

複数の検出基準による画像内の類似物体検出

内田 有人^{1,a)} 丸山 一貴^{1,b)}

概要：機械学習により一枚の画像から高精度で物体を検出できるようになった。しかし、学習を用いる手法では、検出したい物体画像を事前に学習させておく必要がある。例えば、様々な商品をカメラで一枚の画像に収め、その画像から特定の商品を検出するような場合だと、事前の学習は困難である。また、人がある画像を見て物体を計数したいと考えた場合、どういう基準で同じ物体と定義するかはそのたびに変わる可能性がある。そこで、本研究では一枚の画像から複数の検出基準による類似物体検出を事前の学習なしで行う手法を提案する。本研究を行うことにより、計数したい物体をその場でカメラに収めることで、その物体を複数の基準で検出できるようになる。これにより、商品在庫をデータで管理していない小規模な店舗などでも在庫管理を容易に行うことができるようになる。実際に SIFT、DoG、SIFT と色、DoG と色の 4 つの特徴を使用した類似物体検出方法を検証した。

1. はじめに

近年、機械学習により様々な物体を検出できるようになった [1]。しかし画像の中に検出させたい物体があるとき、その物体があらかじめ学習されている必要がある。また、その物体が車や人といった一般的な物体ではなく、特定の食品のパッケージや特定の硬貨などの限定的なものである場合、あらかじめ学習されている可能性は低い。

本研究では特定の物体の類似物体を検出したい場合を想定する。例えば図 1 のように複数種の硬貨が複数枚ずつ映ってる画像から図 2 の赤色で囲まれている百円硬貨と類似した物体を数えたいとき、青で囲まれている百円硬貨を数えたいのか、緑で囲まれている他の硬貨を含めた全ての硬貨を数えたいのかはコンピュータで判断することは難しい。また汚れて変色した百円硬貨を除きたい場合や、色が同じ硬貨のみを数えたい場合も考えられる。これらを踏まえて、細かい特徴が同じ、輪郭が同じ、細かい特徴と色が同じ、輪郭と色が同じ、の 4 種類の検出基準で物体検出を行うことを目的とする。これにより事前の学習なしで画像から、4 つの基準で類似物体を検出し、計数することが可能となる。

2. 関連研究

横溝ら [2] は Tow Circle Algorithm をベースに一枚の画



図 1 硬貨の画像



図 2 検出基準の例

像から半径が同一である 2 つの実円物体の検出を行った。しかし、この手法では円以外の形状の検出はできないため本研究とは異なる。

藤橋ら [3] は CNN で物体画像を学習させて、画像内のその物体の個数と位置の推定を行った。これにより、あらかじめ登録された商品であれば、画像内から登録された商品

¹ 明星大学情報学部情報学科
Department of Information Science, School of Information
Science, Meisei University

a) 15j5024@stu.meisei-u.ac.jp

b) kazutaka.maruyama@meisei-u.ac.jp

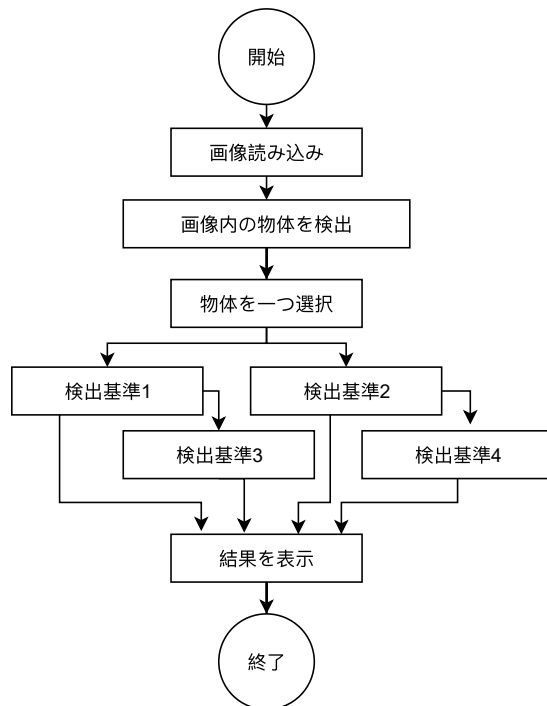


図 3 提案手法の流れ

の個数と位置を特定することができるようになった。しかし、この手法は学習を用いており、事前の登録が必要なため本研究とは異なる。

Tareen[4] は、一枚の画像から特徴を検出するアルゴリズムである SIFT, SURF, KAZE, AKAZE, ORB, BRISK の精度を比較した。SIFT が最も精度が高いという結果が示されていることから、本研究でも特徴を検出するアルゴリズムに SIFT を用いる。

