(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第5227193号
(P5227193)

(45)発行日　平成25年7月3日(2013.7.3)　(24)登録日　平成25年3月22日(2013.3.22)

(51)Int.Cl.			F I		
H04N	***7/32***	***(2006.01)***	H04N	7/137	A
H03M	***7/30***	***(2006.01)***	H03M	7/30	A

請求項の数 10　(全 17 頁)

(21)出願番号	特願2008-555755 (P2008-555755)
(86)(22)出願日	平成19年2月14日 (2007.2.14)
(65)公表番号	特表2009-528715 (P2009-528715A)
(43)公表日	平成21年8月6日 (2009.8.6)
(86)国際出願番号	PCT/EP2007/051451
(87)国際公開番号	WO2007/096288
(87)国際公開日	平成19年8月30日 (2007.8.30)
審査請求日	平成22年2月5日 (2010.2.5)
(31)優先権主張番号	06300164.8
(32)優先日	平成18年2月27日 (2006.2.27)
(33)優先権主張国	欧州特許庁 (EP)

(73)特許権者　501263810
トムソン　ライセンシング
Ｔｈｏｍｓｏｎ　Ｌｉｃｅｎｓｉｎｇ
フランス国，　９２１３０　イツシー　レ　ムーリノー，　ル　ジヤンヌ　ダルク，　１－５
１－５，　ｒｕｅ　Ｊｅａｎｎｅ　ｄ’Ａｒｃ，　９２１３０　ＩＳＳＹ　ＬＥＳ　ＭＯＵＬＩＮＥＡＵＸ，　Ｆｒａｎｃｅ
(74)代理人　100070150
弁理士　伊東　忠彦
(74)代理人　100091214
弁理士　大貫　進介
(74)代理人　100107766
弁理士　伊東　忠重

最終頁に続く

(54)【発明の名称】ＳＶＣデコーダにおけるパケットロス検出および仮想パケット生成のための方法および装置

(57)【特許請求の範囲】

【請求項１】

受信されたパケット化ビデオ・データにおける誤りを検出する方法であって、前記ビデオ・データは基本層と少なくとも一つの向上層とを含み、各層は、それぞれのヘッダおよびピクチャー・グループ（GOP）構造をもって、複数のフレームまたはスライスに編成されており、前記GOP構造が複数の完全なGOPおよび一つの不完全なGOPを含み、完全なGOPはエンコード前にあらかじめ決定されている事前固定された一定長をもち、前記不完全なGOPは完全なGOPより少ないピクチャーまたはフレームを有しており、不完全なGOPについてはGOPサイズ、参照フレーム数および不完全なGOPの時間的レベルについてのパラメータが実際のGOPサイズに適応されており、

当該方法が、

・ある層に関係するヘッダ情報を読むステップと；

・前記ヘッダ情報からスキームを生成するか、前記ヘッダ情報に従ってあらかじめ定義されたスキームを選択し、ここで前記スキームは、各層のGOP内の諸フレームについての少なくとも表示順番号（POC）、参照フレーム数（frame_num）および時間的レベルを含み、時間的レベルとは時間的向上層の階層レベルを指す、ステップと；

・受信したビデオ・データ・フレームから表示順番号（POC）、参照フレーム数（frame_num）および時間的レベルを含むパラメータを抽出するステップと；

・一つまたは複数のフレームまたはパケットが欠けていることを検出し、ここで、受信されたパラメータが前記スキームに基づいて期待されるフレームのパラメータと比較され、

受信されたパラメータと期待されるパラメータとの間の不一致によって欠けているパケットまたはフレームが判別される、ステップと；
・欠けているフレームについて、それが完全なGOPに属するか不完全なGOPに属するかを判定するステップと、ここで、少なくとも一つのその後受信されるパケットが完全なGOPのその後期待されるパケットと比較され、前記少なくとも一つのその後受信されるパケットが完全なGOPのフレーム・シーケンスにおいて期待されていない場合には前記欠けているフレームは不完全なGOPに割り当てられ；
・不完全なGOPに属する一つまたは複数の欠けているフレームの場合、前記不完全なGOPの最大GOPサイズを推定し、前記推定されたGOPサイズまでの表示順番号（POC）をもつ、前記不完全なGOPの欠けているフレームについては、デフォルト・パケットを生成し、ここで、各デフォルト・パケットは各期待されるパケットの表示順番号、参照フレーム数および時間的レベルを与えられ、前記推定されたGOPサイズを超える表示順番号（POC）をもつ欠けているフレームについてはデフォルト・パケットを生成しないステップとを有する、
方法。
【請求項２】
・欠けているパケットまたはフレームのそれぞれについて、各期待されるパケットの表示順番号（POC）、参照フレーム数（frame_num）および時間的レベルを与えられたデフォルト・パケットを生成するステップと；
・欠けているパケットのそれぞれについて、生成されたデフォルト・パケットを、データ・ストリームがパースされる前にデータ・ストリーム中に挿入するステップとをさらに有する、
請求項１記載の方法。
【請求項３】
デフォルト・パケットを挿入する前記ステップが、完全なGOPに属するすべての欠けているフレームについても実行される、請求項１または２記載の方法。
【請求項４】
請求項１ないし３のうちいずれか一項記載の方法であって、不完全なGOPの最大GOPサイズを推定する前記ステップが、
・奇数の表示順番号（POC）をもつ受信フレーム（f_c）であってframe_numパラメータ値が$f_{c,frame_num}$であるフレーム（f_c）を選択するステップと；
・前記GOPサイズについての上限を、
$GopSize_{abn,max}=((f_{c,frame_num}-1)\%(GopSize>>1)<<1)+1$
に従って計算するステップとを有しており、
前記の計算された上限が推定された最大GOPサイズとして使われる、
方法。
【請求項５】
前記ヘッダ情報が、所定の空間的層の相続くフレームの表示順番号間の差の指標を含んでいる、請求項１ないし４のうちいずれか一項記載の方法。
【請求項６】
前記指標が二つの相続いて表示されるピクチャーの間の表示順番号の増加を計算するために使われる、請求項５記載の方法。
【請求項７】
受信されたパケット化ビデオ・データにおける誤りを検出する装置であって、前記ビデオ・データは基本層および少なくとも一つの向上層を含み、各層はそれぞれのヘッダおよびピクチャー・グループ（GOP）構造をもって複数のフレームまたはスライスに編成されており、前記GOP構造が複数の完全なGOPおよび一つの不完全なGOPを含み、完全なGOPはエンコード前にあらかじめ決定されている事前固定された一定長をもち、前記不完全なGOPは完全なGOPより少ないピクチャーまたはフレームを有しており、不完全なGOPについてはGOPサイズ、参照フレーム数（frame_num）および不完全なGOPの時間的レベルについてのパラメータが実際のGOPサイズに適応されており、

