

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2010-136037

(P2010-136037A)

(43) 公開日 平成22年6月17日 (2010.6.17)

(51) Int. Cl. F I テーマコード (参考)  
 HO4N 7/26 (2006.01) HO4N 7/13 Z 5C059  
 5C159

審査請求 有 請求項の数 8 O L (全 15 頁)

(21) 出願番号	特願2008-309277 (P2008-309277)	(71) 出願人	00004226 日本電信電話株式会社
(22) 出願日	平成20年12月4日 (2008.12.4)	(74) 代理人	100087848 弁理士 小笠原 吉義
		(72) 発明者	中村 健 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
		(72) 発明者	清水 淳 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
		(72) 発明者	谷田 隆一 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内

最終頁に続く

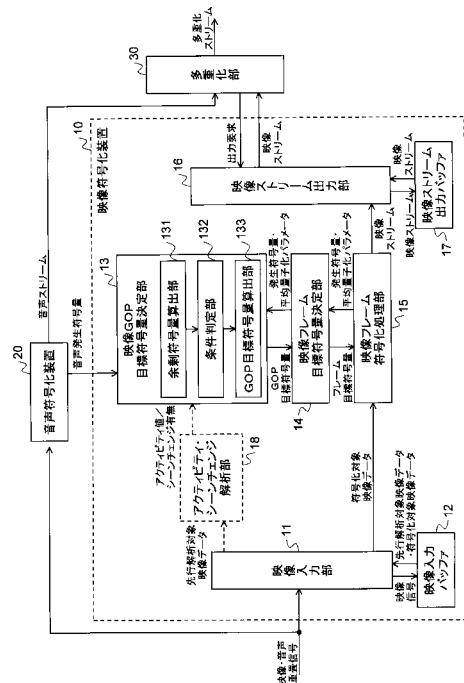
(54) 【発明の名称】 映像符号量制御方法、映像符号化装置、映像符号量制御プログラムおよびその記録媒体

(57) 【要約】

【課題】 映像符号化において、音声の余剰符号量の利用を図り、全体としては画質の変動を目立たせなくしながら、主観的な画質の向上を実現する。

【解決手段】 余剰符号量算出部131は、これから符号化するGOPの目標符号量を決定する際に、音声発生符号量からGOPの符号化に利用することができる音声の余剰符号量を算出する。条件判定部132により、予め定められた余剰符号量の利用条件を判定し、条件に合致した場合に、GOP目標符号量算出部133は、音声の余剰符号量の全部または一部を、本来のGOP目標符号量に加算する。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

映像符号化データと音声符号化データとを多重化した符号化ストリームが所定のビットレートの範囲内になるように、映像の符号量を制御する映像符号量制御方法であって、符号化済みの音声発生符号量から映像符号化制御単位の期間における音声の余剰符号量を算出する過程と、

前記余剰符号量を映像符号化制御単位における目標符号量に加えるかどうかの、あらかじめ定められた映像の発生符号量に影響する映像の特徴または符号化の条件を判定する過程と、

前記条件が満たされると判定された場合に、前記余剰符号量の一部または全部を映像符号化制御単位における目標符号量に加え、前記条件が満たされないと判定された場合に、前記余剰符号量を映像符号化制御単位における目標符号量に加えないで、次に符号化する映像符号化制御単位における目標符号量を決定する過程と、

前記決定された目標符号量に従って、映像符号化制御単位における映像信号を符号化する過程とを有する

ことを特徴とする映像符号量制御方法。

**【請求項 2】**

請求項 1 に記載の映像符号量制御方法において、

前記条件は、過去の所定数の映像フレームの平均量子化パラメータが所定の閾値以上であるという条件、または過去の所定数の映像フレームの平均量子化ステップが所定の閾値以上であるという条件、またはこれから符号化する映像符号化制御単位のアクティビティ平均値が所定の閾値以上であるという条件、またはこれから符号化する映像符号化制御単位にシーンチェンジがあるという条件、または過去の所定数の映像フレームの目標符号量と発生符号量との差の総和が所定の閾値以上であるという条件である

ことを特徴とする映像符号量制御方法。

**【請求項 3】**

請求項 1 または請求項 2 に記載の映像符号量制御方法において、

前記音声の余剰符号量を、前記映像符号化制御単位と同一表示期間の音声符号化制御単位群における、各音声符号化制御単位の所定の最大発生符号量と実際の発生符号量との差の総和から算出する

ことを特徴とする映像符号量制御方法。

**【請求項 4】**

請求項 1 または請求項 2 に記載の映像符号量制御方法において、

前記音声の余剰符号量を、前記映像符号化制御単位より所定の映像フレーム数だけ前の前記映像符号化制御単位と同じ長さの期間の映像フレーム群と同一表示期間の音声符号化制御単位群における、各音声符号化制御単位の所定の最大発生符号量と実際の発生符号量との差の総和から算出する

ことを特徴とする映像符号量制御方法。

**【請求項 5】**

映像符号化データと他の符号化データとを多重化した符号化ストリームが所定のビットレートの範囲内になるように、映像の符号量を制御する映像符号量制御方法であって、符号化済みの前記他の符号化データの発生符号量から映像符号化制御単位の期間における前記他の符号化データの余剰符号量を算出する過程と、

前記余剰符号量を映像符号化制御単位における目標符号量に加えるかどうかの、あらかじめ定められた映像の発生符号量に影響する映像の特徴または符号化の条件を判定する過程と、

前記条件が満たされると判定された場合に、前記余剰符号量の一部または全部を映像符号化制御単位における目標符号量に加え、前記条件が満たされないと判定された場合に、前記余剰符号量を映像符号化制御単位における目標符号量に加えないで、次に符号化する映像符号化制御単位における目標符号量を決定する過程と、

10

20

30

40

50

前記決定された目標符号量に従って、映像符号化制御単位における映像信号を符号化する過程とを有する

ことを特徴とする映像符号量制御方法。

【請求項 6】

映像符号化データと音声符号化データとを多重化した符号化ストリームが所定のビットレートの範囲内になるように、映像の符号量を制御し、映像信号を符号化する映像符号化装置であって、

