

## 3次元回転不変文字認識

成田 了<sup>†</sup> 大山 航<sup>†</sup> 若林 哲史<sup>†</sup> 木村 文隆<sup>†</sup>

<sup>†</sup> 三重大学大学院工学研究科 〒 514-8507 三重県津市栗真町屋町 1577

E-mail: †{narita,ohyama,waka,kimura}@hi.info.mie-u.ac.jp

あらまし 本論文では、カメラで撮影された文字を認識するための新たな文字認識手法を提案する。提案手法では、あらかじめ計算機上で3次元の回転処理を施して生成した学習サンプルで識別器を構成し、文字の認識を行う。印刷英数字(全62字種、各字種約600サンプル)を実験データとして認識性能の評価実験を行った結果、回転文字を学習させても、回転していない文字の認識にさほど悪影響を与えることなく高精度な認識が実現できることを確認した。また、実際にカメラで撮影した文字に対しても高い認識精度を得ることができ、文字単位であれば3次元の回転を学習することで透視投影の影響を受けて変形した文字の認識にもほぼ対応できることを確認した。さらに、文字情報から平面の法線ベクトルを検出するといったシーン解析への応用を検討した。

キーワード カメラベース文字認識, 回転不変, 加重方向指数ヒストグラム, MQDF, シーン解析

### 1. はじめに

過去10年で、デジタルカメラの性能は大きく向上した。最近では、携帯電話のカメラでさえ1300万画素を超えるものが登場している。画素数増加の他にも、手ぶれ補正や顔認識などの周辺技術が搭載されるようになり、誰もが手軽に美しい写真を撮影できるようになった。このような状況で、デジタルカメラで撮影した画像中の文字を自在に認識することができれば、標識や看板の情報を利用したナビゲーションシステムや博物館の展示物説明文の翻訳システムなど様々な応用例が考えられるため、デジタルカメラを入力手段とした文字認識への期待が高まっている。

しかし、デジタルカメラで撮影した画像中の文字認識には、解決しなければならない問題が数多く存在する[1][2]。その一つとして、射影変換による文字の変形がある。カメラの視軸と文字平面との角度・位置関係によって、撮影された画像は、回転、せん断、透視投影の影響を受ける。また、最初から文字が回転した状態で記述されている場合も多いため、カメラで撮影した文字と学習サンプルとを単純に照合して認識を行うことはできず、何らかの方法で傾きや回転を補正してからマッチングする手法が用いられるのが一般的である。更に、デジタルカメラの利便性を損なわずに実用化するためには、処理時間を考慮に入れた単純かつ効果的な文字認識手法が必要である。

これまでも、デジタルカメラで撮影した文字を認識対象とした研究がいくつか行われている。その一つに、撮影した画像から文字行を抽出し、行全体の傾きを補正してから文字を認識する手法[3]がある。しかし、この手法では、文字が行に沿って印刷されていない場合や回転している場合には、認識できないことがある。一方で、

回転や射影変換による変形に強い手法として、アフィン不変な特徴[4]が提案され、レイアウトに依存せず実時間で認識可能な手法[5]も提案されている。

本研究では、スキャナで取得した回転や歪みのない文字に対して3次元回転処理を施すことで人工的に生成された回転文字を学習サンプルとして識別器を構成し、デジタルカメラで撮影された画像中の文字を認識する、シンプルかつ効果的な文字認識手法を提案する。提案手法では、認識時に入力文字に対する回転などの補正を行わないため、変形のない文字を認識する場合に比べて処理時間を増加させることなく認識を行える。

あらかじめ回転させた文字を学習する手法には、パラメトリック固有空間法[6]を応用した長谷らによる回転文字認識手法[7]がある。しかし、長谷らの手法では平面上での回転のみ考慮しており、3次元空間中の回転を考慮し、射影変換による歪みにも頑強な本手法とは異なる。

### 2. 提案手法

本節では、本研究で提案する文字認識手法について述べる。提案手法は回転文字生成、特徴抽出、辞書作成、認識の大きく四つの処理から成る。

はじめに、標準文字に対して計算機上で3次元の回転処理を施し、2次元平面に平行投影した回転文字を生成する。そして得られた回転文字から特徴抽出を行い、文字のクラスごとに学習辞書を作成する。一方で、認識対象の文字に対しては回転などの補正は行わずに特徴抽出を行い、得られた特徴ベクトルを回転文字を学習した辞書と照合することで認識を行う。ここで、標準文字とは図1(a)に示すような正面から撮影した回転や歪みのない文字のことを言い、回転文字とは図1(b)に示すような回転のある文字のことを言う。以降、各処理内容の詳細について順に説明していく。



図1 (a)標準文字と(b)回転文字.

