

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2013-178659

(P2013-178659A)

(43) 公開日 平成25年9月9日(2013.9.9)

|                            |                |             |
|----------------------------|----------------|-------------|
| (51) Int.Cl.               | F I            | テーマコード (参考) |
| <b>G06K 9/20 (2006.01)</b> | G06K 9/20 340K | 5B029       |
|                            | G06K 9/20 360C |             |
|                            | G06K 9/20 340B |             |

審査請求 未請求 請求項の数 12 O L (全 9 頁)

(21) 出願番号 特願2012-42085 (P2012-42085)  
 (22) 出願日 平成24年2月28日 (2012.2.28)

特許法第30条第1項適用申請有り

(71) 出願人 800000068  
 学校法人東京電機大学  
 東京都足立区千住旭町5番  
 (74) 代理人 100110928  
 弁理士 速水 進治  
 (72) 発明者 中村 明生  
 東京都千代田区神田錦町2丁目2番地 学  
 校法人東京電機大学内  
 (72) 発明者 淵田 正隆  
 東京都千代田区神田錦町2丁目2番地 学  
 校法人東京電機大学内  
 Fターム(参考) 5B029 AA01 AA02 CC25 CC28 DD04  
 EE04 EE13 EE16 EE18

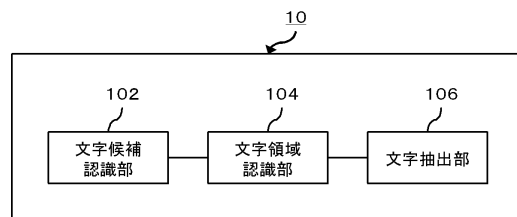
(54) 【発明の名称】 文字認識装置、文字認識方法およびプログラム

(57) 【要約】

【課題】 指先でなぞった部分の文字をより高い精度で認識する、汎用的な文字認識装置、文字認識方法およびプログラムを提供することにある。

【解決手段】 文字認識装置10は、文字候補認識部102と、文字列領域認識部104と、文字列抽出部106を有する。文字候補認識部102は、1つの文字と推定される領域である文字候補領域を認識する。文字列領域認識部104は、文字候補認識部102が認識した文字候補領域毎に、文字と推定される色を有する領域を膨張させる膨張処理を施し、膨張処理を施した文字候補領域同士で繋がった領域を、文字列領域112として認識する。文字列抽出部106は、文字列領域112に基づき、画像から文字を抽出する。

【選択図】 図1



## 【特許請求の範囲】

## 【請求項 1】

2 値化された画像から、1つの文字と推定される領域である文字候補領域を認識する文字候補認識手段と、

前記文字候補認識手段が認識した前記文字候補領域毎に、文字と推定される色を有する領域を膨張させる膨張処理を施し、前記膨張処理を施した前記文字候補領域同士で繋がった領域を文字列領域として認識する文字列領域認識手段と、

前記文字列領域に基づき、前記画像から文字列を抽出する文字列抽出手段を有する文字認識装置。

## 【請求項 2】

10

請求項 1 に記載の文字認識装置において、  
前記文字列抽出手段は、

前記画像全体における 2 次元座標値を保持しており、前記画像全体における前記 2 次元座標値と前記文字列領域の前記 2 次元座標値とを比較して、前記画像から前記文字列を抽出する文字認識装置。

## 【請求項 3】

請求項 1 に記載の文字認識装置において、  
前記文字列抽出手段は、

前記膨張処理を施した画像をマスク情報として別途保持し、前記画像と前記マスク情報の論理積から、前記文字列を抽出する文字認識装置。

20

## 【請求項 4】

請求項 1 ~ 3 のいずれか一項に記載の文字認識装置において、  
前記文字候補認識手段は、

前記画像内で移動する指標の位置に基づき、前記画像から前記文字候補領域を認識する文字認識装置。

## 【請求項 5】

コンピュータが、

2 値化された画像から、1つの文字と推定される領域である文字候補領域を認識し、前記文字候補領域毎に、文字と推定される色を有する領域を膨張させる膨張処理を施し、前記膨張処理を施した前記文字候補領域同士で繋がった領域を文字列領域として認識し

