

修士論文

エッジ方向成分射影法による
多値文字画像からの特徴抽出

吉村 浩至

2000年2月14日

奈良先端科学技術大学院大学
情報科学研究科 情報システム学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士(工学) 授与の要件として提出した修士論文である。

吉村 浩至

審査委員： 横矢 直和 教授
二矢田 勝行 教授
千原 國宏 教授
栄藤 稔 助教授

エッジ方向成分射影法による 多値文字画像からの特徴抽出*

吉村 浩至

内容梗概

Gabor フィルタを用いたエッジ方向成分射影による多値文字パターン認識手法を提案する. 本手法は, 文字画像に対して複数の Gabor フィルタを適用し局所領域内でエッジ方向成分を累積することにより, 低解像文字に対しても有効で, 位置ずれに頑健な特徴抽出を可能にする. 本手法を映画字幕認識に適用し, 映画 7 本から手動で切り出した文字 865 文字に対して文字認識実験を行った. まず多値画像からの認識と 2 値画像からの認識を比較実験し, 平均認識率が 2 値画像認識で 77.0% に対し, 多値画像認識で 84.4% という結果を得た. これにより映像内の文字認識において多値画像による認識が有効であることが確認できた. またエッジ方向成分射影法を局所領域内で適用することにより, 認識率を 90.8% まで高めることができた. さらに位置ずれがある文字画像に対しても認識実験を行い, 累積方法を変更することによって文字幅の 5% 程度までは位置ずれに頑健であることが確認できた. これにより, 文字認識による動画像インデクシングへの応用が期待される.

キーワード

文字認識, 多値文字画像, Gabor フィルタ, エッジ特徴, ビデオインデクシング

*奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻 修士論文, NAIST-IS-MT9851129, 2000年 2月 14日.

Feature Extraction by Directional Edge Component Projection Method From Gray-scale Character Images*

Hiroshi Yoshimura

Abstract

We propose a gray-scale character recognition method that uses Gabor features. To form a feature vector to be classified, we accumulate the extracted Gabor features along projection lines in local regions, and then categorize them with a standard LVQ classifier. The projective accumulation provides robustness under character deformation caused by variation of font types or imprecise segmentation. We compare the proposed method by experiments with a typical OCR method, for which correct binarization is advantageously given. The proposed method attains 84.4% correct recognition rate, and furthermore 90.8% when using improved feature vectors, while the compared method does 77.0%. The robustness under character deformation of 5% of character size is confirmed by an experiment. The proposed method solves difficulties of character recognition in video indexing; binarization against a complex background and low resolution.

Keywords:

character recognition, gra-scale character image, Gabor filter, edge feature, video indexing

*Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT9851129, February 14, 2000.

目次

1. はじめに	1
2. 文字認識の従来手法と本研究の位置づけ	4
3. エッジ方向成分射影法による特徴抽出	7
3.1 Gabor フィルタによるエッジ抽出	7
3.2 エッジ方向成分射影法	9
3.3 局所領域内で適用するエッジ方向成分射影法	11
3.3.1 エッジ方向成分射影による特徴ベクトルの生成	11
3.3.2 特徴ベクトル要素の分散比計算による累積方法の検討	14
4. 多値文字画像認識と2値文字画像認識の比較実験	17
4.1 実験設定	17
4.2 実験データ	18
4.2.1 学習データ	18
4.2.2 評価データ	20
4.3 実験結果および考察	23
5. エッジ方向成分射影法による文字認識実験	25
5.1 実験設定	25
5.2 実験結果および考察	26
6. 位置ずれがある場合の認識実験	32
6.1 実験設定	32
6.2 実験結果および考察	32
7. 特徴ベクトルの次元削減	36
7.1 主成分分析による特徴ベクトルの次元削減	36
7.2 実験設定	36
7.3 実験結果および考察	37

8. むすび	42
謝辞	43
参考文献	45

目 次

1	映像中の文字認識	2
2	2値化処理による情報量低下	3
3	Gabor フィルタ出力画像の例	8
4	エッジ方向成分射影法による特徴抽出手順	10
5	エッジ方向成分射影法での輝度累積	13
6	輝度の累積方向の説明	15
7	特徴ベクトル累積方向別の分散比	16
8	学習データの例	19
9	評価データの例	21
10	2値化の例	22
11	2値画像認識における誤認識文字の例	24
12	分割数に対する認識率の変化（累積方法：全方向を累積）	29
13	分割数に対する認識率の変化（累積方法：エッジに平行な方向以外を累積）	29
14	分割数に対する認識率の変化（累積方法：エッジに垂直な方向のみを累積）	30
15	分割数に対する認識率の変化（累積方法：方向累積なしの単純累積）	30
16	累積方法の変更による認識率の変化（フィルタ方向：4方向）	31
17	累積方法の変更による認識率の変化（フィルタ方向：6方向）	31
18	位置ずれデータの例	34
19	位置ずれ文字画像の認識結果	35
20	主成分分析における固有値の計算結果	39
21	主成分分析における累積寄与率の計算結果	39
22	主成分分析による次元数削減した特徴ベクトルに対する認識実験の結果	41

表 目 次

1	多値画像認識と2値画像認識の比較実験の結果	24
2	エッジ方向成分射影法による文字認識実験の結果	28
3	主成分分析による次元数削減した特徴ベクトルに対する認識実験 の結果	40

1. はじめに

膨大な映像データから必要なデータを抽出する手段として、映像内の文字をキーとする検索が有効である [1]. そのために映像内の文字認識が必要となる.(図 1 参照)しかし映像内の文字はスキャナで読み込んだ文字に比べて低解像度であるため、図 2 に示すように 2 値化した場合、輝度の中間値の情報量が減ることにより文字識別に必要な特徴が失われ文字認識が困難になる. そこで、2 値化による情報量低下を防ぐため、切り出した文字画像を 2 値化することなく多値画像から直接認識することにより認識率の向上を考える. 文書画像を認識対象とした多値文字画像認識を大まかに分類すると、構造的特徴による手法 [2] [3] [4] [5], フィルタを用いたエッジ抽出による手法 [6] の 2 つに分類される. 構造的特徴による手法は文字ストロークの変化に頑健で、手書き文字に多く適用されているが、ビデオ映像内の文字のような低解像画像やノイズが存在する場合にノイズに対して脆弱である. 本研究では上で述べたように低解像度でノイズがある文字画像を対象としており、手書き文字に比べて、映像内の文字は文字データ毎の文字ストロークの変化が少ないことから、フィルタを用いたエッジ抽出による手法で特徴抽出を行う.

また映像内の文字は、文書画像に比べて背景やノイズのために文字切出しが困難であり、切り出しが成功した場合でもノイズによる位置ずれが生じる場合がある. そこでエッジ抽出のためのフィルタとして Gabor フィルタ [7] を用いる. Gabor フィルタは、直交条件を満たした 2 つのフィルタ応答を利用しているため、エッジ位置に対して位相不変な出力を得ることができ、文字の位置ずれに対して頑健になる [8]. 2 値文字画像に対しては、Gabor フィルタを用いた特徴抽出法がいくつか提案されているが、印刷漢字認識 [9], 手書き漢字認識 [10], 手書き数字認識 [11] に対して検討が行われており、映像内の文字認識に関しては十分な考察は行われていない.

そこで映像内の文字認識手法として、これまで榮藤が提案したエッジ方向成分射影による特徴 [12] [13] を局所領域内で用いる多値文字パターン認識手法を提案する. 本論文では、2 章で文字認識における従来手法を紹介し、本研究の位置づけを述べる. 3 章でエッジ方向成分射影法について詳細を述べる. また、多値文



図 1 映像中の文字認識

字画像認識の有効性を検証した結果および考察を 4 章で述べ、5 章で本手法を映画字幕から切り出した文字画像に対して行った認識実験について述べる。本手法を位置ずれがある文字に対して行った評価実験を 6 章で示し、7 章では特徴ベクトルの次元削減について述べる。最後にまとめと今後の課題について 8 章で述べる。

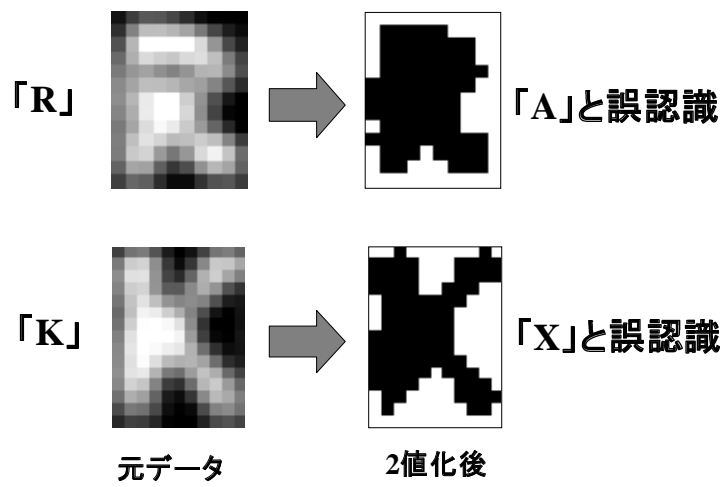


図 2 2 値化処理による情報量低下

2. 文字認識の従来手法と本研究の位置づけ

文書画像解析や文字認識は30年以上に渡って研究されてきたテーマであり、高品質の文書に対してはすでに高い認識精度を達成している [9] [14]. 認識技術が多種多様な実際の応用・利用環境で用いられるにつれて、低品質の文書や文字の認識という難しい技術的要求が出現してきた. 紙の文書はまずスキャナなどの入力機器を通してデジタル画像に変換される. スキャナを通して得られた画像は、各画素ごとに濃淡値表現された輝度を持っている. そのような濃淡値画像をしきい値処理によって2値化し、次に2値画像から認識に有効な特徴を抽出するのが通常である.

文字認識の分野では、2値化された文書画像に対しては、数多くの特徴抽出法が提案されており、文字線の方向に関する特徴について、方向寄与度 [14], 方向指数 [15], ストロークと背景図形からの方向性抽出方向線素 [16] などの特徴抽出法があげられる. しかし、どの特徴抽出法もスキャナで読み込んだ高解像度の文書画像を対象としており、認識対象も印刷漢字、手書き漢字、手書き数字でありビデオ映像やデジタルカメラ画像内の文字のような低解像度文字に対しては十分な考察はされていない.

映像内文字のような低解像度文字に対して2値化処理を行う際、十分な情報なしに2値化処理を行い、しきい値が不適切に選択されると、文字ストロークが途中で切れたり、アルファベットの「R」や「A」などの文字内にある穴がふさがるといった文字識別に重要な位相情報を失う可能性がある. それが原因で文字認識の際、誤認識が発生する.

