

知っておきたいキーワード

Mobile Visual Search

～携帯端末による画像認識・検索～

(正会員) 島村 潤†

† NTT メディアインテリジェンス研究所

"Mobile Visual Search" by Jun Shimamura (NTT Media Intelligence Laboratories, Yokosuka)

キーワード: Mobile Visual Search, 携帯端末, 画像認識, 画像検索

Mobile Visual Search とは

Mobile Visual Searchとは、スマートフォンなどの携帯端末で撮影された画像から物を認識して、関連する情報を検索・提示するサービス・技術¹⁾のことです(図1)。認識対象は、現在のところ、建築物やワイン、書籍などが代表的です。また検索・提示対象は、関連するWebページや被写体の名称・位置、類似物の画像などです。物の名前を調べたい時や、ガイドンス・

口コミなど関連する情報を調べたいときに、その場で携帯端末を対象にかざ

して検索するようなユースケースが考えられます。

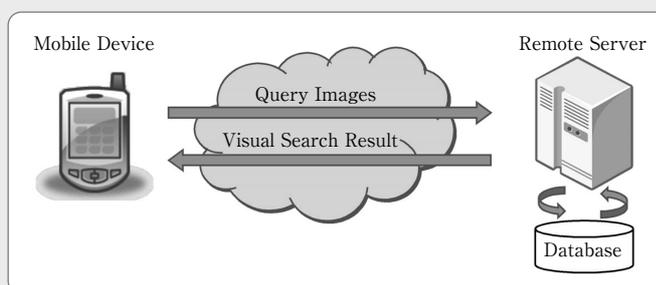


図1 Mobile Visual Search

代表サービス

イメージを掴んでいただくために、まず代表的なMobile Visual Searchサービスを以下で紹介したいと思います。なお、著作権上の問題で、本稿では各サービスのスナップショットを載せることはできませんが、URLを記載しておきますので、気になったサービスがありましたら、是非ご自身で確認してみてください。

(1) Google goggles

<http://www.google.com/mobile/goggles/>

絵画、建築物、本/DVD/CDなどの

製品を認識できます。認識できると、これをクエリとしたWeb検索結果の表示も行ってくれます。また、画像内の文字認識にも対応しており、被写体自体を認識できない時でも、被写体上に印字された文字を使ってWeb検索結果を表示したり、異なる言語への翻訳もしてくれたりします。なお文字認識は、本原稿執筆時点では日本語には対応していません。本サービスでは、画像撮影された数独パズルを解くといったユニークな機能もあります。また、バーコードやQRコード読取りにも対応しているなど、非常に多機能なサービスとなっています。

(2) CamFind

<http://camfindapp.com/>

Google gogglesと同様のサービスです。認識結果をクエリとして、同じ製品の価格比較情報やオンラインショッピングの情報までも併せて提示してくれます。認識精度が高いと評判のサービスです。

(3) kooaba

<http://www.kooaba.com/>

Mobile Visual Search向けの画像認識プラットフォームサービスです。このサービスの提供元会社は、高速・高精度な局所特徴量SURFの開発者であるHerbert Bayによって創設 

☞ されました。本サービスでは、Mobile Visual Searchアプリケーションを考える事業者向けにAPIを提供しています。事業者はエンドユーザに認識させたい対象を自分でデータベース

登録することができます。登録対象は任意ですが、前者二つのサービス同様に、印刷物やCD、DVD、書籍などの商品、建築物などが認識可能とされています。このサービスではAPI利用に

加えて、オプションでメディアカバー画像データベースやワインラベルデータベースも提供されています。

技術

Mobile Visual Searchサービスの核となる画像認識技術は現在も精力的に研究開発されています。画像認識のフレームワークは、最も一般的な例としては図2に示すように、(a) 特徴量抽出部と (b) データベース画像とのマッチング部から構成されます。初めに入力されたクエリ画像から特徴量 (Descriptor) を抽出し、次いで、予めデータベースに格納された訓練画像とのマッチングを行うことで、クエリ画像がどのデータベース画像に類似するかを算出し、これをランキング形式で出力します。

