

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4467984号
(P4467984)

(45) 発行日 平成22年5月26日 (2010. 5. 26)

(24) 登録日 平成22年3月5日 (2010. 3. 5)

(51) Int. Cl. F I
G 1 0 L 19/02 (2006.01) G 1 0 L 19/02 1 5 0

請求項の数 18 (全 35 頁)

(21) 出願番号	特願2003-561255 (P2003-561255)	(73) 特許権者	590000248
(86) (22) 出願日	平成14年12月12日 (2002. 12. 12)		コーニンクレッカ フィリップス エレク
(65) 公表番号	特表2005-515511 (P2005-515511A)		トロニクス エヌ ヴィ
(43) 公表日	平成17年5月26日 (2005. 5. 26)		オランダ国 5 6 2 1 ベーアー アイン
(86) 国際出願番号	PCT/IB2002/005417		ドーフエン フルーネヴァウツウェッハ
(87) 国際公開番号	W02003/061299		1
(87) 国際公開日	平成15年7月24日 (2003. 7. 24)	(74) 代理人	100087789
審査請求日	平成17年11月16日 (2005. 11. 16)		弁理士 津軽 進
(31) 優先権主張番号	02075215.0	(74) 代理人	100092048
(32) 優先日	平成14年1月18日 (2002. 1. 18)		弁理士 沢田 雅男
(33) 優先権主張国	欧州特許庁 (EP)	(72) 発明者	アブレア, ヤヴィエル, エフ.
			オランダ国 5 6 5 6 アー アー アイ
			ンドーフエン プロフホルストラーン 6
			最終頁に続く

(54) 【発明の名称】 オーディオのコード化

(57) 【特許請求の範囲】

【請求項 1】

エンコード時に定義されたシーケンス $F(j)$ の、長さ F が可変のオーディオ・フレームを設けることにより、オーディオ・フレーム平均長



10

を、整数 M 個のビデオ・フレームに渡ってビデオ・フレーム長 $1/f_v$ に等しくさせる、オーディオ・データとビデオ・データを運ぶストリームを得るためのオーディオ・エンコード方式。

【請求項 2】

逐次的なオーディオ・フレーム間のオーバーラップ O を変えることにより、前記フレーム長 F が調整される、請求項 1 に記載のエンコード方式。

【請求項 3】

前記値 $F(j)$ が、 j (一連のフレームを定義する $F(j)$ の周期性) に基づいて周期的に

20

繰り返す、請求項 1 または請求項 2 に記載のエンコード方式。

【請求項 4】

シーケンスごとに M 個のビデオ・フレームと N 個のオーディオ・フレームを有し、各オーディオ・フレームが、t 個のサンプルから成るブロック k 個分により構成されている、請求項 3 に記載のエンコード方式。

【請求項 5】

前記シーケンス内のフレーム間のオーバーラップの合計 O_T が、 $O_T = p \times O + q \times (O + 1)$ に等しく、式中 O が、

$$p \in \mathbb{N} \wedge q \in \mathbb{N} \wedge O \in \mathbb{N} \wedge O_T \in \mathbb{N}$$

10

となる、ブロック内のオーバーラップ長である、請求項 4 に記載のエンコード方式。

【請求項 6】

特定のビデオ・フレームに対応するオーディオ・フレームしかオーバーラップされない、請求項 5 に記載のエンコード方式。

【請求項 7】

$p = (N - M) \times (O + 1) - O_T$ および $q = (N - M) - p$ である、請求項 6 に記載のエンコード方式。

【請求項 8】

特定のビデオ・シーケンスに対応するオーディオ・フレームしかオーバーラップされない、請求項 5 に記載のエンコード方式。

20

【請求項 9】

$p = (N - 1) \times (O + 1) - O_T$ および $q = (N - 1) - p$ である、請求項 8 に記載のエンコード方式。

【請求項 10】

隣接したオーディオ・フレームがどれもオーバーラップされる、請求項 5 に記載のエンコード方式。

【請求項 11】

$p = N \times (O + 1) - O_T$ および $q = N - p$ である、請求項 10 に記載のエンコード方式。

30

【請求項 12】

$$\exists n \in \mathbb{N}^+ : n \times t = M \times \left(\frac{f_A}{f_V} \right)$$

40

である、請求項 4 ～ 請求項 11 の何れかのエンコード方式。

【請求項 13】

ビデオとある程度マッチングした N 個のフレームのオーディオ・サンプルが、準可変のオーバーラップと共にフレーム内でエンコードされ、このことにより前記オーディオ・フレームの有効長が、M 個の一連のビデオ・フレームの長さと合致する（M と N は正の整数である）、オーディオ・データとビデオ・データとをエンコードしたストリームを得るためのオーディオ・エンコード方式。

【請求項 14】

有効長 F が、エンコード時に定義されたシーケンス $F(j)$ になるように、オーディオ・フレームにより可変のオーバーラップを設けることにより、オーディオ・フレームの平均

50

有効長



が、整数M個のビデオ・フレームとN個のオーディオ・フレームに渡って前記ビデオ・フレーム長 $1/f_v$ に等しくなるように、可変長のオーディオ・フレームを作り出す、 10

オーディオをコード化して、オーディオ・データとビデオ・データを運ぶストリームを得るための、オーディオ・エンコーダ。

【請求項 15】

長さOの短いオーバーラップと、シーケンス状態の合計q個の長いオーバーラップと、を有するストリームをコード化し、N個のフレームの後に反復するアルゴリズムを用いて、先頭のオーバーラップを計算する、

請求項 14 に記載のオーディオ・エンコーダ。

【請求項 16】

入来するフレームの予想されるフレーム有効長を計算し、 20

前記入来するフレームの実際の長さを調整して、この長さを前記予想されるフレーム長に等しくさせ、

受信されたフレーム内のブロックが、冗長ブロックまたは非冗長ブロックかを決定し、前記非冗長ブロックをサブバンド・サンプル上にマッピングする、

オーディオ・データとビデオ・データをエンコードするストリームをデコードするための、オーディオ・デコーダであって、

前記デコードするストリームは、請求項 1 に記載のオーディオ・エンコード方式によって得られるストリームである、オーディオ・デコーダ。

【請求項 17】

各ブロックに1組のブロック演算子を1つ以上適用することにより、前記データ・ストリーム内のブロックの前記オーバーラップ状態を変更する、請求項 16 に記載のオーディオ・デコーダ。 30

【請求項 18】

前記 1 組の演算子が、

ブロックの状態を変化させない演算子であるNOPと、

先頭のオーバーラップからの最初の非冗長ブロックを冗長ブロックに変化させる演算子であるDropと、

末尾のオーバーラップからの最初の冗長ブロックを非冗長ブロックに変化させる演算子であるAPPENDと、

drop演算子とappend演算子の両方を組み合わせた演算子であるShiftと、 40

を1つ以上含む、

請求項 17 に記載のオーディオ・デコーダ。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、他のデータ・ストリームに同期化されている箇所それが編集可能となるように、オーディオ信号をデータ・ストリームにコード化することに関する。本発明は、特に、関連するビデオ・フレーム境界の圧縮領域内でオーディオを非破壊的にスプライシングする必要があるデジタル・テレビ伝送方式を用途とするが、この用途には限定されない。

【背景技術】

【0002】

デジタル・テレビジョン (DTV) システムにより、バンド幅が限られたチャンネル上で、幾つかのプログラムをブロードキャストすることが可能となる。これらのプログラムの各々は、ビデオ・コンテンツとオーディオ・コンテンツを有する。これらのプログラムの幾つかは、高品質な多重チャンネル（例えば、家庭用映画システムにより再生可能な5つのチャンネル）のオーディオを含むことができる。DTVの製作サイト、ネットワーク、および加盟局は、一般的に、ビデオ・テープ・レコーダと、すべてのオーディオ・コンテンツを運ぶための伝送線とを用いる。このインフラストラクチャの多くは、2つの未圧縮オーディオ・チャンネルに対する容量しか有していないので、多重チャンネルは、通常、軽く圧縮されかつフォーマットされてから、記録または伝送される。放出（すなわち、エンドユーザへのブロードキャスト）の前にプログラム・ストリームは、強く圧縮される。

10

【0003】

DTV製作でのコントリビューション段と配信段では、元のストリームは、プログラムの編集またはプログラムのスイッチング（例えば、ローカルなコンテンツを生中継のネットワーク・フィードに挿入すること）のために、スプライスされなければならない。このようなスプライシングは、コンテンツ・ストリーム内のビデオ・フレームの境界で行われる。

【0004】

ブロードキャストされたストリームのオーディオ・コンテンツは、幾つかの要件を満たすことが出来なければならない。特に、プログラムが家庭用映画システムなどの高品質の再生機器を用いて再生される場合、DTV視聴者は、受信されたプログラムが高い知覚オーディオ品質を有していることを期待する可能性がある。例えば、エンコーディング段と多重のデコーディング段とを多重にカスケードさせることによる可聴アーチファクトがあってはならず、かつプログラムをスイッチングしている間に知覚可能な音の中断があってはならない。最も重要な点は、再生されたプログラムがリップ・シンクでなければならないこと、つまりオーディオ・ストリームが、対応するビデオ・ストリームと同期していないことである。これらの目的を、手頃なコストで、すなわち、既存の（2チャンネルの）インフラストラクチャを用いて達成するには、オーディオ・プログラムを圧縮領域内でスプライスしなければならない。

20

30

【0005】

【非特許文献1】「ドルビー・デジタル・ブロードキャスト実施ガイドライン第91549号 (Dolby Digital Broadcast Implementation Guidelines Part No. 91549)」 (Dolby Laboratories, Version 2, Dolby Laboratories, 1998)

