

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

電気通信大学 情報理工学系研究科 総合情報学専攻
秋山 瑞樹
柳井 啓司

背景

- ◆ 「物体認識」

- ◆ カテゴリー認識 - 物体・シーン認識、顔検出

一般物体認識

- ◆ 同一物体認識 - 登録物体の検索、人物名認識

特定物体認識

背景

- ◆ 特定物体認識

- ◆ 大量の学習画像 + 特徴点マッチング
- ◆ 剛体は高い精度で認識可能

- ◆ 一般物体認識

- ◆ 少数の学習画像 + 機械学習
- ◆ カテゴリー内分散が大きく認識が難しい

- ◆ 特定の物体を大量に集めることで一般物体認識は可能か？

目的

- ◆ 一般物体認識を実験
 - ◆ 特定物体認識手法による特徴点マッチングベース
 - ◆ 約15万枚の学習画像を用いて分類実験
 - ◆ データ量で一般物体認識を実現

The chair

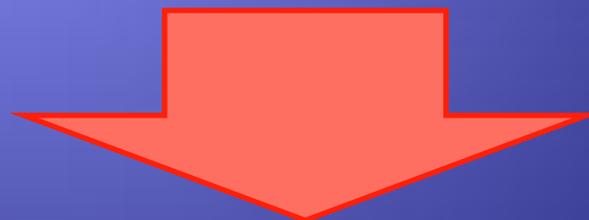


A chair



関連研究

- ◆ 大量画像を用いた物体認識
 - ◆ 20億枚の画像を用いた画像への自動タグ付け [Xin-Jing et al CVPR(2010)]
 - ◆ ほとんど同一な画像を探索する。



- ◆ ランキングに含まれる、マッチする特徴を持つ画像を用いる

方法概要

画像収集
+
特徴抽出

- Webから約15万枚収集
- 局所特徴抽出

データベース
化

- 学習画像特徴をデータベース化

特徴点
マッチング

- 未知画像特徴に対してデータベースの特徴とマッチング
- データベースの特徴を持つ学習画像に投票

認識

- 投票数による未知画像認識

方法詳細

◆ 画像収集

- ◆ 独自に決めた25クラスに関して計15万枚収集
- ◆ Google, Yahoo!, Flickr
- ◆ 単純に収集するのでノイズを含む



方法詳細

◆ 特徴抽出

◆ SIFT

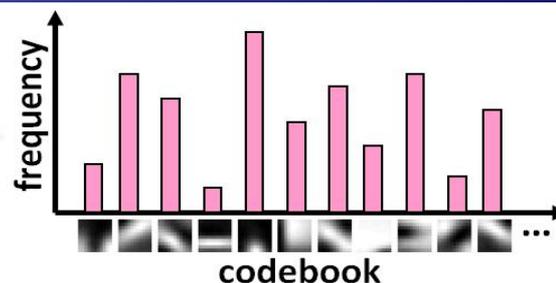
- ◆ 128次元,画像変化に強い特徴

◆ PCA-SIFT

- ◆ SIFTを主成分分析により36次元に圧縮

◆ Bag-of-Features

- ◆ 画像を特徴の出現頻度で表現したもの
- ◆ コードブックサイズを変化



方法詳細

- ◆ データベース化
 - ◆ 学習画像特徴IDと学習画像の対応をとる
 - ◆ 学習画像特徴をデータベース化
 - ◆ 特徴から画像名を引けるようにする

```
1228 animal/cat/cat00000.jpg  
1881 animal/cat/cat00001.jpg  
2604 animal/cat/cat00002.jpg  
3502 animal/cat/cat00003.jpg  
3878 animal/cat/cat00004.jpg  
5005 animal/cat/cat00005.jpg  
5783 animal/cat/cat00006.jpg
```

