

Motion JPEG 2000 における最適解像度変換法

A study on image resampling for Motion JPEG 2000

田邊 集† Shu TANABE 渡辺 裕‡ Hiroshi WATANABE 富永 英義‡ Hideyoshi TOMINAGA

† 早稲田大学大学院 理工学研究科 ‡ 早稲田大学大学院 国際情報通信研究科

† Graduate School of Sci., and Eng., WASEDA University ‡ Graduate School of GITS, WASEDA University

Abstract : Motion JPEG 2000 is regarded as a promising coding system for ultra-high quality images like Digital Cinema and HDTV. When we distribute such images, it is necessary to resample the image to the resolution suitable for user environment. We can realize spatial scalability easily by Motion JPEG 2000. However, only by Motion JPEG 2000, we cannot get any resolution images. Our goal is to propose the best resampling method for Motion JPEG 2000.

1. はじめに

Motion JPEG 2000 はデジタルシネマや HDTV など、いわゆる超高精細映像の符号化方式として期待されている。超高精細映像を配信する場合、配信先の環境に合わせた解像度変換が必要になる。Motion JPEG 2000 はロスレス画像に対応し、スケーラビリティ機能を有している [1][2]。これは超高精細映像の蓄積、配信に最も適している符号化方式のひとつといえる。しかし、Motion JPEG 2000 で得られる空間スケーラビリティ機能は、 $1/2^n$ の解像度しか実現できない。また、ロスレスに用いられるフィルタを解像度スケーラビリティに用いると、性能が悪いという特徴もある。

我々は以上の問題の解決を目的とし、Motion JPEG 2000 に適した解像度変換法の提案を行う。

2. フィルタ性能比較

解像度変換 (縮小) の主要要素は、ローパスフィルタとサンプリングである。Motion JPEG 2000 における解像度スケーラビリティは、DWT (Discrete Wavelet Transform) によって得られる低域成分をデコードすることによって実現される。DWT はローパスフィルタとハイパスフィルタを組み合わせることで実現されている。Motion JPEG 2000 で用いられているローパスフィルタは 2 種類ある。それぞれの係数を表 1 に示す。

ここでは、Motion JPEG 2000 で用いられるローパスフィルタと、従来画像縮小に用いられるデシメーションフィルタの性能比較を行う。比較に用いたデシメーションフィルタ (Lanczos3 filter and Gaussian filter) を 3.1, 3.2 に示す [3][4][5]。

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

$$Lanczos3(x) = \begin{cases} \frac{\sin(\pi x)}{\pi x} \frac{\sin(\frac{\pi x}{3})}{\frac{\pi x}{3}} & |x| < 3 \\ 0 & |x| \geq 3 \end{cases} \quad (2)$$

2.1 周波数応答特性

ここでは例として、縮小率 $1/2$ 時の各フィルタの周波数応答を図 1 に示す。この時のカットオフ周波数は $\pi/2$ であり、図 1 を見ると、カットオフ周波数において、wavelet フィルタよりデシメーションフィルタのほうがゲインが低いことがわかる。ただし、カットオフ周波数より低周波領域のゲインロスについては、画質劣化を生じることになる。よって、一概にこの値のみでフィルタの性能を評価することはできない。

表 1 The coefficient of LPF in DWT

I	Integer	Daubechies
0	6/8	0.6029490182363579
±1	2/8	0.2668641184428723
±2	-1/8	-0.07822326652898785
±3		-0.01686411844287495
±4		0.02674875741080946

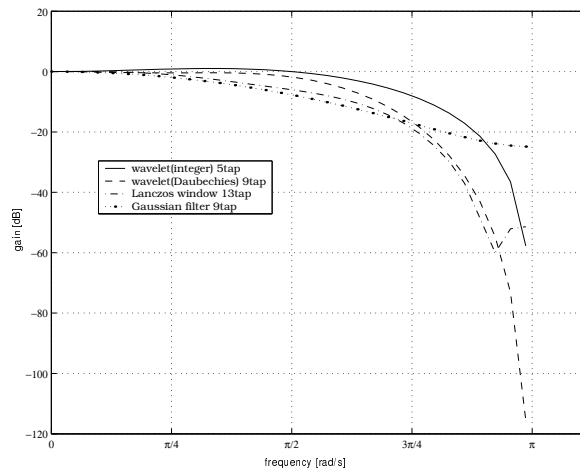


図 1 frequency characteristic(1/2)

