

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2012 年 11 月 22 日(22.11.2012)



(10) 国際公開番号
WO 2012/157516 A1

- (51) 国際特許分類:
G06T 19/00 (2011.01)
- (21) 国際出願番号: PCT/JP2012/061996
- (22) 国際出願日: 2012 年 5 月 10 日(10.05.2012)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2011-110160 2011 年 5 月 17 日(17.05.2011) JP
- (71) 出願人 (米国を除く全ての指定国について):
シャープ株式会社(SHARP KABUSHIKI KAISHA)
[JP/JP]; 〒5458522 大阪府大阪市阿倍野区長池町
2 2 番 2 2 号 Osaka (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 大津 誠
(OHTSU, Makoto), 瀬戸 幹生(SETO, Mikio).
- (74) 代理人: 岡田 宏之(OKADA, Hiroyuki); 〒2310041
神奈川県横浜市中区吉田町 7 2 番地サリュート
ビル 9 F なぎさ特許事務所 Kanagawa (JP).

- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

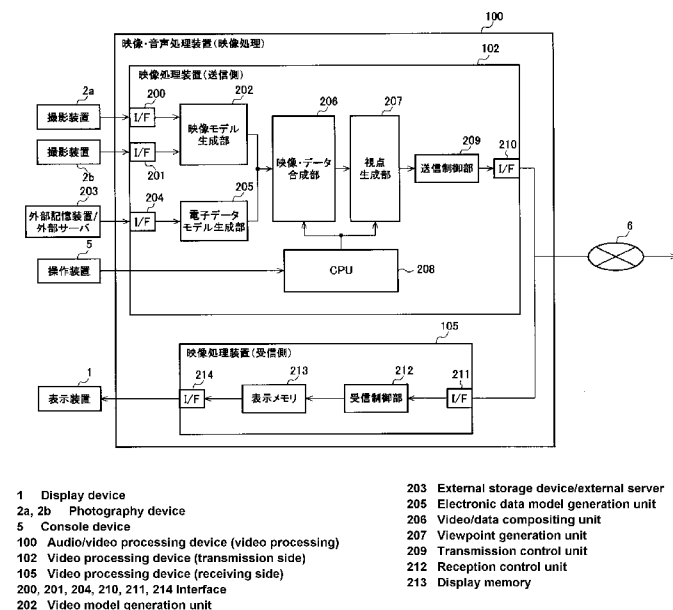
添付公開書類:

— 国際調査報告 (条約第 21 条(3))

(54) Title: VIDEO PRESENTATION SYSTEM, VIDEO PRESENTATION METHOD, PROGRAM, AND RECORDING MEDIUM

(54) 発明の名称: 映像提示システム、映像提示方法、プログラム及び記録媒体

[図3]



(57) Abstract: An objective of the present invention is to allow carrying out work while viewing an electronic work object in a shared positional relationship wherein workers in remote locations appear to be near one another. A video presentation system allows communication between remote locations by receiving an input of photographed video which is acquired by photography devices (2a, 2b), compositing electronic data with the photographed video, and transmitting and receiving same. This system comprises: a video model generating unit (202) which generates, based on depth information of the scope of the photography, a video model from the depth information and video information; an electronic data model generation unit (205) which generates an electronic data model on the basis of reference information of the inputted electronic data; a video/data compositing unit (206) which disposes a common coordinate system and composites the video model and the electronic data model within the common coordinate system; and a viewpoint generation unit (207) which installs a virtual viewpoint between common coordinate spaces and generates a video of the viewpoint. The video which is acquired with the viewpoint generation unit is transmitted.

(57) 要約:

[続葉有]

WO 2012/157516 A1



遠隔地作業者が隣にいるような位置関係で共通の電子作業オブジェクトを見ながら作業を行うことができるようにする。映像提示システムは、撮影装置（２ａ）、（２ｂ）によって取得された撮影映像を入力し、その撮影映像に電子データを合成して送受信することにより、遠隔地間のコミュニケーションを可能とする。本システムは、撮影する範囲の奥行き情報を基に、奥行き情報と映像情報とからなる映像モデルを生成する映像モデル生成部（２０２）と、入力した電子データの基準情報に基づいて電子データモデルを生成する電子データモデル生成部（２０５）と、共通の座標系を設けて、映像モデルと電子データモデルを共通の座標系内で合成する映像・データ合成部（２０６）と、共通の座標空間に仮想的な視点を設置して、該視点の映像を生成する視点生成部（２０７）とを有し、視点生成部で取得した映像を送信する。

明 細 書

発明の名称：

映像提示システム、映像提示方法、プログラム及び記録媒体

技術分野

[0001] 本発明は、映像提示システム、映像提示方法、プログラム及び記録媒体に関し、より具体的には、撮影画像に対してリアルタイムに電子データを合成表示することで、遠隔地のシステム間で、コミュニケーションをとって共同作業を行うことができるようにした映像提示システム、及び該システムの機能を実現する映像提示方法とプログラム及び記録媒体に関する。

背景技術

[0002] 複数の作業者によるフェイスツーフェイスの共同作業には様々な形態がある。例えば、フェイスツーフェイスの共同作業として、実在の物を使った作業（モックアップを用いたプロダクトデザイン確認、作業リーダによる製造装置の使用方法に関するレクチャー等）や、作業者がお互いに横に座りながら、あるいは側に立ち、共通の画面を見ながら計算機を使った作業（ペアプログラミング、作業画面を使った状況報告等）等があり、これらの作業は頻繁に行われている。

[0003] 撮影された実写映像に計算機で利用することができる電子データを合成表示することで、上記実際の作業に近いコミュニケーションを遠隔地間で実現することができる。さらに撮影された実写映像の3次元的な空間や位置関係を考慮して電子データを合成できると、2次元平面に単純に電子データを貼り付ける場合に比べて、視認性や物体位置認識効果を高めることができる。これらの研究は、拡張現実（AR：Augmented Reality）・複合現実（MR：Mixed Reality）と呼ばれ、多くの実施例がある。

[0004] 例えば、特許文献1には、送信者側の映像を受信者側に転送し、仮想空間内で送信者と受信者のアバターと呼ばれる人物映像を仮想空間の背景画像と合成しリアルタイムに表示することで、コミュニケーションや協調作業を行

うための装置について開示されている。送信者の映像は、送信者と受信者の仮想空間内の位置関係から、複数の角度に設置された撮影装置の中から適切な位置の映像が選択されて伝送されるような構成になっている。しかし、双方の利用者は仮想空間内の位置を考慮して、アバターを適切に配置する必要がある。

[0005] 例えば、非特許文献1には、利用者の前に設置された特殊な表示面、具体的には偏光された光がその後方に透過できない半透過の投影面を設置し、上部より偏光された遠隔地の映像をその投影面に映し出す。さらに、半透過投影面の背後に撮影カメラを設置し、透過してきた利用者の映像を撮影し遠隔地に送信する。また、遠隔地の映像内に電子データを重畳することで、双方の間に電子的な作業オブジェクトを表示させて協調作業を行うための装置について開示されている。

先行技術文献

特許文献

[0006] 特許文献1：特開2004-178036号公報

非特許文献

[0007] 非特許文献1：“Integration of Interpersonal Space and Shared Workspace: ClearBoard Design and Experiments” ACM Transaction on Information Systems, Vol. 11, No. 4, October 1993, Pages 349-375

発明の概要

発明が解決しようとする課題

[0008] 特許文献1で開示されている方法は、利用者自身が、仮想空間内の位置や向いている方向を常に意識する必要があることと、双方が同一の役割で利用できるシステムになっていないという課題がある。

非特許文献2で開示されているシステムは、特殊な投影面が必要であり、かつ投影面の上部に偏光した映像を投影するための装置と投影面の後方に利用者を撮影する装置を設置する必要がある、システムとして非常に大がかり

になる課題がある。

[0009] 本発明は、上記特許文献 1 の課題にあるような、利用者が仮想的な 3 次元空間位置を考慮する必要なく、共通に利用する電子データの 3 次元的な位置を介して、双方の利用者の位置及び表示装置の相対的な位置を確定させ、その情報に基づいて双方のシステムに送る映像を生成することで、遠隔地作業者が隣にいるような位置関係で共通の電子作業オブジェクトを見ながら作業を行うことができるようにすること、また、そのような作業を実現するシステムを汎用的なディスプレイや撮影装置を用いて実現することを目的とするものである。

課題を解決するための手段

[0010] 上記課題を解決するために、本発明の第 1 の技術手段は、撮影装置によって取得された撮影映像を入力し、該入力した撮影映像に、画像として生成された電子データを加工して合成し、該合成した映像を送受信することにより、遠隔地間でコミュニケーションを行なうことを可能とした映像提示システムであって、撮影する範囲の奥行き情報と、該奥行き情報を基に該奥行情報に貼り付けた映像情報とからなる映像モデルを生成する映像モデル生成部と、電子データを入力し、該入力した電子データの位置、大きさ、向きの情報を含む基準情報に基づいて電子データモデルを生成する電子データモデル生成部と、共通の座標系を設けて、前記映像モデルと前記電子データモデルを共通の座標系内で合成する映像・データ合成部と、該映像・データ合成部で合成されたモデルを入力し、前記共通の座標空間に仮想的な視点を設置して、該視点の映像を生成する視点生成部と、該視点生成部で取得した映像を送信する送信部と、を有することを特徴としたものである。

[0011] 第 2 の技術手段は、第 1 の技術手段において、前記視点生成部が、前記電子データの 3 次元的な位置と前記共通の座標系における利用者の相対的な位置関係に基づいて算出された視点位置の映像を生成することを特徴としたものである。

[0012] 第 3 の技術手段は、第 2 の技術手段において、利用者による操作入力を受

け付ける操作装置と、前記電子データの位置及び利用者の視点位置を設定するための表示画面を表示装置に表示させる制御手段とを有し、前記視点生成部は、前記表示画面に対する前記操作装置の操作入力に従って、前記電子データの３次元的位置及び利用者の視点位置を決定することを特徴としたものである。

[0013] 第４の技術手段は、第１～第３のいずれか１の技術手段において、利用者の３次元上の位置を計測する３次元位置計測装置を有し、前記視点生成部は、前記３次元位置計測装置による計測結果に従って、利用者の位置をリアルタイムに変更することを特徴としたものである。

[0014] 第５の技術手段は、第１～第４のいずれか１の技術手段において、前記撮影装置を複数備え、前記映像モデル生成部が、複数の前記撮影装置で撮影された映像から視差画像を生成し、該生成した視差画像に基づいて撮影された対象のデプスを示すデプスマップを生成し、該デプスマップから前記撮影装置の固有の座標系における奥行き情報を生成し、該撮影装置の固有の座標系を前記映像提示システムで扱う共通座標系に変換して、奥行き情報と該奥行き情報に貼り付けた映像情報とからなる前記映像モデルを生成することを特徴としたものである。

[0015] 第６の技術手段は、撮影装置によって取得された撮影映像を入力し、該入力した撮影映像に、画像として生成された電子データを加工して合成し、該合成した映像を送受信することにより、遠隔地間でコミュニケーションを行なうことを可能とした映像提示方法であって、撮影する範囲の奥行き情報と、該奥行き情報を基に該奥行き情報に貼り付けた映像情報とからなる映像モデルを生成する第１の過程と、電子データを入力し、該入力した電子データの位置、大きさ、向きの情報を含む基準情報に基づいて該電子データモデルを生成する第２の過程と、共通の座標系を設けて、前記映像モデルと前記電子データモデルを共通の座標系内で合成する第３の過程と、該第３の過程で合成されたモデルを入力し、前記共通の座標空間に仮想的な視点を設置して、該視点の映像を生成する第４の過程と、該第４の過程で取得した映像を送

