

(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
2025年2月13日(13.02.2025)



(10) 国際公開番号
WO 2025/032985 A1

- (51) 国際特許分類:
G06T 11/80 (2006.01)
- (21) 国際出願番号: PCT/JP2024/022627
- (22) 国際出願日: 2024年6月21日(21.06.2024)
- (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 2023-131665 2023年8月10日(10.08.2023) JP
- (71) 出願人: **Fotographer AI** 株式会社
(FOTOGRAPHER AI, INC.) [JP/JP]; 〒1410022
東京都品川区東五反田1丁目10番9
号STビル5階Tokyo (JP).
- (72) 発明者: 鈴木 麟太郎 (**SUZUKI Rintaro**);
〒1410022 東京都品川区東五反田1丁目10
番9号STビル5階 **Fotographer AI** 株式会
社内Tokyo (JP). **カンイ
ブラヒマ(KANE Ibrahima)**; 〒1410022 東京都
品川区東五反田1丁目10番9号STビル5
階 **Fotographer AI** 株式会

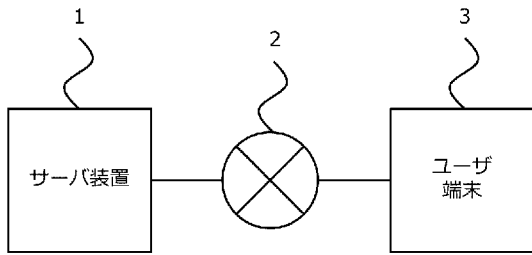
内Tokyo (JP). **カン サリユー(KANE Saliou)**;
〒1410022 東京都品川区東五反田1丁目10
番9号STビル5階 **Fotographer AI** 株式会
社内Tokyo (JP).

(74) 代理人: **One ip** 弁理士法人 (**ONE IP
PATENT FIRM**); 〒1040031 東京都中央区
京橋二丁目7番14号 **ビュレックス**
京橋208 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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC,
EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR,
HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, KE, KG, KH,
KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,
MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ,
NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,
QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, WS, ZA, ZM, ZW.

(54) **Title:** GENERATION SUPPORT DEVICE, GENERATION SUPPORT PROGRAM, AND GENERATION SUP-
PORT METHOD

(54) 発明の名称: 生成支援装置、生成支援プログラム、生成支援方法



1 Server apparatus
3 User terminal

(57) **Abstract:** Provided is a generation support device for support-
ing generation of a result image, the generation support device be-
ing characterized by including: a generation information acquisition
unit for acquiring, from a user, generation information including at
least style information relating to a style of the result image and an
element image constituting a part of the result image; and a result
image generation unit for inputting text information generated on
the basis of the generation information to a generative model, and
generating the result image on the basis of output information output
from the generative model.

(57) 要約: 【課題】 【解決手段】 結果画像の生成を支援す
る生成支援装置であって、ユーザより、少なくとも、前記
結果画像のスタイルに関するスタイル情報と、前記結果
画像の一部を構成する要素画像と、を含む生成情報を取
得する生成情報取得部と、前記生成情報をもとに生成し
たテキスト情報を生成モデルに入力し、前記生成モデル
から出力した出力情報に基づいて、前記結果画像を生成す
る結果画像生成部と、を備えることを特徴とする。

(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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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).

規則4.17に規定する申立て:

- 一 不利にならない開示又は新規性喪失の例外に関する申立て (規則4.17(v))

添付公開書類:

- 一 国際調査報告 (条約第21条(3))

明 細 書

発明の名称：生成支援装置、生成支援プログラム、生成支援方法

技術分野

[0001] 本発明は、生成支援装置、生成支援プログラム、生成支援方法に関する。

背景技術

[0002] 近年、様々な方法により画像生成が行われている。

先行技術文献

特許文献

[0003] 特許文献1：特許第7169027号公報

発明の概要

発明が解決しようとする課題

[0004] 例えば、特許文献1では、機械学習を利用してキャラクタ画像を生成する技術が提案されている。

[0005] しかしながら、特許文献1の技術では、任意の姿勢をしたキャラクタの画像を生成することしかできず、様々な画像生成には適用できない。

[0006] 本発明はこのような背景を鑑みてなされたものであり、ユーザが目的とする画像を容易に生成することを目的とする。

課題を解決するための手段

[0007] 上記の課題を解決するため、本開示にかかる、生成支援装置は、ユーザより、少なくとも、結果画像のスタイルに関するスタイル情報と、前記結果画像の一部を構成する要素画像と、を含む生成情報を取得する生成情報取得部と、前記生成情報をもとに生成したテキスト情報を生成モデルに入力し、前記生成モデルから出力した出力情報に基づいて、前記結果画像を生成する結果画像生成部と、を備えることを特徴とする。

[0008] その他本願が開示する課題やその解決方法については、発明の実施形態の欄及び図面により明らかにされる。

発明の効果

[0009] 本発明によれば、ユーザが目的とする画像を容易に生成することができる。

図面の簡単な説明

- [0010] [図1]本発明の一実施形態に係る評価システムの全体構成例を示す図である。
[図2]同実施形態に係るサーバ装置1のハードウェア構成例を示す図である。
[図3]同実施形態に係るサーバ装置1の機能構成例を示す図である。
[図4]生成情報記憶部131に記憶される基本情報の例を示す図である。
[図5]生成情報取得部111が部分画像と素材画像の位置情報を取得する画面の一例である。
[図6]同実施形態に係るサーバ装置1の処理の例を示す図である。

発明を実施するための形態

[0011] <発明の概要>

[項目1]

結果画像の生成を支援する生成支援装置であって、
ユーザより、少なくとも、前記結果画像のスタイルに関するスタイル情報と、前記結果画像の一部を構成する要素画像と、を含む生成情報を取得する生成情報取得部と、
前記生成情報をもとに生成したテキスト情報を生成モデルに入力し、前記生成モデルから出力した出力情報に基づいて、前記結果画像を生成する結果画像生成部と、
を備えることを特徴とする、生成支援装置。