そこで本研究では、多値の文字画像からの直接特徴量を抽出する手法を考える. 低品質の文書画像に関しては、本研究と同じ視点で、2値化による情報量の低下を防ぐため多値文字認識手法が提案されている. 大きく分けて構造的特徴による手法とフィルタを用いたエッジ抽出による手法とに分けられる. 以下にそれぞれ代表的な手法について説明する.

1. 構造的特徴による手法

- L.Wang, Pavlidisの手法 [2]

濃淡値曲面の尾根を辿ることにより文字中心線が抽出できるという仮定の下に、微分幾何学の演算を用いて、尾根、谷、鞍点の特徴を抽出する手法。この手法の問題点は、手書き文字のような十分に細い線からは尾根を抽出することが可能であるが、文字線が太くなるにつれて尾根検出が困難になることである。しかし本研究で対象とする映像内文字は、文字のスクロール時の可読性を確保するため、あるいは文字以外の背景と区別がつくように文字線は太く表示される。したがってこの手法は尾根検出が困難になることから映像内文字認識に適していないと考えられる。

- Nishidaの手法 [5]

濃淡値曲面を三角形近似し、こう配が急な三角形の辺を連結して文字境界線を抽出して2値画像とし、2値画像認識を行う。手書き文字を対象としており、文字画像の縦横の画素数は、50画素～120画素と高解像度文字を対象としている。しかし、本研究の対象とする映像内文字の縦横の画素数は10画素～20画素程度であるため、低解像度文字に適用すると三角形近似が困難となり映像内文字に適用できないと考えられる。

2. フィルタを用いたエッジ抽出による手法

- G Srikantanの手法 [6]

多値文字パターンに Sobel オペレータを重畳することで、2値化により掠れる文字からも安定した輪郭方向成分を得る特徴抽出法掠れや潰れのある低品質の手書き文字、印刷文字を対象にしているが、文字画像はスキャナで読み込んだ高解像度文字であり、低解像度文字に対しては考察が行われていない。

1章でも述べたように構造的特徴による手法は文字ストロークの変化に頑健で、手書き文字に多く適用されているが、ビデオ映像内の文字のような低解像画像や

ノイズが存在する場合にノイズに対して脆弱である。本研究では上で述べたように低解像度でノイズがある文字画像を対象としており、手書き文字に比べて映像内の文字は文字データ毎の文字ストロークの変化が少ないことから、フィルタを用いたエッジ抽出による手法のアプローチをとる。

3. エッジ方向成分射影法による特徴抽出

本章では、多値文字画像からのエッジ抽出に用いる Gabor フィルタについて説明し、栄藤のエッジ方向成分射影法 [12] の概要を述べた後、提案手法である局所領域内で適用するエッジ方向成分射影法について説明する。

3.1 Gabor フィルタによるエッジ抽出

本節では、エッジ方向成分射影法でのエッジ抽出に用いる Gabor フィルタについて述べる。Gabor フィルタは、Gauss フィルタに正弦波を畳み込んだフィルタであり [7]、2次元 Gabor フィルタは、2次元 Gauss フィルタに2次元面上の1方向へ伝わる正弦波を畳み込んだものである。本論文では、この2次元 Gabor フィルタを、Gabor フィルタと呼ぶことにする。このフィルタ応答は、大脳1次視覚野の線分やエッジに対する応答に近似していることが確かめられており [17]、このフィルタ出力を特徴とすることにより、人間の知覚に近い認識を実現できることが期待される。Gabor フィルタは、ガウス曲面における標準偏差 σ_x , σ_y 、平面波の進行方向 (Gabor フィルタ方向) θ および平面波の周波数 ω により決定される。

Gabor フィルタの方向を θ とすると、Gabor フィルタは以下の式 $h(x, y)$ で表される。

$$\begin{aligned}h(x, y) &= g(x', y') \exp(2\pi i \omega x') \\g(x, y) &= \frac{\exp\{ -[(x/\sigma_x)^2 + (y/\sigma_y)^2]/2 \}}{2\pi\sigma_x\sigma_y} \\(x', y') &= (x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta)\end{aligned}$$

上式の $h(x, y)$ と文字画像上の各画素の輝度をたたみ込み、虚部と実部の2乗和をとることにより、Gabor フィルタ出力を得る。Gabor フィルタは、このように直交条件を満たした2つのフィルタ応答を利用しているため、エッジ位置に対して位相不変な出力を得ることができ、文字の位置ずれに対して頑健になる [8]。本研究では、上記の式のパラメータは $\omega = 0.12$, $\sigma_x = 4.0$, $\sigma_y = 4.0$ として実験を行った。

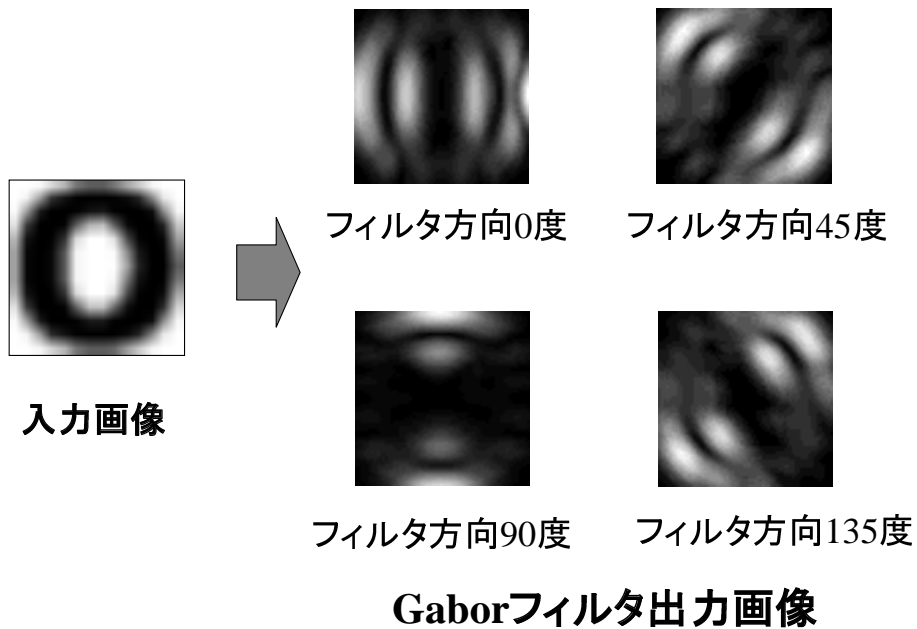


図 3 Gabor フィルタ出力画像の例

文字画像に対して Gabor フィルタを適用した例の $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ とした場合のフィルタ出力画像を図3に示す. 図3からわかるように, この Gabor フィルタを文字画像に適用することにより, 文字ストロークのエッジ成分を抽出することができる. この出力画像から輝度を累積することにより, 文字識別のための特徴を得る.

3.2 エッジ方向成分射影法

榮藤の提案したエッジ方向成分射影法は、Gaborフィルタによって文字のエッジ成分を抽出し、画像全体に対してエッジ成分を複数の方向から累積することによって特徴を得る。以下に特徴抽出手順を示す。

1. 64画素×64画素の入力画像に対して4方向のGaborフィルタにより $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ のエッジ成分を出力し、そのフィルタ出力の実部と虚部の2乗和を出力画像とする。
2. 各方向の出力画像から輝度を累積する。このとき抽出されるエッジに対して平行な方向の累積は行わない(図4参照)。エッジに平行な方向の累積はパターン位置の変動に頑健でないと考えられるからである。累積する画素を2ライン間引き一方向につき16の特徴を得る。これにより4(フィルタ出力方向)×3(累積方向)×16=192次元の特徴ベクトルを得る。
3. 特徴ベクトルの大きさと特徴ベクトルの各要素を正規化する。

本研究では、このエッジ方向成分射影法を局所領域で用いることにより、局所領域内での輝度の単純累積よりも認識性能を向上させ、位置ずれに頑健な特徴抽出を行う。

したがって、局所領域内で適用するエッジ方向成分射影法の手順は、文字画像サイズの正規化、Gaborフィルタ適用、部分領域分割、エッジ方向成分累積、特徴ベクトルの正規化からなる。次節で局所領域内で適用するエッジ方向成分射影法について特徴ベクトル抽出手順を示す。

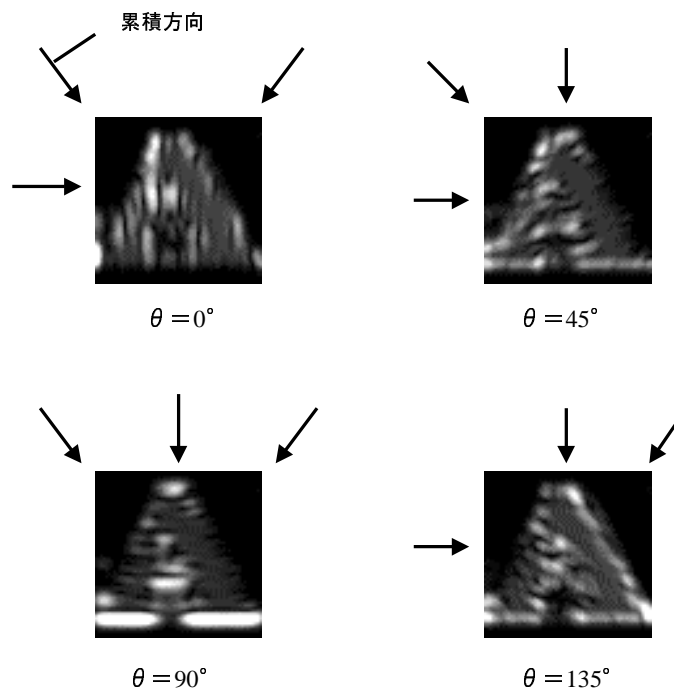


図 4 エッジ方向成分射影法による特徴抽出手順

3.3 局所領域内で適用するエッジ方向成分射影法

3.3.1 エッジ方向成分射影による特徴ベクトルの生成

局所領域内で適用するエッジ方向成分射影法の特徴ベクトルを得る手順は以下の通りである.

1. 文字画像の大きさを 64 画素×64 画素に正規化する.
2. 正規化した文字画像に n 方向の Gabor フィルタを適用し, 文字ストロークのエッジ成分を出力する.
3. 得られた n 枚の各フィルタ出力画像を $k \times k$ の k^2 領域に t 画素の重複を持たせて分割する.
4. 各分割領域を累積方向に m 分割し l 方向の累積を行う. 1つの分割領域から m 個の特徴が得られるので, フィルタ出力 n 方向×部分領域 k^2 領域×累積 l 方向× m 特徴 = $n \times k^2 \times l \times m$ 次元の特徴の特徴ベクトルを得る.
5. 特徴ベクトルの大きさを各要素を正規化する.

