

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第4606987号
(P4606987)

(45) 発行日 平成23年1月5日 (2011.1.5)

(24) 登録日 平成22年10月15日 (2010.10.15)

(51) Int. Cl.	F I
HO 4 N 1/41 (2006.01)	HO 4 N 1/41 B
HO 4 N 7/30 (2006.01)	HO 4 N 7/133 Z
GO 6 T 9/00 (2006.01)	GO 6 T 9/00

請求項の数 15 外国語出願 (全 11 頁)

(21) 出願番号	特願2005-290007 (P2005-290007)	(73) 特許権者	597067574
(22) 出願日	平成17年10月3日 (2005.10.3)		ミツビシ・エレクトリック・リサーチ・ラ
(65) 公開番号	特開2006-166413 (P2006-166413A)		ボラトリーズ・インコーポレイテッド
(43) 公開日	平成18年6月22日 (2006.6.22)		アメリカ合衆国、マサチューセッツ州、ケ
審査請求日	平成20年8月8日 (2008.8.8)		ンブリッジ、ブロードウェイ 201
(31) 優先権主張番号	11/002, 817		201 BROADWAY, CAMBR
(32) 優先日	平成16年12月2日 (2004.12.2)		IDGE, MASSACHUSETTS
(33) 優先権主張国	米国 (US)		O2139, U. S. A.
		(74) 代理人	100110423
			弁理士 曾我 道治
		(74) 代理人	100084010
			弁理士 古川 秀利
		(74) 代理人	100094695
			弁理士 鈴木 憲七
最終頁に続く			

(54) 【発明の名称】 スケーラブルな画像トランスコーディング方法

(57) 【特許請求の範囲】

【請求項 1】

画像の符号化入力ビットストリームを部分的に復号化して、前記符号化入力ビットストリームの構造を得ること、
前記画像の関心領域の座標を指定すること、
プログレッシブパラメータを規定すること、
及び
部分的に復号化された入力ビットストリームを前記構造、前記座標、及び前記プログレッシブパラメータに従って、符号化出力ビットストリームとして符号化することを含み、
前記符号化入力ビットストリームは、複数の画質レイヤを含み、各画質レイヤは複数の解像度レベルを含み、各解像度レベルは複数のコンポーネントを含み、各コンポーネントは複数のプレシントを含み、各プレシントはパケット分割位置、ヘッダ長、及び本文の長さを含むとともに、
プログレッシブパラメータmが前記画像の背景の画質を指定し、パラメータnが前記関心領域の画質を指定し、
前記プログレッシブパラメータm及びnを、
前記関心領域の場合、
 $0 \leq m < n \leq \text{最大画質}$ とし、さもなければ
前記背景の場合、
 $0 \leq m \leq \text{最大画質}$ とし、

- 背景なしの場合、
 $m = 0$ とする
 に従って制約することをさらに含む、
 スケーラブルな画像トランスコーディング方法。
- 【請求項 2】
 前記符号化入力ビットストリーム及び前記符号化出力ビットストリームは、ともに J P E G 2 0 0 0 ビットストリームである
 請求項 1 記載の方法。
- 【請求項 3】
インデックス番号を用いて前記プレシントをインデックス付けすることをさらに含む 10
 請求項 1 記載の方法。
- 【請求項 4】
符号化中に前記インデックス番号を用いて前記構造にランダムにアクセスすることをさ
らに含む
 請求項 3 記載の方法。
- 【請求項 5】
 符号化中に前記インデックス番号を用いて前記構造に直接アクセスすることをさらに含
 む
 請求項 3 記載の方法。
- 【請求項 6】 20
トランスコーディング中にユーザによって前記座標を規定することをさらに含む
 請求項 1 記載の方法。
- 【請求項 7】
 トランスコーディング中に外部ソースによって前記座標を規定することをさらに含む
 請求項 1 記載の方法。
- 【請求項 8】
前記外部ソースは、オブジェクト認識を用いて前記関心領域の前記パラメータを求める
請求項 7 記載の方法。
- 【請求項 9】 30
前記外部ソースは、パターン認識を用いて前記関心領域の前記パラメータを求める
請求項 7 記載の方法。
- 【請求項 10】
前記パラメータを外側の、前記複数のプレシントの最も近い境界に丸めることをさら
に含む
 請求項 1 記載の方法。
- 【請求項 11】
トランスコーディング中にユーザによって前記プログレッシブパラメータを指定するこ
とをさらに含む
 請求項 1 記載の方法。
- 【請求項 12】 40
前記符号化入力ビットストリームは、関心領域パケットと背景パケットを含み、
 $m = 0$ の場合に全ての背景パケットを前記符号化出力ビットストリームにおいて空パケ
ットに変換することをさらに含み、1つの空パケットは1バイトのパケットヘッダを有し
、1番目のビットがゼロに設定され、ペイロードがなくなる
 請求項 1 記載の方法。
- 【請求項 13】
前記関心領域パケットの場合に、最も低い画質レイヤ1から画質レイヤnまでの関心領
域パケットのみを抽出し、前記抽出した関心領域パケットを前記符号化出力ビットストリ
ームにおいて符号化するとともに、他の全ての関心領域パケットを前記符号化出力ビット
ストリームにおいて空パケットに変換すること、 50

