Vol.2013-CVIM-187 No.35
2013/5/30

# Diverse Density を用いた画像検索用キーポイントの削減法

湯浅圭太† 和田俊和† 渡辺顕司†

我々は, SIFT や SURF などの画像の局所特徴量を用いた画像検索システムを FPGA 上に構築するためのプロトタイプを設計している．この設計では Bag of Features (BoF)などのコードブックを用いた特徴記述を行うことは煩雑であるため，局所特徴量そのものを用いた画像検索を行う．この際，一枚の 画像を特徴付けるキーポイントが満たすべき条件としては，１）他の画像には含まれない弁別性の高い局所特徴量を持つこと，２）画像の回転やスケール変化等の変換を受けても検出されやすいこと，という２つが重要である．これらの条件を満足する少数のキーポイントを用いて登録する画像を記述しておけば，省メモリかつ高速で，しかもロバストな画像検索が行えるはずである．このようなキーポイントを抽出するために Diverse Density を用いたキーポイントの絞込み法を提案する．実験では，提案手法により一枚の画像を表すキーポイントの最大数を 100～1500 個に削減したものと，画像から抽出されたキーポイントをすべて利用したものとを比較したところ，キーポイントを削減したほうが，安定な検索が行えることを確認した．また，キーポイントを削減することにより計算コストの削減も実現していることがわかる．

# Keypoint Selection based on Diverse Density for Image Retrieval

KEITA YUASA† TOSHIKAZU WADA†
KENJI WATANABE†

We are planning to construct an image retrieval system using FPGA. For designing this system, we developed a prototype system, which retrieves the image having maximum number of matched local image features with a query image. The reason why we don’t use Bag of Features (BoF) is that codebook referencing may consume considerable time and computational resources on FPGA. For this purpose, the local features describing a stored image should satisfy the following conditions: 1) they should have strong discrimination power from other images, 2) they should be robust against observation distortions including rotation, scaling, and so on. In order to maximize the number of stored images, the number of local features describing a stored image should be minimized. For selecting such “good local features” from all local features, we propose a method based on Diverse Density. In the experiment, we confine the max number of local features describing a single image from 100 to 1500, and our method outperforms the method that uses all local features. Also computing cost improves by reducing the local features.

## 1. はじめに

我々は，画像検索システムをFPGA上に構築することを計画しており，そのためのプロトタイプを設計している．この設計では，FPGA上でBag of Features(BoF)などのコードブックを用いた特徴記述を行うことは煩雑であるため，SIFTやSURFなどの画像の局所特徴量を直接用いて検索を行う．具体的には，クエリ画像から検出した局所特徴量のうちの何個が，登録された画像のインデックスである局所特徴量とマッチするかを調べ，最も多くマッチした画像の検索を行うというものである．

通常，局所特徴量の数，すなわちキーポイントの個数は，１枚の画像当たり数百から数千と膨大となるため，計算能力の限られたシステム上に,数百枚の画像を登録する場合，画像のインデックスとして用いる特徴の個数を削減することが望ましい．この削減が行えれば，登録画像枚数を増やすことができ，しかも検索速度も向上しやすくなる．しかし，画像上のキーポイントをランダムに選び，その場所の局所特徴量を用いるだけでは，キーポイントの個数を減らせば減らすほど，検索の精度は低下してしまう．この問題を解決するには，その画像を検索するのに適したキーポイントをインデックスとして採用すれば良い．

本研究では，画像を検索するのに適したキーポイントとは，次の２つの条件を満足するものであると見なす．

(1)他の画像には含まれない弁別性の高い局所特徴量を持つこと

(2)画像の回転やスケール変化等の変換を受けても検出されやすいこと

という２つの性質が重要である．

これらの条件を満足する範囲内であれば，少数のキーポイントで求められる特徴を用いて登録する画像を記述しても，精度の低下は起きにくいはずである．むしろ，画像登録時に全キーポイントを用いると，他の画像と区別する上で本質的ではなく，しかも，検出されにくいキーポイントも多数含まれることになると考えられる．このため，キーポイントの削減は，単にデータベースへの登録画像枚数の増加や，検索速度の向上だけでなく，精度の向上にも寄与するものと考えられる．

このようなキーポイントを抽出するために，本報告ではDiverse Density (DD)を用いたキーポイントの絞り込み法を提案する．DDは，特徴空間内の点で計算される量であり，その点が，ポジティブラベルを付けられた画像集合に含まれる画像特徴とどれほど近く，またネガティブラベルを付

† 和歌山大学
Wakayama University

けられた画像特徴からどれほど離れているかを表した量である．DDの値が高い特徴は，ポジティブラベルが付けられた画像集合にとって共通性が高く，ネガティブラベルの付けられた画像には含まれない特徴であるということが出来る．

