

超低遅延映像符号化に向けた 2パス符号化構成法に関する一検討

小林 大祐^{†1} 横張 和也^{†1} 新田 高庸^{†1}
上倉 一人^{†1} 如澤 裕尚^{†1}

自然な映像コミュニケーションサービスを実現するためには、映像符号化の低遅延化が重要となる。本研究では、10msec以下のコーデック遅延を実現する超低遅延映像符号化の実現に向け、符号化を行う処理単位を細分化して遅延を短くすることに着目した。本報告では、複数マクロブロックライン単位ごとに対して第一符号化→復号(解析)→第二符号化を行う2パスアルゴリズムを用いる方式において、細分化した処理範囲内での符号量制御・画質向上を行う処理の予備検討を行った。ソフトウェアシミュレーションにより、従来の1フレーム単位から更に細分化された符号化処理においても、同一ビットレートにおける画質が向上することが確認された。

A Study on Ultra Low Delay Video Encoder with 2-pass Coding Method

DAISUKE KOBAYASHI,^{†1} KAZUYA YOKOHARI,^{†1}
KOYO NITTA,^{†1} KAZUTO KAMIKURA^{†1}
and HIROHISA JOZAWA ^{†1}

In order to realize natural video communication services, it is necessary to shorten delay on video encoding. To realize an ultra low delay video codec whose delay is less or equal than 10msec, we are focused on shortening delay by subdividing the processing units to perform encoding. In this paper we propose a 2-pass algorithm processed in every some macroblock lines. By computer simulations, we confirm our proposed algorithm is able to improve image quality at the same bitrate even if encode processes are subdivided.

1. はじめに

近年、テレビ電話に代表されるような映像コミュニケーションサービスに注目が集まっている。特に企業においては、環境負荷低減、出張費・移動時間の削減、パンデミック対策として重要度が上がっている。また、遠隔医療・遠隔教育等での利用も行われるなど、近年様々な用途で用いられるようになってきている。

従来の映像コミュニケーションサービスで用いられる映像コーデックの例として、165msecの低遅延での会話が可能なソフトウェアコーデック¹⁾がある。通常の会話においては、165msecの遅延は問題にならない遅延量であるが、遠隔協調作業(ジャンケンや演奏など、相手との身体的動作の連動が必要な作業)を行うには、この遅延量でも問題がある。

一方、最近になってコーデック遅延が10msecのH.264/AVC^{*1}超低遅延エンコーダが登場し始めている²⁾³⁾⁴⁾⁵⁾。ただし、これらの製品は放送素材伝送用途を対象とした製品であり、高ビットレートでの運用を想定したものである。つまり、主に低レートでの利用を想定している家庭やオフィスの場における遠隔協調作業が可能な双方向コミュニケーション用途での利用分野は対象としていない。従って、家庭やオフィスで遠隔協調作業が可能な双方向コミュニケーションを実現するには、低レート域(～十数Mbps)での超低遅延映像コーデックが要求される。しかし、低レート域では画質と遅延のトレードオフがより厳しくなる。

そこで本研究では、遅延10msec以下、ビットレート十数Mbpsの超低遅延映像符号化における高画質化を目指す。

本報告は以下のように構成される。まず、第2章でコーデック処理において着目したバッファ遅延と、この遅延を削減するための従来方式およびその際の問題点について述べる。第3章では、本報告での提案手法を述べる。第4章では、提案方式の有効性を実験により示す。最後に第5章では、今後の課題および本報告のまとめを行う。

2. コーデック処理におけるバッファ遅延

NGNなどの帯域保証型のネットワークでは、映像の伝送レートには固定ビットレート(Constant Bitrate, 以下CBR)が用いられる。CBR制御では、エンコーダが生成したビッ

^{†1} 日本電信電話株式会社 NTT サイバースペース研究所

NTT Cyber Space Laboratories Nippon Telegraph and Telephone Corporation

^{*1} 動画符号化方式の標準の一つで、Blu-ray Disc やワンセグ放送などの標準符号化方式として採用されており、高い圧縮性能を誇っている。

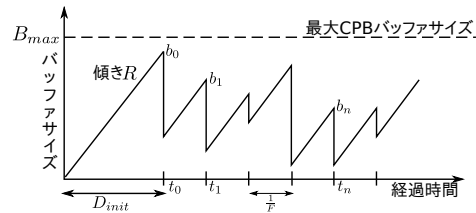


図1 CPB バッファモデルの概念図.
Fig.1 Concept of CPB buffer model.