本研究の実験で変化させた各パラメータは,

- フィルタ方向数 $n = 4$ および 6 方向
 - ・ 4 方向の場合 : $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$
 - ・ 6 方向の場合 : $\theta = 0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ$
- 分割領域数 $k = 2, 3, 4$, すなわち, 2×2 の 4 領域, 3×3 の 9 領域, 4×4 の 16 領域
- 分割領域内累積方向数 l
 - ・ フィルタ方向の全方向 n 方向
 - ・ エッジに対して平行な方向以外 $(n-1)$ 方向
 - ・ エッジに対して垂直な方向のみ 1 方向

- 分割領域内で一方向累積につき得られる特徴数 m
 - $k = 2, 3$ の場合 : $m = 6$ 特徴
 - $k = 4$ の場合 : $m = 4$ 特徴

の場合について行った。分割領域間で画素数 t の重複を持たせたが、これは分割領域間にまたがる位置ずれに対しても頑健にするためである。予備実験により、重複させる画素は分割領域が 4 領域、9 領域の場合に 4 画素、16 領域で 2 画素とした。輝度の累積方法は、「全方向累積」「エッジに平行な方向以外の累積」「エッジに垂直な方向のみの累積」「方向累積なしの単純累積」4 つの方法で実験を行った。累積方法に関する検討に関しては次節で述べる。

局所領域内で適用するエッジ方向成分射影法の特徴抽出の例として、

- フィルタ方向数 : 4 方向 ($n = 4$)
- 分割領域数 : 4×4 の 16 領域 ($k = 4$)
- 累積方法 : エッジに垂直のみ 1 方向 ($l = 1$)
- 分割領域内での一方向累積につき得られる特徴数 : 4 特徴 ($m = 4$)

とすると、フィルタ 4 方向 \times 16 領域 \times 累積 4 方向 \times 各分割領域 4 特徴 = 256 次元の特徴ベクトルが得られる。この場合の特徴抽出手順を図 5 に示す。

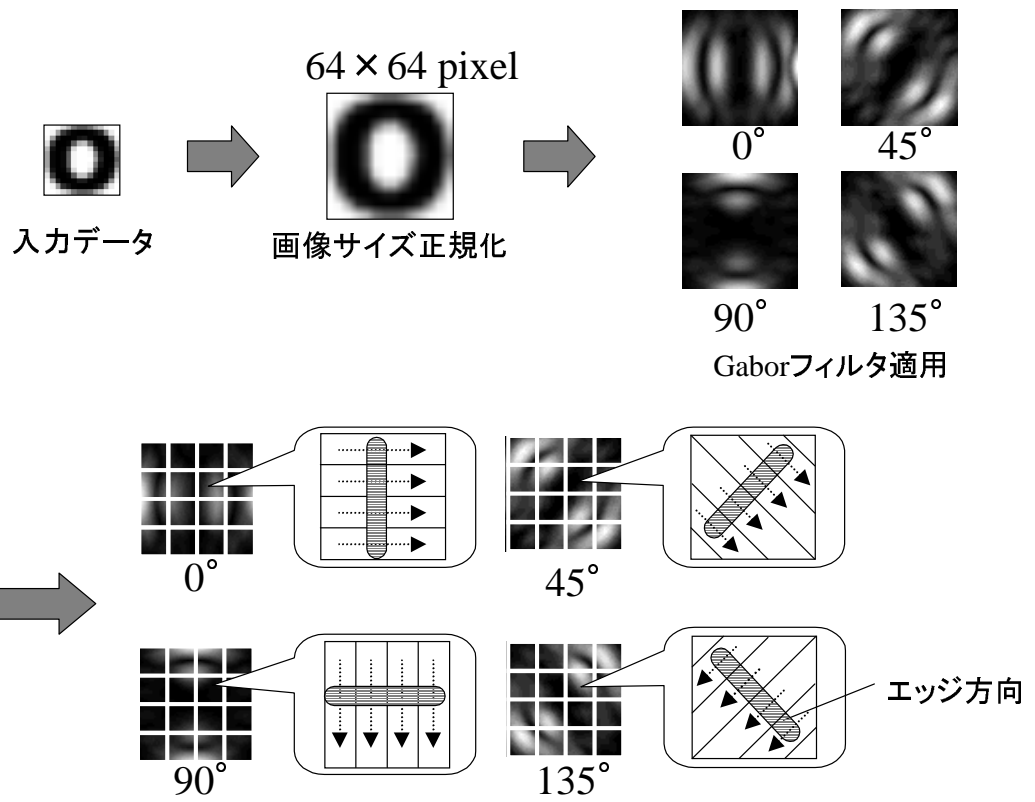


図 5 エッジ方向成分射影法での輝度累積

3.3.2 特徴ベクトル要素の分散比計算による累積方法の検討

本手法でエッジ方向成分を累積する方法は「全方向累積」だけでなく「エッジに平行な方向以外の累積」と「エッジに垂直な方向のみ」を行った。理由は以下の通りである。

- エッジに平行な方向の累積は、累積方向に対して垂直に位置ずれが生じた場合に、累積した値が大きく変動してしまい、文字認識の際に、識別するための特徴として重要でない。よってエッジに平行以外の累積を行うことで認識性能向上を目指す。
- エッジに垂直な方向の累積は、エッジに垂直な位置ずれが生じた場合でも各領域内で一定の特徴が得られるため、位置ずれに頑健な特徴が得られる。この累積のみ用いることにより位置ずれに強い特徴を得る。

上述の累積方向を図で表したものを図 6 に示す。

上述の仮定を検証するために予備実験を行った。フォントの異なる文字に対して榮藤のエッジ方向成分射影法を適用して特徴ベクトルを抽出し、各特徴ベクトル要素の分散比計算を行い、どの累積方向が有効であるかを調べた。

分散比は、各カテゴリ間の分散を「カテゴリ間分散」、各カテゴリ内での分散を「カテゴリ内分散」とすると、

$$\text{分散比} = \frac{\text{カテゴリ間分散}}{\text{カテゴリ内分散}}$$

で表される。カテゴリ間の特徴ベクトル要素の分散が大きいとその要素はカテゴリ間でばらつきが大きいため識別の特徴としては有効である。一方、カテゴリ内の特徴ベクトル要素の分散が大きいとその要素は同一カテゴリ内でのばらつきが大きいため識別のための特徴としては有効ではない。したがって、カテゴリ間分散が大きく、カテゴリ内分散が小さい場合、すなわち分散比が大きければその特徴は有効であると言える。

実験で用いた文字データは、学習データと同様 Windows フォントから作成した。文字数は、5 フォント × 26 カテゴリ = 130 文字、各文字画像の縦横画素数は、30 画素前後である。

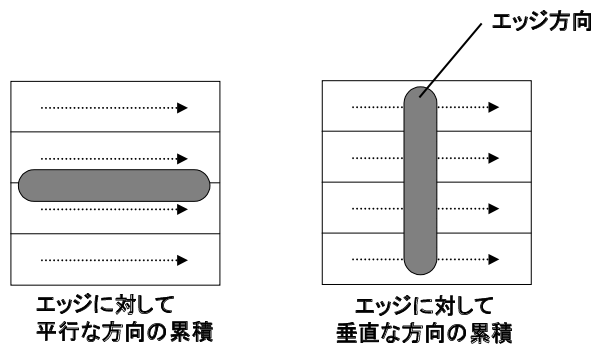


図 6 輝度の累積方向の説明

分散比の計算結果をグラフで表したものを図 7 に示す。グラフ内で黒色の棒がエッジ方向に対して垂直な方向の累積であり、白色の棒がエッジ方向に対して平行な累積を表している。

図 7 のグラフよりエッジに垂直な累積は、分散比が他の要素に比べて比較的高く識別の特徴として有効であることが確認できる。またエッジに平行な累積は、他の要素に比べて分散比が低い傾向があり識別の特徴としては有効でないことを示している。

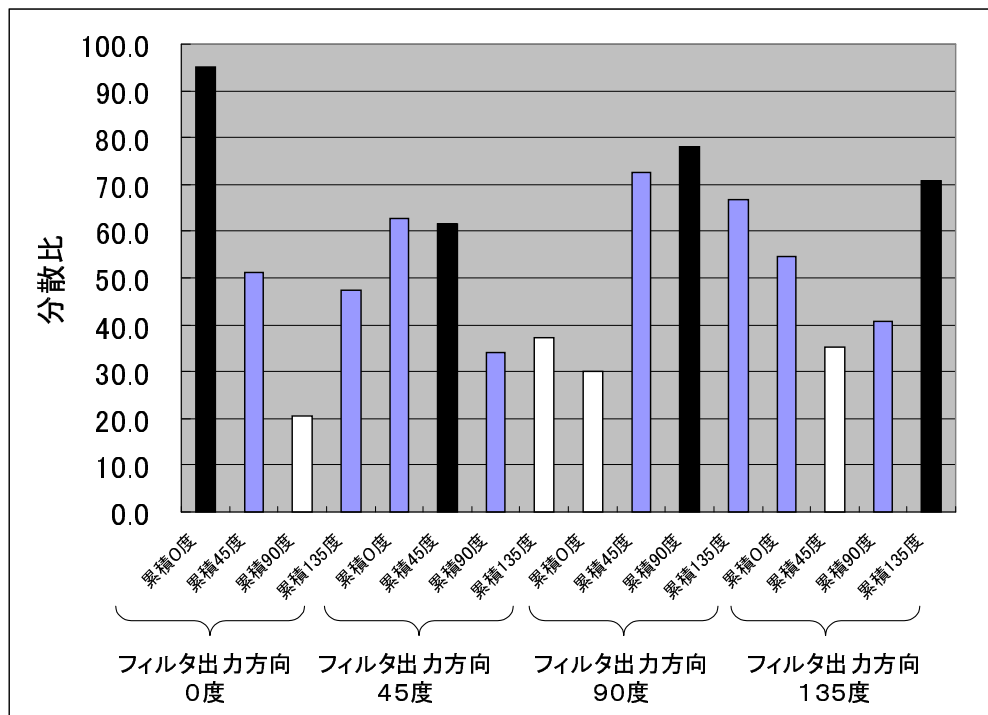


図 7 特徴ベクトル累積方向別の分散比

4. 多値文字画像認識と2値文字画像認識の比較実験

4.1 実験設定

本章では、フィルタを用いたエッジ抽出による多値文字画像認識の有効性を検証するために、本手法の簡略形として、Gaborフィルタ特徴の局所累積のみを特徴とする手法で認識実験を行った。認識対象は、映画のエンドロールから手動で切り出したアルファベット大文字の26カテゴリである。

本実験での多値画像認識の特徴抽出法は、各文字画像サイズを正規化し、 $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ の4方向のGaborフィルタを適用する。4枚のフィルタ出力画像を 4×4 の16領域に分割した後、各分割領域の輝度を単純に累積する。これによりフィルタ方向4方向 \times 16領域の64次元の特徴ベクトルが得られる。

