

低レート静止画圧縮時の文字認識性能品質評価

田中 陽平[†]杉浦 彰彦[‡]静岡大学大学院情報学研究科[†]静岡大学大学院情報学研究科[‡]

1. はじめに

近年、静止画像での画像認識技術への関心が高まっている。静止画像の符号化方式では、人間の視覚特性にマッチングした方式が採用されており、データ量が多い画像での画像認識精度は高い。しかし、携帯電話などの高性能端末の登場により、少ないデータ量で高い画像認識精度を得られることが求められるようになった。少ないデータ量では、人間の視覚特性に優れた符号化方式が画像認識に効果的だとは限らない。そこで本研究では、人間の視覚特性に優れた符号化方式での低レート圧縮時に文字認識性能に与える影響を、符号化方式別と認識方法別で比較検証する。

2. 原理

本研究では、画像の符号化方式として JPEG と JPEG2000 を用いる。また、画像認識のアルゴリズムとしてフーリエ記述子を用いた認識プログラムとテンプレートマッチングを用いた認識プログラムを用いる。

2.1 JPEG

JPEG とは、ISO と ITU-TS の合同組織が定めたデジタル静止画像の圧縮符号化方式で、Joint Photographic Experts Group of CCITT-ISO/IEC の略称である。特徴としては、圧縮率が非常に高く、圧縮率と復号される画像の品質をユーザが調整でき、プラットフォームやアプリケーションに依存せず、国際標準であり世界中で用いられているということが挙げられる。

また JPEG は、自然画像には原色があまり現れないことに着目し、元の画像から、もともと少ない情報である特異な色差情報を削っても、人間の目にはその差はわかりにくいという人間の視覚特性に優れたものである。この削る程度をユーザが調整しているため、画像

の品質と圧縮率を調整できる。

JPEG には四つの処理方式があるが、本研究では最も基本的な方式である基本 DCT 方式を用いる。

2.2 JPEG2000

ISO と ITU-TS の合同組織が定めたデジタル静止画像の圧縮符号化方式で、JPEG の後継にあたる。特徴としては、高圧縮時の画質が JPEG より顕著に良いことや、圧縮率をユーザが調整できること、可逆符号化方式であることなどが挙げられる。

JPEG2000 の処理方式は 10 を越える種類があるが、本研究では最も基本的な Part1 を用いる。

2.3 フーリエ記述子を用いた文字認識

二次元の平面上の図形でも、境界線上の座標や偏角をある点からの位置の関数とみなすと一次元波形となる。この一次元波形をフーリエ級数展開し、有限個のフーリエ係数を用いて線図形の特徴を表現することができる。このフーリエ係数をフーリエ記述子といい、本研究ではこのフーリエ記述子を用いた文字認識アルゴリズムを用いて認識率の算出を行う。あらかじめ登録した教師データとの相関係数を算出し、最も高いものを認識結果とする。

2.4 テンプレートマッチングを用いた文字認識

ある特定のパターンをあらかじめ登録しておき、入力画像中にそれと同じパターンが存在するかどうか調べ、その位置を特定する操作をテンプレートマッチングと呼ぶ。これを応用し、本研究では、テンプレート画像を作成し、そのテンプレート画像と入力画像がどの程度類似しているかを求め、その値によって認識結果を算出するプログラムを用いる。

テンプレート画像は、ある画素に対し、全ての入力画像の濃度値の平均をとり、これを全画素に対して行ない作成する。テンプレート画像と入力画像の類似度を最小二乗法を用いて算出し、最も小さい数字のものを認識結果とする。

Quality assessment of character recognition in low rate image compression.