符号化済みの音声発生符号量を取得し、映像符号化制御単位の期間における音声の余剰符号量を算出する手段と、

前記余剰符号量を映像符号化制御単位における目標符号量に加えるかどうかの、あらかじめ定められた映像の発生符号量に影響する映像の特徴または符号化の条件を判定する手段と、

前記条件が満たされると判定された場合に、前記余剰符号量の一部または全部を映像符号化制御単位における目標符号量に加え、前記条件が満たされないと判定された場合に、前記余剰符号量を映像符号化制御単位における目標符号量に加えないで、次に符号化する映像符号化制御単位における目標符号量を決定する手段と、

前記決定された目標符号量に従って、映像符号化制御単位における映像信号を符号化する手段とを備える

ことを特徴とする映像符号化装置。

【請求項 7】

請求項 1 から請求項 5 までのいずれか 1 項に記載の映像符号量制御方法を、コンピュータに実行させるための映像符号量制御プログラム。

【請求項 8】

請求項 1 から請求項 5 までのいずれか 1 項に記載の映像符号量制御方法を、コンピュータに実行させるための映像符号量制御プログラムを記録したコンピュータ読み取り可能な記録媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音声ストリームのビットレートと映像ストリームのビットレートとの和が所定のビットレートの範囲内になるように、音声の発生符号量に応じて映像の符号量を制御する映像符号量制御方法、映像符号化装置、映像符号量制御プログラムおよびその記録媒体に関するものである。

【背景技術】

【0002】

一般に、映像ストリームと、音声やデータなどの他のストリームとを多重化して送るような伝送システムでは、映像は、他のストリームとは独立にビットレート制御を行っている（非特許文献 1，2 参照）。

【0003】

一般に映像・音声符号化では、複雑で変化の多い入力信号に対しては符号量が多くなり、単調で変化の少ない入力信号に対しては符号量が少なくなる。

【0004】

ロッシェ符号化ストリームは、固定ビットレート（C B R : Constant Bit Rate）で伝送するために、信号の劣化の程度を変化させることにより、発生符号量を所定の範囲内に収めるように制御する。一方、ロスレス符号化ストリームでは、発生符号量の変化を制御することができないため、一般に可変ビットレート（V B R : Variable Bit Rate）で伝送される。

【非特許文献 1】ISO/IEC 13818-2 Annex C, ITU-T Recommendation H.264 Annex C

【非特許文献 2】映像情報メディア学会編，“総合マルチメディア選書 M P E G”，社団法人映像情報メディア学会，（株）オーム社発行，1996.4.20

10

20

30

40

50

## 【発明の開示】

## 【発明が解決しようとする課題】

## 【0005】

MPEG-4 ALSのようなロスレス音声符号化ストリームと映像符号化ストリーム（以下、符号化ストリームを単に「ストリーム」ともいう）とを多重化する場合、音声伝送ビットレートの変動が映像伝送ビットレートに比べて無視できないほど大きいことが多い。

## 【0006】

したがって、映像と音声の多重化ストリームにおいて、固定で割り当てられていた音声のビットレートのうち、音声符号化データの発生符号量が小さい余剰符号量分を映像符号化に利用することができれば、ビットレートに無駄が生じることがなく、画質の向上に寄与すると考えられる。

## 【0007】

しかし、このように音声符号化データの発生符号量の変動分を、映像符号化に有効に利用することを考えた場合、映像の全体としての画質は向上するものの、音声符号化データの発生符号量が多い場面では画質の劣化が大きく、音声符号化データの発生符号量が少ない場面では画質が鮮明になり過ぎることがあり、そのため音声符号化データの変動に応じて画質の変動が目立ってしまうことがあるという、新たな問題が生じる。

## 【0008】

図9は、本発明の課題を説明する図である。従来、符号化ビットレートが固定の音声符号化ストリームと映像符号化ストリームを多重化する場合、図9(A)に示すように、映像は、音声のストリームとは独立にビットレート制御が行われていた。なお、GOP(Group Of Pictures)は、一般に十数フレームから数十フレームのピクチャ群からなる映像符号化データの単位であり、ここで、GOP(n)はn番目のGOPを表している。

## 【0009】

ロスレス音声符号化ストリームの場合、音声の発生符号量は変動し、音声符号化データのデータ量が、与えられた音声ビットレートで送信可能なデータ量よりもかなり少なくなることがある。このビットレートの余裕分の符号量を、ここでは余剰符号量という。図9(B)に斜線を付して示している部分が余剰符号量である。

## 【0010】

図9(C)は、図9(B)に示した音声の余剰符号量を映像符号化データの伝送に利用することを考えた場合の音声ビットレートと映像ビットレートとの関係を示している。音声の余剰符号量を映像符号化データの伝送に利用することにより、映像符号化時におけるGOP単位の目標符号量を、元の映像ビットレートによる目標符号量よりも大きくすることができる。

## 【0011】

図9(C)に示すように、音声の余剰符号量分を映像符号化データの伝送に利用したとすると、映像の発生符号量を増加させることができるので、映像の画質が向上すると考えられるが、音声ビットレートの変動に応じて映像ビットレートが変動することになり、画質の変動が目につく可能性があるという、新たな問題が発生する。

## 【0012】

本発明は、上記課題の解決を図り、音声の余剰符号量分を映像符号化に有効に利用し、画質を向上させるとともに、全体として画質の変動を目立たせなくすることを目的とする。

## 【課題を解決するための手段】

## 【0013】

本発明は、上記課題を解決するため、ロスレス音声符号化ストリームやデータストリーム等の符号化ビットレートを制御できないストリームと、ロッキー映像符号化ストリームとをCBRに多重化する場合に、次のように音声の発生符号量に応じて映像の符号量を制御する。

10

20

30

40

50

## 【 0 0 1 4 】

( 1 ) 映像符号化時の目標符号量と発生符号量の差，量子化ステップサイズ，アクティビティなどのパラメータが，所定の条件に合致した場合にのみ，音声の余剰符号量を映像符号化に利用する。

## 【 0 0 1 5 】

( 2 ) 映像符号化に利用可能な音声の余剰符号量は，その時点で符号化済みの音声符号化データの余剰符号量や伝送タイミングから求める。また，余剰符号量の利用判断，利用量を定めるタイミングは，GOPの先頭またはフレーム単位とする。

