

サイヤスカメラ: 大規模特定物体認識のショッピングサービスへの適用 SAIYASU CAMERA: An Approach of Large-Scale Object Recognition for Online Shopping

山口 修司[†] 日暮 立[†] 西村 修平[†] 鈴木 智絵[†] 岩崎雅二郎[†]
Shuji Yamaguchi Tatsuru Higurashi Shuhei Nishimura Chie Suzuki Masajirou Iwasaki

1. はじめに

近年においてもインターネットショッピングにおける商品検索は商品分類やテキスト検索が主流であるが、スマートフォンによって撮影された商品画像を用いた検索が新たな検索手段として注目されている。商品画像を用いた検索手法として、特定物体認識は、画像から局所特徴量を抽出しデータベースから類似度が高い物体を検索し物体を特定する手法として知られている^[1]。この手法では事前にデータベースを構築するため、認識する物体数が大規模になるほどデータベース構築コストが増大し、かつ実用的な速度での検索が困難になるという課題がある。本論文では、ヤフーショッピング^[2]の商品検索アプリ "サイヤスカメラ" を事例として、数億単位の局所特徴量を対象とする物体認識システムの構築事例について述べる。

2. 大規模商品への特定物体認識の適用

特定物体認識により商品を確認するには認識対象とするすべての商品の画像から抽出した特徴点座標および局所特徴量をデータベース(特徴量データベース)化して、認識時には撮影された画像の局所特徴量を基にデータベースを検索する。本システムにおける認識時の流れは以下の通りである。

1. 撮影画像から特徴点座標および局所特徴量を抽出
2. 撮影画像からの個々の局所特徴量により特徴量データベースを検索
3. 検索された商品画像ごとに投票
4. 投票数上位の商品画像について撮影画像との特徴点座標の照合

単一の画像から数百の特徴量が抽出されるので、上記流れで最も時間を要する処理が局所特徴量の検索であり、商品数が多い場合には現実的な時間では処理できなくなる。そこで、以下の施策を行った。

1. 登録検索処理の高速化のための特徴量データベース(商品認識サーバ)の分割
2. 各特徴量データベースにおける特徴量検索の並列化
3. 特徴量検索にグラフ型インデックスの導入

近年、高次元データの近傍検索技術が発展してきており、我々が開発したグラフ型検索インデックス技術^{[3][4]}(以下、NGT)と直積量子化手法^[5]を組み合わせることで大規模な特徴量の近傍検索を実現している。

3. 商品検索システム

“サイヤスカメラ”のシステム構成を図1に示す。図中の商品認識サーバは大規模特定物体認識機能を提供するサーバである。商品データベースサーバは商品名等の商品情報を特徴量データベースと共通の商品ID体系で関連付けたデータベースである。フロントエンドサーバは商品認識サーバ

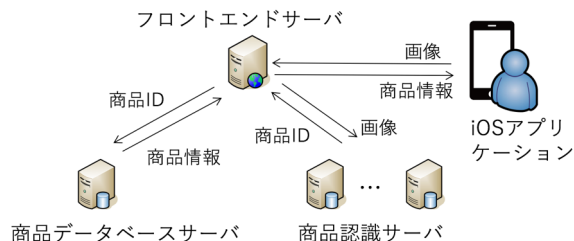


図1 システム構成

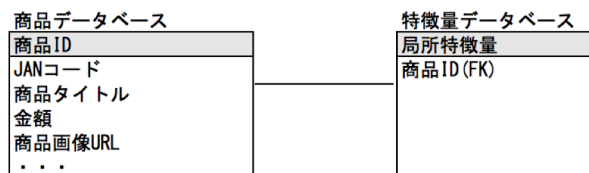


図2 データベース ER 図

と商品データベースサーバへのリクエストをまとめる機能を有したWEB APIである。iOSアプリケーションは、スマートフォンのアプリケーションであり、画像の撮影や認識の実行等を行う。

3.1 商品認識サーバ

商品認識サーバは、送信された画像を受け、特定物体認識を実行し、結果を返却するWEB APIとして機能する。大規模特定物体認識には、NGTを用いて作成した特徴量データベースを利用している。特徴量データベースは商品データベースと商品名等の商品情報を共通の商品ID体系で関連付けている。両データベースのER図を図2に示す。特徴量データベースと商品データベースは以下の手順で作成する。

1. 対象とする全ての商品情報、商品画像を取得
 2. 主キーとして商品IDを発行しながら商品データベースに全商品を登録
 3. 商品画像から局所特徴量を抽出し商品IDと関連付け
 4. 特徴量データベースに局所特徴量・商品IDを登録
- なお、認識処理時間および精度向上のために特徴量データベースを分割し、それに対応して複数の商品認識サーバを設置する。

3.2 フロントエンドサーバ

複数の商品認識サーバと商品データベースサーバをまとめ、iOSアプリケーションからのリクエストを受ける機能としてフロントエンドサーバを定義する。フロントエンドサーバの処理の流れを以下に示す。

