

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7271099号
(P7271099)

(45)発行日 令和5年5月11日(2023.5.11)

(24)登録日 令和5年4月28日(2023.4.28)

(51)国際特許分類		F I	
H 0 4 N	5/92 (2006.01)	H 0 4 N	5/92 0 1 0
H 0 4 N	5/765(2006.01)	H 0 4 N	5/765
G 0 6 T	19/00 (2011.01)	G 0 6 T	19/00 A
G 0 6 T	1/00 (2006.01)	G 0 6 T	1/00 2 0 0 C
G 0 6 T	5/50 (2006.01)	G 0 6 T	1/00 5 0 0 Z
請求項の数 26 (全26頁) 最終頁に続く			
(21)出願番号	特願2018-136204(P2018-136204)	(73)特許権者	000001007 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
(22)出願日	平成30年7月19日(2018.7.19)	(74)代理人	110001243 弁理士法人谷・阿部特許事務所
(65)公開番号	特開2020-14159(P2020-14159A)	(72)発明者	伊藤 裕尚 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
(43)公開日	令和2年1月23日(2020.1.23)	(72)発明者	小沼 和文 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
審査請求日	令和3年7月19日(2021.7.19)	(72)発明者	前田 充 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
		審査官	鈴木 順三 最終頁に続く

(54)【発明の名称】 ファイルの生成装置およびファイルに基づく映像の生成装置

(57)【特許請求の範囲】

【請求項1】
複数の視点で撮影された画像に基づく仮想視点映像の生成に用いる素材データを取得する取得手段と、
前記取得手段により取得された素材データと前記素材データの種別を示す種別情報とを格納したファイルを生成するファイル生成手段と、
前記ファイル生成手段により生成された前記ファイルを出力する出力手段とを備え、
前記ファイル生成手段は、前記ファイルの生成において、前記ファイルに格納される各素材データのそれぞれに対しメタデータを付与して、前記ファイルに格納する、
ことを特徴とするファイル生成装置。

【請求項2】
前記素材データは、前記撮影を行うカメラの位置、姿勢、画角の情報を少なくとも含むカメラパラメータを含み、
前記種別情報は、前記素材データが前記カメラパラメータを含むことを示す情報を含むことを特徴とする請求項1に記載のファイル生成装置。

【請求項3】
前記素材データは、前記撮影を行うカメラの撮影画像に含まれる前景オブジェクトを示す前景画像を含み、
前記種別情報は、前記素材データが前景画像を含むことを示す情報を含むことを特徴とす

る請求項 1 又は 2 に記載のファイル生成装置。

【請求項 4】

前記素材データは、前記撮影を行うカメラの撮影画像に含まれる前景オブジェクトの 3 次元形状を表す前景 3 D モデルを含み、

前記種別情報は、前記素材データが前景 3 D モデルを含むことを示す情報を含むことを特徴とする請求項 1 乃至 3 のいずれか 1 項に記載のファイル生成装置。

【請求項 5】

前記素材データは、前記撮影を行うカメラの撮影画像に含まれる背景オブジェクトの 3 次元形状を表す背景 3 D モデルを含み、

前記種別情報は、前記素材データが背景 3 D モデルを含むことを示す情報を含むことを特徴とする請求項 1 乃至 4 のいずれか 1 項に記載のファイル生成装置。

10

【請求項 6】

前記ファイル生成手段は、前記複数の視点で撮影を行う各カメラの撮影画像を、前記素材データとしてファイルを生成することを特徴とする請求項 1 乃至 5 のいずれか 1 項に記載のファイル生成装置。

【請求項 7】

前記ファイル生成手段は、前記各カメラの撮影画像のデータを、フレーム単位又はカメラ単位で前記ファイルに格納することを特徴とする請求項 6 に記載のファイル生成装置。

【請求項 8】

複数の視点で撮影された画像に基づく仮想視点映像の生成に用いる素材データを取得する取得手段と、

20

前記取得手段により取得された素材データと前記素材データの種別を示す種別情報とを格納したファイルを生成するファイル生成手段と、

前記ファイル生成手段により生成された前記ファイルを出力する出力手段と、
を備え、

前記取得手段は、前記複数の視点で撮影を行うカメラの台数の情報、及び、各カメラの ID と紐付けられた各カメラの位置、姿勢、画角の情報を少なくとも含むカメラパラメータを生成して取得し、

前記ファイル生成手段は、前記ファイルの生成において、前記カメラパラメータを前記素材データとして前記ファイルに格納する

30

ことを特徴とするファイル生成装置。

【請求項 9】

前記カメラパラメータには、さらに、前記各カメラの ID と紐付けられた各カメラの撮影に関する情報であって、撮影画像の幅、高さ、絞り値、シャッタースピード、ISO 感度、レンズの焦点距離、被写体までの距離の少なくとも一つ以上の情報を含む、ことを特徴とする請求項 8 に記載のファイル生成装置。

【請求項 10】

複数の視点で撮影された画像に基づく仮想視点映像の生成に用いる素材データを取得する取得手段と、

前記取得手段により取得された素材データと前記素材データの種別を示す種別情報とを格納したファイルを生成するファイル生成手段と、

40

前記ファイル生成手段により生成された前記ファイルを出力する出力手段と、
を備え、

前記取得手段は、前記複数の視点で撮影を行う各カメラの撮影画像に対し前景背景分離を行い、

前記ファイル生成手段は、前記ファイルの生成において、前記前景背景分離の結果の少なくとも一部を、前記素材データとして前記ファイルに格納する

ことを特徴とするファイル生成装置。

【請求項 11】

複数の視点で撮影された画像に基づく仮想視点映像の生成に用いる素材データを取得する

50

取得手段と、

前記取得手段により取得された素材データと前記素材データの種別を示す種別情報とを格納したファイルを生成するファイル生成手段と、

前記ファイル生成手段により生成された前記ファイルを出力する出力手段と、
を備え、

前記取得手段は、前記複数の視点で撮影を行う各カメラの撮影画像における前景シルエット画像と、各カメラの位置及び姿勢を少なくとも含むカメラパラメータとに基づき、前景オブジェクトの3次元形状を表す前景3Dモデルを生成して取得し、

前記ファイル生成手段は、前記ファイルの生成において、前記前景3Dモデルを前記素材データとして前記ファイルに格納する

ことを特徴とするファイル生成装置。

【請求項12】

前記ファイル生成手段は、前記前景3Dモデルのデータを、モデル単位で前記ファイルに格納することを特徴とする請求項4又は11に記載のファイル生成装置。

【請求項13】

複数の視点で撮影された画像に基づく仮想視点映像の生成に用いる素材データを取得する取得手段と、

前記取得手段により取得された素材データと前記素材データの種別を示す種別情報とを格納したファイルを生成するファイル生成手段と、

前記ファイル生成手段により生成された前記ファイルを出力する出力手段と、
を備え、

前記ファイル生成手段は、前記ファイルの生成において、前記複数の視点で撮影された画像における背景オブジェクトの3次元形状を表す背景3Dモデルを、前記素材データとして前記ファイルに格納する

ことを特徴とするファイル生成装置。

【請求項14】

複数の視点で撮影された画像に基づく仮想視点映像の生成に用いる素材データと、前記素材データの種別を示す種別情報と、を格納したファイルを入力する入力手段と、

前記ファイルに格納された素材データを用いた前記仮想視点映像の生成が可能かどうかを、前記種別情報に基づいて判定する判定手段と、

前記判定手段で生成可能と判定された場合に、前記入力手段により入力されたファイルに格納された素材データを用いて仮想視点映像を生成する生成手段と、

を備えたことを特徴とする映像生成装置。

【請求項15】

前記判定手段は、前記種別情報で特定される素材データが、前記生成手段の機能に応じて要求される素材データを満たす場合、仮想視点映像が生成可能であると判定することを特徴とする請求項14に記載の映像生成装置。

【請求項16】

前記判定手段における判定結果をユーザに通知する通知手段をさらに備えたことを特徴とする請求項14又は15に記載の映像生成装置。

【請求項17】

複数の視点で撮影された画像に基づく仮想視点映像の生成に用いる素材データと、前記素材データの種別を示す種別情報と、を格納したファイルを入力する入力手段と、

前記入力手段により入力されたファイルに格納された素材データを用いて仮想視点映像を生成する生成手段と、

を備え、

前記生成手段は、前記複数の視点で撮影を行う各カメラの撮影画像における前景オブジェクトの3次元形状を表す前景3Dモデルに対し色づけする機能を少なくとも有する、

ことを特徴とする映像生成装置。

【請求項18】

10

20

30

40

50

前記前景 3 D モデルの種類は、点群で表現された 3 D モデル、メッシュで表現された 3 D モデル、ビルボード形式の 3 D モデルのうちいずれかであることを特徴とする請求項 1 7 に記載の映像生成装置。

【請求項 1 9】

前記生成手段は、

前記ファイルに複数種類の前景 3 D モデルが含まれる場合、所定の基準に従って、使用する前景 3 D モデルを決定し

決定した前記前景 3 D モデルを用いて、前記仮想視点映像を生成する

ことを特徴とする請求項 1 8 に記載の映像生成装置。

【請求項 2 0】

前記所定の基準は、前記前景 3 D モデルの種類について予め定めた優先順であることを特徴とする請求項 1 9 に記載の映像生成装置。

【請求項 2 1】

前記所定の基準は、前記仮想視点映像の生成モードに応じてどの種類の前景 3 D モデルを使用するかを定めた基準であることを特徴とする請求項 1 9 に記載の映像生成装置。

【請求項 2 2】

前記生成手段は、

前記ファイルに格納された素材データの中に含まれる、前記各カメラの撮影画像における前景シルエット画像と、前記各カメラの位置及び姿勢を少なくとも含むカメラパラメータとを用いて前景 3 D モデルを生成し、

生成した前記前景 3 D モデルに対して色づけを行なって、前記仮想視点映像を生成する

ことを特徴とする請求項 1 7 乃至 2 1 のいずれか 1 項に記載の映像生成装置。

【請求項 2 3】

前記生成手段は、

前記ファイルに格納された素材データの中に含まれる前記各カメラの撮影画像に対し前景背景分離を行って前景シルエット画像を生成し、

前記ファイルに格納された素材データの中に含まれる、前記各カメラの位置及び姿勢を少なくとも含むカメラパラメータと、生成した前記前景シルエット画像とを用いて、前景 3 D モデルを生成し、

生成した前記前景 3 D モデルに対して色づけを行って、前記仮想視点映像を生成する

ことを特徴とする請求項 1 7 乃至 2 1 のいずれか 1 項に記載の映像生成装置。

【請求項 2 4】

前記生成手段は、

設定された仮想視点に基づいて、前記仮想視点映像の生成に必要な素材データを決定し、

決定した前記必要な素材データを前記ファイルの中から取得し、

取得した素材データを用いて、前記仮想視点映像を生成する

ことを特徴とする請求項 1 4 乃至 2 3 のいずれか 1 項に記載の映像生成装置。

【請求項 2 5】

複数の視点で撮影された画像に基づく仮想視点映像の生成に用いる素材データを取得する取得ステップと、

前記取得ステップにより取得された素材データと前記素材データの種別を示す種別情報とを格納し、格納される各素材データのそれぞれに対しメタデータを付与したファイルを生成するファイル生成ステップと、

前記ファイル生成ステップにより生成された前記ファイルを出力する出力ステップと、
を含むことを特徴とするファイル生成方法。

【請求項 2 6】

コンピュータを、請求項 1 乃至 1 3 のいずれか 1 項に記載のファイル生成装置、又は請求項 1 4 乃至 2 4 のいずれか 1 項に記載の映像生成装置として機能させるためのプログラム。

