

9:30-11:30 企画セッション

画像処理への機械学習の応用

特徴量変換とは

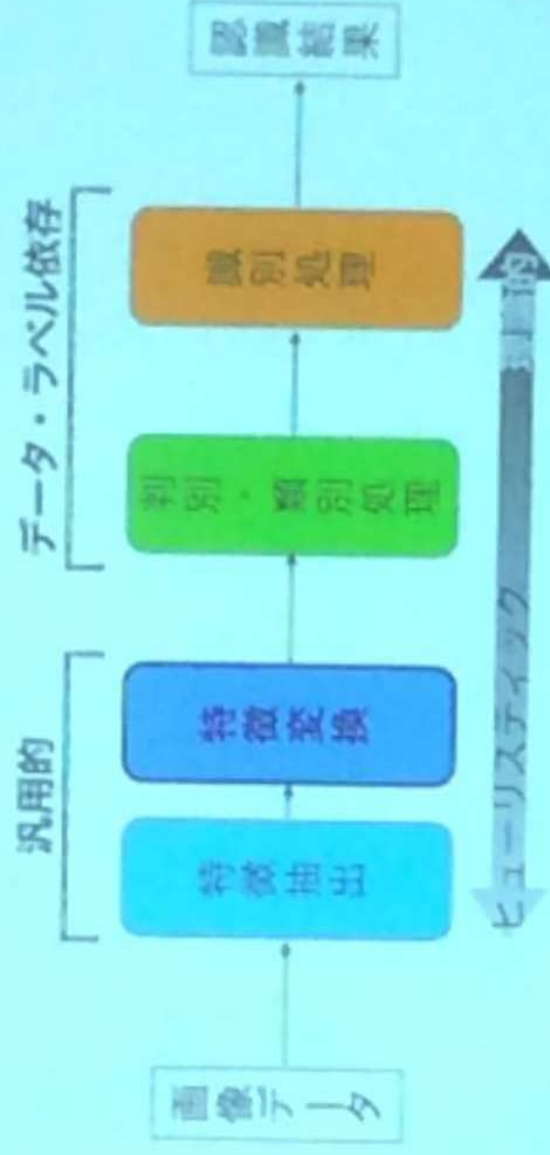
2

パターン認識における特徴量変換の位置付け



特徴量変換とは

パターン認識における特徴量変換の位置付け

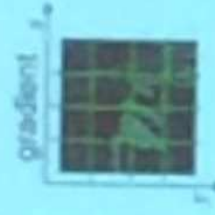


特徴変換：
・ 汎用性（データに依って処理を変えない）
・ 弁別性（特徴の判別力を向上）



人為的特徴量 = Hand-crafted特徴

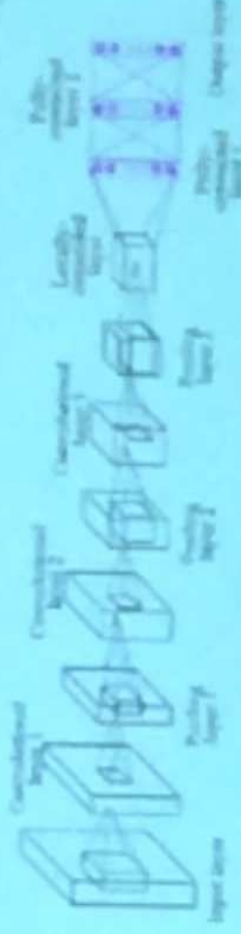
例: image patch



勾配方向ヒストグラム
(SIFT)

学習的特徴量 = 畳み込みニューラルネット(CNN)特徴

例:



転移された (pre-trained on ImageNet) CNNモデルを用いる

人工的特徴量 = Hand-crafted特徴

T. Kobayashi,

"Structured Feature Similarity with Explicit Feature Map",
CVPR2016

学習的特徴量 = 畳み込みニューラルネット(CNN)特徴

T. Kobayashi,

"Learning Additive Kernel For Feature Transformation
and Its Application to CNN Features",
BMVC2016

