

変量と不変量と文字の正規化に基づく文書画像の射影歪み補正

丹羽 亮[†] 堀松 晃[†] 岩村 雅一[†] 黄瀬 浩一[†] 内田 誠一^{††}

大町真一郎^{†††}

[†] 大阪府立大学大学院工学研究科 〒 599-8531 堺市中区学園町 1-1

^{††} 九州大学大学院システム情報科学研究所 〒 819-0395 福岡市西区元岡 744

^{†††} 東北大学大学院工学研究科 〒 980-8579 仙台市青葉区荒巻字青葉 6-6-05

E-mail: [†]{niwa,horimatsu}@m.cs.osakafu-u.ac.jp, [†]{masa,kise}@cs.osakafu-u.ac.jp,

^{††}uchida@is.kyushu-u.ac.jp, ^{†††}machi@ecei.tohoku.ac.jp

あらまし 近年、デジタルカメラをスキャナの代わりに用いて文書画像処理を行う試みが盛んに研究されている。しかし、既存の文書画像処理技術の多くはスキャナで取得した文書画像を対象としているため、デジタルカメラで斜めから撮影した文書画像には適用できない。このような場合、文書画像に生じた射影歪みを補正し、スキャナで取得したかのような真正面の文書画像を得る必要がある。我々は既に変量と不変量に基づく射影歪みからアフィン歪みへの補正を提案している。これは、デジタルカメラの種類や文書の枠の有無、文字行の平行性を問題にしない補正へのアプローチである。本稿では手法の改良による精度向上、さらに書体情報を学習することで正対画像の復元の実現を実験結果とともに報告する。

キーワード 文書画像, 変量, 不変量, 正規化, 射影歪み, アフィン歪み

Rectification of Distorted Document Images Using Variants, Invariants and Normalization

Ryo NIWA[†], Akira HORIMATSU[†], Masakazu IWAMURA[†], Koichi KISE[†], Seiichi UCHIDA^{††},
and Shinichiro OMACHI^{†††}

[†] Graduate School of Engineering, Osaka Prefecture University
1-1 Gakuencho, Naka, Sakai, 599-8531 Japan

^{††} Faculty of Information Science and Electrical Engineering, Kyushu University
744 Motoooka, Nishi, Fukuoka, 819-0395 Japan

^{†††} Graduate School of Engineering, Tohoku University
6-6-05 Aoba, Aramaki, Aoba, Sendai, 980-8579 Japan

E-mail: [†]{niwa,horimatsu}@m.cs.osakafu-u.ac.jp, [†]{masa,kise}@cs.osakafu-u.ac.jp,

^{††}uchida@is.kyushu-u.ac.jp, ^{†††}machi@ecei.tohoku.ac.jp

Abstract For user convenience, document image processing captured with a digital camera instead of a scanner has been researched. However, existing methods of document image processing are not usable for a perspective distorted document image because most of them require a neat image captured by a scanner. Thus, we have to rectify perspective distortion of document images. For the purpose, we have proposed a layout free rectification method based on *variants* and *invariants*. In this paper, we reports two improvements on the method. The first one is to increase the stability of the method. The second one is to eliminate remaining affine distortions using prior knowledge. We confirm the effectiveness of the improvements by experiments.

Key words document image, variants, invariants, normalization, perspective distortion, affine distortion

1. はじめに

近年、文字認識や文書画像処理にデジタルカメラを用いる試みが盛んに行われている [1], [2]. これは、文書画像の取得にスキャナを用いず、代わりにデジタルカメラを用いるものである。デジタルカメラを使用する利点は文書画像取得の可搬性や簡便性である。スキャナとは違い、デジタルカメラは手軽に持ち歩くことができるため、思い立ったときに手軽に撮影することができる。また、ポスターや看板など、大きな物や移動できない物を撮影することができるなど、これまでにない新しい使用形態に発展する可能性を秘めている。

このようにデジタルカメラを用いた文書画像処理は優れたアプリケーションを生み出す可能性がある反面、その実現は容易ではない。その理由は、既存の文書画像処理技術の多くがスキャナで取得した文書画像を対象としていることである。つまり、デジタルカメラで取得した文書画像に既存の技術を適用するためには、デジタルカメラで取得した文書画像を補正し、スキャナで取得したかのような画像を得る必要がある。本稿では、デジタルカメラを用いることによって生じる文書画像の劣化のうち、射影歪みに焦点を当て、文書画像の射影歪み補正を行う。