及び、

前記背景パケットの場合に、前記最も低い画質レイヤ1から画質レイヤmまでの背景パケットのみを抽出し、前記抽出した背景パケットを前記符号化出力ビットストリームにおいて符号化するとともに、他の全ての背景パケットを前記符号化出力ビットストリームにおいて空パケットに変換することをさらに含む

請求項12記載の方法。

【請求項14】

前記符号化入力ビットストリームは、複数の画像を含むビデオであり、前記復号化するステップと、前記指定するステップと、前記規定するステップと、前記符号化するステップとは画像毎に実行される

請求項1記載の方法。

【請求項15】

前記符号化入力ビットストリームは、複数のパケットを含み、前記復号化することは前記パケットのヘッダのみに適用される

請求項1記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、包括的には画像及びビデオの符号化に関し、特に関心領域を含む画像のトランスコーディングに関する。

【背景技術】

【0002】

ビデオ監視、携帯電話、デジタルカメラ、プリンタ、スキャナ、ファクシミリ、コピー機、医用画像、衛星画像、インターネット、及び複合文書により、画像及びビデオの用途に対する要求が高まった。しかし、帯域幅、記憶装置及びプロセッサのような資源には限りがあるため、高画質の画像は不可能である場合が多い。画像の画質は、画像の画素数と、各画素に割り当てられるビット数とに依存する。例えば、1画素が24ビットで表される 1024×1024 画素の画像は25Mbの高画質カラー画像となり、1画素が1ビットで表される 10×10 画素の画像は100ビットの低画質の白黒「サムネイル」画像となる。

【0003】

1つの解決策は、画像中の関心領域(ROI)を背景(BG)と区別する。ROIの符号化にはBGよりも多くのビットが用いられる。ROIにBGよりも多くのビットを割り当てることにより、符号化画像中のROIの知覚される解像度及び画質を低下させることなく、画像の符号化に用いられる総ビット数を減らすことができる。ビット数が少なくなれば必要な資源が削減される。

【0004】

1つのROI符号化方法は、ROIのウェーブレット変換係数を選択的に増大させる。Atsumi等著「階層木における集合分割法(SPIHT)に基づく有損失/無損失の関心領域の画像符号化(Lossy/lossless region-of-interest image coding based on set partitioning in hierarchical trees)」IEEE Proceeding of ICIP, October 1998。ROIはまた、より高い優先度で転送される。しかし、スケーリング値によっては、ROIがBGに溶け込んで見える可能性がある。したがって、デコーダは、ROIをBGと区別するために形状情報も必要とする。

【0005】

JPEG2000規格は、ROI符号化のためのmax-shift法を規定する。ISO/IEC 15444-1「情報技術-JPEG2000画像符号化方式-パート1:基本符号化方式(Information technology - JPEG 2000 image coding system - Part 1: Core coding system)」1st Ed., 2000。JPEG2000規格は、色変換、量子化、ウェーブレット変換、プログレッシブなビットプレーン符号化、及びエントロピー符号化を用いる。符号

10

20

30

40

50

化画像は、階層パケットストリームとして転送される。J P E G 2 0 0 0では、符号化中に出力画像のサイズ及び画質が選択される。m a x - s h i f t法は、R O Iを重複のないビットプレーンにスケーリングすることによって、R O Iを背景から分離する。Skodras等著「J P E G 2 0 0 0の静止画圧縮規格(The JPEG 2000 still image compression standard)」IEEE Signal Processing Magazine, September, 2001。スケーリング値は、R O Iに関連する最小係数が背景の最大係数よりも確実に大きくなるように十分に大きい。デコーダは、スケーリング値を受け取ると、R O I係数をその大きさに識別する。m a x - s h i f t法により、R O Iの形状情報をデコーダに明示的に送信することなく、任意形状のR O Iの符号化が可能になる。しかし、m a x - s h i f t符号化では、R O Iの境界の画定に必要な余分な符号ブロックによりオーバーヘッドが増加する。

