

特徴点削減のため Diverse Density の近似高速化

向井 祐一郎^{†1,a)} 和田 俊和^{†1,b)}

概要: 局所特徴を用いた Content-based image retrieval(CBIR) は, 2 手法に分類することができる. 1 つは, 画像を, 局所特徴を統合した, 1 つの Bag of Features(BoF) ベクトルとして表現して利用する手法である. もう 1 つは, 個々の局所特徴を統合せずに利用する手法である. 我々は, 後者の手法を Multiple instance image retrieval(MIIR) と呼んでいる. MIIR では, Diverse density(DD) を用いた重要度評価による特徴点 (インスタンス) 選択手法が提案されており, 特徴の安定性と弁別性を同時に考慮した検索が可能になっている. しかし, 重要度の評価で用いる DD は, ある特徴点とデータベース内のすべての画像のインスタンスとの距離計算と等しく, 膨大な計算時間を要する. そこで本研究では, 最近傍探索を用いた, DD の高速化を提案する. Nister の画像セットを用いた実験結果から, Image retrieval の精度は保つつ, DD の計算を約 520 倍高速化できることを確認した.

キーワード: Diverse density, Nearest neighbors search

Accelerating Diverse Density for Keypoint Reduction

Abstract: Content Based Image Retrieval (CBIR) using local features can be classified into two types. One is the method using single vector representation that integrates local features detected in an image into a single Bag of Feature (BoF) vector. The other uses multiple instance representation without integration. We call the latter method Multiple Instance Image Retrieval (MIIR). In MIIR, a method for reducing database indexes has been proposed. This method employs the framework of Diverse Density (DD) to represent the importance measure, which means the stability as well as the discriminative power of the feature (instance). This reduction reduces the memory usage and the retrieval accuracy. The computational cost of DD, however, is considerably big, because it has to compute all distances between all combinations of instances. This report presents the approximate computation of DD for MIIR using nearest neighbor search. We confirmed through the experiments that the computational speed of DD becomes 520 times faster on Nister's database while keeping the accuracy.

Keywords: Diverse density, Nearest neighbors search

1. はじめに

画像検索 (Content-based image retrieval: CBIR) に関する手法は, 物体識別, 実世界への情報埋め込みなど多くのアプリケーションで利用されている. SIFT や SURF などの局所特徴量を用いた CBIR の手法として, Bag of Features を用いる手法 [1] があるが, この手法ではあらかじめ全特徴量について, クラスタリングを行う必要があるため, 計算コストが高い.

一方で, Bag of Features を用いることなく, 個々の特徴

からインデックスへの近接性に基づく投票を行い, 最多得票を集めた画像を検索結果とする, Number of match を用いた画像検索の手法 [2] がある. この手法は, 事前にクラスタリングなどの前計算を行う必要が無くシンプルのため, 広く用いられている. また, Number of match の検索精度を向上させるため, Diverse density (DD) を用いて表現する Importance measure を基準とした局所特徴の削減を行う手法が提案されており [3], より検索精度の高い Number of match を行うことが可能になっている. ここで, Importance measure とは, インスタンスが所望の対象を表現できるという安定性と, 分類・識別のための弁別性の高さを表現する尺度である.

^{†1} 現在, 和歌山大学 Wakayama University

^{a)} yuichirw@vrl.sys.wakayama-u.ac.jp

^{b)} twada@ieee.org

しかし、DD は、入力画像のあるインスタンスと、データベース内の全インスタンスとの距離計算を行う必要があり、大規模なデータベースにおいては、その計算が困難である。

そこで本稿では、Radius search を用いて DD の計算コストを削減した、Accelerating diverse density(ADD) を提案する。提案手法は、DD と比べ、計算速度が高速であり、大規模なデータベースに対しても Importance measure の計算が可能であることを示す。

また、実験結果では、計算が高速化されただけでなく、キーポイントの削減を行っても検索精度が低下しにくく、Importance measure としての性質も改善されることを示す。

