

ウェーブレット変換係数の二次元可変長符号化に関する一検討

嵯峨田 淳 八島 由幸 小林 直樹

NTT サイバースペース研究所

〒 239-0847 横須賀市光の丘 1-1
TEL: 0468-59-4673, FAX: 0468-59-2829
E-mail: sagata@nttvdt.hil.ntt.co.jp

従来のゼロツリー符号化はゼロツリールートの出現確率が極めて高いため、単純に可変長符号化した場合、そのエントロピーと比較して効率が低下する。このため、各シンボルの符号化には、一般的に算術符号化が用いられる。また、ゼロツリー符号化はその符号体系自体に冗長性があり、親子係数間の相関を効率的に除去できないという問題がある。本稿では、ある変換係数の直接の子供係数のうち、この子供係数が ZTR である個数と、変換係数のレベルの絶対値からなる二次元可変長符号を用いてエントロピー符号化を行う。本手法により、解像度スケーラビリティを有しつつ、符号化効率を損なうことなく、可変長符号化を用いてウェーブレット変換係数を符号化することができる。

A Study on Two Dimensional Variable Length Coding of Wavelet Coefficients

Atsushi SAGATA Yoshiyuki YASHIMA Naoki KOBAYASHI

NTT Cyber Space Laboratories

1-1 Hikarino-oka, Yokosuka, Kanagawa 239-0847, JAPAN
TEL: +81-468-59-4673, FAX: +81-468-59-2829
E-mail: sagata@nttvdt.hil.ntt.co.jp

This paper describes an efficient and effective entropy coding strategy for the compression of wavelet transformed, motion compensated prediction residuals. The proposed entropy coding algorithm consists of two dimensional variable length coding (2D-VLC), which encodes both the level of transformed coefficient and the number of ZTRs in its direct children-coefficients simultaneously. At this time, 2×2 cluster of adjacent wavelet coefficients are jointly coded with the 2D-VLC to exploit the spatial correlation of neighboring coefficients. Along with the 2D VLCs of each cluster, the proposed zerotree-map code is appended at the head of them to remove the redundancies of any conventional zerotree encoding algorithms. Simulation results show that the proposed coding method is superior to conventional arithmetic coding, or variable length coding with the length ZTR runs.

1 はじめに

従来の静止画像又は動画画像圧縮符号化には、離散コサイン変換が広く用いられてきた。しかしながら、離散コサイン変換による画像符号化は、ある程度以上の圧縮率では、変換処理単位であるブロックの境界が不連続になるブロック歪みや、急峻なエッジ部周辺に発生するモスキートノイズなどの特有の

画質劣化が主観画質上問題となっていた。これらの問題点を回避する方法としてウェーブレット変換 [1] を利用した符号化方式が種々検討されている。

また、低周波係数から高周波係数に向けてウェーブレット変換係数がピラミッド状に生成される、すなわち解像度スケーラビリティを有する特性を利用して、ウェーブレット変換を用いた符号化方式を QoS 非保証のネットワークを介した映像伝送 [2] に

利用する提案もなされている。

ウェーブレット符号化において、その変換係数をエントロピー符号化する方法として、ゼロツリー符号化 [3] が提案されている。ゼロツリー符号化は特定のシンボルの出現確率が極めて高いため、これを単純に可変長符号化した場合、そのエントロピーと比較して著しく符号化効率が低下する。このため、ゼロツリー符号化におけるウェーブレット変換係数のエントロピー符号化には、一般的に算術符号化が用いられる。

本稿では、スケーラビリティ機能を有し、符号化効率を損なうことなく、ウェーブレット変換係数を可変長符号化する方法を検討する。

2 ウェーブレット変換係数符号化

2.1 ゼロツリー符号化

多くのウェーブレット変換を用いた符号化方式では、対応する周波数帯域間で相似な場所の係数間にツリーを生成し、これを低周波係数から高周波係数にかけて順に符号化する。ウェーブレット変換係数に対しエントロピー符号化を行う方式として、ゼロツリー符号化が提案されている。ゼロツリー符号化では、ある変換係数 C_0 をルート（幹）とするサブツリー（部分木）内の変換係数を量子化した結果、量子化値が全てゼロとなる場合、このサブツリーをゼロツリーと呼び、変換係数 C_0 を一つのゼロツリールート（Zero Tree Root, ZTR）符号に置き換えて符号化することでデータ圧縮を行う。この時点で ZTR で置換されたサブツリーは符号化済みであり、後段で変換係数 C_0 の子孫係数 $D(C_0)$ を符号化