10

20

30

40

50

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、仮想視点映像のファイルの生成及びその取扱いに関する。

【背景技術】

【0002】

複数のカメラを用いてオブジェクトを多方向から撮像して得られた複数視点映像から、ユーザが指定した任意の仮想視点から見た映像を再現する技術として、仮想視点映像生成技術がある。仮想視点映像の生成には、複数視点映像、前景画像、前景シルエット画像、前景3Dモデル、カメラパラメータ、背景画像、背景3Dモデルといった様々な映像素材が用いられる。これら映像素材は生成後にデータベースに蓄積され、ユーザが指定した仮想視点に基づいてデータベースから適切な素材データを取得して、仮想視点映像が生成される（特許文献1を参照）。

10

【先行技術文献】

【特許文献】

【0003】

【文献】特開2017-211828号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

20

しかしながら、データベースに格納される素材データのフォーマットを解釈できない装置は、当該素材データに基づく仮想視点映像を再生することが困難である。

【0005】

例えば、データベースに格納される素材データのフォーマットが各素材を生成する装置毎にオリジナルのものが定義され、その取得方法も素材によって異なっていた場合、世の中に広く普及した端末（例えばスマートフォンなど）が仮想視点映像を生成することが困難であった。

【0006】

本発明は、上記の課題に鑑みてなされたものであり、より多くの装置が仮想視点映像を生成できるようにすることを目的とする。

30

【課題を解決するための手段】

【0007】

本発明に係るファイル生成装置は、複数の視点で撮影された画像に基づく仮想視点映像の生成に用いる素材データを取得する取得手段と、前記取得手段により取得された素材データと前記素材データの種別を示す種別情報とを格納したファイルを生成するファイル生成手段と、前記ファイル生成手段により生成された前記ファイルを出力する出力手段とを備え、前記ファイル生成手段は、前記ファイルの生成において、前記ファイルに格納される各素材データのそれぞれに対しメタデータを付与して、前記ファイルに格納する、ことを特徴とする。

【発明の効果】

40

【0008】

本発明によれば、より多くの装置が仮想視点映像を生成できるようになる。

【図面の簡単な説明】

【0009】

【図1】実施形態1に係る、ファイル生成装置の機能ブロック図

【図2】実施形態1に係る、VvVmファイル生成処理の流れを示すフローチャート

【図3】(a)はVvVmファイルのデータ構造の一例を示す図、(b)は格納データ情報のビットアサインの一例を示す図

【図4】(a)は各カメラの撮影画像のメタデータの一例を示す図、(b)はカメラパラメータのメタデータの一例、(c)はそのビットアサインを示す図、(d)は背景3Dモ

50

デルのメタデータの一例、(e)はそのビットアサインを示す図

【図5】(a)及び(b)は、ボックスの構造の一例を示す図

【図6】V v v mファイルの一例を示す図

【図7】(a)はc i m gボックスの一例、(b)はそのシンタックスを示す図

【図8】(a)はc a m pボックスの一例、(b)はそのシンタックスを示す図

【図9】(a)はf i m gボックスの一例、(b)はそのシンタックスを示す図

【図10】(a)はb i m gボックスの一例、(b)はそのシンタックスを示す図

【図11】(a)はf m d lボックスの一例、(b)はそのシンタックスを示す図

【図12】(a)はb m d lボックスの一例、(b)はそのシンタックスを示す図

【図13】(a)は実施形態1に係る仮想視点映像生成装置の機能ブロック図、(b)は映像化部の内部構成を示すブロック図

10

【図14】実施形態1に係る、V v v mファイルから仮想視点映像を生成する処理の流れを示すフローチャート

【図15】V v v mファイルのデータ構造のバリエーションを示す図

【図16】(a)、(c)、(e)はメタデータの一例を示す図であり、(b)、(d)、(f)はそれらのビットアサインを示す図

【図17】V v v mファイルのデータ構造のバリエーションを示す図

【図18】V v v mファイルのデータ構造のバリエーションを示す図

【図19】実施形態2に係る、ファイル生成装置の機能ブロック図

【図20】(a)は実施形態2に係る仮想視点映像生成装置の機能ブロック図、(b)は映像化部の内部構成を示すブロック図、(c)は前景色づけ部の内部構成を示すブロック図
【発明を実施するための形態】

20

【0010】

以下、本発明の実施形態について、図面を参照して説明する。なお、以下の実施形態は本発明を限定するものではなく、また、本実施形態で説明されている特徴の組み合わせの全てが本発明の解決手段に必須のものとは限らない。

【実施形態1】

【0011】

本実施形態の仮想視点映像生成システムは、仮想視点映像の素材データを包含したファイルを生成するファイル生成装置と、当該ファイルを用いて仮想視点映像を生成する仮想視点映像生成装置とで構成される。本明細書においては、仮想視点映像(Virtual viewpoint video)の素材(material)データを包含するファイルを、「V v v mファイル」と呼ぶこととする。なお、仮想視点映像とは、エンドユーザ及び/又は選任のオペレータ等が自由に仮想カメラの位置及び姿勢を操作することによって生成される映像であり、自由視点映像や任意視点映像などとも呼ばれる。また、仮想視点映像は、動画であっても、静止画であっても良い。以下では、仮想視点映像が動画である場合の例を中心に説明する。

30

【0012】

ファイル生成装置は、複数視点映像やカメラパラメータといった仮想視点映像の生成に必要な素材データを格納したV v v mファイルを生成する。この際、V v v mファイルには、格納されている素材データを特定する情報である格納データ情報が付加される。仮想視点映像生成装置は、V v v mファイルに含まれる格納データ情報に基づき、自装置において仮想視点映像の生成が可能と分かれば、V v v mファイル内の素材データを用いて仮想視点映像を生成する。

40

【0013】

(ファイル生成装置)

図1は、本実施形態に係るファイル生成装置10の機能ブロック図である。ファイル生成装置10は、データ入力部11、カメラパラメータ生成部12、ファイル生成部13、ファイル出力部14で構成される。なお、ファイル生成装置10は、CPU、RAM、ROM、HDDといった、演算やデータ記憶のためのハードウェアを備えた情報処理装置であり、例えばPCなどである。図1に示す各部の機能は、CPUが、所定のプログラムを

50

R O M又はH D Dから読み込んでR A Mに展開し、これをC P Uが実行するなどして実現される。以下、図 1 に示す各部の役割について説明する。

【 0 0 1 4 】

データ入力部 1 1 は、V v v mファイルに格納される各種の素材データ或いはその元になるデータの入力を、外部インタフェース（不図示）を介して受け付ける。本実施形態の場合、撮影シーンに設置された複数のカメラに関するカメラ情報、当該複数のカメラで撮影された複数視点映像、スタジアム等の施設（背景オブジェクト）の 3 次元形状を表す背景 3 Dモデルのデータが、外部装置から入力される。例えば、カメラ情報は各カメラのキャリブレーションを行う P C から、複数視点映像データは複数のカメラから、背景 3 Dモデルのデータはこれを生成する P C からそれぞれ入力される。カメラ情報には、各カメラの位置、姿勢（視線方向）、画角といった情報が含まれる。撮影途中でカメラの台数や各カメラの位置、姿勢、画角が変わった場合には、カメラ情報が再取得される。また、複数視点映像を構成するカメラ毎の撮影画像は、各カメラに付加されたカメラ I D と関連づけられ、どのカメラで撮影した画像かを判別できるようになっている。入力されたカメラ情報はカメラパラメータ生成部 1 2 へ送られ、複数視点映像及び背景 3 Dモデルの各データは、ファイル生成部 1 3 に送られる。

10

【 0 0 1 5 】

カメラパラメータ生成部 1 2 は、データ入力部 1 1 から受け取ったカメラ情報に基づき、共通の注視点を向いた複数のカメラからなるカメラ群単位で、カメラパラメータを生成する。このカメラパラメータには、各カメラ群を構成するカメラの台数の情報、各カメラの I D（カメラの識別番号等）と紐付けられた、各カメラの位置、姿勢、画角の情報が少なくとも含まれる。これらに加え、シャッタースピード、絞り値、I S O感度、色温度、レンズの焦点距離、撮影ターゲットまでの距離などの情報を含めてもよい。生成されたカメラパラメータは、ファイル生成部 1 3 に送られる。

20

【 0 0 1 6 】

ファイル生成部 1 3 は、データ入力部 1 1 から受け取った複数視点映像と背景 3 Dモデルのデータ、カメラパラメータ生成部 1 2 から受け取ったカメラパラメータを、共通フォーマットで 1 つのファイルにまとめる。そして、素材データとして、カメラパラメータ、背景 3 Dモデル、複数視点映像が格納されていることを示す格納データ情報を生成し、当該ファイルに素材データと共に格納して、上述の V v v mファイルを生成する。なお、撮影途中で 1 又は複数のカメラの位置、姿勢、画角等が変更された場合には、変更後の内容が時刻情報と共にそのカメラパラメータに付加され、V v v mファイル内に格納されることになる。こうすることで、格納されている複数視点映像を撮影しているカメラの位置、姿勢、画角等が途中で変わっていた場合においても、その事実と内容を V v v mファイル内のカメラパラメータから把握することが可能となる。また、上記ファイル生成の際には、例えばファイル転送時の容量制限等を考慮して、複数の V v v mファイルに分割してもよい。このような分割によるファイル生成を行った際には、本来は 1 つのファイルになるはずだった複数の分割ファイル同士を相互に関連付けるといった処理も併せてなされる。

30

【 0 0 1 7 】

ファイル出力部 1 4 は、ファイル生成部 1 3 で生成された V v v mファイルを、仮想視点映像を生成可能な各種装置に出力する。出力態様としては、例えば、P C への有線送信、スマートフォンやタブレットといった携帯端末への無線送信、ネットワークを介してのサーバへのアップロード、携帯可能な記憶媒体への格納などが考えられる。

40

【 0 0 1 8 】

以上が、ファイル生成装置 1 0 の主要機能を担う各部の説明である。

【 0 0 1 9 】

（ V v v mファイルの生成 ）

続いて、ファイル生成装置 1 0 で V v v mファイルを生成する処理の流れを説明する。図 2 は、本実施形態に係る、V v v mファイル生成処理の流れを示すフローチャートである。なお、各処理の冒頭における記号「 S 」はステップを意味する。

50

【 0 0 2 0 】

S 2 0 1では、データ入力部 1 1 が、背景 3 Dモデルのデータ及び撮影を行う各カメラの情報を、外部 P C 等から受信する。受信した背景 3 Dモデルのデータはファイル生成部 1 3 に送られ、カメラ情報はカメラパラメータ生成部 1 2 に送られる。続く S 2 0 2 では、カメラパラメータ生成部 1 2 が、データ入力部 1 1 から受け取ったカメラ情報に基づき、上述のカメラパラメータを生成する。生成したカメラパラメータは、ファイル生成部 1 3 に送られる。

【 0 0 2 1 】

S 2 0 3では、ファイル生成部 1 3 が、データ入力部 1 1 から受け取った背景 3 Dモデルと、カメラパラメータ生成部 1 2 から受け取ったカメラパラメータを、V v v mファイルに格納する。

10

【 0 0 2 2 】

S 2 0 4では、データ入力部 1 1 が、複数視点映像の取得を開始する。例えば、ディジチェーン方式等で接続された複数のカメラに対し撮影開始信号が送信され、これに応じて各カメラの撮影画像のデータがフレーム単位で順次受信される。撮影開始信号の発信元は、ファイル生成装置 1 0 でもよいし、別の装置でもよい。取得された複数視点映像のデータは、ファイル生成部 1 3 に順次送られる。

【 0 0 2 3 】

S 2 0 5では、ファイル生成部 1 3 が、データ入力部 1 1 から受け取った複数視点映像のデータを V v v mファイル内に格納する。格納は例えばフレーム単位で実行され、複数視点映像を構成する各撮影画像がどのカメラで撮影されたのかが識別可能なように格納される。

20

【 0 0 2 4 】

S 2 0 6では、データ入力部 1 1 が、複数視点映像の取得が完了したか否かを判定する。例えば、撮影を行っている複数のカメラに対し撮影停止信号が送信され、撮影された全フレーム分のデータの受信が完了していれば、複数視点映像の取得完了と判定して S 2 0 7に進む。一方、撮影が継続中で未受信のフレームデータがあれば、S 2 0 5に戻って複数視点映像の取得とその V v v mファイルへの格納を続行する。

