

商標図形の複雑さを表わす特徴量について

3M-3

長嶋 秀世 早乙女 拓栄
工学院大学

1 まえがき

現在、特許庁に登録された商標の数は数百万件に及んでおり、文字列的、幾何学的、イラスト的、写実的などのさまざまな種類のものがある。これらの商標図形の類似性には、図形の形、内部の模様、意味・概念などの要因に影響されるものと考えられるが、その中でも特に、図形の形状による影響は大きいと思われる。このことから、本研究では、商標図形の形状による分類を行なうことによって、類似商標検索システムの精度と効率の向上を意図している。

商標図形の形状には、円や四角形などの単純なものから、動物や人を描いた複雑なものまで、さまざまな種類があり、図形の形状による分類を行うとき、形の複雑さは重要な特徴である。

本論文では、人の主観を反映した特徴量の一検討として、商標図形の形状の複雑さを表わす特徴量について検討したので報告する。

2 形状の複雑さを表わす特徴量

まず、図1にさまざまな形状の商標図形の例を示す。



図1. 商標図形

これらの図形を見たとき、形が複雑であると感じるかどうかは人によって異なると考えられる。そこで、アンケートを用いて人の平均的な主観を抽出し、これを目的変数として重回帰分析を行なうことによって、人の主観に対応した特徴量を作成する。以下にこの方法を述べる。

2.1 商標図形の形状の複雑さの評価

m 個のサンプル図形を用いて、 n 人の被験者に複雑な形状か否かを問うアンケートを行なう。商標図形 i を複

雑であると判断した人数を a_i として、指数 S_i を次式によって求める。

$$S_i = \frac{a_i}{n} \quad (1)$$

指数 S_i は被験者のうち、図形に特徴が表れていると判断した被験者の割合を表わしている。

2.2 重回帰分析を用いた主観の導入

(1) 集合 G に含まれる n 個の図形から m 個の画像特徴量 f_{ij} ($i = 1, 2, \dots, n; j = 1, 2, \dots, m$)を抽出し、これらを説明変数とする。この説明変数の行列式 F' を

$$F' = \begin{bmatrix} 1 & f_{11} & f_{12} & \cdots & f_{1m} \\ 1 & f_{21} & f_{22} & \cdots & f_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & f_{n1} & f_{n2} & \cdots & f_{nm} \end{bmatrix} \quad (2)$$

と表す。

(2) 集合 G に含まれる商標に対し、画像特徴量 f_{ij} を説明変数、指数 S_i を目的変数として線形重回帰分析を行う。ここで、ベクトル S 、回帰係数ベクトル a 、誤差ベクトル e をそれぞれ式(3)~(5)で表すと、回帰式は式(6)のように表される。

$$S = [S_1, S_2, \dots, S_n]^t \quad (3)$$

$$a = [a_0, a_1, \dots, a_m]^t \quad (4)$$

$$e = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n]^t \quad (5)$$

$$S = F'a + e \quad (6)$$

(4) 得られた回帰係数 a_0, a_1, \dots, a_m を用いて指数の予測式 \hat{S} をたてる。

$$\hat{S} = a_0 + a_1 f_1 + a_2 f_2 + \cdots + a_m f_m \quad (7)$$

(5) 予測値 \hat{S} を計算し、この値を形状の複雑さを表わす特徴量とする。

2.3 使用する特徴量

説明変数には形状の複雑さと関係が強い特徴量を用いる必要がある。形状の複雑さには、さまざまな要因が影響すると考えられるが、図形の凹凸が入り組んでおり、凹んでいる領域が多い図形ほど複雑であると考えられる。