科研費「パターン認識のための特徴量変換に関する研究」15K00261



距離尺度

主にヒストグラム特徴に対する手法が提案されてきた

- χ^2 (カイ2乗)-distance
- Earth Mover's distance (EMD)
- Faster EMD (高速版EMD)
- SiftDist (SIFTでのEMD)
- Diffusion distance (ヒストグラム上の拡散過程)

ここでは特徴の物理的構造に着目した特徴変換を考える

ポイント

1. 特徴量のテンソル構造
2. SSIMに基づく距離尺度



特徴量のテンソル構造 (1)

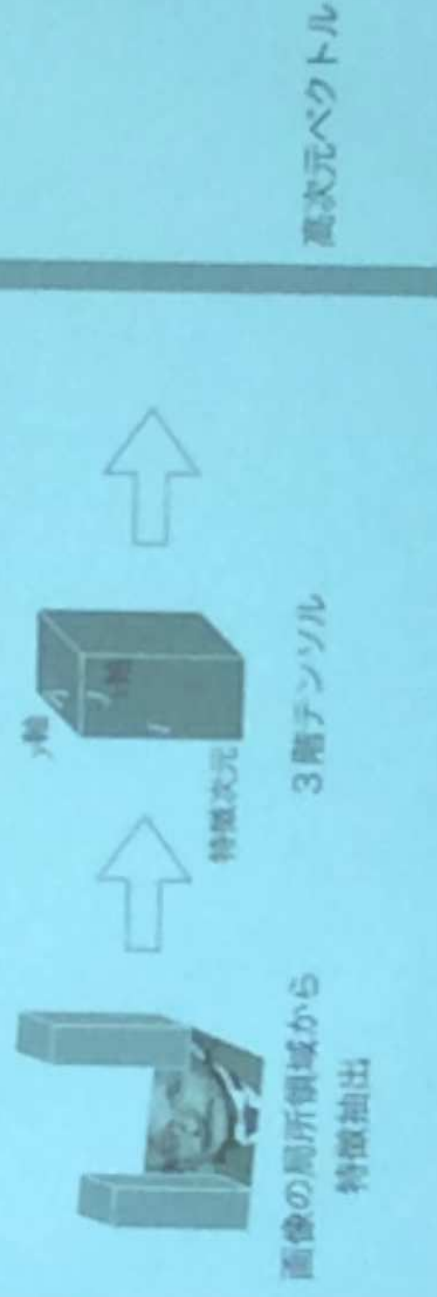
9

画像特徴量は多くの場合に**テンソル構造**を内包する

画像 ($x \times y$: 2次元) から**特徴量** (1次元) を抽出

→ 3階テンソル ($x \times y \times \text{特徴}$)

