

3M-7 空間的構造特徴を用いた類似度学習 — 類似画検索への応用 —

加藤 俊一^{*)} 栗田多喜夫^{*)} 下垣 弘行^{**)}
^{*)} 電子技術総合研究所 ^{**)} 共同印刷(株)

1. はじめに

画像データベースで個々の利用者の主観に合致したマルチメディア検索を実現するためには、(a) 手書き例示画でのパターンの歪みや、(b) 視覚心理に基づく主観的類似性を判断することが、重要な課題である。

我々は、商標・意匠図形を対象に、類似画検索等のマルチメディア対話が可能な画像データベースシステム TRADEMARK を開発している [1]。本稿では、図形の濃淡信号、空間周波数、コントラスト等の空間的構造特徴に注目した図形視のモデルを提案し、このモデル上で類似度学習の有効性を示す。判別分析による学習で、類似画検索に98%以上の適合率を実現した。

2. 図形視のモデルと空間的構造特徴

マルチメディア対話を実現するためには、視覚系の網膜のレベルから主観的な判断のレベルまでを扱う必要がある。ここでは、商標図形等の画像データベースを対象とした図形視のモデルを考える。

図形視の規則性を調べるため、各被験者(11名)に、商標図形の集合(227種)を形状的な類似性に着目して自由に分類させた。その結果、平均約50グループへの分類の内、図形の全体的な印象(信号の特徴 58.9%)と共に、部分図形の形状や隣接関係(構造的な特徴 15.9%, 15.3%)に注目して判断したと推定された [1]。

これより、図形視のモデルとして、図形の全体的な印象は濃淡信号や空間周波数で、また、図形の階層性や2次元性等の構造は、局所的な信号の相関等で表現するものとした(図1)。この仮定は、視覚系の自然なモデル化である [2]。採用した特徴量を次に示す。

(a) 濃度分布 Gray8: 対象図形を8×8個のメッシュに分割して、各メッシュ内の黒画素数 m_i で濃淡信号を表わす(64次元ベクトル)。

(b) 周波数分布 RunW: 対象図形を縦・横に短冊型のメッシュに4分割して、メッシュ内のラン長のヒストグラムで空間周波数を近似表現する(128次元)。

(c) 局所相関 Corr4: 対象図形を4×4個のメッシュに分割して、隣接するメッシュ間での黒画素数の相関 $Corr(m_i, m_j) = m_i \cdot m_j$ (1)

で隣接関係を表現する(24次元)。

(d) 局所コントラスト Cont4: 対象図形を4×4個のメッシュに分割し、隣接メッシュ間の黒画素数に関し $Cont(m_i, m_j) = (m_i - m_j) / (m_i + m_j)$ (2)

で隣接関係を表現する(24次元)。

3. 空間的構造特徴による内容検索

3.1 例示画検索と再現率

手書き等の例示画検索に関して、再現率を評価した。各特徴量が表現する図形特徴空間GF上のユークリッド距離に基づき、候補図形を検索する。1600件のデータ

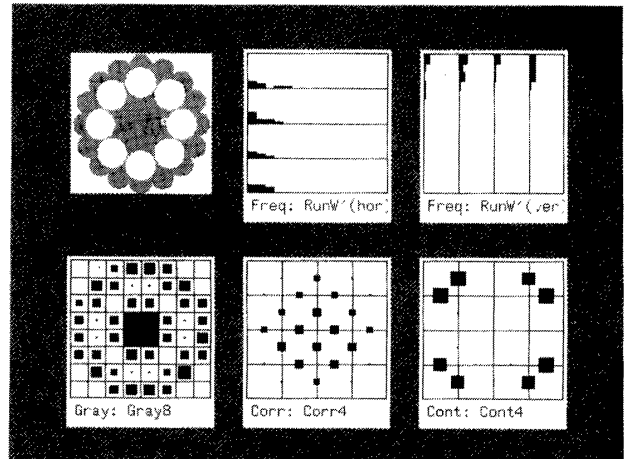


図1 図形視と図形の空間的構造特徴

ベースに対し、清書・手書き・ラフスケッチ各100種を例示した場合の再現率を表1に示す。いずれの特徴量も、清書に対して高い再現率(90%程度以上)を示した。一方、歪みの大きいラフスケッチでは低下した。

