(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2004-320667
(P2004-320667A)
(43)公開日　平成16年11月11日(2004.11.11)

(51)Int.Cl.[7]	F I		テーマコード（参考）
H04N 7/24	H04N 7/13	Z	5C053
H04N 5/91	H04N 7/173	610Z	5C059
H04N 7/173	H04N 5/91	N	5C064

審査請求 有　請求項の数 9　ＯＬ　（全 12 頁）

(21)出願番号　特願2003-115165(P2003-115165)
(22)出願日　平成15年4月21日(2003.4.21)

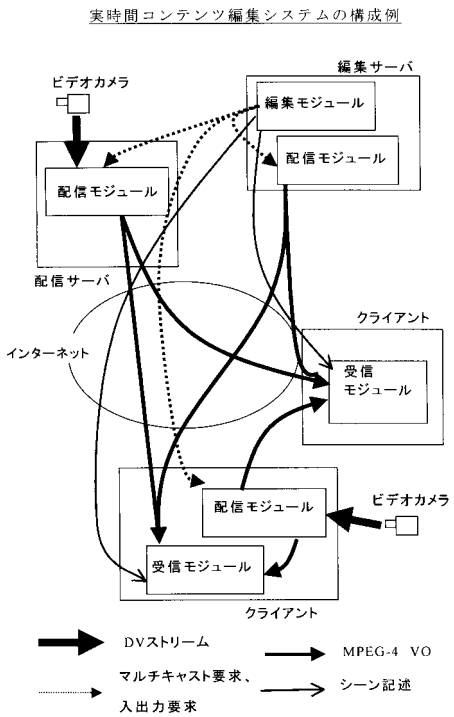
(71)出願人　301022471
独立行政法人情報通信研究機構
東京都小金井市貫井北町４－２－１
(74)代理人　100108660
弁理士　大川　譲
(74)代理人　100097836
弁理士　福井　國敞
(72)発明者　三浦　康之
東京都小金井市貫井北町４－２－１　独立行政法人通信総合研究所内
(72)発明者　勝本　道哲
東京都小金井市貫井北町４－２－１　独立行政法人通信総合研究所内

Ｆターム(参考)　5C053 FA14 GA11 GB22 GB26 GB28 GB37 KA24 LA01 LA06 LA15

最終頁に続く

(54)【発明の名称】実時間コンテンツ編集方法及びシステム並びにプログラム

(57)【要約】
【課題】ＣＰＵパワーに応じた符号化処理を行うことで、リアルタイムなＭＰＥＧ－４　ＶＯへの符号化を実現する。
【解決手段】本発明は、インターネット上に散在するライブ映像を含む多数の画像及び／又は音声を素材にした編集処理を行い、複数のユーザに対して配信する。本発明は、入力装置としての複数のビデオカメラと、複数の符号化アルゴリズムの内の１つを選択して符号化可能の符号化規格を用いて、前記ビデオカメラによる入力画像を符号化して配信する複数の配信モジュール、画像を受信して表示する受信モジュール、および少なくとも一つの編集モジュールとを備えている。配信モジュールは使用するマシンの性能に応じて、使用するビデオオブジェクトプレーンＶＯＰの種類および使用頻度を変更することにより、圧縮効率の良い符号化アルゴリズムを選択する。
【選択図】　図１



【特許請求の範囲】
【請求項１】
インターネット上に散在するライブ映像を含む多数の画像及び／又は音声を素材にした編集処理を行い、複数のユーザに対して配信する実時間コンテンツ編集方法において、
入力装置としての複数のビデオカメラと、複数の符号化アルゴリズムの内の１つを選択して符号化可能の符号化規格を用いて、前記ビデオカメラによる入力画像を符号化して配信する複数の配信モジュール、画像を受信して表示する受信モジュール、および少なくとも一つの編集モジュールとを備えて、
配信モジュールは使用するマシンの性能に応じて、使用するビデオオブジェクトプレーンＶＯＰの種類および使用頻度を変更することにより、圧縮効率の良い符号化アルゴリズムを選択する、
ことから成る実時間コンテンツ編集方法。
【請求項２】
前記ビデオオブジェクトプレーンＶＯＰの種類および使用頻度の変更は、入力画像の符号化処理を基本処理と付加的処理に分類し、分類した各種処理の符号化実行時間を計測し、その計測結果に基づいて行う請求項１に記載の実時間コンテンツ編集方法。
【請求項３】
インターネット上に散在するライブ映像を含む多数の画像及び／又は音声を素材にした編集処理を行い、複数のユーザに対して配信する実時間コンテンツ編集システムにおいて、
入力装置としての複数のビデオカメラと、複数の符号化アルゴリズムの内の１つを選択して符号化可能の符号化規格を用いて、前記ビデオカメラによる入力画像を符号化して配信する複数の配信モジュール、画像を受信して表示する受信モジュール、および少なくとも一つの編集モジュールから構成されて、
配信モジュールは使用するマシンの性能に応じて、使用するビデオオブジェクトプレーンＶＯＰの種類および使用頻度を変更することにより、圧縮効率の良い符号化アルゴリズムを選択する、
ことから成る実時間コンテンツ編集システム。
【請求項４】
前記マシンの性能は、プロセスの実行時間の監視を行うことによって行い、監視の結果に基づいて複数の符号化アルゴリズムから適切なものを選択する請求項３に記載の実時間コンテンツ編集システム。
【請求項５】
前記プロセスの実行時間の監視は、システム内でビデオオブジェクトプレーンＶＯＰの符号化時間を測定し、あらかじめ定められた平均フレームレートにもとづいてフレーム間圧縮を行うか否か選択する請求項４に記載の実時間コンテンツ編集システム。
【請求項６】
前記符号化規格は、ＭＰＥＧ－４規格である請求項３に記載の実時間コンテンツ編集システム。
【請求項７】
前記編集モジュールは、配信サーバに対する画像及び／又は音声のマルチキャスト要求を行い、クライアントへ送信するシーン記述言語の生成およびマルチキャストを担当する請求項３に記載の実時間コンテンツ編集システム。
【請求項８】
前記選択された符号化アルゴリズムによる符号化処理を段階的に行い、一定時間経過後に必要最小限の符号化を完了し、その後に解像度や圧縮率を高めた付加的な符号化処理を行い、制限時間がすぎた時点で付加的な符号化処理が完了しなかった場合、その時点で処理を打ち切り、直前の段階における符号化処理の結果を配信する請求項３に記載の実時間コンテンツ編集システム。
【請求項９】
インターネット上に散在するライブ映像を含む多数の画像及び／又は音声を素材にした編

