

動画像ビットストリームにおけるサイズ変換処理の検討

笠井 裕之 児玉 明 富永 英義

早稲田大学 理工学部 電子・情報通信学科

〒169-8555 東京都新宿区大久保 3-4-1

TEL: (03) 5286-3385

TAX: (03) 3200-6735

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

高速な動画像表示、部分情報の抽出、複数画像の転送、狭帯域ネットワークへの動画像転送などの要求から、高速かつ簡易な動画像情報サイズ変換処理方式の開発が期待されている。本稿ではMPEG-2ビットストリームを画像情報に復号することなく、高速にサイズ変換することを目的として、MPEG-2ビットストリームにおけるサイズ変換処理方式を提案する。特に、DCT係数領域内ダウンサンプリング手法を利用することによる提案方式実現手法について検討する。その結果、MPEG-2符号化時に4マクロブロックタイプの共通化という制限を与えることにより実現できることを示し、その際の符号化器、及びサイズ変換器を示す。最後に、提案方式の処理量に関して理論的検証を行い、提案方式の有効性を示す。

A Study on Size Transformation Processing Method for Video Bitstream

Hiroyuki KASAI Mei KODAMA Hideyoshi TOMINAGA

Dept. of Electronics, Information and Communication Engineering, WASEDA University

3-4-1 Ohkubo Shinjuku-ku, Tokyo, 169-8555 JAPAN

TEL: (03) 5286-3385

TAX: (03) 3200-6735

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

In this paper, we propose the Size Transformation Processing Method for Video Bitstream. This method has no the DCT/IDCT and the digital down-sampling filter in spatial domain. In particular, this method uses the down-sampling method in DCT domain. At first, we indicate the en-coding codition for the realization of our proposed method. And we indicate the special encoder and size transformer, and state the processing flow. At last, we verified the amount of process, and indicate the usefulness of our proposed method.

1. はじめに

将来のマルチメディア情報サービス及びマルチメディア情報流通を考えた場合、特にマルチメディア情報の簡易な変換処理、編集処理が可能であることが挙げられる。特に高速な動画表示、部分情報の抽出、複数画像の転送、狭帯域ネットワークへの動画画像転送などの要求から、より高速かつ簡易な動画画像情報のサイズ変換処理が期待されている。しかしながら、符号化ビットストリームを画像情報まで復号し、フィルタリング処理を施すことにより実現する手法が、現在中心となっている。そこで、スケーラビリティのように画像情報が階層構造は有しないものの、標準フォーマットに準拠した画像情報作成の際に制限を与えることで、通常提供不可能な変換処理を可能とすることができると考えられる。

本稿では動画標準符号化方式(以下、MPEG-2⁽¹⁾と呼ぶ)を対象として、符号化時に制限を与えることにより、MPEG-2 ビットストリーム上でサイズ変換処理を実現する方式を提案する。本提案サイズ変換処理方式は、MPEG-2 ビットストリームを画像情報に復号することなく、高速にサイズ変換することを目的とする。特に本稿では、DCT 係数領域内ダウンサンプリング手法を利用することによる提案方式実現手法とその有効性について検討する。まず、DCT 係数領域内ダウンサンプリング手法について説明する。次に、サイズ変換処理を実現可能なビットストリームを作成する符号化器について、その制御方式と処理手順を明らかにする。そしてサイズ変換処理方式について述べる。最後に本提案方式の処理量について理論的検証を行ない、本提案手法の有効性を明らかにする。

2. サイズ変換処理方式の提案

本節では、デジタル画像情報の提供サービス形態を考え、従来のサイズ変換処理手法の問題を説明し、ビットストリーム上でのサイズ変換処理手法を提案する。

2.1 提案の概念

デジタル画像情報の提供サービスを考える。ユーザ要求への対応能力、柔軟性を考えた場合、デジタル画像情報をそのままの形態で蓄積・保有することが最適であると考えられるが、情報提供者(Information Provider)側のデータベースメモリの有効利用及びネットワークの有効利用を考慮