## 【 0 0 1 6 】

( 3 ) 映像のGOP符号化開始時に，GOP期間分音声を先行して符号化した場合には，同一GOP期間の，そうでない場合には直前のGOP期間の音声の余剰符号量分を求める。余剰符号量の利用が決定された場合，その余剰符号量を上限とする。画質改善のために必要な符号量は，量子化ステップサイズ，アクティビティから計算する。

10

## 【 0 0 1 7 】

以上の点を踏まえて，本発明は，次のように映像の符号量を制御する。まず，符号化済みの音声発生符号量から映像符号化制御単位（例えばGOPまたはフレーム）の期間における音声の余剰符号量を算出する。その余剰符号量を映像符号化制御単位における目標符号量に加えるかどうかの，あらかじめ定められた映像の発生符号量に影響する映像の特徴または符号化の条件を判定する。条件が満たされる場合には，余剰符号量の一部または全部を映像符号化制御単位における目標符号量に加え，条件が満たされない場合には，余剰符号量を映像符号化制御単位における目標符号量に加えないで，次に符号化する映像符号化制御単位における目標符号量を決定する。決定された目標符号量に従って，映像符号化制御単位における映像信号を符号化する。

20

## 【 0 0 1 8 】

前記条件として，例えば次のような条件のいずれかを用いる実施が好適である。

- ・過去の所定数の映像フレームの平均量子化パラメータが所定の閾値以上であるという条件，
- ・過去の所定数の映像フレームの平均量子化ステップが所定の閾値以上であるという条件，
- ・これから符号化する映像符号化制御単位のアクティビティ平均値が所定の閾値以上であるという条件，
- ・これから符号化する映像符号化制御単位にシーンチェンジがあるという条件，
- ・過去の所定数の映像フレームの目標符号量と発生符号量との差の総和が所定の閾値以上であるという条件。

30

## 【 0 0 1 9 】

映像符号化に利用する音声の余剰符号量の算出方法としては，次のような方法を用いることができる。

- ・音声の余剰符号量を，映像符号化制御単位と同一表示期間の音声符号化制御単位群における，各音声符号化制御単位の所定の最大発生符号量と実際の発生符号量との差の総和から算出する方法，
- ・音声の余剰符号量を，映像符号化制御単位より所定の映像フレーム数だけ前の映像符号化制御単位と同じ長さの期間の映像フレーム群と同一表示期間の音声符号化制御単位群における，各音声符号化制御単位の所定の最大発生符号量と実際の発生符号量との差の総和から算出する方法。

40

## 【 0 0 2 0 】

本発明は，映像符号化データと，映像符号化データ以外の符号化データとを多重化して伝送する場合に，前述した音声符号化データの余剰符号量と同様に，映像符号化データ以外の符号化データの余剰符号量をもとに，追加目標符号量を決定し，映像符号化の符号量を制御することもできる。

## 【 発明の効果 】

50

## 【 0 0 2 1 】

本発明によれば、映像と音声とを多重化した符号化ストリームが所定のビットレートの範囲内になるように映像を符号化する場合に、映像の符号化が困難なシーンでは画質を向上させ、全体としては画質の変動が目立たせなくすることで、主観的な画質を向上させることができるようになる。

## 【 発明を実施するための最良の形態 】

## 【 0 0 2 2 】

以下、本発明の実施の形態を図面を用いながら詳細に説明する。

## 【 0 0 2 3 】

図 1 は、本発明の実施例に係る装置の構成例を示す図である。映像符号化装置 10 は、映像・音声重畳信号のうち映像信号を入力し、映像信号を符号化して映像ストリームを出力する。音声符号化装置 20 は、映像・音声重畳信号のうち音声信号を入力し、音声ストリームを出力する。多重化部 30 は、音声ストリームと映像ストリームとを所定のビットレートとなるように多重化して出力する。

10

## 【 0 0 2 4 】

本実施例において、音声符号化装置 20 は、フレーム毎もしくは逐次、音声発生符号量を映像符号化装置 10 に伝達する機能を持つほかは従来技術と同様に構成できるので、その内部構成についての詳しい説明は省略する。

## 【 0 0 2 5 】

映像符号化装置 10 は、映像信号を入力する映像入力部 11、入力した映像信号を蓄積する映像入力バッファ 12、GOP 毎の映像の目標符号量を決定する映像 GOP 目標符号化量決定部 13、映像フレーム毎の目標符号量を決定する映像フレーム目標符号量決定部 14、映像信号をフレーム目標符号量に従って符号化する映像フレーム符号化処理部 15、映像ストリームを出力する映像ストリーム出力部 16、映像ストリームを一時的に蓄積する映像ストリーム出力バッファ 17 を備える。

20

## 【 0 0 2 6 】

映像 GOP 目標符号量決定部 13 は、GOP 毎に音声発生符号量から映像符号化に利用することができる余剰符号量を算出する余剰符号量算出部 131 と、算出した余剰符号量を映像符号化に利用するかどうかを判定する条件判定部 132 と、条件の判定結果に従って GOP 目標符号量を算出する GOP 目標符号量算出部 133 とを備える。

30

## 【 0 0 2 7 】

アクティビティ・シーンチェンジ解析部 18 は、入力映像信号のうち符号化に先立って解析される先行解析対象映像データについて、アクティビティ値の算出もしくはシーンチェンジの有無を検出するものであり、後述する実施例 3、4 で用いられる。他の実施例では省略することができる。

## 【 0 0 2 8 】

## 〔 実施例 1 〕

図 2 は、本発明の実施例 1 の映像符号化処理のフローチャートである。実施例 1 では、過去のいくつかの数の映像フレームの平均量子化パラメータ  $QP_{ave}(n)$  が、ある閾値  $TH_{qp}$  以上の場合に、音声の余剰符号量を映像符号化に利用する。これは、映像符号化に利用できる符号量の不足により平均量子化パラメータ  $QP_{ave}(n)$  がある値以上になると、画質の粗くなる程度が大きくなり、画質向上の必要性が増すとともに、音声の余剰符号量の追加による画質の変動が目立ちにくくなるからである。