集処理を行い、複数のユーザに対して配信する実時間コンテンツ編集プログラムにおいて、
入力装置としての複数のビデオカメラと、複数の符号化アルゴリズムの内の１つを選択して符号化可能の符号化規格を用いて、前記ビデオカメラによる入力画像を符号化して配信する複数の配信モジュール、画像を受信して表示する受信モジュール、および少なくとも一つの編集モジュールとを備えて、
配信モジュールは使用するマシンの性能に応じて、使用するビデオオブジェクトプレーンＶＯＰの種類および使用頻度を変更することにより、圧縮効率の良い符号化アルゴリズムを選択する、
各手順を実行することから成る実時間コンテンツ編集プログラム。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

本発明は、インターネット上に散在するライブ映像を含む多数の画像及び／又は音声を素材にした編集処理を行い、複数のユーザに対して配信する実時間コンテンツ編集方法及びシステム並びにプログラムに関する。

【０００２】

【従来の技術】

近年の広帯域なネットワークの普及により、大容量の動画像コンテンツの配信が可能となっている。２００１年以降、ＡＤＳＬ（Ａｓｙｍｍｅｔｒｉｃ Ｄｉｇｉｔａｌ Ｓｕｂｓｃｒｉｂｅｒ Ｌｉｎｅ）の加入者が大幅に増加したことにより、ブロードバンド環境が整いつつある。現在（２００２年３月現在）、日本では、７．３％の世帯がＡＤＳＬまたはＣＡＴＶに加入し、１～１０Ｍｂｐｓ程度のデータ受信および数百キロｂｐｓ程度の送信が可能となっている。また、今後ＦＴＴＨ（Ｆｉｂｅｒ Ｔｏ Ｔｈｅ Ｈｏｍｅ）の普及が見込まれており、これにより最大で１００Ｍｂｐｓ程度の通信が可能となり、ユーザ同士で大容量の動画像を送受信することが可能になると予想される。

【０００３】

それにともない、広帯域ネットワークを通じて画像・音声を配信するサービスやツールが数多く提案されている。広帯域を使用した高画質映像配信システムとして、Ｒｕｆｆ Ｓｙｓｔｅｍｓ（非特許文献１参照）、ＤＶＴＳ（非特許文献２参照）などのシステムが提案されている。Ｒｕｆｆ Ｓｙｓｔｅｍｓは、非圧縮のＤ１／ＨＤＴＶ映像やビデオカメラから取得したＤＶストリームを、ＴＣＰ／ＩＰ上で配信するシステムである。非圧縮画像やＤＶ画像を使用しているため大きな帯域を必要とするが、高画質の配信が可能という特徴を有する。狭帯域に対応したＭＰＥＧ動画やＨ．３２３規格を利用しシステムやツールは多数存在する（非特許文献３～６参照）。これらはいずれもＰＣ上からカメラ画像をインターネット経由で受信することが可能であるが、単一の動画を対象としたアプリケーションであり、多数の動画を扱うには不向きである。

