

画像圧縮処理の中身と画質への影響

—— 開発システムに合った画像エンコーダ選択の手引き

清 恭二郎

デジタル画像を圧縮・伝送・復号化するシステムにおいて、画像の質を決定づけるのは、テレビに近い受信側ではなく、むしろ画像を圧縮するエンコーダ側にあると言われる。ここでは、その理由と、画像エンコーダの選択のヒントを述べる。

(編集部)

例えばビデオ・カメラを想定した場合、イメージ・センサからの出力は、大きく以下の三つのモジュールを経由してビット・ストリームになります。

- A-Dコンバータなどのアナログ部
- 補間処理、色変換、ガンマ補正などを行うデジタル・フロントエンド
- MPEG, H.264 など、圧縮のためのデジタル・バックエンド

ここではおもに画像の圧縮処理に注目し、「アルゴリズムに起因する画質^{注1}」を決定づけるカギは何であるのか、また、それを最適化するためにはどういふふうが必要なのか、を掘り下げていきます。

1. 動き予測と符号量制御が圧縮処理の肝

図1はMPEGなどの圧縮処理のブロック構成を簡単に示したものです。圧縮処理は、基本的にDCT(離散コサイン変換)やDHT(ハフマン・テーブル定義)といった直交変換を利用した「画面内テクスチャ符号化」と、前後のフレーム

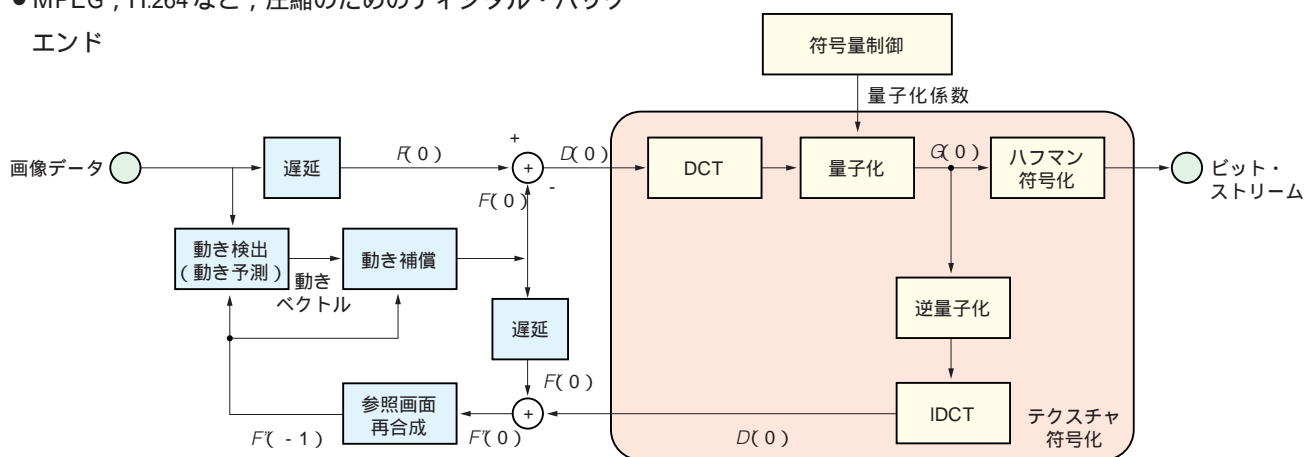


図1 圧縮処理のブロック構成

画質を左右する重要なモジュールは、「符号量制御」と「動き予測」の二つだけ

注1：ここでいう画質とは、「圧縮処理を施していない画像」と「符号化し、さらに復号したあとの画像」の見た目の違いを表したもので、圧縮処理を施していない画像に近いほど、質が高いとする。

Keyword 動き予測, 符号量制御, 画面内テクスチャ符号化, フレーム・ベース, マクロブロック・ベース, 1pass, 2pass, CBR, VBR, GOP, 探索アルゴリズム, 探索範囲, ダイヤモンド・サーチ, 3ステップ・サーチ, サーチ・レンジ

から動きを検出・補償したうえでテキスト符号化を行う「フレーム間符号化」から構成されています。

図1において、入力された画像はマクロブロックと呼ばれる16×16画素を一つの単位として、あらかじめ蓄積されていた前後のフレームからの動きを検出し、その結果に基づいて動き補償を行います。その動き補償後のデータと元画像との差信号をテキスト符号化するのが、フレーム間符号化と呼ばれる処理です。

