

特徴点選択とペア化による

Naive-Bayes Nearest-Neighbor手法 の改良

電気通信大学情報理工学専攻

秋山 瑞樹

柳井 啓司

背景

- **BoF(Bag of Features)**

- 量子化を行う
- SVMやBoostingなどの機械学習により量子化誤差を補う
- 近年は、量子化誤差を減らす手法も提案されている



クマ

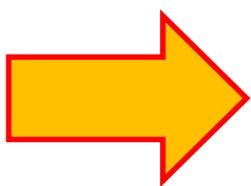
- **特徴点マッチング**

- 量子化を行わない
- 特定物体認識でのみ使われてきた
 - カテゴリー分布の広がりに対応できないため

Naive-Bayes Nearest-Neighbor(NBNN)

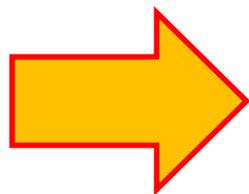
[Oren Boiman et al. 2008]

特徴点マッチングでありながら、カテゴリー認識が可能



目的

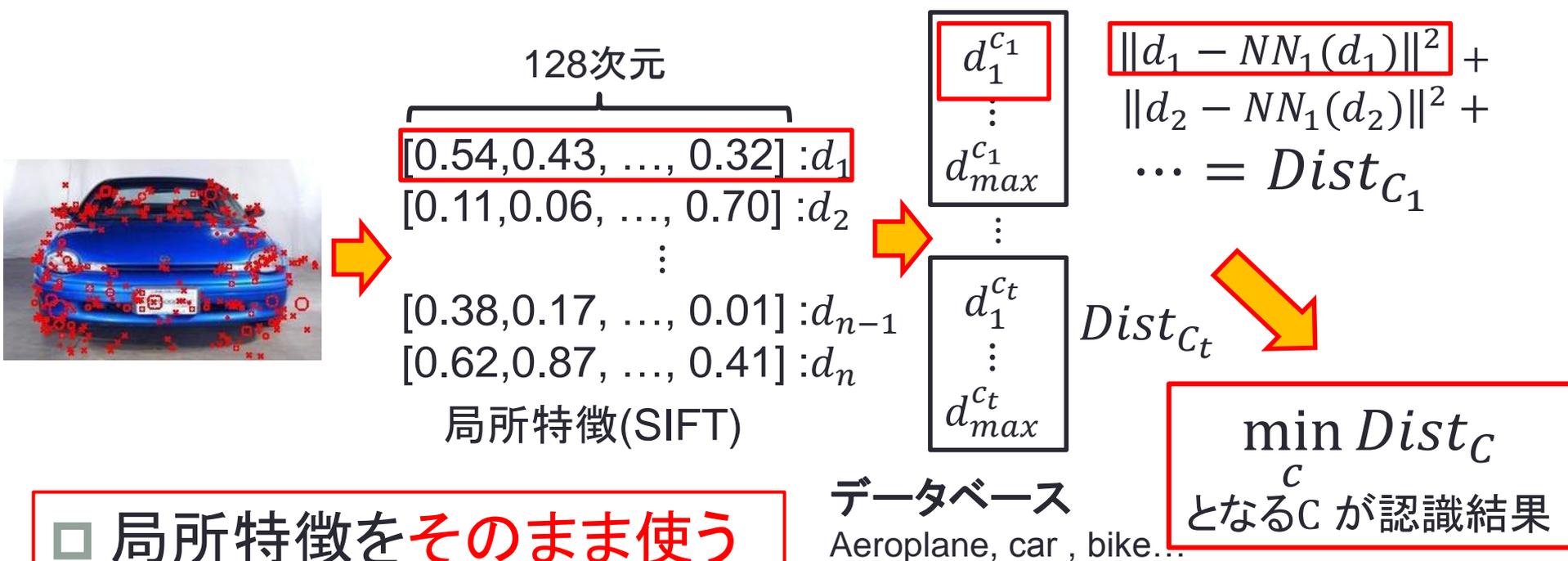
- **特徴点マッチングベース手法**の認識精度向上を目指す
 - NBNNの改良を行う



1. SaliencyMapによる特徴の重み付け
2. mi-SVMによる特徴の重み付け
3. 特徴のペア化

関連研究: NBNN

- Naive-Bayes Nearest-Neighbor(NBNN) [Oren Boiman et al. 2008]



□ 局所特徴をそのまま使う

$$\hat{c} = \arg \min_c \sum_{i=1}^n \|d_i - NN_c(d_i)\|^2$$

\hat{c} : テスト画像の推定されたクラス
 n : テスト画像の特徴数
 d_i : テスト画像の特徴量
 $NN_c(d_i)$: クラス c に属する d_i の最近傍特徴量

関連研究: NBNN

□ ノンパラメトリックな“Image to Class”の認識

Image to Image(NN法)



最短距離の画像

Image to Class(NBNN)

Image to Class距離



Aeroplane

最短距離のクラス



car

car

認識結果



motorbike

データベース
局所特徴の集合

関連研究

- Towards Optimal Naïve Bayes Nearest Neighbor [Regis Behmo et al. CVPR 2010]
 - NBNNにパラメータを導入し、各クラス毎の異なる分布を正規化する距離計算を行えるようなパラメータを学習する。

$$\hat{C} = \arg \min_c \alpha^c \left(\sum_{i=1}^n \|d_i - NN_c(d_i)\|^2 + \beta^c \right)$$

関連研究

- Image-To-Class Distance Metric Learning for Image Classification. [Zhengxiang Wang et al. CVPR 2010]
 - NBNNの距離計算をマハラノビス距離で行うことで精度向上

$$\hat{C} = \arg \min_c \sum_{i=1}^n \|d_i - NN_c(d_i)\|^T M_c \|d_i - NN_c(d_i)\|$$

- Multi-label learning by image-to-class distance for scene classification and image annotation. [Zhengxiang Wang et al. ACM 2010]
 - データベース特徴にマージン最大化によって重みをつける

