

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5508534号  
(P5508534)

(45) 発行日 平成26年6月4日(2014.6.4)

(24) 登録日 平成26年3月28日(2014.3.28)

(51) Int.Cl. F I  
H04N 19/00 (2014.01) H04N 7/13 Z

請求項の数 20 (全 16 頁)

(21) 出願番号	特願2012-527954 (P2012-527954)	(73) 特許権者	310021766 株式会社ソニー・コンピュータエンタテインメント 東京都港区港南1丁目7番1号
(86) (22) 出願日	平成22年8月30日 (2010.8.30)	(74) 代理人	100105924 弁理士 森下 賢樹
(65) 公表番号	特表2013-504256 (P2013-504256A)	(72) 発明者	リー、ファンージュ アメリカ合衆国、カリフォルニア州 94404、フォスター・シティー、イースト・ヒルスデイル・ブルバード 919
(43) 公表日	平成25年2月4日 (2013.2.4)	審査官	坂東 大五郎
(86) 国際出願番号	PCT/US2010/047178	(56) 参考文献	特開平06-022298 (JP, A)
(87) 国際公開番号	W02011/028666		最終頁に続く
(87) 国際公開日	平成23年3月10日 (2011.3.10)		
審査請求日	平成24年4月26日 (2012.4.26)		
(31) 優先権主張番号	12/553,069		
(32) 優先日	平成21年9月2日 (2009.9.2)		
(33) 優先権主張国	米国 (US)		

(54) 【発明の名称】 シーン切替検出

(57) 【特許請求の範囲】

【請求項1】

- a) 現在のピクチャ内の所与の部分領域における統計量を計算するステップと、
- b) 先行するピクチャ内の対応する位置にある部分領域周辺の2以上の部分領域のウィンドウを規定するステップと、
- c) 前記ウィンドウ内の2以上の部分領域における統計量の総和を計算するステップと、
- d) 前記総和と前記所与の部分領域における統計量との差分を計算するステップと、
- e) 前記総和と前記所与の部分領域における前記統計量との差分を用いて前記所与の部分領域がシーン切替領域か否かを決定するステップと、
- f) 前記現在のピクチャについてシーン切替領域の数を決定するステップと、
- g) 前記シーン切替領域の数から前記現在のピクチャがシーン切替ピクチャか否かを決定するステップと、
- h) 前記現在のピクチャがシーン切替ピクチャか否かを示す情報を格納または転送するステップとを含むことを特徴とする、1以上のデジタルピクチャを符号化する際のシーン切替を検出する方法。

【請求項2】

- i) 前記現在のピクチャがシーン切替ピクチャか否かを示す情報にしたがって、エンコーダを用いて前記現在のピクチャを符号化するステップをさらに含むことを特徴とする請求項1に記載の方法。

## 【請求項 3】

前記 i ) のステップは、

前記現在のピクチャがシーン切替ピクチャであることを前記情報が示す場合、前記現在のピクチャを I フレームとして符号化し、現在のピクチャがシーン切替ピクチャではないことを前記情報が示す場合、前記現在のピクチャを B フレームまたは P フレームとして符号化するステップを含むことを特徴とする請求項 2 に記載の方法。

## 【請求項 4】

前記現在のピクチャ内の複数の部分領域について、前記 a ) のステップから e ) のステップを繰り返すステップをさらに含むことを特徴とする請求項 1 に記載の方法。

## 【請求項 5】

各部分領域はマクロブロックであることを特徴とする請求項 1 に記載の方法。

## 【請求項 6】

前記 a ) のステップは、現在のピクチャ内の各部分領域における画素値の平均値または画素値の分散を計算するステップを含むことを特徴とする請求項 1 に記載の方法。

## 【請求項 7】

前記ウィンドウは、先行するピクチャ内の対応する位置にある部分領域に隣接する 1 以上の部分領域を含むことを特徴とする請求項 1 に記載の方法。

## 【請求項 8】

前記 c ) のステップは、ウィンドウ内の各部分領域における重みと統計量とを決定し、  
前記総和は、前記重みと前記各部分領域における統計量との積の総和で与えられ、  
前記ウィンドウ内の全ての部分領域に渡って前記総和を計算するステップを含むことを  
特徴とする請求項 1 に記載の方法。

## 【請求項 9】

前記ウィンドウ内の各部分領域における重みを決定するステップは、運動方向を決定し、前記ウィンドウ内の対応する部分領域が前記運動方向に沿って位置するか否かにしたがって各重みを割り当てるステップとを含むことを特徴とする請求項 8 に記載の方法。

## 【請求項 10】

前記 e ) のステップは、前記総和と前記統計量との差分を閾値と比較するステップを含むことを特徴とする請求項 1 に記載の方法。

## 【請求項 11】

前記 g ) のステップは、前記シーン切替領域の数を閾値と比較するステップを含むことを特徴とする請求項 1 に記載の方法。

## 【請求項 12】

前記 g ) のステップは、前記シーン切替領域の数が前記現在のピクチャ内の部分領域の過半数か否かを決定するステップを含むことを特徴とする請求項 1 に記載の方法。

## 【請求項 13】

プロセッサと、

前記プロセッサと接続するメモリと、

前記メモリ内に具現化され前記プロセッサが実行可能な命令であって、1 以上のデジタルピクチャを符号化する際のシーン切替を検出する方法を実装するように構成された命令を含み、前記方法は、

a ) 現在のピクチャ内の所与の部分領域における統計量を計算するステップと、

b ) 先行するピクチャ内の対応する位置にある部分領域周辺の 2 以上の部分領域のウィンドウを規定するステップと、

c ) 前記ウィンドウ内の部分領域全面における統計量の総和を計算するステップと、

d ) 前記総和と前記所与の部分領域における統計量との差分を計算するステップと、

e ) 前記総和と前記所与の部分領域における前記統計量との差分を用いて前記所与の部分領域がシーン切替領域か否かを決定するステップと、

f ) 前記現在のピクチャについてシーン切替領域の数を決定するステップと、

g ) 前記シーン切替領域の数から前記現在のピクチャがシーン切替ピクチャか否かを決

10

20

30

40

50

定するステップと、

h) 前記現在のピクチャがシーン切替ピクチャか否かを示す情報を格納または転送するステップと、

i) 前記現在のピクチャがシーン切替ピクチャか否かを示す情報にしたがって、エンコーダを用いて前記現在のピクチャを符号化するステップとを含むことを特徴とする、1以上のデジタルピクチャを符号化する際のシーン切替を検出するシステム。

【請求項14】

前記現在のピクチャがシーン切替ピクチャか否かを示す情報にしたがって、前記現在のピクチャを符号化するように構成されたエンコーダをさらに含むことを特徴とする請求項13に記載のシステム。

10

【請求項15】

前記命令はさらに、前記現在のピクチャ内の複数の部分領域について前記a)のステップからe)のステップを繰り返すように構成されていることを特徴とする請求項13に記載のシステム。

