

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7669162号
(P7669162)

(45)発行日 令和7年4月28日(2025.4.28)

(24)登録日 令和7年4月18日(2025.4.18)

(51)国際特許分類		F I	
G 0 6 T	7/90 (2017.01)	G 0 6 T	7/90 B
G 0 6 F	40/58 (2020.01)	G 0 6 F	40/58
G 0 9 G	5/02 (2006.01)	G 0 9 G	5/02 B
G 0 9 G	5/22 (2006.01)	G 0 9 G	5/22 6 6 0 Z
G 0 9 G	5/377(2006.01)	G 0 9 G	5/377
請求項の数 20 (全18頁)			
(21)出願番号	特願2021-40014(P2021-40014)	(73)特許権者	516014409
(22)出願日	令和3年3月12日(2021.3.12)		ライン プラス コーポレーション
(65)公開番号	特開2021-163482(P2021-163482 A)		L I N E P l u s C o r p o r a t i o n
(43)公開日	令和3年10月11日(2021.10.11)		大韓民国 1 3 5 9 1 キョンギード ソ
審査請求日	令和6年3月5日(2024.3.5)		ンナムーシ ブンダンーグ ファンセウル
(31)優先権主張番号	10-2020-0038295		ーロ 3 6 0 ボンーギル 4 2 トウェン
(32)優先日	令和2年3月30日(2020.3.30)		ディース フロア
(33)優先権主張国・地域又は機関	韓国(KR)	(74)代理人	100107766
			弁理士 伊東 忠重
		(74)代理人	100070150
			弁理士 伊東 忠彦
		(74)代理人	100135079
			弁理士 宮崎 修
		(72)発明者	ベク スンフン
最終頁に続く			

(54)【発明の名称】 画像内の文字色と背景色を抽出して提供する方法、システム、およびコンピュータプログラム

(57)【特許請求の範囲】

【請求項1】

コンピュータシステムが実行する方法であって、

前記コンピュータシステムは、メモリに含まれるコンピュータ読み取り可能な命令を実行するように構成された少なくとも1つのプロセッサを含み、

前記方法は、

前記少なくとも1つのプロセッサが、与えられた画像内で文字が含まれた第1領域を検出する段階、

前記少なくとも1つのプロセッサが、前記第1領域で前記文字を代表する文字代表色と前記第1領域の背景を代表する背景代表色を抽出する段階、および

前記少なくとも1つのプロセッサが、前記文字に対する翻訳結果が含まれた第2領域を前記画像の上にオーバーレイさせるが、このとき、前記第2領域の文字色と背景色として前記文字代表色と前記背景代表色を適用する段階

を含み、

前記抽出する段階は、

前記第1領域の各ピクセルに対するヒストグラムを抽出する段階、

前記第1領域の色分布に基づき文字基本色を抽出する段階、および

前記第1領域の各ピクセルに対する前記ヒストグラムと前記文字基本色との距離値に基づき前記文字代表色を選定する段階、

を含む、方法。

【請求項 2】

前記検出する段階は、

OCRを利用して前記第1領域を検出する、請求項1に記載の方法。

【請求項 3】

前記第1領域で前記文字を代表する前記文字代表色と前記第1領域の背景を代表する前記背景代表色を抽出する段階は、

前記第1領域の前記色分布に基づき背景基本色を抽出する段階、および

前記第1領域の各ピクセルに対する前記ヒストグラムと前記背景基本色との距離値に基づき前記背景代表色を選定する段階

を含む、請求項1に記載の方法。

10

【請求項 4】

前記第1領域で前記文字を代表する前記文字代表色と前記第1領域の背景を代表する前記背景代表色を抽出する段階は、

画像透視変換アルゴリズムを利用して前記第1領域を一定の形態に変換する段階

をさらに含む、請求項1に記載の方法。

【請求項 5】

前記第1領域で前記文字を代表する前記文字代表色と前記第1領域の背景を代表する前記背景代表色を抽出する段階は、

前記ヒストグラムを大きさ順に整列する段階

をさらに含む、請求項1に記載の方法。

20

【請求項 6】

前記ヒストグラムを抽出する段階は、

前記第1領域の各ピクセルに対して色成分の組み合わせによる3Dヒストグラムを抽出する段階

を含む、請求項1に記載の方法。

【請求項 7】

前記ヒストグラムを抽出する段階は、

前記第1領域のすべてのピクセル値を特定の倍数に量子化する段階

を含む、請求項1に記載の方法。

【請求項 8】

前記文字基本色を抽出する段階と前記背景基本色を抽出する段階は、

K-平均アルゴリズムを利用して前記第1領域の前記色分布を2つの色にクラスタリングする段階、

前記第1領域のフレーム領域を対象に前記2つの色の色ごとにピクセルの数を合算する段階、および

前記2つの色のうちでピクセルの数が多い色を前記背景基本色として選定し、残りの色を前記文字基本色として選定する段階

を含む、請求項3に記載の方法。

30

【請求項 9】

前記文字代表色を選定する段階と前記背景代表色を選定する段階は、

前記文字基本色との距離値が閾値以下のピクセルの色を前記文字代表色として選定する段階、および

前記背景基本色との距離値が前記閾値以下のピクセルの色を前記背景代表色として選定する段階

を含む、請求項3に記載の方法。

40

【請求項 10】

前記文字代表色を選定する段階と前記背景代表色を選定する段階は、

前記閾値以下のピクセルが存在しない場合、前記閾値を増加させる段階

を含む、請求項9に記載の方法。

【請求項 11】

50

前記方法は、

前記少なくとも1つのプロセッサが、画像オブジェクトの代表色を抽出し、前記画像オブジェクトに関連するUIを構成する色として適用する段階
をさらに含む、請求項1に記載の方法。

【請求項12】

前記方法は、

前記少なくとも1つのプロセッサが、アイコンの代表色を抽出し、背景テーマ色として適用する段階
をさらに含む、請求項1に記載の方法。

【請求項13】

前記方法は、

前記少なくとも1つのプロセッサが、動画に対して以前フレームと現在フレームそれぞれから代表色を抽出し、前記以前フレームの代表色と前記現在フレームの代表色との差によって場面の変わり目を検知する段階
をさらに含む、請求項1に記載の方法。

【請求項14】

請求項1～13のうちのいずれか一項に記載の方法を前記コンピュータシステムに実行させる、コンピュータプログラム。

【請求項15】

コンピュータシステムであって、

メモリに含まれるコンピュータ読み取り可能な命令を実行するように実現される少なくとも1つのプロセッサ

を含み、

前記少なくとも1つのプロセッサは、

与えられた画像内で文字が含まれた第1領域を検出する文字検出部、

前記第1領域で前記文字を代表する文字代表色と前記第1領域の背景を代表する背景代表色を抽出する色抽出部、および

前記文字に対する翻訳結果が含まれた第2領域を前記画像の上にオーバーレイさせるが、このとき、前記第2領域の文字色と背景色として前記文字代表色と前記背景代表色を適用して提供する翻訳結果提供部