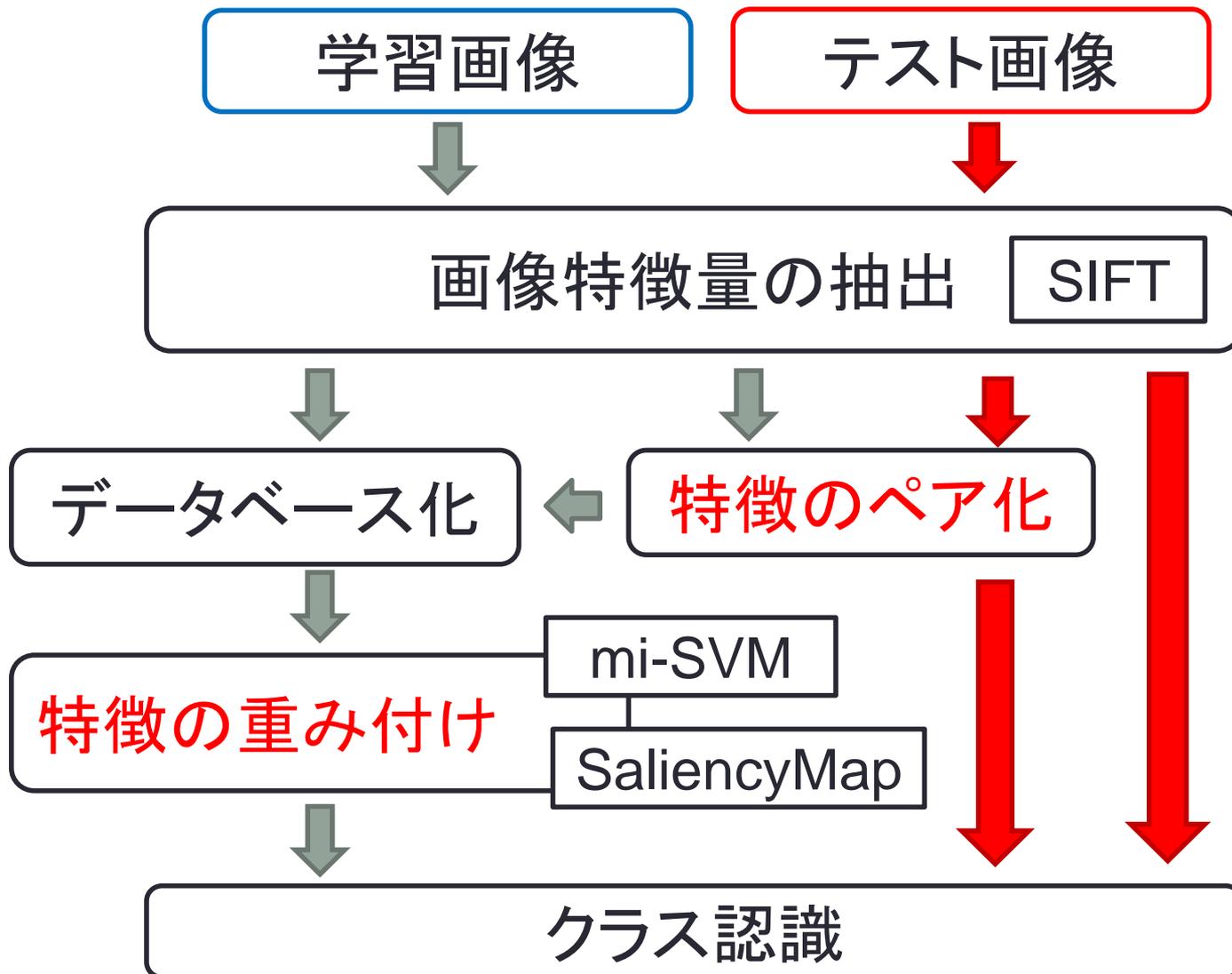
$$\hat{C} = \arg \min_c \sum_{i=1}^n \|W_{NN_c(d_i)} \cdot d_i - NN_c(d_i)\|^2$$

関連研究

- The nbnn kernel.[T.Tuytelaars et al. ICCV 2011]
 - NBNNをカーネル化することで、SVMを用いる学習ベース手法と組み合わせることを可能にした。
- 本手法の位置付け

	パラメータ 学習	距離空間 の変化	重み付け	カーネル化	ペア化
Boiman (CVPR2008)	×	×	×	×	×
Behmo (ICCV2010)	○	×	×	×	×
Wang(ECCV2010)	○	○	×	×	×
Wang(ACM2010)	○	×	○	×	×
Tuytelaars(ICCV2011)	×	×	×	○	×
提案手法	○	×	○	×	○

提案手法概要

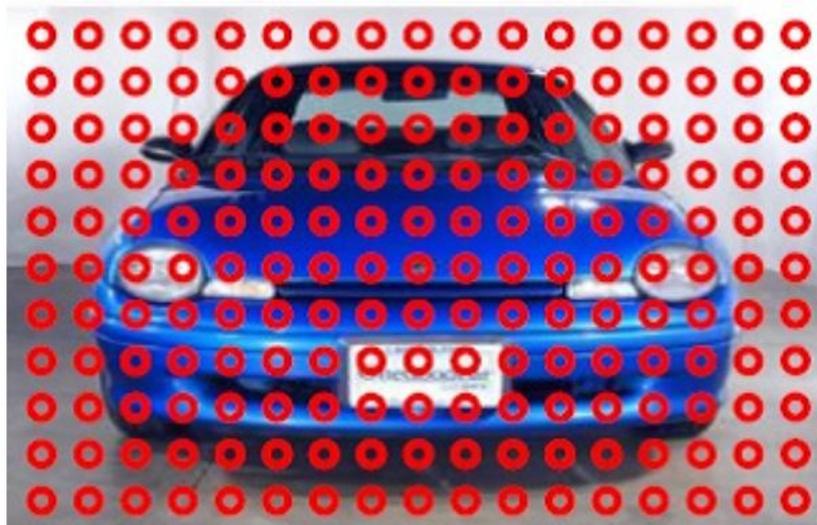


提案手法詳細

- 画像特徴量の抽出

- SIFT特徴量.