2. Importance measure による特徴選択

本章では、Importance measure による インスタンスの選択の手法と、その手法を大規模なデータベースに適用する上での問題点について述べる。

2.1 Diverse Density

Diverse Density(DD) は、Multiple instance learning の枠組みで提案されたサンプル選択手法である [4]。

DD は、ネガティブバッグ(たとえば、非選択対象となる画像集合) B^- 中の、すべてのインスタンスから離れ、ポジティブバッグ(たとえば、選択対象となる画像集合) B^+ 中に 1 つでも近いインスタンスが存在すると、DD の出力値(ポテンシャル)は大きくなる。ここで、インスタンス B_{ij} を重心とした、未知のインスタンス x を式 (1) とする。

$$P(x|B_{ij}) \equiv \exp\left(-\frac{\|B_{ij} - x\|^2}{\sigma^2}\right) \quad (1)$$

ここで、 σ はスケールパラメータである。ポテンシャル $P(x|B_{ij})$ は、 x と B_{ij} の L2 ノルムで決まり、距離が遠いと、 B_{ij} が x に与える影響は小さくなる。

式 (1) を用いて、DD における、ポジティブバッグ B_i^+ 、およびネガティブバッグ B_i^- のポテンシャルは、それぞれ式 (2)、および (3) で表される。

$$P(x|B_i^+) \equiv 1 - \prod_{B_{ij} \in B_i^+} (1 - P(x|B_{ij})) \quad (2)$$

$$P(x|B_i^-) \equiv \prod_{B_{ij} \in B_i^-} (1 - P(x|B_{ij})) \quad (3)$$

上式から、DD では x とバッグ内における全インスタンスとのポテンシャルを、NOR 的に統合している。

各バッグのポテンシャルから、DD における x のポテンシャルは式 (4) で表される。

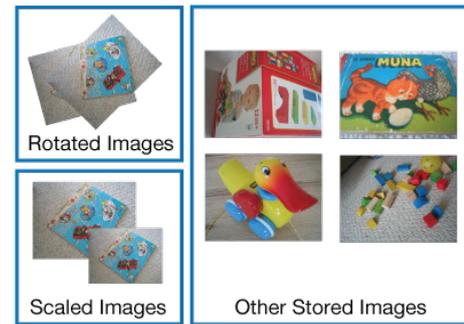


図 1 Importance Measure におけるポジティブバッグとネガティブバッグ

$$DD(x) \equiv \prod_i^m P(x|B_i^+) \prod_i^n P(x|B_i^-) \quad (4)$$

つまり、DD のポテンシャルは、各バッグのポテンシャルを積で統合し計算した値となっている。

2.2 Importance measure による局所特徴の選択

Importance measure としての DD[3] は、データベース中の画像から得られるインスタンスについて、以下の2つの性質を満たすと、値が大きくなる指標である。

- その画像固有に表れる特徴である(弁別性)
- スケール変化や回転に対して頑健である(安定性)

この Importance measure では、2つの性質を同時に考慮するため、DD の各バッグを以下のように割り当てている(図 1)。

- ポジティブバッグ: スケール・回転変換を行って得られる注目画像
- ネガティブバッグ: データベース内の注目画像以外の画像

まず、ポジティブバッグを、スケール・回転変換を行った注目画像とすることで、ポジティブバッグのポテンシャルが高いインスタンスは、変換後に得られる画像から抽出されたインスタンスとも近く、これにより安定性を得る。

次に、ネガティブバッグを、データベース中の注目画像以外とすることで、ネガティブバッグのポテンシャルが大きい局所特徴は、他の画像から得られるすべての局所特徴から遠く、これにより弁別性を得る。

DD は、ポジティブバッグとネガティブバッグのポテンシャルの積で計算されるため、Importance measure が大きい値場合は、安定性と弁別性を同時に満足していることになる。

以上の Importance measure を基準とした Number of match による画像検索では、図 2 に示す通り、検索精度を向上できることが報告されている。

