(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2005-500721
(P2005-500721A)
(43)公表日　平成17年1月6日(2005.1.6)

(51)Int.Cl.[7]	F I		テーマコード（参考）
H04N 13/00	H04N 13/00		5C061
G09G 5/00	G09G 5/00	510Z	5C082
G09G 5/393	G09G 5/36	530E	
G09G 5/395	G09G 5/36	530F	

審査請求 未請求　予備審査請求 有　（全 105 頁）

(21)出願番号　特願2003-508064(P2003-508064)
(86)(22)出願日　平成13年12月21日(2001.12.21)
(85)翻訳文提出日　平成15年12月25日(2003.12.25)
(86)国際出願番号　PCT/US2001/049287
(87)国際公開番号　WO2003/001803
(87)国際公開日　平成15年1月3日(2003.1.3)
(31)優先権主張番号　09/891,733
(32)優先日　平成13年6月25日(2001.6.25)
(33)優先権主張国　米国(US)

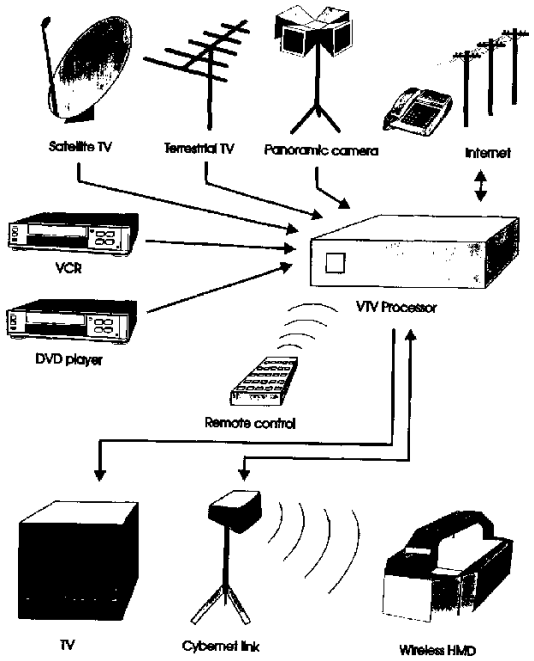
(71)出願人　504000018
アンガス　リチャーズ
アメリカ合衆国　カリフォルニア州　９００６６，ロサンゼルス，ケリー　ストリート　５０１６
(74)代理人　100089635
弁理士　清水　守
(72)発明者　アンガス　リチャーズ
アメリカ合衆国　カリフォルニア州　９００６６，ロサンゼルス，ケリー　ストリート　５０１６

Fターム(参考)　5C061 AB01 AB08 AB14
5C082 AA37 BA12 BA46 BB02 BB03 BB15 BB22 BD07 CA81 DA53 DA63 DA86 DA89 MM07

最終頁に続く

(54)【発明の名称】ＶＴＶシステム

(57)【要約】
拡張型立体テレビジョン様の視聴体験を生みだすハードウェア構成全体が開示される（図１）。本システムの場合、通常のテレビとは異なり、視聴者は映画の動きと関連して視聴者のビューイング方向と相対位置との制御（遠隔制御）を行うことが可能である（図６）。特定のハードウェア構成（Ｓ０、Ｓ１、Ｓ２、Ｓ３）に加えて、本発明はまた、このバーチャルリアリティ様の体験を可能にする新しいビデオフォーマット（無線HMD）にも関するものである。



【特許請求の範囲】
【請求項１】
ソースメディアを供給する従来方式のビデオ装置を利用して拡張型立体テレビジョン様の視聴体験を生み出す方法において、
前記ソースメディアを供給するための従来方式のビデオ装置を利用して、ページベースで画像球面全体をリアルタイムで更新するステップと、
複数のさらに小さな画像のサブセクションから構成されるパノラマ画像を描くグラフィック画像を生成するステップとを有することを特徴とする方法。
【請求項２】
請求項１記載の方法において、前記パノラマ画像が３６０°の水平視野であることを特徴とする方法。
【請求項３】
請求項１記載の方法において、上記ソースメディア内に存在する付加情報によりページ順を確定するステップをさらに有することを特徴とする方法。
【請求項４】
請求項１記載の方法において、視聴者が、バーチャルスクリーンの幅を水平方向に最大３６０°及び垂直方向に１８０°まで拡張できるワイドスクリーンフォーマットで、予め記録された視聴覚メディアを視ることができることを特徴とする方法。
【請求項５】
請求項１記載の方法において、視聴者が、ＴＶ、投影型ＴＶまたはコンピュータ画面を含む従来方式のスクリーンをベースとする表示装置で、予め録画された視聴覚マテリアルを視ることができ、前記表示装置が上記３６０°のフルパノラマ画像のビューポートまたはサブセットを表示するように為すことを特徴とする方法。
【請求項６】
娯楽システムにおいて、
（ＶＣＲ、ＤＶＤプレーヤ、衛星受信機などのような）別のメディア供給装置と、
上記メディア供給装置から供給されるビデオデータからパノラマビデオ画像を生成する電子装置（ＶＴＶプロセッサ）と、
従来方式の平面スクリーン型テレビや、ヘルメット装着型表示装置（ＨＭＤ）とＶＴＶプロセッサ間で上記ビデオ情報を通信する追加的な無線データ通信ネットワークと共に、追加的な単一の視聴装置またはパノラマビデオ取得装置を備えたＨＭＤあるいはその他のバーチャルリアリティ表示装置を含む表示装置と、を有することを特徴とする娯楽システム。
【請求項７】
視聴覚効果を生成する方法において、
現行のテレビジョン規格と互換性を保ちながら、従来のテレビジョン規格を改変して新しい視聴覚規格を形成するステップを有し、該新しい視聴覚規格が異なるフレームを処理し、前記フレームが画像データ、サウンド情報、制御情報またはこれらの組み合わせを含むことを特徴とする方法。
【請求項８】
請求項７記載の方法において、上記新しい視聴覚規格が、ＶＴＶグラフィックプロセッサ用の制御用パラメータと画像処理データとを定義する情報を提供するための、追加のデジタルコード化データ、アナログコード化データまたはこれらのデータの組み合わせを標準的ビデオ画像の少なくとも１つの走査線内に含むことを特徴とする方法。
【請求項９】
請求項８記載の方法において、上記単複の走査線が、４またはそれ以上のオーディオトラックをリアルタイムで生成するための情報を提供する追加のデジタルデータとアナログコード化データとをさらに含むことを特徴とする方法。
【請求項１０】
時分割多重化と共に周波数圧縮と伸長の合成（コンパウンディング）を行って、標準的な

２トラック視聴覚録音システム上へ／から４またはそれ以上のオーディオトラックの符号化及びその後の復号化を行う方法において、上記視聴覚データストリームのビデオ成分（Ｖ同期、Ｈ同期、奇数／偶数など）の中に存在する同期信号を利用して、遠隔クロックの生成および／または同期を行って、時分割多重化済み視聴覚データの符号化及びその後の復号化を行うことを特徴とする方法。

【請求項１１】

ソースメディアを供給する従来方式のビデオ装置を利用して拡張型立体テレビジョン様の視聴体験を形成する電子装置。

【請求項１２】

上記“バーチャル”スクリーンの幅が水平方向に最大３６０°及び垂直方向に１８０°まで拡張できるようなワイドスクリーン（パノラマ）フォーマットで、視聴者がライブの視聴覚メディアまたは予め録画された視聴覚メディアを視ることが可能となることを特徴とする電子装置。

【請求項１３】

パノラマ画像を描くグラフィック画像を形成して、３６０°のワイドパノラマの場合に上記パノラマ画像全体（“画像球面”）が、複数のさらに小さな画像サブセクション（“ページ”）から構成されるように為すことを特徴とする電子装置。

【請求項１４】

グラフィック画像を形成する、請求項１３記載の電子装置において、上記ソースメディアを供給する従来方式のビデオ装置を利用して、ページベースで上記画像球面全体をリアルタイムで更新するように為すことを特徴とする電子装置。

【請求項１５】

請求項１３記載の電子装置において、上記ソースメディア内に存在する付加情報によりページ順を確定することを特徴とする電子装置。

【請求項１６】

請求項１２記載の電子装置において、視聴者が、従来方式のスクリーンをベースとする表示装置（ＴＶ、投影型ＴＶ、コンピュータ画面）で生のマテリアルまたは予め録画された視聴覚マテリアルを視ることができ、上記表示装置がフルパノラマ画像のビューポートまたはサブセットを表示するように為すことを特徴とする電子装置。

【請求項１７】

請求項１２記載の電子装置において、視聴者が、ヘルメット装着型表示装置（ＨＭＤ）またはその他のバーチャルリアリティ型表示装置を用いて、生の視聴覚マテリアルまたは予め録画された視聴覚マテリアルを視ることができ、上記表示装置が上記フルパノラマ画像のビューポートまたはサブセットを表示するように為すことを特徴とする電子装置。

【請求項１８】

（ＶＴＶ画像規格）

多様な異なる“フレーム”を処理する新しい視聴覚規格であって、現行のテレビジョン規格に対応する改変から構成される新しい視聴覚規格（バーチャルテレビ（ＶＴＶ）規格）において、現行のテレビジョン規格（ＮＴＳＣ、ＰＡＬ、ＨＤＴＶなど）とそのまま互換性を保ちながら、上記フレームが画像データ、サウンド情報、制御情報またはこれらの組み合わせを含むことができるように為すことを特徴とする新しい視聴覚規格。

【請求項１９】

請求項１８記載の新しい視聴覚規格において、標準的ビデオ画像の１以上の走査線内に、上記ＶＴＶグラフィックプロセッサ用の制御用パラメータと画像処理用データとを定義する情報を提供する追加のデジタルコード化データおよび／またはアナログコード化データを含むことを特徴とする新しい視聴覚規格。

【請求項２０】

請求項１８記載の新しい視聴覚規格において、３６０°の水平方向×Ｘ°の垂直方向の円筒形または先端を切り落とした球面である“画像球面”というコンセプトを含み、各画像球面が複数のサブセクションや“ページ”から構成されるように為すことを特徴とする新

しい視聴覚規格。
【請求項２１】
請求項２０記載の新しい視聴覚規格において、標準的ビデオ画像の１以上の走査線内に、３６０°の水平方向×Ｘ°の垂直方向の“画像球面”の範囲内の現在のページ（ビデオフィールドまたはフレーム）の相対的配置位置に関する情報を提供する追加のデジタルコード化データまたはアナログコード化データを含むことを特徴とする新しい視聴覚規格。
【請求項２２】
請求項１８記載の新しい視聴覚規格において、４またはそれ以上のオーディオトラックをリアルタイムで生成するための情報を提供する追加のデジタルコード化データとアナログコードデータ（ハイブリッドコード化データ）とを、標準的ビデオ画像の１以上の走査線内にさらに含むことを特徴とする新しい視聴覚規格。
【請求項２３】
請求項１８記載の新しい視聴覚規格において、標準的な２トラック視聴覚録音システム上へ４またはそれ以上のオーディオトラックの符号化及びその後の復号化を行うために、時分割多重化と共に周波数圧縮と伸長の合成（コンパウンディング）を行うことを特徴とする新しい視聴覚規格。
【請求項２４】
請求項２３記載の新しい視聴覚規格において、上記視聴覚データストリームのビデオ成分（Ｖ同期、Ｈ同期、奇数／偶数など）の中に存在する同期信号のうちの１以上の信号を利用して、遠隔クロックの生成および／または同期を行って時分割多重化済み視聴覚データの符号化及びその後の復号化を行うことを特徴とする新しい視聴覚規格。
【請求項２５】
請求項１８記載の新しい視聴覚規格であって、標準的ビデオ画像の１以上の走査線内に、オーディオトラック数と、オーディオデータフォーマットと、オーディオサンプリングレートと、トラック同期とに関する情報を提供する追加のデジタルコード化データまたはアナログコード化データを含む視聴覚規格において、上記ＶＴＶグラフィックプロセッサが、請求項２２と２３に記載のように上記オーディオ情報を復号化して、（位置と方角に敏感な）空間サウンドに変えることができることを特徴とする新しい視聴覚規格。
【請求項２６】
請求項１８記載の新しい視聴覚規格において、標準的ビデオ記憶システム及び伝送システムと共に使用するためのマルチトラックオーディオの符号化を可能にし、その後、専用ハードウェア（ＶＴＶプロセッサ）によりこの情報を復号化して、（ヘッドホンやスピーカシステム用の）左右のオーディオチャネルを形成し、ユーザの左右の耳に対して空間的に正しい３Ｄサウンドを生みだすように、上記オーディオチャネルをミックスする（数学的に合成する）ことができるようにし、上記数学的合成に影響を与える前記パラメータは、（４トラックオーディオシステムの場合）主として視聴者またはビューポートの方位角であり、（８トラックオーディオシステムの場合）視聴者またはビューポートの方位角と高度方位角の双方であることを特徴とする新しい視聴覚規格。
【請求項２７】
請求項１８記載の新しい視聴覚規格において、標準的ビデオ画像の１以上の走査線内に、上記画像を撮影したカメラの絶対方角（方位角または方位角と高度及び追加として回転）並びに追加的に前記カメラのＸ、Ｙ、Ｚ位置に関する情報を提供する追加のデジタルコード化データまたはアナログコード化データを含むことを特徴とする新しい視聴覚規格。（ＶＴＶプロセッサ）
【請求項２８】
請求項１１乃至１７記載の電子装置において、上記パノラマの範囲内の上記ビューポートの方位角または方位角と高度の双方を（画像を表示しながら）実行時にユーザが動的に制御することが可能であることを特徴とする電子装置。
【請求項２９】
請求項１１乃至１７記載の電子装置において、上記ビューポートの方位角と高度とが視聴

者の頭部の方位角または方位角と高度の双方を連続的に測定し、（画像を表示しながら）実行時に上記ビューポートの位置を適宜調整するトラッキング装置によって自動的に制御されることを特徴とする電子装置。

【請求項３０】

請求項１１乃至１７記載の電子装置において、上記ビューポートのサイズ（視野レベルから見た幅と高度）を（画像を表示しながら）実行時にユーザが動的に制御することが可能であり、上記表示画像に対する“ズーム”制御機能を効果的に行うようにすることを特徴とする電子装置。

【請求項３１】

請求項１１乃至１７記載の電子装置において、３Ｄ“バーチャルな環境”（すなわち上記画像を含むオブジェクトが含まれる３Ｄバーチャル空間内での視聴者の視点）を（画像を表示しながら）実行時にユーザが動的に制御することが可能であることを特徴とする電子装置。

【請求項３２】

請求項１１乃至１７記載の電子装置において、３Ｄ“バーチャルな環境”（すなわち上記画像を含むオブジェクトが含まれる３Ｄバーチャル空間内での視聴者の視点）が“リアルワールド座標”で視聴者の物理位置を連続的に測定するトラッキング装置によって自動的に制御されることを特徴とする電子装置。

【請求項３３】

請求項１１乃至１７記載の電子装置において、（方位角または方位角と高度の双方を利用して）視聴者の頭部の方向により制御される方角に敏感なサウンドがリアルタイムで提供されることを特徴とする電子装置。

【請求項３４】

請求項１１乃至１７記載の電子装置において、（方位角または方位角と高度の双方を利用して）上記パノラマ内のビューポートの座標により制御される方角に敏感なサウンドがリアルタイムで提供されることを特徴とする電子装置。

【請求項３５】

請求項１１乃至１７記載の電子装置において、上記パノラマ内のビューポートの座標と、３Ｄ“バーチャルな環境”内の視点のバーチャル位置とによって制御される方角と位置とに敏感なサウンドがリアルタイムで提供されることを特徴とする電子装置。

【請求項３６】

請求項１１乃至１７記載の電子装置において、単純な単一視聴システムまたはパノラマカメラシステムを利用してリアルタイムで取得された“リアルワールド画像”と、予め録画されたコンピュータグラフィックや生の画像との合成を行う能力を持ち、それによって、拡張リアリティ体験を生みだすように為すことを特徴とする電子装置。

【請求項３７】

請求項１１乃至１７記載の電子装置において、単純な単一視聴システムまたはパノラマカメラシステムによって取得された“リアルワールド”画像を解析する能力を持ち、さらに、差分画像処理技術および／またはその他の画像処理技術を利用することにより、上記背景の“リアルワールド”場面を自動的に取り除き、この場面を、ビデオ装置（ＶＣＲＤＶＤプレーヤなど）から供給される合成画像または予め録画された画像と置き換える能力を持つことを特徴とする電子装置。

【請求項３８】

請求項１１乃至１７記載の電子装置において、“リアルワールド”画像あるいは予め録画された画像かのいずれかの画像を選択的に合成し、幾何学的に変更する能力を持ち、それによってコンポジットな拡張リアリティ体験をつくりだすように為すことを特徴とする電子装置。

【請求項３９】

請求項１１乃至１７記載の電子装置において、同じビデオソースにより前景情報と背景情報とを供給することが可能なクロマキー技法であって、さらに、特別の制御フレームある

いは上記ビデオ画像の各走査線の部分のいずれかとしてコード化されたクロマキーカラーのアナログサンプルまたはデジタルサンプルを供給することにより、上記クロマキーカラーを１画像内で動的に変えることができるクロマキー技法を利用して、“前景”と“背景”の予めレンダーされたビデオ情報を合成することができることを特徴とする電子装置。

【請求項４０】

請求項４９乃至５４に記載のような単純な単一表示システムまたはパノラマカメラシステムにより取得された画像を解析し、リアルタイムで２次元または３次元オブジェクトとして上記画像を解釈することが可能な電子装置（“バーチャライザ”）。

【請求項４１】

請求項４０記載の電子装置において、格納またはライブ転送を行うために、さらに、別のＶＴＶプロセッサ及び適切な表示装置とによる後続の復号化によりグラフィック画像への変換を行うために、上記３次元オブジェクトをＶＲＭＬまたはその他の適切な言語などの“ユニバーサル画像記述言語”に変換するように為すことを特徴とする電子装置。

【請求項４２】

図８乃至１０に図示の電子装置（もしくは上記ＶＴＶグラフィックプロセッサとして周知の装置）であって、

１以上のビデオデジタル化用モジュールと、

１以上の以下のメモリ領域であって、

拡張リアリティメモリ（ＡＲＭ）と、

バーチャルリアリティメモリ（ＶＲＭ）と、

変換メモリ（ＴＭ）とからなるメモリ領域と、

ＡＲＭおよび／またはＶＲＭの中に保持される上記データのアドレスマッピングを変更して、ＡＲＭおよび／またはＶＲＭ領域内の１つの位置から別の位置へグラフィック情報を効果的に動かすようにすることが可能な少なくとも“ピクセル変換エンジン”（ワープエンジン）、並びに、上記ソースと宛先位置との双方からデータを数学的に合成し、このデータを変更し、それによって画像合成、スケーリング、幾何学的変換、透かし（ｔｒａｎｓｐａｒｅｎｃｉｅｓ）などを実行する能力を達成する能力を含むデジタル処理用モジュールと、

１以上のビデオ生成用モジュールと、から構成されることを特徴とする電子装置。

【請求項４３】

請求項４２記載の電子装置において、上記拡張リアリティメモリ（ＡＲＭ）を“マップ”して、上記バーチャルリアリティメモリ（ＶＲＭ）よりも狭い垂直方向の視野を占め、さらに、変換メモリ（ＴＭ）を“マップ”して、高品質画像をそのまま維持しながら、上記メディアを供給するためのデータ要件を最小化するように為すことを特徴とする電子装置。

【請求項４４】

請求項４２記載の電子装置において、上記拡張リアリティメモリ（ＡＲＭ）と、バーチャルリアリティメモリ（ＶＲＭ）と、変換メモリ（ＴＭ）とを、異なる解像度で“マップする”ことも可能である（各メモリ領域のピクセルが、異なる度合いの角偏差を表すことできる）ことを特徴とする電子装置。

【請求項４５】

請求項４２記載の電子装置において、拡張リアリティメモリ（前景メモリ）の中に最初に“リアルワールド”ビデオ情報を入れ、ビデオ供給装置（ＶＣＲ、ＤＶＤプレーヤなど）からのソース情報をバーチャルリアリティメモリの中へ入れ、次いで、（平面スクリーン型表示装置やＨＭＤなどの）出力装置に表示する前に、変換メモリ（ワープエンジンの一部）に保持されているデータパターンに従って、これら２つの画像ソースを合成して“コンポジット画像”に変換することにより、請求項３６乃至３８に記載のように画像を表示することを特徴とする電子装置。

【請求項４６】

請求項４２記載の電子装置において、ビデオ供給装置（ＶＣＲ、ＤＶＤプレーヤなど）か

らの前景ビデオ情報を最初に拡張リアリティメモリの中へ入れ、次いで、ビデオ供給装置（ＶＣＲ、ＤＶＤプレーヤなど）からの背景ビデオ情報をバーチャルリアリティメモリの中へ入れ、次いで、（平面スクリーン型表示装置やＨＭＤなどの）出力装置に表示する前に、ソースメディア（”変換マップ”）により提供される付加情報に基づいてこれら２つの画像ソースを合成して“コンポジット画像”に変換することにより、もしくは、請求項４０に記載のように、上記オブジェクト認識装置“バーチャライザ”により提供される“リアルワールド”オブジェクトジオメトリを利用することにより、（平面スクリーン表示やＨＭＤなどの）出力装置に表示する前に“コンポジット画像”を生みだすためにワープエンジンと連携して、請求項３７乃至３９記載のように、画像を表示することを特徴とする電子装置。

【請求項４７】

時分割多重化と共に周波数圧縮と伸長の合成（コンパウンディング）を組み込んで、標準的な２トラック視聴覚録音システム上へ／から４またはそれ以上のオーディオトラックの符号化及びその後の復号化を行うことを特徴とする電子装置。

【請求項４８】

請求項４７記載の電子装置において、上記視聴覚データストリームのビデオ成分（Ｖ同期、Ｈ同期、奇数／偶数など）の中に存在する同期信号のうちの１以上の信号を利用して、遠隔クロックの生成および／または同期を行って時分割多重化済み視聴覚データの符号化及びその後の復号化を行うことを特徴とする電子装置。

【請求項４９】

（Ｐａｎ－Ｃａｍシステム）

複数の電子画像取得装置（ビデオカメラ、ＨＤＴＶカメラ、デジタル静止カメラなど）から構成される光電式アセンブリであって、オーバーラップする水平方向の視野によって構成されて、上記オーバーラップする水平方向の視野が最大３６０°を一体にカバーするように為すことを特徴とする光電式アセンブリ。

【請求項５０】

請求項４９に記載のアセンブリにより作成された個々の画像（ページ）をクロップし、アラインメントを行って、上記個々のページ間の無視できる歪みとオーバーラップとを持つ３６０°のパノラマ画像全体を形成するように為すことを特徴とする電子装置。

【請求項５１】

請求項５０記載の電子装置において、上記別個の画像のクロッピングとアラインメントとを行って、切れ目のない３６０°のパノラマ画像を形成するステップに加えて、上記画像に対して歪み補正も適用して、結果として得られる３６０°のパノラマ画像が横軸で数学的に“平らに”なる（上記画像の横軸の各ピクセルがカメラに対して等しい角度で対する）ように為すことを特徴とする電子装置。

【請求項５２】

請求項５１記載の電子装置において、上記画像に対して歪み補正も適用して、結果として得られる３６０°のパノラマ画像が縦軸で数学的に“平らに”なる（上記画像の縦軸の各ピクセルがカメラに対して等しい角度で対する）ように為すことを特徴とする電子装置。

【請求項５３】

請求項４９乃至５０記載の電子装置において、パノラマカメラアセンブリにより形成される３６０°のパノラマ画像を含む個々のページの位置を記述する“ページ識別情報”を送出用ビデオストリームの中へ挿入することを特徴とする電子装置。

【請求項５４】

請求項４９乃至５０記載の電子装置において、上記パノラマカメラアセンブリの現在の方角（方位角と高度）と、オプションとして上記パノラマカメラアセンブリの位置（すなわち方位角と高度、並びに、オプションとして上記カメラアセンブリの回転及びＸ、Ｙ、Ｚ位置）も記述する“トラッキング情報”を上記ビデオストリームの中へ挿入することを特徴とする電子装置。

【請求項５５】

電子装置であって、図１６、１７、２３に図示のように、また、パラグラフ４．１～４．８でさらに完全に説明されているように、１以上のビデオ取得装置（ビデオカメラなど）からの受信データを利用し、自己相関などの単純な一連の画像解析処理を実行することにより、（ＨＭＤベースのカメラアセンブリの場合、視聴者の）カメラの方位角での相対運動を計算することを特徴とする電子装置。

【請求項５６】

電子装置であって、図１６、１８、２３に図示のように、また、パラグラフ４．１～４．８でさらに完全に説明されているように、１以上のビデオ取得装置（ビデオカメラなど）からの受信データを利用し、自己相関などの単純な一連の画像解析処理を実行することにより、（ＨＭＤベースのカメラアセンブリの場合、視聴者の）カメラの高度での相対運動を計算することを特徴とする電子装置。

【請求項５７】

電子装置であって、図１６、１９、２３に図示のように、また、パラグラフ４．１～４．８でさらに完全に説明されているように、１以上のビデオ取得装置（ビデオカメラなど）からの受信データを利用し、自己相関などの単純な一連の画像解析処理を実行することにより、（ＨＭＤベースのカメラアセンブリの場合、視聴者の）カメラの回転時における相対運動を計算することを特徴とする電子装置。

【請求項５８】

電子装置であって、図１６、図２０～２３に図示のように、また、パラグラフ４．１～４．８でさらに完全に説明されているように、１以上のビデオ取得装置（ビデオカメラなど）からの受信データを利用し、自己相関などの単純な一連の画像解析処理を実行することにより、Ｘ軸、Ｙ軸、Ｚ軸のいずれかの軸またはこれらの軸の任意の組み合わせにおける（ＨＭＤベースのカメラアセンブリの場合、視聴者の）カメラの物理（空間）位置の相対運動を計算することを特徴とする電子装置。

【請求項５９】

請求項５５乃至５８記載の電子装置において、一定のまたは同期ストロボ発光が行われる軸上の光源と共に、所定の“リアルワールド”座標を持つ複数の逆反射作用を生じるターゲットを利用して、上記装置により確定された相対角データと相対空間データとを変換して絶対角データと絶対空間データとに変えることを目的として、絶対角基準値／空間基準値を確定することを特徴とする電子装置。

【請求項６０】

請求項５９記載の電子装置において、上記ＨＭＤベースのパノラマカメラまたは遠隔パノラマカメラのビデオ取得レートと同期する制御可能な軸上の光源と共に、上記逆反射作用を生じるターゲットに対するカラーフィルタの組み合わせを利用して、上記個々の逆反射作用を生じるターゲットのトラッキングを正しく識別し、該トラッキングを維持するシステムの能力の向上を図るように為すことを特徴とする電子装置。

【請求項６１】

請求項５９記載の電子装置において、上記ＨＭＤベースのパノラマカメラまたは遠隔パノラマカメラのビデオ取得レートと同期する制御可能な軸上の光源と共に、上記逆反射作用を生じるターゲットの組み合わせを利用して、上記個々の逆反射作用を生じるターゲットのトラッキングを正しく識別し、該トラッキングを維持するシステムの能力の向上を図るように為すことを特徴とする電子装置。

【請求項６２】

請求項５９記載の電子装置において、上記ＨＭＤベースのパノラマカメラまたは遠隔パノラマカメラのビデオ取得レートと同期するカラー制御可能な軸上の光源と共に、上記逆反射作用を生じるターゲットに対するカラーフィルタの組み合わせを利用して、上記個々の逆反射作用を生じるターゲットのトラッキングを正しく識別し、該トラッキングを維持するシステムの能力の向上を図るように為すことを特徴とする電子装置。

【請求項６３】

請求項５５乃至５８記載の電子装置において、複数の“アクティブな光ビーコン”（上記

ＨＭＤベースのビデオ取得レートまたは遠隔パノラマカメラのビデオ取得レートと同期する制御可能な光源）を利用して、パルス出力タイミングと、光の色および／またはこれらの組み合わせを用いて上記ビーコンの“リアルワールド”座標を上記ＨＭＤまたは遠隔パノラマカメラへ送信して、上記装置により確定された相対角データと相対空間データとを変換して絶対角データと絶対空間データとに変えることを目的として、絶対角基準値／空間基準値を確定するように為すことを特徴とする電子装置。

【請求項６４】

請求項５５乃至５８記載の電子装置において、上記ＨＭＤと上記ビーコンとの間で一意のＩＤコードの通信を行う複数の“双方向赤外線ビーコン”を利用し、上記ＨＭＤが“既知の空間基準点”のうちのいずれかの下を通過する度に１回だけこの較正が行われるように為すことを特徴とする電子装置。

【請求項６５】

単一の光画像処理装置を利用して、天井のパターンをモニタする電子装置であって、残りの角方向（ピッチと回転）を確定するための流体チルトセンサなどの別の角度トラッキングシステムと共に、請求項５５乃至５８に記載の同様の画像処理技術を利用して、相対空間運動と方位角とを確定するように為すことを特徴とする電子装置。

【請求項６６】

請求項６５記載の電子装置において、請求項５９乃至６４に記載のような較正システムのうちのいずれかのシステムを利用して、上記装置により確定された相対空間データを絶対空間データに変換することを目的として絶対基準値を確定することを特徴とする電子装置。

【請求項６７】

複数の電子サウンド取得装置（マイクロフォン）から構成されるアセンブリであって、オーバーラップする水平方向のカバーエリアフィールドにより構成されて、上記オーバーラップする水平方向のカバーエリアフィールドが水平方向に最大３６０°を一体にカバーするように為すことを特徴とするアセンブリ。

【請求項６８】

請求項６７に記載のような複数の電子サウンド取得装置（マイクロフォン）から構成されるアセンブリであって、オーバーラップする垂直方向のカバーエリアフィールドで追加して構成されて、上記オーバーラップする垂直方向のカバーエリアフィールドが最大値＋／－９０°までを一体にカバーするように為すことを特徴とするアセンブリ。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は画像視聴システムとサウンドシステムに関する。

【背景技術】

【０００２】

本願は、２００１年6月２５日にＡｎｇｕｓ　Ｄｕｎｃａｎ　Ｒｉｃｈａｒｄｓにより出願された“ＶＴＶシステム”という名称の米国特許出願番号第０９／８９１，７３３号により優先権を主張するものである。

【０００３】

通常のテレビ画面や映画スクリーンは現代の娯楽の一般的な形態であるが、これらの従来のシステムは、映画の動きと関連して視聴者のビューイング方向及び相対位置のユーザによる制御を可能にするものではない。さらに、バーチャルリアリティの表示がポピュラーになりつつあるが、３次元画像を生成する従来のシステムは複雑な場合が多く、メディアグラフィックは作成にコストのかかるものである。

【発明の開示】

【課題を解決するための手段】

【０００４】

以下の発明は、拡張型立体テレビジョン様の視聴体験を生成するハードウェア構成全体に

関する。通常のテレビとは異なり、本システムの場合、視聴者は映画の動きと関連して視聴者のビューイング方向と相対位置との制御を行うことが可能となる。専用のハードウェア構成に加えて、本発明はまた、このバーチャルリアリティ様の体験を可能にする新しいビデオフォーマットにも関する。さらに、上記目標の実現を容易にするいくつかの独自のビデオ圧縮規格も定義される。ＶＴＶシステムは、従来方式の２次元の画像撮影法と真のバーチャルリアリティとの間の中継技術として設計される。

【０００５】

ＶＴＶシステムの進化にはいくつかの段階が存在するが、その最も基本的な形であるパノラマ表示方式から、その最も優れた形式である、アニメーション化されたテクスチャマップを利用するフルオブジェクトベースのバーチャルリアリティを特徴化（ｆｅａｔｕｒｅ）し、さらに、“視聴環境を意識した”拡張リアリティシステム全体で実物の俳優および／またはコンピュータで生成したキャラクタを特徴化する方式の範囲にまで及ぶ段階がある。

【発明を実施するための最良の形態】

【０００６】

添付図面と関連して以下に記載の詳細な説明は、本発明の現在の好ましい実施例の説明を目的とするものであり、本発明の構成および／または利用の唯一の形態を表すことを意図するものではない。本説明は、例示される実施形態と関連して本発明を構成し、作動させる機能及び一連のステップについて記載するものである。しかし、本発明の精神と範囲内に含まれることが意図される異なる実施形態によっても、同じあるいは同等の機能及びシーケンスの達成が可能であると理解されたい。

【０００７】

１．１）以下の発明は拡張型立体テレビジョン様の視聴体験を生成するハードウェア構成全体に関するものである。通常のテレビとは異なり、本システムの場合、視聴者は映画の動きと関連して視聴者のビューイング方向と相対位置との制御を行うことが可能となる。専用のハードウェア構成に加えて、本発明はまた、このバーチャルリアリティ様の体験を可能にする新しいビデオフォーマットにも関するものである。さらに、上記目標の実現を容易にするいくつかの独自のビデオ圧縮規格も定義される。ＶＴＶシステムは、従来方式の２次元の映画撮影法と真のバーチャルリアリティとの間の中継技術として設計される。ＶＴＶシステムの進化にはいくつかの段階が存在するが、その最も基本的な形であるパノラマ表示方式から、その最も優れた形式である、アニメーション化されたテクスチャマップを利用するフルオブジェクトベースのバーチャルリアリティを特徴化し、さらに、“視聴環境を意識した”拡張リアリティシステム全体で実物の俳優および／またはコンピュータで生成したキャラクタを特徴化する方式の範囲にまで及ぶ段階がある。

