

文書からの書き込み抽出の精度向上と客観的評価

岩田 和将[†] 中居 友弘[†] 黄瀬 浩一[†]

[†] 大阪府立大学大学院工学研究科

〒 599-8531 大阪府堺市中区学園町 1-1

E-mail: †{iwata,nakai}@m.cs.osakafu-u.ac.jp, †kise@cs.osakafu-u.ac.jp

あらまし 文書からの書き込みの抽出手法に、書き込みを行う前の文書画像と、書き込みを行った後の文書画像の差異を見るものがある。従来手法の一つには、印刷とスキャンによる画像劣化に対し、書き込みを行う前の文書にも同じ処理を加えることで対処をしているものがある。しかし、画像劣化を等しくするために、同じ印刷機やスキャナを用いる必要があり、利便性に問題がある。本稿で提案する手法では、両画像に画像処理を加えることにより、印刷やスキャンによる差異を補正し、書き込みの抽出精度を向上させる。提案手法では、ディザリング、濃淡の変化、局所的な位置あわせの3つの処理を行う。また、実験結果の正確な比較をするため、客観的な評価手法を提案する。実験により、従来手法と比較して、再現率で 12.6%、適合率で 12.3%の抽出精度の向上が示された。

キーワード 文書画像, 画像劣化, 書き込み抽出, 客観的評価

Accuracy Improvement and Objective Evaluation of Annotation Extraction from Printed Documents

Kazumasa IWATA[†], Tomohiro NAKAI[†], and Koichi KISE[†]

[†] Graduate School of Engineering, Osaka Prefecture University

1-1 Gakuencho, Naka, Sakai, Osaka, 599-8531 Japan

E-mail: †{iwata,nakai}@m.cs.osakafu-u.ac.jp, †kise@cs.osakafu-u.ac.jp

Abstract There is an approach of annotation extraction from printed documents in which a document image of an annotated document and its original document image are compared. In one of the previous methods, the original image is processed in the same manner as the annotated image to deal with image degradations by printing and scanning. However, the previous method has a serious problem in convenience which is that we must use the same printer and scanner to equalize the image degradations. In this report, we propose a method in which both the original and the annotated images are processed with some image processings. The proposed method consists of three steps of processing which are for detherring, for color change, and for local displacement. We also propose an objective evaluation of extracted annotations to compare the experiment results accurately. Experimental results have shown that the extraction accuracy has improved by 12.6% at the recall rate and 12.3% at the precision rate.

Key words Document image, Image degradation, Annotation extraction, Objective evaluation

1. はじめに

私たちは日常的に、訂正、メモ書き等、文書への書き込みをする。ここで、書き込まれる情報はユーザにとってなにかしらの重要な情報である。従って、文書から書き込みを取り出し処理をすることで、文書に付与された情報が得られると考えられる。例えば、同じ文書に複数の人による書き込みが行われた時、書き込みを抽出することで、多くの書き込みがなされている箇所を特定することができる。そして、その箇所は文書の中で重

要視されていると考えられる。また、あるユーザの様々な書き込みを抽出することで、ユーザの書き込みの傾向を得ることができる。

文書からの書き込み抽出手法の一つに、書き込みを行う前の文書画像(元画像)と、書き込みを行った後の文書画像(書き込み画像)の差分を見る手法がある。しかし、元画像と書き込み画像に書き込み以外の差異があると、多くのノイズが発生する可能性がある。これらのノイズの原因は、印刷とスキャンによる画像劣化や位置のずれなどである。したがって、書き込みの抽

出をするためには、ノイズの原因を除去する必要がある。

このような問題を回避するため、例えば中居らの手法 [1] では、画像劣化に対しては、元画像を印刷やスキャンしたものを利用することで対処し、位置のずれには、位置合わせの処理を行っている。しかし、印刷機やスキャナによって、発生するノイズが異なるため、元画像と書き込み画像の作成に同じ印刷機やスキャナが必要となる制限がある。また、単一のパラメータによる画像全体の位置合わせを行うため、微小な位置のずれが残る問題点がある。

そこで、本研究では、劣化のない元画像と劣化した書き込み画像に処理を加えることによって、画像劣化と位置のずれによる問題を解決する手法を提案する。画像劣化に対する処理では、劣化の種類ごとに処理を加え、元画像と書き込み画像の差異を減少させる。これは、印刷やスキャンによる画像劣化には、複数の種類の劣化が混ざっているためである。位置のずれに対する処理は、局所的な位置のずれを補正する処理を提案する。これは、画像全体の位置合わせにおける、微小な位置のずれが残る問題点に対処するものである。これらの画像処理により、抽出精度を向上させる。

また、結果の評価が主観的に行われると、他の手法と正確に比較することができない。そこで、正解となる画像（正解画像）を用意し、抽出結果（抽出画像）と比較する客観的な評価システムを構築し、結果を評価する。

2. 関連研究

文書からの書き込み抽出の手法は大きく 2 種類に分類できる。1 つは元画像を必要とせず抽出を行う手法 [2]~[7] である。もう 1 つは、書き込み画像と元画像との差分を取って、書き込みの抽出を行う手法 [1] である。

