

J-036

JPEG2000 におけるテキスト分離符号化方式に関する検討

A Study on Text-separation Coding for JPEG2000

石川 孝明[†]
Takaaki ISHIKAWA

富永 英義[‡]
Tominaga HIDEYOSHI

渡辺 裕[†]
Hiroshi WATANABE

1. まえがき

JPEG2000 で用いられている EBCOT は、算術符号化をベースとした 1 パス符号量制御が可能なアルゴリズムであり、ロシー/ロスレス圧縮を同一のアルゴリズムで扱うことが可能である。しかし、画像から高周波成分を除去することで高圧縮を実現している点では、JPEG などと同じである。このような圧縮手法は、自然画像のような静止画像には非常に適した圧縮手法であると言える。しかし、画像中にテキストのような周波数特性の異なる領域が含まれている場合には、必ずしも適切な圧縮手法であるとは言えない。この考え方に基づいた標準として、MRC(MixedRasterContent)がある [1]。

JPEG2000 の標準化過程においても、MRC の概念が継承されており、Part6 に定義されている [2]。しかし、いずれの標準もテキスト部分の再圧縮や解像度変換による画質劣化については特に明記されていない。

そこで本稿では、画像中のテキスト領域を一度画像から分離し、画像のみを再圧縮もしくは解像度変換したのち、テキストと画像を多重化し、最終的に JPEG2000 として復号可能なビットストリームを構成するための基礎検討を行う。

2. テキスト領域の画質低下

JPEG を含め、自然画像に対するデータ圧縮を行う場合、人間の視覚特性を考慮し、画像の高周波成分を除去する手法が用いられる。具体的には、JPEG では高周波成分である AC 成分を DC 成分よりも優先的に量子化し、JPEG2000 においても HH,HL,LH 成分の情報を捨てることにより高い圧縮率を実現している [3]。しかし、現状の圧縮手法では、オブジェクトのエッジ周辺にモスキートノイズが発生し、これが画質低下の要因となっている。

2.1 予備実験

テキスト中心の画像とグラフィクス中心の画像に対し、JPEG 符号化および JPEG2000 符号化を施し、高周波成分を除去する圧縮手法がどれほどテキスト領域の画質低下の原因となるかについて、視覚的に理解するための予備実験を行った。実験では、テキストおよびグラフィクス画像を、JPEG、JPEG2000 符号化方式で圧縮し、符号量と PSNR を比較し符号化ひずみを検出した。

実験に使用した画像およびプログラムを表 1 に示す。

2.2 考察

図 1 にグラフィクス画像の各符号化方式による符号化ひずみを、図 2 にテキスト画像における符号化ひずみを示した。同図より、テキストとグラフィクスの例外なく、JPEG と比較し JPEG2000 の符号化特性が優れていることがわかる。また、図 3 より、グラフィクスの画質低下は、テキスト領域のひずみにくらべ緩やかであるのに対し、テキスト領域のひずみの勾配はグラフィクス

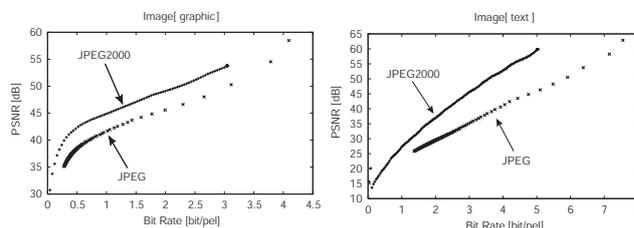


図 1: The rate distortion of graphic
図 2: The rate distortion of text

のそれよりも急であることがわかる。したがって、同じ符号量で各画像を圧縮した場合、テキスト領域の PSNR は相対的にも低下することがわかる。

図 3 中の点 A,B,C,D における復号画像を図 4 に示した。同図より、視覚的にも圧縮によりテキスト領域のひずみが強く現れていることがわかる。これは、高周波成分を多く含むテキスト領域がテクスチャに近い特性を持っていることに加え、このような領域が、Wavelet 係数 LSB 側の数値で表現され、ポスト量子化において容易に除去されてしまうためである [4]。

ここで、図 3 中のグラフィックのひずみとテキストのひずみを表すグラフの間の領域を S_{g-t} とするならば、 $S_{g-t} = 0$ となるような符号化を行うことで、テキスト領域の画質低下を視覚的に低減することが出来ると考えられる。本稿では、このような符号化を行う手法として JPEG2000 におけるテキスト分離符号化を提案する。

3. 性質の異なる領域を含む画像の符号化

前節で示した問題点に対する解決手法の 1 つとして、WITCH がある [4]。WITCH(Wavelet-based Image/Texture Coding Hybrid) は、圧縮により高周波成分が失われた画像に対し、それを補間する手法である。同手法は、次のようなシステムである。(1) 失われた情報をテクスチャ情報のモデル化を行い、パラメータで表現する。(2) 復号後に画像を補間するために、パラメータを付加情報としてヘッダに埋め込む。(3) 専用の Decoder を利用して復号することで失われた高周波成分を補い、テクスチャ領域の画質を向上させる。また、同手法は、JPEG2000Part1 準拠でありヘッダ情報を利用せずとも最低限の復号が可能である。このように、Wavelet 係数

表 1: 使用した画像およびプログラム

グラフィクス画像	JIS X 9201「らん」1280 × 1024 Y 成分
テキスト画像	フォント [Tahoma] 1280 × 1024 Y 成分
JPEG2000 圧縮	KaKaDu 4.0.2
JPEG 圧縮	IJG jpeg software 6b

