

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3958305号  
(P3958305)

(45) 発行日 平成19年8月15日(2007.8.15)

(24) 登録日 平成19年5月18日(2007.5.18)

(51) Int. Cl.

F I

H O 4 N 7/32 (2006.01)

H O 4 N 7/137 Z

H O 4 N 7/173 (2006.01)

H O 4 N 7/173 6 1 O Z

請求項の数 7 (全 18 頁)

(21) 出願番号 特願2004-171990 (P2004-171990)  
 (22) 出願日 平成16年6月10日(2004.6.10)  
 (65) 公開番号 特開2005-20729 (P2005-20729A)  
 (43) 公開日 平成17年1月20日(2005.1.20)  
 審査請求日 平成16年6月10日(2004.6.10)  
 (31) 優先権主張番号 10/603428  
 (32) 優先日 平成15年6月24日(2003.6.24)  
 (33) 優先権主張国 米国(US)

(73) 特許権者 503003854  
 ヒューレット・パカード デベロップメン  
 ト カンパニー エル. ピー.  
 アメリカ合衆国 テキサス州 77070  
 ヒューストン 20555 ステイト  
 ハイウェイ 249  
 (74) 代理人 110000039  
 特許業務法人アイ・ピー・エス  
 (72) 発明者 ボー・シェン  
 アメリカ合衆国カリフォルニア州フレモン  
 ト ボガルーサコート685

審査官 岩井 健二

最終頁に続く

(54) 【発明の名称】 ストリーミングメディアを処理する (servicing) 方法およびシステム

(57) 【特許請求の範囲】

【請求項1】

ストリーミングメディアを処理する方法であって、  
 前記ストリーミングメディアを受け取ること(601)と、  
 利用可能な、前記ストリーミングメディアのトランスコーディングに用いられる資源お  
 よび前記トランスコーディングされたストリーミングメディアの記憶に用いられる資源の  
 配分を決定すること(603)と、

前記決定された配分に応じて、前記ストリーミングメディアに対して複数の段階でトラ  
 ンスコーディングを実行すること(605)と、

前記配分が最も有利となるよう、前記複数の段階の中から1つの段階を選択し、前記選  
 択した段階からの中間トランスコード結果をキャッシングすること(607)と

を含むストリーミングメディアを処理する方法。

【請求項2】

フレームレート、ビットレートおよび解像度の少なくとも1つを低減させる複数の演算  
 が行われる

請求項1記載のストリーミングメディアを処理する方法。

【請求項3】

前記ストリーミングメディアのトランスコーディングに用いられる資源は、トランスコ  
 ードであり、前記トランスコーディングされたストリーミングメディアの格納に用いられ  
 る資源は、複数のキャッシュである

10

20

請求項 1 記載のストリーミングメディアを処理する方法。

【請求項 4】

前記複数の段階の中の最後の段階からの最終トランスコード結果が出力されること  
をさらに含む請求項 1 記載のストリーミングメディアを処理する方法。

【請求項 5】

前記キャッシングすることは、

前記ストリーミングメディアの出力ストリームのターゲットビットレートが前記ストリーミングメディアのデータキャッシュレートよりも大きい場合、前記ストリーミングメディアの前記出力ストリームの前記中間トランスコード結果をキャッシングすること

を含む

請求項 1 記載のストリーミングメディアを処理する方法。

【請求項 6】

前記中間トランスコード結果は、

ピクセル、ブロック、マクロブロック、ピクチャおよびシーケンスからなる群から選択されるメタデータ

を含む請求項 5 記載のストリーミングメディアを処理する方法。

【請求項 7】

前記複数の演算は、動きベクトル発生器、ビットレートコントローラ、およびパーサからなる群から選択される資源によって実行される

請求項 2 記載のストリーミングメディアを処理する方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明の実施形態は、包括的にストリーミングメディアに関する。

特に、本発明の実施形態は、ストリーミングメディアを処理する方法およびシステムに関する。

【背景技術】

【0002】

ストリーミングメディアの処理には、記憶資源および演算資源両方を消費する必要がある。

こういった資源に固有の制約により、処理プロセスにおけるボトルネックは確実に、これら制約がシステム資源の適宜配分によって調停されない限り緩和されない。

システム資源の制約は、システム資源の誤った配分に起因する非効率性によって悪化する。

これら非効率性によって助長されるボトルネックは、ストリーミングメディア処理システムの演算資源および記憶資源への過負荷および全体のパフォーマンスの悪化を助長する役割を果たす恐れがある。

【0003】

従来のシステムは、上記処理プロセスボトルネック問題を克服しない。

実際、多くの従来のシステムは、問題を悪化させる可能性のある処理システムの出力からとられる完全に処理済みのデータのキャッシングに的を絞っている。

容易に認められるように、このように完全に処理済みのデータのキャッシングは、記憶資源に対して深刻な負荷をかける。

これは、このような方法が過度の量のシステム記憶資源を消費し、それによって、それ以降の処理を行うのに必要な演算資源の消費が減るためである。

しかし、記憶資源のこのような誤った配分に由来する記憶可用性の欠如に起因するボトルネックは蔓延している。

【非特許文献 1】 Exploring Aggregate Effect with Weighted Transcoding Graphs for Efficient Cache Replacement in Transcoding Proxies, Chang et al., Proceedings of the 18th International Conference on Data Engineering (ICDE '02), 1063-6382/02

10

20

30

40

50

, 2002, IEEE

【発明の開示】

【発明が解決しようとする課題】

【0004】

データの処理および送出行う従来システムの主な欠点は、一種のシステム資源に対する緊張が、別種のシステム資源が豊富に存在すると同時に過度に伸張されることである。

このようなメディア処理の誤った配分に起因する非効率性は、システムパフォーマンスを劣化させるボトルネックを助長する。

【課題を解決するための手段】

10

【0005】

ストリーミングメディアを処理する方法およびシステムが開示される。

一実施形態において、本方法は、ストリーミングメディアを受け取ること、利用可能な処理資源およびメモリ資源の配分を決定することを含む。

さらに、本方法は、ストリーミングメディアに対して多段階処理を実行すること、多段階プロセスの中の1つの段階からの中間結果をキャッシングすることを含む。

中間結果は利用可能な処理資源およびメモリ資源に従って選択される。

【0006】

本明細書に組み込まれて本明細書の一部をなす添付図面は、本発明の実施形態を示し、説明とともに本発明の原理を説明する役割を果たす。

20

【発明を実施するための最良の形態】

【0007】

これより、例を添付図面に示す本発明の実施形態を詳細に参照する。

本発明についてこれら実施形態と併せて説明するが、本発明をこれら実施形態に限定する意図がないことを理解されたい。

反対に、本発明は、添付の特許請求の範囲によって規定される本発明の精神および範囲内に包含し得る代替形態、変更形態、および同等物を保護範囲に含むことを意図する。

さらに、以下の本発明の詳細な説明では、本発明の完全な理解を提供するために多くの特定の詳細が述べられる。

しかし、本発明はこれら特定の詳細なしでも実施し得ることが当業者に認められよう。

30

一方、既知の方法、手順、構成要素、および回路については、本発明の態様を不必要に曖昧にしないように詳細に説明しなかった。

【0008】

[発明の実施形態による例示的なシステム]