元画像を必要としない手法として、手法 [2], [3] では、書き込みに用いることのできるペンの色を定め、画素の色によって書き込みを判別している。手法 [4], [5] では、連結成分を手書き文字と印刷された文字に分類し、書き込みを判別している。手法 [6], [7] では、枠で囲まれた領域に注目した書き込みの抽出を行っている。これらの手法は、書き込みの色や文字の形状、および特定の位置条件を利用して抽出を行うため、精度のよい抽出ができる。しかし、書き込みをする際に、色や文字の形、および抽出される領域が制限される問題点がある。一方で、元画像を必要とする手法は、抽出する書き込みに制限を設けない利点がある。しかし、電子媒体で文書を所持していなければならない問題点がある。また、書き込み画像に発生する様々な画像の変化に対応しなければ、書き込みの抽出を行うことができない問題点もある。

ここで、後者の元画像を必要とする手法に注目する。画像の変化についての研究には、位置のずれに関するものと、濃淡の変化に関するものがある。位置のずれへの処理として [1] では、相似変換に不変な特徴点の対応付けを用いる位置合わせを行っている。しかし、微小な位置のずれが残る問題点がある。また、[8] では、印刷とスキャンによる濃淡の変化に対する処理が行われている。これは、元画像の画素値が印刷およびスキャン

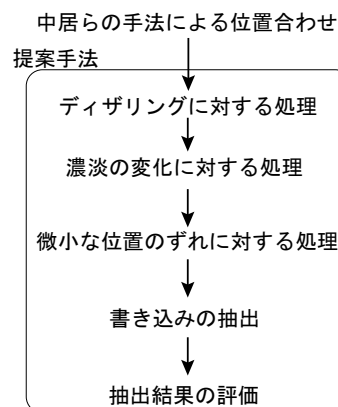


図 1 処理の流れ

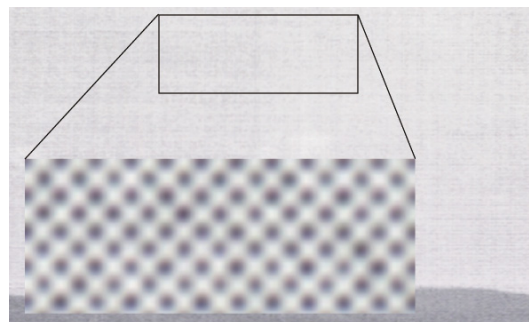


図 2 ディザ

でどのように変化しているかを調べ、そこから濃淡の変化モデルを作り、元画像の濃淡を変化させることで、劣化の再現をしており、有用な手法である。

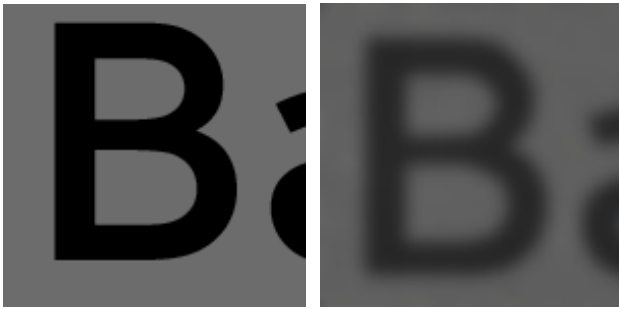
3. 提案手法

提案手法では、グレースケールの画像を対象とし、画像処理により画像劣化や位置のずれによる元画像と書き込み画像の差異をなくすことで、書き込み抽出の精度を向上させる。処理の流れを図 1 に示す。まず、提案手法の処理の前に中居らの手法の位置合わせを用い、大きな位置のずれに対する位置合わせをする。次に、ディザリングによる画素値の変化、濃淡の変化、位置のずれの 3 点に対する処理をする。そして、画素値の差分を用いて書き込みを判定し、書き込みの抽出、結果の評価をする。結果の評価として、正解画像と抽出結果を画素単位で比較し、評価をする手法を提案する。

3.1 ディザリングに対する処理

ディザリングとは、印刷の際に限られた色の組み合わせで中間階色を表現するための処理である。この処理によって、中間階色は図 2 のような限られた色のドットパターンによって表現される。この状態では、画素単位では元画像の中間階色とは異なっており、画素単位の比較では元画像と書き込み画像の差分が正しく取れない。そのため、ディザに対する処理が必要となる。

そこで、ディザリングに対する処理として、ガウシアンフィルタを用いて書き込み画像を平滑化する。平滑化により、ディザリングによるドットパターンが滑らかになり、元画像の中間



(a) 元画像 (b) 平滑化された画像

図 3 平滑化による問題点

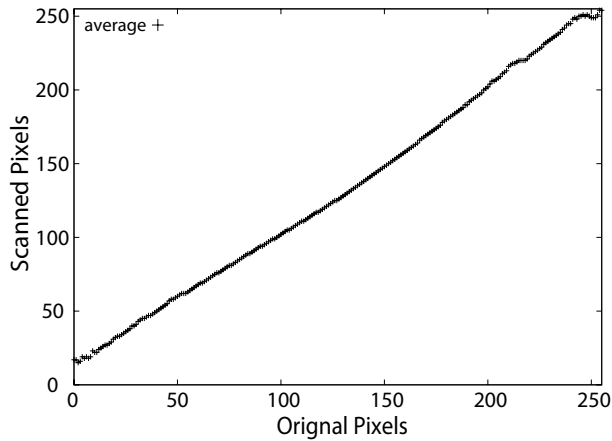


図 4 濃淡の変化モデル

階色に近づけることができる。しかし、平滑化を画像全体に施すと、図 3(a) と図 3(b) のようにディザが発生している領域以外にも変化が起り、元画像と書き込み画像に新たな差異が発生する。そこで、平滑化によって発生する差異を打ち消すように、元画像に対しても同程度の平滑化を施す。

