

モバイル環境における画像検索システムの開発

3E-5

野田 真樹子† 園部 博崇‡ 高木 佐恵子‡ 吉本 富士市‡

†和歌山大学大学院システム工学研究科 ‡和歌山大学システム工学部

1. はじめに

今日、大量の画像データを管理し、検索するために画像データベースを利用した検索システムが使用されている。現在、画像の検索方法は画像の内容を表すキーワードを検索キーとして検索する方法と、画像そのものを検索キーとして検索する方法の2つに大別される[1]。画像に写っているものが何かということを検索する場合、画像そのものを検索キーとして画像から特徴を抽出する方法が有効である。斎藤ら[2]は、花と葉の特徴を用いて野草の自動認識を行っている。しかし、このシステムはモバイル環境での利用には対応していない。

本稿では、モバイル環境で利用できる、花の画像検索システムを提案する。提案システムでは、花が咲いている場所から移動しなくても、利用者はモバイル機器を用いて、花の画像やその特徴から花の名前を検索することができる。

2. モバイル画像検索システム

2.1 モバイル画像検索システムの概要

本稿で提案するシステムの全体図を、図1に示す。

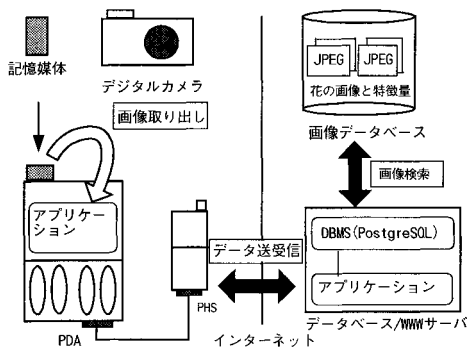


図1 モバイル画像検索システム

利用者がデジタルカメラで撮影した花の画像と、その花の特徴情報をPDAとPHSを用いて検索サーバに送信すると、サーバは受信した画像から特徴を抽出して検索を行う。検索結果はWebページに表示され、利用者はそのページをPDAのブラウザ

で見て結果を確認する。

PDA上に作成したアプリケーションは、記憶媒体内の画像ファイルをリストにして表示する機能、花の特徴を指定させる機能、画像ファイルと特徴情報をサーバに送信する機能を持っている。利用者は検索したい画像をリストから選択し、花の特徴を指定する。利用者が指定できる花の特徴は、花びらの色、花の種類、花が咲く季節である。これらの特徴情報と、花の画像をPHSを介してサーバに送信する。

サーバでは受信画像から形状特徴と色の特徴を抽出する。そしてデータベースに登録されている画像の特徴量と受信画像の特徴量とを比較し、類似しているものから順にデータベース内の画像を結果画像として取り出す。その後、結果画像を表示するためのHTMLファイルを作成し、そのWebページのURLを利用者に電子メールで送信する。利用者は電子メールの受信後、PDAのブラウザでそのページを見て結果を確認する。

2.2 検索キーとなる画像

検索に利用する特徴を抽出しやすくするために、検索対象画像に以下の3つの条件を設定した。

- (1) 画像の大きさが320×240ピクセルである。
- (2) 画像中央の160×160ピクセルの範囲内に花が一輪だけ写っている。
- (3) 花は160×160ピクセルの範囲に入る大きさである。

ただし、(2)の花の位置は、後に述べる特徴抽出の際に位置調整できるので多少ずれていてもよい。

2.3 花の抽出

花の特徴を抽出する際に、画像から花だけを取り出す必要があり、画像の2値化を行った。このとき、利用者によって指定された花びらの色を使用する。検索対象となる画像は2.2節で述べた条件を満たしているため、画像中央の160×160ピクセルの範囲のみを2値化する。まず、その範囲内にある全画素のCIE-LAB値を求め、利用者が指定した色との色差を算出する。その色差のヒストグラムをとり、モード法によってしきい値を決定し、色差の値がしきい値以下のものを白、それ以外を黒にする。この2値化処理により白となった領域を花とみなす。

2.4 形状特徴・色の特徴の抽出

検索に利用する花の形状特徴には、複雑度[3]を用いている。複雑度はある形状がどれだけ円に近いを示す値で、以下の式で求められる。

$$(\text{複雑度}) = (\text{周囲長の2乗}) / (\text{面積})$$

ここで、花の面積は2値画像中の白画素数である。また、周囲長は境界線追跡アルゴリズム[3]を用いて求める。

色の特徴には HSV 値の H(色相)の値を用いている。花には、中心の色と花びらの色が異なっているものもあるので、図2に示すように花びらの4点と花の中心の1点、合計5点の色相の値を特徴として抽出する。

