

(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2010-503280  
(P2010-503280A)

(43) 公表日 平成22年1月28日(2010.1.28)

(51) Int.Cl.		F I				テーマコード (参考)
HO4N 7/15 (2006.01)		HO4N 7/15	610			5C159
HO4N 7/26 (2006.01)		HO4N 7/13	Z			5C164

審査請求 未請求 予備審査請求 未請求 (全 41 頁)

(21) 出願番号 特願2009-526799 (P2009-526799)  
 (86) (22) 出願日 平成19年8月15日 (2007. 8. 15)  
 (85) 翻訳文提出日 平成21年2月25日 (2009. 2. 25)  
 (86) 国際出願番号 PCT/US2007/075963  
 (87) 国際公開番号 W02008/027724  
 (87) 国際公開日 平成20年3月6日 (2008. 3. 6)  
 (31) 優先権主張番号 11/511, 749  
 (32) 優先日 平成18年8月29日 (2006. 8. 29)  
 (33) 優先権主張国 米国 (US)

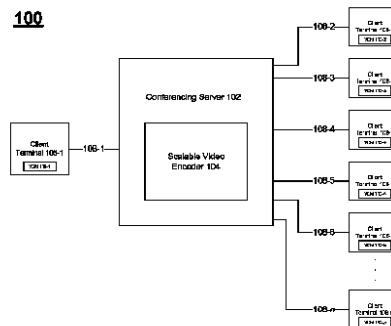
(71) 出願人 500046438  
 マイクロソフト コーポレーション  
 アメリカ合衆国 ワシントン州 9805  
 2-6399 レッドモンド ワン マイ  
 クロソフト ウェイ  
 (74) 代理人 100077481  
 弁理士 谷 義一  
 (74) 代理人 100088915  
 弁理士 阿部 和夫  
 (72) 発明者 ウォーレン ブイ. バークレイ  
 アメリカ合衆国 98052 ワシントン  
 州 レッドモンド ワン マイクロソフト  
 ウェイ マイクロソフト コーポレーシ  
 ョン インターナショナル パテンツ内

最終頁に続く

(54) 【発明の名称】 マルチメディア音声会議向け視覚的構成の管理技術

(57) 【要約】

マルチメディア音声会議向けの視覚的構成を管理する技術を説明する。装置はプロセッサを備えることができ、そのプロセッサは、全てのディスプレイオブジェクトに対するディスプレイオブジェクトビットレートの合計が入力ビットレートの合計以下になるように複数のディスプレイオブジェクトにディスプレイオブジェクトビットレートを割り当て、複数のビデオストリームからビデオ情報を復号化する。それらの複数のビデオストリームの各々は、様々なレベルの空間解像度、時間解像度および品質を2つ以上のディスプレイオブジェクトに対して持つ様々なビデオ階層を有する。他の実施形態も説明および主張してある。



**【特許請求の範囲】****【請求項 1】**

符号化ビデオ情報を、様々なレベルの空間解像度、時間解像度および品質を持つ様々なビデオ階層を各々が有する複数のビデオストリームで受信すること、

複数のディスプレイオブジェクトに対する入力ビットレートの合計を決定すること、

ディスプレイオブジェクトビットレートを、全てのディスプレイオブジェクトに対するディスプレイオブジェクトビットレートの合計が前記入力ビットレートの合計以下であるように複数のディスプレイオブジェクトに割り当てること、および

前記割り当てに従って、ビデオ情報を前記ディスプレイオブジェクトに対する前記の様々なビデオ階層から復号化すること、

を含む方法。

10

**【請求項 2】**

サブスクリプションメッセージを会議サーバに送信して音声会議向けの様々なビデオ階層を取得することを含む、請求項 1 に記載の方法。

**【請求項 3】**

ディスプレイフレームに対するディスプレイオブジェクトビットレートを、ディスプレイ上の前記ディスプレイフレームのサイズまたは前記ディスプレイフレームの位置に基づいて修正することを含む、請求項 1 または 2 に記載の方法。

**【請求項 4】**

ディスプレイフレームに対するディスプレイオブジェクトビットレートを、通信リンクに対する瞬間的なチャンネル容量に基づいて修正することを含む、請求項 1 から 3 の任意の 1 つに記載の方法。

20

**【請求項 5】**

前記復号化ビデオ情報を各ディスプレイフレームに描画して音声会議向けの視覚的構成を作成することを含む、請求項 1 から 4 の任意の 1 つに記載の方法。

**【請求項 6】**

クライアント端末で、前記符号化ビデオ情報を、音声会議向けの様々なレベルの空間解像度、時間解像度および品質を各々が持つ様々なビデオ階層を有する前記ビデオストリームで受信することを含む、請求項 1 から 5 の任意の 1 つに記載の方法。

**【請求項 7】**

前記符号化ビデオ情報を前記ビデオストリームで受信することであって、各ビデオストリームは、第 1 のレベルの空間解像度および第 1 のレベルの時間解像度を有するベース階層と、前記第 1 のレベルの空間解像度または前記第 1 のレベルの時間解像度を増やす拡張階層とを含む様々なビデオ階層を有することを含む、請求項 1 から 6 の任意の 1 つに記載の方法。

30

**【請求項 8】**

アクティブな話者に基づいてディスプレイフレームに対するディスプレイオブジェクトビットレートを修正することを含む、請求項 1 から 7 の任意の 1 つに記載の方法。

**【請求項 9】**

符号化ビデオ情報を、様々なレベルの空間解像度、時間解像度および品質を持つ様々なビデオ階層を各々が有する複数のビデオストリームから受信および解析するための一連のパーサ、

40

前記パーサに接続された視覚的構成モジュールであって、複数のディスプレイオブジェクトに対する入力ビットレートの合計を決定し、全てのディスプレイオブジェクトに対するディスプレイオブジェクトビットレートの合計が前記入力ビットレートの合計以下であるように、ディスプレイオブジェクトビットレートを複数のディスプレイオブジェクトに割り当てる前記視覚的構成モジュール、および

前記割り当てに従って、ビデオ情報を前記ディスプレイオブジェクトに対する各ビデオストリーム内の前記の様々なビデオ階層から復号化するための、前記視覚的構成モジュールに接続された一連のビデオ復号器、

50

を備えるクライアント端末。

【請求項 10】

前記視覚的構成モジュールはサブスクリプションメッセージを会議サーバに送信して音声会議向けの様々なビデオ階層を取得する、請求項 9 に記載のクライアント端末。

【請求項 11】

前記視覚的構成モジュールはディスプレイフレームに対するディスプレイオブジェクトビットレートを、ディスプレイ上の前記ディスプレイフレームのサイズまたは前記ディスプレイフレームの位置に基づいて修正する、請求項 9 または 10 に記載のクライアント端末。

【請求項 12】

前記視覚的構成モジュールはディスプレイフレームに対するディスプレイオブジェクトビットレートを、通信リンクに対する瞬間的なチャンネル容量に基づいて修正する、請求項 9 から 11 の任意の 1 つに記載のクライアント端末。

【請求項 13】

ディスプレイに接続された描画モジュールであって、前記復号化ビデオ情報を各ディスプレイフレームに描画して音声会議向けの視覚的構成を前記ディスプレイ上に作成する前記描画モジュールを備える、請求項 9 から 12 の任意の 1 つに記載のクライアント端末。

【請求項 14】

符号化ビデオ情報を複数のビデオストリームで受信する受信機であって、各ビデオストリームが、第 1 のレベルの空間解像度および第 1 のレベルの時間解像度を有するベース階層と、前記第 1 のレベルの空間解像度または前記第 1 のレベルの時間解像度を増やす拡張階層とを含む様々なビデオ階層を有する受信機、および

前記受信機に接続された適応スケジューリングモジュールであって、前記の様々なビデオ階層を様々な時刻に受信クライアント端末に送信する前記適応スケジューリングモジュールを備える会議サーバ。

【請求項 15】

前記適応スケジューリングモジュールは前記ベース階層の送信後まで前記拡張階層の送信を遅らせる、請求項 14 に記載の会議サーバ。

【請求項 16】

前記適応スケジューリングモジュールは、新たな主要話者を決定する際に前記ベース階層を第 1 の時刻に前記受信クライアント端末に送信する、請求項 14 または 15 に記載の会議サーバ。

【請求項 17】

前記適応スケジューリングモジュールは、新たな主要話者を決定する際に前記ベース階層を第 1 の時刻に前記受信クライアント端末に送信し、所定の待機時間後の第 2 の時刻に前記拡張階層を送信する、請求項 14 から 16 の任意の 1 つに記載の会議サーバ。

【請求項 18】

前記適応スケジューリングモジュールに接続された送信機であって、前記ベース階層と前記拡張階層とを異なる予定時刻に送信する前記送信機を備える、請求項 14 から 17 の任意の 1 つに記載の会議サーバ。

【請求項 19】

データ処理装置での実行時に、請求項 1 から 8 の任意の 1 つに記載の方法を実装し、および / または請求項 9 から 13 の任意の 1 つに記載のクライアント端末を実装し、および / または請求項 14 から 18 の任意の 1 つに記載の会議サーバを実装するコンピュータプログラム要素を備えるコンピュータプログラム。

【請求項 20】

請求項 19 に記載のコンピュータプログラムを保持するコンピュータプログラムキャリア媒体。

【発明の詳細な説明】

【背景技術】

10

20

30

40

50

## 【0001】

マルチメディア音声会議では一般に、複数のエンドポイント間の音声通信、ビデオ、および/またはデータ情報が関与する。データネットワークが拡大し、マルチメディア会議では従来の回線交換ネットワークよりパケットネットワークが使用されつつある。パケットネットワーク上でのマルチメディア音声会議を設定するため、一般には会議サーバが動作して音声会議を調整および管理する。会議サーバはビデオストリームを送信側の参加者から受信し、そのビデオストリームを音声会議の他の参加者にマルチキャストする。

## 【0002】

ビデオ音声会議向けのデジタルビデオのようなマルチメディア情報の送信に関する1つの問題は、デジタルビデオにより(場合により、埋め込まれたデジタルオーディオと組み合わせられて)記憶容量と送信容量が多くの場合大量に消費されることである。一般的な生のデジタルビデオシーケンスには毎秒15、30または60フレーム(フレーム/秒)が含まれる。各フレームには無数の画素が含まれる可能性がある。各画素は微小な画像要素を表す。生の形態では、画素は一般にコンピュータにより例えば24ビットで表現される。したがって、一般的な生のデジタルビデオシーケンスのビットレートまたは1秒当りのビット数は毎秒5百万ビット(ビット/秒)以上の大きさになる可能性がある。ほとんどのメディア処理装置および通信ネットワークには生のデジタルビデオを処理するリソースが不足している。このため、メディア通信システムでは情報源圧縮(コーディングまたは符号化とも呼ばれる)を用いてデジタルビデオのビットレートを低減する。復元(または復号化)は圧縮の逆を行う。

## 【0003】

一般に、所与の処理装置および/または通信ネットワークに対して特定の種類のビデオ圧縮を選択することには設計のトレードオフが存在する。例えば、圧縮は、高いビットレートは得られないがビデオの品質の高さが維持されるロスレス、またはビデオの品質は落ちるがビットレートの減少がより劇的であるロッシーであることができる。ほとんどのシステム設計では、一連の所与の設計制約および性能要件に基づいて品質とビットレートとの間で何らかの妥協がなされる。結果として、所与のビデオ圧縮技術は一般に様々な種類のメディア処理装置および/または通信ネットワークには適していない。このことは、1つまたは複数の受信装置が複数のディスプレイフレーム、ウィンドウまたは他のオブジェクトを利用してマルチメディア音声会議の様々な参加者向けにビデオ情報を表示する際に特に問題になる恐れがある。この問題は、様々な話者集団を収容するための様々なディスプレイウィンドウに様々な参加者が現れるとさらに悪化する。

## 【発明の概要】

## 【0004】

本発明の概要は選択した概念を簡潔な形で導入するために提供する。本概要は以下の発明を実施するための形態でさらに詳細に説明する。本概要により請求の範囲に記載されている主題の主要な特徴または本質的な特徴を特定することは意図しておらず、本概要を用いて請求の範囲に記載されている主題の範囲を限定することも意図していない。

## 【0005】

様々な実施形態は一般に、ビデオ、イメージ、画像等のようなデジタルメディアコンテンツのデジタル符号化、復号化および処理を対象としている。幾つかの実施形態では、デジタルメディアコンテンツのデジタル符号化、復号化および処理は、SMPTE(Society of Motion Picture and Television Engineers)標準421M(「VC-1」)のビデオコーデックシリーズ標準およびその変形に基づくことができる。より詳細には、幾つかの実施形態は、複数の解像度ビデオ符号化および復号化技術、ならびに下位互換性を崩さずにその技術をVC-1ビットストリームで可能とする方法を対象としている。例えば、1つの実施形態では、デジタルビデオ情報を拡張SMPTE VC-1ビデオストリームまたはビットストリームに圧縮または符号化するように配置されたビデオ符号器を機器に含めることができる。ビデオ符号器は、ベース階層ならびに1つまたは複数の空間および/または時間拡張階層のような複

10

20

30

40

50

数の階層の形でデジタルビデオ情報を符号化することができる。ベース階層は所定の最低限の空間解像度とベースレベルの時間解像度を提供することができる。1つまたは複数の拡張階層は符号化ビデオ情報を含むことができ、その符号化ビデオ情報を用いて、ベース階層に符号化されたビデオ情報に対するベースレベルの空間解像度および/またはベースレベルの時間解像度を増やすことができる。

【0006】

様々な実施形態では、ビデオ復号器はビデオ情報をベース階層および1つまたは複数の拡張階層から選択的に復号化して、所望の品質レベルでビデオ情報を再生することができる。同様に、オーディオビデオマルチポイント制御装置(AVMCU)は、現在利用可能なネットワーク帯域幅と受信機の復号化能力のような情報に基づいて、ビデオ情報をベース階層および1つまたは複数の拡張階層から会議参加者に転送することを選択することができる。

10

【0007】

幾つかの実施形態では特に、ビデオ情報をビデオストリームのベース階層および1つまたは複数の拡張階層から選択的に復号化して、音声会議に用いる視覚的構成向けの様々なレベルのビデオ解像度および品質でビデオ情報を再生するようにビデオ復号器を配置することができる。視覚的構成には一般にディスプレイウィンドウのような複数のディスプレイオブジェクトが含まれ、各々は音声会議の様々な参加者のビデオ情報を表示する。例えば、1つの実施形態では、受信クライアント端末のようなクライアントは処理システム、メモリおよびディスプレイを含むことができる。処理システムは、音声会議向け視覚的構成に対する複数のディスプレイオブジェクトにディスプレイオブジェクトビットレートを割り当てるように配置されたプロセッサを含むことができる。全てのディスプレイオブジェクトに対するディスプレイオブジェクトビットレートの合計がクライアント端末のようなクライアントに対する入力ビットレートの合計以下になるように上記の割り当てを行うことができる。クライアント端末は次いで、2つ以上のディスプレイオブジェクトに対して様々なレベルの空間解像度、時間解像度および品質を各々が有する様々なビデオ階層に対するサブスクリプション要求を、会議サーバまたは送信クライアント端末に対する上記の割り当てに基づいて送信することができる。このように、クライアント端末はその入力ビットレート割当量を効率的に使用することができる。あるいは、クライアント端末はスケラブルビデオストリームを受信することができ、ビデオ復号器は適切なビデオ情報をビデオストリームの様々なビデオ階層から復号化することができる。このように、クライアント端末はその計算リソースを効率的に使用することができる。上記の復号化ビデオ情報を各ディスプレイフレームに描画して音声会議向けの視覚的構成を上記ディスプレイ上に作成するように、描画モジュール(例えば、ディスプレイチップセット)を配置することができる。

20

30

【0008】

様々な実施形態は、会議サーバまたはAVMCU向けの適応スケジューリング技術も対象とすることができる。複数のビデオストリーム内の符号化ビデオ情報を受信するように配置された受信機を機器に含めることができる。その複数のビデオストリームの各々は、第1のレベルの空間解像度および第1のレベルの時間解像度を有するベース階層と、その第1のレベルの空間解像度または第1のレベルの時間解像度を増やす拡張階層とを含む様々なビデオ階層を有する。上記の機器は、上記の受信機に接続された適応スケジューリングモジュールをさらに含むことができる。音声会議において主要またはアクティブな話者の変化に応じて様々なビデオ階層を様々な時点で受信クライアント端末に送信するように上記の適応スケジューリングモジュールを配置することができる。他の実施形態も説明および主張してある。