【請求項16】

前記c)のステップは、ウィンドウ内の各部分領域における重みと統計値とを決定し、前記総和は、前記重みと前記各部分領域における統計量との積の総和で与えられ、前記ウィンドウ内の全ての部分領域に渡って前記総和を計算するステップを含むことを特徴とする請求項13に記載のシステム。

【請求項17】

前記ウィンドウ内の各部分領域における重みを決定するステップは、運動方向を決定し、前記ウィンドウ内の対応する部分領域が前記運動方向に沿って位置するか否かにしたがって各重みを割り当てるステップを含むことを特徴とする請求項16に記載のシステム。

20

【請求項18】

前記エンコーダは、前記現在のピクチャがシーン切替ピクチャであることを前記情報が示す場合、前記現在のピクチャをIフレームとして符号化し、現在のピクチャがシーン切替ピクチャではないことを前記情報が示す場合、前記現在のピクチャをBフレームまたはPフレームとして符号化するように構成されていることを特徴とする請求項14に記載のシステム。

【請求項19】

コンピュータプログラム命令が格納されているコンピュータ読み取り可能な記録媒体であって、前記コンピュータプログラム命令は実行されたときに、1以上のデジタルピクチャを符号化する際のシーン切替を検出する方法を実装するように構成されており、前記方法は、

30

a) 現在のピクチャ内の所与の部分領域における統計量を計算するステップと、

b) 先行するピクチャ内の対応する位置にある部分領域周辺の2以上の部分領域のウィンドウを規定するステップと、

c) 前記ウィンドウ内の部分領域全面における統計量の総和を計算するステップと、

d) 前記総和と前記所与の部分領域における統計量との差分を計算するステップと、

e) 前記総和と前記所与の部分領域における前記統計量との差分を用いて前記所与の部分領域がシーン切替領域か否かを決定するステップと、

40

f) 前記現在のピクチャについてシーン切替領域の数を決定するステップと、

g) 前記シーン切替領域の数から前記現在のピクチャがシーン切替ピクチャか否かを決定するステップと、

h) 前記現在のピクチャがシーン切替ピクチャか否かを示す情報を格納または転送するステップと、

i) 前記現在のピクチャがシーン切替ピクチャか否かを示す情報にしたがって、エンコーダを用いて前記現在のピクチャを符号化するステップとを含むことを特徴とするコンピュータ読み取り可能な記録媒体。

【請求項20】

50

前記コンピュータプログラム命令はさらに、前記現在のピクチャがシーン切替ピクチャか否かを示す情報にしたがって、前記現在のピクチャを符号化するように構成されていることを特徴とする請求項19に記載の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】

本発明の実施の形態はビデオ画像のデジタル符号化に関し、特に、ビデオ画像の符号化の際のシーン切替検出に関する。

【0002】

本出願は、2009年9月2日に出願された「シーン切替検出」と題する米国出願第12/553,069号の優先権を主張し、開示内容全体を参照によりここに完全に組み入れる。

10

【0003】

本出願は、同一出願人による同時係属中の米国特許出願12/553,070(「ビデオ符号化のためのピクチャレベル比率制御」、2009年9月2日出願、代理人ドケット番号:SC EA 08073US00)に関連し、開示内容全体を参照によりここに完全に組み入れる。

【0004】

本出願は、同一出願人による同時係属中の米国特許出願12/553,073(「並列デジタルピクチャ符号化」、2009年9月2日出願、代理人ドケット番号:SC EA 08077US00)に関連し、開示内容全体を参照によりここに完全に組み入れる。

20

【0005】

本出願は、同一出願人による同時係属中の米国特許出願12/553,075(「ビデオ符号化において高速な動き予測を達成するための閾値の利用および早期中止」、2009年9月2日出願、代理人ドケット番号:SC EA 08078US00)に関連し、開示内容全体を参照によりここに完全に組み入れる。

【背景技術】

【0006】

デジタル信号圧縮は、多くのマルチメディアアプリケーションや装置で広く用いられている。符号化器/復号器(コーデック)を用いるデジタル信号圧縮によって、音声やビデオ信号のようなストリーミングメディアをインターネットを介して送信したり、コンパクトディスクに格納したりすることができるようになる。H.261、H.263、DV、MPEG-1、MPEG-2、MPEG-4、VC1、およびAVC(H.264)を含むデジタルビデオ圧縮における多くの異なる標準が出現している。これらの標準は、他のビデオ圧縮技術と同様に、あるビデオフレームピクチャと連続するピクチャとの間における空間的および時間的な冗長性を取り除くことによって、そのビデオフレームピクチャを効率良く表現しようとする。このような圧縮標準を用いることを通して、ビデオコンテンツは高度に圧縮されたビデオビットで伝えることが可能となり、それゆえ効率よくディスクに格納したりネットワークを介して送信したりすることができる。

30

【0007】

MPEG-4 AVC(Advanced Video Coding; H.264としても知られる)はビデオ圧縮標準であり、先行する標準よりも明らかに高い圧縮を提供する。H.264標準は、従前のMPEG-2標準の最大2倍の圧縮を提供することが期待される。H.264標準は、視覚的な質の向上を提供することも期待されている。その結果、ますます多くのビデオコンテンツがAVC(H.264)で符号ストリームの形態で配信されている。ふたつの競合するDVDフォーマット、すなわちHD-DVDフォーマットとBlu-Ray(登録商標)ディスクフォーマットは、H.264/AVCのハイプロファイルデコードを任意のプレイヤーの機能としてサポートしている。AVC(H.264)符号化は、"Draft of Version 4 of H.264/AVC (ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG-4 part 10) Advanced Video Coding)" by Gary Sullivan, Thomas Wiegand and Ajay

40

50

Luthra, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6), 14th Meeting: Hong Kong, CH 18-21 January, 2005, に詳細に記載されており、あらゆる目的のために開示内容全体を参照によりここに完全に組み入れる。

【 0 0 0 8 】

MPEG - 2、MPEG - 4、およびH. 264のような最新のビデオ符号化/復号(コーデック)は、一般にビデオフレームを、イントラフレーム、プレディクティブフレーム、およびバイプレディクティブフレームとして知られる3つの基本的な種類に分割する。これらの3つのフレームはそれぞれ、通常はIフレーム、Pフレーム、およびBフレームと呼ばれる。

10

【 0 0 0 9 】

Iフレームは、Iフレーム自身を除くと、いかなるピクチャも参照することなく符号化される。Iフレームは、ランダムアクセスおよび他のPフレームまたはBフレームを復号するために用いられる。Iフレームは、ランダムアクセスのポイントを生成するため(デコーダに所与のピクチャ位置において0から適切に復号を開始できるようにするため)にエンコーダによって生成されてもよい。Iフレームは、画像の細部の違いが、効果的なPまたはBフレームを生成するのに妨げとなるときに生成されてもよい。Iフレームは完全なピクチャを含むので、Iフレームは通常、PフレームやBフレームよりも符号化するために多くのビットが必要となる。

