

(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
2017年1月19日(19.01.2017)



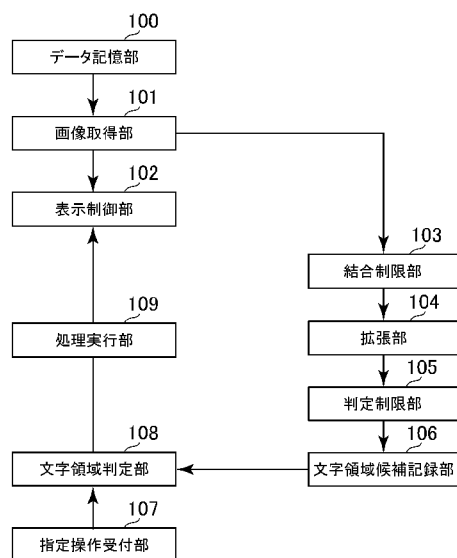
(10) 国際公開番号
WO 2017/010351 A1

- (51) 国際特許分類:
G06T 7/40 (2006.01)
- (21) 国際出願番号: PCT/JP2016/069884
- (22) 国際出願日: 2016年7月5日(05.07.2016)
- (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:
PCT/JP2015/069955 2015年7月10日(10.07.2015) JP
- (71) 出願人: 楽天株式会社(RAKUTEN, INC.) [JP/JP]; 〒1580094 東京都世田谷区玉川一丁目14番1号 Tokyo (JP).
- (72) 発明者: 劉 栄(LIU, Rong); 〒1580094 東京都世田谷区玉川一丁目14番1号 楽天株式会社内 Tokyo (JP).
- (74) 代理人: 特許業務法人はるか国際特許事務所 (HARUKA PATENT & TRADEMARK ATTORNEYS); 〒1600023 東京都新宿区西新宿三丁目1番4号 ウエル新都心ビル4階 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, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, 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).

[続葉有]

(54) Title: IMAGE PROCESSING DEVICE, IMAGE PROCESSING METHOD, AND PROGRAM

(54) 発明の名称: 画像処理装置、画像処理方法、及びプログラム



- 100 Data storage unit
101 Image acquisition unit
102 Display control unit
103 Concatenation restriction unit
104 Expansion unit
105 Determination restriction unit
106 Unit that records character region candidates
107 Designation operation accepting unit
108 Character region determination unit
109 Process execution unit

(57) Abstract: The purpose of the invention is to quickly specify a region of a group of characters included in an image. An image acquisition means (101) of this image processing device (1) acquires image data stored in a means (100) that stores the image data. An expansion means (104) expands each of a plurality of regions indicating an object included in an image indicated by the image data acquired by the image acquisition means (101), and concatenates the regions. A character region determination means (108) determines whether or not all or some of the regions expanded and concatenated by the expansion means (104) are character regions. A process execution means (109) executes a predetermined process on the basis of regions determined by the character region determination means (108) as being character regions.

(57) 要約: 画像に含まれる文字群の領域を迅速に特定する。画像処理装置(1)の画像取得手段(101)は、画像データを記憶する手段(100)に記憶された前記画像データを取得する。拡張手段(104)は、画像取得手段(101)により取得された画像データが示す画像に含まれる物体を示す複数の領域の各々を拡張して結合させる。文字領域判定手段(108)は、拡張手段(104)により拡張して結合した全部又は一部の領域が文字領域であるか否かを判定する。処理実行手段(109)は、文字領域判定手段(108)により文字領域であると判定された領域に基づいて、所定の処理を実行する。

WO 2017/010351 A1

添付公開書類:

— 国際調査報告 (条約第 21 条(3))

明 細 書

発明の名称：画像処理装置、画像処理方法、及びプログラム

技術分野

[0001] 本発明は、画像処理装置、画像処理方法、及びプログラムに関する。

背景技術

[0002] 従来、光学文字認識（OCR：Optical Character Recognition）のように、画像内の文字群の各文字に対して個別に文字認識を行う技術が知られている。例えば、特許文献1には、2値化画像内を縦方向及び横方向に繰り返し走査して、黒画素が存在しない白ラインで挟まれた領域を抽出して、各領域に対して文字が含まれるか否かを判定する技術が記載されている。

先行技術文献

特許文献

[0003] 特許文献1：特開平5－40849号公報

発明の概要

発明が解決しようとする課題

[0004] しかしながら、従来の技術では、例えば、画像内の文字群をひとかたまりとして特定したいような場合であっても、当該文字群に含まれる各文字の領域を抽出して、当該領域に対してOCRを利用して個別に文字認識をすることになるので、画像内の文字群の領域を特定するのに時間がかかっていた。

[0005] 本発明は上記課題に鑑みてなされたものであって、その目的は、画像に含まれる文字群の領域を迅速に特定することである。

課題を解決するための手段

[0006] 上記課題を解決するために、本発明に係る画像処理装置は、画像データを記憶する手段に記憶された前記画像データを取得する画像取得手段と、前記画像取得手段により取得された画像データが示す画像に含まれる物体を示す複数の領域の各々を拡張して結合させる拡張手段と、前記拡張手段により拡張して結合した全部又は一部の領域が文字領域であるか否かを判定する文字

領域判定手段と、前記文字領域判定手段により文字領域であると判定された領域に基づいて、所定の処理を実行する処理実行手段と、を含む。

[0007] 本発明に係る画像処理方法は、画像データを記憶する手段に記憶された前記画像データを取得する画像取得ステップと、前記画像取得ステップにより取得された画像データが示す画像に含まれる物体を示す複数の領域の各々を拡張して結合させる拡張ステップと、前記拡張ステップにより拡張して結合した全部又は一部の領域が文字領域であるか否かを判定する文字領域判定ステップと、前記文字領域判定ステップにより文字領域であると判定された領域に基づいて、所定の処理を実行する処理実行ステップと、を含む。

[0008] 本発明に係るプログラムは、画像データを記憶する手段に記憶された前記画像データを取得する画像取得手段、前記画像取得手段により取得された画像データが示す画像に含まれる物体を示す複数の領域の各々を拡張して結合させる拡張手段、前記拡張手段により拡張して結合した全部又は一部の領域が文字領域であるか否かを判定する文字領域判定手段、前記文字領域判定手段により文字領域であると判定された領域に基づいて、所定の処理を実行する処理実行手段、としてコンピュータを機能させる。

[0009] 本発明に係る情報記憶媒体は、上記記載のプログラムが記憶されたコンピュータ読み取り可能な情報記憶媒体である。

[0010] また、本発明の一態様によれば、前記画像処理装置は、前記画像に含まれる物体を示す領域の輪郭又は当該領域自身の大きさが基準以上である場合、当該領域と、前記拡張手段により拡張した領域と、が結合することを制限する結合制限手段を更に含む、ことを特徴とする。

[0011] また、本発明の一態様によれば、前記結合制限手段は、前記画像に含まれる物体を示す領域の輪郭又は当該領域自身の大きさが基準以上である場合、当該領域を背景色で塗りつぶすことによって、前記拡張手段により拡張した領域と結合することを制限する、ことを特徴とする。

[0012] また、本発明の一態様によれば、前記画像処理装置は、前記拡張手段により拡張した領域の輪郭又は当該領域自身の大きさが基準未満である場合、当

該領域が前記文字領域判定手段の判定対象となることを制限する判定制限手段を更に含む、ことを特徴とする。

[0013] また、本発明の一態様によれば、前記判定制限手段は、前記拡張手段により拡張した領域の輪郭又は当該領域自身の大きさが基準未満である場合、当該領域を背景色で塗りつぶすことによって、前記文字領域判定手段の判定対象となることを制限する、ことを特徴とする。

[0014] また、本発明の一態様によれば、前記画像処理装置は、前記拡張手段により拡張して結合した一部の領域の指定操作を受け付ける指定操作受付手段を更に含み、前記文字領域判定手段は、前記指定操作受付手段により受け付けられた指定操作により指定された一部の領域が文字領域であるか否かを判定する、ことを特徴とする。

[0015] また、本発明の一態様によれば、前記処理実行手段は、前記画像が表示された表示手段において、前記文字領域判定手段により文字領域であると判定された領域内を拡大して表示させる、ことを特徴とする。

発明の効果

[0016] 本発明によれば、画像に含まれる文字群の領域を迅速に特定することが可能になる。

図面の簡単な説明

[0017] [図1]本実施形態における電子書籍表示装置のハードウェア構成を示す図である。

[図2]電子書籍表示装置で実現される機能の一例を示す機能ブロック図である。

[図3]電子書籍画像の一例を示す図である。

[図4]領域の大きさの説明図である。

[図5]2値化され反転された電子書籍画像を示す図である。

[図6]図5に示す電子書籍画像から抽出された輪郭を示す図である。

[図7]基準以上の大きさの輪郭線を拡大した後の各輪郭線の状態を示す図である。

[図8]拡大した輪郭線が重畳された電子書籍画像を示す図である。

[図9]拡張部の処理内容の説明図である。

[図10]拡張部の処理内容の説明図である。

[図11]拡張部により領域が拡張された後の状態を示す図である。

[図12]輪郭の大きさが基準未満である領域が背景色で塗りつぶされた後の状態を示す図である。

[図13]電子書籍画像内の文字領域候補の一例を示す図である。

[図14]学習データを用いた文字領域の判定処理の説明図である。

[図15]文字群が拡大表示される様子を示す図である。

[図16]電子書籍表示装置において実行される処理の一例を示すフロー図である。

発明を実施するための形態

[0018] [1. 電子書籍表示装置のハードウェア構成]

以下、本発明に係る実施形態を図面に基づき詳細に説明する。図1は、本実施形態における電子書籍表示装置（画像処理装置の一例）のハードウェア構成を示す図である。電子書籍表示装置1は、ユーザが操作するコンピュータであり、例えば、携帯電話機（スマートフォンを含む）、携帯情報端末（タブレット型コンピュータを含む）、電子書籍リーダー端末、又はパーソナルコンピュータ等である。図1に示すように、電子書籍表示装置1は、制御部10、記憶部11、通信部12、操作部13、表示部14、及び入出力部15を含む。

[0019] 制御部10は、例えば、一又は複数のマイクロプロセッサを含む。制御部10は、記憶部11に記憶されたプログラムやデータに従って処理を実行する。記憶部11は、主記憶部及び補助記憶部を含む。例えば、主記憶部はRAMであり、補助記憶部は、ハードディスク又はソリッドステートドライブ等である。通信部12は、ネットワークカード等の通信インタフェースである。

[0020] 操作部13は、一般的な入力デバイスであり、例えば、タッチパネルやマ

ウス等のポインティングデバイスである。操作部 13 は、ユーザの操作内容を制御部 10 に伝達する。表示部 14 は、例えば、液晶表示部又は有機 EL 表示部等である。表示部 14 は、制御部 10 の指示に従って画面を表示する。入出力部 15 は、外部機器とのデータの入出力を行う入出力インタフェースである。例えば、入出力部 15 は、コンピュータ読み取り可能な情報記憶媒体（例えば、光ディスクやメモリカード等）からデータやプログラムを読み取る。

[0021] なお、記憶部 11 に記憶されるものとして説明するプログラム及びデータは、ネットワークに接続されたサーバコンピュータから、通信部 12 を介して記憶部 11 に供給されるようにしてもよいし、情報記憶媒体から、入出力部 15 を介して記憶部 11 に供給されるようにしてもよい。また、電子書籍表示装置 1 のハードウェア構成は、上記の例に限られず、種々のコンピュータのハードウェアを適用可能である。

[0022] [2. 電子書籍表示装置において実現される機能]

図 2 は、電子書籍表示装置 1 で実現される機能の一例を示す機能ブロック図である。図 2 に示すように、電子書籍表示装置 1 では、データ記憶部 100、画像取得部 101、表示制御部 102、結合制限部 103、拡張部 104、判定制限部 105、文字領域候補記録部 106、指定操作受付部 107、文字領域判定部 108、及び処理実行部 109 が実現される。データ記憶部 100 は、記憶部 11 を主として実現され、他の各機能は、制御部 10 を主として実現される。

[0023] [2-1. データ記憶部]

データ記憶部 100 は、電子書籍などの画像に関する各種データを記憶する。例えば、データ記憶部 100 は、電子書籍などの画像の画像データを記憶する。電子書籍は、表示部 14 に電子的に表示される書籍であればよく、例えば、漫画、絵本、又は雑誌である。電子書籍は、少なくとも文字を含む。電子書籍は、文字以外に、絵・写真・図形等を含むようにしてよい。なお、ここでの絵とは、人が手で描画した絵をスキャナで読み込んで電子化した

ものに限らず、人がコンピュータを使って描画したCGを含む意味である。画像データのデータ形式は、一般的に電子書籍で用いられているデータ形式を用いればよい。

[0024] 電子書籍は、複数のページにより構成されてもよいし、1ページのみから構成されてもよい。電子書籍が複数のページから構成される場合には、ページごとに画像データが存在してもよいし、これら複数ページをひとまとめにした1つの画像データのみが存在してもよい。また、電子書籍は、カラーであってもよいし、モノクロ又はグレースケールであってもよい。本実施形態では、複数ページから構成されるグレースケールの漫画を、電子書籍の一例として説明する。

[0025] なお、データ記憶部100は、画像データ以外のデータを記憶していてもよい。例えば、データ記憶部100は、画像データに付帯する付帯データを記憶してもよい。付帯データは、各ページのページ数等である。他にも例えば、データ記憶部100は、電子書籍リーダーのアプリケーション（ビューワアプリケーション）を記憶してもよい。また例えば、データ記憶部100は、後述する文字領域の判定のための学習データを記憶してもよい。

[0026] [2-2. 画像取得部]

画像取得部101は、画像データを記憶するデータ記憶部100に記憶された画像データを取得する。本実施形態のように、電子書籍が複数のページから構成される場合、画像取得部101は、これら全てのページの画像データを一度に取得してもよいし、一部のページの画像データのみを取得してもよい。画像取得部101が一部のページの画像データのみを取得する場合、画像取得部101は、表示対象のページの画像データのみを取得してもよいし、表示対象のページとその前後のページの画像データを取得してもよい。なお、表示対象となるのは、1ページだけであってもよいし、見開き等のように複数ページであってもよい。更に、本実施形態のように電子書籍が漫画の場合、1ページの中の1コマのみが表示対象となってもよい。本実施形態では、複数ページのうちの何れか1ページが表示対象になる場合を例に挙げ

て説明する。

[0027] [2-3. 表示制御部]

