

Motion JPEG2000 における最適解像度変換法の検討

田邊 集[†] 渡辺 裕^{††} 富永 英義^{††}

[†] 早稲田大学 大学院 理工学研究科

〒 1169-8555 東京都新宿区大久保 3-4-1

^{††} 早稲田大学 大学院 国際情報通信研究科

〒 169-0051 東京都新宿区西早稲田 1-3-10

E-mail: †shu@tom.comm.waseda.ac.jp

あらまし デジタルシネマや HDTV などの超高解像度映像に対する符号化方式として Motion JPEG2000 の期待が高まっている。Motion JPEG2000 では wavelet 変換を用いているため、解像度スケラビリティが容易に実現できるといわれている。しかし、実際に用いられている wavelet フィルタは、現在解像度変換に用いられているデシメーションフィルタと比べると性能が悪く、また任意サイズに対応した解像度変換が実現できないといった問題がある。本研究では、最適解像度変換法の提案を目的とし、フィルタの性能比較や画質比較などの基礎検討を行った。

キーワード Motion JPEG2000, 解像度変換, デシメーションフィルタ, wavelet 変換

A Study on image resampling for Motion JPEG2000

Shu TANABE[†], Hiroshi WATANABE^{††}, and Hideyoshi TOMINAGA^{††}

[†] Graduate School of Electronics, Information and Communication Engineering Studies, Waseda University

Okubo 3-4-1, Shinjuku-ku, Tokyo, 169-8555 Japan

^{††} Graduate School of Global Information and Telecommunication Studies, Waseda University

Nishiwaseda 1-3-10, Shinjuku-ku, Tokyo, 169-0051 Japan

E-mail: †shu@tom.comm.waseda.ac.jp

Abstract Motion JPEG2000 is regarded as a promising coding system for an ultra high quality image like digital-cinema and HDTV. Wavelet transform used in Motion JPEG2000 enables to realize spatial scalability easily. However, the performance of the wavelet filter used in Motion JPEG2000 is worse than those of decimation filters used in image resampling and resizing. What is worse, the wavelet filter can only resize images to half resolution. We aim to provide of the best resampling method for Motion JPEG2000. In this paper, we investigate the performance of filters and its relationship with image quality.

Key words Motion JPEG2000, image resampling, decimation filter, wavelet transform

1. はじめに

近年、デジタルシネマやHDTVなど、いわゆる超高精細画像に対する符号化方式として、wavelet変換を用いた符号化方式の適用が議論されている。wavelet変換を用いた符号化方式としてMotion JPEG2000が挙げられる。

Motion JPEG2000は次世代の動画圧縮規格として、JPEG2000と同様、国際標準化組織ISO/IEC JTC1/SC29/WG1で規格化作業が進められてきた[1][2][3]。

Motion JPEG2000の利点としては(1)従来のJPEGやMPEG1,2,4などで用いられているDCTの代わりに、wavelet変換を用いることで、符号化効率の向上と高画質化を実現できる、(2)階層符号化によるスケーラビリティ機能の実現(Spatial/SNR scalability)(3)可逆圧縮と非可逆圧縮の統合、などがある。

本研究では、このうち空間スケーラビリティ機能に注目した。デジタルシネマなどの高解像度の画像を符号化した場合、配信時にユーザの環境に合わせた解像度にダウンサンプリングする必要がある。Motion JPEG2000ではデータ構造が階層構造をとっており、サブバンドごとの低域成分を用いることで、低解像度の画像を得ることが出来る。

しかし、Motion JPEG2000で用いられるwaveletフィルタはダウンサンプリングを目的としたフィルタではないため、得られる画質は最良とはいえない。また、wavelet変換によって得られた低域成分は、サブバンド分割の性質上、元画像の画像サイズの $1/2^n$ のサイズしか実現できない。

現在の主要メディアの解像度を表1に示した。これを見ると、アスペクト比も解像度比もメディアによって異なることがわかる。映像配信時の解像度変換はこれに合わせて行う必要があり、Motion JPEG2000が持つ空間スケーラビリティ機能だけでは不十分であることは明らかである。

