

(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
2017年8月3日(03.08.2017)



(10) 国際公開番号

WO 2017/131071 A1

- (51) 国際特許分類:
G06T 19/00 (2011.01) A63B 69/00 (2006.01)
- (21) 国際出願番号: PCT/JP2017/002665
- (22) 国際出願日: 2017年1月26日(26.01.2017)
- (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 2016-014487 2016年1月28日(28.01.2016) JP
特願 2016-014488 2016年1月28日(28.01.2016) JP
特願 2016-107729 2016年5月30日(30.05.2016) JP
特願 2016-107730 2016年5月30日(30.05.2016) JP
特願 2016-171420 2016年9月2日(02.09.2016) JP
特願 2016-212800 2016年10月31日(31.10.2016) JP
- (71) 出願人: 日本電信電話株式会社(NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目5番1号 Tokyo (JP).
- (72) 発明者: 高橋 康輔(TAKAHASHI, Kosuke); 〒1808585 東京都武蔵野市緑町三丁目9番11号 N T T 知的財産センタ内 Tokyo (JP). 三上 弾

(MIKAMI, Dan); 〒1808585 東京都武蔵野市緑町三丁目9番11号 N T T 知的財産センタ内 Tokyo (JP). 五十川 麻理子(ISOGAWA, Mariko); 〒1808585 東京都武蔵野市緑町三丁目9番11号 N T T 知的財産センタ内 Tokyo (JP). 小島 明(KOJIMA, Akira); 〒1808585 東京都武蔵野市緑町三丁目9番11号 N T T 知的財産センタ内 Tokyo (JP). 木全 英明(KIMATA, Hideaki); 〒1808585 東京都武蔵野市緑町三丁目9番11号 N T T 知的財産センタ内 Tokyo (JP). 松本 鮎美(MATSUMOTO, Ayumi); 〒1808585 東京都武蔵野市緑町三丁目9番11号 N T T 知的財産センタ内 Tokyo (JP).

(74) 代理人: 中尾 直樹, 外(NAKAO, Naoki et al.); 〒1600022 東京都新宿区新宿三丁目1番22号 新宿NSビル6階 Tokyo (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, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN,

[続葉有]

(54) Title: VIRTUAL ENVIRONMENT CONSTRUCTION DEVICE, VIDEO PRESENTATION DEVICE, MODEL LEARNING DEVICE, OPTIMUM DEPTH DETERMINATION DEVICE, METHOD THEREFOR, AND PROGRAM

(54) 発明の名称: 仮想環境構築装置、映像提示装置、モデル学習装置、最適奥行き決定装置、それらの方法、およびプログラム

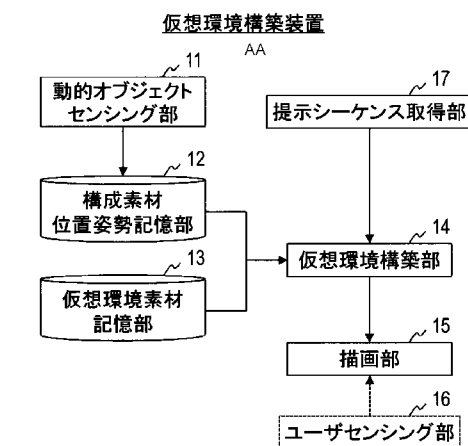


図5

- 11 Dynamic-object-sensing unit
12 Constituent material position/posture storage unit
13 Virtual environment material storage unit
14 Virtual environment construction unit
15 Drawing unit
16 User-sensing unit
17 Presentation sequence acquisition unit
AA Virtual environment construction device

(57) Abstract: The present invention enables a game to be experienced from a player's standpoint in advance. A virtual environment material storage unit 13 stores a virtual environment material for reproducing a dynamic object and a static object in a virtual environment. A dynamic object sensing unit 11 measures the position and posture of the dynamic object in a real environment in time series, and generates position/posture information comprising one exercise motion. A presentation sequence acquisition unit 17 acquires a presentation sequence that includes position/posture information pertaining to multiple kinds of dynamic objects differing from each other. A virtual environment construction unit 14 combines the virtual environment material of the dynamic objects and the virtual environment material of the static object on the basis of the presentation sequence and constructs a virtual environment.

(57) 要約: 事前に選手目線での試合体験を可能にする。仮想環境素材記憶部13に、動的オブジェクトと静的オブジェクトとを仮想環境上に再現する仮想環境素材が記憶されている。動的オブジェクトセンシング部11は、現実環境における動的オブジェクトの位置および姿勢を時系列に計測して1つの運動動作からなる位置姿勢情報を生成する。提示シーケンス取得部17は、相異なる複数種類の動的オブジェクトの位置姿勢情報を含む提示シーケンスを取得する。仮想環境構築部14は、提示シーケンスに基づいて動的オブジェクトの仮想環境素材と静的オブジェクトの仮想環境素材とを合成して仮想環境を構築する。



MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, ST, 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, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

— 国際調査報告 (条約第 21 条(3))

明 細 書

発明の名称：

仮想環境構築装置、映像提示装置、モデル学習装置、最適奥行き決定装置、それらの方法、およびプログラム

技術分野

[0001] この発明は、現実の環境を撮影した映像に基づいて仮想的かつ移動可能な視点からの映像を合成する仮想環境構築技術に関する。

背景技術

[0002] スポーツの試合において、事前に体験していない初めての状況に遭遇すると、十分なパフォーマンスを発揮できないことがある。例えば、野球では「初物に弱い」という表現を用いて初対戦の投手を苦手とする状況が語られることがある。この状況に対し、実際の試合と同じ環境において選手を目線で事前に投手の投球を体験することが有効であると考えられる。

[0003] しかしながら、実際に次の対戦相手の投球を事前に体験することは実施の観点から困難である。この対策としてスタンドなどから撮影した対戦相手の過去の投球映像を視聴することが一般的であるが、ここで視聴する映像は実際にバッターボックスに立っている選手が見ている映像と比べて視点や視野の観点から異なるため、試合状況の事前体験として十分ではない。

[0004] 非特許文献1には、選手を目線における事前体験の実現方法として、バッターボックス内の選手を目線の位置に360度のパノラマ画像を撮影できるカメラ（以下、全天球カメラと呼ぶ）を設置し、その全天球カメラで撮影した選手目線の映像をヘッドマウントディスプレイにより視聴する方法が記載されている。

先行技術文献

非特許文献

[0005] 非特許文献1：Ochi, D., Kunita, Y., Kameda, A., Kojima, A., Iwaki, S.,
“Live streaming system for omnidirectional video”, Virtual Reality

(VR), 2015 IEEE , pp. 349-350, March 2015.

発明の概要

発明が解決しようとする課題

- [0006] しかしながら、非特許文献 1 に記載の方法では、試合中にバッターボックス内にカメラを設置する必要があるため実施の観点から実現は困難である。加えて、このようにカメラで撮影された映像を視聴する場合、カメラを設置した視点のみの事前体験しかできないという課題がある。
- [0007] この発明の目的は、上記のような事情を鑑み、仮想的に試合環境を構築することで、事前に選手目線での試合体験を可能にする技術を提供することである。

課題を解決するための手段

- [0008] 上記の課題を解決するために、この発明の仮想環境構築装置は、動的オブジェクトと静的オブジェクトとが存在する現実環境に基づいてユーザに閲覧させる仮想環境を構築する仮想環境構築装置であって、動的オブジェクトと静的オブジェクトとを仮想環境上に再現する仮想環境素材を記憶する仮想環境素材記憶部と、現実環境における動的オブジェクトの位置および姿勢を時系列に計測して 1 つの運動動作からなる位置姿勢情報を生成する動的オブジェクトセンシング部と、相異なる複数種類の動的オブジェクトの位置姿勢情報を含む提示シーケンスを取得する提示シーケンス取得部と、提示シーケンスに基づいて動的オブジェクトの仮想環境素材と静的オブジェクトの仮想環境素材とを合成して仮想環境を構築する仮想環境構築部と、を含む。

発明の効果

- [0009] この発明の仮想環境構築技術によれば、仮想的に試合環境を構築することで、事前に選手目線での試合体験が可能となる。

図面の簡単な説明

- [0010] [図1]第一実施形態の仮想環境構築装置の機能構成を例示する図である。
- [図2]位置姿勢テーブルのデータ構造を例示する図である。

- [図3]構成素材組み込みリストのデータ構造を例示する図である。
- [図4]第二実施形態の仮想環境構築装置の機能構成を例示する図である。
- [図5]第三実施形態の仮想環境構築装置の機能構成を例示する図である。
- [図6]提示シーケンステーブルのデータ構造を例示する図である。
- [図7]提示シーケンステーブルの具体例を示す図である。
- [図8]仮想環境構築部の処理手続きを例示する図である。
- [図9]付与効果テーブルの具体例を示す図である。
- [図10]投球コースの具体例を示す図である。
- [図11]第四実施形態の仮想環境構築装置の機能構成を例示する図である。
- [図12]第五実施形態の仮想環境構築装置の機能構成を例示する図である。
- [図13]第五実施形態の仮想環境構築方法の処理手続きを例示する図である。
- [図14]動作データベースを説明するための図である。
- [図15]コースの分類の一例を説明するための図である。
- [図16]コースの分類の一例を説明するための図である。
- [図17]球種の分類を説明するための図である。
- [図18]シーケンスの生成方法を説明するための図である。
- [図19]確率分布の変形例を説明するための図である。
- [図20]第六実施形態の映像提示装置の機能構成を例示する図である。
- [図21]第六実施形態の映像提示方法の処理手続きを例示する図である。
- [図22]ユーザ状態の具体例を示す図である。
- [図23]ユーザ状態の具体例を示す図である。
- [図24]ユーザ状態の具体例を示す図である。
- [図25]第七実施形態の仮想環境構築装置の構成を示すブロック図である。
- [図26]第七実施形態の仮想環境構築装置の動作を示すフローチャートである。
- 。
- [図27]仮想素材位置姿勢情報テーブルを例示する図である。
- [図28]仮想特殊動的素材生成部の動作の例を示すフローチャートである。
- [図29]同期部の動作の例を示すフローチャートである。

[図30]バッターボックスに立つバッターの視点を再現するためのカメラアングルの例を示す模式図である。

[図31]第八実施形態の仮想環境構築装置の構成を示すブロック図である。

[図32]第八実施形態の仮想環境構築装置の事前準備フェーズにおける動作を示すフローチャートである。

[図33]映像情報テーブルを例示する図である。

[図34]第八実施形態の仮想環境構築装置の仮想環境生成フェーズにおける動作を示すフローチャートである。

[図35]仮想素材位置姿勢情報テーブルを例示する図である。

[図36]類似映像検索部の動作の例を示すフローチャートである。

[図37]仮想特殊動的素材生成部の動作の例を示すフローチャートである。

[図38]同期部の動作の例を示すフローチャートである。

[図39]第九実施形態の仮想環境構築装置の機能構成を例示する図である。

[図40]野球における具体例を説明するための図である。

[図41]動的オブジェクト構築部の処理手続きを例示する図である。

[図42]観測角度と視聴角度のずれを説明するための図である。

[図43]テニスにおける具体例を説明するための図である。

[図44]テニスにおける具体例を説明するための図である。

[図45]原画像または映像とマスク情報の例を示す図である。

[図46]第十実施形態の映像提示装置の機能構成を例示する図である。

[図47]第十実施形態の映像提示方法の処理手続きを例示する図である。

[図48]実験結果を示す図である。

[図49]画像処理システムのシステム構成を示す図である。

[図50]合成情報テーブルの具体例を示す図である。

[図51]最適奥行き情報テーブルの具体例を示す図である。

[図52]コントラストの算出結果の具体例を示す図である。

[図53]第十一実施形態の画像処理装置の処理手続きを例示する図である。

[図54]第十二実施形態の画像処理装置の機能構成を例示する図である。

[図55]第十二実施形態の画像処理装置の処理手続きを例示する図である。

[図56]第十三実施形態の画像処理装置の機能構成を例示する図である。

[図57]第十三実施形態の決定部の具体的な処理について説明するための図である。

[図58]第十三実施形態の画像処理装置の処理手続きを例示する図である。

[図59]第十四実施形態の画像処理装置の機能構成を例示する図である。

[図60]第十四実施形態の画像処理装置の処理手続きを例示する図である。

[図61]従来システムにおいて仮想全天球画像を得るためのシステムを示す図である。

[図62]画像処理システムにおける画像処理の流れを説明するための図である。

発明を実施するための形態

[0011] 以下、この発明の実施の形態について詳細に説明する。なお、図面中において同じ機能を有する構成部には同じ番号を付し、重複説明を省略する。

[0012] [第一実施形態]

最初に、第一実施形態の仮想環境構築装置の概要を説明する。まず、動的オブジェクトセンシング部が事前体験の対象である動的なオブジェクト（例えば、野球の投球やテニスのサーブの軌道など）に関する各時刻の位置および姿勢を求め、その位置姿勢情報を記憶しておく。時間変化をしない静的なオブジェクト（例えば、スタジアムやグラウンドなど）に関する位置姿勢情報は予め記憶されているものとする。以下では、ボールやスタジアムのように仮想環境を構成するためのオブジェクトを構成素材と呼ぶ。また、構成素材を仮想環境に再現するためのコンピュータ・グラフィックス（CG: Computer Graphics）データを仮想環境素材と呼び、これらも予め記憶されているものとする。構成素材のうち、その位置が時間経過により遷移し、かつ動く素材を動的素材と呼び、その位置が時間経過により遷移せず動かない構成素材を静的素材と呼び、その位置は遷移しないものとみなされるがその位置に留まったまま動く構成素材を特殊動的素材と呼ぶ。例えば野球であれば、動的

素材は野球ボールであり、静的素材はスタジアムであり、特殊動的素材はピッチャーである。動的オブジェクトは、動的素材と特殊動的素材とからなる。静的オブジェクトは、静的素材である。なお、静的オブジェクトに関する仮想環境素材はCGデータに限定されず、実際に撮影した映像から切り出した動画または静止画データであってもよい。仮想環境構築部は各構成素材の位置姿勢情報に基づいて仮想環境素材を配置することで仮想環境を構築する。最後に、描画部がその仮想環境を全天球映像などの映像として出力する。

[0013] 第一実施形態の仮想環境構築装置は、図1に示すように、動的オブジェクトセンシング部11、構成素材位置姿勢記憶部12、仮想環境素材記憶部13、仮想環境構築部14、および描画部15を備える。この仮想環境構築装置が後述する各ステップの処理を行うことにより第一実施形態の仮想環境構築方法が実現される。

[0014] 仮想環境構築装置は、例えば、中央演算処理装置（CPU: Central Processing Unit）、主記憶装置（RAM: Random Access Memory）などを有する公知又は専用のコンピュータに特別なプログラムが読み込まれて構成された特別な装置である。仮想環境構築装置は、例えば、中央演算処理装置の制御のもとで各処理を実行する。仮想環境構築装置に入力されたデータや各処理で得られたデータは、例えば、主記憶装置に格納され、主記憶装置に格納されたデータは必要に応じて読み出されて他の処理に利用される。また、仮想環境構築装置の各処理部の少なくとも一部が集積回路等のハードウェアによって構成されていてもよい。仮想環境構築装置が備える各記憶部は、例えば、RAM（Random Access Memory）などの主記憶装置、ハードディスクや光ディスクもしくはフラッシュメモリ（Flash Memory）のような半導体メモリ素子により構成される補助記憶装置、またはリレーショナルデータベースやキーバリューストアなどのミドルウェアにより構成することができる。仮想環境構築装置が備える各記憶部は、それぞれ論理的に分割されていればよく、一つの物理的な記憶装置に記憶されていてもよい。

[0015] 動的オブジェクトセンシング部11は、現実環境における動的オブジェクト

トの位置および姿勢を計測して1つの運動動作からなる位置姿勢情報を生成する。生成された動的オブジェクトの位置姿勢情報は構成素材位置姿勢記憶部12へ記憶される。具体的には、仮想環境構築装置に接続されたセンサ（図示せず）で動的オブジェクトの1つの運動動作における位置および姿勢を計測し、その値を入力として動的オブジェクトの3次元位置姿勢情報を推定する。動的オブジェクトとは、仮想環境において時間経過とともに3次元位置が変化する物体、例えば野球であれば野球ボールやピッチャーなどの相手選手を、テニスであればテニスボールやサーバーなどの相手選手を指す。1つの運動動作とは、例えば野球ボールであればピッチャーの投球一球分の動作であり、テニスボールであればサーブ一球分の動作である。

[0016] 動的オブジェクトの3次元位置姿勢を算出するためのセンサとして複数カメラが利用できる。複数カメラを利用してボールの3次元位置情報を算出する手法の具体例として、ステレオ法を用いた手法がある。具体的には、予めカメラキャリブレーションを実施したカメラ群で撮影された映像に対し、画像上のボールの位置を求める。このとき、画像上のボールの位置は手作業で与えてもよいし、テンプレートマッチングなどの検出手法を用いてボールを検出し、その中心位置をボールの位置としてもよい。これら画像上で検出された位置とカメラキャリブレーションで求めたカメラパラメータとを用いてステレオ法を適用することで、ボールの3次元位置を求めることができる。また、姿勢に関しては、例えば、撮影したボールの縫い目の模様から基準の姿勢に対してどの程度回転しているかを求めることができる。動的オブジェクトの位置姿勢を求めるセンサの他の例としては、ドップラーレーダーなどを挙げることができる。このセンサは公知の技術であるため詳細な説明は省略する。その他、ボールの3次元位置姿勢情報を取得できるセンサおよび手法であればどのようなものであってもよい。

[0017] 構成素材位置姿勢記憶部12は、各構成素材に関する位置姿勢情報を記憶する。動的オブジェクトに関する位置姿勢情報は動的オブジェクトセンシング部11により生成されて記憶されたものである。静的オブジェクトに関する

る位置姿勢情報は予め用意されて記憶されたものである。具体的には、各時刻における各構成素材の3次元位置姿勢情報を表すレコードによって構成される構成素材位置姿勢テーブルが記憶されている。

[0018] 構成素材位置姿勢情報テーブルは、図2に示すように、構成素材を一意に識別する「構成素材ID」と、仮想環境素材を一意に識別する「仮想環境素材ID」と、各「時刻」におけるオブジェクトの「位置」および「姿勢」の各値を有する。位置は3次元の並進ベクトルとして表され、姿勢は各軸に対する回転角を表す3次元ベクトルとして表される。例えば、野球の投球を想定した場合、構成素材 $0i$ がストレートボールを表すとすると、構成素材 $0i$ に対応する仮想環境素材 Ei は、仮想環境素材記憶部13に記憶された野球ボールのCGデータである。このとき、各時刻に対応する位置および姿勢はその時刻にボールがどの位置にどの姿勢で存在しているかということを表している。より具体的には、位置 $Ti(0)$, $Ti(1)$, ..., $Ti(ti)$ はボールの軌道を、姿勢 $Ri(0)$, $Ri(1)$, ..., $Ri(ti)$ はボールがどのように回転しているかを表している。なお、図2において、時刻に「-」が格納されている構成素材は、スタジアムのように時間経過に伴い位置や姿勢が変化しない静的オブジェクトであることを表す。このような静的オブジェクトの位置姿勢情報は予め記憶されているものとする。また、各構成素材に対して定められている時刻は、その構成素材が後述する仮想環境構築部14において仮想環境に組み込まれた時点からの経過時刻を表す。具体的には、仮想環境において時刻 ta の時点で構成素材 $0i$ が組み込まれた場合、時刻 $(ta+0)$ の時点の構成素材 $0i$ の位置および姿勢はそれぞれ $Ti(0)$, $Ri(0)$ であり、時刻 $(ta+1)$ の時点の構成素材 $0i$ の位置および姿勢はそれぞれ $Ti(1)$, $Ri(1)$ である。

[0019] 仮想環境素材記憶部13は、各構成素材に対応する仮想環境素材を記憶する。仮想環境素材は仮想環境を構築するための素材であり、例えばCGデータや実写データなどである。仮想環境素材記憶部13は仮想環境素材IDに関連付けて各仮想環境素材を記憶する。仮想環境素材の例としては、野球であればスタジアムや野球ボールのCGデータ、テニスであればテニスコートや

テニスボールのCGデータである。これらは事前に用意され記憶されているものとする。

[0020] 仮想環境構築部14は、構成素材位置姿勢記憶部12から各構成素材の位置姿勢情報を読み出し、仮想環境素材記憶部13から各構成素材に対応する仮想環境素材のデータを読み出し、動的オブジェクトの仮想環境素材と静的オブジェクトの仮想環境素材とを合成して仮想環境のデータを構築する。構築した仮想環境のデータは描画部15へ送られる。仮想環境のデータとは、ある仮想的な空間において一つ以上の仮想環境素材のデータが配置された仮想空間を指す。具体例として、野球における仮想環境のデータを作成する手順を説明する。まず、仮想環境を構築するための構成素材組み込みリストを読み込む。この構成素材組み込みリストは、図3に示すように、構成素材を一意に識別する「構成素材ID」と、その構成素材を仮想環境に組み込む「時刻」を持つ。構成素材組み込みリストは予め人手で用意されているものとする。例えば野球における仮想環境のデータを作成する場合、スタジアムやピッチャー、ボールなどに関する構成素材IDおよび時刻が構成素材組み込みリストに記載されている。次に、各時刻について描画すべき構成素材が存在する場合、その構成素材に対応する仮想環境素材を仮想環境に組み込む。このとき、構成素材の位置姿勢を構成素材位置姿勢記憶部12から読み出し、その位置姿勢に基づいて仮想空間に配置する。構成素材組み込みリストに記載されている全ての構成素材を仮想環境に組み込んだら処理を終了する。

[0021] 描画部15は、構築した仮想環境のデータをユーザが視聴できる形態として出力する。出力する映像は、仮想環境のある視点から一方向を見るような映像でもよいし、全天球映像のように全方位を見ることができ映像でもよい。例えば、ある視点から一方向を見るような映像を出力する場合は、ある位置に仮想的に透視投影カメラCを設置し、仮想環境の空間中に存在するCGモデルをカメラCの画像平面に投影すればよい。具体的には、TをカメラCの外部パラメータである位置とし、Rを姿勢とし、Kを内部パラメータとすると、仮想環境のデータが存在する空間中のある点pは、以下のようにカメラCの画

像平面上に点 q として投影される。

[0022] [数1]

$$q = K(Rp + T)$$

[0023] ただし、内部パラメータ K はカメラ C の焦点距離や画像中心からなる 3×3 の行列である。

[0024] ある視点から全方位を見渡せる全天球映像を出力する場合は、ある視点に仮想的に全天球カメラ 0 を設置し、仮想環境の空間中に存在するCGモデルをカメラ 0 の画像平面に投影すればよい。具体的には、 T をカメラ 0 の外部パラメータである位置とし、 R を姿勢とし、 W を投影する画像平面の幅とし、 H を高さとする、仮想環境のデータが存在する空間中のある点 p は、以下のようにカメラ 0 の画像平面上に点 $q=(q_x, q_y)^T$ として投影される。

[0025] [数2]

$$p' = Rp + T = (p'_x p'_y p'_z)^T$$

$$q_x = \left(\frac{\theta W}{2\pi} \right) + \frac{W}{2}$$

$$q_y = \frac{\phi H}{\pi}$$

ただし、

$$\theta = \arctan \left(\frac{p'_z}{p'_x} \right),$$

$$\phi = \arccos \left(\frac{p'_y}{|p'|} \right)$$

[0026] この例では全天球映像を2次元画像で表現する図法として正距円筒図法を用いているが、メルカトル図法など他の図法を用いてもよい。

[0027] [第二実施形態]

第一実施形態では、描画部において仮想環境内の予め定めた視点からの映像のみを出力している。そのため、仮想環境内を自由に移動するような体験の実現が困難である。全ての地点における映像を出力するということも考え

られるが、処理量の観点から現実的ではない。第二実施形態では第一実施形態の構成に対してユーザの位置および姿勢をセンシングするユーザセンシング部を追加し、ユーザが視聴したい位置および方向をセンシングすることで、仮想環境内を自由に移動しながら映像を視聴できるようにする。

[0028] 第二実施形態の仮想環境構築装置は、図4に示すように、動的オブジェクトセンシング部11、構成素材位置姿勢記憶部12、仮想環境素材記憶部13、仮想環境構築部14、および描画部15を第一実施形態と同様に備え、ユーザセンシング部16をさらに備える。この仮想環境構築装置が後述する各ステップの処理を行うことにより第二実施形態の仮想環境構築方法が実現される。

[0029] 以下、第二実施形態の仮想環境構築方法の処理手続きを、第一実施形態との相違点を中心に説明する。

[0030] ユーザセンシング部16は、ユーザの頭部の位置および姿勢を計測してユーザ位置姿勢情報を生成する。生成されたユーザ位置姿勢情報は描画部15へ送られる。具体的には、仮想環境構築装置が備えるセンサ（図示せず）でユーザの頭部の位置および姿勢を計測し、その値を入力としてユーザの頭部の3次元位置姿勢情報を推定する。

[0031] ユーザの頭部の3次元位置姿勢を算出するためのセンサはどのようなセンサでもよい。このセンサの具体例として、Optitrack社に代表されるモーションセンサーが挙げられる。また、Oculus社に代表されるヘッドマウントディスプレイ（HMD: Head Mount Display）と付属の赤外線（IR: Infrared）カメラを用いてもよい。これらは公知の技術を用いて実現されるため詳細な説明は省略する。

[0032] 描画部15は、ユーザセンシング部16より受け取ったユーザ位置姿勢情報に基づいて、仮想環境のデータをユーザが視聴できる形態として出力する。このとき、仮想環境におけるユーザの視線方向の映像を出力してもよいし、全方位映像を出力してもよい。

[0033] [第三実施形態]

一般に、どのような能力を向上させたいかに応じてトレーニングの内容は変化する。例えば、テニスのサーブの感覚を体に覚えさせたい場合には重点的にサーブの練習をすることが効果的であると考えられる。また、サーブの威力を増すことを目的とする場合には体の使い方の矯正や筋力トレーニングが効果的であると考えられる。第三実施形態では、第一実施形態に対して提示シーケンス取得部を追加することで、ユーザの練習意図に応じて構築する仮想環境を変化させることを可能にする。

[0034] 第三実施形態の仮想環境構築装置は、図5に示すように、動的オブジェクトセンシング部11、構成素材位置姿勢記憶部12、仮想環境素材記憶部13、仮想環境構築部14、および描画部15を第一実施形態と同様に備え、提示シーケンス取得部17をさらに備える。この仮想環境構築装置が後述する各ステップの処理を行うことにより第三実施形態の仮想環境構築方法が実現される。

[0035] 以下、第三実施形態の仮想環境構築方法の処理手続きを、第一実施形態との相違点を中心に説明する。

[0036] 提示シーケンス取得部17は、構成素材位置姿勢記憶部12に記憶されている、相異なる複数種類の動的オブジェクトの位置姿勢情報を含む、ユーザに閲覧させたい場面を表現した提示シーケンスを取得する。取得した提示シーケンスは仮想環境構築部14へ送られる。具体的には、ユーザからの入力を取得して、予め用意した提示シーケンステーブルからユーザの入力に対応する提示シーケンスに関するレコードを抽出して出力する。ユーザからの入力を取得する方法はどのような方法でもよく、例えば、キーボードでキーを入力する方法、マウスでクリックする方法、またはゲーム用のコントローラを用いてキーを入力する方法などが挙げられる。また、提示シーケンステーブルは、図6に示すように、「練習意図」と、その練習意図に対する「構成素材ID」群と、各構成素材が仮想環境に組み込まれる「時刻」によって構成される。例えば、図6の例では「〇〇の能力を向上させたい」という練習意図に関しては、時刻t1に構成素材01が組み込まれ、時刻t2に構成素材02, 03

が組み込まれる。ただし、ここでの時刻は仮想環境構築部 14 に出力された時点からの経過時刻である。すなわち、ある時刻 t に「〇〇の能力を向上させたい」という練習意図に関するレコード（提示シーケンス）が出力された場合、構成素材01は時刻 $(t+t1)$ に、構成素材02, 03は時刻 $(t+t2)$ に仮想環境へ組み込まれる。また、この提示シーケンステーブルに関し、あるユーザの入力に対して各提示シーケンスが対応づいており、その対応は予め定められているものとする。すなわち、あるキーが押下された場合には「〇〇の能力を向上させたい」という練習意図に関するレコード（提示シーケンス）が出力されるということである。図 7 は野球を例としたときの提示シーケンステーブルの具体例である。

[0037] 仮想環境構築部 14 は、仮想環境素材のデータ、各構成素材の位置姿勢情報、および提示シーケンステーブルを入力とし、仮想環境のデータを構築する。図 8 は本形態の仮想環境構築部 14 の処理手続きを例示する図である。なお、仮想環境構築部 14 は図 3 に示した構成素材組み込みリストを予め保持しているものとする。

[0038] ステップ S 11 において、仮想環境構築部 14 は、ある時刻 t において提示シーケンステーブルの入力があったか否かを判定する。入力があった場合（YES）、ステップ S 12 へ処理を進め、入力がなかった場合（NO）、ステップ S 13 へ処理を進める。ステップ S 12 において、構成素材組み込みリストに提示シーケンステーブルを統合する。テーブルの統合は時刻をキーとして行う。ただし、ある時刻 t に図 6 に示す提示シーケンステーブルの入力があった場合、構成素材01は時刻 $(t+t1)$ に、構成素材02は時刻 $(t+t2)$ に仮想環境へ組み込まれる。ステップ S 13 において、時刻 t について組み込むべき構成素材があるか否かを判定する。組み込むべき構成素材がある場合（YES）、ステップ S 14 へ処理を進め、組み込むべき構成素材がなかった場合（NO）、ステップ S 15 へ処理を進める。ステップ S 14 において、その構成素材を仮想環境へ組み込む。ステップ S 15 において、現在構築している仮想環境を描画部 15 へ出力する。その後、ステップ S 16 において、 $t=t$

+1としてステップS 1 1へ戻る。

[0039] 以下、野球の打者を例として向上させたい能力とその向上方法について、より具体的に説明する。なお、本具体例では、提示シーケンス取得部 1 7 において利用する提示シーケンステーブルを作成するときに、どのような順番でどのような投球を実施すべきかを直感的にわかりやすくするため、図 9 に示すように、各投球に対して「球速」、「コース」、「球種」などの属性を追加して保持する。なお、これらの属性は、構成素材位置姿勢記憶部 1 2 に記憶されている各構成素材の位置および姿勢の変化から求めることができる。例えば、球速はある時間内に移動した距離をその時間で除算することで求めることができる。また、コースは図 1 0 に示すようにホームベース上の平面を分割し、ボールがその平面上を通過する際にどの領域に最も近いかで求めることができる。球種は、リリース直後とキャッチャーミットに収まる時のボールの速度ベクトル（それぞれ v_1 、 v_2 とする）を用いることで求めることができる。例えば、 v_1 と v_2 とが概ね同じ方向であればストレート、 v_1 と v_2 とを比較して v_2 の方向が鉛直下方向に向いている場合はフォーク、斜め下方向に向いている場合はカーブ、といった具合である。これらの属性を投球に対応する構成素材に対して保持しておく。なお、これらは野球の投球における属性の例であり、サッカーやテニスの場合には各競技の特性に合わせて異なる属性とすればよい。

[0040] (a) 緩急に対する能力を向上させる場合

速いボールの後に遅いボールを投げられる、あるいは遅いボールの後に速いボールを投げられると、直前のボールのタイミングに影響されて適切なタイミングでバットを振れない場合がある。これに対し、速いボールと遅いボールをそれぞれランダムな順番で体験させることで、緩急に対する能力が向上すると考えられる。この場合、提示シーケンス取得部 1 7 では予めランダムに投球を行うように構成素材ID群を用意することで、ユーザはこの緩急に対する能力を向上させるトレーニングを行うことができる。このトレーニングのための配球例を以下に示す。

[0041] 配球例 (i) : ストレート (球速140km/h) –カーブ (球速120km/h)
–ストレート (球速150km/h) –スローボール (球速110km/h) …

配球例 (ii) : ストレート (球速140km/h) –ストレート (球速145km/h) –カーブ (球速120km/h) –ストレート (球速135km/h) –ストレート (球速140km/h) …

[0042] なお、上記配球例の各投球ではコースはどのようなコースでもよい。配球例 (i) の提示シーケンステーブルは図7の具体例のとおりである。また、これらの配球は手動で決めてもよいし、自動で決めてもよい。自動で決める場合は、例えば全ての投球の中からランダムで選び出してもよいし、予め用意した閾値thよりも球速が速い場合はグループ1とし、遅い場合はグループ2としてこれらのグループから選び出される割合が予め定めた割合になるように (例えばグループ1から10球、グループ2から5球など) ランダムに選び出してもよい。

[0043] (b) 選球眼 (ストライクとボールを見分ける能力) を向上させる場合
ストライクゾーンとボールゾーンの境界付近に投げられたボールに対して、適切にジャッジすることは難しい。これに対し、同じような軌道でストライク判定の投球とボール判定の投球を交互に繰り返し体験させることで選球眼が向上すると考えられる。また、頻繁に用いられる配球例として一度バッターの胸元に投げてから外角低めに投げることもある。これはバッターの目線を胸元に慣れさせておいて、そこから最も遠い外角低めに投げることでバッターの反応を遅らせる、あるいはストライクとボールの判断を鈍らせる効果があると言われている。これに対し、ストライクゾーンの一方の隅への投球と対角線上の他方の隅への投球とを交互に繰り返し体験させることで選球眼が向上すると考えられる。提示シーケンス取得部17ではこれらの投球を実施するように構成素材ID群を用意することでユーザは選球眼を向上させるトレーニングを行うことができる。

[0044] 配球例 (iii) : ストレート (コース2-A) –ストレート (コース2-B)
–ストレート (コース1-B) –ストレート (コース2-B) …

配球例 (iv) : ストレート (コース4-D) – ストレート (コース2-B)
– ストレート (コース4-D) – ストレート (コース2-D) …

[0045] なお、上記配球例の各投球では球速はどのような球速でもよい。配球例 (i ii) の提示シーケンステーブルは図7の具体例のとおりである。

[0046] なお、第三実施形態の仮想環境構築装置は、第一実施形態の構成を基本として構成および処理手続きを説明したが、第二実施形態の構成を基本として構成してもよい。すなわち、第三実施形態の仮想環境構築装置は、ユーザセンシング部16をさらに備えるように構成してもよい。

[0047] [第四実施形態]

第一実施形態から第三実施形態までは実際に得られた軌跡、すなわち現実と同じ環境を構築してトレーニングを行うことを想定している。しかしながら、トレーニングにおいては現実よりも厳しい環境を体験することで本番でのパフォーマンスをより良くするという取り組みがある。例えば、野球において速い球に目を慣れさせるために実際のマウンドよりも近いところからの投球を体験させることや、ネクストバッターズサークルにおいて通常のバットよりも重いマスコットバットを用いて素振りをするなどが挙げられる。第四実施形態では効果付与部を追加することで、現実よりも厳しい環境を体験するための仮想環境を実現する。

[0048] 第四実施形態の仮想環境構築装置は、図9に示すように、動的オブジェクトセンシング部11、構成素材位置姿勢記憶部12、仮想環境素材記憶部13、仮想環境構築部14、描画部15、および提示シーケンス取得部17を第三実施形態と同様に備え、効果付与部18をさらに備える。この仮想環境構築装置が後述する各ステップの処理を行うことにより第四実施形態の仮想環境構築方法が実現される。

[0049] 以下、第四実施形態の仮想環境構築方法の処理手続きを、第三実施形態との相違点を中心に説明する。

[0050] 本形態の提示シーケンス取得部17は、提示シーケンスに加えて、付与効果に関する情報を出力する。付与効果に関する情報は提示シーケンスとともに

に効果付与部 18 へ送られる。付与効果とは、構成素材IDに対応する構成素材に付与する効果の内容である。例えば、ユーザに対する視覚的負荷を高める効果が挙げられる。具体的には、構成素材の位置姿勢の変化率を変更したり（例えば、球速アップ、ダウン、回転量増加など）、構成素材を途中で透明にするなどの見え方に関する特殊効果を与えたり、といった効果が挙げられる。なお、あるボタンをクリックすればボールのスピードがアップするといったように、あるユーザ入力に対応する付与効果は予め定められていることとする。

[0051] 効果付与部 18 は、仮想環境素材のデータ、各構成素材の位置姿勢情報、および付与効果に関する情報を入力とし、付与効果に関する情報に基づいて、仮想環境素材のデータおよび各構成素材の位置姿勢情報に対して効果を与えて出力する。

[0052] 以下、第三実施形態と同様に野球を例として効果を付与する方法について、より具体的に説明する。

[0053] (a) 速球に対するバッティング能力を向上させる場合

速い球に目や体が慣れていると、それより遅い球には余裕を持って適切なタイミングでバットを振ることができると考えられる。しかしながら、実際に投手が投げたボールの軌跡を獲得する場合に、必ずしもその投手がどの投手よりも速い球を投げられるとは限らない。このような場合、提示シーケンス取得部 17 において「球速を10km/hアップする」という付与効果の指示を出す。このとき、効果付与部 18 に入力された投球 $0i$ の球速が $v1$ であったとすると、 $0i$ の各単位時間あたりの位置姿勢の変化量を $((v1+10)/v1)$ 倍することで球速を10km/hアップした投球の位置姿勢を求めることができる。これを用いて速球に対するバッティング能力を向上させるためのトレーニングを実施することができる。

[0054] (b) 変化球に対するバッティング能力を向上させる場合

変化量の大きい変化球に目や体が慣れていると、変化量の少ない球には余裕を持って適切なタイミングおよび位置にバットを振ることができると考え

られる。しかしながら、実際に投手が投げたボールの位置姿勢を獲得する場合に、必ずしもその投手が変化量のよい変化球を投げられるとは限らない。このような場合、提示シーケンス取得部 17 において「ボールの変化量を Δa %アップする」という付与効果の指示を出す。このとき効果付与部 18 に入力された投球 0i の加速度がある地点で a であった場合、その加速度に Δa を加えて再度位置姿勢を計算することで、より変化量の増した変化球の位置姿勢を作成することができる。これにより変化球に対するバッティング能力を向上させるためのトレーニングを実施することができる。

- [0055] (c) 投球の初期の軌道から全体の軌道を把握する能力を向上させる場合
球速が 130km/h の場合ボールがリリースされてからキャッチャーミットに収まるまで 0.5 秒程度しかなく、コースや球種の判断をギリギリまで待っている場合にバットを振り出すタイミングが遅れてしまうという問題がある。このような場合、提示シーケンス取得部 17 において「ボールを見えにくくする」という付与効果の指示を出す。このとき、効果付与部 18 では入力として与えられた投球に関してリリースポイントから d メートル離れた地点までは通常の野球ボールの CG を組み込み、d メートル離れた地点を過ぎてから野球ボールを見えにくくする効果を加える（例えば、透明にしたり、徐々に透明にしたりする）。このときの d の値や加える効果は適宜設定すればよい。提示シーケンス取得部 17 ではこのような見えにくくする効果を加えた投球と、全く同一の投球で効果を加えていない投球を交互に組み込むことで、ユーザは投球の初期の軌道、すなわちリリースポイントから d メートル離れた場所までのボールの位置姿勢の変化から、その投球の全体の軌道を予測するためのトレーニングを実施することができる。

- [0056] なお、第四実施形態の仮想環境構築装置は、第一実施形態の構成を基本として構成および処理手続きを説明したが、第二実施形態の構成を基本として構成してもよい。すなわち、第四実施形態の仮想環境構築装置は、ユーザセンシング部 16 をさらに備えるように構成してもよい。

- [0057] [第五実施形態]

例えばスポーツ等のトレーニングにおいて、試合と類似の体験をバーチャルに生成して選手が事前に経験することで、試合でのパフォーマンス向上を目指すシステムが提案されている（例えば、参考文献１）。参考文献１には、フットボールの選手を指導するにあたって、仮想現実（VR: Virtual Reality）技術を用いて試合と同じ環境を再現した選手目線からの３６０度映像をヘッドマウントディスプレイで視聴させるシステムが記載されている。

[0058] 〔参考文献１〕STRIVR Labs, Inc, “STRIVR | Immersive Training Solutions”、[online]、[平成29年1月4日検索]、インターネット<URL: <http://www.strivrlabs.com/home/about-us>>

[0059] スポーツの練習では、試合でのパフォーマンス向上を目的として、実際の試合よりも厳しい環境を体験することがしばしば行われている。例えば、女子チームが男子チームを練習相手としたり、野球においてピッチャーがマウンドよりも手前から投球することでより速いボールを体験したり、バッターが試合で使用するよりも重いバットを振ったりすることがある。しかしながら、参考文献１に記載された従来技術は、あらかじめ撮影した現実の状況をありのまま提示するものであり、実際よりも厳しい環境を体験することはできない。

[0060] 第五実施形態の目的は、上記のような点に鑑みて、現実の環境から取得した情報に基づいて現実よりも厳しい仮想環境を体験することができる仮想環境構築技術を提供することである。

[0061] 上記の課題を解決するために、第五実施形態の第一の態様の仮想環境構築装置は、相手が飛来物を発出する現実環境に基づいてユーザに体験させる仮想環境を構築する仮想環境構築装置であって、飛来物のシーンの事前確率と複数シーンの繋がりを表すダイナミクスとを合わせた確率に基づいてユーザへ提示するシーンの時系列を表す提示シーケンスを生成する提示シーケンス生成部と、提示シーケンスに含まれる各シーンに対応する飛来物の動きを表す提示軌跡を決定する提示軌跡決定部と、提示シーケンスに含まれる各シーンに対応する相手の動きを表す提示動作を決定する提示動作決定部と、提示

軌跡と提示動作とに基づいてユーザに体験させる仮想環境を合成する体験合成部と、を含む。

[0062] 第五実施形態の第二の態様の仮想環境構築装置は、相手が飛来物を発出する現実環境に基づいてユーザに体験させる仮想環境を構築する仮想環境構築装置であって、ユーザへ提示するシーンの時系列を表す提示シーケンスを生成する提示シーケンス生成部と、提示シーケンスに含まれる各シーンに対応する飛来物の動きを表す提示軌跡を決定する提示軌跡決定部と、提示シーケンスに含まれる各シーンに対応して現実環境とは異なる相手の動きを表す提示動作を決定する提示動作決定部と、提示軌跡と提示動作とに基づいてユーザに体験させる仮想環境を合成する体験合成部と、を含む。

[0063] 第五実施形態は、バーチャルな体験によるスポーツのパフォーマンス向上を目指すものである。ここでは、スポーツのある局面で相手の動作とそれに伴うボールの移動が発生し、それに対処する動作をトレーニングするシチュエーションを想定する。例えば、野球においてピッチャーが投げたボールをバッターが打つ、サッカーのペナルティキックにおいてキッカーが蹴ったボールをキーパーが捕球する、テニスにおいてサーバーが打ったボールをレシーバーが返球する、バレーボールにおいてサーバーが打ったボールをレシーバーがレセプションする、などのシチュエーションを想定する。以降、上記のピッチャー、キッカー、サーバーなどボールを発出する人物を「相手選手」または単に「相手」と呼ぶ。また、一球分の映像を「シーン」と呼び、複数シーンの連続を「シーケンス」と呼ぶ。また、ボールの動きの情報を「軌跡」と呼び、相手選手の動きの情報を「動作」と呼ぶ。「軌跡」および「動作」は現実の環境で撮影した映像であってもよいし、主体となる物体もしくは人物の三次元位置情報を時系列に並べたものであってもよい。

[0064] 第五実施形態を適用可能なスポーツは、上述したようなボールを用いる球技に限定されるものではない。例えば、フリスビーのように相手選手から何らかの物体が発出されるスポーツや状況全般に利用可能である。以降、上記のボールを含めて相手選手から発出される物体を「飛来物」と呼ぶ。

[0065] 第五実施形態の仮想環境構築装置は、図 1 2 に示すように、軌跡記憶部 2 1、動作記憶部 2 2、練習意図取得部 2 3、提示シーケンス生成部 2 4、提示軌跡決定部 2 5、提示動作決定部 2 6、体験合成部 2 7、および体験提示部 2 8 を備える。この仮想環境構築装置が図 1 3 に示す各ステップの処理を行うことにより第五実施形態の仮想環境構築方法が実現される。

[0066] 仮想環境構築装置は、例えば、中央演算処理装置（CPU: Central Processing Unit）、主記憶装置（RAM: Random Access Memory）などを有する公知又は専用のコンピュータに特別なプログラムが読み込まれて構成された特別な装置である。仮想環境構築装置は、例えば、中央演算処理装置の制御のもとで各処理を実行する。仮想環境構築装置に入力されたデータや各処理で得られたデータは、例えば、主記憶装置に格納され、主記憶装置に格納されたデータは必要に応じて読み出されて他の処理に利用される。また、仮想環境構築装置の各処理部の少なくとも一部が集積回路等のハードウェアによって構成されていてもよい。仮想環境構築装置が備える各記憶部は、例えば、RAM（Random Access Memory）などの主記憶装置、ハードディスクや光ディスクもしくはフラッシュメモリ（Flash Memory）のような半導体メモリ素子により構成される補助記憶装置、またはリレーショナルデータベースやキーバリューストアなどのミドルウェアにより構成することができる。仮想環境構築装置が備える各記憶部は、それぞれ論理的に分割されていればよく、一つの物理的な記憶装置に記憶されていてもよい。

[0067] 図 1 3 を参照して、第五実施形態の仮想環境構築方法の処理手続きを説明する。

[0068] 軌跡記憶部 2 1 には、動作 I D、相手選手 I D、および軌跡情報を関連付けた軌跡データを蓄積した軌跡データベースが記憶されている。動作 I D は相手選手の動作を一意に特定する識別子である。相手選手 I D は相手選手を一意に特定する識別子である。軌跡情報は相手選手が動作した後の飛来物の軌跡を表す情報であり、ここでは飛来物の三次元位置情報が時系列に並んでいるものとする。

- [0069] 動作記憶部 22 には、動作 I D、相手選手 I D、および動作映像を関連付けた動作データを蓄積した動作データベースが記憶されている。動作映像は奥行きを含めて三次元の動作情報を保持していてもよいし、二次元（すなわち通常の映像）であってもよい。
- [0070] 図 14 に動作データベースの一例を示す。動作データベースには、動作 I D および相手選手 I D に対して時刻 1 ～ T に対応するフレーム画像 1 ～ T を設定することで動作映像を記憶する（図 14 の動作 I D = 1、相手選手 I D = 1 参照）。また、動作 I D および相手選手 I D に対して時刻 1 ～ T における、相手選手の各部位（手や腕など）の三次元座標など、動的に動作映像を生成できる情報を設定しておいてもよい（図 14 の動作 I D = 2、相手選手 I D = 1 参照）。
- [0071] 軌跡データベースに記憶されている軌跡データと動作データベースに記憶されている動作データとは一対一で対応付けがなされており、互いに参照可能であるものとする。すなわち、動作 I D と相手選手 I D の組み合わせにより動作データおよび軌跡データは一意に定まるものとする。
- [0072] ステップ S 21 において、練習意図取得部 23 は、図示しない入力部、例えば、キーボード等から練習意図および練習強度のうち、少なくとも練習意図を含む情報を取得し、出力する。練習意図は、ユーザに体験させたい場面を表現した情報である。練習強度は、練習の難易度を表現した情報である。練習意図および練習強度は、あらかじめ用意された複数の項目を、例えばディスプレイに表示する等の方法によりユーザへ提示し、ユーザのタッチパネル操作やキーボード入力等によりいずれか一つもしくは複数の項目を選択させることで取得する。練習意図および練習強度は都度選択する必要はなく、事前にユーザが設定したものを読み込むことで取得してもよい。以降の処理において練習強度を用いない構成とすることは可能であり、その場合には練習意図のみを取得すればよい。
- [0073] 練習意図は、具体的には、例えば、

1. コース見極め強化、

2. スピード強化、
3. 速度変化強化、
などの項目から選択する。

[0074] 1. コース見極め強化に対しては、さらに、
1-1. 高低、
1-2. 外角のストライク／ボール、
1-3. 内角のストライク／ボール、
1-4. 狙い球／それ以外、
などの選択肢が用意される。

[0075] 2. スピード強化に対しては、さらに、
2-1. ○○km/h～△△km/h
2-2. △△km/h～□□km/h
...

などの選択肢が用意される。

[0076] 3. 速度変化強化に対しては、さらに、
3-1. 2種類（ストレート、カーブ）、
3-2. 3種類（ストレート、カーブ、スライダー）、
3-3. 4種類（ストレート、カーブ、スライダー、チェンジアップ）
、
...

などの選択肢が用意される。

[0077] 練習強度は、例えば、1から6までの各整数で設定するものとし、
1. トレーニングするボールを投げた本物のピッチャー映像で練習する、
2. 本物のピッチャーをスティックピクチャに置き換えた映像で練習する、
3. 複数用意されたテンプレートピッチャーのいずれかの映像で練習する、

4. すべて同じピッチャーの同じ映像で練習する、
5. ピッチャーの手元が見えないように加工した映像で練習する、
6. ピッチャーがいるべき場所に何もいない映像で練習する、

などと定義する。上記の例では練習強度の値が大きくなるほど難易度が高くなる例を挙げたが、必ずしもそのように定義する必要はない。

[0078] なお、練習強度が入力された場合には、練習意図取得部 2 3 は、体験合成部 2 7 へ練習強度を出力する。

[0079] ステップ S 2 2 において、提示シーケンス生成部 2 4 は、練習意図取得部 2 3 から練習意図を受け取り、その練習意図に基づいてユーザへ提示するシーンの時系列を表す提示シーケンスを生成し、出力する。

[0080] 提示シーケンスは、例えばコース、速度、球種などの情報を含むシーン情報が、提示するシーン数だけ並ぶ。ここで、シーン数は、例えば野球であれば球数のことであり、映像のフレーム数ではない。シーン情報は、例えば、以下のように設定される。

[0081] 提示No.1: コース = (1, 3)、速度 = 140km/h、球種 = ストレート

コースは、例えば、図 1 5 に示すように、ストライクゾーンを横方向と縦方向の 3 × 3 の 9 マスに分類し、ストライクゾーンから上下左右に外れたボールゾーンを含めて横方向を x 軸、縦方向を y 軸とする二次元で表現する。図 1 5 は、太枠で囲んだストライクゾーンのバッター側から見て左下隅を原点とした場合のコース分類の例である。例えば、コース = (1, 3) は、ストライクゾーンの 9 マスのうち左から 1 番目、下から 3 番目のマス（右バッターから見て、内角高めのストライクゾーン）を通過する球であることを示す。ボールゾーンを指定する場合には 0 以下または 4 以上の値とすればよい。例えば、コース = (4, 0) は、ストライクゾーンから右に 1 個、下に 1 個外れたマス（右バッターから見て、外角低めのボールゾーン）を通過する球であることを示す。ただし、これに限るものではなく、例えば、図 1 6 に示すように、ホームベースの上方、ストライクゾーンを含む空間を「上」「真ん中左」「真ん中右」「左下」「右下」の 5 つの位置に分類する 5 分類などの別

の分類方法であってもよい。ボールゾーンを指定する場合には、ストライクゾーンから上下左右に外れた空間を、例えば「上外（「上」のさらに上の空間）」「上左外（「上」の左の空間）」「上右外（「上」の右の空間）」「中左外（「真ん中左」のさらに左の空間）」「中右外（「真ん中右」のさらに右の空間）」「下左外（「左下」のさらに左の空間）」「下右外（「右下」のさらに右の空間）」「下外（「左下」および「右下」のさらに下の空間）」などと分類すればよい。

[0082] 球種の分類方法は任意である。例えば、図17に示すように、球種の分類を速度の範囲と変化の大きさとで定義すればよい。図17の例では、速度を横軸、変化を縦軸とするグラフ上に球種の分類が定義されており、グラフ上の距離が近いほど球種が似ていることを表している。一般的には、ストレート、スライダー、チェンジアップ、カーブ、フォークの順に球種が近いと考えられている。球種の類似度を考慮すると、より柔軟な練習が可能となる。例えば、類似度の高い球種は所定の回数以上連続しないように制御したり、類似度が所定値以上離れたボールだけが連続するように制御したりすることができるようになる。

[0083] 分類は、同じクラスに分類されたものの距離は0、そうでないものを1とするものと言える。図17のように分類を定義すると、球種間の類似度を連続的な距離として取り扱えるため、後述するダイナミクスを考える上で柔軟な制御が可能となる。

[0084] 分類や類似度は相手選手（例えば野球であれば投手、テニスであればサーバー）ごとに固有のものを用意してもよい。

[0085] 提示シーケンス生成部24は、事前に求めた事前確率 $p(s)$ およびダイナミクス $p(s_{t+1}|s_t)$ を保持している。ここで、 t はシーンのインデックスを表す番号である。また、 s はあるシーンにおける飛来物の飛来に関するパラメータであり、例えばコース、速度、球種のいずれかである。各パラメータの事前確率は、例えば、様々な投手が投げた様々な球種の投球を記録したデータベースから平均を算出して求めればよい。各パラメータのダイナミクスは、例えば

、各投手が投げた連続する投球における各パラメータの遷移（前の投球と次の投球との間の変化）から求めればよい。

[0086] 提示シーケンスの作成方法を以下に示す。全く同じシーンを連続して視聴すると練習効果が薄まることが考えられるため、提示シーケンスでは練習意図に応じて練習が効果的に進む配球を生成する。提示シーケンス作成部 24 では、提示シーケンスの原型を作成する。例えば、コースは以下に説明する方法で作成し、速度と球種についてはランダムに決めて出力する。

[0087] 図 18 に示すように、まず、事前確率 $p(s)$ からのランダムサンプリングにより、1 球目のコース (x_1, y_1) を決定する。コースはシーン情報で用いたコース分類と同様の定義に基づいて決定する。ここでは、図 15 に示すように、ストライクゾーンを横方向と縦方向の 3×3 の 9 マスに分類したときの、バッター側から見て左下隅を原点とし、横方向を x 軸とし、縦方向を y 軸として、各マスの二次元位置を示す座標値とする。下添え字はシーンのインデックスを表す。次に、コースの確率分布 $p(s)$ とコースの遷移のダイナミクス $p(s_t | s_{t-1})$ とを用いて、 $p(s)p(s_2 | s_1)$ からのランダムサンプリングにより、2 球目のコース (x_2, y_2) を決定する。以降、同様にして、コースの確率分布 $p(s)$ とコースの遷移のダイナミクス $p(s_t | s_{t-1})$ とを用いて、 $p(s)p(s_t | s_{t-1})$ からのランダムサンプリングにより、 t 球目のコース (x_t, y_t) を決定する。

[0088] 速度と球種については、提示軌跡決定部 25 において軌跡データベースを参照して提示軌跡を決定する際に同時に決定するとして、提示シーケンス生成部 24 では決定しないものとしてもよいし、コースと同様に事前確率とダイナミクスに基づいて提示シーケンス生成部 24 で決定するものとしてもよい。

[0089] 提示シーケンス生成部 24 がランダムサンプリングにより飛来物の飛来に関する複数のパラメータ（例えばコース、速度、および球種）のすべてを決定した場合、軌跡データベースを参照して、そのシーンに類似した軌跡データが存在するかを確認してもよい。ランダムサンプリングにより決定したシーンに類似する軌跡データが存在しない場合、飛来物の飛来に関する複数の

パラメータのうちの一部を優先してランダムサンプリングし、類似した軌跡データが存在するように各パラメータを決定してもよい。このとき、各パラメータのいずれを優先するかは練習意図によりあらかじめ定めておく。このように構成することにより、実際の投球に近いシーンを生成することができ、体験のリアリティを高めることができる。

[0090] 提示シーケンス生成部24が練習意図取得部23から受け取った練習意図に基づいて、事前確率 $p(s)$ とダイナミクス $p(s_{t+1}|s_t)$ とを制御することも可能である。これにより、例えば、1. コース見極め練習、2. スピード変化強化練習、3. 球種変化強化練習を実現することができる。以下、それぞれ詳細に説明する。

[0091] 1. コース見極め練習は、例えば、外角のストライク／ボールを見極めるための練習である。例えば、図19に示すように、事前確率 $p(s)$ の確率分布をストライク（手を出して良いコース）とボール（手を出してはいけないコース）の二峰性とする。ダイナミクス $p(s_{t+1}|s_t)$ は、まずコースのみを決定し、 $p(\text{球種, 速度}|\text{コース})$ に基づくランダムサンプリングにより、球種および速度を決定する。

[0092] 2. スピード変化強化練習は、ダイナミクス $p(s_{t+1}|s_t)$ を速度で規定し、 $p(s_{t+1}|s_t)$ として速度を決定した後で、 $p(\text{コース, 球種}|\text{速度})$ に基づくランダムサンプリングにより、コースおよび球種を決定する。

[0093] 3. 球種変化強化練習は、ダイナミクス $p(s_{t+1}|s_t)$ を球種で規定し、 $p(s_{t+1}|s_t)$ として球種を決定した後で、 $p(\text{コース, 速度}|\text{球種})$ に基づくランダムサンプリングにより、コースおよび速度を決定する。ただし、一般的に速度と球種は相関が高いため、実効的にはスピード変化強化練習と類似度が高くなる。

[0094] 事前確率分布とダイナミクスは、実際に投げられたボールに基づいて決定してもよい。例えば、特定の選手のボールの平均から事前確率分布を求め、当該選手のボールの遷移からダイナミクスを決めてもよいし、すべての投手のボールの平均から事前確率分布を求め、当該選手のボールの遷移からダイナミクスを決めてもよい。

[0095] 提示シーケンス生成部24は、上記のようにして各シーン情報を生成し、これをシーン数（球数）だけ繰り返すことで、提示シーケンスを生成する。生成した提示シーケンスは提示軌跡決定部25へ送られる。

[0096] ステップS23において、提示軌跡決定部25は、提示シーケンス生成部24から提示シーケンスを受け取り、その提示シーケンスに含まれる各シーンについて、軌跡データベースから最も類似した軌跡データを選択することにより、飛来物の動きを表す提示軌跡を決定し、決定した提示軌跡を提示シーケンスと対応づけて出力する。提示軌跡DB_Iは、以下の式を用いて選択する。

[0097] [数3]

$$DB_I = \underset{i}{\operatorname{argmin}} E(\text{コース} q - \text{コース} DB_i) \\ + E(\text{球種} q - \text{球種} DB_i) + E(\text{速度} q - \text{速度} DB_i)$$

[0098] ここで、qは提示シーケンスに含まれるシーン情報を表し、DB_iは軌跡データベースに含まれる軌跡データを表す。すなわち、「コースq」はあるシーンのコースであり、「コースDB_i」はi番目の軌跡データにおけるコースである。提示軌跡決定部25は、提示軌跡DB_Iと提示シーケンスとが対応付けられた情報（以下、軌跡付与後提示シーケンスと呼ぶ）を出力する。

[0099] ステップS24において、提示動作決定部26は、提示軌跡決定部25から軌跡付与後提示シーケンスを受け取り、その軌跡付与後提示シーケンスに含まれる各シーンについて、動作データベースから動作データを選択することにより、相手選手の動きを表す提示動作を決定する。動作データベースに記憶されている動作データと、軌跡データベースに記憶されている軌跡データは一対一で対応付けされているため、提示軌跡DB_Iに対応付けられている動作IDおよび相手選手IDにより特定される動作データを提示動作として選択する。具体的には、提示動作決定部26は、提示軌跡DB_Iと提示シーケンスとが対応付けられた情報を入力とし、各DB_Iに対応する動作情報（動作ID、相手選手ID、および動作映像）を動作データベースから取り出して軌跡付与後提示シーケンスと対応付け、動作データを対応づけた軌跡付与後

提示シーケンスを軌跡動作付与後提示シーケンスとして出力する。軌跡動作付与後提示シーケンスは、体験合成部 27 へ送られる。

[0100] ステップ S 25 において、体験合成部 27 は、提示動作決定部 26 から軌跡動作付与後提示シーケンスを受け取り、その軌跡動作付与後提示シーケンスに含まれる提示軌跡と提示動作とに基づいてユーザに体験させる仮想環境を構築し、出力する。

[0101] 仮想環境の構築方法は、公知のいかなる技術を用いてもよいが、例えば、以下のように行うことができる。まず、ボールやスタジアムのように仮想環境を構成するためのオブジェクトを構成素材と呼び、各構成素材を仮想環境に再現するためのコンピュータ・グラフィックス (CG: Computer Graphics) データを仮想環境素材として予め記憶しておく。仮想環境素材は CG データに限定されず、実際に撮影した映像から切り出した動画または静止画データであってもよい。また、構成素材に関する位置姿勢情報も予め記憶しておく。このとき、位置姿勢情報は提示軌跡と同じ座標系における位置情報とその座標系における姿勢情報とを含む情報とする。体験合成部 27 は、仮想環境に配置される仮想環境素材を読み出し、時間変化をしない静的オブジェクトの仮想環境素材を、その位置姿勢情報に従って仮想環境に配置する。同時に、相手選手および飛来物などの時間変化を伴う動的オブジェクトの仮想環境素材は、軌跡動作付与後提示シーケンスに対応付けられた提示軌跡および提示動作に基づいて仮想環境に配置する。例えば、提示軌跡が三次元位置情報の時系列であれば、飛来物に対応する仮想環境素材を読み込んで、仮想環境上の三次元位置に任意の姿勢で配置する。また、提示動作が現実環境の映像から切り出した映像であれば、仮想環境上の所定の位置に相手選手の映像を合成する。このとき、モーションセンサー等を用いてユーザの頭部の位置および姿勢を計測し、そのユーザの位置姿勢情報に基づいて仮想環境内の視線を制御するように構成することも可能である。

[0102] 体験合成部 27 では、練習意図取得部 23 から練習強度を入力として受け取り、受け取った練習強度に基づいて、映像を変化させてもよい。練習強度

では、練習の難易度を決定する。ただし、練習強度の数が大きくなると単純に難しくなるものとは限らない。

[0103] 例えば、練習強度を1から6の数字で分類し、各練習強度を以下のように定義する。

1. 変化を加えない。
2. 投手映像をスティックピクチャで表示する。人間をスティックピクチャで提示すると、動作がゆっくりと見えるとの知見がある（参考文献2参照）。一方、実写映像で提示する場合と比較して臨場感が低下することになる。

[0104] 〔参考文献2〕松本鮎美，三上弾，ウ小軍，川村春美，小島明，“多視点階層型GPDMを用いた単眼カメラからの視点非依存3次元人物姿勢推定”，電子情報通信学会論文誌，Vol. J97-D，No. 7，Jul. 2014.