10

【0006】

別の方法は、R O Iの相対的な重要度に合わせてプレーン毎にビットをシフトする。Wang等著「ビットプレーン毎のシフト(B b B S h i f t) - J P E G 2 0 0 0における関心領域の画像符号化に対する提案(Bitplane-by-bitplane shift (BbBShift) - A suggestion for JPEG 2000 Region of Interest image coding)」IEEE Signal Processing Letters, Vol. 9, No. 5, May 2002。しかし、B b B S h i f t法はJ P E G 2 0 0 0規格に準拠していない。

【0007】

別の方法は「部分的な上位ビットプレーンシフト(partial significant bit-planes shift)」(P S B S h i f t)と呼ばれる。Liu等著「J P E G 2 0 0 0における関心領域の新たな画像符号化方法：部分的な上位ビットプレーンシフト(A new JPEG 2000 region-of-interest image coding method: partial significant bitplanes shift)」IEEE Signal Processing Letters, Vol. 10, No. 2, February 2003。P S B S h i f t法は、R O Iの高画質を維持しようとする。P S B S h i f t法もやはりJ P E G 2 0 0 0規格に準拠していない。

20

【0008】

上記のR O I符号化方法はすべて静的符号化を用いる。すなわち、R O Iは「符号化」中に規定される。これは、R O I情報が「復号化」中にしか利用できない場合に問題となる。例えば、閲覧者がR O Iを指定することを望む。これもまた、R O I情報が外部ソースによって動的に供給される場合に問題となる。例えば、オブジェクト追跡のような外部プロセスは、画像を復号化前に解析してR O Iを判定する。

30

【0009】

動的なR O I符号化方法が、Rosenbaum等著「J P E G 2 0 0 0におけるフレキシブルで、動的で、且つコンプライアントな関心領域の符号化(Flexible, dynamic and compliant region of interest coding in JPEG 2000)」IEEE Proceeding of ICIP, Rochester, New York, September, 2002に記載されている。この方法は、インタラクティブ環境における動的なR O I情報に対処する。この方法は、J P E G 2 0 0 0規格によって規定されるプレシント／レイヤ構造を用いて、各レイヤにおけるプレシントの優先度を定める。この方法はレイヤを動的に挿入する。R O Iパケットは同じレイヤに残り、他のパケットは1つ上のレイヤにシフトされる。しかし、動的なレイヤ挿入にはパケットヘッダの再符号化が必要になる。これはレート歪みの再計算を必要とするが、この特徴はリアルタイム画像伝送用途には望ましくない。さらに、この方法はJ P E G 2 0 0 0規格に準拠している。

40

【発明の開示】

【発明が解決しようとする課題】

【0010】

したがって、従来技術による符号化方法の問題のために、パケットヘッダの再符号化を回避するとともに、R O Iの符号化をフレキシブルで、動的で、且つ計算の複雑度が低いものにする新たな符号化機構を提供することが望ましい。

【課題を解決するための手段】

50

【0011】

1つの方法では、符号化入力ビットストリームの形態の画像又はビデオをトランスコードする。パケットの形態の入力ビットストリームは画質レイヤ（quality layer）を含み、各画質レイヤは解像度レベルを含み、各解像度レベルはコンポーネントを含み、各コンポーネントはプレシントを含み、各プレシントはパケット分割位置、ヘッダ長、及び本文の長さを含む。

【0012】

入力ビットストリームは部分的に復号化されて、インデックス付きの符号化入力ビットストリーム構造が得られる。画像中の関心領域の座標が指定されて、プログレッシブパラメータが規定される。

10

【0013】

部分的に復号化された入力ビットストリームは次に、構造、座標、及びプログレッシブパラメータに従って符号化され、符号化出力ビットストリームが得られる。

【0014】

トランスコーディング中、入力ビットストリームの関心領域パケット及び背景パケットがプログレッシブパラメータに従って出力ビットストリームに選択的に変換される。例えば、出力ビットストリームに背景パケットは含まれないか、関心領域と、低画質レイヤからの背景パケットのみが含まれる。

【発明の効果】

【0015】

20

本発明は、符号化ビットストリームからの画像の復元を任意数の異なる方法において任意の所望の空間解像度及び画質で可能にする。本発明によるトランスコーディングは適応的且つスケラブルである。本発明は、関心領域の高画質及び高解像度を維持することができる。画像の残りの部分は、画質を落とすか、又は完全に削除して、所望の帯域幅を達成することができる。本発明による方法は、従来技術の方法に比べて複雑度が低く効率が低い。

