

動き補償予測誤差画像に対するサブバンド Matching Pursuits 符号化

Matching Pursuits Sub-band Coding for Motion Compensated Prediction Error Image

今村 幸祐[†], 石崎 智範^{††}, 正会員 橋本 秀雄[†]

Kousuke Imamura[†], Tomonori Ishizaki^{††} and Hideo Hashimoto[†]

Abstract Most recent video coding methods like MPEG, are types of hybrid coding that use motion compensation(MC) and discrete cosine transform(DCT). In DCT, because images are partitioned into macro blocks, the efficiency decreases for a very low bitrate. We propose a new video coding method that uses matching pursuits(MP). MP represents a signal approximately by using a dictionary to obtain a high coding efficiency for a very low bitrate. In this paper, we propose a matching pursuits coding method that uses a dictionary suitable for the sub-band frequency characteristic for motion compensated prediction error images, and will demonstrate how this method can improve coding efficiency.

キーワード: Matching Pursuits, 動き補償予測, サブバンド分割

1. ま え が き

現在, MPEG や H.26x 等の動画画像符号化方式は動き補償(MC)と離散コサイン変換(DCT)を組合せた方法が主流となっている。DCT 符号化は, 画面全体をブロック単位に分割して符号化を行うため, ビットレートが非常に低い場合に効率が低下しやすい。近年, Matching Pursuits (MP) と呼ばれる, 符号化対象信号に適した辞書を用いて, 信号を近似構成する手法が提案されている¹⁾。MP は信号中の電力が最も大きい信号成分から近似を行っていくため, 特に低いビットレート時に高い符号化効率を示す。この MP を DCT の代わりに用いた動画画像符号化方式も提案されている^{2)~5)}。

MP の符号化特性は, 符号化に用いる辞書に大きく依存する。本論文では, MC + MP による動画画像符号化システムにおける, MP の符号化対象である予測誤差画像をサブバンド分割し, 各バンドの特徴を考慮した辞書を設計することで符号化効率の改善を行う。

2. Matching Pursuits

2.1 Matching Pursuits の基本原理

Matching Pursuits(MP) は, 非直交で overcomplete な基底ベクトル $g_\gamma(t)$ によって構成された辞書 \mathcal{D} を用意し, 符号化対象信号 $f(t)$ を, 辞書に登録された波形の線形結合で近似する手法である。

MP による近似信号 $\hat{f}(t)$ は式 (1) で表される。

$$\hat{f}(t) = \sum_{k=1}^m p_k \cdot g_{\gamma_k}(t - \tau_k) \quad (1)$$

ここで τ_k , γ_k , p_k はそれぞれ基底ベクトルの位置, 種類, 係数を表し, これらの3種類のパラメータの組 (τ_k, γ_k, p_k) をアトム(atom)と呼ぶ。式(1)を構成するアトムの集合は, 信号 $f(t)$ を基底ベクトル $g_\gamma(t)$ へ直交射影することで決定される。

まず, 信号 $f(t)$ をある基底ベクトル $g_{\gamma_1}(t - \tau_1)$ へ射影した場合, 残差信号 $Rf(t)$ を用いて

$$f(t) = p_1 \cdot g_{\gamma_1}(t - \tau_1) + Rf(t) \quad (2)$$

$$p_1 = \langle f(t), g_{\gamma_1}(t - \tau_1) \rangle \quad (3)$$

と表せる。 p_1 は信号と基底ベクトルの内積値を示す。また基底ベクトル $g_\gamma(t)$ が正規化されているものとする, $Rf(t)$ と $g_{\gamma_1}(t - \tau_1)$ は互いに直交し, 以下の関係式が成立する。

2003年11月10日受付, 2004年1月9日採録

[†] 金沢大学工学部

(〒920-8667 石川県金沢市小立野 2-40-20, 076-234-4894)

^{††} ソニー株式会社

(〒141-0032 品川区大崎 2-10-14, 03-3495-3013)

[†] Faculty of Engineering, Kanazawa University

(2-4-20 Kodatsuno Kanazawa-shi, Ishikawa, 920-8667 Japan)

^{††} Sony Corporation

(2-10-14 Osaki Shinagawa-ku, Tokyo, 141-0032 Japan)

$\alpha\beta$	s	ξ	ϕ
1	2	0	0
2	3	0	0
3	4	0	0
4	5	0	0
5	6	0	0
6	8	0	0
7	10	0	0
8	11	0	0
9	1	1	$\pi/2$
10	5	1	$\pi/2$
11	11	2	$\pi/2$
12	10	3	0
13	8	2	0
14	4	2	0
15	4	2	$\pi/4$
16	6	4	$\pi/4$