【非特許文献2】ISO/IEC 11172「デジタル格納媒体のための動画および関連付けられたオーディオの最大約1.5 Mbit/sまでのコード化（第3部）：オーディオ（1993～08）("Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s" Part 3: Audio (1993-08))」

【非特許文献3】「カスケードされた音響心理学的なコード化におけるオーディオ品質の維持 (Maintaining Audio Quality in Cascaded Psychoacoustic Coding)」 (Warner R. Th. ten Kate 101st AES Convention 1996 November 8-11)

40

【非特許文献4】IEC61937規格「PCMによりエンコードされた非線形のオーディオ・ビット・ストリームのためのIEC 60958を適用したインターフェイス」 (Interface for non-linear PCM encoded audio bitstreams applying IEC 60958)

【発明の開示】

【課題を解決するための手段】

【0006】

既存のメザニン・エンコード方式には、最大8チャンネルまでのエンコードされたオーディオおよび多重化されたメタデータを、AES-3ペアを介して配信するための、ドルビー研究所 (Dolby Laboratories) の「ドルビー・デジタル・ブロードキャスト実施ガイドライ

50

ン第91549号 (Dolby Digital Broadcast Implementation Guidelines Part No. 91549)」
(Dolby Laboratories, Version 2, 1998年) に定義されているドルビー E (r. t. m.)
が含まれる。間もなく導入される (NAB 1999) DP571 ドルビー E エンコーダおよびDP572
ドルビー E デコーダにより、エンコードされたオーディオの編集とスイッチングを、ミ
ュートまたはグリッチを最小限に抑えて行うことが可能となるはずである。さらに、これ
らのエンコーダとデコーダにより、可聴な劣化のないカスケードイングが可能となる。ド
ルビー Eは20ビットのサンプル・サイズを用いることにより、ビット・レートを2:1 ~
5:1 に低減させる。

【0007】

英国放送協会外は、MPEG-2ビデオ・ビット・ストリームのスイッチングと編集を柔軟に
行う方法を、ACTS ATLANTICプロジェクトを通じて提案している。このシームレスなコン
カチネーション・アプローチは、デコーディングと再エンコーディングを側面情報と共に
行うことを利用して、カスケードイングによる劣化をなくしている。しかしながら、この
方式は、MPEG-2レイヤIIと AES/EBUインターフェイスを用いた用途に限定されている。さ
らに、このオーディオ・データは、編集箇所に対してスライドさせることができるので、
タイム・オフセットをもたらしてしまう。したがって、逐次的に編集を行うことによって
、オーディオ情報とビデオ情報との間のタイム・オフセットが大きくなってしまう可能性
がある。

【0008】

ビデオ・ストリームとオーディオ・ストリームは、ブロードキャスト・チェーン全体に
渡りリップ・シンク状態を維持しなければならない。すなわち、オーディオは、対応する
ビデオとの同期を保たなければならない。配信サイトは、放出前に、(例えば、ローカル
・コンテンツを含ませるために) オーディオ・ストリームとビデオ・ストリームをスプラ
イス(例えば、スイッチング、編集、またはミキシング)することができる。スプライ
シングの後、ビデオ・フレームとオーディオ・フレームの境界が合致しない場合(これは、
ほとんどのオーディオ・コード化方式で発生する)、オーディオがビデオに対してスリッ
プしているので、リップ・シンクを自動的に保証することはできない。特別な措置が講じ
られない極端な場合、このことにより、ミュートまたはグリッチなどのオーディオ・アー
チファクトが発生してしまう可能性がある。非対応のオーディオ・ストリームのデコード
化を試みた結果、グリッチが発生してしまう場合もあるが、これらのグリッチをなくすた
めにミュートを適用してもよい。本発明の目的は、ミュート、グリッチ、またはスリップ
などのオーディオ・アーチファクトをもたらすことなくスプライスさせることができるオー
ディオ・ストリームを得るためのエンコード方式を提供することである。

【0009】

本発明の別の目的は、品質損失を最小限に抑えて圧縮と解凍のカスケードイングを行う
ことができるエンコード方式を提供することである。

【0010】

本発明は、第一態様により、

エンコード時に定義されたシーケンス

$F(j)$

の長さ

F

が可変のオーディオ・フレームを設けることにより、オーディオ・フレーム平均長

$\bar{F}$

を、整数

M

10

20

30

40

50

個のビデオ・フレームに渡ってビデオ・フレーム長

$$\frac{1}{f_v}$$

に等しくさせる、

オーディオ・データとビデオ・データを運ぶストリームを得るためのオーディオ・エンコード方式、
を提供する。

【0011】

この方式により、オーディオ情報を劣化させずに、ストリームを少なくともビデオ・フレームごとに編集することが確実に可能となる。逐次的なオーディオ・フレーム間のオーバーラップ

10

O

を変えることにより、フレーム長

F

を調整できることが好ましい。

【0012】

本発明を具現化する方式の場合、値

20

$F(j)$

を

j

(一連のフレームを定義する周期性)に基づいて周期的に繰り返してもよい。通常、

M

個のビデオ・フレームと

N

30

個のオーディオ・フレームがシーケンスごとに存在し、各オーディオ・フレームが

k

個のブロックから構成される。シーケンス内のフレーム間のオーバーラップ

O_T

の合計を、例えば、

$$O_T = p \times O + q \times (O + 1)$$

40

に等しくしてもよい(ここで、

O

は、ブロック内のオーバーラップ長である)。

【0013】

本発明の範囲内にある一方式の場合、特定のビデオ・フレームに対応するオーディオ・フレームしかオーバーラップされない。このような方式の場合、

p

と

50

q

の値は、等式、

$$p = (N - M) \times (O + 1) - O_T$$

および

$$q = (N - M) - p$$

10

を満たしてもよい。

【0014】

一代替方式の場合、特定のビデオ・シーケンスに対応するオーディオ・フレームしかオーバーラップされない。このような方式の場合、

 p

と

 q

の値は、等式、

20

$$p = (N - 1) \times (O + 1) - O_T$$

および

$$q = (N - 1) - p$$

を満たしてもよい。

【0015】

さらなる一代替方式の場合、隣接するどのオーディオ・フレームもオーバーラップされる。このような好ましい方式の場合、

30

 p

と

 q

の値は、等式、

$$p = N \times (O + 1) - O_T$$

40

および

$$q = N - p$$

を満たしてもよい。この後者の方式は、一連のビデオ・フレーム

 M

に最適なオーバーラップの値を、

$$\exists n \in \mathbb{N}^+ : n \times t = M \times \left(\frac{f_A}{f_V} \right)$$

となるように提供することができる。

【0016】

ビデオ・シーケンスとは、毎秒のビデオ・フレームのレートが

f_V

10

である、整数の（かつ可能なら有限数の）ビデオ・フレーム（すなわち

M

）であって、各ビデオ・フレームが、等しい整数

N

個の（圧縮された）オーディオ・フレームを含み、各オーディオ・フレームが整数

k

20

個のブロックを含み、各ブロックが、毎秒

f_A

個のサンプルというサンプリング・レートの整数

t

個のオーディオ・サンプルを表すビデオ・フレームである、と定義付ける。ビデオ・フレームの数に、オーディオ周波数とビデオ周波数との間の比率を掛けたものを、（圧縮された）オーディオのブロック毎のオーディオ・サンプル数で除算した余りをゼロに等しくさせることにより、

30

M

は整数となることが保証される。したがって、

N

も整数となる。この結果、オーバーラップしているブロックの総数も整数となるので、各単一のオーバーラップも整数となる。ほとんどの場合、オーバーラップするブロック数が整数であることは必須である。サンプルのブロックは、基礎を成しているコーデックにより扱われる情報の最小単位である。

【0017】

40

本発明は、第二態様により、ビデオとある程度マッチングした

N

個のフレームのオーディオ・サンプルが、準可変のオーバーラップと共にフレーム内でエンコードされ、このことにより前記オーディオ・フレームの有効長が

M

個の一連のビデオ・フレームの長さと合致する（

M

50

と
 N

は正の整数である)、エンコードされたオーディオ・データとビデオ・データとを運ぶストリームを得るためのオーディオ・エンコード方式を提供する。

【0018】

本発明は、本発明の前述した何れかの態様の方式に従ってエンコードされたデータ・ストリームを提供する。このようなストリームには、オーディオ・フレームのサイズを示すタグが各々付けられたオーディオ・フレームを含めてもよい。ブロックが冗長ブロックであるか否かを示すタグをブロックに付けてもよい。

10

【0019】

本発明は、別の態様により、本発明の第一態様に従ってオーディオ・ストリームをエンコードするための(例えば、ソフトウェア・コンポーネントまたはハードウェア回路として実施してもよい)オーディオ・エンコーダを提供し、さらに本発明の第一態様に従ってオーディオ・ストリームをデコードするためのオーディオ・デコーダを提供する。

【0020】

本発明のこの態様によるオーディオ・デコーダは、1組のブロック演算子の内の1つ以上を各ブロックに適用することによって、データ・ストリーム内のブロックの冗長状態を変更することにより動作する。これは、