3. 提案手法

まずはじめに、検出したい物体が写っている画像に対して Selective Search を行い、物体を検出する。その後、ユーザは検出された物体から一つを選択する。選択した物体と類似する物体を以下の 4 つの基準で検出する。

検出基準 1：特徴が同じ物体の検出

検出基準 2：輪郭が同じ物体の検出

検出基準 3：特徴が同じで色も同じ物体の検出

検出基準 4：輪郭が同じで色も同じ物体の検出

これらの基準で類似物体検出を行い、その結果を提示する。全体のフローチャートを図 3 に示す。

3.1 物体検出

類似物体を検出する前処理として、画像内の全ての物体の検出を行う。本研究では物体検出に Selective Search[5] を用いる。Selective Search とはピクセル単位で類似する領域をグルーピングしていくことで候補領域を選出するアルゴリズムである。

3.2 検出基準 1

検出基準 1 では特徴が同じ物体を検出する。特徴とは物体の模様やパッケージに描かれた絵などを想定する。特徴の検出には SIFT[6] を用いる。SIFT とは特徴点の検出、特徴量の記述を行うアルゴリズムである。特徴点の検出には、異なるガウシアンフィルタで平滑化された画像の差分により輪郭を検出する、DoG(Difference-of-Gaussian) を使用する。特徴量の記述には、その周辺領域の勾配強度と Gaussian 重みの勾配方向ヒストグラムを用いる。この処理により、スケールと回転に不変な 128 次元の特徴量を得る。その後、検出した特徴量同士のマッチングにより、同じ特徴を持つ物体同士を類似物体として検出する。

3.3 検出基準 2

検出基準 2 では輪郭が類似する物体の検出を行う。輪郭の検出には DoG を用いる。DoG とは前節で述べたように異なるガウシアンフィルタで平滑化された画像の差分により輪郭を検出するアルゴリズムである。DoG で求めた輪郭同士の形や面積を比較して類似度を求める。これにより類似した輪郭と大きさを持つ物体を検出する。

3.4 検出基準 3 と検出基準 4

検出基準 3 と検出基準 4 は、それぞれ検出基準 1 と検出基準 2 に加えて、色も類似する物体を検出する。色を使った検出には、物体の代表色を用いる。代表色は物体画像を 1×1 にリサイズして求める。リサイズには OpenCV の INTER_AREA のアルゴリズムを用いる。これにより得られた代表色の値を比べて類似度を求める。

4. 実装

4.1 物体検出

画像全体を物体として検出するのを防ぐため、全体画像の 50%以上の面積の物体を除外した。またノイズの検出を防ぐため、全体画像の 1%以下の面積の物体除外した。検出した物体同士が 90%以上が重なっていて、高さや幅がそれぞれ 90%以上同じである場合は重複する物体とみなし除外した。また、物体を検出する際、物体の端が見切れて検出されることがあるため、検出した物体の高さと幅を 1.2 倍した。

4.2 検出基準 1

Selective Search で検出したそれぞれの物体画像から SIFT 特徴量を抽出し、特徴量同士を比較することで対応点マッチングを行う。その後、Homography 行列で対応点 4 組を算出する。対応点 4 組の取得の際には、RANSAC[7] を用いた外れ値除去を行うことで、Homography 行列算出時のロバスト性を向上させる。算出した Homography 行列を用いて類似物体の位置を特定する。



図 4 食品の画像

4.3 検出基準 2

Selective Search で検出したそれぞれの物体画像から DoG を求めて輪郭を抽出する。その後、形を比較するため物体画像を縦に分割する。本研究では 20 分割とした。分割されたそれぞれの画像の中で物体の両端の長さを求める。それぞれの長さ同士を比較して何%一致しているかを求める。同様に横に分割した場合の比較も行い何%一致しているかを求める。これらの一致度の平均で類似度を求める。本研究では 80% 以上一致する場合を類似物体とみなす。

4.4 検出基準 3 と検出基準 4

Selective Search で検出したそれぞれの物体画像に対して第 3.4 節で説明した手法で代表の色を決める。類似度は代表色の HSV で比べる。これにより、画像の色相、彩度、明度の 3 つの成分からそれぞれ閾値を決める。色の類似度を比べる上で、彩度や明度は照明などに左右されやすいため、許容範囲を大きくする。また、色相は照明によって左右されにくい成分のため、許容範囲を狭くする。本研究では H を 0~180, S を 0~255, V を 0~255 までの範囲で表し, H, S, V の差の絶対値がそれぞれ 10, 80, 80 未満であることを類似の条件とした。

