

# 論文の内容の要旨

論文題目「コンピュータ筆跡鑑定に関する研究」

学位申請者 常盤 公德

キーワード：科学捜査，筆跡鑑定，筆者異同識別，コンピュータ，日本語文字

科学捜査分野では、遺言書、脅迫文、委任状、契約書、カード署名など、主に紙面に残された文字を対象に鑑定人とよばれる人によって筆跡鑑定が行われている。その鑑定内容は、①他人の文字をまねた模倣や筆跡個性を隠した<sup>とうかい</sup>韜晦などの記載方法の識別と、②同一人が筆記したものか否かを識別する筆者異同識別の2つに分けられる。本研究は、筆者異同識別に関するものである。

現在なされている鑑定人による筆者異同識別では、字画構成、字画形態、運筆状態について、ルーペなどを用いた目視検査や既存の筆跡標本に基づく出現頻度の検査などが実施されている。しかし、それらはいくまで鑑定人の学識と経験による鑑定であり、その客観性についてはときに裁判などで問題となることがある。さらに、裁判員裁判が2009年に開始され、裁判員に対して説明し易い客観的・定量的鑑定の必要性が増しており、この問題はコンピュータによる筆者異同識別により解決が可能であると考えられる。

コンピュータによる筆者異同識別は、バイオメトリックス認証としての筆者照合や筆者認識の研究としてなされているが、次のような問題点がある。第1に、電子ペンなどを用いて書字行動そのものを扱うオンライン型筆者照合に研究が偏っており、その研究成果の多くは、オフライン型すなわち書字行動の結果として得られる筆跡を扱う科学捜査分野の筆者異同識別での利用が困難である。第2に、科学捜査分野への適用性と得られる異同識別精度が両立していない。たとえば、対象筆跡の書体が楷書体で統一されていればある程度の高い識別精度が保たれるが、行書体などの文字が混ざると識別精度は大きく低下する。また、記載状態についても筆記媒体（紙など）が無地で統一されていれば識別精度は保たれるが、模様や重ね書きがあったり、下線や枠線などがあると対応できなくなる。筆記具についても一種類で統一されていれば識別精度が保たれるが、鉛筆、ボールペン、サインペンなどが混ざると識別精度は低下する。さらに、識別法については識別過程が不明確な場合が多々あり、たとえば、ニューラルネットワーク識別法のように学習次第で識別精度の向上は可能であるが、なぜそのような識別結果が得られたかについて説明しづらいという問題がある。

そこで本研究では、科学捜査分野における筆者異同識別を様々な条件の鑑定対象筆跡に安定的に実用レベルの識別精度で実施でき、かつ、筆者異同識別の過程と根拠を定量的に説明できるコンピュータによる筆跡鑑定法を開発することを目的とした。

第1章では、筆跡鑑定の現状、コンピュータによる筆者異同識別の必要性と問題点、そして、本研究の目的について論じた。

第2章では、科学捜査分野の筆跡鑑定で最もよく使用される漢字を対象筆跡として、筆者異同識別法の基本的な枠組みを構築した。その骨子は、次の4点からなる。第1に様々な状態にある鑑定対象筆跡から安定的に抽出可能な字画の始筆部・転折部・終筆部からなる特徴点を人が採取し、二つの筆跡間の幾何学的マッチング残差を字形の違いを表す特徴量とした。第2に、筆跡そのものの位置・大きさ・回転を安定的・高精度に一致させる新しい正規化法を考案した。第3に、従来の研究ではその発想すらなかった書きむらすなわち個人内変動の抑制処理を考案した。第4に、識別過程が明確で説明し易く、統計的に合理的な結果が安定的に得られる代表的な統計的識別法の一つである最大尤度法を採用した。科学捜査分野において要求される識別精度は、最高レベルで99.9%が1つの目安となっている。構築した筆者異同識別法は4文字を使用したときこの最高レベルの要求識別精度99.9%を達成することを被験者320人の識別実験により示した。