平滑化のパラメータは、実験的に定める。書き込み画像のガウシアンカーネルの直径 g_a については、数枚の書き込み画像を平滑化し、その中のディザが目立たなくなるような値に設定する。そして、元画像のガウシアンカーネルの直径 g_o については、平滑化された書き込み画像と同程度に見えるような値に設定する。

3.2 濃淡の変化に対する処理

濃淡の変化は、印刷の濃さ、スキャナの感度に影響される。また、ディザを平滑化しても元画像と完全に一致する画素値には復元されない。

書き込み抽出は、画素値の差分によって行う。しかし、濃淡の変化によって差分が 0 になることはないと考えられるために、抽出をする際には、書き込みと判別するための閾値を設けることになる。元画像に対する濃淡の変化が大きくなると、ノイズを削減するために閾値を大きくすることになる。しかし、背景に似た色で書き込みを行ったり、鉛筆などで色の薄い書き込みを行うと、誤って書き込みが削除される可能性も大きくなるため、閾値を低くする必要がある。そこで、元画像に対する書き込み画像の濃淡の変化を解析し、元画像に変化を与えることで、

Meeting 資料		Meeting 資料		発行 期日
2007年5月31日(水曜日)				
1 今回までの予定				
<ul style="list-style-type: none"> 正解プログラムの作成 				
2 今回までに行ったこと				
<ul style="list-style-type: none"> 正解プログラムの作成 <ul style="list-style-type: none"> コーンテスト 正解画像の書き込み画像と見比べてみられる 抽出結果と正解画像を比べるとずれが生じる ローレンツ曲線(階調) スキャナで抽出された画像に階調を平滑化して書き込み画像を生成する 画像抽出プログラムの作成 <ul style="list-style-type: none"> 画像抽出に想定していた書き込みと異なる可能性がある 30ピクセル以上の濃淡では書き込みはなし、元画像をコピーして合成ではずれあり 濃度、黒・赤・緑・青の階調については合成で見ればはずれなし 濃淡がある領域のみ抽出して書き込み画像を生成する 				
3 次回までの予定				
<ul style="list-style-type: none"> 画像抽出プログラムの改良 <ul style="list-style-type: none"> 抽出してしまっている部分を修正する 抽出結果のずれについて考える 濃度抽出は抽出画像にのみ書き込みし直したものを残す ローレンツ曲線と階調を比較して調子 抽出プログラムの書き込み 				

図 5 領域分割

閾値を低く保つ。

この濃淡の変化に対応するために、[8] を参考とした濃淡の変化モデルを作成する。濃淡の変化モデルとは、元画像におけるそれぞれの画素値が、書き込み画像ではどのように変化しているかを見るものである。濃淡の変化モデルの例を図 4 に示す。濃淡の変化モデルは、対となる元画像と書き込み画像で画素値の比較を行い、画素値ごとの色の変化の平均値を求めることで作成する。

なお、濃淡の変化に対する処理を行う段階では、中居らの手法により位置の補正がされているが、微細なずれが残っている。このような微細なずれのある状態では、印刷やスキャンによる濃淡の変化ではなく、本来は別の画素である位置を読み込み、濃淡の変化として扱ってしまうことになる。これは、エッジの出るような領域では大きな変化となる可能性が考えられる。そこで、閾値 c を設定し、元画像と書き込み画像の画素値の差が c より大きいときは、位置のずれが影響していると判断し、濃淡の変化モデルを作成する際にその画素を除外する。除外する画素の閾値 c は、元画像とスキャン画像において、対応している画素を比較し、どの程度画素値が変化するかを調べ、その上限を基準に設定する。

3.3 微小な位置のずれに対する処理

書き込み画像は、印刷やスキャンを通して画像化されるため、位置のずれが発生する。元画像に対し、書き込み画像がずれていると差分による抽出が困難となる。そこで、中居らの手法では、相似変換に不変な特徴量を用いる位置合わせをしている。しかし、画像全体を単一の相似変換パラメータで変換しているため、微小な位置のずれが残っており、抽出精度が低下する一因となっている。そこで、微小な位置のずれを補正することで、さらなる抽出精度の向上を図る。

