

(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  
2022年10月6日(06.10.2022)



(10) 国際公開番号  
**WO 2022/209362 A1**

- (51) 国際特許分類:  
G06T 15/20 (2011.01)
- (21) 国際出願番号: PCT/JP2022/005748
- (22) 国際出願日: 2022年2月14日(14.02.2022)
- (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 2021-061676 2021年3月31日(31.03.2021) JP
- (71) 出願人: 富士フイルム株式会社 (FUJIFILM CORPORATION) [JP/JP]; 〒1068620 東京都港区西麻布2丁目2番30号 Tokyo (JP).
- (72) 発明者: 宮田 真彦 (MIYATA, Masahiko); 〒3319624 埼玉県さいたま市北区植竹町1丁目324番地 富士フイルム株式会社内 Saitama (JP). 青木 貴嗣 (AOKI, Takashi); 〒3319624 埼玉県さいたま市北区植竹町1丁目324番

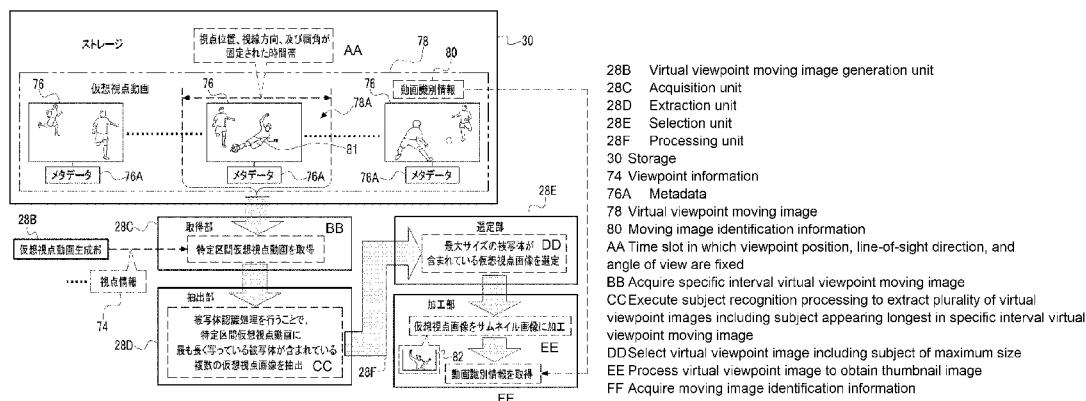
地 富士フイルム株式会社内 Saitama (JP). 林健吉 (HAYASHI, Kenkichi); 〒3319624 埼玉県さいたま市北区植竹町1丁目324番地 富士フイルム株式会社内 Saitama (JP). 村上泰規 (MURAKAMI, Yasunori); 〒3319624 埼玉県さいたま市北区植竹町1丁目324番地 富士フイルム株式会社内 Saitama (JP).

(74) 代理人: 特許業務法人太陽国際特許事務所 (TAIYO, NAKAJIMA & KATO); 〒1600022 東京都新宿区新宿4丁目3番17号 Tokyo (JP).

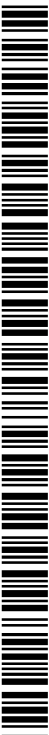
(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ,

(54) Title: IMAGE PROCESSING DEVICE, IMAGE PROCESSING METHOD, AND PROGRAM

(54) 発明の名称: 画像処理装置、画像処理方法、及びプログラム



(57) Abstract: This image processing device comprises a processor, and a memory that is connected to or embedded in the processor. The processor uses a plurality of pieces of viewpoint information and a plurality of photographed images obtained by photographing a photography region to acquire a representative image corresponding to a virtual viewpoint moving image generated on the basis of the plurality of pieces of viewpoint information and the plurality of photographed images, and outputs data for displaying, on a display, the representative image of a size different from the size of the virtual viewpoint moving image.



WO 2022/209362 A1

NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,  
QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,  
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,  
US, UZ, VC, VN, WS, 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))
- 

(57) 要約 : 画像処理装置は、プロセッサと、プロセッサに接続又は内蔵されたメモリと、を備える。プロセッサは、撮像領域が撮像されることで得られる複数の撮像画像と複数の視点情報とに基づいて生成される仮想視点動画に対応する代表画像を、複数の撮像画像及び複数の視点情報に基づいて取得し、仮想視点動画とは異なる大きさを代表画像をディスプレイに表示させるためのデータを出力する。

## 明 細 書

発明の名称：画像処理装置、画像処理方法、及びプログラム

### 技術分野

[0001] 本開示の技術は、画像処理装置、画像処理方法、及びプログラムに関する。

### 背景技術

[0002] 特開2018-046448号公報には、複数のカメラを用いて撮影した多視点映像から、仮想カメラから見た映像である自由視点映像を生成する画像処理装置において、仮想カメラの移動の軌跡を示すカメラパス及び仮想カメラが注視する先である注視点の移動の軌跡を示す注視点パスを、ユーザが指定するためのユーザインタフェースと、ユーザインタフェースを介して指定されたカメラパス及び注視点パスに基づいて、自由視点映像を生成する生成手段と、を備え、ユーザインタフェースは、多視点映像の撮影シーンを俯瞰的に捉えた2次元画像を用いたUI画面に、多視点映像のうち自由視点映像の生成対象となるタイムフレームにおける被写体の時系列の変化を表示し、ユーザが、2次元画像に対する入力操作によって軌跡を描くことで、カメラパス及び注視点パスが指定されるように構成されることを特徴とする画像処理装置が開示されている。また、特開2018-046448号公報に記載の画像処理装置では、2次元画像は静止画であり、ユーザインタフェースは、タイムフレームを一定の間隔でサンプリングした所定のフレームにおける各被写体を時間軸方向に異なる態様で静止画に重ねて表示することで、被写体の時系列の変化を表示するように構成される。また、特開2018-046448号公報に記載の画像処理装置において、ユーザインタフェースは、ユーザが指定したカメラパスに沿って、時間軸方向に一定の間隔で仮想カメラから見た場合のサムネイル画像を配置し、サムネイル画像に対するユーザの入力操作を介して、仮想カメラの経路、高度、移動速度が調整されるように構成される。

[0003] 特開 2017-212592 号公報には、被写体を複数の方向から撮影するための複数のカメラによる撮影に基づく画像データに基づいて、画像生成装置により仮想視点画像が生成されるシステムにおいて、仮想視点画像の生成に係る視点を指定するためのユーザによる指示を受け付ける受け付け手段と、受け付け手段により受け付けられる指示に基づく視点の指定が制限される制限領域であって、システムに含まれる装置の動作状態及び画像データに関するパラメータの少なくとも何れかに応じて変化する制限領域を特定するための情報を取得する取得手段と、取得手段が取得した情報に基づいて、制限領域に応じた表示制御に基づく画像を表示部に表示させる表示制御手段と、を有することを特徴とする制御装置が開示されている。

[0004] 特開 2014-126906 号公報には、自由視点再生処理において、動画像の再生開始前に、ユーザにより選択された何れか一の撮像装置の表示制御部が、複数の撮像装置により撮像された動画像に対応するサムネイル画像を一覧表示し、この中でユーザにより選択されたものから再生を開始するようにしても良いことが記載されている。

### 発明の概要

[0005] 本開示の技術に係る一つの実施形態は、観賞者に対して仮想視点動画に対応する代表画像を見せることができる画像処理装置、画像処理方法、及びプログラムを提供する。

### 課題を解決するための手段

[0006] 本開示の技術に係る第 1 の態様は、プロセッサと、プロセッサに接続又は内蔵されたメモリと、を備え、プロセッサが、撮像領域が撮像されることで得られる複数の撮像画像と複数の視点情報とに基づいて生成される仮想視点動画に対応する代表画像を、複数の撮像画像及び複数の視点情報に基づいて取得し、仮想視点動画とは異なる大きさに代表画像をディスプレイに表示させるためのデータを出力する画像処理装置である。

[0007] 本開示の技術に係る第 2 の態様は、代表画像が、仮想視点動画において撮像領域内の第 1 被写体が含まれている複数のフレームのうちの第 1 フレーム

に関する画像である、第1の態様に係る画像処理装置である。

[0008] 本開示の技術に係る第3の態様は、第1被写体が、仮想視点動画に含まれる時間に基づいて決められた被写体である、第2の態様に係る画像処理装置である。

[0009] 本開示の技術に係る第4の態様は、第1フレームが、仮想視点動画内での第1被写体のサイズに基づいて決められたフレームである、第2の態様又は第3の態様に係る画像処理装置である。

[0010] 本開示の技術に係る第5の態様は、プロセッサが、複数の視点情報の編集結果に基づいて代表画像を取得する、第1の態様から第4の態様の何れか1つの態様に係る画像処理装置である。

[0011] 本開示の技術に係る第6の態様は、複数の視点情報が、複数の視点パスを有し、編集結果が、複数の視点パスに対して行われた編集の結果を含む、第5の態様に係る画像処理装置である。

[0012] 本開示の技術に係る第7の態様は、プロセッサが、複数の視点情報間の相違度に基づいて代表画像を取得する、第1の態様から第6の態様の何れか1つの態様に係る画像処理装置である。

[0013] 本開示の技術に係る第8の態様は、複数の視点情報が、複数の視点パスを有し、相違度が、複数の視点パス間の相違度である、第7の態様に係る画像処理装置である。

[0014] 本開示の技術に係る第9の態様は、複数の視点情報が、複数の視点パスを有し、プロセッサが、複数の視点パスの位置関係に基づいて代表画像を取得する、第1の態様から第8の態様の何れか1つの態様に係る画像処理装置である。

[0015] 本開示の技術に係る第10の態様は、位置関係が、撮像領域内の第2被写体に対する複数の視点パスの位置関係である、第9の態様に係る画像処理装置である。

[0016] 本開示の技術に係る第11の態様は、プロセッサが、複数の仮想視点動画から、与えられた検索条件に適合する検索条件適合仮想視点動画を検索し、

検索条件適合仮想視点動画に基づいて代表画像を取得する、第1の態様から第10の態様の何れか1つの態様に係る画像処理装置である。

[0017] 本開示の技術に係る第12の態様は、代表画像が、撮像領域内の第3被写体の状況に応じて決められた画像である、第1の態様から第11の態様の何れか1つの態様に係る画像処理装置である。

[0018] 本開示の技術に係る第13の態様は、代表画像が、仮想視点動画に関わる人物の属性に応じて決められた画像である、第1の態様から第12の態様の何れか1つの態様に係る画像処理装置である。

[0019] 本開示の技術に係る第14の態様は、代表画像が、仮想視点動画の内容を示す画像である、第1の態様から第13の態様の何れか1つの態様に係る画像処理装置である。

[0020] 本開示の技術に係る第15の態様は、複数の視点情報が、視点が異なる第1視点情報及び第2視点情報を含み、第1視点情報及び第2視点情報が、異なる時刻に関する情報を有する、第1の態様から第14の態様の何れか1つの態様に係る画像処理装置である。

[0021] 本開示の技術に係る第16の態様は、プロセッサが、代表画像を第1ディスプレイに対して表示させるための第1データを出力し、第1ディスプレイに表示されている代表画像が選択されたことに応じて代表画像に対応する仮想視点動画を第1ディスプレイ及び第2ディスプレイのうちの少なくとも一方に対して表示させるための第2データを出力する、第1の態様から第15の態様の何れか1つの態様に係る画像処理装置である。

[0022] 本開示の技術に係る第17の態様は、プロセッサが、代表画像と仮想視点動画とを関連付けた状態でメモリに記憶させる、第1の態様から第16の態様の何れか1つの態様に係る画像処理装置である。

[0023] 本開示の技術に係る第18の態様は、プロセッサと、プロセッサに接続又は内蔵されたメモリと、を備え、プロセッサが、撮像領域が撮像されることで得られる複数の撮像画像と複数の視点情報とに基づいて生成される仮想視点動画に対応する代表画像を、複数の撮像画像及び複数の視点情報に基づい

て取得し、複数の画像が表示される画面に代表画像を表示させるためのデータを出力する画像処理装置である。

[0024] 本開示の技術に係る第19の態様は、撮像領域が撮像されることで得られる複数の撮像画像と複数の視点情報とに基づいて生成される仮想視点動画に対応する代表画像を、複数の撮像画像及び複数の視点情報に基づいて取得すること、並びに、仮想視点動画とは異なる大きさを代表画像をディスプレイに表示させるためのデータを出力することを含む画像処理方法である。

[0025] 本開示の技術に係る第20の態様は、コンピュータに、撮像領域が撮像されることで得られる複数の撮像画像と複数の視点情報とに基づいて生成される仮想視点動画に対応する代表画像を、複数の撮像画像及び複数の視点情報に基づいて取得すること、並びに、仮想視点動画とは異なる大きさを代表画像をディスプレイに表示させるためのデータを出力することを含む処理を実行させるためのプログラムである。

### 図面の簡単な説明

[0026] [図1]画像処理システムの構成の一例を示す概念図である。

[図2]ユーザデバイスの電気系のハードウェア構成の一例を示すブロック図である。

[図3]画像処理装置のCPUの要部機能の一例を示すブロック図である。

[図4]受付画面生成部の処理内容の一例、及び、ユーザデバイスのディスプレイの表示内容の一例を示す概念図である。

[図5]ユーザデバイスの動作モードが視点設定モードである場合の受付画面の表示態様の一例を示す画面図である。

[図6]ユーザデバイスの動作モードが注視点設定モードである場合の受付画面の表示態様の一例を示す画面図である。

[図7]視点情報の内容の一例、及び、視点情報がユーザデバイスから画像処理装置に送信される態様の一例を示すブロック図である。

[図8]仮想視点動画生成部の処理内容の一例を示す概念図である。

[図9]取得部、抽出部、選定部、及び加工部の処理内容の一例を示す概念図で

ある。

[図10]加工部及び一覧画面生成部の処理内容の一例を示す概念図である。

[図11]画面生成処理の流れの一例を示すフローチャートである。

[図12]画像処理装置のCPUの要部機能の一例を示すブロック図である。

[図13]視点パスが編集される態様の一例を示す概念図である。

[図14]視点情報の内容の一例、及び、視点情報がユーザデバイスから画像処理装置に送信される態様の一例を示すブロック図である。

[図15]仮想視点動画生成部の処理内容の一例を示す概念図である。

[図16]編集結果処理部の処理内容の一例を示す概念図である。

[図17]取得部、抽出部、選定部、及び加工部の処理内容の一例を示す概念図である。

[図18]加工部及び一覧画面生成部の処理内容の一例を示す概念図である。

[図19]画像処理装置のCPUの要部機能の一例を示すブロック図である。

[図20]第1視点パス及び第2視点パスがユーザによって指定される態様の一例を示す概念図である。

[図21]第1視点パス情報の内容及び第2視点パス情報の内容の一例を示す概念図である。

[図22]第1視点パス情報及び第2視点パス情報がユーザデバイスから画像処理装置に送信される態様の一例を示すブロック図である。

[図23]仮想視点動画生成部の処理内容の一例を示す概念図である。

[図24]ストレージに第1仮想視点動画及び第2仮想視点動画が記憶される態様の一例を示す概念図である。

[図25]相違度算出部の処理内容の一例を示す概念図である。

[図26]第1仮想視点動画が取得部、抽出部、選定部、及び加工部によって処理される態様の一例を示す概念図である。

[図27]第2仮想視点動画が取得部、抽出部、選定部、及び加工部によって処理される態様の一例を示す概念図である。

[図28]画像処理装置のCPUの要部機能の一例を示すブロック図である。

[図29]第1視点パス情報及び第2視点パス情報がユーザデバイスから画像処理装置に送信される態様の一例を示すブロック図である。

[図30]被写体位置特定部の処理内容の一例を示す概念図である。

[図31]視点位置特定部の処理内容の一例を示す概念図である。

[図32]第1仮想視点動画が取得部及び加工部によって処理される態様の一例を示す概念図である。

[図33]第2仮想視点動画が取得部及び加工部によって処理される態様の一例を示す概念図である。

[図34]画像処理装置のCPUの要部機能の一例を示すブロック図である。

[図35]検索条件付与部及び取得部の処理内容の一例を示す概念図である。

[図36]画像処理装置のCPUの要部機能の一例を示すブロック図である。

[図37]状況認識部及び取得部の処理内容の一例を示す概念図である。

[図38]画像処理装置のCPUの要部機能の一例を示すブロック図である。

[図39]人物属性被写体認識部及び取得部の処理内容の一例を示す概念図である。

[図40]第1視点パス情報の内容及び第2視点パス情報の内容の一例を示す概念図である。

[図41]記憶媒体に記憶されている画面生成処理プログラムが画像処理装置のコンピュータにインストールされる態様の一例を示す概念図である。

### 発明を実施するための形態

[0027] 添付図面に従って本開示の技術に係る画像処理装置、画像処理方法、及びプログラムの実施形態の一例について説明する。

[0028] 先ず、以下の説明で使用される文言について説明する。

[0029] CPUとは、“Central Processing Unit”の略称を指す。GPUとは、“Graphics Processing Unit”の略称を指す。TPUとは、“Tensor processing unit”の略称を指す。RAMとは、“Random Access Memory”の略称を指す。SSDとは、“Solid State Drive”の略称を指す。HDDとは、“Hard Disk Drive”の略称を指す。EEPROMとは、“Electrically Erasable

and Programmable Read Only Memory”の略称を指す。I/Fとは、“Interface”の略称を指す。ASICとは、“Application Specific Integrated Circuit”の略称を指す。PLDとは、“Programmable Logic Device”の略称を指す。FPGAとは、“Field-Programmable Gate Array”の略称を指す。SoCとは、“System-on-a-chip”の略称を指す。CMOSとは、“Complementary Metal Oxide Semiconductor”の略称を指す。CCDとは、“Charge Coupled Device”の略称を指す。ELとは、“Electro-Luminescence”の略称を指す。LANとは、“Local Area Network”の略称を指す。USBとは、“Universal Serial Bus”の略称を指す。“HMD”とは、“Head Mounted Display”の略称を指す。LTEとは、“Long Term Evolution”の略称を指す。5Gとは、“5th generation (wireless technology for digital cellular networks)”の略称を指す。TDMとは“Time-Division Multiplexing”の略称を指す。AIとは、“Artificial Intelligence”の略称を指す。また、本明細書において、画像（静止画像及び動画像を含む意味合いでの画像）に含まれる被写体とは、画像に像（例えば、電子像）として含まれる被写体を指す。

[0030] [第1実施形態]

一例として図1に示すように、画像処理システム2は、画像処理装置10及びユーザデバイス12を備えている。

[0031] 本第1実施形態では、画像処理装置10の一例として、サーバが適用されている。サーバは、例えば、メインフレームによって実現される。但し、これはあくまでも一例に過ぎず、例えば、サーバは、クラウドコンピューティング、フォグコンピューティング、エッジコンピューティング、又はグリッドコンピューティング等のネットワークコンピューティングによって実現されてもよい。また、画像処理装置10は、複数のサーバであってもよいし、ワークステーションであってもよいし、パーソナルコンピュータであってもよいし、少なくとも1つのワークステーションと少なくとも1つのパーソナルコンピュータとを組み合わせた装置であってもよいし、少なくとも1つの

ワークステーションと少なくとも1つのパーソナルコンピュータと少なくとも1つのサーバとを組み合わせた装置等であってもよい。

[0032] また、本第1実施形態では、ユーザデバイス12の一例として、スマートフォンが適用されている。但し、スマートフォンは、あくまでも一例に過ぎず、例えば、パーソナルコンピュータであってもよいし、タブレット端末又はHMD等の携帯型の多機能端末であってもよい。

[0033] また、本第1実施形態において、画像処理装置10及びユーザデバイス12は、例えば、基地局（図示省略）を介して通信可能に接続されている。基地局で使用する通信規格には、5G規格及び／又はLTE規格等を含む無線通信規格と、WiFi（802.11）規格及び／又はBluetooth（登録商標）規格を含む無線通信規格と、TDM規格及び／又はイーサネット（登録商標）規格を含む有線通信規格とが含まれる。

[0034] 画像処理装置10は、画像を取得し、取得した画像をユーザデバイス12に送信する。ここで、画像とは、例えば、撮像されることで得られた撮像画像64（図4参照）、及び撮像画像64（図4等参照）に基づいて生成された画像等を指す。撮像画像（図4参照）に基づいて生成された画像の一例としては、仮想視点画像76（図8等参照）が挙げられる。

[0035] ユーザデバイス12は、ユーザ14によって使用される。ユーザデバイス12は、タッチパネルディスプレイ16を備えている。タッチパネルディスプレイ16は、ディスプレイ18及びタッチパネル20によって実現される。ディスプレイ18の一例としては、ELディスプレイ（例えば、有機ELディスプレイ又は無機ELディスプレイ）が挙げられる。なお、ELディスプレイに限らず、液晶ディスプレイ等の他の種類のディスプレイであってもよい。

[0036] タッチパネルディスプレイ16は、ディスプレイ18の表示領域に対してタッチパネル20を重ね合わせることによって、又は、ディスプレイ18の内部にタッチパネル機能が内蔵されたインセル型にすることによって形成されている。なお、インセル型は、あくまでも一例に過ぎず、アウトセル型又

はオンセル型であってもよい。

[0037] ユーザデバイス12は、タッチパネル20等によってユーザから受け付けた指示に応じた処理を実行する。例えば、ユーザデバイス12は、タッチパネル20等によってユーザから受け付けた指示に従って、画像処理装置10との間で各種情報の授受を行う。

[0038] ユーザデバイス12は、画像処理装置10から送信された画像を受信し、受信した画像をディスプレイ18に対して表示させる。ユーザ14は、ディスプレイ18に表示された画像を観賞する。

[0039] 画像処理装置10は、コンピュータ22、送受信装置24、及び通信I/F26を備えている。コンピュータ22は、本開示の技術に係る「コンピュータ」の一例であり、プロセッサ28、ストレージ30、及びRAM32を備えている。画像処理装置10は、バス34を備えており、プロセッサ28、ストレージ30、及びRAM32は、バス34を介して接続されている。図1に示す例では、図示の都合上、バス34として1本のバスが図示されているが、複数のバスであってもよい。また、バス34には、シリアルバス、又は、データバス、アドレスバス、及びコントロールバス等で構成されるパラレルバスが含まれていてもよい。

[0040] プロセッサ28は、本開示の技術に係る「プロセッサ」の一例である。プロセッサ28は、画像処理装置10の全体を制御する。例えば、プロセッサ28は、CPU及びGPUを有しており、GPUは、CPUの制御下で動作し、画像処理の実行を担う。

[0041] ストレージ30は、各種パラメータ及び各種プログラム等を記憶している。ストレージ30の一例としては、EEPROM、SSD、及び／又はHDDが挙げられる。ストレージ30は、本開示の技術に係る「メモリ」の一例である。RAM32には、各種情報が一時的に記憶される。RAM32は、プロセッサ28によってワークメモリとして用いられる。

[0042] 送受信装置24は、バス34に接続されている。送受信装置24は、通信用プロセッサ（図示省略）及びアンテナ等を含む装置であり、プロセッサ2

8の制御下で、基地局（図示省略）を介してユーザデバイス12との間で各種情報の送受信を行う。すなわち、プロセッサ28は、送受信装置24を介してユーザデバイス12との間で各種情報の授受を行う。

[0043] 通信1/F26は、例えば、FPGAを有するデバイスによって実現される。通信1/F26は、LANケーブル（図示省略）を介して複数の撮像装置36に接続されている。撮像装置36は、CMOSイメージセンサを有する撮像用のデバイスであり、光学式ズーム機能及び／又はデジタルズーム機能が搭載されている。なお、CMOSイメージセンサに代えてCCDイメージセンサ等の他種類のイメージセンサを採用してもよい。

[0044] 例えば、複数の撮像装置36は、サッカースタジアム（図示省略）に設置されており、サッカースタジアム内の被写体を撮像する。撮像装置36によって被写体が撮像されることで得られた撮像画像64（図4参照）は、例えば、仮想視点画像76（図8等参照）の生成に用いられる。そのため、複数の撮像装置36の各々は、サッカースタジアム内にて、互いに異なる箇所、すなわち、仮想視点画像76（図8等参照）を生成可能な複数の撮像画像64（図4参照）が得られる箇所に設置されている。ここで、複数の撮像画像64は、本開示の技術に係る「複数の撮像画像」の一例である。また、サッカースタジアムは、本開示の技術に係る「撮像領域」の一例である。

[0045] サッカースタジアムは、サッカーフィールドと、サッカーフィールドを取り囲むように建設された観戦席とを含む3次元領域であり、ユーザ14の観察対象とされている。観察者、すなわち、ユーザ14は、観戦席、又は、サッカースタジアム外の場所から、ユーザデバイス12のディスプレイ18によって表示される画像を通して、サッカースタジアム内を観察することができる。

[0046] なお、ここでは、一例として、複数の撮像装置36が設置される場所として、サッカースタジアムが例示されているが、本開示の技術はこれに限定されず、複数の撮像装置36が設置される場所は、野球場、ラグビー場、カーリング場、陸上競技場、競泳場、コンサートホール、野外音楽場、及び演劇

会場等のように、複数の撮像装置 36 が設置可能な場所であれば、如何なる場所であってもよい。

[0047] 通信 I / F 26 は、バス 34 に接続されており、プロセッサ 28 と複数の撮像装置 36 との間で各種情報の授受を司る。例えば、通信 I / F 26 は、プロセッサ 28 の要求に従って複数の撮像装置 36 を制御する。通信 I / F 26 は、複数の撮像装置 36 の各々によって撮像されることで得られた撮像画像 64 (図 4 参照) をプロセッサ 28 に出力する。なお、ここでは、通信 I / F 26 は、有線通信 I / F として例示されているが、高速無線 LAN 等の無線通信 I / F であってもよい。

[0048] ストレージ 30 は、画面生成処理プログラム 38 を記憶している。画面生成処理プログラム 38 は、本開示の技術に係る「プログラム」の一例である。プロセッサ 28 は、ストレージ 30 から画面生成処理プログラム 38 を読み出し、画面生成処理プログラム 38 を RAM 32 上で実行することで画面生成処理 (図 11 参照) を行う。

[0049] 一例として図 2 に示すように、ユーザデバイス 12 は、ディスプレイ 18、コンピュータ 40、撮像装置 42、送受信装置 44、スピーカ 46、マイクロフォン 48、及び受付デバイス 50 を備えている。コンピュータ 40 は、プロセッサ 52、ストレージ 54、及び RAM 56 を備えている。ユーザデバイス 12 は、バス 58 を備えており、プロセッサ 52、ストレージ 54、及び RAM 56 は、バス 58 を介して接続されている。

[0050] 図 2 に示す例では、図示の都合上、バス 58 として 1 本のバスが図示されているが、複数のバスであってもよい。また、バス 58 には、シリアルバス、又は、データバス、アドレスバス、及びコントロールバス等で構成されるパラレルバスが含まれていてもよい。

[0051] プロセッサ 52 は、ユーザデバイス 12 の全体を制御する。プロセッサ 52 は、例えば、CPU 及び GPU を有しており、GPU は、CPU の制御下で動作し、画像処理の実行を担う。

[0052] ストレージ 54 は、各種パラメータ及び各種プログラム等を記憶している

。ストレージ54の一例としては、EEPROMが挙げられる。RAM56には、各種情報が一時的に記憶される。RAM56は、プロセッサ52によってワークメモリとして用いられる。プロセッサ52は、ストレージ54から各種プログラムを読み出し、各種プログラムをRAM56上で実行することで、各種プログラムに応じて処理を行う。

[0053] 撮像装置42は、CMOSイメージセンサを有する撮像用のデバイスであり、光学式ズーム機能及び／又はデジタルズーム機能が搭載されている。なお、CMOSイメージセンサに代えてCCDイメージセンサ等の他種類のイメージセンサを採用してもよい。撮像装置42は、バス58に接続されており、プロセッサ52は、撮像装置42を制御する。撮像装置42によって撮像されることで得られた撮像画像は、バス58を介してプロセッサ52によって取得される。

[0054] 送受信装置44は、バス58に接続されている。送受信装置44は、通信用プロセッサ（図示省略）及びアンテナ等を含む装置であり、プロセッサ52の制御下で、基地局（図示省略）を介して画像処理装置10との間で各種情報の送受信を行う。すなわち、プロセッサ52は、送受信装置44を介して画像処理装置10との間で各種情報の授受を行う。

[0055] スピーカ46は、電気信号を音に変換する。スピーカ46は、バス58に接続されている。スピーカ46は、プロセッサ52から出力された電気信号を、バス58を介して受信し、受信した電気信号を音に変換し、電気信号を変換して得た音をユーザデバイス12の外部に出力する。

[0056] マイクロフォン48は、収集した音を電気信号に変換する。マイクroフォン48は、バス58に接続されている。マイクroフォン48によって収集された音を変換されて得られた電気信号は、バス58を介してプロセッサ52によって取得される。

[0057] 受付デバイス50は、ユーザ14等からの指示を受け付ける。受付デバイス50の一例としては、タッチパネル20及びハードキー（図示省略）等が挙げられる。受付デバイス50は、バス58に接続されており、受付デバイ

ス50によって受け付けられた指示は、プロセッサ52によって取得される。

[0058] 一例として図3に示すように、画像処理装置10において、プロセッサ28は、ストレージ30から画面生成処理プログラム38を読み出し、画面生成処理プログラム38をRAM22C上で実行することで、受付画面生成部28A、仮想視点動画生成部28B、取得部28C、抽出部28D、選定部28E、加工部28F、及び一覧画面生成部28Gとして動作する。以下、受付画面生成部28A、仮想視点動画生成部28B、取得部28C、抽出部28D、選定部28E、加工部28F、及び一覧画面生成部28Gによる処理内容の一例について説明する。

[0059] 一例として図4に示すように、ユーザデバイス12のタッチパネルディスプレイ16には、受付画面66及び仮想視点動画画面68が表示される。図4に示す例では、タッチパネルディスプレイ16に、受付画面66及び仮想視点動画画面68が並べて表示されている。なお、これは、あくまでも一例に過ぎず、タッチパネルディスプレイ16に対してユーザ14から与えられた指示に従って受付画面66と仮想視点動画画面68とが切り替えられて表示されてもよいし、受付画面66と仮想視点動画画面68とが異なる表示デバイスによって個別に表示されるようにしてもよい。

[0060] また、図4に示す例では、ユーザデバイス12のタッチパネルディスプレイ16に受付画面66が表示されているが、本開示の技術はこれに限定されず、例えば、仮想視点動画78（図8参照）を作成したり編集したりする者が使用するデバイス（例えば、ワークステーション及び／又はパーソナルコンピュータ等）に接続されたディスプレイに受付画面66が表示されるようにしてもよい。

[0061] ユーザデバイス12は、画像処理装置10と通信を行うことで、画像処理装置10から、仮想視点動画78（図8参照）を取得する。タッチパネルディスプレイ16の仮想視点動画画面68には、ユーザデバイス12によって画像処理装置10から取得された仮想視点動画78（図8参照）が表示され

る。図4に示す例では、仮想視点動画画面68に仮想視点動画78は表示されていない。

[0062] ユーザデバイス12は、画像処理装置10と通信を行うことで、画像処理装置10から、受付画面66を示す受付画面データ70を取得する。タッチパネルディスプレイ16には、ユーザデバイス12によって画像処理装置10から取得された受付画面データ70により示される受付画面66が表示される。

[0063] 受付画面66には、俯瞰映像画面66A、案内メッセージ表示領域66B、決定キー66C、及び取消キー66Dが含まれており、仮想視点動画78（図8参照）の生成に必要な各種情報が表示される。ユーザ14は、受付画面66を参照してユーザデバイス12に対して指示を与える。ユーザ14からの指示は、例えば、タッチパネルディスプレイ16によって受け付けられる。

[0064] 俯瞰映像画面66Aには、俯瞰映像72が表示される。俯瞰映像72は、サッカースタジアム内を俯瞰で観察した場合の態様を示す動画像であり、複数の撮像装置36のうちの少なくとも1つによって撮像されることで得られる複数の撮像画像64に基づいて生成される。俯瞰映像72の一例としては、録画映像及び／又は生中継映像等が挙げられる。

[0065] 案内メッセージ表示領域66Bには、ユーザ14に要求される操作の内容を示す各種メッセージが表示される。ユーザ14に要求される操作とは、例えば、仮想視点動画78（図8参照）の生成に必要な操作（例えば、視点を設定する操作、及び注視点を設定する操作等）を指す。

[0066] 案内メッセージ表示領域66Bの表示内容は、ユーザデバイス12の動作モードに応じて切り替わる。例えば、ユーザデバイス12は、動作モードとして、視点を設定する視点設定モードと、注視点を設定する注視点設定モードと、を有しており、視点設定モードの場合と注視点設定モードの場合とで案内メッセージ表示領域66Bの表示内容が異なる。

[0067] 決定キー66C及び取消キー66Dは、何れも、ソフトキーである。決定

キー 66C は、受付画面 66 によって受け付けられた指示を決定する場合にユーザ 14 によってオンされる。取消キー 66D は、受付画面 66 によって受け付けられた指示を取り消す場合にユーザ 14 によってオンされる。

[0068] 受付画面生成部 28A は、複数の撮像装置 36 から複数の撮像画像 64 を取得する。撮像画像 64 には、撮像条件情報 64A が含まれている。撮像条件情報 64A とは、撮像条件を示す情報を指す。撮像条件の一例としては、撮像装置 36 の設置位置を特定可能な 3次元座標、撮像装置 36 による撮像方向、及び撮像装置 36 による撮像で用いられる画角、及び撮像装置 36 に対して適用されているズーム倍率等が挙げられる。

[0069] 受付画面生成部 28A は、複数の撮像装置 36 から取得した複数の撮像画像 64 に基づいて俯瞰映像 72 を生成する。そして、受付画面生成部 28A は、受付画面データ 70 として、俯瞰映像 72 を含む受付画面 66 を示すデータを生成する。

[0070] 受付画面生成部 28A は、受付画面データ 70 を送受信装置 24 に出力する。送受信装置 24 は、受付画面生成部 28A から入力された受付画面データ 70 をユーザデバイス 12 に送信する。ユーザデバイス 12 は、送受信装置 24 から送信された受付画面データ 70 を送受信装置 44 (図 2 参照) で受信する。送受信装置 44 によって受信された受付画面データ 70 により示される受付画面 66 は、タッチパネルディスプレイ 16 に表示される。

[0071] 一例として図 5 に示すように、ユーザデバイス 12 の動作モードが視点設定モードの場合、受付画面 66 の案内メッセージ表示領域 66B には、メッセージ 66B1 が表示される。メッセージ 66B1 は、仮想視点動画 78 (図 8 参照) の生成に用いられる視点の指示をユーザ 14 に対して促すメッセージである。ここで、視点とは、サッカースタジアム内を観察する仮想的な視点を指す。仮想的な視点とは、例えば、被写体を撮像する物理カメラ (例えば、撮像装置 36) のように実在するカメラが設置される位置ではなく、被写体を撮像する仮想的なカメラが設置される位置を指す。

[0072] タッチパネルディスプレイ 16 は、案内メッセージ表示領域 66B にメッ

ページ66B1が表示されている状態で、ユーザ14からの指示を受け付ける。この場合、ユーザ14からの指示とは、視点の指示を指す。視点は、俯瞰映像72内の画素の位置に対応している。俯瞰映像72内の画素の位置は、サッカースタジアム内の位置に対応している。視点の指示は、タッチパネルディスプレイ16を介して俯瞰映像72内の画素の位置がユーザ14によって指示されることで行われる。なお、視点は俯瞰映像72内の3次元位置に対応する3次元の座標であってもよい。3次元位置を指示する方法は任意の手法を用いることができる。例えば、ユーザ14が3次元の座標位置を直接入力してもよいし、互いに垂直な2つの平面からサッカースタジアムを見た2つの画像を表示し、それぞれ画素位置を指定することで3次元の座標位置を指定してもよい。