- 128 次元特徴ベクトルで表現し、回転・スケール・変化・照明変化に対して耐性のある特徴である

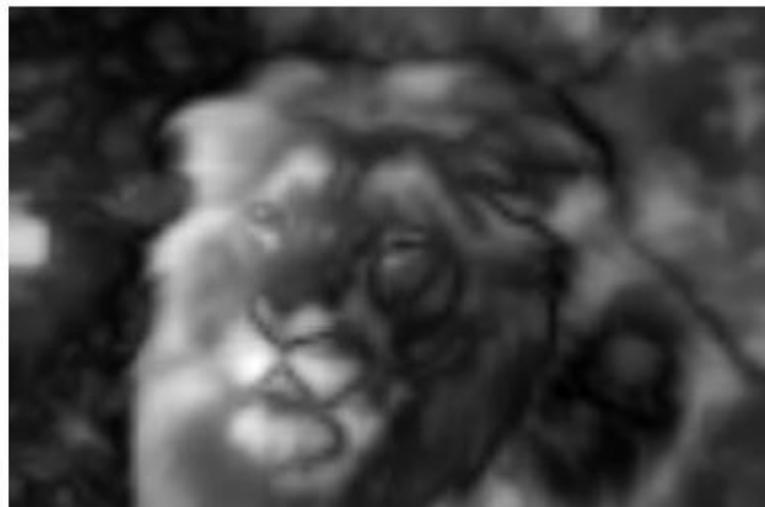


提案手法詳細

- 特徴の重み付け

1. SaliencyMapによる重み付け.

画像の顕著性を考慮することで、物体から得られた特徴に対する重要度を上げる。各ピクセルごとに評価値が得られる

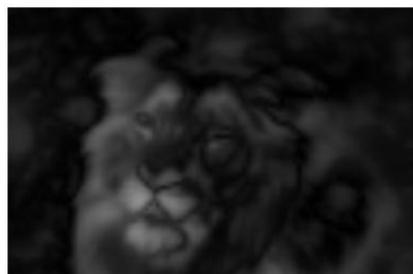


Saliency Map

1. ガウシアンピラミッドによる9スケール画像
2. 輝度、色相、方向成分の抽出
3. 各成分ごとにスケール間差分の計算



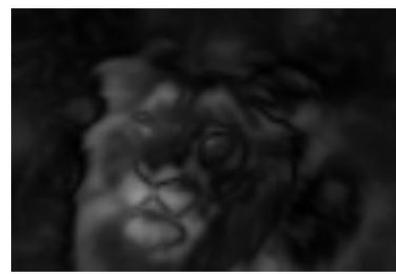
スケール画像



輝度成分



色相成分



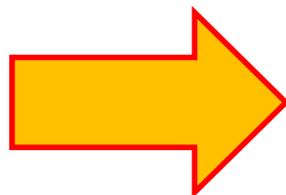
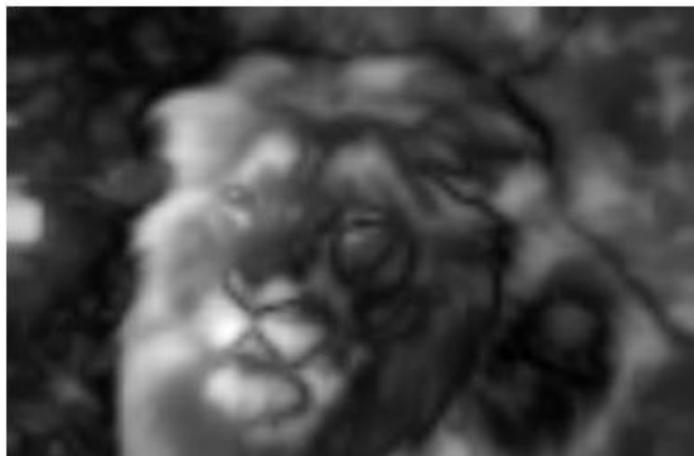
方向成分

4. 各成分の統合



Saliency Map

Saliency Map

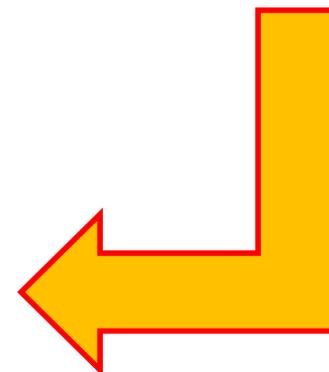


SaliencyMap 評価値

[0.04, 0.83, ..., 0.03, 0.01]
[0.11, 0.96, ..., 0.15, 0.10]
⋮
[0.18, 0.37, ..., 0.22, 0.24]
[0.22, 0.17, ..., 0.31, 0.41]

データベース特徴

SIFT特徴量	SaliencyMap重み
[0.54, 0.43, ..., 0.32]	0.23
[0.11, 0.16, ..., 0.12]	0.67
⋮	
[0.38, 0.17, ..., 0.01]	0.89
[0.62, 0.87, ..., 0.41]	0.78