NOP: ブロックの状態を変化させない演算子と、

20

Drop: 先頭のオーバーラップからの非冗長な第一ブロックを冗長ブロックに変化させる演算子と、

APPEND: 末尾のオーバーラップからの最初の冗長ブロックを非冗長ブロックに変化させる演算子と、

Shift: drop演算子とAPPEND演算子との組み合わせである演算子と、
を1つ以上含む1組の演算子により達成される。

【0021】

本発明は、特に、
長さ

F

30

がエンコード時に定義されたシーケンス

$F(j)$

になるように、オーディオ・フレームにより可変のオーバーラップを設けることにより、オーディオ・フレームの平均有効長

$\bar{F}$

が整数

M

40

個のビデオ・フレームに渡ってビデオ・フレーム長

$\frac{1}{f_v}$

に等しくなるように、可変長のオーディオ・フレームを作り出す、

オーディオをコード化して、オーディオ・データとビデオ・データをエンコードしたストリームを得るための、オーディオ・エンコーダ、
を提供する。

【0022】

50

このようなオーディオ・エンコーダは、

N

個のオーディオ・フレームの後に繰り返すアルゴリズムを用いて先頭のオーバーラップを計算して、長さ

O

の短いオーバーラップと、シーケンス状態の合計

q

10

個の長いオーバーラップとを有するようにストリームをコード化してもよい。

【0023】

本発明は、さらなる態様により、エンコードされたオーディオ・データとビデオ・データを運ぶストリームをデコードするための、（例えば、ソフトウェア・コンポーネントまたはハードウェア回路として実施してもよい）オーディオ・デコーダを提供する。このデコーダは、入来するフレームの予想されるフレーム長

F

を、できる限り円形シフトされたシーケンス

$F(j)$

20

で計算し、この入来するフレームの実際の長さを、予想されたフレーム長に等しくなるように調整し、受信されたフレーム内のブロックの何れかが、冗長ブロックかまたは非冗長ブロックかを決定し、この非冗長ブロックをサブバンドのオーディオ・サンプル上にマッピングする。

【0024】

本発明を具現化するシステムの場合、サンプル・レート変換などの余分なオーディオ操作は通常行われない。さらに、受信されたストリームを正確にデコードするために必要なすべての情報は、ほとんどの場合、エンコーダで加えられることが通常なので、この情報を編集の間に変更する必要はない。したがって、既存のインフラストラクチャを用いて、変更を加えることなく編集を行うことができる。さらに、デコーディングを可能にするためにストリームに加える必要がある余分な情報は、非常に僅かとなる。最後に忘れてはならないことは、放出形式としてMPEGを用いる場合、MPEGに適した形式も用いて伝送を行うと、便利となる場合があることである。

30

【発明を実施するための最良の形態】

【0025】

次に、本発明の実施例の詳細を、添付の図面を参照しながら単に例として説明する。

【0026】

以下の説明では、以下の記号が全体に渡り用いられている。

40

f_A, f_V	オーディオ・サンプリング周波数、ビデオ・フレーム・レート	
t_A, t_V	オーディオ・フレームとビデオ・フレームの継続時間の長さ	
s	オーディオ・フレームごとのサンプル	
k	オーディオ・フレームごとのサンプルのブロック	
t	ブロックごとのサンプル	
$O, O_T, \overline{O}$	短いオーバーラップ、オーバーラップ合計、および平均オーバーラップ	
M, N	シーケンスごとのビデオ・フレームの量とオーディオ・フレームの量	10
p	シーケンスごとの短いオーバーラップの量	
q	シーケンスごとの長いオーバーラップの量	
j	フレーム指数	
$F(j), G(j)$	フレームの有効長	
$H(j), T(j)$	フレーム先頭のオーバーラップ、フレーム末尾のオーバーラップ	
$X(j), \overline{X}(j)$	累算された有効長、累算された平均有効長	
$\overline{F}$	平均有効長	20
b	短いフレームの長さ	
B	ビデオ・シーケンス内のブロック総数	
φ	位相	
$\mathbb{N}^+$	$\{1, 2, 3, \dots, \infty\}$	
Q	ヌル・パディング	
$A(j)$	アペンド演算トグル (append operation toggle)	
$OP(j)$	演算子	30
$\varepsilon(j)$	同期化エラー	
δ	同期化エラーの合計	
u, v	補助変数	

【0027】

第一に、図1を参照すると、典型的なDTVブロードキャスト・システムが、コントリビューション段10と、配信段12と、放出段14とを含んだチェーン状になっている。

【0028】

コントリビューション段内では、コンテンツが1つ以上の製作サイト20で生じ、かつ配信ネットワーク22によりブロードキャスト・ネットワーク・サイト24へ転送される。ブロードキャスト・ネットワーク24は、このコンテンツを含んだプログラム・ストリームを作り出し、かつこのプログラム・ストリームを、DTH (direct-to-home) 衛星ブロードキャスト32、地上ブロードキャスト34、ケーブル・テレビのプロバイダ36などの加盟局に、配信ネットワーク30を介して配信する。その後、加入者40は、これらの加盟局の内の1局の出力から、プログラム・ストリームを受信することができる。

【0029】

製作サイト内では、幾つかのタイプのコンテンツを製作し、かつ複数の異なる媒体に格納することができる。例えば、第一スタジオ50は、生放送型コンテンツを製作し、かつ第二スタジオ52は、記録型コンテンツ（例えば、コマーシャル広告）を製作することができる。いずれの場合にも、コンテンツには、ビデオ・コンポーネントとオーディオ・コンポ

ーメントが含まれている。各スタジオ50からの出力は、各々のエンコーダ54により同様に処理され、オーディオ・コンテンツとビデオ・コンテンツをエンコードした基本ストリームが生成される。生放送でブロードキャストされる第一スタジオ50からのコンテンツは、次に（適切な処理の後）無線リンクにより配信ネットワーク22へ伝送される。第二スタジオのコンテンツの場合、時間は重要ではないので、このコンテンツをテープ56上に記録し、かつ適切な方法で配信ネットワーク22に送信することができる。エンコーダ54、およびエンコーダ54が作り出す基本ストリームが、本発明の態様の実施例である。

【0030】

図3に示されているように、ネットワーク・サイト24内では、スプライサ60によって、様々なソースからのコンテンツがスプライスされて、プログラム出力が構成される。スプライサ60への入力、製作ユニット20からの無線リンク経由、テープ56、またはローカル・スタジオ64などの様々なソースから導出される同様のタイプの基本ストリームから導出される。スプライサ60の出力も、同じく基本ストリームであり、どの所与の時刻でも、これらの入力ストリームの内の選択されたストリームである。スプライサ60は、出力ストリームのオーディオ・コンポーネントとビデオ・コンポーネントとが途切れなく再生することが可能となるように入力ストリーム間のスイッチングを行うように、動作させることができる。次に、スプライサ60の出力はパケタイザ62により処理されて、トランスポート・ストリームが形成される。次に、このトランスポート・ストリームは変調されて、無線リンクにより加盟局に伝送され、加入者に配信される。

【0031】

本発明を具現化する基本ストリーム内のエンコードされたビデオ・コンテンツは、通常、一連のスキャンされたビデオ・フレームを有するであろう。このようなフレームは、プログレッシブ・スキャン型ビデオ・フレーム (progressive scanning video frames) でもよい。この場合、各フレームは、完全な静止画像である。このような場合、これらのビデオ・フレームは、フレーム・レート

$$f_v$$

を有し、かつ継続時間

$$t_v = 1/f_v$$

を有する。これに代えて、これらのフレームを、インターレース・スキャン型のフレームとしてもよい。この場合、各フレームは、2つの逐次的にインターレースされたフィールドから構築され、フィールド周波数は、上記で紹介した表記では

$$2f_v$$

となる。フレーム・レートとスキャン・タイプは、ストリームが目的とするテレビ・システムにより定義される。基本的なテレビ規格であるPALとNTSCは、これらの規格が用いられている国の送電線の周波数から、フレーム・レートを導出する。カラーテレビの導入により、NTSCは、率にして1000 / 1001だけ変更された。さらに、映画は24 Hzを用いているが、これを同じ率だけ変更してもよい。さらに、コンピュータ・モニタは、最大96 Hzまでの幾つかのフレーム・レートで走ることができる。以下の表1には、

$$f_v$$

の典型的な値が与えられている。

(表1)

ビデオ・フレーム・レート [Hz]	t_f [ms]	用 途
23.976	41.71	3-2 プルダウン NTSC
24	41.67	映画
25	40	PAL, SECAM
29.97	33.37	NTSC, PAL-M, SECAM-M
30	33.33	ドロップフレーム NTSC
50	20	ダブルレート PAL
59.94	16.68	ダブルレート NTSC
60	16.67	ダブルレート、ドロップフレーム NTSC

10

【 0 0 3 2 】

オーディオ信号は、ある周波数

$$f_A$$

(例えば、48 kHz) でサンプリングされた、時間的に連続したパルスコード変調 (PCM) 信号である。以下の表2には、

$$f_A$$

20

の値の例が与えられている。

(表2)

オーディオ・サンプリング周波数 [kHz]	用途
24	DAB
32	DAT, DBS
44.1	CD, DA-88, DAT
48	プロ用オーディオ, DA-88, DVD
96	DVD

30

【 0 0 3 3 】

これらの周波数の他にも、44.1 kHzと48 kHzを、率にして1000 / 1001 だけ変更して (例えば、44.056 kHz、44.144 kHz、47.952 kHz、および48.04 8 kHzにし)、オーディオを、映画からNTSCへのプルアップとプルダウンによる変換に適合させてもよい。さらに、映画からPAL方式へ変換するために、24 / 25という率を適用してもよい (例えば、42.336 kHz、45.937 kHz、46.08 kHz、および50 kHzが得られる)。さらに、DABは24 kHzと48 kHzを用い、DVDオーディオは44.1 kHz、88.2 kHz、176.4 kHz、48 kHz、9 6kHz、および19 2 kHzを用い、DVDビデオは48 kHzと96 kHzを用いてもよい。DATは32 kHz、44.1 kHz、お

40

【 0 0 3 4 】

サンプルの幅は、通常、16ビット、20ビット、または24ビットである。

【 0 0 3 5 】

圧縮の前に、オーディオ・ストリームは、継続時間が

$$t_A = s / f_A$$

のオーディオ・フレームに分割される。ここで、

50

S

は、オーディオ・フレームごとのサンプル数である（例えば、MPEG-2レイヤIIの場合は、

$$S = 1,152$$

個のサンプルであり、AC-3の場合は、

$$S = 1,536$$

個のサンプルである）。以下の表3には、様々なコード化方式で用いられるフレーム・レート
の例が示されている。

（表3）

10

コード化方式	使用するもの	フレーム長 [サンプル]	t_A [ms] @ 48 kHz
MPEG-1 レイヤ I	DCC	384	8
MPEG-1 レイヤ II	DAB, DVB, DVD-V	1,152	24
MPEG-1 レイヤ III	ISDN, MP3	1,152	24
MPEG-2 レイヤ II	DVB, DVD	1,152	24
MPEG-2 AAC		1,024	21.33
Dolby AC-3	DVD	1,536	32
Sony ATRAC	MiniDisc	512	n.a.