[0073] 図5に示す例では、視点の一例として、被写体を観察するパスである視点パスP1が示されている。視点パスP1は、始点P1sから終点P1eにかけて複数の視点が線状に並べられた集合体である。視点パスP1は、タッチパネル20の全領域のうちの俯瞰映像72の表示領域に対応する領域上でユーザ14が自身の指先14Aをスライド（スワイプ）させた経路（図5に示す例では、始点P1sから終点P1eまでの蛇行した経路）に沿って規定されている。また、視点パスP1からの観察時間（例えば、2つの異なる視点間で観察する時間、及び／又は、ある1点で静止して観察する時間等）は、タッチパネルディスプレイ16を介して視点パスP1を形成する場合にタッチパネルディスプレイ16に対して行われるスライドの速度及び視点パスP1上の1つの視点に留まる時間（例えば、ロングプレスの時間）等によって規定されている。

[0074] 図5に示す例において、視点パスP1を確定する場合、決定キー66Cがオンされ、視点パスP1を取り消す場合、取消キー66Dがオンされる。

[0075] なお、図5に示す例では、視点パスP1のみが設定されているが、これはあくまでも一例に過ぎず、複数の視点パスが設定されていてもよい。また、視点パスに限らず、複数の不連続な視点であってもよいし、1つの視点であ

ってもよい。

[0076] 一例として図6に示すように、ユーザデバイス12の動作モードが注視点設定モードの場合、受付画面66の案内メッセージ表示領域66Bには、メッセージ66B2が表示される。メッセージ66B2は、仮想視点動画78（図8参照）の生成に用いられる注視点の指示をユーザ14に対して促すメッセージである。ここで、注視点とは、サッカースタジアム内を視点から観察する場合に仮想的に注視する点を指す。視点及び注視点が設定されると、仮想的な視線方向（仮想的なカメラの撮像方向）も一意に定まる。仮想的な視線方向とは、視点から注視点に向かう方向を指す。

[0077] タッチパネルディスプレイ16は、案内メッセージ表示領域66Bにメッセージ66B2が表示されている状態で、ユーザ14からの指示を受け付ける。この場合、ユーザ14からの指示とは、注視点の指示を指す。注視点は、俯瞰映像72内の画素の位置に対応している。俯瞰映像72内の画素の位置は、サッカースタジアム内の位置に対応している。注視点の指示は、タッチパネルディスプレイ16を介して俯瞰映像72内の画素の位置がユーザ14によって指示されることによって行われる。図6に示す例では、注視点GPが示されている。注視点GPは、タッチパネルディスプレイ16の全領域のうちの俯瞰映像72の表示領域に対応する領域上でユーザ14が自身の指先14Aをタッチした箇所に従って規定されている。図6に示す例において、注視点GPを確定する場合、決定キー66Cがオンされ、注視点GPを取り消す場合、取消キー66Dがオンされる。なお、注視点は俯瞰映像72内の3次元位置に対応する3次元の座標であってもよい。3次元位置を指示する方法は、視点位置の指示と同様に任意の手法を用いることができる。

[0078] なお、図6に示す例では、注視点GPのみが指定されているが、これはあくまでも一例に過ぎず、複数の注視点であってもよいし、複数の注視点が線状に並べられたパス（注視点パス）であってもよい。注視点パスは1つであってもよいし、複数であってもよい。

[0079] 一例として図7に示すように、ユーザデバイス12のプロセッサ52は、

視点パス P 1 及び注視点 G P に基づいて複数の視点情報 7 4 を生成する。複数の視点情報 7 4 は、本開示の技術に係る「複数の視点情報」の一例である。

[0080] 視点情報 7 4 は、仮想視点動画 7 8 (図 8 参照) の生成に用いられる情報である。視点情報 7 4 は、視点位置情報 7 4 A、視線方向情報 7 4 B、画角情報 7 4 C、移動速度情報 7 4 D、及び経過時間情報 7 4 E を含んでいる。

[0081] 視点位置情報 7 4 A は、視点の位置 (以下、「視点位置」とも称する) を特定可能な情報である。視点位置とは、例えば、上述した仮想カメラの位置を指す。ここでは、視点位置の一例として、視点設定モードで確定した視点パス P 1 (図 5 参照) に含まれる 1 つの視点の俯瞰映像 7 2 内での画素の位置が適用されている。視点パス P 1 の俯瞰映像 7 2 内での画素の位置を特定する情報の一例としては、俯瞰映像 7 2 内での視点パス P 1 の画素の位置を特定する座標が挙げられる。

[0082] 視点パス P 1 には始点 P 1 s 及び終点 P 1 e (図 5 参照) が含まれている。従って、視点パス P 1 に含まれる全ての視点を示す複数の視点位置情報 7 4 A には、始点 P 1 s の位置を特定可能な始点位置情報 (以下、単に「始点位置情報」とも称する)、及び終点 P 1 e の位置を特定可能な終点位置情報 (以下、単に「終点位置情報」とも称する) も含まれている。始点位置情報の一例としては、俯瞰映像 7 2 内での始点 P 1 s の画素の位置を特定可能な座標が挙げられる。終点位置情報の一例としては、俯瞰映像 7 2 内での終点 P 1 e の画素の位置を特定可能な座標が挙げられる。

[0083] 視線方向情報 7 4 B は、視線方向を特定可能な情報である。視線方向とは、例えば、視点パス P 1 に含まれる視点から注視点 G P に向かって被写体を観察する方向を指す。視線方向情報 7 4 B は、例えば、視点パス P 1 に含まれる全ての視点を示す複数の視点位置情報 7 4 A から特定される視点毎に定められており、視点の位置を特定可能な情報 (例えば、俯瞰映像 7 2 内での視点の画素の位置を特定可能な座標)、及び注視点設定モードで確定した注視点 G P の位置を特定可能な情報 (例えば、俯瞰映像 7 2 内での注視点 G P

の画素の位置を特定する座標)によって規定されている。

- [0084] 画角情報74Cは、画角(以下、単に「画角」とも称する)を示す情報である。ここで、画角とは、視点パスP1上で被写体を観察する視野角を指す。本第1実施形態では、画角は、既定角度(例えば、100度)に固定されている。但し、これは、あくまでも一例に過ぎず、画角は、移動速度に応じて定められてもよい。ここで、移動速度とは、視点パスP1上で被写体を観察する視点位置が移動する速度を指す。移動速度の一例としては、タッチパネルディスプレイ16を介して視点パスP1を形成する場合にタッチパネルディスプレイ16に対して行われるスライドの速度が挙げられる。
- [0085] 画角が移動速度に応じて定められる場合、例えば、画角の上限(例えば、150度)と下限(例えば、15度)とが定められた範囲内において、移動速度が低くなるほど画角は狭くなる。また、移動速度が高くなるほど画角を狭めてもよい。
- [0086] また、画角は、視点位置に対応する経過時間(以下、単に「経過時間」とも称する)に応じて定められてもよい。ここで、経過時間とは、例えば、視点パスP1上のある視点位置で視点が静止している時間を指す。
- [0087] 画角が経過時間に応じて定められる場合、例えば、経過時間が第1既定時間(例えば、3秒)を超えた場合に画角を最小にしたり、経過時間が第1既定時間を超えた場合に画角を最大にしたりすればよい。
- [0088] また、画角は、受付デバイス50によって受け付けられた指示に従って定められてもよい。この場合、視点パスP1上で画角を変更する視点位置、及び、変更後の画角に関する指示が受付デバイス50によって受け付けられるようにすればよい。
- [0089] 移動速度情報74Dは、上述した移動速度(以下、単に「移動速度」とも称する)を示す情報であり、視点パスP1内の対応する視点毎に対応付けられている。経過時間情報74Eは、経過時間を示す情報である。
- [0090] プロセッサ52は、複数の視点情報74を送受信装置44に出力する。送受信装置44は、プロセッサ52から入力された複数の視点情報74を画像

処理装置 10 に送信する。画像処理装置 10 の送受信装置 24 は、送受信装置 44 から送信された複数の視点情報 74 を受信する。画像処理装置 10 の仮想視点動画生成部 28 B は、送受信装置 24 によって受信された複数の視点情報 74 を取得する。

[0091] 一例として図 8 に示すように、仮想視点動画生成部 28 B は、複数の視点情報 74 (例えば、図 5 に示す視点パス P 1 を特定可能な複数の視点情報 74) に従って、仮想視点画像 76 の生成に用いる複数の撮像画像 64 (図 4 参照) を選定する。すなわち、仮想視点動画生成部 28 B は、複数の視点情報 74 に従って被写体を観察した場合の被写体の態様を示す画像である仮想視点画像 76 の生成に用いる複数の撮像画像 64 (図 4 参照) を、複数の撮像装置 36 (図 1 及び図 4 参照) によって撮像されることで得られた複数の撮像画像 64 から選定する。

[0092] 仮想視点動画生成部 28 B は、複数の視点情報 74 及び複数の撮像画像 64 に基づいて仮想視点動画 78 を生成する。すなわち、仮想視点動画生成部 28 B は、複数の視点情報 74 に従って選定した複数の撮像画像 64 に基づいて、複数の視点情報 74 (例えば、図 5 に示す視点パス P 1 を特定可能な複数の視点情報 74) により特定される視点から被写体を観察した場合の被写体の態様を示す動画である仮想視点動画 78 を生成する。

[0093] 例えば、仮想視点動画生成部 28 B は、視点パス P 1 (図 5 参照) に従って複数フレームの仮想視点画像 76 を生成する。すなわち、仮想視点動画生成部 28 B は、視点パス P 1 上の視点毎に仮想視点画像 76 を生成する。仮想視点動画生成部 28 B は、複数フレームの仮想視点画像 76 を時系列で並べることで仮想視点動画 78 を生成する。このようにして生成された仮想視点動画 78 は、ユーザデバイス 12 のタッチパネルディスプレイ 16 に対して表示させるためのデータである。仮想視点動画 78 がタッチパネルディスプレイ 16 に表示される時間は、複数の視点情報 74 (例えば、図 1 に示す視点パス P 1 を示す複数の視点情報 74) に応じて定められる。

[0094] 仮想視点動画生成部 28 B は、仮想視点動画 78 に含まれる複数フレーム

の仮想視点画像 76 の各々に対してメタデータ 76 A を付与する。メタデータ 76 A は、例えば、仮想視点画像 76 の生成に用いられた撮像画像 64 に含まれる撮像条件情報 64 A (図 4 参照) に基づいて仮想視点動画生成部 28 B によって生成される。メタデータ 76 A には、仮想視点画像 76 が生成された時刻、及び、撮像条件情報 64 A に基づく情報が含まれる。

[0095] 仮想視点動画生成部 28 B は、仮想視点動画 78 を生成する毎に、仮想視点動画 78 に対して動画識別情報 80 を付与する。動画識別情報 80 は、仮想視点動画 78 に対して固有に割り当てられた識別子を含み、仮想視点動画 78 の特定に用いられる。また、動画識別情報 80 には、仮想視点動画 78 が生成された時刻、及び／又は、仮想視点動画 78 の総再生時間等のメタデータが含まれる。

[0096] 仮想視点動画生成部 28 B は、生成した仮想視点動画 78 をストレージ 30 に記憶させる。ストレージ 30 には、例えば、視点パス P1 を含めた複数の視点パスについて仮想視点動画生成部 28 B によって生成された仮想視点動画 78 が記憶される。

[0097] 一例として図 9 に示すように、取得部 28 C は、仮想視点動画生成部 28 B によって仮想視点動画 78 (図 9 に示す例では、ストレージ 30 に記憶されている仮想視点動画 78) の生成で用いられた複数の視点情報 74 を仮想視点動画生成部 28 B から取得する。取得部 28 C は、ストレージ 30 に記憶されている仮想視点動画 78 から特定区間仮想視点動画 78 A を取得する。特定区間仮想視点動画 78 A は、仮想視点動画 78 のうちの視点位置、視線方向、及び画角が固定された時間帯 (例えば、視点パス P1 に含まれる複数の視点位置のうちの視点が静止している時間が最も長い視点位置に関する視点情報 74 から特定される時間帯) の仮想視点動画である。すなわち、仮想視点動画 78 のうちの視点位置、視線方向、及び画角が固定された時間帯の仮想視点動画とは、例えば、複数の視点情報 74 のうち、最も長い経過時間を示す経過時間情報 74 E が含まれている視点情報 74 に従って仮想視点動画生成部 28 B によって生成された仮想視点動画 (すなわち、複数フレー

ムの仮想視点画像)を指す。

- [0098] 抽出部28Dは、仮想視点動画78に含まれる時間(図9に示す例では、視点位置、視線方向、及び画角が固定された時間帯)に基づいて決められた対象被写体81を特定する。ここで、対象被写体81は、本開示の技術に係る「第1被写体」の一例である。
- [0099] 仮想視点動画78に含まれる時間の第1例としては、被写体が写っている時間の長さが挙げられる。また、仮想視点動画78に含まれる時間の第2例としては、仮想視点動画78の総再生時間のうちの最初及び／又は最後の時間帯(例えば、数秒間)が挙げられる。また、仮想視点動画78に含まれる時間の第3例としては、時点が挙げられる。
- [0100] 本第1実施形態において、抽出部28Dは、取得部28Cによって取得された特定区間仮想視点動画78Aに含まれる全ての仮想視点画像76に対してA1方式の被写体認識処理を行うことで、特定区間仮想視点動画78Aに最も長く写っている被写体を対象被写体81として特定する。そして、抽出部28Dは、特定した対象被写体81が含まれている複数フレームの仮想視点画像76を特定区間仮想視点動画78Aから抽出する。
- [0101] なお、ここでは、A1方式の被写体認識処理が行われる形態例を挙げて説明しているが、これは、あくまでも一例に過ぎず、テンプレートマッチング方式の被写体認識処理が行われてもよい。また、仮想視点動画78に含まれる全ての仮想視点画像76に含まれている被写体に対して、被写体を特定する識別子(以下、「被写体識別子」と称する)が事前に付与されており、抽出部28Dが、被写体識別子を参照して各仮想視点画像76に含まれている被写体を特定するようにしてよい。
- [0102] 選定部28Eは、抽出部28Dによって抽出された複数フレームの仮想視点画像76での対象被写体81のサイズに基づいて決められた1つのフレームの仮想視点画像76を選定する。例えば、選定部28Eは、抽出部28Dによって抽出された複数フレームの仮想視点画像76から、最大サイズの対象被写体81が含まれている1フレームの仮想視点画像76を選定する。例

例えば、抽出部 28D によって A1 方式の被写体認識処理が行われた場合、選定部 28E は、A1 方式の被写体認識処理で用いられるバウンディングボックスのサイズを参照することで最大サイズの対象被写体 81 が含まれている仮想視点画像 76 を特定する。

[0103] ここで、抽出部 28D によって抽出された複数フレームは、本開示の技術に係る「仮想視点動画において撮像領域内の第 1 被写体が含まれている複数のフレーム」の一例である。また、最大サイズの対象被写体 81 が含まれている 1 フレームの仮想視点画像 76 は、本開示の技術に係る「第 1 フレームに関する画像」の一例である。また、「最大サイズ」は、本開示の技術に係る「第 1 被写体のサイズ」の一例である。

[0104] なお、ここでは、最大サイズの対象被写体 81 を例示しているが、これは、あくまでも一例に過ぎず、最大サイズ以外の指定されたサイズ（例えば、最大サイズの次に大きなサイズ）の対象被写体 81 であってもよいし、予め定められたサイズ範囲（例えば、受付デバイス 50 等によって受け付けられた指示に応じて定められたサイズ範囲）内で最大サイズの対象被写体 81 であってもよいし、受付デバイス 50 等によって受け付けられた指示に応じて定められたサイズの対象被写体 81 であってもよい。

[0105] 加工部 28F は、仮想視点動画 78 を、仮想視点動画 78 とは異なる大きさの画像に加工する。仮想視点動画 78 とは異なる大きさの画像としては、例えば、仮想視点動画 78 よりもデータ量が少ない画像（例えば、少なくとも 1 フレーム分の画像）、仮想視点動画 78 が間引かれた画像（例えば、コマ送り画像）、仮想視点動画 78 に含まれる少なくとも 1 フレーム分の仮想視点画像 76 の表示サイズを縮小した画像、及び／又は、仮想視点動画 78 に含まれる少なくとも 1 フレーム分の仮想視点画像 76 内の画素を間引いて得た画像等が挙げられる。

[0106] 加工部 28F は、仮想視点動画 78 に含まれる全ての仮想視点画像 76 のうちの 1 フレームの仮想視点画像 76 に関する画像を生成する。1 フレームの仮想視点画像 76 に関する画像は、例えば、仮想視点動画 78 の内容を示

す画像である。ここで、1フレームの仮想視点画像76に関する画像は、本開示の技術に係る「第1フレームに関する画像」の一例である。1フレームの仮想視点画像76に関する画像としては、例えば、1フレームの仮想視点画像76の全体そのもの、1フレームの仮想視点画像76から切り出した一部、及び／又は、1フレームの仮想視点画像76が加工された画像が挙げられる。

[0107] 加工部28Fは、複数の撮像画像64及び複数の視点情報74に基づいて、仮想視点動画78に対応するサムネイル画像82を取得する。サムネイル画像82は、本開示の技術に係る「代表画像」の一例である。すなわち、加工部28Fは、仮想視点動画78に含まれる全ての仮想視点画像76のうちの代表的な1フレームの仮想視点画像76をサムネイル化する。例えば、加工部28Fは、選定部28Eによって選定された仮想視点画像76をサムネイル画像82に加工する。仮想視点画像76をサムネイル画像82に加工する方法としては、上記の仮想視点動画78を、仮想視点動画78とは異なる大きさの画像に加工する方法を用いることができる。また、加工部28Fは、サムネイル化される前の仮想視点画像76に付与されているメタデータ76Aをサムネイル画像82に対して関連付ける。また、加工部28Fは、サムネイル化された仮想視点画像76が含まれる仮想視点動画78から動画識別情報80を取得する。

[0108] 一例として図10に示すように、加工部28Fは、仮想視点画像76をサムネイル化して得たサムネイル画像82に対して動画識別情報80を関連付ける。

[0109] 一覧画面生成部28Gは、メタデータ76A及び動画識別情報80が関連付けられたサムネイル画像82を加工部28Fから取得する。一覧画面生成部28Gは、メタデータ76A及び／又は動画識別情報80に基づいて参照情報86Aを生成し、サムネイル画像82に対応付ける。一覧画面生成部28Gは、参照情報86Aが対応付けられたサムネイル画像82を含めた一覧画面86を示す一覧画面データ84を生成する。一覧画面データ84は、サ

ムネイル画像 82 をユーザデバイス 12 のタッチパネルディスプレイ 16 に表示させるためのデータである。一覧画面生成部 28G は、生成した一覧画面データ 84 を送受信装置 24 に出力し、かつ、ストレージ 30 に記憶させる。これにより、ストレージ 30 には、動画識別情報 80 と関連付けられたサムネイル画像 82 が記憶される。すなわち、動画識別情報 80 は、仮想視点動画 78 に対して固有に割り当てられた識別子なので、ストレージ 30 には、サムネイル画像 82 と仮想視点動画 78 とが関連付けられた状態で記憶される。

[0110] 一覧画面データ 84 は、本開示の技術に係る「データ」及び「第 1 データ」の一例である。また、タッチパネルディスプレイ 16 は、本開示の技術に係る「ディスプレイ」及び「第 1 ディスプレイ」の一例である。

[0111] 一覧画面生成部 28G によってサムネイル画像 82 に対応付けられる参照情報 86A としては、例えば、文字情報が挙げられる。文字情報としては、例えば、仮想視点動画 78 が生成された時刻（例えば、図 4 に示す撮像条件情報 64A から特定される時刻）、サムネイル画像 82 に含まれる対象被写体 81 に関連する情報（例えば、対象被写体 81 の名称、及び／又は、対象被写体 81 の所属チーム等）、仮想視点動画 78 の総再生時間、仮想視点動画 78 のタイトル、及び／又は、仮想視点動画 78 の作成者の名称等を示す文字情報が挙げられる。

[0112] ストレージ 30 に一覧画面データ 84 が記憶されている状態で、加工部 28F によって、サムネイル画像 82 が生成され、かつ、生成されたサムネイル画像 82 に対してメタデータ 76A 及び動画識別情報 80 が関連付けられると、一覧画面生成部 28G は、ストレージ 30 から一覧画面データ 84 を取得し、一覧画面データ 84 を更新する。すなわち、一覧画面生成部 28G は、加工部 28F から、メタデータ 76A 及び動画識別情報 80 が関連付けられたサムネイル画像 82 を取得し、参照情報 86A を生成する。一覧画面生成部 28G は、生成した参照情報 86A をサムネイル画像 82 に対応付ける。そして、一覧画面生成部 28G は、参照情報 86A が対応付けられたサ

ムネイル画像 82 を一覧画面 86 に含めることで一覧画面データ 84 を更新する。一覧画面生成部 28G は、更新した一覧画面データ 84 を送受信装置 24 に出力し、かつ、ストレージ 30 に記憶させる。

[0113] 更新された一覧画面データ 84 により示される一覧画面 86 には、複数のサムネイル画像 82 が含まれる。また、更新された一覧画面データ 84 により示される一覧画面 86 において、複数のサムネイル画像 82 の各々には参照情報 86A が対応付けられている。

[0114] 送受信装置 24 は、一覧画面生成部 28G から入力された一覧画面データ 84 をユーザデバイス 12 に送信する。ユーザデバイス 12 において、送受信装置 44 は、画像処理装置 10 から送信された一覧画面データ 84 を受信する。プロセッサ 52 は、送受信装置 44 によって受信された一覧画面データ 84 を取得し、取得した一覧画面データ 84 により示される一覧画面 86 をタッチパネルディスプレイ 16 に表示させる。一覧画面 86 には、複数の画像が並行して表示される。図 10 に示す例では、一覧画面 86 に、複数のサムネイル画像 82 が参照情報 86A と共に表示されている。すなわち、参照情報 86A は、サムネイル画像 82 との関連性が視覚的に把握可能な態様（例えば、参照情報 86A とサムネイル画像 82 とが 1対1 の関係にあることが視覚的に把握可能に整列された態様）で一覧画面 86 に表示される。

[0115] なお、ここでは、複数のサムネイル画像 82 が一覧画面 86 に表示される形態例を挙げているが、一覧画面 86 にサムネイル画像 82 が 1つだけ表示されていてもよい。また、複数のサムネイル画像 82 は、必ずしも並行して表示される必要はなく、複数のサムネイル画像 82 が視覚的に把握可能な態様であればどのような表示でもよい。

[0116] タッチパネルディスプレイ 16 に一覧画面 86 が表示されている状態で、ユーザ 14 は、タッチパネルディスプレイ 16 を介して一覧画面 86 内の何れかのサムネイル画像 82 をタップすることでサムネイル画像 82 を選択する。サムネイル画像 82 が選択されると、これに応じて、画像処理装置 10 のプロセッサ 28（図 1 及び図 3 参照）は、仮想視点動画 78 をタッチパネ

ルディスプレイ 16 に対して表示させるためのデータをユーザデバイス 12 に出力する。

[0117] 例えば、ユーザ 14 によってタッチパネルディスプレイ 16 を介してサムネイル画像 82 が選択されると、ユーザデバイス 12 のプロセッサ 52 は、選択されたサムネイル画像 82 に関連付けられている動画識別情報 80 を、送受信装置 44 を介して画像処理装置 10 に送信する。画像処理装置 10 では、送受信装置 24 によって動画識別情報 80 が受信される。画像処理装置 10 のプロセッサ 28 は（図 1 及び図 3 参照）は、送受信装置 24 によって受信された動画識別情報 80 に対応する仮想視点動画 78 をストレージ 30 から取得し、取得した仮想視点動画 78 を、送受信装置 24 を介してユーザデバイス 12 に送信する。ユーザデバイス 12 では、画像処理装置 10 から送信された仮想視点動画 78 が送受信装置 44 によって受信される。ユーザデバイス 12 のプロセッサ 52 は、送受信装置 44 によって受信された仮想視点動画 78 をタッチパネルディスプレイ 16 に表示させる。例えば、仮想視点動画 78 は、タッチパネルディスプレイ 16 の仮想視点動画画面 68（図 4 参照）に表示される。

[0118] なお、ここでは、仮想視点動画 78 がタッチパネルディスプレイ 16 に表示される形態例を挙げているが、これは、あくまでも一例に過ぎず、例えば、タッチパネルディスプレイ 16 に代えて、又は、タッチパネルディスプレイ 16 と共に、画像処理装置 10 に直接的又は間接的に接続されているディスプレイに対して仮想視点動画 78 が表示されるようにしてもよい。この場合、画像処理装置 10 に直接的又は間接的に接続されているディスプレイは、本開示の技術に係る「第 2 ディスプレイ」の一例である。

[0119] また、ここでは、一覧画面 86 内の何れかのサムネイル画像 82 がタップされることでサムネイル画像 82 が選択される形態例を挙げて説明したが、これは、あくまでも一例に過ぎず、タッチパネルディスプレイ 16 を介してサムネイル画像 82 に対してフリック、スワイプ、及び／又はロングプレス等が行われることによってサムネイル画像 82 が選択されるようにしてもよ

いし、マイクロフォン48によって取得された音声に対する音声認識処理が行われることによってサムネイル画像82が選択されるようにしてもよいし、マウス及び／又はキーボード等の操作によってサムネイル画像82が選択されるようにしてもよい。

[0120] 次に、本第1実施形態に係る画像処理装置10の動作について図11を参照しながら説明する。

[0121] なお、図11には、画像処理装置10のプロセッサ28によって行われる画面生成処理の流れの一例が示されている。図11に示す画面生成処理の流れは、本開示の技術に係る画像処理方法の一例である。

[0122] 図11に示す画面生成処理では、まず、ステップST10で、仮想視点動画生成部28Bは、ユーザデバイス12から複数の視点情報74（例えば、視点パスP1に対応する複数の視点情報74）を取得する（図7参照）。ステップST10の処理が実行された後、画面生成処理はステップST12へ移行する。

[0123] ステップST12で、仮想視点動画生成部28Bは、ステップST10で取得した複数の視点情報74に従って複数の撮像画像64を選定する（図8参照）。ステップST12の処理が実行された後、画面生成処理はステップST14へ移行する。

[0124] ステップST14で、仮想視点動画生成部28Bは、ステップST12で選定した複数の撮像画像64に基づいて仮想視点動画78を生成し、生成した仮想視点動画78をストレージ30に記憶させる（図8参照）。ステップST14の処理が実行された後、画面生成処理はステップST16へ移行する。

[0125] ステップST16で、取得部28Cは、仮想視点動画生成部28Bによる仮想視点動画78の生成に用いられた複数の視点情報74に従って、仮想視点動画78のうち、視点位置、視線方向、及び画角が固定された時間帯の仮想視点動画を特定区間仮想視点動画78Aとしてストレージ30から取得する（図9参照）。ステップST16の処理が実行された後、画面生成処理は

ステップS T 1 8へ移行する。

- [0126] ステップS T 1 8で、抽出部2 8 Dは、特定区間仮想視点動画7 8 Aに対してA 1方式の被写体認識処理を行うことで、特定区間仮想視点動画7 8 Aに最も長く写っている対象被写体8 1が含まれている複数の仮想視点画像7 6を特定区間仮想視点動画7 8 Aから抽出する（図9参照）。ステップS T 1 8の処理が実行された後、画面生成処理はステップS T 2 0へ移行する。
- [0127] ステップS T 2 0で、選定部2 8 Eは、ステップS T 1 8で抽出された複数の仮想視点画像7 6から、最大サイズの対象被写体8 1が含まれている仮想視点画像7 6を選定する（図9参照）。ステップS T 2 0の処理が実行された後、画面生成処理はステップS T 2 2へ移行する。
- [0128] ステップS T 2 2で、加工部2 8 Fは、ステップS T 2 0で選定された仮想視点画像7 6をサムネイル画像8 2に加工する（図9及び図10参照）。サムネイル画像8 2には、ステップS T 2 0で選定された仮想視点画像7 6のメタデータ7 6 Aが加工部2 8 Fによって付与される。ステップS T 2 2の処理が実行された後、画面生成処理はステップS T 2 4へ移行する。
- [0129] ステップS T 2 4で、加工部2 8 Fは、ステップS T 2 2で得られたサムネイル画像8 2に対応する仮想視点画像7 6が含まれる仮想視点動画7 8に関する動画識別情報8 0をストレージ3 0から取得し（図9参照）、取得した動画識別情報8 0をサムネイル画像8 2に対して関連付ける（図10参照）。ステップS T 2 4の処理が実行された後、画面生成処理はステップS T 2 6へ移行する。
- [0130] ステップS T 2 6で、一覧画面生成部2 8 Gは、メタデータ7 6 A及び動画識別情報8 0が関連付けられたサムネイル画像8 2を含む一覧画面8 6を示す一覧画面データ8 4を生成し、生成した一覧画面データ8 4をストレージ3 0及び送受信装置2 4に出力する（図10参照）。これにより、ストレージ3 0に一覧画面データ8 4が記憶され、送受信装置2 4によって一覧画面データ8 4がユーザデバイス1 2に送信される。ユーザデバイス1 2では、送受信装置2 4から送信された一覧画面データ8 4により示される一覧画

面 8 6 がプロセッサ 5 2 によってタッチパネルディスプレイ 1 6 に表示される (図 1 0 参照)。ステップ S T 2 6 の処理が実行された後、画面生成処理はステップ S T 2 8 へ移行する。

[0131] ステップ S T 2 8 で、一覧画面生成部 2 8 G は、画面生成処理が終了する条件 (以下、「終了条件」と称する) を満足したか否かを判定する。終了条件の一例としては、画面生成処理を終了させる指示がタッチパネルディスプレイ 1 6 等の受付デバイスによって受け付けられた、との条件が挙げられる。ステップ S T 2 8 において、終了条件を満足していない場合は、判定が否定されて、画面生成処理はステップ S T 1 0 へ移行する。ステップ S T 2 8 において、終了条件を満足した場合は、判定が肯定されて、画面生成処理が終了する。

[0132] 以上説明したように、本第 1 実施形態に係る画像処理装置 1 0 では、複数の撮像画像 6 4 及び複数の視点情報 7 4 に基づいて生成される仮想視点動画 7 8 に対応するサムネイル画像 8 2 が、複数の撮像画像 6 4 及び複数の視点情報 7 4 に基づいて取得される。そして、サムネイル画像 8 2 をユーザデバイス 1 2 のタッチパネルディスプレイ 1 6 に表示させるためのデータとして一覧画面データ 8 4 がユーザデバイス 1 2 に送信される。ユーザデバイス 1 2 では、一覧画面データ 8 4 により示される一覧画面 8 6 がタッチパネルディスプレイ 1 6 に表示される。従って、本構成によれば、ユーザ 1 4 に対して仮想視点動画 7 8 に対応するサムネイル画像 8 2 を見せることに寄与することができる。

[0133] また、本第 1 実施形態に係る画像処理装置 1 0 では、仮想視点動画 7 8 に含まれる特定区間仮想視点動画 7 8 A が取得される。そして、特定区間仮想視点動画 7 8 A に含まれる複数フレームの仮想視点画像 7 6 のうちの 1 フレームの仮想視点画像 7 6 に対応するサムネイル画像 8 2 が取得される。そして、サムネイル画像 8 2 をユーザデバイス 1 2 のタッチパネルディスプレイ 1 6 に表示させるためのデータとして一覧画面データ 8 4 がユーザデバイス 1 2 に送信される。従って、本構成によれば、特定区間仮想視点動画 7 8 A

に含まれる複数フレームの仮想視点画像 76 のうちの 1 フレームの仮想視点画像 76 に対応するサムネイル画像 82 をユーザ 14 に見せることに寄与することができる。

[0134] また、本第 1 実施形態に係る画像処理装置 10 では、仮想視点動画 78 に含まれる時間に基づいて決められた対象被写体 81 が含まれている 1 フレームの仮想視点画像 76 に対応するサムネイル画像 82 が取得される。そして、サムネイル画像 82 をユーザデバイス 12 のタッチパネルディスプレイ 16 に表示させるためのデータとして一覧画面データ 84 がユーザデバイス 12 に送信される。従って、本構成によれば、仮想視点動画 78 に含まれる時間に基づいて決められた対象被写体 81 が含まれている 1 フレームの仮想視点画像 76 に対応するサムネイル画像 82 をユーザ 14 に見せることに寄与することができる。

[0135] また、本第 1 実施形態に係る画像処理装置 10 では、特定区間仮想視点動画 78 A 内での対象被写体 81 のサイズに基づいて決められた 1 フレームの仮想視点画像 76 に対応するサムネイル画像 82 が取得される。そして、サムネイル画像 82 をユーザデバイス 12 のタッチパネルディスプレイ 16 に表示させるためのデータとして一覧画面データ 84 がユーザデバイス 12 に送信される。従って、本構成によれば、対象被写体 81 のサイズに基づいて決められた 1 フレームの仮想視点画像 76 に対応するサムネイル画像 82 をユーザ 14 に見せることに寄与することができる。

[0136] また、本第 1 実施形態に係る画像処理装置 10 では、タッチパネルディスプレイ 16 に表示されているサムネイル画像 82 が選択されたことに応じて、選択されたサムネイル画像 82 に対応する仮想視点動画 78 をタッチパネルディスプレイ 16 に表示させるためのデータとして一覧画面データ 84 がユーザデバイス 12 に送信される。従って、本構成によれば、ユーザ 14 に対して、選択されたサムネイル画像 82 に対応する仮想視点動画 78 を観賞させることに寄与することができる。

[0137] また、本第 1 実施形態に係る画像処理装置 10 では、サムネイル画像 82

と仮想視点動画 78 とが関連付けられた状態でストレージ 30 に記憶される。従って、本構成によれば、サムネイル画像 82 と仮想視点動画 78 とが関連付けられていない場合に比べ、サムネイル画像 82 から仮想視点動画 78 を迅速に得ることができる。

[0138] また、本第 1 実施形態に係る画像処理装置 10 では、複数の画像が並行して表示される一覧画面データ 84 にサムネイル画像 82 を表示させるためのデータが一覧画面データ 84 としてユーザデバイス 12 に送信される。従って、本構成によれば、ユーザ 14 に対して、複数の画像とサムネイル画像 82 とを一覧させることに寄与することができる。

[0139] なお、上記実施形態では、視点位置、視線方向、及び画角が固定された時間帯の仮想視点動画を特定区間仮想視点動画 78 A としたが、本開示の技術は、これに限定されない。例えば、仮想視点動画 78 のうち、ユーザ 14 等によって指定された時間帯の仮想視点動画を特定区間仮想視点動画 78 A としてもよいし、複数の視点情報 74 のうち、既定の速度範囲内の移動速度を示す移動速度情報 74 D を有する少なくとも 1 つの視点情報 74 から特定された仮想視点動画を特定区間仮想視点動画 78 A としてもよいし、特定の視点位置、特定の視線方向、及び／又は特定の画角に対応する少なくとも 1 つの視点情報 74 から特定された仮想視点動画を特定区間仮想視点動画 78 A としてもよい。

[0140] [第 2 実施形態]

本第 2 実施形態では、上記第 1 実施形態で説明した構成要素については、同一の符号を付し、説明を省略し、上記第 1 実施形態と異なる部分について説明する。

[0141] 一例として図 12 に示すように、本第 2 実施形態に係る画像処理装置 10 のプロセッサ 28 は、図 3 に示すプロセッサ 28 に比べ、画面生成処理プログラム 38 を実行することで編集結果取得部 28 H として更に動作する点が異なる。

[0142] 一例として図 13 に示すように、視点パス P1 は、ユーザ 14 による指示

がタッチパネルディスプレイ 16 によって受け付けられることによって編集される。図 13 に示す例では、視点パス P1 の編集前後で始点 P1s 及び終点 P1e は共通であり、始点 P1s から終点 P1e に至るまでのパスが異なる。

