

(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2009-282986
(P2009-282986A)

(43) 公開日 平成21年12月3日(2009.12.3)

(51) Int.Cl. F I テーマコード (参考)
G06K 9/62 (2006.01) G06K 9/62 610C 5B064

審査請求 有 請求項の数 14 O L (全 51 頁)

(21) 出願番号 特願2009-123421 (P2009-123421)
(22) 出願日 平成21年5月21日 (2009. 5. 21)
(31) 優先権主張番号 12/126, 561
(32) 優先日 平成20年5月23日 (2008. 5. 23)
(33) 優先権主張国 米国 (US)

(71) 出願人 000005049
シャープ株式会社
大阪府大阪市阿倍野区長池町2番2号
(74) 代理人 110000338
特許業務法人原謙三国際特許事務所
(72) 発明者 アーメット マフィット ファーマン
アメリカ合衆国 ワシントン州 9860
7, カマス, ノースウェスト パシフィッ
ク リム プールバード 5750 シャ
ープ ラボラトリーズ オブ アメリカ
インコーポレイテッド内

最終頁に続く

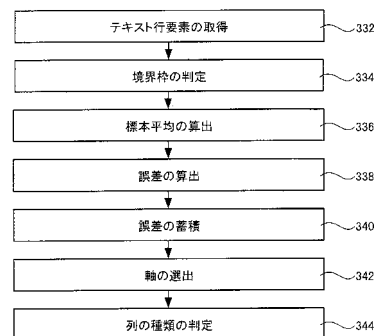
(54) 【発明の名称】 デジタル画像内の数字を検知する方法

(57) 【要約】

【課題】 デジタル画像からアラビア数字を容易に抽出することが可能な方法を提供する。

【解決手段】 画像データから、文字列を抽出し(332)、当該文字列に含まれる非背景画素の連結成分の各々について外接矩形枠を決定する(334)。そして、文字列に含まれる複数の連結成分について、外接矩形枠の各辺の座標の変動量として、平均値からの誤差量の平均を求める(336, 338, 340)。そして、当該誤差量の平均に基づいて、文字列がアラビア数字からなる数字列の候補であるか否かを判定する(344)。

【選択図】 図 2 1



【特許請求の範囲】

【請求項 1】

デジタル画像内の数字連結成分を検知する方法であって、

a) デジタル画像内の文字列の候補である文字列要素を取得する工程と、

b) 上記文字列要素に含まれる非背景画素の連結成分の各々について、所定の第 1 軸に平行な第 1 辺および第 2 辺と当該第 1 軸に垂直な第 2 軸に平行な第 3 辺および第 4 辺とを有する外接矩形枠を決定する工程と、

c) 上記連結成分の各々について、上記第 1 辺の上記第 2 軸上の座標である第 1 辺座標と、上記第 2 辺の上記第 2 軸上の座標である第 2 辺座標と、上記第 3 辺の上記第 1 軸上の座標である第 3 辺座標と、上記第 4 辺の上記第 1 軸上の座標である第 4 辺座標とを求める工程と、

10

d) 上記文字列要素に含まれる複数の上記連結成分について、上記第 1 辺座標の変動量を示す第 1 変動量と、上記第 2 辺座標の変動量を示す第 2 変動量と、上記第 3 辺座標の変動量を示す第 3 変動量と、上記第 4 辺座標の変動量を示す第 4 変動量とを求める変動量算出工程と、

e) 上記第 1 変動量と上記第 2 変動量との和が上記第 3 変動量と上記第 4 変動量との和よりも小さい場合、上記第 1 変動量が所定の第 1 閾値よりも小さく、かつ、上記第 2 変動量が所定の第 2 閾値よりも小さいときに、上記文字列要素をアラビア数字からなる文字列の候補である数字列候補として分類し、上記第 1 変動量が上記第 1 閾値以上であるか、または、上記第 2 変動量が上記第 2 閾値以上であるときに、上記文字列要素をアルファベットの

20

大文字と小文字との混合文字列として分類し、
上記第 1 変動量と上記第 2 変動量との和が上記第 3 変動量と上記第 4 変動量との和よりも大きい場合、上記第 3 変動量が所定の第 3 閾値よりも小さく、かつ、上記第 4 変動量が所定の第 4 閾値よりも小さいときに、上記文字列要素をアラビア数字からなる文字列の候補である数字列候補として分類し、上記第 3 変動量が上記第 3 閾値以上であるか、または、上記第 4 変動量が上記第 4 閾値以上であるときに、上記文字列要素をアルファベットの大文字と小文字との混合文字列として分類する数字列候補分類工程と、
を含む方法。

【請求項 2】

上記変動量算出工程は、

30

上記文字列要素に含まれる複数の連結成分の第 1 辺座標の代表値である第 1 代表値と、当該複数の連結成分の第 2 辺座標の代表値である第 2 代表値と、当該複数の連結成分の第 3 辺座標の代表値である第 3 代表値と、当該複数の連結成分の第 4 辺座標の代表値である第 4 代表値とを求める工程と、

上記第 1 変動量として、上記複数の連結成分の各々の第 1 辺座標と上記第 1 代表値との誤差量の平均値を算出し、上記第 2 変動量として、上記複数の連結成分の各々の第 2 辺座標と上記第 2 代表値との誤差量の平均値を算出し、上記第 3 変動量として、上記複数の連結成分の各々の第 3 辺座標と上記第 3 代表値との誤差量の平均値を算出し、上記第 4 変動量として、上記複数の連結成分の各々の第 4 辺座標と上記第 4 代表値との誤差量の平均値を算出する工程とを含む、

40

請求項 1 に記載のデジタル画像内の数字連結成分を検知する方法。

【請求項 3】

上記第 1 代表値は、上記複数の連結成分の第 1 辺座標の平均値であり、

上記第 2 代表値は、上記複数の連結成分の第 2 辺座標の平均値であり、

上記第 3 代表値は、上記複数の連結成分の第 3 辺座標の平均値であり、

上記第 4 代表値は、上記複数の連結成分の第 4 辺座標の平均値である、

請求項 2 に記載のデジタル画像内の数字連結成分を検知する方法。

【請求項 4】

上記誤差量は、平均絶対誤差、二乗平均誤差、二乗平均平方根誤差から選ばれる誤差量である、請求項 2 または 3 に記載のデジタル画像内の数字連結成分を検知する方法。

50

【請求項 5】

上記数字列候補分類工程にて分類された数字列候補に含まれる複数の連結成分の各々について、当該連結成分の上記外接矩形枠の縦横比を算出する工程と、

上記複数の連結成分の上記縦横比の平均値を算出する工程と、

上記複数の連結成分の上記縦横比の分散値または標準偏差を算出する工程と、

上記縦横比の平均値が所定の第 5 閾値よりも大きく、かつ、上記縦横比の分散値または標準偏差が所定の第 6 閾値よりも小さい場合に、上記数字列候補を、アラビア数字からなる数字列として決定し、上記縦横比の平均値が上記第 5 閾値以下である、もしくは、上記縦横比の分散値または標準偏差が上記第 6 閾値以上である場合に、上記数字列候補を、非数字文字を含む文字列として決定する工程と、をさらに含む、請求項 1 から 4 のいずれか 1 項に記載のデジタル画像内の数字連結成分を検知する方法。

10

【請求項 6】

上記数字列候補分類工程にて分類された数字列候補に含まれる複数の連結成分の各々について、当該連結成分の上記外接矩形枠の縦横比を算出する工程と、

上記複数の連結成分の上記縦横比の平均値を算出する工程と、

上記複数の連結成分の上記縦横比の分散値または標準偏差を算出する工程と、

上記縦横比の平均値が所定の第 5 閾値よりも大きく、かつ、上記縦横比の分散値または標準偏差が所定の第 6 閾値よりも小さく、かつ、上記数字列候補に含まれる連結成分の個数が所定の第 7 閾値より小さい場合に、上記数字列候補を、アラビア数字からなる数字列として決定し、上記縦横比の平均値が上記第 5 閾値以下である、もしくは、上記縦横比の分散値または標準偏差が上記第 6 閾値以上である、もしくは、上記数字列候補に含まれる連結成分の個数が上記第 7 閾値以上である場合に、上記数字列候補を、非数字文字を含む文字列として決定する工程と、をさらに含む、請求項 1 から 4 のいずれか 1 項に記載のデジタル画像内の数字連結成分を検知する方法。

20

【請求項 7】

上記数字列候補分類工程にて分類された数字列候補に含まれる連結成分の個数が所定の第 7 閾値より小さい場合に、上記数字列候補をアラビア数字からなる数字列として決定し、上記数字列候補に含まれる連結成分の個数が上記第 7 閾値以上である場合に、上記数字列候補を、非数字文字を含む文字列として決定する工程、をさらに含む、請求項 1 から 4 のいずれか 1 項に記載のデジタル画像内の数字連結成分を検知する方法。

30

【請求項 8】

デジタル画像内の数字連結成分を検知する方法であって、

a) デジタル画像内の文字列の候補である文字列要素を取得する工程と、

b) 上記文字列要素に含まれる非背景画素の連結成分の各々について、当該連結成分の上記外接矩形枠の縦横比を算出する工程と、

c) 上記複数の連結成分の上記縦横比の平均値を算出する工程と、

d) 上記縦横比の平均値が所定の第 1 閾値よりも大きい場合に、上記文字列要素を、アラビア数字からなる数字列として決定し、上記縦横比の平均値が上記第 1 閾値以下である場合に、上記文字列要素を、非数字文字を含む文字列として決定する工程と、を含むデジタル画像内の数字連結成分を検知する方法。

40

【請求項 9】

デジタル画像内の数字連結成分を検知する方法であって、

a) デジタル画像内の文字列の候補である文字列要素を取得する工程と、

b) 上記文字列要素に含まれる非背景画素の連結成分の各々について、当該連結成分の上記外接矩形枠の縦横比を算出する工程と、

c) 上記複数の連結成分の上記縦横比の分散値または標準偏差を算出する工程と、

d) 上記縦横比の分散値または標準偏差が所定の第 2 閾値よりも小さい場合に、上記文字列要素を、アラビア数字からなる数字列として決定し、上記縦横比の分散値または標準偏差が上記第 2 閾値以上である場合に、上記文字列要素を、非数字文字を含む文字列として決定する工程と、

50

を含むデジタル画像内の数字連結成分を検知する方法。

【請求項 10】

デジタル画像内の数字連結成分を検知する方法であって、

a) デジタル画像内の文字列の候補である文字列要素を取得する工程と、
b) 上記文字列要素に含まれる非背景画素の連結成分の各々について、当該連結成分の上記外接矩形枠の縦横比を算出する工程と、

c) 上記複数の連結成分の上記縦横比の平均値を算出する工程と、

d) 上記複数の連結成分の上記縦横比の分散値または標準偏差を算出する工程と、

e) 上記縦横比の平均値が所定の第 1 閾値よりも大きく、かつ、上記縦横比の分散値または標準偏差が所定の第 2 閾値よりも小さい場合に、上記文字列要素を、アラビア数字からなる数字列として決定し、上記縦横比の平均値が上記第 1 閾値以下である、もしくは、上記縦横比の分散値または標準偏差が上記第 2 閾値以上である場合に、上記文字列要素を、非数字文字を含む文字列として決定する工程と、

を含むデジタル画像内の数字連結成分を検知する方法。

【請求項 11】

デジタル画像内の数字連結成分を検知する方法であって、

a) デジタル画像内の文字列の候補である文字列要素を取得する工程と、

b) 上記文字列要素に含まれる非背景画素の連結成分の各々について、当該連結成分の上記外接矩形枠の縦横比を算出する工程と、

c) 上記複数の連結成分の上記縦横比の平均値を算出する工程と、

d) 上記縦横比の平均値が所定の第 1 閾値よりも大きく、かつ、上記文字列要素に含まれる連結成分の個数が所定の第 3 閾値より小さい場合に、上記文字列要素を、アラビア数字からなる数字列として決定し、上記縦横比の平均値が上記第 1 閾値以下である、もしくは、上記文字列要素に含まれる連結成分の個数が上記第 3 閾値以上である場合に、上記文字列要素を、非数字文字を含む文字列として決定する工程と、

を含むデジタル画像内の数字連結成分を検知する方法。

【請求項 12】

デジタル画像内の数字連結成分を検知する方法であって、

a) デジタル画像内の文字列の候補である文字列要素を取得する工程と、

b) 上記文字列要素に含まれる非背景画素の連結成分の各々について、当該連結成分の上記外接矩形枠の縦横比を算出する工程と、

c) 上記複数の連結成分の上記縦横比の分散値または標準偏差を算出する工程と、

d) 上記縦横比の分散値または標準偏差が所定の第 2 閾値よりも小さく、かつ、上記文字列要素に含まれる連結成分の個数が所定の第 3 閾値より小さい場合に、上記文字列要素を、アラビア数字からなる数字列として決定し、上記縦横比の分散値または標準偏差が上記第 2 閾値以上である、もしくは、上記文字列要素に含まれる連結成分の個数が上記第 3 閾値以上である場合に、上記文字列要素を、非数字文字を含む文字列として決定する工程と、

を含むデジタル画像内の数字連結成分を検知する方法。

【請求項 13】

デジタル画像内の数字連結成分を検知する方法であって、

a) デジタル画像内の文字列の候補である文字列要素を取得する工程と、

b) 上記文字列要素に含まれる非背景画素の連結成分の各々について、当該連結成分の上記外接矩形枠の縦横比を算出する工程と、

c) 上記複数の連結成分の上記縦横比の平均値を算出する工程と、

d) 上記複数の連結成分の上記縦横比の分散値または標準偏差を算出する工程と、

e) 上記縦横比の平均値が所定の第 1 閾値よりも大きく、かつ、上記縦横比の分散値または標準偏差が所定の第 2 閾値よりも小さく、かつ、上記文字列要素に含まれる連結成分の個数が所定の第 3 閾値より小さい場合に、上記文字列要素を、アラビア数字からなる数字列として決定し、上記縦横比の平均値が上記第 1 閾値以下である、もしくは、上記縦横

10

20

30

40

50

比の分散値または標準偏差が上記第2閾値以上である、もしくは、上記文字列要素に含まれる連結成分の個数が上記第3閾値以上である場合に、上記文字列要素を、非数字文字を含む文字列として決定する工程と、
を含むデジタル画像内の数字連結成分を検知する方法。

【請求項14】

デジタル画像内の数字連結成分を検知する方法であって、

a) デジタル画像内の文字列の候補である文字列要素を取得する工程と、

b) 上記文字列要素に含まれる非背景画素の連結成分の個数が所定の第1閾値より小さい場合に、上記文字列要素を、アラビア数字からなる数字列として決定し、上記文字列要素に含まれる連結成分の個数が上記第1閾値以上である場合に、上記文字列要素を、非数字文字を含む文字列として決定する工程と、
を含むデジタル画像内の数字連結成分を検知する方法。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、デジタル画像内の数字を検知する方法及びシステム、特に、連結成分を数字と非数字とに分類する方法及びシステムが含まれる。

【背景技術】

【0002】

数字などの文字種別を判定する方法としては、特許文献1・2などに開示されている。例えば、特許文献1には、文字媒体の行領域を行方向と垂直な方向に分割し、分割領域より特徴量（例えば、英文字列のアッセンダーが存在する領域、アッセンダーを除く残りの領域における文字線量）を抽出し、抽出された特徴量に基づいて、文字行の文字種（日本文字列、英文字列）の判別を行うことが記載されている。

20

【0003】

また、特許文献2には、文書画像を縮小した後、黒画素連結成分を抽出し、それらを統合して文字領域を生成する。生成した文字領域について、連結成分の長さを基にその成分を分類し、分類結果の集計値を基に日本語領域であるか英語領域であるかを判別することが記載されている。

【先行技術文献】

30

【特許文献】

【0004】

【特許文献1】特開平3-290773号公報（1991年12月20日公開）

【特許文献2】特開2006-107527号公報（2006年4月20日公開）

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、特許文献1・2に記載の技術では、日本語文字か英語文字かを判別するものであり、アラビア数字を抽出するものではない。

【0006】

40

数字の検知は、光学式文字認識（OCR）やその他の原稿処理分析システム及び方法に不可欠である。

【0007】

本発明は、上記課題を解決するためになされたものであり、デジタル画像からアラビア数字を容易に抽出することが可能な方法を提供することを目的とする。

【課題を解決するための手段】

【0008】

本発明に係る方法は、デジタル画像内の数字連結成分を検知する方法であって、a) デジタル画像内の文字列の候補である文字列要素を取得する工程と、b) 上記文字列要素に含まれる非背景画素の連結成分の各々について、所定の第1軸に平行な第1辺および第2

50

辺と当該第1軸に垂直な第2軸に平行な第3辺および第4辺とを有する外接矩形枠を決定する工程と、c)上記連結成分の各々について、上記第1辺の上記第2軸上の座標である第1辺座標と、上記第2辺の上記第2軸上の座標である第2辺座標と、上記第3辺の上記第1軸上の座標である第3辺座標と、上記第4辺の上記第1軸上の座標である第4辺座標とを求める工程と、d)上記文字列要素に含まれる複数の上記連結成分について、上記第1辺座標の変動量を示す第1変動量と、上記第2辺座標の変動量を示す第2変動量と、上記第3辺座標の変動量を示す第3変動量と、上記第4辺座標の変動量を示す第4変動量とを求める変動量算出工程と、e)上記第1変動量と上記第2変動量との和が上記第3変動量と上記第4変動量との和よりも小さい場合、上記第1変動量が所定の第1閾値よりも小さく、かつ、上記第2変動量が所定の第2閾値よりも小さいときに、上記文字列要素をアラビア数字からなる文字列の候補である数字列候補として分類し、上記第1変動量が上記第1閾値以上であるか、または、上記第2変動量が上記第2閾値以上であるときに、上記文字列要素をアルファベットの大文字と小文字との混合文字列として分類し、上記第1変動量と上記第2変動量との和が上記第3変動量と上記第4変動量との和よりも大きい場合、上記第3変動量が所定の第3閾値よりも小さく、かつ、上記第4変動量が所定の第4閾値よりも小さいときに、上記文字列要素をアラビア数字からなる文字列の候補である数字列候補として分類し、上記第3変動量が上記第3閾値以上であるか、または、上記第4変動量が上記第4閾値以上であるときに、上記文字列要素をアルファベットの大文字と小文字との混合文字列として分類する数字列候補分類工程と、を含む。

10

【0009】

20

アルファベットの大文字と小文字とが混合する文字列の場合、当該文字列に含まれる文字の上部の位置または下部の位置は、大文字と小文字とで異なるし、小文字の中でも「a」「g」「b」で異なることとなる。そのため、アルファベットの大文字と小文字とが混在する場合、当該文字列に含まれる文字の上部の位置と下部の位置との変動量は大きくなる。一方、アラビア数字からなる数字列である場合、当該数字列に含まれる文字の上部の位置または下部の位置は一定である。本発明はこの点に着目したものである。

【0010】

上記の構成によれば、文字列要素に含まれる非背景画素の連結成分の各々について、所定の第1軸に平行な第1辺および第2辺と当該第1軸に垂直な第2軸に平行な第3辺および第4辺とを有する外接矩形枠を決定する工程と、上記連結成分の各々について、上記第1辺の上記第2軸上の座標である第1辺座標と、上記第2辺の上記第2軸上の座標である第2辺座標と、上記第3辺の上記第1軸上の座標である第3辺座標と、上記第4辺の上記第1軸上の座標である第4辺座標とを求める工程と、上記文字列要素に含まれる複数の上記連結成分について、上記第1辺座標の変動量を示す第1変動量と、上記第2辺座標の変動量を示す第2変動量と、上記第3辺座標の変動量を示す第3変動量と、上記第4辺座標の変動量を示す第4変動量とを求める変動量算出工程と、を含む。

30

【0011】

文字列要素は、文字が第1軸の方向に並んでいる場合と、文字が第2軸の方向に並んでいる場合とがある。ただし、文字が第1軸方向に並んでいる場合、文字列要素に含まれる連結成分の各々の第3辺座標および第4辺座標は、大きく変化することとなる。そのため、第3変動量および第4変動量は大きな値をとる。また、文字が第2軸方向に並んでいる場合、文字列要素に含まれる連結成分の各々の第1辺座標および第2辺座標は、大きく変化することとなる。そのため、第1変動量および第2変動量は大きな値をとる。よって、第1変動量と第2変動量との和が第3変動量と第4変動量との和よりも小さい場合、文字列要素に含まれる連結成分は、第1軸方向に並んでいるものとみなすことができる。また、そこで、第1変動量と第2変動量との和が第3変動量と第4変動量との和よりも大きい場合、文字列要素に含まれる連結成分は、第2軸方向に並んでいるものとみなすことができる。

40

【0012】

そして、上記数字列候補分類工程に従って分類を行う。上述したように、アルファベッ

50

トの大文字と小文字とが混合する文字列と、アラビア数字のみを含む文字列とでは、当該文字列に含まれる文字の上端部分と下端部分との変動量に差がある。そのため、アルファベットとアラビア数字からなる原稿画像を示すデジタル画像データが入力されたとき、上記の数字列候補分類工程によれば、アルファベットの大文字と小文字との混合文字列と、数字列からなる数字列の候補とを精度よく分類することができる。すなわち、デジタル画像からアラビア数字を容易に抽出することができる。

【0013】

また、本発明の方法において、上記変動量算出工程は、上記文字列要素に含まれる複数の連結成分の第1辺座標の代表値である第1代表値と、当該複数の連結成分の第2辺座標の代表値である第2代表値と、当該複数の連結成分の第3辺座標の代表値である第3代表値と、当該複数の連結成分の第4辺座標の代表値である第4代表値とを求める工程と、上記第1変動量として、上記複数の連結成分の各々の第1辺座標と上記第1代表値との誤差量の平均値を算出し、上記第2変動量として、上記複数の連結成分の各々の第2辺座標と上記第2代表値との誤差量の平均値を算出し、上記第3変動量として、上記複数の連結成分の各々の第3辺座標と上記第3代表値との誤差量の平均値を算出し、上記第4変動量として、上記複数の連結成分の各々の第4辺座標と上記第4代表値との誤差量の平均値を算出する工程とを含むことが好ましい。これにより、変動量を容易に算出することができる。