[項目2]

前記生成情報取得部は、更に前記要素画像の前記結果画像における位置情報を取得すること、
を特徴とする、項目1に記載の生成支援装置。

[項目3]

前記スタイル情報はウェブアドレスの情報を含み、

前記生成情報取得部は、前記ウェブアドレスで指定されるウェブサイトに含まれるウェブ情報をもとに判定した前記スタイル情報を前記スタイル情報とすること、

を特徴とする、項目 1 または 2 に記載の生成支援装置。

[項目 4]

前記情報取得部は、前記要素画像のアップロードまたは選択を受け付け、前記結果画像の枠における前記要素画像の配置により前記位置情報を取得すること、

を特徴とする、項目 2 に記載の生成支援装置。

[項目 5]

前記情報取得部は、前記生成情報を前記ユーザによるチャット形式での入力により取得すること、

を特徴とする、項目 1 または 2 に記載の生成支援装置。

[項目 6]

前記情報取得部は、前記チャット形式での入力を取得する場合に、前記生成情報として必要となる情報の示唆を前記ユーザに提示すること、

を特徴とする、項目 5 に記載の生成支援装置。

[項目 7]

結果画像の生成を支援する生成支援プログラムであって、プロセッサに、

ユーザより、少なくとも、前記結果画像のスタイルに関するスタイル情報と、前記結果画像の一部を構成する要素画像と、を含む生成情報を取得する生成情報取得ステップと、

前記生成情報をもとに生成したテキスト情報を生成モデルに入力し、前記生成モデルから出力した出力情報に基づいて、前記結果画像を生成する画像生成ステップと、

を実行させる、生成支援プログラム。

[項目 8]

結果画像の生成を支援する生成支援方法であって、
プロセッサが、

ユーザより、少なくとも、前記結果画像のスタイルに関するスタイル情報
と、前記結果画像の一部を構成する要素画像と、を含む生成情報を取得する
生成情報取得ステップと、

前記生成情報をもとに生成したテキスト情報を生成モデルに入力し、前記
生成モデルから出力した出力情報に基づいて、前記結果画像を生成する画像
生成ステップと、

を実行させる、生成支援方法。

[0012] 図1は、本発明の一実施形態に係る評価システムの全体構成例を示す図である。本実施形態の生成支援システムは、サーバ装置1を含んで構成される。サーバ装置1は、ユーザ端末3と、通信ネットワーク2を介して通信可能に接続される。通信ネットワーク2は、たとえばインターネットであり、公衆電話回線網や携帯電話回線網、無線通信路、イーサネット（登録商標）などにより構築される。

[0013] ==サーバ装置1==

サーバ装置1は、例えば、ワークステーションやパーソナルコンピュータのような汎用コンピュータとしてもよいし、或いはクラウド・コンピューティングによって論理的に実現されてもよい。本実施形態においては、説明の便宜上1台を例示しているが、これに限定されず、複数台であってもよい。

[0014] ==ユーザ端末3==

ユーザ端末3は、画像の生成を行うユーザが扱うコンピュータである。例えば、スマートフォンやタブレットコンピュータ、パーソナルコンピュータなどである。ユーザは、たとえばユーザ端末3で実行されるアプリケーションやWebブラウザによりサーバ装置1にアクセスすることができる。

[0015] 図2は、サーバ装置1のハードウェア構成例を示す図である。なお、図示された構成は一例であり、これ以外の構成を有していてもよい。サーバ装置

1は、プロセッサ101、メモリ102、記憶装置103、通信インターフェース104、入力装置105、出力装置106を備える。記憶装置103は、各種のデータやプログラムを記憶する、例えばハードディスクドライブやソリッドステートドライブ、フラッシュメモリなどである。通信インターフェース104は、通信ネットワーク2に接続するためのインターフェースであり、例えばイーサネット（登録商標）に接続するためのアダプタ、公衆電話回線網に接続するためのモデム、無線通信を行うための無線通信機、シリアル通信のためのUSB（Universal Serial Bus）コネクタやRS232Cコネクタなどである。入力装置105は、データを入力する、例えばキーボードやマウス、タッチパネル、ボタン、マイクロフォンなどである。出力装置106は、データを出力する、例えばディスプレイやプリンタ、スピーカなどである。なお、後述するサーバ装置1の各機能部はプロセッサ101が記憶装置103に記憶されているプログラムをメモリ102に読み出して実行することにより実現され、サーバ装置1の各記憶部はメモリ102及び記憶装置103が提供する記憶領域の一部として実現される。

[0016] 図3は、サーバ装置1の機能構成を示している。図3に示すように、サーバ装置1は、生成情報記憶部131と、結果画像情報記憶部132と、の各記憶部と、生成情報取得部111と、結果画像生成部112と、の各処理部を備える。

[0017] 生成情報記憶部131と、結果画像情報記憶部132と、の各記憶部の説明を記載する。

[0018] 生成情報記憶部131は、図4に一例を示すように、結果画像（サーバ装置1が生成する画像）の生成に用いる情報（以下、生成情報と記す）を記憶する。生成情報は、一例として、スタイルに関する情報（テキスト情報、ウェブアドレス、ウェブ情報等を含む）等の情報を含んでもよい。また、生成情報は、結果画像の一部の構成のもととなる要素画像を含んでもよい。要素画像は、例えば、結果画像の主題（例えば人物や物であって、以下に記す生成

情報取得部 1 1 1 が生成モデルに入力するテキストを生成する際に、画像のメインの内容とするように記載する対象) の画像である部分画像であって、部分画像は主題が物の画像であれば、例えば製品の画像、製品を含む画像、製品の容器、外箱などの外見の画像を含んでよい。また、要素画像には、結果画像の主題ではない、素材画像を含んでもよい。生成情報は、例えば、結果画像における部分画像、素材画像の位置の情報、を含んでもよいが、これらに限定されない。

