

局所特徴の位置関係を用いた情景画像中の文字認識

小林 拓也[†] 岩村 雅一^{††} 黄瀬 浩一^{††}

[†] 大阪府立大学工学部 〒 599-8531 大阪府堺市中区学園町 1-1

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

E-mail: kobayashi@m.cs.osakafu-u.ac.jp, {masa,kise}@cs.osakafu-u.ac.jp

あらまし カメラで撮影した情景画像中の文字を全て認識し、その中から重要な情報を得られれば非常に有用であると考えられる。それを実現するためには、認識対象の画像中から文字の部分を正確に識別し、認識することが必要となる。本稿では局所特徴の位置関係を用いることで画像中から文字部分を見つけ、その文字が何であるか認識することを考える。提案手法を用いて行った実験では、正面から撮影した情景画像中の日本語を 84.3% の再現率と 97.0% の適合率で認識することができた。

キーワード 文字認識, 局所特徴量, SIFT, ANN, RANSAC

Recognition of Characters in Scenes by Using Geometric Contexts of Local Features

Takuya KOBAYASHI[†], Masakazu IWAMURA^{††}, and Koichi KISE^{††}

[†] School of Engineering, Osaka Prefecture University

1-1 Gakuencho, Naka, Sakai, Osaka, 599-8531 Japan

^{††} Graduate School of Engineering, Osaka Prefecture University

1-1 Gakuencho, Naka, Sakai, Osaka, 599-8531 Japan

E-mail: kobayashi@m.cs.osakafu-u.ac.jp, {masa,kise}@cs.osakafu-u.ac.jp

Abstract If you can get useful information from a character image in scene you take with your camera, all you have to do is release the shutter and you can save time. In order to realize such a system, we have to propose a character recognition system such that it can correctly find the character area from the scene image and recognize it. In this paper, we propose two methods to achieve it, which consider geometric contexts of local features. We had a utmost 84.3% recall and 97.0% precision in character recognition experiments we had.

Key words Characters Recognition, Local Feature, SIFT, ANN, RANSAC

1. はじめに

近年カメラ付き携帯端末の性能が向上していることから、それによって情景画像を撮影することで、その中にある文字を認識し、有用な情報を取り出すサービスが考えられる。具体的な応用例として、情景画像中の全ての文字を認識し、看板や標識の情報を読み取りそれを音声に変換することで視覚障害者に伝える応用が考えられる [1]。別の応用としては、気になった単語を撮影することで、その単語をキーワードとして Web ページにアクセスしたり、動画を再生したり、その関連情報を得るサービスなども考えられる [2]。

そのようなシステムの実現を考えた際に必要になるのは、情景画像中から正確に文字の部分の切り出し、認識することであ

る。文字は必ずしも正面から撮影されるとは限らず、撮影する距離や角度によって画像が変わるため、射影変換に対する頑健性を有することが望ましい。

情景中の射影変換に頑健な既存の文字認識手法として、文献 [1, 3-6] などが挙げられる。このうち、草地らの手法 [4] は頑健な認識が可能であるが、検索用のインデックス作成を目的としてあらゆる文字領域の可能性を考慮するため、処理に時間がかかり前述のような実時間性が求められるアプリケーションには向かない。草地らの手法 [4] 以外の手法ではまず入力画像から文字領域を切り出し、その後の認識処理によって文字認識を行っているため、切り出しに失敗した時点で文字認識にも失敗してしまうという問題が存在する。そのため、複雑な背景を持つ文字をうまく認識できない可能性がある。

そこで本稿では射影変換と複雑な背景に頑健な文字認識の実現を目指して、SIFT [7] などの局所特徴量を予め登録しておいた参照画像と入力画像から抽出し、それを照合することで文字の切り出しと認識を同時に行う方法を検討する。この方法であれば、複雑な背景の影響で文字の一部が切り出せず、既存手法では認識できないような文字も、文字の局所から特徴を得ることで認識できる可能性がある。また、局所特徴量は認識に有効な領域のみから抽出されるため、文献 [4] のようにあらゆる場所に文字領域が存在する可能性を考慮する処理よりは高速と考えられる。文字認識に局所特徴量が使われている例として、手書き文字を対象としたものには内田らの手法 [8] があり、フォント文字を対象としたものでは Campos らの手法 [9] や Wu らの手法 [10] などが挙げられる。しかし、いずれも認識対象は切り出された文字であるため、情景画像に含まれる切り出されていない文字の認識には使えない。

