

氏 名	後 藤 英 賢
授 与 学 位	博 士 (工 学)
学位授与年月日	平成7年3月24日
学位授与の根拠法規	学位規則第4条第1項
研究科, 専攻の名称	東北大学大学院工学研究科 (博士課程) 情報工学専攻
学 位 論 文 題 目	文書画像理解システムに関する研究
指 導 教 官	東北大学教授 阿曾 弘具
論 文 審 査 委 員	東北大学教授 阿曾 弘具 東北大学教授 丸岡 章 東北大学教授 白鳥 則郎

論 文 内 容 要 旨

第1章 序 論

近年の情報処理機器の普及に伴い、紙の上に書かれた文書を計算機に自動的に読みとらせ、その内容を計算機上で再利用することが望まれている。文書の入力を自動化するためには、文字や図形、写真といった個々の要素の認識の他に、これらの要素が文書上にどのように配置されているか、すなわち文書の幾何構造を自動的に抽出し、さらに要素間の関係を表す論理構造を抽出する処理が必要となる。

文書の幾何構造と論理構造の抽出の処理を合わせたものは文書画像解析と呼ばれており、従来、主に様々な文書のフォーマット(構造)にも対応できるという汎用性と、処理速度の向上を目指した手法が提案されている。しかし、従来手法の多くは、例えば論文や新聞のように文章領域が矩形として切り出せるような比較的単純な構造をもつ文書を処理の対象としており、汎用性の追求は十分とはいえない。また、実際の文書画像には大きなシミや汚れ、ゆがみ、文字行の接触や湾曲などが含まれることがあるが、これらの外乱への積極的な対策がとられておらず、外乱を含む文書に対しては処理が不安定になりやすいという問題があった。

本研究では、外乱を含む文書も安定に処理できるという性質を処理の頑健性と呼んで重視するとともに、従来の手法では困難とされた複雑な構造(レイアウト)をもつ文書も取り扱えるように、高い汎用性と頑健性を有する文書画像解析処理を実現することを目的とする。そして、文書画像解析の中でも特に重要な文書の幾何構造の抽出の部分に焦点をあて、その中で必要とされる文字行と

罫線の抽出手法について検討，提案し，実験によりその有効性を示す。

第2章 文書画像解析処理の構造方法

本章では，従来の文書画像解析手法の問題点を明らかにするとともに，高い汎用性と頑健性を有する処理を実現するためには文書画像解析処理をどのように構成すべきかについて議論し，その枠組みを与えている。文書画像解析処理は，処理の内容に応じて，1. 画像処理，2. 幾何構造の抽出，3. 論理構造の抽出の3つの段階に分けることができる。汎用性や頑健性の向上を阻害する原因は主に2の幾何構造の抽出にあるので，本研究ではこの部分を研究の対象としている。

文書は，文字（文章）や図形，写真，罫線（フィールドセパレータ）といった要素から成り立っている。文字領域に注目するとその基本となる要素は個々の文字であり，文字が並ぶことにより文字行が，文字行が並ぶことにより文章ブロックが構成されるというように，要素が階層的な構造をなしている。文字行の局所的な形状や文字の概形は矩形とみなすことができるが，形状に関する自由度が小さいことから，これを低レベルな構造と呼ぶことにする。一方，文章ブロックや写真，図形などの形状は自由度が大きく，これを高レベルな構造と呼ぶ。

従来のほとんどの手法では，「文書中の各領域は直線によって分割できる」とか「各領域の概形は矩形である」といった高レベルな構造に関する知識（高レベルな知識）が用いられている。しかし，システムに蓄えられている知識と実際の文書構造との矛盾はしばしば高レベル側で発生し，また，文字行の湾曲，傾きなどの外乱は高レベル側の構造の乱れに顕著に現れる。従って，文書画像解析処理の汎用性と頑健性を向上させるためには，高レベルな知識に依存しない要素（文字行，罫線）の抽出処理が必要である。