[0143] 一例として図 14 に示すように、ユーザデバイス 12 において、プロセッサ 52 は、上記第 1 実施形態で説明した複数の視点情報 74、すなわち、編集前の視点パス P1 に関する複数の視点情報 74 を編集前視点パス情報 88 として送受信装置 44 を介して画像処理装置 10 に送信する。また、プロセッサ 52 は、編集後の視点パス P1 及び注視点 GP (図 6 参照) に基づいて編集後視点パス情報 90 を生成する。編集後視点パス情報 90 は、編集後の視点パス P1 に関する複数の視点情報 74 を有する。プロセッサ 52 は、視点パス P1 が編集されたことに応じて編集後視点パス情報 90 を生成し、生成した編集後視点パス情報 90 を、送受信装置 44 を介して画像処理装置 10 に送信する。

[0144] 一例として図 15 に示すように、仮想視点動画生成部 28B は、編集後視点パス情報 90 (図 14 参照) に従って、仮想視点画像 92 の生成に用いる複数の撮像画像 64 (図 4 参照) を選定する。すなわち、仮想視点動画生成部 28B は、編集後視点パス情報 90 に従って被写体を観察した場合の被写体の態様を示す画像である仮想視点画像 92 の生成に用いる複数の撮像画像 64 (図 4 参照) を、複数の撮像装置 36 (図 1 及び図 4 参照) によって撮像されることで得られた複数の撮像画像 64 (図 4 参照) から選定する。

[0145] 仮想視点動画生成部 28B は、編集後視点パス情報 90 及び複数の撮像画像 64 に基づいて仮想視点動画 94 を生成する。すなわち、仮想視点動画生成部 28B は、編集後視点パス情報 90 に従って選定した複数の撮像画像 64 に基づいて、編集後視点パス情報 90 (例えば、図 13 に示す編集後の視点パス P1 を特定可能な複数の視点情報 74) により特定される視点から被写体を観察した場合の被写体の態様を示す動画である仮想視点動画 94 を生成する。

- [0146] 例えば、仮想視点動画生成部 28B は、図 14 に示す編集後の視点パス P1 に従って複数フレームの仮想視点画像 92 を生成する。すなわち、仮想視点動画生成部 28B は、編集後の視点パス P1 上の視点毎に仮想視点画像 92 を生成する。仮想視点動画生成部 28B は、複数フレームの仮想視点画像 92 を時系列で並べることで仮想視点動画 94 を生成する。このようにして生成された仮想視点動画 94 は、ユーザデバイス 12 のタッチパネルディスプレイ 16 に対して表示させるためのデータである。仮想視点動画 94 がタッチパネルディスプレイ 16 に表示される時間は、編集後視点パス情報 90 に含まれる複数の視点情報 74（例えば、図 13 に示す編集後の視点パス P1 を示す複数の視点情報 74）に応じて定められる。
- [0147] 仮想視点動画生成部 28B は、仮想視点動画 94 に含まれる複数フレームの仮想視点画像 92 の各々に対してメタデータ 92A を付与する。メタデータ 92A は、例えば、仮想視点画像 92 の生成に用いられた撮像画像 64 に含まれる撮像条件情報 64A（図 4 参照）に基づいて仮想視点動画生成部 28B によって生成される。メタデータ 92A には、仮想視点画像 92 が生成された時刻、及び、撮像条件情報 64A に基づく情報が含まれる。
- [0148] 仮想視点動画生成部 28B は、仮想視点動画 94 を生成する毎に、仮想視点動画 94 に対して動画識別情報 96 を付与する。動画識別情報 96 は、仮想視点動画 94 に対して固有に割り当てられた識別子を含み、仮想視点動画 94 の特定に用いられる。また、動画識別情報 96 には、仮想視点動画 94 が生成された時刻、及び／又は、仮想視点動画 94 の総再生時間等のメタデータが含まれる。
- [0149] 仮想視点動画生成部 28B は、生成した仮想視点動画 94 をストレージ 30 に記憶させる。ストレージ 30 には、例えば、編集後の視点パス P1 を含めた複数の視点パスについて仮想視点動画生成部 28B によって生成された仮想視点動画 94 が記憶される。
- [0150] 一例として図 16 に示すように、編集結果取得部 28H は、編集前視点パス情報 88 及び編集後視点パス情報 90 を参照して、視点パス P1 を編集し

た結果である編集結果98を取得する。編集結果98の第1例としては、視点パスP1を編集した部分（以下、「編集部分」とも称する）が挙げられる。編集部分は、例えば、編集後視点パス情報90に含まれる複数の視点位置情報74Aのうち、編集前視点パス情報88に含まれる複数の視点位置情報74Aと一致していない少なくとも1つの視点位置情報74Aから特定される。編集結果98の第2例としては、視点パスP1に対する編集の頻度が既定頻度（例えば、3回）よりも高い部分（以下、「編集高頻度部分」とも称する）が挙げられる。編集高頻度部分は、例えば、編集後視点パス情報90に含まれる複数の視点位置情報74Aのうち、編集の頻度が既定頻度を上回る少なくとも1つの視点位置情報74Aから特定される。編集結果98の第3例としては、編集後の視点パスP1のうち、編集前の視点パスP1からの差分が大きい部分（以下、「差分部分」とも称する）が挙げられる。差分部分は、例えば、編集後視点パス情報90に含まれる複数の視点位置情報74Aのうち、編集前視点パス情報88に含まれる複数の視点位置情報74Aからの距離が既定距離（例えば、俯瞰映像72内での数十ピクセル）以上の少なくとも1つの視点位置情報74Aから特定される。

[0151] 一例として図17に示すように、取得部28Cは、編集結果取得部28Hから編集結果98を取得する。取得部28Cは、ストレージ30に記憶されている仮想視点動画94から特定区間仮想視点動画94Aを取得する。特定区間仮想視点動画94Aは、仮想視点動画94のうち、取得部28Cによって取得された編集結果98から特定された時間帯（例えば、編集部分、編集高頻度部分、又は差分部分）の仮想視点動画である。

[0152] 抽出部28Dは、仮想視点動画94に含まれる時間（図17に示す例では、編集結果98から特定された時間帯）に基づいて決められた対象被写体100を特定する。ここで、対象被写体100は、本開示の技術に係る「第1被写体」の一例である。

[0153] 仮想視点動画94に含まれる時間としては、例えば、被写体が写っている時間の長さ、仮想視点動画94の総再生時間のうちの最初及び／又は最後の

時間帯（例えば、数秒間）又は時点が挙げられる。

- [0154] 本第2実施形態において、抽出部28Dは、取得部28Cによって取得された特定区間仮想視点動画94Aに含まれる全ての仮想視点画像92に対してA1方式の被写体認識処理を行うことで、特定区間仮想視点動画94Aに最も長く写っている被写体を対象被写体100として特定する。そして、抽出部28Dは、特定した対象被写体100が含まれている複数フレームの仮想視点画像92を特定区間仮想視点動画94Aから抽出する。
- [0155] なお、ここでは、A1方式の被写体認識処理が行われる形態例を挙げて説明しているが、これは、あくまでも一例に過ぎず、テンプレートマッチング方式の被写体認識処理が行われてもよい。また、仮想視点動画94に含まれる全ての仮想視点画像92に含まれている被写体に対して、被写体を特定する識別子（以下、「被写体識別子」と称する）が事前に付与されており、抽出部28Dが、被写体識別子を参照して各仮想視点画像92に含まれている被写体を特定するようにしてよい。
- [0156] 選定部28Eは、抽出部28Dによって抽出された複数フレームの仮想視点画像92での対象被写体100のサイズに基づいて決められた1つのフレームの仮想視点画像92を選定する。例えば、選定部28Eは、抽出部28Dによって抽出された複数フレームの仮想視点画像92から、最大サイズの対象被写体100が含まれている1フレームの仮想視点画像92を選定する。例えば、抽出部28DによってA1方式の被写体認識処理が行われた場合、選定部28Eは、A1方式の被写体認識処理で用いられるバウンディングボックスのサイズを参照することで最大サイズの対象被写体100が含まれている仮想視点画像92を特定する。
- [0157] ここで、抽出部28Dによって抽出された複数フレームは、本開示の技術に係る「仮想視点動画において撮像領域内の第1被写体が含まれている複数のフレーム」の一例である。また、最大サイズの対象被写体100が含まれている1フレームの仮想視点画像92は、本開示の技術に係る「第1フレームに関する画像」の一例である。また、「最大サイズ」は、本開示の技術に

係る「第1被写体のサイズ」の一例である。

[0158] なお、ここでは、最大サイズの対象被写体100を例示しているが、これは、あくまでも一例に過ぎず、最大サイズ以外の指定されたサイズ（例えば、最大サイズの次に大きなサイズ）の対象被写体100であってもよいし、予め定められたサイズ範囲（例えば、受付デバイス50等によって受け付けられた指示に応じて定められたサイズ範囲）内で最大サイズの対象被写体100であってもよいし、受付デバイス50等によって受け付けられた指示に応じて定められたサイズの対象被写体100であってもよい。

[0159] 加工部28Fは、仮想視点動画94を、仮想視点動画94とは異なる大きさの画像に加工する。仮想視点動画94とは異なる大きさの画像としては、例えば、仮想視点動画94よりもデータ量が少ない画像（例えば、少なくとも1フレーム分の画像）、仮想視点動画94が間引かれた画像（例えば、コマ送り画像）、仮想視点動画94に含まれる少なくとも1フレーム分の仮想視点画像92の表示サイズを縮小した画像、及び／又は、仮想視点動画94に含まれる少なくとも1フレーム分の仮想視点画像92内の画素を間引いて得た画像等が挙げられる。

[0160] 加工部28Fは、仮想視点動画94に含まれる全ての仮想視点画像92のうち1フレームの仮想視点画像92に関する画像を生成する。1フレームの仮想視点画像92に関する画像は、例えば、仮想視点動画94の内容を示す画像である。ここで、1フレームの仮想視点画像92に関する画像は、本開示の技術に係る「第1フレームに関する画像」の一例である。1フレームの仮想視点画像92に関する画像としては、例えば、1フレームの仮想視点画像92の全体そのもの、1フレームの仮想視点画像92から切り出した一部、及び／又は、1フレームの仮想視点画像92が加工された画像が挙げられる。

[0161] 加工部28Fは、複数の撮像画像64及び複数の視点情報74に基づいて、仮想視点動画94に対応するサムネイル画像102を取得する。本第2実施形態において、加工部28Fは、複数の視点情報74を編集した結果に相

当する編集結果 98 に基づいてサムネイル画像 102 を取得する。サムネイル画像 102 は、本開示の技術に係る「代表画像」の一例である。すなわち、加工部 28F は、仮想視点動画 94 に含まれる全ての仮想視点画像 92 のうちの代表的な 1 フレームの仮想視点画像 92 をサムネイル化する。

[0162] 例えば、加工部 28F は、選定部 28E によって選定された仮想視点画像 92 をサムネイル画像 102 に加工する。仮想視点画像 92 をサムネイル画像 102 に加工する方法としては、上記の仮想視点動画 94 を、仮想視点動画 94 とは異なる大きさの画像に加工する方法を用いることができる。また、加工部 28F は、サムネイル化される前の仮想視点画像 92 に付与されているメタデータ 92A をサムネイル画像 102 に対して関連付ける。また、加工部 28F は、サムネイル化された仮想視点画像 92 が含まれる仮想視点動画 94 から動画識別情報 96 を取得する。

[0163] 一例として図 18 に示すように、加工部 28F は、仮想視点画像 92 をサムネイル化して得たサムネイル画像 102 に対して動画識別情報 96 を関連付ける。

[0164] 一覧画面生成部 28G は、メタデータ 92A 及び動画識別情報 96 が関連付けられたサムネイル画像 102 を加工部 28F から取得する。一覧画面生成部 28G は、メタデータ 92A 及び／又は動画識別情報 96 に基づいて参照情報 104A を生成し、サムネイル画像 102 に対応付ける。一覧画面生成部 28G は、参照情報 104A が対応付けられたサムネイル画像 102 を含めた一覧画面 104 を示す一覧画面データ 106 を生成する。一覧画面データ 106 は、サムネイル画像 102 をユーザデバイス 12 のタッチパネルディスプレイ 16 に表示させるためのデータである。一覧画面生成部 28G は、生成した一覧画面データ 106 を送受信装置 24 に出力し、かつ、ストレージ 30 に記憶させる。これにより、ストレージ 30 には、動画識別情報 96 と関連付けられたサムネイル画像 102 が記憶される。すなわち、動画識別情報 96 は、仮想視点動画 94 に対して固有に割り当てられた識別子なので、ストレージ 30 には、サムネイル画像 102 と仮想視点動画 94 とが

関連付けられた状態で記憶される。一覧画面データ106は、本開示の技術に係る「データ」及び「第1データ」の一例である。

[0165] 一覧画面生成部28Gによってサムネイル画像102に対応付けられる参照情報104Aとしては、例えば、文字情報が挙げられる。文字情報としては、例えば、仮想視点動画94が生成された時刻（例えば、図4に示す撮像条件情報64Aから特定される時刻）、サムネイル画像102に含まれる対象被写体100に関連する情報（例えば、対象被写体100の名称、及び／又は、対象被写体100の所属チーム等）、仮想視点動画94の総再生時間、仮想視点動画94のタイトル、及び／又は、仮想視点動画94の作成者の名称等が挙げられる。

[0166] ストレージ30に一覧画面データ106が記憶されている状態で、加工部28Fによって、サムネイル画像102が生成され、かつ、生成されたサムネイル画像102に対してメタデータ92A及び動画識別情報96が関連付けられると、一覧画面生成部28Gは、ストレージ30から一覧画面データ106を取得し、一覧画面データ106を更新する。すなわち、一覧画面生成部28Gは、加工部28Fから、メタデータ92A及び動画識別情報96が関連付けられたサムネイル画像102を取得し、参照情報104Aを生成する。一覧画面生成部28Gは、生成した参照情報104Aをサムネイル画像102に対応付ける。そして、一覧画面生成部28Gは、参照情報104Aが対応付けられたサムネイル画像102を一覧画面104に含めることで一覧画面データ106を更新する。一覧画面生成部28Gは、更新した一覧画面データ106を送受信装置24に出力し、かつ、ストレージ30に記憶させる。

[0167] 更新された一覧画面データ106により示される一覧画面104には、複数のサムネイル画像102が含まれる。また、更新された一覧画面データ106により示される一覧画面104において、複数のサムネイル画像102の各々には参照情報104Aが対応付けられている。

[0168] 送受信装置24は、一覧画面生成部28Gから入力された一覧画面データ

106をユーザデバイス12に送信する。ユーザデバイス12において、送受信装置44は、画像処理装置10から送信された一覧画面データ106を受信する。プロセッサ52は、送受信装置44によって受信された一覧画面データ106を取得し、取得した一覧画面データ106により示される一覧画面104をタッチパネルディスプレイ16に表示させる。一覧画面104には、複数の画像が並行して表示される。図18に示す例では、一覧画面104に、複数のサムネイル画像102が参照情報104Aと共に表示されている。なお、ここでは、複数のサムネイル画像102が一覧画面104に表示される形態例を挙げているが、一覧画面104にサムネイル画像102が1つだけ表示されていてもよい。また、複数のサムネイル画像102は、必ずしも並行して表示される必要はない。

[0169] タッチパネルディスプレイ16に一覧画面104が表示されている状態で、ユーザ14は、タッチパネルディスプレイ16を介して一覧画面104内の何れかのサムネイル画像102をタップすることでサムネイル画像102を選択する。サムネイル画像102が選択されると、これに応じて、画像処理装置10のプロセッサ28（図1及び図12参照）は、仮想視点動画94をタッチパネルディスプレイ16に対して表示させるためのデータをユーザデバイス12に出力する。

[0170] 例えば、ユーザ14によってタッチパネルディスプレイ16を介してサムネイル画像102が選択されると、ユーザデバイス12のプロセッサ52は、選択されたサムネイル画像102に関連付けられている動画識別情報96を、送受信装置44を介して画像処理装置10に送信する。画像処理装置10では、送受信装置24によって動画識別情報96が受信される。画像処理装置10のプロセッサ28は（図1及び図12参照）は、送受信装置24によって受信された動画識別情報96に対応する仮想視点動画94をストレージ30から取得し、取得した仮想視点動画94を、送受信装置24を介してユーザデバイス12に送信する。ユーザデバイス12では、画像処理装置10から送信された仮想視点動画94が送受信装置44によって受信される。

ユーザデバイス 12 のプロセッサ 52 は、送受信装置 44 によって受信された仮想視点動画 94 をタッチパネルディスプレイ 16 に表示させる。例えば、仮想視点動画 94 は、タッチパネルディスプレイ 16 の仮想視点動画画面 68 (図 4 参照) に表示される。

[0171] なお、ここでは、仮想視点動画 94 がタッチパネルディスプレイ 16 に表示される形態例を挙げているが、これは、あくまでも一例に過ぎず、例えば、タッチパネルディスプレイ 16 に代えて、又は、タッチパネルディスプレイ 16 と共に、画像処理装置 10 に直接的又は間接的に接続されているディスプレイに対して仮想視点動画 94 が表示されるようにしてもよい。

[0172] また、ここでは、一覧画面 104 内の何れかのサムネイル画像 102 がタップされることでサムネイル画像 102 が選択される形態例を挙げて説明したが、これは、あくまでも一例に過ぎず、タッチパネルディスプレイ 16 を介してサムネイル画像 102 に対してフリック、スワイプ、及び／又はロングプレス等が行われることによってサムネイル画像 102 が選択されるようにしてもよいし、マイクロフォン 48 によって取得された音声に対する音声認識処理が行われることによってサムネイル画像 102 が選択されるようにしてもよいし、マウス及び／又はキーボード等によってサムネイル画像 102 が選択されるようにしてもよい。

[0173] 以上説明したように、本第 2 実施形態に係る画像処理装置 10 では、視点パス P1 に対して行われた編集に伴って得られた編集結果 98 に基づいてサムネイル画像 102 が取得される。すなわち、仮想視点動画 94 に含まれる複数の仮想視点画像 92 から編集結果 98 に基づいて特定された仮想視点画像 92 に対応するサムネイル画像 102 が取得される。そして、画像処理装置 10 によって取得されたサムネイル画像 102 を含む一覧画面 104 がユーザデバイス 12 のタッチパネルディスプレイ 16 に表示される。従って、本構成によれば、編集結果 98 に基づいて得られたサムネイル画像 102 をユーザ 14 に見せることに寄与することができる。

[0174] なお、上記第 2 実施形態では、編集結果 98 の一例として、視点パス P1

のみを編集した結果を示したが、本開示の技術は、これに限定されない。編集結果98には、視点パスP1のみならず、複数の仮想視点動画を示す複数の視点パスに対して行われた編集の結果が含まれていてもよい。この場合、複数の視点情報74は、複数の視点パスを有する。すなわち、複数の視点パスは、複数の視点情報74によって規定される。そして、プロセッサ28は、複数の視点パスのうちの少なくとも1つの視点パスに対して行われた編集の結果に基づいて少なくとも1つの仮想視点画像（すなわち、少なくとも1つの仮想視点動画から得られた少なくとも1つの仮想視点画像）を特定する。プロセッサ28は、特定した少なくとも1つの仮想視点画像に対応する少なくとも1つのサムネイル画像を生成し、生成したサムネイル画像を含む一覧画面104を生成する。これにより、複数の視点パスに対して行われた編集の結果に基づいて得られた少なくとも1つの仮想視点画像に対応する少なくとも1つのサムネイル画像をユーザ14に見せることに寄与することができる。

[0175] [第3実施形態]

本第3実施形態では、上記第1及び第2実施形態で説明した構成要素については、同一の符号を付し、説明を省略し、上記第1及び第2実施形態と異なる部分について説明する。

[0176] 一例として図19に示すように、本第3実施形態に係る画像処理装置10のプロセッサ28は、図12に示すプロセッサ28に比べ、画面生成処理プログラム38を実行することで相違度算出部281として更に動作する点が異なる。

[0177] 本第3実施形態では、説明の便宜上、一例として図20に示すように、複数の視点パスから、処理対象の視点パスとして、互いに異なる位置に存在する第1視点パス108及び第2視点パス110が、タッチパネルディスプレイ16を介してユーザ14によって指定されたことを前提として説明する。

[0178] 一例として図21に示すように、ユーザデバイス12において、プロセッサ52は、第1視点パス108（図20参照及び第1注視点（例えば、図6

に示す注視点G P)に基づいて第1視点パス情報112を生成する。第1視点パス情報112には、上記第1及び第2実施形態で説明した複数の視点情報74が含まれている。また、プロセッサ52は、第2視点パス110(図20参照)及び第2注視点(例えば、図6に示す注視点G P)に基づいて第2視点パス情報114を生成する。第2視点パス110には、上記第1及び第2実施形態で説明した複数の視点情報74が含まれている。第1視点パス情報112に含まれる複数の視点情報74は、第1視点パス108の特徴を示しており、第2視点パス情報114に含まれる複数の視点情報74は、第2視点パス110の特徴を示している。従って、第1視点パス情報112に含まれる複数の視点情報74と第2視点パス情報114に含まれる複数の視点情報74とは内容が異なっている。

[0179] 一例として図22に示すように、ユーザデバイス12のプロセッサ52は、第1視点パス情報112及び第2視点パス情報114を、送受信装置44を介して画像処理装置10に送信する。画像処理装置10において、送受信装置24は、画像処理装置10から送信された第1視点パス情報112及び第2視点パス情報114を受信する。仮想視点動画生成部28B及び相違度算出部281は、送受信装置24によって受信された第1視点パス情報112及び第2視点パス情報114を取得する。

[0180] 一例として図23に示すように、仮想視点動画生成部28Bは、第1視点パス情報112(図21及び図22参照)に従って、仮想視点画像116の生成に用いる複数の撮像画像64(図4参照)を選定する。すなわち、仮想視点動画生成部28Bは、第1視点パス情報112に従って被写体を観察した場合の被写体の態様を示す画像である仮想視点画像116の生成に用いる複数の撮像画像64(図4参照)を、複数の撮像装置36(図1及び図4参照)によって撮像されることで得られた複数の撮像画像64(図4参照)から選定する。

[0181] 仮想視点動画生成部28Bは、第1視点パス情報112及び複数の撮像画像64に基づいて第1仮想視点動画118を生成する。すなわち、仮想視点

動画生成部 28B は、第 1 視点パス情報 112 に従って選定した複数の撮像画像 64 に基づいて、第 1 視点パス情報 112 により特定される視点から被写体を観察した場合の被写体の態様を示す動画である第 1 仮想視点動画 118 を生成する。

[0182] 例えば、仮想視点動画生成部 28B は、第 1 視点パス 108 (図 20 参照) に従って複数フレームの仮想視点画像 116 を生成する。すなわち、仮想視点動画生成部 28B は、第 1 視点パス 108 上の視点毎に仮想視点画像 116 を生成する。仮想視点動画生成部 28B は、複数フレームの仮想視点画像 116 を時系列で並べることで第 1 仮想視点動画 118 を生成する。このようにして生成された第 1 仮想視点動画 118 は、ユーザデバイス 12 のタッチパネルディスプレイ 16 に対して表示させるためのデータである。第 1 仮想視点動画 118 がタッチパネルディスプレイ 16 に表示される時間は、第 1 視点パス情報 112 に含まれる複数の視点情報 74 (図 21 参照) に応じて定められる。

[0183] 仮想視点動画生成部 28B は、第 1 仮想視点動画 118 に含まれる複数フレームの仮想視点画像 116 の各々に対して第 1 メタデータ (図示省略) を付与する。第 1 仮想視点動画 118 に含まれる複数フレームの仮想視点画像 116 の各々に対して付与される第 1 メタデータの技術的意義は、上記第 1 実施形態で説明したメタデータ 76A 及び上記第 2 実施形態で説明したメタデータ 92A と同じである。

[0184] 仮想視点動画生成部 28B は、第 1 仮想視点動画 118 を生成する毎に、第 1 仮想視点動画 118 に対して第 1 動画識別情報 120 を付与する。第 1 動画識別情報 120 は、第 1 仮想視点動画 118 に対して固有に割り当てられた識別子を含み、第 1 仮想視点動画 118 の特定に用いられる。また、第 1 動画識別情報 120 には、第 1 仮想視点動画 118 が生成された時刻、及び／又は、第 1 仮想視点動画 118 の総再生時間等のメタデータが含まれる。

[0185] 仮想視点動画生成部 28B は、第 2 視点パス情報 114 (図 21 及び図 2

2参照)に従って、仮想視点画像122の生成に用いる複数の撮像画像64(図4参照)を選定する。すなわち、仮想視点動画生成部28Bは、第2視点パス情報114に従って被写体を観察した場合の被写体の態様を示す画像である仮想視点画像122の生成に用いる複数の撮像画像64(図4参照)を、複数の撮像装置36(図1及び図4参照)によって撮像されることで得られた複数の撮像画像64(図4参照)から選定する。

[0186] 仮想視点動画生成部28Bは、第2視点パス情報114及び複数の撮像画像64に基づいて第2仮想視点動画124を生成する。すなわち、仮想視点動画生成部28Bは、第2視点パス情報114に従って選定した複数の撮像画像64に基づいて、第2視点パス情報114により特定される視点から被写体を観察した場合の被写体の態様を示す動画である第2仮想視点動画124を生成する。

[0187] 例えば、仮想視点動画生成部28Bは、第2視点パス110(図20参照)に従って複数フレームの仮想視点画像122を生成する。すなわち、仮想視点動画生成部28Bは、第2視点パス110上の視点毎に仮想視点画像122を生成する。仮想視点動画生成部28Bは、複数フレームの仮想視点画像122を時系列で並べることで第2仮想視点動画124を生成する。このようにして生成された第2仮想視点動画124は、ユーザデバイス12のタッチパネルディスプレイ16に対して表示させるためのデータである。第2仮想視点動画124がタッチパネルディスプレイ16に表示される時間は、第2視点パス情報114に含まれる複数の視点情報74(図21参照)に応じて定められる。

[0188] 仮想視点動画生成部28Bは、第2仮想視点動画124に含まれる複数フレームの仮想視点画像122の各々に対して第2メタデータ(図示省略)を付与する。第2仮想視点動画124に含まれる複数フレームの仮想視点画像122の各々に対して付与される第2メタデータの技術的意義は、上記第1実施形態で説明したメタデータ76A及び上記第2実施形態で説明したメタデータ92Aと同じである。

- [0189] 仮想視点動画生成部28Bは、第2仮想視点動画124を生成する毎に、第2仮想視点動画124に対して第2動画識別情報126を付与する。第2動画識別情報126は、第2仮想視点動画124に対して固有に割り当てられた識別子を含み、第2仮想視点動画124の特定に用いられる。また、第2動画識別情報126には、第2仮想視点動画124が生成された時刻、及び／又は、第2仮想視点動画124の総再生時間等のメタデータが含まれる。
- [0190] 一例として図24に示すように、仮想視点動画生成部28Bは、生成した第1仮想視点動画118をストレージ30に記憶させる。また、仮想視点動画生成部28Bは、生成した第2仮想視点動画124もストレージ30に記憶させる。
- [0191] 一例として図25に示すように、相違度算出部281は、第1視点パス情報112と第2視点パス情報114との間の相違度128を算出する。相違度128は、第1視点パス情報112に含まれる複数の視点情報74と第2視点パス情報114に含まれる複数の視点情報74との間の相違度とも言える。相違度128の一例としては、第1視点パス108の分割エリア108Aと第2視点パス110の分割エリア110Aとのずれ量が挙げられる。相違度128は、本開示の技術に係る「相違度」の一例である。
- [0192] 分割エリア108Aは、第1視点パス108の始点から終点までがN等分されることで得られたエリアである。分割エリア110Aは、第2視点パス110の始点から終点までがN等分されることで得られたエリアである。ここで、“N”は、2以上の自然数であり、例えば、受付デバイス50等によって受け付けられた指示に応じて定められる。“N”は、固定値であってもよいし、受付デバイス50によって受け付けられた指示、及び／又は、各種情報（例えば、撮像条件）に応じて変更される可変値であってもよい。
- [0193] 本第3実施形態では、相違度算出部281は、第1視点パス108の始点から終点に至るまでの複数の分割エリア108Aと第2視点パス110の始点から終点に至るまでの複数の分割エリア110Aとの間の分割エリア間で

のずれ量を相違度128として算出する。すなわち、相違度128は、第1視点パス108の複数の分割エリア108Aと第2視点パス110の複数の分割エリア110Aとの対応する分割エリア間でのずれ量を、始点から終点にかけて各分割エリア108A及び各分割エリア110Aに対応付けた情報である。

[0194] 一例として図26に示すように、取得部28Cは、相違度算出部28Iから相違度128を取得する。取得部28Cは、ストレージ30に記憶されている第1仮想視点動画118から第1特定区間仮想視点動画118Aを取得する。第1特定区間仮想視点動画118Aは、第1仮想視点動画118のうち、取得部28Cによって取得された相違度128から特定された時間帯の仮想視点動画である。ここで、相違度128から特定された時間帯とは、例えば、相違度128によって表される複数のずれ量のうちの最大のずれ量に対応付けられている分割エリア108A（図25参照）に対応する時間帯を指す。ここでは、最大のずれ量を例示しているが、最小のずれ量であってもよいし、中央値のずれ量であってもよいし、最頻値のずれ量であってもよい。

[0195] 抽出部28Dは、第1仮想視点動画118に含まれる時間（図26に示す例では、相違度128に応じて定められた時間帯）に基づいて決められた対象被写体130を特定する。ここで、対象被写体130は、本開示の技術に係る「第1被写体」の一例である。

[0196] 第1仮想視点動画118に含まれる時間としては、例えば、被写体が写っている時間の長さ、第1仮想視点動画118の総再生時間のうちの最初及び／又は最後の時間帯（例えば、数秒間）又は時点が挙げられる。

[0197] 本第3実施形態において、抽出部28Dは、取得部28Cによって取得された第1特定区間仮想視点動画118Aに含まれる全ての仮想視点画像116に対してA1方式の被写体認識処理を行うことで、第1特定区間仮想視点動画118Aに最も長く写っている被写体を対象被写体130として特定する。そして、抽出部28Dは、特定した対象被写体130が含まれている複

数フレームの仮想視点画像 116 を第 1 特定区間仮想視点動画 118 A から抽出する。

[0198] なお、ここでは、A1方式の被写体認識処理が行われる形態例を挙げて説明しているが、これは、あくまでも一例に過ぎず、テンプレートマッチング方式の被写体認識処理が行われてもよい。また、第 1 仮想視点動画 118 に含まれる全ての仮想視点画像 116 に含まれている被写体に対して、被写体を特定する識別子（以下、「被写体識別子」と称する）が事前に付与されており、抽出部 28 D が、被写体識別子を参照して各仮想視点画像 116 に含まれている被写体を特定するようにしてよい。

[0199] 選定部 28 E は、抽出部 28 D によって抽出された複数フレームの仮想視点画像 116 での対象被写体 130 のサイズに基づいて決められた 1 つのフレームの仮想視点画像 116 を選定する。例えば、選定部 28 E は、抽出部 28 D によって抽出された複数フレームの仮想視点画像 116 から、最大サイズの対象被写体 130 が含まれている 1 フレームの仮想視点画像 116 を選定する。例えば、抽出部 28 D によって A1方式の被写体認識処理が行われた場合、選定部 28 E は、A1方式の被写体認識処理で用いられるバウンディングボックスのサイズを参照することで最大サイズの対象被写体 130 が含まれている仮想視点画像 116 を特定する。

[0200] ここで、抽出部 28 D によって抽出された複数フレームは、本開示の技術に係る「仮想視点動画において撮像領域内の第 1 被写体が含まれている複数のフレーム」の一例である。また、最大サイズの対象被写体 130 が含まれている 1 フレームの仮想視点画像 116 は、本開示の技術に係る「第 1 フレームに関する画像」の一例である。また、「最大サイズ」は、本開示の技術に係る「第 1 被写体のサイズ」の一例である。

[0201] なお、ここでは、最大サイズの対象被写体 130 を例示しているが、これは、あくまでも一例に過ぎず、最大サイズ以外の指定されたサイズ（例えば、最大サイズの次に大きなサイズ）の対象被写体 130 であってもよいし、予め定められたサイズ範囲（例えば、受付デバイス 50 等によって受け付け

られた指示に応じて定められたサイズ範囲) 内で最大サイズの対象被写体 130 であってもよいし、受付デバイス 50 等によって受け付けられた指示に応じて定められたサイズの対象被写体 130 であってもよい。

[0202] 加工部 28F は、第 1 仮想視点動画 118 を、第 1 仮想視点動画 118 とは異なる大きさの画像に加工する。第 1 仮想視点動画 118 とは異なる大きさの画像としては、例えば、第 1 仮想視点動画 118 よりもデータ量が少ない画像 (例えば、少なくとも 1 フレーム分の画像)、第 1 仮想視点動画 118 が間引かれた画像 (例えば、コマ送り画像)、第 1 仮想視点動画 118 に含まれる少なくとも 1 フレーム分の仮想視点画像 116 の表示サイズを縮小した画像、及び/又は、第 1 仮想視点動画 118 に含まれる少なくとも 1 フレーム分の仮想視点画像 116 内の画素を間引いて得た画像等が挙げられる。

[0203] 加工部 28F は、第 1 仮想視点動画 118 に含まれる全ての仮想視点画像 116 のうちの 1 フレームの仮想視点画像 116 に関する画像を生成する。1 フレームの仮想視点画像 116 に関する画像は、例えば、第 1 仮想視点動画 118 の内容を示す画像である。ここで、1 フレームの仮想視点画像 116 に関する画像は、本開示の技術に係る「第 1 フレームに関する画像」の一例である。1 フレームの仮想視点画像 116 に関する画像としては、例えば、1 フレームの仮想視点画像 116 の全体そのもの、1 フレームの仮想視点画像 116 から切り出した一部、及び/又は、1 フレームの仮想視点画像 116 が加工された画像が挙げられる。

[0204] 加工部 28F は、複数の撮像画像 64 及び複数の視点情報 74 に基づいて、第 1 仮想視点動画 118 に対応するサムネイル画像 132 を取得する。本第 3 実施形態において、加工部 28F は、複数の視点情報 74 間 (ここでは、一例として、第 1 視点パス情報 112 と第 2 視点パス情報 114 との間) の相違度 128 に基づいてサムネイル画像 132 を取得する。サムネイル画像 132 は、本開示の技術に係る「代表画像」の一例である。すなわち、加工部 28F は、第 1 仮想視点動画 118 に含まれる全ての仮想視点画像 11

6のうちの代表的な1フレームの仮想視点画像116をサムネイル化する。

[0205] 例えば、加工部28Fは、選定部28Eによって選定された仮想視点画像116をサムネイル画像132に加工する。仮想視点画像116をサムネイル画像132に加工する方法としては、上記の第1仮想視点動画118を、第1仮想視点動画118とは異なる大きさの画像に加工する方法を用いることができる。また、加工部28Fは、サムネイル化される前の仮想視点画像116に付与されている第1メタデータ（図示省略）をサムネイル画像132に対して関連付ける。また、加工部28Fは、サムネイル化された仮想視点画像116が含まれる第1仮想視点動画118から第1動画識別情報120を取得する。

[0206] なお、このように、加工部28Fによって取得されたサムネイル画像132、サムネイル画像132に関連付けられた第1メタデータ、及び加工部28Fによって取得された第1動画識別情報120に対してプロセッサ28によって行われる処理は、例えば、上記第2実施形態で説明したサムネイル画像102、メタデータ92A、及び動画識別情報96に対してプロセッサ28によって行われる処理と同じである（図18参照）。