図1 Gabor関数のパラメータと辞書
Parameters of Gabor function and dictionary.

$$\|f(t)\|^2 = p_1^2 + \|Rf(t)\|^2 \quad (4)$$

したがって、内積 p_1 の絶対値を最大とするパラメータを求めることにより、残差信号のエネルギーの減少量 $\Delta e = \|f(t)\|^2 - \|Rf(t)\|^2$ を最大とするアトムを決めることができる。

また、残差信号 $Rf(t)$ は

$$Rf(t) = p_2 \cdot g_{\gamma_2}(t - \tau_2) + R^2 f(t) \quad (5)$$

のように展開することができることから、 m 個のアトムにより展開された信号 $f(t)$ は

$$f(t) = \sum_{k=1}^m p_k \cdot g_{\gamma_k}(t - \tau_k) + R^m f(t) \quad (6)$$

と表すことができる。

2.2 Matching Pursuits における辞書

Matching Pursuits で用いる辞書は式(8)に示す Gabor 関数から生成される2次元分離可能な辞書である。

$$G_{\alpha,\beta}(i, j) = g_{\alpha}(i) \otimes g_{\beta}(j) \quad (7)$$

\otimes : 直積

$$g_{\gamma}(i) = K_{\gamma} g\left(\frac{i - \frac{N}{2} + 1}{s}\right) \cos\left(\frac{2\pi\xi(i - \frac{N}{2} + 1)}{N} + \phi\right) \quad (8)$$

$g(\cdot)$ は Gaussian 関数であり、 K_{γ} は正規化係数である。この Gabor 関数は $\gamma = (s, \xi, \phi)$ により決まる。

図1に経験的に用いられている Gabor 関数のパラメータとそれから作成される辞書を示す。

3. サブバンド Matching Pursuits を用いた動画画像符号化

サブバンド Matching Pursuits を用いた動画画像符号化システム (MC+SBMP 方式) を図2に示す。

この符号化システムでは、動き補償予測の後、予測誤差フレームをサブバンド分割し、各バンドに対して Matching Pursuits を適用するものである。

Matching Pursuits 部においては、符号化効率向上のために、各サブバンドに適した辞書を設定する。サブバンド画像は、1/2 にサブサンプリングされた画像であることか

研究速報 □ 動き補償予測誤差画像に対するサブバンド Matching Pursuits 符号化

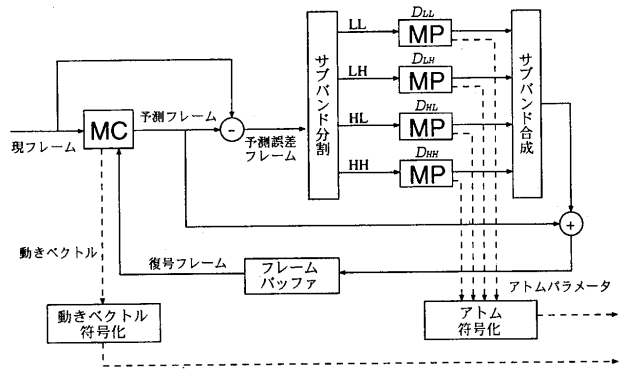


図2 サブバンド Matching Pursuits を用いた動画画像符号化システム

Video coding system using Sub-band Matching Pursuits.

表1 サブバンド用辞書パラメータ
Dictionary parameters for sub-band.

(a) LL バンド

$\tilde{\gamma}_{LL}$	s	ξ	ϕ	$\tilde{\gamma}_{LL}$	s	ξ	ϕ
0	1	0	0	8	0.5	1	$\pi/2$
1	1.5	0	0	9	2.5	1	$\pi/2$
2	2	0	0	10	5.5	1	$\pi/2$
3	2.5	0	0	11	5	3	0
4	3	0	0	12	4	2	0
5	4	0	0	13	2	2	0
6	5	0	0	14	2	2	$\pi/4$
7	5.5	0	0	15	3	4	$\pi/4$

(b) 低周波数バンド

$\tilde{\gamma}_L$	s	ξ	ϕ
0	1	0	0
1	2	0	0
2	3	0	0
3	4	0	0
4	2	1	0
5	2	1	$\pi/2$
6	4	1	0
7	4	1	$\pi/2$