[0019] 前記スタイルとは、生成支援装置が生成する結果画像のデザインにおけるスタイルであって、例えば、結果画像の要件（画像に含まれる物体、人物、風景などの要素など）、また、コンセプト（ターゲット、物語性など）、色彩、質感（視覚的な感覚で感じられる画像表面のテクスチャなど）、レイアウト（要素の配置や相対的な位置関係など）、フォント、形状（シャープな角度や丸みのある形状、直線、曲線など）などの美的な属性や特徴を表すが、これらに限定されない。

[0020] 前記素材画像とは、手や顔、植物、日用品、台などの画像、丸、三角、四角などの幾何学的な、または幾何学的な規則に拘束されない自由な形状の図形などであって、結果画像の一部の基となる画像である。また、前記素材画像には、生成する画像の背景の基となる画像（テンプレート）も含んでよいが、これらに限定されない。

[0021] 結果画像情報記憶部 1 3 2 は、結果画像生成部 1 1 2 が生成した結果画像を記憶する。

[0022] 以下に、生成情報取得部 1 1 1 と、結果画像生成部 1 1 2 と、の各処理部の説明を記載する。

[0023] 生成情報取得部 1 1 1 は、一例として、通信ネットワーク 2 を介して、ユーザ端末 3 から、結果画像生成に必要な、前記結果画像のスタイルに関するスタイル情報と、前記結果画像を構成する要素画像と、前記要素画像の前記結果画像における位置情報と、を含む生成情報を取得する。生成情報取得部 1 1 1 は、取得した生成情報を、生成情報記憶部 1 3 1 に記憶する。当該送

受信における通信は、有線、無線のいずれでもよく、また、互いの通信が実行できるのであれば、どのような通信プロトコルを用いてもよい。

[0024] 生成情報取得部 111 は、生成情報をテキスト情報で取得してもよい。生成情報取得部 111 は、生成する結果画像を示す文章をユーザの入力操作により取得してもよいし、1以上の単語を取得してもよい。また、生成情報取得部 111 は、生成する結果画像のスタイルを表す文章または単語をユーザに提示し、ユーザにより選択された文章または単語を生成情報として取得してもよい。

[0025] 生成情報取得部 111 は、生成情報としてウェブアドレス（URL等）の情報を取得してもよい。生成情報取得部 111 は、前記ウェブアドレスで指定されるウェブサイトに含まれるウェブ情報を取得し、スタイルを判定し、生成情報とすればよい。ウェブ情報としては、当該ウェブサイトに含まれるテキスト情報、画像情報、映像情報、コード情報（ウェブサイトを構成するコードであって、例えば、HTML、CSS、JavaScriptなどの形式であってよいが、これらに限定されない）等であってよい。生成情報取得部 111 は、例えば当該テキスト情報を基にターゲットやコンセプト等のスタイルを判定してもよい。この場合、生成情報取得部 111 は、例えば当該テキスト情報を形態素解析し、含まれる単語やその数の情報を基に、ターゲットやコンセプト等のスタイルを判定すればよいが、これらの方法に限定されない。生成情報取得部 111 は、当該画像情報や映像情報、またはコード情報などから色彩や質感、フォント、形状等のスタイルを判定してもよい。この場合、生成情報取得部 111 は、当該画像情報や映像情報を分析し、多く含まれる色彩や質感、フォント、形状等からスタイルを判定してもよいし、コード情報に含まれる、ウェブ背景画像の色彩や質感、形状、ウェブで用いられているフォント等の情報をもとに、スタイルを判定してもよいが、これらの方法に限定されない。

[0026] 生成情報取得部 111 は、結果画像の一部の構成のもととなる要素画像を取得する。生成情報取得部 111 は、要素画像のアップロードを受け付けて

もよい。また、生成情報取得部111は、例えば図5に一例を示すように、素材画像（例えば図5の201）をサーバ装置1に記憶しておきユーザ端末3に提示し、ユーザによる当該素材画像の選択操作を受け付け、ユーザが選択した素材画像を生成情報として取得してもよい。

[0027] 生成情報取得部111は、一例として、結果画像における要素画像の位置情報を取得する。位置情報は、結果画像における要素画像の座標を示し、例えば、結果画像の特定の角などの所定の位置を原点としたXY座標であって、要素画像の中心などのXY座標であればよい。この場合、生成情報取得部111は、例えば図5に一例を示すように、結果画像の形状に応じた枠（例えば202。当該枠は横長、正方形、縦長などの形態であってよいが、これらに限定されない）をユーザ端末3に提示する。生成情報取得部111は、当該枠において、ユーザ端末3上でのユーザの操作の情報を取得し、結果画像における素材画像の位置情報を取得する。この場合、生成情報取得部111は、例えばユーザによるドラッグアンドドロップ操作を受け付け、部分画像（203）または素材画像（204）の配置情報を取得し、要素画像の位置情報を取得すればよい。また、生成情報取得部111は、要素画像の拡大、縮小、回転、反転、変形なども併せて受け付けてもよい。更に、生成情報取得部111は、位置情報として、複数の要素画像の前後関係（例えば、レイヤー情報であってよい）を取得してもよい。その他、位置情報は、要素画像同士の位置関係（部分画像の下部に素材画像がある、など）の情報であってもよい。

[0028] 生成情報取得部111は、チャット形式で生成情報を取得してもよい。この場合、生成情報取得部111は、チャット形式で取得したテキスト情報を形態素解析等により単語二分割し、単語の情報を生成情報として取得すればよい。この場合、生成情報取得部111は、ユーザから取得する生成情報に関して、「主題の画像をアップしてください」、「参考となるウェブサイトをお教えください」などの案内をユーザに提示し、必要な生成情報をユーザが認識しやすくなる支援を行ってもよい。この場合、生成情報取得部111