そこで、複数の文字が書かれた画像から各文字の領域を特定するために Philbin らによる建造物認識の手法 [11] を参考にする。この手法では入力画像中の建造物を効率よく探索するために、建造物領域から抽出される特徴点の位置関係が保存されることを利用する。具体的には、参照画像と入力画像の特徴点の対応を求め、建物領域に正しく対応付いた複数の特徴点を RANSAC [12] で選定し、リランキングに使用する。本稿ではこのアイデアを文字領域の特定と認識に適用することを考える。しかしこの手法のように RANSAC を用いるだけでは文字領域の正しい対応を検証することはできない。なぜなら、文章中には普通同じ文字が何度も出現するが、Philbin らの手法は、一つの画像中に認識対象となる建造物は一つしか映っていないという仮定を用いているためである。これは RANSAC の制約による。すなわち、RANSAC は入力画像と参照画像の全ての特徴点の位置関係から妥当な対応を求める手法であるため、入力画像中に複数の文字が存在すると、妥当な対応が求まらない。

そこで本稿では RANSAC の点の選択範囲を限定することで特定の一文字から抽出された特徴のみで位置関係を計算し、文字領域の特定と認識を行う手法を提案する。提案手法の性能を評価するため、風景画像上に書かれた文字の認識実験を行う。本稿では、日本語で書かれた画像を認識対象とする。

2. 関連手法

本節ではまず文字切り出しに基づく文字認識手法として、既存手法を説明する。その際に、そのような手法の問題点を示し、本稿で用いる解決策について述べる。その後、本稿で文字認識に用いる関連手法について述べる。まず、局所特徴量抽出・記述法である SIFT について述べた後、近似最近傍探索を行うための手法である ANN について説明する。

2.1 文字切り出しに基づく既存の文字認識手法

情景画像中の文字を認識する際に、まず入力画像から文字領域を切り出し、それを認識処理にかけると提案されている。Myers らの手法 [3] は複数の閾値を用いる 2 値化処理によって画像中に含まれる文字行領域の切り出しを行った後、それを分割して文字領域を特定している。この手法では文字行単

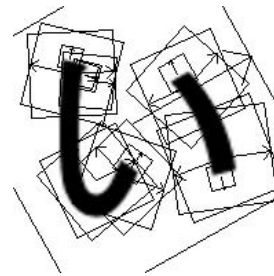


図 1 文字から抽出された SIFT 特徴の例

位の切り出しを想定しているため、文字が一つずつ点している場合などにはこの手法を用いることができない。辻らの手法 [1] では、画像の 2 値化により文字を構成する連結成分を切り出すことで、Myers らの手法で扱えなかった点状文字が認識可能になったが、認識対象として白い背景に黒い文字で書かれたものを想定しているため、情景画像中の文字認識にそのまま用いることが出来ない。また、Li らや Chen らはそれぞれ文献 [5, 6] で、標識や看板に書かれた文字を認識する手法を提案している。認識対象が射影変換を受けていても頑健に文字の切り出しと認識を行うことを、Li らは切り出された文字の連結成分を射影不変な特徴量を用いて記述することで実現している。一方 Chen らもアフィン変換で射影歪みを補正し、高精度な認識を行なっている。これらの手法は共通の問題として、文字切り出しを行った後に認識処理に移るため、うまく切り出せなかった文字に関しては正確な認識を行えない。

2.2 SIFT

Scale-Invariant Feature Transform (SIFT) [7] は David Lowe によって考案された局所特徴量抽出・記述法である。SIFT は主にイメージモザイクのマッチングや物体認識の分野で用いられている [13]。この手法は画像の局所的な部分から識別性の高い特徴点を抽出し、その特徴点の周辺の情報が 128 次元のベクトルとして記述される。得られた特徴は回転、スケール変化に不変であり、ある程度の照明変化と射影変換にも耐えることができる。図 1 に SIFT によって文字から抽出された特徴の例を示す。描画されている矩形と矢印はそれぞれ特徴点のスケールとオリエンテーションを表している。SIFT の処理は主に 2 つのステップからなる。まず最初のステップは特徴点の抽出である。SIFT は特徴点の位置とそのスケールを、DoG 処理によるスケール空間の極値探索によって決定する。この処理では、ある画素がスケール空間内の近傍点と比較して最大または最小であればその画素を特徴点とする。この処理によりスケール変化に不変な特徴点を選択できる。その後、得られた特徴点の中から、エッジ部分のようにノイズなどの影響を受けやすい点を除去する。