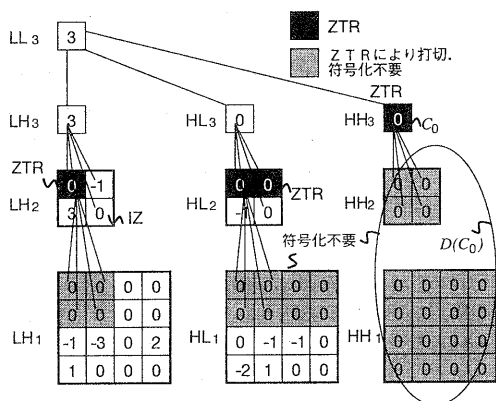


図 1: ゼロツリー符号化

する必要はない。ゼロツリー符号化の一例を図 1 に示す。

また、ZTR 符号とは逆に、ある変換係数の量子化値自体はゼロであるものの、その子孫係数に一つでも有意係数がある場合は、変換係数を孤立ゼロ (Isolated Zero, IZ) 符号として符号化する。

2.2 EZW 符号化

一方、量子化の結果変換係数が有意係数である場合は、ZTR・IZ 符号のいずれの条件にも合致しない。この場合、図 2 に示す EZW 符号化 [3] を行うのが一般的である。EZW 符号化では、まず閾値 T を設定し、最低周波帯域から高周波帯域に向け順に走査しながら各変換係数の絶対値が閾値より大きいかを調べる。入力変換係数 C_0 の絶対値が閾値より小さい場合は、先に述べた ZTR あるいは IZ として符号化する。大きい場合はこれを重要係数と定義し、 C_0 の符号が正のときは POS, 負のときは NEG を出力する。同時にその近似精度を高める符号を出力する。

以上の低周波帯域から順に走査しながら符号化する操作を、閾値 T を半分の値にして再帰的に行い、 T が 1 になるまで繰り返す。この時、直前の操作までで重要係数とされた係数に関しては、近似精度を高める符号のみを出力する。

一般的に T の初期値は、変換係数の最大の絶対値を越える最小の 2 の階乗の値とするため、EZW 符号化では有意係数は実質的にビットプレーン展開される。MSB より順にゼロツリー符号化されることになる。このため、画品質スケーラビリティを利用することが可能となる。

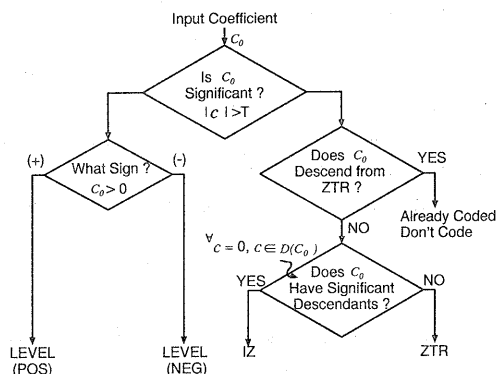


図 2: EZW 符号化

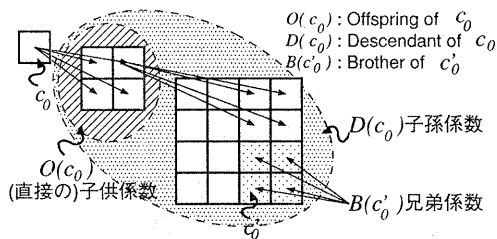


図 3: ツリー構造での係数の定義

この画品質スケーラビリティに加えて、多くのウェーブレット変換符号化方式では、低周波係数から高周波係数にかけて順に符号化・伝送するため、解像度スケーラビリティの機能を有する。静止画像の符号化では、解像度・画品質スケーラビリティを組み合わせた符号化データを生成する事により、復号器では、既に受信済みの符号化データだけを用いて漸次復号し、時間の経過とともに段階的に映像を出力する事が可能となる。

2.3 EZW 符号化を利用しない有意係数符号化方法

動画像の符号化の場合、1/30 秒程度の周期でフレームが切り替わるため、EZW 符号化のような細分化されたスケーラビリティは必要がない。さらに EZW 符号化はビットプレーン展開による分割損が大きいため符号化効率が低下する。このため、動画像すなわち予測差分信号の符号化の場合には、変換係数をスカラ量子化したものを直接符号化 [4] する方式も多く提案されている。