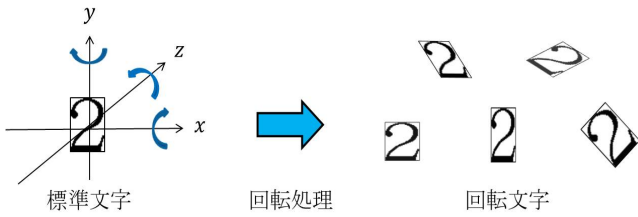


図2 回転文字生成の流れ.

## 2.1 回転文字生成

図2のように標準文字を3次元直交座標系(左手系)の $xy$ 平面に置いて外接枠の中心を原点に対応させ、計算機上で3次元の回転処理を施して回転文字を生成する. 回転は $x$ 軸まわり,  $y$ 軸まわり,  $z$ 軸まわりの順に行うものとする. 図2中の文字には, 変形がわかりやすいように外接枠を表示してある. また, 回転文字生成を行う際には, 画像の回転に伴うノイズを低減するために画素値の線形補間を施している.

## 2.2 特徴抽出

本研究では特徴ベクトルとして加重方向指数ヒストグラム(392次元)を使用する[8]. 特徴抽出の手順は次の通りである.

(1) 入力として与えられた文字に対して8近傍で境界追跡を行いチェーンコードを求め, 隣接する2つのベクトル和から向きの違いを含めて16方向に量子化を行う.

(2) 文字の外接枠を169個(縦13×横13)の小領域に分割し, 領域別, 方向別に輪郭画素数を集計し, 2704次元(縦13×横13×16方向)のヒストグラムを得る.

(3) 2次元ガウスフィルタを縦横1領域おきに施し, 領域数を13×13から7×7に削減する. 同様に[1 2 1]の加重フィルタを1方向おきに施し, 方向量子化数を16方向から8方向に削減し, 392次元(縦7×横7×8方向)の特徴ベクトル(加重方向指数ヒストグラム)を得る.

(4) 文字外接枠の縦横サイズの大きい方で特徴量を正規化する.

(5) 変数変換( $y = x^{0.5}$ )により, 特徴ベクトルの分布を正規分布に近づける.

## 2.3 辞書作成

特徴抽出によって得られた特徴ベクトルからクラスごとに平均ベクトル, 共分散行列の固有値, 固有ベクトルを算出し, それらを学習辞書の中身とする.



図3 実験データとして使用したアルファベット“A”の一例.

## 2.4 認識

本研究では識別関数として式(1)で表されるMQDF(Modified Quadratic Discriminant Function)を使用する[9]. この識別関数は, 分布パラメータのうち母集団の共分散行列が未知の正規分布に対する最適識別関数[10]から導出された近似式で, 識別精度を損うことなく計算量を大幅に削減できるのが特徴である.

$$g(X) = \frac{1}{\alpha\sigma^2} \left[ \|X - M\|^2 - \sum_{i=1}^k \frac{(1-\alpha)\lambda_i}{(1-\alpha)\lambda_i + \alpha\sigma^2} \{\Phi_i^T(X - M)\}^2 \right] + \sum_{i=1}^k \ln(1-\alpha)\lambda_i + \alpha\sigma^2 \quad (1)$$

ここで,  $X$ は入力文字の $n$ 次元特徴ベクトル,  $M$ は母集団の平均ベクトル,  $\Phi_i, \lambda_i$ はそれぞれ標本共分散行列の第 $i$ 固有ベクトルと第 $i$ 固有値,  $k$ は識別に用いる固有ベクトル数である. 式中の $\sigma^2$ は特徴ベクトル $X$ の事前確率分布を球状と仮定した場合の分散であり,  $\alpha$ は $\sigma^2$ の信頼度を表す定数で信頼度定数と呼ぶべきものである.

