

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7528951号
(P7528951)

(45)発行日 令和6年8月6日(2024.8.6)

(24)登録日 令和6年7月29日(2024.7.29)

(51)国際特許分類		F I			
H 0 4 N	5/64 (2006.01)	H 0 4 N	5/64	5 1 1 A	
G 0 9 G	5/00 (2006.01)	G 0 9 G	5/00	5 5 0 C	
G 0 9 G	5/377(2006.01)	G 0 9 G	5/377		

請求項の数 4 (全11頁)

(21)出願番号	特願2021-561065(P2021-561065)	(73)特許権者	000002185 ソニーグループ株式会社 東京都港区港南1丁目7番1号
(86)(22)出願日	令和1年11月28日(2019.11.28)	(74)代理人	110002147 弁理士法人酒井国際特許事務所
(86)国際出願番号	PCT/JP2019/046514	(72)発明者	徳武 健司 東京都品川区東品川4丁目12番3号 ソニーモバイルコミュニケーションズ株式会社内
(87)国際公開番号	WO2021/106136	(72)発明者	月岡 正明 東京都品川区東品川4丁目12番3号 ソニーモバイルコミュニケーションズ株式会社内
(87)国際公開日	令和3年6月3日(2021.6.3)	審査官	鈴木 圭一郎
審査請求日	令和4年11月16日(2022.11.16)		

最終頁に続く

(54)【発明の名称】 表示端末装置

(57)【特許請求の範囲】

【請求項1】

メモリと、

現実空間の画像である第二画像を撮影する撮影部と、

前記第二画像に基づいて、現実空間における仮想オブジェクトの配置位置をソフトウェア処理により決定し、前記仮想オブジェクトの画像である第一画像と前記配置位置を示す情報とを前記メモリへ出力するCPUと、

前記メモリに記憶されている最新の前記情報に基づいて、前記第一画像と最新の前記第二画像とをハードウェア処理により合成することにより合成画像を生成する合成器と、

前記合成器と直接接続され、前記合成画像を表示するディスプレイと、

を具備し、

前記撮影部は、第一ラインを介して前記CPUと接続される一方で、第二ラインを介して前記合成器と接続され、同一の時点で撮影した同一の前記第二画像を、前記第一ラインを介して前記CPUへ出力する一方で、前記第二ラインを介して前記合成器へ出力し、

ハードウェアである前記合成器は、ハードウェアである前記ディスプレイと第三ラインを介して直接接続され、

前記撮影部で撮影された前記第二画像は、前記合成器に前記第三ラインを介して直接接続された前記ディスプレイへ、前記CPUによるソフトウェア処理を施されること無く出力される、

表示端末装置。

【請求項 2】

前記合成器は、前記第二画像の水平方向の1ライン毎に前記第一画像と前記第二画像とを合成する、

請求項1に記載の表示端末装置。

【請求項 3】

前記撮影部及び前記合成器を有するカメラモジュール、を具備し、

前記カメラモジュール及び前記ディスプレイの双方がM I P I規格に準拠している、

請求項1に記載の表示端末装置。

【請求項 4】

前記CPUは、前記第二画像に基づいてS L A Mを行うことにより環境マップとポーズグラフとを生成し、前記環境マップ及び前記ポーズグラフに基づいて前記配置位置を決定する、

10

請求項1に記載の表示端末装置。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、表示端末装置に関する。

【背景技術】

【0002】

A R (Augmented Reality) 技術を利用したサービスを実現するための表示端末装置が開発されている。表示端末装置の一例としてH M D (Head Mounted Display) が挙げられ、H M Dには、例えば、光学シースルー型H M D (Optical See-Through type H M D) とビデオシースルー型H M D (Video See-Through type H M D) とがある。

