論文の内容の要旨

論文題目　　文書画像中の手書き文字と活字文字の判別

氏名　　小山　純平

本論文では，文書画像中の手書き文字と活字文字の判別を取り扱う．
序論では，文書解析の一分野である手書き文字と活字文字の判別の重要性と，それが現在抱えている問題点，そしてわれわれの目的について述べる．文書解析とは，スキャナなどで取り込まれた文書画像のレイアウトや論理構造を計算機が扱える形で表現する技術である．電子データの普及によって，紙文書を使用する機会は減少した．しかし紙文書はその使用方法の柔軟性から，いまだに必要不可欠なメディアである．文書解析は，これらのメディアの橋渡しとなる．われわれは文書解析の一つである，手書き文字と活字文字の判別技術に注目する．文書画像中の手書き文字と活字文字を判別する技術は様々な利点を生む．代表的な例として，現在人手で行っている光学文字認識（optical character recognition: OCR）システム選択の自動化や，文書画像中の属性値の抽出などが挙げられる．
従来の手書き文字と活字文字の判別手法は，文字または文字列の形状を利用した手法が主体であった．これらの手法は，前処理として文字や文字列の切り出しが共通して必要である．しかし，筆記体のような繋がり文字や，重なり線分が存在すると，現在の技術でも正確に文字や文字列を切り出すことは難しい．そのため，フォーマットが整っている帳票などの文字や文字列を切り出すことはできるが，フォーマットが定まっていない自由文書ではその切り出し精度が大きく落ちる．このような文字，文字列切り出しの不正確さは，その後の処理の精度を悪化させる原因となる．
そこでわれわれは，人間の視覚に注目した．人間の手書き文字と活字文字の判別は，瞬間的で，ノイズに強く，文字種に関係なく，さらに文字や文字列の切り出し処理に依存しない．このような人間の視覚をモデル化し文字種の判別を行った研究が存在する．人間の初期視覚が網膜に映った映像の空間周波数と方向を知覚できることから，それらは文書画像をテクスチャとみなし，そのテクスチャを解析することで文字種の判別を行った．しかし，判別のために文書サイズのデータが必要であるため，それらの手法では局所的な手書き文字と活字文字の判別を行うことは難しい．
そこで本論文では，文字や文字列の切り出し処理を必要とせず，さらに局所的な手書き文字と活字文字の判別が可能な手法を提案することを目標とする．

第二章では，われわれの提案するスペクトルに基づく局所ゆらぎ検出法の基本的アイデ

ィアについて述べる．
われわれの提案手法は人間の視覚システムを参考にしている．人間の視覚システムは，一次視覚野において，物体の発光や反射光が網膜に形成する空間周波数とその方向を，局所領域ごとに検知する．この局所領域の様々な空間周波数とその方向は，人間の視覚システムが利用する基礎的な情報である．加えて，人間のテクスチャ解析は，その視覚システムの初期的な部位で主に行われていることが知られている．そこで提案手法では，人間の視覚と同様に，局所領域ごとに空間周波数とその方向を検出する．そのためにわれわれは二次元高速フーリエ変換を用い，文書中の局所領域のパワースペクトルを得る．ウィーナー・ヒンチンの定理から，パワースペクトルは文書画像中の局所領域のテクスチャを良く表現する．また，テクスチャは自己相関性の統計的な性質であるため，テクスチャ解析は局所領域に含まれる文字線分を，その絶対的な位置ではなく，線分の相対的な位置関係によって評価することと同等である．つまり，パワースペクトルを解析することで，文字線分をその絶対的な位置によらず評価することが可能である．これにより，文字や文字列の切り出しを必要としない，手書き文字と活字文字の判別手法が実現される．
さらに，われわれは手書き文字と活字文字の判別のために，手書きに起因する文字線分のゆらぎに注目する．手で字を書く以上，人間はこのゆらぎを避けることができない．われわれはこの手書き文字に含まれるゆらぎを定量化し，それを基に手書き文字と活字文字を判別する．この手法は文字線分という，文字よりも小さい単位に含まれるゆらぎを評価するため，欠損がある文字や文字線分のみでも手書きと活字を判別できる．