10

【0014】

なお、例えば、上記第1代表値は、上記複数の連結成分の第1辺座標の平均値であり、上記第2代表値は、上記複数の連結成分の第2辺座標の平均値であり、上記第3代表値は、上記複数の連結成分の第3辺座標の平均値であり、上記第4代表値は、上記複数の連結成分の第4辺座標の平均値である。

20

【0015】

また、上記誤差量は、例えば、平均絶対誤差、二乗平均誤差、二乗平均平方根誤差から選ばれる誤差量である。

【0016】

さらに、本発明の方法は、上記数字列候補分類工程にて分類された数字列候補に含まれる複数の連結成分の各々について、当該連結成分の上記外接矩形枠の縦横比を算出する工程と、上記複数の連結成分に対して算出された上記縦横比の平均値を算出する工程と、上記複数の連結成分に対して算出された上記縦横比の分散値または標準偏差を算出する工程と、上記縦横比の平均値が所定の第5閾値よりも大きく、かつ、上記縦横比の分散値または標準偏差が所定の第6閾値よりも小さい場合に、上記数字列候補を、アラビア数字からなる数字列として決定し、上記縦横比の平均値が上記第5閾値以下である、もしくは、上記縦横比の分散値または標準偏差が上記第6閾値以上である場合に、上記数字列候補を、非数字文字を含む文字列として決定する工程と、をさらに含むことが好ましい。

30

【0017】

上記の数字列候補の中には、アラビア数字のみからなる数字列以外にも、アルファベットの大文字のみからなる大文字列も含まれる可能性がある。

【0018】

しかしながら、上記の構成によれば、数字列候補に含まれる複数の連結成分の各々について、当該連結成分の外接矩形枠の縦横比を算出する。そして、縦横比の平均値が所定の第5閾値よりも大きく、かつ、上記縦横比の分散値または標準偏差が所定の第6閾値よりも小さい場合に、上記数字列候補を、アラビア数字からなる数字列として決定し、上記縦横比の平均値が上記第5閾値以下である、もしくは、上記縦横比の分散値または標準偏差が上記第6閾値以上である場合に、上記数字列候補を、非数字文字を含む文字列として決定する。

40

【0019】

ここで、アルファベットの大文字の多くは、アラビア数字と比べて、縦横比が小さい。また、アルファベットの大文字には、縦横比の小さい文字が多数であるが、縦横比の大き

50

な「I」なども含まれる。そのため、アルファベットの大文字からなる大文字列の場合、当該大文字列に含まれる文字の縦横比の分散値または標準偏差は相対的に大きくなる。一方、アラビア数字の0～9は、縦横比がほぼ同じであり、数字列の場合、当該数字列に含まれる文字の縦横比の分散値または標準偏差は相対的に小さい。

【0020】

そのため、上記の構成によれば、大文字からなる大文字列を非数字文字を含む文字列として決定することができ、アラビア数字からある数字列を精度よく検出することができる。

【0021】

もしくは、本発明の方法は、上記数字列候補分類工程にて分類された数字列候補に含まれる複数の連結成分の各々について、当該連結成分の上記外接矩形枠の縦横比を算出する工程と、上記複数の連結成分の上記縦横比の平均値を算出する工程と、上記複数の連結成分の上記縦横比の分散値または標準偏差を算出する工程と、上記縦横比の平均値が所定の第5閾値よりも大きく、かつ、上記縦横比の分散値または標準偏差が所定の第6閾値よりも小さく、かつ、上記数字列候補に含まれる連結成分の個数が所定の第7閾値より小さい場合に、上記数字列候補を、アラビア数字からなる数字列として決定し、上記縦横比の平均値が上記第5閾値以下である、もしくは、上記縦横比の分散値または標準偏差が上記第6閾値以上である、もしくは、上記数字列候補に含まれる連結成分の個数が上記第7閾値以上である場合に、上記数字列候補を、非数字文字を含む文字列として決定する工程と、をさらに含んでもよい。

10

20

【0022】

アラビア数字のみからなる数字列の多くは、当該数字列に含まれる数字の個数が少ない場合が多い。そのため、上記の構成によれば、より精度良く数字列を検出することができる。

【0023】

もしくは、本発明の方法は、上記数字列候補分類工程にて分類された数字列候補に含まれる連結成分の個数が所定の第7閾値より小さい場合に、上記数字列候補をアラビア数字からなる数字列として決定し、上記数字列候補に含まれる連結成分の個数が上記第7閾値以上である場合に、上記数字列候補を非数字文字を含む文字列として決定する工程、をさらに含んでもよい。

30

【0024】

上記の構成によれば、数字列を精度よく、かつ、簡単な処理で検出することができる。

【0025】

また、本発明の方法は、デジタル画像内の数字連結成分を検知する方法であって、a) デジタル画像内の文字列の候補である文字列要素を取得する工程と、b) 上記文字列要素に含まれる非背景画素の連結成分の各々について、当該連結成分の上記外接矩形枠の縦横比を算出する工程と、c) 上記複数の連結成分の上記縦横比の平均値を算出する工程と、d) 上記縦横比の平均値が所定の第1閾値よりも大きい場合に、上記文字列要素を、アラビア数字からなる数字列として決定し、上記縦横比の平均値が上記第1閾値以下である場合に、上記文字列要素を、非数字文字を含む文字列として決定する工程と、を含むものであってもよい。

40

【0026】

また、本発明の方法は、デジタル画像内の数字連結成分を検知する方法であって、a) デジタル画像内の文字列の候補である文字列要素を取得する工程と、b) 上記文字列要素に含まれる非背景画素の連結成分の各々について、当該連結成分の上記外接矩形枠の縦横比を算出する工程と、c) 上記複数の連結成分の上記縦横比の平均値を算出する工程と、d) 上記縦横比の平均値が所定の第1閾値よりも大きく、かつ、上記文字列要素に含まれる連結成分の個数が所定の第3閾値より小さい場合に、上記文字列要素を、アラビア数字からなる数字列として決定し、上記縦横比の平均値が上記第1閾値以下である、もしくは、上記文字列要素に含まれる連結成分の個数が上記第3閾値以上である場合に、上記文字

50

列要素を、非数字文字を含む文字列として決定する工程と、を含むものであってもよい。

【0027】

ここで、アルファベットの大文字および小文字の多くは、アラビア数字と比べて、縦横比が小さい。そのため、アルファベットの大文字からなる大文字列や大文字および小文字の混在文字列の場合、当該文字列に含まれる文字の縦横比の平均値は相対的に小さくなる。一方、アラビア数字の0～9の全ては、縦横比が相対的に大きく、数字列の場合、当該数字列に含まれる文字の縦横比の平均値は相対的に大きくなる。よって上記の構成によれば、容易に数字列を検出することができる。

【0028】

また、本発明の方法は、デジタル画像内の数字連結成分を検知する方法であって、a) デジタル画像内の文字列の候補である文字列要素を取得する工程と、b) 上記文字列要素に含まれる非背景画素の連結成分の各々について、当該連結成分の上記外接矩形枠の縦横比を算出する工程と、c) 上記複数の連結成分の上記縦横比の分散値または標準偏差を算出する工程と、d) 上記縦横比の分散値または標準偏差が所定の第2閾値よりも小さい場合に、上記文字列要素を、アラビア数字からなる数字列として決定し、上記縦横比の分散値または標準偏差が上記第2閾値以上である場合に、上記文字列要素を、非数字文字を含む文字列として決定する工程と、を含むものであってもよい。

10

【0029】

また、本発明の方法は、デジタル画像内の数字連結成分を検知する方法であって、a) デジタル画像内の文字列の候補である文字列要素を取得する工程と、b) 上記文字列要素に含まれる非背景画素の連結成分の各々について、当該連結成分の上記外接矩形枠の縦横比を算出する工程と、c) 上記複数の連結成分の上記縦横比の分散値または標準偏差を算出する工程と、d) 上記縦横比の分散値または標準偏差が所定の第2閾値よりも小さく、かつ、上記文字列要素に含まれる連結成分の個数が所定の第3閾値より小さい場合に、上記文字列要素を、アラビア数字からなる数字列として決定し、上記縦横比の分散値または標準偏差が上記第2閾値以上である、もしくは、上記文字列要素に含まれる連結成分の個数が上記第3閾値以上である場合に、上記文字列要素を、非数字文字を含む文字列として決定する工程と、を含むものであってもよい。

20

【0030】

ここで、アルファベットの大文字および小文字には、縦横比の小さい文字が多数であるが、縦横比の大きな「I」「b」なども含まれる。そのため、アルファベットの大文字からなる大文字列や大文字および小文字の混在文字列の場合、当該文字列に含まれる文字の縦横比の分散値または標準偏差は相対的に大きくなる。一方、アラビア数字の0～9は、縦横比がほぼ同じであり、数字列の場合、当該数字列に含まれる文字の縦横比の分散値または標準偏差は相対的に小さい。そのため上記の構成によれば、容易に数字列を検出することができる。

30

【0031】

また、本発明の方法は、デジタル画像内の数字連結成分を検知する方法であって、a) デジタル画像内の文字列の候補である文字列要素を取得する工程と、b) 上記文字列要素に含まれる非背景画素の連結成分の各々について、当該連結成分の上記外接矩形枠の縦横比を算出する工程と、c) 上記複数の連結成分の上記縦横比の平均値を算出する工程と、d) 上記複数の連結成分の上記縦横比の分散値または標準偏差を算出する工程と、e) 上記縦横比の平均値が所定の第1閾値よりも大きく、かつ、上記縦横比の分散値または標準偏差が所定の第2閾値よりも小さい場合に、上記文字列要素を、アラビア数字からなる数字列として決定し、上記縦横比の平均値が上記第1閾値以下である、もしくは、上記縦横比の分散値または標準偏差が上記第2閾値以上である場合に、上記文字列要素を、非数字文字を含む文字列として決定する工程と、を含むものであってもよい。

40

【0032】

また、本発明の方法は、デジタル画像内の数字連結成分を検知する方法であって、a) デジタル画像内の文字列の候補である文字列要素を取得する工程と、b) 上記文字列要素

50

に含まれる非背景画素の連結成分の各々について、当該連結成分の上記外接矩形枠の縦横比を算出する工程と、c)上記複数の連結成分の上記縦横比の平均値を算出する工程と、d)上記複数の連結成分の上記縦横比の分散値または標準偏差を算出する工程と、e)上記縦横比の平均値が所定の第1閾値よりも大きく、かつ、上記縦横比の分散値または標準偏差が所定の第2閾値よりも小さく、かつ、上記文字列要素に含まれる連結成分の個数が所定の第3閾値より小さい場合に、上記文字列要素を、アラビア数字からなる数字列として決定し、上記縦横比の平均値が上記第1閾値以下である、もしくは、上記縦横比の分散値または標準偏差が上記第2閾値以上である、もしくは、上記文字列要素に含まれる連結成分の個数が上記第3閾値以上である場合に、上記文字列要素を、非数字文字を含む文字列として決定する工程と、を含むものであってもよい。

10

【0033】

上述したように、アルファベットの大文字からなる大文字列や大文字および小文字の混在文字列の場合、当該文字列に含まれる文字の縦横比の平均値は相対的に小さく、分散値または標準偏差は相対的に大きくなる。一方、アラビア数字の0~9は、縦横比がほぼ同じであり、数字列の場合、当該数字列に含まれる文字の縦横比の平均値は相対的に大きく、分散値または標準偏差は相対的に小さい。そのため上記の構成によれば、容易に数字列を検出することができる。

【0034】

また、本発明の方法は、デジタル画像内の数字連結成分を検知する方法であって、a)デジタル画像内の文字列の候補である文字列要素を取得する工程と、b)上記文字列要素に含まれる非背景画素の連結成分の個数が所定の第1閾値より小さい場合に、上記文字列要素を、アラビア数字からなる数字列として決定し、上記文字列要素に含まれる連結成分の個数が上記第1閾値以上である場合に、上記文字列要素を、非数字文字を含む文字列として決定する工程と、を含むものであってもよい。

20

【0035】

上述したように、数字列に含まれる文字数は相対的に小さい場合が多い。そのため上記の構成によれば、デジタル画像からアラビア数字を容易に抽出することができる。

【発明の効果】**【0036】**

本発明によれば、デジタル画像からアラビア数字を容易に抽出することができるという効果を奏する。

30

【図面の簡単な説明】**【0037】**

【図1】図1(A)は、「読みやすい」向きにあるテキスト原稿の一例を示し、図1(B)は、「読みやすい」向きから時計回りに90度回転した向きにあるテキスト原稿の一例を示し、図1(C)は、「読みやすい」向きから時計回りに180度回転した向きにあるテキスト原稿の一例を示し、図1(D)は、「読みやすい」向きから反時計回りに90度回転した向きにあるテキスト原稿の一例を示す。

【図2】図2は、連結成分の候補が数字テンプレートと照合される前に、2値化され、サイズが正規化され、距離算出処理が実行される原稿向き判定処理を含む、本発明の実施形態の例を示す図である。

40

【図3】図3は、推定向きと関連する信頼度の値の判定を含む、本発明の実施形態の例を示す図である。

【図4】図4は、サイズが正規化された2値連結成分を平均化することにより数字テンプレートを作り出す数字テンプレート作成工程を含む、本発明の実施形態の例である。

【図5】図5は、4つの基本の向きにおける10個のアラビア数字のための数字テンプレートの例を示す図である。

【図6】図6は、連結成分の候補をある向きにおける数字テンプレートと比較し、当該向きにおける少なくとも一つの比較結果が十分な一致を示したら、当該向きと関連する一致カウンタのカウンタ数を増やす工程を含む、本発明の実施形態の例を示す図である。

50

【図 7】図 7 は、図 6 に示す本発明の実施形態に従って判定される推定向きと関連する信頼度の値を作り出す工程を含む、本発明の実施形態の例を示す図である。

【図 8】図 8 は、向きの組合せの一致カウンタの率に基づいて信頼度が判定される、本発明の実施形態の例を示す図である。

【図 9】図 9 は、連結成分の候補を複数の向きにおける数字テンプレートと比較し、もっともよく一致した結果が一致の基準を満たしているときは、その一致が得られた向きと関連する一致カウンタのカウント数を増やす工程を含む、本発明の実施形態の一例を示す図である。

【図 10】図 10 は、図 9 に示す本発明の実施形態に従って判定される推定向きと関連する信頼度の値を作り出す工程を含む、本発明の実施形態の例を示す図である。

10

【図 11】図 11 は、2 値連結成分のサイズを正規化し、距離算出処理を行い、平均化することによって数字テンプレートを作り出す数字テンプレート作成工程を含む、本発明の実施形態の例を示す図である。

【図 12】図 12 は、連結成分の候補をある向きにおける数字テンプレートと比較し、当該向きにおける少なくとも一つの比較結果が十分な一致を示したら、当該向きと関連する一致カウンタのカウント数を増やす工程を含む、本発明の実施形態の例を示す図である。

【図 13】図 13 は、図 12 に示す本発明の実施形態に従って判定される推定向きと関連する信頼度の値を作り出す工程を含む、本発明の実施形態の例を示す図である。

【図 14】図 14 は、連結成分の候補を複数の向きにおける数字テンプレートと比較し、もっともよく一致した結果が一致の基準を満たしているときは、その一致が得られた向きと関連する一致カウンタのカウント数を増やす工程を含む、本発明の実施形態の例を示す図である。

20

【図 15】図 15 は、図 14 に示す本発明の実施形態に従って判定される推定向きと関連する信頼度の値を作り出す工程を含む、本発明の実施形態の例を示す図である。

【図 16】図 16 は、連結成分の候補が数字テンプレートと照合される前に 2 値化されサイズを正規化される原稿向き判定を含む、本発明の実施形態の例を示す図である。

【図 17】図 17 は、推定向きと関連する信頼度の値の判定を含む、本発明の実施形態の例を示す図である。

【図 18 A】図 18 A は、テキスト文字とテキスト文字境界枠の例を示す図である。

【図 18 B】図 18 B は、テキストオブジェクトとテキストオブジェクト境界枠の例を示す図である。

30

【図 19】図 19 は、文字列と文字列境界枠の例を示す図である。

【図 20 A】図 20 A は、大文字と小文字からなる文字列の一例を示す図である。

【図 20 B】図 20 B は、大文字のみからなる文字列の一例を示す図である。

【図 20 C】図 20 C は、数字のみからなる文字列の一例を示す図である。

【図 21】図 21 は、文字の境界に関わる標本平均に関連する文字境界変動性に基づく数字列の識別を含む、本発明の実施形態を示す図である。

【図 22】図 22 は、実施形態 1 または 2 における、類似度の求め方を示す図である。

【図 23】図 23 は、実施形態 3 から 4 における、類似度の求め方を示す図である。

【図 24】図 24 は、文字列の標本平均を示す図である。

40

【図 25】原稿方向判定システムの構成を示すブロック図である。

【発明を実施するための形態】

【0038】

本発明のいくつかの実施形態は、デジタル画像内の数字を検知する方法及びシステムを含む。

【0039】

本発明のいくつかの実施形態は、複数の連結成分を含む文字列成分の軸のうち一つの軸の方向に広がる連結成分の広さの変動性に基づいて、デジタル画像内の数字を検知する方法及びシステムを含む。本発明のこれらの実施形態のいくつかでは、上記変動性としては、上記広さの代表的な尺度に関連する誤差量がある。

50

【0040】

本発明のいくつかの実施形態は、文字列成分内の連結成分の縦横比に基づいてデジタル画像内の数字を検知する方法及びシステムを含む。これらの実施形態のいくつかでは、縦横比の第一のモーメントが計算される。これらの実施形態のいくつかでは、縦横比の第二のモーメントが計算される。

【0041】

本発明のいくつかの実施形態には、文字列成分内における連結成分の数に基づいて、デジタル画像内の数字を検知する方法及びシステムが含まれる。

【0042】

本発明のいくつかの実施形態では、連結成分オブジェクトの候補を4つの主要な向きにおける数字テンプレートと照合し、原稿画像の向きを推定する。本発明のいくつかの実施形態では、トレーニングデータを用いて上記数字テンプレートを作成し、上記数字テンプレートは共通の、サイズとフォントが不変の数字の代表例である。

10

【0043】

本発明のいくつかの実施形態では、原稿要素は、サイズが正規化され、距離算出処理が実行され、照合数字テンプレートと照合される。照合結果は累積され、累積された照合結果に基づいて推定向きが作り出される。これらの実施形態のいくつかでは、上記数字テンプレートは、サイズが正規化され平均化された数字の例と関連している。

【0044】

本発明の他の実施形態では、原稿要素は、サイズが正規化され数字テンプレートと照合される。照合結果は蓄積され、蓄積された照合結果に基づいて推定向きが作り出される。これらの実施形態のいくつかでは、上記数字テンプレートは、サイズが正規化され、距離が算出され、平均化された数字の例と関連している。

20

【0045】

本発明のいくつかの実施形態では、信頼度あるいは確実性度は、推定向きと関連して判定される。

【0046】

本発明の上記および他の目的、特徴、利点は、添付の図面とともに以下に述べる説明を考慮することにより、よりよく理解できるであろう。

【0047】

本発明の実施形態は、図面を参照することによってもっともよく理解されるであろう。すべての図面において、同じ部材は同じ番号を振ってある。上記の図面は本実施形態の一部に明確に含まれる。

30

【0048】

本実施形態の図面において概略的に述べられ描かれた本発明の諸要素は、さまざまな異なった構成を取ることができることは容易に理解できるであろう。本発明の方法及びシステムの実施形態を以下により詳細に述べるが、それらは本発明の範囲を限定するものではなく、本発明の現在のところ好ましい実施形態の代表例にすぎない。

【0049】

本発明の実施形態の構成要素はハードウェア、ファームウェア、及び/またはソフトウェアで具体化することができる。ここで述べられる実施形態の例ではこれらの形態のうち一つだけが述べられるけれども、当業者ならば、本発明の範囲内で、これらの形態のいずれにおいても実施形態の構成要素を達成できると考えられる。

40

【0050】

電子原稿画像のページの向き（原稿画像の向きとも見なされる）は、スキャンする方向や、原稿載置台上での元の原稿の向きなどの要因により、元の原稿のページの向きとは一致しない場合がある。元の原稿のページの向きは、名目上のページの向きあるいは読みやすい向きともいう。すなわち、名目上の向き（読みやすい向き）とは、原稿上の文字を読むのに適した向きのことである。電子原稿画像のページの向きと名目上のページの向きとのこうした食い違いは、当該電子原稿を処理する際に、望ましくない結果や、予想外の結

50

果や、最適ではない結果や、その他の不満足な結果をもたらす場合がある。例えば、電子原稿を印刷したものに最終的な処理を行う際、原稿の向きが違っていると、望ましくない結果が生じる場合がある。最終的な処理の例としては、製本や綴じなどがある。また、ある種の画像処理、例えば光学式文字認識（OCR）を満足できる正確さで行うために、入力データの向きを具体的に指定する必要がある場合がある。さらに、名目上のページの向きに対する電子原稿のページの向きがわからないと、コンピュータのモニタや手持ちディスプレイなどのような表示装置に表示する際、適切な向きで表示することができない場合がある。

【0051】

本発明のいくつかの実施形態は、電子原稿におけるテキストの最も有力な向きの自動的な検知に関するものである。テキストの向きは上記名目上のページの向きと関連している。

10

【0052】

スキャンやコピーやその他のデジタル画像処理を行う際、原稿を画像装置の原稿載置台や原稿フィーダーに普通とは違う置き方をし、その結果、電子的な再生（スキャンともいう）が通常の読む向きから反転したり回転したりする場合がある。