[†]早稲田大学大学院 国際情報通信研究科 渡辺研究室

[‡]早稲田大学大学院 国際情報通信研究科 富永研究室

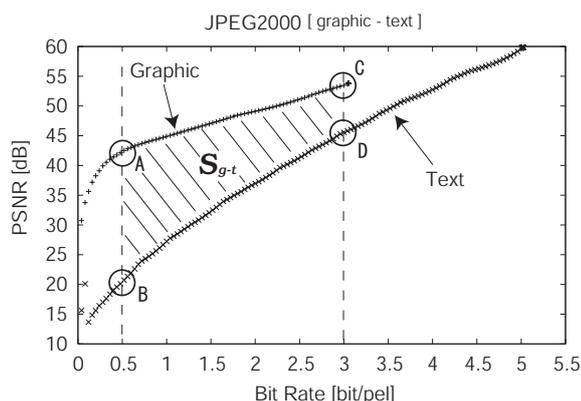


図 3: The rate distortion of Graphic and Text

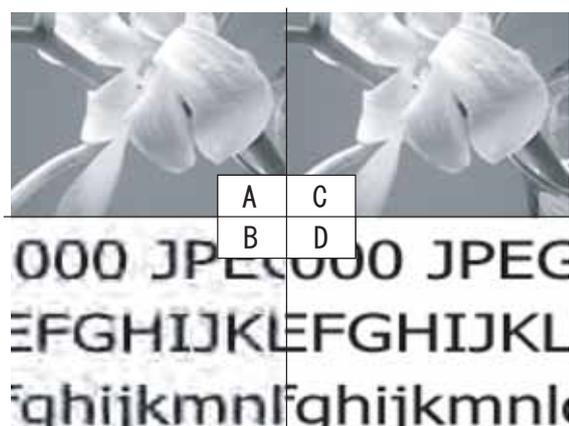


図 4: Graphic/Text Rate Distortion

の特性を利用した画質改善手法がある。

また、従来の適応圧縮手法もテキスト領域のような性質の異なる領域に対する画質向上を実現する一つの手法と考えることが出来る。適応圧縮は、JPEG2000 符号化においても有効である [5]。

4. テキスト分離符号化

静止画像に含まれる特性の異なる領域を考慮した符号化方法として、前節にあげたような手法がこれまで考えられてきた。しかし、テキスト領域の画質向上とコンテンツの2次利用性向上、この2つの要求条件を同時に実現する手法としては、あらかじめテキストを分離し符号化する手法が最も有効な符号化手段の1つであると考えられる。

分離後のテキストは、JPEG2000 Part2 および Part6 にて定義されている不透明レイヤを用いて合成することができる。

4.1 JPEG2000 Part2 [6]

JPEG2000 Part2 は、すべての Part のベースとなる Part1 の拡張として定義されている。Part2 Annex M では、各アプリケーションが、JPEG2000 の Codestream を透過画像として扱うために、チャンネルを1つのコンポーネントとして定義している。同ファイルフォーマットは、JPX として定義されている。

本稿で提案する分離符号化を Part2 で実現するのであれば、最終的なコードストリームは、グラフィック領域を示すコンポーネント (R_g, G_g, B_g) とテキスト領域を示す

(R_t, G_t, B_t)、さらにチャンネルを加えた計7つのコンポーネントを持つことになる。これは、分離符号化を実現する上で符号量の増加という問題を引き起こす。

4.2 JPEG2000 Part6 [2]

JPEG2000 Part6 は、MRC の概念を継承した分離符号化を実現する標準である [1]。Part2 では、チャンネルを1つのコンポーネントとして扱ったが、Part6 では、各オブジェクトを“Image”と“Mask”で表現し、オブジェクトごとに最適な符号化を施すことができる。Part6 ではこれらの符号化方式を JPM として定義している。JPM は、テキスト領域には JBIG2、グラフィック領域には JPEG2000 と言うように異なる符号化方式が混在した画像フォーマットである。

Part6 を利用しテキスト分離符号化を実現するためには、“Mask”の定義を明確にする必要がある。

5. テキスト分離符号化のメリット

5.1 テキスト領域解像度変換

テキスト領域の画質低下は、解像度変換を行った場合にも顕著に見られる。分離符号化をおこなうことで、なめらかな階調を持つグラフィック領域とエッジ成分を多く含むテキスト領域それぞれに対し、最適な解像度変換を施すことができる。

5.2 グラフィック領域の再圧縮

第2章で述べたように、グラフィック領域とテキスト領域では、圧縮特性が異なる。そこで、テキスト分離をおこなうことで、グラフィック領域だけを再圧縮し、相対的にテキスト領域の画質を保つことができる。

6. まとめと今後の課題

本稿では、静止画像に対し、テキスト分離符号化がテキスト領域の画質低下を防ぐ根本的な解決手法の1つであると共に、コンテンツの2次利用性を向上させる手法であることを述べた。また、同符号化方式を実現するためには、符号量増大、テキスト領域の解像度変換、文字認識など、様々な問題が存在することを示した。

今後は、テキスト領域のモデル化を行い解像度変換手法を検討する。

参考文献

- [1] ISO/IEC FDIS 16486 : Mixed Raster Content (MRC), 2000.
- [2] ISO/IEC FDIS 15444-6 : Compound Image File Format, December 2002.
- [3] ISO/IEC FDIS 15444-1 : Core Coding System, August 2000.
- [4] J.Nadenau, Reichel, Kunt, “Visually Improved Image Compression by Combining a Conventional Wavelet-Codec With Texture Modeling”, IEEE Transaction on Image Processing, Vol.11, No.11, November 2002.
- [5] 内山, 高橋, “JPEG2000 における適応圧縮に向けたタイル分類についての一検討”, 信学技報, ITS2001-49, IE2001-188, January 2002.
- [6] ISO/IEC FDIS 15444-2 : Extensions, December 2001.