

(19) 日本国特許庁(JP)

## (12) 公表特許公報(A)

(11) 特許出願公表番号

特表2006-509463  
(P2006-509463A)

(43) 公表日 平成18年3月16日(2006.3.16)

(51) Int.CI.

HO4N 7/32 (2006.01)  
HO3M 7/36 (2006.01)

F 1

HO4N 7/137  
HO3M 7/36

テーマコード(参考)

5C059  
5J064

審査請求 未請求 予備審査請求 有 (全 15 頁)

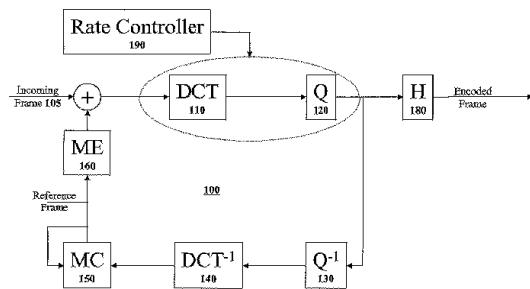
(21) 出願番号 特願2004-559042 (P2004-559042)  
 (86) (22) 出願日 平成15年8月7日 (2003.8.7)  
 (85) 翻訳文提出日 平成17年5月27日 (2005.5.27)  
 (86) 國際出願番号 PCT/US2003/024953  
 (87) 國際公開番号 WO2004/054257  
 (87) 國際公開日 平成16年6月24日 (2004.6.24)  
 (31) 優先権主張番号 10/313,773  
 (32) 優先日 平成14年12月6日 (2002.12.6)  
 (33) 優先権主張国 米国(US)

(71) 出願人 505012737  
 アップル コンピュータ、インコーポレイテッド  
 アメリカ合衆国、95014 カリフォルニア州、クーペルティーノ、メイル スト ップ: 38-ピーエーティ、ワン インフィニット ループ  
 (74) 代理人 100086461  
 弁理士 齋藤 和則  
 (74) 代理人 100086287  
 弁理士 伊東 哲也  
 (72) 発明者 ハスケル、バリン、ジー。  
 アメリカ合衆国、94040-3960  
 カリフォルニア州、マウンテン ヴュー、  
 フェアブルック ドクター。 1190  
 最終頁に続く

(54) 【発明の名称】除算演算の要件が低減されたデジタル・ビデオ符号化の可変精度画像間タイミング指定の方法および装置

## (57) 【要約】

デジタル・ビデオ・システムで動き予測を実行するための方法および装置が開示される(図1)。特に、本発明は、極めて効率的な方法で予測動きベクトルを迅速に計算するシステムを開示する(図1、項目160)。一実施形態では、第1のビデオ画像と第2のビデオ画像との第1の表示時間差に2の累乗値を乗算することで第1の被乗数が決定される(図1、項目150、160)。このステップで比率の分子が拡大する(図1、項目120)。次に、システムはその変倍された分子を前記第2のビデオ画像と第3のビデオ画像の第2の表示時間差で除算することで変倍された比率を決定する。次に、変倍された比率は記憶され、動きベクトル予測の計算を行う。変倍された比率を記憶することすべての予測動きベクトルは高い精度で迅速に計算できる。これは、変倍された比率が有効ビットを保存し、比率の低減が簡単なため送りで実行されるためである(図1、項目180)。



**【特許請求の範囲】****【請求項 1】**

デジタル・ビデオ・システムで動き予測を実行する方法であって、

第1のビデオ画像と第2のビデオ画像との第1の表示時間差に2の累乗値を乗算することで第1の被乗数を決定するステップと、

前記被乗数を前記第2のビデオ画像と第3のビデオ画像の第2の表示時間差で除算することで変倍された比率を決定するステップと、

動きベクトル予測の計算のために前記変倍された比率を記憶するステップとを含む方法。  
。

**【請求項 2】**

前記変倍された比率に前記第1のビデオ画像に関連する動きベクトルを乗算することで変倍された動きベクトルを決定するステップをさらに含む請求項1に記載のデジタル・ビデオ・システムで動き予測を実行する方法。

**【請求項 3】**

前記変倍された動きベクトルを前記2の累乗値で除算するステップをさらに含む請求項2に記載のデジタル・ビデオ・システムで動き予測を実行する方法。

**【請求項 4】**

前記除算が前記変倍された動きベクトルのビットを受けた送りすることで実行される請求項3に記載のデジタル・ビデオ・システムで動きベクトル予測を実行する方法。

**【請求項 5】**

前記2の累乗値が256を含む請求項1に記載のデジタル・ビデオ・システムで動き予測を実行する方法。

**【請求項 6】**

前記2の累乗値が通信チャネルを通して周期的に送信される請求項1に記載のデジタル・ビデオ・システムで動き予測を実行する方法。

**【請求項 7】**

除算によって前記変倍された比率を決定する前記ステップがビデオ画像あたり1回だけ実行される請求項1に記載のデジタル・ビデオ・システムで動き予測を実行する方法。

**【請求項 8】**

前記変倍された比率が切り捨てされる請求項1に記載のデジタル・ビデオ・システムで動き予測を実行する方法。

**【請求項 9】**

前記変倍された比率が最も近い整数に丸められる請求項1に記載のデジタル・ビデオ・システムで動き予測を実行する方法。

**【請求項 10】**

前記2の累乗値が定数を含む請求項1に記載のデジタル・ビデオ・システムで動き予測を実行する方法。

**【請求項 11】**

コンピュータ可読媒体であって、該コンピュータ可読媒体が動き予測を実行する一組のコンピュータ命令を有し、前記一組のコンピュータ命令が、  
40

第1のビデオ画像と第2のビデオ画像との第1の表示時間差に2の累乗値を乗算することで第1の被乗数を決定するステップと、

前記被乗数を前記第2のビデオ画像と第3ビデオ画像の第2の表示時間差で除算することで変倍された比率を決定するステップと、

動きベクトル予測の計算のために前記変倍された比率を記憶するステップとを含む一組のステップを実施するコンピュータ可読媒体。

**【請求項 12】**

前記コンピュータ命令が、

前記変倍された比率に前記第1のビデオ画像に関連する動きベクトルを乗算することで変倍された動きベクトルを決定するステップをさらに実行する請求項11に記載のコンピュ  
50