30

、前記文字列領域に基づき、前記画像から文字列を抽出する文字認識方法。

## 【請求項 6】

請求項 5 に記載の文字認識方法において、  
前記コンピュータが、

前記画像全体における 2 次元座標値を保持しており、前記画像全体における前記 2 次元座標値と前記文字列領域の前記 2 次元座標値とを比較して、前記画像から前記文字列を抽出する文字認識方法。

## 【請求項 7】

請求項 5 に記載の文字認識方法において、  
前記コンピュータが、

前記膨張処理を施した画像をマスク情報として別途保持し、前記画像と前記マスク情報の論理積から、前記文字列を抽出する文字認識方法。

40

## 【請求項 8】

請求項 5 ~ 7 のいずれか一項に記載の文字認識方法において、  
前記コンピュータが、

前記画像内で移動する指標の位置に基づき、前記画像から前記文字候補領域を認識する文字認識方法。

## 【請求項 9】

コンピュータを、

50

2 値化された画像から、1つの文字と推定される領域である文字候補領域を認識する手段、

前記文字候補領域毎に、文字と推定される色を有する領域を膨張させる膨張処理を施し、前記膨張処理を施した前記文字候補領域同士で繋がった領域を文字列領域として認識する手段、

前記文字列領域に基づき、前記画像から文字列を抽出する手段として機能させるためのプログラム。

【請求項 10】

請求項 9 に記載のプログラムにおいて、

前記コンピュータを、

前記画像全体における 2 次元座標値を保持する手段、

前記画像全体における前記 2 次元座標値と前記文字列領域の前記 2 次元座標値とを比較して、前記画像から前記文字列を抽出する手段としてさらに機能させるためのプログラム。

【請求項 11】

請求項 9 に記載のプログラムにおいて、

前記コンピュータを、

前記膨張処理を施した画像をマスク情報として別途保持する手段、

前記画像と前記マスク情報の論理積から、前記文字列を抽出する手段としてさらに機能させるためのプログラム。

【請求項 12】

請求項 9 ~ 11 のいずれか一項に記載のプログラムにおいて

前記コンピュータを、

前記画像内で移動する指標の位置に基づき、前記画像から前記文字候補領域を認識する手段としてさらに機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は文字認識装置、文字認識方法およびプログラムに関する。

【背景技術】

【0002】

視覚障害者は、主に点字や音声といった情報によって周囲から情報を取得している。しかし、商品の容器や張り紙など、生活において重要な情報を含む文字情報には、これらの情報が備わっていないことが多い。

【0003】

点字や音声の情報を備えない文字情報から、音声情報を作成するシステムとして、例えば特許文献 1 や非特許文献 1 に記載の装置がある。特許文献 1 に記載のシステムは、装置への入力情報とする画像の中で、ある特性をもつ色に囲まれた領域を切り出し、その領域に存在する文字を抽出する。そして抽出した文字を文字コードに変換して音声合成部で読み上げるものである。

【0004】

非特許文献 1 に記載の装置は、入力画像上の文字を指先でなぞり、なぞった部分にある文字を認識して読み上げるものである。

【先行技術文献】

【特許文献】

【0005】

【特許文献 1】特開平 4 - 2 4 8 8 5 号公報

【非特許文献】

【0006】

【非特許文献 1】淵田正隆、中村明生 著、「指先指定文字抽出の検討」、第 17 回画像

10

20

30

40

50

センシングシンポジウム、横浜、2011年6月、IS3-08-1~IS3-08-4

【発明の概要】

【発明が解決しようとする課題】

【0007】

しかしながら、特許文献1では、当該装置用に予めマークが施された資料でなければ、文字を抽出することができず、一般に流通する商品や張り紙のように、予めこのようなマークをつけることが難しいものでは、文字情報は認識できない。