【0016】

従来技術と異なり、本発明はパケットの完全な復号化及び再符号化を行わない。その代わりに、本発明は、パケットを選択的に削除するか、又は、空パケットを用いて、ROIの優先度を効果的に上げる。これにより、リアルタイムでのトランスコーディング用途が可能になる。

30

【発明を実施するための最良の形態】

【0017】

図1は、本発明による、関心領域（ROI）を含む画像をトランスコードするシステム及び方法を示す。ROIは画像中の矩形領域である。本発明のシステムに対する入力は、符号化ビットストリーム101、例えばJPEG2000ビットストリームである。このビットストリームはデータパケットシーケンスの形態である。画像は、有損失技法又は無損失技法を用いて圧縮することができる。

【0018】

ビットストリーム101の符号化中、JPEG2000規格においてウェーブレット変換を使用して、画像を4つのサブバンド画像として解析する。サブバンド画像は、それらの異なる空間周波数特性を記述する係数を含む。周波数の最も低いサブバンド画像をさらに4つのより小さなサブバンドに分解する。このプロセスは、所望の画像解像度に対して必要に応じて繰り返すことができる。各サブバンド画像は、「符号ブロック」と呼ばれる重複のない矩形ブロックに分割される。各符号ブロックは個別に符号化されて、最終的な符号化ビットストリーム101が生成される。

40

【0019】

ビットストリームを効率的に編成するために、符号ブロックを各解像度レベル内で「プレシント」にグループ分けする。プレシント分割は、画像の特定の空間領域に対応するウェーブレット係数へのアクセスを容易にする。各プレシントはビットストリーム中

50

の1パケットを生成する。SNR-プログレッシブ精緻化を可能にするために、1つのパケットに含まれる情報を複数のレイヤに分散させる。各レイヤは検討されるプレシントからの一定量のデータを含む。

【0020】

ビットストリーム中の各画質レイヤ及び各解像度レベルから所望のパケットを抽出するために、本発明は、符号ブロックの算術的復号化を行わずにパケットヘッダ情報を部分的に復号化するパーサ110を提供する。こうして、ROIトランスコーディング120により、最小の計算複雑度で符号化出力ビットストリーム104を生成することが可能になる。

【0021】

図2に示すように、パーサ110は、タグツリーデコーダを用いて、入力ビットストリームの階層データ構造200を得る。

【0022】

ビットストリーム101中のパケットを部分的に復号化110して構造200を得る。階層構造は画質レイヤ(Qレイヤn)201を含み、各レイヤは解像度レベル(Rレベルm)202を含み、各解像度レベルはコンポーネント203を含み、各コンポーネントはプレシント204を含み、各プレシントはパケット分割位置205、ヘッダ長206、及び本文の長さ207を含む。インデックス番号により、プレシント情報205~207に直接アクセスすることができる。

【0023】

JPEG2000ビットストリームの複雑なアーキテクチャにより、全てのプレシントにアクセスすることは高くつく。したがって、本発明ではデータ構造200を、インデックス番号によりランダムに直接アクセスできるように設計する。

【0024】

図1に示すように、ROI座標112及びプログレッシブパラメータ113も例えばユーザ又は外部ソースによって指定及び規定される。

【0025】

構造200、ROI座標112、及びプログレッシブパラメータ113はトランスコーダ120に渡される。従来技術とは対照的に、ROI座標112は、符号化中ではなくトランスコーディング中に指定される。例えば、ユーザがROIをマーク付けするか、又は、オブジェクト認識又はパターン認識等の他の技法によって座標が供給される。認識器は、画像中のオブジェクト又はパターンを見つけ、そのオブジェクトの周囲にROIを当てはめて座標を得る。

【0026】

ROI座標はプレシント位置と一致しない場合がある。その場合、トランスコーディング120は、ROI座標を外側の最も近いプレシント境界に丸める。すなわち、矩形のROIを、プレシントによって規定される境界に合わせて大きくする。プログレッシブパラメータは、ユーザがトランスコーディングの前に規定して、背景(BG)を削除できるかを指示するか、又は、BG及びROIについてトランスコードする必要のある画質レベル数を指示することもできる。プログレッシブパラメータの使用については後述する。

【0027】

トランスコーダは、構造情報200及び座標112を用いて、入力ビットストリーム101中のBGパケット及びROIパケットを識別する。パケットをROI座標112及びプログレッシブパラメータ113に従って、符号化出力ビットストリーム104のパケットに再編成する。最終的な符号化出力ビットストリーム104は、JPEG2000規格に完全準拠する。

