

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号
特開2017-102784
(P2017-102784A)

(43) 公開日 平成29年6月8日(2017. 6. 8)

(51) Int.Cl.	F I	テーマコード (参考)
G06T 3/00 (2006.01)	G O 6 T 3/00 7 8 0	5 B 0 5 7
H04N 5/232 (2006.01)	H O 4 N 5/232 Z	5 C 1 2 2
H04N 5/225 (2006.01)	H O 4 N 5/225 Z	

審査請求 未請求 請求項の数 8 O L (全 26 頁)

(21) 出願番号	特願2015-236706 (P2015-236706)	(71) 出願人	000004226
(22) 出願日	平成27年12月3日 (2015. 12. 3)		日本電信電話株式会社
			東京都千代田区大手町一丁目5番1号
		(74) 代理人	110001634
			特許業務法人 志賀国際特許事務所
		(72) 発明者	高橋 康輔
			東京都千代田区大手町一丁目5番1号 日
			本電信電話株式会社内
		(72) 発明者	三上 弾
			東京都千代田区大手町一丁目5番1号 日
			本電信電話株式会社内
		(72) 発明者	五十川 麻理子
			東京都千代田区大手町一丁目5番1号 日
			本電信電話株式会社内
		最終頁に続く	

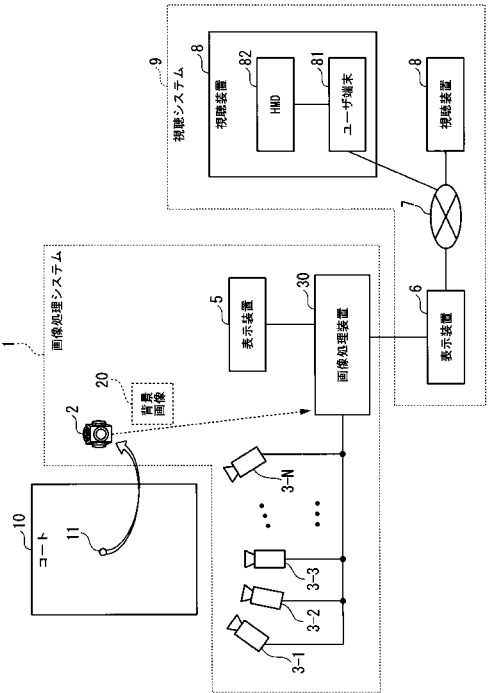
(54) 【発明の名称】 画像処理装置、画像処理方法及び画像処理プログラム

(57) 【要約】

【課題】 視聴品質の低下を抑制した仮想全天球画像を生成する画像処理装置を提供する。

【解決手段】 第1の映像から取得された画像である第1画像と、第2の映像から取得された画像であって前記第1画像と撮像領域の一部が共通する画像である第2画像との2つの画像を結合する際に、共通する部分において前記2つの画像を結合するための結合線を決定する画像処理装置であって、前記画像を見ている者が見ている可能性が高い領域を避けつつ、前記結合線を配置することを決定する結合線決定手段を備える。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

第 1 の映像から取得された画像である第 1 画像と、第 2 の映像から取得された画像であって前記第 1 画像と撮像領域の一部が共通する画像である第 2 画像との 2 つの画像を結合する際に、共通する部分において前記 2 つの画像を結合するための結合線を決定する画像処理装置であって、

前記画像を見ている者が見ている可能性が高い領域を避けつつ、前記結合線を配置することを決定する結合線決定手段を備える画像処理装置。

【請求項 2】

第 1 の映像から取得された画像である第 1 画像と、第 2 の映像から取得された画像であって前記第 1 画像と撮像領域の一部が共通する画像である第 2 画像との 2 つの画像を結合する際に、共通する部分において前記 2 つの画像を結合するための結合線を決定する画像処理装置であって、

前記共通する部分において、人間の目が敏感に感じる画素はそのエネルギーが大きいと定義される第 1 のエネルギー画像を生成する第 1 のエネルギー画像生成手段と、

前記第 1 のエネルギー画像のエネルギーが小さい画素を探索しながら前記共通部分において、前記エネルギーが小さい一連の画素を求めることにより前記結合線を配置することを決定する結合線決定手段と

を備え、

前記結合線決定手段は、

前記画像を見ている者が見ている可能性が高い領域を避けつつ、前記結合線を決定する画像処理装置。

【請求項 3】

前記画像の前記共通する部分において、前記画像を見ている者が見ている領域を検出する領域検出手段と、

前記領域検出手段による検出結果に基づき、検出した前記領域のエネルギーが大きくなる第 2 のエネルギー画像を生成する第 2 のエネルギー画像生成手段と、

前記第 1 のエネルギー画像と第 2 のエネルギー画像を加算して第 3 のエネルギー画像を生成する第 3 のエネルギー画像生成手段をさらに備え、

前記結合線決定手段は、前記第 3 のエネルギー画像に基づき、前記エネルギーが小さい一連の画素を求めることにより前記結合線を決定する請求項 2 に記載の画像処理装置。

【請求項 4】

前記結合線決定手段は、シームカービング処理によって求めた前記共通する部分の前記 2 つの画像を結合する際に並べた方向と直交する方向のシームを結合線とする請求項 3 に記載の画像処理装置。

【請求項 5】

前記 2 つの画像は、所定の位置に設けた仮想視点から見た仮想全天球画像を生成する際に用いる請求項 1 から 4 のいずれか 1 項に記載の画像処理装置。

【請求項 6】

第 1 の映像から取得された画像である第 1 画像と、第 2 の映像から取得された画像であって前記第 1 画像と撮像領域の一部が共通する画像である第 2 画像との 2 つの画像を結合する際に、共通する部分において前記 2 つの画像を結合するための結合線を決定する画像処理装置が行う画像処理方法であって、

前記画像を見ている者が見ている可能性が高い領域を避けつつ、前記結合線を配置することを決定する結合線決定ステップを有する画像処理方法。

【請求項 7】

第 1 の映像から取得された画像である第 1 画像と、第 2 の映像から取得された画像であって前記第 1 画像と撮像領域の一部が共通する画像である第 2 画像との 2 つの画像を結合する際に、共通する部分において前記 2 つの画像を結合するための結合線を決定する画像処理装置が行う画像処理方法であって、

前記共通する部分において、人間の目に敏感に感じる画素はそのエネルギーが大きいと定義される第1のエネルギー画像を生成するエネルギー画像生成ステップと、

前記第1のエネルギー画像のエネルギーが小さい画像を探索しながら前記共通部分における前記結合線を決定する結合線決定ステップと

を有し、

前記結合線決定ステップは、

前記画像を見ている者が見ている可能性が高い領域を避けつつ、前記結合線を決定する画像処理方法。

【請求項8】

コンピュータを、請求項1から5のいずれか1項に記載の画像処理装置として機能させるための画像処理プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、複数のカメラからの画像データを処理する画像処理装置、画像処理方法及び画像処理プログラムに関する。

【背景技術】

【0002】

近年、周囲360度を含む全天の画像である全天球画像を撮影できるカメラ（以下、全天球カメラという）及びその全天球画像の視聴において利用者が向いた方向を視聴することができるヘッドマウントディスプレイ（HMD）が普及し始めている。そして、ネットワークを介して全天球画像を配信するサービスが注目を集めている。上記のような全天球画像は、HMDで視聴することで高い臨場感を得ることができ、スポーツやアーティストのライブ等のコンテンツの視聴における利用が期待されている。

【0003】

一般に、これらの全天球画像は、所望の視点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、競技中のサッカーコートの中やバスケットコートの中は、全天球カメラを設置しようとする競技者の邪魔となるため、全天球カメラを設置することができない。しかし、競技中のサッカーコートの中やバスケットコートの中に立っているかのような映像を視聴してみたいという要望がある。そこで、通常では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、仮想視点を含む領域を撮影する複数のカメラを設置し、それらのカメラからの画像を合成することにより、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得る技術が考案されている（例えば、非特許文献1参照）。以下の説明において、仮想視点における全天球画像を、仮想全天球画像という。

【0004】

仮想全天球画像を複数のカメラからの画像の合成によって得る画像処理システムの具体例について説明する。図20は、従来の仮想全天球画像を得るための画像処理システムを示す図である。図20に示すように、画像処理システム1は、全天球カメラ2と、N台（ $N \geq 2$ ）の複数のカメラ3-1、3-2、3-3、…、3-N（以下、カメラ群3とする。）と、画像処理装置4と、表示装置5とを備える。画像処理システム1は、フットサルのコート10内に仮想視点11を設定した場合に、コート10外に設置したカメラ群3からの画像の合成によって仮想視点11における仮想全天球画像を得る。

【0005】

全天球カメラ2は、全天球画像を撮影するカメラである。全天球カメラ2は、試合が行われる前のタイミングでコート10内の仮想視点11の位置に設置される。全天球カメラ2は、予め、仮想視点11の位置から仮想全天球画像の背景となる背景画像20を撮影する。全天球カメラ2で撮影された全天球画像である背景画像20は、画像処理装置4に入力されて蓄積される。

【0006】

10

20

30

40

50

コート10の周囲には、カメラ群3が設置されている。図20においてNは4以上の自然数である。カメラ群3を構成するカメラの数は、多ければ多いほどよいが、最低数は2である。カメラ群3は、それぞれ仮想視点11を含む画角となるようにコート10の周囲に設置されている。画像処理装置4は、背景画像20に対して合成するためカメラ群3の各カメラが出力する前景画像を含む切り出し画像に対して画像処理を行う。画像処理装置4は、全天球カメラ2より取得した背景画像20に画像処理後の部分画像を合成して仮想全天球画像を生成する。表示装置5は、画像処理装置4で生成した仮想全天球画像を表示する装置であり、液晶ディスプレイ等である。

【0007】

画像処理システム1における画像処理の具体例を説明する。図21は、画像処理システム1における画像処理される画像の具体例を示す図である。図21(A)は、仮想視点11の位置に設置された全天球カメラ2で撮影された背景画像20の例を示す図である。仮想視点11を中心とする360度の画像となっている。背景画像20は、競技開始前に撮影される画像であるのでコート10内に競技を行う選手等は映っていない。

【0008】

図21(B)は、左からカメラ3-1で撮影した部分画像21と、カメラ3-2で撮影した部分画像22と、カメラ3-3で撮影した部分画像23とを示している。画像処理装置4は、部分画像21~23のそれぞれから仮想視点11を含み、かつ、フットサルの選手を含む領域211、221、231を切り出す。画像処理装置4は、切り出した領域211、221、231の画像に対して、画像処理を行うことで背景画像20に貼り付け可能な部分画像211a、221a、231aを生成する。

【0009】

画像処理装置4は、背景画像20に対して部分画像211a、221a、231aを合成することで、仮想全天球画像24を生成する。図21(C)は、画像処理装置4が生成する仮想全天球画像24の例を示す図である。図21(C)に示すように、仮想全天球画像24は、所定の領域に部分画像211a、221a、231aを貼り付けているので、コート10上で競技を行っているフットサルの選手が映っている画像である。

【0010】

従来の画像処理システム1は、合成に用いているカメラ群3の光学中心及び仮想視点11において想定する仮想全天球カメラの光学中心はそれぞれ異なる。このため、合成された仮想全天球画像24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4は、部分画像211a、221a、231aを、仮想視点11からの距離を示す奥行の一点で整合性が保たれるよう画像処理を行い背景画像20に貼り付ける必要がある。しかしながら、整合性が保たれる奥行に存在せずに別の奥行に存在している物体(例えば、競技中の選手)の部分画像を貼り付ける場合には、画像処理により奥行の整合性を保つことができない。このような奥行に整合性のない物体は、仮想全天球画像24において、その画像が分身(多重像)したり、消失したりする現象が発生する。