図1Aは、本発明の一実施形態によるトランスコードシステムの機能ブロック図100である。

機能ブロックは、図示および考察を明確にするために、図1において別個に示されていることが理解される。

しかし、これらブロックは、別個の要素として存在しなくてもよく、実用では、1つの装置または複数の装置として具現し得る。

40

一般に、本発明のトランスコードシステムは、これから説明する機能を提供する1つまたは複数(の装置)を含むことができる。

【0009】

本発明の実施形態は、データ(たとえば、ストリーミングメディアデータ)の処理およびデータ(たとえば、ストリーミングメディアデータ)の記憶に配分されるシステム資源間の最適なバランスを見つけることによって、システム動作を最適化する。

図1は、入力ストリーム101、メタトランスコード103、出力ストリーム105、資源バランサ107、メタキャッシュ109、正規のキャッシュ111、メタデータ113、オリジナルコンテンツサーバ115、およびクライアント装置(複数可)117を示す。

50

## 【 0 0 1 0 】

入力ストリーム 1 0 1 は、クライアント装置 1 1 7 (たとえば、ラップトップ、デスクトップ、PDA、携帯電話等)に送出することができるストリーミングメディアを搬送する。

一実施形態によれば、入力ストリーム 1 0 1 は、システム資源を最適化する様式でトランスコード 1 0 3 によってトランスコーディングされる。

一実施形態によれば、入力ストリーム 1 0 1 はコンテンツサーバ 1 1 5 (たとえば、ビデオコンテンツプロバイダ等)によって供給される。

他の実施形態によれば、入力ストリームは他のソース(たとえば、ローカル記憶媒体、リモート記憶媒体)から提供される。

10

## 【 0 0 1 1 】

メタトランスコード(たとえば、1 0 3)は、前のトランスコードセッションにおいて生成されキャッシングされたメタデータを現在のトランスコードセッションに採用するトランスコードである。

さらに、メタトランスコードは、現在のセッションにおいて、それ以降の同様のトランスコードセッションに使用されるメタデータを生成することができる。

一実施形態によれば、本発明のメタトランスコード 1 0 3 は、メタデータ(メタキャッシュ 1 0 9 から検索される中間トランスコードデータ)およびデータ全体(たとえば、コンテンツサーバ 1 1 5 によって供給されるものなど)を両方とも入力として受け取ることができることに留意されたい。

20

## 【 0 0 1 2 】

一実施形態では、メタトランスコード 1 0 3 は、入力ストリーム 1 0 1 を提供するオリジナルコンテンツサーバ 1 1 5 と、出力ストリーム 1 0 5 を供給することができるクライアント 1 1 7 のグループとの間に配置される。

メタトランスコード 1 0 3 は、メタキャッシュ 1 0 9 および正規のキャッシュ 1 1 1 に回復することができるキャッシュを備える。

処理されるデータは、トランスコーディングの最終結果および「メタデータ」1 1 3 と呼ばれるトランスコーディングの中間結果の両方からなることができる。

これらデータ構成要素それぞれの最適な使用が、システムパフォーマンスの最適化につながり得る。

30

## 【 0 0 1 3 】

メタトランスコード 1 0 3 によって実行されるトランスコーディングは、多段階プロセスである。

プロセスの1つの段階の結果は通常、少なくとも他の1つの段階への入力として使用される。

本明細書において使用する「中間結果」は、トランスコーディングの最終段階前の各種段階の出力前の各種段階の出力を指し、「最終結果」は、すべての段階を完了したデータを指す。

## 【 0 0 1 4 】

メタキャッシュ 1 0 9 は、メタデータ 1 1 3 のキャッシングに使用される。

40

メタデータ 1 1 3 は、トランスコードプロセスの中間結果と定義することができる。

メタキャッシュ 1 0 9 は、中間トランスコード結果を表すメタデータ 1 1 3 をキャッシングする。

たとえば、データが a、b、および c を加算することによってトランスコーディングされて「X」を得る場合を考える。

中間結果 a + b をメタデータ(たとえば、1 1 3)とみなすことができる。

## 【 0 0 1 5 】

a + b をメタトランスコード 1 0 3 にキャッシングするという決定を行うこともできる。

したがって、「X」を得るための後続のトランスコーディングに必要なのは、2 つ(た

50

たとえば、 $a + b$ の加算、そしてこれに $c$ を加算)ではなく1つの追加処理ステップ(たとえば、 $a + b$ への加算)のみである。

分かるように、これは、処理の節約により $a + b$ の格納に追加のキャッシュ空間の使用が保証される場合に有利である。

本発明は、各種実施形態において、メタトランスコードが利用できる処理資源および記憶資源の最適なバランスを実現する。

#### 【0016】

正規のキャッシュ111は、トランスコードプロセスの最終結果のキャッシングに使用される。

一実施形態によれば、これら結果は、キャッシングに消費する記憶空間が中間結果よりも少ない場合に格納することができる。 10

本実施形態によれば、本発明は、従来のシステムによって提供される単一レベルではなく複数レベルのキャッシュを導入する。

#### 【0017】

本実施形態では、資源バランサ107がキャッシュ資源と処理資源との間の配分を指示する。

資源バランサは、キャッシュ資源よりも多いまたは少ないトランスコード(処理)資源を採用することができるか否かを決定する。

中間トランスコード結果であれ、最終トランスコード結果であれ、いずれもキャッシングすることができる。 20

資源バランサ107は、利用可能なキャッシュ空間および利用可能な演算サイクル等の要因に基づいて決定を行う。

一般に、システムが十分なアイドル演算サイクルを有するが十分なキャッシュ空間を有していない場合、最終トランスコード結果ではなく中間トランスコード結果(たとえば、メタデータ)をキャッシングすることができる。