40

【図面の簡単な説明】

【0009】

【図1】マルチメディア会議システムの実施形態を示す図である。

【図2】視覚的構成の実施形態を示す図である。

50

【図3】コンピューティング環境の実施形態を示す図である。

【図4】クライアント端末の実施形態を示す図である。

【図5】論理フローの実施形態を示す図である。

【図6】ビデオキャプチャ再生システムの実施形態を示す図である。

【図7】一般的なビデオ符号化システムの実施形態を示す図である。

【図8】一般的なビデオ復号化システムの実施形態を示す図である。

【図9】ビデオ階層構造の実施形態を示す図である。

【図10】適応スケジューリング技術の第1の図を示す図である。

【図11】適応スケジューリング技術の第2の図を示す図である。

【図12】適応スケジューリング技術の第3の図を示す図である。

10

【発明を実施するための形態】

【0010】

様々な実施形態は、マルチメディア音声会議向け視覚的構成の管理を対象とすることができる。マルチパーティのビデオ会議では、各受信クライアント端末はビデオストリームを、音声会議に参加し自身のビデオストリームを送信する他のクライアント端末の各々から受信する。受信クライアント端末は他の受信クライアント端末からの複数のビデオストリームを視覚的構成の形でディスプレイスクリーン上に配置することができる。視覚的構成は全てまたは一部の参加者からのビデオストリームを、所与のクライアント端末のディスプレイ装置上にモザイクとして描画または表示する。例えば、視覚的構成は現在アクティブな話者のビデオ情報を表示するためのトップディスプレイオブジェクトを有することができ、他の参加者のパノラマ表示をそのトップディスプレイオブジェクトの下に位置する小型のディスプレイオブジェクト集合により表示することができる。

20

【0011】

各視覚的構成には異なる伝達要件がある。例えば、小型ディスプレイオブジェクトおよびピクチャインピクチャディスプレイには大型のディスプレイオブジェクトより低い空間解像度という要件があるかもしれない。同様に、あまりアクティブでない参加者のビデオ情報には、よりアクティブな参加者のビデオより低い時間解像度という要件があるかもしれない。一般に空間および/または時間解像度が低いと、SNR (signal-to-noise ratio) または他の測定基準で測定されることが多い所与の画像品質に対するビットレートは遅い。所与の時空間解像度に対して、一般に画像品質が低いとビットレートは遅く、一般に画像品質が高いとビットレートは速い。視覚的構成の中には、参加者の一部または全部に対して画像品質が低いという要件を有するものがある。

30

【0012】

各クライアント端末は一般に、全入力ビットレート割当量、または制約、および全出力ビットレート割当量を有する。そのため、その入力ビットレートおよび出力ビットレートの割当量を効率的に利用することが1つの設計目標である。したがって、様々な実施形態ではスケラブルビデオ表現を実装して、所与のクライアント端末がマルチメディア音声会議向けの視覚的構成を描画または表示するための全入力ビットレート割当量および/または全出力ビットレート割当量に対する利用および効率性を改善することができる。各ビデオストリームに対して複数の空間解像度、時間解像度、および品質レベルを利用できるので、クライアント端末は、視覚的構成に必要なビデオ情報のみを選択的に受信および/または復号化することにより、任意の所与の構成に対する自身の入力ビットレート割当量を効率的に利用することができる。

40

【0013】

様々な実施形態では、視覚的構成モジュールをクライアント端末、会議サーバ、または音声会議で用いる任意の他の装置に実装することができる。視覚的構成モジュールはスケラブルビデオストリームの全体または一部を受信することができ、それに応じて拡張復号化および視覚的構成表示の動作を実施することができる。様々な実施形態では、視覚的構成モジュールは所与のクライアント端末の複数のディスプレイオブジェクトに対する入力ビットレートの合計を受信することができる。いったん視覚的構成モジュールがクライ

50

アント端末に対する入力ビットレート割当量の合計を受信すると、視覚的構成モジュールはディスプレイオブジェクトビットレートを、クライアント端末の視覚的構成に用いる各ディスプレイオブジェクトに動的に割り当てることができる。視覚的構成モジュールはディスプレイオブジェクトビットレートを、後述のように任意数の因子に基づいて所与のディスプレイオブジェクトに割り当てることができる。例えば、幾つかの実施形態では、視覚的構成モジュールはディスプレイオブジェクトのサイズ、ディスプレイオブジェクトの位置、および所与の通信リンクまたはメディアチャネルに対する瞬間的なチャネル容量に基づいてディスプレイオブジェクトビットレートを割り当てることができる。

#### 【0014】

割り当て動作中は、視覚的構成モジュールはディスプレイオブジェクトビットレートの割り当てを、クライアント端末に対する入力ビットレートの合計以下である、全てのディスプレイオブジェクトに対するディスプレイオブジェクトビットレートの合計に制限する。視覚的構成モジュールは、アクティブな話者の変化、ディスプレイオブジェクトのサイズの変化、所与のディスプレイオブジェクト内のビデオ情報に対する動き量の変化、状態の変化（一時停止ビデオまたはストリーミングビデオ）等のような変化条件に基づいてディスプレイオブジェクトビットレートの割り当てを動的に変化させることができる。視覚的構成モジュールは、拡張ビデオ情報を拡張ビデオ符号器から復号可能なスケーラブルビデオ復号器にディスプレイオブジェクトビットレートの割り当てを出力することができる。スケーラブルビデオ復号器はディスプレイオブジェクトビットレートの割り当てを視覚的構成モジュールから受信し、スケーラブル復号化動作を開始して、そのディスプレイオブジェクトビットレートの割り当てに従ってビデオ情報を各ディスプレイオブジェクトに対する様々なビデオ階層から復号化することができる。所与のビデオ情報およびディスプレイオブジェクトの集合に対し、スケーラブルビデオ復号器は様々なレベルの空間解像度、時間解像度および品質を復号化することができる。あるいは、視覚的構成モジュールはサブスクリプションメッセージを会議サーバに送信して、視覚的構成内の各ディスプレイオブジェクトに対する所望のレベルの解像度および品質を有する様々なビデオ階層を要求することができる。このように、スケーラブルビデオ符号器/復号器および/または視覚的構成モジュールは、マルチメディア音声会議の複数の参加者に対応する複数のディスプレイオブジェクトで視覚的構成を描画する際に、所与のクライアント端末に対する入力ビットレートの効果的な使用を改善することができる。

#### 【0015】

様々な実施形態は、マルチメディア音声会議向けの適応スケジューリング技術も対象とすることができる。新たな主要またはアクティブな話者が所与の送信クライアント端末から話し始めると、会議サーバまたはAVMCUは、任意の受信クライアント端末がその新たな主要話者でディスプレイオブジェクトを描画し始めることができるように、新たなビデオキーフレームに対するキーフレーム要求を送信することができる。しかしながら、キーフレームは相対的に大きく、したがって他のビデオフレームと比較して送信時間が非常に長い。その結果、ビデオ遅延がより大きく、その新たな主要話者を参加者が見ることができるまでに数秒かかる。

#### 【0016】

様々な実施形態では、適応スケジューリングモジュールを用いてこれらおよび他の問題を解決することができる。送信クライアント端末のような別の装置の代わりに時間内にビデオ階層の送信を適応的にスケジューリングできるように、適応スケジューリングモジュールを配置することができる。その結果、主要またはアクティブな話者が話し始め、自身のビデオを送信し始める際の応答時間を改善することができる。低ビデオ階層を最初に送信し、追加の階層を徐々に送信して、長時間に渡ってビデオ品質を改善する。このように、新たな主要話者が話し始める際に、ある空間または時間解像度からより精度の良いものへスムーズに遷移し、それによりディスプレイオブジェクト内のスイッチを起動して新たな主要話者のビデオ情報を示すような視覚的構成を描画することができる。ビデオ階層の送信を適応的にスケジューリングすることで、ちらつき、ブランキング、および視覚的構

10

20

30

40

50

成内で主要話者と対応するディスプレイオブジェクトとの間を遷移することにより発生する他の副作用を軽減することができる。

【0017】

#### マルチメディア会議システム

図1にマルチメディア会議システム100のブロック図を示す。マルチメディア会議システム100により、様々な実施形態を実装するのに適した一般的なシステムアーキテクチャを表現することができる。マルチメディア会議システム100は複数の要素を備えることができる。要素は、ある特定の動作を実施するように配置された任意の物理または論理構造を備えることができる。各要素を、所与の一連の設計パラメータまたは性能制約に対して望ましいように、ハードウェア、ソフトウェア、またはそれらの任意の組み合わせで実装することができる。ハードウェア要素の例には、装置、コンポーネント、プロセッサ、マイクロプロセッサ、回路、回路要素(例えば、トランジスタ、レジスタ、キャパシタ、インダクタ、等)、集積回路、ASIC(application specific integrated circuit)、PLD(programmable logic device)、DSP(digital signal processor)、FPGA(field programmable gate array)、記憶装置、論理ゲート、レジスタ、半導体素子、チップ、マイクロチップ、チップセット、等を含めることができる。ソフトウェアの例には、任意のソフトウェアコンポーネント、プログラム、アプリケーション、コンピュータプログラム、アプリケーションプログラム、システムプログラム、マシンプログラム、オペレーティングシステムソフトウェア、ミドルウェア、ファームウェア、ソフトウェアモジュール、ルーチン、サブルーチン、関数、メソッド、インタフェース、ソフトウェアインタフェース、API(application program interface)、命令セット、計算コード、コンピュータコード、コードセグメント、コンピュータコードセグメント、ワード、値、シンボル、またはそれらの任意の組み合わせを含めることができる。図1に示すように、マルチメディア会議システム100はある特定のトポロジで有限数の要素を有するが、マルチメディア会議システム100は所与の実装に対して望ましいように代替的なトポロジでそれより多いかまたは少ない要素を含んでもよいことは理解できるだろう。本実施形態はこの文脈に限定されない。

10

20

【0018】

様々な実施形態では、メディア情報および制御情報のような様々な種類の情報を送信、管理または処理するようにマルチメディア会議システム100を配置することができる。メディア情報の例には一般に、音声情報、ビデオ情報、オーディオ情報、イメージ情報、テキスト情報、数値情報、英数字記号、グラフィック、等のようなユーザ向けのコンテンツを表現する任意のデータを含めることができる。制御情報は自動化システム向けのコマンド、命令または制御ワードを表現する任意のデータを指すことができる。例えば、制御情報を使用して、システムを通してメディア情報を経路付けして装置間の接続を確立し、所定の方法でメディア情報を処理するように装置に指示することができる、等である。ビデオフレームをビデオ情報から選択的に除去してビデオビットレートを下げる文脈で幾つかの実施形態を具体的に説明しているが、様々な実施形態は画像、イメージ、データ、音声、音楽またはそれらの任意の組み合わせのような任意の種類 of 所望のメディア情報の使用を網羅することに留意されたい。

30

40

【0019】

様々な実施形態では、マルチメディア会議システム100は会議サーバ102を含むことができる。会議サーバ102は、クライアント端末106-1-m間のマルチメディア音声会議を管理または制御するように配置された任意の論理または物理エンティティを備えることができる。ここで、mは会議内の端末数を表す。様々な実施形態では、会議サーバ102は、コンピュータ、サーバ、ルータ、スイッチ、ブリッジ、等のような処理装置またはコンピューティング装置を備えるか、または会議サーバ102をそれらとして実装することができる。会議サーバ102に対する具体的な実装は、会議サーバ102に用い

50



られる一連の通信プロトコルまたは標準により様々であってよい。1つの例では、会議サーバ102をITU(International Telecommunication Union)H.323シリーズの標準および/またはその変形に従って実装することができる。H.323標準は音声会議の動作を調整するためのMCU(multipoint control unit)を定義する。特に、MCUはH.245信号送信を扱うMC(multipoint controller)、およびデータストリームを混合および処理するための1つまたは複数のMP(multipoint processor)を含む。別の例では、会議サーバ102をIETF(Internet Engineering Task Force)MMUSIC(Multiparty Multimedia Session Control)作業部会のSIP(Session Initiation Protocol)シリーズの標準および/またはその変形に従って実装することができる。SIPは、ビデオ、音声、インスタントメッセージング、オンラインゲーム、および仮想現実のようなマルチメディア要素が関与する対話的なユーザセッションを開始、修正、および終了する提案標準である。H.323およびSIP標準は両方とも本質的に、VoIP(Voice over Internet Protocol)またはVOP(Voice Over Packet)マルチメディア音声会議の動作に対するシグナリングプロトコルである。しかしながら、他のシグナリングプロトコルを会議サーバ102に対して実装することができ、そのプロトコルが本実施形態の範囲内であることは理解できるだろう。本実施形態はこの文脈に限定されない。

10

20

**【0020】**

様々な実施形態では、マルチメディア会議システム100は、1つまたは複数の通信リンク108-1-n上で会議サーバ102と接続するための1つまたは複数のクライアント端末106-1-mを含むことができる。ここで、mおよびnは正の数を表し、これらは必ずしも一致する必要はない。例えば、クライアントアプリケーションは、各々が別個の会議を表す幾つかのクライアント端末を同時にホストすることができる。同様に、クライアントアプリケーションは複数のメディアストリームを受信することができる。例えば、全てまたは一部の参加者からのビデオストリームを、参加者のディスプレイのトップウィンドウにモザイクとして、現在アクティブな話者のビデオで表示することができ、他の参加者のパノラマ表示を他のウィンドウに表示することができる。クライアント端末106-1-mは、会議サーバ102が管理するマルチメディア音声会議に参加または従事するように配置された任意の論理的または物理的エンティティを備えることができる。クライアント端末106-1-mを、プロセッサおよびメモリ(例えば、メモリユニット110-1-p)を含む処理システム、1つまたは複数のマルチメディア入出力(I/O)コンポーネント、ならびに無線および/または有線ネットワーク接続をその最も基本的な形において含む任意の装置として実装することができる。マルチメディアI/Oコンポーネントの例には、オーディオI/Oコンポーネント(例えば、マイクロフォン、スピーカ)、ビデオI/Oコンポーネント(例えば、ビデオカメラ、ディスプレイ)、触覚(I/O)コンポーネント(例えば、パインレータ)、ユーザデータ(I/O)コンポーネント(例えば、キーボード、サムボード、キーパッド、タッチスクリーン)、等を含めることができる。クライアント端末106-1-mの例には、電話、VoIPまたはVOP電話、PSTN(Packet Switched Telephone Network)上で動作するように設計されたパケット電話、インターネット電話、ビデオ電話、携帯電話、PDA(personal digital assistant)、携帯電話とPDAの組み合わせ、モバイルコンピューティング装置、スマートフォン、単方向ページャ、双方向ページャ、メッセージング装置、コンピュータ、PC(personal computer)、デスクトップコンピュータ、ラップトップコンピュータ、ノートブックコンピュータ、ハンドヘルドコンピュータ、ネットワーク機器、等を含めることができる。本実施形態はこの文脈に限定されない。

30

40

**【0021】**

動作モードにより、クライアント端末106-1-mを送信クライアント端末または受

50

信クライアント端末と呼ぶことができる。例えば、ビデオストリームを会議サーバ102に送信するように動作する場合、所与のクライアント端末106-1-mを送信クライアント端末と呼ぶことができる。別の例では、例えば送信クライアント端末からのビデオストリームのようなビデオストリームを会議サーバ102から受信するように動作する場合、所与のクライアント端末106-1-mを受信クライアント端末と呼ぶことができる。後述する様々な実施形態では、クライアント端末106-1を送信クライアント端末として説明し、クライアント端末106-2-mを受信クライアント端末として説明するが、これは例に過ぎない。クライアント端末106-1-mのうち任意のものは音声会議の経過を通じて送信または受信クライアント端末として動作することができ、音声会議における様々な時点でモードを頻繁に切り替えることができる。本実施形態はこの点に限定されない。

10

#### 【0022】

様々な実施形態では、マルチメディア会議システム100は、有線通信システム、無線通信システム、またはその両方の組み合わせを備えるか、またはそれらの一部を形成することができる。例えば、マルチメディア会議システム100は、1つまたは複数の種類の有線通信リンク上で情報を送信するように配置された1つまたは複数の要素を含むことができる。有線通信リンクは、有線、ケーブル、バス、PCB (printed circuit board)、イーサネット (登録商標) 接続、P2P (peer-to-peer) 接続、バックプレーン、スイッチファブリック、半導体材料、ツイストペア線、同軸ケーブル、光ファイバ接続、等を含むことができるがこれらに限らない。マルチメディア会議システム100は、1つまたは複数の種類の無線通信リンク上で情報を送信するように配置された1つまたは複数の要素を含むこともできる。無線通信の例は、無線チャネル、赤外線チャネル、RF (radio-frequency) チャネル、WiFi (Wireless Fidelity) チャネル、RFスペクトルの一部、および/あるいは1つまたは複数の認可周波数帯域または自由周波数帯域を含むことができるがこれらに限らない。

