

登録物体数の違いが Bloomier Filter を用いた 特定物体認識手法に与える影響

井上 勝文^{†,††} 黄瀬 浩一[†]

†大阪府立大学大学院 工学研究科 ††日本学術振興会
inoue@m.cs.osakafu-u.ac.jp kise@cs.osakafu-u.ac.jp

1 はじめに

コンピュータビジョンの分野において、物体認識は近年最も盛んに研究されている分野の一つである。特に、SIFT 特徴量 [1] に代表される局所特徴量は、高精度な照合が可能のため、物体認識の分野で広く用いられている。本研究では、物体認識の分野の中でも、物体のインスタンスを認識する特定物体認識に焦点を当てる。

一般的に、特定物体を高精度に認識するには、大量の局所特徴量をデータベースに登録しておく必要がある。このため、局所特徴量を用いた特定物体認識手法には、これらを記録するのに莫大なメモリ量が必要となる問題がある。この問題を解決するために、我々は空間効率の良い Bloomier Filter [2] と呼ばれるデータ構造を用いた手法 [3] を提案し、400 物体を用いた実験より、メモリ使用量における従来手法に対する有効性を示している [4]。しかし、[4] ではさらに物体数を大規模化した場合の有効性について十分な議論がなされていない [5]。そこで、本研究では、さらに物体数を増加させ、物体数の違いが Bloomier Filter を用いた物体認識手法に与える影響について検証する。

2 Bloomier Filter を用いた特定物体認識手法

2.1 Bloomier Filter

Bloomier Filter [2] は、連想配列の一種で、これに登録される要素に関連する値を連想できる。本研究では、データベースに登録される特徴量に対する物体 ID を連想するのに用いる。Bloomier Filter は、Bloom Filter [6] と呼ばれるデータ構造を複数個用いた構成となっている。この Bloom Filter は、自身にある特徴量が登録されているか否かを判断できる。Bloomier Filter は、この特徴量を利用し、特徴量の登録されている Bloom Filter の組合せにより、特徴量に関連する値を連想する。

2.2 データベース作成処理

まず、データベース作成用画像より、PCA-SIFT [7] を用いて特徴ベクトルを抽出する。得られる特徴ベクトル

の次元数は 36 である。次に、得られた特徴ベクトルの各次元の中央値を基準に、特徴ベクトルをビットベクトル化し、これに対する MD5 値を求める。本研究では、128 bit の MD5 値を特徴量とする。

次に、Bloomier Filter と物体 ID の関係について述べる。Bloomier Filter では、特徴量の登録されている Bloom Filter の組合せによってその特徴量に対応する物体 ID を連想する。本研究では、物体 ID を二進数表現し、桁毎に 0 と 1 を連想する Bloom Filter を用意する。ここで、二進数表現された物体 ID を以後“*bit ID*”と呼ぶ。もう少し具体的に説明すると、まず、物体 ID を ν 桁の *bit ID* に変換する。次に、 2ν 個の Bloom Filter を用意する。ここで、*bit ID* の“0”を記録する Bloom Filter を $B_1^{(0)}, B_2^{(0)}, \dots, B_\nu^{(0)}$ 、“1”を記録する Bloom Filter を $B_1^{(1)}, B_2^{(1)}, \dots, B_\nu^{(1)}$ とする。そして、物体 ID i の *bit ID* の j 番目の桁が $\tau (\tau \in \{0, 1\})$ の場合、 $B_j^{(\tau)}$ にこの物体から得られる特徴量を登録する。

2.3 認識処理

まず、データベース作成処理と同様に、検索質問画像から PCA-SIFT を用いて特徴ベクトルを抽出し、これらに対する MD5 値を求める。次に、上記の処理で求めた各特徴量に対して、 2ν 個の Bloom Filter のどれに登録されているかを調べる。このとき、 $B_j^{(\tau)}$ ($j \in \{1, 2, \dots, \nu\}, \tau \in \{0, 1\}$) に登録されていると判断されると、*bit ID* の j 番目の桁を τ とする。そして、求めた *bit ID* を持つ物体に対して、 $s_i \leftarrow s_i + 1/\sqrt{n_i}$ のようにスコアを更新する。ここで、 s_i は物体 ID i の物体のスコア、 n_i は物体 ID i の物体から得られる特徴ベクトル数をそれぞれ意味する。この処理を繰り返し、最大スコアを持つ物体を認識結果とする。

3 実験

3.1 実験条件

本研究では、登録物体数の違いが Bloomier Filter を用いた物体認識手法に与える影響を検証するために、独自に収集した 1002 個の三次元物体を用いて実験を



