

Cosmos: モバイル機器で利用できる花の画像検索システム

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

現在、多くの画像検索システムがあるが、それらをモバイル環境で利用することは難しい。そこで、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。提案システムは、デジタルカメラで撮影した花の画像と、利用者が指定した簡単な特徴情報を、モバイル機器を用いて画像検索サーバに送信し、サーバで検索を行った結果をモバイル機器で確認するというシステムである。サーバでの検索の際には、利用者が指定した特徴と、デジタルカメラで撮影した花の画像から抽出した形状や、色の特徴を使用する。本システムで検索実験を行った結果、目的とする画像が検索結果の第1位から第3位に入ったものが約92%であり、その結果は満足できる時間内に取得できた。

Cosmos: Convenient Image Retrieval System of Flowers for Mobile Computing Situations

Makiko Noda[†] Hirotaka Sonobe^{††} Saeko Takagi^{††} Fujiichi Yoshimoto^{††}

[†] Graduate School of Systems Engineering, Wakayama University

^{††} Department of Computer and Communication Sciences, Wakayama University

There are many image retrieval systems today. These systems, however, are difficult to use outdoors. In this paper, we propose an image retrieval system of flowers that can be used for mobile computing environment. In this system, we photograph a flower with a digital camera and specify some simple characteristics of the flower. Then we transmit the flower image and the characteristics to an image retrieval server with PDA and PHS. In the image retrieval server, images similar to the received image are retrieved. For the purpose of retrieving the images, we use the characteristics specified and characters of shape and color extracted from the images. The result of our experiment shows that the percentage of getting the objective image from the first to the third places was about 92%. The objective flower-data was obtained in satisfactory time.

1 はじめに

野山や道端、また、人家の庭先などに咲く花には様々な種類があり、その名前を知りたいと思うときはよくある。そのような場合、既存の植物図鑑で調べる方法や、Web上の検索システムで調べる方法などがある。しかし、このような方法は、屋外では利用しにくく、時間がかかる。

その解決方法として、モバイル機器の利用が考えられる。モバイル環境で利用できる検索システムならば、屋外で花が咲いている場所で直ちに使用できる。現在、デジタルカメラや、携帯電話、PDAなどのモバイル機器は一

般に広く普及してきている。また、モバイル機器の発展は目覚しく、大容量のデータを高速に伝送できるようになってきている[1]。それにともない、モバイル機器を利用したシステムも多く提案されている[2, 3, 4]。しかし、モバイル機器を用いて、花が咲いている場所で直ちに花の特徴から花の名前を調べるというシステムは開発されていない。斎藤ら[5]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、花の名前や花に関する情報を、花が咲いている場所で調べるということには行っていない。

本稿では、モバイル環境で利用できる、花

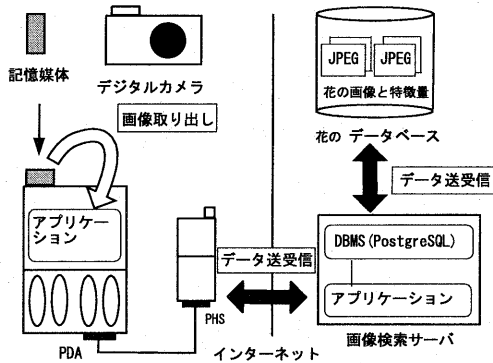


図 1: モバイル画像検索システムの全体図

の画像検索システムを提案する。提案システムでは、利用者は花が咲いている場所から移動しなくても、モバイル機器を用いて直ちに、花の画像やその簡単な特徴から花の名前やその花に関する情報を検索することができる。

2 提案システムの概要

2.1 モバイル画像検索システム

本稿で提案するシステムをモバイル画像検索システム Cosmos と呼び、その全体図を図 1 に示す。PDA は計算環境が貧弱であるため、PDA 上で可能な処理には限度がある。そこで、本システムでは画像からの特徴抽出やその特徴比較など計算負荷の多い処理は画像検索サーバで行うようにした。

本システムでは、まず、利用者は名前の分からない花をデジタルカメラで撮影し、花の画像を作成する。その画像を PDA に取り込み、PDA 上のアプリケーションで花の簡単な特徴を指定する。そして、花の画像とその特徴情報を、PHS を介して検索サーバに送信する。サーバでは受信した画像から形や色の特徴を抽出して検索を行う。検索結果は Web ページに表示される。その後、サーバから利用者に、検索結果が表示されている Web ページの URL が電子メールで送信されるので、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。