は、事前に準備された、画像の生成に必要な生成情報のリストから、ユーザから取得していない情報、または取得した生成情報からは判定できなかった情報について、ユーザに対して案内を提示すればよいが、この方法に限定されない。

[0029] 生成情報取得部111は、取得した生成情報をもとに、画像生成モデルに入力するプロンプトまたは前提となる条件等（本明細書においては纏めてプロンプト情報と記載する）を生成する。前提となる条件は、例えば画像のサイズや枠の形状、ファイルサイズ、解像度等の情報を含んでもよいが、これらに限定されない。生成情報取得部111が生成するプロンプトは、少なくともスタイルを表すテキストを含む。生成情報取得部111は、例えば、特徴抽出モジュールと言語モデルを用いて、プロンプト情報を生成する。なお、生成情報取得部111は、プロンプト情報を1以上生成し、ユーザに提示をして、プロンプトの選択、または編集を受け付けてもよい。

[0030] 生成情報取得部111は、プロンプトを生成する際に、結果画像生成部112が画像生成に用いる生成モデルの種類によって、生成するプロンプトの構造を変えてもよい。生成情報取得部111は、例えば、文章型のプロンプトを生成してもよいし、単語を並べる形のプロンプトを生成してもよい。また、例えば、重要な単語をカッコで囲むことや、単語の順番をプロンプトの冒頭に登場させること、重要な単語を複数含めること、などの単語の重要度を示す方法により、重要な単語を生成モデル側に認識させる形のプロンプトを生成してもよい。

[0031] 生成情報取得部111が生成するプロンプトは、少なくともスタイルを表すテキストを含む。また、生成情報取得部111は、プロンプトを複数生成してもよいし、プロンプトに含まれるテキストに一定のランダム性を持たせてもよい。例えば、生成情報取得部111は、スタイルを表すテキストとの意味合いの距離や類似度などにより、プロンプトに含まれるテキストに一定のランダム性を持たせる。具体的には、生成情報取得部111は、ユーザの入力等により取得した生成情報に、例えば「海」にまつわるスタイルの情報

が複数含まれている場合に、「海」と意味合いの近い、または類似度の高い他の単語をプロンプトに含めて、プロンプトを生成する。後述する結果画像生成部112が、これらのプロンプトを用いて結果画像を生成することにより、ユーザが生成を希望する画像により近い結果画像の生成を行うことができる。逆に、生成情報取得部111は、ユーザの入力等により取得した生成情報に、「海」にまつわるスタイルの情報が少ない場合に、「海」と意味合いの遠い、または類似度の低い他の単語をプロンプトに含めて、プロンプトを生成する。後述する結果画像生成部112が、これらのプロンプトを用いて結果画像を生成することにより、ユーザがまだ海のイメージが湧いていない場合などに、結果画像のスタイルの方向性を検討し易くすることができる。なお、プロンプトに含まれるテキストに一定のランダム性を持たせることの効果は、他にあってよい。

[0032] 生成情報取得部111は、結果画像生成部112が生成した第1の結果画像に対して、追加で生成情報を取得してもよい。生成情報取得部111が追加で取得した生成情報は、第1の結果画像を生成した場合に用いたプロンプトの改変、追記等に用いられ、結果画像生成部112が第2の結果画像を生成する際に用いる。

[0033] 結果画像生成部112は、一例として、スタイル情報と、要素画像と、位置情報と、の少なくともいずれかをもとに結果画像を生成する。結果画像生成部112は、例えば、スタイル情報と、要素画像と、位置情報と、の少なくともいずれかの情報をもとに生成情報取得部111が生成したプロンプト情報を生成モデルに入力し、生成モデルが出力した画像を取得する。結果画像生成部112は、生成モデルが出力した画像を結果画像としてもよいし、出力した画像を基に加工等を行い、結果画像を生成してもよい。結果画像生成部112は、生成した結果画像をユーザに提示する。ユーザは、提示された画像をダウンロードすることができる。

[0034] 結果画像生成部112が結果画像の生成に用いる生成モデルは、サーバ装置1に実装されていてもよいし、通信ネットワーク2を通じてアクセスでき

る他のサーバに実装されていてもよいが、これらに限定されない。このため、生成モデルがサーバ装置 1 に実装されている場合には、結果画像生成部 1 1 2 がプロンプト情報を生成モデルに入力し、生成モデルが他のサーバに実装されている場合には、結果画像生成部 1 1 2 はプロンプト情報を、通信ネットワーク 2 を介して生成モデルに送信する構成となる。本明細書においては、プロンプト情報を生成モデルに送信する場合も含めて、プロンプト情報を生成モデルに入力する、と表現している。

[0035] 前記生成モデルは、例えば、入力として与えられた特定の入力ベクトルやランダムなノイズを受け取り、それらの情報から画像を生成するモデルであればよい。前記生成モデルは、例えば、生成器 (Generator) を備える。生成器は、入力された情報を適切なフィーチャーやパターンに変換し、それを画像に変換する。生成器は、例えば、畳み込みニューラルネットワーク (Convolutional Neural Network, CNN) やトランスフォーマー (Transformer)、またはその他のディープラーニングアーキテクチャを使用して構築されるが、他のアーキテクチャも利用可能である。また、前記生成モデルは、例えば、識別器 (Discriminator) を備える。識別器は、画像が本物の画像か、生成器によって生成された偽の画像かを識別する。識別器は、例えば、CNNなどのネットワークを使用して構築されるが、これらに限定されない。前記生成モデルは、例えば、敵対的ネットワーク (GAN) を備える。敵対的ネットワークは、生成器がよりリアルな画像を生成するように学習し、同時に識別器が本物の画像と偽の画像を区別する能力を高めるように学習する。