信する第5の過程と、を有することを特徴としたものである。

[0016] 第7の技術手段は、撮影装置によって取得された撮影映像を入力し、該入力した撮影映像に、画像として生成された電子データを加工して合成し、該合成した映像を送受信することにより、遠隔地間でコミュニケーションを行なうことを可能とした映像提示装置のコンピュータに、撮影する範囲の奥行き情報と、該奥行き情報を基に該奥行き情報に貼り付ける映像情報とからなる映像モデルを生成する手順と、電子データを入力し、該入力した電子データの位置、大きさ、向きの情報を含む基準情報に基づいて該電子データモデルを生成する手順と、共通の座標系を設けて、前記映像モデルと前記電子データモデルを共通の座標系内で合成する手順と、該合成されたモデルを入力し、前記共通の座標空間に仮想的な視点を設置して、該視点の映像を生成する手順と、該生成された映像を送信する手順を実行させるための映像提示プログラムである。

[0017] 第8技術手段は、第7の技術手段の映像提示プログラムを、コンピュータ読み取り可能に記録した記録媒体である。

発明の効果

[0018] 遠隔コミュニケーションにおいて、双方がお互いに隣にいるような位置関係で、共通の電子データを見ながらコミュニケーションを行うことが可能になる。

図面の簡単な説明

[0019] [図1]本発明の映像提示システムの第1の実施形態を模式的に示す図である。

[図2]本発明の映像提示システムの第1の実施形態における映像・音声処理機能を示すブロック図である。

[図3]図2の映像処理機能を更に詳しく説明するためのブロック図である。

[図4]カメラ座標系に変換されたデプスマップから3次元的な位置に復元する方法を説明する図である。

[図5]本発明の映像提示システムの第1の実施形態において、システムと利用者の位置関係及び座標系を、従来のテレビ会議システムにおける位置関係を

例にして説明する図である。

[図6]本発明の映像提示システムの第1の実施形態において、利用者と電子データの仮想的な位置関係を説明する利用形態の一例を示す図である。

[図7]図6における利用形態において、双方の利用者が見る映像の模式図である。

[図8]本発明の映像提示システムの第1の実施形態において、遠隔地に送る映像を生成するための視点位置を算出する方法を説明するために用いる図である。

[図9]利用者が容易に位置関係を設定するためのグラフィカルなインタフェース図である。

[図10]本発明の映像提示システムの第1の実施形態において、システム全体の制御フローを示すフローチャート図である。

[図11]図10における、データ送信スレッド処理のフローチャート図である。

[図12]図11における、送信フレーム情報生成処理のフローチャート図である。

[図13]図10における、データ受信スレッド処理のフローチャート図である。

[図14]図10における、電子データ制御スレッド処理のフローチャート図である。

[図15]電子データ制御スレッド内で生成・伝送される電子データ属性情報の構造を示す図である。

[図16]データ送信スレッド内で生成・伝送される送信フレーム情報の構造を示す図である。

発明を実施するための形態

[0020] 以下、図面を参照しながら本発明の実施形態について説明する。

[実施形態1]

<システム構成>

図 1 は映像と音声を双方向通信できる本発明に係る映像提示システムの第 1 の実施形態を模式的に示す図である。広域のネットワーク 6 より先の遠隔地にも同じ映像提示システムが設置されており、映像提示システム間で、映像と音声を使ったコミュニケーションをリアルタイムに行うことができる。

[0021] 図 1 について詳しく説明する。本発明に係る映像提示システムの実施形態は、計算機 4 で実行され、計算機 4 の有する機能により映像・音声処理が行われる。計算機 4 は、例えば P C などの情報処理装置を適用することができる。

計算機 4 には、遠隔地の利用者 7 の映像、及び遠隔地の相手と一緒に作業を行っている電子データ 8 を表示する液晶ディスプレイなどの表示装置 1 と、遠隔地の音声を再生する音声再生装置 3 とが接続される。計算機 4 は、遠隔地から伝送されてきた映像と音声の情報を受け取り、表示装置 1 と音声再生装置 3 に映像や音声を提示するためのものである。

[0022] 表示装置 1 の上部には、利用者の映像と音声を一度に取得するマルチメディア情報取得装置 2 が設置されている。映像や音声を取得するためのマルチメディア情報取得装置 2 は必ずしも一体化している必要はなく、映像の取得と音声の取得を分離した装置で実現する構成も可能である。例えば、マルチメディア情報取得装置 2 をマイク及び撮影装置（カメラ）によって構成し、音声を取得するためのマイクを利用者の近くに置き、撮影カメラのみを表示装置 1 の上部に設置した構成としてもよい。また、マルチメディア情報取得装置 2 は、複数の、例えば 2 個のカメラを備えている。マルチメディア情報取得装置 2 は、上記遠隔地からの映像と音声を提示するための計算機 4 に接続されている。

[0023] 計算機 4 は、上記受信側の処理とは別に、遠隔地に映像や音声を伝送するための送信制御も行っている。

また、計算機 4 には、キーボードなどの操作装置 5 が接続されており、利用者が作業形態等の情報を入力したり電子作業オブジェクトの属性を変更することが可能で、システムを制御することが可能である。

計算機 4 は、ネットワーク 6 を介して、遠隔地の同一の映像提示システムと繋がっており、映像や音声の送受信を行うことができる。利用者 9 はこの映像提示システムを利用することで、遠隔地に設置された同システムの利用者とリアルタイムに電子データ 8 を見ながらコミュニケーションを行うことができる。

[0024] <ブロック構成>

(システム基本ブロックの説明)

計算機 4 の内部で行われる基本の処理について、図 2 のブロック図を用いて説明する。図 2 において、広域のネットワーク 6 より左側が上記利用者 9 の側を示し、右側が遠隔地の側を示している。

[0025] 利用者 9 側の映像提示システムにおける送信機能についてはじめに説明する。カメラやマイクで構成されるマルチメディア情報取得装置 2 で取得された映像と音声の信号は、映像・音声処理装置 100 の内部にある送信装置 101 に取り込まれる。映像・音声処理装置 100 は、図 1 の計算機 4 と等価な装置である。ここでは、映像信号は送信装置 101 内の映像処理装置（送信側）102 に、音声信号は音声処理装置（送信側）103 に入力され、遠隔地に送信するための信号処理が施された後に、外部に出力され、広域のネットワーク 6 を経由して遠隔地に伝送される。

[0026] 続いて、映像提示システムの受信機能について説明する。広域のネットワーク 6 を経由して遠隔地から送られてきた映像と音声の信号が、映像・音声処理装置 100 の内部の受信装置 104 に取り込まれる。映像信号は受信装置 104 内の映像処理装置（受信側）105 に入力され、音声信号は音声処理装置（受信側）106 に入力される。そしてこれら映像信号及び音声信号は、映像や音声を再現できる形に信号処理され、ディスプレイ（表示装置）やスピーカで構成されるマルチメディア情報提示装置 10 に出力される。

[0027] 遠隔地の処理についても同様で、映像・音声処理装置 107 は映像・音声処理装置 100 に、受信装置 109 は受信装置 104 に、送信装置 108 は送信装置 101 に、マルチメディア情報提示装置 12 はマルチメディア情報

提示装置 10 に、マルチメディア情報取得装置 11 はマルチメディア情報取得装置 2 に、それぞれ対応している。

[0028] <処理>

(映像・音声処理装置 100 の詳細説明)

映像・音声処理装置 100 について、図 3 のブロック図を用いて具体的に説明する。但し、音声処理については、支障がない限り説明を省略する。以下、上記利用者 9 が使用する映像提示システムにおいて、撮影された利用者の撮像画像に作業で扱う電子データを合成（重畳）した映像を遠隔地に伝送する流れと、遠隔地から送られてきた映像を受信して表示するまでの流れについて説明する。もう一方の遠隔地側の流れについては、利用者 9 のシステムと同じであるため説明は省略する。

[0029] 撮影装置 2a、2b は、マルチメディア情報取得装置 2 に含まれる装置であり、本映像提示システムを使う利用者 9 を含めて撮影する。この撮影装置 2a、2b は、Web カメラや HD (High-Definition) 解像度の撮影カメラなどの汎用の撮像デバイスを利用することが可能で、レンズ、露光調整機構、CCD (Charge Coupled Device) など基本的な撮像デバイス構成要素を有している。撮影装置 2a、2b は、それぞれ外部入力端子 (I/F) 200、201 によって映像・音声処理装置 100 に接続されている。撮影装置 2a、2b では、撮影対象を含む視野からの光が、レンズ・露光調整機構を通して CCD 上に結像される。CCD は入力された受光量に応じた信号を生成し出力する。

[0030] CCD で生成された信号は、カメラ用の所定の信号処理が施された後に、画像信号（映像信号）として映像処理装置（送信側）102 の内部に取り込まれる。撮影装置 2a、2b は、動画像を扱うため一定間隔で、お互いに同期して映像を取得し、逐次映像信号を映像処理装置（送信側）102 に供給する。

[0031] 映像モデル生成部 202 は、物理的に位置の異なる撮影装置 2a と撮影装置 2b で取得した映像を、外部入力端子 (I/F) 200 と外部入力端子 (

1/F) 201 を介して入力し、ある時刻において取り出したフレーム（画像）信号（2a、2bの画像）を用いて、視差画像信号を生成する。視差画像信号は、撮影装置2aのある時刻のフレーム（画像信号： $I_{base}(x, y)$ ）を基準画像とし、これと同一時刻の撮影装置2bのフレーム（画像信号： $I_{ref}(x, y)$ ）との間で、ブロックマッチング法によって画素の対応点を求めることによって生成できる。この場合、撮影装置2bの画像信号を、参照画像と呼ぶ。

[0032] （視差画像生成）

映像モデル生成部202は、ブロックマッチングを行う際に、基準画像を所定の大きさのブロックに分割を行う。分割の大きさは、例えば16画素×16画素である。当該分割されたブロックと参照画像の間の指標値を領域毎に算出し、算出した指標値が最小となる領域を探し出す（対応点）。映像モデル生成部202は、基準画像の分割ブロックを画素単位でずらしていき、すべての画素において対応点を算出する。

[0033] ブロックマッチングで用いる指標値は、画像信号間の相関性や類似性を示すものであればよい。映像モデル生成部202は、例えば、分割されたブロックに含まれる画素の輝度値と参照画像のある領域における輝度値の差の絶対値総和（SAD；Sum of Absolute Difference）を用いる。基準画像信号から分割されたブロック（例えば、大きさが $N \times N$ 画素）と参照画像信号のブロックとの間のSADは次の式で表される。

[0034] [数1]

$$SAD(p, i_0, j_0) = \sum_{i=-N/2}^{N/2} \sum_{j=-N/2}^{N/2} |I_{base}(i_0+i, j_0+j) - I_{ref}(i_0+i+p, j_0+j)| \quad \cdots (1)$$

[0035] 式（1）において、 $I_{base}(i_0+i, j_0+j)$ は基準画像の座標（ i_0+i, j_0+j ）における輝度値、（ i_0, j_0 ）は視差算出する処理対象画素座標を示す。 $I_{ref}(i_0+i+p, j_0+j)$ は参照画像の座標（ i_0+i+

$p, j_0 + j$) における輝度値で、 p は当該分割されたブロックの中心座標 (i_0, j_0) を基準とした x 軸方向 (水平方向) のシフト量 (視差) である。撮影装置 2 a と撮影装置 2 b を水平になるように設置した場合、対応する画素は水平な直線上であり、 y 軸方向 (垂直方向) のパラメータは削除できる。

映像モデル生成部 202 は、ブロックマッチングにおいて、パラメータ p 毎に $SAD(p)$ を算出し、 $SAD(p)$ を最小とする p を探し出す。 p は基準画像から当該分割されたブロックと当該参照領域の位置までの視差を表す。

[0036] [数2]

$$V_{diff}(x, y) = \arg \min_p SAD(p, x, y) \quad \dots (2)$$

[0037] 式 (2) において、右辺の
[数A]

$$\arg \min_p \{ \bullet \}$$

は、評価式 ($\{ \}$ 内の数式) を最小にする p (視差) 値を抽出する関数である。以上により、映像モデル生成部 202 は、視差画像を画素単位で求めることができる。