また比較手法として2値画像認識の特徴抽出には、高解像の文書画像の認識に用いられている手法の一つである輪郭方向密度特徴 [19] を用いた、輪郭方向密度特徴は、文字画像を 4×4 の16領域に分割し各領域に含まれる文字輪郭の各点の方向性が $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ のいずれに属するかを求め、分割領域ごとに求めた各方向の密度（計64次元）を特徴とするものである。

すなわち多値画像認識、2値画像認識両方の特徴抽出において、 4×4 の16領域の分割領域内に $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ の4方向の特徴がどの程度含まれているかを識別のための特徴としている。

比較手法では領域分割する際、文字画像の輝度値から文字部分の重心を求め、その重心から分割位置を決定するという重心分割が用いられているが、本実験では、純粋に2値画像認識と多値画像認識の比較実験を行うため重心分割を省いて実験を行った。

本研究で行った文字認識実験における認識結果の判定方法を以下に示す。

1. 学習データに対して各特徴抽出法を適用して特徴ベクトルを出力する。
2. 学習ベクトル量子化法 (LVQ) [20] により学習データの特徴ベクトルを学習し、各カテゴリごとの参照ベクトルを出力する。
3. 評価データに対しても各特徴抽出法を適用し、得られた評価データの特徴

ベクトルと各カテゴリの参照ベクトルとのユークリッド距離を計算し、最も距離が近いカテゴリを認識結果として出力する。

4. 出力されたカテゴリが正解のカテゴリと一致すれば認識成功とする。

学習データの学習に用いた学習ベクトル量子化法 (LVQ) [20] とは、最適なパターン分類できるように、競合学習によって参照ベクトルの更新を行う学習法である。入力特徴ベクトル、参照ベクトルにカテゴリを割り振り、入力特徴ベクトルと参照ベクトルのカテゴリを比較し、一致していれば参照ベクトルの距離を近づけ、一致しなければ遠ざける操作を繰り返すことで学習を行う。このような操作によって、理論的に最適な識別境界を与えるベイズ識別境界を形成することができる。本実験での LVQ の参照ベクトルの個数は、予備実験により 1 カテゴリにつき 81 個とした。

また実験に用いた学習データおよび評価データは、次節に示すデータを用いた。

4.2 実験データ

4.2.1 学習データ

学習に用いた文字データは、Windows フォントから作成した。認識対象である映像内文字に形状を似せるため以下の処理を行って、一つのフォントから複数の文字画像を作成した。

作成手順は

1. 文字画像取得し、回転を付加
(-0.8 度, 回転なし, 0.8 度の 3 種類)
2. ガウシアンフィルタで平滑化
3. 解像度変換 (10 画素前後, 20 画素前後, 30 画素前後の 3 種類)

である。作成したデータ数は 1 カテゴリにつき回転 3 種類 \times 解像度 3 種類 \times 18 フォント = 162 種類, アルファベット大文字 26 カテゴリで 4212 文字である。

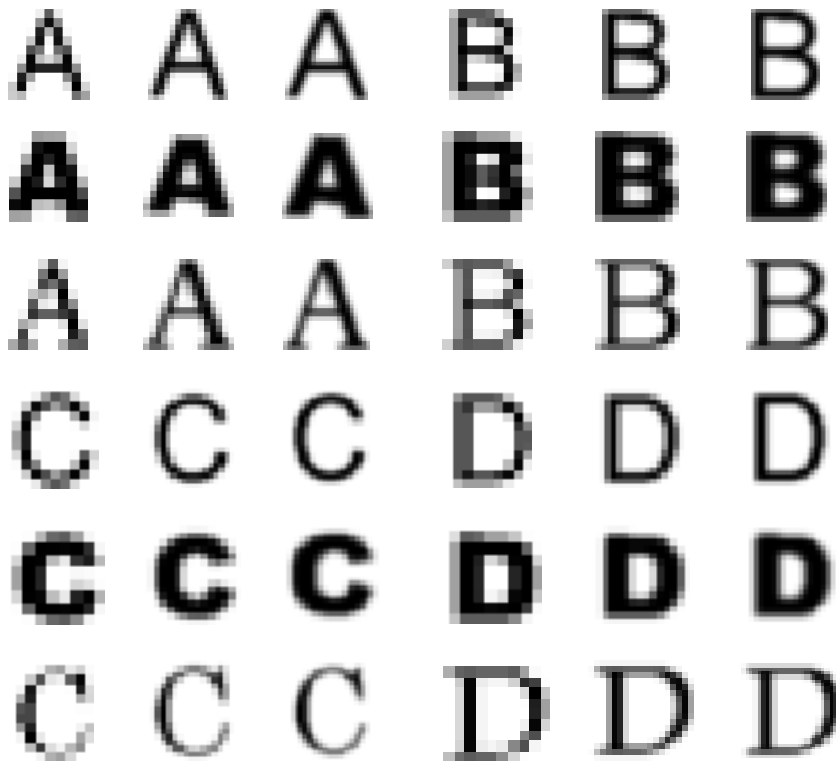


図 8 学習データの例

なお学習に用いた Windows フォントは, "Arial" "Arial Black" "Arial Narrow" "Book Antiqua" "Bookman Old Style" "Calisto MT" "Century" "Copperplate Gothic Bold" "Copperplate Gothic Light" "Courier New" "Garamond" "Impact" "Lucida Console" "Lucida Sans Unicode" "News Gothic MT" "Tahoma" "Times New Roman" "Verdana" の 18 種である.

学習データ例を図 8 に示す.

4.2.2 評価データ

7種類の外国映画のエンドロールから名前の部分を手動で切り出し評価データとした。評価データに関しての詳細を以下に示す。

- 7映画から総文字数 865 文字
- カラー画像をグレースケール画像に変換した1ピクセル当たり8ビットの多値画像
- 各画像は縦横のサイズが10画素～25画素の低解像度文字画像
- カテゴリはアルファベット大文字の26カテゴリ

評価データ例を図9に示す。

また比較のための2値画像認識に用いる文字画像の2値化には、大津の自動しきい値選定法 [18] を用いた。この方法は、画像全体の濃淡値分布の判別分析により、文字と背景の2つのクラスの分離度を最大にする最適なしきい値を決定するものである。自動しきい値選定法により2値化処理した文字画像例を図10に示す。

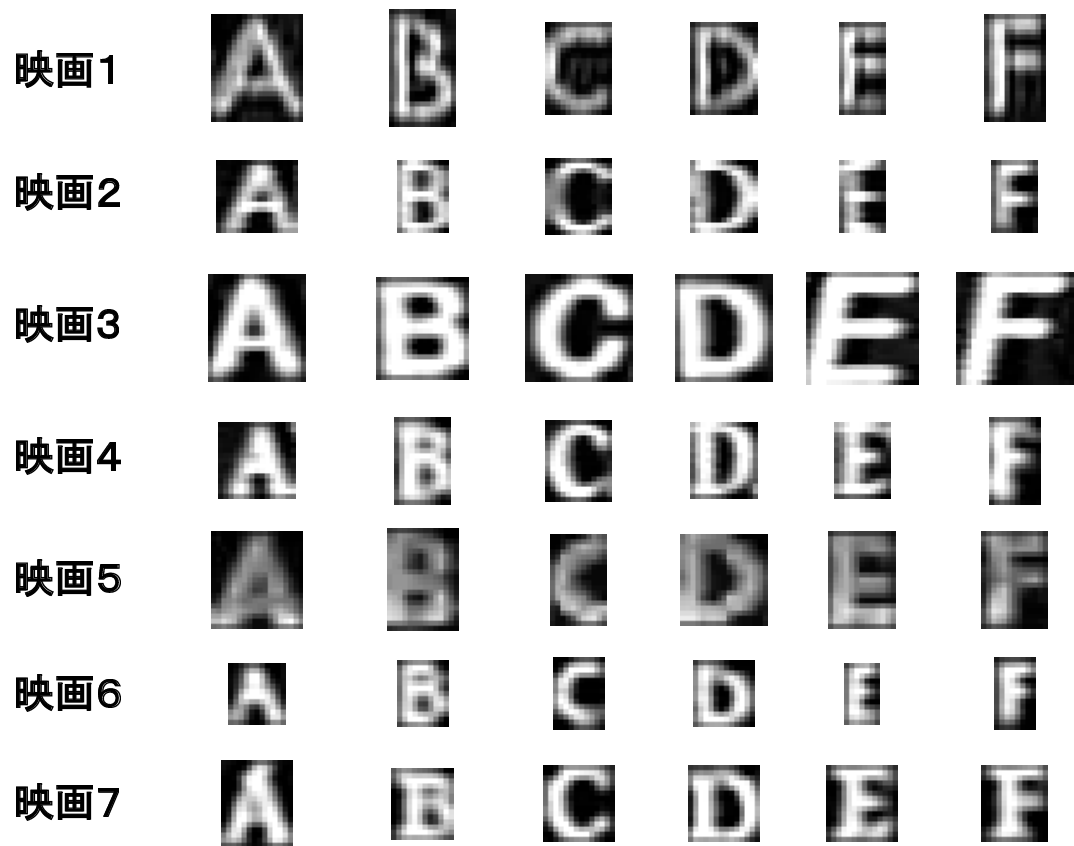


図9 評価データの例

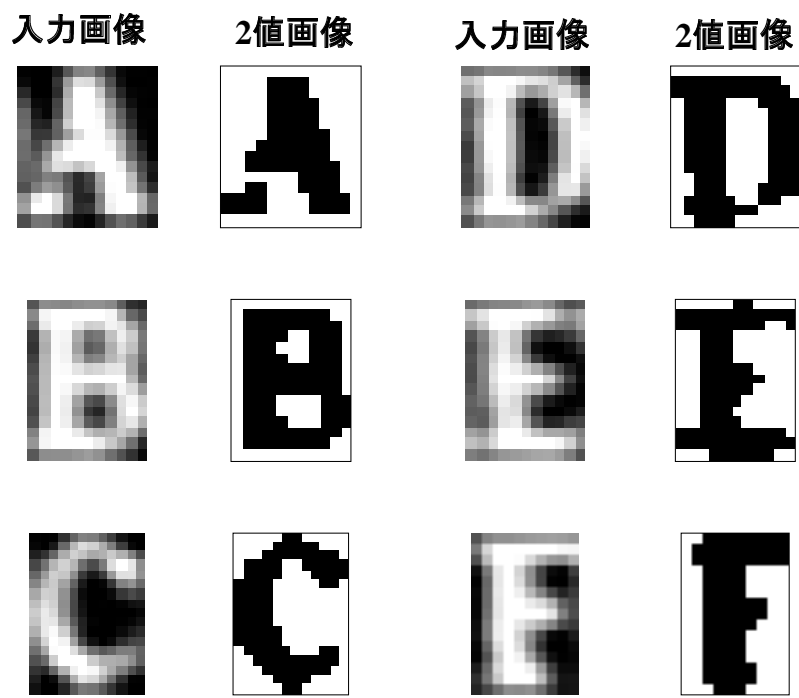


図 10 2 値化の例

4.3 実験結果および考察

2値画像認識と多値画像認識の比較実験の結果を表1に示す。表に示す通り平均認識率は、多値画像認識で84.4%、2値画像認識で77.0%という結果を得た。これより4節の特徴抽出法においては、2値画像認識よりも多値画像認識の方が有効であることが確認できた。各映画毎の認識率を見てみると、図9に示したデータのうち特に低解像度である映画2、映画6、映画7に関しては、多値画像認識が2値画像認識よりも約15%認識率が高くなるという結果を得た。これにより特に低解像度画像に対しては多値画像認識は2値画像認識より有効であることを示している。