一 タ可読媒体。

【請求項 1 3】

前記コンピュータ命令が、

前記変倍された動きベクトルを前記 2 の累乗値で除算するステップをさらに実行する請求項 1 2 に記載のコンピュータ可読媒体。

【請求項 1 4】

前記除算が前記変倍された動きベクトルのビットを受けた送りすることで実行される請求項 1 3 に記載のコンピュータ可読媒体。

【請求項 1 5】

前記 2 の累乗値が 2 5 6 を含む請求項 1 1 に記載のコンピュータ可読媒体。 10

【請求項 1 6】

前記 2 の累乗値が通信チャネルを通して周期的に送信される請求項 1 1 に記載のコンピュータ可読媒体。

【請求項 1 7】

除算によって前記変倍された比率を決定する前記ステップがビデオ画像あたり 1 回だけ実行される請求項 1 1 に記載のコンピュータ可読媒体。

【請求項 1 8】

前記変倍された比率が切り捨てされる請求項 1 1 に記載のコンピュータ可読媒体。

【請求項 1 9】

前記変倍された比率が最も近い整数に丸められる請求項 1 1 に記載のコンピュータ可読媒体。 20

【請求項 2 0】

前記 2 の累乗値が定数を含む請求項 1 1 に記載のコンピュータ可読媒体。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

(関連出願)

本特許出願は、2002年12月6日出願の米国特許出願番号 10 / 313,773 について、米国特許法第 120 条により先の出願日の利益を主張する。 30

【0 0 0 2】

(技術分野)

本発明は、マルチメディア圧縮システムの分野に関する。特に、本発明は、プロセッサ集約型除算演算の要件が低減された可変精度画像間タイミング指定の方法およびシステムを開示する。

【背景技術】

【0 0 0 3】

デジタル・ベースの電子媒体フォーマットはついに大幅にアナログ電子媒体フォーマットに取って代わる段階に達した。デジタル・コンパクト・ディスク (CD) はかなり以前にアナログのビニール製レコードに取って代わっている。アナログ磁気カセット・テープはますます稀になっている。ミニディスクおよび MP3 (MPEG オーディオ・レイヤ 3) などの第 2 世代および第 3 世代デジタル・オーディオ・システムは、コンパクト・ディスクの第 1 世代デジタル・オーディオ・フォーマットから市場シェアを奪いつつある。 40

ビデオ媒体フォーマットは、オーディオ媒体と比較してデジタル記憶装置およびデジタル伝送フォーマットへの移行が遅れている。このデジタルへの移行の遅さの理由は、許容できる品質の映像をデジタル形式で正確に表現するのに必要な大量のデジタル情報と圧縮された映像を符号化するのに必要な高速処理能力とに主として起因する。映像を正確に表現するのに必要な大量のデジタル情報は、極めて大容量のデジタル記憶システムおよび高帯域伝送システムを必要とする。

【0 0 0 4】

50

しかし、映像はデジタル記憶装置および伝送フォーマットに急速に移行しつつある。より高速のコンピュータ・プロセッサ、高密度記憶システム、および新しい効率的な圧縮および符号化アルゴリズムによって、デジタル・ビデオ伝送および記憶装置は消費者価格にまで引き下げる。デジタル・ビデオ・システムであるDVD（デジタル多用途ディスク）は、数年にわたって最もよく売れている消費者向け電子製品の1つである。DVDは、その高品質の映像、極めて高品質の音響、利便性および他の特徴によって、好んで選ばれる事前録画されたビデオ再生システムとして、ビデオカセット・レコーダ（VCR）に急速に取って代わっている。古くなったアナログNTSC（米国テレビジョン方式委員会）のビデオ伝送システムは現在、デジタルATSC（高度テレビジョン方式委員会）のビデオ伝送システムに移行しつつある。

10

コンピュータ・システムは何年もの間、さまざまな異なるデジタル・ビデオ符号化フォーマットを使用してきた。特に、コンピュータ・システムはデジタル映像をそれぞれ圧縮および符号化、または圧縮解除および復号化する異なるビデオ符号器／復号器の方法を使用してきた。ハードウェアまたはソフトウェア実施態様でのビデオ符号器／復号器の方法は、一般に「コーデック」と呼ばれる。

#### 【0005】

コンピュータ・システムによって使用してきた最良のデジタル・ビデオ圧縮および符号化システムには、頭字語MPEGで一般に周知である動画映像専門家グループ（Motion Pictures Expert Group）が支援するデジタル・ビデオ・システムが含まれる。MPEGの3つの最も知られ、使用されているデジタル・ビデオ・フォーマットは、MPEG-1、MPEG-2、およびMPEG-4である。ビデオCD（VCD）および消費者グレードのデジタル・ビデオ編集システムは初期のMPEG-1デジタル・ビデオ符号化フォーマットを使用する。デジタル多用途ディスク（DVD）およびディッシュ・ネットワーク・ブランドの直接放送衛星（DBS）テレビジョン放送システムは高品質MPEG-2デジタル・ビデオ圧縮および符号化システムを使用する。MPEG-4符号化システムは、最新のコンピュータ・ベースのデジタル・ビデオ・エンコーダおよびそれに関連するデジタル・ビデオ・プレーヤに急速に採用されている。

20

#### 【0006】

MPEG-2およびMPEG-4標準は一連のビデオ・フレームまたはビデオ・フィールドを圧縮し、次に圧縮されたフレームまたはフィールドをデジタル・ビットストリームに符号化する。ビデオ・フレームまたはフィールドをMPEG-2およびMPEG-4システムで符号化する時には、ビデオ・フレームまたはフィールドは画素ブロックの矩形の格子に分割される。各画素ブロックは独立して圧縮および符号化される。

30

ビデオ・フレームまたはフィールドを圧縮する時には、MPEG-4標準はフレームまたはフィールドを、イントラフレーム（Iフレーム）、片方向予測フレーム（Pフレーム）、または双方向予測フレーム（Bフレーム）という3つのタイプの圧縮フレームまたはフィールドのうちの1つに圧縮できる。イントラフレームは他のビデオ・フレームを参照せずに完全に独立して個別のビデオ・フレームを符号化する。Pフレームは、単一の以前に表示されたビデオ・フレームを参照してビデオ・フレームを定義する。Bフレームは、現在のフレーム以前に表示されたビデオ・フレームと現在のフレーム以後に表示されるビデオ・フレームの両方を参照してビデオ・フレームを定義する。冗長ビデオ情報を効率的に使用することで、PフレームおよびBフレームは最良の圧縮を提供する。