[0105] 3. 映像をランダムに入れ替える。この場合、ボールが投げられる前の事前の動作に影響を受けないようにする練習になる。

4. すべて同じ映像を提示する。練習強度3と類似した効果が生じる。
5. 画像処理によって一部（例えばボールの出元）を見えにくくする。これにより選手は自身の手元により近い位置でボールを判断する必要性が生じる。
6. 映像を提示しない。

[0106] 具体的には、

1. トレーニングする球を投げた本物のピッチャー映像でトレーニング
2. 本物のピッチャーをスティックピクチャに置き換えてトレーニング
3. 複数用意されたテンプレートピッチャーのいずれかの映像でトレーニング
4. どの球もすべて同じピッチャーの同じ映像でトレーニング
5. ピッチャーの手元が見えないように加工した映像でトレーニング
6. ピッチャーがいるべき場所に何も無い(球だけ)の映像でトレーニングとする。

[0107] 例えば、スティックピクチャで提示することを考える。この場合、動作データベースに登録された各動作について姿勢推定を行う。単眼画像からの姿

勢推定は、参考文献2に記載された手法などにより実現可能である。参考文献2によれば、スティックピクチャで提示することによって、動きがゆっくり感じられることが指摘されている。これは、人の動きが、棒のみで単純化されたことにより認知的な負荷が軽減されたためと推測されている。これは、ある側面では練習強度が下がっていると言える。一方、映像からスティックピクチャになることで、投手の表情や細かな手の動きなどが視聴できなくなる。これは、より削減された情報で対峙することになるので練習が難しくなる。

[0108] 練習強度3、4、5、6のコンセプトは同じである。実際に投じる、あるいは、サーブを打つ前の動きに、投じるボールに関する情報がある可能性がある。例えば、フォームで球種が分かる場合がある。しかしながら、投手は全く同じフォームで多くの球種を投げることができるように練習しているため、事前のフォームに左右されることなく、ボールの軌道だけで判断し反応できる様になることが望ましい。そのために、ランダムなフォームから投げられるボール、あるいは全く同じフォームから投げられるボール、さらにはフォームが全く視聴できないように映像を全く提示せずに突然飛び出してくるボール、などで練習することが有効となる。

[0109] 練習強度3の場合、動作データベースに記憶されている当該投手の動作データから一つの動作をランダムに選択し提示動作とする。練習強度4の場合、動作データベースに記憶されている当該投手の動作データから一つの動作をランダムに選択し、その後の練習の間は常に同一の動作データを選択する。そのため、練習強度4が設定された場合、同一投手の映像を視聴している間常に同一の映像を選択する。練習強度5の場合、いくつかの方法が検討できる。第一の方法は、投手を遮蔽するオブジェクトを仮想的に出現する方法である。第二の方法は、あらかじめ画像処理により動作データベースに登録されている投手動作について、ブラーをかけるなどの処理を行っておく方法である。

[0110] ステップS26において、体験提示部28は、体験合成部27から構築し

た仮想環境のデータを受け取り、ユーザが視聴できる形態として出力する。出力する映像は、仮想環境のある視点から一方向を見るような映像でもよいし、全天球映像のように全方位を見ることができ映像でもよい。例えば、ある視点から一方向を見るような映像を出力する場合は、ある位置に仮想的に透視投影カメラを設置し、仮想環境の空間中に存在するCGモデルを透視投影カメラの画像平面に投影すればよい。また、ある視点から全方位を見渡せる全天球映像を出力する場合は、ある視点に仮想的に全天球カメラを設置し、仮想環境の空間中に存在するCGモデルを全天球カメラの画像平面に投影すればよい。

[0111] 上記のように構成することにより、第五実施形態の仮想環境構築技術によれば、相手からの飛来物があるスポーツや状況のトレーニングにおいて、練習意図に基づいて現実の環境よりも厳しい環境をユーザへ体験させるための仮想環境を構築することができる。これにより、現実には厳しい環境を用意させなくとも練習意図に沿った環境を容易にユーザへ体験させることができ、練習の効率を向上することができる。

[0112] 第五実施形態の仮想環境構築技術によれば、現実の環境から取得した情報に基づいて現実よりも厳しい環境を体験することができる仮想環境を構築することが可能となる。これにより、選手は事前に現実環境よりも厳しい環境を体験することができ、試合でのパフォーマンスを向上することができる。

[0113] [第六実施形態]

スポーツにおいて、試合の前に実際に試合を行っている選手の視点位置からの映像を高い臨場感で体験しておくことは、実戦でのパフォーマンス向上にとって非常に重要である。しかしながら、試合中に選手のいる場所からビデオを撮影することは競技の妨げとなり困難である。

[0114] 上記の参考文献1には、スポーツ等のトレーニングにおいて、試合と類似の環境をバーチャルに生成して選手が事前に経験することで、試合でのパフォーマンス向上を目指すシステムが記載されている。参考文献1に記載されたシステムは、フットボールの選手を指導するにあたって、仮想現実（VR：V

irtual Reality) 技術を用いて試合と同じ環境を再現した選手目線からの360度映像を、ヘッドマウントディスプレイで視聴させるシステムである。

[0115] 従来のバーチャルなトレーニング方法では、予め用意された映像からコーチや選手などが手動で選択したシーンの映像を視聴するものであった。しかしながら、効率的なトレーニングを行うためには、選手の動きに応じて重点的に練習するシーンを決定することが重要である。

[0116] 第六実施形態の目的は、上述のような点に鑑みて、ユーザの反応に基づいてユーザに提示するシーンを決定することができる映像提示技術を提供することである。

[0117] 上記の課題を解決するために、第六実施形態の映像提示装置は、相手が飛来物を発出する映像をユーザに提示する映像提示装置であって、ユーザの身体的な挙動を表すユーザ状態と相手が飛来物を発出するシーンを表すシーン情報とを関連付けて記憶するユーザ状態記憶部と、ユーザに提示されたシーンに対する反応として現れるユーザ状態であるユーザ反応を取得するユーザ状態取得部と、ユーザ状態記憶部からユーザ反応に関連付けられたシーン情報である提示シーン情報を取得する提示マネジメント部と、提示シーン情報に基づいてユーザに提示する映像を生成する映像提示部と、を含む。

[0118] 第六実施形態は、バーチャルな体験によるスポーツのパフォーマンス向上を目指すものである。ここでは、スポーツのある局面で相手の動作とそれに伴うボールの移動が発生し、それに対処する動作をトレーニングするシチュエーションを想定する。例えば、野球においてピッチャーが投げたボールをバッターが打つ、サッカーのペナルティキックにおいてキッカーが蹴ったボールをキーパーが捕球する、テニスにおいてサーバーが打ったボールをレシーバーが返球する、バレーボールにおいてサーバーが打ったボールをレシーバーがレセプションする、などのシチュエーションを想定する。以降、上記のピッチャー、キッカー、サーバーなどボールを発出する人物を「相手選手」または単に「相手」と呼ぶ。また、一球分の映像を「シーン」と呼ぶ。「シーン」は現実の環境で撮影した映像であってもよいし、主体となる物体も

しくは人物の三次元位置情報を時系列に並べた情報に基づいてコンピュータ・グラフィックス（CG: Computer Graphics）で作成した仮想的な映像であってもよい。もしくは、実写映像とCG映像を組み合わせで合成した仮想環境上の映像であってもよい。

[0119] 第六実施形態を適用可能なスポーツは、上述したようなボールを用いる球技に限定されるものではない。例えば、フリスビーのように相手から何らかの物体が発出されるスポーツや状況全般に適用可能である。以降、上記のボールやフリスビーなどの相手から発出される物体を総称して「飛来物」と呼ぶ。

[0120] 第六実施形態の映像提示装置は、例えば、野球の場合、ユーザがキーボード等の入力装置から次のシーンを要求するための信号を入力すると、ピッチャーが一球投球するシーンをユーザへ提示する情報処理装置である。テニスの場合でも同様に、ユーザがキーボード等の入力装置から次のシーンを要求する信号を入力すると、サーバーが一球サーブするシーンをユーザへ提示する。ユーザはヘッドマウンドディスプレイを装着して、映像提示装置から提示されるシーンを視聴する。ヘッドマウンドディスプレイに提示される映像は、ユーザの頭部の位置および姿勢などのユーザの身体的な状態に合わせて決定される。

[0121] 第六実施形態の映像提示装置は、図20に示すように、ヘッドマウンドディスプレイ31、ユーザ状態取得部32、ユーザ状態記憶部33、提示マネジメント部34、映像提示部35、およびシーン情報記憶部36を備える。映像提示装置は、ユーザ評価取得部37およびユーザ状態提示部38をさらに備えていてもよい。この映像提示装置が、図21に示す各ステップの処理を行うことにより第六実施形態の映像提示方法が実現される。

[0122] 映像提示装置は、例えば、中央演算処理装置（CPU: Central Processing Unit）、主記憶装置（RAM: Random Access Memory）などを有する公知または専用のコンピュータに特別なプログラムが読み込まれて構成された特別な装置である。映像提示装置は、例えば、中央演算処理装置の制御のもとで各処

理を実行する。映像提示装置に入力されたデータや各処理で得られたデータは、例えば、主記憶装置に格納され、主記憶装置に格納されたデータは必要に応じて読み出されて他の処理に利用される。また、映像提示装置の各処理部の少なくとも一部が集積回路等のハードウェアによって構成されていてもよい。映像提示装置が備える各記憶部は、例えば、RAM (Random Access Memory) などの主記憶装置、ハードディスクや光ディスクもしくはフラッシュメモリ (Flash Memory) のような半導体メモリ素子により構成される補助記憶装置、またはリレーショナルデータベースやキーバリューストアなどのミドルウェアにより構成することができる。映像提示装置が備える各記憶部は、それぞれ論理的に分割されていればよく、一つの物理的な記憶装置に記憶されていてもよい。

[0123] ヘッドマウントディスプレイ 31 は、ユーザが頭部に装着して映像や音を視聴するためのデバイスである。ヘッドマウントディスプレイ 31 は、例えば、ユーザが装着したときの視界に対応する範囲にディスプレイが設けられ、両耳に対応する位置にヘッドセットが設けられる。ヘッドマウントディスプレイ 31 には、加速度センサや変位センサなどの各種センサが搭載されており、ユーザが装着したときの頭部の位置および姿勢を認識することが可能である。これにより、ヘッドマウントディスプレイ 31 は、ユーザの視点に合わせた映像を提示することが可能となっている。

[0124] ユーザ状態記憶部 33 には、ユーザ状態データベースが記憶されている。ユーザ状態データベースは、ユーザの身体的な挙動を表すユーザ状態と、ユーザに提示するシーンを表すシーン情報とを関連付けて蓄積したデータベースである。ユーザ状態データベースは、具体的には、次表に示すように、視聴 ID、シーン ID、視聴回数、およびユーザ状態からなるテーブルである。

[0125]

[表1]

視聴 I D	シーン I D	視聴回数	ユーザ状態
1	2	1	a
2	1	1	b
3	2	2	c
4	2	3	d
5	1	2	e
...	...	...	...

[0126] 視聴 I D は、一つのシーンを一度視聴するたびに付与されるユニークな通し番号である。シーン I D は、シーンごとに付与された、シーンを特定するための番号である。視聴回数は、そのシーンについて何度目の視聴かを表す数値である。ユーザ状態は、そのシーンを視聴したときのユーザの身体的な挙動を表す時系列情報である。

[0127] ユーザ状態について、より詳細に説明する。ユーザ状態は、例えば、頭部の位置および姿勢、全身姿勢、体重移動などである。頭部の位置および姿勢は、ヘッドマウントディスプレイ 31 から取得することができる。全身姿勢は、例えば、VICON 社製品のようなモーションキャプチャシステムを用いて取得することができる（参考文献 3 参照）。体重移動は、フォースプレートなどのバランス評価システムを用いて取得することができる。上述のように、ヘッドマウントディスプレイ 31 が映像を提示するためには頭部の位置および姿勢は必須の情報であるため、ユーザ状態は少なくとも頭部の位置および姿勢を含む。全身姿勢や体重移動などのその他の身体的な挙動は、より詳細なユーザの状態を取得するために任意に追加することができる。

[0128] [参考文献 3] Vicon Motion Systems Ltd., “Motion Capture Systems | VICON”、[online]、[平成 28 年 10 月 20 日検索]、インターネット<URL: <https://www.vicon.com/>>

[0129] シーン情報記憶部 36 には、シーン情報データベースが記憶されている。シーン情報データベースは、ユーザに提示するシーンの詳細を表すシーン情報を蓄積したデータベースである。シーン情報データベースは、具体的には

、次表に示すように、シーンID、球種、球速、軌道、キーバインド、および練習確率からなるテーブルである。

[0130] [表2]

シーンID	球種	球速	軌道	キーバインド	練習確率
1	ストレート	150km/h	内角高め	“z”	0.1
2	カーブ	110km/h	外角低め	“x”	0.08
3	ストレート	135km/h	真ん中低め	“c”	0.03
...	...	...	...	...	...

[0131] シーンIDは、シーンごとに付与された、シーンを特定するための番号である。シーンIDは、ユーザ状態データベースのシーンIDとリンクしている。球種は、そのシーンにおいて投げられるボールの球種である。球速は、そのシーンにおいて投げられるボールの球速である。軌道は、そのシーンにおいて投げられるボールのコースである。キーバインドは、後述する通常モードにおいて、ユーザが押下するキーの種類である。練習確率は、後述するフィードバックモードにおいて、提示するシーンを選択するときに用いる確率値である。練習確率は、データベース全体での総和が1となるように設定され、初期状態では均等となるように初期化されている。

[0132] 以下、図21を参照しながら、第六実施形態の映像提示装置が実行する映像提示方法の処理手続きを説明する。

[0133] ステップS31において、映像提示装置のユーザ状態取得部32は、ユーザの身体的な挙動を表すユーザ状態を取得して出力する。同時に、ユーザ状態取得部32は、直前に提示したシーンを表すシーン情報と、そのシーンをユーザが視聴している際に取得したユーザ状態を関連付けて、ユーザ状態記憶部33に記憶されたユーザ状態データベースに蓄積する。以下、直前に提示したシーンの視聴中にユーザが反応として示したユーザ状態をユーザ反応と呼ぶ。取得したユーザ反応は提示マネジメント部34へ送られる。

[0134] ユーザ状態データベースにおいて、あるシーンに関連付いたユーザ状態が複数存在する場合、そのシーンに対するユーザ状態が安定していない（すなわち、そのシーンに対する練習が不足している）ことを意味するため、その

シーンがより多く提示されるように、シーン情報データベースの練習確率を上昇させる。練習確率を上昇させる方法はどのようなものでもよい。一例としては、練習確率を現在の値よりも予め定めた割合 a だけ上昇させる。練習確率を上昇させるシーン情報の現在の練習確率が x だとすると、上昇後の練習確率は ax となる。このとき、すべてのシーンの練習確率の合計は $1+ax-x$ となるため、すべてのシーンの練習確率を $1/(1+ax-x)$ 倍して合計が1となるようにする。なお、 a は例えば1.1程度とする。

[0135] ステップS33において、映像提示装置の提示マネジメント部34は、ユーザに提示するシーンを表すシーン情報（以下、提示シーン情報と呼ぶ）を決定して出力する。提示マネジメント部34が提示シーン情報を決定する方法は、通常モードの場合とフィードバックモードの場合とで異なる。なお、モードの設定は、ユーザがキーボード等の入力装置を手動で操作することにより行う。

[0136] 通常モードでは、ユーザがキーボード等の入力装置から次のシーンを要求するためのキー入力を行うと、提示マネジメント部34が入力されたキーに予め対応付けられたシーン情報をシーン情報記憶部36に記憶されたシーン情報データベースから取得して出力する。例えば、表2に示したシーン情報がシーン情報データベースに記憶されているとして、ユーザが“z”キーを入力した場合、時速150kmのストレートが内角高めに投じられるシーンを表すシーン情報が出力される。このとき、一つのキーに対して複数のシーンを設定し、一回のキー入力に対して複数球のシーンの時系列を提示するように設定してもよい。

[0137] フィードバックモードでは、提示マネジメント部34は、ユーザ状態取得部32が取得したユーザ反応を入力とし、そのユーザ反応に関連づいたシーン情報をユーザ状態記憶部33に記憶されたユーザ状態データベースから取得して出力する。例えば、表1に示したユーザ状態がユーザ状態データベースに記憶されており、表2に示したシーン情報がシーン情報データベースに記憶されているとして、“c”というユーザ反応が入力された場合、時速110k

mのカーブが外角低めに投げられるシーンがユーザに提示される。

[0138] ユーザ情報データベースにおいて、あるユーザ状態に対して複数のシーンが関連付けられている場合、シーン情報データベースの練習確率に基づいて、提示シーン情報を選択する。このとき、シーンの繋がりを考慮して、直前に提示された複数シーンを合わせてシーンの出現確率を変えるように設定してもよい。例えば、時速140kmのボールは、時速160kmのボールを見た後には遅く感じられるが、時速100kmのボールを見た直後では速く感じられる。このような連続性を考慮して複数のシーンをパッケージとして提示することで、より効率的なトレーニングを行うことが可能となる。

[0139] ステップS34において、映像提示装置の映像提示部35は、提示マネジメント部34が取得した提示シーン情報に基づいて、その提示シーン情報に対応するシーンの映像を生成して、ヘッドマウントディスプレイ31へ出力する。これにより、ユーザはヘッドマウントディスプレイ31を介して映像提示装置が提示するシーンを視聴することができる。提示する映像は、実際にピッチャーが投球するシーンをバッターボックスの位置から撮影した実写映像であってもよいし、以下のようにしてシーン情報に基づいて構成素材を合成した仮想環境の映像であってもよい。まず、ボールやスタジアムのように仮想環境を構成するためのオブジェクトを構成素材と呼び、各構成素材を仮想環境に再現するためのCGデータを仮想環境素材として予め記憶しておく。仮想環境素材はCGデータに限定されず、実際に撮影した映像から切り出した動画または静止画データであってもよい。また、構成素材に関する位置姿勢情報も予め記憶しておく。映像提示部35は、仮想環境に配置される仮想環境素材を読み出し、時間変化をしない静的オブジェクトの仮想環境素材を、その位置姿勢情報に従って仮想環境に配置する。同時に、相手選手および飛来物などの時間変化を伴う動的オブジェクトの仮想環境素材は、シーン情報で指定された球種・球速・軌道に基づいて仮想環境に配置する。このとき、ユーザ状態取得部32で取得した頭部の位置および姿勢に基づいて仮想環境内の視線を制御するように構成することも可能である。

[0140] <第六実施形態の変形例1>

変形例1では、各シーンに対するユーザの自己評価を取得して、提示するシーンの選択に用いることで、ユーザが苦手と考えるシーンをより重点的に練習できるように構成する。変形例1の映像提示装置は、実施形態の映像提示装置が備える各処理部に加えて、ユーザ評価取得部37をさらに備えるものとする。

[0141] ステップS32において、映像提示装置のユーザ評価取得部37は、ユーザがキーボード等の入力装置から入力する直前のシーンに対する自己評価を取得する。自己評価の取り得る値は、例えば、うまくできた／できなかった、のような2段階評価でもよいし、1から5までの各整数を用いて値が大きいほどうまくできたことを表す5段階評価としてもよい。ユーザ評価取得部37は、取得した自己評価に基づいて、シーン情報データベースの練習確率を調整する。例えば、自己評価の低いシーンについては練習確率を上昇させ、自己評価の高いシーンについては練習確率を低下させる。これにより、自己評価が低くユーザが苦手と考えているシーンは提示される頻度が高くなり、自己評価が高くユーザが得意と考えているシーンは提示される頻度が低くなる。

[0142] ユーザ評価取得部37は、ユーザ状態取得部32が取得したユーザ状態に基づいて評価値を推定することで、ユーザからの自己評価の入力を不要とすることもできる。例えば、予めユーザが手動で自己評価を与えたユーザ状態を学習データとして分類器を学習しておき、その分類器を用いて入力されたユーザ状態を分類して評価値を推定する。この場合、学習が完了する程度の自己評価が収集できればよいので、ユーザによる自己評価の入力は一定時間で終わることができる。この一定時間は、取得したユーザ状態を学習したときの誤差がしきい値以下に収束するまでとしてもよいし、最初の30球と決めておいてもよい。ユーザ状態は時系列の遷移であるため、リカレントニューラルネットワークなどの時系列情報を扱うことができる学習器を用いてもよいし、一定数をリサンプリングしてサポートベクターマシン（SVM: Support

t Vector Machine)などで学習してもよい。例えば、捕球前0.4秒の間のユーザ状態を30fpsでリサンプリングを行い(つまり、12サンプル)、a次元のユーザ状態(例えば、頭部位置姿勢なら6次元)を入力として、SVMを用いて良／悪の二値分類することが考えられる。これにより、ユーザからの自己評価の入力がなくても、ユーザ状態から自己評価を推定することが可能となる。

[0143] <第六実施形態の変形例2>

変形例2では、取得したユーザ状態を提示して、ユーザが自らの状態を確認できるように構成する。変形例2の映像提示装置は、実施形態の映像提示装置が備える各処理部に加えて、ユーザ状態提示部38をさらに備えるものとする。

[0144] ステップS35において、映像提示装置のユーザ状態提示部38は、ユーザ状態取得部32が取得したユーザ状態、もしくは、ユーザ状態記憶部33に蓄積されたユーザ状態を、ヘッドマウントディスプレイ31、もしくは、映像提示装置に接続した液晶ディスプレイ等の外部の表示部に出力する。ユーザ状態の表示方法はどのようなものでもよい。例えば、1回分のユーザ状態の時間遷移を時系列グラフで提示する(方法1)、予め定められたN回分のユーザ状態の時間遷移を時系列グラフで提示する(方法2)、予め定められたN回分のユーザ状態の時間遷移を平均と分散により提示する(方法3)、といった方法が挙げられる。

[0145] 図22(a)は上記方法1によりユーザ状態(頭部姿勢)の推移をグラフ化した例である。縦軸は絶対座標値であり、横軸は時間である。図22では、「リリース」で示す時刻に相手選手が飛来物を発出したことを表している。このような可視化により、例えば野球であれば打席内での挙動を可視化して、他の選手との比較を行うことが可能となる。上記方法2のように複数回分を提示する場合、図22(a)のようなグラフが複数個連続することになる。図22(b)は上記方法3により複数回分のユーザ状態(頭部姿勢)の推移をグラフ化した例である。同じシーンに対する複数回分のユーザ状態を収集し、それらの平均と分散をグラフ化したものである。例えば、分散が大

きい場合は、挙動にぶれが大きく、安定していない、というような分析をすることが可能となる。

[0146] 図23および図24は、ユーザ状態を頭部の位置および姿勢としたときの、詳細なユーザ状態の提示例である。頭部の位置はx軸、y軸、z軸の座標値それぞれについてグラフ化し（図23）、頭部の姿勢はx軸、y軸、z軸の回転角それぞれについてグラフ化する（図24）。これら6次元のユーザ状態を、異なる二人の選手（選手Aおよび選手B）について並べることで、比較および分析することが可能となる。

[0147] 第六実施形態の映像提示技術は、第一実施形態から第五実施形態までの仮想環境構築技術と組み合わせることが可能である。

[0148] 第六実施形態によれば、あるシーンに対するユーザの反応に基づいてユーザに提示するシーンを決定することができるため、効率的なトレーニングを行うことが可能となる。

[0149] [第七実施形態]

スポーツの試合において、事前に体験していない初めての状況に遭遇すると、十分なパフォーマンスを発揮できないことがある。野球では「初物に弱い」という表現を用いて初対戦の投手を苦手とする状況が語られることがある。この状況に対し、実際の試合と同じ環境において選手の目線で事前に投手の投球を体験することが有効であると考えられる。

[0150] しかしながら、次の対戦相手の投球を完全な形で事前体験することは事実上困難である。次の対戦相手の投球を事前体験するために、スタンドなどから撮影した対戦相手の過去の投球映像を視聴する方法が知られているが、この方法において視聴される映像は実際にバッターボックスに立っている選手が見ている風景とは視点や視野が異なり、事前体験として十分ではない。

[0151] 上記の非特許文献1には、選手の目線における事前体験の実現方法としてバッターボックス内の選手の目線の位置に360度のパノラマ画像を撮影できるカメラ（以下、全天球カメラと呼ぶ）を設置し、全天球カメラで撮影した選手目線の映像をヘッドマウントディスプレイにより視聴する方法が開示

されている。

[0152] しかしながら、試合中にバッターボックス内に全天球カメラを設置することは事実上困難であるから、非特許文献 1 の方法は実現が難しい。

[0153] 非特許文献 1 以外の事前体験の実現方法として、例えば実環境を全て C G で再現した仮想環境を生成することも考えられる。しかしながら、プレー中の選手の C G を正確に作るためには実際にプレーしている選手に高精度なセンサを装着させる必要があり、実環境を全て C G で再現することはかなり高コストになる。一方、簡易な C G データで選手を再現することも考えられるが、再現された選手の C G データは不自然な動きとなり臨場感に欠けるため、事前体験として十分ではない。

[0154] そこで第七実施形態では臨場感の高い仮想環境を簡易に生成できる仮想環境構築装置を提供することを目的とする。

[0155] 位置が遷移する動く素材である動的素材と、位置が遷移せず動かない素材である静的素材と、位置が遷移せず動く素材である特殊動的素材が実環境に存在するものとする。第七実施形態の仮想環境構築装置は、このような実環境を再現する仮想環境を生成する仮想環境構築装置である。第七実施形態の第一の態様の仮想環境構築装置は、仮想素材取得部と、仮想動的素材位置姿勢検出部と、仮想特殊動的素材生成部と、同期部と、仮想環境生成部を含む。

[0156] 仮想素材取得部は、動的素材である物体を仮想環境上に再現するための仮想動的素材と、静的素材を仮想環境上に再現するための仮想静的素材とを取得する。仮想動的素材位置姿勢検出部は、仮想動的素材の位置および姿勢を検出してその時系列データを取得する。仮想特殊動的素材生成部は、実環境を撮影した映像に基づいて特殊動的素材である人間を仮想環境上に再現するための仮想特殊動的素材を生成する。同期部は、人間と物体とが離隔する瞬間の前後を高フレームレートで撮影した映像と、人間と物体とが離隔する瞬間の時刻とのペアを学習したモデルへ、人間と物体とが離隔する前を撮影した離隔前画像もしくは人間と物体とが離隔した後を撮影した離隔後画像の少

なくともいずれかを入力することで人間と物体とが離隔する瞬間の時刻を推定し、推定された時刻に基づいて仮想動的素材と仮想特殊動的素材とを同期する。仮想環境生成部は、仮想静的素材と、同期された仮想動的素材と、同期された仮想特殊動的素材に基づいて仮想環境を生成する。

[0157] 第七実施形態の第二の態様の仮想環境構築装置は、仮想素材取得部と、仮想動的素材位置姿勢検出部と、仮想特殊動的素材生成部と、同期部と、仮想環境生成部を含む。仮想素材取得部は、動的素材を仮想環境上に再現するための仮想動的素材と、静的素材を仮想環境上に再現するための仮想静的素材とを取得する。仮想動的素材位置姿勢検出部は、仮想動的素材の位置および姿勢を検出してその時系列データを取得する。仮想特殊動的素材生成部は、実環境を撮影した映像に基づいて特殊動的素材を仮想環境上に再現するための仮想特殊動的素材を生成する。同期部は、特殊動的素材と動的素材とが離隔する瞬間の前後を高フレームレートで撮影した映像と、特殊動的素材と動的素材とが離隔する瞬間の時刻とのペアを学習したモデルへ、特殊動的素材と動的素材とが離隔する前を撮影した離隔前画像もしくは特殊動的素材と動的素材とが離隔した後を撮影した離隔後画像の少なくともいずれかを入力することで特殊動的素材と動的素材とが離隔する瞬間の時刻を推定し、推定された時刻に基づいて仮想動的素材と仮想特殊動的素材とを同期する。仮想環境生成部は、仮想静的素材と、同期された仮想動的素材と、同期された仮想特殊動的素材に基づいて仮想環境を生成する。

[0158] 第七実施形態の第三の態様のモデル学習装置は、人間と物体とが離隔する瞬間の前後を高フレームレートで撮影した映像と、人間と物体とが離隔する瞬間の時刻とのペアを用いて、人間と物体とが離隔する前を撮影した離隔前画像もしくは人間と物体とが離隔した後を撮影した離隔後画像との少なくともいずれかを入力として人間と物体とが離隔する瞬間の時刻を出力するモデルを学習する。

[0159] 第七実施形態の第四の態様のモデル学習装置は、位置が遷移せず動く素材である特殊動的素材と位置が遷移する動く素材である動的素材とが離隔する

瞬間の前後を高フレームレートで撮影した映像と、特殊動的素材と動的素材とが離隔する瞬間の時刻とのペアを用いて、特殊動的素材と動的素材とが離隔する前を撮影した離隔前画像もしくは特殊動的素材と動的素材とが離隔した後を撮影した離隔後画像との少なくともいずれかを入力として特殊動的素材と動的素材とが離隔する瞬間の時刻を出力するモデルを学習する。

[0160] 本実施形態ではピッチャーやボールやスタジアムのように現実存在する人、物を素材と呼ぶ。素材のうちその位置が（時間経過により）遷移し、かつ動く素材（例えば、ボール）を動的素材、その位置が遷移せず動かない素材（例えば、スタジアム、スコアボード、観客席）を静的素材、その位置は遷移しないものとみなされるがその位置に留まったまま動く素材（例えば、野球のピッチャーやサーブを打つテニスプレイヤー）を特殊動的素材と呼ぶ。

[0161] 素材を仮想環境に再現するためのコンピュータ・グラフィクスデータ（CGデータ）を仮想素材と呼ぶ。仮想素材のうち、動的素材を仮想環境上に再現するための仮想素材を仮想動的素材、静的素材を仮想環境上に再現するための仮想素材を仮想静的素材、特殊動的素材を仮想環境上に再現するための仮想素材を仮想特殊動的素材と呼ぶ。

[0162] 図25に示すように第七実施形態の仮想環境構築装置4は、仮想素材取得部41と、仮想素材記憶部41Aと、仮想動的素材位置姿勢検出部42と、仮想素材位置姿勢記憶部42Aと、仮想特殊動的素材生成部43と、仮想特殊動的素材記憶部43Aと、同期部44と、仮想環境生成部45と、仮想素材組み込みリスト記憶部45Aと、描画部46とを含む。

[0163] 仮想素材取得部41は、仮想動的素材と仮想静的素材を取得する（S41）。仮想動的素材、仮想静的素材は、仮想環境を生成するために必要なCGデータであるため、例えば人手で仮想素材取得部41に予め入力される。ステップS41で取得された仮想動的素材、仮想静的素材は仮想素材記憶部41Aに記憶される。

[0164] 動的素材（例えば野球やテニスのボール）は、例えば図25に示す第1カ

メラ４９１、第２カメラ４９２で撮影される。仮想動的素材位置姿勢検出部４２は、例えば図２５に示す第１カメラ４９１、第２カメラ４９２で動的素材（例えば野球やテニスのボール）を撮影した映像に基づいて、動的素材に対応する仮想動的素材の位置および姿勢（例えばリリースされた野球ボールのＣＧデータの仮想空間上の軌道、サーブされたテニスボールのＣＧデータの仮想空間上の軌道）を検出してその時系列データを取得する（Ｓ４２）。ステップＳ４２で検出された仮想動的素材の位置および姿勢は仮想素材位置姿勢記憶部４２Ａに記憶される。仮想素材位置姿勢記憶部４２Ａには、予め仮想静的素材、仮想特殊動的素材の位置および姿勢（例えばスタジアム、スコアボード、ピッチャーのＣＧデータの位置、姿勢）が記憶されているものとする。仮想静的素材、仮想特殊動的素材の位置および姿勢は、予め人手で入力されることとしてもよい。

[0165] 特殊動的素材（例えば野球のピッチャーやサーブを打つテニスプレイヤー）は、例えば図２５に示す第３カメラ４９３で撮影される。仮想特殊動的素材生成部４３は、例えば図２５に示す第３カメラ４９３で特殊動的素材を撮影した映像に基づいて仮想特殊動的素材を生成する（Ｓ４３）。仮想特殊動的素材生成部４３は、特殊動的素材を撮影した映像を用いて簡易なＣＧデータを作成するものとする。ステップＳ４３で生成されるＣＧデータの詳細については後述する。ステップＳ４３で生成されたＣＧデータ（仮想特殊動的素材）は、仮想特殊動的素材記憶部４３Ａに記憶される。

[0166] 同期部４４は、仮想動的素材と仮想特殊動的素材とを同期する（Ｓ４４）。ステップＳ４４の詳細については後述する。ステップＳ４４における同期の結果は、仮想素材位置姿勢記憶部４２Ａに記憶される。

[0167] 仮想環境生成部４５は、仮想静的素材と、同期された仮想動的素材と、同期された仮想特殊動的素材に基づいて仮想環境を生成する（Ｓ４５）。仮想環境生成部４５は、各仮想素材の位置姿勢情報、同期情報に基づいてＣＧデータを配置することで仮想環境を生成する。

[0168] 描画部４６は、生成された仮想環境を描画し、全天球映像などの映像とし

て出力する（S 4 6）。以下、各構成要件の入力、出力、処理内容について詳細に説明する。

[0169] [仮想素材取得部 4 1]

入力：仮想静的素材、仮想動的素材（入力元：外部装置など、図示せず）

出力：仮想静的素材、仮想動的素材（出力先：仮想素材記憶部 4 1 A）

処理内容：仮想素材取得部 4 1 は仮想静的素材、仮想動的素材を外部装置などから取得し、取得した仮想静的素材、仮想動的素材を仮想素材記憶部 4 1 A に記憶する。仮想静的素材、仮想動的素材は事前に用意され、仮想素材取得部 4 1 に入力されるものとする。

[0170] [仮想素材記憶部 4 1 A]

入力：仮想静的素材、仮想動的素材（入力元：仮想素材取得部 4 1）

出力：仮想静的素材、仮想動的素材（出力先：仮想環境生成部 4 5）

処理内容：仮想素材記憶部 4 1 A は仮想素材取得部 4 1 が取得した仮想静的素材、仮想動的素材を仮想素材 I D と対応付けて記憶し、仮想環境生成部 4 5 の要求に応じて仮想静的素材、仮想動的素材を出力する。仮想静的素材の具体例として、野球であればスタジアムの C G データ、テニスであればテニスコートの C G データなどが挙げられる。仮想動的素材の具体例として、野球であれば野球ボール、テニスであればテニスボールの C G データなどが挙げられる。

[0171] [仮想動的素材位置姿勢検出部 4 2]

入力：センサ（カメラ）の出力値（入力元：例えば第 1 カメラ 4 9 1、第 2 カメラ 4 9 2）

出力：仮想動的素材の各時刻の位置および姿勢（出力先：仮想素材位置姿勢記憶部 4 2 A）

処理内容：仮想動的素材位置姿勢検出部 4 2 は、センサ（カメラ）の出力値に基づいて実環境における動的素材の位置および姿勢を推定し、推定結果に基づいて動的素材に対応する仮想動的素材の各時刻の仮想空間上の位置および姿勢を検出し、その時系列データを取得する。

[0172] 前述したように動的素材の３次元の位置および姿勢を推定するためのセンサとして、複数カメラ（例えば第１カメラ４９１、第２カメラ４９２）が利用できる。複数カメラを利用して物体の３次元の位置情報を算出する手法の具体例として、ステレオ法を用いた方法がある。具体的には、予めカメラキャリブレーションを実施したカメラ群で撮影された映像に対し、画像上の物体の位置を求める。このとき、画像上の物体の位置は手作業で与えてもよいし、テンプレートマッチングなどの検出手法を用いて物体を検出し、その中心位置を物体の位置としてもよい。これら画像上で検出された位置とカメラキャリブレーションで求めたカメラパラメータを用いてステレオ法を適用することで物体の３次元位置を求めることができる。

[0173] また、姿勢情報に関しては、撮影した物体表面の詳細（例えばボールの縫い目、模様）から基準の姿勢に対してどの程度回転しているのかを求めることができる。動的素材の位置および姿勢を求めるセンサの他の例としては、ドップラーレーダーなどを挙げることができる。ドップラーレーダーは公知の技術であるため、詳細な説明は省略する。その他、物体の３次元位置姿勢情報を取得できるセンサおよび手法であればどのような方法でもよい。本ステップＳ４２において推定された動的素材（例えば野球のボール）の位置および姿勢は、本ステップ内において対応する仮想動的素材（例えば野球のボールのＣＧデータ）の仮想空間上の位置および姿勢に変換され、仮想動的素材の位置および姿勢に関する時系列データが出力されるという点に留意する。

[0174] [仮想素材位置姿勢記憶部４２Ａ]

入力：仮想素材の位置姿勢情報、同期情報（入力元：仮想動的素材位置姿勢検出部４２、同期部４４）

出力：仮想素材の位置姿勢情報、同期情報（出力先：同期部４４、仮想環境生成部４５）

処理内容：仮想素材位置姿勢記憶部４２Ａは、各仮想素材の位置姿勢情報、同期情報を記憶する。仮想素材位置姿勢記憶部４２Ａは、仮想動的素材位置

姿勢検出部 4 2 により取得された時系列データを、仮想動的素材の各時刻における位置および姿勢として記憶する。一方、スタジアム、グラウンド、ピッチャーなどのように時間経過とともに位置が遷移しない静的素材、特殊動的素材に関する位置姿勢情報については、仮想素材位置姿勢記憶部 4 2 A は、予め用意されたデータを記憶するものとする。具体的には、仮想素材位置姿勢記憶部 4 2 A には、各時刻における各仮想素材の 3 次元の位置姿勢情報を表すレコードによって構成される仮想素材位置姿勢情報テーブル（図 2 7 参照）が記憶されている。

[0175] 仮想素材位置姿勢情報テーブルは、図 2 7 に示すように、素材を一意に識別する「素材 ID」と、仮想素材を一意に識別する「仮想素材 ID」と、各「時刻」における仮想素材の「位置」および「姿勢」の各値を有する。位置は 3 次元の並進ベクトルとして表され、姿勢は各軸に対する回転角を表す 3 次元ベクトルとして表される。例えば、野球の投球を想定した場合、素材 O_i がストレートボールを表すとする、素材 O_i に対応する仮想素材 E_i は、仮想素材記憶部 4 1 A に記憶された野球ボールの CG データである。このとき、各時刻に対応する位置および姿勢はその時刻に野球ボールの CG データが仮想空間上のどの位置にどの姿勢で存在しているかということを表している。より具体的には、位置 $T_i(1), \dots, T_i(t_i)$ は野球ボールの CG データ（仮想素材 E_i ）の仮想空間上の軌道を、姿勢 $R_i(1), \dots, R_i(t_i)$ は野球ボールの CG データ（仮想素材 E_i ）が仮想空間上をどのように回転していくかを表している。なお、図 2 7 において、時刻に「ー」が格納されている仮想素材は、スタジアムやピッチャーのように時間経過に伴い位置や姿勢が遷移しない仮想静的素材、仮想特殊動的素材であることを表す。このような位置や姿勢が遷移しない仮想素材の位置姿勢情報は予め記憶されているものとする。また、各仮想素材に対して定められている時刻は、その仮想素材が後述する仮想環境生成部 4 5 によって仮想環境に組み込まれた時点からの経過時刻を表す。具体的には、仮想環境生成部 4 5 が時刻 t_a の時点で仮想素材 E_i を仮想環境に組み込んだ場合、時刻 $(t_a + 1)$

の時点の仮想素材 E_i の位置および姿勢はそれぞれ $T_i(1)$, $R_i(1)$ であり、時刻 $(t_a + 2)$ の時点の仮想素材 E_i の位置および姿勢はそれぞれ $T_i(2)$, $R_i(2)$ である。

[0176] [仮想特殊動的素材生成部 43]

入力：映像（入力元：例えば第3カメラ 493）

出力：仮想特殊動的素材（出力先：仮想特殊動的素材記憶部 43A）

処理内容：仮想特殊動的素材生成部 43 は、カメラ（例えば第3カメラ 493）で撮影した映像を入力とし、その映像の一部または全部を予め用意した CG データにテクスチャとして貼り付けることで、実写に基づく CG データを作成し、当該 CG データを仮想特殊動的素材として出力する。

[0177] 具体例として、野球のピッチャーの仮想特殊動的素材を生成する場合の仮想特殊動的素材生成部 43 の動作を図 28 に示す。

[0178] 図 28 の例では、映像を貼り付けるベースとなる CG データとして長方形の平面（ビルボード）を利用する。まず仮想特殊動的素材生成部 43 は、カメラ（例えば第3カメラ 493）で撮影した映像 V を読みだす（S131）。なお、この映像 V にはピッチャーが含まれるものとする。次に仮想特殊動的素材生成部 43 は、映像 V から部分映像 V' を切り出す（S132）。なお、切り出す領域の例としては、例えばピッチャーが存在する領域を含む矩形領域が挙げられる。この矩形領域の大きさや位置は映像 V を見た人が予め人手で与えてもよいし、自動的に与えられてもよい。矩形領域の大きさや位置を自動的に与える方法としては、例えば背景差分法などがある。背景差分法を用いることでピッチャーが存在する領域を求めることができるので、この領域を含むように矩形領域を定めればよい。また、切り出す領域は矩形領域に限らず、例えばピッチャーが存在する領域のみを切り出してもよい。ピッチャーが存在する領域のみを切り出す方法として、前述の背景差分法などを用いることができる。なお、映像 V から部分映像 V' を切り出す処理は行っても行わなくてもよい。部分映像 V' を切り出す処理を行わない場合は、便宜上、映像 V と同じ大きさの矩形領域を切り出して映像 V' を得たものと

して取り扱う。最後に仮想特殊動的素材生成部 4 3 は、映像 V' を長方形の平面（ビルボード）にテクスチャとして貼り付け（S 1 3 3）、その CG データを仮想特殊動的素材として仮想特殊動的素材記憶部 4 3 A に出力する。

[0179] [仮想特殊動的素材記憶部 4 3 A]

入力：仮想特殊動的素材（入力元：仮想特殊動的素材生成部 4 3）

出力：仮想特殊動的素材（出力先：同期部 4 4、仮想環境生成部 4 5）

処理内容：仮想特殊動的素材記憶部 4 3 A は、ステップ S 4 3 で生成された仮想特殊動的素材を記憶し、各構成要件の要求に応じて各構成要件に仮想特殊動的素材を出力する。

[0180] [同期部 4 4]

入力：仮想特殊動的素材（入力元：仮想特殊動的素材記憶部 4 3 A）、仮想動的素材の位置姿勢情報（入力元：仮想素材位置姿勢記憶部 4 2 A）、仮想特殊動的素材の位置姿勢情報（入力元：仮想素材位置姿勢記憶部 4 2 A）

出力：同期情報（出力先：仮想素材位置姿勢記憶部 4 2 A）

処理内容：同期部 4 4 は、仮想特殊動的素材（例えばピッチャーの CG データ）と、仮想特殊動的素材に関連する仮想動的素材（例えばピッチャーが投げたボールの CG データ）の同期を取る。上記組み合わせの他の例としてはテニス選手の CG データと、その選手がサーブしたテニスボールの CG データなどが該当する。どの仮想特殊動的素材とどの仮想動的素材が関連するかに関しては、予め定められているものとする。同期とは、図 2 7 の仮想素材位置姿勢情報テーブルの同期素材の同期開始時刻を求めることに相当する。

[0181] 具体例としてピッチャーの CG データと、そのピッチャーが投げたボールの CG データの同期を取る場合を考える。この場合、ビルボードにテクスチャとして貼り付けられた映像 V' において、ピッチャーの手からボールがリリースされた瞬間の時刻を、同期素材であるボールの CG データの同期開始時刻とすればよい。ここで、手からボールがリリースされた瞬間の時刻はどのような方法を用いて求めてもよい。

[0182] 例えば、テクスチャ（映像 V'）の毎時刻に対してボールのテンプレート

マッチングを行い、最初に検出された時刻をボールがリリースされた瞬間の時刻、すなわち同期開始時刻としてもよい。あるいはテクスチャ（映像 V' ）の毎時刻に対してピッチャーのリリース時の手や腕の形のテンプレートをを用いてテンプレートマッチングを行い、最初に検出された時刻を同期開始時刻としてもよい。

[0183] 上記の方法ではピッチャーがボールをリリースする瞬間が映像 V' に撮影されていることが前提となるが、ピッチャーの腕の振りの速度は非常に速いため、映像の撮影に用いられる一般的なフレームレート（例えば30fps, 60fpsなど）では、撮影した映像 V' にリリースの瞬間が収められていない場合がある。リリースの瞬間が映像 V' に収められていない場合、上記のようにボールや手や腕のテンプレートをマッチングしてその時刻をリリース時刻とするのではなく、リリース前後の画像からリリース時刻を推定してもよい。

[0184] リリース前後の画像からリリース時刻を推定する方法としては、例えば学習を用いる手法がある。具体的には、リリースの瞬間を含むような高フレームレート（例えば1000fps）で撮影した画像群からリリースの瞬間（時刻）を求め、リリースの瞬間（時刻）を基準として所定のフレーム過去または未来にシフトした時刻に該当するリリース前画像、リリース後画像からなる画像ペアを、シフトするフレーム数を異ならせて複数個作成する。上記の複数個の画像ペアとリリースの瞬間（時刻）を教師データとして、画像ペアからリリースの瞬間（時刻）を求めるモデルを学習する。このように学習されたモデルを用いることにより、低フレームレートにおけるリリース前の画像（第1画像ともいう）、低フレームレートにおけるリリース後の画像（第2画像ともいう）の画像ペアが与えられた場合に、リリースの瞬間（時刻）を推定することが可能となる。モデルの学習方法はどのような方法でもよい。例えばリリース前画像、リリース後画像における手の座標値や、腕の方向、リリース後のボールの位置などを特徴量として用いて線形識別器を適用してもよいし、特徴抽出も含めてDNNを適用してもよい。

- [0185] ステップS 4 4 をより一般的な表現で再定義するとすれば、同期部 4 4 は、人間（例えばピッチャー）と物体（例えばボール）とが離隔（例えばリリース）する前を撮影した画像である第 1 画像（リリース前の画像）と、人間（例えばピッチャー）と物体（例えばボール）とが離隔（例えば、リリース）した後を撮影した画像である第 2 画像（リリース後の画像）とからなる画像ペアを、実環境を撮影した映像（例えば映像 V' ）から抽出し、画像ペアを予め学習されたモデルと照合することで人間（例えばピッチャー）と物体（例えばボール）とが離隔する時刻（例えば、リリースされる瞬間）を推定し、推定された時刻に基づいて仮想動的素材（例えばボールの CG データ）と仮想特殊動的素材（例えばピッチャーの CG データ）を同期する。
- [0186] なお、リリース前後の画像からリリース時刻を推定するために用いるモデルは、同期部 4 4 が予め学習して内部に記憶しておいてもよいし、仮想環境構築装置とは異なる外部のモデル学習装置により予め学習して、同期部 4 4 の内部に記憶しておいてもよい。
- [0187] 上記の方法を用いたリリース時刻検出の一例を図 2 9 に示す。同期部 4 4 は、ピッチャーの CG データに貼られたテクスチャ（映像 V' ）の全てのフレーム（毎時刻）に対してリリース判定を行う。具体的には、同期部 4 4 は、ピッチャーのリリース時の手の形のテンプレートを用いてテンプレートマッチングを実行し、マッチングスコアを算出する（S 1 4 1）。マッチングスコアが閾値 t_{h1} 以上となる場合（S 1 4 2 Y、すなわち類似の場合）、その時刻をボールの同期開始時刻とする（エンド）。同期部 4 4 はステップ S 1 4 1 の処理を映像 V' の全てのフレーム（毎時刻）に対して行い、映像 V' の全てのフレームにおいて閾値 t_{h1} 以上のスコアが得られなかった場合（S 1 4 3 Y）、同期部 4 4 はボールのテンプレートを用いて最初のフレーム（時刻）からテンプレートマッチングを実行し、マッチングスコアを算出する（S 1 4 4）。マッチングスコアが閾値 t_{h2} 以上となった場合に（S 1 4 5 Y）、同期部 4 4 は、閾値 t_{h2} 以上となった最初の時刻 T_a をリリース後の時刻とし、その 1 時刻前の時刻 $T_a - 1$ をリリース前の時刻とす

る（S 1 4 6）。同期部 4 4 は、時刻 T_{a-1} における第 1 画像と、時刻 T_a における第 2 画像よりなる画像ペアと予め学習したモデルを用いてリリース時刻を推定する（S 1 4 6）。閾値 t_{h1} および閾値 t_{h2} は予め定められているものとする。同期部 4 4 は、以上の処理で推定した時刻を図 2 7 に示した仮想素材位置姿勢情報テーブルにおける同期素材の同期開始時刻として設定する。

[0188] 上記フローチャートのステップ S 1 4 1～S 1 4 6 の動作をより一般的な表現で再定義すれば、同期部 4 4 は、人間（例えばピッチャー）と物体（例えばボール）とが離隔する瞬間における人間（例えばピッチャー）の身体のテンプレートとして予め用意された第 1 のテンプレート（例えば手のテンプレート）と、人間（例えばピッチャー）と物体（例えばボール）とが離隔する瞬間における物体（例えばボール）のテンプレートとして予め用意された第 2 のテンプレート（例えばボールのテンプレート）を用い、実環境を撮影した映像（例えば映像 V' ）に対してテンプレートマッチングを実行して人間（例えばピッチャー）と物体（例えばボール）とが離隔（例えば、リリース）する時刻を推定し、推定された時刻に基づいて仮想動的素材と仮想特殊動的素材とを同期する。なお、上記第 1 のテンプレート（例えば手のテンプレート）と第 2 のテンプレート（例えばボールのテンプレート）はどちらか一方のみを用いてもよい。