位置合わせは、元画像と書き込み画像の位置のずれを打ち消すように画素を平行移動させることで行われる。しかし、紙面

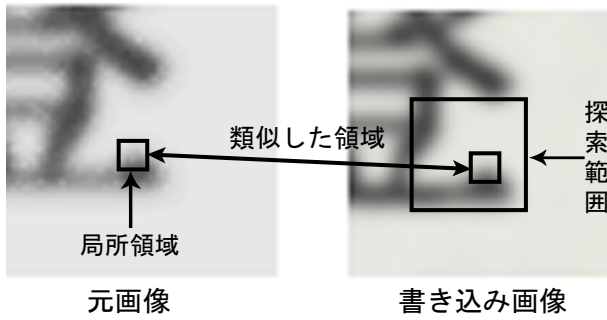


図 6 局所領域を利用した位置合わせ

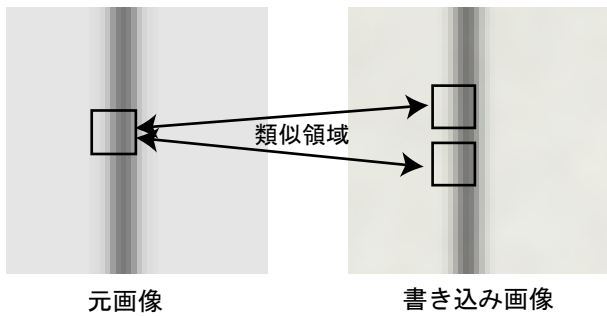


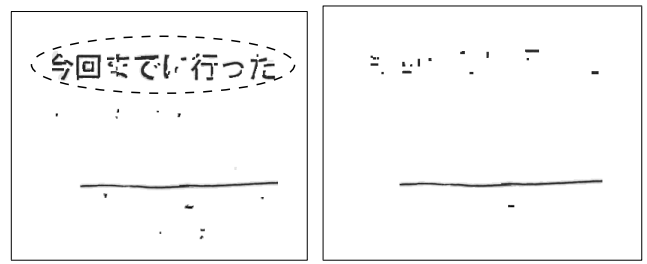
図 7 誤認識領域の例

の歪みなどにより、画像全体で位置のずれが一定でないため、画像全体で1つの位置合わせのパラメータを決めることはできない。そこで、図5のように、画像をピクセル単位で $l \times l$ となる複数の領域に分割し、その領域ごとに位置合わせのパラメータを求める。ここで、 l を大きくしすぎると、分割が粗くなるため細かい歪みの違いに対応することができない。そこで、領域を分割する大きさ l は、いくつかのパラメータで実験を行い、もっともよい値とする。

位置合わせのパラメータの算出には、局所領域の画素値を利用する。図6に示すように、元画像において、局所領域の画素値を取得する。そして、書き込み画像において探索範囲を設定し、画素値の配置パターンが類似した領域を探索する。類似した局所領域を発見したら、対応した領域の位置関係から位置のずれの大きさを求める。その後、図5のように分割した領域ごとに、領域内の画素から求めた位置合わせのパラメータの中央値を、その領域の位置合わせのパラメータとする。

ここで、全ての画素に対して、対応点の探索を行うと、膨大な処理時間となる。また、余白部分や、塗りつぶし部分では、画素値の配置パターンが似ている領域が続くため、間違った対応をとる元となる。そこで、画像の中でエッジとなって出る画素に対してのみ、対応点の探索をする。これは、エッジとなる画素とは画素値の変化が大きい領域周辺にあり、特徴を持った画素値の配列パターンが得られると考えられるためである。

ただし、エッジに注目するだけでは、図7のような領域で位置合わせパラメータを求めると、正しい位置合わせのパラメータを得ることができない。これは、同じパターンの画素値配列が続き、元画像と書き込み画像で正確に対応点を取ることができないためである。そこで、隣接する領域の位置合わせパラメータを利用することで、この問題に対処する。図5のように



(a) 修正前 (b) 修正後

図 8 ノイズの発生と対処

分割した領域において、位置合わせのパラメータは、隣接した領域で大きく変化しないと考えられる。したがって、周囲の領域に対して位置合わせのパラメータが大きく違う領域は、正確に位置合わせのパラメータが算出できていない可能性がある。そこで、このような領域の位置合わせのパラメータは、隣接する上下左右の領域の中央値とする。この処理によって、間違っていた位置合わせのパラメータを排除する。

3.4 書き込みの抽出

本研究では、元画像と書き込み画像の画素値の差分を見ることが書き込みを抽出する。抽出の際には、閾値を設け、閾値より差分が大きい場合を書き込み画素とみなし、書き込み画像の画素値を結果画像に出力する。この閾値は、元画像と補正された書き込み画像の対応する画素を比較することで、定めることができる。そのため、どの程度の画素値の差が残っているかを調査し、設定する。

抽出の際の問題として、位置合わせ処理を加えても依然として微小なずれが残っている場合があることが挙げられる。そのような場合、元画像と書き込み画像において、同位置の画素値の差分だけで判断をすると、図8(a)の破線枠内のようなノイズとなる画素が発生する。そこで、微小なずれに対応し、ノイズを減少させるために、探索する範囲の大きさ s を設定する。書き込み画像で注目している画素に対して、元画像では同座標の画素を中心とした周囲の $s \times s$ ピクセルを探索する。そして、探索した中で最も近い画素値の差分が閾値以上であれば、書き込み画素であると判断する。その結果、図8(b)のようなノイズの削減された抽出結果が得られる。

このような処理をすると、書き込みの近くに印刷された文字が存在する部分で、書き込み画素と印刷された文字の画素が対応し、書き込みが消える場合がある。この問題は、画素値の比較により書き込み画素の判断を行っているため、印刷された文字と書き込みの区別ができない場合に発生する。そこで、書き込み画素の周囲には書き込み画素が存在すると考え、書き込みと判断された画素を中心とした抽出する範囲の大きさ e を設ける。そして、書き込みと判断された画素を中心とした周囲 $e \times e$ ピクセルを書き込み画像から抽出する。この処理により、消えた書き込みが復元でき、抽出精度が向上すると考える。

3.5 抽出結果の客観的評価手法

結果の適切な評価をするため、客観的評価手法を提案する。

本研究では、正解画像と抽出画像を比較することで、抽出結果に対して客観的な評価を与える。そこで、書き込みのみを正確に取り出した正解画像が必要となる。しかし、書き込まれた文書から正確に書き込みのみを分離し、正解画像を作成することは容易ではない。そこで、書き込みのみの紙面を用意し、スキャナで取り込み正解画像とする。そして、書き込み対象の文書のスキャナで取り込んだ画像(スキャン画像)と合成することで、書き込み画像を得る。