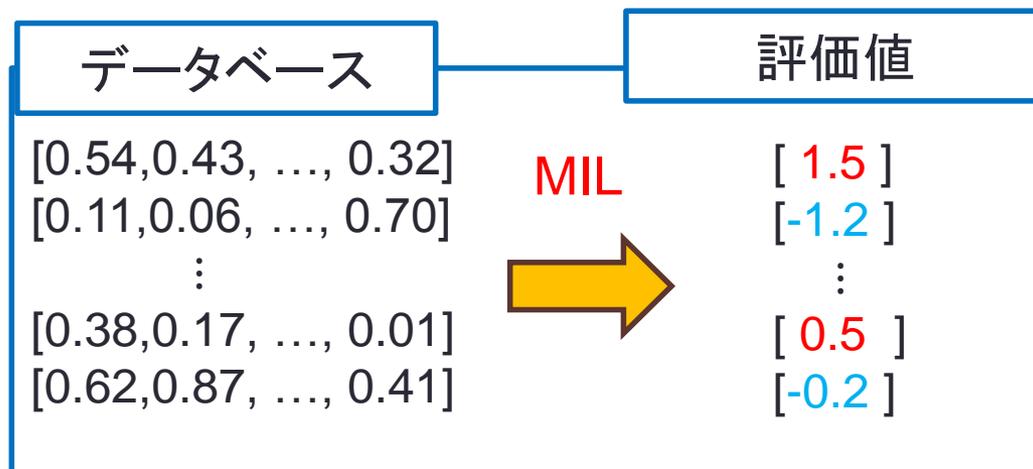
対応した座標の局所特徴
に対する重みとする

提案手法詳細

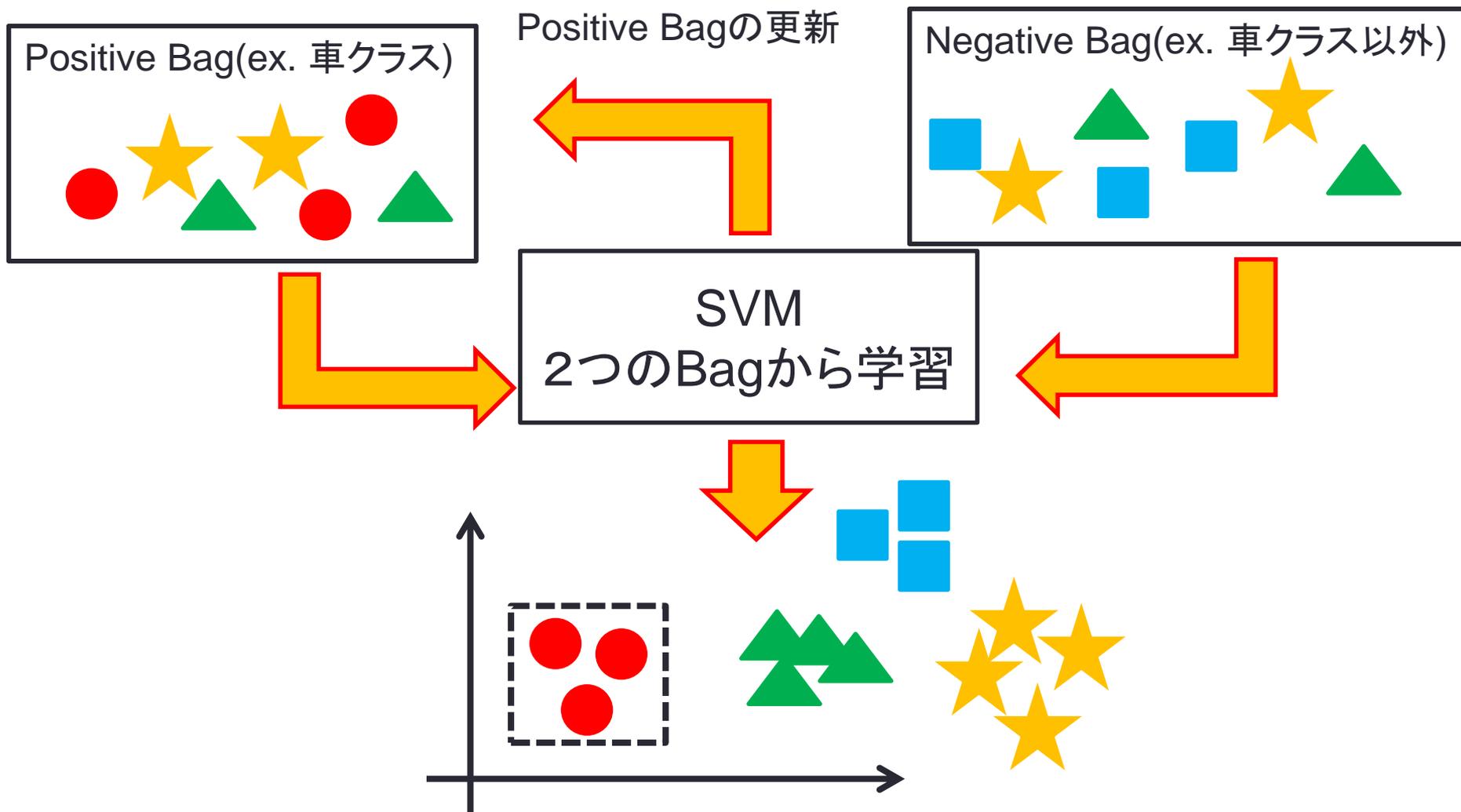
• 特徴の重み付け

2. mi-SVMによる重み付け.

MIL(Multiple Instance Learning)手法の一種で、あるクラスのデータベースだけに現れる画像特徴と、他のクラスにも現れる画像特徴を識別することができる。



mi-SVM



ポジティブ画像にだけ現れる重要な特徴に重み付け

mi-SVM

Positive データセット

(ex. 車クラス)

[0.54,0.43, ..., 0.32]

[0.11,0.06, ..., 0.70]

⋮

[0.38,0.17, ..., 0.01]

[0.62,0.87, ..., 0.41]

Negative データセット

(ex. 車クラス以外)

[0.04,0.35, ..., 0.54]

[0.11,0.37, ..., 0.53]

⋮

[0.10,0.85, ..., 0.29]