40

#### 【発明の開示】

#### 【0007】

ビデオ・コーデックで動き予測を実行する方法および装置を開示する。特に、本発明は、過剰な除算回数を必要とせずに極めて効率的な方法で予測動きベクトルを迅速に計算するシステムを開示する。

一実施形態では、第1のビデオ画像と第2のビデオ画像との第1の表示時間差に2の累乗値を乗算することで第1の被乗数が決定される。このステップで比率の分子が拡大する。次に、システムは、その変倍された分子を前記第2のビデオ画像と第3のビデオ画像の

50

第 2 の表示時間差で除算することで変倍された比率を決定する。次に、変倍された比率は記憶され、後で動きベクトル予測の計算に使用される。変倍された比率を記憶することですべての予測動きベクトルは高い精度で迅速に計算できる。これは、変倍された比率が有効ビットを保存し、倍率の低減が簡単なけた送りで実行されるために時間がかかる除算演算の必要を解消するためである。

本発明の他の目的、特徴、および利点は、添付の図面および以下の詳細な説明から明らかになるだろう。

#### 【 0 0 0 8 】

##### ( 図面の簡単な説明 )

本発明の目的、特徴、および利点は、以下の詳細な説明を読めば当業者に明らかになるだろう。

10

図 1 は、1 つの可能なデジタル・ビデオ・エンコーダ・システムの高レベル・ブロック図である。

図 2 は、異なる画像を結ぶ矢印が動き補償を用いて作成される画像間依存を示す順序で画像が表示される一連のビデオ画像を示す図である。

図 3 は、異なる画像を結ぶ矢印が動き補償を用いて作成される画像間依存を示す画像の好みの送信順序で図 2 の画像をリストしたビデオ画像を示す図である。

図 4 は、互いを参照するビデオ画像間の距離が 2 の累乗になるように選択される一連のビデオ画像を示す概念図である。

20

#### 【 発明を実施するための最良の形態 】

#### 【 0 0 0 9 】

除算演算の要件が低減されたマルチメディア圧縮および符号化システムの可変精度画像間タイミング指定の方法およびシステムを開示する。以下の説明では、説明を分かりやすくするために、特定の命名法を用いて本発明の徹底的な理解を図る。しかし、本発明を実施する際にこれらの特定の詳細内容が必要ではないことは当業者には明らかであろう。例えば、本発明を M P E G マルチメディア圧縮および符号化システムに関して記載している。しかし、他のタイプの圧縮および符号化システムにも同じ技法を容易に適用することができる。

#### 【 0 0 1 0 】

##### マルチメディア圧縮および符号化の概要

30

図 1 は、当業技術分野で周知である典型的なデジタル・ビデオ・エンコーダ 1 0 0 の高レベル・ブロック図である。デジタル・ビデオ・エンコーダ 1 0 0 は、ブロック図の左側でビデオ・フレームの着信ビデオ・ストリーム 1 0 5 を受信する。デジタル・ビデオ・エンコーダ 1 0 0 は、各ビデオ・フレームを画素ブロックの格子に分割する。画素ブロックは別々に圧縮される。さまざまな異なるサイズの画素ブロックを異なるビデオ符号化システムが使用できる。例えば、異なる画素ブロック解像度は、 $8 \times 8$ 、 $8 \times 4$ 、 $16 \times 8$ 、 $4 \times 4$  等を含む。さらに、画素ブロックは時折「マクロブロック」と呼ばれる。本明細書では、いかなるサイズの画素のいかなるブロックを指す場合でも画素ブロックという用語を使用する。

離散コサイン変換 ( D C T ) ユニット 1 1 0 はビデオ・フレーム内の各画素ブロックを処理する。フレームは独立して処理でき（イントラフレーム）、または動き補償ユニットから受信した他のフレームからの情報に関連して処理できる（インターフレーム）。次に、量子化 ( Q ) ユニット 1 2 0 は離散コサイン変換ユニット 1 1 0 からの情報を量子化する。最後に、量子化されたビデオ・フレームはエントロピー・エンコーダ ( H ) ユニット 1 8 0 によって符号化され、符号化されたビットストリームを生成する。エントロピー・エンコーダ ( H ) ユニット 1 8 0 は可変長符号化 ( V L C ) システムを使用できる。

40

#### 【 0 0 1 1 】

フレーム間符号化ビデオ・フレームは他の近くのビデオ・フレームに関連して定義されるので、デジタル・ビデオ・エンコーダ 1 0 0 は、インターフレームが符号化できるように、各復号化されたフレームがデジタルビデオ・デコーダ内でどのように表示されるかの

50

複製を作成する必要がある。それ故、デジタル・ビデオ・エンコーダ100の下の部分は実際にデジタル・ビデオ・デコーダ・システムになる。特に、逆量子化器( $Q^{-1}$ )ユニット130はビデオ・フレーム情報の量子化を反転し、逆離散コサイン変換(DCT $^{-1}$ )ユニット140はビデオ・フレーム情報の離散コサイン変換を反転する。逆離散コサイン変換(DCT $^{-1}$ )ユニット140からすべてのDCT係数が再構成されると、動き補償ユニットが動きベクトルと共にこの情報を使用して符号化されたビデオ・フレームを再構成する。次に、再構成されたビデオ・フレームは以後のフレームの動き予測の基準ビデオ・フレームとして使用される。

#### 【0012】

次に、復号化されたビデオ・フレームを用いて復号化されたビデオ・フレーム内の情報について定義されたインターフレーム(PフレームまたはBフレーム)を符号化することができる。特に、動き補償(MC)ユニット150および動き予測(ME)ユニット160を用いて動きベクトルを決定し、インターフレームの符号化に用いる差分値を生成することができる。

レート・コントローラ190はデジタル・ビデオ・エンコーダ100内の多くの異なる構成要素から情報を受信し、この情報を用いてビデオ・フレームごとにビット・バジェットを割り当てる。レート・コントローラ190は、指定された一組の制約事項に適合する最高品質のデジタル・ビデオ・ビット・ストリームを生成するようにビット・バジェットを割り当てなければならない。特に、レート・コントローラ190は、バッファのオーバーフロー(記憶できる以上の情報を送信することでビデオ・デコーダ内に利用可能なメモリの量を超過すること)、またはバッファのアンダフロー(ビデオ・フレームの送信が遅く、ビデオ・デコーダが表示するビデオ・フレームが不足すること)を引き起こすことなく、最高品質の圧縮ビデオ・ストリームを生成しようとする。