[0189] [仮想環境生成部 4 5]

入力：仮想静的素材（入力元：仮想素材記憶部 4 1 A）、仮想動的素材（入力元：仮想素材記憶部 4 1 A）、仮想素材の位置姿勢情報（入力元：仮想素材位置姿勢記憶部 4 2 A）、仮想特殊動的素材（入力元：仮想特殊動的素材記憶部 4 3 A）、仮想素材組み込みリスト（入力元：仮想素材組み込みリスト記憶部 4 5 A）

出力：仮想環境のデータ（出力先：描画部 4 6）

処理内容：仮想環境生成部 4 5 は、仮想素材記憶部 4 1 A から仮想静的素材、仮想動的素材を読み出し、仮想素材位置姿勢記憶部 4 2 A から各仮想素材

の位置姿勢情報を読みだし、仮想特殊動的素材記憶部43Aから仮想特殊動的素材を読み出し、これらを仮想素材組み込みリストに従って統合し、仮想環境のデータを生成する(S45)。

[0190] 生成された仮想環境のデータは描画部46に送られる。仮想環境のデータとは、ある仮想的な空間において一つ以上の仮想素材が配置された仮想空間のデータを指す。仮想環境生成部45は、仮想環境を生成するための仮想素材組み込みリストを仮想素材組み込みリスト記憶部45Aから読み出す。この仮想素材組み込みリストは、仮想素材を一意に識別する「仮想素材ID」と、その仮想素材を仮想環境に組み込む時刻である「組み込み開始時刻」を持つ。仮想素材組み込みリストは予め人手で用意され、仮想素材組み込みリスト記憶部45Aに予め記憶されているものとする。仮想環境生成部45は仮想素材組み込みリストに従って、仮想素材を仮想環境に組み込む。このとき、仮想素材の位置姿勢を仮想素材位置姿勢記憶部42Aから読み出し、その位置姿勢に基づいて仮想空間上に配置する。仮想環境生成部45は、仮想素材組み込みリストに記載されている全ての仮想素材を仮想環境に組み込んだら処理を終了する。

[0191] [描画部46]

入力：仮想環境のデータ（入力元：仮想環境生成部45）

出力：映像（出力先：映像表示デバイス（図示せず））

処理内容：描画部46は、ステップS45で生成された仮想環境のデータを描画し、ユーザが視聴できる状態、例えば映像として出力する(S46)。描画部46が出力する映像は、仮想環境のある視点から一方向を見るような映像でもよいし、全天球映像のように全方位を見ることができる映像でもよい。例えば、ある視点から一方向を見るような映像を出力する場合は、ある位置に仮想的に透視投影カメラCを設置し、仮想環境の空間中に存在するCGデータをカメラCの画像平面に投影すればよい。具体的には、TをカメラCの外部パラメータである位置とし、Rを姿勢とし、Kを内部パラメータとすると、仮想環境のデータが存在する空間中のある点pは、以下のようにカ

メラCの画像平面上に点qとして投影される。

[0192] [数4]

$$q = K(Rp + T)$$

[0193] ただし、内部パラメータKはカメラCの焦点距離や画像中心からなる 3×3 の行列である。

[0194] ある視点から全方位を見渡せる全天球映像を出力する場合は、ある視点に仮想的に全天球カメラOを設置し、仮想環境の空間中に存在するCGデータをカメラOの画像平面に投影すればよい。具体的には、TをカメラOの外部パラメータである位置とし、Rを姿勢とし、Wを投影する画像平面の幅とし、Hを高さとすると、仮想環境のデータが存在する空間中のある点pは、以下のようにカメラOの画像平面上に点 $q = (q_x, q_y)$ として投影される。

[0195] [数5]

$$p' = Rp + T = (p'_x p'_y p'_z)^T$$

$$q_x = \left(\frac{\theta W}{2\pi} \right) + \frac{W}{2}$$

$$q_y = \frac{\phi H}{\pi}$$

ただし、

$$\theta = \arctan \left(\frac{p'_z}{p'_x} \right),$$

$$\phi = \arccos \left(\frac{p'_y}{|p'|} \right)$$

[0196] この例では全天球映像を2次元画像で表現する図法として正距円筒図法を用いているが、メルカトル図法など他の図法を用いてもよい。

[0197] 第七実施形態の仮想環境構築装置によれば、臨場感の高い仮想環境を簡易に生成できる。

[0198] [第八実施形態]

スポーツの試合において、事前に体験していない初めての状況に遭遇する

と、十分なパフォーマンスを発揮できないことがある。野球では「初物に弱い」という表現を用いて初対戦の投手を苦手とする状況が語られることがある。この状況に対し、実際の試合と同じ環境において選手の目線で事前に投手の投球を体験することが有効であると考えられる。

[0199] しかしながら、次の対戦相手の投球を完全な形で事前体験することは事実上困難である。次の対戦相手の投球を事前体験するために、スタンドなどから撮影した対戦相手の過去の投球映像を視聴する方法が知られているが、この方法において視聴される映像は実際にバッターボックスに立っている選手が見ている風景とは視点や視野が異なり、事前体験として十分ではない。

[0200] 上記の非特許文献 1 には、選手の目線における事前体験の実現方法としてバッターボックス内の選手の目線の位置に 360 度のパノラマ画像を撮影できるカメラ（以下、全天球カメラと呼ぶ）を設置し、全天球カメラで撮影した選手目線の映像をヘッドマウントディスプレイにより視聴する方法が開示されている。

[0201] しかしながら、試合中にバッターボックス内に全天球カメラを設置することは事実上困難であるから、非特許文献 1 の方法は実現が難しい。

[0202] 非特許文献 1 以外の事前体験の実現方法として、例えば実環境を全て CG で再現した仮想環境を生成することも考えられる。しかしながら、プレー中の選手の CG を正確に作るためには実際にプレーしている選手に高精度なセンサを装着させる必要があり、実環境を全て CG で再現することはかなり高コストになる。上記参考文献 4 には、簡易に臨場感の高い仮想環境を作り出す方法として、ピッチャーのように複雑な動きをする素材に対しては、実写映像を CG データ（ビルボード）にテクスチャとして貼り付ける方法が開示されている。

[0203] 例えば図 30 に示すように、野球の右バッターボックス 471 に立つバッター 482 の体験を事前体験するために、仮想環境を生成する場合を想定する。この場合、ピッチャー 481 はカメラで撮影した映像に基づいたビルボード表現を用いて CG 化することとすると、ピッチャー 481 を撮影するカ

メラのアングルはバッター４８２が見ている方向と極力一致することが望ましい。すなわち、カメラ４９３aのようにピッチャー４８１とバッター４８２を通る直線上に存在し、かつその撮影方向がピッチャー４８１とバッター４８２を通る直線と平行な方向であることが望ましいと考えられる。しかしながら、実際の試合においてカメラ４９３aの位置からピッチャー４８１を撮影しようとする、バッター４８２がオクルージョンになってしまうため、カメラ４９３aで撮影した映像を用いてCGデータを作成すると違和感が大きいという課題があった。また、バッター４８２のオクルージョンを回避する方法としてカメラ４９３bのように別の角度から撮影した映像を用いる方法も考えられるが、カメラ４９３bとカメラ４９３aは視点が異なるため、カメラ４９３bで撮影した映像を用いてCGデータを作成してもピッチャー４８１の向きに関して違和感が大きいという課題があった。

[0204] そこで第八実施形態では違和感の少ない仮想環境を生成できる仮想環境構築装置を提供することを目的とする。

[0205] 位置が遷移する動く素材である動的素材と、位置が遷移せず動かない素材である静的素材と、位置が遷移せず動く素材である特殊動的素材が実環境に存在するものとする。第八実施形態の仮想環境構築装置は、このような実環境を再現する仮想環境を生成する仮想環境構築装置である。第八実施形態の仮想環境構築装置は、仮想素材取得部と、仮想動的素材位置姿勢検出部と、事前準備部と、類似映像検索部と、仮想特殊動的素材生成部と、同期部と、仮想環境生成部を含む。

[0206] 仮想素材取得部は、動的素材を仮想環境上に再現するための仮想動的素材と、静的素材を仮想環境上に再現するための仮想静的素材とを取得する。仮想動的素材位置姿勢検出部は、仮想動的素材の位置および姿勢を検出してその時系列データを取得する。事前準備部は、予め定めた事前準備フェーズにおいて取得された時系列データと、これに関連する特殊動的素材の映像とを対応付ける。類似映像検索部は、予め定めた仮想環境生成フェーズにおいて取得された時系列データと類似する事前準備フェーズにおける時系列データ

を検索し、検索された時系列データと対応付けられた特殊動的素材の映像を出力する。仮想特殊動的素材生成部は、出力された特殊動的素材の映像に基づいて、特殊動的素材を仮想環境上に再現するための仮想特殊動的素材を生成する。同期部は、仮想動的素材と仮想特殊動的素材とを同期する。仮想環境生成部は、仮想静的素材と、同期された仮想動的素材と、同期された仮想特殊動的素材に基づいて仮想環境を生成する。

[0207] 第八実施形態ではピッチャーやボールやスタジアムのように現実存在する人、物を素材と呼ぶ。素材のうちその位置が（時間経過により）遷移し、かつ動く素材（例えば、ボール）を動的素材、その位置が遷移せず動かない素材（例えば、スタジアム、スコアボード、観客席）を静的素材、その位置は遷移しないものとみなされるがその位置に留まったまま動く素材（例えば、野球のピッチャーやサーブを打つテニスプレイヤー）を特殊動的素材と呼ぶ。

[0208] 素材を仮想環境に再現するためのコンピュータ・グラフィクスデータ（CGデータ）を仮想素材と呼ぶ。仮想素材のうち、動的素材を仮想環境上に再現するための仮想素材を仮想動的素材、静的素材を仮想環境上に再現するための仮想素材を仮想静的素材、特殊動的素材を仮想環境上に再現するための仮想素材を仮想特殊動的素材と呼ぶ。

[0209] 第八実施形態の仮想環境構築装置5の構成を図31に示す。第八実施形態では仮想環境構築装置5の外部または内部に1台以上のカメラを有するものとする。少なくとも1台のカメラは図30のカメラ493aのように、事前体験をする対象（以下、体験対象。例えば図30におけるバッター482など）と実写映像を撮影する対象（以下、撮影対象。例えば図30におけるピッチャー481など）を結ぶ直線上またはその付近に設置され、撮影対象の方向を向いていることとする。図31の例では、動的素材を撮影するための第1カメラ491、第2カメラ492、特殊動的素材を撮影するための第3カメラ493を仮想環境構築装置5の外部に備えるものとし、第3カメラ493は、図30のカメラ493aのように体験対象と撮影対象を結ぶ直線上

またはその付近に設置されているものとする。なお詳細は後述するが、第1カメラ491、第2カメラ492は他のセンサに代替可能である。

[0210] 第八実施形態の仮想環境構築装置5は、事前準備部40と、映像記憶部40Aと、仮想素材取得部41と、仮想素材記憶部41Aと、仮想動的素材位置姿勢検出部42と、仮想素材位置姿勢記憶部42Aと、類似映像検索部47と、仮想特殊動的素材生成部43と、仮想特殊動的素材記憶部43Aと、同期部44と、仮想環境生成部45と、仮想素材組み込みリスト記憶部45Aと、描画部46とを含む。

[0211] 以下、第八実施形態の仮想環境構築装置5の動作の概略を説明する。以下では野球を例に挙げ、具体的には図30に示したように右バッターボックス471にバッター482が立っている状態で、右バッターボックス471における試合の事前体験が可能な仮想環境を生成することを想定して説明を行う。本実施例において、仮想環境構築装置5の動作の流れは大きく2つのフェーズに分けることができる。1つは事前準備フェーズである。事前準備フェーズでは、ピッチャー481の映像とピッチャー481が投げたボールの軌道に対応付けて保存する。もう一つは仮想環境生成フェーズである。仮想環境生成フェーズでは、各仮想素材を組み合わせることで仮想環境を生成する。

[0212] <事前準備フェーズ>

事前準備フェーズにおいて、動的素材（ボール）は、例えば図31に示す第1カメラ491、第2カメラ492で撮影される。仮想動的素材位置姿勢検出部42は、例えば図31に示す第1カメラ491、第2カメラ492で動的素材（ボール）を撮影した映像に基づいて、動的素材に対応する仮想動的素材の位置および姿勢（リリースされたボールのCGデータの仮想空間上の軌道）を検出してその時系列データを取得する（S52A）。

[0213] 特殊動的素材（ピッチャー）は、例えば図31に示す第3カメラ493で撮影される。事前準備部40は、例えば図31に示す第3カメラ493で特殊動的素材（ピッチャー）を撮影した映像を取得し、ステップS42Aにおいて取得された時系列データと、これに関連する特殊動的素材（ピッチャー

）の映像とを対応付け、映像記憶部４０Ａに記憶する（Ｓ５０）。

[0214] ステップＳ５０では、オクルージョンを含まない特殊動的素材（ピッチャー）の映像が選別される。事前準備部４０が映像を自動的に選別することとしてもよいし、人手によりオクルージョンを含まない映像が選別されてもよい。例えばオクルージョンを含まない映像は左バッターボックス４７２にバッター４８２がいる時に、カメラ４９３ａの角度で投球を撮影することで取得することができる。ステップＳ４２Ａでは、オクルージョンを含まない特殊動的素材（ピッチャー）に対応した時系列データが取得され、ステップＳ５０において、これらが関連付けられて映像記憶部４０Ａに記憶される。

[0215] ＜仮想環境生成フェーズ＞

仮想環境生成フェーズにおいて、まず仮想素材取得部４１は、仮想動的素材と仮想静的素材を取得する（Ｓ５１）。仮想動的素材、仮想静的素材は、仮想環境を生成するために必要なＣＧデータであるため、例えば人手で仮想素材取得部４１に予め入力される。ステップＳ５１で取得された仮想動的素材、仮想静的素材は仮想素材記憶部４１Ａに記憶される。

[0216] 動的素材（ボール）は、前述と同様に、第１カメラ４９１、第２カメラ４９２で撮影される。仮想動的素材位置姿勢検出部４２は、ステップＳ５２Ａと同様に、仮想動的素材の位置および姿勢（野球ボールのＣＧデータの仮想空間上の軌道）を検出してその時系列データ（以下、軌道データともいう）を取得する（Ｓ５２Ｂ）。ステップＳ５２Ｂで検出された仮想動的素材の位置および姿勢（時系列データ、軌道データ）は仮想素材位置姿勢記憶部４２Ａに記憶される。この時の時系列データ（軌道データ）はクエリとして、後述のステップＳ４７で利用される。仮想素材位置姿勢記憶部４２Ａには、予め仮想静的素材、仮想特殊動的素材の位置および姿勢（例えばスタジアム、スコアボード、ピッチャーのＣＧデータの位置、姿勢）が記憶されているものとする。仮想静的素材、仮想特殊動的素材の位置および姿勢は、予め人手で入力されることとしてもよい。

[0217] 類似映像検索部47は、ステップS52Bにおいて取得された時系列データ（軌道データ）をクエリとして、これに類似するステップS52Aにおける時系列データ（軌道データ）を、映像記憶部40Aから検索し、検索された時系列データ（軌道データ）と対応づけられた特殊動的素材（ピッチャー）の映像を出力する（S57）。

[0218] 仮想特殊動的素材生成部43は、ステップS57で出力された特殊動的素材の映像に基づいて仮想特殊動的素材を生成する（S53）。仮想特殊動的素材生成部43は、特殊動的素材を撮影した映像を用いて簡易なCGデータを作成するものとする。ステップS53で生成されるCGデータの詳細については後述する。ステップS53で生成されたCGデータ（仮想特殊動的素材）は、仮想特殊動的素材記憶部43Aに記憶される。

[0219] 同期部44は、仮想動的素材と仮想特殊動的素材とを同期する（S54）。ステップS54の詳細については後述する。ステップS54における同期の結果は、仮想素材位置姿勢記憶部42Aに記憶される。

[0220] 仮想環境生成部45は、仮想静的素材と、同期された仮想動的素材と、同期された仮想特殊動的素材に基づいて仮想環境を生成する（S55）。仮想環境生成部45は、各仮想素材の位置姿勢情報、同期情報に基づいてCGデータを配置することで仮想環境を生成する。

[0221] 描画部46は、生成された仮想環境を描画し、全天球映像などの映像として出力する（S56）。なお、上述の例におけるピッチャーとボールの関係は、テニス、バレーボール、卓球やバドミントンなどの様々な球技における選手とボール（あるいは羽など）に置き換えて考えることができる。以下、各構成要件の入力、出力、処理内容について詳細に説明する。

[0222] [事前準備部40]

入力：事前準備フェーズにおいて取得された時系列データ（入力元：仮想動的素材位置姿勢検出部42）、映像V（入力元：第3カメラ493）

出力：事前準備フェーズにおいて取得された時系列データ（出力先：映像記憶部40A）、映像V（出力先：映像記憶部40A）

処理内容：事前準備部40は、事前準備フェーズにおいて取得された仮想動的素材（ボールのCGデータ）の時系列データ（軌道データ）と、これに関連する特殊動的素材（ピッチャー）の映像Vとを関連付けて映像記憶部40Aに記憶する。映像Vには撮影対象である特殊動的素材（ピッチャー）が含まれているものとする。なお、本実施例では、時系列データ（軌道データ）を仮想動的素材（ボールのCGデータ）の仮想空間上の軌道を表すものとしているが、これに限らず、時系列データ（軌道データ）を現実存在する動的素材（現実のボール）の実空間上の軌道を表すものとしてもよい。後述するように、実空間上の軌道も、仮想空間上の軌道もステップS52A（S52B）で検出可能である。

[0223] [映像記憶部40A]

入力：事前準備フェーズにおいて取得された時系列データ（入力元：事前準備部40）、映像V（入力元：事前準備部40）

出力：事前準備フェーズにおいて取得された時系列データ（出力先：類似映像検索部47）、映像V（出力先：類似映像検索部47）

処理内容：映像記憶部40Aは、事前準備部40から時系列データ（軌道データ）と映像Vとを取得し、これらを関連付けて記憶する。映像記憶部40Aは、類似映像検索部47の要求（クエリ）に応じて、対応する映像Vを出力する。このとき、クエリと類似する時系列データを共に出力してもよい。

[0224] 具体的には、映像記憶部40Aには、各映像に関するレコードによって構成される映像情報テーブル（図33参照）が記憶されている。図33に示すように、映像情報テーブルは、映像を一意に識別する「映像ID」と、各映像に関連する動的素材のIDである「関連動的素材ID」と、各「時刻」における「位置」および「姿勢」の各値を有する。時刻、位置および姿勢の表記に関しては後述する仮想素材位置姿勢情報テーブルと同じである。映像とそれに関連する動的素材の例としては、映像の撮影対象をピッチャーとした場合、これに関連する動的素材はピッチャーが投げたボールに該当する。これらの対応は生成したい仮想環境に応じて予め定めておくものとする。

[0225] [仮想素材取得部 4 1]

入力：仮想静的素材、仮想動的素材（入力元：外部装置など、図示せず）

出力：仮想静的素材、仮想動的素材（出力先：仮想素材記憶部 4 1 A）

処理内容：仮想素材取得部 4 1 は仮想静的素材、仮想動的素材を外部装置などから取得し、取得した仮想静的素材、仮想動的素材を仮想素材記憶部 4 1 A に記憶する。仮想静的素材、仮想動的素材は事前に用意され、仮想素材取得部 4 1 に入力されるものとする。

[0226] [仮想素材記憶部 4 1 A]

入力：仮想静的素材、仮想動的素材（入力元：仮想素材取得部 4 1）

出力：仮想静的素材、仮想動的素材（出力先：仮想環境生成部 4 5）

処理内容：仮想素材記憶部 4 1 A は仮想素材取得部 4 1 が取得した仮想静的素材、仮想動的素材を仮想素材 I D と対応付けて記憶し、仮想環境生成部 4 5 の要求に応じて仮想静的素材、仮想動的素材を出力する。仮想静的素材の具体例として、野球であればスタジアムの C G データ、テニスであればテニスコートの C G データなどが挙げられる。仮想動的素材の具体例として、野球であれば野球ボール、テニスであればテニスボールの C G データなどが挙げられる。

[0227] [仮想動的素材位置姿勢検出部 4 2]

入力：センサ（カメラ）の出力値（入力元：例えば第 1 カメラ 4 9 1、第 2 カメラ 4 9 2）

出力：仮想動的素材の各時刻の位置および姿勢（出力先：仮想素材位置姿勢記憶部 4 2 A）

処理内容：仮想動的素材位置姿勢検出部 4 2 は、センサ（カメラ）の出力値に基づいて実環境における動的素材の位置および姿勢を推定し、推定結果に基づいて動的素材に対応する仮想動的素材の各時刻の仮想空間上の位置および姿勢を検出し、その時系列データを取得する。

[0228] 前述したように動的素材の 3 次元の位置および姿勢を推定するためのセンサとして、複数カメラ（例えば第 1 カメラ 4 9 1、第 2 カメラ 4 9 2）が利

用できる。複数カメラを利用して物体の3次元の位置情報を算出する手法の具体例として、ステレオ法を用いた方法がある。具体的には、予めカメラキャリブレーションを実施したカメラ群で撮影された映像に対し、画像上の物体の位置を求める。このとき、画像上の物体の位置は手作業で与えてもよいし、テンプレートマッチングなどの検出手法を用いて物体を検出し、その中心位置を物体の位置としてもよい。これら画像上で検出された位置とカメラキャリブレーションで求めたカメラパラメータを用いてステレオ法を適用することで物体の3次元位置を求めることができる。

[0229] また、姿勢情報に関しては、撮影した物体表面の詳細（例えばボールの縫い目、模様）から基準の姿勢に対してどの程度回転しているのかを求めることができる。動的素材の位置および姿勢を求めるセンサの他の例としては、ドップラーレーダーなどを挙げることができる。ドップラーレーダーは公知の技術であるため、詳細な説明は省略する。その他、物体の3次元位置姿勢情報を取得できるセンサおよび手法であればどのような方法でもよい。本ステップS52Bにおいて推定された動的素材（例えば野球のボール）の位置および姿勢は、本ステップ内において対応する仮想動的素材（例えば野球のボールのCGデータ）の仮想空間上の位置および姿勢に変換され、仮想動的素材の位置および姿勢に関する時系列データが出力されるという点に留意する。

[0230] [仮想素材位置姿勢記憶部42A]

入力：仮想素材の位置姿勢情報、同期情報（入力元：仮想動的素材位置姿勢検出部42、同期部44）

出力：仮想素材の位置姿勢情報、同期情報（出力先：同期部44、仮想環境生成部45）

処理内容：仮想素材位置姿勢記憶部42Aは、各仮想素材の位置姿勢情報、同期情報を記憶する。仮想素材位置姿勢記憶部42Aは、仮想動的素材位置姿勢検出部42により取得された時系列データを、仮想動的素材の各時刻における位置および姿勢として記憶する。一方、スタジアム、グラウンド、ピ

ッチャーなどのように時間経過とともに位置が遷移しない静的素材、特殊動的素材に関する位置姿勢情報については、仮想素材位置姿勢記憶部42Aは、予め用意されたデータを記憶するものとする。具体的には、仮想素材位置姿勢記憶部42Aには、各時刻における各仮想素材の3次元の位置姿勢情報を表すレコードによって構成される仮想素材位置姿勢情報テーブル（図35参照）が記憶されている。

[0231] 仮想素材位置姿勢情報テーブルは、図35に示すように、素材を一意に識別する「素材ID」と、仮想素材を一意に識別する「仮想素材ID」と、各「時刻」における仮想素材の「位置」および「姿勢」の各値を有する。位置は3次元の並進ベクトルとして表され、姿勢は各軸に対する回転角を表す3次元ベクトルとして表される。例えば、野球の投球を想定した場合、素材 O_i がストレートボールを表すとする、素材 O_i に対応する仮想素材 E_i は、仮想素材記憶部11Aに記憶された野球ボールのCGデータである。このとき、各時刻に対応する位置および姿勢はその時刻に野球ボールのCGデータが仮想空間上のどの位置にどの姿勢で存在しているかということを表している。より具体的には、位置 $T_i(1)$, ..., $T_i(t_i)$ は野球ボールのCGデータ（仮想素材 E_i ）の仮想空間上の軌道を、姿勢 $R_i(1)$, ..., $R_i(t_i)$ は野球ボールのCGデータ（仮想素材 E_i ）が仮想空間上をどのように回転していくかを表している。なお、図35において、時刻に「ー」が格納されている仮想素材は、スタジアムやピッチャーのように時間経過に伴い位置や姿勢が遷移しない仮想静的素材、仮想特殊動的素材であることを表す。このような位置や姿勢が遷移しない仮想素材の位置姿勢情報は予め記憶されているものとする。また、各仮想素材に対して定められている時刻は、その仮想素材が後述する仮想環境生成部45によって仮想環境に組み込まれた時点からの経過時刻を表す。具体的には、仮想環境生成部45が時刻 t_a の時点で仮想素材 E_i を仮想環境に組み込んだ場合、時刻 $(t_a + 1)$ の時点の仮想素材 E_i の位置および姿勢はそれぞれ $T_i(1)$, $R_i(1)$ であり、時刻 $(t_a + 2)$ の時点の仮想素材 E_i の位置および姿勢はそれぞれ

れ $T_i(2)$, $R_i(2)$ である。

[0232] [類似映像検索部 47]

入力：事前準備フェーズにおいて取得された時系列データ（入力元：映像記憶部 40A）、映像 V（入力元：映像記憶部 40A）

出力：映像 V（出力先：仮想特殊動的素材生成部 43）

処理内容：類似映像検索部 47 の処理フローの例を図 36 に示す。類似映像検索部 47 は、仮想素材位置姿勢記憶部 42A から時系列データ（軌道データ）を読み出す（S151）。類似映像検索部 47 は、映像記憶部 40A から時系列データ（軌道データ）を読みだす（S152）。類似映像検索部 47 は、2つの時系列データ（軌道データ）の類似度を計算する（S153）
 なお、ステップ S153 における時系列データ（軌道データ）の類似度はどのように計算してもよい。例えば軌跡データ i から算出したボールの初速を $v_s(i)$ 、終速を $v_e(i)$ 、初速を計算した位置を $P_s(i)$ 、終速の位置を計算した位置を $P_e(i)$ 、そして加速度を $A(i)$ としたとき、軌跡データ i と軌跡データ j の類似度 S_{ij} を以下のように定める。

$$S_{ij} = k_{vs} * |v_s(i) - v_s(j)| + k_{ve} * |v_e(i) - v_e(j)| + k_{ps} * |P_s(i) - P_s(j)| + k_{pe} * |P_e(i) - P_e(j)| + k_a * |A(i) - A(j)|$$

なお、 $v_s(i)$, $v_e(i)$, $P_s(i)$, $P_e(i)$, $A(i)$ は全て 3次元ベクトルであり、 $|\cdot|$ はノルムを表す。さらに、 k_{vs} , k_{ve} , k_{ps} , k_{pe} , および k_a は各三次元ベクトルに対する係数であり、予め定める。

[0233] 次に、類似映像検索部 47 は、全ての時系列データ（軌道データ）を読みだしたか否かを判定する（S154）。全ての時系列データ（軌道データ）を読みだしていなければ（S154N）、処理はステップ S152 に戻る。全ての時系列データ（軌道データ）を読みだしていれば（S154Y）、これまでに繰り返し実行されたステップ S153 により計算した類似度に基づいて一つの軌道データを選択し、その軌道データに対応する映像 V を映像記憶部 40A から読み出す（S155）。なお、軌道データを選択する手段は

どのような手段でもよい。例えば、類似度が最も高い軌道データを選択してもよい。また、ある閾値を定め、その閾値以上の類似度を持つ軌道データの中からランダムな軌道データを選択してもよい。最後に、類似映像検索部47は読みだした映像Vを仮想特殊動的素材生成部43に出力する（S156）。

[0234] [仮想特殊動的素材生成部43]

入力：映像V（入力元：類似映像検索部47）

出力：仮想特殊動的素材（出力先：仮想特殊動的素材記憶部43A）

処理内容：仮想特殊動的素材生成部43は、類似映像検索部47が出力した映像Vを入力とし、その映像の一部または全部を予め用意したCGデータにテクスチャとして貼り付けることで、実写に基づくCGデータを作成し、当該CGデータを仮想特殊動的素材として出力する。

[0235] 具体例として、野球のピッチャーの仮想特殊動的素材を生成する場合の仮想特殊動的素材生成部43の動作を図37に示す。

[0236] 図37の例では、映像を貼り付けるベースとなるCGデータとして長方形の平面（ビルボード）を利用する。まず仮想特殊動的素材生成部43は、類似映像検索部47が出力した映像Vを取得する（S161）。次に仮想特殊動的素材生成部43は、映像Vから部分映像V'を切り出す（S162）。なお、切り出す領域の例としては、例えばピッチャーが存在する領域を含む矩形領域が挙げられる。この矩形領域の大きさや位置は映像Vを見た人が予め人手で与えてもよいし、自動的に与えられてもよい。矩形領域の大きさや位置を自動的に与える方法としては、例えば背景差分法などがある。背景差分法を用いることでピッチャーが存在する領域を求めることができるので、この領域を含むように矩形領域を定めればよい。また、切り出す領域は矩形領域に限らず、例えばピッチャーが存在する領域のみを切り出してもよい。ピッチャーが存在する領域のみを切り出す方法として、前述の背景差分法などを用いることができる。なお、映像Vから部分映像V'を切り出す処理は行っても行わなくてもよい。部分映像V'を切り出す処理を行わない場合は

、便宜上、映像Vと同じ大きさの矩形領域を切り出して映像V'を得たものとして取り扱う。最後に仮想特殊動的素材生成部43は、映像V'を長方形の平面（ビルボード）にテクスチャとして貼り付け（S163）、そのCGデータを仮想特殊動的素材として仮想特殊動的素材記憶部43Aに出力する。

[0237] [仮想特殊動的素材記憶部43A]

入力：仮想特殊動的素材（入力元：仮想特殊動的素材生成部43）

出力：仮想特殊動的素材（出力先：同期部44、仮想環境生成部45）

処理内容：仮想特殊動的素材記憶部43Aは、ステップS43で生成された仮想特殊動的素材を記憶し、各構成要件の要求に応じて各構成要件に仮想特殊動的素材を出力する。

[0238] [同期部44]

入力：仮想特殊動的素材（入力元：仮想特殊動的素材記憶部43A）、仮想動的素材の位置姿勢情報（入力元：仮想素材位置姿勢記憶部42A）、仮想特殊動的素材の位置姿勢情報（入力元：仮想素材位置姿勢記憶部42A）

出力：同期情報（出力先：仮想素材位置姿勢記憶部42A）

処理内容：同期部44は、仮想特殊動的素材（例えばピッチャーのCGデータ）と、仮想特殊動的素材に関連する仮想動的素材（例えばピッチャーが投げたボールのCGデータ）の同期を取る。上記組み合わせの他の例としてはテニス選手のCGデータと、その選手がサーブしたテニスボールのCGデータなどが該当する。どの仮想特殊動的素材とどの仮想動的素材が関連するかに関しては、予め定められているものとする。同期とは、図35の仮想素材位置姿勢情報テーブルの同期素材の同期開始時刻を求めることに相当する。

[0239] 具体例としてピッチャーのCGデータと、そのピッチャーが投げたボールのCGデータの同期を取る場合を考える。この場合、ビルボードにテクスチャとして貼り付けられた映像V'において、ピッチャーの手からボールがリリースされた瞬間の時刻を、同期素材であるボールのCGデータの同期開始時刻とすればよい。ここで、手からボールがリリースされた瞬間の時刻はど

のような方法を用いて求めてもよい。

[0240] 例えば、テクスチャ（映像 V' ）の毎時刻に対してボールのテンプレートマッチングを行い、最初に検出された時刻をボールがリリースされた瞬間の時刻、すなわち同期開始時刻としてもよい。あるいはテクスチャ（映像 V' ）の毎時刻に対してピッチャーのリリース時の手や腕の形のテンプレートを用いてテンプレートマッチングを行い、最初に検出された時刻を同期開始時刻としてもよい。

[0241] 上記の方法ではピッチャーがボールをリリースする瞬間が映像 V' に撮影されていることが前提となるが、ピッチャーの腕の振りの速度は非常に速いため、映像の撮影に用いられる一般的なフレームレート（例えば30fps, 60fpsなど）では、撮影した映像 V' にリリースの瞬間が収められていない場合がある。リリースの瞬間が映像 V' に収められていない場合、上記のようにボールや手や腕のテンプレートをマッチングしてその時刻をリリース時刻とするのではなく、リリース前後の画像からリリース時刻を推定してもよい。

[0242] リリース前後の画像からリリース時刻を推定する方法としては、例えば学習を用いる手法がある。具体的には、リリースの瞬間を含むような高フレームレート（例えば1000fps）で撮影した画像群からリリースの瞬間（時刻）を求め、リリースの瞬間（時刻）を基準として所定のフレーム過去または未来にシフトした時刻に該当するリリース前画像、リリース後画像からなる画像ペアを、シフトするフレーム数を異ならせて複数個作成する。上記の複数個の画像ペアとリリースの瞬間（時刻）を教師データとして、画像ペアからリリースの瞬間（時刻）を求めるモデルを学習する。このように学習されたモデルを用いることにより、低フレームレートにおけるリリース前の画像（第1画像ともいう）、低フレームレートにおけるリリース後の画像（第2画像ともいう）の画像ペアが与えられた場合に、リリースの瞬間（時刻）を推定することが可能となる。モデルの学習方法はどのような方法でもよい。例えばリリース前画像、リリース後画像における手の座標値や、腕の方

向、リリース後のボールの位置などを特徴量として用いて線形識別器を適用してもよいし、特徴抽出も含めてDNNを適用してもよい。

[0243] ステップS54をより一般的な表現で再定義するとすれば、同期部44は、人間（例えばピッチャー）と物体（例えばボール）とが離隔（例えばリリース）する前を撮影した画像である第1画像（リリース前の画像）と、人間（例えばピッチャー）と物体（例えばボール）とが離隔（例えば、リリース）した後を撮影した画像である第2画像（リリース後の画像）とからなる画像ペアを、実環境を撮影した映像（例えば映像V'）から抽出し、画像ペアを予め学習されたモデルと照合することで人間（例えばピッチャー）と物体（例えばボール）とが離隔する時刻（例えば、リリースされる瞬間）を推定し、推定された時刻に基づいて仮想動的素材（例えばボールのCGデータ）と仮想特殊動的素材（例えばピッチャーのCGデータ）を同期する。

[0244] なお、リリース前後の画像からリリース時刻を推定するために用いるモデルは、同期部44が予め学習して内部に記憶しておいてもよいし、仮想環境構築装置とは異なる外部のモデル学習装置により予め学習して、同期部44の内部に記憶しておいてもよい。

[0245] 上記の方法を用いたリリース時刻検出の一例を図38に示す。同期部44は、ピッチャーのCGデータに貼られたテクスチャ（映像V'）の全てのフレーム（毎時刻）に対してリリース判定を行う。具体的には、同期部44は、ピッチャーのリリース時の手の形のテンプレートを用いてテンプレートマッチングを実行し、マッチングスコアを算出する（S171）。マッチングスコアが閾値 t_{h1} 以上となる場合（S172Y、すなわち類似の場合）、その時刻をボールの同期開始時刻とする（エンド）。同期部44はステップS171の処理を映像V'の全てのフレーム（毎時刻）に対して行い、映像V'の全てのフレームにおいて閾値 t_{h1} 以上のスコアが得られなかった場合（S173Y）、同期部44はボールのテンプレートを用いて最初のフレーム（時刻）からテンプレートマッチングを実行し、マッチングスコアを算出する（S174）。マッチングスコアが閾値 t_{h2} 以上となった場合に（

S 1 7 5 Y)、同期部44は、閾値 t_{h2} 以上となった最初の時刻 T_a をリリース後の時刻とし、その1時刻前の時刻 T_{a-1} をリリース前の時刻とする(S 1 7 6)。同期部44は、時刻 T_{a-1} における第1画像と、時刻 T_a における第2画像よりなる画像ペアと予め学習したモデルを用いてリリース時刻を推定する(S 1 7 6)。閾値 t_{h1} および閾値 t_{h2} は予め定められているものとする。同期部44は、以上の処理で推定した時刻を図35に示した仮想素材位置姿勢情報テーブルにおける同期素材の同期開始時刻として設定する。

[0246] 上記フローチャートのステップS 1 7 1～S 1 7 6の動作をより一般的な表現で再定義すれば、同期部44は、人間(例えばピッチャー)と物体(例えばボール)とが離隔する瞬間における人間(例えばピッチャー)の身体のテンプレートとして予め用意された第1のテンプレート(例えば手のテンプレート)と、人間(例えばピッチャー)と物体(例えばボール)とが離隔する瞬間における物体(例えばボール)のテンプレートとして予め用意された第2のテンプレート(例えばボールのテンプレート)を用い、実環境を撮影した映像(例えば映像 V') に対してテンプレートマッチングを実行して人間(例えばピッチャー)と物体(例えばボール)とが離隔(例えば、リリース)する時刻を推定し、推定された時刻に基づいて仮想動的素材と仮想特殊動的素材とを同期する。なお、上記第1のテンプレート(例えば手のテンプレート)と第2のテンプレート(例えばボールのテンプレート)はどちらか一方のみを用いてもよい。

[0247] [仮想環境生成部45]

入力：仮想静的素材(入力元：仮想素材記憶部41A)、仮想動的素材(入力元：仮想素材記憶部41A)、仮想素材の位置姿勢情報(入力元：仮想素材位置姿勢記憶部42A)、仮想特殊動的素材(入力元：仮想特殊動的素材記憶部43A)、仮想素材組み込みリスト(入力元：仮想素材組み込みリスト記憶部45A)

出力：仮想環境のデータ(出力先：描画部46)

処理内容：仮想環境生成部45は、仮想素材記憶部41Aから仮想静的素材、仮想動的素材を読み出し、仮想素材位置姿勢記憶部42Aから各仮想素材の位置姿勢情報を読みだし、仮想特殊動的素材記憶部43Aから仮想特殊動的素材を読み出し、これらを仮想素材組み込みリストに従って統合し、仮想環境のデータを生成する（S55）。

生成された仮想環境のデータは描画部46に送られる。仮想環境のデータとは、ある仮想的な空間において一つ以上の仮想素材が配置された仮想空間のデータを指す。仮想環境生成部45は、仮想環境を生成するための仮想素材組み込みリストを仮想素材組み込みリスト記憶部45Aから読み出す。この仮想素材組み込みリストは、仮想素材を一意に識別する「仮想素材ID」と、その仮想素材を仮想環境に組み込む時刻である「組み込み開始時刻」を持つ。仮想素材組み込みリストは予め人手で用意され、仮想素材組み込みリスト記憶部45Aに予め記憶されているものとする。仮想環境生成部45は仮想素材組み込みリストに従って、仮想素材を仮想環境に組み込む。このとき、仮想素材の位置姿勢を仮想素材位置姿勢記憶部42Aから読み出し、その位置姿勢に基づいて仮想空間上に配置する。仮想環境生成部45は、仮想素材組み込みリストに記載されている全ての仮想素材を仮想環境に組み込んだら処理を終了する。

[0248] [描画部46]

入力：仮想環境のデータ（入力元：仮想環境生成部45）

出力：映像（出力先：映像表示デバイス（図示せず））

処理内容：描画部46は、ステップS55で生成された仮想環境のデータを描画し、ユーザが視聴できる状態、例えば映像として出力する（S56）。

描画部46が出力する映像は、仮想環境のある視点から一方向を見るような映像でもよいし、全天球映像のように全方位を見ることができ映像でもよい。例えば、ある視点から一方向を見るような映像を出力する場合は、ある位置に仮想的に透視投影カメラCを設置し、仮想環境の空間中に存在するCGデータをカメラCの画像平面に投影すればよい。具体的には、Tをカメラ

Cの外部パラメータである位置とし、Rを姿勢とし、Kを内部パラメータとすると、仮想環境のデータが存在する空間中のある点pは、以下のようにカメラCの画像平面上に点qとして投影される。

[0249] [数6]

$$q = K(Rp + T)$$

[0250] ただし、内部パラメータKはカメラCの焦点距離や画像中心からなる 3×3 の行列である。

[0251] ある視点から全方位を見渡せる全天球映像を出力する場合は、ある視点に仮想的に全天球カメラOを設置し、仮想環境の空間中に存在するCGデータをカメラOの画像平面に投影すればよい。具体的には、TをカメラOの外部パラメータである位置とし、Rを姿勢とし、Wを投影する画像平面の幅とし、Hを高さとすると、仮想環境のデータが存在する空間中のある点pは、以下のようにカメラOの画像平面上に点 $q = (q_x, q_y)$ として投影される。

[0252] [数7]

$$p' = Rp + T = (p'_x p'_y p'_z)^T$$

$$q_x = \left(\frac{\theta W}{2\pi} \right) + \frac{W}{2}$$

$$q_y = \frac{\phi H}{\pi}$$

ただし、

$$\theta = \arctan \left(\frac{p'_z}{p'_x} \right),$$

$$\phi = \arccos \left(\frac{p'_y}{|p'|} \right)$$

[0253] この例では全天球映像を2次元画像で表現する図法として正距円筒図法を用いているが、メルカトル図法など他の図法を用いてもよい。

[0254] 第八実施形態の仮想環境構築装置によれば、違和感の少ない仮想環境を生成できる。

[0255] 〔第九実施形態〕

スポーツにおいて、試合の前に実際に試合を行っている最中の選手の視点位置からプレーを高い臨場感で体験しておくことは、実戦でのパフォーマンス向上にとって非常に重要である。しかし、試合中に選手のいる場所からビデオを撮影することは競技の妨げとなり難しい。

[0256] これを解決するために、試合中にフィールド外に設置されたカメラによって撮影された映像に基づいて、選手の位置から見た映像を合成する試みが行われている（例えば、参考文献5参照）。参考文献5には、そのアプローチの一つとして、空間の一部を3次元モデルにより作成し、他の一部をビルボードモデルとそこに貼りつけた映像により作成し、任意視点からの映像を合成する取り組みが記載されている。

[0257] 〔参考文献5〕高橋康輔，三上弾，五十川麻理子，小島明，“複数カメラ映像を用いた仮想全天球映像合成に関する検討”，信学技報，vol. 115，no. 76，pp. 43-48，2015年6月

[0258] 合成する空間の一部をコンピュータ・グラフィックス（CG: Computer Graphics）で作成し、他の部分を実写映像で作成し、その空間内に設定した仮想視点からの映像を合成する場合、CGで作成したオブジェクトが実写で作成した領域内にも含まれている可能性がある。この場合、仮想視点から視聴すると、そのオブジェクトがCGで作成されたものと実写で撮影されたものと両方が描画されてしまい不自然である。

[0259] 解決策としては、例えば、1. セグメンテーションなどの手法により実写領域から背景部分を除去し前景オブジェクトのみにすることで実写領域に写らないようにする、2. 実写領域から不要な実写オブジェクト部分を除去および修復する、などの方法が考えられる。1の方法は一定の効果が得られると考えられるが、実写領域内で実写オブジェクトよりも仮想視点に近い位置に他のオブジェクトが存在する場合、解決できない。

[0260] 第九実施形態の目的は、このような点に鑑みて、CGで作成されたオブジェクトと実写で撮影されたオブジェクトとを合成して構築する仮想環境にお

いてリアリティを向上することである。

[0261] 上記の課題を解決するために、第九実施形態の仮想環境構築装置は、複数のオブジェクトが存在する現実環境に基づいてユーザに閲覧させる仮想環境を構築する仮想環境構築装置であって、複数のオブジェクトのうち少なくともいずれかのオブジェクトを対象オブジェクトとして、対象オブジェクトのセンシング難易度とCG描画難易度の少なくともいずれかに基づいてオブジェクトをCGによりモデル構築するか実写映像によりモデル構築するかを示すモデル構築方法を決定するモデル構築方法決定部と、現実環境における対象オブジェクトの状態を計測する状態センシング部と、時間変化を伴わない静的オブジェクトについてモデル構築方法に従って仮想環境中に描画するための環境モデルを構築する環境モデル構築部と、時間変化を伴う動的オブジェクトについてモデル構築方法に従って仮想環境中に描画するためのモデルを構築して環境モデルに合成する動的オブジェクト構築部と、を含む。

[0262] 第九実施形態の仮想環境構築装置は、図39に示すように、モデル構築方法決定部60、状態センシング部61、状態記憶部62、環境モデル構築部63、動的オブジェクト構築部64、視点位置方向取得部65、映像レンダリング部66、および映像提示部67を備える。この仮想環境構築装置が図41に示す各ステップの処理を行うことにより第九実施形態の仮想環境構築方法が実現される。

[0263] 仮想環境構築装置は、例えば、中央演算処理装置（CPU: Central Processing Unit）、主記憶装置（RAM: Random Access Memory）などを有する公知又は専用のコンピュータに特別なプログラムが読み込まれて構成された特別な装置である。仮想環境構築装置は、例えば、中央演算処理装置の制御のもとで各処理を実行する。仮想環境構築装置に入力されたデータや各処理で得られたデータは、例えば、主記憶装置に格納され、主記憶装置に格納されたデータは必要に応じて読み出されて他の処理に利用される。また、仮想環境構築装置の各処理部の少なくとも一部が集積回路等のハードウェアによって構成されていてもよい。仮想環境構築装置が備える各記憶部は、例えば、RAM（

Random Access Memory) などの主記憶装置、ハードディスクや光ディスクもしくはフラッシュメモリ (Flash Memory) のような半導体メモリ素子により構成される補助記憶装置、またはリレーショナルデータベースやキーバリューストアなどのミドルウェアにより構成することができる。仮想環境構築装置が備える各記憶部は、それぞれ論理的に分割されていればよく、一つの物理的な記憶装置に記憶されていてもよい。

[0264] モデル構築方法決定部 60 は、構築対象とする空間内に存在するいずれかのオブジェクト（以下、対象オブジェクトと呼ぶ。）について、センシング難易度と CG 描画難易度の少なくともいずれかに基づいて、その対象オブジェクトに関するモデル構築方法を決定する。モデル構築方法決定部 60 は、センシング難易度と CG 描画難易度とに加えて、観測角度と視聴角度のずれに基づいて、その対象オブジェクトに関するモデル構築方法を決定してもよい。モデル構築方法とは、その対象オブジェクトを CG によりモデル構築するか、実写映像によりモデル構築するか、を示す情報である。センシング難易度とは、その対象オブジェクトの状態を計測する際の困難さを示す指標である。ここで、状態とは、その対象オブジェクトを CG 描画するために必要となるものを指し、位置や姿勢などが含まれる。CG 描画難易度とは、その対象オブジェクトを CG で描画する際の困難さを示す指標である。なお、本形態における CG とは 3 次元 CG であるものとする。観測角度と視聴角度のずれとは、実写映像を撮影した際のカメラの位置とユーザへ提示する映像の仮想視点の位置とのずれである。

[0265] 仮想環境においてオブジェクトを 3 次元 CG で描画する場合と実写映像を利用して描画する場合とでそれぞれメリットとデメリットがある。3 次元 CG で描画する場合のメリットは、センシングした情報を元に描画するため、視点設定位置の自由度が高く、野球のボールをバッターの視点で見るとともに、ユーザが視点位置・向きを変化させる中で、位置が短時間で急激に変化するようなオブジェクトを描画するようなケースでも比較的容易に描画できることである。一方、デメリットは、オブジェクトによってはセンシ

グが難しく、合成された映像の雰囲気 unnatural になる（例えば、頭髮、衣服のはためき、遠方からセンシングした人物の姿勢、等）ことである。実写映像を利用して描画する場合のメリットは、合成された映像の雰囲気が自然になることである。一方、デメリットは、位置が短時間で急激に変化するオブジェクトをきれいに合成しようとする、多数のカメラが必要となることである。モデル構築方法決定部 60 は、これらのメリットとデメリットを勘案し、オブジェクトごとに最適なモデル構築方法を決定する。これにより最終的に構築される仮想環境においてリアリティを向上することができる。

[0266] 具体的には、モデル構築方法決定部 60 は、以下の 7 ステップによりモデル構築方法を決定する。

- [0267] ステップ 1. オブジェクト存在範囲を設定する
 ステップ 2. カメラ位置を設定する
 ステップ 3. 仮想視点位置を設定する
 ステップ 4. 観測角度と視聴角度のずれの最大値を求める
 ステップ 5. センシング難易度を決定する
 ステップ 6. CG 描画難易度を決定する
 ステップ 7. モデル構築方法を判定する

[0268] ステップ 1 において、構築対象の空間における各オブジェクトの存在範囲を設定する。多くのスポーツではルールなどの制約でオブジェクトごとに存在する領域が定まっているため、ルールに基づく存在範囲を設定することが可能である。また、ルールに定められていなくとも典型的なオブジェクトの存在範囲を設定することも可能である。あるいは、撮影された後にモデル構築する場合には、これから構築するシーンに合わせて設定しても構わない。

[0269] ここで、構築対象とする空間を野球グラウンドとして説明する。図 40 は、構築対象とする野球グラウンドを鉛直上方位置の視点から見た図である。このような構築対象の空間を模したイラスト上でユーザがマウスクリックなどによりオブジェクトの存在範囲を指定することで、ステップ 1 における各オブジェクトの存在範囲を設定できる。あるいは、システム利用時に指定し

たものを読み込むことで設定するようにしてもよい。また、ルールに基づいてオブジェクト存在範囲を記載したファイルをあらかじめ用意し、システム利用時に指定することでそのファイルを読み込むようにしても構わない。例えば、野球において投球のシーンを構築する場合、オブジェクトには、例えば「投手」「ボール」などが含まれる。この場合、各オブジェクトに、「投手」「ボール」といったオブジェクトの「種類」、「剛体」「多関節物体」といったオブジェクトの「属性」（用途は後述する）、各オブジェクトが動的オブジェクトであるか静的オブジェクトであるかを示す「動的／静的の種別」（属性と紐づけておくことで、省略してもよい。用途は後述する）、位置情報で表されたオブジェクトの「存在範囲」など、オブジェクトの情報を設定したファイルを予め用意しておく。モデル構築方法決定部60は、当該ファイルを読み込み、オブジェクトの存在範囲を設定する。ファイル中の内容は人手で設定しておく。例えば、処理対象オブジェクトが投手であれば、図40に示すように、マウンド周辺の領域をオブジェクトの存在範囲とする。また、処理対象オブジェクトがボールであれば、図40に示すように、マウンドとホームベースとを結ぶ矩形の領域をオブジェクトの存在範囲とする。図40の例では、バッターボックスの中心を原点とし、バックスクリーン方向をy軸とし、y軸と直交する方向をx軸として、イラスト上の2次元位置座標を位置情報としている。

[0270] ステップ2において、構築対象の空間外に設置された複数のカメラについて、それぞれの位置を取得する。カメラ位置の取得は、オブジェクトの存在範囲と同様に、図40のように空間を模したイラスト上でユーザが指定したものを読み込むことができる。また、別途あらかじめ用意したカメラ位置が記載されたファイルをモデル構築方法決定部60が読み込んでカメラ位置を設定するようにしても構わない。

[0271] ステップ3において、仮想視点として設定される可能性のある範囲を読み込む。仮想視点として設定される可能性のある範囲の取得は、オブジェクトの存在範囲と同様に、図40のようなイラスト上でユーザまたはシステム運

用者が指定したものを読み込むことができる。また、別途あらかじめ用意した仮想視点位置が記載されたファイルをモデル構築方法決定部60が読み込んで仮想視点として設定される可能性のある範囲を設定するようにしても構わない。

[0272] ステップ4において、観測角度と視聴角度とのずれを計算する。オブジェクトに対する観測角度および視聴角度は、ステップ1で取得したオブジェクト存在範囲と、ステップ2で取得したカメラ位置と、ステップ3で取得した仮想視点位置とを用いて求めることができる。例えば、図42(A)は、野球の投球に対する観測角度と視聴角度とのずれ $\theta_{\max}$ を示している。L1はホームベース上を通過するボールとカメラとを結ぶ直線であり、L2はバッターボックスの中心にある仮想視点とボールとを結ぶ直線であり、直線L1と直線L2とのなす角度が観測角度と視聴角度とのずれ $\theta_{\max}$ である。図42(B)は、マウンド上にいる投手に対する観測角度と視聴角度とのずれ $\theta_{\max}$ を示している。

[0273] ステップ5において、オブジェクトごとにセンシング難易度を設定する。センシング難易度は、例えば、センシング難易度テーブルから読み込む。センシング難易度テーブルは、次表に示すように、オブジェクトの属性に応じてあらかじめ定めたセンシング難易度が記載されたテーブルである。次表の例では、剛体であるか否か、多関節物体であるか否か、可塑性物体であるか否か、といった属性に基づいて、センシング難易度の値を設定している。なお、各オブジェクトの属性は、ステップ1で説明したように、オブジェクトごとに「種類」「属性」「動的／静的の種別」「存在範囲」を設定したファイルから読み込んで取得されているものとする。あるいは、各オブジェクトの属性は、構築するオブジェクトごとにユーザがマウスクリックなどによりオブジェクトの属性を指定することで取得してもよい。

[0274]

[表3]

形状複雑度	剛体	多関節物体	可塑性物体
シンプル	a	b	c
複雑	d	e	f

[0275] ステップ6において、オブジェクトごとにCG描画難易度を設定する。CG描画難易度は、例えば、CG描画難易度テーブルから読み込む。CG描画難易度テーブルは、次表に示すように、オブジェクトの属性に応じてあらかじめ定めたCG描画難易度が記載されたテーブルである。次表の例では、剛体であるか否か、多関節物体であるか否か、可塑性物体であるか否か、といった属性に基づいて、CG描画難易度の値を設定している。各オブジェクトの属性の取得方法は、センシング難易度の設定で説明したのと同様である。

[0276] [表4]

形状複雑度	剛体	多関節物体	可塑性物体
シンプル	g	h	i
複雑	j	k	l

[0277] ステップ7において、センシング難易度、CG描画難易度、および観測角度と視聴角度のずれに基づいて、そのオブジェクトのモデル構築方法を判定する。例えば、SEをセンシング難易度とし、CMをCG描画難易度とし、 $\theta_{\max}$ を観測角度と視聴角度のずれとして、

$$\gamma \theta_{\max} / (\alpha SE * \beta CM) > th$$

が成り立つならばCGによりモデル構築し、成り立たなければ実写映像によりモデル構築する、などと判定すればよい。ここで、thは閾値であり、 α 、 β 、 γ は所定のパラメータである。上記の数式は一例であって、センシング難易度、CG描画難易度、および観測角度と視聴角度のずれのうち、ひとつ以上を用いて閾値との比較により判定するのであればどのように変形してもよい。