これとは反対に、システムが十分なキャッシュ空間を有するが十分な演算サイクルを有していない場合、中間結果(たとえば、メタデータ)ではなく最終結果をキャッシングすることができる。

#### 【0018】

一実施形態によれば、メタトランスコーディングはメタデータに基づくトランスコーディングと定義される。 30

現在のトランスコードセッション中、トランスコードはメタデータを生成することができ、そこから、続けて行われ得る同様のトランスコードセッション中に恩恵を受け得ることを理解されたい。

さらに、トランスコードセッション中、トランスコードは、前のトランスコードセッションにおいて生成されたメタデータの恩恵を受けることができる。

#### 【0019】

図1Bは、本発明の一実施形態による例示的なトランスコード構造のブロック図である。

トランスコードセッションは複数の離散したトランスコード段階(たとえば、図1B中の段階1~段階Nとして示す)を含むことができ、それぞれトランスコード完了の中間(非最終)段階に存在するメタデータを出力することを理解されたい。 40

一実施形態によれば、任意のトランスコード段階それぞれから出力されるメタデータは、後続の同様のトランスコードセッションにおいてそれ以降に使用するためにキャッシングすることができる。

例示的な実施形態によれば、メタデータがキャッシングされるトランスコードプロセス段階は、最も有利な演算と記憶のバランスを提供するものが選択されることを理解されたい。

図1Bは、入力ストリーム101、メタトランスコード103、出力ストリーム105、および出力1~出力Nと示された出力を有する(図1B参照)トランスコード段階1~ 50

Nを示す。

【0020】

図1Bを参照すると、入力ストリーム101はオリジナルコンテンツサーバからメタトランスコード103に供給される。

図1Bに示すように、ストリーミングビデオは、最終トランスコード結果の出力ストリーム105が生成される（たとえば、クライアント装置に供給するためなどに）までにN個のトランスコード段階（たとえば、図1Bには段階1～段階Nと示される）を通過しなければならない。

各トランスコード段階のいずれの出力も（たとえば、図1Bでは出力1～出力Nと示される）キャッシングし、メタトランスコードへの入力として使用することができ、それによってそれ以降の同様のセッションにおける演算負荷が低減することを理解されたい。

10

しかし、メタデータが実際にキャッシングされる段階は、最も有利な演算と記憶のバランスを提供するものが選択される。

【0021】

〔キャッシングするメタデータの様々なレベル〕

一実施形態によれば、キャッシングするメタデータ（たとえば、113）の選択には、トランスコードセッションの処理ボトルネックを考慮に入れる必要がある。

文脈として、圧縮されたMPEG（モーションピクチャエキスパートグループ）ビデオストリームに対するビデオトランスコーディングを使用して、以下の図2を参照して行う考察において、これをどのように実現することができるかについて説明する。

20

図2は、本発明の一実施形態によりシーケンスレベルからピクセルレベルにキャッシングすることができる5つの中間レベルのメタデータの間の関係を示すグラフ200である。

図2は、「キャッシュヌル」201、「キャッシュフル」213、演算軸205、記憶軸207、シーケンス209、ピクチャ212、マクロブロック215、ブロック217、およびピクセル219を示す。

【0022】

本発明の一実施形態によれば、図2に示す配分法に従ってトランスコードセッション中に得ることができる空間レベルは2つある。

すなわち、（1）「キャッシュヌル」201、および（2）「キャッシュフル」213である。

30

「キャッシュヌル」201レベルは、最終結果も、また図2に示す5つのレベルのいずれも、トランスコード中の特定のビデオストリーム部分にキャッシングされない場合に得られる。

「キャッシュヌル」201の場合、ヌルキャッシュをもたらしたトランスコードセッション中に生成されたデータに対してそれ以降に要求がある場合、処理システムは、使用可能な最終結果も、またメタデータもキャッシングされていないため、100%の演算資源を要し得る新しいトランスコードセッションを開始する。

しかし、新しいトランスコードセッションでは、必要とされる記憶資源は0%である。

これは、前のトランスコードセッション中に、この空間を占有するものが何もキャッシングされなかったためである。

40

【0023】

「キャッシュフル」213レベルは、トランスコードセッションの最終結果がキャッシングされる場合に得られる。

「キャッシュフル」213の場合、フルキャッシュをもたらしたトランスコードセッション中に生成されたデータに対してそれ以降に要求がある場合、トランスコードは、キャッシュ記憶装置がフルであるため、100%の記憶資源を要し得る新しいトランスコードセッションを開始することができる。

しかし、新しいトランスコードセッションに必要な演算資源は0%であり、これは、前のトランスコードセッション中にキャッシングされた最終トランスコード結果を利用する

50

ことができるためである。

【0024】

シーケンスレベル209メタデータは、ヘッダ情報、フレーム総数等を含むことができるが、これらに限定されない。

一般に、シーケンスレベル209メタデータは、ビデオストリームのセグメント全体の全体的な特徴を記述し、通常、限られた量の記憶空間を占める。

一実施形態によれば、システムが中間データをシーケンスレベル209にキャッシングする場合、このデータに対するそれ以降の要求に応答してトランスコーディングの完了に必要な演算資源は、「キャッシュヌル」201の場合と比較して少ない。

【0025】

トランスコードセッションが開始されると、CPU（中央演算処理装置）に対する負荷は、シーケンスレベル209メタデータ等、キャッシングされた中間メタデータを使用することによって低減することができる。

たとえば、キャッシングされたシーケンスレベル209メタデータを使用することにより、ビデオの初期部分をパースしてそこに提供される情報を得る必要なく、新しいトランスコードセッションを実行することができ、初期化プロセスの簡易化につながる。

さらに、使用されるキャッシング資源も低減される。

【0026】

図2に示す中間データレベルをピクチャレベルメタデータ212からピクセルレベルメタデータ219に通過するとき、トランスコードプロセスに関わるシステム記憶資源の量は、ポイントに位置するメタデータが記憶軸207に沿って達する縦方向の大きさが大きくなるにつれて増大し得る。

しかし、関わる記憶資源の量は、処理資源の利用が多くなるにつれて低減していく。

【0027】

図2に示す中間データレベルをピクセルレベルメタデータ219からピクチャレベルメタデータ212に通過するとき、トランスコードプロセスに必要なシステム演算資源の量は、ポイントに位置するメタデータ（たとえば、113）が演算軸205に沿って達する横方向の大きさが大きくなるにつれて増大し得る。