#### 【0013】

##### 画素ブロックによるデジタル・ビデオ符号化

一部のビデオ信号では、連続するビデオ画像(フレームまたはフィールド)の間の時間は一定ではない。(注:本明細書ではビデオ画像という用語で一般にビデオ・フレームまたはビデオ・フィールドを指す。)例えば、伝送帯域幅の制約によって一部のビデオ画像が欠落することがある。さらに、カメラの不規則性あるいはスローモーションまたはファストモーションのような特殊効果のためにビデオ・タイミングも異なる場合がある。一部のビデオ・ストリームでは、元のビデオ・ソースは設計による不均一の画像間時間を単に有することがある。例えば、コンピュータ・グラフィック・アニメーションなどの合成ビデオは、ビデオ・カメラ・システムなどの均一のタイミングのビデオ・キャプチャー・システムによって任意のビデオ・タイミングを課されないために、不均一のタイミングを有する場合がある。柔軟なデジタル・ビデオ符号化システムは不均一のビデオ画像タイミングを扱うことができなければならない。

前述したように、大半のデジタル・ビデオ符号化システムはビデオ画像を画素ブロックの矩形の格子に分割する。ビデオ画像内の各々の個々の画素ブロックは別々に圧縮され符号化される。例えば、ISO MPEGまたはITU H.264などのいくつかのビデオ符号化標準は、別のタイプの予測画素ブロックを用いてビデオ画像を符号化する。1つのシナリオでは、画素ブロックは以下の3つのタイプのうちのいずれかである。

1. I画素ブロック - イントラ(I)画素ブロックは符号化に際して他のビデオ画像の情報を使用しない(イントラ画素ブロックは完全に自己定義される)。

2. P画素ブロック - 片方向予測(P)画素ブロックは以前のビデオ画像の画像情報を参照する。

3. B画素ブロック - 双方向予測(B)画素ブロックは以前および将来のビデオ画像の情報を参照する。

#### 【0014】

ビデオ画像内の画素ブロックのすべてがイントラ画素ブロックの場合、ビデオ画像はイントラフレームである。ビデオ画像が片方向予測マクロブロックまたはイントラ画素ブロ

10

20

30

40

50

ックのみを含む場合、ビデオ画像は P フレームと呼ばれる。ビデオ画像が双方向予測画素ブロックを含む場合、ビデオ画像は B フレームと呼ばれる。話を分かりやすくするために、本明細書では所与の画像内のすべての画素ブロックが同じタイプである場合を想定する。

符号化されるビデオ画像のシーケンスの例は以下のように表現される。

I<sub>1</sub> B<sub>2</sub> B<sub>3</sub> B<sub>4</sub> P<sub>5</sub> B<sub>6</sub> B<sub>7</sub> B<sub>8</sub> B<sub>9</sub> P<sub>10</sub> B<sub>11</sub> P<sub>12</sub> B<sub>13</sub> I<sub>14</sub> . . .

ここで、文字 (I、P、またはB) は、デジタル・ビデオ画像が I フレーム、P フレーム、または B フレームのいずれかであるか、番号はビデオ画像のシーケンス内のビデオ画像のカメラの順序を表す。カメラ順序はカメラがそのビデオ画像を記録した順序で、ビデオ画像を表示する順序 (表示順) でもある。

10

### 【0015】

上記の一連のビデオ画像の例を図 2 に概念図で示す。図 2 を参照すると、矢印は記憶された画像 (この場合、I フレームまたは P フレーム) の画素ブロックが他の画像の動き補償予測で使用されることを示す。

図 2 のシナリオでは、他の画像の情報はイントラフレーム・ビデオ画像 I<sub>1</sub> の符号化で使用されない。ビデオ画像 P<sub>5</sub> は、符号化に以前のビデオ画像 I<sub>1</sub> 内のビデオ情報を使用する P フレームであり、したがって、ビデオ画像 I<sub>1</sub> からビデオ画像 P<sub>5</sub> に矢印が引かれている。ビデオ画像 B<sub>2</sub>、ビデオ画像 B<sub>3</sub>、ビデオ画像 B<sub>4</sub> はすべてその符号化にビデオ画像 I<sub>1</sub> およびビデオ画像 P<sub>5</sub> 両方の情報を使用する。したがって、矢印がビデオ画像 I<sub>1</sub> およびビデオ画像 P<sub>5</sub> からビデオ画像 B<sub>2</sub>、ビデオ画像 B<sub>3</sub>、およびビデオ画像 B<sub>4</sub> に引かれている。上記のように、画像間時間は一般に同じではない。

B 画像は将来の画像 (後で表示される画像) の情報を使用するので、送信順は普通表示順とは異なる。特に、他のビデオ画像を構成するのに必要なビデオ画像は最初に送信しなければならない。上記のシーケンスでは、送信順は以下の通りである。

I<sub>1</sub> P<sub>5</sub> B<sub>2</sub> B<sub>3</sub> B<sub>4</sub> P<sub>10</sub> B<sub>6</sub> B<sub>7</sub> B<sub>8</sub> B<sub>9</sub> P<sub>12</sub> B<sub>11</sub> I<sub>14</sub> B<sub>13</sub> . . .

### 【0016】

図 3 は図 2 のビデオ画像の以前の送信順を示す。さらに、図中の矢印は記憶されたビデオ画像 (この場合、I または P) の画素ブロックが他のビデオ画像の動き補償予測に使用されることを示す。

図 3 を参照すると、システムは最初に他のいかなるフレームにも依存しない I フレーム I<sub>1</sub> を送信する。次に、システムはビデオ画像 I<sub>1</sub> に依存する P フレーム・ビデオ画像 P<sub>5</sub> を送信する。次に、ビデオ画像 B<sub>2</sub> がビデオ画像 P<sub>5</sub> の前に表示されるとしても、システムはビデオ画像 P<sub>5</sub> の後に B フレーム・ビデオ画像 B<sub>2</sub> を送信する。これは、ビデオ画像 B<sub>2</sub> を復号化する時点では、デコーダはビデオ画像 B<sub>2</sub> を復号化するのに必要なビデオ画像 I<sub>1</sub> および P<sub>5</sub> 内の情報を受信し、記憶しているからである。同様に、ビデオ画像 I<sub>1</sub> および P<sub>5</sub> をただちに用いて次のビデオ画像 B<sub>3</sub> およびビデオ画像 B<sub>4</sub> を復号化することができる。受信機 / デコーダは正しく表示するためにビデオ画像シーケンスを並び替える。この動作で、I および P 画像は記憶画像と呼ばれることが多い。