【０００８】

１．２）図１に見られるように、本ＶＴＶシステム全体は、中央画像処理装置（ＶＴＶプロセッサ）と、各種ビデオ入力装置（ＤＶＤ、ＶＣＲ、衛星、地上テレビ、遠隔ビデオカメラ）、赤外線リモートコントロール、デジタルネットワーク接続部、及びいくつかの出力装置接続部とから構成される。本ＶＴＶユニットは、図２に図示するようなその最も基本的な構成では、従来型のテレビ装置へ画像を出力することになる。このような構成では、（おそらく赤外線方式の）リモートコントロール装置を用いて、このＶＴＶ環境内で視聴者の所望のビューイング方向と位置の制御が行われることになる。この“基本システム構成”の利点として、現在の視聴覚技術を利用してこの構成の実現が可能であるという点が挙げられる。ＶＴＶ画像規格は、標準的ビデオのレイヤー（ｌａｙｅｒ）の上位“レイヤー”と考えることができる（上位）互換性を有する画像規格である。すなわち、従来のビデオは新しいＶＴＶ画像規格のサブセットを表すものとなる。この規格が互換性を有する結果、テレビメーカーおよび／または視聴覚機器メーカーは仕様の大きな変更をまったく必要とせずＶＴＶを導入することが可能となる。さらに、ＶＴＶ互換テレビ復号化用ユニットは従来方式のテレビ送信と本質的に互換性を有するものである。

【０００９】

１．３）図３に図示するようなさらに優れた構成では、ＶＴＶシステムは表示装置として無線ＨＭＤを使用する。このような構成では、単純な画像表示に加えてトラッキング装置として無線ＨＭＤを使用することが可能である。このトラッキング情報は、最も基本的な形では、単にビューイング方向の制御から成るものであってもよい。さらに優れたシステムでは、バーチャルな環境において視聴者のビューイング方向と位置の双方の確定が可能となる。最終的に、最も優れた実施構成では、ＨＭＤの遠隔カメラによって、リアルワールド画像がＶＴＶシステムに入力され、ＶＴＶシステムはこのリアルワールド画像を解釈してこれを空間オブジェクトに変換し、次いで、この空間オブジェクトはバーチャルなオブジェクトと置き換えられ、それによって、“視聴環境を意識した”拡張リアリティシステムの提供が可能となる。

【００１０】

１．４）無線ＨＭＤは、無線データリンク“サイバーネットリンク”によってＶＴＶプロセッサに接続される。このリンクは、その最も基本的な形ではＶＴＶプロセッサからＨＭＤへビデオ情報を送信し、ＨＭＤからＶＴＶプロセッサへトラッキング情報を送信する能力を有する。サイバーネットリンクは、その最も優れた形では、ＨＭＤからＶＴＶプロセッサへのトラッキング情報の転送に加えてＨＭＤとサイバーネットリンク間でのビデオ情報の送受信も行うことになる。さらに、ＶＴＶプロセッサのある構成要素を遠隔ＨＭＤの中に組み込むようにし、それによってサイバーネットリンクを介するデータ転送要件を提言するようにしてもよい。この無線データリンクは、（未圧縮あるいはデジタル圧縮のいずれかのフォーマットで）アナログデジタルビデオ送信またはデジタルビデオ送信のいずれかを利用する多様な異なる方法で、トラッキング情報用の、二次のデジタルで符号化したデータストリームを用いて実現することができる。あるいは、これらのチャネルの双方を運ぶ純粋にデジタルな単方向のまたは双方向のデータリンクを組み込むこともできる。データ転送用の実際のメディアはおそらくマイクロウェーブや光である。しかしいずれの転送メディアも適切なものとして利用可能である。本システムの好ましい実施形態は、ＨＭＤのボード上やおそらくＶＴＶ基地局上の画像解析用ハードウェアと共にＨＭＤに取り付けられたオンボードのパノラマカメラを利用してリアルタイムのトラッキング情報を出力する形態となる。システム精度のさらなる向上を図るために、逆反射作用を有するマーカーを”リアルワールド環境で”利用することも可能である。このような構成では、オンボードカメラの光軸の近くに配置される切り替え可能な光源をこれらのカメラと共に利用して“差分画像解析”システムが形成される。このようなシステムは、直接ビデオ画像を単独で利用する場合よりも著しく高い認識精度を特徴とする。

【００１１】

１．５）最終的に、本ＶＴＶシステムは“ユニバーサル画像規格（ｕｎｉｖｅｒｓａｌ　ｇｒａｐｈｉｃｓ　ｓｔａｎｄａｒｄ）”を利用して画像情報の転送を行うことになる。このような標準規格では、オブジェクトベースの画像記述言語が組み込まれ、この言語によってサブシステム間での“共通の画像知識ベース”による高圧縮が達成されることになる。本発明は、基本的な用語で、この画像言語の進化における３つのレベルの漸進的精巧化（ｓｏｐｈｉｓｔｉｃａｔｉｏｎ）について記載するものである。

【００１２】

１．６）本特許の目的として以下の３つの圧縮規格について説明する：

ａ）ｃ－ｃｏｍ

ｂ）ｓ－ｃｏｍ

ｃ）ｖ－ｃｏｍ

１．７）ＶＴＶシステムは、その最も基本的なフォーマットにおいて、視聴者を取り囲む３６０°のパノラマ表示スクリーンと考えることができる。

【００１３】

１．８）上記“バーチャル表示画面”は多数の“ビデオページ”から構成される。“ページキーコード”がビデオ画像内で符号化され、この“ページキーコード”によって、上記“バーチャル表示画面”内の特定の位置内へ画像情報を入力するようにＶＴＶプロセッサ

への指示が行われる。動的に画像を入力する上記能力の結果、高解像度レートと高いフレームレートのいずれにも著しい犠牲を生じることなく、有効な同等の高解像度レートと高いフレームレートの双方の達成が可能となる。例えば、高速で変化している画像セクションのみが高速の画像更新を必要とするのに対して、画像の大部分は一般に静的である。（一般に動いている）重要な要素が主要場面に配置される従来方式の画像撮影法とは異なり、パノラマ画像の大部分は一般に静的である。

【００１４】

（ＶＴＶ画像規格：）

２．１）ＶＴＶ画像規格は、その最も基本的な形において、バーチャルな３６０°パノラマ表示画面から構成され、ＶＣＲ、ＤＶＤ、衛星、カメラあるいは地上テレビ受信装置のような外部のビデオソースから得られるビデオ画像をこのパノラマ表示画面上に描画することが可能であり、その結果、各ビデオフレームには、ビデオ情報だけでなく、バーチャルな表示画面内における各ビデオフレームの位置を定義する情報も含まれる。このようなシステムは、可変解像度画像だけでなく、フレームレートに依存しない画像も出力するので、著しく汎用性が高い。すなわち、ある特定のバーチャルな画像（バーチャル表示画面全体）内の実際の更新率は表示画面自体の範囲内で変動すると考えられる。この更新は、自身のバーチャル位置情報を含む各フレームによって達成される。これによって、ほとんどあるいはまったく変化しない画像上の非更新セクションの名目上の認知を代償にしてバーチャル画像のアクティブ領域の更新を迅速に行うことが可能となる。このようなシステムが図４に図示されている。

【００１５】

２．２）画像のリアリズムのさらなる改善を図るために、図５に図示のフォーマットに合わせて基本ＶＴＶシステムの拡張を行うことができる。この構成では、円筒形状のバーチャル表示画面が先端を切り落とした球面としてＶＴＶプロセッサにより解釈される。ＶＴＶプロセッサのデジタル処理ハードウェアコンポーネント内のジオメトリトランスレータ、すなわち”ワープエンジン”をしようすることによって、上記効果を容易に生み出すことが可能となる。

【００１６】

２．３）絶対基準平面が絶えず変動することから、（ＨＭＤベースまたはＰａｎ－Ｃａｍベースのいずれかの）移動カメラ用アプリケーションでは、ＶＴＶグラフィックエンジンが画像を正しく復号化できるように、カメラシステムの方位角と高度に関連するトラッキング付加情報が視覚情報と共に含まれることが求められる。このようなシステムでは、カメラの絶対方位角と高度が画像フレーム情報の一部となる。この絶対基準データを解釈するいくつかの可能性のある技法が存在する。まず、座標データを用いてメモリ書き込み処理中にメモリ内に像面の原点を画定することができる。残念ながら、上記アプローチでは、異なるアラインメント値を持つ前回のフレームからの残りの画像フラグメントが結果としてメモリ内に残される傾向がある。さらに実際的な解決方法として、単に、０方位角、０高度の想定基準点と共にビデオ情報をメモリの中へ書き込む方法がある。次いで、カメラアングルの補正値に対応する表示ビューポートの補正により上記ビデオ情報の正しい表示が行われる。このようなシステムのデータフォーマットが図１１に図示されている。

【００１７】

【表１】

制御フィールド	ビット	値	割当てキー
フレームフリップ	FF		フリップメモリフレーム
フィールドタイプ	F2-F0	0	前景ビデオ（ＡＲＭ）
		1	背景ビデオ（ＶＲＭ）
		2	デジタルハイブリッド（ＴＭ）
		3	デジタルビットストリームフレーム（ＴＭ）
		4	予約
		5	予約
		6	予約
		7	デジタル制御フレーム
ページ番号	P3-P0	0-15	（ＭＥＭレイアウトに依存）
オーディオ同期	AS		オーディオバッファを０にリセット
高度補正	E8-E0	(±45°)	カメラの高度
方位角補正	A10-A0	(±180°)	カメラ方位角
オーディオトラック	T1-T0	0	オーディオトラックなし
		1	４オーディオトラック
		2	８オーディオトラック
		3	オブジェクトベースのオーディオ
オーディオサンプルレート	S1-S0	0	２／４回線（１５Ｋ　Ｓ／Ｓ）
		1	３／６回線（２３Ｋ　Ｓ／Ｓ）
		2	４／８回線（３１Ｋ　Ｓ／Ｓ）
		3	５／１０回線（３８Ｋ　Ｓ／Ｓ）

（オーディオ規格：）

２．４）３６０°パノラマビデオに加えて、４トラック（４チャネル録音再生方式）または８トラック（８チャネル録音再生方式）のいずれかの立体的サウンドもＶＴＶ規格によってサポートされる。４トラックシステムのバーチャルな表現が図６に示されている。単純な４トラックオーディオシステムの場合、サウンドシステムの左右スピーカ（またはＨＭＤベースのシステムの場合、ヘッドホン）を通る音は、ビューポート（ＶＲ環境内での視方向）の方位角に基づいてスケーリングされる。８トラックオーディオシステムの場合、サウンドシステムの左右スピーカ（またはＨＭＤベースのシステムの場合、ヘッドホン）を通る音は、システムのバーチャルな表現（図７）に図示のように、ビューポートの方位角と高度の双方に基づいてスケーリングされる。

【００１８】

２．５）その最も基本的な形では、図１２に図示するようなデジタル／アナログハイブリッドフォーマットにおいて、ＶＴＶ規格によって、ビデオ情報の一部としてマルチトラックオーディオチャネルの符号化が行われる。その結果、現行機器とビデオとの互換性の達成が可能となる。この例図に見られるように、オーディオデータは、アナログコード化された圧縮フォーマットで格納されているため、各ビデオ走査線には５１２本のオーディオサンプルが含まれるようになる。このアナログコード化オーディオ情報に加えて、各オーディオ走査線には、オーディオ情報の“プレスケーリング”に用いられる３ビットデジタルコードが含まれる。すなわち、実際のオーディオサンプル値はＸ＊Ｓであり、ここでＸはプレスケール番号であり、Ｓはサンプル値である。上記２種の符号化方式を用いて、オーディオシステムのダイナミックレンジを約４３ｄＢから６０ｄＢ以上まで拡げることが

可能となる。第２に、このダイナミックレンジの拡張は、オーディオ品質に比して比較的“低コスト”で行われる。なぜなら、信号レベル全体が高ければ、人は音の歪みに対して相対的に敏感ではないからである。スタートビットはこのシステムでは重要な構成要素である。スタートビットの機能は、走査線に対して最大レベル（例えば１００％レベルすなわち白レベル）を設定することである。黒レベルと関連するこのレベル（このレベルはカラーバーストの直後にサンプリングすることができる）は、各ラインに対して０％と１００％のレンジを形成する。ラインベースで各ラインに対する０％と１００％のマークを動的に調整することにより、走査線のデジタル成分の復号化の精度が向上することに加えて、ビデオサブモジュールのＡＣ結合および／またはビデオメディアの録画と再生に起因して生じる黒レベルの変動に対するシステム感度がずっと低くなる。

【００１９】

２．６）このデジタル情報のプレスケーリングに加えて、オーディオ制御ビット（ＡＲ）が（ライン２１の）各フィールドに含まれる。この制御ビットによってオーディオバッファシーケンスは設定時０に設定される。この設定によって、４または８トラックオーディオ情報が同期されて、その結果、ビデオページの更新シーケンスに関係なく常に現在のデータから正しいトラックの更新が行われる。

【００２０】

２．７）別の構成では、ビデオ録画再生システム（典型的にはデジタルオーディオトラックや線形アナログオーディオトラック）により出力される従来方式のステレオオーディオを用いることによって、立体的サウンドの提供が可能となる。最善でも、このような従来方式のシステムでは、２つのオーディオトラックだけに４８ｋＨｚのサンプルレート（すなわち、２チャネル２４ｋＨｚオーディオ）が与えられるのみである。しかし、オーディオデータを圧縮することによって同じ記憶メディア上へさらに多数のオーディオトラック（例えば４または８オーディオトラック）をタイムドメインで符号化し、次いで、その結果得られるデータの時分割多重化を行うことが可能となる。受信側では、まず時分割多重化されたデータを復号化し、次いでオーディオデータを減圧縮（ｄｅｃｏｍｐｒｅｓｓｉｎｇ）することによって上記プロセスが逆に行われて、一連の連続するオーディオストリームが形成される。高いオーダーのノッチフィルタまたはハイパスフィルタを組み込んで、時分割多重化／逆多重化処理手順により生じるサンプリングノイズをすべて取り除くようにすることが好ましい。このようなフィルタは、おそらくデジタル設計のフィルタであり、単にＶＴＶプロセッサのデジタル処理用ハードウェアの一部にすぎない。このようなオーディオ符号化システムのための代表的な構成が図１４～１５に図示されている。

【００２１】

このようなシステムが簡単に実現可能であることは明らかである。オーディオチャネル数の増加による“パフォーマンス代価”は、システムの周波数応答が減少することである。この減少はチャネル数の上昇に直接比例する。例えば、チャネル数が２から４へ増加すると、周波数応答が２４ｋＨｚから１２ｋＨｚへ減少することになる。この減少レベルは容認できる範囲のレベルであり、最終的にこのオーディオの多重化システムは、比較的“ロバスト（ｒｏｂｕｓｔ）”で、かつ、ＶＴＶフォーマットマテリアルの実施と編集が容易であるという点から見て、顕著な利点を与えるものであるという点で、前述したビデオ符号化されたオーディオシステムにとって好ましいものとなり得るものである。

【００２２】

任意の時分割多重化システムの有効な実施において最も重要なファクタの１つとして、時分割符号化クロック及び復号化クロックの安定した精度の高い生成が挙げられる。これは絶対に重要なファクタである。なぜなら、復号化されたデータパケットのいかなるミスアラインメントも非常に顕著なチャネルクロストークに関連するノイズを結果として引き起こすことになるからである。本ＶＴＶシステムでは、ビデオストリームの中に同期信号（Ｈ同期やＶ同期、奇数／偶数などのいずれか）を用いて、ローカル復号化クロックを生成させることで、この問題に対する新規かつ非常に有効な解決方法を組み込む。これらの信号は一般に非常に正確であり、デジタル記録機器をＶＴＶデータの保存に利用する場合、

特に正確である。
【００２３】
好ましい実施形態では、時分割復号化ノイズ（ＮＴＳＣに対して毎秒６０スパイク）が相対的にほとんど発生しないこと、このノイズがデジタルノッチフィルタを使用することで（必要な場合）容易に除去できると考えられることを考慮して、復号化クロックとしておそらくＶ同期が使用されることになる。一方、（Ｈ同期などの）高周波時分割クロックを選択した場合、さらに多くの不連続インスタンス（毎秒約１５，０００）が発生する可能性があるが、同様に、デジタルフィルタ（この場合ローパスフィルタ）によってこれらのインスタンスを除去することが可能である。ＨＤＴＶやその他の高い走査速度（３１ｋＨｚまたはそれ以上の）のＴＶ規格を用いる場合、高速の復号化クロックの方が好ましいものになると考えられる。なぜなら、このような状況では、フィルタリングの要件が非常に簡単でかつ効果的になるためである。さらに、高速の復号化クロックを利用するシステムのメモリ要件が大幅に減少する。
【００２４】
２．８）コンピュータ用ＡＶファイル及びデジタルテレビ送信などのさらに優れたマルチメディア用データフォーマットでは、さらに効率的であるか、あるいは利点の多い別の方法でこれらの付加的オーディオトラックを格納することも可能である。
【００２５】
２．９）視聴覚装置としての利用に加えて、この空間でオーディオシステム／規格を、好適でコンパクトなトラッキング装置と１組のコードレスヘッドホンとを組み合わせることによって、オーディオ専用モードで使用して、高性能のハイファイ機器用空間オーディオシステムの実現を図ることも可能であることに注意されたい。
【００２６】
（拡張機能）
２．１０）上記の単純化した画像規格に加えて、単独でまたは基本ＶＴＶ画像規格と共に使用可能な多数の拡張機能がある。後続する発明で、ａ）ｃ－ｃｏｍ、ｂ）ｓ－ｃｏｍ、ｃ）ｖ－ｃｏｍとして知られる３つの画像規格について詳細に説明するが、これは本発明の目的のために行うものである。
【００２７】
２．１１）最初の２つの規格が空間画像オブジェクトの定義に関するものであるのに対して、第３の画像規格は、サブセットとして第１の規格を利用し、追加の環境定義と制御アルゴリズムとを組み込む完全なＶＲ環境定義言語に関する規格である。
【００２８】
２．１２）（最も基本的な形での）ＶＴＶ画像規格は、従来方式のビデオ規格（ＮＴＳＣ、ＰＡＬなど）の層の上位にある制御層と考えることができる。したがって、ＶＴＶ画像規格は従来方式のアナログビデオ送信規格に制限されることは全くない。基本的に同一の技術を用いて、本ＶＴＶ規格は、ＨＤＴＶ規格並びにコンピュータグラフィック及び業界の視聴覚規格の多くと共に機能することができる。
【００２９】
２．１３）表１と図１１で概説したデジタル制御情報は、本システムの一つの可能な構成を表すものにすぎず、現行のアナログＴＶシステムとの互換性を有する仕様の一般性を減じるものではない。制御ビットの別の構成および／またはこの情報を符号化してアナログＴＶなどのようなビデオフォーマットに変換する別の方法が存在し、そのような構成および／または方法も同様に有効である。例えば、ＭＪＰＥＧやＭＰＥＧ-ＩＩなどの多くのデジタル符号化処理によって、連続するビデオ走査線が平均化されることが多く、図１１に図示するような単一ラインとして符号化される場合、デジタル情報が損なわれる傾向があることが知られている。しかし、一連のラインまたは連続するラインとしてこの情報を符号化すれば、この問題は解決される。このような構成では、データは、一連の直線分ではなく、結果として得られる画像内の一連の小さな矩形として符号化されるデータと考えることができる。このような変更は予想されるものであり、別のビデオフォーマット、特

にＨＤＴＶなどのようなデジタル符号化されたフォーマットを用いてパフォーマンス効率の向上を図る、さらに別の適合化が予想される。

【００３０】

（ＶＴＶプロセッサ：）

３．１）ＶＴＶグラフィックプロセッサはＶＴＶシステムの核心部である。このモジュールは、その最も基本的な形で、表示装置（従来方式のＴＶ／ＨＤＴＶあるいはＨＭＤのいずれか）へ出力画像をリアルタイムで生成する役割を果たすものである。ＶＣＲ、ＤＶＤ、衛星、カメラあるいは地上テレビ受信機のようなビデオメディア供給装置から入力される生の画像情報のデジタル化に加えて、このモジュールのさらに優れたバージョンでは、インターネットやその他のネットワーク接続を介して渡される“ユニバーサル画像言語（ｕｎｉｖｅｒｓａｌ　ｇｒａｐｈｉｃｓ　ｌａｎｇｕａｇｅ）”から得られる画像をリアルタイムで描画することも可能である。このデジタル化タスクと描画タスクに加えて、ＶＴＶプロセッサは画像解析を実行することも可能である。本システムの初期のバージョンは、ＨＭＤの座標のトラッキングを確定することを目的としてこの画像解析機能を利用するものである。このモジュールのさらに優れたバージョンでは、上記トラッキング情報の提供の他に、ＨＭＤから得られるリアルワールド画像の物理的３次元オブジェクトとしての解釈も行われる。これらの３次元オブジェクトはユニバーサル画像言語で定義され、次いで、それを記録し、インターネットやその他のネットワークを介して同様の遠隔地にある表示装置へ通信したり、あるいは、類似する物理的サイズの別のバーチャルなオブジェクトにより置き換えたりすることが可能であり、それによって、真の拡張リアリティ体験が創り出される。

【００３１】

３．２）ＶＴＶハードウェア自体は以下のようなサブモジュール群から構成される。：

ａ）ビデオデジタル化モジュール

ｂ）拡張リアリティメモリ（ＡＲＭ）

ｃ）バーチャルリアリティメモリ（ＶＲＭ）

ｄ）変換メモリ（ＴＭ）

ｅ）デジタル処理ハードウェア

ｆ）ビデオ生成モジュール

３．３）これらのモジュールの正確な構成は他の外部ハードウェアによって決まる。例えば、デジタルビデオソースを用いる場合、ビデオデジタル化モジュールは比較的単純なものとなり、１グループのラッチバッファまたはＦＩＦＯバッファのみから構成される。しかし、コンポジットビデオ入力やＹ／Ｃビデオ入力を利用する場合、これらの信号をデジタル形式へ変換するための追加ハードウェアが必要となる。さらに、ビデオ入力ソースとしてデジタルＨＤＴＶ信号を使用する場合、（ＨＤＴＶ信号は圧縮されたフォーマットでは処理できないため）システムのフロントエンドとしてＨＤＴＶデコーダが必要となる。

【００３２】

３．４）アナログＴＶなどのようなフィールドベースのビデオシステムの場合、ＶＴＶグラフィックエンジンの基本処理は以下のようになる：

ａ）ビデオ情報がデジタル化され、０°方位角、０°高度の絶対ページ基準を想定して、フィールドベースで拡張リアリティメモリ内に入れられ、各ページの原点がページ番号ビット（Ｐ３～Ｐ０）の状態により確定される。

【００３３】

ｂ）背景および／または床／天井マップに関連する補助ビデオ情報は、“フィールドタイプ”ビット（Ｆ３～Ｆ０）とページ番号ビット（Ｐ３～Ｐ０）の状態によって決まる。フィールドベースでバーチャルリアリティメモリ内にロードされるか、セクション３．１２に記載のような拡張リアリティメモリ内に入れられるビデオ情報の解釈によりリアルタイムで生成される。

【００３４】

ｃ）デジタル処理用ハードウェアは、拡張リアリティメモリとバーチャルリアリティメモ

リの中に保持された上記情報を解釈し、ジオメトリ処理用エンジン（ワープエンジン）、デジタル減算画像処理、及び“ブルースクリーニング（ｂｌｕｅ－ｓｃｒｅｅｎｉｎｇ）”の新しい汎用性の高い形式との組み合わせを利用してこのデータを変換し、選択的にこのデータを合成して、ビデオマテリアルの撮影時に、パノラマカメラの在る位置と同じ位置に視聴者が立っていた場合に、その目に映る画像とほぼ類似の画像に変換する。この画像と、従来方式のビデオ技法を利用して入手可能な画像との間の主要な相違点として、前者の画像は、３６０°のパノラマ画像であるだけでなく、複合した没入型の拡張リアリティ体験を形成できるように一体に融合されたバーチャルリアリティと“リアルワールド”画像の双方の要素を備えることも可能であるという点が挙げられる。

【００３５】

ｄ）バーチャルリアリティと“リアルワールド画像”とを合成する正確な方法はＶＴＶプロセッサの作動モードによって決まるが、この方法については、本明細書の後のセクションでさらに詳細に説明する。この特別のＶＴＶプロセッサモードは、ソースメディアの中に存在する追加の制御情報により決定される。したがって、ＶＴＶメディアのソースを表示しながら、処理モードと表示モードとを動的に変更することが可能である。

【００３６】

ｅ）次いで、ビデオ生成モジュールは、従来方式のテレビまたはＨＭＤ表示装置に表示される単一のビデオ画像または一対のビデオ画像を生成する。（マルチスピンＤＶＤ装置が画像メディアとして使用されない場合）ＶＴＶ画像フィールドは最大フレームレート未満で更新されるが、画像の描画は、立体的サウンドの更新の場合のように、そのまま最大のビデオフレームレートで行われる。これが可能である理由として、各“画像球面”が、視聴者の任意の方角（方位角と高度）に関連するビデオとオーディオ双方の必要な情報のすべてを含むということが挙げられる。

【００３７】

３．５）図９に見ることができるように、ＶＴＶプロセッサのメモリの書き込み側は（ＡＤＣの）２つの別個のビデオ入力段を示す。ＡＤＣ－０は一般に生のパノラマビデオの送出に用いられ、ＡＤＣ－２は一般に予め描画されたビデオマテリアルからのバーチャルリアリティビデオの送出に用いられるものではあるが、双方のビデオ入力段は、拡張リアリティメモリとバーチャルリアリティメモリの双方へ完全にアクセスする（すなわち、これらのメモリではメモリプールが用いられる）ことに注意されたい。上記ハードウェア構成では、設計時にさらに多くの汎用性がサポートされ、いくつかの通常のものとは異なる表示モードの実施が可能である（これについては後のセクションでさらに詳述する）。同様に、ビデオ出力段（ＤＡＣ－０とＤＡＣ－１）はバーチャルリアリティメモリと拡張リアリティメモリの双方へ完全にアクセスする。

【００３８】

３．６）２つの入力段と２つの出力段とを備えることにより、設計の汎用性の向上が図られるが、メモリプールスタイルの設計により、１つ乃至２つの入力段のいずれかおよび／または出力段によってシステムが機能することが可能になり（但し機能は低下するが）、したがって、ある特定の実施構成において１つまたは２つの入力段あるいは出力段のいずれかが存在すれば、仕様の一般性が制限を受けるものではない。

【００３９】

３．７）設計の容易さのために、高速スタティックＲＡＭがプロトタイプ装置のビデオメモリとして利用された。しかし、この設計仕様の一般性を制限することなく別のメモリ技術の利用も可能である。

【００４０】

３．８）好ましい実施形態では、デジタル処理ハードウェアは１以上のフィールドプログラマブル論理アレイやカスタムＡＳＩＣの形をとる。Ｄｅｉｄｐｒｏｇｒａｍｍａｂｔｅｉｏｇｉｃアレイを用いる利点としては、ハードウェアを何時でも更新可能であるという点が挙げられる。この技術の主な欠点としては、ＡＳＩＣほど高速ではないという点が挙げられる。あるいは、従来方式の高速デジタルプロセッサを利用して上記の画像解析およ

び／または画像生成タスクを実行することも可能である。
【００４１】
３．９）前述したように、おそらくＶＴＶハードウェア全体が携帯用ＨＭＤ装置内に存在するようになる程度まで、このハードウェアの特定のセクションをＨＭＤの中に組み込むことも可能である。このような場合、ＶＴＶ基地局のハードウェアは、ＨＭＤとインターネットやその他のネットワークとの間のリンクとしてのみ機能することになり、すべてのグラフィック画像生成、画像解析及び空間オブジェクト認識はＨＭＤ自体の中で行われることになる。
【００４２】
３．１０）注：ビューポートアドレスジェネレータの下位ビットは、生成された画像に対してバレル歪曲をかけるＸＹ画像軸のルックアップテーブルアドレス変換装置を介して実行される。この変換装置によって、ビューポートの現在の視野の正確な画像の歪曲が示される。このハードウェアは、おそらくＦＰＧＡやＡＳＩＣ論理回路内に実装され、ビューポートアドレスジェネレータの機能ブロックの一部を含むことになるため、図１０には明示されていない。同様の方法で最終画像の同じような回転が実現される。
【００４３】
３．１１）ビューポート－０だけが変換エンジン（ワープエンジン）により影響を受け、ビューポート－１は歪みを受けずに読み出されることに注意されたい。スーパーインポーズ及びオーバーレイ拡張リアリティモードを使用する場合、この読み出しが必要となる。なぜなら、記憶装置から再生されるＶＲビデオマテリアルは格納される前に予め“平坦化”（すなわち、ピンクッション歪み）されているのに反して、ＨＭＤに備えられたパノラマカメラから得られる生のビデオ画像は、本システムによって拡張リアリティモードで表示される前に歪み補正を必要とするからである。上記予備的歪み補正を行った後、ＨＭＤのパノラマカメラにより録画された画像は、それ自体新しいＶＲマテリアルとして、幾何学的に正確でかつ格納に適したものにすることが望ましい（すなわち、これらの画像はＶＲマテリアルになることができる）。この場合、ワープエンジンの主要な役割の１つとしてはＨＭＤのパノラマカメラのジオメトリ補正とトリミングとがある。この中にはカメラの視野間で切れ目のない移行を行うという複雑なタスクが含まれる。
【００４４】
（例外処理）
３．１２）図４、５に見られるように、ＶＴＶ画像フレームは円筒形または先端を切り落とした球面のいずれかから構成される。この空間は、視聴者に対して有限の対頂角（プロトタイプで＋／－４５°）でしか対していない。これは、ビデオ格納装置の利用可能なデータ帯域と送信メディアとを最大限に活用できるように設計された意図的な制限であり、これによって、現行ビデオシステムとの互換性が保たれる。しかし、この妥協の結果、ビューポートが画像データの範囲を超える状況が生じる可能性がある。第一に、この例外的状況を処理できるいくつかの異なる方法がある。この例外的状況を処理する最も単純な方法として、単に境界線外のビデオデータを黒にする方法が挙げられる。これによって、天井と床が黒い部屋に居る外観が得られる。上記とは別に、さらに優れたアプローチとして、“床”と“天井”（すなわちＶＴＶ画像フレームの底部と最上部に存在する画像情報に基づく、ＶＴＶ画像フレームの下方及び上方の空間）を形成するアプローチがある。特定の一実施構成では、ＶＴＶ画像フレームの最上部と底部“リム”に対して赤、緑、青のピクセルの平均値を用いて、純色表面または好適にはグラデーションを施した色表面のいずれかを形成しその環境の“天井”と”床”とを表すようにすることも可能である。戸外で生成されるＶＴＶビデオの状況では、通常天井が青のシェードで、床はほぼ地面の色という結果が生じることになる。ビューポートの高度に伴ってシェーディングを変動させることによって生成される“床／天井”の現実感が改善され、その結果一般にビューポートの高度が＋／－９０°の極値に近づくにつれて、輝度が上昇する。この効果は、比較的単純な数学を利用して簡単に生成することができ、おそらくＶＴＶグラフィックエンジンの一部としてハードウェア内で実現される。

【００４５】
しかしさらに別の好ましい構成として、低い解像度で最大３６０°＊１８０°の背景画像マップを格納する２次的ビデオメモリによる格納を利用する構成がある。このメモリ域はバーチャルリアリティメモリ（ＶＲＭ）として知られている。拡張リアリティメモリ及びバーチャルリアリティメモリの両方を利用するシステム用の基本メモリマップが（変換メモリと併せて）図８に示されている。上記例図でわかるように、この変換メモリ域は最大３６０°＊１８０°をカバーできる程の十分なレンジを持つ必要があり、理想的には、（３６０°＊９０°をカバーする）拡張リアリティメモリバンクの角度分解能と同じ角度分解能を持つ必要がある。このような構成によって、床及び天井に対する例外処理と、前景の窓を通して見たり、窓の後ろの背景を示したりするような変動する透明画像との両方の提供が可能となる。これらの背景は静的あるいは動的のいずれかであることができ、ページ化されたフォーマットの利用により、前景（拡張リアリティメモリ）と基本的に同じ方法で更新が可能である。

【００４６】
（オペレーションモード）
３．１３）ＶＴＶシステムには２つの基本的オペレーションモードがある。これら２つのモード内にはいくつかのサブモードも存在する。この２つの基本モードは以下の通りである：
ａ）拡張リアリティモード
ｂ）バーチャルリアリティモード
（拡張リアリティモード１：）
３．１４）拡張リアリティモード１では、“リアルワールド画像”の選択成分がバーチャルリアリティの背景の上にオーバーレイされる。一般に、この処理では、まず“リアルワールド”画像から背景成分をすべて除去することが必要となる。これは、差分画像処理技術を用いて、すなわち現在の“リアルワールド”画像を前にとった格納済みのコピーと比較し、この２つの差分を検出することにより簡単に行うことができる。２つの画像の正確なアラインメントを行った後、異なる領域が新規のオブジェクトまたは前景のオブジェクトとなり、同じ状態のまま変化しない領域は静的な背景オブジェクトとなる。これは拡張リアリティモードの中で最も単純なモードであり、背景のほとんどは処理時に除去されることになるので、一般に十分な関心の対象となるものではない。移動式Ｐａｎ－Ｃａｍ（テレプレゼンス）や拡張リアリティモードで作動する場合、拡張リアリティメモリは、一般に、ランダムなページ更新ではなく、シーケンシャルなページ順に（すなわちシステムフレーム全体で）更新が行われることに注意されたい。このような更新を行う理由としては、撮影中のパノラマカメラシステムの位置と方角が常に変動するため、これらの変動を別々に処理した場合、画像ページの不一致が生じる可能性が高いということが挙げられる。

【００４７】
（拡張リアリティモード２：）
３．１５）拡張リアリティモード２は、前景の自動抽出と、オブジェクトの移動と、人工的背景環境におけるこれらの配置とに加えて、システムがワープエンジンも利用して追加の“リアルワールド”オブジェクトを背景の中へ“押し込む”という点で、モード１とは異なるものである。これらの“リアルワールド”オブジェクトを単純にバーチャルな環境の中へ追加することに加えて、ワープエンジンは、これらのオブジェクトのスケーリングと変換とを行う能力も有するため、さらに有効にこれらのオブジェクトがバーチャルな環境の中へマッチするようになる。これらのオブジェクトは、不透明なオーバーレイあるいはトランスペアレンシとして処理することができる。

【００４８】
（拡張リアリティモード３：）
３．１６）拡張リアリティモード３の場合、ワープエンジンを用いて、背景オブジェクトを前景の中へ“引き込み”、“リアルワールド”オブジェクトを置き換えるという点でモ

