PSNRの値はあまり変化せず、その変化幅は約1[dB]程度と微小であり、視覚的には変化が少ない。

(3) Lanczos3フィルタの性質により、縮小率が上がるにつれ伝達関数のタップ長が長くなり、演算量が増加する。

これらの性質を考慮すると、ソースがデジタルシネマのような高い解像度であればあるほど、出力画像を得るための縮小率が小さくなるため、低域画像を用いないデコードパスでは、演算量の増加を招く事が考えられる。逆に、演算量の増加にもかかわらず、出力される画像の品質は増加しない傾向にあるため、有効タップ数による推定結果を出力するシステムは、最適なシステムであるとは必ずしも言えない。したがって、これらの性質を考慮したシステム構成について、以下で述べる。

4.1 Lanczos フィルタとの組み合わせによる画質の変化

自然画像におけるデコード後の画像品質を左右する要因が、デシメーションフィルタにあることを示す。これは、オクターブ分割を繰り返すサブバンド符号化の本質的な問題点である。

Motion JPEG 2000により符号化されたソースを利用して、任意の解像度の映像を取得する本システムは、図3のようなマルチステージシステムとして表す事ができる。図3はソースが9/7-Daubechies フィルタで解析されている場合を示して

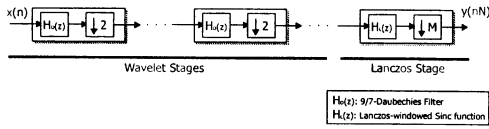


図3 Multi-stage rate conversion

いる。以降、Wavelet フィルタを使用するデシメータ部分を“Wavelet ステージ”、Lanczos フィルタを使用するデシメータ部分を“Lanczos ステージ”と呼ぶ。

本来、サブバンド符号化では解析過程に生じるエイリアシング成分がサブバンドの合成過程により打ち消されるが、図3のような、マルチレートシステムの Wavelet フィルタによる解析過程においては、ダウンサンプリングによりエイリアス成分が除去されず蓄積される [6]。したがって、Wavelet フィルタによる分割数が多ければ、信号は劣化する傾向にあると言える。

これに対し、マルチステージシステムにおける最終段に Lanczos フィルタによるデシメータを挿入することで、Wavelet ステージで蓄積されたエイリアス成分が、ある程度打ち消され、DFT による解像度変換により出力されるエイリアス成分が存在しない理想的な信号と比較し、PSNR の値が改善されていると考えられる。これは、Lanczos フィルタの特性が、図4に示されたように 9/7-Daubechies, 5/3-Integer よりもカットオフ周波数付近で急峻な特性を持つフィルタであり、不要なエイリアス成分が除去されるためであると考えられる。

4.2 Lanczos フィルタの演算コスト

前章で示したように、本システムの出力する画像の品質は、Lanczos フィルタによるエイリアシング成分の抑制に依存していると予想される。したがって、より高次の Lanczos フィルタを用いることで画質が改善されると考えられる。しかし、高次のフィルタは、多くの演算を必要とし、そのままではシステムに導入できない。そこで、任意のローブ数で構成される Lanczos フィルタを使用する場合の演算コストを定式化する必要がある。

4.2.1 Lanczos フィルタによるデシメータ

ローブ数が3である、Lanczos3(3-lobed Lanczos-windowed

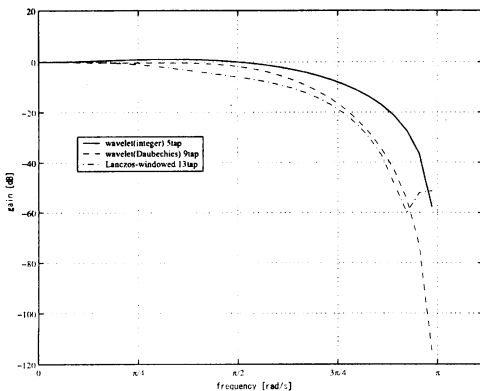


図4 Frequency response



図5 noble identity

sinc function [4]) フィルタは、区間 $-3 \leq x \leq 3$ に縮小率に応じた数の係数を持つ。フィルタの係数は、理想的なローパスフィルタである SINC 関数に対し、式 (5) の窓関数を掛けた合わせ、式 (3) により求められるタップ数で関数の値が 0 で無い区間 $(-3 \leq x \leq 3)$ を等分割することで得られる。

$$Window_{Lanczos}(x) = \begin{cases} \sin(\pi \frac{x}{3}) / \pi \frac{x}{3}, & |x| < 3 \\ 0, & otherwise \end{cases} \quad (5)$$

