

ウェーブレット変換による画像の特徴抽出と文字領域

金子 俊介[†] 久保田 光一[‡]
中央大学大学院 理工学研究科 情報工学専攻

要約: 本論文ではウェーブレット変換によるエッジ抽出処理を利用して、情景画像中の文字領域を推定する方法を提案する。提案手法では、ウェーブレット変換によって得た高周波成分の輝度値分布から画像中の文字領域を推定する。評価方法には予め用意した正解画像データと推定結果後の画像データを比較して、正解率を求めた。画像によっては推定結果の良い画像を得ることができ、本手法の有効性を示した。

キーワード: ウェーブレット変換, 文字領域

1 背景

情景画像に含まれる文字を解析する前段階として、情景画像中の文字領域を特定して背景と文字領域を分離する研究が行われている。色空間についてクラスタリングすることで文字領域と背景を分離する研究 [1] や、文字周辺の空間周波数の高さから文字領域を特定する研究 [2] などが行われている。特に空間周波数を用いた研究において、ウェーブレット変換を利用する手法 [3] があるが、対象とする画像に限られていたり、画像中の様々なノイズにより領域特定が困難になるといった問題がある。

2 目的

本論文ではウェーブレット変換を用いて画像から文字領域を推定する手法を提案する。提案手法では、情景画像を小領域に分割する。それぞれの領域において、ウェーブレット変換によって得られるエッジ抽出画像の輝度値の分布から、文字領域の条件に当てはまる領域を文字領域と推定する。推定手法の実験においては、小領域のサイズやフィルタによる前処理、輝度値の分布を分ける閾値の変更など、様々なパラメータの組み合わせを試行して推定精度を評価する。実験結果から文字領域推定の精度を向上するパラメータを見つけ出す。

3 ウェーブレット変換

ウェーブレット変換は周波数解析の手法の一つとして、近年ますますその応用分野が広がっている。ウェーブレット変換は入力波形データをマザーウェーブレットと呼ばれる基本参照波で解析する。マザーウェーブレットを拡大あるいは縮小し、移動させながら入力波形を特徴的な波形ごとに分解していく。波形は連続的なデータであるが、コンピュータ上では離散値を取る。ここから、入力データに画像の輝度値を離散データとして与えることで、画像についてもウェーブレット変換が可能になる。

離散ウェーブレット変換

ウェーブレット変換の基本となるものは連続ウェーブレット変換で、関数 $h(x)$ に対する変換は次式で定義される。これ以降の変換式は [4] より引用した。

$$(W_{\psi}h)(b, a) = \frac{1}{\sqrt{a}} \cdot \int_{-\infty}^{\infty} \overline{\psi\left(\frac{x-b}{a}\right)} \cdot h(x) dx$$

ここで $\psi(x)$ はマザーウェーブレットを表す。 a, b はマザーウェーブレットのパラメータで、 a は $\psi(x)$ を拡大縮小させ、 b は平行移動させる。離散ウェーブレット変換は a, b をそれぞれ、 $a = 2^{-j}, b = 2^{-j}k$ として離散化したものである。離散ウェーブレット変換の式を以下に示す。

$$d_k^{(j)} = 2^{\frac{j}{2}} \cdot \int_{-\infty}^{\infty} \overline{\psi(2^j x - k)} \cdot h(x) dx$$

多重解像度解析 (MRA)

離散ウェーブレット変換を二次元へと拡張することで画像に対してウェーブレット変換を行える。変換により画像は周波数成分ごとに分解される。このとき、特定の周波数帯域で画像を複数の小さい画像データに分割することを多重解像度解析 (MRA) という。MRA により画像を高周波成分ごとに分割した画像を出力できる。図 1 は簡単な画像に対し、MRA を適用したものである。

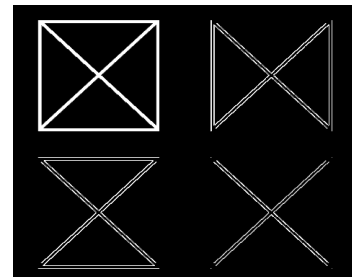


図 1 多重解像度解析

図 1 の左上は元画像を平均化した画像である。また、右上は垂直方向の、左下は水平方向、右下は対角方向の高周波成分を取り出した画像である。各方向においてエッジが抽出されることが分かる。MRA を用いて情景画像からエッジを抽出した場合、文字以外のエッジも多く抽出される。エッジの特徴から文字領域を判別する。

4 提案手法

はじめに、ウェーブレット変換した画像を、一辺 n ピクセルの正方形領域に分割する。この正方形領域ごとに文字領域か否かを推定する。以下に示す条件に当てはまる領域を文字領域、条件に一致しない領域は非文字領域と推定する。

1. 輝度値 t 以上のピクセルをカウントする。カウント数を p とする。
2. p について、 $u \cdot n^2 < p < v \cdot n^2$ となるような p を持つ領域を文字領域とする (n^2 は領域の総ピクセル数)。ただし、 $0 < u < v < 1$ とする。

u, v は、領域における高輝度ピクセルの疎密を調整するパラメータである。高周波成分が顕著に現れる領域は輝度値の高いピクセルが密集した領域である。高輝度のピクセルが密集した領域は、元画像においてエッジが多い領域であり、文字以外のノイズが多いとして文字領域から外す必要がある。また、高輝度ピクセルが少ない領域は、元画像においてエッジの少ない平坦な領域であり、文字領域ではない可能性が高いため外した。また、 n の値は 12、 t の値は 127 で固定とした [5][6]。

