

高階調信号に対する映像符号化方式の基礎検討 A study on the image coding method to the high gradation signal

山田 悦久
Yoshihisa Yamada

井須 芳美
Yoshimi Isu

1. はじめに

近年、デジタルカメラや液晶・プラズマなどの映像入出力デバイスの高機能化・高品質化が急速に進んでいる。高機能化の一つとして、色再現性の高品質化、すなわち画素あたりの高階調化が挙げられる。従来、デジタル映像信号は1色信号あたり8ビットで表現され、YCbCrやRGBなどの3つの色空間で表現されることにより1画素あたり24ビット、1678万色で表されることが多かった。放送局内などの素材レベルでの編集作業においては、業務用のVCRをはじめとしてこれまでも1色信号を10ビットで扱う機材もすでにあったが、家庭内に10ビットのAV機器が普及してきたのはこの数年のことである。

このように画素階調が入出力デバイスに対して高められつつある状況に対して、映像符号化の分野においては、JPEG・MPEG等のいずれの国際標準化においても、8ビットで検討・評価された符号化ツール・符号化方式をベースにして、10ビット・12ビット等に拡張しているに過ぎず、映像信号の特性そのものに立ち入った評価は十分に行われていない。

本研究では、高階調映像信号の符号化方式を検討することを最終の目的として、静止画像信号を対象として下位ビットの信号そのものの特性等について評価・検討した結果を報告する。

2. 画像信号の下位ビット信号の特性

評価用画像として、映像情報メディア学会のハイビジョン用デジタル標準画像データ(1920画素×1080ライン、GBRおよびYPbPr色空間の10ビットの静止画)[1]に収録されているものから「輝度信号(Y)」を使用した。本報告では、緩やかなグラデーションを含む画像で特に下位ビットの影響が顕著に見られる画像「エッフェル塔」を用いる。



原画像 (エッフェル塔)

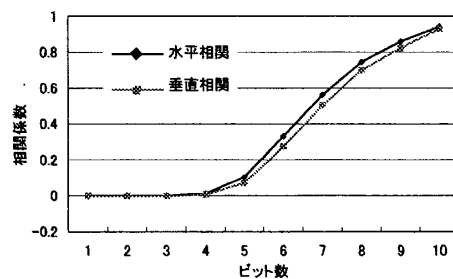
2.1 下位ビット信号の相関

10ビットの画像信号に対して、下位側から順に1~10ビットの信号を取り出し、隣接画素間(水平方向・垂直方向それぞれ)の相互相関を以下の計算式により測定した。

$$Corr_{Horizontal} = \frac{\sum((pel_{(x,y)} - m) \times (pel_{(x+1,y)} - m))}{\sum(pel_{(x,y)} - m)^2}$$

m は画面全体における各画素の下位ビットの平均値である。

2.1.1 画面全体に対する結果と考察



結果を以下に示す。

下位の1~4ビット目までは非常に相関が低く、雑音に近い信号であり、5ビット目からビットが増えるごとに徐々に相関が高くなっていることが読み取られる。

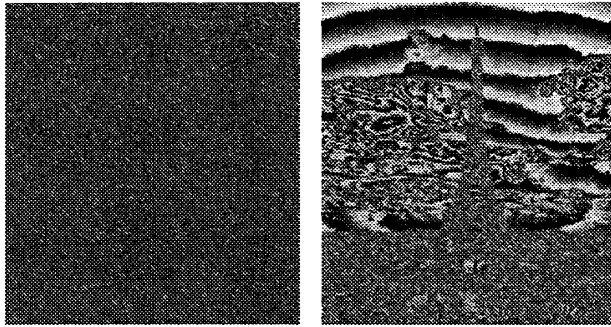
実際にこの結果を上位ビット側から考察すると、上位6ビットあたりまでは画素間の相関が強く、被写体に影響される何らかのテクスチャ情報を画像信号自体が含んでおえり、一方上位から見て7ビット目あたりから下位のビットには情報は少なくなって雑音信号に近づいている、と考えることができる。これは、上位6ビットまでの画像では、青空の部分に擬似輪郭が容易に認められるが、7ビット以上の画像ではその影響を認めにくい(例えばモニタの設定条件や周辺環境をある程度調整する必要がある)、という主観評価の結果とも親和性が高い。

2.1.2 領域単位に対する考察

次に画像内の領域単位に主観により評価を行う。下位ビットの信号だけを取り出して評価を行うと、「エッフェル塔」の画像の場合、青空や雲の部分については下位3ビットまでの信号では擬似輪郭に相当する縞模様の構造が認められないが、4ビット目から縞模様が認められるため、上位6ビットの信号だけでは不足していることが読み取れる。一方、中央の塔の領域や左右下部にある木々の領域については、下位6ビットの信号でも雑音同様の信号状態であり、下位7ビット目でようやく雑音とは異なる構造が読み取れる。すなわち、テクスチャの細かな領域では、上位4ビッ