- [0278] 上述のような各ステップを経ず、オブジェクトごとにCGまたは実写映像の利用をあらかじめ指定したファイルを用意しておき、そのファイルを読み込むことでモデル構築方法を決定することも可能である。
- [0279] 状態センシング部61は、構築対象とする空間内に存在する各オブジェクトについて、その空間におけるオブジェクトの状態をセンシングする。取得した各オブジェクトの状態は状態記憶部62へ記憶する。センシングの方法は、対象とするオブジェクトの属性およびそのオブジェクトのモデル構築方法により異なる。例えば、野球のボールやテニスのボールなどをCG描画するために状態として位置情報をセンシングする場合、野球であれば公知のPITCHf/xシステム、テニスであれば公知のHawk-Eyeシステムなどによりボールの位置情報を取得できることが広く知られている。
- [0280] 環境モデル構築部63は、状態記憶部62に記憶された、各オブジェクトの状態を入力とし、各オブジェクトのうち位置や姿勢の時間変化がほとんどない静的オブジェクトについて、モデル構築方法決定部60により決定されたモデル構築方法に従って、仮想環境中に描画するための環境モデルを構築する。構築した環境モデルは動的オブジェクト構築部64へ送られる。CGによりモデル構築を行うと決定されたオブジェクトについては、3次元形状およびテクスチャからなるモデルを構築する。3次元形状は実際に計測しても構わないし、多くのスポーツにおいてはルールにより競技場等の大きさが定められているため、ルールで定められた典型的な形状を利用しても構わない。実写映像によりモデル構築を行うと決定されたオブジェクトについては、カメラで撮影した実写映像を利用してモデルを構築する。
- [0281] 動的オブジェクト構築部64は、状態記憶部62に記憶された、各オブジェクトの状態を入力とし、各オブジェクトのうち位置や姿勢が時間変化を伴う動的オブジェクトについて、状態記憶部62に記憶された状態に基づいて、モデル構築方法決定部60により決定されたモデル構築方法に従って、仮想環境中に描画するためのモデルを構築する。構築した動画オブジェクトのモデルは、環境モデル構築部63により構築された環境モデルと合成される

。これにより、ユーザへ提示する仮想環境が構築される。構築された仮想環境は映像レンダリング部66へ送られる。

[0282] 具体的には、動的オブジェクト構築部64は、図41に示す6ステップにより動的オブジェクトのモデルを構築する。ステップS61において、各オブジェクトのうち未処理のオブジェクトを処理対象オブジェクトとして設定する。ステップS62において、処理対象オブジェクトがCGによりモデル構築される場合（YES）、ステップS63へ移行する。そうでない場合（NO）、ステップS66へ移行する。ステップS63において、CGによりモデル構築される処理対象オブジェクトが実写映像により構築される領域に含まれているか否かを判定する。含まれる場合（YES）、ステップS64へ移行する。そうでない場合（NO）、ステップS66へ移行する。ステップS64において、CGでモデル構築される処理対象オブジェクトが映りこんでいる領域を推定する。ステップS65において、処理対象オブジェクトが映りこんでいると推定された領域を対象とした修復処理を実施する。ステップS66において、すべてのオブジェクトが処理済みの場合（YES）、処理を終了する。そうでない場合（NO）、ステップS61に戻る。

[0283] ステップS64の修復領域推定については、1. 画像処理によりオブジェクトを検出する方法、2. 競技空間座標との間でキャリブレーションされたカメラを利用して、センシングされた競技空間座標中のオブジェクト位置を外部カメラ位置に変換する方法、が考えられる。以下、それぞれの方法について、具体例を挙げて説明する。

[0284] 上記1の方法は以下のように修復領域を推定する。以下、野球においてバッターボックスからの視線で映像を合成するシーンを想定し、ボールをCGによりモデル構築するものと仮定する。このとき、実写映像によりモデル構築される投手の領域内に実際に撮影されたボールが存在し、これを映像修復技術により除去および補完する状況を説明する。この場合、ボールのテンプレートを用意し、各フレームにおいてテンプレートマッチングによりオブジェクトの位置および大きさを検出することが考えられる。あるいは、ボール

のようにカメラ画像内での形状が円に限定される対象では、エッジ検出後に円形ハフ変換を行うことによって位置および大きさを検出することが可能である。

- [0285] 上記2の方法は以下のように修復領域を推定する。この手法では、センサ座標においてセンシングされたオブジェクトの位置がカメラ座標系に変換され、そのカメラ座標がさらに画面座標系へと変換される。センサ座標 $(x, y, z)^T$ からカメラ座標 $(p, q, r)^T$ への変換行列をA、B、画面座標 $(s, t)^T$ への変換行列をC、Dとすると、

$$(p, q, r)^T = A(x, y, z)^T + B$$

$$(s, t)^T = C(p, q, r)^T + D$$

となる。これにより、映像内での除去対象とするオブジェクト領域を推定することができる。

- [0286] 修復領域の推定について、テニスの場合を例として具体的に説明する。テニスにおいてプレーヤー位置の視点での映像合成を行う場合、図43-44に示すように、プレーヤーを実写映像によりモデル構築し、ボールおよびコートはCGによりモデル構築することとなる。このとき、プレーヤーを撮影するカメラに写り込んでいるボールおよびコート（ネット、ライン）を除去することになる。ボールの軌跡は公知のHawk-Eyeシステムなどにより容易に取得することができる。コートの形状はルールにより既知である。したがって、CGによりモデル構築するオブジェクトの3次元形状は取得可能である。外部カメラとコートの世界座標がキャリブレーションできていれば、コートとボールがどこに撮影されるかわかるので、同様にして修復領域の推定が可能である。

- [0287] ステップS65における修復処理には、例えば、下記参考文献6に記載された画像映像修復技術が利用可能である。

- [0288] 〔参考文献6〕五十川麻理子，三上弾，小島明，“特徴量の削減と合成に基づく体感品質に優れた画像修復”，信学技報，vol. 114，no. 239，MVE2014-37，pp. 37-42，2014年10月

[0289] 視点位置方向取得部65は、ユーザに提示する映像における仮想的な視点位置および視点方向を取得する。仮想視点の位置はモデル構築決定部60のステップ3で設定した範囲内で任意に設定することができる。視点位置方向取得部65は、設定され外部から入力されたユーザに提示する映像における仮想的な視点位置および視点方向を取得し、映像レンダリング部66へ送る。

[0290] 映像レンダリング部66は、動的オブジェクト構築部64により合成された、環境モデルと動的オブジェクトのモデルとを合成して構築された仮想環境、および、視点位置方向取得部65により取得された仮想視点の位置および方向とを用いて、ユーザへ提示する仮想環境の映像をレンダリングする。レンダリングされた映像は映像提示部67へ送られる。

[0291] 映像提示部67は、映像レンダリング部66によりレンダリングされた映像をユーザへ提示する。映像提示部67は、仮想環境構築装置に各種のインターフェースを経由して接続される周辺機器であってもよく、例えば、一般的な液晶ディスプレイやプロジェクタ、ヘッドマウントディスプレイ（HMD: Head Mount Display）などを用いることができる。

[0292] 以下、野球においてバッターボックスからの視線で仮想環境の映像を構築する具体例を説明する。この場合、環境モデルはスタジアムであり、動的オブジェクトは投手およびボールが想定される。スタジアムをさらに細分化すると、打者に近いものから、バッターボックス、ホームベース、ファールライン、スタンドなどがある。これらの各オブジェクトについて、モデル構築方法決定部60で説明した7ステップを行うことによりモデル構築方法を決定する。ここでは、投手が投げるボールを対象のオブジェクトとする具体例を説明する。

[0293] （ステップ1）ボールは、バッターボックスから18.44メートル前方にあるピッチャープレートからピッチャーによって投じられ、ホームベースの少し後方にいるキャッチャーによって捕球される。つまり、バッターボックス内でのおよその打者の位置を原点とした場合、 $(0.5, 18.44) - (0.5, -1)$ がボ

ールのオブジェクト存在範囲となる。

[0294] (ステップ2) カメラはバックネット側に存在するものとし、カメラ位置を(1, -20)として設定する。

[0295] (ステップ3) 仮想視点位置はバッターボックスに立つバッターの視点とし、仮想視点位置を(0, 0)として設定する。

[0296] (ステップ4) ステップ1から3で設定された情報に基づいて、カメラでボールを観測したときの角度(観測角度)と仮想視点位置からボールを観測したときの角度(視聴角度)との間に生じるずれの最大値 $\theta_{\max}$ を決定する。 $\theta_{\max}$ は、実写でのモデル構築の難しさと正の相関を有する。以下、ボールbに関する観測角度と視聴角度のずれは、 $\theta_{\max}(b)$ とする。

[0297] (ステップ5) ボールに関するセンシングの難易度を推定する。これは、センシング難易度テーブルに基づいて決めるものとする。ボールは剛体であり、回転してもアピアランスに変化が生じないため、センシング難易度の低いオブジェクトである。公知のPITCHf/xなど、精度良くボールの状態をセンシング可能なシステムが導入されている場合には、さらにセンシング難易度は低いと考えられる。以下、ボールbに関するセンシング難易度のスコアは、SE(b)とする。

[0298] (ステップ6) ボールに関するCG描画難易度を推定する。これは、CG描画難易度テーブルに基づいて決めるものとする。ボールは歪みが存在せず、球体であるため、CGモデルの作成は容易である。以下、ボールbに関するCG描画難易度のスコアは、CM(b)とする。

[0299] (ステップ7) ステップ4から6で取得した情報を用いて、下記の判定式を検証する。判定式が成り立てば実写映像によりモデル構築するものとし、成り立たなければCGによりモデル構築するものとする。

[0300]
$$SE(b) * CM(b) / \theta_{\max}(b) \geq th$$

他の例として、投手を処理対象オブジェクトとする具体例を説明する。

[0301] (ステップ1) 投手は、投球にあたり1歩前に踏み出す程度の移動しか行わない。そのため、(0.5, 18.44)-(0.5, 17.44)を投手のオブジェクト存在範

囲として設定する。

[0302] (ステップ2) カメラ位置はボールの例と変わらないものとする。すなわち、(1, -20)として設定する。ただし、オブジェクトごとに観測用のカメラを用意することも可能である。

[0303] (ステップ3) 仮想視点位置はボールの場合と変わらないものとする。すなわち、(0, 0)として設定する。

[0304] (ステップ4) 投手pに関する観測角度と視聴角度のずれ $\theta_{\max}(p)$ を設定する。図42(A)～(B)からわかるように、投手pの場合の $\theta_{\max}(p)$ はボールbの場合の $\theta_{\max}(b)$ と比較して小さい値となる。

[0305] (ステップ5) 投手(つまり人間)の場合、多関節物体であるため、センシングが困難である。以下、投手pに関するセンシング難易度のスコアは、 $SE(p)$ とする。

[0306] (ステップ6) 投手の場合、ユニフォームの非剛体かつ複雑な変形に起因して、リアルなCGモデル作成には大きなコストがかかる。以下、投手pに関するCG描画難易度のスコアは、 $CM(p)$ とする。

[0307] (ステップ7) ステップ4から6で取得した情報を用いて、下記の判定式を検証する。判定式が成り立てば実写映像によりモデル構築するものとし、成り立たなければCGによりモデル構築するものとする。

[0308] $SE(p) * CM(p) / \theta_{\max}(p) \geq th$

[0309] このようにして、各々のオブジェクトについて、CGによりモデル構築を行うか、実写映像によりモデル構築を行うかを判定する。各オブジェクトにとって最適なモデル構築方法を取ることができるため、リアリティのある仮想環境を構築することが可能となる。

[0310] 他のスポーツの具体例を簡潔に説明する。前述のテニスの例では、選手および観客席を実写映像によりモデル構築し、ボールおよびコートはCGによりモデル構築することとした。ボールの軌跡は公知のHawk-Eyeシステムにより取得することが可能である。選手を撮影するカメラはコートの外に設置すればよい。サッカーの例では、観客席を実写映像によりモデル構築し、グラ

サウンドをCGによりモデル化する。選手については状況により実写とCGとに分けられる。選手の位置は公知のTRACABシステムなどにより取得可能である。

[0311] 第九実施形態の仮想環境構築技術では、各オブジェクトにとって最適な方法によりモデル構築を行うため、CGで作成されたオブジェクトと実写で撮影されたオブジェクトとを合成して構築する仮想環境においてリアリティが向上する。

[0312] [第十実施形態]

写真や動画を撮影すると、撮影したいものに被さって不要なものが映りこんでしまうことがある。それにより撮影した写真や動画の視聴の体感品質を大きく損なう恐れがあるため、このような不要な映り込みを画像処理によって見た目に違和感なく除去するという手法への需要は極めて高い。なお、以降では、不要な映り込みなどの除去したい画像内の領域および遮蔽などで観測されず補完したい領域を「補完対象領域」と呼ぶ。また、補完対象領域がマスクで与えられた画像を入力して、補完対象領域が補完対象領域外の領域との見た目に違和感なく補完された画像を出力する処理を「補完処理」と呼ぶ。

[0313] 補完対象領域の位置を示すマスクは、静止画であるか動画であるかに関わらず、手動または既存手法（例えば、参考文献7）によって与えられる。なお、マスクとは画像処理において、ある領域が処理を施す対象であるか否かを示す情報である。例えば、図45（1）に示すように、入力画像とは別に処理を施す部分とそうでない部分を二値で与えるマスクや、図45（2）に示すように、画像内で使用されておらず判別が容易な色を入力画像に重畳して処理する領域を明示し、該当色で示された領域以外は処理を施さないようにするものなどが挙げられる。

[0314] [参考文献7] Bai, X., Wang, J., Simons, D., Sapiro, G., “Video sn apcut: robust video object cutout using localized classifiers,” ACM Transactions on Graphics (TOG), vol. 28, Issue 3, August 2009

[0315] 画像および映像の補完処理を実現するための手法としては、パッチ単位で画像中から修復に適した領域を探索し、そのパッチに基づいて逐次補完を行う方法（例えば、参考文献8）や、パッチ単位で画像をどの程度平行移動させれば類似領域が存在するかを調べ、その情報に基づいて補完を行う方法（例えば、参考文献9）等が挙げられる。また、映像に対しては、同一の画像内だけでなく、時系列方向も考慮した三次元パッチ単位で補完を行う方法（例えば、参考文献10）等が用いられている。

[0316] [参考文献8] A. Criminisi, P. Perez, and K. Toyama, “Region filling and object removal by exemplar-based inpainting,” *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1200–1212, Sept. 2004.

[参考文献9] Kaiming He, and Jian Sun, “Image Completion Approaches Using the Statistics of Similar Patches”, *IEEE Transactions on pattern analysis and machine intelligence*, vol. 36, no. 12, pp. 2423–2435, 2014.

[参考文献10] Alasdair Newson, Andres Almansa, Matthieu Fradet, Yann Gousseau, Patrick Perez, “Video Inpainting of Complex Scenes,” *SIAM Journal of Imaging Science* 2014 7:4, 1993–2019

[0317] 映像に映りこむ不要な映り込みなどを除去して違和感がないように補完する映像補完処理において、映像を通してカメラが固定されている条件下にあり、かつ、補完したい領域が背景領域であれば、補完対象領域の画素位置に対応する同一映像内の他フレームの画素値を用いて補完を行うと、高速かつ安定した画質で補完が可能である。このような処理を「貼り付け合成」と記載する。しかし、補完したい領域がフレーム内の他のオブジェクト（例えば人物など）に隣接している場合には、上記の方法を用いるとオブジェクトが欠損してしまうため、良好に補完することができない。そのため、従来はフレーム毎に参考文献8、9のように画像の縦軸・横軸を含む二次元のパッチ単位で補完に適した領域を探索して補完を行う方法を採用するか、参考文献10のように時系列方向も考慮した三次元パッチ単位で補完を行う方法等が用い

られてきた。上記の方法を「コンプリーション処理」または「パッチ探索に基づくコンプリーション処理」と記載する。

[0318] しかしながら、従来の補完手法はパッチ探索に膨大な計算時間を要する。そのため、実用的には、長時間にわたる補完時間を許容するか、ユーザが補完手法や補完対象となるフレームを手動で入力するなどにより高速化を図る必要がある。これにより、長大な待ち時間が発生する、ユーザによる手動入力により著しくユーザビリティを損ねる、などといった課題があった。

[0319] 第十実施形態の目的は、不要な映り込みなどにより補完が必要な領域を含む映像を効率的に補完することができる映像提示技術を提供することである。

[0320] 上記の課題を解決するために、第十実施形態の第一の態様の映像提示装置は、複数フレームからなる映像シーケンスと映像シーケンス中の補完対象領域を指定するマスク情報とを入力として取得する映像入力部と、映像シーケンスに含まれる複数フレーム間の差分を表す二値画像に基づいて前景領域と背景領域とに分離し、補完対象領域の周縁から一定距離以内に存在する前景領域に属する画素数に基づいて補完対象領域の補完方法をパッチ探索に基づくコンプリーションまたは貼り付け合成のいずれかに判定する補完方法判定部と、補完方法に従って補完対象領域を補完する映像補完部と、を含む。

[0321] 第十実施形態の第二の態様の映像提示装置は、複数フレームからなる映像シーケンスと映像シーケンス中の補完対象領域を指定するマスク情報とを入力として取得する映像入力部と、補完対象領域内に存在する画素の平均色と補完対象領域の周縁から一定距離以内に存在する画素の平均色の差分に基づいて、補完対象領域の補完方法をパッチ探索に基づくコンプリーションまたは貼り付け合成のいずれかに判定する補完方法判定部と、補完方法に従って補完対象領域を補完する映像補完部と、を含む。

[0322] 第十実施形態の第三の態様の映像提示装置は、複数フレームからなる映像シーケンスと映像シーケンス中の補完対象領域を指定するマスク情報とを入力として取得する映像入力部と、映像シーケンス中の各フレームをセグメン

テーションして予め定めた特徴を有する前景セグメントと背景セグメントとに分離し、補完対象領域の周縁から一定距離以内に存在する前景セグメントに属する画素数に基づいて補完対象領域の補完方法をパッチ探索に基づくコンプリーションまたは貼り付け合成のいずれかに判定する補完方法判定部と、補完方法に従って補完対象領域を補完する映像補完部と、を含む。

[0323] 第十実施形態では、上記の課題を解決する方法として、前後のフレームの該当する位置の画素を貼り付ける貼り付け合成処理によって高速かつ良好な補完結果が得られやすいという固定カメラ映像の特性を活用する。すなわち、補完対象領域がフレーム内の他のオブジェクト（例えば人物など）と重複しておらず、近傍フレームの該当位置の画素値がそのまま使える場合には他フレームの補完したい画素位置に該当する画素値をそのまま貼り付けることにより補完を行う。一方、補完対象領域がフレーム内の他のオブジェクトと重複しており、他フレームを用いた貼り付け合成処理が難しい場合にはパッチ探索に基づくコンプリーション処理を行う。このように、補完対象領域の特性に応じて適切な補完方法を自動的に判定することで、高速かつ自動で効率的な映像補完を行う。

[0324] 第十実施形態の映像提示装置 70 は、図 46 に示すように、映像入力部 71、補完方法判定部 72、映像補完部 73、および映像出力部 74 を備える。この映像提示装置が図 47 に示す各ステップの処理を行うことにより第十実施形態の映像提示方法が実現される。

[0325] 映像提示装置は、例えば、中央演算処理装置（CPU: Central Processing Unit）、主記憶装置（RAM: Random Access Memory）などを有する公知又は専用のコンピュータに特別なプログラムが読み込まれて構成された特別な装置である。映像提示装置は、例えば、中央演算処理装置の制御のもとで各処理を実行する。映像提示装置に入力されたデータや各処理で得られたデータは、例えば、主記憶装置に格納され、主記憶装置に格納されたデータは必要に応じて読み出されて他の処理に利用される。また、映像提示装置の各処理部の少なくとも一部が集積回路等のハードウェアによって構成されていてもよ

い。

[0326] 図47を参照して、第十実施形態の映像提示方法の処理手続きを説明する。

[0327] ステップS71において、映像入力部71は、複数フレームからなる映像シーケンス V_{orig} と映像シーケンス V_{orig} 中の補完対象領域を指定するマスク情報Dとを入力として受け取る。映像シーケンス V_{orig} およびマスク情報Dは共に、 x （フレーム内の x 軸情報）、 y （フレーム内の y 軸情報）、 t （時間軸情報）の三次元情報を持ち、 $V_{orig}(x, y, t)$ もしくは時間軸情報のみを指定して $V_{orig}(t)$ とも表す。マスク情報Dは、補完対象として指定されたピクセルの位置の集合 $\{(d_{x1}, d_{y1}, d_{t1}), (d_{x2}, d_{y2}, d_{t2}), \dots, (d_{xn}, d_{yn}, d_{tn})\}$ で表してもよいし、映像シーケンス V_{orig} と同じ大きさの映像シーケンスで、欠損領域を1、そうでない領域を0とした二値動画として保持しても構わない。映像シーケンス V_{orig} およびマスク情報Dは、補完方法判定部72へ送られる。

[0328] ステップS72において、補完方法判定部72は、映像入力部71から映像シーケンス V_{orig} およびマスク情報Dを受け取り、映像シーケンス V_{orig} を構成する各フレーム t （ $1 < t < n$ 、 n は映像シーケンス V_{orig} の総フレーム数）において、 $V_{orig}(t)$ を補完するための補完方法 $P(t)$ と、補完に用いる参照領域 $R(t)$ を決定する。補完方法 $P(t)$ は、フレーム t に対する補完方法を表し、パッチ探索に基づくコンプリーション処理および貼り付け合成処理のいずれかを0、1などで示す。参照領域 $R(t)$ は、フレーム t に対する参照フレームベクトルを t' として、 $R(t)=t'$ のように表す。参照フレームベクトル t' は要素数1以上の参照フレームのフレームIDをもつスカラーベクトルである。

[0329] 具体的な参照フレームの決定方法としては、例えば以下のようなものが挙げられる。ただし、本形態では、参照フレームを決定できる手法であれば何でもよく、以下に挙げる方法に限定されるものではない。

[0330] 第一の参照フレーム決定方法は、前景らしさに基づいて決定する方法である。これは、補完したい対象となるフレーム t の補完したい領域（補完対象領域）がフレーム内の人物などの前景らしいオブジェクトに隣接しているか否

かを判定し、隣接していればフレーム内を参照領域としてコンプリーション処理を行い、そうでない場合は補完対象領域の画素位置が前景オブジェクトと隣接しないような他フレームを用いて貼り付け合成処理を行う方法である。

- [0331] 前景らしさに基づいて判定を行う理由を説明する。補完対象領域がフレーム内の人物などの前景らしいオブジェクトに隣接している場合、貼り付け合成処理を用いて他フレームの該当位置画素を合成すると、人物などのオブジェクトの一部が背景画素で合成され、欠損してしまう。一方で、パッチ探索に基づくコンプリーション処理は画像中の他の領域から欠損を埋めるのに適した領域を探索し、その情報を元に補完処理を行うため、より見た目に違和感のない補完処理が行われる可能性が高い。そのため、前景らしい領域に補完対象領域が隣接するか否かにより判定を行うと効果的である。
- [0332] まず、補完対象フレーム t の補完対象領域がフレーム内の他オブジェクトに隣接しているか否かの判定を行う。この判定を行うにあたり、フレーム内の前景らしい領域の画素位置を知る必要があるため、何らかの方法により背景と前景とを分離し、前景画像 S を生成する。本形態では、背景前景分離手法を特に限定するものではないが、例えば以下の方法により分離処理を行うことが可能である。
- [0333] まず、映像シーケンス V_{orig} 内のキーフレーム F を k 枚用意する。なお、「キーフレーム」とは、前景背景分離処理を行う際に、基準として用いたい映像シーケンス内のフレームのことを指す。 k は、 $0 < k < n$ （ただし、 n は映像シーケンス V_{orig} の総フレーム数）を満たすものであれば任意の枚数を設定してよいが、実用的には映像シーケンス V_{orig} の総フレーム数の10%程度の枚数を設定するとよい。キーフレーム F の取得方法は、例えば、ランダムに取得する方法や、一定間隔で取得する方法が考えられるが、用途に応じて決定すればよい。次に、補完対象フレーム t とキーフレーム F を構成する全キーフレームとの差分画像を求め、二値化する。この処理により、前景と背景とを1、0（どちらを0にするか1にするかは任意である）の二値で表すことができる。この二値

画像の総数は、キーフレームの総数と同じく k 枚である。さらに、 k 枚すべての二値画像上で前景を示す画素のみを前景とし（具体的には、 k 枚の二値画像の論理積（AND）を取る）、前景画像 S を生成する。

[0334] 前景画像 S が生成されたら、補完対象フレーム t の補完対象領域から一定距離 d 以内に存在する、前景画像 S 上での前景を示す画素数をカウントし、その総数を C とする。 C が一定閾値 C_{th} よりも小さければ、補完対象領域の近傍に他オブジェクトが存在しないと判定し、 $R(t)=t$ としてフレーム内を参照領域としたコンプリーション処理を施す。そうでなければ、参照フレームを映像シーケンス V_{orig} 中のフレームから決定し、そのフレームを参照領域とした貼り付け合成処理を行う。なお、距離 d は0以上、かつ、画像の縦横の解像度より小さい値であれば任意の値を設定して構わないが、隣接する前景の有無を調べることが目的であるため、 $0 \leq d < 5[\text{pixel}]$ などの比較的小さい値を設定するとよい。特に、厳密に隣接するか否かを調べるのであれば、 $d=0$ とすればよい。また、閾値 C_{th} は、0以上255以下であれば任意の値を設定して構わないが、実用的には50～150程度に設定するのがよい。ここでは、映像シーケンス V_{orig} から参照フレームを決定する方法を特に限定するものではないが、例えば以下のような方法が挙げられる。

[0335] 補完対象フレーム t の時系列軸的に近傍であるフレーム（ t' とする）を順に探索し（時系列的な前を先に探索しても後ろを先に探索してもよい）、フレーム毎に前述した前景背景分離処理を探索対象のフレーム t' に対して行い、前景画像 S' を得る。さらに、補完対象フレーム t の補完対象領域から一定距離 d 以内に存在する、前景画像 S' 上での前景を示す画素数をカウントし、その総数を C とする。 C が一定閾値 C_{th} よりも小さければ、補完対象領域の近傍に他オブジェクトが存在しないと判定し、 $R(t)=t'$ としてフレーム t' を参照領域としたコンプリーション処理を施す。そうでなければ、前述の条件を満たすフレーム t' を発見するまで探索を続ける。

[0336] 第二の参照フレーム決定方法は、周辺との色特徴量の違いにより決定する方法である。これは、方針としては第一の参照フレーム決定方法と同様であ

るが、色特微量のみに基づいて、以下のように参照フレームおよび補完方法の決定を行う方法である。

[0337] まず、修復したいオブジェクトがフレーム内に存在する他のオブジェクトと位置的に重複しているか否かを色特微量に基づいて調査する。

[0338] あるフレーム t の補完対象領域 $R(t)$ 内に相当する画素位置における映像シーケンス $V_{orig}(t)$ の平均色を U_{in} とする。また、補完対象領域 $R(t)$ から一定距離 d 以内に存在する近傍画素の平均色を U_{out} とする。距離 d は上記の方法と同様に設定すればよい。 U_{in} , U_{out} は次式のように求められる。

[0339] [数8]

$$U_{in}(t) = \frac{\sum_{(x,y) \in \Omega} u(x,y,t)}{\sum_{(x,y) \in \Omega} 1},$$

$$U_{out}(t) = \frac{\sum_{(x,y) \in \bar{\Omega} \cap \Theta} u(x,y,t)}{\sum_{(x,y) \in \bar{\Omega} \cap \Theta} 1}$$

[0340] ただし、 Ω は参照領域 $R(t)$ のうちマスク情報 D が示す領域を示し、 Θ は参照領域 $R(t)$ のうちマスク情報 D が示す領域から一定距離 d 以内にある領域を示す。また、 $u(x, y, t)$ は $V_{orig}(x, y, t)$ の色情報（例えば、RGB）を示す。

[0341] こうして求めた平均色 U_{in} , U_{out} の差 U_{diff} があらかじめ定められた閾値 Th よりも小さければ、参照領域は現フレームのみとし、 $R(t)=t$ としてフレーム内を参照領域としたコンプリーション処理を行う。そうでない場合は、 U_{diff} が閾値 Th よりも小さくなる U_{out} を出力するフレーム t' のうち、最も現フレーム t に時系列軸上で近いフレームを選定し、 $R(t)=t'$ としてフレーム t' を参照領域とした貼り付け合成処理を行う。 Th は0以上255以下であれば任意の値を設定して構わないが、実用的には50～150程度に設定するとよい。 t' が複数フレーム存在する場合は、ランダムに決定する等によりいずれか一つのみ選択してもよいし、複数フレームを参照領域として保持してもよい。

[0342] 第三の参照フレーム決定方法は、セグメンテーションを用いる方法である。画像・映像内の画素を分割する手法として、参考文献11のようなセグメンテーション手法が知られている。セグメンテーションとは、画像を入力と

して、類似した色や構造に基づいて類似した領域を一つのまとまった領域として分割する技術である。このような手法を用いることにより、画像内で多くの画素を占めるセグメントを特定することができ、フレーム t における背景画素の画像内位置を知ることが可能である。例えば、シーケンスの先頭フレームの各セグメントに対し、はじめにユーザが手動でタグ付け（例えば、マウンド、選手、フェンス等）をし、残りのフレームではそのタグと同一セグメントに分類されたものを調べれば、該当セグメントが背景であるか前景であるかを知ることが可能である。

[0343] 〔参考文献 11〕 Chen, L. Cao, Y. Wang, J. Liu, and X. Tang, “Image Segmentation by MAP-ML Estimations,” IEEE Transactions on Image Processing (TIP), vol. 19, no. 9, pp. 2254-2264, 2010.

[0344] このセグメンテーションを利用し、あるフレーム t の補完対象領域 $R(t)$ から一定距離 d 以内に存在する近傍画素が背景であった場合にはフレーム内情報のみからの補完が容易であると判断して、 $R(t)=t$ としてフレーム内を参照領域としたコンプリーション処理を行う。そうでない場合には、他フレームの補完対象領域 $R(t)$ に相当する画素位置における画素が背景となるフレーム t' のうち、最も現フレーム t に時系列軸上で近いフレームを選定し、 $R(t)=t'$ としてフレーム t' を参照領域とした貼り付け合成処理を施す。距離 d は上記の方法と同様に設定すればよい。 t' が複数フレーム存在する場合は、ランダムに決定する等によりいずれか一つのみ選択してもよいし、複数フレームを参照領域として保持してもよい。

[0345] ステップ S 7 3 において、映像補完部 7 3 は、補完方法判定部 7 2 から受け取った修復方法 P がパッチ探索に基づくコンプリーション処理を示している場合、補完方法判定部 7 2 から映像シーケンス V_{orig} 、マスク情報 D 、および参照領域 R を受け取り、映像シーケンス V_{orig} 中のマスク情報 D が示す補完対象領域を、参照領域 R を用いてパッチ探索に基づくコンプリーション処理により補完し、補完後の映像シーケンス V_{out} を生成する。補完後の映像シーケンス V_{out} は、映像出力部 7 4 へ送られる。

[0346] ステップS 7 4 において、映像補完部 7 3 は、補完方法判定部 7 2 から受け取った修復方法Pが貼り付け合成処理を示している場合、補完方法判定部 7 2 から映像シーケンス V_{orig} 、マスク情報D、および参照領域Rを受け取り、映像シーケンス V_{orig} 中のマスク情報Dが示す補完対象領域を、参照領域Rを用いて貼り付け合成処理により補完し、補完後の映像シーケンス V_{out} を生成する。補完後の映像シーケンス V_{out} は、映像出力部 7 4 へ送られる。

[0347] ステップS 7 5 において、映像出力部 7 4 は、映像補完部 7 3 から補完後の映像シーケンス V_{out} を受け取り、外部の映像出力手段を通して出力する。映像出力手段としては、例えば、映像提示装置の映像出力端子に接続されたヘッドマウントディスプレイや、液晶ディスプレイ、プロジェクタなどが挙げられる。

[0348] [第十実施形態の実験結果]

図 4 8 に実際の計算時間に関する実験結果を示しながら、従来技術（フレーム内コンプリーション処理および貼り付け合成処理）単体による補完と比較して、特に計算時間・画質が向上することを説明する。なお、図 4 8 中で（１）～（４）で示される結果は、以下のとおりである。（１）（He+ 2014）は、参考文献 9 に記載されたフレーム内コンプリーション処理に基づく実験結果である。（２）（貼り付け合成）は、貼り付け合成処理による実験結果である。（３）（Newson+ 2014）は、参考文献 1 0 に記載された三次元パッチに基づく動画コンプリーション処理による実験結果である。（４）（本技術）は、第十実施形態の映像提示装置を用いた実験結果である。また、図 4 8 中で（Ａ）（Ｂ）で示される結果は、それぞれ以下のような映像シーケンスに対する実験結果である。（Ａ）シーケンス内に修復対象領域であるボールが、前景領域であるピッチャーに隣接するフレームが含まれており、貼り付け合成処理が困難であるが、一方で背景が一様なためフレーム内コンプリーションを施した結果に違和感が出にくい映像である。（Ｂ）シーケンス内に修復対象領域であるボールが、前景領域であるピッチャーに隣接または重複するフレームが含まれておらず貼り付け合成処理が比較的容易であるが

、背景にフェンスなどの複雑なテクスチャが含まれるためフレーム内コンプリーションが難しい映像である。

[0349] 図48(A)の計算時間を実験により算出した際の、画質について以下に説明する。(A)は補完対象領域であるボールが人物(ピッチャー)に隣接するフレームを含んでおり、該当フレームでは貼り付け合成処理による補完結果が良好に行なえないが、一方で背景が一様なためフレーム内コンプリーションでも違和感が少ないシーンを用いた実験結果である。背景がマウンドの土などで複雑なテクスチャを含まない場合、フレーム内コンプリーションを施した際の違和感は少ないため、このシーンでは(1)のフレーム内コンプリーション処理が有効である。一方で、(2)の貼り付け合成処理では除去対象領域が前景に隣接した場合、他フレームからの画素の貼り付けにより、ピッチャーの身体の一部がマウンドの土の領域の画素で補完されてしまい、著しく画質が悪化してしまう。(3)の動画コンプリーション処理も有効であるが、計算時間が膨大である(図48参照)。(4)の本技術は、(1)および(2)のどちらが適切な補完手法であるかフレーム毎に判定しており、(2)の貼り付け合成処理において著しく画質が劣化したフレームに関しては(1)のフレーム内コンプリーション処理が適切であると判定され、それ以外のフレームではより高速な(2)が選択された。そのため、(2)と比較して大幅に画質を向上することができた。

[0350] また、図48(B)の計算時間を実験により算出した際の、画質について以下に説明する。(B)は、補完対象領域であるボールが人物に隣接せず貼り付け合成処理を用いた補完でも違和感が少ないが、一方で背景が複雑なテクスチャ構造を含んでおりフレーム内コンプリーションが失敗しやすいシーンを用いた実験結果である。背景がフェンスであるなどの複雑なテクスチャ構造の場合、(1)のフレーム内コンプリーションでは背景テクスチャの補完に失敗しやすく、フェンスの線形構造が欠損して補完されてしまうフレームがあった。そのため、時系列一貫性が保たれず動画シーケンスとして視聴した際に著しく違和感を生じる原因となる。一方で、(2)の貼り付け合成

処理では他フレームに存在する正しいテクスチャ構造を用いることができるため、違和感の少ない合成が可能であった。(A)の例と同じく、(3)の動画コンプリーション処理も有効であるが、計算時間が膨大である(図48参照)。(4)の本技術は、補完対象領域であるボールが前景(ピッチャー)に隣接していない場合には(2)の方法を用いて補完を行うため、(1)の手法においてフェンスの線形構造が欠損されて補完されたフレームに関しても(2)の他フレームから補完を行う手法が適切であると判定された。したがって、(1)と比較して大幅に画質が向上した。

[0351] 図48(A)(B)の例では、それぞれ貼り付け合成処理またはコンプリーション処理のいずれかが失敗し、結果として違和感のある補完の原因となっている。本技術では、この2つの手法をフレームに応じて自動判定により使い分けることで、従来補完が困難であったシーンであっても自動的に良好な補完を行うことが可能である。

[0352] 図48に、(A)(B)に示したシーンそれぞれにおける実行時間の比較結果を示す。(2)の貼り付け合成処理は、いずれの場合でも最も実行時間が少ないが、上記で述べたように画質の面で不十分な場合がある。一方、(4)に示す本技術は、画質を向上させつつ計算時間も(2)に次いで少ない((A)では(1)の31.1%の実行時間、(B)では(1)の2.5%の実行時間)。したがって、本技術は計算時間を短縮しつつ画質も向上させられるという点で効果的であると言える。

[0353] 上記のように構成することにより、第十実施形態によれば、補完対象領域の近傍に前景オブジェクトがあればパッチ探索に基づくコンプリーション処理により補完し、補完対象領域の近傍に前景オブジェクトがなければ貼り付け合成処理により補完するため、不要な映り込みなどにより補完が必要な領域を含む映像を効率的に補完することができる。

[0354] [第十一実施形態]

近年、360度のパノラマ画像を撮影できるカメラ(以下、「全天球カメラ」という。)が普及し始めている。全天球カメラによって撮影されたパノ

ラマ画像（以下、「全天球画像」という。）は、所望の視点位置に全天球カメラを設置することで撮影することができる。しかしながら、競技中の競技者の邪魔となるためサッカーコートやバスケットコートなどの競技用コートの中には全天球カメラを設置することができない。そのため、競技用コートの中の所望の視点位置における競技中の全天球画像を撮影することができない。

[0355] そこで、全天球カメラを設置することのできない場所に仮想的な視点である仮想視点を設定して、この仮想視点において全天球カメラで撮影したかのような全天球画像を、コートの外側に設置された複数のカメラによって撮影された画像を合成することによって得る技術が提案されている（例えば、参考文献5参照）。以下の説明において、仮想視点における全天球画像を、仮想全天球画像と記載する。

[0356] 仮想全天球画像を複数のカメラによって撮影された画像の合成によって得るシステムの具体例について説明する。

[0357] 図61は、従来システムにおいて仮想全天球画像を得るためのシステムを示す図である。図61に示すように、画像処理システム900は、全天球カメラ92と、複数のカメラ93-1、93-2、93-3、・・・、93-N（以下、「カメラ群93」という。）（Nは4以上の整数）と、画像処理装置94と、表示装置95とを備える。画像処理システム900は、競技用コート90内に仮想視点91を設定した場合に、競技用コート90外に設置したカメラ群93によって撮影された画像の合成によって仮想視点91における仮想全天球画像を得る。

[0358] 全天球カメラ92は、全天球画像を撮影するカメラである。全天球カメラ92は、競技が行われる前のタイミングで競技用コート90内の仮想視点91の位置に設置される。全天球カメラ92は、仮想視点91の位置から、仮想全天球画像の背景となる画像（以下、「背景画像」という。）を撮影する。全天球カメラ92で撮影された背景画像は、画像処理装置94に入力されて蓄積される。このように、画像処理装置94は、予め背景画像を蓄積する

。

[0359] 競技用コート90の周囲には、カメラ群93が設置されている。カメラ群93の各カメラ93-1、93-2、93-3、・・・、93-Nは、それぞれ仮想視点91を含む画角となるように競技用コート90の周囲に設置されている。カメラ群93は、仮想視点91を含む領域を撮影する。画像処理装置94は、カメラ群93の各カメラ93-1、93-2、93-3、・・・、93-Nによって撮影された画像に対して画像処理を施して、背景画像に画像処理後の画像を合成して仮想全天球画像を生成する。表示装置95は、液晶ディスプレイ、有機EL (Electro Luminescence) ディスプレイ、CRT (Cathode Ray Tube) ディスプレイ等の画像表示装置である。表示装置95は、画像処理装置94で生成した仮想全天球画像を表示する。

[0360] 次に、画像処理システム900における画像処理の具体例について図62を用いて説明する。

[0361] 図62は、画像処理システム900における画像処理の流れを説明するための図である。図62(A)は、背景画像920の具体例を示す図である。背景画像920には、仮想視点91を中心として全方位(360度)の被写体が撮影されている。背景画像920は、競技用コート90内に人物がいない状態で撮影される画像であるので競技用コート90内には人物が撮影されない。

[0362] 図62(B)は、各カメラ93-1、93-2及び93-3で撮影された画像を示す図である。図62(B)には、左からカメラ93-1で撮影された画像921と、カメラ93-2で撮影された画像922と、カメラ93-3で撮影された画像923とが示されている。画像処理装置94は、画像921~923のそれぞれから仮想視点91を含む領域9211、9221、9231を抽出する。画像処理装置94は、抽出した領域9211、9221、9231の画像に対して、画像処理を行うことで背景画像920に合成可能な部分画像9211a、9221a、9231aを生成する。

[0363] 画像処理装置94は、背景画像920に対して部分画像9211a、92

21a、9231aを合成することによって仮想全天球画像924を生成する。図62(C)は、画像処理装置94が生成する仮想全天球画像924の例を示す図である。図62(C)に示すように、仮想全天球画像924の所定の領域には部分画像9211a、9221a、9231aが合成されている。そのため、仮想全天球画像924として、競技用コート90上に被写体（例えば、人物）が撮影されている画像が生成される。従来の画像処理システム900は、合成に用いているカメラ群93の光学中心及び仮想視点91において想定する全天球カメラ92の光学中心がそれぞれ異なる。そのため、合成された仮想全天球画像924は幾何学的に正しくない画像を含む。これを防ぐためには、画像処理装置94は、部分画像9211a、9221a、9231aを、仮想視点91からの距離を示す奥行き的一点で整合性が保たれるよう画像処理を行い背景画像920に合成する必要がある。

[0364] しかしながら、整合性が保たれる奥行きに存在せずに別の奥行に存在している被写体（例えば、人物）の部分画像を背景画像920に合成する場合には、画像処理により奥行きの整合性を保つことができない。このような奥行きに整合性のない被写体は、仮想全天球画像924において、その画像が多重像になってしまったり、消失したりする現象が発生する。このように、被写体の奥行きが不明な場合には、適切な合成が出来なくなってしまうという問題があった。なお、このような問題は、複数の撮影装置のそれぞれで撮影された画像を合成する場合全てに共通する問題である。

[0365] 上記事情に鑑み、第十一実施形態は、複数の撮影装置のそれぞれで撮影された画像の最適な奥行きを決定することができる技術の提供を目的としている。

[0366] 第十一実施形態の一態様は、所定の位置を含む領域が撮影範囲となるように前記所定の位置を含む領域の周囲に設置された少なくとも2つの撮像装置それぞれが撮影した複数の入力画像を用いて、前記所定の位置を仮想的な視点である仮想視点として前記仮想視点からの距離を表す複数の異なる奥行きの合成画像を生成し、生成した複数の合成画像それぞれからコントラストを

算出し、算出した前記コントラストに基づいて複数の合成画像のうちコントラストが最も大きい合成画像の奥行きを各入力画像の最適奥行きに決定する決定部を備える最適奥行き決定装置である。

[0367] 以下、第十一実施形態を、図面を参照しながら説明する。

[0368] 図49は、第十一実施形態における画像処理システム800のシステム構成を示す図である。

[0369] 画像処理システム800は、全天球カメラ92、複数のカメラ93-1～93-M（Mは2以上の整数）及び画像処理装置80を備える。なお、以下の説明では、カメラ93-1～93-Mについて特に区別しない場合には、カメラ93と記載する。

[0370] 全天球カメラ92は、撮影対象領域81内の仮想視点82の位置に設置される。撮影対象領域81は、例えばサッカーコートやバスケットコートなどの競技用コートなどである。仮想視点82は、所定の領域（本実施形態では、撮影対象領域81）内に仮想的に設定された視点である。全天球カメラ92は、仮想視点82の位置における全天球画像を撮影する。本実施形態における全天球画像は、仮想視点82を中心として撮影対象領域81全体を含む。全天球カメラ92による処理は、画像処理装置80による処理の開始前に行われる。全天球カメラ92は、撮影した全天球画像を背景画像として画像処理装置80に出力する。

[0371] M台のカメラ93-1、93-2、・・・、93-Mは、撮影対象領域81の外側に設けられ、画像を動画（映像）で撮影するカメラであり、仮想視点82を含む領域を撮影する。M台のカメラ93-1、93-2、・・・、93-Mのそれぞれで撮影された動画は、複数フレームの画像により構成される。図49に示すように、カメラ93-1には仮想視点82の位置上を通過する光線83-1が入力され、カメラ93-2には仮想視点82の位置上を通過する光線83-2が入力される。以下、カメラ93に入力される光線を実光線と記載する。図49では示していないが、カメラ93は撮影対象領域81の周囲に設置される。つまり、カメラ93は、それぞれ仮想視点82

を含む画角となるように撮影対象領域 8 1 の周囲を取り囲むように設置される。図 4 9 において M は、2 以上の整数であり、同程度の画質の仮想全天球画像を得ようとするのであれば撮影対象領域 8 1 が大きいほど大きな値となる。また、撮影対象領域 8 1 の大きさが同じであれば、M の値が大きい程、合成領域（仮想全天球画像において、M 台のカメラ 9 3 からの画像を合成した領域）の面積が大きくなり、あるいは合成領域の大きさが同じであれば合成領域における画質が向上する仮想全天球画像の画質を高いものにしようとするほど大きな値となる。

[0372] 画像処理装置 8 0 は、M 台のカメラ 9 3 - 1、9 3 - 2、・・・、9 3 - M のそれぞれで撮影されたそれぞれの動画から入力画像を事前を取得する。撮影されたそれぞれの動画は複数フレームの画像で構成されており、本実施形態における画像処理装置 8 0 は処理対象となるフレームの画像を入力画像として取得する。画像処理装置 8 0 は、全天球カメラ 9 2 によって撮影された全天球画像と、M 台のカメラ 9 3 - 1、9 3 - 2、・・・、9 3 - M のそれぞれで撮影された動画からそれぞれ取得された入力画像とに基づいて仮想全天球画像を生成する。具体的には、画像処理装置 8 0 は、入力画像から最適な奥行きを決定し、決定した奥行きに基づいて入力画像から領域を抽出し、抽出した領域を全天球画像に重畳することによって仮想全天球画像を生成する。

[0373] 第十一実施形態では、画像処理システム 8 0 0 に 2 台のカメラ 9 3（カメラ 9 3 - 1 及び 9 3 - 2）が隣接して備えられている場合を例に説明する。また、以下の説明では、隣接するカメラ 9 3 の組（例えば、カメラ 9 3 - 1 と 9 3 - 2 との組）をカメラペアと記載する。

[0374] 画像処理装置 8 0 は、バスで接続された CPU（Central Processing Unit）やメモリや補助記憶装置などを備え、画像処理プログラムを実行する。画像処理プログラムの実行によって、画像処理装置 8 0 は、入力画像記憶部 8 0 1、合成情報記憶部 8 0 2、決定部 8 0 3、最適奥行き記憶部 8 0 4、背景画像記憶部 8 0 5、画像合成部 8 0 6 を備える装置として機能する。なお

、画像処理装置 80 の各機能の全て又は一部は、ASIC (Application Specific Integrated Circuit) や PLD (Programmable Logic Device) や FPGA (Field Programmable Gate Array) 等のハードウェアを用いて実現されてもよい。また、画像処理プログラムは、コンピュータ読み取り可能な記録媒体に記録されてもよい。コンピュータ読み取り可能な記録媒体とは、例えば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である。また、画像処理プログラムは、電気通信回線を介して送受信されてもよい。

[0375] 入力画像記憶部 801 は、磁気ハードディスク装置や半導体記憶装置などの記憶装置を用いて構成される。入力画像記憶部 801 は、各カメラ 93 を識別するためのカメラ ID に関連付けて、カメラ 93 毎の入力画像を時系列順に記憶する。入力画像は、撮影時刻及び動画の画像データを含む。

[0376] 合成情報記憶部 802 は、磁気ハードディスク装置や半導体記憶装置などの記憶装置を用いて構成される。合成情報記憶部 802 は、合成情報テーブルを記憶する。合成情報テーブルは、背景画像に対して画像を重畳するために用いられる情報（以下、「合成情報」という。）を表すレコード（以下、「合成情報レコード」という。）によって構成される。

[0377] 図 50 は、合成情報テーブルの具体例を示す図である。

[0378] 合成情報テーブルは、合成情報レコードを複数有する。合成情報レコードは、カメラ ID 及び合成情報の各値を有する。カメラ ID の値は、カメラ 93 を識別するための識別情報を表す。例えば、図 50 におけるカメラ ID “C₁” で識別されるカメラ 93 はカメラ 93-1 であり、カメラ ID “C₁” で識別されるカメラ 93 はカメラ 93-1 である。

[0379] 合成情報の値は、同じ合成情報レコードのカメラ ID で識別されるカメラ 70 で撮影された画像（入力画像）から生成される画像を背景画像に重畳するために用いられる情報を表す。合成情報の具体例として、奥行き、抽出領域情報及び変換情報がある。

- [0380] ある合成情報レコードの奥行き値は、カメラIDで識別されるカメラ93の仮想視点82からの距離を表す。図50に示す例では、合成情報テーブルには、カメラID毎にN段階の奥行きが登録されている。なお、Nは2以上の整数である。奥行き1～奥行きNは、仮想視点82から撮影対象領域81の端までの間で設定される。設定される奥行きの数は、最適奥行きの決定精度がある程度保障される数であることが好ましい。
- [0381] ある合成情報レコードの抽出領域情報の値は、カメラIDで識別されるカメラ93で撮影された画像（入力画像）から抽出する領域（以下、「抽出領域」という。）に関する情報を表す。抽出領域情報の具体例として、左上座標、幅及び高さがある。左上座標は、抽出領域の左上の座標を表す。幅は、抽出領域の幅を表す。高さは、抽出領域の高さを表す。なお、幅及び高さは、抽出領域の左上座標を基準とし、かつ、仮想視点82を含む範囲に設定される。抽出領域は、背景画像に重畳された画像上において、隣接するカメラ93の画像の間に隙間ができないような領域に設定されることが望ましい。なお、以下の説明では、抽出領域情報に応じて入力画像から抽出された部分領域の画像を部分領域画像と記載する。
- [0382] ある合成情報レコードの変換情報の値は、抽出領域情報に応じて抽出された部分領域画像を部分画像に変換するための情報を表す。部分画像は、部分領域画像を背景画像の対応領域に違和感なく重畳するために、部分領域画像に対して変換情報に応じて拡大、縮小、回転等の変形処理を行うことによって生成される。この変形処理は、例えば、画像に対してアフィン変換を施すことによって行う。画像に対してアフィン変換を施す場合の変換情報は、例えばアフィン変換行列である。以下、部分領域画像に対して行う変形処理としてアフィン変換を用いる例を示すが、変形処理はアフィン変換に限定される必要はなく、変換情報に応じて拡大、縮小、回転等による画像の変換を行う処理であればどのような処理であってもよい。なお、アフィン変換行列には、背景画像において部分画像を重畳する領域を示す情報（以下、「重畳情報」という。）が含まれる。

- [0383] アフィン変換行列は、以下に示す方法により予め取得して合成情報記憶部 802 に記憶される。例えば、仮想視点 82 から複数種類の距離（奥行き）（図 50 の例では、奥行き 1 ～奥行き N）の位置それぞれに格子模様のチェスボードを設置して、仮想視点 82 に設置した全天球カメラ 92 で撮影したチェスボードを含む画像と、カメラ 93 で撮影したチェスボードを含む画像とを比較する。そして、チェスボードの各格子について、全天球カメラ 92 で撮影した画像中のチェスボードの格子と、カメラ 93 で撮影した画像中のチェスボードの格子が対応するように画像を変換するアフィン変換行列を求める。このようにして、チェスボードを設置した奥行き 1 ～奥行き N それぞれに対応したアフィン変換行列を求める。
- [0384] 図 50 に示されるように、合成情報テーブルには、カメラ ID 毎に N 段階の奥行きが登録され、各奥行き 1 ～N それぞれに対して変換情報が登録されている。図 50 において、合成情報テーブルの最上段に登録されている合成情報レコードは、カメラ ID の値が “C₁”、奥行きの値が “奥行き 1”、左上座標の値が “(A, B)”、幅の値が “C”、高さの値が “D”、変換情報の値が “A_{1j1}” である（j は 1 以上の整数）。すなわち、カメラ ID “C₁” で識別されるカメラ 93-1 において奥行き 1 の場合には、入力画像から左上座標 (A, B)、幅 C、高さ D で表される部分領域を抽出し、抽出された部分領域に対して “A_{1j1}” の変形処理を施すことが表されている。
- [0385] 図 49 に戻って、画像処理装置 80 の説明を続ける。
- [0386] 決定部 803 は、入力画像記憶部 801 に記憶されているカメラペアの入力画像と、合成情報記憶部 802 に記憶されている合成情報テーブルとを入力とする。決定部 803 は、入力されたカメラペアの入力画像と合成情報テーブルとに基づいて、各カメラ 93 の各入力画像の最適奥行きを決定する。ここで、最適奥行きとは、入力画像に撮像されている被写体に関して多重像や消失が他の奥行きに比べて視認しにくい奥行きを表す。
- [0387] 次に、決定部 803 の具体的な処理について説明する。
- [0388] まず決定部 803 は、合成情報テーブルの 1 つの奥行き（例えば、奥行き

1) に対する抽出領域情報に基づいて、入力画像それぞれから部分領域を抽出することによって各入力画像の部分領域画像を生成する。次に、決定部 803 は、生成した各入力画像の部分領域画像に対して、合成情報テーブルの 1 つの奥行き（例えば、奥行き 1）に対する変換情報のアフィン変換行列に基づく変形処理を行うことによって各入力画像の部分画像を生成する。次に、決定部 803 は、生成した各入力画像の部分画像を用いて、各入力画像の最適な奥行きを決定するための合成画像（以下、「奥行き決定用合成画像」という。）を生成する。例えば、決定部 803 は、変換情報に含まれる重畳情報に基づいて奥行き決定用合成画像を生成する。決定部 803 は、生成した奥行き決定用合成画像の重複領域に対するコントラストを算出する。本実施形態では、コントラストを輝度差で求める場合を例に説明する。なお、コントラストを輝度差で求める場合には後述する式 1 が用いられる。決定部 803 は、上記の処理を全ての奥行き 1 ～ N に対して行う。そして、決定部 803 は、全ての奥行き 1 ～ N において算出したコントラストの中でコントラストが最も大きい奥行き決定用合成画像の奥行きをカメラペアの各入力画像の最適奥行きに決定する。決定部 803 は、決定したカメラペアの各入力画像の最適奥行きから各カメラ 93 の各入力画像の最適奥行きを決定する。なお、決定部 803 は、本実施形態のようにカメラペアが 1 つの場合には、カメラペアの各入力画像の最適奥行きを各カメラ 93 の各入力画像の最適奥行きに決定してもよい。カメラペアが複数ある場合の説明は後述する。決定部 803 は、決定した各カメラ 93 の各入力画像の最適奥行き情報を最適奥行き記憶部 804 に記憶させる。決定部 803 は、以上の処理を各時刻におけるカメラペアの入力画像に対して行う。

[0389] 最適奥行き記憶部 804 は、磁気ハードディスク装置や半導体記憶装置などの記憶装置を用いて構成される。最適奥行き記憶部 804 は、最適奥行き情報テーブルを記憶する。最適奥行き情報テーブルは、各カメラ 93 の各入力画像の最適奥行きに関する情報を表すレコード（以下、「最適奥行き情報レコード」という。）によって構成される。

[0390] 図51は、最適奥行き情報テーブルの具体例を示す図である。

[0391] 最適奥行き情報テーブルは、最適奥行き情報レコードを複数有する。最適奥行き情報レコードは、時刻及び最適奥行きの各値を有する。時刻の値は、入力画像が撮影された時刻を表す。ある最適奥行き情報レコードの最適奥行きの値は、ある時刻に撮影された各カメラ93の入力画像の最適奥行きを表す。

[0392] 図51に示される例では、最適奥行き情報テーブルには各時刻における最適奥行きの情報が登録されている。図51において、最適奥行き情報テーブルの最上段に登録されている最適奥行き情報レコードは、時刻の値が“ t ”、カメラ93-1の入力画像の最適奥行きの値が“ $d1(t)$ ”、カメラ93-2の入力画像の最適奥行きの値が“ $d2(t)$ ”である。すなわち、時刻 t にカメラ93-1及び93-2によって撮影された画像（入力画像）それぞれの最適奥行きが“ $d1(t)$ ”、“ $d2(t)$ ”であることが表されている。

[0393] 背景画像記憶部805は、磁気ハードディスク装置や半導体記憶装置などの記憶装置を用いて構成される。背景画像記憶部805は、全天球カメラ92によって撮影された全天球画像を背景画像として記憶する。