当該装置は、
・ある層に関係するヘッダ情報を読む手段と；
・前記ヘッダ情報からスキームを生成するまたは前記ヘッダ情報に従ってあらかじめ定義されたスキームを選択する手段であって、ここで前記スキームは各層のGOP内の諸フレームについての少なくとも表示順番号（POC）、参照フレーム数（frame_num）および時間的レベルを含み、時間的レベルとは時間的向上層の階層レベルを指す、手段と；
・受信したビデオ・データ・フレームから、表示順番号（POC）、参照フレーム数（frame_num）および時間的レベルを含むパラメータを抽出する手段と；
・一つまたは複数のフレームまたはパケットが欠けていることを検出する手段であって、受信されたパラメータが前記スキームに基づいて期待されるフレームのパラメータと比較され、受信されたパラメータと期待されるパラメータとの間の不一致によって欠けているパケットまたはフレームが判別される、手段と；
・欠けているフレームについて、それが完全なGOPに属するか不完全なGOPに属するかを判定する手段と、ここで、少なくとも一つのその後受信されるパケットが完全なGOPのその後期待されるパケットと比較され、前記少なくとも一つのその後受信されるパケットが完全なGOPのフレーム・シーケンスにおいて期待されていない場合には前記欠けているフレームは不完全なGOPに割り当てられ；
・不完全なGOPの場合、前記不完全なGOPの最大GOPサイズを推定する手段と；
・前記推定されたGOPサイズまでの表示順番号（POC）をもつ、前記不完全なGOPの欠けているフレームについては、デフォルト・パケットを生成し、ここで、各デフォルト・パケットは各期待されるパケットの表示順番号、参照フレーム数および時間的レベルを与えられ、不完全なGOPについて、前記推定されたGOPサイズを超える表示順番号（POC）をもつ欠けているフレームについてはデフォルト・パケットが生成されない、手段とを有する、装置。
【請求項８】
・欠けているパケットまたはフレームのそれぞれについて、各期待されるパケットの表示順番号、参照フレーム数および時間的レベルを与えられたデフォルト・パケットを生成する手段と；
・欠けているパケットのそれぞれについて、生成されたデフォルト・パケットを、データ・ストリームがパースされる前にデータ・ストリーム中に挿入する手段とをさらに有する、
請求項７記載の装置。
【請求項９】
デフォルト・パケットの前記挿入が、完全なGOPに属するすべての欠けているフレームについても実行される、請求項７または８記載の装置。
【請求項１０】
前記ヘッダ情報が、所定の空間的層の相続くフレームの表示順番号間の差の指標を含んでいる、請求項７ないし９のうちいずれか一項記載の装置。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、スケーラブル・ビデオ符号化（SVC: scalable video coding）に関する。特に、本発明は、SVCビットストリーム用デコーダにおけるパケットロス検出および仮想パケット生成のための方法および装置に関する。
【背景技術】
【０００２】
誤り隠蔽（error concealment）の諸方法がJSVM4（非特許文献１）に導入されている。例示的に、JVT-Q046（非特許文献２）には４つの方法が記載されている。誤り隠蔽のためのこの既知の解決策は、固定GOP（ピクチャー・グループ［group-of-pictures］）サイズを対象とし、基本層はH.264準拠で、空間的およびFGS（細粒度スケーラビリティ［Fine

Grain Scalability］）スケーラブル層をもつ。SVCにおける時間的スケーラビリティは、階層的なBフレームGOP構造によってサポートされる。この構造では、GOP内の複数の双方向予測されるフレーム（Bフレーム）は、所定のスキームに従った異なる時間的レベル（temporal level）をもつ。しかしながら、誤り隠蔽の堅牢さおよび正確さをさらに改善する必要がある。さらに、任意の種類のスケーラビリティおよび組み合わされたスケーラビリティについての誤り隠蔽をサポートすることが望ましい。

【非特許文献１】“Scalable Video Coding―Joint Draft 4”, J. Reichel, H. Schwarz, M. Wien, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6), Draft Output Document from JVT JVT-Q202-JSVM4, 17th Meeting: Nice, France, October, 2005

【非特許文献２】Chen Ying, Jill Boyce, Xie Kai “Frame Loss Error Concealment for SVC”, JVT-Q046, JVT of ISO/IEC MPEG & ITU-T VCEG, Nice, October 2005

【発明の開示】

【発明が解決しようとする課題】

【０００３】

本発明は、受信されたパケット化ビデオ・データにおける誤りを検出する方法を提供する。ここで、ビデオ・データは空間的または空間時間的な基本層（BL: base-layer）および少なくとも一つの向上層（EL: enhancement-layer）を含み、各層はそれぞれヘッダおよびGOP構造をもって複数のフレームまたはスライスに編成されており、諸層のうちには異なる時間的レベルが存在する。パケットは、一つのフレームまたはスライスを含むと想定される。

【課題を解決するための手段】

【０００４】

本発明に基づく方法は、ある層に関係するヘッダ情報を読み／抽出し、前記ヘッダ情報から、少なくとも表示順番号（すなわちピクチャー順カウント［picture order count］、POC）、参照フレーム数（すなわちframe_num）および各層の完全なGOP内の諸フレームについての時間的レベルを含むスキームを生成し、受信したビデオ・データ・フレームからPOC番号、frame_numおよび時間的レベルを含むパラメータを抽出し、一つまたは複数のフレームまたはパケットが欠けていることを検出するステップを有しており、ここで、受信されたパラメータが前記スキームに従って期待されるフレームのパラメータと比較され、受信されたパラメータと期待されるパラメータとの間の不一致によって欠けているパケットまたはフレームが判別される。

【０００５】

前記ヘッダ情報からスキームを生成する代わりに、あらかじめ定義された複数のスキームのうちから前記ヘッダ情報に従って一つのあらかじめ定義されたスキームを選択することも可能である。そのようなあらかじめ定義されたスキームは保存されていてもよいし、あるいは適切なアルゴリズムを使ってオンザフライで生成されてもよい。

