

ピクトグラムマッチングのための 輪郭情報を取り入れた局所形状記述子

上西 くるみ^{†1} 青木 輝勝^{†1†2}

概要: 標識など身近に多く存在しているピクトグラム（単純な絵記号）のマッチングには広く実用的な応用が考えられる。しかし、既存の自然画像マッチング技術では単純な画像の認識は大変難しく、撮影条件の変化はそれをさらに困難にしている。本稿では、この問題を解決するために、輪郭と凸包の関係をを用いた撮影条件の変化に頑強な局所形状記述子を提案する。局所特徴量にすることで撮影条件の変化に柔軟な形状記述子とし、輪郭と凸包の関係をを用いることによって個々のピクトグラム独自の特徴量記述を可能にした。提案方式は、既存手法と比較してピクトグラムのマッチ率、計算時間をともに大幅に向上させるものである。

キーワード: 画像マッチング、ピクトグラム、形状記述子、局所特徴量

Local Shape Descriptor with Contour Information for Pictogram Matching

KURUMI KAMINISHI^{†1} TERUMASA AOKI^{†1†2}

Abstract: Pictogram, which is a simple picture-based symbol, is widely used to indicate important facilities (such as “rest room”) or important rules (such as “no smoking”). A lot of new application will be generated if a practicable pictogram matching method is generated. Unfortunately current natural image matching methods are not suited for this purpose because pictogram contains less information than natural images. Furthermore, the matching results strongly depend on photographing conditions. In this paper, we present a novel pictogram matching method which is robust to the changes of photographing conditions. Our method uses relationship between contour and convex hull, and extracts local features from a pictogram to represent shape descriptors. The experiments show our method improves the matching rate and computation time for existing methods.

Keywords: Image Matching, Pictogram, Shape Descriptor, Local Feature

1. はじめに

画像マッチングはコンピュータビジョンにおいて基盤となる技術の一つである[1]。画像検索、物体認識、画像分類、3D オブジェクトの作成など、画像マッチングの用途は幅広い。画像マッチングは対象とする2つの画像間の相関関係を見つける技術であるが、撮影環境が異なる画像間の照合には様々な課題が残されている。

本研究の対象であるピクトグラム (pictogram) とは、「絵文字」という意味であり、簡単な絵の記号のことを示す。一般に、ピクトグラムは単調な色・単純な図形で表されているため、特徴の少ない画像である。そのため、人の目で見て確認することは容易であるが、画像マッチングにおいては含まれる情報が著しく少ないために逆に困難度を高めている。ピクトグラムのマッチングが可能になれば、交通

標識、企業ロゴ、看板の絵記号などがコンピュータ上で読み取ることができるようになる。この技術により、カーナビゲーションや携帯カメラとの連携などへの応用が実現する。携帯カメラで企業のロゴや商品を表す記号を読み取ることができると、QR (Quick Response) コードの代わりにピクトグラムでデザイン性を損なわずに URL 利用ができるようになる。これらの例のように、ピクトグラムマッチングは応用先が多い重要な技術である。

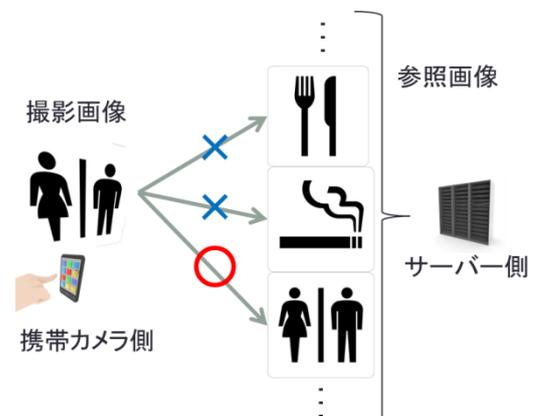


図1 ピクトグラムマッチングの概要

^{†1} 東北大学情報科学研究科
TOHOKU University Graduate School of Information Sciences
^{†2} 東北大学未来科学技術共同研究センター
TOHOKU University New Industry Creation Hatchery Center

自然画像を対象とした場合、マッチングを行うにあたり、現在 SIFT[2]や SURF[3]などの局所特徴量が広く用いられている。これらの局所特徴量は複雑な画像（自然画像）に対応できる記述子を持っている。しかし、既存の局所特徴量をピクトグラムに適応させると、ピクトグラムの特徴の少なから著しくマッチング精度が劣化してしまう。