→ (従来は) ベクトル



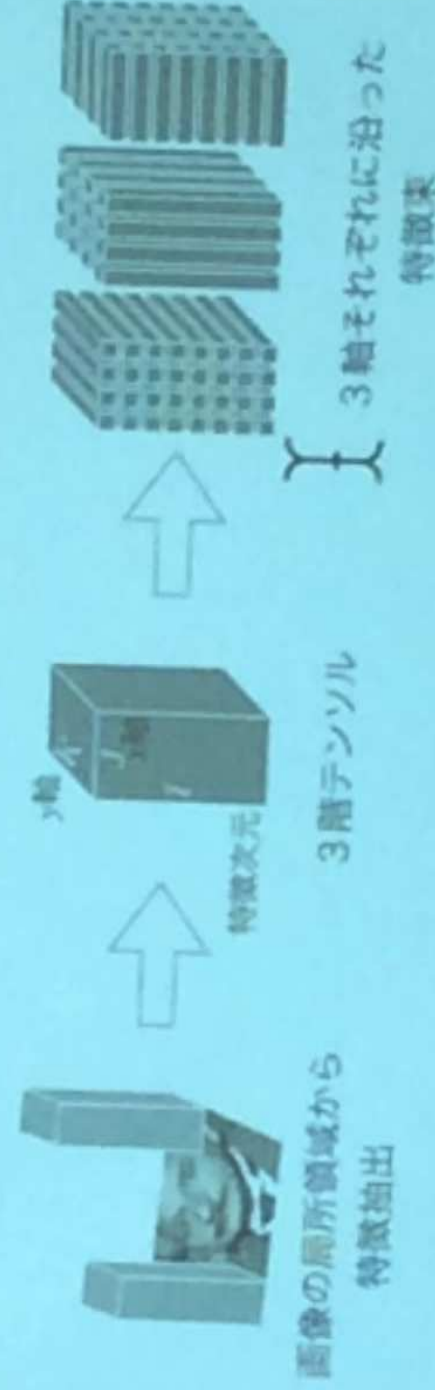
特徴量のテンソル構造 (2)

10

テンソル構造を保持した表現を考える

テンソルの各軸 (x, y , 特徴) に沿った特徴束を一つの単位として、そこでの距離尺度を考える

- 特徴束を大きくしすぎると、元のベクトル表現
- 特徴束を小さくしすぎると、特徴要素1個ずつ



SSIMに基づく距離尺度 (1)

Structural Similarity Index Measure (SSIM)

画像の品質評価のための定量的指標として用いられる



Z. Wang and A. C. Bovik, Mean squared error: Love it or leave it? - a new look at signal fidelity measures, *IEEE Signal Processing Magazine*, 26(1):96-117, January 2008

SSIMに基づく距離尺度 (1)

Structural Similarity Index Measure (SSIM)

画像の品質評価のための定量的指標として用いられる



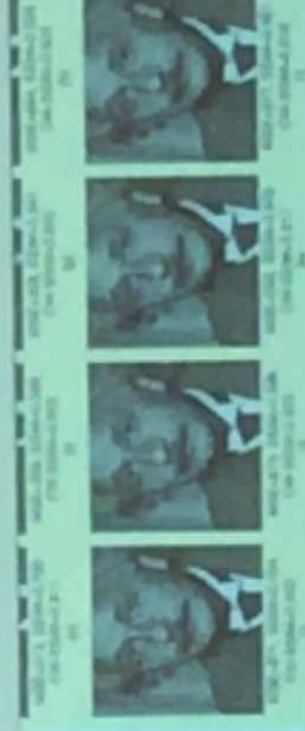
SSIM formulation. 画像 x, y に対して

$$S_{\text{org}}(x, y) = \underbrace{\mathcal{M}(x, y) \times \mathcal{V}(x, y) \times \mathcal{C}(x, y)}_{\text{人}}$$

明るさ
の類似性

コントラスト
の類似性

構造パターン
の類似性



Z. Wang and A. C. Bovik, Mean squared error: Love & lose it? a new look at signal fidelity measures, *IEEE Signal Processing Magazine*, 26(1):96–119, January 2009

SSIMに基づく距離尺度 (2)

特徴束においてSSIMを距離尺度として採用

特徴束 x y に対して

$$S_{\text{org}}(x, y) = \underbrace{M(x, y)}_{\text{明るさ}} \times \underbrace{V(x, y)}_{\text{コントラスト}} \times \underbrace{C(x, y)}_{\text{構造パターン}} \rightarrow C(x, y) = \frac{(x - u(x))^T (y - u(y))}{\|x - u(x)\|_2 \|y - u(y)\|_2} \quad \text{相関係数}$$

$$k(u(x), u(y)), \quad k(q(x), q(y)), \quad \text{where 類似度関数 } k(a, b) = \frac{2ab}{a^2 + b^2}$$