【０００４】

【非特許文献１】

ｈｔｔｐ：／／ｗｗｗ．ｃｎｄ．ｔｅｌ．ｃｏ．ｊｐ／ｐｒｏｄｕｃｔ／ｓｙｏ６４．ｈｔｍｌ

【非特許文献２】

ｈｔｔｐ：／／ ｗｗｗ．ｓｆｃ．ｗｉｄｅ．ａｄ．ｊｐ／ＤＶＴＳ

【非特許文献３】

ｏｈｐｈｏｎｅ： ｈｔｔｐ：／／ｗｗｗ．ｏｐｅｎｈ３２３．ｏｒｇ／ｄｏｃｓ／ｏｈｐｈｏｎｅ_ｍａｎ．ｈｔｍｌ

【非特許文献４】

ＦＦｍｐｅｇ： ｈｔｔｐ：／／ｆｆｍｐｅｇ．ｓｏｕｒｃｅｆｏｒｇｅ．ｎｅｔ／

【非特許文献５】

Ｃｏｒｉａｎｄｅｒ： ｈｔｔｐ：／／ｗｗｗ．ｔｅｌｅ．ｕｃｌ．ａｃ．ｂｅ／ＰＥＯ

ＰＬＥ／ＤＯＵＸＣＨＡＭＰＳ／ｉｅｅｅ１３９４／ｃｏｒｉａｎｄｅｒ／
【非特許文献６】
三浦康之，勝本道哲「実時間環境におけるビジュアル符号化のための実証実験」，情報処理学会研究報告，　２００２－ＤＰＳ－１００，　ｐｐ．２５－３０，　２００２
【０００５】
【発明が解決しようとする課題】
ビデオカメラから取得した映像をその時点で配信するライブ配信を可能とするためには、符号化処理を実時間処理で行うことが求められる。実時間処理とは、外部からの連続的な入力に対し、定められた時間内に出力を返すことが要求される処理である。一秒間にｆフレームの動画を配信する実時間処理システムでは、１フレームに対する編集処理およびビットストリームのパケット化を平均１／ｆ秒以内に完了し、その場で配信しなければならない。
【０００６】
上記のような目標を達成するためには、以下のアプローチが有効である。
・符号化処理を高速化し、制限時間以内に符号化処理を完了する。
・性能の異なるマシンの上で正しく動作するために、プロセスの実行時間の監視を行い、監視の結果に基づいて複数の符号化アルゴリズムから適切なものを選択する。
・符号化処理を段階的に行う。一定時間経過後に必要最小限の符号化を完了し、その後に解像度や圧縮率を高めた付加的な符号化処理を行う。制限時間がすぎた時点で付加的な符号化処理が完了しなかった場合、その時点で処理を打ち切り、直前の段階における符号化処理の結果を配信する。
【０００７】
上記の考え方に基づいて、本発明の実時間コンテンツ編集システムでは、システム内でＶＯＰ（ビデオオブジェクトプレーン：Ｖｉｄｅｏ　Ｏｂｊｅｃｔ　Ｐｌａｎｅ、フレームに相当）の符号化時間を測定し、あらかじめ定められた平均フレームレートにもとづいてフレーム間圧縮を行うか否か選択する。
【０００８】
本発明者は個々のストリームを対象とした配信システムに対して、複数の画像や音声を入力とし、リアルタイムで編集・コード化して配信する作業を行う実時間コンテンツ編集システムを提案する。本システムは、ＭＰＥＧ－４のビデオオブジェクト（ＶＯ）および編集用言語の符号化および復号を分散環境において処理するもので、従来の配信システムに対して以下のような特長を有する。
▲１▼　離れた複数地点からのライブ映像を用いたコンテンツ編集作業が可能
▲２▼　各地点からの映像配信者や映像受信者は、システムの稼動中に自由に参加・離脱が可能
個々のユーザのニーズに沿ったコンテンツを提供するためには、ユーザ側に多数のビデオオブジェクト（ＶＯ）を送りコンテンツを構築するため、ユーザ側に多数のビデオオブジェクト（ＶＯ）が集中する。そのため個々のビデオオブジェクト（ＶＯ）に多くの帯域を割くことが困難となる。また、送信者がＡＤＳＬのような非対称な通信サービスを使用する場合、上り側の帯域幅は下り側のそれに比べて小さなものにとどまるため、やはり　ビデオオブジェクト（ＶＯ）に多くの帯域を割くことが困難となる。Ｒｕｆｆ　Ｓｙｓｔｅｍｓのように、Ｄ１非圧縮データやＤＶ圧縮データ等の大きなデータをそのまま伝送する方式を採用した場合、必要帯域が大きいために多数のビデオオブジェクト（ＶＯ）を同時に送受信するシステムでは莫大な量の帯域が必要となる。そのため、高い圧縮率を持ち、さまざまな帯域に対応したＭＰＥＧ－４ビデオオブジェクト（ＭＰＥＧ－４　ＶＯ）等への符号化が必要になるが、ＭＰＥＧ－４　ＶＯ符号化にはフレーム間圧縮に多大なＣＰＵパワーを必要とするため、時間的制約の厳しい実時間環境下でスペックの低いマシンを用いてライブ映像の符号化を実行する場合、複雑な符号化は困難となる。個人ユーザが自宅から、あるいは持ち運び可能な程度の機材を用いて出張先等から、ライブ映像を送信する

等の場合、家庭用ＰＣ程度の機材しか使用できないことが多い。他方、比較的高性能なサーバマシンを使用した場合、スペックの低いマシンを使用する場合に比べ、複雑な符号化が可能になる。このように、使用するサーバの性能により可能な処理の範囲が変わってくる。

【０００９】

そこで、本発明は、係る問題点を解決して、ＣＰＵパワーに応じた符号化処理を行うことで、リアルタイムなＭＰＥＧ－４　ＶＯへの符号化を実現することを目的としている。

【００１０】

【課題を解決するための手段】

本発明は、複数の画像・音声を入力とし、リアルタイムでＭＰＥＧ－４ビデオオブジェクト（ＶＯ）への符号化・編集・配信を一体的に行う実時間コンテンツ編集システムを提供する。本システムは、離れた複数地点からのデジタルビデオによるライブ映像を用いたコンテンツを提供することが可能である。