トストリームをデコーダが正しく処理できることを保証するために、仮想デコーダモデルを導入している。この時の仮想デコーダで用いるバッファをCPB (Coded Picture Buffer) バッファと呼んでいる。CPB バッファは、エンコーダ側でエンコードされた各ピクチャの発生符号量を平滑化させ、CBR 伝送を保証する役割を持っている。ここで、このCPB バッファの動作をモデル化したものを図1に示す。

ここで、図1の横軸は経過時間、縦軸はバッファサイズを表している。図1における各パラメータは以下の通りである。

- D_{init} : 初期遅延
- B_{max} : 最大 CPB バッファサイズ
- b_i : i 番目のピクチャの符号量 ($i = 0, 1, 2, \dots, n, \dots$)
- t_i : i 番目のピクチャをデコードする時刻 ($i = 0, 1, 2, \dots, n, \dots$)
- F : フレーム (or フィールド) レート ($1/F$: フレーム (or フィールド) 間隔)
- R : ターゲットビットレート

このCPB バッファを破綻させない範囲内で、従来から様々な遅延削減手法およびそのための符号量制御手法が検討されている。

2.1 従来の遅延削減手法および符号化方式

2.1.1 Long GOP 方式

一般的なエンコーダにおいては、I/P/B の3つのピクチャ(ピクチャ: フレームまたはフィールドとする)で構成される Long GOP (GOP: Group of Pictures) 構造が広く用いられている。ここで、I/P/B ピクチャはそれぞれ発生符号量が大きく異なるため、固定ビットレート制御における符号量を平滑化するためには多量のCPB バッファが必要であった。更に、Bピクチャのための並べかえ遅延が発生するために、コーデック遅延は短いものでもおよそ500msec相当となっている。

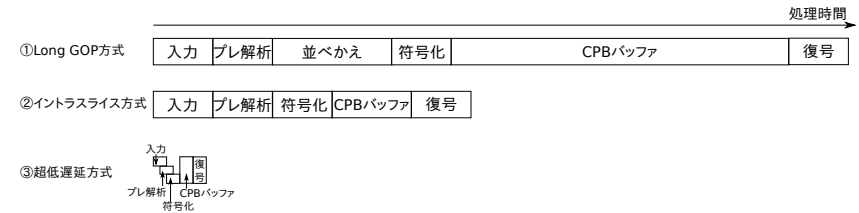


図2 符号化方式における処理遅延の内訳の違い
Fig.2 Breakdown of processing delays in each encoding scheme.

2.1.2 イントラスライス方式

各ピクチャ間の発生符号量を均一化させることによってバッファサイズを削減し、低遅延を実現する方式として、符号量の多いIピクチャをより細かい単位であるスライスに分割し、Pピクチャにイントラスライスを挿入するイントラスライス方式が提案されている。ここでGOP構造はPPPP...となり、イントラ符号化による予測誤差をリフレッシュするため、イントラスライスを各Pピクチャごとにずらしながら周期的に配置する。このイントラスライス方式の採用により、符号量が多く伝送に時間を要するIピクチャが不要になり、CPB バッファサイズの削減を行うことができる。

2.1.3 符号化処理単位の細分化による追いかかけ処理

図1に示したCPB バッファモデルはピクチャ単位にバッファサイズが変化することを前提としていたが、必ずしも1ピクチャ単位でバッファを制御する必要は無い。つまり、今まで1ピクチャ単位で行っていた処理を更に細分化させ、マクロブロック (Macro Block; 以下,MB) ラインごとに符号化やストリーム出力を行う、いわゆる「追いかかけ処理」を行うことで、遅延を1フィールド以下に縮めることが可能である。超低遅延映像符号化を実現する上では、この追いかかけ処理が基本的な考え方となる。