ここで、ローブの大きさを N 、縮小率 r を出力系列の数と入力系列の数を用いて表し、Lanczos フィルタのタップ数を一般化した関係式を式 (6) に示す。

$$TAPS_{LANC}(N) = \frac{Src}{Dst} \cdot N \cdot 2 + 1 \quad (6)$$

Dst : 出力信号数列の総数, Src : 入力信号数列の総数

一方、縮小率 r の Lanczos フィルタのカットオフ周波数は、 N の値によらず、 $\omega_c = \pi r$ で与えられる。したがって、任意のカットオフ周波数を与える Lanczos フィルタのタップ数は、式 (7) により表される。

$$TAPS_{LANC}(N) = \frac{\pi}{\omega_c} \cdot N \cdot 2 + 1 \quad (7)$$

このように、最終ステージのデシメータにおいて、Lanczos フィルタをローパスフィルタとして用いた場合、フィルタのタップ数が、要求されるカットオフ周波数 ω_c (もしくは縮小率 r) に依存するため、縮小率が上がることでタップ数が増加し、それに伴い演算量も増加する事が分かる。

4.2.2 Lanczos デシメータの演算量の定式化

縮小率と Lanczos フィルタのタップ数の関係を利用して、ローブ数が N の Lanczos フィルタの処理コストを定式化する。入力信号に対するローパスフィルタとして、カットオフ周波数が ω_c の Lanczos フィルタを利用する。これは、一般的な FIR フィルタと同じ回路図で表現される。ダウンサンプリング前の FIR フィルタの回路に、図5に示すノーブル恒等式 [6] を適用することで、通常のローパスフィルタよりも低レートで処理を行う事ができ、演算量を低減する事ができる。さらに、Lanczos フィルタは SINC 関数が元になっているため、偶数列の FIR フィルタであると言える。したがって、デシメータの演算コストは、乗算処理と加算処理に分けて考えることで、以下のように導出される。ただし、簡略化のため、Lanczos フィルタのローブ数は、水平成分と垂直成分で区別せず、 N としている。また、本システムでは、2次元信号に対するフィルタリング処理を分離型システムとして扱っている。

まず、デシメータが係数を1つ計算するために必要とする乗算回数と加算回数は、それぞれ式 (8) と式 (9) で表される。

$$Cost_{multi_coeff} = \frac{TAPS_{LANC}(N) - 1}{2} + 1$$

$$= N \frac{Src}{Dst} + 1 \quad (8)$$

$$\begin{aligned} Cost_{add_coeff} &= TAPSLANC(N) - 1 \\ &= 2N \frac{Src}{Dst} \end{aligned} \quad (9)$$

次に、分離型 2 次元システムであることを考慮し、水平成分の画素数、および垂直成分の画素数を考慮すると、デシメータ全体での処理量は、式 (8) と式 (9) を用いて式 (10) および式 (11) で表される。

$$Cost_{multi} = XY(N(M_V + M_H) + 2) \quad (10)$$

$$Cost_{add} = 2XY(N(M_V + M_H)) \quad (11)$$

X : 出力 2 次元信号の水平画素数、

Y : 出力 2 次元信号の垂直画素数、

M_V : 入力系列の水平画素数/出力系列の水平画素数、

M_H : 入力系列の垂直画素数/出力系列の垂直画素数、

ピクセルアスペクト比の変換を行わない場合、 $M_V = M_H (= \frac{1}{2})$ が成立する。そのため、式 (10) および式 (11) は、 $\omega_c = \pi T$ であることを用いて、式 (12)、式 (13) として表され、これはカットオフ周波数の値が演算コストを直接表すことを示している。

$$Cost_{multi} = 2XY \left(\frac{\pi N}{\omega_c} + 1 \right) \quad (12)$$

$$Cost_{add} = 4XY \frac{\pi N}{\omega_c} \quad (13)$$

5. 演算コストを考慮した提案システム

自然画像を対象としたマルチ解像度変換システムにおける諸性質をまとめると、出力画像の品質は、ポストフィルタリング (Lanczos フィルタ) の周波数特性に影響され、カットオフ周波数の値によりフィルタのタップ数が決定される Lanczos-lobed Sinc 関数では、周波数特性と処理コストの間に、式 (12)、(13) で定義される関係があることが分かった。