【 0 0 1 0 】

Pフレームを復号するためには、前もって他の1またはいくつかのピクチャの復号が必要となる。Pフレームは通常、Iフレームの符号化に要するビットよりも少ないビットを要する。Pフレームは、復号の順序において先にあるIフレームに対する相違点に関する符号化された情報を含む。Pフレームは通常、GoP(Group of Pictures)内の先行するIフレームを参照する。Pフレームは、画像データと動きベクトルの変化との両方、およびそれらふたつの組み合わせを含む。MPEG - 2のようないくつかの標準コーデックにおいてPフレームは、復号中に、既に復号されているピクチャをただひとつだけ参照として用いる。ここでPフレームは、符号時に参照するピクチャは、表示の順序においてもPフレームよりも前にあることを要求する。H. 264においてPフレームは、復号中に、複数の既に復号されているピクチャを参照として用いることができ、また予測に用いられるピクチャに対して任意の表示順序となることができる。

20

30

【 0 0 1 1 】

Bフレームは、復号するために、前もってIフレームかPフレームのいずれか一方の復号が必要となる。Pフレームと同様に、Bフレームは、画像データと動きベクトルの変化との両方、および/またはそれらふたつの組み合わせを含んでもよい。Bフレームは、動き領域(例えばマクロブロックやそれより小さな領域のようなフレームのセグメント)の予測を、前もって復号されているふたつの異なる参照領域を用いて得られる予測を平均することで形成するいくつかの予測モードを含んでもよい。MPEG - 2のようないくつかのコーデックでは、Bフレームは他のピクチャを予測するための参照としては使われることはない。その結果、これらのBフレームにはより低品質の符号化(そうでない場合よりもより少ないビットの使用)を用いることができる。なぜなら、細部の喪失が後続のピクチャの予測品質を害さないからである。H. 264のような他のコーデックでは、Bフレームは、Bフレームは他のピクチャの復号のための参照に使われてもよいし使われなくてもよい(エンコーダの裁量による)。MPEG - 2のようないくつかのコーデックは、復号中に、先に復号されたピクチャを参照として厳密にふたつ使い、このうち一方のピクチャは表示順序でBフレームに先行し、もう一方は追従する。H. 264のような他のコーデックでは、Bフレームは、復号中に、1、2、または2よりも多くの先に復号されたピクチャを参照として用いることができ、また予測に用いられるピクチャに対して任意の表示順序となることができる。Bフレームは通常、IフレームまたはPフレームのいずれのフレームよりも、符号化に要求されるビットは少ない。

40

50

## 【 0 0 1 2 】

本明細書において、Iフレーム、Bフレーム、およびPフレームという用語は、例えば上述したストリーミングビデオの文脈におけるIフレーム、Bフレーム、およびPフレームと同様の性質を持つ任意のストリーミングデータの単位に適用される。

## 【 発明の概要 】

## 【 発明が解決しようとする課題 】

## 【 0 0 1 3 】

ビデオの符号化は、多くの場合、所与の映像シーン内においてそのシーンのビジュアルコンテンツのいくつかの要素が比較的静止したままでいる傾向があるという事実を利用する。それ故、所与のピクチャを、参照として用いられる先行ピクチャとそのピクチャとの差分の観点から符号化することによって、ビデオ信号の符号化に必要なデータを抑制することが可能となる。しかしながら、もしビデオシーケンスがシーン切替を含んでいると、参照として利用可能な先行ピクチャが存在しなくなりうる。したがって、シーン切替を検出することができる、符号化プログラムにとって利用価値がある。シーン切替は符号化プロセスに影響を及ぼすことができるからである。従来のシーン切替検出アルゴリズムは、ビデオフレームのコンテンツの解析に基づくものである。

## 【 課題を解決するための手段 】

## 【 0 0 1 4 】

本発明の実施の形態は、この文脈の中で生まれたものである。

## 【 0 0 1 5 】

本発明の教示は、添付の図面とともに下記の詳細な説明を考慮することにより容易に理解されよう。

## 【 図面の簡単な説明 】

## 【 0 0 1 6 】

【 図 1 A 】 二つのビデオフレーム間のシーン切替を説明する図である。

【 図 1 B 】 符号化のためのビデオピクチャの分割のあり得るひとつの候補を説明する図である。

【 図 2 A 】 本発明の実施の形態に係るシーン切替検出アルゴリズムの一例を説明するフロー図である。

【 図 2 B 】 本発明の実施の形態に係るマクロブロックに基づくシーン切替アルゴリズムを説明するブロック図である。

【 図 2 C 】 本発明の実施の形態に係るマクロブロックに基づくシーン切替アルゴリズムを説明するブロック図である。

【 図 3 A 】 重み付きフィルタウィンドウのためのウィンドウの形を示す図である。

【 図 3 B 】 重み付きフィルタウィンドウのためのウィンドウの別の形を示す図である。

【 図 3 C 】 重み付きフィルタウィンドウのためのウィンドウのさらに別の形を示す図である。

【 図 4 A 】 4つの8×8ピクセルから構成されるマクロブロックを示す図である。

【 図 4 B 】 4つの8×8ピクセルから構成されるマクロブロックを示す図である。

【 図 5 A 】 移動方向とともにマクロブロックを示す図である。

【 図 5 B 】 移動方向とともにマクロブロックを示す図である。

【 図 6 】 本発明の実施の形態に係るビデオピクチャ中のシーン切替検出を実装するための装置を説明するブロック図である。

【 図 7 】 本発明の実施の形態に係るビデオピクチャ中のシーン切替検出を実装する装置のセルプロセッサの実装の一例を説明するブロック図である。

【 図 8 】 本発明の実施の形態に係るピクチャレベル比率制御の実装のためのコンピュータ読み取り可能な命令を含むコンピュータ読み取り可能な記録媒体を示す図である。

## 【 発明を実施するための形態 】

## 【 0 0 1 7 】

下記の詳細な説明は、例示を目的として多くの特定の詳細を含むが、下記の詳細に対す

10

20

30

40

50

る多くの変形及び置換が本発明の範囲に入ることは、当業者に認識されることである。したがって、下記に説明する本発明の実施の形態は、請求された発明の一般性を失わせることなく、また限定を付加することなく説明される。

**【0018】**

## イントロダクション

シーン切替検出の問題の特性は図1Aを参照することで理解されよう。図1Aは、第1ビデオフレーム101と第2ビデオフレーム102とをそれぞれ示している。フレーム101と102とは、ビデオシーケンスの中で連続するフレームである。第1フレーム101は、特定の背景105に対して特定のアングルから撮像したふたりのキャラクタ103および104の写真を示す。第2フレーム102は同じふたりのキャラクタ103および104を示すが、第1フレーム101と第2フレーム102との間でキャラクタ103と104との相対位置が変わっている。その上、第1フレーム101においてはキャラクタ103は笑っているが、第2フレーム102においてはキャラクタ103は顔をしかめている。