[†]TANAKA Yohei, Shizuoka University Graduate School of Informatics

[‡]SUGIURA Akihiko, Shizuoka University Graduate School of Informatics

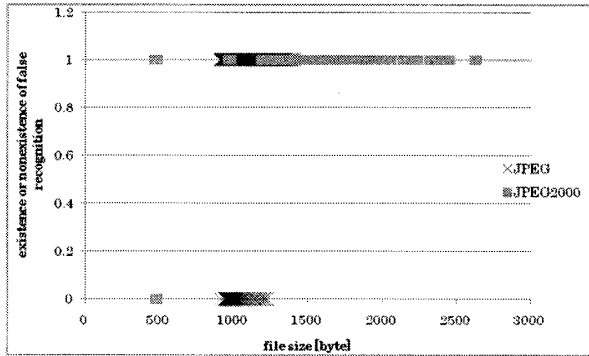


図 1: Result of Character Recognition using Fourier Descriptor

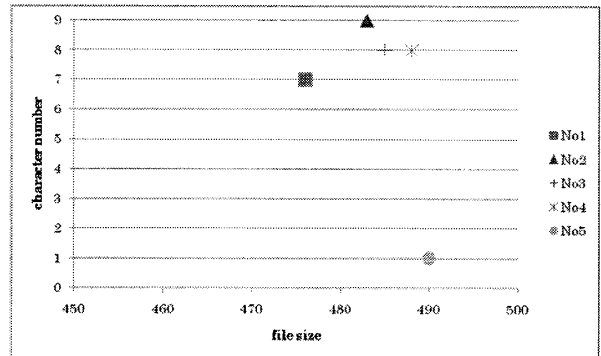


図 3: Classification of False Recognition image using JPEG2000

3. 実験

実験で用いる入力画像は、男女 10 人に 0 から 9 までの 10 文字を書いてもらい、これをスキャンしグレースケール化して作成する。画像のサイズは縦 128 画素×横 128 画素とする。画像符号化方式として、JPEG と JPEG2000 を用いる。入力画像の低レート圧縮は、量子化テーブルの変更等を行わずに通常の圧縮を行う。100 段階の圧縮率から、下位 20 段階を用い、1 つの画像についてそれぞれで圧縮し、合計 20 枚の画像を得る。画像の合計枚数は 2000 枚となる。

3.1 フーリエ記述子による文字認識

まず、低レート圧縮時のフーリエ記述子による文字認識性能への影響を調べる。前述の画像に対し、フーリエ記述子を用いた認識プログラムにて認識率の測定を行った。

3.2 テンプレートマッチングによる文字認識

次に低レート圧縮時の、テンプレートマッチングによる文字認識性能への影響を調べる。実験 3.1 と同じ画像に対し、テンプレートマッチングを用いた認識プログラムにて認識率の測定を行った。

4. 結果

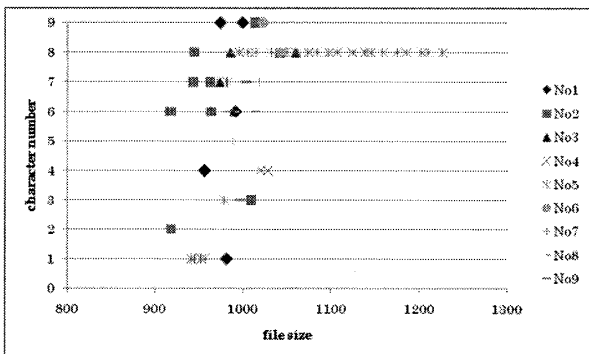


図 2: Classification of False Recognition image using JPEG

実験 3.1 の結果のグラフを図 1 に示す。横軸が画像のファイルサイズ、縦軸が認識結果での誤認識の有無を表し、誤認識がある場合は 0、ない場合は 1 である。JPEG では全体的に誤認識が起きることが多く、誤認識の有無の境界を見分けづらいたことが分かる。JPEG2000 では全体的に誤認識が起きておらず、500byte 近辺以外のファイルサイズでは誤認識がないことが分かる。

また、実験 3.1 の結果、誤認識が起きた JPEG 画像を、筆者と文字の種類でグラフに分布したものを図 2 に示す。文字では 8、筆者では No2 と No5 の人物の文字に対して誤認識をすることが多い可能性があることが分かる。

実験 3.1 の結果、誤認識が起きた JPEG2000 画像を、筆者と文字の種類でグラフに分布したものを図 3 に示す。誤認識が起きた数が少なく、筆者・文字共にばらばらであることが分かる。

5. まとめ

フーリエ記述子による文字認識の実験において、二つの符号化方式の間に文字認識性能の差を確認した。JPEG2000、つまり人間の視覚特性に優れた符号化方式の方が認識結果も高い値であり、誤認識を生じる境界も明確であった。今後は、符号化方式別で生じる認識性能の差の原因を調査し、様々な認識アルゴリズムで実験を行うことで、より正確な影響を調査したい。また、文字の筆者の数や認識を行う画像の圧縮率のパターンを増やし、低レートの幅を広げ、文字認識性能に与える影響を調査したい。

Reference

- [1] 酒井幸市：画像処理とパターン認識入門，森北出版株式会社，2006 年
- [2] アズウィ，橋本晋之介：JPEG 概念から C++での実装まで，ソストバンクパブリッシング株式会社，1998 年