【0011】

以下に、図面を用いて仮想全天球画像24において、物体の画像が分身したり、消失したりする現象について説明する。図22は、画像処理システム1における課題を説明するための図である。図22において、撮影範囲41は、カメラ3-1の撮影範囲において図21(B)に示した領域211の撮影範囲を示す。撮影範囲42は、カメラ3-2の撮影範囲において図21(B)に示した領域221の撮影範囲を示す。撮影範囲43は、カメラ3-3の撮影範囲において図21(B)に示した領域231の撮影範囲を示す。また、仮想視点11からの距離(奥行)が異なる3つの被写体(選手)49~51が存在する。

【0012】

図22において破線で示している仮想視点11からの第1の距離を示す奥行46は、各撮影範囲41~43が、重なりなく並んでいる。このような奥行46に位置する被写体49は、その画像が分身したり消失したりすることがなく、奥行に整合性のある被写体49である。仮想視点11からの第2の距離を示す奥行47は、各撮影範囲41~43が、横

10

20

30

40

50

線部分 4 4 に示すように重なっている。このような奥行 4 7 に位置する被写体 5 0 は、その画像が分身してしまうので、奥行に整合性のない被写体 5 0 となる。仮想視点 1 1 からの第 3 の距離を示す奥行 4 8 は、各撮影範囲 4 1 ~ 4 3 の間が斜線部分 4 5 に示すように空いている。このような奥行 4 8 に位置する被写体 5 1 は、その画像の一部が消失してしまうので、奥行に整合性のない被写体 5 1 となる。

【先行技術文献】

【非特許文献】

【0 0 1 3】

【非特許文献 1】高橋康輔、外 3 名、「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、2015 年 06 月 01 日、vol.115, no.76、MVE2015-5、p.43-48

10

【発明の概要】

【発明が解決しようとする課題】

【0 0 1 4】

以上のように、仮想全天球画像において被写体がある領域は、ユーザが注視する領域である視聴領域である可能性が高く、その視聴領域において被写体の分身や消失が発生すると、仮想全天球画像の視聴品質が低下するという問題がある。また、画像合成する際に、画像の継ぎ目が目に付いてしまうとこれもまた、仮想全天球画像の視聴品質が低下するという問題がある。

【0 0 1 5】

本発明は、このような事情に鑑みてなされたもので、視聴品質の低下を抑制した仮想全天球画像を生成する画像処理装置、画像処理方法及び画像処理プログラムを提供することを目的とする。

20

【課題を解決するための手段】

【0 0 1 6】

本発明の一態様は、第 1 の映像から取得された画像である第 1 画像と、第 2 の映像から取得された画像であって前記第 1 画像と撮像領域の一部が共通する画像である第 2 画像との 2 つの画像を結合する際に、共通する部分において前記 2 つの画像を結合するための結合線を決定する画像処理装置であって、前記画像を見ている者が見ている可能性が高い領域を避けつつ、前記結合線を配置することを決定する結合線決定手段を備える画像処理装置である。

30

【0 0 1 7】

本発明の一態様は、第 1 の映像から取得された画像である第 1 画像と、第 2 の映像から取得された画像であって前記第 1 画像と撮像領域の一部が共通する画像である第 2 画像との 2 つの画像を結合する際に、共通する部分において前記 2 つの画像を結合するための結合線を決定する画像処理装置であって、前記共通する部分において、人間の目に敏感に感じる画素はそのエネルギーが大きいと定義される第 1 のエネルギー画像を生成する第 1 のエネルギー画像生成手段と、前記第 1 のエネルギー画像のエネルギーが小さい画素を探索しながら前記共通部分において、前記エネルギーが小さい一連の画素を求めることにより前記結合線を配置することを決定する結合線決定手段とを備え、前記結合線決定手段は、前記画像を見ている者が見ている可能性が高い領域を避けつつ、前記結合線を決定する画像処理装置である。

40

【0 0 1 8】

本発明の一態様は、前記画像処理装置であって、前記画像の前記共通する部分において、前記画像を見ている者が見ている領域を検出する領域検出手段と、前記領域検出手段による検出結果に基づき、検出した前記領域のエネルギーが大きくなる第 2 のエネルギー画像を生成する第 2 のエネルギー画像生成手段と、前記第 1 のエネルギー画像と第 2 のエネルギー画像を加算して第 3 のエネルギー画像を生成する第 3 のエネルギー画像生成手段をさらに備え、前記結合線決定手段は、前記第 3 のエネルギー画像に基づき、前記エネルギーが小さい一連の画素を求めることにより前記結合線を決定する。

【0 0 1 9】

50

本発明の一態様は、前記画像処理装置であって、前記結合線決定手段は、シームカービング処理によって求めた前記共通する部分の前記２つの画像を結合する際に並べた方向と直交する方向のシームを結合線とする。

【００２０】

本発明の一態様は、前記画像処理装置であって、前記２つの画像は、所定の位置に設けた仮想視点から見た仮想全天球画像を生成する際に用いる。

【００２１】

本発明の一態様は、第１の映像から取得された画像である第１画像と、第２の映像から取得された画像であって前記第１画像と撮像領域の一部が共通する画像である第２画像との２つの画像を結合する際に、共通する部分において前記２つの画像を結合するための結合線を決定する画像処理装置が行う画像処理方法であって、前記画像を見ている者が見ている可能性が高い領域を避けつつ、前記結合線を配置することを決定する結合線決定ステップを有する画像処理方法である。

【００２２】

本発明の一態様は、第１の映像から取得された画像である第１画像と、第２の映像から取得された画像であって前記第１画像と撮像領域の一部が共通する画像である第２画像との２つの画像を結合する際に、共通する部分において前記２つの画像を結合するための結合線を決定する画像処理装置が行う画像処理方法であって、前記共通する部分において、人間の目に敏感に感じる画素はそのエネルギーが大きいと定義される第１のエネルギー画像を生成するエネルギー画像生成ステップと、前記第１のエネルギー画像のエネルギーが小さい画像を探索しながら前記共通部分における前記結合線を決定する結合線決定ステップとを有し、前記結合線決定ステップは、前記画像を見ている者が見ている可能性が高い領域を避けつつ、前記結合線を決定する画像処理方法である。

【００２３】

本発明の一態様は、コンピュータを、前記画像処理装置として機能させるための画像処理プログラムである。

【発明の効果】

【００２４】

本発明によれば、２つの画像を結合する際に、画像の継ぎ目を気づきにくい位置に配置するようにしたため、視聴品質の低下を抑制した仮想全天球画像を生成することができるという効果が得られる。

【図面の簡単な説明】

【００２５】

【図１】本発明の一実施形態による画像処理装置の構成を示すブロック図である。

【図２】画像処理装置３０の基本構成例を示す図である。

【図３】オブジェクト情報格納部３０３に格納するオブジェクト情報の一例を示す図である。

【図４】隣り合う部分画像間の境界領域において重複が発生する場合の具体例を示す図である。

【図５】画像処理システム１において１フレームの仮想全天球画像を作成する動作を示すフロー図である。

【図６】画像処理装置３０が動画の仮想全天球画像を作成する動作について説明する図である。

【図７】仮想全天球画像の生成処理を示す模式図である。

【図８】シームを求める動作を示す模式図である。

【図９】２つの部分画像を結合する処理を示す説明図である。

【図１０】前景の進行方向に基づいてエネルギー画像を生成して処理を行う例を示す図である。

【図１１】エネルギー画像の生成方法を示す説明図である。

【図１２】エネルギー画像の生成方法を示す説明図である。

10

20

30

40

50

【図 1 3】図 9 に示す 2 つの部分画像を結合する動作を示すフローチャートである。

【図 1 4】向いている方向で重み付けを行う例を示す図である。

【図 1 5】中心からの距離で重み付けを行う例を示す図である。

【図 1 6】エネルギー画像の生成方法を示す説明図である。

【図 1 7】前景の動きの速さに基づいてエネルギーをも求める例を示す図である。

【図 1 8】仮想全天球画像を生成した例を示す図である。

【図 1 9】カメラの入力画像から切り出す部分画像の例を示す説明図である。

【図 2 0】従来の仮想全天球画像を得るための画像処理システムを示す図である。

【図 2 1】画像処理システム 1 における画像処理される画像の具体例を示す図である。

【図 2 2】画像処理システム 1 における課題を説明するための図である。

10

【発明を実施するための形態】

【0026】

以下、図面を参照して、本発明の一実施形態による画像処理装置を説明する。図 1 は同実施形態による仮想全天球画像を視聴するためのシステム構成を示すブロック図である。この図において、図 20 に示す従来の装置と同一の部分には同一の符号を付し、その説明を簡単に行う。仮想全天球画像を視聴するためのシステムは、画像処理システム 1 及び視聴システム 9 を備えている。

【0027】

図 1 に示すように、画像処理システム 1 は、全天球カメラ 2 と、N 台 ($N \geq 2$) の複数のカメラ 3-1、3-2、3-3、…、3-N (以下、カメラ群 3 とする。) と、画像処理装置 30 と、表示装置 5 とを備える。画像処理システム 1 は、フットサルのコート 10 内に仮想視点 11 を設定した場合に、コート 10 外に設置したカメラ群 3 からの画像の合成によって仮想視点 11 における仮想全天球画像を得る。

20

【0028】

全天球カメラ 2 は、全天球画像を撮影するカメラである。全天球カメラ 2 は、競技が行われる前のタイミングでコート 10 内の仮想視点 11 の位置に設置される。全天球カメラ 2 は、予め、仮想視点 11 の位置から仮想全天球画像の背景となる背景画像 20 を撮影する。全天球カメラ 2 で撮影された背景画像 20 は、画像処理装置 4 に入力されて蓄積される。全天球カメラ 2 は、競技中も仮想視点 11 に設置したままだと競技の支障となるため、競技開始前に仮想視点 11 の位置から取り除かれる。

30

【0029】

コート 10 の周囲には、カメラ群 3 が設置されている。カメラ群 3 の各カメラ 3-1、3-2、3-3、…、3-N は、背景画像 20 に対して合成する前景を含む動画 (映像) で撮影するカメラであり、それぞれ仮想視点 11 を含む画角となるようにコート 10 の周囲を取り囲むように設置されている。図 1 において N は、4 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コート 10 が大きいほど大きな値となり、コート 10 の大きさが同じであれば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。

【0030】

画像処理装置 30 は、各カメラ 3-1、3-2、3-3、…、3-N で撮影される動画を構成する画像である入力画像に対して画像処理を施して、全天球カメラ 2 より取得した背景画像 20 に画像処理後の部分画像を合成する処理を行う。表示装置 5 は、画像処理装置 30 で生成した仮想全天球画像を表示する装置であり、液晶ディスプレイ、ヘッドマウントディスプレイ (HMD) 等である。

40

【0031】

視聴システム 9 は、画像サーバ 6 と、ネットワーク 7 と、複数の視聴装置 8 とを備える。画像サーバ 6 は、ネットワーク 7 を介して画像処理装置 30 が生成した仮想全天球画像を配信するサーバである。ネットワーク 7 は、例えばインターネット等の通信網である。視聴装置 8 は、ネットワーク 7 に接続可能なユーザ端末 81 と、ユーザ端末 81 に接続された HMD 82 とから構成される装置である。ユーザ端末 81 は、ネットワーク 7 を介し

50

て画像サーバ 6 が配信する仮想全天球画像を受信する機能と、受信した仮想全天球画像を HMD 8 2 で視聴可能な映像信号に変換して HMD 8 2 へ出力する機能とを備える。

【0032】

HMD 8 2 は、ユーザ端末 8 1 から映像信号等を受信する受信部と、受信部を介して受信した映像信号を表示する液晶ディスプレイ等で構成される画面と、視聴者の頭の動きを検出する検出部と、検出部が検出した結果をユーザ端末 8 1 に送信する送信部とを備える。HMD 8 2 の画面に表示される映像は、仮想全天球画像に基づいた仮想全天球映像の一部であり視野と呼ぶ。HMD 8 2 は、検出部が検出した視聴者の頭の動きに応じて表示する映像の範囲である視野を変更する機能を有する。