平均値の類似度 標準偏差の類似度

特徴変動への頑健性を高めるため乗法的表現から加法的表現へ

- 上記の元表現 S_{org} は、 M, V, C のわずかな変化にも敏感

乗法的: $S_{\text{org}} = M \times V \times C$

加法的: $S_{\text{add}} = w_M M + w_V V + w_C C$

$(w_M, w_V, w_C) = (2, 2, 1)$
 $M \quad V \quad C$

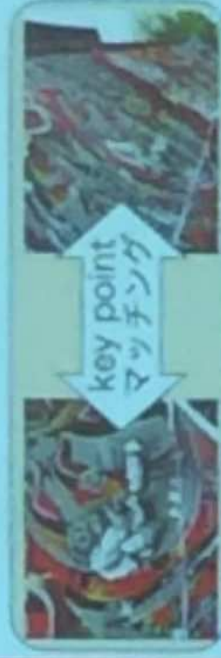
non-neg. feat.	$[0, +1]$	$[0, +1]$	$[-1, +1]$
real feat.	$[-1, +1]$	$[0, +1]$	$[-1, +1]$

 value range
 $(w_M, w_V, w_C) = (1, 2, 1)$

実験結果

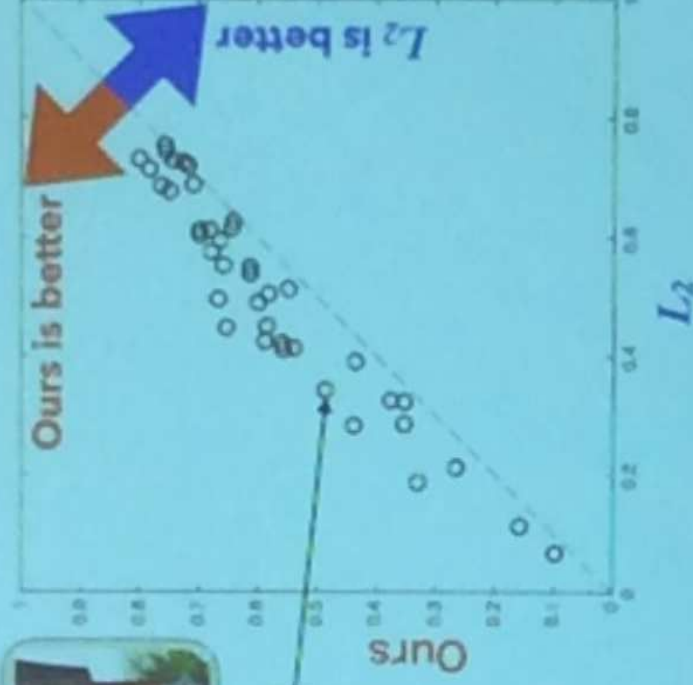
19

keypoint マッチング@INRIA データセット



(Average Precision by Ours

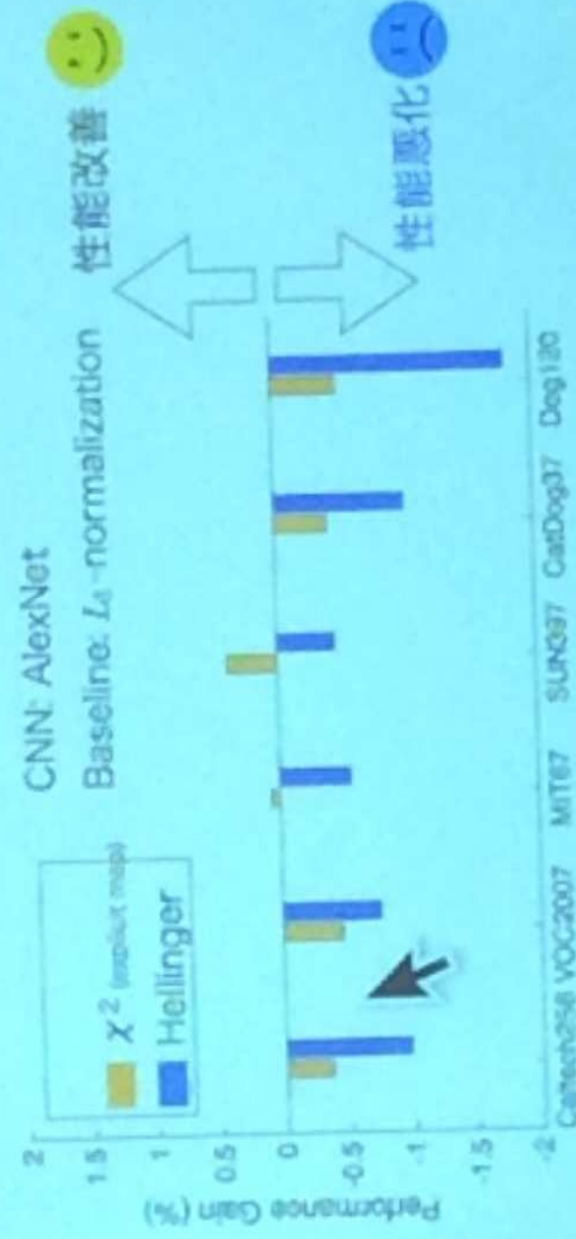
Average Precision by L_2)



学習的特徴量では？

ヒストグラムなどの人為的特徴量では特徴変換はうまくいくが、学習的特徴量（CNN特徴）でもうまくいくか？

- ・ 実はあまりうまくいかない...



なぜか？