【 0 0 3 6 】

オーディオ・エンコーダの内部では、オーディオ・フレームは、さらに

20

 t

個のサンプルから成るブロック

 k

個に分割される（例えば、MPEG-2 レイヤ IIの場合、32個のサンプルから成るブロックが36個存在する）。ブロックとは、オーディオの最小処理単位である。これは、

$$s = k \times t$$

として表現することができる。以下の表4は、様々なコード化方式で用いられているフレーム
細分類を示す。

30

（表4）

コード化方式	$k \times t$ [ブロック×サンプル]
MPEG レイヤ I	12×32
MPEG レイヤ II	36×32
MPEG レイヤ III	2×576
Dolby AC-3	6×256

【 0 0 3 7 】

40

ブロードキャスト・チェーン全体に渡り、ビデオ・ストリームとオーディオ・ストリームをリップ・シンク状態に維持しなければならない。すなわち、オーディオは、対応するビデオと同期に保たなければならない。配信サイトは、放出前に（例えば、ローカル・コンテンツを含めるために）オーディオ・ストリームとビデオ・ストリームをスプライス（例えば、スイッチング、編集、またはミキシング）してもよい。

【 0 0 3 8 】

スプライシングの後、ビデオ・フレームとオーディオ・フレームの境界が合致しない場合（このことは、ほとんどのオーディオ・コード化方式で発生する）、リップ・シンクを自動的に保証することはできない。特別な手段が講じられない極端な場合、このことによりミュートまたはスリップなどのオーディオ・アーチファクトが発生してしまう。

50

【0039】

本発明の様々な実施例は、(MPEG-1とMPEG-2などの)既存の規格に関連するエンコーディングを行うことができるが、これらの実施例は、必ずしもこれらの既存の規格との後方互換性を有するとは限らない。

【0040】

実施例の基本原理

本実施例のコード化方式の場合、ビデオがある程度マッチングしている

N

個のフレーム内のオーディオ・サンプルは、準可変のオーバーラッピングにより、一連の 10

M

個のビデオ・フレームと合致する。本発明の実施例に従ってエンコードを行うと直ちに、各ビデオ・フレームには、等しい整数個のオーディオ・フレームが含まれる。したがって、ビデオ・フレーム境界で編集を行ってもよい。デコード後直ちに、冗長なサンプルを破棄してもよい。

【0041】

オーディオ・フレームが、

t

20

個のサンプルから成るブロック

k

個に分割されると仮定すると、ブロック内のオーバーラップの合計

O_T

は、下式により計算することができる。

$$O_T = (k \times N) - \left(\frac{M}{t} \times \frac{f_A}{f_V} \right) \quad \text{式 1}$$

30

式中、

M, N, k

および

t

は正の整数であり、かつ

f_A

と

f_V

は、周波数をHzで表し、

f_A / f_V

は有理数である。

【0042】

50

デコーダの再構築フィルタ内の編集されたオーディオ・ストリーム間でクロスフェードを行うために、オーバーラップの合計

O_T

は、下式に与えられているように整数のブロックと合致するように選択される。

$$O_T = p \times O + q \times (O + 1) \quad \text{式 2}$$

式中、

p, q

10

および

O

は、負ではない整数である。

【0043】

本発明の様々な実施例の範囲内で、オーディオ・フレームの端から端に至るまでのオーバーラップを全体的に広げるための様々なアプローチを採用することができる。すなわち、様々な異なる制約を課すことにより、これらの実施例の異なる実施態様を得てもよい。本明細書では、これらのアプローチとして、以下の3種類が参照されている。

アプローチ1—ビデオ・フレーム内でオーバーラップを行う。

20

アプローチ2—ビデオ・フレームのシーケンス内でオーバーラップを行う。

アプローチ3—ビデオ・ストリーム全体に渡りオーバーラップを行う。

【0044】

アプローチ3は、2つの隣接するオーディオ・フレーム間に、可能な最小のオーバーラップを常に設け、多くの場合、シーケンスごとのビデオ・フレーム数が最小となることを示すことができる。したがって、多くの用途の場合、このアプローチが他のアプローチよりも好ましい。しかしながら、特定の用途次第では、このことが常に当てはまるとは限らない。

【0045】

アプローチ1

30

図4の場合のように、1つのビデオ・フレーム内にしかオーバーラップが存在しない場合、ブロック内の平均オーバーラップ

$\bar{O}$

は、下式により与えられる。

$$\bar{O} = \frac{O_T}{N - M} \quad \text{式 3}$$

この式は、

$$p = (N - M) \times (O + 1) - O_T \quad \text{式 4}$$

40

つまり、長さ

O

のブロックのオーバーラップと、

$$q = (N - M) - p \quad \text{式 5}$$

つまり、長さ

$(O+1)$

のブロックのオーバーラップ、
として実施してもよい。

【0046】

アプローチ2

図5のように、オーバーラップが1つのシーケンス内にしか存在しない場合、ブロック内の平均オーバーラップ

$\bar{O}$

10

は、下式により与えられる。

$$\bar{O} = \frac{O_T}{N-1}$$

式 6

式6は、

$$p = (N-1) \times (O+1) - O_T$$

式 7

つまり、長さ

O

20

のブロックのオーバーラップと、

$$q = (N-1) - p$$

式 8

つまり、長さ

$(O+1)$

のブロックのオーバーラップ、
として実施してもよい。

30

【0047】

アプローチ3

図6のように、オーバーラップが複数のシーケンス内に存在する場合、ブロック内の平均オーバーラップ

$\bar{O}$

は、下式により与えられる。

$$\bar{O} = \frac{O_T}{N}$$

式 9

40

式9は、

$$p = N \times (O+1) - O_T$$

式 10

つまり、長さ

O

のブロックのオーバーラップと、

$$q = N - p$$

式 11

つまり、長さ

50

$(O+1)$

のブロックのオーバーラップ、
として実施してもよい。オーバーラップの長さ

O

は、下式のように表してもよい。

$$O = \lfloor \overline{O} \rfloor \quad \text{式 12} \quad 10$$

式12は、最後のアプローチの場合、下式のように書くことができる。

$$O = \left\lfloor k - \frac{\left(\frac{f_A}{f_V} \right)}{\left(\frac{N}{M} \right) \times t} \right\rfloor \quad \text{式 13}$$

M

20

は、下式を満たすように選択される。

$$\exists n \in \mathbb{N}^+ : n \times t = M \times \left(\frac{f_A}{f_V} \right) \quad \text{式 14}$$

また、ビデオ・フレームごとのオーディオ・フレームのレート

$\frac{N}{M}$

30

は、下式のように書いてもよい。

$$\frac{N}{M} = \left\lfloor \frac{\left(\frac{f_A}{f_V} \right)}{k \times t} \right\rfloor \quad \text{式 15}$$

【0048】

クロスフェード

40

ISO/IEC 11172の「デジタル格納媒体のための動画および関連付けられたオーディオの最大約1.5 Mbit/sまでのコード化（第3部）：オーディオ（1993-08）("Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s" Part 3: Audio (1993-08)）」に定義されている、MPEG-1デコーダ内の再構築フィルタは、オーバーラップ型のフィルタ・バンクである。スプライシングが、サブバンド領域（すなわちブロック）内で行われる場合、デコード後直ちに、約512個のオーディオ・サンプルがクロスフェードする。

【0049】

一般的なコード化規格に基づく実施例の実施

本発明を実施するための基礎として、様々なエンコード方式が考慮されてきた。特に、

50

MPEG-1とMPEG-2のレイヤ I とレイヤIIが考慮されてきたが、可能な方式がこのリストに限定されるわけではない。本発明を具現化する方式は、既存の規格と同様のコード化方式を用いているが、オーバーラッピングのためこれらの規格からは逸脱することを、ここで述べておかなければならない。

【0050】

当業者にはよく知られているように、MPEG-2は、MPEG-1との後方互換性を有するように多重チャンネル・オーディオをエンコードのための規格方式である。他方、後方互換性を有せずにMPEG-1規格を多重チャンネルに拡張したものは、実施態様を単純化することができる。さらに、レイヤ IIは、レイヤ I よりも効率的である。他方、レイヤ I は、ブロック数が少ないので、エンコードの冗長性が少ない。本発明の実施例の場合、MPEG-1 レイヤ I に基づく方式により、低い冗長性と実施態様の単純性とが、最良に組み合わせられる。

【0051】

MPEG-2レイヤII

MPEG-2 レイヤ IIを、エンコード方式の基礎として用いる場合、

$$k = 36$$

および

$$t = 32$$

である。

【0052】

表5は、実施例が上述したアプローチ1に基づいている場合の、オーディオ・サンプル周波数とビデオ・フレーム・レートの様々な組み合わせに対するオーバーラップ・シーケンスの例を、幾つか示している。