を含み、

前記色抽出部は、

前記第1領域の各ピクセルに対するヒストグラムを抽出し、

前記第1領域の色分布に基づき文字基本色を抽出し、

前記第1領域の各ピクセルに対する前記ヒストグラムと前記文字基本色との距離値に基づき前記文字代表色を選定する、

コンピュータシステム。

【請求項16】

前記色抽出部は、

前記第1領域の前記色分布に基づき背景基本色を抽出し、

前記第1領域の各ピクセルに対する前記ヒストグラムと前記背景基本色との距離値に基づき前記背景代表色を選定する

請求項15に記載のコンピュータシステム。

【請求項17】

前記色抽出部は、

前記第1領域の各ピクセルに対して色成分の組み合わせによる3Dヒストグラムを抽出する

請求項15に記載のコンピュータシステム。

【請求項18】

前記色抽出部は、

10

20

30

40

50

前記第 1 領域のすべてのピクセル値を特定の倍数に量子化する

請求項 15 に記載のコンピュータシステム。

【請求項 19】

前記色抽出部は、

K-平均アルゴリズムを利用して前記第 1 領域の色分布を 2 つの色にクラスタリングし、

前記第 1 領域のフレーム領域を対象に前記 2 つの色の色ごとにピクセルの数を合算し、

前記 2 つの色のうちでピクセルの数が多き色を前記背景基本色として選定し、残りの色

を前記文字基本色として選定する

請求項 16 に記載のコンピュータシステム。

【請求項 20】

前記色抽出部は、

前記文字基本色との距離値が閾値以下のピクセルの色を前記文字代表色として選定し、

前記背景基本色との距離値が前記閾値以下のピクセルの色を前記背景代表色として選定

する

請求項 16 に記載のコンピュータシステム。

【発明の詳細な説明】

【技術分野】

【0001】

以下の説明は、画像内の文字検出および翻訳機能を提供する技術に関する。

【背景技術】

【0002】

一般的に、文字認識技術は、カメラが撮影した画像に含まれる文字を背景から分離して特徴を抽出した後、予め格納された文字の特徴パターンと比較し、最も類似するパターンの文字を認識結果として提供する。

【0003】

このような一般的な文字認識技術は、限定された範囲の文字や数字を使用する、例えば、自動車のナンバープレートや郵便番号などの認識には高い文字認識率を示し、実用化されている。

【0004】

一方、世界には多様な言語が存在するが、他国の言語を理解するには限界があるため、理解できない言語を理解可能な言語に翻訳して提供する自動翻訳システムの研究が活発化している。

【0005】

例えば、特許文献 1（公開日 2020 年 1 月 30 日）には、外国語で記載された画像に対して翻訳サービスを提供する技術が開示されている。

【先行技術文献】

【特許文献】

【0006】

【文献】韓国公開特許第 10-2020-0009205 号公報

【発明の概要】

【発明が解決しようとする課題】

【0007】

画像に対して文字検出および翻訳機能を提供するにあたり、翻訳結果領域を該当の画像にオーバーレイ（overlay）させて提供する。

【0008】

画像内の文字検出領域から文字色と背景色を抽出し、抽出された文字色と背景色を翻訳結果領域の文字と背景に適用する。

【課題を解決するための手段】

【0009】

コンピュータシステムが実行する方法であって、前記コンピュータシステムは、メモリ

10

20

30

40

50

に含まれるコンピュータ読み取り可能な命令を実行するように構成された少なくとも1つのプロセッサを含み、前記方法は、前記少なくとも1つのプロセッサが、与えられた画像内で文字が含まれた第1領域を検出する段階、前記少なくとも1つのプロセッサが、前記第1領域で前記文字を代表する文字代表色と前記第1領域の背景を代表する背景代表色を抽出する段階、および前記少なくとも1つのプロセッサが、前記文字に対する翻訳結果が含まれた第2領域を前記画像の上にオーバーレイさせるが、このとき、前記第2領域の文字色と背景色として前記文字代表色と前記背景代表色を適用する段階を含む方法を提供する。

【0010】

一側面によると、前記検出する段階は、OCR (optical character reader) を利用して前記第1領域を検出してよい。

10

【0011】

他の側面によると、前記抽出する段階は、前記第1領域の各ピクセルに対するヒストグラムを抽出する段階、前記第1領域の色分布に基づいて文字基本色と背景基本色を抽出する段階、前記第1領域の各ピクセルに対する前記ヒストグラムと前記文字基本色との距離値に基づき前記文字代表色を選定する段階、および前記第1領域の各ピクセルに対する前記ヒストグラムと前記背景基本色との距離値に基づき前記背景代表色を選定する段階を含んでよい。

【0012】

また他の側面によると、前記抽出する段階は、画像透視変換 (image perspective transform) アルゴリズムを利用して前記第1領域を一定の形態に変換する段階をさらに含んでよい。

20

【0013】

また他の側面によると、前記抽出する段階は、前記ヒストグラムを大きさ順に整列する段階をさらに含んでよい。

【0014】

また他の側面によると、前記ヒストグラムを抽出する段階は、前記第1領域の各ピクセルに対して色成分 (color channel) の組み合わせによる3Dヒストグラムを抽出する段階を含んでよい。

【0015】

また他の側面によると、前記ヒストグラムを抽出する段階は、前記第1領域のすべてのピクセル値を特定の倍数に量子化 (quantization) する段階を含んでよい。

30

【0016】

また他の側面によると、前記文字基本色と前記背景基本色を抽出する段階は、K-平均 (mean) アルゴリズムを利用して前記第1領域の色分布を2つの色にクラスタリングする段階、前記第1領域のフレーム領域を対象に前記2つの色の色ごとにピクセルの数を合算する段階、および前記2つの色のうちでピクセルの数が多い色を前記背景基本色として選定し、残りの色を前記文字基本色として選定する段階を含んでよい。

【0017】

また他の側面によると、前記文字代表色と前記背景代表色を選定する段階は、前記文字基本色との距離値が閾値 (threshold) 以下のピクセルの色を前記文字代表色として選定する段階、および前記背景基本色との距離値が前記閾値以下のピクセルの色を前記背景代表色として選定する段階を含んでよい。

40

【0018】

また他の側面によると、前記文字代表色と前記背景代表色を選定する段階は、前記閾値以下のピクセルが存在しない場合、前記閾値を増加させる段階を含んでよい。

【0019】

また他の側面によると、前記方法は、前記少なくとも1つのプロセッサが、画像オブジェクトの代表色を抽出し、前記画像オブジェクトと関連するUIを構成する色として適用する段階をさらに含んでよい。