【0033】

頭を上下左右に動かすことに応じて視聴している映像が変化するので、HMD 8 2 を頭に装着した視聴者は、仮想視点 1 1 の位置から競技を見ているかのような映像を視聴することができる。このように、HMD 8 2 を装着した視聴者は、あたかも仮想視点 1 1 に立って競技を観戦しているかのような臨場感のある映像を視聴することができる。

【0034】

画像処理システム 1 において処理される画像は、図 20 に示した従来の画像処理システム 1 で処理される画像と同様であるので、図 20 を用いて画像処理システム 1 の動作について簡単に説明する。全天球カメラ 2 は、コート 10 内の仮想視点 1 1 に設置されて、図 21 (A) に示す背景画像 20 を競技開始前に撮影する。競技が開始されるとカメラ群 3 の各カメラが撮影を開始する。例えば、カメラ群 3 内のカメラ 3-1、3-2、3-3 は、図 21 (B) に示す部分画像 21~23 を撮影する。

【0035】

画像処理装置 30 は、撮影された部分画像 21~23 のそれぞれから仮想視点 1 1 を含み、かつ、競技中の選手を含む領域 211、221、231 を切り出す。画像処理装置 30 は、切り出した領域 211、221、231 の画像に対して、画像処理を行うことで背景画像 20 に貼り付け可能な部分画像 211a、221a、231a を生成する。画像処理装置 30 は、背景画像 20 に対して部分画像 211a、221a、231a を合成することで、図 21 (C) に示すような仮想全天球画像 24 を生成する。

【0036】

なお、視聴システム 9 は、図 1 に示す構成に限定されるものではない。視聴システム 9 は、画像処理装置 30 が生成した仮想全天球画像を編集してから画像サーバ 6 へ出力する編集装置を備える構成等、仮想全天球画像をネットワーク 7 経由で配信可能な構成であればよい。視聴装置 8 の構成は、ネットワーク 7 を介して受信した仮想全天球画像を利用者が視聴できる構成であれば、どのような構成であってもよい。

【0037】

次に、図 1 に示す画像処理装置 30 の構成について説明する。図 2 は、画像処理装置 30 の基本構成例を示す図である。図 2 に示すように、画像処理装置 30 は、オブジェクト解析部 31 と、奥行取得部 32 と、合成情報取得部 33 と、画像入力部 34 と、画像切り出し部 35 と、画像合成部 36 と、表示処理部 37 と、キーボードやマウス等で構成され、奥行に関する情報を入力する入力部 38 と、貼り合わせる部分画像の境界（継ぎ目）を決定する境界決定部 39 と、カメラ群 3 の各カメラが撮影した前景画像を含む部分画像を格納する前景画像格納部 301 と、背景画像 20 を格納する背景画像格納部 302 と、オブジェクト情報格納部 303 と、合成情報テーブル 304 とを備える。

【0038】

オブジェクト解析部 31 は、前景画像格納部 301 に格納されている部分画像を入力とし、部分画像に含まれるオブジェクトを抽出して、出力する。ここでオブジェクトとは、背景画像 20 に含まれていないが部分画像に含まれている人物、物体（例えばボール）等である。オブジェクト解析部 31 は、抽出したオブジェクトに対して当該オブジェクトを識別するための識別子である ID を付与する。

【0039】

10

20

30

40

50

カメラ群 3 の各カメラで撮影される部分画像は、所定のフレーム周期を有する動画像であり、各フレームには撮影時間が関連付けられている。オブジェクト解析部 31 は、時間方向に一連のフレームから抽出した同一オブジェクトに対して同じ ID を付与する。オブジェクト情報格納部 303 は、オブジェクトを抽出する対象とした部分画像のフレーム毎の撮影時刻に関連付けてオブジェクト解析部 31 が付与した ID を含むオブジェクトに関する情報を格納する。

【0040】

例えば、オブジェクト解析部 31 は、カメラ 3-1 が撮影した撮影時刻 t 、 $t+1$ 、 $t+2$ 、…の一連のフレームである部分画像 21 から抽出したオブジェクトには、ID1 の識別子を付与する。同様に、オブジェクト解析部 31 は、カメラ 3-2 が撮影した撮影時刻 t 、 $t+1$ 、 $t+2$ 、…の一連のフレームである部分画像 22 から抽出したオブジェクトには、ID2 の識別子を付与し、カメラ 3-3 が撮影した撮影時刻 t 、 $t+1$ 、 $t+2$ 、…の一連のフレームである部分画像 23 から抽出したオブジェクトには、ID3 の識別子を付与する。

【0041】

オブジェクト解析部 31 は、部分画像を解析してオブジェクトを抽出する際に、オブジェクトの属性を示すラベルと、オブジェクトのコート 10 上の空間における 3 次元的位置情報である 3 次元位置情報とを取得する。ラベルの具体例としては、人物であることを示す「人」、ボールであることを示す「ボール」、物体 A であることを示す「物体 A」、物体 B であることを示す「物体 B」、…等のカメラ群 3 の撮影範囲を移動する可能性のある物体を識別する情報を用いる。

【0042】

オブジェクト解析部 31 は、オブジェクトを抽出するために部分画像を解析処理することで、オブジェクトが「人」、「ボール」、「物体 A」、「物体 B」のいずれに該当するのかを解析・判定して、その判定結果をラベルとして出力する。なお、オブジェクトが「人」、「ボール」、「物体 A」、「物体 B」のいずれに該当するのかを解析・判定する手法としては、公知の画像解析技術を用いる。例えば、画像の解析により人を検出する技術を開示する文献として以下の公知文献 1 がある。

公知文献 1：山内悠嗣、外 2 名、「[サーベイ論文] 統計的学習手法による人検出」、電子情報通信学会技術研究報告、vol.112、no.197、PRMU2012-43、pp.113-126、2012 年 9 月

【0043】

また、オブジェクト解析部 31 は、部分画像内におけるオブジェクトの位置、オブジェクトを撮影したカメラ群 3 内の複数のカメラの位置及びその複数のカメラの撮影範囲（撮影方向及び画角）等の情報に基づいて、コート 10 上の空間におけるオブジェクトの 3 次元位置を取得する。このオブジェクトの 3 次元位置を取得する手法としては、公知の技術を用いる。また、取得位置情報は、2 次元位置の情報であってもよい。

【0044】

オブジェクト情報格納部 303 は、オブジェクト解析部 31 が抽出したオブジェクトに関する情報であるオブジェクト情報を入力とし、オブジェクト情報をその撮影時刻に関連付けて格納する。オブジェクト情報は、オブジェクトを識別する ID と、オブジェクトの属性を示すラベルと、オブジェクトの 3 次元位置とを含む。

【0045】

図 3 は、オブジェクト情報格納部 303 に格納するオブジェクト情報の一例を示す図である。図 3 に示すように、部分画像の各フレームの撮影時刻を示す時刻 t 、 $t+1$ 、 $t+2$ 、…に関連付けて複数のオブジェクト情報を格納している。時刻 t においては、オブジェクト 1 のオブジェクト情報として、ID1、ラベル 1、3 次元位置情報 1 が格納され、オブジェクト 2 のオブジェクト情報として、ID2、ラベル 2、3 次元位置情報 2 が格納されている。時刻 $t+1$ 、時刻 $t+2$ においても、同じ情報が格納されている。

【0046】

奥行取得部 3 2 は、オブジェクト情報格納部 3 0 3 よりオブジェクト情報を読み出して、各撮影時刻において、複数のオブジェクトの中から重要なオブジェクトである主オブジェクトを特定して、仮想視点 1 1 から特定した主オブジェクトまでの距離である奥行に関する奥行情報を取得し、出力する。重要なオブジェクトとは、例えば、仮想全天球画像の中で視聴者が注視する領域に存在するオブジェクトである。

【0047】

奥行取得部 3 2 は、予め各撮影時刻における主オブジェクトを特定しておく。具体的には、仮想全天球画像を作成するコンテンツ作成者が、各撮影時刻において視聴者が注視すると推定される領域または視聴者が注視すると推定されるオブジェクトを特定する情報を入力部 3 8 から入力する。これにより、奥行取得部 3 2 は、入力された情報に基づいて各撮影時刻における主オブジェクトを特定する。奥行取得部 3 2 において、主オブジェクトを特定する方法は、上述した方法に限定されるものではなく、色々な方法を用いてよい。例えば、撮影した部分画像における視聴者の興味の高さを領域別に表したマップである Saliency Map を求めて奥行取得部 3 2 に入力する。奥行取得部 3 2 では、入力された Saliency Map に基づいて視覚的に顕著な領域に存在するオブジェクトを主オブジェクトとして特定してもよい。また、予め被験者に部分画像である動画を視聴させ、各撮影時刻においてどの領域を見ていたかという視聴ログを獲得し、その視聴ログを奥行取得部 3 2 に入力し、入力された視聴ログに基づいて主オブジェクトを特定してもよい。

【0048】

なお、Saliency Map の求め方は公知の技術であり、例えば、以下の公知文献 2 に記載の技術を用いてもよい。

公知文献 2 : Laurent Itti, Christof Koch, and Ernst Niebur, "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(11):1254-1259 (1998)

【0049】

合成情報テーブル 3 0 4 は、部分画像から仮想視点 1 1 を含む領域を切り出すための切り出し領域に関する情報である切り出し領域情報と、その切り出し領域に応じて切り出した画像を部分画像に変換するための情報である変換情報とを含む合成情報を格納する。部分画像は、切り出した画像を背景画像 2 0 の対応領域に違和感なく貼り付けるために、切り出した画像に対して上記変換情報に応じて拡大、縮小、回転等の変形処理を行って生成される。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。合成情報テーブル 3 0 4 は、カメラ群 3 において処理対象となる部分画像を撮影したカメラを特定するカメラコードと、仮想視点 1 1 からの奥行と、その奥行に応じたアフィン変換行列である変換情報と、その奥行に応じた切り出し領域情報とを対応づけて格納するテーブルである。

【0050】

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テーブル 3 0 4 に記憶しておく。例えば、仮想視点 1 1 から複数種類の距離（奥行）の位置に格子模様のチェスボードを設置して、仮想視点 1 1 に設置した全天球カメラ 2 で撮影したチェスボードを含む画像と、カメラ群 3 で撮影したチェスボードを含む画像とを比較する。そして両画像において、撮影したチェスボードの各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に対応したアフィン変換行列を求める。

【0051】

切り出し領域情報は、以下に示す方法により予め取得して合成情報テーブル 3 0 4 に記憶しておく。例えば、カメラ群 3 の内の隣接する 2 つのカメラで撮影された部分画像に同一の

被写体（チェスボード）が存在する重複している領域がある場合は、一方の領域のみ残るように双方のカメラの画像に対する切り出し領域を設定する。切り出し領域は、仮想視点11から被写体（チェスボード）まで複数種類の距離（奥行）について、カメラ群3に含まれるカメラ毎に求める。なお、双方のカメラの画像において、数画素～数十画素の幅の重複領域を残すように切り出し領域を設定してもよい。

【0052】

合成情報取得部33は、奥行取得部32が取得した奥行を入力とし、奥行に基づいて、合成情報テーブル304から、カメラ群3の各カメラで撮影された部分画像に対応する切り出し領域及びアフィン変換行列を含む合成情報を取得して、出力する。なお、合成情報テーブル304に格納されている奥行は数種類～数十種類なので、奥行取得部32が取得した奥行と同じ値の奥行のテーブルが無い場合が想定される。このような場合は、合成情報取得部33は、奥行取得部32が取得した奥行の前後の値となる合成情報テーブル304に記録済の2つの奥行の値に対応する合成情報（切出領域情報及び変換情報）を用いて、奥行取得部32が取得した奥行に対応する合成情報を算出する。具体的には、上記記録済の2つの奥行の値に対応する切出領域情報の切り出し領域の座標値を線形補間することにより、その中間に位置する切り出し領域を特定する。上記記録済の2つの奥行の値に対応するアフィン変換行列の各係数を線形補間することにより、その中間値となるアフィン変換行列を算出する。