【００１１】

本システムは、編集・配信・受信の３種類のモジュールにより構成され、配信・受信の各モジュールはシステムの稼動中に自由に参加・離脱が可能である。また、配信モジュールは使用するマシンの性能により使用するＶＯＰの種類および頻度を変更することにより、圧縮効率の良い符号化法を自動で選択することが可能である。

【００１２】

本発明の実時間コンテンツ編集方法及びシステム並びにプログラムは、インターネット上に散在するライブ映像を含む多数の画像及び／又は音声を素材にした編集処理を行い、複数のユーザに対して配信する。本発明は、入力装置としての複数のビデオカメラと、複数の符号化アルゴリズムの内の１つを選択して符号化可能の符号化規格を用いて、前記ビデオカメラによる入力画像を符号化して配信する複数の配信モジュール、画像を受信して表示する受信モジュール、および少なくとも一つの編集モジュールとを備えている。配信モジュールは使用するマシンの性能に応じて、使用するビデオオブジェクトプレーンＶＯＰの種類および使用頻度を変更することにより、圧縮効率の良い符号化アルゴリズムを選択することを特徴としている。

【００１３】

【発明の実施の形態】

（実時間コンテンツ編集システム構成）

実時間コンテンツ編集システムは、インターネット上に散在する、ライブ映像を含む多数の画像・音声を素材にした編集処理を行い、複数のユーザに対して配信するシステムである。本システムは、既存のさまざまな環境に柔軟に対応するため、各種回線・マシンによらず動作することを目指している。そのため、画像の符号化として、高い圧縮率を持ちさまざまな帯域に対応したＭＰＥＧ－４規格を用いる。さらに、複数の画像・音声の編集のために、シーン記述言語を使用する。

【００１４】

図１に、構築するシステムの概念図を示す。本システムは、入力装置として複数のビデオカメラと、それらによる入力画像を符号化して配信する複数の配信モジュール、画像を受信して表示する受信モジュール、および少なくとも一つの編集モジュールから構成される。また、例示のシステムは、１つの配信サーバと、１つの編集サーバと、２つのクライアントがインターネットを介して接続されるものとして示している。編集モジュールは、編集サーバに配置され、かつ、配信モジュールは、配信サーバに配置するだけでなく、編集サーバにも配置されるものとして例示している。受信モジュールは、クライアントのそれぞれに配置され、かつ、一方のクライアントには、受信モジュールに加えて、配信モジュールも配置されるものとして例示している。図中、ＤＶ（Ｄｉｇｉｔａｌ　Ｖｉｄｅｏ）ストリーム、ＭＰＥＧ－４　ＶＯ、マルチキャスト要求、入出力要求、シーン記述は、それぞれ線の種類を異にする矢印によって示している。

【００１５】

編集モジュールはオペレータからの指示に従い配信サーバに対する画像・音声のマルチキャスト要求を実際に行い、クライアントへ送信するシーン記述言語の生成およびマルチキャストを担当する。配信モジュールは、入力装置からＤＶ（Ｄｉｇｉｔａｌ　Ｖｉｄｅｏ）形式で取り込まれた画像・音声のＭＰＥＧ－４符号化を行い、複数のクライアントに対しインターネットを介してパーソナライズメディアストリーム配信によるマルチキャストを行う。受信モジュールを持つクライアントでは、複数のＭＰＥＧ－４画像音声データおよびシーン記述に基づいて動画像を表示する。ビデオカメラと、配信モジュールを含む装置は、ＩＥＥＥ１３９４ポートを介して接続される。ＩＥＥＥ１３９４ポートは、多くのＤＶ機器に備えられており、高速な上に一定周期でデータを送受信するアイソクロナス転送をサポートしているため、映像音声データの送受信に適している。

【００１６】

配信モジュール、受信モジュール、編集モジュールとしては任意の形態が考えられる。例えば、汎用のサーバマシン上のプログラムとして配置することもできるし、家庭用ＰＣ上で動作させても構わない。また、ＰＣに対する組み込みシステムとして、専用のＰＣカードの形態を取ることも考えられる。

【００１７】

（配信モジュールにおけるＤＶ－ＭＰＥＧ変換器）

１）　構成

実時間処理の可能なコンテンツ編集システムでは、送信者側からデジタルビデオカメラを使用して手持ちのＰＣを配信サーバとして画像を配信することも、職場の会議室などから高スペックのサーバマシンを使用して配信することも可能である。前者の場合、後者に比べて低スペックなＰＣを使用することが多く、複雑な符号化処理を行うのは困難である。後者の場合、より複雑な符号化処理が可能であるが、サーバマシンのスペックにより利用可能な符号化の種類が異なる。また、送信者が使用する回線の帯域による制限も存在する。送信側でＡＤＳＬのような非対称な回線を使用していた場合、上り回線の帯域が数百キロｂｐｓと、下り回線に比べて大きく制限される。以上のような理由から、配信側の実情に合った符号化を行う必要がある。

【００１８】

多くの場合、動画質の維持の観点から一定以上のフレームレートを保つことが求められる。そのため本システムではＲ（ｆｌａｍｅｓ／ｓ）（Ｒ：所定数）以上のフレームレートを維持するという拘束条件が課せられるものとする。