2.2 サーバと PDA 間のデータの送受信

画像検索サーバと PDA 間のデータの送受信の様子を図 2 に示す。図 2 において黒色の矢印は、花の画像や花の特徴情報の移動を表し、

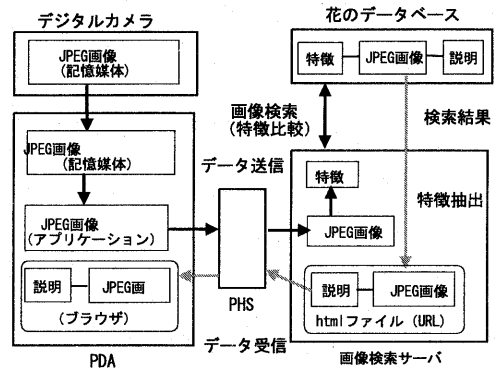


図 2: 提案システムにおけるデータの送受信

灰色の矢印は検索結果の移動を表す。

PDA 上に作成したアプリケーションは、次の 3 つの機能をもつ。その 1 つ目の機能は、記憶媒体内の画像ファイルをリストにして表示し、利用者に画像を選択させる機能である。記憶媒体内に保存されている画像は、PDA 上のアプリケーションによって撮影時間の新しい順にリストアップされる。利用者は、その中の画像をスタイラスペンでタップすることで検索したい画像を選択したことになる。

2 つ目の機能は、花の簡単な特徴を指定させる機能である。利用者が指定する項目は、花びらの色、花のタイプ、花の咲く時期である。花びらの色はあらかじめ用意してある 9 種類の色の中から近い色を選択し、併わせてその濃淡も選択する。続いて、花のタイプとしてその花が木に咲く花か、または草の花か、という簡単な選択を行う。花のタイプをこれら 2 つに分けたのは、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。最後に、花の咲く時期として、その花が現在咲いているかどうかを指定する。これらの特徴はいずれも、利用者が簡単に得ることができるものである。しかし、画像検索サーバ側の処理でこれらの特徴を得ることは難しいため、使用者に指定してもらうことにした。

3 つ目の機能は、検索したい花の情報をサーバ側に送信する機能である。選択した花の画像ファイルと利用者が指定した花の特徴情報を、あらかじめ登録してあるアクセスポイントへ PHS を介して接続し、画像検索サーバに送信する。以上の 3 つの機能によって利用者

は検索したい花の情報をサーバに送信できる。

画像検索サーバでは、受信した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、結果画像とその花の情報を表示するための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。

3 花の検索方法

3.1 検索キーとなる画像の条件

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。

- (i) 画像の大きさが 320×240 ピクセルである。
- (ii) 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。
- (iii) 花は 160×160 ピクセルの範囲に入る大きさである。

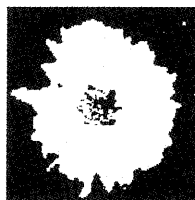
ただし、(ii)の条件の花の位置は、後に述べる特徴抽出の際に位置を調整できるので、花がその範囲から多少ずれていてもよい。

3.2 花の抽出

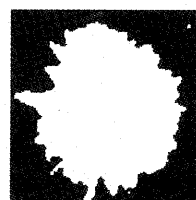
画像から花の特徴を抽出するためには、画像から花だけを取り出す必要がある。そこで、画像を2値化する方法をとった。このとき、利用者によって指定された花びらの色を使用する。検索対象となる画像は、3.1節で述べた条件を満たしているので、画像中央の 160×160 ピクセルの範囲のみを2値化する。まず、その範囲内にある全画素のCIE-LAB値を求め、利用者が指定した色との色差を算出する。その色差のヒストグラムをとると、画素が花びらの色に近い部分と、背景部分に分かれて集まるので、その間の値をしきい値とする。そして、色差がしきい値以下のものを白、それ以上のものを黒にする。この2値化処理により白くなった領域を花とみなす。2値化処理を行った画像を図3に示す。



(a) 原画像



(b) 2値画像



(c) 中心を白く塗った
2値画像

図3: 画像の2値化

図3(a)の原画像のように、花びらの色と花の中心部分の色が異なると、2値化の際に、図3(b)のように中心部分が黒く残ってしまう。このようになると、後に説明する形状特徴を求めるときに適当でないため白く塗りつぶすようにする。中心部分を白く塗った画像は図3(c)のようになる。

4 特徴抽出と比較

4.1 形状特徴の抽出

花の形状特徴には、花の大きさや花全体の形などがある。しかし、撮影者や撮影条件によって画像中の花の大きさは異なるため、大きさに左右されない特徴を用いる必要がある。そこで、検索に利用する花の形状特徴には、複雑度[6]を用いることにした。複雑度はある形状がどれだけ円に近いかを示す値であり、以下の式で求められる。

$$\text{複雑度} = \frac{\text{周囲長の2乗}}{\text{花の面積}} \quad (1)$$

ここで、花の面積は2値画像中の白画素の数となる。2値画像中に、花のほかにも小さな白画素の領域が生じる場合もある。このような場合は、面積が本来より少し大きな値になる