50

【0020】

また他の側面によると、前記方法は、前記少なくとも1つのプロセッサが、アイコンの代表色を抽出し、背景テーマ色として適用する段階をさらに含んでよい。

【0021】

さらに他の側面によると、前記方法は、前記少なくとも1つのプロセッサが、動画の以前フレームと現在フレームそれぞれから代表色を抽出し、前記以前フレームの代表色と前記現在フレームの代表色との差によって場面の変わり目を検知する段階をさらに含んでよい。

【0022】

前記方法を前記コンピュータシステムに実行させるために非一時的なコンピュータ読み取り可能な記録媒体に記録される、コンピュータプログラムを提供する。

10

【0023】

コンピュータシステムであって、メモリに含まれるコンピュータ読み取り可能な命令を実行するように実現される少なくとも1つのプロセッサを含み、前記少なくとも1つのプロセッサは、与えられた画像内で文字が含まれた第1領域を検出する文字検出部、前記第1領域で前記文字を代表する文字代表色と前記第1領域の背景を代表する背景代表色を抽出する色抽出部、および前記文字に対する翻訳結果が含まれた第2領域を前記画像の上にオーバーレイさせるが、このとき、前記第2領域の文字色と背景色として前記文字代表色と前記背景代表色を適用して提供する翻訳結果提供部を含む、コンピュータシステムを提供する。

20

【図面の簡単な説明】

【0024】

【図1】本発明の一実施形態における、ネットワーク環境の例を示した図である。

【図2】本発明の一実施形態における、電子機器およびサーバの内部構成を説明するためのブロック図である。

【図3】本発明の一実施形態における、サーバのプロセッサが含むことのできる構成要素の例を示した図である。

【図4】本発明の一実施形態における、サーバが実行することのできる方法の例を示したフローチャートである。

【図5】本発明の一実施形態における、画像内の文字検出および翻訳機能に対する使用シナリオの例を示した図である。

30

【図6】本発明の一実施形態における、色抽出過程の具体的な例を示した図である。

【図7】本発明の一実施形態における、ヒストグラム抽出過程を説明するための例示図である。

【図8】本発明の一実施形態における、基本色選定過程を説明するための例示図である。

【図9】本発明の一実施形態における、基本色選定過程を説明するための例示図である。

【図10】本発明の一実施形態における、基本色選定過程を説明するための例示図である。

【図11】本発明の一実施形態における、ヒストグラム整列過程を説明するための例示図である。

【図12】本発明の一実施形態における、代表色選定過程を説明するための例示図である。

40

【発明を実施するための形態】

【0025】

以下、本発明の実施形態について、添付の図面を参照しながら詳しく説明する。

【0026】

本発明の実施形態は、画像内の文字検出および翻訳機能を提供する技術に関する。

【0027】

本明細書で具体的に開示される事項を含む実施形態は、画像に対して文字検出および翻訳機能を提供するにあたり、文字検出領域で抽出された文字色と背景色からなる翻訳結果領域を該当の画像にオーバーレイさせて提供することにより、原本画像と最も類似する色が適用されるようになり、翻訳結果領域に対する違和感を最小限にすることができる。

50

【0028】

図1は、本発明の一実施形態における、ネットワーク環境の例を示した図である。図1のネットワーク環境は、複数の電子機器110、120、130、140、複数のサーバ150、160、およびネットワーク170を含む例を示している。このような図1は、発明の説明のための一例に過ぎず、電子機器の数やサーバの数が図1のように限定されることはない。

【0029】

複数の電子機器110、120、130、140は、コンピュータシステムによって実現される固定端末や移動端末であってよい。複数の電子機器110、120、130、140の例としては、スマートフォン、携帯電話、ナビゲーション、PC (personal computer)、ノート型PC、デジタル放送用端末、PDA (Personal Digital Assistant)、PMP (Portable Multimedia Player)、タブレット、ゲームコンソール、ウェアラブルデバイス、IoT (Internet of Things) デバイス、VR (Virtual Reality) デバイス、AR (Augmented Reality) デバイスなどがある。一例として、図1では、電子機器110の例としてスマートフォンを示しているが、本発明の実施形態において、電子機器110は、実質的に無線または有線通信方式を利用し、ネットワーク170を介して他の電子機器120、130、140および／またはサーバ150、160と通信することのできる多様な物理的なコンピュータシステムのうちの1つを意味してよい。

【0030】

通信方式が限定されることはなく、ネットワーク170が含むことのできる通信網（一例として、移動通信網、有線インターネット、無線インターネット、放送網、衛星網など）を利用する通信方式だけではなく、機器間の近距離無線通信が含まれてもよい。例えば、ネットワーク170は、PAN (personal area network)、LAN (local area network)、CAN (campus area network)、MAN (metropolitan area network)、WAN (wide area network)、BBN (broadband network)、インターネットなどのネットワークのうちの1つ以上の任意のネットワークを含んでよい。さらに、ネットワーク170は、バスネットワーク、スターネットワーク、リングネットワーク、メッシュネットワーク、スターバスネットワーク、ツリーまたは階層的ネットワークなどを含むネットワークトポロジのうちの任意の1つ以上を含んでもよいが、これらに限定されることはない。

【0031】

サーバ150、160それぞれは、複数の電子機器110、120、130、140とネットワーク170を介して通信して命令、コード、ファイル、コンテンツ、サービスなどを提供する1つ以上のコンピュータ装置によって実現されてよい。例えば、サーバ150は、ネットワーク170を介して接続した複数の電子機器110、120、130、140に第1サービスを提供するシステムであってよく、サーバ160もネットワーク170を介して接続した複数の電子機器110、120、130、140に第2サービスを提供するシステムであってよい。より具体的な例として、サーバ150は、複数の電子機器110、120、130、140においてインストールされて実行されるコンピュータプログラムであるアプリケーションを通じ、該当のアプリケーションが目的とするサービス（一例として、画像ベースの翻訳サービスなど）を第1サービスとして複数の電子機器110、120、130、140に提供してよい。他の例として、サーバ160は、上述したアプリケーションのインストールおよび実行のためのファイルを複数の電子機器110、120、130、140に配布するサービスを第2サービスとして提供してよい。

【0032】

図2は、本発明の一実施形態における、電子機器およびサーバの内部構成を説明するためのブロック図である。図2では、電子機器に対する例として電子機器110の内部構成

およびサーバ１５０の内部構成について説明する。また、他の電子機器１２０、１３０、１４０やサーバ１６０も、上述した電子機器１１０またはサーバ１５０と同一または類似の内部構成を有してよい。