そこで本研究では、図1に示すようなMotion JPEG2000で符号化された画像に対する最適な解像度変換法の提案を目的とし、Motion JPEG2000で実際に用いられているwaveletフィルタによる解像度変換と、既存の解像度変換法との比較を行った。

2. 解像度変換

画像の解像度変換は一般的に次の手順で行われる。

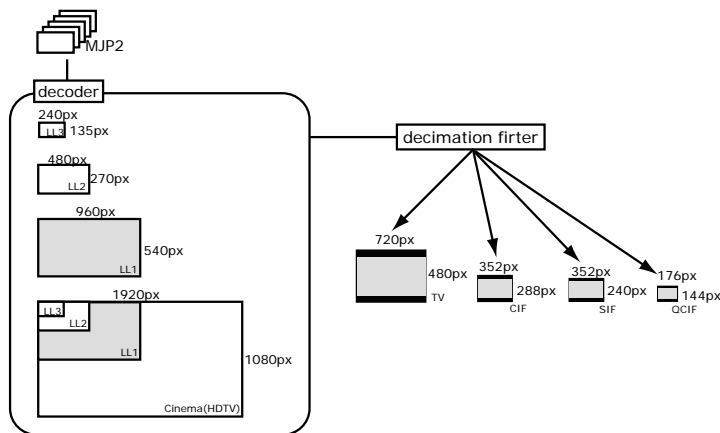


図1 提案システム図

- (1) プレフィルタ
- (2) 座標変換
- (3) 補間処理

座標変換は、縮小後の画素が、縮小前のどの位置に対応するか求めることである。求められた対応位置の画素値を求めるのが補間処理である。補間処理には、nearest neighbor法、bi-linear法、cubic convolution法[6]などが用いられる。ただし、画像サイズを小さくすると、縮小後のサンプリング周波数では表現できない高周波成分が折り返しノイズとして現れる。そこでプレフィルタにローパスフィルタを用いて高周波成分をカットし、折り返しノイズを抑えた縮小を可能にしている。

2.1 デシメーションフィルタ

解像度変換の中で最も重要な部分がプレフィルタ、つまりローパスフィルタの性能である。ローパスフィルタは解像感を高めようとタップ数を増やせばリングングノイズが発生し、リングングノイズを減らそうとタップ数を減らせばエイリアスノイズが発生する。また、エイリアスノイズを減らそうと縮小サイズよりも小さいカットオフ周波数を指定すれば解像感が失われる。このようなトレードオフの関係が存在するので、縮小処理に合わせたフィルタの設計が必要となる。そこで、縮小率に応じてフィルタのタップ数を変化させ、最適なローパスフィルタの効果をもたせながら縮小処理を行うように考え出されたのがデシメーションフィルタである。

デシメーションフィルタの例としては、次のようなものがある[4][5]。

- 平均画素法
- Gaussianフィルタ
- Lanczos2,3フィルタ

平均画素法は元画像をいったん拡大し、画素の平均から縮小画像の画素値を求める手法である。Gaussianフィルタは画像解析などによく使用されるガウシアン関数(1)を用いたフィルタである。Lanczos2,3フィルタは理想ローパスフィルタにそれぞれの窓関数(2)(3)をかけた手法である。

$$Gaussian(x) = 2^{-4x^2} \tag{1}$$

$$Lanczos2(x) = \begin{cases} \frac{\sin(\pi x)}{\pi x} \frac{\sin(\frac{\pi x}{2})}{\frac{\pi x}{2}} & |x| < 2 \\ 0 & |x| \geq 2 \end{cases} \tag{2}$$

$$Lanczos3(x) = \begin{cases} \frac{\sin(\pi x)}{\pi x} \frac{\sin(\frac{\pi x}{3})}{\frac{\pi x}{3}} & |x| < 3 \\ 0 & |x| \geq 3 \end{cases} \tag{3}$$