しかし、必要な演算資源の量は一般に、キャッシング資源の利用が多くなるにつれて低減していく。

【0028】

一実施形態において上に述べたように、本発明は、従来のキャッシングシステムによって提供される1つのレベルではなく、いくつかのキャッシングレベルを提供する。

本発明の実施形態は、提供される様々なキャッシングレベルを利用し、資源の配分を適応的に調整する。

たとえば、同じトランスコード処理に対して複数の要求がある状況に応答して、システムは、より多くのメタデータ（または演算要件が低いメタデータ）を格納するよう選択することにより、または極端な場合では、最終結果 - 出力ストリーム（たとえば、105）をキャッシングすることによって適応的に応答することができる。

こうして促進される細かい粒度により、処理プラットフォーム上で利用可能なすべての資源（たとえば、演算および記憶）をより効率的に利用することができる。

【0029】

図2に示すように、本発明の一実施形態によるストリーミングメディア（たとえば、ビデオ、テキスト、スピーチ等）トランスコードシステムをサポートするために、いくつかの種類の中間データ（たとえば、メタデータ）が、キャッシングの際に識別される。

これらの種類の中間データのキャッシングにより、以下述べる記憶と演算の良好なトレードオフの実現を容易にすることができる。

【0030】

[フレームレート低減のメタキャッシング]

図3は、本発明の一実施形態による、MPEGビデオストリームに対してフレームレ

10

20

30

40

50

ト低減トランスコーディングを実行するシステムに採用される機能ブロックを示すブロック図300である。

一実施形態によれば、フレームレートの低減は、フレームドロップ等の対策を採用することによって実現することができる。

たとえば、2/3のフレームレート低減は、GOP（ピクチャグループ）構造IBBPBBPBBを有するMP EGビデオストリーム中の「B」ピクチャを落とすことによって実現される。

フレームを「通す」かそれとも「落とす」かについての決定に関わるメタデータをキャッシングすることにより、図1のメタトランスコード103にかかる演算負荷は、このキャッシングデータにアクセスする後続のトランスコードセッションにおいて低減される。

10

図3は、入力ビデオストリーム301、パーサ303、フレームレートコントローラ307、出力ビデオストリーム305、およびメタキャッシュ309を示す。

さらに、図3は、本発明の図3の実施形態のメタトランスコード（たとえば、103）の構成要素（たとえば、303および307）の周囲に線を描いた破線311の図を示す。

#### 【0031】

本発明の実施形態は、図3に示すビデオストリームパーズ機能ブロック303に存在し得るMP EGビデオストリームトランスコードプロセスにおけるボトルネックを軽減する。

一実施形態によれば、パーズ機能の目的は、ピクチャタイプおよび各フレームに関連するバイト境界データを確認することである。

20

この情報はフレームレートコントローラ307に供給され、特定のフレームを「通す」かそれとも「落とす」かを判断する際に使用することができる。

先に触れたように、フレームを「通す」かそれとも「落とす」かの判断に関わるメタデータをキャッシングすることにより、このデータが使用される後続のトランスコードセッションにおいて、メタトランスコードに対する演算負荷が低減する。

#### 【0032】

一実施形態によれば、フレームレート低減トランスコーディングの場合、フレームレートfps（たとえば、フレーム/秒）を有する入力ストリームが与えられる場合、メタデータのデータキャッシングのデータレートしきい値は、以下のように計算される。

30

$$R = 40f \text{ ビット/秒}$$

#### 【0033】

上記式は、以下のように導き出される。

キャッシングビットレートは、ピクチャタイプ（たとえば、I、B、またはP）の記録に2ビット、およびピクチャ境界のバイトオフセットの記録に38ビット（組み合わせて5バイト、すなわち40ビット）を使用して計算することができる。

2ビットは4つのピクチャタイプを記録するに十分であり、38ビットは最大で512GB（たとえば、ギガバイト）サイズのメディアファイルにおけるオフセットを記録するに十分であることを理解されたい。

これは、現在利用可能なメディアファイルのタイプのすべてではなくともその大半をカバーするのに十分すぎると予期される。

40

#### 【0034】

本明細書における考察では、Rは「キャッシングレートしきい値」を指す。

一実施形態では、出力ストリームのターゲットビットレートがRよりも小さい場合、最終トランスコード結果（たとえば、「キャッシュフル」213）は、中間結果（たとえば、キャッシュメタ）とは対照的にキャッシングされる。

これは、メタデータのキャッシングに必要なデータレートがしきい値「R」を満たさないためである。

一実施形態によれば、資源バランサ（たとえば、107）が採用されて、最終トランスコード結果がキャッシングされるか、それとも中間トランスコード結果がキャッシングさ

50



れるかが決定される。

【0035】

以下の例に、フレームレート30fpsを有するビデオストリーム301に関してこれを示す。

上に提供した式（たとえば、 $R = 40 \text{ fビット/秒}$ ）を使用すると、キャッシングビットレートは $40 \times 30 = 1200 \text{ ビット/秒}$ （たとえば、1.2キロビット/秒）と計算することができる。

したがって、出力ストリーム305のターゲットビットレートがRよりも小さい場合、メタキャッシュ（たとえば、309）へのキャッシングではなく「キャッシュフル」を選択すべきである。

10

上で参照した式中の定数「40」は、MPEGビデオストリームが処理されている場合に特別に導出されたものであることを理解されたい。

他の場合（たとえば、テキストからスピーチ、スピーチからテキストへのトランスコーディング等）では、他の定数を導出することができる。

【0036】

1メガバイト/秒のビデオを約1Kbpsのレートでキャッシングすることができるため、必要な記憶資源は1000分の1に低減することができる。

さらに、ピクチャのタイプおよび関連するバイトオフセットはここではメタキャッシュ309において容易に利用することができるため、メタトランスコードは、パーサ303なしで、後続の同じ種類のトランスコード要求を容易に処理することができる。

20

したがって、メタトランスコードにかかる演算負荷が大幅に低減する。

【0037】

キャッシングされたピクチャタイプおよびバイト境界メタデータは、図2に示すグラフに表されるメタデータスキームのピクチャレベルで見つけられる。

したがって、シーケンスレベル（たとえば、209）にあるメタデータと比較して、このようなメタデータをキャッシングするにあたってより多くの記憶資源が使用され得るが、それ以降のトランスコード（たとえば、処理）セッションに必要な演算資源は大幅に低減することになる。

【0038】

要約すると、一実施形態では、フレームレートfftpを有する入力ストリーム（たとえば、301）が与えられる場合、ピクチャタイプおよびバイト境界メタデータのキャッシングに必要なデータレートは $R = 40 \text{ fビット/秒}$ である。