実験では,  $M$ は標本の平均ベクトルで代用し,  $\sigma^2$ の値としては, 全字種, 全固有値の平均を用いる.  $g(X)$ が最小となるクラスが認識結果となり, 計算時間と記憶容量のオーダーは $O(kn)$ となる.

## 3. 認識性能の評価実験

提案手法の有効性を確認するために, 標準文字のみを学習させた場合と回転文字(標準文字を含む)を学習させた場合の比較実験を行った. 以下, 実験1では文字画像サイズを正規化せずに, 実験2では3種類の正規化手法で文字画像サイズを正規化して認識実験を行い, 認識率がどう変化するかを調査した.

### 3.1 実験データ

実験データには図3のように様々な大きさや形をした印刷英数字(全62字種, 各字種約600サンプル)を使用し, 通し番号が偶数の約300サンプルを学習用データ, 奇

表 1 実験 1 の認識率 (%) .

(a) すべての字種を区別したとき .

学習 \ 評価	回転あり	回転なし
回転あり	92.11	92.86
回転なし	48.21	94.80

(b) アルファベットの大文字・小文字を区別しないとき .

学習 \ 評価	回転あり	回転なし
回転あり	97.15	97.48
回転なし	54.51	98.29

(c) 類似文字を同一字種としたとき .

学習 \ 評価	回転あり	回転なし
回転あり	99.18	99.35
回転なし	56.54	99.80

数の約 300 サンプルを評価用データとした . 回転角度は  $x$  軸と  $y$  軸に関しては  $-45^\circ \sim 45^\circ$  ,  $z$  軸に関しては  $-30^\circ \sim 30^\circ$  の範囲でそれぞれ  $15^\circ$  間隔で変化させた . そのため , 一字種あたり約 147,000 サンプル ( $600 \times 7 \times 7 \times 5 = 147,000$ ) 生成される .

### 3.2 実験 1

実験 1 では , 文字画像サイズの正規化を行わずに認識実験を行った .

実験結果を表 1(a) に示す . 標準文字のみを学習した辞書で回転文字を認識したところ 48.21 % であった認識率が , 提案手法により 92.11 % に向上した . また , 標準文字の認識に関しても提案手法では 92.86 % という認識率が得られ , 標準文字のみ学習させた場合の 94.80 % に比べて低下はわずかであり , 回転文字を学習させることが標準文字の認識にさほど悪影響を与えないことがこの実験で確認できた . ここで , 誤認識した文字の内訳を調べたところ , C と c , S と s など , アルファベットの大文字・小文字が類似した字種の間違いが多く見られた . そこで , アルファベットの大文字・小文字を同一クラスとして認識率を算出したところ , 表 1(b) のような結果が得られた . さらに , l と I (アイ) と l (エル) , 0 と O (オー) と o (オー) の誤認識に関しても字体によっては人が見ても正しく区別できないほど類似していたため , これらも同一クラスとして認識率を算出した . その結果を表 1(c) に示す . これにより , 提案手法では入力として回転文字が与えられた場合でも 99 % を超える高い認識率が得られ , 標準文字の認識率低下もわずかであることがわかった .

### 3.3 実験 2

実験 1 では文字画像サイズの正規化は行わずに文字外接枠の縦横サイズの大きいほうで特徴ベクトルの要素を除算して実験を行ったが , 実験データとして使用してい



図 4 各正規化手法で正規化したアルファベット “F” の例 .

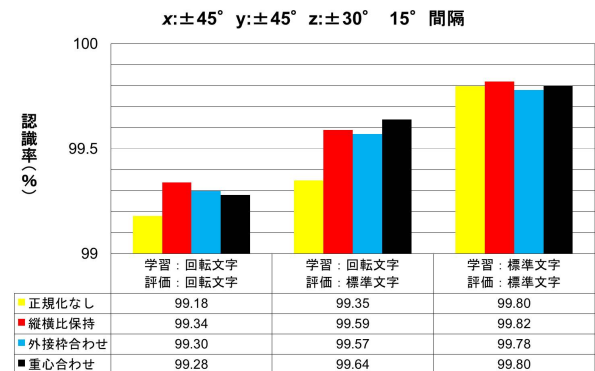


図 5 各正規化手法による認識率の比較 .