20

【0003】

光学シースルー型H M Dでは、例えば、ハーフミラーや透明な導光板を用いた虚像光学系がユーザの眼前に保持され、その虚像光学系の内側に画像が表示される。このため、光学シースルー型H M Dを装着したユーザは、虚像光学系の内側に表示された画像を見ている間も、ユーザの周囲の風景を視界に入れることが可能となる。よって、光学シースルー型H M DにA R技術を適用することで、光学シースルー型H M Dの位置や姿勢に応じて、現実空間に存在するオブジェクトの光学像に対して、テキスト、アイコンまたはアニメーション等の様々な態様の仮想的なオブジェクト(以下では「仮想オブジェクト」と呼ぶことがある)の画像(以下では「仮想オブジェクト画像」と呼ぶことがある)を合成することが可能となる。

30

【0004】

一方で、ビデオシースルー型H M Dは、ユーザの眼を覆うようにユーザに装着され、ビデオシースルー型H M Dが有するディスプレイがユーザの眼前に保持される。また、ビデオシースルー型H M Dは、ユーザ前方の風景を撮影するためのカメラモジュールを有し、カメラモジュールにより撮影された風景の画像がディスプレイに表示される。このため、ビデオシースルー型H M Dを装着したユーザは、ユーザ前方の風景を直接視界に入れることは困難であるが、ディスプレイに表示される画像により、ユーザ前方の風景を確認することができる。また、ビデオシースルー型H M DにA R技術を適用することで、ユーザ前方の風景の画像を現実空間の背景の画像(以下では「背景画像」と呼ぶことがある)にし、ビデオシースルー型H M Dの位置や姿勢に応じて、仮想オブジェクト画像を背景画像に合成することが可能となる。以下では、背景画像に仮想オブジェクト画像が合成された画像を「合成画像」と呼ぶことがある。

40

【先行技術文献】

【特許文献】

【0005】

【文献】特表2018-517444号公報

【文献】特開2018-182511号公報

50

【発明の概要】**【発明が解決しようとする課題】****【0006】**

ここで、ビデオシースルー型HMDに用いられるAR技術では、背景画像に対する仮想オブジェクト画像の合成は、背景画像の分析等を含む、比較的時間を要するソフトウェア処理により行われている。このため、ビデオシースルー型HMDでは、背景画像が撮影された時点から、その背景画像を含む合成画像が表示される時点までの間に生じる遅延が大きくなってしまふ。また、背景画像は、ビデオシースルー型HMDの移動に伴って随時変化する画像である。

【0007】

よって、ビデオシースルー型HMDを装着したユーザの顔の向きが変化した場合に、ディスプレイに表示される背景画像の更新のスピードがユーザの顔の向きの変化のスピードに追従できないことがある。よって例えば図1に示すように、ビデオシースルー型HMDを装着したユーザの顔の向きが向きD1から向きD2に変化した場合に、向きD1の時点で撮影された背景画像BIが向きD2の時点でもディスプレイに表示されることがある。このため、ユーザの顔の向きが向きD2になった時点でディスプレイに表示される背景画像BIがユーザ前方の実際の風景FVと異なってしまふので、ユーザの違和感が大きくなってしまふ。

【0008】

また、合成画像に含まれる背景画像及び仮想オブジェクト画像のうち、仮想オブジェクト画像は背景画像に合成される画像であるのに対し、背景画像は、上記のように、ビデオシースルー型HMDの移動に伴って変化する画像である。このため、ビデオシースルー型HMDの移動があった場合に、背景画像が撮影された時点から、その背景画像に合成される仮想オブジェクト画像が表示または更新される時点までの間の遅延はユーザに認識され難いのに対し、背景画像の更新の遅延はユーザに認識され易い。つまり、ユーザは、仮想オブジェクト画像の表示遅延には鈍感である一方で、背景画像の更新遅延には敏感である。よって、背景画像の更新遅延が大きくなると、ユーザの違和感が大きくなってしまふ。