24

距離尺度 = 非線形 (kernel) 関数は、対象の特徴量の特性に沿って決めていた。

CNN特徴のようないくつかからない特徴に対して強引に適用するのは困難であるらしい。

そこで、学習的特徴量に対する距離尺度も学習する
ポイント

1. 距離尺度のデータからの学習
2. 弁別性・汎化性の高い変換方式

加法的カーネル表現

加法的カーネルから出発

D 次元特徴ベクトル $x, y \in \mathbb{R}^D$ に対して

加法的カーネル: $\bar{k}(x, y) = \sum_{i=1}^D k(x_i, y_i)$, 要素毎のカーネル関数の和

カーネル関数の隔写像展開:

$$\begin{aligned} k(x, y) &= \int \phi(\gamma, x) \phi(\gamma, y) d\lambda \approx \sum_{l=1}^K \phi(\gamma_l, x) \phi(\gamma_l, y) \Delta_l \\ &= \phi(x)^T \phi(y) \end{aligned}$$

加法的カーネルの学習 \rightarrow 関数 $\phi(x)$ の学習

3種類のCNNモデル



データセット

Video: HMDB51, UCF101/50, Hollywood2 (動作認識)
 Image: Caltech256, VOC2007, (物体認識)
 MIT67, SUN397, (シーン認識)
 CatDog37, Dog120 (詳細な種別認識)