本研究では、画素単位での評価を行う。これは、正解画像と抽出画像で対応する座標の画素値を比較し、一致する画素の割合を評価値とするものである。一方、評価手法としては連結成分単位のものも考えられる。これは、つながった書き込みを一つの連結成分とし、連結成分ごとに抽出の成功と失敗を判定し、抽出に成功した連結成分の割合を評価値とするものである。しかし、連結成分を用いた評価では、連結成分中の一部の画素が抽出されないときに、どのような評価を与えるかという問題がある。例えば、連結成分の抽出が成功したかを、抽出できた画素の割合を用いて評価することを考える。このとき、抽出できた割合が同じでも、書き込みの線が細くなるかのように全体から少しずつ書き込みが失われるときと、書き込みの一部が消えるかのように局所的に失われるときでは、主観的には評価が違ふ可能性も考えられる。このように、連結成分単位の評価では機械的な判断が難しい事例があることから、問題を単純化するため、画素単位で抽出できたかを評価する。

3.5.1 正解画像の作成

正解画像を作成する際の課題として、印刷された文書へ直接書き込むときの差異を可能な限りなくすることが挙げられる。具体的には、鉛筆とボールペンなど筆記具による差、書き込む際に起きるかすれなどを残せる形で正解画像を作成する必要がある。タブレットを用いれば、書き込みのみを分離することは容易であるが、筆記具による差を表現することが難しくなる。また、白紙に書くだけでは文書内容に合わせた書き込みをすることができないため、不自然な書き込みになる。これらの問題を解決するため本研究では、普通紙ではなく半透明であるトレーシングペーパーを用いる。これにより、書き込み対象の文書に対応した正解画像が作成できる。

正解画像の作成は以下のようなものである。まず、元画像を印刷した文書にトレーシングペーパーを重ね、書き込みを行う。その後、トレーシングペーパーをスキャナで読み取ることで、正解画像を作成する。

3.5.2 書き込み画像の作成

この処理ではスキャン画像と正解画像を合成する。正解画像の各画素について背景部分である画素かを判断し、背景部分でないと判断された画素についてその画素値を元画像の対応する画素へ上書きする。この処理において、問題点として次の2点が考えられる。

1つ目は、文書においては背景部分に見える画素でもスキャナで正解画像として読み込んだとき、紙の汚れやスキャナの汚れなどで、画素値が目視では判別できない程度変化している可能性があることである。この問題点は、判断するために閾値

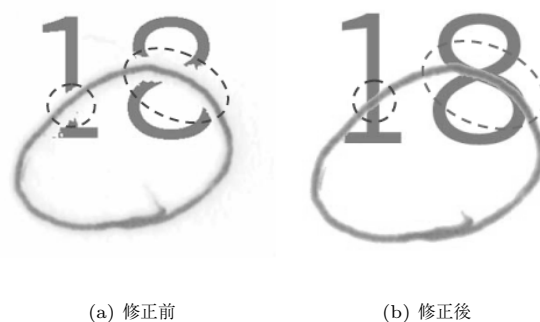


図9 画素の削れ

を設け、背景とみなす範囲を広げることで解決する。具体的には、画素値のRGBの平均が閾値より大きいときは、その画素は「背景部分である」と判定する。閾値は実際にスキャナで取り込んだ画像の画素を見て、ボールペン等で書き込んだ際のにじみ等書き込みの一部が消えないような値に設定する。

2つ目は、スキャン画像の印刷部分と正解画像の書き込み部分が重なったときに、どのような画素値を出力するかである。図9(a)は、この点を考慮せず、正解画像をスキャン画像に上書きをした結果の一部を拡大したものである。破線部に注目すると、スキャン画像が正解画像によって不自然に削られてしまっている。これは、正解画像での薄い色も、すべて上書きをしているためである。そこで、両画像の画素値に重み付けをして和を取ることで解決する。修正を加えた結果が図9(b)である。この時、書き込み画像の画素値は計算によって求められているので、正解画像とは一致しない。この状態で抽出を行い、正解画像と比較すると不一致と判断される。従って、正解画像側にも計算した画素値を反映させる。

3.5.3 抽出結果の客観的評価

この処理では正解画像と抽出画像を比較することで、抽出画像に対して客観的な評価を与える。評価値として再現率 R と適合率 P を用いる。 A を抽出画像中で正解画像と一致した画素数、 B を抽出画像の画素数、 C を正解画像の画素数とすると、 R と P は

$$R = \frac{A}{C}, \quad P = \frac{A}{B},$$

で定義される。ここで、2つの式中の画素数は、背景部分と判断される画素を除いたものである。再現率が高いとき、書き込みの画素が十分に抽出されていることを示す。適合率が高いとき、ノイズが少ないことを示す。

画素の比較方法としては、対応画素のRGBすべてにおいて値が完全一致するかを用いる。本研究では、すべての画像が同じサイズであり、画像の変化による色の変化が起きないため、完全一致で評価できる。なお、全画素でなく背景部分でないと判定した画素のみに注目するのは、文書は背景によって多くが占められているため、背景に再現率と適合率の値が支配され、結果として意味の薄い値しか得られないからである。