【0009】

そこで、本開示では、AR技術が適用されたビデオシースルー型HMD等の表示端末装置を装着したユーザの違和感を軽減することができる技術を提案する。

【課題を解決するための手段】**【0010】**

本開示によれば、表示端末装置は、CPUと、撮影部と、合成器と、ディスプレイとを有する。CPUは、現実空間における仮想オブジェクトの配置位置をソフトウェア処理により決定し、仮想オブジェクトの画像である第一画像と配置位置を示す情報とを出力する。撮影部は、現実空間の画像である第二画像を撮影する。合成器は、配置位置に基づいて第一画像と第二画像とをハードウェア処理により合成することにより合成画像を生成する。ディスプレイは、合成器と直接接続され、合成画像を表示する。

【図面の簡単な説明】**【0011】**

【図1】本開示の課題の説明に供する図である。

【図2】本開示の実施形態に係る表示端末装置の構成例を示す図である。

【図3】本開示の実施形態に係る表示端末装置における処理手順の一例を示す図である。

【図4】本開示の実施形態に係る画像合成処理の説明に供する図である。

【図5】本開示の実施形態に係る画像合成処理の説明に供する図である。

【図6】本開示の技術の効果の説明に供する図である。

【発明を実施するための形態】**【0012】**

以下に、本開示の実施形態について図面に基づいて説明する。なお、以下の実施形態において、同一の部位または同一の処理には同一の符号を付すことにより重複する説明を省

10

20

30

40

50

略することがある。

【 0 0 1 3 】

また、以下に示す項目順序に従って本開示の技術を説明する。

< 表示端末装置の構成 >

< 表示端末装置における処理手順 >

< 画像合成処理 >

[開示の技術の効果]

【 0 0 1 4 】

< 表示端末装置の構成 >

図 2 は、本開示の実施形態に係る表示端末装置の構成例を示す図である。図 2 において、表示端末装置 1 は、カメラモジュール 1 0 と、CPU (Central Processing Unit) 2 0 と、ディスプレイ 3 0 と、センサーモジュール 4 0 と、メモリ 5 0 とを有する。カメラモジュール 1 0 は、撮影部 1 1 と、メモリ 1 2 と、合成器 1 3 とを有する。表示端末装置 1 は、表示端末装置 1 のユーザの眼を覆うようにしてユーザに装着される。表示端末装置 1 の一例として、ビデオシーラー型 HMD や、スマートフォン、タブレット端末等のスマートデバイスが挙げられる。表示端末装置 1 がスマートデバイスである場合は、スマートデバイス用の頭部装着器具を用いて、スマートデバイスが、スマートデバイスのユーザの眼を覆うようにしてユーザに装着される。

10

【 0 0 1 5 】

カメラモジュール 1 0 は、ライン L 1 , L 2 , L 3 , L 4 を有する。撮影部 1 1 は、ライン L 1 を介して CPU 2 0 と接続される一方で、ライン L 4 を介して合成器 1 3 と接続される。メモリ 1 2 は、ライン L 3 を介して CPU 2 0 と接続される。合成器 1 3 は、ライン L 2 を介してディスプレイ 3 0 と接続される。

20

【 0 0 1 6 】

撮影部 1 1 は、レンズユニットとイメージセンサーとを有し、表示端末装置 1 を自身の眼を覆うように装着したユーザの前方の風景の画像を背景画像として撮影し、撮影した背景画像を合成器 1 3 及び CPU 2 0 へ出力する。撮影部 1 1 は、所定のフレームレートで背景画像を撮影する。撮影部 1 1 は、同一の時点で撮影した同一の背景画像を、一方ではライン L 4 を介して合成器 1 3 へ出力し、他方ではライン L 1 を介して CPU 2 0 へ出力する。つまり、カメラモジュール 1 0 は、カメラモジュール 1 0 によって撮影された背景画像がカメラモジュール 1 0 から CPU 2 0 へ出力されるライン L 1 を有する。