P フレーム画像の符号化は、通常、動きベクトルが画像内の画素ブロックごとに計算される動き補償を使用する。計算された動きベクトルを用いて、前述の以前の画像の画素を並進して予測画素ブロック (P 画素ブロック) を形成できる。次に、P フレーム画像内の実際の画素ブロックと予測画素ブロックとの差が符号化されて送信される。

### 【0017】

#### P 画像

P 画像の符号化は、通常、以前の画像のある位置をポイントする動きベクトル (MV) が現在の画像内の画素ブロックごとに計算される動き補償 (MC) を使用する。動きベクトルを用いて、前述の以前の画像の画素を並進して予測画素ブロックを形成できる。次に、P 画像内の実際の画素ブロックと予測画素ブロックとの差が符号化されて送信される。

また、各動きベクトルは予測符号化によって送信できる。例えば、動きベクトル予測は近くの動きベクトルを用いて形成できる。その場合、実際の動きベクトルと予測動きベク

20

30

40

50

トルとの差が符号化されて送信される。

### 【0018】

#### B 画像

各 B 画素ブロックは、前述した以前のビデオ画像を参照する第 1 の動きベクトルと、以後のビデオ画像を参照する第 2 の動きベクトルという 2 つの動きベクトルを使用する。これら 2 つの動きベクトルから 2 つの予測画素ブロックが計算される。次に、2 つの予測画素ブロックはある機能を用いて結合され、最終予測画素ブロックが形成される。上記と同様、B フレーム画像の実際の画素ブロックと最終予測画素ブロックとの差が符号化され送信される。

P 画素ブロックの場合、予測符号化を用いて B 画素ブロックの各動きベクトル (MV) を送信することもできる。特に、予測動きベクトルは近くの動きベクトルを用いて形成される。次に、実際の動きベクトルと予測動きベクトルとの差が符号化され送信される。

しかし、B 画素ブロックの場合、最も近くの記憶画像画素ブロック内の動きベクトルから動きベクトルを補間する余地がある。そのような動きベクトル補間はデジタル・ビデオ・エンコーダとデジタル・ビデオ・デコーダの両方で実行される。

この動きベクトル補間は固定された背景をゆっくりとパンするカメラによって作成されるビデオ・シーケンスのビデオ画像で特に有効である。実際、そのような動きベクトル補間は単独で使用できる。特に、これは、補間を用いて符号化された上記の B 画素ブロック動きベクトルに関して差分情報を計算したり送信したりする必要がないということを意味する。

### 【0019】

さらに説明を深めるために、上記のシナリオで、画像 i と j との間の画像間表示時間を  $D_{i,j}$  と表す。すなわち、画像の表示時間が、それぞれ、 $T_i$  および  $T_j$  の場合、

$$D_{i,j} = T_i - T_j \quad \text{ここから}$$

$$D_{i,k} = D_{i,j} + D_{j,k}$$

$$D_{i,k} = -D_{k,i}$$

ただし、 $D_{i,j}$  は場合によって負である。

それ故、 $MV_{5,1}$  が  $I_1$  を参照する P<sub>5</sub> 画素ブロックの動きベクトルの場合、B<sub>2</sub>、B<sub>3</sub> および B<sub>4</sub> 内の対応する画素ブロックでは、I<sub>1</sub> および P<sub>5</sub> を参照する動きベクトルはそれぞれ以下の式により補間される。

$$MV_{2,1} = MV_{5,1} * D_{2,1} / D_{5,1}$$

$$MV_{5,2} = MV_{5,1} * D_{5,2} / D_{5,1}$$

$$MV_{3,1} = MV_{5,1} * D_{3,1} / D_{5,1}$$

$$MV_{5,3} = MV_{5,1} * D_{5,3} / D_{5,1}$$

$$MV_{4,1} = MV_{5,1} * D_{4,1} / D_{5,1}$$

$$MV_{5,4} = MV_{5,1} * D_{5,4} / D_{5,1}$$

動きベクトル予測に表示時間の比率が使用されるので、絶対表示時間は必要ないことに留意されたい。それ故、 $D_{i,j}$  画像間表示時間値には相対表示時間を使用することができる。

このシナリオは、例えば、H.264 標準内で生成できる。生成過程で、P または B 画像はその動きベクトル予測に任意の以前に送信された画像を使用できる。それ故、上記の場合、画像 B<sub>3</sub> はその予測に画像 I<sub>1</sub> および画像 B<sub>2</sub> を使用できる。さらに、動きベクトルは補間だけでなく外挿してもよい。それ故、この場合、以下のようになる。

$$MV_{3,1} = MV_{2,1} * D_{3,1} / D_{2,1}$$

このような動きベクトルの外挿（または補間）は動きベクトルの予測符号化の予測プロセスでも使用できる。

### 【0020】

#### 画像間表示時間の符号化

10

20

30

40

50

ビデオ・シーケンスの可変画像間表示時間は、極めて高い符号化効率を得られるように符号化し送信しなければならない。表示時間はビデオ・デコーダの要件を満たすように選択可能な精度を有する。理想的には、符号化システムは、比較的簡単なコンピュータ・システムがデジタル・ビデオを復号化できるように、デコーダのタスクを簡単化すべきである。

可変画像間表示時間は、差分動きベクトル、直接モード動きベクトル、および／または暗黙的B予測ブロック重み付けを計算するためにいくつかの異なるビデオ符号化システムで潜在的に必要である。

ビデオ・シーケンス内の可変画像間表示時間という問題は時間的基準の使用とからみあっている。理想的には、ビデオ・コーデックの出力画像の正確な画素値の導出はその画像が復号化または表示される時間から独立していかなければならない。したがって、タイミングの問題と時間基準とはコーデック・レイヤ外で解決しなければならない。