方法詳細

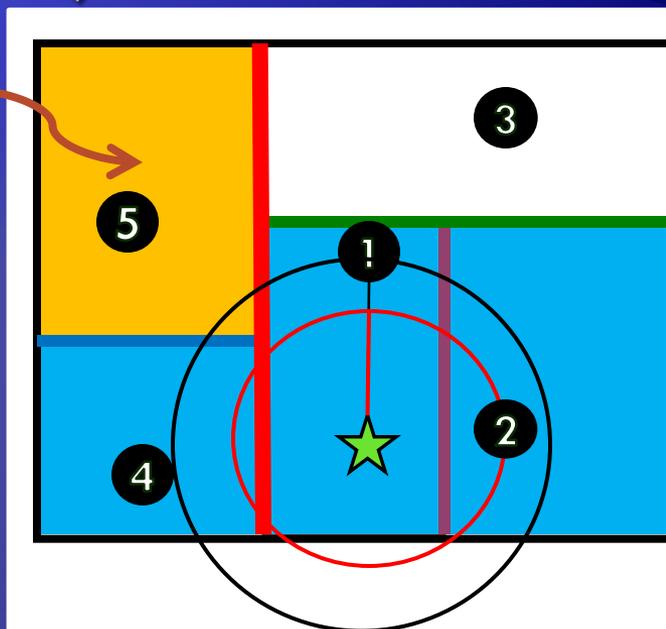
◆ 特徴マッチング

- ◆ 単純な線形探索では探索コストが大きすぎる。
- ◆ Approximate Nearest Neighbor(ANN)
 - ◆ Kd-tree(木構造)を用いた近似最近傍探索手法
 - ◆ 未知画像の特徴に対して近傍特徴を n 位まで探索する
 - ◆ データベースを参照し、近傍特徴を持つ学習画像に投票

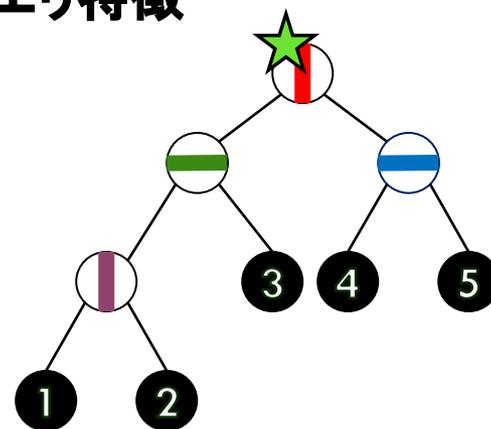
近似最近傍探索(ANN)

1. 学習画像特徴を読み込み
2. セルを作成
3. kd-Tree作成
4. クエリ特徴に対してセルを選択
5. セルの特徴との距離を半径とした円を考え、円に含まれるセルを選択
6. 選択されたセルの特徴とクエリ特徴に関して距離を計算
 - ◆ 最近傍探索：オレンジ
 - ◆ 近似最近傍探索：ブルー
7. 近傍特徴 n 位までを投票

セル



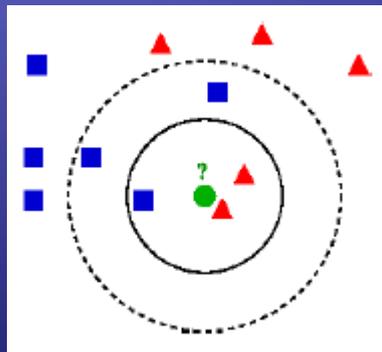
★クエリ特徴



方法詳細

◆ 認識

- ◆ 特徴マッチングによる学習画像の投票数ランキングをえる
- ◆ ランキングに関してクラスの多数決によって最も多い学習画像のクラスを未知画像のクラスとする。
- ◆ k-Nearest Neighbor
 - ◆ ランキング上位 k 位までに対して多数決をとる



Bag-of-Features手法

- ◆ BoF表現は非常にスパース
 - ◆ 非ゼロ要素のみを記録
 - ◆ 転置インデックス
 - ◆ 同じコードブックを持つ画像リスト
- ◆ コードブックベクトルをANNの学習特徴とする
- ◆ 近傍特徴として選ばれたコードブックベクトルに対して転置インデックスを用いて投票