この問題に対して我々は文献 [3] において、変量と不変量を用いて射影歪みからアフィン歪みに補正する手法を提案している。この手法の特徴は、文書の傾きに応じて値が変化する「変量」と変化しない「不変量」を組み合わせることである。変量と不変量はレイアウトに関係なく求めることができるため、文書の枠が長方形になるように変換を施す手法 [4] や、文字行の平行性を利用する手法 [4] のように、レイアウトの制約を受けない。また、複数のカメラ [5] もしくは動画 [6] を用いる手法のようにステレオを使用しないため、1枚の静止画のみを用いて補正することができる。

本稿ではこの手法に2つの改良を加える。1つ目は不変量の改良による安定性の向上である。文献 [3] では複数の不変量を要素とする「不変量ベクトル」を用いて性能向上を図っている。しかし、文字が少ない場合には性能が落ちるという問題がある。そこで本稿では、文献 [7] で用いられている「DE-パラメータ」を応用して、その制約をなくす方法を提案する。

2つ目は射影歪みを完全に補正することである。文献 [8] にも示されているように、アフィン変換を受けた文書画像を事前知識無しで完全に復元することはできない。そのため、文献 [3] では射影歪みをアフィン歪みに復元するに留まっていた。本稿では、対象に関する知識として書体情報の学習を導入し、射影歪みを完全に補正する。具体的には、文献 [9] のアフィン不変な文字認識手法を利用して、文字認識と同時に文書画像のアフィン歪みのパラメータを推定し、正面を向いた画像への変換を得る。

2. 文献 [3] の手法

2.1 黒画素の面積と奥行き

デジタルカメラで撮像した文書画像中の各文字に注目すると、本来同じ字種であっても位置によって大きさが変化している。例えば、画像中からアルファベットの a のような特定の字種だ

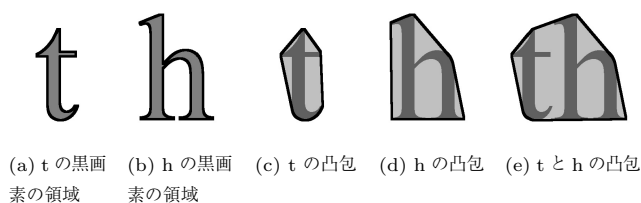


図 1 2文字から取り出せる領域

けを取り出したときを考える。このとき、文字の黒画素の面積は射影歪みによって、カメラから近い方が大きく、遠い方が小さくなる。この面積の変化から文書の奥行き情報を求めることができるので、文献 [3] ではこれを基に文書画像の補正を行っている。

面積と奥行きの関係について、遠近法としてよく知られた事実を利用する。これ以降、カメラからの光軸方向の距離を奥行きと呼ぶことにする。文字の面積は図形を置いた位置の奥行きの2乗に反比例する。画像中から特定の字種を取り出したとして、その中の j 番目の文字の奥行き Z_j は、文字の面積 s_j とカメラの焦点距離 f 、比例定数 K を用いて、

$$Z_j = fKs_j^{-\frac{1}{2}} \quad (1)$$

と表すことができる。(詳細は文献 [3] に記載している。)

2.2 面積比によるクラスタリング

2.1 で述べた、面積と奥行きの関係は同種の文字で成り立っている。さらに、本来面積が同じである字種 (フォントに依るが鏡像の文字 b と d など) の間でも成り立つ。b と d のような、同じ奥行きにあるとき面積が同じであることを以後“等価”と表現する。しかし、a と b のような等価でない字種間では成り立たない。文書中には様々な字種が混在しているため、2.1 の関係を用いて補正を行うためにはあらかじめ文字をその種類ごとに分ける必要がある。