Long GOP 方式による符号化処理から超低遅延動作による符号化処理における、処理遅延の内訳の概念図を図2に示す。

2.2 符号量制御

符号化遅延の短縮と高画質化を両立させた符号化方式として、著者らはランタイム2パス符号化方式を提案している¹⁰⁾。ランタイム2パス符号化方式では、ピクチャごとに処理を行った1パス目の符号化結果をもとに、各MBのRD (Rate-Distortion) 特性が良くなる方向への量子化パラメータ制御を行い、第二符号化を行う。第一符号化ブロックが終わったピクチャについての解析情報を、第二符号化ブロックで即時参照してピクチャ内の符号

量再配分を行う。第二符号化を行う際の量子化パラメータ制御情報をフレーム単位の情報に元に出しているため、MBライン単位で処理を行う必要がある超低遅延環境下ではそのまま用いることが出来ない。また、文献¹⁰⁾は主に高圧縮化・高画質化を目指したものである。そのため第二符号化における量子化パラメータ算出においては、符号化効率を最大にすることを主眼としており、CPBバッファ状態を加味した量子化パラメータ決定は行っていない。

2.3 符号化処理単位をピクチャ未満にした際の問題点

ピクチャの上側と下側で映像の複雑さが異なる場合などで空間的位置によって発生符号量が異なる場合、バッファ破綻を起こさないためには、発生符号量が最も多くなる領域分を見越したバッファサイズを用意しておく必要がある。

また、局所的な発生符号量増によるバッファ破綻を回避するためには、符号量制御において目標とするビットレートを、設定ビットレートよりも低く見積もって符号化を行えば良い(足りない分は、スタッフィングビットを挿入する)。ただし、通常遅延での同一ビットレートにおける画質と比べて、デコード画像の画質が悪くなるという欠点がある。

3. 提案手法

3.1 アプローチ

前述の通り、超低遅延映像符号化を実現するには、符号化・復号の各処理を細分化し、処理の終わった部分を順次エンコーダから送出し、デコーダ側で順次復号をしていく必要がある。そのため、符号量制御処理についても、この細分化した領域で閉じた処理を行う必要がある。そこで本報告では、図3に示すように、1ピクチャよりも更に細分化した単位であるMBラインブロックで2パス符号化を行うことで、超低遅延環境下での映像の高画質化を行う手法を提案する。通常の1パス符号化による方法と比較すると、2パス符号化を行うことで、1パス目(第一符号化)の発生符号量・量子化パラメータを利用することにより画質や符号量制御の調整の柔軟性の高さが期待されるが、デメリットとして符号化を2回行う分の遅延がかかる。本手法においては、通常の1パス符号化を行うMBライン数よりも更に半分程度のMBライン数で符号化処理を行うことで、2パス符号化に起因する符号化遅延の解消を目指す。本手法による符号化処理の概要を図4に示す。

本手法による処理の流れは以下ようになる。各MBラインブロックに対して、以下の処理を繰り返す。

- (1) 第一符号化を実行する。
- (2) 符号化を行なったMBラインブロックに対して復号処理を行い、MB毎に発生符号

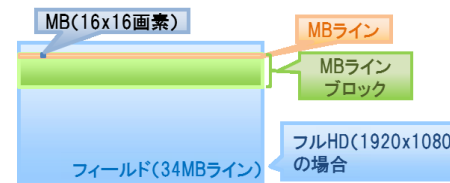


図3 MBラインブロックの模式図
Fig.3 Concept of MB line block.

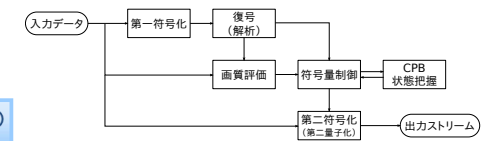


図4 提案方式のブロック図
Fig.4 Block diagram of proposal method.

量、量子化パラメータ ($QP^{(1)}$)、MBモード情報を得る。

- (3) 所定のMBラインブロック内の全MB符号化データから、画質・発生符号量を考慮して、第二符号化時に用いる量子化パラメータ ($QP^{(2)}$) を計算する。
- (4) $QP^{(2)}$ を用いて、第二符号化を実行する。
- (5) 第二符号化が終わったMBラインブロックは送信バッファに入れられた後に順次送信し、送信バッファから取り除く。
- (6) 次のMBラインブロックを対象にして、(1)からの処理を行う。