次のステップでは、得られた特徴点周りの情報から特徴ベクトルを記述する。その際には回転に不変な特徴を得るために、特徴点におけるオリエンテーションを算出し、方向に関する正規化を行う。その後、得られたオリエンテーションと周辺の情報から、128 次元の特徴ベクトルが計算される。そして、照明変化への頑健性を得るために特徴ベクトルの各要素が取りうる

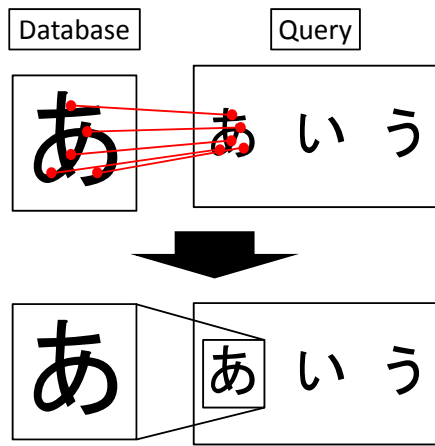


図 2 文字領域特定処理の概略．上段:入力画像と参照画像から局所特徴を抽出し、その対応を求める．下段: 対応付いた複数の特徴点の位置関係を用いて文字領域の推定を行う．

値についても正規化が行われる．

2.3 ANN

Approximate Nearest Neighbor(ANN) [14] は近似最近傍探索手法の 1 つであり、木構造を用いることで高速な処理を可能にしている．本稿では SIFT によって入力画像と参照画像から得られたそれぞれの特徴ベクトルから、最も距離が近いものを探索するために ANN を用いる．ANN では特徴空間の分割により 2 分木を作成し、それぞれの葉に 1 つの特徴ベクトルを対応付ける．特徴ベクトルは特徴空間の分割により、それぞれ 1 つの超長方形のセルに属する．

クエリとして特徴ベクトルが与えられると、まずクエリの特徴ベクトルが特徴空間内のどのセルに属するかを求める．そしてセル内の特徴ベクトルとの距離を計算し、その距離を半径とする超球を探索することで、真の最近傍点を求める．その際、近似パラメータとして ϵ を用いることで、探索する範囲を限定し処理の高速化が図れる．しかし探索する範囲を限定し過ぎると、高速化とともに探索の精度が低くなるというトレードオフの関係が存在する．

3. 文字領域特定法

本節では本稿で提案する文字領域特定法について説明する．提案手法では、局所特徴量の位置関係を用いることで文字領域の特定と認識を同時に行う．図 2 にその処理の概略を示す．まず参照画像と入力画像の双方から SIFT によって局所特徴量を抽出・記述する．そして ANN を用いた近似最近傍探索により、入力画像から得られたそれぞれの特徴点がどの参照画像 (文字) のどの特徴点に最も近いかを求める．これにより、入力画像中の各特徴点と各参照画像の特徴点との位置的な対応関係が分かる．そして、ある参照画像に対応付いた特徴点を 3 つ用いることで、入力画像中の 3 つの特徴点の座標の組と参照画像中の対応する 3 点の座標から、それぞれの画像間の座標の対応関係を計算するためのアフィン変換行列を求めることができる．そのアフィン変換行列を用いて、参照画像中の文字領域を囲む矩形

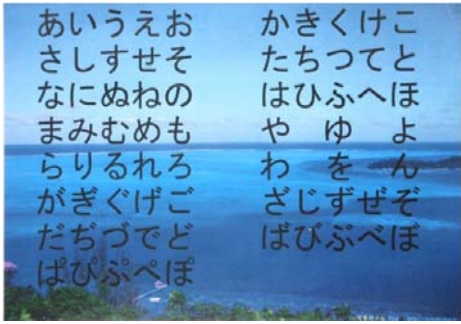
の各頂点の座標を入力画像上の座標へ射影し、新たな矩形を作成することで、それがその文字の領域であると特定すると同時にその文字であると認識する．その際、その矩形領域の中からその参照画像に対応付いた特徴点の数を数え、それをその参照画像から得られた特徴点数で割ったものを得点として与える．この処理を対応関係の得られた全ての 3 点の組に対して行う．入力画像中で、射影された全ての矩形について、中心の (x,y) 座標の差がそれぞれ 20 ピクセルより小さく、面積の差が 10000 より小さかった場合は、それぞれの得点を比較し、最高のもののみをその領域の認識結果とする．