40

## 【 0 0 2 9 】

$n$  番目の GOP (Group Of Pictures) を、 $GOP(n)$  とする。映像と音声の符号化は並行して行われる。ただし、本実施例では、 $GOP(n)$  の映像フレーム符号化開始より前に、 $GOP(n)$  の映像フレームと同一表示期間の音声フレームの符号化が完了しているものとする。

## 【 0 0 3 0 】

まず、映像 GOP 目標符号量決定部 13 は、以下のステップ S101 ~ S105 の処理

50

を実行する。音声符号化装置 20 から音声発生符号量の情報を受け取り，GOP (n) の映像フレームと同一表示期間の音声フレーム  $j \sim j + M - 1$  について，GOP (n) 期間の余剰符号量  $G_r (n)$  を，次式によって求める (ステップ S101)。

【0031】

$$G_r (n) = \sum_{x=j}^{j+M-1} \{ S_{amax} - S_a (x) \}$$

ここで，M は，GOP (n) 期間の音声フレーム数， $S_{amax}$  は，各音声フレームの最大発生符号量， $S_a (x)$  は，x 番目の音声フレームの発生符号量である。 $S_{amax}$  としては，あらかじめ定められた値を用いてよい。

【0032】

次に，過去の数映像フレームの平均量子化パラメータ  $QP_{ave} (n)$  が，ある閾値  $TH_{qp}$  以上かどうかを判定する (ステップ S102)。平均をとる映像フレーム数は任意の数でよく，あらかじめ定められる。平均量子化パラメータ  $QP_{ave} (n)$  が，ある閾値  $TH_{qp}$  以上の場合には，余剰符号量  $G_r (n)$  を超えない範囲で，平均量子化パラメータ  $QP_{ave} (n)$  の閾値  $TH_{qp}$  超過分に比例して，追加 GOP 目標符号量  $G_d (n)$  を決定する (ステップ S103)。この場合の  $G_d (n)$  は，次式によって求められる。

【0033】

$$G_d (n) = \min (G_r (n), a \cdot (QP_{ave} (n) - TH_{qp}))$$

$\min ()$  は，最小値を返す関数であり，a は予め定められた定数である。

【0034】

一方，平均量子化パラメータ  $QP_{ave} (n)$  が，ある閾値  $TH_{qp}$  以上でない場合には，追加 GOP 目標符号量  $G_d (n)$  を 0 とする (ステップ S104)。

【0035】

次に，GOP 目標符号量  $G (n)$  を，GOP 内フレーム数  $N \times$  基準映像ビットレート  $R$  / 映像フレームレート  $P$  と，追加 GOP 目標符号量  $G_d (n)$  と，前の GOP からの繰越符号量の和で求める (ステップ S105)。GOP 目標符号量  $G (n)$  は，次のようになる。

【0036】

$$G (n) = N \cdot R / P + G_d (n) + \text{前の GOP からの繰越符号量}$$

なお，基準映像ビットレート  $R$  とは，音声フレームの発生符号量が常に最大の場合に確保できる映像ビットレートを指す。

【0037】

次に，映像フレーム目標符号量決定部 14 および映像フレーム符号化処理部 15 は，下に説明するステップ S106 ~ S109 の処理を，GOP (n) 内のフレームの符号化が終了するまで繰り返す (ステップ S110)。

【0038】

映像フレーム目標符号量決定部 14 は，GOP (n) の目標符号量を GOP 内の各フレームに対して配分し，次フレームの目標符号量  $T (i)$  を求める (ステップ S106)。GOP (n) の目標符号量から，次フレームの目標符号量  $T (i)$  を算出する方法としては，例えば次の参考文献 1 に開示されている方法などがあり，周知技術であるため，ここでの詳細な説明は省略する。

[参考文献 1] International Organisation for Standardisation, Test Model Editing Committee, 1993, Test Model 5, April, ISO-IEC/JTC1/SC29/WG11/N0400

次に，映像フレーム目標符号量決定部 14 は，伝送先の映像復号装置におけるデコーダ受信バッファのバッファサイズや転送済みのデータ量から算出したバッファ占有量をもとに，デコーダ受信バッファが破綻しない限界の符号量を求め，目標符号量はその値を超えないようにクリッピングする (ステップ S107)。なお，デコーダ受信バッファのバッファサイズが十分大きい場合や，デコードまでの遅延時間を大きくしてもよい場合には，このクリッピングの処理を省略する実施も可能である。

【0039】

映像フレーム符号化処理部 15 は，映像フレーム目標符号量決定部 14 が算出した目標

10

20

30

40

50

符号量に従って、フレーム  $i$  を符号化する（ステップ S 1 0 8）。なお、インターレースの指定があればフィールド単位で符号化するが、説明を簡単にするため、フィールド単位の符号化の場合も単にフレームの符号化として説明する。

【 0 0 4 0 】

フレームの符号化が終了すると、映像フレーム目標符号量決定部 1 4 は、GOP (  $n$  ) の目標符号量から発生符号量  $S ( i )$  を減算する（ステップ S 1 0 9）。その後、GOP (  $n$  ) 内の全フレームについて符号化が終了したかどうかを判定し、終了していなければ、ステップ S 1 0 6 に戻って、全フレームについて符号化が終了まで同様に処理を繰り返す（ステップ S 1 1 0）。

【 0 0 4 1 】

映像 GOP 目標符号量決定部 1 3 は、GOP 開始時の GOP 目標符号量  $G ( n )$  と、GOP 発生符号量との差から、次の GOP への繰越符号量を決定する（ステップ S 1 1 1）。以下、ステップ S 1 0 1 に戻って、次の GOP について同様に符号化処理を行う。

【 0 0 4 2 】

〔実施例 2〕

図 3 は、本発明の実施例 2 の映像符号化処理のフローチャートである。実施例 2 では、過去のいくつかの数の映像フレームの平均量子化ステップ  $Q_{ave} ( n )$  が、ある閾値  $TH_q$  以上の場合に、音声の余剰符号量を映像符号化に利用する。これは、映像符号化に利用できる符号量の不足により平均量子化ステップ  $Q_{ave} ( n )$  がある値以上になると、画質の粗くなる程度が大きくなり、画質向上の必要性が増すとともに、音声の余剰符号量の追加による画質の変動が目立ちにくくなるからである。