3.2 符号量再調整

本方式の第二符号化の際の符号量再調整処理においては、以下のことを検討する。ここで、当該MBラインブロックにおけるj番目の第一符号化時QP値を $QP_j^{(1)}$ 、第二符号化時QP値を $QP_j^{(2)}$ とする。

第一符号化後のMBラインブロックのMB発生符号量を元に、CPBバッファステータスの更新を行う。更新直後のCPBバッファステータスを元に、第二符号化時の量子化パラメータ $QP^{(2)}$ は大まかに以下のように決定される。ここで、 $\Delta \geq 0$ は量子化パラメータ調整変数である。

- (1) CPBバッファが破綻している場合(アンダーフロー): 画質が悪い場合は、 $QP^{(2)} = QP^{(1)} - \Delta$ となるような処理を行うことで、画質の向上を目指すような処理を行う。または、予定よりも符号量が出ていない場合は、スタッフィングビットによってCBR性を確保する。
- (2) CPBバッファが破綻している場合(オーバーフロー): 符号化対象MBラインブロックの絵柄が複雑な場合、固定ビットレート伝送において予定している発生符号量よりも実際の発生符号量が多くなってしまい、バッファ破綻の原因となる。その為、符号量再調整においては、何よりもバッファ破綻を起こさない符号量にまで第二符号化の

際の発生符号量を削減する必要がある。つまり、この状態では予定よりも符号量が出過ぎていて、CPB バッファ破綻の回避を最優先とするように、全体的に $QP^{(2)} = QP^{(1)} + \Delta$ となるような処理を行う。

- (3) CPB バッファが破綻していない場合：CPB バッファが破綻しない範囲内で、画質向上のため、MB ごとに $QP^{(2)} = QP^{(1)} \pm \Delta$ となるような処理を行う。大まかに言うと、割り当て符号量を少し多くすれば画質が良くなりそうな MB に対しては $-\Delta$ をし、割り当て符号量を少なくしても画質にほとんど影響がなさそうな MB に対しては $+\Delta$ となるような処理を行う。

3.3 符号化効率を考慮した目標発生符号量の決定

本方式では、文献¹⁰⁾にあるピクチャ単位での RD 特性近似曲線を用いて目標発生符号量を決定する方法を、MB ラインブロック単位に応用したものをを用いることにする。そこで、第二符号化時の各 MB ラインブロックの符号化効率を E_{MBLines} を次のように定義し、この E_{MBLines} を最大にするような目標発生符号量を求める。

まず、ある MB において、 x を MB 最大発生符号量 (3200bit) で正規化した MB 発生符号量、 $f(x)$ を画質評価指標である SSIM とし、この MB における発生符号量と SSIM の関係を表す RD 特性近似曲線を以下のように近似する。

$$f(x) = 1 - (1 - x)^a \quad (1)$$

このとき、符号化効率 $E(x)$ は、RD 曲線の傾きとして、以下で与えられる。

$$E(x) = f'(x) = a(1 - x)^{a-1} \quad (2)$$

ここで、 a は MB ごとに異なる変数であり、 j 番目の MB について、式変形により次のように求められる。

$$a_j = \frac{\ln(1 - f(x_j))}{\ln(1 - x_j)} \quad (3)$$

また、式 (2) より MB ラインブロックの符号化効率 E_{MBLines} は以下で定義される。

$$E_{\text{MBLines}} = a_{\text{MBLines}} (1 - \text{AVE}(b_{\text{MBLines}}))^{a_{\text{MBLines}} - 1} \quad (4)$$

ここで、

$$a_{\text{MBLines}} = \frac{\ln(1 - \text{AVE}(\text{SSIM}_{\text{MBLines}}))}{\ln(1 - \text{AVE}(b_{\text{MBLines}}))} \quad (5)$$

とする。このとき、 $\text{AVE}(\text{SSIM}_{\text{MBLines}})$ は MB ラインブロックの平均 SSIM、 $\text{AVE}(b_{\text{MBLines}})$ は MB ラインブロックの平均発生符号量を表す。

よって、式 (2) より第二符号化における j 番目の MB のターゲット符号量 M_{Tj} を次のよ

うに定義する。

$$M_{Tj} = 1 - \left(\frac{E_{\text{MBLines}}}{a_j} \right)^{\frac{1}{a_j - 1}} \quad (6)$$