以下、本稿においてゼロツリー符号化と述べる場合は、EZW 符号化を行わず、ZTR・IZ 符号・有意係数の量子化番号（以下 レベル）の絶対値 $|Q[c]|$ を可変長符号化する方式を示すものとする。

2.4 ゼロツリー符号化の問題点

以上に示したゼロツリー符号化は有効な符号化方式であるが、ゼロツリー符号化にはいくつかの問題点があると考えられる。

(a) 符号化体系の冗長性

ある変換係数 C_0 が IZ 符号として符号化される場合、(1) 変換係数 C_0 のレベルがゼロであり、(2) 子孫係数 $D(C_0)$ の中に一つ以上有意係数が存在することを示している。従って、図 3 に示すように変換係数 C_0 の直接の子供係数を $O(C_0)$ と定義した場合、 $c \in O(C_0)$ が全て ZTR であることはあり得

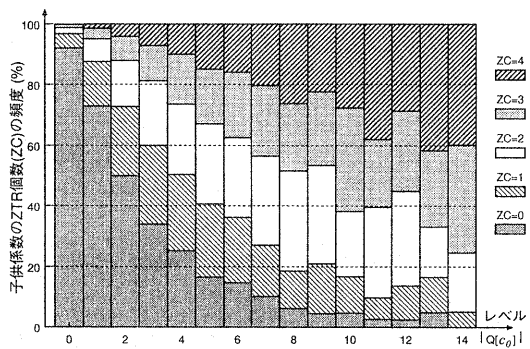


図 4: レベル絶対値と子供係数の ZTR 個数の相関

えない。しかし、実際には $c \in O(C_0)$ が全て ZTR となっている符号化データを、復号器は復号することができる。符号化器から出力される可能性のない符号を復号器が復号可能と言う事実は、符号体系に冗長性があることを示している。

(b) 絶対値と子供係数の相関を未利用

ある変換係数 C_0 のレベル ($Q[C_0]$) がゼロである場合、子孫係数 $D(C_0)$ のレベル ($Q[c]$, $c \in D(C_0)$) が全てゼロになる確率が高くなる特性を利用して、これを一つの ZTR 符号に置換して圧縮を行う。しかし実際には、図 4 に示すように、レベル $Q[C_0]$ がゼロの場合だけではなく、レベルの絶対値 $|Q[C_0]|$ が小さい程、4 つの子供係数 $c \in O(C_0)$ を幹とするサブツリーがそれぞれ ZTR となる確率が高くなり、逆に $Q[C_0]$ の絶対値が大きい程、 $c \in O(C_0)$ がそれぞれ ZTR となる確率が低くなる。

(c) ZTR の出現確率が極めて高い

2.3 節に説明した通り、本稿のゼロツリー符号化において、実際に可変長符号化を行う対象のシンボルは、ZTR・IZ・各 $|Q[C_0]|$ である。しかし実際には、符号化ビットレートが低い程、ZTR の出現確率が増加し、一般的に 50% 以上となる。このため、可変長符号化を行う入力シンボルの出現確率をそれぞれ $1/2^n$ で近似することができず、平均符号長が、そのエントロピーと比較して著しく増大する。