映画5に関しては、多値画像認識よりも2値画像認識の方が認識率は高くなっている。図9の映画5のデータを見てもわかるように、映画5は文字データのコントラストが低い。多値画像認識の場合、文字ストロークのエッジ抽出を行う際に、文字データのコントラストが低いと正確なエッジの抽出ができないため、識別に必要な特徴が得られなかったことが認識率低下の原因と考えられる。これは、Gaborフィルタを適用する前に、文字画像の輝度のコントラストを正規化することにより、改善されると思われる。

図11に多値文字画像認識で成功し、2値文字画像認識で誤認識した例を示す。この図より、2値化処理が行われる際に、輝度の中間値が削除されるため本来の文字の形が変わっていることが確認できる。このため2値文字画像認識においては識別に有効な特徴が得られず誤認識が発生したと考えられる。一方、多値画像では2値化処理による情報量の低下がないため、低解像度文字でも認識可能であったと考えられる。

表 1 多値画像認識と2値画像認識の比較実験の結果

評価映画	多値画像認識率 (%)	2値画像認識率 (%)
映画1	93.2	86.3
映画2	90.4	74.4
映画3	96.7	86.8
映画4	74.8	70.9
映画5	64.5	79.3
映画6	82.1	68.3
映画7	89.2	73.1
平均認識率	84.4	77.0

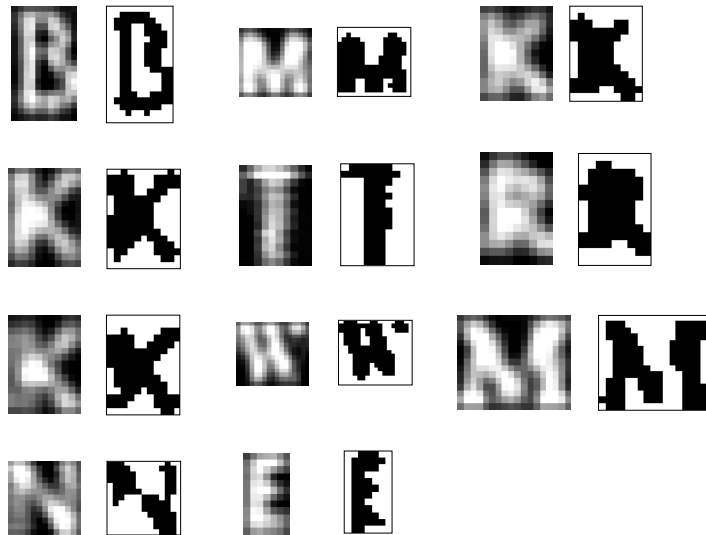


図 11 2値画像認識における誤認識文字の例

5. エッジ方向成分射影法による文字認識実験

5.1 実験設定

3.3節で述べたエッジ方向成分射影法のパラメータである方向数 n , 分割領域数 k^2 を変化させて映像内文字に対する認識実験を行った. 4章の多値画像認識と2値画像認識の比較実験において, 多値画像認識の手法では, フィルタ方向数4方向, 分割領域数 4×4 の16領域として実験を行ったが, フィルタ方向数を増やすことにより, より文字ストロークに対して正確なエッジが抽出でき, 認識率向上が期待されるため, フィルタ4方向に加えて6方向に対しても実験を行った.

また分割領域数に関しては, 4節の多値画像認識と2値画像認識の比較実験において, 分割領域数を 4×4 の16領域で実験を行ったが, 今回の文字認識対象がアルファベットであり, 文字構造が簡単なことから分割領域数を減らすことによって次元数を減らすことを考える. 分割領域数を減らすことによる認識率低下がなければ, より次元数が少ない特徴の方が特徴抽出時間および学習時間の短縮につながる. そこで分割領域数をいくらにすれば最も高い認識率が得られるか検証するために分割領域数16領域に加えて, 4領域, 9領域の場合でも実験を行った.

学習データおよび評価データは, 4章の多値文字認識と2値文字認識の比較実験と同様, 4.2節の文字データを用いた

今回の実験で変化させたパラメータをまとめると以下の通りである.

- フィルタ方向数 $n = 4$ および 6
 - ・ 4方向の場合 : $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$
 - ・ 6方向の場合 : $\theta = 0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ$
- 分割領域数 $k = 2, 3, 4$, すなわち, 4領域, 9領域, 16領域
- 分割領域内累積方法
 - ・ フィルタ方向の全方向 n 方向
 - ・ エッジに平行以外 $(n-1)$ 方向
 - ・ エッジに垂直のみ1方向

- ・方向を変化させての累積なし（単純累積）

以上のパラメータについて認識実験を行った。

5.2 実験結果および考察

エッジ方向成分射影法による映像内文字認識実験の実験結果を表 2に示す。表中の特徴数は、各分割領域内で累積を1方向行ったときに得られる特徴数を表している。そのため次元数は、フィルタ方向数×分割領域数×累積方向数×特徴数で求められる。

表 2より最も認識率が高く映像内文字認識に有効である特徴抽出法は、フィルタ方向数4方向、分割領域数16領域、累積方法がエッジに平行以外3方向の場合で、認識率90.8%という結果を得た。これにより方向累積を行わずに単純累積する場合よりもエッジ方向成分射影法を適用した方が、認識性能は向上することがわかった。また表 2の実験結果から各パラメータを変更したときの認識率変化を見るために横軸に分割領域数、縦軸に認識率を取り、各フィルタ方向数毎にプロットしたグラフを図 12, 図 13, 図 14, 図 15に示す。

- 分割領域数に関する考察

図 12, 図 13, 図 14, 図 15より、累積方法にかかわらず分割領域数が4領域、9領域の場合よりも16領域の方が認識率が高くなっている。これは16領域よりも荒く分割して次元数を減らすよりも、16領域分割のままの方が、認識性能は高いことを示している。これより、本手法に用いる分割領域数は16領域が適当であると考えられる。

- フィルタ方向数に関する考察

図 12, 図 13, 図 14, 図 15より、図 12の累積方法が全方向累積で分割領域数9領域の場合を除いて、フィルタ方向が4方向がフィルタ方向6方向に比べて同等かそれ以上の認識率になっている。これはフィルタ方向を6方向にして細かくエッジ成分を抽出する必要がないことを示しており、本手法に用いるフィルタ方向数は4方向が適当であると考えられる。

- 累積方法に関する考察

輝度の累積方法を変更した際の認識率変化を見るために、横軸に累積方法、縦軸に認識率をプロットしたグラフを図 16, 図 17に示す。図 16, 図 17より、フィルタ方向4方向, 6方向どちらの場合でも、累積方法が「全方向累積」よりも「エッジに平行以外累積」を行う場合が最も認識率が高くなっている。これはエッジに平行な累積によって得られる特徴ベクトル要素は、累積方向に対して垂直に位置ずれが生じた場合に、特徴要素の値が大きく変動してしまい、文字認識の際に、識別するための特徴として重要でないからと考えられる。

表 2 エッジ方向成分射影法による文字認識実験の結果

フィルタ方向数	分割領域数	累積方法	特徴数	次元数	認識率 (%)
4 方向	4 領域	累積全方向 4 方向	6 特徴	384	83.7
4 方向	4 領域	累積平行以外 3 方向	6 特徴	288	89.5
4 方向	4 領域	累積垂直のみ 1 方向	6 特徴	96	80.6
4 方向	4 領域	累積方向なし (単純累積)	1 特徴	16	71.4
4 方向	9 領域	累積全方向 4 方向	6 特徴	864	84.1
4 方向	9 領域	累積平行以外 3 方向	6 特徴	648	90.0
4 方向	9 領域	累積垂直のみ 1 方向	6 特徴	216	86.3
4 方向	9 領域	累積方向なし (単純累積)	1 特徴	36	85.6
4 方向	16 領域	累積全方向 4 方向	4 特徴	1024	86.2
4 方向	16 領域	累積平行以外 3 方向	4 特徴	768	90.8
4 方向	16 領域	累積垂直のみ 1 方向	4 特徴	256	88.8
4 方向	16 領域	累積方向なし (単純累積)	1 特徴	64	87.3
6 方向	4 領域	累積全方向 6 方向	4 特徴	576	82.6
6 方向	4 領域	累積平行以外 5 方向	4 特徴	480	87.7
6 方向	4 領域	累積垂直のみ 1 方向	4 特徴	96	77.1
6 方向	4 領域	累積方向なし (単純累積)	1 特徴	24	71.5
6 方向	9 領域	累積全方向 6 方向	4 特徴	1296	87.3
6 方向	9 領域	累積平行以外 5 方向	4 特徴	1080	88.9
6 方向	9 領域	累積垂直のみ 1 方向	4 特徴	216	83.8
6 方向	9 領域	累積方向なし (単純累積)	1 特徴	54	84.7
6 方向	16 領域	累積全方向 6 方向	4 特徴	2304	86.6
6 方向	16 領域	累積平行以外 5 方向	4 特徴	1920	90.0
6 方向	16 領域	累積垂直のみ 1 方向	4 特徴	384	85.4
6 方向	16 領域	累積方向なし (単純累積)	1 特徴	96	86.9