30

【 0 0 1 7 】

センサーモジュール 4 0 は、表示端末装置 1 の位置や姿勢の変化を検出するために、表示端末装置 1 の加速度及び角速度を検出し、検出した加速度及び角速度を示す情報 (以下では「センサー情報」と呼ぶことがある) を CPU 2 0 へ出力する。センサーモジュール 4 0 の一例として、IMU (Inertial measurement Unit) が挙げられる。

【 0 0 1 8 】

CPU 2 0 は、所定の周期で、背景画像とセンサー情報とに基づいて SLAM (Simultaneous Localization and Mapping) を行う。すなわち、CPU 2 0 は、背景画像とセンサー情報とに基づいて SLAM における環境マップとポーズグラフとを生成し、表示端末装置 1 が存在する現実空間を環境マップにより認識し、認識した現実空間における表示端末装置 1 の位置及び姿勢をポーズグラフにより認識する。また、CPU 2 0 は、生成した環境マップ及びポーズグラフに基づいて、現実空間における仮想オブジェクトの配置位置、つまり、背景画像における仮想オブジェクト画像の配置位置 (以下では「仮想オブジェクト配置位置」と呼ぶことがある) を決定し、決定した仮想オブジェクト配置位置を示す情報 (以下では「配置位置情報」と呼ぶことがある) を仮想オブジェクト画像に対応付けてメモリ 1 2 へ出力する。CPU 2 0 は、仮想オブジェクト画像及び配置位置情報をライン L 3 を介してメモリ 1 2 へ出力する。

40

【 0 0 1 9 】

メモリ 5 0 は、CPU 2 0 によって実行されるアプリケーションや CPU 2 0 が用いる

50

データを記憶する。例えば、メモリ50は、仮想オブジェクトのデータ（例えば、仮想オブジェクトの形や色を再現するためのデータ）を記憶し、CPU20は、メモリ50に記憶されている仮想オブジェクトのデータを用いて仮想オブジェクト画像を生成する。

【0020】

メモリ12は、CPU20から所定の周期で入力される仮想オブジェクト画像及び配置位置情報を所定の時間だけ記憶する。

【0021】

合成器13は、メモリ12に記憶されている仮想オブジェクト画像及び配置位置情報のうち、最新の仮想オブジェクト画像及び配置位置情報に基づいて、背景画像に仮想オブジェクト画像を合成して合成画像を生成する。つまり、合成器13は、撮影部11から入力される最新の背景画像に対して、配置位置情報によって示される位置に、最新の仮想オブジェクト画像を合成することにより合成画像を生成する。合成器13は、生成した合成画像をラインL2を介してディスプレイ30へ出力する。つまり、カメラモジュール10は、カメラモジュール10によって生成された合成画像がカメラモジュール10からディスプレイ30へ出力されるラインL2を有する。

10

【0022】

合成器13は、ハードウェアとして実現され、例えばワイヤードロジックを用いて作成された電子回路により実現される。つまり、合成器13は、背景画像と仮想オブジェクト画像とをハードウェア処理により合成することにより合成画像を生成する。また、合成器13とディスプレイ30とはラインL2によりハードウェア同士で直接接続される。

20

【0023】

ディスプレイ30は、合成器13から入力される合成画像を表示する。これにより、背景画像に仮想オブジェクト画像が重畳された合成画像が、表示端末装置1を装着したユーザの眼前に表示される。

【0024】

ここで、カメラモジュール10及びディスプレイ30の双方は、同一のインターフェース規格に準拠しており、例えばMIP I (Mobile Industry Processor Interface) 規格に準拠する。カメラモジュール10及びディスプレイ30の双方がMIP I規格に準拠する場合、撮影部11によって撮影された背景画像はMIP I規格によるCSI (Camera Serial Interface) を用いて合成器13へシリアル伝送され、合成器13によって生成された合成画像はMIP I規格によるDSI (Display Serial Interface) を用いてディスプレイ30へシリアル伝送される。