[0038] (デプスマップ生成)

映像モデル生成部 202 は、上記算出した視差画像を、式 (3) を用いて実際の距離を示すデプスマップに変換する。

[0039]

[数3]

$$z = \frac{f \cdot b}{d + \Delta d} \quad \dots (3)$$

[0040] ここで、 z はデプス値、 d は視差、 f と b はそれぞれカメラの焦点距離とカメラ間のベースライン距離、 Δd はカメラオフセットを示している。

[0041] (カメラ座標における3次元モデル生成)

次に、映像モデル生成部202は、得られたデプスマップから、カメラ座標における奥行き情報(3次元モデル)を生成する。カメラ座標とは、撮影したカメラ固有の座標系である。

その前に、式(3)で得られたデプスマップを、画像座標系 X_{img} から前記カメラ座標系 X_{cam} に変換する必要がある。画像座標系とは、画像の画素を扱うための座標系である。その変換式は、以下の通りである。

[0042] [数4]

$$X_{cam} = A^{-1} \tilde{X}_{img} \quad \dots (4)$$

[0043] 式(4)において、

[数B]

$$\tilde{X}_{img}$$

は画像座標系 X_{img} の斉次座標を表しており、2次元の画像座標(i, j)に1次元加えた3次元の座標($i, j, 1$)である。斉次座標とは、一般的に、空間の次元を1次上げ、これらの比によって空間を表現する座標の取り方である。

A は、前記基準カメラの撮像面に投影された前記カメラ座標系の平面画像

を前記画像座標系に変換するマトリクスであり、カメラの内部行列と呼ばれている。 A^{-1} は、その逆行列を示している。内部行列は、前記撮影装置のカメラ特性を示すパラメータ（内部パラメータ）を含むマトリクスで、具体的には以下のとおりである。

[0044] [数5]

$$A = \begin{pmatrix} fk_u & fk_v & u_0 \\ 0 & fk_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad \cdots (5)$$

[0045] ここで、 f は焦点距離、 (u_0, v_0) は画像中心の座標、 k_u および k_v は u 軸（ x 軸）、 v 軸（ y 軸）方向のスケールファクタで、 k_s は直角性を保たないような歪みを表すパラメータである。

[0046] 図4を用いて、3次元空間の奥行き情報を復元する方法について説明する。デプスマップ上のある座標 (i, j) が、式(4)によってカメラ座標系の画像平面 $(z = f)$ 上の1点 $m(x, y, f)$ に変換されたとし、座標 (i, j) におけるデプス値が z である場合、投影元の3次元空間内の点 m_c の座標は、図4に示す x_c 軸と z_c 軸で表せる平面上にある三角形と、 y_c 軸と z_c 軸で表せる平面上にある三角形の相似の関係から以下ようになる。

[0047] [数6]

$$m_c(z \frac{x}{f}, z \frac{y}{f}, z) \quad \cdots (6)$$

[0048] (ワールド座標における3次元モデル生成)

図3に戻り、映像モデル生成部202は、続いて座標系の変換を行う。具体的には、上記撮影装置で撮影されたカメラの座標系を、本システムで扱う共通の座標系（以下、ワールド座標系 X_{world} ）に変換する必要がある。ワー

ルド座標系は、図 1 の表示装置の中心を原点に、ディスプレイ面に向かって右の方向に x 軸の正の方向、ディスプレイ面の上部の方向に y 軸の正の方向、ディスプレイ面から奥の方向に z 軸の正の方向とする。

座標系の変換は、一般的に以下の式によって行うことができる。

[0049] [数7]

$$X_{to} = RX_{from} + T \quad \dots (7)$$

[0050] ここで、 X_{from} は変換前の座標系、 X_{to} は変換後の座標系を表し、R と T は変換前の座標系から変換後の座標系への回転と併進を示すマトリクスである。回転と併進のマトリクスを具体的に示すと、以下のとおりである。

[0051] [数8]

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_x & -\sin \theta_x \\ 0 & \sin \theta_x & \cos \theta_x \end{pmatrix} \begin{pmatrix} \cos \theta_y & 0 & \sin \theta_y \\ 0 & 1 & 0 \\ -\sin \theta_y & 0 & \cos \theta_y \end{pmatrix} \begin{pmatrix} \cos \theta_z & -\sin \theta_z & 0 \\ \sin \theta_z & \cos \theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \dots (8)$$

[0052] ここで θ_x 、 θ_y 、 θ_z は、それぞれ X 軸周り、Y 軸周り、Z 軸周りの回転を示す。

[0053] [数9]

$$T = \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix} \quad \dots (9)$$

[0054] ここで T_x 、 T_y 、 T_z は、それぞれ X 軸方向、Y 軸方向、Z 軸方向の併進を示す。

上記、カメラの特性や位置を表す行列 A、行列 R や行列 T は、カメラ校正によって一般的に求めることができる。

[0055] 以上により、映像モデル生成部 202 は、上記ワールド座標系における 3

次元空間上の位置を取得でき、映像モデルを生成することができる。映像モデルとは、３次元的な奥行き情報（位置情報）と、それに貼り付ける映像（テクスチャ）情報で構成されるものである。テクスチャを奥行き情報に貼り付ける処理は、テクスチャマッピングとも呼ばれる。

[0056] 図３に戻り、電子データを取得する方法について説明する。

共同作業で用いる電子データは、映像・音声処理装置１００に接続された、外部記憶装置／外部サーバ２０３に格納されており、外部入力端子（Ｉ／Ｆ）２０４を通じて、電子データモデル生成部２０５に入力される。外部記憶装置／外部サーバは、ネットワークに接続されており、遠隔地に設置された映像提示システムからも参照することが可能である。

[0057] 電子データモデル生成部２０５は、入力した電子データの基準情報（位置、大きさ、向き）を基に電子データモデルを生成し、映像・データ合成部２０６に出力する。電子データは、コンピュータで扱うことができる、テキストによる文字情報、あるいは画像（映像）や図形などのビジュアル情報によって構成されたオフィスドキュメントである。一般的に、電子データモデルは、矩形状の面で形成される。具体的には、電子データモデルは、３次元の位置情報を有した最小単位の三角形（ポリゴン）を２枚組み合わせ（三角形の長辺を接合し）、上記オフィスドキュメントの画像（映像）を貼り付けたモデルである。

なお、電子データの位置や角度は、後述する電子データ制御方法によって、システム稼働中に可変しても良い。なお、基準設定については、後述する作業スタイルと利用者位置毎に準備し、作業内容に応じて切り替えることもできる。

[0058] 映像・データ合成部２０６及び視点生成部２０７は、前記ワールド座標系でモデルデータを処理するための３次元モデル処理機能を有しており、具体的にはＣＧなどの３次元モデルを扱うＧＰＵ（Graphics Processing Unit）によってその機能を実現することができる。

[0059] （３次元位置に基づいたレンダリング方法）

映像・データ合成部206は、GPUで実現される一般的な3DCGレンダリング技術によって、映像モデル生成部202から入力する映像モデルと電子データモデル生成部205から入力する電子データモデルを共通の座標系（ワールド座標系）内に配置させ、お互いを合成させる。

[0060] 具体的には、供給される3次元位置を持った複数の平面、あるいは最小単位である三角形（ポリゴン）を複数組み合わせることで表現する立体的な面を仮想的な3次元空間内（ワールド座標系）に配置することでシーンが形成され、さらに、それらの面に直接撮影した映像（電子データの絵）を貼り付けたり、あるいは撮影位置からそれらの面に向けて映像（電子データの絵）を投影（投影テクスチャリング）することで、3次元的な奥行き構造を持ったモデルに映像を貼り付けることができる。各モデル間の前後関係の処理（奥のオブジェクトが、手前のオブジェクトによって隠れる領域については表示させないための処理）は、GPUの機能によって実現される。

[0061] 視点生成部207は、映像・データ合成部206から映像と電子データを合成したモデルを入力し、GPUで実現される一般的な3DCGレンダリング技術によって、上記算出したワールド座標系において想定する利用者位置（詳細は後述する）に仮想的なカメラ（仮想視点）を設置し、そのカメラによって3次元空間に再現されたシーンを撮影する。

具体的には、視点生成部207は、ビュー変換処理によって想定する観察者位置から見た画像を切り出すことができる。ビュー変換処理は、ワールド座標系から観察者位置に置かれた仮想カメラを基準とする座標に変換することである。仮想カメラのX軸を示すベクトルを $R = (r_x, u_x, v_x)$ 、仮想撮像装置のY軸を示すベクトルを $U = (r_y, u_y, v_y)$ 、仮想撮像装置のZ軸を示すベクトルを $V = (r_z, u_z, v_z)$ 、仮想撮像装置の位置を示すベクトルを $P = (x, y, z)$ としたときに、ビュー変換行列は式(10)のようになる。

[0062]

[数10]

$$\begin{pmatrix} r_x & u_x & v_x & 0 \\ r_y & u_y & v_y & 0 \\ r_z & u_z & v_z & 0 \\ -(R \cdot p) & -(U \cdot p) & -(V \cdot p) & 1 \end{pmatrix} \quad \dots (10)$$

[0063] 続いて、ビュー変換されたシーンに対して遠くのもの小さく、近くのもの大きく見えるように透視投影処理が実施される。透視投影処理では、撮像装置に近いところ（前方投影面）と、遠いところ（後方投影面）に対して2つの投影面が導入され、撮像装置原点を中心とした錘状になるように設定される。プロジェクション変換によって、近いところの面と遠いところの面の間の空間が立方体になるように変換することによって、近いところが大きく、遠いところが小さく見えるような変換がなされる。このプロジェクション変換によって、射影座標系に変換される。水平方向の視野角を w 、垂直方向の視野角を h とし、仮想撮像装置の位置から前方投影面までの距離を z_n 、後方投影面までの距離を z_f とすると、以下の行列によって射影座標系に変換することができる。

[0064] [数11]

$$\begin{pmatrix} w & 0 & 0 & 0 \\ 0 & h & 0 & 0 \\ 0 & 0 & z_f / (z_f - z_n) & 1 \\ 0 & 0 & -z_n \times z_f / (z_f - z_n) & 0 \end{pmatrix} \quad \dots (11)$$

[0065] さらに、表示装置に表示するために、射影座標系からスクリーン座標系への変換が行なわれる。スクリーン座標系では、表示装置の左上を原点に、またY軸の向きが逆になるような変換が行なわれる。表示装置の水平方向の画素数を $Width$ 、垂直方向の画素数を $Height$ とすると変換行列は、以下のようになる。

[0066]

[数12]

$$\begin{pmatrix} Width/2 & 0 & 0 & 0 \\ 0 & -Height/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ Width/2 & Height/2 & 0 & 1 \end{pmatrix} \quad \dots (12)$$

[0067] 以上3DCGレンダリング技術により利用者の位置から見た画像が生成される。

[0068] 続いて、本発明に関わる、遠隔地に居る利用者が隣にいるような位置関係で、共通の作業オブジェクトを見ながらコミュニケーションを行うための視点制御方法について、図5～図9を用いて説明する。

図5は、本システムを一般的な対面のテレビ会議システムのように利用した場合に、システム全体を上部から見下ろした図（上面図）である。図の500及び501は、図1に示した1～5の構成を含む映像提示システムである。500は、図1で示した利用者9側の映像提示システムで、501は遠隔地に設置してある映像提示システムである。利用者は9と7で、7は遠隔地に居る利用者である。

[0069] システム間は、ネットワーク6等を介して接続され、お互いが通信できるようになっている。504及び505は、それぞれ利用者9の側のシステムのワールド座標系 X_{world_A} と利用者7の側のシステムのワールド座標系 X_{world_B} を示している。ワールド座標系 X_{world_A} は、ディスプレイ面に向かって右の方向にx軸の正の方向、ディスプレイ面の奥に向かってz軸の正の方向となり、また図示していないy軸の正の方向は上方向に向かっているものとする。