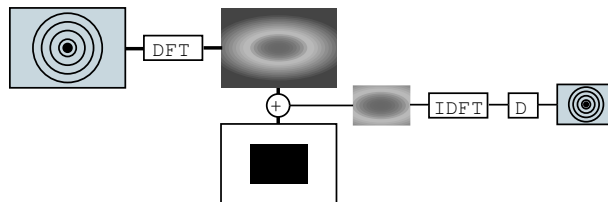


図 2 ideal image resampling

2.2 画質評価

次に、PSNR による画質比較を行った。テスト画像にはサーキュラゾンプレート (1920x1080) を用いた。この画像は周波数成分の影響を視覚的に見ることができる。

ここでは、それぞれのフィルタによる解像度変換と、理想的な解像度変換によって得られる画像の比較を行った。理想的な解像度変換とは、図 2 に示すように、2次元 DFT とダウンサンプリングによるものと定義する。

表 2 に PSNR 値を示す。ここから Daubechies filter の性能がよいことがわかる。また、周波数成分の結果とは異なっていることがわかる。しかし、いずれの場合も Motion JPEG 2000 でのロスレス符号化に用いられる Integer filter は性能が悪いことがわかる。

3. 最適解像度変換への提案

表 2 は縮小率が $1/2^n$ の結果である。これ以外の縮小率は DWT では実現できない。よって、様々な映像フォーマットに対応するためには、デシメーションフィルタと組み合わせて出力する必要がある。

表 2 comparison by PSNR(ratio $1/2^n$)

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]

Motion JPEG 2000 では、デフォルトで 5 回の DWT を行う。よって、5 種類の低域成分が得られることになる。ここでは、縮小率にあわせてどの低域成分を用いるかを選択する手法を提案する。想定システムを図 3 に示す。

ここでは、選択判断要素のひとつとして、“effective tap length”を用いる。DWT での低域成分は、ローパスフィルタを多段にかけることによって得られる。また、ここで用いるデシメーションフィルタである Lanczos3 フィルタは、縮小率によってタップ数が変動する。これらのフィルタを組み合わせて用いた場合、最終的に縮小後の 1 つの画素を構成するのに用いた画素数のことを “effective tap length” と定義する。ただし、係数に関しては考慮していない。まず、Motion JPEG 2000 中で、DWT によって得られる低域成分に関して、“effective tap length” は以下の式によって得られる。

$$A_k = 3A_{k-1} - 2 \quad A_1 = t \quad (3)$$

$$\therefore A_n = 3^{n-1}(t-1) + 1 \quad (4)$$

t : タップ長, n : 繰り返し回数

また、Lanczos3 フィルタは以下の式によってタップ数が決まる。

$$D_{(r)} = \text{Int}\left\{\frac{1}{r} * 3\right\} * 2 + 1 \quad (5)$$

r : 縮小率

よって、最終的な縮小率が r 、Lanczos3 フィルタをある低域成分に対して用いた場合（低域成分は t タップのローパスフィルタを n 回多段にかけたものとする ($n \leq 5$)), “effective tap length” は (6) によって与えられる。

$$T_{(r,n,t)} = \begin{cases} \text{Int}\left\{\frac{1}{r} * 3\right\} * 2 + 1 & (n=0) \\ \text{Int}\left\{\frac{1}{r * 2^n} * 3\right\} * 2 + 3^{n-1}(t-1) + 1 & (n \geq 1) \end{cases} \quad (6)$$

表 3, 4 に “effective tap length” の例を示す。これは 1920x1080 の HDTV の映像から SDTV, CIF, QCIF への解像度変換を行った場合の “effective tap length” である。また、それぞれの方法で解像度変換を行い、理想解像度変換を行った画像との PSNR を表 5, 6 に示す。