近年、上述の局所特徴量とは別の流れとして、物体認識の重要な手がかりである形状に関する形状記述子の開発に大きな注目が向けられている。形状記述子は物体の形状の特徴化を目的としている特徴記述子であり、領域ベースの手法[4][5][6][7]、輪郭ベースの手法[8][9][10][11][12][13]の主に2つのカテゴリーに大別される[6]。輪郭ベースの手法では形状の輪郭や輪郭周辺の情報から特徴量を得る。最新の論文[12][13]では節変換やノイズ、射影変換に強い形状記述子が提案されている。しかし、輪郭ベースの手法は1つの閉じた輪郭上から情報を得るアルゴリズムのため、複数の輪郭の集合から成っているピクトグラムや、輪郭は同じで内部情報のみ異なる交通標識などに用いることができない。一方、領域ベースの手法では形状の領域全体から得られるグローバルな情報を用いる。この分野の代表的な手法である、Cross Ratio Spectrum (CRS)[7]、Characteristic Number (CN) [6]はいずれも形状全体で1つの特徴ベクトルを算出する大域特徴量である。これらの手法は射影変換に頑強であることを目指しているが、大域特徴量は局所特徴量と異なり、オクルージョンにより認識精度が大きく落ちてしまうという問題点がある。ピクトグラムマッチングを実用化するにあたり、撮影条件の変化に頑強であることは極めて重要な要素である。さらに、領域ベースの手法では輪郭情報を用いていない。輪郭情報は元々情報が少ないピクトグラムの貴重な情報であることから、輪郭情報と内部情報の両方を取り入れた形状記述子が、ピクトグラム独自の特徴を記述するのに適していると考えられる。

我々は、以上の異なる撮影条件下でのピクトグラムマッチングを可能にするために、輪郭情報として輪郭と凸包の関係を取り入れた局所形状記述子を提案する。提案手法は、形状の輪郭・内部領域全体の情報を得ながら、それらを局所特徴量として記述し、局所情報毎のマッチングを行うことによりピクトグラムマッチングを実現する。

2. 既存研究とその課題

2.1. Cross Ratio Spectrum (CRS)

Cross Ratio Spectrum (CRS)[7]は、Li と Tan によって提案された形状記述子である。ピクトグラム1つを、複比によって求められる1つのスペクトルで表現する。文献[7]では、CRS が様々な射影変換に強く、標識などの形状や色の構成が似ているピクトグラムにも適応できるということが示さ

れている。

CRS は、ピクトグラムの凸包上のある点 P_i と点 P_{i+1} から点 P_{i-1} までの凸包上の全ての点による直線を引き、その直線とピクトグラムの交点の初めの2点を用いて複比により特徴量を計算する。そして、射影不変量である複比によって得られた特徴量によって作られるスペクトル Cross Ratio Spectrum (CRS)によってピクトグラム特有の形状を記述する。

しかし、CRS には2つの問題がある。1つは、内部情報利用が不十分なため認識精度が落ちてしまうということである。上で説明したように、CRS ではピクトグラムの凸包内部の交点を始めの2点のみ用いる。凸包に近い2点の交点のみを使用するため、ピクトグラムの凸包内が複雑である場合、ピクトグラムの中心部の構造情報が特徴量に組み込まれない。このことから、内部構造が複雑なピクトグラムや中心部だけ異なる全体が似たピクトグラムについて認識率が低くなってしまふ。2つ目の問題は、計算時間が非常に大きいということである。CRS は、特徴量であるスペクトルを比較するためにダイナミックタイムワーピング (DTW) を用いている。DTW による計算コストは、対象とするピクトグラムの凸包上のピクセルの数に依存し、画素数が大きいほどより大変大きな時間がかかってしまう。以上の2つの問題から、CRS を実用アプリケーションに応用することは難しい。

2.2. Characteristic Number (CN)

Characteristic Number (CN) [6]は、Luo らによって提案された形状記述子である。1つのピクトグラムを、多くの内部構造情報を組み込んだ1つの特徴ベクトルで表現する。文献[6]では、このCNが従来手法であるSC(shape context) [9]や、前節で説明したCRS (cross ratio spectrum) [7]よりも射影変換に頑強で、実行時間も早いことが述べられている。

CN は、凸包上の等間隔の点を P_i 、線分 P_iP_j と内部構造との交点を $Q_i^{(j)}$ として、

$$Q_i^{(j)} = a_i^{(j)}P_i + b_i^{(j)}P_j \quad (1)$$

$$CN(P_i, P_j, P_k) = \prod_{i=1}^3 \prod_{j=1}^N \left(\frac{a_i^{(j)}}{b_i^{(j)}} \right) \quad (2)$$

と定義される。ここで、Nは各3線分上の交点の数のうち、一番少ない数である。文献[6]では、線分 P_iP_j (j は i 以外の任意の数) に対し、 $N=2$ の時、CNの値は点 P_i 、 $Q_i^{(1)}$ 、 $Q_i^{(2)}$ 、 P_j からなる複比に等しいとし、CNは複比の拡張であるため、射影不変であると述べている。

CN 特徴量の具体的な算出アルゴリズムは以下の通りである。

1. 対象の凸包上に反時計回りで n 個の等間隔のサンプル点 P_i を取る。

2. 三角形を形成するように点 P_i, P_j, P_k を選ぶ。
ここで、3点在同一線上にあるとき $CN=0$ とする。
3. 三角形の各辺が内部形状と交差する点 $Q_i^{(j)}$ を抽出し、それらの点から CN 値を計算する。
4. 全ての三角形の組み合わせで CN 値を計算し、CN 値を連結することで特徴ベクトルとする。

$$\text{Descriptor} = \left(CN(P_i, P_j, P_k) \right)_{1 \times n C_3} \quad (3)$$