第三章では，水平・鉛直方向に注目したスペクトルに基づく局所ゆらぎ検出法を論ずる．活字の文字線分に注目すると，それは完全にまっすぐである．そのため，その文字線分はパワースペクトルの二次元周波数座標上に，完全にまっすぐな，大きいパワーを持つ点列を構成する．便宜上われわれはその点列を基軸と呼ぶ．それに対して手書き文字は，手書きに起因するゆらぎを含むため，完全な直線ではありえない．そのため，そのパワースペクトルに表れる基軸は傾きやぼやけを含む．このように，ゆらぎによって，手書き文字から得られるパワースペクトルは，活字文字にはない性質を多く有する．この傾きやぼやけを定量化し，それを基にわれわれは手書き文字と活字文字の判別を行う．まず、われわれは水平・鉛直方向の線分に注目した．活字文字の水平・鉛直方向の線分は，パワースペクトルにも完全に水平・鉛直方向を向いた基軸を誘起する．手書きの文字線分はゆらぎを含むため，誘起される基軸は水平・鉛直からの傾きや基軸自体のぼけを含む．そのため，われわれはパワースペクトルの二次元周波数座標上で水平・鉛直方向から外れたパワーを手書きのゆらぎに誘起されたものと捉え，その定量化を行った．
文字画像サンプルと文書画像サンプルを用意し，提案手法の有効性を検証する実験を行った．われわれは文字種として漢字，かな文字，英数字を考え，それぞれに複数の手書

き文字と活字文字を用意したデータベースを作成し，手書き文字と活字文字の判別実験を行った．得られた結果は，提案手法が手書き文字と活字文字の判別のために有効な特徴量を提供することを示した．特に水平・鉛直方向の文字線分を多く含む漢字において提案手法は有効であった．次に，われわれは企業活動で使用されている文書画像と，その文書の手書き文字を全て活字文字に置き換えた比較データを用意した．それらの文書に提案手法を適用し，比較実験を行った．そして，われわれの提案手法が文字や文字列の切り出し処理を必要とせずに，文書中の手書き文字と活字文字の判別のために有効な特徴量を提供することを示した．

第四章では，方向の制限を取り除いた，スペクトルに基づく局所ゆらぎ検出法を論ずる．水平・鉛直方向に注目したスペクトルに基づいた局所ゆらぎ検出法は，主に水平・鉛直方向の線分を持つ文字の手書き文字と活字文字の判別に非常に有効であった．しかし斜線や曲線を持つ文字ではその有効性が大きく下がってしまう．そこでわれわれは，水平・鉛直に限らず，より多くの方向の文字線分を観察し，それらの手書きに起因するゆらぎを定量化する手法を提案した．提案手法は学習ステップと判別ステップに分けられる．学習ステップでは，手書き文字と活字文字の特徴を学習する．文字画像をパワースペクトルに変換し，そこで文字線分の方向を表す特徴量を定量化する．得られた特徴量を多層パーセプトロンで学習し，手書き文字と活字文字の判別器を作製する．活字文字はいくつかの限られたパターンの文字線分の太さや方向を持つ．逆に手書き文字はゆらぎを含むため，同じパターンであってもその太さや方向にばらつきが含まれる．多層パーセプトロンは，活字文字の線分のパターンと，手書き文字の線分のゆらぎを学習する．判別ステップは，文書画像の局所領域ごとに手書きであるか活字であるかの判別を行う．まず文書画像を局所領域ごとにパワースペクトルに変換する．そこから，学習時と同様の手法で抽出した特徴量を判別器に入力し，局所領域が手書きであるか活字であるかを判別する．

まずわれわれは文字画像の学習・判別実験を行った．われわれは産業技術総合研究所の ETL 文字データベースを用いた．そのデータベースは英数字，ひらがな，カタカナ，漢字の手書き文字と活字文字を持つ．そこから学習データセットと判別データセットを作成し，学習・判別実験を行った．得られた判別率は 97%であり，これは十分実用に耐えうる数値である．次に，提案手法が文字や文字列の切り出し処理を必要としないことを示すために，われわれは ETL 文字データベースの文字をランダムに配置した文書中の手書き文字と活字文字を判別する実験を行った．そして，文字や文字列を切り出すことなく，文書中の 94%の手書き文字と活字文字の領域を判別することができた．最後にわれわれは実際に企業活動などで使用されている文書中の手書き文字と活字文字の判別実験を行い，提案手法が実際の文書に適用できることを示した．

第五章では，手書き・活字文字領域自動判別光学文字認識装置について述べる．われわれは，提案手法の実際の使用形態を意識し，文書画像中の手書き文字と活字文字を判別し，その結果に基づいて光学文字認識装置による認識実験を行った．そしてその認識率を検討した．

第六章で本論文の纏めを行う．本論文では，文書中の手書き文字と活字文字の判別を論じた．われわれはスペクトルに基づく局所ゆらぎ検出法を提案し，文字や文字列の切り出しを必要としない手書き文字と活字文字の判別を実現した．本手法は人間の視覚モデルに着想を得ている．人間の視覚モデルが解き明かされていくにつれ，われわれのモデルはより洗練されていき，手書き文字と活字文字の判別だけではなく，より幅広い分野で貢献できるようになるだろう．