DDのこのような性質から，本研究の場合には，図1に示すように，データベースに含む画像自体はポジティブ，データベース中の他の画像はネガティブとして登録して，DDの値が高い特徴を求めれば良いことが分かる．しかし，登録する画像は１枚であるため，ポジティブラベルをつけた画像は１枚だけになってしまう．この結果，DDの特性の一つである共通性を持つ特徴を探すことができなくなってしまう．

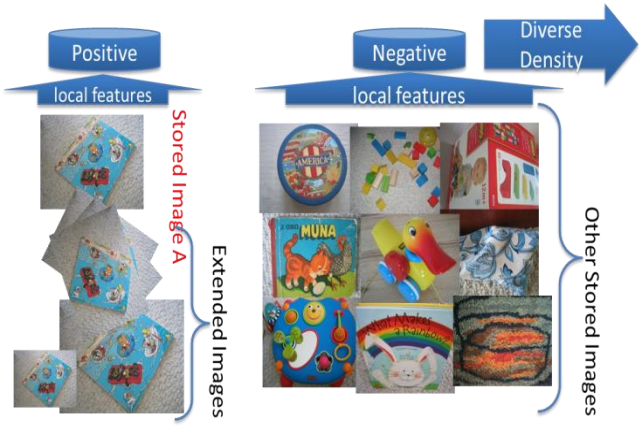


図 1 本研究の基本的アイデア

そこで，図2を見て分かる通り，SIFTやSURF，特にSURFは45度の回転やスケール変化をかけると精度が下がっている． そこで，この問題を解消し，様々な変換に対してロバストな特徴を得るため，登録する画像に対して様々な劣化を与えて局所特徴量を求め，これらをポジティブバッグとして用いる．これによって，ポジティブバッグから抽出される共通性の高い特徴は，劣化に対してもほとんど変化しない安定な局所特徴であるということが出来る．

しかも，登録する他の画像をネガティブとしているため，弁別性能も高くなるはずである．

以下，関連研究について述べ，その後，提案手法，実験の順に述べていく．

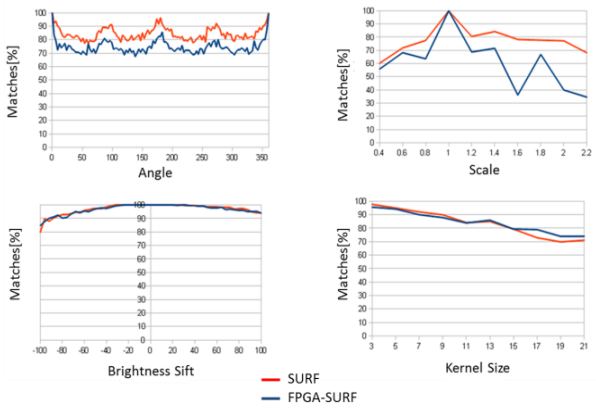


図 2 各変化に対する SURF の精度評価
(参考文献[12]より引用)

## 2. 関連研究

画像検索の分野では，D. Lowe[1]によるSIFTやH. Bayら[2]らによるSURFなどの局所特徴量が，よく用いられる．

具体的には，局所特徴量をクラスタリングしたVisual Wordをコードブックとして与え，この出現頻度を画像の特徴ベクトルとするBag-of-Features[3][4]や，k-means クラスタリングを繰り返して得られる階層的なコードブック（Vocabulary tree）を用いて，これを高速化する方法[5]などが提案されてきた．また，黄瀬ら[6]はハッシュを用いて局所特徴量を投票することで高速な探索を実現している．これらの手法では，画像の位置合わせを必要とせず，高速な，検索を実現するために，幾何学的な変形に対して比較的頑健な局所特徴量がしばしば用いられる．しかし，これらの局所特徴量はキーポイント検出に依存しており，検出に失敗してしまうと，特徴の抽出が全く行われないという問題がある．この問題に対処する方法としては，キーポイント検出を行わず，画像上のグリッド点，あるいはランダムな点で特徴抽出を行う方法がある[11].