30

フレームレートの低減は通常、ビットレートの低減につながる。

出力ストリーム305のターゲットビットレートがRよりも小さい場合、メタキャッシュ（たとえば、309）へのキャッシングではなく「キャッシュフル」が選択される。

【0039】

〔ビットレート低減トランスコーディングのメタキャッシング〕

図4は、本発明の一実施形態による、MPEGビデオストリームに対してビットレート低減トランスコーディングを実行するシステムに採用される機能ブロックを示すブロック図400である。

40

一実施形態によれば、出力ストリーム（たとえば、409）のビットレートのみが、ビットレート低減トランスコーディングのために低減される。

一実施形態によれば、トランスコードビデオ信号のビデオフレームレートおよびビデオフレーム解像度等のパラメータが、ビットレート低減トランスコーディングのために一定レベルに維持される。

さらに、図4は、本発明の図4の実施形態のメタトランスコード（たとえば、103）の構成要素（例えば、403、405、407、および411）の周囲に線を描いた破線415の図を示す。

【0040】

ビットレート低減を実現する1つの方法は、変換領域係数の再量子化を通してのもので

50

ある。

考え得る従来のプロセスにはいくつかのボトルネックがあるが、再量子化係数 (M q u a n t) のキャッシングは、記憶資源の使用と演算資源の使用の間の良好なトレードオフを実現する。

再量子化係数は、最終トランスコード結果よりもキャッシングが容易な中間トランスコード結果である。

一実施形態によれば、この中間トランスコード結果のキャッシングにより、このようなデータを再生する必要なく、それ以降のトランスコードセッションを実行することが可能になる。

図4は、入力ビデオストリーム401、VLC(可変長コード)デコーダ403、逆量子化405、量子化407、ビットレートコントローラ411、メタキャッシュ413、および出力ストリーム409を示す。

#### 【0041】

一実施形態によれば、メタキャッシュ413は、再量子化係数 (M q u a n t) を格納するため、再量子化係数をビットレートコントローラ411によって生成する必要なく、メタキャッシュ413から得ることができる。

キャッシングされたM q u a n t データは、それ以降のトランスコードセッションをスピードアップする役割を果たすことができる。

これは、このようなそれ以降のトランスコードセッション中にこのデータを生成する必要であった処理を報告する必要がないためである。

先に触れたように、中間トランスコード結果は、メモリ消費と残りの処理量の間のトレードオフに応じてキャッシングされる。

#### 【0042】

一実施形態によれば、ビットレートトランスコーディングの場合、解像度  $w \times h$  およびフレームレート  $f f p s$  を有する入力ストリーム(たとえば、401)が与えられる場合、メタデータのデータキャッシングのデータレートしきい値は、以下のように計算される。

$$R = w h f / 32 \text{ b p s (ビット/秒)}$$

#### 【0043】

出力ストリーム409のターゲットビットレートがRよりも小さい場合、中間結果(たとえば、メタキャッシュにおけるキャッシュ)ではなく最終トランスコード結果(たとえば、キャッシュフル)をキャッシングすべきである。

これは、最終結果のキャッシングが同じくらい容易に対応することができる(たとえば、メタデータのキャッシングに必要なデータレートがしきい値「R」を満たさない)ためである。

一実施形態によれば、資源バランサ(たとえば、107)が採用されて、最終トランスコード結果をキャッシングするか、それとも中間トランスコード結果をキャッシングするかが決定される。

#### 【0044】

以下の例にこのポイントを示す。

一実施形態によれば、30秒の  $352 \times 288$ 、30  $f p s$  のビデオクリップの場合、M q u a n t データの格納に必要なキャッシュのサイズは356400バイトになり、M q u a n t メタデータのキャッシングに必要なデータレートは95040  $b p s$  になる(たとえば、毎秒95040ビットのM q u a n t データがキャッシングされる)。

したがって、ターゲットビットレートが95 K  $b p s$  以上の場合、M q u a n t メタデータは、このようなメタデータをキャッシングしたほうが有利なため(たとえば、メタデータのキャッシングを促進するのに必要なデータレートが存在し、データにより、それ以降の同様のトランスコードセッションのスピードアップが図られる)、キャッシングされることになる。

しかし、ターゲットビットレートが95 K  $b p s$  未満の場合、M q u a n t メタデータ

10

20

30

40

50

をキャッシングしないほうが有利である。

これは、最終トランスコード結果のキャッシングに同じくらい容易に対応することができる（たとえば、ターゲットデータレートが  $Mq u a n t$  メタデータのキャッシングに必要なデータレートよりも低い）ためである。

さらに、ランレングス符号化を使用して、後で  $Mq u a n t$  データを圧縮する、または記憶空間の半分のバイトを個々の  $Mq u a n t$  データに使用する（それにより、必要な空間が半分だけになる）等、キャッシュ空間の節約に他のキャッシュ記憶空間節約技法を採用してもよい。

キャッシングされるメタデータは、図 2 に示すグラフの（M P E G の場合、 $16 \times 16$  サイズの）マクロブロックレベル 215 にあることを理解されたい。

10

#### 【0045】

要約すると、一実施形態では、解像度  $w \times h$  およびフレームレート  $f f p s$  を有する出力ビデオストリーム 401 が与えられる場合、再量子化メタデータのキャッシングに必要なデータレートは、 $R = w h f / 32$  ビット/秒である。

出力ストリームのターゲットビットレートが  $R$  よりも小さい場合、「キャッシュメタ」413 ではなく「キャッシュフル」（たとえば、213）が選択される。

#### 【0046】

図 4 に示す他の構造（たとえば、V L C デコーダ 403、逆量子化 405、量子化 407 機能ブロック）は、ビデオ処理分野において既知の構成要素であり、本明細書においてこれ以上説明しない。

20

#### 【0047】

解像度低減トランスコーディングのメタキャッシング

図 5 は、本発明の一実施形態による、M P E G ビデオストリームに対して空間または解像度低減トランスコーディングを実行するシステムに採用される機能ブロックを示すブロック図 500 である。

空間解像度の低減は、各フレームの解像度を低減し、その結果をより低いビットレートでコーディングすることによって実現することが可能である。

従来のプロセスにはいくつかのボトルネックが存在し得るが、動き情報および再量子化係数のキャッシングにより、記憶資源と演算資源の間の良好なトレードオフが実現される。

30

図 5 は、入力ストリーム 501、動きベクトル発生器 507、レート制御モジュール 521、出力ストリーム 523、動きベクトル（529 および 531）およびメタキャッシュ 527 を示す。