【００１９】

Ｉ－ＶＯＰのみを使用して符号化を行う場合、フレーム間圧縮に関わる部分である動き補償が不必要になるため、直前にエンコードした画像が不必要になる。したがって、符号化後の画像に対する逆量子化や逆ＤＣＴ（離散コサイン変換）を省略することができるため、大幅に符号化時間を短縮することができる（非特許文献６参照）が、符号化後の符号量が大きくなることや任意形状の符号化が行えないことが問題となる。なお、Ｉ－ＶＯＰとは、入力画像をブロックに分け、ブロック毎にＤＣＴを施してから量子化し、この量子化したＤＣＴ係数を可変長符号化（Ｉｎｔｒａ　符号化）したＶＯＰを表している。

【００２０】

符号量が大きくなる問題については、ＭＰＥＧ－４規格では量子化スケールを設定できるため符号量の調節が可能だが、量子化スケールを大きく設定するに従って画質が悪化するため、変換器を搭載するホストのＣＰＵパワーに余力がある場合にはフレーム間圧縮も行うことが望ましい。そのため、フレーム間圧縮の可能なＰ－ＶＯＰ（前方予測符号化されたＶＯＰ）やＢ－ＶＯＰ（双方向予測符号化されたＶＯＰ）を含めた符号化の可能な変換器を構築する。

【００２１】

図２に、ＤＶデータをＭＰＥＧ符号化データに変換する変換器の構成を示す。この変換器構成自体、及びその動作は周知である。本発明は、符号化処理を行うに際して、図２中のすべての符号化処理を、基本処理と付加的処理に分類する。うち基本処理とは最低限必要

な符号化処理、付加的処理は処理時間に余裕がある時にのみ行う処理である。図２中、太枠で示された各処理、「可変長復号」「逆量子化」「逆ＤＣＴ」「再構成」「ＤＣＴ」「量子化」「可変長符号化」は基本処理に該当する。これらは、全ＶＯＰをＩ－ＶＯＰとした時に必要となる最小限度の処理である。これに対して、「ＡＣ／ＤＣ予測」「逆量子化」「逆ＤＣＴ」「動き補償」「動き情報符号化」は付加的処理と定義される。これらのうち「ＡＣ／ＤＣ予測」はＩ－ＶＯＰおよびＰ－ＶＯＰに対して符号量を節約するために用いられ、他の各処理はＰ－ＶＯＰおよびＢ－ＶＯＰによるフレーム間圧縮のために用いられる。

【００２２】

２）　配信モジュールにおける符号化処理

例示のシステムでは、Ｐ－ＶＯＰやＢ－ＶＯＰを使用するか否かと、それらの使用頻度を決定するため、各種処理の符号化実行時間を計測して１ＶＯＰあたりの平均符号化時間を予測し、定められたアルゴリズムに従って判定を行う。図３は、このような配信モジュールにおける符号化処理を説明する図である。

【００２３】

ステップＳ１：

上述したように、符号化処理を、基本処理と付加的処理に分類する。

【００２４】

ステップＳ２：

配信モジュール起動時にＩＢＰＢＩＢＰＢＩＢＰＢ…．の順に入力画像の符号化を行って各種ＶＯＰについて複数個分の処理時間計測を行い、各種処理に要する平均処理時間を測定する。ここで、Ｉ、Ｂ、Ｐはそれぞれ、Ｉ－ＶＯＰ、Ｂ－ＶＯＰ、Ｐ－ＶＯＰをそれぞれ表している。

【００２５】

この処理時間計測では、ＤＶデコーダにおける「可変長復号」「逆量子化」「逆ＤＣＴ」「ブロック再構成」、ＭＰＥＧ－４エンコーダにおける「ＤＣＴ」「量子化」「可変長符号化」「ＡＣ／ＤＣ予測」「逆量子化」「逆ＤＣＴ」「動き補償」「動き情報符号化」の１ＶＯＰあたりの処理時間を計測して平均値を求める。これらのうち、「動き補償」はＰ－ＶＯＰとＢ－ＶＯＰで異なる処理を行っており処理時間が大きく異なるため、それぞれ別個に平均値を求める。このようにして、各種処理の符号化実行時間を計測する。

【００２６】

各ＶＯＰの符号化に必要な時間について、さらに説明する。全ＶＯＰの符号化をＩ－ＶＯＰのみを使用して実行する場合、最短の時間で符号化することが可能であるが、Ｐ－ＶＯＰやＢ－ＶＯＰを使用した符号化を行う場合、前後のＶＯＰと動き補償を行うための参照画像を用意する必要があるため、逆量子化や逆ＤＣＴが必要になる。したがって、同一ＶＯ内でＰ－ＶＯＰやＢ－ＶＯＰを使用するか否かにより、Ｉ－ＶＯＰの符号化に必要な時間が異なる。評価に必要な各パラメータを以下のように定義する。