【 0 0 2 5 】

S 2 0 7では、ファイル生成部 1 3 が、V v v mファイルに格納された素材データの中身を示す格納データ情報を生成し、V v v mファイルに付加する。完成した V v v mファイルは、ファイル出力部 1 4 に送られる。

30

【 0 0 2 6 】

そして、S 2 0 8では、ファイル出力部 1 4 が、ユーザ等が指定する出力先としての仮想視点映像生成装置 2 0 に、V v v mファイルを出力する。

【 0 0 2 7 】

以上が、ファイル生成装置 1 0 における、V v v mファイル生成処理の流れである。なお、図 2 のフローでは、ファイル生成部 1 3 が全フレームデータの取得処理が終了した後、1つの V v v mファイルにまとめているが、これに限定されない。例えば、フレーム単位で生成したファイルを順次出力し、出力先の外部装置で1つのファイルにまとめても構

40

【 0 0 2 8 】

(V v v mファイルのデータ構造)

図 3 (a) は、本実施形態に係る、V v v mファイルのデータ構造の一例を示す図である。図 3 (a) に示す V v v mファイル 3 0 0 は、データバージョン 3 0 1、格納データ情報 3 0 2、カメラパラメータ 3 0 3、背景 3 Dモデル 3 0 4、フレームデータ 3 0 5 から構成される。フレームデータ 3 0 5 は、第 1 フレームから第 N フレームまで存在し、1つのフレームデータには M 台のカメラによる撮影画像が含まれている。この場合において、N 及び M は 1 以上の自然数である。

【 0 0 2 9 】

50

V v v mファイル3 0 0にどのような素材データが格納されているかは、格納データ情報3 0 2に示される。図3 (b)は格納データ情報3 0 2のビットアサインの一例を示している。いま、格納データ情報3 0 2は3 2ビットの値を持ち、各ビットは、“ 1 ”であればその対象データが格納されていることを示し、“ 0 ”であればその対象データが格納されていないことを示す。図3 (a)で示すV v v mファイル3 0 0には、各カメラの撮影画像（複数視点映像）、カメラパラメータ、背景3 Dモデルの3種類の素材データが格納される。この場合、b 0、b 1、b 5の各ビットが“ 1 ”、他のビットは“ 0 ”となる。b 1 0 ~ b 1 3は独自に定義することができる。例えば、データ名称をデプスマップ或いは距離画像とした、前景画像と同じフォーマットで各画素値がカメラから被写体までの距離を表す画像を追加してもよい。また、仮想視点映像の生成を高速化するためのデータや高画質化するためのデータを追加してもよい。

10

【 0 0 3 0 】

また、格納される素材データそれぞれについても、それがどのようなデータであるかを示す付属情報（メタデータ）が付与される。図4 (a)は各カメラの撮影画像のメタデータの一例を示している。図4 (b)はカメラパラメータのメタデータの一例を示し、同 (c)は格納カメラパラメータのビットアサインを示している。図4 (d)は背景3 Dモデルのメタデータの一例を示し、同 (e)は格納背景3 Dモデルのビットアサインを示している。なお、カメラパラメータ3 0 3の先頭（最初の4 byte）には設置カメラの台数を示す情報が含まれ、この台数情報に続いてカメラパラメータに関するメタデータが、カメラの台数分だけ続くことになる。

20

【 0 0 3 1 】

以上のようにして、仮想視点映像の生成に必要な様々な素材データが共通のフォーマットで定義され、まとめてファイル化される。なお、カメラパラメータ3 0 3は、1つのV v v mファイルに対して1つでもよいが、フレーム単位でカメラパラメータを持つようにしてもよい。フレーム単位でカメラパラメータを持つことで、撮影途中でカメラの位置・姿勢が変化するケースや、一部のカメラが故障してカメラ群を構成するカメラ台数が変化するケースにも対応することができる。

【 0 0 3 2 】

（ V v v mファイルの具体例 ）

続いて、I S O B M F F（ISO Base Media File Format ISO/IEC 14496-12 MPE G-4 Part 12）規格に準拠した、本実施形態のV v v mファイルの具体例を説明する。I S O B M F F規格では、「ボックス」という単位でファイルが構成される。図5 (a)はボックスの構造を示す図であり、サイズ情報を格納する領域、タイプ情報を格納する領域、及びデータを格納する領域で構成される。また、図5 (b)に示すように、ボックスの中にデータとしてさらにボックスを含む構造とすることも可能である。

30

【 0 0 3 3 】

図6は、I S O B M F Fに準拠したV v v mファイルの内部構造の一例を示している。I S O B M F Fに準拠したV v v mファイル6 0 0は、f t y p 6 0 1、m o o v 6 0 2、c i m g 6 0 3、c a m p 6 0 4、f i m g 6 0 5、b i m g 6 0 6、f m d l 6 0 7、b m d l 6 0 8及びm d a t 6 0 9の各ボックスで構成されている。以下、各ボックスについて説明する。

40

【 0 0 3 4 】

f t y pボックス（File Type Compatibility Box）6 0 1は、ファイル内の最初に配置されるボックスである。f t y pボックス6 0 1には、ファイルフォーマットの情報、ボックスのバージョンを示す情報、他のファイルフォーマットとの互換性に関する情報、ファイルを作成したメーカー名称の情報などが記載される。V v v mファイル内に格納される各素材データの種別を示す上述の格納データ情報3 0 2は、このf t y pボックス6 0 1に格納されても良い。また、m o o vボックス（Movie Box）6 0 2は、ファイル内にどんなデータがどのように格納されているかを明らかにするボックスであり、メディアデータを管理するための時間軸やアドレスなどの情報が入る。そして、m d a tボックス（

50

Media Data Box) 609には、動画や音声といった、仮想視点映像を生成する際に使用するメディアデータ(素材データ)が入る。mdatボックス609にどのようにデータを格納したのかをmoovボックス602に記載しておくことで、これらメディアデータへのアクセスが可能になる。ftypボックス601、moovボックス602及びmdatボックス609は、ISO BMFF準拠のファイルにおいて共通で設けられるボックスである。これに対し、camp604、img605、bimg606、fmdl607及びbmdl608の各ボックスは、Vvwmファイルに特有のボックスである。以下、Vvwmファイルに特有のボックスについて具体例を交えて説明する。

【0035】

cimgボックス(Camera-captured Image Data Box)603は、カメラ毎に設定可能で、各カメラの撮影画像のフォーマット情報が入る。図7(a)はcimgボックス603の一例であり、同(b)はそのシンタックスである。

10

【0036】

campボックス(Camera Parameter Data Box)604は、カメラ毎に設定可能で、カメラパラメータとしてどのようなデータが格納されているかを示す情報とそれに対応するカメラパラメータの各値が入る。図8(a)はcampボックス604の一例であり、同(b)はそのシンタックスである。

【0037】

imgボックス(Foreground Image Data Box)605は、前景画像としてどのような画像データが格納されているかを示す情報と各前景画像のフォーマット情報が入る。図9(a)はimgボックス605の一例であり、同(b)はそのシンタックスである。「前景フォーマット」901は、RAW、LOG、RGB、YUVといった前景画像の保存形式とビット深度を示す。「前景シルエットフォーマット」902は、前景シルエット画像のビット深度を示し、例えば、1ビット前景なのか背景なのかを示したり、8ビットで前景シルエットとしての確からしさを表現することができる。「切り出し前景フォーマット」903も「前景フォーマット」901と同様の情報を示す。

20

【0038】

bimgボックス(Background Image Data Box)606は、背景画像としてどのような画像データが格納されているかを示す情報と各背景画像のフォーマット情報が入る。図10(a)はbimgボックス606の一例であり、同(b)はそのシンタックスである。「背景フォーマット」1001は、上述の「前景フォーマット」901と同様、背景画像の保存形式とビット深度の情報を示す。「背景テクスチャフォーマット」1002は、その保存形式やビット深度の情報に加え、対応する背景3Dモデルの頂点座標の数を示す。「切り出し背景フォーマット」1003は「背景フォーマット」1001と同様の情報を示す。

30

【0039】

fmdlボックス(Foreground Model Data Box)607は、前景3Dモデルとしてどのようなデータが格納されているかを示す情報と各前景3Dモデルのフォーマット情報が入る。図11(a)はfmdlボックス607の一例であり、同(b)はそのシンタックスである。「点群フォーマット」1101は、各点は何ミリのボクセルの座標を表現するかを示す。「メッシュモデルフォーマット」1102は、メッシュの頂点数などを示す。「ビルボードフォーマット」1103は、ビルボードを立てる座標の単位を示す。「バウンディングボックスフォーマット」1104は、バウンディングボックスが代表頂点2点で示されるのか、代表頂点1点と、幅、奥行き、高さで示されるのか、といった表現形式を示す。

40

【0040】

bmdlボックス(Background Model Data Box)608は、背景3Dモデルとしてどのようなデータが格納されているかを示す情報と各背景3Dモデルのフォーマット情報が入る。図12(a)はbmdlボックス608の一例であり、同(b)はそのシンタックスである。「点群フォーマット」1201や「メッシュモデルフォーマット」1202

50

は、上述の前景 3 D モデルに関する、「点群フォーマット」1 1 0 1 や「メッシュモデルフォーマット」1 1 0 2 と同様の情報が示される。

【 0 0 4 1 】

本実施形態では、I S O B M F F 規格を例に説明したが、V v v m ファイルのフォーマットはこれに限定されない。例えば、I S O B M F F と互換性のある H E I F (High Efficiency Image File Format) や M i A F (Multi-Image Application Format) など他の規格であってもよい。或いは独自のフォーマットや今後新たに登場するフォーマットであっても構わない。また、それぞれのパラメータの値や表現も上述の例に限定されない。また、図 6 に示す c i m g 6 0 3、c a m p 6 0 4、f i m g 6 0 5、b i m g 6 0 6、f m d 1 6 0 7、b m d 1 6 0 8 の各ボックスのうち、少なくとも 1 つが m o o v ボックス 6 0 2 に格納されるようにしてもよい。

10

【 0 0 4 2 】

(仮想視点映像生成装置)

続いて、図 3 (a) で示したデータ構造を持つ V v v m ファイル 3 0 0 を用いて、仮想視点から見た映像を生成する仮想視点映像生成装置について説明する。図 1 3 (a) は、本実施形態に係る、仮想視点映像生成装置 2 0 の内部構成を示す機能ブロック図である。仮想視点映像生成装置 2 0 は、ファイル入力部 2 1、生成可否判定部 2 2、前景背景分離部 2 3、仮想視点パス入力部 2 4、映像化部 2 5、映像出力部 2 6 で構成される。なお、仮想視点映像生成装置 2 0 は、C P U、R A M、R O M、H D D といった、演算やデータ記憶のためのハードウェアを備えた情報処理装置であり、例えば P C などである。図 1 3 (a) に示す各部の機能は、C P U が、所定のプログラムを R O M 又は H D D から読み込んで R A M に展開し、これを C P U が実行することで実現される。以下、図 1 3 (a) に示す各部の役割について説明する。

20

【 0 0 4 3 】

ファイル入力部 2 1 は、ファイル生成装置 1 0 で生成された V v v m ファイル 3 0 0 の入力を受け付け、各処理に必要なデータを読み出して、各処理に対応した機能ブロックに渡す。例えば、格納データ情報 3 0 2 を読み出して、生成可否判定部 2 2 に送る。また、各カメラのカメラパラメータ 3 0 3 を読み出して、前景背景分離部 2 3 と映像化部 2 5 に送る。また、背景 3 D モデル 3 0 4 を読み出して、映像化部 2 5 に送る。また、仮想視点入力部 2 4 から受け取った仮想視点パスの情報に基づき、設定された仮想視点に対応する時刻のフレームデータ 3 0 5 を読み出し、前景・背景分離部 2 3 に送る。