Mobile Visual Searchにおいては、その場で所望の情報を提示するために、携帯端末で撮影されたら即座に回答を返す高速なレスポンス性が求められます。この観点で (a)、(b) それぞれの技術動向を以下に述べます。

(a) 画像認識の分野においては、特徴量として画像中の特徴的な点近傍の

勾配情報から特徴量記述を行うSIFTがしばしば用いられますが、精度が高い一方、1点あたりの情報量が多い、特徴点抽出・記述に計算量を要するなどの課題があります。このためMobile Visual Searchや組込み端末処理向けに、省メモリ、高速演算可能な局所特徴量の研究開発が近年盛んになっています。また、局所特徴量そのものをネットワークを介してリモートサーバに送信することを想定して、この圧縮方式や、Mobile Visual Searchアプリケーション間での相互運用性を検討するMPEG標準化も行われつつあります²⁾。局所特徴量については、いくつかチュートリアルや性能比較文献も出

されています^{3) 4)}ので、さらに詳しいことをお知りになりたい方は、そちらも参照してください。

(b) マッチング部ではクエリ画像から抽出した特徴量に最も距離の近い特徴量を有する訓練画像を探索しますが、これは大量の局所特徴量を扱う処理となるため、高速な近傍探索処理が必要となってきます。Mobile Visual Searchなど大規模なデータベース画像を検索対象とする場合には、多少の探索誤りを許容する近似最近傍探索がよく用いられ、これには木構造を用いるもの (ANNなど) や、ハッシュ法を用いるもの (LSHなど) が代表的です。

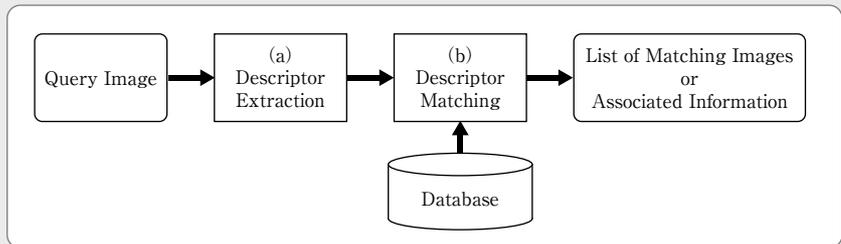


図2 画像認識のフレームワーク

むすび

撮影画像から物を認識して、関連する情報を検索・提示するMobile Visual Searchについて、そのサービス例や求められる技術について紹介しました。

近年、スマートフォンなどの携帯端末の高性能化と、ネットワークのブロードバンド化に伴い、簡便な入力インタフェースとしての画像利用サービスが増えつつあります。その一つとしてMobile Visual Searchが広がっていくことと、これを支える画像認識技術の研究開発が進められて精度・速度がさらに向上していくことが今後期待されます。

(2014年1月8日受付)

参考文献

- 1) O. Marques: "Mobile Visual Search", <http://www.slideshare.net/foerdereverein/mobile-visual-search> (2012)
- 2) CDVS (Compact Descriptor for Visual Search), <http://mpeg.chiariglione.org/standards/mpeg-7/compact-descriptors-visual-search>
- 3) 藤吉: "画像局所特徴量と特定物体認識 - SIFTと最近のアプローチ -", <http://www.vision.cs.chubu.ac.jp/CVTutorial/>
- 4) J. Heinly, E. Dunn and J. Frahm: "Comparative Evaluation of Binary Features", European Conference on Computer Vision (2012)



しまむら じゅん
島村 潤

1998年、大阪大学工学部情報システム工学科卒業。2000年、奈良先端大情報科学研究科修士課程修了。同年、NTTサイバースペース研究所入社。現在、NTTメディアインテリジェンス研究所主任研究員。主にコンピュータビジョン、複合現実感に関する研究開発に従事。博士(工学)、正会員。