【0053】

図1Aないし1Dは、読みやすい向き（「上」向きともいう）に対する原稿のいくつかの向きを描くものである。図1Aは、読みやすい向き（0度の向きともいう）にあるテキスト原稿2の一例を示す。図1Bは、270度回転した向き（時計回りに90度回転した向き、反時計回りに270度回転した向き、右に90度回転した向き、左に270度回転した向きともいう）にあるテキスト原稿4の一例を示す。図1Cは、180度回転した向き（反転した向きともいう）にあるテキスト原稿6の一例を示す。図1Dは、90度回転した向き（時計回りに270度回転した向き、反時計回りに90度回転した向き、右に270度回転した向き、左に90度回転した向きともいう）にあるテキスト原稿8の一例を示す。

20

【0054】

電子ページが回転していると、コピーの場合には元原稿を物理的に回転させることによって、スキャンの場合にはドキュメントビューアーや他の画像/原稿処理アプリケーションにおけるページをデジタル的に反転または回転させることによって、ユーザーは手動で向きを修正しなければならない。

30

【0055】

本発明のいくつかの実施形態では、電子原稿が読みやすい向きにあるか（以下、第1の向きという）、90度回転しているか（以下、第2の向きという）、180度回転しているか（以下、第3の向きという）、それとも270度回転しているか（以下、第4の向きという）を判定する。これらの4つの向きは、4つの基本方向（または基本の向き）（cardinal direction）ともいう。

【0056】

本発明のいくつかの実施形態では、原稿画像の向きは、連結成分オブジェクトと向きを特定した一組の数字テンプレートとを比較することにより判定される。ここで、連結成分オブジェクトとは、2値化画像において、非背景画素（例えば黒画素：画素値が1の画素）が連結したひとまとまりの領域の一つをいう。本発明のいくつかの実施形態では、距離を測定することによりかたちが似ているかどうかを判定する。テンプレートと一致した結果は、原稿画像全体にわたって累積してもよいし、原稿画像の一部において累積してもよい。一致した回数が一番多かった向きが、推定向きとして選ばれる。本発明のいくつかの実施形態では、推定向きの確実性を反映する信頼度（確実性度ともいう）が判定される。これらの実施形態のいくつかでは、信頼度があらかじめ定められた基準を満たしているときは、推定向きは拒絶される。

40

【0057】

（実施形態1）

50

本発明に係る一実施形態について、図2を参照しながら説明する。本実施形態では、まず、原稿要素のサイズが正規化される(10)。ここで、原稿要素とは、2値化された電子原稿画像データにおいて、非背景画素が連結したひとまとまりの連結成分である。なお、ここでは説明を省略するが、数字部分と推定される連結成分が選択され、当該連結成分に対してのみ、符号10以降の処理が実行される。次に、サイズ正規化された原稿要素において、各画素の画素値を文字画素からの距離を示す値に変換する距離算出処理が実行される(12)。その後、距離算出処理がされた原稿要素は、参照テンプレートに対して照合される(14)。そして、照合結果が累積され(16)、累積された照合結果に基づいて、推定向きが求められる(18)。ここで、推定向きとは、照合結果から推定される電子原稿画像データの向きのことである。本実施形態において、参照テンプレートは向きを特定した数字テンプレートを含んでいる。この数字テンプレートの詳細については後述する。

10

【0058】

(実施形態2)

また、本発明に係る別の実施形態について、図3を参照しながら説明する。本実施形態でも、原稿要素は、サイズが正規化され(20)、距離算出処理が実行され(22)、参照テンプレートに対して照合される(24)。なお、本実施形態でも、原稿要素とは、2値化された電子原稿画像データにおいて、黒画素が連結したひとまとまりの連結成分である。また、ここでは説明を省略するが、数字部分と推定される連結成分が選択され、当該連結成分に対してのみ、符号10以降の処理が実行される。そして、照合結果が累積され(26)、累積された照合結果に基づいて、推定向きが求められる(28)。また、この実施形態では、推定向きの信頼性を示す信頼度が求められる(30)。本実施形態でも、参照テンプレートとしては、向きを特定した数字テンプレートがある。

20

【0059】

(実施形態1・2における数字テンプレートの作成方法)

上記の実施形態1・2では、連結成分オブジェクトの候補を原稿要素として4つの基本方向の数字テンプレートと照合し、原稿画像の向きを推測する。数字テンプレートは、トレーニングデータ(training data)を用いて作り出してもよい。また、数字テンプレートは、共通の、サイズ及びフォント不変の数字の代表例(representation)であってもよい。

30

【0060】

以下、数字テンプレートの作成方法の具体例について、図4を参照しながら説明する。まず、トレーニングデータを作成する。このトレーニングデータの作成方法は、以下のとおりである。すなわち、様々な種類のフォント、サイズのサンプル原稿を準備し、当該サンプル原稿をスキャンする。そして、当該スキャンにより得られたスキャンデータを2値化する。そして、2値化されたスキャンデータの中から、数字を含み、当該数字の外接矩形の部分画像を切り取ることでトレーニングデータを作成することができる。ここで、外接矩形は、原稿画像の縦方向および横方向に平行な辺を有し、連結成分に外接する矩形である。なお、このスキャンデータでは、読むのに適した向きになるように数字が配置されているようにしている。このように、様々な種類のフォント、サイズのサンプル原稿から得られた複数のトレーニングデータをまとめたものをトレーニングセットという。

40

【0061】

次に、各数字に対応する複数のトレーニングデータの各々のサイズを正規化し(32)、お互いのあいだで平均化(34)する。これにより、当該数字の平均化されたテンプレートを作成することができる。

【0062】

例えば、各数字について、トレーニングデータである2値化された数字要素サンプルの各々は、サイズ正規化処理(32)として、32画素×32画素の格子状にする。そして、サイズ正規化された複数のトレーニングデータについて、画素ごとに平均値を求める。当該平均値を画素値とする数字テンプレートを作成する。なお、スキャンデータでは、読

50

むのに適した向きになるように数字が配置されているため、このようにして作成された数字テンプレートは、第1の向きに向いたものとなる。すなわち、上記の処理フローでは、第1の向きに対応した数字テンプレートが作成される。

【0063】

そして、回転されたテンプレート、つまり、第2～第4の向きに向いた数字テンプレート（以下、第2～第4の向きに対応する数字テンプレートともいう）は、上記第1の向きに対応した数字テンプレートを回転させることが容易に作成することができる。

【0064】

なお、第1の向きに向いた数字を含む原稿、第2の向きに向いた数字を含む原稿、第3の向きに向いた数字を含む原稿、第4の向きに向いた数字を含む原稿の各々をスキャンして、第1～第4の向きに対応するスキャンデータを作成してもよい。なお、各原稿は、様々な種類のフォント、サイズの数字を含むものとする。この場合、各数字について、第1～第4の向きの各々に対応する複数のトレーニングデータを得ることができる。そして、第1～第4の向きの各々について、トレーニングデータから上記のように直接数字テンプレートを作成してもよい。

10

【0065】

また、上記の説明では、トレーニングデータを、2値化されたスキャンデータから切り出すことで作成されるものとした。しかしながら、テンプレートを作成するためのトレーニングデータは、さまざまな方法でコンパイルしてよい。例えば、ワードプロセッシングプログラムを用いて、数字符号だけを含んでいる原稿データを作り出し、当該原稿データからトレーニングデータを作成してもよい。また、一つの数字に対して、異なるフォントとサイズの当該数字を含む割付原稿を作成してもよい。各原稿のページは画像フォーマットに変換してもよい。画像は2値化され、連結成分ラベリングアルゴリズムによって分析したりしてよい。連結成分の一組（2値マップともいう）が、数字に対応するトレーニングデータを形成してもよい。

20

【0066】

以下、数字テンプレートの具体的な算出方法の一例を述べる。まず、一つの数字 n に対応し、かつ、向き k に対応するトレーニングセット $[n, k]$ に含まれる複数のトレーニングデータの各々は、32画素×32画素のブロックサイズに正規化してフォントタイプやサイズの違いの影響を除外する。ここで、トレーニングセット $[n, k]$ に含まれる複数のトレーニングデータは、様々な種類のフォント、サイズの数字 n を向き k で配置した原稿をスキャンすることで得られるスキャンデータを2値化し、数字部分を切り出すことで得られるものである。正規化されたトレーニングデータである2値化マップの各々の画素値は、画素ごとに累積される。そして、向き k における数字 n のための平均化されたテンプレート（ t_n^k で表される）を作り出す。

30

【0067】

【数1】

$$t_n^k(i, j) = \frac{1}{N} \sum_{l=1}^N b_{n_l}(i, j), \quad n = 1, \dots, 10, \quad k \in \{0, 90, 180, 270\}$$

40

【0068】

上記式において、 $b_{n_l}(i, j)$ は、トレーニングセットに含まれる l 番目のトレーニングデータをサイズ正規化し、座標 (i, j) で示される画素の画素値（2値化要素）を表している。また、 N は、トレーニングセットに含まれるトレーニングデータの総数である。この具体例では、トレーニングセットとして、全ての数字0～9について、4つの基本の向きである 0° 、 90° 、 180° 、 270° に対応するセットを準備しておき、各トレーニングセット $[n, k]$ から、 t_n^k が直接算出される。

【0069】

50

ただし、読みやすい向きである 0° の向きに配置された数字をスキャンすることで得られるトレーニングセット $[n,0]$ から、テンプレート t^0_n を算出し、他の向きに対応するテンプレート t^k_n は、当該テンプレート t^0_n を回転させることにより作成してもよい。なお、本具体例では、4つの基本の向き (0° 、 90° 、 180° 、 270°) に対応するテンプレート t^k_n が予め作成され、記憶するものとする。ただし、テンプレートとして、ただ一つの向きに対応するものだけを記憶してもよい。

【0070】

図5に、4つの基本の向き0度(符号50で示される)、90度(符号52で示される)、180度(符号55で示される)、270度(符号56で示される)における「0」から「9」までのアラビア数字に対応するテンプレート例40~49を示す。

10

【0071】

(実施形態1の具体的処理例1-1)

上述した実施形態1の具体的な処理例の一つを、図6を参照しつつ説明する。この具体例においては、原稿画像60は、向き推定の処理を行う前に、前処理がされている。この前処理として、まず、原稿画像60の2値化処理を行う(62)。次に、2値化された原稿画像に連結成分ラベリング処理を行う(64)。

【0072】

本具体例においては、2値化処理(62)では閾値処理を行う。閾値処理においては、原稿画像60の各画素の画素値を予め定められた閾値と比較し、当該閾値より小さい画素値を2値「0」と「1」のうちどちらか一方に割り振り、閾値より大きいか閾値と等しい原稿画素値を当該2値のうち他方に割り振る。例えば、白い背景と黒い文字の原稿画像に閾値処理を施すと、すべての文字画素が「1」で表され、すべての非文字画素が「0」で表される2値マスク画像が作り出される。

20

【0073】

また、連結成分ラベリング処理は、公知の技術であり、例えば、2値画像データにおいて、同じ画素値(例えば1)を有する画素が連結してひとまとまりになっている領域を連結成分として抽出し、各連結成分に対してラベリング処理(番号付け)を行う処理である。

【0074】

すなわち、隣接する共通の画素を介して連結された画素一つ一つは、連結成分ラベリング処理(64)で一つにまとめられる。連結成分は原稿画像における非背景内容に相当する。非背景内容としては、文字や、写真領域や、線画や、局所的な背景領域や、その他の非背景内容が挙げられる。本発明のいくつかの実施形態では、連結領域それぞれは、当該連結領域の非背景画素をすべて含む最小の境界枠によって表される。本発明のいくつかの実施形態においては、要素の性質を計算して非文字の連結成分を除外する。具体的には、(数字文字の検出について)にて後述するように、数字以外の連結成分を除外する。

30

【0075】

次に、各連結成分を調査対象として、数字テンプレートとの比較処理を行う。まずは、調査していない連結成分が存在するか否か判断する(66)。まだ調べられていない連結成分がある場合(67)、次の連結成分を調査することになる。次の未調査の連結成分について、当該連結成分の外接矩形を決定し、当該外接矩形の部分画像を切り出す。ここで、外接矩形は、原稿画像の縦方向および横方向に平行な辺を有し、連結成分に外接する矩形である。そして、切り出した部分画像、つまり、連結成分を含む外接矩形領域の画像に対して、サイズ正規化を行う(68)。そして、この正規化された、2値化の連結成分を含む矩形の部分画像に対して距離算出処理が行われる(70)。

40

【0076】

本具体例では、2値化された連結成分を含む部分画像は、サイズ正規化処理として、32画素×32画素のサイズに変換される。

【0077】

そして、距離算出処理として以下のような処理を行う。すなわち、サイズ正規化処理が

50

された部分画像（連結成分を含む矩形の部分画像）の各画素について、非背景画素（例えば、白背景に黒文字が形成されている画像である場合、画素値「1」を有する画素）からの距離を示す値を画素値として割り振る。具体的には、非背景画素（ここでは、白背景に黒文字が形成されている画像である場合、画素値「1」を有する文字画素）に対して画素値「0」を割り振り、背景画素（ここでは、画素値「0」を有する画素）に対しては、最も近傍の非背景画素との距離を示す値を画素値として割り振る。距離測定の例としては、 L_1 距離、 L_2 距離、市街地距離測定、ユークリッド距離測定、重み付け市街地距離測定、及び当技術分野で周知のその他の距離測定が挙げられる。本具体例では、2値画像を文字画素からの距離に応じて、背景画素を距離（例えば、市街地距離）で置き換える変換処理（grassfire transform）を行う。図22の左側は、距離算出処理後の部分画像の一例を示している。なお、この図では、6画素×7画素にサイズ正規化し、文字「4」の連結成分を含む部分画像が切り出された場合の例である。

10

【0078】

なお、本具体例では、符号66で示す処理の後に、サイズ正規化処理（68）および距離算出処理（70）を行うものとしたが、先に、全ての連結成分に対して、サイズ正規化処理（68）および距離算出処理（70）を行ってから、一つ一つの連結成分の候補が後述する数字テンプレートとの比較調査が行われたかを判定する一連の処理（72, 76, 80, 82, 86）を実行してもよい。

【0079】

距離算出処理が実行された部分画像は、4つの基本の向きのそれぞれにおける数字テンプレートと比較される。まず、すべての向きについて比較処理が実行されたかどうかを最初に判定される（72）。もしすべての向きについて比較処理が実行された場合（73）、未調査の連結成分が存在するかが判定される（66）。もしまだ比較処理が実行されていない向きが残っていたら（74）、まだ調べられていない数字テンプレートがあるかどうか判定される（76）。ある向きにおいて調べられていない数字テンプレートがなかったら（77）、次の向きが調べられる（72）。まだ調べられていない数字テンプレートがあったら（78）、距離算出処理が実行された部分画像と、現在調べられている向きの数字テンプレートとのあいだで類似度測定が行われる（80）。

20

【0080】

本具体例では、調査対象である連結成分を含む矩形の部分画像 c_m とテンプレート t_n^k （下付き文字 n は数字を表し、上付き文字 k は向きを表す）とのあいだの類似度 $s_k(m, n)$ は、下記式に従って計算される（80）。

30

【0081】

【数2】

$$s_k(m, n) = \sum_{i=1}^I \sum_{j=1}^J c_m(i, j) t_n^k(i, j),$$

【0082】

上記式において、部分画像とテンプレートの正規化されたサイズは I 画素 × J 画素である。本具体例では、 $I = 32$ 、 $J = 32$ である。すなわち、同じサイズである部分画像および数字テンプレートに対して、同じ位置の画素について画素値同士を積算し、全ての画素の積算値を足し合わせた値を類似度とする。図22は、類似度を求める方法を示す図である。

40

【0083】

本具体例において、距離算定処理では、非背景画素（ここでは、白背景に黒文字が形成されている画像である場合、画素値「1」を有する文字画素）に対して画素値「0」を割り振り、背景画素（ここでは、画素値「0」を有する画素）に対しては、最も近傍の非背景画素との距離を示す値を画素値として割り振る。そのため、 $s_k(m, n)$ の値が低い

50

ことは、部分画像 c_m とテンプレート t_n^k がよく一致していることを示している。

【0084】

次に、測定した類似度を所定の基準と比較することで、部分画像 c_m とテンプレート t_n^k とが一致しているか否かの判定が実行される(82)。一致している場合(84)、現在調べられている向きに対応する向き一致カウンタのカウント数を増やす(インクリメントする)(86)。そして、次のまだ調べられていない向きが検討される(72)。もし部分画像 c_m とテンプレート t_n^k とが一致していないと見なされると(83)、次のまだ調べられていない数字が検討される(76)。

【0085】

本具体例では、 $s_k(m, n)$ の値が小さいことは、部分画像 c_m とテンプレート t_n^k とがよく一致していることを示している。そこで、 $s_k(m, n) \geq T_{match}$ である場合に、部分画像 c_m とテンプレート t_n^k とが一致していると見なされる。ここで、 T_{match} は予め定められた閾値である。なお、 $s_k(m, n) < T_{match}$ である場合に、部分画像 c_m とテンプレート t_n^k とが一致していると見なされてもよい。

【0086】

また、類似度 $s_k(m, n)$ として、値が大きくなるにつれ、部分画像 c_m とテンプレート t_n^k とがよく一致していることを示すものを算出してもよい。例えば、距離算出処理として、文字画素の画素値を「10」に変換し、背景画素の画素値を、文字画素からの距離を示す値を「10」から差し引いた値に変換する処理を行うことが考えられる。そして、上記の数2と同様に、類似度 $s_k(m, n)$ を求める。この場合、 $s_k(m, n)$ が大きいほど、部分画像 c_m とテンプレート t_n^k とがよく一致していることを示すことになる。このような場合では、 $s_k(m, n) \geq T_{match}$ である場合に、部分画像 c_m とテンプレート t_n^k とが一致していると見なされる。ここで、 T_{match} は予め定められた閾値である。もしくは、 $s_k(m, n) > T_{match}$ である場合に、部分画像 c_m とテンプレート t_n^k とが一致していると見なされてもよい。

【0087】

調べていない連結成分が残っていないとき(90)、推定向きが決定される(92)。本具体例では、最大のカウンタを持つ向き一致カウンタに対応する向きが、原稿画像60の推定向きとして決定される。

【0088】

(実施形態2の具体的な処理例2-1)

次に、上述した実施形態2の具体的な処理例の一つを、図7を参照しつつ説明する。図7に示されるように、本具体例では、図6に示した具体例の処理に加えて、推定向きが決定されたあと(92)、当該推定向きに関する信頼度(確実性度)を求める(94)。

【0089】

本具体例では、それぞれの向きに対応する向き一致カウンタを分析して、数字に基づいた推定向きの確実性度が判定される(94)。この判定処理について図8を参照して述べる。まず、向き一致カウンタを分析して、向きの組合せ(90度と270度の組合せ対0度と180度の組合せ)のうち、どちらの組合せが最大の値を持つかを判断する(100)。この分析により、原稿が±90度回転しているのか(102)、あるいはまっすぐな向き/反転した向きにあるか(101)が判定される。次に、選ばれた組合せにおける大きな値と小さな値との比率が計算される(103、104)。

【0090】

もし0度と180度の向きの組合せが、90度と270度の向きの組合せよりも大きな値を持っているならば(101)、上記比率(R_{cnt_r} と表記される)は下記式に従って決定される(103)。

【0091】

10

20

30

40

【数3】

$$R_{cntr} = \frac{\max(cntr0, cntr180)}{\min(cntr0, cntr180)}$$

【0092】

上記式において、 $cntr0$ は0度の向きに対応する一致カウンタの値、 $cntr180$ は180度の向きに対応する一致カウンタの値である。

10

【0093】

もし90度と270度の向きの組合せが、0度と180度の向きの組合せよりも大きな値を持っているならば(102)、上記比率(R_{cntr} と表記される)は下記式に従って決定される(104)。

【0094】

【数4】

$$R_{cntr} = \frac{\max(cntr90, cntr270)}{\min(cntr90, cntr270)}$$

20

【0095】

上記式において、 $cntr90$ は0度の向きに対応する一致カウンタの値、 $cntr270$ は270度の向きに対応する一致カウンタの値である。

【0096】

比率 R_{cntr} は、原稿画像における数値候補の数とテキスト要素の候補の総数との比率(R_{comp} と表される)と乗算される(105)。得られる尺度($M_{reliability}$ と表される)は、入力原稿の推定向きの信頼度を表す。そして、本具体例においては、信頼度は閾値(T_{conf} と表される)と比較される。信頼度 $M_{reliability}$ が閾値 T_{conf} より大きいときは(107)、推定向きは信頼できると見なされる(108)。信頼度数 $M_{reliability}$ が閾値 T_{conf} 以下のときは(109)、推定向きは信頼できないと見なされ拒否される(110)。本具体例においては、 T_{conf} の値は0.6である。

30

【0097】

なお、推定向きは信頼できると見なされると(108)、当該推定向きを出力する。そして、外部の画像処理部は、当該推定向きに従った処理を実行する。例えば、画像処理部は、表示部に原稿画像を表示する場合、当該推定向きに応じた角度だけ原稿画像を回転処理し、回転後の原稿画像を表示する。なお、推定向きが0度である場合は、回転処理を行わず、推定向きが90度(反時計周りに90度)である場合は、時計周りに90度だけ回転させ、推定向きが180度である場合は、180度だけ回転させ、推定向きが270度(反時計周りに270度)である場合は、時計周りに270度だけ回転させる。この回転角度は、推定向きに対応付けて記憶部に格納されており、画像処理部は、当該記憶部から推定向きに対応する角度を読み出せばよい。

40

【0098】

一方、推定向きは信頼できないと見なされ拒否される(110)、当該推定向きを出力しない。この場合、ユーザに対して向きの入力を促す画面を表示し、ユーザ入力に応じて向きを決定してもよい。もしくは、向きを判定できないことを示す情報を出力してもよい。