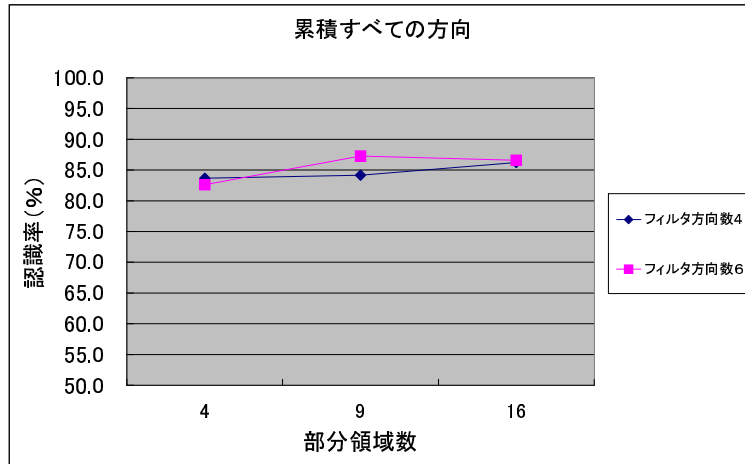


図 12 分割数に対する認識率の変化（累積方法：全方向を累積）

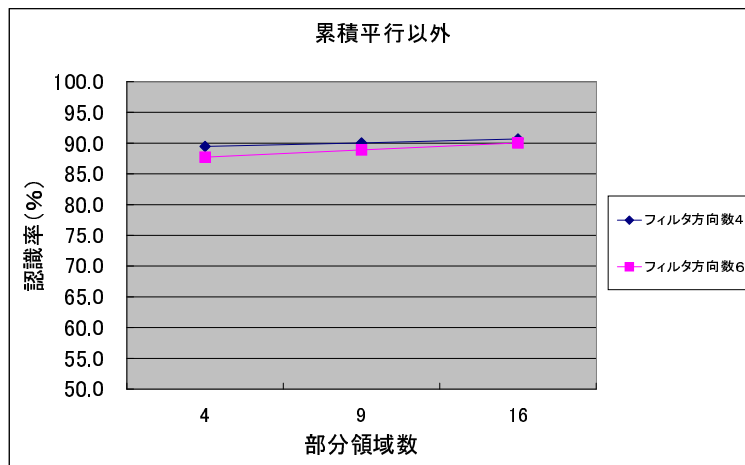


図 13 分割数に対する認識率の変化（累積方法：エッジに平行な方向以外を累積）

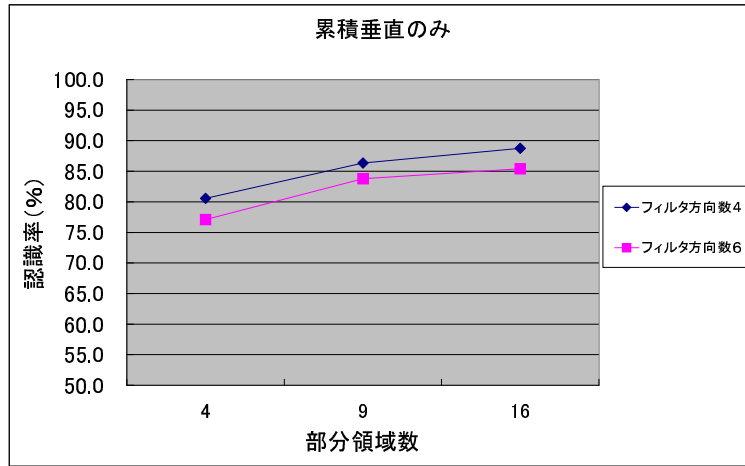


図 14 分割数に対する認識率の変化（累積方法：エッジに垂直な方向のみを累積）

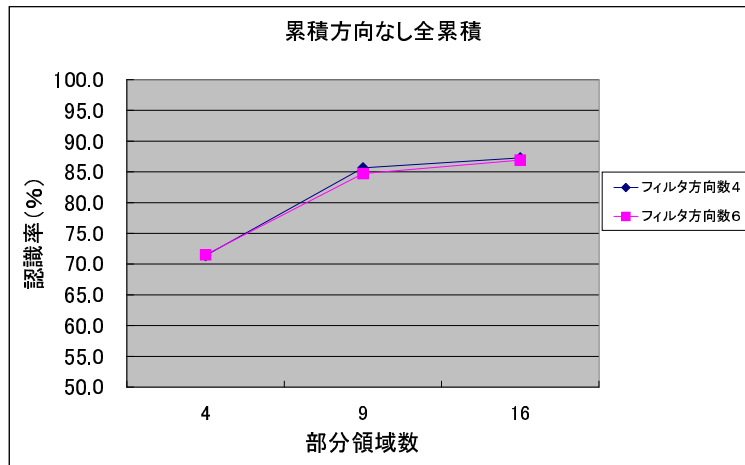


図 15 分割数に対する認識率の変化（累積方法：方向累積なしの単純累積）

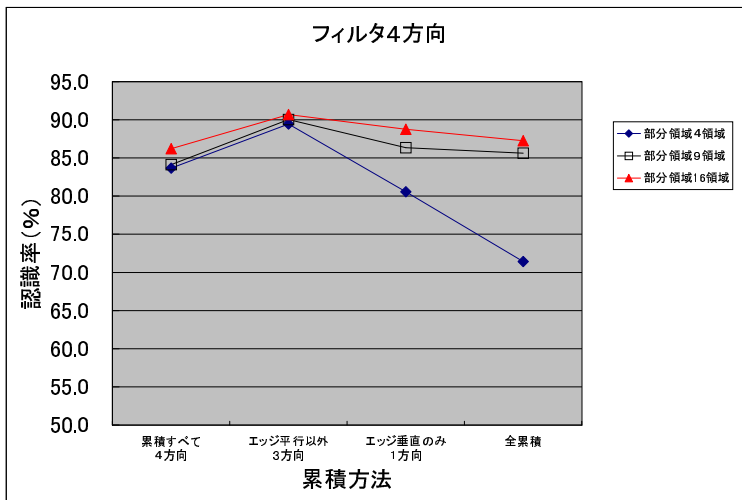


図 16 累積方法の変更による認識率の変化（フィルタ方向：4方向）

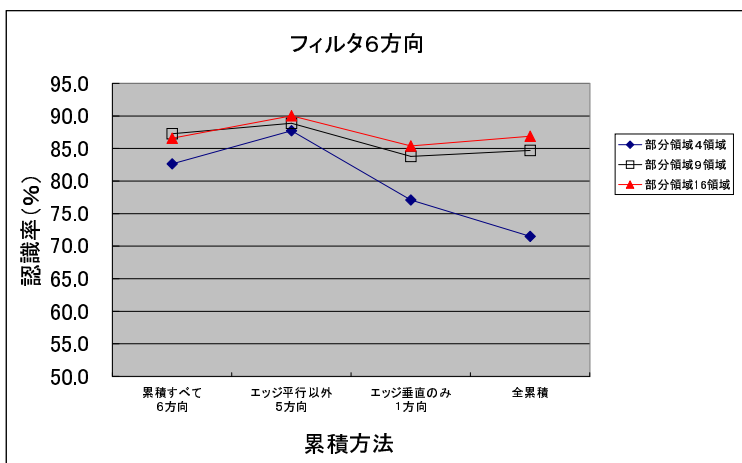


図 17 累積方法の変更による認識率の変化（フィルタ方向：6方向）

6. 位置ずれがある場合の認識実験

6.1 実験設定

提案手法が位置ずれに対してどの程度頑健であるかを検証するための評価実験を行った。4.2.2節の評価データに人工的に位置ずれを持たせた文字画像を認識させることにより、文字画像のサイズの何%までの位置ずれまで保証するかを検証する。本実験のデータは4.2のすべての評価データに対し、文字サイズの5%から25%までの大きさの背景部分を5%毎に文字データの上下左右に付加することによって作成した。作成した位置ずれがある場合のデータ例を図18に示す。学習データは、4.2.1と同様の学習データを用いた。また学習データの学習は前述の実験同様LVQで行った。学習ベクトルの参照ベクトル数は1カテゴリにつき81個である。

累積方法は3.3節で述べた「全方向累積」、「エッジに平行以外累積」、「エッジに垂直のみ累積」、「方向累積なしの単純累積」の4つの累積方法に対して行い、各累積方法がどの程度の位置ずれまで保証するか検証する。本実験で用いたエッジ方向成分射影法の各パラメータは、5章のエッジ方向成分射影法の認識実験で最も認識率が高かったフィルタ方向数4方向、分割領域数16領域で認識実験を行った。また参考のため、4節の2値画像認識の比較手法である輪郭方向密度特徴に対しても実験を行った。

6.2 実験結果および考察

位置ずれがある場合の認識実験の実験結果を図19に示す。方向累積なしの場合は、10%程度までの位置ずれが起こった場合においても認識率低下が起こっていない。これは分割領域内でどの方向に位置ずれが起きても一定の特徴が得られるため、各分割領域間にまたがる位置ずれが起きなければ、一定の認識性能が得られるからと考えられる。

方向累積なしに次いで位置ずれに強い累積方法は、エッジに垂直のみ累積の場合である。この累積方法では、5%程度までの位置ずれが生じても認識率低下は起こらなかった。これは、エッジに垂直な方向のみを累積した場合、3.3節で述べた

通り，エッジに垂直な方向に対して位置ずれが生じても各領域内で一定の特徴が得られるため，残りの2つの累積方法よりも位置ずれに頑健な特徴抽出が可能であるからと考えられる。

残りの全方向累積，エッジに平行以外累積および比較手法の輪郭方向密度特徴の場合では，5%の位置ずれがある場合でも認識率低下が起こり，位置ずれに対しては弱いことが確認できた。

また位置ずれが生じていない場合の認識率は，5章の実験結果である表2のフィルタ方向数が4方向，分割領域数が16領域の認識結果から

- エッジに平行以外累積 : 90.8%
- エッジに垂直のみ累積 : 88.8%
- 方向累積なし単純累積 : 87.3%
- 全方向累積 : 86.2%

であり，エッジに平行以外の累積の場合が最も高い認識率となっている。以上の結果から，以下のことが結論付けられる。

- エッジ平行以外の累積：認識率は最も高いが，位置ずれに弱い
- エッジに垂直のみ累積：認識率は2番目に高く，位置ずれに対しては文字サイズの5%の位置ずれまで認識率は低下しない
- 方向累積なし単純累積：認識率は上の累積方法に比べて低いが，位置ずれに対しては最も頑健で文字サイズの10%の位置ずれまで認識率は低下しない