字種を判別する方法として、文字認識が考えられるが、射影歪みを受けている場合は処理が難しい。また、ここでは字種を分類して奥行きを計算することが目的なので、文字認識のように文字にラベルを付ける必要はない。そこで文献 [3] では、アフィン変換を受けても変化しない量 (アフィン不変量) である面積比を用いた分類を行う。アフィン不変量は射影変換に対して不変ではないが、局所領域において射影変換をアフィン変換に近似できることから、文字領域のように面積の小さな領域の面積比を射影不変量のように扱うことができる。また、面積比を計算する領域の抽出には文字などの連結成分の凸包を利用する。

1文字から計算できる面積の種類は限られているため、文献 [3] では最近傍の2文字を組み合わせ、2文字から計算できる複数の面積比を用い、面積比の判別性能を向上させる。例えば t と h の2文字を選んだとき、文字の黒画素領域と凸包領域から図 1 の5種類の領域が得られる。これらを組み合わせることで、面積比を複数作り出すことができる。得られた m 個の面積比は m 次元の不変量ベクトルとして用いる。そして、以降の処理も2文字毎に行う。

クラスタリングを行うと、得られたそれぞれのクラスタには等価な文字が含まれると期待されるため、文書中の文字を不変量を用いて分類する。文献[3]ではクラスタリングに k -means 法を用いる。

2.3 平面への当てはめ

2.2のようにクラスタリングを行うと、各クラスタにおいて2.1で述べた平面への当てはめを考えることができる。紙面の傾きを高精度に推定するためには同一字種が文書中で分散していることが望ましいが、このような状況は必ずしも期待できない。そこで、1つの平面に当てはめることでクラスタ(字種)毎に推定した紙面の傾き情報の統合を考える。その際に問題になるのは同一距離に戻したときの各字種の面積の比である。それぞれの傾きが等しくなるように、式(1)の K をクラスタ毎に設定する。このとき、 K はクラスタ間で相対的に決める。従って、当てはめる平面の式 $Z = aX + bY + c$ のパラメータ a, b, c のうち、 c の値は定まらない。そこで、カメラ座標 $(X Y Z)^T$ の Z を $1/c$ 倍した Z' で置き換え、 c を消去した座標系 $(X Y Z')^T$ で平面の式を表現し

$$Z' = a'X + b'Y + 1 \quad (2)$$

とする。ただし、 $a' = a/c, b' = b/c$ である。

2.4 回転

最後に画像中の紙面を回転して、正面に向ける。これは紙面の正面に視点を移すことと等価であるので、傾いた紙面の法線ベクトルを求め、延長上に視点を移す。座標を X 軸周りと Y 軸周りで回転させ、紙面の法線を Z 軸と平行にする。しかし、現段階では c を推定していないので、暫定的に $c = 1$ とした。この場合、射影歪み補正後に、本来は長方形である図形が平行四辺形になるアフィン歪みが残る。

カメラ座標系において R によって回転させた像は、画像平面上へ射影させる。

3. 不変量ベクトルの改良

文献[3]では2つの文字を組み合わせて不変量を多次元化している。これにより判別性能を向上させているが、画像の復元に必要な文字数が膨大になるという欠点がある。例えば、 $a \sim z$ の26文字を区別する問題に対して、2文字の組合せでは、 $26 \times 26 = 676$ 通りの組から頻出度の高い文字の組合せで分類することになるからである。本稿では、文献[3]における文字分類の精度向上を目的に、「DE-パラメータ」[7]を基に新しい不変量を提案する。

3.1 DE-パラメータ

文献[7]は、射影変換を受けたテキストチャから同一の図形を探し出し、これの面積を基に消失線を求める。そのため2つの図形が同じ図形かを調べる方法について言及しており、そのうちのひとつがDE-パラメータである。文献[7]では、ある図形に対して、同じ共分散行列を持つ楕円をその図形の「A-楕円」と呼んでいる。図2(a)の図形を対象にした場合、A-楕円は図2(b)である。このときDE-パラメータは図2(c)が示すA-楕円との不一致部分の面積を、図形の面積で正規化したものである。