【０００６】

本発明のある実施形態では、本方法はさらに、
欠けているパケットまたはフレームのそれぞれについて、各期待されるパケットの表示順番号（POC）、参照フレーム数（frame_num）および時間的レベルを与えられたデフォルト・パケットを生成し、
欠けているパケットのそれぞれについて、生成されたデフォルト・パケットを、データ・ストリームがパースされる（parsed）前にデータ・ストリーム中に挿入する、
ステップを有する。

【０００７】

ある実施形態では、GOP構造は複数の完全なGOPおよび高々一つの不完全GOP（これは、他よりも少ないピクチャー／フレームをもつGOPである）を含む。本発明に基づく方法はさらに、欠けているフレームについて、それが完全なGOPに属するか不完全なGOPに属する

かを判定するステップを有する。ある実施形態では、デフォルト・パケットを挿入する前記ステップは、完全なGOPに属する欠けているフレームについてのみ実行される。

【０００８】

前記実施形態における欠けているフレームについて、それが完全なGOPに属するか不完全なGOPに属するかを判定する前記ステップは、前記不完全なGOPについての最大GOPサイズを推定し、推定されたGOPサイズを超える表示順番号（POC）については不完全なGOPについてのデフォルト・パケットを生成しないまたは抑制するステップを有することにおいてさらに特定されることができる。

【０００９】

ある実施形態では、本発明に基づく方法は、前記ヘッダ情報が、所定の空間時間的または空間的層の相続くフレームのPOC番号間の差の指標を含んでいることにおいてさらに特定されることができる。

【００１０】

前記実施形態は、前記所定の空間時間的または空間的層がBLであることにおいてさらに特定されることができる。

【００１１】

ある実施形態では、前記指標は前記差の2を底とした対数である。

【００１２】

さらに、本発明は、パケット化ビデオ・データを含む信号を提供する。ここで、ビデオ・データはフレームまたはスライスに編成されており、それぞれのヘッダおよび複数のGOPを含むピクチャー・グループ（GOP）構造をもち、各ビデオ・フレームまたはスライスは、少なくとも、そのPOC番号の指標をもち、前記ヘッダは、相続くフレームの表示順番号（POC）の間の差の指標（log2_min_poc_increase）を含んでいる。

【００１３】

前記ビデオ信号は、空間的または空間時間的基本層および少なくとも一つの向上層をもつということにおいてさらに特定されることができる。相続くフレームの表示順番号（POC）の間の差の前記指標は、それらの層のうちのある所定の層の諸フレームのことをいっている。

【００１４】

さらに、本発明はそのような信号を生成するための装置を提供する。

【００１５】

受信されたパケット化ビデオ・データにおける誤りを検出する装置であって、ビデオ・データは基本層および少なくとも一つの向上層を含み、各層はそれぞれヘッダおよびピクチャー・グループ（GOP）構造をもって複数のフレームまたはスライスに編成されており、本装置は、ある層に関係するヘッダ情報を読み、抽出する手段と、前記の抽出されたヘッダ情報からスキームを生成する手段または前記の抽出されたヘッダ情報（SEI）に従って一つのあらかじめ定義されたスキームを選択する手段であって、ここで前記スキームは少なくとも表示順番号（POC）、参照フレーム数（frame_num）および各層の完全なGOP内の諸フレームについての時間的レベルを含む、手段と、受信したビデオ・データ・フレームからPOC番号、frame_numおよび時間的レベルを含むパラメータを抽出する手段と、一つまたは複数のフレームまたはパケットが欠けていることを検出する手段とを有しており、ここで、受信されたパラメータが前記スキームに従って期待されるフレームのパラメータと比較され、受信されたパラメータと期待されるパラメータとの間の不一致によって欠けているパケットまたはフレームが判別される。

【００１６】

本発明のある実施形態では、本装置はさらに、欠けているパケットまたはフレームのそれぞれについて、各期待されるパケットの表示順番号（POC）、参照フレーム数（frame_num）および時間的レベルを与えられたデフォルト・パケットを生成する手段と、
欠けているパケットのそれぞれについて、生成されたデフォルト・パケットを、データ・ストリームがパースされる（parsed）前にデータ・ストリーム中に挿入する手段とを有す

る。
【００１７】
ある実施形態では、GOP構造は複数の完全なGOPおよび高々一つの不完全GOPを含む。本発明に基づく装置はさらに、欠けているフレームについて、それが完全なGOPに属するか不完全なGOPに属するかを判定する手段と、完全なGOPに属する欠けているフレームについてのみデフォルト・パケットの前記挿入を実行する手段とを有する。
【００１８】
本発明の有利な実施形態は、従属請求項、以下の記述および図面において開示される。
【００１９】
本発明の例示的な実施形態について、付属の図面を参照しつつ述べる。
【発明を実施するための最良の形態】
【００２０】
図１は、SVCビットストリームのNAL（network abstraction layer［ネットワーク抽象化層］）パケット構造を示している。パケットのシーケンスは、補足向上情報（SEI: Supplementary Enhancement Information）、シーケンス・パラメータ・セット（SPS: sequence parameter set）、スケーラブルな拡張におけるシーケンス・パラメータ・セット（SPS_S）、ピクチャー・パラメータ・セット（PPS: picture parameter set）、スケーラブルな拡張におけるピクチャー・パラメータ・セット（PPS_S）、IDRピクチャーの符号化されたスライス（SLICE_IDR）およびスケーラブルな拡張におけるIDRピクチャーの符号化されたスライス（SLICE_IDR_S）を含む。一般に、ピクチャーは二つ以上のスライスから構築されてもよいが、ここでは例示的に、ピクチャーは一枚のスライスを含んでいると想定する。ピクチャー・データをもつ各NALパケットは、スライスのデータを含み、よってピクチャーのデータを含む。IDRピクチャーの符号化されたスライス（SLICE_IDR）をもつパケットはIDRピクチャー（IDR）を含み、これがキー・ピクチャーのはたらきをする。
【００２１】
これらのパケットには、いくつかのピクチャー・グループ（GOP）が続く。ここで、各ピクチャーは（例示的に）別個のパケットに入れられている。ピクチャーは通常の非IDRピクチャー（SLICE_NOIDR）またはスケーラブルな拡張における非IDRピクチャーピクチャー（SLICE_NOIDR_S）である。図１の例では、ビットストリームは二つの空間的層をもち、GOPサイズは16である。よって、1ライン（16個のBLパケット＋16個のELパケット）が完全なGOPをなす。ストリーム中の完全なGOPの数は、たとえば数千など、非常に大きいこともありうる。各GOP中の最後のピクチャーはキー・ピクチャーとしてはたらくIフレームまたはPフレーム（イントラ符号化されるまたは予測される）であり、他のピクチャーはBフレーム（双方向予測される）である。
【００２２】
図２は、異なるサイズの複数の階層的なBフレームGOP内での時間的レベルを示している。最低の時間的レベルTL0上のフレームはキー・ピクチャーと呼ばれ、IフレームまたはPフレームである。その一つ（0と記される）は実際には直前のGOPに属する。一般に、TL0より高い時間的レベルのフレームは二つ以上の参照フレームから双方向予測される。参照フレームは、より低い時間的レベル上のフレームまたは（最高の時間的レベルにあるのではないフレームについては）同じ時間的レベルの以前にデコードされたフレームであることができる。時間的レベルは昇順にデコードされる。すなわちまずTL0、次いでTL1、次いでTL2といった具合である。
【００２３】
GopSize＝2のGOPでは、図２ａ）に示されるように、二つの時間的レベルTL0、TL1しか存在せず、一つのフレーム１がより高い時間的レベルTL1に、他のフレーム２が最低の時間的レベルTL0にある。フレーム１は、先にデコードされるその参照フレーム０および２から予測される。
【００２４】