さらに、図 5 は、本発明の図 5 の実施形態のメタトランスコーダ（たとえば、103）の構成要素（たとえば、図 5 において破線 537 によって囲まれている構成要素を参照）の周囲に線を描いた破線 537 の図を示す。

#### 【0048】

空間解像度低減トランスコードプロセスでは、多くの演算集約的なブロックがあり得ることを理解されたい。

しかし、一実施形態によれば、2つの機能ブロック、すなわち、動きベクトル発生器 507 およびレート制御モジュール 521 が、メタキャッシングを適用して良好な記憶と演算とのトレードオフを実現することができる、トランスコードプロセスにおけるポイントとして識別されている。

40

#### 【0049】

ビデオ符号化プロセスでは、動きベクトル（529 および 531）の生成が、演算集約的な作業である。

一実施形態によれば、動きベクトル発生器 507 のトランスコード結果をキャッシングすることにより（出力ストリーム 523 ではなく中間結果）、後続の同様の処理（たとえば、トランスコードセッション）に対する要求に、より少ない演算で応えることができる。

50

## 【 0 0 5 0 】

一実施形態によれば、空間解像度低減トランスコーディングの場合、解像度  $w \times h$  およびフレームレート  $f$  [fps]、GOP サイズ  $n$  を有する入力ストリームが与えられる場合、メタデータのデータキャッシングのデータレートしきい値は、以下のように計算される。

$$R = w h f (n - 1) / 8 n \text{ (ビット/秒)}$$

## 【 0 0 5 1 】

出力ストリーム 5 2 3 のターゲットビットレートが  $R$  よりも小さい場合、中間結果（たとえば、キャッシュメタ 5 2 7）ではなく最終トランスコード結果（たとえば、「キャッシュフル」 2 1 3）をキャッシングすべきである。

これは、最終結果のキャッシングが同じくらい容易に対応することができる（たとえば、メタデータのキャッシングに必要なデータレートが存在しない）ためである。

10

一実施形態によれば、資源バランサ（たとえば、1 0 7）が採用されて、最終トランスコード結果をキャッシングするか、それとも中間トランスコード結果をキャッシングするかが決定される。

## 【 0 0 5 2 】

以下の例にこのポイントを示す。

MPEG 標準によれば、ピクチャ間にあるマクロブロックには、最大で 4 つの動きベクトル（たとえば、5 2 9、5 3 1）が関連付けられる。

各動きベクトルは 1 バイトで表すことが可能である。

出力ストリーム 5 2 3 が、ビットレート 1 Mbps およびフレームレート 30 fps を有するサイズ  $352 \times 240$  の典型的なウェブビデオであると仮定すると、メタデータのキャッシングに必要なデータレートは  $2 * 14 * 330 * 4 * 8 = 288 \text{ Kbps}$  であり、これは最終結果のキャッシングに必要なレートの 3 分の 1 に近い。

20

## 【 0 0 5 3 】

したがって、288 Kbps 以上のターゲットビットレートの場合、動きベクトルメタデータはキャッシングされる。

これは、このようなメタデータのキャッシングがそれ以降の同様のセッションに有利であるためである（たとえば、メタデータのキャッシングの促進に必要なデータレートが存在し、そのデータにより、それ以降の同様のトランスコードセッションのスピードアップが図られる）。

30

しかし、288 Kbps 以下のターゲットビットレートの場合、動きベクトルメタデータのキャッシングは有利ではない可能性がある。

これは、このような場合、最終トランスコード結果のキャッシングに同じくらい容易に対応することができるためである（たとえば、ターゲットデータレートが、動きベクトルメタデータのキャッシングに必要なデータレートよりも低い）。

## 【 0 0 5 4 】

キャッシングされた動きベクトルメタデータは、図 2 に示す（MPEG の場合、 $16 \times 16$  の）マクロブロックレベル 2 1 5 にあることを理解されたい。

したがって、キャッシングし得る他のレベルのメタデータ（たとえば、ピクチャ、シーケンス）と比較してより多くの記憶資源を使用し得るが、それ以降の要求の処理に必要な演算資源は少なくなる。

40

他の事項として、空間解像度低減トランスコーディングのパフォーマンスは通常、ビットレートの低減につながることに留意されたい。

## 【 0 0 5 5 】

一実施形態によれば、レート制御モジュール 5 2 1 からの中間トランスコード結果は、空間解像度低減トランスコーディングの一貫として、図 4 を参照して上述したビットレート低減トランスコーディングの文脈の中で行われた方法と同様の方法でキャッシングすることができる。

これら結果のキャッシングにより、すでに生成されキャッシングされている量子化係数を再生成する必要をなくすることができる。

50

図4を参照して考察したものと同一解析がこの文脈にも当てはまることを理解されたい。

【0056】

要約すると、一実施形態では、解像度 $w \times h$ 、フレームレート $f \text{ fps}$ 、およびGOPサイズ $n$ を有する出力ビデオストリーム523が与えられる場合、動きベクトルメタデータ（たとえば、529）のキャッシングに必要なデータレートは、 $R = whf(n - 1) / 8n$ ビット/秒である。

出力ストリーム523のターゲットビットレートが $R$ よりも小さい場合、「キャッシュメタ」527ではなく「キャッシュフル」（たとえば、213）が選択される。

【0057】

図5に示す他の構造（たとえば、VLCデコーダ503、逆量子化505、量子化517、VLCエンコーダ519、順方向量子化器（forward quantizer）517、FDCT（高速離散コサイン変換）515、IDCT（逆離散コサイン変換）507、動き補償509および513、ダウンサンプリング511、ならびにフレームバッファ533および535機能ブロック）は、ビデオ処理技術分野において既知の構成要素であり、本明細書においてこれ以上説明しない。

【0058】

[本発明の実施形態による動作例]

図6は、本発明の一実施形態によるストリーミングメディアを処理するコンピュータ制御方法を表すフローチャート600を示す。

フローチャート600は、一実施形態では、コンピュータ可読かつコンピュータ実行可能な命令の制御下でプロセッサおよび電子構成要素によって実行されるプロセスを含む。

コンピュータ可読かつコンピュータ実行可能な命令は、たとえば、コンピュータ使用可能な揮発性メモリおよび/またはコンピュータ使用可能な不揮発性メモリ等、データ記憶機構に常駐する。

