

花画像データベースのための花領域抽出手法

岡村 知明[†] 精廬 幹人[†] 岩崎 雅二郎[‡] 中村 真吾^{††} 青木 義満^{‡‡} 橋本 周司[†]

早稲田大学理工学部応用物理学科[†] 株式会社リコーソフトウェア研究所[‡]

(株)ベルテック・ジャパン^{††} 芝浦工業大学工学部情報工学科^{‡‡}

1. はじめに

自然環境下で撮影された花画像を入力として花画像データベースから類似画像検索を検索する際、対象画像から花の領域だけを抽出する必要がある。しかしこれらの画像には通常、葉や土などの背景部分が含まれている。ここでは複雑な背景の花画像から花領域だけを抽出する新たな手法を検討した結果を報告する。

2. 花画像の選出

花画像の収集のために東京都港区の国立科学博物館付属自然教育園内に咲く花をデジタルスチルカメラ(DSC-F505K, SONY)を用いて撮影を行った。画像はRGB各8bitで解像度は320×240ピクセルである。本手法による花領域の抽出精度を測定するにあたり、この画像の中から次の4つの条件を満たす花画像100枚を選出した。

- 花が画像のほぼ中央に位置する。
- 画像全体に対し、花が十分な領域を占める。
- 画像中に目的とする花以外の花が含まれない。
- 明るい状況下において撮影された花である。

3. 花領域抽出手法

3.1 従来の花領域抽出手法[1][2]

まず画像をRGB値からHLS値に変換する。この値を用いて花の色と思われる色を検出することによって花領域を抽出する。HLS空間において9つ(赤、黄、茶、緑、水色、青、紫、白、暗)にクラスタリングを行い、この9つの中から、

- 葉の色等背景に多いと思われる茶、緑、暗以外である。

- 全体に対してある程度の画素数を持っている。
- 画素の画像中心からの平均距離が最も近い。という条件を満たした上位2つの色領域に属するピクセルを花領域として抽出する。

この手法は高速に実現できるが色の閾値処理のみであるため、背景を拾ってしまうなど問題点が多かった。

3.2 提案手法

画像をRGB値からHLS値と比べて人の色感覚により近いLab値に変換し、この値を用いてラベリング処理を行う。このとき、領域結合の判断基準として注目領域の平均色と隣接ピクセルの色のLab色空間における色差を用いる。2つの色 $(L_\alpha, a_\alpha, b_\alpha)$ 、 $(L_\beta, a_\beta, b_\beta)$ の色差は、

$$\Delta = \{(L_\alpha - L_\beta)^2 + (a_\alpha - a_\beta)^2 + (b_\alpha - b_\beta)^2\}^{1/2}$$

で表される。この色差が閾値Tより小さければピクセルは隣接するラベル領域に結合され、大きければ新たなラベル領域となる。このようにして画像の(0,0)座標から(320,240)座標まで処理を行い、全ピクセルに対してラベルを振り分ける。また、Lab色空間においてLを10等分、a、bを12等分して花の色と背景の色の分類を行った花領域判定テーブルを用意する。ラベルの平均色とこのテーブルを比較することにより各ラベルが花領域か背景かを分類し、花領域を抽出する。図1は画像から花領域を抽出した結果の例である。

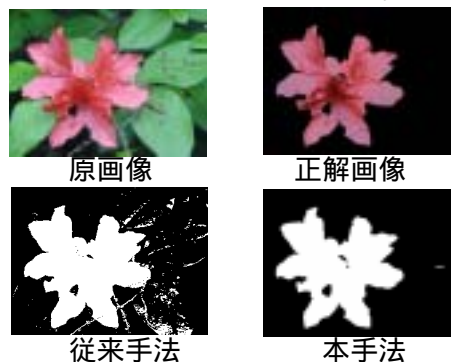


図1 花領域抽出画像

The method of flower area extraction for flower image database
[†]Tomoaki Okamura, Mikito Toguro, and Shuji Hashimoto, Department of Applied Physics, Waseda University

[‡]Masajiro Iwasaki, RICOH COMPANY, LTD.

^{††}Shingo Nakamura, BERTEC JAPAN CO., LTD.