以上が文字認識処理の流れであるが、文字領域の特定処理では、用いる 3 点の組の座標の対応関係がどれほど正確かによって、文字領域を囲む矩形の正確さが変化する．そこで、入力画像中からそのような 3 点を選択する方法として 2 つの手法を提案した．

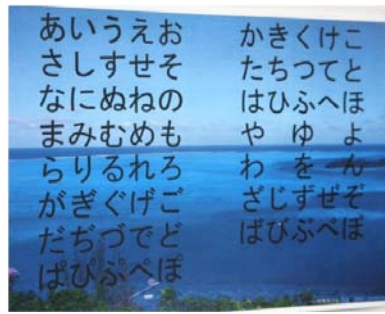
3.1 提案手法 1: 局所 RANSAC

1 つ目の提案手法は射影の精度を上げるために RANSAC [12] を用いる方法である．RANSAC はロバストなパラメータ推定法の 1 つであり、今回はアフィン変換行列の推定精度向上のために利用した．処理の流れを説明すると、まず入力画像中の、ある参照画像 (文字) に対応付いた特徴点の中から 3 点をランダムに選択し、それに対応する参照画像中の 3 点とでアフィン変換行列を算出する．次に、同じ参照画像と対応付いた点のうち、選んだ 3 点とは別の点をそのアフィン変換行列により射影する．もし対応関係が正しいのであれば、射影された点の近傍には、その点と対応付いた点とが実際に存在しているはずなので、その距離を計測し、閾値以下なら投票を行う．これを最初に選んだ 3 点以外の対応付いた点全てに対して行い、得票数が一定以上ならばその変換行列は妥当であると判断する．この処理を規定回数繰り返し、最も得票数が多かったアフィン変換行列を求める．そしてその得票数が決められた閾値を上回っていた場合、これを用いて文字領域を囲む矩形を射影する．このようにして文字領域特定の精度を向上することができる．

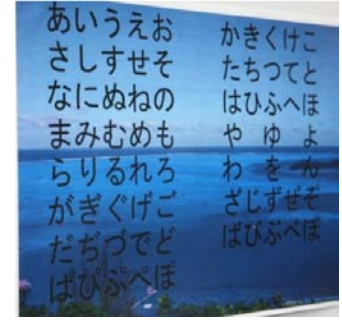
しかし RANSAC を入力画像中の文字認識に適用することを考えた場合、問題が発生する．RANSAC は画像全体の特徴点の中から、ある文字と対応付いた特徴点を探索し、そこからランダムに特徴点を選択する．しかし実際の文章中には何度も同じ文字が出現するため、RANSAC をそのまま用いると、複数個ある同一の文字からランダムに特徴点を選択してしまうことになる．その結果、誤った対応関係が求まる可能性が急増し、また、複数ある文字の中で最も良い対応関係を得られた文字のみに票が固まり、結果として全ての文字を認識できなくなるケースが考えられる．そこで我々は、同一文字が複数存在するケースでも正確に文字領域を特定するために、特徴点の選択範囲を限定した局所 RANSAC を提案する．局所 RANSAC では、ある特徴点を基準とし、その特徴点と同じ文字に対応付いた点をその近傍から探索する．もしその範囲内で見つからなかった場合は、更に探索範囲を広げて探索を続ける．この処理を画像の端に突き当たるか予め決めた数の特徴点が見つかるまで行う．そして見つかった点について RANSAC を行い、文字領域を特