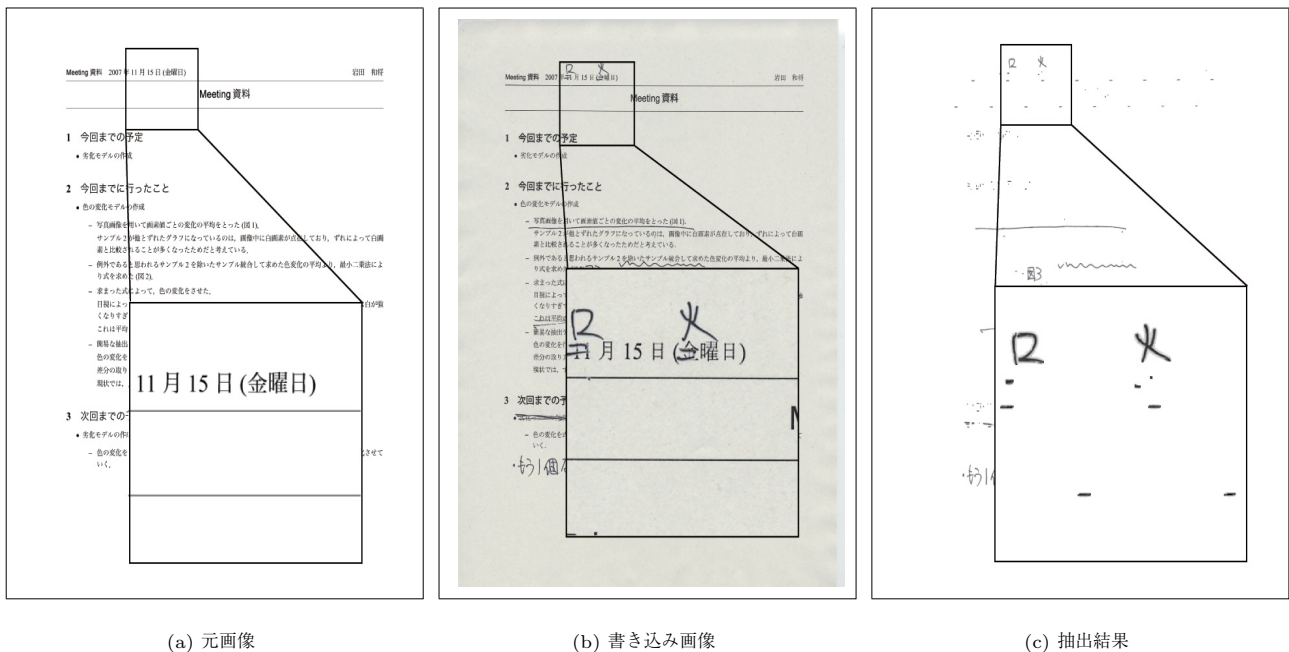


図 10 使用した画像例

表 1 全ての画像における実験結果

	再現率	適合率
従来手法	68.3%	73.3%
提案手法	80.9%	85.6%

表 2 位置合わせに成功した画像における実験結果

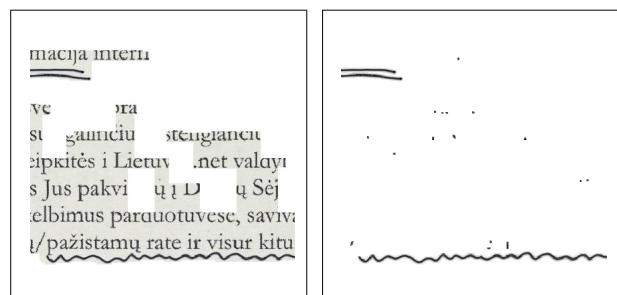
	再現率	適合率
従来手法	67.9%	79.2%
提案手法	81.0%	91.7%

4. 実験結果

本実験では、従来手法である中居らの手法と提案手法の抽出結果を比較した。文書として、100枚のPDF文書を用い、それぞれに対応する正解画像を作成した。元画像はImage Magickに付属しているconvertコマンドを用いて、PDFファイルから200dpiで変換した。書き込み画像として用いる画像は、A4用紙に印刷した後、スキャナにて200dpiで取り込んだ。その後、正解画像と合成することで、書き込み画像を作成した。図10に元画像、書き込み画像、抽出結果の例を示す。

実験結果は、様々なパラメータによって、再現率と適合率の関係が変化する。本実験では、いくつかあるパラメータの中で、3.4において設定する探索範囲と抽出範囲の2つを変化させた。ここで、表1に提案手法におけるF値が最も高い実験結果と従来手法の実験結果の比較を示す。表1において、従来手法に対して提案手法は、再現率において12.6%、適合率において12.3%の結果向上が得られた。ここで、抽出精度が向上した例を図11に示す。このように、従来手法において発生していたノイズを除外することができ、適合率が向上した。