10

**【0019】**

コンテンツの観点（従来技術のアプローチ）から見ると、第1フレーム101と第2フレーム102との間の相違はシーン切替と解釈されるであろう。しかしながら、符号化の観点から見ると、シーンの要素（例えば図1Aの例で言えば、キャラクタ103、104、および背景105）は実質的に同一であるため、これはシーン切替ではない。それ故、第1フレーム101と第2フレーム102が根底となる動き探索の範囲内にある限り、第1フレーム101は第2フレーム102の参照として用いることができる。このことは、残念なことに、コンテンツベースのシーン切替検出を用いる従来技術に係るビデオ符号化スキームでは認識することができず、第2フレーム102を独立フレーム（Iフレーム）として符号化するであろう。これはフレーム101、102の符号化における非効率の原因となりうる。

20

**【0020】**

この問題を克服するために、本発明の実施の形態は、ビデオシーン切替検出アルゴリズムを実装する。本明細書で提案するアルゴリズムは、ビデオシーケンス内に効果的なシーンカットを設置するため、続くビデオ処理モジュール（例えば、ビデオ圧縮/符号化）はその有益な情報を利用することができる、ビデオ符号化アプリケーションの領域における典型的なアプリケーションは、ピクチャフレームのタイプを決定する。例えば、シーンカットのあるピクチャフレームはIピクチャフレームと分類され、MPEGビデオエンコーダはこのフレームを過去のピクチャフレームを参照することなく独立に符号化することができる。

30

**【0021】**

本明細書に記載されるシーン切替検出アルゴリズムは、多くの観点から他のシーン検出アルゴリズムと相違する。特に、提案するアルゴリズムは概念的にはビデオ符号化と結びつけられるが、必ずしも実装においてビデオ符号化と結びつける必要はない。すなわち、現存する多くのシーン切替アルゴリズムは動き推定モジュールの中に埋め込まれているが、本アルゴリズムはビデオ符号化プログラムの一部である必要はない。本アルゴリズムは、ビデオ符号化エンジンから分離してもよく、符号化のためのビデオシーケンスの前処理の形とみなすこともできる。

40

**【0022】**

提案するアルゴリズムでは、あるピクチャフレームがシーン切替であるか否かの決定は、ピクチャのコンテンツに加えて符号化の効率に基づく。例えば、あるピクチャフレームがその隣接するフレームに対して非常に類似した背景の上に異なる小さな前景のオブジェクトがあることを示しているとする。コンテンツの観点から見れば、このピクチャはシーン切替フレームであると分類されるであろう。しかしながら、符号化の効率の観点から見れば、ピクチャフレームの大きな部分（例えば、背景オブジェクト）は冗長であり、先行するフレームの背景と類似する。この冗長な情報は、内在する動き推定/圧縮モジュール

50

によって抽出され除去される。

【 0 0 2 3 】

提案するアルゴリズムはマクロブロックベースのアプローチであり、典型的な M P E G ベースのビデオ符号化に適合する。提案するアルゴリズムは、重み係数付きウィンドウフィルタリングも利用して、現在のピクチャ内のマクロブロックと先行するピクチャにおいて対応する位置にあるマクロブロックおよびそれと隣接するマクロブロックとの差異を決定する。ウィンドウベースのフィルタリングを用いて、統計的差異を計算して、最終的な決定のために収集する。

【 0 0 2 4 】

ビデオピクチャは符号化および復号のために適切な大きさの処理単位に分割される。例えば図 1 B に示すように、あるひとつのピクチャ 1 1 0 は 1 以上の部分領域に分割される。上述のように、本明細書において「部分領域 (section)」とは、ひとつのピクチャ内の 1 以上のピクセルの集まりを意味する。部分領域は、ひとつのピクチャ内の 1 ピクセルからピクチャ全体までの範囲を取り得る。部分領域についての限定しない例としては、スライス 1 1 2、マクロブロック 1 1 4、サブマクロブロック 1 1 6、ブロック 1 1 8、および個々のピクセル 1 2 0 を含む。図 1 B に示すように、各スライス 1 1 2 は、1 以上のマクロブロック 1 1 4 の列を含む。ひとつの列内のマクロブロックの数は、マクロブロックの大きさとピクチャ 1 1 0 の大きさおよび解像度に依存する。例えば、各マクロブロックが  $16 \times 16$  ピクセルである場合、それぞれの列におけるマクロブロックの数は、ピクチャ 1 1 0 の幅 (ピクセル数) を 16 で除算することで決定される。各マクロブロック 1 1 4 は、いくつかのサブブロック 1 1 6 に分割される。各サブマクロブロック 1 1 6 は、いくつかのブロック 1 1 8 に分割され、各ブロックはいくつかのピクセル 1 2 0 を含む。本発明を限定しない例として、通常のビデオ符号化スキームにおいて、ピクセルのマクロブロック 1 1 4 それぞれは、 $8 \times 8$  ピクセルのサブマクロブロック 1 1 6 に分割される。各サブマクロブロックは、4 つのブロック 1 1 8 に分割され、各ブロックは 16 個のピクセル 1 2 0 を含む 4 つの  $4 \times 4$  ピクセルの配置を含む。

【 0 0 2 5 】

アルゴリズムの詳細

図 2 A は、シーン切替検出アルゴリズム 2 0 0 の一例を示す。上述したように、ビデオストリーム内の各ピクチャは複数の部分領域に分割される。本発明の実施の形態においてシーン切替の決定は、部分領域毎における前のピクチャに対する現在のピクチャの分析に基づく。限定ではなく例示の目的として、この分析は上述した各マクロブロックベースで説明される。このアルゴリズムは部分領域の大きさを他の大きさに一般化できることは当業者には理解されよう。

【 0 0 2 6 】

符号 2 0 2 で示すように、現在のピクチャ内の各マクロブロックを特徴付けるために、各マクロブロックの統計量  $\mu_m$  が計算される。有用な統計量の例は、各マクロブロックのピクセルの値 (例えばクロミナンスや輝度) の平均値やそれらの分散である。各マクロブロックの統計量  $\mu_m$  の値はマクロブロック基準でメモリに格納される。

【 0 0 2 7 】

図 2 B - 2 C に描写するように、現在のビデオピクチャ N は、先行するピクチャ N - 1 と同様にマクロブロックに分割される。アルゴリズム 2 0 0 は、先行するピクチャ N - 1 および現在のピクチャ N 内のマクロブロックの解析に基づいて現在のフレームがシーン切替フレームか否かを決定するために先行するフレームを用いる。特に、先行するフレーム N - 1 内の各マクロブロックの統計量  $\mu_m$  の値が計算され、以下に説明するシーン切替決定のためのコンピュータメモリに格納される。