[0207] 一例として図27に示すように、取得部28Cは、相違度算出部28Iから相違度128を取得する。取得部28Cは、ストレージ30に記憶されている第2仮想視点動画124から第2特定区間仮想視点動画124Aを取得する。第2特定区間仮想視点動画124Aは、第2仮想視点動画124のうち、取得部28Cによって取得された相違度128から特定された時間帯の仮想視点動画である。ここで、相違度128から特定された時間帯とは、例えば、相違度128によって表される複数のずれ量のうちの最大のずれ量に対応付けられている分割エリア110A（図25参照）に対応する時間帯を指す。ここでは、最大のずれ量を例示しているが、最小のずれ量であってもよいし、中央値のずれ量であってもよいし、最頻値のずれ量であってもよい。

[0208] 抽出部28Dは、第2仮想視点動画124に含まれる時間（図27に示す

例では、相違度 1 2 8 に応じて定められた時間帯) に基づいて決められた対象被写体 1 3 4 を特定する。ここで、対象被写体 1 3 4 は、本開示の技術に係る「第 1 被写体」の一例である。

[0209] 第 2 仮想視点動画 1 2 4 に含まれる時間としては、例えば、被写体が写っている時間の長さ、第 2 仮想視点動画 1 2 4 の総再生時間のうちの最初及び／又は最後の時間帯（例えば、数秒間）又は時点が挙げられる。

[0210] 本第 3 実施形態において、抽出部 2 8 D は、取得部 2 8 C によって取得された第 2 特定区間仮想視点動画 1 2 4 A に含まれる全ての仮想視点画像 1 2 2 に対して A 1 方式の被写体認識処理を行うことで、第 2 特定区間仮想視点動画 1 2 4 A に最も長く写っている被写体を対象被写体 1 3 4 として特定する。そして、抽出部 2 8 D は、特定した対象被写体 1 3 4 が含まれている複数フレームの仮想視点画像 1 2 2 を第 2 特定区間仮想視点動画 1 2 4 A から抽出する。

[0211] なお、ここでは、A 1 方式の被写体認識処理が行われる形態例を挙げて説明しているが、これは、あくまでも一例に過ぎず、テンプレートマッチング方式の被写体認識処理が行われてもよい。また、第 2 仮想視点動画 1 2 4 に含まれる全ての仮想視点画像 1 2 2 に含まれている被写体に対して、被写体を特定する識別子（以下、「被写体識別子」と称する）が事前に付与されており、抽出部 2 8 D が、被写体識別子を参照して各仮想視点画像 1 2 2 に含まれている被写体を特定するようにしてよい。

[0212] 選定部 2 8 E は、抽出部 2 8 D によって抽出された複数フレームの仮想視点画像 1 2 2 での対象被写体 1 3 4 のサイズに基づいて決められた 1 つのフレームの仮想視点画像 1 2 2 を選定する。例えば、選定部 2 8 E は、抽出部 2 8 D によって抽出された複数フレームの仮想視点画像 1 2 2 から、最大サイズの対象被写体 1 3 4 が含まれている 1 フレームの仮想視点画像 1 2 2 を選定する。例えば、抽出部 2 8 D によって A 1 方式の被写体認識処理が行われた場合、選定部 2 8 E は、A 1 方式の被写体認識処理で用いられるバウンディングボックスのサイズを参照することで最大サイズの対象被写体 1 3 4

が含まれている仮想視点画像 1 2 2 を特定する。

[0213] ここで、抽出部 2 8 D によって抽出された複数フレームは、本開示の技術に係る「仮想視点動画において撮像領域内の第 1 被写体が含まれている複数のフレーム」の一例である。また、最大サイズの対象被写体 1 3 4 が含まれている 1 フレームの仮想視点画像 1 2 2 は、本開示の技術に係る「第 1 フレームに関する画像」の一例である。また、「最大サイズ」は、本開示の技術に係る「第 1 被写体のサイズ」の一例である。

[0214] なお、ここでは、最大サイズの対象被写体 1 3 4 を例示しているが、これは、あくまでも一例に過ぎず、最大サイズ以外の指定されたサイズ（例えば、最大サイズの次に大きなサイズ）の対象被写体 1 3 4 であってもよいし、予め定められたサイズ範囲（例えば、受付デバイス 5 0 等によって受け付けられた指示に応じて定められたサイズ範囲）内で最大サイズの対象被写体 1 3 4 であってもよいし、受付デバイス 5 0 等によって受け付けられた指示に応じて定められたサイズの対象被写体 1 3 4 であってもよい。

[0215] 加工部 2 8 F は、第 2 仮想視点動画 1 2 4 を、第 2 仮想視点動画 1 2 4 とは異なる大きさの画像に加工する。第 2 仮想視点動画 1 2 4 とは異なる大きさの画像としては、例えば、第 2 仮想視点動画 1 2 4 よりもデータ量が少ない画像（例えば、少なくとも 1 フレーム分の画像）、第 2 仮想視点動画 1 2 4 が間引かれた画像（例えば、コマ送り画像）、第 2 仮想視点動画 1 2 4 に含まれる少なくとも 1 フレーム分の仮想視点画像 1 2 2 の表示サイズを縮小した画像、及び／又は、第 2 仮想視点動画 1 2 4 に含まれる少なくとも 1 フレーム分の仮想視点画像 1 2 2 内の画素を間引いて得た画像等が挙げられる。

[0216] 加工部 2 8 F は、第 2 仮想視点動画 1 2 4 に含まれる全ての仮想視点画像 1 2 2 のうちの 1 フレームの仮想視点画像 1 2 2 に関する画像を生成する。1 フレームの仮想視点画像 1 2 2 に関する画像は、例えば、第 2 仮想視点動画 1 2 4 の内容を示す画像である。ここで、1 フレームの仮想視点画像 1 2 2 に関する画像は、本開示の技術に係る「第 1 フレームに関する画像」の一

例である。1フレームの仮想視点画像122に関する画像としては、例えば、1フレームの仮想視点画像122の全体そのもの、1フレームの仮想視点画像122から切り出した一部、及び／又は、1フレームの仮想視点画像122が加工された画像が挙げられる。

[0217] 加工部28Fは、複数の撮像画像64及び複数の視点情報74に基づいて、第2仮想視点動画124に対応するサムネイル画像136を取得する。本第3実施形態において、加工部28Fは、複数の視点情報74間（ここでは、一例として、第1視点パス情報112と第2視点パス情報114との間）の相違度128に基づいてサムネイル画像136を取得する。サムネイル画像136は、本開示の技術に係る「代表画像」の一例である。すなわち、加工部28Fは、第2仮想視点動画124に含まれる全ての仮想視点画像122のうちの代表的な1フレームの仮想視点画像122をサムネイル化する。

[0218] 例えば、加工部28Fは、選定部28Eによって選定された仮想視点画像122をサムネイル画像136に加工する。仮想視点画像122をサムネイル画像136に加工する方法としては、上記の第2仮想視点動画124を、第2仮想視点動画124とは異なる大きさの画像に加工する方法を用いることができる。また、加工部28Fは、サムネイル化される前の仮想視点画像122に付与されている第2メタデータ（図示省略）をサムネイル画像136に対して関連付ける。また、加工部28Fは、サムネイル化された仮想視点画像122が含まれる第2仮想視点動画124から第2動画識別情報126を取得する。

[0219] なお、このように、加工部28Fによって取得されたサムネイル画像136、サムネイル画像136に関連付けられた第2メタデータ、及び加工部28Fによって取得された第2動画識別情報126に対してプロセッサ28によって行われる処理は、例えば、上記第2実施形態で説明したサムネイル画像102、メタデータ92A、及び動画識別情報96に対してプロセッサ28によって行われる処理と同じである（図18参照）。

[0220] 以上説明したように、本第3実施形態に係る画像処理装置10では、第1

視点パス108と第2視点パス110との相違度（例えば、第1視点パス情報112と第2視点パス情報114との相違度）として相違度128が算出され、算出された相違度128に基づいてサムネイル画像132が取得される。すなわち、第1仮想視点動画118に含まれる複数の仮想視点画像116から相違度128に基づいて特定された仮想視点画像116に対応するサムネイル画像132が取得される。また、本第3実施形態に係る画像処理装置10では、相違度128に基づいてサムネイル画像136が取得される。すなわち、第2仮想視点動画124に含まれる複数の仮想視点画像122から相違度128に基づいて特定された仮想視点画像122に対応するサムネイル画像136が取得される。そして、画像処理装置10によって取得されたサムネイル画像132及び136を含む一覧画面がユーザデバイス12のタッチパネルディスプレイ16に表示される。従って、本構成によれば、第1視点パス108と第2視点パス110との相違度として算出された相違度128に基づいて得られたサムネイル画像102をユーザ14に見せることに寄与することができる。

[0221] なお、上記第3実施形態では、第1視点パス108と第2視点パス110との相違度として相違度128が算出され、算出された相違度128に基づいて、サムネイル化される仮想視点画像が特定される形態例を挙げたが、本開示の技術は、これに限定されない。1つの視点に対応する1つの視点情報74と第1視点パス108又は第2視点パス110に含まれる複数の視点情報74のうちの少なくとも1つとの相違度に基づいて、サムネイル化される仮想視点画像が特定されるようにしてもよい。

[0222] また、上記第3実施形態では、第1視点パス108及び第2視点パス110という2本の視点パス間の相違度として相違度128が算出されるようにしたが、本開示の技術はこれに限定されず、3本以上の視点パス間の相違度が算出されるようにしてもよい。この場合、3本以上の視点パスのうちの少なくとも1つの視点パスに対応する仮想視点動画に含まれる少なくとも1つの仮想視点画像に対応するサムネイル画像が生成されるようにすればよい。

## [0223] [第4実施形態]

本第4実施形態では、上記第1～第3実施形態で説明した構成要素については、同一の符号を付し、説明を省略し、上記第1～第3実施形態と異なる部分について説明する。

[0224] 一例として図28に示すように、本第4実施形態に係る画像処理装置10のプロセッサ28は、図19に示すプロセッサ28に比べ、画面生成処理プログラム38を実行することで被写体位置特定部28J及び視点位置特定部28Kとして更に動作する点異なる。プロセッサ28は、仮想視点動画生成部28B、取得部28C、加工部28F、被写体位置特定部28J、及び視点位置特定部28Kとして動作することで、複数の視点パスの位置関係に基づいてサムネイル画像を取得する。位置関係とは、特定被写体138（図30参照）に対する複数の視点パスの位置関係（図31参照）を指す。ここで、特定被写体138は、本開示の技術に係る「第2被写体」の一例である。

[0225] 一例として図29に示すように、ユーザデバイス12のプロセッサ52は、第1視点パス情報112及び第2視点パス情報114を、送受信装置44を介して画像処理装置10に送信する。画像処理装置10において、送受信装置24は、送受信装置44から送信された第1視点パス情報112及び第2視点パス情報114を受信する。仮想視点動画生成部28B及び視点位置特定部28Kは、送受信装置24によって受信された第1視点パス情報112及び第2視点パス情報114を取得する。

[0226] 一例として図30に示すように、ストレージ30には、上記第3実施形態と同様に、第1仮想視点動画118及び第2仮想視点動画124が記憶されている。被写体位置特定部28Jは、ストレージ30から第1仮想視点動画118を取得する。被写体位置特定部28Jは、第1仮想視点動画118に対してA1方式の被写体認識処理を行うことで、第1仮想視点動画118に含まれる特定被写体138を認識する。ここで、特定被写体138とは、例えば、ユーザ14等によって予め指定された被写体を指す。被写体位置特定

部 28 J は、仮想視点画像 116 に含まれる特定被写体 138 の仮想視点画像 116 内での位置を特定可能な情報として、特定被写体 138 が含まれる仮想視点画像 116 内での特定被写体 138 の座標（以下、「第 1 画像内座標」とも称する）を取得する。被写体位置特定部 28 J は、第 1 画像内座標を、俯瞰映像 72（図 4 参照）内で対応する位置を特定可能な座標（以下、「第 1 俯瞰映像内座標」とも称する）を変換する。

[0227] また、被写体位置特定部 28 J は、ストレージ 30 から第 2 仮想視点動画 124 を取得する。被写体位置特定部 28 J は、第 2 仮想視点動画 124 に対して A1 方式の被写体認識処理を行うことで、第 2 仮想視点動画 124 に含まれる特定被写体 138 を認識する。被写体位置特定部 28 J は、仮想視点画像 122 に含まれる特定被写体 138 の仮想視点画像 122 内での位置を特定可能な情報として、特定被写体 138 が含まれる仮想視点画像 122 内での特定被写体 138 の座標（以下、「第 2 画像内座標」とも称する）を取得する。被写体位置特定部 28 J は、第 2 画像内座標を、俯瞰映像 72（図 4 参照）内で対応する位置を特定可能な座標（以下、「第 2 俯瞰映像内座標」とも称する）に変換する。

[0228] 一例として図 31 に示すように、視点位置特定部 28 K は、俯瞰映像 72 内での特定被写体 138 の座標として被写体位置特定部 28 J によって得られた第 1 俯瞰映像内座標を取得する。視点位置特定部 28 K は、第 1 俯瞰映像内座標及び第 1 視点パス情報 112（図 21 参照）に基づいて、第 1 視点パス 108 に含まれる複数の視点位置のうち、特定被写体 138 が最も大きく見える視点位置 108 B を特定する。そして、視点位置特定部 28 K は、特定した視点位置 108 B に対応する視点情報 74 を第 1 視点パス情報 112 から取得する。

[0229] また、視点位置特定部 28 K は、俯瞰映像 72 内での特定被写体 138 の座標として被写体位置特定部 28 J によって得られた第 2 俯瞰映像内座標を取得する。視点位置特定部 28 K は、第 2 俯瞰映像内座標及び第 2 視点パス情報 114（図 21 参照）に基づいて、第 2 視点パス 110 に含まれる複数

の視点位置のうち、特定被写体 138 が最も大きく見える視点位置 110B を特定する。そして、視点位置特定部 28K は、特定した視点位置 110B に対応する視点情報 74 を第 2 視点パス情報 114 から取得する。

[0230] 視点位置特定部 28K によって第 1 視点パス情報 112 から取得された視点情報 74、及び視点位置特定部 28K によって第 2 視点パス情報 114 から取得された視点情報 74 は、視点位置特定部 28K による特定結果である。以下では、説明の便宜上、視点位置特定部 28K によって第 1 視点パス情報 112 から取得された視点情報 74 を「第 1 特定結果」とも称し、視点位置特定部 28K によって第 2 視点パス情報 114 から取得された視点情報 74 を「第 2 特定結果」とも称する。

[0231] 一例として図 32 に示すように、取得部 28C は、視点位置特定部 28K から第 1 特定結果を取得する。取得部 28C は、ストレージ 30 に記憶されている第 1 仮想視点動画 118 から、第 1 特定結果に対応する仮想視点画像 116 を第 1 視点位置仮想視点画像 140 として取得する。第 1 視点位置仮想視点画像 140 は、第 1 視点パス 108 (図 31 参照) 上で特定被写体 138 が最も大きく見える視点位置 108B に対応する仮想視点画像 116、すなわち、視点位置 108B に対応する視点情報 74 に従って生成された仮想視点画像 116 である。

[0232] 加工部 28F は、取得部 28C によって取得された第 1 視点位置仮想視点画像 140 をサムネイル化する。すなわち、加工部 28F は、第 1 視点位置仮想視点画像 140 をサムネイル画像 142 に加工する。また、加工部 28F は、サムネイル化される前の第 1 視点位置仮想視点画像 140 に付与されている第 1 メタデータ (図示省略) をサムネイル画像 142 に対して関連付ける。また、加工部 28F は、サムネイル化された第 1 視点位置仮想視点画像 140 が含まれる第 1 仮想視点動画 118 から第 1 動画識別情報 120 を取得する。

[0233] なお、このように、加工部 28F によって取得されたサムネイル画像 142、サムネイル画像 142 に関連付けられた第 1 メタデータ、及び加工部 2

8Fによって取得された第1動画識別情報120に対してプロセッサ28によって行われる処理は、例えば、上記第2実施形態で説明したサムネイル画像102、メタデータ92A、及び動画識別情報96に対してプロセッサ28によって行われる処理と同じである（図18参照）。

[0234] 一例として図33に示すように、取得部28Cは、視点位置特定部28Kから第2特定結果を取得する。取得部28Cは、ストレージ30に記憶されている第2仮想視点動画124から、第2特定結果に対応する仮想視点画像122を第2視点位置仮想視点画像144として取得する。第2視点位置仮想視点画像144は、第2視点パス110（図31参照）上で特定被写体138が最も大きく見える視点位置110Bに対応する仮想視点画像122、すなわち、視点位置110Bに対応する視点情報74に従って生成された仮想視点画像116である。

[0235] 加工部28Fは、取得部28Cによって取得された第2視点位置仮想視点画像144をサムネイル化する。すなわち、加工部28Fは、第2視点位置仮想視点画像144をサムネイル画像146に加工する。また、加工部28Fは、サムネイル化される前の第2視点位置仮想視点画像144に付与されている第2メタデータ（図示省略）をサムネイル画像146に対して関連付ける。また、加工部28Fは、サムネイル化された第2視点位置仮想視点画像144が含まれる第2仮想視点動画124から第2動画識別情報126を取得する。

[0236] なお、このように、加工部28Fによって取得されたサムネイル画像146、サムネイル画像146に関連付けられた第2メタデータ、及び加工部28Fによって取得された第2動画識別情報126に対してプロセッサ28によって行われる処理は、例えば、上記第2実施形態で説明したサムネイル画像102、メタデータ92A、及び動画識別情報96に対してプロセッサ28によって行われる処理と同じである（図18参照）。

[0237] 以上説明したように、本第4実施形態に係る画像処理装置10では、第1視点パス108と第2視点パス110との位置関係に基づいてサムネイル画

像142及び146が取得される。例えば、第1仮想視点動画118に含まれる全ての仮想視点画像116のうち、第1視点パス108上で特定被写体138が最も大きく見える視点位置108Bに対応する第1視点位置仮想視点画像140のサムネイル画像142が得られる。また、第2仮想視点動画124に含まれる全ての仮想視点画像122のうち、第2視点パス110上で特定被写体138が最も大きく見える視点位置110Bに対応する第2視点位置仮想視点画像144のサムネイル画像146が得られる。そして、画像処理装置10によって取得されたサムネイル画像142及び146を含む一覧画面がユーザデバイス12のタッチパネルディスプレイ16に表示される。従って、本構成によれば、第1視点パス108と第2視点パス110との位置関係に基づいて得られたサムネイル画像142及び146をユーザ14に見せることに寄与することができる。

[0238] また、本第4実施形態に係る画像処理装置10では、特定被写体138に対する第1視点パス108と第2視点パス110との位置関係に基づいてサムネイル画像142及び146が取得される。従って、本構成によれば、特定被写体138に対する第1視点パス108と第2視点パス110との位置関係に基づいて得られたサムネイル画像142及び146をユーザ14に見せることに寄与することができる。

[0239] なお、上記第4実施形態では、第1視点パス108上から特定被写体138が最も大きく見える視点位置108Bと、第2視点パス110上から特定被写体138が最も大きく見える視点位置110Bとを例示したが、本開示の技術はこれに限定されず、例えば、第1視点パス108上から、ユーザ14等によって予め定められたサイズ範囲内で特定被写体138が最も大きく見える視点位置と、第2視点パス110上からユーザ14等によって予め定められたサイズ範囲内で特定被写体138が最も大きく見える視点位置とを適用してもよい。

[0240] また、上記第4実施形態では、第1視点パス108及び第2視点パス110という2本の視点パスを例示したが、本開示の技術はこれに限定されず、

3本以上の視点パスの位置関係に基づいてサムネイル化される仮想視点画像が特定されるようにしてもよい。

[0241] [第5実施形態]

本第5実施形態では、上記第1～第4実施形態で説明した構成要素については、同一の符号を付し、説明を省略し、上記第1～第4実施形態と異なる部分について説明する。

[0242] 一例として図34に示すように、本第5実施形態に係る画像処理装置10のプロセッサ28は、図28に示すプロセッサ28に比べ、画面生成処理プログラム38を実行することで検索条件付与部28Lとして更に動作する点異なる。

[0243] 一例として図35に示すように、ストレージ30には、複数の仮想視点動画78が記憶されている。検索条件付与部28Lは、検索条件148を取得部28Cに付与する。検索条件148とは、複数の仮想視点動画78から、サムネイル化される仮想視点画像76が含まれる仮想視点動画を検索する条件を指す。検索条件148の一例としては、メタデータ76Aに含まれる各種情報（例えば、仮想視点画像76が生成された時刻等）、及び／又は、動画識別情報80等が挙げられる。検索条件148は、受付デバイス50等によって受け付けられた指示、及び／又は、各種条件（例えば、撮像条件等）に従って決定される。最初に決定された検索条件148は、固定されてもよいし、受付デバイス50等によって受け付けられた指示、及び／又は、各種条件（例えば、撮像条件等）に応じて変更されてもよい。

[0244] 取得部28Cは、ストレージ30に記憶されている複数の仮想視点動画78から、検索条件付与部28Lによって付与された検索条件148に適合する仮想視点動画78である検索条件適合仮想視点動画150を検索する。ここで、「適合」の意味には、検索条件148との完全一致の他、許容される誤差内での一致の意味も含まれる。本第5実施形態に係る画像処理装置10では、取得部28Cによって検索されて得られた検索条件適合仮想視点動画150に対して上記第1～第4実施形態で説明したプロセッサ28による処

理が行われる。

[0245] このように、本第5実施形態に係る画像処理装置10では、複数の仮想視点動画78から、与えられた検索条件148に適合する検索条件適合仮想視点動画150が検索され、検索されて得られた検索条件適合仮想視点動画150に基づいて上記第1～5実施形態で説明したサムネイル画像が取得される。従って、本構成によれば、与えられた検索条件に適合する仮想視点動画78に基づいて得られたサムネイル画像をユーザ14に見せることに寄与することができる。

[0246] 本第5実施形態の変形例として、例えば、何らかの方法により生成され、それぞれ動画に関連付けられた複数のサムネイル画像がディスプレイに一覧表示されている場合に、ユーザ14が検索条件148を入力すると、検索条件適合仮想視点動画150に関連付けられているサムネイル画像は、入力された検索条件に応じて変更されても良い。例えば、ユーザ14が特定の人物（例えば、特定の人物の名称）を検索条件として入力した場合、検索条件として入力された特定の人物が映った検索条件適合仮想視点動画150に関連付けられているサムネイル画像は、特定の人物のサムネイル画像に変更されて表示される。この場合、例えば、検索条件適合仮想視点動画150の中で特定の人物が最も大きく映ったフレームを変更後のサムネイル画像とする。これにより、ユーザ14は、検索条件として入力した特定の人物がそれぞれの動画においてどのように映っているかを一覧で確認することができる。

[0247] [第6実施形態]

本第6実施形態では、上記第1～第5実施形態で説明した構成要素については、同一の符号を付し、説明を省略し、上記第1～第5実施形態と異なる部分について説明する。

[0248] 一例として図36に示すように、本第6実施形態に係る画像処理装置10のプロセッサ28は、図34に示すプロセッサ28に比べ、画面生成処理プログラム38を実行することで状況認識部28Mとして更に動作する点が異なる。

[0249] 一例として図37に示すように、状況認識部28Mは、ストレージ30に記憶されている仮想視点動画78に含まれる複数の仮想視点画像76（例えば、指定された時間帯に含まれる複数の仮想視点画像76、及び／又は、仮想視点動画78に含まれる全ての仮想視点画像76等）対してA1方式の被写体認識処理を行うことで、特定状況に関する仮想視点画像76を特定する。ここで、特定状況としては、例えば、単位面積あたりに既定人数以上の人物被写体が存在している状況、サッカーフィールド内のペナルティエリア内にサッカーボールと複数の人物被写体とが存在している状況、ボールをキープしている人物被写体を複数の人物被写体を取り囲んでいる状況、及び／又は、ゴールキーパーの指先にサッカーボールが触れている状況等が挙げられる。なお、サッカーフィールド内に存在する人物被写体は、本開示の技術に係る「第3被写体」の一例であり、特定状況は、本開示の技術に係る「第3被写体の状況」の一例である。

[0250] 取得部28Cは、状況認識部28Mによって特定された仮想視点画像76を特定状況仮想視点画像152として仮想視点動画78から取得する。本第6実施形態に係る画像処理装置10では、取得部28Cによって取得された特定状況仮想視点画像152に対して上記第1～第5実施形態で説明したプロセッサ28による処理が行われる。

[0251] このように、本第6実施形態に係る画像処理装置10では、特定状況に応じて決められた仮想視点画像76がサムネイル化される。すなわち、状況認識部28Mによって特定された特定状況仮想視点画像152が取得部28Cによって取得され、特定状況仮想視点画像152に対応するサムネイル画像が生成される。従って、本構成によれば、特定状況に応じて決められたサムネイル画像をユーザ14に見せることができる。

[0252] [第7実施形態]

本第7実施形態では、上記第1～第6実施形態で説明した構成要素については、同一の符号を付し、説明を省略し、上記第1～第6実施形態と異なる部分について説明する。

- [0253] 一例として図38に示すように、本第7実施形態に係る画像処理装置10のプロセッサ28は、図36に示すプロセッサ28に比べ、画面生成処理プログラム38を実行することで人物属性被写体認識部28Nとして更に動作する点異なる。
- [0254] 一例として図39に示すように、人物属性被写体認識部28Nは、ストレージ30に記憶されている仮想視点動画78に含まれる複数の仮想視点画像76（例えば、指定された時間帯に含まれる複数の仮想視点画像76、及び／又は、仮想視点動画78に含まれる全ての仮想視点画像76等）対してAI方式の被写体認識処理を行うことで、特定人物の属性に関する仮想視点画像76を特定する。ここで、特定人物とは、例えば、仮想視点動画78を観賞している人物、及び／又は、仮想視点動画78の製作に関わっている人物等の仮想視点動画78に関わる人物を指す。属性としては、例えば、性別、年齢、住所、職業、人種、及び／又は、課金状況等が挙げられる。
- [0255] 人物属性被写体認識部28Nは、特定人物の属性毎に応じた被写体認識処理を行うことで特定人物の属性に関する仮想視点画像76を特定する。この場合、例えば、先ず、人物属性被写体認識部28Nは、特定人物の種類及び属性を入力とし、仮想視点動画78に含まれる被写体を特定可能な被写体特定情報を出力とする導出テーブル（図示省略）から、外部（例えば、ユーザデバイス12等）から与えられた特定人物の種類及び属性に対応する被写体特定情報を導出する。そして、人物属性被写体認識部28Nは、導出テーブルから導出した被写体特定情報から特定される被写体が含まれる仮想視点画像76を、仮想視点動画78に対して被写体認識処理を行うことで特定する。
- [0256] 取得部28Cは、人物属性被写体認識部28Nによって特定された仮想視点画像76を人物属性仮想視点画像154として仮想視点動画78から取得する。本第7実施形態に係る画像処理装置10では、取得部28Cによって取得された人物属性仮想視点画像154に対して上記第1～第6実施形態で説明したプロセッサ28による処理が行われる。

[0257] このように、本第7実施形態に係る画像処理装置10では、仮想視点動画78に関わる人物の属性に応じて決められた仮想視点画像76がサムネイル化される。すなわち、人物属性被写体認識部28Nによって特定された人物属性仮想視点画像154が取得部28Cによって取得され、人物属性仮想視点画像154に対応するサムネイル画像が生成される。従って、本構成によれば、仮想視点動画78に関わる人物の属性に応じて決められたサムネイル画像をユーザ14に見せることができる。

[0258] なお、上記各実施形態では、互いに視点が異なる複数の視点情報74の各々に視点位置情報74A、視線方向情報74B、画角情報74C、移動速度情報74D、及び経過時間情報74Eが含まれる形態例を挙げて説明したが、本開示の技術はこれに限定されず、互いに視点が異なる複数の視点情報74に、互いに異なる時刻に関する情報が含まれていてもよい。例えば、図40に示すように、第1視点パス情報112に含まれる複数の視点情報74が、互いに異なる時刻に関する情報である時刻情報74Fを有し、第2視点パス情報114に含まれる複数の視点情報74も、互いに異なる時刻に関する情報である時刻情報74Fを有するようにしてもよい。これにより、仮想視点動画78に対応するサムネイル画像として、互いに異なる視点及び互いに異なる時刻に基づいて得られた画像をユーザ14に見せることに寄与することができる。

[0259] また、上記各実施形態では、サムネイル画像として1フレームの仮想視点画像がサムネイル化された静止画像を例示したが、本開示の技術は、これに限定されず、複数フレームの仮想視点画像がサムネイル化されることで得られた動画像を適用してもよい。この場合、上記各実施形態で説明した要領で仮想視点動画からサムネイル化される仮想視点画像として特定された基準仮想視点画像と、基準仮想視点画像に対して時間的に前及び／又は後の少なくとも1フレームの仮想視点画像とがサムネイル化されることで得られた複数のサムネイル画像に基づいて動画像が生成されるようにしてもよい。このようにサムネイル化された複数の基準仮想視点画像がディスプレイに一覧表示

され、ユーザ 14 が例えばマウス操作により何れかの基準仮想視点画像の上にカーソルを移動させた場合に、カーソルを移動させた先の基準仮想視点画像に対応する動画像が再生されるようにしてもよい。

[0260] なお、複数の撮像画像及び複数の視点情報に基づいて代表画像を取得する方法は、上記の方法に限られない。複数の撮像画像 64 及び複数の視点情報 74 を使用して代表画像を取得していればどのような方法で代表画像を決めてもよい。また、上記の通り、代表画像とは、例えば一覧画面に表示される画像である。

[0261] また、上記各実施形態では、画像処理装置 10 のコンピュータ 22 によって画面生成処理が実行される形態例を挙げて説明したが、本開示の技術はこれに限定されず、ユーザデバイス 12 のコンピュータ 40 によって画面生成処理が実行されるようにしてもよいし、画像処理装置 10 のコンピュータ 22 とユーザデバイス 12 のコンピュータ 40 とによる分散処理が行われるようにしてもよい。

[0262] また、上記各実施形態では、コンピュータ 22 を例示したが、本開示の技術はこれに限定されない。例えば、コンピュータ 22 に代えて、ASIC、FPGA、及び／又はPLDを含むデバイスを適用してもよい。また、コンピュータ 22 に代えて、ハードウェア構成及びソフトウェア構成の組み合わせを用いてもよい。ユーザデバイス 12 のコンピュータ 40 についても同様である。

[0263] また、上記の例では、ストレージ 30 に画面生成処理プログラム 38 が記憶されているが、本開示の技術はこれに限定されず、一例として図 41 に示すように、非一時的記憶媒体であるSSD又はUSBメモリなどの任意の可搬型の記憶媒体 200 に画面生成処理プログラム 38 が記憶されていてもよい。この場合、記憶媒体 200 に記憶されている画面生成処理プログラム 38 がコンピュータ 22 にインストールされ、プロセッサ 28 は、画面生成処理プログラム 38 に従って、画面生成処理を実行する。

[0264] また、通信網（図示省略）を介してコンピュータ 22 に接続される他のコ

ンピュータ又はサーバ装置等のメモリに画面生成処理プログラム38を記憶させておき、画像処理装置10の要求に応じて画面生成処理プログラム38が画像処理装置10にダウンロードされるようにしてもよい。この場合、ダウンロードされた画面生成処理プログラム38に従って、画面生成処理がコンピュータ22のプロセッサ28によって実行される。

[0265] また、上記の例では、プロセッサ28を例示したが、プロセッサ28に代えて、又は、プロセッサ28と共に、他の少なくとも1つのCPU、少なくとも1つのGPU、及び／又は、少なくとも1つのTPUを用いるようにしてもよい。

[0266] 画面生成処理を実行するハードウェア資源としては、次に示す各種のプロセッサを用いることができる。プロセッサとしては、例えば、上述したように、ソフトウェア、すなわち、プログラムに従って画面生成処理を実行するハードウェア資源として機能する汎用的なプロセッサであるCPUが挙げられる。また、他のプロセッサとしては、例えば、FPGA、PLD、又はASICなどの専用の処理を実行させるために専用に設計された回路構成を有するプロセッサである専用電気回路が挙げられる。何れのプロセッサにもメモリが内蔵又は接続されており、何れのプロセッサもメモリを使用することで画面生成処理を実行する。

[0267] 画面生成処理を実行するハードウェア資源は、これらの各種のプロセッサのうちの1つで構成されてもよいし、同種または異種の2つ以上のプロセッサの組み合わせ（例えば、複数のFPGAの組み合わせ、又はCPUとFPGAとの組み合わせ）で構成されてもよい。また、画面生成処理を実行するハードウェア資源は1つのプロセッサであってもよい。

[0268] 1つのプロセッサで構成する例としては、第1に、クライアント及びサーバなどのコンピュータに代表されるように、1つ以上のCPUとソフトウェアの組み合わせで1つのプロセッサを構成し、このプロセッサが、画面生成処理を実行するハードウェア資源として機能する形態がある。第2に、SOCなどに代表されるように、画面生成処理を実行する複数のハードウェア資

源を含むシステム全体の機能を1つのICチップで実現するプロセッサを使用する形態がある。このように、画面生成処理は、ハードウェア資源として、上記各種のプロセッサの1つ以上を用いて実現される。

[0269] 更に、これらの各種のプロセッサのハードウェア的な構造としては、より具体的には、半導体素子などの回路素子を組み合わせた電気回路を用いることができる。

[0270] また、上述した画面生成処理はあくまでも一例である。従って、主旨を逸脱しない範囲内において不要なステップを削除したり、新たなステップを追加したり、処理順序を入れ替えたりしてもよいことは言うまでもない。

[0271] 以上に示した記載内容及び図示内容は、本開示の技術に係る部分についての詳細な説明であり、本開示の技術の一例に過ぎない。例えば、上記の構成、機能、作用、及び効果に関する説明は、本開示の技術に係る部分の構成、機能、作用、及び効果の一例に関する説明である。よって、本開示の技術の主旨を逸脱しない範囲内において、以上に示した記載内容及び図示内容に対して、不要な部分を削除したり、新たな要素を追加したり、置き換えたりしてもよいことは言うまでもない。また、錯綜を回避し、本開示の技術に係る部分の理解を容易にするために、以上に示した記載内容及び図示内容では、本開示の技術の実施を可能にする上で特に説明を要しない技術常識等に関する説明は省略されている。

[0272] 本明細書において、「A及び／又はB」は、「A及びBのうちの少なくとも1つ」と同義である。つまり、「A及び／又はB」は、Aだけであってもよいし、Bだけであってもよいし、A及びBの組み合わせであってもよい、という意味である。また、本明細書において、3つ以上の事柄を「及び／又は」で結び付けて表現する場合も、「A及び／又はB」と同様の考え方が適用される。

[0273] 本明細書に記載された全ての文献、特許出願及び技術規格は、個々の文献、特許出願及び技術規格が参照により取り込まれることが具体的かつ個々に記された場合と同程度に、本明細書中に参照により取り込まれる。

## 請求の範囲

- [請求項1]            プロセッサと、  
                         前記プロセッサに接続又は内蔵されたメモリと、を備え、  
                         前記プロセッサは、  
                         撮像領域が撮像されることで得られる複数の撮像画像と複数の視点  
                         情報とに基づいて生成される仮想視点動画に対応する代表画像を、前  
                         記複数の撮像画像及び前記複数の視点情報に基づいて取得し、  
                         前記仮想視点動画とは異なる大きさに前記代表画像をディスプレイ  
                         に表示させるためのデータを出力する  
                         画像処理装置。
- [請求項2]            前記代表画像は、  
                         前記仮想視点動画において前記撮像領域内の第1被写体が含まれて  
                         いる複数のフレームのうちの第1フレームに関する画像である  
                         請求項1に記載の画像処理装置。
- [請求項3]            前記第1被写体は、前記仮想視点動画に含まれる時間に基づいて決  
                         められた被写体である  
                         請求項2に記載の画像処理装置。
- [請求項4]            前記第1フレームは、前記仮想視点動画内での前記第1被写体のサ  
                         イズに基づいて決められたフレームである  
                         請求項2又は請求項3に記載の画像処理装置。
- [請求項5]            前記プロセッサは、前記複数の視点情報の編集結果に基づいて前記  
                         代表画像を取得する  
                         請求項1から請求項4の何れか一項に記載の画像処理装置。
- [請求項6]            前記複数の視点情報は、複数の視点パスを有し、  
                         前記編集結果は、前記複数の視点パスに対して行われた編集の結果  
                         を含む  
                         請求項5に記載の画像処理装置。
- [請求項7]            前記プロセッサは、前記複数の視点情報間の相違度に基づいて前記