GopSize＝4のGOPでは（図２ｂ））、三つの時間的レベルTL0、TL1、TL2が存在し、一つのフレーム４が最低の時間的レベルTL0に、一つのフレーム２が第一の時間的レベルTL1に、二つのフレーム１、３が最高レベルTL2にある。まず、フレーム２がその参照フレーム０および４から予測され、次いでフレーム１および３がその参照フレーム０、２および４から予測される。TL2が最高の時間的レベルなので、その同じ参照フレームが両方に使われる。

【００２５】

GopSize＝8のGOPでは（図２ｃ））、四つの時間的レベルTL0、…、TL3が存在する。まず、フレーム８がフレーム０から、次いでフレーム４がフレーム０および８から、次いでフレーム２がフレーム０、４、８から、次いでフレーム６がフレーム０、２、４、８から、次いでフレーム１、３、５、７がフレーム０、２、４、６、８から予測される。

【００２６】

たとえば図２ｄ）のGopSize＝16のようなより大きなGopSizeをもつGOPも、対応して構築される。

【００２７】

本発明は、パケットロス検出および仮想フレーム生成を、特に非IDRピクチャー（図１ではSlice_NOIDR／SLICE_NOIDR_Sと示されている）をもつパケットについて提供する。SEIメッセージ、SPS、PPSおよび最初のIDRピクチャーに属するパケットの送信は保証されていると想定する。本発明によれば、パケットロス検出の第一の段階に、紛失パケット解析、仮想パケットの生成および挿入、パースおよびデコードが続く。

【００２８】

パケットロス検出は、picture_order_count（POC）値、frame_num値および時間的レベル値を使う。POCは、図２ａ）～ｄ）で使われるような、諸フレームの表示順の番号付けである。frame_numは、フレームがもつ（潜在的な）参照フレームの数を示す、フレームのパラメータである。パケットロス検出のためには、二つの相次いで表示されるピクチャーの間のPOC増も重要である。それをここではPOC_Increaseと呼ぶ。通例、最高の空間的層について、POC_Increaseが1、すなわちすべてのフレームが表示されると想定する。しかしながら、時にこれが真でないこともある。たとえば、入力フレームレートが30fpsで空間的基本層がたった15fpsでエンコードされる場合、二つに一つのピクチャーしかエンコードされないので、基本層のPOC_Increaseは2である。

【００２９】

デコーダでは、POC_IncreaseはSEIメッセージからは知られない。これはパケットロス検出について問題を引き起こすであろう。最初、パーサーが二つに一つのパケットが紛失していると検出しうる。若干のパースのあと、奇数POC値をもつパケットがみつかったことが認識されることがあり、よって、POC_Increaseが2であると結論されることがありうるが、遅すぎるかもしれない。すなわち、バッファがオーバーフローするかもしれず、他の問題が生じることがありうる。これは最適な解決策ではない。この問題を克服するため、本発明の一つの側面は、SEIメッセージに最小のPOC_Increaseの指標log2_min_poc_increaseを加えることである。これは例示的に表１に示され、後述される。

【００３０】

【表１】

scalability_info(payloadSize) {	C	Descriptor
num_layers_minus1	5	ue(v)
log2_min_poc_increase	5	ue(v)
for (i = 0; i <= num_layers_minus1; i++) {		
layer_id[i]	5	u(8)
…		
…		
}		
}		

表１：SEIパケットへのlog2_min_poc_increaseの挿入

同じ依存性レベル（空間解像度層）に属するすべての層の最大temporal_levelをMTLと定義すると、POC_Increaseは、最大のMTL（MTL_{max}）をもつ空間的層については、$2^{log2_min_poc_increase}$である。$MTL_{Layer}$という（より低い）最大temporal_levelをもつ別の空間的層については、POC_Increaseは

【数１】

$$POC_Increase = 2^{MTL_{max}-MTL_{Layer}}.2^{log2_min_poc_increase} \quad (式1)$$

すなわち、$POC_Increase = 2^{MTLmax-MTLlayer+log2_min_poc_increase}$である。

【００３１】

log2_min_poc_increase値が挿入されたSEIメッセージがあれば、パケットロス検出を扱うのが、上記の諸例でも、より簡単になる。たった15fps（frames per second［フレーム毎秒］）のフレームレートで基本層が符号化される場合、log2_min_poc_increaseパラメータの値は1であり、フレームレート30fpsをもつ空間的向上層も符号化される場合、log2_min_poc_increaseは0である。よって、最小のPOC_Increaseを決定する問題は、基本層しか含まないビットストリームにおけるパケットロス検出については解決される。

【００３２】

パケットロスを検出したのち、パーサーおよびデコーダのために仮想パケットが生成されることができる。これはデコーダにおいて、実際のパースの前になされる。パケットがパーサーに送られるときは常に、それが正しく受信されたパケットか仮想パケットかに関わりなく、デコードされたピクチャー数（decoded_index）はインクリメントされることになる。したがって、パーサー前に仮想パケットを挿入して、その後のパケットの番号付けが正しくなるようにすることが有利である。