(a) 0 度

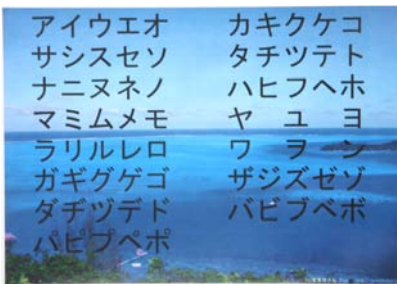


(b) 30 度

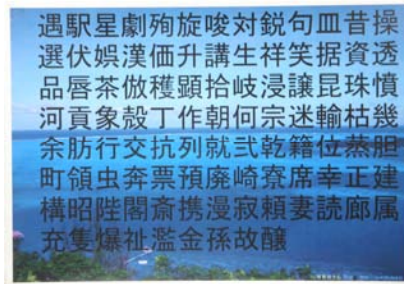


(c) 45 度

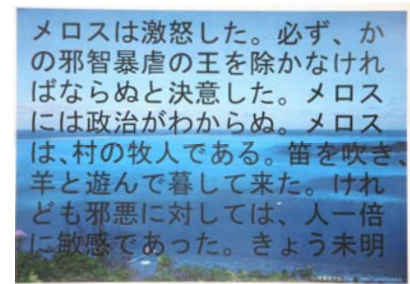
図 3 角度別の認識対象の例 (ひらがな)



(a) カタカナ



(b) 漢字



(c) 物語冒頭

図 4 0 度の場合の認識対象の例

定する．この処理を入力画像中の全ての特徴点について行う．

3.2 提案手法 2: 近傍点探索

2 つ目の提案手法として、近傍点探索を提案する．これはある特徴点を基準としてその近傍から特徴点を探索するという点では局所 RANSAC と同じであるが、特徴点が近傍から 2 点見つかれば 3 点の組が得られた時点で探索処理を打ち切り、その 3 点でアフィン変換行列を求めるといったものである．

以上の 2 手法を提案手法とし、今回は更に文字領域特定の精度を向上させるために、参照文字を囲む矩形の各頂点を射影させてできる矩形について、その短辺の長さが長辺の $3/4$ より短い場合には射影を行わないこととした．これは、文字を囲む矩形はある程度正方形に近い形状であるという仮定に基づくものである．

4. 実 験

4.1 実験条件

3 節で提案した認識手法の性能を評価するため、風景画像上に書かれた文字列の認識実験を行った．今回は参照画像用の文字として日本語のひらがな 71 字種とカタカナ 71 字種 (それぞれ濁音半濁音を含む) と常用漢字 1945 字種の合計 2088 字種を選択した．参照画像はそれぞれの文字が白い背景に黒い文字で書かれた画像であり、フォントは MS ゴシックで文字のサイズは 60pt とした．各文字は 97×97 ピクセルの白い背景の中央に

配置した．特徴量の抽出の際には上記の参照画像のサイズを 2 倍に拡大したものをを用いた．認識対象には、風景画像の上に黒い文字が複数書かれたものを A4 用紙に印刷し、その用紙を正面に対して 0 度、30 度、45 度の角度をつけてカメラで撮影して用いた．画像サイズは 4368×2912 である．認識対象として異なる文字が書かれた 10 枚の画像を用意した．まずひらがな 71 字種が書かれたものとそれをカタカナにしたものの 2 枚である．また、常用漢字からランダムに重複のないよう 100 字種選択したものを 5 枚と、物語の冒頭から 100 字程度引用したものを 3 枚用意した．図 3 にそれぞれの角度の認識対象の例 (ひらがな) を示す．また、図 4 には 0 度の場合のひらがな以外の認識対象の例を示す．なお、フォントは MS ゴシック、文字のサイズは 72pt である．

評価方法には、再現率と適合率を用いた．再現率は画像に書かれている文字数のうち、何文字を正しく認識できたかによって求め、適合率は、誤認識も含めて認識された文字数のうち、何文字を正しく認識できたかによって求めた．また、今回の実験では、ANN の近似パラメータである ϵ は 2 とし、RANSAC の繰り返し回数は 100 回とした．RANSAC の得票数の閾値は 2 とした．