5. 実験

5.1 実験条件

撮影は照明の点いた屋内で行った。撮影する機材は ZenFone 4 Pro ZS551KL (ASUS_Z01GD) を使用した。撮影する物体は百円硬貨、五十円硬貨、十円硬貨がそれぞれ 4 枚ずつ並べられているものと、3 種類の食品が 4 個ずつ並べられているものを、それぞれ 1 枚に納まるように真上から撮影する。実際に実験で使用した硬貨の写真を図 1、食品の写真を図 4 に示す。

5.2 物体検出

Selective Search で物体検出を行った結果を図 5、図 6 に示す。赤枠は Selective Search で検出した部分を表す。この処理ではすべての物体を一回ずつ検出することを想定する。しかし、図 5、図 6 とも、画像内にある全ての物体

を検出したが、同じ物体を複数回検出する結果となった。

5.3 類似物体検出

類似物体検出では、検出させたい物体を Selective Search で検出した物体の中から一つ選び、その物体の類似物体を検出する。本実験では、硬貨と食品の両方とも一番左上の物体をユーザが選択した、という設定で検出を行った。4 つの基準を用いた検出結果を以下の節で示す。また、以下の節で説明する図の赤枠で囲われている部分は検出した物体を示す。

5.3.1 検出基準 1

検出基準 1 による検出結果を図 7、図 8 に示す。この手法では SIFT を使い、同じ模様やパッケージの絵が描かれているものを検出することを想定する。

図 8 の類似する食品の検出では、選択した物体と同じパッケージのものを全て検出し、期待通りの結果となった。しかし、図 7 の類似する硬貨の検出では、裏向き（数字の 100 が印字された面）の百円硬貨全てを検出することを想定していたが、上段右から 2 番目の百円硬貨を検出できなかった。

5.3.2 検出基準 2

検出基準 2 による検出結果を図 9、図 10 に示す。この手法では DoG を使い、同じ輪郭を持つ物体の検出をすることを想定する。

図 9 の類似する硬貨の検出では硬貨全てを検出し、期待通りの結果となった。しかし、第 4.3 節で述べた手法は回転する物体の検出は想定していないため、図 10 の類似する食品の検出で右一列の回転して置かれている食品を検出できなかった。

5.3.3 検出基準 3 と検出基準 4

検出基準 3 による検出結果を図 11、図 12 に示し、検出基準 4 による検出結果を図 13、図 14 に示す。この手法ではそれぞれ検出基準 1、検出基準 2 の検出結果から色が類似しないものを除外するような結果を想定する。

今回の実験では SIFT 特徴量が同じで色が異なる物体がなかったため、検出基準 3 による検出結果は検出基準 1 と同じ検出結果であることが想定され、実際に同じ検出結果となった。

図 14 の類似する食品の検出では、検出基準 1 の検出結果と比べ異なる絵柄の食品が除外され、期待通りの結果となった。しかし図 13 の類似する硬貨の検出では、選択した百円硬貨とは異なる色の十円硬貨は除外されたが、百円硬貨と類似する色である五十円硬貨が 4 枚中 2 枚が除外される結果となった。