Video Datasets

33

HMDB51



shoot
gun



sit



situp



smile



smoke

UCF101/50



Backstroke



Swimming Diving



Barrel Press



Shooting



Rolling



Pushing



Charm and Love



Charm and Love



Charm and Love



Charm and Love



Charm and Love



Charm and Love

Hollywood2

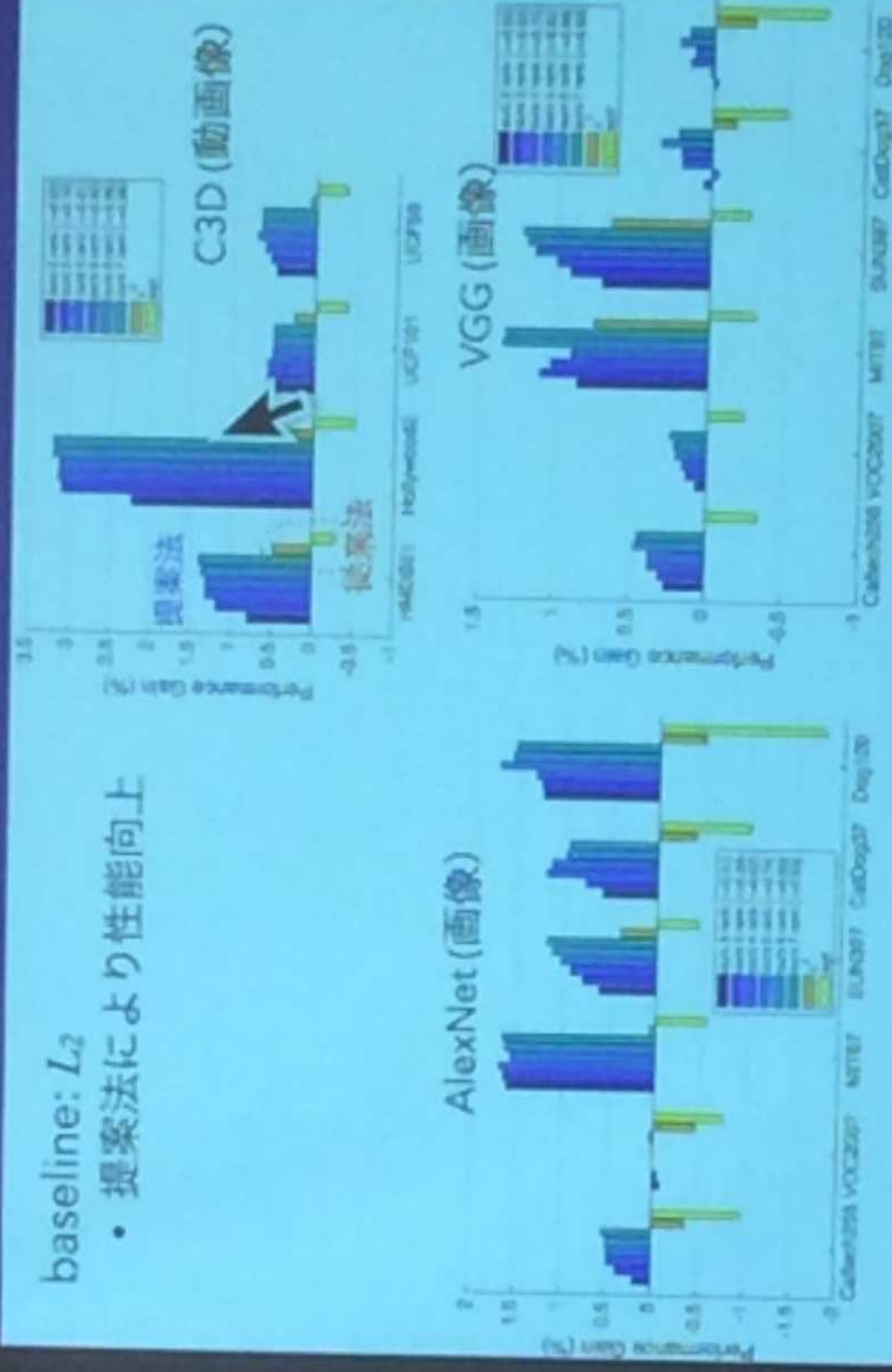


実験結果

35

baseline: L_2

- 提案法により性能向上



まとめ

人為的特徴量と学習的特徴量に対しての特徴変換

人為的特徴 = hand-crafted特徴 : SIFT, HOGなど

学習的特徴 = CNN特徴 : AlexNet, VGGなど

人為的特徴量に対しては内在しているテンソル構造に着目し、学習的特徴量に対しては距離尺度の学習を通して、それぞれの特徴変換手法を示した。

学習されたカーネル関数（の類似性）は、CNN特徴の

特性解析に使えるかも...

