適応型多レベルブロックランケーション符号化による小型モバイル表示器用画像圧縮手法の提案

Mobile Multi-level Adaptive Block Truncation Coding for Small-size Displays

佐々木 尚[†], 南 崎 浩 徳[†], 正会員 奥 村 治 彦^{††}

Hisashi Sasaki[†], Hironori Minamizaki[†] and Haruhiko Okumura^{††}

Abstract We developed mobile multi-level adaptive block truncation coding (BTC) that reduces the size of the mobile source driver's embedded RAM by 3/4 (for VGA screen with QVGA-size memory). The BTC suppresses border artifacts, especially in images shown in windows of PCs, and it corrects jagged edges in natural images. It also improves the PSNR by approximately 6.6 dB compared with that of conventional quad decimation for natural images. Our BTC enables the suppression of artifacts using: (1) trimmed averaging for three-level representative values and (2) adaptive scheme selection (two- or three-level).

キーワード: BTC, 小型表示器, 画像圧縮, 高精細, 内蔵RAM, ドライバIC

1. ま え が き

携帯電話などモバイル表示器においてさらなる高精細化の要求が強くなっている。携帯電話は小さい画面であるからQVGAでもう充分との見解もあるが、インターネットでのフルブラウジングを考えた場合、VGAでもまだまだ不足である。高精細になるとそれを駆動するためのソースドライバICとしては、その中に搭載される内蔵RAMのサイズ増加がコストアップとなり、特に問題となる。低コスト化の一つの手法として、RAMをすべてなくして、画像データを毎回転送する方式があるが、消費電力が大きくなる。特に静止画の場合には、画像が変わらないのであるから、ドライバICに内蔵RAMを持たせることで、消費電力を削減できる方式がモバイル機器には有効である。しかし、このような内蔵RAMサイズの増加に関していえば、VGAはQVGAの4倍も大きく、ドライバICの面積サイズを2倍に増加させてしまい、実用化の壁となっている。そこで現状は、VGAデータを4分の1にまびいて部分保存するパ-

シャルRAM方式が採用されている。しかしこれでは、表示器がVGAであることを十分に生かしきれていない。そこで、本報告はこのような内蔵RAMの圧縮手法を考える。

画像圧縮はその応用を考えると大きく二つの技術分野に分けることができる。第一の技術はDPCMやBTCといったアプローチで、実現ハードウェアサイズが小さく、周辺装置での利用に適している。DPCMは差分により画像データを符号化するので、前から順番に処理するため、ランダムアクセスできない。一方、BTCはブロック単位で処理を行うので、その単位でランダムアクセス可能である。第二の技術はJPEGなど変換符号化というアプローチで、ハードウェアはDPCMやBTCと比較すると100倍以上も大きくなるが、圧縮率が高いので、画像転送のメインストリームとして適している。内蔵RAMに保存される画像は、その変化部分だけをデータ変更することで、さらに消費電力を削減できる。この部分変更に、ランダムアクセス機能が不可欠である。そこで我々は、ランダムアクセス機能を必要とすること、周辺機器中に内蔵可能なようにハードサイズが小さいことから、BTCを選択した。BTCはブロック処理のためにエッジで画質劣化するという問題があり、これを改善する手法を提案する。

2. 新規ブロックランケーション符号化 (BTC) の提案

我々は、MOMA-BTC (Mobile Multi-level Adaptive Block Truncation Coding) と呼ぶ、新しい多レベル適応型

2006年12月7日, IDW'06で発表

[†] 株式会社東芝 セミコンダクター社

(〒247-8585 横浜市栄区笠間 2-5-1, TEL045-890-2635)

^{††} 株式会社東芝 研究開発センター

(〒212-8582 川崎市幸区小向東芝町 1, TEL044-549-2440)

[†] Toshiba Corporation, Semiconductor Company

(2-5-1, Kasama, Sakae-ku, Yokohama 247-8585, Japan)

^{††} Toshiba Corporation, Corporate Research and Development Center

(1, Komukai-Toshiba-cho, Saiwai-ku, Kawasaki 212-8582, Japan)

図 1 BTCの原理
Principle of BTC.