6. 考察

実験結果をまとめたものを表 1、表 2 に示す。検出基準 1 を使った類似物体検出では、食品の検出では精度 100% と

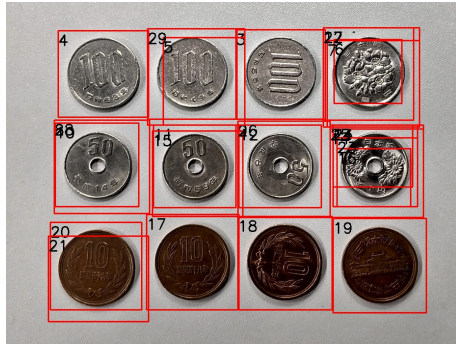


図 5 硬貨を Selective Search で物体検出した結果

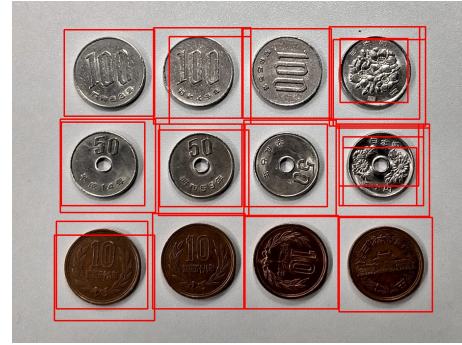


図 9 検出基準 2 (DoG) を使って硬貨を類似物体検出した結果



図 6 食品を Selective Search で物体検出した結果



図 10 検出基準 2 (DoG) を使って食品を類似物体検出した結果



図 7 検出基準 1 (SIFT) を使って硬貨を類似物体検出した結果



図 11 検出基準 3 (SIFT と色) を使って硬貨を類似物体検出した結果



図 8 検出基準 1 (SIFT) を使って食品を類似物体検出した結果



図 12 検出基準 3 (SIFT と色) を使って食品を類似物体検出した結果



図 13 検出基準 4 (DoG と色) を使って硬貨を類似物体検出した結果



図 14 検出基準 4 (DoG と色) を使って食品を類似物体検出した結果

表 1 硬貨を検出した結果

	想定した検出数	実際の検出数	精度
検出基準 1	3	2	67%
検出基準 2	12	12	100%
検出基準 3	3	2	67%
検出基準 4	8	6	75%

期待通りの結果となったが、硬貨の検出では百円硬貨を全て検出することはでき精度は 67%となった。その理由として、食品は多彩な色で絵柄が印刷されているのに対して、硬貨の文字は同色で記されているため特徴が取りにくかったためだと考えられる。また、硬貨は光を強く反射するため、光の当たり方によって特徴が変わってしまうという要因も考えられる。このことから、SIFT 以外の特徴量抽出方法なども検討していく必要があると考えられる。

検出基準 2 を使った類似物体検出では、円形である硬貨は精度 100%であったが、四角形である食品は回転したものを検出しなかった。これは実装段階で回転を想定していなかったためである。解決策として、類似度を求めるとき、片方の物体を回転させながら複数回求め、一番類似度が高いものを採用することで回転に対応できると考える。

検出基準 3, 検出基準 4 を使った検出では、硬貨などの光を反射する物体では期待通りの結果にならないことが分かった。これについては類似度を求める際のパラメータを再度調整していく必要があると考えられる。

表 2 食品を検出した結果

	想定した検出数	実際の検出数	精度
検出基準 1	4 個	4 個	100%
検出基準 2	12 個	9 個	75%
検出基準 3	4 個	4 個	100%
検出基準 4	4 個	3 個	75%

7. まとめ

本論文では、一枚の画像から 4 つの検出基準を用いて類似物体検出する手法を提案し、実際に SIFT, DoG, SIFT と色, DoG と色の 4 つの手法を使った類似物体検出を検証した。その結果、SIFT を使った類似物体検出では食品の類似物体を正しく検出した。DoG を使った類似物体検出では類似した形状の硬貨を正しく検出した。

今後の改善点としては、光を反射する物体の特徴を抽出する手法の模索、DoG を使った検出方法を回転に対応させることなどが上げられる。また Selective Search では重複する物体を完全には除外できなかったため、パラメータの再調整などを行う必要があると考えられる。

参考文献

- [1] 秋山 瑞樹, 柳井 啓司: 特定物体認識手法による大量画像を用いた一般物体認識, 第 13 回画像の認識・理解シンポジウム (MIRU2010), pp.1348–1355, 2010.
- [2] 溝上 尚輝, 呉 海元, 陳 謙, 鈴木一正: 一枚の画像から半径が同一である 2 つの実円物体の検出, 情報処理学会研究報告, Vol.2016-CVIM-202, No.16, 2018.
- [3] 藤橋 一輝, 木村 雅之, 金崎 朝子, 小沢 順: 半教師あり学習による商品画像の個数と位置の同時推定, 人工知能学会第 32 年全国大会, 4M2-04, pp.1–3, 2018.
- [4] Shaharyar Ahmed, Khan Tareen: A Comparative Analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK, 2018 International Conference on Computing, Mathematics and Engineering Technologies, pp.1–10, 2018.
- [5] Koen E. A. van de Sande, Jasper R. R. Uijlings, Theo Gevers, and Arnold W. M. Smeulders.: Segmentation as selective search for object recognition, in Proceedings of the 2011 International Conference on Computer Vision, ICCV '11, pp.1879–1886, 2011.
- [6] D. G. Lowe: Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, Vol.60, pp.91–110, 2004.
- [7] M. Fischler and R. Boles, Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, Common. ACM, Vol.24, No.6, pp.381–395, 1981.