【0099】

50

なお、信頼度の求め方は、上記に限定されない。例えば、推定向きの妥当性は各一致カウンタの最大値に基づいている。そのため、一致カウンタの最大値を信頼度として求め、当該最大値があらかじめ定められた閾値 (T_{MaxVal} と表される) を越える場合にのみ、推定向きが信頼できるものとみなされ、出力されてもよい。もしこの条件が満たされないならば、数字に基づいた推定向きは信頼できないと見なされ、用いられない。

【0100】

その他、 T_{Conf} と T_{MaxVal} それぞれに関する条件双方が満たされて初めて、推定向きが信頼できると見なされてもよい。そして、これら信頼性の条件の一方が満たされないならば、推定向きは信頼できないと見なされてもよい。もしくは、これらの条件の一方だけが調べられてもよい。

【0101】

なお、 T_{MaxVal} の値は例えば 25 に設定される。

【0102】

(実施形態 1 の具体的処理例 1 - 2)

上記の (実施形態 1 の具体的処理例 1 - 1) では、図 6 を参照して述べるように、ある向きに対応する一致カウンタは、当該向きに関連付けられた少なくとも一つのテンプレートが、連結成分を含む部分画像と一致しているに、カウンタ数が増やされる。しかしながら、本発明はこれに限定されず、部分画像ごとに、全ての数字および向きに対応する数字テンプレートとの一致度を求め、最も一致度が高かった向きに対応する向き一致カウンタのカウンタ数を増やすようにしてもよい。本具体例は、このような方式の例である。本具体例について、図 9 を参照しつつ述べる。

【0103】

この具体例においては、原稿画像 118 は、向き推定の処理を行う前に、前処理がされている。この前処理として、まず、原稿画像 118 の 2 値化処理を行う (120)。次に、2 値化された原稿画像に連結成分ラベリング処理を行う (122)。

【0104】

本具体例においては、2 値化処理 120 では閾値処理を行う。閾値処理においては、原稿画像 118 の各画素の画素値を予め定められた閾値と比較し、当該閾値より小さい画素値を 2 値「0」と「1」のうちどちらか一方に割り振り、閾値より大きいか閾値と等しい原稿画素値を当該 2 値のうち他方に割り振る。例えば、白い背景と黒い文字の原稿画像に閾値処理を施すと、すべての文字画素が「1」で表され、すべての非文字画素が「0」で表される 2 値マスク画像が作り出される。

【0105】

また、連結成分ラベリング処理は、公知の技術であり、例えば、2 値画像データにおいて、同じ画素値 (例えば 1) を有する画素が連結してひとまとまりになっている領域を連結成分として抽出し、各連結成分に対してラベリング処理 (番号付け) を行う処理である。

【0106】

すなわち、隣接する共通の画素を介して連結された画素一つ一つは、連結成分ラベリング処理 (122) で一つにまとめられる。連結成分は原稿画像における非背景内容に相当する。非背景内容としては、文字や、写真領域や、線画や、局所的な背景領域や、その他の非背景内容が挙げられる。本発明のいくつかの実施形態では、連結領域それぞれは、当該連結領域の非背景画素をすべて含む最小の境界枠によって表される。本発明のいくつかの実施形態においては、要素の性質を計算して非文字の連結成分を除外する。具体的には、(数字文字の検出について) にて後述するように、数字以外の連結成分を除外する。

【0107】

次に、各連結成分を調査対象として、数字テンプレートとの比較処理を行う。まずは、調査していない連結成分が存在するか否か判断する (124)。まだ調べられていない連結成分がある場合 (123)、次の連結成分を調査することになる。次の未調査の連結成分について、当該連結成分の外接矩形を決定し、当該外接矩形の部分画像を切り出す。こ

10

20

30

40

50

ここで、外接矩形は、原稿画像の縦方向および横方向に平行な辺を有し、連結成分に外接する矩形である。そして、切り出した部分画像、つまり、連結成分を含む外接矩形領域の画像に対して、サイズ正規化を行う(125)。そして、この正規化された、2値化の連結成分を含む矩形の部分画像に対して距離算出処理が行われる(126)。

【0108】

本具体例では、2値化された連結成分を含む部分画像は、サイズ正規化処理として、32画素×32画素のサイズに変換される。

【0109】

そして、距離算出処理として以下のような処理を行う。すなわち、サイズ正規化処理がされた部分画像(連結成分を含む矩形の部分画像)の各画素について、非背景画素(例えば、白背景に黒文字が形成されている画像である場合、画素値「1」を有する画素)からの距離を示す値を画素値として割り振る。具体的には、非背景画素(ここでは、白背景に黒文字が形成されている画像である場合、画素値「1」を有する文字画素)に対して画素値「0」を割り振り、背景画素(ここでは、画素値「0」を有する画素)に対しては、最も近傍の非背景画素との距離を示す値を画素値として割り振る。距離測定の例としては、 L_1 距離、 L_2 距離、市街地距離測定、ユークリッド距離測定、重み付け市街地距離測定、及び当技術分野で周知のその他の距離測定が挙げられる。本具体例では、2値画像を文字画素からの距離に応じて、背景画素を距離(例えば、市街地距離)で置き換える変換処理(grassfire transform)を行う。

10

【0110】

なお、本具体例では、符号124で示す処理の後に、サイズ正規化処理(125)および距離算出処理(126)を行うものとしたが、先に、全ての連結成分に対して、サイズ正規化処理(125)および距離算出処理(126)を行ってから、一つ一つの連結成分の候補が後述する数字テンプレートとの比較調査が行われたかを判定する一連の処理(128, 132, 136, 138, 142)を実行してもよい。

20

【0111】

距離算出処理が実行された部分画像は、4つの基本の向きのそれぞれにおける数字テンプレートと比較される。まず、すべての向きについて比較処理が実行されたかどうかを最初に判定される(128)。もしすべての向きについて比較処理が実行された場合(129)、部分画像と数字テンプレートとのあいだで求めた類似度に基づいて、最もよく類似している数字テンプレートの向きに対応する向き一致カウンタのカウンタ値を増やす(144)。例えば、類似しているほど類似度が小さい値をとる場合、類似度が最小の数字テンプレートの向きに対応するカウンタ値を増やせばよい。続いて、次の連結成分が調べられる(124)。

30

【0112】

もしまだ比較処理が実行されていない向きが残っていたら(130)、まだ調べられていない数字テンプレートがあるかどうかを判定される(132)。ある向きにおいて調べられていない数字テンプレートがなかったら(133)、次の向きが調べられる(128)。まだ調べられていない数字テンプレートがあったら(134)、距離算出処理が実行された部分画像と、現在調べられている向きの数字テンプレートとのあいだで類似度測定が行われる(136)。

40

【0113】

本具体例では、調査対象である連結成分を含む矩形の部分画像 c_m とテンプレート t_n^k (下付き文字 n は数字を表し、上付き文字 k は向きを表す) とのあいだの類似度 $s_k(m, n)$ は、下記式に従って計算される(136)。

【0114】

【数5】

$$s_k(m, n) = \sum_{i=1}^I \sum_{j=1}^J c_m(i, j) t_n^k(i, j)$$

【0115】

上記式において、部分画像とテンプレートの正規化されたサイズはI画素×J画素である。本具体例では、I = 32、J = 32である。すなわち、同じサイズである部分画像および数字テンプレートに対して、同じ位置の画素について画素値同士を積算し、全ての画素の積算値を足し合わせた値を類似度とする。

10

【0116】

本具体例において、距離算定処理では、非背景画素（ここでは、白背景に黒文字が形成されている画像である場合、画素値「1」を有する文字画素）に対して画素値「0」を割り振り、背景画素（ここでは、画素値「0」を有する画素）に対しては、最も近傍の非背景画素との距離を示す値を画素値として割り振る。そのため、 $s_k(m, n)$ の値が低いことは、部分画像 c_m とテンプレート t_n^k がよく一致していることを示している。

【0117】

次に、測定した類似度を所定の基準と比較することで、部分画像 c_m とテンプレート t_n^k とが一致しているか否かの判定が実行される（138）。一致している場合（140）、その類似度を一致度とし、当該一致度が前回記録された最もよい一致度（最良一致度）よりもより一致していることを示している場合、当該一致度により、最良一致度インディケータが更新される（142）。すなわち、ここでの一致度は、各部分画像に対して算出された4方向の類似度の中で、最も類似していることを示す類似度を一致度とすることを意味している。そして、次のまだ調べられていない数字が検討される（132）。部分画像 c_m とテンプレート t_n^k とが一致していないと見なされると（139）、次のまだ調べられていない数字が検討される（132）。

20

【0118】

本具体例では、 $s_k(m, n)$ の値が小さいことは、部分画像 c_m とテンプレート t_n^k とがよく一致していることを示している。そこで、 $s_k(m, n) \geq T_{match}$ である場合に、部分画像 c_m とテンプレート t_n^k とが一致していると見なされる。ここで、 T_{match} は予め定められた閾値である。なお、 $s_k(m, n) < T_{match}$ である場合に、部分画像 c_m とテンプレート t_n^k とが一致していると見なされてもよい。

30

【0119】

また、類似度 $s_k(m, n)$ として、値が大きくなるにつれ、部分画像 c_m とテンプレート t_n^k とがよく一致していることを示すものを算出してもよい。例えば、距離算出処理として、文字画素の画素値を「10」に変換し、背景画素の画素値を、文字画素からの距離を示す値を「10」から差し引いた値に変換する処理を行うことが考えられる。そして、上記の数2と同様に、類似度 $s_k(m, n)$ を求める。この場合、 $s_k(m, n)$ が大きいほど、部分画像 c_m とテンプレート t_n^k とがよく一致していることを示すことになる。このような場合では、 $s_k(m, n) \geq T_{match}$ である場合に、部分画像 c_m とテンプレート t_n^k とが一致していると見なされる。ここで、 T_{match} は予め定められた閾値である。もしくは、 $s_k(m, n) > T_{match}$ である場合に、部分画像 c_m とテンプレート t_n^k とが一致していると見なされてもよい。

40

【0120】

調べていない要素が残っていないときは（146）、推定向きが判定される（148）。本具体例では、最大のカウンタ値を持つ一致カウンタに対応する向きが、原稿画像118の推定向きとして決定される。

【0121】

50

なお、上記の説明では、各連結成分について、全ての向きおよび数字に対応する数字テンプレートについて順番に当該連結成分を含む部分画像との類似度を求めていく。そして、符号142で示されるように、これまで求めた類似度よりも高い類似度が算出されるたびに、最良類似度を更新するようにした。しかしながら、これに限らず、各連結成分について、全ての向きおよび数字に対応する数字テンプレートの各々と当該連結成分を含む部分画像との類似度を求め、その中から最もよく類似していることを示す類似度のものを選択して、当該最良類似度を記憶部に記録してもよい。すなわち、各連結成分ごとに類似度を求めた後、カウントアップ処理および方向を算出してもよい。

【0122】

(実施形態2の具体的処理例2-2)

次に、上述した実施形態2の具体的な別の処理例の一つを、図10を参照しつつ説明する。図10に示されるように、本具体例では、図9に示した具体例の処理に加えて、推定向きが決定されたあと(148)、当該推定向きに関する信頼度(確実性度)を求める(150)。

【0123】

本具体例において、信頼度は、図8を参照して上述した説明に従い、求められる(150)。

【0124】

(別の実施形態における数字テンプレートの作成方法)

以下、別の実施形態における数字テンプレートの作成方法について説明する。本実施形態においても、連結成分オブジェクトの候補は、4つの基本の向きの数字テンプレートと照合され、原稿画像の向きを推測する。数字テンプレートは、トレーニングデータ(training data)を用いて作り出してもよい。また、数字テンプレートは、共通の、サイズ及びフォント不変の数字の代表例(representation)であってもよい。

【0125】

以下、本実施形態における数字テンプレートの作成方法の具体例について、図11を参照しながら説明する。

まず、トレーニングデータを作成する。このトレーニングデータの作成方法は、以下のとおりである。すなわち、様々な種類のフォント、サイズのサンプル原稿を準備し、当該サンプル原稿をスキャンする。そして、当該スキャンにより得られたスキャンデータを2値化する。そして、2値化されたスキャンデータの中から、数字を含み、当該数字の外接矩形の部分画像を切り取ることでトレーニングデータを作成することができる。ここで、外接矩形は、原稿画像の縦方向および横方向に平行な辺を有し、連結成分に外接する矩形である。なお、このスキャンデータでは、読むのに適した向きになるように数字が配置されているようにしている。このように、様々な種類のフォント、サイズのサンプル原稿から得られた複数のトレーニングデータをまとめたものをトレーニングセットという。

【0126】

次に、各数字に対応する複数のトレーニングデータの各々のサイズを正規化し(160)、正規化されたトレーニングデータの各々に対して、距離算出処理を実行する(162)。そして、距離算出処理が実行された複数のトレーニングデータのあいだで平均化(164)する。これにより、当該数字の平均化されたテンプレートを作成することができる。なお、スキャンデータでは、読むのに適した向きになるように数字が配置されているため、このようにして作成された数字テンプレートは、第1の向きに向いたものとなる。すなわち、上記の処理フローでは、第1の向きに対応した数字テンプレートが作成される。

【0127】

そして、回転されたテンプレート、つまり、第2~第4の向きに向いた数字テンプレート(以下、第2~第4の向きに対応する数字テンプレートともいう)は、上記第1の向きに対応した数字テンプレートを回転させることが容易に作成することができる(166)。

【0128】

50

なお、第1の向きに向いた数字を含む原稿、第2の向きに向いた数字を含む原稿、第3の向きに向いた数字を含む原稿、第4の向きに向いた数字を含む原稿の各々をスキャンして、第1～第4の向きに対応するスキャンデータを作成してもよい。なお、各原稿は、様々な種類のフォント、サイズの数字を含むものとする。この場合、各数字について、第1～第4の向きの各々に対応する複数のトレーニングデータを得ることができる。そして、第1～第4の向きの各々について、トレーニングデータから上記のように直接数字テンプレートを作成してもよい。

【0129】

本具体例では、一つの数字 n に対応し、かつ、向き 0° に対応するトレーニングセット $[n, 0]$ に含まれる複数のトレーニングデータの各々は、 32 画素 \times 32 画素のブロックサイズに正規化してフォントタイプやサイズの違いの影響を除外する。

10

【0130】

そして、距離算出処理として以下のような処理を行う。すなわち、サイズ正規化処理がされたトレーニングデータ（数字を含む矩形の画像）の各画素について、文字画素（例えば、画素値「1」を有する画素）からの距離を示す値を画素値として割り振る。具体的には、文字画素に対して画素値「0」を割り振り、背景画素（ここでは、画素値「0」を有する画素）に対しては、最も近傍の文字画素との距離を示す値を画素値として割り振る。距離測定の例としては、 L_1 距離、 L_2 距離、市街地距離測定、ユークリッド距離測定、重み付け市街地距離測定、及び当技術分野で周知のその他の距離測定が挙げられる。本具体例では、2値画像を文字画素からの距離に応じて、背景画素を距離（例えば、市街地距離）で置き換える変換処理（grassfire transform）を行う。図23の右側は、距離算出処理後の数字テンプレートの一例を示している。なお、この図では、6画素 \times 7画素にサイズ正規化され、数字「4」に対応するテンプレートの例である。

20

【0131】

（実施形態3）

上記（別の実施形態における数字テンプレートの作成方法）で記載された数字テンプレートを用いた一実施形態を、図12を参照しつつ説明する。この実施形態においては、原稿画像170、向き推定の処理を行う前に、前処理がされている。この前処理として、まず、原稿画像170の2値化処理を行う（172）。次に、2値化された原稿画像に連結成分ラベリング処理を行う（174）。

30

【0132】

本実施形態においては、2値化処理（172）では閾値処理を行う。閾値処理においては、原稿画像170の各画素の画素値を予め定められた閾値と比較し、当該閾値より小さい画素値を2値「0」と「1」のうちどちらか一方に割り振り、閾値より大きいか閾値と等しい原稿画素値を当該2値のうち他方に割り振る。例えば、白い背景と黒い文字の原稿画像に閾値処理を施すと、すべての文字画素が「1」で表され、すべての非文字画素が「0」で表される2値マスク画像が作り出される。

【0133】

また、連結成分ラベリング処理は、公知の技術であり、例えば、2値画像データにおいて、同じ画素値（例えば1）を有する画素が連結してひとまとまりになっている領域を連結成分として抽出し、各連結成分に対してラベリング処理（番号付け）を行う処理である。

40

【0134】

すなわち、隣接する共通の画素を介して連結された画素一つ一つは、連結成分ラベリング処理（174）で一つにまとめられる。連結成分は原稿画像における非背景内容に相当する。非背景内容としては、文字や、写真領域や、線画や、局所的な背景領域や、その他の非背景内容が挙げられる。本発明のいくつかの実施形態では、連結領域それぞれは、当該連結領域の非背景画素をすべて含む最小の境界枠によって表される。本発明のいくつかの実施形態においては、要素の性質を計算して非文字の連結成分を除外する。なお、ここでは説明を省略するが、数字部分と推定される連結成分が選択され、当該連結成分に対し

50

でのみ、符号 10 以降の処理が実行される。

【0135】

次に、各連結成分を調査対象として、数字テンプレートとの比較処理を行う。まずは、調査していない連結成分が存在するか否か判断する(176)。まだ調べられていない連結成分があったら(177)、次の連結成分を調査することになる。次の未調査の連結成分について、当該連結成分の外接矩形を決定し、当該外接矩形の部分画像を切り出す。ここで、外接矩形は、原稿画像の縦方向および横方向に平行な辺を有し、連結成分に外接する矩形である。そして、切り出した部分画像、つまり、連結成分を含む外接矩形領域の画像に対して、サイズ正規化を行う(178)。2値化された連結成分を含む部分画像は、サイズ正規化処理として、32画素×32画素のサイズに変換される。

10

【0136】

なお、本実施形態では、符号176で示す処理の後に、サイズ正規化処理(178)を行うものとしたが、先に、全ての連結成分に対して、サイズ正規化処理(178)を行ってから、一つ一つの連結成分の候補が後述する数字テンプレートとの比較調査が行われたかを判定する一連の処理(180, 184, 188, 190, 194)を実行してもよい。

【0137】

2値化され正規化された、連結成分を含む部分画像は、4つの基本の向きのそれぞれにおける数字テンプレートと比較される。まず、すべての向きについて比較処理が実行されたかどうか最初に判定される(180)。もしすべての向きについて比較処理が実行された場合(181)、未調査の連結成分が存在するかが判定される(176)。もしまだ比較処理が実行されていない向きが残っていたら(182)、まだ調べられていない数字テンプレートがあるかどうか判定される(184)。もしある向きにおいて調べられていない数字テンプレートが残っていなかったら(185)、次の向きが調べられる(180)。もしまだ調べられていない数字テンプレートが残っていたら(186)、正規化された部分画像と現在調べられている向きにおける数字テンプレートとのあいだで類似度が算出される(188)。

20

【0138】

本実施形態では、調査対象である連結成分を含む矩形の部分画像 c_m とテンプレート g_n^k (下付き文字 n は数字を表し、上付き文字 k は向きを表す) とのあいだの類似度 $s_k(m, n)$ は、下記式に従って計算される(188)。

30

【0139】

【数6】

$$s_k(m, n) = \sum_{i=1}^I \sum_{j=1}^J c_m(i, j) g_n^k(i, j)$$

【0140】

上記式において、部分画像とテンプレートの正規化されたサイズは I 画素 × J 画素である。本具体例では、 $I = 32$ 、 $J = 32$ である。すなわち、同じサイズである部分画像および数字テンプレートに対して、同じ位置の画素について画素値同士を積算し、全ての画素の積算値を足し合わせた値を類似度とする。図23は、類似度を求める方法を示す図である。

40

【0141】

本実施形態において、テンプレートは、距離算出処理が実行されたものであり、部分画像は正規化された2値の画素値を含んでおり、類似度は下記式に従って算出される。

【0142】

【数 7】

$$s_k(m,n) = \sum_{i=1}^I \sum_{j=1}^J \left\{ \begin{array}{ll} g_n^k(i,j), & \text{if } c_m(i,j)=1 \\ 0, & \text{otherwise} \end{array} \right\}$$

【0143】

本実施形態において、テンプレートに適用される距離算出処理では、文字画素に対して画素値「0」を割り振り、背景画素（ここでは、画素値「0」を有する画素）に対しては、最も近傍の文字画素との距離を示す値を画素値として割り振る。そのため、 $s_k(m, n)$ の値が低いことは、部分画像 c_m とテンプレート g_n^k がよく一致していることを示している。

10

【0144】

次に、測定した類似度を所定の基準と比較することで、部分画像 c_m とテンプレート g_n^k とが一致しているか否かの判定が実行される（190）。一致している場合（192）、現在調べられている向きに対応する向き一致カウンタのカウント数を増やす（インクリメントする）（194）。そして、次のまだ調べられていない向きが検討される（180）。もし部分画像 c_m とテンプレート g_n^k とが一致していないと見なされると（191）、次のまだ調べられていない数字が検討される（184）。

20

【0145】

本発明のいくつかの実施形態では、 $s_k(m, n)$ の値が小さいことは、部分画像 c_m とテンプレート g_n^k とがよく一致していることを示している。そこで、 $s_k(m, n) < T_{match}$ である場合に、部分画像 c_m とテンプレート g_n^k とが一致していると見なされる。ここで、 T_{match} は予め定められた閾値である。なお、 $s_k(m, n) < T_{match}$ である場合に、部分画像 c_m とテンプレート g_n^k とが一致していると見なされてもよい。