文字行と罫線はいずれも直線状の要素とみなせる。文字領域の抽出において，高レベルな知識を用いずに文章ブロックを抽出したり，個々の文字を抽出することは困難である。また，個々の文字行の全体を矩形とみなす処理では，文字行の湾曲や入り組みに対処できない。人間の読書という行為においても主に局所的な情報が利用されていることを考えると，文字行や罫線の局所的な形状に注目すれば，多少の湾曲も吸収できる頑健な処理が実現できると考えられる。第3章の文字行抽出処理と第5章の罫線抽出処理では，いずれも局所的な情報を重視したアルゴリズムを用いる。

第3章 頑健性と高速性を重視した文字行抽出手法

本章では，処理の汎用性はもちろん，特に頑健性も重視して，高レベルな知識に依存しない文字行抽出処理について検討し，複雑な構造の文書からも頑健・高速に文字行を抽出できる新しいアルゴリズム「区分直線連結法」を提案している。

従来，画像の局所的な情報を利用する文書画像解析の手法として，文書画像から得られる外接矩形を局所的な処理で結合して文章ブロックに組み上げる手法などが提案されている。しかし，これらの手法では矩形間や画素間の距離を基準とした単純な結合条件を用いており，異なる文章ブロックが狭い間隔で隣接している場合には文字行の大部分の抽出に失敗するという欠点があった。また，文書画像解析においてしばしば用いられる外接矩形は，黒画素の連結成分の抽出を伴うために生成

に時間を要し、罫線に接触した文字が図形要素の方に含まれてしまうという問題があった。本研究では処理の頑健性と処理速度の向上を目的とした新しい中間データ表現「基本矩形」を導入し、提案する文字行抽出処理ではこれを利用した。

区分直線連結法では、処理対象は横書きと縦書きの文字行が主体の文書であるが、文字行の傾きや接触はある程度許容する。区分直線連結法のアルゴリズムの概要は次のとおりである。

文書画像はまずバンドと呼ばれる幅一定の帯状の領域に分割される。横方向の文字行を抽出するためには横方向の分割を行なう。次にこのバンドの中で基本矩形の連結成分を作り、同時に基本矩形の中心点の射影分布のピークを解析して連結成分の直線性（文字行らしさ）を評価する。続いて隣合うバンドの間で連結成分を結合し、文字行候補を生成する。最後に、文字行間の結合を強制的に分離する処理—ブリッジ分割処理—を適用し、結合した複数の文字行を個々の文字行として分離・抽出できるようにする。ブリッジ分割処理は、文字行候補中の背の高い基本矩形を強制的に分割することで実現される。

未知文書 61件を用いた評価実験によって、従来手法では抽出が困難とされた局所的に傾いた（ $\pm 7^\circ$ 程度）文字行や湾曲して入り組んだ文字行、接触した文字行などが安定に抽出できることが確認され、提案手法が高い汎用性と頑健性を有することが示された。また処理速度については、本手法が従来手法より高速であり、ディザ画像を含んだ入力に対しても著しい速度の低下がない（安全性が高い）ことが示された。

第 4 章 Hough 変換の効率的なパラメータ空間の設計法

ハフ (Hough) 変換は画像中の様々な図形を検出するための一手法であり、効率的で頑健な手法としてコンピュータビジョンの分野ではしばしば利用されている。しかし、パラメータ空間の選び方やサンプリング間隔については設計上曖昧な部分が残されていた。本章では直線検出用のハフ変換をとりあげ、画像の量子化による誤差とハフ変換による誤差を理論的に解析して、要求される精度を保って効率的な変換を行うための設計手法を示した。その結果、正方格子・正方形の画像に対しては、直線のモデル式を $y=ax+b$ とする変換や和田らの提案した $\gamma-\omega$ ハフ変換の方が $\rho-\theta$ 表現よりも効率的 (64%) であること、パラメータのサンプリング間隔の上限値が S_{valbe} の $a-d$ パラメータによるハフ変換のそれと一致することなどが明らかになり、従来の高精度化・効率化手法が誤差解析という観点から統一的に理解できることが示された。