表1 主要メディア、規格解像度一覧

画像, 映像	解像度	アスペクト比
映画フィルム(35mm)	1828x1332	1.37
	3656x2664	1.37
PC	1600x1200	4:3
	⋮	5:4 も任意
video(SD)	720x486	4:3,16:9
video(HD)	1920x1080	16:9
	1280x720	16:9
SIF	352x240	22:15
CIF	352x288	11:9
QCIF	176x144	11:9

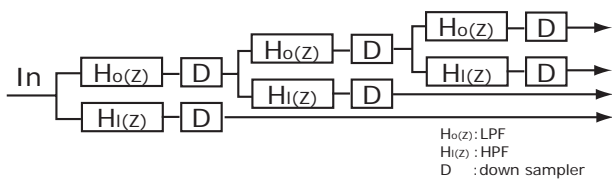


図2 2分割フィルタバンクのオクターブ構成

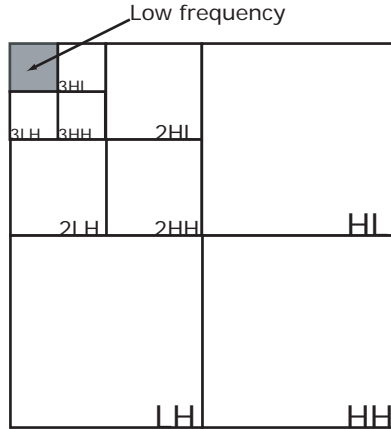


図3 2次元サブバンド分割

表2 DWTにおけるローパスフィルタ係数

I	整数 DWT	実数 DWT
0	6/8	0.6029490182363579
±1	2/8	0.2668641184428723
±2	-1/8	-0.07822326652898785
±3		-0.01686411844287495
±4		0.02674875741080946

2.2 Motion JPEG2000における解像度変換

それに対して、Motion JPEG2000における解像度スケラビリティの機能を示す。Motion JPEG2000における wavelet 変換は離散 wavelet 変換 (DWT: Discrete Wavelet Transform) で実現されている。これは分離型フィルタの考え方に基づき、1次元の2分割フィルタバンクを、低域成分に繰り返し用いて2次元DWTを実現する。このことをJPEG2000ではオクターブ分割と呼ぶ。2分割フィルタバンクはローパスフィルタ、ハイパスフィルタ、ダウンサンブラから成る。このシステム図を図2に示した。

Motion JPEG2000の空間スケラビリティ機能は、図3に示すように、サブバンド分割によって得られた低域成分のみを用いてデコードを行い、縮小サイズでの再生を可能にする。よって、Motion JPEG2000の解像度変換の性能はこの2分割フィルタバンクに用いられているローパスフィルタの性能に因るといことができる。Motion JPEG2000中のDWTは可逆変換が可能な整数DWTと不可逆であるが性能の高い実数DWTの2種類があり、それぞれのローパスフィルタの係数は異なる(整数フィルタ5tap, 実数フィルタ9tap)。表2にフィルタの係数を示す。

3. 実験結果

Motion JPEG2000の映像を配信する際の最適な解像度変換法を提案するための基礎実験として、Motion JPEG2000で用いられるDWTフィルタの特性比較、及び、実際の解像度変換の結果、得られる画質の比較を行った。

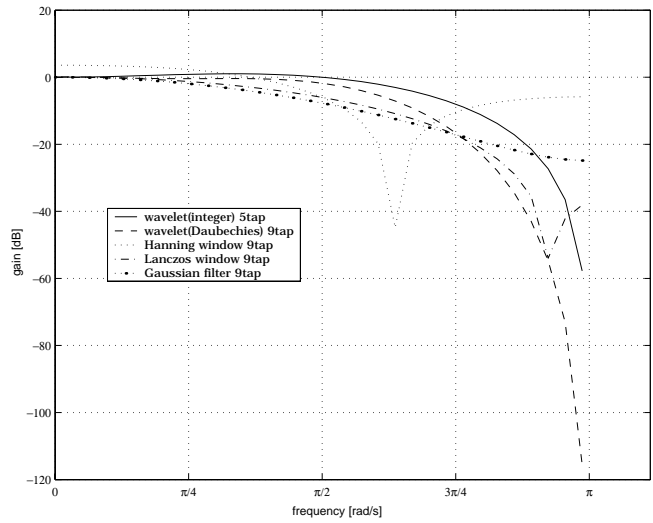


図4 フィルタの周波数応答 (9tap)