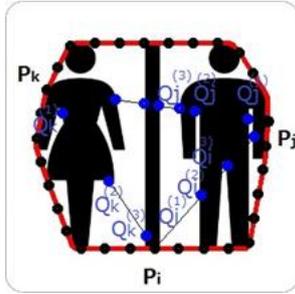


図2 CNの計算例

CNにおける2枚の画像の類似度は、ヒストグラム交差法により算出する。つまり、形状 Q と T の類似度 S は、正規化された特徴量 $\tilde{D}(Q), \tilde{D}(T)$ を用いて以下の式で求められる。

$$S = \text{sum}(\min(\tilde{D}(Q), \tilde{D}(T))) \quad (4)$$

以上の手法から、DTWを用いるCRSと比較し、ベクトル記述の位置を合わせるために、凸包上のサンプル点の始点のみの位置合わせを行えば良いので、CRSと比較して計算時間が少ない。また、内部情報を多く取り入れることができるため、CRSよりも精度が高く、射影変換を伴うピクトグラム認識に有効であると考えられている。

2.3. CNの問題点

2.2ではCNについて概説したが、この記述方式は大きな問題点が3つある。

A. CN値

1つ目の問題は、文献[6]で提案されているCNの特徴量は、複比の定義と異なっており、射影不変量でないということである。CNは射影変換に対応するために複比を用いることを目指しているが、実はCNでは複比ではなく、一直線上の3点間の距離の比が用いられている。このため、射影変換を伴ったピクトグラムにCNを適応させると、著しくマッチング精度が劣化してしまう。式(1)と式(2)に注目すると、辺の比 $a_i^{(j)}, b_i^{(j)}$ を算出するために用いられている点は、点 $Q_i^{(j)}, P_i, P_j$ の3点である。一直線上の3点間の距離の比はアフィン不変量であり、射影不変量ではない。そのため、CNの定義にしたがった特徴量(CN値)は射影不変量ではない。

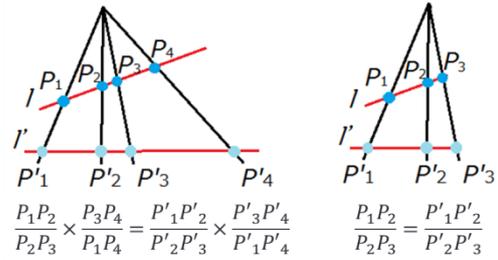


図3 真の複比の定義(左)と

CNで用いられている複比の定義(右)

B. 大域特徴量

2つ目の問題は、1つのピクトグラムを1つの特徴量で表す大域特徴量であるため、射影変換・オクルージョンに弱いことである。射影変換について具体的には、凸包上のサンプル点 P_i の位置が射影不変でない。CN特徴量は対象の凸包上に反時計回りで等間隔にとったサンプル点 P_i を用いている。このサンプル点の位置は、射影変換により対象が変形することで簡単に位置が変わってしまう。よって、正しく照合することができない場合が多く生じる。例えば、図4のように大きな射影変換が起こってしまうと、サンプル点 P_i の位置がずれ、それにより得られる交点の数と交点とサンプル点間の距離が変化してしまう。このことが原因で、違ったサンプル点の位置のまま算出されたCN値とそれを連結した大域特徴量である特徴ベクトル全体が射影変換前後で大きく異なったものになってしまう。



図4 参照画像(左)と射影変換画像(右)の
サンプル点(凸包上黒点)の例

また、オクルージョンのようにピクトグラム内にノイズが入ってしまうと特徴量全体に影響が及び、異なる値になってしまうため、認識することができない。

C. 計算時間

認識率以外の問題として、計算コストが高いことがあげられる。CNは大域特徴量ということができるので、回転変換でサンプル点の位置が変わってしまったら、特徴ベクトルの記述の順番が変わってしまうことで特徴量全体が大きく変化する。これに対応するために全てのサンプル点を始点としてマッチングを複数回行うため、計算時間が大きくなってしまいます。文献[6]では、サンプル点の始点の位置合わせのみで比較が可能となり、CRSより大幅に計算時間が早くなったと述べられているが、凸包上のサンプル点の数(n)が大きくなればなるほど始点の候補と比較回数が増え、特徴量比較に多くの時間がかかってしまう。さらに、特徴

量算出のためにピクトグラム内に三角形を形成していることで、三角形の各1辺のみに注目すると何度も同じ直線上の交点による特徴量が計算されているため、計算時間が冗長である。論文[6]の実験結果からも SIFT の計算時間に対して約2倍の計算時間となっていることが分かる。

以上の理由から、CN は射影変換やオクルージョンなどの撮影条件伴ったピクトグラムの認識に弱く、計算時間が多くかかることから実用的なアプリケーションに用いることは難しい。