20

#### 【0023】

メディア処理向けの様々な標準および/またはプロトコルに従って動作するようにマルチメディア会議システム100を配置することもできる。メディア処理標準の例には、SMPT E (Society of Motion Picture and Television Engineers) 421M (「VC-1」) シリーズの標準およびその変形、MICROSOFT (登録商標) WINDOWS (登録商標) MEDIA VIDEOバージョン9 (WMV-9) シリーズの標準およびその変形、DVB-T (Digital Video Broadcasting Terrestrial) 放送標準、ITU/IEC H.263標準、低ビットレート通信向けビデオコーディング、2000年11月公開のITU-T勧告H.263v3および/またはITU/IEC H.264標準、超低速ビットレート通信向けビデオコーディング、2003年5月公開のITU-T勧告H.264、MPEG (Motion Picture Experts Group) 標準 (例えば、MPEG-1、MPEG-2、MPEG-4)、および/またはHiperLAN (High performance radio Local Area Network) 標準が含まれるがこれらに限らない。メディア処理プロトコルの例には、SDP (Session Description Protocol)、RTSP (Real Time Streaming Protocol)、RTP (Real-time Transport Protocol)、SMIL (Synchronized Multimedia Integration Language) プロトコル、および/またはISMA (Internet Streaming Media Alliance) プロトコルが含まれるがこれらに限らない。本実施形態はこの文脈に限定されない。

30

40

#### 【0024】

例えば、1つの実施形態では、マルチメディア会議システム100の会議サーバ102

50

およびクライアント端末 106-1-m を、1つまたは複数の H.323 シリーズの標準および/またはその変形に従って動作する H.323 システムの一部として実装することができる。H.323 は、コンピュータ、機器、およびサービス品質を保証しないネットワーク上でのマルチメディア通信サービスに関する仕様を提供する ITU 標準である。H.323 のコンピュータおよび機器はリアルタイムビデオ、オーディオ、およびデータ、またはこれら要素の任意の組み合わせを有することができる。この標準は、コールシグナリングプロトコル、ならびにデータおよびオーディオビジュアル通信向けプロトコルを加えた、IETF の RTP および RTCP プロトコルに基づく。H.323 はオーディオおよびビデオ情報をネットワーク上での送信向けにフォーマットおよびパッケージする方法を定義する。標準的なオーディオおよびビデオコーデック/復号器(コーデック)はオーディオおよびビデオソースからの入力/出力をノード間の通信向けに符号化および復号化する。コーデックはオーディオまたはビデオ信号をアナログおよびデジタル形式の間で変換する。さらに、H.323 はデータ通信および会議向けの T.120 サービスを H.323 セッション内および H.323 セッションの次に行うことを規定する。T.120 サポートサービスは、所与の実装に対して望ましいように、データ処理が H.323 オーディオおよびビデオと関連して行われるか、または別個に行われるかのいずれかの可能性があることを意味する。

10

#### 【0025】

一般的な H.323 システムに従って、会議サーバ 102 を、H.323 ゲートウェイ、H.323 ゲートキーパ、1つまたは複数の H.323 端末 106-1-m、ならびに(例えば、ローカルエリアネットワーク上の)パーソナルコンピュータ、サーバおよび他のネットワーク装置のような複数の他の装置に接続された MCU として実装することができる。H.323 装置を H.323 シリーズの標準またはその変形に準拠して実装することができる。以下でさらに説明するように、H.323 クライアント端末 106-1-m を各々“エンドポイント”と考える。H.323 エンドポイントは、メディアチャネルの使用をネゴシエーションするための H.245 制御シグナリング、コールシグナリングおよびコールセットアップのための Q.931 (H.225.0)、H.225.0 RAS (Registration, Admission, and Status)、ならびにオーディオおよびビデオパケットを配列するための RTP/RTCP をサポートする。H.323 エンドポイントは、様々なオーディオおよびビデオコーデック、T.120 データ会議プロトコルおよびある特定の MCU 機能をさらに実装することができる。幾つかの実施形態を例としてのみ H.323 システムの文脈で説明できるが、マルチメディア会議システム 100 を 1つまたは複数の IETF SIP シリーズの標準および/またはその変形、ならびに他のマルチメディアシグナリング標準に従って実装してもよく、それらはなお本実施形態の範囲内にあることは理解できるだろう。本実施形態はこの文脈に限定されない。

20

30

#### 【0026】

一般的な動作では、マルチメディア会議システム 100 をマルチメディア音声会議に使用することができる。マルチメディア音声会議には一般に、音声、ビデオ、および/またはデータ情報を複数のエンドポイント間で送信することが含まれる。例えば、公衆パケットネットワークまたはプライベートパケットネットワークを、オーディオ音声会議、ビデオ音声会議、オーディオ/ビデオ音声会議、共同文書共有および編集 (collaborative document sharing and editing)、等に使用することができる。パケットネットワークを、回線交換情報とパケット情報との間で変換するように配置された 1つまたは複数の適切な VoIP ゲートウェイを介して PSTN に接続することもできる。パケットネットワーク上でマルチメディア音声会議を設定するために、各クライアント端末 106-1-m は、例えば低帯域幅 PSTN 電話接続、中間帯域幅 DSL モデム接続またはケーブルモデム接続、および LAN (local area network) 上の高帯域幅イントラネット接続のような、様々な接続速度または帯域幅で動作する様々な種類の有線または無線の通信リンク 108-1-n を用いて会議サ

40

50

ーバ102に接続することができる。

【0027】

マルチパーティのビデオ会議では、各受信クライアント端末106-2-mはビデオストリームを、音声会議に参加し自身のビデオストリームを送信する他のクライアント端末の各々から受信する。受信クライアント端末106-2-mは、他の受信クライアント端末106-2-mからの複数のビデオストリームをディスプレイスクリーン上に視覚的構成の形で配置することができる。これを例えば、それぞれクライアント端末106-1-mの一部として実装された視覚的構成モジュール110-1-pを用いて実現することができる。視覚的構成の代表例を、図2を参照して説明することができる。

【0028】

図2は視覚的構成の例示的な実施形態を示す。図2は、所与のクライアント端末106-1-mのディスプレイ上にモザイクとして表示された全てまたは一部の参加者からのビデオストリームを有する視覚的構成200を示す。図2に示すように、トップディスプレイオブジェクト202は、現在アクティブな話者のビデオ情報を表示するように配置されたディスプレイウィンドウを備えることができ、他の参加者のパノラマ表示を、トップディスプレイオブジェクトの下に位置する小型のディスプレイオブジェクト集合204-1-vにより表示することができる。アクティブな話者が小型のディスプレイオブジェクト集合204-1-vの1つに表示されている参加者の1人になると、小型のディスプレイオブジェクト集合204-1-vの1つからのアクティブな話者のビデオ情報をトップディスプレイオブジェクト202に表示することができ、逆もまた同様である。視覚的構成200は視覚的構成の一例に過ぎず、所与の実装または特定の音声会議に対して望ましいように、様々な数のディスプレイオブジェクトおよび様々なサイズのディスプレイオブジェクトを有する他の視覚的構成を使用してもよいことは理解できるだろう。例えば、ディスプレイオブジェクト202、204を（例えば、背景ありまたは背景なしの）“ヘッドアンドショルダ”カットアウト、他のオブジェクトを重ね合わせることができる透過オブジェクト、遠近法での長方形領域、等として実装することができる。本実施形態はこの文脈に限定されない。

【0029】

図2に示すように、視覚的構成200は、現在アクティブな話者を含むメインウィンドウ、他の参加者の複数の小型ウィンドウ、および半透過的に重ね合わせたような他の要素を含むことができる。さらに、視覚的構成は動的であることができる。アクティブな話者は変わる可能性があるので、音声会議内の他の参加者をメインウィンドウおよびピクチャインピクチャを通して交代させることができる。場合によっては、全ての参加者が常に見えていなくともよい。可視である参加者の集合は時間とともに変わってもよい。

【0030】

幾つかの実施形態では、視覚的構成は複数の会議に関与することができる。参加者は、互いに対する関係と相対的な重要性に従って、各音声会議を適切に配置させたいかもしれない。原則として、これらの会議は互いに完全に独立している可能性がある。しかしながら、場合によっては、それらはメイン会議のサブ会議であろう。例えば、補助会議は主要会議の別の参加者とのサイドチャットであるかもしれない。

【0031】

各クライアント端末106-1-mは独自の一意的な視覚的構成を構築することを選択できる。一般に、ビデオ情報を構成から完全に削除するかまたは特別な位置に置くなど、クライアント端末106-1-m上に表示する際にクライアント端末106-1-mのユーザに対してビデオ情報を描画することには特別な処置が存在する。

【0032】

各視覚的構成には異なる伝達要件がある。例えば、小型ディスプレイオブジェクトおよびピクチャインピクチャディスプレイには、大型のディスプレイオブジェクトより低い空間解像度という要件があるかもしれない。同様に、あまりアクティブでない参加者のビデオ情報には、よりアクティブな参加者のビデオより低い時間解像度という要件があるかも

10

20

30

40

50

しれない。

【0033】

空間解像度は一般に、測定されている空間の細部に関する精度の測定値を指すことができる。デジタルビデオの文脈では、空間解像度をフレーム、画像またはイメージ内の画素数で測定または表現することができる。例えば、640×480画素のデジタルイメージサイズは326,688個の個々の画素と等価である。一般に、高い空間解像度を有するイメージは低い空間解像度のイメージより多い数の画素で構成される。空間解像度はとりわけ、ビデオフレーム、画像、またはイメージの画像品質に影響を及ぼす可能性がある。

【0034】

時間解像度は一般に、時間に関する特定の測定精度を指すことができる。デジタルビデオの文脈では、時間解像度をフレームレート、または15フレーム/秒、30フレーム/秒、60フレーム/秒のような秒ごとにキャプチャされたビデオ情報のフレーム数、等で測定または表現することができる。一般に、高い時間解像度は、低い時間解像度のものよりも多い1秒当りのフレーム数を指す。時間解像度はとりわけ、ビデオイメージまたはフレームのシーケンスに対する動き表現に影響を及ぼす可能性がある。ビデオストリームまたはビットストリームは、オーディオおよび/またはビデオ情報を表す連続的なセグメント(例えば、ビットまたはバイト)のシーケンスを指すことができる。

【0035】

一般に空間および/時間解像度が低いと、SNR(signal-to-noise ratio)または他の測定基準で測定されることが多い所与の画像品質に対するビットレートは遅い。所与の時空間解像度に対して、一般に画像品質が低いとビットレートは遅く、一般に画像品質が高いとビットレートは速い。視覚的構成の中には、参加者の一部または全部に対して画像品質が低いという要件を有するものがあってもよい。

【0036】

各クライアント端末106-1-mは一般に、全入力ビットレート割当量、または制約、および全出力ビットレート割当量を有する。そのため、その入力ビットレートおよび出力ビットレートの割当量を効率的に利用することが1つの設計目標である。したがって、様々な実施形態ではスケラブルビデオ符号器104を実装して、所与のクライアント端末106-1-mがマルチメディア音声会議向けの視覚的構成を描画または表示するための全入力ビットレート割当量および/または全出力ビットレート割当量に対する利用および効率性を改善することができる。各ビデオストリームに対して複数の空間解像度、時間解像度、および品質レベルを利用できるので、クライアント端末106-1-mは、視覚的構成に必要なビデオ情報のみを選択的に受信および/または復号化することにより、任意の所与の構成に対する自身の入力ビットレート割当量を効率的に使用することができる。

【0037】

様々な実施形態では、埋め込みコーディングまたは階層コーディングと呼ばれることがある1つまたは複数のスケラブルコーディング技術および復号技術を用いて動作するように、スケラブルビデオ符号器104を実装することができる。スケラブルコーディングは、会議サーバ102のような送信機が複数の空間解像度、時間解像度、および品質レベルを生成して、全出力ビットレート割当量を効率的に利用しつつこれらの複数のレベルを送信するための効率的な方法である。対照的に、従来は、複数バージョンの同一ビデオは独立した符号化情報として生成され、全て並列に送信される。この技術は“サイマルキャスト”と呼ばれることもある。サイマルキャスト技術は一般に全入力ビットレート割当量および/または全出力ビットレート割当量を非効率的に利用する。マルチメディア会議システム100を一般に、およびスケラブルビデオ符号器104およびビデオ構成モジュール110-1-pを特に、図3を参照して説明することができる。

【0038】

図3はコンピューティング環境300のブロック図を示す。コンピューティング環境300を、会議サーバ102および/またはクライアント端末106-1-mのような装置

10

20

30

40

50

、または装置の一部として実装することができる。幾つかの実施形態では、ソフトウェア 310 を実行するようにコンピューティング環境 300 を実装することができる。ソフトウェア 310 の例には、スケーラブルビデオ符号器 104 および / または視覚的構成モジュール 110 - 1 - p を含めることができる。例えば、コンピューティング環境 300 を会議サーバ 102 の一部として実装する際は、ソフトウェアプログラム 310 にスケーラブルビデオ符号器 104 ならびに / または視覚的構成モジュール 110 - 1 - p およびその付随データを含めることができる。別の例では、コンピューティング環境 300 をクライアント端末 106 - 1 - m の一部として実装する際は、ソフトウェアプログラム 310 にスケーラブルビデオ符号器 104 ならびに / または視覚的構成モジュール 110 - 1 - p およびその付随データを含めることができる。さらに別の例では、コンピューティング環境 300 を会議サーバ 102 および / またはクライアント端末 106 - 1 - m の一部として実装する際は、ソフトウェアプログラム 310 に、一般に電気、電子、および / または電気機械装置向けに実装されるオペレーティングシステムまたは他のシステムソフトウェアを含めることができる。コンピューティング環境 300 により記憶および実行されるソフトウェアとして実装されたスケーラブルビデオ符号器 104 および / または視覚的構成モジュール 110 - 1 - p 向けの動作で幾つかの実施形態を説明できるが、ソフトウェアモジュール 104、110 に対する動作を専用ハードウェア、ソフトウェア、またはそれらの任意の組み合わせを用いて実装してもよいことは理解できるだろう。本実施形態はこの文脈に限定されない。

#### 【0039】

最も基本的な構成では、コンピューティング環境 300 は一般に、少なくとも 1 つの処理装置 302 とメモリ 304 とを備える処理システム 308 を含む。処理装置 302 は、汎用目的プロセッサ、専用プロセッサ、メディアプロセッサ、コントローラ、マイクロコントローラ、組み込みプロセッサ、DSP (digital signal processor)、等のような、ソフトウェアを実行できる任意の種類のプロセッサであることができる。メモリ 304 を、揮発性および不揮発性メモリの両方を含む、データを記憶可能な任意の機械読取可能またはコンピュータ読取可能媒体を用いて実装することができる。例えば、メモリ 304 は ROM (read-only memory)、RAM (random-access memory)、DRAM (dynamic RAM)、DDRAM (Double-Data-Rate DRAM)、SDRAM (synchronous DRAM)、SRAM (static RAM)、PROM (programmable ROM)、EPROM (erasable programmable ROM)、EEPROM (electrically erasable programmable ROM)、フラッシュメモリ、強誘電体ポリマメモリのようなポリマメモリ、オポニックスメモリ、相変化または強誘電体メモリ、SONOS (silicon-oxide-nitride-oxide-silicon) メモリ、磁気または光カード、あるいは情報の記憶に適した任意の他種の媒体を含むことができる。

#### 【0040】

図 3 に示すように、メモリ 304 は、スケーラブルビデオ符号器 104、視覚的構成モジュール 110、および付随データのような様々なソフトウェアプログラム 310 を記憶することができる。スケーラブルビデオ符号器 104 に関する場合のような幾つかの場合、1 つのメディアストリームを一度に処理するよう設計すると、ソフトウェアプログラム 310 をメモリ内で複製しなければならない可能性がある。同様に、ホストシステムがマルチコアのマルチプロセッサベースコンピューティングプラットフォームである場合、プロセッサ 302 およびスケーラブルビデオ符号器 104 を数回複製することができる。メモリ 304 は、様々な種類のオペレーティングシステムソフトウェア、アプリケーションプログラム、ビデオコーデック、オーディオコーデック、呼制御ソフトウェア、ゲートキーパソフトウェア、マルチポイントコントローラ、マルチポイントプロセッサ、等のような、会議サーバ 102 の様々な態様を実装するための他のソフトウェアプログラムを記憶することもできる。あるいは上記の動作を、所与の実装に対して望ましいように専用ハー

ドウェア（例えば、DSP、ASIC、FPGA、等）あるいはハードウェア、ファームウェアおよび/またはソフトウェアの組み合わせの形で実装することができる。本実施形態はこの文脈に限定されない。