【0146】

また、類似度 $s_k(m, n)$ として、値が大きくなるにつれ、部分画像 c_m とテンプレート g_n^k とがよく一致していることを示すものを算出してもよい。例えば、テンプレートを求める際に実行される距離算出処理として、文字画素の画素値を「10」に変換し、背景画素の画素値を、文字画素からの距離を示す値を「10」から差し引いた値に変換する処理を行うことが考えられる。そして、上記の数6と同様に、類似度 $s_k(m, n)$ を求める。この場合、 $s_k(m, n)$ が大きいほど、部分画像 c_m とテンプレート g_n^k とがよく一致していることを示すことになる。このような場合では、 $s_k(m, n) < T_{match}$ である場合に、部分画像 c_m とテンプレート g_n^k とが一致していると見なされる。ここで、 T_{match} は予め定められた閾値である。もしくは、 $s_k(m, n) > T_{match}$ である場合に、部分画像 c_m とテンプレート g_n^k とが一致していると見なされてもよい。

30

【0147】

調べていない要素が残っていないときは（196）、推定向きが決定される（198）。本具体例では、最大のカウンタを持つ向き一致カウンタに対応する向きが、原稿画像170の推定向きとして決定される。

40

【0148】

(実施形態4)

次に、上記（別の実施形態における数字テンプレートの作成方法）で記載された数字テンプレートを用いた別の実施形態の具体的な処理例の一つを、図13を参照しつつ説明する。図13に示されるように、本実施形態では、図12に示した具体例の処理に加えて、推定向きが決定されたあと（198）、当該推定向きに関する信頼度（確実性度）を求める（200）。

50

【0149】

本実施形態において、信頼度は、図8を参照して上述した説明に従い、求められる(200)。

【0150】

(実施形態5)

上記の実施形態3では、図12を参照して述べるように、ある向きに対応する一致カウンタは、当該向きに関連付けられた少なくとも一つのテンプレートが、連結成分を含む部分画像と一致しているに、カウンタ数が増やされる。しかしながら、本発明はこれに限定されず、部分画像ごとに、全ての数字および向きに対応する数字テンプレートとの一致度を求め、最も一致度が高かった向きに対応する向き一致カウンタのカウンタ数を増やすようにしてもよい。本実施形態は、このような方式の例である。本実施形態について、図14を参照しつつ述べる。

10

【0151】

この実施形態においては、原稿画像210、向き推定の処理を行う前に、前処理がされている。この前処理として、まず、原稿画像210の2値化処理を行う(212)。次に、2値化された原稿画像に連結成分ラベリング処理を行う(214)。

【0152】

本実施形態においては、2値化処理(212)では閾値処理を行う。閾値処理においては、原稿画像210の各画素の画素値を予め定められた閾値と比較し、当該閾値より小さい画素値を2値「0」と「1」のうちどちらか一方に割り振り、閾値より大きいか閾値と等しい原稿画素値を当該2値のうち他方に割り振る。例えば、白い背景と黒い文字の原稿画像に閾値処理を施すと、すべての文字画素が「1」で表され、すべての非文字画素が「0」で表される2値マスク画像が作り出される。

20

【0153】

また、連結成分ラベリング処理は、公知の技術であり、例えば、2値画像データにおいて、同じ画素値(例えば1)を有する画素が連結してひとまとまりになっている領域を連結成分として抽出し、各連結成分に対してラベリング処理(番号付け)を行う処理である。

【0154】

すなわち、隣接する共通の画素を介して連結された画素一つ一つは、連結成分ラベリング処理(214)で一つにまとめられる。連結成分は原稿画像における非背景内容に相当する。非背景内容としては、文字や、写真領域や、線画や、局所的な背景領域や、その他の非背景内容が挙げられる。本発明のいくつかの実施形態では、連結領域それぞれは、当該連結領域の非背景画素をすべて含む最小の境界枠によって表される。本発明のいくつかの実施形態においては、要素の性質を計算して非文字の連結成分を除外する。

30

【0155】

なお、ここでは説明を省略するが、数字部分と推定される連結成分が選択され、当該連結成分に対してのみ、符号10以降の処理が実行される。

【0156】

次に、各連結成分を調査対象として、数字テンプレートとの比較処理を行う。まずは、調査していない連結成分が存在するか否か判断する(216)。まだ調べられていない連結成分があったら(217)、次の連結成分を調査することになる。次の未調査の連結成分について、当該連結成分の外接矩形を決定し、当該外接矩形の部分画像を切り出す。ここで、外接矩形は、原稿画像の縦方向および横方向に平行な辺を有し、連結成分に外接する矩形である。そして、切り出した部分画像、つまり、連結成分を含む外接矩形領域の画像に対して、サイズ正規化を行う(218)。2値化された連結成分を含む部分画像は、サイズ正規化処理として、32画素×32画素のサイズに変換される。

40

【0157】

なお、本実施形態では、符号216で示す処理の後に、サイズ正規化処理(218)を行うものとしたが、先に、全ての連結成分に対してサイズ正規化処理(218)を行って

50

から、その後に数字テンプレートとのマッチング調査を行なってもよい。

【0158】

2値化され正規化された、連結成分を含む部分画像は、4つの基本の向きそれぞれのにおける数字テンプレートと比較される。まず、すべての向きについて比較処理が実行されたかどうか最初に判定される(222)。もしすべての向きについて比較処理が実行された場合(223)、部分画像と数字テンプレートとのあいだで求めた類似度の中で、最もよく類似していることを示す類似度(最良類似度)を特定し、当該最良類似度となる数字テンプレートの向きに対応する向き一致カウンタのカウンタ値を増やす(220)。それから、次の要素が調べられる(216)。

【0159】

もしまだ比較処理が実行されていない向きが残っていたら(224)、まだ調べられていない数字テンプレートがあるかどうか判定される(226)。ある方向において調べられていない数字テンプレートがなかったら(227)、次の向きが調べられる(222)。まだ調べられていない数字テンプレートがあったら(228)、正規化された部分画像と現在調べられている向きにおける数字テンプレートとのあいだで類似度が算出される(230)。

【0160】

本実施形態では、調査対象である連結成分を含む矩形の部分画像 c_m とテンプレート g_n^k (下付き文字 n は数字を表し、上付き文字 k は向きを表す) とのあいだの類似度 s_k (m, n) は、下記式に従って計算される(230)。

【0161】

【数8】

$$s_k(m, n) = \sum_{i=1}^I \sum_{j=1}^J c_m(i, j) g_n^k(i, j)$$

【0162】

上記式において、部分画像とテンプレートの正規化されたサイズは I 画素 \times J 画素である。本具体例では、 $I = 32$ 、 $J = 32$ である。すなわち、同じサイズである部分画像および数字テンプレートに対して、同じ位置の画素について画素値同士を積算し、全ての画素の積算値を足し合わせた値を類似度とする。図23は、類似度を求める方法を示す図である。

【0163】

本実施形態において、テンプレートは、距離算出処理が実行されたものであり、部分画像は正規化された2値の画素値を含んでおり、類似度は下記式に従って算出される。

【0164】

【数9】

$$s_k(m, n) = \sum_{i=1}^I \sum_{j=1}^J \left\{ \begin{array}{ll} g_n^k(i, j), & \text{if } c_m(i, j) = 1 \\ 0, & \text{otherwise} \end{array} \right\}$$

【0165】

本実施形態において、テンプレートに適用される距離算出処理では、文字画素に対して画素値「0」を割り振り、背景画素(ここでは、画素値「0」を有する画素)に対しては、最も近傍の文字画素との距離を示す値を画素値として割り振る。そのため、 $s_k(m, n)$ の値が低いことは、部分画像 c_m とテンプレート g_n^k がよく一致していることを示している。

10

20

30

40

50

【0166】

次に、測定した類似度を所定の基準と比較することで、部分画像 c_m とテンプレート g_n^k とが一致しているか否かの判定が実行される(232)。一致している場合(234)、その類似度を一致度とし、当該一致度が前回記録された最大の一致度より大きいならば、当該一致度により、最上一致インディケータが更新される(236)。そして、次のまだ調べられていない数字が検討される(226)。もし部分画像 c_m とテンプレート g_n^k とが一致していないと見なされると(233)、次のまだ調べられていない数字が検討される(226)。

【0167】

本実施形態では、 $s_k(m, n)$ の値が小さいことは、部分画像 c_m とテンプレート g_n^k とがよく一致していることを示している。そこで、 $s_k(m, n) < T_{match}$ である場合に、部分画像 c_m とテンプレート t_n^k とが一致していると見なされる。ここで、 T_{match} は予め定められた閾値である。なお、 $s_k(m, n) < T_{match}$ である場合に、部分画像 c_m とテンプレート g_n^k とが一致していると見なされてもよい。

10

【0168】

また、類似度 $s_k(m, n)$ として、値が大きくなるにつれ、部分画像 c_m とテンプレート g_n^k とがよく一致していることを示すものを算出してもよい。例えば、テンプレートを求める際に実行される距離算出処理として、文字画素の画素値を「10」に変換し、背景画素の画素値を、文字画素からの距離を示す値を「10」から差し引いた値に変換する処理を行うことが考えられる。そして、上記の数6と同様に、類似度 $s_k(m, n)$ を求める。この場合、 $s_k(m, n)$ が大きいほど、部分画像 c_m とテンプレート g_n^k とがよく一致していることを示すことになる。このような場合では、 $s_k(m, n) < T_{match}$ である場合に、部分画像 c_m とテンプレート g_n^k とが一致していると見なされる。ここで、 T_{match} は予め定められた閾値である。もしくは、 $s_k(m, n) > T_{match}$ である場合に、部分画像 c_m とテンプレート g_n^k とが一致していると見なされてもよい。

20

【0169】

調べていない要素が残っていないときは(238)、推定向きが決定される(240)。本実施形態では、最大のカウント値を持つ一致カウンタに対応する向きが、原稿画像210の推定向きとして決定される。

30

【0170】

なお、上記の説明では、各連結成分について、全ての向きおよび数字に対応する数字テンプレートについて順番に当該連結成分を含む部分画像との類似度を求めていく。そして、符号236で示されるように、これまで求めた類似度よりも高い類似度が算出されるたびに、最良類似度を更新するようにした。しかしながら、これに限らず、各連結成分について、全ての向きおよび数字に対応する数字テンプレートの各々と当該連結成分を含む部分画像との類似度を求め、その中から最もよく類似していることを示す類似度(最良類似度)を選択して、当該最良類似度を記憶部に記録してもよい。すなわち、各連結成分ごとに類似度を求めた後、カウントアップ処理および方向を算出してもよい。

40

【0171】

(実施形態6)

次に、上記(別の実施形態における数字テンプレートの作成方法)で記載された数字テンプレートを用いた別の実施形態の具体的な処理例の一つを、図15を参照しつつ説明する。図15に示されるように、本実施形態では、図14に示した具体例の処理に加えて、推定向きが決定されたあと(240)、当該推定向きに関する信頼度(確実性度)が判定される(242)。

【0172】

本実施形態において、信頼度は、図8を参照して上述した説明に従い、求められる(242)。

【0173】

50

(変形例)

上記の実施形態 1 - 6 は、これに限定されるものではなく、種々の変更が可能である。以下、この変形例について説明する。

【0174】

上記の実施形態 1 - 6 の説明では、十個のアラビア数字に対応する数字テンプレートを用いるものとした。しかしながら、本発明はこれに限定されるものではない。すなわち、十個のアラビア数字の中から選択された 10 個未満のアラビア数字の組合せのみを用いてもよい。この場合、高い向き特性を持つ数字を含むように、より少ない組合せのテンプレートが選ばれる。例えば、十個のアラビア数字のうち、2、3、4、5、7 に対応する数字テンプレートのみを用いても良い。これにより、推定向きをより正確に決定できるとともに、処理時間の短縮化を図ることができる。

10

【0175】

本発明のいくつかの実施形態では、複数のテスト原稿を分析して、部分集合を決定する。いくつかの実施形態では、上記分析は自動的である。他の実施形態では、上記分析は人間によって行われる。本発明の他の実施形態では、上記部分集合は専門家によって発見的 (heuristically) に割り振られる。

【0176】

なお、上記の実施形態では、4つの基本の向きに対応する数字テンプレートを記憶しており、連結成分を含む部分画像について、全ての向きに対応する数字テンプレートとの照合を行うものとした。しかしながら、本発明はこれに限定されず、向き 0 度の数字テンプレートのみを記憶していてもよい。この場合、部分画像を 90 度ずつ回転させながら、向き 0 度の数字テンプレートとの照合を行えばよい。

20

【0177】

また、上記の説明では、4つの基本の向きに対応する数字テンプレートと部分画像とを照合するものとして説明した。しかしながら、数字テンプレートとしては、4つに限られない。例えば、2つであってもよい。

【0178】

具体的には、縦長の原稿に対して、横方向に文字が記載された画像データのみを判定対象とする場合には、当該画像データから得られた部分画像が 90 度または 270 度の向きの数字テンプレートと一致することがない。そのため、このような画像データのみを判定対象とすることが予めわかっている場合には、0 度と 180 度の向きに対応する数字テンプレートのみを記憶しておき、この 2 つの数字テンプレートと照合すればよい。

30

【0179】

また、上記の実施形態では、各向きに対応する一致カウンタは、一ずつカウンタ値を増すものとした。しかしながら、必ずしも「1」ずつ増やさなくてもよい。例えば、数字ごとに予め登録された本発明の他の実施形態では、数字ごとに登録した加算値だけカウンタ値を増やすようにしてもよい。すなわち、数字テンプレートと加算値とを対応付けたテーブルを記憶しておき、部分画像と数字テンプレートとが一致するとみなされたときに、当該数字テンプレートに対応する加算値を上記テーブルから読みだし、当該加算値だけ一致カウンタのカウンタ値を増やす。これにより、数字により異なる加算値となる。すなわち、加算値は、向き特性によって重み付けされる。この変形例では、一致加算器に対する貢献は、一致が向き特有の明確な特性を持つ数字を伴うとき、より大きくなる。つまり、高い向き特性を持つ数字については、加算値を大きな値とし、それ以外の数字については加算値を小さな値とする。これにより、推定向きの決定の精度を高めることができる。なお、加算値の求め方(つまり、特性重み付け)は、一組のテスト原稿を分析して自動的なプロセスによって決められてもよいし、専門家によって開発された発見的方法に従って割り振られてもよい。各数字に対応する加算値は、対称性の小さい数字、つまり、向き特性の高い数字に対応する加算値が相対的に大きくなるように、予め定められている。

40

【0180】

(数字文字の検出について)

50

以下、上記の実施形態 1 - 6 で説明を省略した数字文字の検出について説明する。すなわち、上記の実施形態 1 - 6 では、原稿画像から抽出した全ての連結成分の中から、数字部分と推定される連結成分を選択し、当該連結成分に対してのみ数字テンプレートとの比較処理を行う。

【0181】

この数字文字の検出を用いた実施形態について、図 16 を参照しつつ説明する。本形態では、全ての連結成分をフィルタリングして (250)、調査対象連結成分の候補を決定する。そして、連結成分の候補は、サイズ正規化処理が実行され (252)、距離算出処理が実行され (254)、参照テンプレート (数字テンプレート) に対して照合される (256)。本形態では、全ての連結成分は、2 値の連結成分を含んでおり、フィルタリング (250) によって、数字と推定される連結成分のみが抽出される。その後、一致カウンタ値が累積され (258)、累積された一致カウンタ値に基づいて、推定向きが作られる (260)。

10

【0182】

また、数字文字の検出を用いた実施形態の別の例について、図 17 を参照しつつ説明する。本形態では、全ての連結成分をフィルタリングして (250)、調査対象連結成分の候補を決定する。そして、連結成分の候補は、サイズ正規化処理が実行され (252)、距離算出処理が実行され (254)、参照テンプレート (数字テンプレート) に対して照合される (256)。本形態では、全ての連結成分は、2 値の連結成分を含んでおり、フィルタリング (250) によって、数字と推定される連結成分のみが抽出される。その後、一致カウンタ値が累積され (258)、累積された一致カウンタ値に基づいて、推定向きが作られる (260)。そして、本形態では、この後、推定向きの妥当性を示す信頼度が判定される (262)。本発明のこれらの実施形態では、参照テンプレートとしては向き特有数字テンプレートがある。

20

【0183】

本形態は、原稿画像の中の数字文字の位置を特定する方法とシステムを含んでいる。これらの方法とシステムは、連結成分のフィルタリングを実施するために用いられる。この形態では、電子原稿の入力画像から、2 値テキストマップが作られる。個々のテキスト文字は、上記 2 値テキストマップの連続した画素の組 (連結成分) として表される。

【0184】

本発明のいくつかの実施形態では、デジタル原稿画像の個々のテキスト文字は集められて文字列 (テキスト行という場合もある) になる。ここで、文字列とは、当該文字列に含まれる文字を読むのに適した方向に並べられた列であり、図 1 (A) または図 1 (C) に示される場合は、文字列は、横方向であり、図 1 (B) または図 1 (D) に示される場合は、文字列は、縦方向である。図 18 A に示すように、個々のテキスト文字 270 は、関連境界枠 271 によって表現される。本形態では、テキスト文字境界枠 271 とは、関連テキスト文字 270 が実質的に囲まれる境界枠 (外接矩形枠) のことである。もしくは、テキスト文字境界枠 271 とは、関連テキスト文字 270 が完全に含まれる境界枠 (外接矩形枠) のことである。境界枠 271 は、2 つの向かい合う角の座標、たとえば、境界枠 271 の、 (x_1, y_1) と表記される左上の角 272 と (x_2, y_2) と表記される右下の角 273 によって特徴付けられるか、あるいは、第一の角たとえば (x_1, y_1) と表記される左上の角 272 と第一の角からお互いに直角な 2 つの方向に伸びる境界枠の延長 d_x 、 d_y によって特徴付けられるか、あるいはデジタル原稿画像において境界枠 271 のサイズと位置を特定する他のいかなる方法によってでも特徴付けられる。

30

40

【0185】

一つ以上のテキスト文字を含んでいるテキストオブジェクトは、テキストオブジェクト境界枠によって表される。図 18 B に、テキストオブジェクト 274 とテキストオブジェクト境界枠 275 の一例を示す。本発明のいくつかの実施形態では、テキストオブジェクト境界枠 275 は、関連するテキストオブジェクト 274 が実質的に囲まれる境界枠 (外接矩形枠) である。もしくは、テキストオブジェクト境界枠 275 は、関連するテキスト

50

オブジェクト 274 が完全に含まれる境界枠（外接矩形枠）である。境界枠 275 は、2つの向かい合う角の座標、たとえば、境界枠 275 の、 (x_1, y_1) と表記される左上の角 276 と (x_2, y_2) と表記される右下の角 277 によって特徴付けられるか、あるいは、第一の角たとえば (x_1, y_1) と表記される左上の角 276 と第一の角からお互いに直角な 2つの方向に伸びる境界枠の延長 dx 、 dy によって特徴付けられるか、あるいはデジタル原稿画像において境界枠 275 のサイズと位置を特定する他のいかなる方法によってでも特徴付けられる。

【0186】

図 19 に示すように、文字列 280 は、関連する文字列境界枠 282 によって表される。本発明のいくつかの実施形態では、文字列境界枠 282 は、関連する文字列 280 が実質的に囲まれる境界枠（外接矩形枠）である。もしくは、文字列境界枠 282 は、関連テキストオブジェクト 280 が完全に含まれる境界枠（外接矩形枠）である。文字列境界枠 282 は、左端 284 の x 座標 (x_L) と表記、右端 285 の x 座標 (x_R) と表記、下端 286 の y 座標 (y_B) と表記、上端 287 の y 座標 (y_T) と表記によって表されるか、あるいはデジタル原稿画像における文字列境界枠 282 のサイズと位置を表す他のいかなる方法によっても表される。

10

【0187】

本発明のいくつかの実施形態では、文字列境界枠 282 は、文字列 280 の構成要素であるテキスト文字やテキストオブジェクトの個々の境界枠から、下記式に従って決定される。

20

【0188】

【数 10】

$$y_T = \min\{y_1(i)\}, i = 1, \dots, N,$$

$$y_B = \max\{y_2(i)\}, i = 1, \dots, N,$$

$$x_L = \min\{x_1(i)\}, i = 1, \dots, N,$$

$$x_R = \max\{x_2(i)\}, i = 1, \dots, N,$$

30

【0189】

上記式において、 N は文字列におけるテキスト文字やテキストオブジェクトの数、 $y_1(i)$ と $y_2(i)$ は i 番目のテキスト文字やテキストオブジェクトや境界枠の y_1 と y_2 座標の値であり、 $x_1(i)$ と $x_2(i)$ は i 番目のテキスト文字やテキストオブジェクトや境界枠の x_1 と x_2 座標の値である。

【0190】

このような文字列は、例えば、特許文献 1 に記載のような公知技術を用いて抽出することができる。また、文字列の中の各テキストも、特許文献 1 に記載のような公知技術を用いて抽出することができる。

40

【0191】

図 20 A、20 B、20 C を参照しつつ以下に述べる本発明の実施形態では、印刷関連用語を用いる。この用語は、ラテン文字、ギリシャ文字、キリル文字、デーヴァナーガリー文字その他のアルファベットを用いる言語を含む（ただしこれらに限定されない）文字言語のテキスト文字（文字および符号ともいう）に関するものである。図 20 A に、大文字と小文字の双方を含むラテンアルファベットの文字列を示す。図 20 B に、大文字のみを含むラテンアルファベットの文字列を示す。図 20 C に、アラビア数字のみの文字列を示す。