【0008】

また、非特許文献1に記載の文字認識装置では、2値化した原画像全体に膨張処理を施して繋がった領域を文字領域として認識する。しかし、原画像全体に膨張処理を施してしまうと文字でない背景領域のノイズが大きくなるため、文字領域として誤って認識されてしまい、精度が落ちる恐れがある。

【0009】

本発明の目的は、上記問題に鑑み、指先でなぞった部分の文字をより高い精度で認識する、汎用的な文字認識装置、文字認識方法およびプログラムを提供することにある。

【課題を解決するための手段】

【0010】

本発明によれば、

2値化された画像から、1つの文字と推定される領域である文字候補領域を認識する文字候補認識手段と、

前記文字候補認識手段が認識した前記文字候補領域毎に、文字と推定される色を有する領域を膨張させる膨張処理を施し、前記膨張処理を施した前記文字候補領域同士で繋がった領域を文字列領域として認識する文字列領域認識手段と、

前記文字列領域に基づき、前記画像から文字列を抽出する文字列抽出手段を有する文字認識装置が提供される。

【0011】

本発明によれば、

コンピュータが、

2値化された画像から、1つの文字と推定される領域である文字候補領域を認識し、

前記文字候補領域毎に、文字と推定される色を有する領域を膨張させる膨張処理を施し、前記膨張処理を施した前記文字候補領域同士で繋がった領域を文字列領域として認識し、

前記文字列領域に基づき、前記画像から文字列を抽出する文字認識方法が提供される。

【0012】

本発明によれば、

コンピュータを、

2値化された画像から、1つの文字と推定される領域である文字候補領域を認識する手段、

前記文字候補領域毎に、文字と推定される色を有する領域を膨張させる膨張処理を施し、前記膨張処理を施した前記文字候補領域同士で繋がった領域を文字列領域として認識する手段、

前記文字列領域に基づき、前記画像から文字列を抽出する手段として機能させるためのプログラムが提供される。

【発明の効果】

【0013】

本発明によれば、指先でなぞった部分の文字をより高い精度で認識することができる、汎用的な文字認識装置、文字認識方法およびプログラムが提供される。

【図面の簡単な説明】

【0014】

【図1】本発明の第1の実施形態に係る文字認識装置の構成を示すブロック図である。

10

20

30

40

50

【図 2】本発明の第 1 の実施形態に係る文字認識装置の処理の流れを示すフローチャートである。

【図 3】生成される矩形の例を示す図である。

【図 4】矩形の配置例を示す図である。

【図 5】矩形の拡大方法を説明する図である。

【図 6】文字候補領域を認識する流れを説明する図である。

【図 7】文字候補認識部が文字列領域部に送信する情報の例を示す図である。

【図 8】文字列領域認識部が送信する文字列領域の例を示す図である。

【図 9】文字列抽出部が抽出する文字列の例を示す図である。

【図 10】文字と推定される色を決定する流れを示す図である。

10

【発明を実施するための形態】

【0015】

以下、本発明の実施の形態について、図面を用いて説明する。尚、すべての図面において、同様な構成要素には同様の符号を付し、適宜説明を省略する。

【0016】

(第 1 の実施形態)

図 1 は、本発明の第 1 の実施形態に係る文字認識装置 10 の構成を示すブロック図である。文字認識装置 10 は、文字候補認識部 102 と、文字列領域認識部 104 と、文字列抽出部 106 を有する。

【0017】

20

文字候補認識部 102 は、1 つの文字と推定される領域である文字候補領域を認識する。

【0018】

文字列領域認識部 104 は、文字候補認識部 102 が認識した文字候補領域毎に、文字と推定される色を有する領域を膨張させる膨張処理を施し、膨張処理を施した文字候補領域同士で繋がった領域を、文字列領域 112 として認識する。