ト程度もあれば十分であり、平坦な領域（およびその境界）については6ビット以上必要である、ということが、画像「エッフェル塔」については読み取れる。

以上の結果より、今回評価に使用した10ビット信号の場合、下位1～3ビットについては画面全体に対して雑音に近く、4～6ビットについては平坦な領域でのみ相関の高い信号である、ということがわかる。



左：下位3ビット信号、右：下位6ビット信号
(画面の中央部分を切り出して強調したもの)

2.2 JPEG 符号化における下位ビットの効果

下位ビットが意味を持つ符号化処理はどのようなときに可能であるかを評価するために、DCTを基底とするJPEG符号化を用いて実験を行った。本稿では、JPEG符号化としてBaseline processを用いることとし、10ビットの原画像を一旦8ビットに丸めた上で実験・評価を行った。なお、量子化テーブルは、JPEG規格のAnnex Kに記されているものではなく、周波数に重み付けを行わない平坦なものを使用した。

2.2.1 画面全体に対する結果と考察

原画像と符号化画像との間で、誤差のないもの、誤差が±1の範囲に収まるもの（8ビット目にはのみ誤差が含まれるもの）、誤差が±3の範囲に収まるもの（7ビット目には誤差が含まれるもの）、という分類で、画素単位に百分率を測定した。

圧縮率 (bit/pel)	PSNR	誤差		
		0	±1以内	±3以内
4.24	58.91	91.646	100.000	100.000
3.60	49.67	47.898	94.275	99.995
2.32	46.75	35.393	81.598	99.695
1.28	43.36	27.537	66.958	94.884
0.80	39.31	20.615	55.096	84.465
0.48	34.95	13.208	42.835	73.790

圧縮率が1/10を超えるJPEG画像では、半数の画素信号で±1を超える符号化誤差を含んでしまうことから、9ビット以上の画像信号を表現できる可能性はかなり低くなることが想定される。圧縮率が1/2～1/3程度のJPEG画像では、80%以上の画素が符号化誤差を±1以内でしか含んでいないため、より高階調な復号画像を提供できる可能性がある。

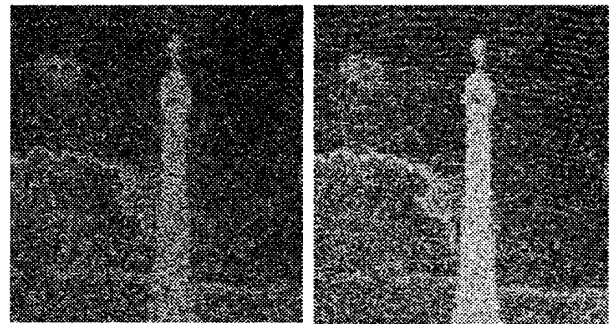
2.2.2 領域単位に対する考察

JPEG符号化画像、および原画と符号化画像との差分画像を主観により観察した結果を以下に記す。

青空の部分について観察すると、画面全体に対する2.2.1節の評価結果と同様、圧縮率が2.32bit/pelの場合（DCT係数の量子化処理としては、12ビットで得られる係数を10ビットで量子化した場合に相当する）には、符号化画像・差分画像ともに特異なテクスチャ構造は見られないが、圧縮率が1.28bit/pelの場合（DCT係数を9ビットに量子化した場合）には疑似輪郭が認められた。

一方、中央の塔の領域や左右下部にある木々の領域については圧縮率0.48bit/pelの場合（DCT係数を7ビットに量子化した場合）でも、符号化画像・差分画像ともに特異な構造は認められない。

2.1.1節での考察結果と同様、青空と塔・木々の領域では空間領域（画素レベル）・周波数領域（DCT係数）のいずれにおいても3ビット程度の差があることがわかる。



左：圧縮率2.32bit/pel、右：圧縮率1.28bit/pel
(画面の上中央部分を切り出して強調したもの)

3. まとめ

高階調映像信号の符号化方式を検討することを最終の目的として、HDTV静止画像信号を対象として上位ビット・下位ビットの信号特性に対する評価を行った。

今回の実験結果や考察については、入力デバイスや被写体・撮影条件等によって大きく変わってくることは容易に類推できるが、大まかな傾向としては今回考察した結果と同様であろうと推測できる。

今後は、今回の輝度信号に対する評価に加えて色信号に対する評価を行うとともに、評価結果から得られた信号特性を利用することにより、本来の目的である高階調映像信号に対する符号化方式のアルゴリズムの考案を進めていく。

参考文献

[1]映像情報メディア学会「ハイビジョン用デジタル標準画像データ」http://www.ite.or.jp/shuppan/testchart_index.html