30

【 0 0 4 4 】

生成可否判定部 2 2 は、ファイル入力部 2 1 から受け取った格納データ情報 3 0 2 を参照して、V v v m ファイル 3 0 0 内の素材データによって仮想視点映像を生成可能であるかを判定する。すなわち、格納データ情報で特定される素材データが、自装置が持つ機能に応じて要求される素材データを満たす場合、仮想視点映像が生成可能であると判定する。仮想視点映像を生成する手法にはいくつかあるが、仮想視点映像生成装置 2 0 は、例えば、前景画像、背景画像、前景 3 D モデル、背景 3 D モデルの 4 つの素材データを用いることによって、モデルベースレンダリングに基づく仮想視点映像を生成できる。ここで、本実施形態の V v v m ファイル 3 0 0 に格納される素材データは、各カメラで撮影された画像 (複数視点映像)、カメラパラメータ、背景 3 D モデルの 3 種類であるとする。この例において、仮想視点映像生成装置 2 0 は、各カメラの撮影画像から前景と背景を分離する機能、さらに、前景 3 D モデルを生成する機能を持っていれば、仮想視点映像を生成できる。生成可否判定部 2 2 は、自装置が採用する仮想視点映像の生成手法を前提として、自装置が有する仮想視点映像生成のための諸機能を踏まえた対応可否の判定基準 (必要な素材データを示した判定用テーブル等) を予め作成し保持しておく。そして、当該判定基準に示された素材データが V v v m ファイル 3 0 0 に含まれるか否かを、格納データ情報 3 0 2 を用いて確認することで、仮想視点映像が生成可能かどうかを判定する。以下に、対応可否の判定基準の一例として、バリエーション A ~ C を示す。

40

	前景／背景分離機能	前景 3 Dモデル生成機能	必要な素材データ
判定基準 A	あり	あり	・複数視点映像 ・カメラパラメータ ・背景 3 Dモデル
判定基準 B	なし	あり	・前景画像 ・背景画像 ・カメラパラメータ ・背景 3 Dモデル
判定基準 C	なし	なし	・前景画像 ・背景画像 ・前景 3 Dモデル ・背景 3 Dモデル

10

【 0 0 4 5 】

格納データ情報 3 0 2 を用い上記判定基準に照らして得られた判定結果は、映像出力部 2 6 を介して、仮想視点映像生成装置 2 0 が備える不図示のモニタ等に表示するなどして、ユーザに通知される。

【 0 0 4 6 】

20

前景背景分離部 2 3 は、各カメラの撮影画像を前景と背景とに分離して前景シルエット画像を生成する。例えば、まず、カメラ毎に複数フレーム分の撮影画像を用いて背景画像を生成する。具体的には、同一カメラで撮影された複数フレーム間の画像比較により、動きのある領域と動きの無い領域とを検出し、動きのない領域を背景画像とする。背景画像を得ると、次に、各カメラの撮影画像と該生成した背景画像との間で対応する画素同士を比較し、画素値の差が所定の閾値以上である画素を前景画素とする。これにより、各カメラの撮影画像と同じ画像サイズであって、画素値が“ 1 ”を前景画素、画素値が“ 0 ”を背景画素とした前景シルエット画像が得られる。こうして得られた、背景画像と前景シルエット画像は、映像化部 2 5 に送られる。なお、撮影画像を前景と背景とに分離する手法は、上述の例に限定されない。例えば、撮影画像から前景 3 Dモデルを生成する処理の中で、機械学習技術等を用いて前景部分を分離することで、前景シルエット画像を生成してもよい。

30

【 0 0 4 7 】

仮想視点設定部 2 4 は、不図示の UI 等を介したユーザ指定に基づいて、或いは予め決められた仮想視点の設定情報を読み込んで、仮想視点を設定する。動画による仮想視点映像を生成する場合に設定される仮想視点は、その生成対象時間（例えば 1 0 秒間、複数視点映像が 6 0 f p s で撮影されていれば 6 0 0 フレーム分）の間で移動するように設定されることが多い。そのようなユーザ指示等に基づき設定された仮想視点の移動経路（以下、「仮想視点パス」と呼ぶ。）の情報は、ファイル入力部 2 1 と映像化部 2 5 に送られる。仮想視点として設定されるパラメータは、例えば仮想カメラの位置、姿勢、画角等であり、さらに焦点位置や絞り値等を含んでもよい。

40

【 0 0 4 8 】

映像化部 2 5 は、仮想視点設定部 2 4 から受け取った仮想視点パスの情報に基づいて仮想視点映像を生成する。映像化部 2 5 の詳細については後述する。映像化部 2 5 で生成した仮想視点映像のデータは、映像出力部 2 6 に送られる。

【 0 0 4 9 】

映像出力部 2 6 は、映像化部 2 5 から受け取った仮想視点映像のデータを、不図示のモニタに表示したり、或いは外部装置に送信するなどして出力する。

【 0 0 5 0 】

（映像化部の詳細）

50

図 1 3 (b) は、本実施形態に係る、映像化部 2 5 の詳細を示す機能ブロック図である。本実施形態の映像化部 2 5 は、前景 3 D モデル生成部 2 5 1、前景色づけ部 2 5 2、背景色づけ部 2 5 3、合成部 2 5 4 で構成される。以下、各部について説明する。

【 0 0 5 1 】

前景 3 D モデル生成部 2 5 1 は、前述の前景背景分離部 2 3 で生成された前景シルエット画像とファイル入力部 2 1 から受け取ったカメラパラメータとに基づき、前景 3 D モデルを生成する。例えば、視体積交差法による前景 3 D モデルの生成手順は、以下の 1) ~ 3) のとおりである。

【 0 0 5 2 】

1) 予め設定された前景 3 D モデルの生成対象空間に一定の大きさをもった直方体 (ボクセル) を敷き詰める。生成対象空間は、仮想視点パスに基づき、仮想カメラから見える範囲を算出して決定する。

【 0 0 5 3 】

2) 各カメラのカメラパラメータを用いて、各直方体を三次元的にカメラの前景シルエット画像上に射影し、前景画素と重なるか確認する。前景画素と重ならない場合、該直方体は、前景 3 D モデルを形成する直方体でないと判定する。全カメラのうち 1 台でも前景 3 D モデルを形成する直方体でないと判定された場合、該直方体を削除する。

【 0 0 5 4 】

3) 上記 2) の処理をすべての直方体に対して行い、残った直方体の重心座標の点の集合を前景 3 D モデルとする。

【 0 0 5 5 】

なお、ここでは視体積交差法を用いた前景 3 D モデルの生成方法を説明したが、3 D モデルの生成方法はこれに限定されない。そして、上述のようにして生成した前景 3 D モデルは、前景色づけ部 2 5 2 に送られる。

【 0 0 5 6 】

前景色づけ部 2 5 2 は、前景 3 D モデル生成部 2 5 1 で生成された前景 3 D モデルに対し、仮想視点設定部 2 4 から受け取った仮想視点パスに基づき色づけ処理を行って、前景色付き画像を生成する。色づけ処理の具体的な手順は、以下の 1) ~ 3) のとおりである。

【 0 0 5 7 】

1) 前景 3 D モデルの表面上の各点がどのカメラから見えるかの判定 (可視性判定) を行う。この可視性判定では、まず、各点を全カメラの前景シルエット画像に射影する。この際、複数の点が、前景シルエット画像上の同じ画素に射影された場合は、カメラからみて最も距離が短い位置にある点が見える点、その他の点は見えない点と判定する。この処理を前景 3 D モデルの表面上の全点に対して行って、各点に色づけするためのカメラを特定する。

【 0 0 5 8 】

2) 仮想視点パスに基づいて、前景 3 D モデルを仮想視点映像となる画像に射影し、各画素に写る点を特定する。

【 0 0 5 9 】

3) 画素毎に、上記 1) の可視性判定で射影された点が見えると判定されたカメラの撮影画像を用いて、当該点が射影された画素の色を決定する。この際、複数のカメラから見える点の場合は、仮想視点の向きと各カメラの向きを比較し、より角度の小さい 2 つのカメラの撮影画像を適宜重み付けるなどして、色を決定すればよい。

【 0 0 6 0 】

上述のようにして生成した前景色付き画像のデータは、合成部 2 5 4 に送られる。

【 0 0 6 1 】

背景色づけ部 2 5 3 は、各カメラの撮影画像から抽出された背景画像を用いて、背景 3 D モデルに貼り付ける背景テクスチャを生成する。具体的には、以下の 1) ~ 5) の手順で、背景テクスチャが生成される。

【 0 0 6 2 】

10

20

30

40

50

1) 背景 3 D モデルにおいて代表点となる頂点を設定する。

【 0 0 6 3 】

2) 上記 1) で設定された頂点を、仮想視点に近い 2 台のカメラ (カメラ 1、カメラ 2 とする) のカメラパラメータを用いて、仮想視点映像となる画像上に射影する。

【 0 0 6 4 】

3) 仮想視点とカメラ 1 の対応点、仮想視点とカメラ 2 の対応点を用いて、仮想視点とカメラ 1 間の射影行列 1、及び、仮想視点とカメラ 2 間の射影行列 2 を算出する。

【 0 0 6 5 】

4) 仮想視点映像となる画像における各画素の座標を、射影行列 1 と射影行列 2 を用いて、カメラ 1 の背景画像とカメラ 2 の背景画像にそれぞれ射影し、当該座標が射影された位置の画素値をそれぞれ取得する。

10

【 0 0 6 6 】

5) 得られた 2 つの画素値の平均値を、仮想視点映像となる画像における背景領域の画素値とする。

【 0 0 6 7 】

さらに、背景色づけ部 2 5 3 は、生成した背景テクスチャを、V v v m ファイル内に格納されていた背景 3 D モデルに対し貼り付け、仮想視点から見たときの背景画像 (仮想背景画像) を生成する。上述のようにして生成した仮想背景画像のデータは、合成部 2 5 4 に送られる。

【 0 0 6 8 】

20

合成部 2 5 4 は、前景色づけ部 2 5 2 から受け取った前景色付き画像と、背景色づけ部 2 5 3 から受け取った仮想背景画像とを合成し、仮想視点からの見えを表す画像を生成する。具体的には、前景色付き画像の各画素に対し、仮想背景画像における対応する画素の画素値を上書きする処理を行なう。

【 0 0 6 9 】

図 1 4 は、本実施形態に係る V v v m ファイル 3 0 0 から、仮想視点映像生成装置 2 0 が仮想視点映像を生成し出力するまでの処理の流れを示すフローチャートである。なお、各処理の冒頭における記号「S」はステップを意味する。以下、図 1 4 のフローに沿って説明する。

【 0 0 7 0 】

30

S 1 4 0 1 では、ファイル入力部 2 1 が、V v v m ファイル 3 0 0 から格納データ情報 3 0 2 を読み出し、生成可否判定部 2 2 に出力する。続く S 1 4 0 2 では、生成可否判定部 2 2 が、予め用意された生成可否判定基準の情報を読み出し、格納データ情報 3 0 2 を参照して、V v v m ファイル 3 0 0 内に格納された素材データを用いて仮想視点映像を生成可能であるかを判定する。判定の結果、仮想視点映像を生成可能な場合は、S 1 4 0 3 に進む。この際、映像出力部 2 6 が、仮想視点映像を生成可能であることを示すメッセージを不図示のモニタに表示するなどの通知を行ってもよい。一方、仮想視点映像を生成不可能な場合は、S 1 4 0 9 に進む。

【 0 0 7 1 】

S 1 4 0 3 では、ファイル入力部 2 1 が、V v v m ファイル 3 0 0 からカメラパラメータ 3 0 3 と背景 3 D モデル 3 0 4 を読み出し、カメラパラメータ 3 0 3 を前景背景分離部 2 3 に、背景 3 D モデル 3 0 4 を映像化部 2 5 にそれぞれ出力する。続く S 1 4 0 4 では、仮想視点設定部 2 4 は、ユーザ指定等に基づき仮想視点パスを設定し、その情報をファイル入力部 2 1 と映像化部 2 5 に出力する。