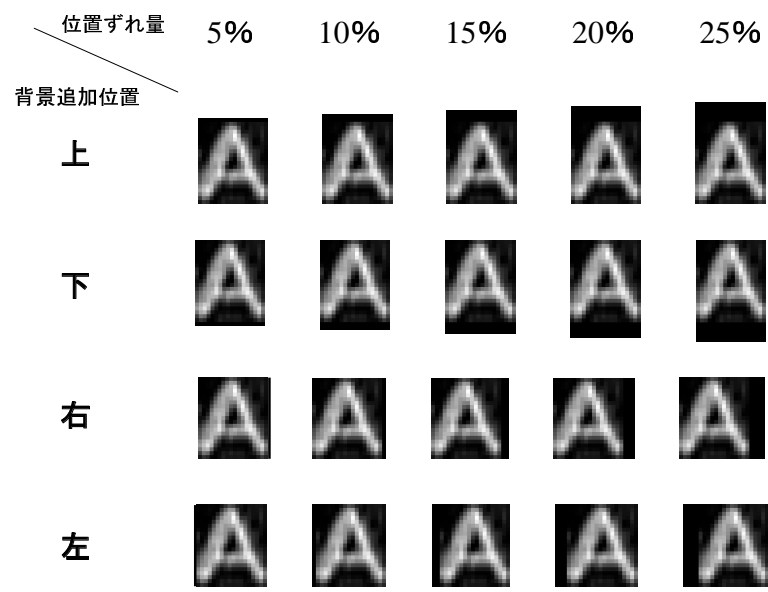


図 18 位置ずれデータの例

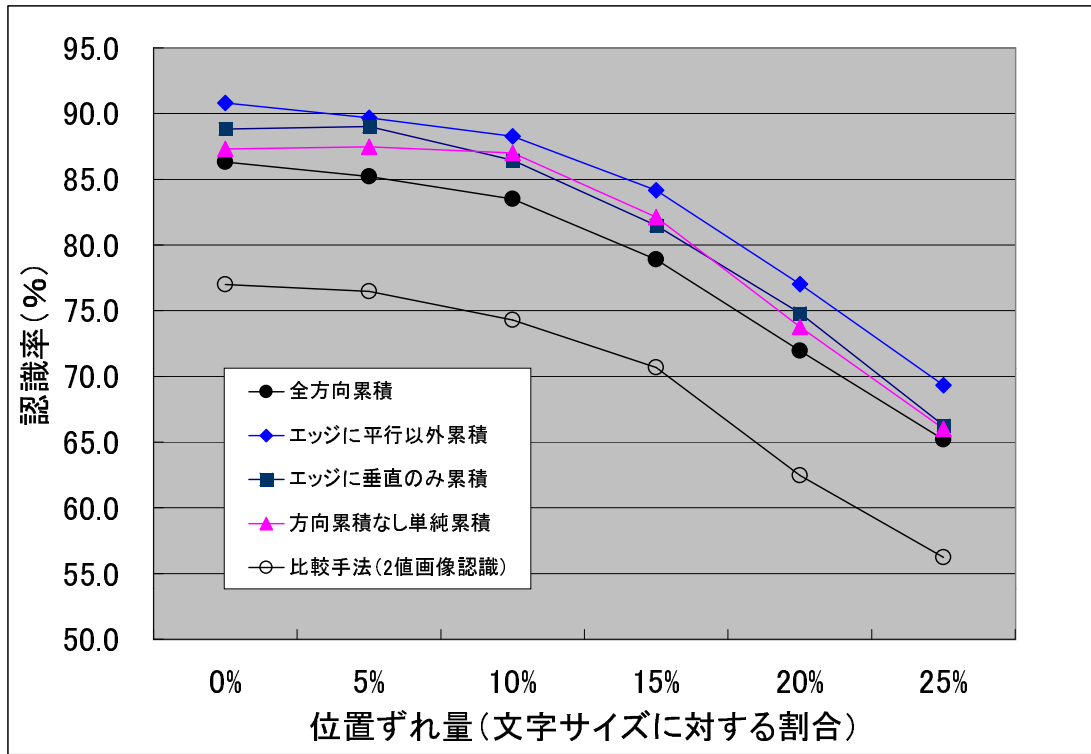


図 19 位置ずれ文字画像の認識結果

7. 特徴ベクトルの次元削減

7.1 主成分分析による特徴ベクトルの次元削減

5章において、映像内の文字認識に対して認識率の向上を計った。しかし、識別に有効な特徴量を増加させるために特徴ベクトルの次元数が増加し、文字認識に必要な計算量やメモリ使用量が増大する問題が生じる。この問題を解決するためには、認識率を低下させることなく特徴ベクトルの次元数を減少させる必要がある。本章では、主成分分析を特徴ベクトル次元数の削減に適用し、次元削減した特徴ベクトルに対して行った認識実験について述べる。

主成分分析 [22] による次元数削減は、まず学習データから得られた n 次元の特徴ベクトル X から全共分散行列 S_t を計算し、次式を満たす固有ベクトル行列 Φ と固有値行列 Λ を求める。

$$S_t \Phi = \Phi \Lambda \quad (1)$$

Λ の対角要素を λ_i ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$)、対応する固有ベクトルを Φ_i ($\Phi = [\Phi_1 \Phi_2 \dots \Phi_n]$) とすると線形結合により n 個の主成分 $z_i = \Phi_i^T X$, ($i = 1, 2, \dots, n$) が得られる。固有値（全分散）の大きい方から n' 個の主成分を選択することにより n' 次元のベクトル $Z = (z_1, z_2, \dots, z')$ が得られる。この Z を次元削減後の特徴ベクトルとして認識に用いる。

7.2 実験設定

次元削減の対象とした特徴は、5章のエッジ方向成分射影法の認識実験で述べた特徴抽出法のうち以下の特徴について実験を行った。

- フィルタ方向数：4 方向
- 分割領域数：16 領域
- 累積方法：全方向累積
- 特徴ベクトル次元数：1024 次元

学習データおよび評価データは、4章、5章の認識実験と同様、4.2節の文字データを用いた。

7.3 実験結果および考察

主成分分析における固有値（全分散）および累積寄与率の計算結果を図20、図21に示す。寄与率とは一つの主成分が元の全変数が持っている情報の何%を説明できるかということを表す指標であり、第1主成分から第 m 主成分までの寄与率の合計を累積寄与率という。累積寄与率は識別に適用する次元数を決定するのに用いられる。図21より、次元数が600次元で累積寄与率がほぼ100%になっている。つまり600次元の主成分で全変数が持つ情報の大部分を説明できることを示しており、認識に必要な次元数は600次元程度であることが確認できる。

主成分分析によって次元削減し、各次元数に対して行った認識実験の結果を表3、図22に示す。表3、図22の各次元数における認識率結果で示すとおり、元の次元数1024次元における認識率が87.3%に対し、主成分分析によって半分の次元数である512次元に削減した場合でも認識率は87.6%という結果が得られた。よって次元数を半分に削減した場合でも認識率が低下しないことが確認できた。この次元数を削減した特徴ベクトルを文字認識に採用することにより、計算量やメモリ使用量を削減することができる。また適切な次元削減を行うことによって、識別に有効な別の特徴量を付加することができるので、さらなる認識率向上も可能である。

5章では累積方向を変更することによって次元削減を行い、全方向累積1024次元で86.2%であるのに対し、平行以外の累積768次元では90.8%という結果が得られた（フィルタ方向数:4方向、分割領域数16領域の場合）。1024次元から識別に有効でない特徴を削除することで認識率が向上したと考えられる。見方を変えれば、1024次元には有効な特徴768次元が含まれているにもかかわらず、識別に有効でない次元が含まれているために認識率が低下したと見ることができる。本章で行った主成分分析による次元削減によって識別に有効でない特徴次元を削減することになり認識率の向上が期待されたが、表3、図22の実験結果から最高の認識率は700次元で88.1%と、5章での最高認識率である90.8%には及ばなかつ

た. これは今回行った全分散の大きい主成分から選択する方法では, 必ずしも識別に有効な特徴軸が選択されていないためと考えられる. また, 特徴軸は同一であっても全分散の大きい主成分から特徴を選択するのではなく, 各主成分のクラス内分散とクラス間分散を求め, 分散比 (F 比) の大きい主成分から順に選択する方法も考えられる [21]. この方法はクラス内分散も考慮するため, 今回行った方法よりも識別に有効な特徴が得ることが可能であり, 認識率の向上が期待できる.

また他の方法として, 複数のクラスに対して分散比 (F 比) を最大にする基底軸を求める手法である正準判別分析 [23] による次元削減が考えられる. この方法によって得られる特徴ベクトルは“カテゴリ数-1”と少ないため, 本論文で対象としているアルファベット大文字の26カテゴリの場合, 得られる特徴ベクトルは25次元までと少ないため有効でない. しかし, 多くのカテゴリがある漢字認識の場合には有効であると考えられるので, 映像内の漢字認識を行う場合の特徴ベクトル次元数削減には正準判別分析は有効であると考えられる.

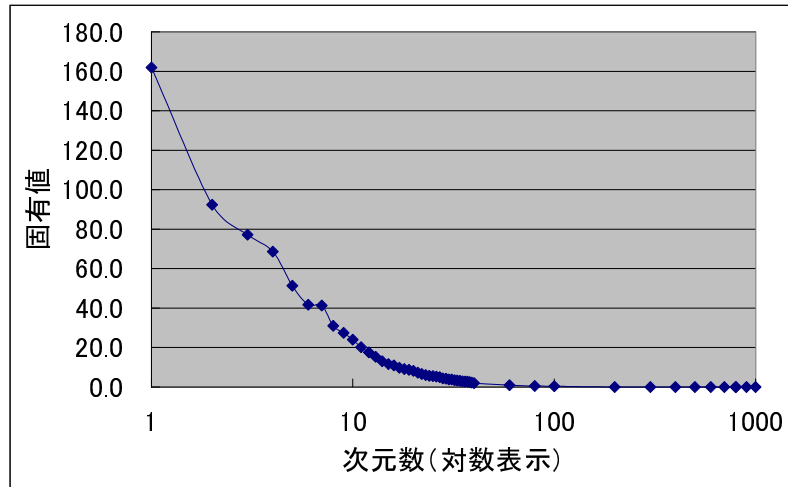


図 20 主成分分析における固有値の計算結果

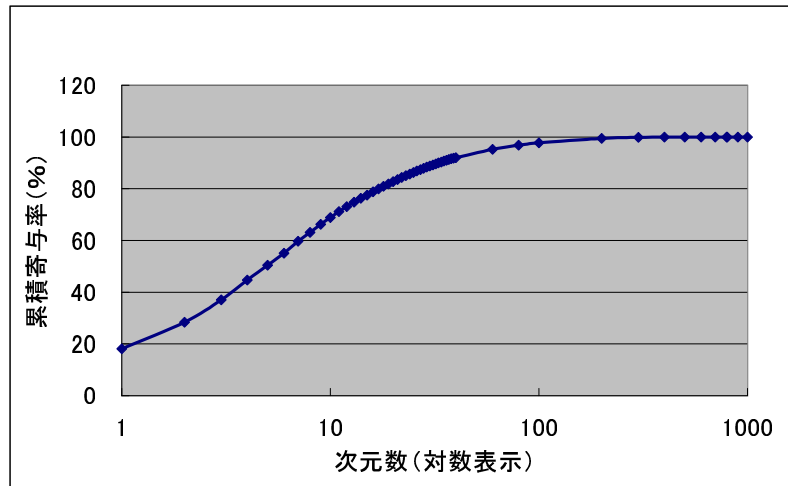


図 21 主成分分析における累積寄与率の計算結果

表 3 主成分分析による次元数削減した特徴ベクトルに対する認識実験の結果

特徴ベクトルの次元数	認識率(%)
1024	87.3
900	87.6
794	87.6
768	87.0
700	88.1
650	87.5
600	87.3
512	87.6
400	86.3
300	85.3
256	83.8
100	81.0
64	79.1