【 0 0 4 3 】

前述した実施例 1 と実施例 2 との違いは、音声の余剰符号量を映像符号化に利用するかどうかの条件の判定を、平均量子化パラメータ  $Q_{P_{ave}} ( n )$  で行うか平均量子化ステップ  $Q_{ave} ( n )$  で行うかという点である。量子化パラメータと量子化ステップの換算式は、符号化規格によって異なるが、一般に「(符号量) =  $a / (量子化ステップ) + b$ 」( $a$ ,  $b$  は定数)で近似できるため、実施例 1 より実施例 2 のほうが追加目標符号量の計算精度がよい。

【 0 0 4 4 】

実施例 2 では、図 2 に示すフローチャートのうち、ステップ S 1 0 2 ~ S 1 0 4 が、図 3 のステップ S 2 0 2 ~ S 2 0 4 に置き換わり、他の部分については同様であるので、以下では、ステップ S 2 0 2 ~ S 2 0 4 の部分についてだけ説明する。

【 0 0 4 5 】

図 2 に示すステップ S 1 0 1 を実行した後、映像 GOP 目標符号量決定部 1 3 の条件判定部 1 3 2 は、過去の数映像フレームの平均量子化ステップ  $Q_{ave} ( n )$  が、ある閾値  $TH_q$  以上かどうかを判定する（ステップ S 2 0 2）。GOP 目標符号量算出部 1 3 3 は、平均量子化パラメータ  $Q_{ave} ( n )$  が閾値  $TH_q$  以上の場合には、余剰符号量  $G_r ( n )$  を超えない範囲で、平均量子化パラメータ  $Q_{ave} ( n )$  の閾値  $TH_q$  超過分に応じて、追加 GOP 目標符号量  $G_d ( n )$  を決定する（ステップ S 2 0 3）。この場合の  $G_d ( n )$  は、次式によって求められる。

【 0 0 4 6 】

$$G_d ( n ) = \min ( G_r ( n ) , a \cdot ( 1 / TH_q - 1 / Q_{ave} ( n ) ) )$$

$\min ( )$  は、最小値を返す関数であり、 $a$  は予め定められた定数である。

【 0 0 4 7 】

一方、平均量子化パラメータ  $Q_{ave} ( n )$  が閾値  $TH_q$  以上でない場合には、追加 GOP 目標符号量  $G_d ( n )$  を 0 とする（ステップ S 2 0 4）。

【 0 0 4 8 】

以降の処理は、実施例 1 における図 2 のステップ S 1 0 5 ~ S 1 1 1 と同様である。

【 0 0 4 9 】

〔実施例 3〕

10

20

30

40

50



図4は、本発明の実施例3の映像符号化処理のフローチャートである。実施例3では、これから符号化するGOPのアクティビティ平均値  $act(n)$  が、ある閾値  $TH_{act}$  以上の場合に、音声の余剰符号量を映像符号化に利用する。これは、アクティビティ平均値  $act(n)$  が大きいと発生符号量が増え、画質が粗くなり、画質向上の必要性が増すとともに、音声の余剰符号量の追加による画質の変動が目立ちにくくなるからである。

【0050】

実施例3では、図1に示す映像符号化装置10において、アクティビティ・シーンチェンジ解析部18が設けられ、アクティビティ・シーンチェンジ解析部18は、映像符号化より1GOP分先行して、映像のアクティビティ解析を行い、アクティビティ平均値  $act(n)$  を事前に算出する。ここで、アクティビティとは、符号化対象領域の画素値分布の特徴を示すものであり、画素値の変動具合を表す値である。例えばアクティビティは、マクロブロックをさらに分割した小ブロックの分散値の最小値などで定義される。小ブロックの分散値とは、小ブロック内の輝度値の平均値とその小ブロック内の各画素の輝度値との差分の絶対値の総和を取ったものである。

10

【0051】

実施例3では、図2に示すフローチャートのうち、ステップS102～S104が、図4のステップS302～S304に置き換わり、他の部分については同様であるので、以下では、ステップS302～S304の部分についてだけ説明する。

【0052】

図2に示すステップS101を実行した後、映像GOP目標符号量決定部13の条件判定部132は、これから符号化するGOPのアクティビティ平均値  $act(n)$  が、ある閾値  $TH_{act}$  以上かどうかを判定する(ステップS302)。GOP目標符号量算出部133は、アクティビティ平均値  $act(n)$  が閾値  $TH_{act}$  以上の場合には、余剰符号量  $G_r(n)$  を超えない範囲で、アクティビティ平均値  $act(n)$  の閾値  $TH_{act}$  超過分に比例して、追加GOP目標符号量  $G_d(n)$  を決定する(ステップS303)。この場合の  $G_d(n)$  は、次式によって求められる。

20

【0053】

$$G_d(n) = \min(G_r(n), a \cdot (act(n) - TH_{act}))$$

$\min()$  は、最小値を返す関数であり、 $a$  は予め定められた定数である。

【0054】

一方、アクティビティ平均値  $act(n)$  が閾値  $TH_{act}$  以上でない場合には、追加GOP目標符号量  $G_d(n)$  を0とする(ステップS304)。

30

【0055】

以降の処理は、実施例1における図2のステップS105～S111と同様である。

【0056】

〔実施例4〕

図5は、本発明の実施例4の映像符号化処理のフローチャートである。実施例4では、これから符号化するGOPにシーンチェンジがあった場合に、音声の余剰符号量を映像符号化に利用する。これは、これから符号化するGOPにシーンチェンジがあると、シーンチェンジ直後のフレームにおいて動き補償が当たらず、同じ画質を保つためには、通常より多くの符号量を必要とするためである。

40

【0057】

実施例4では、図1に示す映像符号化装置10において、アクティビティ・シーンチェンジ解析部18が設けられ、アクティビティ・シーンチェンジ解析部18は、映像符号化より1GOP分先行して、映像のアクティビティ解析およびシーンチェンジ解析を行い、シーンチェンジの有無と、シーンチェンジがある場合のシーンチェンジフレームのアクティビティ  $act_{sc}$  を事前に求める。