表示制御部102は、画像取得部101により取得された画像データに基づいて、電子書籍画像を表示部14に表示させる。図3は、電子書籍画像の一例を示す図である。本実施形態では、複数ページからなる漫画を電子書籍の一例として説明するので、図3に示すように、電子書籍画像20は、漫画の各ページである。電子書籍画像20は、1又は複数のコマ21を含む。コマ21は、ページ内で絵や文字が描画される区画である。図3の例では、電子書籍画像20は、2つのコマ21（これらをコマ21A、21Bとも記載する）を含む。図3の例では、各コマ21は、枠で囲われているが、枠で囲われていなくてもよい。

[0028] 各コマ21は、絵、吹き出し線、文字等の物体を含む。吹き出し線は、漫画のキャラクタの台詞を表す。キャラクタは、漫画の登場人物であり、人間以外にも動物・ロボット・架空の生物であってもよい。吹き出し線の内側には、キャラクタの台詞を表す1又は複数の文字が含まれている。以降、複数の文字のまとまりのことを文字群という。図3の例では、コマ21Aには、キャラクタ22A、吹き出し線23A、文字群24A、及び太陽25Aが描かれており、コマ21Bには、キャラクタ22B、22C、吹き出し線23B、文字群24B、及び食器25Bが描かれている。なお、キャラクタ22A～22C、太陽25A、及び食器25Bは、上記説明した絵の一例である。吹き出し線23Aは、キャラクタ22Aの台詞を示し、吹き出し線23Bは、キャラクタ22Cの台詞を示す。

[0029] なお、本実施形態のように表示対象が1ページの場合、表示制御部102は、図3のようにページの全体を表示させてもよいし、ページ内の一部のみを表示させてもよい。ページ内の一部のみを表示させる場合、表示制御部102は、ユーザが操作部13を用いて指示した部分を表示させることになる。

[0030] [2-4. 結合制限部]

結合制限部 103 は、電子書籍画像 20 に含まれる物体を示す領域の輪郭又は当該領域自身の大きさが基準以上である場合、当該領域と、後述する拡張部 104 により拡張した領域と、が結合することを制限する。例えば、文字群 24 A が拡張して吹き出し線 23 A に結合すると、文字群 24 A を文字のかたまりとして検出できなくなってしまう。このため、結合制限部 103 は、下記に説明するようにして、これらの結合を制限することで、文字群 24 A の検出の精度を高めるようにしている。

[0031] なお、本実施形態では、説明の簡略化のため、結合制限部 103、拡張部 104、判定制限部 105、及び文字領域候補記録部 106 による画像処理の対象となるのは、表示部 14 に表示された電子書籍画像 20 である場合を説明するが、表示部 14 に表示される前の電子書籍画像 20 が画像処理の対象となってもよい。この場合、表示部 14 に表示されたページの前後のページの電子書籍画像 20 が画像処理の対象となるようにしてよい。

[0032] 領域とは、電子書籍画像 20 のうち物体が描かれている部分（画素）であり、背景色（例えば、白）以外の色（例えば、黒又はグレー）の部分である。例えば、キャラクタ 22 A～22 C、吹き出し線 23 A、23 B、文字群 24 A、24 B、太陽 25 A、及び食器 25 B 等の物体が描かれている部分が、本実施形態でいう領域に相当する。なお、物体は、電子書籍画像 20 の背景以外の部分に表れている描画物といえる。

[0033] 輪郭は、領域を縁取ったものであり、本実施形態では、電子書籍表示装置 1 が線として表現したものを輪郭線と記載する。領域は、輪郭の内部、又は、輪郭と輪郭で挟まれた部分ということもできる。例えば、文字群 24 A であれば、「L」の文字は、輪郭の内部が領域であり、「O」の文字は、外周の輪郭と内周の輪郭とに挟まれた部分が領域である。

[0034] 輪郭の大きさは、例えば、輪郭の幅、輪郭線の長さ、輪郭の内側の面積等であってよいが、本実施形態では、輪郭の外接矩形の大きさを、輪郭の大きさとして用いる。領域の大きさは、例えば、領域の幅であってよいが、本実施形態では、領域の外接矩形の大きさを、領域の大きさとして用いる。

- [0035] 図4は、輪郭と領域の大きさの説明図である。図4の X_s 軸及び Y_s 軸は、スクリーン座標系の座標軸である。ここでは、領域の一例として吹き出し線23Aを挙げて、吹き出し線23Aの輪郭の大きさと、吹き出し線23A自身の大きさと、について説明する。図4では、吹き出し線23Aの外周輪郭線と内周輪郭線を破線で描き、吹き出し線23A自身を実線で描いている。なお、図4では、説明の都合上、実線と破線の間に間隔を設けているが、実際には、これらに間隔はない。
- [0036] 例えば、輪郭線の外接矩形は、輪郭線のうち、 X_s 座標の最小値と及び最大値の間であり、かつ、 Y_s 座標の最小値と最大値の間の領域である。図4に示す例では、外周輪郭線の外接矩形は、点 $P_{1A} \sim P_{4A}$ を頂点とする長方形なので、この長方形の大きさが外周輪郭線の大きさに相当する。一方、内周輪郭線の外接矩形は、点 $P_{1B} \sim P_{4B}$ を頂点とする長方形なので、この長方形の大きさが内周輪郭線の大きさに相当する。
- [0037] また例えば、吹き出し線23Aの外接矩形は、吹き出し線23Aが占める画素のうち、 X_s 座標の最小値と最大値の間であり、かつ、 Y_s 座標の最小値と最大値の間の領域である。吹き出し線23Aの外接矩形は、点 $P_{1C} \sim P_{4C}$ を頂点とする長方形なので、この長方形の大きさが吹き出し線23A自身の大きさに相当する。
- [0038] 結合制限部103は、輪郭の大きさと領域自身の大きさの何れを用いてもよいが、本実施形態では、輪郭の大きさを用いる場合を説明する。結合制限部103は、電子書籍画像20内の各領域の輪郭の外接矩形を抽出し、外接矩形の大きさが基準以上であるか否かを判定する。なお、大きさを示す数値としては、外接矩形の縦幅及び横幅の少なくとも一方が用いられてもよいし、外接矩形の面積が用いられてもよい。他にも例えば、外接矩形の対角線の長さが、大きさを示す数値として用いられてもよい。本実施形態では、大きさを示す数値として、外接矩形の縦幅と横幅の両方が用いられる場合を例に挙げて説明する。
- [0039] 大きさの基準を示す閾値 T_1 は、固定値であってもよいし、可変値であって

もよい。閾値 T_1 が固定値である場合には、データ記憶部100にその値が記憶されており、閾値 T_1 が可変値である場合には、データ記憶部100に記憶された数値をもとに計算されることになる。本実施形態では、閾値 T_1 が固定値である場合を説明する。例えば、結合制限部103は、外接矩形の縦幅と横幅がそれぞれ閾値 T_1 以上であるか否かを判定することによって、各輪郭の大きさが基準以上であるか否かを判定する。

[0040] 本実施形態では、結合制限部103は、領域の輪郭の大きさが基準以上である場合、当該領域を背景色で塗りつぶすことによって、後述する拡張部104により拡張した領域と結合することを制限する場合を説明する。背景色は、画像に対して指定された所定の色であり、例えば、画像内で占める割合が最も多い色である。例えば、背景色は、白又は黒である。なお、本実施形態における結合とは、ある領域が他の領域と接触又は重畳して1つの領域となることである。

[0041] 例えば、吹き出し線23Aの輪郭は比較的大きいので、結合制限部103は、吹き出し線23Aの輪郭を基準以上の大きさであると判定し、吹き出し線23Aを背景色で塗りつぶすことになる。その結果、吹き出し線23Aと、後述する拡張部104により拡張された文字群24Aと、が結合することを防止することができるようになっている。具体的には、本実施形態では、結合制限部103は、下記のように、(1)2値化反転処理、(2)輪郭抽出処理、(3)輪郭線拡大処理、(4)拡大輪郭線重畳処理を実行することによって、吹き出し線23A等を塗りつぶすことになる。

[0042] [2値化反転処理]

まず、結合制限部103は、電子書籍画像20に対して2値化処理と反転処理を実行する。図5は、2値化され反転された電子書籍画像20を示す図である。結合制限部103は、図3に示す電子書籍画像20の各画素の画素値と閾値に基づいて2値化処理を実行し、各画素を白又は黒に設定する。2値化処理自体は、公知の種々のアルゴリズムを適用可能であり、例えば、Pタイル法やモード法によって閾値を定めてもよい。結合制限部103は、2

値化した各画素の画素値を反転させ、白の画素を黒にして黒の画素を白にすることによって、図4に示す電子書籍画像20を、図5に示す状態にする。反転処理自体も、公知の種々の手法を適用可能であり、例えば、各色の画素値の最大値（8ビットグレースケールなら255）から画素値の現在値を引くことで反転後の画素値を計算してもよい。

[0043] [輪郭抽出処理]

結合制限部103は、2値化して反転させた電子書籍画像20（図5）に対して輪郭抽出処理を実行し、閉領域の輪郭を抽出する。図6は、図5に示す電子書籍画像20から抽出された輪郭線を示す図である。図6では、結合制限部103が抽出した輪郭線を破線で示している。輪郭線は、所定の太さの線（例えば、1ピクセル）であればよい。また、輪郭線の色は、予め定められた色であればよいが、ここでは黒とする。これは、2値化して反転させた電子書籍画像20（図5）内の吹き出し線23A、23Bが白で描かれているため、これを背景色である黒で塗りつぶすためである。

[0044] なお、輪郭抽出処理自体は、公知の種々の輪郭抽出アルゴリズムを適用可能であり、例えば、微分フィルタ、Prewittフィルタ、又はSobelフィルタ等を用いて画像内のエッジを検出することによって、各領域の輪郭を抽出するようにしてよい。また、図5に示す電子書籍画像20では、キャラクタ22Aの左側頭部と左肩がコマ21Aの輪郭に触れているので、図6に示すように、キャラクタ22Aとコマ21Aの輪郭線は一連のものになっている。

[0045] [輪郭線拡大処理]

結合制限部103は、輪郭抽出処理によって抽出した輪郭線のうち、基準以上の大きさの輪郭線を拡大する。結合制限部103は、各輪郭線の外接矩形を抽出して、その大きさが閾値 T_1 以上であるか否かを判定することになる。先述したように、本実施形態では、外接矩形の縦幅と横幅を大きさとして用いるので、結合制限部103は、外接矩形の縦幅と横幅がそれぞれ閾値 T_1 以上であるか否かを判定することによって、各輪郭線の大きさが基準以上であるか否かを判定する。

[0046] 結合制限部 103 は、基準以上の大きさの輪郭線を拡大する。輪郭線を拡大するとは、輪郭線を太くすることである。輪郭線の拡大の程度（線を太らせるピクセル数）は、固定値であってもよいし、可変値であってもよい。固定値である場合には、データ記憶部 100 にその値が記憶されており、可変値である場合には、データ記憶部 100 に記憶された数値をもとに計算されることになる。例えば、結合制限部 103 は、輪郭線の位置や輪郭線の大きさに基づいて、当該輪郭線の拡大の程度を計算してもよい。この場合、輪郭線が画像の中央付近にあるほど、その輪郭線が吹き出し線 23A、23B のものである蓋然性が高いので、輪郭線の拡大の程度を大きくして、より確実に塗りつぶせるようにしてもよい。また例えば、輪郭線が大きいほど、その輪郭線が吹き出し線 23A、23B のものである蓋然性が高いので、輪郭線の拡大の程度を大きくして、より確実に塗りつぶせるようにしてもよい。本実施形態では、拡大の程度が固定値であり、所定ピクセル分だけ線を太らせる場合を説明する。

[0047] 図 7 は、基準以上の大きさの輪郭線を拡大した後の各輪郭線の状態を示す図である。図 7 に示す例では、コマ 21A、21B の輪郭線、キャラクタ 22A～22C の一部の輪郭線、吹き出し線 23A、23B の輪郭線、及び太陽 25A の一部の輪郭線が所定ピクセル分だけ太くなっている。図 7 に示すように、結合制限部 103 が輪郭線を拡大させると、近くにある輪郭線同士が結合して 1 つの太い輪郭線になることがある。例えば、図 7 に示す例では、吹き出し線 23A、23B の外周の輪郭線と内周の輪郭線がそれぞれ拡大し、これらが互いに結合して 1 つの太い輪郭線となっている。

[0048] [拡大輪郭線重畳処理]

結合制限部 103 は、2 値化して反転させた電子書籍画像 20（図 5）に、拡大した輪郭線（図 7）を重畳させる。図 8 は、拡大した輪郭線が重畳された電子書籍画像 20 を示す図である。図 8 に示すように、拡大して太い 1 本になった輪郭線で吹き出し線 23A、23B を重畳したことによって、吹き出し線 23A、23B が黒く塗りつぶされている。このため、後述する処

理によって、拡張部 104 が文字群 24 A, 24 B を拡張したとしても、吹き出し線 23 A, 23 B に結合しないことになる。

[0049] 上記のようにして、吹き出し線 23 A, 23 B を背景色で塗りつぶすことができる。なお、吹き出し線 23 A, 23 B を背景色で塗りつぶす方法は、上記説明した方法に限られない。例えば、結合制限部 103 は、基準以上の大きさの輪郭線で挟まれた領域（例えば、吹き出し線 23 A, 23 B）を選択して、その選択範囲についてのみ、もう一度色の反転処理をしてもよい。このようにすれば、例えば、図 5 に示す吹き出し線 23 A, 23 B の色が白から黒に戻るので、吹き出し線 23 A, 23 B を黒く塗りつぶすことができる。他にも例えば、結合制限部 103 は、基準以上の大きさの輪郭線で挟まれた領域であり、かつ、当該輪郭線の間隔が一定距離未満である領域を、背景色で塗りつぶすようにしてもよい。更に、例えば、結合制限部 103 は、輪郭線で挟まれた領域が一定距離以上続いている区間を、背景色で塗りつぶすようにしてもよい。結合制限部 103 の処理対象となる領域に、このような条件を追加することによって、当該条件によって特定される領域が、吹き出し線 23 A, 23 B のような線が描かれた部分である蓋然性が高まる。このため、より確実に、吹き出し線 23 A, 23 B のような線の部分を塗りつぶし、文字群 24 A, 24 B と結合することを防止することができる。

[0050] また例えば、吹き出し線 23 A, 23 B と、文字群 24 A, 24 B と、を結合させないようにする方法は、吹き出し線 23 A, 23 B を背景色で塗りつぶす方法に限られない。例えば、結合制限部 103 は、基準以上の大きさの領域の位置（領域を構成する画素の位置）をデータ記憶部 100 に記録しておき、後述する拡張部 104 により拡張された領域が当該画素に含まれても、その画素は同じ領域とみなさないようにしてもよい。同じ領域とみなさないためには、例えば、同一領域を抽出するラベリング処理において、同じ番号を振らないようにすればよい。他にも例えば、結合制限部 103 は、基準以上の大きさの輪郭線の方角には、拡張部 104 が領域を拡張しないように制限をするようにしてもよい。このようにすることでも、例えば、文字群