所望の時間独立の基礎となる符号化関連およびシステム関連の両方の理由がある。ビデオ・コーデックでは、以下の2つの目的で時間基準が使用される。

(1) 基準画像選択の順序の確立

(2) 画像間の動きベクトルの補間

基準画像選択の順序を確立するには、相対位置値を送信するだけでよい。例えば、復号化順のフレーム位置Nと表示順のフレーム位置Mの差、すなわち、N - Mである。そのような実施形態では、タイム・スタンプまたは他の時間基準は必要ない。動きベクトルを補間する場合、時間的な距離が補間距離に関連するならば、時間的な距離が有用である。しかし、動きが非直線の場合にはこれはあてはまらない。したがって、動きベクトル補間の時間情報以外のパラメータを送信することがより適当と思われる。

#### 【0021】

システムに関しては、通常のビデオ・コーデックはビデオ・コーデックが他のビデオ（およびオーディオ）・コーデックと共に存する、より大きいシステムの一部であると予測できる。そのようなマルチ・コーデック・システムでは、優れたシステム・レイヤ形成および設計は、タイミングなどのコーデックに論理的に依存しない一般機能がコーデック外のレイヤによって処理されることを必要とする。各コーデックによる個別のタイミング管理ではなく、システムによるタイミング管理は、同期化などの共通機能の一貫した処理を達成するのに欠かせない。例えば、ビデオ／オーディオ・プレゼンテーションなどの同時に複数のストリームを処理するシステムでは、異なるストリームを同期状態に保つためにタイミング調整がストリーム内に必要になることがある。同様に、異なるクロックを備えたリモート・システムのストリームを処理するシステムでは、リモート・システムとの同期を保つためにタイミング調整が必要である。そのようなタイミング調整はタイム・スタンプを用いて達成できる。例えば、送信機からの「送信側レポート」によってリンクされ、ストリームごとに RTP レイヤ内で RTP に供給されるタイム・スタンプを同期化に使用できる。これらの送信側レポートは、以下の形式をとることができる。

ビデオ RTP タイム・スタンプ X が基準タイム・スタンプ Y と位置合わせされている。

オーディオ RTP タイム・スタンプ W が基準タイム・スタンプ Z と位置合わせされている。

ここで、基準タイム・スタンプのクロック速度は知られており、2つのストリームを位置合わせすることができる。しかし、これらのタイム・スタンプ基準は2つのストリームに関して周期的かつ個別に到着し、2つのストリームの何らかの必要な再位置合わせを引き起こす。これは一般にビデオ・ストリームをオーディオに合わせて調整するかその逆の方法で達成される。タイム・スタンプをシステムが処理する場合、表示される画素値に影響してはならない。より一般的には、時間情報のシステムによる処理はコーデック外部で実行しなければならない。

#### 【実施例】

#### 【0022】

前節で述べたように、不均一の画像間時間の場合の問題は、画像間表示時間値  $D_{i,j}$

10

20

30

40

50

を効率的にデジタル・ビデオ受信機に送信することである。この目標を達成する1つの方法は、第1の画像の後に、画像ごとに現在の画像と最も最近送信された記憶画像との表示時間差をシステムに送信させるというものである。エラー許容力については、画像内で送信を数回繰り返してもよい。例えば、表示時間差をMPEGまたはH.264標準のスライス・ヘッダ内で繰り返すことができる。すべてのスライス・ヘッダが失われた場合、損失画像に依存して復号化情報を得る他の画像もおそらくは復号化できない。

したがって、前節の例に関して、システムは以下の画像間表示時間値を送信する。

$D_{5,1} D_{2,5} D_{3,5} D_{4,5} D_{10,5} D_{6,10} D_{7,10} D_{8,10} D_{9,1}$   
 $D_{12,10} D_{11,12} D_{14,12} D_{13,14} \dots$

動きベクトル予測のために、画像によって、画像間表示時間  $D_{i,j}$  は変化する。例えば、単一のBフレーム画像  $B_6$  のみが2つのPフレーム画像  $P_5$  と  $P_7$  との間に存在する場合、以下の情報のみを送信すれば済む。

$$D_{7,5} = 2 \text{ および } D_{6,7} = -1$$

ここで、 $D_{i,j}$  画像間表示時間値は相対時間値である。

ビデオ画像  $B_6$  がビデオ画像  $P_5$  とビデオ画像  $P_7$  との距離のわずか  $1/4$  の位置にある場合、送信する適当な  $D_{i,j}$  画像間表示時間値は以下のようになる。

$$D_{7,5} = 4 \text{ および } D_{6,7} = -1$$

上記の両方の例では、ビデオ画像  $B_6$  とビデオ画像  $P_7$  との間の表示時間（画像間表示時間  $D_{6,7}$ ）は表示時間の「単位」値として使用されていることに留意されたい。最近の例では、ビデオ画像  $P_5$  とビデオ画像  $P_7$  との間の表示時間差（画像間表示時間  $D_{6,7}$ ）は、4表示時間「単位」( $4 * D_{6,7}$ )である。

### 【0023】

#### 復号化の効率の改善

一般に、動きベクトル予測計算は、除数が2の累乗の場合に大幅に簡略化される。これは、2つの記憶された画像間の  $D_{i,j}$ （画像間時間）が図4に示すように2の累乗として選択される場合に本発明者の実施態様で容易に達成される。あるいは、予測手順はすべての除数を2の累乗に切り捨てるか丸めるように定義できる。

画像間時間が2の累乗の場合、画像間時間の値そのものの代わりに(2の)累乗を送信しさえすればデータ・ビットの数を低減できる。図4は、画像間の距離が2の累乗になるように選択される場合を示すグラフである。この場合、ビデオ画像  $P_1$  とビデオ画像  $P_3$  との間の  $D_{3,1}$  表示時間値2が  $1 (2^1 = 2)$  のためとして送信され、ビデオ画像  $P_1$  とビデオ画像  $P_3$  との間の  $D_{7,3}$  表示時間値4が  $2 (2^2 = 4)$  のためとして送信できる。

あるいは、外挿演算の動きベクトル補間は、分母が2の累乗であるような方法で変倍して任意の所望の精度に近似することができる。（分母に2の累乗がある除算は除算する値のビットをけた送りするだけで実行できる。）例えば、

$$D_{5,4} / D_{5,1} \sim Z_{5,4} / P$$

ただし、値  $P$  は2の累乗で、 $Z_{5,4} = P * D_{5,4} / D_{5,1}$  は最も近い整数に丸められるか切り捨てされる。 $P$  の値はシステムの定数として周期的に送信されるか設定される。一実施形態では、 $P$  の値は  $P = 2^8 = 256$  に設定される。

