

画像をキーとした撮影ポイント推定アルゴリズム

高橋 琢磨† 須藤 智† 恩田 憲一†
 尚美学園大学 芸術情報学部 情報表現学科†

1 はじめに

携帯電話が普及するにつれて、付加されている機能も高性能化の一途をたどっている。特にカメラ機能に関しては今や 100 万画素も珍しく無く、通信性能も向上していることから、携帯電話をインターフェースとして使用しサーバーと連動した高機能なアプリケーションの作成が可能となってきた。このようなシステム構成において、サーバー群を用いた画像処理を想定した場合は、並列処理可能なアルゴリズムを用いる等して、従来は不可能と思われていた計算コストの高い処理も短時間で実行することが可能であると考えられる。

本研究では、携帯電話使用者が撮影した画像を送信することで、システムが自動的に建物を特定し対象物の情報を返信するというサービスを想定している。これを行うためには位置情報と視点方向の推定が必要となるが、位置情報については、撮影時に自動的にGPS情報を記録することにより取得が可能である。

視点の方向を推定する方法として、携帯電話に内蔵されている電子コンパスを利用するものがあるが[1]、GPSに比べ普及率が低く汎用性が少ないことや、密集した地域において判断を誤ることが少なくない。そこで、本研究においては僅かな誤差でも的確に指定できる様、視点推定における処理の提案を行う。

実際にシステムとして運用する際には前処理として、使用する視点データベースに主要な建物を様々な角度から撮影した画像を予め準備しておき、建物を特定する際は、視点データベース内の画像と、検索原画像を解像度別階層画像生成を行った後の画像である検索対象画像とマッチングを行う。

2 手法

本論文では図1の手順によりマッチングを行っていく。

(1) 解像度別階層画像生成

入力された画像に対して、それぞれX・Y軸ゾーベルフィルタを施しエッジ強度画像を生成、次に2値化を行いエッジ画像を生成する。そしてこれを基に、4分木を用いた階層画像をLnまで生成する。このとき、1つ下の階層画像を見た際、4画素中1つでもエッジノードが存在すれば、対応する階層の親ノードはエッジノードとする。また逆に4画素中1つもない場合は非エッジノードとなる。この作業の流れを解像度別階層画像生成とする。

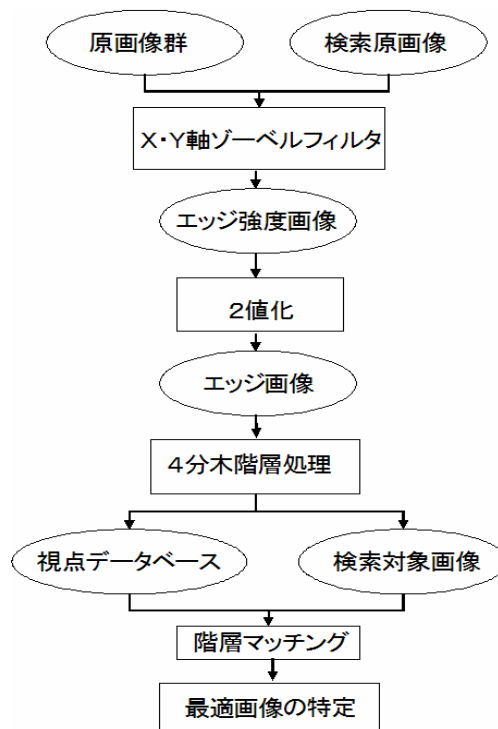


図1 処理手順の流れ

(2) 視点データベース構築

想定される視点位置を網羅的に撮影することにより取得した原画像群に対して、解像度別階層画像生成を行い、視点データベースを作成する。これによりデータ量の少ない上位階層での絞込みが可能となり、探索時に使用するデータ量を減らした高速なマッチングが実現できる。

(3) 検索対象画像

撮影した検索原画像に対して、視点データベースとのマッチングのため、解像度別階層画像生成を行う。これを検索対象画像とする。

(4) 階層マッチング

視点データベースと検索対象画像とのマッチングを行う。探索は階層上位のLn階層から行き最終的にL1階層に至るまで行う。また指標として正誤率を計算することで、各階層が終わるごとに一定値以上の誤差が出た場合、視点データベースの検索対象から外すことが可能となる。これにより明らかに異なる画像を幾度も検索の対象とすることが回避され、絞込みによる、検索コストの削減が期待できる。階層マッチングで最も結果のよかつ

An Image Retrieval Algorithm for Viewpoint Search

†Takahashi Takuma †Sudou Satoshi †Norikazu Onda

†Shobi University
 Faculty of Informatics for Arts
 Department of Digital Expression

た画像を最適画像とする。

今回、正誤率の算出には式①を用いた。なお、この定義での正誤率は低いほど誤差がないということを示している。

$(Z1 + Z2) \times 100 / (\text{white1} + \text{white2}) \dots\dots ①$
 式①にある white1 とは検索対象画像のエッジ数であり、white2 は視点データベース画像のエッジ数である。また white1・white2 を比較したとき、同じ場所にある白エッジの個数を数えたものをマッチ数とし、Z1とZ2は、それぞれ white1・white2 からマッチ数を引いたものである。

3 実験

(1) エッジ検出処理

図2に対して解像度別階層画像生成を行った結果を例として図3に表す。



図2 検索原画像

	縦エッジ画像	横エッジ画像
L0階層		
L1階層		
L2階層		
L3階層		

図3 階層画像

(2) マッチング正誤率結果

先に述べた階層マッチングの結果を示す。図4は縦エッジ画像のL3階層画像における正誤率を表し、図5は横エッジ画像のL3階層画像における正誤率を表している。結果は、図4・5を見てわかる通り、この段階で40%の以下の正誤率なのは8の画像番号1点のみであり、想定していた結果である画像が最適な画像結果であることが判断できた。以下の図6と図7がそれぞれの原画像となる。

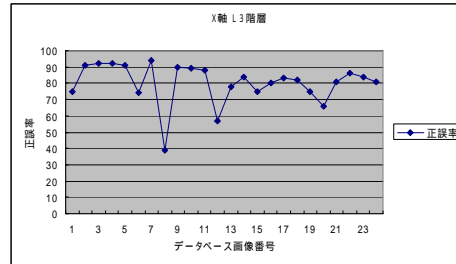


図4 縦軸正誤率表

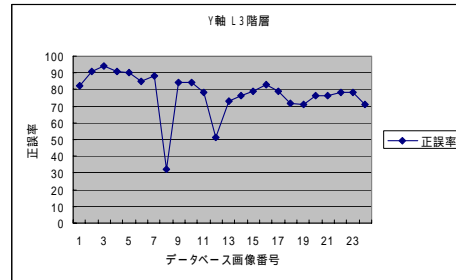


図5 横軸正誤率表



図6 検索原画像

図7 最適結果画像

4 まとめ

本論文ではデジタルカメラでの検証を行い、サイズ640×480画素の画像を使用し、あらかじめ用意しておいた複数の建物での視点データベースを基に階層マッチングの検索検証を行った。

また撮影者の位置などの拘束条件から、ある程度撮影できるポイントを制限できたため、想定していたよりも比較的少ない枚数での視点データベース化も可能となった。

今後はGPS機能と連動させた画像検索の選定や、画像が送信された際の、解像度別階層画像生成と階層マッチングの自動処理、また携帯電話での撮影画像時におけるマッチングの検証などが課題となる。

参考文献

[1]森 俊勇、堀内 成朗、森川 宗ら：”方位・方角情報つき写真送信システム (View Spot System) の開発について”，砂防学会研究発表会, 18年度, O-98