代表画像を取得する

請求項 1 から請求項 6 の何れか一項に記載の画像処理装置。

[請求項8]

前記複数の視点情報は、複数の視点パスを有し、

前記相違度は、前記複数の視点パス間の相違度である

請求項 7 に記載の画像処理装置。

[請求項9]

前記複数の視点情報は、複数の視点パスを有し、

前記プロセッサは、前記複数の視点パスの位置関係に基づいて前記代表画像を取得する

請求項 1 から請求項 8 の何れか一項に記載の画像処理装置。

[請求項10]

前記位置関係は、前記撮像領域内の第 2 被写体に対する前記複数の視点パスの位置関係である

請求項 9 に記載の画像処理装置。

[請求項11]

前記プロセッサは、

複数の前記仮想視点動画から、与えられた検索条件に適合する検索条件適合仮想視点動画を検索し、

前記検索条件適合仮想視点動画に基づいて前記代表画像を取得する

請求項 1 から請求項 10 の何れか一項に記載の画像処理装置。

[請求項12]

前記代表画像は、前記撮像領域内の第 3 被写体の状況に応じて決められた画像である

請求項 1 から請求項 11 の何れか一項に記載の画像処理装置。

[請求項13]

前記代表画像は、前記仮想視点動画に関わる人物の属性に応じて決められた画像である

請求項 1 から請求項 12 の何れか一項に記載の画像処理装置。

[請求項14]

前記代表画像は、前記仮想視点動画の内容を示す画像である

請求項 1 から請求項 13 の何れか一項に記載の画像処理装置。

[請求項15]

前記複数の視点情報は、視点が異なる第 1 視点情報及び第 2 視点情報を含み、

前記第 1 視点情報及び前記第 2 視点情報は、異なる時刻に関する情

報を有する

請求項 1 から請求項 1 4 の何れか一項に記載の画像処理装置。

[請求項16]

前記プロセッサは、

前記代表画像を第 1 ディスプレイに対して表示させるための第 1 データを出力し、

前記第 1 ディスプレイに表示されている前記代表画像が選択されたことに応じて前記代表画像に対応する前記仮想視点動画を前記第 1 ディスプレイ及び第 2 ディスプレイのうちの少なくとも一方に対して表示させるための第 2 データを出力する

請求項 1 から請求項 1 5 の何れか一項に記載の画像処理装置。

[請求項17]

前記プロセッサは、前記代表画像と前記仮想視点動画とを関連付けた状態で前記メモリに記憶させる

請求項 1 から請求項 1 6 の何れか一項に記載の画像処理装置。

[請求項18]

プロセッサと、

前記プロセッサに接続又は内蔵されたメモリと、を備え、

前記プロセッサは、

撮像領域が撮像されることで得られる複数の撮像画像と複数の視点情報とに基づいて生成される仮想視点動画に対応する代表画像を、前記複数の撮像画像及び前記複数の視点情報に基づいて取得し、

複数の画像が表示される画面に前記代表画像を表示させるためのデータを出力する

画像処理装置。

[請求項19]

撮像領域が撮像されることで得られる複数の撮像画像と複数の視点情報とに基づいて生成される仮想視点動画に対応する代表画像を、前記複数の撮像画像及び前記複数の視点情報に基づいて取得すること、並びに、

前記仮想視点動画とは異なる大きさに前記代表画像をディスプレイに表示させるためのデータを出力すること

を含む画像処理方法。

[請求項20]

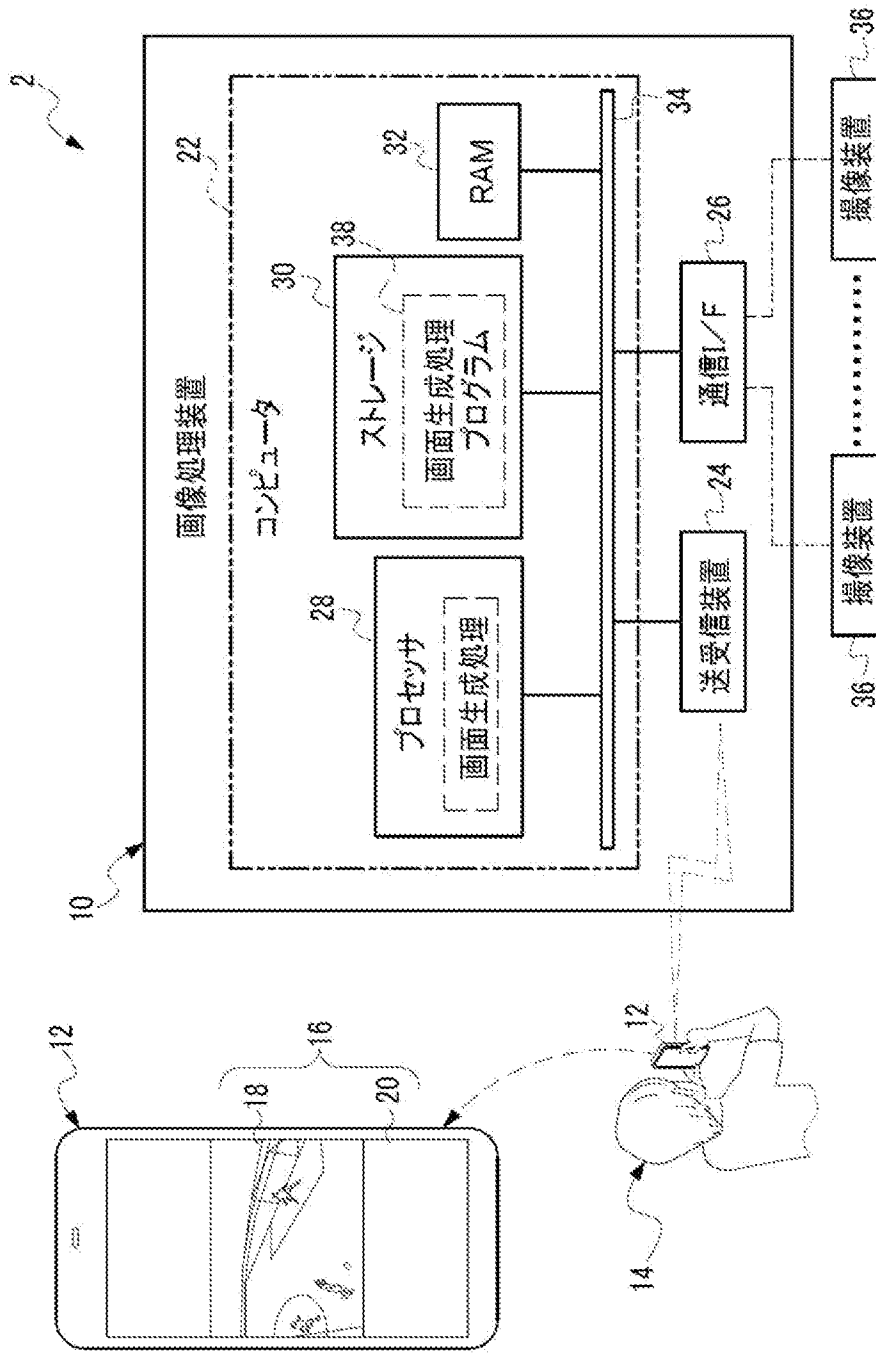
コンピュータに、

撮像領域が撮像されることで得られる複数の撮像画像と複数の視点情報とに基づいて生成される仮想視点動画に対応する代表画像を、前記複数の撮像画像及び前記複数の視点情報に基づいて取得すること、並びに、

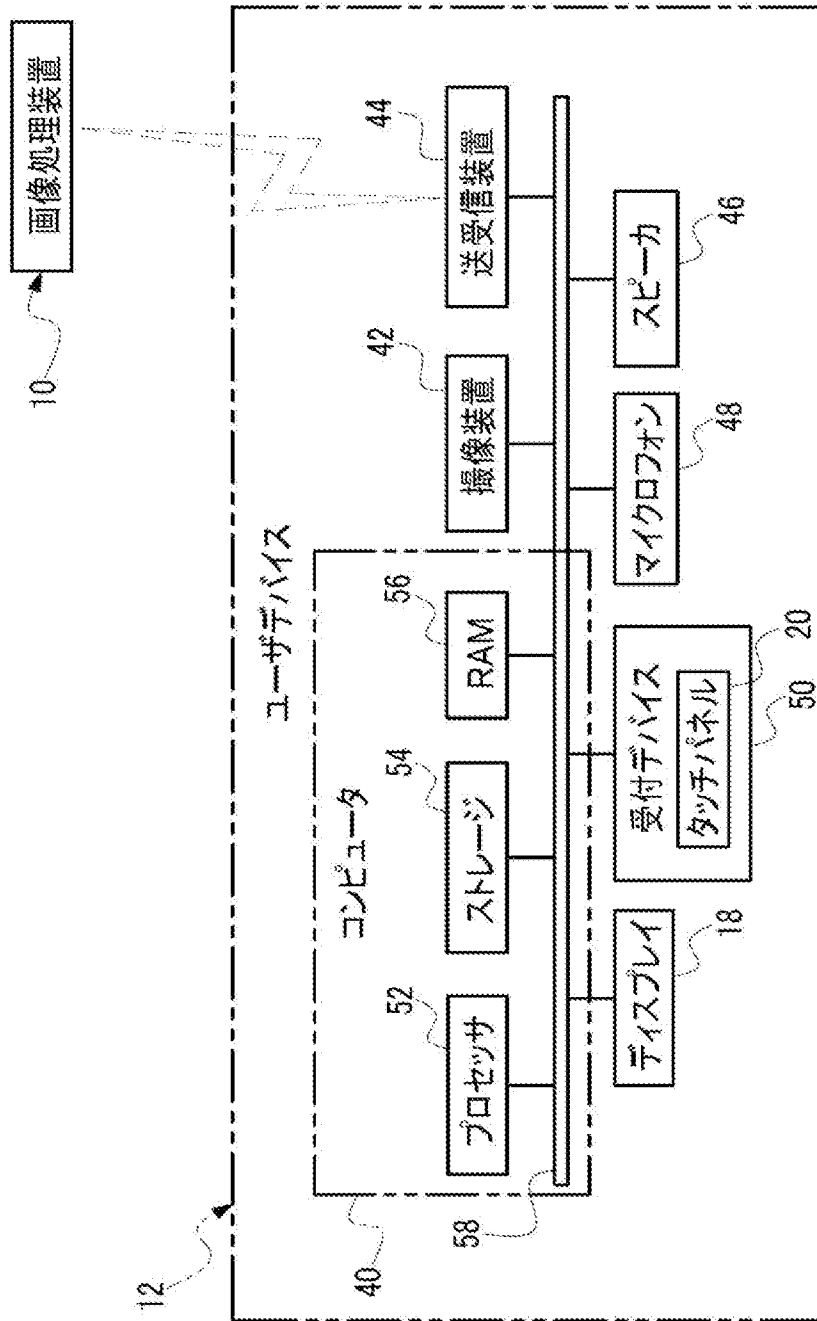
前記仮想視点動画とは異なる大きさに前記代表画像をディスプレイに表示させるためのデータを出力すること

を含む処理を実行させるためのプログラム。

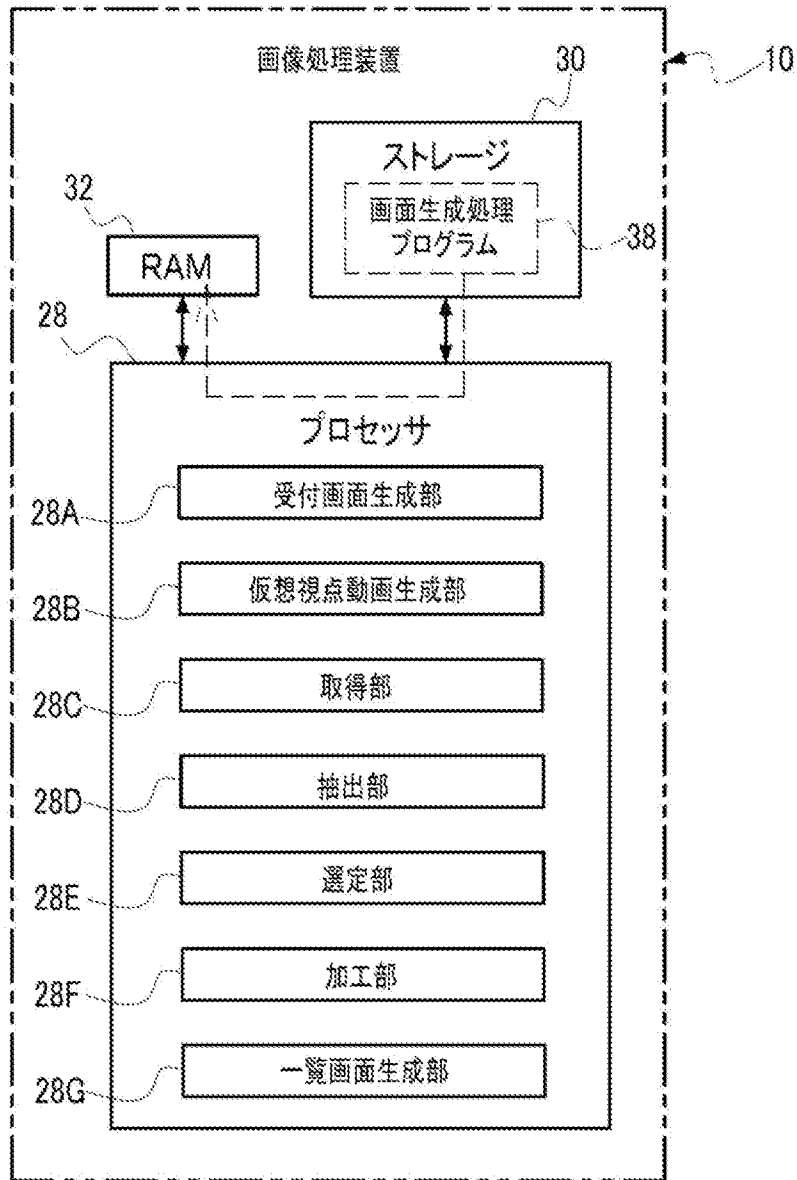
[図1]



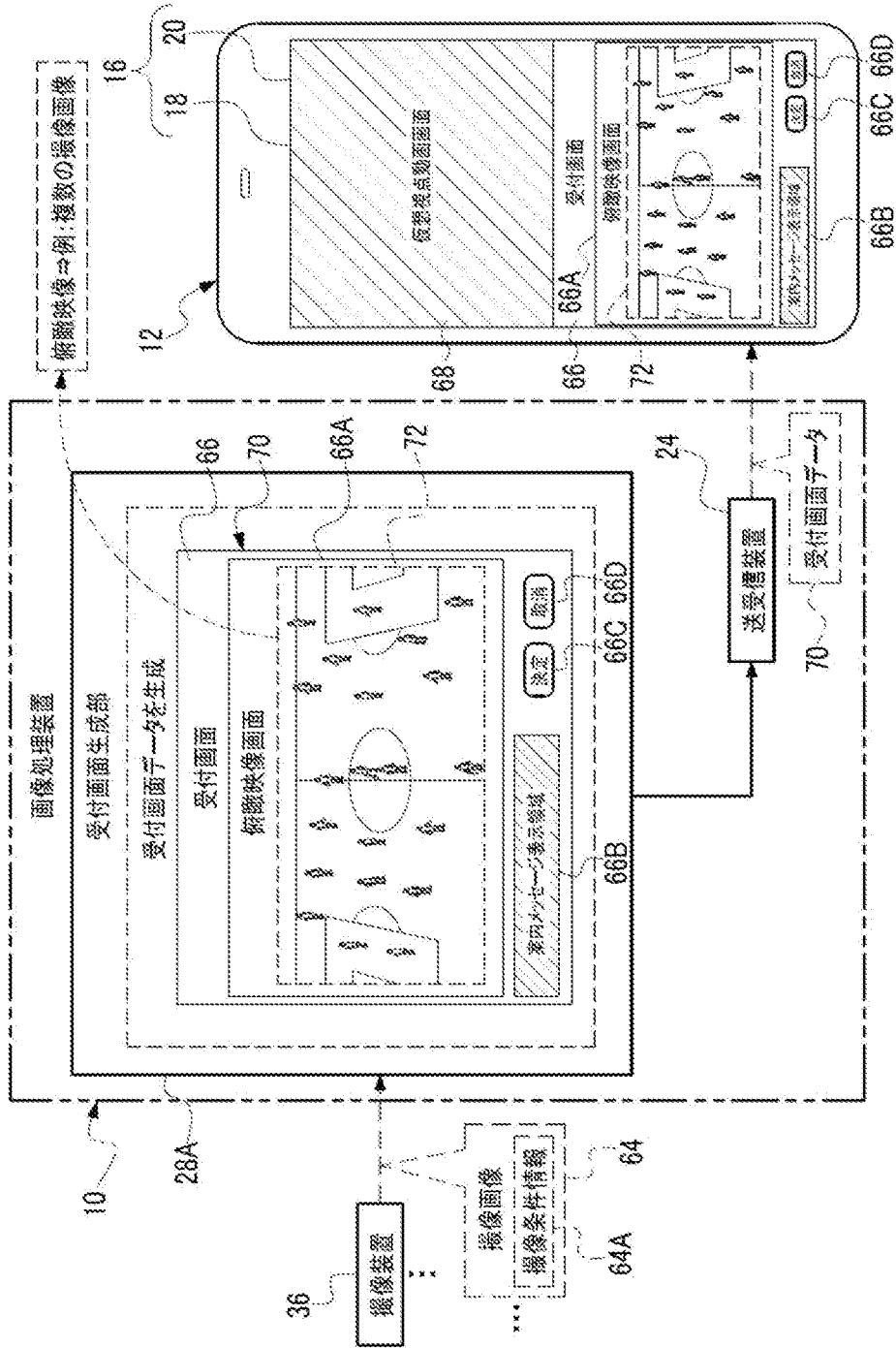
[図2]



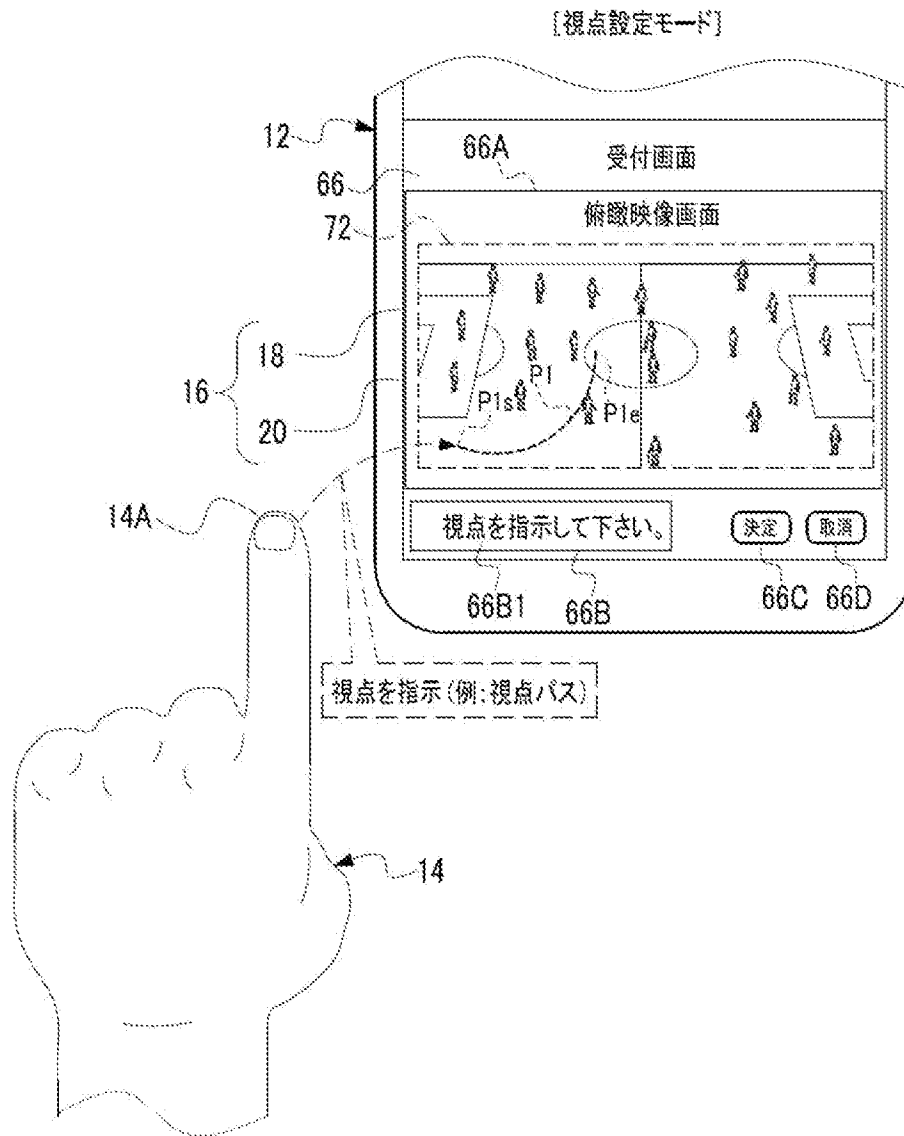
[図3]



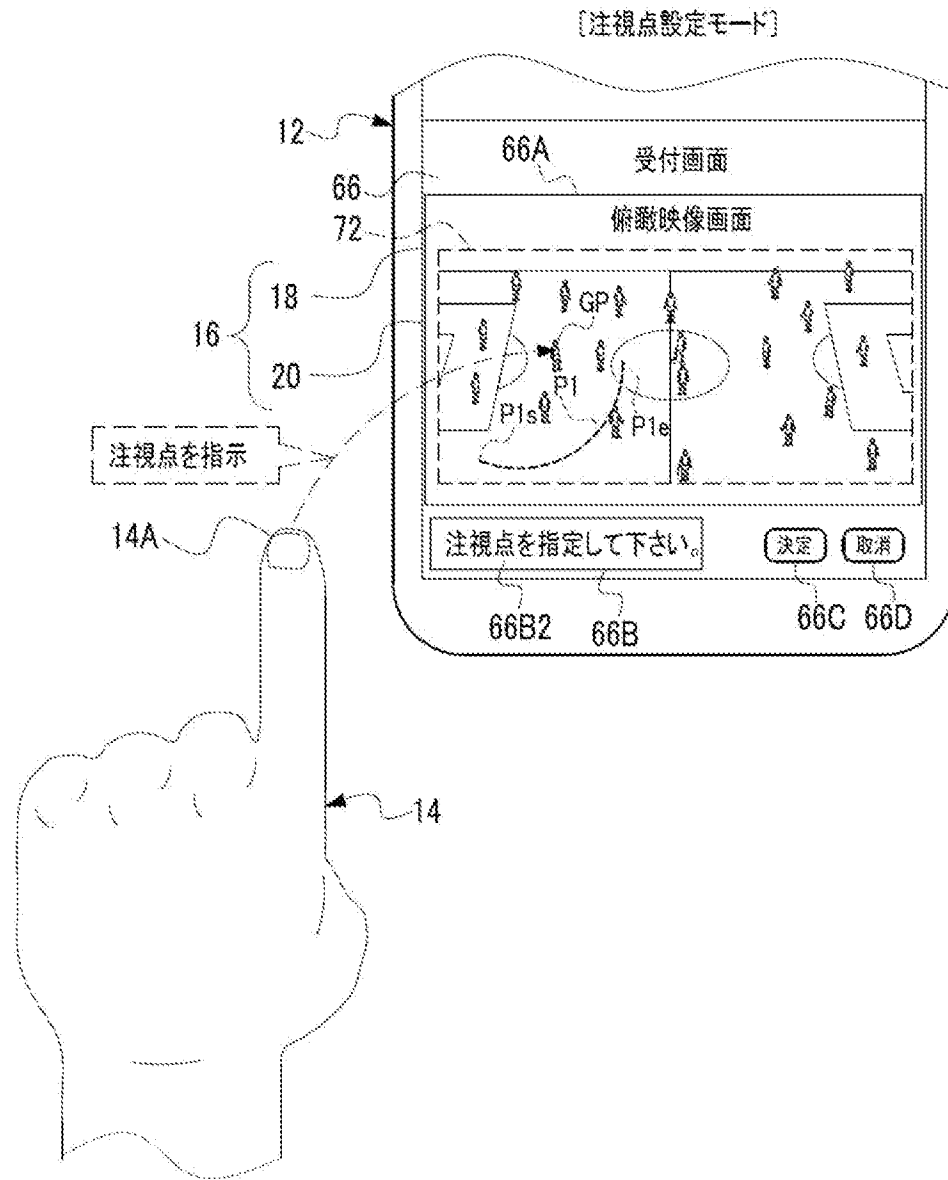
[図4]



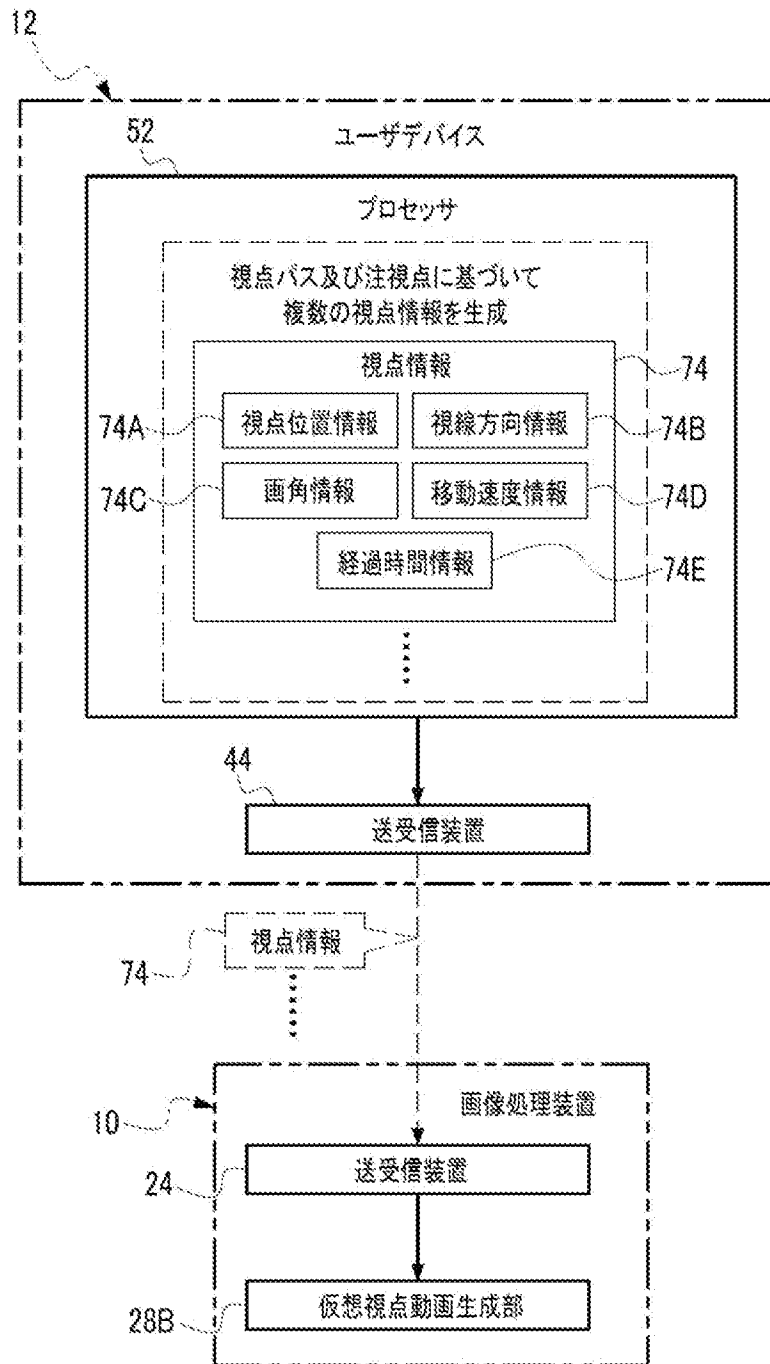
[図5]



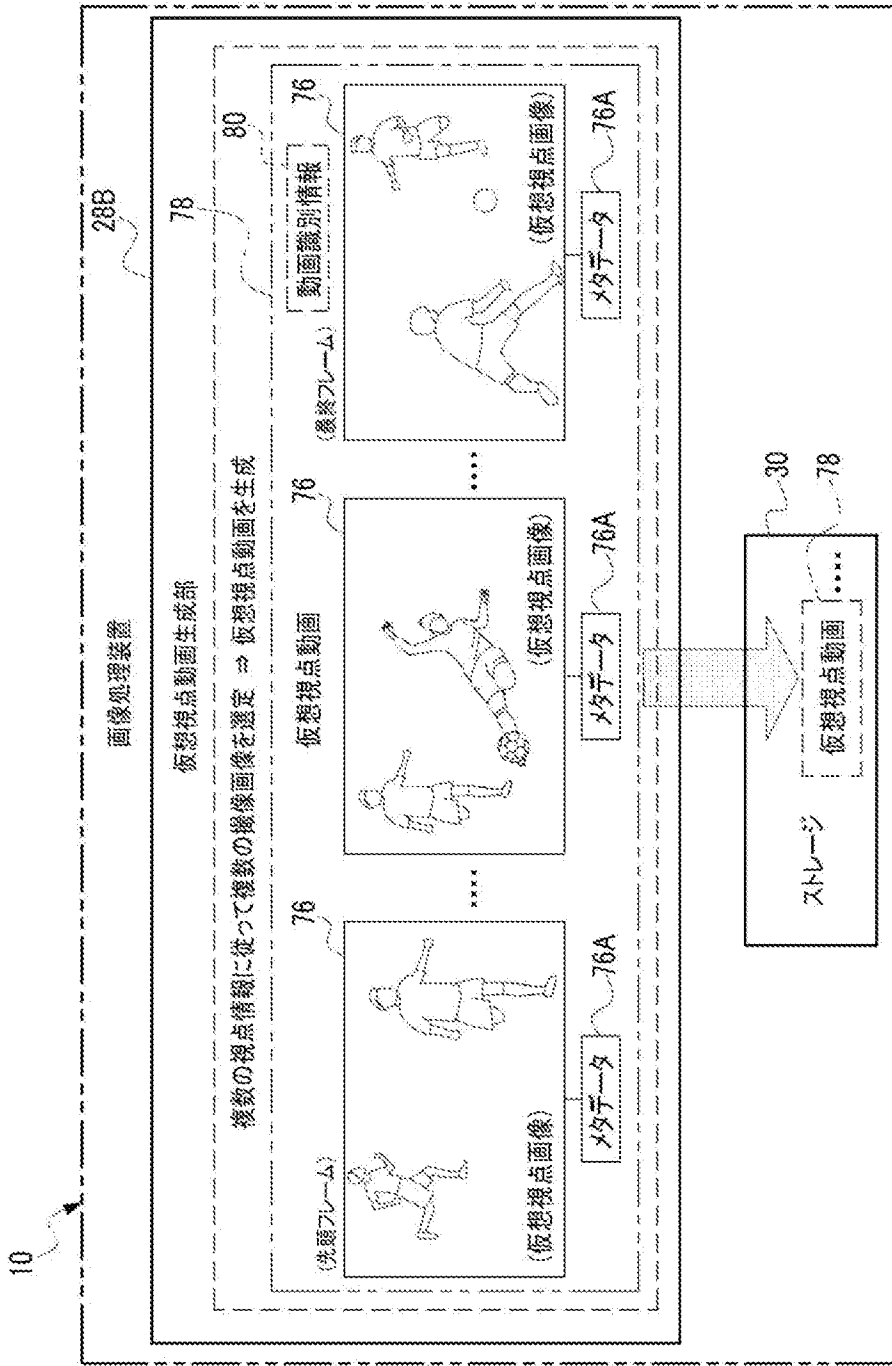
[図6]



[図7]

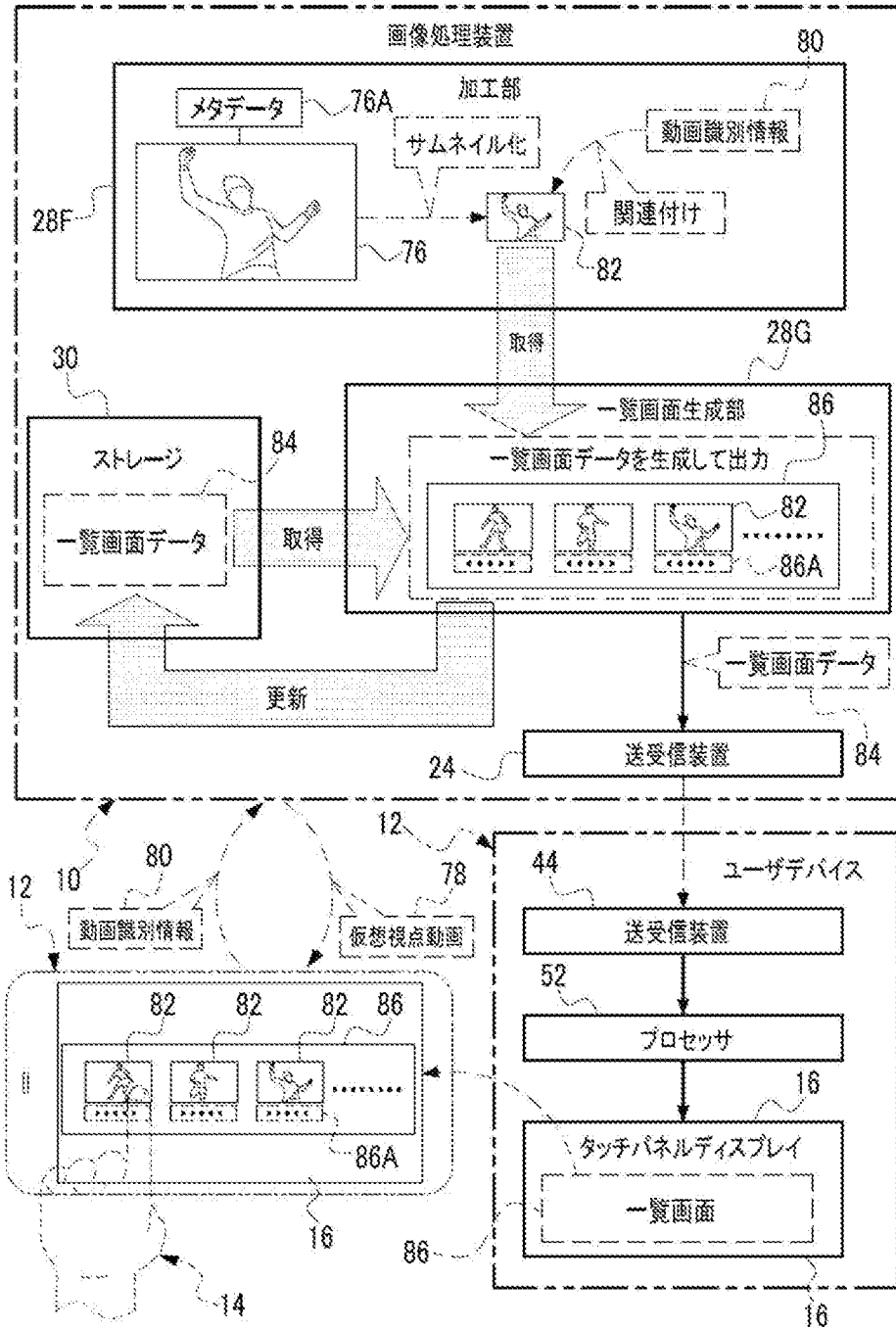


[図8]