3. 提案手法

3.1. 提案局所形状記述子の概要

本章では、提案特徴量である輪郭情報を取り入れた局所形状記述子について説明する。以下、図5に示す提案手法の概要に沿って、撮影ピクトグラムの入力から結果の出力までについて述べる。なお、提案手法を構成する3つの提案手法、“CRN (Cross Ratio Number)”、“ピクトグラムに有効な局所特徴量”、“輪郭と凸包の関係による特徴化”については次節以降で詳しく説明する。

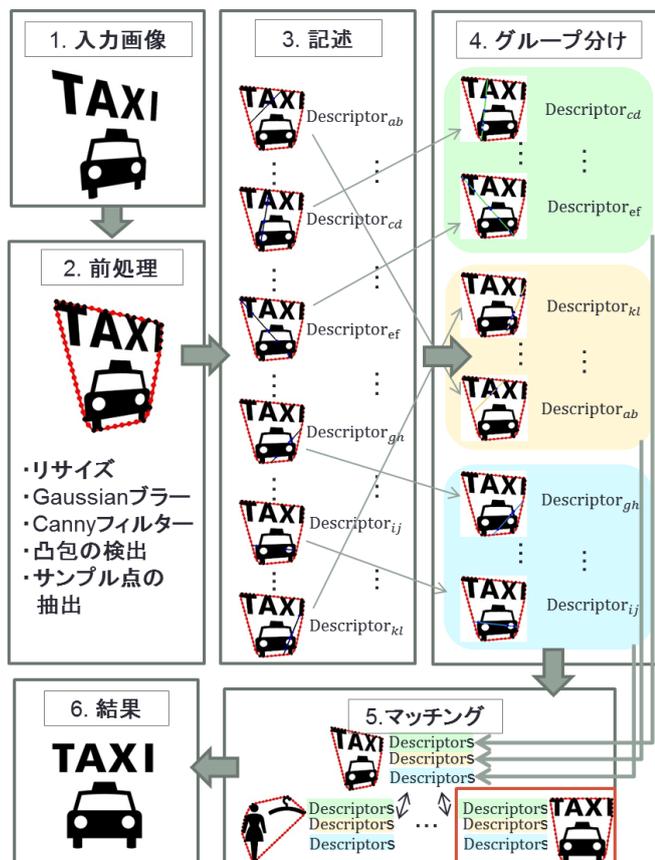


図5 提案手法の概要

1. 入力画像

光学カメラでピクトグラムを撮影することを前提とする。この時、自然画像からピクトグラム領域のみを取り出したものを入力画像とする。

2. 前処理

スケールの正規化のために、リサイズを行う。次に、細かいノイズを除去するためにガウシアンブラーをかける。そして、内部構造の検出のためにキャニーフィルタによりエッジを抽出する。このエッジを基に凸包を描き、凸包上に等間隔の n 個のサンプル点を取る。この時、サンプル点が位置する凸包がピクトグラムの輪郭に接しているか接していないかで場合分けを行う (図5の凸包上のサンプル点の色が赤: 輪郭にサンプル点が接していない、黒: 輪郭にサンプル点が接している)。

3. 記述

異なるサンプル点を2点選択し、その2点をつなぐ直線を引く。その直線と内部構造の交点を基に CRN 値を算出する (3.2節参照)。複数の交点から求めた複数の CRN 値を連結したものが1つの局所特徴量となる (3.3節参照)。これを異なるサンプル点2点を選択する全ての組み合わせで行い、 nC_2 個の局所特徴量を記述する。

4. グループ分け

輪郭と凸包の関係による特徴化のために、特徴ベクトルのグループ分けを行う (3.4節参照)。凸包上のサンプル点が輪郭に接している数 (0 から 2) によって3つのグループに分ける。

5. マッチング

参照画像補の局所特徴量と入力画像の局所特徴量についてそれぞれのグループ毎にマッチングを行い、全てのグループでのマッチング数の合計を参照画像と入力画像のマッチング数の結果とする。マッチングは入力画像の特徴量と参照画像の特徴量の双方から距離が一番近いものを選択するクロスマッチングとし、距離は L1 ノルムで算出する。

6. 結果

全ての参照画像で5と6を行い、入力画像の特徴量とのマッチング数が一番大きい参照画像を結果として出力する。

表1に提案特徴量と既存手法 CN との違いをまとめる。

表1 提案特徴量と既存特徴量の構成

	提案特徴量	CN[6]
特徴量記述方式	局所特徴量	大域特徴量
特徴値	CRN	CN
特徴量算出に用いる構造	直線	三角形

3.1. Cross Ratio Number (CRN)

2.3 節の A の問題を解決し、射影変換に不変な特徴量とするために、正しい複比を基礎とした特徴値 CRN (Cross Ratio Number) を定義する[14]。

$$CRN(Q_{ij}^{(m)}, Q_{ij}^{(n)}) = \frac{P_i Q_{ij}^{(m)}}{Q_{ij}^{(m)} Q_{ij}^{(n)}} \times \frac{Q_{ij}^{(n)} P_j}{P_i P_j} \quad (5)$$