4. 考察

表 3, 4 と表 5, 6 を比べると、タップ数の多さと PSNR の高さが対応している。しかし、誤って判断される箇所もある。ここから、この “effective tap length” は有効な判断材料にはなりえるが、これだけで判断するのは十分でないといえる。多段のフィルタ周波数応答などの要素を合わせて検討する必要がある。

5. まとめ

本研究では、Motion JPEG 2000 における最適な解像度変換法の提案を目的とし、“effective tap length” により、解像度変換手法の選択を行うことを提案し、評価実験を行った。“effective tap length” は選択には有用なものの、十分な判断要素としては言えないことがわかった。

今後の検討課題を以下に示す。

- より高解像度 (4kx2k) での実験
- 判断要素の追加, 検討
- 動画としての画質評価, 量子化の影響の評価

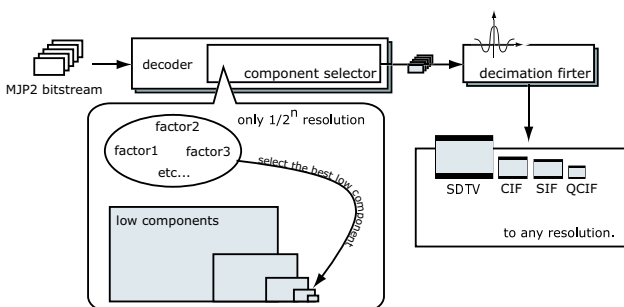


図 3 resampling process

表 3 effective tap length (Integer)

low component	to SD	to CIF	to QCIF	to 88x60
directly Lanczos3	17	31	61	121
use LL1(960x540)	13	19	35	65
use LL2(480x270)	-	19	27	43
use LL3(240x135)	-	-	43	51
use LL4(120x67)	-	-	-	115
use LL5(60x34)	-	-	-	-

表 4 effective tap length (Daubechies)

low component	to SD	to CIF	to QCIF	to 88x60
directly Lanczos3	17	31	61	121
use LL1(960x540)	17	23	39	69
use LL2(480x270)	-	31	39	55
use LL3(240x135)	-	-	79	87
use LL4(120x67)	-	-	-	223
use LL5(60x34)	-	-	-	-

表 5 PSNR(Integer)

low component	to SD	to CIF	to QCIF	to 88x60
directly Lanczos3	31.66	35.92	40.56	45.71
use LL1(960x540)	29.88	35.50	40.42	45.47
use LL2(480x270)	-	33.47	40.03	45.55
use LL3(240x135)	-	-	38.19	45.36
use LL4(120x67)	-	-	-	45.61
use LL5(60x34)	-	-	-	-

表 6 PSNR(Daubechies) PSNR[dB]

low component	to SD	to CIF	to QCIF	to 88x60
directly Lanczos3	31.67	35.92	40.51	45.45
use LL1(960x540)	31.17	36.00	40.58	45.51
use LL2(480x270)	-	35.53	40.70	45.81
use LL3(240x135)	-	-	40.44	46.25
use LL4(120x67)	-	-	-	46.58
use LL5(60x34)	-	-	-	-

PSNR[dB]

謝辞

本研究は、通信放送機構からの委託研究「通信ネットワーク利用放送技術の研究開発」に基づき行われた。

参考文献

- [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 JPEG 2000 Final Committee Draft 1.0,” ISO/IEC JTC 1/SC 29/WG1, N2117, Mar. 2001.
- [3] Ken Turkowski, “Filters for Common Resampling Tasks,” Graphics Gems I, Academic Press, 1990, pp. 147-165.
- [4] Stephen R. Marschner and Richard J. Lobb, “An Evaluation of Reconstruction Filters for Volume Rendering,” Program of Computer Graphics...Proceedings of Visualization '94, 1994
- [5] Philippe Thevenaz et al, “Image Interpolation and Resampling,” Handbook of Medical Imaging, Processing and Analysis, I.N. Bankman, Ed., Academic Press, San Diego CA, USA, pp. 393-420, 2000.

早稲田大学大学院 理工学研究科

〒169-8555 東京都新宿区大久保 3-4-1 55 号館 N-06-02

Phone : 03-5236-9839

E-mail : shu@tom.comm.waseda.ac.jp