3 提案方式

3.1 子供係数の ZTR 数を用いた二次元可変長符号

以上の問題を解決するため、ある変換係数 C_0 を符号化するにあたり、子供係数 $O(C_0)$ を幹とするサ

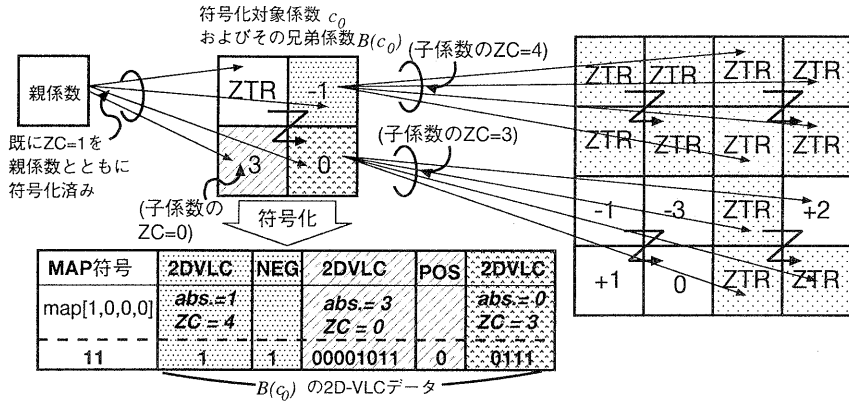


図 5: 係数およびその兄弟係数の符号化

ツリーがゼロツリーである個数 (ZC) と、変換係数 C_0 のレベルの絶対値 $|Q[C_0]|$ からなる二次元可変長符号を用いてエントロピー符号化する。これにより、問題点 (b) に示した親係数のレベルと子供係数の ZTR の相関を効率的に利用することができる。

また、問題点 (a) を解決するため、変換係数 C_0 を符号化するにあたり、変換係数 C_0 の兄弟係数 $B(C_0)$ を一単位として走査し、それぞれ符号化データに出力する。 $B(C_0)$ 内の ZTR である係数は、親係数を符号化した際に出力済みの $B(C_0)$ の ZC と、3.2 節に後述するゼロツリーマップにより、どの係数が ZTR であるかが一意に定まる。このため $B(C_0)$ 内の ZTR 符号を単独で出力する必要はない。

以上の方式を用いて、図 1 の LH_2 帯域内の 4 係数を提案方式を用いて符号化する例を図 5 に示す。4 係数 $B(C_0)$ を符号化するにあたり、左上の係数は ZTR なので符号化不要。右上・左下の係数は、それぞれ $|LEVEL|=1$, $ZC=4$ と $|LEVEL|=3$, $ZC=0$ の提案二次元可変長符号を出力したのちに、前者はマイナスを示す NEG、後者はプラスを示す POS の 1 ビット符号を出力する。右下の係数は、同様に $|LEVEL|=0$, $ZC=3$ の二次元可変長符号を出力するが、レベルがゼロであるため符号の 1 ビットは必要ない。

3.2 ゼロツリーマップ

兄弟係数内の ZTR の符号化のため、兄弟係数を一単位として符号化するにあたり、3.1 節で説明した二次元可変長符号を符号化データに出力する直前に、 $B(C_0)$ 中のどの変換係数が ZTR であることを示すゼロツリーマップを出力する。

通常、4 係数の ON/OFF を表すためには 4 ビット

必要である。提案するゼロツリーマップの符号化では、 $B(C_0)$ の ZC が 1, 2, 3 の場合に対応する符号化テーブルを切替えて符号化する。これにより、符号量は兄弟係数あたり 2, 3 ビットで済む。復号器では、 C_0 の親係数を符号化したときに符号化データに出力済みである $B(C_0)$ の ZC を元に、符号化器・復号器の間で共有される対応する符号化テーブルを用いて復号する。

$B(C_0)$ の ZC が 0 あるいは 4 の場合は、ゼロツリーマップを出力する必要はない。0 の場合は $B(C_0)$ 内の全係数が ZTR でないこと、4 の場合は全係数が ZTR であることを一意に示すためである。ゼロツリーマップの符号化テーブルの一例を表 1 に示す。

表 1: ゼロツリーマップの符号化テーブル

ZC	$B(c)$ 内の ZTR の位置	VLC
0	符号化不要	—
1	[—, —, —, ZTR]	00
	[—, —, ZTR, —]	01
	[—, ZTR, —, —]	10
	[ZTR, —, —, —]	11
2	[—, —, ZTR, ZTR]	01
	[—, ZTR, —, ZTR]	100
	[—, ZTR, ZTR, —]	101
	[ZTR, —, —, ZTR]	110
	[ZTR, —, ZTR, —]	111
	[ZTR, ZTR, —, —]	00
3	[—, ZTR, ZTR, ZTR]	00
	[ZTR, —, ZTR, ZTR]	01
	[ZTR, ZTR, —, ZTR]	10
	[ZTR, ZTR, ZTR, —]	11
4	符号化不要	—