ード２とは異なるものである。モード２の場合のように、これらのオブジェクトの変換及びスケーリングを行うことが可能であり、不透明なオーバーレイあるいはトランスペアレンシのいずれかとしてこれらのオブジェクトの処理を行うことが可能である。これによって、ユーザは“リアルワールド”オブジェクトの物理的サイズと位置をバーチャルなオブジェクトと「一致」させることができるようになる。これを行うことにより、ユーザは、“リアルワールド”環境の場合と同じように、拡張リアリティ環境内でリアルタイムで交信し、ナビゲーションを行うことが可能となる。上記モードは、ユーザ自身の居間の中へハリウッド製作の映画を持ち込むことを可能にするものとして、娯楽目的やゲーム目的のためにおそらく最も利用が予想されるモードである。

【００４９】

（拡張機能）

３．１７）拡張リアリティモード２及び３を有効に作動させるキーとなるのは、高速かつ正確な光トラッキングシステムであることは明らかである。理論的には、ＶＴＶプロセッサが“リアルワールド”オブジェクトをリアルタイムで識別し、追跡することが可能である。しかし、特に、“リアルワールド”環境内における視聴者の物理位置の変化と共にオブジェクトジオメトリが大幅に変化するため、この識別と追跡は複雑なものとなる。したがって、単純な自動相関型のトラッキング技法が有効に機能しないことになる。このような状況では、逆反射作用を生じるいくつかのターゲットを当該オブジェクトのキー要素上に配置することにより、トラッキング精度の大幅な向上を図ることが可能である。比較的単純な差分画像処理技術の利用によりこのような逆反射作用を生じるターゲットを容易に識別することが可能である。

【００５０】

（バーチャルリアリティモード：）

３．１８）バーチャルリアリティモードは前述の拡張リアリティモードよりも機能的に単純なモードである。このモードでは、“予め撮影された”またはコンピュータで生成された画像がランダムなページベースで拡張リアリティメモリの中へロードされる。これが可能となるのはバーチャルカメラの基準平面が固定されているためである。前述の例の場合のように、バーチャルリアリティのメモリは固定した背景や動的背景についてはさらに低い解像度でロードされる。前景像面及び背景像面の双方を使用することで、運動視差などのさらに優れた画像処理技術が可能となる。

【００５１】

（拡張機能）

３．１９）拡張された“ブルースクリーニング”の形を利用することにより、バーチャルリアリティメモリ（背景メモリ）の汎用性の向上を図ることが可能となる。このようなシステムでは、“クロマキー”カラーのサンプルが背景フィールドにおける各走査線の開始時に与えられる。これによって、画像内で任意の色が許容可能となる汎用性の高いシステムが得られる。したがって、“透明な”クロマキーカラーによって個々のオブジェクトを囲繞することで、ワープエンジンによるこのオブジェクトの“カット＆ペースト”に関連する問題と不正確さとが大幅に減少する。さらに、前景のバーチャルリアリティ画像内の“透明な”クロマキー済みの領域を使用することにより、付加情報オーバーヘッドを伴うことなく、複雑な鋭く尖った領域および／または動的前景領域の簡単な生成が可能となる。

【００５２】

（カメラシステム：）

４．１）画像規格の定義でわかるように、移動式Ｐａｎ－ＣａｍやＨＭＤベースのビデオシステムにより取得される画像を正しく配置し、次いで表示するためには、追加のページ配置及びトラッキング情報が必要となる。さらに、立体的サウンドをリアルタイムで記録する場合、ビデオストリームの一部として上記情報の符号化も行わなければならない。コンピュータ生成画像の場合、描画段階でこの追加ビデオ情報を簡単に挿入することができる。これに対して、生のビデオを取得する場合、録画の前に上記追加トラッキング情報と

オーディオ情報とをビデオストリームの中へ挿入しなければならない。これは、本明細書で以後ＶＴＶエンコーダモジュールと呼ぶ画像処理用モジュールにより効果的に達成することができる。

【００５３】

（画像取得）

４．２）移動式パノラマカメラシステムにより収集される画像の場合、画像はまずＶＴＶエンコーダモジュールにより処理される。この装置によってビデオ歪み補正が行われ、ビデオページ情報、方角トラッキングデータ及び立体的サウンドデータがビデオストリームの中へ挿入される。ビデオ規格を変更せずに上記挿入が可能なため、現行の記録再生装置との互換性が保たれる。このモジュールはＶＴＶプロセッサ内へ組み込むこともできるが、最終的ビデオ情報を格納するか、何らかの形で無線ネットワークを介して送信するかのいずれかを必要とする遠隔カメラ用アプリケーションで使用するのには、このモジュールを別個の構成要素として設けるのが好都合である

（トラッキングシステム：）

４．３）“Ｐａｎ－Ｃａｍ”やＨＭＤベースのカメラシステムのようないずれの移動式パノラマカメラシステムの場合も、“絶対”方位角と高度座標系を維持するために、トラッキング情報は、結果として得られるビデオストリームの一部を含む必要がある。コンピュータ生成画像の場合、カメラ方向は描画時にコンピュータシステムが認知する理論上の構成物であるため、上記データは必要ではない。

【００５４】

（基本システム：）

４．４）ＶＴＶ　ＨＭＤの基本トラッキングシステムでは、オンボードのパノラマビデオカメラを利用して、囲繞するリアルワールド環境に関する３６０°の必要な視覚情報が取得される。次いで、この情報は、自動相関のような計算上集中的な、しかしアルゴリズムの点からは比較的単純な技法を利用して、ＶＴＶプロセッサ（ＨＭＤ内に存在するか基地局ユニットとして存在するかの如何にかかわらず）により解析される。可能なアルゴリズムの例が図１６～２２に示されている。

【００５５】

４．５）図１６～２２に概説される単純なトラッキングシステムは位置と方角の変化のみを検出するシステムである。差分画像処理技術を用いることで容易に背景画像から識別できるいくつかの逆反射作用を生じるターゲットを追加することにより、絶対基準点を得ることが可能となる。このような絶対基準点は、ＶＴＶハードウェアがこれらのマーカーのリアルワールド座標を認知するかぎり、現実の環境内のどこに置かれるようなことがあっても、環境領域の末端部（すなわち、ユーザスペースの境界部）に位置することになると考えられる。（画像解析データから得られる）上記絶対基準点と差分運動との組み合わせによって、最大のビデオフレームレートでの絶対リアルワールド座標情報の生成が可能となる。所定の空間座標における逆反射作用を生じるターゲット配置の代替物として、アクティブな光ビーコンを用いることも可能である。これらの装置は、ビデオ取得レートと同期してストロボが発光するように構成され、結果として得られる画像に対する差分ビデオ解析が可能となるという点で、逆反射作用を生じるターゲットと同様の方法で作動する装置である。しかし、逆反射作用を生じるパッシブなターゲットとは異なり、アクティブな光ビーコンは、ビデオ取得と同期してストロボを発光することに加えて、上記ターゲットのリアルワールド座標を記述する付加情報をＨＭＤへ送信することも可能である。その結果、このシステムは、上記データを“リアルタイムで”取り出すことも可能であるため、上記ビーコンの位置を明白に認知する必要がなくなる。このようなシステムは、単純な逆反射作用を生じる構成よりもある程度頑丈な非常に汎用性の高いシステムとなる。パッシブなまたはアクティブなビーコンのいずれかの利用に加えて、双方の技法の組み合わせを用いることで、非常に汎用性の高いかつ強力なトラッキングシステムを形成することも可能である。逆反射作用を生じるターゲットの光源に対して“逆相”でアクティブな光ビーコンにそのストロボ光を発光させることにより、光点が互いに１８０°位相を異にしたも

のになるという事実に基づいて、差分画像解析技術を利用して、背景環境からの画像抽出と、（光を出射する）アクティブビーコンか（光を反射する）パッシブビーコンのいずれかから発するような光点の効果的類別との双方を行うようにすることが可能となる。ＶＴＶプロセッサによってアクティブビーコンのタイミングを生成することが可能であるが、パッシブビーコン用として使用されるＨＭＤからの光ストロボパルスの検出により、あるいは、ＨＭＤからの赤外線パルスのような２次的光タイミング信号の利用により、自動的にタイミングを発生させるほうが好ましいであろう。このような構成では、アクティブビーコンをバッテリで駆動することも可能であり、追加の配線は何ら必要ではなくなる。何らかの光タイミング信号が存在しないときにこれらのアクティブビーコンがパワーダウンするように設計されていれば、作動していない時間中上記アクティブビーコンを停止させることさえ不要となるかも知れず、さらに、バッテリ消費量の大幅な増加を伴うことなく上記アクティブビーコンを永久に作動状態のままにすることも可能である。

【００５６】

アクティブビーコンは、単にその位置を単に示すことに加えて、移動ＨＭＤシステムへ付加情報を送信することが可能である。この情報は、作成される光パルスの色の変化とこれらの色の変化するタイミングとによって符号化される。最も単純な形では、各アクティブビーコンが各々様々な色の光を発生させることにより、簡単な分離と識別とを行うことが可能であるが、さらに優れた符号化技法もまた可能である。例えば、一連のカラーフラッシュによって或る特定の数を示すことが可能である。この数によって、ビーコンのコード番号を表すことも可能であり、あるいはさらに有用なこととして、ビーコンの物理的リアルワールド座標の符号化も可能である。このようなシステムは、シミュレーションの一部としてＨＭＤへ転送しなければならない“リアルワールド”の“ビーコンマップ”を必要とせず、代わりにアクティブビーコンシステムを含む任意の空間から必要な情報をピックアップすることが可能なので、非常に汎用性の高いシステムとなる。

【００５７】

４．６）注：図２３は、ある特定の運動の有無が自己相関器により単純に検出されるトラッキングハードウェアの単純化した図を示す。実用的システムでは、各クラスの運動用の複数の自己相関器を組み込むことになるであろう（例えば、１６個またはそれ以上の、水平運動を検出する別々の自己相関器が存在してもよい）。このようなシステムでは、すべての方向で異なるレベルの運動や異なる量の運動を検出することが可能となる。

【００５８】

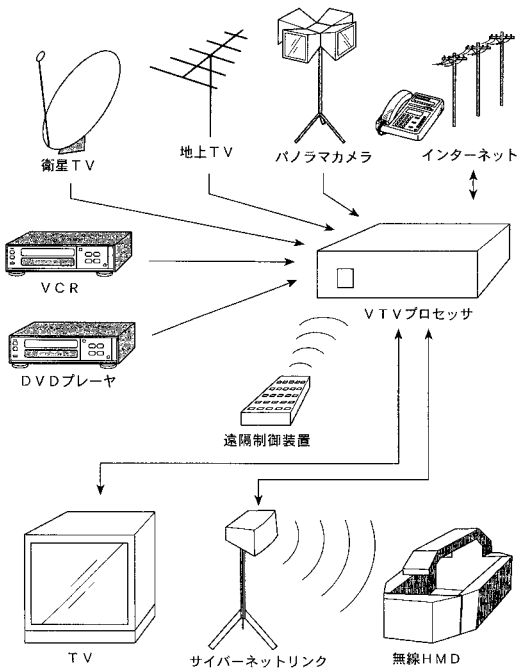
（その他の構成：）

４．７）空間測位情報を達成するために天井のパターンのトラッキングが行われる同様の画像解析技術と、ＨＭＤ/Ｐａｎ－Ｃａｍシステムの角度方向を検出する単純な“チルトセンサ”とを利用する上記トラッキングシステムの代替実施構成が可能である。本システムの利点として、前述した最大６本の軸線を有する光トラッカに比べて著しく構成がシンプルで、低コストで済むという点が挙げられる。天井がＨＭＤから一定の距離と所定の方角にあるという事実に起因して、光学系、必要な画像処理装置の品質、及び後続する画像解析の複雑さとが大幅に単純化される。前述の６軸線光トラッキングシステムの場合のように、上記空間測位情報は本質的に相対運動の形のみの情報である。しかし、“絶対基準点”を追加することによって、このようなシステムによってその絶対基準値を再較正して絶対座標系全体を達成することが可能となる。いくつかの異なる技法を利用して、比較的簡単に上記絶対基準点の較正を行うことが可能となる。第１のそしておそらく最も単純な技法として、前述したような色に敏感に感応する逆反射スポットの利用がある。上記とは別に、（ＬＥＤビーコンのような）アクティブな光ビーコンの利用も可能である。利用可能なさらに別の絶対基準較正システムとして双方向赤外線ビーコンをベースとするシステムもある。このようなシステムにより、ＨＭＤとビーコンとの間で一意のＩＤコードの通信が行われ、それによって、ＨＭＤが上記“既知の空間基準点”のうちのいずれかの下を通過する毎に１回だけ較正が行われることになる。これは、複数回の原点リセットに起因して生じる、較正ビーコンの近傍内にある“不感トラッキング領域”を回避するのに必要

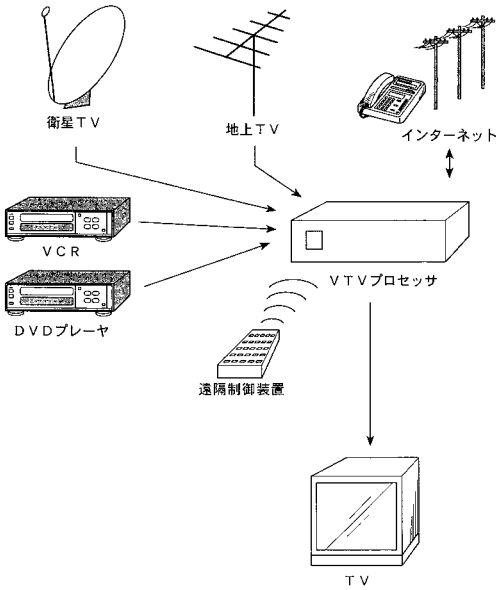
である。
【００５９】
（単純化：）
画像内の運動位置の検出に利用される基本的な自動相関技法を単純化して、かなり簡単な画像処理ステップにすることが可能である。まず、回転検出を単純化して、画像中心（カメラの光軸）の周りで左右対称を成す１グループの片寄り（上、下、左、右）にすることが可能である。さらに、横方向運動に関連するこれらの“サンプルポイント”は必ずしも非常に大きなものである必要はない。しかしこれらの“サンプルポイント”は固有の映像情報を含む必要がある。例えば、特徴のない空白な壁は有用なトラッキング情報をもたらさないが、オブジェクトの端縁や高輝度の強調点のような高いコントラスト領域を持つ画像は比較的簡単にトラックされる。この考え方をさらに一歩進めて、最初に画像全体を縮減して強調点／端縁に変えるようにすることが可能である。次いで、一連の水平方向及び垂直方向のストリップとして画像処理を行い、強調点／端縁間に自動相関領域の境界を設けるようにすることができる。さらに、現在の画像と前回の画像フレームとの比較を行い、これら画像間の“最も近い可能なフィット部分”（すなわち強調点の最小運動部分）を確定することによって、狭い強調領域のトラッキングを非常に簡単に行うことが可能となる。このような技法は、画像前処理のオーバーヘッドの若干をハードウェアにより処理すれば比較的簡単であり、適度の処理スピードを持つマイクロプロセッサの能力の範囲内に十分含まれる。
【００６０】
好ましい実施形態や特定の実施形態と関連して本発明を説明したが、本発明の範囲、あるいは本発明の発明概念から逸脱することなく、種々の変更及び追加の変形を行うことが可能であり、本発明の要素の代わりに同等の要素を用いることが可能であることを理解されたい。さらに、本発明の本質的範囲から逸脱することなく多くの改変を行って、特定の状況やマテリアルを本発明の教示に適合させることも可能である。したがって、本願明細書に開示された、本発明を実施するための特定の実施形態に本発明を限定することを意図するものではなく、本発明は添付の請求項の範囲に属するすべての実施形態例を含むものである。
【産業上の利用可能性】
【００６１】
本発明は、ユーザが映画の動きと関連して視聴者のビューイング方向と相対位置とを制御することを可能とする視聴覚システムを提供することを目的とする。
本発明の上記及び他の目的、利点及び産業上の利用可能性は添付の明細書と図面の再査により明らかである。
【図面の簡単な説明】
【００６２】
【図１】ＶＴＶシステム全体の概略図である。
【図２】基本構成に基づくＶＴＶシステムの概略図である。
【図３】高度の構成に基づくＶＴＶシステムの概略図である。
【図４】円筒形状のバーチャル表示フィールドを示す例図である。
【図５】先端を切り落とした球面のバーチャル表示フィールドを示す例図である。
【図６】４トラックサウンドシステムのバーチャル表現を示す例図である。
【図７】８トラックサウンドシステムのバーチャル表現を示す例図である。
【図８】拡張リアリティメモリとバーチャルリアリティメモリの双方のメモリを利用するシステムを表すＶＴＶメモリマップの描写である。
【図９】ＶＴＶプロセッサのデータ書き込みサイドを示すＶＴＶグラフィックエンジンの図である。
【図１０】ＶＴＶプロセッサのデータ読み出しサイドを示すＶＴＶグラフィックエンジンの図である。
【図１１】任意のアナログビデオと互換性を持つＶＴＶ符号化ビデオラインの例である。

【図１２】任意のアナログビデオと互換性を持つ音声データを含むことを示すＶＴＶ符号化ビデオラインの例である。
【図１３】周波数圧縮符号化を利用するＶＴＶにより符号化された音声を示す図である。
【図１４】周波数圧縮符号化を行うための結線図である。
【図１５】周波数圧縮復号化を行うための結線図である。
【図１６】位置と方角の変化を検出する光トラッキングシステムの図である。
【図１７】方角の方位角の変化を検出する光トラッキングシステムの図である。
【図１８】方角の高度の変化を検出する光トラッキングシステムの図である。
【図１９】方角の回転の変化を検出する光トラッキングシステムの結線図である。
【図２０】前後方向の位置変化を検出する光トラッキングシステムの結線図である。
【図２１】左右の位置変化を検出する光トラッキングシステムの結線図である。
【図２２】上下の位置変化を検出する光トラッキングシステムの結線図である。
【図２３】単純化したバージョンによる光トラッキングシステム用ハードウェアのブロック図である。

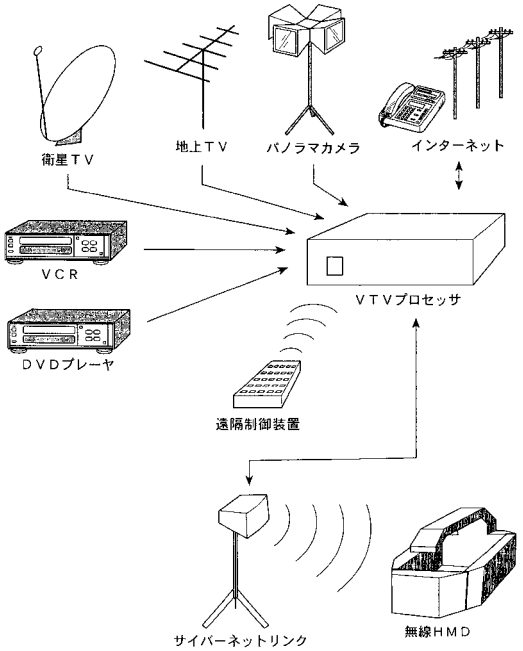
【図１】



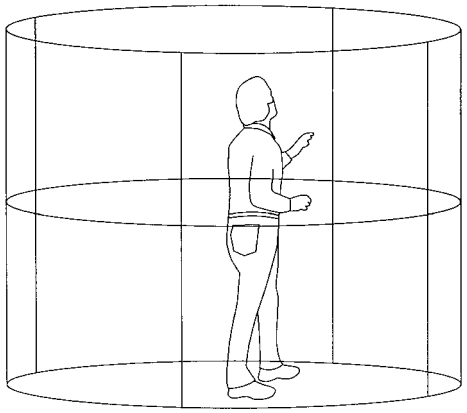
【図２】



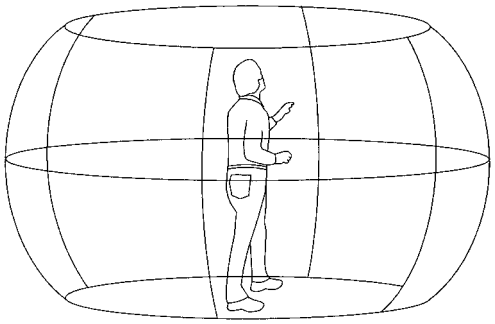
【図３】



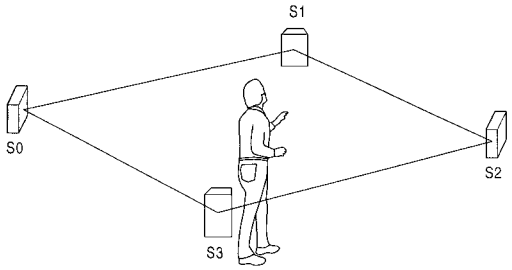
【図４】



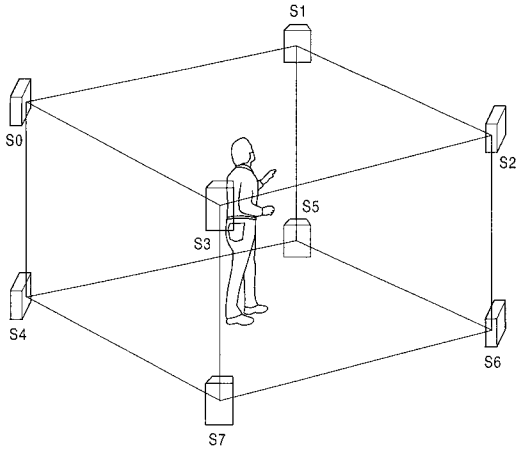
【図５】



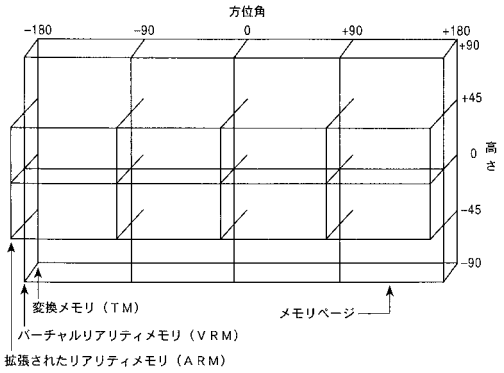
【図６】



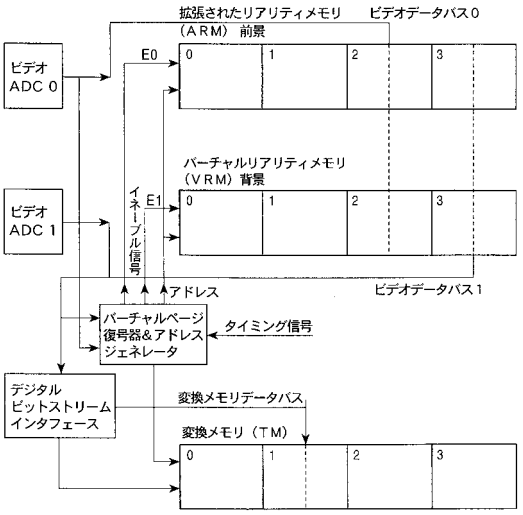
【図７】



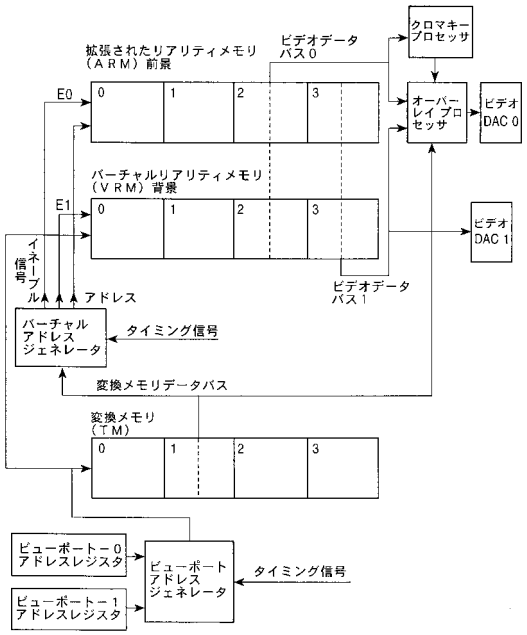
【図８】



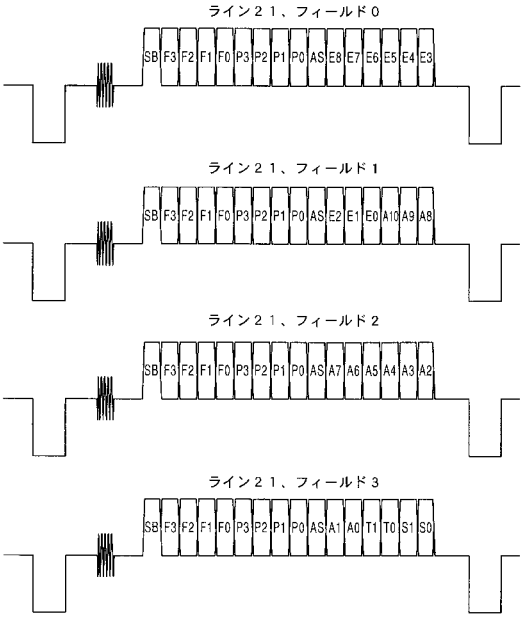
【図９】



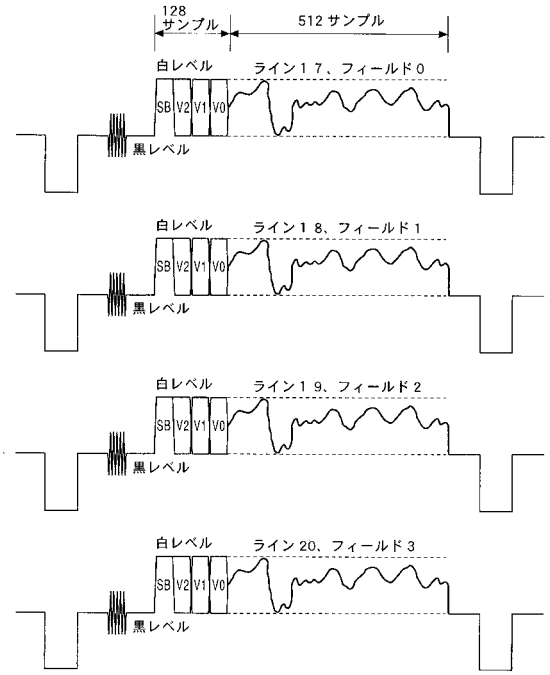
【図１０】



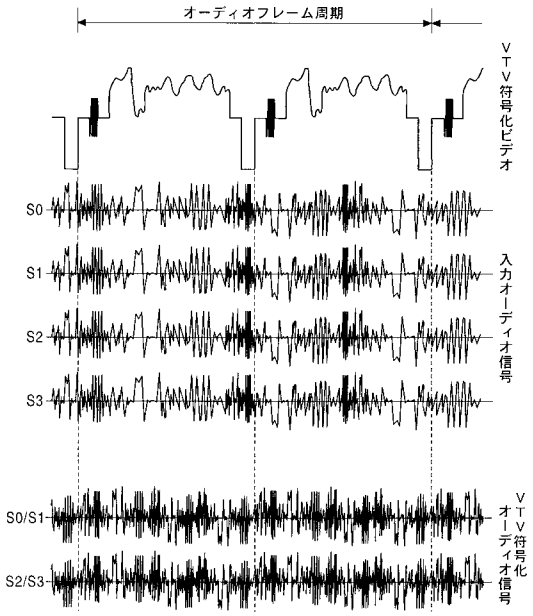
【図１１】



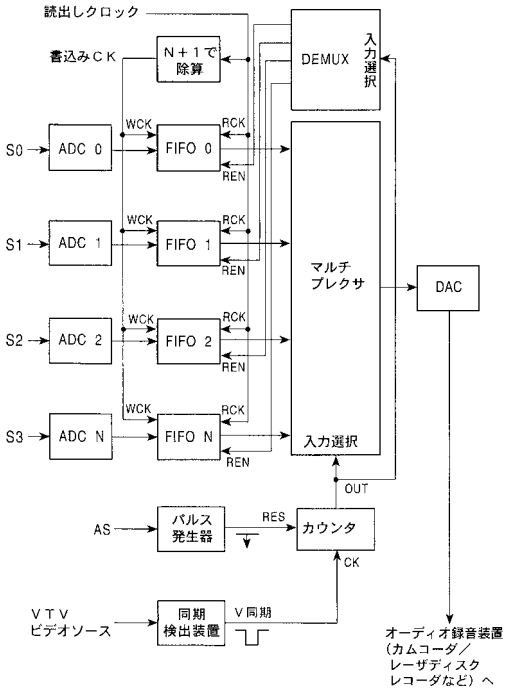
【図１２】



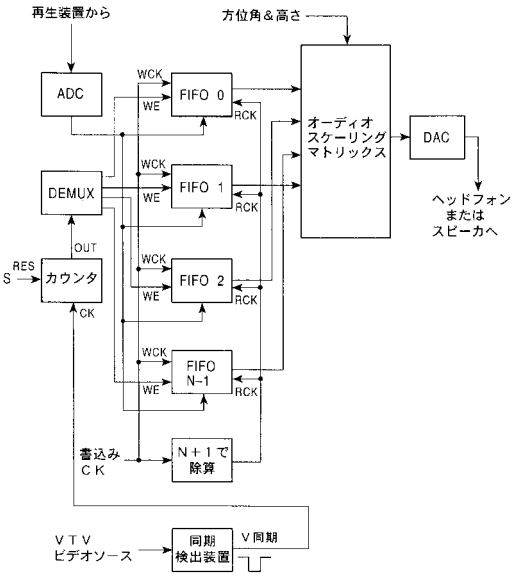
【図１３】



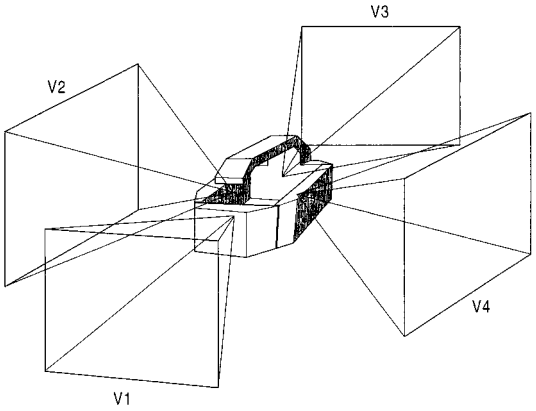
【図１４】



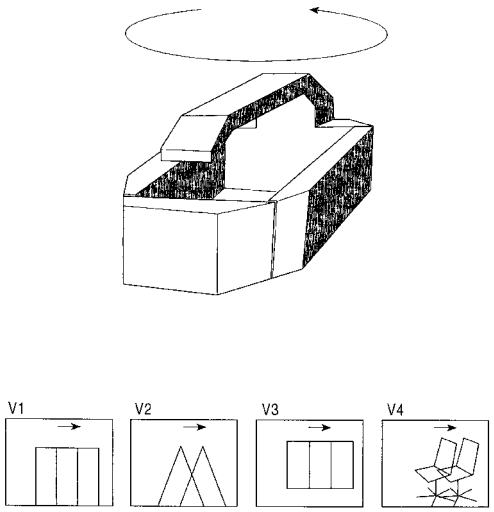
【図１５】



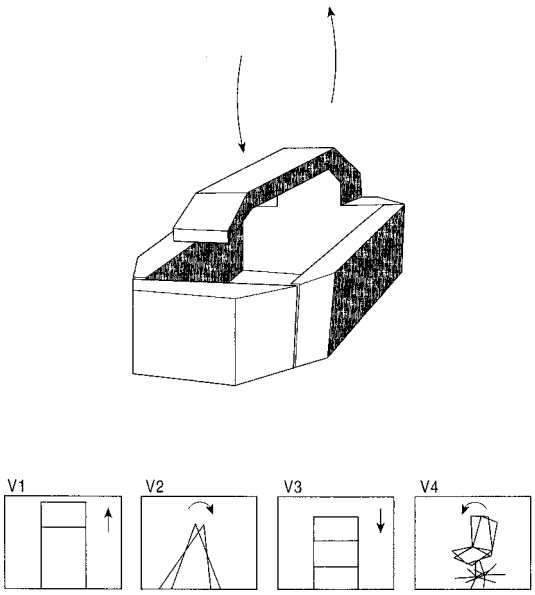
【図１６】



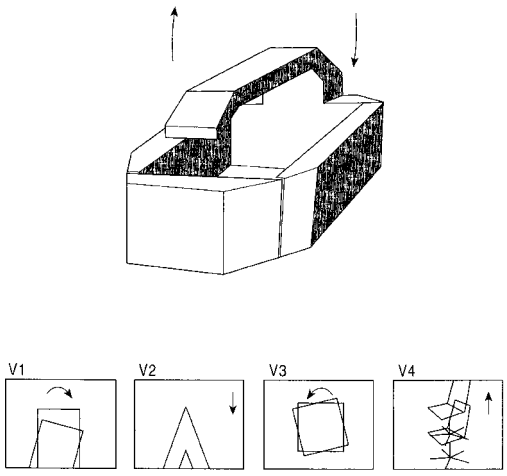
【図１７】



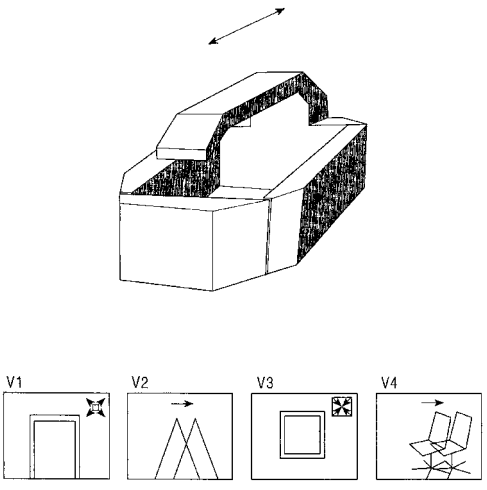
【図１８】



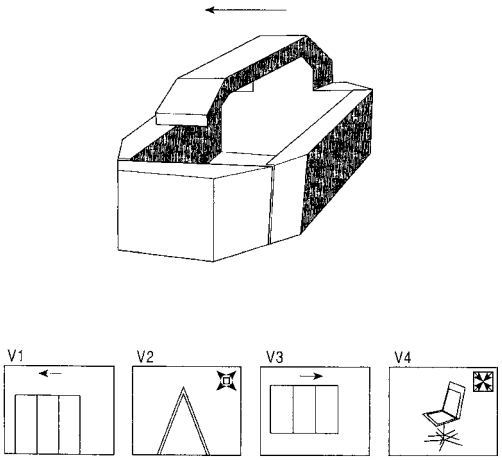
【図１９】



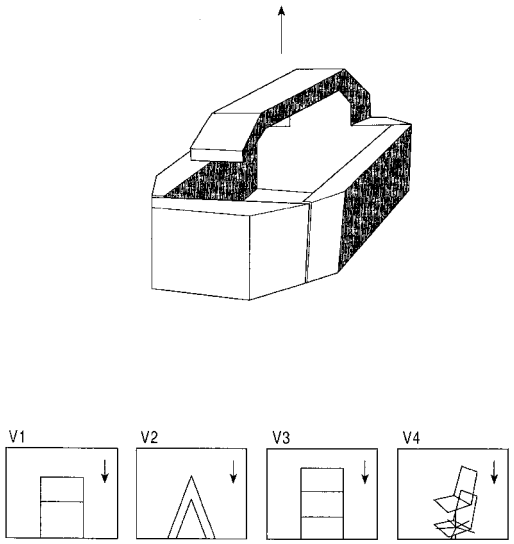
【図２０】



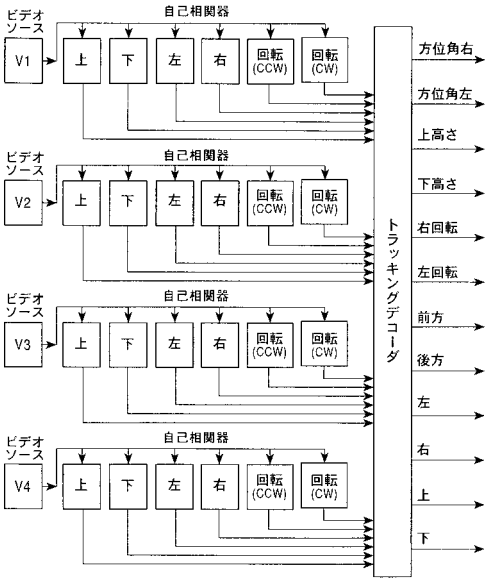
【図２１】


V1
V2
V3
V4

【図２２】


V1
V2
V3
V4

【図２３】


ビデオソース
V1
V2
V3
V4
自己相関器
上
下
左
右
回転(CCW)
回転(CW)
トラッキングデコーダ
方位角右
方位角左
上高さ
下高さ
右回転
左回転
前方
後方
左
右
上
下

【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 January 2003 (03.01.2003)

PCT

(10) International Publication Number
WO 03/001803 A1

(51) International Patent Classification[7]: H04N 7/00, 9/43, 9/64, G09G 5/00, G02B 27/22, 5/08, 5/20

(21) International Application Number: PCT/US01/49287

(22) International Filing Date:
21 December 2001 (21.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/891,733 25 June 2001 (25.06.2001) US

(71) Applicant and
(72) Inventor: RICHARDS, Angus [/US]; 5016 Kelly Street, Los Angeles, CA 90066 (US).