[0394] 画像合成部806は、入力画像記憶部801に記憶されている各カメラ93の入力画像と、合成情報記憶部802に記憶されている合成情報テーブルと、最適奥行き記憶部804に記憶されている最適奥行き情報テーブルと、背景画像記憶部805に記憶されている背景画像とを入力とする。画像合成部806は、入力された各カメラ93の入力画像と、合成情報テーブルと、最適奥行き情報テーブルと、背景画像とに基づいて仮想全天球画像を生成する。具体的には、画像合成部806は、最適奥行き情報テーブルを参照して、ある時刻（例えば、時刻 t ）におけるあるカメラ93（例えば、カメラ93-1）の入力画像の最適な奥行きの情報を取得する。画像合成部806は、取得したカメラ93の入力画像の最適な奥行きに基づいて、合成情報テーブルからあるカメラ93（例えば、カメラ93-1）の最適な奥行きに対応する抽出領域情報を取得する。画像合成部806は、取得した抽出領域情報

に基づいて、入力画像から部分領域を抽出することによって部分領域画像を生成する。

[0395] 画像合成部806は、合成情報テーブルからあるカメラ93（例えば、カメラ93-1）の最適な奥行きに対応する変換情報を取得する。画像合成部806は、生成した部分領域画像に対して、取得した変換情報のアフィン変換行列に基づく変形処理を行うことによって入力画像の部分画像を生成する。画像合成部806は、部分画像生成までの処理をある時刻における各カメラ93の入力画像に対して行う。その後、画像合成部806は、アフィン変換行列に含まれる重畳情報に基づいて、生成した部分画像を背景画像に重畳することで仮想全天球画像を生成する。より具体的には、画像合成部806は、部分画像の画素値で、背景画像上の部分画像を重畳する領域の画素値を置き換えることによって仮想全天球画像を生成する。画像合成部806は、仮想全天球画像生成の処理を各時刻の入力画像に対して行うことによって各時刻における仮想全天球画像を生成する。画像合成部806は、生成した仮想全天球画像を出力する。

[0396] 以下の説明では、画像合成部806による上記の全ての処理を画像合成処理ステップと記載する。

[0397] 図52は、コントラストの算出結果の具体例を示す図である。

[0398] 図52において、縦軸はコントラストを表し、横軸は奥行きを表す。奥行きの値が0に近いほど仮想視点82からの距離が近い奥行きを表し、奥行きの値が高くなるほど仮想視点82からの距離が遠い奥行きを表す。図52では、円990で示される奥行きdのコントラストが最も大きい。この場合、決定部803は、奥行きdをカメラペアの各入力画像の最適奥行きに決定する。

[0399] 図53は、画像処理装置80の処理の流れを示すフローチャートである。

[0400] 決定部803は、カメラペアの同時刻における入力画像を入力画像記憶部801から読み出す（ステップS101）。次に、決定部803は、読み出した入力画像と、合成情報記憶部802に記憶されている合成情報テーブル

とに基づいて、１つの奥行きに対する変換情報を用いて奥行き決定用合成画像を生成する（ステップＳ１０２）。１つの奥行きの選択方法は、処理が行われていない奥行きを選択できればどのような方法であってもよい。例えば、決定部８０３は、仮想視点８２に近い奥行きから順番に奥行きを選択してもよいし、ランダムに奥行きを選択してもよいし、その他の方法で奥行きを選択してもよい。次に、決定部８０３は、生成した奥行き決定用合成画像の重複領域のコントラストを以下の式１に基づいて算出する（ステップＳ１０３）。なお、式１はコントラストの求め方の一例である。また、式１の L_{min} は画像中の重複領域の最小輝度値を表し、 L_{max} は画像中の重複領域の最大輝度値を表す。

[0401] [数9]

$$\text{コントラスト} = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}} \quad \dots(\text{式 } 1)$$

[0402] その後、決定部８０３は、全ての奥行きに対してコントラストを算出する処理を行ったか否か判定する（ステップＳ１０４）。全ての奥行きに対してコントラストを算出する処理を行っていない場合（ステップＳ１０４－ＮＯ）、決定部８０３はステップＳ１０２以降の処理を繰り返し実行する。

[0403] 一方、全ての奥行きに対してコントラストを算出する処理を行った場合（ステップＳ１０４－ＹＥＳ）、決定部８０３は算出した全ての奥行きに対するコントラストに基づいてカメラペアの各入力画像の最適奥行きを決定する（ステップＳ１０５）。

[0404] なお、以下の説明では、図５３のステップＳ１０２からステップＳ１０５までの処理をカメラペア奥行き決定処理と記載する。

[0405] 以上のように構成された画像処理装置８０によれば、複数の撮影装置のそれぞれで撮影された画像の最適な奥行きを決定することが可能になる。以下、この効果について詳細に説明する。

[0406] 画像処理装置８０は、カメラペアの入力画像を用いて全ての奥行きに対する奥行き決定用合成画像を生成する。次に、画像処理装置８０は、生成した全ての奥行きに対する奥行き決定用合成画像それぞれからコントラストを算

出し、算出したコントラストが最も大きい奥行き決定用合成画像を選択する。コントラストが低い場合には画像がぼやけている可能性が高い。つまり、合成による多重像や消失が発生している画像である可能性が高い。したがって、画像処理装置80は、コントラストが最も大きい奥行き決定用合成画像を選択することによって、多重像や消失が発生していない可能性の高い画像を選択することができる。そのため、複数の撮影装置のそれぞれで撮影された画像の最適な奥行きを決定することが可能になる。

[0407] <第十一実施形態の変形例>

画像処理装置80における入力画像記憶部801、合成情報記憶部802、決定部803及び最適奥行き記憶部804は、最適奥行き決定装置として構成されてもよい。

[0408] 本実施形態では、カメラペアが1つの場合について説明したが、カメラペアが複数ある場合には各カメラ93には2つの最適奥行きが存在することになる。例えば、あるカメラ C_i (i は2以上の整数)に関しては (C_{i-1}, C_i) のペアと (C_i, C_{i+1}) のペアがあり、それぞれのカメラペアに関して最適奥行きが決定される。そこで、決定部803は、カメラペアが複数ある場合にはそれぞれのペア(例えば、 (C_{i-1}, C_i) のペアと (C_i, C_{i+1}) のペア)で決定された最適奥行きに基づいてカメラ C_i の最適奥行きを決定する。この方法としては、2つの最適奥行きのうち仮想視点82に近い奥行をカメラ C_i の最適奥行きと決定する方法などがある。決定部803は、この処理を全てのカメラペアに対して行うことによってカメラペアが複数ある場合においても各カメラ93の各入力画像の最適奥行きを決定することができる。そして、決定部803は、決定した各カメラ93の各入力画像の最適奥行きの情報を最適奥行き記憶部804に記憶させる。その後、画像合成部806は、画像合成処理ステップを実行してもよいし、合成画像の生成指示が入力されたタイミングで画像合成処理ステップを実行してもよいし、その他のタイミングで画像合成処理ステップを実行してもよい。

[0409] また、カメラペアを選ぶ際に、 (C_{i-1}, C_i) のペアと (C_{i+1}, C_{i+2}) の

ペアとするなど1台のカメラ93には2つの最適奥行きが存在しないようにする方法をとってもよい。

[0410] [第十二実施形態]

第十二実施形態では、第十一実施形態よりも処理の高速化を実現する。具体的には、第十一実施形態ではN個の奥行き全てに対してコントラストを算出する処理が行なわれているのに対して、第十二実施形態では、画像処理装置は、第十一実施形態よりも少ない奥行きに基づいてカメラペアの各入力画像の最適奥行きを決定する。より具体的には、第十二実施形態では、画像処理装置は、処理を繰り返す度に奥行き候補群に含まれる各奥行きの間隔が狭くなるように新たな奥行き候補群を作成し、所定の回数繰り返した時点の最終的な奥行き候補群に含まれる奥行きの中から最適な奥行きを決定する。また、第十二実施形態では、第十一実施形態と同様に、画像処理システム800に2台のカメラ93（カメラ93-1及び93-2）がカメラペアとして備えられている場合を例に説明する。

[0411] 図54は、第十二実施形態における画像処理装置80aの機能構成を表す概略ブロック図である。画像処理装置80aは、バスで接続されたCPUやメモリや補助記憶装置などを備え、画像処理プログラムを実行する。画像処理プログラムの実行によって、画像処理装置80aは、入力画像記憶部801、合成情報記憶部802、決定部803a、最適奥行き記憶部804、背景画像記憶部805、画像合成部806を備える装置として機能する。なお、画像処理装置80aの各機能の全て又は一部は、ASICやPLDやFPGA等のハードウェアを用いて実現されてもよい。また、画像処理プログラムは、コンピュータ読み取り可能な記録媒体に記録されてもよい。コンピュータ読み取り可能な記録媒体とは、例えば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である。また、画像処理プログラムは、電気通信回線を介して送受信されてもよい。

[0412] 画像処理装置80aは、決定部803に代えて決定部803aを備える点

で画像処理装置 80 と構成が異なる。画像処理装置 80 a は、他の構成については画像処理装置 80 と同様である。そのため、画像処理装置 80 a 全体の説明は省略し、決定部 803 a について説明する。

[0413] 決定部 803 a は、入力画像記憶部 801 に記憶されているカメラペアの入力画像と、合成情報記憶部 802 に記憶されている合成情報テーブルとを入力とする。決定部 803 a は、入力されたカメラペアの入力画像と合成情報テーブルとに基づいて、各カメラ 93 の各入力画像の最適奥行きを決定する。

[0414] 次に、決定部 803 a の具体的な処理について説明する。

[0415] まず決定部 803 a は、N 個の奥行きのうち特定の奥行き（例えば、S 個の奥行き）を選択する。ここで、特定の奥行きの選択方法としては、等間隔（ N/S 個置き）に奥行きを選択する方法などが挙げられる。以下、このような N 個の奥行きのうち特定の奥行きを選択する処理を初期処理と記載する。初期処理で選択された奥行きを初期奥行き候補群と記載する。次に、決定部 803 a は、合成情報テーブルを参照して、初期奥行き候補群に含まれる奥行きそれぞれについて奥行き決定用合成画像を生成し、生成した奥行き決定用合成画像のコントラストを算出する。奥行き決定用合成画像の生成及びコントラストの算出に関する処理については第 11 実施形態と同様である。この処理によって、決定部 803 a は、初期奥行き候補群に含まれる各奥行きの奥行き決定用合成画像の生成及びコントラストの算出を行う。そして、決定部 803 a は、算出したコントラストの中でコントラストが最も大きい奥行きを含む新たな奥行き候補群を作成する。なお、決定部 803 a は、新たな奥行き候補群を作る際には、コントラストが最も大きかった奥行きが中心になるような範囲から新たな奥行き候補群を作成する。中心になるような範囲は、予め定められる。

[0416] 以下、コントラストが最も大きい奥行きを含む新たな奥行き候補群を作成する処理について説明する。この説明では、コントラストが最も大きかった奥行きが中心になるような範囲の定め方の一例を示しながら新たな奥行き候

補群を作成する処理について説明する。一例として、5個の奥行き（ $N = 5$ 。S1、S2、S3、S4、S5（ $S1 < S2 < S3 < S4 < S5$ ）の5つの奥行き）があるとする。

[0417] 決定部803aは、コントラストが最も大きい奥行きと、コントラストが最も大きい奥行きに隣接する奥行きとの間において所定の数 T の奥行きを選択する。例えば、コントラストが最も大きい奥行きがS3であれば、隣接する奥行きはS2及びS4となる。このように、決定部803aは、コントラストが最も大きい奥行きに隣接する奥行きの間（例えば、S2とS4の間）を「コントラストが最も大きかった奥行きが中心になるような範囲」とする。そして、決定部803aは、「コントラストが最も大きかった奥行きが中心になるような範囲」の中（例えば、S2とS4の間）から T （ T は1以上の整数）個の奥行きを選択する。例えば、 T 個の奥行きは、合成情報テーブルに記憶されている、コントラストが最も大きい奥行きに隣接する奥行きの間（コントラストが最も大きかった奥行きが中心になるような範囲）の奥行きの中から等間隔に選択されてもよい。その後、決定部803aは、コントラストが最も大きい奥行き（例えば、S3）と、コントラストが最も大きい奥行きに隣接する奥行きの間（例えば、S2及びS4の間）の奥行きの中から選択された T 個の奥行きと、を含む新たな奥行き候補群を作成する。つまり、決定部803aは、 $T + 1$ 個の奥行きを含む新たな奥行き候補群を作成する。

[0418] 決定部803aは、作成した新たな奥行き候補群に含まれる奥行きそれぞれについて奥行き決定用合成画像を生成し、生成した奥行き決定用合成画像のコントラストを算出する。なお、この際、決定部803aは、既にコントラストを算出している奥行きに対しては処理を行わなくてもよい。決定部803aは、上記処理を所定の回数行う。そして、所定の回数実行後の奥行き候補群の中でコントラストが最も大きい奥行きをカメラペアの各入力画像の最適奥行きに決定する。決定部803aは、以上の処理を各時刻におけるカメラペアの入力画像に対して行う。

[0419] 図55は、画像処理装置80aの処理の流れを示すフローチャートである。

[0420] 決定部803aは、カメラペアの同時刻における入力画像を入力画像記憶部801から読み出し、入力する（ステップS201）。次に、決定部803aは、N個の奥行きのうちS個の奥行きを選択することによって、初期奥行き候補群を作成する（ステップS202）。具体的には、決定部803aは、予め設定されている複数の奥行きの中から特定の数の奥行きを選択することによって初期奥行き候補群を作成する。なお、初期奥行き候補群を作成する処理は、初期処理にのみ行われる。決定部803aは、読み出した入力画像と、合成情報記憶部802に記憶されている合成情報テーブルとに基づいて、奥行き候補群に含まれる1つの奥行きに対する変換情報を用いて奥行き決定用合成画像を生成する（ステップS203）。決定部803aは、生成した奥行き決定用合成画像の重複領域のコントラストを上記の式1に基づいて算出する（ステップS204）。その後、決定部803aは、奥行き候補群に含まれる奥行きの全てに対してコントラストを算出する処理を行ったか否か判定する（ステップS205）。奥行き候補群に含まれる奥行きの全てに対してコントラストを算出する処理を行っていない場合（ステップS205-NO）、決定部803aは奥行き候補群に含まれる奥行きの全てに対してコントラストを算出する処理を行うまでステップS203及びステップS204の処理を繰り返し実行する。

[0421] 一方、奥行き候補群に含まれる奥行きの全てに対してコントラストを算出する処理を行った場合（ステップS205-YES）、決定部803aは最新の奥行き候補群の中でコントラストが最も大きい奥行きを含む、新たな奥行き候補群を作成する。なお、新たな奥行き候補群を作成する場合の奥行き候補群の選択方法は前述の通りである。つまり、決定部803aは、コントラストが最も大きい奥行きと、選択されたT個の奥行きとを含む新たな奥行き候補群を作成する。その後、決定部803aは、作成した新たな奥行き候補群に含まれる各奥行きに対してステップS203から205までの処理を

行う（ステップS206）。その後、決定部803aは、ステップS206の処理を所定の回数行なったか否か判定する（ステップS207）。ステップS206の処理を所定の回数行っていない場合（ステップS207-N）、決定部803aはステップS206の処理を所定の回数に達するまで繰り返し実行する。

[0422] 一方、ステップS206の処理を所定の回数行なった場合（ステップS207-Y E S）、決定部803aは最終的な奥行き候補群の中でコントラストが最も大きい奥行きをカメラペアの各入力画像の最適奥行きに決定する（ステップS208）。ここで、最終的な奥行き候補群とは、ステップS206の処理を所定の回数行なった時点の奥行き候補群を表す。

[0423] 以上のように構成された画像処理装置80aによれば、第十一実施形態と同様の構成を得ることができる。

[0424] また、画像処理装置80aは、第十一実施形態と比べて少数の奥行きに基づいた推定を行いながら最適奥行きを決定する。そのため、第十一実施形態に比べて処理を高速化することが可能になる。

[0425] <第十二実施形態の変形例>

第十二実施形態は、第十一実施形態と同様に変形されてもよい。

[0426] 処理を繰り返す度に新たな奥行き候補群を作成し、作成した新たな候補群から最適解（本実施形態では、最適奥行き）を決定する処理として、C o a r s e - t o - f i n eな方法や、シンプレックス法などが用いられてもよい。

[0427] 本実施形態では、ステップS206の処理が所定の回数行なわれた場合にステップS208の処理が実行される例を示したが、ステップS208の処理は記憶されている奥行き候補の最小単位に行き着いた場合に行われてもよい。

[0428] 本実施形態では、合成情報テーブルに登録されている奥行きの中から選択して奥行き候補群を作成する構成を示したが、決定部803aは合成情報テーブルに登録されている奥行き以外の奥行きを含めて奥行き候補群を作成す

るように構成されてもよい。例えば、仮想視点 8 2 からの距離が 1 である奥行き α （奥行き $\alpha = 1$ ）と、仮想視点 8 2 からの距離が 3 である奥行き β （奥行き $\beta = 3$ ）とが合成情報テーブルに登録されている場合、決定部 8 0 3 a は 2 つの奥行きの間の奥行きである仮想視点 8 2 からの距離が 2（奥行き α と奥行き β の間の奥行き $\gamma = 2$ ）の奥行き γ を含めて奥行き候補群を作成する。このように、合成情報テーブルに登録されている奥行き以外の奥行きを奥行き候補群に含めて用いる場合、決定部 8 0 3 a は合成情報テーブルに登録されている奥行き以外の奥行きの変換情報を登録されている奥行きの変換情報から補間するなどして求める。例えば、変換情報がアフィン変換行列であり、奥行き $\alpha = 1$ のときのアフィン変換行列 $A_{1 \ j_1}$ 、奥行き $\beta = 3$ のときのアフィン変換行列 $A_{1 \ j_2}$ であるとする、奥行き α と奥行き β の間の奥行き γ のアフィン変換行列の各要素は、 $A_{1 \ j_1}$ の行列の各要素と、 $A_{1 \ j_2}$ の行列の各要素とを平均するなどして求めればよい。なお、この例は一例であり、決定部 8 0 3 a は 2 つの奥行きの間（例えば、奥行き α と奥行き β の間）の奥行きとして複数の奥行きを含めて奥行き候補群を作成してもよい。

[0429] 決定部 8 0 3 a は、以下の方法で新たな奥行き候補群を作成してもよい。上記の 5 個の奥行き（ $N = 5$ 。S 1、S 2、S 3、S 4、S 5 の 5 つの奥行き）を例に説明する。

[0430] （第 1 の方法）

第 1 の方法は、コントラストが最も大きい奥行きと、コントラストが最も大きい奥行きに隣接する奥行きそれぞれとの間においてそれぞれ所定の数の奥行きを選択し、選択したそれぞれの所定の数の奥行きと、コントラストが最も大きい奥行きと、を含む新たな奥行き候補群を作成する方法である。以下、具体的に説明する。

[0431] まず決定部 8 0 3 a は、上記のように「コントラストが最も大きかった奥行きが中心になるような範囲」を定める。つまり、コントラストが最も大きい奥行きが S 3 であれば、隣接する奥行きは S 2 及び S 4 となり、S 2 と S 4 の間が「コントラストが最も大きかった奥行きが中心になるような範囲」

になる。ここで、コントラストが最も大きい奥行き（例えば、S 3）とコントラストが最も大きい奥行きに隣接する奥行きの一方（例えば、S 2）との間を第1範囲とし、コントラストが最も大きい奥行きとコントラストが最も大きい奥行きに隣接する奥行きの他方（例えば、S 4）との間を第2範囲とする。次に、決定部803aは、第1範囲においてU（Uは1以上の整数）個の奥行きを選択し、第2範囲においてV（Vは1以上の整数）個の奥行きを選択する。例えば、U個の奥行きは、合成情報テーブルに記憶されている、第1範囲内の奥行きの中から等間隔に選択されてもよい。また、例えば、V個の奥行きは、合成情報テーブルに記憶されている、第2範囲内の奥行きの中から等間隔に選択されてもよい。なお、UとVは同じ値であってもよいし、異なる値であってもよい。その後、決定部803aは、コントラストが最も大きい奥行き（例えば、S 3）と、第1範囲の奥行きの中から選択されたU個の奥行きと、第2範囲の奥行きの中から選択されたV個の奥行きと、を含む新たな奥行き候補群を作成する。つまり、決定部803aは、 $U + V + 1$ 個の奥行きを含む新たな奥行き候補群を作成する。

[0432] （第2の方法）

第2の方法は、コントラストが最も大きい奥行きと、コントラストが最も大きい奥行きに隣接する奥行きの一方との間において所定の数の奥行きを選択し、選択した所定の数の奥行きと、コントラストが最も大きい奥行きと、を含む新たな奥行き候補群を作成する方法である。以下、具体的に説明する。

[0433] まず決定部803aは、上記のように「コントラストが最も大きかった奥行きが中心になるような範囲」を定める。つまり、コントラストが最も大きい奥行きがS 3であれば、隣接する奥行きはS 2及びS 4となり、S 2とS 4の間が「コントラストが最も大きかった奥行きが中心になるような範囲」になる。ここで、コントラストが最も大きい奥行き（例えば、S 3）とコントラストが最も大きい奥行きに隣接する奥行きの一方（例えば、S 2）との間を第3範囲とする。次に、決定部803aは、第3範囲においてZ（Zは

1以上の整数)個の奥行きを選択する。例えば、Z個の奥行きは、合成情報テーブルに記憶されている、第3範囲内の奥行きの中から等間隔に選択されてもよい。その後、決定部803aは、コントラストが最も大きい奥行き(例えば、S3)と、第3範囲の奥行きの中から選択されたZ個の奥行きと、を含む新たな奥行き候補群を作成する。つまり、決定部803aは、Z+1個の奥行きを含む新たな奥行き候補群を作成する。

[0434] [第十三実施形態]

第十一実施形態及び第十二実施形態では、各時刻における各カメラ93の各入力画像の最適奥行きを決定する構成について説明した。このように決定された最適奥行きに基づいて生成される合成画像(仮想全天球画像)を映像として再生した場合、各時刻における最適奥行きが異なるため違和感のある映像になってしまうおそれがある。そこで、第十三実施形態では、画像処理装置は、各時刻における各カメラ93の各入力画像の最適奥行きが時間方向に滑らかに変化するように平滑化(スムージング)処理を行う。また、第十三実施形態では、画像処理システム800に複数のカメラペアが存在する場合を例に説明する。

[0435] 図56は、第十三実施形態における画像処理装置80bの機能構成を表す概略ブロック図である。画像処理装置80bは、バスで接続されたCPUやメモリや補助記憶装置などを備え、画像処理プログラムを実行する。画像処理プログラムの実行によって、画像処理装置80bは、入力画像記憶部801、合成情報記憶部802、決定部803b、最適奥行き記憶部804、背景画像記憶部805、画像合成部806を備える装置として機能する。なお、画像処理装置80bの各機能の全て又は一部は、ASICやPLDやFPGA等のハードウェアを用いて実現されてもよい。また、画像処理プログラムは、コンピュータ読み取り可能な記録媒体に記録されてもよい。コンピュータ読み取り可能な記録媒体とは、例えば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である。また、画像処理プログラムは、

電気通信回線を介して送受信されてもよい。

[0436] 画像処理装置 80b は、決定部 803 に代えて決定部 803b を備える点で画像処理装置 80 と構成が異なる。画像処理装置 80b は、他の構成については画像処理装置 80 と同様である。そのため、画像処理装置 80b 全体の説明は省略し、決定部 803b について説明する。

[0437] 決定部 803b は、入力画像記憶部 801 に記憶されているカメラペアの入力画像と、合成情報記憶部 802 に記憶されている合成情報テーブルとを入力とする。決定部 803b は、入力されたカメラペアの入力画像と合成情報テーブルとに基づいて第十一実施形態における決定部 803 と同様の処理を行うことによって各カメラ 93 の各入力画像の最適奥行きを決定する。決定部 803b は、決定した最適奥行きを暫定最適奥行きとして最適奥行き記憶部 804 に記憶させる。決定部 803b は、最適奥行き記憶部 804 に記憶されている暫定最適奥行きを入力とする。決定部 803b は、入力された暫定最適奥行きに基づいて、各カメラ 93 の暫定最適奥行きを時間方向にスムージングすることによって各時刻における各カメラ 93 の各入力画像の最適奥行きを決定する。決定部 803b の具体的な処理について図 57 を用いて説明する。

[0438] 図 57 は、決定部 803b の具体的な処理について説明するための図である。

[0439] 図 57 において、縦軸は奥行きを表し、横軸は時刻を表す。図 57 に示す複数の円は各時刻における暫定最適奥行きを表す。なお、図 57 では、あるカメラ 93 の各時刻における暫定最適奥行きの例を示している。各時刻において暫定最適奥行きが決定されるため、図 57 に示す円 992 のように前後の暫定最適奥行きと大幅にずれている暫定最適奥行きが存在する場合がある。このような場合には、映像として再生すると違和感のある映像になってしまう可能性がある。そこで、決定部 803b は、各時刻における暫定最適奥行きの値を時間軸上に並べてなめらかになるようにスムージングする。例えば、決定部 803b は、多項式近似することによって各時刻における暫定最

適奥行きから、映像として再生した場合に違和感が軽減されるような最適奥行きの推定値を求める。図57に示される線993が多項式近似によって得られた各時刻における最適奥行きの推定値である。決定部803bは、この処理によって得られた各時刻における最適奥行きの推定値をあるカメラ93の各時刻における最適奥行きとして決定する。そして、決定部803bは、決定した最適奥行きで、最適奥行き記憶部804に記憶されている暫定最適奥行きを上書きすることによって最適奥行きを更新する。

[0440] 図58は、画像処理装置80bの処理の流れを示すフローチャートである。

[0441] 決定部803bは、あるカメラペアのある時刻における入力画像を入力画像記憶部801から読み出す（ステップS301）。次に、決定部803bは、読み出した入力画像と、合成情報記憶部802に記憶されている合成情報テーブルとに基づいてカメラペア奥行き決定処理を行う（ステップS302）。決定部803bは、全てのカメラペアの各入力画像の最適奥行きを決定したか否か判定する（ステップS303）。全てのカメラペアの各入力画像の最適奥行きを決定していない場合（ステップS303-NO）、決定部803bは他のカメラペアを選択する（ステップS304）。ここで、選択される他のカメラペアは、カメラペア奥行き決定処理が行われていないカメラペアである。その後、決定部803bは、選択した他のカメラペアに対してステップS301からステップS302までの処理を実行する。

[0442] 一方、全てのカメラペアの各入力画像の最適奥行きを決定した場合（ステップS303-YES）、決定部803bは決定したカメラペアの各入力画像の最適奥行きに基づいて各カメラ70の各入力画像の最適奥行きを決定する（ステップS305）。その後、決定部803bは、ステップS305の処理で決定した各カメラ93の各入力画像の最適奥行きを暫定最適奥行きとして最適奥行き記憶部804に記憶させる（ステップS306）。決定部803bは、ステップS301からステップS306までの処理を所定の時刻分行ったか否か判定する（ステップS307）。ステップS301からステ

ップS 3 0 6までの処理を所定の時刻分に行っていない場合（ステップS 3 0 7－N O）、決定部8 0 3 bはステップS 3 0 1以降の処理を繰り返し実行する。

[0443] 一方、ステップS 3 0 1からステップS 3 0 6までの処理を所定の時刻分に行った場合（ステップS 3 0 7－Y E S）、決定部8 0 3 bは最適奥行き記憶部8 0 4に記憶されているあるカメラ9 3の所定の時刻分の暫定最適奥行きに関して時間方向にスムージングを行う（ステップS 3 0 8）。例えば、決定部8 0 3 bは、多項式近似することによって所定の時刻分の暫定最適奥行きから、あるカメラ9 3の最適奥行きの推定値を求める。その後、決定部8 0 3 bは、あるカメラ9 3の最適奥行きの推定値をあるカメラ9 3の各時刻における最適奥行きとして最適奥行き記憶部8 0 4に記憶させる（ステップS 3 0 9）。決定部8 0 3 bは、ステップS 3 0 8及びステップS 3 0 9の処理を全てのカメラ9 3に対して行ったか否か判定する（ステップS 3 1 0）。ステップS 3 0 8及びステップS 3 0 9の処理を全てのカメラ9 3に対して行っていない場合（ステップS 3 1 0－N O）、決定部8 0 3 bはステップS 3 0 8及びステップS 3 0 9の処理を全てのカメラに対して行うまで処理を繰り返し実行する。

[0444] 一方、ステップS 3 0 8及びステップS 3 0 9の処理を全てのカメラ9 3に対して行った場合（ステップS 3 1 0－Y E S）、画像合成部8 0 6は画像合成処理ステップを実行する（ステップS 3 1 1）。

[0445] 以上のように構成された画像処理装置8 0 bによれば、各時刻における入力画像を用いてそれぞれ生成された仮想全天球画像を映像で出力する場合に生じる可能性がある違和感を軽減することが可能になる。以下、この効果について詳細に説明する。

[0446] 画像処理装置8 0 bは、各時刻における暫定最適奥行きの情報を用いて、各時刻における最適奥行きの値が滑らかになるようにスムージングを行う。これにより、画像処理装置8 0 bは、ある時刻における最適奥行きが他の最適奥行きと大幅にずれている場合であっても前後の時刻における最適奥行き

から、大幅にずれている時刻における最適奥行きを補正することができる。そのため、各時刻における入力画像を用いてそれぞれ生成された仮想全天球画像を映像で出力する場合に生じる可能性がある違和感を軽減することが可能になる。

[0447] <第十三実施形態の変形例>

画像処理装置80bにおける入力画像記憶部801、合成情報記憶部802、決定部803b及び最適奥行き記憶部804は、最適奥行き決定装置として構成されてもよい。

[0448] 本実施形態では、決定部803bが決定部803と同様の処理を行うことによって各カメラ70の各入力画像の最適奥行きを決定する構成を示したが、決定部803bは決定部803aと同様の処理を行うことによって各カメラ93の各入力画像の最適奥行きを決定するように構成されてもよい。

[0449] [第十四実施形態]

第十四実施形態では、仮想全天球画像をリアルタイムに映像として出力する場合について説明する。具体的には、第4実施形態では、ある時刻 t におけるカメラ93の入力画像の最適奥行きを直前の時刻（例えば、時刻 $t-1$ 、 $t-2$ など）の最適奥行きに基づいて決定する。また、第4実施形態では、画像処理システム800に複数のカメラペアが存在する場合を例に説明する。

[0450] 図59は、第十四実施形態における画像処理装置80cの機能構成を表す概略ブロック図である。画像処理装置80cは、バスで接続されたCPUやメモリや補助記憶装置などを備え、画像処理プログラムを実行する。画像処理プログラムの実行によって、画像処理装置80cは、入力画像記憶部801、合成情報記憶部802、決定部803c、最適奥行き記憶部804c、背景画像記憶部805、画像合成部806を備える装置として機能する。なお、画像処理装置80cの各機能の全て又は一部は、ASICやPLDやFPGA等のハードウェアを用いて実現されてもよい。また、画像処理プログラムは、コンピュータ読み取り可能な記録媒体に記録されてもよい。コンピ

ユーザ読み取り可能な記録媒体とは、例えば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である。また、画像処理プログラムは、電気通信回線を介して送受信されてもよい。

[0451] 画像処理装置80cは、決定部803及び最適奥行き記憶部804に代えて決定部803c及び最適奥行き記憶部804cを備える点で画像処理装置80と構成が異なる。画像処理装置80cは、他の構成については画像処理装置80と同様である。そのため、画像処理装置80c全体の説明は省略し、決定部803c及び最適奥行き記憶部804cについて説明する。

[0452] 最適奥行き記憶部804cは、磁気ハードディスク装置や半導体記憶装置などの記憶装置を用いて構成される。最適奥行き記憶部804cは、最適奥行き情報テーブルを記憶する。第十一実施形態から第十三実施形態における最適奥行き記憶部804が記憶する最適奥行き情報テーブルには各時刻における各カメラ93の各入力画像の最適奥行き情報が登録されていた。それに対して、第4実施形態における最適奥行き記憶部804cが記憶する最適奥行き情報テーブルには、処理開始時には現時刻（例えば、リアルタイムに出力する仮想全天球画像を生成する時刻） t における各カメラ93の各入力画像の最適奥行き情報は登録（記憶）されておらず、現時刻 t よりも前の時刻（例えば、時刻 $t-1$ 、 $t-2$ など）における各カメラ93の各入力画像の最適奥行き情報が登録（記憶）されている。

[0453] 決定部803cは、入力画像記憶部801に記憶されているカメラペアの入力画像と、合成情報記憶部802に記憶されている合成情報テーブルとを入力とする。決定部803bは、入力されたカメラペアの入力画像と合成情報テーブルとに基づいて第十一実施形態における決定部803と同様の処理を行うことによって各カメラ93の各入力画像の最適奥行きを決定する。また、決定部803cは、現時刻 t よりも前の時刻（例えば、時刻 $t-1$ 、 $t-2$ など）における最適奥行き情報に基づいて、現時刻 t における最適奥行きを決定する。なお、第4実施形態では、現時刻 t よりも前の時刻の最適奥行き

情報に基づいて決定する方法として時系列フィルタの一つであるパーティクルフィルタによる実現方法を説明する。なお、以下の処理において作成するパーティクルの数は P であるとする。

[0454] 図60は、画像処理装置80cの処理の流れを示すフローチャートである。

[0455] 決定部803cは、時刻 t に撮影されたカメラペアの入力画像を入力画像記憶部801から読み出す（ステップS401）。なお、処理開始時には、時刻 t は初期時刻である。初期時刻とは全ての時刻における最初の時刻を設定することが一般的であるが、今回導入しているような時系列フィルタリングの考え方は誤差が蓄積して望ましくない値に収束する場合がある（ドリフトと呼ばれている）。このような場合、適当な条件を満たした時、例えば時刻 t で最大尤度が閾値以下になった時に時刻 $t+1$ を初期時刻としてもよい。

[0456] 決定部803cは、読み出した入力画像に関して、全ての奥行き（例えば、奥行き1～奥行き N ）の中から P 個の奥行き候補群を選択する（ステップS402）。初期時刻において全ての奥行きの中から P 個の奥行きを選択する方法として、ランダムに P 個選択する方法、又は、全ての奥行きを P 等分して P 個選択する方法がある。初期時刻以外では、基本的に前の時刻において選択した奥行き候補群を用いてもよい。決定部803cは、奥行き候補群のある奥行きについて予測モデルを用いて奥行きの予想値を算出する（ステップS403）。予測モデルには、等速直線運動や等加速度運動、あるいはランダムウォークなどがある（これは、一般的なパーティクルフィルタの「予測」にあたるステップである。）。

[0457] その後、決定部803cは、ステップS303の処理で算出した奥行きの予想値に対して尤度を算出する（ステップS404）。尤度を求める方法として、ある奥行きに対応する合成情報に基づいて生成された奥行き決定用合成画像の重複領域に対するコントラストなどがある。この場合、コントラストが大きいほど尤度が大きいとする（これは、一般的なパーティクルフィル

タの「重み付け」にあたるステップである。)。決定部803cは、全ての奥行き候補群に対して尤度を算出したか否か判定する(ステップS405)。全ての奥行き候補群に対して尤度を算出していない場合(ステップS405-NO)、決定部803cは全ての奥行き候補群に対して尤度を算出するまでステップS402からステップS404までの処理を繰り返し実行する。

[0458] 一方、全ての奥行き候補群に対して尤度を算出した場合(ステップS405-YES)、決定部803cはステップS404の処理で算出した尤度に基づいて時刻 t におけるカメラペアの各入力画像の最適奥行きを決定する(ステップS406)。最適奥行きを求める方法として、尤度の最も大きい奥行きを最適奥行きとして求める方法、尤度の重み付き平均値を最適奥行きとする方法などがある(これは、一般的なパーティクルフィルタの「状態推定」にあたるステップである。)。次に、決定部803cはステップS404の処理で算出した尤度に基づいて時刻 $t+1$ の奥行き候補群を選択する(ステップS407)。尤度に基づいて時刻 $t+1$ の奥行き候補群を選択する方法として、尤度の大きさを用いる方法がある。具体的には、各奥行きについて尤度の大きさに基づいてその奥行き付近で選択する候補の数を変動させる。例えば、尤度の小さい奥行きの近くでは候補をほとんど選択せず、尤度の大きい奥行きの近くでは多くの候補を選択する。この時、候補の合計が P 個になるようにする(これは、一般的なパーティクルフィルタの「リサンプリング」にあたるステップである。))。

[0459] 決定部803cは、ステップS401からステップ407までの処理を全てのカメラペアに対して行ったか否か判定する(ステップS408)。ステップS401からステップ407までの処理を全てのカメラペアに対して行っていない場合(ステップS408-NO)、決定部803cはステップ401以降の処理を全てのカメラペアに対して行う。

[0460] 一方、ステップS401からステップ407までの処理を全てのカメラペアに対して行った場合(ステップS408-YES)、決定部803cは時

刻 t におけるカメラペアの各入力画像の最適奥行きから各カメラ 93 の各入力画像の最適奥行きを決定する（ステップ S 409）。その後、決定部 803c は、時刻 t における各カメラ 93 の各入力画像の最適奥行きを最適奥行き記憶部 804c に記憶させる。その後、画像合成部 806 は、画像合成処理ステップを実行する（ステップ S 410）。

[0461] 以上のように構成された画像処理装置 80c によれば、リアルタイムに各カメラ 93 の入力画像の最適奥行きを決定することが可能になる。以下、この効果について詳細に説明する。

[0462] 画像処理装置 80c は、時刻 t における各カメラ 93 の入力画像の最適奥行きを過去の時刻（例えば、直前の時刻）の入力画像の最適奥行きから推定する。具体的には、まず画像処理装置 80c は、時刻 t におけるカメラペアの各入力画像の奥行きを決定するための奥行き候補群に含まれる奥行きについて予測モデルを用いて奥行きの予想値を算出する。次に、画像処理装置 80c は、算出した奥行きの予想値の尤度を算出する。そして、画像処理装置 80c は、算出した尤度に基づいて時刻 t におけるカメラペアの入力画像の最適奥行きを決定する。その後、画像処理装置 80c は、同様の処理を全てのカメラペアの入力画像に対して行う。そのため、リアルタイムに各カメラ 93 の入力画像の最適奥行きを決定することが可能になる。

[0463] <第十四実施形態の変形例>

画像処理装置 80c における入力画像記憶部 801、合成情報記憶部 802、決定部 803c 及び最適奥行き記憶部 804c は、最適奥行き決定装置として構成されてもよい。

[0464] 以下、第十一実施形態から第十四実施形態までの各実施形態に共通する変形例について説明する。

[0465] これらの実施形態は、全天球画像の場合に限らず、複数のカメラ 93 で撮影された画像を背景画像に重畳することによって合成画像を生成する場合に適用可能である。

[0466] これらの実施形態では、合成情報テーブルに登録されている幅の値が全て

のカメラIDで同じであるが、幅の値はカメラID毎に異なってもよいし、一部のカメラIDで異なってもよい。

[0467] 決定部803は、1つの入力画像の所定の領域毎にそれぞれ最適奥行きを決定してもよい。

[0468] 第十一実施形態から第十四実施形態までの各実施形態により、複数の撮影装置のそれぞれで撮影された画像の最適な奥行きを決定することが可能となる。

[0469] 以上、この発明の実施の形態について説明したが、具体的な構成は、これらの実施の形態に限られるものではなく、この発明の趣旨を逸脱しない範囲で適宜設計の変更等があっても、この発明に含まれることはいうまでもない。実施の形態において説明した各種の処理は、記載の順に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的にあるいは個別に実行されてもよい。

[0470] [プログラム、記録媒体]

上記実施形態で説明した各装置における各種の処理機能をコンピュータによって実現する場合、各装置が有すべき機能の処理内容はプログラムによって記述される。そして、このプログラムをコンピュータで実行することにより、上記各装置における各種の処理機能がコンピュータ上で実現される。

[0471] この処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、例えば、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリ等のようなものでもよい。

[0472] また、このプログラムの流通は、例えば、そのプログラムを記録したDVD、CD-ROM等の可搬型記録媒体を販売、譲渡、貸与等することによって行う。さらに、このプログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することにより、このプログラムを流通させる構成としてもよい。

[0473] このようなプログラムを実行するコンピュータは、例えば、まず、可搬型記録媒体に記録されたプログラムもしくはサーバコンピュータから転送されたプログラムを、一旦、自己の記憶装置に格納する。そして、処理の実行時、このコンピュータは、自己の記録媒体に格納されたプログラムを読み取り、読み取ったプログラムに従った処理を実行する。また、このプログラムの別の実行形態として、コンピュータが可搬型記録媒体から直接プログラムを読み取り、そのプログラムに従った処理を実行することとしてもよく、さらに、このコンピュータにサーバコンピュータからプログラムが転送されるたびに、逐次、受け取ったプログラムに従った処理を実行することとしてもよい。また、サーバコンピュータから、このコンピュータへのプログラムの転送は行わず、その実行指示と結果取得のみによって処理機能を実現する、いわゆるASP (Application Service Provider) 型のサービスによって、上述の処理を実行する構成としてもよい。なお、本形態におけるプログラムには、電子計算機による処理の用に供する情報であってプログラムに準ずるもの（コンピュータに対する直接の指令ではないがコンピュータの処理を規定する性質を有するデータ等）を含むものとする。

[0474] また、この形態では、コンピュータ上で所定のプログラムを実行させることにより、本装置を構成することとしたが、これらの処理内容の少なくとも一部をハードウェア的に実現することとしてもよい。

請求の範囲

[請求項1] 動的オブジェクトと静的オブジェクトとが存在する現実環境に基づいてユーザに閲覧させる仮想環境を構築する仮想環境構築装置であって、

 上記動的オブジェクトと上記静的オブジェクトとを上記仮想環境上に再現する仮想環境素材を記憶する仮想環境素材記憶部と、

 上記現実環境における上記動的オブジェクトの位置および姿勢を時系列に計測して1つの運動動作からなる位置姿勢情報を生成する動的オブジェクトセンシング部と、

 相異なる複数種類の上記動的オブジェクトの位置姿勢情報を含む提示シーケンスを取得する提示シーケンス取得部と、

 上記提示シーケンスに基づいて上記動的オブジェクトの仮想環境素材と上記静的オブジェクトの仮想環境素材とを合成して上記仮想環境を構築する仮想環境構築部と、

 を含む仮想環境構築装置。

[請求項2] 請求項1に記載の仮想環境構築装置であって、

 上記提示シーケンスに含まれる動的オブジェクトの位置姿勢情報または仮想環境素材に対して、上記ユーザに対する視覚的負荷を高める効果を与える効果付与部をさらに含む、

 仮想環境構築装置。

[請求項3] 請求項1または2に記載の仮想環境構築装置であって、

 上記提示シーケンス取得部は、上記動的オブジェクトを野球ボールとし、球速が早い投球から生成した位置姿勢情報と球速が遅い投球から生成した位置姿勢情報とを少なくとも含む提示シーケンス、または、ストライクゾーンの一方向の隅への投球から生成した位置姿勢情報と対角線上の他方の隅への投球から生成した位置姿勢情報とを少なくとも含む提示シーケンスのいずれかを取得するものである、

 仮想環境構築装置。

- [請求項4] 請求項2に記載の仮想環境構築装置であって、
 上記効果付与部は、上記動的オブジェクトを野球ボールとし、投球から生成した位置姿勢情報に対して所定の割合で球速を増減させる効果、変化球から生成した位置姿勢情報に対して所定の割合で変化量を増減させる効果、または、投球から生成した位置姿勢情報に対して時間経過に伴って視認しにくい表示とする効果のいずれかを与えるものである、
 仮想環境構築装置。
- [請求項5] 請求項1から4のいずれかに記載の仮想環境構築装置であって、
 上記ユーザの頭部の位置および姿勢を計測してユーザ位置姿勢情報を生成するユーザセンシング部と、
 上記ユーザ位置姿勢情報に基づいて上記仮想環境における上記ユーザの視線方向の映像を出力する描画部と、
 をさらに含む仮想環境構築装置。
- [請求項6] 動的オブジェクトと静的オブジェクトとが存在する現実環境に基づいてユーザに閲覧させる仮想環境を構築する仮想環境構築方法であって、
 仮想環境素材記憶部に、上記動的オブジェクトと上記静的オブジェクトとを上記仮想環境上に再現する仮想環境素材が記憶されており、
 動的オブジェクトセンシング部が、上記現実環境における上記動的オブジェクトの位置および姿勢を時系列に計測して1つの運動動作からなる位置姿勢情報を生成し、
 提示シーケンス取得部が、相異なる複数種類の上記動的オブジェクトの位置姿勢情報を含む提示シーケンスを取得し、
 仮想環境構築部が、上記提示シーケンスに基づいて上記動的オブジェクトの仮想環境素材と上記静的オブジェクトの仮想環境素材とを合成して上記仮想環境を構築する、
 仮想環境構築方法。

[請求項7] 相手が飛来物を発出する現実環境に基づいてユーザに体験させる仮想環境を構築する仮想環境構築装置であって、

上記飛来物のシーンの事前確率と複数シーンの繋がりを表すダイナミクスとを合わせた確率に基づいて上記ユーザへ提示するシーンの時系列を表す提示シーケンスを生成する提示シーケンス生成部と、

上記提示シーケンスに含まれる各シーンに対応する上記飛来物の動きを表す提示軌跡を決定する提示軌跡決定部と、

上記提示シーケンスに含まれる各シーンに対応する上記相手の動きを表す提示動作を決定する提示動作決定部と、

上記提示軌跡と上記提示動作とに基づいて上記ユーザに体験させる上記仮想環境を合成する体験合成部と、

を含む仮想環境構築装置。

[請求項8] 相手が飛来物を発出する現実環境に基づいてユーザに体験させる仮想環境を構築する仮想環境構築装置であって、

上記ユーザへ提示するシーンの時系列を表す提示シーケンスを生成する提示シーケンス生成部と、

上記提示シーケンスに含まれる各シーンに対応する上記飛来物の動きを表す提示軌跡を決定する提示軌跡決定部と、

上記提示シーケンスに含まれる各シーンに対応して上記現実環境とは異なる上記相手の動きを表す提示動作を決定する提示動作決定部と、

上記提示軌跡と上記提示動作とに基づいて上記ユーザに体験させる上記仮想環境を合成する体験合成部と、

を含む仮想環境構築装置。

[請求項9] 請求項8に記載の仮想環境構築装置であって、

上記提示シーケンス生成部は、上記飛来物のシーンの事前確率と複数シーンの繋がりを表すダイナミクスとを合わせた確率に基づいて上記提示シーケンスを生成するものである、

仮想環境構築装置。

[請求項10]

請求項7または9に記載の仮想環境構築装置であって、

上記提示シーケンス生成部は、上記飛来物のコース、速度、および球種のすべてについて上記飛来物の上記事前確率と上記ダイナミクスとを合わせた確率に基づいて上記提示シーケンスを生成し、上記現実環境から取得した上記飛来物の軌跡の中に当該提示シーケンスに該当するものが存在しない場合にはユーザに体験させたい場面を想定して予め定めた練習意図に従ってコース、速度、および球種のいずれかを優先して上記飛来物の上記事前確率と上記ダイナミクスとを合わせた確率に基づいて上記提示シーケンスを生成するものである、

仮想環境構築装置。

[請求項11]

請求項7または9に記載の仮想環境構築装置であって、

上記提示シーケンス生成部は、上記飛来物の飛来に関する複数のパラメータのすべてについて上記飛来物の上記事前確率と上記ダイナミクスとを合わせた確率に基づいて上記提示シーケンスを生成し、上記現実環境から取得した上記飛来物の軌跡の中に当該提示シーケンスに該当するものが存在しない場合にはユーザに体験させたい場面を想定して予め定めた練習意図に従って上記複数のパラメータのうちの一部のパラメータを優先して上記飛来物の上記事前確率と上記ダイナミクスとを合わせた確率に基づいて上記提示シーケンスを生成するものである、

仮想環境構築装置。

[請求項12]

相手が飛来物を発出する現実環境に基づいてユーザに体験させる仮想環境を構築する仮想環境構築方法であって、

提示シーケンス生成部が、上記飛来物のシーンの事前確率と複数シーンの繋がりを表すダイナミクスとを合わせた確率に基づいて上記ユーザへ提示するシーンの時系列を表す提示シーケンスを生成し、

提示軌跡決定部が、上記提示シーケンスに含まれる各シーンに対応

する上記飛来物の動きを表す提示軌跡を決定し、

提示動作決定部が、上記提示シーケンスに含まれる各シーンに対応する上記相手の動きを表す提示動作を決定し、

体験合成部が、上記提示軌跡と上記提示動作とに基づいて上記ユーザに体験させる上記仮想環境を合成する、

仮想環境構築方法。

[請求項13]

相手が飛来物を発出する現実環境に基づいてユーザに体験させる仮想環境を構築する仮想環境構築方法であって、

提示シーケンス生成部が、上記ユーザへ提示するシーンの時系列を表す提示シーケンスを生成し、

提示軌跡決定部が、上記提示シーケンスに含まれる各シーンに対応する上記飛来物の動きを表す提示軌跡を決定し、

提示動作決定部が、上記提示シーケンスに含まれる各シーンに対応する上記現実環境とは異なる上記相手の動きを表す提示動作を決定し、

体験合成部が、上記提示軌跡と上記提示動作とに基づいて上記ユーザに体験させる上記仮想環境を合成する、

仮想環境構築方法。

[請求項14]

相手が飛来物を発出する映像をユーザに提示する映像提示装置であって、

上記ユーザの身体的な挙動を表すユーザ状態と上記相手が飛来物を発出するシーンを表すシーン情報とを関連付けて記憶するユーザ状態記憶部と、

上記ユーザに提示された上記シーンに対する反応として現れるユーザ状態であるユーザ反応を取得するユーザ状態取得部と、

上記ユーザ状態記憶部から上記ユーザ反応に関連付けられたシーン情報である提示シーン情報を取得する提示マネジメント部と、

上記提示シーン情報に基づいて上記ユーザに提示する上記映像を生

成する映像提示部と、

を含む映像提示装置。

[請求項15]

請求項14に記載の映像提示装置であって、

上記ユーザ状態は、少なくとも上記ユーザの頭部の位置および姿勢を含むものである、

映像提示装置。

[請求項16]

請求項15に記載の映像提示装置であって、

上記ユーザ状態は、上記ユーザの全身の姿勢または体重移動をさらに含むものである、

映像提示装置。

[請求項17]

請求項14から16のいずれかに記載の映像提示装置であって、

上記ユーザ反応に対する評価値を取得するユーザ評価取得部をさらに含み、

上記提示マネジメント部は、上記ユーザ反応に関連付けられた複数のシーン情報から上記評価値に基づいて上記提示シーン情報を選択するものである、

映像提示装置。

[請求項18]

請求項14から17のいずれかに記載の映像提示装置であって、

上記ユーザ反応の時間遷移を時系列に提示するユーザ状態提示部をさらに含む、

映像提示装置。

[請求項19]

相手が飛来物を発出する映像をユーザに提示する映像提示方法であって、

ユーザ状態記憶部に、上記ユーザの身体的な挙動を表すユーザ状態と上記相手が飛来物を発出するシーンを表すシーン情報とが関連付けて記憶されており、

ユーザ状態取得部が、上記ユーザに提示された上記シーンに対する反応として現れるユーザ状態であるユーザ反応を取得し、

提示マネジメント部が、上記ユーザ状態記憶部から上記ユーザ反応に関連付けられたシーン情報である提示シーン情報を取得し、

映像提示部が、上記提示シーン情報に基づいて上記ユーザに提示する上記映像を生成する、

映像提示方法。

[請求項20]

位置が遷移する動く素材である動的素材と、位置が遷移せず動かない素材である静的素材と、位置が遷移せず動く素材である特殊動的素材が実環境に存在するものとし、前記実環境を再現する仮想環境を生成する仮想環境構築装置であって、

前記動的素材である物体を仮想環境上に再現するための仮想動的素材と、前記静的素材を仮想環境上に再現するための仮想静的素材とを取得する仮想素材取得部と、

前記仮想動的素材の位置および姿勢を検出してその時系列データを取得する仮想動的素材位置姿勢検出部と、

前記実環境を撮影した映像に基づいて前記特殊動的素材である人間を仮想環境上に再現するための仮想特殊動的素材を生成する仮想特殊動的素材生成部と、

前記人間と前記物体とが離隔する瞬間の前後を高フレームレートで撮影した映像と、前記人間と前記物体とが離隔する瞬間の時刻とのペアを学習したモデルへ、前記人間と前記物体とが離隔する前を撮影した離隔前画像もしくは前記人間と前記物体とが離隔した後を撮影した離隔後画像の少なくともいずれかを入力することで前記人間と前記物体とが離隔する瞬間の時刻を推定し、推定された時刻に基づいて前記仮想動的素材と前記仮想特殊動的素材とを同期する同期部と、

前記仮想静的素材と、同期された前記仮想動的素材と、同期された前記仮想特殊動的素材に基づいて前記仮想環境を生成する仮想環境生成部と、

を含む仮想環境構築装置。

- [請求項21] 請求項20に記載の仮想環境構築装置であって、
前記同期部は、前記モデルへ前記離隔前画像および前記離隔後画像の両方を入力することで前記人間と前記物体とが離隔する瞬間の時刻を推定するものである、
仮想環境構築装置。
- [請求項22] 請求項20または21に記載の仮想環境構築装置であって、
前記同期部は、前記人間と前記物体とが離隔する前または後における前記人間の身体または前記物体のテンプレートを用いて前記実環境を撮影した映像に対してテンプレートマッチングを実行して前記離隔前画像または前記離隔後画像を抽出し、前記モデルへ前記離隔前画像または前記離隔後画像を入力することで前記人間と前記物体とが離隔する瞬間の時刻を推定するものである、
仮想環境構築装置。
- [請求項23] 位置が遷移する動く素材である動的素材と、位置が遷移せず動かない素材である静的素材と、位置が遷移せず動く素材である特殊動的素材が実環境に存在するものとし、前記実環境を再現する仮想環境を生成する仮想環境構築装置であって、
前記動的素材を仮想環境上に再現するための仮想動的素材と、前記静的素材を仮想環境上に再現するための仮想静的素材とを取得する仮想素材取得部と、
前記仮想動的素材の位置および姿勢を検出してその時系列データを取得する仮想動的素材位置姿勢検出部と、
前記実環境を撮影した映像に基づいて前記特殊動的素材を仮想環境上に再現するための仮想特殊動的素材を生成する仮想特殊動的素材生成部と、
前記特殊動的素材と前記動的素材とが離隔する瞬間の前後を高フレームレートで撮影した映像と、前記特殊動的素材と前記動的素材とが離隔する瞬間の時刻とのペアを学習したモデルへ、前記特殊動的素材

と前記動的素材とが離隔する前を撮影した離隔前画像もしくは前記特殊動的素材と前記動的素材とが離隔した後を撮影した離隔後画像の少なくともいずれかを入力することで前記特殊動的素材と前記動的素材とが離隔する瞬間の時刻を推定し、推定された時刻に基づいて前記仮想動的素材と前記仮想特殊動的素材とを同期する同期部と、

前記仮想静的素材と、同期された前記仮想動的素材と、同期された前記仮想特殊動的素材に基づいて前記仮想環境を生成する仮想環境生成部と、

を含む仮想環境構築装置。

[請求項24]

人間と物体とが離隔する瞬間の前後を高フレームレートで撮影した映像と、前記人間と前記物体とが離隔する瞬間の時刻とのペアを用いて、前記人間と前記物体とが離隔する前を撮影した離隔前画像もしくは前記人間と前記物体とが離隔した後を撮影した離隔後画像との少なくともいずれかを入力として前記人間と前記物体とが離隔する瞬間の時刻を出力するモデルを学習する、

モデル学習装置。

[請求項25]