【0028】

図3は、ROI302と、残りの背景(BG)303とを含む例示的な画像301を示す。ROIの座標は、原点(O)並びにX軸及びY軸に対して規定される。

【0029】

図4に示すように、画像301は、ビットストリーム101の5つの「画質」レイヤ411～415からなるBGパケット401及びROIパケット402に符号化される。

【0030】

図5は、ビットストリーム101の5つのレイヤ411～415におけるROIパケット及びBGパケットの、いわゆる「レイヤー解像度－コンポーネント位置（LRCP）」プログレッションを用いた配置を1つの解像度レベル及び1つのコンポーネントについて示す。

【0031】

図6は、3つのレベル601～603を有する解像度マップ600を示す。プレシントは、特定の解像度レベルにおける画像中の空間領域を記述する。JPEG2000による分解では、各方向におけるプレシントのサイズは2のべき乗である。高解像度レベルの画像領域は、2で割ることによってその下の解像度レベルに伝わる。この手順は解像度レベル毎に繰り返される。

10

【0032】

図7は、LRCPプログレッションを用いた対応するビットストリーム700を示す。図7から、結果として得られるパケットは、所定の解像度レベルでの画像の特定の領域に関する情報を含むことが分かる。

【0033】

所望のROIを抽出し、BGとROIに対してスケーラブルな画質を指定するために、トランスコーダ120は2つのプログレッシブパラメータm及びnを用いる。パラメータmはBG303の所望の「画質」を指定し、パラメータnはROI302の所望の画質を指定する。パラメータm及びnは以下の制約を有する。

20

ROIの場合、

$0 \leq m < n \leq \text{最大画質}$ とし、さもなければ、

背景の場合、

$0 \leq m \leq \text{最大画質}$ とし、

背景なしの場合、

$m = 0$ とする。

【0034】

最後の制約の下では、背景全体が空パケットに変換され、ROIパケットのみがトランスコードされる。1つの空パケットは1バイトのパケットヘッダを有し、1番目のビットがゼロに設定され、ペイロードがなくなる。この場合、トランスコードされた出力ビットストリームに対する帯域幅要件が大幅に削減される。

30

【0035】

ROIパケットの場合、トランスコーダ120は、入力ビットストリームから、最も低い画質レイヤ1から画質レイヤnまでのROIパケットのみを抽出し、これらのパケットを符号化して出力ビットストリームを生成する。nより高いレイヤからのROIパケットは、符号化出力ビットストリームにおいて空パケットに変換される。

【0036】

BGパケットの場合、 $m = 0$ であれば、全てのBGパケットを空パケットに変換する。 $m > 0$ であれば、トランスコーダ120が入力ビットストリームの最も低い画質レイヤ1から画質レイヤmまでのBGパケットを抽出し、これらのパケットを符号化して出力ビットストリームを生成する。mより高いレイヤのBGパケットは全て、出力ビットストリームにおいて空パケットに変換される。

40

【0037】

パラメータm及びnの組み合わせを変えることで、ROI及びBGについてさまざまな画質のプログレッシブな結果を得ることができる。ユーザは、この特徴を利用して、視覚要件と通信帯域容量の間で調整を行うことができる。

【0038】

50

図 8 は、プログレッシブパラメータを $m = 1$ 及び $n = 4$ としてトランスコードする例示的な画像 800 を示す。画像 800 は、空パケット 801 と、ROI パケット 802 と、BG パケット 803 とを有する。

【0039】

図 9 は、対応する出力ビットストリーム 900 のパケットを示す。

【0040】

本発明を好ましい実施形態の例として説明してきたが、本発明の精神及び範囲内でさまざまな他の適応及び変更を行うことができることが理解される。したがって、添付の特許請求の範囲の目的は、本発明の真の精神及び範囲に入るそのような変形及び変更をすべて網羅することである。