【0058】

シーンチェンジの検出方法は、例えば次の参考文献2に記載されているような方法など、従来から種々の方法が知られているので、ここでの詳しい説明は省略する。

50

[参考文献2]大辻,外村,“映像カット自動検出方式の検討”,社団法人映像情報メディア学会,テレビジョン学会技術報告ITEJ Technical Report, Vol.16, No.43(19920710), pp.7-12

実施例4では,図2に示すフローチャートのうち,ステップS102~S104が,図5のステップS402~S404に置き換わり,他の部分については同様であるので,以下では,ステップS402~S404の部分についてだけ説明する。

【0059】

図2に示すステップS101を実行した後,映像GOP目標符号量決定部13の条件判定部132は,アクティビティ・シーンチェンジ解析部18による解析結果から,これから符号化するGOPにシーンチェンジがあるかどうかを判定する(ステップS402)。GOP目標符号量算出部133は,シーンチェンジがある場合には,余剰符号量 $G_r(n)$ を超えない範囲で,シーンチェンジフレームのアクティビティ $act_{sc}$ に比例して,追加GOP目標符号量 $G_d(n)$ を決定する(ステップS403)。この場合の $G_d(n)$ は,次式によって求められる。

【0060】

$$G_d(n) = \min(G_r(n), a \cdot act_{sc} + b)$$

$\min()$ は,最小値を返す関数であり, $a, b$ は予め定められた定数である。

【0061】

一方,これから符号化するGOPにシーンチェンジがない場合には,追加GOP目標符号量 $G_d(n)$ を0とする(ステップS404)。

【0062】

以降の処理は,実施例1における図2のステップS105~S111と同様である。

【0063】

〔実施例5〕

図6は,本発明の実施例5の映像符号化処理のフローチャートである。実施例5では,過去のいくつかの数の映像フレームの目標符号量と発生符号量との差の総和 $B$ が,ある閾値 $TH_B$ 以上の場合に,音声の余剰符号量を映像符号化に利用する。これは,目標符号量と発生符号量との差が大きい状態が続く場合には,符号量が不足していると考えられるためである。

【0064】

実施例5では,図2に示すフローチャートのうち,ステップS102~S104が,図6のステップS502~S504に置き換わり,他の部分については同様であるので,以下では,ステップS502~S504の部分についてだけ説明する。

【0065】

図2に示すステップS101を実行した後,映像GOP目標符号量決定部13の条件判定部132は,過去の数映像フレームの目標符号量と実際の発生符号量との差の総和 $B$ が,ある閾値 $TH_B$ 以上かどうかを判定する(ステップS502)。GOP目標符号量算出部133は,差の総和 $B$ が閾値 $TH_q$ 以上の場合には,余剰符号量 $G_r(n)$ を超えない範囲で,閾値 $TH_B$ 超過分に比例して,追加GOP目標符号量 $G_d(n)$ を決定する(ステップS503)。この場合の $G_d(n)$ は,次式によって求められる。

【0066】

$$G_d(n) = \min(G_r(n), a \cdot (B - TH_B))$$

$\min()$ は,最小値を返す関数であり, $a$ は予め定められた定数である。

【0067】

一方,差の総和 $B$ が閾値 $TH_B$ 以上でない場合には,追加GOP目標符号量 $G_d(n)$ を0とする(ステップS504)。

【0068】

以降の処理は,実施例1における図2のステップS105~S111と同様である。

【0069】

〔実施例6〕

10

20

30

40

50

図7は、本発明の実施例6の映像符号化処理のフローチャートである。実施例6と、前述した実施例1～5との違いは、余剰符号量算出部131におけるGOP(n)期間の余剰符号量 $G_r(n)$ の算出方法である。

【0070】

実施例6においても、映像と音声の符号化は並行して行われる。ただし、実施例1～5では、GOP(n)の映像フレームと同一表示期間の音声フレームの発生符号量をもとにGOP(n)期間の余剰符号量 $G_r(n)$ を算出するのに対し、実施例6では、GOP(n)の最後の映像フレームよりA個の映像フレーム分だけ前の期間と、同じ長さの同一表示期間の音声フレームの発生符号量をもとに、GOP(n)期間の余剰符号量 $G_r(n)$ を算出する。ここで、Aは予め定められた非負の整数である。

10

【0071】

したがって、本実施例では、GOP(n)の映像フレーム符号化開始より前に、音声符号化装置20により、このGOP(n)の最後の映像フレームのA映像フレーム前の期間(長さはGOP(n)期間と同じ)と同一表示期間の音声フレームの符号化が完了しているものとする。

【0072】

実施例6は、GOP(n)の映像符号化期間と音声発生符号量算出期間との時間的なずれにより、前述した実施例と比べて余剰符号量 $G_r(n)$ の算出精度が多少悪くなる場合があるが、音声を先行して符号化する期間を短くすることが可能になるため、符号化の遅延時間を前述した実施例に比べて短縮することができるようになる。

20

【0073】

実施例6では、図2に示すフローチャートのうち、ステップS101が、図7のステップS601に置き換わる。

【0074】

ステップS601では、映像GOP目標符号量決定部13の余剰符号量算出部131が、音声符号化装置20から通知された音声発生符号量の情報をもとに、GOP(n)の映像フレーム期間よりA映像フレーム前の同一長期間と同じ表示期間の音声フレーム $j \sim j + M - 1$ (Mは前記表示期間の音声フレーム数)について、GOP(n)期間の余剰符号量 $G_r(n)$ を、次式によって求める。

【0075】

$$G_r(n) = \sum_{x=j}^{j+M-1} \{ S_{amax} - S_a(x) \}$$

30

ここで、 $S_{amax}$ は、各音声フレームの最大発生符号量、 $S_a(x)$ は、x番目の音声フレームの発生符号量である。

【0076】

以降の処理は、実施例1における図2のステップS102～S111と同様であるが、余剰符号量利用条件の判定を、前述した実施例2～5の方法を用いて実施してもよい。この場合、図2のステップS102～S104が、実施例2の場合には図3のステップS202～S204に、実施例3の場合には図4のステップS302～S304に、実施例4の場合には図5のステップS402～S404に、実施例5の場合には図6のステップS502～S504に、さらに置き換わることになる。