【0053】

前景画像格納部301は、各カメラを特定するカメラコードに関連付けてカメラ群3の各カメラで撮影した前景画像を含む部分画像を格納する。部分画像は、撮影時刻及び動画の画像データを含む。前景画像格納部301は、例えば、図21（B）に示す部分画像21を、カメラ3-1を特定するカメラコードに関連付けて格納し、部分画像22を、カメラ3-3を特定するカメラコードに関連付けて格納し、部分画像23を、カメラ3-3を特定するカメラコードに関連付けて格納する。

【0054】

背景画像格納部302は、全天球カメラ2で撮影した全天球画像である背景画像20を格納する。背景画像格納部302は、例えば、コート10内の仮想視点11に設置した全天球カメラ2で撮影した図21（A）に示す背景画像20を格納する。格納する背景画像20は、1フレーム分の画像データでも所定時間分の動画の画像データでもよい。所定時間分の画像データを格納する場合は、背景画像20において周期的に変化する部分（例えば電光掲示板が映っている部分があり、かつ、電光掲示板の表示内容が周期的に変化している部分。）があれば、その周期に応じた時間分の画像データを背景画像20として格納すればよい。

【0055】

画像処理装置30が全天球カメラ2から背景画像20を取得する構成はどのような構成であってもよい。例えば、画像処理装置30が全天球カメラ2と有線または無線で通信可能な通信部を備えて、その通信部を介して背景画像20を取得する構成であってもよい。また、全天球カメラ2に着脱可能な記録媒体を用いて当該記録媒体に背景画像20を記録して、記録後の記録媒体を画像処理装置30に接続して、画像処理装置30が記録媒体から背景画像20を読み出す構成により、背景画像20を取得する構成であってもよい。また、画像処理装置30が、カメラ群3から部分画像を取得する構成も全天球カメラ2の場合と同様にどのような構成であってもよい。

【0056】

画像入力部34は、部分画像格納部301から部分画像を取得し、背景画像格納部302から背景画像20を取得して、部分画像を画像切り出し部35へ出力し、背景画像20を画像合成部36へ出力する。画像切り出し部35は、合成情報取得部33が取得した合成情報に含まれる切出領域情報に基づいて、カメラ群3の各カメラからの部分画像に対応する切り出し領域を特定し、部分画像から特定した切り出し領域を切り出して、切り出した画像を画像合成部36へ出力する。画像切り出し部35は、例えば、図21（B）に示

す部分画像 2 1 ~ 2 3 のそれぞれから切り出し領域 2 1 1、2 2 1、2 3 1 を切り出す処理を行う。

【0057】

画像合成部 3 6 は、画像切り出し部 3 5 が切り出した画像と合成情報取得部 3 3 が取得した合成情報と、背景画像を入力とし、画像切り出し部 3 5 が切り出した画像に対して、合成情報取得部 3 3 が取得した合成情報に含まれる変換情報のアフィン変換行列に基づいて変形処理を行い、部分画像を生成する。画像合成部 3 6 は、生成した部分画像をアフィン変換行列に基づいて背景画像 2 0 に貼り付けて合成することで仮想全天球画像を生成し、出力する。なお、アフィン変換行列は、背景画像 2 0 において部分画像を貼り付ける領域を示す情報を含む。画像合成部 3 6 は、生成した仮想全天球画像を画像サーバ 6 へ送信する機能を有する。

10

【0058】

画像合成部 3 6 は、例えば、図 2 1 (B) に示す部分画像 2 1 ~ 2 3 のそれぞれから切り出し領域 2 1 1、2 2 1、2 3 1 を切り出した画像に対して、アフィン変換行列に基づいた変形処理を行うことで、部分画像 2 1 1 a、2 2 1 a、2 3 1 a を生成する。画像合成部 3 6 は、例えば、背景画像 2 0 に対して、部分画像 2 1 1 a、2 2 1 a、2 3 1 a を所定の領域に貼り付けて合成することで図 2 1 (C) に示す仮想全天球画像 2 4 を生成する。

【0059】

部分画像を背景画像 2 0 に貼り付けて仮想全天球画像 2 4 を生成した際に、隣り合う部分画像間の境界領域において重複が発生する場合がある。図 4 は、隣り合う部分画像間の境界領域において重複が発生する場合の具体例を示す図である。図 4 に示すように、仮想全天球画像 2 4 に貼り付けた部分画像 2 1 1 b と部分画像 2 2 1 b とが境界領域 2 5 において重複している。なお、図 4 に示す部分画像 2 1 1 b と部分画像 2 2 1 b が、図 2 1 (C) に示した部分画像 2 1 1 a 及び部分画像 2 2 1 a と比較して異なる点は、両画像に重複する領域がある点である。

20

【0060】

図 4 に示すように、部分画像 2 1 1 b と部分画像 2 2 1 b とが境界領域 2 5 において重複している場合には、画像合成部 3 6 は、重複している境界領域 2 5 に対して以下に示すブレンディング (Blending) 処理を行う。画像合成部 3 6 は、Blending

30

$$g(x, y) = \alpha I_i(x, y) + (1 - \alpha) I_{i+1}(x, y) \quad \cdots \quad (\text{式 } 1)$$

【0061】

(式 1) において、 x 、 y は、仮想全天球画像 2 4 上における水平方向、垂直方向の座標である。 $g(x, y)$ は、境界領域 2 5 内の座標 (x, y) の画素値の値である。 $I_i(x, y)$ と $I_{i+1}(x, y)$ は、カメラ群 3 内のカメラ 3 - i 及びカメラ 3 - ($i + 1$) によって撮影された部分画像に基づいて生成された部分画像の座標 (x, y) の画素値の値を表す。また、この α の値は重複領域 2 5 で一定であるが、以下の (式 2) に示すように変化させてもよい。

$$\alpha(x) = (x - x_s) / (x_e - x_s) \quad \cdots \quad (\text{式 } 2)$$

40

(式 2) において、 x_s 及び x_e は、図 4 に示すように重複領域 2 5 の両端の x 座標であり、 $x_s < x_e$ である。

【0062】

表示処理部 3 7 は、画像合成部 3 6 が出力する仮想全天球画像を入力とし、仮想全天球画像を表示装置 5 において表示可能な映像信号に変換して出力する。仮想全天球画像 2 4 は、図 2 1 (C) に示す通り、歪みを含む画像であり、かつ、仮想視点 1 1 を中心とする 3 60 度の景色を含む画像であるので、表示処理部 3 7 は、仮想全天球画像から表示装置 5 に表示させる範囲の画像を切り出して、切り出した画像の歪みを補正する機能を有する。

【0063】

50

画像処理装置 30 は、前景画像格納部 301 及び背景画像格納部 302 を備える構成としたが、これに限定されるものではない。例えば、前景画像格納部 301 及び背景画像格納部 302 を備える画像格納装置を別に設け、画像処理装置 30 は、画像格納装置から前景画像格納部 301 及び背景画像格納部 302 を取得する構成であってもよい。

【0064】

境界決定部 39 は、画像合成部 36 が出力する仮想全天球画像と合成情報を入力とし、前述した境界領域に対して、ブレンディング処理を行うのではなく、目に付きにくい境界線とすることにより、自然な画像合成を行うための境界領域を決定し、出力する。

【0065】

次に、画像処理システム 1 において 1 フレームの仮想全天球画像を作成する動作について説明する。図 5 は、画像処理システム 1 において 1 フレームの仮想全天球画像を作成する動作を示すフロー図である。図 5 に示す動作は、各撮影時刻における仮想全天球画像を生成する処理の前に、予めオブジェクト情報、合成情報、背景画像 20 及び部分画像を取得する処理も含まれる。

【0066】

仮想視点 11 に全天球カメラ 2 を設置し、仮想視点 11 から所定の距離（奥行）にチェスボードを設置した後に、全天球カメラ 2 は、チェスボードを含む全天球画像を撮影する（ステップ S101）。全天球カメラ 2 を仮想視点 11 から取り去って、カメラ群 3 の各カメラで、仮想視点 11 及びチェスボードを含む撮影範囲を撮影し、全天球カメラ 2 で撮影された全天球画像に含まれるチェスボードと、カメラ群 3 内の一つのカメラで撮影された画像に含まれるチェスボードとを対応させるための合成情報を求める（ステップ S102）。なお、ステップ S101、102 におけるチェスボードの撮影は、仮想視点 11 から複数種類の距離にチェスボードを設置して行われる。

【0067】

仮想視点 11 に全天球カメラ 2 を設置した後に、全天球カメラ 2 は、背景画像 20 を撮影する（ステップ S103）。撮影された背景画像 20 は、背景画像格納部 302 に格納される。全天球カメラ 2 を仮想視点 11 から取り去った後であって、例えば競技開始と共に、カメラ群 3 は撮影を開始する。これにより、画像処理装置 30 は、カメラ群 3 が撮影した部分画像を前景画像格納部 301 に格納する。オブジェクト解析部 31 は、前景画像格納部 301 から部分画像を読み出して解析処理し、解析結果をオブジェクト情報格納部 303 に格納する。奥行取得部 32 は、オブジェクト情報格納部 303 に格納されているオブジェクトの中から、入力部 38 から入力された情報に基づいて主オブジェクトを特定する。奥行取得部 32 は、仮想視点 11 から特定した主オブジェクトまでの奥行情報を取得する（ステップ S104）。

【0068】

合成情報取得部 33 は、奥行取得部 32 が取得した奥行を入力とし、奥行に基づいて、合成情報テーブル 304 から、各部分画像に対応する切り出し領域及びアフィン変換行列を含む合成情報を取得して、出力する（ステップ S105）。ステップ S105 において、合成情報取得部 33 は、奥行取得部 32 が取得した奥行と同じ値の奥行のテーブルが無い場合は、奥行取得部 32 が取得した奥行の前後の値となる奥行に対応する合成情報に基づいて、奥行取得部 32 が取得した奥行に対応する合成情報を求める。

【0069】

画像切り出し部 35 は、合成情報取得部 33 が取得した合成情報に含まれる切出領域情報を入力とし、切出領域情報に基づいて、カメラ群 3 の各カメラからの部分画像に対応する切り出し領域を特定し、部分画像から特定した切り出し領域を切り出して、切り出した画像を画像合成部 36 へ出力する。画像合成部 36 は、画像切り出し部 35 が切り出した画像と合成情報取得部 33 が取得した合成情報と背景画像を入力とし、画像切り出し部 35 が切り出した画像に対して、合成情報に含まれる変換情報のアフィン変換行列に基づいて変形処理を行い、部分画像を生成する。画像合成部 36 は、生成した部分画像をアフィン変換行列に基づいて背景画像 20 に貼り付けて合成することで仮想全天球画像を生成し

、出力する（ステップ S 1 0 6）。

【0070】

画像合成部 3 6 は、背景画像 2 0 に貼り付ける 2 つの部分画像間の境界領域において重複している場合には、重複している境界領域に対してブレンディング処理を行う（ステップ S 1 0 7）。

【0071】

次に、画像処理装置 3 0 が動画の仮想全天球画像を作成する基本動作について説明する。図 6 は、画像処理装置 3 0 が動画の仮想全天球画像を作成する動作について説明する図である。図 6 の動作においては、図 5 に示したステップ S 1 0 1 ～ステップ S 1 0 4 における部分画像の撮影までの処理は既に終わっているものとする。図 6 に示すように、画像処理装置 3 0 は、最初の撮影時刻のフレームに対する処理を開始する（ステップ S 2 0 1）。

10

【0072】

画像入力部 3 4 は、前景画像格納部 3 0 1 から部分画像を取得し、背景画像格納部 3 0 2 から背景画像 2 0 を取得して、部分画像を画像切り出し部 3 5 へ出力し、背景画像 2 0 を画像合成部 3 6 へ出力する（ステップ S 2 0 2）。奥行取得部 3 2 は、オブジェクト情報格納部 3 0 3 に格納されているオブジェクトの中から、入力部 3 8 から入力された情報に基づいて主オブジェクトを特定して、特定した主オブジェクトまでの奥行を取得する（ステップ S 2 0 3）。