位置が遷移せず動く素材である特殊動的素材と位置が遷移する動く素材である動的素材とが離隔する瞬間の前後を高フレームレートで撮影した映像と、前記特殊動的素材と前記動的素材とが離隔する瞬間の時刻とのペアを用いて、前記特殊動的素材と前記動的素材とが離隔する前を撮影した離隔前画像もしくは前記特殊動的素材と前記動的素材とが離隔した後を撮影した離隔後画像との少なくともいずれかを入力として前記特殊動的素材と前記動的素材とが離隔する瞬間の時刻を出力するモデルを学習する、

モデル学習装置。

[請求項26]

位置が遷移する動く素材である動的素材と、位置が遷移せず動かない素材である静的素材と、位置が遷移せず動く素材である特殊動的素材が実環境に存在するものとし、前記実環境を再現する仮想環境を生

成する仮想環境生成装置が実行する仮想環境構築方法であって、

前記動的素材である物体を仮想環境上に再現するための仮想動的素材と、前記静的素材を仮想環境上に再現するための仮想静的素材とを取得するステップと、

前記仮想動的素材の位置および姿勢を検出してその時系列データを取得するステップと、

前記実環境を撮影した映像に基づいて前記特殊動的素材である人間を仮想環境上に再現するための仮想特殊動的素材を生成するステップと、

前記人間と前記物体とが離隔する瞬間の前後を高フレームレートで撮影した映像と、前記人間と前記物体とが離隔する瞬間の時刻とのペアを学習したモデルへ、前記人間と前記物体とが離隔する前を撮影した離隔前画像もしくは前記人間と前記物体とが離隔した後を撮影した離隔後画像の少なくともいずれかを入力することで前記人間と前記物体とが離隔する瞬間の時刻を推定し、推定された時刻に基づいて前記仮想動的素材と前記仮想特殊動的素材とを同期するステップと、

前記仮想静的素材と、同期された前記仮想動的素材と、同期された前記仮想特殊動的素材に基づいて前記仮想環境を生成するステップと、
を含む仮想環境構築方法。

[請求項27]

位置が遷移する動く素材である動的素材と、位置が遷移せず動かない素材である静的素材と、位置が遷移せず動く素材である特殊動的素材が実環境に存在するものとし、前記実環境を再現する仮想環境を生成する仮想環境構築装置が実行する仮想環境構築方法であって、

前記動的素材を仮想環境上に再現するための仮想動的素材と、前記静的素材を仮想環境上に再現するための仮想静的素材とを取得するステップと、

前記仮想動的素材の位置および姿勢を検出してその時系列データを

取得するステップと、

前記実環境を撮影した映像に基づいて前記特殊動的素材を仮想環境上に再現するための仮想特殊動的素材を生成するステップと、

前記特殊動的素材と前記動的素材とが離隔する瞬間の前後を高フレームレートで撮影した映像と、前記特殊動的素材と前記動的素材とが離隔する瞬間の時刻とのペアを学習したモデルへ、前記特殊動的素材と前記動的素材とが離隔する前を撮影した離隔前画像もしくは前記特殊動的素材と前記動的素材とが離隔した後を撮影した離隔後画像の少なくともいずれかを入力することで前記特殊動的素材と前記動的素材とが離隔する瞬間の時刻を推定し、推定された時刻に基づいて前記仮想動的素材と前記仮想特殊動的素材とを同期するステップと、

前記仮想静的素材と、同期された前記仮想動的素材と、同期された前記仮想特殊動的素材に基づいて前記仮想環境を生成するステップと、

、

を含む仮想環境構築方法。

[請求項28]

人間と物体とが離隔する瞬間の前後を高フレームレートで撮影した映像と、前記人間と前記物体とが離隔する瞬間の時刻とのペアを用いて、前記人間と前記物体とが離隔する前を撮影した離隔前画像もしくは前記人間と前記物体とが離隔した後を撮影した離隔後画像との少なくともいずれかを入力として前記人間と前記物体とが離隔する瞬間の時刻を出力するモデルを学習する、

モデル学習方法。

[請求項29]

位置が遷移せず動く素材である特殊動的素材と位置が遷移する動く素材である動的素材とが離隔する瞬間の前後を高フレームレートで撮影した映像と、前記特殊動的素材と前記動的素材とが離隔する瞬間の時刻とのペアを用いて、前記特殊動的素材と前記動的素材とが離隔する前を撮影した離隔前画像もしくは前記特殊動的素材と前記動的素材とが離隔した後を撮影した離隔後画像との少なくともいずれかを入力

として前記特殊動的素材と前記動的素材とが離隔する瞬間の時刻を出力するモデルを学習する、

モデル学習方法。

[請求項30]

位置が遷移する動く素材である動的素材と、位置が遷移せず動かない素材である静的素材と、位置が遷移せず動く素材である特殊動的素材が実環境に存在するものとし、前記実環境を再現する仮想環境を生成する仮想環境構築装置であって、

前記動的素材を仮想環境上に再現するための仮想動的素材と、前記静的素材を仮想環境上に再現するための仮想静的素材とを取得する仮想素材取得部と、

前記仮想動的素材の位置および姿勢を検出してその時系列データを取得する仮想動的素材位置姿勢検出部と、

予め定めた事前準備フェーズにおいて取得された前記時系列データと、これに関連する前記特殊動的素材の映像とを対応付ける事前準備部と、

予め定めた仮想環境生成フェーズにおいて取得された前記時系列データと類似する前記事前準備フェーズにおける前記時系列データを検索し、検索された前記時系列データと対応付けられた前記特殊動的素材の映像を出力する類似映像検索部と、

出力された前記特殊動的素材の映像に基づいて、前記特殊動的素材を仮想環境上に再現するための仮想特殊動的素材を生成する仮想特殊動的素材生成部と、

前記仮想動的素材と前記仮想特殊動的素材とを同期する同期部と、

前記仮想静的素材と、同期された前記仮想動的素材と、同期された前記仮想特殊動的素材に基づいて前記仮想環境を生成する仮想環境生成部と、

を含む仮想環境構築装置。

[請求項31]

位置が遷移する動く素材である動的素材と、位置が遷移せず動かない

い素材である静的素材と、位置が遷移せず動く素材である特殊動的素材が実環境に存在するものとし、前記実環境を再現する仮想環境を生成する仮想環境生成装置が実行する仮想環境構築方法であって、

前記動的素材を仮想環境上に再現するための仮想動的素材と、前記静的素材を仮想環境上に再現するための仮想静的素材とを取得するステップと、

前記仮想動的素材の位置および姿勢を検出してその時系列データを取得するステップと、

予め定めた事前準備フェーズにおいて取得された前記時系列データと、これに関連する前記特殊動的素材の映像とを対応付けるステップと、

予め定めた仮想環境生成フェーズにおいて取得された前記時系列データと類似する前記事前準備フェーズにおける前記時系列データを検索し、検索された前記時系列データと対応付けられた前記特殊動的素材の映像を出力するステップと、

出力された前記特殊動的素材の映像に基づいて、前記特殊動的素材を仮想環境上に再現するための仮想特殊動的素材を生成するステップと、

前記仮想動的素材と前記仮想特殊動的素材とを同期するステップと、

前記仮想静的素材と、同期された前記仮想動的素材と、同期された前記仮想特殊動的素材に基づいて前記仮想環境を生成するステップと、

を含む仮想環境構築方法。

[請求項32] 複数のオブジェクトが存在する現実環境に基づいてユーザに閲覧させる仮想環境を構築する仮想環境構築装置であって、

上記複数のオブジェクトのうち少なくともいずれかのオブジェクトを対象オブジェクトとして、上記対象オブジェクトのセンシング難易

度とCG描画難易度の少なくともいずれかに基づいて上記オブジェクトをCGによりモデル構築するか実写映像によりモデル構築するかを示すモデル構築方法を決定するモデル構築方法決定部と、

上記現実環境における上記対象オブジェクトの状態を計測する状態センシング部と、

時間変化を伴わない静的オブジェクトについて上記モデル構築方法に従って上記仮想環境中に描画するための環境モデルを構築する環境モデル構築部と、

時間変化を伴う動的オブジェクトについて上記モデル構築方法に従って上記仮想環境中に描画するためのモデルを構築して上記環境モデルに合成する動的オブジェクト構築部と、

を含む仮想環境構築装置。

[請求項33]

請求項32に記載の仮想環境構築装置であって、

上記モデル構築方法決定部は、SEを上記センシング難易度とし、CMを上記CG描画難易度とし、 $\theta_{\max}$ を観測角度と視聴角度のずれとし、 th をあらかじめ定めた閾値とし、 α 、 β 、 γ をあらかじめ定めたパラメータとし、次式に基づいて上記モデル構築方法を決定するものである、

$$\gamma \theta_{\max} / (\alpha SE * \beta CM) > th$$

仮想環境構築装置。

[請求項34]

請求項32または33に記載の仮想環境構築装置であって、

上記動的オブジェクト構築部は、CGによりモデル構築する動的オブジェクトが実写映像によりモデル構築するオブジェクトに映り込んでいる場合に、上記実写映像中の上記動的オブジェクトが映り込んでいる領域について、上記動的オブジェクト含まない映像となるように上記実写映像を加工したものを、上記実写映像に代えて用いるものである、

仮想環境構築装置。

[請求項35] 複数のオブジェクトが存在する現実環境に基づいてユーザに閲覧させる仮想環境を構築する仮想環境構築方法であって、

モデル構築方法決定部が、上記複数のオブジェクトのうち少なくともいずれかのオブジェクトを対象オブジェクトとして、上記対象オブジェクトのセンシング難易度とCG描画難易度の少なくともいずれかに基づいて上記オブジェクトをCGによりモデル構築するか実写映像によりモデル構築するかを示すモデル構築方法を決定し、

状態センシング部が、上記現実環境における上記対象オブジェクトの状態を計測し、

環境モデル構築部が、時間変化を伴わない静的オブジェクトについて上記モデル構築方法に従って上記仮想環境中に描画するための環境モデルを構築し、

動的オブジェクト構築部が、時間変化を伴う動的オブジェクトについて上記モデル構築方法に従って上記仮想環境中に描画するためのモデルを構築して上記環境モデルに合成する、

仮想環境構築方法。

[請求項36] 複数フレームからなる映像シーケンスと上記映像シーケンス中の補完対象領域を指定するマスク情報とを入力として取得する映像入力部と、

上記映像シーケンスに含まれる複数フレーム間の差分を表す二値画像に基づいて前景領域と背景領域とに分離し、上記補完対象領域の周縁から一定距離以内に存在する前景領域に属する画素数に基づいて上記補完対象領域の補完方法をパッチ探索に基づくコンプリーションまたは貼り付け合成のいずれかに判定する補完方法判定部と、

上記補完方法に従って上記補完対象領域を補完する映像補完部と、
を含む映像提示装置。

[請求項37] 複数フレームからなる映像シーケンスと上記映像シーケンス中の補完対象領域を指定するマスク情報とを入力として取得する映像入力部

と、

上記補完対象領域内に存在する画素の平均色と上記補完対象領域の周縁から一定距離以内に存在する画素の平均色の差分に基づいて、上記補完対象領域の補完方法をパッチ探索に基づくコンプリーションまたは貼り付け合成のいずれかに判定する補完方法判定部と、

上記補完方法に従って上記補完対象領域を補完する映像補完部と、
を含む映像提示装置。

[請求項38]

複数フレームからなる映像シーケンスと上記映像シーケンス中の補完対象領域を指定するマスク情報とを入力として取得する映像入力部と、

上記映像シーケンス中の各フレームをセグメンテーションして予め定めた特徴を有する前景セグメントと背景セグメントとに分離し、上記補完対象領域の周縁から一定距離以内に存在する前景セグメントに属する画素数に基づいて上記補完対象領域の補完方法をパッチ探索に基づくコンプリーションまたは貼り付け合成のいずれかに判定する補完方法判定部と、

上記補完方法に従って上記補完対象領域を補完する映像補完部と、
を含む映像提示装置。

[請求項39]

請求項37に記載の映像提示装置であって、

上記補完方法判定部は、 Ω を上記補完対象領域とし、 Θ を上記補完対象領域の周縁から一定距離以内の領域とし、 $u(x, y, t)$ はフレーム t の (x, y) 座標に存在する画素の色情報とし、次式により求まる $U_{in}(t)$ と $U_{out}(t)$ との差 U_{diff} が予め定めた閾値以下または閾値より小さい場合に上記補完方法をパッチ探索に基づくコンプリーションと判定し、上記 U_{diff} が上記閾値以上または閾値より大きい場合に上記補完方法を貼り付け合成と判定するものである、

[数10]

$$U_{in}(t) = \frac{\sum_{(x,y) \in \Omega} u(x,y,t)}{\sum_{(x,y) \in \Omega} 1},$$

$$U_{out}(t) = \frac{\sum_{(x,y) \in \bar{\Omega} \cap \Theta} u(x,y,t)}{\sum_{(x,y) \in \bar{\Omega} \cap \Theta} 1}$$

映像提示装置。

[請求項40]

映像入力部が、複数フレームからなる映像シーケンスと上記映像シーケンス中の補完対象領域を指定するマスク情報とを入力として取得し、

補完方法判定部が、上記映像シーケンスに含まれる複数フレーム間の差分を表す二値画像に基づいて前景領域と背景領域とに分離し、上記補完対象領域の周縁から一定距離以内に存在する前景領域に属する画素数に基づいて上記補完対象領域の補完方法をパッチ探索に基づくコンプリーションまたは貼り付け合成のいずれかに判定し、

映像補完部が、上記補完方法に従って上記補完対象領域を補完する、

映像提示方法。

[請求項41]

映像入力部が、複数フレームからなる映像シーケンスと上記映像シーケンス中の補完対象領域を指定するマスク情報とを入力として取得し、

補完方法判定部が、上記補完対象領域内に存在する画素の平均色と上記補完対象領域の周縁から一定距離以内に存在する画素の平均色の差分に基づいて、上記補完対象領域の補完方法をパッチ探索に基づくコンプリーションまたは貼り付け合成のいずれかに判定し、

映像補完部が、上記補完方法に従って上記補完対象領域を補完する、

映像提示方法。

[請求項42]

映像入力部が、複数フレームからなる映像シーケンスと上記映像シ

ーケンス中の補完対象領域を指定するマスク情報とを入力として取得し、

補完方法判定部が、上記映像シーケンス中の各フレームをセグメンテーションして予め定めた特徴を有する前景セグメントと背景セグメントとに分離し、上記補完対象領域の周縁から一定距離以内に存在する前景セグメントに属する画素数に基づいて上記補完対象領域の補完方法をパッチ探索に基づくコンプリーションまたは貼り付け合成のいずれかに判定し、

映像補完部が、上記補完方法に従って上記補完対象領域を補完する、

映像提示方法。

[請求項43]

所定の位置を含む領域が撮影範囲となるように前記所定の位置を含む領域の周囲に設置された少なくとも2つの撮影装置それぞれが撮影した複数の入力画像を用いて、前記所定の位置を仮想的な視点である仮想視点として前記仮想視点からの距離を表す複数の異なる奥行き合成画像を生成し、生成した複数の合成画像それぞれからコントラストを算出し、算出した前記コントラストに基づいて複数の合成画像のうちコントラストが最も大きい合成画像の奥行きを各入力画像の最適奥行きに決定する決定部を備える最適奥行き決定装置。

[請求項44]

請求項43に記載の最適奥行き決定装置であって、

前記決定部は、コントラストが最も大きい合成画像の奥行きと、前記コントラストが最も大きい合成画像の奥行きの近傍の奥行きとの間において所定数の奥行きを求め、前記コントラストが最も大きい合成画像の奥行きと、選択した所定数の奥行きとを含む新たな奥行き候補群の中からコントラストが最も大きい合成画像の奥行きを決定する処理を実行し、前記処理を所定の回数繰り返し行った時点におけるコントラストが最も大きい合成画像の奥行きを前記最適奥行きに決定する、

最適奥行き決定装置。

[請求項45]

請求項43または44に記載の最適奥行き決定装置であって、

前記決定部は、前記撮影装置によって撮影された各時刻における入力画像の最適奥行きを決定し、決定した前記各時刻における最適奥行きの値を暫定最適奥行きとして時系列順に並べ、前記各時刻の暫定最適奥行きの値を時間方向に平滑化することによって得られる各時刻の最適奥行きの推定値を前記撮影装置によって撮影された各時刻における入力画像の最適奥行きに決定する、

最適奥行き決定装置。

[請求項46]

請求項43から45のいずれかに記載の最適奥行き決定装置であって、

前記決定部は、ある時刻において前記撮影装置それぞれで撮影された各入力画像の最適奥行きを、前記ある時刻よりも前の時刻における前記撮影装置の入力画像の最適奥行きに基づいて決定する、

最適奥行き決定装置。

[請求項47]

所定の位置を含む領域が撮影範囲となるように前記所定の位置を含む領域の周囲に設置された少なくとも2つの撮影装置それぞれが撮影した複数の入力画像を用いて、前記所定の位置を仮想的な視点である仮想視点として前記仮想視点からの距離を表す複数の異なる奥行きの合成画像を生成し、生成した複数の合成画像それぞれからコントラストを算出し、算出した前記コントラストに基づいて複数の合成画像のうちコントラストが最も大きい合成画像の奥行きを各入力画像の最適奥行きに決定する決定ステップを有する最適奥行き決定方法。

[請求項48]

請求項1から5、7から11、20から23、30、32から34のいずれかに記載の仮想環境構築装置、請求項14から18、36から39のいずれかに記載の映像提示装置、請求項24または25に記載のモデル学習装置、もしくは請求項43から46のいずれかに記載の最適奥行き決定装置としてコンピュータを機能させるためのプログ

ラム。

[図1]

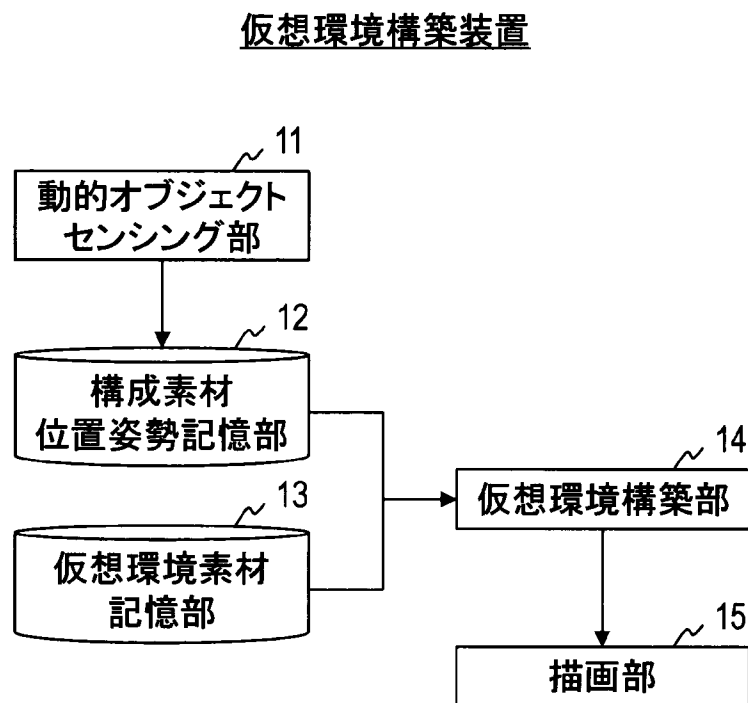


図 1

[図2]

構成素材ID	仮想環境素材ID	時刻	位置	姿勢
Oi	Ei	0	Ti(0)	Ri(0)
		1	Ti(1)	Ri(1)
		...	...	...
		ti	Ti(ti)	Ri(ti)
Oj	Ej	0	Tj(0)	Rj(0)
		1	Tj(1)	Rj(1)
		...	...	...
		tj	Tj(tj)	Rj(tj)
...	...	...	...	...
Ok	Ek	-	Tk	Rk
...	...	...	...	...

図2

[図3]

時刻	構成素材ID	例
1	O1	スタジアムを組み込む
2	-	
...	...	...
t1	O2	ピッチャーを組み込む
...	...	...
t2	O3	ボールを組み込む

図3

[図4]

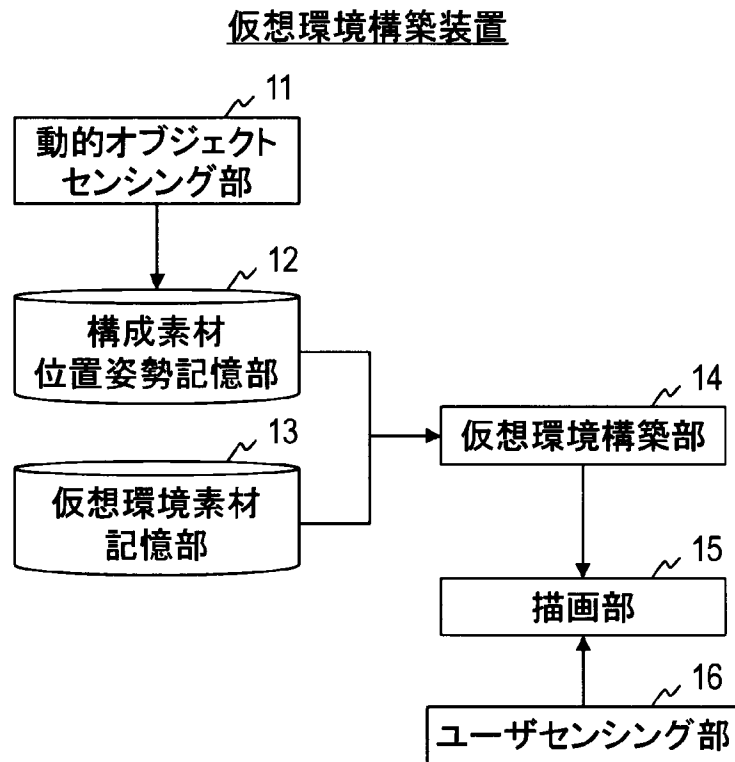


図4

[図5]

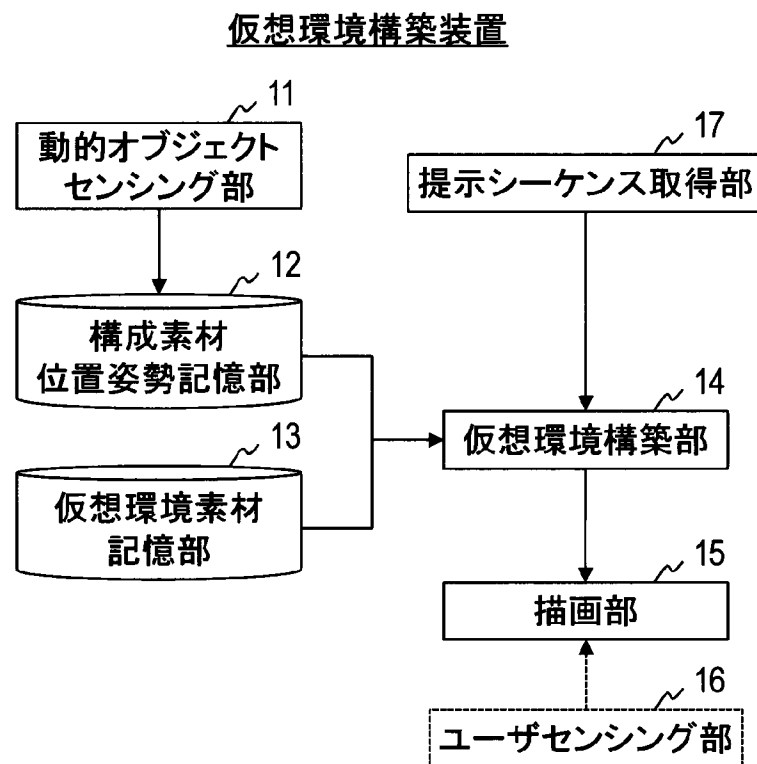


図5

[図6]

練習意図	時刻	構成素材ID
〇〇の能力を向上させたい	1	-
	2	-
	...	...
	t1	O1
	...	...
	t2	O2, O3
△△の能力を向上させたい	1	O4
	2	O5
	...	
	tj	O6
...	...	...

図6

[図7]

練習意図	時刻	構成素材ID
緩急に対する能力を向上させたい	1	O1 (140km/hのストレート)
	2	-
	...	...
	t1	O2 (110km/hのカーブ)
	...	...
	t2	O3 (150km/hのストレート)
選球眼を向上させたい	1	O4 (コース2-Aのストレート)
	2	-
	...	...
	tj	O5 (コース2-Bのストレート)
...	...	...

図7

[図8]

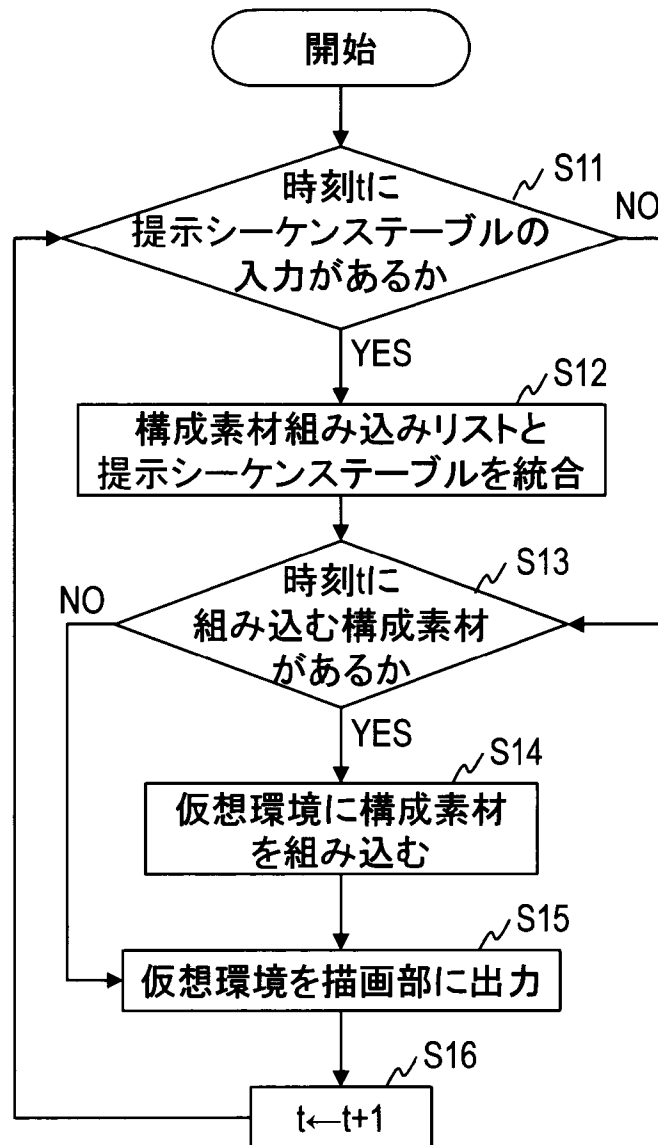


図8

[図9]

構成素材ID	球速 (km/h)	コース	球種
O1	140	1-A	ストレート
O2	120	3-B	カーブ
...	...	...	...
Oi	135	2-C	フォーク
...	...	...	...

図9

[図10]

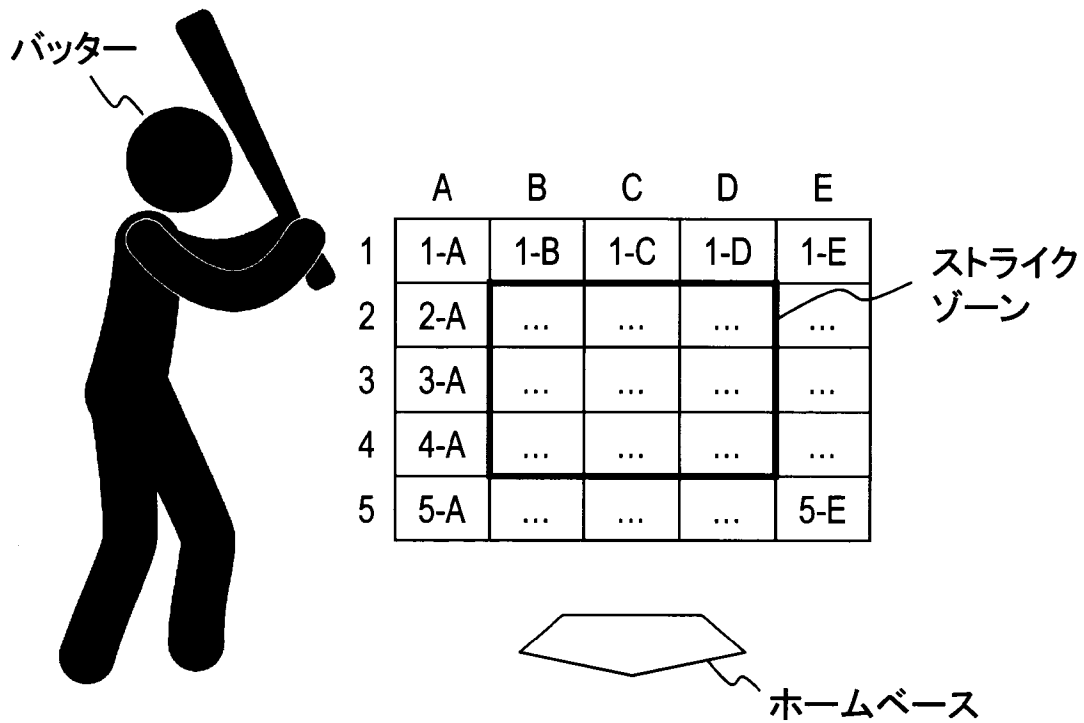


図10

[図11]

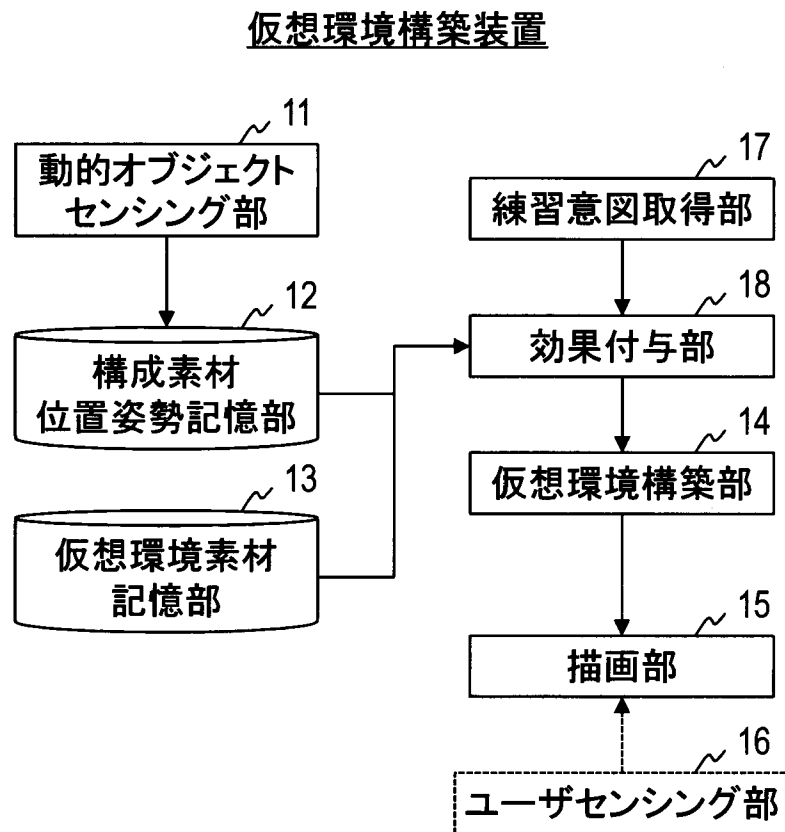


図11

[図12]

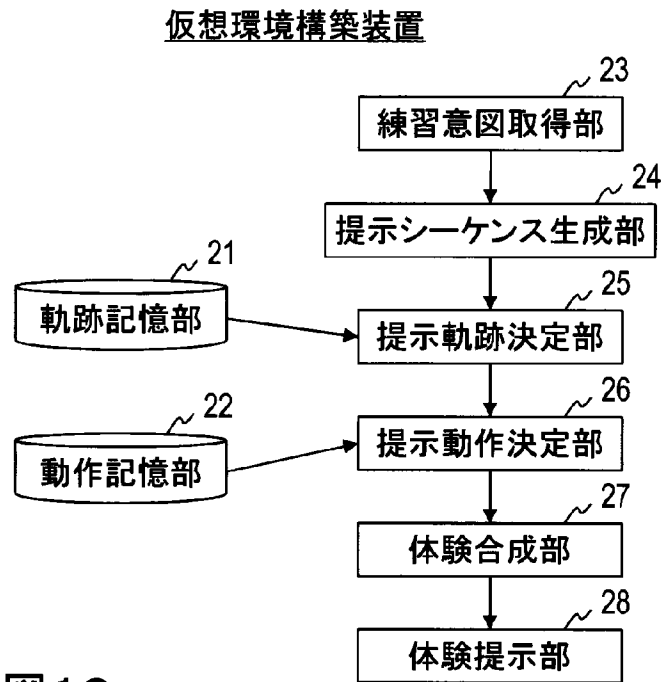


図12

[図13]

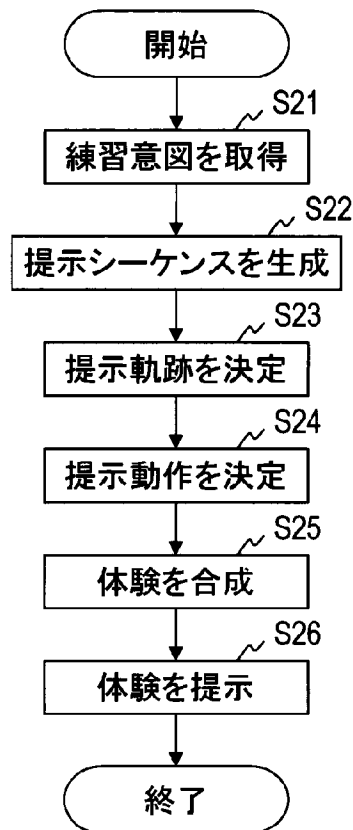


図13

[図14]

動作ID	相手選手ID	時刻	動作
1	1	1	フレーム画像1
		2	フレーム画像2
		⋮	⋮
		T	フレーム画像T
2	1	1	手(x_{11}, y_{11}, z_{11})、腕(x_{21}, y_{21}, z_{21})
		2	手(x_{12}, y_{12}, z_{12})、腕(x_{22}, y_{22}, z_{22})
		⋮	⋮
		T	手(x_{1T}, y_{1T}, z_{1T})、腕(x_{2T}, y_{2T}, z_{2T})
⋮	⋮	⋮	⋮
n	1	⋮	⋮
1	2	⋮	⋮
2	2	⋮	⋮
⋮	⋮	⋮	⋮

図14

[図15]

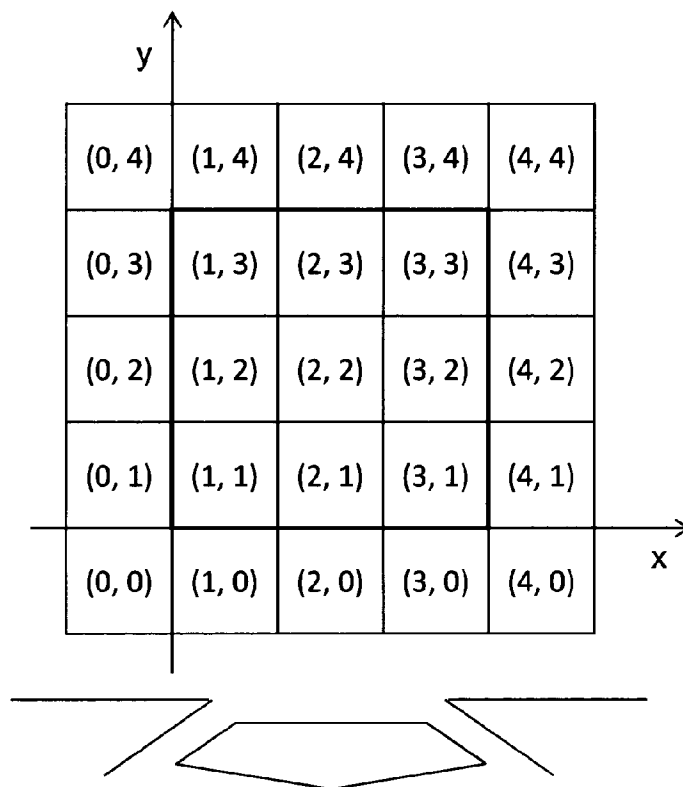


図15

[図16]

上左 外	上外		上右 外
	上		
中左 外	真ん中 左	真ん中 右	中右 外
下左 外	左下	右下	下右 外
	下外		

図16

[図17]

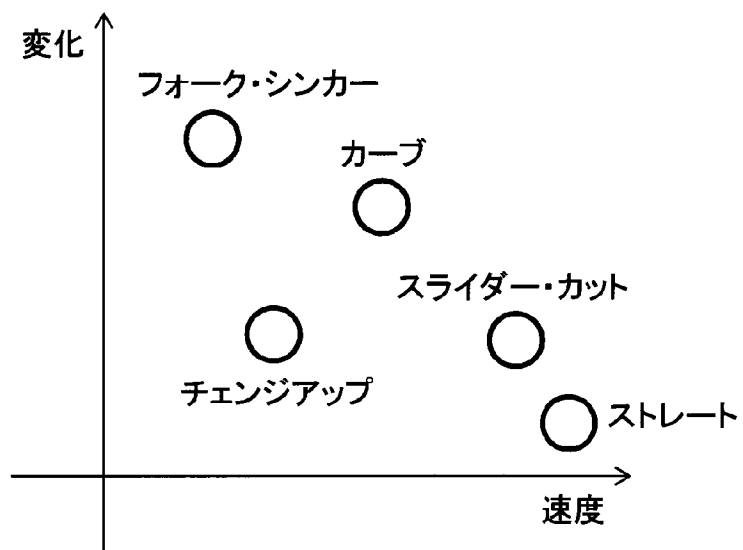


図17

[図18]

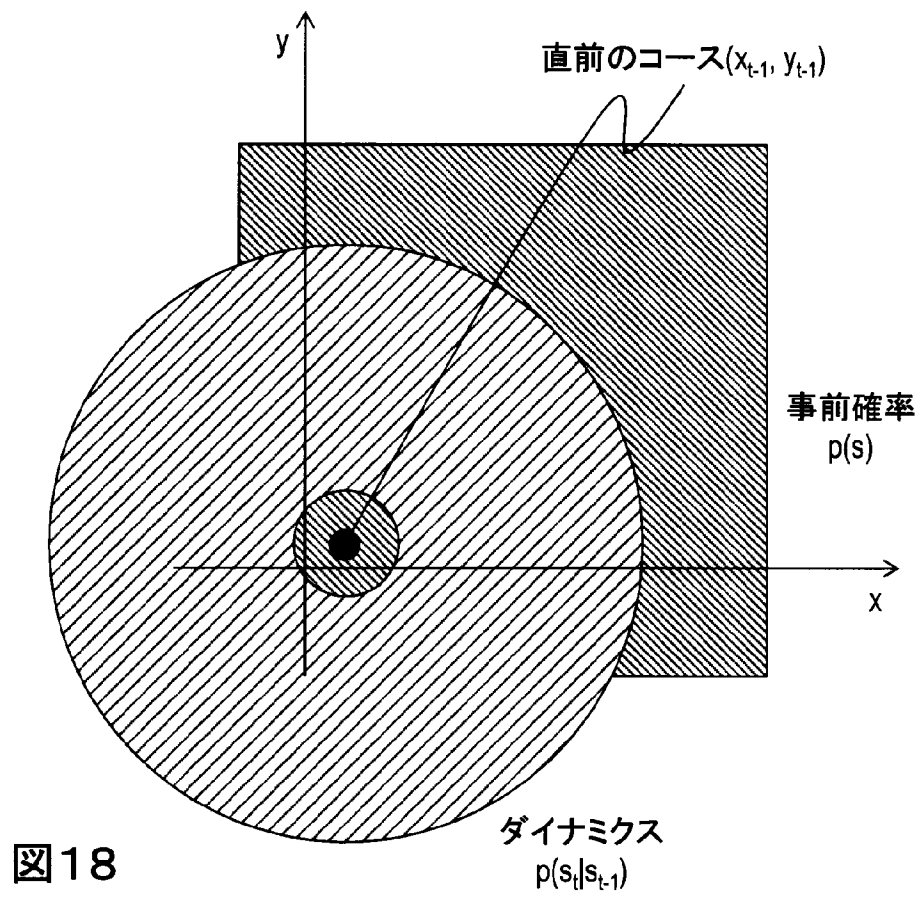


図18

[図19]

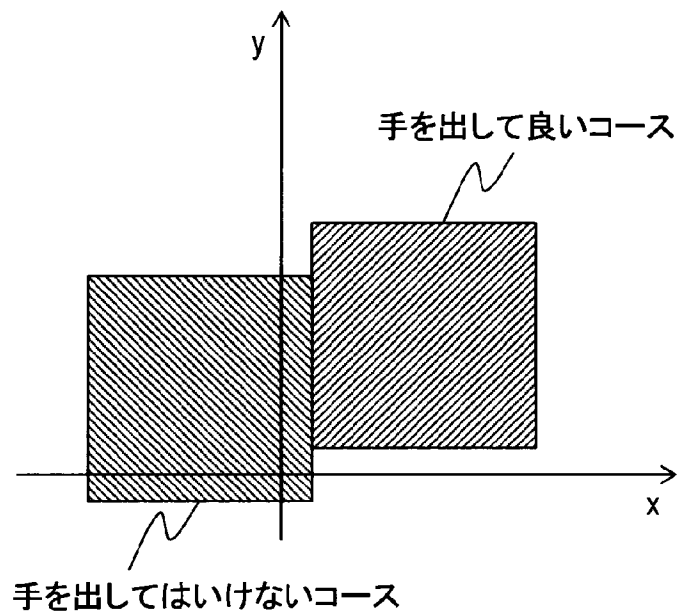


図19

[図20]

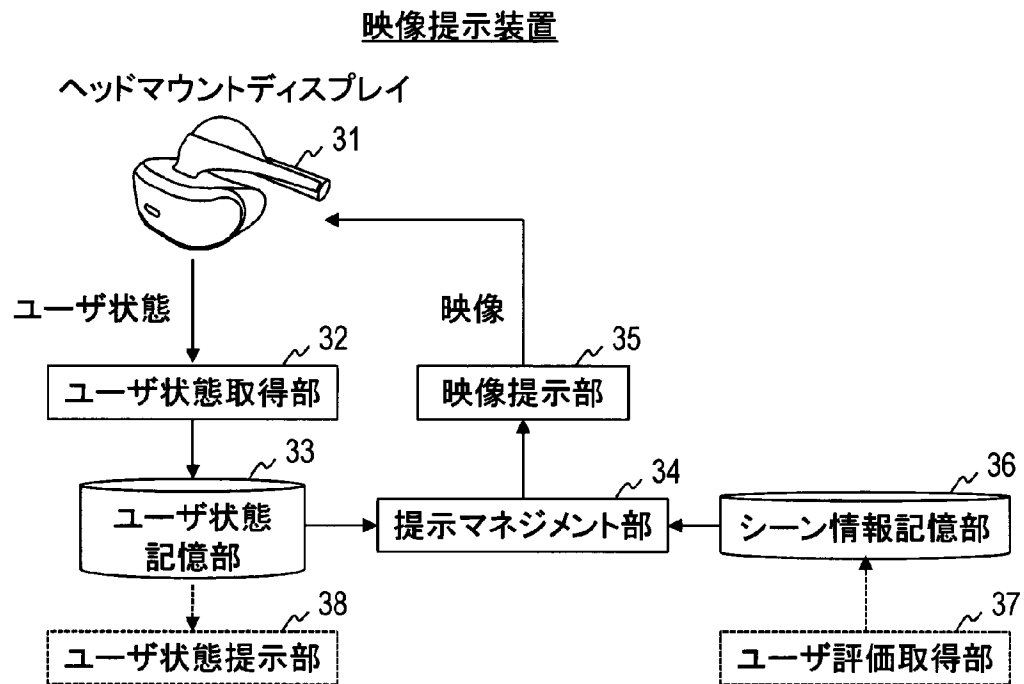


図20

[図21]

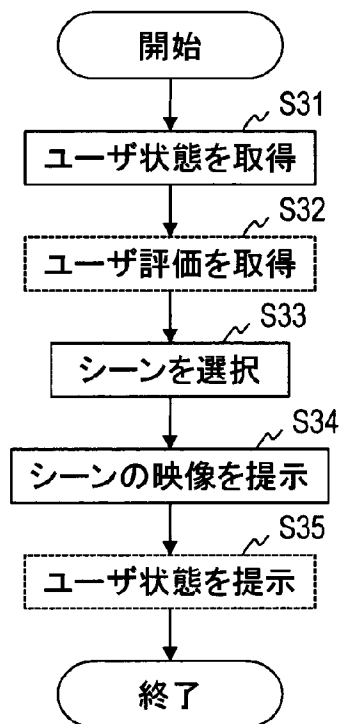


図21

[図22]

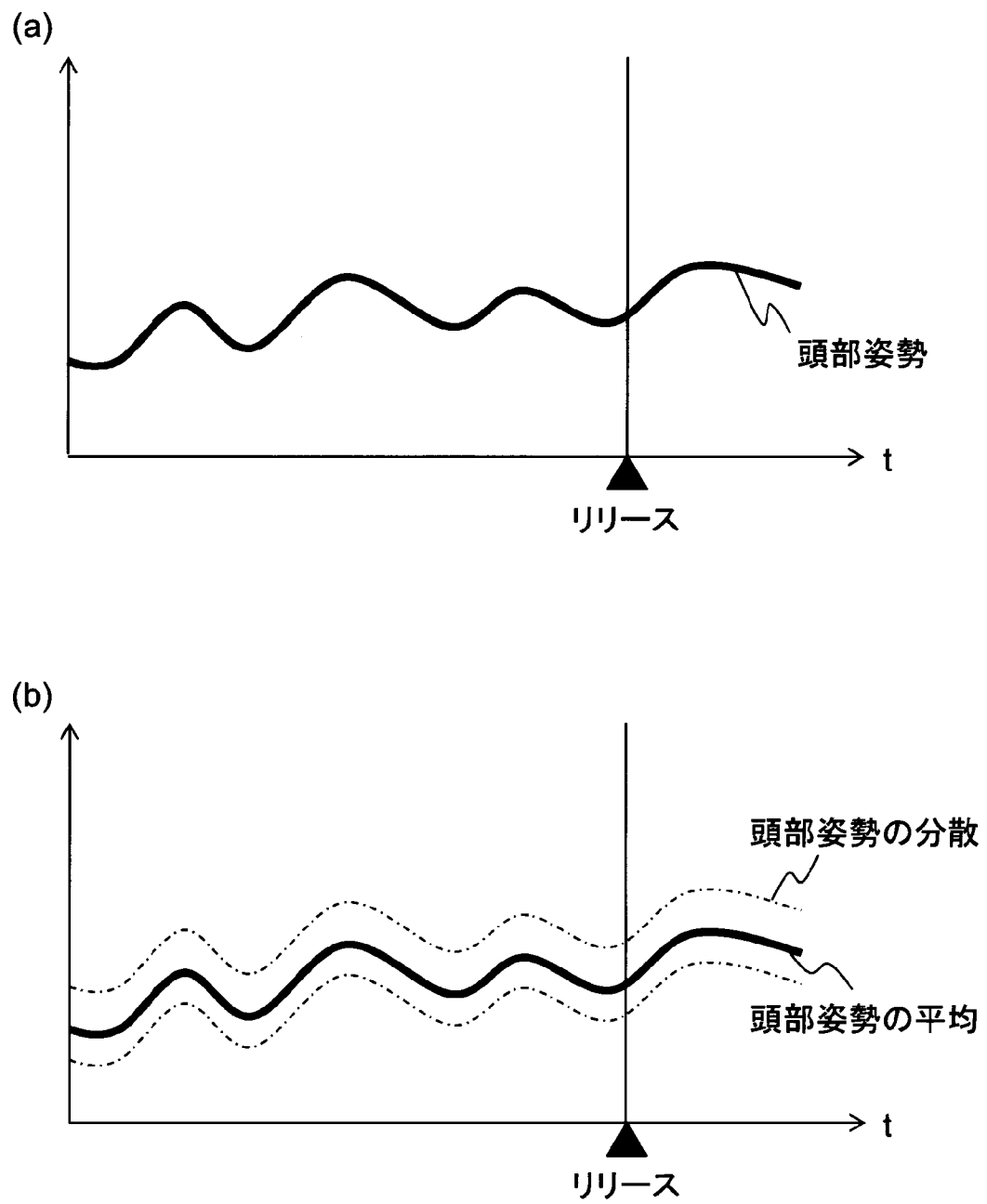


図22

[図23]

選手A

選手B

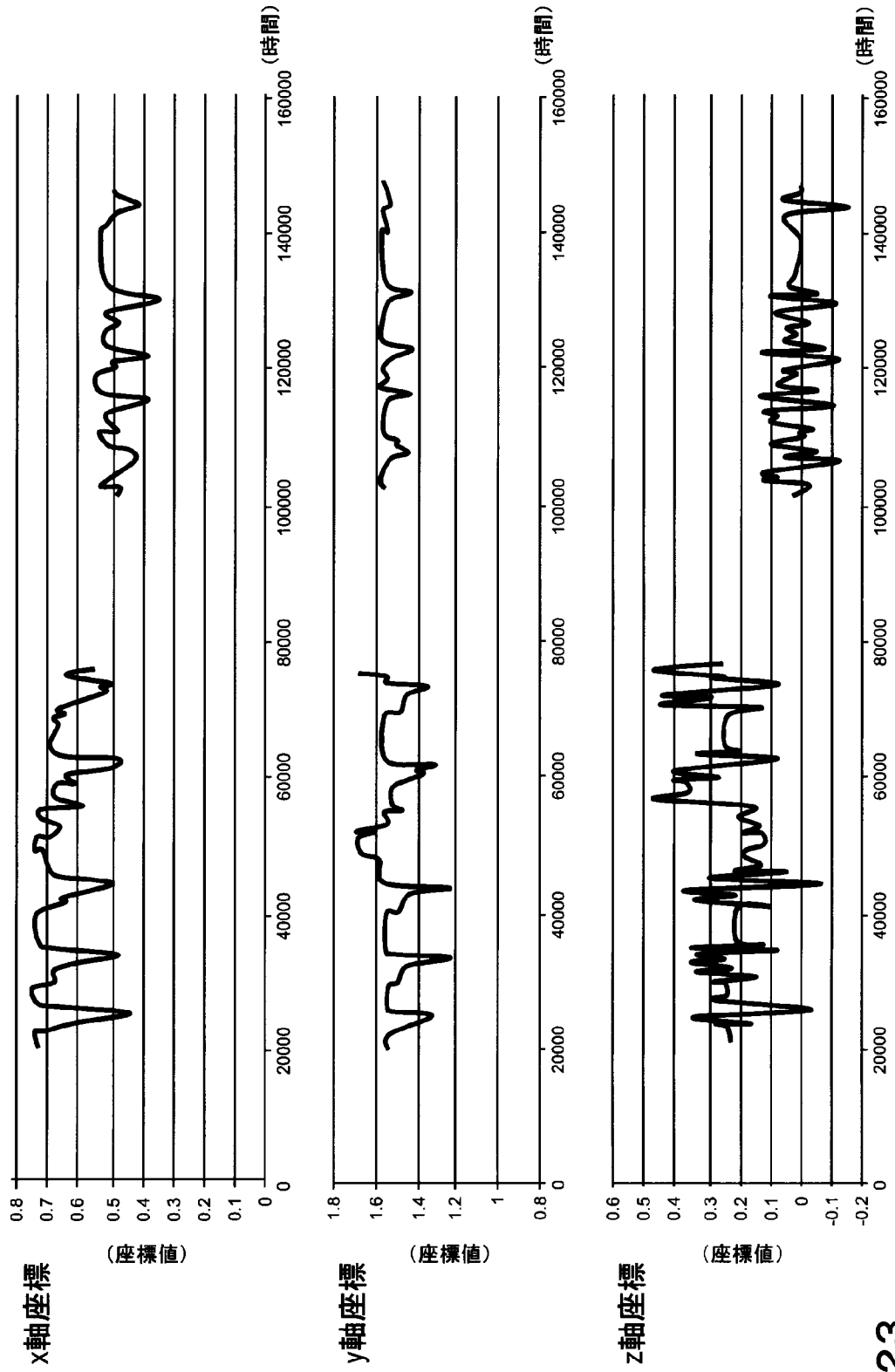


図23

[図24]

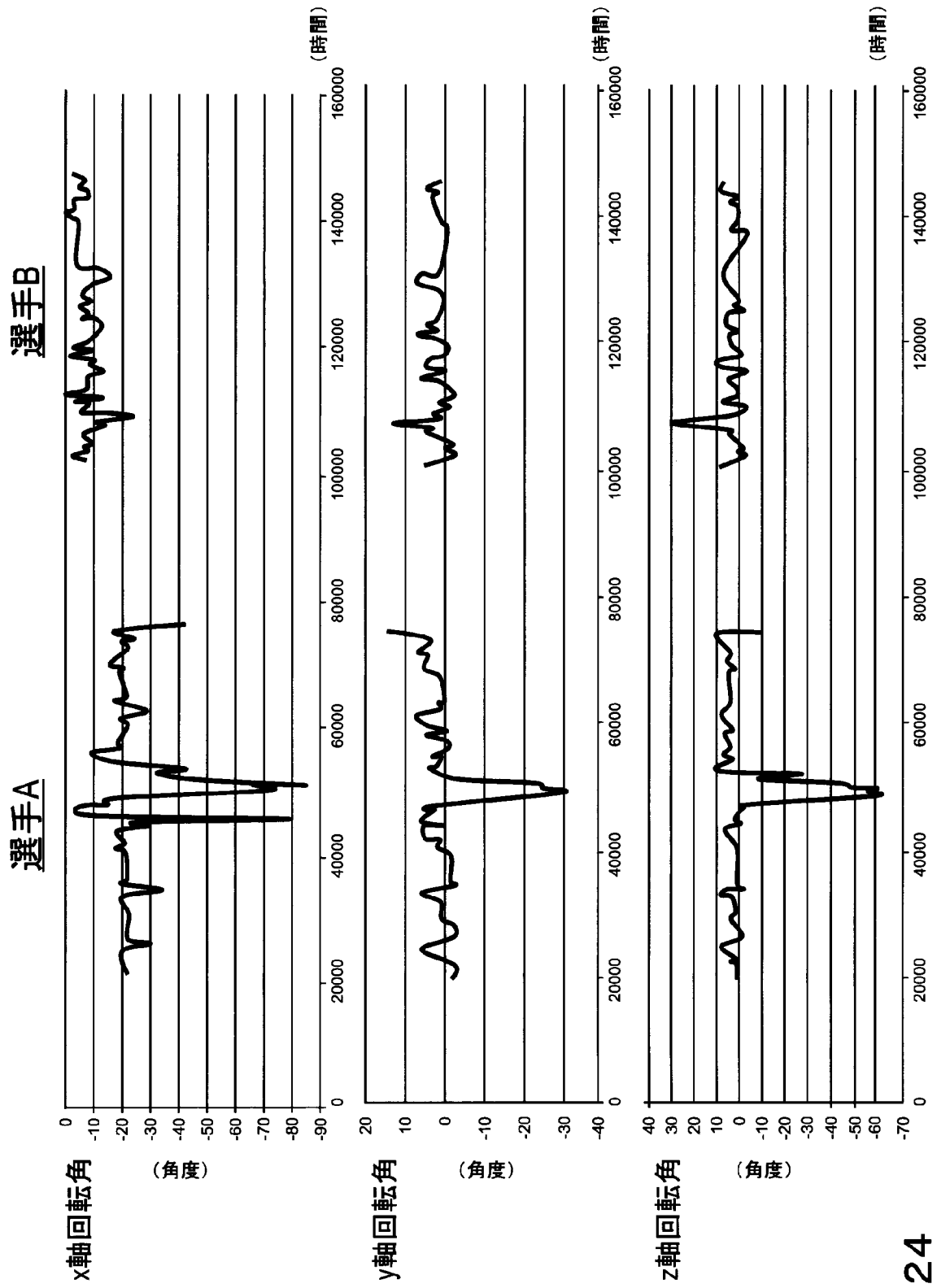


図24

[図25]

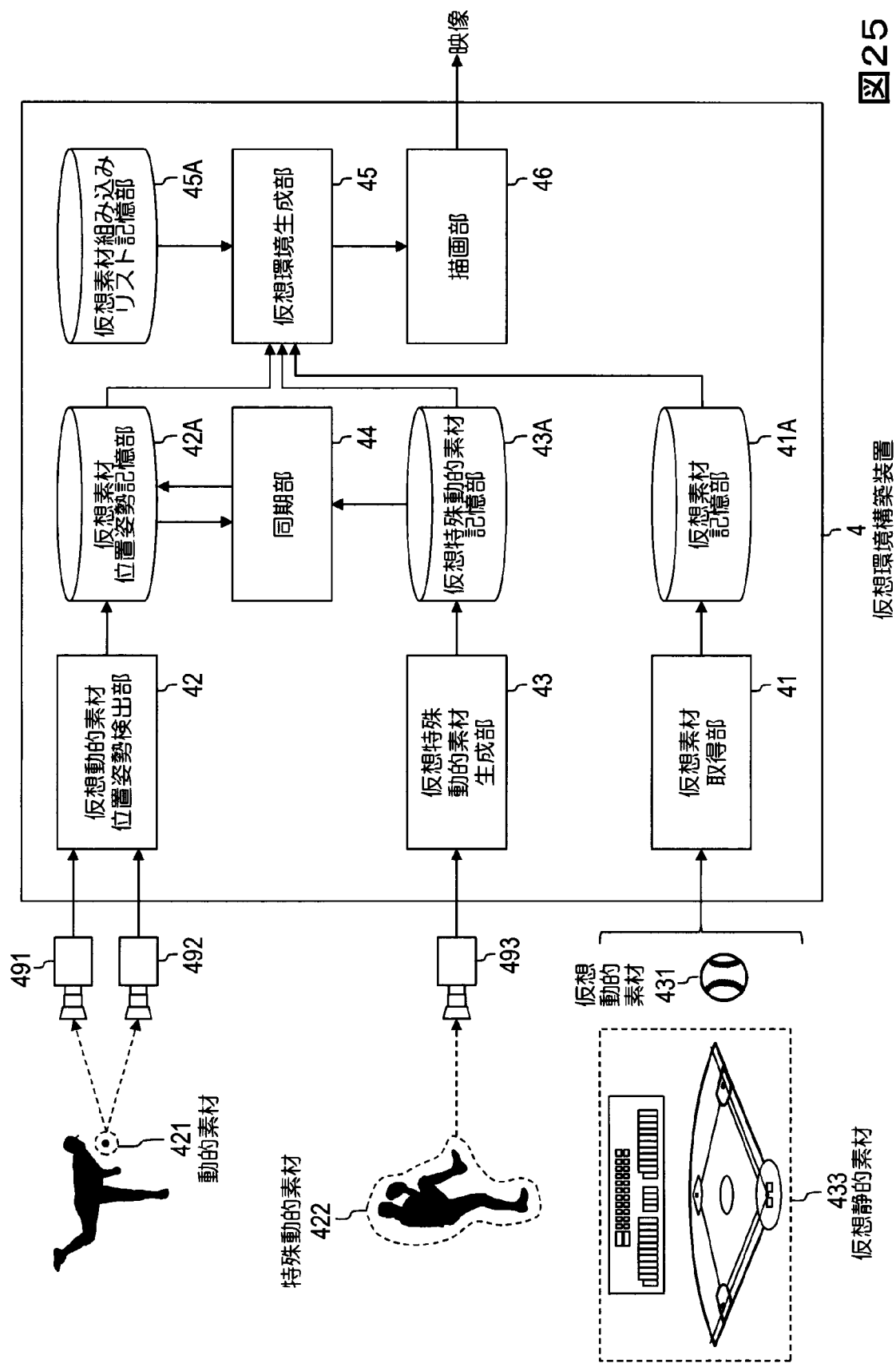


図25

[図26]

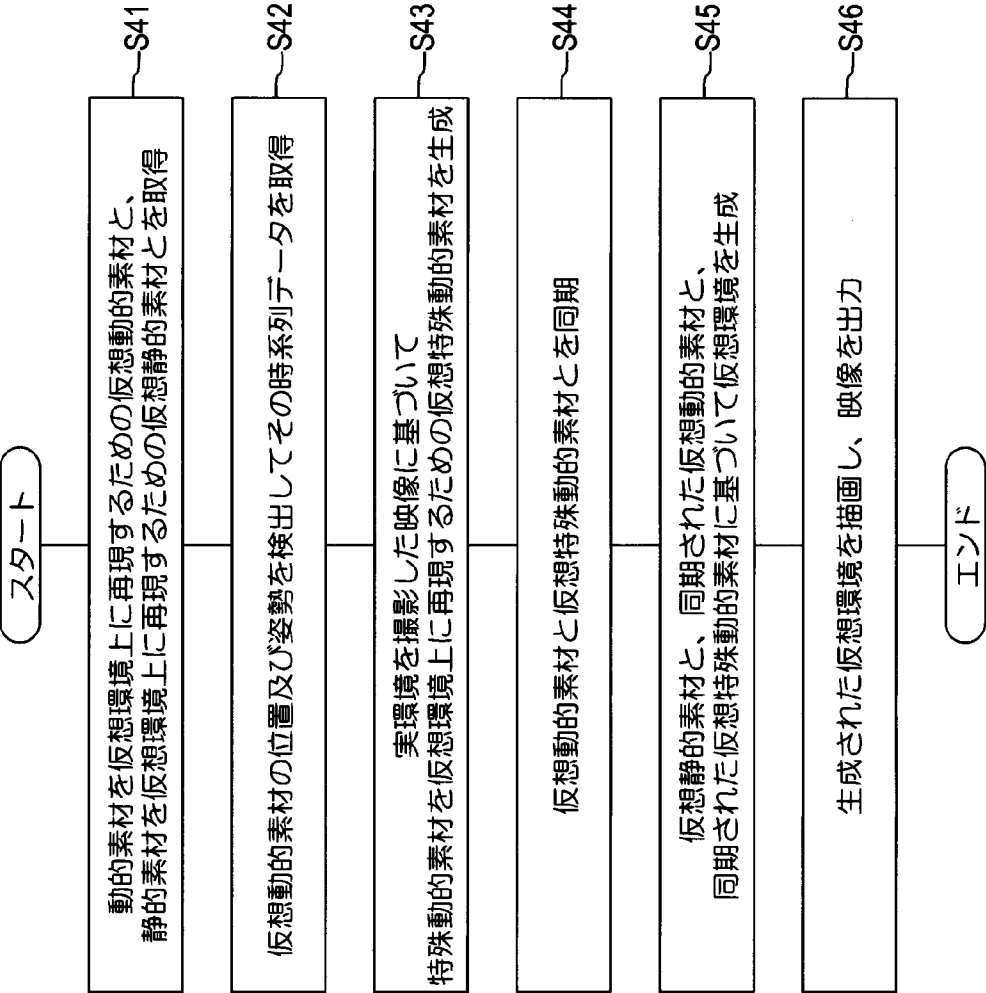


図26

[図27]

仮想素材位置姿勢情報テーブル									
素材ID	仮想素材ID	時刻	位置	姿勢	同期素材ID (1)	同期開始時刻 (1)	同期素材ID (2)	同期開始時刻 (2)	...
Oi	Ei	1	Ti(1)	Ri(1)	-	-	-	-	...
		2	Ti(2)	Ri(2)					
		...	...	...					
		ti	Ti(ti)	Ri(ti)					
Oj	Ej	1	Tj(1)	Rj(1)	Oi	Tji	-	-	...
		2	Tj(2)	Rj(2)					
		...	...	...					
		tj	Tj(tj)	Rj(tj)					
...	...	...	...	...	...	...	...	...	...
Ok	Ek	-	Tk	Rk	Oa	Tka	Ob	Tkb	...
...	...	...	...	...	...	...	...	...	...