### 【0024】

この手法の利点は、デコーダが画像あたり  $Z_{5,4}$  を1回だけ計算すれば済むか、または多くの場合にデコーダが  $Z$  値を事前計算して記憶できるという点である。これによって、デコーダは画像内の動きベクトルごとに  $D_{5,1}$  で除算する必要がなくなり、はるかに効率的に動きベクトル補間が実行できる。例えば、通常の動きベクトル計算は以下のようになる。

$$MV_{5,4} = MV_{5,1} * D_{5,4} / D_{5,1}$$

しかし、 $Z_{5,4}$ （ただし  $Z_{5,4} = P * D_{5,4} / D_{5,1}$ ）を計算し記憶する場合、

$$MV_{5,4} = MV_{5,1} * Z_{5,4} / P$$

$P$  値は2の累乗になるように選択されているので、 $P$  による除算は単にビットのけた送り

10

20

30

40

50

である。したがって、ビデオ画像の Z 値が計算されたら、以降の画素ブロックの動きベクトルを計算するのに 1 回の乗算と 1 回のけた送りだけが必要である。さらに、計算時に有効ビットが失われないように、システムはすべての除算を最後に実行することで精度を高く保つことができる。こうして、デコーダはエンコーダと同様に動きベクトル補間を正確に実行して、この方法を探らなければ発生する可能性がある不一致の問題を回避することができる。

除算（2 の累乗による除算以外の）は、加算または乗算よりもデジタル・コンピュータ・システムにとってはるかに計算が集中するタスクであるため、この手法は動きベクトル補間または外挿を使用する画像を再構築するために必要な計算を大幅に削減することができる。10

ある場合には、動きベクトル補間を使用しなくてもよい。ただし、受信機／プレーヤ・システムがビデオ画像を正しい順序で表示するように、受信機／プレーヤ・システムにビデオ画像の表示順序を送信することがなお必要である。この場合、実際の表示時間に関わらず、 $D_{i,j}$  は単に符号付き整数値で足りる。いくつかの応用例では、画像順序を再構成するために符号（正または負）しか必要でない。

#### 【0025】

画像間時間  $D_{i,j}$  は単一の符号付き整数値として送信できる。ただし、 $D_{i,j}$  値の符号化に多くの方法を使用してさらに圧縮をかけることができる。例えば、符号ビットに続けて可変長符号化の大きさを指定すると比較的実施しやすく、符号化の効率が上がる。

使用可能なそのような可変長符号化システムは UVLC（ユニバーサル可変長符号）として知られる。UVLC 可変長符号化システムは以下のコード・ワードで与えられる。20

```

1 =      1
2 =    0 1 0
3 =    0 1 1
4 =  0 0 1 0 0
5 =  0 0 1 0 1
6 =  0 0 1 1 0
7 =  0 0 1 1 1
8 = 0 0 0 1 0 0 0 . .

```

画像間時間を符号化する別の方法は算術符号化の使用である。通常、算術符号化はデータ・ビットの極めて高い圧縮を行う条件付き確率を使用する。30

以上のように、本発明は、動きベクトル予測で使用する画像間表示時間の簡単で強力な符号化および伝送方法とこれら画像間表示時間の復号化方法とを紹介する。画像間表示時間の符号化は、可変長符号化または算術符号化を用いて極めて効率的にすることができる。さらに、ビデオ・コーデックの必要のみを満たすように所望の精度を選択することができる。

以上、マルチメディア圧縮および符号化システムで可変精度画像間タイミングを指定するシステムについて説明した。本発明の範囲を逸脱することなく、本発明の要素の材料および構成をさまざまな変更および修正できることは、当業者には明らかであろう。

#### 【図面の簡単な説明】

#### 【0026】

【図1】1つの可能なデジタル・ビデオ・エンコーダ・システムの高レベル・ブロック図である。

【図2】異なる画像を結ぶ矢印が動き補償を用いて作成される画像間依存を示す順序で画像が表示される一連のビデオ画像を示す図である。

【図3】異なる画像を結ぶ矢印が動き補償を用いて作成される画像間依存を示す画像の好みの送信順序で図2の画像をリストしたビデオ画像を示す図である。

【図4】互いを参照するビデオ画像間の距離が2の累乗になるように選択される一連のビデオ画像を示す概念図である。

#### 【符号の説明】

10

20

40

50

【 0 0 2 7 】

- 1 0 0 デジタル・ビデオ・エンコーダ  
 1 0 5 着信ビデオ・ストリーム  
 1 1 0 離散コサイン変換 (DCT) ユニット  
 1 2 0 量子化 (Q) ユニット  
 1 3 0 逆量子化器 ( $Q^{-1}$ ) ユニット  
 1 4 0 逆離散コサイン変換 ( $DCT^{-1}$ ) ユニット  
 1 5 0 動き補償 (MC) ユニット  
 1 6 0 動き予測 (ME) ユニット  
 1 8 0 エントロピー・エンコーダ (H) ユニット  
 1 9 0 レート・コントローラ