40

【0077】

〔実施例7〕

図8は、本発明の実施例7の映像符号化処理のフローチャートである。実施例7と、前述した実施例1～6との違いは、前述した実施例1～6では、音声の余剰符号量を映像符号化に利用していたのに対し、実施例7では、符号化ストリームに重畳する映像ストリーム以外のデータストリームの余剰符号量を、余剰符号量算出部131におけるGOP(n)期間の余剰符号量 $G_r(n)$ として算出している点である。

【0078】

本実施例では、GOP(n)の映像フレーム符号化開始より前に、GOP(n)の映像フレームと同一デコード時刻の符号化ストリーム中に多重化するデータのデータビット量

50

が判明しているものとする。この場合の基準映像ビットレートは、データフレームのビット量が常に最大の場合に確保できる映像ビットレートを指す。

【0079】

実施例7では、図2に示すフローチャートのうち、ステップS101が、図8のステップS701に置き換わる。

【0080】

ステップS701では、映像GOP目標符号量決定部13の余剰符号量算出部131が、GOP(n)の映像フレームのデコード時刻と同一時刻に受信側で必要とされるデータフレームj~j+M-1(MはGOP(n)と同一期間のデータフレーム数)について、GOP(n)期間の余剰符号量 $G_r(n)$ を、次式によって求める。

10

【0081】

$$G_r(n) = \sum_{x=j}^{j+M-1} \{ S_{dmax} - S_d(x) \}$$

ここで、 $S_{dmax}$ は、各データフレームの最大発生符号量、 $S_d(x)$ は、x番目のデータフレームの発生符号量である。

【0082】

以降の処理は、実施例1における図2のステップS102~S111と同様であるが、実施例6の場合と同じように、余剰符号量利用条件の判定を、実施例2~5の方法を用いて行ってもよい。

【0083】

以上の映像符号化の処理は、コンピュータとソフトウェアプログラムとによって実現することができ、そのプログラムをコンピュータ読み取り可能な記録媒体に記録することも、ネットワークを通して提供することも可能である。

20

【図面の簡単な説明】

【0084】

【図1】本発明の実施例に係る装置の構成例を示す図である。

【図2】本発明の実施例1の映像符号化処理のフローチャートである。

【図3】本発明の実施例2の映像符号化処理のフローチャートである。

【図4】本発明の実施例3の映像符号化処理のフローチャートである。

【図5】本発明の実施例4の映像符号化処理のフローチャートである。

【図6】本発明の実施例5の映像符号化処理のフローチャートである。

30

【図7】本発明の実施例6の映像符号化処理のフローチャートである。

【図8】本発明の実施例7の映像符号化処理のフローチャートである。

【図9】本発明の課題を説明する図である。

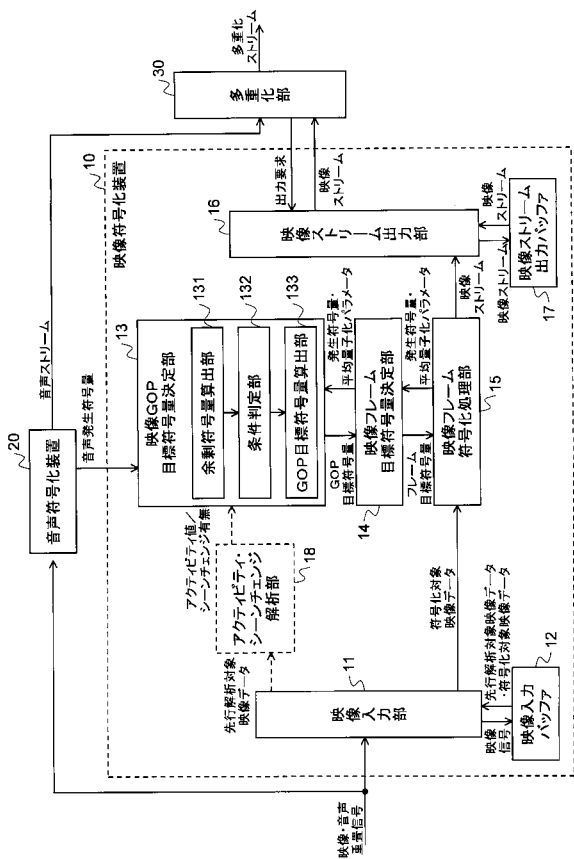
【符号の説明】

【0085】

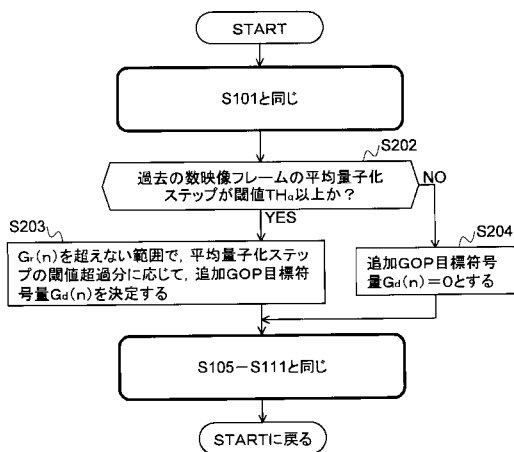
- 10 映像符号化装置
- 11 映像入力部
- 12 映像入力バッファ
- 13 映像GOP目標符号化量決定部
- 131 余剰符号量算出部
- 132 条件判定部
- 133 GOP目標符号量算出部
- 14 映像フレーム目標符号量決定部
- 15 映像フレーム符号化処理部
- 16 映像ストリーム出力部
- 17 映像ストリーム出力バッファ
- 18 アクティビティ・シーンチェンジ解析部
- 20 音声符号化装置
- 30 多重化部