【００３３】

電子機器１１０およびサーバ１５０は、メモリ２１１、２２１、プロセッサ２１２、２２２、通信モジュール２１３、２２３、および入力／出力インタフェース２１４、２２４を含んでよい。メモリ２１１、２２１は、非一時的なコンピュータ読み取り可能な記録媒体であって、RAM (random access memory)、ROM (read only memory)、ディスクドライブ、SSD (solid state drive)、フラッシュメモリ (flash memory) などのような永続的大容量記録装置を含んでよい。ここで、ROM、SSD、フラッシュメモリ、ディスクドライブのような永続的大容量記録装置は、メモリ２１１、２２１とは区分される別の永続的記録装置として電子機器１１０やサーバ１５０に含まれてもよい。また、メモリ２１１、２２１には、オペレーティングシステムと、少なくとも１つのプログラムコード（一例として、電子機器１１０においてインストールされて実行されるブラウザや、特定のサービスの提供のために電子機器１１０にインストールされたアプリケーションなどのためのコード）が記録されてよい。このようなソフトウェア構成要素は、メモリ２１１、２２１とは別のコンピュータ読み取り可能な記録媒体からロードされてよい。このような別のコンピュータ読み取り可能な記録媒体は、フロッピー（登録商標）ドライブ、ディスク、テープ、DVD／CD-ROMドライブ、メモリカードなどのコンピュータ読み取り可能な記録媒体を含んでよい。他の実施形態において、ソフトウェア構成要素は、コンピュータ読み取り可能な記録媒体ではない通信モジュール２１３、２２３を通じてメモリ２１１、２２１にロードされてもよい。例えば、少なくとも１つのプログラムは、開発者またはアプリケーションのインストールファイルを配布するファイル配布システム（一例として、上述したサーバ１６０）がネットワーク１７０を介して提供するファイルによってインストールされるコンピュータプログラム（一例として、上述したアプリケーション）に基づいてメモリ２１１、２２１にロードされてよい。

【００３４】

プロセッサ２１２、２２２は、基本的な算術、ロジック、および入出力演算を実行することにより、コンピュータプログラムの命令を処理するように構成されてよい。命令は、メモリ２１１、２２１または通信モジュール２１３、２２３によって、プロセッサ２１２、２２２に提供されてよい。例えば、プロセッサ２１２、２２２は、メモリ２１１、２２１のような記録装置に記録されたプログラムコードにしたがって受信される命令を実行するように構成されてよい。

【００３５】

通信モジュール２１３、２２３は、ネットワーク１７０を介して電子機器１１０とサーバ１５０とが互いに通信するための機能を提供してもよいし、電子機器１１０および／またはサーバ１５０が他の電子機器（一例として、電子機器１２０）または他のサーバ（一例として、サーバ１６０）と通信するための機能を提供してもよい。一例として、電子機器１１０のプロセッサ２１２がメモリ２１１のような記録装置に記録されたプログラムコードにしたがって生成した要求が、通信モジュール２１３の制御にしたがってネットワーク１７０を介してサーバ１５０に伝達されてよい。これとは逆に、サーバ１５０のプロセッサ２２２の制御にしたがって提供される制御信号や命令、コンテンツ、ファイルなどが、通信モジュール２２３とネットワーク１７０を経て電子機器１１０の通信モジュール２１３を通じて電子機器１１０に受信されてよい。例えば、通信モジュール２１３を通じて受信されたサーバ１５０の制御信号や命令、コンテンツ、ファイルなどは、プロセッサ２１２やメモリ２１１に伝達されてよく、コンテンツやファイルなどは、電子機器１１０がさらに含むことのできる記録媒体（上述した永続的記録装置）に記録されてよい。

【００３６】

入力／出力インタフェース２１４は、入力／出力装置２１５とのインタフェースのため

10

20

30

40

50

の手段であってよい。例えば、入力装置は、キーボード、マウス、マイクロフォン、カメラなどの装置を、出力装置は、ディスプレイ、スピーカ、触覚フィードバックデバイスなどのような装置を含んでよい。他の例として、入力／出力インタフェース214は、タッチスクリーンのように入力と出力のための機能が1つに統合された装置とのインタフェースのための手段であってもよい。入力／出力装置215は、電子機器110と1つの装置で構成されてもよい。また、サーバ150の入力／出力インタフェース224は、サーバ150に接続するかサーバ150が含むことのできる入力または出力のための装置（図示せず）とのインタフェースのための手段であってよい。より具体的な例として、電子機器110のプロセッサ212がメモリ211にロードされたコンピュータプログラムの命令を処理するにあたり、サーバ150や電子機器120が提供するデータを利用して構成されるサービス画面やコンテンツが、入力／出力インタフェース214を通じてディスプレイに表示されてよい。

10

【0037】

また、他の実施形態において、電子機器110およびサーバ150は、図2の構成要素よりも多くの構成要素を含んでもよい。しかし、大部分の従来技術の構成要素を明確に図に示す必要はない。例えば、電子機器110は、上述した入力／出力装置215のうちの少なくとも一部を含むように実現されてもよいし、トランシーバ、GPS（Global Positioning System）モジュール、カメラ、各種センサ、データベースなどのような他の構成要素をさらに含んでもよい。より具体的な例として、電子機器110がスマートフォンである場合、一般的にスマートフォンが含んでいる加速度センサやジャイロセンサ、カメラモジュール、物理的な各種ボタン、タッチパネルを利用したボタン、入力／出力ポート、振動のための振動器などのような多様な構成要素が、電子機器110にさらに含まれるように実現されてよい。

20

【0038】

以下では、画像内の文字色と背景色を抽出して提供する方法およびシステムの具体的な実施形態について説明する。

【0039】

図3は、本発明の一実施形態における、サーバのプロセッサが含むことのできる構成要素の例を示したブロック図であり、図4は、本発明の一実施形態における、サーバが実行することのできる方法の例を示したフローチャートである。

30

【0040】

本実施形態に係るサーバ150は、画像ベースの翻訳サービスを提供するプラットフォームの役割を担うものである。サーバ150には、コンピュータで実現された翻訳結果提供システムが構成されてよい。サーバ150は、クライアント（client）である複数の電子機器110、120、130、140を対象に画像ベースの翻訳サービスを提供するものであり、電子機器110、120、130、140上にインストールされた専用アプリケーションやサーバ150と関連するウェブ／モバイルサイトに接続することにより、画像に対する文字検出および翻訳機能を提供してよい。特に、サーバ150は、画像内の文字検出領域で抽出された文字色と背景色からなる翻訳結果領域を該当の画像にオーバーレイさせて提供してよい。

40

【0041】

サーバ150のプロセッサ222は、図4に係る翻訳結果提供方法を実行するための構成要素として、図3に示すように、文字検出部310、色抽出部320、および翻訳結果提供部330を含んでよい。実施形態によって、プロセッサ222の構成要素は、選択的にプロセッサ222に含まれても除外されてもよい。また、実施形態によって、プロセッサ222の構成要素は、プロセッサ222の機能の表現のために分離されても併合されてもよい。

【0042】

このようなプロセッサ222およびプロセッサ222の構成要素は、図4の翻訳結果提供方法が含む段階S410～S430を実行するようにサーバ150を制御してよい。例

50

例えば、プロセッサ 222 およびプロセッサ 222 の構成要素は、メモリ 221 が含むオペレーティングシステムのコードと、少なくとも 1 つのプログラムのコードとによる命令 (instruction) を実行するように実現されてよい。