【0019】

文字列抽出部 106 は、文字列領域 112 に基づき、画像から文字を抽出する。

【0020】

なお、各図に示した文字認識装置 10 の各構成要素は、ハードウェア単位の構成ではなく、機能単位のブロックを示している。文字認識装置 10 の各構成要素は、任意のコンピュータの CPU、メモリ、メモリにロードされた本図の構成要素を実現するプログラム、そのプログラムを格納するハードディスクなどの記憶メディア、ネットワーク接続用インタフェースを中心にハードウェアとソフトウェアの任意の組合せによって実現される。そして、その実現方法、装置には様々な変形例がある。

30

【0021】

本実施形態における処理の流れを、図 2 ~ 9 を用いて説明する。

【0022】

まず、文字候補認識部 102 は、2 値化された画像（以下、2 値化画像とする。）を取得する（S102）。そして、文字認識手段 102 は、2 値化画像から文字候補領域の認識処理を行う（S104）。

40

【0023】

まず、文字候補認識部 102 は、図 3 に示すように、画像の中の指標 108 に基づき文字候補領域を特定するための矩形 110 を 2 値化画像上に生成する。ここで、矩形 110 と画像上の文字との関係は、例えば、図 4 (a) に示すように、文字候補領域が矩形 110 内に収まる場合と、図 4 (b)、(c) に示すように、文字候補領域が矩形 110 内に収まらない場合とが考えられる。文字候補認識部 102 は、図 4 (a) のような場合は、矩形 110 をそのまま文字候補領域としてよいが、図 4 (b)、(c) のような場合は、図 5 に示すように矩形 110 を拡大して文字候補領域を特定する。以下で矩形 110 を拡大する処理の流れについて説明する。

50

## 【 0 0 2 4 】

文字候補認識部 1 0 2 は、図 5 に示すように矩形 1 1 0 の外周を走査し、矩形 1 1 0 の外周と接する画素が文字と推定される色であるか否かを確認する。文字と推定される色は、2 値化画像のどちらの色が文字と推定される色かを示す情報を、図示しない記憶部で予め記憶しておくことなどで判断できる。文字と推定される色の画素が存在した場合、文字候補認識部 1 0 2 は、矩形 1 1 0 を一定量拡大する。ここで、文字候補認識部 1 0 2 は、外周の縦方向を走査していた場合は横方向の幅を拡大し、外周の横方向を走査していた場合は縦方向の幅を拡大する。例えば、図 5 の A 地点では、文字候補認識部 1 0 2 は、矩形 1 1 0 の横幅を左方向に一定量拡大する。文字候補認識部 1 0 2 は、文字と推定される色の画素が矩形 1 1 0 の外周に存在しなくなるまでこの処理を繰り返す。そして、拡大した

10

## 【 0 0 2 5 】

文字候補認識部 1 0 2 は、図示しない記憶部で記憶する、指標 1 0 8 が通過した位置の座標に基づき、上述の処理を繰り返し実行して複数の文字候補領域を特定する。最終的に、文字候補認識部 1 0 2 は、例えば図 6 に示すように、複数の文字候補領域を特定する。文字候補認識部 1 0 2 は、特定したすべての文字候補領域を文字列領域認識部 1 0 4 へ送信する。図 7 は、文字列領域認識部 1 0 4 へ送信される情報の例を示す図である。