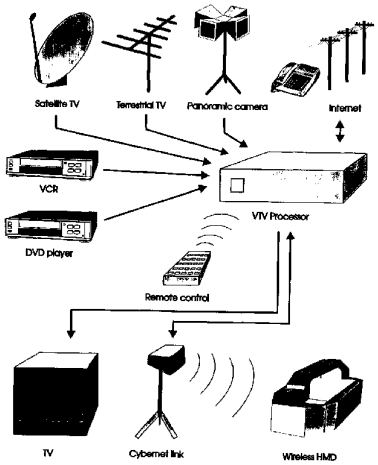
(74) Agents: CISLO, Daniel, M. et al.; Cislo & Thomas LLP, Suite 900, 233 Wilshire Boulevard, Santa Monica, CA 90401-1211 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: VTV SYSTEM



(57) Abstract: An overall hardware configuration that produces an enhanced spatial television-like viewing experience is disclosed (Fig. 1). Unlike normal television, with this system the viewer is able to control (remote control) both the viewing direction and relative position of the viewer with respect to the movie action (Fig. 6). In addition to a specific hardware configuration (S0, S1, S2, S3), this patent also relates to a new video format which makes possible this virtual reality like experience (Wireless HMD).

WO 03/001803 A1

WO 03/001803 A1

Published:
— *with international search report*
— *with amended claims*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

VTV SYSTEM

REFERENCE TO RELATED APPLICATION

This application claims priority from US utility application 09/891,733 entitled "VTV System" filed June 25, 2001 by Angus Duncan Richards.

TECHNICAL FIELD

This invention relates to image viewing and sound systems.

BACKGROUND ART

While normal television or movie screens are a common form of modern entertainment, these conventional systems do not allow a user to control the viewing direction and relative position of the viewer with respect to the movie action. Additionally, while virtual reality viewing is becoming increasingly popular, conventional systems for producing three dimensional images are often complex and the media graphics are costly to generate.

DISCLOSURE OF INVENTION

The following invention relates to an overall hardware configuration that produces an enhanced spatial television-like viewing experience. Unlike normal television, with this system the viewer is able to control both the viewing direction and relative position of the viewer with respect to the movie action. In addition to a specific hardware configuration, this invention also relates to a new video format which makes possible this virtual reality like experience. Additionally, several proprietary video compression standards are also defined which facilitate this goal. The VTV system is designed to be an intermediary technology between conventional two-dimensional cinematography and true virtual reality.

There are several stages in the evolution of the VTV system ranging from, in its most basic form, a panoramic display system to, in its most sophisticated form featuring full object based virtual reality utilizing animated texture maps and featuring live actors and/or computer-generated characters in a full "environment aware" augmented reality system.

BRIEF DESCRIPTION OF DRAWINGS

Fig. 1 is a schematic diagram of an overall VTV system.

Fig. 2 is a schematic diagram of of a VTV system according to its basic configuration.

Fig. 3 is a schematic diagram of a VTV system according to an advanced configuration.

Fig. 4 is an illustration of a cylindrical virtual display field.

Fig. 5 is an illustration of a truncated spherical virtual display field.

Fig. 6 is an illustration of a virtual representation of a 4 track sound system.

Fig. 7 is an illustration of a virtual representation of an 8 track sound system.

Fig. 8 is a dipiction of a VTV memory map for a system utilizing both augmented reality memory and virtual reality memory.

Fig. 9 is a VTV graphics engine diagram showing the data write side of the VTV processor.

Fig. 10 is a VTV graphics engine diagram showing the data read side of the VTV processor.

Fig. 11 is an example of any analogue video compatible VTV encoded video line..

Fig. 12 is an example of any analogue video compatible VTV encoded video line shown containing audio data.

Fig. 13 is diagram showing the VTV encoded audio using frequency compression encoding.

Fig. 14 is a diagram for the frequency compression encoding.

Fig. 15 is a diagram for the frequency compression decoding.

Fig. 16 is a diagram of an optical tracking system for detecting changes in position and orientation.

Fig. 17 is a diagram of an optical tracking system for detecting azimuth changes in orientation.

Fig. 18 is a diagram of an optical tracking system for detecting elevation changes in orientation.

Fig. 19 is a diagram of an optical tracking system for detecting roll changes in orientation.

Fig. 20 is a diagram of an optical tracking system for detecting forwards/backwards changes in position.

Fig. 21 is a diagram of an optical tracking system for detecting left/right changes in position.

Fig. 22 is a diagram of an optical tracking system for detecting up/down changes in position.

Fig. 23 is a block diagram of hardware for an optical tracking system according to a simplified version.

MODE(S) FOR CARRYING OUT THE INVENTION

The detailed description set forth below in connection with the appended drawings is intended as a description of presently-preferred embodiments of the invention and is not intended to represent the only forms in which the present invention may be constructed and/or utilized. The description sets forth the functions and the sequence of steps for constructing and operating the invention in connection with the illustrated embodiments. However, it is to be understood that the same or equivalent functions and sequences may be accomplished by different embodiments that are also intended to be encompassed within the spirit and scope of the invention.

1.1) The following patent relates to an overall hardware configuration that produces an enhanced spatial television-like viewing experience. Unlike normal television, with this system the viewer is able to control both the viewing direction and relative position of the viewer with respect to the movie action. In addition to a specific hardware configuration, this patent also relates to a new video format which makes possible this virtual reality like experience. Additionally, several proprietary video compression standards are also defined which facilitate this goal. The VTV system is designed to be an intermediary technology between conventional two-dimensional cinematography and true virtual reality. There are several stages in the evolution of the VTV system ranging from, in its most basic form, a panoramic display system to, in its most sophisticated form featuring full object based virtual reality utilizing animated texture maps and featuring live actors and/or computer-generated characters in a full "environment aware" augmented reality system.

1.2) As can be seen in fig 1 the overall VTV system consists of a central graphics processing device (the VTV processor), a range of video input devices (DVD, VCR, satellite, terrestrial television, remote video cameras),

infrared remote control, digital network connection and several output device connections. In its most basic configuration as shown in fig 2, the VTV unit would output imagery to a conventional television device. In such a configuration a remote control device (possibly infrared) would be used to control the desired viewing direction and position of the viewer within the VTV environment. The advantage of this "basic system configuration" is that it is implementable utilizing current audiovisual technology. The VTV graphics standard is a forwards compatible graphics standard which can be thought of as a "layer" above that of standard video. That is to say conventional video represents a subset of the new VTV graphics standard. As a result of this standard's compatibility, VTV can be introduced without requiring any major changes in the television and/or audiovisual manufacturers specifications. Additionally, VTV compatible television decoding units will inherently be compatible with conventional television transmissions.

1.3) In a more sophisticated configuration, as shown in fig. 3, the VTV system uses a wireless HMD as the display device. In such a configuration the wireless HMD can be used as a tracking device in addition to simply displaying images. This tracking information in the most basic form could consist of simply controlling the direction of view. In a more sophisticated system, both direction of view and position of the viewer within the virtual environment can be determined. Ultimately, in the most sophisticated implementation, remote cameras on the HMD will provide to the VTV system, real world images which it will interpret into spatial objects, the spatial objects can then be replaced with virtual objects thus providing an "environment aware" augmented reality system.

1.4) The wireless HMD is connected to the VTV processor by virtue of a wireless data link "Cybernet link". In its most basic form this link is capable of transmitting video information from the VTV processor to the HMD and transmitting tracking information from the HMD to the VTV processor. In its most sophisticated form the cybernet link would transmit video information both to and from the HMD in addition to transferring tracking information from the HMD to the VTV processor. Additionally certain components of the VTV processor may be incorporated in the remote HMD thus reducing the data transfer requirement through the cybernet link. This wireless data link can be implemented in a number of different ways utilizing either analog or digital video transmission (in either an un-compressed or a digitally compressed format) with a secondary digitally encoded data stream for tracking information. Alternately, a purely digital uni-directional or bi-directional data link which carries both of these channels could be incorporated. The actual medium for data transfer would probably be microwave or optical. However either transfer medium may be utilized as appropriate. The preferred embodiment of this system is one which utilizes on-board panoramic cameras fitted to the HMD in conjunction with image analysis hardware on board the HMD or possibly on the VTV base station to provide real-time tracking information. To further improve system accuracy, retroflective markers may also the utilized in the "real world environment". In such a configuration, switchable light sources placed near to the optical axis of the on-board cameras would be utilized in conjunction with these cameras to form a "differential image analysis" system. Such a system features considerably higher recognition accuracy than one utilizing direct video images alone.

1.5) Ultimately, the VTV system will transfer graphic information utilizing a "universal graphics standard". Such a standard will incorporate an object based graphics description language which achieves a high degree of

compression by virtue of a "common graphics knowledge base" between subsystems. This patent describes in basic terms three levels of progressive sophistication in the evolution of this graphics language.

1.6) These three compression standards will for the purpose of this patent be described as:

a) c-com
b) s-com
c) v-com

1.7) In its most basic format the VTV system can be thought of as a 360 Degree panoramic display screen which surrounds the viewer.

1.8) This "virtual display screen" consists of a number of "video Pages". Encoded in the video image is a "Page key code" which instructs the VTV processor to place the graphic information into specific locations within this "virtual display screen". As a result of this ability to place images dynamically it is possible to achieve the effective equivalent to both high-resolution and high frame rates without significant sacrifice to either. For example, only sections of the image which are rapidly changing require rapid image updates whereas the majority of the image is generally static. Unlike conventional cinematography in which key elements (which are generally moving) are located in the primary scene, the majority of a panoramic image is generally static.

VTV GRAPHICS STANDARD:

2.1) In its most basic form the VTV graphics standard consists of a virtual 360 degree panoramic display screen upon which video images can be rendered from an external video source such as VCR, DVD, satellite, camera or terrestrial television receiver such that each video frame contains not only the video information but also information that defines its location within the virtual display screen. Such a system is remarkably versatile as it provides not only variable resolution images but also frame rate independent imagery. That is to say, the actual update rate within a particular virtual image (entire virtual display screen) may vary within the display screen itself. This is inherently accomplished by virtue of each frame containing its virtual location information. This allows active regions of the virtual image to be updated quickly at the nominal perception cost of not updating sections on the image which have little or no change. Such a system is shown in fig 4.

2.2) To further improve the realism of the imagery, the basic VTV system can be enhanced to the format shown in fig 5. In this configuration the cylindrical virtual display screen is interpreted by the VTV processor as a truncated sphere. This effect can be easily generated through the use of a geometry translator or "Warp Engine" within the digital processing hardware component of the VTV processor.

2.3) Due to constant variation of absolute planes of reference, mobile camera applications (either HMD based or Pan-Cam based) require additional tracking information for azimuth and elevation of the camera system to be included with the visual information in order that the images can be correctly decoded by the VTV graphics engine. In such a system, absolute camera azimuth and elevation becomes part of the image frame information. There are several possible techniques for the interpretation of this absolute reference data. Firstly, the coordinate data could be used to define the origins of the image planes within the memory during the memory writing process. Unfortunately this approach will tend to result in remnant image fragments being left in memory from previous frames with different alignment values. A more practical solution is simply to write the video information into memory with an assumed reference point of 0 azimuth, 0 elevation. This video information is then correctly displayed by correcting the display viewport for the camera angular offsets. The data format for such a system is shown in fig 11.

CONTROL FIELD	BITS	VALUE	ASSIGNMENT KEY
FRAME FLIP	FF		FLIP MEMORY FRAMES
FIELD TYPE	F2-F0	0	FOREGROUND VIDEO (ARM)
		1	BACKGROUND VIDEO (VRM)
		2	DIGITAL HYBRID (TM)
		3	DIGITAL BIT STREAM FRAME (TM)
		4	RESERVED
		5	RESERVED
		6	RESERVED
		7	DIGITAL CONTROL FRAME
PAGE NUMBER	P3-P0	0-15	(DEPENDANT UPON MEM LAYOUT)
AUDIO SYNC	AS		RESET AUDIO BUFFER TO ZERO
ELEVATION CORRECTION	E8-E0	(+/- 45 DEG)	CAMERA ELEVATION
AZIMUTH CORRECTION	A10-A0	(+/- 180 DEG)	CAMERA AZIMUTH
AUDIO TRACKS	T1-T0	0	NO AUDIO TRACKS
		1	4 AUDIO TRACKS
		2	8 AUDIO TRACKS
		3	OBJECT BASED AUDIO
AUDIO SAMPLE RATE	S1-S0	0	2/4 LINES (15K S/S)
		1	3/6 LINES (23K S/S)
		2	4/8 LINES (31k S/S)
		3	5/10 LINES (38K S/S)

Table 1

AUDIO STANDARDS:

2.4) In addition to 360 Degree panoramic video, the VTV standard also supports either 4 track (quadraphonic) or 8 track (octaphonic) spatial audio. A virtual representation of the 4 track system is shown in fig 6. In the case of the simple 4 track audio system sound through the left and right speakers of the sound system (or headphones, in the case of an HMD based system) is scaled according to the azimuth the of the view port (direction of view within the VR environment). In the case of the 8 track audio system sound through the left and right speakers of the sound system (or headphones, in the case of an HMD based system) is scaled according to both the azimuth and elevation of the view port, as shown in the virtual representation of the system, fig 7.

2.5) In its most basic form, the VTV standard encodes the multi-track audio channels as part of the video information in a digital/analogue hybrid format as shown in fig 12. As a result, video compatibility with existing equipment can be achieved. As can be seen in this illustration, the audio data is stored in a compressed analogue coded format such that each video scan line contains 512 audio samples. In addition to this analogue coded audio information, each audio scan line contains a three bit digital code that is used to "pre-scale" the audio information. That is to say that the actual audio sample value is X*S where X is the pre-scale number and S is

the sample value. Using this dual-coding scheme the dynamic range of the audio system can be extended from about 43dB to over 60dB. Secondly, this extending of the dynamic range is done at relatively "low cost" to the audio quality because we are relatively insensitive to audio distortion when the overall signal level is high. The start bit is an important component in the system. It's function is to set the maximum level for the scan line (i.e. the 100% or white level) This level in conjunction with the black level (this can be sampled just after the colour burst) forms the 0% and 100% range for each line. By dynamically adjusting the 0% and 100% marks for each line on a line by line basis, the system becomes much less sensitive to variations in black level due to AC-coupling of video sub modules and/or recording and play back of the video media in addition to improving the accuracy of the decoding of the digital component of the scan line.

2.6) In addition to this pre-scaling of the digital information, an audio control bit (AR) is included in each field (at line 21). This control bit sets the audio buffer sequence to 0 when it is set. This provides a way to synchronize the 4 or 8 track audio information so that the correct track is always being updated from the current data regardless of the sequence of the video Page updates.

2.7) In an alternate configuration spatial audio can be provided by using the conventional stereo audio provided by video recording and playback systems (typically digital or linear analogue audio tracks). At best such conventional systems provide only two audio tracks with a 48Khz sample rate (ie 2 channel 24 Khz audio). However it is possible to encode a greater number or audio tracks (for example 4 or 8 audio tracks) onto the same storage medium by compressing the audio data in the time domain and then time division multiplexing the resultant data. At the receive side the process is reversed by first decoding the time division multiplexed data and then by decompressing the audio data to form a series of continuous audio streams. It is preferable to incorporate a high order notch filter or high pass filter to remove any sampling noise introduced by the time division multiplexing/de-multiplexing procedure. Such a filter would probably be of digital design and would be simply part of the digital processing hardware of the VTV Processor. A typical configuration for such an audio encoding system is shown in figs. 13-15.

Clearly such a system is easily implementable. The "performance cost" of increasing the number of audio channels is a reduction in the frequency response of the system. This reduction is directly proportional to the increase in channels. For example, increasing the number of channels from 2 to 4 will reduce the frequency response from 24 Khz to 12 Khz. This level of reduction is acceptable and ultimately this Audio Multiplexing system may be preferable to the previously described Video encoded Audio system in that it is relatively "robust" and offers significant advantages in terms of ease of implementation and editing of VTV format material.

One of the most critical factors in the successful implementation of any time division multiplexing system is the generation of a stable and accurate time division encoding and decoding clock. This is absolutely critical because any miss-alignment of the decoded data packets will result in a very noticeable channel crosstalk related noise. The VTV system incorporates a novel and very effective solution to this problem by using the Sync signals in the video stream (either H_Sync or V_Sync, Odd/Even etc.) to generate the local decode clock. These signals are generally very accurate, particularly if digital recording equipment is utilized for the storage of the

VTV data.

The preferred embodiment will probably use V_Sync as the decode clock given that it produces relatively little time division decoding noise (60 spikes per second for NTSC) and that this noise can probably easily removed through the use of a digital notch filter (if required). If on the other hand a high frequency time division clock is chosen (such as H_Sync) more instances of discontinuity are possible (approx. 15,000 per second) but once again, these can be removed by a digital filter (a low pass filter in this case). If HDTV or other high scan rate (31 Khz or above) TV standard is used a fast decode clock would probably become preferable because the filtering requirement becomes very easy and effective in such a circumstance. In addition, the memory requirements for a system utilizing a fast decode clock are significantly reduced.

2.8) In more sophisticated multimedia data formats such as computer AV. files and digital television transmissions, these additional audio tracks could be stored in other ways which may be more efficient or otherwise advantageous.

2.9) It should be noted that, in addition to it's use as an audiovisual device. this spatial audio system/standard could also be used in audio only mode by the combination of a suitable compact tracking device and a set of cordless headphones to realize a spatial-audio system for advanced hi-fi equipment.

ENHANCEMENTS:

2.10) In addition to this simplistic graphics standard, There a are number of enhancements which can be used alone or in conjunction with the basic VTV graphics standard. These three graphics standards will be described in detail in subsequent patents, however for the purpose of this patent, they are known as:

a) c-com
b) s-com
c) v-com

2.11) The first two standards relate to the definitions of spatial graphics objects where as the third graphics standard relates to a complete VR environment definition language which utilizes the first standards as a subset and incorporates additional environment definitions and control algorithms.

2.12) The VTV graphic standard (in its basic form) can be thought of as a control layer above that of the conventional video standard (NTSC, PAL etc.). As such, it is not limited purely to conventional analog video transmission standards. Using basically identical techniques, the VTV standard can operate with the HDTV standard as well as many of the computer graphic and industry audiovisual standards.

2.13) The digital control information outlined in Table 1 and fig 11 represent only one possible configuration of the system for compatibility with existing analogue TV systems and should not reduce the generality of the

specification. Other configurations of control bits and/or ways to encode this information into video formats such as analogue TV exist and are equally valid. For example, it has been found that many digital encoding processes such as MJPEG and MPEG-II often average consecutive video scan lines and tend to damage the digital information when encoded as a single line as shown in fig 11. However if this information is encoded as a series or consecutive lines this problem is overcome. In such a configuration the data can be thought of as having been encoded as a series of small rectangles in the resultant image, rather than a series of line segments. Such modifications are expected and still further adaptations are expected to improve performance efficiency with other video formats, particularly the digitally encoded formats such as HDTV etc.

VTV PROCESSOR:

3.1) The VTV graphics processor is the heart of the VTV system. In its most basic form this module is responsible for the real-time generation of the graphics which is output to the display device (either conventional TV/HDTV or HMD). In addition to digitizing raw graphics information input from a video media provision device such as VCR, DVD, satellite, camera or terrestrial television receiver. More sophisticated versions of this module may real-time render graphics from a "universal graphics language" passed to it via the Internet or other network connection. In addition to this digitizing and graphics rendering task, the VTV processor can also perform image analysis. Early versions of this system will use this image analysis function for the purpose of determining tracking coordinates of the HMD. More sophisticated versions of this module will in addition to providing this tracking information, also interpret the real world images from the HMD as physical three-dimensional objects. These three-dimensional objects will be defined in the universal graphics language which can then be recorded or communicated to similar remote display devices via the Internet or other network or alternatively be replaced by other virtual objects of similar physical size thus creating a true augmented reality experience.

3.2) The VTV hardware itself consists of a group of sub modules as follows:

a) video digitizing module
b) Augmented Reality Memory (ARM)
c) Virtual Reality Memory (VRM)
d) Translation Memory (TM)
e) digital processing hardware
f) video generation module

3.3) The exact configuration of these modules is dependent upon other external hardware. For example, if digital video sources are used then the video digitizing module becomes relatively trivial and may consist of no more than a group of latch's or FIFO buffer. However, if composite or Y/C video inputs are utilized then additional hardware is required to convert these signals into digital format. Additionally, if a digital HDTV signal is used as the video input source then an HDTV decoder is required as the front end of the system (as HDTV signals cannot be processed in compressed format).

3.4) In the case of a field based video system such as analogue TV, the basic operation of the VTV graphics engine is as follows:

a) Video information is digitized and placed in the augmented reality memory on a field by field basis assuming an absolute Page reference of 0 degree azimuth, 0 degree elevation with the origin of each Page being determined by the state of the Page number bits (P3-P0).

b) Auxiliary video information for background and/or floor/ceiling maps is loaded into the virtual reality

memory on a field by field basis dependent upon the state of the "field type" bits (F3-F0) and Page number bits (P3-P0) or generated in real time from an interpretation of the video information placed in the augmented reality memory as described in section 3.12.

c) The digital processing hardware interprets this information held in augmented reality and virtual reality memory and utilizing a combination of a geometry processing engine (Warp Engine), digital subtractive image processing and a new versatile form of "blue-screening", translates and selectively combines this data into an image substantially similar to that which would be seen by the viewer if they were standing in the same location as that of the panoramic camera when the video material was filmed. The main differences between this image and that available utilizing conventional video techniques being that it is not only 360 degree panoramic but also has the ability to have elements of both virtual reality and "real world" imagery melded together to form a complex immersive augmented reality experience.

d) The exact way in which the virtual reality and "real world imagery" is combined depends upon the mode that the VTV processor is operating in and is discussed in more detail in later sections of this specification. The particular VTV processor mode is determined by additional control information present in the source media and thus the processing and display modes can change dynamically while displaying a source of VTV media.

e) The video generation module then generates a single or pair of video images for display on a conventional television or HMD display device. Although the VTV image field will be updated at less than full frame rates (unless multi-spin DVD devices are used as the image media) graphics rendering will still occur at full video frame rates, as will the updates of the spatial audio. This is possible because each "Image Sphere" contains all of the required information for both video and audio for any viewer orientation (azimuth and elevation).

3.5) As can be seen in fig 9. The memory write side of the VTV processor shows two separate video input stages (ADC's). It should be noted that although ADC-0 would generally be used for live panoramic video feeds and ADC-2 would generally be used for virtual reality video feeds from pre-rendered video material, both video input stages have full access to both augmented reality and virtual reality memory (i.e. they use a memory pool). This hardware configuration allows for more versatility in the design and allows several unusual display modes (which will be covered in more detail in later sections). Similarly, the video output stages (DAC-0 and DAC-1) have total access to both virtual and augmented reality memory.

3.6) Although having two input and two output stages improves the versatility of the design, the memory pool style of design means that the system can function with either one or two input and/or output stages (although with reduced capabilities) and as such the presence of either one or two input or output stages in a particular implementation should not limit the generality of the specification.

3.7) For ease of design, high-speed static RAM was utilized as the video memory in the prototype device. However, other memory technologies may be utilized without limiting the generality of the design specification.

3.8) In the preferred embodiment, the digital processing hardware would take the form of one or more field programmable logic arrays or custom ASIC. The advantage of using field programmable logic arrays is that the hardware can be updated at anytime. The main disadvantage of this technology is that it is not quite as fast as an ASIC. Alternatively, high-speed conventional digital processors may also be utilized to perform this image analysis and/or graphics generation task.

3.9) As previously described, certain sections of this hardware may be incorporated in the HMD, possibly even to the point at which the entire VTV hardware exists within the portable HMD device. In such a case the VTV base station hardware would act only as a link between the HMD and the Internet or other network with all graphics image generation, image analysis and spatial object recognition occurring within the HMD itself.

3.10) Note: The low order bits of the viewport address generator are run through a look up table address translator for the X and Y image axies which impose barrel distortion on the generated images. This provides the correct image distortion for the current field of view for the viewport. This hardware is not shown explicitly in fig 10 because it will probably be implemented within an FPGA or ASIC logic and thus comprises a part of the viewport address generator functional block. Likewise roll of the final image will likely be implemented in a similar fashion.

3.11) It should be noted that only viewport-0 is affected by the translation engine (Warp Engine), Viewport-1 is read out undistorted. This is necessary when using the superimpose and overlay augmented reality modes because VR-video material being played from storage has already been "flattened" (i.e. pincushion distorted) prior to being stored whereas the live video from the panoramic cameras on the HMD require distortion correction prior to being displayed by the system in Augmented Reality mode. After this preliminary distortion, images recorded by the panoramic cameras in the HMD should be geometrically accurate and suitable for storage as new VR material in their own right (i.e. they can become VR material). One of the primary roles of the Warp Engine is then to provide geometry correction and trimming of the panoramic camera's on the HMD. This includes the complex task of providing a seamless transition between camera views.

EXCEPTION PROCESSING:

3.12) As can be seen in figs. 4,5 a VTV image frame consists of either a cylinder or a truncated sphere. This space subtends only a finite vertical angle to the viewer (+/- 45 degrees in the prototype). This is an intentional limitation designed to make the most of the available data bandwidth of the video storage and transmission media and thus maintain compatibility with existing video systems. However, as a result of this compromise, there can exist a situation in which the view port exceeds the scope of the image data. There are several different ways in which this exception can be handled. Firstly, the simplest way to handle this exception is to simply make out of bounds video data black. This will give the appearance of being in a room with a black ceiling and floor. Alternatively, a more sophisticated approach is to produce a "floor" and "ceiling" ie space below and above the VTV image frame based upon image information present in the bottom and top of the

VTV image frame. In one particular implementation, an average of the red, green and blue pixels for the top and bottom "rim" of the VTV image frame could be used to produce either a pure or preferably a graded colour surface to represent the "ceiling" and "floor" of the environment. In a situation of VTV video generated outdoors this would typically result in the ceiling being a shade of blue and the floor being approximately the colour of the ground. The realism of the "floor/ceiling" generation is improved by varying the shading with viewport elevation, typically so that the brightness increases as the elevation of the viewport gets closer to the extremeness of +/- 90 degrees. This effect can easily be generated through the use of relatively simple mathematics and would probably be implemented in hardware as part of the VTV graphics engine

However, a further alternative and preferable configuration is to use a secondary video memory store to store a full 360 degree * 180 degree background image map at reduced resolution. This memory area is known as Virtual reality memory (VRM). The basic memory map for the system utilizing both augmented reality memory and virtual reality memory (in addition to translation memory) is shown in fig 8. As can be seen in this illustration, The translation memory area must have sufficient range to cover a full 360 degree * 180 degrees and ideally have the same angular resolution as that of the augmented reality memory bank (which covers 360 degree * 90 degree). With such a configuration, it is possible to provide both floor and ceiling exception handling and variable transparency imagery such as looking through windows in the foreground and showing the background behind them. The backgrounds can be either static or dynamic and can be updated in basically the same way as foreground (augmented reality memory) by utilizing a Paged format.

MODES OF OPERATION:

3.13) The VTV system has two basic modes of operation. Within these two modes there also exist several sub modes. The two basic modes are as follows:

a) Augmented reality mode
b) Virtual reality mode

AUGMENTED REALITY MODE 1:

3.14) In augmented reality mode 1, selective components of "real world imagery" are overlaid upon a virtual reality background. In general, this process involves first removing all of the background components from the "real world" imagery. This can be easily done by using differential imaging techniques. I.e. by comparing current "real world" imagery against a stored copy taken previously and detecting differences between the two. After the two images have been correctly aligned, the regions that differ are new or foreground objects and those that remain the same are static background objects. This is the simplest of the augmented reality modes and is generally not sufficiently interesting as most of the background will be removed in the process. It should be

noted that, when operated in mobile Pan-Cam (telepresense) or augmented reality mode the augmented reality memory will generally be updated in sequential Page order (i.e. updated in whole system frames) rather than random Page updates. This is because constant variations in the position and orientation of the panoramic camera system during filming will probably cause mis-matches in the image Pages if they are handled separately.

AUGMENTED REALITY MODE 2:

3.15) Augmented reality mode 2 differs from mode 1 in that, in addition to automatically extracting foreground and moving objects and placing these in an artificial background environment, the system also utilizes the Warp Engine to "push" additional "real world" objects into the background. In addition to simply adding these "real world" objects into the virtual environment the Warp Engine is also capable of scaling and translating these objects so that they match into the virtual environment more effectively. These objects can be handled as opaque overlays or transparencies.

AUGMENTED REALITY MODE 3:

3.16) Augmented reality mode 3 differs from the mode 2 in that, in this case, the Warp Engine is used to "pull" the background objects into the foreground to replace "real world" objects. As in mode 2, these objects can be translated and scaled and can be handled as either opaque overlays or transparencies. This gives the user to the ability to "match" the physical size and position of a "real world" object with a virtual object. By doing so, the user is able to interact and navigate within the augmented reality environment as they would in the "real world" environment. This mode is probably the most likely mode to be utilized for entertainment and gaming purposes as it would allow a Hollywood production to be brought into the users own living room.

ENHANCEMENTS:

3.17) Clearly the key to making augmented reality modes 2 and 3 operate effectively is a fast and accurate optical tracking system. Theoretically, it is possible for the VTV processor to identify and track "real world" objects in real-time. However, this is a relatively complex task, particularly as object geometry changes greatly with changes in the viewer's physical position within the "real world" environment, and as such, simple auto correlation type tracking techniques will not work effectively. In such a situation, tracking accuracy can be greatly improved by placing several retroflective targets on key elements of the objects in question. Such retroflective targets can easily be identified by utilizing relatively simple differential imaging techniques.

VIRTUAL REALITY MODE:

3.18) Virtual reality mode is a functionally simpler mode than the previous augmented reality modes. In this mode "pre-filmed" or computer-generated graphics are loaded into augmented reality memory on a random Page by Page basis. This is possible because the virtual camera planes of reference are fixed. As in the previous examples, virtual reality memory is loaded with a fixed or dynamic background at a lower resolution. The use of both foreground and background image planes makes possible more sophisticated graphics techniques such as motion parallax.

ENHANCEMENTS:

3.19) The versatility of virtual reality memory (background memory) can be improved by utilizing an enhanced form of "blue-screening". In such a system, a sample of the "chroma-key" color is provided at the beginning of each scan line in the background field. This provides a versatile system in which any color is allowable in the image. Thus, by surrounding individual objects with the "transparent" chroma-key color, problems and inaccuracies associated with the "cutting and pasting" of this object by the Warp Engine are greatly reduced. Additionally, the use of "transparent" chroma-keyed regions within foreground virtual reality images allows easy generation of complex sharp edged and/or dynamic foreground regions with no additional information overhead.