ここで、画質という観点からこの処理を見た場合、画質を左右する重要なモジュールは、「符号量制御」と「動き予測」の二つだけといってよいでしょう。そのほかの構成要素は基本的に定型的な算術演算処理を行っており、画質の差は発生しません。そのため、各チップ・メーカーやIPコア・ベンダによるこれらの基本性能の差は、処理速度、またはチップ面積からくるコストの差と考えられます。

一方、符号量制御と動き予測が画質に与える影響には多岐にわたるものがあり、MPEG 開発技術者の腕の見せどころでもあります。A社のチップは低ビット・レートでは高画質だが、高ビット・レートではB社のほうが画質が良い、あるいはC社のIPコアは動きが早い画像でも破たんしないがゆっくりした動きの画像ではD社のほうが良さそうだ、などさまざまです。そこで、この符号量制御と動き予測の2点について、さらに詳しく説明します。

● 符号量制御のアルゴリズムは大きく三つ

符号量制御とは、直交変換したデータを量子化する際に係数を決定するための制御処理で、画像の複雑さやすでに生成された符号量を考慮して最適化していきます。量子化係数が大きくなれば生成符号量は小さくなり、ビット・レートを下げられますが、再生する際に元画像との差が大きくなり、モスキート・ノイズやブロック・ノイズと言われるデジタル・ノイズが増えてしまいます。量子化係数を小さくすれば、それだけ損失がなくなるので画質は向上しますが、生成符号量は増え、ビット・レートが上がります。符号量制御のアルゴリズムの基本は、次の3種類に分けて考えられます。

- 1) フレーム・ベース/マクロブロック・ベース
- 2) 1pass/2pass
- 3) CBR/VBR

次に詳しく説明します。

1) フレーム・ベース/マクロブロック・ベース

MPEG や H.264 では、マクロブロックと呼ばれる16×16画素を基本単位として処理を行います。符号量は、マクロブロック周期、フレーム周期、GOP (group of picture) 周期で制御します。

このGOPは、十数フレームから数十フレームを一単位とした画像で、GOPの中にIフレームが一つ存在し、残りをPフレームやBフレームで構成しています。Iフレームは、PフレームやBフレームよりもたくさんの符号が必要なので、GOP周期で符号量を平滑化する制御を行います。そして、フレーム単位でI、P、Bフレームの種類に応じて符号量を割り当てます。

ここでは符号量制御ループをマクロブロック単位で行うことを、マクロブロック・ベースと呼びます。マクロブロック・ベースの場合、あるマクロブロックの動きベクトルが検出されると、直ちにそのマクロブロックの動き補償を行い、テキスト符号化を行います。したがって、そのピクチャ全体を把握することなく符号化していくことになります。いわば、「木を見て森を見ない」的な手法です。

一方、フレーム・ベースでは、マクロブロック単位で動き予測を行っても、そのフレーム処理が終了するまではテキスト符号化を行いません。一般的に動き予測回路では、予測を検出していく過程でそのフレームの複雑さを解析していくことができます。あまり変化のない単調な画像か、あるいは非常に細かいテキストチャが含まれた画像か、もしくは画面の上半分は簡単だが下半分はかなり複雑である、といった特徴を解析できます。そのため、そのフレームの予測処理が終了したとき、画像の内容を理解したうえで、その時点までに生成した符号量も考慮して最適な符号配分を決めることができます。これがフレーム・ベースです。「森」全体を見ている手法です。画質を論じる場合は、フレーム・ベースのほうがはるかに良い手法と言えるでしょう。

マクロブロック・ベースではメモリ使用量や回路など、ハードウェア資源を最小限にできることがメリットと言えます。MPEGのIPベンダでは、このタイプのアルゴリズムを用いている例がいくつか見られます。超低ビット・レートの携帯電話などでは有効でしょう。

2) 1pass/2pass

1pass/2passは、フレーム・ベースをさらに進めたものと言えます。2passとは圧縮を2回に分けて行うことを意味し、1回目で圧縮したい画像全体の解析を行い、2回目