第3章では、前章において実用レベルの識別精度を達成するには個人内変動の抑制が必須であることが判明したことから、個人内変動の抑制に最適な幾何変換モデルについて検討した。検討した幾何変換は、2次等角変換、アフィン変換、疑似アフィン変換、射影変換、2次多項式、3次多項式の6種類である。個人内変動は射影歪みでその多くを説明でき、それゆえ射影変換により個人内変動を大幅に低減できること、また同時に筆跡個性すなわち個人間変動を小さくするような悪影響を与えないことを明らかにした。結果として、検討した6種類の内、射影変換が個人内変動を抑制する最適な幾何変換であることを示した。

第4章では、構築した筆者異同識別法によって得られる異同識別精度の信頼性について論じた。前章までは全標本法とよばれる母集団推定と識別に同じサンプルを用いる方法で精度推定を行った。この方法は識別精度が高めに推定され易いことが知られていることから、全標本法の他に交差検証法、4重交差検証法、一点除外法を用いて識別精度を推定した。いずれの精度検証法においても、4文字を使用したときに99.9%という実用レベルの識別精度が得られることを確認した。また、99.9%という識別精度は統計的信頼度95%で充分有意であることを確認した。さらに、ROCカーブとよばれる照合精度特性を求め、その解析から最大尤度法が決定する識別境界は、統計的に平均識別誤差が最小となる境界となるはずであるが、実際にはそうになっていない、言い換えると、平均識別精度の向上の余地があることがわかった。

第5章では、前章において識別精度向上の余地があることが明確となったことから、さらなる異同識別精度向上の可能性を探った。前章の異同識別精度の信頼性を吟味する過程において、平均識別誤差が最小となる識別境界が実際には選定されていないことがわかった。本章では、まずその原因が、母集団が正規分布をしているという最大尤度法の仮定が実際には満足されていないことにあることを示した。次に、母集団分布に対する仮定を必要としないノンパラメトリックな識別法の代表例であるk近傍法を適用し、得られる識別精度について検討した。識別に2文字以上、特に2, 3文字を使用した場合、k近傍法の方が予想通り高い識別精度が得られ、3文字使用で99.9%を超えることを確認した。k近傍法の適用には、十分な数の筆跡サンプルの収集と最適なkの探索が必要となるが、これらが行えた場合、k近傍法によって識別精度のさらなる向上が見込めることを示した。

第6章では、科学捜査分野の筆跡鑑定で扱われているひらがな・カタカナ・数字を含む日本語文字全般に適用できるように提案する筆者異同識別法を拡張した。さらに、字種間での識別特性について論じた。従来、ひらがな・カタカナ・数字は字画構成が漢字に比べて単純なため、鑑定に使用しづらいと言われていたが、いずれの文字種でも3~4文字を識別に使用すれば99.9%を超える実用的な識別精度が得られることを示した。また、従来、定性的にいわれていた書きむらや筆跡個性を個人内変動分布と個人間変動分布として定量的に表すことを可能とし、その解析から①日本語文字は、漢字とひらがなからなるグループと、カタカナと数字からなるグループの2つのグループに分けられる、②ひらがなは漢字と同程度の識別精度が得られるばかりでなく、カタカナと数字は漢字よりもむしろより高い識別精度が得られる、など従来の人による鑑定で経験的にいわれている定説とはかなり異なる新しい知見を得た。

第7章は結論である。本研究の内容を総括し結びとした。

本研究で提案したコンピュータによる筆者異同識別法は、科学捜査分野で要求される適用性と高い異同識別性能を日本語文字全般に対して合わせ持つばかりではなく、従来の鑑定人の学識・経験によるアプローチとは全く異なる識別過程が合理的で説明し易い客観的・定量的鑑定手段を与えるものである。その意味で本研究は、科学捜査分野における筆跡鑑定のあり方を変革し、大きく貢献するものと期待される。