^{‡‡}Yoshimitu Aoki, Department of Information Science and Engineering, Shibaura Institute Technology

4. 領域抽出実験

前述の撮影条件を満たす花画像100枚を用いて花領域の抽出実験を行った。100枚の花画像について、あらかじめ手作業で花領域を抽出した正解画像を用意し、提案する花領域抽出方法によって作成した抽出画像と比較することにより抽出精度を測定した。提案手法では色差の閾値を10から50まで5ずつ変えて実験を行った。抽出精度の評価には次式の再現率と適合率を使用した。

$$\text{再現率} = \frac{\text{正解したピクセル数}}{\text{正解画像の花領域のピクセル数}} \quad (1)$$

$$\text{適合率} = \frac{\text{正解したピクセル数}}{\text{抽出画像の花領域のピクセル数}} \quad (2)$$

表1 閾値による抽出精度の変化の平均

閾値	再現率(%)	適合率(%)	再現率の分散	適合率の分散
10	83.92	78.87	0.0397	0.0542
15	85.26	80.31	0.0455	0.0515
20	85.57	81.18	0.0516	0.0528
25	87.86	81.22	0.0552	0.0623
30	87.66	82.46	0.0645	0.0683
35	88.20	84.77	0.0679	0.0617
40	86.05	85.27	0.0845	0.0624
45	80.27	83.41	0.1220	0.0802
50	70.79	78.34	0.1686	0.1154
従来手法	75.17	69.24	0.0903	0.1116

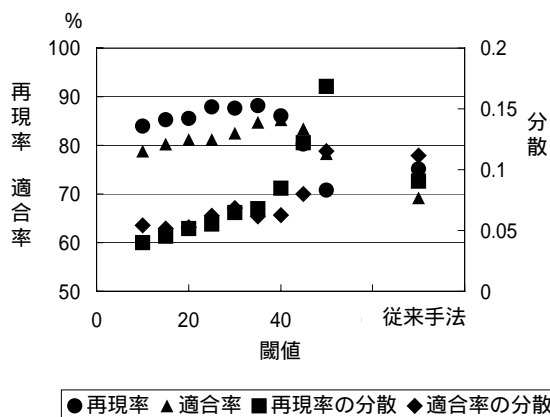


図2 閾値による抽出精度の変化の平均

図2は100枚の画像における再現率と適合率の平均値を示したものである。これより閾値50を除いて従来手法より抽出精度が高く、さらに閾値35の場合では再現率88.20%、適合率84.77%と共に高い抽出精度であることが分かる。

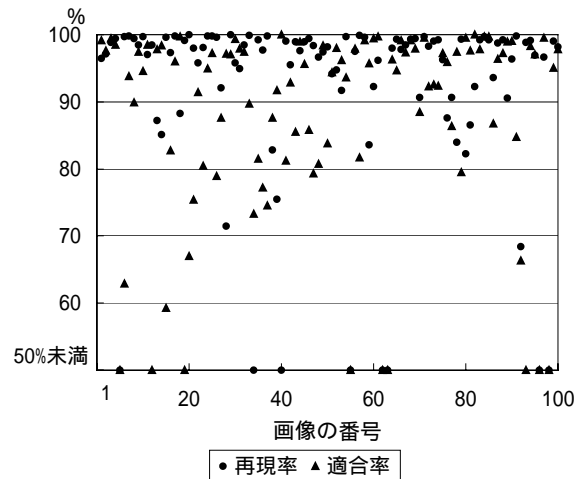


図3 閾値35の各画像における抽出精度

図3は閾値35における各画像の抽出精度を表す。ほとんどの画像で高い抽出精度となっているのがわかる。しかし、花と背景の色差が少ない画像に関しては、花領域が背景と定義され抽出精度が大幅に低下する。このため図2で閾値が高くなると分散が大きくなるのである。したがって、抽出したピクセル数が極端に低い画像に関しては、閾値を下げても抽出し直すことによって抽出精度は改善されると考えられる。

5. まとめ

自然環境下で撮影された花画像から花の領域のみを抽出する手法を提案し、実際に抽出実験を行い多くの画像で高い抽出精度が得られることを示した。今後の抽出精度の向上の方法としては例えば各ラベルの位置情報にも注目し、画像中心から離れたラベルを背景として除去したり、ピクセル数の極端に少ないラベルを除去すること等が考えられる。現在、抽出した画像によって類似画像検索実験を行うことにより、花領域抽出が検索精度の向上にどの程度貢献しているのかを検証中である。

6. 参考文献

- [1] 岡村知明、精廬幹人、中村真吾、青木義満、橋本周司、“画像特徴とインデックスデータによる検索機能を備えた花画像データベースの構築、”画像電子学会第196回研究会、講演予稿、pp.29-32, Sep.2002.
- [2] S. Nakamura, M. Sawada, Y. Aoki, P. Hartono, S. Hashimoto, “Flower Image Database Construction and its Retrieval,” 7th Korea - Japan Joint Workshop on Computer Vision, FCV2001, pp37-42, 2001