データセット

◆ データセット

- 動物
- 車
- 花
- 食べ物
- 楽器

動物



車



花



食べ物



楽器



バラ
プリウス
バラ
ピザ
オリン

実験

- ◆ 5、25クラス分類
 - ◆ テスト画像1,250枚で分類実験
- ◆ SIFT,PCA-SIFT,BoFの3手法で実験
 - ◆ ANNの近傍特徴数 n 位と投票数ランキングk-NN k 位の変化による認識精度
 - ◆ コードブックサイズ変化による認識精度
 - ◆ 学習画像数変化による認識精度

実験

◆ 学習画像数

- ◆ メモリ32GBのマシン

	画像数	クラスあたりの画像数	特徴数	メモリ使用量
SIFT	26,250	1,050	1,500万	20GB
PCA-SIFT	73,500	2,940	5,350万	25GB
BoF	145,000	5,800	-	5GB

◆ 評価

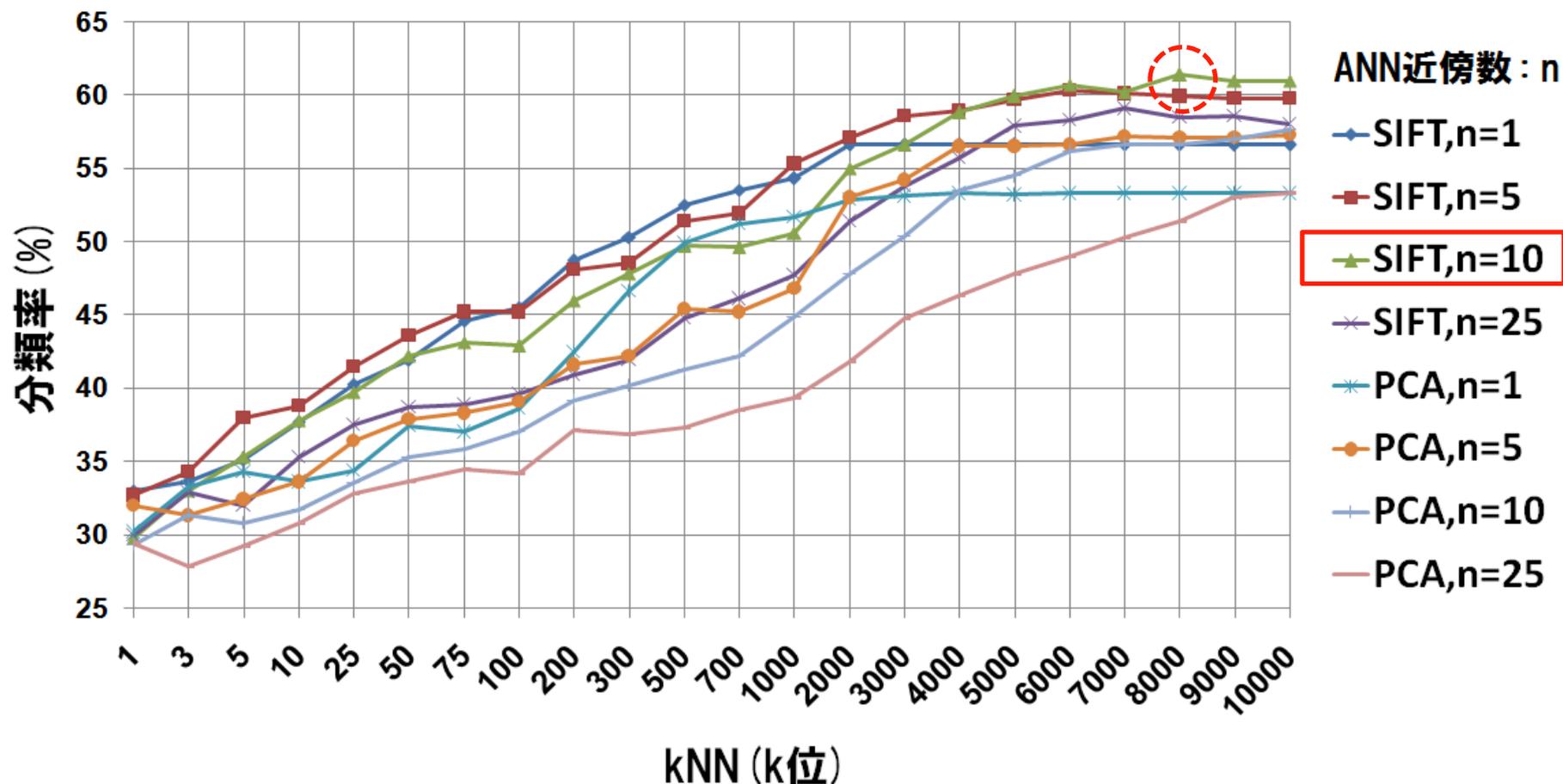
- ◆ 再現率、適合率、分類率で評価
- ◆ ベースラインとしてBag-of-Features + サポートベクターマシンを使用