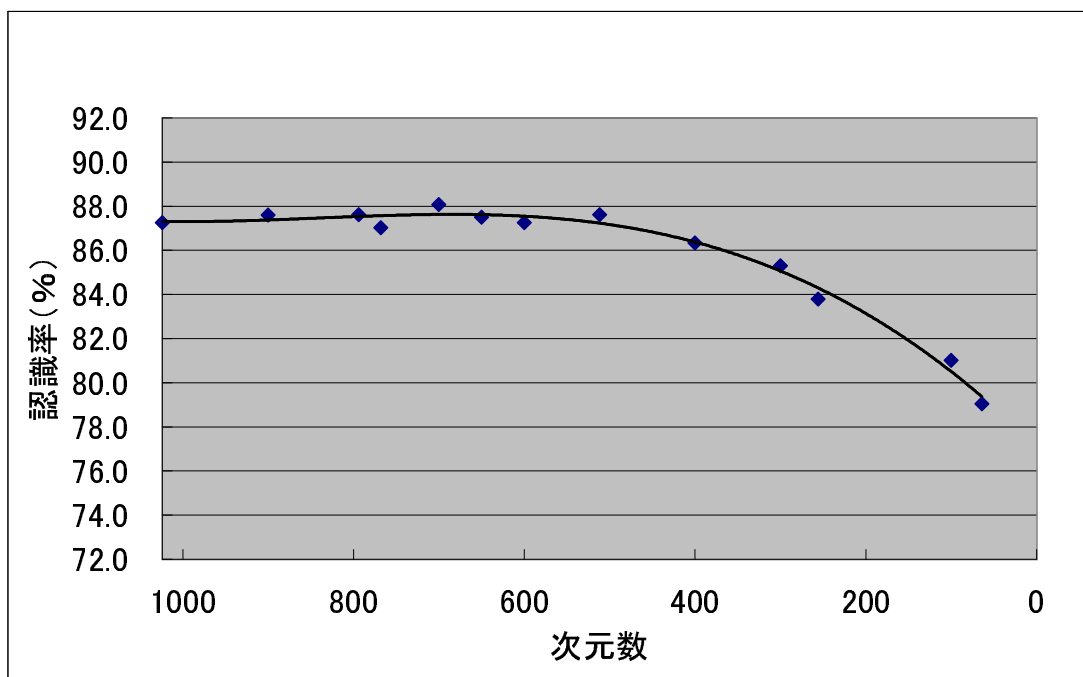


図 22 主成分分析による次元数削減した特徴ベクトルに対する認識実験の結果

8. むすび

本論文では、多値の文字画像からの直接特徴を得る特徴抽出法を提案した。本手法は、多値文字画像に対してエッジ方向成分射影法を局所領域内で適用することにより、低解像画像に対しても有効で、位置ずれに頑健な特徴抽出を可能にする。

本手法の映像内文字に対する有効性を検証するため、映画7本のエンドロールから手動で切り出した文字865文字に対して文字認識実験を行った。これにより以下のことが確認できた。

- 多値画像からの認識と2値画像からの認識を比較実験し、平均認識率が2値画像認識で77.0%に対し、多値画像認識で84.4%という結果を得た。これにより映像内の文字認識において多値画像による認識が有効であることが確認した。
- エッジ方向成分射影法を局所領域内で適用することにより、認識率を90.8%まで高めることが可能である。
- 位置ずれがある文字画像に対しても認識実験を行い、累積方法を変更することによって文字幅の5%程度までは位置ずれに頑健にできることを確認した。
- 本手法で得られた特徴ベクトルに対して主成分分析による次元削減を行い、認識率を低下させずに特徴ベクトル次元数を半分にすることが可能であることを確認した。

これにより本手法を映像内の文字認識の特徴抽出に適用して動画像のインデクシング等に応用が期待される。

謝辞

本研究は、本学および松下電器産業（株）の協力により開設された連携講座：人間・ネットワーク系学講座（二矢田 勝行 教授，栄藤 稔 助教授，松下電器産業（株）先端技術研究所）で行われたものです。筆者は基幹講座として本学のソフトウェア基礎講座にも所属し終始暖かい御指導・御助言を頂きました。

本研究を進めるにあたり、主指導教官として様々な御指導・御助言を頂きましたソフトウェア基礎講座 横矢 直和 教授 に深く感謝いたします。

非常に恵まれた環境で研究する機会を与えてくださり、研究において御助言を頂きました人間・ネットワーク系学講座（松下電器産業（株）先端技術研究所）二矢田 勝行 教授に厚く御礼申し上げます。

また副指導教官として適切な御指導・御助言を頂きました像情報処理講座 千原 國宏 教授 に深く感謝致します。

本研究への貴重な助言や指針を与えて頂き、研究に対する姿勢や論文の書き方を懇切丁寧に御指導頂いた人間・ネットワーク系学講座（松下技研（株）情報ネットワーク研究所）栄藤 稔 助教授に深く感謝の意を表します。

本研究において様々な御助言・御指導を頂いたソフトウェア基礎講座 竹村 治雄 助教授，岩佐 英彦 助手，山澤 一誠 助手に深く感謝致します。

日々の研究の中で常に適切な研究方針を示していただき、基礎から懇切丁寧に御指導頂きました松下電器産業（株）先端技術研究所 近藤 堅司 氏に深く感謝の意を表します。

本研究を進めるにあたり暖かく多大な御助言・御指導を頂きました松下電器産業（株）先端技術研究所 目方 強司 氏，今川 太郎 氏，松川 善彦 氏，芳澤 伸一 氏に深く感謝いたします。

本研究において、プログラム作成，論文執筆の基礎を御指導頂きました松下電器産業（株）宗續 敏彦 氏，畑 幸一 氏，野口 栄治 氏，青木 利道 氏に深く感謝いたします。

本研究を進めるにあたり多大な御指導・御助力を頂いたソフトウェア基礎講座の佐藤 哲 氏，神原 誠之 氏，松宮 雅敏 氏に深く感謝致します。

日々の研究室活動を支えていただいた，福永 博美 女史に深く感謝致します。

この2年間互いに励まし合い、有意義な研究生活を共に過ごすことのできたソフトウェア基礎講座の戎野 聡一 氏，島村 潤 氏，林 一成 氏，藤井 博文 氏，町田 貴史 氏，松川 昌洋 氏，安居 亮 氏，山下 敏行 氏，山口 晃一郎 氏，吉森 勇人 氏に深く感謝致します。

本研究を進めるにあたり様々な面で支えていただいたソフトウェア基礎講座のOBの皆様，M1 諸氏に深く感謝いたします。

参考文献

- [1] R. Lienhart : “Automatic text recognition for video indexing”, Proc. ACM Multimedia 96, pp. 11-20, 1996.
- [2] L. Wang and T. Pavlidis: “Direct gray-scale extraction of features for character recognition”, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 15, No. 10, pp. 1053-1067, 1993.
- [3] J. C. Pettier and J. Camillerapp: “Script representation by a generalized skeleton”, Proc. 2nd. International Conference on Document Analysis and Representation, pp. 850-853, 1993.
- [4] S. W. Lee and Y. J. Kim: “Direct extraction of topographic features for gray scale character recognition”, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 17, No. 7, pp. 724-729, July, 1995.
- [5] H. Nishida: “ Boundary extraction from gray-scale document images based on surface data structures”, Graphical Models Image Process. , Vol. 60, No. 1, Jan. , pp. 35-45, 1998.
- [6] G. Srikantan, S.W. Lam, and S.N. Srihari: “Gradient-based contour encoding for character recognition”, Pattern Recognition, Vol. 29, No. 7, pp. 1147-1160, 1996.
- [7] D.Gabor :“ Theory of communication ”, J. Institute of Elec. Eng. , Vol. 93, pp. 429-457, 1946.
- [8] W. Freeman, E. Adelson: “ Steerable filters for early vision ” Image Analysis and Wavelet Decomposition, pp.406-415, Proc. 3rd International Conference on Computer Vision, 1990.
- [9] 山田敬嗣, 津雲淳 :“Gabor 特徴の安定性考察と文字認識”, 信学技報, PRU92-112, pp. 75-82, Jan. , 1993.

- [10] 浜本義彦, 政水克典, 内村俊二, 富田眞吾 : “手書き漢字認識のための Gabor 特徴”, 信学論 (D-II), Vol. J79-D-II, No. 2, pp. 202-209, Feb. , 1996.
- [11] 渡辺政典, 安田哲也, 浜本義彦, 富田眞吾 : “Gabor 特徴を用いた手書き数字認識”, 信学技報, PRMU96-27, Jun. , 1996.
- [12] 栄藤稔 : “パターン位置決め装置とパターン類別装置”, 日本国特許公開 H04-329488, 1991.
- [13] 吉村浩至, 近藤堅司, 栄藤稔, 横矢直和 : “字幕パターン認識におけるエッジ方向成分射影法の有効性検討”, 情報処理学会第 59 回全国大会論文集(2), p. 215, 1999.
- [14] 萩田紀博, 内藤誠一郎, 増田功 : “外郭方向寄与度特徴による手書き漢字の識別”, 信学論 (D), Vol. J66-D, no. 10, pp. 1185-1192, Oct. , 1983.
- [15] 鶴岡真治, 栗田昌徳, 原田智夫, 木村文隆, 三宅康二 : “加重方向指数ヒストグラム法による手書き漢字・ひらがな認識”, 信学論 (D), Vol. J77-D, No. 7, pp. 1390-1397, July, 1987.
- [16] 孫寧, 田原透, 阿曾弘具, 木村正行 : “方向線素特徴量を用いた高精度文字認識”, 信学論 (D-II), Vol. J74-D-II, No. 3, pp. 330-339, March, 1991.
- [17] J.G.Daugman : “Complete discrete 2-D gabor transforms by neural networks for image analysis and compression”, IEEE Trans. on Acoustics Speech and Signal Processing, Vol. 36, no. 7, pp. 1169-1179, July, 1988.
- [18] 大津展之 : “判別および最小 2 乗規準に基づく自動しきい値選定法”, 信学論, Vol. J63-D No. 4, 1980.
- [19] 小島良宏, 香田敏行, 丸野進, 坂上茂生, 山本浩司, 木泰治 : “構造化ニューラルネットワーク (NARA) によるマルチフォント漢字認識”, 信学技報, NC91-31, pp. 9-16, 1991.

- [20] T.Kohonen : “Learning vector quantization for pattern recognition”, Helsinki University of Technology, Report TKK-F-A601, Nov. , 1986.
- [21] 若林哲史, 鶴岡信治, 木村文隆, 三宅康二 : “手書き数字認識における特徴選択に関する考察”, 信学論 (D-II), Vol. J78-D-II, No. 11, pp. 1627-1638, Nov. , 1995.
- [22] 高根芳雄, 柳井晴夫 : “現代人の統計 2 多変量解析法”, 朝倉書店
- [23] K.Fukunaga : “Introduction to Statistical Pattern Recognition, Second Edition”, Academic Press, 1990.