#### 【0041】

コンピューティング環境300は構成308以外の追加の特徴および/または機能性を有することもできる。例えば、コンピューティング環境300は記憶装置312を含むことができ、記憶装置312は様々な種類の取り外し可能または取り外し不能記憶装置を備えることができる。記憶装置312を、前述のように様々な種類の機械読取可能またはコンピュータ読取可能媒体のうち任意のものをを用いて実装することができる。コンピューティング環境300はキーボード、マウス、ペン、音声入力装置、タッチ入力装置、等のような1つまたは複数の入力装置314を有することもできる。ディスプレイ装置、スピーカ、プリンタ、等のような1つまたは複数の出力装置316も同様にコンピューティング環境300に含めることができる。

10

#### 【0042】

コンピューティング環境300はさらに1つまたは複数の通信接続318を含むことができ、それによりコンピューティング環境300は通信リンク108-1-nを介して他の装置と通信することができる。通信接続318は、1つまたは複数の通信インターフェース、ネットワークインターフェース、NIC (network interface card)、無線、無線送信機/受信機(トランシーバ)、有線および/または無線通信媒体、物理接続、等のような様々な種類の標準的な通信要素を含むことができる。通信媒体は一般にコンピュータ読取可能命令、データ構造、プログラムモジュールまたは他のデータを、搬送波または他の伝送メカニズムのような変調データ信号で具現化し、任意の情報配信媒体を含む。「変調データ信号」という用語は、1つまたは複数のその特性集合を有し、または信号内の情報を符号化するように変化する信号を意味する。限定ではなく例として、通信媒体には前述のように有線通信媒体および無線通信媒体の両方が含まれる。本明細書で用いる機械読取可能媒体およびコンピュータ読取可能媒体という用語は、記憶媒体と通信媒体の両方を含む。

20

#### 【0043】

様々な実施形態では、コンピューティング環境300をクライアント端末106-1-mの一部または全部として実装することができる。特に、コンピューティング環境300をソフトウェアプログラム310で実装して、1つまたは複数の視覚的構成モジュール110-1-pを含めることができる。マルチパーティのビデオ会議では、各受信クライアント端末106-1-mはビデオストリームを、音声会議に参加し自身のビデオストリームを送信する他のクライアント端末の各々から受信する。所与のクライアント端末106-1-mに対し、視覚的構成モジュール110-1-pは、他のクライアント端末106-1-mからの複数のビデオストリームを例えば視覚的構成200のようなディスプレイスクリーン上の視覚的構成内に配置することができる。

30

#### 【0044】

様々な実施形態では、各視覚的構成には異なる伝達要件がある。例えば、小型ディスプレイオブジェクトおよびピクチャインピクチャディスプレイには大型のディスプレイオブジェクトより低い空間解像度という要件があるかもしれない。同様に、あまりアクティブでない参加者のビデオ情報には、よりアクティブな参加者のビデオより低い時間解像度という要件があるかもしれない。一般に空間および/時間解像度が低いと、SNRまたは他の測定基準で測定されることが多い所与の画像品質に対するビットレートは遅い。所与の時空間解像度に対して、一般に画像品質が低いとビットレートは遅く、一般に画像品質が高いとビットレートは速い。視覚的構成の中には、参加者の一部または全部に対して画像品質が低いという要件を有するものがあるかもしれない。

40

#### 【0045】

各クライアント端末106-1-mは一般に、全入力ビットレート割当量、または制約、および全出力ビットレート割当量を有する。そのため、その入力ビットレートおよび出

50

力ビットレートの割当量を効率的に利用することが1つの設計目標である。様々な実施形態ではスケラブルビデオ符号器104を会議サーバ102または様々なクライアント端末106-1-mで実装して、所与のクライアント端末106-1-mがマルチメディア音声会議向けの視覚的構成を描画または表示するための全入力ビットレート割当量および/または全出力ビットレート割当量に対する利用および効率性を改善することができる。各ビデオストリームに対して複数の空間解像度、時間解像度、および品質レベルを利用できるので、クライアント端末106-1-mは、視覚的構成内の様々なディスプレイオブジェクトに必要なビデオ情報のみを選択的に受信および/または復号化することにより、任意の所与の視覚的構成に対する自身の入力ビットレート割当量を効率的に使用することができる。クライアント端末106-1-mおよび対応する視覚的構成モジュール110-1-pを、図4を参照してさらに説明することができる。

#### 【0046】

図4は代表的なクライアント端末の実施形態を示す。図4では、クライアント端末106-1-mの任意の1つを代表するクライアント端末のより詳細なブロック図を提供している。図4に示すように、クライアント端末106は、1つまたは複数の符号化ビデオストリーム402-1-oを入力として受信するように配置された有線または無線パーサ404-1-pを備えることができる。例えば、1つの実施形態では、符号化ビデオストリーム402-1-oを、会議サーバ102の一部として実装されたスケラブルビデオ符号器104により生成することができる。符号化ビデオストリーム402-1-oは、様々なビデオ階層で符号化されたビデオ情報を含むことができる。各ビデオ階層は様々なレベルの空間解像度、時間解像度および品質を有することができる。スケラブルビデオ符号器104は様々なビデオ階層を符号化ビデオストリーム402-1-oに多重化し、ビデオストリーム402-1-oを1つまたは複数の通信接続318を介して通信リンク108上で送信することができる。スケラブルビデオ符号器104および符号化ビデオストリーム402-1-oを、図6~9を参照して以下でさらに詳細に説明することができる。

#### 【0047】

様々な実施形態では、符号化ビデオストリーム402-1-oを1つまたは複数のパーサ401-1-pで受信することができる。パーサ404-1-pは受信したビデオストリーム402-1-oを、各々をパーサ401-1-pと通信可能に接続された1つまたは複数のスケラブルビデオ復号器406-1-rに出力することができる。パーサ404-1-pは受信したビデオストリーム402-1-oおよび/またはスケラブルインジケータを、やはりパーサ404-1-pに通信可能に接続された視覚的構成モジュール110に出力することもできる。

#### 【0048】

様々な実施形態では、視覚的構成モジュール110はビデオストリーム402-1-oおよび/またはスケラブルインジケータをパーサ404-1-pから受信することができる。両方の場合において、視覚的構成モジュール110はビデオストリーム402-1-oまたはスケラブルインジケータを使用して、ビデオストリーム402-1-oが、様々なレベルの空間解像度、時間解像度、および/または品質を有する様々なビデオ階層を含むかどうかを判断することができる。ビデオストリーム402-1-oまたはスケラブルインジケータの値を検査して非スケラブルビデオストリームがあることが分かると、ビデオ復号器406-1-qは通常通り復号化動作および視覚的構成の表示動作を行うことができる。視覚的構成モジュール110によりビデオストリーム402-1-oがスケラブルビデオストリームであると判断されると、ビデオ復号器406-1-qはそれに応じて拡張復号化を行うことができる。後者の場合、ビデオストリーム内で復号化される空間または時間スケールの量、およびビデオストリーム内で空間または時間スケールのどちらを復号化するかを視覚的構成モジュール110により決定する。いずれの場合も、同一の視覚的構成モジュール110を使用して、全入力ビットレート割当量または復号化性能のような別の制約を超過しないように、全ての構成ウィンドウに渡っ



て空間、時間解像度の割り当てを調整することができる。

【0049】

例えば、1つの実施形態では、視覚的構成モジュール110は、クライアント端末106の複数のディスプレイオブジェクトに対する入力ビットレートの合計を受信することができる。入力ビットレート値の合計をメモリ304から静的に受信するか、または描画モジュール408-1-r、通信インタフェース、トランシーバ、オペレーティングシステム、等から動的に受信することができる。入力ビットレート値の合計により、クライアント端末106に対する入力ビットレート割当量の合計を示すことができる。入力ビットレート割当量の合計は、通信リンク108の瞬間的な帯域幅、処理装置302の処理速度、メモリ304のメモリサイズ、メモリ帯域幅（例えば、メモリバス速度、アクセス時間、等）、ユーザが選択した品質および解像度基準、ディスプレイオブジェクトのサイズ、ディスプレイオブジェクトの位置、ディスプレイオブジェクトに対するビデオフレームシーケンス内の動き量、コーディングビットレート、グラフィックバス速度、等のような多数の因子に従って変化してもよい。さらに、入力ビットレート割当量の合計は1つまたは複数のこれらの因子に対する条件の変化に応じて、時間とともに変化してもよい。

10

【0050】

いったん視覚的構成モジュール110がクライアント端末106に対する入力ビットレート割当量の合計を受信すると、視覚的構成モジュール110はディスプレイオブジェクトビットレートを、クライアント端末106での視覚的構成に用いられる各ディスプレイオブジェクトに割り当てることができる。視覚的構成モジュール110は、入力ビットレート割当量の合計に関して前述した任意数の因子に基づいて、ディスプレイオブジェクトビットレートを所与のディスプレイオブジェクトに割り当てることができる。例えば、幾つかの実施形態では、視覚的構成モジュール110はディスプレイオブジェクトビットレート値を、ディスプレイオブジェクトのサイズ、ディスプレイオブジェクトの位置、および通信リンク108に対する瞬間的なチャンネル容量に基づいて割り当てることができる。

20

【0051】

割り当て動作中は、視覚的構成モジュール110はディスプレイオブジェクトビットレートの割り当てを、クライアント端末106に対する入力ビットレートの合計以下である、全てのディスプレイオブジェクトに対するディスプレイオブジェクトビットレートの合計に制限する。視覚的構成モジュール110は、アクティブな話者の変化、ディスプレイオブジェクトのサイズの変化、所与のディスプレイオブジェクト内のビデオ情報に対する動き量の変化、状態の変化（一時停止ビデオまたはストリーミングビデオ）等のような変化条件に基づいて、ディスプレイオブジェクトビットレートの割り当てを動的に変化させることができる。視覚的構成モジュール110はディスプレイオブジェクトビットレートの割り当てを、視覚的構成モジュール110に通信可能に接続されたスケーラブルビデオ復号器406-1-qに出力することができる。

30

【0052】

様々な実施形態では、クライアント端末106は一連のスケーラブルビデオ復号器406-1-qを、各ビデオストリーム402-1-oに対して1つのビデオ復号器406に含むことができる。スケーラブルビデオ復号器406-1-qは各々、それぞれの符号化ビデオストリーム402-1-oから拡張ビデオ情報を復号化可能なビデオ復号器を備えることができる。一般に、スケーラブルビデオ復号器406-1-qは、会議サーバ102で実装されたスケーラブルビデオ符号器104に適合するので、同様なスケーラブル復号化アルゴリズムに対応し、そのアルゴリズムを使用することができる。例えば、1つの実施形態では、スケーラブルビデオ復号器406-1-qは各々、ディスプレイオブジェクトビットレートの割り当てを視覚的構成モジュール110から受信し、スケーラブル復号化動作を開始して、そのディスプレイオブジェクトビットレートの割り当てに従ってビデオ情報を各ディスプレイオブジェクトに対する様々なビデオ階層から復号化することができる。所与のビデオビットストリームおよびディスプレイオブジェクトの集合に対し、スケーラブルビデオ復号器406-1-qは様々なレベルの空間解像度、時間解像度およ

40

50

び品質を復号化することができる。スケラブルビデオ復号器 406-1-q は復号化したビデオ情報を、スケラブルビデオ復号器 406-1-q に通信可能に接続された描画モジュール 408-1-r に出力することができる。

#### 【0053】

描画モジュール 408-1-r は所与のディスプレイオブジェクトに対する復号化ビデオ情報に対応するスケラブルビデオ復号器 406-1-q から受信し、描画動作を開始して復号化ビデオ情報をディスプレイオブジェクト内に描画し、音声会議向けの視覚的構成をディスプレイ 410 のようなディスプレイ上に作成することができる。描画モジュール 408-1-r を全てのビットストリーム復号器間で共有することができ、描画モジュール 408-1-r は予め設定された配置またはユーザ設定に従ってディスプレイオブジェクトを配置する。描画モジュール 408-1-r の例には、グラフィックチップセット、オペレーティングシステム向けグラフィックユーザインタフェースモジュール、等を含めることができる。さらに、所与のディスプレイオブジェクトに対するディスプレイオブジェクトのサイズまたはビデオ情報コンテンツにおける変化のような、自動またはユーザによる手動のいずれかにより視覚的構成が変化すると、描画モジュール 408-1-r はディスプレイオブジェクト変更情報を視覚的構成モジュール 110 に伝えることができる。あるいは、ディスプレイオブジェクト変更情報をオペレーティングシステムまたはグラフィックユーザインタフェースモジュールを介して視覚的構成モジュール 110 に送信してもよい。それに応じて、視覚的構成モジュール 110 はディスプレイオブジェクト変更情報を使用してディスプレイオブジェクトビットレート割り当てを変化させることができる。描画モジュール 408-1-r は、例えばディスプレイ 410 のような 1 つまたは複数のディスプレイ装置に表示情報を出力することができる。ディスプレイ 410 は視覚的構成 200 または別の視覚的構成と同様な視覚的構成を描画または表示することができる。視覚的構成はディスプレイオブジェクトのモザイクを備えることができ、各ディスプレイオブジェクトは音声会議の様々な参加者のビデオ情報を含み、クライアント端末 106 を用いる参加者のディスプレイオブジェクトを含む。

10

20

#### 【0054】

様々な実施形態では、描画モジュール 408-1-r および / またはディスプレイ 410 は、例えば階層マネージャ 412 から受信した階層情報に従って復号化ビデオ情報を描画することができる。階層マネージャ 412 は一連のディスプレイレイアウト情報に対するユーザ入力またはデフォルトの設定を受信することができる。ディスプレイレイアウト情報は、視覚的構成におけるディスプレイオブジェクトのレイアウトに関する情報を備えることができ、その情報にはディスプレイオブジェクトのサイズ、ディスプレイオブジェクトの位置、ディスプレイオブジェクト番号、各ディスプレイオブジェクトに対するビデオ解像度の優先度、等が含まれる。階層マネージャ 412 は、ビデオ構成内の各ビデオディスプレイオブジェクトに対する空間解像度パラメータおよびフレームレートパラメータを視覚的構成モジュール 110 に出力することができる。

30

#### 【0055】

視覚的構成モジュール 110 を再度参照すると、空間解像度および / またはフレームレートパラメータを階層マネージャ 412 から受信することができる。視覚的構成モジュール 110 は各ビデオディスプレイオブジェクトに対する空間解像度および / またはフレームレートパラメータを組み合わせて、規定のグローバル入力ビットレートを超過せずに各ビデオストリーム 402-1-o に対して改良した空間的および時間的組み合わせを計算する。視覚的構成モジュール 110 は、クライアント端末 106 に対する様々な種類のリソース利用を示す様々な入力パラメータも受信することができる。例えば、システムリソース（例えば、計算サイクル、メモリ割り当てまたは帯域幅、等）の場合、視覚的構成モジュール 110 は 1 つまたは複数のディスプレイオブジェクトに用いるスケールを修正することができる。別の例では、視覚的構成モジュール 110 は通信チャネル帯域幅の推定値を受信し、それに応じて様々なスケールを修正することができる。本実施形態はこの文脈に限定されない。

40

50

## 【0056】

幾つかの実施形態では、所与のクライアント端末106は「受信機駆動型階層化マルチキャスト」と呼ばれる技術を用いて、会議サーバ102に信号を伝達して、音声会議に対する別のビデオ階層を示すかまたは取得することができる。これは例えば、所与の装置（例えば、クライアント端末106、会議サーバ102、オーバーレイネットワーク内のノード、等）に対する入力ビットレート割当量の合計が符号化ビデオストリーム402-1-0に対して必要とされる帯域幅未満であるときに望ましい可能性がある。例えば、クライアント端末106が、スケーラブルコーデモジュール104により生成される符号化ビデオストリーム402-1-0より少ない入力ビットレート割当量の合計を有する場合、クライアント端末106の視覚的構成モジュール110は要求414を送信して、符号化ビデオストリーム402-1-0に対する帯域幅の合計をクライアント端末106に対する入力ビットレート割当量の合計以下に減らすように特定のビデオ階層に要求することができる。要求414は、各ビデオ入力に対する特定の空間および時間スケールの集合を会議サーバ102に送信させる要求を備えることもできる。その集合は、時間とともに変化することができる。さらに、要求414は、音声会議に対する別のビデオ階層を取得するための、会議サーバ102に対するサブスクリプションメッセージを備えることができる。本実施形態はこの文脈に限定されない。

10

## 【0057】

サブスクリプションメッセージを用いると、視覚的構成モジュール110の様々な要素を修正または削除することができる。例えば、視覚的構成モジュール110は必ずしもパーサ404-1-pを必要としなくともよい。むしろ、パーサ404-1-pを会議サーバ102により実装することができる。会議サーバ102がサブスクリプションメッセージをクライアント端末106から受信すると、会議サーバ102はパーサ404-1-pを使用して、会議サーバ102が受信した入力ビデオストリームを解析し、適切なビデオ階層をクライアント端末106に転送することができる。この場合、クライアント端末106で実装された視覚的構成モジュール110はパーサ404-1-pを回避または省略してもよい。本実施形態はこの文脈に限定されない。