実験結果

	SIFT 再現率(%) (提案手法)	BoF+SVM 再現率(%) (ベースライン)	SIFT 適合率(%) (提案手法)	BoF+SVM 適合率(%) (ベースライン)
ピザ	40	64	33	48
ラーメン	18	60	33	35
ドラム	0	22	0	42
トラ	78	70	63	53
ピアノ	70	46	60	46
ギター	58	34	76	35

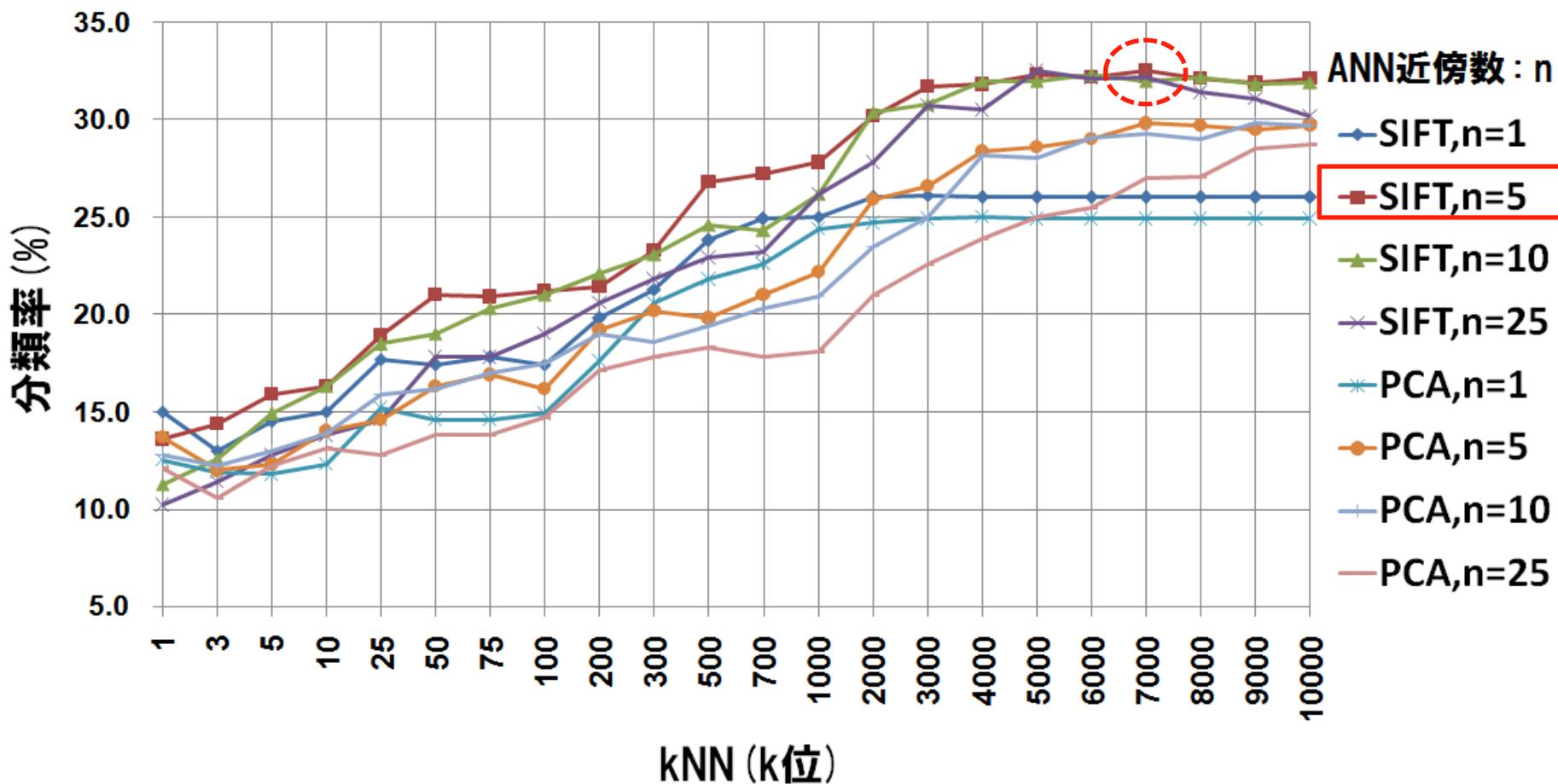
実験結果(SIFT,PCA,5クラス分類)