[0.21,0.43, ..., 0.94]

mi-SVM



Positive データセット

(ex. 車クラス)

mi-SVM重み

[0.54,0.43, ..., 0.32] 0.98

[0.11,0.06, ..., 0.70] -1.56

⋮

[0.38,0.17, ..., 0.01] 1.78

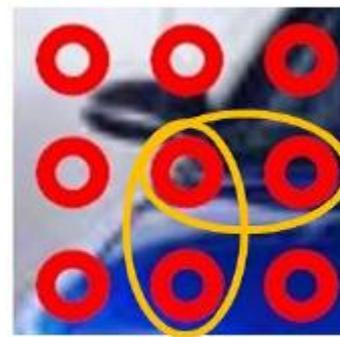
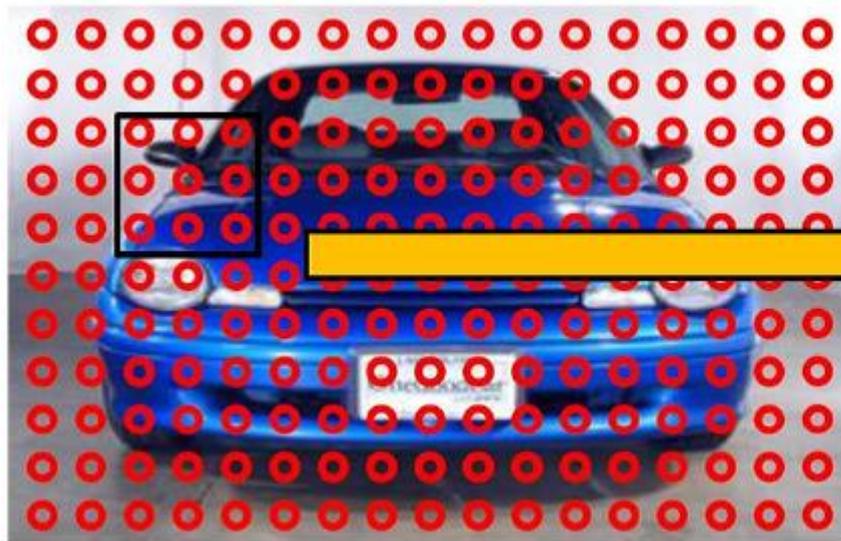
[0.62,0.87, ..., 0.41] -0.67

提案手法詳細

• 特徴のペア化

- 隣接特徴5個までをペアにする。
- 同じスケール内で隣接特徴を選択する
- 特徴数を増やすことができる

$$[x_1, y_1, x_2, y_2, f_1, f_2]$$



評価実験

1. Caltech101クラスから10クラスを選択して実験
 - 各クラス学習画像15枚、テスト画像15枚
 - 3 fold - crossvalidationで評価
 - 8ピクセルグリッド、5スケールによりSIFT特徴を抽出
 - クラスあたりの特徴数 30,000(ペア化なし)、150,000(ペア化あり)

barrel	bass	car side	cougar body	gareld
headphone	joshua tree	laptop	saxphone	watch

評

2. 狹

- V
- 像
- 各
- D
- 夕

• 評



評価実験(10クラス分類)

- ベースライン

- NBNN[Boiman 2008]

$$\hat{c} = \arg \min_c \sum_{i=1}^n \|d_i - NN_c(d_i)\|^2 + \alpha \left\| l_{d_i} - l_{NN_c(d_i)} \right\|^2$$

l_x : x の座標(x, y)

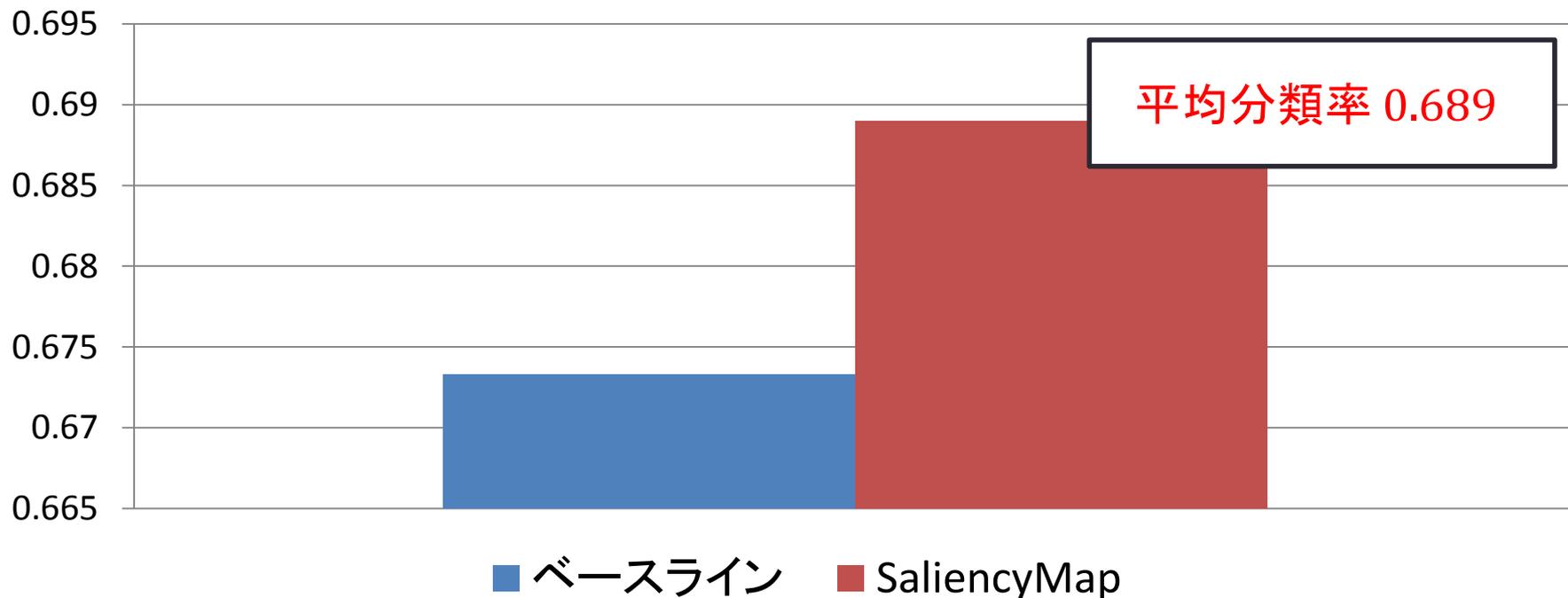
- α を変化させて一番精度がよかった値をベースラインとする
 - **平均分類率 0.673**
- 提案手法に対しても、 α を変化させて実験を行った。