T_{min}：最小限の符号化に必要とされる時間

T_{ACDC}（ｎ）：ｎ番目のＶＯＰのＡＣ／ＤＣ予測に要する時間

T_{IQ}：逆量子化に必要な時間

T_{IDCT}：逆ＤＣＴに必要な時間

Ｔｃ：動き情報符号化に要する時間

T_M（ｎ）：　ｎ番目のＶＯＰの動き補償に要する時間

【００２７】

上記のうち、T_MとT_{ACDC}を除く各数値は使用するマシンのスペックや対象画像の大きさに対して固定値を取る。これに対し、途中で動き補償の処理を中断しても中断結果を用いた符号化が可能なため、T_Mは可変値に設定することが可能である。また、ＡＣ／ＤＣ予測はフレームごとに行う／行わないの決定が可能なため、やはり可変値を取る。具体的には、T_MはＰ－ＶＯＰでは$0 \leqq T_M(n) \leqq T_{M-Pmax}$、Ｂ－ＶＯＰでは$0 \leqq T_M(n) \leqq T_{M-Bmax}$のいずれかの値、$T_{ACDC}$は、０もしくは$T_{ACDCmax}$

のいずれかの値を取る。ここで、T_{M-Pmax}、T_{M-Bmax}、$T_{ACDCmax}$はそれぞれ、Ｐ－ＶＯＰおよびＢ－ＶＯＰにおいてすべてのマクロブロックに対する動き補償を行う際に要する時間、およびＡＣ／ＤＣ予測を実行した場合に１ＶＯＰあたりのＡＣ／ＤＣ予測に要する時間と定義される。上記の各パラメータは、以下の処理時間を計測し、その平均値を算出することにより得られる。

【００２８】

T_{min}：　ＤＶデコーダにおける「可変長復号」「逆量子化」「逆ＤＣＴ」「ブロック再構成」、ＭＰＥＧ－４エンコーダにおける「ＤＣＴ」「量子化」「可変長符号化」の処理時間の合計

$T_{ACDCmax}$：　「ＡＣ／ＤＣ予測」の処理時間

T_{IQ}：　ＭＰＥＧ－４エンコーダにおける「逆量子化」の処理時間

T_{IDCT}：　ＭＰＥＧ－４エンコーダにおける「逆ＤＣＴ」の処理時間

Ｔｃ：　「動き情報符号化」の処理時間

T_{M-Pmax}：　Ｐ－ＶＯＰにおける「動き補償」の処理時間

T_{M-Bmax}：　Ｂ－ＶＯＰにおける「動き補償」の処理時間

なお、上記のうち「動き補償」は、完全に処理を実行した際の処理時間を計測するものとしている。

【００２９】

各ＶＯＰの符号化時間は、以下のようになる。

I-VOPの符号化

・　他のVOPでP-VOPやB-VOPを使用しない場合

$$T(n)=T_{min}+T_{ACDC}(n) \qquad (1)$$

・　他のVOPでP-VOPやB-VOPを使用する時

$$T(n)=T_{min}+T_{ACDC}(n)+T_{IQ}+T_{IDCT} \qquad (2)$$

P-VOP、B-VOPの符号化

$$T(n)=T_{min}+T_{ACDC}(n)+T_{IQ}+T_{IDCT}+Tc+T_{M}(n) \qquad (3)$$

N個のVOP中に出現するI-VOP 、P-VOP、B-VOPの数をそれぞれNIVOP、NPVOP、NBVOPと置くと、N個のVOPにおける符号化時間の合計は

$$\begin{aligned}Ttotal&=\Sigma_{I\text{-}VOP}\,T(n)+\Sigma_{P\text{-}VOP}\,T(n)+\Sigma_{B\text{-}VOP}\,T(n)\\&=N_{IVOP}\,(T_{min}+T_{IQ}+T_{IDCT})+\Sigma_{I\text{-}VOP}T_{ACDC}(n)\\&+N_{PVOP}\,(T_{min}+T_{IQ}+T_{IDCT}+Tc)+\Sigma_{P\text{-}VOP}\,(T_{ACDC}(n)+T_{M}(n))\\&+N_{BVOP}\,(T_{min}+T_{IQ}+T_{IDCT}+Tc)+\Sigma_{B\text{-}VOP}\,(T_{ACDC}(n)+T_{M}(n))\\&=N\times(T_{min}+T_{IQ}+T_{IDCT})+(N_{PVOP}+N_{BVOP})\,Tc\\&+\Sigma_{all\text{-}VOP}T_{ACDC}(n)+\Sigma_{P\text{-}VOP}T_{M}(n)+\Sigma_{B\text{-}VOP}T_{M}(n) \qquad (4)\end{aligned}$$

その際の１フレームの平均符号化時間は

$$\begin{aligned}Tave&=Ttotal/N=T_{min}+T_{IQ}+T_{IDCT}+Tc\,(N_{PVOP}+N_{BVOP})/N\\&+(\Sigma_{all\text{-}VOP}T_{ACDC}(n)+\Sigma_{P\text{-}VOP}T_{M}(n)+\Sigma_{B\text{-}VOP}T_{M}(n))/N \qquad (5)\end{aligned}$$

T_{M}、T_{ACDC}の各値が可変であることから、（５）式のうち、Σで示された各項目を自由に定めることができる。$T_{ACDCmax}$はT_{M-Pmax}の数分の一程度であり、T_{M-Pmax}はT_{M-Bmax}の数分の一程度であることから、ＡＣ／ＤＣ予測、Ｐ－ＶＯＰの動き補償、Ｂ－ＶＯＰの動き補償の順に優先的に符号化時間を配分する。実際に