20

## 【0058】

幾つかの実施形態では、会議サーバ102は、必ずしもサブスクリプションメッセージを所与のクライアント端末106-1-mから受信せずに、パーサ404-1-pを使用することができる。例えば、1つの実施形態では、会議サーバ102は各クライアント端末106-1-mに対する視覚的構成プロファイルを有することができる。各視覚的構成プロファイルは、対応するクライアント端末106-1-mまたは所与のマルチメディア音声会議が用いる視覚的構成に関する様々な視覚的構成パラメータを記憶することができる。例えば、視覚的構成パラメータは、主要話者が常に高解像度フォーマットであり、一方で全ての他の話者（所与のクライアント端末106-1-mのユーザは除く）は低解像度フォーマットであるべきであることを示すパラメータを含むことができる。視覚的構成パラメータを所与のクライアント端末106-1-m、会議サーバ102、音声会議の管理者、等により定義することができる。視覚的構成パラメータはさらにデフォルトの視覚的構成パラメータを含むことができる。デフォルトの視覚的構成パラメータを、所与のマルチメディア音声会議のある特定の特徴といった、様々な数の因子に基づいて選択することができる。上記特徴の例には、多数の参加者、各クライアント端末106-1-mに関連する能力情報、音声会議に用いられるメディア情報の種類（例えば、オーディオ、ビデオ、オーディオ/ビデオ、共有文書編集、等）、サブスクリプションレベル、サービスレベル品質、優先度レベル、地理的制約、帯域幅制約、等が含まれてもよい。このように、会議サーバ102は視覚的構成プロファイルおよび視覚的構成パラメータを使用して、1つまたは複数のクライアント端末106-1-mにより使用される視覚的構成を決定することができる。本実施形態はこの文脈に限定されない。

30

40

## 【0059】

上述の実施形態に対する動作を以下の図面およびそれに付随する例を参照してさらに説

50

明することができる。図面の一部は論理フローを含むことがある。本明細書で提示する上記図面は特定の論理フローを含むことがあるが、論理フローは本明細書で説明する一般的な機能性を実装できる方法の例を提供するに過ぎないことは理解できるだろう。さらに、特に明記しない限り、所与の論理フローを必ずしも提示した順序で実行する必要はない。さらに、所与の論理フローをハードウェア要素、プロセッサにより実行されるソフトウェア要素、またはそれらの任意の組み合わせで実装することができる。本実施形態はこの文脈に限定されない。

#### 【0060】

図5は論理フロー500の1つの実施形態を示す。論理フロー500は、マルチメディア会議システム100、会議サーバ102、スケーラブルビデオ符号器104、クライアント端末106、および/または視覚的構成モジュール110のような、本明細書で説明した1つまたは複数の実施形態により実行される動作の代表例であることができる。図5に示すように、ブロック502で、符号化ビデオ情報を、様々なレベルの空間解像度、時間解像度および品質を持つ様々なビデオ階層を各々が有する複数のビデオストリームで受信することができる。複数のディスプレイオブジェクトに対する入力ビットレートの合計をブロック504で決定することができる。ブロック506で、ディスプレイオブジェクトビットレートを複数のディスプレイオブジェクトに割り当てることができる。この場合、全てのディスプレイオブジェクトに対するディスプレイオブジェクトビットレートの合計は入力ビットレートの合計以下である。ブロック508で、ビデオ情報をディスプレイオブジェクトビットレートの割り当てに従ってディスプレイオブジェクトに対する様々なビデオ階層から復号化することができる。復号化ビデオ情報を各ディスプレイフレーム内で描画して、音声会議に対する視覚的構成を作成することができる。本実施形態はこの文脈に限定されない。

10

20

#### 【0061】

例えば、1つの実施形態では、サブスクリプションメッセージを会議サーバに送信して、音声会議に対する様々なビデオ階層を取得することができる。本実施形態はこの文脈に限定されない。

#### 【0062】

例えば、1つの実施形態では、ディスプレイフレームに対するディスプレイオブジェクトビットレートを、ディスプレイ上のディスプレイフレームのサイズまたはディスプレイフレームの位置に基づいて修正することができる。ディスプレイフレームに対するディスプレイオブジェクトビットレートを、通信リンクに対する瞬間的なチャネル容量に基づいて修正することもできる。ディスプレイフレームに対するディスプレイオブジェクトビットレートをアクティブな話者に基づいてさらに修正することができる。本実施形態はこの文脈に限定されない。

30

#### 【0063】

例えば、1つの実施形態では、符号化ビデオ情報を、様々なビデオ階層を有するビデオストリームで受信することができる。各ビデオ階層は、クライアント端末側での音声会議に対して様々なレベルの空間解像度、時間解像度および/または品質を含むことができる。例えば、1つの実施形態では、様々なビデオ階層は第1のレベルの空間解像度および第1のレベルの時間解像度を有するベース階層と、その第1のレベルの空間解像度または第1のレベルの時間解像度を増やす1つまたは複数の拡張階層とを含むことができる。本実施形態はこの文脈に限定されない。

40

#### 【0064】

##### スケーラブル符号化および復号化

前述のように、様々な実施形態は複数の解像度符号化および復号化技術を実装することができる。図1を再度参照すると、デジタルビデオ情報を共通ビデオストリーム内の複数のビデオ階層として符号化するようにスケーラブルビデオ符号器104を配置することができ、この場合、各ビデオ階層は1つまたは複数のレベルの空間解像度および/または時間解像度を提供する。図1に示すように、スケーラブルビデオ符号器104を会議サーバ

50

102の一部として実装することができる。これは例えば、1つまたは複数のクライアント端末106-1-mにスケラブルビデオ符号化機能が実装されていないときに望ましい可能性がある。その結果、会議サーバ102は、ビデオ情報を様々な受信ビデオストリームからアップサンプリングまたはダウンサンプリングすることでクライアント端末106-1-mの代わりにスケラブルビデオ符号化技術を実装することができる。しかしながら、幾つかの実施形態では、スケラブルビデオ符号器104をクライアント端末106-1-mの各々において実装することができる。この場合、会議サーバ102向けに実装されたスケラブルビデオ符号化機能を、サブスクリプションメッセージに応じて受信ビデオストリームを解析して適切なビデオ階層をクライアント端末106-1-mに転送することに限定するか、または完全に省略することができる。本実施形態はこの文脈に限定されない。

10

**【0065】**

スケラブルビデオ符号器104は、ベース階層および拡張階層のような複数のビデオ階層に対するデジタルビデオ情報を単一の共通ビデオストリームに多重化することができる。スケラブルビデオ符号器104により符号化されたビデオ情報を復号化するようにスケラブルビデオ復号器406-1-qを配置することができる。例えば、共通ビデオストリームからのビデオ情報を逆多重化または選択的に復号化して、ベース階層および1つまたは複数の拡張階層からビデオ情報を取り出し、例えば一般にSNR値で定義される所望のレベルの品質でビデオ情報を再生するように、スケラブルビデオ復号器を配置することができる。スケラブルビデオ復号器406-1-qは、各ビデオ階層に対して定義されたように様々な開始符号を用いてビデオ情報を選択的に復号化することができる。同様に、AVMCU(例えば、会議サーバ102)は、現在利用可能な帯域幅および復号器機能のような情報に基づいて1人または複数の参加者にベース階層と拡張階層のサブセットのみを転送することを選択することができる。AVMCUはビデオビットストリーム内の開始符号を用いて階層を選択する。

20

**【0066】**

例えば、1つの実施形態では、スケラブルビデオ符号器104はデジタルビデオ情報をベース階層ならびに1つまたは複数の時間および/または空間拡張階層として符号化することができる。ベース階層はデジタルビデオ情報に対する基本レベルまたは最低レベルの空間解像度および/または時間解像度を提供することができる。時間および/または空間拡張階層はデジタルビデオ情報に対する拡大した拡張レベルのビデオ空間解像度および/または時間解像度を提供することができる。様々な種類のエントリポイントおよび開始符号を定義して、ビデオストリーム内の様々なビデオ階層を区別することができる。このように、単一のスケラブルビデオ符号器は複数レベルの空間解像度および/または時間解像度を単一の共通ビデオストリーム(例えば、符号化ビデオストリーム402-1-o)で提供および多重化することができる。

30

**【0067】**

様々な実施形態では、多数の様々なビデオ復号器は符号化ビデオストリームの所与のビデオ階層から選択的にデジタルビデオ情報を復号化して、望ましいレベルの空間解像度および/または時間解像度を所与のメディア処理装置に提供することができる。例えば、ある種のビデオ復号器はベース階層をビデオストリームから復号化することができ、別の種類のビデオ復号器はベース階層および1つまたは複数の拡張階層をビデオストリームから復号化することができる。メディア処理装置は各ビデオ階層から復号化されたデジタルビデオ情報を様々な方法で組み合わせて、空間解像度および/または時間解像度の意味で様々なレベルのビデオ品質を提供することができる。メディア処理装置は次いで復号化したデジタルビデオ情報を、選択したレベルの空間解像度および時間解像度で1つまたは複数のディスプレイ上に再生することができる。

40

**【0068】**

スケラブルまたは複数解像度のビデオ符号器および復号器により、従来のビデオ符号器および復号器に対して幾つかの利点を与えることができる。例えば、単一のスケラブ

50

ルビデオ符号器および1つまたは複数の種類のビデオ復号器を用いて、様々な拡大または差分化したデジタルビデオサービスを提供することができる。レガシービデオ復号器は必ずしも拡張階層へのアクセスを有していなくてもデジタルビデオ情報をビデオストリームのベース階層から復号化することができ、拡張ビデオ復号器は同一のビデオストリーム内でベース階層および1つまたは複数の拡張階層の両方にアクセスすることができる。別の例では、異なる暗号化技術を各階層に使用し、それにより各階層へのアクセスを制御することができる。同様に、様々なデジタル権限を各階層に割り当てて各階層に対するアクセスを許可することができる。さらに別の例では、空間および/または時間解像度のレベルを、ビデオソースの種類、ビデオ圧縮技術の種類、通信リンクに対して利用可能な帯域幅またはプロトコル、所与の受信装置に対して利用可能な処理またはメモリリソース、デジタルビデオの再生に用いられるディスプレイ装置の種類、等に基づいて上げ下げすることができる。

#### 【0069】

特に、この改良した可変ビデオ符号化解像度の実装には、ディスプレイ解像度の次元をビデオストリーム内で指定するパラメータを持つという利点がある。ビデオの一部に対する符号化解像度はエントリポイントレベルで送信される。エントリポイントはビデオシーケンスの1つまたは複数の画像サブシーケンスまたは画像グループの近傍にあるかまたはそれらに隣接する。上記のビデオシーケンスは、イントラ符号化フレーム（「Iフレーム」とも呼ばれる）で始まり、そのイントラ符号化フレームに対して予測符号化した1つまたは複数の予測符号化フレーム（「Pフレーム」または「Bフレーム」とも呼ばれる）を含むこともできる。所与のエントリポイントで送信された符号化解像度はしたがって、ベース階層でIフレームを含み、そのIフレームを参照するPフレームまたはBフレームを含む画像グループに適用される。

#### 【0070】

様々な実施形態は、ビデオシーケンスの一部を異なる解像度で可変に符号化可能とする、改良した可変解像度符号化技術の実装を対象とすることができる。本技術は例えばビデオコーデックシステムに適用される。したがって、可変解像度符号化技術を、符号化ビットストリームシンタックスを利用する例示的なビデオ符号器/復号器の文脈で説明する。特に、改良した可変解像度符号化技術の実装の1つを、SMPTE標準421M（VC-1）ビデオコーデックシリーズの標準およびその変形のアドバンスドプロファイルに準拠するビデオコーデックにおいて説明する。あるいは、本技術を様々なビデオコーデック実装および標準に組み込むことができ、それらは後述する例示的なビデオコーデックおよびシンタックスとは細部で異なる可能性がある。

#### 【0071】

図6はビデオキャプチャおよび再生システム600に対する実装を示す。図6はビデオコーデックを採用するビデオキャプチャおよび再生システム600を示し、その中で様々な解像度符号化技術が一般的な適用または使用シナリオで実装される。ビデオキャプチャおよび再生システム600を、例えば会議サーバ102および/またはクライアント端末106-1-mのようなマルチメディア会議システム100内の多数の様々な要素で実装することができる。

#### 【0072】

ビデオキャプチャおよび再生システム600は一般に、ビデオコンテンツを入力デジタルビデオソース610からキャプチャして通信チャネル640上の圧縮ビデオビットストリームに符号化するビデオソース/符号器620と、ビデオをチャネルから受信および復号化してビデオディスプレイ670上でそのビデオを表示するビデオプレイヤー/復号器650とを含む。可変解像度符号化を有する後述のビデオコーデックを実装できる上記システムの幾つかの例には、ビデオキャプチャ、符号化、復号化および再生が全て単一のマシン内で実施されるシステム、およびこれらの動作が別個の、地理的に離れたマシン上で実施されるシステムが含まれる。例えば、デジタルビデオレコーダ、またはTVチューナカード付きパーソナルコンピュータは、ビデオ信号をキャプチャしてそのビデオをハード

ライブに符号化し、ビデオをハードドライブからモニタ上に読み戻し、復号化および表示することができる。別の例として、商業出版者またはビデオ放送者は、ビデオ送信（例えば、デジタル衛星チャンネル、またはWebビデオストリーム）を生成するためのビデオ符号器あるいは符号化ビデオを保持する記憶装置（例えば、テープまたはディスク）を取り込んだビデオマスタシステムを使用することができる。ビデオマスタシステムは、ビデオをユーザの復号器および再生マシン（例えば、パーソナルコンピュータ、ビデオプレイヤー、ビデオレシーバ、等）に配信するために使用される。

#### 【0073】

図示したシステム600では、ビデオソース/符号器620は情報源プリプロセッサ622、情報源圧縮符号器624、マルチプレクサ626およびチャンネル符号器628を含む。プリプロセッサ622は未圧縮のデジタルビデオを、ビデオカメラ、アナログテレビキャプチャ、または他の情報源のようなデジタルビデオソース610から受信し、そのビデオを処理して圧縮符号器624に入力する。圧縮符号器624は、図7を参照して説明されるビデオ符号器700の例であり、ビデオの圧縮および符号化を実施する。マルチプレクサ626は結果として生じる圧縮ビデオビットストリームをパケット化して、チャンネル符号器628に配信して通信チャンネル640上に符号化する。通信チャンネル640はデジタルテレビ放送、衛星または他の無線送信のようなビデオ送信、あるいはケーブル、電話または他の有線送信、等であることができる。通信チャンネル640はコンピュータハードドライブまたは他の記憶ディスクのような記録ビデオ媒体、テープ、光ディスク(DVD)または他の取り外し可能記録媒体であることもできる。チャンネル符号器628は圧縮ビデオビットストリームをファイルコンテナ、送信搬送波信号等に符号化する。

10

20

#### 【0074】

ビデオプレイヤー/復号器650では、チャンネル復号器652は圧縮ビデオビットストリームを通信チャンネル640上で復号化する。デマルチプレクサ654は圧縮ビデオビットストリームを逆多重化してチャンネル復号器から圧縮復号器656に配信する。圧縮復号器656の例は、図8を参照して説明するビデオ復号器800である。圧縮復号器は次いでビデオを圧縮ビデオビットストリームから復号化および再構築する。最後に、ポストプロセッサ658はビデオを処理してビデオディスプレイ670上に表示する。後処理動作の例には、デブロック、デリングまたは他のアーチファクト除去、レンジリマッピング、色変換および他の同様な動作が含まれる。

30

#### 【0075】

図7は一般化したビデオ符号器700のブロック図であり、図8は一般化したビデオ復号器800のブロック図であり、それらの符号器および復号器に可変解像度符号化技術を組み込むことができる。符号器および復号器内のモジュール間に示した関係は符号器および復号器内の情報の主な流れを示し、明確にするために他の関係は省略してある。特に、図7および8は通常、ビデオシーケンス、フレーム、マクロブロック、ブロック、等を使用される符号器設定、モード、テーブル、等を示す副情報は示さない。上記副情報を、一般にはその副情報のエントロピー符号化後に、出力ビットストリームで送信する。出力ビットストリームのフォーマットは例えば、SMPTE VC-1フォーマット、リアルタイム通信に適応されたSMPTE VC-1フォーマット、H.263フォーマット、H.264フォーマットまたは他のビデオフォーマットであることができる。