すると、蓄積情報形態としてデジタル動画の符号化ビットストリームが有効である。符号化ビットストリームの提供サービスを考えた場合、どのようなユーザ要求に対しても、同一フォーマットの形態でユーザ端末(User Equipment)にストリームを送り、ユーザ端末内において符号化データの復号、デジタルフィルタ処理を施し、ユーザの要求するデジタル画像情報を提供するシステムが考えられる(図1(a))。一方、ネットワークの有効利用及び利用者端末内処理量の軽減を考慮すると、情報提供者側において、ユーザ要求に応じて符号化データを変換し、ユーザ端末へ伝送するシステムが考えられる(図1(b))。本システムは、従来ユーザ端末内で行われていた、復号・デジタルフィルタ処理といった一連の変換処理を情報提供者側が行なうものである。しかしながら、情報提供者側の処理量が増大するといった問題が一方で生じてしまう。そこで、情報提供者側処理量の軽減可能な、符号化ビットストリームレベルでの高速変換処理器による情報提供サービスシステムが要求される(図1(c))。

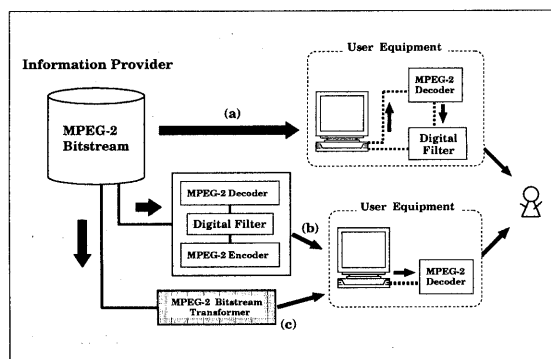


図 1: デジタル動画画像情報提供サービス

2.2 従来方式の課題

特にサービス要求として、画像情報のサイズ変換機能が挙げられる。従来、MPEG-2 符号化情報からサイズの異なる MPEG-2 符号化情報を作成する方法として、空間領域まで復号し、空間領域においてダウンサンプリングする方式(図2)が存在する。図2に示すように、逆量子化された DCT 係数に対して IDCT 変換を行ない、空間領域信号に復号する。復号された信号に対してデジタルフィルタリング処理であるダウンサンプリング処理を行ない、

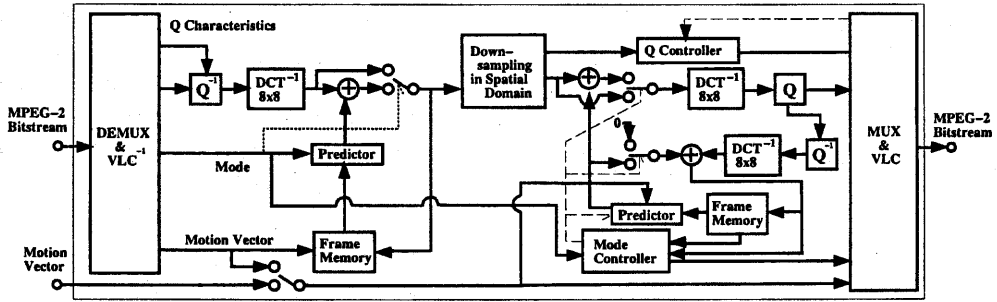


図 2: 動画情報のデジタルダウンサンプリングフィルタによるサイズ変換処理手法

ダウンサンプリングされた信号に対して、改めてDCT変換を施し、量子化、可変長符号化する。本方式の長所は、動き補償を再度行なうことが可能であり、それにより復号時の品質を向上することが可能である点である。しかし、動き補償は非常に処理量が大いため、高速なサイズ変換においては望ましくない。よって、動き補償処理は行ないことを前提とした場合、本方式では動き補償処理に次いで処理量が大いと考えられるDCT/IDCT処理を行なうため、処理量が増大してしまい、高速なトランスコーディングが望めない。よってIDCT/DCT処理を行わず、DCT係数領域内でダウンサンプルすることにより高速化を図ることが可能なサイズ変換処理方式が望まれる。