る文字画像の大きさは一辺の長さが 1 から 135 ピクセルとまちまちであったため , 文字画像のサイズを統一することで認識率が向上するのではないかと考えた . そこで実験 2 では学習用・評価用の両方の実験データに対して , (1) 縦横比保持 (2) 外接枠合わせ (3) 外接枠・重心合わせ (文字の重心を正規化後の外接枠の中心に合わせる) の 3 種類の正規化手法で文字画像サイズを正規化して認識実験を行い , 認識率を比較した . 各正規化手法で正規化した文字の例を図 4 に示す .

実験結果を図 5 に示す . 表 1(c) と同じく , 類似文字に関しては同一クラスとして認識率を算出した . 文字画像を正規化したほうがしない場合に比べて認識率が高く , 特に回転文字の認識においては縦横比を保持して正規化した場合に一番高い認識率 99.34 % が得られた . そのため , 以降は学習用データに関しては回転処理の前後に , 評価用データに関しては特徴抽出を行う前に縦横比を保持した正規化を行うものとする . 正規化サイズは  $52 \times 52$  とする . 52 は特徴抽出の際に文字の外接枠を分割するブロック数 13 の倍数となるように定めたものであり , 実験データの平均サイズに近い値でもある .

図 6 に各正規化手法における誤認識の内訳を示す . 画像の正規化を行っても , l と I (アイ) と l (エル) , 0 と O (オー) と o (オー) の誤認識数は減少せず , 誤認識に占める割合が増えていることがわかる . これらの字種は文字単位で識別することが困難であり , 画像正規化を行っても誤認識は減少しないが , 他の字種については誤認識が減少し , 結果として全体の認識率が向上したものと考えられる . また , 「その他」に分類されている主な誤認識の内訳を調べたところ , g を 9 と誤認識した数が 2077 個と一番多く , 次いで a を q と誤認識した数が 1879 個 , j を i と誤認識した数が 1830 個といった結果であった .

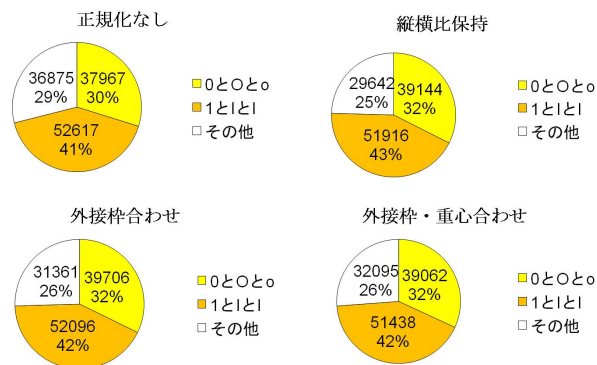


図 6 各正規化手法における誤認識の内訳（上段：文字数，下段：割合）。

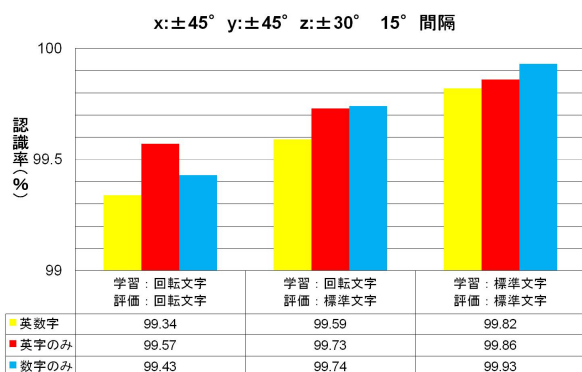


図 7 縦横比を保持した場合の認識率のまとめ。

参考までに縦横比を保持した正規化を行った場合に、英字のみ、数字のみを実験データとしたときの認識率を算出したものを図 7 に示す。実験データを英字のみとした際の認識率が一番高く、次いで数字のみ、英数字という結果であった。

#### 4. 実画像を用いた自動認識実験

ここまで、標準文字に対して 3 次元の回転処理を施して人工的に生成した回転文字を実験データとして認識性能の評価実験を行い、高い認識率が得られることを確認した。次に、提案手法に基づいて作成した学習辞書を用いて、実際に携帯電話のカメラやデジタルカメラで撮影した英数字画像から文字の検出と認識を自動で行うシステムを作成し認識実験を行った。