40

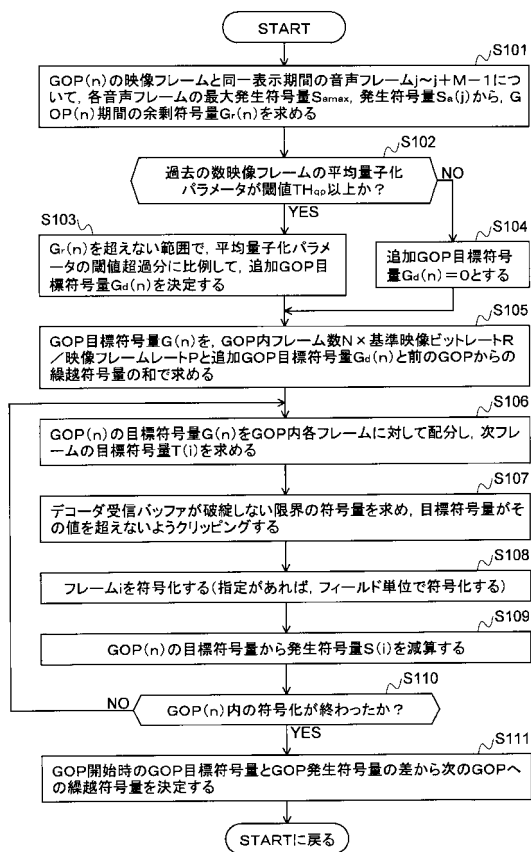
【図 1】



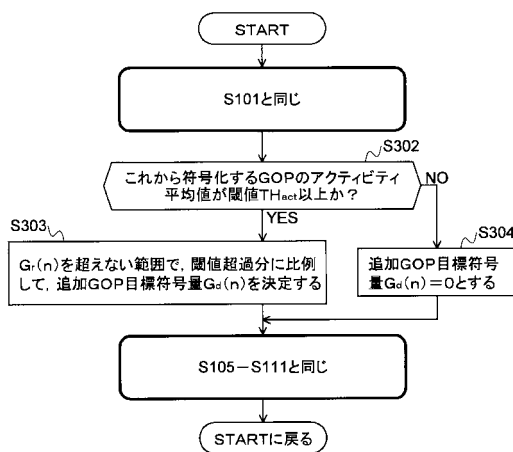
【図 3】



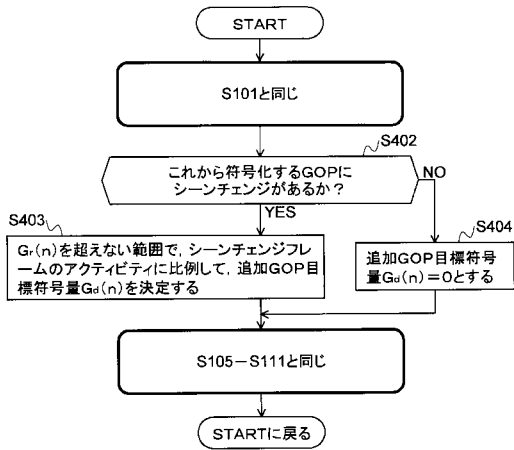
【図 2】



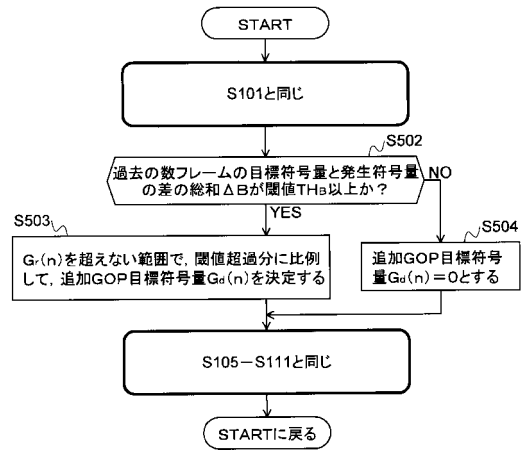
【図 4】



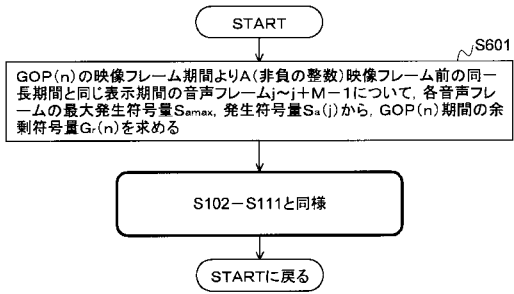
【 図 5 】



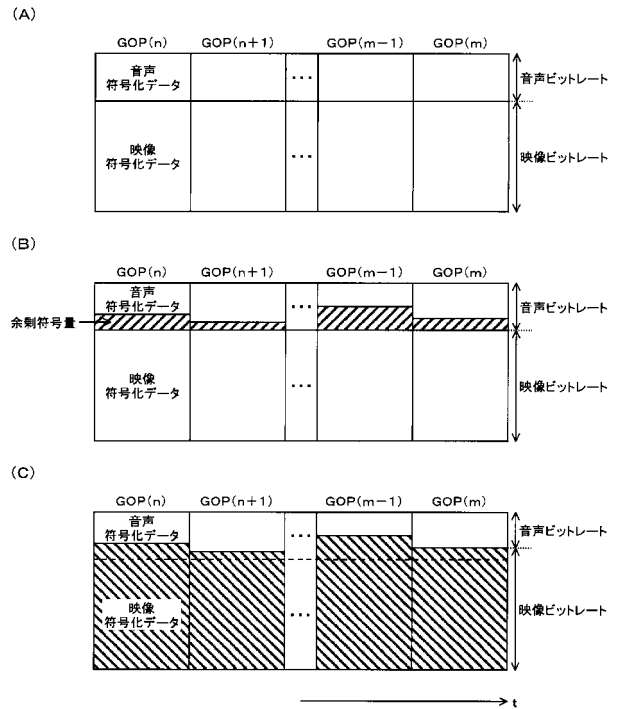
【 図 6 】



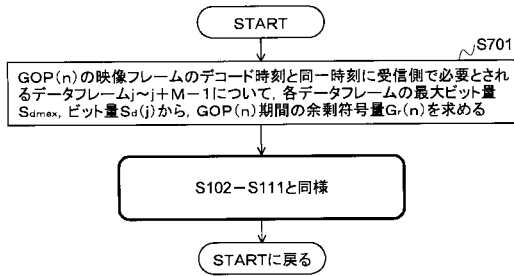
【 図 7 】



【 図 9 】



【 図 8 】



---

フロントページの続き

(72)発明者 原田 登

東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内

Fターム(参考) 5C059 RC32 TA60 TC00 TC10 TC14 TC18 TC38 TD03 TD05 TD12

UA02

5C159 RC32 TA60 TC00 TC10 TC14 TC18 TC38 TD03 TD05 TD12

UA02