(c) 高周波数バンド

$\tilde{\gamma}_H$	s	ξ	ϕ
0	2	2	0
1	2	2	$\pi/2$
2	4	2	0
3	4	2	$\pi/2$
4	2	3	0
5	2	3	$\pi/2$
6	4	3	0
7	4	3	$\pi/2$

ら、辞書についても基底長が半分の Gabor 関数を用い、水平垂直共に 1/2 サイズのブロックの辞書を作成した。予測誤差画像を LL, LH, HL, HH にサブバンド分割することにより、LH バンドに水平エッジ、HL バンドには垂直エッジといった特徴的な成分があらわれる。各サブバンドに適した辞書作成のために、表1に示すようにパラメータを設定した。LL バンドについては、サブバンド画像が 1/2 の解像度の画像であることを考慮して、パラメータ s を 1/2 に設定した。また、LH, HL, HH バンドについては、周波数パラメータ ξ を低い値に設定した低周波数バンド用パラメータ ($\tilde{\gamma}_L$) と ξ を高い値に設定した高周波数バンド用パラメータ ($\tilde{\gamma}_H$) を用い、式(9)に示すような組合せで、各バンドに適した辞書を作成した。

$$\begin{aligned} D_{LL} &= g_{\tilde{\gamma}_{LL}} \otimes g_{\tilde{\gamma}_{LL}}, & D_{LH} &= g_{\tilde{\gamma}_L} \otimes g_{\tilde{\gamma}_H} \\ D_{HL} &= g_{\tilde{\gamma}_H} \otimes g_{\tilde{\gamma}_L}, & D_{HH} &= g_{\tilde{\gamma}_H} \otimes g_{\tilde{\gamma}_H} \end{aligned} \quad (9)$$

4. シミュレーション実験

本システムの評価のためにシミュレーション実験を行った。

実験には、テスト画像として標準動画画像 (QCIF : Y 信号のみ, 1~150 フレーム, 10Hz) Akiyo, Mother&Daughter(略記:Mother), Carphone, Foreman を用いた。提案方式である MC+SBMP 方式の他に、比較対象として MC+MP 方式 (従来法), 低ビットレート動画画像符号化の国際標準方式である H.263⁹⁾ についてもシミュレーションを行った。提案法におけるサブバンド分割には 8 タップ QMF を用い、H.263 のシミュレーションには “TMNcoder version 3.2” を使用した。

4.1 符号化特性比較

図 3 に、テスト画像における符号化特性の比較結果を示す。提案法と従来法の比較より、どのテスト画像においても、提案法は比較的ビットレートが低い時に従来法とほぼ同様の特性を示し、ビットレートが高い時に特性が改善された。

また、H.263 との比較では、従来法で大きく劣っていた高ビットレート時の特性が改善され、提案法では、H.263 を低ビットレートで上回り、高ビットレートでも近い符号化特性を得ることができた。

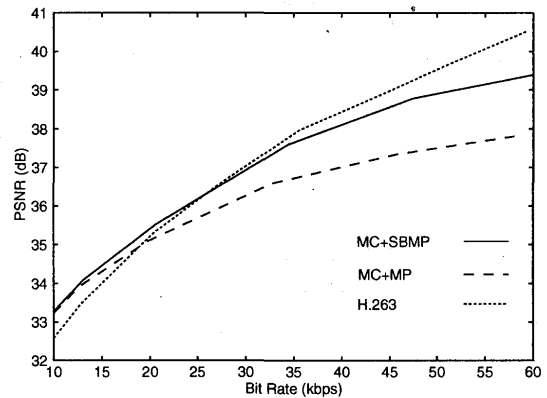
次に、ビットレートに対する特性の違いを解析する。アトム数を 100, 200, 300, 400 に設定した場合の各サブバンドへのアトム割当て比率を表 2 に示す。表よりアトム数が増加すると、LL バンド以外への割当て比率が増加していることがわかる。これは、低ビットレート時に、誤差成分が集中する LL バンドに多くアトム割当てを行い、高ビットレート時に、LH, HL, HH バンドに割当てるといった適切なアトム割当て制御が行われていることをあらわしている。

また、高ビットレート時に特性改善効果が大きいのは、ビットレートが高い場合には、割当てアトム数が多く、LL バンド以外にも多くアトム割当てが行われることで、特徴に合った辞書が利用される割合が多くなり、それによって符号化効率が向上したためである。