【００３３】

各層のtemporal_levelはSVCスケーラビリティ情報SEIメッセージ・シンタックス（JVT_Q201参照）で定義されている。同じ依存性レベル（空間解像度層）に属するすべての層の最大temporal_levelをMTLと定義する。よって、正常なGOPについてのサイズは$GopSize = 2^{MTL}$である。正常なGOPサイズがあれば、GOP内の相対POCを構築できる。GOP内の相対POCは、階層的なBフレーム構造を使って構築される。相対POC、相対frame_numおよび時間的レベル値が下記の諸表においてデコード順に挙げられる。図２ａ）～ｄ）を参照。それらは、特定のGOPサイズについて簡単に生成できる。

【００３４】

あるいはまた、下記の諸スキームはあらかじめ定義されていることができ、受信されたGOPSizeに対応するスキームが選択される。

【００３５】

【表２】

デコード順	0	1
POC	0	1
frame_num	0	1
temporal_level	0	0

表２ａ：GopSize＝1についての重要な値

【００３６】
【表３】

GopSize	2		
デコード順	0	1	2
POC	0	2	1
frame_num	0	1	2
temporal_level	0	0	1

表２ｂ：GopSize＝2についての重要な値

【００３７】
【表４】

GopSize	4				
デコード順	0	1	2	3	4
POC	0	4	2	1	3
frame_num	0	1	2	3	3
temporal_level	0	0	1	2	2

表２ｃ：GopSize＝4についての重要な値

【００３８】
【表５】

GopSize	8								
デコード順	0	1	2	3	4	5	6	7	8
POC	0	8	4	2	6	1	3	5	7
frame_num	0	1	2	3	4	5	5	5	5
temporal_level	0	0	1	2	2	3	3	3	3

表２ｄ：GopSize＝8についての重要な値

【００３９】
【表６】

GopSize	16																
デコード順	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
POC	0	16	8	4	12	2	6	10	14	1	3	5	7	9	11	13	15
frame_num	0	1	2	3	4	5	6	7	8	9	9	9	9	9	9	9	9
temp_lvl	0	0	1	2	2	3	3	3	3	4	4	4	4	4	4	4	4

表２ｄ：GopSize＝16についての重要な値

【００４０】
【表７】

GopSize	32																
デコード順	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
POC	0	32	16	8	24	4	12	20	28	2	6	10	14	18	22	26	30
frame_num	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
temp_lvl	0	0	1	2	2	3	3	3	3	4	4	4	4	4	4	4	4

表２ｅ１：GopSize＝32についての重要な値

GopSize	32															
デコード順	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
POC	1	3	5	7	9	11	13	15	17	19	21	23	25	27	39	31
frame_num	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17
temp_lvl	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5

表２ｅ２：GopSize＝32についての重要な値（承前）

より厳密な記述のために、二つの事項が想定される必要がある。まず、実際の応用では、真のframe_numが、実際には最大フレーム数を用いた法の演算（modulo operation）によって修正されることがありうる。

【００４１】
第二に、デコード順における最初のフレームは実際には必ずしも順（order）0ではない。その代わり、通例、このフレームは、直前のGOPの最初に符号化されているので、デコード順値－(GopSize－1)をもつほうが精確である。

【００４２】
フレームまたはフレーム情報を順序付けるのは、たとえばPOCによって、デコード順によって、frame_numによってなど、実際の応用ではさまざまな基準でできる。この特徴は、パケットロス検出および仮想パケット生成のために役に立つ。

【００４３】
エンコーダにおける入力フレームレートに基づいて、二つの相続いて表示されるフレームのPOC増は1または2、あるいはさらには空間的なスケーラブルな層の数が3以上であり各層が異なるフレームレートをもつときには2^nでありうる。フレーム番号を数えるときは、POC増が考慮される必要がある。

【００４４】
frame_numは1増されるべきである（最高時間的レベルの諸フレームを除いて）。値POC/POC_Increaseが奇数であれば、そのフレームはframe_numを増加させない。それをここでは奇数フレームと呼ぶ。通例、正常なGOPでは、これらの奇数フレームが最高時間的レベルをもつ。より多くの空間的層があり、異なる層についての入力フレームレートが同じでない場合、POC_Increaseはより大きくてもよい。一般に、それは、nを現在の層のMTLと最高時間的層のMTLとの間の差として、2^nである。通例、最高の（空間的／空間時間的）層は時間的レベルの最大数をもつ。

【００４５】
たとえば、基本層（BL）および向上層（EL）をもつ「二つの空間的層」の場合、BLはELの半分のフレームレートをもつことができる。BLがたとえば5個の時間的層をもつ場合、BLのMTLは5である。例示的に、ELのMTLは6である。よって、ELについてはPOC_Increaseは1であり、BLについてはPOC_Increaseは$2^{6-5}=2$である。BLおよびELについてのGOPサイズはそれぞれ2^5および2^6である。