【0043】

ここで、プロセッサ 222 の構成要素は、サーバ 150 に記録されたプログラムコードが提供する命令にしたがってプロセッサ 222 によって実行される、互いに異なる機能 (different functions) の表現であってよい。例えば、サーバ 150 が画像から文字を検出するように上述した命令にしたがってサーバ 150 を制御するプロセッサ 222 の機能的表現として、文字検出部 310 が利用されてよい。

【0044】

プロセッサ 222 は、サーバ 150 の制御と関連する命令がロードされたメモリ 221 から必要な命令を読み取ってよい。この場合、前記読み取られた命令は、プロセッサ 222 が以下で説明する段階 S410～S430 を実行するように制御するための命令を含んでよい。以下で説明する段階 S410～S430 は、図 4 に示した順とは異なる順に実行されることもあるし、段階 S410～S430 のうちの一部が省略されたり追加の過程が含まれたりすることもある。

【0045】

以下で説明する実施形態では、図 4 の翻訳結果提供方法がサーバ 150 側で実行されるものと説明しているが、これに限定されてはならず、翻訳結果提供方法が含む段階 S410～S430 のうちの少なくとも一部が電子機器 110 で実行されることも可能である。

【0046】

図 4 を参照すると、段階 S410 で、文字検出部 310 は、電子機器 110 がアップロードした画像を対象に、画像内から文字に該当する領域 (以下、「文字検出領域」とする) を検出してよい。プロセッサ 222 は、画像ベースの文字翻訳機能を提供するものであって、OCR (optical character reader) のような文字認識技術を利用して画像内から翻訳対象となる文字部分を検出してよい。文字検出部 310 は、画像内の文字部分を四角形の枠領域で指定してよく、このとき、文字検出領域に対する画像上の座標値を含む文字枠情報を出力してよい。例えば、文字枠情報は、左上座標、右上座標、左下座標、右下座標を含んでよい。

【0047】

他の例として、上述した文字検出機能が電子機器 110 で実行されるときに、文字検出部 310 は、電子機器 110 から、画像とともに文字検出結果として文字枠情報を受信してよい。

【0048】

段階 S420 で、色抽出部 320 は、文字検出領域から、文字を代表する色 (以下、「文字代表色」とする) と背景を代表する色 (以下、「背景代表色」とする) を抽出してよい。色抽出部 320 は、画像内の特定の領域、すなわち、文字検出領域に対する一連の画像プロセス過程により、最適な文字代表色と背景代表色を導き出してよい。色抽出部 320 には画像と該当の画像内の文字検出領域の座標が入力され、色抽出対象領域として該当の座標の領域を調査しながら文字代表色と背景代表色を抽出してよい。具体的な色抽出過程については、以下で詳しく説明する。

【0049】

段階 S430 で、翻訳結果提供部 330 は、画像内の文字に対する翻訳結果が含まれた領域 (以下、「翻訳結果領域」とする) を該当の画像の上にオーバーレイさせて提供してよい。翻訳結果提供部 330 は翻訳エンジンを含んでよく、翻訳エンジンを利用して、画像から検出された文字を事前に定められた翻訳言語に翻訳してよい。このとき、翻訳エンジンは、統計による機械翻訳またはニューラルネットワークによる機械翻訳を提供するものであってよく、ディープラーニング (deep learning) などの機械学習 (machine learning) 方式で学習して形成されたものであってよい。実施形態によっては、サーバ 150 と連動可能な他のサーバ (一例として、サーバ 160) で

10

20

30

40

50

翻訳を実行することも可能であるが、この場合、画像から検出された文字に対する翻訳の要請がサーバ160に伝達されてよく、サーバ160はサーバ160に含まれた翻訳エンジンによって翻訳を行った後、翻訳結果を翻訳結果提供部330に提供してよい。翻訳結果提供部330は、原本画像の上に翻訳結果領域をオーバーレイさせてよく、これにより、電子機器110のユーザは、翻訳結果領域が含まれた画像を確認することができる。

【0050】

特に、翻訳結果提供部330は、画像の上に翻訳結果領域をオーバーレイさせて表示するにあたり、段階S420で抽出された文字代表色と背景代表色を翻訳結果領域の文字色と背景色に適用してよい。翻訳結果提供部330は、画像内の文字検出領域の文字を代表する色と背景を代表する色を、画像の上にオーバーレイさせる翻訳結果領域の文字色と背景色として利用してよい。

10

【0051】

一例として、図5を参照すると、色抽出部320には、画像500に対する画像ファイルと文字検出領域501の座標値を含む文字枠情報が入力されてよく、このとき、文字枠情報に基づいて画像500内の文字検出領域501を調査し、一連の画像処理によって文字代表色と背景代表色を抽出してよい。これにより、翻訳結果提供部330は、画像500の上にオーバーレイされる翻訳結果領域502の文字色と背景色を、文字検出領域501から抽出された文字代表色と背景代表色で表示してよい。画像500にオーバーレイされる翻訳結果領域502は、文字および背景色と同じように、その位置と大きさも文字検出領域501と同じように構成されてよい。

20

【0052】

したがって、翻訳結果提供部330は、原本画像500と最も類似する（文字と背景）色を翻訳結果領域502に適用することにより、翻訳結果領域502に対する違和感を最小限にすることができる。

【0053】

他の例として、上述した翻訳結果提供機能が電子機器110で実行されるときには、電子機器110は、サーバ150から、翻訳結果とともに文字代表色と背景代表色を受信して翻訳結果領域を画像の上にオーバーレイさせるが、翻訳結果領域に文字代表色と背景代表色を適用して表示してよい。

【0054】

図6は、本発明の一実施形態における、色抽出過程の具体的な例を示した図である。

30

【0055】

図6を参照すると、色抽出部320には、画像500に対して、画像ファイル601と文字検出領域501の座標値を含む文字枠情報602とが入力されてよい。

【0056】

色抽出部320は、画像ファイル601から色抽出のための領域を探索するものであって、文字枠情報602に基づいて文字検出領域501を探索してよい（S61）。ユーザが入力した画像ファイル601と文字枠情報602が色抽出部320に入力され、文字枠情報602に対応する文字検出領域501を色抽出が必要な関心領域として設定してよい。このとき、色抽出部320は、画像透視変換（*image perspective transform*）アルゴリズムを利用して文字検出領域501を長方形に変換してよい。透視変換アルゴリズムは、傾きや捩じれのある領域を所望する形状（例えば、長方形）に変換するコンピュータビジョンアルゴリズムに該当するものである。

40

【0057】

色抽出部320は、文字検出領域501を対象に3Dヒストグラムを抽出してよい（S62）。従来の2Dヒストグラムは、それぞれの色成分（R、G、B）に対して計算を行うため、色の組み合わせの分布を把握するには不適であるが、3Dヒストグラムは、色の組み合わせが1つの要素となり、その個数が総合される。例えば、図7に示すように、色抽出部320は、文字検出領域501のすべてのピクセルに対して色の組み合わせに対するヒストグラムを抽出してよい。