3.4 バッファ状態を考慮した目標発生符号量決定

局所的な発生符号量増によるバッファ破綻を回避するため、第一符号化後のバッファサイズを考慮した目標発生符号量の決定を行う。これにより、ターゲットレート・ターゲット遅延に合わせたバッファサイズでバッファ破綻を起こす場合には、第二符号化時にバッファ破綻を起こさないような目標発生符号量を設定すれば良く、バッファ破綻回避の目的でバッファサイズに必要な以上の余裕を持たせる必要は無くなる。

2パス符号化における第二符号化時の目標発生符号量を次のように定義する。ピクチャ内の MB ライン数を N_{Picture} 、MB ラインブロック内の MB ライン数を N_{MBLines} としたとき、もしも第一符号化の結果、 j 番目の MB ラインブロックの発生符号量が b 、デコード直後のバッファサイズ B_j が $B_j < 0$ となると、すなわち $B_{j-1} + R \cdot \frac{N_{\text{MBLines}}}{N_{\text{Picture}}} - b < 0$ のとき、第二符号化時の MB ラインブロックの目標発生符号量 M_T を次式で定義する。

$$M_T = B_{j-1} + R \cdot \frac{N_{\text{MBLines}}}{N_{\text{Picture}}} - B_{\text{offset}} \quad (7)$$

とする。ここで、 B_{offset} は、バッファ調整変数であり、 $B_{\text{offset}} \geq 0$ である。

これより、MB ラインブロック内の j 番目の MB のターゲット符号量 M_{Tj} を次のように定義する。

$$M_{Tj} = c_j \cdot M_{1j} \cdot \frac{B_T}{b} \quad (8)$$

ここで、 M_{1j} は第一符号化における MB ラインブロック内の j 番目の MB 発生符号量、 j 番目の MB の c_j はスケール変数である。

3.5 第二符号化時の QP 決定

様々な MB に対して QP を 0 から 51 まで設定した時の QP と発生符号量の関係を見ると、QP-発生符号量は次のように二次方程式として近似することが出来る¹⁰⁾。

$$g(q) = \alpha q^2 = \beta q + \gamma \quad (9)$$

ここで、 q は量子化パラメータ QP、 $g(q)$ は MB 発生符号量である。また、 γ は MB の特性ごとに異なる値となる。そこで、この γ は第一符号化の結果を元に決定する。つまり、

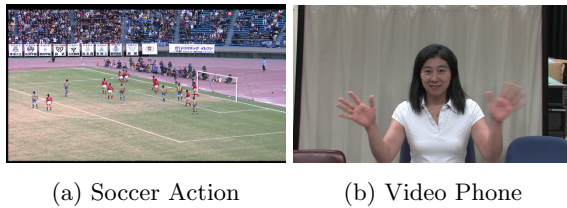


図 5 テストシーケンス
Fig. 5 Test sequences.

$$\gamma = g(QP^{(1)}) - \alpha \cdot QP^{(1)2} - \beta \cdot QP^{(1)} \quad (10)$$

から, j 番目の MB を第二符号化する際の $QP_j^{(2)}$ は, 式 (9)~(10) から次のように定義する.

$$QP_j^{(2)} = \frac{-\beta \pm \sqrt{\beta^2 - 4 \cdot \alpha \cdot (\gamma - M_{Tj})}}{2 \cdot \alpha} \quad (11)$$

ここで, M_{Tj} は第二符号化における目標発生符号量である.

QP の値は H.264/AVC の規格により $0 \leq QP \leq 51$ で定義されるため, この値域に含まれる QP 値を第二符号化で用いる QP として設定する.