【図面の簡単な説明】

【0041】

【図 1】 本発明による、画像をトランスコードするシステム及び方法のブロック図である。

【図 2】 本発明による階層ビットストリームの構造のブロック図である。

【図 3】 本発明による、関心領域を含む画像のブロック図である。

【図 4】 本発明による、トランスコードする階層パケットのブロック図である。

【図 5】 本発明による入力ビットストリームのブロック図である。

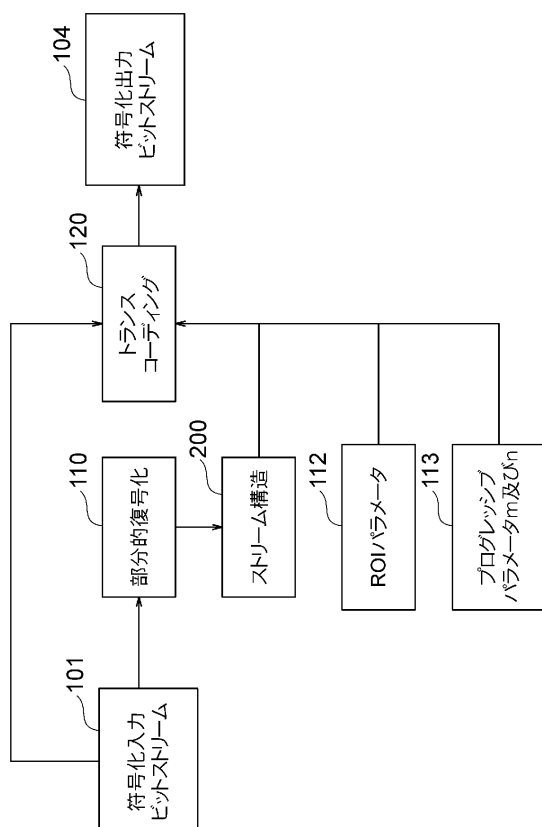
【図 6】 本発明による解像度レベルのブロック図である。

【図 7】 図 6 の解像度レベルに対応するビットストリームのブロック図である。

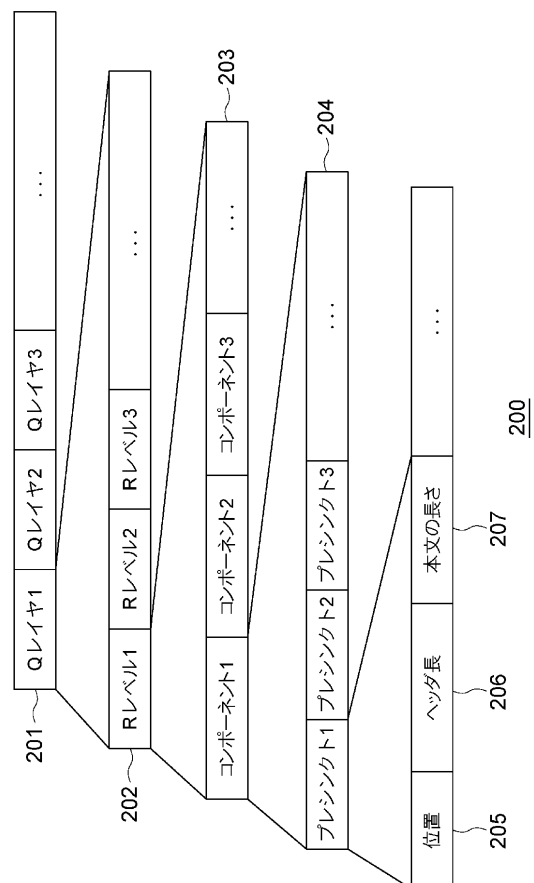
【図 8】 本発明による、トランスコードする画像のブロック図である。

【図 9】 本発明による出力ビットストリームのブロック図である。

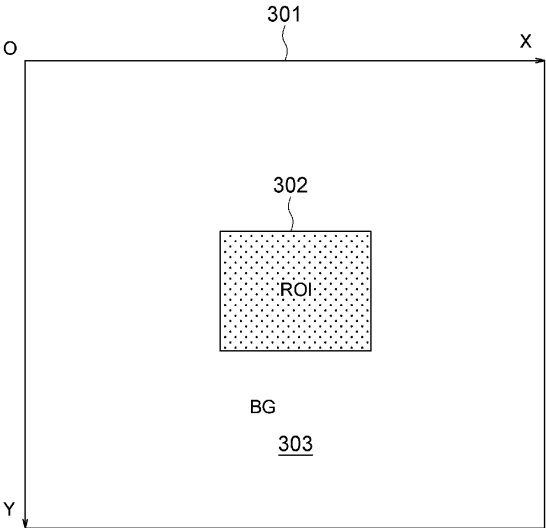
【図 1】