3.1 フィルタ特性の比較

まず、Motion JPEG2000中のDWTで用いられるローパスフィルタの性能を評価する。比較対象として、次式(4)(5)からFIRローパスフィルタを設計した。 ω_L がカットオフ周波数、 S がタップ数である。また、窓関数としてハニング関数(6)、Lanczos3関数(lanczos3)を用いた。また、DWTでは、サブバンドごとに1/2にダウンサンプリングすることから、縮小フィルタとして見た時のカットオフ周波数は $\pi/2$ となる。よって、比較対象に設計したFIRローパスフィルタのカットオフ周波数も $\pi/2$ とした。また(1)からガウシアンフィルタを設計し、比較対象とした。

$$H_0 = \frac{\omega_L}{\pi} \quad (4)$$

$$H_m = \frac{\sin(m\omega_L)}{m\pi} \quad (5)$$

$$\mathcal{W}(m) = 0.5(1 + \cos(\frac{m\pi}{n})) \quad (6)$$

$$m = \pm 1, \pm 2, \dots, \pm \frac{S-1}{2}$$

それぞれのフィルタの周波数特性を図4に示す。ここでは wavelet 実数フィルタに合わせ全て9tapのフィルタを設計し、比較を行った。周波数が $\pi/2$ 時のgainが低いほど、高性能のフィルタということになる。これを見るとDWTで用いられているフィルタに比べ、FIRローパスフィルタ、ガウシアンフィルタの方が優れた特性を示している。ただし、DWTのローパスフィルタは、周波数が高くなればなるほど、gainが小さくなるのに対し、FIRローパスフィルタは周波数が高くなると連続的な応答を見せなかった。ガウシアンフィルタは $\pi/2$ でのgainは最も低いものの、特性曲線がなだらかであるため、高周波の折り返しが存在することがわかる。

3.2 解像度比較

次に、実際に縮小処理を行った画像の画質比較を行った。

実験画像としては、図7に示すようなサーキュラゾーンプレートの画像を用いた。サーキュラゾーンプレートの算出式を(7)に示す。これは低域から高

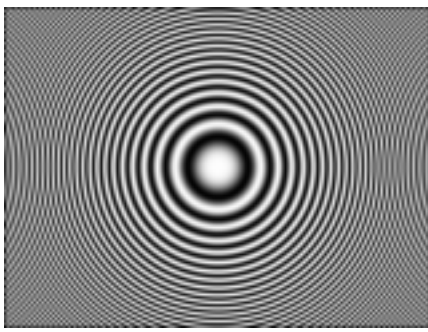


図 5 ゾーンプレート

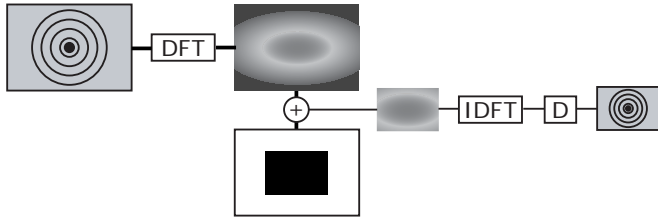


図 6 基準画像生成法

域までの周波数成分の影響を調べるために、周波数成分の分布が視覚的にもわかりやすい画像を用いたためである。画像サイズは HDTV の 1920x1080 のものを用いた。

$$N(x, y) = A \sin\left(\frac{\pi}{2} + \left(\frac{x-a}{\alpha} + \frac{x-b}{\beta}\right) + \theta\right) + B \quad (7)$$