はＡＣ／ＤＣ予測の符号化全体に占める時間割合は小さな値のため、通常はＡＣ／ＤＣ予測を実行するものとして説明する。

【００３０】

ステップＳ３：

ステップ２で測定された平均処理時間の値を用いて、ＶＯＰ判定アルゴリズムに従って使用するＶＯＰの種類および頻度を決定する。

【００３１】

目標フレームレートをＲと置くと、１フレームの目標平均符号化時間は１／Ｒとなる。（５）式より、Ｎ個のＶＯＰ中においてN_{PVOP}個のＰ－ＶＯＰを使用して完全に動き補償を実行する際の平均符号化時間は、

$$T_{ave}=T_{min}+T_{IQ}+T_{IDCT}+(T_c+T_{M-Pmax})N_{PVOP}/N+T_{ACDCmax} \quad (6)$$

となる。したがって、Ｎ個のＶＯＰ中においてN_{PVOP}個のＰ－ＶＯＰを使用するための最低条件は

$$1/R \geqq T_{min}+T_{IQ}+T_{IDCT}+(T_c+T_{M-Pmax})N_{PVOP}/N+T_{ACDCmax} \quad (7)$$

となる。また、Ｎ個のＶＯＰ中においてN_{PVOP}個のＰ－ＶＯＰおよびN_{BVOP}個のＢ－ＶＯＰを使用して、Ｐ－ＶＯＰおよびＢ－ＶＯＰ中のすべてのマクロブロックに対して完全に動き補償を行った場合、平均符号化時間は

$$T_{ave}=T_{min}+T_{IQ}+T_{IDCT}+(T_c+T_{M-Pmax})N_{PVOP}/N+(T_c+T_{M-Bmax})N_{BVOP}/N+T_{ACDCmax} \quad (8)$$

となる。したがって、Ｎ個のＶＯＰ中においてN_{PVOP}個のＰ－ＶＯＰおよびN_{BVOP}個のＢ－ＶＯＰを使用するための条件は

$$1/R \geqq T_{min}+T_{IQ}+T_{IDCT}+(T_c+T_{M-Pmax})N_{PVOP}/N+(T_c+T_{M-Bmax})N_{BVOP}/N+T_{ACDCmax} \quad (9)$$

となる。

【００３２】

実際には各ＶＯＰの比率は自由に変更できるわけではなく、２個のＩ－ＶＯＰ間に挟まれるＰ－ＶＯＰの数や、２個のＩ－ＶＯＰもしくはＰ－ＶＯＰの間に挟まれるＢ－ＶＯＰの数を変更することにより比率が決定される。連続する２個のＩ－ＶＯＰの間に挟まれるＰ－ＶＯＰの数に１を加えたものをＮｐ、連続する２個のＩ－ＶＯＰもしくは Ｐ－ＶＯＰの間に挟まれるＢ－ＶＯＰの数に１を加えたものをＮｂと定義すると、Ｎ個の連続したＶＯＰに含まれる各ＶＯＰの数は以下のようになる。

$$N_{IVOP}=N/N_p\ N_b \quad (10)$$

$$N_{PVOP}=(N_p-1)N/N_p\ N_b \quad (11)$$

$$N_{BVOP}=(N_b-1)N/N_b \quad (12)$$

このようにすると、（７）式および（９）式は以下のようになる。

$$1/R \geqq T_{min}+T_{IQ}+T_{IDCT}+(T_c+T_{M-Pmax})(N_p-1)/N_p N_b+T_{ACDCmax} \quad (13)$$

$$1/R \geqq T_{min}+T_{IQ}+T_{IDCT}+(T_c+T_{M-Pmax})(N_p-1)/N_p N_b+(T_c+T_{M-Bmax})(N_b-1)/N_b+T_{ACDCmax} \quad (14)$$

適正なＮＢＶＯＰ、ＮＰＶＯＰを決定するためのアルゴリズムを図４に示す。アルゴリズムの入力ＴおよびＲは、それぞれ各符号化処理の処理時間および目標フレームレートである。うちＴは、全ての符号化処理時間を含み、Ｔ＝［T_{min}，T_{ACDC}，T_{IQ}，T_{IDCT}，T_c，T_{M-Pmax}，T_{M-Bmax}］で示される。

【００３３】

始めに、N_{BVOP}＝１、N_{PVOP}＝２で（１３）式の判定を行う。これで全ＶＯＰの動き補償が可能なようならN_{PVOP}を二倍ずつ増やして判定する。これを、あらかじめ設定されている最大値である$N_{PVOPmax}$に達するまで判定を続ける。この途中で（１３）式を満たさない条件が得られた場合、N_{PVOP}を直前の値に戻して、Ｉ－ＶＯＰとＰ－ＶＯＰを使用した符号化を行う。

【００３４】

ＮＰＶＯＰ＝$N_{PVOPmax}$に達してもなお全ＶＯＰの動き補償が可能なようならN_{BVOP}＝２とし、以後二倍ずつ増やして（１４）式の判定を行い、$N_{BVOPmax}$に達するまで判定を続ける。この途中で（１４）式を満たさない条件が得られた場合、N_{BVOP}を直前の値に戻して、全種類のＶＯＰを使用した符号化を行う。その際、N_{BVOP}を増やしてゆくに従ってＩ－ＶＯＰが減るので、Ｉ－ＶＯＰの比率を一定に保つため、N_{BVOP}を二倍にするたびにN_{PVOP}を二分の一にする。