4.2 結果・考察

図 5 に局所 RANSAC による 0 度の場合の 4 種類の認識結果を示す．それぞれの字の周りの矩形が射影された文字領域を

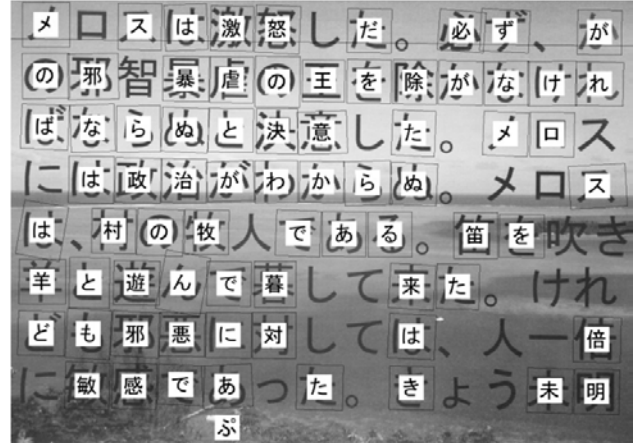
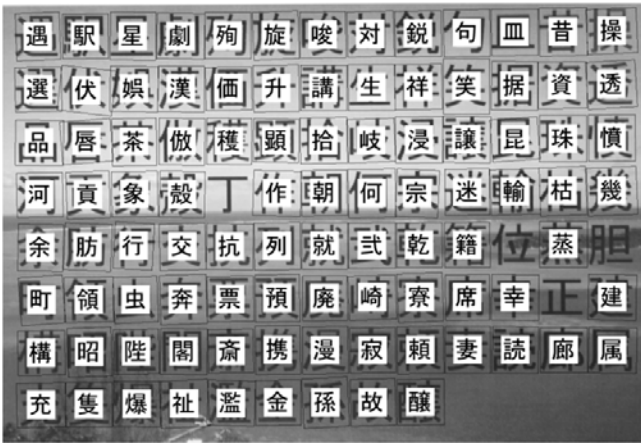
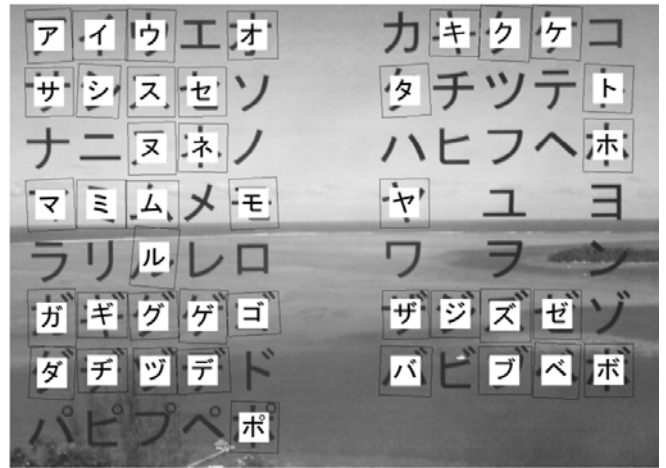
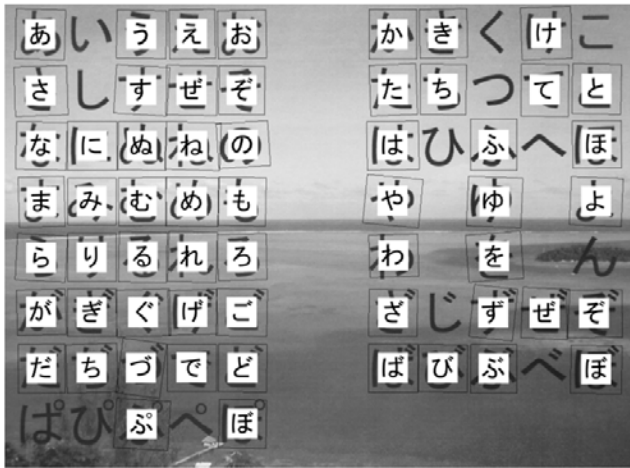


図 5 0度で局所 RANSAC を用いた場合の認識結果
(左上:ひらがな 右上:カタカナ 左下:漢字 右下:物語冒頭)

表 1 再現率 [%]

手法	0度	30度	45度
局所 RANSAC	84.3	79.5	29.0
近傍点探索	83.6	76.1	29.5

表 2 適合率 [%]

手法	0度	30度	45度
局所 RANSAC	97.0	97.6	78.9
近傍点探索	87.6	88.4	62.4

表 3 分類別再現率 (0度)[%]