ブロックランケーション符号化を提案する. MoMA-BTCの主たる特徴は: (1) コデック (符号化器, 復号化器) のハードウェアサイズが小さいわりには, データサイズを $1/4$ に圧縮でき, VGAサイズをQVGAサイズに圧縮するのに都合がよいこと, (2) PC画面のウィンドウ境界と文字といったシャープなエッジを持ったデスクトップ画面で現れる特有のBTCアーティファクト (具体的には後述する図11のアーティファクト) を改善する機能を増強したことである. また自然画像での直線エッジに関するギザギザ感も同時に改善される.

2.1 BTCの原理

まず, 従来技術として, (2レベルの) BTCの原理を説明する. BTCは画素ブロックを処理単位とした画像圧縮技術であり, この単位でランダムアクセスできる.

図1の 4×4 のブロックには例えばY信号の画素値が示されているとする. 1画素8ビットとすると, $4 \times 4 \times 8 = 128$ ビットを必要とする. BTCでは, まず平均値 $= 162.75$ を求め, これより上の値を持つ画素 (上グループ) と, 下の値を持つ画素 (下グループ) の二つのグループに分ける. さらに上グループ, 下グループそれぞれをひとつの「代表値」で近似するため, それらの平均値 (下代表値 114 と上代表値 191) を求める. また, このとき, グループ分けを示すために, 4×4 の画素に上グループか下グループかを示すビットを考える. これを「ビットマップ」と呼ぶ. このようなデータに変換することで, $2 \times$ 代表値8ビット+ビットマップ16ビット $= 32$ ビットとなり, 128から32ビットにデータ量を削減できる. 復元は, ビットマップの値を参照して, 二つの代表値のうちのどちらかを選択する*. フルブラウザ利用などのPC画面利用を想定し, このBTCアルゴリズムの改良として, 圧縮率を維持しつつ, 代表値の個数を3に増やすのが本提案である.

2.2 MoMA-BTCの提案

以下, 本提案を説明する. 図2と図4は本技術の特徴づ

* なお, BTC圧縮した画像を復元するときに, 画質改善のためフィルタ処理することも一般に可能であるが, PC画面での文字劣化を考慮して, ここではそのような処理は一切行っていない.

図 2 MoMA-BTC 符号化メインフロー
The Main Flow of MoMA-BTC Encoding.

図 3 MoMA-BTC の3レベルBTCの量子化器
The 3-level Quantizer for MoMA-BTC.

けるアルゴリズムを示す. 図3で本提案の量子化器を説明する. 図5は圧縮データの構成を示す. 図6は本BTCの適応選択における活動量閾値の経験的な決定を示す.

図2は本符号化のメインフローであり, 活動量によるBTCスキームの選択を示す. 平坦な場合 (活動量が小さい場合) に2レベルBTCを選択する. 変化が激しい場合 (活動量が大きい場合) に3レベルBTCを選択する. 活動量とは 4×4 ブロック中でYCC信号がどの程度変化するかを測定するパラメータであり, 垂直および水平方向で隣合う 4×4 ブロック中での2画素からなる24ペアに関する差分の絶対値の総和である. 垂直も水平も同じ重みで, 絶対値の和を加算するので, 方向による区別はない. 活動量閾値は図6のように経験的に決める. なお, 2レベルBTCは従来BTCと同じであり, 平坦部分での8ビット精度を可能にしている.

図3は3レベルBTCの量子化器の原理を説明する. 本3レベルBTC実行での重要な点は, ここでの閾値決定である. 3レベルでの区間および代表値を考えるときに, なにも考えないで等間隔にグループ化する量子化器を $A1s \ a \ k$

a¹⁾が提案している。このような単純な予測は特にPC画面に関して好ましくない。詳細は図11で説明する。4x4ブロックでの最小値をMIN, 最大値をMAXとする。本量子化器は前もってMINとMAX画素を意図的に排除して平均を求める(排除の許容範囲設置も可)。この結果得られた平均AVE'はよりよい予測を与える。具体的には, 二つの量子化閾値を, $Th\_low = MIN + \Delta 1 / 2$, $Th\_high = MAX - \Delta 2 / 2$ とし量子化を実行する。ここで, $\Delta 1 = AVE' - MIN$, $\Delta 2 = MAX - AVE'$ と定義する。