2.3 ビットストリーム上でのサイズ変換処理方式の提案

そこで、MPEG-2ビットストリームのサイズ変換変換処理方式を提案する。本提案方式は、DCT領域上でのダウンサンプリング手法⁽²⁾を利用することにより、MPEG-2ビットストリームレベルでサイズ変換処理を実現し、且つ、MPEG-2準拠のビットストリーム出力することを目的とする。

提案変換処理実現にあたり、以下の2つのステップで検討する。

- (i) サイズ変換可能なビットストリームを出力するMPEG-2符号化制御方式
- (ii) サイズ変換処理方式

以下では、最初にDCT係数領域内ダウンサンプリング手法について説明し、上記(i)(ii)の順で検討する。

3. DCT 係数領域内ダウンサンプリング手法

本節では、DCT係数領域におけるダウンサンプリング手法について説明する⁽²⁾。

空間領域における $N \times N$ ブロックの2次元データを $[g]$ 、そのDCT成分を $[G]$ とすると、DCT変換及び逆DCT変換はそれぞれ $[G] = [T] [g]$
 $[T]^{-1}$ 、 $[g] = [T]^{-1} [G] [T]$ で表される。ここで、 $[T]$ は、 $N \times N$ DCT変換マトリクスであり、 $[T]^{-1} = [T]^T$ である。

ここで、DCT係数領域におけるダウンサンプリングを、横方向の2:1ダウンサンプリングについて考える。ただしダウンサンプリングされたブロックを $[g^{(n)}]$ と表すこととする。

空間領域において横方向に隣接した2つの $N \times N$ ブロックを $\left[\begin{matrix} [g^{(n)}] & [g^{(n+1)}] \end{matrix} \right]$ とすると、横方向にダウンサンプリングされた $N \times N$ のブロック $[g^{(n,n+1)}]$ は、式(1)で表される。

$$\begin{aligned} [g^{(n,n+1)}] &= \left[\begin{matrix} [g^{(n)}] & [g^{(n+1)}] \end{matrix} \right] \cdot [S] \\ &= \left[\begin{matrix} [g^{(n)}] & [g^{(n+1)}] \end{matrix} \right] [S] \end{aligned} \quad (1)$$

ここで $[S]$ は $2N \times N$ のダウンサンプリングマトリクスである。上式より、空間領域において横方向にダウンサンプリングされた $[g^{(n,n+1)}]$ のDCT変換は式(2)で表現される。

$$\begin{aligned} [G^{(n,n+1)}] &= [T] \left[\begin{matrix} [g^{(n)}] & [g^{(n+1)}] \end{matrix} \right] \times [S] [T]^{-1} \\ &= [T] \left[\begin{matrix} [g^{(n)}] & [g^{(n+1)}] \end{matrix} \right] [\tilde{T}_2]^{-1} \\ &\quad \times [\tilde{T}_2] [S] [T]^{-1} \\ &= \left[\begin{matrix} [G^{(n)}] & [G^{(n+1)}] \end{matrix} \right] \times [\tilde{T}_2] [S] [T]^{-1} \end{aligned} \quad (2)$$

ここで、 $[\tilde{T}_2]$ は、式 (3) で示される $2N \times 2N$ ブロック DCT マトリクスである。

$$[\tilde{T}_2] = \begin{bmatrix} [T] & [0] \\ [0] & [T] \end{bmatrix} \quad (3)$$

よって式 (2) は最終的に式 (4) で表現される。

$$[G^{(n,n+1)}]_s = \left[[G^{(n)}] \mid [G^{(n+1)}] \right] [F] \quad (4)$$

ここで、 $[F] = [\tilde{T}_2] [S] [T]^{-1}$ であり、 $[F]$ は、 $2N \times N$ の行列である。結果として、式 (4) は、DCT 係数領域において横方向に隣接した 2 つの $N \times N$ ブロック $\left[[G^{(n)}] \mid [G^{(n+1)}] \right]$ に、行列 $[F]$ を掛けることにより、横方向にダウンサンプリングされた $N \times N$ の DCT ブロックが得られることを示している。また縦方向のダウンサンプリングに関しても同様に考えることができる。