40

【 0 0 7 2 】

S 1 4 0 5 では、ファイル入力部 2 1 が、入力された仮想視点パスに対応するフレームデータ 3 0 5 を V v v m ファイル 3 0 0 から読み出し、前景背景分離部 2 3 に出力する。以降の S 1 4 0 6 及び S 1 4 0 7 は、フレーム単位で実行される。

【 0 0 7 3 】

S 1 4 0 6 では、前景背景分離部 2 3 が、各カメラの撮影画像に対して前景背景分離処

50

理を行って、前景シルエット画像と背景画像を生成し、映像化部 25 に出力する。続く S 1407 では、映像化部 25 が、入力された仮想視点パス、前景シルエット画像及び背景画像に基づき仮想視点映像を生成する。

【0074】

S 1408 では、映像出力部 26 が、S 1408 で生成された仮想視点映像のデータを出力する。また、S 1409 では、映像出力部 26 が、仮想視点映像を出力できない旨のメッセージを不図示のモニタに表示するなどの通知を行う。

【0075】

以上が、V v v m ファイル 300 から仮想視点映像を生成し出力するまでの処理の流れである。

【0076】

< V v v m ファイルのバリエーション >

共通のフォーマットで素材データをファイル化した V v v m ファイルのデータ構造は、前述の図 3 (a) に示す例に限定されない。以下に、V v v m ファイルのデータ構造のバリエーションを示す。

【0077】

図 15 に示す V v v m ファイル 1500 は、データバージョン 1501、格納データ情報 1502、カメラパラメータ 1503、背景 3D モデル 1504、フレームデータ 1505 から構成される。そして、フレームデータ 1505 は、第 1 フレームから第 N フレームまで存在し、1 つのフレームデータには、前景画像群、前景 3D モデル群、背景画像群が含まれている。そして、前景画像群には、M 台のカメラに対応した前景画像と前景シルエット画像が含まれる。また、前景 3D モデル群には、O 個の前景 3D モデルが含まれる。また、背景画像群には、P 個の背景画像が含まれる。この場合において、N、M、O 及び P は、いずれも 1 以上の自然数である。V v v m ファイル 1500 にどのような素材データが格納されているかは、格納データ情報 1502 によって示される。そして、図 16 (a) は前景画像に関するメタデータの一例を示し、同 (b) はそのビットアサインを示している。ここで、前景切り出し画像は、カメラの撮影画像から前景領域を切り出した画像を意味し、前景画像をメインプロファイルとするときの拡張プロファイルに相当する。V v v m ファイルに格納されているのが前景切り出し画像のみで前景画像自体は格納されていない場合、仮想視点映像生成装置側が前景切り出し画像に対応していないと仮想視点映像を生成できないということになる。例えば、あるバージョンのメインプロファイルには対応可能であるが、拡張プロファイルには対応できないという装置の場合、前景画像に関するメタデータを確認することで、自装置では仮想視点映像の生成ができないと判断することができる。このように、素材データに付与されるメタデータによって、仮想視点映像の生成可否をさらに詳細に判断することが可能になる。また、図 16 (c) は背景画像に関するメタデータの一例を示し、同 (d) は格納背景画像のビットアサインを示している。そして、図 16 (e) は前景 3D モデルのメタデータの一例を示し、同 (f) は格納前景 3D モデルのビットアサインを示している。前景 3D モデルや背景 3D モデルには、点の集合で構成される点群モデルや、多角形の面を集めて構成されるメッシュモデルなど、いくつか種類がある。仮想視点映像を生成する装置は、V v v m ファイルに格納されている前景 / 背景の 3D モデルの種類が、自装置で対応可能な種類であるかを、これらメタデータを参照して判断することができる。図 15 に示すデータ構造を持つ V v v m ファイル 1500 を生成し、それを用いて仮想視点映像を生成する態様については、次の実施形態 2 おいて説明することとする。

【0078】

図 17 に示す V v v m ファイル 1700 は、データバージョン 1701、格納データ情報 1702、背景 3D モデル 1703、フレームデータ 1704 から構成される。そして、フレームデータ 1704 は、第 1 フレームから第 N フレームまで存在し、1 つのフレームデータには、前景画像群、前景 3D モデル群、可視性情報、背景テクスチャ群が含まれている。ここで、可視性情報は、前述の可視性判定の結果に相当する情報である。前景画

10

20

30

40

50

像群には、M台のカメラに対応した前景画像が含まれる。また、前景3Dモデル群には、O個の前景3Dモデルが含まれる。また、可視性情報には、Q個の可視性情報が含まれる。また、背景テクスチャ群には、背景3Dモデルに貼り付けるR個の背景テクスチャが含まれる。ここでのQ及びRも1以上の自然数である。また、R個の背景テクスチャは背景画像を合成して生成するため、R=Pとなる。そして、Vvvvmファイル1700にどのような素材データが格納されているかは、格納データ情報1702によって示される。メタデータとビットアサインについては省略する。図17に示すVvvvmファイル1700の場合、前景3Dモデルに加えて、可視性情報と背景テクスチャのデータも含まれている。したがって、前景色づけ部252における可視性判定や、背景色づけ部253における背景テクスチャの生成を省くことができ、仮想視点映像装置側の負荷を軽くすることができる。

10

【0079】

図18に示すVvvvmファイル1800は、データバージョン1801、格納データ情報1802、背景3Dモデル1803、フレームデータ1804から構成される。そして、フレームデータ1804は、第1フレームから第Nフレームまで存在し、1つのフレームデータには、テクスチャ付き前景3Dモデル群と背景テクスチャ群とが含まれている。ここで、テクスチャ付き前景3Dモデルは、前述の前景色づけ処理が施された前景3Dモデルを意味する。テクスチャ付き前景3Dモデル群には、O個のテクスチャ付き前景3Dモデルが含まれる。また、背景テクスチャ群には、R個の背景テクスチャが含まれる。そして、Vvvvmファイル1800にどのような素材データが格納されているかは、格納データ情報1802によって示される。メタデータとビットアサインについては省略する。図18に示すVvvvmファイル1800の場合、既に色づけ処理が済んだ前景3Dモデルが含まれていることから、図17のVvvvmファイル1700よりも仮想視点映像装置側の負荷をさらに軽くすることができる。

20

【0080】

なお、図3(a)、図15、図17におけるフレームデータを構成する画像データ(各カメラの撮影画像、前景画像、背景画像)は、カメラに付された通し番号順に格納されるのが通常であるが、これに限定されない。例えば各素材データのヘッダにカメラ番号を付与した上で、順不同で格納してもよい。また、本実施形態では、各カメラの撮影画像や前景画像などのデータを、フレーム単位で読み出し可能なように格納しているが、カメラ単位で読み出し可能なように格納してもよい。

30

【0081】

また、図3(a)、図15、図17、図18で例示したデータ構造においては、どのVvvvmファイルにも背景3Dモデルが含まれているが必須ではない。例えば、スタジアムでのサッカーの試合といったように、撮影シーンが同一施設で行われる同種イベントである場合は、背景3Dモデルは、予め用意したものを共通で繰り返し使用可能である。受信したVvvvmファイルに背景3Dモデルが含まれていなくても、仮想視点映像生成装置がその撮影シーンに対応する背景3Dモデルのデータを他の情報処理装置から取得するなどして、仮想視点映像を生成してもよい。この際、対応可否の判定基準がそれに応じたものになることは言うまでもない。

40

【0082】

以上のとおり本実施形態によれば、仮想視点映像を生成するために必要な素材データが共通のフォーマットでまとめてファイル化される。これにより、共通フォーマットのVvvvmファイルに対応した端末やアプリケーションであれば容易に仮想視点映像を生成できるようになり、仮想視点映像の利用環境が広がる。

【実施形態2】

【0083】

次に、前述の図15に示すデータ構造を持つVvvvmファイル1500を生成し、それを用いて仮想視点映像を生成する態様を、実施形態2として説明する。なお、実施形態1と共通する内容については説明を省略ないしは簡略化することとし、以下では差異点を中

50

心に説明を行うものとする。

【 0 0 8 4 】

前述のとおり、図 1 5 に示す V v v m ファイル 1 5 0 0 は、データバージョン 1 5 0 1、格納データ情報 1 5 0 2、カメラパラメータ 1 5 0 3、背景 3 D モデル 1 5 0 4、フレームデータ 1 5 0 5 から構成される。このため、本実施形態に係るファイル生成装置は、前景背景分離を実施し、前景 3 D モデルの生成を行って、前景画像、前景シルエット画像、前景 3 D モデル、背景画像の各データをフレーム単位で格納する。そして、V v v m ファイル 1 5 0 0 には、これらの素材データが格納されていることを示す格納データ情報 1 5 0 2 が付加される。また、本実施形態に係る仮想視点映像生成装置は、格納データ情報 1 5 0 2 を参照して生成可否の判定を行って仮想視点映像の生成が可能であれば、当該ファイル内の前景画像、背景画像、前景 3 D モデルの各素材データを用いて仮想視点映像を生成する。

10

【 0 0 8 5 】

(ファイル生成装置)

図 1 9 は、実施形態 1 の図 1 に対応する図であり、本実施形態に係るファイル生成装置 1 0 ' の機能ブロック図である。ファイル生成装置 1 0 ' は、データ入力部 1 1 '、カメラパラメータ生成部 1 2 '、ファイル生成部 1 3 '、ファイル出力部 1 4 に加え、前景背景分離部 1 9 0 1 と前景 3 D モデル生成部 1 9 0 2 を有する。以下、図 1 9 に示す各部の役割について説明する。

【 0 0 8 6 】

20

データ入力部 1 1 ' は、その機能は、実施形態 1 のデータ入力部 1 1 と同じであるが、入力されたデータの出力先が異なる。すなわち、カメラ情報、複数視点映像、背景 3 D モデルの入力を受け付けると、カメラ情報はカメラパラメータ生成部 1 2 ' に、背景 3 D モデルはファイル生成部 1 3 ' に、複数視点映像のデータは前景背景分離部 1 9 0 1 にそれぞれ送られる。

【 0 0 8 7 】

カメラパラメータ生成部 1 2 ' も、その機能は実施形態 1 のカメラパラメータ生成部 1 2 と同じであるが、生成したカメラパラメータの出力先が異なる。本実施形態の場合、生成されたカメラパラメータは、前景背景分離部 1 9 0 1、前景 3 D モデル生成部 1 9 0 2 及びファイル生成部 1 3 ' に送られる。

30

【 0 0 8 8 】

前景背景分離部 1 9 0 1 は、実施形態 1 の仮想視点映像生成装置 2 0 における前景背景分離部 2 3 と同様、各カメラの撮影画像を前景と背景とに分離する処理を行い、前景画像、前景シルエット画像、背景画像を生成する。そして、得られた前景シルエット画像を前景 3 D モデル生成部 1 9 0 2 に、前景画像と背景画像をファイル生成部 1 3 ' にそれぞれ送る。

【 0 0 8 9 】

前景 3 D モデル生成部 1 9 0 2 は、実施形態 1 の仮想視点映像生成装置 2 0 の映像化部 2 5 における前景 3 D モデル生成部 2 5 1 と同様、前景シルエット画像と仮想視点パスとに基づき前景 3 D モデルを生成する。そして、得られた前景 3 D モデルを、ファイル生成部 1 3 ' に送る。

40

【 0 0 9 0 】

ファイル生成部 1 3 ' は、データ入力部 1 1 ' から背景 3 D モデル、前景背景分離部 1 9 0 1 から前景画像、前景シルエット画像及び背景画像、前景 3 D モデル生成部 1 9 0 2 から前景 3 D モデル、カメラパラメータ生成部 1 2 ' からカメラパラメータを取得する。そして、取得したこれらデータを共通フォーマットでファイル化する。この際、前景画像、前景シルエット画像、前景 3 D モデル及び背景画像は、フレーム単位で格納される。さらに、前景画像と前景シルエット画像はカメラ単位でも識別可能となるように、前景 3 D モデルはモデル単位でも識別可能となるように、それぞれ格納される。さらに、素材データとしてこれらが格納されていることを示す格納データ情報 1 5 0 2 を付加して、図 1 5 に示