50

【0058】

実施形態によっては、色抽出部320は、3Dヒストグラムを抽出するときの計算量を減らすために、すべてのピクセル値を特定の倍数に量子化（quantization）して処理してよい。例えば、R：123、G：234、B：5であるピクセルのヒストグラムを求めるとき、R：123、G：234、B：5をR：124、G：236、B：8のように4の倍数単位に変更してよい。これにより、量子化を経る場合にヒストグラムを抽出するための計算量と時間を短縮することができる。

【0059】

再び図6を参照すると、色抽出部320は、文字検出領域501内に分布するすべての色のうちから優勢となる2つの基本色を抽出してよい（S63）。一例として、図8を参照すると、色抽出部320は、K-平均（mean）アルゴリズムを利用して、文字検出領域501に該当する原本画像の色分布を2つの色にクラスタリングしてよい。クラスタリング結果として文字検出領域501の原本画像と同じ大きさの画像が生成され、該当の画像には一種のラベル（label）値が含まれる。2つの色にクラスタリングする場合、1または2のラベルが含まれる。

【0060】

色抽出部320は、2つの基本色のうちの1つを背景基本色として選定し、残りの1つを文字基本色として選定してよい。一例として、図9を参照すると、色抽出部320は、文字検出領域501をなす四面のフレーム領域91を検索しながら、ラベルが1であるピクセルの数とラベルが2であるピクセルの数をそれぞれ合算する。このとき、ピクセルの1つの大きさのフレーム領域91を検索することはもちろん、事前に定められた大きさ、すなわち、2つ以上のピクセルの大きさのフレーム領域91を検索することも可能である。文字検出領域501をなす四面のフレーム領域91を検索することの他に、四面のうちの一部のフレーム領域（例えば、左上、左下、右上、右下など）を検索することも可能である。色抽出部320は、合算されたピクセルの数がより多いラベルの色を背景基本色として選定し、残りを文字基本色として選定してよい。文字検出領域501のフレーム領域91を検索して背景基本色を選定する理由は、フレーム領域91には文字に該当するピクセルよりも背景に該当するピクセルの数が多くと予測されるためである。フレーム領域91に最も多く分布する色を背景基本色として選定する方法の他にも、フレーム領域91内の事前に定められた特定の位置のピクセルの色を背景基本色として選定することも可能である。例えば、図10に示すように、フレーム領域91のうちの左枠の最初のピクセル102の色を背景基本色として選定してもよい。また、フレーム領域91のうちの右枠の最初のピクセル、左枠の最後のピクセル、右枠の最後のピクセル、上枠の中央のピクセル、下枠の中央のピクセルなどの色を背景基本色として選定してもよい。

【0061】

再び図6を参照すると、色抽出部320は、段階62で抽出されたヒストグラムを整理してよい（S64）。図11に示すように、色抽出部320は、3Dヒストグラムを大きさ順に整理してよく、このとき、整理されたヒストグラムは、文字色用ヒストグラムと背景色用ヒストグラムとして利用されてよい。

【0062】

再び図6を参照すると、色抽出部320は、段階S63で抽出された基本色と段階S64で整理されたヒストグラムに基づき、背景代表色と文字代表色を選定してよい（S65）。色抽出部320は、段階S64で整理されたヒストグラムと段階S63で抽出された基本色を使用して色距離（color distance）値を求め、これに基づいて最善の色を選択してよい。一例として、色距離を求める方法としてユークリッド方法を利用してよく、ユークリッド方法によってヒストグラムの色の組み合わせと基本色との距離を求めてよい。

【0063】

図12に示すように、色抽出部320は、文字基本色と背景基本色それぞれに対し、ヒストグラムの整理順にヒストグラムの色の組み合わせと基本色との距離を計算する。色抽

10

20

30

40

50

出部 320 は、文字基本色との距離が最も短いか事前に定められた閾値 (threshold) (例えば、40) 以下のピクセルの色 121 を文字代表色として選定し、背景基本色との距離が最も短いか閾値以下のピクセルの色 122 を背景代表色として選定してよい。代表色の特性は、画像内で多くの色分布を示し、基本色と類似の色を有する。基本色との距離が閾値以下のピクセルが見つからない場合は閾値を増加させ、上述した過程を再度実行する。

【0064】

したがって、色抽出部 320 は、文字検出領域の原本画像に対するヒストグラムと支配的な基本色を利用することにより、該当の画像に最適な文字代表色と背景代表色を得ることができる。

【0065】

本発明に係る色抽出方法は、与えられた画像内の文字を翻訳して提供する環境の他にも、多様な環境での活用可能である。一例として、電子機器 110 で撮影中のリアルタイム映像内の文字を翻訳し、適した色を探索して翻訳結果の上にオーバーレイさせてよい。他の例として、画像オブジェクトの代表となる色を探索し、該当のオブジェクトと関連する UI (user interface) を構成する色として活用してよい。例えば、電子機器 110 の画面に表示されるアイコンの代表となる色を探索し、背景テーマ色などとして活用してよい。このとき、アイコンの代表となる 2 つの色により、アイコンの外郭領域と類似する色とそうでない色が抽出されてよい。また他の例として、本発明に係る色抽出方法は、動画内の場面の変わり目を検知するための用途として活用されてもよい。例えば、以前フレームと現在フレームそれぞれから代表色を抽出し、フレームの代表色の差が一定の大きさ以上になる場合に場面が切り換わったと判断してよい。さらに、代表色を抽出する方法によって場面の変わり目として判断される時点、付加コンテンツ (例えば、広告など) の挿入位置として設定してよい。

【0066】

このように、本発明の実施形態によると、画像内の文字に対する翻訳結果を該当の画像にオーバーレイさせて提供するにあたり、文字検出領域から文字代表色と背景代表色を抽出し、抽出された文字代表色と背景代表色を翻訳結果領域の文字と背景に適用することにより、画像上の翻訳結果領域に対する違和感を最小限にすることができる。

【0067】

上述した装置は、ハードウェア構成要素、ソフトウェア構成要素、および／またはハードウェア構成要素とソフトウェア構成要素との組み合わせによって実現されてよい。例えば、実施形態で説明された装置および構成要素は、プロセッサ、コントローラ、ALU (arithmetic logic unit)、デジタル信号プロセッサ、マイクロコンピュータ、FPGA (field programmable gate array)、PLU (programmable logic unit)、マイクロプロセッサ、または命令を実行して応答することができる様々な装置のように、1 つ以上の汎用コンピュータまたは特殊目的コンピュータを利用して実現されてよい。処理装置は、オペレーティングシステム (OS) および OS 上で実行される 1 つ以上のソフトウェアアプリケーションを実行してよい。また、処理装置は、ソフトウェアの実行に応答し、データにアクセスし、データを記録、操作、処理、および生成してもよい。理解の便宜のために、1 つの処理装置が使用されるとして説明される場合もあるが、当業者であれば、処理装置が複数の処理要素および／または複数種類の処理要素を含んでもよいことが理解できるであろう。例えば、処理装置は、複数のプロセッサまたは 1 つのプロセッサおよび 1 つのコントローラを含んでよい。また、並列プロセッサのような、他の処理構成も可能である。