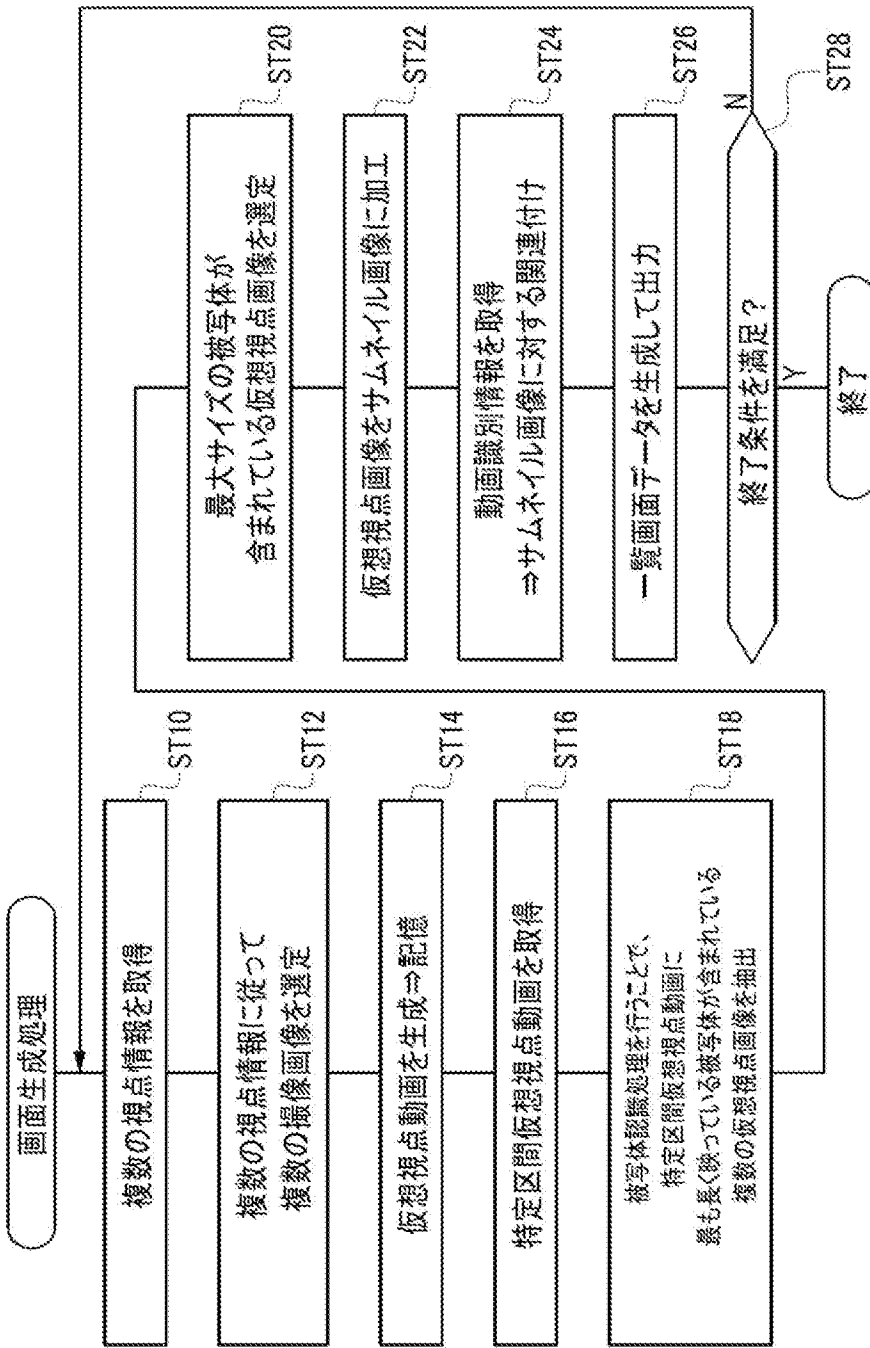




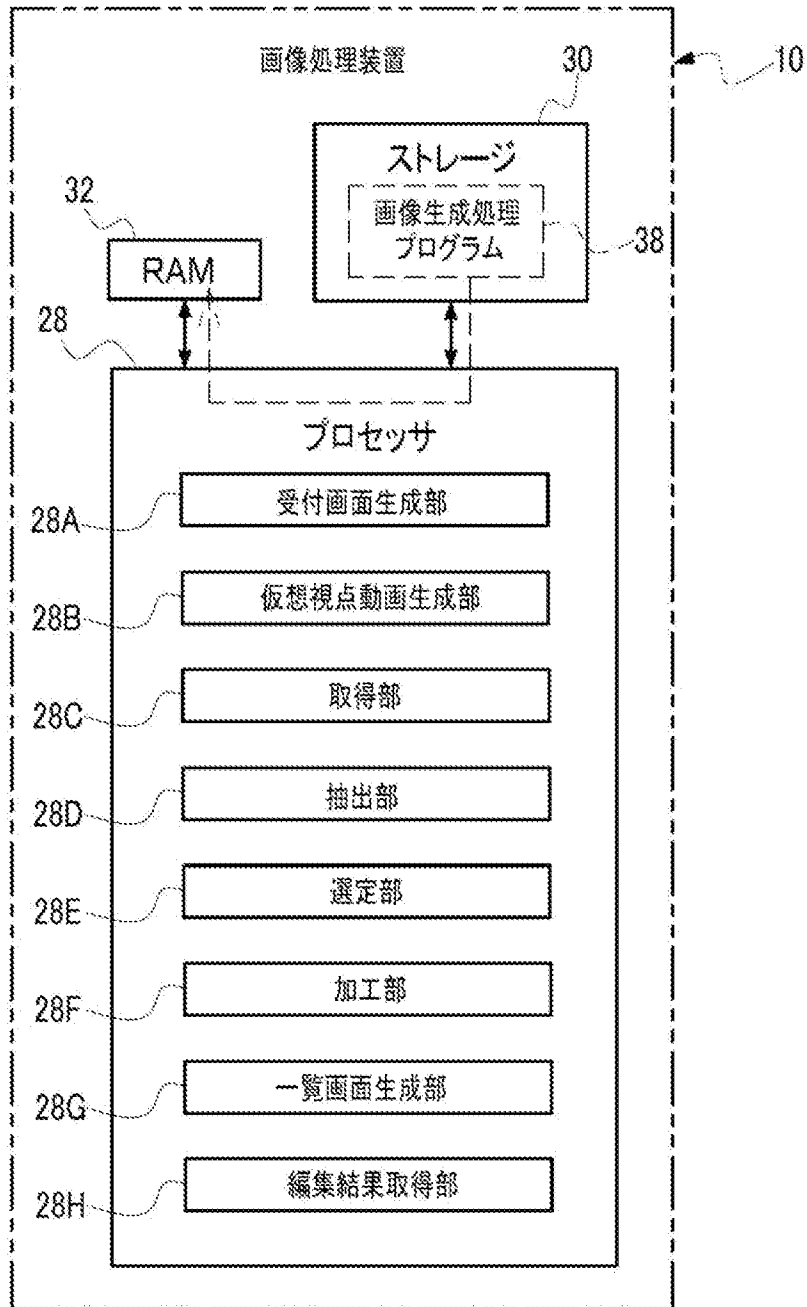
[図10]



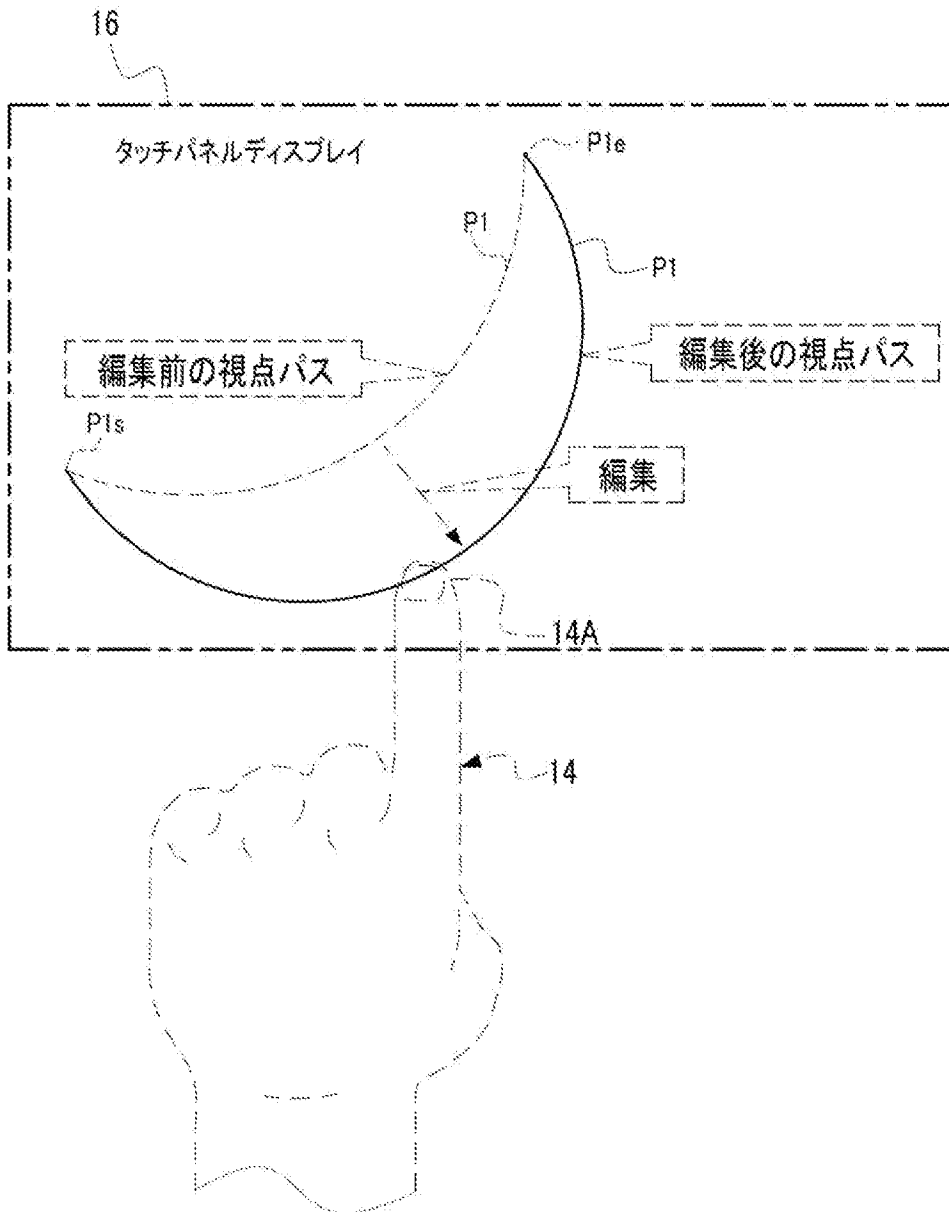
[図11]



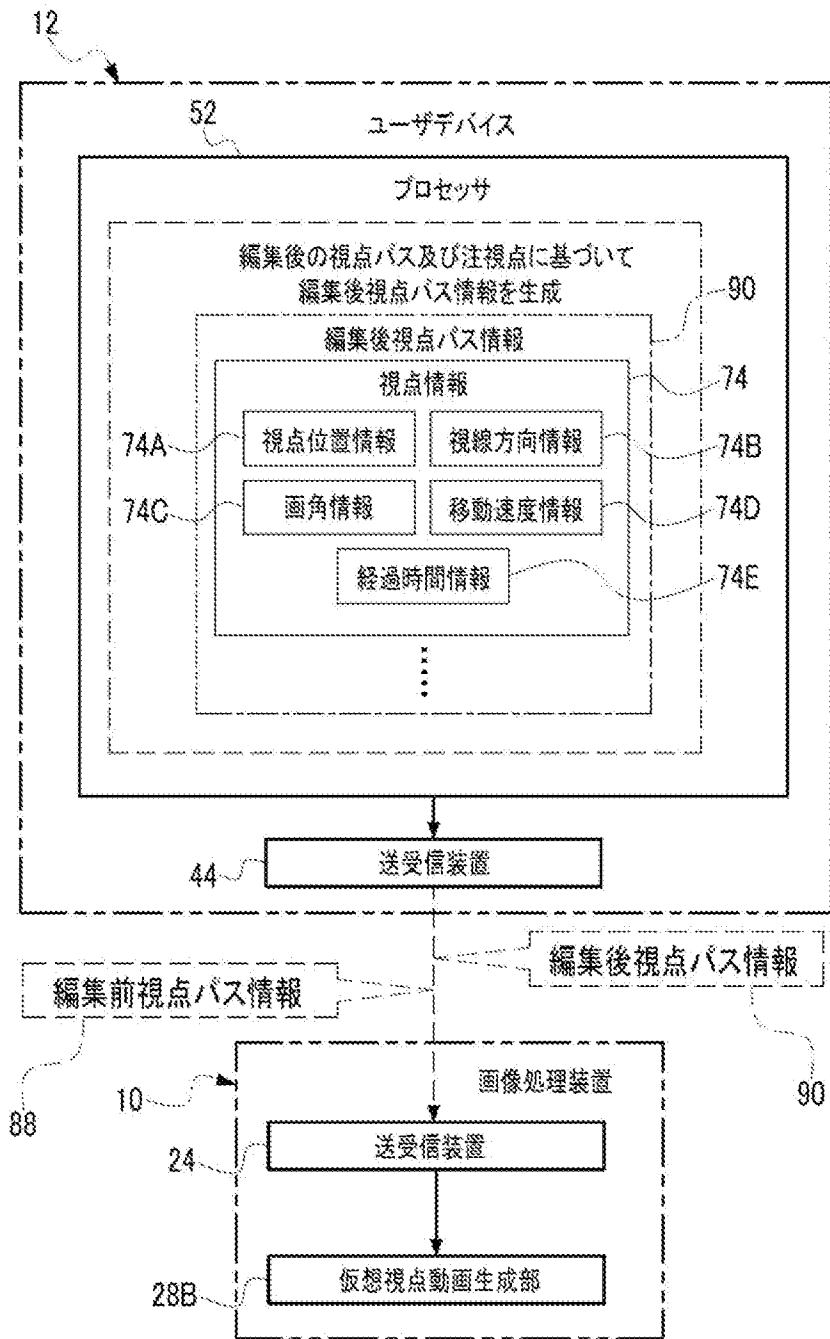
[図12]



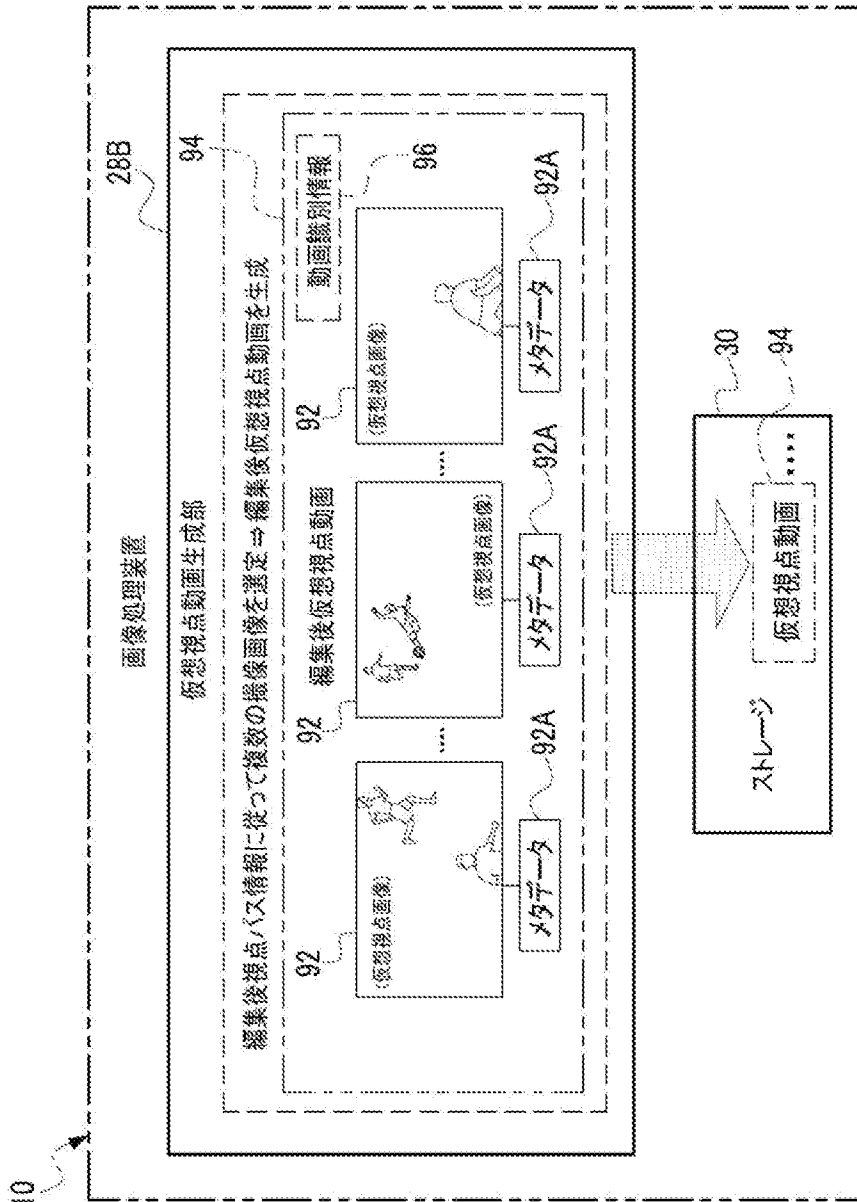
[図13]



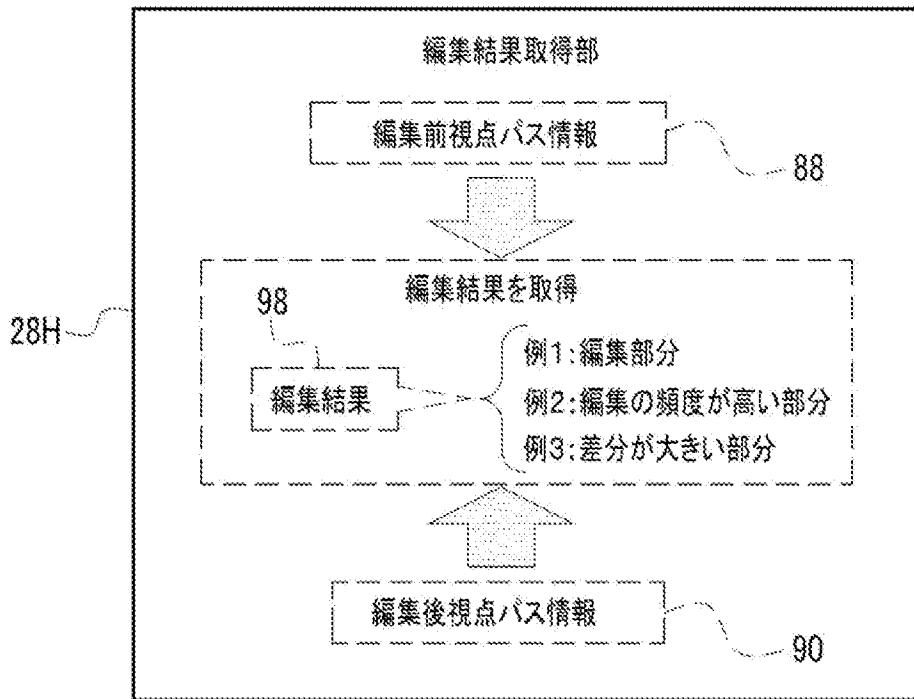
[図14]



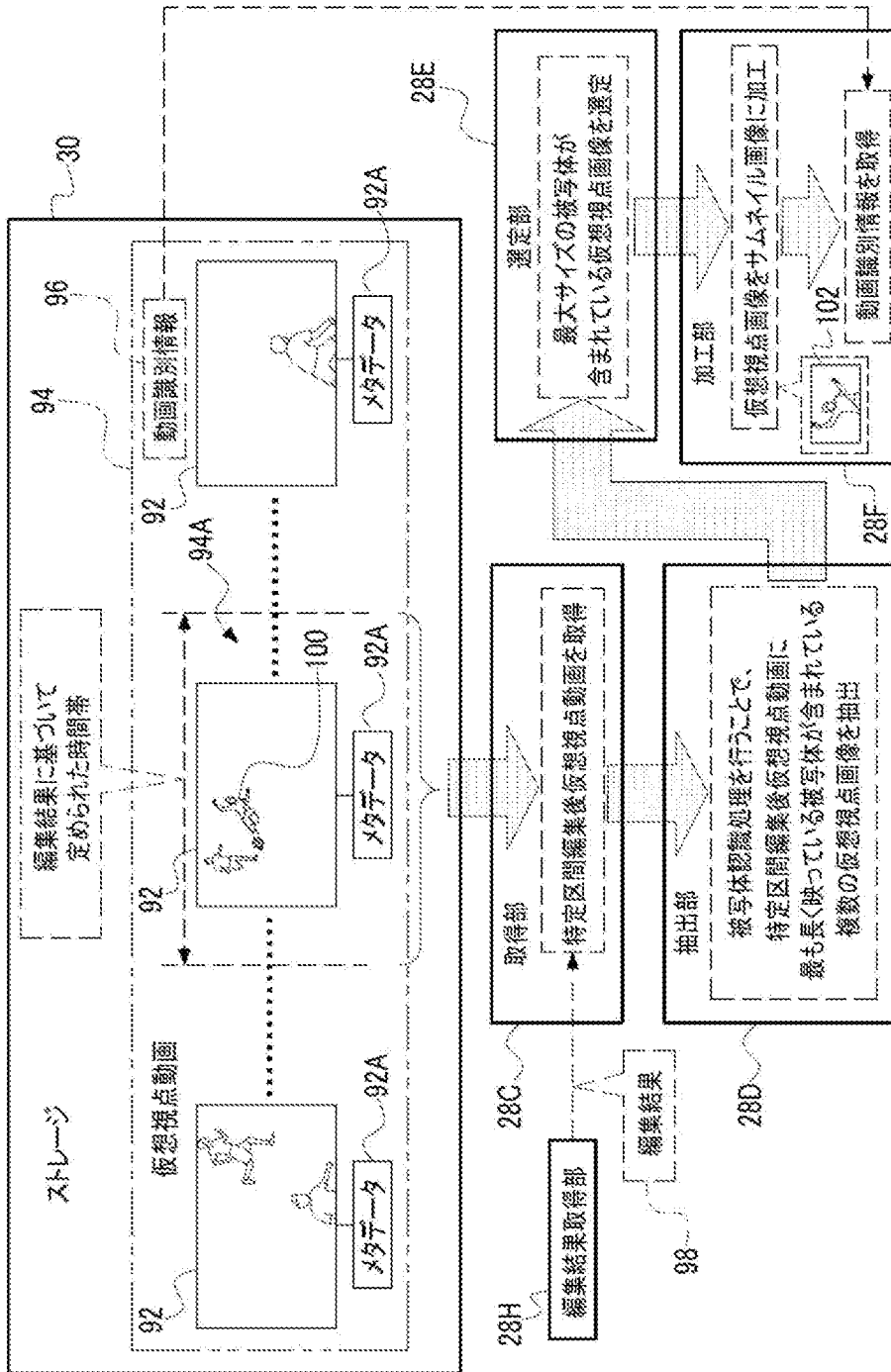
[図15]



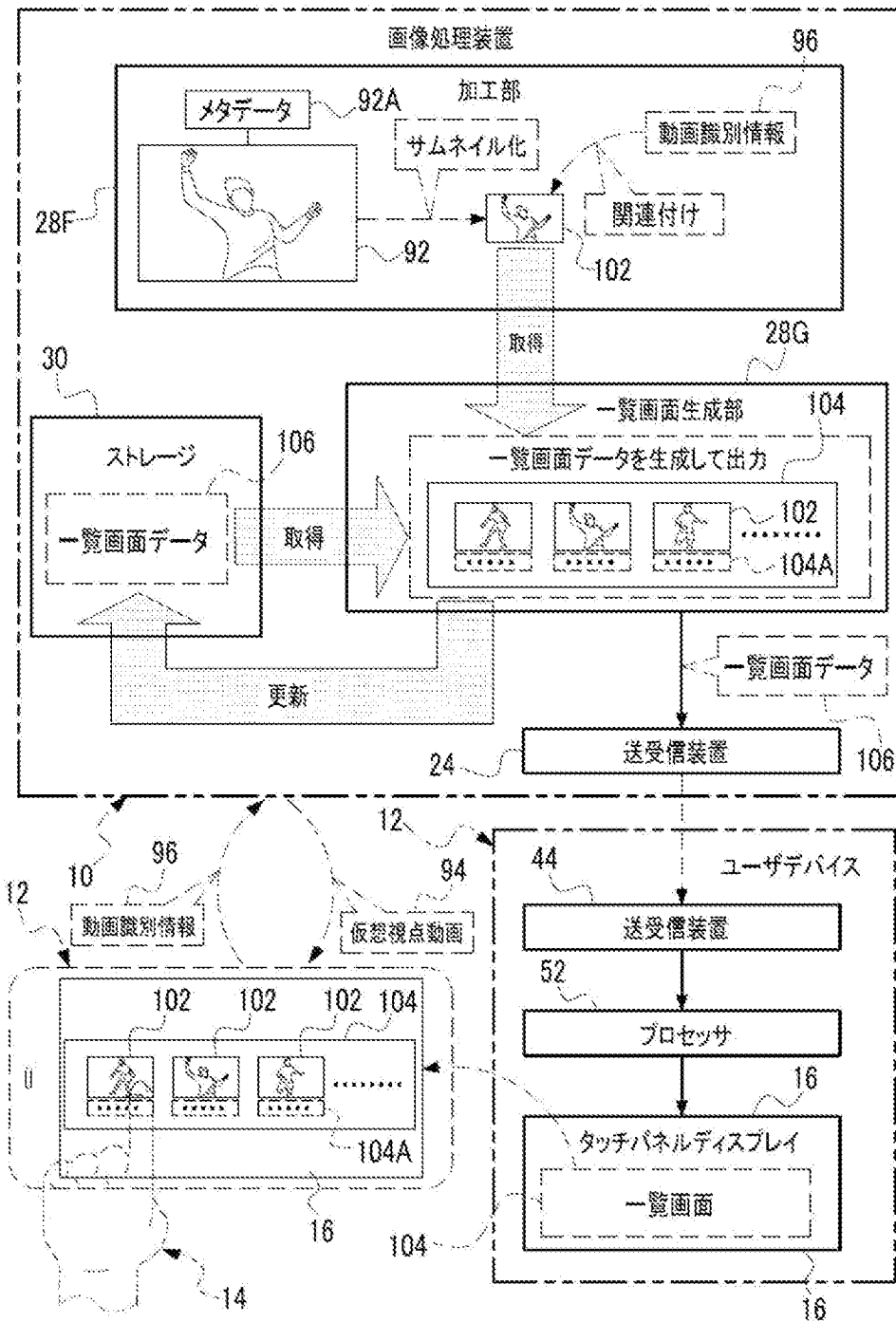
[図16]



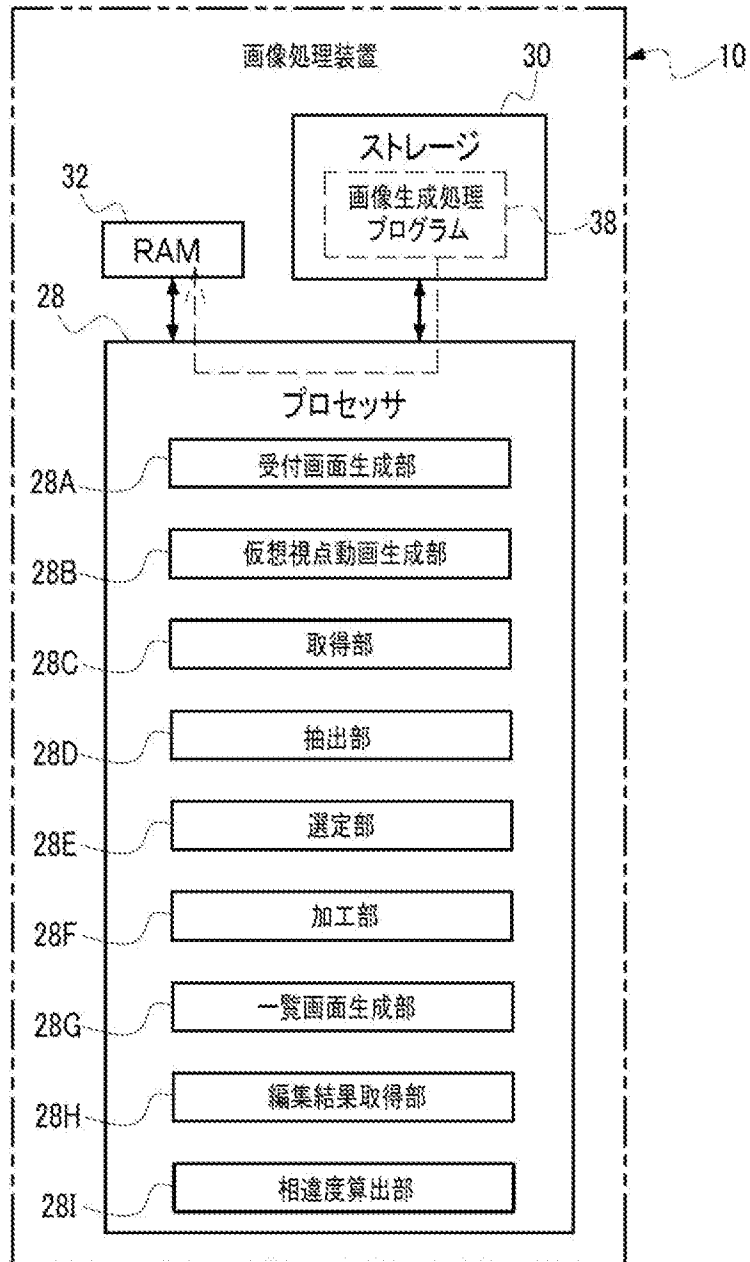
[図17]



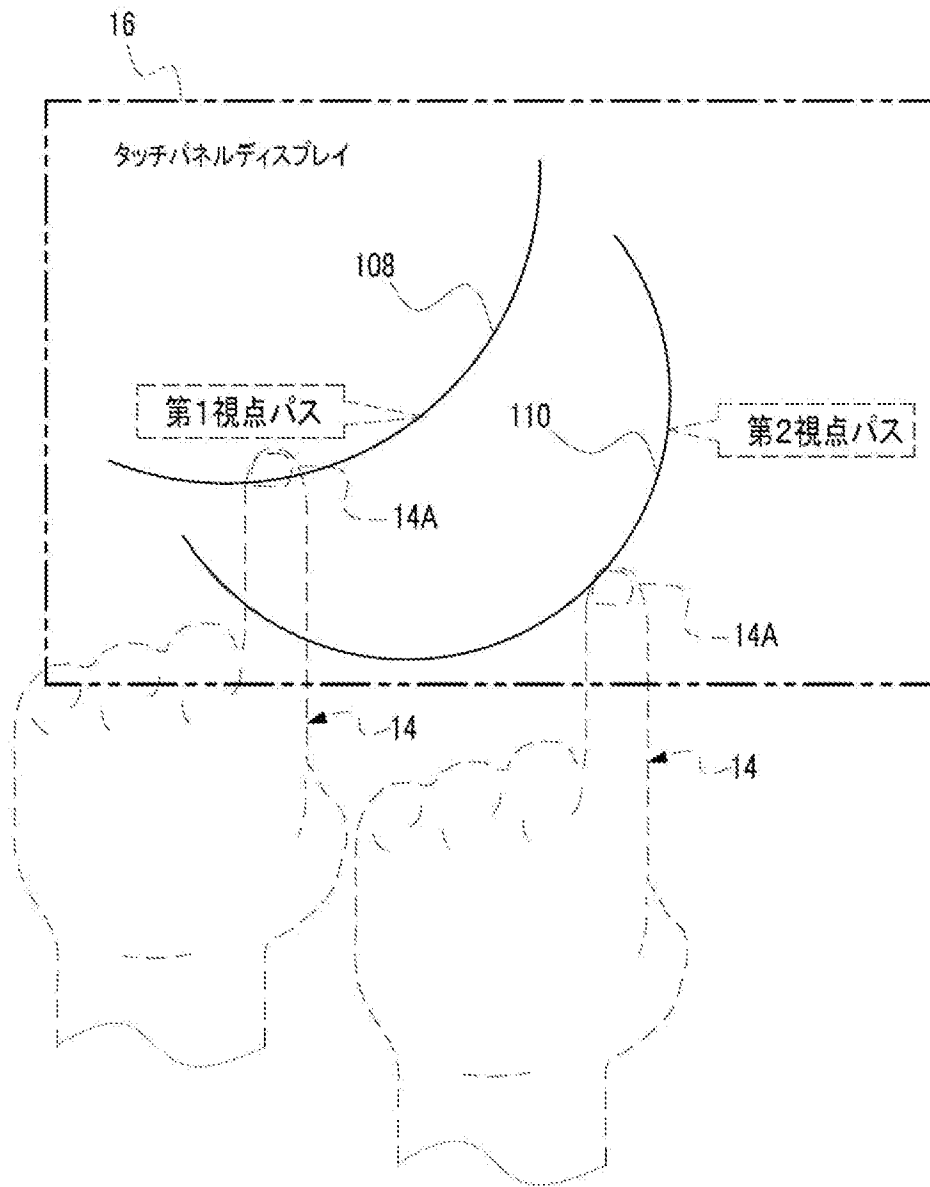
[図18]



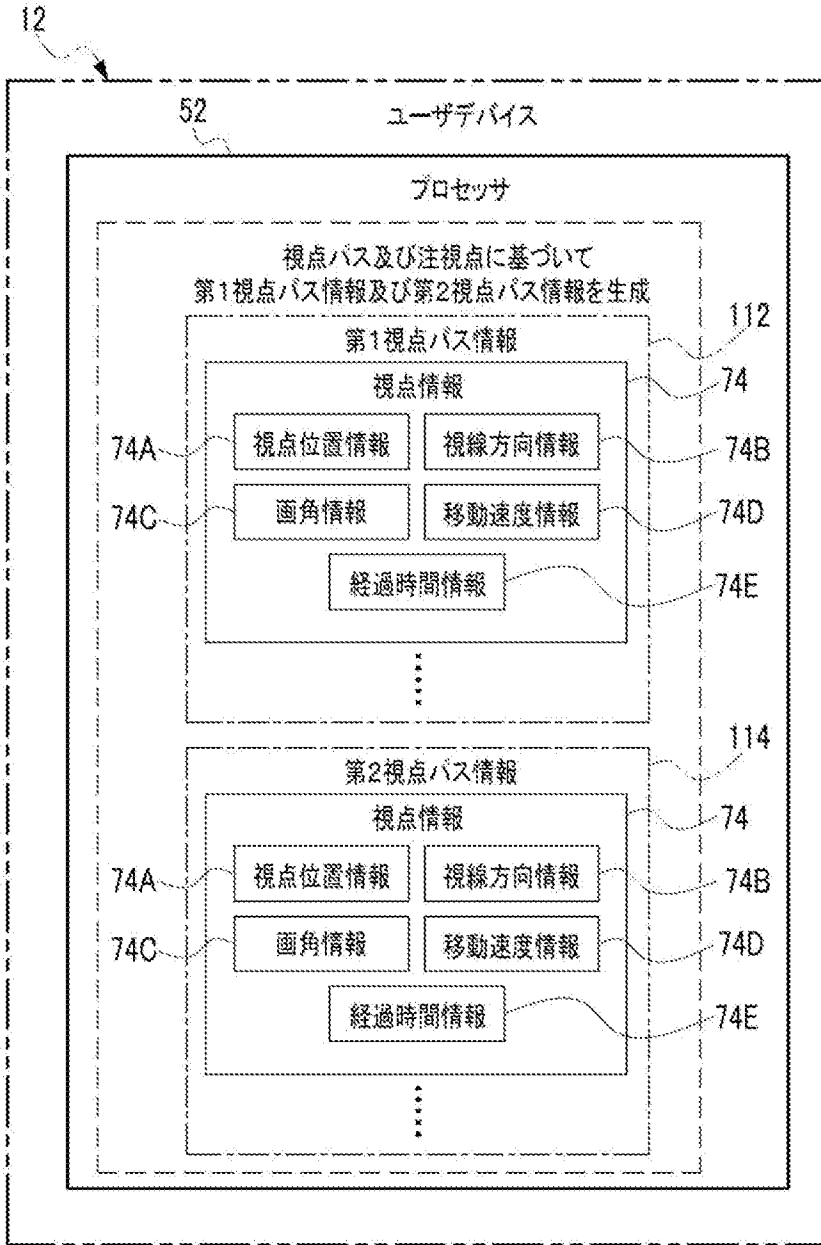
[図19]



