RITSUMEIKAN

陳研究室 CHEN LABORATORY

2013年度NIFS画像計測研究会

マイクロTextonにおけるパラメトリック確率モデルおよび画像理解への応用

韓　先花　　　　　陳　延偉

立命館大学

Intelligent Image Processing Lab.

研究概要

統計解析手法

(1) 主成分分析
(2) 独立成分分析
(3) 多様体学習 （LLE, LPP等）
(4) Sparse Coding

拡張

①非線形（kernel関数）
② テンソルベース（多重線形）統計解析

Deep Learning

超解像度技術
画質改善

機械学習・統計解析

画像認識・理解

（1） 肝硬変診断支援；
（2） 医用画像検索による医療疾患診断支援

医用画像処理
医療診断支援

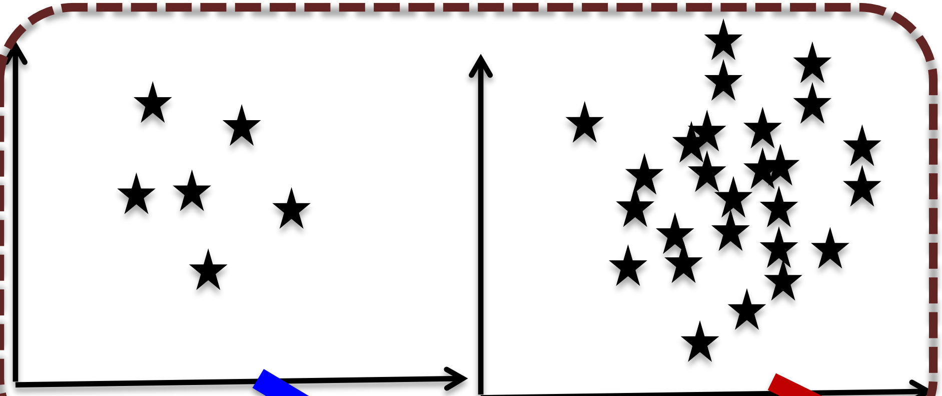
（3）医用データの画質改善、エッジ抽出
（4）学習ベース医用画像超解像度

発表内容

目的: 画像認識・理解

データの集約
コアー情報の抽出

Data 1　Data 2



画像 1　画像 2

類似度

アウトライン

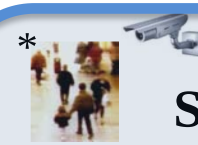
- Introduction
- Related Works
- Micro-Texton space for local structure representation
- Parametric probability model and high-order statistics
- Experimental Results
  - HEp-2 Cell pattern recognition
  - Food image classification
  - Texture classification
- Conclusions

Introduction

Image (Object/Scene/Texture) understanding

Applications

Digital libraries, Space science, Web searching, Geographic information systems, Biomedicine, Surveillance and sensor system, Commerce, Education



Security

Robot vision

Medical diagnosis

……

- Difficulties
  - Variable & uncontrolled image conditions
  - Complex and hard-to-describe objects in image
  - Objects occluding other objects