粗い信号特徴を表現する Gray8, RunB/W, RunWは、単独で用いても比較的高い再現率を示した。これに対して、構造的特徴を表わす Corr4, Cont4は、信号特徴の場合に比べて低い。複数の特徴量を併用した場合、比較的大きい手書きに対しても、ほぼ実用的な再現率(Gray8, RunB/W: 94%)を実現した。

表1 例示画検索での再現率

特徴量	次元数	清書(%)		手書き(%)		ラフ(%)	
		1~10	10~100	1~10	10~100	1~10	10~100
Gray8	64	100	100	80	95	72	86
Edge8	64	100	100	68	85	57	79
RunB/W	256	100	100	74	89	58	80
RunW	128	96	100	61	85	47	70
Corr4	24	83	91	33	69	27	47
Cont4	24	92	97	36	68	31	54
G8, RB/W	320	100	100	92	99	79	94
G8, E8, RW	256	100	100	91	98	80	95

3.2 類似度学習と適合率

知的な類似画検索を実現するためには、個々の利用者の主観的な類似度を学習する必要がある。本稿では、利用者から得られた商標図形のグループ分けの情報をもとに、判別分析により主観的類似度を反映する空間(主観特徴空間SF)を構成した [3]。SF空間上の表現 y は、GF空間上の表現 x の線形変換で与えられる。

$$y = {}^tAx \quad (3)$$

線形写像Aは、以下の固有値問題から求められる。

$$\Sigma_B A = \Sigma_W A \Lambda, \quad {}^tA \Lambda A = I \quad (4)$$

Σ_B : グループ間共分散行列

Σ_W : グループ内共分散行列

表2 類似画検索での適合率

特徴量	被験者ET ¹⁾	ET ²⁾	MK ²⁾	平均 ²⁾
Gray8	61.1%(30)	70.9%	86.6%(12)	76.9%
RunW	95.5 (37)	100	96.4 (14)	93.1
Corr4	47.9 (16)	42.9	70.1 (8)	51.0
Cont4	46.7 (14)	41.8	62.9 (10)	50.7
G8, RW, Cr4	99.9 (37)	100	99.1 (13)	99.5
G8, RW, Ct4	99.7 (34)	100	98.7 (14)	99.5
グループ	89	89	17	49.5

- 1) 同じグループ内の全要素を検索する確率、(次元数)
- 2) 同じグループ内の1要素を検索する確率、(次元数)

類似画検索は、SF空間上でのユークリッド距離に基づいて、候補図形を選択する。227種の学習サンプルに対する類似画検索での適合率を表2に示す。

粗い信号特徴を表現するRunWは、単独で用いても比較的高い適合率(11名の平均 93.1%)で、Gray8 もそれに次ぐ適合率(76.9%)を示した。これに対し、構造的特徴を表わすCorr4, Cont4は、相対的に低い。これは信号レベルの情報がかかり失われるためであろう。

複数の特徴量を併用すると(Gray8, RunW, Cont4等)、非常に高い適合率(99.5%)を達成した。

Gray8 を用いて、類似画検索した結果を図2に示す。図2で第1, 2, 3, 4, 8, 10候補が、被験者によって同じグループに分類されていたものである。

4. 図形視モデルとの対比

本稿で提案した図形視モデルを、判別学習の結果に基づき評価する。図3(a)に学習サンプル(227種)の平均的性質を示す。Gray8 の分布にみるように、多くは円状で、矩形状の図形等も一部含まれている。

GF空間からSF空間への線形変換Aの各要素は、分類過程で各特徴量をどの程度参照するかを表わす尺度となる。ある被験者について、分類の第1軸, 第2軸に相当するAの要素を図3(b), (c)に示す。この例では、(Gray8, RunW, Cont4)の3種の特徴量を用いて判別学習した。(b), (c)で値の大きな特徴量(の要素)を主に