本システムは文字候補となる連結成分切出部と切り出された連結成分から認識対象の文字を検出し、文字種の認識までを行う文字検出部から成る。

##### 4.1 実験データ

携帯電話のカメラやデジタルカメラを用いて 50 枚の英数字画像を撮影し、実験データとした。このとき、普段カメラで対象物を撮影するときと同様に、解像度やズームの調整は固定せずに画像毎に変化させて撮影を行った。画像中には認識対象となる文字が合計で 437 文字含まれており、歪み補正などは行わずにそのまま実験データと



図 8 携帯電話のカメラやデジタルカメラで撮影した英数字画像例。

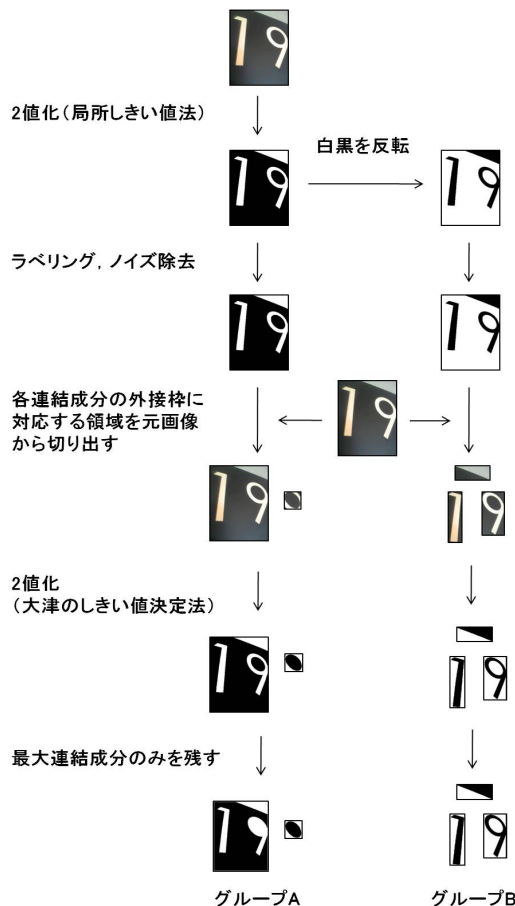


図 9 連結成分の切り出しの流れ。

して使用した。実際に撮影した英数字画像を図 8 に示す。ただし、学習した範囲を超えて回転している文字を含んだ画像やノイズにより文字が劣化している画像、過度に装飾された文字を含んだ画像は実験データとして使用しないこととした。また、 $i$  と  $j$  は二つの連結成分から成り、切り出す際に特別な処理が必要となるため、今回の実験では文字として扱わないこととした。

##### 4.2 連結成分切出部

連結成分切出部では図 8 に示すような英数字画像から、文字候補となる連結成分の切り出しを行う。連結成分を切り出すまでの流れを図 9 に示す。

はじめに、入力画像に対して局所しきい値法 [11] で 2 値化を行い、得られた 2 値画像に対してラベリング、ノ



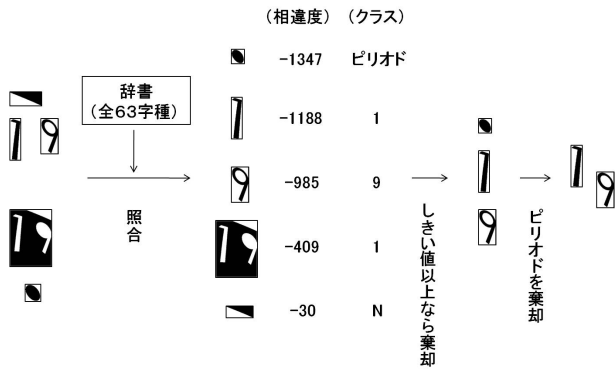


図 10 文字の検出の流れ .