10

【図1】

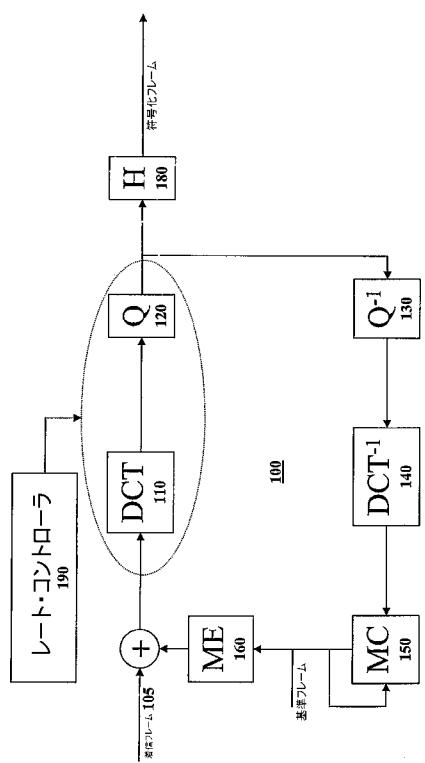


Figure 1

【図2】

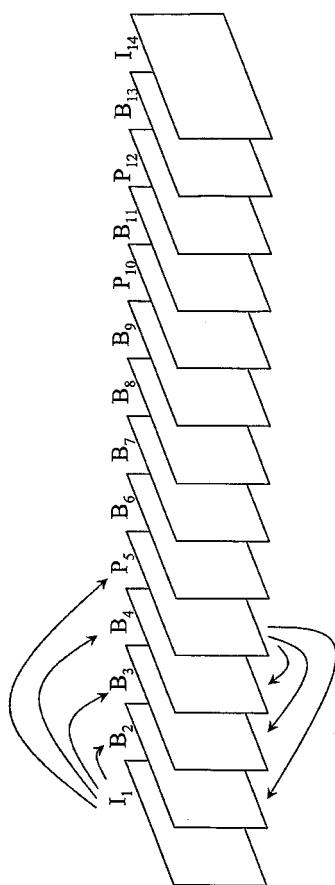


Figure 2

【図3】

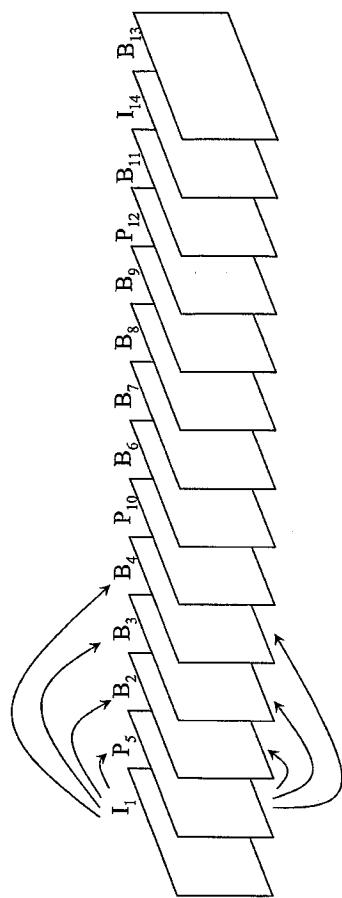


Figure 3

【図4】

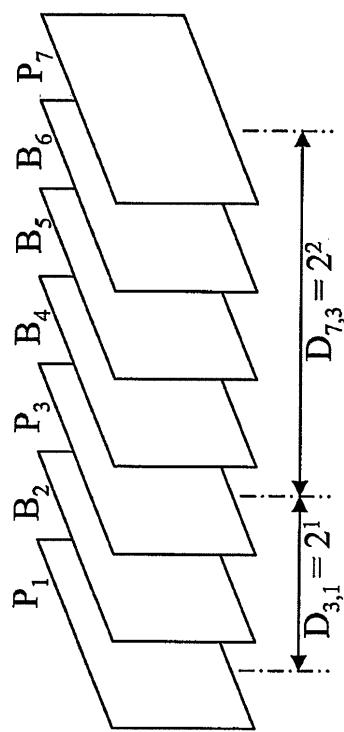


Figure 4

## 【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US03/24953															
<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC(7) : H04N 7/12, 9/74; G06K 9/36 US CL : 375/240.15, 240; 348/581, 584; 382/236 According to International Patent Classification (IPC) or to both national classification and IPC																	
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) U.S. : 375/240.15, 240; 348/581, 584; 382/236																	
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched																	
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)																	
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">Category *</th> <th style="text-align: left;">Citation of document, with indication, where appropriate, of the relevant passages</th> <th style="text-align: left;">Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>US 5,737,023 A (LINZER) 07 April 1998 (07.04.1998), column 1, lines 63-67, and col. 2, lines 1-10.</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>US 6,108,047 A (CHEN) 22 August 2000 (22.08.2000), column 1, lines 45-62.</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>US 6,072,834 A (KIM et al) 06 June 2000 (06.06.2000), column 3, lines 66-67 and col. 4, lines 1-31.</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>US 6,297,852 A (LAKSONO et al) 02 October 2001 (02.10.2001), column 2, lines 37-62.</td> <td>1-20</td> </tr> </tbody> </table>			Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	A	US 5,737,023 A (LINZER) 07 April 1998 (07.04.1998), column 1, lines 63-67, and col. 2, lines 1-10.	1-20	A	US 6,108,047 A (CHEN) 22 August 2000 (22.08.2000), column 1, lines 45-62.	1-20	A	US 6,072,834 A (KIM et al) 06 June 2000 (06.06.2000), column 3, lines 66-67 and col. 4, lines 1-31.	1-20	A	US 6,297,852 A (LAKSONO et al) 02 October 2001 (02.10.2001), column 2, lines 37-62.	1-20
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.															
A	US 5,737,023 A (LINZER) 07 April 1998 (07.04.1998), column 1, lines 63-67, and col. 2, lines 1-10.	1-20															
A	US 6,108,047 A (CHEN) 22 August 2000 (22.08.2000), column 1, lines 45-62.	1-20															
A	US 6,072,834 A (KIM et al) 06 June 2000 (06.06.2000), column 3, lines 66-67 and col. 4, lines 1-31.	1-20															
A	US 6,297,852 A (LAKSONO et al) 02 October 2001 (02.10.2001), column 2, lines 37-62.	1-20															
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.																	
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "B" earlier application or patent published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed																	
Date of the actual completion of the international search  05 January 2004 (05.01.2004)	Date of mailing of the international search report  27 JAN 2004																
Name and mailing address of the ISA/US  Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (703)305-3230	Authorized officer <i>Jean-Sante Hely</i> Gims S Philippe Telephone No. (703) 305-4700																

---

フロントページの続き

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HU,IE,IT,LU,MC,NL,PT,RO,SE,SI,SK,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,PL,PT,RO,RU,SD,SE,SG,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,UZ,VN,YU,ZA,ZW

(72)発明者 シンガー、デヴィッド、ダブリュー。

アメリカ合衆国、94127-1328 カリフォルニア州、サンフランシスコ、ワオナ ストリート 268

(72)発明者 ドゥミトラス、アドリアーナ

アメリカ合衆国、94087 カリフォルニア州、サニーヴェイル、アパートメント 2403、ダブリュー・エル 250 キャミーノ リアル

(72)発明者 ブリ、アチュール

アメリカ合衆国、95014-0632 カリフォルニア州、クーペルティーノ、#4203 ブルンエッジ アヴェニュー 19500

F ターム(参考) 5C059 KK15 KK19 MA00 MA05 MA14 MA23 MC11 MC38 ME01 ME11  
NN01 NN16 NN19 NN21 PP05 PP06 PP07 RB09 RC03 RC04  
RC16 SS20 UA02  
5J064 AA03 BA09 BA16 BB03 BC01 BC09 BC16 BD02 BD03