【００４６】
本願で特定の空間的層のPOC値が言及されるとき、POCは実際にはPOC/POC_Increaseである。パケットロス検出後およびパケット仮想化の最終ステップにおいて、POC値がPOC_Increaseを乗算してマッピングし戻されることになる。
【００４７】
エラー・フリーのビットストリームにおいて、各空間的またはFGS層はGopSize・n＋d個のパケットによって構成されることになる。n個の完全なGOPがあり、サイズdの不完全なGOPがあるかもしれない。ここで、dはGopSizeより小さい。しかしながら、「不完全」というのは単に、「完全な」GOPよりも少ないフレームを含んでいるという事実を言及するのみであり、「不完全な」GOPはシンタックス上は正しく、各フレームについての属性（後述）は実際のGOPサイズに合わせて適応されている。完全なGOPのことを正常GOPと称し、最後の不完全なGOPのことを異常GOPと称する。ビットストリーム中に異常GOPは高々一つしかない。
【００４８】
各フレームについて、表２ａ）～２ｅ）にも挙げられているように、三つの重要な属性：POC番号、frame_numおよびtemporal_levelがある。パケットロス検出および仮想パケット生成をより簡単に記述するため、フレーム情報の抽象子（abstractor）として、3D要素f＝(Poc,frame_num,temporal_level)を導入する。これらのフレームの集合は$A=\{f_i \mid i=0 \cdots GopSize-1\}$である。集合Aにおいて、GopSizeの数|A|＝GopSizeのフレームf_iの系列がある。本稿ではAのことを検出フレーム情報リストと呼ぶ。
【００４９】
Aはデコード順によって、POCによって、frame_numによって順序付けされることができ、あるいはさらには、まずtemporal_levelによって順序付けされ、同じtemporal_level内ではPOCによって順序付けされることもできる。
【００５０】
パケットのデコード後、A内の諸要素は、通例、POCをGopSizeだけ増し、frame_numをGopSize/2だけ増すことによって更新される。
【００５１】
Aがデコード順によって順序付けられているとき、集合Aおよび直前のデコードされたフレームを用いて、現在の要素f′の正しい予測が$f' = f_{(decoded_index\%GopSize)+1}$として実行できる。現在のパケットを得て、スライス・ヘッダからデコードされたPOC、frame_numおよびtemporal_level値をもつ現在のフレームf_cを生成するとき、f′およびf_cの比較をすることができる。
【００５２】
二つの仮想フレーム要素は、そのPOC値およびframe_num値が等しいときに、等しいと定義する。
【００５３】
f′がf_cに等しい場合、すなわち、受信されたパケットが期待されるパケットと等しい場合、失われたパケットはなく、decoded_indexは1インクリメントされる。そうでない場合、一つまたは複数のパケットが失われたに違いない。
【００５４】
次のステップで、受信されたフレームが現在のGOPのフレームのシーケンスにおいてのちに現れるかどうかが判定される。f″＝f_cとなるようなf″が存在する場合、それは正常なGOPであり、一つまたは複数のパケットが失われている。そうでない場合、異常GOPであり、よってビットストリームの終わりにある。
【００５５】
例示的に、ビットストリームは、図２ｄ）に示されるようにGopSize＝16のGOPを含む。上記の表２ｄ）によれば、規定されるデコード順は$A_{POC}=\{0,16,8,4,12,2,6,10,14,1,3,5,7,9,11,13,15\}$である。図３ａ参照。期待されないPOC番号、たとえばPOC＝2のあとにPOC＝14をもつパケットが受信されるとき、受信されたパケット（POC＝14）が、最後に正し

く受信されたパケット（POC＝2）よりあとに残っている期待されるパケットの一つであるかどうかが検査される。ここではこれがあてはまる。さらに、その後のパケットの（いくつかの）POC番号が顧慮されてもよい。この例では、これは{1,3,5,7,9,11,13,15}である。これは、そのGOPのその後に期待されるパケットに一致する。したがって、正常なGOPであり、欠けているのは二つのパケット、すなわち6および10だけだと結論できる。図３ｂに示されるように、次いでこれらのパケットは仮想パケットi1、i2によって置き換えられる。

【００５６】

図４ｂに描かれる別の例では、異常GOPであるGOPは9フレームしか含んでいない。その際、まずPOC＝8（このGOPにおける最低時間的レベルをもつ）のフレームが符号化され、次いでGOP＝4のフレーム（このGOPにおける二番目の時間的レベル；このGOPはPOC＝12のフレームをもたないことに注意）、次いでフレーム{2,6}（三番目の時間的レベル）そして{1,3,5,7,9}となる。よって、正しく受信される順序は{8,4,2,6,1,3,5,7,9}である。

【００５７】

ここで、この異常GOPの検出について述べる。記述される異常GOPのパケットの相対frame_num値は{1,2,3,4,5,5,5,5,5}である。もしたとえばPOC＝8またはPOC＝4のパケットが欠けているがPOC＝2のパケットは受信される場合、このパケットの相対frame_numは3である。これは正常なGOPの相対frame_num（これはPOC＝3については5であるべきである）に一致しない。frame_numの期待される値と受信された値の間のこの不一致は、POC値との関連で、異常GOPを検出するために使用されることができる。同じことはフレームのtemporal_level属性およびその他のパラメータ（たとえばnal_ref_idc）についてあてはまる。

【００５８】

以下では、正常GOPについての仮想パケットの構築（パケット仮想化）について例示的に述べる（図３参照）。

【００５９】

Aもデコード順によって順序付けし、Aの終わりをAの始まりにつなぐと、環を構築できる。f′とf_cとの間のギャップに対応するパケットを挿入し、同時にdecoded_indexを増す。たとえば、f″からf_cまでの環において（図３ではf″＝2、f_c＝14）、f_aおよびf_b（図３ではf_a＝6、f_b＝10）を有する場合、二つのパケットi1、i2をデコードのためのビットストリームに挿入し、decoded_indexを2だけ増す。これらのパケットi1、i2は今では、f_aからf_bまで、少なくとも三つの正しい値：POC番号、frame_numおよびtemporal_levelをもつ。何らかのスライス・レベルの予備的な値をもつf_aおよびf_bがデコーダに送られ、デコーダは誤り隠蔽技術を用いてユーザー攪乱を低減させる。

【００６０】

ここで、異常GOPについての仮想パケットの構築について述べる。

【００６１】

A内に、現在フレームf_c（たとえば、上の例ではPOC＝3、frame_num＝3）に等しいf″を見出せない場合、異常GOPのデコード過程にある。異常GOPは、長いビットストリームのまさに最後の部分であり、正常なGOPより短い。しかしながら、これが扱いが最も複雑な部分である。

【００６２】

ひとたび異常GOPにおけるパケットロスが検出されると、以下の「異常GOPプロセス」にはいってもよい。

【００６３】

まず、正常GOPについて上記したようにして、正常GOPの失われたパケットすべてが挿入される。これは図４に示されている。図４では例示的に、GOPはPOC＝9のあとで終わっており、スライス中のGOPの正常サイズはGopSize＝16であった。これは、初期の、期待されるフレーム相対順を決定するという目的をもつ。すなわち、不完全なGOPのサイズとは独立に、含まれるフレームの相対順が固定されるのである。それは、完全なGOPについてと同じである。たとえば（図４参照）、不完全なGOPがPOC＝{8}およびPOC＝{4}のフレーム

を有する場合、{8}は常に{4}より低い時間的レベルにあり、先にデコードされる。

【００６４】

GopSizeパラメータを変更する前に、decoded_indexがnext_decoded_index＝cur_decoded_index%GopSize（ここで、%は法の演算）、あるいは短縮記法では
decoded_index%=GopSize
によって更新される。

【００６５】

次いで、異常GOPについてGopSizeは異なっているので、これが変更される。デコード過程全体はビットストリームの終わりに近いので、GopSize値はいずれにせよまもなく変更される。