【0073】

20

合成情報取得部 3 3 は、奥行取得部 3 2 が取得した奥行を入力とし、奥行に基づいて、合成情報テーブル 3 0 4 から、各部分画像に対応する合成情報を取得して、出力する（ステップ S 2 0 4）。画像切り出し部 3 5 は、合成情報取得部 3 3 が取得した合成情報を入力とし、合成情報に基づいて、部分画像から切り出し領域を切り出して、切り出した画像を画像合成部 3 6 へ出力する。画像合成部 3 6 は、画像切り出し部 3 5 が切り出した画像と合成情報取得部 3 3 が取得した合成情報と背景画像を入力とし、画像切り出し部 3 5 が切り出した画像に対して、合成情報に含まれるアフィン変換行列に基づいて変形処理を行い、部分画像を生成する。画像合成部 3 6 は、生成した部分画像をアフィン変換行列に基づいて背景画像 2 0 に貼り付けて合成して、仮想全天球画像を生成して出力する（ステップ S 2 0 5）。画像処理装置 3 0 は、次の撮影時刻の部分画像があればステップ S 2 0 1 に戻りループを継続し、次の撮影時刻の部分画像がなければ、ループを終了する（ステップ S 2 0 6）。

30

【0074】

以上に説明したように画像処理装置 3 0 は、視聴者が注目する主オブジェクトに対応した奥行を求めて、求めた奥行に対応した部分画像の生成し、生成した部分画像を背景画像 2 0 に貼り付けることで仮想全天球画像を生成することができる。これにより、画像処理装置 3 0 は、仮想全天球画像に含まれる主オブジェクトである被写体において分身が起こったり、消失が起こったりすることを抑制することができる。画像処理装置 3 0 は、視聴者の注目する被写体の奥行に応じた合成処理を行うことで、仮想全天球画像に含まれる視聴者の注目する被写体における分身の発生を抑制することができ、視聴品質の低下を抑制した仮想全天球画像を視聴者に提供することができる。

40

【0075】

<第 1 の実施形態>

次に、本発明の第 1 の実施形態による画像処理装置を説明する。第 1 の実施形態は、前述した合成処理中に行う境界処理について変形を加えたものである。第 1 の実施形態では、前述したように、ブレンディング処理を行うのではなく、部分画像の合成時に目に付きにくい境界とするものである。ここで、図 7 を参照して仮想全天球画像の生成処理について簡単に説明する。図 7 は、仮想全天球画像の生成処理を示す模式図である。まず、カメラ C_{i-1} 、 C_i 、 C_{i+1} によって撮影されたそれぞれの動画（映像）からそれぞれ取得された入力画像を事前に獲得する。カメラによって撮影された動画（映像）は、複数フ

50

レームの画像により構成されるものであり、ここでは処理対象フレームの画像を入力画像として獲得する。また、エネルギー画像の生成のために、処理対象フレームよりも過去の1つまたは複数のフレームの画像も事前に獲得する。そして、得られたそれぞれの入力画像から前景となる切り出し画像 S_{i-1} 、 S_i 、 S_{i+1} をそれぞれ切り出す。

【0076】

次に、切り出し画像 S_{i-1} 、 S_i 、 S_{i+1} に対して、予め求めてあるアフィン変換行列 A_{i-1} 、 A_i 、 A_{i+1} によって画像変換を行い、部分画像 S'_{i-1} 、 S'_i 、 S'_{i+1} を生成する。そして、予め撮影してあった全天球画像 B と合成処理を行う。このとき、部分画像を重ね合うように並べて合成するのではなく、シームカービングの処理を適用して、部分画像と部分画像の継ぎ目が目に付かないようにして合成を行う。このように合成するようにすることにより、仮想視点 P_v から見た仮想全天球画像を生成することが可能となる。この仮想全天球画像を $HDM82$ によって、ユーザが見たい場面の方向へ視線を向けることにより、あたかもコート10内の仮想視点11のからフットサルの試合を観戦することが可能となる。

【0077】

ここで、本実施形態において用いるシームカービングについて説明する。シームの定義は、視覚的に重要でない画素を連結した画像列であり、これをシームと呼ぶ。画像の上端から下端を結ぶように連結したシームを縦シームという。シームカービングでは横シームもあるが、ここでは、縦シームのみを用いるため、縦シームについて説明する。シームカービングは、画像を拡大縮小する際に用いる技術であり、シームを削除することで、画像を縮小することができる。また、このシームを複製することにより画像の拡大にも適用可能である。シームカービングでは、視覚的に重要でないシームを検出することが重要である。これを行うために、画像のエネルギーという考えを導入する。基本的なアイデアは、人間の目に敏感に感じる画素はそのエネルギーが大きいとすることである。各画素について、エネルギーの値を付与した画像をエネルギー画像という。

【0078】

図8は、シームを求める動作を示す模式図である。まずシームが定義されていない部分画像に対して、エネルギー計算（図8（B））を行う。この結果、エネルギー画像（図8（C））が得られることになる。このエネルギー画像に対して、公知のシームカービング処理のうち、縦シームを求める処理を行うことによって縦シームが得られる。シームカービングの処理については公知の技術（公知文献3参照）であるため、ここでは詳細な説明を省略するが、一般に累積エネルギーの低い境界がシームとして求められることになる。

公知文献3: Avidan, Shai, and Ariel Shamir. "Seam carving for content-aware image resizing." ACM Transactions on graphics (TOG). Vol. 26. No. 3. ACM, 2007.

【0079】

そして、エネルギー画像上において求めたシームを元の部分画像上に重ね合わせる（図8（A））ことによって、継ぎ目として用いることができるシームを定義することが可能となる。

【0080】

次に、図9を参照して、2つの部分画像を結合する処理について説明する。図9は、2つの部分画像を結合する処理を示す説明図である。前述したように、2つの部分画像1、2は、その一部が重複している領域（これを境界領域という）を有している。この境界領域について、部分領域1、2のいずれかの境界領域部分を用いて、シームカービング処理の縦シームを求める処理のみを行うことによって、縦シームを求める。そして、求めた縦シームを2つの部分画像1、2を結合するための結合線として用いる。ここでは、縦シームを例にして説明したが、結合する画像が上下に並べられている場合は、横シームを求めることになる。すなわち、2つの画像を結合する際に並べた方向と直交する方向のシームを求めることになる。このようにすることにより、前述したブレンディング処理が不要になるとともに、目で気づきにくい場所に結合線を設定するようにしたため、2つの部分画像1、2を結合する際に自然に結合を行うことが可能となる。

【0081】

しかし、シームカービングによって、単に結合線（継ぎ目）を求めて結合しただけでは、以下のような問題が生じる。すなわち、視聴者が注目している領域に部分画像の境界（継ぎ目）があると目に付きやすいという問題である。そこで、本実施形態では、視聴者は人などの前景が動いていく方向を見ているはずであるという仮定の基に、人などの前景画像であるオブジェクトが移動する方向にはシームを作らないようにした。図10は、前景の進行方向に基づいてエネルギー画像を生成して処理を行う例を示す図である。視聴者は、図10（A）に示すように、画像に映っている人の進行方向の領域（破線で囲んだ領域）に視線が向くため、その領域にシームを作ってしまうとシームに基づいて結合した結合線が目立つことになる。

10

【0082】

これは、図10（B）に示す対象の部分画像から単にエネルギー画像を生成したため、画像に映っている人の進行方向の領域にシームができてしまったためである。これに対して、図10（C）のように、新たに第2のエネルギー画像を生成し、この第2のエネルギー画像を加算することで前景の進行方法にシームを作らないようにする。第2のエネルギー画像は、前景の動きに基づいて、シームを作りたくない領域（ここでは、前景が進行しようとしている方向の領域）のエネルギーが高くなるようにしたエネルギー画像である。

【0083】

ここで、エネルギー画像の生成方法について説明する。図11、図12は、エネルギー画像の生成方法を示す説明図である。ある境界領域に関し、過去 s （ ≥ 1 ）フレームの画像を用いて各ピクセルに関してオプティカルフローを算出する（図11（A））。オプティカルフローを算出する手法については公知であるため、ここでは詳細な説明を省略する。

20

【0084】

次に、先に求めたオプティカルフローについて、位置と方向に関して、K-meansなどを用いてクラスタリングする（図11（B））。続いて、各クラスタについて、エネルギー画像を求める。エネルギー画像は境界領域と同じサイズの画像であるとする。図12のようにクラスタ K_i （ j はクラス番号。 $1 \leq j$ ）に含まれるオプティカルフローの平均をそのクラスタのオプティカルフローを f_i とし、 K_i の重心を g_i とすると、平均（ $g_i + \Delta v$ ）の x 座標、分散1の正規分布にしたがってエネルギー画像を作成する。

30

【0085】

なお、 $\Delta v = \alpha f_i$ であり、正規分布の横軸は図12のように f_i と平行であるとする。 α は係数であり、適当な数値を与えてよい。ただし、クラスタ K_i 付近に存在する移動物体が次時刻に $g_i + \Delta v$ 付近に来るような値を与えることが望ましい。また、エネルギーの最大値 e_{max} は予め定め（255など）、正規分布で得られるエネルギーの最大値が e_{max} と等しくなるように正規化する。なお、 e_{max} は $e_{max} = N f_i + P f_i$ のように、 K_i に含まれるオプティカルフローの数（ $N f_i$ ）やオプティカルフローの大きさ（ $P f_i$ ）によって変化させてもよい（図11（C））。

【0086】

次に、各クラスタで作成したエネルギー画像を足し合わせ、境界領域のエネルギー画像とする。なお、図11（C）で求めた f_i はクラスタ K_i に含まれるオプティカルフローの最大値、最小値、中央値などでもよい。

40

【0087】

次に、図13を参照して、図9に示す2つの部分画像を結合する動作を説明する。図13は、図9に示す2つの部分画像を結合する動作を示すフローチャートである。まず、境界決定部39は、エネルギー画像を前述した方法によって生成する（ステップS301）。

【0088】

次に、境界決定部39は、2つのエネルギー画像を足し合わせて、第3のエネルギー画像を生成する（ステップS302）。そして、境界決定部39は、シームカービング処理

50

によって、足し合わされたエネルギー画像から縦シームを求める（ステップS303）。この処理によって、図9（B）に示す縦シームが定義されることになる。

【0089】

次に、境界決定部39は、求めた縦シームを用いて境界を決定する（ステップS304）。続いて、境界決定部39は、求めた境界に基づいて、部分画像の切り出しを行う（ステップS305）。この画像切り出しは、2つの部分画像に対して行う。そして、境界決定部39は、2つの結合線を重ね合わせるようにして2つの部分画像を結合する（ステップS306）。

【0090】

次に、境界決定部39は、未だ処理していない境界領域があるか否かにより処理が終了であるか否かを判定し（ステップS307）。未処理の境界領域があれば処理を繰り返す。

【0091】

このような処理によって、人などの前景が移動する方向に部分画像の継ぎ目ができることを防止することができるため、視聴品質の低下を抑制した仮想全天球画像を生成することができる。

【0092】

<第2実施形態>

次に、本発明の第2の実施形態による画像処理装置を説明する。第2の実施形態では、第1の実施形態と同様に、ブレンディング処理を行うのではなく、部分画像の合成時に目に付きにくい境界とするものである。前述したように、シームカービングによって、単に継ぎ目を求めて結合しただけでは、以下のような問題が生じる。すなわち、視聴者が注目している領域（視線方向）に部分画像の境界（継ぎ目）があると目に付きやすいという問題である。そこで、本実施形態では、視聴者が注目している方向（視聴者が見ている可能性が高い領域であり、視野のほぼ中央のことである）にはシームを作らないようにする。図14は、向いている方向で重み付けを行う例を示す図である。視聴者は、図14（A）に示すように、視線の方向にシームを作ってしまうとシームに基づいて結合した結合線が目立つことになる。