40

#### 【0076】

例えば、1つの実施形態では、符号器700および復号器800はブロックベースであり、各々が4つの8×8のルミナンスブロック（場合によって、1つの16×16のマクロブロックとして扱われる）と2つの8×8のクロミナンスブロックを含む4:2:0のマクロブロックフォーマットを使用する。あるいは、符号器700および復号器800はオブジェクトベースであり、異なるマクロブロックまたはブロックフォーマットを使用し、または8×8のブロックおよび16×16のマクロブロックとは異なるサイズまたは構成の画素集合上で動作を実施する。マクロブロックを使用してプログレッシブまたはインタレースビデオコンテンツのいずれかを表現することができる。

50

## 【 0 0 7 7 】

様々な実施形態におけるスケラブルビデオ符号化および復号化技術およびツールをビデオ符号器および/または復号器において実装することができる。ビデオ符号器および復号器はそれらの中に様々なモジュールを含むことができ、それらの様々なモジュールは多様な方法で、互いに関連および通信することができる。後述するモジュールおよび関係は例であり限定ではない。所望の実装および圧縮の種類により、ビデオ符号器またはビデオ復号器のモジュールを追加、省略、複数モジュールへ分割、他のモジュールと組み合わせ、および/または同様なモジュールで置換することができる。代替的な実施形態では、異なるモジュールおよび/または他のモジュール構成を有するビデオ符号器またはビデオ復号器により、1つまたは複数の説明した技術を実施することができる。

10

## 【 0 0 7 8 】

一般に、ビデオ圧縮技術にはフレーム内圧縮とフレーム間圧縮が含まれる。フレーム内圧縮技術は、一般にIフレーム、キーフレーム、または参照フレームと呼ばれる個々のフレームを圧縮する。フレーム間圧縮技術は、先行および/または後続フレームを参照してフレームを圧縮し、一般に予測フレームと呼ばれる。予測フレームの例には予測(P)フレーム、超予測(SP)フレーム、および両予測または両方向(B)フレームが含まれる。予測フレームは1つまたは複数の他のフレームからの動き補償予測(または差分)の点から表される。予測残差は予測したものと元のフレームとの間の差分である。対照的に、Iフレームまたはキーフレームは他のフレームを参照せずに圧縮される。

20

## 【 0 0 7 9 】

ビデオ符号器は一般に、カレントフレームを含むビデオフレームシーケンスを受信し、圧縮ビデオ情報を出力として生成する。符号器は予測フレームとキーフレームを圧縮する。符号器のコンポーネントの多くはキーフレームと予測フレームの両方を圧縮するために用いられる。これらのコンポーネントにより実施される厳密な動作は、圧縮されている情報の種類により様々であることができる。

## 【 0 0 8 0 】

図7は一般的なビデオ符号化システム700のブロック図である。符号化システム700は、カレントフレーム705を含むビデオフレームシーケンスを受信し、圧縮ビデオ情報795を出力として生成する。ビデオ符号器の特定の実施形態では一般に、一般化符号器700を変形または補完したものが用いられる。

30

## 【 0 0 8 1 】

符号化システム700は予測フレームとキーフレームを圧縮する。表現の便宜上、図7は符号化システム700を通るキーフレーム用の経路と前方予測フレーム用の経路とを示す。符号化システム700のコンポーネントの多くはキーフレームおよび予測フレームの両方を圧縮するために用いられる。これらのコンポーネントにより実施される厳密な動作は、圧縮されている情報の種類により様々であることができる。

## 【 0 0 8 2 】

予測フレーム(例えば、Pフレーム、SPフレーム、およびBフレーム)は1つまたは複数の他のフレームからの予測(または差分)の点から表わされる。予測残差は予測したものと元のフレームとの間の差分である。対照的に、キーフレーム(例えば、Iフレーム)は他のフレームを参照せずに圧縮される。

40

## 【 0 0 8 3 】

カレントフレーム705が前方予測フレームである場合、動き推定器710はカレントフレーム705のマクロブロックまたは他の画素集合(例えば、16×8、8×16または8×8のブロック)の動きを参照フレームに対して推定する。参照フレームは、フレーム記憶720にバッファされた再構築後の前フレーム725である。代替的な実施形態では、参照フレームは後続フレームであり、またはカレントフレームは両方向に予測される。動き推定器710は動きベクトルのような動き情報715を副情報として出力する。動き補償器730は動き情報715を再構築後の前フレーム725に適用して、動き補償されたカレントフレーム735を形成する。しかしながら、予測が完全であることは稀であ

50



り、動き補償されたカレントフレーム735と元のカレントフレーム705との間の差分が予測残差745である。あるいは、動き推定器および動き補償器では別の種類の動き推定/補償が適用される。

#### 【0084】

周波数変換器760は空間領域ビデオ情報を周波数領域(すなわち、空間)データに変換する。ブロックベースのビデオフレームに対して、周波数変換器760は、DCT(discrete cosine transform)と同様の特性を有する、以下のセクションで説明する変換を適用する。幾つかの実施形態では、周波数変換器760は周波数変換を、キーフレームに対する空間予測残差ブロックに適用する。周波数変換器760は8×8、8×4、4×8、または他のサイズの周波数変換を適用することができる。

10

#### 【0085】

量子化器770は次いでスペクトルデータ係数のブロックを量子化する。量子化器は均一なスカラ量子化を、フレーム毎または他の基準で変化するステップサイズでスペクトルデータに適用する。あるいは、量子化器は別の種類の量子化、例えば不均一量子化、ベクトル量子化、または非適応量子化をスペクトルデータ係数に適用するか、または周波数変換を使用しない符号化システムで空間領域データを直接量子化する。適応量子化に加えて、符号器700はフレーム除去、適応フィルタリング、またはレート制御向けの他の技術を使用することができる。

#### 【0086】

再構築されたカレントフレームが後続の動き推定/補償に必要であると、逆量子化器776が、量子化されたスペクトルデータ係数に対して逆量子化を実施する。逆周波数変換器766は次いで周波数変換器760の動作の逆を実施し、(予測フレームに対する)再構築された予測残差または再構築されたキーフレームを生成する。カレントフレーム705がキーフレームであった場合、再構築されたキーフレームを、再構築されたカレントフレームと捉える。カレントフレーム705が予測フレームであった場合、再構築された予測残差を動き補償されたカレントフレーム735に加えて再構築後のカレントフレームを形成する。フレーム記憶720は、次のフレームの予測に用いるため、再構築されたカレントフレームをバッファする。幾つかの実施形態では、符号器は再構築フレームにデブロッキングフィルタを適用して、フレームブロック内の不連続性を適応的に平滑化する。

20

#### 【0087】

エントロピー符号器780は量子化器770の出力およびある特定の副情報(例えば、動き情報715、量子化ステップサイズ)を圧縮する。一般的なエントロピー符号化技術には、算術符号化、差分符号化、ハフマン符号化、ランレングス符号化、LZ符号化、辞書符号化、およびそれらの組み合わせが含まれる。エントロピー符号器780は一般に様々な種類の情報(例えば、DC係数、AC係数、様々な種類の副情報)に対して様々な符号化技術を使用し、特定の符号化技術における複数の符号表から選択することができる。

30

#### 【0088】

エントロピー符号器780は圧縮ビデオ情報795をバッファ790に置く。バッファレベルインジケータをビットレート適応モジュールにフィードバックする。圧縮ビデオ情報795はバッファ790から一定のビットレートまたは相対的に一定のビットレートで消費され、そのビットレートで後のストリーミング向けに記憶される。あるいは、符号器700が圧縮ビデオ情報を圧縮直後に流す。

40

#### 【0089】

バッファ790の前後に、ネットワーク上で送信するために圧縮ビデオ情報795をチャンネル符号化することができる。チャンネル符号化は誤り検出および補正データを圧縮ビデオ情報795に適用することができる。

#### 【0090】

図8は一般的なビデオ復号化システム800のブロック図である。復号化システム800は圧縮されたビデオフレームシーケンスに対する情報895を受信し、再構築フレーム805を含む出力を生成する。ビデオ復号器の特定の実施形態では一般に、一般化復号器

50

800を変形または補完したものを使用する。

【0091】

復号化システム800は予測フレームおよびキーフレームを復元する。表現の便宜上、図8は復号化システム800を通るキーフレーム用の経路と前方予測フレーム用の経路とを示す。復号化システム800のコンポーネントの多くはキーフレームおよび予測フレームの両方を圧縮するために用いられる。これらのコンポーネントにより実施される厳密な動作は、圧縮されている情報の種類により様々であることができる。

【0092】

バッファ890は圧縮ビデオシーケンスに対する情報895を受信し、受信した情報をエントロピー復号器880が利用できるようにする。バッファ890は一般に、長時間に渡ってほぼ一定の速度で情報を受信し、帯域幅または送信における短期変動を平滑化するためのジッタバッファを含む。バッファ890は再生バッファおよび他のバッファも含むことができる。あるいは、バッファ890は情報を可変速度で受信する。バッファ890の前後に、誤り検出および補正のために圧縮ビデオ情報をチャンネル復号化および処理することができる。

【0093】

エントロピー復号器880は、一般に符号器で実施されたエントロピー符号化の逆を適用して、エントロピー符号化された量子化データおよびエントロピー符号化された副情報（例えば、動き情報、量子化ステップサイズ）を復号化する。エントロピー復号化技術には、算術復号化、差分復号化、ハフマン復号化、ランレングス復号化、LZ復号化、辞書復号化、およびそれらの組み合わせが含まれる。エントロピー復号器880は様々な種類の情報（例えば、DC係数、AC係数、様々な種類の副情報）に対して様々な復号化技術を使用し、特定の復号化技術における複数の符号表から選択することができる。

【0094】

再構築すべきフレーム805が前方予測フレームである場合、動き補償器830は動き情報815を参照フレーム825に適用して、再構築されているフレーム805の予測835を形成する。例えば、動き補償器830はマクロブロック動きベクトルを用いて参照フレーム825内のマクロブロックを見つける。予測835はしたがって、前に復号化したビデオフレームから動き補償された一連のビデオブロックである。フレームバッファ820は参照フレームとして用いるために、以前の再構築フレームを記憶する。あるいは、動き補償器は別の種類の動き補償を適用する。動き補償器による予測が完全であることは稀であり、したがって復号器800は予測残差も再構築する。

【0095】

復号器が後続の動き補償に対して再構築フレームを必要とすると、フレーム記憶820は次のフレームの予測に用いるためにその再構築フレームをバッファする。幾つかの実施形態では、符号器はデブロッキングフィルタを再構築フレームに適用してフレームブロック内の不連続性を適応的に平滑化する。

【0096】

逆量子化器870はエントロピー復号化データを逆量子化する。一般に、逆量子化器は均一なスカラ逆量子化を、フレーム毎または他の基準で変化するステップサイズでエントロピー復号化データに適用する。あるいは、逆量子化器は別の種類の逆量子化、例えば不均一量子化、ベクトル量子化、または非適応量子化をデータに適用するか、または逆周波数変換を使用しない復号化システムで空間領域データを直接逆量子化する。

【0097】

逆周波数変換器860は量子化された周波数領域データを空間領域ビデオ情報に変換する。ブロックベースのビデオフレームに対して、逆周波数変換器860は以下のセクションで説明する逆変換を適用する。幾つかの実施形態では、逆周波数変換器860は逆周波数変換を、キーフレームに対する空間予測残差ブロックに適用する。逆周波数変換器860は8×8、8×4、4×8、または他のサイズの逆周波数変換を適用することができる。

。

10

20

30

40

50

## 【 0 0 9 8 】

可変解像度符号化技術により復号器は所望のビデオディスプレイ解像度を維持することができ、一方で符号器は、1つまたは複数のビデオ部分を複数レベルの符号化解像度で符号化することを選択する柔軟性を有することができる。その符号化解像度は、ディスプレイ解像度とは異なってもよい。符号器はビデオシーケンスの一部の画像を低符号化解像度で符号化して低い符号化ビットレート、ディスプレイサイズまたはディスプレイ品質を実現することができる。低符号化解像度を使用したいときは、符号器は画像（単数または複数）を低解像度にフィルタおよびダウンサンプルする。復号化時、復号器はディスプレイ解像度で表示するためにこれらのビデオストリーム部分を低符号化解像度で選択的に復号化する。復号器は、大型画素のアドレス指定能力によりスクリーン上に表示する前に、低解像度のビデオをアップサンプルすることもできる。同様に、符号器はビデオシーケンスの一部の画像を高符号化解像度で符号化して高い符号化ビットレート、ディスプレイサイズまたはディスプレイ品質を実現することができる。高解像度符号化を使用したいときは、符号器フィルタにより元のビデオ解像度の大部分を維持する。これは一般に、高解像度のビデオとその高解像度ビデオのサイズにマッチするように補間された低解像度階層のビデオとの間の差分を表す追加の階層を符号化することで行われる。例えば、元のビデオは、それぞれ640画素および480画素の水平および垂直画素解像度を有することができる。符号化ベース階層は320×240画素を有することができる。第1の空間拡張階層は640×480画素の解像度を提供することができる。元のビデオを水平および垂直解像度に沿って2の倍数でダウンサンプルすることで、この空間拡張階層を得ることができる。640×480のビデオと、第1の拡張階層の640×480の解像度にマッチするように水平および垂直に2の倍数で補間した320×240のベース階層との間の差分を計算することで、上記の空間拡張階層を符号化する。復号化時は、ディスプレイの解像度に関わらず、復号器はディスプレイ解像度で表示するためにこれらのビデオストリーム部分をベース空間解像度符号化および高空間解像度符号化で選択的に復号化するか、またはビデオに相当な細部を提供する。

10

20

## 【 0 0 9 9 】

様々な実施形態では、ビデオ符号器700は可変解像度符号化をフレーム毎または他の基準で提供することができる。様々なレベルの解像度符号化を複数のビデオ階層の形で構成ことができ、各ビデオ階層は所与のビデオ情報集合に対して様々なレベルの空間解像度および/または時間解像度を提供する。例えば、ビデオ情報をベース階層および拡張階層でビデオストリームに符号化するようにビデオ符号器700を配置することができる。例えば、ビデオ情報は1つまたは複数のフレームシーケンス、フレーム、イメージ、画像、静止画、ブロック、マクロブロック、画素集合、または他の定義されたビデオデータ集合（集合的に「フレーム」と呼ぶ）を含むことができる。ベース階層は第1のレベルの空間解像度と第1のレベルの時間解像度を有することができる。拡張階層は第1のレベルの空間解像度、第1のレベルの時間解像度、またはその両方を増やすことができる。複数の拡張階層を置いて、所与のビデオ情報集合に対して空間解像度または時間解像度を増やす際に、所望のレベルの粒度を提供することができる。ビデオ階層を、図9を参照してより詳細に説明することができる。

30

40

## 【 0 1 0 0 】

図9はビデオ階層構造の例示的な実施形態を示す。図9はビデオストリームにおける符号化デジタルビデオの複数の独立ビデオ階層900を階層的に表している。図9に示すように、ビデオ階層900はベース階層（BL）を備えることができる。BLはベースレベルの空間解像度とベースレベルの時間解像度（例えば、フレームレート）のビデオストリームを表すことができる。例えば、1つの実施形態では、ベースレベルの時間解像度はT=15フレーム/秒としてTフレーム/秒を備えることができる。ビデオの符号化は、後続のBLビデオフレームの復号化が同一階層からの以前のビデオフレーム（例えば、ベース階層内の1つまたは複数のP、SPまたはBフレーム）にのみ左右されるようなものである。

50

## 【 0 1 0 1 】

ビデオ階層 9 0 0 は 1 つまたは複数の拡張階層も備えることができる。例えば、拡張階層は、第 1 の空間拡張階層 ( S L 0 )、第 2 の空間拡張階層 ( S L 1 )、および第 3 の空間拡張階層 ( S L 2 ) のような 1 つまたは複数の空間拡張階層を含むことができる。S L 0 は、B L に追加して高解像度ビデオを B L シーケンスと同一のフレームレート ( 例えば、15 フレーム / 秒 ) で提供することができる空間拡張階層を表す。S L 1 は、B L に追加して高解像度ビデオを B L シーケンスより高速の中間フレームレートで提供することができる空間拡張階層を表す。例えば、1 つの実施形態では、中間フレームレートは  $T / 2$  フレーム / 秒であることができ、 $T = 30$  フレーム / 秒である。S L 2 は、B L に追加して高解像度ビデオを B L シーケンスよりさらに高速の高フレームレートで提供することができる空間拡張階層である。例えば、1 つの実施形態では、高フレームレートは  $T$  フレーム / 秒であることができ、 $T = 60$  フレーム / 秒である。T に与えた値は例に過ぎず限定ではないことは理解できるだろう。

10

## 【 0 1 0 2 】

拡張階層は、第 1 の時間拡張階層 ( T L 1 ) および第 2 の時間拡張階層 ( T L 2 ) のような 1 つまたは複数の時間拡張階層を含むこともできる。T L 1 は、B L に追加して B L と同一の低解像度ビデオを生成することができるが、フレームレートは B L フレームに対して 2 倍のフレームレートであることができる空間拡張階層を表す。その結果、このシーケンスで動き描画が改善される。T L 2 は B L および T L 1 のフレームレートを 2 倍にする時間拡張階層を表す。このレベルでの動き描画は B L または T L 1 より良好である。