4. サイズ変換処理を実現する符号化制御方式

本節では、3. を利用したサイズ変換処理を実現するための、ビットストリーム符号化器について述べる。本符号化器は、MPEG-2 MP@ML ビットストリームを出力するが、提案方式実現のための制御条件に基づいて符号化処理を行う。以下では、MPEG-2 ビットストリーム中に記述される符号化パラメータに着目し、制御条件について考察する。

4.1 MPEG-2 符号化パラメータ

提案方式は、DCT 係数上でダウンサンプリングを行うため、ダウンサンプリング後のデータ単位が、MPEG-2 で符号化処理可能な最小データ単位である必要がある。通常 MPEG-2 の符号化処理はマクロブロック ($16[\text{pel}] \times 16[\text{line}]$) 単位での処理を行なうため、本稿では、サンプリング処理対象となるデータ単位を 4 つのマクロブロックとして考えていく。

- (1) 上述したように、4 つのマクロブロックを単位として処理するので、4 つマクロブロックのマクロブロックタイプが同一である必要がある。つまり、マクロブロックの動き補償予測モード (順 / 逆 / 双方向予測)、MC/Non MC などのタイプ情報が共通である必要がある。本条件は、提案変換処理を実現するための必須条件であるといえる。

- (2) マクロブロックタイプと共通化と同様に、4 マクロブロックの DCT タイプを共通化する必要がある。MPEG-2 のフレーム構造では、フレーム DCT、フィールド DCT の 2 種類の DCT 符号化モードを、マクロブロック単位に切替えることが可能である。DCT 符号化モードの共通化も必須条件である。
- (3) MPEG-2 では、マクロブロック単位で量子化値を決定することが可能である。本提案方式では、量子化値の共通化は必須条件ではないと考えられるが、ダウンサンプル対象となる各 4 マクロブロックの品質が極端に異なる場合には、誤差が他のマクロブロックに拡散するため、量子化値の共通化は望ましい。さらにサイズ変換処理器で、符号化器で用いた量子化値を用いた場合、復号器では、符号化前処理と同等の復号処理を実現することが可能となり、画質劣化を軽減することが可能である。
- (4) サイズ変換処理時に符号化する動きベクトルは、各 4 つのマクロブロック情報として復号されたベクトルを、平均化したものを用いることが可能であるので、共通化の必要はないといえる。一方、変換処理された画像情報の復号を考えた場合、最も画質劣化を引き起こす要因となるものが、予測参照画像の不一致によるドリフト誤差の発生であると考えられる。そこで、符号化器では、ダウンサンプルした画像を用いて動き推定をし、そこで求めた動きベクトルの 2 倍のベクトルを 4 マクロブロックで共通に使用する。そしてサイズ変換処理器では、そのベクトルを半分にしたものをストリーム化することにより、符号化側で用いた動きベクトルを、復号器で用いることが可能である。

4.2 提案符号化器

以上の条件を基にした符号化器を図 3 に示す。

図 3 に示すように、ダブルループの符号化器であり、MPEG-2 符号化のための前処理部と MPEG-2 符号化部から構成される。以下、処理過程について述べる。

まず入力された画像情報は、空間領域でダウンサンプリングされ 1/4 サイズの画像に変換される。ダウンサンプリングされた画像は、仮想的な MPEG-2 符号化処理が施される。ここでマクロブロックの動きベクトル、動き補償予測方向 (順 / 逆 / 両方向)、MC/NoMC などのマクロブロックタイプが