[0036] 結果画像生成部 1 1 2 は、2 以上の結果画像を生成してもよい。また、結果画像生成部 1 1 2 は、生成した結果画像をユーザに提示する。

[0037] 結果画像生成部 1 1 2 は、生成した複数の結果画像をユーザ端末 3 に提示した場合、ユーザ端末 3 において、ユーザより、複数の画像から生成したい結果画像に近い画像、または外れている画像の選択操作を受け付け、当該選択操作により選択された結果画像をもとに、更に結果画像を生成してもよい

。その場合、結果画像生成部 112 は、例えば、画像から生成したい結果画像に近いものとして選択された結果画像 A の特徴をもとに、結果画像 A に似た結果画像 B を生成してもよい。具体的な処理としては、生成情報取得部 111 は、当該選択操作により選択された結果画像 A を生成する際に生成モデルに入力したプロンプトの情報を改変、または選択された結果画像 A に似た画像やバリエーションの再生成を示すプロンプトを生成し、これらのプロンプトを再度生成モデルに入力して結果画像を生成すればよいが、この方法に限定されない。

[0038] 結果画像生成部 112 は、生成した結果画像（第 1 の結果画像）に対して、生成情報取得部 111 がユーザより追加情報を取得した場合に、第 2 の結果画像を生成してもよい。この場合、結果画像生成部 112 は、前記第 1 の結果画像を生成する際に前記生成モデルに入力したプロンプト情報と、当該追加情報をもとに生成情報取得部 111 が生成したプロンプト情報を生成モデルに入力し、前記生成モデルが出力した出力情報に基づいて、第 2 の結果画像を生成すればよい。

[0039] 図 6 は、本実施形態の生成支援装置の処理の例を説明する図である。

[0040] サーバ装置 1 は、ユーザより生成情報を取得する（1001）。サーバ装置 1 は、取得した生成情報を基にプロンプトを生成する（1002）。サーバ装置 1 は、プロンプトを生成モデルに入力する（1003）。サーバ装置 1 は、生成モデルの出力情報（結果画像）を取得する（1004）。サーバ装置 1 は、出力情報をユーザに提示する（1005）。

[0041] 以下、その他の例を記載する。

[0042] サーバ装置 1 は、例えば、生成情報取得部 111 が取得した部分画像の前処理を行ってもよい。サーバ装置 1 は、例えば、部分画像の主題を判定し、主題部分以外の背景を除去してもよい。また、生成情報取得部 111 は、例えば部分画像から主題を強調してもよい。

[0043] サーバ装置 1 は、前処理として、例えば、部分画像に含まれる主題について、主題とカメラの位置の関係を判定してカメラアングルを検出し、それに

合わせて生成情報取得部 111 が生成するプロンプトを生成してもよい。当該プロンプトは、例えば、結果画像において、主題を表示する角度を指定するものなどを含むが、この例に限定されない。

[0044] なお、生成情報取得部 111 は、上述した前処理を行った後の部分画像について、生成する画像における位置情報を取得してもよいし、プロンプトを生成してもよい。

[0045] サーバ装置 1 は、マーケティング情報をもとに、ユーザにスタイルを提案してもよい。当該マーケティング情報は、事前に取得した、結果画像の主題となる製品等の情報や、業界の情報、マーケティング調査等の結果などの情報や、ユーザから取得した、主題となる製品の過去の製品販売の実績、類似製品の売れ行き等の情報を含んでいてもよい。サーバ装置 1 は、例えば、生成する画像の主題の製品について、類似する製品の販売実績等の情報を元に、販売数の多い類似製品の販売ウェブサイトや広告画像などから判定されるスタイルを、ユーザに提案する。この場合、サーバ装置 1 は、販売数の多い類似製品の販売ウェブサイトのウェブアドレスの情報や、スタイルとして生成情報取得部 111 が生成するプロンプトに含めるテキスト情報（例えば「高級感」、「ナチュラル」など）をユーザに提示してもよいし、画像生成に用いるプロンプトに含めるようにすればよい。

[0046] サーバ装置 1 は、ユーザが過去にサーバ装置 1 を用いて生成した結果画像、または当該結果画像を生成する際に用いたプロンプトの情報を元に、ユーザにスタイルのレコメンドを行ってもよい。例えば、サーバ装置 1 は、ユーザが過去に生成した結果画像を分析、またはプロンプトに含まれるテキストの情報によりスタイルを判定し、多く検出されたスタイルを、ユーザ端末 3 に提示し、当該スタイルを結果画像生成に用いるかどうかの選択操作を取得すればよい。具体的には、サーバ装置 1 は、例えば、当該ユーザが過去にリアルな画像しか生成していないと判定した場合に、「リアルな画像を生成しますか？はい いいえ」などの質問をチャット等によりユーザ端末 3 に提示し、ユーザの選択操作を取得し、ユーザが選択した答えにより、プロンプト

を生成すればよい。

[0047] サーバ装置1は、画像に限らず、製品販売に関わる情報を生成してもよい。サーバ装置1が生成する情報は、例えば、製品やキャンペーンなどを宣伝するバナー広告用画像、商品やブランドの特長を簡潔に表現した効果的なキャッチフレーズ・キャッチコピー、商品の詳細な説明や特長を記述するテキスト情報であるプロダクトディスクリプション、製品を販売するECサイトのトップページなどで使われるデザインやレイアウト、商品カテゴリーページのデザインや表示方法であるカテゴリーページデザイン、特定のキャンペーンや商品を強調するためのランディングページのデザイン、ソーシャルメディアや広告プラットフォーム向けの画像やキャッチフレーズなどを含むが、これらに限定されない。サーバ装置1が上述した情報を生成する場合、生成情報取得部111が取得した情報をもとにプロンプトを生成し、画像やデザインであれば結果画像生成部112が画像生成モデルにプロンプトを入力し、テキスト情報であればテキスト生成モデル（例えば、ChatGPTのような大規模言語モデル）にプロンプトを入力し、生成すればよい。