比較対象画像の作成法を図 6 に示す。画像に対し、フーリエ変換で求められた周波数成分を帯域制限することで、理想的なローパスフィルタを実現し、これをダウンサンプリングすることで理想的な縮小処理を実現した。これは処理に膨大な時間がかかり、提案システムにおける縮小処理には不適切であるが、画質の面では理想的であるため、比較対象として用いた。

表 3 に、理想的な縮小画像と、それぞれの縮小アルゴリズムで解像度変換を行った画像との PSNR を示す。ここでは、wavelet フィルタに合わせて、1/2(960 × 540), 1/4(480 × 270), 1/8(240 × 135) の縮小結果に対して測定を行った。また、縮小率に合わせたフィルタの周波数特性を図 7, 8, 9 に示す。

表 3 を見ると、小さなサイズへの縮小をする場合のほうが PSNR の値は高く、フィルタの中では Daubechies の wavelet 実数フィルタが画質的に優れていた。これは wavelet フィルタを多段に用いることで、実質のフィルタのタップ数が増加し、高い性能を示したためと考えられる。

ただし、Lanczos3 に代表されるデシメーションフィルタは、縮小率によってフィルタのタップ数が変化するという特徴がある。そこで、1/2ⁿ の縮小率以外

表 3 解像度変換後の比較 (縮小率 1/2ⁿ)

size	Lanczos3	Gaussian	integer	Daubechies
960x540	30.43(13tap)	30.99(9tap)	29.88(5tap)	34.33(9tap)
480x270	34.50(25tap)	35.95(17tap)	33.28(5tap)	40.66(9tap)
240x135	39.14(49tap)	44.83(33tap)	38.01(5tap)	46.81(9tap)

PSNR[dB]

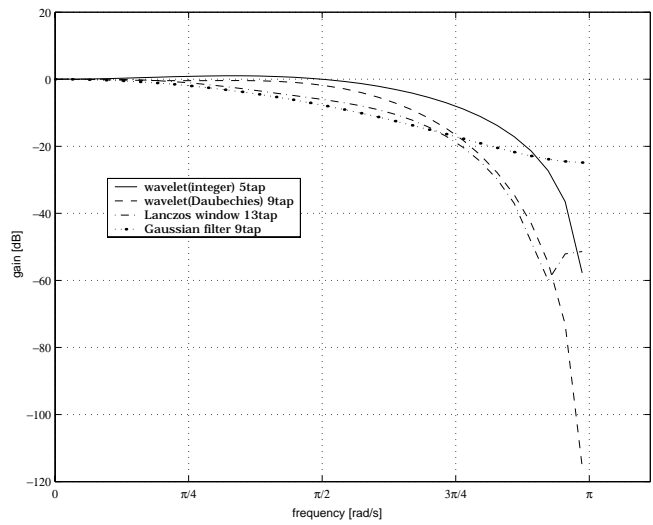


図 7 フィルタの周波数応答 (縮小率 1/2)

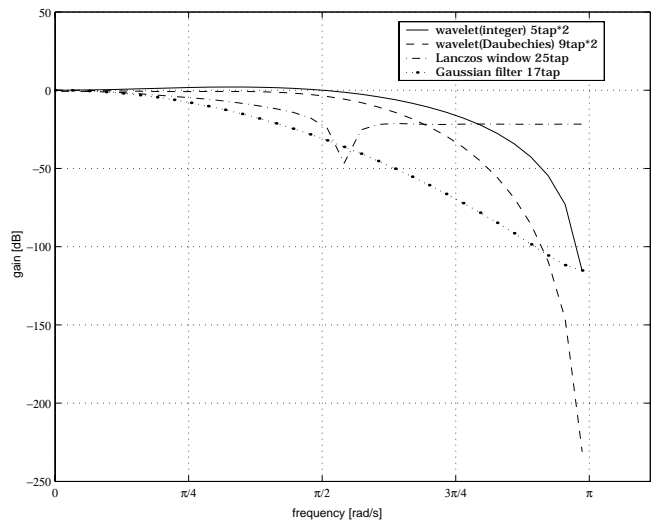


図 8 フィルタの周波数応答 (縮小率 1/4)