実験結果(10クラス分類)

- Saliency Mapによる重み付けの結果

$$\hat{c} = \arg \min_c \sum_{i=1}^n \exp(-w_{SaliencyMap}/\beta) \|d_i - NN_c(d_i)\|^2$$

- 各クロスバリデーションで β の値を変化させて実験を行った平均の分類率

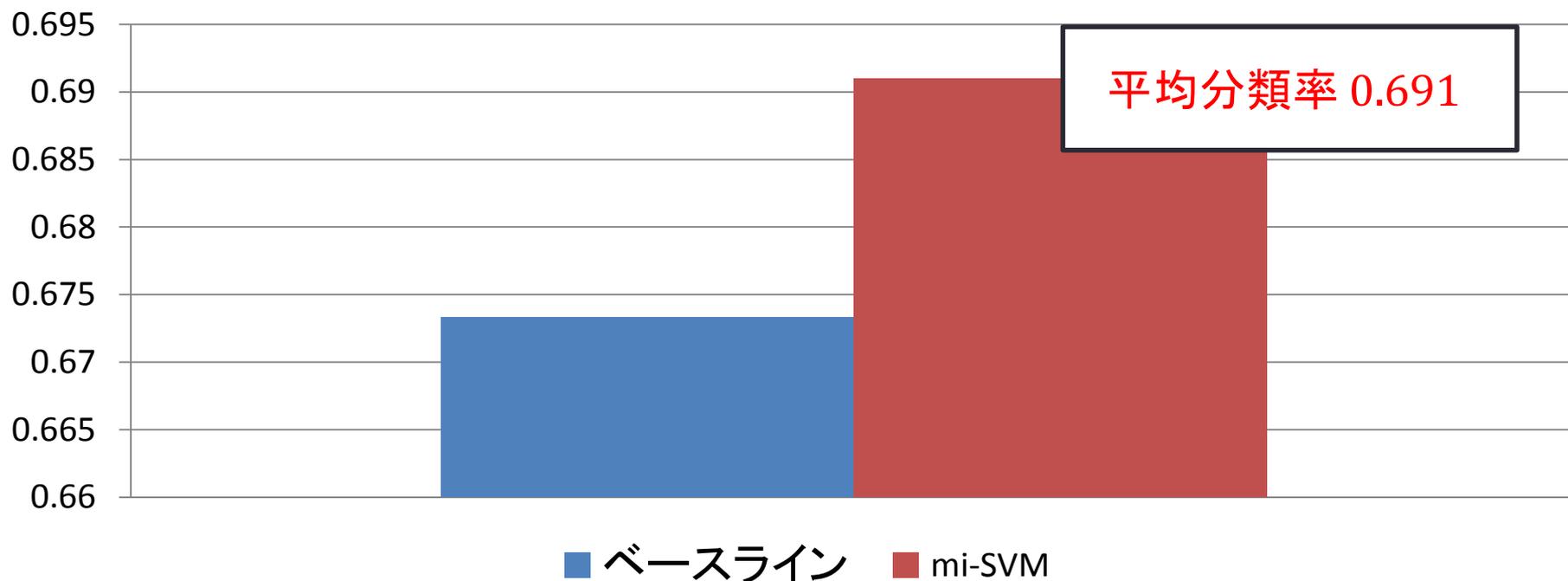


実験結果(10クラス分類)