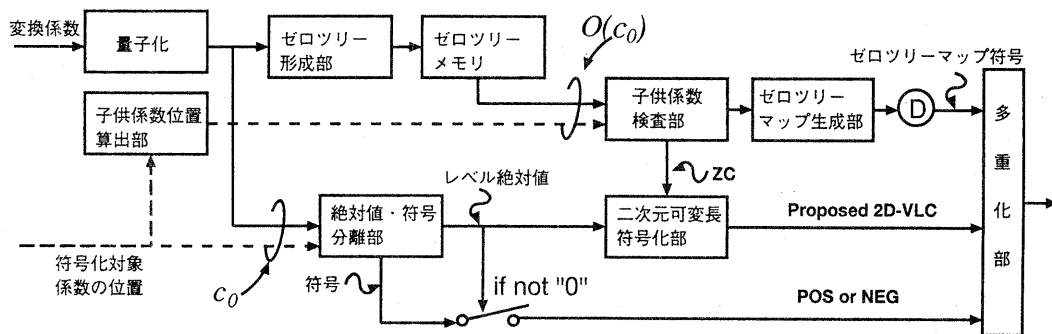


図 6: 提案手法の回路図

3.3 兄弟係数の走査

一般的に水平方向の帯域は、水平方向に高域フィルタ、垂直方向に低域フィルタをかけた信号であるため、垂直方向のエッジ成分が現れやすく、同様に垂直方向の帯域は水平方向に、対角方向の帯域は斜め方向にエッジ成分が現れやすいことが知られている。

兄弟係数を一単位に符号化する際、この周波数帯域内の変換係数間の相関を利用し、相関の高い方向に優先して各兄弟係数内を走査する。具体的には垂直方向の帯域内の兄弟係数を走査するときは左上・右上・左下・右下の順に、水平方向のときは左上・左下・右上・右下の順に、対角方向のときは左上・右下・右上・左下の順に兄弟係数を走査する。この走査により、走査順に1, 2番目の変換係数と、3, 4番目の変換係数に大きな相関が生まれる。表1に示すように、ZCが2のときのこれらのパターンの符号長は他より短くしてあるため、ゼロツリーマップの平均符号長を効率的に短くする事ができる。

3.4 最高周波帯域の符号化

最高周波帯域である LH_1 , HL_1 , HH_1 帯域内の係数には子供係数が存在しないため、レベルの絶対値と ZC の二次元可変長符号を用いて符号化することができない。したがって、最高周波帯域の有意係数に関しては、従来のゼロツリー符号化と同様に、レベルの絶対値を一次元可変長符号化する。ただし、ゼロツリー符号化とは異なり、親係数を符号化したときに最高周波数帯域内の各兄弟係数の ZC は符号化データに記述済みであるため、兄弟係数の一次元可変長符号の前に同様にゼロツリーマップを出力する事で、最高周波帯域内のゼロ係数を符号化する。

4 計算機シミュレーション

4.1 シミュレーション条件

提案方式の効果を、シミュレーションにより確認した。比較した従来ゼロツリー方式では、ZTRのランレングス・IZ・各レベルの絶対値を一次元可変長符号化した。有意係数のレベル符号化に用いた符号化テーブルは、提案方式の最高周波数帯域のレベルの符号化のときに用いる符号化テーブルと同様のものを用いた。なお、ZTRにランレングス符号化を用いたのは、ZTRの出現確率が圧倒的に高いため、ZTRを単独で可変長符号化するのは比較の対象として公平でないためである。

同時に、ランレングス符号化を用いず、ZTR・IZ符号・各有意係数の絶対値からなる入力シンボルに対して単純に算術符号化を施す手法の発生符号量の参考値も算出した。この参考値は上記入力シンボルのそれぞれの出現確率からエントロピーを算出し、これを変換係数の発生符号量とした値である。

シミュレーションに用いた映像は、標準画像 Foreman と、Silent Voice (QCIF, 30 f/s) であり、64 ~ 300 kb/s 程度の符号化ビットレートで比較を行った。フレーム間予測には、動き補償を伴う時間方向ハル変換 [5] を二階層に施し、4 フレームを1 ユニットとして符号化した。各ユニットの平均画像間の予測にはオーバーラップ動き補償 [6] を用いた。ウェーブレット変換には7-9 フィルタ [1] を用いて、周波数帯域分割を三回行った。提案方式に用いた二次元可変長符号は、上記の二つの標準画像を除いた複数の映像を上記のレートで符号化し、ZC とレベルの出現確率から作成した。係数の量子化にはセンターデッドゾーン付きのスカラ量子化器 [7] を用いた。