20

## 【 0 0 2 6 】

文字列領域認識部 1 0 4 は、文字候補認識部 1 0 2 より受信した情報に基づき、文字列領域 1 1 2 を認識する ( S 1 0 6 )。まず、文字列領域認識部 1 0 4 は、認識した各文字候補領域において、文字と推定される色の領域を膨張させる膨張処理を施す。そして、文字列領域認識部 1 0 4 は、各文字候補領域同士でいくつか形成される連結領域の中から、ノイズを除去して文字列領域 1 1 2 を特定する。文字列領域認識部 1 0 4 は、連結領域と対応する文字候補領域の形状に基づき、当該連結領域がノイズか否かを判別する。文字列領域認識部 1 0 4 は、当該連結領域と対応する文字候補領域の上下左右の対称性、2 値化した色の密度比、2 値化した色で形成される形状の複雑さや大きさなどの情報に基づき、

30

## 【 0 0 2 7 】

なお、元となる 2 値化画像が傾いていると、文字候補領域は傾いた状態で連結される可能性があるため、文字列領域認識手段 1 0 4 は、ノイズを除去する前に、連結領域の慣性主軸を計算して傾き角度を求め、アフィン変換などで傾きを補正してもよい。傾きを補正することで、文字列領域認識手段 1 0 4 が文字列とノイズを判別する精度が向上する。

40

## 【 0 0 2 8 】

文字列抽出部 1 0 6 は、文字列領域認識部 1 0 4 から受信した文字列領域 1 1 2 に基づき、2 値化された画像から文字列を抽出する ( S 1 0 8 )。例えば、文字列領域 1 1 2 の座標値を受信した場合、文字列抽出部 1 0 6 は、その座標値に該当する領域を文字列として抽出する。あるいは、マスク情報を受信した場合、2 値化された画像とマスク情報の論理積により、該当する領域を文字列として抽出する。図 9 は、文字列抽出部 1 0 6 が抽出する文字列の一例を示す図である。

## 【 0 0 2 9 】

以上、本実施形態において、指標の通過した座標値に基づいて 1 つの文字と推定される文字候補領域を認識し、各文字候補領域に膨張処理を施し、連結した領域の中から文字列

50

領域 1 1 2 を認識する構成を取る。本構成により、背景のノイズを抑えて文字列を精度良く抽出することができる。

【 0 0 3 0 】

なお、本実施形態において、例えば指先そのものを指標とすることもできるし、指先に特定波長の光を反射する物質をつけ、その物資を指標とすることもできる。

【 0 0 3 1 】

( 第 2 の実施形態 )

本実施形態は、第 1 の実施形態で認識した文字情報を音声として読み上げる構成を取る。

【 0 0 3 2 】

本実施形態では、第 1 の実施形態の文字認識装置 1 0 に、図示しない画像入力部および音声合成部をさらに有する。画像入力部は、カメラなどの入力装置からカラー画像を取得する。音声合成部は、文字候補認識部 1 0 2、文字列領域認識部 1 0 4、文字列抽出部 1 0 6 を用いて、商品パッケージの表面などから抽出した文字列をテキスト化し、音声として読み上げる。

【 0 0 3 3 】

画像入力部は、例えば小型カメラなどから、処理対象とするカラー画像を原画像として取得する。そして、画像入力部は、原画像を「文字と推定される色」と「文字と推定される色以外の色」とで構成される 2 値化画像に変換する。図 1 0 は、文字と推定される色を決定する流れを示す図である。画像入力部は、原画像に平滑化処理を施した画像のヒストグラムを生成し、極大値、極小値などを基準に、ヒストグラムの切り分け位置を決定する。そして、切り分け位置に基づいた色空間を作成する。画像入力部は、指標 1 0 8 の座標値に基づき、矩形 1 1 0 を生成し、矩形 1 1 0 内に存在する色を先の色空間に投影する。そして、画像入力部は、色空間内で最も投票数が多いマスに該当する色を、「文字と推定される色」の候補とする。そして、画像入力部は、候補とした「文字と推定される色」で 2 値化を行い、矩形 1 1 0 内の「文字と推定される色」の密度や、「文字と推定される色」の領域の個数などを判定基準とし、矩形 1 1 0 の「文字と推定される色」を決定する。最も投票数が多い色が判定基準を満たさない場合、画像入力部は、次に投票数が多いマスに該当する色で同様の判定を行う。そして、画像入力部は、「文字と推定される色」と「文字と推定される色以外の色」とで 2 値化した画像を作成して、文字候補認識手段 1 0 2