24 A, 24 Bが吹き出し線23 A, 23 Bの方向には拡張しないので、吹き出し線23 A, 23 Bと、文字群24 A, 24 Bと、が結合することを防止することができる。

[0051] [2-5. 拡張部]

拡張部104は、画像取得部101により取得された画像データが示す電子書籍画像20に含まれる物体を示す複数の領域の各々を拡張して結合させる。なお、拡張部104は、図8に示す電子書籍画像20内の全ての領域を拡張してもよいし、結合制限部103が抽出した輪郭の大きさが基準以上の領域は拡張せず、輪郭の大きさが基準未満の領域を拡張するようにしてもよい。即ち、例えば、拡張部104は、キャラクタ22 A, 22 B, 22 Cの目・鼻・口等、文字群24 A, 24 Bの各文字、太陽25 Aの一部、食器25 Bのように、輪郭の大きさが基準未満のもののみを拡張してもよい。

[0052] なお、領域を拡張するとは、領域を太らせることであり、領域が占める面積を大きくすることである。領域の拡張の程度は、固定値であってもよいし、可変値であってもよい。固定値である場合には、データ記憶部100にその値が記憶されており、可変値である場合には、データ記憶部100に記憶された数値をもとに計算されることになる。例えば、拡張部104は、画像における領域の位置や大きさに基づいて、当該領域の拡張の程度を計算してもよい。この場合、領域が画像の中央付近にあるほど、その領域が文字である蓋然性が高いので、領域の拡張の程度を大きくして、より確実に他の文字と結合させるようにしてもよい。また、他の領域との距離が近いほど、その領域が文字群の1文字である蓋然性が高いので、領域の拡張の程度を大きくして、より確実に他の文字と結合させるようにしてもよい。また例えば、領域の大きさが一定範囲（文字と推測される程度の広さ）におさまっていれば、その領域が文字である蓋然性が高いので、領域の拡張の程度を大きくして、より確実に他の文字と結合させるようにしてもよい。本実施形態では、拡張の程度が固定値であり、所定ピクセル分だけ領域を太らせる場合を説明する。

[0053] 図9及び図10は、拡張部104の処理内容の説明図である。ここでは、拡張部104が、文字群24Aの各文字を拡張する場合を例に挙げて説明する。例えば、図9に示すように、拡張部104は、「H」を所定ピクセル分だけ拡張する。別の言い方をすれば、拡張部104は、「H」から所定距離以内にある画素（即ち、「H」の周囲の画素）を白で塗りつぶす。拡張部104は、文字群24Aの「HELLO! HOW ARE YOU?」の「H」以外の文字も、「H」と同様に所定ピクセル分だけ拡張する。拡張部104が文字群24Aの各文字を拡張させると、図10に示すように、文字群24Aの各文字が結合して、文字群24Aを1つの領域とすることができる。

[0054] 図11は、拡張部104により領域が拡張された後の状態を示す図である。図11に示すように、ここでは、文字群24A、24Bに含まれる各文字の領域が拡張して互いに結合し、1つの領域になっている。他にも例えば、食器25Bの各領域も拡張して互いに結合している。他の領域については、拡張はされたものの、付近に別の領域が無かったため結合していない。

[0055] [2-6. 判定制限部]

判定制限部105は、拡張部104により拡張した領域（図12）の輪郭又は当該領域自身の大きさが基準未満である場合、当該領域が文字領域判定部108の判定対象となることを制限する。なお、判定制限部105の説明でも、結合制限部103と同様、輪郭の大きさが用いられる場合を説明する。

[0056] 上記の基準を示す閾値 T_2 は、閾値 T_1 よりも小さくてよい。なお、閾値 T_2 が固定値であってもよいし、可変値であってもよい点については、閾値 T_1 と同様である。本実施形態では、外接矩形の縦幅と横幅を大きさとして用いるので、判定制限部105は、外接矩形の縦幅と横幅がそれぞれ閾値 T_2 以上であるか否かを判定することによって、拡張部104により拡張した領域の輪郭が基準未満であるか否かを判定する。

[0057] 本実施形態では、判定制限部105は、拡張部104により拡張した領域

の輪郭の大きさが基準未満である場合、当該領域を背景色で塗りつぶすことによって、文字領域判定部 108 の判定対象となることを制限する場合を説明する。図 12 は、輪郭の大きさが基準未満である領域が背景色で塗りつぶされた後の状態を示す図である。例えば、キャラクタ 22A, 22C の目・鼻・口等、太陽 25A の一部の輪郭が基準未満の大きさであるため、図 12 に示すように、判定制限部 105 は、これらを黒で塗りつぶしている。

[0058] なお、基準未満の大きさの領域を背景色で塗りつぶす以外の方法で、文字領域判定部 108 による判定処理の対象外とするようにしてもよい。例えば、判定制限部 105 は、基準未満の大きさの領域の位置をデータ記憶部 100 に記録しておき、ユーザがその領域を指定しても文字領域判定部 108 による判定処理を実行しないようにしてもよい。

[0059] [2-7. 文字領域候補記録部]

文字領域候補記録部 106 は、電子書籍画像 20 内の文字領域候補を示すデータをデータ記憶部 100 に記録する。文字領域候補は、文字領域判定部 108 による文字領域判定部 108 による判定対象となる領域である。別の言い方をすれば、文字領域候補は、文字が含まれている可能性のある領域である。本実施形態では、判定制限部 105 による処理が実行された後の電子書籍画像 20 (図 12) 内の白画素の領域が、文字領域候補となる。

[0060] 図 13 は、文字領域候補の一例を示す図である。文字領域候補記録部 106 は、図 12 に示す電子書籍画像 20 に対して輪郭抽出処理を実行して、白画素の領域の輪郭 (図 13 に破線で示す) を抽出する。輪郭抽出処理自体は、結合制限部 103 による処理と同様であってよい。ここでは、文字領域候補記録部 106 は、閉領域の輪郭を抽出するので、図 12 に示すキャラクタ 22A については輪郭を抽出しないことになる。文字領域候補記録部 106 は、抽出した輪郭線の位置を示すデータをデータ記憶部 100 に記録する。

[0061] なお、ここでは文字領域候補を識別するデータとして、輪郭線の位置を示すデータを例に挙げるが、文字領域候補を識別可能なデータであればよい。例えば、文字領域候補記録部 106 は、図 12 に示す電子書籍画像 20 に対

してラベリング処理を実行し、白画素が連続する画素に同じ番号を付与することによって、文字領域候補を識別するデータを記録するようにしてもよい。

[0062] [2-8. 指定操作受付部]

指定操作受付部107は、拡張部104により拡張して結合した一部の領域の指定操作を受け付ける。指定操作は、電子書籍画像20内の領域を指定するための操作であればよく、例えば、電子書籍画像20内の位置を指定する操作である。本実施形態では、操作部13により表示部14の表示画面内の位置を指定する操作が指定操作に相当する。操作部13により指定された位置を含む領域が、指定操作により指定されたことになる。本実施形態では、指定操作は、文字領域候補のうちの何れかを指定する操作といえる。

[0063] [2-9. 文字領域判定部]

文字領域判定部108は、拡張部104により拡張して結合した全部又は一部の領域が文字領域であるか否かを判定する。文字領域は、1つ以上の文字を含む画像内の領域である。例えば、文字領域判定部108による判定対象となるのは、文字領域候補記録部106により記録された文字領域候補の全てであってもよいし、一部のみであってもよい。例えば、文字領域判定部108は、指定操作受付部107により受け付けられた指定操作により指定された一部の領域が文字領域であるか否かを判定する。本実施形態では、指定操作受付部107は、文字領域候補の指定を受け付けるので、文字領域判定部108は、文字領域候補内に文字が含まれるか否かを判定することになる。なお、本実施形態では、文字領域判定部108は、図3に示す状態の電子書籍画像20を用いて判定処理を実行する場合を説明するが、使用する画像の状態に応じて文字領域の判定方法を変えるようにすれば、図5や図8に示す状態の電子書籍画像20を用いてもよい。

[0064] 例えば、文字パターンの学習データをデータ記憶部100に記憶させておき、文字領域判定部108は、文字領域候補と、学習データが示す文字パターンと、を比較することによって、文字領域候補が文字領域であるか否かを

判定する。学習データには、複数の判定項目が定義されており、例えば、下記のような判定項目によって文字領域であるかの判定が行われる。

[0065] 図14は、学習データを用いた文字領域の判定処理の説明図である。まず、第1の判定項目として、文字領域判定部108は、文字領域候補の全体の大きさが基準以上であるか否かを判定する。文字領域候補の大きさは、外接矩形の大きさを用いてもよいし、文字領域候補のピクセル数をカウントしてもよい。ここでは、外接矩形を用いる場合を説明する。この基準を示す閾値 T_3 は、閾値 T_1 、 T_2 と同じであってもよいし、異なってもよい。また、閾値 T_3 は、閾値 T_1 、 T_2 と同様、固定値であってもよいし、可変値であってもよい。文字領域判定部108は、外接矩形の大きさが基準以上である場合、文字領域と判定する。

[0066] 図14に示すように、文字領域判定部108は、外接矩形を n マス $\times$ m マス(n 、 m は、ともに2以上の整数。ここでは、 $n=m=10$ とする。)の矩形領域に区切る。そして、矩形領域の垂直ライン(図14では1列目~10列目の各ライン)及び水平ライン(図14では1行目~10行目の各ライン)のそれぞれについて、第2の判定項目~第5の判定項目について判定する。なお、全てのマスについて判定処理を行ってもよいし、一部のマス(例えば、外周の一部のマスを除いたマス)についてのみ判定処理を行ってもよい。

[0067] 第2の判定項目として、文字領域判定部108は、各ラインの白ピクセルの数が基準数以上であるか否かを判定する。この基準数は、固定値であってもよいし、可変値であってもよい。文字領域判定部108は、白ピクセルの数が基準数以上である場合、文字領域と判定する。第3の判定項目として、文字領域判定部108は、各ラインの画素を端から走査し、白画素と黒画素とが反転した数をカウントし、当該カウント数が基準数以上であるか否かを判定する。この基準数も、固定値であってもよいし、可変値であってもよい。文字領域判定部108は、カウント数が基準数以上である場合、文字領域と判定する。

[0068] 第4の判定項目として、文字領域判定部108は、各ラインの画素を端から走査し、最初に白の画素を発見した位置が所定の位置であるか否かを判定する。例えば、文字領域判定部108は、最初に白の画素を発見した位置が端から基準距離以内であれば、文字領域と判定する。第5の判定項目として、文字領域判定部108は、各ラインの画素を端から走査し、最後に白の画素を発見した位置が所定の位置であるか否かを判定する。例えば、文字領域判定部108は、最後に白の画素を発見した位置が端から基準距離以内であれば、文字領域と判定する。

[0069] 文字領域判定部108は、学習パターンに定義された第1の判定項目～第5の判定項目の全てについて文字領域であると判定された場合に、文字領域候補が文字領域であると判定するようにしてもよいし、所定数以上（例えば、3つ以上）の判定項目について文字領域であると判定された場合に、文字領域候補が文字領域であると判定するようにしてもよい。

[0070] なお、文字領域を判定する方法は、上記説明した方法に限られず、公知の種々の手法を適用可能である。例えば、文字領域判定部108は、文字の形状のパターンと文字領域候補とを比較して、互いの類似度を計算することによって、文字領域であるか否かを判定してもよい。他にも例えば、教師ありの機械学習手法（SVM：Support Vector Machine）を用いて、文字領域であるか否かを判定してもよい。この手法では、学習データを用意しておき、複数の項目をパラメータとして判定アルゴリズムに入力すると、判定アルゴリズムは、入力されたパラメータと学習データとを用いて、文字領域候補が文字領域であるか否かの出力を得ることができるようになっている。このような手法を用いることで、例えば、種々の形状の文字や絵が含まれる漫画のような画像であっても、より確実に文字領域を特定することができる。

[0071] [2-10. 処理実行部]

処理実行部109は、文字領域判定部108により文字領域であると判定された領域に基づいて、所定の処理を実行する。本実施形態では、所定の処理の一例として、画像を拡大する処理である場合を説明する。即ち、例えば

、処理実行部１０９は、電子書籍画像２０が表示された表示部１４において、文字領域判定部１０８により文字領域であると判定された領域内を拡大して表示させる。図１５は、文字群２４Ａが拡大表示される様子を示す図である。図１５に示すように、ユーザが文字群２４Ａに対応する文字領域候補を指定すると、文字領域判定部１０８により文字領域の判定処理が行われるので、処理実行部１０９は、文字領域候補内を切り出した拡大画像２６を表示させる。図１５の例では、処理実行部１０９は、図８に示す電子書籍画像２０の文字領域候補内を切り出して拡大させて表示させるので、拡大画像２６は、色が反転した状態で文字が表示されている。

[0072] なお、処理実行部１０９は、図３に示す電子書籍画像２０の文字領域候補内を切り出して拡大させて表示させてもよい。この場合、拡大画像２６における文字の色は反転しないことになる。また、拡大画像２６に切り出す範囲は、図１５のような長方形に限られず、楕円形等の任意の形状であってよい。また、拡大画像２６の表示位置は、ランダムに決定されてもよいし、ユーザが指定した位置に基づいて定まってもよい。また、拡大画像２６における文字の拡大率は、文字領域候補の大きさに基づいて決まってもよいし、所定の拡大率であってもよい。

[0073] [３．電子書籍表示装置において実行される処理]

図１６は、電子書籍表示装置１において実行される処理の一例を示すフロー図である。図１６に示す処理は、制御部１０が、記憶部１１に記憶されたプログラムに従って動作することによって実行される。本実施形態では、下記に説明する処理が実行されることにより、図２に示す機能ブロックが実現される。例えば、ユーザが操作部１３を用いて、プログラムの起動を指示した場合に、下記の処理が実行される。

[0074] 図１６に示すように、制御部１０は、記憶部１１に記憶されたプログラムを起動し（Ｓ１）、操作部１３からの入力に基づいて、ユーザが指定した電子書籍を特定する（Ｓ２）。例えば、Ｓ１において起動するプログラムは、電子書籍リーダーのアプリケーションである。Ｓ１においては、制御部１０

は、記憶部 11 に画像データが記憶された電子書籍の一覧を表示部 14 に表示させ、S2において、制御部 10 は、一覧の中からユーザにより指定された電子書籍を特定することになる。