2.3 問題点

前節で説明したとおり、DD を用いた Importance mea-

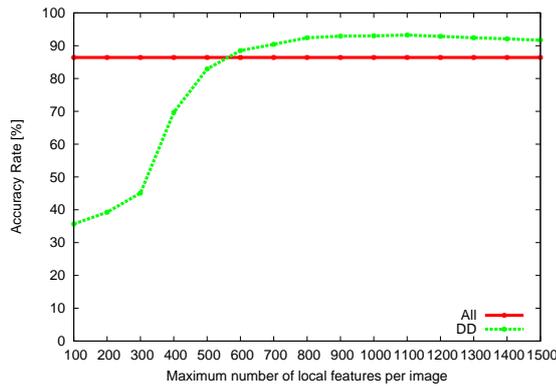


図 2 Importance Measure による特徴削減を用いた時の画像検索の精度。データセットと実験条件は, [3] のものと同様のものを用いた。

sure は, DD のポテンシャルの算出のため, 各バッグの全インスタンスとの比較が必要となる。したがって, その計算量はデータベース内の画像枚数, およびインスタンス数に依存して増加するた, 大規模なデータベースを対象とすると, Importance measure の算出自体が困難となる。

3. 提案手法

本章では, DD の計算で用いるインスタンスの中で, ポテンシャルに与える影響の大きいインスタンス集合を Radius search で列挙し, その集合でポテンシャルを計算することで高速化を図る Accelerating diverse density を提案する。また, Accelerating diverse density を計算する際に, Radius search に必要なインデックスの構築を毎回行う手法と, 各バッグごとにインデックスを構築し, ポテンシャルの計算ではそのインデックスを使いまわす手法の, 2つの手法を提案する。

3.1 Accelerating Diverse Density

Accelerating diverse density (ADD) は, DD のポテンシャルに与える影響の大きいインスタンスのみを列挙し, それらを用いて計算を行うことで, DD の計算の高速化を図る手法である。

まず, DD のポテンシャルに大きく影響を与えないインスタンスについて述べる。 x における, インスタンス B_{ij} のポテンシャルは式 (1) で与えられる。つまり, x における各インスタンスのポテンシャルは L2 ノルムで決まり, インスタンスからの距離が遠いとポテンシャルは小さくなる。そのため, x との距離が遠いインスタンスを計算から省いたとしても, DD のポテンシャルは大きく変化しない。

そこで, x を中心とした半径 R の超球内のインスタンスは DD のポテンシャルへの影響が大きいものとし, それらのインスタンスのみを用いて DD の計算を行う。半径 R の超球内のインスタンスを列挙するために, 以下の Radius search を用いる。

Radius search は, クエリ q について, 以下のように定義

される。

$$RNN(q, P, R) \equiv \{p \in P, d(p, q) < R\} \quad (5)$$

ここで, P は, 検索対象となるデータである。

3.2 Tree Accelerating Diverse Density

Tree tccelerating diverse density(TADD) は, ADD のポテンシャルの計算の時に, 毎回インデックスの構築と Radius search を行う手法である。アルゴリズムを以下に示す。

Algorithm 1 Tree Accelerating Diverse Density

```

1: procedure TADD ( $x; B, R$ )
2:   バッグ集合  $B$  内のすべてのインスタンスでインデックス構築。
3:   for all  $B \in B$  do
4:      $P(B) \leftarrow 1$ 
5:   end for
6:   for all  $I \in RNN(x; B, R)$  do
7:      $B \leftarrow$  インスタンス  $I$  の属するバッグ
8:      $P(B) \leftarrow P(B) \times \left(1 - \exp\left(-\frac{\|I-x\|^2}{\sigma^2}\right)\right)$ 
9:   end for
10:   $p \leftarrow 1$ 
11:  for all  $B$  in  $B$  do
12:    if  $B$  がポジティブバッグ then
13:       $p \leftarrow p \times (1 - P(B))$ 
14:    else
15:       $p \leftarrow p \times P(B)$ 
16:    end if
17:  end for
18:  if  $RNN(x; B, R)$  内に, 各ポジティブバッグのインスタンスが 1 つ以上含まれる then return  $p$ 
19:  elsereturn 0
20:  end if
21: end procedure

```

ここで, B はバッグ集合で, すべてのポジティブバッグとネガティブバッグの和集合である。

手法 1 の 6 行目から 9 行目では, Radius search によって列挙された影響の高いインスタンス集合を元に, 点 x における各バッグのポテンシャルを計算しており, つまり, 式 (2), および (3) 中の, $\exp\left(-\frac{\|B_{ij}-x\|^2}{\sigma^2}\right)$ を求めていることになる。