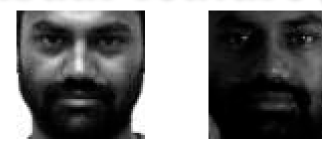
Variations in semantic

Similar in global shape

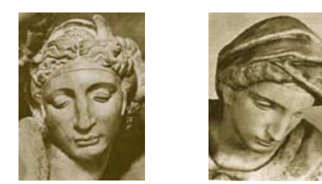
Different semantic

Variations in Visual feature

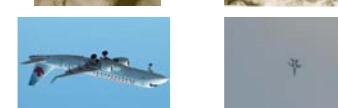
(1) Illumination



(2) View point



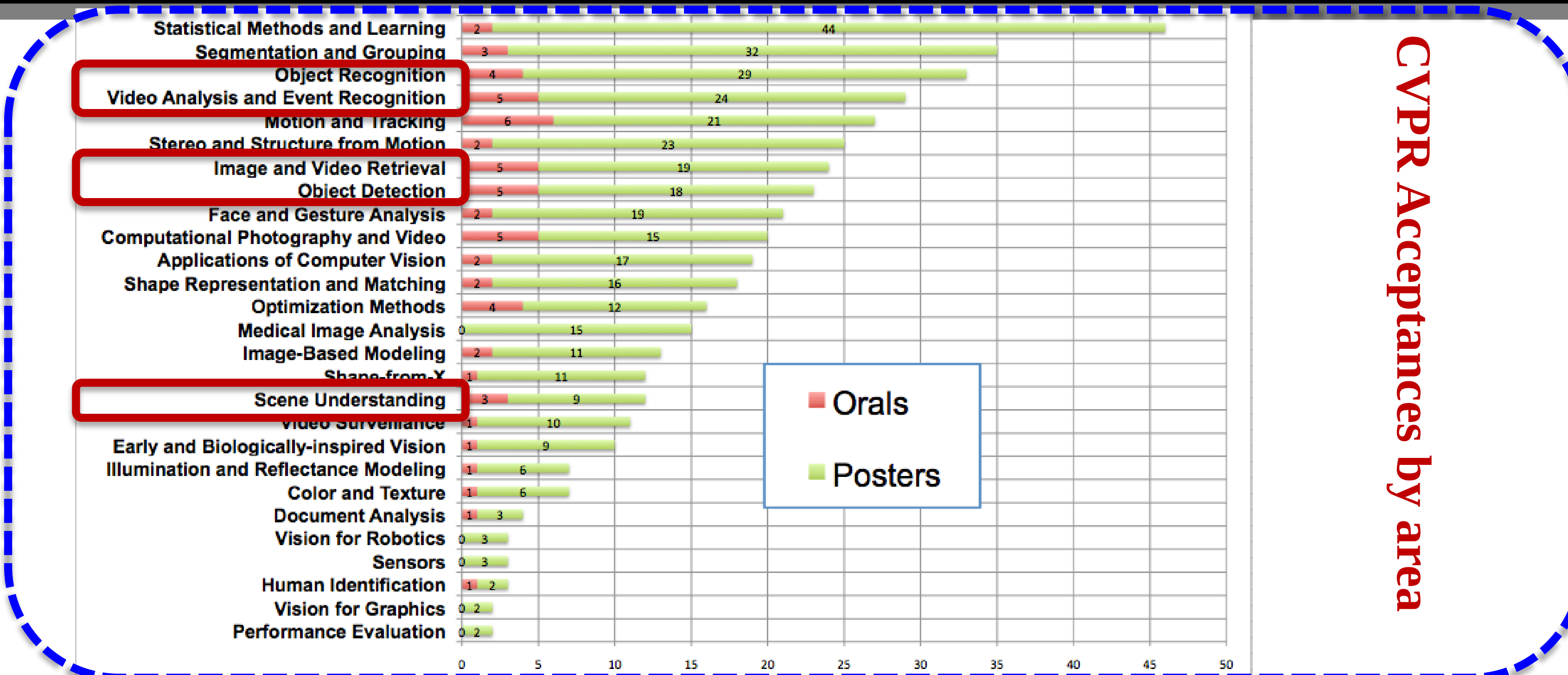
(3) Scale



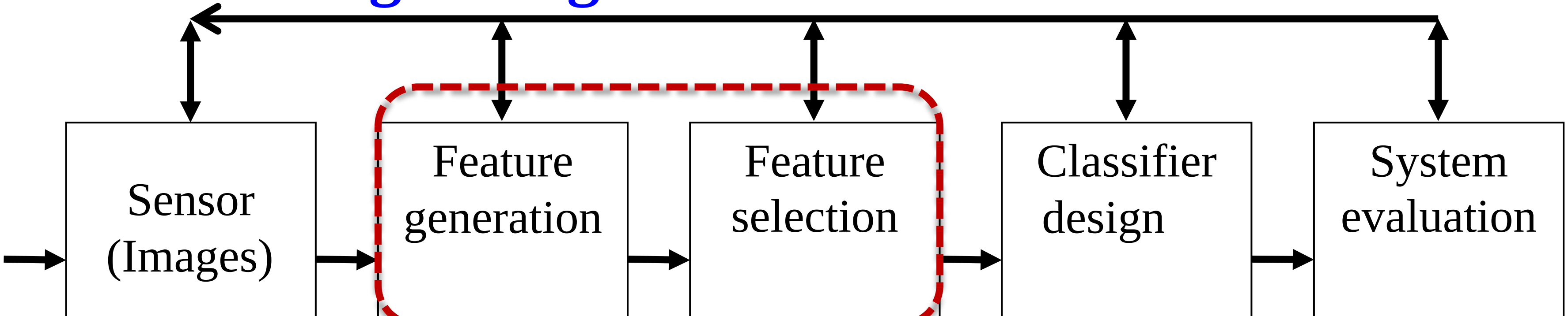
(4) Intra-class

……

Introduction

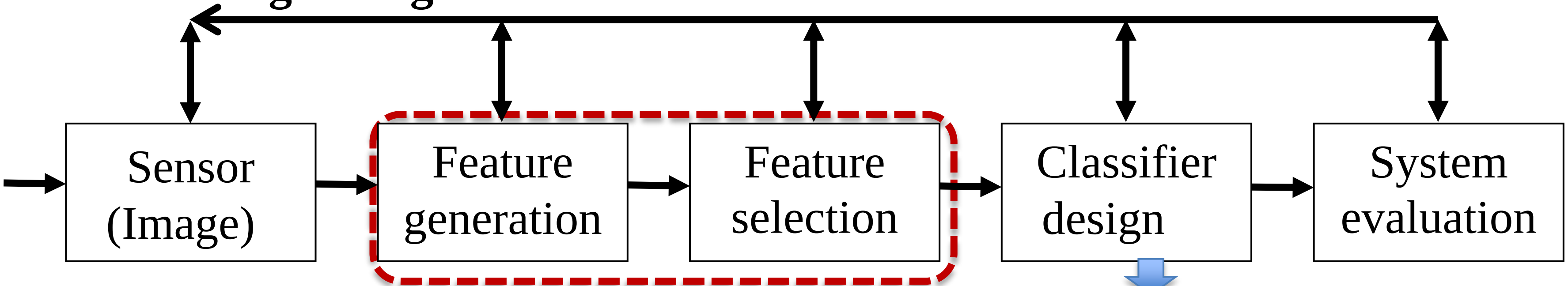


General image recognition flowchart



Related works

General image recognition flowchart



SVM, **Random Forest etc.**

Feature Generation

(1) Global features:

A. Intensity-based feature (Alignment)