以上の性質に基づき、本稿では、従来手法と同等の処理量がそれ以下の演算量で、より高品質な出力画像を得るシステムを提案する。

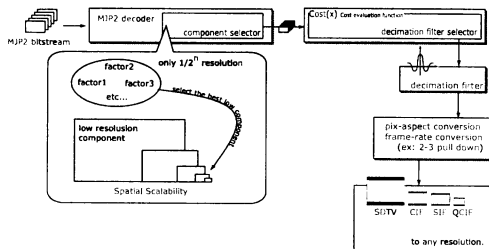


図 6 Proposed resolution conversion system

5.1 提案システム

従来システムに、図 6 のように演算コスト評価部を組み込む。これにより、従来システムよりも低い演算量のフィルタリングを行うことを保証した上で、処理負荷に合わせた Lanczos フィルタの次数を選択することが可能となり、同時に画質改善が期待される。

5.2 コスト関数

演算コスト評価部で用いるコスト関数を式 (10) および式 (11) から導く。まず、Lanczos ステージに入力される信号の画素数を、使用する低域成分を得るために必要な Wavelet ステージの分割数 L を用いて式 (14) のように表す。

$$\text{Lanczos ステージへの入力画素数} = \frac{\text{Wavelet ステージの出力画素数}}{2^L} \quad (14)$$

次に、式 (14) より、分割数 L の場合の処理コスト、 M の場合の処理コストを式 (15) と式 (16) に示す。ただし、 M の場合は、使用する Lanczos フィルタのローブ数を α (正の整数) だけ増加させると仮定し、 $L < M$ が成立するとする。

$$\begin{aligned} Cost(L)_{multi} &= XY \left(N \left(\frac{M_V}{2^L} + \frac{M_H}{2^L} \right) + 2 \right) \\ Cost(L)_{add} &= 2XY \left(N \left(\frac{M_V}{2^L} + \frac{M_H}{2^L} \right) \right) \end{aligned} \quad (15)$$

$$\begin{aligned} Cost(M)_{multi} &= XY \left((N + \alpha) \left(\frac{M_V}{2^M} + \frac{M_H}{2^M} \right) + 2 \right) \\ Cost(M)_{add} &= 2XY \left((N + \alpha) \left(\frac{M_V}{2^M} + \frac{M_H}{2^M} \right) \right) \end{aligned} \quad (16)$$

式 (15) および式 (16) において、 $Cost(L)_{multi} > Cost(M)_{multi}$ かつ $Cost(L)_{add} > Cost(M)_{add}$ が成立する α が存在する場合、従来システムと比較し、同等の演算量でよりローブ数の高い Lanczos フィルタを使用する事ができる。式 (17) に α の定義式を示し、選択可能な α の値を表 5 に示す。

$$N(2^{M-L} - 1) > \alpha \quad (17)$$

表 5 は、任意の解像度を得るために選択可能なデコードパスにおいて、フルデコードを必要とするデコードパスの処理コストを基準として、それ以下の処理量において選択可能な Lanczos フィルタのローブ数を示している。

本システムのデコード処理は、図 3 に示したように、リフティング構成による逆 Wavelet 変換後の係数に対する演算と Lanczos フィルタリングの FIR フィルタリングに対する演算の 2 つにより構成されるが、簡略化のため、本稿では逆 Wavelet 変換にかかる演算コストを考慮しない。

表 5 Max value of alpha

L \ M	1	2	3	4	5
0	3	9	21	45	93
1		3	9	21	45
2			3	9	21
3				3	9
4					3

5.3 画質改善

本システムでは、マルチステージシステムの最終段にあたるデシメータで用いるローパスフィルタの急峻性を改善することで、画質改善が期待される。Lanczos フィルタの急峻性は、窓関数の大きさに依存するため、従来システムで用いていた Lanczos フィルタのロープ数よりも大きい値を使用する。ロープ数を N とした場合の Lanczos フィルタを、式 (18) に示す。

式 (18) において N の値を無限大にすれば、Lanczos フィルタは SINC 関数と近い特性となる。しかし、それに伴い処理量が増加するため、ロープ数の拘束条件として式 (17) を用いる。

$$LanczosN(x) = \begin{cases} \frac{\sin(\pi x)}{\pi x} \frac{\sin(\frac{\pi x}{N})}{\frac{\pi x}{N}}, & |x| < N \\ 0, & otherwise \end{cases} \quad (18)$$