[0070] 同様に、ワールド座標系 X_{world_B} は、利用者7がディスプレイに向かっている方向を基準として、ディスプレイの右方向（図面では左方向）をx軸の正の方向、ディスプレイ面の奥の方向（図面では下方向）をz軸の正の方向とし、また図示していないy軸の正の方向は上方向に向かっているものとする。

[0071] 続いて、本発明に関わる作業の形態（利用者と電子データの位置関係）を図示したものが図6である。利用者9と利用者7が隣に居て、お互いに作業物600（電子データを明示的に示したもの）を見ることができる位置関係を模式的に示している。

図7は、図6の位置関係を実現した場合に、表示装置に映る遠隔地の映像を模式的に示した図である。図7（A）は利用者9が見る映像で、図7（B）は利用者7が見る映像である。

[0072] 図6に示したように、利用者9にとって、作業物600は左側で利用者7は右側に位置する。従って、利用者9が見る遠隔地の映像506は、作業物600は左に見え、利用者7の映像700は右側に見える。一方、利用者7にとって、作業物600は右側で、利用者9は左側に位置する。従って、利用者7が見る遠隔地の映像507は、作業物600は右に見え、利用者9の映像701は左側に見える。

以上、図7に示した映像を提示することによって、利用者は図6に示したような位置関係で、遠隔地間で共通の電子データと相手の映像を見ながらコミュニケーションを行うことができる。

[0073] 具体的に、上記図7に示したような視点の映像を生成する方法について、図8を用いて説明する。初めに、利用者の位置関係を明確にしておく必要がある。これは、利用者は2地点に居て、図6あるいは図7に示したように、利用者9であるか、利用者7であるかによって、提示される映像内の相手の映像と電子データの表示位置が異なるためである。

[0074] 利用者の位置関係を確定させる方法として、例えば、予め双方の位置（図6におけるどちらの利用者であるか）を決めておく方法がある。利用者は、映像・音声処理装置100に接続された操作装置5を使って、図9に示すユーザ表示に従い、お互いの位置を設定することができる。設定結果は、映像・音声処理装置100の内部にあるCPU208に保持される。CPU208は、図示していない通信手段を用いて、遠隔地にある映像提示システムと連携して、図9に示す設定値が排他的（一方が選択した項目は、相手側では

選択できない) になるように制御しても良い。以上の方法により、双方の位置が明確になる。

[0075] 続いて、映像の取得方法について説明する。図8において、利用者7の側のシステム501に表示する遠隔地の映像は、利用者9の側の映像である。その映像の取得方法は、利用者9及びシステム500の側に、図6に示すような位置関係になるように利用者7が存在していると仮定して、その利用者7の位置にカメラを設置して取得する。つまり、実際には存在していない遠隔地の利用者が仮に隣に居たとして、そこから撮影した映像を相手側に伝送することで実現できる。そのためには、お互いの利用者位置をもう一方のシステムにおける仮想的な位置に変換する処理、すなわち、一方のシステムの座標系をもう一方のシステムの座標系に変換することが必要である。この変換によって、もう一方のシステムの座標系における位置を同じ座標系として扱うことが可能で、例えば、利用者9の側のシステムにおける、利用者7の仮想的な位置、すなわち、利用者7の側のシステムに送る映像を取得するカメラ設置位置を算出することが可能になる。しかしながら、システム500とシステム501は物理的に接合しておらず、双方の座標系を関連付けることができない(座標系を変換する基準となる物理的な位置関係が不明確である)。本発明では、共通に利用する電子データ600を座標系変換の基準とすることで、双方の座標系を関連付けて、座標系変換を実現する。利用者9の側のシステム500に表示する遠隔地の映像の取得方法も同様である。

[0076] CPU208は、映像提示システムに設定されたワールド座標系(図8の500のシステムの場合 X_{world_A} (504)、501のシステムの場合 X_{world_B} (505))内に配置された電子データの基本情報のうち、各ワールド座標系における位置(500のシステムの場合805、501のシステムの場合806)とその方向ベクトル(500のシステムの場合807、501のシステムの場合808)を用いて、以下の方法によって遠隔地システムのための映像を取得するカメラ設置位置を決める。

[0077] CPU208は、遠隔地のシステムのワールド座標系 X_{world_B} を、電子

データの座標系を経由して、利用するシステムの座標系 X_{world_A} に変換する。ここで、システム 500 の座標系 X_{world_A} における電子データ中心の 3 次元位置 805 を P_{E_A} 、方向ベクトル 807 を $\rightarrow H_{E\_A}$ とし、システム 501 の座標系 X_{world_B} における電子データ中心の 3 次元位置 806 を P_{E_B} 、方向ベクトル 808 を $\rightarrow H_{E\_B}$ とする。CPU 208 は、式 (7) を用いて、システム 501 の座標系 X_{world_B} を、一旦システム 501 における電子データの座標系に変換する。式 (7) に適用する、回転マトリクス $R_{W\_BE\_B}$ と併進マトリクス $T_{W\_BE\_B}$ は以下の様に求める。

[0078] 回転マトリクス $R_{W\_BE\_B}$ は、システム 501 の座標系 X_{world_B} から電子データ 600 の座標系への回転であり、システム 501 の表示装置の向いている方向から電子データ 600 の向いている方向への回転と同じである。ここで、システム 501 の表示装置の向き 803 を $\rightarrow H_{D\_B}$ (座標系 X_{world_B} を用いると $\rightarrow H_{D\_B} = (0, 0, -1)$ である)、電子データ 600 の向きを $\rightarrow H_{E\_B}$ とすると、 $\rightarrow H_{D\_B}$ と $\rightarrow H_{E\_B}$ 間において、内積の公式を展開して X 軸周り、Y 軸周り、Z 軸周りの回転角度 ($\theta_{x\_DE\_B}$, $\theta_{y\_DE\_B}$, $\theta_{z\_DE\_B}$) を求める。以上の回転角度を、式 (8) に代入して回転マトリクス $R_{W\_BE\_B}$ を算出する。

併進マトリクス $T_{W\_BE\_B}$ は、システム 501 の座標系 X_{world_B} の原点 O_{W_B} から電子データ 600 の中心位置 P_{E_B} に向かうベクトルから構成され、以下の式のとおりである。

[0079] [数13]

$$T_{W\_BE\_B} = P_{E\_B} - O_{W\_B} \quad \dots (13)$$

[0080] 続いて、CPU 208 は、システム 500 において、上記方法と同様に電子データの座標系からワールド座標系 X_{world_A} に変換するための回転マトリクス $R_{E\_AW\_A}$ と併進マトリクス $T_{E\_AW\_A}$ を求める。

CPU 208は、以下の式を用いて、遠隔地のシステム501の座標系を利用するシステム500の座標系に変換する。

[0081] [数14]

$$X_{world\_A} = R_{E\_AW\_A}(R_{W\_BE\_B}X_{world\_B} + T_{W\_BE\_B}) + T_{E\_AW\_A} \quad \cdots(14)$$

[0082] この変換式によって、CPU 208は、遠隔地のシステム501におけるワールド座標系 X_{world_B} 内の座標を、利用するシステム500の座標系 X_{world_A} における座標に変換することができる。

CPU 208は、遠隔地の利用者7の位置を式(14)に従って変換し、変換された座標位置に遠隔地システムのための映像を取得するカメラを設置する。CPU 208は、そのカメラによって取得した映像を、遠隔地のシステム501に伝送する。

利用者9及び利用者7の位置は、適正視距離、例えば3H（ディスプレイ面の垂直の長さの3倍）によって、設定可能である。あるいは、例えば利用者の位置をトラッキングする装置を装着して、利用者位置をリアルタイムに算出しても良い。

[0083] 以上により、利用者が3次元的な位置に関する設定を行うことなく、簡易な手法で、遠隔地の人が隣に居るような位置関係で共通の電子データを見ることができる視点の映像を取得することができる。双方にこの映像を伝送し合うことで、上記位置関係でコミュニケーションをすることができる。なお、上記視点の制御は、システム開始時に一度設定するだけにしても良いし、随時設定を変えるような仕組みにしても良い。なお、仮想カメラの撮影画角については、上記システムが既知、つまり表示装置の大きさが限定できる場合は、設定を変える必要はない。また、システムごとに表示装置の大きさが変わる場合には、相手の表示装置に応じて撮影画角を変えるような仕組みを設けても良い。

[0084] <フローチャート>

次に、本実施形態に係る映像・音声処理装置（映像処理）１００が行う処理について説明する。図１０～図１４は、本実施形態に係る映像・音声処理装置（映像処理）１００が行う処理を示すフローチャートである。なお、以下のフローチャート及び、その説明において、音声信号の処理は必要なところのみ記述する。

映像・音声処理装置（映像処理）１００は、図示していないシステム全体を制御する制御部によって、システム処理を開始すると、ステップＳ１００に進む。

[0085] （ステップＳ１００）映像・音声処理装置（映像処理）１００は、前記制御部によって、システム全体の稼働を監視するための処理（図１０の左側の処理、Ｓ１０１～Ｓ１０４）と通信を行うための処理（図１０の右側の処理、Ｓ１０５～Ｓ１０９）を同時に開始させる。双方の処理を開始させると、ステップＳ１０１とＳ１０５に進む。

[0086] 始めに、システム全体の稼働を監視するための処理について説明する。

（ステップＳ１０１）映像・音声処理装置（映像処理）１００は、前記制御部によって、システムが終了したか否かを常に監視し、終了していない場合はステップＳ１０１の処理を繰り返す。終了した場合には、ステップＳ１０２に進む。

（ステップＳ１０２）映像・音声処理装置（映像処理）１００は、前記制御部によって、前記通信を行うための処理において実行するスレッドの状態を監視し、スレッドが実行中である場合はステップＳ１０３に進み、スレッドが全て完了している場合はステップＳ１０４に進む。

[0087] （ステップＳ１０３）映像・音声処理装置（映像処理）１００は、前記制御部によって、前記通信を行うための処理において実行しているスレッドを強制的に終了させ、ステップＳ１０４に進む。

（ステップＳ１０４）映像・音声処理装置（映像処理）１００は、前記制御部によって、全ての処理が停止していることを確認して、システムを終了さ

せる。

続いて、通信を行うための処理について説明する。

[0088] (ステップS 1 0 5) 映像・音声処理装置(映像処理) 1 0 0は、前記制御部によって、常に遠隔地との間で通信を開始できる状態にあるかを監視し、通信を開始できる状態であればステップS 1 0 6に進み、通信を開始できる状態でない場合はステップS 1 0 5の処理を繰り返す。

(ステップS 1 0 6) 映像・音声処理装置(映像処理) 1 0 0は、前記制御部によって、後述するデータ送信スレッドを開始させ、ステップS 1 0 7に進む。

(ステップS 1 0 7) 映像・音声処理装置(映像処理) 1 0 0は、前記制御部によって、後述するデータ受信スレッドを開始させ、ステップS 1 0 8に進む。

[0089] (ステップS 1 0 8) 映像・音声処理装置(映像処理) 1 0 0は、前記制御部によって、後述する電子データ制御スレッドを開始させ、ステップS 1 0 9に進む。

(ステップS 1 0 9) 映像・音声処理装置(映像処理) 1 0 0は、前記制御部によって、上記開始させたスレッドの状態を常に監視し、全てのスレッドの処理が完了していることを検知した場合はステップS 1 0 5に戻り上記処理を繰り返す。一方監視しているスレッドが全て終了していない場合は、ステップS 1 0 9に戻り処理を繰り返す。

[0090] 次に、データ送信スレッドについて、図 1 1 を用いて説明する。データ送信スレッドは、上記システム処理によって起動される。

(ステップS 2 0 0) 映像処理装置(送信側) 1 0 2は、後述する送信フレーム情報生成処理によって、送信するフレーム情報を生成し、ステップS 2 0 1に進む。