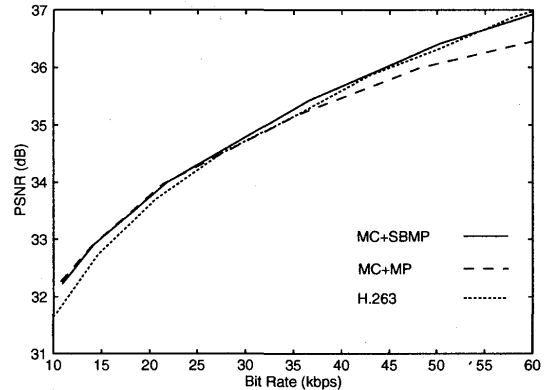
次に、特性改善効果が高いビットレート 60 kbps に設定した場合の、PSNR が高いフレームと低いフレームの各サブバンドへのアトム割当て比率を表 3 に示す。提案法の従来法からの PSNR の改善量という観点で、アトム割当て比率をみると、PSNR の改善効果が高いほど、LL バンド以外に割当てられたアトムの比率が高い。これからも、LL バンド以外へのアトム割当て比率が高いほど、よい特性を示すという結果が得られることが確認できた。

また、表 3 から、各テスト画像における特性の違いをみると、比較的動きが小さい Akiyo において、LL バンド以外へのアトム割当てが多く、改善効果が高いという結果が得られた。これは、動きが小さいため予測誤差が少なく、LL 成分の全体量が少なかったため、結果として LL バンド以外への割当てが増えたものと考えられる。

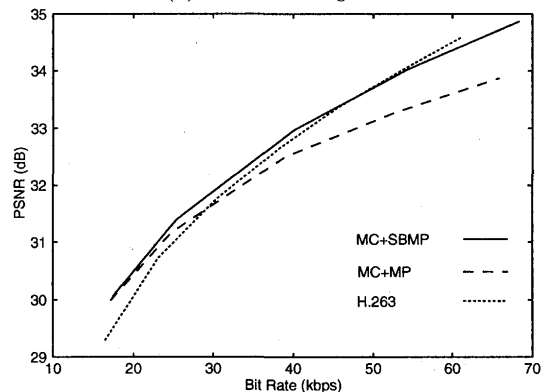
これらの結果より、本提案法である MC+SBMP 方式は、高いビットレート時または、動きが小さい画像において効



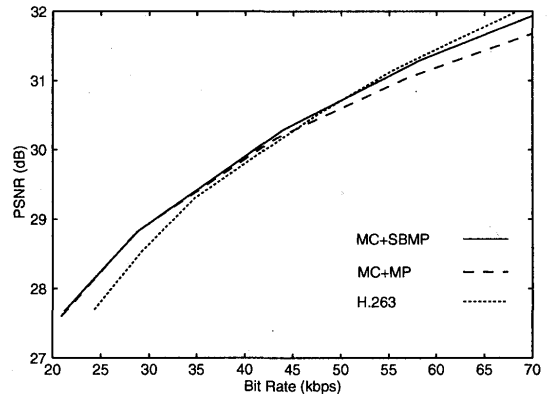
(a) Akiyo



(b) Mother & daughter



(c) Carphone



(b) Foreman

図 3 符号化特性比較

Comparison of performance.

表 2 各バンドに対するアトム割当て比率
Ratios of assigned Atom to each band.

atom/frame	アトム割当て [LL(%):LH(%):HL(%):HH(%)]			
	Akiyo	Mother	Carphone	Foreman
100	61:17:20:2	81: 8:10:1	77:13: 9:1	89: 9:2:0
200	56:18:22:4	71:11:15:3	68:16:13:3	81:13:4:2
300	51:20:23:6	65:13:18:4	62:18:15:5	75:16:6:3
400	50:20:23:7	60:15:19:6	59:19:16:6	71:17:8:4

表 3 PSNR に対するアトム割当て比較 (60kbps)
PSNR to Ratios of assigned Atom (60kbps).

画像名	Frame	PSNR(従来法)	PSNR(提案法)	アトム割当て
Akiyo	73	37.31	39.24	49:20:25:6
	130	38.72	40.82	51:23:19:7
Mother	64	35.84	36.18	75:11:13:1
	43	37.45	38.38	50:12:24:14
Carphone	124	33.04	33.50	70:14:13:3
	148	35.04	36.20	54:27:14:5
Foreman	85	30.00	30.07	81:13:6:0
	109	32.40	32.92	61:24:10:5

表 4 処理の内訳 (Akiyo, 200 atom/frame)
Processing items (Akiyo, 200 atom/frame).