[0048] サーバ装置1は、結果画像の一部を構成する要素画像を取得する場合に、指定されたウェブアドレスが特定するウェブサイトに含まれる画像を、要素画像として取得してもよい。サーバ装置1は、当該ウェブサイトに含まれる画像を全て要素画像として取得し、生成情報記憶部121に記憶してもよいし、当該ウェブサイトに含まれる画像の中から、生成情報として取得する画像についてユーザによる選択操作を受け付けて、選択された画像を要素画像として記憶してもよい。

[0049] 以上、添付図面を参照しながら本開示の好適な実施形態について詳細に説明したが、本開示の技術的範囲はかかる例に限定されない。本開示の技術分野における通常の知識を有する者であれば、特許請求の範囲に記載された技術的思想の範疇内において、各種の変更例または修正例に想到し得ることは明らかであり、これらについても、当然に本開示の技術的範囲に属するものと了解される。

[0050] 本明細書において説明した装置は、単独の装置として実現されてもよく、一部または全部が通信ネットワーク2で接続された複数の装置（例えばクラウドサーバ）等により実現されてもよい。例えば、サーバ装置1のプロセッサ101および記憶装置103は、互いに通信ネットワーク2で接続された異なるサーバにより実現されてもよい。

[0051] 本明細書において説明した装置による一連の処理は、ソフトウェア、ハードウェア、およびソフトウェアとハードウェアとの組合せのいずれを用いて実現されてもよい。本実施形態に係るサーバ装置1の各機能を実現するためのコンピュータプログラムを作製し、PC等を実装することが可能である。また、このようなコンピュータプログラムが格納された、コンピュータで読み取り可能な記録媒体も提供することができる。記録媒体は、例えば、磁気ディスク、光ディスク、光磁気ディスク、フラッシュメモリ等である。また、上記のコンピュータプログラムは、記録媒体を用いずに、例えば通信ネットワーク2を介して配信されてもよい。

[0052] また、本明細書において説明した処理は、必ずしも説明した順序で実行されなくてもよい。いくつかの処理ステップは、並列的に実行されてもよい。また、追加的な処理ステップが採用されてもよく、一部の処理ステップが省略されてもよい。

[0053] また、本明細書に記載された効果は、あくまで説明的または例示的なものであって限定的ではない。つまり、本開示に係る技術は、上記の効果とともに、または上記の効果に代えて、本明細書の記載から当業者には明らかな他の効果を奏しうる。

符号の説明

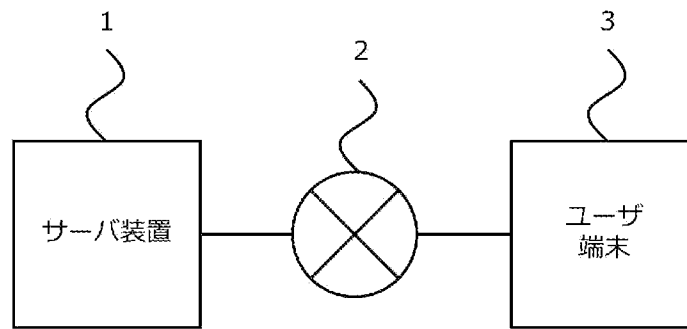
- [0054]
- 1 サーバ装置
 - 2 通信ネットワーク
 - 3 ユーザ端末
 - 101 CPU
 - 102 メモリ

- 1 0 3 記憶装置
- 1 0 4 通信インタフェース
- 1 0 5 入力装置
- 1 0 6 出力装置
- 1 1 1 生成情報取得部
- 1 1 2 結果画像生成部
- 1 3 1 生成情報記憶部
- 1 3 2 画像情報記憶部