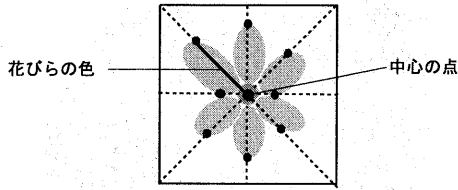


図 4: 特徴とする色を得る点の位置

が、複雑度に影響を与えるほどではないと考えられる。また、周囲長は境界線追跡アルゴリズム [5] を用いて求めた。境界線追跡アルゴリズムでは、まず、2 値画像中の画素をラスタ走査し、まだ追跡されていない境界点を探す。その境界点が見つければ、その点の周りの 8 画素から次の境界点を調べ、境界線を追跡する。そして、追跡されていない境界点がなくなればアルゴリズムを終了するというものである。このようにして 2 値画像中にある、全ての白画素の領域について周囲長を求め、その最大値を花の周囲長としている。

4.2 色の特徴の抽出

提案システムでは、形状特徴の他に色の特徴も使用している。色の特徴には HSV 値の H(色相) の値と S(彩度) の値を用いている。

特徴とする色の求め方は次のようになっている。花の形状からほぼ重心となる点をモーメント [6] を用いて求める。そして、その点と画像から切り出した 160×160 ピクセルの画像の中心の点を合わせる。図 4 に示すように、2 値画像の中心に向かって 8 本の直線を引き、その直線と交わる花びらの境界点 8 点を求める。その 8 点と中心の点との間の長さが一番長いものを求める。この直線上に含まれる画素の色相の値をヒストグラムにとり、色相の値における画素数の分布を調べて、最も多く含まれるものから順に 4 つの色相の値を求める。同様に、彩度の値も 4 つ求め、色相の値 4 つと、彩度の値 4 つを花びらの色の特徴とする。

さらに、画素数の分布から花びらに何色の色が含まれているかを求め、花びらの色の数としている。花びらの色の数も特徴としている。

4.3 特徴の比較

表 1 に示すように検索に利用する形状特徴と、色の特徴を定義する。これらを用いて検

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
花びらの色相	$P_c(i)$ ($i=1,2,3,4$)	$P_s(i)$ ($i=1,2,3,4$)
花びらの彩度	$S_c(i)$ ($i=1,2,3,4$)	$S_s(i)$ ($i=1,2,3,4$)
花びらの色の数	N_c	N_s

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 を以下のように定義する。

$$M = |E_c - E_s| + \sum_{k=1}^4 |P_c(k) - P_s(k)| + \sum_{k=1}^4 |S_c(k) - S_s(k)| + |N_c - N_s| \quad (2)$$

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。

複雑度は、 4π 以上の数で、色相と彩度の値は $0.0 \sim 1.0$ までの値である。したがって、このままでは色相の差の値と彩度の差の値が小さすぎて、照合度に色の特徴があまり反映されない。そこで、色相の差の値と彩度の差の値を 10 倍している。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。

このように、花の形状特徴と、色の特徴を用いて検索を行っているが、それだけでは十分な結果を得ることが難しい場合もある。このため、本システムでは絞り込み検索を行っている。絞り込み検索は、利用者が指定した花のタイプと花の咲く時期の 2 つの特徴を用いて行う。花のタイプと花の咲く時期の特徴は 2.2 節で述べたものである。絞り込みは以下の順で行う。



図 5: 検索結果表示ページの例



図 6: 詳細ページの例

- (i) 花のタイプと花の咲く時期の両方が該当するもの。
 - (ii) 花のタイプだけが該当するもの。
 - (iii) 花の咲く時期だけが該当するもの。
 - (iv) 両方該当しないもの。
- (i) ~ (iv) の順に形状や、色の特徴を比較することによって絞り込み検索を行う。

5 検索結果の表示

利用者が PDA の画面で見たときの検索結果表示ページの例を図 5 に示す。まず、検索結果の第 1 位から第 10 位の花のサムネイル画像と、花の名前が表示される。利用者は、そこから目的とする花を選んでその詳細を見ることができる。

図 6 は詳細ページの例である。利用者が、画像をみただけでは目的とする花の情報が得られたかどうか分からない場合もありうる。そこで、図 6 のように花の画像と名前だけでなく、その花の属する科名、花が咲く環境、分布している場所、高さを説明として加えている。さらに、ワンポイントとして、その花特有の情報を提示している。