イズ除去を行う。次に、各連結成分の外枠の座標値を算出し、その座標値に対応する領域を元画像から切り出す。切り出された画像に対して今度は大津のしきい値決定法で求めたしきい値で画像全体の2値化を行い、最大連結成分のみを残すことで文字候補となる連結成分の切り出しを行う。この流れで切り出した連結成分を仮にグループ A と呼ぶ。同様な流れで、今度は入力画像に対して局所しきい値法で2値化を行った後に白画素と黒画素を反転させて連結成分の切り出しを行う。この流れで切り出した連結成分を仮にグループ B と呼ぶ。これにより文字でない連結成分も多く残るが、白地黒地判定失敗による文字連結成分の切り出しの失敗を無くしている。

実際にこの流れで英数字画像 50 枚から連結成分を切り出したところ、437 個の文字連結成分を含む 1026 個の連結成分が切り出された。

### 4.3 文字検出部

この段階では切り出された連結成分の中に文字でない連結成分が多数含まれているため、文字の検出を行う。文字の検出の流れを図 10 に示す。

各連結成分を対象にして、学習辞書（英数字、全 62 字種）に新たにピリオドを加えた辞書を用いた認識を行い、相違度を記憶しておく。この相違度は文字らしさを表しており、値が小さいほど文字の可能性が高いことを示している。この値がしきい値より大きい場合には文字でないと判断し棄却する。実験ではしきい値を（最小相違度 + 最大相違度）÷ 2 に設定し、相違度がしきい値以上の連結成分を棄却した。辞書にピリオドを学習させる理由は、小さなノイズをピリオドに分類し、ピリオドを棄却することでノイズ除去を容易にするためである。i と j についてもピリオドと認識された連結成分は棄却されるが、今回は i と j を文字として扱わないこととしているので問題ない。最後に、連結成分の切り出しの際に仮にグループ A とグループ B と呼んだ各グループに含まれる連結成分の数をカウントし、数が多いグループに属する連結成分のみを残すようにした。

これにより英数字画像 50 枚中に存在する 437 個の文字を全て検出し文字種を正しく認識できたが、一方で

表 2 検出および認識結果 (個数) .

		分類	
		文字	非文字
正解	文字	437	0
	非文字	51	538



図 11 方向の情報も必要な標識の例 .

文字でない連結成分の誤検出数は 51 個であった。これらの結果をまとめたものを表 2 に示す。そこから適合率 (precision) と再現率 (recall) を算出すると、それぞれ

$$precision = \frac{437}{437 + 0} \times 100 = 100(\%) \quad (2)$$

$$recall = \frac{437}{437 + 51} \times 100 = 89.5(\%) \quad (3)$$

となる。文字だけを正しく検出し認識することはできなかったが、一枚の画像から検出される文字でない連結成分の数を平均で一つに抑えることができた。処理時間は、Intel Xeon の 2.67GHz の CPU を用いて 1 コアで処理をさせた場合、1026 個の連結成分に対して合計 37.748 秒かかり、1 秒当たり約 27 個の連結成分を処理できる性能であった。

実験の結果、文字単位であれば 3 次元空間中で回転させて 2 次元平面に平行投影した文字を学習しておくだけで、透視投影の影響を受けて変形した文字の認識にもほぼ対応できることがわかった。今回の実験では i と j は切り出しの際に特別な処理が必要となるため認識対象外としたが、実際にカメラで撮影したいいくつかの英数字画像から手動で切り出した 8 個の i と 3 個の j を評価用データとして認識実験を行った結果、すべて正しく認識できることを確認した。

### 5. シーン解析への応用

前節において、学習したフォント・回転角度の範囲内であれば、デジタルカメラで撮影した文字に関しても高い認識精度が得られることを確認した。そこで提案手法を応用し、カメラで撮影した文字をただ認識するだけでなく、文字の回転角度まで推定する取り組みを行った。これにより文字情報から画像中の平面の法線ベクトルを検出することが可能となれば、図 11 のような標識が示す方向の情報を得たり、文字情報から物体の形状を復元したりするといった応用が考えられる。

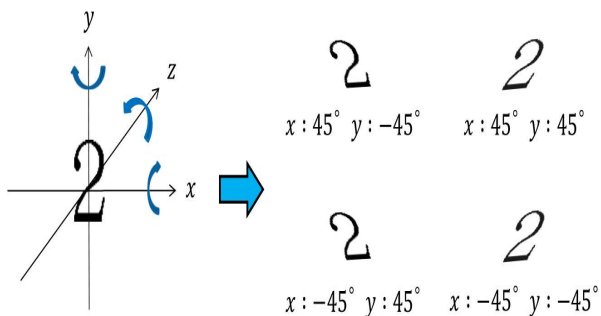


図 12  $x$  軸と  $y$  軸まわりに  $\pm 45^\circ$  回転させた文字の例 .