【0093】

これは、図14（B）に示す対象の部分画像から単にエネルギー画像を生成したため、画像に映っている人の左側にシームができてしまったためである。これに対して、図14（C）のように、新たに第2のエネルギー画像を生成し、この第2のエネルギー画像を加算することで視線の方向にシームを作らないようにする。第2のエネルギー画像は、シームを作りにくい領域（ここでは、視聴者が見ている可能性が高い領域のことであり、視野のほぼ中央の領域）のエネルギーが高くなるようにしたエネルギー画像である。

【0094】

ここで、エネルギー画像の生成方法について説明する。向いている方向を表す情報として画像上の点 $D = (d_x, d_y)$ が既知であり、点Dは境界領域に含まれているとする。この点Dは、HMD82の方向の情報から取得する。境界決定部39は、平均 d_x 、分散1の正規分布にしたがってエネルギー画像を作成する。このとき、正規分布の横軸は水平であるとする。なお、エネルギーの最大値 e_{max} は予め定め、正規分布で得られるエネルギーの最大値が e_{max} と等しくなるように正規化する。

【0095】

第2実施形態における図9に示す2つの部分画像を結合する動作は、図13に示す処理動作と同様であるのでここでは詳細な説明を省略する。

【0096】

このような処理によって、視聴者の視線方向に部分画像の継ぎ目ができることを防止することができるため、視聴品質の低下を抑制した仮想全天球画像を生成することができる。

【0097】

10

20

30

40

50

<第3の実施形態>

次に、本発明の第3の実施形態による画像処理装置を説明する。第3の実施形態では、第1の実施形態と同様に、ブレンディング処理を行うのではなく、結合部（継ぎ目）が部分画像の中心から離れると幾何的整合性が崩れてしまうことを防止するための結合方法である。前述したように、シームカービングによって、単に継ぎ目を求めて結合しただけでは、以下のような問題が生じる。すなわち、結合部が部分画像の中心から離れた位置に生成されてしまうと、画像の幾何的整合性が崩れてしまうという問題である。そこで、本実施形態では、整合性が崩れない位置に結合部を設けるようにする。図15は、中心からの距離で重み付けを行う例を示す図である。本手法の性質上、中心に結合部（継ぎ目）を設定した方が幾何学的に正しい画像となる（図15（A））。

10

【0098】

したがって、従来の手法によって結合部（継ぎ目）を設定すると、画像の中心から離れた位置に結合部（継ぎ目）ができてしまう（図15（B））。これに対して、図15（C）のように、新たに第2のエネルギー画像を生成し、この第2のエネルギー画像を加算することで中心から離れた位置にシームを作らないようにする。第2のエネルギー画像は、シームを作りたくない領域（ここでは、中心から離れた位置の領域）のエネルギーが高くなるようにしたエネルギー画像である。

【0099】

ここで、エネルギー画像の生成方法について説明する。図16は、エネルギー画像の生成方法を示す説明図である。図16に示すようにカメラ C_i と C_{i+1} で撮影した部分画像中で、仮想視点の位置をそれぞれ v_i 、 v_{i+1} とし、 v_{i+1} と v_i の中点を $c = (c_x, c_y)$ とする。平均 c_x 、分散1の正規分布にしたがってエネルギー画像を作成する。このとき、正規分布の横軸は水平であるとする。なお、エネルギーの最大値 e_{max} は予め定め、正規分布で得られるエネルギーの最大値が e_{max} と等しくなるように正規化する。

20

【0100】

第3実施形態における図9に示す2つの部分画像を結合する動作は、図13に示す処理動作と同様であるのでここでは詳細な説明を省略する。

【0101】

このような処理によって、部分画像の継ぎ目が画像の中心付近に設定されるようになるため、幾何的整合性が崩れることを防止することができるため、視聴品質の低下を抑制した仮想全天球画像を生成することができる。

30

【0102】

<第4の実施形態>

次に、本発明の第4の実施形態による画像処理装置を説明する。第4の実施形態では、第1の実施形態と同様に、ブレンディング処理を行うのではなく、結合部（継ぎ目）が部分画像の中心から離れると幾何的整合性が崩れてしまうことを防止するための結合方法である。前述したように、シームカービングによって、単に継ぎ目を求めて結合しただけでは、以下のような問題が生じる。すなわち、結合部が部分画像の中心から離れた位置に生成されてしまうと、画像の幾何的整合性が崩れてしまうという問題である。そこで、本実施形態では、整合性が崩れない位置に結合部を設けるようにする。図17は、前景の動きの速さに基づいてエネルギーをも求める例を示す図である。本手法の性質上、中心に結合部（継ぎ目）を設定した方が幾何学的に正しい画像となる。

40

【0103】

したがって、従来の手法によって結合部（継ぎ目）を設定すると、画像の中心から離れた位置に結合部（継ぎ目）ができてしまう（図17（B））。これに対して、図17（C）のように、新たに第2のエネルギー画像を生成し、この第2のエネルギー画像を加算することで中心から離れた位置にシームを作らないようにする。第2のエネルギー画像は、シームを作りたくない領域（ここでは、ゆっくり動く前景は避け、速く動く前景は避けないようにする）のエネルギーが高くなるようにしたエネルギー画像である。速く動いてい

50

る前景（例えば、ボール）は二重像や消失が発生しても気づきにくいことを利用している。

【0104】

ここで、エネルギー画像の生成方法について説明する。ある境界領域に関し、過去 s (≥ 1) フレームの画像から各ピクセルに関してオプティカルフローを算出する。そして、求めたオプティカルフローについて、位置と方向に関して $K-means$ などを適用してクラスタリングする。

【0105】

次に、各クラスタについて、第3の実施形態と同様にクラスタ K_j に含まれるオプティカルフローの平均を f_j とする。このとき、 f_j の大きさ ($|f_j|$) がしきい値 t_v 以下の場合（すなわちゆっくり動く物体の場合）、前述の処理によって求めた時と同様にエネルギー画像を求める。一方、 f_j の大きさ ($|f_j|$) が t_v 以上の場合（すなわち速く動く物体の場合）、このクラスタ K_j に対するエネルギー画像は求めない。そして、各クラスタで作成したエネルギー画像を足し合わせ、重複領域のエネルギー画像とする。

【0106】

第4実施形態における図9に示す2つの部分画像を結合する動作は、図13に示す処理動作と同様であるのでここでは詳細な説明を省略する。

【0107】

このような処理によって、部分画像の継ぎ目が画像の中心付近に設定されるようになるため、幾何的整合性が崩れることを防止することができるため、視聴品質の低下を抑制した仮想全天球画像を生成することができる。特に、ゆっくり動く物体（前景）は避け、速く動く物体（前景）は、避けないようにしたため、視聴品質の低下を抑制した仮想全天球画像を生成することができる。

【0108】

以上説明したように、2つの部分画像を結合する際に、画像の継ぎ目を視聴者の目に気づかれない位置に設定するようにしたため、視聴品質の低下を抑制した仮想全天球画像を生成することができる。また、2つの部分画像を結合する際に、幾何的整合性が崩れることを防止するようにしたため、同様に視聴品質の低下を抑制した仮想全天球画像を生成することができる。

【0109】

前述した方法によって2つの部分画像を結合する処理を繰り返すことにより、図18に示すような仮想全天球画像を得ることができるようになる。図18は仮想全天球画像を生成した例を示す図である。図18左は、従来技術による仮想全天球画像の例を示す。図18右は、シームカービングの処理を適用して、結合部（継ぎ目）の位置を最適化したため、質の高い仮想全天球画像が生成できているのが分かる。

【0110】

図19は、カメラの入力画像から切り出す部分画像の例を示す説明図である。図19に示すように、左図では、被写体である人が3台のカメラ全てに映っていたため、3重像が発生してしまう。一方、右図のように、シームカービングの処理によって部分画像の領域を大きくすることができるようになるため、1つの部分画像内に被写体である人を収めることができ、3重像が発生することを防止することができる。

【0111】

前述した実施形態における画像処理装置の全部または一部をコンピュータで実現するようにしてもよい。その場合、この機能を実現するためのプログラムをコンピュータ読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、実行することによって実現してもよい。なお、ここでいう「コンピュータシステム」とは、OSや周辺機器等のハードウェアを含むものとする。また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM、CD-ROM等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。さらに「コンピュータ読み取り可能な記録媒体」とは、インター

10

20

30

40

50

ネット等のネットワークや電話回線等の通信回線を介してプログラムを送信する場合の通信線のように、短時間の間、動的にプログラムを保持するもの、その場合のサーバやクライアントとなるコンピュータシステム内部の揮発性メモリのように、一定時間プログラムを保持しているものも含んでもよい。また上記プログラムは、前述した機能の一部を実現するためのものであってもよく、さらに前述した機能をコンピュータシステムに既に記録されているプログラムとの組み合わせで実現できるものであってもよく、P L D (Programmable Logic Device) や F P G A (Field Programmable Gate Array) 等のハードウェアを用いて実現されるものであってもよい。

【0112】

以上、図面を参照して本発明の実施の形態を説明してきたが、上記実施の形態は本発明の例示に過ぎず、本発明が上記実施の形態に限定されるものではないことは明らかである。したがって、本発明の技術思想及び範囲を逸脱しない範囲で構成要素の追加、省略、置換、その他の変更を行ってもよい。

【産業上の利用可能性】

【0113】

2つの画像を結合する際に、画像の継ぎ目を気づきにくい位置に配置するようにしたため、視聴品質の低下を抑制した仮想全天球画像を生成することが不可欠な用途にも適用できる。

【符号の説明】

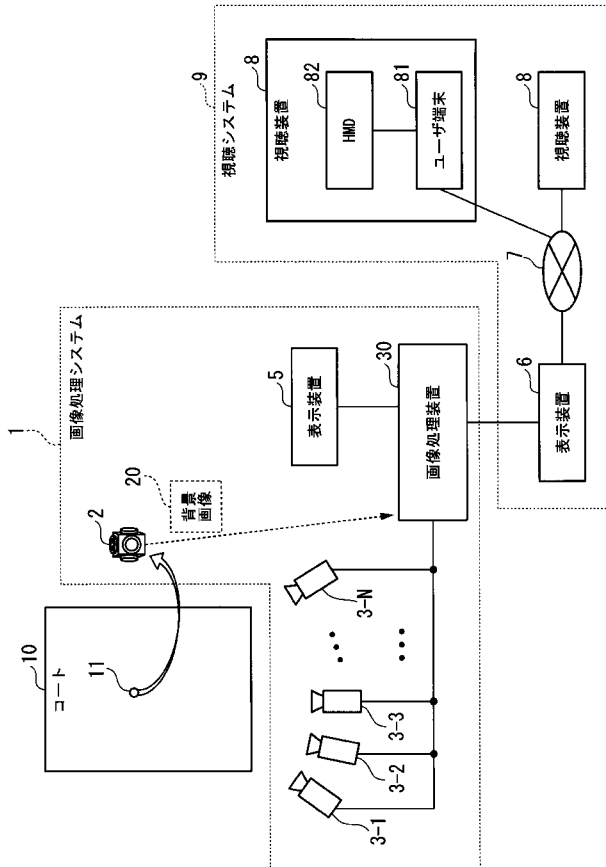
【0114】

10・・・コート、11・・・仮想視点、1・・・画像処理システム、2・・・全天球カメラ、3・・・カメラ群、5・・・表示装置、30・・・画像処理装置、20・・・背景画像、6・・・画像サーバ、7・・・ネットワーク、8・・・視聴装置、81・・・ユーザ端末、82・・・HMD、9・・・視聴システム、31・・・オブジェクト解析部、32・・・奥行取得部、33・・・合成情報取得部、34・・・画像入力部、35・・・画像切り出し部、36・・・画像合成部、37・・・表示処理部、38・・・入力部、39・・・境界決定部、301・・・前景画像格納部、302・・・背景画像格納部、303・・・オブジェクト情報格納部、304・・・合成情報テーブル