10 行目から 17 行目では, 点 x における DD のポテンシャルを計算している。これは, 式 (4) で与えられるように, 各バッグのポテンシャルを積の形で統合計算を行っている。式 (2), および (3) から, ポジティブバッグとネガティブバッグのポテンシャルの計算方法が異なるため, バッグのラベルによって分岐を行い, ポテンシャルを計算している。

ここで, Radius search で列挙されたインスタンス集合のうち, ポジティブバッグに属するインスタンス数が 0 個の場合は, DD の定義より, ポジティブバッグのポテンシャル $P(x|B_i)$ は 0 となるため, ADD のポテンシャルとして 0 を返却する。

TADD の計算量は、バッグ集合 B 中のインスタンスから Radius search のインデックスを構築するための計算量を $f_b(B)$, 半径 R での Radius search の計算量を $f_s(B, R)$ とすると, $\mathcal{O}(f_b(B) + f_s(B, R))$ となる。

3.3 Forest Accelerating Diverse Density

TADD は、バッグ集合 B に対してインデックス構築と Radius search を行い、各バッグ $B \in B$ のポテンシャル $P(x|B)$ を計算しているが、各バッグ B ごとにインデックス構築と Radius search を行い、ポテンシャル $P(x|B)$ を求めても、計算としては等価である。

その事を利用して、Forest accelerating diverse density(FADD) を提案する。FADD は、各バッグごとのインデックスをあらかじめ構築しておき、ADD のポテンシャルの計算では、対象となるバッグについてそれぞれ Radius search を行う手法である。アルゴリズムを以下に示す。

Algorithm 2 Forest Accelerating Diverse Density

```

1: procedure TADD ( $x; B, R$ )
2:   前計算: 各バッグ  $B \in B$  についてインデックスを構築.
3:   for all  $B$  in  $B$  do
4:      $P(B) \leftarrow 1$ 
5:   end for
6:   for all  $B$  in  $B$  do
7:     for all  $I \in RNN(x; B, R)$  do
8:        $P(B) \leftarrow P(B) \times (1 - \exp(-\frac{\|I-x\|^2}{\sigma^2}))$ 
9:     end for
10:  end for
11:   $p \leftarrow 1$ 
12:  for all  $B$  in  $B$  do
13:    if  $B$  がポジティブバッグ then
14:       $p \leftarrow p \times (1 - P(B))$ 
15:    else
16:       $p \leftarrow p \times P(B)$ 
17:    end if
18:  end for
19:  if すべてのポジティブバッグ  $B$  について,  $RNN(x; B, R)$ 
   の要素数が 1 つ以上である then return  $p$ 
20:  elsereturn 0
21:  end if
22: end procedure

```

対象となるバッグ集合 B が異なれば、インデックスの再利用は出来ないため、TADD は ADD のポテンシャル計算の時には毎回インデックスの構築を行う。一方 FADD では、各バッグごとにインデックスを構築するため、事前に各バッグについてインデックスを構築しておけば、それらのインデックスを再利用できる。

FADD は、バッグ集合の要素数が $|B|$, それぞれのバッグを $B \in B$ としたとき、前計算の計算量は $\mathcal{O}(|B|f_b(B))$ であり、ADD のポテンシャルの計算量は $\mathcal{O}(|B|f_s(B, R))$ である。

表 1 に、2 つの手法の計算量を示す。

表 1 ADD の各手法の計算量

手法	前計算	ポテンシャル計算
TADD	-	$\mathcal{O}(f_b(B) + f_s(B, R))$
FADD	$\mathcal{O}(B f_b(B))$	$\mathcal{O}(B f_s(B, R))$



図 3 Recognition Benchmark Images

4. 評価実験

本章では、ADD の有用性を評価するために評価実験を行う。評価実験には、2 章で説明した Importance measure を用いた特徴削減と同様の実験を行い、ADD の性能を検証した。