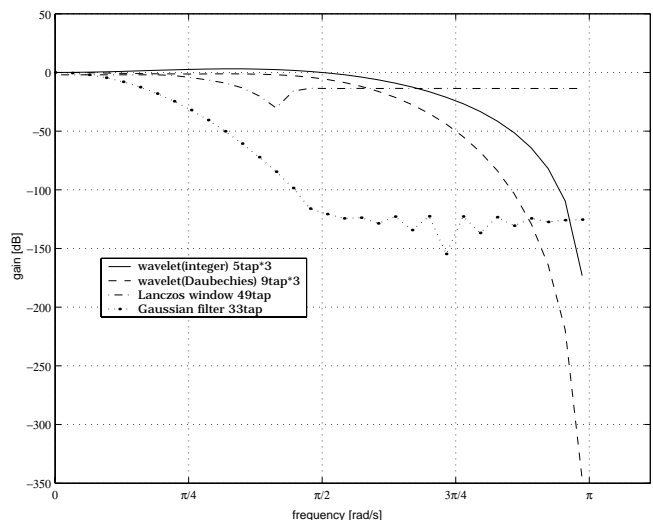


図 9 フィルタの周波数応答 (縮小率 1/8)

の場合についても比較を行った。

1920x1080 から、720x405 への変換 (縮小率 3/8) に対して PSNR を測定した。測定結果を表 4 に示す。この縮小は wavelet フィルタだけでは実現不可能なので、直接デシメーションフィルタを用いる方法と、

表 4 720x405 への解像度変換後の比較

method	PSNR
Lanczos3	31.66(21tap)
Gaussian	32.80(15tap)
LL(integer)+Lanczos3	34.59(5tapx9tap)
LL(integer)+Gaussian	35.32(5tapx7tap)
LL(Daubechies)+Lanczos3	36.64(9tapx9tap)
LL(Daubechies)+Gaussian	35.22(9tapx7tap)

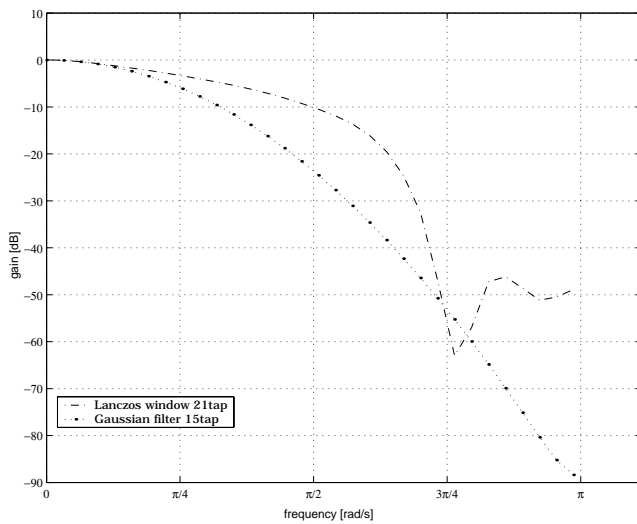


図 10 フィルタの周波数応答 (縮小率 3/8)

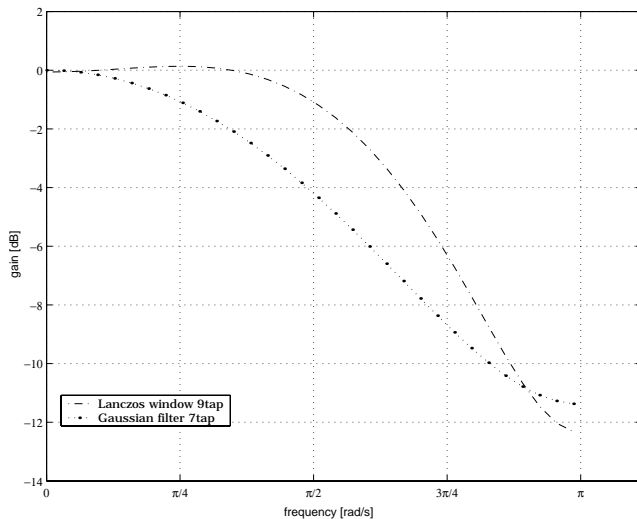


図 11 フィルタの周波数応答 (縮小率 3/4)