4. 評価

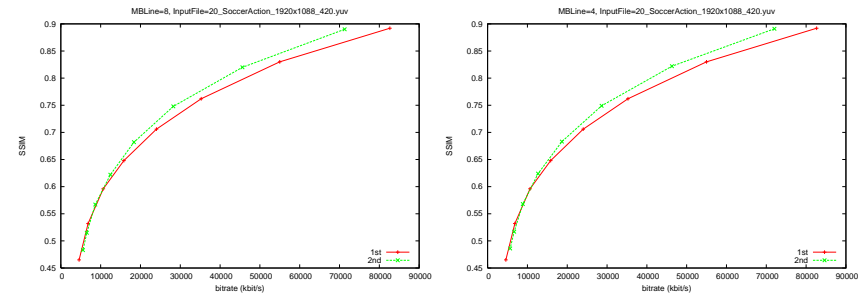
4.1 MB ラインブロック内での 2 パス符号化時における画質向上の検証

4.1.1 実験条件

MB ラインブロック内における 2 パス符号化の有効性を確認するため, ソフトウェアシミュレーションを行なった. ここでは, 第一符号化には固定 QP を用い, 特にレート制御は行わずに画質向上の度合いの確認をする. 符号化対象として, 図 5 に示す 8bit/pixel の 4:2:0 YUV フォーマット画像 (画像サイズ:1920x1088) を用いた. 図 5(a) は, ITE 標準動画画像である「Soccer Action」, 図 5(b) は, 民生カメラで撮影した「Video Phone」である. 第一符号化, 第一復号, 第二符号化には, H.264/AVC の参照ソフトウェア (JM 15.1) を用いた. 符号化における GOP 構造は ALL-I, 符号化フレーム数は 30 フレームで実験を行った. また, 式 (11) で用いる α , β は, 様々な MB に対して $g(q)$ を求めた予備実験結果から, $\alpha = 103.78$, $\beta = -107.66$ とした.

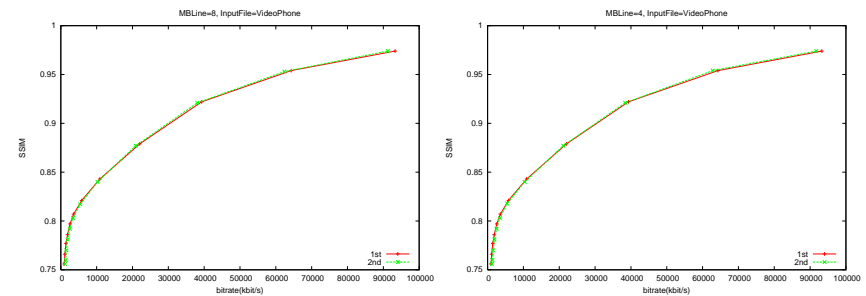
4.1.2 実験結果・考察

2 パス符号化を行う単位となる MB ラインブロック数を変化させた場合の, 1 パス目と 2



(a) 8 MB Lines (b) 4 MB Lines

図 6 MB ライン数を変えた時の RD 曲線の変化 (Soccer Action)
Fig. 6 RD curves in each MB lines. (Soccer Action)



(a) 8MB Lines (b) 4MB Lines

図 7 MB ライン数を変えた時の RD 曲線の変化 (Video Phone)
Fig. 7 RD curves in each MB lines. (Video Phone)

パス目の RD 曲線の比較を行った. その結果を図 6, 図 7 に示す. 図 6, 図 7 から, 1 ピクチャよりも細かい処理単位においても, 2 パス符号化を行うことで符号化効率が向上していることが分かる. また, 符号化効率の向上の度合いは高域のビットレートになるにつれて大きくなっている. ただし, MB ライン数を減らしていくと, 符号化効率の改善度合いが低下していることが分かる. これは, 処理単位の MB ラインを減らすにつれて, 符号量制御処理で MB ごとの符号量再配分を狭い範囲で行うことになり, 効率的な符号量配分が出来なくなってしまうためと考えられる.

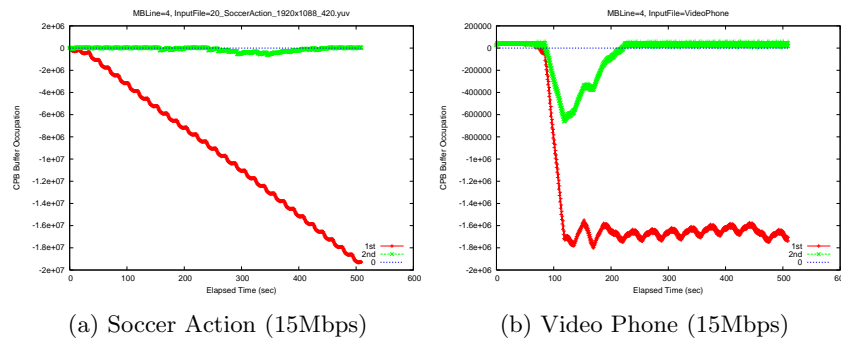


図8 4MBラインで2パス符号化を行った際のCPBバッファ遷移
Fig.8 CPB transition in 2-pass encoding with 4MB lines.