図27

[図28]

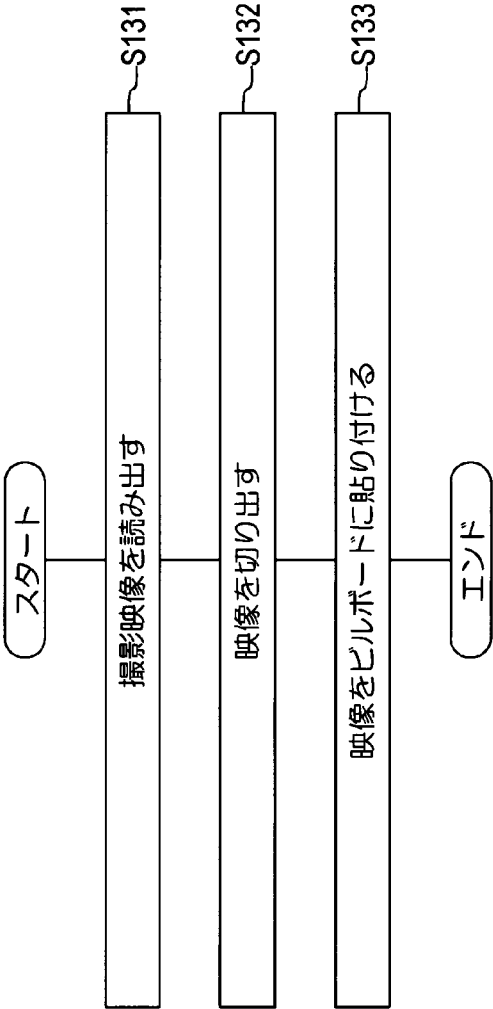


図28

[図29]

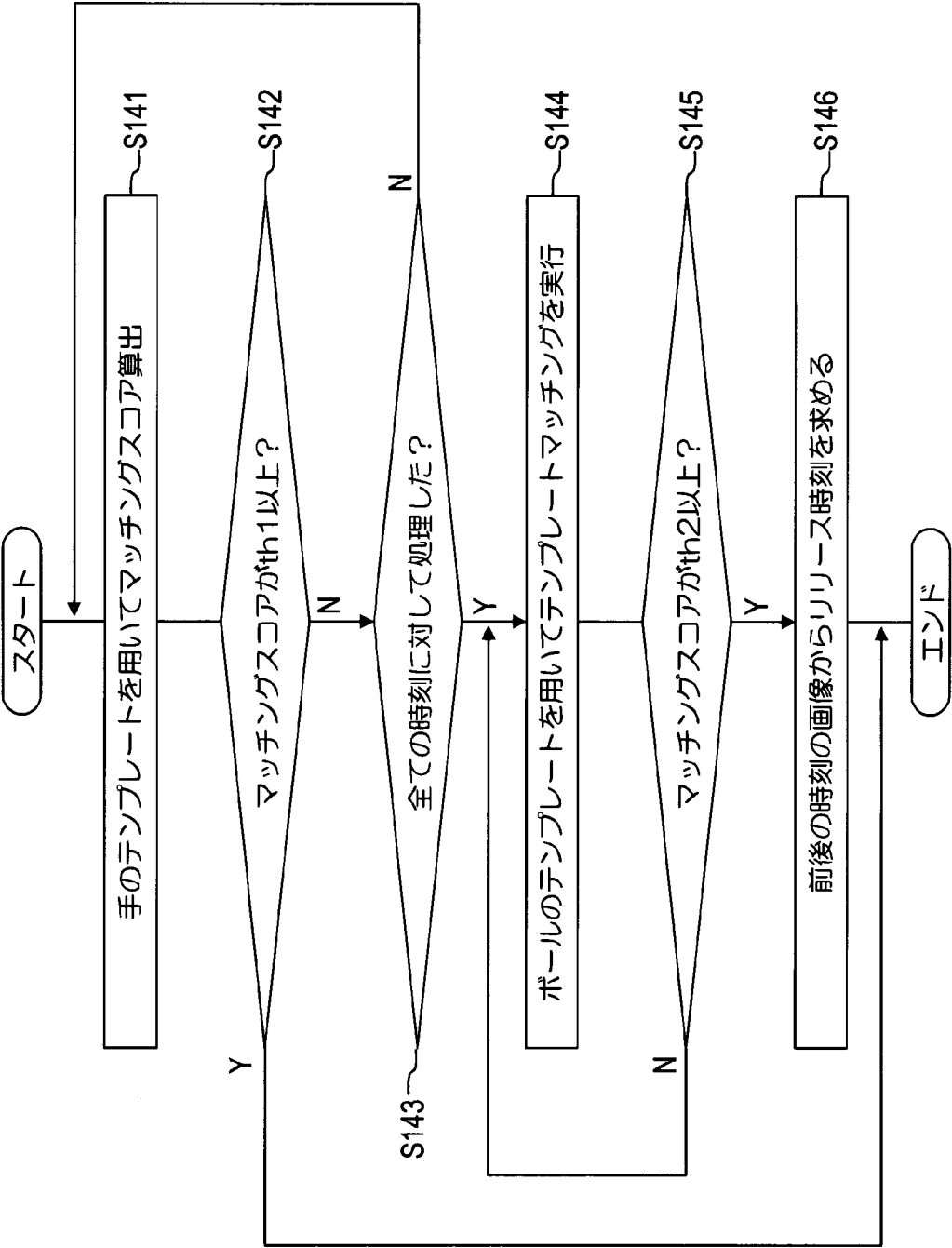
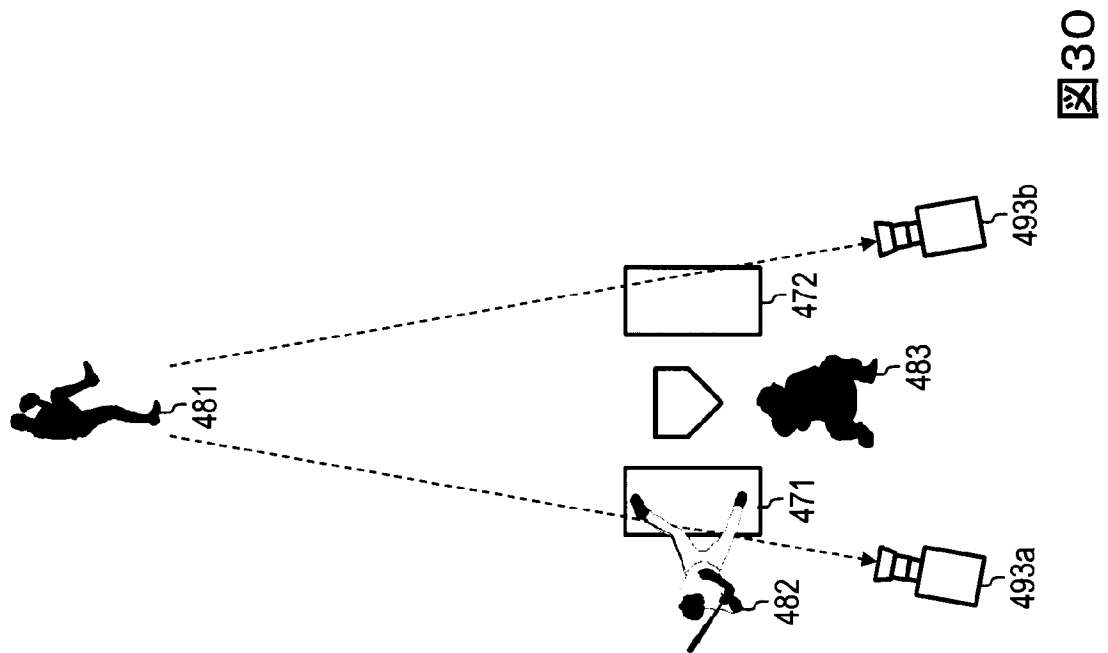


図29

[図30]



[図30]

[図31]

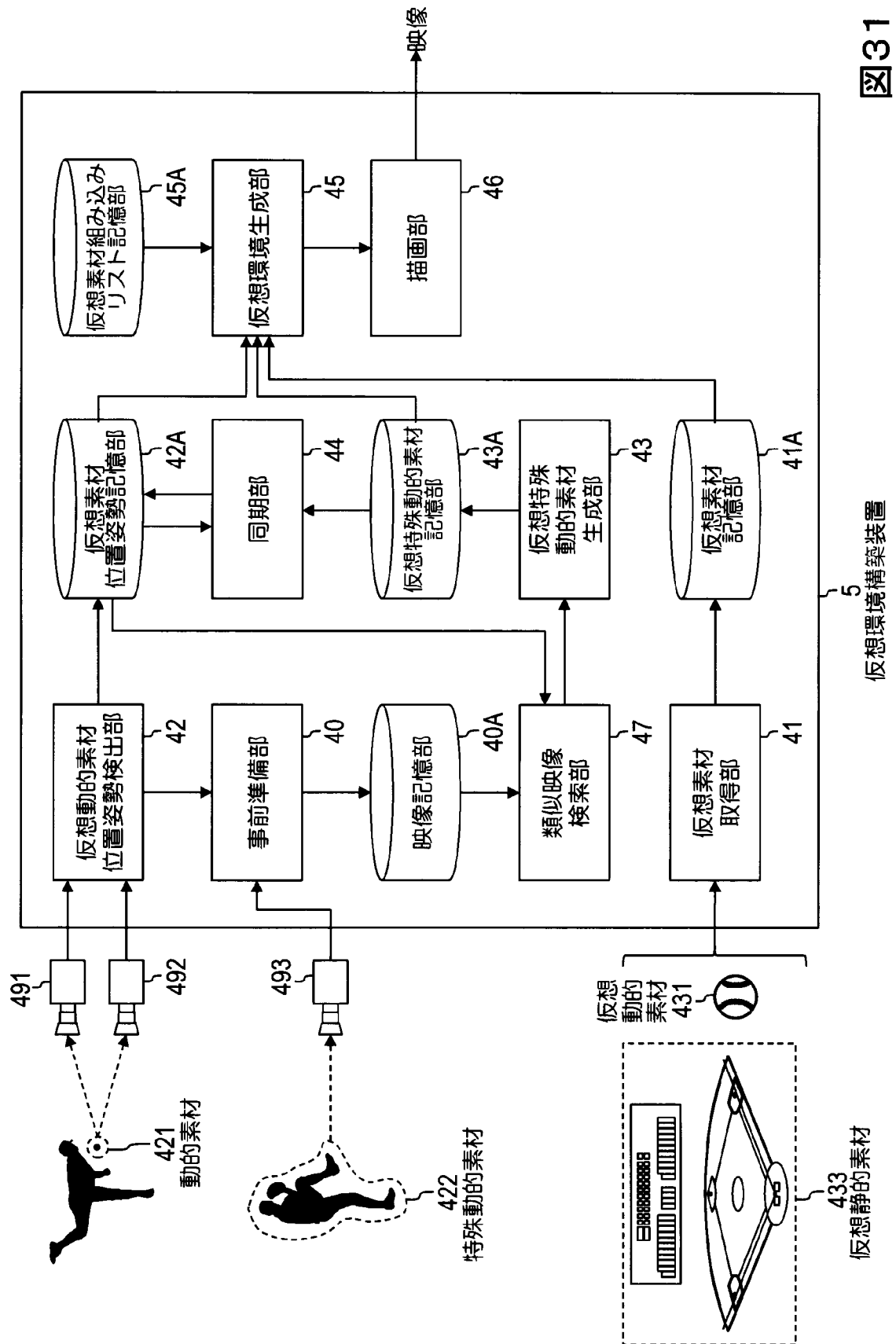


図31

[図32]

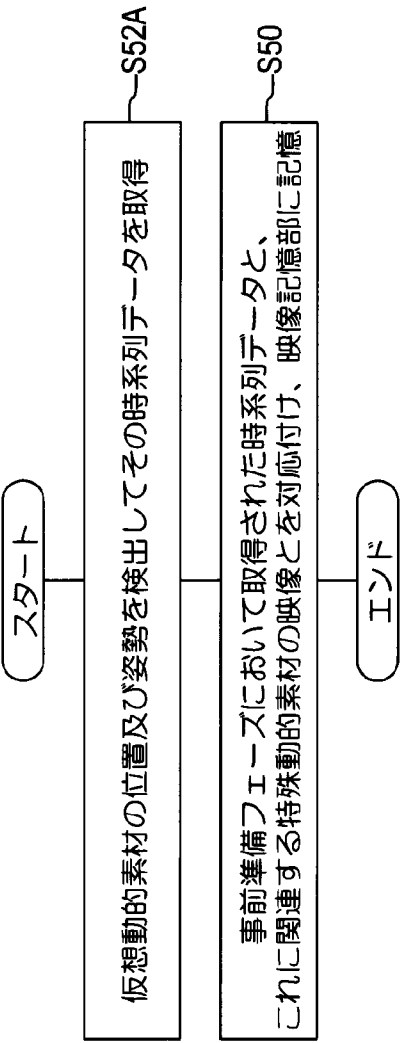


図32

[図33]

映像情報テーブル

映像ID	関連 動的素材ID (1)	時刻 (1)	位置 (1)	姿勢 (1)	関連 動的素材ID (2)	時刻 (2)	位置 (2)	姿勢 (2)	...
Vi	Di1	1	Ti1(1)	Ri1(1)	Di2	1	Ti2(1)	Ri2(1)	...
		2	Ti1(2)	Ri1(2)		2	Ti2(2)	Ri2(2)	...
		...	...	...		...	...	...	...
		ti	Ti1(ti)	Ri1(ti)		ti	Ti2(ti)	Ri2(ti)	...
...	...	...	...	...	...	...	...	...	...

図33

[図34]

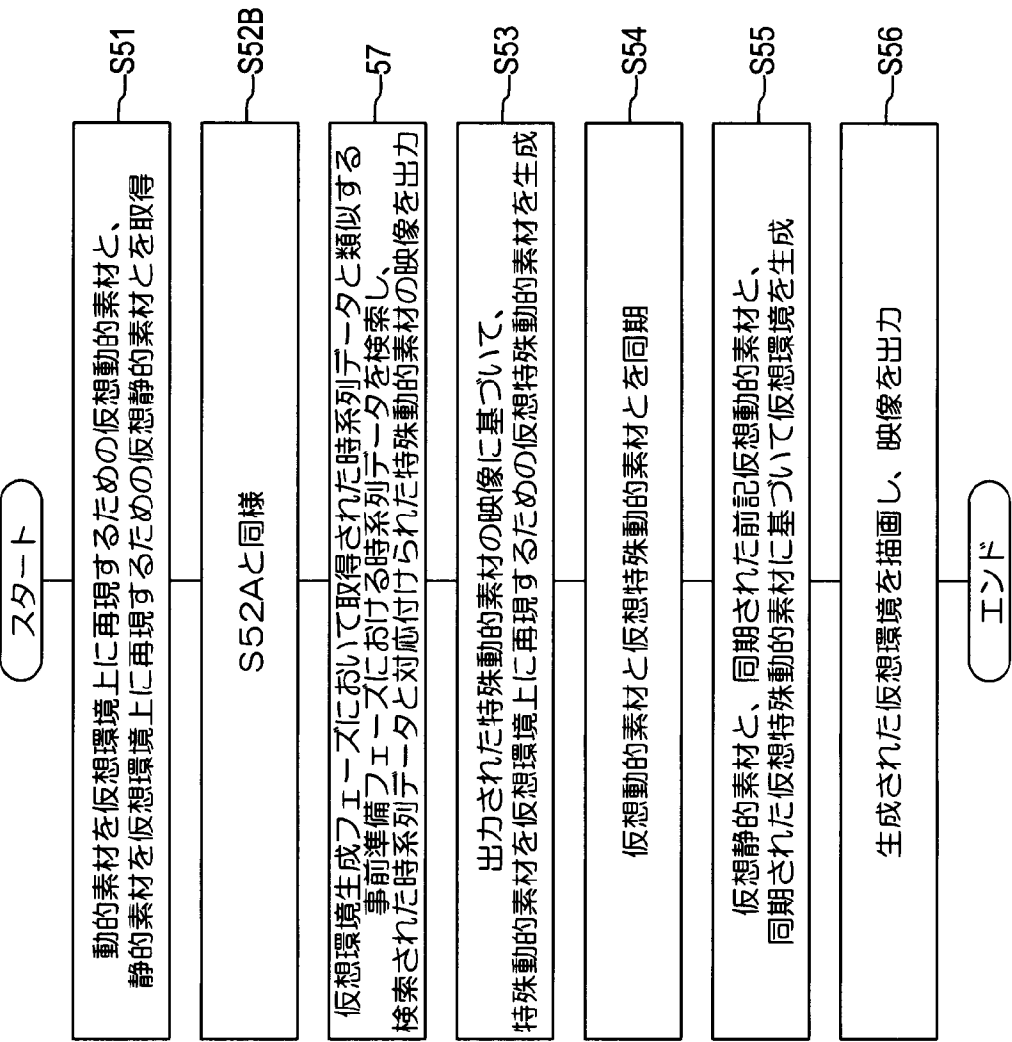


図34

[図35]

仮想素材位置姿勢情報テーブル

素材ID	仮想素材ID	時刻	位置	姿勢	同期素材ID (1)	同期開始時刻 (1)	同期素材ID (2)	同期開始時刻 (2)	...
Oi	Ei	1	Ti(1)	Ri(1)	-	-	-	-	...
		2	Ti(2)	Ri(2)					
		...	...	...					
		ti	Ti(ti)	Ri(ti)					
Oj	Ej	1	Tj(1)	Rj(1)	Oi	Tji	-	-	...
		2	Tj(2)	Rj(2)					
		...	...	...					
		tj	Tj(tj)	Rj(tj)					
...	...	...	...	...	...	...	...	...	...
Ok	Ek	-	Tk	Rk	Oa	Tka	Ob	Tkb	...
...	...	...	...	...	...	...	...	...	...

図35

[図36]

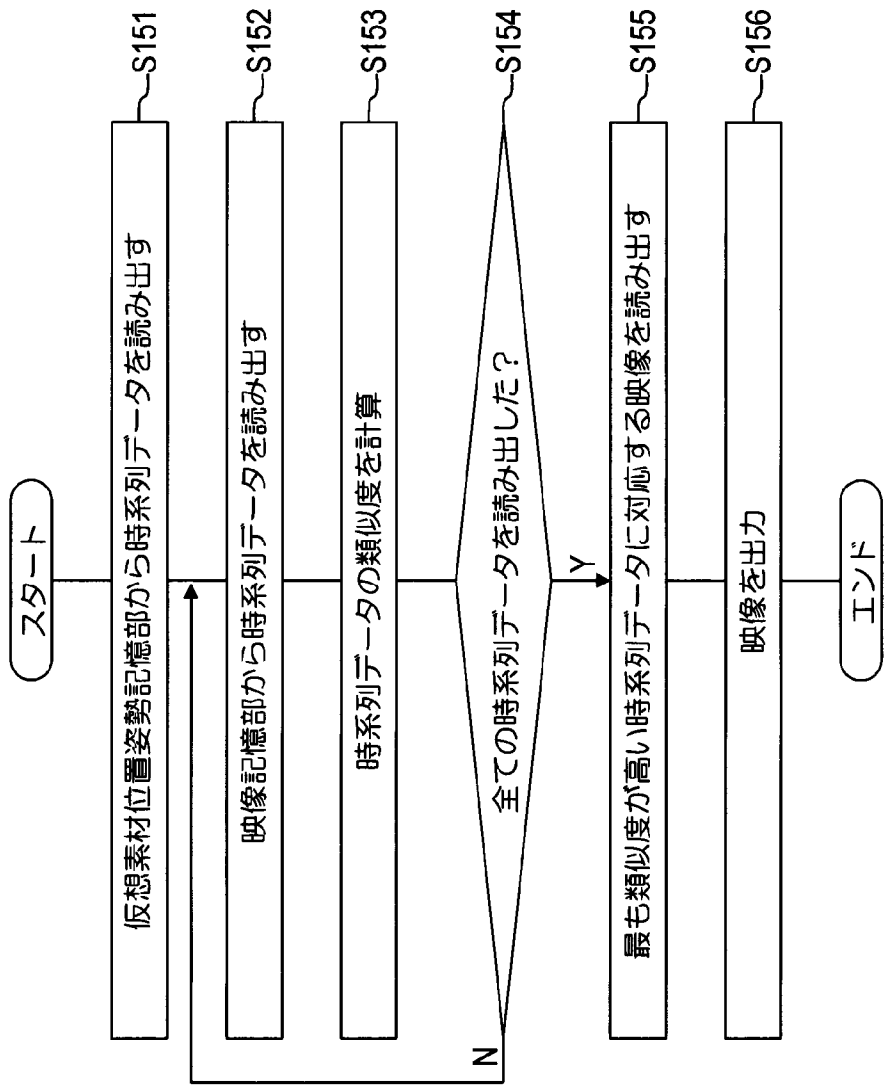


図36

[図37]

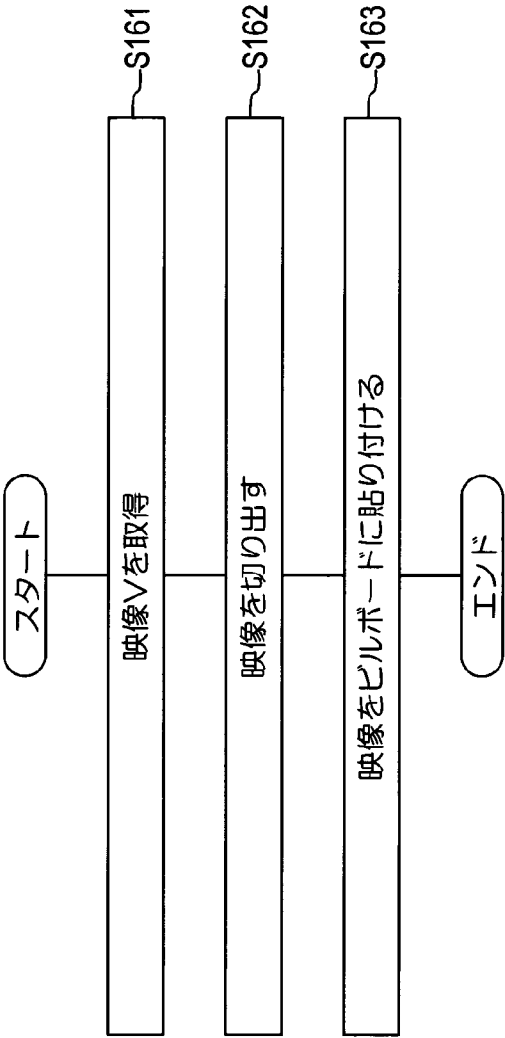


図37

[図38]

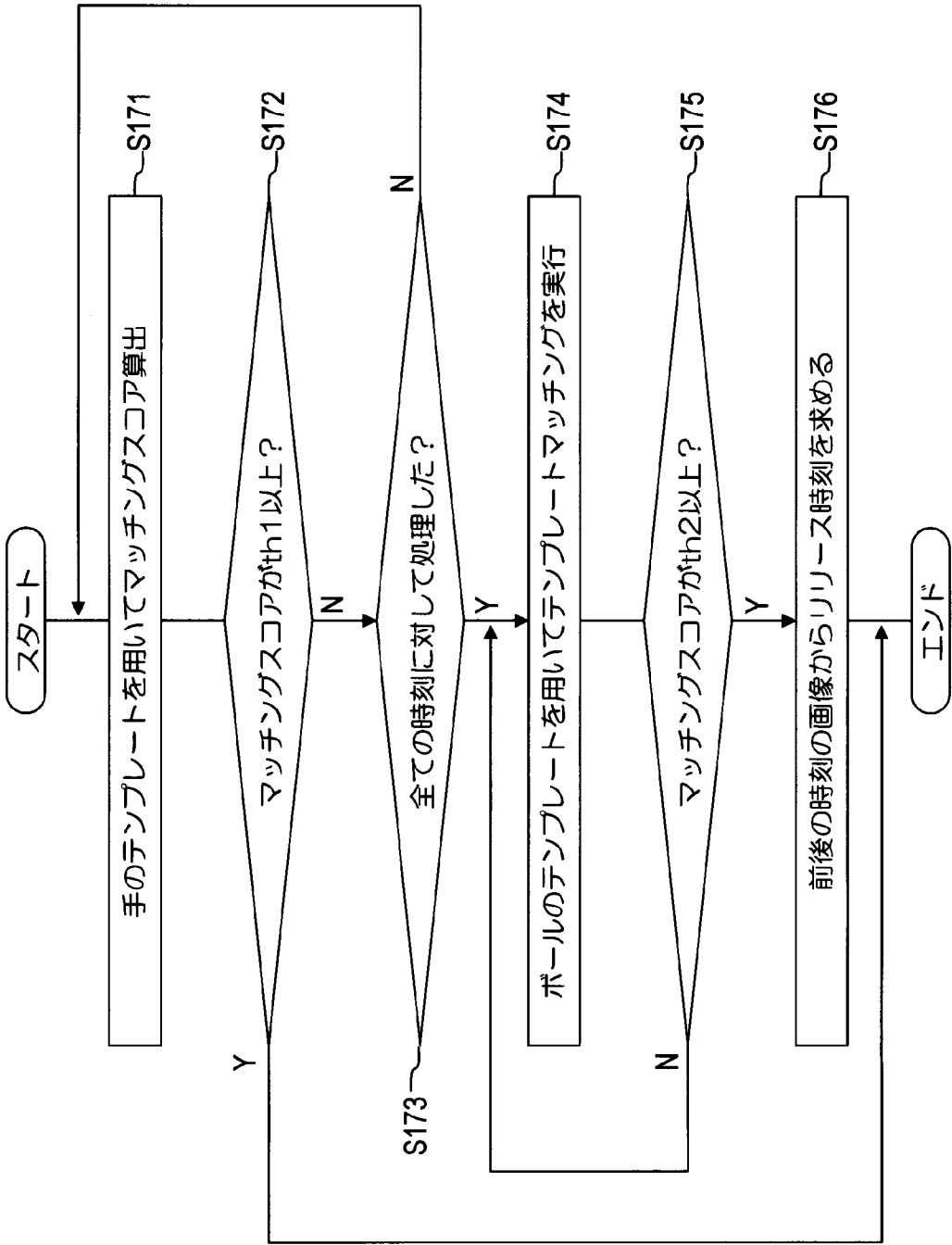


図38

[図39]

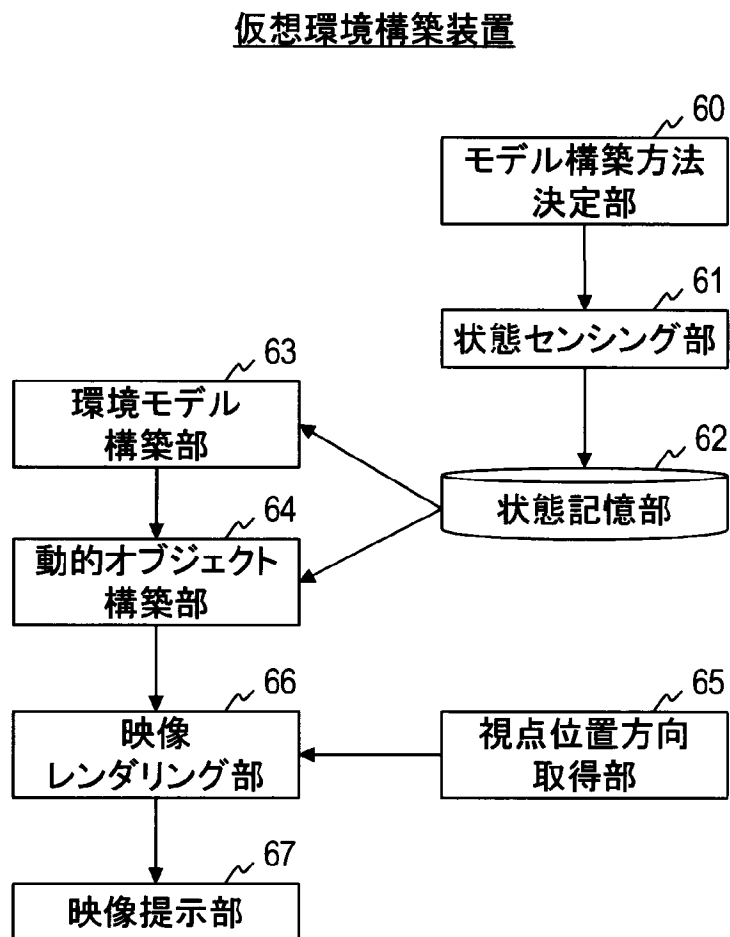


図39

[図40]

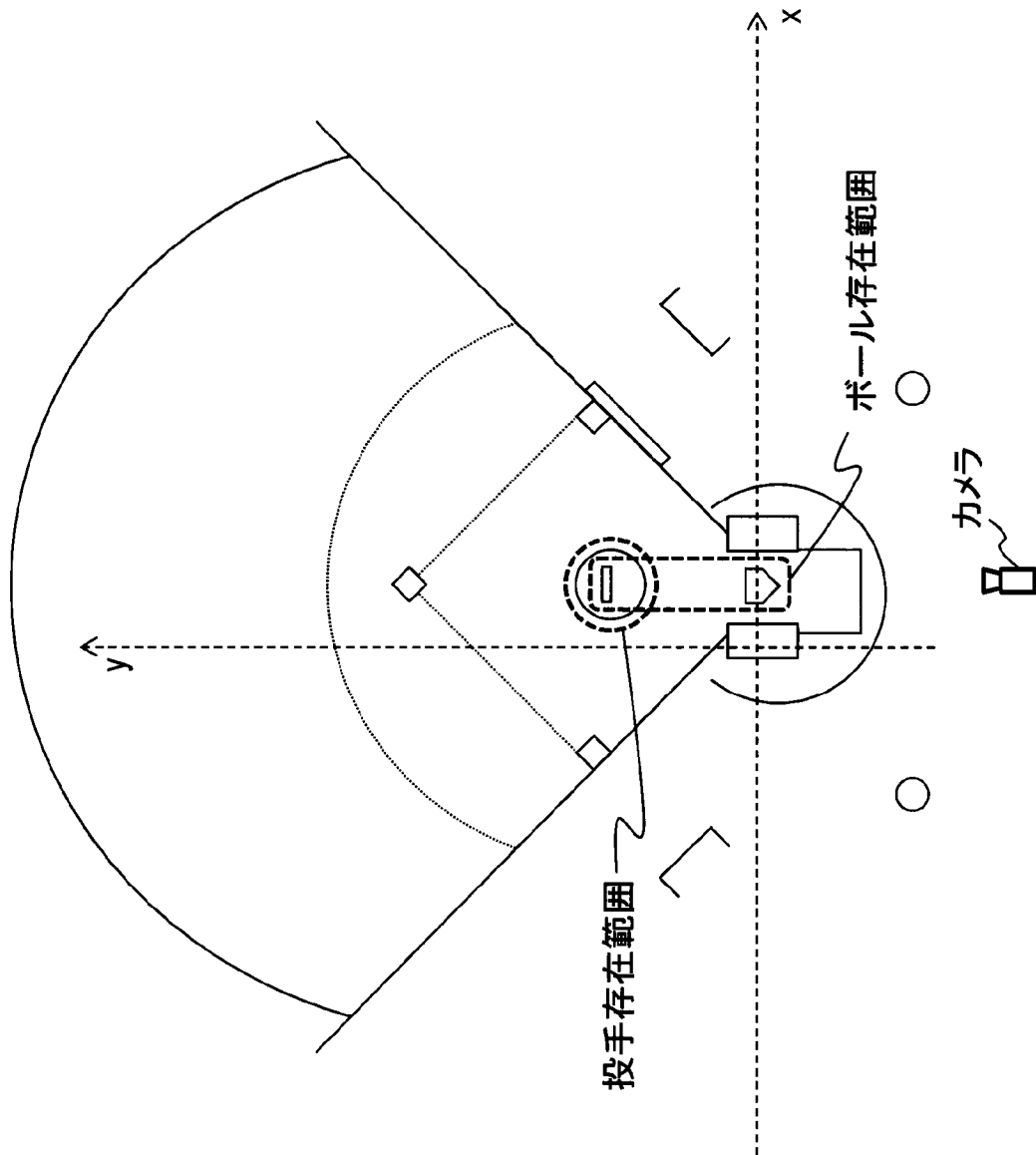


図40

[図41]

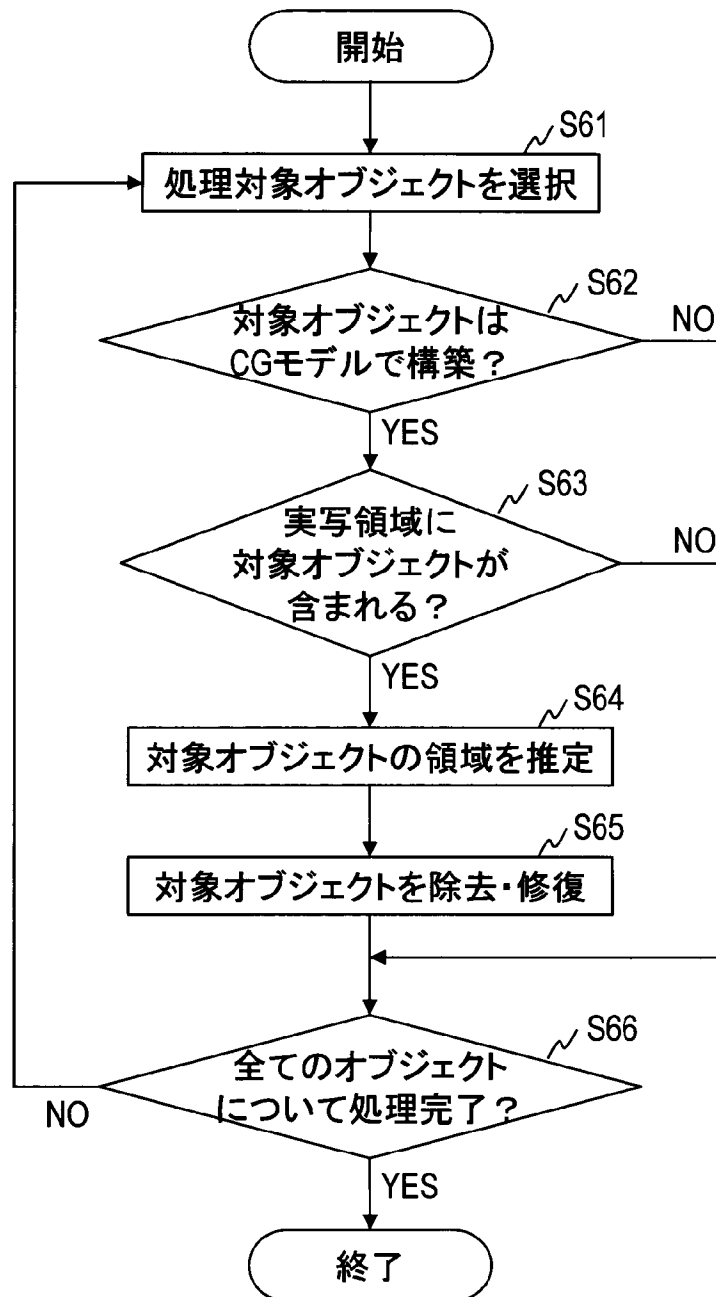


図41

[図42]

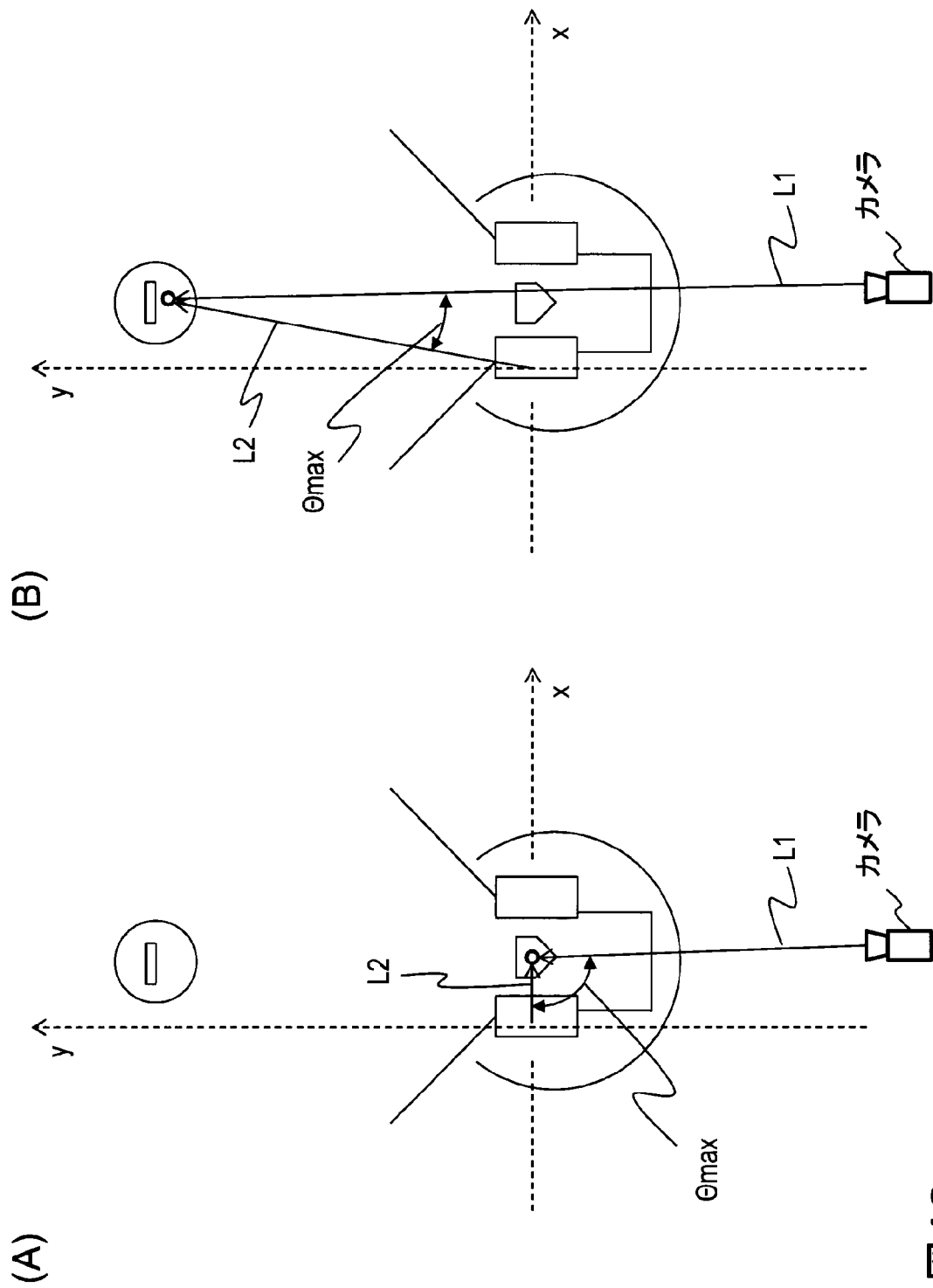


図42

[図43]

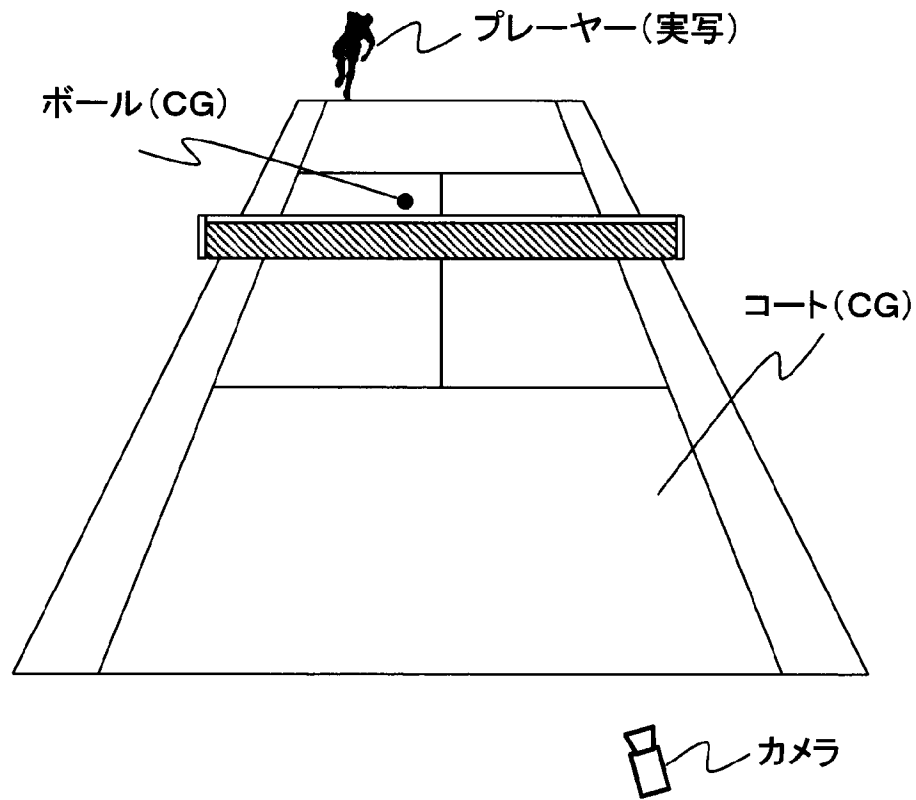


図43

[図44]

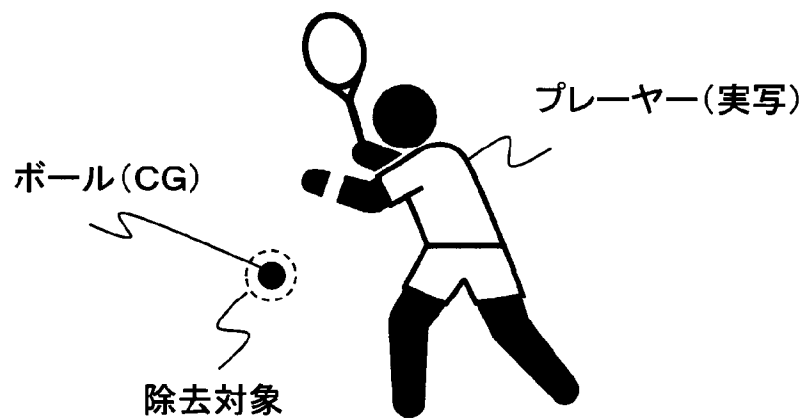
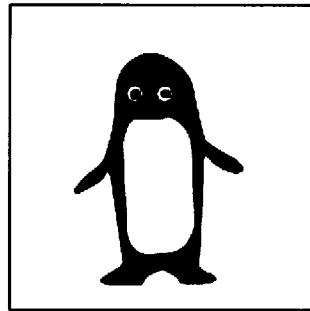


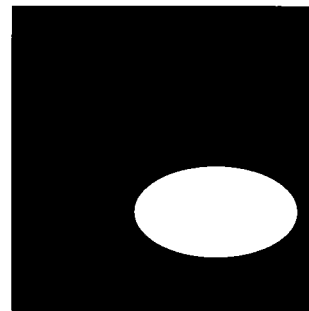
図44

[図45]

(1) 原画像・映像とマスクが別々に与えられた場合

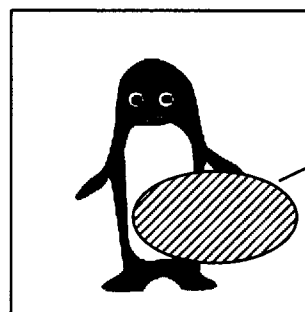


原画像・映像



マスク

(2) 原画像・映像にマスクが重畳されて与えられた場合



原画像・映像

マスク

図45

[図46]

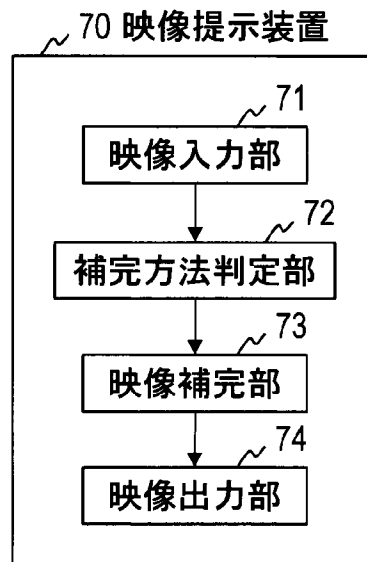


図46

[図47]

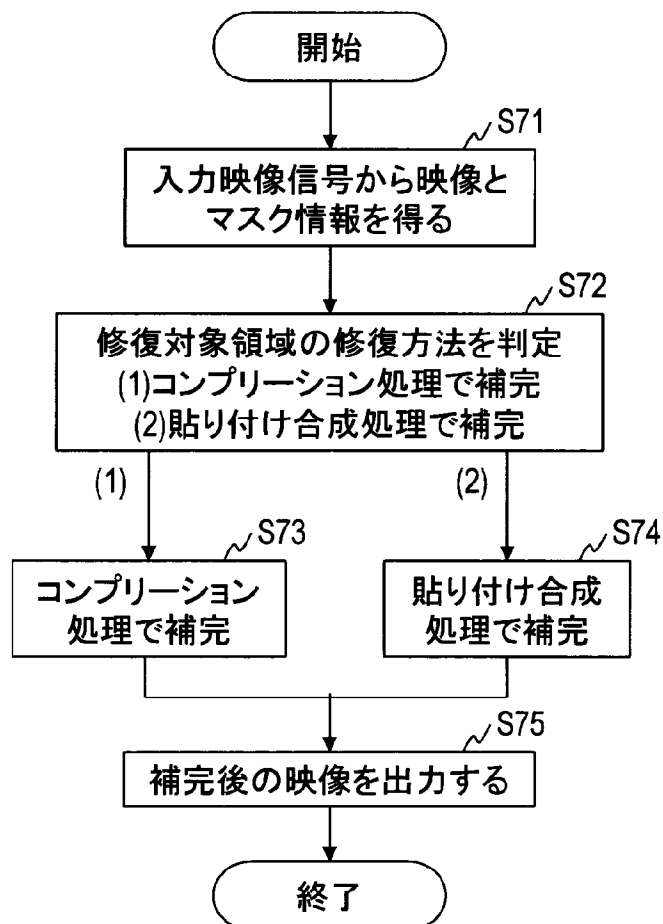
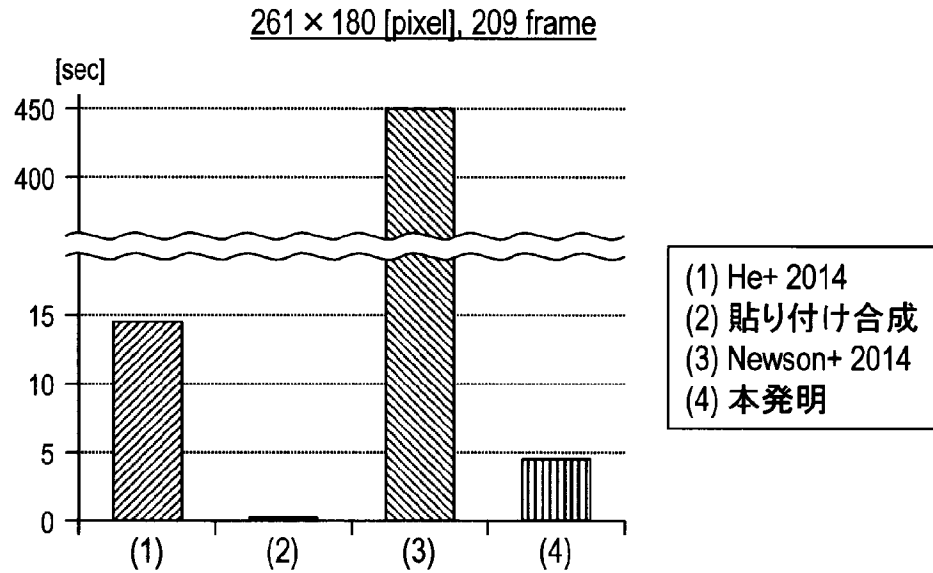