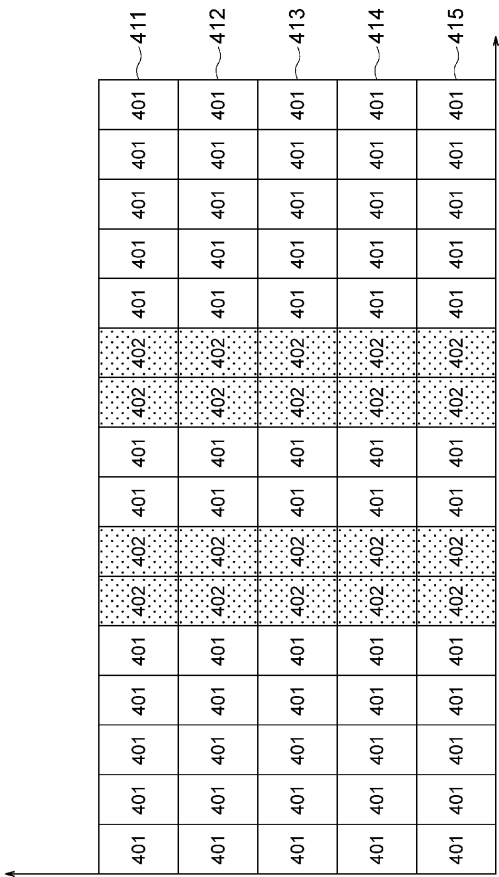
【図 2】



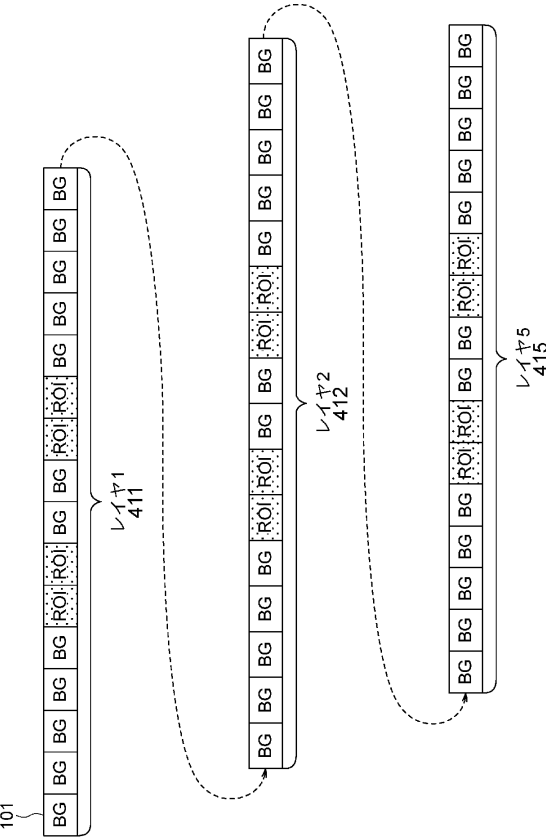
【図 3】



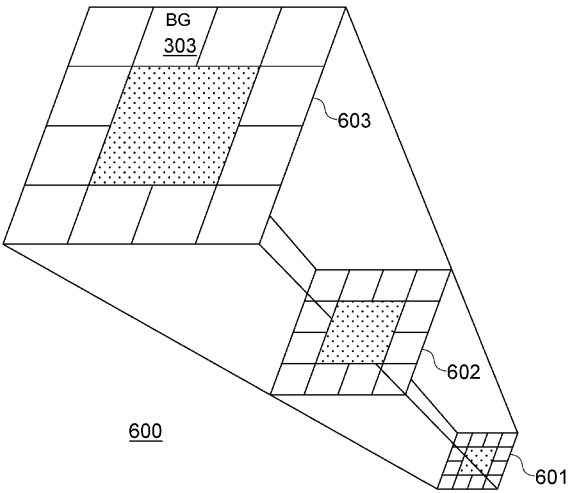
【図 4】



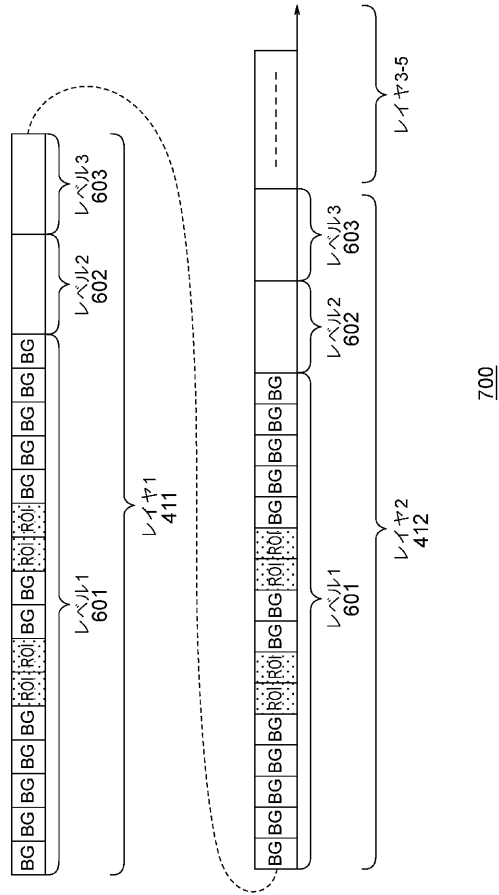
【図 5】



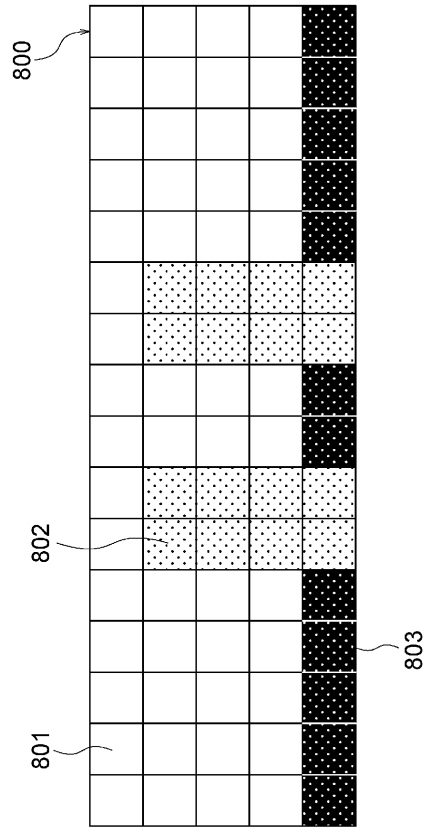
【図 6】



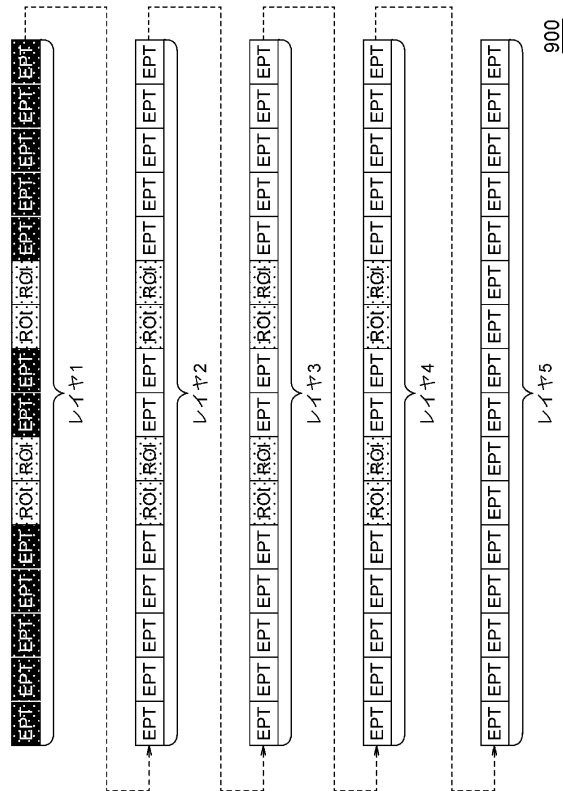
【図 7】



【図 8】



【図 9】



フロントページの続き

(74)代理人 100111648

弁理士 梶並 順

(72)発明者 ハオーソン・コン

アメリカ合衆国、マサチューセッツ州、ニュートン、リバービュー・アベニュー 209、アパートメント 36

(72)発明者 アンソニー・ヴェトロ

アメリカ合衆国、マサチューセッツ州、ケンブリッジ、ハーバード・ストリート 353、ユニット 45

(72)発明者 秦 淑彦

東京都千代田区丸の内二丁目2番3号 三菱電機株式会社内

(72)発明者 桑原 直樹

東京都千代田区丸の内二丁目2番3号 三菱電機株式会社内

審査官 富永 達朗

(56)参考文献 特開2004-242287 (JP, A)

特開2004-072655 (JP, A)

特開2004-134938 (JP, A)

特開2001-313944 (JP, A)

特開2001-218062 (JP, A)

特開2004-282345 (JP, A)

Rosenbaum R et al, FLEXIBLE, DYNAMIC AND COMPLIANT REGION OF INTEREST CODING IN JPEG2000, Image Processing. 2002. Proceedings. 2002 International Conference on, 2002年, vol.3, 101-104

(58)調査した分野(Int.Cl., DB名)

H04N 1/41

G06T 9/00

H04N 7/30