### 5.1 回転角度推定手法と推定成功率

提案手法を応用することで、どの程度文字の回転角度を推定できるかを確認する実験を行った。回転角度を推定する流れは提案手法と同様であり、まず文字種を特定し、各字種の回転角度毎に別々に作成した学習辞書とマッチングすることで回転角度の推定を行う。実験では第 3 節の認識性能の評価実験と同じ実験データを使用し、回転角度についても同じ範囲・間隔で変化させた。そのため、字種毎に 245 個 ( $7 \times 7 \times 5 = 245$ ) の学習辞書が生成される。特徴ベクトルについても同様に、392 次元の加重方向指数ヒストグラムを使用した。識別関数に関してはカテゴリ増加に伴う認識時間の増大を抑えることを視野に入れ、MQDF の他に計算時間のオーダーが  $O(n)$  であるユークリッド距離と線形識別関数も用いて実験を行い、識別関数毎に回転角度の推定成功率を算出した。

実験の結果、ユークリッド距離では 58.01 %、線形識別関数では 78.90 %、MQDF では 84.82 % という結果が得られ、線形識別関数と MQDF においては比較的高い推定成功率が得られた。ただし、提案手法では回転後に平行投影を行っていることから  $x$  軸のみ、 $y$  軸のみの回転については文字が縦や横に縮むだけであり、図 12 のように  $x$  軸と  $y$  軸の回転角度の正負を反転させて回転させた場合には結果として同じ画像が学習されるため、これらは同一角度クラスとして推定成功率を算出した。

### 5.2 実画像中の文字の回転角度推定実験

続いて、実際にデジタルカメラで撮影した英数字画像中の文字の回転角度を推定する実験を行った。識別関数には前節で一番高い推定成功率が得られた MQDF を用いた。ただし、あらかじめカメラと撮影対象の文字の回転角度を測定して撮影するのは困難であるため、今回の実験では画像の中央に最も近い文字の回転角度を推定し、推定した回転角度をもとに画像全体の補正を行い、正面から撮影したものに近い画像が得られるかどうかで正しく回転角度が推定できているかを判断した。提案手法では  $x$  軸と  $y$  軸の回転角度の正負を反転させた二つの回転角度が推定されるため、今回は一方を手動で選ぶこととした。また、透視投影による変形を視点と文字平面まで



図 13 補正前の画像 .



図 14 回転角度を推定し、補正した画像 .



図 15 撮影した英数字画像に対して法線ベクトルを描画した結果画像 .

の距離を一定値に固定した逆変換により補正している。

実験の結果、図 13 の画像は図 14 のように補正され、全体的に安定して回転角度が推定できていることが確認された。処理時間は Intel Xeon の 2.67GHz の CPU を用いて 1 コアで処理させた場合に、1 秒当たり約 6.9 個の文字の回転角度が推定できる性能であった。

また、 $x$  軸、 $y$  軸、 $z$  軸まわりの推定回転角度より文字平面の法線ベクトルを計算することができる。元画像上に法線ベクトルを描画した結果を図 15 に示す。図中の矢印がそれぞれの文字に対して算出された法線ベクトルである。この結果より、文字情報から文字平面を検出するという応用が実現可能であるということがわかった。

しかし一方で、現在は  $15^\circ$  間隔で回転角度を推定しているため十分な推定精度があるとは言えず、回転角度の推定精度をより高めていく必要がある。そのための対策として、より小さな回転角度間隔で生成した学習データによる識別器の構成、各回転角度クラス間の補間や複数の文字における推定結果の利用が有効であると考えている。さらに、処理の高速化のために、MQDF で用いる固有ベクトル数の調整や、ユークリッド距離による大分類導入を考えている。

## 6. ま と め

本研究では、標準文字に対し 3 次元の回転処理を施して人工的に生成した回転文字で識別器を構成し、カメラで撮影された文字を認識するというシンプルかつ効果的

な文字認識手法を提案した。

印刷英数字を認識対象として提案手法による認識性能の評価実験を行った結果、入力として回転文字が与えられたとき、標準文字のみを学習させた場合には 55.20%であった認識率が回転文字を学習させることで 99.34%に向上した。一方で、標準文字に対する認識率は 99.82%が 99.59%になる程度であった。また、評価用データとして携帯電話のカメラやデジタルカメラで撮影した英数字画像 50 枚中の 437 文字を用いた実験では、全ての文字を正しく認識でき、提案手法の有効性を確認できた。