手法	ひらがな	カタカナ	漢字	物語冒頭
局所 RANSAC	76.1	56.3	96.2	72.7
近傍点探索	73.2	57.7	95.8	71.7

表し、貼り付けられた文字がその認識結果を表している。正しい認識結果が各文字の中央部に貼り付けられていれば正解とした。図 5 から、認識に成功している文字に関しては、ほとんど完璧な位置に文字領域を囲む矩形が現れていることが分かる。表 1, 2 にそれぞれ再現率、適合率を示す。表 3 は 0 度の場合の分類別再現率である。表 1 から、SIFT によって得られる局所特徴量が風景画像上に書かれた文字の認識に有効であることが確認できた。表 3 より、分類別に再現率を見た場合、漢字が 95% 以上となっており、高い精度が得られた。これは、ほとんどの漢字が複雑な形状を持つため、局所から安定した特徴点が多く得られ、文字領域の特定に有利に働いたことが大きな要因であると考えられる。撮影角度の変化に対する頑健性を見ると、30 度までであればほとんど再現率を低下させずに認識が行えた。この結果から、ある程度の角度をつけて撮影した文字も正確に認識出来ることが確認できた。

次に手法間の比較を行う。再現率ではどちらの手法もほぼ同

じ精度となったが、表 2 を見ると近傍点探索が局所 RANSAC と比べ低い適合率となっていることが分かる。そこで、比較のために近傍点探索による 0 度の場合の物語冒頭の認識結果を図 6 に示す。近傍点探索では誤った位置に認識結果の文字が多数貼り付けられていたり、文字領域でないところにも矩形が描かれていることが分かる。これは、近傍点探索が RANSAC による正確な変換行列の推定を行っていないため、不適切な矩形も射影してしまったことが原因だと考えられる。そのため、多数の誤認識が発生し、適合率の低下につながった。そのような誤認識の結果が除外されずに残っていた場合、今後単語認識などの処理を実装することを考えた際に不都合が生じる可能性が高い。以上から、RANSAC は正確な文字領域特定に非常に有

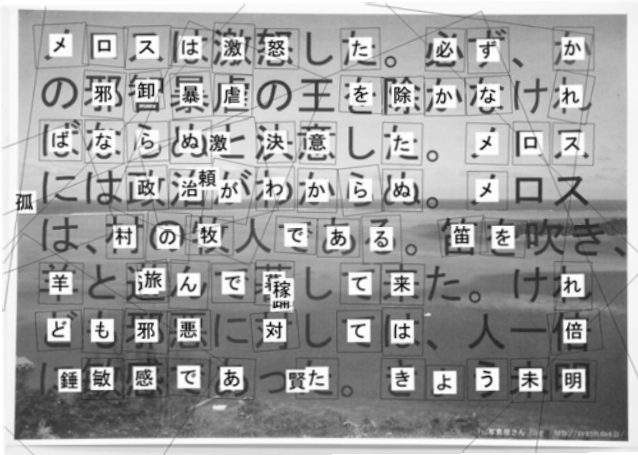


図 6 0度で近傍点探索を用いた場合の認識結果 (物語冒頭)

効であると言える。

また、図 5 右下の物語冒頭の認識結果を見ると、画像中に同じ文字が複数存在していても、その多くが正しく認識できていることが分かる。この結果から、局所的に RANSAC を用いることにより、本来 RANSAC では認識が困難な同一文字の重複を含む場合にも正確な認識を行えることが確認できた。

最後に RANSAC アルゴリズムが有効に働かなかった場合について考察する。RANSAC は、対応関係の得られた 3 点をランダムに選んでアフィン変換行列を算出した後、その妥当性を評価するために最初の 3 点以外の点を用いている。今回の実験では RANSAC による評価の際、閾値として最初の 3 点以外の 2 点で変換が妥当でない場合は不適と判断した。そのため、得られる特徴点数が絶対的に少ない文字や、射影変換を受けて特徴点数が減った文字などでは正しく対応付く特徴点の数が 5 個に満たず、特徴点数の少なさから RANSAC がうまく機能しないという問題が発生した。この問題を解決するためには、特徴点数が少ない場合には RANSAC を使わないようにするか、あるいは文字からより多くの特徴点を得られるような局所特徴量抽出・記述法を用いるといった方法が考えられる。この問題に関連して、表 3 に示されるように、ひらがなとカタカナの認識率が漢字に比べて低いという問題がある。これは文字の複雑さ故に漢字と比べてひらがなとカタカナから得られる特徴点数が少ないという事に起因している。同様に漢字の中にも“了”や“一”のように形状が単純で認識に失敗した例が見られた。