THE CAMERA SYSTEM:

4.1) As can be seen in the definition of the graphic standard, additional Page placement and tracking information is required for the correct placement and subsequent display of the imagery captured by mobile Pan-Cam or HMD based video systems. Additionally, if Spatial audio is to be recorded in real-time then this information must also be encoded as part of the video stream. In the case of computer-generated imagery this additional video information can easily be inserted at render-stage. However, in the case of live video capture, this additional tracking and audio information must be inserted into the video stream prior to recording. This can effectively be achieved through a graphics processing module herein after referred to as the VTV encoder module.

IMAGE CAPTURE:

4.2) In the case of imagery collected by mobile panoramic camera systems, the images are first processed by a VTV encoder module. This device provides video distortion correction and also inserts video Page information, orientation tracking data and spatial audio into the video stream. This can be done without altering the video standard, thereby maintaining compatibility with existing recording and playback devices. Although this module could be incorporated within the VTV processor, having this module as a separate entity is advantageous for use in remote camera applications where the video information must ultimately be either stored or transmitted through some form of wireless network

TRACKING SYSTEM:

4.3) For any mobile panoramic camera system such as a "Pan-Cam" or HMD based camera system, tracking information must comprise part of the resultant video stream in order that an "absolute" azimuth and elevation coordinate system be maintained. In the case of computer-generated imagery this data is not required as the camera orientation is a theoretical construct known to the computer system at render time.

THE BASIC SYSTEM:

4.4) The basic tracking system of the VTV HMD utilizes on-board panoramic video cameras to capture the required 360 degree visual information of the surrounding real world environment. This information is then analyzed by the VTV processor (whether it exists within the HMD or as a base station unit) utilizing computationally intensive yet relatively algorithmically simple techniques such as auto correlation. Examples of a possible algorithm are shown in figs 16-22.

4.5) The simple tracking system outlined in figs 16-22 detects only changes in position and orientation. With the addition of several retroflective targets, which can be easily distinguished from the background images using differential imaging techniques, it is possible to gain absolute reference points. Such absolute reference points would probably be located at the extremities of the environmental region (i.e. confines of the user space) however they could be placed anywhere within the real environment, provided the VTV hardware is aware of the real world coordinates of these markers. The combination of these absolute reference points and differential movement (from the image analysis data) makes possible the generation of absolute real world coordinate information at full video frame rates. As an alternative to the placement of retroflective targets at known spatial coordinates, active optical beacons could be employed. These devices would operate in a similar fashion to the retroflective targets in that they would be configured to strobe light in synchronism with the video capture rate thus allowing differential video analysis to be performed on the resultant images. However, unlike passive retroflective targets, active optical beacons could, in addition to strobing in time with the video capture, transmit additional information describing their real world coordinates to the HMD. As a result, the system would not have to explicitly know the locations of these beacon's as this data could be extracted "on the fly". Such as system is very versatile and somewhat more rugged than the simpler retroflective configuration. In addition to utilizing either passive or active beacons a combination of both techniques could be used to produce a very versatile and powerful tracking system. By making the active optical beacons strobe their light in "anti-phase" with the light source of retroflective targets differential image analysis techniques can be used to both extract the images from the background environment and to effectively categorize the points of light as originating from either active beacons (those emitting light) and passive beacons (those reflecting light) by virtue of the fact that the points of light will be 180 degrees out of phase with each other. Timing of the active beacons can be generated by the VTV Processor but it is probably preferable to generate the timing automatically by detecting the light strobe pulses from the HMD used for the passive beacons or alternately by utilizing a secondary optical timing signal such as infra-red pulses from the HMD. In such a configuration, the active beacons could be battery powered and would not require any additional wiring. If these active beacons were designed to power-

down in the absence of any optical timing signals they may not even be required to be shut down during periods of inactivity and could remain permanently enabled without significantly increasing battery consumption.

Active beacons can, in addition to simply indicating their position, transfer additional information to the mobile HMD system. This information is encoded through changing the color of the light pulses produced and the timing of these color changes. In its simplest form each active beacons could each produce a different color of light to allow easy separation and identification, however more sophisticated encoding techniques are also possible. For example, a sequence of colour flashes can indicate a particular number. This number could represent the beacon code number, or more usefully, the physical real-world coordinates of the beacon could be encoded. Such a system would then be very versatile as it would not require a "beacon map" of the "real-world" to be transferred to the HMD as part of the simulation but rather could pick up the required information from any space which contained active beacon systems.

4.6) Note: fig 23 shows a simplistic representation of the tracking hardware in which the auto correlators simply detect the presence or absence of a particular movement. A practical system would probably incorporate a number of auto correlators for each class of movement (for example there may be 16 or more separate auto correlators to detect horizontal movement). Such as system would then be able to detect different levels or amounts of movement in all of the directions.

ALTERNATE CONFIGURATIONS:

4.7) An alternative implementation of this tracking system is possible utilizing a similar image analysis technique to track a pattern on the ceiling to achieve spatial positioning information and simple "tilt sensors" to detect angular orientation of the HMD/Pan-Cam system. The advantage of this system is that it is considerably simpler and less expensive than the full six axis optical tracker previously described. The fact that the ceiling is at a constant distance and known orientation from the HMD greatly simplifies the optical system, the quality of the required imaging device and the complexity of the subsequent image analysis. As in the previous six-axis optical tracking system, this spatial positioning information is inherently in the form of relative movement only. However, the addition of "absolute reference points" allows such a system to re-calibrate its absolute references and thus achieve an overall absolute coordinate system. This absolute reference point calibration can be achieved relatively easily utilizing several different techniques. The first, and perhaps simplest technique is to use color sensitive retroflective spots as previously described. Alternately, active optical beacon's (such as LED beacon's) could also be utilized. A further alternative absolute reference calibration system which could be used is based on a bi-directional infrared beacon. Such as system would communicate a unique ID code between the HMD and the beacon, such that calibration would occur only once each time the HMD passed under any of these "known spatial reference points". This is required to avoid "dead tracking regions" within the vicinity of the calibration beacons due to multiple origin resets.

SIMPLIFICATIONS:

The basic auto correlation technique used to locate movement within the image can be simplified into reasonably straightforward image processing steps. Firstly, rotation detection can be simplified into a group of lateral shifts (up, down, left, right) symmetrical around the center of the image (optical axis of the camera). Additionally, these "sample points" for lateral movement do not necessarily have to be very large. They do however have to contain unique picture information. For example a blank featureless wall will yield no useful tracking information However an image with high contrast regions such as edges of objects or bright highlight points is relatively easily tracked. Taking this thinking one step further, it is possible to first reduce the entire image into highlight points/edges. The image can then be processed as a series of horizontal and vertical strips such that auto correlation regions are bounded between highlight points/edges. Additionally, small highlight regions can very easily be tracked by comparing previous image frames against current images and determining "closest possible fit" between the images (i.e., minimum movement of highlight points). Such techniques are relatively easy and well within the capabilities of most moderate speed micro-processors, provided some of the image pre-processing overhead is handled by hardware.

While the present invention has been described with reference to a preferred embodiment or to particular embodiments, it will be understood that various changes and additional variations may be made and equivalents may be substituted for elements thereof without departing from the scope of the invention or the inventive concept thereof. In addition, many modifications may be made to adapt a particular situation or material to the teachings of the invention without departing from the essential scope thereof. Therefore, it is intended that the invention not be limited to particular embodiments disclosed herein for carrying it out, but that the invention includes all embodiments falling within the scope of the appended claims.

INDUSTRIAL APPLICABILITY

It is an object of the present invention to provide an audiovisual system which allows a user to control the viewing direction and relative position of the viewer with respect to the movie action.

These and other objects, advantages, and the industrial utility of the present invention will be apparent from a review of the accompanying specification and drawings.

CLAIMS

What is claimed is:

1) A method for producing an enhanced spatial television like viewing experience utilizing conventional video devices for the provision of the source media comprising:

updating an overall image sphere on a page by page basis in real-time utilizing conventional video devices for a provision of the source media,

producing graphical imagery depicting a panoramic image composed of a plurality of smaller image subsections.

2) The method of Claim 1 wherein said panoramic image is a 360 degree horizontal view.

3) The method of Claim 1 further comprising determining a page order by additional information present in the source media.

4) The method of Claim 1 wherein a viewer can view prerecorded audiovisual media in a wide screen format such that the width of the virtual screen can extend to a full 360 degrees horizontally and up to 180 degrees vertically.

5) The method of Claim 1 wherein a viewer can view prerecorded audiovisual material on a conventional screen based display device including TV, projection TV, or computer screen, such that the display device represents a viewport or subset of the full 360 degree panoramic image.

6) An entertainment system comprising:

a range of alternative media provision devices, the alternative media provision devices comprising a VCR, DVD, or satellite receiver,

an electronic device, (VTV processor), which generates panoramic video imagery from video data provided from the devices, and

a display device such comprising a conventional flat screen television, a helmet mounted display device (HMD), or other virtual reality display device, fitted with an optional single view or panoramic video capture device, in conjunction with a wireless data communication network to communicate the video information between the HMD and the VTV processor.

7) The system of Claim 6 wherein the electronic device is a VTV processor.

8) A method for producing audiovisual effects comprising:

modifying conventional television standard to produce a new audiovisual standard, the new audiovisual standard allowing for a variety of different frames, the frames containing graphical data, sound information, control information, or a combination thereof, while maintaining compatibility with existing television standards.

9) The method of Claim 8 wherein the new audiovisual standard includes, within at least one scan line of a standard video image, additional digital coded data, analogue coded data, or a combination thereof for providing information which defines control parameters and image manipulation data for a VTV graphics processor.

10) The method of Claim 8 wherein the scan line further includes additional digital data and analogue coded data which provides information to generate four or more audio tracks in real-time.

11) An electronic device that produces an enhanced spatial television like viewing experience utilizing conventional video devices for the provision of the source media.

12) An electronic device that produces graphical imagery depicting a panoramic (360 degree horizontal view) image such that this overall panoramic image ("Image Sphere") is composed of a number of smaller image subsections ("Pages").

13) An electronic device that produces graphical imagery as described in claims 11-12, such that the overall Image Sphere is updated on a Page by Page basis in real-time utilizing conventional video devices for the provision of the source media.

14) An electronic device that is described in claims 11-13 in which the Page order is determined by additional information present in the source media.

15) An electronic device as described in claims 11-14, which allows the viewer to view prerecorded audiovisual media in a wide screen format such that the width of the "virtual" screen can extend to a full 360 degrees horizontally and up to 180 degrees vertically.

16) An electronic device as described in claims 11-15, which allows the viewer to view prerecorded audiovisual material on a conventional screen based display device (TV, projection TV, computer screen) such that the display device represents a viewport or subset of the full 360 degree panoramic image.

17) An entertainment system consisting of; a range of alternative media provision devices (such as VCR, DVD, satellite receiver etc.) an electronic device (VTV processor), which generates panoramic video imagery from video data provided from the aforementioned devices and a display device such as a conventional flat screen television or helmet mounted display device (HMD) or other virtual reality display device, fitted with an optional single view or panoramic video capture device, in conjunction with a wireless data communication network to communicate this video information between the HMD and the VTV processor as shown figs. 1-3

(VTV graphics standard)

18) A new audiovisual standard (the virtual television or VTV standard) which consists of a modification to the existing television standard which allows for a variety of different "Frames", such that these Frames may

contain graphical data, sound or control information while still maintaining compatibility with the existing television standards.

(NTSC, PAL, HDTV etc.)

19) A new audiovisual standard as described in claim 18 which, includes within one or more scan lines of a standard video image, additional digital and/or analog coded data which provides information which define control parameters and image manipulation data for the VTV graphics processor.

20) A new audiovisual standard as described in claim 18 which, includes within one or more scan lines of a standard video image, additional digital and analog coded data (hybrid coded data) which provides information to generate 4 or more audio tracks in real-time.

21) A new audiovisual standard as described in claim 18 which, includes within one or more scan lines of a standard video image, additional digital or analog coded data which provides information as to absolute orientation (azimuth or azimuth and elevation) of the camera that filmed the imagery.

22) A new audiovisual standard as described in claim 18 which, includes within one or more scan lines of a standard video image, additional digital or analog coded data which provides information as to the relative placement position of the current Page (video field or frame) within the 360 degree horizontal by X degree vertical "Image Sphere".

23) A new audiovisual standard as described in claims 18,20, which, includes within one or more scan lines of a standard video image, additional digital or analog coded data which provides information as to the number of audio tracks, the audio sampling rate and the track synchronization which allows the VTV graphics processor to decode the audio information as described in claim 10 into spatial (position and orientation sensitive) sound.

24) A new audiovisual standard based around the concept of "Image Spheres" which are 360 degree horizontal by X degree vertical cylinders or truncated spheres, such that each Image Sphere consists of a number of subsections or "Pages".

25) A new audiovisual standard as described in claim 18 which makes possible the encoding of multi-track audio for use with standard video storage and transmission systems such that this information can be subsequently decoded by specific hardware (the VTV processor) to produce a left and right audio channel (for headphones or speaker systems) such that the audio channels are mixed (mathematically combined) in such a way as to produce spatially correct audio for the left and right ears of the user. The parameters affecting this mathematical combination being primarily azimuth (in the case of a 4 track audio system) and both azimuth and elevation azimuth (in the case of an 8 track audio system).

25(a) A new audiovisual standard as described in claim 18 which, includes a combination of frequency compression and expanding (Companding) in conjunction with time division multiplexing to encode 4 or more

audio tracks onto a standard two track audiovisual recording system.

25(b) A new audiovisual standard as described in claim 25(A) which utilizes one or more of the sync signals present in the video component of the audiovisual data stream (V_Sync, H_sync, Odd/Even etc.) to generate and/or synchronize a local clock for the encoding and subsequent decoding of the Time division multiplexed audiovisual data.

(VTV processor)

26) An electronic device as described in claims 11-16, which allows the viewer to view prerecorded audiovisual material using a helmet mounted display (HMD) or other virtual reality type display device such that the display device represents a viewport or subset of the full 360 degree horizontal panoramic image.

27) An electronic device as described in claims 11-16, 26, such that the horizontal direction of view within the 360 degree by X degree vertically "virtual environment" is dynamically controllable by the user at runtime (while the images being displayed).

28) An electronic device as described in claims 11-16,26-27, such that both the azimuth and elevation of the viewport within the 360 degree horizontal by X degree vertical "virtual environment" is dynamically controllable by the user at runtime (while the images being displayed).

29) An electronic device as described in claims 11-16, 26-28, in which the direction of view is automatically controlled by virtue of a tracking device which continuously measures the azimuth or both azimuth and elevation of the viewer's head.

30) An electronic device as described in claims 11-16, 26-29, in which the virtual camera position within "virtual environment" (i.e. the viewpoint of the viewer) is dynamically controllable by the user at runtime (while the images are being displayed).

31) An electronic device as described in claims 11-16, 26-30, in which the virtual camera position within "virtual environment" (i.e. the viewpoint of the viewer) is automatically controlled by virtue of a tracking device which continuously measures the physical position of viewer's head in "real world coordinates".

32) An electronic device, in which orientation sensitive audio is provided in real-time, which is controlled by the direction of the viewers head (azimuth and elevation).

33) An electronic device as described in claims 11-16,26-31, in which orientation sensitive audio is also provided in real-time, which is controlled by the direction of the viewport within the 360 degree Image Sphere ("virtual environment").

34) An electronic device as described in claims 11-16, 26-31, in which orientation and position sensitive

audio is also provided in real-time, which is controlled by the direction of the viewport within the 360 degree Image Sphere and virtual position within the "virtual environment".

35) An electronic device as described in claims 11-16, 26-34, which is capable of displaying prerecorded computer graphic or live imagery in a 360 degree Image Sphere format to produce a virtual reality experience which is capable of being provided from standard video storage and transmission devices (VCR, DVD, satellite transmission etc.)

36) An electronic device as described in claims 11-16, 26-35 which is capable of combining prerecorded computer graphic or live imagery with "real world imagery" captured utilizing a simple single view or panoramic camera system in real-time to produce an augmented reality experience.

37) An electronic device as described in claims 11-16, 26-36, which is capable of selectively combining and geometrically altering either "real world" or prerecorded imagery to create a composite augmented reality experience.

38) An electronic device as described in claims 11-16, 26-37, which is capable of analyzing "real world" images captured by a simple single view or panoramic camera system and by utilizing differential imaging techniques and/or other image processing techniques, is capable of automatically removing the background "real world" scenery and replacing this with synthetic or prerecorded imagery provided from a video device (such as VCR DVD player etc.)

39) An electronic device as described in claims 11-16, 26-35, which is capable of combining "foreground" and "background" pre-rendered video information utilizing chroma-keying techniques in which the foreground and background information may be provided by the same video source and which additionally the chroma-key color is dynamically variable within an image by providing an analog or digital sample of the chroma-key color coded either as a special control frame, as part of each scan line of the video image.

40) An electronic device which is capable of performing both of the functions described in clams 38 and 39.

41) An electronic device which is capable of analyzing images captured by a simple single view or panoramic camera system as described in claims 49-54 and interpreting the imagery as three-dimensional objects in real-time.

42) An electronic device as described in claim 41, which converts the three-dimensional objects into a "universal graphics description language" such as VRML or other appropriate language for storage or live transmission and subsequent decoding into graphical imagery by another VTV processor and appropriate display device.

43) An electronic device (otherwise known as the VTV graphics processor) described in claims 11-16, 26-42, shown in figs 8-10, and who's functionality is described in paragraphs 3.1-3.18, which is comprised of; one or

more video digitizing modules, three areas of memory, known as augmented reality memory (ARM), virtual reality memory (VRM), and translation memory (TM), a digital processing module and one or more video generation modules.

44) An electronic device as described in claim 43, In which the augmented reality memory (ARM) is "mapped" to occupy a smaller vertical field of view than the virtual reality memory (VRM), and translation memory (TM) so as to minimize the data requirement for the provision of the media whilst still maintaining a high-quality image.

45) An electronic device as described in claims 43-44, In which the augmented reality memory (ARM), virtual reality memory (VRM), and translation memory (TM) may be "mapped" at different resolutions (i.e. pixels in each memory region can represent a different degree of angular deviation.)

46) An electronic device as described in claims 43-45, which displays imagery as described in claims 36-38, by first placing the "real world" video information in augmented reality memory (foreground memory), source information from video provision device (VCR, DVD player etc.) into virtual reality memory and then combining these two sources of imagery according to the pattern of data held in translation memory (part of the Warp Engine) into a "composite image" before displaying on the output device (such as a flat screen display or HMD).

47) An electronic device as described in claims 43-45, which displays imagery as described in claims 35,39, by first placing the foreground video information from a video provision device (VCR, DVD player etc.) into augmented reality memory, and then by placing background video information from a video provision device (VCR, DVD player etc.) into virtual reality memory and then combining these two sources of imagery according to the pattern of data held in translation memory (part of the Warp Engine) into a "composite image" before displaying on the output device (such as a flat screen display or HMD).

48) An electronic device as described in claims 47, which in addition to using the Warp Engine for image combination also relies on chroma-keying information present in the video media to determine foreground and background priority for final combination and display.

48(a) An electronic device as described in claims 1 – 6 which , incorporates a combination of frequency compression and expanding (Commanding) in conjunction with time division multiplexing to encode 4 or more audio tracks onto a standard two track audiovisual recording system.

48(b) An electronic device as descried in claim 48(a) which utilizes one or more of the sync signals present in the video component of the audiovisual data stream (V_Sync, H_sync, Odd/Even etc.) to generate and/or synchronize a local clock for the encoding and subsequent decoding of the Time division multiplexed audiovisual data.

(Pan-Cam system)

49) An electro-optical assembly which consists of a plurality of electronic image capture devices (video cameras, HDTV cameras, digital still cameras etc.) which are configured with overlapping horizontal fields of view such that collectively the overlapping horizontal fields of view cover a full 360 degrees.

50) An electronic device which crops and aligns the individual images (Pages) produced by the assembly described in claims 49 to produce an overall 360 degree panoramic image with negligible distortion and overlap between the individual Pages.

51) An electronic device as described in claim 50, which in addition to cropping and aligning the separate images to produce a seamless 360 degree panoramic image, also applies distortion correction to the images so that the resulting 360 degree panoramic image is mathematically "flat" in the horizontal axis. (i.e. each pixel in the horizontal axis of the image subtends an equal angle to the camera.)

52) An electronic device as described in claims 50-51, which also applies distortion correction to the images so that the resulting 360 degree panoramic image is mathematically "flat" in the vertical axis. (i.e. each pixel in the vertical axis of the image subtends an equal angle to the camera.)

53) An electronic device as described in claims 50-52, which additionally, Inserts "Page identification information" which describe the location of the individual Pages that comprise the 360 degree panoramic image produced by the panoramic camera assembly, into the outgoing video stream.

54) An electronic device as described in claims 50-53, which additionally, Inserts "tracking information" which describe the current orientation of the panoramic camera assembly (azimuth and elevation) into the video stream.

55) An electronic device which utilizing data received from one or more video capture devices (video cameras etc.) and by performing a series of simple image analysis processes such as autocorrelation calculates relative movement in the azimuth of the camera (of the viewer in the case of an HMD based camera assembly) as shown in figs. 16,19 and more completely described in paragraphs 4.1-4.8.

56) An electronic device which utilizing data received from one or more video capture devices (video cameras etc.) and by performing a series of simple image analysis processes such as autocorrelation calculates relative movement in the elevation of the camera (of the viewer in the case of an HMD based camera assembly) as shown in figs. 26,28 and more completely described in paragraphs 4.1-4.8.

57) An electronic device which utilizing data received from one or more video capture devices (video cameras etc.) and by performing a series of simple image analysis processes such as autocorrelation calculates relative movement in the roll of the camera (of the viewer in the case of an HMD based camera assembly) as shown in figs. 26,29 and more completely described in paragraphs 4.1-4.8.

25

58) An electronic device which utilizing data received from one or more video capture devices (video cameras etc.) and by performing a series of simple image analysis processes such as autocorrelation calculates relative movement in the physical (spatial) position of the camera (of the viewer in the case of an HMD based camera assembly) in either or any combination of the X, Y or Z axes as shown in figs. 16,20-21 and more completely described in paragraphs 4.1-4.8.

59) An electronic device as described in claims 55-58, which utilizes a number of retroflective targets with known "real world" coordinates in conjunction with constant or strobed on-axis light sources to determine absolute angular/spatial references for the purposes of a converting the relative angular and spatial data determined by devices described in claims 55-58 into absolute angular and spatial data.

60) An electronic device as described in claim 59, which utilizes a combination of color filters over the retroflective targets in conjunction with controllable on-axis light sources which are synchronized to the video capture rate of the HMD based or remote panoramic cameras to improve the ability of the system to correctly identify and maintain tracking of the individual retroflective targets.

61) An electronic device as described in claims 59-60, which utilizes a combination of retroflective targets in conjunction with color controllable on-axis light sources which are synchronized to the video capture rate of the HMD based or remote panoramic cameras to improve the ability of the system to correctly identify and maintain tracking of the individual retroflective targets.

62) An electronic device as described in claims 59-61, which utilizes a combination of color filters over the retroflective targets in conjunction with color controllable on-axis light sources which are synchronized to the video capture rate of the HMD based or remote panoramic cameras to improve the ability of the system to correctly identify and maintain tracking of the individual retroflective targets.

63) An electronic device as described in claims 55-58, which utilizes a number of "active optical beacons" (controllable light sources which are synchronized to the video capture rate of the HMD based or remote panoramic cameras) such that pulse timing, color of light and/or combinations of these are used to transmit the "real world" coordinates of the beacon to the HMD or remote panoramic camera to determine absolute angular/spatial references for the purposes of a converting the relative angular and spatial data determined by devices described in claims 55-58, into absolute angular and spatial data.

64) An electronic device as described in claims 545-58, which utilizes a number of "bi-directional infrared beacons" which communicate a unique ID code between the HMD and the beacon such that this calibration would occur only once each time the HMD passed under any of these "known in spatial reference points."

65) An electronic device which utilizes a single optical imaging device to monitor a pattern on the ceiling and utilizing similar image processing techniques as described in claims 55-58, determines relative spatial movement and azimuth, in conjunction with an alternative angular tracking system such as fluid level sensors to determine the remaining angular orientations (pitch and roll).

26

66) An electronic device as described in claim 65 which utilizes any of the calibration systems as described in claims 59-64 to determine absolute references for the purposes of converting the relative spatial data determined by the device described in claim 65, into absolute spatial data.

67) An assembly which consists of a plurality of electronic audio capture devices microphones) which are configured with overlapping horizontal fields of coverage such that collectively the overlapping horizontal fields of coverage cover a full 360 degrees horizontally.

68. An assembly which consists of a plurality of electronic audio capture devices (microphones) such as described in claim 67 which are additionally configured with overlapping vertical fields of coverage such that collectively the overlapping vertical fields of coverage cover up to a maximum of +/- 90 degrees vertically.

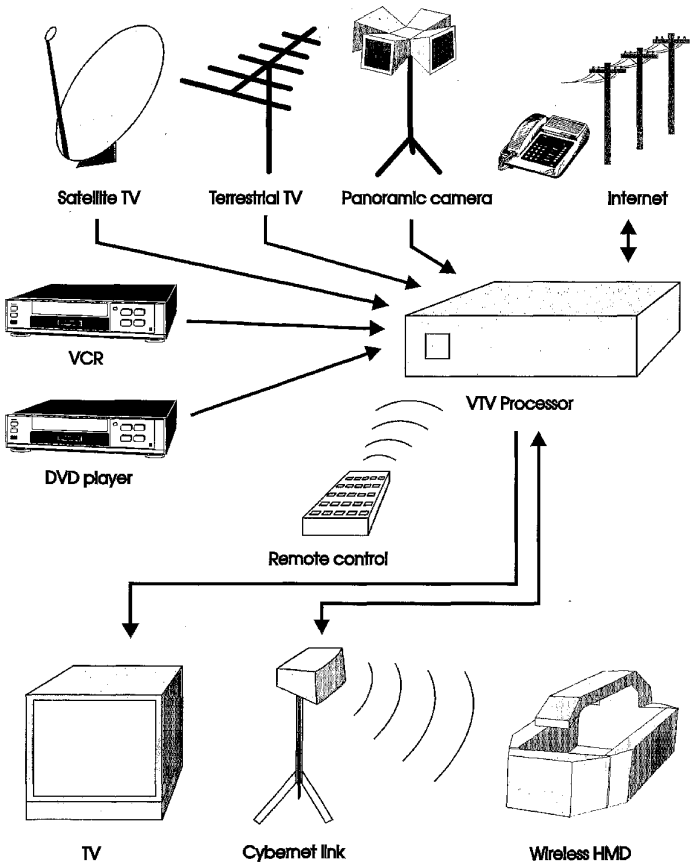

Satellite TV
Terrestrial TV
Panoramic camera
Internet
VCR
VTV Processor
DVD player
Remote control
TV
Cybernet link
Wireless HMD

Fig. 1

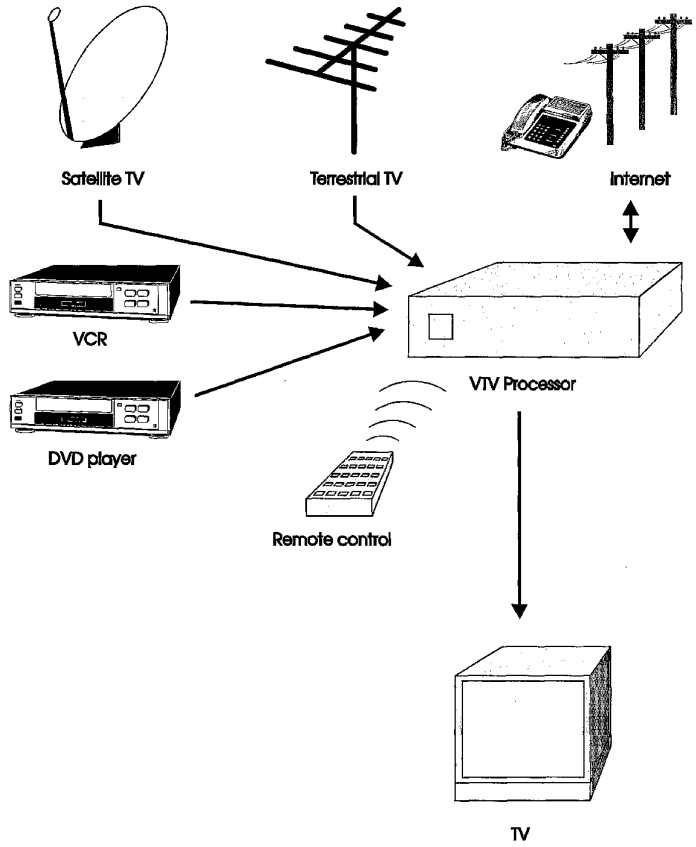

Satellite TV
Terrestrial TV
Internet
VCR
DVD player
VTV Processor
Remote control
TV

Fig. 2

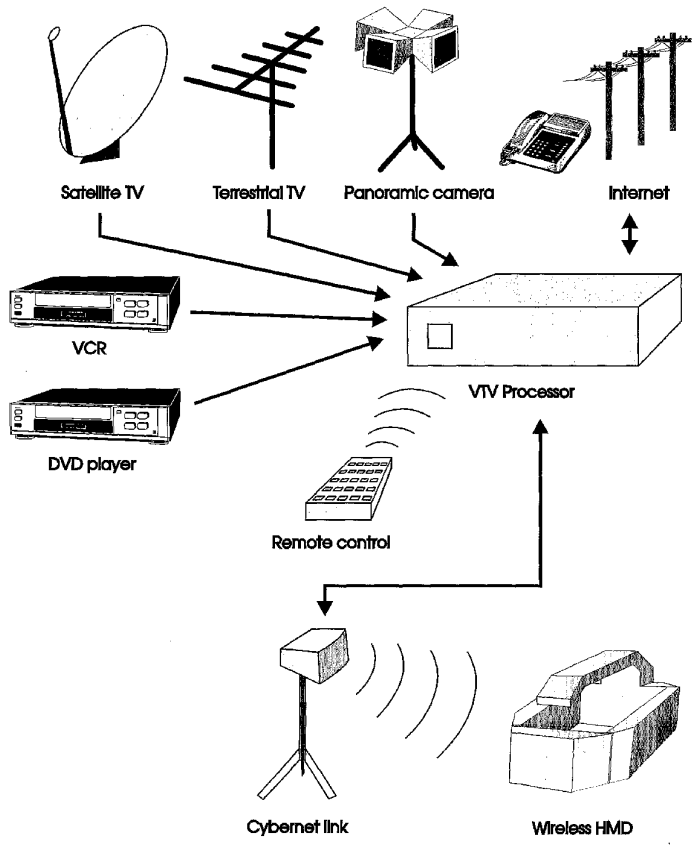

Satellite TV
Terrestrial TV
Panoramic camera
Internet
VCR
VTV Processor
DVD player
Remote control
Cybernet link
Wireless HMD

Fig. 3

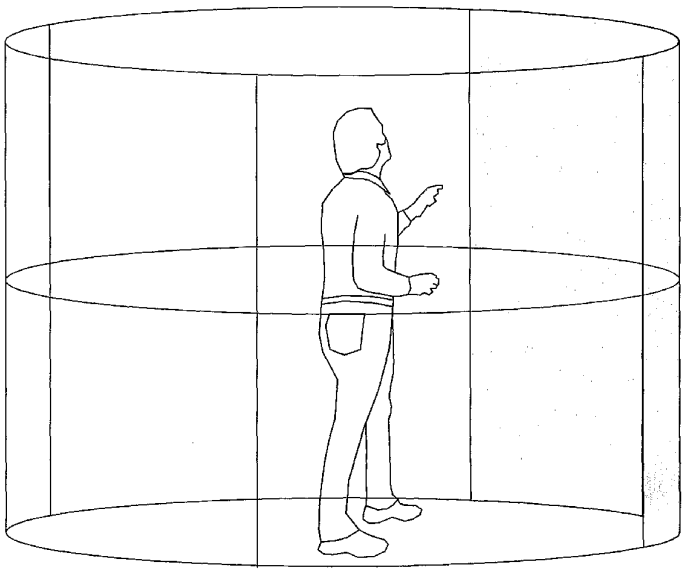


Fig. 4

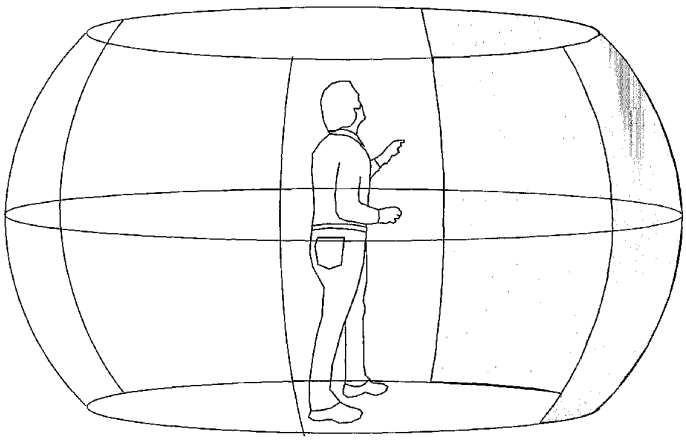


Fig. 5

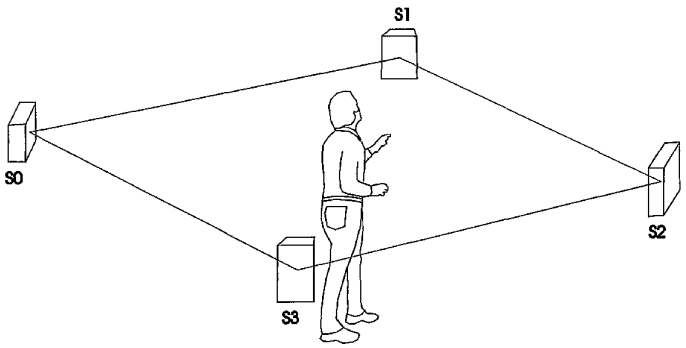


Fig. 6

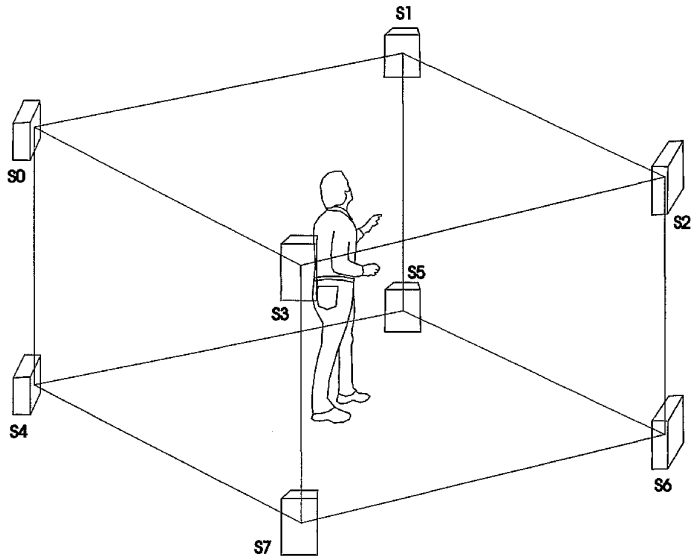


Fig. 7

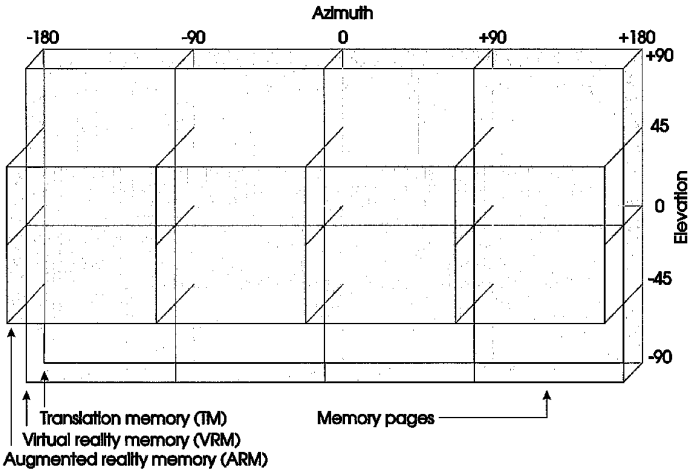

Azimuth
-180
-90
0
+90
+180
+90
45
0
-45
-90
Elevation
Translation memory (TM)
Virtual reality memory (VRM)
Augmented reality memory (ARM)
Memory pages