処理		割合 (%)
MC 処理		6.3
MP 処理	内積演算	86.9
	その他	6.0
その他		0.8

表 5 フレーム当たりの符号化実測時間の比較 (200 atom/frame)
Comparison of real processing time per frame (200 atom/frame).

画像	従来法 [sec]	提案法 [sec]
akiyo	59	29
mother	60	31
carphone	59	30
foreman	60	32

果が高いといえる。

4.2 演算量比較

次に提案法の演算量の解析を行う。提案方式の処理は、動き補償における動きベクトル探索処理と Matching Pursuits におけるアトム探索処理が大部分を占める。特にアトム探索処理は、各ブロックについて多くの辞書との内積演算処理が必要となり、多くの演算量が必要となる。

表 4 は、プロファイラ (gprof) により測定した、従来法 (akiyo, 200 atom/frame) の各処理部の演算量の割合を示したものである。表 4 から、アトム探索における内積演算が大きな割合を占めていることがわかる。

提案法では、サブバンド画像の水平および垂直解像度は 1/2 となる。そのため辞書の基底長も 1/2 になるため、内積演算における計算量の削減が見込める。表 5 に、従来法および提案法を計算機で実行した場合の実測時間を示す。両方式におけるアトム数は、200 atom/frame に設定した。測定は

CPU : Pentium III 700MHz
Memory : 128MB
OS : FreeBSD 4.7R
Compiler : gcc-2.95.4

で行い、スタンドアロンで、10 回測定した平均時間を処理時間としている。表 5 より、提案法は従来法と比較して、約 2 倍の実行速度を得ることができた。これは辞書のサイズが小さくなったことにより、MP 処理の大部分を占める内積演算における計算量が削減されたためと考えられる。

5. まとめ

本論文では、サブバンド分割された予測画像に MP 符号化を行う動画像符号化方式の提案を行った。各バンドに対して特徴を考慮した辞書を適用し、符号化を行うことで、符号化特性の改善を行った。

サブバンド分割を行わない従来法と提案法との特性比較により、比較的高いビットレートで符号化効率が改善されることを示した。また演算量についても、提案法が従来法より高速に符号化処理を行うことができることを示した。H.263 との比較では、高いビットレートではほぼ同等の、低いビットレートではこれを上回る符号化特性を得ることができた。

〔文 献〕

- 1) S.Mallat, Z.Zhang: "Matching Pursuits With Time-Frequency Dictionaries", IEEE Trans. on Signal Processing, **76**, 376, pp.3397-3415, (1993)
- 2) R.Neff, A.Zakhor: "Very Low Bit-Rate Video Coding Based on Matching Pursuits", IEEE Trans. on Circuits and Systems for Video Technology, **7**, 1, pp.158-171, (1997)
- 3) R.Banham, C.Brailean, "A Selective Update Approach to Matching pursuits Video Coding", IEEE Trans. on Circuits and Systems for Video Technology, **7**, 1, pp.119-129, (1997)
- 4) 松田一朗, 伊東晋: "MC と MP を統合した動画像符号化", 信学技報, **101**, 301, pp.7-14, (2001)
- 5) 石崎智範, 今村幸祐, 橋本秀雄: "サブバンド分割された MC 誤差画像の Matching Pursuits 符号化", 信学技報, IE2002-260, pp.7-12, (2003)
- 6) ITU-T Study Group 15, "Draft Recommendation H.263", (1995)



いまむら こうすけ
今村 幸祐 1995 年, 長崎大学工学部電気情報工学科卒業。1997 年, 同大学院工学研究科修了。2000 年, 同大学院海洋生産科学研究所修了。同年より, 金沢大学工学部情報システム工学科助手。画像信号の高効率符号化などの研究に従事。工博。



いしざき とものり
石崎 智範 2001 年, 金沢大学工学部電気・情報工学科卒業。2003 年, 同大学院自然科学研究科博士前期課程修了。同年, ソニー株式会社入社。



はしもと ひでお
橋本 秀雄 1968 年, 大阪大学工学部卒業。1970 年, 同大学院修士課程修了。1975 年, 同大学院博士課程修了。同年, 電電公社 (現 NTT) 電気通信研究所入所。以来, 画像符号化方式, 画像通信方式の研究開発に従事。現在, 金沢大学工学部情報システム工学科教授。工博。正会員。