5. まとめ

本稿では、情景画像中の文字認識を目的として、局所特徴の位置関係を利用することでこれを達成することを考えた。今回は局所特徴量抽出・記述法として SIFT を用いて、参照画像と入力画像とで対応する点を求め、その位置関係を利用して文字領域の特定と認識を同時に行った。その際には提案手法として、局所的に RANSAC を用いることで文字領域特定の精度を上げることが考えた。風景画像上に書かれた文字の認識実験により、実際にその有効性が確認でき、漢字の認識では、95% を越える認識率が得られた。今後の課題として、実際の情景画像中に書

かれた文字で認識実験を行うことが挙げられる。また、SIFT によって得られる局所特徴量が射影変換を受けた文字の認識にもある程度有効であることが確認できた。実験の結果から、認識率低下の主な原因として単純な形状の文字から得られる特徴点数の不足が挙げられ、今後は文字からより多くの頑健な特徴を得るための手法を考案することが課題となる。他にも、フォントの変化に対する頑健性も調査の必要があると思われる。また、実時間性を持たせるための処理の高速化も今後の課題として挙げられる。

謝辞 本研究の一部は科研費補助金 (21700202) ならびに科学技術振興機構 CREST の補助による。

文 献

- [1] 辻 智彦, 岩村雅一, 黄瀬浩一, “カメラで撮影した単語画像の実時間認識,” 画像の認識・理解シンポジウム (MIRU2010) 論文集, pp.247–254, jul 2010.
- [2] 嶺 竜治, 亀山達也, 高橋寿一, 古賀昌史, 緒方日佐男, “文字認識と単語レイアウト解析を用いた紙文書とデジタルデータの情報リンク手法,” 電子情報通信学会論文誌, vol.J92-D, no.6, pp.868–875, 2009.
- [3] G.K. Myers, R.C. Bolles, Q.-T. Luong, J.A. Herson, and H.B. Aradhye, “Rectification and recognition of text in 3-d scenes,” IJ DAR, vol.7, no.2–3, pp.147–158, 2004.
- [4] Y. Kusachi, A. Suzuki, N. Ito, and K. Arakawa, “Kanji recognition in scene images without detection of text fields—robust against variation of viewpoint, contrast, and background texture—,” Proc. ICPR2004, 2004.
- [5] L. Li and C.L. Tan, “Recognizing planar symbols with severe perspective deformation,” IEEE TPAMI, vol.32, no.4, pp.755–762, April 2010.
- [6] X. Chen, J. Yang, and A. Waibel, “Automatic detection and recognition of signs from natural scenes,” IEEE Trans. Image Processing, vol.13, no.1, pp.87–99, Jan. 2004.
- [7] D.G. Lowe, “Distinctive image features from scale-invariant keypoints,” IJCV, vol.60, no.2, pp.91–110, 2004.
- [8] S. Uchida and M. Liwicki, “Part-based recognition of hand-written characters,” Proc. ICFHR2010, pp.545–550, 2010.
- [9] T.E. de Campos, B.R. Babu, and M. Varma., “Character recognition in natural images,” VISAPP, pp.273–280, Feb. 2009.
- [10] T. Wu, K. Qi, Q. Zheng, K. Chen, J. Chen, and H. Guan, “An improved descriptor for chinese character recognition,” Proc. Third International Symposium on Intelligent Information Technology Application, pp.400–403, 2009.
- [11] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman, “Object retrieval with large vocabularies and fast spatial matching,” Proc. CVPR2007, pp.1–8, 2007.
- [12] M.A. Fischler and R.C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” Commun. ACM, vol.24, no.6, pp.381–395, June 1981.
- [13] M. Brown and D. Lowe, “Automatic panoramic image stitching using invariant features,” IJCV, vol.74, no.1, pp.59–73, 2007.
- [14] S. Arya, D.M. Mount, R. Silverman, and A.Y. Wu, “An optimal algorithm for approximate nearest neighbor searching in fixed dimensions,” Journal of the ACM, vol.45, no.6, pp.891–923, Nov. 1998.