(表5) MPEG-2レイヤ II とアプローチ1

f_v [Hz]	f_A [kHz]	M	N	O_t	$\overline{O}$	$p \times O + q \times (O+1)$
23.976	48	16	32	151	9.437...	$9 \times 9 + 7 \times 10$
	44.1	2,560	5,120	37,173	14.520...	$1,227 \times 14 + 1,333 \times 15$
	32	24	48	727	30.291...	$17 \times 30 + 7 \times 31$
24	48	2	4	19	9.5	$1 \times 9 + 1 \times 10$
	44.1	64	128	933	14.578...	$27 \times 14 + 37 \times 15$
	32	3	6	91	30.333...	$2 \times 30 + 1 \times 31$
25	48	1	2	12	12	$1 \times 12 + 0 \times 13$
	44.1	8	16	135	16.875	$1 \times 16 + 7 \times 17$
	32	1	2	32	32	$1 \times 32 + 0 \times 33$
29.97	48	20	40	439	21.95	$1 \times 21 + 19 \times 22$
	44.1	3,200	6,400	83,253	26.016...	$3,147 \times 26 + 53 \times 27$
	32	n/a	n/a	n/a	n/a	n/a

表6は、実施例が上述したアプローチ2に基づいている場合の、オーディオ・サンプル周波数とビデオ・フレーム・レートの多様な組み合わせに対するオーバーラップ・シーケンスの例を、幾つか示している。

(表6) MPEG-2レイヤ II とアプローチ2

f_v [Hz]	f_A [kHz]	M	N	O_t	$\overline{O}$	$p \times O + q \times (O+1)$
23.976	48	16	32	151	4.870...	$4 \times 4 + 27 \times 5$
		32	64	302	4.793...	$13 \times 4 + 50 \times 5$
		48	96	453	4.768...	$22 \times 4 + 73 \times 5$
	44.1	2,560	5,120	37,173	7.261...	$3,779 \times 7 + 1,340 \times 8$
	32	24	48	727	15.468...	$25 \times 15 + 22 \times 16$
		48	96	1,454	15.305...	$66 \times 15 + 29 \times 16$
		72	144	2,181	15.251...	$107 \times 15 + 36 \times 16$
	24	48	2	4	19	$2 \times 6 + 1 \times 7$
			10	20	95	$19 \times 5 + 0 \times 6$
			48	96	456	$19 \times 4 + 76 \times 5$
		44.1	64	128	933	$83 \times 7 + 44 \times 8$
			128	256	1,866	$174 \times 7 + 81 \times 8$
			192	384	2,799	$265 \times 7 + 118 \times 8$
		32	3	6	91	$4 \times 18 + 1 \times 19$
			6	12	182	$5 \times 16 + 6 \times 17$
			24	48	728	$24 \times 15 + 23 \times 16$
25	48	1	2	12	12	$1 \times 12 + 0 \times 13$
		2	4	24	8	$3 \times 8 + 0 \times 9$
		7	14	84	6.461...	$7 \times 6 + 6 \times 7$
	44.1	8	16	135	9	$15 \times 9 + 0 \times 10$
		72	144	1,215	8.496...	$72 \times 8 + 71 \times 9$
	32	1	2	32	32	$1 \times 32 + 0 \times 33$
		2	4	64	21.333...	$2 \times 21 + 1 \times 22$
		17	34	544	16.484...	$17 \times 16 + 16 \times 17$
	29.97	48	20	40	439	$29 \times 11 + 10 \times 12$
			40	80	878	$70 \times 11 + 9 \times 12$
			220	440	4,829	$439 \times 11 + 0 \times 12$
		44.1	3200	6,400	83,253	$6,333 \times 13 + 66 \times 14$
			6400	12,800	166,506	$12,680 \times 13 + 119 \times 14$
		32	30	30	79	$8 \times 2 + 21 \times 3$
			60	60	158	$19 \times 2 + 40 \times 3$
			90	90	237	$30 \times 2 + 59 \times 3$

表7は、実施例が上述したアプローチ3に基づいている場合の、オーディオ・サンプル周波数とビデオ・フレーム・レートの様々な組み合わせに対するオーバーラップ・シーケンスを、幾つか示している。

(表7) MPEG-2レイヤIIとアプローチ3

10

20

30

f_v [Hz]	f_A [kHz]	M	N	O_t	$\overline{O}$	$p \times O + q \times (O+1)$
23.976	48	16	32	151	4.718...	$9 \times 4 + 23 \times 5$
	44.1	2,560	5,120	37,173	7.260...	$3,787 \times 7 + 1,333 \times 8$
	32	24	48	727	15.145...	$41 \times 15 + 7 \times 16$
24	48	2	4	19	4.75	$1 \times 4 + 3 \times 5$
	44.1	64	128	933	7.289...	$91 \times 7 + 37 \times 8$
	32	3	6	91	15.166...	$5 \times 15 + 1 \times 16$
25	48	1	2	12	6	$2 \times 6 + 0 \times 7$
	44.1	8	16	135	8.437...	$9 \times 8 + 7 \times 9$
	32	1	2	32	16	$2 \times 16 + 0 \times 17$
29.97	48	20	40	439	10.975	$1 \times 10 + 39 \times 11$
	44.1	3200	6400	83,253	13.008...	$6,347 \times 13 + 53 \times 14$
	32	30	30	79	2.633...	$11 \times 2 + 19 \times 3$

10

【 0 0 5 3 】

MPEG-2レイヤ I

MPEG-2 レイヤ I をエンコード方式として用いる場合、

 $k=12$

および

20

 $t=32$

である。アプローチ3を用いることにより、表8に示されているシーケンスが得られる。

(表8) MPEG-2レイヤ I とアプローチ3

f_v [Hz]	f_A [kHz]	M	N	O_t	$\overline{O}$	$p \times O + q \times (O+1)$
23.976	48	16	96	151	1.572...	$41 \times 1 + 55 \times 2$
	44.1	2,560	12,800	6,453	0.504...	$6,347 \times 0 + 6,453 \times 1$
	32	24	96	151	1.572...	$41 \times 1 + 55 \times 2$
24	48	2	12	19	4.75	$5 \times 1 + 7 \times 2$
	44.1	64	384	933	2.429...	$219 \times 2 + 165 \times 3$
	32	3	12	19	1.583...	$5 \times 1 + 7 \times 2$
25	48	1	5	0	0	$5 \times 0 + 0 \times 1$
	44.1	8	40	39	0.975	$1 \times 0 + 39 \times 1$
	32	1	4	8	2	$4 \times 2 + 0 \times 3$
29.97	48	20	100	199	1.99	$1 \times 1 + 99 \times 2$
	44.1	3,200	12,800	6,453	0.504...	$6,347 \times 0 + 6,453 \times 1$
	32	30	90	79	0.877...	$11 \times 0 + 79 \times 1$

30

【 0 0 5 4 】

平均冗長は、レイヤ IIを用いた場合よりも、かなり小さくなることに留意すべきである。

40

【 0 0 5 5 】

MPEG-1

各実施例に適用可能な別の単純化した例は、MPEG-1をエンコード方式の基礎として用いたものである。この場合、MPEG-1の2本の（例えば、ステレオ）チャンネル上の上限は、

 n

チャンネルまで拡張できる。したがって、各チャンネルには、利用可能なビット合計と、チャンネルごとのオーディオ・コンテンツとに依存したビット割り付けが可能である。

【 0 0 5 6 】

アルゴリズム

50

以下のセクションでは、アプローチ3によるオーバーラップの計算に適用可能なアルゴリズムを説明する。

【0057】

エンコーディング

実施例のストリームを作り出すためのエンコーダは、予め決定されている構造の一連のフレームを作り出す。各フレーム

j

は、以下の表9に示されている構造を有する。表中、

k

はブロック総数であり、

$H(j)$

は先頭のオーバーラップ内のブロック数であり、かつ

$T(j)$

は末尾のオーバーラップ内のブロック数である。

(表9)

$H(j)$	$k - [H(j) + T(j)]$	$T(j)$
--------	---------------------	--------

$T(j) = H(j+1)$

であることに留意されたい。

【0058】

エンコーダは、

N, O

および

q

の値を知っていれば、以下のアルゴリズムを用いて、先頭のオーバーラップを正確に計算することができる。

```
while (new frame) {
    if (counter >= N || counter == 0) {
        overlap = O + 1;
        counter = counter % N;
    }
    else overlap = O;
    return (overlap);
    counter = counter + q;
}
```

MPEG-2レイヤ IIの場合、

$f_v = 24Hz$

および

$f_A = 48kHz$

であり、表7から、

10

20

30

40

50

$$N = 4, O = 4$$

および

$$q = 3$$

が得られる。これにより、先頭のオーバーラップのシーケンスが、5、4、5、および5、またはこの任意の円形シフトとして生成される。

【0059】

あらゆるオーディオ・フレームには、そのサイズを示すためのタグを付けられなければならない。上述した方式の場合、先頭のオーバーラップは、長さが

10

O

または

$$O + 1$$

になることしかできない。したがって、1ビットのタグを用いて、短いフレームと長いフレームとを区別することができる。ビデオ・シーケンス内のフレーム

j

の有効サイズ

20

$$F(j)$$

は、下式により与えられる。

【0060】

$$F(j) = k - H(j+1)$$

式 16

【0061】

あらゆるブロックには、その冗長度を示すタグを付けられなければならない。上述の方式の場合、ブロックは、冗長または非冗長の何れかにしかなることができない。したがって、1ビットのタグを用いて、冗長ブロックと非冗長ブロックとを区別することができる。

30

【0062】

記録と伝送

冗長な情報はエンコードされなければならないが、すべてを伝送する必要はない。このことにより、伝送されるストリーム内のビット・レートが節約される。ビデオ・シーケンス内で記録または伝送されるブロックの最小の総数