実験を通じて、

(1) 回転文字を学習させることが回転していない文字の認識にさほど悪影響を与えない、

(2) 文字単位の認識においては、3次元空間中で回転させた文字を学習することで、透視投影の影響を受けて変形した文字の認識にもほぼ対応できる、

(3) 提案手法では入力として与えられた文字に対して回転などの補正を行わないため、回転文字に対して処理時間を増加させることなく認識を行える、ことを確認した。

今回の認識実験では英数字のみを認識対象としたが、本手法は他の全ての言語に対しても適用可能であると考えられる。さらに、用途を限定すれば、学習に用いるフォントや回転角度の範囲及び分解能を選んで学習させることで、認識精度をさらに高めることができると考えられる。

また、提案手法により、デジタルカメラで撮影した情景中の文字の回転角度を推定し、文字平面の法線ベクトルを求めることでシーン解析への応用の可能性を検討した。その結果、安定して回転角度を推定できることを確認し、文字情報から画像中の平面を検出するといったシーン解析への応用が可能であることがわかった。

今後の課題としては、複数の文字情報を使って回転角度の推定精度を向上させたり、ユークリッド距離による大分類を行ってから MQDF による詳細分類を行うなど、分類手法の工夫を行う。また、本研究のように文字の変動を吸収させて認識する手法 [12] もいくつか提案されているため、そういった手法との比較実験についても今後取り組んでいきたいと考えている。

## 文 献

- [1] J. Liang, D. Doermann and H.Li, "Camera-based analysis of text and documents image analysis," International Journal on Document Analysis and Recognition, vol.7, no.2-3, pp.84-104, 2005.
- [2] 黄瀬浩一, 大町真一郎, 内田誠一, 岩村雅一, "カメラを用いた文字認識・文章画像解析の現状と課題," 信学技報, PRMU2004-246, 2005.
- [3] G.K. Myers, R.C. Bolles, Q.-T. Luong, J.A. Heron, and H.B. Aradhye, "Rectification and recognition of text in 3-d scenes," International Journal on Document Analysis and Recognition, vol.7, no.2-3, pp.147-158, 2004.

- [4] 堀松晃, 丹羽亮, 岩村雅一, 黄瀬浩一, 内田誠一, 大町真一郎, "アフィン不変な文字認識手法とその高速化," 画像の認識・理解シンポジウム (MIRU2008) 論文集, IS5-10, pp.1450-1455, 2008.
- [5] M. Iwamura, T. Tsuji, K. Kise, "Memory-based Recognition of Camera-Captured Characters," Proceedings of the 9th IAPR International Workshop on Document Analysis Systems, pp.89-96, 2010.
- [6] 村瀬 洋, S.K.Nayar, "2次元照合による3次元物体認識-パラメトリック固有空間法-," 信学論, J77-D-II, 11, pp.2179-2187, 1994.
- [7] H. Hase, T. Shinokawa, M. Yoneda, C.Y. Suen, "Recognition of Rotated Characters by Eigen Space," ICDAR2003, PII-2, pp.731-735, 2003.
- [8] F. Kimura, T. Wakabayashi, S. Tsuruoka, Y. Miyake, "Improvement of Handwritten Japanese Character Recognition Using Weighted Direction Code Histogram," Pattern Recognition, vol.30, no.8, pp.1329-1337, 1997.
- [9] F. Kimura, K. Takashina, S. Tsuruoka, Y. Miyake, "Modified quadratic discriminant functions and the application to Chinese character recognition," IEEE Trans. Pattern Anal. Mach. Intell. PAMI-9(1), 149-153, 1987.
- [10] D.G. Keehn, "A note on learning for Gaussian properties," IEEE Trans. Inform. Theory, vol.IT-11, no.1, pp.126-132, 1965.
- [11] 多田光博, 大山航, 若林哲史, 木村文隆, "SVMによるナンバープレート領域抽出," 電気関係学会東海支部連合大会, O-122, 2009.
- [12] 草地良規, 鈴木章, 伊藤直己, 荒川賢一, 安野貴之, "景観画像中の文字候補群によるインデクシング及び検索技術," 信学論 (D), vol.J90-D, no.9, pp.2562-2572, 2007.