30

【0025】

<表示端末装置における処理手順>

図3は、本開示の実施形態に係る表示端末装置における処理手順の一例を示す図である。

【0026】

図3に示すカメラモジュールドライバ、センサーモジュールドライバ、SLAMアプリケーション、及び、ARアプリケーションはメモリ50に記憶されており、CPU20によって実行されるソフトウェアである。一方で、カメラモジュール10、センサーモジュール40及びディスプレイ30は、ハードウェアである。図3に示すカメラモジュールドライバはカメラモジュール10用のドライバであり、図3に示すセンサーモジュールドライバはセンサーモジュール40用のドライバである。

40

【0027】

図3において、ステップS101では、カメラモジュール10が背景画像をCPU20へ出力し、ステップS103では、CPU20へ入力された背景画像が、カメラモジュールドライバを介してSLAMアプリケーションへ渡される。

【0028】

また、ステップS101の処理と並行して、ステップS105では、センサーモジュール40がセンサー情報をCPU20へ出力し、ステップS107では、CPU20へ入力されたセンサー情報が、センサーモジュールドライバを介してSLAMアプリケーション

50

へ渡される。

【0029】

次いで、ステップS109では、SLAMアプリケーションが、背景画像とセンサー情報とに基づいてSLAMを行って、SLAMにおける環境マップとポーズグラフとを生成する。

【0030】

次いで、ステップS111では、SLAMアプリケーションが、ステップS109で生成した環境マップ及びポーズグラフをARアプリケーションへ渡す。

【0031】

次いで、ステップS113では、ARアプリケーションが、環境マップ及びポーズグラフに基づいて仮想オブジェクト配置位置を決定する。

10

【0032】

次いで、ステップS115では、ARアプリケーションが、仮想オブジェクト画像と配置位置情報とをカメラモジュール10へ出力し、カメラモジュール10に入力された仮想オブジェクト画像及び配置位置情報は互いに対応付けられてメモリ12に記憶される。

【0033】

ステップS117では、カメラモジュール10は、メモリ12に記憶されている仮想オブジェクト画像及び配置位置情報のうち、最新の仮想オブジェクト画像及び配置位置情報に基づいて、背景画像に仮想オブジェクト画像を合成して合成画像を生成する。

【0034】

次いで、ステップS119では、カメラモジュール10は、ステップS117で生成した合成画像をディスプレイ30へ出力する。

20

【0035】

次いで、ステップS121では、ディスプレイ30は、ステップS119で入力された合成画像を表示する。

【0036】

<画像合成処理>

図4及び図5は、本開示の実施形態に係る画像合成処理の説明に供する図である。

【0037】

図4に示すように、合成器13は、各フレームの背景画像BIの水平方向（行方向）の1ライン毎に背景画像BIに仮想オブジェクト画像VIを合成することにより合成画像CIを生成する。

30

【0038】

例えば、撮影部11、合成器13及びディスプレイ30は、垂直同期信号vsyncと、水平同期信号hsyncとに基づいて、図5に示すように動作する。図5において、「vsync+1」は、垂直同期信号vsync0の次に入力される垂直同期信号を示し、「vsync-1」は、垂直同期信号vsync0の1つ前に入力される垂直同期信号を示す。また、図5には、1つの垂直同期信号vsyncに対して、5つの水平同期信号hsyncが入力される場合を一例として示す。

【0039】

図5において、撮影部11は、水平同期信号hsyncに従って、背景画像BIの各ライン毎のYUVデータ（1ラインYUV）を合成器13へ出力する。

40

【0040】

合成器13は、撮影部11から入力されたYUVデータをRGBデータに変換する。また、合成器13は、水平同期信号hsync及び配置位置情報に従って、1ライン毎に、背景画像BIのRGBデータに対して仮想オブジェクト画像VIのRGBデータ（VIRGB）を重畳する。よって、仮想オブジェクト画像VIが存在するラインでは、合成画像のRGBデータ（合成RGB）が合成器13からディスプレイ30へ出力されて表示され、仮想オブジェクト画像VIが存在しないライン（No image）では、背景画像BIのRGBデータ（1ラインRGB）がそのまま合成器13からディスプレイ30へ出