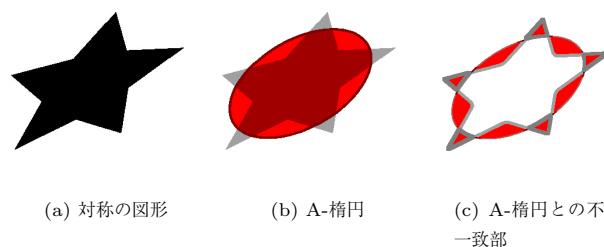


図2 DE-パラメータ算出に用いる領域

DE-パラメータを式で表すと、

$$\text{DE-パラメータ} = \frac{\text{A-楕円と図形の不一致部の面積}}{\text{図形の面積}}$$

である。

3.2 DE-パラメータの文字への適用のための実験

文献[7]では丸や三角といった単純な5種類の図形をDE-パラメータで分類しているが、アルファベットや数字など、より複雑な図形については検討されていない。そこで実験により検証を行う。

実験は数字10字種とアルファベットの大文字26字種、そしてアルファベットの小文字 $a \sim z$ の26字種のうち輪郭が複数ある i と j を除いた24字種の、合計60字種を対象に行う。図3のように、文字を描いた一辺の長さが96ピクセルの画像を用意し、これらの画像に対して様々なアフィン変換を加える。ここで、アフィン変換行列 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ を次のように分解する。

$$A = L(\beta)R(\theta)S(\varphi)A(\alpha) \quad (3)$$

$$= \begin{pmatrix} \beta & 0 \\ 0 & \beta \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} 1 & \tan \varphi \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & 1/\alpha \end{pmatrix}$$

$$\alpha = \pm \sqrt{\frac{a^2 + c^2}{ad - bc}} \quad (4)$$

$$\varphi = \tan^{-1} \frac{ab + cd}{ad - bc} \quad (5)$$

$$\theta = \cos^{-1} \frac{\pm a}{\sqrt{a^2 + c^2}} \quad (6)$$

$$\beta = \pm \sqrt{ad - bc} \quad (7)$$

このうち拡大縮小のパラメータ β を除いた3パラメータ α, φ, θ を変化させてアフィン変換を加えた図形を実験に利用する。各パラメータを $\alpha = \{1, 2, 3, 4\}$, $\varphi = \{-1.0, -0.9, \dots, 1.0\}$, $\theta = \{0.0, 0.1, \dots, 1.4\}$ と変化させ、 $4 \times 21 \times 15 = 1260$ 通りのアフィン変換を加える。

実験の結果を図4に示す。図は各字種に対して1260個のデータの分布を赤のヒストグラムで表している。ヒストグラムは値を縦方向、度数を横方向にとっている。ただし、DE-パラメータの値は平均0、分散1に正規化して表記している。DEパラメータは図のように字種ごとに安定した値が得られた。しかし、複数の字種が同じ値をとるため、単独では字種を完全に判別する能力は持たないことがわかる。