上位5クラス分類 (SIFT,PCA-SIFT)



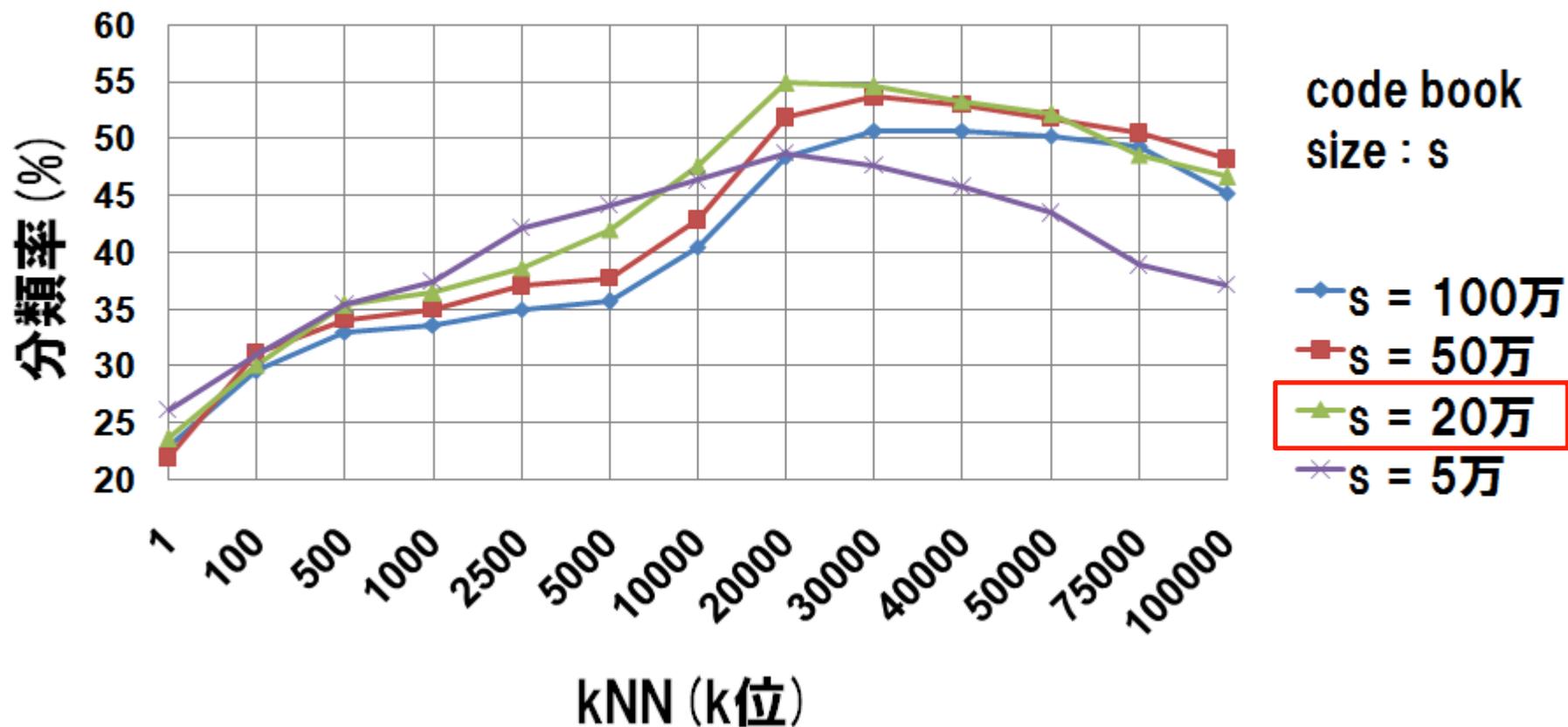
実験結果(SIFT,PCA,25クラス分類)