50

力されて表示される。

【 0 0 4 1 】

以上、本開示の技術の実施形態について説明した。

【 0 0 4 2 】

なお、図 2 には、表示端末装置 1 の構成として、カメラモジュール 1 0 がメモリ 1 2 及び合成器 1 3 を有する構成を示した。しかし、表示端末装置 1 は、メモリ 1 2 及び合成器 1 3 の何れか一方または双方をカメラモジュール 1 0 外に有する構成を採ることも可能である。

【 0 0 4 3 】

[開示の技術の効果]

以上のように、本開示に係る表示端末装置（実施形態に係る表示端末装置 1）は、CPU（実施形態に係る CPU 2 0）と、撮影部（実施形態に係る撮影部 1 1）と、合成器（実施形態に係る合成器 1 3）と、ディスプレイ（実施形態に係るディスプレイ 3 0）とを有する。CPU は、現実空間における仮想オブジェクトの配置位置（実施形態に係る仮想オブジェクト配置位置）をソフトウェア処理により決定し、仮想オブジェクトの画像である第一画像（実施形態に係る仮想オブジェクト画像）と、配置位置を示す情報（実施形態に係る配置位置情報）とを出力する。撮影部は、現実空間の画像である第二画像（実施形態に係る背景画像）を撮影する。合成器は、配置位置に基づいて第一画像と第二画像とをハードウェア処理により合成することにより合成画像を生成する。ディスプレイは、合成器と直接接続され、合成画像を表示する。

【 0 0 4 4 】

例えば、撮影部及び合成器を有するカメラモジュールは、第一画像がカメラモジュールから CPU へ出力される第一ライン（実施形態に係るライン L 1）と、合成画像がカメラモジュールからディスプレイへ出力される第二ライン（実施形態に係るライン L 2）とを有する。

【 0 0 4 5 】

また例えば、合成器は、第二画像の水平方向の 1 ライン毎に第一画像と第二画像とを合成する。

【 0 0 4 6 】

また例えば、カメラモジュール及びディスプレイの双方が M I P I 規格に準拠している。

【 0 0 4 7 】

また例えば、CPU は、第二画像に基づいて S L A M を行うことにより環境マップとポーズグラフとを生成し、環境マップ及びポーズグラフに基づいて配置位置を決定する。

【 0 0 4 8 】

以上の構成によれば、撮影部で撮影された背景画像は、合成器に直接接続されたディスプレイへ、CPU によるソフトウェア処理を施されることなく出力されるため、撮影部で撮影されてから即座にディスプレイに表示される。このため、背景画像が撮影された時点から、その背景画像を含む合成画像が表示される時点までの間に生じる遅延を減少させることができる。このため、本開示に係る表示端末装置を装着したユーザの顔の向きが変化した場合に、ディスプレイに表示される背景画像の更新が、ユーザの顔の向きの変化に追従できるようになる。よって例えば図 6 に示すように、本開示に係る表示端末装置を装着したユーザの顔の向きが向き D 1 から向き D 2 に変化した場合に、ユーザの顔の向きが向き D 2 になった時点で撮影された背景画像 B I が向き D 2 の時点でディスプレイに表示される。このため、ユーザの顔の向きが向き D 2 になった時点でディスプレイに表示される背景画像 B I と、ユーザ前方の実際の風景 F V との相違は、ユーザが認識困難な程度にまで減少する。よって、以上の構成によれば、表示端末装置を装着したユーザの違和感を軽減することができる。