4.1 実験条件

評価実験として、[3] と同様の実験を行った。以下に詳細の実験条件を示す。

4.1.1 データセット

データセットには、D. Nister の Recognition Benchmark Images[5] を用いる。これは、図 3 のように、各オブジェクトについて、撮影方向・照明条件などが異なる 4 枚の画像から構成されるデータセットで、オブジェクト数は 2 5 5 0 個である。

本実験では、上から 1 0 0 個のオブジェクト (4 0 0 枚の画像) を利用し、各オブジェクトについて 1 枚ずつ画像を取り出して学習画像とし、残りの画像すべてをテスト画像としてデータセットを作成し、実験を行う (図 4)。評価では、各オブジェクトにつき 4 枚ずつ画像があるため、4 組のデータセットを作成してそれぞれ実験を行い、識別精度・計算速度ともその平均を算出した。

4.1.2 特徴量

本実験には、局所特徴として OpenCV の拡張モジュールとして提供されている 64 次元の SURF を用いる [6]。

4.1.3 ADD のパラメータ

ADD と、比較として用いる DD のスケールパラメータ σ は 0.1 とした。また、ADD の Radius search には FLANN [7] に実装されている、Randomized KD Trees と Hierarchical Kmeans Tree のインデックスを利用した。この時のインデックスのパラメータは、FLANN で提供されている手法を用いて決定されたパラメータを用いた。



図 4 Recognition Benchmark Images から学習画像とテスト画像を作成する方法. 青枠がそれぞれ学習画像で, 黄枠が対応するテスト画像となる.