50

すV v v mファイル1 5 0 0を生成する。

【0 0 9 1】

(仮想視点映像生成装置)

図2 0 (a) 及び (b) は、実施形態1の図1 3 (a) 及び (b) に対応する図であり、本実施形態に係る仮想視点映像生成装置2 0 ' と映像化部2 5 ' の内部構成をそれぞれ示している。図2 0 (a) には前景背景分離部2 3 に対応する処理ブロックが存在せず、図2 0 (b) には前景3 Dモデル生成部2 5 1 に対応する処理ブロックが存在しない。以下、図2 0 (a) 及び (b) に示す各部の機能について説明する。

【0 0 9 2】

ファイル入力部2 1 ' は、ファイル生成装置1 0 ' で生成されたV v v mファイルの入力を受け付け、各処理に必要なデータを読み出して、各処理に対応したブロックに渡す。図1 5 に示すV v v mファイル1 5 0 0 が入力された場合は、格納データ情報1 5 0 2 を読み出して、それを生成可否判定部2 2 ' に送る。また、カメラパラメータ1 5 0 3 と背景3 Dモデル1 5 0 4 を読み出して、映像化部2 5 ' に送る。また、仮想視点設定部2 4 から受け取った仮想視点パスの情報に基づき、設定された仮想視点パスに対応するフレームデータ1 5 0 5 を読み出し、映像化部2 5 ' に送る。

【0 0 9 3】

生成可否判定部2 2 ' は、ファイル入力部2 1 ' から受け取った格納データ情報を用いて、入力されたV v v mファイルに基づき仮想視点映像を生成可能であるかを判定する。前述の通り、本実施形態の仮想視点映像装置2 0 ' は、前景背景分離機能と前景3 Dモデル生成機能とを有しない。よって、入力されたV v v mファイルが、図1 5 (a) に示すV v v mファイル1 5 0 0 であれば生成可能と判断されるが、例えば図3 (a) に示すV v v mファイル3 0 0 であれば生成不可能と判断されることになる。

【0 0 9 4】

映像化部2 5 ' は、仮想視点設定部2 4 から受け取った仮想視点パスの情報に基づいて仮想視点映像を生成する。映像化部2 5 ' の詳細については後述する。そして、映像出力部2 6 が、映像化部2 5 ' にて生成された仮想視点映像のデータを、不図示のモニタに表示したり、或いは外部端末等に送信するなどして出力する。

【0 0 9 5】

(映像化部の詳細)

図2 0 (b) に示す通り、本実施形態の映像化部2 5 ' は、前景色づけ部2 5 2 ' 、背景色づけ部2 5 3、合成部2 5 4 とで構成される。背景色づけ部2 5 3 と合成部2 5 4 は、実施形態1と異なるところはない。図2 0 (c) は、本実施形態に係る前景色づけ部2 5 2 ' の内部構成を示すブロック図である。前景色づけ部2 5 2 ' は、使用モデル決定部2 0 0 0 と、複数種類の前景3 Dモデルに対応した色づけ部、具体的には、メッシュモデル色づけ部2 0 0 1、ビルボード色づけ部2 0 0 2、点群モデル色づけ部2 0 0 3 とで構成される。使用モデル決定部2 0 0 0 によって決定された色づけ部において、前景オブジェクトの3 Dモデルに色づけがなされ、仮想視点から見た前景色つき画像が生成される。例えば、入力された前景3 Dモデルのタイプが、ビルボード形式であった場合は、ビルボード色づけ部2 0 0 2 が、当該前景3 Dモデルに適用する色づけ部として決定される。ビルボードの場合はそれを立てる座標位置のみが示される (オブジェクトの三次元形状は持たない) ので、指定された座標位置に描画対象のオブジェクトがあるときの仮想視点から見える方向と大きさを計算し、仮想視点から見た前景を生成する。

【0 0 9 6】

モデル決定部2 0 0 0 は、前景3 Dモデルに付与されたメタデータに基づいて、フレームデータ1 5 0 5 に含まれる前景3 Dモデルの種類を特定し、3つの色づけ部2 0 0 1 ~ 2 0 0 3 のうちどれを使用して色づけ処理を行うかを決定する。この際、フレームデータ1 5 0 5 内に複数種類の前景3 Dモデルが格納されていれば、予め定めた優先順 (例えば、点群モデル色づけ部2 0 0 3、メッシュモデル色づけ部2 0 0 1、ビルボード色づけ部2 0 0 2 の順) に使用する色づけ部を決定する。また、仮想視点の設定時にユーザが仮想

10

20

30

40

50

視点映像生成モードを指定した場合には、当該指定に係る生成モードの内容に応じて決定するようにしてもよい。例えば、高画質モードが指定された場合は点群モデル色づけ部 2003、通常画質モードが指定された場合はメッシュモデル色づけ部 2001、高速モードが指定された場合はビルボード色づけ部 2002、といった具合である。このように、予め定めた基準によって、前景 3D モデルに対する色づけ処理に用いる処理部を決定することにより、目的や用途に応じた適切な前景を生成することができる。

【0097】

ここで、例えば、前景色づけ部 252' が点群モデル色づけ部 2003 を有しておらず、仮想視点映像生成装置 20' に入力された V v v m ファイル 1500 のフレームデータ 1505 に格納された前景 3D モデルが点群モデルのみであったと仮定する。この場合、生成可否判定部 22' は、前景 3D モデルに関するメタデータを参照して、フレームデータ 1505 に含まれる前景 3D モデルが点群モデルだけであることを特定する。そして、自装置が点群モデルに対応した色づけ部を持たないことにより、仮想視点映像を生成することはできないと判定する。このように、格納データ情報 1502 だけでなく、それぞれの素材データに付与されたメタデータの内容を確認することにより、さらに詳細に仮想視点映像生成可否を判断することが可能となる。

【0098】

<変形例>

なお、映像化部 25' は、設定された仮想視点パスの下で実際に用いる素材データだけを V v v m ファイル 1500 から取得することで、より効率的に仮想視点映像を生成するようにしてもよい。すなわち、ファイル入力部 21' が、仮想視点設定部 24 から入力された仮想視点パスの情報に基づいて、仮想視点映像の生成に実際に使用する素材データのみを特定して、映像化部 25' に渡すようにする。この特定の際には、対象となるフレームデータ 1505 に含まれる素材の種類（前景画像、前景 3D モデル、背景画像）を指定した上で、さらにカメラ ID やモデル ID といったさらに細かな指定が含まれるようにする。以下に、設定された仮想視点に対応する素材データをどのように特定するのかについて説明する。

【0099】

まず、前景画像と背景画像については、設定された仮想視点位置と近い視点位置のカメラで撮影された前景画像或いは背景画像を、V v v m ファイル内のカメラパラメータに基づいて特定する。この際、選択する前景画像が少ないと、障害物に隠れるオクルージョンにより、色づけができない場合もある。また、選択する背景画像が少ないと、仮想視点に映る背景がすべて色づけできない可能性がある。そのため、複数台（例えば、視点位置が近いカメラから順に 6 台分）のカメラの撮影画像における前景画像や背景画像を選択するようにする。この際、台数を多くしたり、高さの異なるカメラを混ぜたりすることで、オクルージョンの影響を受けづらくしたり、背景が色づけできない可能性を低減したりすることができる。

【0100】

次に、前景 3D モデルについては、V v v m ファイルに含まれる前景 3D モデルのバウンディングボックス情報に基づいて特定する。バウンディングボックス情報は、前景 3D モデルを直方体で囲ったときの、当該直方体の 3 次元座標を示す情報である。このバウンディングボックス情報を用いて、各前景 3D モデルが、設定された仮想視点から見えるか否かを判定する。具体的には、バウンディングボックスを表す直方体の各頂点を、仮想視点映像となる 2 次元画像に向けて射影したときに、各頂点に対応する点が当該画像上に存在するか否かを確認する。こうして、設定された仮想視点から見える前景 3D モデルが特定される。

【0101】

なお、設定された仮想視点パスの下で実際に使用する素材データだけを取得して効率的に仮想視点映像を生成する上述の内容は、図 15 に示すデータ構造の V v v m ファイルに限定されるものではない。他のあらゆるデータ構造の V v v m ファイルに対しても適用可

10

20

30

40

50

能である。

【 0 1 0 2 】

以上のとおり本実施形態の場合、ファイル生成装置が、素材データとして、前景画像、背景画像及び前景 3 D モデルが含まれる V v v m ファイルを生成する。したがって、前景背景分離機能や前景 3 D モデル生成機能を有していない仮想視点映像生成装置においても、V v v m ファイルに対応していれば容易に仮想視点映像を生成することができる。

【 0 1 0 3 】

(その他の実施例)

本発明は、上述の実施形態の 1 以上の機能を実現するプログラムを、ネットワーク又は記憶媒体を介してシステム又は装置に供給し、そのシステム又は装置のコンピュータにおける 1 つ以上のプロセッサがプログラムを読み出し実行する処理でも実現可能である。また、1 以上の機能を実現する回路(例えば、A S I C)によっても実現可能である。

【符号の説明】

【 0 1 0 4 】

1 0 ファイル生成装置

1 3 ファイル化部

1 4 ファイル出力部

2 0 仮想視点映像生成装置

2 2 生成可否判定部

2 5 映像化部

2 6 映像出力部

10

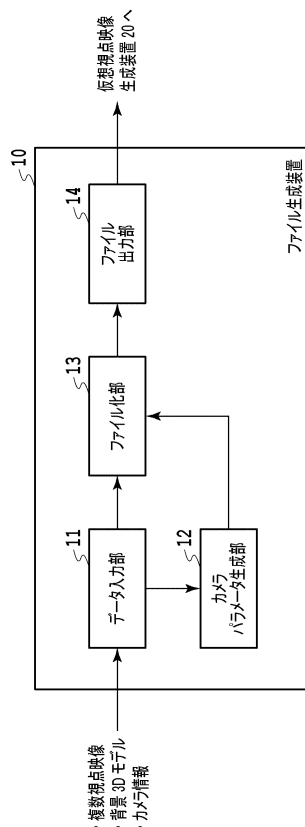
20

30

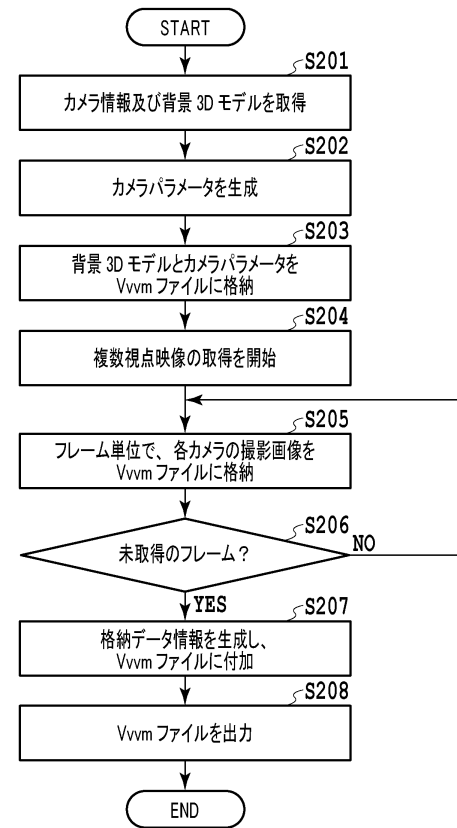
40

50

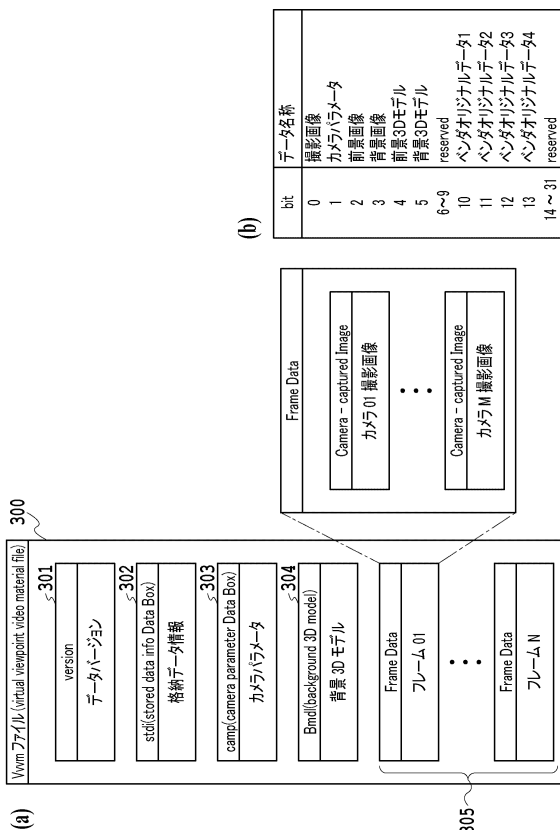
【図面】
【図 1】



【図 2】



【図 3】



【図 4】

タグ名称	Field Name	タイプ	カウント
CIタグのバージョン	CI Version ID	BYTE	4
カメラ ID	Camera ID	LONG	1
フォーマット	format	LONG	1
画像の幅	Image width	LONG	1
画像の高さ	Image height	LONG	1
ビット深度	Bit Depth	LONG	1