6 モバイル画像検索システムの評価実験

実験に用いた花のデータベースは、大学内に咲いている花 12 種類と、CD-ROM [7, 8, 9] などから集めた花 133 種類のデータから構築した。花のデータベースには、この合計 145 種

類の花の画像と、検索に利用している特徴、検索結果を表示する際に必要となる詳細情報を登録した。そして、大学内に咲いている花 12 種類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、絞り込みを行わない検索と、絞り込み検索の両方を行い、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: DoragonballVZ, Clock: 33MHz), サーバ側コンピュータ (CPU: PentiumII, Clock: 266MHz) を使用し、開発言語には、CodeWarrior for ParmOS platform と Java を用いた。また、花のデータベースは、PostgreSQL を用いて作成している。

まず、クチナシを検索した結果を表 2 に示す。絞り込み検索の際には、花のタイプを木、花の咲く時期を 7 月として検索を行った。なお、花のデータベースには、花のタイプが木で、花の咲く時期に 7 月が含まれているものは 10 種類ある。

次に、フェアリーベルを検索した結果を表 3

表 2: クチナシの検索結果

順位	絞り込みなし	絞り込みあり
第 1 位	クチナシ	クチナシ
第 2 位	ウメバチソウ	ステファノテス
第 3 位	スイセン	ニオイバンマツリ

表 3: フェアリーベルの検索結果

順位	絞り込みなし	絞り込みあり
第1位	コウバイ	フェアリーベル
第2位	フェアリーベル	ソバナ
第3位	ソバナ	イワギキョウ

に示す。絞り込みの際には、花のタイプを草、花の咲く時期を7月として検索を行った。なお、花のデータベースには花のタイプが草で、花の咲く時期に7月が含まれているものは75種類ある。

また、検索実験にかかった時間は表4のようになった。

7 考察

実験の結果、利用者が指定した花のタイプや、花の咲く時期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。利用者が指定する花の特徴は簡単なもので、利用者にとってほとんど負担をかけない程度のものであるが、検索結果をより良いものにするために有効であると考えられる。しかし、6節のフェアリーベルの検索のように、絞り込みを行っても候補が75もある場合があるので、形の特徴と色の特徴による検索も重要であるといえる。

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ページで結果を確認するのに、約2分30秒程度かかっている。これは、PDAのブラウザでWebページの画像を表示する際に、画像表示用のプロキシサーバで、画像変換が行われてから画像が表示されているためだと考えられる。この問題に関しては、結果表示機能も独自に開発することで解決可能と考えている。

8 まとめ

本稿では、デジタルカメラ、PDA、PHSを用いて花が咲いている場所で花の名前を検索し、その場で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、花の画像から形状特徴と色の特徴を抽出し、利用した。また、利用者が指定した特徴を用いて絞り込み

表 4: 検索にかかる時間

処理	時間
利用者の操作	約37～70秒
画像検索サーバでの処理	約2～5秒
第1位～第3位の表示	約45～50秒
詳細ページの表示	約1分30秒

検索も行った。本システムで検索実験を行った結果、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ると、目標とする結果が第1位から第3位にはいったものが約92%と良い結果が得られた。

今後の課題として、まず、データベースに登録する花のデータの増加があげられる。また、検索率の向上のために検索に利用する特徴量を追加する必要がある。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] 奥瀬 俊哉: 次世代モバイル・テクノロジーの最新動向, コンピュータ&ネットワークLAN, Vol. 18, No. 10, pp.74-81 (2000).
- [2] 助田 浩子, 佐々木 元, 松尾 仁司, 岡 裕爾: 携帯端末を用いた看護支援システム, 情報処理学会論文誌, Vol. 40, No. 10, pp.3728-3791 (1999).
- [3] 吉野 孝, 宗森 純, 湯ノ口 友友, 泉 裕, 上原 哲太郎, 吉本 富士市: 携帯情報端末を用いた発送一環支援システムの開発と適用, 情報処理学会論文誌, Vol. 41, No. 9, pp. 2382-2393 (2000).
- [4] 藤井 憲作, 杉山 和弘: 携帯端末向け案内図生成システムの開発, 情報処理学会論文誌, Vol. 41, No. 9, pp. 2394-2403 (2000).
- [5] 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, Vol. J84-D-II, No. 7, pp. 1419-1429 (2001).
- [6] 田村 秀行: コンピュータ画像処理入門, 総研出版 (1985).
- [7] 株式会社アスキー: マルチメディア植物図鑑 (1997).
- [8] 株式会社富士通九州システムエンジニアリング: ハイパー植物図鑑 II (1997).
- [9] 株式会社学習研究社: 学研マルチメディア図鑑 野外植物 (1998).