$B_{\min}$

は、下式により与えられる。

【0063】

$$B_{\min} = \left(k - \left\lceil \frac{O_T}{N} \right\rceil \right) \times N + p$$

式 17

【0064】

オーディオ・フレームごとの余分な冗長ブロックは、エンコードされたストリームを編集するために必要となる場合がある。この場合、ビデオ・シーケンス内で記録または伝送される、ブロックの最大総数

$B_{\max}$

50

は、下式により与えられる。

$$B_{MAX} = \left(k - \left\lfloor \frac{O_T}{N} \right\rfloor \right) \times N + p \quad \text{式 18}$$

【0065】

位相

φ

は、ビデオ・シーケンス内の最初のビデオ・フレームに対する、エンコードされたストリームのブロック内の相対的な開始部分を示すように定義してもよい。 10

【0066】

φ

の適切な選択方法は、下式の通りである。

$$\varphi = \left\lceil \frac{O}{2} \right\rceil \quad \text{式 19}$$

【0067】

さらに、エンコーダはヌル・パディング

20

Q

を生成して、IEC61937規格に準じたストリームを完了させる。パディングの長さは、ペイロードの長さに依存するのみならず、ビデオ境界も考慮に入れて、エンコードされたストリーム内にもたらされてしまう累積エラーをなくさなければならない。

【0068】

編集

本実施例に従ってエンコードされたストリームの編集は、フレームを追加、除去、または添付することにより、ビデオ・フレームの境界で行ってもよい。デコーダは、生成されてしまう可能性があるエラーを、デコーダ内で利用可能な (30

f_A

と

f_V

の値などの) 情報、またはエンコーダにより生成される (サイズ・タグなどの) 情報を用いて編集することにより修正する。編集の結果、情報をさらに記録または伝送する必要はなくなる。さらに、デコーダ内の再構築フィルタ・バンクを用いて、編集点でのクロスフェードを設けてもよい。 40

【0069】

デコーディング

ストリームをデコードするためのデコーダは、現在のフレーム

j

に対して予想される有効なサイズ

$F(j)$

を計算する。さらに、このデコーダは、入来するフレームからサイズ・タグを読み取って、実際の有効なサイズ 50

G(j)

を決定する。

【0070】

オーディオ・フレーム内のブロックは、冗長または非冗長という2つの状態の内の1つを有してもよい。非冗長なブロックは、サブバンド・サンプルに記録、伝送、およびデコードされる。（末尾のオーバーラップ内の最初の冗長ブロックなどの）冗長ブロックを記録しかつ伝送して、デコード処理を容易化してもよい。しかしながら、冗長ブロックがサブバンド・サンプルにデコードされることは決してない。

【0071】

オーバーラップ・ブロックの状態を変更させるために、

NOP, DROP, APPEND

および

SHIFT

という4つの演算子が定義されている。

NOP; NOP

演算子は、ブロックの状態を変化させない。

DROP; DROP

演算子は、先頭のオーバーラップからの最初の非冗長ブロックを冗長ブロックに変化させる。

APPEND; APPEND

演算子は、末尾のオーバーラップからの最初の冗長ブロックを非冗長ブロックに変化させる。

SHIFT; SHIFT

演算子は、

DROP

演算子と

APPEND

演算子との両方を組み合わせたものである。

【0072】

本発明を具現化するストリーム内のフレームをサブバンド・サンプルにデコードすることは、マッピングと呼ばれている。非冗長なブロックしかサブバンド・サンプルにマッピングされない。入来するフレームが予想されたよりも大きな場合、

DROP

演算子が適用される。逆に、入来するフレームが予想されたよりも小さい場合、

APPEND

演算子が適用される。実際のサイズが予想されたサイズに等しい場合、デコーダは、直前のフレームを対象にする。直前のフレームがアペンドまたはシフトされている場合、

SHIFT

演算子が適用される。その他の場合、入来するフレームは変更されずにマッピングされる。

【0073】

同期化エラー

本発明を具現化するストリームは、長い（つまりタグを付けられた）フレームと短いフレームをエンコード時に定義されたシーケンスに直すことにより、ビデオ・フレーム長

$$1/f_v$$

に等しいオーディオ・フレーム平均有効長

$$\overline{F}$$

を作り出すことに基づいている。直前に定義された長いフレームと短いフレームのシーケンス

10

$$F(j)$$

を、デコード時に再生するために必要な冗長性は、編集後に入来するフレームの実際の長さ

$$G(j)$$

に関わらず、編集点でフレームをオーバーラップさせることにより得られる。編集時には、フレームを交換することによる、ブロック内の同期化エラー

$$\varepsilon(j)$$

20

は、下式により表すことができる。

$$\varepsilon(j) = \left(j \times \frac{p}{N} \right) - \left\lfloor j \times \frac{p}{N} \right\rfloor$$

式 20

下式のように常に書くことができる。

$$j \times p = u + N \times v$$

式 21

ここでは、

$$u \in \{0, 1, 2, \dots, N-1\}$$

30

および

$$v \in \{0, 1, 2, \dots, p\}$$

である。代入により、下式が得られる。

$$\varepsilon(j) = \frac{u}{N}$$

式 22

40

式中、

$$0 \leq \varepsilon_{MAX} < 1 - 1/N$$

である。デコード後、直ちに、これらの冗長性は、上述した

$$NOP$$

演算子、

$$DROP$$

50

演算子、

APPEND

演算子、および

SHIFT

演算子によって、適切に破棄される。さらに、

DROP

演算または

SHIFT

10

演算により、入来するフレーム

G(j)

を1ブロックだけ遅延させてもよい。したがって、この処理によりもたらされる同期化エラーの合計

δ

は、下式のように定められることを示すことができる

20

$$\Delta t = 0 \Rightarrow \delta \in \left[0, 1 - \frac{1}{N} \right) \wedge \Delta t = -1 \Rightarrow \delta \in \left[-1, -\frac{1}{N} \right) \quad \text{式 23}$$

ここで、端点は、以下の通りである。

【0074】

$$-1 \leq \delta_{MAX} < 1 \quad \text{式 24}$$

【0075】

カスケードニング

損失の多いエンコーディングとデコーディングのカスケードニング・レベルが幾つもあると、信号を劣化させてしまう場合がある。しかしながら、コントリビュション時と配信時における低い圧縮レートの使用、圧縮された信号に関するメタデータの使用、および特殊技術を採用して、この劣化をエンドユーザに知覚されない状態を保つことができる。本発明の技術分野の当業者には、（例えば、「カスケードニングされた音響心理学的なコード化におけるオーディオ品質の維持 (Maintaining Audio Quality in Cascaded Psychoacoustic Coding)」 (Warner R. Th. ten Kate 101st AES Convention 1996 November 8-11) に記述されているような)、MPEGエンコーディングに適用可能な様々な方法が知られている。これらの方法を本発明の実施例に用いて、DTVブロードキャスト・チェーン全体のオーディオ信号の品質を維持してもよい。

30

【0076】

本発明の例

40

ブロック構成

表10には、MPEG-2 レイヤ IIとアプローチ3のオーバーラップとに基づいた、映画とプロ用オーディオのための、本発明の実施例に従ってエンコードされたオーディオ・フレームのシーケンスが示されている。図8には、本発明の別の実施例による、ストリームをデコードした後の可能なブロック構成がすべて示されている。パラメータは、以下の通りである（上記の記号リストを参照されたい）。

ビデオ・フレーム・レート $f_V = 24\text{Hz}$ 、ビデオ・フレーム長 $t_V = 41.67\text{ms}$ ；

オーディオ・サンプリング周波数 $f_A = 48\text{kHz}$ 、オーディオ・フレーム長 $t_A = 24\text{ms}$ ；

$k = 36$ 個のブロック、 $t = 32$ 個のサンプル；

$M = 2$ 個のビデオ・フレーム、 $N = 4$ 個のオーディオ・フレーム；

オーバーラップ： $O_T = 19$ 個のブロック、 $\overline{O} = 4.75$ 個のブロック、 $O = 4$ 個のブロック、 $O+1 = 5$ 個のブロック；

$p = 1$ 個の短いオーバーラップ、 $q = 3$ 個の長いオーバーラップ；

$b = 31$ 個のブロック、 $b+1 = 32$ 個のブロック；

$B_{\min} = 125$ 、 $B_{\max} = 129$ 、 $\varphi = 2$ 個のブロック；

$\varepsilon_{\max} = 0.75$ 個のブロック、 $\delta \in \begin{cases} [0, 0.75) \leftarrow \Delta t = 0 \\ [-1, -0.25) \leftarrow \Delta t = -1 \end{cases}$

(表10)

j	1	2	3	4
$H(j)$	5	4	5	5
$F(j)$	32	31	31	31

【 0 0 7 7 】

IEC61937規格へのシステム適用

本発明を具現化するストリームを伝送するために適切な規格は、IEC61937規格「PCMによりエンコードされた非線形のオーディオ・ビット・ストリームのためのIEC 60958を適用したインターフェイス (Interface for non-linear PCM encoded audio bitstreams applying IEC 60958)」である。図7に示されている前の例に対するストリーム割り付けの場合、以下の通りである。

・ IEC61937によるフレーム長は、

$$(16 / 32) \times 3.072 \text{ Mbit/s} / f_V$$

である。

$$f_V = 24\text{Hz}$$

の場合、これは64,000ビットに対応する。

・ プリアンプル：Pa = F872h、シンクワード 1；Pb = 4E1Fh、シンクワード 2；Pc = バースト情報；Pd = ビット数 < 65,536、長さコード

・ データ・バーストの反復周期は、IEC60958によるフレーム数である。

・ VTRテープ編集後のオーディオとビデオ間の相対的なタイミング確度と、複数のスイッチ・システムのギャップによりもたらされる遅延により、2つのフレーム間に必要な最小限のギャップが決定される。このいわゆるスプライシング・ギャップは、ヌル・フレームをスタッフィングすることにより得てもよい。このことは、以下のように要約することができる。

