

マルチ識別器を用いた花画像検索システムの構築

Construction of the Flower Image Search System Using Multi Classifier

福田恵太
Keita Fukuda

滝口哲也
Tetsuya Takiguchi

有木康雄
Yasuo Ariki

神戸大学工学部 情報知能工学科

Department of Computer and Systems Engineering, Faculty of Technology, Kobe University

1 はじめに

道端，庭，野原に咲く花の情報などを知る手段として植物図鑑やWWW上の検索システムなどがある．しかし，映像，写真から知りたい花の情報を得ることは非常に困難であり，時間もかかる．また従来の手法では花の構造を基に特定の特徴を考慮した研究は少ない．そこで本研究では，花の構造に依存した各特徴量を重視する識別器を複数構築し，Fuzzy C-means法により各識別器への帰属度を求めた重み付きのマルチ識別器を提案する．

2 特徴量抽出

k-means法により減色した花の正面画像に対し，中央でなおかつ画素数の大きいものを花領域として抽出する[1]．得られた花領域に対しHSV色空間において 10×10 の区間に分割したHS領域の二次元ヒストグラムを色の特徴量とする．花領域の重心から輪郭までの距離を一次元グラフとし，さらに周波数領域に変換した低周波成分30次元を形の特徴量とする．

3 マルチ識別器の構成

3.1 Fuzzy C-means

花の構造的な違いに応じて異なる識別機を設計するために教師なしのクラスタリングを行った．この結果，表1に示す4つのクラスタが得られた．これらのクラスタは，円形度と周波数域のエントロピーを主たる特徴量として持っている．検索画像とデータベース内の画像の各識別器への帰属度 $U_i(u_0, u_1, u_2, u_3)$ を基に検索を行うため，Fuzzy C-meansで帰属度を求める．なお， $i = 0$ を検索画像とする．

表1 各識別器の特徴

各識別器	対象とする花の特徴	エントロピー	円形度	重視する周波数領域
F0	円形に近い花	低-高	高	低域
F1	複数花弁を持つ花	高	低	高域
F2	一枚に特徴を持つ花	低	低	ピーク値を含む帯域
F3	以上に属さない花	中	中	30次元全域

3.2 識別の流れ

U_0 から閾値処理により有効な識別器を採用する．次に，採用された識別器への帰属度 U_i が高いデータベース内の画像 i を対象に検索画像との類似度を計算する．画像 i の色の特徴量を C_{ij} ，形の特徴量を S_{ij} とすると k 番目の識別器から出力される検索画像と画像 i の類似度 F_{ki} はヒストグラムインタセクションを用いた以下の式より求められる．

$$F_{ki} = \frac{w_{k1} \times \sum_j \min(C_{0j}, C_{ij}) + w_{k2} \times \sum_j \min(S_{0j}, S_{ij})}{w_{k1} + w_{k2}}$$

w_{k1} と w_{k2} は実験的に決めた色と形の重みである．各識別器の類似度 $F_i(F_{0i}, F_{1i}, F_{2i}, F_{3i})$ から検索画像の帰属度を掛け，線形和を求めることで最終的な類似度を以下の式から求める．

$$E_i = U_0^T F_i$$

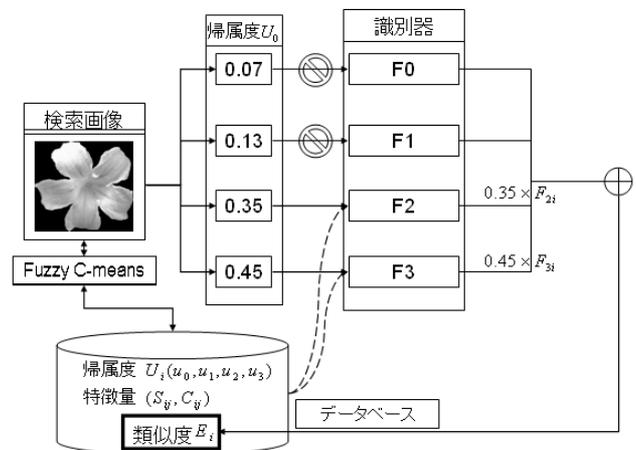


図1 識別器の流れ

4 評価実験

WWW上の花画像92種類をそれぞれ4枚用意し，3枚をデータベース(92 x 3枚)に，1枚をテストデータとし，検索した画像が1位~3位，1位~5位，1位~10位に入る検索率をそれぞれ求めた．評価実験を行った結果を表2にまとめる．特徴を考慮しないF3のみの識別器に比べ，特徴を考慮し統合した識別器では，上記で述べた3タイプの特徴のある花に対して特に高い精度が得られた．

表2 実験結果

	1 - 3位	1 - 5位	1 - 10位
F3のみ	52.1 %	64.1 %	79.3 %
F0~F3	61.5 %	74.6 %	83.3 %

5 まとめ

本稿では，花の構造により重視すべき特徴を考慮したマルチ識別器を用いた花画像検索システムを提案した．今後は集合花など，さらに多くの構造に対応できる識別器を構成する必要がある

参考文献

- [1] 金山，川嶋，青木ら，“植物データを対象とした画像インデクシング”信学技報，PRMU97-172