【 0 0 2 8 】

符号 2 0 4 が示すように、先行するピクチャ内の対応する位置にあるマクロブロックの周囲にウィンドウが規定される。限定ではない例として、現在のピクチャ N 内のマクロブロック A について対応する位置にあるマクロブロック a は、先行するフレーム N - 1 内に

10

20

30

40

50

見出せる。また、マクロブロック a に隣接する 8 つの全てのマクロブロック ( b から i ) も、重み付きフィルタリングウィンドウを形成するために設置される。ここでマクロブロック A および a がそれぞれ異なるピクチャに存在するもののそれぞれのピクチャ内で対応する位置にあるとき、「対応する位置に存在する ( co-located ) 」という。例として、マクロブロック A が、現在のピクチャ内の k 番目のマクロブロック列中の j 番目のマクロブロックであり、対応する位置に存在するマクロブロック a は、先行するピクチャ内の k 番目のマクロブロック列中の j 番目のマクロブロックである。ここで、シーン切替検出のために技術的意義がある限り、異なる形のウィンドウを選ぶこともできることは留意すべきである。図 3 A - 3 C は、本発明の実施の形態に係るフィルタリングウィンドウとして用いることができる代替可能なウィンドウ形状を示す。

10

## 【 0 0 2 9 】

ひとたび先行するピクチャ内の対応する位置に存在するマクロブロックの周囲のフィルタリングウィンドウが選択されると、符号 2 0 6 に示すようにウィンドウ内の各マクロブロックそれぞれの統計量の値に基づいて統計的総和 E が計算される。限定されない例として、統計的総和 E はウィンドウ全体にわたる統計量  $\mu_m$  の重み付き総和であってもよい。ある実施の形態では、ウィンドウ内の各マクロブロックに重み係数が割り当てられ、中心のマクロブロック ( 例えば A および a ) に対する各マクロブロックの関連度が表現される。例えば、マクロブロック A および a が統計的に十分に類似する場合、ウィンドウ内において隣接する他の全てのマクロブロックに等しい重み係数が割り当てられる。A および a が統計的に十分に類似しない場合、関連する移動方向 M が導出される。検出された移動方向 M に基づいて、ウィンドウ内のマクロブロックのうち移動方向 M に沿わないマクロブロックと比較して、移動方向に沿うマクロブロックの方により大きな重み係数が割り当てられる。移動方向 M の導出の仕方はいくつか存在する。限定されない例としては、図 4 A - 4 B に示すように、簡易な  $8 \times 8$  ピクセルのブロック ( 典型的にはひとつのマクロブロックは 4 つの  $8 \times 8$  ブロックを含む ) を移動方向の推定に用いることができる。図 5 A - 5 B に示すように、ブロックの最も大きな差異から移動方向を導くことができる。

20

## 【 0 0 3 0 】

移動方向はあらゆる適切なやり方で導出されてもよい。限定されない例として、移動方向は先行するピクチャ N - 1 内のマクロブロック a と現在のピクチャ内のマクロブロック A との間で最も統計的に異なるブロック間を結ぶラインと平行であってもよい。

30

## 【 0 0 3 1 】

図 5 A - 5 B に示す例では、移動方向 M に沿わないマクロブロック g、b、i、e、および h よりも、移動方向 M に沿うマクロブロック f、c、および d の方に、より重みが割り当てられる。正規化のため、重み係数の総和は定数 1 に設定されてもよい。すなわち、

## 【 数 1 】

$$\sum_{m=a}^i W_m = 1.0,$$

ここで  $W_m$  は所与のマクロブロック m の重み係数であり、 $m = \{ a, b, c, d, e, f, g, h, i \}$ 、 $0.0 \leq W_m \leq 1.0$  である。

40

## 【 0 0 3 2 】

重み付き係数が割り当てられると、重み付き総和が計算される。例えば、マクロブロックのピクセル値の平均値 ( 他の統計データも同様に用いることができる ) を  $\mu_m$  と表記して用いると仮定すると、重み付き総和 E は以下の式で与えられる。

## 【 数 2 】

$$E = \sum_{m=a}^i (W_m \times \mu_m).$$

## 【 0 0 3 3 】

ひとたび重み付き総和 E が計算されると、符号 2 0 8 に示すように、マクロブロック A

50

の平均値  $\mu_A$  と重み付き総和  $E$  との差分値が計算される。差分値が所定のしきい値 ( $TH$ ) を超える場合、マクロブロック  $A$  はシーン切替マクロブロックとしてカウントされ、双でなければそのマクロブロックはシーン切替ではないマクロブロックである。閾値  $TH$  はピクチャのコンテンツに基づいてピクチャ毎に更新してもよいことは留意すべきである。

【 0 0 3 4 】

現在のピクチャ  $N$  内の各マクロブロックについてシーン切替がシーン切替でないかの決定がなされた後、現在のピクチャ  $N$  内のシーン切替マクロブロックの数がカウントされる。シーン切替マクロブロックの数に基づいてピクチャフレームがシーン切替であるか否かを決定することが可能となる。限定されない例として、本アルゴリズムはこの決定に際し、シーン切替マクロブロックの数が現在のピクチャ  $N$  内のマクロブロックの中で過半数 (majority) を占めるか否か、または、シーン切替マクロブロックの数が本ピクチャ用の所定のしきい値よりも大きいか否かに基づいてもよい。図 2 A に図示する例では、シーン切替マクロブロックの数を増加させることで、シーン切替マクロブロックの集計が実行される。例えば、符号 2 1 0 においてあるマクロブロックがシーン切替マクロブロックであると決定されると、符号 2 1 2 において数が増加される。符号 2 1 4 において集計が閾値を超えると、符号 2 1 6 においてそのピクチャはシーン切替ピクチャとみなされ、それ以上のシーン切替マクロブロックの計算は実行されない。符号 2 1 0 においてマクロブロックがシーン切替マクロブロックでないことが分かり、符号 2 1 8 においてマクロブロックが最後のマクロブロックであり、かつ、符号 2 1 4 において集計が閾値以下を維持する場合、符号 2 2 0 において現在のピクチャ  $N$  はシーン切替ピクチャでないと判断される。そうでなければ、現在のピクチャ内の次のマクロブロックについて処理を繰り返す。ひとたび現在のピクチャ内の各マクロブロックについて処理が繰り返されると、後続のピクチャ  $N + 1$  について処理が繰り返す。現在のピクチャ  $N$  内の各マクロブロックについて計算される  $\mu_M$  の値は、ピクチャ  $N + 1$  の先行ピクチャであるピクチャ  $N$  の統計的総和  $E$  の計算のために格納されてもよい。