図4は本3レベルBTC符号化のメインフローを示す。第一ステップは, 4x4ブロック中のMIN, MAXを計算する。第二ステップは, 図2で説明した3レベル量子化器の原理により, 二つの閾値 Th_low と Th_high を決定する。第三ステップは, 次のようにして各区間の代表値を計算する。 $V\_plus = (\text{高区間中にある画素値の総和}) / (\text{高区間中にある画素数})$, すなわち, 高区間に関する平均値である。残りの二つの代表値も同様に計算する。第四ステップは二つの閾値 Th_high と Th_low によりビットマップを決定する。その画素値がどの区間に属するかにしたがって, このビットを選択する。例えば, 低区間なら00, 中央区間なら01, 高区間なら10とする。

図5は圧縮データの構成を示す。一般的に, BTCスキーム(2レベルBTCか3レベルBTCか)が異なると, 圧縮率も異なる。このサイズの不一致は, ランダムアクセス機能にとって, 望ましくない。そこで, 3レベルBTCにおいてYCC420まびきと6ビット/5ビットのまるめを行い, サイズを同じにする*。なお, ここでの画像(VGAで2.6インチ)は300ppiを超えて高精細であり。このような画像に関しては5ビットでも充分に良い画質が得られることはすでに知られている。例えば, 本庄²⁾や次田³⁾を参照すること。

特にPC画面が人工的に一様なデータであるときには, 大きな領域でアーティファクトとなり非常に目立ってしまう(図11)ことが起こる。そこで, 中央区間に関して再度平均値を求めなおすことで, 追加ハード量を少なくしながら代表値の精度を確保しようという発想が, 本アルゴリズムである。

図6は活動量閾値の決定を説明する。図6はテスト画像として図13(と図15)の表計算の列番号(列を示す見出し行)におけるAとBの境界で, 暗ラインと白ラインから成るものを使用している。図中のThは図2のThresholdのことであり, この値を192,128,96,64としたときに, 提

図4 3レベルBTCの符号化フロー
The encoding flow of 3-level BTC for MoMA-BTC.

図5 MoMA-BTCのメモリ構成
The memory configuration for MoMA-BTC.

* 2レベルBTCに関しては, 4x4ブロックに対して96ビットであり, 圧縮率は $(4 \times 4 \times 8 \times 3 + 2 \times 3 \times 8) / 96 = 384 / 96 = 4$ となる。3レベルBTCに関しては, そのサイズは $4 \times 4 \times 2 + 3 \times 6 + 4 \times 2 \times 2 + 3 \times 5 \times 2 = 384$ となる。実際には(BTCスキームを選択するフラグとして)余分に1ビット必要で, 最終的な圧縮率は4より若干低下する。このフラグは圧縮データの復号時に再度利用するので, フラグは各処理ブロックに対してデータとして保存する必要がある。

図 6 活動量閾値の決定
An experimental determination of activity threshold.

案した MoMA-BTC による符号/復号化した画像が数値の左側に示されている。上方と下方にある 4x4 ブロックの拡大は、境界が忠実に再現されているかどうかに注目している。同様の実験を多数行い、最終的に 96 を最適閾値とした。

3. 圧縮画像の評価

従来、画像解析において、画像全体での一つの大域的な P S N R 値 = 40 d B を、標準的な基準として天下りの受け入れてきた。しかしながら、人間の目が平坦領域では非常にデリケートに劣化を検出するが、一方、変化が激しければそう重大なエラーとして検出できない。このことを考慮すると、一律に 40 d B を適用することは、小型画面に対して過剰な要求仕様である*。そこで我々は、以下のような観点から M a t l a b のシミュレーション解析を行った：(1) 画像全体としての大域的解析とブロック毎の画像内の解析**、(2) 一つ一つの画像の解析と画像集団としての解析、(3) 元画像に対する絶対評価と 4 分の 1 まびきに対する相対評価、および 2 レベル B T C に対する相対評価、(4) 自然画像と人工画像：画像対象の特性で区別して解析。これらの観点により、M o M A - B T C の性能が十分に良いことを確認した。

図 7 は一つの画像に関しての種々の解析を示す。A は M o M A - B T C 圧縮画像、B は元画像、C はエラー画像でエラー値を 4 倍してある画像***、D は P S N R = 40 d B を閾値として判定した結果*4、E は低 P S N R (10 から 20

図 7 画像ラッシュモア山の解析
Several analyses for an image "Mt. Rushmore".