第 5 章 罫線・飾り線の自動抽出手法

従来、点線や破線を抽出するための専用の処理が提案されているが、文書中の罫線やフィールドセパレータの形状は自由度が大きいため、任意の形状にも対応できる汎用的な処理が望まれていた。本章では、罫線の規則性に注目し、ハフ変換と自己相関関数を用いて文書画像中の任意の形状の罫線とフィールドセパレータを抽出できる汎用的な罫線抽出処理について検討した。様々な罫線のサンプルと文書画像を用いた実験の結果、本手法では点線や破線はもちろん、従来は抽出が困難だっ

た複雑な形状の罫線・フィールドセパレータも抽出でき、文字と罫線の分離も良いことが示された。

第6章 結 論

本章では本論文を総括し、本研究の成果と今後の課題を述べた。

本研究により、文書画像解析に必要な文字行と罫線の抽出処理が実現された。既存の写真領域の抽出手法と合わせて、文書構成要素から文書の高レベルな幾何構造や論理構造を抽出することが今後の課題である。

審査結果の要旨

膨大な量の文書を計算機に取り込みデータベース化し、文書情報を再利用するためには、文字や図形・写真が混在する文書画像を計算機に入力し認識・理解することが不可欠である。しかし、フォーマットを限定しない文書画像の処理手法は未だ確立されておらず、その開発が望まれていた。著者は、認識・理解の前処理として必須の、文書画像の幾何構造の抽出法について研究し、文字行および罫線図形を抽出する、汎用性と頑健性を兼ね備えたアルゴリズムを開発するとともに、実験によりその有効性を示した。本論文はその成果をとりまとめたもので、全編6章からなる。

第1章は序論である。

第2章では、文書認識・理解システム全体の中で幾何構造抽出処理の位置づけを明確にするとともに、どんな文書画像でも処理対象とできる汎用な処理手法が画像要素の配置などに関する知識を前提としない局所的処理に基づくべきであることを述べている。

第3章では、画像データの処理単位として基本矩形を新しく考案し、文書の構成要素である文字行を抽出する区分直線連結法を提案するとともに、これが汎用性、頑健性、高速性および安全性を有する手法であることを実験により明らかにしている。この手法は、基本矩形の局所的な横方向の並び（区分直線）を疑似文字行として抽出し、それらを連結して文字行を抽出するものである。従来手法ではできなかった、湾曲している文字行や入り組んだ構造の中の文字行が抽出できることを実証している。これは、実用上重要な成果である。

第4章では、画像から直線成分を抽出する手法であるハフ変換について、その誤差解析を行い、要求される精度を保って効率的に直線抽出を行う手法を理論的に明らかにしている。これは、従来の個々の直線モデル式に固有の高精度化・効率化手法が統一的に理解できることを導く新しい知見である。

第5章では、文書画像中の任意形状の罫線図形やフィールドセパレータを抽出する手法を提案し、実験によりその罫線図形抽出能力や文字と罫線図形の分離能力を評価し、その有効性を明らかにしている。この手法では4章で求めたハフ変換の最適化手法を活用し、高速性を実現している。この罫線図形抽出法は個々の画像に関する知識を必要とせず、汎用性をもち、実用上有用な手法である。

第6章は結論である。

以上要するに本論文は、文書画像理解システムの基礎となる幾何構造の理解に焦点をあて、ハフ変換の統一的な誤差解析をもとに効率的な直線抽出法を与えるとともに、効率的で汎用な文字行抽出法と罫線図形抽出法を考案し、その有効性を実証したもので、情報工学の発展に寄与するところが少なくない。

よって、本論文は博士（工学）の学位論文として合格と認める。