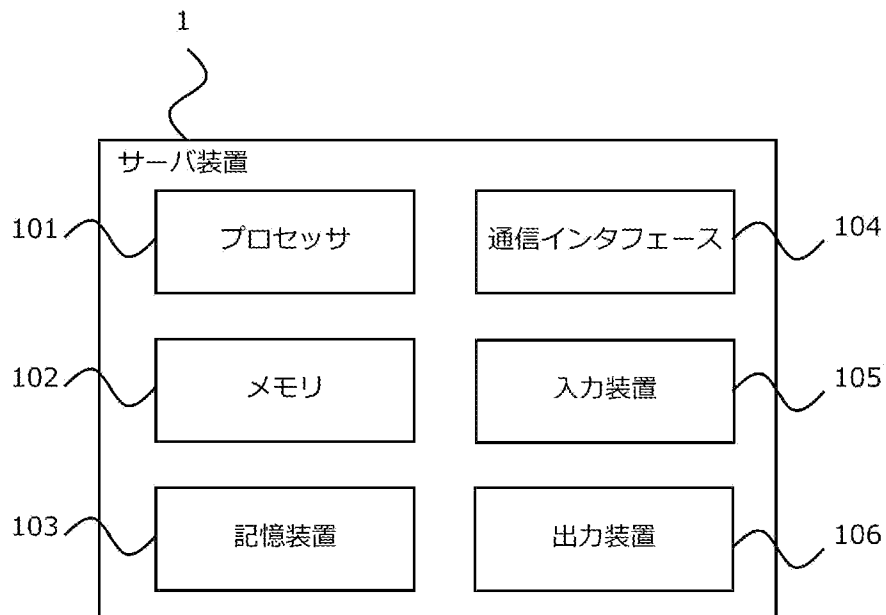
請求の範囲

- [請求項1] 結果画像の生成を支援する生成支援装置であって、
ユーザより、少なくとも、前記結果画像のスタイルに関するスタイル情報と、前記結果画像の一部を構成する要素画像と、を含む生成情報を取得する生成情報取得部と、
前記生成情報をもとに生成したテキスト情報を生成モデルに入力し、前記生成モデルから出力した出力情報に基づいて、前記結果画像を生成する結果画像生成部と、
を備えることを特徴とする、生成支援装置。
- [請求項2] 前記生成情報取得部は、更に前記要素画像の前記結果画像における位置情報を取得すること、
を特徴とする、請求項1に記載の生成支援装置。
- [請求項3] 前記スタイル情報はウェブアドレスの情報を含み、
前記生成情報取得部は、前記ウェブアドレスで指定されるウェブサイトに含まれるウェブ情報をもとに判定した前記スタイル情報を前記スタイル情報とすること、
を特徴とする、請求項1または2に記載の生成支援装置。
- [請求項4] 前記情報取得部は、前記要素画像のアップロードまたは選択を受け付け、前記結果画像の枠における前記要素画像の配置により前記位置情報を取得すること、
を特徴とする、請求項2に記載の生成支援装置。
- [請求項5] 前記情報取得部は、前記生成情報を前記ユーザによるチャット形式での入力により取得すること、
を特徴とする、請求項1または2に記載の生成支援装置。
- [請求項6] 前記情報取得部は、前記チャット形式での入力を取得する場合に、前記生成情報として必要となる情報の示唆を前記ユーザに提示すること、
を特徴とする、請求項5に記載の生成支援装置。

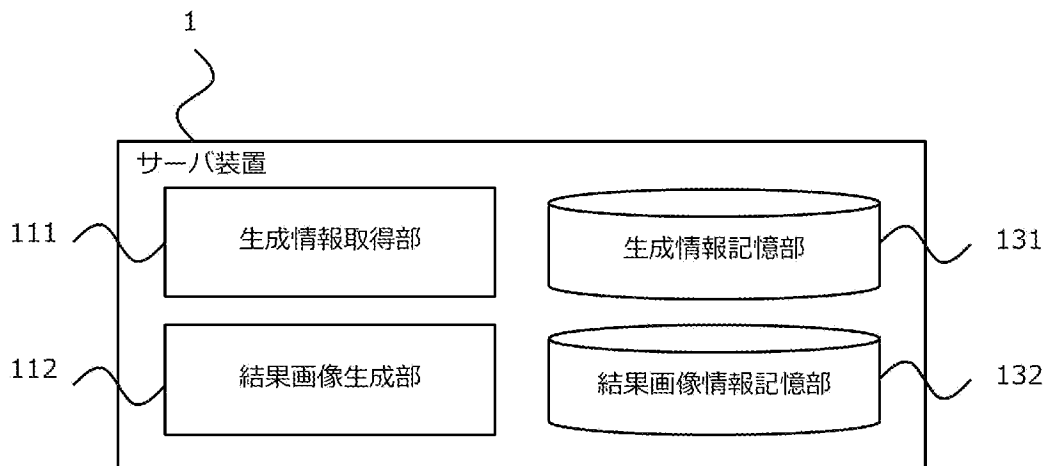
[図1]



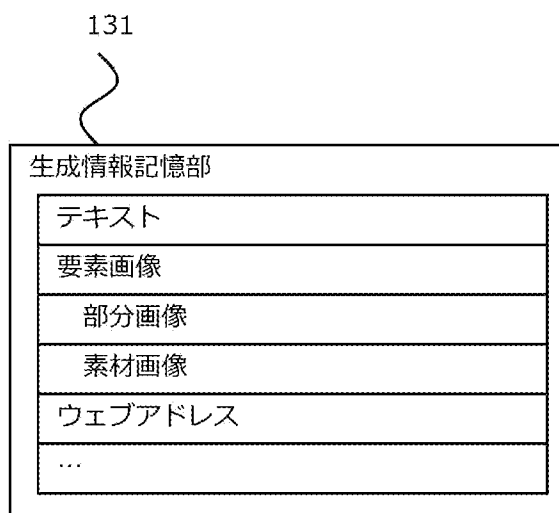
[図2]



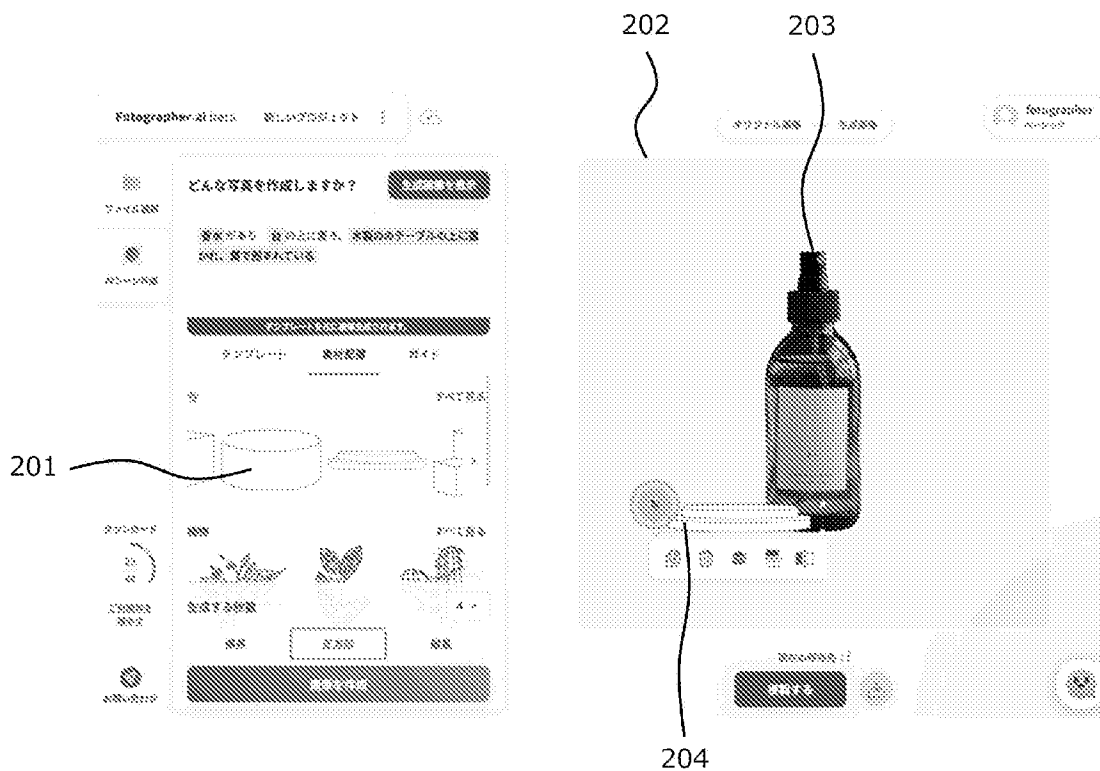
[図3]