これらの手法は，類似画像検索などの用途では，よく用いられ，一定の効果もある．しかし，これらの手法では，データベースに巨大なインデックスを付けるため，FPGAのように比較的小規模なメモリしか持たない装置上での実装には適さない．特に，局所特徴からBoFベクトルを構成する手法では，入力画像から抽出された局所特徴をクエリとしてコードブックに対する最近傍探索を行わなければならないため，処理が煩雑になり，FPGA上での実装には適さない．

これに対し本稿では，BoF ベクトルを作らず，Multiple Instance Learning (MIL) [7]で用いられる Diverse Density [8][9][10]を用いて，弁別性が高く，安定な局所特徴のみを用いて，これらの間の照合だけで画像検索を実現する方法を提案する．

## 3. 提案手法

省メモリ,高速化のためのキーポイントの絞り込み法を提案する．この際,一枚の画像を特徴付けるキーポイントが満たすべき条件として，前述の(1) (2)の条件が重要である．これらの条件を満足するために，MIL で用いる DD の考え方を導入する．

### 3.1 Diverse Density

MIL では，共通性を見出したい画像群に共通に存在し，共通性があってはならない画像には含まれない特徴を求める問題を，特徴空間内のポテンシャルの極値検出問題として扱う．これは，特徴空間内のある点から見て，遠い場所にある特徴からは弱い影響を受け，近い特徴からは強い影響を受けると考えて，それらの影響を合算して，その点のポテンシャルとする．但し，同一視したい特徴からはポジティブな影響を受け，同一視したくない特徴からはネガティブな影響を受けると考える．このようにポテンシャルを設計すると，特徴空間内でポジティブな共通特徴が集まりネガティブな特徴から離れたところに，ポテンシャルの極大点が現れ，それを検出することで，共通性のある特徴を見つけることが出来る．このポテンシャルは，Diverse Density (DD)と呼ばれる．以降，説明で用いる用語と，DD の定義，について説明する．

**バッグ** ：インスタンスの集合．本稿では画像を指し，記号 $\mathcal{B}$で表す．

**ラベル** ：共通性を見出したいバッグに対してはポジティブ，共通性があってはいけないバッグに対してはネガティブのラベルを与える．それぞれ $\mathcal{B}_i^+,(\mathrm{i}=1,\ldots,\mathrm{m})$，$\mathcal{B}_i^-,(\mathrm{i}=1,\ldots,\mathrm{n})$と表す．

**インスタンス** ：個々の画像特徴ベクトル．$\boldsymbol{B}_{ij}^+ \in \mathcal{B}_i^+$，$\boldsymbol{B}_{ij}^- \in \mathcal{B}_i^-$と表す．

まず，特徴空間内のある点$\boldsymbol{x}$におけるインスタンス$\boldsymbol{B}_{ij}$からの影響を以下のように定義する．

$$P(\boldsymbol{x}=\boldsymbol{t}_j|\mathcal{B}_i) = P(\boldsymbol{x}=\boldsymbol{t}_j|\boldsymbol{t}_j \in \mathcal{B}_i) = \exp\left(-\|\boldsymbol{B}_{ij}-\boldsymbol{x}\|^2\right) \quad (1)$$

これは，近ければ最大 1 の影響を受け，遠ければ影響が小さくなることを表している．

つまり，図 3 のように，画像から抽出された特徴$\boldsymbol{B}_{ij}$を特徴空間内に配置し，特徴空間内で 0 から 1 までの値をとるポテンシャルを発生させる．

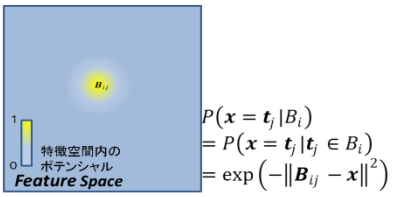


図 3　特徴空間内のポテンシャル

この値をもとにして，あるポジティブバッグ（共通性を見出したいある画像）$\mathcal{B}_i^+$内の全インスタンスから点$\boldsymbol{x}$への影響$P(\boldsymbol{x}|\mathcal{B}_i^+)$を次式のように定義する．

$$P(\boldsymbol{x}|\mathcal{B}_i^+) = 1 - \prod_{\boldsymbol{t}_j \in \mathcal{B}_i^+} (1 - P(\boldsymbol{x}=\boldsymbol{t}_j|\mathcal{B}_i^+)) \quad (2)$$

これは，点$\boldsymbol{x}$と近いインスタンスがバッグ$\mathcal{B}_i^+$内に一つでもあれば，右辺第 2 項の積項が小さくなり，それを 1 から引くという計算である．これは，ポジティブな画像内に$\boldsymbol{x}$と近い特徴が少なくとも一つは含まれている度合いを表している．

つまり，図 4 のように各ポジティブ画像のポテンシャルを発生させる．この時，クエリ$\boldsymbol{x}$が何れかの特徴ベクトルに近ければ，その場所の値は 1 に近くなる．そして，すべてのポジティブ画像のポテンシャルを発生させ，それらを掛け合わせることにより，ポテンシャルを統合する．