下位25クラス分類 (SIFT,PCA-SIFT)



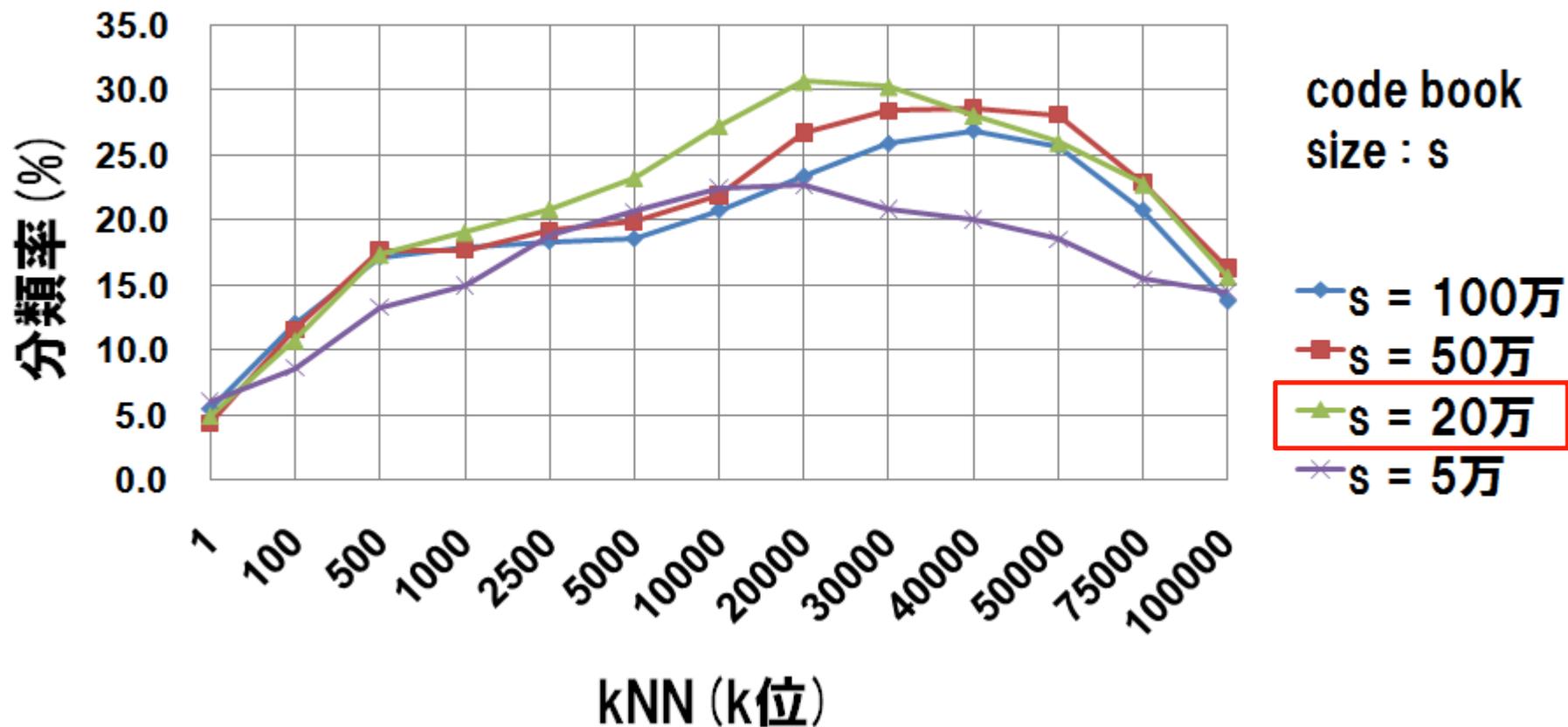
実験結果(BoF,5クラス分類)