【 0 0 3 5 】

ひとたび現在のピクチャがシーン切替ピクチャであるか否かの決定がなされると、現在のピクチャが、例えばソフトウェアエンコーダを用いて符号化される。アルゴリズム 2 0 0 は、現在のピクチャがシーン切替を含んでいるか否かを示す情報を格納するか、転送するか、エンコーダは現在のピクチャを符号化するときにこの情報を用いる。例えば、現在のピクチャがシーン切替ピクチャであることをその情報が示している場合、エンコーダは現在のピクチャを I フレームとして符号化する。そうでない場合、エンコーダはそのピクチャを P フレームまたは B フレームとして符号化する。

【 0 0 3 6 】

上述したアルゴリズムは適切なコンピュータ装置上で実装されてもよい。図 6 は、上述したピクチャ符号化のためのシーン切替検出の実装に用いられるコンピュータ装置 6 0 0 を例示するブロック図である。装置 6 0 0 は概して、プロセッサモジュール 6 0 1 およびメモリ 6 0 2 を含む。プロセッサモジュール 6 0 1 は、1 以上のプロセッサコアを含んでもよい。複数のプロセッサモジュールを用いるプロセッサシステムの一例はセルプロセッサである。セルプロセッサの例は、例えば、[http://www-306.ibm.com/chips/techlib/techlib.nsf/techdocs/1AEEE1270EA2776387257060006E61BA/\\$file/CBEA\\_01\\_pub.pdf](http://www-306.ibm.com/chips/techlib/techlib.nsf/techdocs/1AEEE1270EA2776387257060006E61BA/$file/CBEA_01_pub.pdf) で参照可能な「Cell Broadband Engine Architecture」に詳細が説明されており、参照によりここに組み入れる。

【 0 0 3 7 】

メモリ 6 0 2 は、例えば RAM、DRAM、ROM またはそれと類似する集積回路の形式であってもよい。メモリは、すべてのプロセッサモジュール 6 0 1 からアクセス可能なメインメモリであってもよい。ある実施の形態では、プロセッサモジュール 6 0 1 は、複数のプロセッサコアと、各コアに関連づけられたローカルメモリを含んでもよい。符号化プログラム 6 0 3 は、プロセッサモジュール 6 0 1 上で実行可能なプロセッサ読み取り可能な命令の形式でメインメモリ 6 0 2 内に格納されてもよい。符号化プログラム 6 0 3 は

、ピクチャを圧縮信号データに符号化すること、および/または圧縮信号データを復号するように構成されてもよい。限定されない例として、符号化プログラム603は、同一出願人による同時係属中の米国特許出願公開第20090010338号(参照により開示内容全体をここに組み入れる)に説明されているように構成されてもよい。符号化プログラム603所与のピクチャがシーン切替ピクチャであるか否かの情報を受け取り、上述のように現在のピクチャがシーン切替ピクチャであるか否かに部分的に基づいて、そのピクチャをIフレーム、Pフレーム、またはBフレームとして符号化するように構成されてもよい。

#### 【0038】

シーン切替検出プログラム604もまた、メモリ602に格納されてもよい。シーン切替検出プログラムは、プロセッサモジュール601で実行されたときに、例えば図2Aにしたがって上述したアルゴリズム200としてシーン切替検出を実装する命令を含んでもよい。符号化プログラム603およびシーン切替検出プログラム604は、例えばC、C++、JAV A(登録商標)、アセンブリ言語、MATLAB、FORTRAN、および他のいくつかの言語等の、プロセッサ読み取り可能な適切な言語で記述されてもよい。

#### 【0039】

入力データ607はメモリ602に格納される。そのような入力データは、例えば符号化されたビデオピクチャやそれらの一部などの、バッファリングされたストリーミングデータの一部を含んでもよい。符号化プログラム603および/またはシーン切替検出プログラム604の実行の間、プログラムコードおよび/またはデータ607の一部は、メモリ602中、またはマルチプロセッサコアによって並列処理されるためにプロセッサコアのローカルストアにロードされる。限定されない例として、符号化の前や符号化の中間段階において、入力データ607はビデオピクチャやそれらの部分領域を含んでもよい。これらの様々な部分領域は1以上のバッファに格納されてもよい。具体的には、部分領域はメモリ602内に実装された出力ピクチャバッファ内に格納されてもよい。

#### 【0040】

装置600はさらに、入出力(I/O)装置611、電源(P/S)612、クロック(CLK)613およびキャッシュ614などの公知のサポート機能610を備えてもよい。装置600はオプションとして、プログラムおよび/またはデータを格納するためのディスクドライブ、CD-ROMドライブ、テープドライブなどの大容量記憶装置615を備えてもよい。装置600はまた、オプションとして、装置600とユーザの相互作用を容易にするために、ディスプレイユニット616とユーザインタフェースユニット618を備えてもよい。ディスプレイユニット616は、テキスト、数値、グラフィカルシンボルや画像を表示する陰極線管(CRT)、またはフラットパネルスクリーンの形態であってもよい。ユーザインタフェース618は、キーボード、マウス、ジョイスティック、ライトペンや他の装置を備えてもよく、これらは、グラフィカルユーザインタフェース(GUI)と併せて使われてもよい。装置600はまた、ネットワークインタフェース620を含み、これにより、当該装置がインターネットのようなネットワーク上で他の装置と通信することが可能になる。これらの構成要素はハードウェア、ソフトウェア、ファームウェアまたはこれらの2以上の組み合わせによって実装される。

#### 【0041】

装置600の複数のプロセッサを用いて並列処理を効率化する付加的な方法が多数ある。たとえば、2以上のプロセッサ上でコードを複製することによって、処理ループを「アンロール(unroll)」し、各プロセッサに異なるデータ部分を処理するためにコードを実装させることができる。そのような実装によって、ループ設定に関連するレイテンシを回避することができる。

#### 【0042】

並列処理を実装することができるマルチプロセッサシステムの一つの例は、セルプロセッサとして知られる。セルプロセッサとして分類できる異なるプロセッサアーキテクチャが多数ある。例示であり、限定するものではないが、図7は、セルプロセッサ700の一

10

20

30

40

50

つのありうる構成を示す。セルプロセッサ700は、メインメモリ702、一つのPPE (power processor element) 704、および8個のSPE (synergistic processor element) 706を備える。あるいは、セルプロセッサ700は、任意の数のSPEで構成されてもよい。図7に関し、メモリ702、PPE704、および複数のSPE706は、リング型のエレメント・インターコネクト・バス710上でお互いに通信可能であり、また、I/Oデバイス708とも通信可能である。メモリ702は、上述した入力データ607と同じような特性を持つ入力データ703、上述した符号化プログラム603と同じような特性を持つ符号化プログラム709、および上述したシーン切替検出プログラム604と同じような特性を持つシーン切替検出プログラム711を含む。少なくともSPE706のひとつは、ローカルストア(LS)内にシーン切替検出命令705、および/または、例えば上述したように並列で処理されるためのバッファリングされた入力データの一部を含む。PPEは、L1キャッシュ内に上述した符号化プログラム603と同じ特性を持つコード命令707を含む。命令705およびデータ707は、SPEおよびPPEが必要なときにアクセスするためにメモリ702にも格納されてもよい。