図 8 画像ラッシュモア山の局所的 P S N R ヒストグラム
Local P S N R histogram for an image "Mt. Rushmore"

d B) の詳細、F は高 P S N R (20 から 60 d B*5) の詳細である。なお、20 d B を境界値とする理論的な根拠はなく、経験的に、実際の P S N R 値分布の下限が 10 d B 程度あることをもとに、解析結果の表示のための値として決めた。これで P S N R 値として最悪な場所（主観評価での確認場所）を探した。

青空の部分が 40 d B 以上であり、画質の劣化を感じさせないことがわかる。一方、地面の P S N R は数値的には悪い。地面の画素値分布はバラバラであり活動量も大きく、3 レベル B T C が選択されるが、2 より増えたこの三つの代表値でもバラバラな値をうまく近似できないことが、P S N R 値悪化の原因である*6。しかし、周りの変化も同じように激しく、地面テクスチャとして認識するので、その数値ほどに劣化した印象を与えない。主観評価に関しては、さらに図 13 の主観評価で後述する。

図 8 は一つの画像に関する P S N R のヒストグラムである。左の山は地面に対応している。右の山は青空に対応している。これは、青空のような平坦領域が 40 d B 以上の P S N R を持つことを示している。地面のような活発に変化する領域は低い P S N R を持つことを示している。したがって、大域的 P S N R (ヒストグラムの平均) が 40 d B より低い値にもかかわらず、画質は良いことが再確認できる。

図 9 は、複数の自然画像に対するヒストグラム解析の累積である。集団的な解析でありかつ局所的解析である。図 8 の結果を、複数画像集団に関して再確認できる。

* しかし数値的に何 d B を基準とするかはまだよくわかっていない。局所的 P S N R 値に関してはなおさら不明である。

** ブロック毎の解析は、4 x 4 の小さな領域ごとに P S N R 評価を実行する意味でミクロ的でありかつ局所的な解析である

*** 差分は正負存在するので、差分 0 が表示値 1 2 7 となるように加算。

*4 高いと判断したものを 2 5 5、低いと判断したものを 0 と表示し、R G B

それぞれの解析結果を合成した。

*5 一般的な基準である 40 d B を灰色の中心値 1 2 7 として表示した。青空、地面それぞれでの P S N R 分布を確認する。

*6 有効ビットとしては、Y を 6 ビット、C b と C r をそれぞれ 5 ビットとまるめるのでエラーを増やす方向に働くが、根本的な原因は分布のバラツキである。

図 9 100 画像に対する局所的 P S N R ヒストグラムの累積
Accumulated local PSNR histogram for 100 images.

図 10 ペア T 検定による M o M A - B T C 圧縮の性能改善
The PSNR improvement of MoMA-BTC compared with Quad-decimation by a paired-T test.

図 10 は、M o M A - B T C 圧縮 (V G A) 画像と 4 分の 1 まびき (Q V G A) 画像の大域的 P S N R 差分を 100 画像に関して、ペア T 検定した解析である。M o M A - B T C^{*7} は、現在のパーシャル R A M である 4 分の 1 まびき (Q V G A)^{*8} と比較して、6.6 d B だけ P S N R 値を改善している^{*9}。図中左下の十字マークはペア T 検定での比較の基準、すなわち、まびき画像での平均 P S N R 値の相対位置を示す。また M o M A - B T C と従来の 2 レベル B T C¹⁰ は同等の性能を持つことを確かめた¹¹。つまり、元の P S N R 値を保ちつつエッジに関する改良を行うことができている。

図 11 は、P C 人工画像に関して、アーティファクトを改善する例を示す。従来の 2 レベル B T C では、矢印で示した (文字「グラフ」の上部分の) アーティファクト^{*}が発生す

図 11 アーティファクトの改善 (P C 人工画像)
Improvement of artifact for an artificial PC image.