[0075] 制御部 10 は、記憶部 11 を参照して、ユーザが指定した電子書籍の画像データを取得する (S3)。なお、記憶部 11 には、ユーザが保有する電子書籍の識別情報と、画像データと、が関連付けられて記憶されているものとする。制御部 10 は、S3 で取得した画像データに基づいて、電子書籍の 1 ページ目を表示部 14 に表示させる (S4)。S4 においては、制御部 10 は、画像データが示す電子書籍画像 20 (図 3) を表示部 14 に表示させることになる。

[0076] 制御部 10 は、電子書籍画像 20 に対して、2 値化処理と反転処理を実行する (S5)。なお、先述したように、表示部 14 に表示される電子書籍画像 20 に対して S5 ~ S11 の画像処理が実行される場合を説明するが、表示部 14 に表示される電子書籍画像 20 の前後のページの電子書籍画像 20 に対しても、予め S5 ~ S11 の画像処理を実行しておくようにしてもよい。S5 においては、制御部 10 は、2 値化して反転した電子書籍画像 20 (図 5) のデータを生成して記憶部 11 に記録する。

[0077] 制御部 10 は、S5 で生成した電子書籍画像 20 内に対して、輪郭抽出処理を実行する (S6)。S6 においては、制御部 10 は、電子書籍画像 20 内の輪郭線 (図 6) を示すデータを生成して記憶部 11 に記録する。

[0078] 制御部 10 は、S6 で抽出した輪郭線のうち、基準以上の大きさの輪郭線を拡大させる (S7)。S7 においては、制御部 10 は、拡大した輪郭線を示すデータ (図 7) を生成して記憶部 11 に記録する。なお、制御部 10 は、このデータを新たに生成するのではなく、S6 で記憶部 11 に記録した輪郭線のデータを更新するようにしてもよい。

[0079] 制御部 10 は、S5 で生成した電子書籍画像 20 に、S7 で拡大した輪郭線を重畳させる (S8)。S8 においては、制御部 10 は、重畳後の電子書籍画像 20 (図 8) を示すデータを生成して記憶部 11 に記録する。なお、

制御部１０は、このデータを新たに生成するのではなく、Ｓ５で記憶部１１に記録した電子書籍画像２０のデータを更新するようにしてもよい。

[0080] 制御部１０は、Ｓ８で輪郭線を重畳させた電子書籍画像２０内の領域を拡張させる（Ｓ９）。Ｓ９においては、制御部１０は、拡張後の電子書籍画像２０（図１１）を示すデータを生成して記憶部１１に記録する。なお、制御部１０は、このデータを新たに生成するのではなく、Ｓ５又はＳ８で記憶部１１に記録した電子書籍画像２０のデータを更新するようにしてもよい。

[0081] 制御部１０は、Ｓ９で拡張した電子書籍画像２０の領域のうち、基準未満の大きさの領域を除去する（Ｓ１０）。Ｓ１０においては、制御部１０は、基準未満の大きさの領域を除去した後の電子書籍画像２０（図１２）を示すデータを記憶部１１に記録する。なお、制御部１０は、このデータを新たに生成するのではなく、Ｓ５、Ｓ８、又はＳ９で記憶部１１に記録した電子書籍画像２０のデータを更新するようにしてもよい。

[0082] 制御部１０は、Ｓ１０で領域が除去された電子書籍画像２０に対して、輪郭抽出処理を実行する（Ｓ１１）。Ｓ１１の処理内容は、Ｓ６と同様であり、Ｓ１１で抽出される輪郭は、文字領域候補の輪郭である。Ｓ１１においては、制御部１０は、Ｓ１０で領域が除去された電子書籍画像２０内の輪郭線（図１３）を示すデータを生成して記憶部１１に記録する。

[0083] 制御部１０は、操作部１３からの入力を受け付ける（Ｓ１２）。ここでは、ユーザは、文字領域候補の指定操作、ページ送り操作、又は終了操作の何れかを行うことができるものとする。ユーザが文字領域候補の指定操作をしたと判定された場合（Ｓ１２；指定操作）、制御部１０は、指定された文字領域候補内に文字が含まれるか否かを判定する（Ｓ１３）。Ｓ１３においては、制御部１０は、学習データと文字領域候補内の画素値とに基づいて、文字領域候補内に文字が含まれるか否かを判定する。

[0084] 文字領域候補内に文字が含まれると判定された場合（Ｓ１３；Ｙ）、制御部１０は、指定操作により指定された文字領域候補内を拡大表示させる（Ｓ１４）。Ｓ１４においては、制御部１０は、図８に示す電子書籍画像２０か

ら文字領域候補の内部を切り出して、表示部 14 に拡大して表示させる。

[0085] 一方、ユーザがページ送り操作をしたと判定された場合（S 11；ページ送り操作）、制御部 10 は、次のページを表示部 14 に表示させ（S 15）、S 5 の処理に戻る。なお、表示中のページを示す情報は、記憶部 11 に記憶されているものとする。S 5 に戻った後は、当該次のページに対して S 5 ～S 11 の画像処理が実行されて、文字領域候補が抽出されることになる。

[0086] ユーザが終了操作をしたと判定された場合（S 12；終了操作）、本処理は終了する。なお、S 5 ～S 11 で記憶部 11 に記録されたデータは、本処理の終了時に破棄してもよいし、表示部 14 から該当するページが表示されなくなったときに破棄してもよい。他にも例えば、表示部 14 に表示されているページとのページ差が基準以上になった場合に破棄してもよい。

[0087] 以上説明した電子書籍表示装置 1 によれば、文字群 24 A、24 B の各文字を拡大して互いに結合させ、文字群 24 A、24 B の各文字をひとかたまりにしたうえで文字領域であるかを判定することができるので、OCR のように 1 文字 1 文字を個別に文字認識した後にこれらをかたまりにするような場合に比べて、電子書籍画像 20 に含まれる文字群 24 A、24 B の領域を迅速に特定することができる。例えば、文字領域候補の全体に対して学習データを適用する場合には、文字領域候補の全体的な特徴から文字領域であるかを判定することができるので、OCR に比べて簡易な処理により文字領域であるかを判定することができる。更に、文字を個別に判定する場合には、ある文字と他の文字が互いに関連する 1 つの文字群であるのかを特定することができない可能性があるが、電子書籍表示装置 1 は、文字群 24 A、24 B のように、互いに関連する文字同士を文字群のかたまりとして判定することができる。更に、電子書籍の購入時などに、各文字群の位置を示すデータを予め記憶部 11 に記憶させておくような場合に比べて、電子書籍を表示させる場合に動的に文字領域の判定を行う場合には、メモリ容量を有効活用することができる。

[0088] また、文字群 24 A、24 B が拡大して、吹き出し線 23 A、23 B に結

合してしまうと、文字群をかたまりとして認識することができない可能性があるが、電子書籍表示装置 1 は、吹き出し線 2 3 A, 2 3 B と文字群 2 4 A, 2 4 B とが結合しないように制限するので、文字群をかたまりとして認識することができ、文字領域の判定の精度も向上させることができる。更に、吹き出し線 2 3 A, 2 3 B をノイズとして除去することで、拡大表示した際に吹き出し線 2 3 A, 2 3 B も拡大されてしまうようなことを防止することができる。

[0089] 更に、吹き出し線 2 3 A, 2 3 B の領域を背景色で塗りつぶすことによって、比較的簡易な処理で、文字群 2 4 A, 2 4 B と結合しないようにすることができるので、電子書籍表示装置 1 の処理負荷を軽減することができる。

[0090] また、小さな領域を文字領域の判定の対象外とすることで、ノイズを確実に除去することができ、明らかに文字が含まれていない領域に対しても文字領域であるかの判定をすることがなくなり、無駄な処理を実行することがなくなる。更に、小さな領域を記憶部 1 1 に記憶させず破棄する場合には、メモリ領域の有効活用をすることができる。

[0091] また、小さな領域を背景色で塗りつぶすことによって、比較的簡易な処理で、文字領域の判定の対象外とすることができるので、電子書籍表示装置 1 の処理負荷を軽減することができる。

[0092] また、指定操作により指定された文字領域候補に対して文字領域の判定を行うことによって、ユーザが指定していない文字領域候補以外に対しては文字領域であるかの判定をしないので、無駄な文字領域の判定をする必要がなくなり（例えば、ユーザが特に拡大する必要のないと思う領域については文字領域の判定を行わない）、電子書籍表示装置 1 の処理負荷を軽減することができる。

[0093] また、文字領域であると判定された文字領域候補を拡大表示することによって、OCR のように文字を 1 つずつ認識してからこれらをかたまりにするような場合に比べて、迅速に文字群の拡大表示をすることができる。更に、文字をひとかたまりにしたうえでその領域の全体に対して文字領域の判定を

行うので、OCRのように1文字ずつ認識してからこれらをかたまりにして拡大するのではなく、比較的簡易な処理によって文字のかたまりを拡大表示させることができる。

[0094] [4. 変形例]

なお、本発明は、以上に説明した実施の形態に限定されるものではない。本発明の趣旨を逸脱しない範囲で、適宜変更可能である。

[0095] 例えば、各コマ21に描かれる文字は、キャラクタが発する言葉だけでなく、キャラクタの心情や漫画の効果音を表すものであってもよい。このような文字であっても、拡張部104が領域を拡張することで、文字をかたまりとして認識可能である。また、実施形態では、漫画の台詞が英語で記述されている場合を例に挙げたが、電子書籍表示装置1で文字領域の判定が可能な言語であればよく、他の言語であってもよい。更に、キャラクタの台詞が横書きで記述される場合を説明するが、縦書きであってもよい。この場合、横書き用の学習パターンと、縦書き用の学習パターンと、の2つをデータ記憶部100に用意しておき、文字領域判定部108は、横書き用と縦書き用の学習パターンの何れかで文字領域と判定されれば、文字領域候補を文字領域と判定すればよい。

[0096] また例えば、結合制限部103が行う2値化反転処理は省略してもよい。また、実施形態では、文字が吹き出し線に囲われている場合を説明したが、文字は、吹き出し線に囲われていなくてもよい。この場合、結合制限部103の処理は省略してよい。また、文字は、所定のフォントであってもよいし、手書きであってもよい。手書きの文字であっても、拡張部104の処理により文字をかたまりとして判定可能である。更に、吹き出し・コマ・ページによって、文字のフォントや大きさ・色彩・輝度が異なってもよい。フォントや大きさ・色彩・輝度が異なっていたとしても、拡張部104が画像内の領域を拡張させることで、文字をかたまりとして判定可能である。

[0097] また例えば、処理実行部109が実行する所定の処理の一例として、文字領域判定部108により文字領域であると判定された文字領域候補を拡大表

示する場合を説明したが、文字領域であると判定された文字領域候補に対する処理内容は、拡大表示などの画像処理に限られない。他にも、電子書籍表示装置 1 は、文字領域であると判定された文字領域候補内の文字を翻訳するようにしてもよいし、機械音声を出力するようにしてもよい。即ち、所定の処理は、翻訳処理であってもよいし、音声出力処理であってもよい。更に、電子書籍表示装置 1 は、文字を含むと判定された文字領域候補内を OCR により文字抽出を行い、当該抽出した文字のデータを記憶部 11 に記録して、電子書籍における文字検索を可能なようにしてもよい。即ち、所定の処理は、文字抽出処理や検索処理であってもよい。なお、処理実行部 109 は、文字領域と判定された文字領域候補内を切り出して拡大表示するのではなく、OCR により検出した文字のデータに基づいて拡大表示してもよい。

[0098] また例えば、実施形態では、文字領域判定部 108 は、ユーザが文字領域候補を指定する指定操作をすることを条件として、当該文字領域候補が文字領域であるかを判定する場合を説明した。即ち、文字領域判定部 108 は、ユーザが指定操作をするまでは、文字領域であるかの判定処理を待機する場合を説明したが、当該判定処理は、任意のタイミングで実行されるようにすればよく、判定処理を実行するための条件は、指定操作に限られない。例えば、ユーザが特に操作をしなくても、文字領域判定部 108 は、全ての文字候補領域に対して自動的に文字領域であるかの判定処理を実行するようにしてもよい。

[0099] 例えば、文字領域判定部 108 は、電子書籍画像 20 が表示された場合に、文字領域候補に対して自動的に文字領域であるか否かを判定するようにしてもよい。また例えば、文字領域判定部 108 は、電子書籍画像 20 が表示される前に、文字領域候補に対して自動的に文字領域であるか否かを判定するようにしてもよい。別の言い方をすれば、電子書籍のプログラムが起動してから、電子書籍画像 20 が表示されるまでの間に、文字領域判定部 108 は、文字領域の判定をするようにしてもよい。また例えば、文字領域判定部 108 は、電子書籍のプログラムが起動する前であっても、文字領域の判定

をするようにしてもよい。例えば、電子書籍表示装置 10 がサーバなどから電子書籍の画像データをダウンロードしてから、電子書籍のプログラムを起動するまでの間に、文字領域判定部 108 は、文字領域候補に対して文字領域の判定をしてもよい。なお、この場合、結合制限部 103、拡張部 104、判定制限部 105、及び文字領域候補記録部 106 の各々の処理を事前に実行しておき、予め文字領域候補を抽出しておけばよい。

[0100] 更に、本発明に係る画像処理装置は、電子書籍表示装置 10 以外にも適用可能であり、画像処理装置は、特に画像を表示しない装置であってもよく、表示制御部 102 は省略してもよい。例えば、画像処理装置は、電子書籍の販売前において、電子書籍画像 20 に対して予め文字領域を判定し、電子書籍の画像データとともに、当該判定した文字領域を示すデータをユーザに提供するようにしてもよい。他にも例えば、画像処理装置は、サーバコンピュータで実現されてもよく、例えば、電子書籍の画像データが画像処理装置にアップロードされた場合に、文字領域判定部 108 は、文字領域を判定するようにしてもよい。この場合も、結合制限部 103、拡張部 104、判定制限部 105、及び文字領域候補記録部 106 の各々の処理を事前に実行しておき、予め文字領域候補を抽出しておけばよい。

[0101] また例えば、実施形態では、電子書籍の一例として漫画を説明したが、文字を含む電子書籍であればよく、漫画以外にも本発明は適用可能である。例えば、文字の周りに一定のスペースが存在する絵本や雑誌等に対しても、実施形態と同様の文字領域の抽出が可能なため、絵本や雑誌等に適用してもよい。更に、本発明は、電子書籍以外の画像に対しても適用可能である。例えば、文字が描かれた紙などをスキャンした画像データに対して、文字領域判定部 108 が文字領域の判定を行い、処理実行部 109 が所定の処理を実行するようにしてもよい。また例えば、文字と絵などが混在する資料、論文、チラシなどの画像データに対して、文字領域判定部 109 は文字領域の判定をするようにしてもよい。本発明に係る処理は、文字を含む種々の画像に対して適用可能である。