【 0 0 3 4 】

文字候補認識手段 1 0 2 が 2 値化画像を受信してから文字列を抽出するまでの処理の流れは、第 1 の実施形態と同様のため省略する。文字抽出部 1 0 6 は、抽出した文字列を音声合成部へ送信する。

【 0 0 3 5 】

音声合成部は、文字抽出部 1 0 6 から受信した文字列を O C R ( Optical Character Reader ) などによってテキスト化し、既存の読み上げソフトなどを用いて音声情報として読みあげる。

【 0 0 3 6 】

以上、本実施形態において、さらに画像入力部と音声合成部を有する。本構成により、画像入力部で取得した画像内の文字情報を、第 1 の実施形態のとおり認識して、音声合成部で音声情報として提供することが可能となる。本構成により、例えば商品パッケージ上の商品名を読み上げるなど、視覚障害者の生活を支援する装置を提供することができる。

【 0 0 3 7 】

以上、図面を参照して本発明の実施形態について述べたが、これらは本発明の例示であり、上記以外の様々な構成を採用することもできる。

【 符号の説明 】

【 0 0 3 8 】

1 0 文字認識装置

10

20

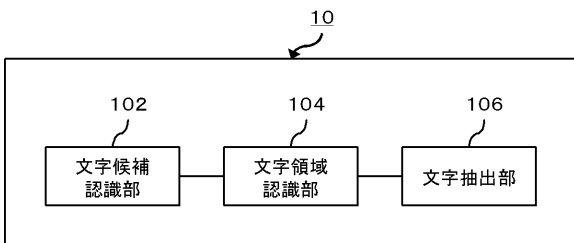
30

40

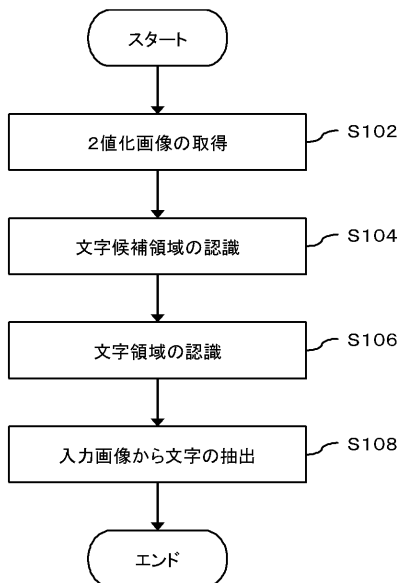
50

- 102 文字候補認識部
- 104 文字列領域認識部
- 106 文字列抽出部
- 108 指標
- 110、110' 矩形
- 112 文字列領域

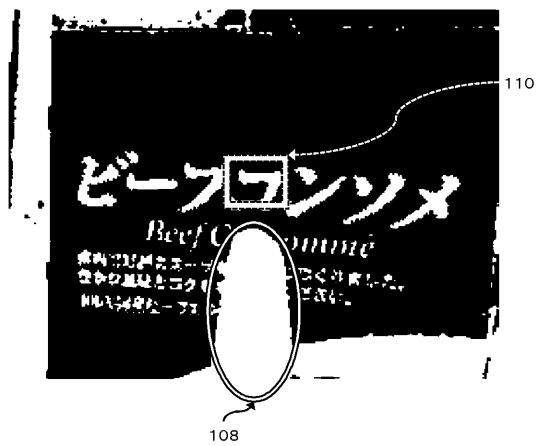
【図1】



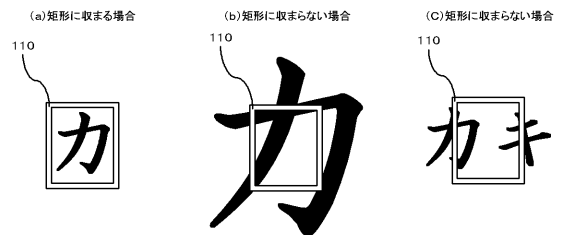
【図2】



【図3】

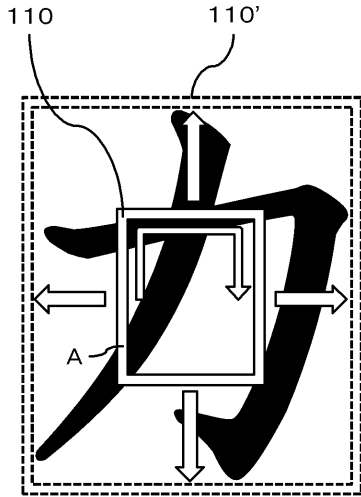


【図4】

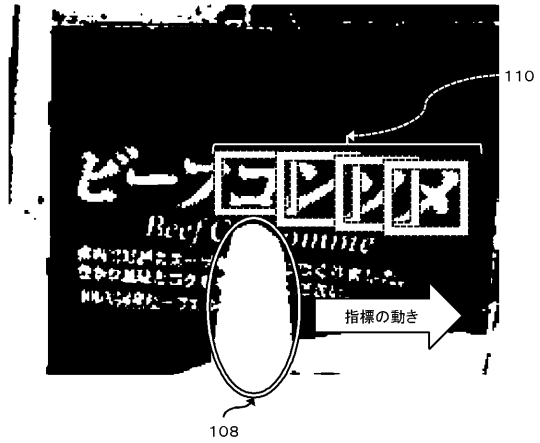




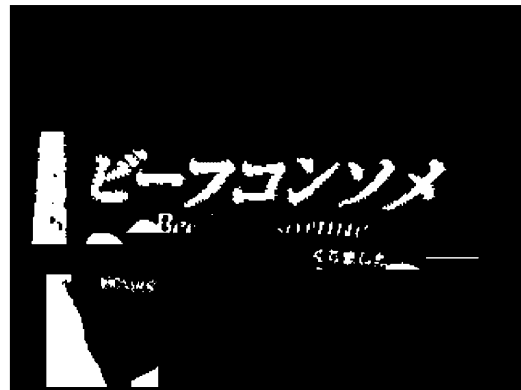
【図5】



【図6】



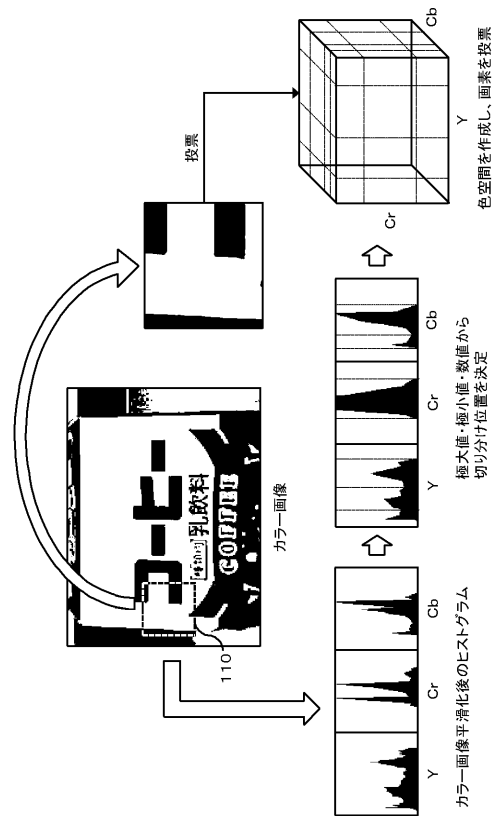
【図7】



【図8】



【図10】



【図9】