図 1: 1002 個の三次元物体の例

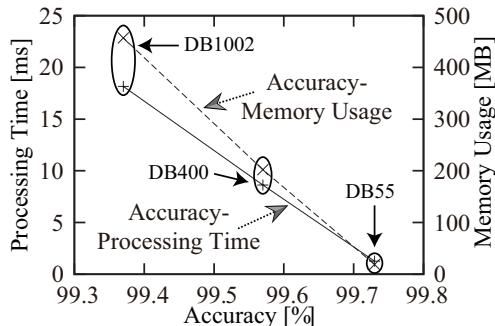


図 2: 登録物体数の違いによる認識率，処理時間，メモリ使用量の関係

行った．まず，我々が独自に収集した 1002 個の三次元物体データセットについて述べる．図 1 に 1002 個の三次元物体の例を示す．本実験では，各物体を 10° 毎 ($10^\circ, 20^\circ, \dots, 360^\circ$) に回転させ，仰角 $0^\circ, 15^\circ, 30^\circ$ の角度からウェブカメラ¹で撮影した画像を，データベース作成用画像とした．一物体あたりの画像枚数は，108 枚である．また，検索質問画像は以下のように用意した．まず，1002 個の三次元物体より 55 個を選び，データベース作成用画像と同様に，各物体を 10° 毎 ($5^\circ, 15^\circ, \dots, 355^\circ$) に回転させ，仰角 $0^\circ, 15^\circ, 30^\circ$ の角度から同じウェブカメラで撮影した画像を用いた．一物体あたりの検索質問画像数は 108 枚である．

本実験では，データベースに登録する物体数を 55 物体，400 物体，1002 物体と変化させ，登録物体数の違いが手法に与える影響について検証する．また，本実験では検索質問に用いた 55 個の三次元物体を 11 物体毎のグループにわけ，5-fold CV により求めた平均認識率・平均処理時間・平均メモリ使用量より，物体数と手法の関係について述べる．以下の実験結果で示す処理時間は，1 枚の検索質問画像の認識に必要な平均の処理時間であり，特徴ベクトルの抽出時間等は含んでいない．

3.2 実験結果

図 2 に登録物体数を 55 物体，400 物体，1002 物体と変化させた場合の結果の内，最も認識率の高かった結果をそれぞれ示す．結果より，登録物体数を増加させると，徐々に認識精度が低下することが分かった．ここで，誤認識した検索質問画像に着目すると，物体に対してテ

クスチャの乏しい方向から撮影されていることが分かった．このことから，さらに認識精度を向上させるためには，テクスチャの豊富さに左右されない特徴抽出手法を用いる必要がある．

次に，処理時間に着目すると，徐々に処理時間が増加していることが分かった．これは，認識処理に用いられる Bloom Filter の数が， $2 \log_2 N$ (N は登録物体数) 個と登録物体数に依存するためである．

最後に，メモリ使用量に着目すると，物体数が増加するにつれてメモリ使用量も徐々に増加することが分かった．実験結果より，さらに物体数を増加させた場合のメモリ使用量を推定すると，10 万物体でおよそ 100 GB 程度のメモリ容量が必要となることが分かった．このことから，さらに大規模な物体認識を実現するには，さらなるメモリ削減法が必要となる．

4 おわりに

本稿では，データベースの登録物体数の違いが，Bloomier Filter を用いた物体認識手法に与える影響について検証した．実験結果より，物体数が増加するにつれて認識精度が低下するものの，1002 個の三次元物体データベースに対して，99.3%以上の認識率を得ることができた．また，処理時間・メモリ使用量に着目すると，どちらも物体数に依存して増加することが分かった．今後の課題として，さらなる物体数の大規模化や，手法の理論的解析といったことが挙げられる．

謝辞 本研究の一部は，日本学術振興会科学研究費補助金 基盤研究 (B) (22300062)，特別研究員奨励費 (22・8970) の補助による．

参考文献

- [1] D. Lowe: “Distinctive Image Features from Scale-Invariant Keypoints”, *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91–110 (2004).
- [2] B. Chazelle, J. Kilian, R. Rubinfeld and A. Tal: “The Bloomier Filter: An Efficient Data Structure for Static Support Lookup Table”, *Proc. 15th Annual ACM-SIAM SODA*, pp. 30–39 (2004).
- [3] 井上，黄瀬：“特定物体認識における Bloomier Filter を用いたメモリ削減法とその実験的評価”，*電子情報通信学会論文誌 D, J93-D*, 8, pp. 1407–1416 (2010).
- [4] 井上，黄瀬：“大規模特定物体認識における bloomier filter を用いたメモリ削減法の有効性の検証”，平成 22 年 電気関係学会関西連合大会講演論文集, pp. 922–925 (2010).
- [5] 井上，黄瀬：“データセットの違いが物体認識に与える影響の解析—特徴ベクトルの一致検索を用いた認識手法の場合—”，*電気学会論文誌 (C)*, **131**, 11 (2011) (掲載予定).
- [6] B. H. Bloom: “Space/Time Trade-offs in Hash Coding with Allowable Errors”, *Commun. ACM*, Vol. 13, No. 7, pp. 422–426 (1970).
- [7] Y. Ke and R. Sukthankar: “PCA-SIFT: A more distinctive representation for local image descriptors”, *Proc. of CVPR2004*, Vol. 2, pp. 506–513 (2004).

¹Logicool Webcam Pro 9000，解像度：640×480