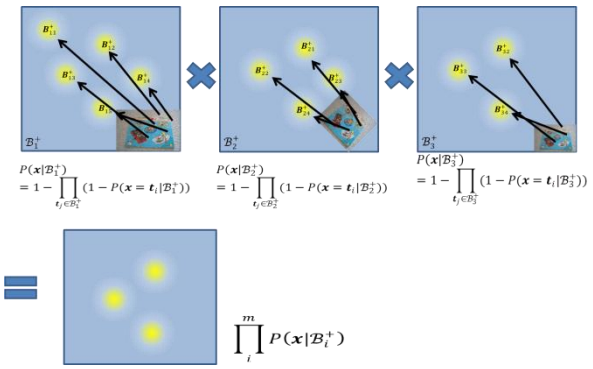


図 4　ポジティブバッグのポテンシャル

逆に，ネガティブバッグ$\mathcal{B}_i^-$に含まれる全インスタンスから点$\boldsymbol{x}$への影響は，次式で定義される．

$$P(\boldsymbol{x}|\mathcal{B}_i^-) = \prod_{\boldsymbol{t}_j \in \mathcal{B}_i^-} (1 - P(\boldsymbol{x}=\boldsymbol{t}_j|\mathcal{B}_i^-)) \quad (3)$$

これは，点$\boldsymbol{x}$と近いインスタンスがバッグ$\mathcal{B}_i^-$内に一つでもあれば小さくなる値であり，ネガティブな画像内に点$\boldsymbol{x}$と近い特徴が含まれない時に大きな値になる．

つまり，図 5 のように各ネガティブ画像のポテンシャルを発生させる．この時，先ほどとは異なり，クエリ$\boldsymbol{x}$が何

れかの特徴ベクトルに近ければ，その場所の値は 0 に近くなる．そして，すべてのネガティブ画像のポテンシャルを発生させ，それらを掛け合わせることにより，ポテンシャルを統合する．

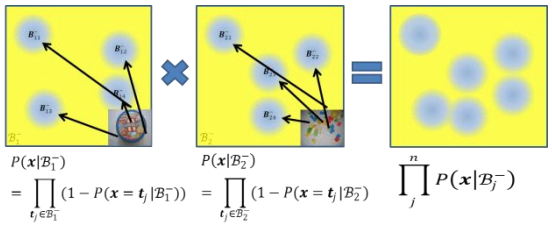


図 5　ネガティブバッグのポテンシャル

点$\boldsymbol{x}$ における Diverse Density $DD(\boldsymbol{x})$は，全ポジティブバッグ，全ネガティブバッグに関して式(3)(4)を掛けあわせた量であり，次式のように定義される．

$$DD(\boldsymbol{x}) = \prod_i^m P(\boldsymbol{x}|\mathcal{B}_i^+) \prod_j^n P(\boldsymbol{x}|\mathcal{B}_j^-) \qquad (4)$$

つまり，図 6 のように，全ポジティブバッグのポテンシャルと全ネガティブバッグのポテンシャルを掛け合わせることにより， $DD(\boldsymbol{x})$を得ることができる．

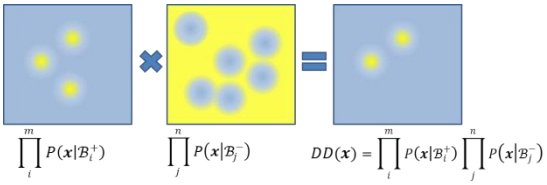


図 6　$DD(\boldsymbol{x})$のポテンシャル

$DD(\boldsymbol{x})$の値が大きければ，その点は各ポジティブバッグに類似した特徴を含み，どのネガティブバッグにも類似した特徴が含まれないということを表している．

### 3.2 局所特徴量抽出

DD を計算する前処理として，まずは，データベースに登録するすべての画像,およびそれらに様々な変換を加えた画像から局所特徴量を抽出する．

本提案手法で用いる局所特徴量のアルゴリズムは吉岡ら[12]の SURF をさらに高速化した整数化 SURF を用いる．そして，整数化 SURF により抽出した 64 次元の特徴ベクトルを DD の計算に用いる各インスタンスとする．

#### 3.2.1 SURF(Speeded Up Robust Features)

SURF とは，画像から撮影条件の影響を受けにくい局所特徴量を高速に抽出するアルゴリズムである．SURF は，キーポイントを位置，向き，スケールなどから検出し，それらパラメータを用いて，正規化された高次元特徴量の抽出を行う．抽出される特徴量は，局所的な勾配分布であるが，キーポイントの位置，向き，スケールを用いて正規化されているため，これらの変化に対して不変な特徴となっている．また，画像の勾配分布であることから，照明の一様な変化の影響も受けにくい．このように，幾何学的変化や照明などの環境変化の影響を受けにくいため，物体認識や類似画像検索，画像の貼り合わせなど非常に多くのアプリケーションで用いられている．

### 3.3 DD の計算

ここで，簡単な例として，図 9 のような数字が同じで，マークが異なる 4 枚のトランプ画像について考えるとする．ここで，ダイヤのエースを画像 A，その他のハート，スペード，クローバーをそれぞれ画像 B，C，D とした時，画像 A の各インスタンスの$DD(\boldsymbol{x})$を求める場合，前述したように，画像 A，および，画像 A に様々な変換を加えた画像をポジティブバッグとし(図 7), 他の画像 B,C,D をネガティブバッグとする(図 8).