請求の範囲

- [請求項1] 画像データを記憶する手段に記憶された前記画像データを取得する画像取得手段と、
- 前記画像取得手段により取得された画像データが示す画像に含まれる物体を示す複数の領域の各々を拡張して結合させる拡張手段と、
- 前記拡張手段により拡張して結合した全部又は一部の領域が文字領域であるか否かを判定する文字領域判定手段と、
- 前記文字領域判定手段により文字領域であると判定された領域に基づいて、所定の処理を実行する処理実行手段と、
- を含むことを特徴とする画像処理装置。
- [請求項2] 前記画像処理装置は、前記画像に含まれる物体を示す領域の輪郭又は当該領域自身の大きさが基準以上である場合、当該領域と、前記拡張手段により拡張した領域と、が結合することを制限する結合制限手段を更に含む、
- ことを特徴とする請求項1に記載の画像処理装置。
- [請求項3] 前記結合制限手段は、前記画像に含まれる物体を示す領域の輪郭又は当該領域自身の大きさが基準以上である場合、当該領域を背景色で塗りつぶすことによって、前記拡張手段により拡張した領域と結合することを制限する、
- ことを特徴とする請求項2に記載の画像処理装置。
- [請求項4] 前記画像処理装置は、前記拡張手段により拡張した領域の輪郭又は当該領域自身の大きさが基準未満である場合、当該領域が前記文字領域判定手段の判定対象となることを制限する判定制限手段を更に含む、
- ことを特徴とする請求項1～3の何れかに記載の画像処理装置。
- [請求項5] 前記判定制限手段は、前記拡張手段により拡張した領域の輪郭又は当該領域自身の大きさが基準未満である場合、当該領域を背景色で塗りつぶすことによって、前記文字領域判定手段の判定対象となること

を制限する、

ことを特徴とする請求項4に記載の画像処理装置。

[請求項6] 前記画像処理装置は、前記拡張手段により拡張して結合した一部の領域の指定操作を受け付ける指定操作受付手段を更に含み、

前記文字領域判定手段は、前記指定操作受付手段により受け付けられた指定操作により指定された一部の領域が文字領域であるか否かを判定する、

ことを特徴とする請求項1～5の何れかに記載の画像処理装置。

[請求項7] 前記処理実行手段は、前記画像が表示された表示手段において、前記文字領域判定手段により文字領域であると判定された領域内を拡大して表示させる、

ことを特徴とする請求項1～6の何れかに記載の画像処理装置。

[請求項8] 画像データを記憶する手段に記憶された前記画像データを取得する画像取得ステップと、

前記画像取得ステップにより取得された画像データが示す画像に含まれる物体を示す複数の領域の各々を拡張して結合させる拡張ステップと、

前記拡張ステップにより拡張して結合した全部又は一部の領域が文字領域であるか否かを判定する文字領域判定ステップと、

前記文字領域判定ステップにより文字領域であると判定された領域に基づいて、所定の処理を実行する処理実行ステップと、

を含むことを特徴とする画像処理方法。

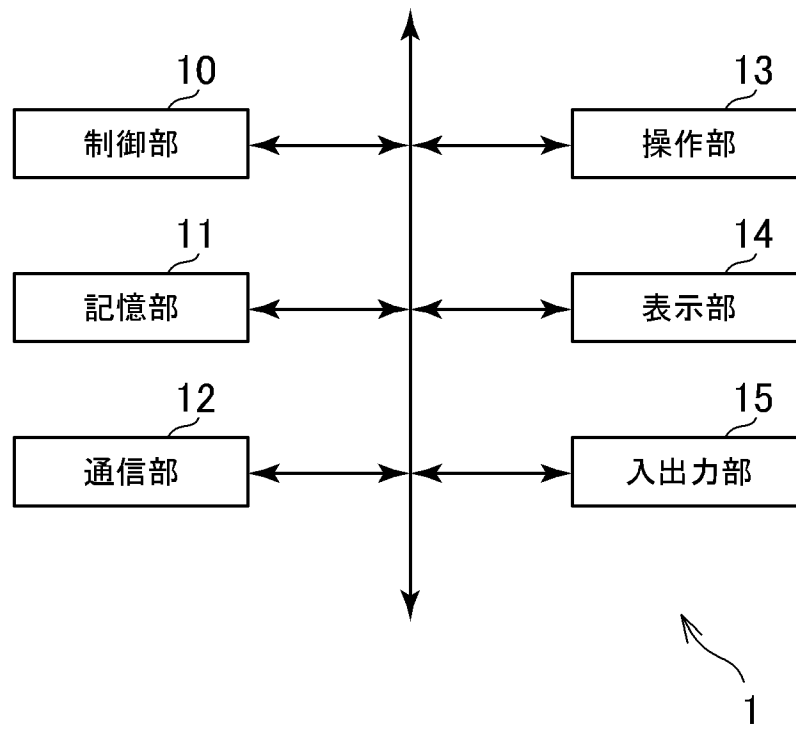
[請求項9] 画像データを記憶する手段に記憶された前記画像データを取得する画像取得手段、

前記画像取得手段により取得された画像データが示す画像に含まれる物体を示す複数の領域の各々を拡張して結合させる拡張手段、

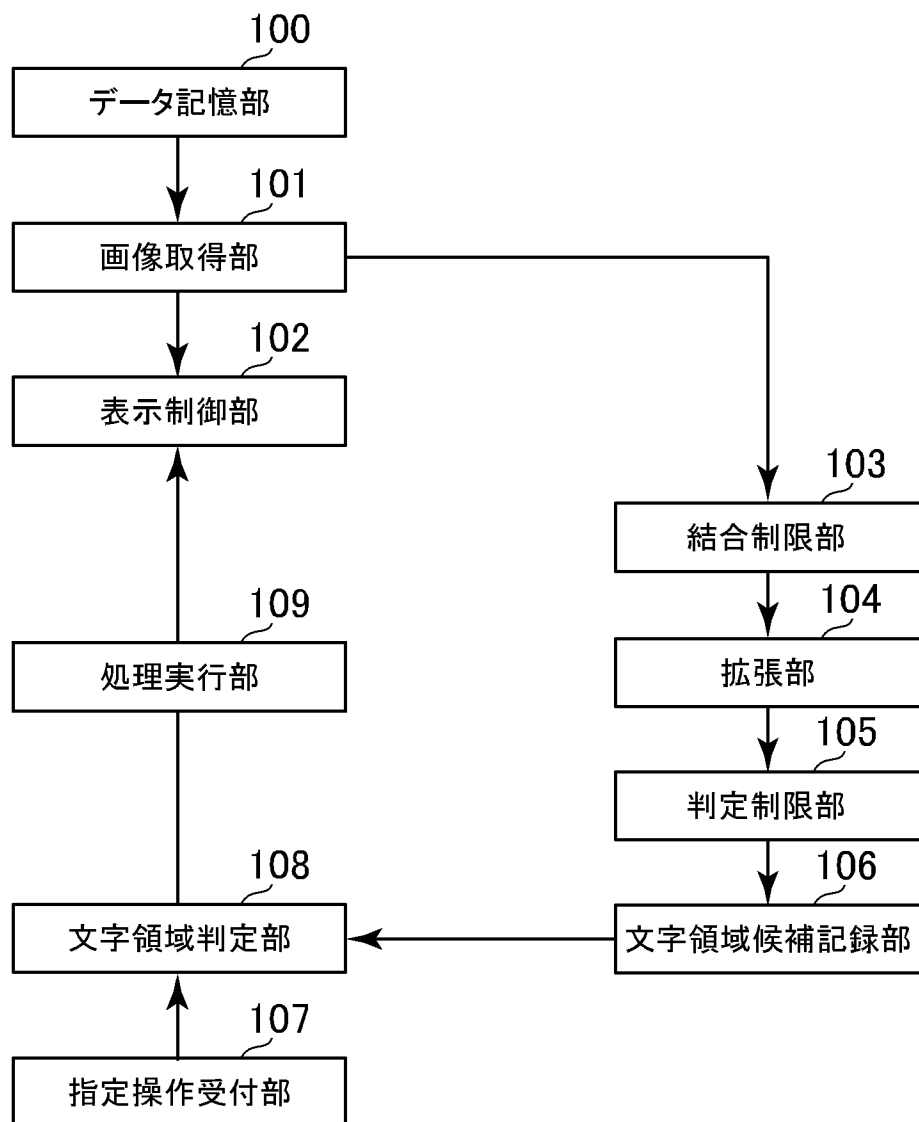
前記拡張手段により拡張して結合した全部又は一部の領域が文字領域であるか否かを判定する文字領域判定手段、

前記文字領域判定手段により文字領域であると判定された領域に基づいて、所定の処理を実行する処理実行手段、
としてコンピュータを機能させるためのプログラム。

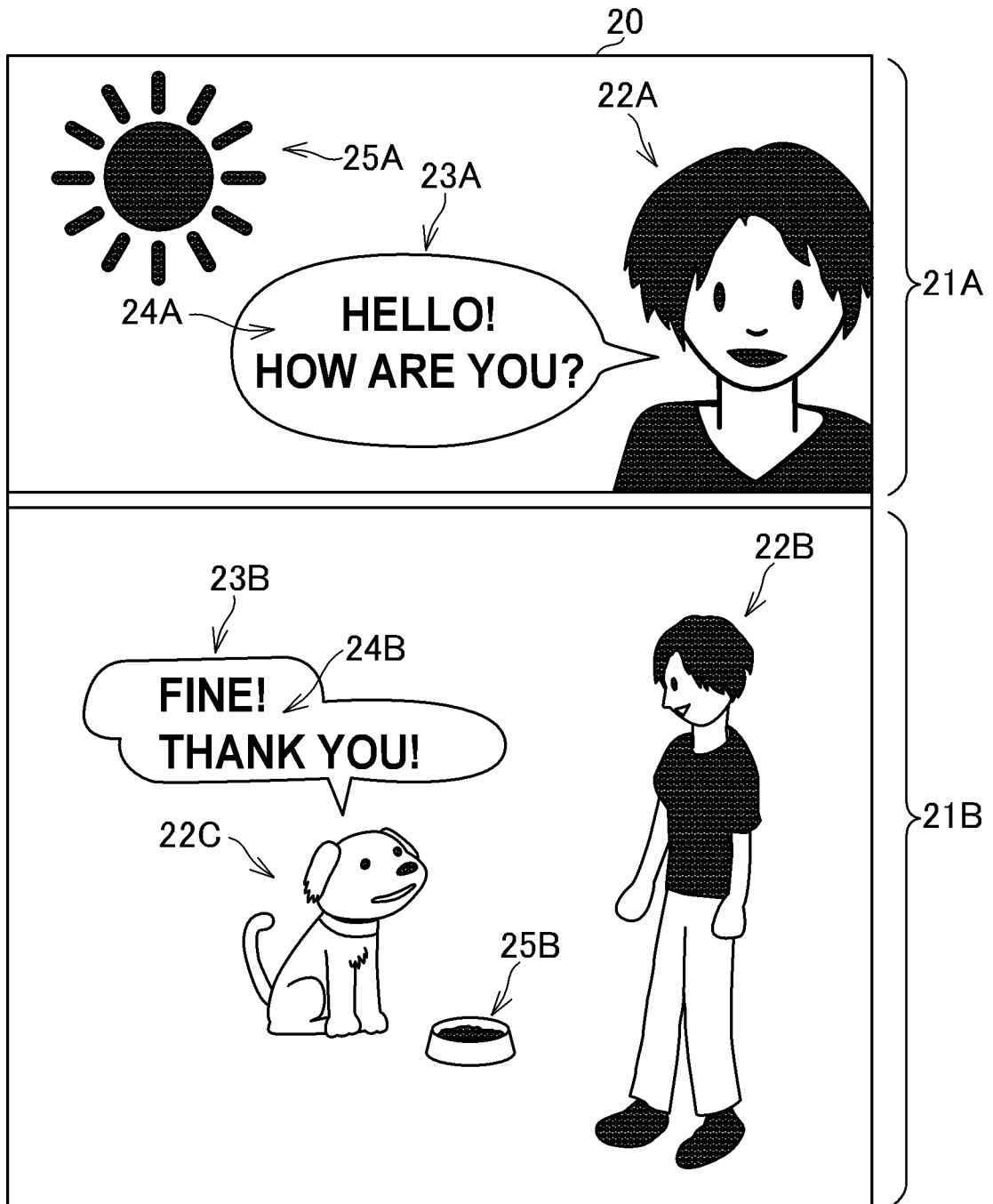
[図1]



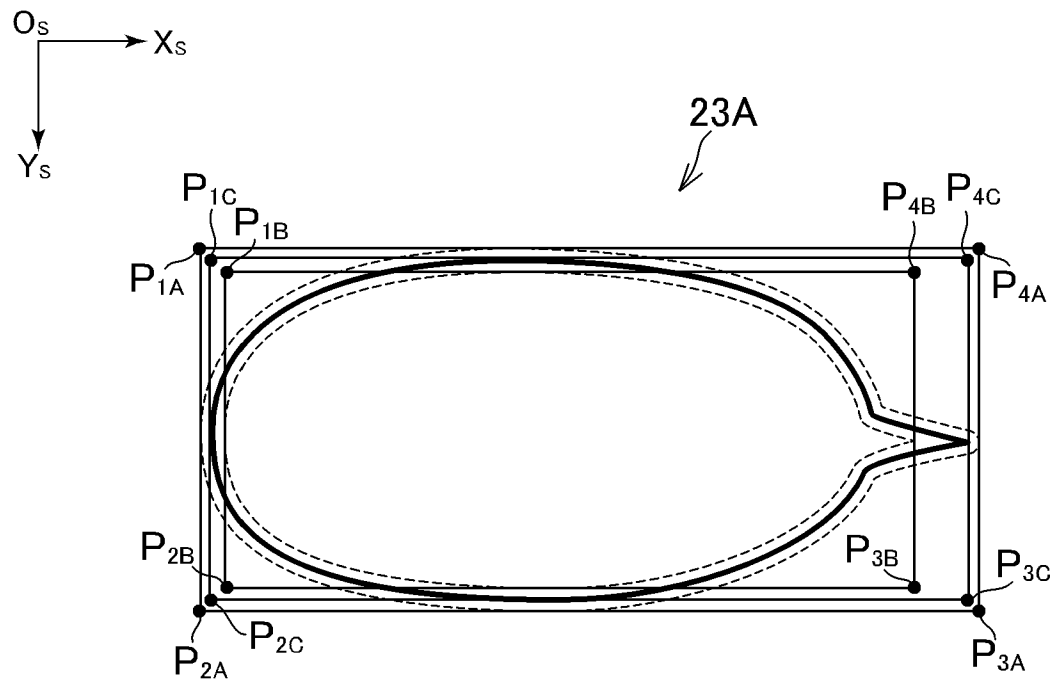
[図2]