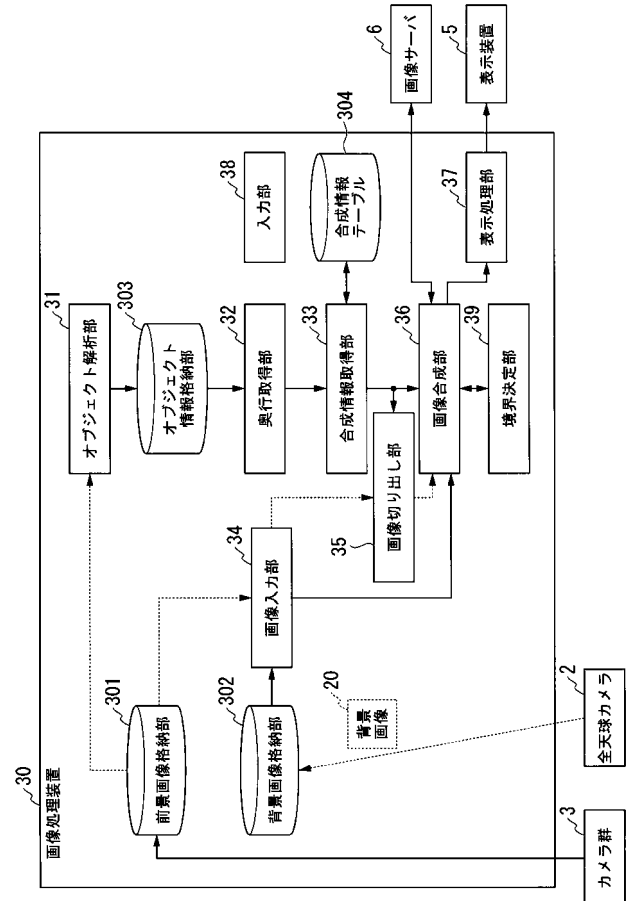
10

20

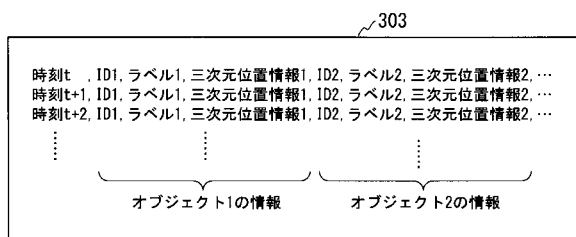
【図 1】



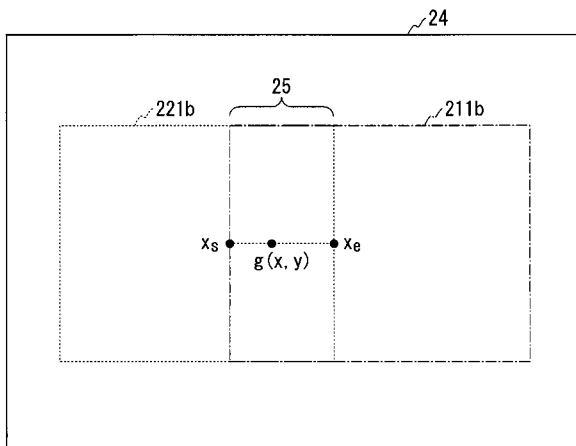
【図 2】



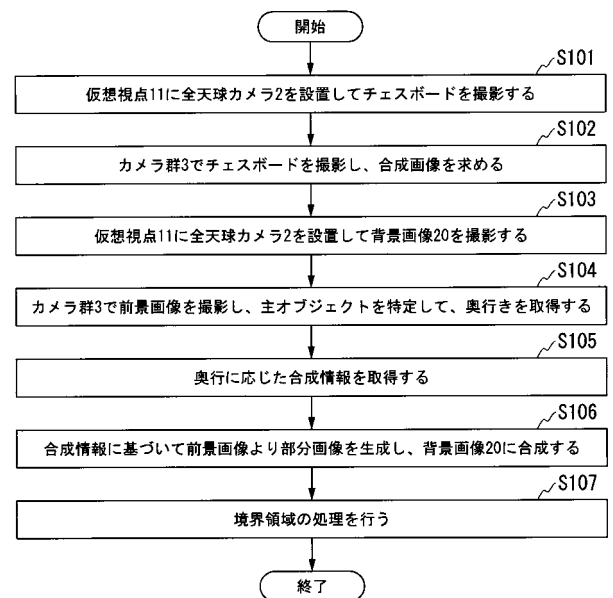
【図 3】



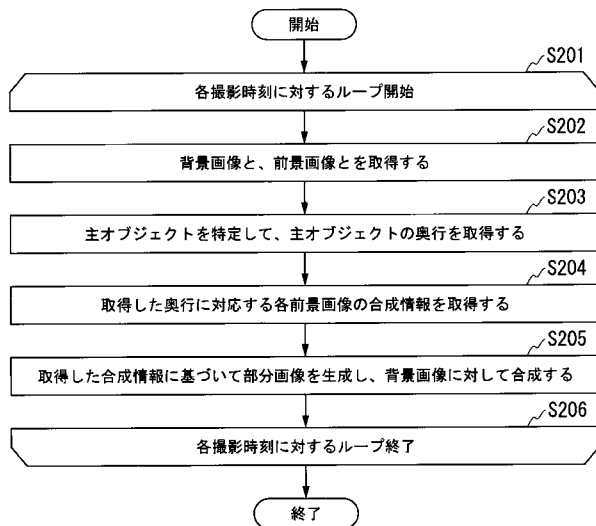
【図 4】



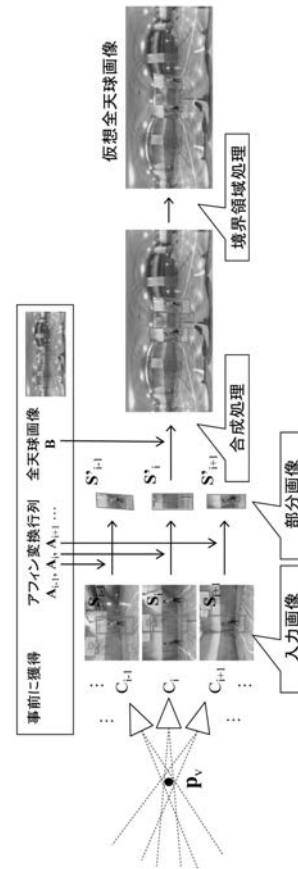
【図 5】



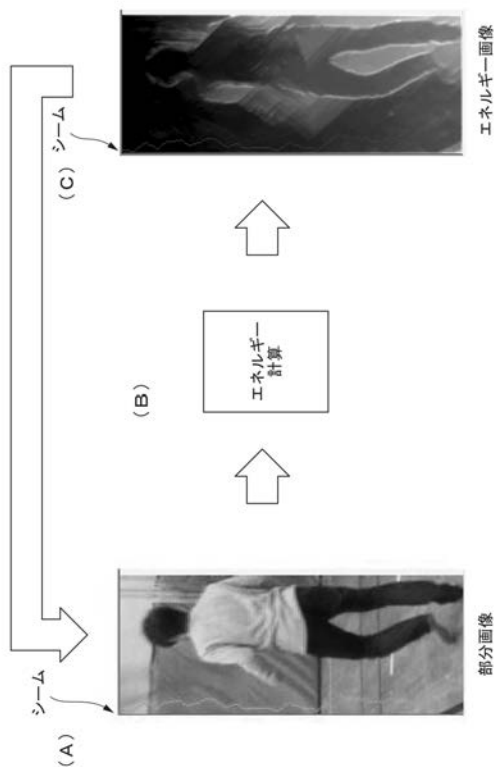
【図 6】



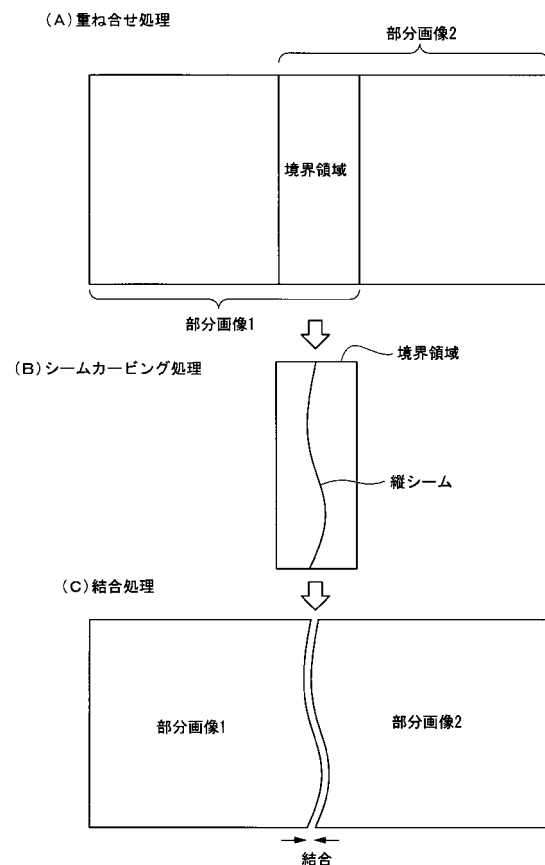
【図 7】



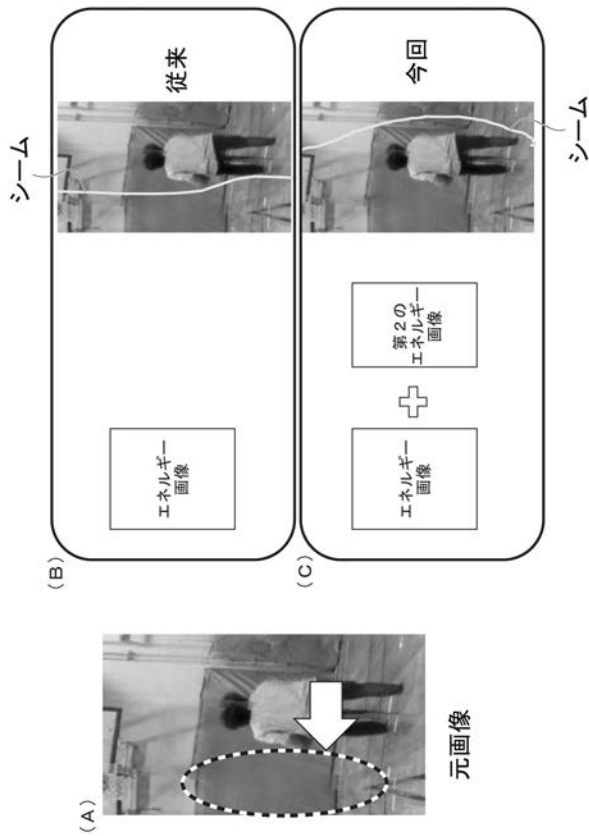
【図 8】



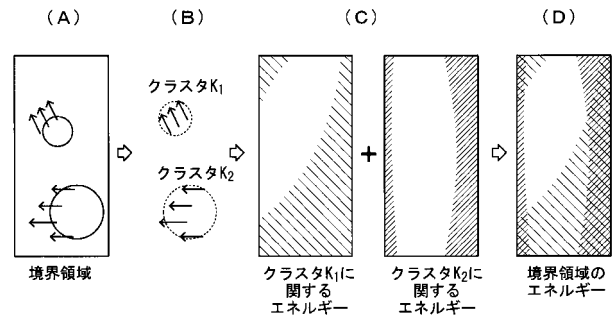
【図 9】



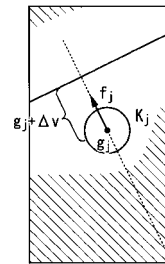
【図 10】



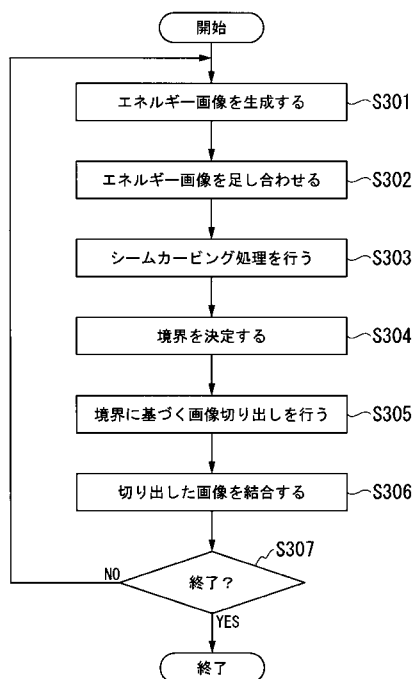
【図 11】



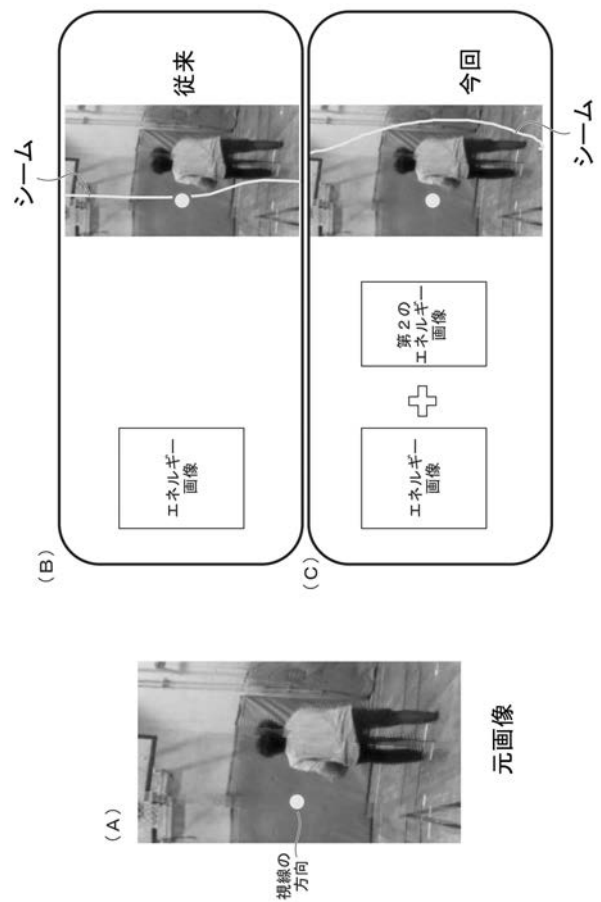
【図 12】