50

【 0 1 9 2 】

並び線という用語は、テキスト文字がその上に位置する線 3 0 0、3 1 0、3 2 0 を指す。ラテンアルファベットのテキストにおいては、これはすべての大文字と大部分の小文字がその上に位置する線である。ディセンドーは、文字やテキスト文字のうち、並び線 3 0 0、3 1 0、3 2 0 より下に延びる部分を指す。ラテンアルファベットでディセンドーを持つ小文字は、“ g ”、“ j ”、“ p ”、“ q ”、“ y ”である。ディセンドー線とは、その線に向かってテキスト文字のディセンドーが伸びる線 3 0 2、3 1 2、3 2 2 を指す。文字のうち、当該文字の主要部分の上に伸びる部分をアセンドーという。ラテンアルファベットでアセンドーを持つ小文字は、“ b ”、“ d ”、“ f ”、“ h ”、“ k ”、“ l ”、“ t ”である。ラテンアルファベットの大文字はアセンドーだと見なされる。アセンドー線とは、その線に向かってテキスト文字のアセンドーが伸びる線 3 0 3、3 1 3、3 2 3 を指す。ラテンアルファベットでアセンドーもディセンドーも持たない小文字の高さ 3 0 4、3 1 4、3 2 4 は、たとえば“ x ”ならば、x 高さと呼ばれる。アセンドーもディセンドーも持たないこれらの文字の頂点を示す線 3 0 5、3 1 5、3 2 5 は、x 線と呼ばれる。大文字の高さ 3 0 6、3 1 6、3 2 6 は、上限高さと呼ばれる。

10

【 0 1 9 3 】

数字の列の持つ性質を利用して、連結成分の候補の数を減らしたり、数字の列を見つけたりできる。小文字と違って、数字列の連結成分の上端と下端にはほとんど変化がない。数字列およびすべての大文字列においては、列境界枠の上端と下端は並び線 3 0 0、3 1 0、3 2 0 およびアセンドー線 3 0 3、3 1 3、3 2 3 のごく近くに並んでいる。このことは図 2 0 B と 2 0 C から見て取れる。大文字と小文字の双方を含む典型的な文字列では、標準サイズのテキストと大文字の符号にアセンドーやディセンドーが混じるために、文字の上端と下端や文字列の境界は変動する。

20

【 0 1 9 4 】

そこで、本形態では、再構築された符号列の上端と下端間の変動量を数量化して、大文字と小文字の双方の含む文字列を検出する。

【 0 1 9 5 】

デジタル原稿画像の中で水平の向きにある文字列 (t と表す) の場合、上限値 ($\text{ceil}(t)$ と表す) と下限値 ($\text{floor}(t)$ と表す) は下記式に従って算出される。

30

【 0 1 9 6 】

【 数 1 1 】

$$\text{ceil}(t) = \frac{1}{N} \sum_{i=1}^N y_1(i) \quad \text{and} \quad \text{floor}(t) = \frac{1}{N} \sum_{i=1}^N y_2(i)$$

【 0 1 9 7 】

上記式において、N は文字列 t におけるテキスト文字の数、 $y_1(i)$ と $y_2(i)$ は i 番目のテキスト文字境界枠の y_1 座標値と y_2 座標値である。上限値は文字境界枠の y_1 座標値の標本平均だと見なされ、下限値は文字境界枠の y_2 座標値の標本平均だと見なされる。

40

【 0 1 9 8 】

デジタル原稿画像の中で垂直の向きにある文字列 (t と表す) の場合、上限値 ($\text{ceil}(t)$ と表す) と下限値 ($\text{floor}(t)$ と表す) は下記式に従って算出される。

【 0 1 9 9 】

【数 1 2】

$$ceil(t) = \frac{1}{N} \sum_{i=1}^N x_1(i) \quad \text{and} \quad floor(t) = \frac{1}{N} \sum_{i=1}^N x_2(i)$$

【0 2 0 0】

上記式において、 N は文字列 t におけるテキスト文字の数、 $x_1(i)$ と $x_2(i)$ は i 番目のテキスト文字境界枠の x_1 座標値と x_2 座標値である。上限値は文字境界枠の x_1 座標値の標本平均だと見なされ、下限値は文字境界枠の x_2 座標値の標本平均だと見なされる。

10

【0 2 0 1】

標本と、対応する標本平均とのあいだの誤差は、テキストの並び線がどこにあるかを示す指標である。上端及び下端の誤差量を計算し、上端及び下端配列特徴として用いる。

【0 2 0 2】

デジタル原稿画像において水平な向きに位置している文字列 (t と表す) の場合は、誤差量の例として、

下記式に従って計算される平均絶対誤差 (Mean Absolute Error (MAE))

【0 2 0 3】

【数 1 3】

$$e_{MAE}^{top}(t) = \frac{1}{N} \sum_{i=1}^N |y_1(i) - ceil(t)|, \quad e_{MAE}^{bottom}(t) = \frac{1}{N} \sum_{i=1}^N |y_2(i) - floor(t)|$$

20

【0 2 0 4】

下記式に従って計算される2乗平均誤差 (Mean-Square Error (MSE))

【0 2 0 5】

【数 1 4】

$$e_{MAE}^{top}(t) = \frac{1}{N} \sum_{i=1}^N (y_1(i) - ceil(t))^2, \quad e_{MAE}^{bottom}(t) = \frac{1}{N} \sum_{i=1}^N (y_2(i) - floor(t))^2$$

30

【0 2 0 6】

下記式に従って計算される2乗平均平方根誤差 (Root Mean-Square Error (RMSE))

【0 2 0 7】

【数 1 5】

$$e_{RMSE}^{top}(t) = \sqrt{e_{MSE}^{top}(t)}, \quad e_{RMSE}^{bottom}(t) = \sqrt{e_{MSE}^{bottom}(t)}$$

40

【0 2 0 8】

およびその他の誤差量が挙げられる。

【0 2 0 9】

デジタル原稿画像において垂直な向きに位置している文字列 (t と表す) の場合は、誤差量の例として、

下記式に従って計算される平均絶対誤差 (Mean Absolute Error (MAE))

50

【 0 2 1 0 】

【 数 1 6 】

$$e_{MAE}^{top}(t) = \frac{1}{N} \sum_{i=1}^N |x_1(i) - \text{ceil}(t)|, \quad e_{MAE}^{bottom}(t) = \frac{1}{N} \sum_{i=1}^N |x_2(i) - \text{floor}(t)|$$

【 0 2 1 1 】

下記式に従って計算される 2 乗平均誤差 (Mean-Square Error (MSE))

【 0 2 1 2 】

【 数 1 7 】

$$e_{MAE}^{top}(t) = \frac{1}{N} \sum_{i=1}^N (x_1(i) - \text{ceil}(t))^2, \quad e_{MAE}^{bottom}(t) = \frac{1}{N} \sum_{i=1}^N (x_2(i) - \text{floor}(t))^2$$

【 0 2 1 3 】

下記式に従って計算される 2 乗平均平方根誤差 (Root Mean-Square Error (RMSE))

【 0 2 1 4 】

【 数 1 8 】

$$e_{RMSE}^{top}(t) = \sqrt{e_{MSE}^{top}(t)}, \quad e_{RMSE}^{bottom}(t) = \sqrt{e_{MSE}^{bottom}(t)}$$

【 0 2 1 5 】

およびその他の誤差量が挙げられる。

【 0 2 1 6 】

数字列の場合は、誤差量は小さいと予期される。図 2 1 に示すように、本形態では、このことを用いて、大文字及び小文字が混ざったテキストの候補を除外している。まず、本形態では、文字の連結成分を含む文字列要素を取得する (3 3 2)。このような文字列要素は、原稿画像のデータについて、例えば、特許文献 1 に記載のような公知技術を用いることで抽出される。

【 0 2 1 7 】

次に、文字列要素において、各連結成分をラベリングし、当該連結成分の非背景画素すべてを含む最小の境界枠を特定する (3 3 4)。ここで、境界枠は、連結成分の外接矩形である。そして、外接矩形は、原稿画像の縦方向および横方向に平行な辺を有し、連結成分に外接する矩形である。

【 0 2 1 8 】

この時点では、連結成分として抽出された文字の向きが不明である。そのため、誤差量において境界枠のどの端の組合せを用いなければならないのかはわからない。そこで、本形態では、標本平均を、x 座標の組合せと y 座標の組合せ双方のために、下記式に従って計算する。

【 0 2 1 9 】

【 数 1 9 】

$$\mu_y^{(1)} = \frac{1}{N} \sum_{i=1}^N y_1(i), \quad \mu_y^{(2)} = \frac{1}{N} \sum_{i=1}^N y_2(i), \quad \text{and} \quad \mu_x^{(1)} = \frac{1}{N} \sum_{i=1}^N x_1(i), \quad \mu_x^{(2)} = \frac{1}{N} \sum_{i=1}^N x_2(i),$$

10

20

30

40

50

【 0 2 2 0 】

上記式において、Nは文字列における要素の数である。図24は、水平方向の文字列における上部y座標の平均値（標本平均）と下部y座標の平均値（標本平均）とを示す図である。

【 0 2 2 1 】

誤差は、各端ごとに、境界枠の端の座標と、対応する標本平均とを用いて算出される（338）。各軸ごとの累積誤差は、当該軸の2つの誤差値の合計として、下記式に従って算出される（340）。

【 0 2 2 2 】

平均絶対誤差のときの累積誤差は、以下のとおりである。

10

【 0 2 2 3 】

【 数 2 0 】

$$Error_{horizontal} = \frac{1}{N} \left(\sum_{i=1}^N |y_1(i) - \text{ceil}(t)| + \sum_{i=1}^N |y_2(i) - \text{floor}(t)| \right) = e^{top} + e^{bottom}$$

【 0 2 2 4 】

【 数 2 1 】

20

$$Error_{vertical} = \frac{1}{N} \left(\sum_{i=1}^N |x_1(i) - \text{ceil}(t)| + \sum_{i=1}^N |x_2(i) - \text{floor}(t)| \right) = e^{top} + e^{bottom}$$

【 0 2 2 5 】

2乗平均誤差のときの累積誤差は、以下のとおりである。

【 0 2 2 6 】

【 数 2 2 】

$$Error_{horizontal} = \frac{1}{N} \left(\sum_{i=1}^N (y_1(i) - \text{ceil}(t))^2 + \sum_{i=1}^N (y_2(i) - \text{floor}(t))^2 \right) = e^{top} + e^{bottom}$$

30

【 0 2 2 7 】

【 数 2 3 】

$$Error_{vertical} = \frac{1}{N} \left(\sum_{i=1}^N (x_1(i) - \text{ceil}(t))^2 + \sum_{i=1}^N (x_2(i) - \text{floor}(t))^2 \right) = e^{top} + e^{bottom}$$

40

【 0 2 2 8 】

2乗平均平方根誤差のときの累積誤差は、以下のとおりである。

【 0 2 2 9 】

【 数 2 4 】

$$Error_{horizontal} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_1(i) - \text{ceil}(t))^2} + \sqrt{\frac{1}{N} \sum_{i=1}^N (y_2(i) - \text{floor}(t))^2} = e^{top} + e^{bottom}$$

【 0 2 3 0 】

50

【数 2 5】

$$Error_{vertical} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_1(i) - ceil(t))^2} + \sqrt{\frac{1}{N} \sum_{i=1}^N (x_2(i) - floor(t))^2} = e^{top} + e^{bottom}$$

【0 2 3 1】

なお、数 2 0 , 2 2 , 2 4 において、 $ceil(t)$ は数 1 9 に示した $\mu_y^{(1)}$ であり、 $floor(t)$ は数 1 9 に示した $\mu_y^{(2)}$ である。一方、数 2 1 , 2 3 , 2 5 において、 $ceil(t)$ は数 1 9 に示した $\mu_x^{(1)}$ であり、 $floor(t)$ は数 1 9 に示した $\mu_x^{(2)}$ である。

10

【0 2 3 2】

次に、水平軸に対応する誤差値と垂直軸に対応する誤差値とのうちの小さい方の誤差値を選択し、選択した誤差値に対応する軸の方向を文字列の方向として決定する (3 4 2)。そして、選択された誤差値が、列の種類の決定のために用いられる (3 4 4)。

【0 2 3 3】

列の種類は、最小の累積誤差に関連する誤差値に基づいて推測される (3 4 4)。本形態では、 e^{top} と e^{bottom} がともに所定の閾値 (T_v と表される) より小さいときは、列の種類が、数字列かすべて大文字の文字列だと見なされる。もしこの条件が満たされないときは、列の種類が、大文字と小文字が混ざり合ったものだと見なされる。

【0 2 3 4】

本形態では、平均絶対誤差量を用いて文字列における変動を測定する場合、例えば、閾値 $T_v = 1.5$ が用いられる。

20

【0 2 3 5】

本形態では、文字列をさらに分析して、大文字のみからなる文字列を除外する。こうした実施形態のいくつかでは、文字列における個々の連結成分の縦横比に対する制限を用いて、大文字のみからなる文字列と数字からなる文字列とを区別する。

【0 2 3 6】

これらの実施形態では、文字列 t における各連結成分 i の縦横比 $AR_t(i)$ は、当該連結成分の境界枠の対角線上に位置する角の座標を用いて、下記式に従って算出される。

【0 2 3 7】

30

【数 2 6】

$$AR_t(i) = \frac{\max(|x_2^{(i)} - x_1^{(i)}|, |y_2^{(i)} - y_1^{(i)}|)}{\min(|x_2^{(i)} - x_1^{(i)}|, |y_2^{(i)} - y_1^{(i)}|)}$$

【0 2 3 8】

そして、文字列に含まれる全ての連結成分の縦横比の平均 μ_{AR} と平方偏差 (分散) σ_{AR}^2 が算出される。本形態では、 $\mu_{AR} > T_{AR\mu}$ かつ $\sigma_{AR}^2 < T_{AR}$ のときは、文字列は数字列としてラベリングされ、そうでないときはそれ以外の列としてラベリングされる。 $T_{AR\mu}$ と T_{AR} はさまざまな方法で決定される閾値である。たとえば、トレーニングデータを分析することによって、専門家によって開発される発見的 (heuristic) 方法に従って、あるいは他の方法によって、予め決定される閾値である。これは、数字は、大文字に比べて、縦横比が相対的に大きく、かつ、その縦横比の数字の種類による変化が小さいことに基づいている。

40

【0 2 3 9】

また、本形態の一具体例では、文字列における連結成分の数の制限を用いて、ある文字列が数字文字列かどうかを判定してもよい。本実施形態では、文字列における連結成分素

50

の数 N が閾値 T_{length} を上回るときは、その列はそれ以上考慮されない。本発明のいくつかの実施形態では、 $T_{length} = 24$ である。これは、数字列では、数字の個数が通常 24 以下であり、24 を越える場合はアルファベットからなるテキストである確率が非常に高いことに基づいている。

【0240】

本形態の一具体例では、文字列とその要素は、以下の条件 1 ~ 4 のすべてが満たされた場合、数字列だと見なしてもよい。

条件 1 . $e^{top} > T_v$ かつ $e^{bottom} < T_v$

条件 2 . $\mu_{AR} > T_{AR}$

条件 3 . $\sigma_{AR}^2 < T_{AR}$

条件 4 . $N < T_{length}$

あるいは、条件 2 および条件 3 のみが満たされた場合に、数字列だと見なしても良い。あるいは、条件 2、条件 3 および条件 4 のみが満たされた場合に、数字列だと見なしても良い。あるいは、条件 1、条件 2 および条件 3 のみが満たされた場合に、数字列だと見なしても良い。あるいは、条件 1 および条件 4 のみが見たされた場合に、数字列だと見なしても良い。あるいは、条件 2、条件 3 および条件 4 のみが満たされた場合に、数字列だと見なしても良い。あるいは、条件 2 のみが満たされた場合に、数字列だと見なしても良い。あるいは、条件 3 のみが満たされた場合に、数字列だと見なしても良い。あるいは、条件 1 および条件 2 のみが満たされた場合に、数字列だと見なしても良い。あるいは、条件 1 および条件 3 のみが満たされた場合に、数字列だと見なしても良い。

【0241】

一般に、アラビア数字の全ては、アルファベットの大文字および小文字に比べて、縦横比が大きい。アルファベットの小文字の中には、「b」「j」のように縦横比の大きなものも存在するが、「a」「c」「s」のように縦横比の小さいものも存在する。そのため、文字列の中の複数の連結成分の縦横比の平均値を取る場合、アルファベットの小文字からなる文字列のときの平均値は、アラビア数字からなる文字列のときの平均値よりも小さくなる。そのため、条件 2 のみであっても、アラビア数字の文字列を抽出できる。

【0242】

また、アラビア数字の全ては、アルファベットの大文字および小文字に比べて、ほぼ同じ縦横比を有する。一方、アルファベットの小文字の中には、「b」「j」のように縦横比の大きなものも存在するが、「a」「c」「s」のように縦横比の小さいものも存在する。また、アルファベットの大文字の中には、「J」「I」のように縦横比の大きなものも存在するが、「M」「C」「W」のように縦横比の小さいものも存在する。そのため、文字列の中の複数の連結成分の縦横比の分散値を取る場合、アルファベットからなる文字列のときの分散値は、アラビア数字からなる文字列のときの分散値よりも大きくなる。そのため、条件 3 のみであっても、アラビア数字の文字列を抽出できる。

【0243】

なお、分散値の代わりに標準偏差を用いても良い。

【0244】

本発明のいくつかの実施形態では、ある文字列についてすべての条件が調べられる。本発明の他の実施形態では、上記条件は漸次調べられる、すなわち現在調べられている条件が満たされたときのみ、次の条件が調べられる。

【0245】

なお、本処理例は、上記のようにアルファベットの大文字・小文字、数字の形状の特徴に着目したものである。そのため、本処理例により数字列抽出の対象となる原稿画像は、アルファベットおよび数字からなる原稿画像であることが好ましい。

【0246】

(本発明の適用例)

次に本発明の適用例について説明する。図 25 は、本発明に係る画像の向き判定方法を適用した原稿方向判定装置 500 および周辺装置を含む原稿方向判定システムの構成を

10

20

30

40

50

示すブロック図である。

【0247】

図25に示されるように、原稿方向判定システムは、数字列抽出装置400と、原稿方向判定装置500と、画像表示装置(画像出力装置)600と、画像形成装置(画像出力装置)700と、画像処理装置(画像出力装置)800とを備える。

【0248】

数字列抽出装置400は、アルファベットおよび数字からなるテキストを含む原稿画像の電子データから、数字列を抽出するものである。すなわち、数字列抽出装置400は、原稿画像の電子データに基づいて、例えば特許文献1に記載された方法に従って文字列(なお列の方向は、原稿画像の縦方向または横方向である)を抽出する。そして、数字列抽出装置400は、抽出した文字列に含まれる各連結成分を特定し、ラベリングする。また、数字列抽出装置400は、各連結成分について、外接矩形である境界枠の位置を特定する。その後、数字列抽出装置400は、数19~25に示されるように、累積誤差を算出し、その誤差値に基づいて、大文字および小文字の双方を含む文字列であるか否かを判定する。次に、数字列抽出装置400は、大文字および小文字の双方を含む文字列ではないと判定した文字列について、上述したように以下の条件1-4の全てを満たす文字列を数字列として抽出する。

条件1 . $e^{top} T_v$ かつ $e^{bottom} T_v$

条件2 . $\mu_{AR} > T_{AR\mu}$

条件3 . $\sigma_{AR}^2 < T_{AR\sigma^2}$

条件4 . $N < T_{length}$

そして、数字列抽出装置400は、抽出した数字列の位置を特定する情報を、原稿方向判定装置500に出力する。なお、数字列抽出装置400の具体的な処理方法は、上記(数字文字の検出について)で述べたとおりである。

【0249】

なお、数字列抽出装置400は、上記の条件1-4の全てを満たす文字列を数字列として抽出する場合に限定されない。すなわち、数字列抽出装置400は、上記の条件13のみを満たす文字列を数字列として抽出してもよい。あるいは、数字列抽出装置400は、上記の条件2および条件3のみを満たす文字列を数字列として抽出してもよい。あるいは、数字列抽出装置400は、上記の条件2、条件3および条件4のみを満たす文字列を数字列として抽出してもよい。あるいは、数字列抽出装置400は、上記の条件1、条件2および条件3のみを満たす文字列を数字列として抽出してもよい。あるいは、数字列抽出装置400は、上記の条件1および条件4のみを満たす文字列を数字列として抽出してもよい。あるいは、数字列抽出装置400は、上記の条件2、条件3および条件4のみを満たす文字列を数字列として抽出してもよい。あるいは、数字列抽出装置400は、上記の条件2のみを満たす文字列を数字列として抽出してもよい。あるいは、数字列抽出装置400は、上記の条件3のみを満たす文字列を数字列として抽出してもよい。あるいは、数字列抽出装置400は、上記の条件1および条件2のみを満たす文字列を数字列として抽出してもよい。あるいは、数字列抽出装置400は、上記の条件1および条件3のみを満たす文字列を数字列として抽出してもよい。

【0250】

原稿方向判定装置500は、入力された原稿画像の電子データについて、原稿の向きを判定するものである。原稿方向判定装置500は、数字テンプレート作成部501、数字テンプレート記憶部502、連結成分抽出部503、サイズ正規化部504、距離算出部505、照合部506、推定部507、信頼度判定部508を備えている。

【0251】

数字テンプレート作成部501は、上記(実施形態1・2における数字テンプレートの作成方法)に従って、数字テンプレートを作成し、数字テンプレート記憶部502に格納するものである。

【0252】

10

20

30

40

50

連結成分抽出部 503 は、上記実施形態 1～6 で説明したように、原稿画像の中から、連結成分の外接矩形で囲まれる部分画像を切り出す処理を行うものである。なお、数字列抽出装置 400 から数字列の位置を特定する情報を受けた場合、当該情報で示される位置からのみ部分画像を切り出すものとする。

【0253】

サイズ正規化部 504 は、上記実施形態 1～6 で説明したように、連結成分抽出部 503 で切り出された部分画像について、数字テンプレート記憶部 502 に格納されている数字テンプレートと同じサイズにサイズ正規化を行うものである。

【0254】

距離算出部 505 は、上記実施形態 1・2 で説明したように、サイズ正規化された部分画像について、距離算出処理を実行し、非背景画素からの距離を示す値を画素値として割り振るものである。