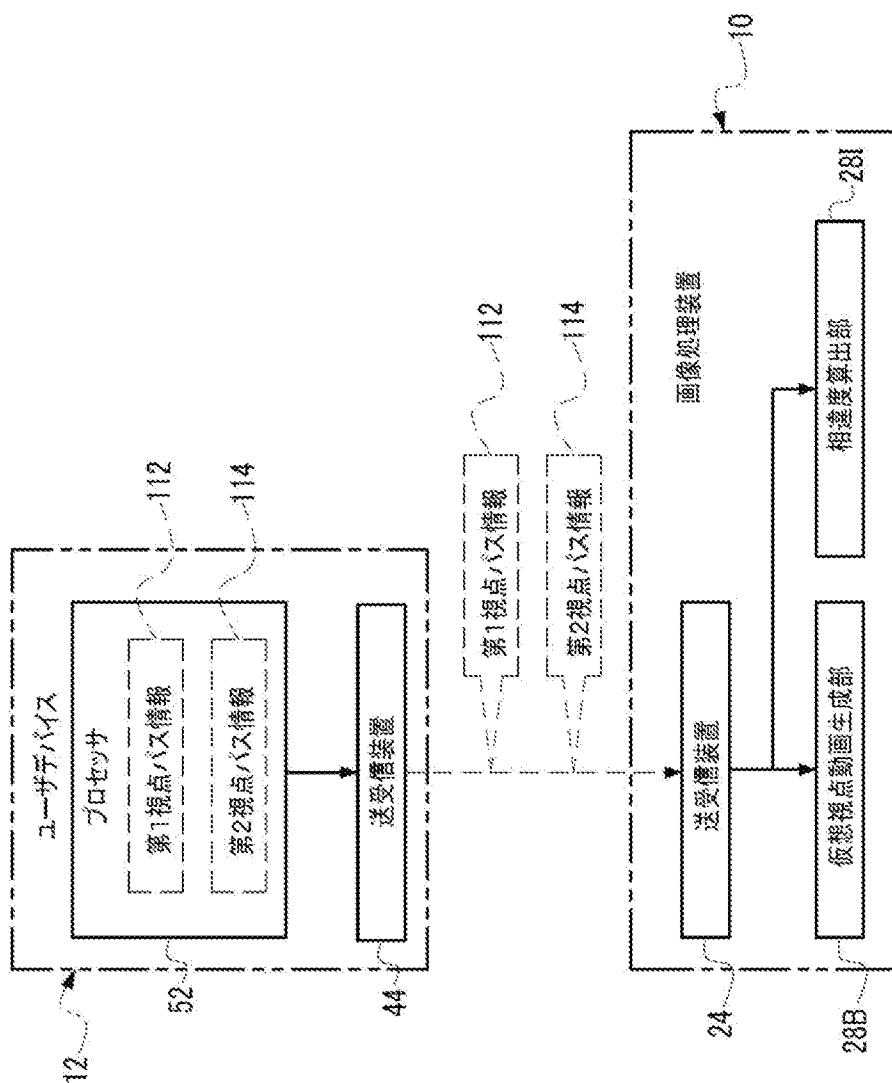
[図20]



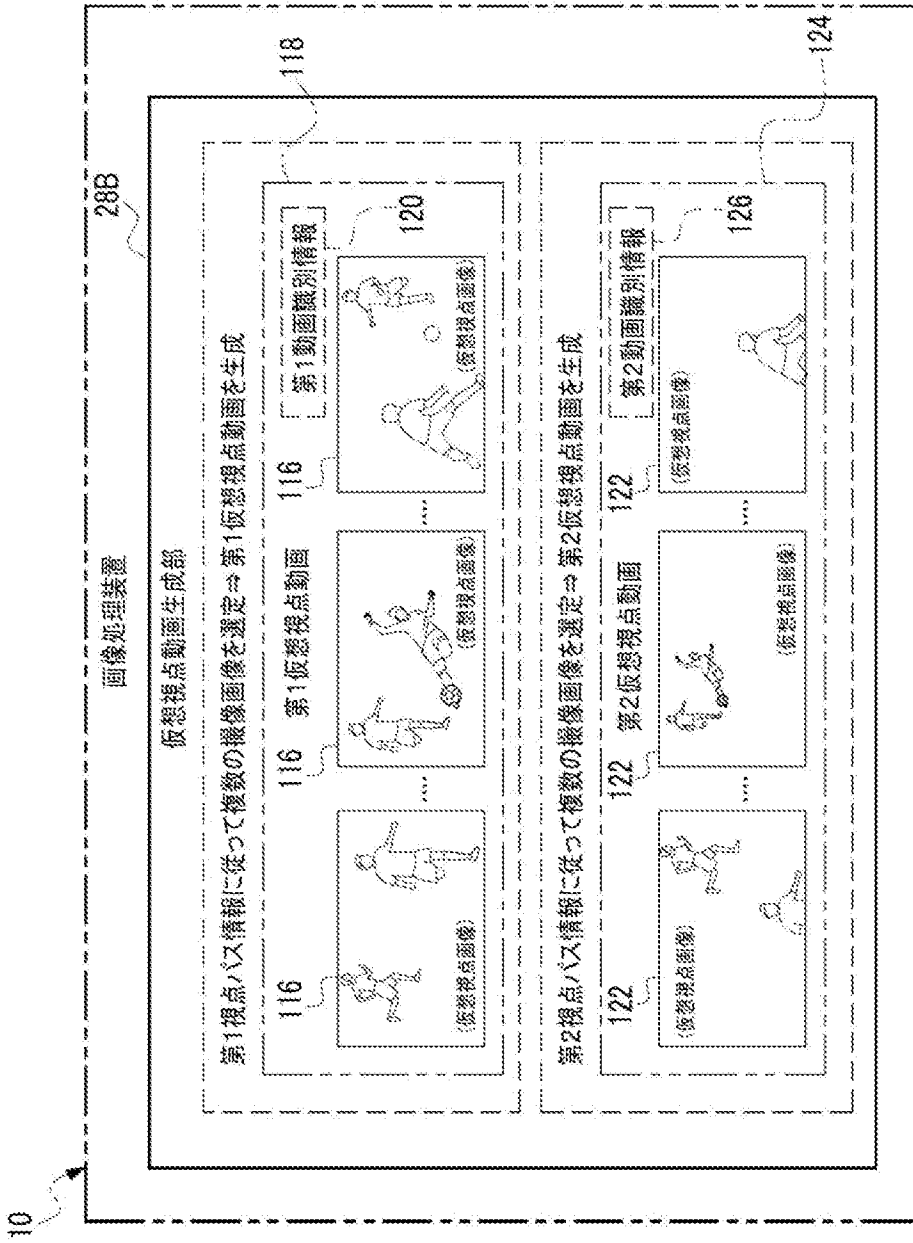
[図21]



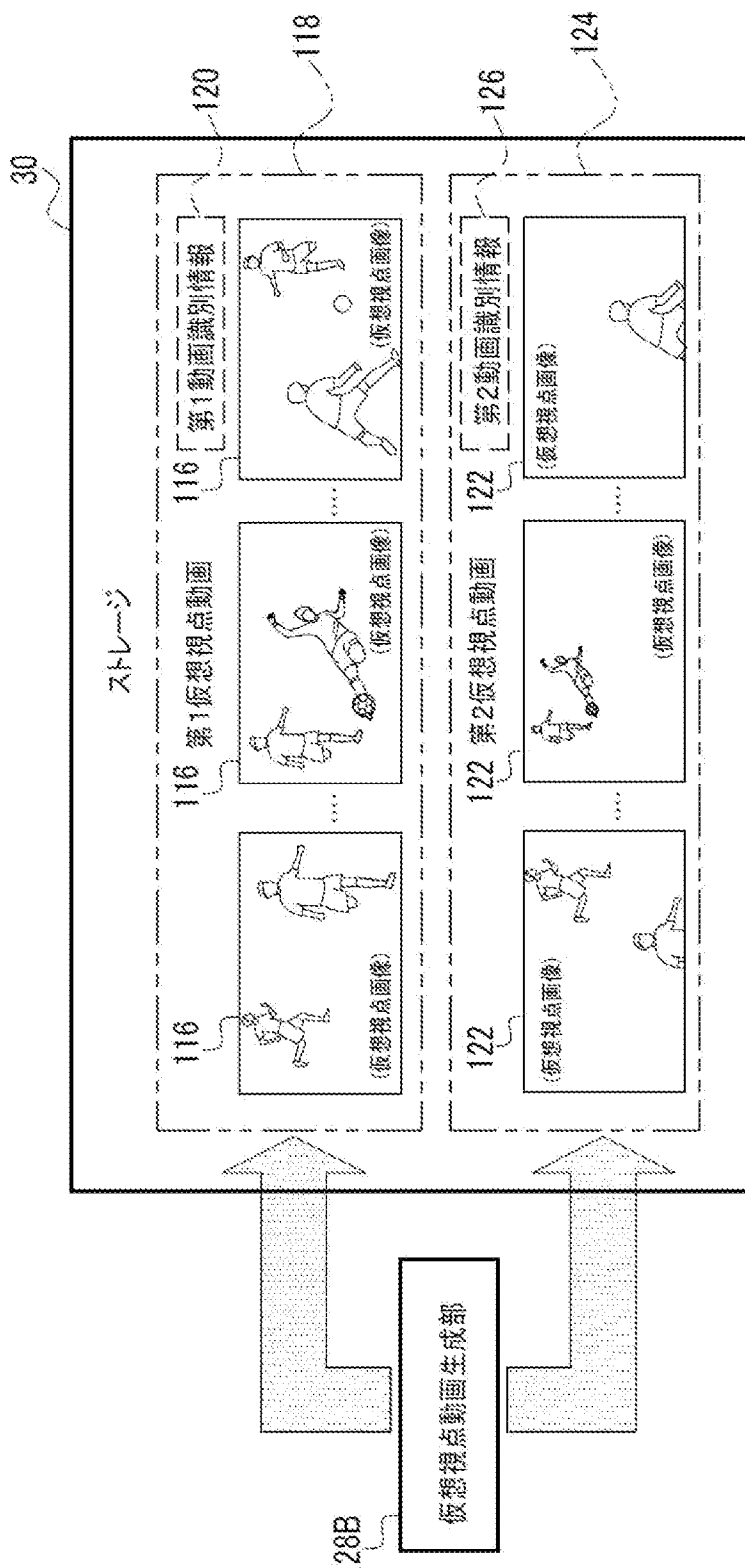
[図22]



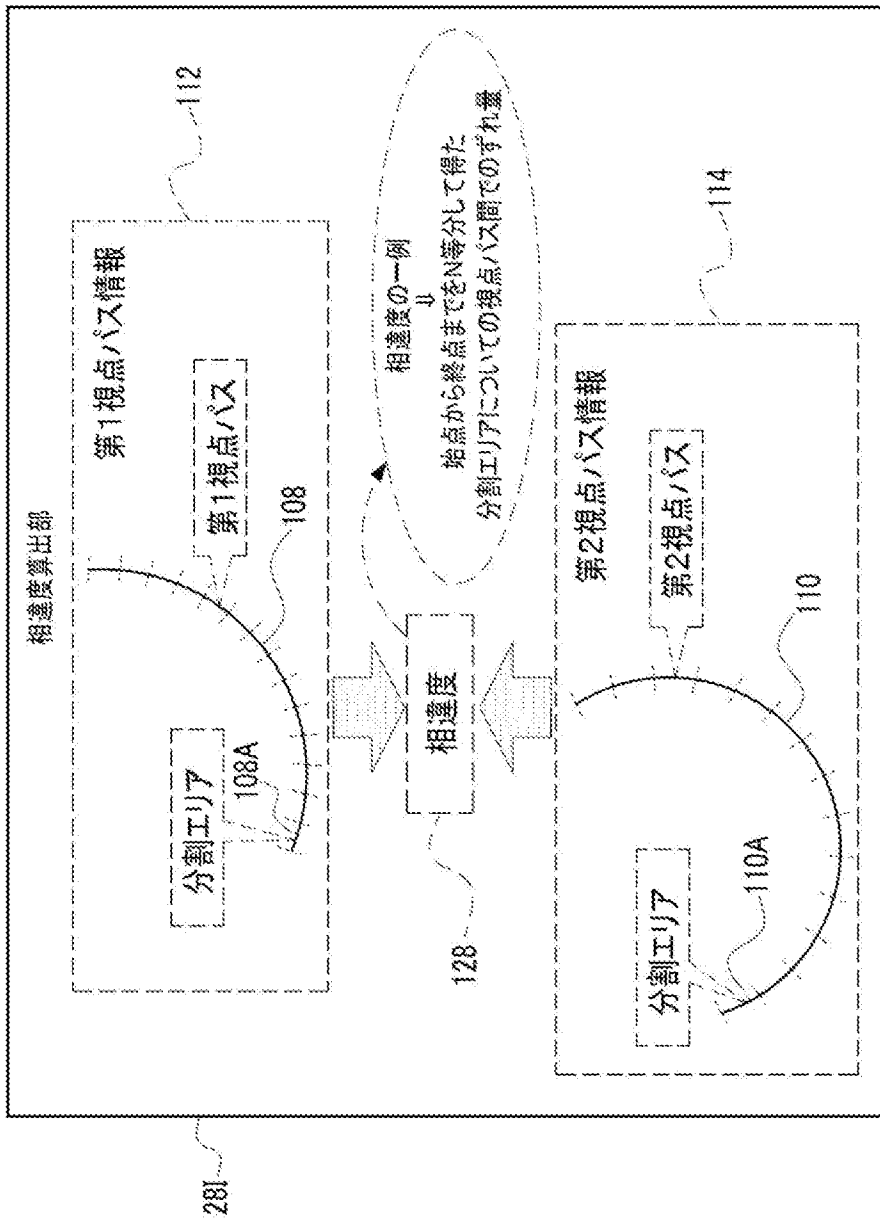
[図23]



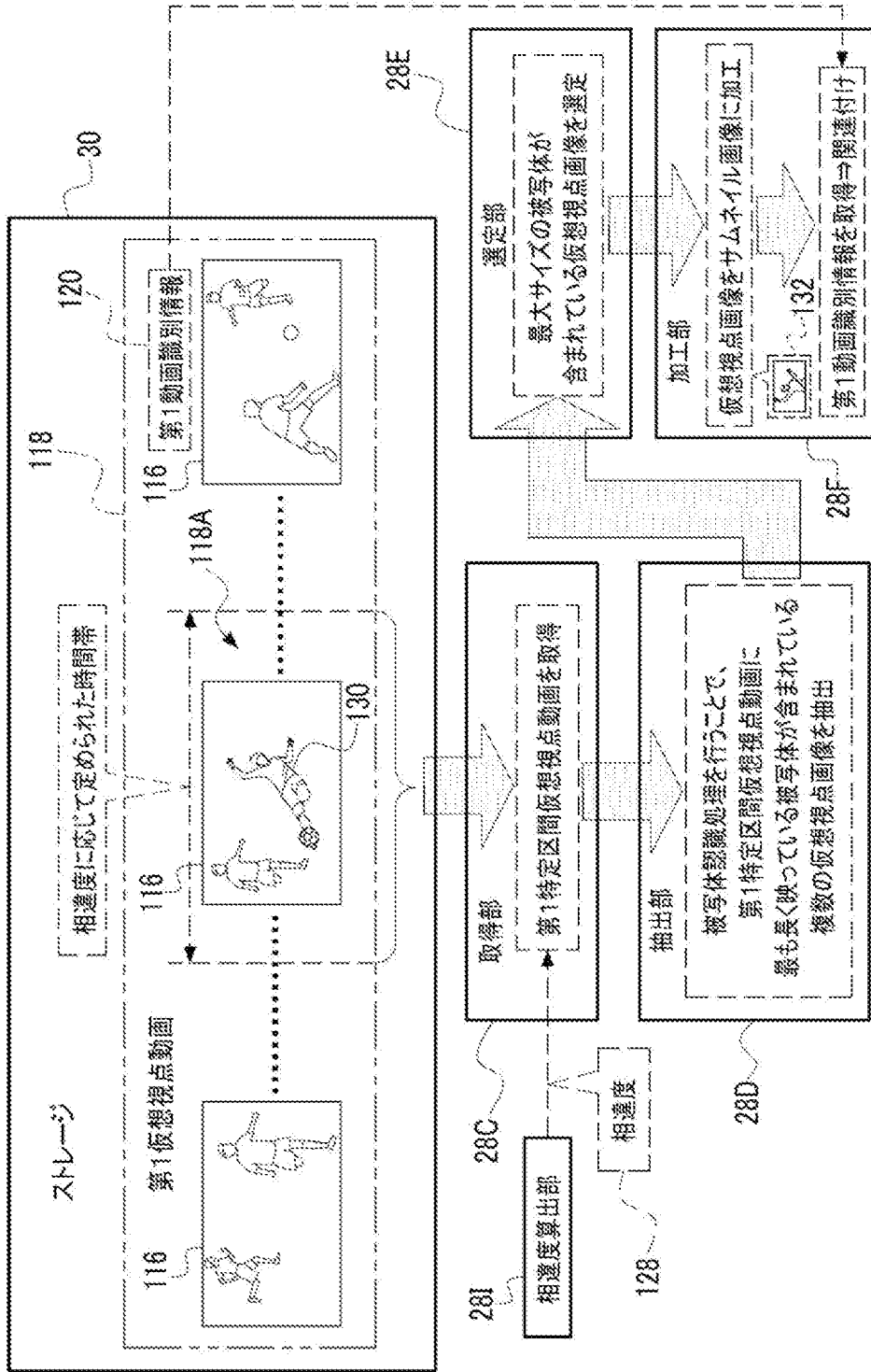
[図24]



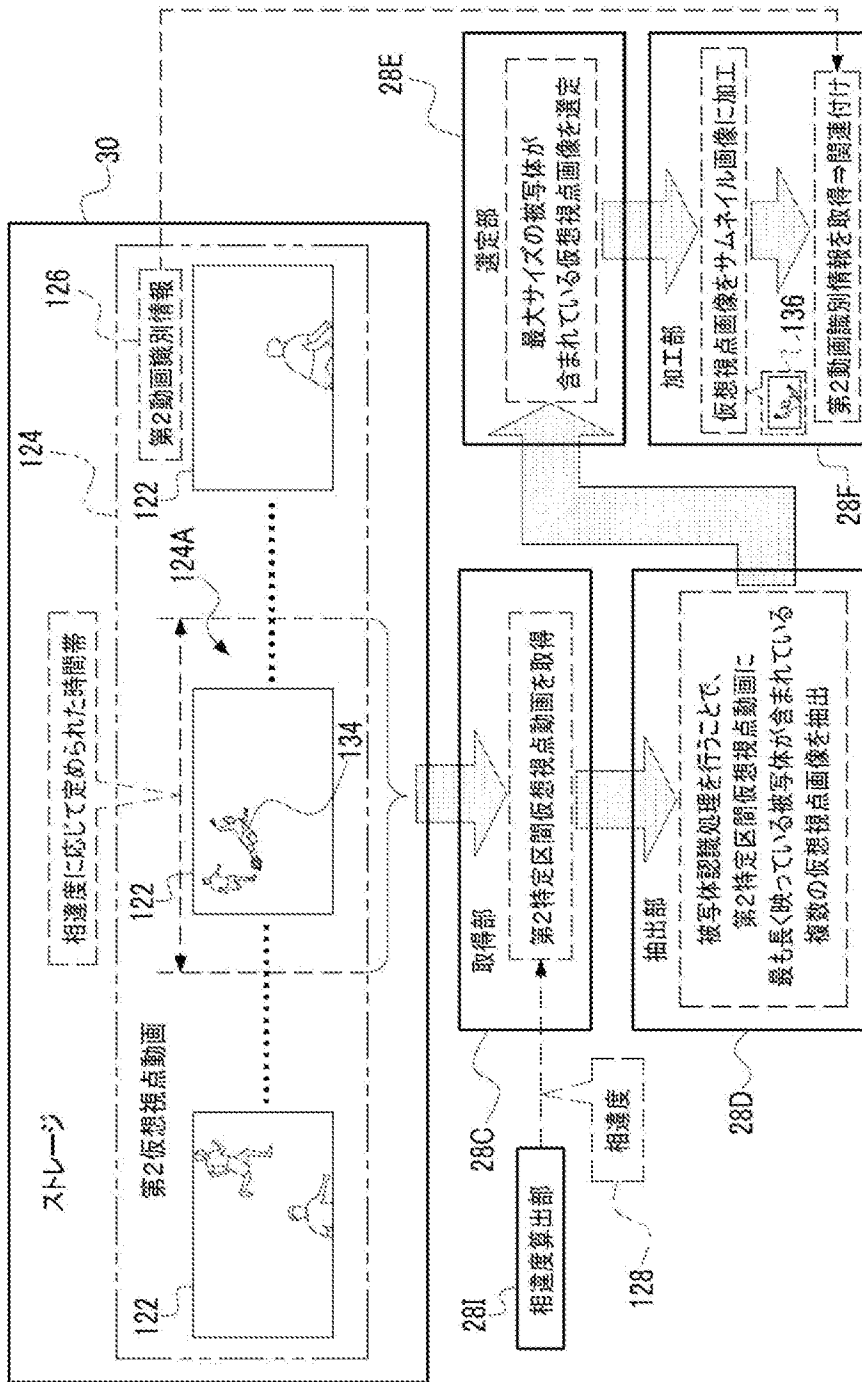
[図25]



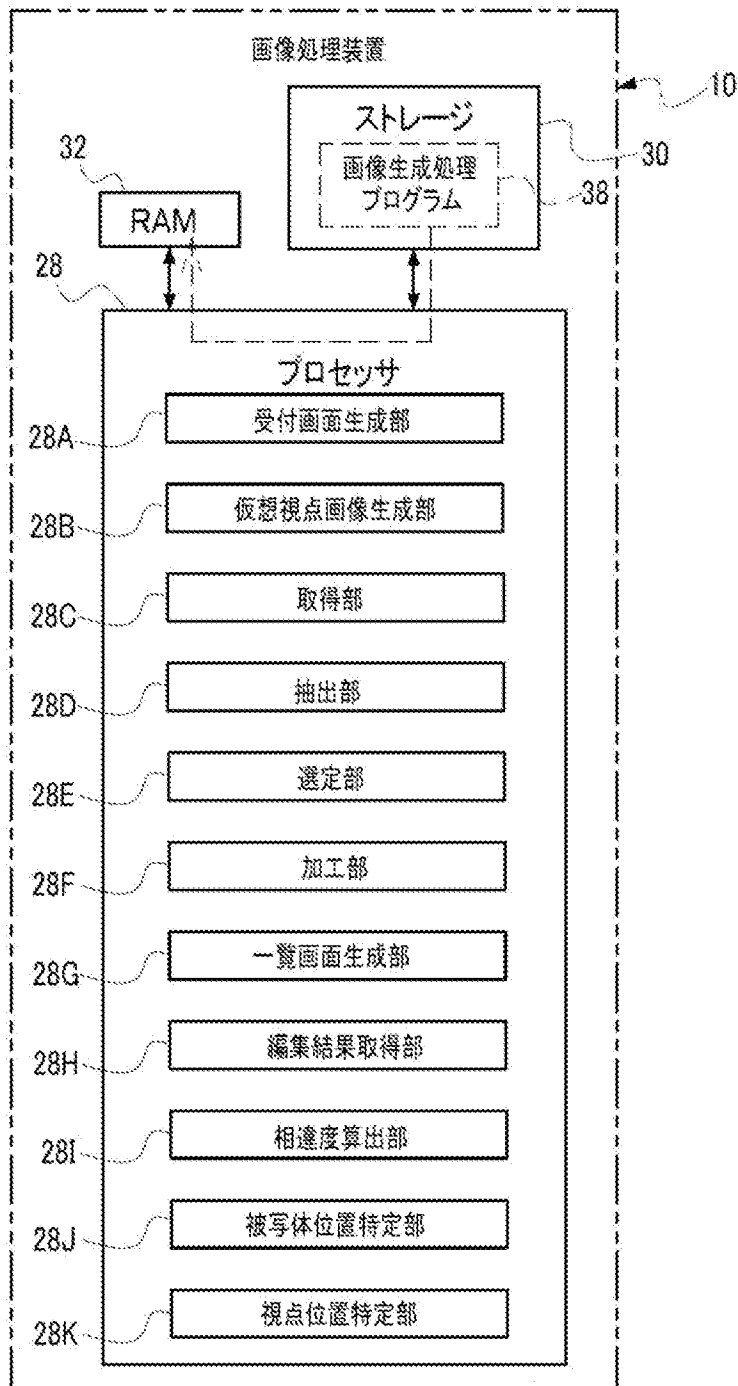
[図26]



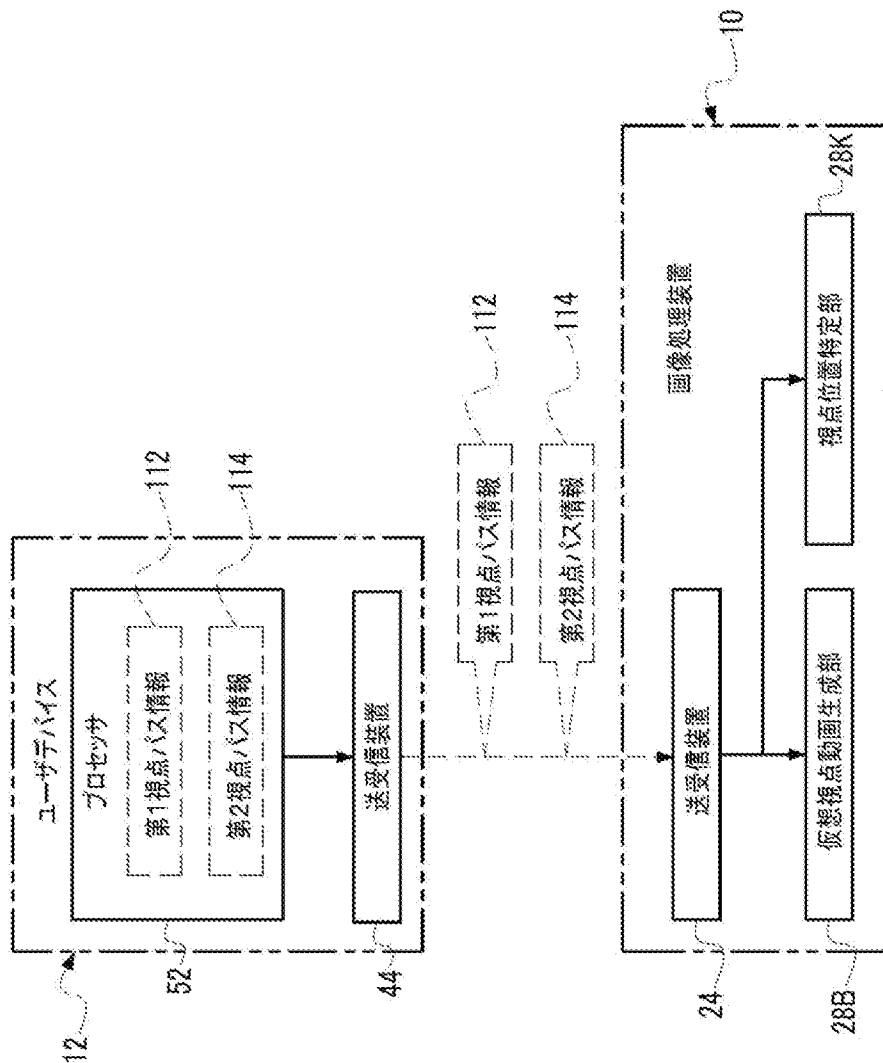
[図27]



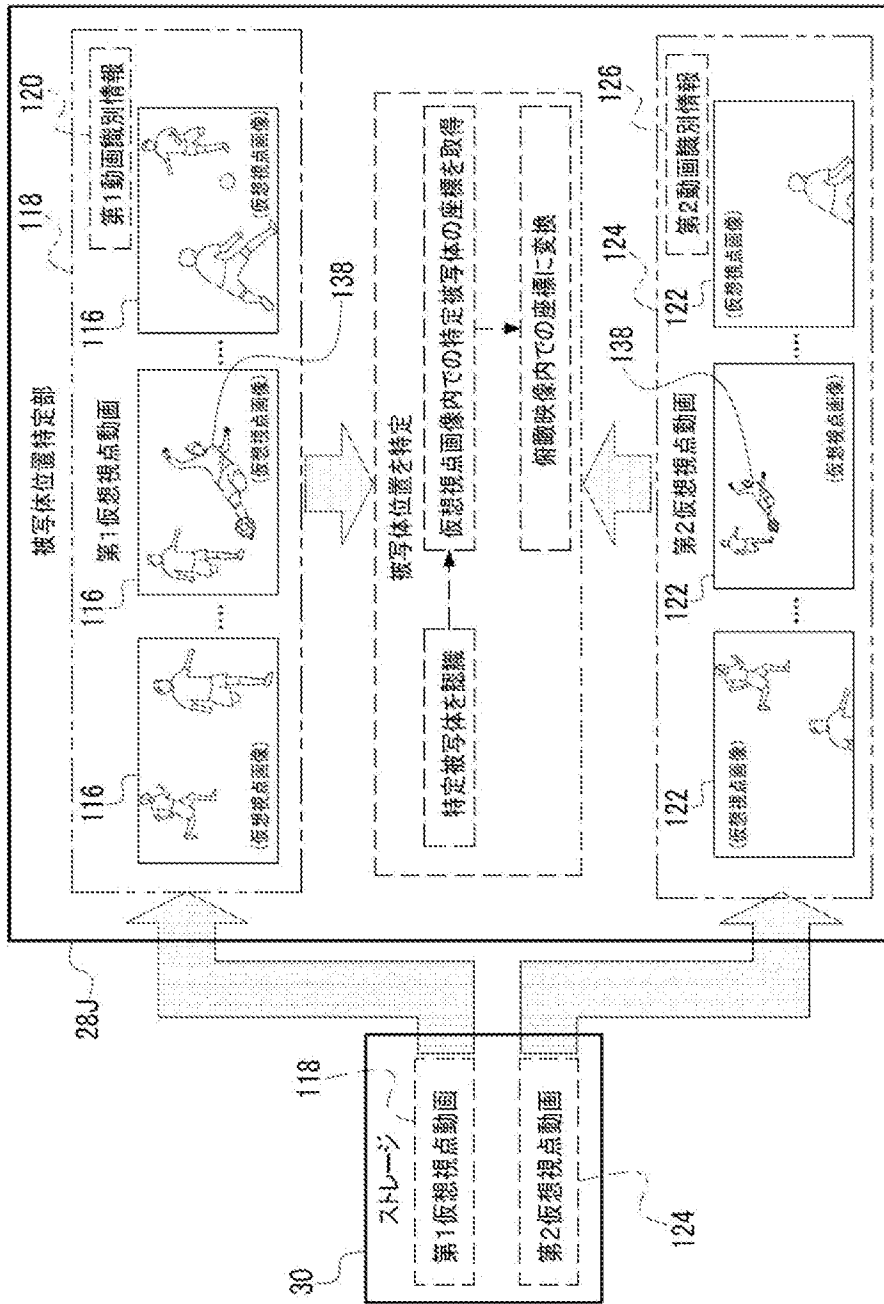
[図28]



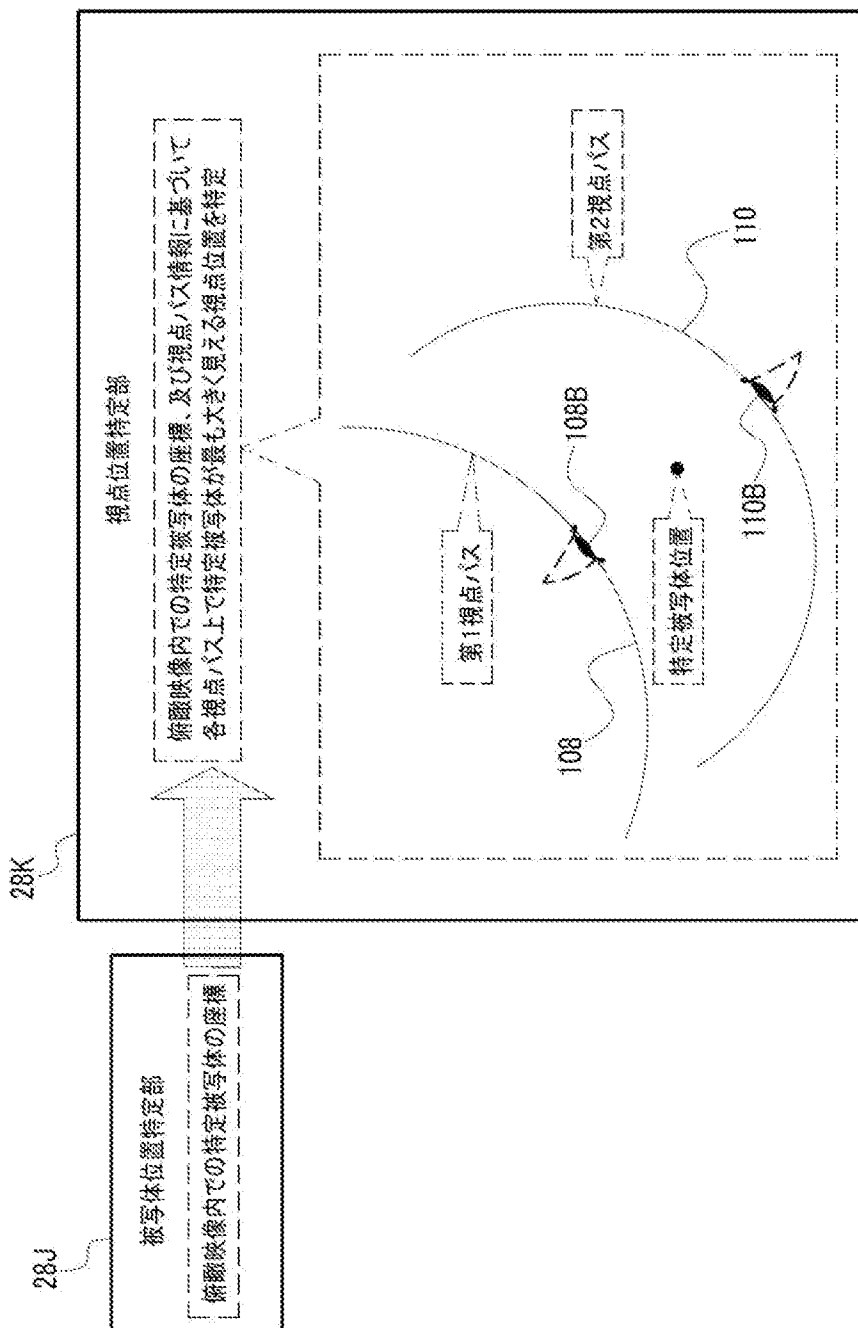
[図29]