・ スタッフィング = スプライシングのギャップ + バースト間隔；スプライシングのギャップ = テープ + スwitchの不正確度；バースト間隔 = $4 \times \text{IEC60958 "0" サブフレーム}$ 、各 $4096 \times \text{IEC60958フレーム}$

・ バーストペイロード：

システム・フレーム = $(N/M) \times [\text{システム・サブフレーム} - \text{先頭のオーバーラップ}]$ ；

$$N = 4; M = 2; N/M = 2$$

【 0 0 7 8 】

本発明を具現化するストリームが、384 kbit/sの5.1チャンネル用のMPEG-2 レイヤ IIに

10

20

30

40

50

基づいている場合、システムは最大で45,504ビット $(2 \times [(1,152 - 4 \times 32) \times 384 / 48 + (2,047 - 4 \times 32 / 1,152 \times 2,047) \times 8] + 0)$ を必要とする。

【0079】

これに代えて、本発明を具現化するストリームが、6チャンネルの場合にはチャンネルごとに192 kbit/sとなる、MPEG-1 レイヤ IIの6チャンネル・バージョンに基づいている場合、最大で49,152ビット $(2 \times (1,152 - 4 \times 32) \times 6 \times 192 / 48 + 0)$ が必要となる。LFEチャンネルがフレームごとに12個のサンプルしか必要としないことを考慮に入れた場合、有効ビット・レートは、チャンネルごとに約230 kbit/sとなることになる。

【図面の簡単な説明】

【0080】

10

【図1】 DTVブロードキャストに伴う典型的なチェーン状の線図である。

【図2】 典型的なDTV製作サイトの主要コンポーネントを示す線図である。

【図3】 典型的なDTVネットワーク・サイトの主要なコンポーネントを示す線図である。

【図4】 本発明の一実施例での第一アプローチによる、エンコードされたストリーム内のオーディオ・フレームとビデオ・フレームの構成を示す線図である。

【図5】 本発明の一実施例での第二アプローチによる、エンコードされたストリーム内のオーディオ・フレームとビデオ・フレームの構成を示す線図である。

【図6】 本発明の一実施例での第三アプローチによる、エンコードされたストリーム内のオーディオ・フレームとビデオ・フレームの構成を示す線図である。

【図7】 IEC61937におけるNTSCと48 kHzのオーディオのための、本発明を具現化するMPEG-2レイヤ II に基づくストリームのビット割り当てである。