【0255】

照合部 506 は、上記実施形態 1・2 で説明したように、距離算出処理が実行された部分画像と数字テンプレートとの照合を行い、類似度を求めるものである。そして、照合部 506 は、類似度に基づいて、各向き的一致カウンタのカウンタ値を増やす処理を行う。

【0256】

推定部 507 は、上記実施形態 1～6 で説明したように、照合部 506 により更新された各向き的一致カウンタのカウンタ値に基づいて、原稿画像の向きを推定するものである。

【0257】

信頼度判定部 508 は、上記実施形態 2・4・6 で説明したように、信頼度を求め、推定部 507 で推定された向き（推定向き）を出力するか否かを判断するものである。

【0258】

画像表示装置 600 は、原稿方向判定装置 500 から出力された推定向きに従って、入力された原稿画像の回転処理を実行し、表示部に表示させるものである。画像表示装置 600 は、原稿画像の天地方向と、表示部の天地方向とが一致するように原稿画像を回転させる。具体的には、画像表示装置 600 は、原稿画像の天地方向と表示部の天地方向とが一致するように、推定向きと回転角度とが予め対応付けられた情報を記憶している。そして、画像表示装置 600 は、当該情報に基づいて、原稿方向判定装置 500 から出力された推定向きに対応する回転角度だけ原稿画像を回転させ、表示させる。これにより、原稿画像の天地方向と、表示部の天地方向とが一致した状態で、原稿画像を表示させることができる。

【0259】

画像形成装置 700 は、原稿方向判定装置 500 から出力された推定向きに従って、入力された原稿画像の回転処理を実行し、用紙に当該原稿画像を印刷するものである。画像形成装置 700 は、用紙に対する原稿画像の天地方向が常に一定になるように原稿画像を回転させる。具体的には、画像形成装置 700 は、用紙に対する原稿画像の天地方向が常に一定になるように、推定向きと回転角度とが予め対応付けられた情報を記憶している。そして、画像形成装置 700 は、当該情報に基づいて、原稿方向判定装置 500 から出力された推定向きに対応する回転角度だけ原稿画像を回転させ、印刷する。これにより、用紙に対する原稿画像の天地方向が常に一定になる。そのため、あるページだけが反転されていても、当該ページの画像だけが回転処理され、印刷物における原稿画像の天地方向は全てのページにおいて同じになる。そのため、ステーブルやパンチングなどの後処理を行う機能を画像形成装置 700 が有していたとしても、全ての用紙に印刷された原稿画像の向きが一致した状態で綴じることができる。

【0260】

画像処理装置 800 は、入力された原稿画像の電子データに対して、所定の画像処理を行い、処理結果を出力するものである。例えば、画像処理装置 800 は、原稿画像の中からテキストデータを抽出する OCR 処理を実行する。このとき、画像処理装置 800 は、原

10

20

30

40

50

稿方向判定装置 500 から出力された推定向きに従って OCR 処理を実行する。これにより、画像処理装置 800 は、適切な方向に従ったテキストデータを生成することができる。

【0261】

なお、原稿方向判定システムは、数字列抽出装置 400 を備えていなくてもよい。この場合、原稿方向判定装置 500 の連結成分抽出部 503 は、原稿画像の中の連結成分を全て抽出すればよい。

【0262】

また、原稿方向判定装置 500 は、距離算出部 505 を備えていなくてもよい。この場合、数字テンプレート作成部 501 は、上記の（別の実施形態における数字テンプレートの作成方法）に従って数字テンプレートを作成すればよい。また、照合部 506 は、上記実施形態 4～6 で説明したように、サイズ正規化された部分画像と数字テンプレートとの照合を行い、類似度を求めればよい。

【0263】

また、原稿方向判定装置 500 は、信頼度判定部 508 を備えていなくてもよい。この場合、推定部 507 は、決定した推定向きを外部装置（画像表示装置 600、画像形成装置 700、画像処理装置 800 など）に出力すればよい。

【0264】

上記の明細書において用いられた用語と表現は本発明を記述するためのものであり、限定するためのものではなく、そうした用語と表現の使用には、提示され記述された諸特徴の同等物またはその一部を除外する意図はなく、本発明の範囲は以下に述べる特許請求の範囲によってのみ定義され限定される。

【0265】

なお、本発明は、以下のようにも表現できる。

【0266】

(1) 本発明は、デジタル画像内の数字連結成分を検知する方法であって、a) デジタル画像内の複数の連結成分を含むテキスト行要素を受け取る工程と、b) 上記複数の連結成分のそれぞれに要素境界枠を決定する工程であって、当該要素境界枠のそれぞれは、第一辺座標、第二辺座標、第三辺座標、第四辺座標を含んでおり、当該第一辺座標と当該第二辺座標は当該境界枠の第一の軸と関連し、当該第三辺座標と当該第四辺座標は当該境界枠の第二の軸と関連する工程と、c) 上記第一辺座標と関連する第一の変動量を判定する工程と、d) 上記第二辺座標と関連する第二の変動量を判定する工程と、e) 上記第三辺座標と関連する第三の変動量を判定する工程と、f) 上記第四辺座標と関連する第四の変動量を判定する工程と、g) 上記第一の変動量と上記第二の変動量との第一の累積を判定する工程と、h) 上記第三の変動量と上記第四の変動量との第二の累積を判定する工程と、i) 上記第一の累積と上記第二の累積とが第一の基準を満たさず場合、i) 上記第一の変動量が第一の閾値基準を満たし、かつ上記第二の変動量が第二の閾値基準を満たさずならば、上記テキスト行要素を数字要素に分類し、ii) 上記第一の変動量が上記第一の閾値基準を満たさないか、上記第二の変動量が上記第二の閾値基準を満たさないならば、上記テキスト行要素を非数字要素に分類する工程と、j) 上記第一の累積と上記第二の累積とが上記第一の基準を満たさない場合、i) 上記第三の変動量が第三の閾値基準を満たし、上記第四の変動量が第四の閾値基準を満たさずならば、上記テキスト行要素を数字要素に分類し、ii) 上記第三の変動量が上記第三の閾値基準を満たさないか、上記第四の変動量が上記第四の閾値基準を満たさないならば、上記テキスト行要素を非数字要素に分類する工程と、を含む。

【0267】

(2) なお、a) 上記第一辺座標と関連する第一の変動量を判定する上記工程は、上記第一辺座標の第一代表値を計算する工程を含み、b) 上記第二辺座標と関連する第二の変動量を判定する上記工程は、上記第二辺座標の第二代表値を計算する工程を含み、c) 上記第三辺座標と関連する第三の変動量を判定する上記工程は、上記第三辺座標の第三代表値を計算する工程を含み、d) 上記第四辺座標と関連する第四の変動量を判定する上記工

10

20

30

40

50

程は、上記第四辺座標の第四代表値を計算する工程を含むことが好ましい。

【0268】

(3) また、a) 上記第一代表値は上記第一辺座標の標本平均であり、b) 上記第二代表値は上記第二辺座標の標本平均であり、c) 上記第三代表値は上記第三辺座標の標本平均であり、d) 上記第四代表値は上記第四辺座標の標本平均であることが好ましい。

【0269】

(4) また、a) 第一の変動量を判定する上記工程は、上記第一の代表値と上記第一辺座標との第一誤差量を算出する工程を含み、b) 第二の変動量を判定する上記工程は、上記第二の代表値と上記第二辺座標との第二誤差量を算出する工程を含み、c) 第三の変動量を判定する上記工程は、上記第三の代表値と上記第三辺座標との第三誤差量を算出する工程を含み、d) 第四の変動量を判定する上記工程は、上記第四の代表値と上記第四辺座標との第四誤差量を算出する工程を含むことが好ましい。

10

【0270】

(5) また、a) 上記第一誤差量は、平均絶対誤差、二乗平均誤差、二乗平均平方根誤差から選ばれる誤差量であり、b) 上記第二誤差量は、平均絶対誤差、二乗平均誤差、二乗平均平方根誤差から選ばれる誤差量であり、c) 上記第三誤差量は、平均絶対誤差、二乗平均誤差、二乗平均平方根誤差から選ばれる誤差量であり、d) 上記第四誤差量は、平均絶対誤差、二乗平均誤差、二乗平均平方根誤差から選ばれる誤差量であることが好ましい。

【0271】

(6) また、上記(1)において、上記テキスト行要素が数字要素に分類された場合、i) 上記複数の連結成分のそれぞれの縦横比を算出し、それによって複数の縦横比を作り出す工程と、ii) 上記複数の縦横比の平均を算出する工程と、iii) 上記複数の縦横比の平方偏差を算出する工程と、iv) 上記平均と上記平方偏差に基づいて、上記テキスト行要素を再分類する工程と、をさらに含むことが好ましい。

20

【0272】

(7) そして、上記(6)において、上記テキスト行要素を再分類する上記工程は、a) 上記平均が第一の平均基準を満たし、かつ上記平方偏差が第一の平方偏差基準を満たすときは、上記テキスト行要素を数字に再分類し、b) 上記平均が上記第一の平均基準を満たさないか、上記平方偏差が上記第一の平方偏差基準を満たさないときは、上記テキスト行要素を非数字に再分類する工程を含むことが好ましい。

30

【0273】

(8) そして、上記(7)において、上記テキスト行要素が数字要素に再分類されたとき、上記複数の連結成分における連結成分の数に基づいて、上記テキスト行要素を再分類する工程をさらに含むことが好ましい。

【0274】

(9) または、上記(1)において、上記テキスト行要素が数字要素に分類されたとき、上記複数の連結成分における連結成分の数に基づいて、上記テキスト行要素を再分類する工程をさらに含むことが好ましい。

【0275】

(10) そして、上記(9)において、上記テキスト行要素を再分類する上記工程は、a) 連結成分の上記数が第一の量基準を満たすときは、上記テキスト行要素を数字要素に再分類し、b) 連結成分の上記数が上記第一の量基準を満たさないときは、上記テキスト行要素を非数字要素に再分類することが好ましい。

40

【0276】

(11) また、本発明は、デジタル画像内の数字連結成分を検知する方法であって、a) デジタル画像内の複数の連結成分を含むテキスト行要素を受け取る工程と、b) 上記複数の連結成分における上記連結成分のそれぞれの縦横比を算出し、それにより複数の縦横比を作り出す工程と、c) 上記複数の縦横比の第一の特徴を算出する工程と、d) 上記第一の特徴が第一の基準を満たすときは、上記テキスト行要素を数字要素に分類する工程と

50

、 e) 上記第一の特徴が第一の基準を満たさないときは、上記テキスト行要素を非数字要素に分類する工程と、を含む。

【0277】

(12) 上記(11)において、上記第一の特徴は、上記複数の縦横比の平均である。

【0278】

(13) もしくは、上記(11)において、上記第一の特徴は、上記複数の縦横比の標準偏差である。

【0279】

(14) もしくは、上記(11)において、上記複数の縦横比の第二の特徴を算出する工程をさらに含み、上記第一の基準は、第二の基準を満たす上記第一の特徴と第三の基準を満たす上記第二の特徴とを含むことが好ましい。

10

【0280】

(15) そして、上記(14)において、上記第二の特徴は、平均と平方偏差から選ばれることが好ましい。

【0281】

(16) また、上記(11)において、上記複数の連結成分における連結成分の数を判定する工程と、上記複数の縦横比の第二の特徴を算出する工程と、をさらに含み、上記第一の基準は、第二の基準を満たす上記第一の特徴と、第三の基準を満たす上記第二の特徴と、量基準を満たす連結成分の上記数と、を含むことが好ましい。

【0282】

(17) そして、上記(16)において、 a) 上記複数の連結成分のそれぞれに要素境界枠を決定する工程であって、当該要素境界枠のそれぞれは、第一辺座標、第二辺座標、第三辺座標、第四辺座標を含んでおり、当該第一辺座標と当該第二辺座標は当該境界枠の第一の軸と関連し、当該第三辺座標と当該第四辺座標は当該境界枠の第二の軸と関連する工程と、 b) 上記第一辺座標と関連する第一の変動量を判定する工程と、 c) 上記第二辺座標と関連する第二の変動量を判定する工程と、 d) 上記第三辺座標と関連する第三の変動量を判定する工程と、 e) 上記第四辺座標と関連する第四の変動量を判定する工程と、 f) 上記第一の変動量と上記第二の変動量との第一の累積を判定する工程と、 g) 上記第三の変動量と上記第四の変動量との第二の累積を判定する工程と、 h) 上記第一の累積と上記第二の累積とが第一の累積基準を満たす場合、 i) 上記第一の変動量と等しいように第一の変動特徴を設定し、 i i) 上記第二の変動量と等しいように第二の変動特徴を設定する工程と、 i) 上記第一の累積と上記第二の累積とが第一の累積基準を満たさない場合、 i) 上記第三の変動量と等しいように上記第一の変動特徴を設定し、 i i) 上記第四の変動量と等しいように上記第二の変動特徴を設定する工程と、をさらに含み、 j) 上記第一の基準は、上記第二の基準を満たす上記第一の特徴と、上記第三の基準を満たす上記第二の基準と、上記量基準を満たす連結成分の上記数と、第四の基準を満たす上記第一の変動特徴と、第五の基準を満たす上記第二の変動特徴と、を含むことが好ましい。

20

30

【0283】

(18) そして、上記(17)において、 a) 第一の変動量を判定する上記工程は、 i) 上記第一辺座標の第一の代表値を算出する工程と、 i i) 上記第一の代表値と上記第一辺座標との第一の誤差量を算出する工程を含み、 b) 第二の変動量を判定する上記工程は、 i) 上記第二辺座標の第二の代表値を算出する工程と、 i i) 上記第二の代表値と上記第二辺座標との第二の誤差量を算出する工程を含み、 c) 第三の変動量を判定する上記工程は、 i) 上記第三辺座標の第三の代表値を算出する工程と、 i i) 上記第三の代表値と上記第三辺座標との第三の誤差量を算出する工程を含み、 d) 第四の変動量を判定する上記工程は、 i) 上記第四辺座標の第四の代表値を算出する工程と、 i i) 上記第四の代表値と上記第四辺座標との第四の誤差量を算出する工程と、を含むことが好ましい。

40

【0284】

(19) また、上記(11)において、 a) 上記複数の連結成分のそれぞれに要素境界枠を決定する工程であって、当該要素境界枠のそれぞれは、第一辺座標、第二辺座標、第

50

三辺座標、第四辺座標を含んでおり、当該第一辺座標と当該第二辺座標は当該境界枠の第一の軸と関連し、当該第三辺座標と当該第四辺座標は当該境界枠の第二の軸と関連する工程と、b)上記第一辺座標と関連する第一の変動量を判定する工程と、c)上記第二辺座標と関連する第二の変動量を判定する工程と、d)上記第三辺座標と関連する第三の変動量を判定する工程と、e)上記第四辺座標と関連する第四の変動量を判定する工程と、f)上記第一の変動量と上記第二の変動量との第一の累積を判定する工程と、g)上記第三の変動量と上記第四の変動量との第二の累積を判定する工程と、h)上記第一の累積と上記第二の累積とが第一の累積基準を満たす場合、i)上記第一の変動量と等しいように第一の変動特徴を設定し、ii)上記第二の変動量と等しいように第二の変動特徴を設定する工程と、i)上記第一の累積と上記第二の累積とが上記第一の累積基準を満たさない場合、i)上記第三の変動量と等しいように上記第一の変動特徴を設定し、ii)上記第四の変動量と等しいように上記第二の変動特徴を設定する工程と、をさらに含み、j)上記第一の基準は、第二の基準を満たす上記第一の特徴と、第三の基準を満たす上記第二の特徴と、第四の基準を満たす上記第一の変動特徴と、第五の基準を満たす上記第二の変動特徴と、を含むことが好ましい。

10

20

30

40

50

【0285】

(20)そして、上記(19)において、a)第一の変動量を判定する上記工程は、i)上記第一辺座標の第一の代表値を算出する工程と、ii)上記第一の代表値と上記第一辺座標との第一の誤差量を算出する工程を含み、b)第二の変動量を判定する上記工程は、i)上記第二辺座標の第二の代表値を算出する工程と、ii)上記第二の代表値と上記第二辺座標との第二の誤差量を算出する工程を含み、c)第三の変動量を判定する上記工程は、i)上記第三辺座標の第三の代表値を算出する工程と、ii)上記第三の代表値と上記第三辺座標との第三の誤差量を算出する工程を含み、d)第四の変動量を判定する上記工程は、i)上記第四辺座標の第四の代表値を算出する工程と、ii)上記第四の代表値と上記第四辺座標との第四の誤差量を算出する工程と、を含むことが好ましい。

【0286】

(21)また、本発明は、デジタル画像内の数字連結成分を検知する方法であって、a)デジタル画像内の複数の連結成分を含むテキスト行要素を受け取る工程と、b)上記複数の連結成分における連結成分の数を判定する工程と、c)連結成分の上記数が量基準を満たすときは、上記テキスト行要素を数字要素に分類する工程と、d)連結成分の上記数が上記量基準を満たさないときは、上記テキスト行要素を非数字要素に分類する工程と、を含む。

【0287】

最後に、原稿方向判定装置500の各ブロック、ハードウェアロジックによって構成してもよいし、次のようにCPUを用いてソフトウェアによって実現してもよい。

【0288】

すなわち、原稿方向判定装置500は、各機能を実現する制御プログラムの命令を実行するCPU(central processing unit)、上記プログラムを格納したROM(read only memory)、上記プログラムを展開するRAM(random access memory)、上記プログラムおよび各種データを格納するメモリ等の記憶装置(記録媒体)などを備えている。そして、本発明の目的は、上述した機能を実現するソフトウェアである原稿方向判定装置500の制御プログラムのプログラムコード(実行形式プログラム、中間コードプログラム、ソースプログラム)をコンピュータで読み取り可能に記録した記録媒体を、上記原稿方向判定装置500に供給し、そのコンピュータ(またはCPUやMPU)が記録媒体に記録されているプログラムコードを読み出し実行することによっても、達成可能である。

【0289】

上記記録媒体としては、例えば、磁気テープやカセットテープ等のテープ系、フロッピー(登録商標)ディスク/ハードディスク等の磁気ディスクやCD-ROM/MO/MD/DVD/CD-R等の光ディスクを含むディスク系、ICカード(メモリカードを含む)/光カード等のカード系、あるいはマスクROM/EPROM/EEPROM/フラッ

シユROM等の半導体メモリ系などを用いることができる。

【0290】

また、原稿方向判定装置500を通信ネットワークと接続可能に構成し、上記プログラムコードを通信ネットワークを介して供給してもよい。この通信ネットワークとしては、特に限定されず、例えば、インターネット、イントラネット、エキストラネット、LAN、ISDN、VAN、CATV通信網、仮想専用網(virtual private network)、電話回線網、移動体通信網、衛星通信網等が利用可能である。また、通信ネットワークを構成する伝送媒体としては、特に限定されず、例えば、IEEE1394、USB、電力線搬送、ケーブルTV回線、電話線、ADSL回線等の有線でも、IrDAやリモコンのような赤外線、Bluetooth(登録商標)、802.11無線、HDR、携帯電話網、衛星回線、地上波デジタル網等の無線でも利用可能である。なお、本発明は、上記プログラムコードが電子的な伝送で具現化された、搬送波に埋め込まれたコンピュータデータ信号の形態でも実現され得る。

10

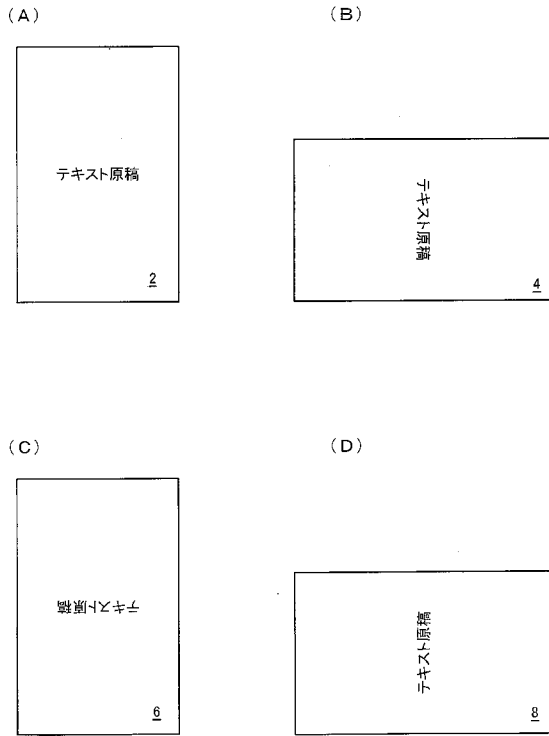
【符号の説明】

【0291】

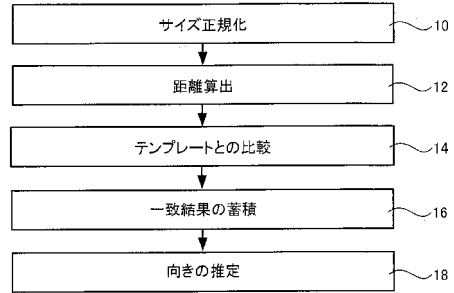
- 2・4・6・8 テキスト原稿
- 400 数字列抽出装置
- 500 原稿方向判定装置
- 501 数字テンプレート作成部
- 502 数字テンプレート記憶部
- 503 連結成分抽出部
- 504 サイズ正規化部
- 505 距離算出部
- 506 照合部
- 507 推定部
- 508 信頼度判定部
- 600 画像表示装置
- 700 画像形成装置
- 800 画像処理装置