しかし、コンピュータ可読かつコンピュータ実行可能な命令は、あらゆるタイプのコンピュータ可読媒体に常駐してよい。

特定のステップがフローチャートに開示されるが、このようなステップは例示である。

すなわち、本発明は、他の様々なステップまたはフローチャートに記されたステップの変形形態の実行に適している。

本実施形態内では、フローチャートのステップはソフトウェア、ハードウェア、またはこれら両方の組み合わせによって実行し得ることを理解されたい。

【0059】

ステップ601において、ストリーミングデータ入力を受け取る。

入力データはリモートサーバ、ローカル記憶媒体、リモート記憶媒体等から供給し得ることを理解されたい。

【0060】

ステップ603において、利用可能な処理資源およびメモリ資源の配分が決定される。

一実施形態によれば、資源の配分は、利用可能なキャッシュ空間および利用可能な演算サイクルに基づく。

キャッシュ空間は中間トランスコード結果（たとえば、メタデータ）に、または最終トランスコード結果に配分可能なことを理解されたい。

【0061】

ステップ605において、ステップ601において受け取ったデータに対して、ステップ603において決定された資源配分に基づいて、多段階処理が実行される。

一実施形態によれば、実行される処理は、ビデオトランスコーディング、テキストからスピーチへのトランスコーディング、およびスピーチからテキストへのトランスコーディングを含むことができるがこれらに限定されない。

【0062】

ステップ607において、上記多段階プロセスの1つの段階からの中間結果がキャッシ

10

20

30

40

50

ングされる。

中間結果は、利用可能な処理資源およびメモリ資源に従って選択される。

一実施形態によれば、中間結果のキャッシングは、ストリーミングデータ出力のターゲットビットレートがレートしきい値を満たしていることに依存し得ることに留意されたい。

一実施形態によれば、このしきい値はキャッシングビットレートしきい値「R」である。

このしきい値が満たされると、資源配分を、一実施形態によれば、資源バランスによって決定することができる。

#### 【0063】

10

本発明の例示的な実施形態を参照して上述されるように、本発明は、ストリーミングメディアを処理する方法およびシステムを提供する。

本方法は、ストリーミングメディアを受け取ること、利用可能な処理資源およびメモリ資源の配分を決定することを含む。

さらに、本方法は、ストリーミングメディアに対して多段階処理を実行すること、多段階プロセスの中の1つの段階からの中間結果をキャッシングすることを含む。

中間結果は利用可能な処理資源およびメモリ資源に従って選択される。

#### 【0064】

本発明の特定の実施形態についての上記説明は、図示および説明を目的として提示されたものである。

20

網羅的である、すなわち本発明を開示した厳密な形態に限定する意図はなく、明らかに、上記教示を鑑みて多くの変更形態および変形形態が可能である。

実施形態は、本発明の原理およびその実際の応用を最良に説明するように選択され説明され、それによって、当業者が、考えられる特定の用途に適した各種変更によって本発明および各種実施形態を最良に利用することができる。

本発明の範囲は添付の特許請求の範囲およびその等価物によって規定されることを意図する。

#### 【図面の簡単な説明】

#### 【0065】

【図1A】本発明の一実施形態によるトランスコードシステムの機能ブロック図である。

30

【図1B】本発明の一実施形態による例示的なトランスコード構造のブロック図である。

【図2】本発明の一実施形態によるシーケンスレベルからピクセルレベルにキャッシングすることができるメタデータの5つの中間レベル間の演算と記憶との間のトレードオフを示すグラフである。

【図3】本発明の一実施形態によるMPEG（モーションピクチャエキスパートグループ）ビデオストリームに対してフレームレート低減トランスコーディングを実行するシステムに採用される機能ブロックを示すブロック図である。

【図4】本発明の一実施形態による、MPEGビデオストリームに対してビットレート低減トランスコーディングを実行するシステムに採用される機能ブロックを示すブロック図である。

40

【図5】本発明の一実施形態による、MPEGビデオストリームに対して空間解像度低減トランスコーディングを実行するシステムに採用される機能ブロックを示すブロック図である。