(ステップS 2 0 1) 映像処理装置(送信側) 1 0 2の内部にある送信制御部 2 0 9は、生成された送信フレーム情報を、よりデータ量の少ない符号に圧縮する方法(例えば、国際標準動画圧縮規格H. 2 6 4等)を用いて符号化

(エンコード) し、ステップ S 2 0 2 に進む。

[0091] (ステップ S 2 0 2) 映像処理装置 (送信側) 1 0 2 の内部にある送信制御部 2 0 9 は、図 2 の音声処理装置 (送信側) 1 0 3 と連携して、符号化された映像信号と音声信号をパッキング (p a c k i n g ; 詰込) し、ステップ S 2 0 3 に進む。

(ステップ S 2 0 3) 映像・音声処理装置 1 0 0 は、生成された符号データを、ネットワークを介して、遠隔地の映像提示装置に向けて伝送する。その後、ステップ S 2 0 4 に進む。

(ステップ S 2 0 4) 映像・音声処理装置 1 0 0 は、前記制御部によって、通信を終了させるか否かを常に監視し、終了させる場合はステップ S 2 0 5 に進み、終了させない場合はステップ S 2 0 0 に戻り上記処理を繰り返す。

(ステップ S 2 0 5) 映像・音声処理装置 1 0 0 は、データ送信スレッドを終了させる。

[0092] 次に、ステップ S 2 0 0 で実施する送信フレーム情報生成処理について、図 1 2 を用いて説明する。

(ステップ S 3 0 0) 映像処理装置 (送信側) 1 0 2 は、外部入力端子 (I / F) 2 0 0 及び 2 0 1 を介して、撮影装置 2 a と 2 b で撮影した映像 (フレーム) を取り込む。その後、ステップ S 3 0 1 に進む。

(ステップ S 3 0 1) 映像モデル生成部 2 0 2 は、前記入力されたマルチ画像を基に、映像モデルを生成する。その後、ステップ S 3 0 2 に進む。

(ステップ S 3 0 2) 電子データモデル生成部 2 0 5 は、後述する電子データ制御スレッドが生成する最新の電子データ属性情報に基づいて電子データのモデル情報を生成する。その後、ステップ S 3 0 3 に進む。

[0093] (ステップ S 3 0 3) 映像・データ合成部 2 0 6 は、映像モデル生成部 2 0 2 から入力した映像モデルと電子データモデル生成部 2 0 5 から入力した電子データモデルをワールド座標系内で合成し、ステップ S 3 0 4 に進む。

(ステップ S 3 0 4) 視点生成部 2 0 7 は、映像・データ合成部 2 0 6 から入力した合成されたモデル情報に基づいて、本発明に関わる前記視点の画像

(フレーム)を生成し、ステップS305に進む。

(ステップS305)視点生成部207は、生成したフレームの信号を、送信制御部209に出力して、送信フレーム情報生成処理を終了する。

[0094] 次に、データ受信スレッドについて、図13を用いて説明する。データ受信スレッドは、上記システム処理によって起動される。

(ステップS400)映像処理装置105の内部にある受信制御部212は、遠隔地にある映像提示システムから送信された符号化データを、ネットワークを介して受信する。その後、ステップS401に進む。

(ステップS401)映像処理装置104の内部にある受信制御部212は、音声処理装置106の内部にある図示していない受信制御部と連携して、受信した符号化データを映像符号化データと音声符号化データに分割する。その後、ステップS402に進む。

[0095] (ステップS402)映像処理装置104の内部にある受信制御部212は、映像符号化データに対して、前記圧縮とは逆の処理を行い、データ量の多い元の信号を復元する。その後、ステップS403に進む。

(ステップS403)受信制御部212は、復元した画像信号を、表示メモリ213に書き込む。表示装置1は、表示メモリ213に書き込まれた画像信号を、定期的に読み出し、画面に表示する。その後、ステップS404に進む。

[0096] (ステップS404)映像・音声処理装置100は、前記制御部によって、通信を終了させるか否かを常に監視し、終了させる場合はステップS405に進み、終了させない場合はステップS400に戻り上記処理を繰り返す。

(ステップS405)映像・音声処理装置100は、データ受信スレッドを終了させる。

[0097] 次に、電子データ制御スレッドについて、図14を用いて説明する。電子データ制御スレッドは、上記システム処理によって起動される。

(ステップS500)映像処理装置(送信側)102内のCPU208は、操作装置5を用いてユーザが設定する電子データの属性(後述)変更を常に

監視し、属性が変更された場合はステップS501に進む。属性が変更されない場合は、ステップS500の処理を繰り返す。

[0098] (ステップS501) CPU208は、電子データの属性情報を更新し、電子データモデル生成部205に最新の属性情報を出力し、前記電子データモデル生成処理にその内容を反映させる。その後、ステップS502に進む。

(ステップS502) CPU208は、更新された電子データ属性情報を図示しないネットワークを介した別の手段によって、遠隔地の映像提示装置にその属性情報を伝送する。その後、ステップS503に進む。

[0099] (ステップS503) 映像・音声処理装置100は、前記制御部によって、通信を終了させるか否かを常に監視し、終了させる場合はステップS504に進み、終了させない場合はステップS500に戻り上記処理を繰り返す。

(ステップS504) 映像・音声処理装置100は、電子データ制御スレッドを終了させる。

[0100] <送信するデータ>

図14のステップS501において生成する電子データ属性情報について、図15を用いて説明する。電子データ属性情報は、例えば電子データの3次元位置と面の向きと大きさを示す情報で、電子データを上記ワールド座標系内の3次元位置に配置させることが可能な情報である。

図12の送信フレーム情報生成処理で生成する送信フレーム情報について、図16を用いて説明する。送信フレーム情報は、例えば画素数(横方向): widthと画素数(縦方向): heightと画素情報から構成された、1枚分の画像を表す情報である。

[0101] (実施形態2)

<ソフトウェア>

上述した実施形態における映像・音声処理装置100の一部、例えば、映像モデル生成部202、電子データモデル生成部205、映像・データ合成部206、視点生成部207、送信制御部209、並びに受信制御部212をコンピュータで実現するようにしても良い。その場合、この制御機能を実

現するためのプログラムをコンピュータ読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、実行することによって実現しても良い。なお、ここでいう「コンピュータシステム」とは、映像・音声処理装置に内蔵されたコンピュータシステム及びDSP（デジタル・シグナル・プロセッサ）であって、OSや周辺機器等のハードウェアを含むものとする。

[0102] また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM、CD-ROM等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。さらに「コンピュータ読み取り可能な記録媒体」とは、インターネット等のネットワークや電話回線等の通信回線を介してプログラムを送信する場合の通信線のように、短時間、動的にプログラムを保持するもの、その場合のサーバやクライアントとなるコンピュータシステム内部の揮発性メモリのように、一定時間プログラムを保持しているものも含んでも良い。また上記プログラムは、前述した機能の一部を実現するためのものであっても良く、さらに前述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるものであっても良い。

[0103] また、上述した実施形態における映像・音声処理装置100の一部、または全部を、LSI（Large Scale Integration）等の集積回路として実現しても良い。映像・音声処理装置100の各機能ブロックは個別にプロセッサ化してもよいし、一部、または全部を集積してプロセッサ化しても良い。また、集積回路化の手法はLSIに限らず専用回路、または汎用プロセッサで実現しても良い。また、半導体技術の進歩によりLSIに代替する集積回路化の技術が出現した場合、当該技術による集積回路を用いても良い。

符号の説明

[0104] 1…表示装置、2…マルチメディア情報取得装置、2a…撮影装置、2b…撮影装置、3…音声再生装置、4…計算機、5…操作装置、6…ネットワーク、7…利用者、8…電子データ、9…利用者、10…マルチメディア情報

提示装置、１１…マルチメディア情報取得装置、１２…マルチメディア情報提示装置、１００…映像・音声処理装置、１０１…送信装置、１０２…映像処理装置、１０３…音声処理装置、１０４…受信装置、１０６…音声処理装置、１０７…映像・音声処理装置、１０８…送信装置、１０９…受信装置、２０２…映像モデル生成部、２０３…外部サーバ、２０５…電子データモデル生成部、２０６…映像・データ合成部、２０７…視点生成部、２０８…ＣＰＵ、２０９…送信制御部、２１２…受信制御部、２１３…表示メモリ、５００…システム、５０１…システム、５０６…映像、５０７…映像、６００…作業物、７００…映像、７０１…映像。

請求の範囲

[請求項1] 撮影装置によって取得された撮影映像を入力し、該入力した撮影映像に、画像として生成された電子データを加工して合成し、該合成した映像を送受信することにより、遠隔地間でコミュニケーションを行なうことを可能とした映像提示システムであって、

 撮影する範囲の奥行き情報と、該奥行き情報を基に該奥行き情報に貼り付けた映像情報とからなる映像モデルを生成する映像モデル生成部と、

 電子データを入力し、該入力した電子データの位置、大きさ、向きの情報を含む基準情報に基づいて電子データモデルを生成する電子データモデル生成部と、

 共通の座標系を設けて、前記映像モデルと前記電子データモデルを共通の座標系内で合成する映像・データ合成部と、

 該映像・データ合成部で合成されたモデルを入力し、前記共通の座標空間に仮想的な視点を設置して、該視点の映像を生成する視点生成部と、

 該視点生成部で取得した映像を送信する送信部と、を有することを特徴とする映像提示システム。

[請求項2] 前記視点生成部は、前記電子データの3次元的位置と前記共通の座標系における利用者の相対的な位置関係に基づいて算出された視点位置の映像を生成することを特徴とする請求項1に記載の映像提示システム。

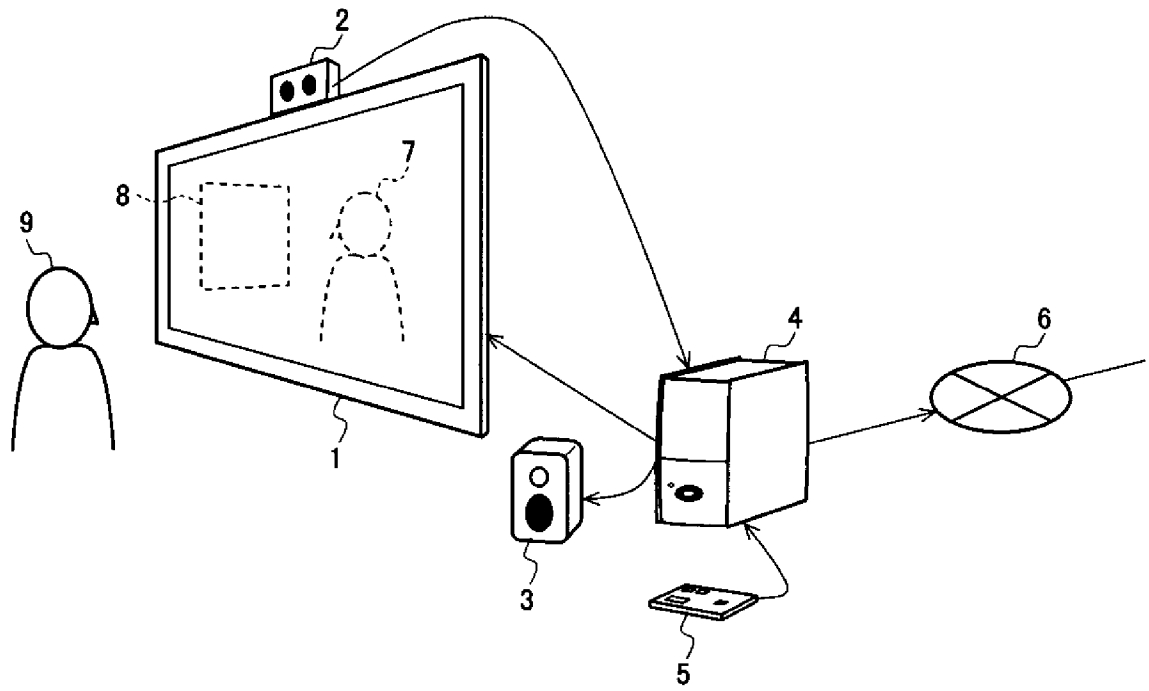
[請求項3] 利用者による操作入力を受け付ける操作装置と、

 前記電子データの位置及び利用者の視点位置を設定するための表示画面を表示装置に表示させる制御手段とを有し、前記視点生成部は、前記表示画面に対する前記操作装置の操作入力に従って、前記電子データの3次元的位置及び利用者の視点位置を決定することを特徴とする請求項2に記載の映像提示システム。