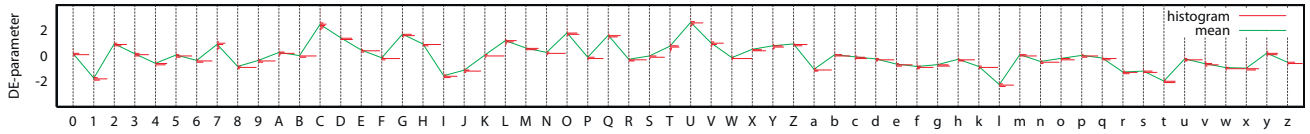


図4 DE-パラメータ

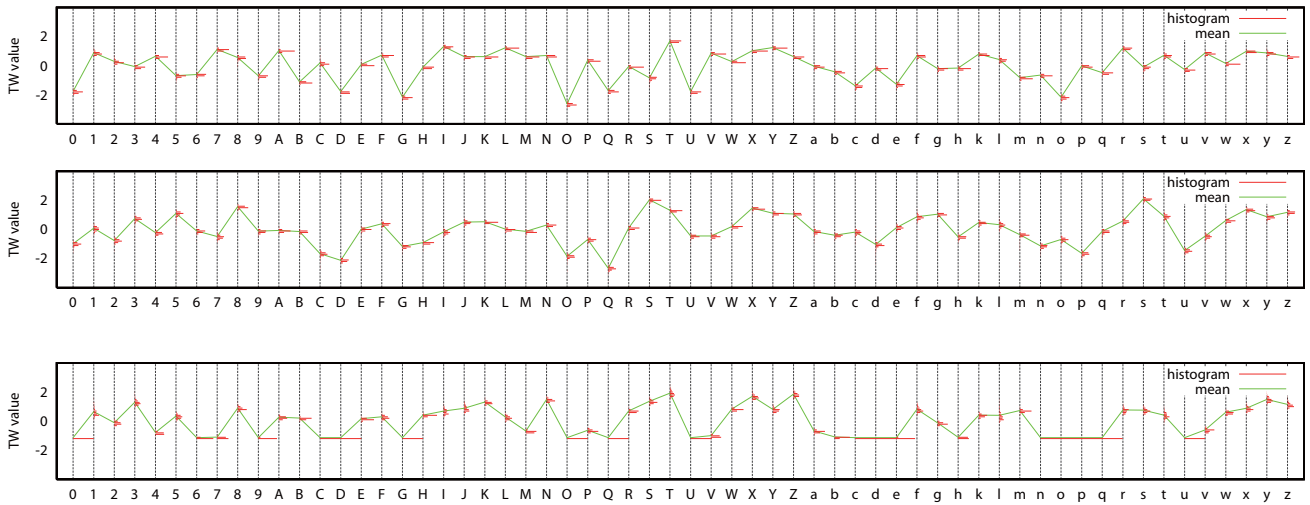


図5 TW 値: 上から順に (a)TW(100, 80), (b)TW(80,60), (c)TW(20,0)



図3 実験データ

3.3 DE-パラメータの改良

DE-パラメータでは前述のように区別できない文字も多いため、本稿では文字を区別するのに適した形に改良したTW値を提案する。まずDE-パラメータを求める際に利用したA-楕円を拡大縮小して、中心が同じで、A-楕円に対して $m\%$ の大きさの楕円と $n\%$ の大きさの楕円を用意する。そして、 $m\%$ の楕円と $n\%$ の楕円で区切った文字領域の面積の割合を $TW(m,n)$ と表す。TW値を式で表すと、

$$TW(m,n) = \frac{m\% \text{ と } n\% \text{ の楕円の間に含まれる領域の面積}}{\text{図形の面積}}$$

である。TW(100,80), TW(80,60), TW(60,40), TW(40,20), TW(20,0)を図6に示す。

TW値が字種を区別できることを実験により確認する。図5は前節のDE-パラメータと同様にTW値について実験した結果である。図より、TW値はアフィン変換に頑健であり、さらに切り取る楕円の大きさによって判別できる文字が異なっていることがわかる。したがって、複数のTW値を組み合わせることにより、字種の判別能力がDE-パラメータより高い。

文献[3]では2文字を組み合わせることにより、不変量を多次元化した。これには画像中に大量の文字を必要とする問題があった。TW値を利用すれば、文字を組み合わせることなく不変量を多次元化できるため、画像中の文字数が同じ場合の判別性能が向上すると期待される。

4. 正規化を利用した正対画像の取得

文献[3]の手法にはアフィン歪みが残ってしまうという問題がある。本稿ではこの問題を解決するために、アフィン不変な文字認識手法を利用する。文字認識を利用するとともに、個々の文字の変形パラメータを求める。さらに変形パラメータを投票することで、紙面の補正に必要な変換パラメータを定め、アフィン歪みを持つ文書画像を補正する。

4.1 アフィン不変な文字認識

本稿ではアフィン不変な文字認識手法[9]を用いる。この手法は、アフィン変換を受けた図形の正規化による認識手法を文字認識に応用したものである。

文献[9]では、正規化のために近似楕円と呼ぶ楕円を利用している。これは、文献[7]のA-楕円と同じものであり、対象の図形と同じ共分散をもつ。認識対象へ、その近似楕円を円に変形させるアフィン変換を加えることを、正規化としている。認識対象を正規化し、学習データを正規化したものと比較することで、対象を認識することができる。