20

## 【 0 1 0 3 】

図 9 内の破線矢印で示すように、ベース階層および拡張階層を用いる多数の組み合わせが利用可能である。限定ではなく例として、幾つかの組み合わせとして以下の組み合わせを含めることができる。

- ・ B L
- ・ B L + S L 0
- ・ B L + T L 1
- ・ B L + T L 1 + T L 2
- ・ B L + S L 0 + T L 1 + S L 1
- ・ B L + S L 0 + T L 1 + S L 1 + T L 2 + S L 2

30

これらおよび他のビデオ階層の組み合わせにより、ビデオ品質が時間内で一貫していることを保証することができる。場合により、ビデオ品質が時間内で一貫するように、全ての時間階層に対して同一数の空間拡張階層を選択することが望ましいかもしれない。

## 【 0 1 0 4 】

以下でさらに十分に説明するように、符号器 7 0 0 は圧縮ビデオビットストリーム 7 9 5 内のシーケンスヘッダで最大解像度を指定する ( 図 7 )。コンテナファイルフォーマット、または送信キャリアフォーマットのヘッダ情報においてのようにビデオビットストリームのシーケンスヘッダで解像度符号化のレベルを符号化すると、ビットストリーム外部で保持されるヘッダ情報と比べて、最大解像度がビデオ復号器により直接復号化可能であるという利点がある。コンテナファイルまたは送信キャリア復号器 ( 例えば、チャンネル復号器 6 5 2 ) により最大解像度を別個にビデオ復号器に渡す必要はない。

40

## 【 0 1 0 5 】

符号器 7 0 0 はさらに、1 つまたは複数の画像グループがビデオビットストリーム内のエントリポイントの後に、エントリポイントヘッダ内の所定のフラグまたは開始符号を用いて低解像度で符号化されることを伝える。幾つかの実施形態では、フラグが低符号化解像度または高符号化解像度を示す場合、符号化サイズをエントリポイントヘッダで同様に符号化することもできる。

## 【 0 1 0 6 】

圧縮ビデオビットストリーム 7 9 5 ( 図 7 ) は圧縮プログレッシブビデオフレームまたは他の画像 ( 例えば、インタレースフレームまたはインタレースフィールドフォーマット

50

画像)のシーケンスに対する情報を含む。ビットストリーム795を、図8の復号器800のような復号器により復号化された幾つかの階層構造に構成する。最高階層はシーケンス階層であり、フレームシーケンス全体に対する情報を有する。さらに、各圧縮ビデオフレームは、3階層構造、すなわち(上から下へ)画像、マクロブロック、およびブロックに構造化されるデータから構成される。様々な解像度符号化技術を採用する代替的なビデオ実装では、様々なシンタックス要素構成を有する他のシンタックス構造を利用することができる。

【0107】

さらに、圧縮ビデオビットストリームは1つまたは複数のエントリポイントを含むことができる。ビットストリーム内の有効なエントリポイントは エレメンタリビットストリーム内の位置である。その位置から、メディア処理システムはビットストリーム内の先行情報(ビット)を全く必要とすることなくビットストリームを復号化または処理することができる。エントリポイントヘッダ(画像グループヘッダとも呼ばれる)は一般に、例えばビデオフレームの水平および垂直方向のサイズ、要求されたエレメンタリストリームバッファの状態および量子化パラメータのような、復号器の重要な初期化情報を含む。先行フレームを参照せずに復号化できるフレームは、独立フレームまたはキーフレームと呼ばれる。

10

【0108】

エントリポイントを、エントリポイントインジケータによりビットストリーム内で送信する。エントリポイントインジケータの目的は、例えば復号化を開始または再開するためのビットストリーム内特殊位置の存在を伝えることであり、エントリポイントインジケータの直後のビデオフレームを復号化する際は、過去に復号化されたビデオフィールドまたはフレームに左右されない。エントリポイントインジケータおよび関連エントリポイント構造をビットストリーム内に一定間隔または不規則間隔で挿入することができる。したがって、符号器は様々なポリシーを採用して、ビットストリームへのエントリポイントインジケータの挿入を管理することができる。エントリポイントインジケータおよび構造をビデオビットストリーム内の一定のフレーム位置に挿入することが一般的な動作であるが、幾つかのシナリオ(例えば、誤り回復または高速チャンネル変更)ではエントリポイント挿入の周期的性質を変更することができる。例えば、VC-1ビデオエレメンタリストリーム内のエントリポイント構造に関する以下の表1を参照されたい。

20

30

【0109】

【表 1】

表 1

アドバンスドプロファイルに対するエントリポイント階層ビットストリーム

エントリポイント階層 () {	ビット数	記述子
BROKEN_LINK	1	u i m s b f
CLOSED_ENTRY	1	u i m s b f
PANSCAN_FLAG	1	u i m s b f
REFDIST_FLAG	1	u i m s b f
LOOPFILTER	1	u i m s b f
FASTUVMC	1	u i m s b f
EXTENDED_MV	1	u i m s b f
DQUANT	2	u i m s b f
VSTRANSFORM	1	u i m s b f
OVERLAP	1	u i m s b f
QUANTIZER	2	u i m s b f
if (HRD_PARAM_FLAG == 1) {		
HRD_FULLNESS ()		
}		
CODED_SIZE_FLAG	1	u i m s b f
if (CODED_SIZE_FLAG == 1) {		
CODED_WIDTH	1 2	u i m s b f
CODED_HEIGHT	1 2	u i m s b f
}		
if (EXTENDED_MV == 1) {		
EXTENDED_DMV	1	u i m s b f
}		
RANGE_MAPY_FLAG	1	u i m s b f
if (RANGE_MAPY_FLAG == 1) {		
RANGE_MAPY	3	u i m s b f
}		
RANGE_MAPUV_FLAG	1	u i m s b f
if (RANGE_MAPUV_FLAG == 1)		
{		
RANGE_MAPUV	3	u i m s b f
}		
}		

10

20

30

40

50

## 【 0 1 1 0 】

様々な実施形態では、エントリポイントインジケータを所与の標準、プロトコルまたはアーキテクチャに従って定義することができる。場合により、エントリポイントインジケータを定義して所与の標準、プロトコルまたはアーキテクチャを拡張することができる。以下の表 1 および 2 では、様々なエントリポイントインジケータを、SMPTE 421 M (VC-1) ビットストリームに埋め込まれたビットストリームセグメントに適した開始符号サフィックスおよびその対応する意味として定義する。開始符号は、ベース階層および 1 つまたは複数の拡張階層のような様々なビデオ階層に対して様々な開始符号で一意的に識別可能であるべきである。しかしながら、開始符号ではビデオ階層間の同様な構造識別子を使用して解析および識別をより容易にすることができる。構造識別子の例にはシーケンスヘッダ、エントリポイントヘッダ、フレームヘッダ、フィールドヘッダ、スライスヘッダ、等を含めることができるがこれらに限らない。さらに、開始符号エミュレーション技術を利用して、所与のビデオ階層に対する開始符号がビデオストリーム内にランダム

に発生する確率を下げるができる。

【0111】

特定の開始符号によっては、各ビデオ階層に対する特定の構造パーサおよび復号器を起動または開始してビデオストリームからビデオ情報を復号化することができる。特定の構造パーサおよび復号器は、必要な参照フレーム、量子化器、速度制御、動き補償モード、等のような、所与のビデオ階層に適切な特定の連続の復号器ツールを実装することができる。本実施形態はこの文脈に限定されない。

【0112】

様々な実施形態では、開始符号サフィックスは最新のVC-1ビットストリームと下位互換であることができ、したがってレガシーVC-1復号器はVC-1ビットストリームが上記の新規セグメントを含む場合でも動作を継続できなければならない。開始符号サフィックスを使用して、SMPTE 421Mビデオビットストリームの最新フォーマットを拡張および構築してスケーラブルビデオ表現をサポートすることができる。

【0113】

【表2】

表2

開始符号サフィックス	意味
0x00	SMPTE予約済
0x01-0x09	SMPTE予約済
0x0A	シーケンスの終わり
0x0B	スライス
0x0C	フィールド
0x0D	フレーム
0x0E	エントリポイントヘッダ
0x0F	シーケンスヘッダ
0x10-0x1A	SMPTE予約済
0x1B	スライスレベルユーザデータ
0x1C	フィールドレベルユーザデータ
0x1D	フレームレベルユーザデータ
0x1E	エントリポイントレベルユーザデータ
0x1F	シーケンスレベルユーザデータ
0x20-0x7F	SMPTE予約済
0x80-0xFF	禁止

【0114】

表2に示す開始符号サフィックスを0x000001の3バイトシーケンスの末尾に追加して様々な開始符号を作成することができる。その開始符号をVC-1ビットストリームに統合することで、ビデオ復号器は自身がビットストリームのどの部分を解析しているかを判断することができる。例えば、シーケンス開始符号はVC-1ビットストリームにおけるシーケンスヘッダの出現を知らせる。開始符号に見えるビットシーケンスの出現を、もう開始符号をエミュレートしない数個のビットストリームに上記シーケンスを分割する開始符号エミュレーション防止により排除することができる。

【0115】

様々な実施形態では、さらなるビデオ階層を表すビットストリーム断片の追加を、ビットストリーム内の拡張階層断片の存在を識別および送信する新規開始符号を追加することで実現する。例えば、図9に示す2つの空間階層と3つの時間階層を用いて、以下の表3に示すように以下のサフィックスを割り当てて、様々な階層ビットストリームセグメントをそれらのサフィックスが保持する内容に関連付けて送信することができる。

【0116】

【表 3】

表 3

開始符号サフィックス	意味
0 x 0 0	SMPTE 予約済
0 x 0 1 - 0 x 0 9	SMPTE 予約済
0 x 0 A	シーケンスの終わり
0 x 0 B	スライス
0 x 0 C	フィールド
0 x 0 D	フレーム
0 x 0 E	エントリポイントヘッダ
0 x 0 F	シーケンスヘッダ
0 x 1 0 - 0 x 1 A	SMPTE 予約済
0 x 1 B	スライスレベルユーザデータ
0 x 1 C	フィールドレベルユーザデータ
0 x 1 D	フレームレベルユーザデータ
0 x 1 E	エントリポイントレベルユーザデータ
0 x 1 F	シーケンスレベルユーザデータ
0 x 2 0	スライスレベル - S L 0
0 x 2 1	スライスレベル - T L 1
0 x 2 2	スライスレベル - S L 1
0 x 2 3	スライスレベル - T L 2
0 x 2 4	スライスレベル - S L 2
0 x 3 0	フィールドレベル - S L 0
0 x 3 1	フィールドレベル - T L 1
0 x 3 2	フィールドレベル - S L 1
0 x 3 3	フィールドレベル - T L 2
0 x 3 4	フィールドレベル - S L 2
0 x 4 0	フィールドレベル - S L 0
0 x 4 1	フィールドレベル - T L 1
0 x 4 2	フィールドレベル - S L 1
0 x 4 3	フィールドレベル - T L 2
0 x 4 4	フィールドレベル - S L 2
0 x 5 0	エントリポイントレベル - S L 0
0 x 5 1	エントリポイントレベル - T L 1
0 x 5 2	エントリポイントレベル - S L 1
0 x 5 3	エントリポイントレベル - T L 2
0 x 5 4	エントリポイントレベル - S L 2
0 x 6 0	シーケンスレベル - S L 0
0 x 6 1	シーケンスレベル - T L 1
0 x 6 2	シーケンスレベル - S L 1
0 x 6 3	シーケンスレベル - T L 2
0 x 6 4	シーケンスレベル - S L 2
0 x 8 0 - 0 x F F	禁止

10

20

30

40

## 【 0 1 1 7 】

断片の挿入については一連の所定の範囲規則に従うべきである。例えば、シーケンスレベル S L 0 の情報はシーケンスレベル B L の情報に従うべきである、等である。

50



## 【 0 1 1 8 】

適応スケジューリング

図 1 0 はビデオのスケラビリティを使用してマルチメディア会議システム 1 0 0 のような R T C ( R e a l T I m e C o m m u n i c a t i o n ) システムにおける遅延を改善するための、適応スケジューリング技術に関する第 1 の論理図である。新たな主要またはアクティブな話者が例えば送信クライアント端末 1 0 6 - 1 から話し始めると、受信クライアント端末 1 0 6 - 2、1 0 6 - 3 がその新たな主要話者でディスプレイオブジェクトを描画し始めることができるように、会議サーバ 1 0 2 は新たなビデオキーフレームに対するキーフレーム要求 1 0 0 2 を送信することができる。しかしながら、キーフレームは相対的に大きく、したがって他のビデオフレームと比較して送信時間が非常に長い。その結果、ビデオ遅延がより大きく、その新たな主要話者を参加者が見るのに数秒かかる。

10

## 【 0 1 1 9 】

様々な実施形態では、適応スケジューリングモジュール 1 0 0 4 を用いてこれらおよび他の問題を解決することができる。送信クライアント端末 1 0 6 - 1 のような別の装置の代わりに時間内にビデオ階層の送信を適応的にスケジューリングできるように、適応スケジューリングモジュール 1 0 0 4 を配置することができる。その結果、主要またはアクティブな話者が話し始め、自身のビデオを送信し始める際の応答時間を改善することができる。低ビデオ階層を最初に送信し、追加の階層を徐々に送信して、長時間に渡ってビデオ品質を改善する。このように、新たな主要話者が話し始める際にスムーズに遷移し、それによりディスプレイオブジェクト内のスイッチを起動して新たな主要話者のビデオ情報を示すような視覚的構成を描画することができる。ビデオ階層の送信を適応的にスケジューリングすることで、ちらつき、ブランキング、および視覚的構成内で主要話者と対応するディスプレイオブジェクトとの間を遷移することにより発生する他の副作用を軽減することができる。

20

## 【 0 1 2 0 】

様々な実施形態では、送信クライアント端末 1 0 6 - 1 をスケラブルビデオ符号器 1 0 4 で実装して複数のビデオ階層を会議サーバ 1 0 2 に送信することができる。新たな主要話者が現れた際に、送信クライアント端末 1 0 6 - 1 を介してベース階層のような低ビデオ階層を最初に受信クライアント端末 1 0 6 - 2、1 0 6 - 3 に送信するように会議サーバ 1 0 2 の適応スケジューリングモジュール 1 0 0 4 を配置することができる。適応スケジューリングモジュール 1 0 0 4 は、1 つまたは複数の拡張階層を受信クライアント端末 1 0 6 - 2、1 0 6 - 3 に送信することで、ビデオストリーム内のフレームレートおよび/または空間解像度を徐々に増やすことができる。受信クライアント端末 1 0 6 - 2、1 0 6 - 3 はそれにより、相対的に迅速に送信される低ビデオ階層を最初に取得する。時間とともに、所与の通信リンク 1 0 8 上で利用可能な帯域幅によっては、受信クライアント端末 1 0 6 - 2、1 0 6 - 3 は高フレームレートおよび/または高解像度のコンテンツを取得することができる。

30

## 【 0 1 2 1 】

図 1 0 に示すように、音声会議で新たな主要話者に切り替える際、会議サーバ 1 0 2 ( または A V M C U ) は、全ての参加者がその話者からビデオを受信し始めることができるように、新規のキーフレームに対する要求を送信する。完全なキーフレームを全ての階層に送信するのと対照的に、スケラブルビデオ符号器 1 0 4 はビデオのベース階層のみを会議サーバ 1 0 2 に送信する。その結果、ビデオを送信および復号化する時間は大幅に削減され、それにより参加者が新たな主要話者を見るのに要する時間が削減される。

40

## 【 0 1 2 2 】

図 1 1 はビデオのスケラビリティを使用してマルチメディア会議システム 1 0 0 における遅延を改善するための適応スケジューリング技術の第 2 の論理図を示す。図 1 1 に示すように、適応スケジューリングモジュール 1 0 0 4 が低ビデオ階層を送信した直後に、適応スケジューリングモジュール 1 0 0 4 は漸進的にさらなる拡張階層を送信して高フレ