図47

[図48]

(A) 貼り付け合成が難しい映像



(B) コンプリーションが難しい映像

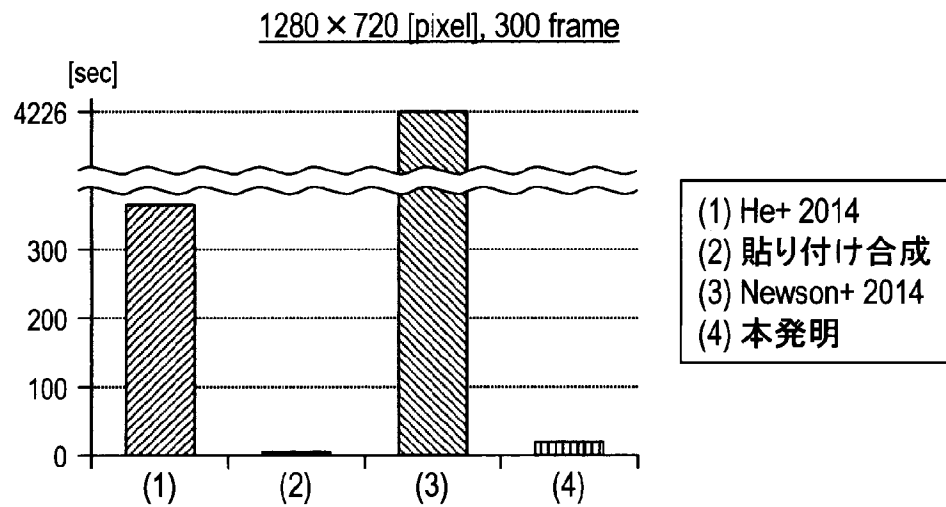


図48

[図49]

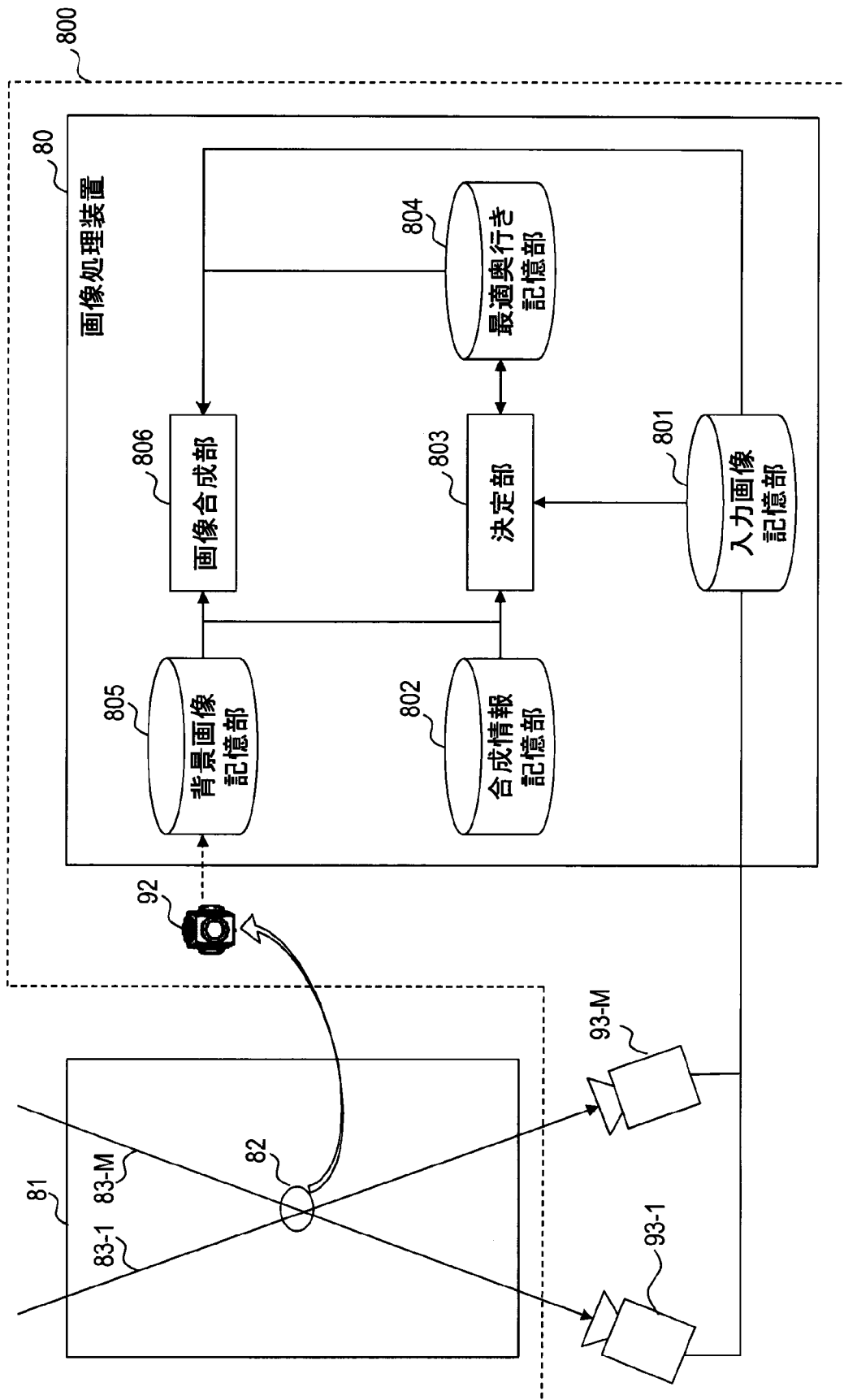


図49

[図50]

カメラID	合成情報				
	奥行き	抽出領域情報			変換情報
		左上座標	幅	高さ	
C ₁	奥行き1	(A, B)	C	D	A1 _{j1}
	・	・	・	・	・
	・	・	・	・	・
	・	・	・	・	・
C ₂	奥行きN	(A, B)	C	D	A1 _{jN}
	奥行き1	(E, F)	C	D	A2 _{j1}
	・	・	・	・	・
	・	・	・	・	・
C ₂	・	・	・	・	・
	奥行きN	(E, F)	C	D	A2 _{jN}

図50

[図51]

時刻	最適奥行き	
	C_1	C_2
t	$d1(t)$	$d2(t)$
$t+1$	$d1(t+1)$	$d2(t+1)$
$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$

図51

[図52]

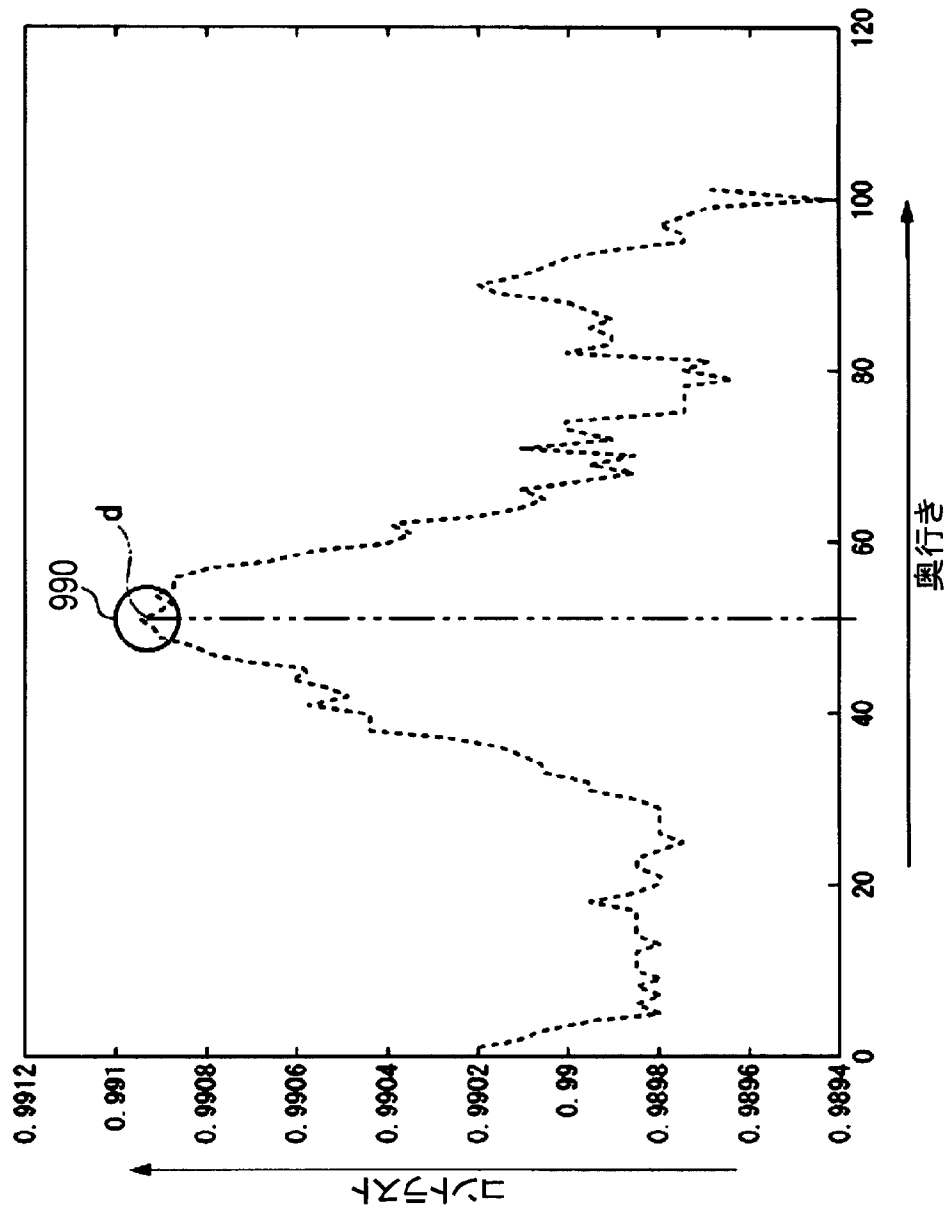


図52

[図53]

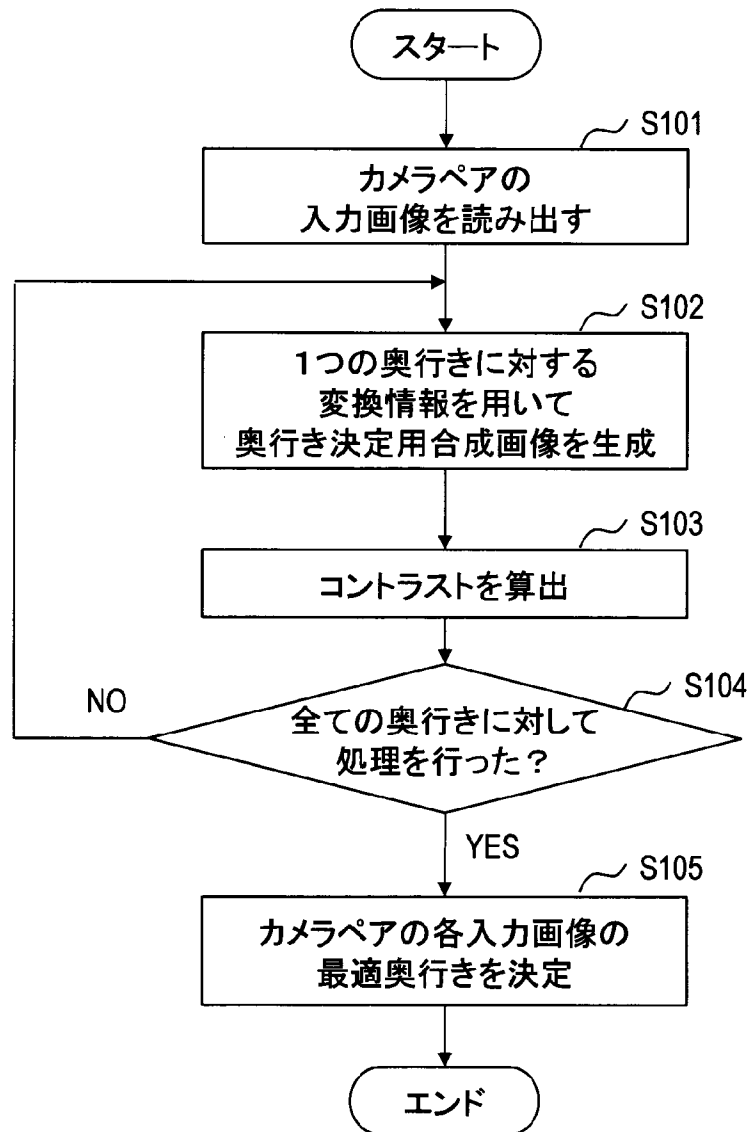


図53

[図54]

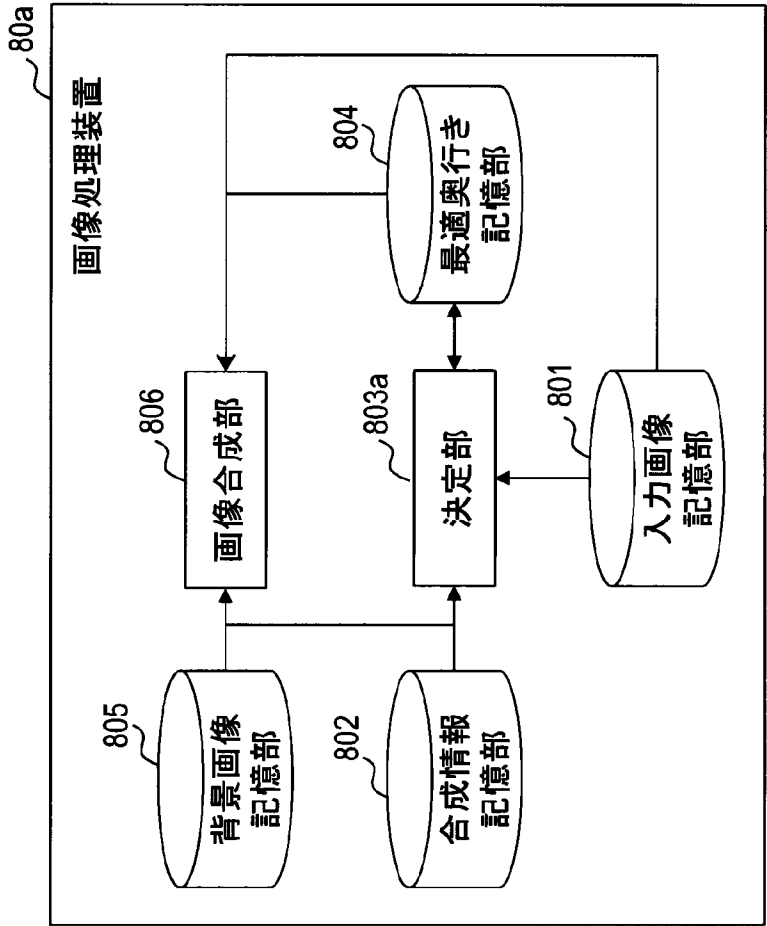


図54

[図55]

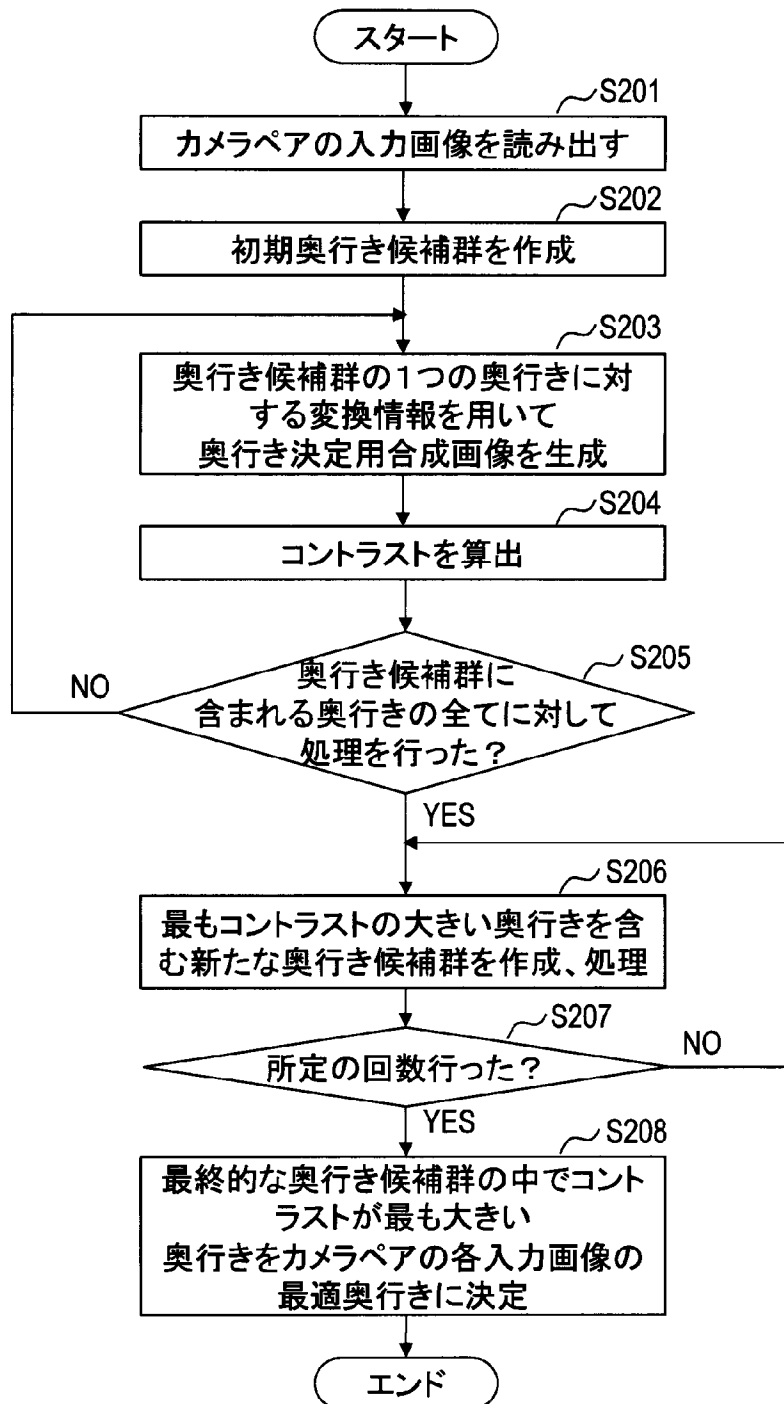


図55

[図56]

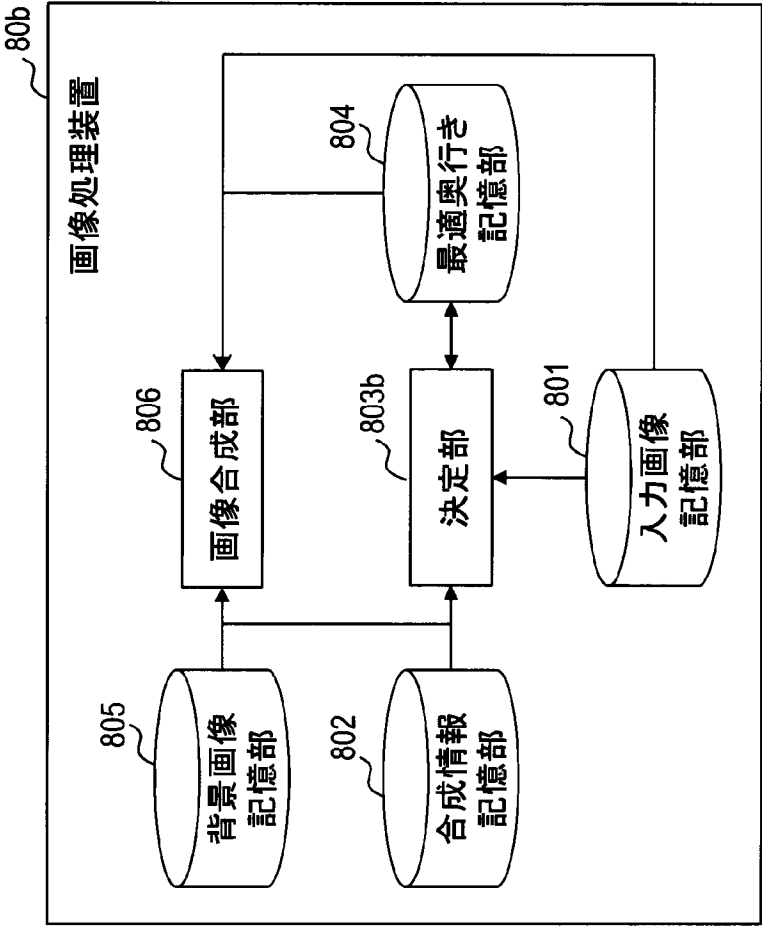


図56

[図57]

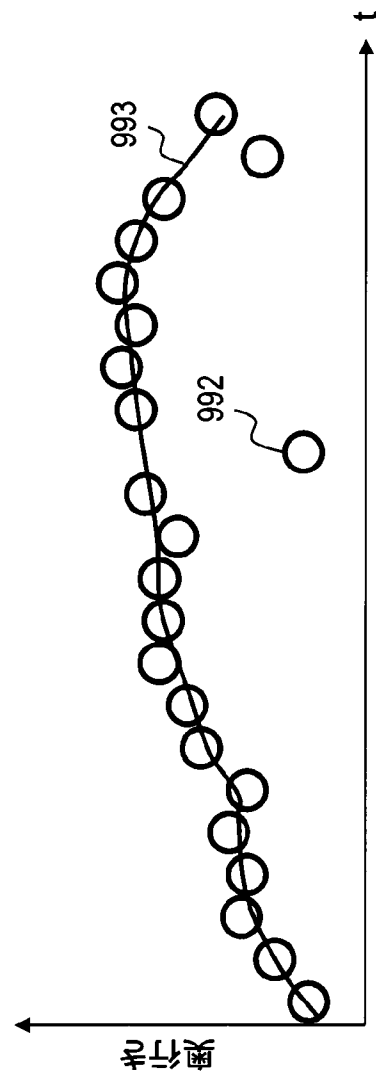


図57

[図58]

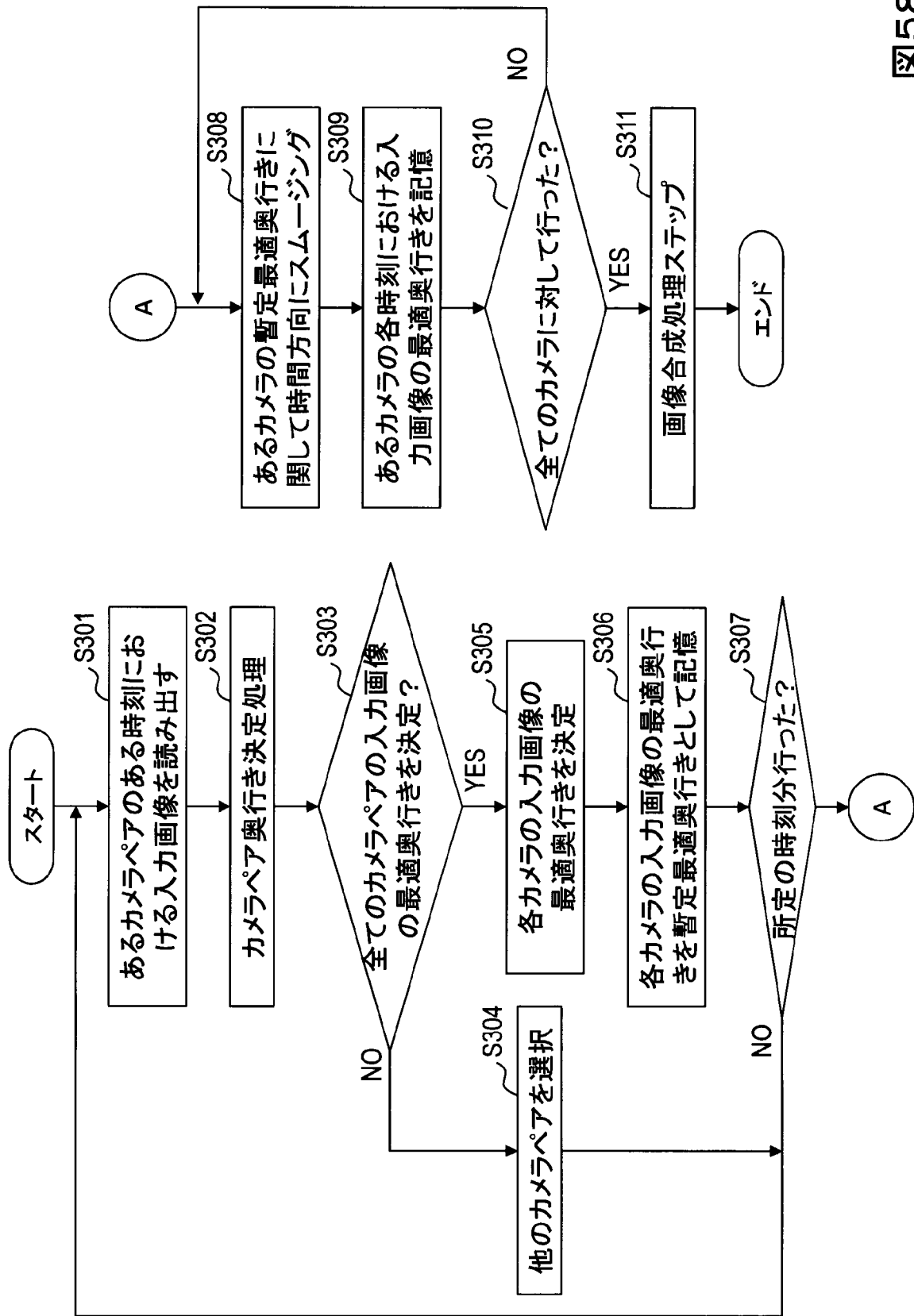


図58

[図59]

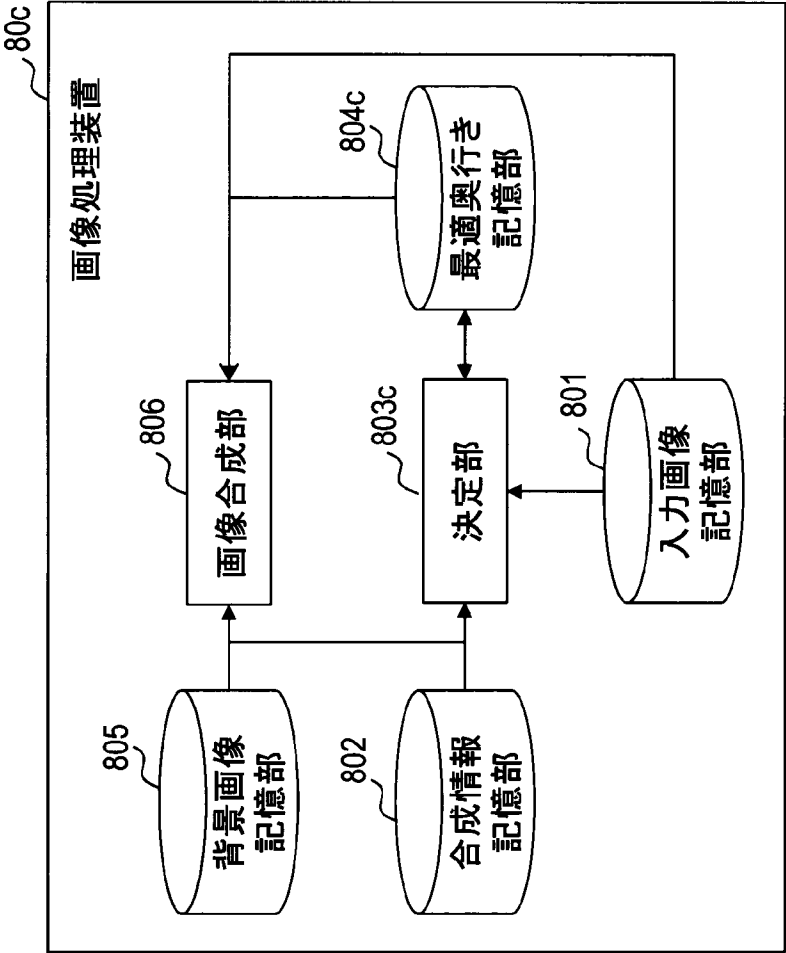


図59

[図60]

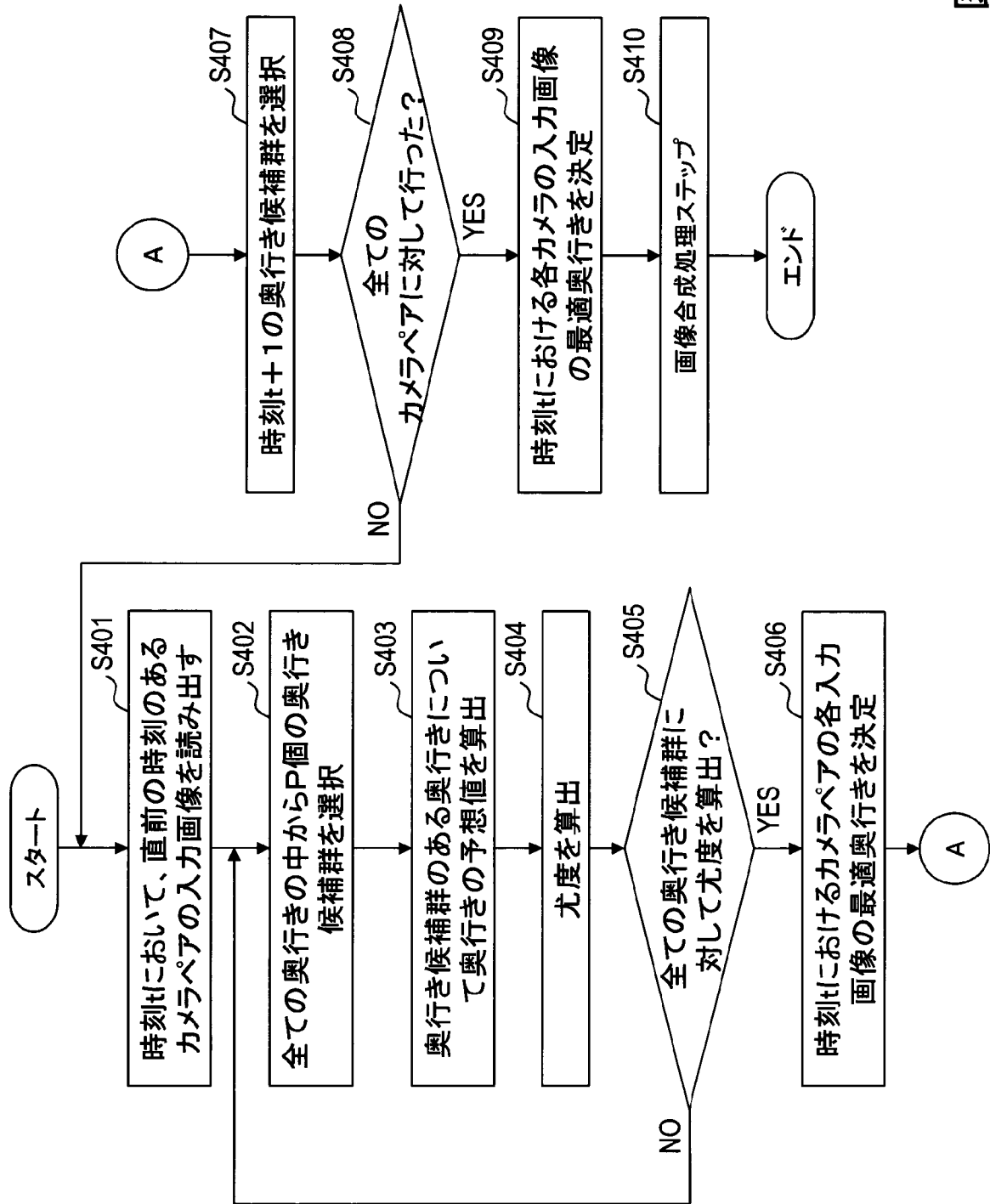


図60

[図61]

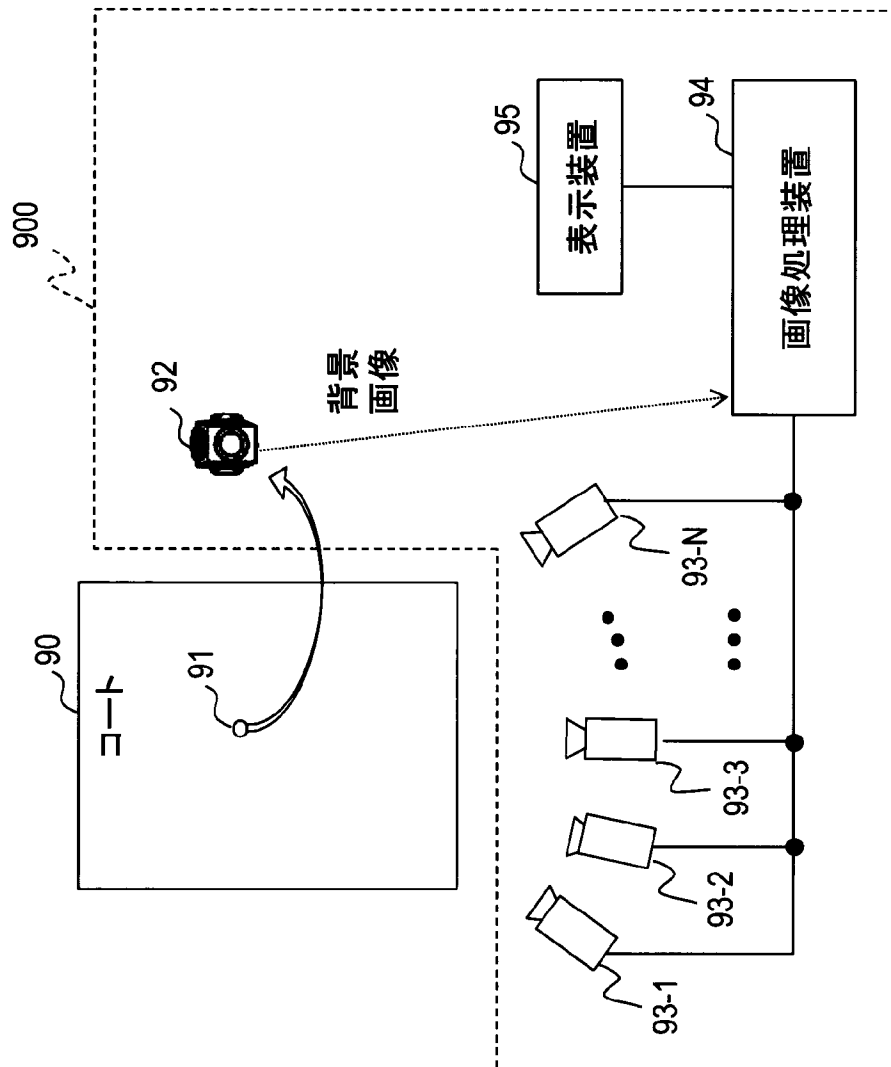


図61

[図62]

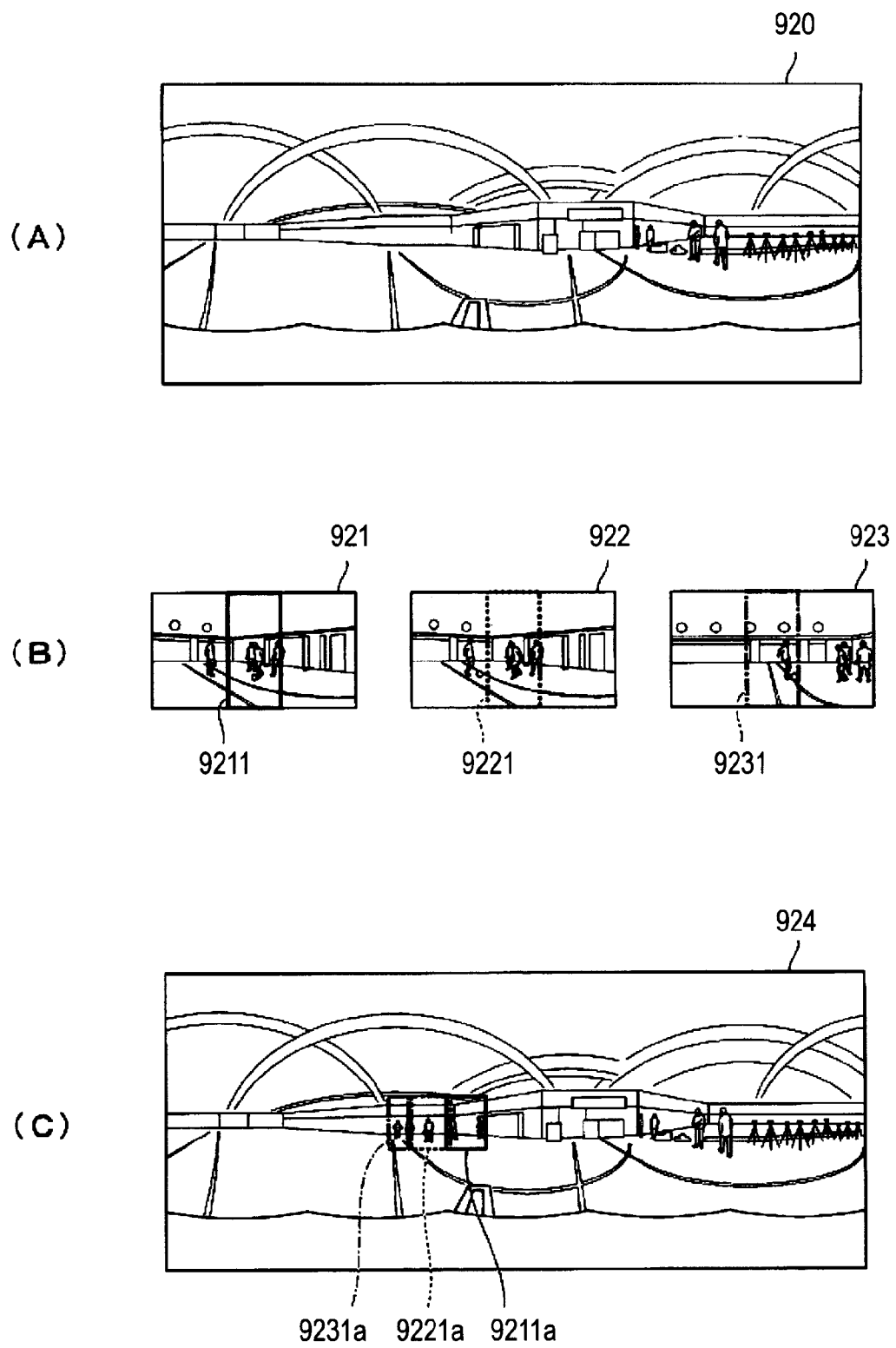


図62

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2017/002665

A. CLASSIFICATION OF SUBJECT MATTER

G06T19/00(2011.01)i, A63B69/00(2006.01)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, A63B69/00

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	1996-2017
Kokai Jitsuyo Shinan Koho	1971-2017	Toroku Jitsuyo Shinan Koho	1994-2017

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	Masaki TAKAGI et al., "Study of New Baseball Batting Practice Method using VR", Dai 19 Kai Proceedings of the Virtual Reality Society of Japan, Annual Conference, 17 September 2014 (17.09.2014), pages 210 to 213, ISSN 1349-5062	1-6, 8, 13-16, 18-19, 30-31, 48
A		7, 9-12, 17, 20-29, 32-47
Y	Takuya INOUE et al., "Kosoku Camera Eizo o Mochiita Yakyu Ball no Kaiten Sokudo Suitei System", SSII2010 Dai 16 Kai Symposium on Sensing via Image Information Koen Ronbunshu, 09 June 2010 (09.06.2010), pages (IS2-22-1) to (IS2-22-7)	1-6, 8, 13-16, 18-19, 30-31, 48
A		7, 9-12, 17, 20-29, 32-47



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

"I"

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

10 April 2017 (10.04.17)

Date of mailing of the international search report

18 April 2017 (18.04.17)

Name and mailing address of the ISA/

Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2017/002665

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Masaki TAKAGI et al., "Suburi o Kokateki ni Jisshi Dekiru Renshu Shien System no Kaihatsu", Dai 20 Kai Proceedings of the Virtual Reality Society of Japan, Annual Conference, 09 September 2015 (09.09.2015), pages 465 to 468	2-5, 14-16, 18-19
Y A	JP 2012-40071 A (Astrodesign, Inc.), 01 March 2012 (01.03.2012), paragraph [0017] (Family: none)	8, 13, 30-31 20-29
A	JP 2015-225552 A (Nippon Telegraph and Telephone Corp.), 14 December 2015 (14.12.2015), claim 1; fig. 1 (Family: none)	1-48
A	JP 2008-307091 A (Akira TOMONO), 25 December 2008 (25.12.2008), paragraphs [0201] to [0214]; fig. 14 (Family: none)	1-48
A	JP 2007-265269 A (Toshiba Corp.), 11 October 2007 (11.10.2007), paragraphs [0061] to [0065] & US 2007/0229529 A1 paragraphs [0099] to [0103]	1-48
A	JP 2001-175863 A (Nippon Hoso Kyokai), 29 June 2001 (29.06.2001), claim 1; paragraph [0026]; fig. 1 (Family: none)	1-48

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2017/002665

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

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

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

This International Searching Authority found multiple inventions in this international application, as follows:
See extra sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

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

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2017/002665

Continuation of Box No.III of continuation of first sheet(2)

(Invention 1) claims 1-6

Claims 1 to 6 are classified into Invention 1 because of having a special technical feature of "a virtual environment construction device including a virtual environment construction unit".

(Invention 2) claims 7-13

It is not considered that claims 7-13 have a special technical feature same as or corresponding to claim 1 classified into Invention 1.

Further, claims 7-13 are not dependent on claim 1.

In addition, claims 7-13 have no relationship such that these claims are substantially same as or equivalent to any claim classified into Invention 1.

Consequently, claims 7-13 cannot be classified into Invention 1.

Further, claims 7 to 13 are classified into Invention 2 because of having a special technical feature of "a virtual environment construction device including a presentation sequence generation unit".

(Invention 3) claims 14-19

It is not considered that claims 14-19 have a special technical feature same as or corresponding to claims 1, 7 classified into Inventions 1-2.

Further, claims 14-19 are not dependent on claims 1, 7.

In addition, claims 14-19 have no relationship such that these claims are substantially same as or equivalent to any claim classified into Inventions 1-2.

Consequently, claims 14-19 cannot be classified into Inventions 1-2.

Further, claims 14 to 19 are classified into Invention 3 because of having a special technical feature of "a video presentation device including a user state storage unit and a user state acquisition unit".

(Invention 4) claims 20-23, 26-27

It is not considered that claims 20-23, 26-27 have a special technical feature which is same as or corresponding to claims 1, 7, 14 classified into Inventions 1-3.

Further, claims 20-23, 26-27 are not dependent on claims 1, 7, 14.

In addition, claims 20-23, 26-27 have no relationship such that these claims are substantially same as or equivalent to any claim classified into Inventions 1-3.

Consequently, claims 20-23, 26-27 cannot be classified into Inventions 1-3.

Further, claims 20 to 23 and 26 to 27 are classified into Invention 4 because of having a special technical feature of "a virtual environment construction device including a synchronization unit".

(Invention 5) claims 24-25, 28-29

It is not considered that claims 24-25, 28-29 have a special technical feature same as or corresponding to claims 1, 7, 14, 20 classified into Inventions 1-4.

Further, claims 24-25, 28-29 are not dependent on claims 1, 7, 14, 20.

In addition, claims 24-25, 28-29 have no relationship such that these claims are substantially same as or equivalent to any claim classified into Inventions 1-4.

Consequently, claims 24-25, 28-29 cannot be classified into Inventions 1-4.

Further, claims 24 to 25 and 28 to 29 are classified into Invention 5 because of having a special technical feature of "a model learning device for learning a model outputting the time at the instant of isolation".

(Continued to next extra sheet)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2017/002665

(Invention 6) claims 30-31

It is not considered that claims 30-31 have a special technical feature same as or corresponding to claims 1, 7, 14, 20, 24 classified into Inventions 1-5.

Further, claims 30-31 are not dependent on claims 1, 7, 14, 20, 24.

In addition, claims 30-31 have no relationship such that these claims are substantially same as or equivalent to any claim classified into Inventions 1-5.

Consequently, claims 30-31 cannot be classified into Inventions 1-5.

Further, claims 30 to 31 are classified into Invention 6 because of having a special technical feature of "a virtual environment construction device including a similar video search unit".

(Invention 7) claims 32-35

It is not considered that claims 32-35 have a special technical feature same as or corresponding to claims 1, 7, 14, 20, 24, 30 classified into Inventions 1-6.

Further, claims 32-35 are not dependent on claims 1, 7, 14, 20, 24, 30.

In addition, claims 32-35 have no relationship such that these claims are substantially same as or equivalent to any claim classified into Inventions 1-6.

Consequently, claims 32-35 cannot be classified into Inventions 1-6.

Further, claims 32 to 35 are classified into Invention 7 because of having a special technical feature of "a virtual environment construction device including a model construction method decision unit".

(Invention 8) claims 36-42

It is not considered that claims 36-42 have a special technical feature same as or corresponding to claims 1, 7, 14, 20, 24, 30, 32 classified into Inventions 1-7.

Further, claims 36-42 are not dependent on claims 1, 7, 14, 20, 24, 30, 32.

In addition, claims 36-42 have no relationship such that these claims are substantially same as or equivalent to any claim classified into Inventions 1-7.

Consequently, claims 36-42 cannot be classified into Inventions 1-7.

Further, claims 36 to 42 are classified into Invention 8 because of having a special technical feature of "a video presentation device including a complementing method determination unit".

(Invention 9) claims 43-48

It is not considered that claims 43-48 have a special technical feature same as or corresponding to claims 1, 7, 14, 20, 24, 30, 32, 36 classified into Inventions 1-8.

Further, claims 43-48 are not dependent on claims 1, 7, 14, 20, 24, 30, 32, 36.

In addition, claims 43-48 have no relationship such that these claims are substantially same as or equivalent to any claim classified into Inventions 1-8.

Consequently, claims 43-48 cannot be classified into Inventions 1-8.

Further, claims 43 to 48 are classified into Invention 9 because of having a special technical feature of "an optimal depth decision device including a decision unit".

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G06T19/00(2011.01)i, A63B69/00(2006.01)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（IPC））

Int.Cl. G06T19/00, A63B69/00

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	1971-2017年
日本国実用新案登録公報	1996-2017年
日本国登録実用新案公報	1994-2017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 A	▲高▼木将樹 ほか、仮想空間を用いた野球打撃練習の効率化手法の提案, 第19回 日本バーチャルリアリティ学会大会 論文集, 2014.09.17, p.210-213, ISSN 1349-5062	1-6, 8, 13-16, 18-19, 30-31, 48 7, 9-12, 17, 20-29, 32-47

☒ 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

10.04.2017

国際調査報告の発送日

18.04.2017

国際調査機関の名称及びあて先

日本国特許庁（ISA/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 A	井上卓也 ほか, 高速カメラ映像を用いた野球ボールの回転速度推定システム, S S I I 2 0 1 0 第 1 6 回 画像センシングシンポジウム講演論文集, 2010.06.09, p. (IS2-22-1)-(IS2-22-7)	1-6, 8, 13-16, 18-19, 30-31, 48 7, 9-12, 17, 20-29, 32-47
Y	▲高▼木将樹 ほか, 素振りを効果的に実施できる練習支援システムの開発, 第 2 0 回 日本バーチャルリアリティ学会大会 論文集, 2015.09.09, p. 465-468	2-5, 14-16, 18-19
Y A	JP 2012-40071 A (アストロデザイン株式会社) 2012.03.01, 段落[0017] (ファミリーなし)	8, 13, 30-31 20-29
A	JP 2015-225552 A (日本電信電話株式会社) 2015.12.14, 請求項 1, 第 1 図 (ファミリーなし)	1-48
A	JP 2008-307091 A (伴野明) 2008.12.25, 段落[0201]-[0214], 第 14 図 (ファミリーなし)	1-48
A	JP 2007-265269 A (株式会社東芝) 2007.10.11, 段落[0061]-[0065], & US 2007/0229529 A1, 段落[0099]-[0103]	1-48
A	JP 2001-175863 A (日本放送協会) 2001.06.29, 請求項 1, 段落[0026], 第 1 図 (ファミリーなし)	1-48

第Ⅱ欄 請求の範囲の一部の調査ができないときの意見（第1ページの2の続き）

法第8条第3項（P C T 17条(2)(a)）の規定により、この国際調査報告は次の理由により請求の範囲の一部について作成しなかった。

1. ☐ 請求項 _____ は、この国際調査機関が調査をすることを要しない対象に係るものである。
つまり、
2. ☐ 請求項 _____ は、有意義な国際調査をすることができる程度まで所定の要件を満たしていない国際出願の部分に係るものである。つまり、
3. ☐ 請求項 _____ は、従属請求の範囲であってP C T 規則6.4(a)の第2文及び第3文の規定に従って記載されていない。

第Ⅲ欄 発明の単一性が欠如しているときの意見（第1ページの3の続き）

次に述べるようにこの国際出願に二以上の発明があるところの国際調査機関は認めた。

特別ページ参照

1. ☐ 出願人が必要な追加調査手数料をすべて期間内に納付したので、この国際調査報告は、すべての調査可能な請求項について作成した。
2. ☒ 追加調査手数料を要求するまでもなく、すべての調査可能な請求項について調査することができたので、追加調査手数料の納付を求めなかった。
3. ☐ 出願人が必要な追加調査手数料を一部のみしか期間内に納付しなかったため、この国際調査報告は、手数料の納付のあった次の請求項のみについて作成した。
4. ☐ 出願人が必要な追加調査手数料を期間内に納付しなかったため、この国際調査報告は、請求の範囲の最初に記載されている発明に係る次の請求項について作成した。

追加調査手数料の異議の申立てに関する注意

- ☐ 追加調査手数料及び、該当する場合には、異議申立手数料の納付と共に、出願人から異議申立てがあった。
- ☐ 追加調査手数料の納付と共に出願人から異議申立てがあったが、異議申立手数料が納付命令書に示した期間内に支払われなかった。
- ☐ 追加調査手数料の納付はあったが、異議申立てはなかった。

< 第 III 欄の続き >

(発明 1) 請求項 1-6

請求項 1-6 は、「仮想環境構築部を含む仮想環境構築装置」という特別な技術的特徴を有しているので、発明 1 に区分する。

(発明 2) 請求項 7-13

請求項 7-13 は、発明 1 に区分された請求項 1 と、同一の又は対応する特別な技術的特徴を有しているとはいえない。

また、請求項 7-13 は請求項 1 の従属請求項ではない。さらに、請求項 7-13 は、発明 1 に区分されたいずれの請求項に対しても実質同一又はそれに準ずる関係にはない。

したがって、請求項 7-13 は発明 1 に区分できない。

そして、請求項 7-13 は、「提示シーケンス生成部を含む仮想環境構築装置」という特別な技術的特徴を有しているので、発明 2 に区分する。

(発明 3) 請求項 14-19

請求項 14-19 は、発明 1-2 に区分された請求項 1,7 と、同一の又は対応する特別な技術的特徴を有しているとはいえない。

また、請求項 14-19 は請求項 1,7 の従属請求項ではない。さらに、請求項 14-19 は、発明 1-2 に区分されたいずれの請求項に対しても実質同一又はそれに準ずる関係にはない。

したがって、請求項 14-19 は発明 1-2 に区分できない。

そして、請求項 14-19 は、「ユーザ状態記憶部及びユーザ状態取得部を含む映像提示装置」という特別な技術的特徴を有しているので、発明 3 に区分する。

(発明 4) 請求項 20-23,26-27

請求項 20-23,26-27 は、発明 1-3 に区分された請求項 1,7,14 と、同一の又は対応する特別な技術的特徴を有しているとはいえない。

また、請求項 20-23,26-27 は請求項 1,7,14 の従属請求項ではない。さらに、請求項 20-23,26-27 は、発明 1-3 に区分されたいずれの請求項に対しても実質同一又はそれに準ずる関係にはない。

したがって、請求項 20-23,26-27 は発明 1-3 に区分できない。

そして、請求項 20-23,26-27 は、「同期部を含む仮想環境構築装置」という特別な技術的特徴を有しているので、発明 4 に区分する。

(発明 5) 請求項 24-25,28-29

請求項 24-25,28-29 は、発明 1-4 に区分された請求項 1,7,14,20 と、同一の又は対応する特別な技術的特徴を有しているとはいえない。

また、請求項 24-25,28-29 は請求項 1,7,14,20 の従属請求項ではない。さらに、請求項 24-25,28-29 は、発明 1-4 に区分されたいずれの請求項に対しても実質同一又はそれに準ずる関係にはない。

したがって、請求項 24-25,28-29 は発明 1-4 に区分できない。

そして、請求項 24-25,28-29 は、「隔離する瞬間の時刻を出力するモデルを学習する、モデル学習装置」という特別な技術的特徴を有しているので、発明 5 に区分する。

特別ページに続く

<第 III 欄の続き>

(発明 6) 請求項 30-31

請求項 30-31 は、発明 1-5 に区分された請求項 1,7,14,20,24 と、同一の又は対応する特別な技術的特徴を有しているとはいえない。

また、請求項 30-31 は請求項 1,7,14,20,24 の従属請求項ではない。さらに、請求項 30-31 は、発明 1-5 に区分されたいずれの請求項に対しても実質同一又はそれに準ずる関係にはない。

したがって、請求項 30-31 は発明 1-5 に区分できない。

そして、請求項 30-31 は、「類似映像検索部を含む仮想環境構築装置」という特別な技術的特徴を有しているので、発明 6 に区分する。

(発明 7) 請求項 32-35

請求項 32-35 は、発明 1-6 に区分された請求項 1,7,14,20,24,30 と、同一の又は対応する特別な技術的特徴を有しているとはいえない。

また、請求項 32-35 は請求項 1,7,14,20,24,30 の従属請求項ではない。さらに、請求項 32-35 は、発明 1-6 に区分されたいずれの請求項に対しても実質同一又はそれに準ずる関係にはない。

したがって、請求項 32-35 は発明 1-6 に区分できない。

そして、請求項 32-35 は、「モデル構築方法決定部を含む仮想環境構築装置」という特別な技術的特徴を有しているので、発明 7 に区分する。

(発明 8) 請求項 36-42

請求項 36-42 は、発明 1-7 に区分された請求項 1,7,14,20,24,30,32 と、同一の又は対応する特別な技術的特徴を有しているとはいえない。

また、請求項 36-42 は請求項 1,7,14,20,24,30,32 の従属請求項ではない。さらに、請求項 36-42 は、発明 1-7 に区分されたいずれの請求項に対しても実質同一又はそれに準ずる関係にはない。

したがって、請求項 36-42 は発明 1-7 に区分できない。

そして、請求項 36-42 は、「補完方法判定部を含む映像提示装置」という特別な技術的特徴を有しているので、発明 8 に区分する。

(発明 9) 請求項 43-48

請求項 43-48 は、発明 1-8 に区分された請求項 1,7,14,20,24,30,32,36 と、同一の又は対応する特別な技術的特徴を有しているとはいえない。

また、請求項 43-48 は請求項 1,7,14,20,24,30,32,36 の従属請求項ではない。さらに、請求項 43-48 は、発明 1-8 に区分されたいずれの請求項に対しても実質同一又はそれに準ずる関係にはない。

したがって、請求項 43-48 は発明 1-8 に区分できない。

そして、請求項 43-48 は、「決定部を含む最適奥行き決定装置」という特別な技術的特徴を有しているので、発明 9 に区分する。