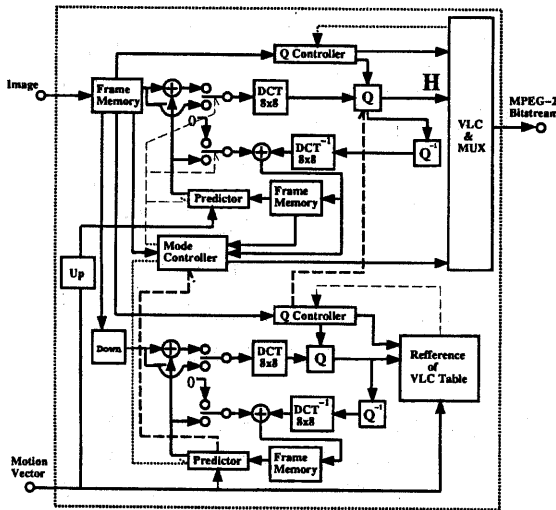


図 3: 提案符号化器

決定される。また量子化係数値も決定される。ここで、量子化された DCT 係数値、動きベクトルなどの情報は VLC テーブルを参照し、発生符号量を仮想的に算出する。ここで発生する情報量を仮想的に算出し、量子化係数値にフィードバックする。よって、本処理で得られる各情報はビットストリーム化されることはない。つまり MPEG-2 符号化処理の前処理をここで行う。

次に、上記処理より得られた各情報を用いて、入力されたフルサイズの画像情報の MPEG-2 符号化処理を行う。量子化は前処理で得られた量子化値を用いることとする。動きベクトル情報は、前処理の動き推定で得られた情報を単純に 2 倍したものを用いる。マクロブロックタイプ情報を含めたマクロブロック階層情報は、前処理で得られているデータを用いる。

本符号化器で用いる制御方式は、サイズ変換処理で得られる画像情報の品質を重視したものであり、特に、4 マクロブロックで共通の動きベクトルを用いることにより、フルサイズでの品質は劣化するものの、変換後の画像情報を復号する際、参照画像の不一致によるドリフト誤差を軽減する方式であるといえる。

5. サイズ変換処理方式

4. で述べた符号化器から出力される MPEG-2 ビッ

トストリームを、DCT 領域上ダウンサンプリング手法により、サイズ変換することが可能なトランスコーディング方式について述べる。

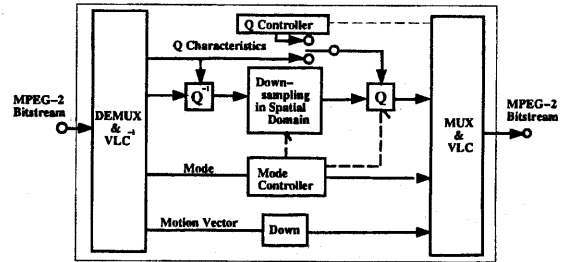


図 4: サイズ変換処理器

図 4 に、提案処理器を示す。以下、処理手順について述べる。

- (1) MPEG-2 ビットストリームを可変長復号する。
- (2) 可変長復号データの内、ピクチャデータ層以上のヘッダ情報に関しては、可変長符号化器へと出力される。ただし、サイズ変換に伴うパラメータの変更として、画像サイズを示すコードの修正が必要である。
- (3) マクロブロック内の量子化係数を逆量子化し、DCT 係数を取得する。
- (4) 4 マクロブロックを処理単位として、得られた DCT 係数をダウンサンプリングし、1 マクロブロックの DCT 係数を得る。
- (5) ダウンサンプリングされた DCT 係数を量子化し、可変長符号化する。この量子化処理は、復号された量子化係数値を用いる場合と、新たに符号量制御を行う場合が考えられる。
- (6) また各マクロブロックで復号された動きベクトル情報は、1/2 サイズにし、可変長符号化する。

6. 提案変換処理方式の理論的検証

本節では、提案トランスコーディング方式の品質及び処理量に関して考察する。ここでは、空間領域まで復号してダウンサンプリングする手法と比較検討する。ここで、提案方式を方式 (a)、空間領域でダウンサンプリングする方式を方式 (b) と定義する。

6.1 DCT 係数領域内ダウンサンプリング手法の評価

DCT 係数信号 $[G]$ を逆 DCT して空間領域の信号 $[g]$ に戻し、空間領域においてダウンサンプルし、さらに DCT により周波数領域に戻す過程を考える。