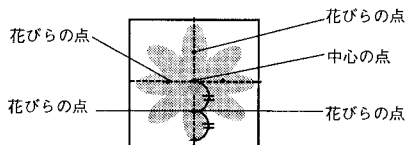


図2 色の特徴を抽出する点の位置

2.5 特徴の比較

利用者が指定する花の特徴は、花の種類と花が咲く季節である。種類としては木か草かを選択し、季節は撮影した時点の月を指定する。種類を木と草に分けたのは、植物の分類について詳しい知識がなくても指定できるようにするためである。撮影した時点から時間が経過した後に行う検索に対応するために、利用者は季節の情報を無効にすることもできる。検索の際には、これらの特徴情報で候補となる画像の絞り込みを行う。

次に、形状特徴、色の特徴を表1のように定義する。

表1 形状特徴と色の特徴の定義

	検索キーとなる画像	データベース内の画像
複雑度	E_c	E_s
中心の色	O_c	O_s
花びらの色	$P_c(i) (i=1,2,3,4)$	$P_s(i) (i=1,2,3,4)$

これらを用いて検索キーとなる画像と、データベース内の画像がどれだけ類似しているかを評価する。そのために照合度 M と呼ぶ値を用いる。 M は以下の式より求める。

$$M = |E_c - E_s| \times W_e + |O_c - O_s| \times 100 \times W_o + \sum_{i=1}^4 |P_c(i) - P_s(i)| \times 100 \times W_p$$

ここで、 W_e , W_o , W_p はそれぞれの特徴につける重みである。照合度 M の値が小さいほど検索キーとなる画像と検索結果画像は類似しているといえる。

3. モバイル画像検索システムの評価実験

学内に咲いている花12種類をデジタルカメラで撮影し、データベースに登録した。この段階で、データベースには合計135種類の花のデータが登録されている。その後、データベースに登録した花と同じ種類であるが、異なる花を撮影し、その場で検索実験を行った。このとき、利用者が指定する花の特徴が検索に有効かどうかの確認を行った。実験に

は、PDA (CPU:DragonballIVZ, Clock:33MHz), サーバ側コンピュータ (CPU:Pentium II, Clock:266MH) を使用し、開発言語には CodeWarrior for PalmOS platform と Java を用いた。また、PostgreSQL を用いて画像データベースを作成している。

季節などの情報を使わずに検索実験を行った結果、検索結果が第1位から第3位に入ったものが約58%であった。同じ画像を、季節などの情報を使って絞り込みを行って検索した結果、検索結果が第1位から第3位にはいったものは約92%であった。一連の実験において、サーバ側での特徴抽出と画像検索にかかった時間は、約2~5秒であった。また、利用者が検索結果第一位のWebページを見るまでは約2~3分かかった。

4. 考察

実験の結果、利用者が指定した季節や花の種類の情報が検索結果に大きな影響を与えていることがわかった。検索結果をより良いものにするために、利用者が花の特徴を指定するという操作は重要であると考えられる。

また、検索にかかった時間については、特徴抽出、画像検索にかかる時間はリアルタイムに結果を返すために十分な時間だといえる。しかし、Webページの確認までに約2~3分と時間がかかっている。これは、電波の状態や実験を行う時間帯にも左右されるが、PDA側での、メール受信及びWebページのデータ受信がスムーズに行われていないことによる。これに関しては、結果表示機能も独自に開発することで解決可能と考えられる。

5. まとめ

本稿では、PDA, PHS を用いて花が咲いている場所で花の名前を検索し、リアルタイムに結果を得ることができるシステムを提案した。検索の際には、利用者から与えられた花の特徴情報と、デジタルカメラで撮影した花の画像から形状特徴と色の特徴を抽出して利用した。本システムで実験を行った結果、画像から抽出した特徴と、利用者が指定した特徴を合わせて利用すれば目標とする結果が第1位から第3位に入ったものが約92%と良い結果が得られた。今後の課題はデータベースに登録する花のデータの増加と、検索の際に使用する花の特徴を増加である。また、PDA側では、ユーザインタフェースの改良、結果画像の表示方法の変更などが今後の課題である。

参考文献

- [1]小早川 倫広, 星 守: 画像内容に基づいた画像検索システム, bit, Vol. 31, No. 10, pp. 23-34 (1999).
- [2]斎藤 剛史, 金子 豊久: 花と葉による野草の自動認識, 電子情報通信学会論文誌, D-II, Vol. J84-D-II, No. 7, pp1419-1429, (2001).
- [3]田村 秀行: コンピュータ画像処理入門, 経研出版 (1985).