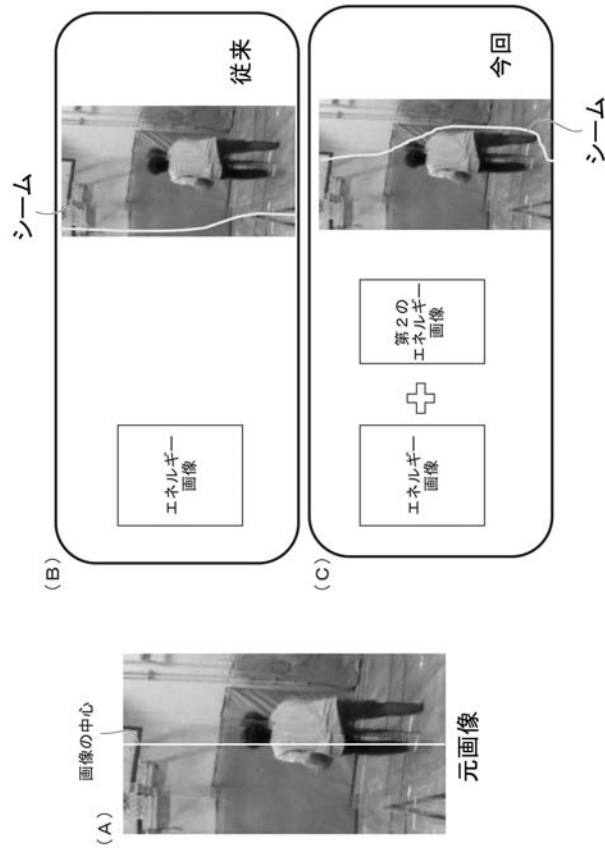
【図 13】



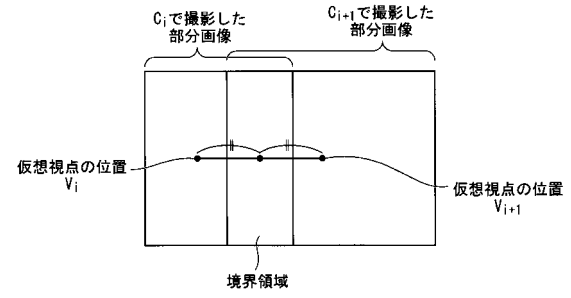
【図 14】



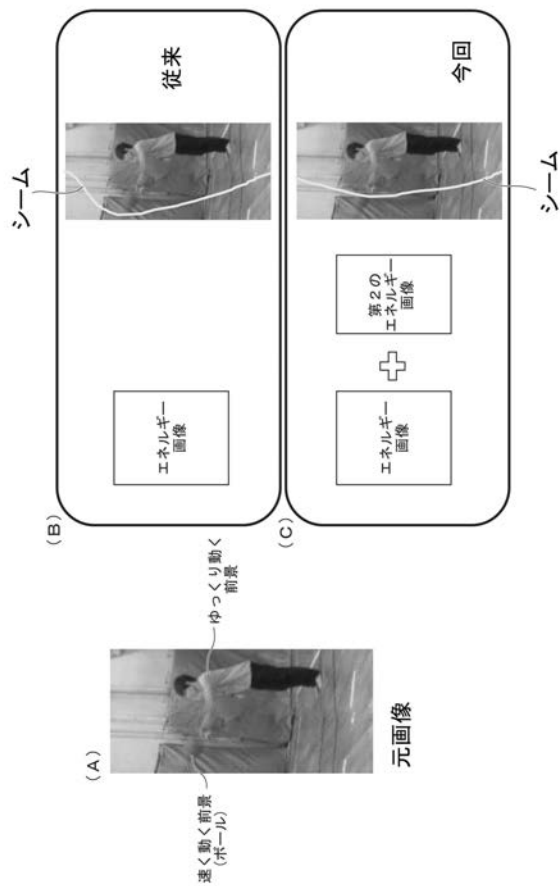
【図 15】



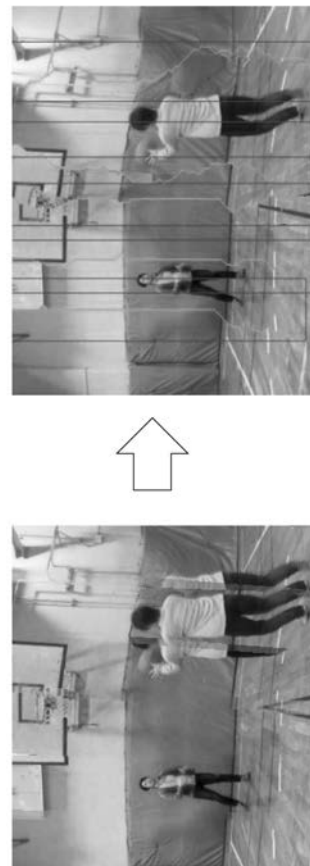
【図 16】



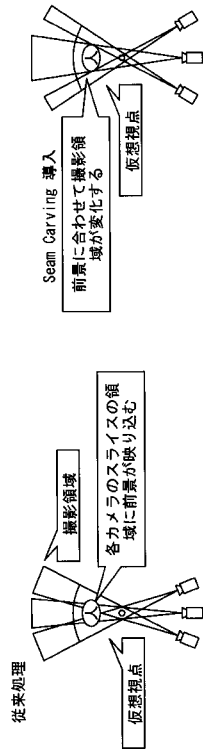
【図 17】



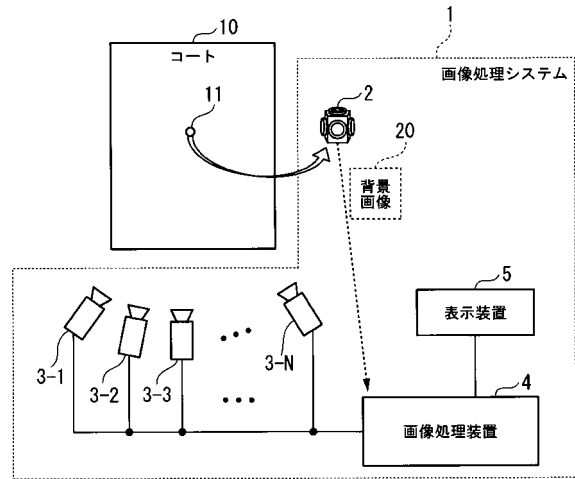
【図 18】



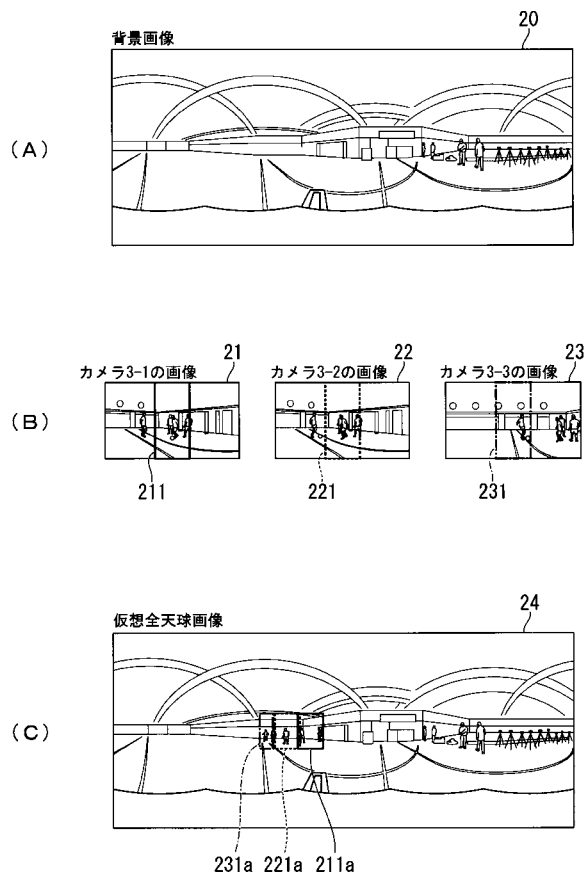
【図 19】



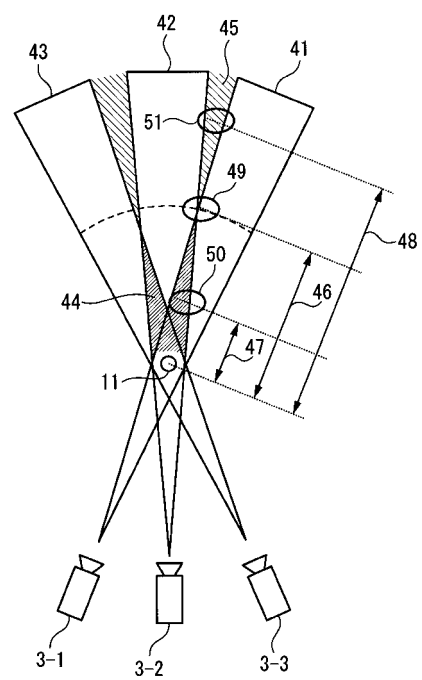
【図 20】



【図 21】



【図 22】



フロントページの続き

(72)発明者 小島 明

東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内

Fターム(参考) 5B057 CA08 CA12 CA16 CB08 CB12 CB16 CE10 CH11 DA07 DA16
DB02 DB09 DC22 DC32
5C122 DA03 EA47 EA59 EA67 FA02 FA18 FH04 FH07 FH09 FH12
FH14 GC02 GC14 HB01