- mi-SVMによる重み付けの結果

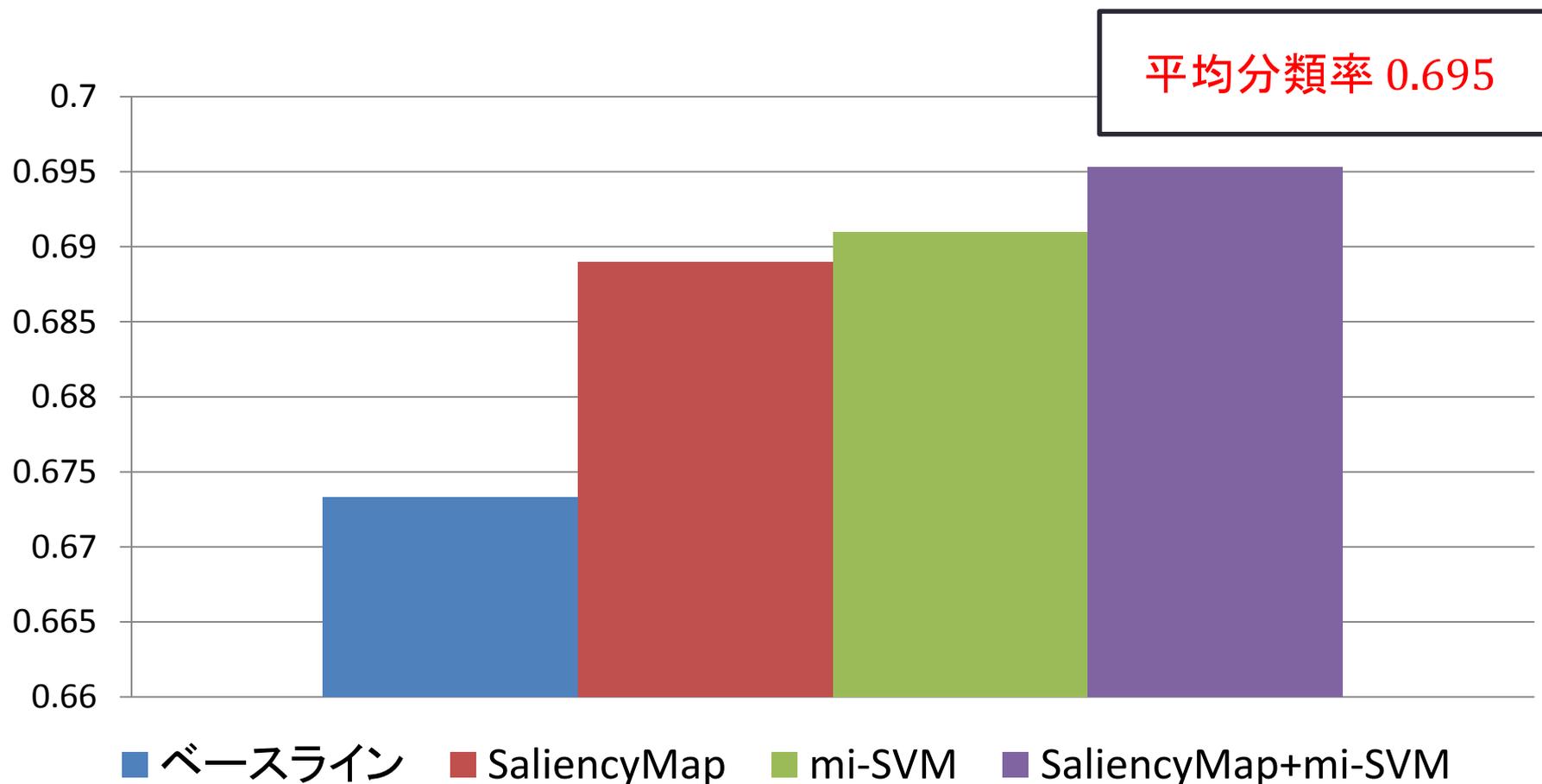
$$\hat{C} = \arg \min_c \sum_{i=1}^n \exp(-w_{mi-svm}/\gamma) \|d_i - NN_c(d_i)\|^2$$

- 各クロスバリデーションで γ の値を変化させて実験を行った平均の分類率



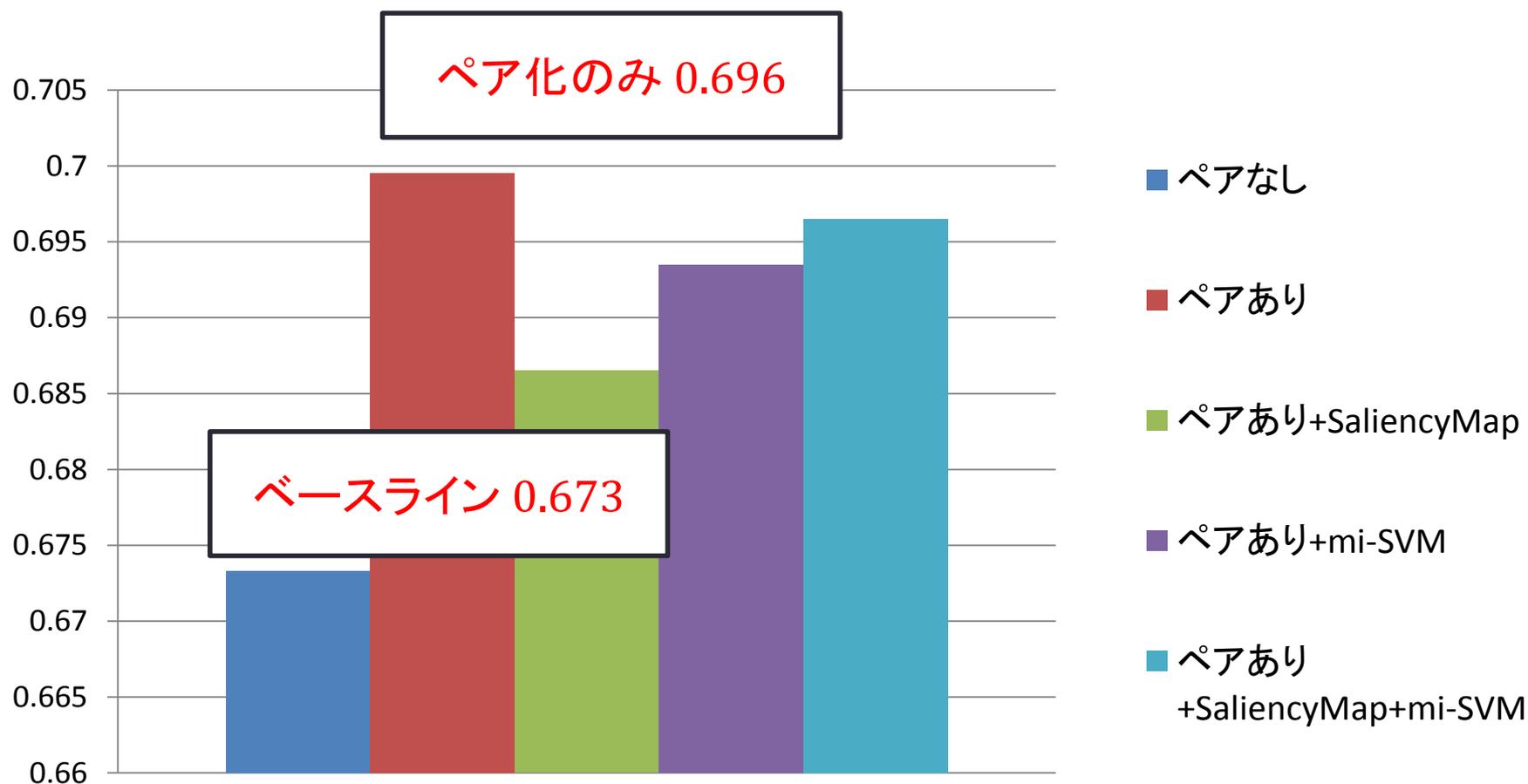
実験結果(10クラス分類)