Fig. 8

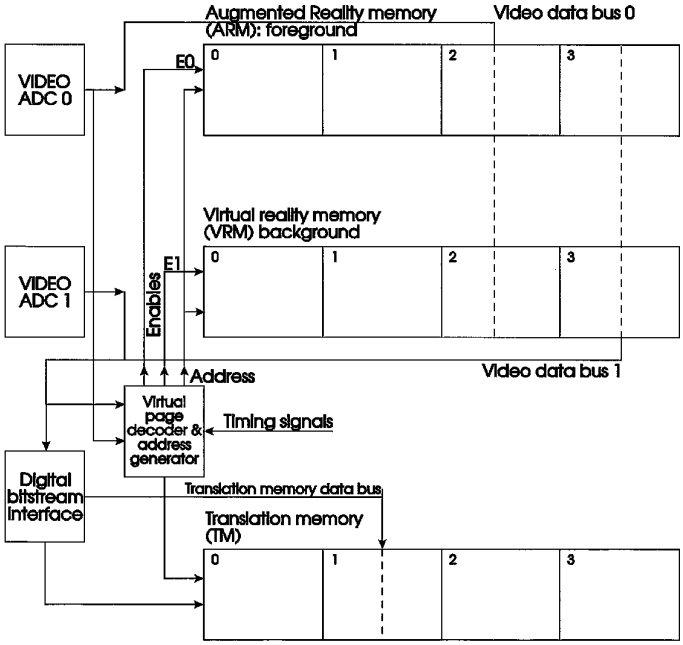

Augmented Reality memory
(ARM): foreground
Video data bus 0
VIDEO
ADC 0
E0
0
1
2
3
Virtual reality memory
(VRM) background
VIDEO
ADC 1
E1
Enables
Address
Video data bus 1
Virtual
page
decoder &
address
generator
Timing signals
Digital
bitstream
interface
Translation memory data bus
Translation memory
(TM)

Fig. 9

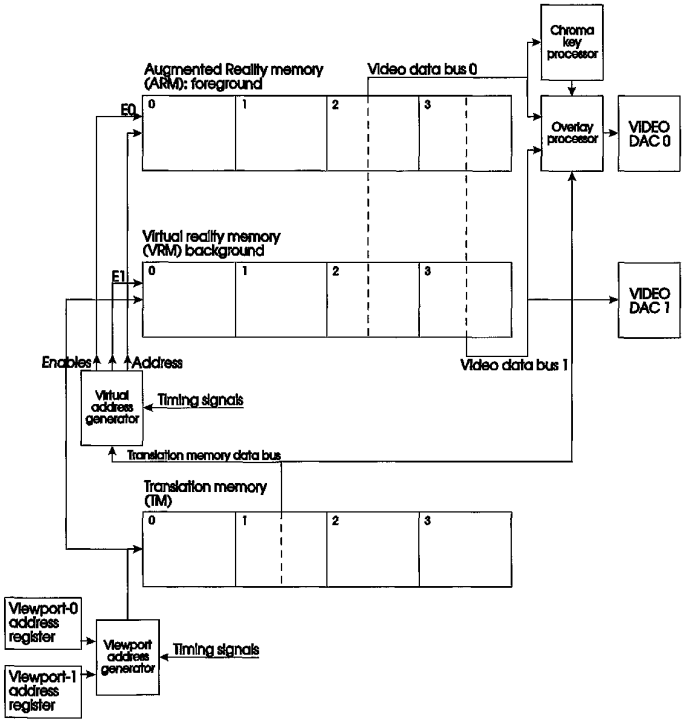

Chroma key processor
Augmented Reality memory (ARM): foreground
Video data bus 0
E0
0
1
2
3
Overlay processor
VIDEO DAC 0
Virtual reality memory (VRM) background
E1
0
1
2
3
VIDEO DAC 1
Enables
Address
Video data bus 1
Virtual address generator
Timing signals
Translation memory data bus
Translation memory (TM)
0
1
2
3
Viewport-0 address register
Viewport-1 address register
Viewport address generator
Timing signals

Fig. 10

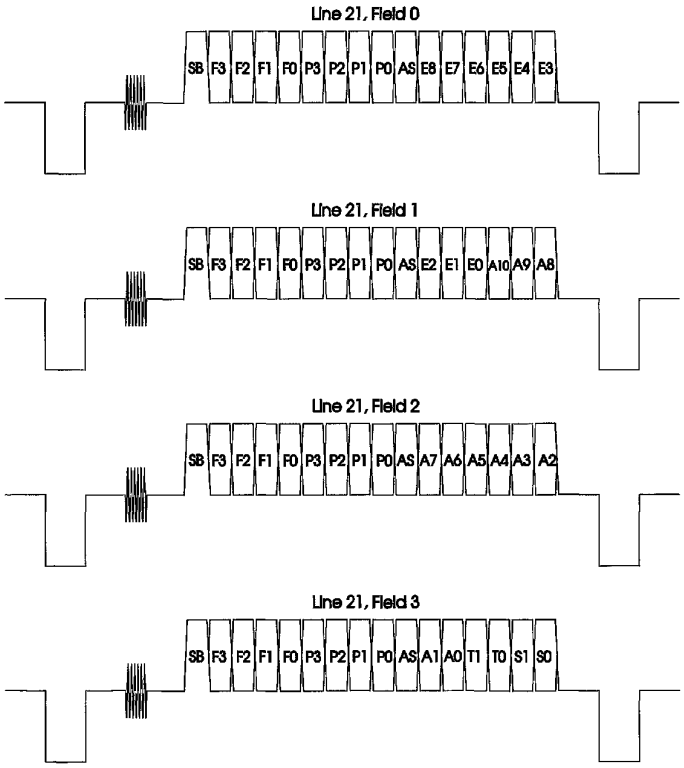

Line 21, Field 0
SB F3 F2 F1 F0 P3 P2 P1 P0 AS E8 E7 E6 E5 E4 E3
Line 21, Field 1
SB F3 F2 F1 F0 P3 P2 P1 P0 AS E2 E1 E0 A10 A9 A8
Line 21, Field 2
SB F3 F2 F1 F0 P3 P2 P1 P0 AS A7 A6 A5 A4 A3 A2
Line 21, Field 3
SB F3 F2 F1 F0 P3 P2 P1 P0 AS A1 A0 T1 T0 S1 S0

Fig. 11

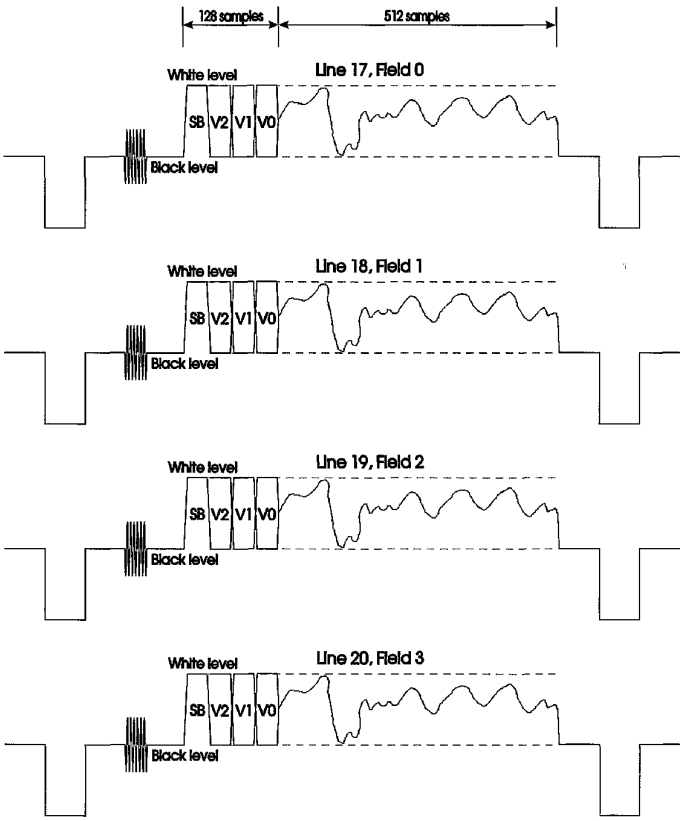

128 samples
512 samples
White level
Line 17, Field 0
SB
V2
V1
V0
Black level
White level
Line 18, Field 1
SB
V2
V1
V0
Black level
White level
Line 19, Field 2
SB
V2
V1
V0
Black level
White level
Line 20, Field 3
SB
V2
V1
V0
Black level

Fig. 12

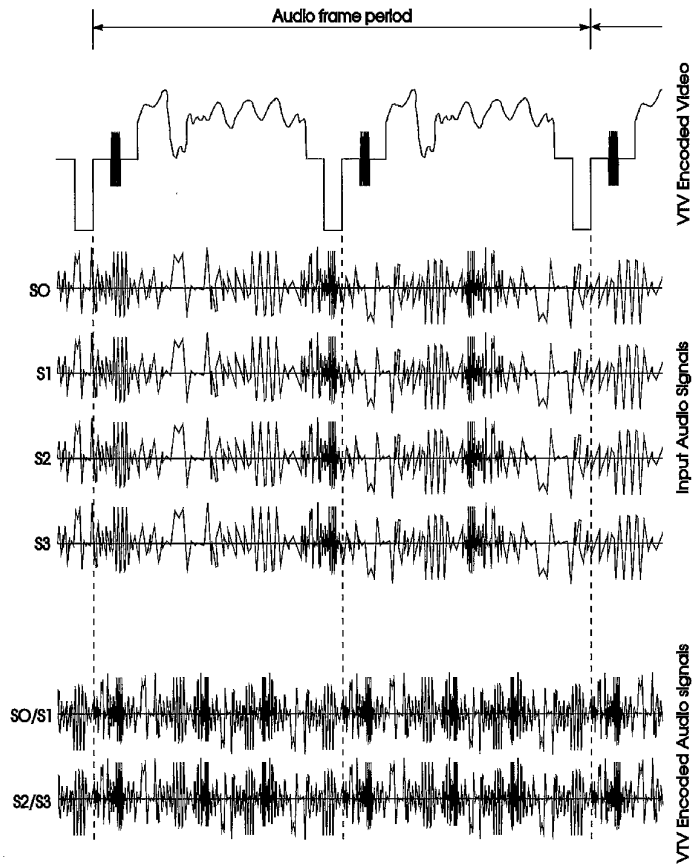

Audio frame period
VTV Encoded Video
S0
S1
S2
S3
Input Audio Signals
S0/S1
S2/S3
VTV Encoded Audio signals

Fig. 13

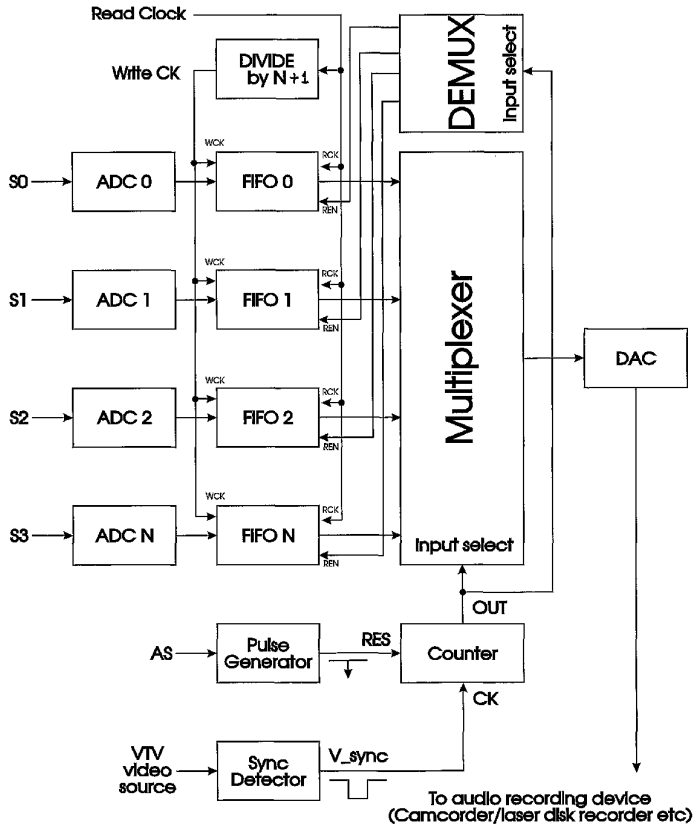

Read Clock
Write CK
DIVIDE by N+1
DEMUX
Input select
S0
ADC 0
FIFO 0
S1
ADC 1
FIFO 1
S2
ADC 2
FIFO 2
S3
ADC N
FIFO N
WCK
RCK
REN
Multiplexer
Input select
DAC
OUT
AS
Pulse Generator
RES
Counter
CK
VTV video source
Sync Detector
V_sync
To audio recording device
(Camcorder/laser disk recorder etc)

Fig. 14

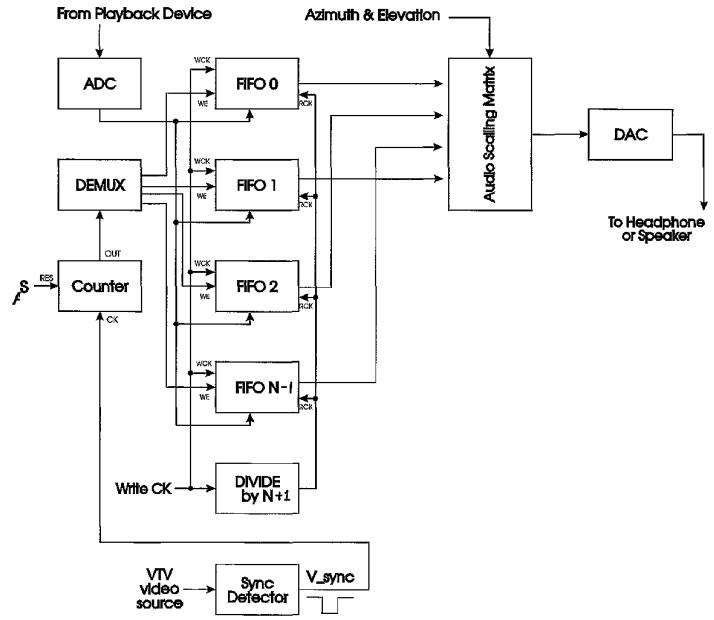

From Playback Device
Azimuth & Elevation
ADC
FIFO 0
DEMUX
FIFO 1
Audio Scalling Matrix
DAC
To Headphone or Speaker
Counter
FIFO 2
FIFO N-1
Write CK
DIVIDE by N+1
VTV video source
Sync Defector
V_sync

Fig. 15

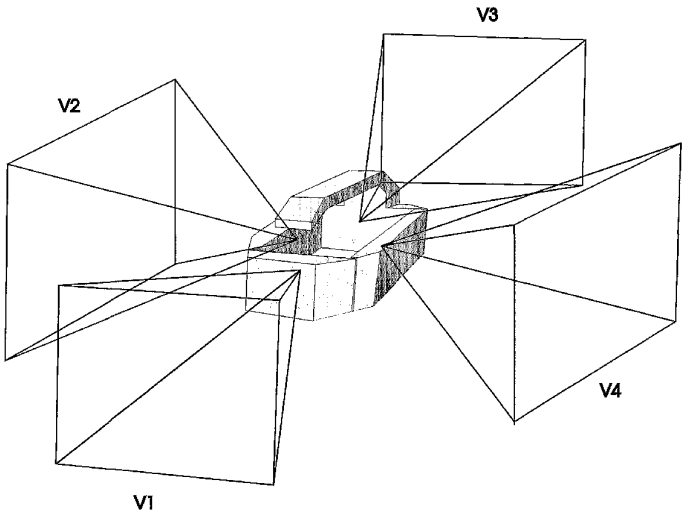


Fig. 16

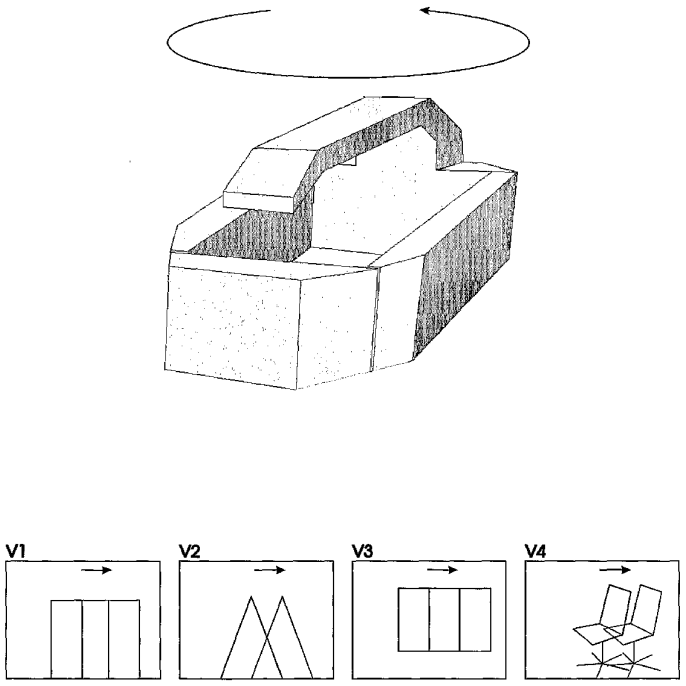


Fig. 17

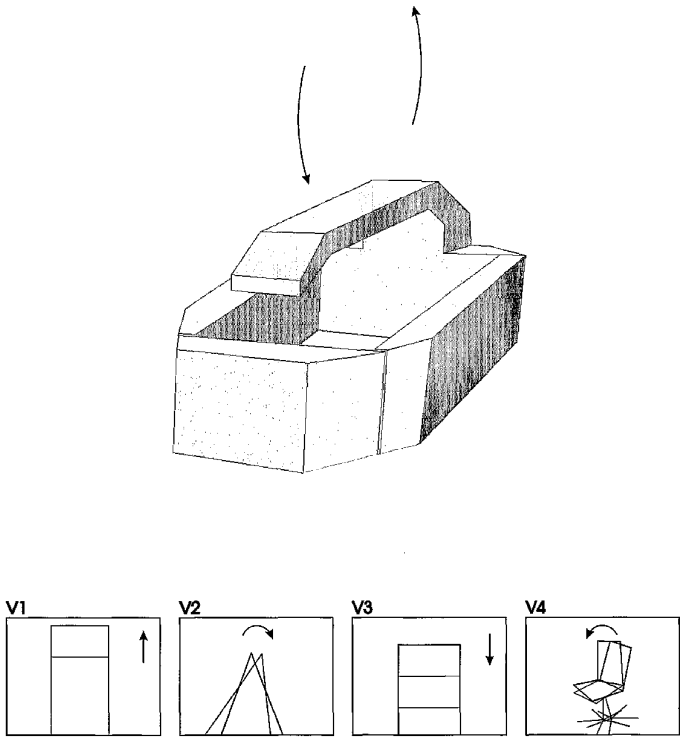
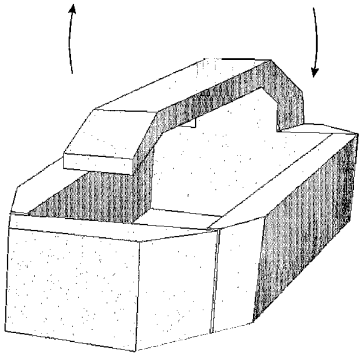


Fig. 18



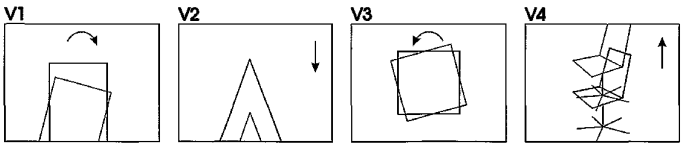


Fig. 19

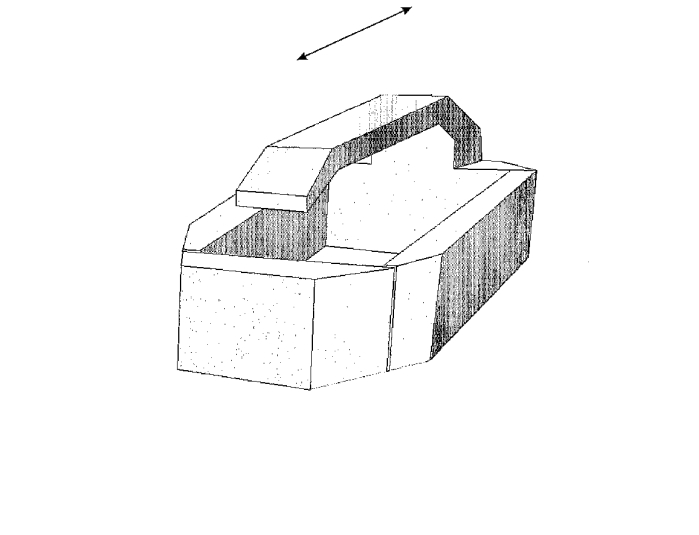


Fig. 20

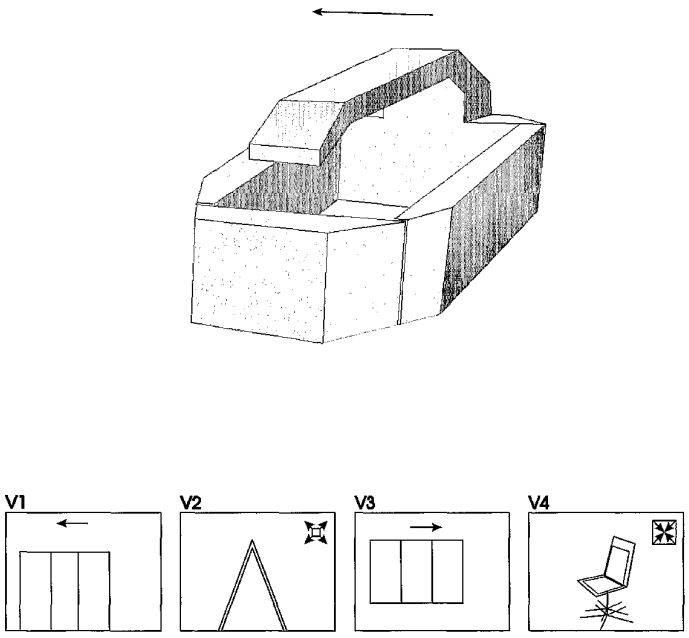


Fig. 21

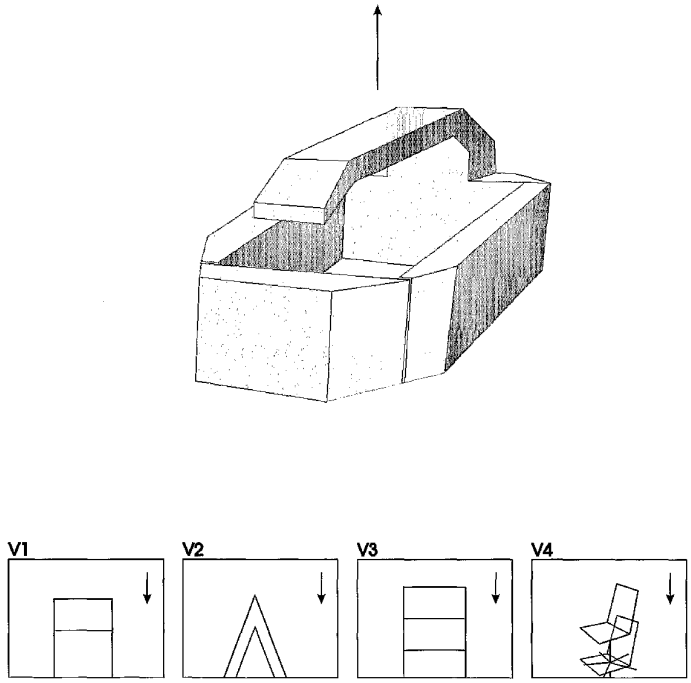


Fig. 22

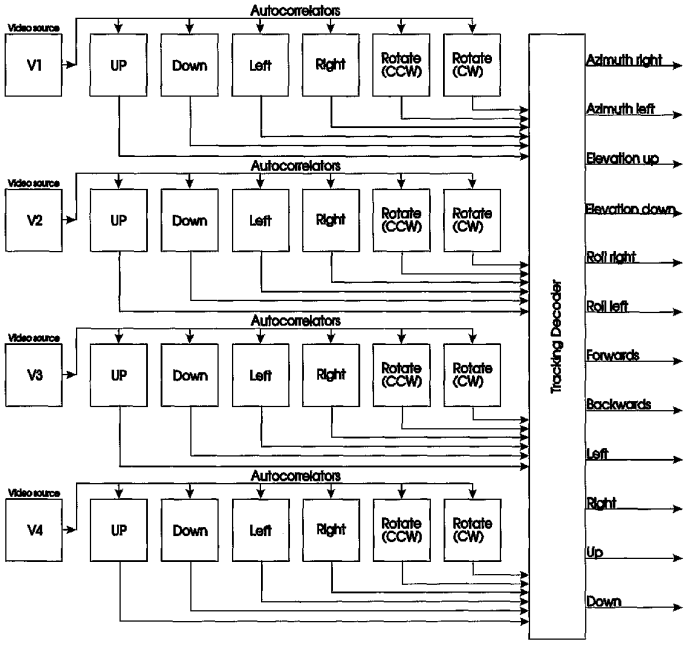

Autocorrelators
Video source
V1
UP
Down
Left
Right
Rotate (CCW)
Rotate (CW)
Autocorrelators
Video source
V2
UP
Down
Left
Right
Rotate (CCW)
Rotate (CW)
Autocorrelators
Video source
V3
UP
Down
Left
Right
Rotate (CCW)
Rotate (CW)
Autocorrelators
Video source
V4
UP
Down
Left
Right
Rotate (CCW)
Rotate (CW)
Tracking Decoder
Azimuth right
Azimuth left
Elevation up
Elevation down
Roll right
Roll left
Forwards
Backwards
Left
Right
Up
Down

Fig. 23

【手続補正書】
【提出日】平成15年2月8日(2003.2.8)
【手続補正１】
【補正対象書類名】特許請求の範囲
【補正対象項目名】全文
【補正方法】変更
【補正の内容】
【特許請求の範囲】
【請求項１】
ソースメディアを供給するための従来方式のビデオ装置を利用して拡張型立体テレビジョン様の視聴体験を生み出す方法において、
（ａ）ソースメディアを供給するための従来方式のビデオ装置を利用して、リアルタイムでページベースで画像球面全体を更新するステップと、
（ｂ）複数のさらに小さな画像のサブセクションから構成されるパノラマ画像を描くグラフィック画像を生成するステップと、を有することを特徴とする方法。
【請求項２】
請求項１記載の方法において、グラフィック画像を生成するステップが、３６０°の水平方向の視野を持つパノラマ画像を供給するステップを有することを特徴とする方法。
【請求項３】
請求項１記載の方法において、ソースメディアの中に存在する付加情報によりページ順を確定するステップをさらに有することを特徴とする方法。
【請求項４】
請求項１記載の方法において上記バーチャルスクリーンの幅を水平方向に最大３６０°及び垂直方向に１８０°まで拡張できるようなワイドスクリーンフォーマットで、予め録画された視聴覚メディアを視聴するステップをさらに有することを特徴とする方法。
【請求項５】
請求項１記載の方法において、従来方式のスクリーンをベースとする表示装置で予め録画された視聴覚マテリアルを視聴するステップをさらに有し、上記表示装置が上３６０°のフルパノラマ画像のビューポートを表示するように為すことを特徴とする方法。
【請求項６】
娯楽システムであって、
（ａ）異なるタイプの複数のメディア供給装置と、
（ｂ）上記メディア供給装置から得られるビデオデータに応答して、そこからパノラマビデオ画像を生成するＶＴＶプロセッサと、
（ｃ）上記ＶＴＶプロセッサから得られるパノラマビデオ画像に応答する表示装置と、を有することを特徴とする娯楽システム。
【請求項７】
視聴覚効果を生成する方法において、オリジナルのテレビプロトコルを改変して、上記オリジナルのプロトコルとの互換性を壊すことなく、グラフィックデータと、サウンド情報と、制御情報とのうちの少なくとも１つのデータの伝送を個々のフレームで調節するように為すステップを有することを特徴とする方法。
【請求項８】
請求項７記載の方法において、上記改変されたプロトコルが、追加のデジタルコード化データとアナログコード化データとのうちの少なくとも一方のデータを処理して、上記オリジナルのプロトコルに従って伝送されるビデオ画像の少なくとも１つの走査線内で、ＶＴＶグラフィックプロセッサ用の制御用パラメータと画像処理データを定義する情報とを提供するように為すことを特徴とする方法。
【請求項９】
請求項８記載の方法において、上記走査線が、４またはそれ以上のオーディオトラックをリアルタイムで生成するための情報を提供する追加のデジタルデータとアナログコード化

データとをさらに含むことを特徴とする方法。
【請求項１０】
時分割多重化と共に信号を周波数圧縮し、次いで、上記信号を周波数伸長するステップを有し、２トラック視聴覚録画システムから出力される４またはそれ以上のオーディオトラックを符号化し、その後該トラックを復号化して、上記視聴覚データストリームのビデオ成分の中に存在する同期信号を用いて、上記時分割多重化済み視聴覚データの符号化とそれに後続する復号化を行うために遠隔クロックの生成と同期のうちの少なくとも一方を実行するように為すことを特徴とする方法。
【請求項１１】
電子装置であって、
（ａ）複数のビデオ画像表示装置と、
（ｂ）前記装置を利用して拡張型立体テレビジョン様の視聴体験を生み出す手段と、
（ｃ）視聴可能領域が水平方向に最大３６０°及び垂直方向に１８０°まで拡張するようなワイドスクリーンのパノラマフォーマットで視聴者が視聴覚メディアを視聴する手段と、を有することを特徴とする電子装置。
【請求項１２】
電子装置であって、
（ａ）パノラマ画像全体が複数のさらに小さな画像ページから構成されるようなパノラマ画像を描くグラフィック画像を生成する手段と、
（ｂ）ページベースで上記パノラマ画像をリアルタイムで更新する手段と、を有することを特徴とする電子装置。
【請求項１３】
請求項１２記載の電子装置において、ソースメディアの中に存在する付加情報により画像ページの順序を確定する手段をさらに有することを特徴とする電子装置。
【請求項１４】
請求項１２記載の電子装置において、スクリーンベースの表示装置をさらに有し、視聴者が上記フルパノラマ画像の一部を該表示装置で観察できるようにすることを特徴とする電子装置。
【請求項１５】
請求項１２記載の電子装置において、バーチャルリアリティ型表示装置をさらに有し、視聴者が上記フルパノラマ画像の一部を該表示装置を通じて観察できるようにすることを特徴とする電子装置。
【請求項１６】
視聴覚伝播信号であって、画像データ、サウンドデータ、及び制御情報のうちの少なくとも１つを個々のフレームが含む複数のフレームを有する視聴覚信号において、少なくとも１つの広く受け入れられているテレビジョン規格と互換性を有することを特徴とする視聴覚伝播信号。
【請求項１７】
請求項１６記載の視聴覚伝播信号において、標準的ビデオ画像の１以上の走査線内にＶＴＶグラフィックプロセッサ用の制御用パラメータを定義する追加のコード化データと、画像処理用データとをさらに有することを特徴とする視聴覚伝播信号。
【請求項１８】
請求項１６記載の視聴覚伝播信号において、パノラマ画像球面のデータをさらに有することを特徴とする視聴覚伝播信号。
【請求項１９】
請求項１６記載の視聴覚伝播信号において、標準的ビデオ画像の１以上の走査線内に上記画像球面内の現在のビデオフィールドまたはフレームの相対的配置位置に関する情報を提供する追加のコード化データをさらに有することを特徴とする視聴覚伝播信号。
【請求項２０】
請求項１６記載の視聴覚伝播信号において、標準的ビデオ画像の１以上の走査線内に４ま