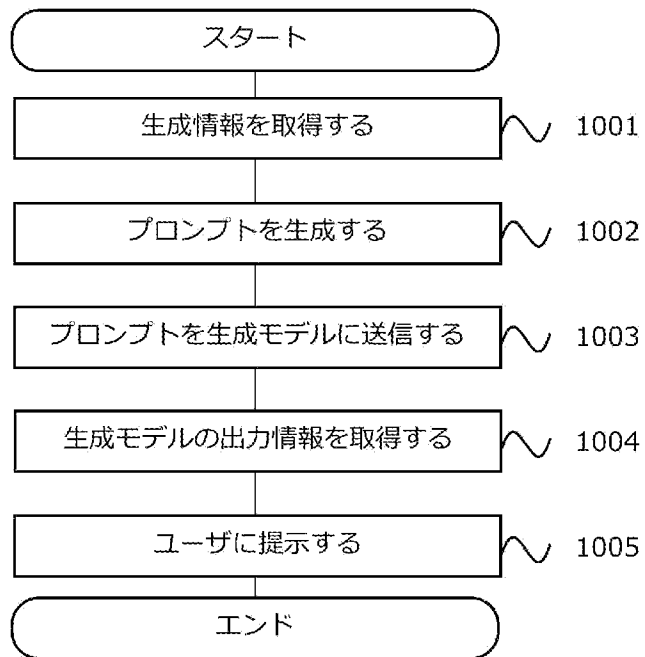
[図4]



[図5]



[図6]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2024/022627

A. CLASSIFICATION OF SUBJECT MATTER		
G06T 11/80(2006.01)i FI: G06T11/80 A		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) G06T11/00 - G06T11/80		
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 1971-2024 Registered utility model specifications of Japan 1996-2024 Published registered utility model applications of Japan 1994-2024		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EC業界を大きく変える画像生成AIサービスの起業から調達、開発の裏側, YouTube [online] [video]. ANOBAKA CH. 26 July 2023, [retrieved on 18 December 2023], Internet: <URL: https://www.youtube.com/watch?v=_WVOISuN-KA >, 1:19-4:45, non-official translation (The start-up to procurement and behind scenes of development of an image generation AI service that make a vast change in the EC industry) 1:19-4:45	1-2, 4-8
A	大谷大, ChatGPTに日本語入力すると画像生成AI用のプロンプトを出力するプラグイン「Photorealistic」が便利すぎる, デレイマニア [online], 25 May 2023, [retrieved on 18 December 2023], Internet: <URL: https://delaymania.com/202305/websevice/chatgpt-plugin-photorealistic/ >, (DELAYMANIA), non-official translation (OTANI, Dai, The plugin, "Photorealistic" which outputs a prompt for image generation AI when entering Japanese into ChatGPT, is extremely useful)	1-8
A	伊藤倫太郎, 人工知能 2023年の人工知能のトレンドについて, Rad Fan. メディカルアイ, 31 March 2023, vol. 21, no. 4, pp. 78-80, (ITO, Rintaro), non-official translation (About trends in artificial intelligence 2023, MEDICAL EYE)	1-8
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input 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 "D" document cited by the applicant in the international application "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 18 July 2024		Date of mailing of the international search report 30 July 2024
Name and mailing address of the ISA/JP Japan Patent Office (ISA/JP) 3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915 Japan		Authorized officer Telephone No.

A. 発明の属する分野の分類（国際特許分類（IPC）） G06T 11/80(2006.01)i FI: G06T11/80 A		
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）） G06T11/00 - G06T11/80		
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 - 2024年 日本国実用新案登録公報 1996 - 2024年 日本国登録実用新案公報 1994 - 2024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	EC業界を大きく変える画像生成AIサービスの起業から調達、開発の裏側, YouTube [online] [video], ANOBAKA CH, 2023.07.26, [検索日 2023.12.18], インターネット: <URL: https://www.youtube.com/watch?v=WV0ISuN-KA>, 1:19~4:45 1:19~4:45	1-2, 4-8
A	大谷大, ChatGPTに日本語入力すると画像生成AI用のプロンプトを出力するプラグイン「Photorealistic」が便利すぎる, ディレイマニア [online], 2023.05.25, [検索日 2023.12.18], インターネット: <URL: https://delaymania.com/202305/websevice/chatgpt-plugin-photorealistic/>	1-8
A	伊藤倫太郎, 人工知能 2023年の人工知能のトレンドについて, Rad Fan, メディカルアイ, 2023.03.31, 第21巻, 第4号, p. 78-80	1-8
<input type="checkbox"/> C欄の続きにも文献が列挙されている。 <input 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 “D” 国際出願で出願人が先行技術文献として記載した文献 “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 18.07.2024	国際調査報告の発送日 30.07.2024	
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） 鈴木 肇 5P 2374 電話番号 03-3581-1101 内線 3539	