【0068】

ソフトウェアは、コンピュータプログラム、コード、命令、またはこれらのうちの 1 つ以上の組み合わせを含んでもよく、思うままに動作するように処理装置を構成したり、独立的または集散的に処理装置に命令したりしてよい。ソフトウェアおよび／またはデータは、処理装置に基づいて解釈されたり、処理装置に命令またはデータを提供したりするた

10

20

30

40

50

めに、いかなる種類の機械、コンポーネント、物理装置、コンピュータ記録媒体または装置に具現化されてよい。ソフトウェアは、ネットワークによって接続されたコンピュータシステム上に分散され、分散された状態で記録されても実行されてもよい。ソフトウェアおよびデータは、1つ以上のコンピュータ読み取り可能な記録媒体に記録されてよい。

【0069】

実施形態に係る方法は、多様なコンピュータ手段によって実行可能なプログラム命令の形態で実現されてコンピュータ読み取り可能な媒体に記録されてよい。ここで、媒体は、コンピュータ実行可能なプログラムを継続して記録するものであっても、実行またはダウンロードのために一時記録するものであってもよい。また、媒体は、単一または複数のハードウェアが結合した形態の多様な記録手段または格納手段であってよく、あるコンピュータシステムに直接接続する媒体に限定されることはなく、ネットワーク上に分散して存在するものであってもよい。媒体の例としては、ハードディスク、フロッピー（登録商標）ディスク、および磁気テープのような磁気媒体、CD-ROMおよびDVDのような光媒体、フロッピカルディスク（*f l o p t i c a l d i s k*）のような光磁気媒体、およびROM、RAM、フラッシュメモリなどを含み、プログラム命令が記録されるように構成されたものであってよい。また、媒体の他の例として、アプリケーションを配布するアプリケーションストアやその他の多様なソフトウェアを供給または配布するサイト、サーバなどで管理する記録媒体または格納媒体が挙げられる。

【0070】

以上のように、実施形態を、限定された実施形態および図面に基づいて説明したが、当業者であれば、上述した記載から多様な修正および変形が可能であろう。例えば、説明された技術が、説明された方法とは異なる順序で実行されたり、かつ／あるいは、説明されたシステム、構造、装置、回路などの構成要素が、説明された方法とは異なる形態で結合されたりまたは組み合わされたり、他の構成要素または均等物によって対置されたり置換されたとしても、適切な結果を達成することができる。