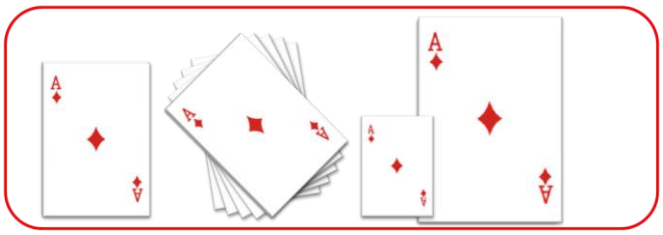

図 7．ポジティブバッグ

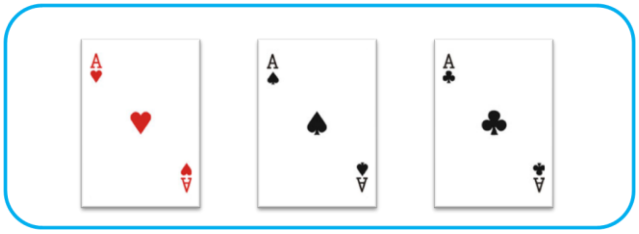

図 8．ネガティブバッグ

そして，このポジティブバッグ，ネガティブバッグを用いて画像 A のすべてのインスタンスに対して，$DD(\boldsymbol{x})$を計算する．ここで，各ポジティブバッグに類似した特徴を含み，全ネガティブバッグに類似した特徴が含まれない点が$DD(\boldsymbol{x})$の値が高いという性質から，$DD(\boldsymbol{x})$の高い点を抽出することで，前述の 2 つの条件を満たすキーポイントの絞り込みを実現する．

### 3.4 DD の最大値

図 9 に 3.4 項で用いた各画像における$DD(\boldsymbol{x})$の最大値を示す．図 9 を見るとわかるように，各画像の特有の領域であるマークの部分で$DD(\boldsymbol{x})$が最大となっていることがわかる．

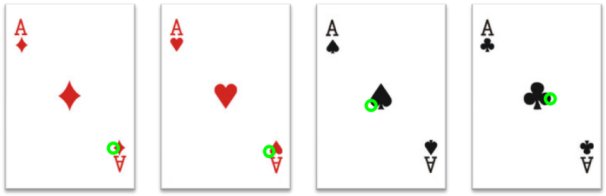

図 9．$DD(\boldsymbol{x})$の最大値

一方，図 10 のようなマークが同じで数字が異なる 4 枚のトランプ画像について考えるとする．このとき，各画像の特有の領域は先ほどとは異なり，数字の部分になる．ここで，先ほどと同様の手順で$DD(\boldsymbol{x})$を計算し，各画像における$DD(\boldsymbol{x})$の最大値を図 10 に示す．図 10 を見ると，確かに数字の部分で$DD(\boldsymbol{x})$が最大となっていることがわかる．

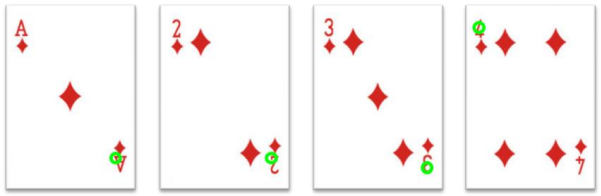

図 10．$DD(\boldsymbol{x})$の最大値

また，これらの点は，前述したようにポジティブバッグに様々な変換を加えた画像を登録しているため，回転やスケール変化などの変換にもロバストな点となっている．

### 3.5 データベース作成

キーポイントの絞り込みをデータベースに登録する全画像に対して行い，抽出した少数の特徴を画像検索用のデータベースとする．こうすることで，画像のインデックスとして用いる特徴の個数を削減することができる．

## 4. 実験

本提案手法の有効性を確認するため，データベース中の局所特徴量と入力画像から検出した局所特徴量の対応付けの実験を行い，その精度を確認する．

ここでは，提案手法により一枚の画像を表す局所特徴量の数を削減したデータベース(DBdd)と各画像から抽出された局所特徴量をそのまま用いたデータベース（DBoriginal）を用いて，比較実験を行う．

### 4.1 実験方法

作成したデータベースの局所特徴量とクエリ画像から抽出された局所特徴量との対応付けは，最近傍探索のライブラリである FLANN(Fast Library for Approximate Nearest Neighbor)[13]を用いる．