上位5クラス分類 (BoF)



実験結果(BoF,25クラス分類)

下位25クラス分類 (BoF)

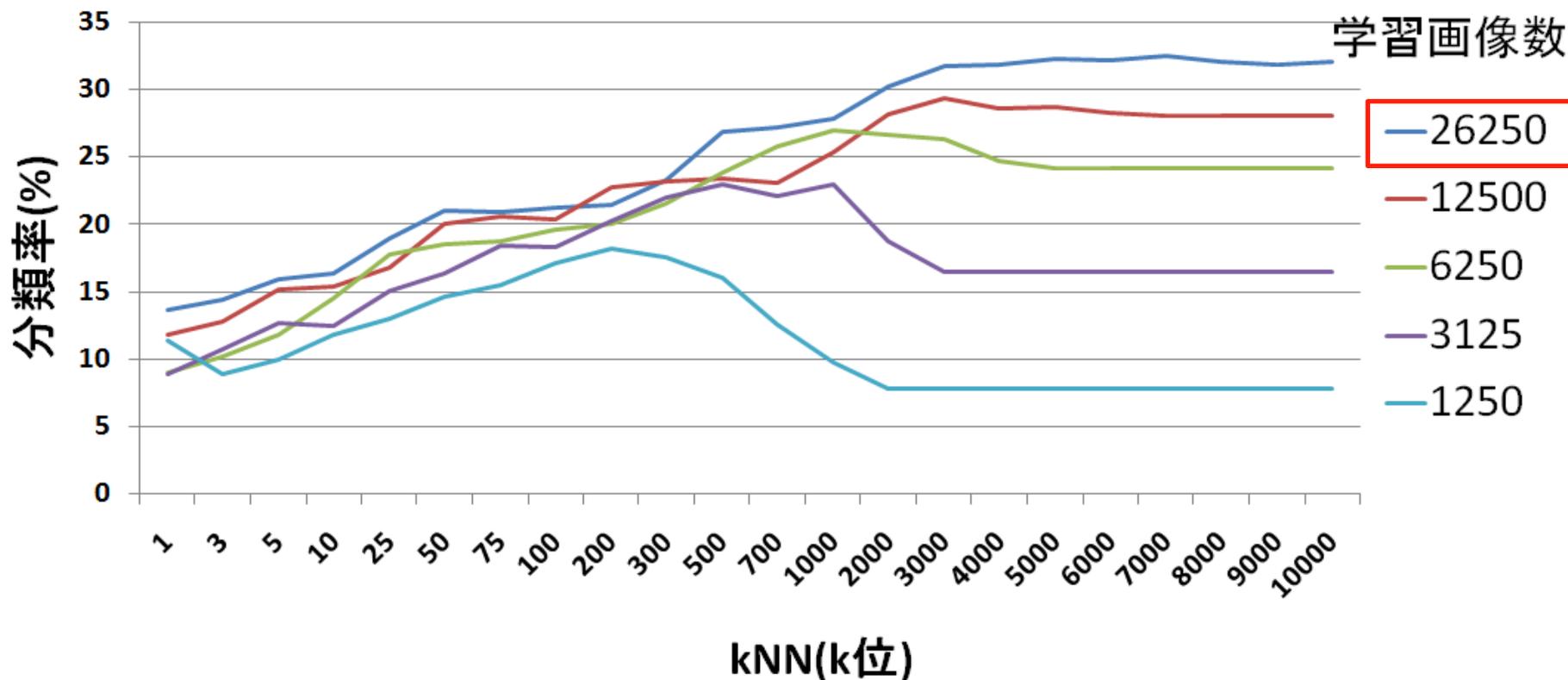


実験結果(SIFT,5クラス分類)

	1	2	3	4	5	再現率(%)
1: 動物	155	14	37	37	7	62
2: 車	10	228	3	4	5	91
3: 花	35	15	150	41	9	60
4: 食べ物	43	24	40	135	8	54
5: 楽器	12	128	10	14	86	34
適合率(%)	61	56	63	58	75	60.3

実験結果(学習画像数による変化)

下位25クラス分類(SIFT,ANN n=5)



まとめ

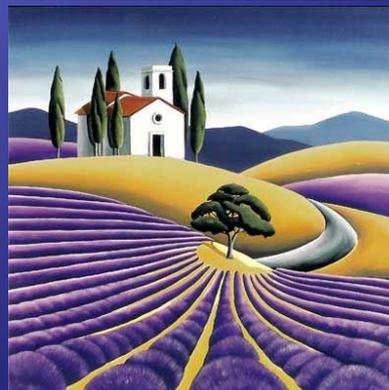
- ◆ 特徴点マッチングに基づく特定物体認識手法で一般物体認識を実験
- ◆ 多数の画像があればある程度分類可能
 - ◆ ベースラインの結果には少し及ばなかった
 - ◆ SIFT, $n=5, k=7000$