ここで、 $CRN(Q_{ij}^{(m)}, Q_{ij}^{(n)})$ は、点 $P_i, P_j, Q_{ij}^{(m)}, Q_{ij}^{(n)}$ の 4 点から得られる CRN 値を示す。点 P_i, P_j は凸包上にとった基準点、点 $Q_{ij}^{(m)}, Q_{ij}^{(n)}$ は線分 $P_i P_j$ と内部構造との m 番目、 n 番目の交点

($m \neq n$) である。既存手法 CN のように一直線上から 3 点選ぶのではなく、一直線上から 4 点選んで 1 つの計算を行うことで、複比の正しい式に適応させることができる。こうすることで、計算に射影不変量が用いられ、射影変換前後で等しい特徴値が得られる。

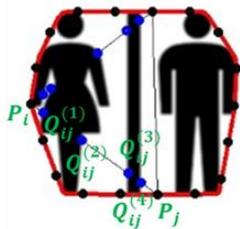


図 6 CRN 算出に用いる記号例

提案手法は、特徴値計算のためのピクトグラム内部交点検出にピクトグラムの凸包上のサンプル点 2 点から成る直線を用いている。一方、CN ではサンプル点 3 点による三角形を特徴値算出に用いているため、2.3 節の C の計算時間に関する問題が生じる。三角形ではなく、直線を用いることで、1 直線あたりの特徴量計算を 1 度のみとすることができる。

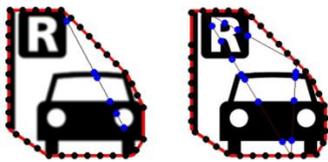


図 7 特徴量算出に用いる構造の例
(左：提案特徴量、右：CN[6])

3.2. ピクトグラムに有効な局所特徴量

提案特徴量における既存の形状記述子と最も異なる点が、局所特徴量だということである。既存形状記述子 CN は計算した特徴量を結合することによって、1 つのピクトグラムから 1 つの特徴ベクトルを得ているが、提案局所特徴量では 1 つのピクトグラムから特徴ベクトルを複数個検出する。局所特徴量として記述し、それぞれの特徴ベクトルごとにマッチングすることによって、射影変換によるサンプル点のずれ、回転変換によるサンプル点の始点のずれ、ピクトグラム内のオクルージョンに関わらず認識が可能に

なる。さらに、局所特徴量にすることで形状の位置合わせが不要になり、計算時間も早くなる。

提案ローカル特徴量は、以下のように定義される。

- 凸包上のサンプル点 2 点をつなぐ直線で 1 つの特徴ベクトルとする。
- 1 つのベクトルを M 次元の特徴ベクトルで表す。
- 直線と内部構造とのある 2 点の交点と、2 点のサンプル点から得られる CRN 値を 1 つの要素とする。
- 直線 $P_i P_j$ の特徴ベクトル Descriptor $_{ij}$ を以下の式で示す。

$$Descriptor_{ij} = (CRN(Q_{ij}^{(1)}, Q_{ij}^{(2)}), \dots, CRN(Q_{ij}^{(a-1)}, Q_{ij}^{(a)}), 0, \dots)_M \quad (6)$$

ここで、直線と内部構造の交点の数を a とすると、特徴ベクトルに入れられる CRN 値の順番は、1 番目の交点と 2 番目から a 番目の交点により得られる CRN 値、その次に 2 番目の交点と 3 番目から a 番目の交点により得られる CRN 値と続け、最後は $(a-1)$ 番目と a 番目の交点により得られる CRN 値が入る。なお、直線と内部構造の交点の組み合わせは最大 M 個使用し、 M 個より少ない場合は余った要素に 0 が入れられる。

式(6)を凸包上のサンプル点から構成される全ての直線に適応させることによって、1 つのピクトグラムにつき C_2^n 個の特徴ベクトルが得られる。これらの集合が提案特徴量である。

以上の 3 つの提案手法を取り入れた、提案特徴量のアルゴリズムを以下に示す。

1. 対象の凸包上に反時計回りで n 個の等間隔のサンプル点 P_i を取る。
2. 直線を形成するように点 P_i, P_j を選ぶ。
3. 直線が内部形状と交差する点 $Q_{ij}^{(m)}, Q_{ij}^{(n)}$ を抽出し、それらの点から CRN 値を計算する (式(5))。
4. 3 で 1 直線から算出した複数の CRN 値をそれぞれ 1 つの要素として特徴ベクトルを記述する(式(6))。
5. 2 から 4 を全ての直線の組み合わせで行い、その集合をピクトグラムの特徴量とする。