FLANN とは，k-近傍探索の高速な近似計算アルゴリズムのコレクションで，巨大なデータセットと高次元特徴量に対して最適化したものである．

また，ここで用いられる ANN とは，k-d 木を用いた木探索による暫定解の決定と，バックトラックしながら精密な探索を行って，k 近傍点を得るアルゴリズムである．

本実験では，データベースの局所特徴量群により，木構造を構築した後，クエリ画像から抽出された全局所特徴量をクエリ点として与え，クエリ画像上の各キーポイントの最近傍点を求め，最近傍点までの距離に対して直接，閾値処理をし，閾値より近ければ，対応付けを行う．ここでの，距離とは，ユークリッド距離の平方根をとっていないものを用いている．また，対応付けの閾値は 8000 に設定している．

つまり，最近傍点までのユークリッド距離が 8000 より近ければ，対応付けるということである．

そして，データベースに登録してある各画像の中で，マッチした点が最も多かった画像を認識画像とする．

### 4.2 実験条件

**データベース作成に用いた画像(図 11)**

Nister の画像データベースから 300 枚．

図 11　データベース作成に用いた画像の一例

**クエリとして用いた画像(図 12)**

データベース作成に用いた画像と同種類で見え方の違う画像 300 枚の 3 セットの計 900 枚の画像．

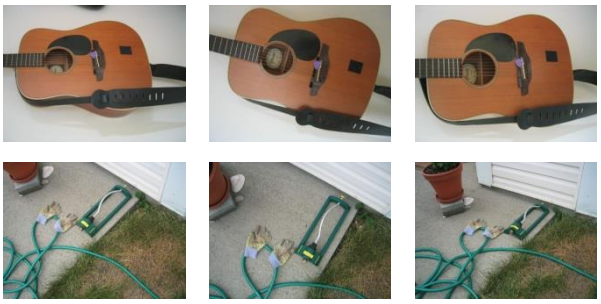

図 12　クエリとした用いた画像の一例

**閾値**

SURF の閾値：300

マッチングの閾値：8000

**DBdd**

一枚の画像を表す局所特徴量の最大数：

100~1500 点（最大数に達しない場合は，抽出されたキーポイントすべてを用いる．）

ポジティブバッグ：

元画像，{9,18,27,36,45}度回転した画像，スケールを 0.75 倍に縮小した画像，スケールを 1.5 倍に拡大した画像の計 8 枚の画像

ネガティブバッグ：

データベース作成に用いる注目画像以外の画像 299 枚

**DBoriginal**

一枚の画像に対して登録する局所特徴量の数：抽出された局所特徴量の数．

### 4.3 実験環境

**CPU**：Intel Core i7 3.50GHz

**メモリ**：16GB

**OS**：Windows 7

### 4.4 実験結果，評価

上記のような実験条件の下，対応付けの実験を行った時の認識率の評価を行い，その結果を図 13 に示す．図 13 は横軸が一枚の画像を表す局所特徴量の最大数を表し，縦軸が認識率を表している．ここで，認識率とは，正しく入力画像を認識した割合である．

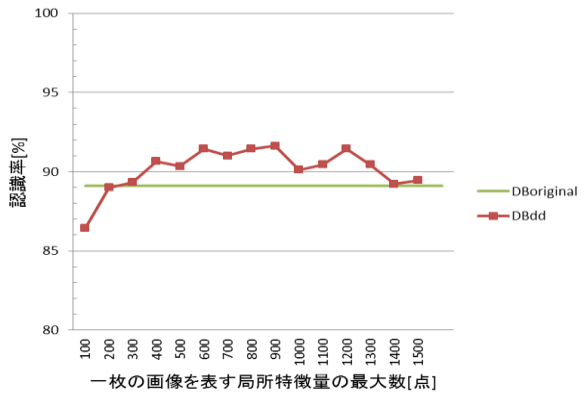


図 13　局所特徴量数に対する認識率の評価

図 13 の DBoriginal は画像から抽出されたキーポイントをすべて利用したもので，横軸に関係なく基準としており，約 89%の認識率である．一方，DBdd は画像から抽出されたキーポイントをすべて利用せず，提案手法によりキーポイントに重要度をつけ，用いるキーポイント数を削減している．図 13 より，DBoriginal よりも多くの範囲で認識率は良い結果となっている．このことから，本提案手法の有効性が確認できる．

また，図 14 に速度の評価を示す．横軸が先ほどと同様に一枚の画像を表す局所特徴量の最大数を表しており，縦軸が検索時間を表している．ここで，検索時間とは，クエリ画像から整数化 SURF で局所特徴量を抽出する時間，クエリ画像から抽出された全キーポイントの第一近傍点を決定する時間，それらによって対応付けを行う時間の合計の時間である．