5.4 シミュレーション

入力画像として、標準画像 harbor と wool を使用した。画像サイズは、デジタルシネマで使用される画像サイズに近い 3840x2160[pe] とした。表 6~表 9 に、各出力画像と DFT による理想解像度変換を行った画像との比較による PSNR の値を示す。画質改善の比較対象として、従来手法において低域成分を使用しないデコードパス ($L=0$) の PSNR を用いる。太字で示した部分が、従来手法よりも画質が改善したデコードパスである。

5.5 考察

表 6~表 9 に示された実験結果では、Wavelet フィルタの種類に依らず、同じ縮小率とデコードパス ($M=1, 2, 3$) において画質が改善されている事が分かる。また、画質が改善されたデコードパスでは、いずれのパスを選択しても出力画像の PSNR 値の変動は微小である。一方で、画質が改善されたデコードパスにおける Lanczos ステージの縮小率に着目すると、すべて 1/2 未満であることが分かった。これは、Wavelet ステージで蓄積されるエイリアス成分が、 $\pi/2$ よりも高い周波数成分に主に影響を与えているためであると考えられるが、正確には入力した画像の周波数特性を考慮する必要がある。

表 6 PSNR (harbor, Daubechies)

	1536x864	768x432	384x216	192x108	96x54
L=0	34.140	36.150	34.320	32.297	30.354
M=1	30.413	31.891	36.006	33.889	31.318
M=2		28.175	30.827	34.864	31.911
M=3			27.018	29.797	32.142
M=4				26.213	28.635

[dB]

表 7 PSNR (harbor, Integer)

	1536x864	768x432	384x216	192x108	96x54
L=0	34.149	36.158	34.326	32.299	30.361
M=1	29.744	31.706	36.051	33.857	31.271
M=2		26.749	30.475	34.687	31.749
M=3			25.036	29.260	31.816
M=4				23.971	28.037

[dB]

表 8 PSNR (wool, Daubechies)

	1536x864	768x432	384x216	192x108	96x54
L=0	37.180	38.360	36.243	34.722	32.158
M=1	33.404	34.160	37.674	35.542	32.647
M=2		30.610	33.691	36.001	32.923
M=3			30.132	32.350	33.048
M=4				29.108	30.026

[dB]

表 9 PSNR (wool, Integer)

	1536x864	768x432	384x216	192x108	96x54
L=0	37.196	38.372	36.244	34.723	32.155
M=1	32.704	33.971	37.735	35.551	32.647
M=2		29.110	33.272	35.915	32.880
M=3			27.896	31.844	32.906
M=4				27.101	29.592

[dB]

6. まとめ

本稿では、Motion JPEG 2000 を利用したマルチ解像度変換システムにおける演算コストの定式化と提案システムによる画質改善について述べた。

まず、処理対象を自然画像とした場合の従来システムの欠点を改良し、出力画像の品質を左右する処理が、Lanczos フィルタによるフィルタリングであることを示した。次に、Lanczos フィルタリングの演算コストを定式化することで、従来システムと同等の処理コストであることを保証した上で、より高次のフィルタリングを行い、画質が改善される場合があることを示した。これにより、符号化データの空間スケーラビリティを利用したマルチ解像度変換システムの有効性を示した。

謝辞

この研究は、情報通信研究機構 (NICT) 委託研究課題 “通信ネットワーク利用放送技術の研究開発” のサポートによって行われた。

文献

- [1] *JPEG 2000 Part1 FDIS*, ISO/IEC JTC1/SC29 WG1, July 2002.
- [2] T. Fukuhara, D. Singer, *Motion JPEG2000 Final Draft International Standard 1.0*, N2250, ISO/IEC JTC1/SC29 WG1, September 2001.
- [3] 田邊集, 渡辺裕, 富永英義, “Motion jpeg 2000 における最適解像度変換法,” 画像符号化シンポジウム *PCSJ2003*, 2003.
- [4] K. Turkowski, “Filters for common resampling tasks.” Apple Computer, Tech. Rep., October 1990.
- [5] 田邊集, 石川孝明, 渡辺裕, 富永英義, “Motion jpeg2000 映像配信のためのマルチ解像度変換法,” 情報処理学会 *AVM* 研究会, no. AVM44-19, March 2004.
- [6] 貴家仁志, *マルチレート信号処理*. 昭晃堂, 1995.