【００６６】

ここで、正常GOPに対する相違は、Aが異なる仕方で更新され、最後の異常GOP（abnormal GOP）についてのGopSizeが推定され、変更されるということである。同時に、Aにおけるframe_num値も変更される。

【００６７】

第一のステップでは、GopSizeの上限が決定される。現在フレームf_cが奇数フレームであれば、GopSizeの上限を、可能性のある誤差1で、知ることができる（<<および>>はシフト演算）：

$GopSize_{abn,max}$＝((frame_num－1)%(GopSize>>1)<<1)+1　(式2)

よって、上限を推定する一つの可能性は、奇数の現在フレームを選択して、この　(式2)を使うことである。

【００６８】

そうでなく、f_cが奇数フレームでない場合には、Aにおいて、f'''がf_cと同じPOC番号をもつようなf'''をみつける。プロセス全体を通じてGopSizeは何らかの正の可能な上限まで減少するのみなので、f'''は常に存在する。

【００６９】

たどるべき逐次反復アルゴリズムは、ともに同じPOC番号をもつf'''とf_cとの間のframe_numの差から出発して、A内で、現在フレームf_cより高いPOC番号をもつがより低いtemporal_levelをもつ各フレームについて、値が0になるまで、値を減少させていくことである。最後のフレームのPOC番号がGOPサイズについての上限を与える：

初期化：

$frame_diff = f^m_{frame_num} - f_{c\ \ frame_num}$

〔$f_{c\ \ frame_num}$はフレームf_cのframe_numを意味する。〕

```
temp_frame_diff＝frame_diff;
for(i＝GopSize－1;  i>=0;  i－－)//AはPOCによって順序付けされている
{
  if (f_i POC＞f_c POC && f_i Temporal_Level＜f_c Temporal_Level)
    temp_frame_diff－－;
  if(temp_frame_diff==0)  break;
}
```

$GopSize_{abn,max}$=$f_{i-1\ \ POC}$（すなわち、フレームi－1のPOC値）

$GopSize_{abn,max}$が現在の空間的層の異常GOPのGOPサイズの上限である。

【００７０】

第二のステップでは、すべての空間的層の異常GOPのGOPサイズの上限が更新される。MTL_{layer}＜MTLであることを注意しておく。

【００７１】

【数２】

$$GopSize_{layer} = GopSize \cdot 2^{MTL_{layer}-MTL} + \left\lfloor 2^{MTL_{layer}-MTL} - 1 \right\rfloor$$

今や、各層について、GopSizeの新しい長さをもつフレーム情報リストAがある。A内の他の要素は削除される。

【００７２】

単に簡単のために、Aを表すのに、空間的層を区別するための添え字を使っていない。これは各層が同様のプロセスをたどるからである。しかしながら、各空間的層が独自の要素リストAを有していてもよい。

【００７３】

第三のステップでは、A内のframe_num値が更新される。直前のGOPの最後のパケットのframe_numから、始まりのstart_frame_numはわかっている。更新されたAにおいて、Aについて新しい順序を与える。すなわち、Aはまず時間的レベルによって順序付けされ、同じ時間的レベル内では、POCによって順序付けられる。実際、この順序は正常GOPのデコード順と同様である。簡単のため、やはりそれをデコード順と呼ぶことにする。しかしながら、順序付け指数tを使う。よって、新たに順序付けされたA＝$\{f_t | t=0,\cdots GopSize-1\}$

$f_{0\ \ frame_num}$＝start_frame_num

$f_{t\ \ frame_num}=f_{t-1\ \ frame_num}+1$　　$f_{t-1\ \ Temporal_level}$＜MTLの場合

$f_{t\ \ frame_num}=f_{t-1\ \ frame_num}$　　それ以外の場合

が得られる。

【００７４】

frame_num値も伝統的な仕方で更新される：

frame_num＝frame_num%Max_Frame_Num

このステップで述べたプロセスは、各空間的層について使われるべきである。

【００７５】

このステップののち、次のパケットを処理する用意をし、第一のステップにジャンプする。

【００７６】

ひとたび異常パケットロス検出の中にはいったら、このプロセスはパースが終了されるまで続けられる。最後のプロセス・ステップは、全パケットが受信されるまで反復される。打ち切られたAのGopSizeがパケットロス検出およびパケット仮想化の次のプロセスのために保存されてもよい。decoded_indexを使うことにより、Aがデコード順によって順序付けされていれば、常にA内に要素：$f_{decoded_index+1}$を見出すことができる。f_cが$f_{decoded_index+1}$に等しい場合、失われたパケットはなく、現在の受信パケットがパーサーに送られる。そうでない場合、仮想パケットが$f_{decoded_index+1}$を使って構築される。いずれの場合にも、decoded_indexは1だけ増される。

【００７７】

さらに、以下に述べるように、BLSkipモードのために、新しい参照リスト構築方法が役に立つ。

【００７８】

まず、仮想シンタックス生成について述べる。pic_order_cnt_lsbおよびframe_numがスライス・ヘッダからわかっているので、仮想パケットについての正しい値が生成されることができる。時間的レベルが0でない場合、仮想パケットは、1または0のnal_ref_idcをもつBフレームとして指定され（該仮想パケットが最高の時間的レベルであれば、nal_ref_idcは0である）、それ以外の場合、nal_ref_idcが3のPフレームとして指定される。仮想パケットのSPS　IDおよびPPS　IDは同じ層の正しく受信されたパケットから設定する。

【００７９】

時間的レベルが0に等しいか否かに基づいて、仮想パケットがキー・ピクチャーである

か否かを指定することもできる。

【００８０】

仮想パケットがキー・ピクチャーである場合、メモリ管理制御動作（MMCO: memory management control operation）コマンドが、既知のJSVMエンコーダがするように、挿入される。これらのMMCOコマンドは、直前のGOPのBフレーム（あるいはむしろキー・ピクチャーでないピクチャー）および直前のキー・ピクチャーの前のキー・ピクチャー（もし存在すれば）を短期リスト（the short term list）から除去するために使われる。これらのコマンドは、キー・ピクチャーのスライス・ヘッダに挿入される。

【００８１】

RPLRおよび参照指数番号も扱われる必要がある。通常、キー・ピクチャーだけが、直前のキー・ピクチャーを参照することをキー・ピクチャーに強制するRPLRコマンドを追加する必要がある。

【００８２】

BlSkipモードについても誤り隠蔽方法の効率を保証するため、ELパラメータnum_ref_idx_l0_active_minus1／num_ref_idx_l1_active_minus1の値および参照リストも指定される。