図 12 ギザギザな直線エッジの改善 (自然画像)
Improvement of jagged line edges for a natural image.

る。一つの 4 x 4 ブロックだけでなく、続けて複数のブロックで同じ問題が起きるので、目立ちやすい。M o M A - B T C は、3 レベルを提供することで、このようなアーティファクトを除去できている^{**}。

図 12 は、自然画像に関して、ギザギザな直線エッジの改善する例を示す。比較に用いる基準としては元画像の絶対基準の他、二つの基準も追加している：(1) 4 分の 1 まびき (パーシャル R A M) との比較は、製品が世の中に出回っているもので、実証的な比較である。(2) 2 レベル B T C は従来の圧縮技術と比較を行うという意味で、理論的な比較である。パーシャル R A M から大きく改善されていること

^{*7} 平均 28.2606, 標準偏差 2.7732, SE 平均 0.2773.

^{*8} 平均 21.5225, 標準偏差 3.2966, SE 平均 0.3297.

^{*9} (MoMA-BTC) - (まびき) に関して平均 6.739, 標準偏差 0.086273, SE 平均 0.08627. 差の平均に対する 95% 下限が 6.59476. 「差の平均 = 0」 (ペア > 0) の T 検定で P 値 = 0.000 となり、大きいことを検証.

¹⁰ 平均 28.2511, 標準偏差 3.0617, SE 平均 0.3062.

¹¹ (MoMA-BTC) - (従来 BTC) に関して平均 0.009391, 標準偏差 0.421156, SE 平均 0.042116. 差の平均に対する 95% の信頼区間は (-0.074175, 0.092957). 「差の平均 = 0」 (ペアが 0 と等しくない) の T 検定で P 値 = 0.824. したがって、「等しくないこと」の帰無仮説「等しい」を棄却できないので、等しいと考える.

^{*} アーティファクト発生状況の補足。グラフの文字の「黒」、背景の「白」、上部背景の「灰色」の 3 色がブロック中に存在している。輝度で 3 レベルが必要となるので、2 レベル B T C ではこの 3 色を表現できない。上代表値を計算するとき、背景の「白」と「灰色」が平均化されて「薄い灰色」になる。このような状況は、ウィンド境界付近で文字が存在する場合に実際しばしば起こる。自然画像の場合でも同様な現象

が起こるのであるが、複数ブロックで継続して起こることは極めて少なく、問題とならない。

^{**} 図 11 の先の脚注に補足する。単純な 2 分割の予測では三つ目の「灰色」の値がいつも正しく予測できるとは限らない：灰色の値が 127 でない場合もありえる。本アルゴリズムでは「黒」と「白」を除いて平均するので、いつでも正しく予想できる。

図 13 実際の主観評価
Subjective image evaluation by human eyes.

がわかる。なお、パーシャルRAMではYCC空間での単純なまびきであり、フィルタ処理はしていない。

図 13 に主観評価の様子を示す。参考に、図 15 に圧縮画像（符号化/復号化した画像）の拡大を示す。精細度VGAで表示できる2.6インチおよび3.8インチの表示器を実際に利用して、VGAサイズの圧縮画像を主観評価した。元画像と圧縮画像を交互に切替えて表示し、目に付くところがないかを確認した（被験者20名）。結果として、元画像と圧縮画像の差を見つけた人はいなかった*。自然画像に関しても同様であった。

このように、画像解析と実際の主観評価の両方により、充分良い画像品質を持つことを確かめた。

図 14 にVGA内蔵RAMで「圧縮なし」と「圧縮あり」、[パーシャルRAM内蔵]、それぞれの場合のドライバICのサイズ見積もりを示す。コデックを入れても従来パーシャルRAMと比較して1割程度の増加で済む。Verilog記述(RTLで5760行)をSynopsysで論理合成し、ゲートサイズ(25KG)から、ランダム論理部分の経験的平均ゲート密度を利用して回路面積の見積もりを行った。

4. む す び

本報告では、小型携帯表示器用に、1/4に圧縮しても画質劣化の少ないBTC方式を提案し、画像解析および主観

評価により十分な画質であることを確かめた。

本報告では内蔵RAM圧縮を目的としたが、送信データが圧縮されれば、転送速度を低くでき、さらなる低消費電力化につながる。したがって、符号化器をドライバICからコントローラに移行させることも今後の課題である。