 - ◆ 5クラス分類率 60.3%, 25クラス分類率 32.5%
 - ◆ トラ, ピアノ, ギターなどの認識には有効



今後の課題

- ◆ 画像数を増やすことで精度が向上した
 - ◆ 複数計算機のメモリを使用
- ◆ 価値のある特徴の探索
 - ◆ 対象物体以外の特徴が他クラスの主な特徴とマッチ
 - ◆ 動物画像の「草木」と花画像の「草木」がマッチ
- ◆ ノイズによる影響
 - ◆ 学習画像のノイズの影響はどのくらいあるのか

学習画像例

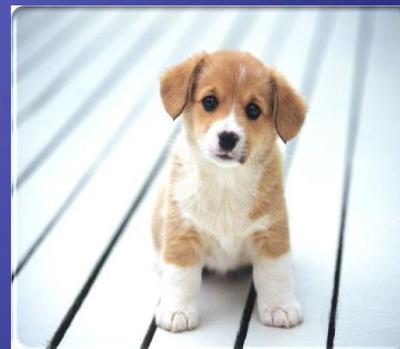


今後の課題

- ◆ 画像数を増やすことで精度が向上した
 - ◆ 複数計算機のメモリを使用
- ◆ 価値のある特徴の探索
 - ◆ 対象物体以外の特徴が他クラスの主な特徴とマッチ
 - ◆ 動物画像の「草木」と花画像の「草木」がマッチ
- ◆ ノイズによる影響
 - ◆ 学習画像のノイズの影響はどのぐらいあるのか

ご清聴ありがとうございました

學習画像例



テスト画像例



テスト画像例