- 2つの重み付けを組み合わせた結果



実験結果

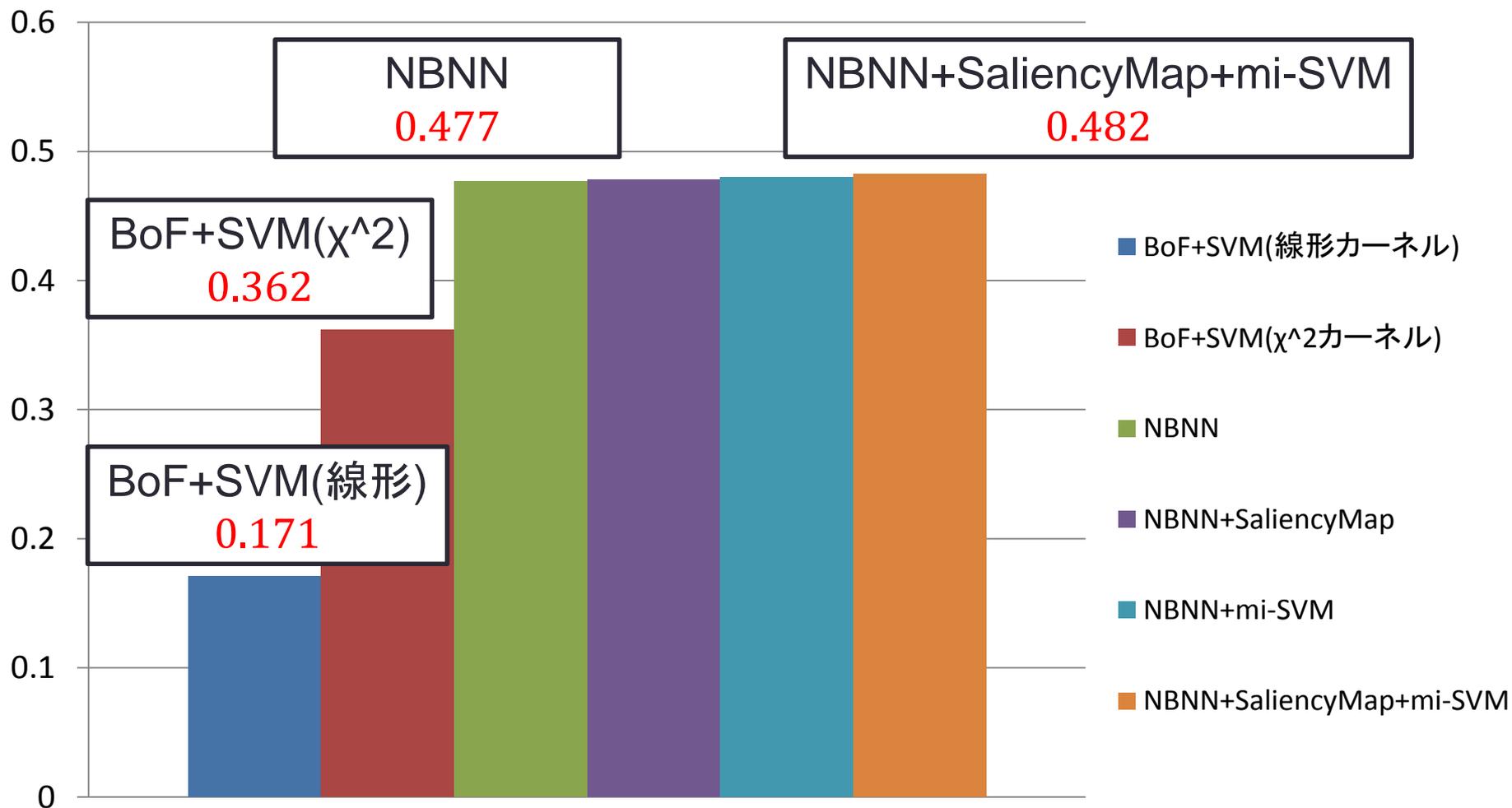
• 特徴量ペア化の結果



実験結果(25クラス分類)

- ベースライン
 - BoF(codebook size = 1000)+サポートベクターマシン(線形カーネル)
 - BoF(codebook size = 1000)+サポートベクターマシン(χ^2 カーネル)
 - NBNN[Boiman 2008]
- 全てにおいて座標情報(α)は利用しない
- またクラスあたりの特徴数が多いので、ペア化は利用しない

実験結果(25クラス分類)



まとめ

• NBNN手法の改良

- SaliencyMap、mi-SVMによる重み付け
- 特徴のペア化

• 評価実験

10クラス分類	NBNN	提案手法
分類率	0.673	0.695

25クラス分類	SVM(χ^2)	NBNN	提案手法
分類率	0.362	0.477	0.482

• データセット

- 関連研究ではノイズが無い綺麗なデータセットでしか、実験をしていないが、NBNN手法は**ノイズが有る**信頼性の低いデータセットに対して、BoFベース手法を上回る精度が期待できることが分かった。

今後の課題

- mi-SVMについて
 - 学習時間の関係から学習特徴数を減らした。
 - ポジティブ5000, ネガティブ30000しか利用していない。
 - より大規模かつ高速な重み付けが可能になるMIL手法を検討する必要がある。
- 有効な利用方法
 - 信頼性の低いデータセットを利用するシーン
 - Web動画、Web画像の自動ランキングなど

ご清聴ありがとうございました

Caltech101における比較結果

Caltech101	
Boiman (CVPR2008)	65%
Wang(ECCV2010)	48%
Tuytelaars (ICCV2011)	62.7%

Caltech 5クラス	
Boiman (CVPR2008)	73.1%
Wang(ECCV2010)	89.8%
BoW + χ^2 -SVM (Wang(ECCV2010))	89.1%