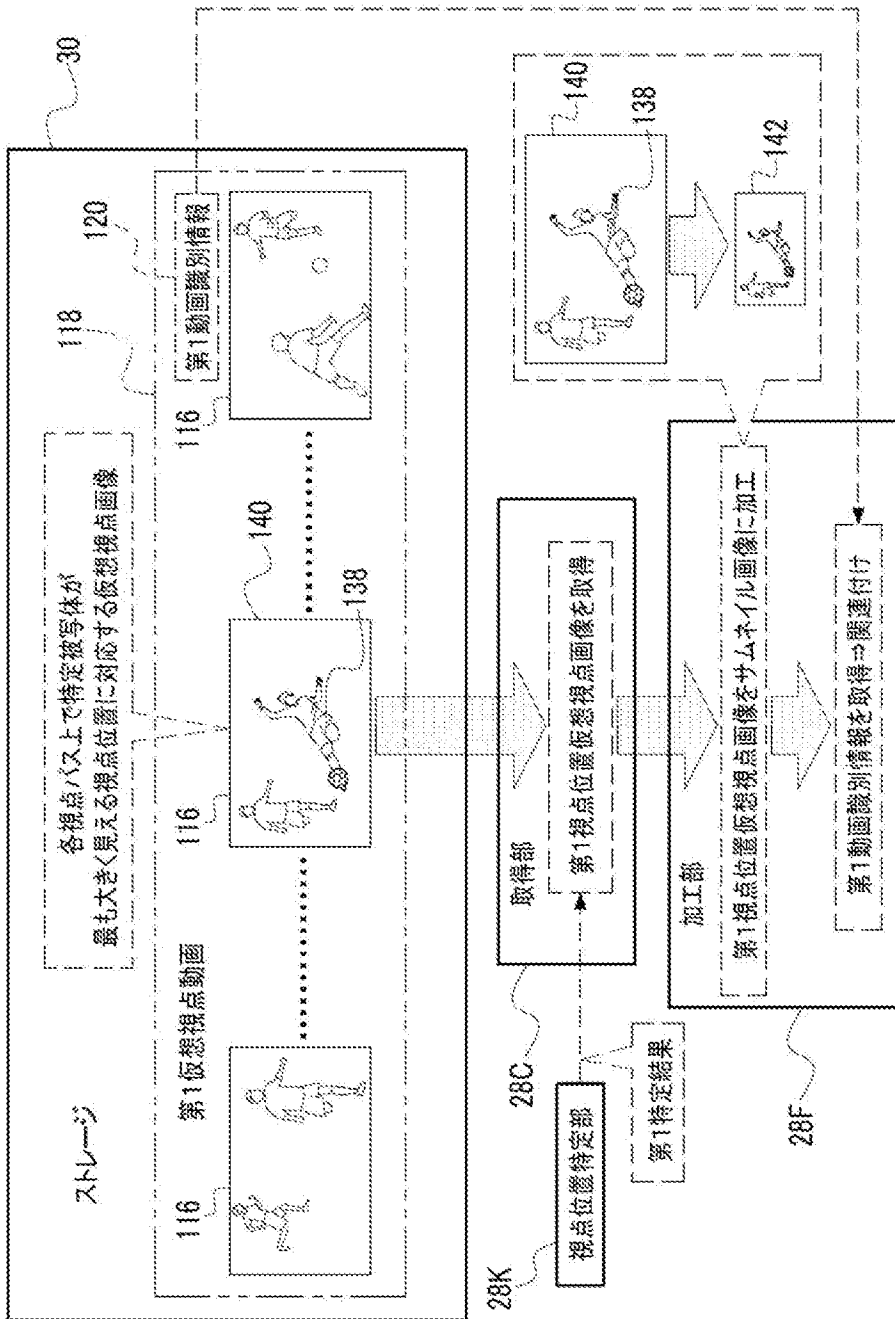
[図30]



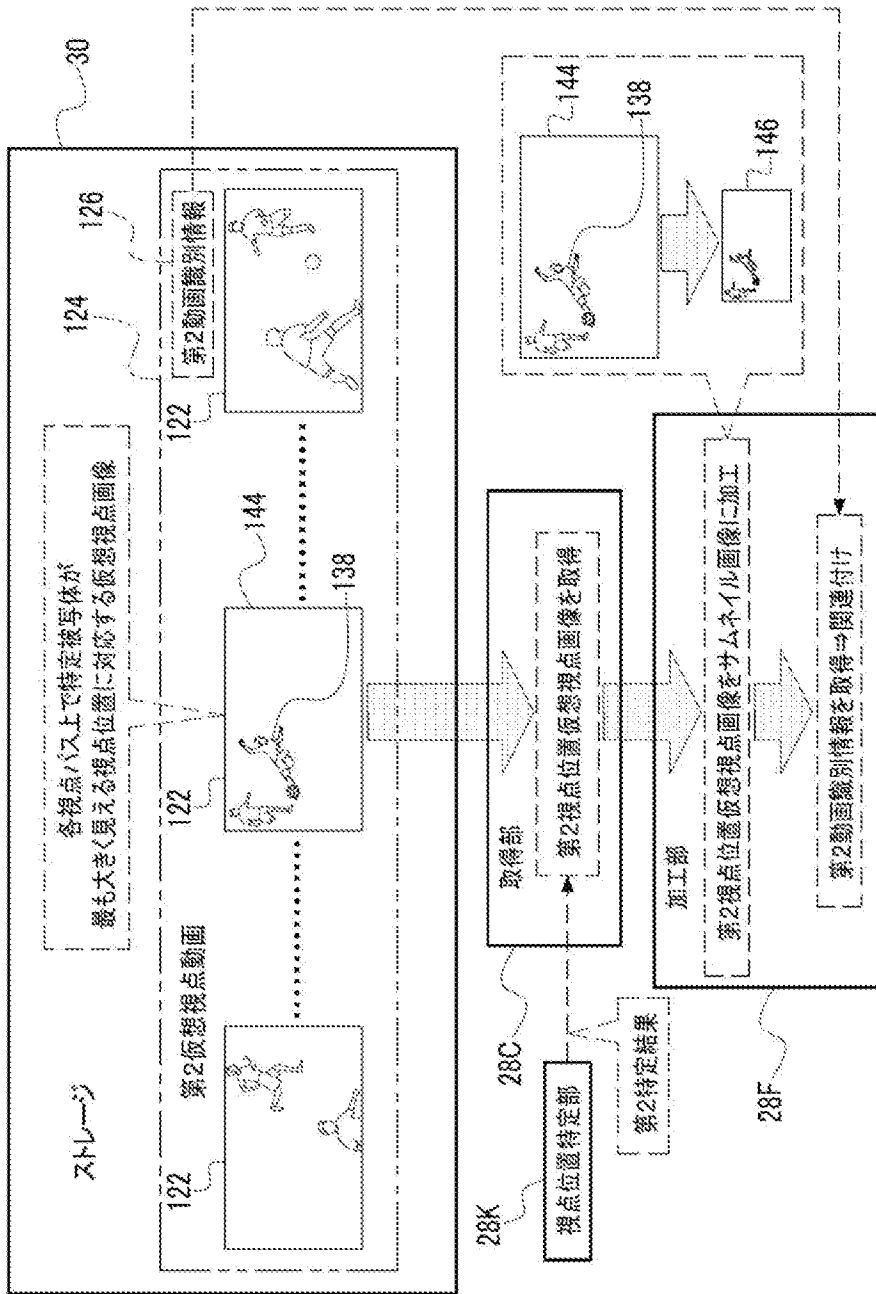
[図31]



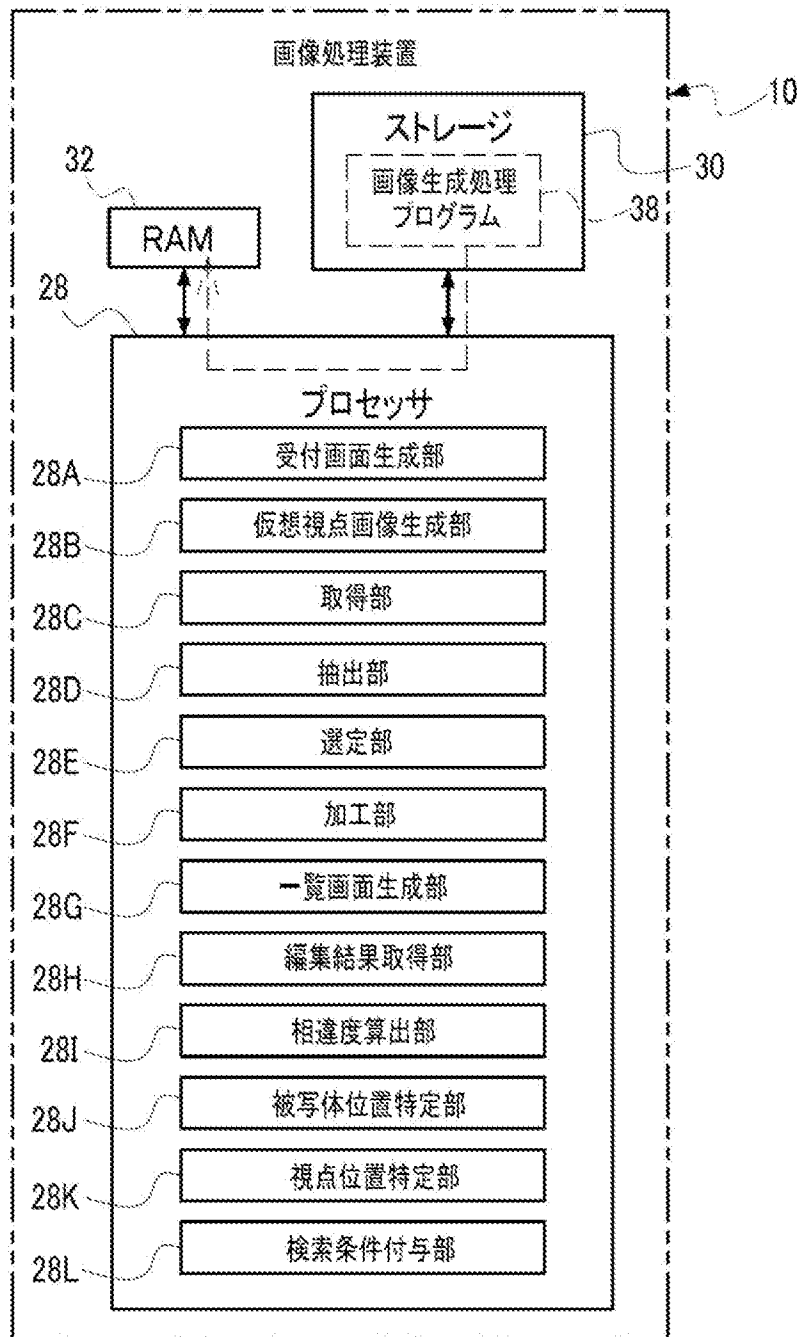
[図32]



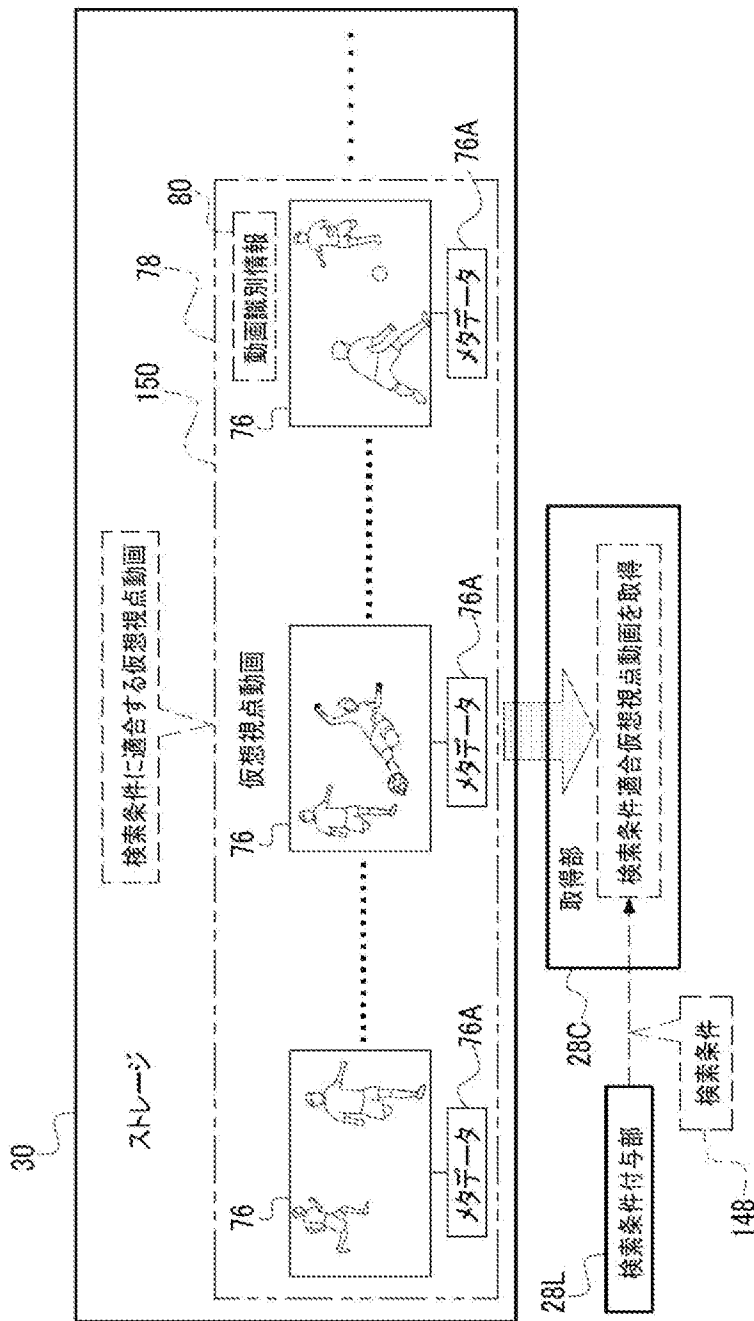
[図33]



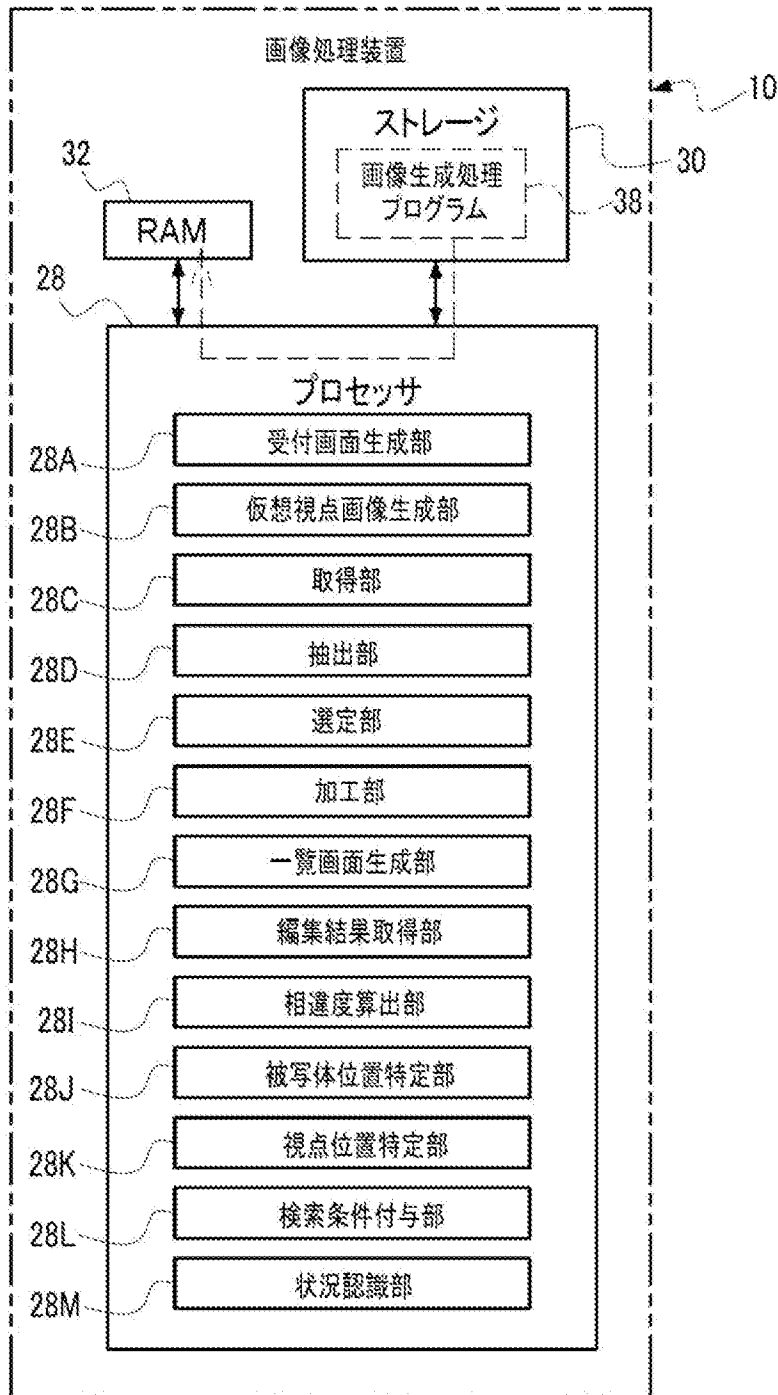
[図34]



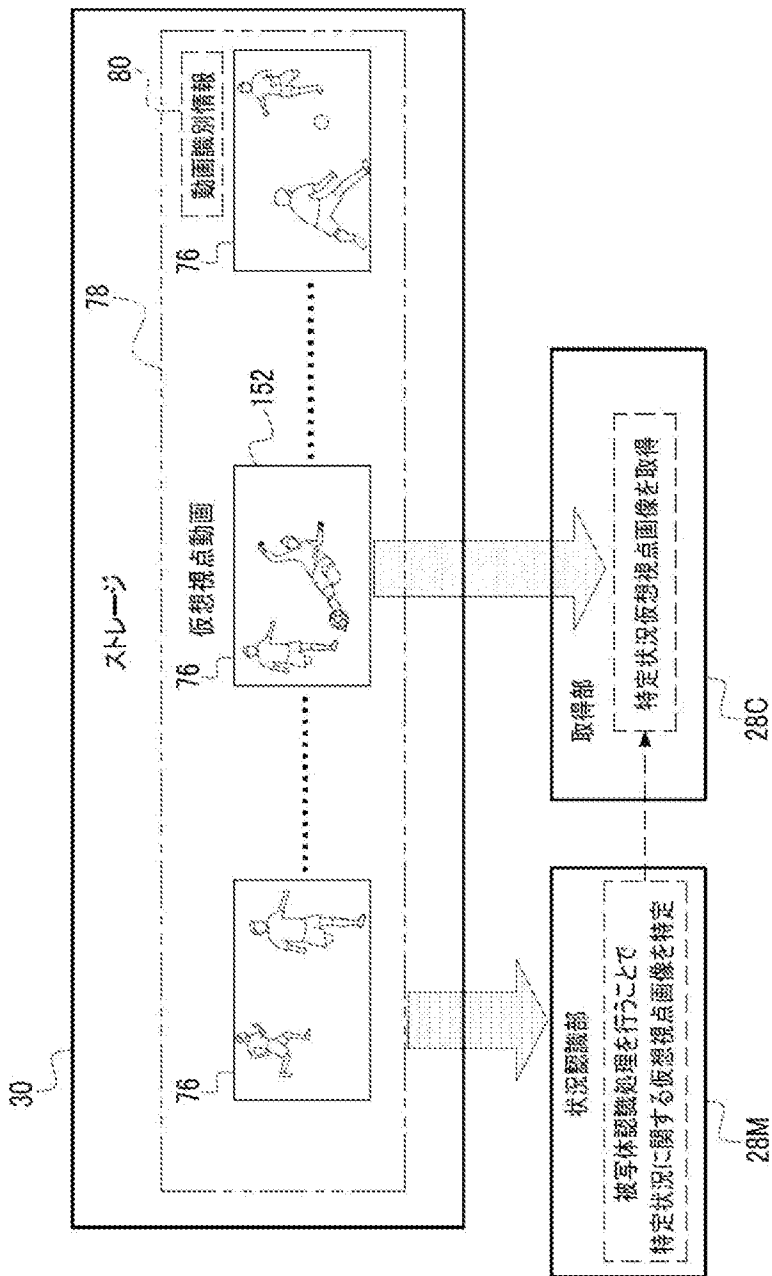
[図35]



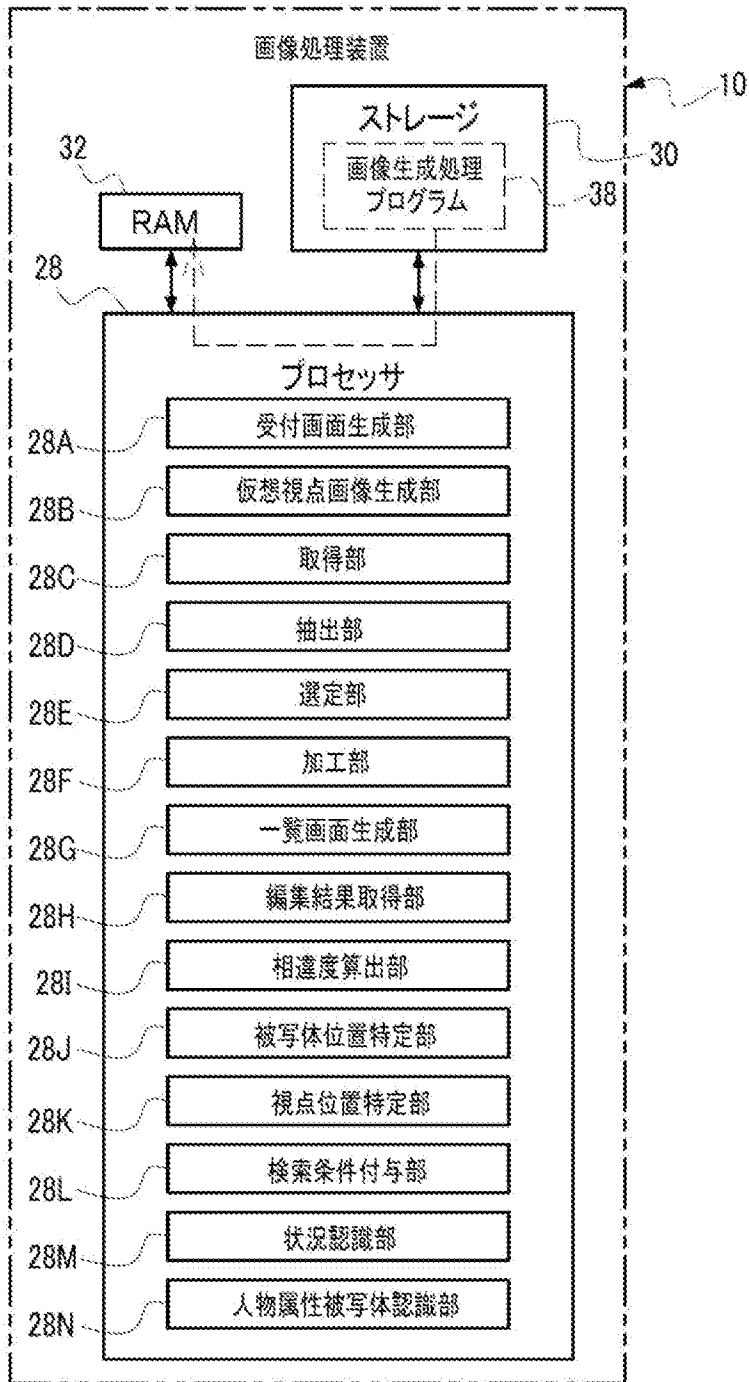
[図36]



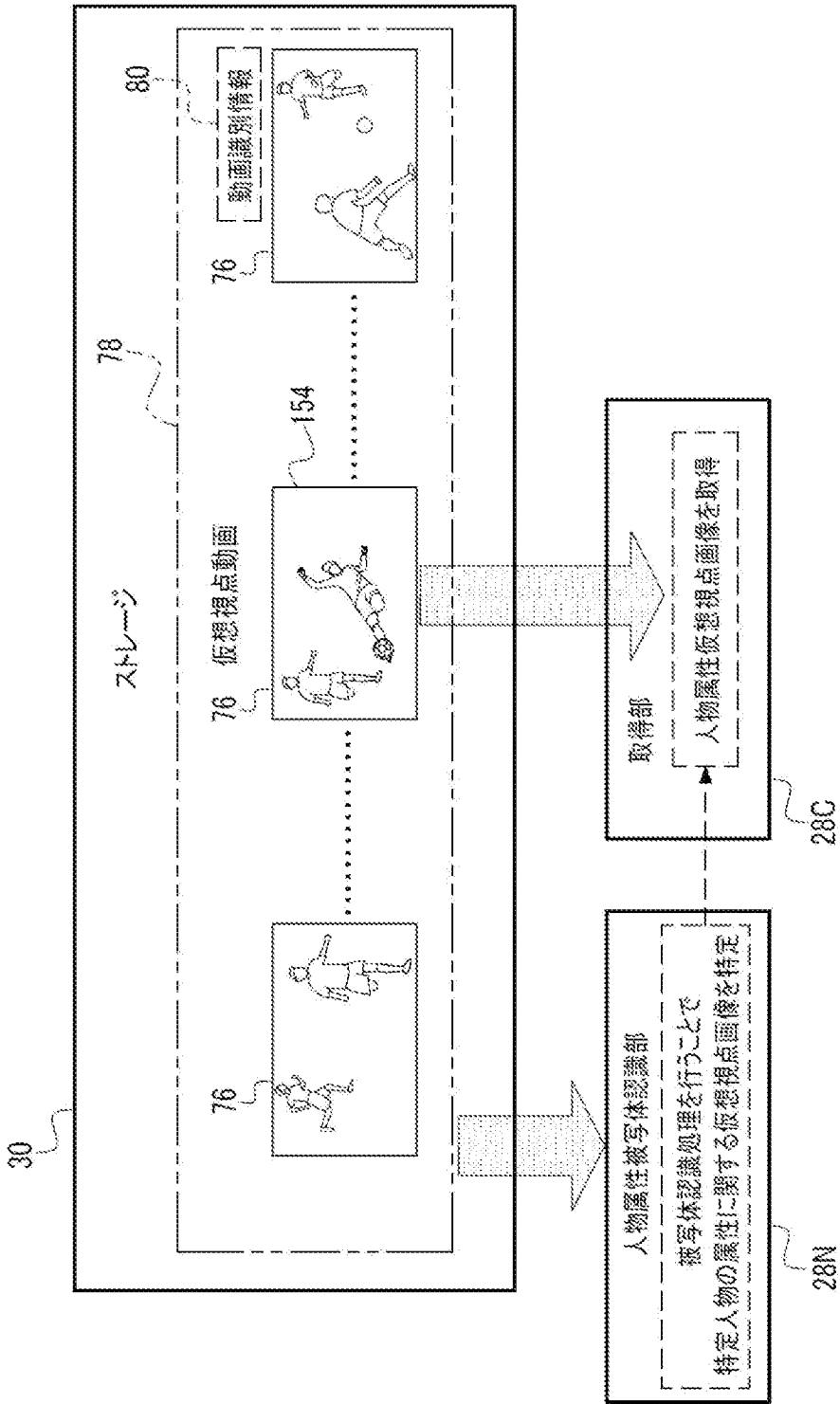
[図37]



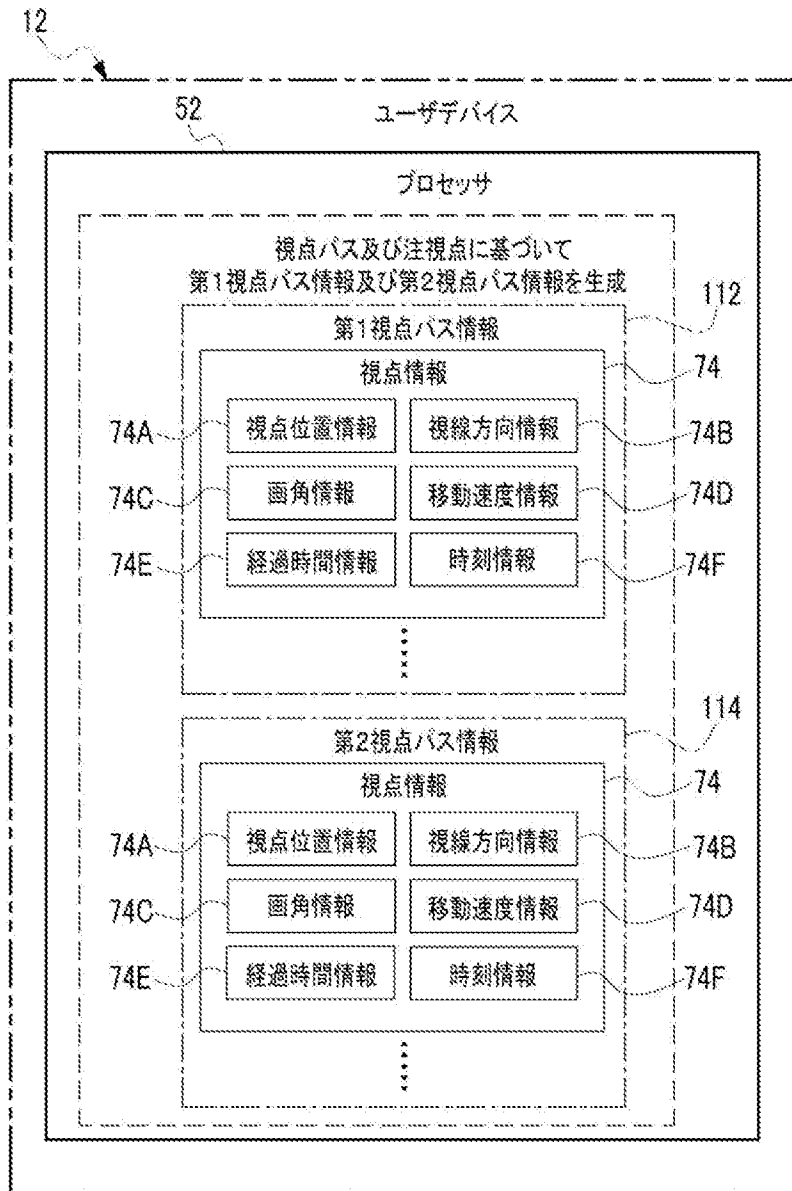
[図38]



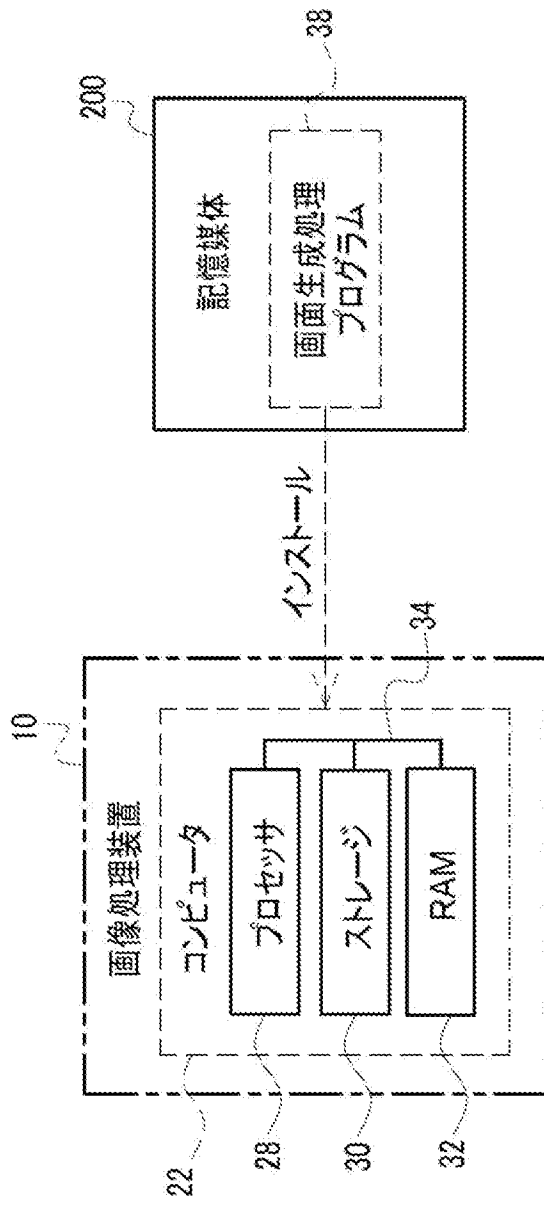
[図39]



[図40]



[図41]



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2022/005748

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
G06T 15/20(2011.01)i FI: G06T15/20 500		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols) G06T15/00-19/20; H04N5/222-5/28; H04N13/00-13/398		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Published examined utility model applications of Japan 1922-1996 Published unexamined utility model applications of Japan 1971-2022 Registered utility model specifications of Japan 1996-2022 Published registered utility model applications of Japan 1994-2022		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2018/051688 A1 (CANON INC.) 22 March 2018 (2018-03-22) paragraphs [0010]-[0051], fig. 3B, 6C	1, 5, 12, 14-20
A	paragraphs [0010]-[0051], fig. 3B, 6C	2-4, 6-11, 13
A	WO 2019/208702 A1 (PANASONIC INTELLECTUAL PROPERTY CORP. OF AMERICA) 31 October 2019 (2019-10-31) fig. 7B	1-20
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search <b>19 April 2022</b>		Date of mailing of the international search report <b>10 May 2022</b>
Name and mailing address of the ISA/JP <b>Japan Patent Office (ISA/JP) 3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915 Japan</b>		Authorized officer  Telephone No.

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No. <b>PCT/JP2022/005748</b>
---

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
WO	2018/051688	A1	22 March 2018	US	2019/0213791	A1	paragraphs [0053]-[0094], fig. 3B, 6C JP 2018-46448 A KR 10-2019-0039774 A CN 109716751 A
WO	2019/208702	A1	31 October 2019	US	2021/0044793	A1	fig. 7B

A. 発明の属する分野の分類（国際特許分類（IPC）） G06T 15/20(2011.01)i FI: G06T15/20 500		
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）） G06T15/00-19/20; H04N5/222-5/28; H04N13/00-13/398 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 1971-2022年 日本国実用新案登録公報 1996-2022年 日本国登録実用新案公報 1994-2022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
X	WO 2018/051688 A1 (キヤノン株式会社) 22.03.2018 (2018-03-22) 段落[0010]-[0051], 図3B, 図6C	1, 5, 12, 14-20
A	段落[0010]-[0051], 図3B, 図6C	2-4, 6-11, 13
A	WO 2019/208702 A1 (パナソニック インテレクチュアル プロパティ コーポレーシ ン オブ アメリカ) 31.10.2019 (2019-10-31) 図7B	1-20
<input type="checkbox"/> C欄の続きにも文献が列挙されている。 <input checked="" type="checkbox"/> 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抵 触するものではなく、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 19.04.2022	国際調査報告の発送日 10.05.2022	
名称及びあて先 日本国特許庁 (ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号	権限のある職員（特許庁審査官） 村松 貴士 5V 9854 電話番号 03-3581-1101 内線 3569	

国際調査報告  
パテントファミリーに関する情報

国際出願番号

PCT/JP2022/005748

引用文献	公表日	パテントファミリー文献	公表日
WO 2018/051688 A1	22.03.2018	US 2019/0213791 A1 段落[0053]-[0094], 図3B, 図6C JP 2018-46448 A KR 10-2019-0039774 A CN 109716751 A	
WO 2019/208702 A1	31.10.2019	US 2021/0044793 A1 図7B	