20

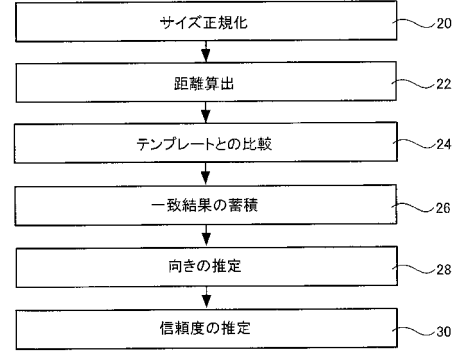
【図1】



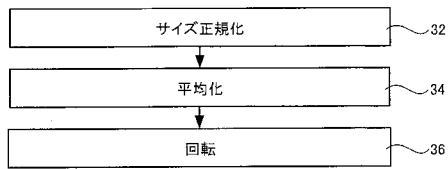
【図2】



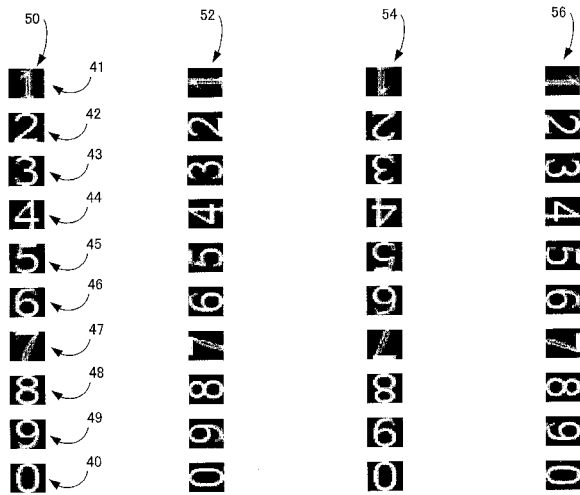
【図3】



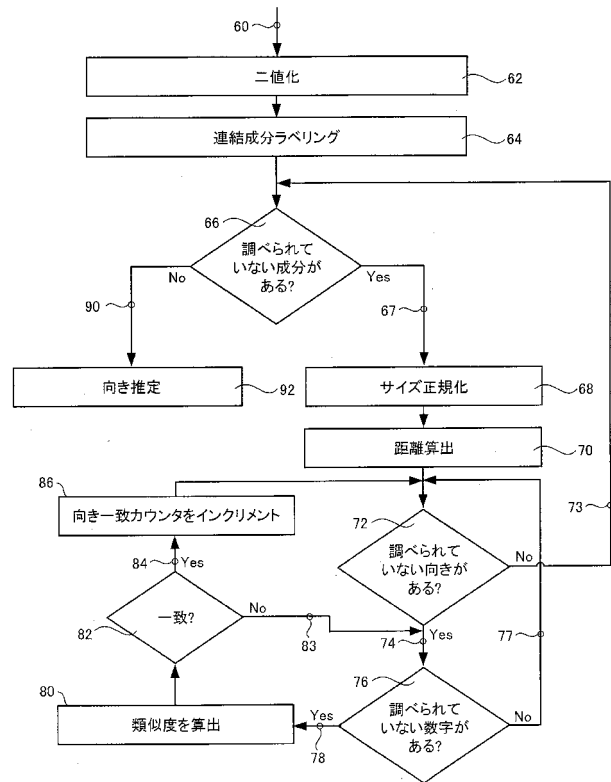
【図4】



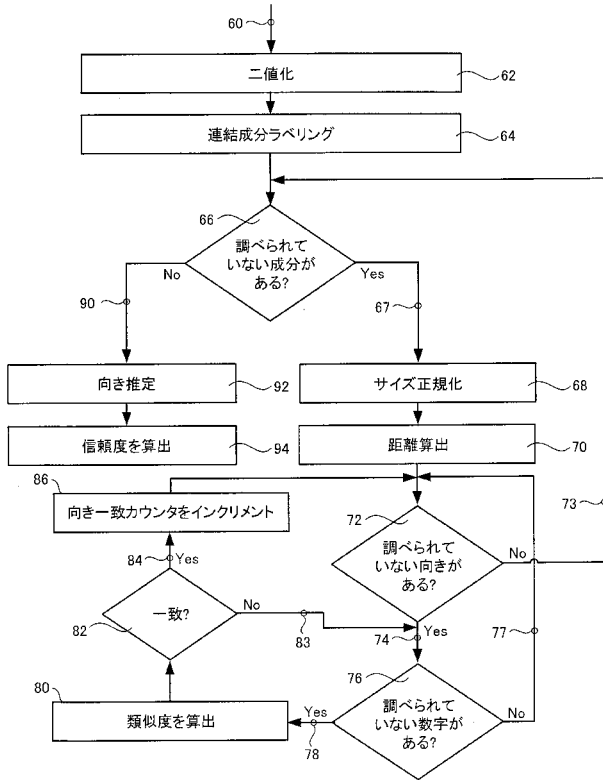
【図5】



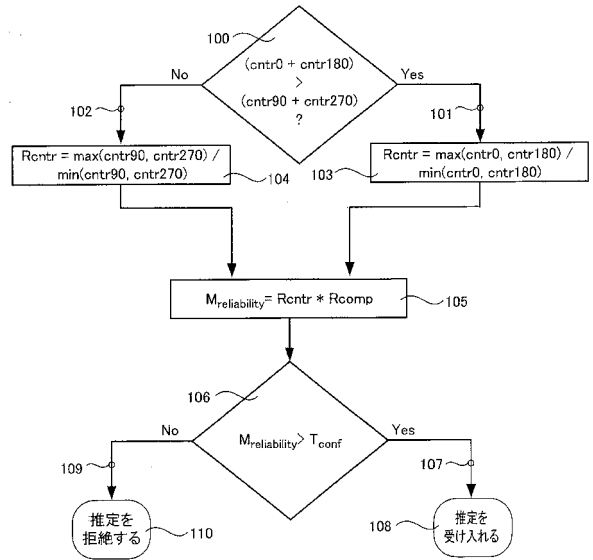
【図6】



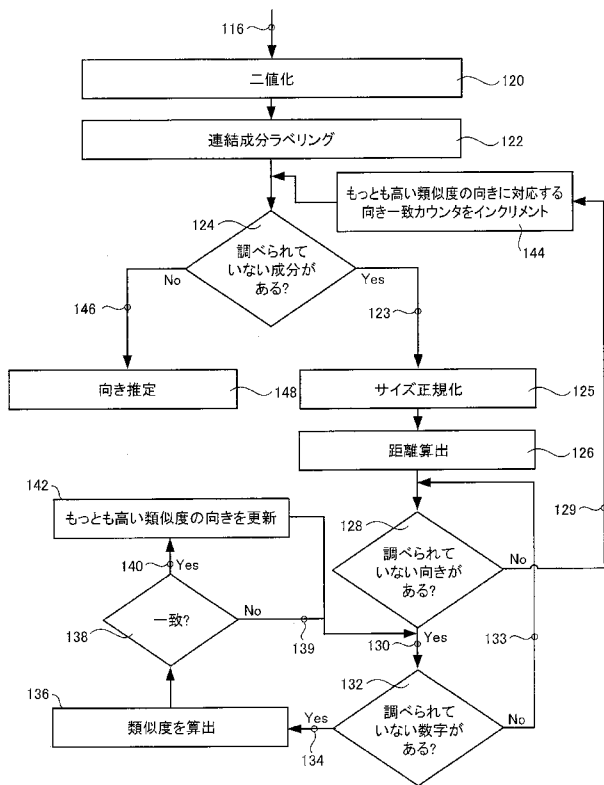
【 図 7 】



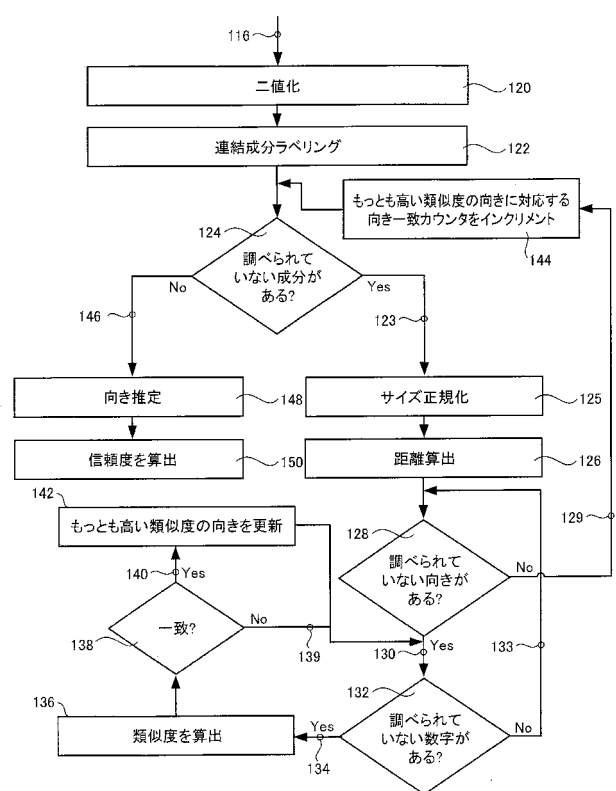
【 図 8 】



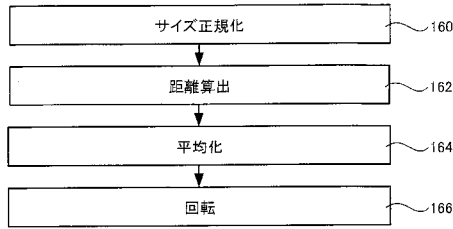
【 図 9 】



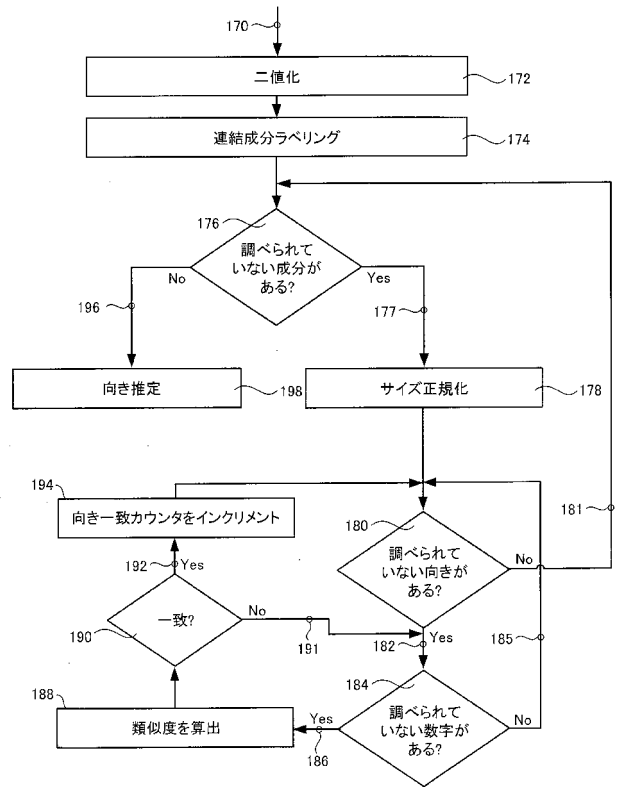
【 図 10 】



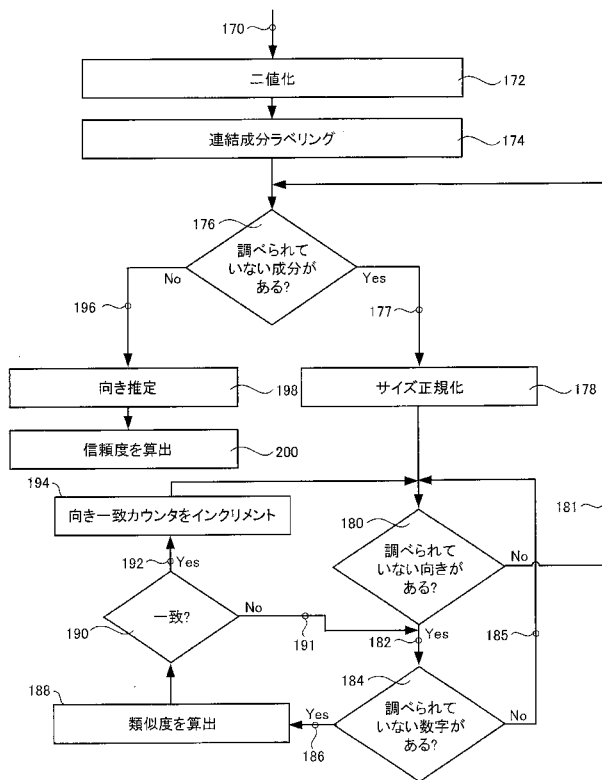
【図 1 1】



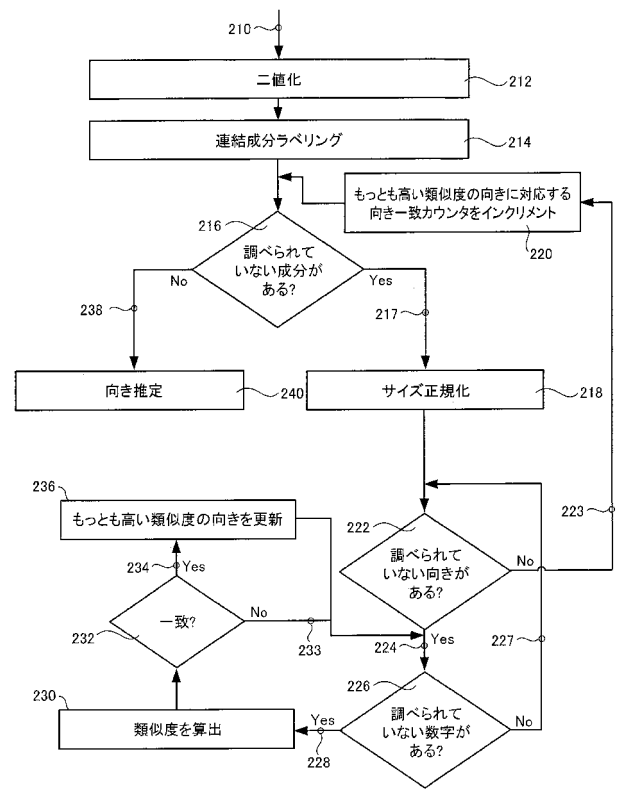
【図 1 2】



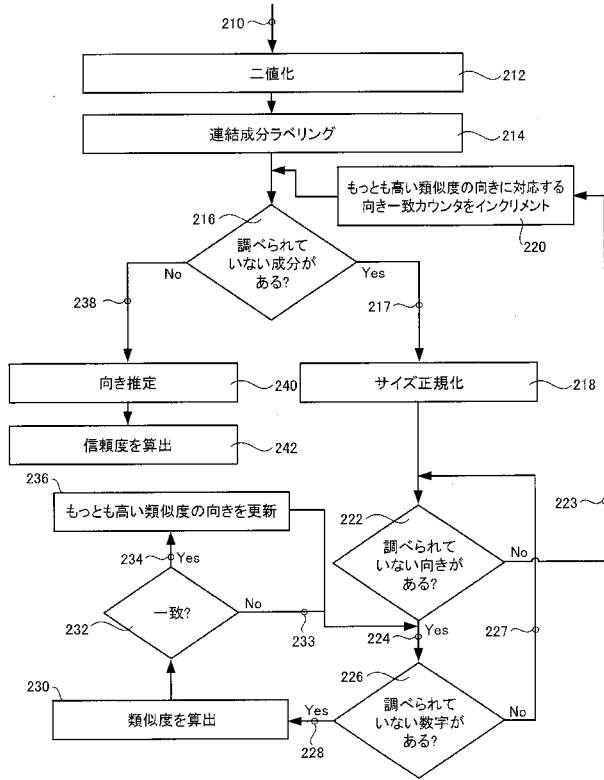
【図 1 3】



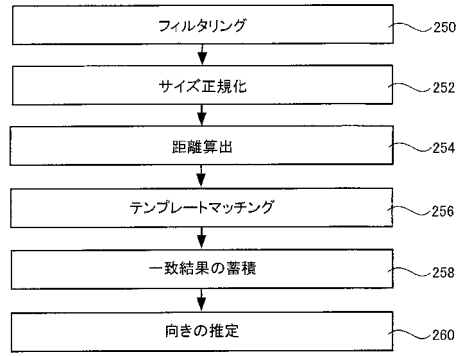
【図 1 4】



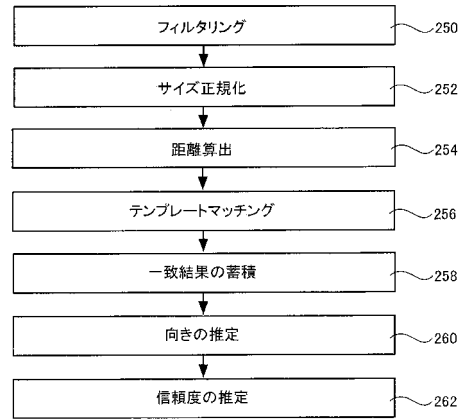
【 図 1 5 】



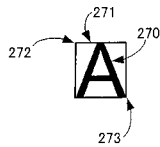
【 図 1 6 】



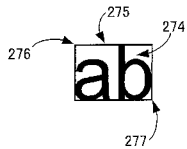
【 図 1 7 】



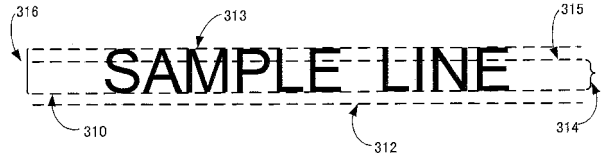
【 図 1 8 A 】



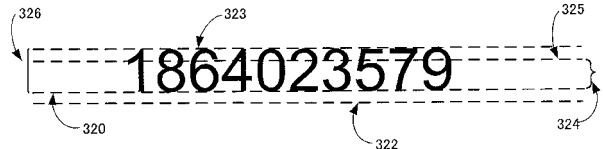
【 図 1 8 B 】



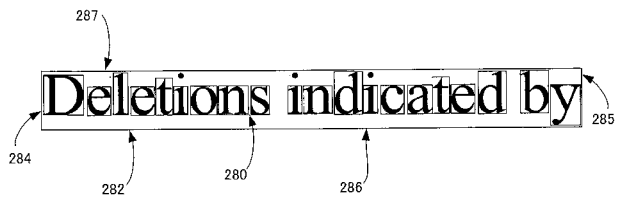
【 図 2 0 B 】



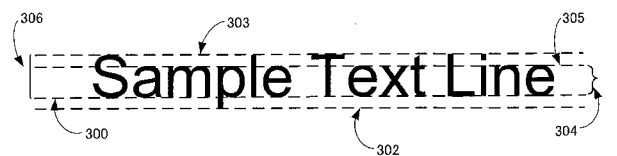
【 図 2 0 C 】



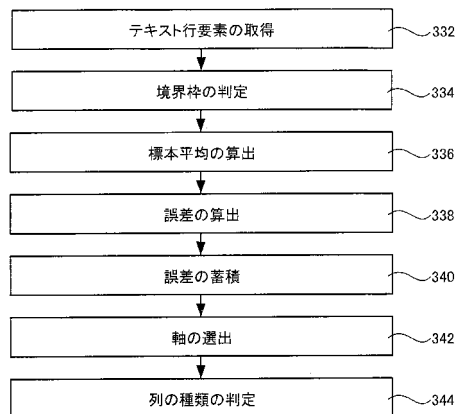
【 図 1 9 】



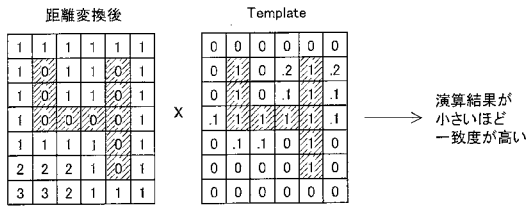
【 図 2 0 A 】



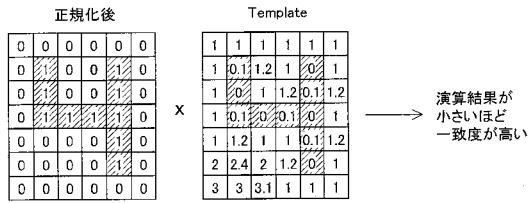
【 図 2 1 】



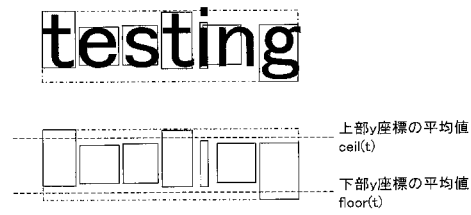
【 図 2 2 】



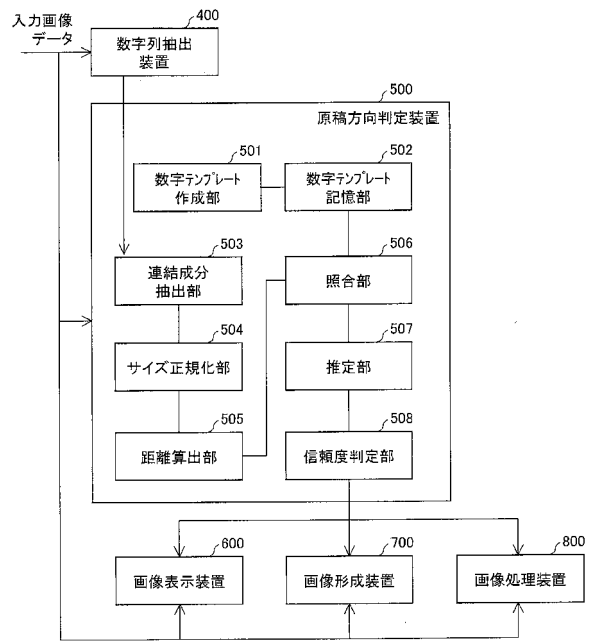
【 図 2 3 】



【 図 2 4 】



【 図 2 5 】



フロントページの続き

(72)発明者 リチャード ジョン キャンベル

アメリカ合衆国 ワシントン州 98607, カマス, ノースウエスト パシフィック リム ブ
ールバード 5750 シャープ ラボラトリーズ オブ アメリカ インコーポレイテッド内

Fターム(参考) 5B064 AA01 AB02 AB13 BA01