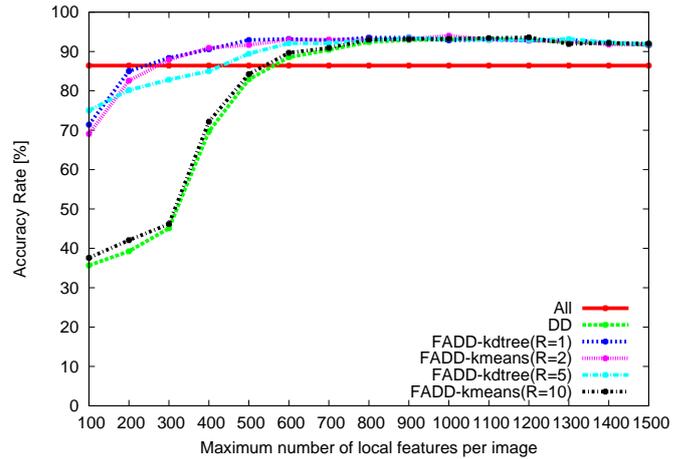


図 6 半径を変化させた時の画像検索の精度

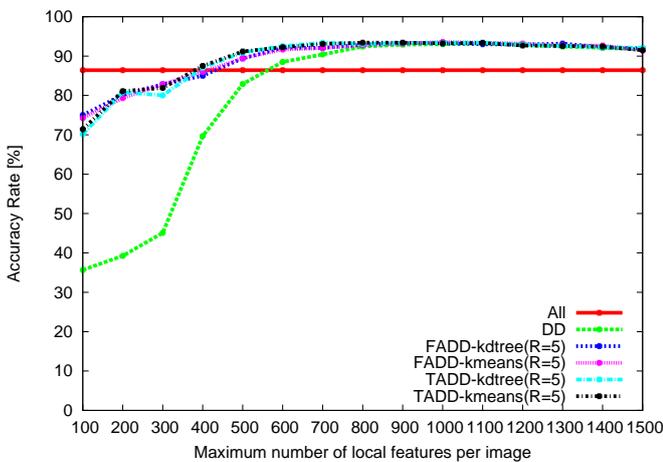


図 5 インデックスを変化させた時の画像検索の精度

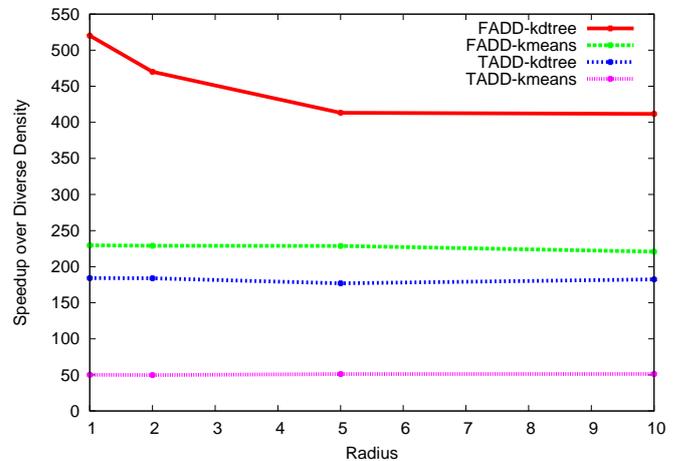


図 7 ADD のポテンシャルの計算速度の比較

4.2 実験結果

4.2.1 Radius search のインデックス

まず, ADD における Radius search のパラメータの, インデックスの種類を変化させた時の画像検索の性能についての実験を行った. 前章で提案した, TADD と FADD, Radius search のインデックスにそれぞれ randomized k-d tree[8] と priority search k-means tree[9] の 2 つを利用した, 計 4 手法について, 検索精度の比較を行った. ここで, 半径は, $R = 5.0\sigma$ とした. 図 5 に結果を示す.

図 5 は, 縦軸に画像検索の精度, 横軸に各データベース画像で用いた局所特徴の最大数を表す. All は, 特徴削減を行わず, すべての局所特徴を用いて Number of match を行ったときのグラフである. また, DD は, 2 章で説明した, Importance measure による特徴削減を行った後, それらの局所特徴を用いて Number of match を行ったグラフである.

実験結果より提案手法による結果は, 特徴数が 800 点より多い場合, DD の結果に収束しているが, 100~700 点では, 提案した 4 手法とも, DD よりも検索精度を上回っていることが確認できた.

4.2.2 Radius search の半径

次に, ADD における Radius search の半径を変化させた時の画像検索の精度について比較を行った. ADD には FADD を, インデックスには Randomized k-d tree を用い, 半径 R を, $1\sigma, 2\sigma, 5\sigma, 10\sigma$ と変化させた. 図 6 に結果を示す.

実験結果より, 半径を大きく取るにつれて, DD による画像検索の精度と結果が一致することが確認できた.

4.2.3 ADD の計算速度

ADD の各手法におけるポテンシャルの計算が, DD のポテンシャルの計算速度と比べ, どの程度高速であるか比較を行った. 図 7 に示す.

図 7 は, 横軸に ADD の Radius search の半径 R を, 縦軸に 1 つの局所特徴につき, DD と比べて ADD が何倍高速であるかを, 表している. この結果から, FADD が TADD に比べて全体的に速度が速いことが確認できた. これは, ポテンシャルを計算する度にインデックスを構築する必要のある TADD と比べ, FADD は, 1 度作ったインデックスを再利用しているため, インデックス構築のコストを大きく削減出来ているためだと考えられる. また, 半径 R を大き

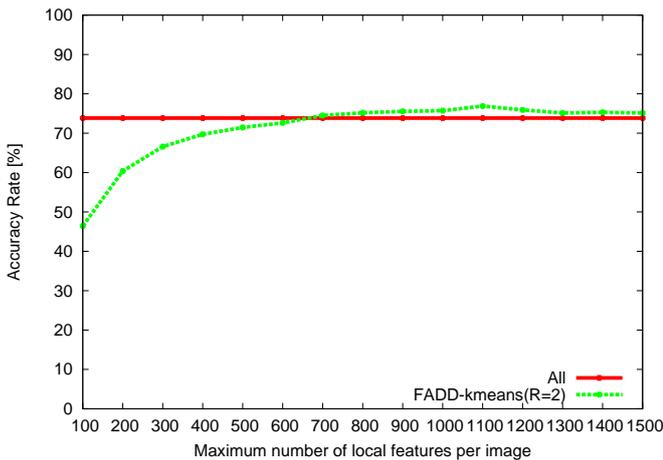


図 8 Nister のデータセットをフルセット用いた時の画像検索の精度

く取るにつれ、計算速度が低下していることも確認できた。

FADD で Randomized k-d tree を用いた時に、DD と比べ、約 5 2 0 倍高速にポテンシャルが計算できることが確認できた。

4.3 Nister のデータセットをフルセット用いた実験

Nister のデータセットを、2 5 5 0 オブジェクトすべて用いた時の、画像検索の精度について比較を行った。ADD のインデックスには Randomized k-d tree を、半径を $R = 2\sigma$ とした。図 8 に結果を示す。