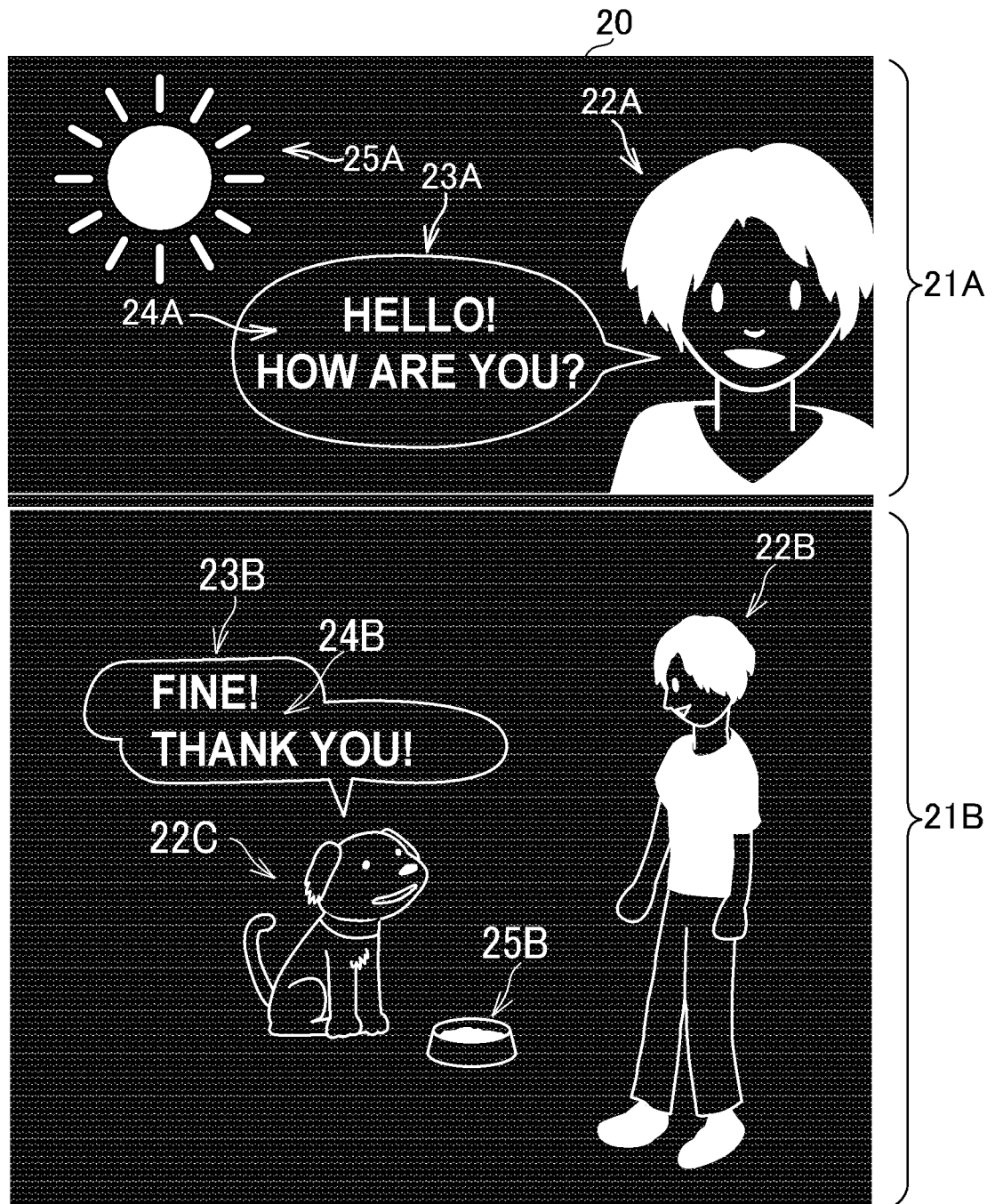
[図3]



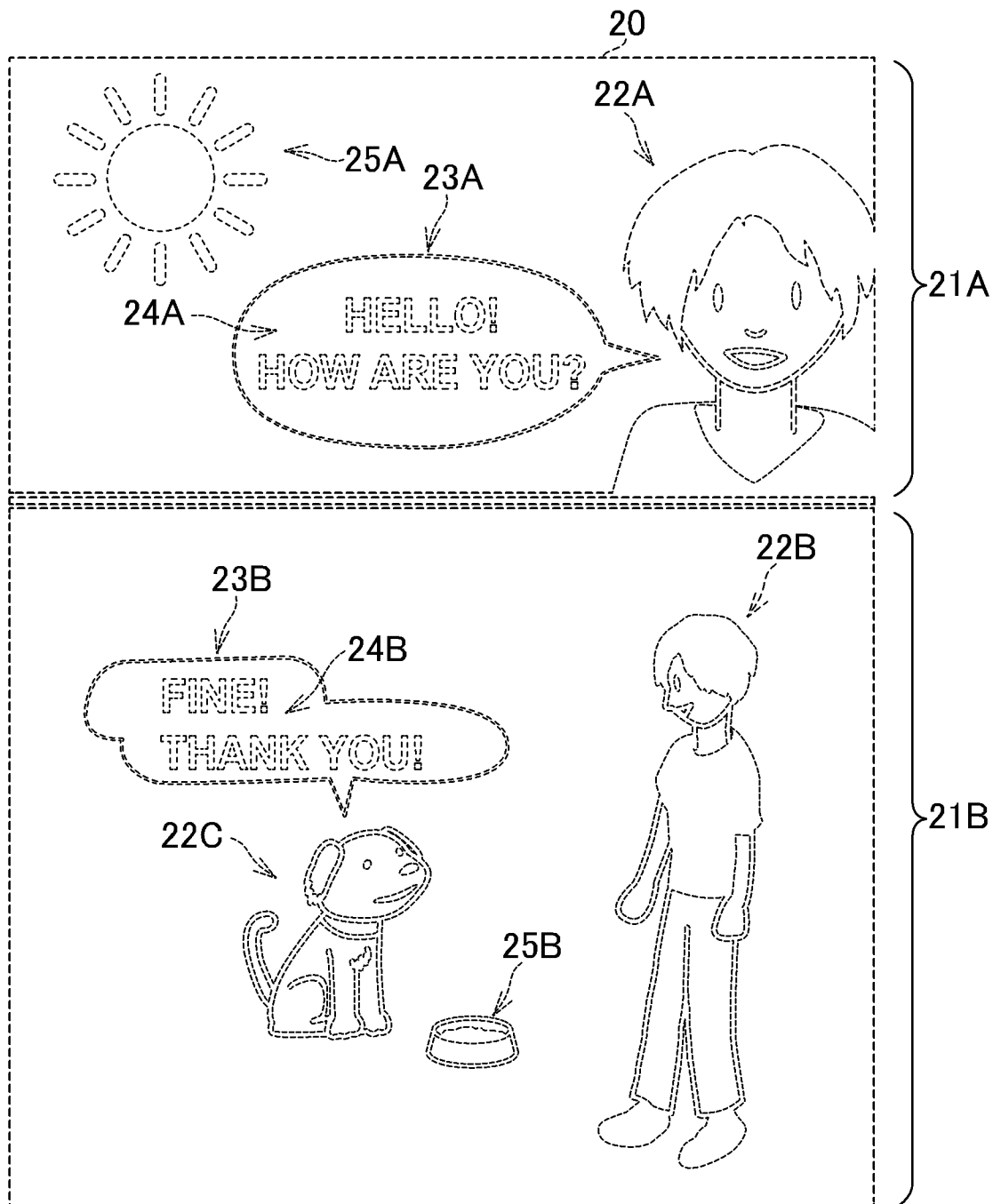
[図4]



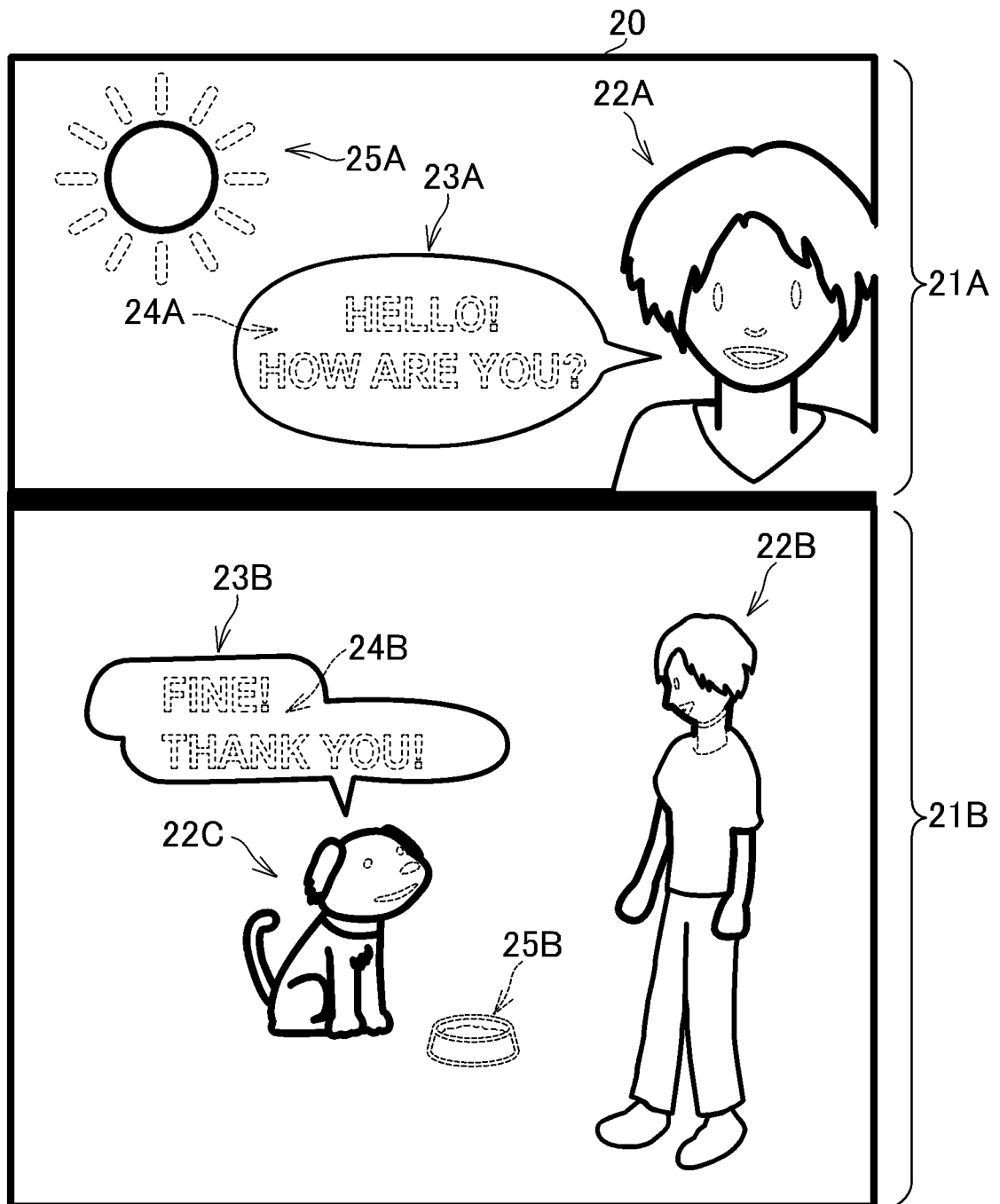
[図5]



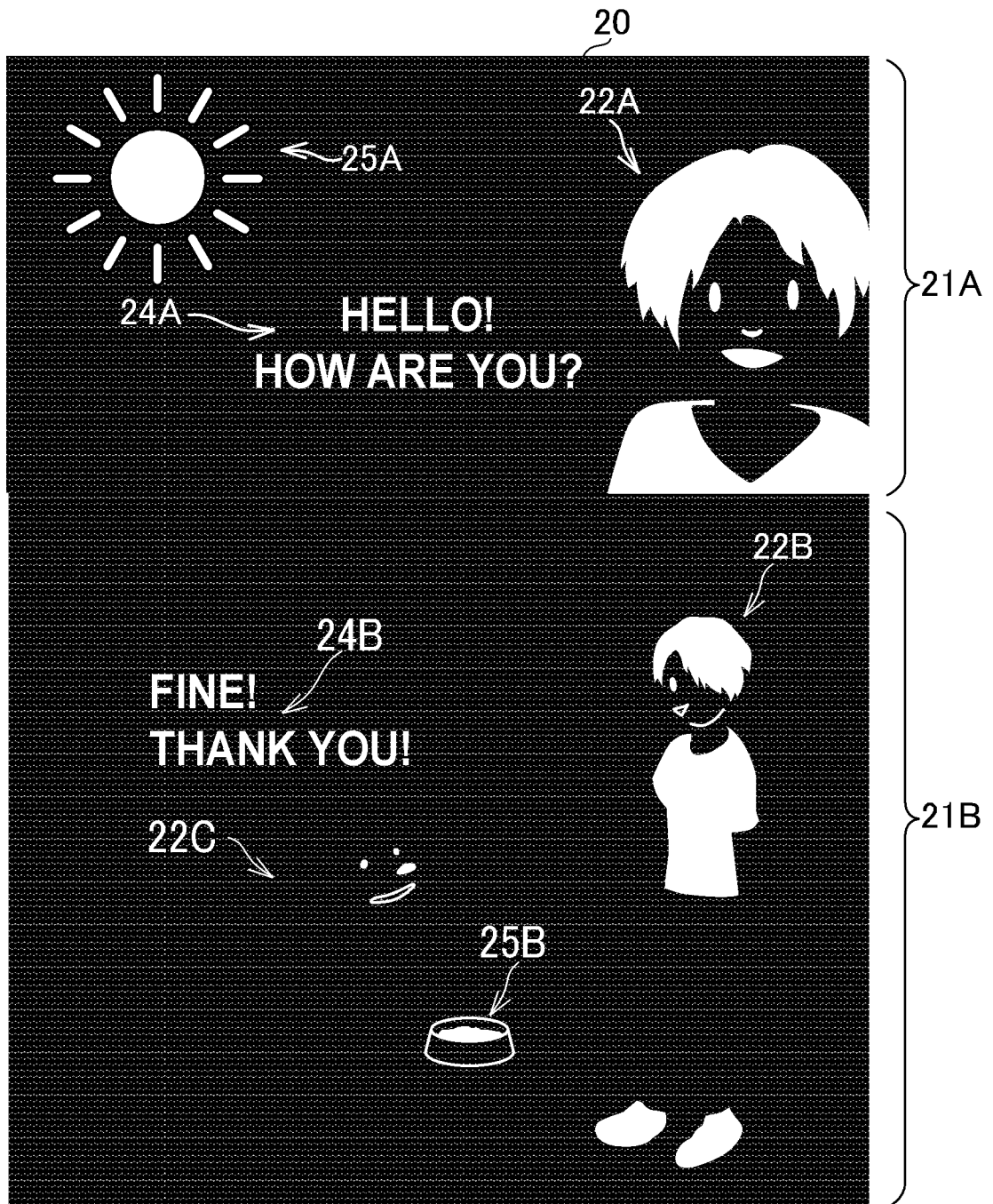
[図6]