【００３５】

ステップＳ４：

ステップＳ３で決定されたＶＯＰの種類及び頻度に従って符号化を続けてゆく。

【００３６】

【発明の効果】

本発明は、複数の画像・音声を扱った編集・配信作業が可能なシステムを、編集・配信・受信の３種類のモジュールにより構成し、その配信モジュールを使用するマシンの性能に応じて使用するＶＯＰの種類を変更することにより、可能な限り圧縮効率の良い符号化法を自動で選択することが可能となる。

【図面の簡単な説明】

【図１】構築するシステムの概念図を示す。

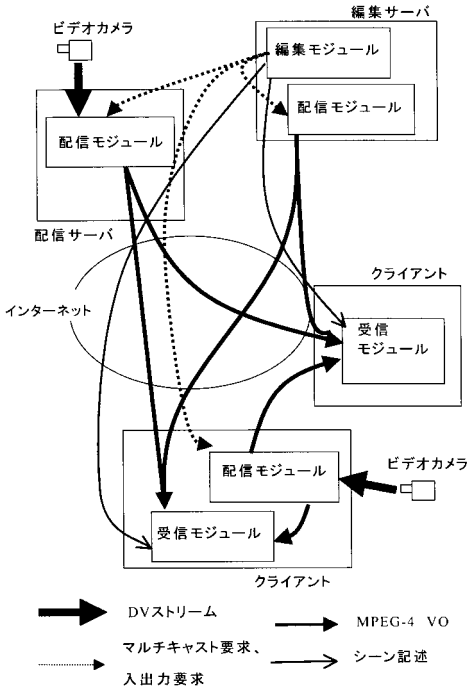
【図２】変換器の構成を示す図である。

【図３】配信モジュールにおける符号化処理を説明する図である。

【図４】適正なＮＢＶＯＰ、ＮＰＶＯＰを決定するためのアルゴリズムを示す図である。

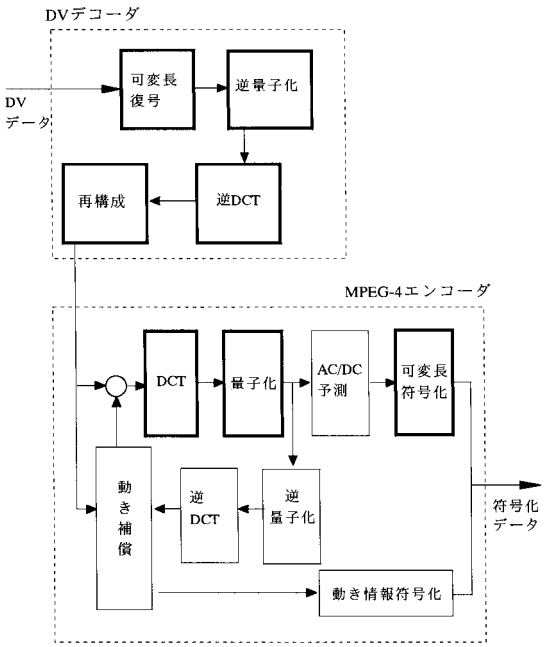
【図１】

実時間コンテンツ編集システムの構成例



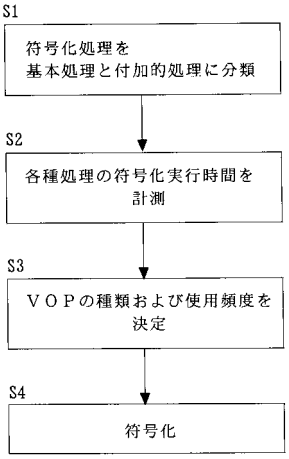
【図２】

ＤＶ－ＭＰＥＧ変換器の構成



【図３】

配信モジュールにおける符号化処理



【図４】

符号化パラメータ決定のためのアルゴリズム

```
Decide_param(T, R){

Nb=1;

for(Np=2;Np<=Npmax;Np=Np*2)
    if(1/R <Tmin+TIQ+TIDCT
        +(Tc+TM-Pmax)(Np-1)/NpNb+TACDCmax){
                Np=Np/2;
                return(Np,Nb);
    }

Np=Npmax;

for(Nb=2;Nb<=Nbmax;Nb=Nb*2){
    Np=Np/2;
    if(1/R <Tmin+TIQ+TIDCT
    +(Tc+TM-Pmax)(Np-1)/NpNb
    +(Tc+TM-Bmax)(Nb-1)/Nb+TACDCmax){
                Nb=Nb/2;
                return(Np,Nb);
    }
}

Nb=Nbmax;
Np=Npmax/Nbmax;
return(Np,Nb);
}
```

フロントページの続き

Ｆターム(参考) 5C059 KK36 MA00 MA05 MA14 MA23 MC11 MC38 ME01 PP05 PP06
PP07 SS08 TA25 TB04 TC00 TD03 TD06 UA02
5C064 BA01 BC10 BC18 BD08