4.2 アフィン変換パラメータと復元

4.1の手法の正規化はアフィン変換である。従って、そのアフィン変換から認識対象の歪みを復元するアフィン変換パラメータを求めることが可能である。これを図7に示す。認識対象の文字画像(入力画像)と、それに一致した学習データ(テンプレート画像)の正規化をそれぞれ行列 A_1 と A_2 で表す。入力画像にはさらに回転 A_3 を加え、両者を一致させる。即ち、入力画像はアフィン変換 $A_2^{-1}A_3A_1$ によってテンプレート画像と同じ形に戻る。

入力画像中の文字が全て同じ向きに揃っていれば常に同

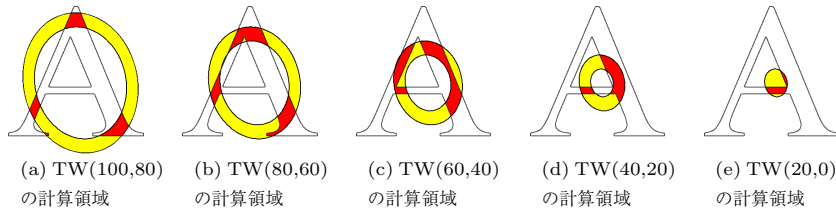


図 6 TW 値の計算領域

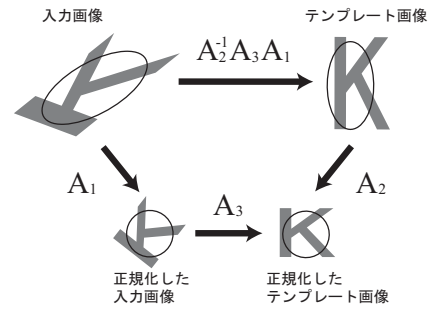


図 7 入力画像とテンプレート画像の正規化

じアフィン変換で正面を向くはずである。しかし、実際に変換を求めるとき、アフィン歪み以外の歪みの影響がある(歪みが残っている)可能性や、計測に誤差が出ること、また、認識に失敗することもある。従って、画像中の全ての文字から共通のアフィン変換を得ることはできない。そこで、最も信頼できる変換を、全ての文字から得られた変換の中から選ぶ。

アフィン変換行列を式(4)のように分解して、 α , φ , θ の3パラメータだけを取り出し、投票処理を行う。特に誤認識を起さにくい字種をあらかじめ選んでおき、画像中の該当文字に対して3パラメータを求め、3次元空間に配置する。そして、 k -近傍法による密度推定[10]を利用して、密度の高い点を選ぶ。すなわち、3次元空間内の全ての点に対し、ユークリッド距離が k 番目に近い点までの距離を測る。「 k 番目までの距離が最も小さい点が最も密度の高い位置にある点」と考え、代表点を選ぶ。この代表点のアフィン変換パラメータで全体の補正を行う。

5. 実験

本稿で行った2つの改良(以下、提案手法)の有効性を検証するため、文献[3]の手法(以下、従来手法)と比較する。

5.1 平行性の復元実験

まず、射影歪みをアフィン歪みに補正する際の精度を定量的に評価する。

実験データは文書を実際にデジタルカメラで撮影して作成した。撮影対象の文書は文字の割合が多い英文雑誌の電子データを58ページ用意し、A4用紙にインクジェットプリンタ EPSON PX-G920 で印刷したものをデジタルカメラで撮影した。撮影に用いたカメラは Canon EOS Digital 5D と Fujifilm Finepix F710 である。また、一眼レフカメラのレンズは EF24-105mm F4L IS USM と 100mm F2.8 マクロ USM を利用した。使用カメラやレンズ、撮影角度や照明を変えて4つの条件で58文書を撮影した232枚の画像を対象に実験を行った。また、手法のパラメータ(フィルタのウィンドウサイズや2値化の閾値、最適化問題の初期値や収束条件)は、実験的に定めた。従来手法は不変量として凸包による面積比を用い、提案手法では凸包による面積比と3.3で定義した TW(100,80), TW(80,60), TW(60,40), TW(40,20) を用いて不変量ベクトルを作成した。

実験に用いる文書にはあらかじめ、21mm 間隔で縦横に9×12個、格子状に小さな印を印刷しておく。この印を用いて、本来長方形であった領域のうち、取り得る最大のものを取り出し、