【００８３】

num_ref_idx_l0_active_minus1およびnum_ref_idx_l1_active_minus1のBLパケット値は、空間的ELパケットについて設定される。

【００８４】

実は、参照指数番号のみならず、ELの参照リストがBL参照リストに正しく一致すべきである。ここで、「一致する（match）」とは、BLおよびEL中のlist_Xの同じref_idx_lx値をもつ場合、BLおよびELの対応する参照ピクチャーは同じ表示時間（またはPOC）をもつということを意味する。通常のJSVMエンコーダは、BLについてのRPLRと、ELにおいてBLに一致するような参照リスト構築の初期化とを使うことによってこれを保証している。よって、失われた非キー・ピクチャーについてのRPLRコマンドを生成することは必要ではないが、それでもキー・ピクチャーについてRPLRを生成することは必要である。

【００８５】

JVT-Q201, SVC WD 4.0のG.8.2.4.2.1“Initialisation process for the reference picture list for P, EP and SP slices in frames”において定義されているデコーダにおける制限が取り除かれる場合、失われた向上層についてのRPLRは、BLSkipモードのパフォーマンスをサポートすることが必要である。たとえば欧州特許出願EP06300020に記載されているプロセスは、空間的ELパケットが失われている場合に、デフォルトの参照ピクチャー・リスト構築プロセスとして、誤り隠蔽のために役に立つことがありうる。

【００８６】

本発明は以下の利点を提供する。

【００８７】

SEIメッセージ、SPS、PPSおよびIDRに属するパケット以外は任意のパケットロスを扱える。

【００８８】

固定GOPサイズをもつ長いシーケンス（8000フレーム超）を扱える。長さはGopSize・n＋d（ここで0≦d≦GopSize）である。

【００８９】

異なる空間的層が同じまたは異なるフレームレートをもつことができる。

【００９０】

各解像度は、一つまたは複数のFGS層をもつことができる。FGS層は、その対応するより低い品質レベルのパケット（同じ依存性レベルおよび時間的レベルをもつもの）が失われているか意図的にドロップされている場合、意図的にドロップされる。

【００９１】

パーサーの前にデフォルト・パケットを挿入することは、パーサーがパケットロス状況を扱う必要がないという利点をもつ。パケットロス状況はたとえば、パケットの番号付けをし直して通常の単純なパーサーを使えるようにすることを要求しうる。そのようなパケットのペイロード内容は重要ではない。

【００９２】

end_of_stream_rbspについての実装は、逐次反復でコードをサポートする。

【００９３】

本発明の開示された概念は、パケットが二つ以上のフレームまたはスライスを含む場合にも使うことができる。

【００９４】

多重ループ・デコードおよび単一ループ・デコードは両方ともサポートされる。スライス・グループ特徴についてのパケットロス検出もサポートできる。

【００９５】

本発明は、ビデオ・デコード製品またはビデオ・エンコード製品のために、特に基本層および少なくとも一つの向上層が含まれ、複数の時間的レベルが定義されている場合、使うことができる。特に、複数の空間的層は異なるフレームレートを有していてもよい。

【図面の簡単な説明】

【００９６】

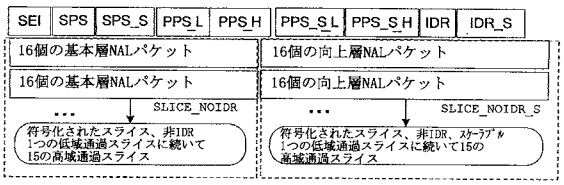
【図１】SVCビデオ・ビットストリームのパケット構造を示す図である。

【図２】種々のサイズのGOP内での時間的レベルを示す図である。

【図３】挿入された仮想パケットをもつGOPを示す図である。

【図４】異常GOPを示す図である。

【図１】



【図２】

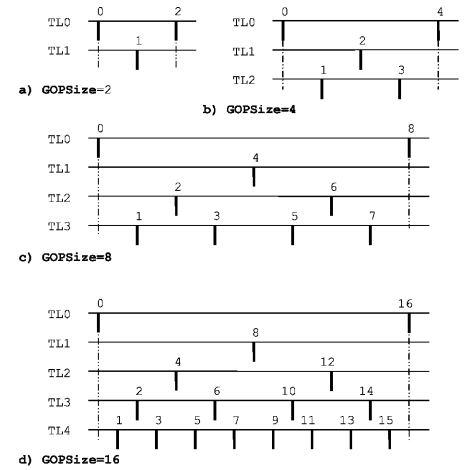


Fig.2

【図３】

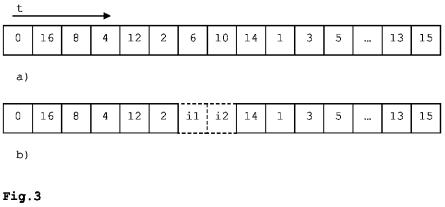


Fig.3

【図４】

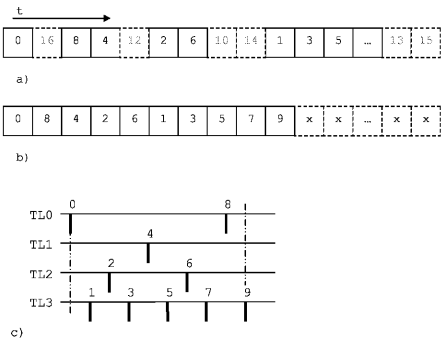


Fig.4

フロントページの続き

(72)発明者　チェン，イン
中華人民共和国，ベイジン　１０００８３，ハイ・ディアン・ディストリクト，シュ・チン・ロード，チェンホワ・ユアン，ビルディング　３　１０３
(72)発明者　シエ，カイ
中華人民共和国，ベイジン　１０００８６，ハイ・ディアン・ディストリクト，シュアンユシュ・ノース・ロード　２０７　ビルディング　４エイ
(72)発明者　パンディット，パーヴィン，ビバス
アメリカ合衆国，ニュージャージー州　０８８２３，フランクリン・パーク，ペア・トゥリー・レーン　２３

審査官　長谷川　素直

(56)参考文献　特開２００６－２９５５６８（ＪＰ，Ａ）
国際公開第２００６／０４６４７６（WO，Ａ１）
特表２００５－５３９４１０（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｈ０４Ｎ　　７／２６－７／６８，
Ｈ０３Ｍ　　７／３０