4.2 シミュレーション結果および考察

4.1節に示した条件に基づいたシミュレーションの結果を、図7, 8に示す。縦軸にPSNR(dB), 横軸に符号化ビットレート (kbps)を示す。

提案方式による発生符号量をエントロピーと比較した場合、中程度の動きのForemanの場合で14.4~23.3%, 動きの少ないSilent Voiceの場合では19.3~26.8%削減されている。

動きの少ないSilent Voiceでは、動き補償が比較的效果があるため、多くの変換係数のレベルがゼロとなる。また、図4より、変換係数のレベルが低い方が、よりZCとレベルの絶対値の相関が高いことが判る。このため、提案方式とエントロピーとの比較に関しては、動きの少ないSilent Voiceの方がより効率が改善したものと考えられる。

また、提案方式による発生符号量を従来手法であるZTRをランレングス符号化した場合と比較すると、Foremanの場合で6.1~9.5%, Silent Voiceの場合で3.4~8.5%符号量が削減された。これは、動きの少ないSilent Voiceの場合、ZTRの出現確率が70.8~92.2%と極めて高いため、ZTRのランレングス符号化でも、ある程度効率的に符号量を削減できたものと考えられる。

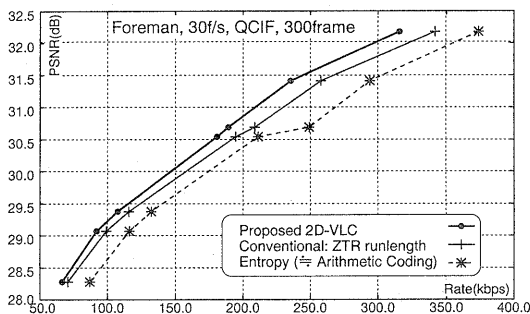


図7: シミュレーション結果 (Foreman)

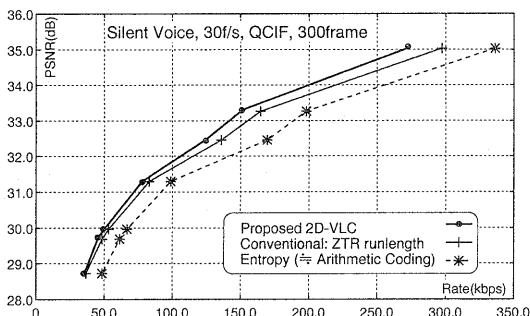


図8: シミュレーション結果 (Silent Voice)

5 まとめ

本稿では、親子係数間の相関を利用した二次元可変長符号を用いた符号化方式を提案した。本提案方式により、スケーラビリティ機能を有しつつ、入力シンボルを単純に算術符号化する場合よりも符号化効率を改善する事を示した。また、出現確率の高いZTRをランレングス符号化する場合よりも良好な符号化効率を得る事も明らかになった。

本提案方式では、同一サブバンド内の空間方向の相関はゼロツリーマップによりある程度取り除かれているものの、ZTRのランレングス符号化方法と比較するとまだ十分にはこの相関を利用できていない。今後は、両方式を統合した、より高効率なスケーラブル符号化方式を検討していく予定である。

参考文献

- [1] I. Daubechies, "Orthonormal bases of compactly supported wavelets," Communications on Pure and Applied Mathematics, vol.41, pp.909-996, 1988
- [2] M. Vishwanath et al. "An efficient algorithm for hierarchical compression of video," Proc. IEEE International Conference on Image Processing, pp.275-279 Nov. 1994
- [3] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," IEEE Trans. Signal Processing, vol.41, no.12, pp.3445-3462, Dec. 1993
- [4] S. A. Martucci et al. "A zerotree wavelet video coder," IEEE Trans. Circuits and Systems for Video Technology, vol.7, no.1, pp.109-118, Feb. 1997
- [5] J.-R. Ohm, "Temporal domain sub-band video coding with motion compensation," Proc. International Conference on Acoustics, Speech and Signal Processing, Vol. 3, pp.229-232, Apr. 1992
- [6] H. Watanabe and S. Singhal, "Windowed motion compensation," Proc. SPIE conf. Visual Communications and Image Processing '91, Vol. 1605, pp.582-589, Nov. 1991
- [7] ITU-T Rec. H.263, "Video coding for low bitrate communication", Dec. 1995.