(a)

タグ名称	Field Name	タイプ	カウント
CPタグのバージョン	CP Version ID	BYTE	4
カメラ ID	Camera ID	LONG	1
格納カメラパラメータ	Stored Camera Parameter	LONG	1
姿勢パラメータ	RotationQuaternion	SLONG	4
位置パラメータ	Translation Vector	FLOAT	3
画角パラメータ	Horizontal Angle	FLOAT	1
撮影対象までの距離	Distance to Target	LONG	1
焦点距離	Focal Length	RATIONAL	1
画像の幅	Image Width	LONG	1
画像の高さ	Image Length	LONG	1
絞り値	Aperture Value	RATIONAL	1
シャッタースピード	Shutter Speed Value	SRATIONAL	1
ISO スピードレート	ISO Speed Ratings	SHORT	1

(b)

bit	データ名称
0	姿勢パラメータ
1	位置パラメータ
2	画角パラメータ
3	撮影対象までの距離
4	焦点距離
5	背景3Dモデル
6	画像の幅
7	画像の高さ
8	絞り値
9	シャッタースピード
10	ISO スピードレート
11 ~ 31	reserved

(c)

タグ名称	Field Name	タイプ	カウント
B3Mタグのバージョン	B3M Version ID	BYTE	4
格納背景3Dモデル情報	Stored Background 3D Model Info	LONG	1
点群フォーマット	Point Cloud Format	LONG	1
メッシュモデルフォーマット	Mesh Model Format	LONG	1
テクスチャとの対応座標	Corresponding coordinates of Texture	LONG	1
ダウンロード URL	Download URL	LONG	1

(d)

bit	データ名称
0	点群モデル
1	メッシュモデル
2	テクスチャとの対応座標
3	ダウンロード URL
4 ~ 31	reserved

(e)

10

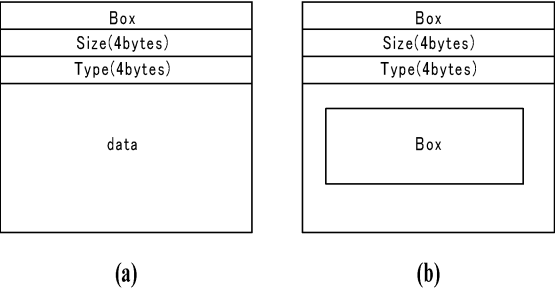
20

30

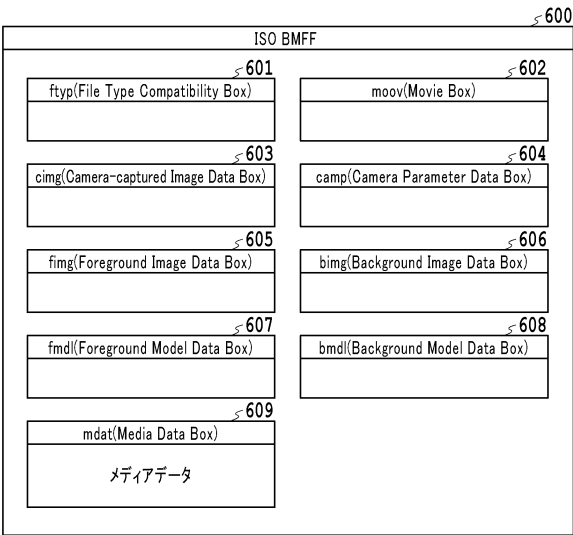
40

50

【 図 5 】



【 図 6 】



【 図 7 】

(a)

Box Type: 'cimg'
Container: Meta box ('meta')
Mandatory: No
Quantity: Zero or one

(b)

```
aligned(8) class ItemLocationBox extends FullBox('cimg',version,0) {
    unsigned int(32) offset_size;
    unsigned int(32) length_size;
    unsigned int(32) base_offset_size;
    if (version == 1) {
        unsigned int(32) index_size;
    } else {
        unsigned int(32) reserved;
    }
    unsigned int(32) camera_id; //カメラID
    unsigned int(32) Format; //フォーマット
    unsigned int(32) Image_Width; //画像の幅
    unsigned int(32) Image_Length; //画像の高さ
    unsigned int(32) bit Depth; //ビット深度
}
```

【 図 8 】

(a)

Box Type: 'camp'
Container: Meta box ('meta')
Mandatory: No
Quantity: Zero or one

(b)

```
aligned(8) class ItemLocationBox extends FullBox('camp',version,0) {
    unsigned int(32) offset_size;
    unsigned int(32) length_size;
    unsigned int(32) base_offset_size;
    if (version == 1) {
        unsigned int(32) index_size;
    } else {
        unsigned int(32) reserved;
    }
    unsigned int(32) camera_id; //カメラID
    unsigned int(32) stored_camera_parameter; //格納カメラパラメータ
    for (i=0, i<4; i++) {
        int(32) Rotetion Quaternion[i]; //姿勢パラメータ
    }
    for (i=0, i<3; i++) {
        float(32) Translation_Vector[i]; //位置パラメータ
    }
    float(32) Horizontal_Angle; //画角パラメータ
    unsigned int(32) Distance_to_Target; //撮影対象までの距離
    unsigned int(32) Focal_Length_Numerator; //焦点距離 分子
    unsigned int(32) Focal_Length_Denominator; //焦点距離 分母
    unsigned int(32) Image_Width; //画像の幅
    unsigned int(32) Image_Length; //画像の高さ
    unsigned int(32) Aperture_Value_Numerator; //絞り値 分子
    unsigned int(32) Aperture_Value_Denominator; //絞り値 分母
    int(32) Shutter_Speed_Numerator; //シャッタースピード 分子
    int(32) Shutter_Speed_Denominator; //シャッタースピード 分母
    short(16) ISO_Speed_Rating; //ISO感度
}
```

10

20

30

40

50

【 9 】

(a)

Box Type: 'fimg'
Container: Meta box ('meta')
Mandatory: No
Quantity: Zero or one

(b)

```
aligned(8) class ItemLocationBox extends FullBox('fimg',version,0) {
    unsigned int(32) offset_size;
    unsigned int(32) length_size;
    unsigned int(32) base_offset_size;
    if (version == 1) {
        unsigned int(32) index_size;
    } else {
        unsigned int(32) reserved;
    }
    unsigned int(32) Stored Foreground Image Info //格納前景情報
    unsigned int(32) Foreground Format //前景フォーマット 901
    unsigned int(32) Foreground Silhouette Format //前景シルエット 902
    フォーマット
    unsigned int(32) Trimming Foreground Format //切り出し前景フォーマット 903
}
```

【 1 0 】

(a)

Box Type: 'bimg'
Container: Meta box ('meta')
Mandatory: No
Quantity: Zero or one

(b)

```
aligned(8) class ItemLocationBox extends FullBox('bimg',version,0) {
    unsigned int(32) offset_size;
    unsigned int(32) length_size;
    unsigned int(32) base_offset_size;
    if (version == 1) {
        unsigned int(32) index_size;
    } else {
        unsigned int(32) reserved;
    }
    unsigned int(32) Stored Background Image Info //格納背景情報
    unsigned int(32) Background Format //背景フォーマット 1001
    unsigned int(32) Background Texture Format //背景テクスチャフォーマット 1002
    unsigned int(32) Trimming Background Format //切り出し背景フォーマット 1003
}
```

【 1 1 】

(a)

Box Type: 'fmdl'
Container: Meta box ('meta')
Mandatory: No
Quantity: Zero or one

(b)

```
aligned(8) class ItemLocationBox extends FullBox('fmdl',version,0) {
    unsigned int(32) offset_size;
    unsigned int(32) length_size;
    unsigned int(32) base_offset_size;
    if (version == 1) {
        unsigned int(32) index_size;
    } else {
        unsigned int(32) reserved;
    }
    unsigned int(32) Stored Foreground 3D Model Info //格納前景 3D モデル情報
    unsigned int(32) Point Cloud Format //点群フォーマット 1101
    unsigned int(32) Mesh Model Format //メッシュモデルフォーマット 1102
    unsigned int(32) Billboard Format //ビルボードフォーマット 1103
    unsigned int(32) Bounding Box Format //バウンディングボックスフォーマット 1104
}
```

【 1 2 】

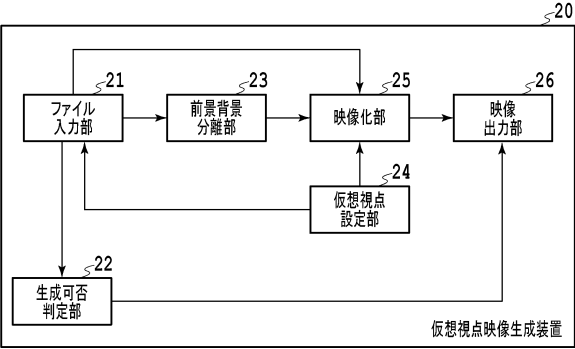
(a)

Box Type: 'bmdl'
Container: Meta box ('meta')
Mandatory: No
Quantity: Zero or one

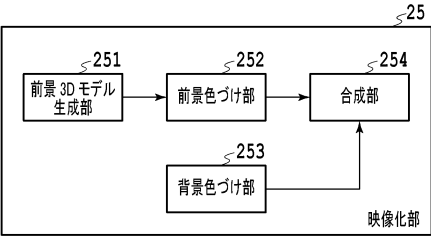
(b)

```
aligned(8) class ItemLocationBox extends FullBox('bmdl',version,0) {
    unsigned int(32) offset_size;
    unsigned int(32) length_size;
    unsigned int(32) base_offset_size;
    if (version == 1) {
        unsigned int(32) index_size;
    } else {
        unsigned int(32) reserved;
    }
    unsigned int(32) Stored Background 3D Model Info //格納背景 3D モデル情報
    unsigned int(32) Point Cloud Format //点群フォーマット 1201
    unsigned int(32) Mesh Model Format //メッシュモデルフォーマット 1202
    unsigned int(32) Corresponding Coordinates of Texture //テクスチャとの対応座標
    unsigned int(32) Download URL //ダウンロードURL
}
```