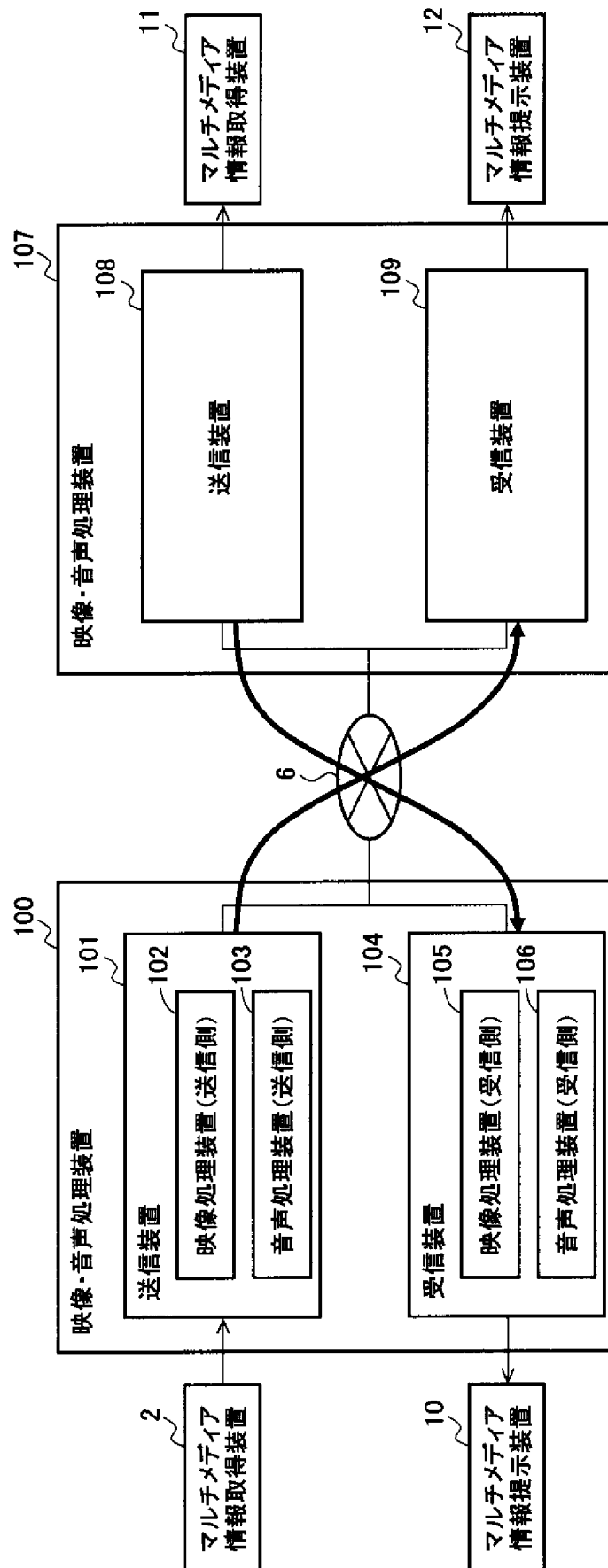
- [請求項4] 利用者の3次元上の位置を計測する3次元位置計測装置を有し、
 前記視点生成部は、前記3次元位置計測装置による計測結果に従って、利用者の位置をリアルタイムに変更することを特徴とする請求項1～3のいずれか1に記載の映像提示システム。
- [請求項5] 前記撮影装置を複数備え、
 前記映像モデル生成部は、複数の前記撮影装置で撮影された映像から視差画像を生成し、該生成した視差画像に基づいて撮影された対象のデプスを示すデプスマップを生成し、該デプスマップから前記撮影装置の固有の座標系における奥行き情報を生成し、該撮影装置の固有の座標系を前記映像提示システムで扱う共通座標系に変換して、奥行き情報と該奥行き情報に貼り付けた映像情報とからなる前記映像モデルを生成することを特徴とする、請求項1～4のいずれか1に記載の映像提示システム。
- [請求項6] 撮影装置によって取得された撮影映像を入力し、該入力した撮影映像に、画像として生成された電子データを加工して合成し、該合成した映像を送受信することにより、遠隔地間でコミュニケーションを行なうことを可能とした映像提示方法であって、
 撮影する範囲の奥行き情報と、該奥行き情報を基に該奥行き情報に貼り付けた映像情報とからなる映像モデルを生成する第1の過程と、
 電子データを入力し、該入力した電子データの位置、大きさ、向きの情報を含む基準情報に基づいて該電子データモデルを生成する第2の過程と、
 共通の座標系を設けて、前記映像モデルと前記電子データモデルを共通の座標系内で合成する第3の過程と、
 該第3の過程で合成されたモデルを入力し、前記共通の座標空間に仮想的な視点を設置して、該視点の映像を生成する第4の過程と、
 該第4の過程で取得した映像を送信する第5の過程と、を有することを特徴とする映像提示方法。

- [請求項7] 撮影装置によって取得された撮影映像を入力し、該入力した撮影映像に、画像として生成された電子データを加工して合成し、該合成した映像を送受信することにより、遠隔地間でコミュニケーションを行なうことを可能とした映像提示装置のコンピュータに、
- 撮影する範囲の奥行き情報と、該奥行き情報を基に該奥行き情報に貼り付けた映像情報とからなる映像モデルを生成する手順と、
- 電子データを入力し、該入力した電子データの位置、大きさ、向きの情報を含む基準情報に基づいて該電子データモデルを生成する手順と、
- 共通の座標系を設けて、前記映像モデルと前記電子データモデルを共通の座標系内で合成する手順と、
- 該合成されたモデルを入力し、前記共通の座標空間に仮想的な視点を設置して、該視点の映像を生成する手順と、
- 該生成された映像を送信する手順を実行させるための映像提示プログラム。
- [請求項8] 請求項7に記載の映像提示プログラムを、コンピュータ読み取り可能に記録した記録媒体。

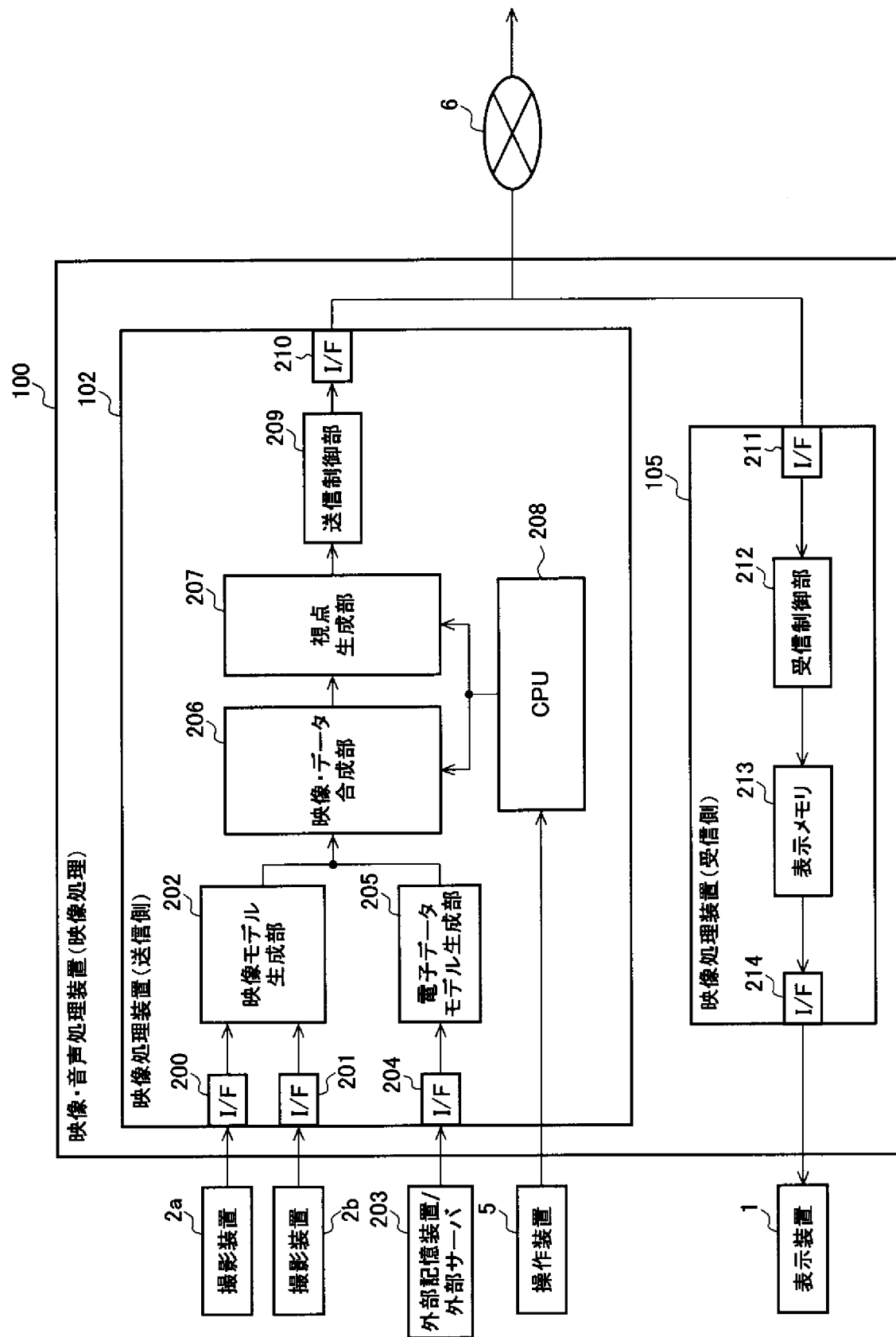
[図1]



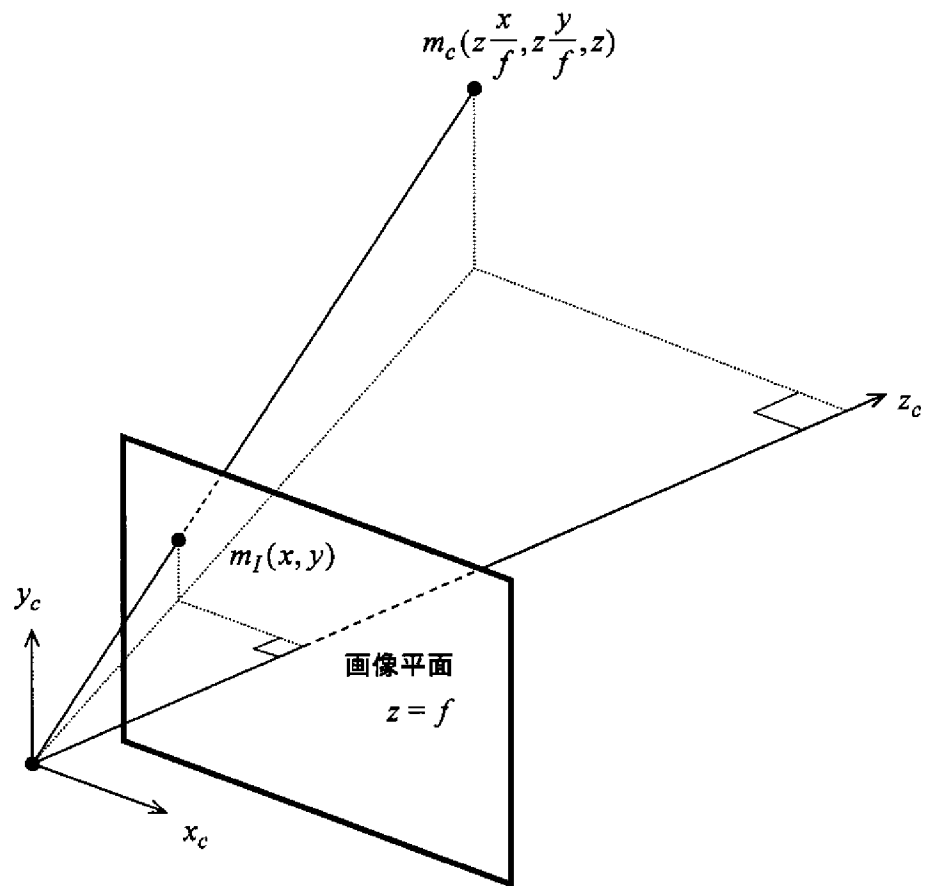
[図2]