$$\begin{aligned}
 & \left[\begin{matrix} [g^{(n)}] & [g^{(n+1)}] \end{matrix} \right] \\
 &= [T]^{-1} \left[\begin{matrix} [G^{(n)}] & [G^{(n+1)}] \end{matrix} \right] [\tilde{T}_2] \\
 & [g^{(n,n+1)}] \\
 &= [T]^{-1} \left[\begin{matrix} [G^{(n)}] & [G^{(n+1)}] \end{matrix} \right] [\tilde{T}_2] [S] \\
 & [G^{(n,n+1)}] \\
 &= [T] [g^{(n,n+1)}] [T]^{-1} \\
 &= [T] [T]^{-1} \left[\begin{matrix} [G^{(n)}] & [G^{(n+1)}] \end{matrix} \right] \\
 & \quad \times [\tilde{T}_2] [S] [T]^{-1} \\
 &= \left[\begin{matrix} [G^{(n)}] & [G^{(n+1)}] \end{matrix} \right] [\tilde{T}_2] [S] [T]^{-1} \\
 &= \left[\begin{matrix} [G^{(n)}] & [G^{(n+1)}] \end{matrix} \right] [F]
 \end{aligned}$$

以上より、DCT 係数領域におけるダウンサンプリングと最終的には同様の式となることが分かる。よって方式(a)と方式(b)は、理論的に等価な変換であると言える。

6.2 処理量の評価

方式(a)と方式(b)の処理量の比較検討を行なう。両方式を比較するに際して、両方式の異なる処理器に着目する。図5、図6に、(a)(b)方式の異なる処理ブロック図を示す。

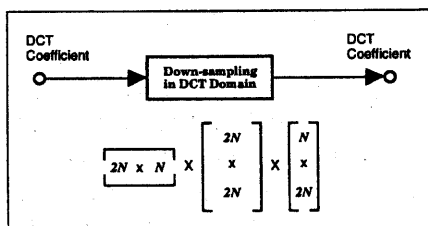


図5: 方式(a)の処理量

図5、図6を用いて行列演算処理量について考察する。ただし、 $(N \times N)$ 行列 $\times (N \times N)$ 行列の行列演算処理量を1とする。まず方式(a)では、DCT 係数領域内ダウンサンプリング処理として、 $(2N \times 2N) \times (N \times 2N)$ が演算処理量4、 $(2N \times N) \times (N \times 2N)$ が演算処理量2であり、合計処理量は6である。一方、方式(b)では、IDCT 演算 $(N \times N) \times (N \times N) \times (N \times N) \times 4$ (ブロック) が

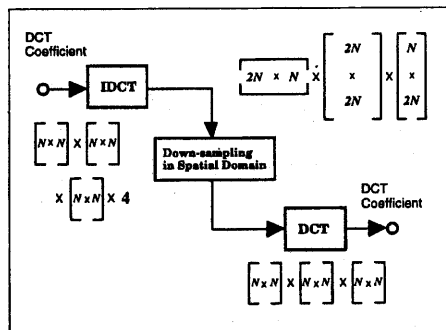


図6: 方式(b)の処理量

処理量8、空間領域内ダウンサンプリング処理として、DCT 係数領域内ダウンサンプリング処理と同様6、またDCT 演算 $(N \times N) \times (N \times N) \times (N \times N)$ が処理量2であり、合計処理量は16である。

以上より、方式(b)は方式(a)と比較して3倍近くの処理量が必要であることが分かり、処理量の観点から方式(a)が有効であるといえる。

7. まとめ

本稿では MPEG-2 ビットストリームを画像情報に復号することなく、高速にサイズ変換することを目的として、MPEG-2 ビットストリームにおけるサイズ変換処理方式を提案した。特に本稿では、DCT 係数領域内ダウンサンプリング手法を利用することによる提案方式実現手法について検討した。その結果、MPEG-2 符号化時に4マクロブロックタイプの共通化という制限を与えることにより実現できることを示した。またその際の符号化器、及びサイズ変換器を示した。最後に、提案方式の処理量に関して、理論的検証を行い、提案方式の有効性を示した。今後は、最適な動きベクトル選択手法、及び符号化制御方式について検討していく予定である。

参考文献

- (1) ISO-IEC 13818-2, IS: "Generic Coding of Moving Picture and Associated Audio Recommendation H.262" (1995).
- (2) A.Neri, G.Russo and P.Talone: "Inter-block filtering and downsampling in DCT domain", Signal Processing Image Communication, pp.303-317 (1994).