なお、表1には、図12のように、中居らの手法の位置合



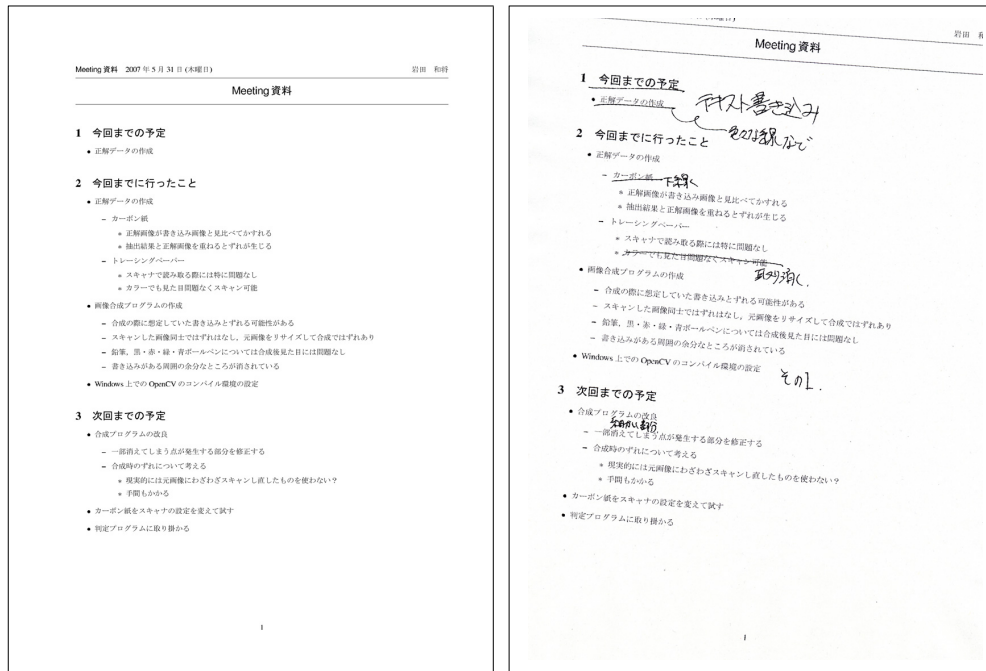
(a) 従来手法 (b) 提案手法

図 11 抽出精度の向上した例

せの段階で位置合わせに失敗している結果も含まれている。提案手法は、ある程度の位置合わせが行われていることを前提とした処理であるため、位置合わせの成否が結果に大きく影響している。そこで、表2に、位置合わせに成功した画像に対する結果を示す。位置合わせに成功した画像は、100枚中90枚の画像であった。これらの画像において、適合率は91.7%となり、提案手法が有用であることが分かる。

また、パラメータを変化させたときの、再現率と適合率の推移について調べる。ここでパラメータを、ディザリングと濃淡の変化、そして位置のずれという画像の劣化に対処する処理で設定するパラメータと、抽出時に設定するパラメータの2種類に分けて考える。

画像の劣化に対処する処理で設定するパラメータ g_a , g_o , c , l は、元画像と書き込み画像を近づけることが目的である。例えば、ディザリングに対処する際のガウシアンカーネルの大きさ g_a は、ディザのドットパターンが滑らかになる値に設定すればよく、ディザが残るような平滑化や極端に強い平滑化をする



(a) 元画像

(b) 書き込み画像位置合わせ後

図 12 位置合わせの失敗例

必要はない。従って、画像劣化の補正処理で設定するパラメータ g_a, g_o, c, l は、元画像と書き込み画像がもっとも近くなる時が最適と考えられる。

抽出時に設定するパラメータには、一致する画素を探索する範囲を決める s と抽出する範囲を決める e がある。探索範囲のパラメータ s を大きくすると、適合率が上昇し、再現率が低下する。また、抽出範囲のパラメータ e を大きくすると、再現率が上昇し、適合率が低下する。このように、抽出時に設定する2つのパラメータ s, e は、再現率と適合率を変化させる役割を持っている。

実験では、画像劣化の補正処理のパラメータは元画像と書き込み画像を近づけるのが目的であるため、目視でもっとも近い画像が得られるパラメータに設定し、固定とした。ここでは、 $g_a = 7, g_o = 7, c = 50, l = 50$ とした。そして、変化させるパラメータを抽出時に設定する2つとし、再現率と適合率の推移を調べた。

実験結果を図 13 に示す。図 13 より、パラメータによっては適合率と再現率どちらかを、90%程度の精度にできることが分かる。図 14 は、抽出時に探索範囲のパラメータを大きくしたために、書き込みが消えた例である。これは、一致する画素の探索範囲を広げたために、書き込み画素と印刷された文字の画素の誤対応が増加するためである。このように、印刷された文字の近傍の書き込みが消えるため、再現率が低下する原因となる。

また、一般に書き込み抽出に求められる再現率と適合率は用途ごとに異なる。書き込み抽出の目的が、書き込み領域の特定であれば、書き込みのかすれなど部分的な消失は問題ないと考えられる。しかし、ノイズが多くなると、書き込みでない領域

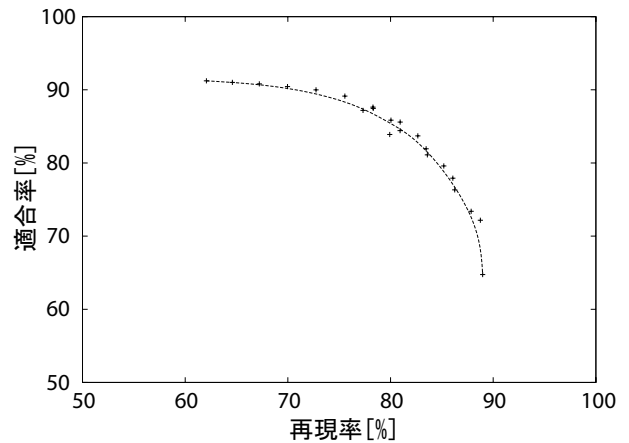
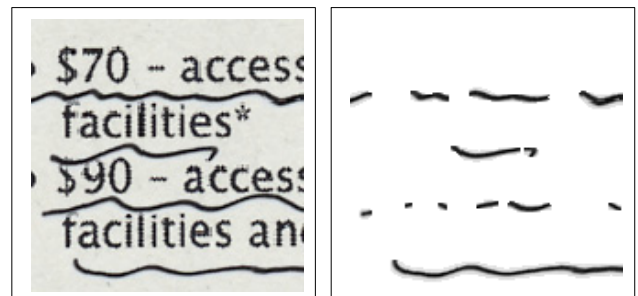