【図 1 3】

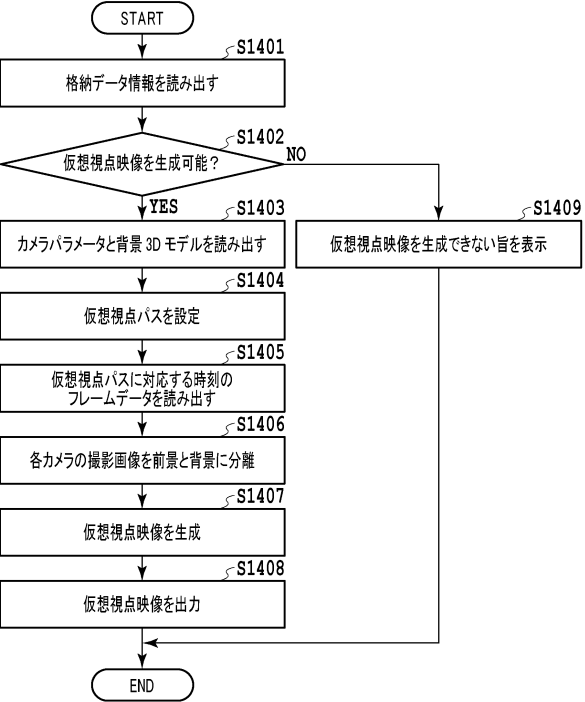


(a)

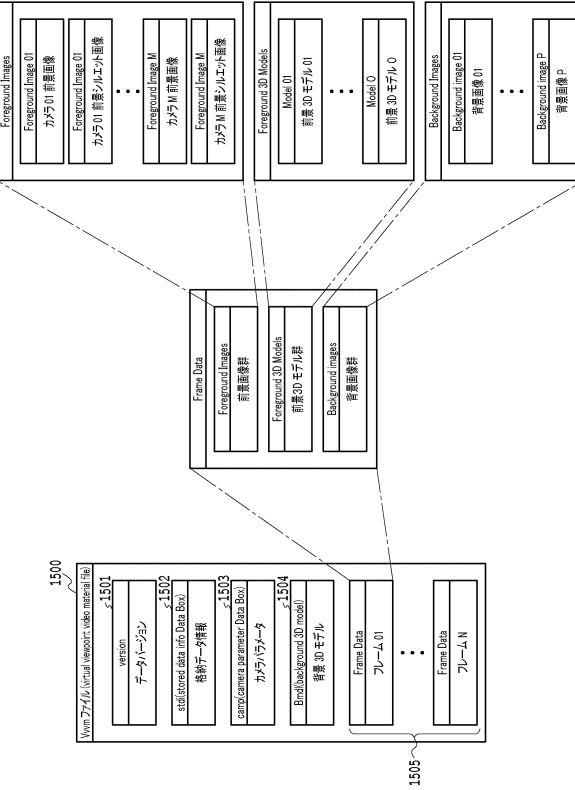


(b)

【図 1 4】



【図 1 5】



【図 1 6】

タグ名称	Field Name	タイプ	カウント	bit	データ名称
FIタグのバージョン	FI Version ID	BYTE	4	0	前景画像
格納前景情報	Stored Foreground Info	LONG	1	1	前景シルエット画像
前景フォーマット	Foreground Format	LONG	1	2	前景切り出し画像
前景シルエットフォーマット	Foreground Silhouette Format	LONG	1	3 ~ 31	reserved
切り出し前景フォーマット	Trimming Foreground Format	LONG	1		

(a)

タグ名称	Field Name	タイプ	カウント	bit	データ名称
BIタグのバージョン	BI Version ID	BYTE	4	0	背景画像
格納背景情報	Stored Background Info	LONG	1	1	背景テクスチャ
背景フォーマット	Background Format	LONG	1	2	背景切り出し画像
背景テクスチャフォーマット	Background Texture Format	LONG	1	3 ~ 31	reserved
切り出し背景フォーマット	Trimming Background Format	LONG	1		

(b)

タグ名称	Field Name	タイプ	カウント	bit	データ名称
F3Mタグのバージョン	F3MI Version ID	BYTE	4	0	点群モデル
格納前景3Dモデル情報	Stored Foreground 3D Model Info	LONG	1	1	メッシュモデル
点群フォーマット	Point Cloud Format	LONG	1	2	ビルボードモデル
メッシュモデルフォーマット	Mesh Model Format	LONG	1	3	バウンディングボックス
ビルボードフォーマット	Billboard Format	LONG	1	4 ~ 31	reserved
バウンディングボックスフォーマット	Bounding Box Format	LONG	1		

(c)

タグ名称	Field Name	タイプ	カウント	bit	データ名称
F3Mタグのバージョン	F3MI Version ID	BYTE	4	0	点群モデル
格納前景3Dモデル情報	Stored Foreground 3D Model Info	LONG	1	1	メッシュモデル
点群フォーマット	Point Cloud Format	LONG	1	2	ビルボードモデル
メッシュモデルフォーマット	Mesh Model Format	LONG	1	3	バウンディングボックス
ビルボードフォーマット	Billboard Format	LONG	1	4 ~ 31	reserved
バウンディングボックスフォーマット	Bounding Box Format	LONG	1		

(d)

タグ名称	Field Name	タイプ	カウント	bit	データ名称
F3Mタグのバージョン	F3MI Version ID	BYTE	4	0	点群モデル
格納前景3Dモデル情報	Stored Foreground 3D Model Info	LONG	1	1	メッシュモデル
点群フォーマット	Point Cloud Format	LONG	1	2	ビルボードモデル
メッシュモデルフォーマット	Mesh Model Format	LONG	1	3	バウンディングボックス
ビルボードフォーマット	Billboard Format	LONG	1	4 ~ 31	reserved
バウンディングボックスフォーマット	Bounding Box Format	LONG	1		

(e)

タグ名称	Field Name	タイプ	カウント	bit	データ名称
F3Mタグのバージョン	F3MI Version ID	BYTE	4	0	点群モデル
格納前景3Dモデル情報	Stored Foreground 3D Model Info	LONG	1	1	メッシュモデル
点群フォーマット	Point Cloud Format	LONG	1	2	ビルボードモデル
メッシュモデルフォーマット	Mesh Model Format	LONG	1	3	バウンディングボックス
ビルボードフォーマット	Billboard Format	LONG	1	4 ~ 31	reserved
バウンディングボックスフォーマット	Bounding Box Format	LONG	1		

(f)

10

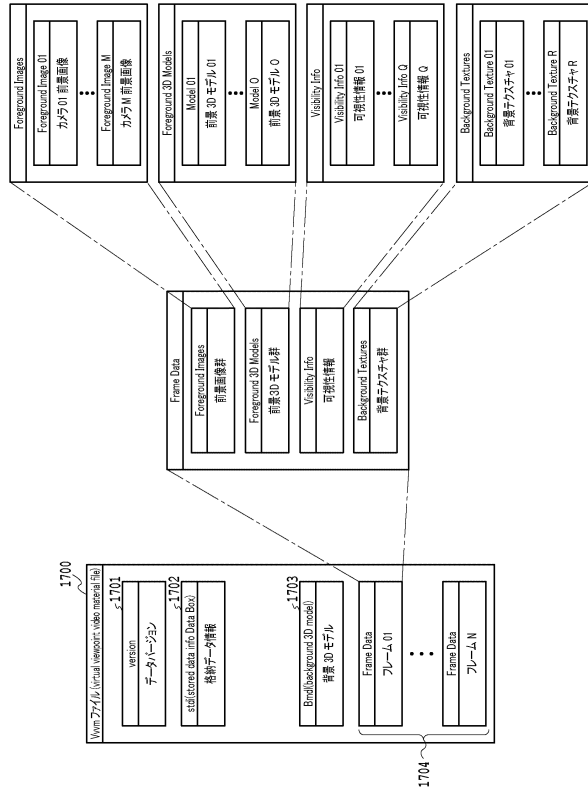
20

30

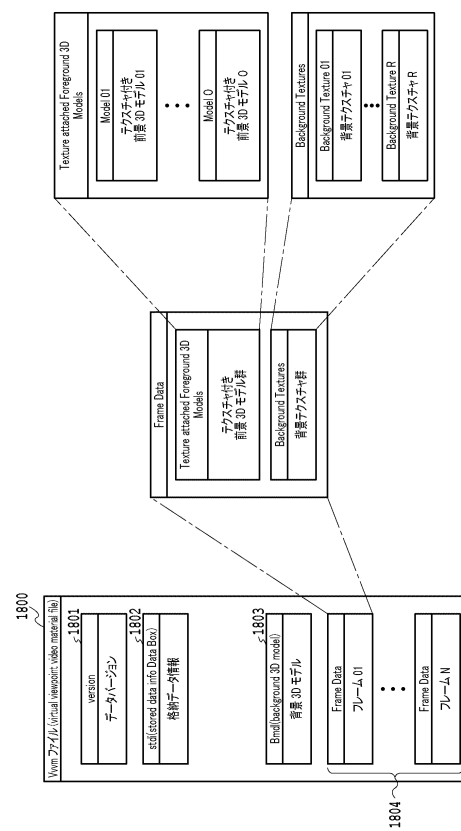
40

50

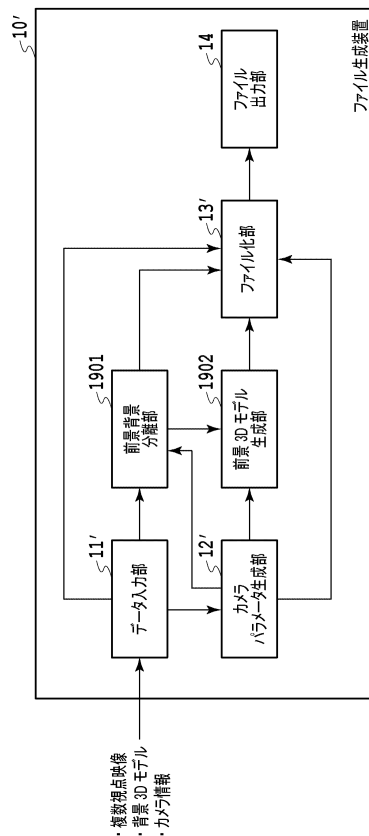
【図 17】



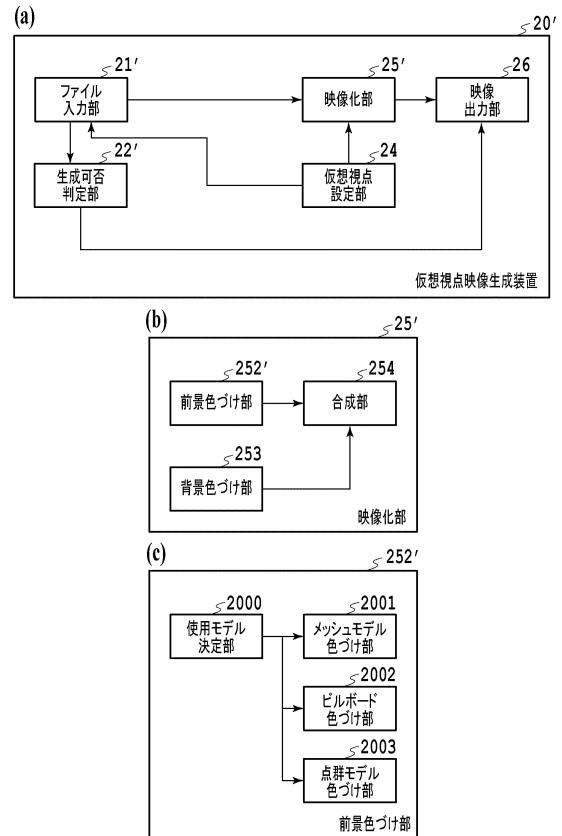
【図 18】



【図 19】



【図 20】



10

20

30

40

50

フロントページの続き

(51)国際特許分類		F I	
		G 0 6 T	5/50
(56)参考文献	特開 2 0 1 4 - 2 1 5 8 2 8 (J P , A)		
	特開 2 0 1 2 - 0 1 0 3 4 4 (J P , A)		
	特開 2 0 1 1 - 0 1 5 2 1 5 (J P , A)		
(58)調査した分野 (Int.Cl. , D B 名)			
		H 0 4 N	5 / 7 6 - 5 / 9 5 6
		G 0 6 T	1 1 / 0 0 - 1 9 / 2 0
		G 0 6 T	1 / 0 0
		G 0 6 T	3 / 0 0 - 5 / 5 0