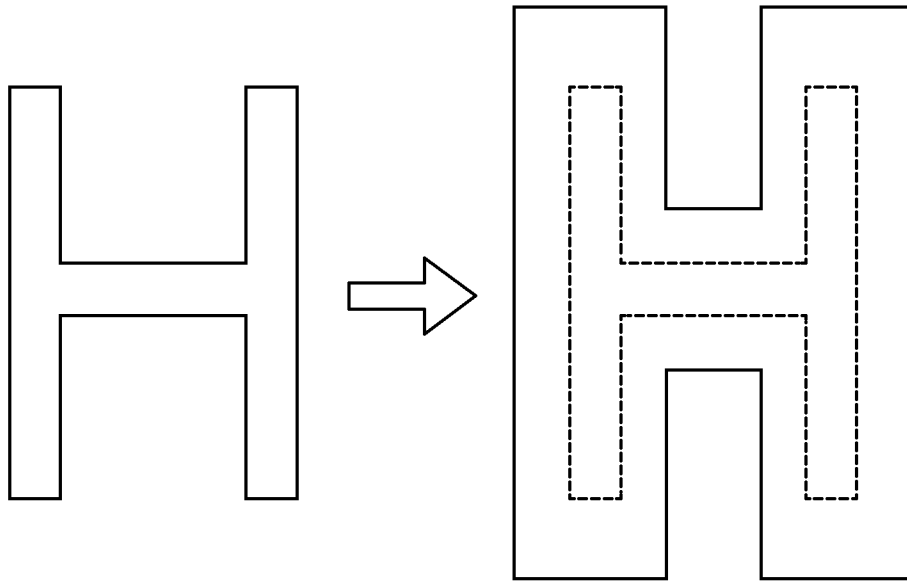
[図7]



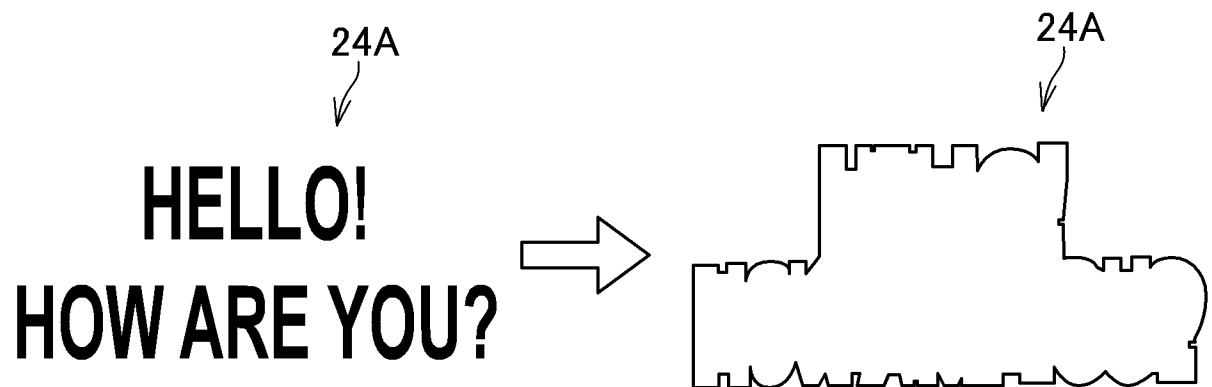
[図8]



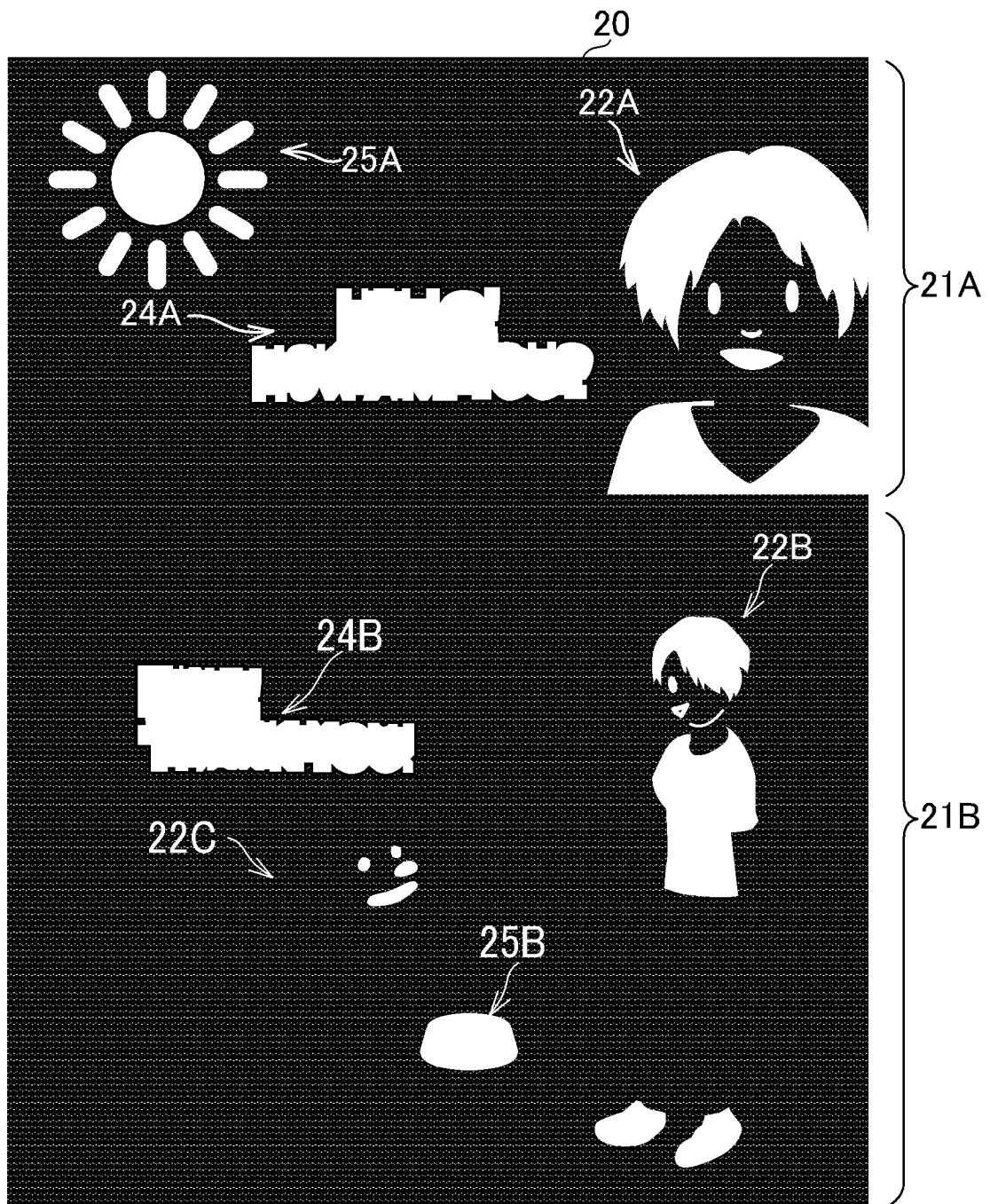
[図9]



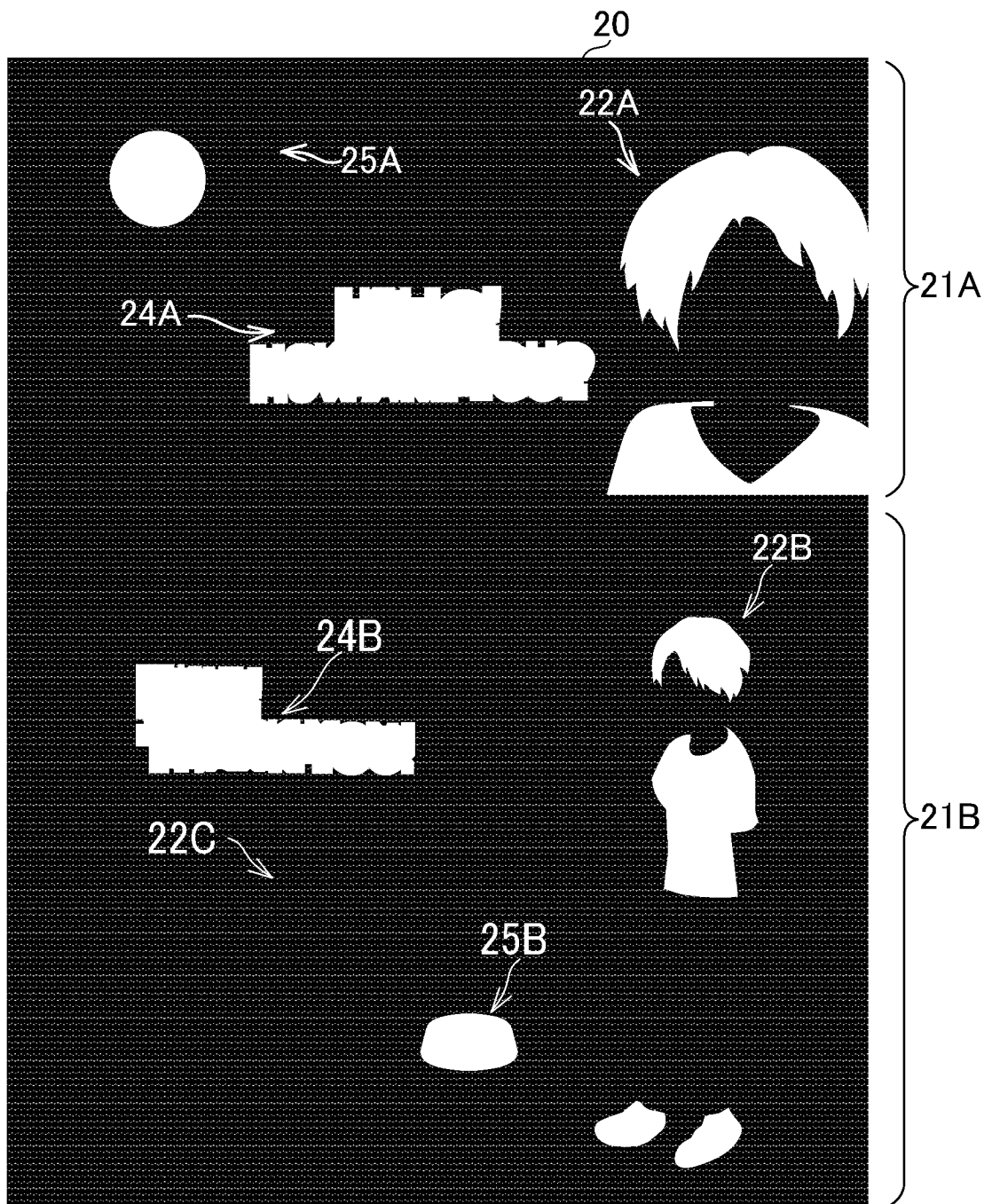
[図10]



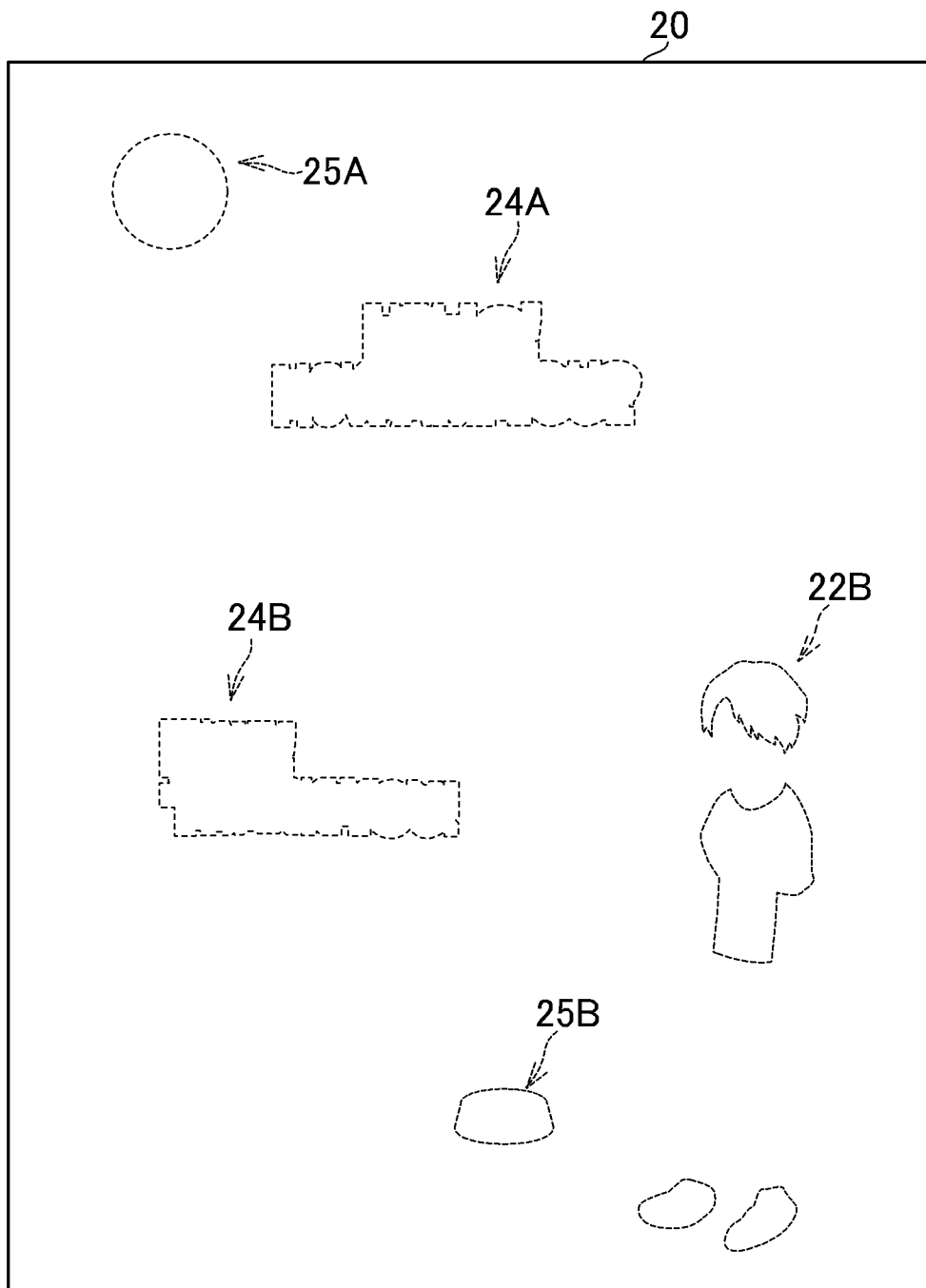
[図11]