4.2 MBラインブロック内での2パス符号化時におけるバッファ状態の検証

4.2.1 実験条件

1パス目と2パス目のバッファ状態の様子を確かめる。第一符号化には、JMオリジナルのレート制御機能を用いた。符号化目標ビットレートは15Mbps、MBライン数は4MBライン、バッファサイズはMBライン数 \times 1.5(byte)として実験を行った。その他は4.1節における実験条件と同様である。

4.2.2 実験結果・考察

第一符号化(JMオリジナル)、第二符号化(提案手法)のそれぞれのCPBバッファ状態の比較を図8に示す。これより、JMオリジナルの第一符号化では、そもそもGOP単位でのレート制御を行なっているために、MBラインブロック単位での処理では完全にバッファ破綻を起こしてしまっている。一方、提案手法においては、バッファ状態がオーバーフローを起こし続ける状態にはなっておらず、バッファ破綻を回避するような動作をしていることが分かった。これは、MBラインブロック毎に、バッファ状態を見ながら第二符号化時の目標発生符号量を狙った量子化パラメータの変更を行なった効果が出たためと考えられる。

5. おわりに

本報告では、超低遅延映像符号化を実現する上で必要となる符号化処理の細分化に着目し、限られた符号化処理領域内での符号量再配分によって復号画質が向上するかどうかの評価を行った。ソフトウェアシミュレーションから、ピクチャを数MBラインブロックに

細分化した状態で2パス符号化を適用することで、主に中高域のビットレートでIピクチャの復号画像の画質が向上することを確認した。また、MBラインブロック毎に対して2パス符号化の際の量子化パラメータの再調整によって、固定ビットレート伝送のためのレート制御にも応用可能なことを確認した。

今後は、これらの画質改善手法及びレート制御手法を組み合わせることによって、超低遅延・低レートでの高画質な符号化方式の実現を目指していきたい。

参考文献

- 1) Takashi Sano, Takayuki Ohnishi, Hiroe Iwasaki, Kazuto Kamikura, Jiro Naganuma, "A Software-based H.264/AVC HDTV Real-time Interactive CODEC Architecture Using Parallel Processing," IEEE ICCE2010, Jan. 2010.
- 2) NEC, "VC・VD-7700/7710H.264 Ultra Low Delay CODEC," <http://www.nec.co.jp/bv/hoso/vcvc7700.html>, (2010-12), 閲覧可能.
- 3) IBEX テクノロジー, "超低遅延 H.264 エンコーダ / H.264 & MPEG-2 デコーダ HLD-3000," <http://www.ibextech.jp/hld3000.html>, (2010-12), 閲覧可能.
- 4) 近藤 雅俊, 高田 昌俊, 大波 雄一, 山口 宗明, "超低遅延 H.264 (High 4:2:2 10bit) コーデックの性能評価," 電子情報通信学会ソサイエティ大会講演論文集, vol.2007, p.150, Oct. 2007.
- 5) 三菱電機, "AVC/H.264 エンコーダ/デコーダ VM-1000E/D, -1100E/D," <http://www.mitsubishielectric.co.jp/info/tenjikai/InterBEE2010/index.html>, (2010-12), 閲覧可能.
- 6) ISO/IEC JTC1/SC29/WG11/N0400, "TEST MODEL5"
- 7) S. Ma, W. Gao, F. Wu, Y. Lu, "Rate control for JVT video coding scheme with HRD considerations," IEEE ICIP2003, vol.III, p.793-796, Sep. 2003.
- 8) Joint Model Reference Software, JM, "<http://iphome.hhi.de/suehring/tml/>," (2010-12), 閲覧可能
- 9) Z. Li, W. Gao, F. Pan, "Adaptive rate control with HRD consideration," Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, JVT-H014, May. 2003.
- 10) 横張 和也, 嵯峨田 淳, 上倉 一人, "マクロブロックのレート・歪み特性に着目した高効率2パス符号化アルゴリズムの検討," 情報処理学会研究報告, Vol.2010-AVM-68 No.2.