【 0 0 4 9 】

なお、本明細書に記載された効果はあくまで例示であって限定されるものではなく、また他の効果があっても良い。

10

20

30

40

50

【 0 0 5 0 】

また、開示の技術は以下のような構成も採ることができる。

(1)

現実空間における仮想オブジェクトの配置位置をソフトウェア処理により決定し、前記仮想オブジェクトの画像である第一画像と前記配置位置を示す情報とを出力するCPUと、現実空間の画像である第二画像を撮影する撮影部と、前記配置位置に基づいて前記第一画像と前記第二画像とをハードウェア処理により合成することにより合成画像を生成する合成器と、前記合成器と直接接続され、前記合成画像を表示するディスプレイと、を具備する表示端末装置。

10

(2)

前記撮影部及び前記合成器を有するカメラモジュール、を具備し、前記カメラモジュールは、前記第一画像が前記カメラモジュールから前記CPUへ出力される第一ラインと、前記合成画像が前記カメラモジュールから前記ディスプレイへ出力される第二ラインと、を有する、前記(1)に記載の表示端末装置。

(3)

前記合成器は、前記第二画像の水平方向の1ライン毎に前記第一画像と前記第二画像とを合成する、前記(1)または(2)に記載の表示端末装置。

20

(4)

前記カメラモジュール及び前記ディスプレイの双方がMIPI規格に準拠している、前記(2)に記載の表示端末装置。

(5)

前記CPUは、前記第二画像に基づいてSLAMを行うことにより環境マップとポーズグラフとを生成し、前記環境マップ及び前記ポーズグラフに基づいて前記配置位置を決定する、前記(1)から(4)の何れか一つに記載の表示端末装置。

【符号の説明】

【 0 0 5 1 】

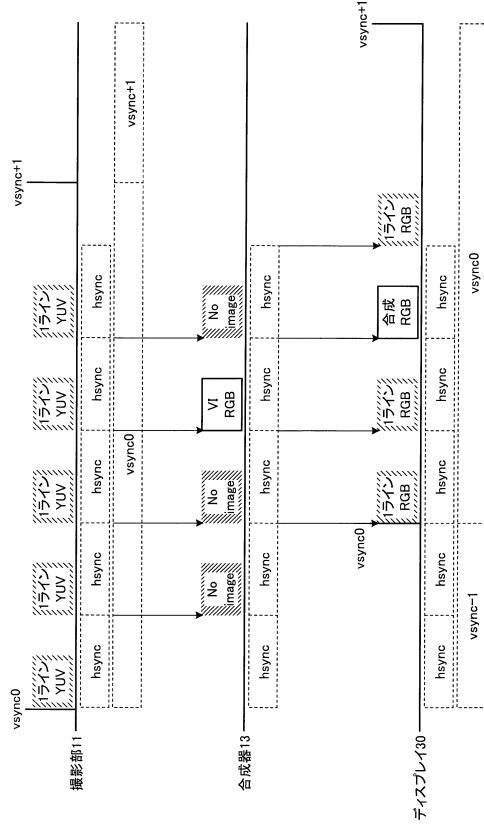
- 1 表示端末装置
- 10 カメラモジュール
- 11 撮影部
- 13 合成器
- 20 CPU
- 30 ディスプレイ
- 40 センサーモジュール

30

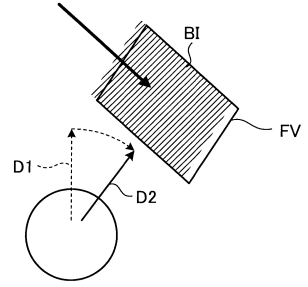
40

50

【図5】



【図6】



10

20

30

40

50

フロントページの続き

(56)参考文献 特開2018-025942(JP,A)
特開2006-106989(JP,A)
特開2017-097573(JP,A)
特表2017-530626(JP,A)
特開2016-019199(JP,A)

(58)調査した分野 (Int.Cl., DB名)
H04N 5/64
G09G 5/00
G09G 5/377