[†]ヤフー株式会社

1. 全ての商品認識サーバにリクエストし認識された商品の商品 ID を取得
2. 1. で得られた全ての商品 ID の商品情報を商品データベースサーバから取得
3. 認識結果として商品情報を xml に整形し返却

3.3 iOS アプリケーション

本実装では、“サイヤスカメラ”は iOS のアプリケーションとした。開発は swift 言語で行い、1)商品を読み込んで画像を取得する機能、2)取得した画像をフロントエンドサーバに送信し認識結果を取得する機能、3)認識結果からヤフーショッピングの購入ページへ移動できる機能、4)認識履歴を保存する機能を実装した。なお、商品棚の多数の商品を一回のスキャン操作により認識できるので、所望の商品を効率良く探索することが可能である。

4. 認識結果および性能

現在公開中のサイヤスカメラで使用しているデータおよびシステムにおける認識結果や性能を示す。

4.1 対象データ

データは、ヤフーショッピングに出品されている商品内で、特定物体認識の特性から形状が変化せず、平面（立体物でも正面がほぼ一意に決まる）物を対象として、ドリンク・水・お酒、食料品の一部、キッチン・日用品の一部、本・雑誌・コミック、そして音楽ソフト・映像ソフトの 5 カテゴリーを選定し利用した。それぞれの件数と得られた局所特徴量数を表 1 に示す。対象の商品について、画像（画像サイズ=250 または 400 ピクセル）と商品情報を取得し、特徴量データベース及び商品データベースを作成した。また、本測定ではカテゴリごとに 1 台、計 5 台の商品認識サーバを構築した。

4.2 認識結果

認識結果の UI 表示の 1 例を図 3 に示す。図 3 は、飲料の商品パッケージを撮影した場合の認識結果であり、上部の画像が商品の撮影画像を示し、下部の 3 商品が認識結果を示すものである。この例中の 3 商品においては、撮影した物体と認識した商品が同一であるため、期待通りの商品情報が取得できていることがわかる。ただし、図 4 に示すような飲料の認識結果の例において、商品ロゴが一致している商品が複数認識されているが商品 C については書籍の結果が認識されている。このようなカテゴリ違いの商品が認識されてしまう点が本システムの課題としてあげられる。

4.3 システム性能

システムの性能試験としてベンチマークテストを実施した。試験には Apache Bench^[6]を利用し、同時並行コネクシ

カテゴリ名	商品数	局所特徴量数
ドリンク・水・お酒	35,950	14,032,637
食料品	30,049	16,941,137
キッチン・日用品	35,402	13,877,068
本・雑誌・コミック	467,008	116,501,656
音楽・映像ソフト	384,949	8,909,9130

表 1 対象データ



図 3 UI 表示の 1 例



図 4 認識結果の 1 例

ョン数を 5、総コネクション数を 100 として評価を実施した。結果、1 商品につき平均で 1.7sec のレスポンスタイムとなり、商品を読み込んで検索する一連の動作を阻害しない実用的なレスポンスタイムが達成できていることがわかった。

5. おわりに

本論文では、ヤフーショッピングの商品検索アプリ “サイヤスカメラ” を事例として、1)特徴量データベースの分割、2)検索の並列化、および、3)グラフィックの導入、により億単位の局所特徴量を用いた商品認識システム構築について述べた。

本システムにより、スマートフォンで撮影された商品画像から実用的な時間内での商品認識が実現できた。したがって、スマートフォンにより効率良く商品が購入可能であると考えられる。一方で、同様の文字が同様のフォントで記載された商品であるがカテゴリが違う場合等に誤認識が起こるパターンが有ることがわかっている。これについては、画像からカテゴリを推測する技術等の組み合わせにより解決する等の対応方法が考えられ、今後の課題としたい。

なお、本手法を用いて開発された“サイヤスカメラ”システムは一般に公開されている^[7]。

参考文献

- [1] Sivic, J. and Zisserman, A., “Video Google: A text retrieval approach to object matching in videos”, *Proc. of IEEE Inter. Conf. on Computer Vision*, pp.1470-1477, 2008
- [2] “ヤフーショッピング”, <https://shopping.yahoo.co.jp>
- [3] “NGT”, <https://research-lab.yahoo.co.jp/software/ngt/>
- [4] Iwasaki, M., “Pruned Bi-directed K-nearest Neighbor Graph for Proximity Search”, *Proc. of Inter. Conf. on Similarity Search and Applications*, pp. 20-33, 2016.
- [5] Jegou, H., Douze, M. and Schmid, C., “Product quantization for nearest neighbor search”, *IEEE trans. on pattern analysis and machine intelligence*, 33(1), pp.117-128, 2011.
- [6] “Apache HTTP server benchmarking tool”, <https://httpd.apache.org/docs/2.4/programs/ab.html>
- [7] “サイヤスカメラ”, <https://itunes.apple.com/jp/app/サイヤスカメラ/id1173810477>