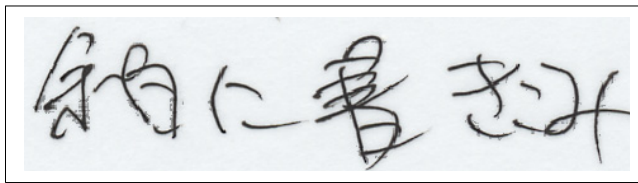
図 13 再現率と適合率



(a) 書き込み画像

(b) 抽出結果

図 14 書き込みが消えた例



(a) 正解画像



(b) 抽出結果

図 15 文字の一部が削れた例

を、書き込み領域と誤認識する可能性が高くなる。そのため、適合率を優先した結果であれば、十分な書き込み抽出ができていと考えられる。

一方、書き込み抽出後に文字認識をすることを考えると、図 15 のように文字の一部が消えると文字認識が困難となる。そのため、文字認識のためには、再現率と適合率がともに高い値となる必要があると考えられる。これを実現するためには、次の 2 つの処理が考えられる。

一つは、再現率を優先するパラメータで抽出を行い、抽出後に適合率を向上させることである。適合率を向上させるためには、ノイズを削除する必要がある。例えば、連結成分が小さなものを削除することを考える。これは、ノイズの多くは細かな点として存在するためである。このような処理を行ったとすると、書き込みであるけれども、細かな点で構成されている文字が誤って削除される可能性が考えられる。

もう一つは、適合率を優先したパラメータで抽出を行い、抽出後に再現率を向上させることである。再現率を向上させるには、抽出に失敗した書き込み部分を復元する必要がある。書き込み部分の復元は、抽出に失敗した周りの書き込みを利用することが考えられる。図 15(b) のように、抽出に失敗した領域の周囲には、抽出に成功した書き込みが存在する。この抽出に成功した書き込みと書き込み画像を用いて、その周りの失った書き込みが復元できると考えられる。そして、ノイズは書き込み部分と比較して固まった領域に多くでないため、書き込みでない部分の復元は発生は少ないと考えられる。

従って、再現率と適合率がともに高い値を得るためには、適合率を優先するパラメータで抽出を行い、抽出後に [7] のような書き込みを復元する処理を加えることで、再現率を向上させる手法が有望であると考えられる。

5. ま と め

本稿では、文書からの書き込み抽出の精度向上と客観的評価

を提案した。書き込み抽出の精度向上は、印刷およびスキャンによる画像の劣化を、画像処理によって補正することで実現させる。また、客観的評価は、正解画像を作成し、抽出結果と画素単位での比較をするものである。

実験の結果、従来手法である中居らの手法よりも、再現率において 12.6%、適合率において 12.3%の精度向上が得られた。

今後の課題として、再現率のさらなる向上が挙げられる。これには、書き込みを抽出した後の処理として、抽出できなかった書き込みの復元をする必要があると考えられる。

謝辞

本研究の一部は、科学研究費補助金(基盤研究(B) 19300062)、(特別研究員奨励費 19-7621)の補助による。

文 献

- [1] 中居友弘, 黄瀬浩一, 岩村雅一, “特徴点の局所的配置に基づく位置あわせを用いた文書からの書き込み抽出法,” 信学技報, PRMU2006-245(2007).
- [2] D. Möri and H. Bunke, “Automatic Interpretation and Execution of Manual Corrections on Text Documents,” In Handbook of Character Recognition and Document Image Analysis, ed. H. Bunke and P. S. P. Wang, pp.679–702, World Scientific, Singapore(1997).
- [3] J. Stevens, A. Gee, and C. Dance, “Automatic Processing of Document Annotations,” In Proc. 1998 British Machine Vision Conf. , Vol. 2, pp.438–448(1998).
- [4] J. K. Guo and M. Y. Ma, “Separating Handwritten Material from Machine Printed Text using Hidden Markov Models,” In Proc. 6th International Conf. on Document Analysis and Recognition, pp.436–443(2001).
- [5] Y. Zheng, H. Li, and D. Doermann, “The Segmentation and Identification of Handwriting in Noisy Document Images,” In Lecture Notes in Computer Science (5th International Workshop DAS2002), vol.2423, pp.95–105(2002).
- [6] S. L. Taylor , and R. Fritzson, “Registration and Region Extraction of Data from Forms,” In Pattern Recognition, 1992. Vol.I. Conf. A: Computer Vision and Applications, Proceedings., 11th IAPR International Conf. , pp173–176(1992).
- [7] Bin Yu, and Anil K. Jain, “A Generic System for Form Dropout,” In Pattern Analysis and Machine Intelligence, Vol.18, No.11, Nov. 1996, pp1127–1134(1996).
- [8] Ching-Yung, Shih-Fu Chang, “Distortion Modeling and Invariant Extraction for Digital Image Print-and-Scan Process, ” International Symposium on Multimedia Information Processing (ISMIP99), Taiwan(1999).