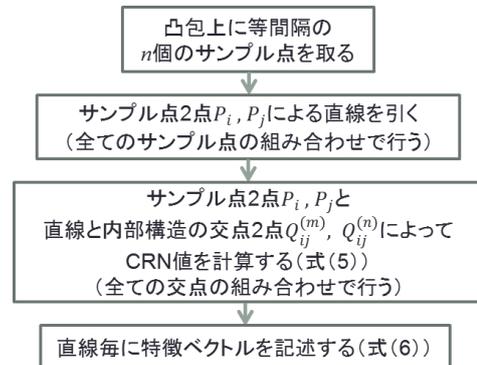


図 8 提案局所特徴量算出アルゴリズム

3.3. 輪郭と凸包の関係による特徴化

局所特徴量を用いることで撮影条件の変化に柔軟な形状記述子となったが、一般に大域特徴量に比べて局所特徴量1つ1つに含まれる情報が少ない。そこで、局所特徴量をピクトグラム毎により特徴化することを狙いとして、特徴量に“凸包上のサンプル点が輪郭と接しているか接していないか”という条件を輪郭情報として付与する。具体的には、1つの直線から記述される1つの局所特徴量を以下の3つのグループに分ける。

- 0: サンプル点が2点とも輪郭に接しない。
- 1: サンプル点が1点だけ輪郭に接する。
- 2: サンプル点が2点とも輪郭に接する。

グループ分け後、参照画像と入力画像について各グループ内でマッチングを行い、各グループのマッチング数の合計を参照画像とテスト画像のマッチング数とする。(図9)。輪郭と凸包の関係を用いることによって、グループ外での誤ったマッチングを減らすことができ、よりピクトグラムを特徴化する局所特徴量とすることができる。さらに、グループ内でのみマッチングを行うため、グループが異なる特徴量の無駄なマッチングを減らすこともできる。

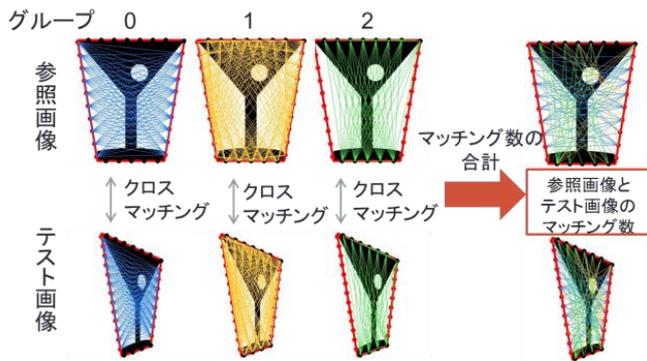


図9 グループ毎のマッチング



図10 グループ分けによるマッチングの様子

4. 評価実験

4.1. 実験準備

A. データセット

評価実験には、ピクトグラム125枚を参照画像として使用した。ピクトグラムは標準案内用図記号[15]の中から同じ意味のものを除く全125枚を用いた。テスト画像については以下の3つの撮影条件の変化を伴うピクトグラムを用いた。

- ・射影変換：画像の中心点を通る、画像水平面上の縦軸、横軸を回転軸にして、水平面を 0° として、それぞれ 20° 、 40° 、 60° に回転させた9通りの射影変換を伴うピクトグラム1125枚(125枚×9通り)
- ・オクルージョン：ピクトグラムを囲む凸包の面積に対してそれぞれ $1/8, 1/16, 1/32$ の大きさの面積の障害物(四角形)を張り付けた3通りのオクルージョンを伴うピクトグラム375枚(125枚×3通り)
- ・回転変換：画像の中心点を回転点にして、画像水平面上を時計回りで 15° 、 30° 、 45° 、 60° 、 75° に回転させた5通りの回転変換を伴うピクトグラム625枚(125枚×5通り)



図11 テスト画像の例

(左から、射影変換、オクルージョン、回転変換)

以上のテスト画像に伴った撮影条件の変化の強さは実際に伴う変換を考慮した十分な範囲と言える。例えば、縦軸横軸共に 60° ずつ射影変換した画像は、およそ地面から6.5mの高さにある看板を、1.5mの高さにあるカメラで斜め方向に約4.2m離れた場所から撮影した場合に相当する変換画像である。

また、参考としてSIFTの自然画像の計算時間を計測するために、自然画像20枚を用意した。尚、全ての画像のスケールは 300×300 ピクセル、拡張子はjpgである。

B. 使用手法

本実験において用いる手法は以下の4つである。

- 手法1) 提案局所特徴量
- 手法2) 既存大域特徴量 CN[6]
- 手法3) 既存局所特徴量 SIFT[2]
- 手法4) 提案特徴量の“輪郭と凸包の情報”をなくした局所特徴量

尚、手法1.2.4における凸包上のサンプル点の数は $n=54$ 、手法1.4における特徴ベクトルの次元数は $M=36$ とした。

1から3の手法を比較することで提案特徴量の有効性を

示す。また、1と4を比較することで、提案特微量における輪郭情報の付与の有効性を示す。

C. 評価基準

本実験の手法評価基準には、マッチ率と計算時間を用いる。マッチ率を計算するため、はじめに、テスト画像と全ての参照画像の類似度を求め、テスト画像との類似度が一番高い参照画像をマッチング結果とする。次に、その結果が正しい画像であるか、異なる画像であるかを判断し、ある撮影条件のマッチ率を算出する。類似度は、局所特微量を用いる手法では参照画像と入力画像の特微量双方から距離が一番近いものを選択するクロスマッチング（手法1、4ではL1ノルムを、手法3ではL2ノルムで距離を算出した）によるマッチング数、大域特微量を用いる手法（手法2）ではヒストグラムインターセクション法を用いて算出する。