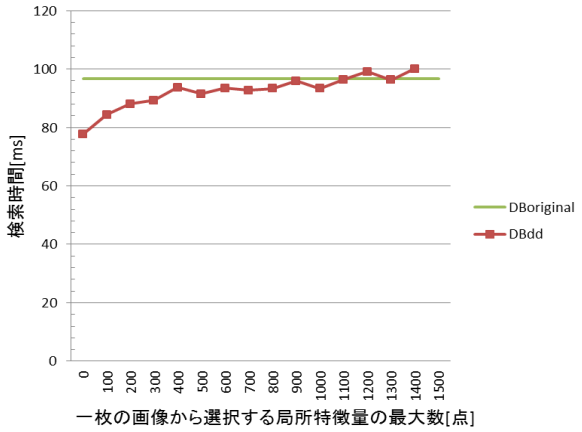


図 14　局所特徴量数に対する検索時間の評価

図中の DBoriginal, DBdd は図 13 と同様である．DBoriginal は基準として，96.73ms を表している．一方，DBdd は DBoriginal よりも大半の範囲では検索時間は早くなっている．これは，参照する局所特徴量の数を減らしていることにより起こる．つまり，画像枚数を増やしていけばいくほど，この差は開いていくと考えることが出来る．

本実験においては，図 13，14 の結果から一枚の画像を表す局所特徴量の数は 400～500 点ぐらいが適切であると考えられる．

また，実際の DBoriginal，DBdd における検索結果を図 15 に示す．ここで，DBdd の局所特徴量の最大数は 600 として，実行した．

図 15 の DBoriginal の上から 2 枚目と 4 枚目の誤対応している．そして，それぞれの画像の対応の様子を見るからに，画像内の雑多な部分，ここでは，草むらの部分が対応してしまい誤認識を引き起こしていると考えることが出来る．一方，DBdd の方を見てみると，草むらの部分からも対応してはいるが，かなり制限されており，その他の画像固有の物体と多くマッチしていることがわかる．