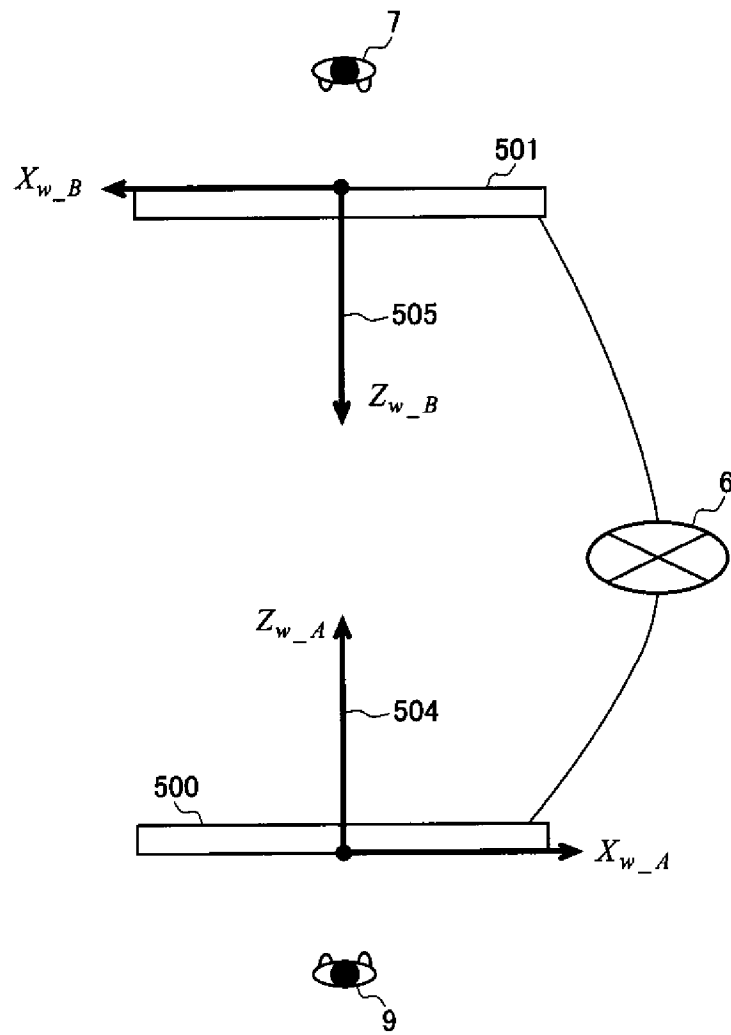
[図3]



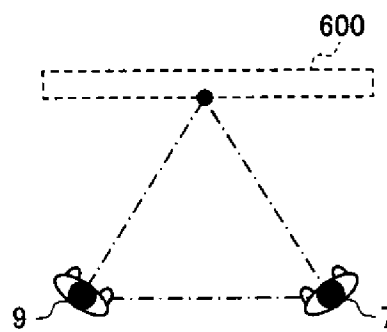
[図4]



[図5]

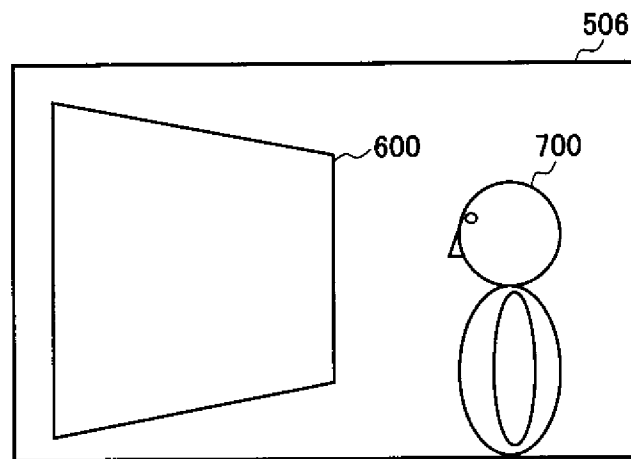


[図6]

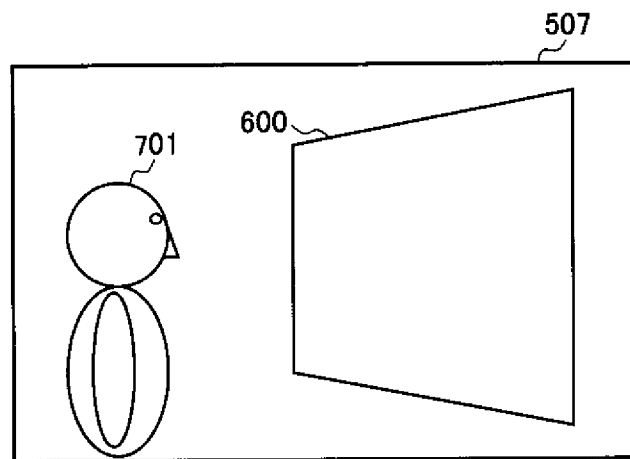


[図7]

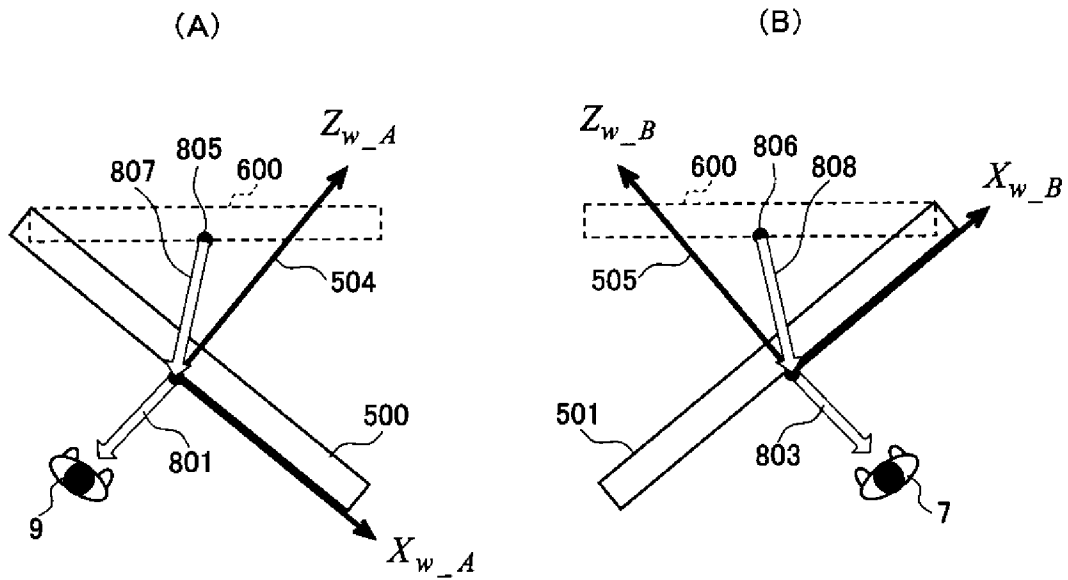
(A)



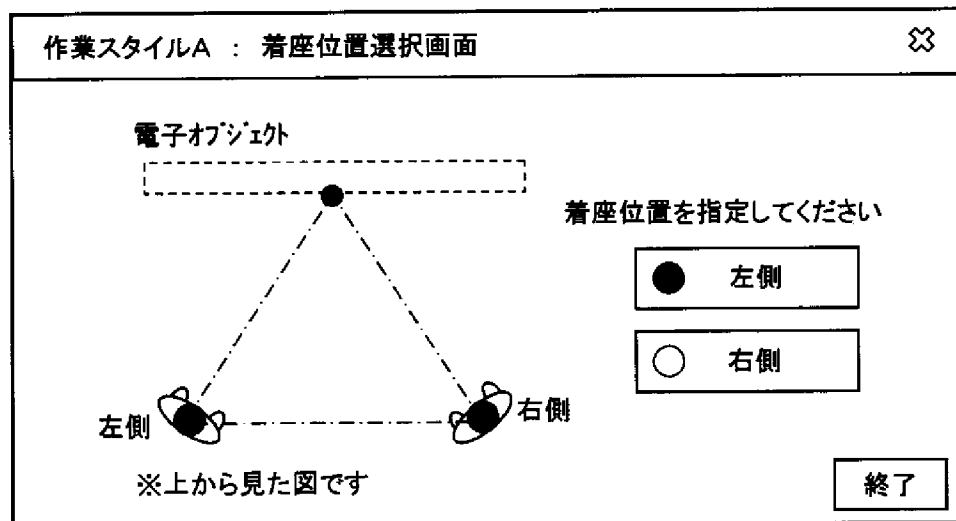
(B)



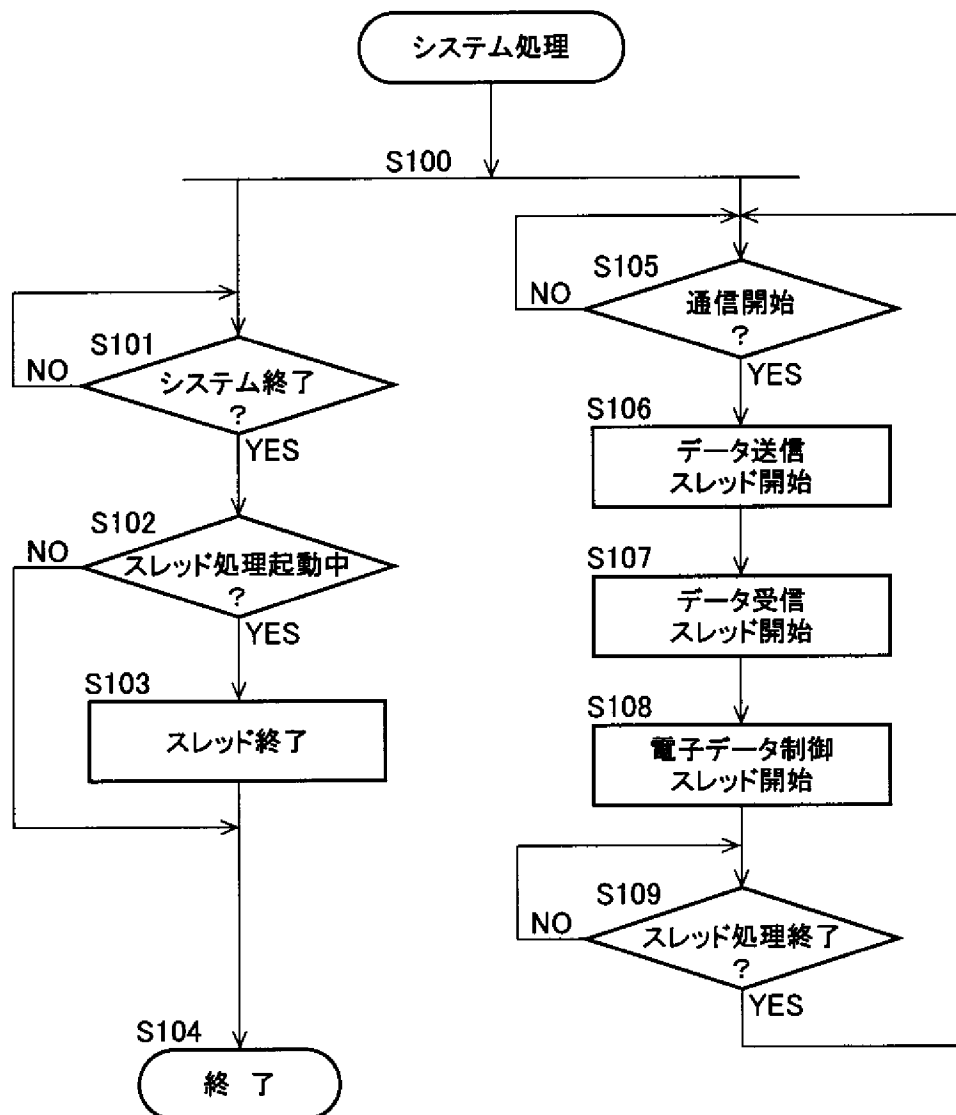
[図8]