#### 【0043】

例示として、PPE704は、関連するキャッシュを持った64ビットのパワーPCプロセッサユニット(PPU)であってもよい。PPE704はオプションとしてベクタマルチメディア拡張ユニットを含んでもよい。各SPE706は、SPU (synergistic processor unit) とローカルストア(LS)を含む。ある実装では、ローカルストアは、たとえば、コードとデータ用のおよそ256キロバイトのメモリ容量をもってもよい。SPUは、典型的にはシステム管理機能を実行しないという点で、PPUよりも単純な計算ユニットである。SPUは、SIMD (single instruction, multiple data) 機能を有し、典型的にはデータ処理を行い、割り当てられたタスクを行うために(PPEにより設定されたアクセス特性にしたがって) 要求されたデータ転送を開始する。SPUにより、システム700は、より高い計算ユニット密度を要求するアプリケーションを実装し、提供された命令セットを効率良く利用することができるようになる。PPE704によって管理されるシステム700の相当数のSPEによって、広範囲のアプリケーションにわたって費用対効果の高い処理が可能になる。例示として、セルプロセッサ700は、セル・ブロードバンド・エンジン・アーキテクチャ(CBEA)準拠のプロセッサとして知られるアーキテクチャで特徴づけられる。CBEA準拠のアーキテクチャにおいて、複数のPPEはPPEグループに結合され、複数のSPEはSPEグループに結合されてもよい。例示のために、セルプロセッサ700は、一つのSPEと一つのPPEをもつ一つのSPEグループと一つのPPEグループだけをもつとして図示している。別の構成として、セルプロセッサは、PPEの複数のグループとSPEの複数のグループを含むことができる。CBEA準拠のプロセッサは、たとえば、「セル・ブロードバンド・エンジン・アーキテクチャ」に詳細に記述されており、これは、[http://www-306.ibm.com/chips/techlib/techlib.nsf/techdocs/1AEEE1270EA2776387257060006E61BA/\\$file/CBEA\\_01\\_pub.pdf](http://www-306.ibm.com/chips/techlib/techlib.nsf/techdocs/1AEEE1270EA2776387257060006E61BA/$file/CBEA_01_pub.pdf)においてオンラインで利用可能であり、その内容を参照によりここに取り込む。

#### 【0044】

本発明の実施の形態は、シーン切替検出およびストリーミングビデオのようなストリーミングデータの符号化の並列処理のためのシステムおよび方法を提供する。このような実施の形態は、ほとんどのビデオエンコーダ、具体的にはH.264/AVCエンコーダおよび特にモジュールとしてビデオエンコーダを持つ製品に適用される。限定はしないが、そのような製品の例はゲームコンソール、DVDプレイヤー、ソフトウェア(PC)ビデオエンコーダ/プレイヤー、携帯電話上のビデオ等を含む。別の実施の形態として、このようなシステムおよび方法を、ビデオ以外の他のストリーミングデータの符号化に適用してもよい。

#### 【0045】

そのような実施の形態に例は、ストリーミング音声データ、グラフィックレンダリングストリーム、静止画およびXML文書の符号化のためのシステムおよび方法を含む。本発

10

20

30

40

50

明の実施の形態は、ビデオゲームを記録したり、携帯ゲーム装置にゲームの画像をストリーミングしたり、ネットワークを介した複数のゲームシステムによってゲームをアップロードや共有したりするのに好ましい。

【0046】

上述したように、本発明の実施の形態によれば、既存のビデオ符号化標準を修正して一連のビデオピクチャ内の所与のピクチャを圧縮および符号化する前に、シーン切替検出をすることが可能となる。ビデオ画像を符号化する際のシーン切替検出の観点为例に説明したが、JPEGのような静止画像の符号化に本発明の実施の形態を利用してよい。

【0047】

他の実施の形態によれば、上述した画像符号化と連動するシーン切替検出を実行する命令は、コンピュータ読み出し可能な記録媒体に格納されてもよい。限定されない例として、図8は、コンピュータ読み出し可能な記録媒体800を例示する。記録媒体は、コンピュータプロセッシング装置から読み出して解釈可能な形で、コンピュータ読み出し可能な命令を含む。限定されない例として、コンピュータ読み出し可能な記録媒体800は、RAM (Random Access Memory)、ROM (Read Only Memory)、ハードディスクドライブ等のコンピュータ読み出し可能なストレージディスク、またはリムーバブルディスクである。さらに、コンピュータ読み出し可能な記録媒体800は、フラッシュメモリ装置、コンピュータ読み出し可能なテープ、CD-ROM、DVD-ROM、Blu-ray、HD-DVD、UMD、またはその他の光学記録媒体であってもよい。

【0048】

記録媒体800は、プロセッサによって実行されるとシーン切替検出を実施するように構成されたシーン切替検出命令801を含む。シーン切替検出命令801は、現在のピクチャ内の各マクロブロックについて統計量 $\mu_m$ の計算を実装する1以上の統計計算命令802を含む。加えて、シーン切替検出命令は、先行するピクチャ内の対応する位置にあるマクロブロック周辺にウィンドウを規定するように構成された1以上のウィンドウ規定命令803を含む。さらに、シーン切替検出命令801は、ウィンドウ内のマクロブロック全体にわたって統計的総和Eを計算する1以上の統計的総和命令804を含む。その上、シーン切替検出命令801は、現在のピクチャ内の所与のマクロブロックにおける $\mu_m$ と統計的総和Eとの差分を計算するように構成された差分計算命令805を含む。シーン切替検出命令801はまた、Eと $\mu_m$ と差分を用いて所与のマクロブロックがシーン切替マクロブロックか否かを決定するシーン切替マクロブロック決定命令806も含む。シーン切替検出命令801はまた、いくつかのシーン切替マクロブロックを用いて現在のピクチャがシーン切替ピクチャであるか否かを決定するシーン切替マクロブロックカウント命令807も含む。

【0049】

記録媒体800は、現在のピクチャがシーン切替ピクチャであると決定されたか否かに部分的に基づいて、そのピクチャをIフレーム、IDRフレーム、Pフレーム、またはBフレームとして符号化するように構成された1以上のピクチャ符号化命令810をオプションとして含んでもよい。

【0050】

上記は、本発明の好ましい実施の形態の完全な説明であるが、種々の代替、変更及び等価物を用いることができる。したがって、本発明の範囲は、上記の説明を参照して決定されるべきではなく、添付された特許請求の範囲をそれらと等価な範囲の全てとともに参照して決定されるべきである。ここで記述される全ての特徴は、好ましいか否かにかかわらず、ここで記述される他の全ての特徴に結合されてもよい。特許請求の範囲において、不定冠詞に続くものは、別段の明示がない限り、1以上の事項の数量をさす。添付された特許請求の範囲は、「～するための手段」という語句を用いて明示的に限定されない限り、ミーンズプラスファンクションの限定を含むものと解釈されるべきではない。