たはそれ以上のリアルタイムのオーディオトラックを生成するための情報を提供する追加のコード化データをさらに有することを特徴とする視聴覚伝播信号。
【請求項２１】
請求項１６記載の視聴覚伝播信号において、周波数圧縮と時分割多重方式により２トラックプロトコルを介して符号化された４またはそれ以上のオーディオトラックデータをさらに有することを特徴とする視聴覚伝播信号。
【請求項２２】
請求項１６記載の視聴覚伝播信号であって、標準的ビデオ画像の１以上の走査線内に追加のコード化データをさらに有し、該コード化データが、
（ａ）４またはそれ以上のリアルタイムのオーディオトラックを生成するためのオーディオ情報と、
（ｂ）使用オーディオトラック数と、使用オーディオデータフォーマットと、使用オーディオサンプリングレートと、トラック同期とを記述するデータとを提供し、これによりＶＴＶグラフィックプロセッサが、上記オーディオ情報を復号化して位置と方角に敏感なサウンドに変換できることを特徴とする視聴覚伝播信号。
【請求項２３】
請求項１６記載の視聴覚伝播信号において、標準的ビデオ画像の１以上の走査線内にカメラの絶対方角とＸ－Ｙ－Ｚ位置とに関する情報を提供する追加のコード化データをさらに有することを特徴とする視聴覚伝播信号。
【請求項２４】
請求項１１乃至１５のいずれかに記載の電子装置において、
（ａ）視聴者の方位角と高度に関する情報を数学的に合成する手段と、
（ｂ）標準的ビデオ格納システムと伝送システムと共に使用するためにマルチトラックオーディオデータを符号化し、上記合成された情報を専用ハードウェアによりその後復号化して、正しい３次元サウンドを視聴者の左右の耳に空間的に生みだす左右のオーディオチャネルを形成できるようにする手段と、をさらに有することを特徴とする電子装置。
【請求項２５】
請求項１１乃至１５のいずれかに記載の電子装置において、視聴覚データストリームのビデオ成分の中に存在する１またはそれ以上の同期信号と同期して作動するクロックをさらに有し、上記同期信号と同期して時分割多重サウンドを生成するように為すことを特徴とする電子装置。
【請求項２６】
請求項１１乃至１５のいずれかに記載の電子装置において、実行時のユーザ制御に応答して上記パノラマ内のビューポートの方位角と高度とを変動させる手段をさらに有することを特徴とする電子装置。
【請求項２７】
請求項１１乃至１５のいずれかに記載の電子装置において、視聴者の頭部の方位角と高度の連続測定に応答して上記パノラマ内のビューポートの方位角と高度とを変動させる手段をさらに有することを特徴とする電子装置。
【請求項２８】
請求項１１乃至１５のいずれかに記載の電子装置において、実行時のユーザ制御に応答して上記パノラマ内で上記ビューポートの角度視野を変動させる手段をさらに有することを特徴とする電子装置。
【請求項２９】
請求項１１乃至１５のいずれかに記載の電子装置において、実行時のユーザ制御に応答して３次元バーチャル空間内で視点を変える手段をさらに有することを特徴とする電子装置。
【請求項３０】
請求項１１乃至１５のいずれかに記載の電子装置において、
（ａ）視聴者の物理位置を連続して測定するためのトラッキング装置と、

（ｂ）上記測定に応答して３次元バーチャル空間内で視点を変える手段とをさらに有することを特徴とする電子装置。
【請求項３１】
請求項１１乃至１５のいずれかに記載の電子装置において、方角に敏感なサウンドをリアルタイムで出力する手段であって、視聴者の頭部の方向により制御される手段をさらに有することを特徴とする電子装置。
【請求項３２】
請求項１１乃至１５のいずれかに記載の電子装置において、方角に敏感なサウンドをリアルタイムで出力する手段であって、上記パノラマ内のビューポートの座標により制御される手段をさらに有することを特徴とする電子装置。
【請求項３３】
請求項１１乃至１５のいずれかに記載の電子装置において、方角及び位置に敏感なサウンドをリアルタイムで出力する手段であって、上記パノラマ内のビューポートの座標と、３次元バーチャル空間内の視点のバーチャル位置とにより制御される手段をさらに有することを特徴とする電子装置。
【請求項３４】
請求項１１乃至１５のいずれかに記載の電子装置において、予め録画されたコンピュータ画像または生の画像を、カメラシステムによりリアルタイムで取得したリアルワールド画像と合成する手段をさらに有し、これにより、拡張リアリティ体験が生み出されることを特徴とする電子装置。
【請求項３５】
請求項１１乃至１５のいずれかに記載の電子装置において、カメラシステムにより、及び、差分画像処理技術を介して取得されたリアルワールド画像を解析して、背景のリアルワールド画像を、ビデオ装置から出力される合成画像や予め録画された画像と自動的に置き換える手段をさらに有することを特徴とする電子装置。
【請求項３６】
請求項１１乃至１５のいずれかに記載の電子装置において、“リアルワールド”画像あるいは予め録画された画像のいずれかの画像を選択的に合成し、幾何学的に変更して、コンポジットな拡張リアリティ体験をつくりだすように為す手段をさらに有することを特徴とする電子装置。
【請求項３７】
請求項１１乃至１５のいずれかに記載の電子装置であって、クロマキー技法を利用して、前景と背景の予めレンダーされたビデオ情報を合成する手段をさらに有することを特徴とする電子装置において、
（ａ）上記前景情報と背景情報とが共通のビデオソースにより与えられ、
（ｂ）上記ビデオ画像の特別の制御フレーム及び各走査線の部分のうちの一方としてコード化されたクロマキーカラーのアナログサンプルまたはデジタルサンプルを供給することにより、１画像内で上記クロマキーカラーが動的に変動可能であることを特徴とする電子装置。
【請求項３８】
ＶＴＶグラフィックプロセッサであって、
（ａ）１以上のビデオデジタル化用モジュールと、
（ｂ）拡張リアリティメモリ（ＡＲＭ）と、バーチャルリアリティメモリ（ＶＲＭ）と、変換メモリ（ＴＭ）とから成るグループから得られる１以上の記憶域と、
（ｃ）デジタル処理手段であって、
（１）ＡＲＭとＶＲＭのうちの少なくとも一方に保持されるデータのアドレスマッピングを変更して、１つの記憶位置から別の記憶位置へ画像情報を効果的に移動させ、
（２）上記ソース記憶位置と宛先記憶位置との双方から得られるデータを数学的に合成し、このデータを変更し、それによって合成と変換の機能を行うデジタル処理手段と、
（ｄ）１以上のビデオ生成モジュールと、を有することを特徴とするＶＴＶグラフィック

プロセッサ。
【請求項３９】
請求項３８記載のＶＴＶグラフィックプロセッサにおいて、ＶＲＭとＴＭよりも狭い垂直方向の視野を占めるようにＡＲＭをマップし、それによって、高品質の画像をそのまま維持しながら、上記メディアの供給に必要なデータを最少化することを特徴とするＶＴＶグラフィックプロセッサ。
【請求項４０】
請求項３８記載のＶＴＶグラフィックプロセッサにおいて、ＡＲＭ、ＶＲＭ、ＴＭを異なる解像度でマップする手段をさらに有し、それによって、各メモリ領域内のピクセルが異なる度数の角度偏差を表すことが可能であることを特徴とするＶＴＶグラフィックプロセッサ。
【請求項４１】
請求項３８記載のＶＴＶグラフィックプロセッサにおいて、
（ａ）画像を表示する手段と、
（ｂ）ＡＲＭ内のリアルワールドビデオ情報と、ビデオ供給装置から得られるソース情報とをＶＲＭの中へ入れる手段と、
（ｃ）ＴＭの中に保持されているデータのパターンに従って、上記２つのソース画像を合成し、表示する前にコンポジット画像に変換する手段と、をさらに有することを特徴とするＶＴＶグラフィックプロセッサ。
【請求項４２】
請求項３８記載のＶＴＶグラフィックプロセッサにおいて、
（ａ）画像を表示する手段と、
（ｂ）上記ビデオ供給装置から得られるソース情報をＡＲＭとＶＲＭとの中へ入れる手段と、
（ｃ）上記ソースメディアの中に含まれる変換マップに従って上記２つの画像ソースを合成する手段と、をさらに有することを特徴とするＶＴＶグラフィックプロセッサ。
【請求項４３】
請求項３８記載のＶＴＶグラフィックプロセッサにおいて、
（ａ）画像を表示する手段と、
（ｂ）ビデオ供給装置から得られるソース情報をＡＲＭとＶＲＭの中へ入れる手段と、
（ｃ）上記画像の幾何学的解釈に従って追加のソース情報に基づき上記２つの画像ソースを合成する手段と、をさらに有することを特徴とするＶＴＶグラフィックプロセッサ。
【請求項４４】
システムであって、
（ａ）オーバーラップする水平方向の視野によって構成される複数の電子画像取得装置であって、上記オーバーラップする水平方向の視野が最大３６０°を一体にカバーするようになっている電子画像取得装置と、
（ｂ）上記装置により作成された個々の画像をクロップし、アラインメントを行って上記個々のページ間の無視できる歪みとオーバーラップとを持つ３６０°のパノラマ画像全体を形成するように為す手段と、を有することを特徴とするシステム。
【請求項４５】
請求項４４記載のシステムにおいて、上記画像に対して歪み補正を適用し、３６０°のパノラマ画像が横軸で数学的に平らになり、上記画像の横軸の各ピクセルが取得装置に対して等しい角度で対する手段をさらに有することを特徴とするシステム。
【請求項４６】
請求項４４記載のシステムにおいて、上記画像に対して歪み補正を適用して、３６０°のパノラマ画像が縦軸で数学的に平らになり、それによって上記画像の縦軸の各ピクセルが取得装置に対して等しい角度で対するように為す手段をさらに有することを特徴とするシステム。
【請求項４７】

請求項４４記載のシステムにおいて、３６０°のパノラマ画像を含む個々の画像の記憶位置を記述するための識別情報を送出用ビデオストリームの中へ挿入する手段をさらに有することを特徴とするシステム。
【請求項４８】
請求項４４記載のシステムにおいて、上記取得装置の現在の方角と位置とを記述するためのトラッキング情報を送出用ビデオストリームの中へ挿入する手段をさらに有することを特徴とするシステム。
【請求項４９】
システムであって、
（ａ）ビデオ取得装置と、
（ｂ）上記ビデオ取得装置から受信したデータを利用して、上記装置の方角の変化を計算するための一連の画像解析処理を実行する手段と、を有することを特徴とするシステム。
【請求項５０】
システムであって、
（ａ）ビデオ取得装置と、
（ｂ）上記ビデオ取得装置から受信したデータを利用して、上記装置の位置の変化を計算するための一連の画像解析処理を実行する手段と、を有することを特徴とするシステム。
【請求項５１】
請求項４９または５０記載のシステムにおいて、
（ａ）所定の座標に配置された複数の逆反射作用を生じるターゲットと、
（ｂ）上記ビデオ取得装置の取得レートと同期してストロボ発光される複数の軸上光源と、
（ｃ）上記所定の座標に基づいて絶対角データと空間データ、及び、上記ビデオ取得装置により決定された相対角データと相対空間データとを計算する手段と、をさらに有することを特徴とするシステム。
【請求項５２】
請求項５１記載のシステムにおいて、上記逆反射作用を生じるターゲット上に配置された複数のカラーフィルタをさらに有し、上記個々の逆反射作用を生じるターゲットのトラッキングを正しく識別し、維持するシステムの能力の向上を図るように為すことを特徴とするシステム。
【請求項５３】
請求項５１記載のシステムにおいて、光源がカラー制御可能であり、これにより、該トラッキングを維持するシステムの能力の向上を図るように為すことを特徴とするシステム。
【請求項５４】
請求項５２記載のシステムにおいて、光源がカラー制御可能であり、これにより、個々の逆反射作用を生じるターゲットのトラッキングを正しく識別し維持するシステムの能力のさらなる向上を図るように為すことを特徴とするシステム。
【請求項５５】
請求項４９または５０記載のシステムにおいて、
（ａ）上記ビデオ取得装置の取得レートと同期する複数の制御可能な光源と、
（ｂ）パルス出力タイミングと光の色とのうちの少なくとも一方を利用して上記ビデオ取得装置へ各ビーコンの空間座標を伝送する手段と、をさらに有し、
これにより、上記装置により相対角データと相対空間データとを確定し、これらのデータを絶対角データと絶対空間データとに変換できることを特徴とするシステム。
【請求項５６】
請求項４９または５０記載のシステムにおいて、上記ビデオ取得装置を用いて一意のＩＤコードを通信するための複数の双方向赤外線ビーコンをさらに有することを特徴とするシステム。
【請求項５７】
請求項４９または５０記載のシステムにおいて、

（ａ）天井の視覚パターンをモニタするための光学的画像処理手段と、
（ｂ）複数の流体チルトセンサと、
（ｃ）上記光学的画像処理手段の出力を介して相対空間運動と方位角とを確定し、上記流体チルトセンサの出力を介してピッチと回転を確定する手段とをさらに有することを特徴とするシステム。
【手続補正２】
【補正対象書類名】明細書
【補正対象項目名】全文
【補正方法】変更
【補正の内容】
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は画像視聴システムとサウンドシステムに関する。
【背景技術】
【０００２】
本願は、２００１年６月２５日にＡｎｇｕｓ　Ｄｕｎｃａｎ　Ｒｉｃｈａｒｄｓにより出願された“ＶＴＶシステム”という名称の米国特許出願番号第０９／８９１，７３３号により優先権を主張するものである。
【０００３】
通常のテレビ画面や映画スクリーンは現代の娯楽の一般的な形態であるが、これらの従来のシステムは、映画の動きと関連して視聴者のビューイング方向及び相対位置のユーザによる制御を可能にするものではない。さらに、バーチャルリアリティの表示がポピュラーになりつつあるが、３次元画像を生成する従来のシステムは複雑な場合が多く、メディアグラフィックは作成にコストのかかるものである。
【発明の開示】
【課題を解決するための手段】
【０００４】
以下の発明は、拡張型立体テレビジョン様の視聴体験を生成するハードウェア構成全体に関する。通常のテレビとは異なり、本システムの場合、視聴者は映画の動きと関連して視聴者のビューイング方向と相対位置との制御を行うことが可能となる。専用のハードウェア構成に加えて、本発明はまた、このバーチャルリアリティ様の体験を可能にする新しいビデオフォーマットにも関する。さらに、上記目標の実現を容易にするいくつかの独自のビデオ圧縮規格も定義される。ＶＴＶシステムは、従来方式の２次元の画像撮影法と真のバーチャルリアリティとの間の中継技術として設計される。
【０００５】
ＶＴＶシステムの進化にはいくつかの段階が存在するが、その最も基本的な形であるパノラマ表示方式から、その最も優れた形式である、アニメーション化されたテクスチャマップを利用するフルオブジェクトベースのバーチャルリアリティを特徴化（ｆｅａｔｕｒｅ）し、さらに、“視聴環境を意識した”拡張リアリティシステム全体で実物の俳優および／またはコンピュータで生成したキャラクタを特徴化する方式の範囲にまで及ぶ段階がある。
【発明を実施するための最良の形態】
【０００６】
添付図面と関連して以下に記載の詳細な説明は、本発明の現在の好ましい実施例の説明を目的とするものであり、本発明の構成および／または利用の唯一の形態を表すことを意図するものではない。本説明は、例示される実施形態と関連して本発明を構成し、作動させる機能及び一連のステップについて記載するものである。しかし、本発明の精神と範囲内に含まれることが意図される異なる実施形態によっても、同じあるいは同等の機能及びシーケンスの達成が可能であると理解されたい。

【０００７】
１．１）以下の発明は拡張型立体テレビジョン様の視聴体験を生成するハードウェア構成全体に関するものである。通常のテレビとは異なり、本システムの場合、視聴者は映画の動きと関連して視聴者のビューイング方向と相対位置との制御を行うことが可能となる。専用のハードウェア構成に加えて、本発明はまた、このバーチャルリアリティ様の体験を可能にする新しいビデオフォーマットにも関するものである。さらに、上記目標の実現を容易にするいくつかの独自のビデオ圧縮規格も定義される。ＶＴＶシステムは、従来方式の２次元の映画撮影法と真のバーチャルリアリティとの間の中継技術として設計される。ＶＴＶシステムの進化にはいくつかの段階が存在するが、その最も基本的な形であるパノラマ表示方式から、その最も優れた形式である、アニメーション化されたテクスチャマップを利用するフルオブジェクトベースのバーチャルリアリティを特徴化し、さらに、“視聴環境を意識した”拡張リアリティシステム全体で実物の俳優および／またはコンピュータで生成したキャラクタを特徴化する方式の範囲にまで及ぶ段階がある。
【０００８】
１．２）図１に見られるように、本ＶＴＶシステム全体は、中央画像処理装置（ＶＴＶプロセッサ）と、各種ビデオ入力装置（ＤＶＤ、ＶＣＲ、衛星、地上テレビ、遠隔ビデオカメラ）、赤外線リモートコントロール、デジタルネットワーク接続部、及びいくつかの出力装置接続部とから構成される。本ＶＴＶユニットは、図２に図示するようなその最も基本的な構成では、従来型のテレビ装置へ画像を出力することになる。このような構成では、（おそらく赤外線方式の）リモートコントロール装置を用いて、このＶＴＶ環境内で視聴者の所望のビューイング方向と位置の制御が行われることになる。この“基本システム構成”の利点として、現在の視聴覚技術を利用してこの構成の実現が可能であるという点が挙げられる。ＶＴＶ画像規格は、標準的ビデオのレイヤー（ｌａｙｅｒ）の上位“レイヤー”と考えることができる（上位）互換性を有する画像規格である。すなわち、従来のビデオは新しいＶＴＶ画像規格のサブセットを表すものとなる。この規格が互換性を有する結果、テレビメーカーおよび／または視聴覚機器メーカーは仕様の大きな変更をまったく必要とせずＶＴＶを導入することが可能となる。さらに、ＶＴＶ互換テレビ復号化用ユニットは従来方式のテレビ送信と本質的に互換性を有するものである。
【０００９】
１．３）図３に図示するようなさらに優れた構成では、ＶＴＶシステムは表示装置として無線ＨＭＤを使用する。このような構成では、単純な画像表示に加えてトラッキング装置として無線ＨＭＤを使用することが可能である。このトラッキング情報は、最も基本的な形では、単にビューイング方向の制御から成るものであってもよい。さらに優れたシステムでは、バーチャルな環境において視聴者のビューイング方向と位置の双方の確定が可能となる。最終的に、最も優れた実施構成では、ＨＭＤの遠隔カメラによって、リアルワールド画像がＶＴＶシステムに入力され、ＶＴＶシステムはこのリアルワールド画像を解釈してこれを空間オブジェクトに変換し、次いで、この空間オブジェクトはバーチャルなオブジェクトと置き換えられ、それによって、“視聴環境を意識した”拡張リアリティシステムの提供が可能となる。
【００１０】
１．４）無線ＨＭＤは、無線データリンク“サイバーネットリンク”によってＶＴＶプロセッサに接続される。このリンクは、その最も基本的な形ではＶＴＶプロセッサからＨＭＤへビデオ情報を送信し、ＨＭＤからＶＴＶプロセッサへトラッキング情報を送信する能力を有する。サイバーネットリンクは、その最も優れた形では、ＨＭＤからＶＴＶプロセッサへのトラッキング情報の転送に加えてＨＭＤとサイバーネットリンク間でのビデオ情報の送受信も行うことになる。さらに、ＶＴＶプロセッサのある構成要素を遠隔ＨＭＤの中に組み込むようにし、それによってサイバーネットリンクを介するデータ転送要件を提言するようにしてもよい。この無線データリンクは、（未圧縮あるいはデジタル圧縮のいずれかのフォーマットで）アナログデジタルビデオ送信またはデジタルビデオ送信のいずれかを利用する多様な異なる方法で、トラッキング情報用の、二次のデジタルで符号化し

たデータストリームを用いて実現することができる。あるいは、これらのチャネルの双方を運ぶ純粋にデジタルな単方向のまたは双方向のデータリンクを組み込むこともできる。データ転送用の実際のメディアはおそらくマイクロウェーブや光である。しかしいずれの転送メディアも適切なものとして利用可能である。本システムの好ましい実施形態は、ＨＭＤのボード上やおそらくＶＴＶ基地局上の画像解析用ハードウェアと共にＨＭＤに取り付けられたオンボードのパノラマカメラを利用してリアルタイムのトラッキング情報を出力する形態となる。システム精度のさらなる向上を図るために、逆反射作用を有するマーカーを”リアルワールド環境で”利用することも可能である。このような構成では、オンボードカメラの光軸の近くに配置される切り替え可能な光源をこれらのカメラと共に利用して“差分画像解析”システムが形成される。このようなシステムは、直接ビデオ画像を単独で利用する場合よりも著しく高い認識精度を特徴とする。

【００１１】

１．５）最終的に、本ＶＴＶシステムは“ユニバーサル画像規格（ｕｎｉｖｅｒｓａｌ ｇｒａｐｈｉｃｓ ｓｔａｎｄａｒｄ）”を利用して画像情報の転送を行うことになる。このような標準規格では、オブジェクトベースの画像記述言語が組み込まれ、この言語によってサブシステム間での“共通の画像知識ベース”による高圧縮が達成されることになる。本発明は、基本的な用語で、この画像言語の進化における３つのレベルの漸進的精巧化（ｓｏｐｈｉｓｔｉｃａｔｉｏｎ）について記載するものである。

【００１２】

１．６）本特許の目的として以下の３つの圧縮規格について説明する：

ａ）ｃ－ｃｏｍ

ｂ）ｓ－ｃｏｍ

ｃ）ｖ－ｃｏｍ

１．７）ＶＴＶシステムは、その最も基本的なフォーマットにおいて、視聴者を取り囲む３６０°のパノラマ表示スクリーンと考えることができる。

【００１３】

１．８）上記“バーチャル表示画面”は多数の“ビデオページ”から構成される。“ページキーコード”がビデオ画像内で符号化され、この“ページキーコード”によって、上記“バーチャル表示画面”内の特定の位置内へ画像情報を入力するようにＶＴＶプロセッサへの指示が行われる。動的に画像を入力する上記能力の結果、高解像度レートと高いフレームレートのいずれにも著しい犠牲を生じることなく、有効な同等の高解像度レートと高いフレームレートの双方の達成が可能となる。例えば、高速で変化している画像セクションのみが高速の画像更新を必要とするのに対して、画像の大部分は一般に静的である。（一般に動いている）重要な要素が主要場面に配置される従来方式の画像撮影法とは異なり、パノラマ画像の大部分は一般に静的である。

【００１４】

（ＶＴＶ画像規格：）

２．１）ＶＴＶ画像規格は、その最も基本的な形において、バーチャルな３６０°パノラマ表示画面から構成され、ＶＣＲ、ＤＶＤ、衛星、カメラあるいは地上テレビ受信装置のような外部のビデオソースから得られるビデオ画像をこのパノラマ表示画面上に描画することが可能であり、その結果、各ビデオフレームには、ビデオ情報だけでなく、バーチャルな表示画面内における各ビデオフレームの位置を定義する情報も含まれる。このようなシステムは、可変解像度画像だけでなく、フレームレートに依存しない画像も出力するので、著しく汎用性が高い。すなわち、ある特定のバーチャルな画像（バーチャル表示画面全体）内の実際の更新率は表示画面自体の範囲内で変動すると考えられる。この更新は、自身のバーチャル位置情報を含む各フレームによって達成される。これによって、ほとんどあるいはまったく変化しない画像上の非更新セクションの名目上の認知を代償にしてバーチャル画像のアクティブ領域の更新を迅速に行うことが可能となる。このようなシステムが図４に図示されている。

【００１５】

２．２）画像のリアリズムのさらなる改善を図るために、図５に図示のフォーマットに合わせて基本ＶＴＶシステムの拡張を行うことができる。この構成では、円筒形状のバーチャル表示画面が先端を切り落とした球面としてＶＴＶプロセッサにより解釈される。ＶＴＶプロセッサのデジタル処理ハードウェアコンポーネント内のジオメトリトランスレータ、すなわち”ワープエンジン”をしようすることによって、上記効果を容易に生み出すことが可能となる。

【００１６】

２．３）絶対基準平面が絶えず変動することから、（ＨＭＤベースまたはＰａｎ－Ｃａｍベースのいずれかの）移動カメラ用アプリケーションでは、ＶＴＶグラフィックエンジンが画像を正しく復号化できるように、カメラシステムの方位角と高度に関連するトラッキング付加情報が視覚情報と共に含まれることが求められる。このようなシステムでは、カメラの絶対方位角と高度が画像フレーム情報の一部となる。この絶対基準データを解釈するいくつかの可能性のある技法が存在する。まず、座標データを用いてメモリ書き込み処理中にメモリ内に像面の原点を画定することができる。残念ながら、上記アプローチでは、異なるアラインメント値を持つ前回のフレームからの残りの画像フラグメントが結果としてメモリ内に残される傾向がある。さらに実際的な解決方法として、単に、０方位角、０高度の想定基準点と共にビデオ情報をメモリの中へ書き込む方法がある。次いで、カメラアングルの補正値に対応する表示ビューポートの補正により上記ビデオ情報の正しい表示が行われる。このようなシステムの１つの可能なデータフォーマットが図１１及び表１に図示されている。

【００１７】

【表１】

制御フィールド	ビット	値	割当てキー
フレームフリップ	FF		フリップメモリフレーム
フィールドタイプ	F2-F0	0	前景ビデオ（ＡＲＭ）
		1	背景ビデオ（ＶＲＭ）
		2	デジタルハイブリッド（ＴＭ）
		3	デジタルビットストリームフレーム（ＴＭ）
		4	予約
		5	予約
		6	予約
		7	デジタル制御フレーム
ページ番号	P3-P0	0-15	（ＭＥＭレイアウトに依存）
オーディオ同期	AS		オーディオバッファを０にリセット
高度補正	E8-E0	(±45°)	カメラの高度
方位角補正	A10-A0	(±180°)	カメラ方位角
オーディオトラック	T1-T0	0	オーディオトラックなし
		1	４オーディオトラック
		2	８オーディオトラック
		3	オブジェクトベースのオーディオ
オーディオサンプルレート	S1-S0	0	２／４回線（１５Ｋ　Ｓ／Ｓ）
		1	３／６回線（２３Ｋ　Ｓ／Ｓ）
		2	４／８回線（３１Ｋ　Ｓ／Ｓ）
		3	５／１０回線（３８Ｋ　Ｓ／Ｓ）

（オーディオ規格：）

２．４）３６０°パノラマビデオに加えて、４トラック（４チャネル録音再生方式）または８トラック（８チャネル録音再生方式）のいずれかの立体的サウンドもＶＴＶ規格によってサポートされる。４トラックシステムのバーチャルな表現が図６に示されている。単純な４トラックオーディオシステムの場合、サウンドシステムの左右スピーカ（またはＨＭＤベースのシステムの場合、ヘッドホン）を通る音は、ビューポート（ＶＲ環境内での視方向）の方位角に基づいてスケーリングされる。８トラックオーディオシステムの場合、サウンドシステムの左右スピーカ（またはＨＭＤベースのシステムの場合、ヘッドホン）を通る音は、システムのバーチャルな表現（図７）に図示のように、ビューポートの方位角と高度の双方に基づいてスケーリングされる。

【００１８】

２．５）その最も基本的な形では、図１２に図示するようなデジタル／アナログハイブリッドフォーマットにおいて、ＶＴＶ規格によって、ビデオ情報の一部としてマルチトラックオーディオチャネルの符号化が行われる。その結果、現行機器とビデオとの互換性の達成が可能となる。この例図に見られるように、オーディオデータは、アナログコード化された圧縮フォーマットで格納されているため、各ビデオ走査線には５１２本のオーディオサンプルが含まれるようになる。このアナログコード化オーディオ情報に加えて、各オーディオ走査線には、オーディオ情報の“プレスケーリング”に用いられる３ビットデジタルコードが含まれる。すなわち、実際のオーディオサンプル値はＸ＊Ｓであり、ここでＸはプレスケール番号であり、Ｓはサンプル値である。上記２種の符号化方式を用いて、オーディオシステムのダイナミックレンジを約４３ｄＢから６０ｄＢ以上まで拡げることが

可能となる。第２に、このダイナミックレンジの拡張は、オーディオ品質に比して比較的“低コスト”で行われる。なぜなら、信号レベル全体が高ければ、人は音の歪みに対して相対的に敏感ではないからである。スタートビットはこのシステムでは重要な構成要素である。スタートビットの機能は、走査線に対して最大レベル（例えば１００％レベルすなわち白レベル）を設定することである。黒レベルと関連するこのレベル（このレベルはカラーバーストの直後にサンプリングすることができる）は、各ラインに対して０％と１００％のレンジを形成する。ラインベースで各ラインに対する０％と１００％のマークを動的に調整することにより、走査線のデジタル成分の復号化の精度が向上することに加えて、ビデオサブモジュールのＡＣ結合および／またはビデオメディアの録画と再生に起因して生じる黒レベルの変動に対するシステム感度がずっと低くなる。

【００１９】

２．６）このデジタル情報のプレスケーリングに加えて、オーディオ制御ビット（ＡＳ）が（ライン２１の）各フィールドに含まれる。この制御ビットによってオーディオバッファシーケンスは設定時０に設定される。この設定によって、４または８トラックオーディオ情報が同期されて、その結果、ビデオページの更新シーケンスに関係なく常に現在のデータから正しいトラックの更新が行われる。

【００２０】

２．７）別の構成では、ビデオ録画再生システム（典型的にはデジタルオーディオトラックや線形アナログオーディオトラック）により出力される従来方式のステレオオーディオを用いることによって、立体的サウンドの提供が可能となる。最善でも、このような従来方式のシステムでは、２つのオーディオトラックだけに４８ｋＨｚのサンプルレート（すなわち、２チャネル２４ｋＨｚオーディオ）が与えられるのみである。しかし、オーディオデータを圧縮することによって同じ記憶メディア上へさらに多数のオーディオトラック（例えば４または８オーディオトラック）をタイムドメインで符号化し、次いで、その結果得られるデータの時分割多重化を行うことが可能となる。受信側では、まず時分割多重化されたデータを復号化し、次いでオーディオデータを減圧縮（ｄｅｃｏｍｐｒｅｓｓｉｎｇ）することによって上記プロセスが逆に行われて、一連の連続するオーディオストリームが形成される。高いオーダーのノッチフィルタまたはハイパスフィルタを組み込んで、時分割多重化／逆多重化処理手順により生じるサンプリングノイズをすべて取り除くようにすることが好ましい。このようなフィルタは、おそらくデジタル設計のフィルタであり、単にＶＴＶプロセッサのデジタル処理用ハードウェアの一部にすぎない。このようなオーディオ符号化システムのための代表的な構成が図１４～１５に図示されている。

【００２１】

このようなシステムが簡単に実現可能であることは明らかである。オーディオチャネル数の増加による“パフォーマンス代価”は、システムの周波数応答が減少することである。この減少はチャネル数の上昇に直接比例する。例えば、チャネル数が２から４へ増加すると、周波数応答が２４ｋＨｚから１２ｋＨｚへ減少することになる。この減少レベルは容認できる範囲のレベルであり、最終的にこのオーディオの多重化システムは、比較的“ロバスト（ｒｏｂｕｓｔ）”で、かつ、ＶＴＶフォーマットマテリアルの実施と編集が容易であるという点から見て、顕著な利点を与えるものであるという点で、前述したビデオ符号化されたオーディオシステムにとって好ましいものとなり得るものである。

【００２２】

任意の時分割多重化システムの有効な実施において最も重要なファクタの１つとして、時分割符号化クロック及び復号化クロックの安定した精度の高い生成が挙げられる。これは絶対に重要なファクタである。なぜなら、復号化されたデータパケットのいかなるミスアラインメントも非常に顕著なチャネルクロストークに関連するノイズを結果として引き起こすことになるからである。本ＶＴＶシステムでは、ビデオストリームの中に同期信号（Ｈ同期やＶ同期、奇数／偶数などのいずれか）を用いて、ローカル復号化クロックを生成させることで、この問題に対する新規かつ非常に有効な解決方法を組み込む。これらの信号は一般に非常に正確であり、デジタル記録機器をＶＴＶデータの保存に利用する場合、

特に正確である。
【００２３】
好ましい実施形態では、時分割復号化ノイズ（ＮＴＳＣに対して毎秒６０スパイク）が相対的にほとんど発生しないこと、このノイズがデジタルノッチフィルタを使用することで（必要な場合）容易に除去できると考えられることを考慮して、復号化クロックとしておそらくＶ同期が使用されることになる。一方、（Ｈ同期などの）高周波時分割クロックを選択した場合、さらに多くの不連続インスタンス（毎秒約１５，０００）が発生する可能性があるが、同様に、デジタルフィルタ（この場合ローパスフィルタ）によってこれらのインスタンスを除去することが可能である。ＨＤＴＶやその他の高い走査速度（３１ｋＨｚまたはそれ以上の）のＴＶ規格を用いる場合、高速の復号化クロックの方が好ましいものになると考えられる。なぜなら、このような状況では、フィルタリングの要件が非常に簡単でかつ効果的になるためである。さらに、高速の復号化クロックを利用するシステムのメモリ要件が大幅に減少する。
【００２４】
２．８）コンピュータ用ＡＶファイル及びデジタルテレビ送信などのさらに優れたマルチメディア用データフォーマットでは、さらに効率的であるか、あるいは利点の多い別の方法でこれらの付加的オーディオトラックを格納することも可能である。
【００２５】
２．９）視聴覚装置としての利用に加えて、この空間でオーディオシステム／規格を、好適でコンパクトなトラッキング装置と１組のコードレスヘッドホンとを組み合わせることによって、オーディオ専用モードで使用して、高性能のハイファイ機器用空間オーディオシステムの実現を図ることも可能であることに注意されたい。
【００２６】
（拡張機能）
２．１０）上記の単純化した画像規格に加えて、単独でまたは基本ＶＴＶ画像規格と共に使用可能な多数の拡張機能がある。後続する発明で、ａ）ｃ－ｃｏｍ、ｂ）ｓ－ｃｏｍ、ｃ）ｖ－ｃｏｍとして知られる３つの画像規格について詳細に説明するが、これは本発明の目的のために行うものである。
【００２７】
２．１１）最初の２つの規格が空間画像オブジェクトの定義に関するものであるのに対して、第３の画像規格は、サブセットとして第１の規格を利用し、追加の環境定義と制御アルゴリズムとを組み込む完全なＶＲ環境定義言語に関する規格である。
【００２８】
２．１２）（最も基本的な形での）ＶＴＶ画像規格は、従来方式のビデオ規格（ＮＴＳＣ、ＰＡＬなど）の層の上位にある制御層と考えることができる。したがって、ＶＴＶ画像規格は従来方式のアナログビデオ送信規格に制限されることは全くない。基本的に同一の技術を用いて、本ＶＴＶ規格は、ＨＤＴＶ規格並びにコンピュータグラフィック及び業界の視聴覚規格の多くと共に機能することができる。
【００２９】
２．１３）表１と図１１で概説したデジタル制御情報は、本システムの一つの可能な構成を表すものにすぎず、現行のアナログＴＶシステムとの互換性を有する仕様の一般性を減じるものではない。制御ビットの別の構成および／またはこの情報を符号化してアナログＴＶなどのようなビデオフォーマットに変換する別の方法が存在し、そのような構成および／または方法も同様に有効である。例えば、ＭＪＰＥＧやＭＰＥＧ-ＩＩなどの多くのデジタル符号化処理によって、連続するビデオ走査線が平均化されることが多く、図１１に図示するような単一ラインとして符号化される場合、デジタル情報が損なわれる傾向があることが知られている。しかし、一連のラインまたは連続するラインとしてこの情報を符号化すれば、この問題は解決される。このような構成では、データは、一連の直線分ではなく、結果として得られる画像内の一連の小さな矩形として符号化されるデータと考えることができる。このような変更は予想されるものであり、別のビデオフォーマット、特