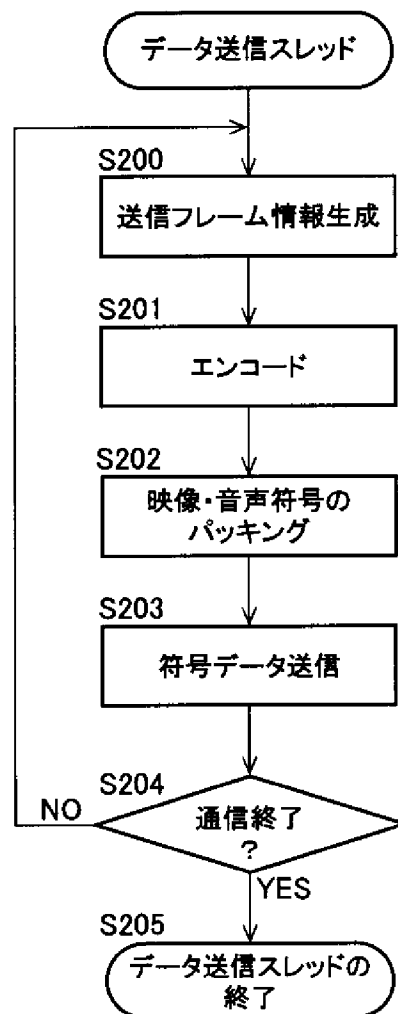
[図9]



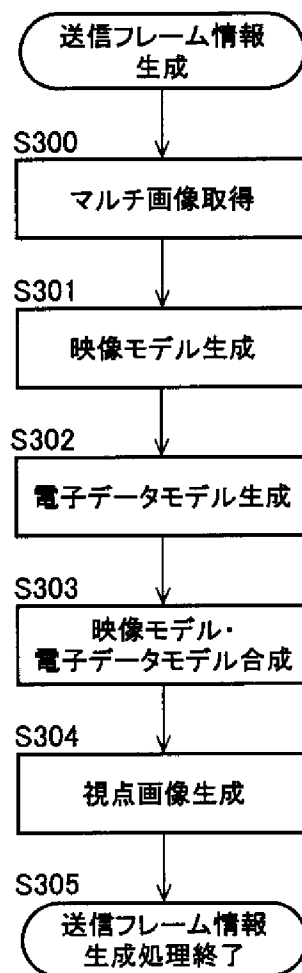
[図10]



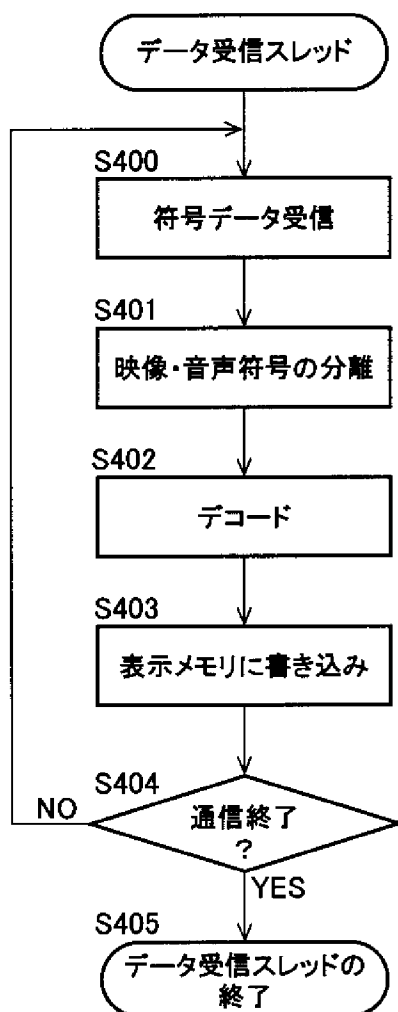
[図11]



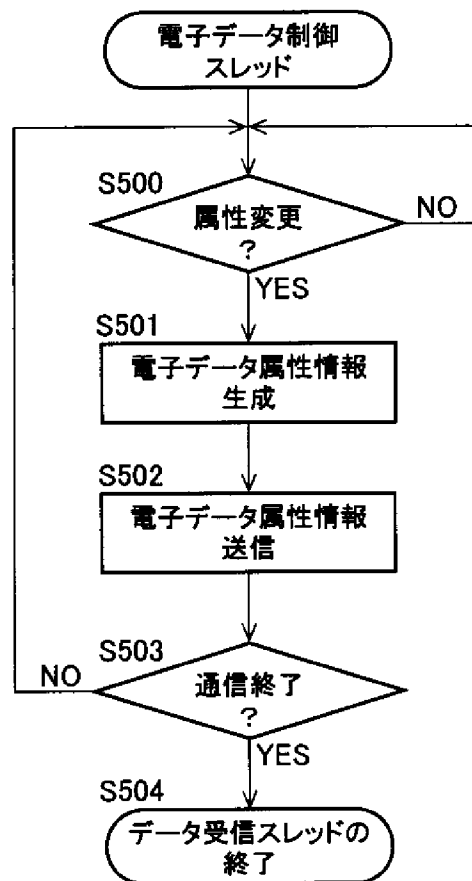
[図12]



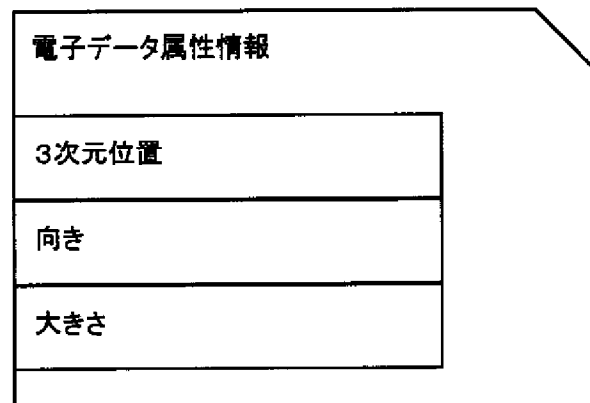
[図13]



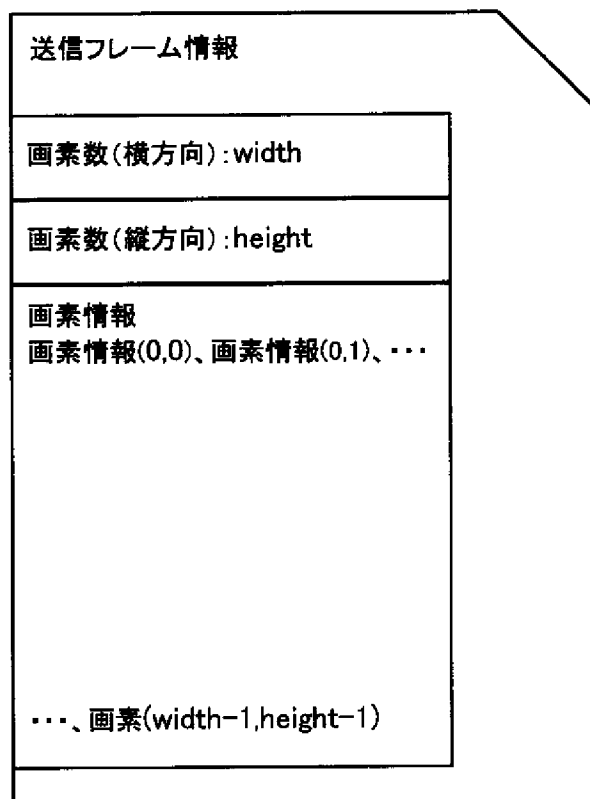
[図14]



[図15]



[図16]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/061996

A. CLASSIFICATION OF SUBJECT MATTER

G06T19/00 (2011.01) i

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)

G06T19/00

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

Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2012

Kokai Jitsuyo Shinan Koho 1971-2012 Toroku Jitsuyo Shinan Koho 1994-2012

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.
Y	JP 2006-293604 A (Canon Inc.), 26 October 2006 (26.10.2006), paragraphs [0079] to [0087]; fig. 6 to 13 & US 2006/0227151 A1 & EP 1710002 A1 & CN 1845064 A	1-8
Y	JP 08-153213 A (Hitachi, Ltd.), 11 June 1996 (11.06.1996), paragraph [0018]; fig. 1 (Family: none)	1-8
Y	Ko SUZUKI et al., "Development of Auto- stereoscopic Display System for Remote Manipulation : The Addition of Motion Parallax to Display Stereoscopic Images", ITE Technical Report, 18 October 2006 (18.10.2006), vol.30, no.52, pages 13 to 16	5

☒ 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 application or patent but 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
18 June, 2012 (18.06.12)Date of mailing of the international search report
26 June, 2012 (26.06.12)Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/061996

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2010-217719 A (Ricoh Co., Ltd.), 30 September 2010 (30.09.2010), paragraphs [0061] to [0064]; fig. 9 to 11 (Family: none)	1-8
A	JP 2005-049996 A (Canon Inc.), 24 February 2005 (24.02.2005), paragraphs [0015] to [0018]; fig. 1 to 2 & US 2005/0024388 A1 & EP 1503348 A2	1-8
A	JP 2004-178036 A (Hitachi, Ltd.), 24 June 2004 (24.06.2004), paragraphs [0009] to [0017]; fig. 1 to 6 (Family: none)	1-8

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. G06T19/00(2011.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（I P C））

Int.Cl. G06T19/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2012年
日本国実用新案登録公報	1996-2012年
日本国登録実用新案公報	1994-2012年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y	JP 2006-293604 A（キヤノン株式会社） 2006.10.26, 段落【0079】-【0087】，第6-13図 & US 2006/0227151 A1 & EP 1710002 A1 & CN 1845064 A	1-8
Y	JP 08-153213 A（株式会社日立製作所） 1996.06.11, 段落【0018】，第1図（ファミリーなし）	1-8

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

18.06.2012

国際調査報告の発送日

26.06.2012

国際調査機関の名称及びあて先

日本国特許庁（I S A/J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

千葉 久博

5H

3991

電話番号 03-3581-1101 内線 3531

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y	鈴木耕, 外 3 名, ” 遠隔作業支援用立体映像システムの開発 ―表示立体映像への運動視差付加―”, 映像情報メディア学会技術報告, 2006.10.18, 第 30 巻, 第 52 号, p.13-16	5
A	JP 2010-217719 A (株式会社リコー) 2010.09.30, 段落【0061】-【0064】, 第 9-11 図 (ファミリーなし)	1-8
A	JP 2005-049996 A (キヤノン株式会社) 2005.02.24, 段落【0015】-【0018】, 第 1-2 図 & US 2005/0024388 A1 & EP 1503348 A2	1-8
A	JP 2004-178036 A (株式会社日立製作所) 2004.06.24, 段落【0009】-【0017】, 第 1-6 図 (ファミリーなし)	1-8