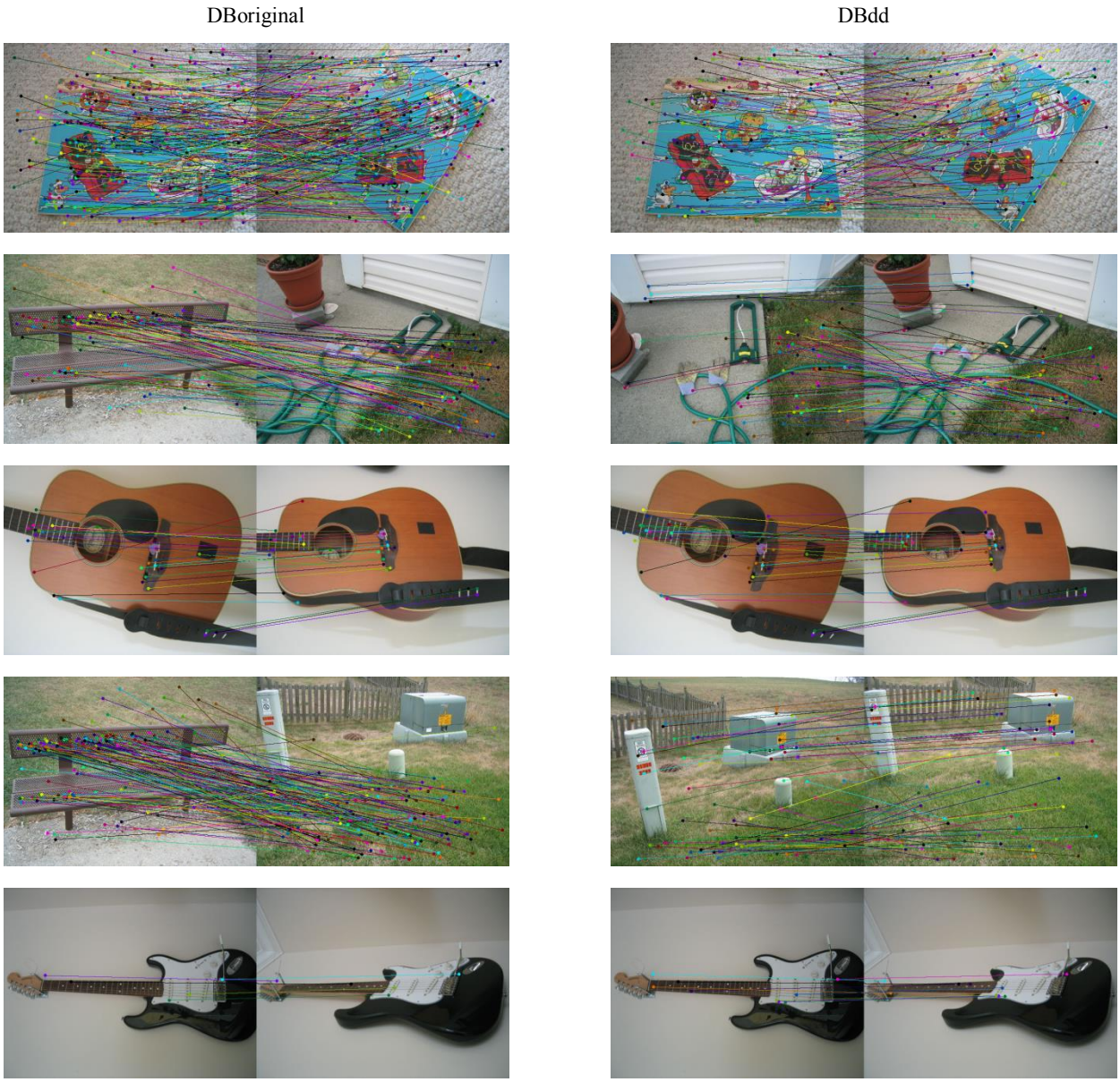


図 13　左）DBoriginal の検索結果の一例　右）DBdd の検索結果の一例

## 5. おわりに

### 5.1 まとめ

本研究では, Multiple-Instance Learning で用いられる Diverse Density を用いたキーポイントの重要度計算とそれによる画像検索に用いるキーポイントの絞り込み法を提案した．また，データベースと入力画像との対応付けの実験では，抽出されたキーポイントをすべて用いた方法よりも，提案手法で削減した方法のほうが大部分で安定な検索が行えることを確認した．また，それによる検索速度の向上も実現した．

これにより，提案手法を用いることによって，回転やスケール変化などの変換に強く，かつ，他の画像との弁別性が高いキーポイントを抽出できていることが分かる．

### 5.2 今後の予定

さらに画像枚数を増やすことでよりユニークなキーポイントを，また，ポジティブバッグとして加える劣化の種類を増やすことでよりロバストなキーポイントを取り出すことが出来るのではないかと考えれる．

また，提案手法を用いて作成したデータベースによる画像検索システムを FPGA 上に構築する．

## 参考文献

1) D.G. Lowe, "Distinctive image features from scale-invariant keypoints," IJCV, Vol. 60, No. 2, pp. 91-110, 2004

2) H. Bay, T. Tiytelaars, and L. J. Van Gool, “SURF: Speeded Up Robust Features,” In ECCV, pp. 404-417, 2006.

3) J. Sivic and A. Zisserman, “Video Google: A text retrieval approach to object matching in videos,” In Proc. of ICCV, Vol.2, pp. 1470-1477, Oct 2003.

4) L. Fei-Fei and P. Perona, “A Bayesian hierarchical model for learning natural scene categories,” In Proc. of CVPR, Vol. 2, pp. 524-531, June 2005.

5) D. Nister and H. Stewenius, “Scalable Recognition with a Vocabulary Tree,” In Proc. of CVPR, June 2006, Vol.2, pp. 2161-2168, June 2006

6) 黄瀬, 岩村, 中居, 野口, “局所特徴量のハッシングによる大規模画像検索,” 日本データベース学会論文誌, Vol. 8, No. 1, pp. 119-124, June, 2009

7) T. G. Dietterich, R. H. Lathrop and T. Lozano-Perez, “Solving the multiple-instance problem with axis-parallel rectangles”, Artificial Intelligence, vol.89, no.1-2, pp31-71, January 1997.

8) O. Maron and T. Lozano-Perez, “*A Framework for Multiple-Instance Learning*”, Advances in Neural Information Processing Systems 1

9) O. Maron and A. Ratan: “Multiple-Instance Learning for Natural Scene Classification”, Proceedings 15th International Conference on Machine Learning, pp341-349, Madison, Wisconsin, USA, July 1998.

10) Q. Zhang, S. A. Goldman: “EM-DD: An Improved Multiple-Instance Learning Technique”, Advances in Neural Information Processing System 14, pp1073-1080, Vancouver, British Columbia, Canada, December 2001.

11) H. Nakamura, T. Harada, Y. Kuniyoshi, “Dense sampling low-level statistics of local features,” In Proc. of CIVR’09, Article No. 17, 2009.

12) 吉岡勇太, 和田俊和,”FPGA 上での画像キーポイント検出と対応付けの並列実装”, MIRU2012 画像の認識・理解シンポジウム, no.IS3-17, 2012

13) Marius Muja and David G. Lowe, **“Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration,”** in International Conference on Computer Vision Theory and Applications (VISAPP'09), 2009.