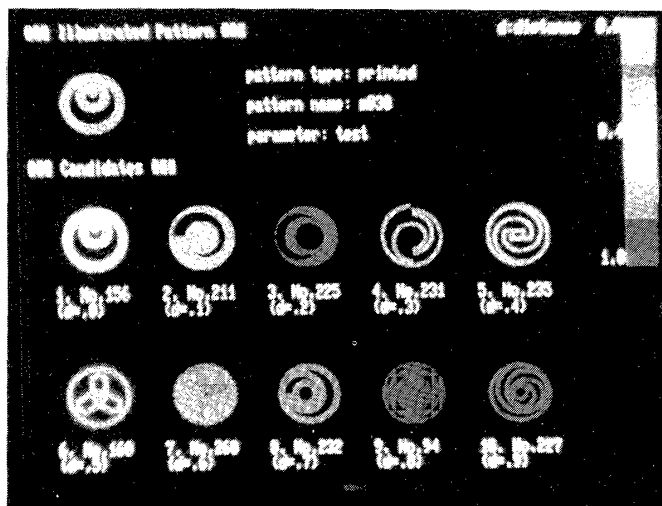


図2 類似画検索の例

参照したといえる。

Gray8 への参照は、主に図形の周辺部に集中しており、円/矩形の分類に使用したと考えられる。RunWへの参照は、分類の第1段階では低い周波数を、第2段階以降では順次高い周波数へと移行しており、全体的な図形の粗さから細かい文様へと注意が向けられたといえる。Cont4への参照は、図形の周辺部から順次中心部に向けられ、Gray8 やRunWを補っている。同様の現象は(Gray8, RunW, Corr4)の組合せでも観察できた。また、細かいグループ分けを行なった他の被験者にも、同様の傾向が見られた。これらのことは分類過程での図形視モデルとして、妥当な結果となっている。

5. まとめ

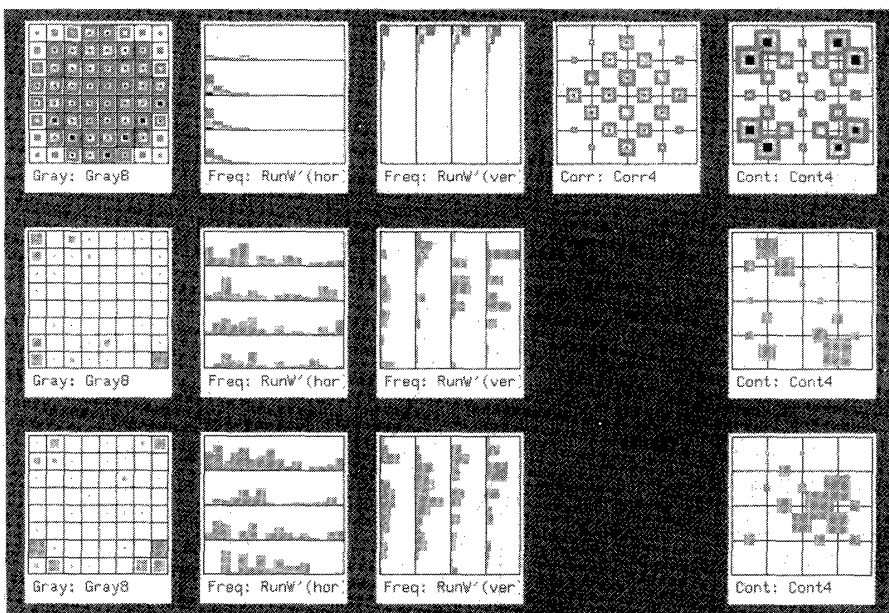
濃淡信号、空間周波数、コントラスト等に着目した図形視モデルと空間的構造特徴を提案し、これがマルチメディア検索に有効なことを実験により示した。

謝辞 コメントを戴いた森晃徳主任研究官に感謝する。

文献 [1] 加藤, 下垣, 藤村: 信学論 Vol. 72D11, No. 4.

[2] リンゼイ, ノーマン: 情報処理心理学入門 I

[3] 栗田, 下垣, 加藤: 情処研資 CV-57-1.



- (a) 学習サンプルの性質
- 赤: $\mu - \sigma$
- 黄: μ
- 緑: $\mu + \sigma$

(b) 線形変換Aの要素(第1軸)

(c) 線形変換Aの要素(第2軸)

図3 学習サンプルの性質と分類の視点