計算時間は”1つのピクトグラムから特微量を計算する時間”を、全参照画像125枚を用いて測定し、その平均時間を手法毎に比較する。

D. 実験環境

評価実験に用いたPCの環境は以下の通りである。

- ・CPU : Intel(R) Core(TM) i3-3240 CPU 3.40GHz
- ・メモリ (RAM) : 12.0GB

4.2. 射影変換を伴うピクトグラムの認識率評価

本実験では、提案特微量と“輪郭と凸包の情報”の付与の、射影変換を伴うピクトグラムに対する有効性を評価し、既存手法CNと比較するために手法1,2,4を用いた。手法3のSIFTは、射影変換に対応できないため除いた。結果を以下に示す。

表2 射影変換を伴うピクトグラムの認識率[%]

手法	縦/横	20°	40°	60°	全画像
提案特微量 (手法1)	20°	100.0	97.6	90.4	94.8
	40°	100.0	97.6	92.8	
	60°	88.8	90.4	95.2	
CN[6] (手法2)	20°	78.4	52.8	25.6	41.7
	40°	55.2	43.2	25.6	
	60°	37.6	32.8	24.0	
"輪郭情報" なし (手法4)	20°	98.4	90.4	76.8	86.5
	40°	91.2	92.0	80.8	
	60°	76.8	83.2	88.8	

以上の結果から、全画像を通して提案手法により、射影変換を伴った画像の認識率が、CNより53.1%と大幅に上がった。特に、既存手法では変換角度の大きいものほど大きく認識率が下がっているが、提案特微量では認識率を88.8%以上に高く保てていることから、射影変換への頑強性があると言える。この結果からCNよりもCRNの射影変換への頑強性が高いことが分かる。また、局所特微量として用いることで、2.3節のBの問題である、射影変換による凸包上のサンプル点の位置のずれの問題も改善されたと

考えられる。手法4では、輪郭と凸包との関係を付与しないことで、ピクトグラムの局所特微量の独自性が高まらず、認識率が提案特微量より悪くなった。この結果から、提案特微量内の“輪郭と凸包の情報”が有効であり、手法1が最も有効であることが示される。

4.3. オクルージョンを伴うピクトグラムの認識率評価

本実験では、オクルージョンを伴うピクトグラムに対して、局所特微量である提案特微量とSIFT、大域特微量であるCNを比較し、“輪郭と凸包の情報”の付与という提案手法の有効性を実証するために手法1,2,3,4を用いた。結果を以下に示す。

表3 オクルージョンを伴うピクトグラムの認識率[%]

手法	オクルージョンの大きさ			全画像
	1/8	1/16	1/32	
提案特微量(手法1)	99.2	100.0	100.0	99.7
CN[6](手法2)	90.4	93.6	92.8	92.3
SIFT[2](手法3)	81.6	85.6	90.4	85.9
"輪郭情報" なし(手法4)	88.8	95.2	98.4	94.1

以上の結果から、全画像を通して提案手法により、オクルージョンを伴った画像の認識率が2つの既存手法より高いことが分かる。提案特微量は局所特微量であることから、オクルージョンによって変化した箇所を無視し、正しい部分のみマッチングできるため、CNより良い結果となった。さらに、手法4の結果が手法1よりも悪いことから、射影変換の実験と同様に、輪郭と凸包の情報の付与によってピクトグラムの局所特微量が特徴化されていると言える。また、SIFTは局所特微量であるが、1章で述べたように、特徴の少ないピクトグラムでは特徴点の数が少なく、異なる特徴点とマッチングさせてしまうという2つの欠点により認識率が低くなった。以上の結果から、“輪郭と凸包の情報”がオクルージョンにおいても効果的であり、提案特微量がオクルージョンを伴うピクトグラムに既存手法よりも有効であることが分かる。

4.4. 回転変換を伴うピクトグラムの認識率評価

本実験では、回転変換を伴うピクトグラムを提案特微量で認識することができるかを確かめるために、手法1を用いた。既存手法CNとSIFTについてはそれぞれの論文で回転変換に対する頑強性が既に示されているため除いた。結果を以下に示す。

表4 回転変換を伴うピクトグラムの認識率[%]

提案特微量 (手法1)	回転角度					全画像
	15°	30°	45°	60°	75°	
	100	96.8	96.0	99.2	100	98.4

以上の結果から、提案手法により全画像を通して高い認識率を得ることができた。回転角度の大きさに認識率が依存していないことから、提案特徴量は回転変換に関しても有効である。

4.5. 計算時間比較

本実験での使用手法は、4.1節の手法1,2,3である。また、通常のSIFTの特徴記述の対象である自然画像20枚も参考のため使用し、SIFTで算出される特徴点の数も測定した。