にＨＤＴＶなどのようなデジタル符号化されたフォーマットを用いてパフォーマンス効率の向上を図る、さらに別の適合化が予想される。
【００３０】
（ＶＴＶプロセッサ：）
３．１）ＶＴＶグラフィックプロセッサはＶＴＶシステムの核心部である。このモジュールは、その最も基本的な形で、表示装置（従来方式のＴＶ／ＨＤＴＶあるいはＨＭＤのいずれか）へ出力画像をリアルタイムで生成する役割を果たすものである。ＶＣＲ、ＤＶＤ、衛星、カメラあるいは地上テレビ受信機のようなビデオメディア供給装置から入力される生の画像情報のデジタル化に加えて、このモジュールのさらに優れたバージョンでは、インターネットやその他のネットワーク接続を介して渡される“ユニバーサル画像言語（ｕｎｉｖｅｒｓａｌ　ｇｒａｐｈｉｃｓ　ｌａｎｇｕａｇｅ）”から得られる画像をリアルタイムで描画することも可能である。このデジタル化タスクと描画タスクに加えて、ＶＴＶプロセッサは画像解析を実行することも可能である。本システムの初期のバージョンは、ＨＭＤの座標のトラッキングを確定することを目的としてこの画像解析機能を利用するものである。このモジュールのさらに優れたバージョンでは、上記トラッキング情報の提供の他に、ＨＭＤから得られるリアルワールド画像の物理的３次元オブジェクトとしての解釈も行われる。これらの３次元オブジェクトはユニバーサル画像言語で定義され、次いで、それを記録し、インターネットやその他のネットワークを介して同様の遠隔地にある表示装置へ通信したり、あるいは、類似する物理的サイズの別のバーチャルなオブジェクトにより置き換えたりすることが可能であり、それによって、真の拡張リアリティ体験が創り出される。
【００３１】
３．２）ＶＴＶハードウェア自体は以下のようなサブモジュール群から構成される。：
ａ）ビデオデジタル化モジュール
ｂ）拡張リアリティメモリ（ＡＲＭ）
ｃ）バーチャルリアリティメモリ（ＶＲＭ）
ｄ）変換メモリ（ＴＭ）
ｅ）デジタル処理ハードウェア
ｆ）ビデオ生成モジュール
３．３）これらのモジュールの正確な構成は他の外部ハードウェアによって決まる。例えば、デジタルビデオソースを用いる場合、ビデオデジタル化モジュールは比較的単純なものとなり、１グループのラッチバッファまたはＦＩＦＯバッファのみから構成される。しかし、コンポジットビデオ入力やＹ／Ｃビデオ入力を利用する場合、これらの信号をデジタル形式へ変換するための追加ハードウェアが必要となる。さらに、ビデオ入力ソースとしてデジタルＨＤＴＶ信号を使用する場合、（ＨＤＴＶ信号は圧縮されたフォーマットでは処理できないため）システムのフロントエンドとしてＨＤＴＶデコーダが必要となる。
【００３２】
３．４）アナログＴＶなどのようなフィールドベースのビデオシステムの場合、ＶＴＶグラフィックエンジンの基本処理は以下のようになる：
ａ）ビデオ情報がデジタル化され、０°方位角、０°高度の絶対ページ基準を想定して、フィールドベースで拡張リアリティメモリ内に入れられ、各ページの原点がページ番号ビット（Ｐ３～Ｐ０）の状態により確定される。
【００３３】
ｂ）背景および／または床／天井マップに関連する補助ビデオ情報は、“フィールドタイプ”ビット（Ｆ３～Ｆ０）とページ番号ビット（Ｐ３～Ｐ０）の状態によって決まる。フィールドベースでバーチャルリアリティメモリ内にロードされるか、セクション３．１２に記載のような拡張リアリティメモリ内に入れられるビデオ情報の解釈によりリアルタイムで生成される。
【００３４】
ｃ）デジタル処理用ハードウェアは、拡張リアリティメモリとバーチャルリアリティメモ

リの中に保持された上記情報を解釈し、ジオメトリ処理用エンジン（ワープエンジン）、デジタル減算画像処理、及び“ブルースクリーニング（ｂｌｕｅ－ｓｃｒｅｅｎｉｎｇ）”の新しい汎用性の高い形式との組み合わせを利用してこのデータを変換し、選択的にこのデータを合成して、ビデオマテリアルの撮影時に、パノラマカメラの在る位置と同じ位置に視聴者が立っていた場合に、その目に映る画像とほぼ類似の画像に変換する。この画像と、従来方式のビデオ技法を利用して入手可能な画像との間の主要な相違点として、前者の画像は、３６０°のパノラマ画像であるだけでなく、複合した没入型の拡張リアリティ体験を形成できるように一体に融合されたバーチャルリアリティと“リアルワールド”画像の双方の要素を備えることも可能であるという点が挙げられる。

【００３５】

ｄ）バーチャルリアリティと“リアルワールド画像”とを合成する正確な方法はＶＴＶプロセッサの作動モードによって決まるが、この方法については、本明細書の後のセクションでさらに詳細に説明する。この特別のＶＴＶプロセッサモードは、ソースメディアの中に存在する追加の制御情報により決定される。したがって、ＶＴＶメディアのソースを表示しながら、処理モードと表示モードとを動的に変更することが可能である。

【００３６】

ｅ）次いで、ビデオ生成モジュールは、従来方式のテレビまたはＨＭＤ表示装置に表示される単一のビデオ画像または一対のビデオ画像を生成する。（マルチスピンＤＶＤ装置が画像メディアとして使用されない場合）ＶＴＶ画像フィールドは最大フレームレート未満で更新されるが、画像の描画は、立体的サウンドの更新の場合のように、そのまま最大のビデオフレームレートで行われる。これが可能である理由として、各“画像球面”が、視聴者の任意の方角（方位角と高度）に関連するビデオとオーディオ双方の必要な情報のすべてを含むということが挙げられる。

【００３７】

３．５）図９に見ることができるように、ＶＴＶプロセッサのメモリの書き込み側は（ＡＤＣの）２つの別個のビデオ入力段を示す。ＡＤＣ－０は一般に生のパノラマビデオの送出に用いられ、ＡＤＣ－２は一般に予め描画されたビデオマテリアルからのバーチャルリアリティビデオの送出に用いられるものではあるが、双方のビデオ入力段は、拡張リアリティメモリとバーチャルリアリティメモリの双方へ完全にアクセスする（すなわち、これらのメモリではメモリプールが用いられる）ことに注意されたい。上記ハードウェア構成では、設計時にさらに多くの汎用性がサポートされ、いくつかの通常のものとは異なる表示モードの実施が可能である（これについては後のセクションでさらに詳述する）。同様に、ビデオ出力段（ＤＡＣ－０とＤＡＣ－１）はバーチャルリアリティメモリと拡張リアリティメモリの双方へ完全にアクセスする。

【００３８】

３．６）２つの入力段と２つの出力段とを備えることにより、設計の汎用性の向上が図られるが、メモリプールスタイルの設計により、１つ乃至２つの入力段のいずれかおよび／または出力段によってシステムが機能することが可能になり（但し機能は低下するが）、したがって、ある特定の実施構成において１つまたは２つの入力段あるいは出力段のいずれかが存在すれば、仕様の一般性が制限を受けるものではない。

【００３９】

３．７）設計の容易さのために、高速スタティックＲＡＭがプロトタイプ装置のビデオメモリとして利用された。しかし、この設計仕様の一般性を制限することなく別のメモリ技術の利用も可能である。

【００４０】

３．８）好ましい実施形態では、デジタル処理ハードウェアは１以上のフィールドプログラマブル論理アレイやカスタムＡＳＩＣの形をとる。フィールドプログラマブル論理アレイを用いる利点としては、ハードウェアを何時でも更新可能であるという点が挙げられる。この技術の主な欠点としては、ＡＳＩＣほど高速ではないという点が挙げられる。あるいは、従来方式の高速デジタルプロセッサを利用して上記の画像解析および／または画像

生成タスクを実行することも可能である。
【００４１】
３．９）前述したように、おそらくＶＴＶハードウェア全体が携帯用ＨＭＤ装置内に存在するようになる程度まで、このハードウェアの特定のセクションをＨＭＤの中に組み込むことも可能である。このような場合、ＶＴＶ基地局のハードウェアは、ＨＭＤとインターネットやその他のネットワークとの間のリンクとしてのみ機能することになり、すべてのグラフィック画像生成、画像解析及び空間オブジェクト認識はＨＭＤ自体の中で行われることになる。
【００４２】
３．１０）注：ビューポートアドレスジェネレータの下位ビットは、生成された画像に対してバレル歪曲をかけるＸＹ画像軸のルックアップテーブルアドレス変換装置を介して実行される。この変換装置によって、ビューポートの現在の視野の正確な画像の歪曲が示される。このハードウェアは、おそらくＦＰＧＡやＡＳＩＣ論理回路内に実装され、ビューポートアドレスジェネレータの機能ブロックの一部を含むことになるため、図１０には明示されていない。同様の方法で最終画像の同じような回転が実現される。
【００４３】
３．１１）ビューポート－０だけが変換エンジン（ワープエンジン）により影響を受け、ビューポート－１は歪みを受けずに読み出されることに注意されたい。スーパーインポーズ及びオーバーレイ拡張リアリティモードを使用する場合、この読み出しが必要となる。なぜなら、記憶装置から再生されるＶＲビデオマテリアルは格納される前に予め“平坦化”（すなわち、ピンクッション歪み）されているのに反して、ＨＭＤに備えられたパノラマカメラから得られる生のビデオ画像は、本システムによって拡張リアリティモードで表示される前に歪み補正を必要とするからである。上記予備的歪み補正を行った後、ＨＭＤのパノラマカメラにより録画された画像は、それ自体新しいＶＲマテリアルとして、幾何学的に正確でかつ格納に適したものにすることが望ましい（すなわち、これらの画像はＶＲマテリアルになることができる）。この場合、ワープエンジンの主要な役割の１つとしてはＨＭＤのパノラマカメラのジオメトリ補正とトリミングとがある。この中にはカメラの視野間で切れ目のない移行を行うという複雑なタスクが含まれる。
【００４４】
（例外処理）
３．１２）図４、５に見られるように、ＶＴＶ画像フレームは円筒形または先端を切り落とした球面のいずれかから構成される。この空間は、視聴者に対して有限の対頂角（プロトタイプで＋／－４５°）でしか対していない。これは、ビデオ格納装置の利用可能なデータ帯域と送信メディアとを最大限に活用できるように設計された意図的な制限であり、これによって、現行ビデオシステムとの互換性が保たれる。しかし、この妥協の結果、ビューポートが画像データの範囲を超える状況が生じる可能性がある。第一に、この例外的状況を処理できるいくつかの異なる方法がある。この例外的状況を処理する最も単純な方法として、単に境界線外のビデオデータを黒にする方法が挙げられる。これによって、天井と床が黒い部屋に居る外観が得られる。上記とは別に、さらに優れたアプローチとして、“床”と“天井”（すなわちＶＴＶ画像フレームの底部と最上部に存在する画像情報に基づく、ＶＴＶ画像フレームの下方及び上方の空間）を形成するアプローチがある。特定の一実施構成では、ＶＴＶ画像フレームの最上部と底部“リム”に対して赤、緑、青のピクセルの平均値を用いて、純色表面または好適にはグラデーションを施した色表面のいずれかを形成しその環境の“天井”と”床”とを表すようにすることも可能である。戸外で生成されるＶＴＶビデオの状況では、通常天井が青のシェードで、床はほぼ地面の色という結果が生じることになる。ビューポートの高度に伴ってシェーディングを変動させることによって生成される“床／天井”の現実感が改善され、その結果一般にビューポートの高度が＋／－９０°の極値に近づくにつれて、輝度が上昇する。この効果は、比較的単純な数学を利用して簡単に生成することができ、おそらくＶＴＶグラフィックエンジンの一部としてハードウェア内で実現される。

【００４５】
しかしさらに別の好ましい構成として、低い解像度で最大３６０°＊１８０°の背景画像マップを格納する２次的ビデオメモリによる格納を利用する構成がある。このメモリ域はバーチャルリアリティメモリ（ＶＲＭ）として知られている。拡張リアリティメモリ及びバーチャルリアリティメモリの両方を利用するシステム用の基本メモリマップが（変換メモリと併せて）図８に示されている。上記例図でわかるように、この変換メモリ域は最大３６０°＊１８０°をカバーできる程の十分なレンジを持つ必要があり、理想的には、（３６０°＊９０°をカバーする）拡張リアリティメモリバンクの角度分解能と同じ角度分解能を持つ必要がある。このような構成によって、床及び天井に対する例外処理と、前景の窓を通して見たり、窓の後ろの背景を示したりするような変動する透明画像との両方の提供が可能となる。これらの背景は静的あるいは動的のいずれかであることができ、ページ化されたフォーマットの利用により、前景（拡張リアリティメモリ）と基本的に同じ方法で更新が可能である。
【００４６】
（オペレーションモード）
３．１３）ＶＴＶシステムには２つの基本的オペレーションモードがある。これら２つのモード内にはいくつかのサブモードも存在する。この２つの基本モードは以下の通りである：
ａ）拡張リアリティモード
ｂ）バーチャルリアリティモード
（拡張リアリティモード１：）
３．１４）拡張リアリティモード１では、“リアルワールド画像”の選択成分がバーチャルリアリティの背景の上にオーバーレイされる。一般に、この処理では、まず“リアルワールド”画像から背景成分をすべて除去することが必要となる。これは、差分画像処理技術を用いて、すなわち現在の“リアルワールド”画像を前にとった格納済みのコピーと比較し、この２つの差分を検出することにより簡単に行うことができる。２つの画像の正確なアラインメントを行った後、異なる領域が新規のオブジェクトまたは前景のオブジェクトとなり、同じ状態のまま変化しない領域は静的な背景オブジェクトとなる。これは拡張リアリティモードの中で最も単純なモードであり、背景のほとんどは処理時に除去されることになるので、一般に十分な関心の対象となるものではない。移動式Ｐａｎ－Ｃａｍ（テレプレゼンス）や拡張リアリティモードで作動する場合、拡張リアリティメモリは、一般に、ランダムなページ更新ではなく、シーケンシャルなページ順に（すなわちシステムフレーム全体で）更新が行われることに注意されたい。このような更新を行う理由としては、撮影中のパノラマカメラシステムの位置と方角が常に変動するため、これらの変動を別々に処理した場合、画像ページの不一致が生じる可能性が高いということが挙げられる。
【００４７】
（拡張リアリティモード２：）
３．１５）拡張リアリティモード２は、前景の自動抽出と、オブジェクトの移動と、人工的背景環境におけるこれらの配置とに加えて、システムがワープエンジンも利用して追加の“リアルワールド”オブジェクトを背景の中へ“押し込む”という点で、モード１とは異なるものである。これらの“リアルワールド”オブジェクトを単純にバーチャルな環境の中へ追加することに加えて、ワープエンジンは、これらのオブジェクトのスケーリングと変換とを行う能力も有するため、さらに有効にこれらのオブジェクトがバーチャルな環境の中へマッチするようになる。これらのオブジェクトは、不透明なオーバーレイあるいはトランスペアレンシとして処理することができる。
【００４８】
（拡張リアリティモード３：）
３．１６）拡張リアリティモード３の場合、ワープエンジンを用いて、背景オブジェクトを前景の中へ“引き込み”、“リアルワールド”オブジェクトを置き換えるという点でモ

ード２とは異なるものである。モード２の場合のように、これらのオブジェクトの変換及びスケーリングを行うことが可能であり、不透明なオーバーレイあるいはトランスペアレンシのいずれかとしてこれらのオブジェクトの処理を行うことが可能である。これによって、ユーザは“リアルワールド”オブジェクトの物理的サイズと位置をバーチャルなオブジェクトと「一致」させることができるようになる。これを行うことにより、ユーザは、“リアルワールド”環境の場合と同じように、拡張リアリティ環境内でリアルタイムで交信し、ナビゲーションを行うことが可能となる。上記モードは、ユーザ自身の居間の中へハリウッド製作の映画を持ち込むことを可能にするものとして、娯楽目的やゲーム目的のためにおそらく最も利用が予想されるモードである。

【００４９】

（拡張機能）

３．１７）拡張リアリティモード２及び３を有効に作動させるキーとなるのは、高速かつ正確な光トラッキングシステムであることは明らかである。理論的には、ＶＴＶプロセッサが“リアルワールド”オブジェクトをリアルタイムで識別し、追跡することが可能である。しかし、特に、“リアルワールド”環境内における視聴者の物理位置の変化と共にオブジェクトジオメトリが大幅に変化するため、この識別と追跡は複雑なものとなる。したがって、単純な自動相関型のトラッキング技法が有効に機能しないことになる。このような状況では、逆反射作用を生じるいくつかのターゲットを当該オブジェクトのキー要素上に配置することにより、トラッキング精度の大幅な向上を図ることが可能である。比較的単純な差分画像処理技術の利用によりこのような逆反射作用を生じるターゲットを容易に識別することが可能である。

【００５０】

（バーチャルリアリティモード：）

３．１８）バーチャルリアリティモードは前述の拡張リアリティモードよりも機能的に単純なモードである。このモードでは、“予め撮影された”またはコンピュータで生成された画像がランダムなページベースで拡張リアリティメモリの中へロードされる。これが可能となるのはバーチャルカメラの基準平面が固定されているためである。前述の例の場合のように、バーチャルリアリティのメモリは固定した背景や動的背景についてはさらに低い解像度でロードされる。前景像面及び背景像面の双方を使用することで、運動視差などのさらに優れた画像処理技術が可能となる。

【００５１】

（拡張機能）

３．１９）拡張された“ブルースクリーニング”の形を利用することにより、バーチャルリアリティメモリ（背景メモリ）の汎用性の向上を図ることが可能となる。このようなシステムでは、“クロマキー”カラーのサンプルが背景フィールド（アクティブな画像領域の外側領域）における各走査線の開始時に与えられる。これによって、画像内で任意の色が許容可能となる汎用性の高いシステムが得られる。したがって、“透明な”クロマキーカラーによって個々のオブジェクトを囲繞することで、ワープエンジンによるこのオブジェクトの“カット＆ペースト”に関連する問題と不正確さとが大幅に減少する。さらに、前景のバーチャルリアリティ画像内の“透明な”クロマキー済みの領域を使用することにより、付加情報オーバーヘッドを伴うことなく、複雑な鋭く尖った領域および／または動的前景領域の簡単な生成が可能となる。

【００５２】

（カメラシステム：）

４．１）画像規格の定義でわかるように、移動式Ｐａｎ－ＣａｍやＨＭＤベースのビデオシステムにより取得される画像を正しく配置し、次いで表示するためには、追加のページ配置及びトラッキング情報が必要となる。さらに、立体的サウンドをリアルタイムで記録する場合、ビデオストリームの一部として上記情報の符号化も行わなければならない。コンピュータ生成画像の場合、描画段階でこの追加ビデオ情報を簡単に挿入することができる。これに対して、生のビデオを取得する場合、録画の前に上記追加トラッキング情報と

オーディオ情報とをビデオストリームの中へ挿入しなければならない。これは、本明細書で以後ＶＴＶエンコーダモジュールと呼ぶ画像処理用モジュールにより効果的に達成することができる。

【００５３】

（画像取得）

４．２）移動式パノラマカメラシステムにより収集される画像の場合、画像はまずＶＴＶエンコーダモジュールにより処理される。この装置によってビデオ歪み補正が行われ、ビデオページ情報、方角トラッキングデータ及び立体的サウンドデータがビデオストリームの中へ挿入される。ビデオ規格を変更せずに上記挿入が可能なため、現行の記録再生装置との互換性が保たれる。このモジュールはＶＴＶプロセッサ内へ組み込むこともできるが、最終的ビデオ情報を格納するか、何らかの形で無線ネットワークを介して送信するかのいずれかを必要とする遠隔カメラ用アプリケーションで使用するのには、このモジュールを別個の構成要素として設けるのが好都合である。

【００５４】

（トラッキングシステム：）

４．３）“Ｐａｎ－Ｃａｍ”やＨＭＤベースのカメラシステムのようないずれの移動式パノラマカメラシステムの場合も、“絶対”方位角と高度座標系を維持するために、トラッキング情報は、結果として得られるビデオストリームの一部を含む必要がある。コンピュータ生成画像の場合、カメラ方向は描画時にコンピュータシステムが認知する理論上の構成物であるため、上記データは必要ではない。

【００５５】

（基本システム：）

４．４）ＶＴＶ　ＨＭＤの基本トラッキングシステムでは、オンボードのパノラマビデオカメラを利用して、囲繞するリアルワールド環境に関する３６０°の必要な視覚情報が取得される。次いで、この情報は、自動相関のような計算上集中的な、しかしアルゴリズムの点からは比較的単純な技法を利用して、ＶＴＶプロセッサ（ＨＭＤ内に存在するか基地局ユニットとして存在するかの如何にかかわらず）により解析される。可能なアルゴリズムの例が図１６～２２に示されている。

【００５６】

４．５）図１６～２２に概説される単純なトラッキングシステムは位置と方角の変化のみを検出するシステムである。差分画像処理技術を用いることで容易に背景画像から識別できるいくつかの逆反射作用を生じるターゲットを追加することにより、絶対基準点を得ることが可能となる。このような絶対基準点は、ＶＴＶハードウェアがこれらのマーカーのリアルワールド座標を認知するかぎり、現実の環境内のどこに置かれるようなことがあっても、環境領域の末端部（すなわち、ユーザスペースの境界部）に位置することになると考えられる。（画像解析データから得られる）上記絶対基準点と差分運動との組み合わせによって、最大のビデオフレームレートでの絶対リアルワールド座標情報の生成が可能となる。所定の空間座標における逆反射作用を生じるターゲット配置の代替物として、アクティブな光ビーコンを用いることも可能である。これらの装置は、ビデオ取得レートと同期してストロボが発光するように構成され、結果として得られる画像に対する差分ビデオ解析が可能となるという点で、逆反射作用を生じるターゲットと同様の方法で作動する装置である。しかし、逆反射作用を生じるパッシブなターゲットとは異なり、アクティブな光ビーコンは、ビデオ取得と同期してストロボを発光することに加えて、上記ターゲットのリアルワールド座標を記述する付加情報をＨＭＤへ送信することも可能である。その結果、このシステムは、上記データを“リアルタイムで”取り出すことも可能であるため、これらビーコンの位置を明白に認知する必要がなくなる。従って、このようなシステムは、単純な逆反射作用を生じる構成よりもある程度頑丈な非常に汎用性の高いシステムとなる。パッシブなまたはアクティブなビーコンのいずれかの利用に加えて、双方の技法の組み合わせを用いることで、非常に汎用性の高いかつ強力なトラッキングシステムを形成することも可能である。逆反射作用を生じるターゲットの光源に対して“逆相”でアクティ

ブな光ビーコンにそのストロボ光を発光させることにより、光点が互いに１８０°位相を異にしたものになるという事実に基づいて、差分画像解析技術を利用して、背景環境からの画像抽出と、（光を出射する）アクティブビーコンか（光を反射する）パッシブビーコンのいずれかから発するような光点の効果的類別との双方を行うようにすることが可能となる。ＶＴＶプロセッサによってアクティブビーコンのタイミングを生成することが可能であるが、パッシブビーコン用として使用されるＨＭＤからの光ストロボパルスの検出により、あるいは、ＨＭＤからの赤外線パルスのような２次的光タイミング信号の利用により、自動的にタイミングを発生させるほうが好ましいであろう。このような構成では、アクティブビーコンをバッテリで駆動することも可能であり、追加の配線は何ら必要ではなくなる。何らかの光タイミング信号が存在しないときにこれらのアクティブビーコンがパワーダウンするように設計されていれば、作動していない時間中上記アクティブビーコンを停止させることさえ不要となるかも知れず、さらに、バッテリ消費量の大幅な増加を伴うことなく上記アクティブビーコンを永久に作動状態のままにすることも可能である。

【００５７】

アクティブビーコンは、単にその位置を単に示すことに加えて、移動ＨＭＤシステムへ付加情報を送信することが可能である。この情報は、作成される光パルスの色の変化とこれらの色の変化するタイミングとによって符号化される。最も単純な形では、各アクティブビーコンが様々な色の光を発生させることにより、簡単な分離と識別とを行うことが可能であるが、さらに優れた符号化技法もまた可能である。例えば、一連のカラーフラッシュによって或る特定の数を示すことが可能である。この数によって、ビーコンのコード番号を表すことも可能であり、あるいはさらに有用なこととして、ビーコンの物理的リアルワールド座標の符号化も可能である。このようなシステムは、シミュレーションの一部としてＨＭＤへ転送しなければならない“リアルワールド”の“ビーコンマップ”を必要とせず、代わりにアクティブビーコンシステムを含む任意の空間から必要な情報をピックアップすることが可能なので、非常に汎用性の高いシステムとなる。

【００５８】

４．６）注：図２３は、ある特定の運動の有無が自己相関器により単純に検出されるトラッキングハードウェアの単純化した図を示す。実用的システムでは、各クラスの運動用の複数の自己相関器を組み込むことになるであろう（例えば、１６個またはそれ以上の、水平運動を検出する別々の自己相関器が存在してもよい）。このようなシステムでは、すべての方向で異なるレベルの運動や異なる量の運動を検出することが可能となる。

【００５９】

（その他の構成：）

４．７）空間測位情報を達成するために天井のパターンのトラッキングが行われる同様の画像解析技術と、ＨＭＤ/Ｐａｎ－Ｃａｍシステムの角度方向を検出する単純な“チルトセンサ”とを利用する上記トラッキングシステムの代替実施構成が可能である。本システムの利点として、前述した最大６本の軸線を有する光トラッカに比べて著しく構成がシンプルで、低コストで済むという点が挙げられる。天井がＨＭＤから一定の距離と所定の方角にあるという事実に起因して、光学系、必要な画像処理装置の品質、及び後続する画像解析の複雑さとが大幅に単純化される。前述の６軸線光トラッキングシステムの場合のように、上記空間測位情報は本質的に相対運動の形のみの情報である。しかし、“絶対基準点”を追加することによって、このようなシステムによってその絶対基準値を再較正して絶対座標系全体を達成することが可能となる。いくつかの異なる技法を利用して、比較的簡単に上記絶対基準点の較正を行うことが可能となる。第１のそしておそらく最も単純な技法として、前述したような色に敏感に感応する逆反射スポットの利用がある。上記とは別に、（ＬＥＤビーコンのような）アクティブな光ビーコンの利用も可能である。利用可能なさらに別の絶対基準較正システムとして双方向赤外線ビーコンをベースとするシステムもある。このようなシステムにより、ＨＭＤとビーコンとの間で一意のＩＤコードの通信が行われ、それによって、ＨＭＤが上記“既知の空間基準点”のうちのいずれかの下を通過する毎に１回だけ較正が行われることになる。これは、複数回の原点リセットに起因

して生じる、較正ビーコンの近傍内にある“不感トラッキング領域”を回避するのに必要である。

【００６０】

（単純化：）

画像内の運動位置の検出に利用される基本的な自動相関技法を単純化して、かなり簡単な画像処理ステップにすることが可能である。まず、回転検出を単純化して、画像中心（カメラの光軸）の周りで左右対称を成す１グループの片寄り（上、下、左、右）にすることが可能である。さらに、横方向運動に関連するこれらの“サンプルポイント”は必ずしも非常に大きなものである必要はない。しかしこれらの“サンプルポイント”は固有の映像情報を含む必要がある。例えば、特徴のない空白な壁は有用なトラッキング情報をもたらさないが、オブジェクトの端縁や高輝度の強調点のような高いコントラスト領域を持つ画像は比較的簡単にトラックされる。この考え方をさらに一歩進めて、最初に画像全体を縮減して強調点／端縁に変えるようにすることが可能である。次いで、一連の水平方向及び垂直方向のストリップとして画像処理を行い、強調点／端縁間に自動相関領域の境界を設けるようにすることができる。さらに、現在の画像と前回の画像フレームとの比較を行い、これら画像間の“最も近い可能なフィット部分”（すなわち強調点の最小運動部分）を確定することによって、狭い強調領域のトラッキングを非常に簡単に行うことが可能となる。このような技法は、画像前処理のオーバーヘッドの若干をハードウェアにより処理すれば比較的簡単であり、適度の処理スピードを持つマイクロプロセッサの能力の範囲内に十分含まれる。

【００６１】

好ましい実施形態や特定の実施形態と関連して本発明を説明したが、本発明の範囲、あるいは本発明の発明概念から逸脱することなく、種々の変更及び追加の変形を行うことが可能であり、本発明の要素の代わりに同等の要素を用いることが可能であることを理解されたい。さらに、本発明の本質的範囲から逸脱することなく多くの改変を行って、特定の状況やマテリアルを本発明の教示に適合させることも可能である。したがって、本願明細書に開示された、本発明を実施するための特定の実施形態に本発明を限定することを意図するものではなく、本発明は添付の請求項の範囲に属するすべての実施形態例を含むものである。

【産業上の利用可能性】

【００６２】

本発明は、ユーザが映画の動きと関連して視聴者のビューイング方向と相対位置とを制御することを可能とする視聴覚システムを提供することを目的とする。

本発明の上記及び他の目的、利点及び産業上の利用可能性は添付の明細書と図面の再査により明らかである。

【図面の簡単な説明】

【００６３】

【図１】ＶＴＶシステム全体の概略図である。

【図２】基本構成に基づくＶＴＶシステムの概略図である。

【図３】高度の構成に基づくＶＴＶシステムの概略図である。

【図４】円筒形状のバーチャル表示フィールドを示す例図である。

【図５】先端を切り落とした球面のバーチャル表示フィールドを示す例図である。

【図６】４トラックサウンドシステムのバーチャル表現を示す例図である。

【図７】８トラックサウンドシステムのバーチャル表現を示す例図である。

【図８】拡張リアリティメモリとバーチャルリアリティメモリの双方のメモリを利用するシステムを表すＶＴＶメモリマップの描写である。

【図９】ＶＴＶプロセッサのデータ書き込みサイドを示すＶＴＶグラフィックエンジンの図である。

【図１０】ＶＴＶプロセッサのデータ読み出しサイドを示すＶＴＶグラフィックエンジンの図である。

【図１１】任意のアナログビデオと互換性を持つＶＴＶ符号化ビデオラインの例である。
【図１２】任意のアナログビデオと互換性を持つ音声データを含むことを示すＶＴＶ符号化ビデオラインの例である。
【図１３】周波数圧縮符号化を利用するＶＴＶにより符号化された音声を示す図である。
【図１４】周波数圧縮符号化を行うための結線図である。
【図１５】周波数圧縮復号化を行うための結線図である。
【図１６】位置と方角の変化を検出する光トラッキングシステムの図である。
【図１７】方角の方位角の変化を検出する光トラッキングシステムの図である。
【図１８】方角の高度の変化を検出する光トラッキングシステムの図である。
【図１９】方角の回転の変化を検出する光トラッキングシステムの結線図である。
【図２０】前後方向の位置変化を検出する光トラッキングシステムの結線図である。
【図２１】左右の位置変化を検出する光トラッキングシステムの結線図である。
【図２２】上下の位置変化を検出する光トラッキングシステムの結線図である。
【図２３】単純化したバージョンによる光トラッキングシステム用ハードウェアのブロック図である。

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/49267

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) :H04N 7/00, 9/43, 9/64; G09G 5/00; G02B 27/22, 5/08, 5/20
US CL :348/32, 36, 37, 38, 39; 345/8, 419; 359/360, 462, 479
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
U.S. : 348/32, 36, 37, 38, 39; 345/8, 419; 359/360, 462, 479

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,703,604 A (MCCUTCHEN) 30 DECEMBER 1997, col. 4, lines 4-63.	1-68
A	US 5,991,085 A (RALLISON et al) 23 NOVEMBER 1999, col. 5, lines 1-19.	1-68
A	US 6,064,423 A (GENG) 16 MAY 2000, col. 3, lines 41-67.	1-68
A	US 5,148,310 A (BATCHKO) 15 SEPTEMBER 1992, col. 3, lines 45-64.	1-68
A	US 3,872,238 A (HERNDON) 18 MARCH 1975, col. 1, lines 31-46.	1-68

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"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
"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
"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
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
08 MARCH 2002	08 APR 2002
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer CHRISTOPHER KELLEY Telephone No. (703) 305-4700

Form PCT/ISA/210 (second sheet) (July 1998)★

フロントページの続き

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