wavelet 変換によって得られた LL 成分に対してデシメーションフィルタを用いる方法を比較した。また、この縮小率についてのフィルタの周波数応答を図 10 に示す。また、DWT の LL 成分を用いた場合、デシメーションフィルタの縮小率は $3/4$ になるので、その場合の周波数応答を図 11 に示す。

表 4 を見ると、最も PSNR 値が高いのは wavelet 実数フィルタで得られた低域成分に Lanczos3 フィルタを用いた場合であった。これは、実質的に 9tap のフィルタを 2 回多段にかけることによって実質的な tap 数が増えたためと考えられる。また、図 10, 11 を見ると、デシメーションフィルタの縮小率 $3/8$ の周波数特性はなだらかで、折り返しが多い。このため、いったん性能の良かった Daubechies のフィルタで $1/2$ に縮小し、その後デシメーションフィルタで $3/4$ に縮小した方が性能が良いことになる。 $3/4$ の縮小率時のデシメーションフィルタの性能は比較的良好で、理想的な画質が得られた。

3.3 考察

今回の実験結果を見ると、 $1/2^n$ の縮小サイズに関して、wavelet 実数フィルタの性能が良いことがわかった。さらに、DWT だけでは対応できないサイズに関しては、直接デシメーションフィルタを用いて縮小するよりも、DWT で求められた低域成分に

関してデシメーションフィルタを用いたほうが高画質を実現できた。デシメーションフィルタは処理時間が問題であることから、実際の映像リアルタイム配信のシステムを考慮に入れても利点があり、有用であるといえる。

4. まとめ

本研究では、Motion JPEG2000 に代表される wavelet 変換を用いて符号化された超高精細画像を、配信時に高画質を保ちながら解像度変換を行うことを目的としている。ここでは基礎検討として、wavelet フィルタと、その他のデシメーションフィルタの性能比較、また、実際の解像度変換画像を用いて画質比較を行った。

今後の課題としては

- 様々なサイズでの比較
- 量子化の影響の調査
- 動画としての画質比較
- 処理時間比較

などがあげられる。リアルタイム配信を目指した処理時間の検討、Motion JPEG2000 の特有の量子化である EBCOT の影響、また、wavelet 変換特有のノイズをいわれている flicker artifact [7] の影響などについて調査する必要がある。

謝辞

この研究は、TAO 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” のサポートによる。

文献

- [1] ISO/IEC FCD 15444-1, "JPEG 2000 Part I Final Committee Draft Version 1.0," ISO/IEC JTC 1/SC 29/WG1, N1646R, Mar.2000.
- [2] Takahiro Fukuhara, David Singer, "Motion JPEG2000 Final Committee Draft 1.0," ISO/IEC JTC 1/SC 29/WG1, N2117, Mar.2001.
- [3] Majid Rabbani, Rajan Joshi, "An overview of the JPEG2000 still image compression standard," Signal Processing: image Communication 17, 2002, pp 3-48.
- [4] Ken Turkowski, "Filters for Common Resampling Tasks," Graphics Gems I, Academic Press, 1990, pp.147-165.
- [5] Stephen R. Marschner and Richard J. Lobb, "An Evaluation of Reconstruction Filters for Volume Rendering," Program of Computer Graphics...Proceedings of Visualization '94, 1994
- [6] 高木幹雄, 下田陽久, "画像解析ハンドブック," 東京大学出版, 1991.
- [7] 久下哲郎, "Wavelet 画像符号化の視覚的歪みに関する考察~ インターレース HDTV・超高精細画像応用の課題~, " 社団法人映像情報メディア学会技術報告, Vol.25, No79, pp33-38, Nov.2001