【図6】本発明の一実施形態によるストリーミングメディアを処理するコンピュータ制御方法を表すフローチャートを示す。

#### 【符号の説明】

#### 【0066】

101・・・入力ストリーム、

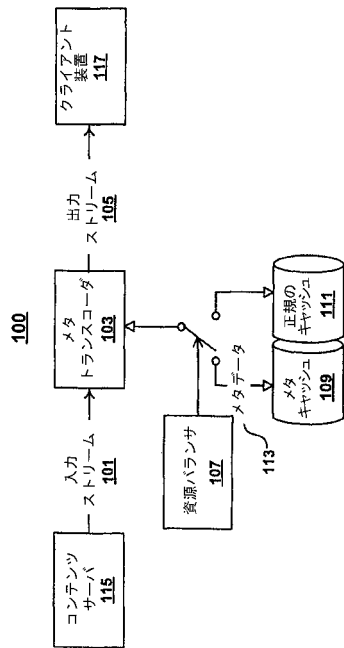
103・・・メタトランスコード、

105・・・出力ストリーム、

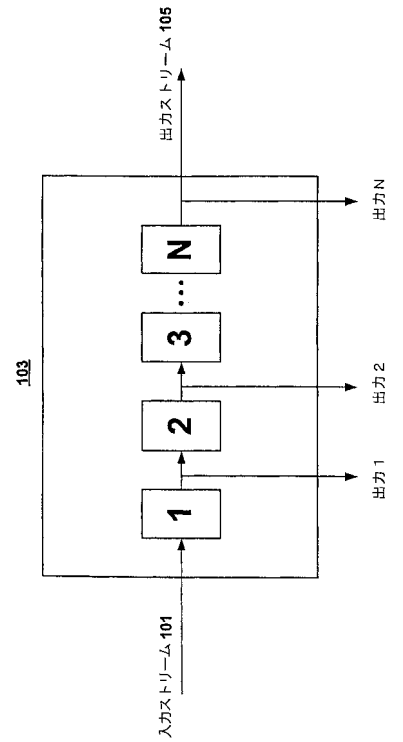
50

1 0 7 . . . 資源バランサ、	
1 0 9 . . . メタキャッシュ、	
1 1 1 . . . 正規のキャッシュ、	
1 1 3 . . . メタデータ、	
1 1 5 . . . コンテンツサーバ、	
1 1 7 . . . クライアント装置、	
3 0 1 . . . 入力ビデオストリーム、	
3 0 3 . . . パーサ、	
3 0 5 . . . 出力ビデオストリーム、	
3 0 7 . . . フレームレートコントローラ、	10
3 0 9 . . . メタキャッシュ、	
4 0 1 . . . 入力ストリーム、	
4 0 3 . . . V L C デコーダ、	
4 0 5 . . . 逆量子化、	
4 0 7 . . . 量子化、	
4 0 9 . . . 出力ストリーム、	
4 1 1 . . . ビットレートコントローラ、	
4 1 3 . . . メタキャッシュ、	
5 0 1 . . . 入力ストリーム、	
5 0 3 . . . V L C デコーダ、	20
5 0 5 . . . 逆量子化器、	
5 0 7 . . . M V 発生器、	
5 0 9 , 5 1 3 . . . 動き補償、	
5 1 1 . . . ダウンサンプリング、	
5 1 7 . . . 順方向量子化器、	
5 1 9 . . . V L C エンコーダ、	
5 2 1 . . . レート制御、	
5 2 3 . . . 出力ストリーム、	
5 2 7 . . . メタキャッシュ、	
5 2 9 . . . 動きベクトル、	30
5 3 1 . . . 新しい動きベクトル、	
5 3 3 , 5 3 5 フレームバッファ、	

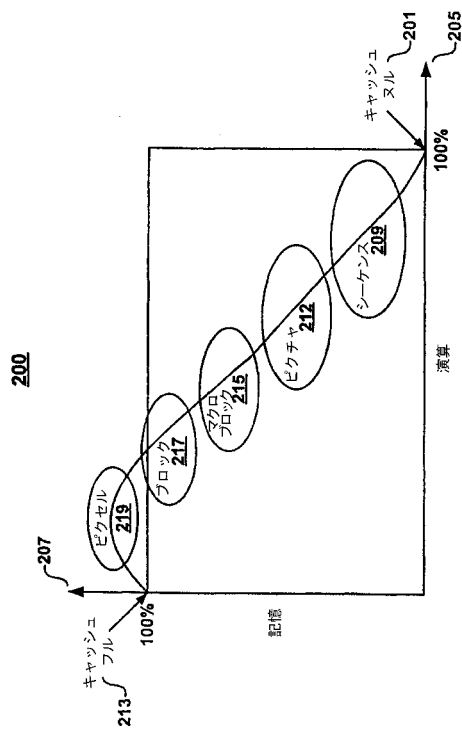
【図 1 A】



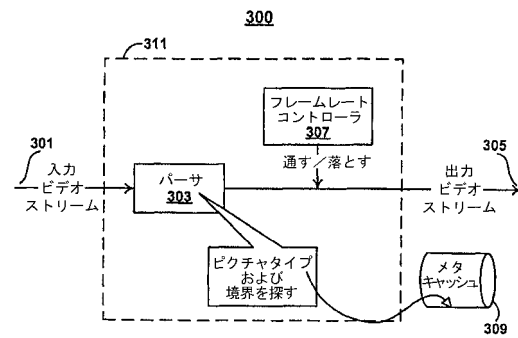
【図 1 B】



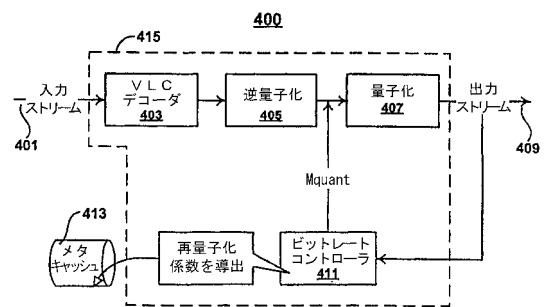
【図 2】



【図 3】

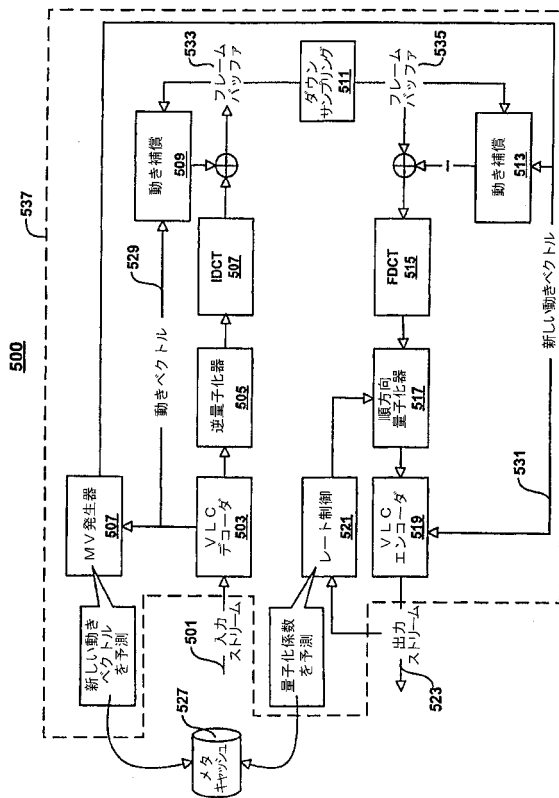


【図 4】

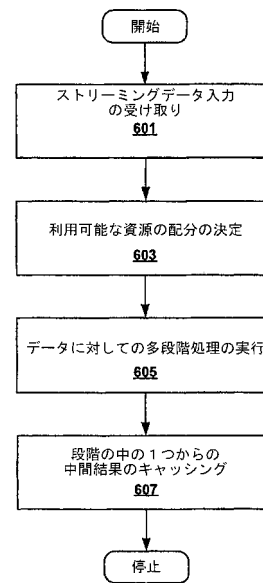




【図 5】



【図 6】



---

フロントページの続き

- (56)参考文献 特開2003-333577(JP,A)  
特開2002-325094(JP,A)  
特開2000-165436(JP,A)  
国際公開第98/043177(WO,A1)  
特表2006-506738(JP,A)  
Sumit Roy et al., A System Architecture for Managing Mobile Streaming Media Services, Proceedings of the 23rd International Conference on Distributed Computer Systems Workshops(ICDCSW'03), IEEE, 2003年 5月, pp.408-413  
Yuko Onoe et al., Network information based Rate Controls on Multimedia Streaming Servers, roceedings of the 23rd International Conference on Distributed Computer Systems Workshops(ICDCSW'03), IEEE, 2003年 5月, pp.543-548  
Cheng-Yue Chang et al., Exploring Aggregate Effect with Weighted Trancoding Graphs for Efficient Cache Replacement in Transcoding Proxies, roceedings of the 18th International Conference on Data Engineering(ICDE'02), IEEE, 2002年, pp.383-392

(58)調査した分野(Int.Cl., DB名)

H04N 7/24 - 7/68  
H04N 7/173