表 1 平行性復元の実験

		従来手法	提案手法
実験に用いた画像数		232	
評価無し回数		1	
補正失敗回数		21	8
長辺 補正前	平均(度)	5.8	
長辺 補正後	平均(度)	7.0	2.9
	標準偏差(度)	8.1	2.4
短辺 補正前	平均(度)	17	
短辺 補正後	平均(度)	7.6	4.6
	標準偏差(度)	9.1	4.2

表 2 補正実験の定量評価

角度の誤差(度)		アスペクト比の復元誤差(%)
垂直方向	水平方向	
4.6	4.3	6.4
15	15	5.1

対辺の角度を補正前と補正後で比較する。補正が成功すれば領域は平行四辺形になるので、補正後の角度が0に近いほど良い。ただし、撮影範囲に全く文字が無い画像は「評価無し」として評価に含めない。また、クラスタリングの失敗などが原因で、射影変換パラメータを求める計算が収束しなかった場合も、定量評価できないために除外する。この場合を「補正失敗」と呼ぶ。「評価無し」と「補正失敗」以外の実験結果については、補正前と補正後の対辺の角度を平均し、評価する。

実験の結果を表1に示す。表より、提案手法のほうが従来手法より平均的に優れていることがわかる。さらに、具体的な例を挙げて実験結果を評価する。従来手法と提案手法で共通するのは文書の広範囲を撮影した画像では安定して良い結果が出やすいことと、複数書体が混じっていたり大きな図がある場合は悪い結果が出やすいことである。提案手法が従来手法に比べて特に優れた結果が出るのは、文字数が少ない文書画像に適用した場合である。従って、不変量の改良により文字数が少ない文書に対しても精度が上がり、従来よりも優れた結果を得ることが確認できた。

5.2 正対画像の復元実験

5.1の実験結果をさらに補正して、正面を向いた画像を得る実験を行う。扱う文書画像中で最も多く用いられている1書体の文字画像に、文献[11]のぼけの劣化モデルを適用し、劣化画

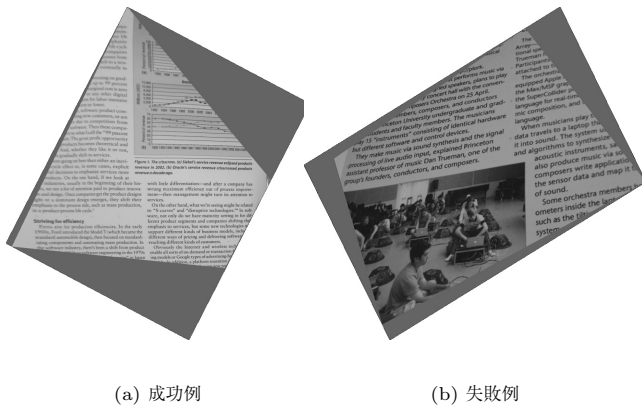


図 8 補正実験の結果

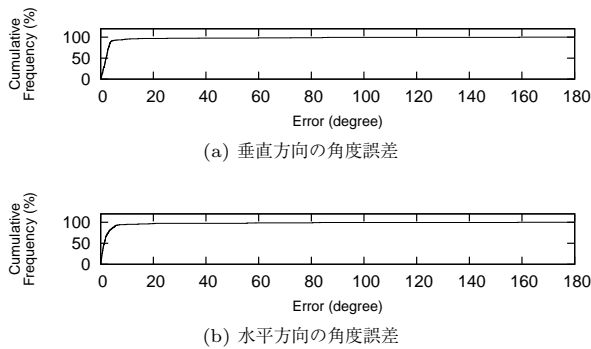


図 9 角度誤差の累積ヒストグラム

像を作成する。これらの文字画像を学習させて 4. の方法で補正する。そして、5.1 の実験で用いた四角形領域の対辺を用いて評価を行う。実験の結果を図 8 と表 2 に示す。