50

ームレートおよび/または高解像度を提供する。図 1 1 は、2つの時間スケールを送信してビデオ情報における動き描画を改善する場合を示す。

【0123】

図 1 2 はビデオのスケラビリティを使用してマルチメディア会議システム 1 0 0 における遅延を改善するための適応スケジューリング技術の第 3 の論理図を示す。図 1 2 に示すように、いったん受信クライアント端末 1 0 6 - 2、1 0 6 - 3 で安定状態が実現されると、適応スケジューリングモジュール 1 0 0 4 は全ての残存するビデオ階層を受信クライアント端末 1 0 6 - 2、1 0 6 - 3 に送信することができる。適応スケジューリングモジュール 1 0 0 4 は、それぞれ通信リンク 1 0 8 - 2、1 0 8 - 3 上でメディアチャネルが利用できる利用可能帯域幅によって、送信するビデオ階層の数を決定する。図 1 2 に示すように、受信クライアント端末 1 0 6 - 2 は帯域幅が十分に広くないので第 2 の空間および時間階層は受信できず、受信クライアント端末 1 0 6 - 3 は通信の帯域幅が制限されているのでベース階層および第 1 の時間スケールのみを受信することができる。

10

【0124】

図 1 0 ~ 1 2 に示すように、適応スケジューリング技術を主要話者間のスムーズな遷移に使用することができる。例えば、1つの実施形態では、受信機は様々なビデオ階層を有するビデオストリーム内の符号化ビデオ情報を受信することができる。上記の様々なビデオ階層には、第 1 のレベルの空間解像度および第 1 のレベルの時間解像度を有するベース階層と、その第 1 のレベルの空間解像度またはその第 1 のレベルの時間解像度を増やす拡張階層とが含まれる。適応スケジューリングモジュール 1 0 0 4 を受信機と通信可能に接続することができる。様々なビデオ階層を様々な時点で受信クライアント端末に送信するように配置することができる。本実施形態はこの文脈に限定されない。

20

【0125】

例えば、1つの実施形態では、ベース階層の送信後まで拡張階層の送信を遅らせるように適応スケジューリングモジュール 1 0 0 4 を配置することができる。本実施形態はこの文脈に限定されない。

【0126】

例えば、1つの実施形態では、新たな主要話者を決定する際にベース階層を第 1 の時刻に受信クライアント端末に送信するように適応スケジューリングモジュール 1 0 0 4 を配置することができる。例えば、新たな主要話者の決定を、主要話者検出器を会議サーバ 1 0 2 で用いるといった任意数の異なる方法で実現することができる。本実施形態はこの文脈に限定されない。

30

【0127】

例えば、1つの実施形態では、新たな主要話者を決定する際にベース階層を第 1 の時刻に受信クライアント端末に送信し、拡張階層を所定の待機時間後の第 2 の時刻に送信するように適応スケジューリングモジュール 1 0 0 4 を配置することができる。所定の待機時間は、所与の実装に望ましいように受信クライアント端末が主要話者間および視覚的構成に対するディスプレイオブジェクト間でスムーズな遷移を実現するのに十分な任意の時間であることができる。一般に、所定の待機時間に、受信クライアント端末が所望のレベルのビデオ解像度および品質で安定状態に安定化および到達するのに十分な時間を含めることができる。本実施形態はこの文脈に限定されない。

40

【0128】

例えば、1つの実施形態では、適応スケジューリングモジュール 1 0 0 4 を送信機と通信可能に接続することができる。送信機を使用してベース階層と拡張階層を異なる予定時刻に送信することができる。本実施形態はこの文脈に限定されない。

【0129】

本明細書では多数の特定の詳細を説明して実施形態の徹底的な理解を提供している。しかしながら当業者は、本実施形態をこれらの特定の詳細なしに実践できることを理解するであろう。他の事例では、本実施形態を不明瞭にしないように公知の動作、コンポーネントおよび回路は詳細に説明していない。本明細書で開示した特定の構造的および機能的詳

50

細は代表的なものであり、本実施形態の範囲を必ずしも限定しないことが理解できるだろう。

【0130】

「1つの実施形態」または「実施形態」への任意の言及は、その実施形態に関連して説明した特定の特徴、構造、または特性が少なくとも1つの実施形態に含まれることを意味することにも留意されたい。本明細書の様々な箇所で見られる句「1つの実施形態では」は必ずしも全て同一の実施形態を指していない。

【0131】

幾つかの実施形態を、「接続された(coupled)」および「接続された(connected)」という表現をそれらの派生語とともに用いて説明することができる。これらの用語は互いに対する同義語を意味しているわけではないことは理解されるべきである。例えば、幾つかの実施形態を用語「接続された(connected)」を用いて説明して、2つ以上の要素が互いに直接、物理的または電氣的に接していることを示すことができる。別の例では、幾つかの実施形態を用語「接続された(coupled)」を用いて説明して、2つ以上の要素が互いに直接、物理的または電氣的に接していることを示すことができる。しかしながら、用語「接続された(coupled)」は、2つ以上の要素が互いに直接は接していないが、互いになお協調または相互作用することをも意味することもできる。本実施形態はこの文脈に限定されない。

10

【0132】

例えば、幾つかの実施形態を、機械により実行した場合に、本実施形態に従ってその機械に方法および/または動作を実施させることが可能な命令または命令集合を記憶できる機械読取可能媒体または製品を用いて実装することができる。上記の機械は例えば、任意の適切な処理プラットフォーム、コンピューティングプラットフォーム、コンピューティング装置、コンピューティング装置、コンピューティングシステム、処理システム、コンピュータ、プロセッサ、等を含むことができ、ハードウェアおよび/またはソフトウェアの任意の適切な組み合わせを用いて実装することができる。機械読取可能媒体または製品は例えば、任意の適切な種類のメモリユニット、メモリ装置、メモリ製品、メモリ媒体、記憶装置、記憶製品、記憶媒体および/または記憶ユニット、例えばメモリ、取り外し可能または取り外し不能媒体、消去可能または消去不能媒体、書き込み可能または再書き込み可能媒体、デジタルまたはアナログ媒体、ハードディスク、フロッピーディスク(登録商標)、CD-ROM、CD-R、CD-RW、光ディスク、磁気媒体、光磁気媒体、取り外し可能メモリカードまたはディスク、様々な種類のDVD、テープ、カセット、等を含むことができる。

20

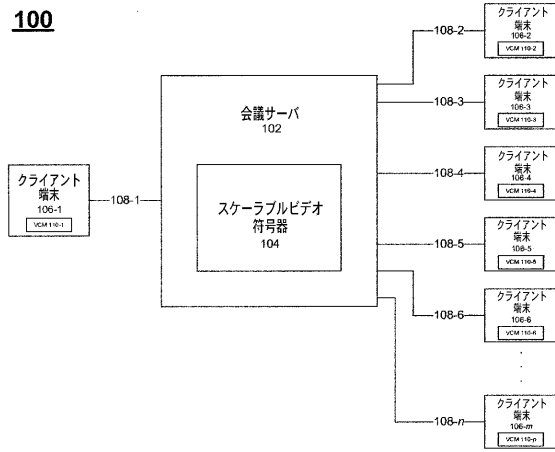
30

【0133】

本主題を構造的特徴および/または方法論的動作に固有な言葉で説明したが、添付の請求項で定義した本主題は必ずしも上述の特定の特徴または動作に限定されないことは理解されるべきである。むしろ、上述の特定の特徴および動作は請求項を実装する形態の例として開示してある。

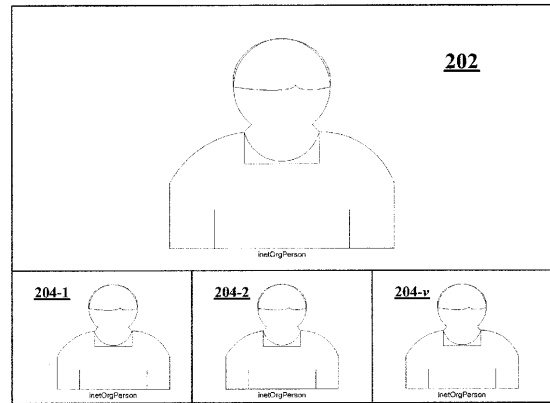
【 図 1 】

100

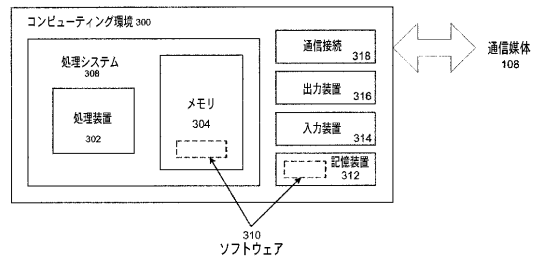


【 図 2 】

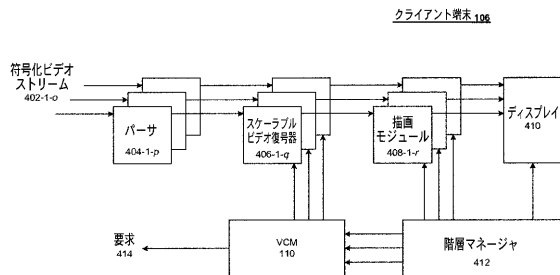
200



【 図 3 】

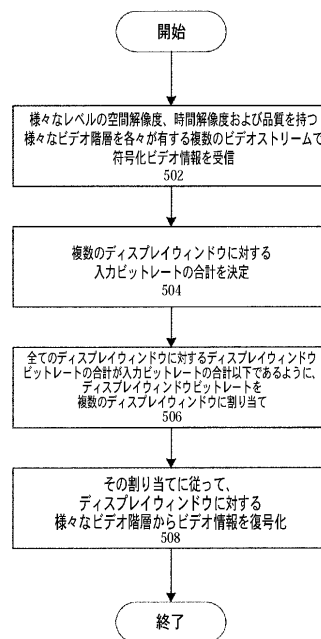


【 図 4 】

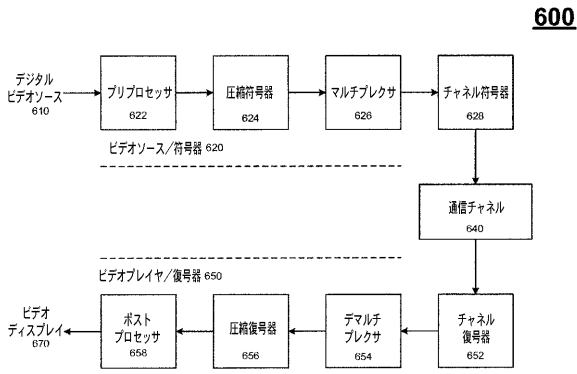


【 図 5 】

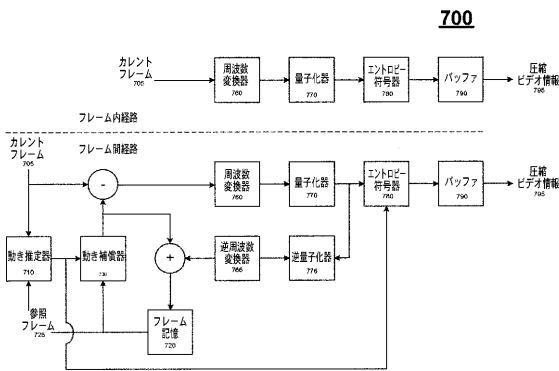
500



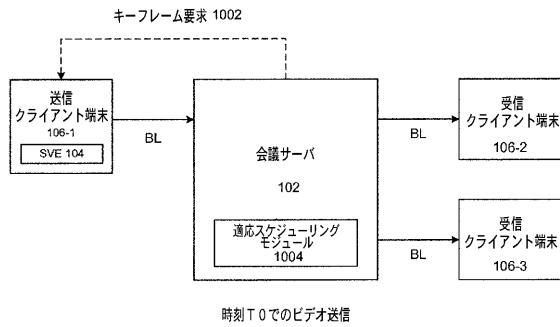
【 図 6 】



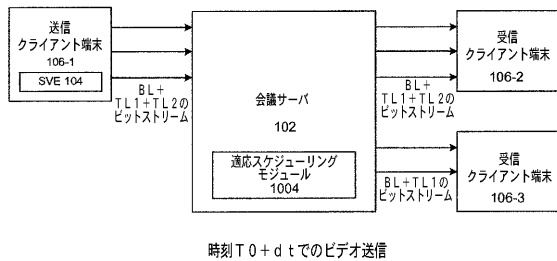
【 図 7 】



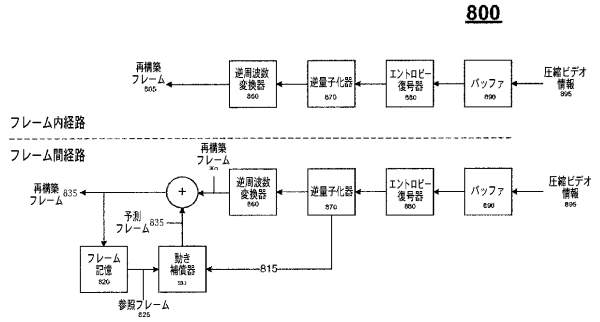
【 図 10 】



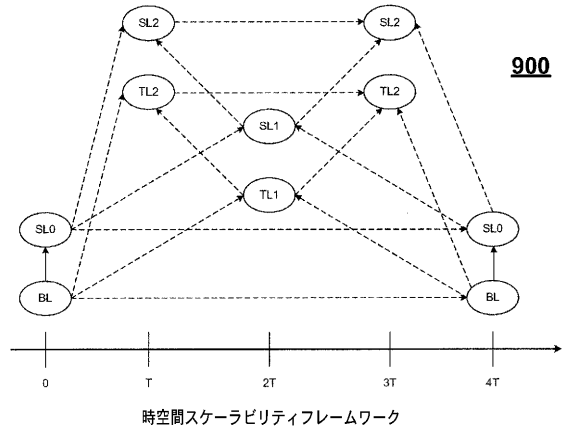
【 図 11 】



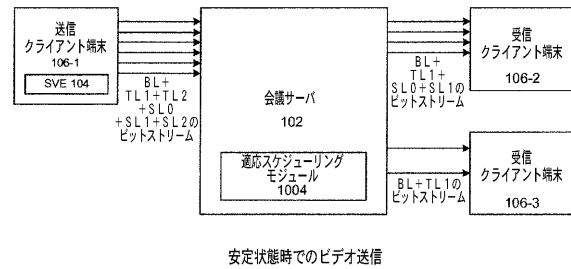
【 図 8 】





【 図 9 】



【 図 12 】



## 【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. <b>PCT/US2007/075963</b>
<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
<b>H04M 3/56(2006.01)i, H04N 7/15(2006.01)i</b>		
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) IPC 8 : H04M 3/56, H04N 7/15		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Korean Utility models and applications for Utility models since 1975 Japanese Utility models and applications for Utility models since 1975		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) eKIPASS(KIPO internal) "multimedia", "conference", "multiple video streams", "encode", "decode"		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 7007098 B1 (SMYTH, JOSEPH et al.) 28 February 2006 See the abstract; claims 1-37; figure 1.	1-20
A	US 6075571 A (KUTHYAR, A. K. et al.) 13 June 2000 See the abstract; claims 1-18; figure 2.	1-20
A	US 6104705 A (ISMAIL, NERMEEN et al.) 15 August 2000 See the abstract; claims 1-11; figure 4.	1-20
A	US 5751791 A (CHEN, MON-MEI et al.) 12 May 1998 See the abstract; claims 1-12; figure 5.	1-20
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
<p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&amp;" document member of the same patent family</p>		
Date of the actual completion of the international search 14 DECEMBER 2007 (14.12.2007)		Date of mailing of the international search report <b>14 DECEMBER 2007 (14.12.2007)</b>
Name and mailing address of the ISA/KR  Korean Intellectual Property Office 920 Dunsan-dong, Seo-gu, Daejeon 302-701, Republic of Korea Facsimile No. 82-42-472-7140		Authorized officer SONG, Hyun Chae Telephone No. 82-42-481-5786 

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2007/075963

**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2.  Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3.  Claims Nos.: 4-8, 12-13, and 17-19  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
  
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
  
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

**INTERNATIONAL SEARCH REPORT**  
Information on patent family membersInternational application No.  
**PCT/US2007/075963**

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 7007098 B1	28.02.2006	None	
US 6075571 A	13.06.2000	None	
US 6104705 A	15.08.2000	EP 0963658 A2 JP 2001516546 A WO 9935852 A2 WO 9935852 A3	15.12.1999 25.09.2001 15.07.1999 16.09.1999
US 5751791 A	12.05.1998	None	



## フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, 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, IS, IT, LT, LU, LV, MC, MT, NL, PL, 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, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(72)発明者 フィリップ エー . チョウ

アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ  
マイクロソフト コーポレーション インターナショナル パテント内

(72)発明者 リージス ジェイ . クリノン

アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ  
マイクロソフト コーポレーション インターナショナル パテント内

(72)発明者 ティム ムーア

アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ  
マイクロソフト コーポレーション インターナショナル パテント内

Fターム(参考) 5C159 LB05 LB07 MA05 MA14 MA32 MA33 PP05 PP06 PP07 SS07

TA07 TA45 TA60 TC21 UA02

5C164 FA10 PA32 PA33 SB02S SC03P VA13P VA24P