Number of match では、精度が 73.85% に対し、FADD では特徴点数 1100 の時、精度が 76.88% であり、Importance measure による特徴削減が、大規模なデータセットにおいても、画像検索に有効であることが確認できた。

4.4 考察

実験結果より、ADD による Importance measure の局所特徴の削減は、DD と比べて、計算速度の向上だけでなく、検索精度の向上も見込めることを、実験を通じて確認できた。

DD の計算において、式 (1) を式 (6) と置き換えることで、ADD のポテンシャルの計算と等価となる。

$$P(\mathbf{x}|B_{ij}) \equiv \begin{cases} \exp\left(-\frac{\|B_{ij}-\mathbf{x}\|^2}{\sigma^2}\right) & (\|B_{ij}-\mathbf{x}\|^2 < R) \\ 0 & (\text{otherwise}) \end{cases} \quad (6)$$

これは、インスタンスの生成するポテンシャルは Gaussian となるが、その裾野を切り落としたような、ポテンシャルに置き換えている事になる。このようなポテンシャル関数の方が、Importance measure の計算に向いている理由は、無限の台を持つ Gaussian では、僅かなポテンシャルでも、遠方に存在するインスタンス数が膨大になれば、大きなバイアスになるためではないかと推測される。

5. おわりに

本稿では、Radius search を用いた DD の近似高速化手法である、ADD を提案した。ADD は、DD による局所特徴の削減に比べ、高速に計算できることが確認できた。また、本来 DD の計算が困難であった大規模なデータセットについても、Importance measure による局所特等の削減を実現でき、さらに、検索精度の向上を確認できた。

今後は、Importance measure だけでなく、DD を用いた様々なアプリケーションに、ADD を適用することを検討している。

参考文献

- [1] Sivic, J. and Zisserman, A.: Video Google: A Text Retrieval Approach to Object Matching in Videos, *Proceedings of the International Conference on Computer Vision*, Vol. 2, pp. 1470–1477 (online), available from (<http://www.robots.ox.ac.uk/vgg>) (2003).
- [2] Nakai, T., Kise, K. and Iwamura, M.: Camera-Based Document Image Retrieval as Voting for Partial Signatures of Projective Invariants, *Proc. 8th International Conference on Document Analysis and Recognition (ICDAR2005)*, pp. 379–383 (2005).
- [3] Yuasa, K. and Wada, T.: Keypoint Reduction for Smart Image Retrieval, *Proceedings of the 2013 IEEE International Symposium on Multimedia, ISM '13*, Washington, DC, USA, IEEE Computer Society, pp. 351–358 (2013).
- [4] Maron, O. and Lozano-Pérez, T.: A Framework for Multiple Instance Learning, *Advances in Neural Information Processing Systems 10*, Cambridge, MA, MIT Press (1998).
- [5] Nister, D. and Stewenius, H.: Recognition Benchmark Images, <http://www.vis.uky.edu/~stewe/ukbench/>.
- [6] itseez: Repository for OpenCV's extra modules, https://github.com/Itseez/opencv_contrib.
- [7] Muja, M. and Lowe, D. G.: FLANN - Fast Library for Approximate Nearest Neighbors, <http://www.cs.ubc.ca/research/flann/>.
- [8] Silpa-Anan, C. and Hartley, R.: Optimised KD-trees for fast image descriptor matching, Vol. 0, Los Alamitos, CA, USA, IEEE Computer Society, pp. 1–8 (online), DOI: 10.1109/cvpr.2008.4587638 (2008).
- [9] Muja, M. and Lowe, D. G.: Scalable Nearest Neighbor Algorithms for High Dimensional Data, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 36 (2014).