また、凹凸の数が多くても、それらが規則的に配置される場合、人が見たときにはあまり複雑と感じないと思われる。これらのことから、使用する特徴量を以下の4つとする。

- 1) 凹凸の数 f_1
- 2) 凸包と原図形の周囲長比 f_2
- 3) 凹凸の周期性 f_3
- 4) 屈折点の数 f_4

以下に説明が必要と考えられる特徴量について、求め方を述べる。

2.3.1 凹凸の数

図のように図形の凹部を抽出し、それぞれの凹部に対して再帰的に凹部を求めていき、凹部がなくなるまで処理を繰り返す。ここで、階層 i で求められた凹部の数を x_i とすると、階層的に求めた凹凸の総数 f_1 は次式のように表わされる。

$$f_1 = \sum_{i=0}^N x_i \quad (i = 0, 1, 2, 3, \dots, N) \quad (8)$$

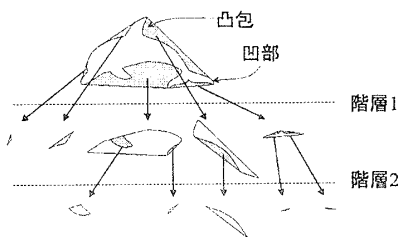


図2 凸包による階層的分解

2.3.2 凹凸の周期性

以下に凹凸の周期性を抽出する手順を示す。

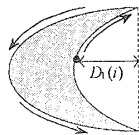


図3 凹部距離

(1) 原図形の輪郭線を追跡し、輪郭線上のすべての点について、凸包から原図形の輪郭線までの最小距離 $D(i)$ を求める。

(2) N を原図形の周囲長としたとき、 $D(i)$ の自己相関関数 $R(i)$ を次式によって求める。

$$R(i) = \frac{1}{R_0 \cdot N} \sum_{j=0}^N (D(j) \cdot D(j+i)) \quad (i = 0, 1, 2, \dots, N) \quad (9)$$

(3) $R(i)$ が極大となる点の相関値 $R_{max}(k)$ を求め、これらの最大値を凹凸の周期性を表わす特徴量 f_3 とする ($k=0,1,2,\dots,n$)。

$$f_3 = \max_{0 \leq k \leq n} R_{max}(k) \quad (10)$$

3 特徴量の評価実験

3.1 評価方法

まず、 S_i に対してしきい値 S_{th} を定め、 $S_i \geq S_{th}$ となる図形とそれ以外の図形に分類する。この結果は、分類の精度を評価する基準となる。

次に、複雑さを表わす特徴量 \hat{S} に対してしきい値 th を定め、この値との大小関係により分類を行なう。クラス i に分類されるべき図形の数を n_i 、特徴量によって正しく分類された図形の数を c_i として、しきい値 th における平均精度 E_{th} を次式によって求める。

$$E_{th} = \frac{1}{2} \sum_{i=0}^N \frac{c_i}{n_i} \quad (11)$$

ここで、しきい値 th を特徴量 \hat{S} の最小値 \hat{S}_{min} から最大値 \hat{S}_{max} まで変化させ、それぞれ平均精度を求める。このうち、最大となる平均精度 E_{th} を最大分類精度 E_{max} とし、これを最終的な評価値とする。

$$E_{max} = \max_{\hat{S}_{min} \leq th \leq \hat{S}_{max}} E_{th} \quad (12)$$

3.2 実験結果

上記の方法を用いて、商標図形480個を対象に、形の複雑さを表わす特徴量の有効性を評価する実験を行なった。アンケート結果から得られる指数に対するしきい値は、 $S_{th} = 0.5$ とした。分類結果の一部を図に示す。最大分類精度は、しきい値が $th = 0.06$ のときで、 $E_{max} = 0.944$ となり、本特徴量は図形の形状の複雑さを表わすのに有効であることが示された。



(a) 複雑な形状 (b) 単純な形状

図 分類結果

4 むすび

ここでは、人の主観を考慮した特徴量の一検討として、形状の複雑さを表わす特徴量について述べた。アンケートから抽出した指数を目的変数として重回帰分析を行なうことによって、人の主観に対応した特徴量を作成することができた。今後、作成した特徴量を用いて図形を分類する方法を検討する。