表 2 は、四角形領域の長辺の角度の平均値と短辺の角度の平均値を計算し、それぞれが垂直方向、水平方向から何度ずれているかを表している。また誤差の累積ヒストグラムを図 9 に示す。5 度未満の誤差にあるものはそれぞれ 9 割以上であることがわかる。一方で 90 度近い大きな誤差が出たデータも存在した。大きな誤差が出た実験データを調べたところ図 8(b) のように、学習させた書体の文字がない場合や、学習させた書体の文字があっても、割合が極端に少ない場合に誤差が大きくなっている。ただし、このように理由付けできないケースもあり、原因についてさらなる追求が必要である。

またアスペクト比が正しく復元されているかを評価する。補正後の四角形の対辺の長さを平均して縦横の長さとしたときの、アスペクト比の復元誤差を表 2 で示している。表より、アスペクト比は若干の誤差はあるものの、おおむね正しく復元できていることがわかる。

以上より、射影歪みを受けた文書画像から正面を向いた画像を得られることを確認した。今後は精度の向上を図るため、データによって誤差が大きくなる理由を解明することが課題となる。

6. まとめ

本稿では、文献[3]で提案した「変量」と「不変量」を組み合わせる文書画像の射影歪みを補正法に 2 つの改良を加えた。まず、不変量の改良によって安定性を向上させた。文献[3]では不変量として面積比のみを用いていたが、本稿では TW 値という字種判別性能の高いアフィン不変量を導入した。従来手法では複数の不変量を計算するために複数の連結成分が必要であったが、提案手法では 1 つの連結成分から計算することができ、画像中に写る文字が少ない場合の性能が大幅に向上した。

さらに、射影歪みを完全に補正した。文献[3]は文書画像の射影歪みをアフィン歪みへ補正することができたが、アフィン歪みを取り除くためには、対象に関する知識が必要であった。そこで本稿では、対象に関する知識として書体情報の学習を導入し、正面を向いた画像を得る方法を提案した。実験を行い、射影歪みを受けた文書画像から正面を向いた画像を得られることを確認した。

謝辞 本研究の一部は、財団法人 大川情報通信基金の補助による。

文 献

- [1] J. Liang, D. Doermann and H. Li, "Camera-based analysis of text and documents: a survey," IJDAR, vol.7, 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・文書画像解析の現状と課題," 信学技報 PRMU2004-246, March 2005.
- [3] 丹羽亮, 岩村雅一, 黄瀬浩一, 内田誠一, 大町真一郎, "変量と不変量に基づく射影歪みからアフィン歪みへの補正," 画像の認識・理解シンポジウム (MIRU2007) 論文集, pp.720-725, July 2007.
- [4] P. Clark and M. Mirmehdi, "Recognising text in real scenes," IJDAR, vol.4, pp.243-257, 2002.
- [5] C. H. Lampert, T. Braun, A. Ulges, D. Keysers and T. M. Breuel, "Oblivious document capture and real-time retrieval," Proc. CBDAR2005, pp.79-86, Aug. 2005.
- [6] 池谷彰彦, 佐藤智和, 池田聖, 神原誠之, 中島昇, 横矢直和, "カメラパラメータ推定による紙面を対象とした超解像ビデオモザイク," 信学論 (D), vol.J88-D, no.8, pp.1490-1498, Aug. 2005.
- [7] 大田友一, 前信潔, 坂井利之, "透視射影されたテキストチャから平面の傾きを得る一方法," 情報処理学会研究報告, 1981-CVIM-16, pp.1-8, Jan. 1982.
- [8] M. Pilu, "Extraction of illusory linear clues in perspective skewed documents," Proc. Computer Vision and Pattern Recognition, 2001 (CVPR '01), Vol. 1, pp.363-368, 2001.
- [9] 堀松晃, 丹羽亮, 岩村雅一, 黄瀬浩一, 内田誠一, 大町真一郎, "アフィン変換と劣化にロバストな文字認識," 画像の認識・理解シンポジウム (MIRU2008) 論文集, July 2008. (掲載予定).
- [10] K. Fukunaga, Introduction to Statistical Pattern Recognition, Academic Press, 177-184 pp., 1972.
- [11] 石田皓之, 高橋友和, 井手一郎, 目加田慶人, 村瀬洋, "携帯カメラ入力型文字認識におけるぼけやぶれに対処するための生成型学習法," 電子情報通信学会論文誌 D, vol.89, no.9, pp.2055-2064, Sept. 2006.