Feature Extraction and Character Areas with Wavelet Transform

[†] Shunsuke KANEKO, Information and System Engineering Course, Graduate School of Science and Engineering, CHUO University

[‡] Koichi KUBOTA, Information and System Engineering Course, Graduate School of Science and Engineering, CHUO University

5 実験

推定の対象となる画像は、正解データのある情景画像 9 枚と、正解データのない画像 251 枚を用意した*。実験における処理の流れを以下に示す。

1. カラー画像入力
2. 前処理
3. ウェーブレット変換
4. 文字領域推定
5. 領域抽出

前処理では、ノイズ除去やエッジ強調のために画像にフィルタをかける。フィルタには Gaussian フィルタ, Sobel フィルタ, 鮮鋭化フィルタをそれぞれ用いて文字領域推定の精度を比較する。ウェーブレット変換ではマザーウェーブレットに Daubechies, Coiflet, Symlet, O-spline を用いて、それぞれ精度を調べる。

評価

正解データを用意した情景画像 9 枚について、推定結果の精度を評価する。推定精度は、推定結果を正解データと比較したとき、正解データと一致している割合で評価する。

まず、推定処理により文字領域とした領域数と、非文字領域と推定した領域数を求め、正解データと比較して更に 4 つの領域に分割する。推定処理で文字領域とした領域数の内、実際に文字領域であるものを A 個とおく。反対に文字領域でないものは B 個とおく。次に、推定処理で非文字領域とした領域数の内、実際には文字領域であるものを C 個とおく。また、実際に文字領域でないものは D 個とおく。表 1 は分類をまとめたものである。

表 1 分類表

	文字領域推定	文字領域	非文字領域
正解データ			
文字領域		A	C
非文字領域		B	D

精度評価の割合は、分類した 4 つの領域数を用いて、以下の 2 つの式を乗算した値 (評価値) とする。

$$\varphi_1 = \frac{A}{A+C}, \quad \varphi_2 = \frac{A+D}{A+B+C+D}$$

φ_1 は、文字領域として推定した領域が正解データの文字領域と等しい割合である。 φ_2 は、すべての領域において文字領域と非文字領域を正しく推定している割合を示す。文字領域推定が正しければ、 φ_1, φ_2 とともに値は 1 へと近づく。しかし、文字領域推定が間違っていればその値は 0 に近い値をとり、推定率が低くなる。実験結果を表 2 に示す。

表 2 より、もっとも評価値の高い画像番号 1 について、元画像を図 2 に、文字領域抽出画像を図 3 に示す。元画像における道路標識の文字領域を抽出していることが分かる。

6 まとめと今後の課題

ウェーブレット変換を用いた文字領域推定方法を提案した。実験によって、特定の画像においては高い推定率を求められることを確認した。前処理による評価値の変動について、エッジ

表 2 実験結果

画像番号	前処理	ウェーブレット	u	v	評価値
0	鮮鋭化	Coiflet	0.25	0.5	0.5020
1	鮮鋭化	Daubechies	0.15	0.35	0.9700
2	鮮鋭化	Coiflet	0.05	0.25	0.6421
3	Sobel	Coiflet	0.15	0.4	0.6681
4	鮮鋭化	Coiflet	0.3	0.5	0.6586
5	鮮鋭化	Daubechies	0.1	0.35	0.7204
6	鮮鋭化	Symlet	0.2	0.5	0.7590
7	鮮鋭化	Daubechies	0.1	0.15	0.3335
8	Sobel	Daubechies	0.1	0.2	0.7145



図 2 元画像



図 3 文字領域抽出

強調フィルタである Sobel フィルタや鮮鋭化フィルタを用いることで評価値を向上できた。また Gaussian フィルタについては、画像を平均化するため、ウェーブレット変換後の値に 0 が多くなる。よって Gaussian フィルタは単独で画像に適用するのではなく他のエッジ抽出フィルタと組み合わせることが有効である。今後の課題として、推定処理の解析領域を指定する n のパラメータを画像ごとに変化させることで、どのような文字の大きさにも対応させることである。

参考文献

- [1] 芦田 和毅, 永井 弘樹, 岡本 正行, 宮尾 秀俊, 山本 博章, 情景画像からの文字抽出, 電子情報通信学会論文誌, Vol.J88-D-II, No.9, pp.1817-1824, 2005 年 9 月.
- [2] 劉 詠梅, 山村 毅, 大西 昇, 杉江 昇, シーン内の文字列領域の抽出について, 電子情報通信学会論文誌, Vol.J81-D-II, No.4, pp.641-650, 1998 年 4 月.
- [3] 米津 宏太郎, ウェーブレット変換による文字領域判定, 中央大学理工学部情報工学科卒業論文, 2005 年 3 月.
- [4] 新井 康平, ウェーブレット解析の基礎理論, 森北出版株式会社, 2000 年 11 月.
- [5] 的場 亮太, ウェーブレット変換を利用した文字領域推定, 情報処理学会創立 50 周年記念 (第 72 回) 全国大会講演論文集, vol.2, pp.609-610, 2010 年 3 月.
- [6] 的場 亮太, ウェーブレット変換を利用した文字領域推定, 中央大学大学院理工学研究科情報工学専攻修士論文, 2010 年 3 月.

* The ICDAR 2003 competitions, <http://algoval.essex.ac.uk/icdar/Datasets.html>