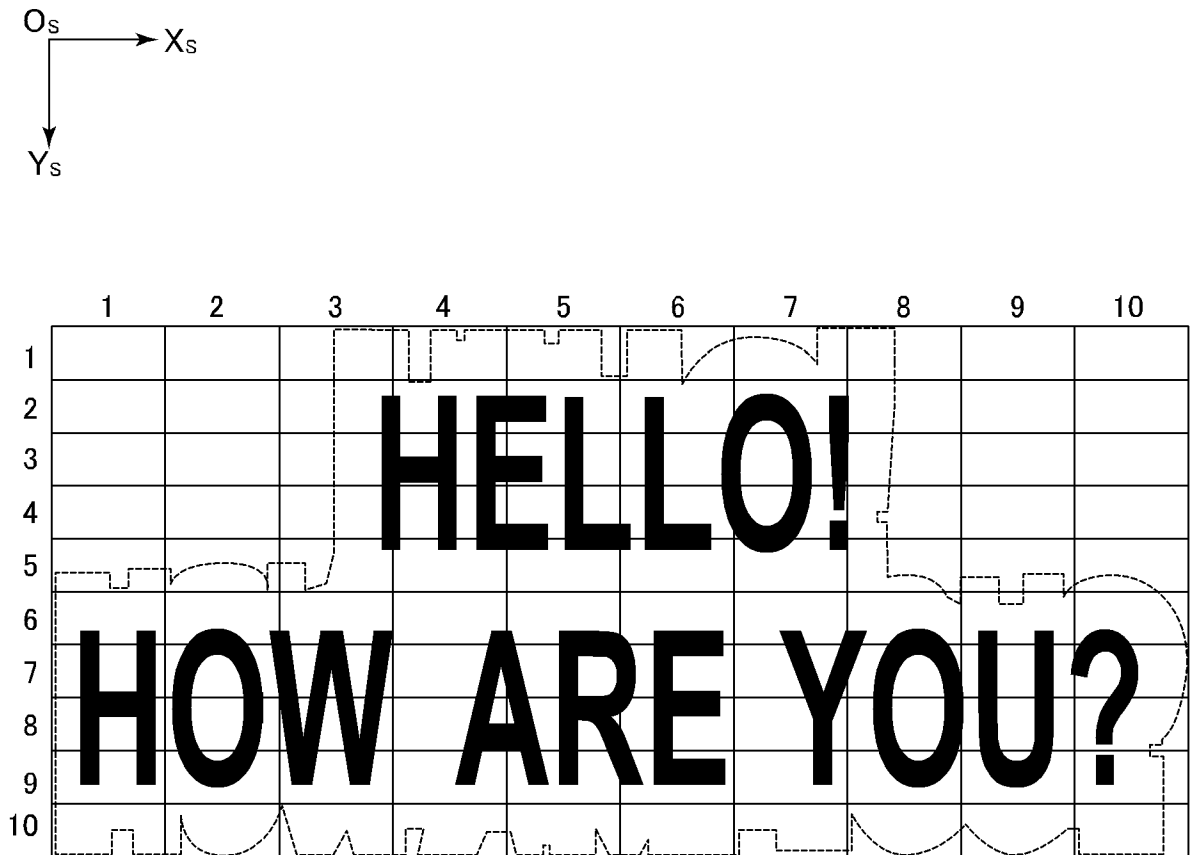
[図12]



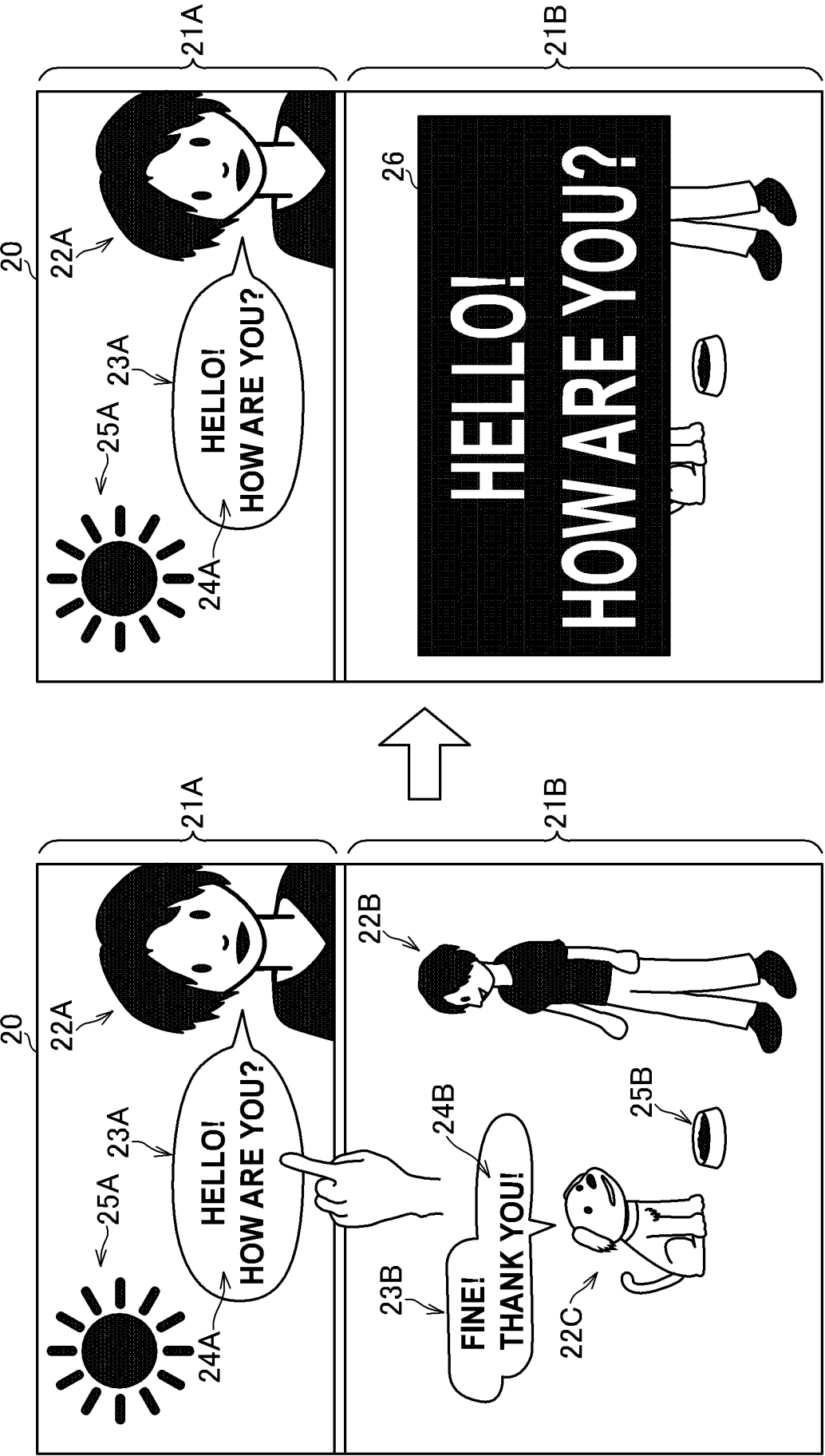
[図13]



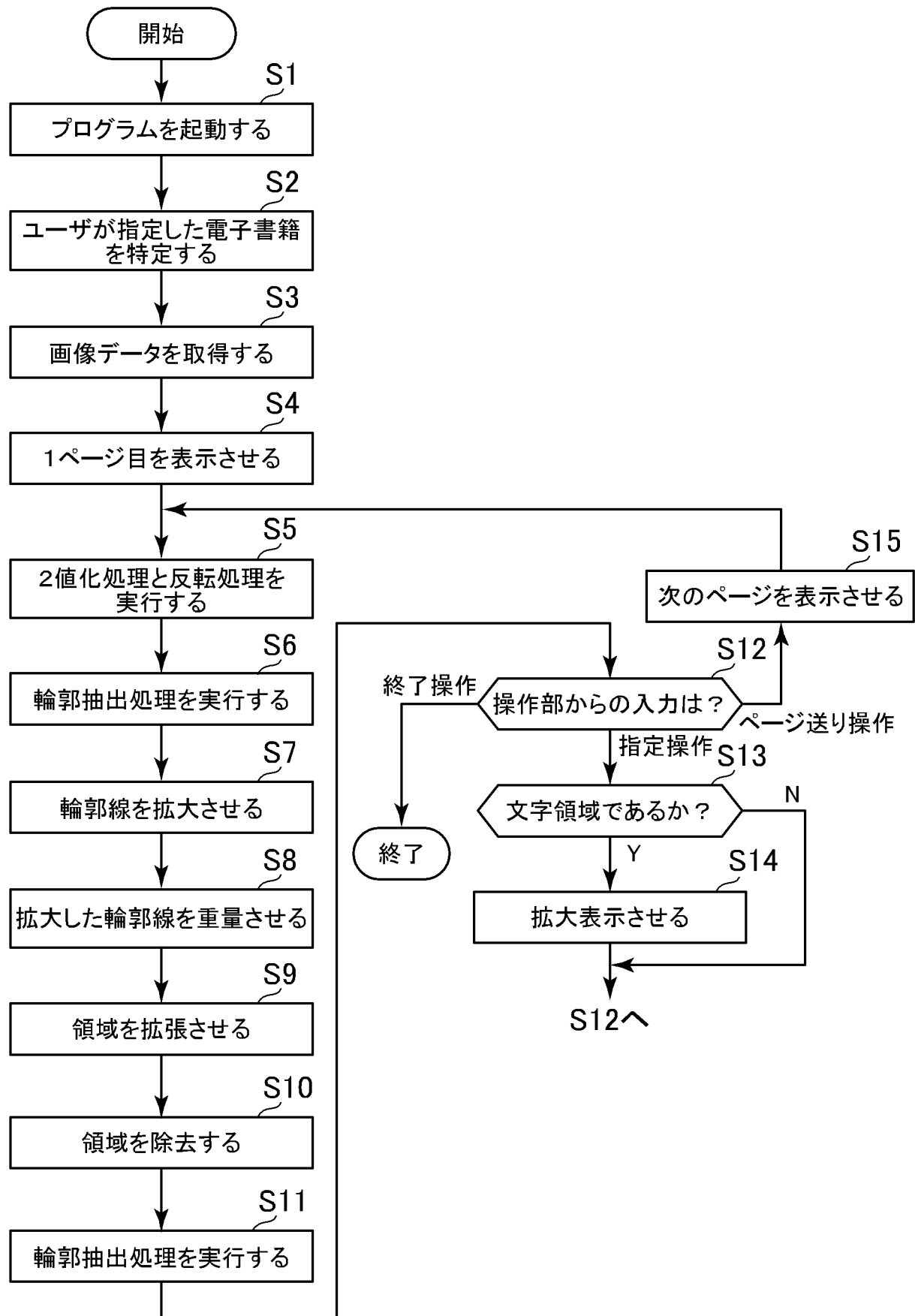
[図14]



[図15]



[図16]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2016/069884

A. CLASSIFICATION OF SUBJECT MATTER

G06T7/40(2006.01) i

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)

G06T7/40, G06K9/20, H04N1/40

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2016

Kokai Jitsuyo Shinan Koho 1971-2016 Toroku Jitsuyo Shinan Koho 1994-2016

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 Y A	JP 2001-143074 A (Minolta Co., Ltd.), 25 May 2001 (25.05.2001), paragraphs [0047] to [0057] (Family: none)	1, 6, 8, 9 4, 5, 7 2, 3
Y A	JP 8-221513 A (Canon Inc.), 30 August 1996 (30.08.1996), paragraphs [0089] to [0097] & US 5920655 A column 8, line 64 to column 9, line 65 & EP 731599 A2	4, 5 1-3, 6-9

☒ Further documents are listed in the continuation of Box C.

☐ 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

"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
15 September 2016 (15.09.16)

Date of mailing of the international search report
27 September 2016 (27.09.16)

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2016/069884

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	JP 2005-328348 A (Ricoh Co., Ltd.), 24 November 2005 (24.11.2005), paragraph [0019] & US 2005/0180645 A1 paragraph [0136] & EP 1555804 A2 & CN 1649384 A	5 1-4, 6-9
Y A	JP 2002-165079 A (Minolta Co., Ltd.), 07 June 2002 (07.06.2002), claims (Family: none)	7 1-6, 8, 9
A	JP 2-071379 A (Fuji Xerox Co., Ltd.), 09 March 1990 (09.03.1990), page 4, lower left column, line 7 to page 5, upper left column, line 20 (Family: none)	1-9

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. G06T7/40(2006.01)i

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（I P C））

Int.Cl. G06T7/40, G06K9/20, H04N1/40

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	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 6 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 6 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 6 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 Y A	JP 2001-143074 A（ミノルタ株式会社）2001.05.25, [0047] - [0057] （ファミリーなし）	1, 6, 8, 9 4, 5, 7 2, 3
Y A	JP 8-221513 A（キヤノン株式会社）1996.08.30, [0089] - [0097] & US 5920655 A, column 8 line 64 - column 9 line 65 & EP 731599 A2	4, 5 1-3, 6-9

☒ C欄の続きにも文献が列挙されている。

☐ 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
「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矛盾するものではなく、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

1 5 . 0 9 . 2 0 1 6

国際調査報告の発送日

2 7 . 0 9 . 2 0 1 6

国際調査機関の名称及びあて先

日本国特許庁（I S A / J P）

郵便番号 1 0 0 - 8 9 1 5

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

松浦 功

5 H

9 1 8 1

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 3 1

C (続き) . 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
Y A	JP 2005-328348 A (株式会社リコー) 2005. 11. 24, [0019] & US 2005/0180645 A1, [0136] & EP 1555804 A2 & CN 1649384 A	5 1-4, 6-9
Y A	JP 2002-165079 A (ミノルタ株式会社) 2002. 06. 07, 特許請求の範囲 (ファミリーなし)	7 1-6, 8, 9
A	JP 2-071379 A (富士ゼロックス株式会社) 1990. 03. 09, 第 4 頁左下 欄第 7 行ー第 5 頁左上欄第 20 行 (ファミリーなし)	1-9