20

【図8】 本発明の実施例によりエンコードされたストリーム状態のブロックの構成に関する線図である。

【符号の説明】

【0081】

10…コントリビューション段

12…配信段

14…放出段

20…製作サイト

22…配信ネットワーク

30

24…ブロードキャスト・ネットワーク

30…配信ネットワーク

32…DTH衛星ブロードキャスト

34…地上ブロードキャスト

36…ケーブル・テレビのプロバイダ

40…加入者

50…第一スタジオ

52…第二スタジオ

54…エンコーダ

56…テープ

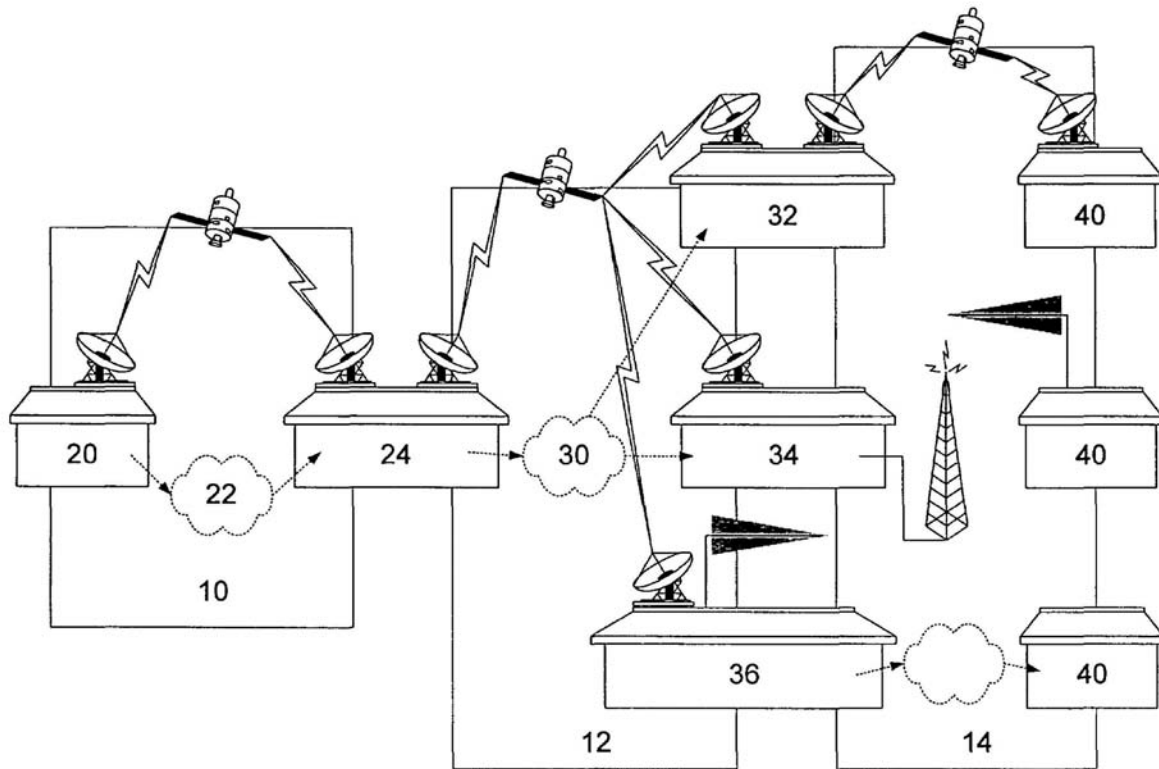
40

60…サプライサ

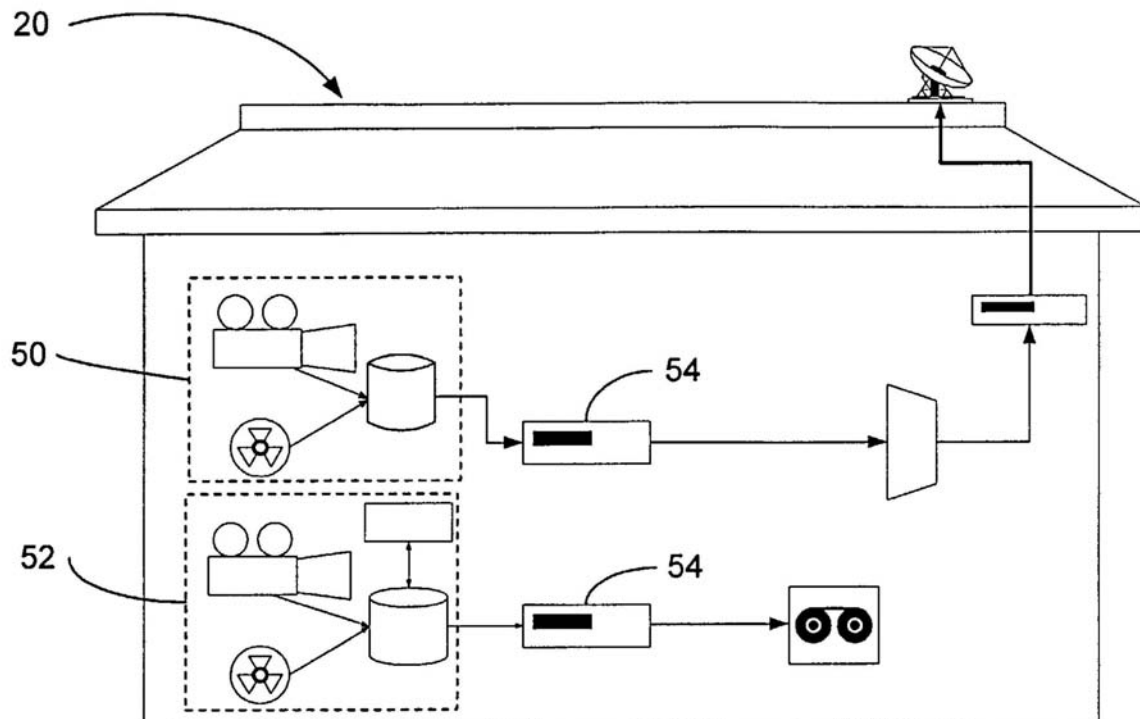
62…パケタイザ

64…ローカル・スタジオ

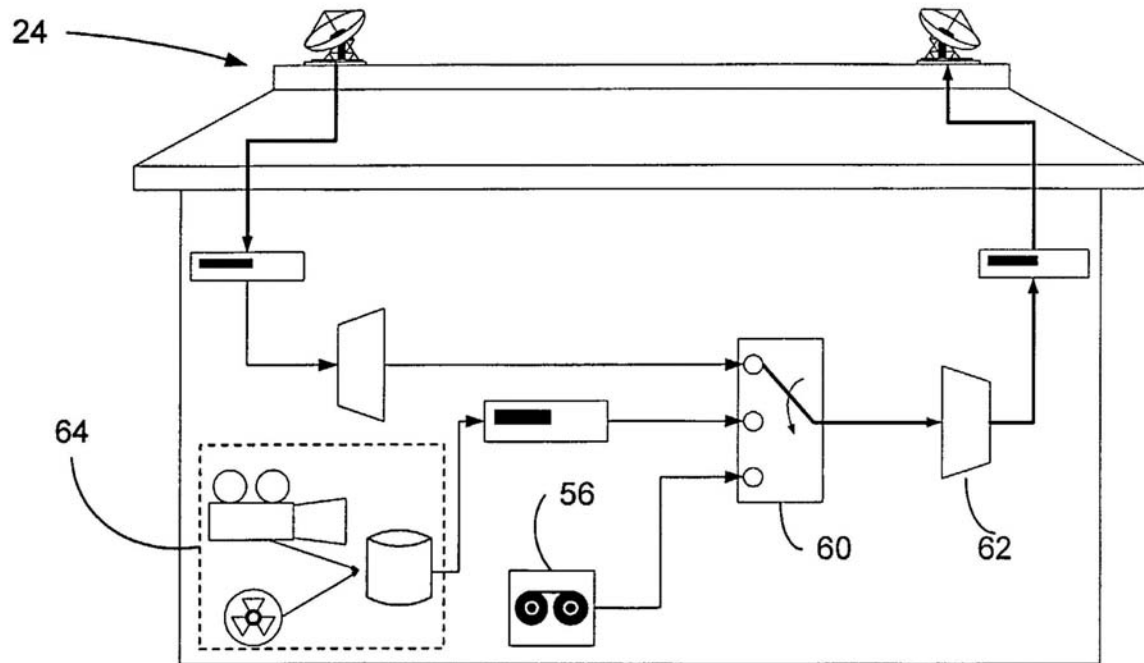
【図 1】



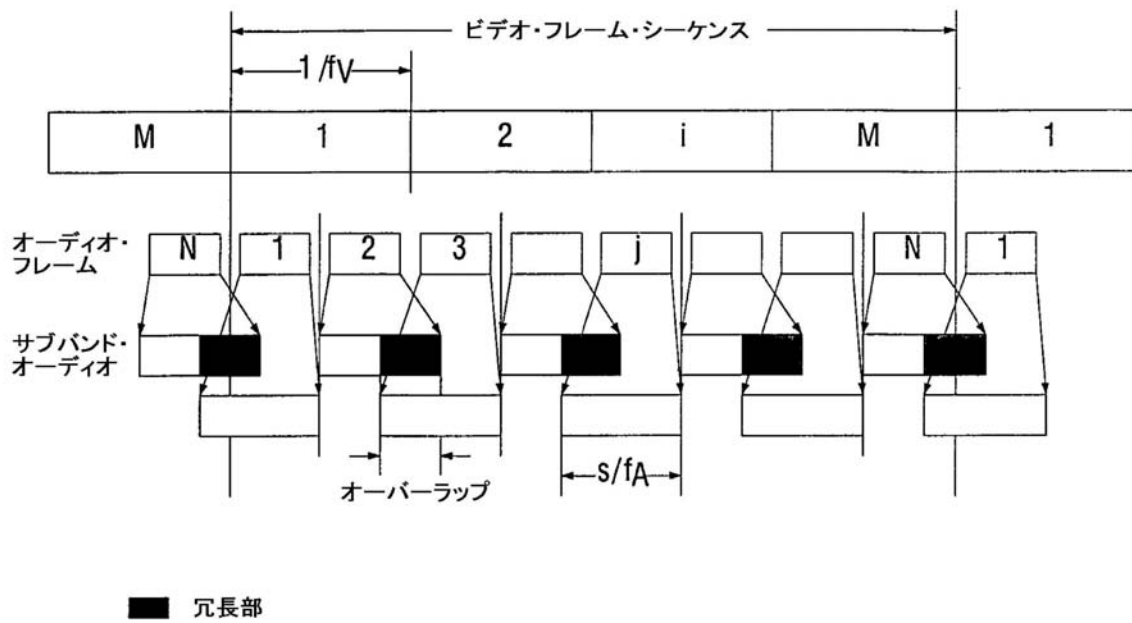
【図 2】



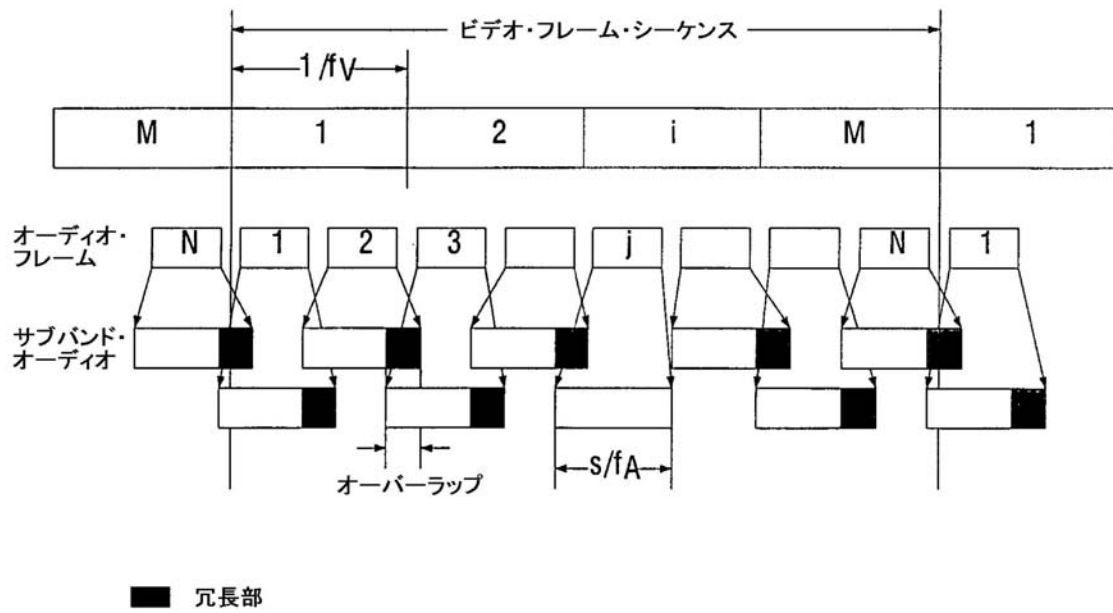
【図 3】



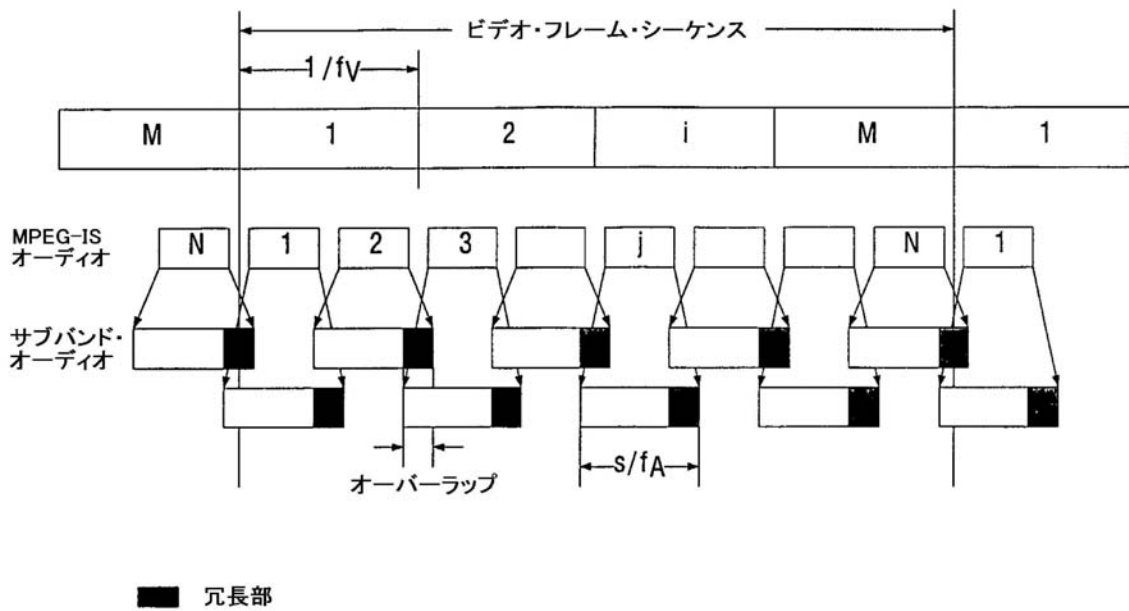
【図 4】



【図 5】



【図 6】



[illegible]

ビデオ・シーケンス

ビデオ・フレーム i

ビデオ・フレーム $i+1$

V1.1 = 32
タグ

V1.2 = 31

V1.3 = 31

V1.4 = 31

V1.5 = 32
タグ

V1.6 = 31

V2.1 = 32
タグ

V2.2 = 31

V1.7 = 31

V1.8 = 31

V1.9 = 32
タグ

V1.10 = 31

V1.11 = 31

V1.12 = 31

V2.3 = 31

V2.4 = 31

-1 0 3 30 34 61 66 92 93 97 123 125

冗長ブロック

ドロップ(先頭のオーバーラップからの非冗長なブロック)

アペンド(末尾のオーバーラップ内の他の冗長ブロック)

シフト = ドロップ + アペンド

フロントページの続き

- (72)発明者 ボルツェ, トマス
オランダ国 5656 アー アー アインドーフェン プロフホルストラーン 6
- (72)発明者 ディレン, パオルス, ハー., アー.
オランダ国 5656 アー アー アインドーフェン プロフホルストラーン 6
- (72)発明者 ヴァン デ ケルコフ, レオン, エム.
オランダ国 5656 アー アー アインドーフェン プロフホルストラーン 6

審査官 井上 健一

- (56)参考文献 特表2004-516505 (JP, A)
特開2000-236262 (JP, A)
特開平09-261070 (JP, A)

- (58)調査した分野(Int.Cl., DB名)
G10L 19/00-19/14