【0071】

したがって、異なる実施形態であっても、特許請求の範囲と均等なものであれば、添付される特許請求の範囲に属する。

【符号の説明】

【0072】

- 222：プロセッサ
- 310：文字検出部
- 320：色抽出部
- 330：翻訳結果提供部

10

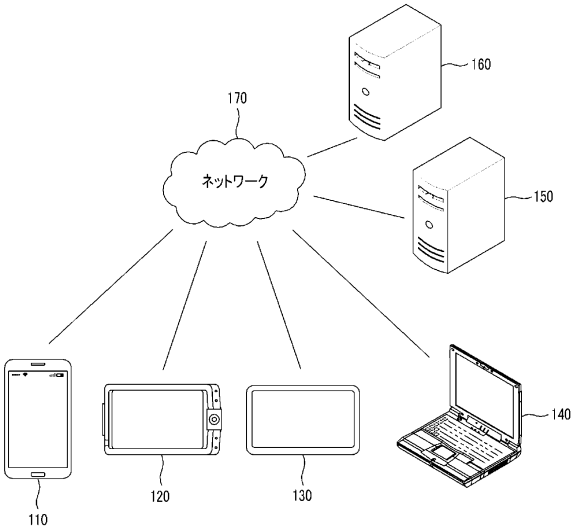
20

30

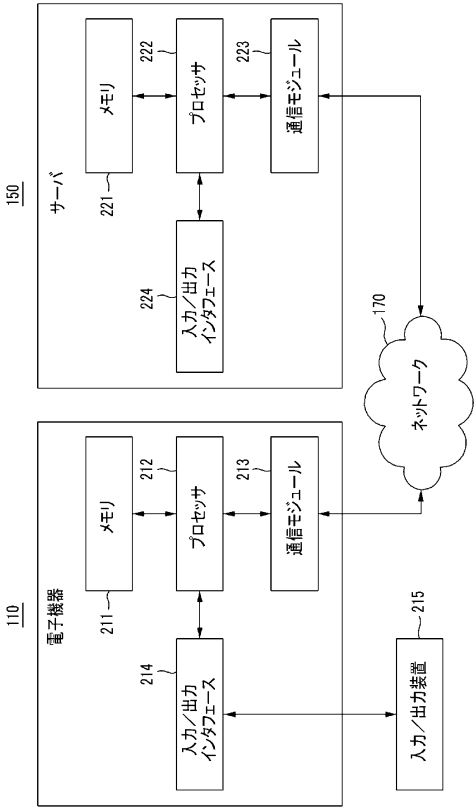
40

50

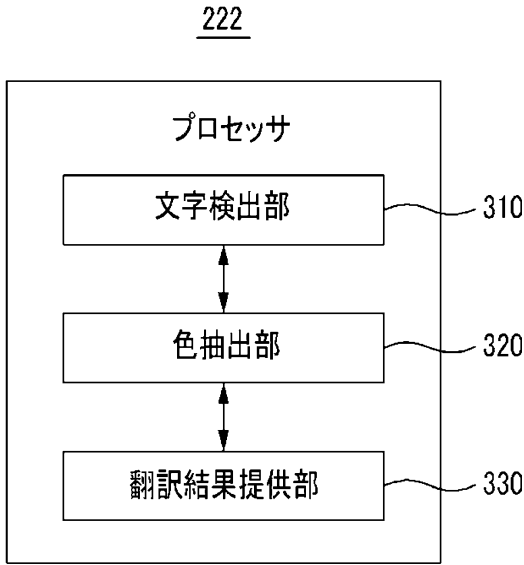
【図面】
【図 1】



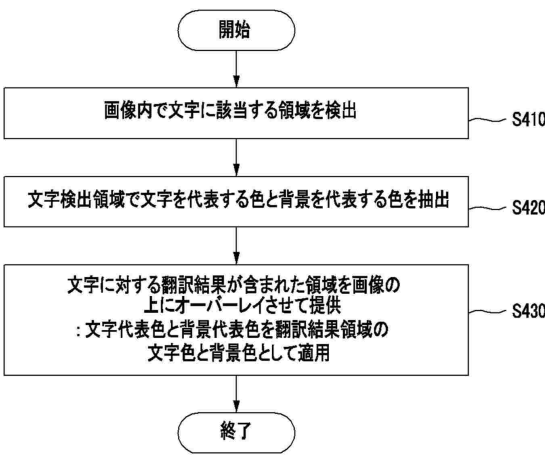
【図 2】



【図 3】



【図 4】



10

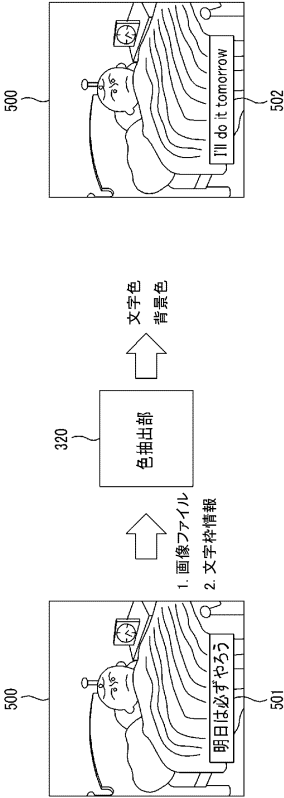
20

30

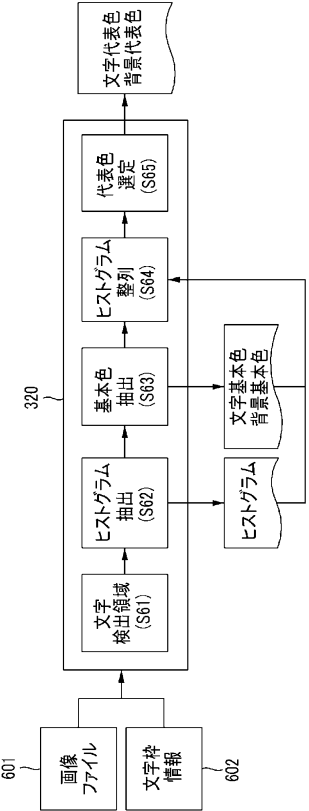
40

50

【図 5】



【図 6】



【図 7】

(R1, G1, B1) = 10234
(R2, G2, B2) = 65
(R3, G3, B3) = 6878
(R4, G4, B4) = 2356
(R5, G5, B5) = 446
...

【図 8】

原本画像(501)

Red	Red	Red	Light Green	Blue	Red
Light Red	Orange	Green	Blue	Blue	Red
Bold Red	Bold Red	Blue	Light Blue	Bold Blue	Blue

クラスタリング結果

1	1	1	1	2	1
1	1	2	2	2	1
1	1	2	2	2	2

1=RED(赤)に関連する色
2=BLUE(青)に関連する色

10

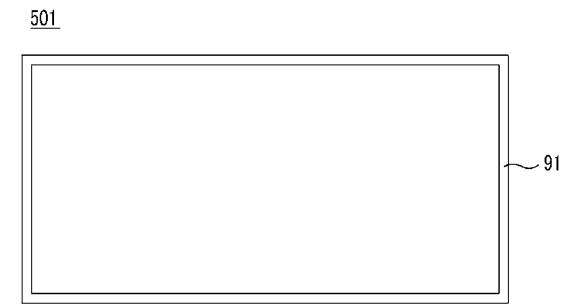
20

30

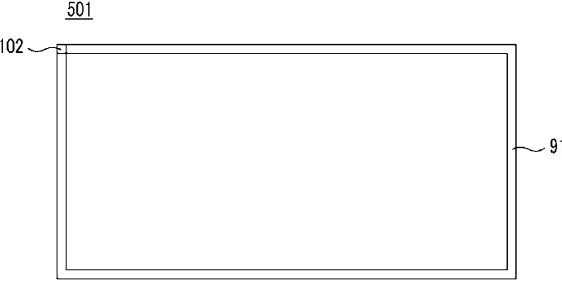
40

50

【図 9】

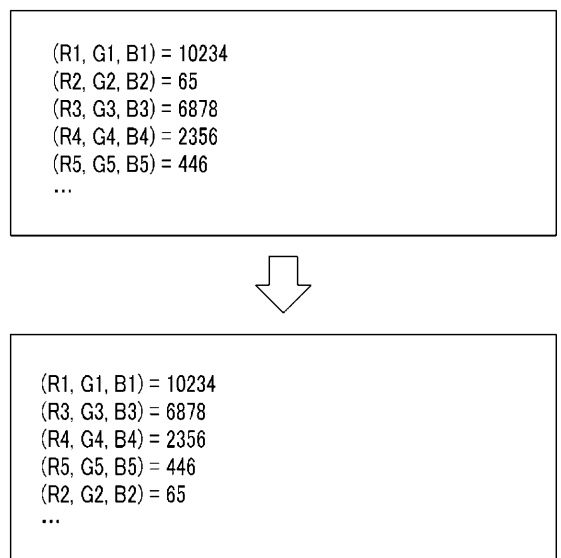


【図 10】

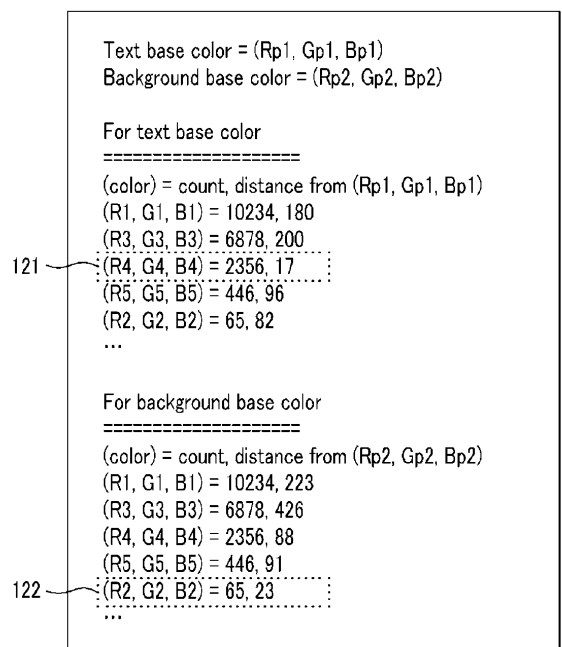


10

【図 11】



【図 12】



20

30

40

50

フロントページの続き

大韓民国 13591 京畿道城南市盆唐区ファンサエル路360番街42 11階 ラインプラス
株式会社内

審査官 菊池 伸郎

- (56)参考文献 特開2012-173785 (JP, A)
特開2008-065803 (JP, A)
特開2012-018490 (JP, A)
米国特許出願公開第2019/0266432 (US, A1)
米国特許出願公開第2018/0018795 (US, A1)
特開2015-069234 (JP, A)

- (58)調査した分野 (Int.Cl., DB名)
G06T 1/00-7/90
G06V 10/00-40/70
G06F 40/58
G09G 5/02
G09G 5/22
G09G 5/377