また、ワンセグなどテレビ動画を扱う場合、液晶の応答遅れによる画質劣化（動きぼけ）を補正するために、オーバドライブする場合、フレームメモリーを利用する必要がある。本技術はこのような応用にも適用できる。

プロセス微細化によりメモリーセルを削減するアプローチもある。本アプローチでは、歩留まり的には枯れたプロセスで低コスト化を狙うので、ドライバICのプロセスは現状のままでよく、多大なプロセス微細化投資は必要でない。なお、プロセス微細化は画像圧縮技術と排他的ではないので、さらなる精細度向上への補完的な技術である。

小型携帯用表示器はVGAで300ppiを超え、多少の歪みがあっても、視覚的に許容できる状況となった。他のロスあり圧縮技術によるさらなる改善も今後の課題である。

〔文 献〕

- 1) Y.A. Alsaka and D.A. Lee: "Three Level Block Truncation Coding", IEEE Southeastcon pp.420-423(1990)
- 2) 本庄: "ハードコピーの階調再現", テレビ学誌, 49, 7, pp.836-841(1995)
- 3) 次田: "階調数、解像度と画質の関係", 電子写真学誌, 25, 1, pp.59-64(1986)
- 4) Jun Someya, Noritaka Okuda, Akihiro Nagase and Hiroaki Sugiyama: "Dynamic Bit-Rate Allocation of High-Compression Feed-forward Driving", SID '06 DIGEST paper, 60-2, pp.1797-1800 (2006)
- 5) Daigo Sasaki and Hiroshi Hayama: "Tiny Code for Decreasing Frame Memories in Display Driver ICs: Smart Pixel-data Codec (SPC)", SID '06 DIGEST paper, P-55, pp.402-405 (2006)

さ さ き ひろし
佐々木 尚 1980年、金沢大学工学部卒業。1986年、東京工業大学大学院理工学研究科博士課程修了。工学博士。1985年東芝セミコンダクター社に入社。1999年IEEE-Std. 1076.1(VHDL-AMS) 言語設計に貢献。

みなみ ひろし
南崎 浩徳 1984年、熊本大学工学部卒業。1986年、熊本大学大学院工学研究科修了。1986年、東芝セミコンダクター社に入社。主としてEL driver IC並びにLCD driver ICに関する開発に従事。

おくむら はるひこ
奥村 治彦 1983年、早稲田大学大学院理工学研究科電気工学専攻修士課程修了。同年東芝に入社。固体撮像装置（デジタルカメラ、ビデオカメラ）の高画質化技術、テレビ会議やテレビ電話の画像圧縮技術に関する研究開発を経て、1988年より、アクティブマトリックス型液晶表示装置の研究開発に従事。現在、ディスプレイ技術の研究主幹。工博、正会員。

* スライドショー機能を用いた切替タイミングを試験者がわからないようにした数秒間隔での自動切替を行ない第一の主観評価とした。エラーの答え（ウィンド右上のクローズ・ボタンやアイコンでのエラー場所）を知らない場合は認識できる人はいなかった。さらに自動切替で、答えを知った場合でも変化タイミングを感じとれない人もあった（半数）。また第二の評価として、被験者自身がボタンを押して切替タイミングがわかる手動切替をして、答えを教えた場合に、注視してもらい変化の違いを確認した。これは認識できた。さらに第三の評価として、虫眼鏡で変化部分を拡大し、本当に変化していることを被験者が確認した後、再度虫眼鏡を除いて肉眼で変化がどう感じられるか感想をきき、改めて変化が簡単にわからないことも被験者に再認識していただいた。なお、エラー探しに30秒から1分の時間を与えた。人間の目は動きに関して敏感であることも考慮しつつ、通常は悪い箇所を先に教えてまで主観評価することはないことも考えると、第一の自動切替の結果の「差はわからない」としても問題ないと思える。

図 15 図 13 の評価画像の拡大
Zoomed evaluation image in the figure 13.

図 14 ドライバICのサイズ見積もり
Size estimation of source driver ICs.