“1つのピクトグラムから特徴量を計算する時間”の平均と、SIFTで検出できる特徴点数の平均の結果を以下に示す。

表5 特徴量計算の平均時間と平均特徴点数

画像	手法	平均時間[ms]	平均特徴点数
ピクトグラム	提案特徴量 (手法1)	223.4	
	CN[6] (手法2)	10074.2	
	SIFT[2] (手法3)	325.8	67.2
自然画像		550.3	477.3

以上の結果から、提案手法による特徴量計算時間が既存形状記述子CNと比較して約1/45という、圧倒的な早さであることが分かる。主な理由としては、3.1節で説明した直線を用いた特徴量算出により、計算の冗長性を取り除いた効果だと考えられる。また、輪郭と凸包の関係の付与によってグループを分けてマッチングを行うことで無駄なマッチング時間も無くすことができたということも理由の1つである。さらに、SIFTについて、ピクトグラムから十分な特徴点数を取ることが難しく、ピクトグラムマッチングの手法としてSIFTが不向きであることが分かる。自然画像におけるSIFTの平均時間よりも提案手法が約2.5倍早い結果となったことから提案手法の計算時間の早さを証明できた。

5. まとめ

本研究では、異なる撮影条件下でのピクトグラムマッチングを可能にするために、輪郭と凸包の関係を取り入れた局所形状記述子を提案した。提案手法は、形状の輪郭・部領域全体の情報を得、局所特徴量として扱うことで条件の変化に柔軟であることを狙いとしている。提案手法によって、射影変換を伴うピクトグラムに対して94.8%、オクルージョンを伴うピクトグラムに対して99.7%と既存手法よりも高い認識率が得られた。また、回転変換を伴うピクトグラムに対しても98.4%の認識率を得られたことから、様々な撮影条件の変化に強く、ピクトグラム固有の情報を持った特徴量だと言える。また、計算時間においても既存形状記述子よりも早く、実用レベルの時間となった。

参考文献

[1]YinanYu, Kaiqi Huang and Wei Chen," A Novel Algorithm for View

and Illumination Invariant Image Matching",IEEE Transaction on Image Processing, vol.21, issue.1, pp.229-240, 2012.

[2] D. G. Lowe, "Object recognition from local scale invariant features", The Proceedings of the Seventh IEEE International Conference on Computer Vision, Sep, 1999.

[3] Herbert Bay, Tinne Tuytelaars, Luc Van Gool, "Speeded-Up Robust Features," Computer Vision and Image Understanding (CVIU), EECV, May,2006

[4] R.B. Yadav, N.K. Nishchal, A.K. Gupta, and V.K. Rastogi, "Retrieval and classification of objects using generic fourier, legendre moment, and wavelet zernike moment descriptors and recognition using joint transform correlator," Optics & Laser Technology, vol. 40, no. 3, pp. 517-527, 2008.

[5] F. Ghorbel, S. Derrode, R. Mezhoud, T. Bannour, S. Dhahbi, "Image reconstruction from a complete set of similarity invariants extracted from complex moments", Pattern recognition letters, vol. 27, no. 12, pp. 1361-1369, 2006.

[6] Zhongxuan Luo, Daiyun Luo, Xin Fan, Xinchun Zhou, and Qi Jia, "A shape descriptor based on new projective invariants", IEEE International Conference on Image Processing (ICIP), 2013

[7] Linlin Li and Chew Lim Tan, "Recognizing Planar Symbols with Severe Perspective Deformation",IEEE Transaction on Pattern Analysis And Machine Intelligence, vol.32, issue.4, pp.755-762,2010

[8] Mirosław Bober, "MPEG-7 visual shape descriptors", IEEE Transactions on Circuits and Systems for Video Technology, pp.716-719, 2001.

[9] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 4, pp. 509-522, 2002.

[10] X.C. He, N.H.C. Yung, "Curvature scale space corner detector with adaptive threshold and dynamic region of support", Proceedings of the 17th International Conference on Pattern Recognition (ICPR). IEEE, vol. 2, pp. 791-794, 2004.

[11] Xiang Bai, Cong Rao, Xinggang Wang, "Shape Vocabulary: A Robust and Efficient Shape Representation for Shape Matching", IEEE Transactions on Image Processing, val.23, issue.9, pp3935-3949, 2014

[12] Haoran Xu, Jianyu Yang, Junsong Yuan, "Invariant multi-scale shape descriptor for object matching and recognition", 2016 IEEE International Conference on Image Processing (ICIP), pp644-648, 2016

[13] Qi Jia, Xin Fan, Yu Liu, Haojie Li, Zhongxuan Luo, He Guo, "Hierarchical projective invariant contexts for shape recognition", Pattern Recognition, Vol 52, pp358-374, 2016

[14] 上西くるみ,青木輝政, "高精度ピクトグラムマッチングのための射影変換に頑強な特徴量", 第202回コンピュータビジョンとイメージメディア研究発表会, 2016

[15] 公益財団法人交通エコロジー・モビリティ財団バリアフリー推進部,標準案内用図記号