10

20

30

40

【図1A】

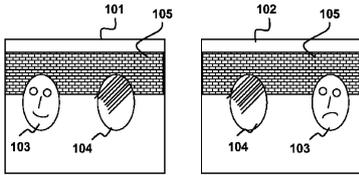


FIG. 1A

【図1B】

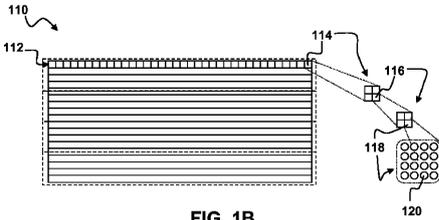
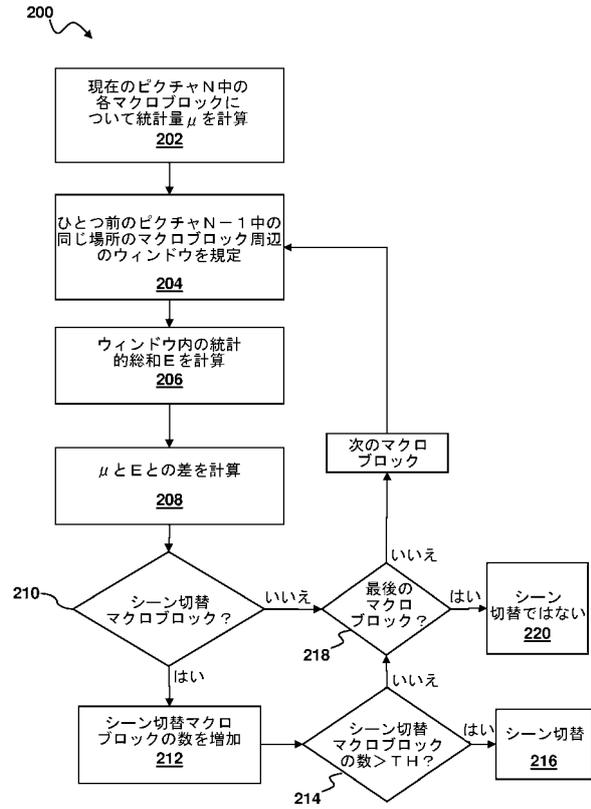
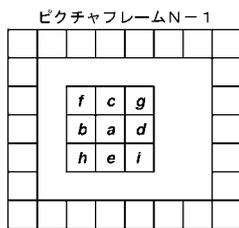


FIG. 1B

【図2A】



【図2B】



【図3B】

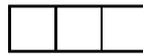
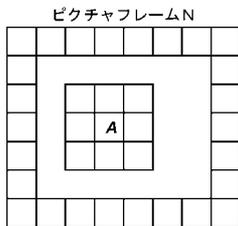


FIG. 3B

【図2C】



【図3C】

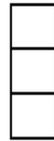


FIG. 3C

【図3A】

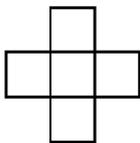


FIG. 3A

【図4A】

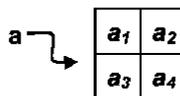


FIG. 4A

【 図 4 B 】

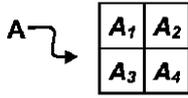


FIG. 4B

【 図 5 A 】

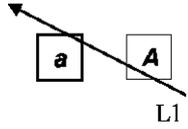


FIG. 5A

【 図 5 B 】

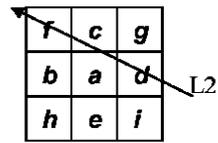
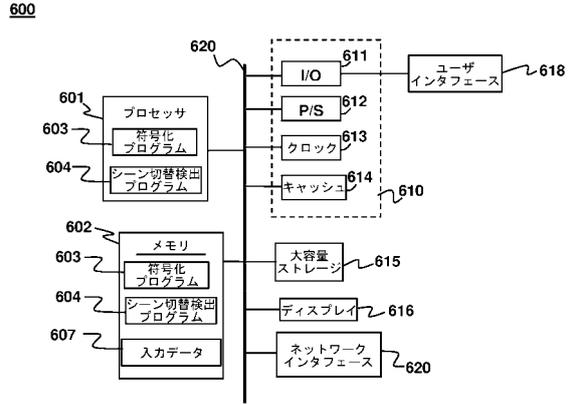
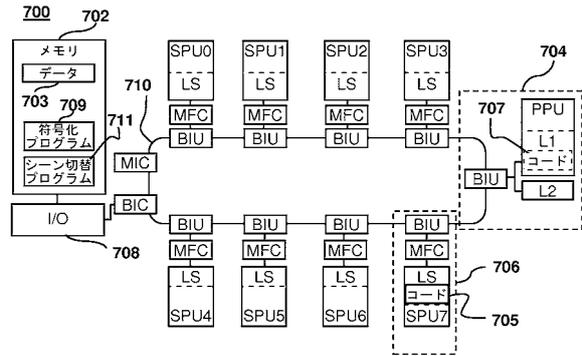


FIG. 5B

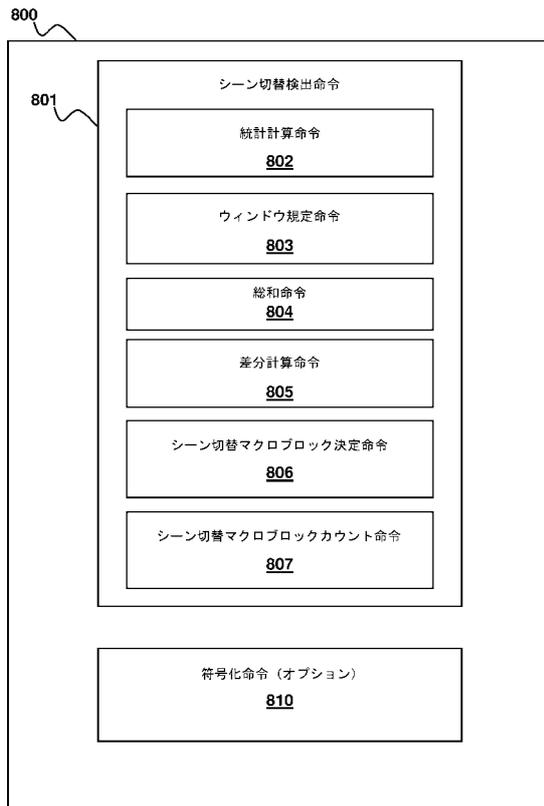
【 図 6 】



【 図 7 】



【 図 8 】



フロントページの続き

(58)調査した分野(Int.Cl. , DB名)

H04N 19/00 - 19/98