B. Color information (histogram, Moment)
C. Shape information (edge histogram, HOG)
D. Texture information (Gist)

(2) Local binary patter… …ttern Histogram
Local descriptor (S… …g-of-Feature (BOF)

Feature Selection

Subspace learning (PCA,ICA)
Manifold learning (LPP, Isomap)
Discriminant analysis (LDA)
……

Popular & Reasonable performance

Related works

(2) Local binary pattern **Integration** Pattern Histogram

200	80	70
23	100	90
56	30	120

Subtraction →

100	-20	-30
-77		-10
-44	-70	20

→

1	0	0
0		0
0	0	1

Binary Number: 10001000

Decimal Number: 138

→

LBP value

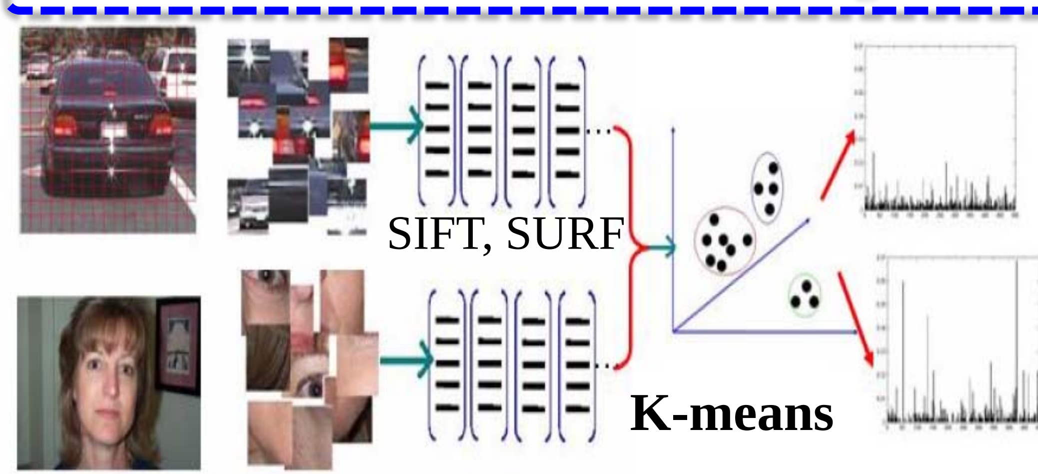
	136	

Drawbacks:

(1) Information lost: **only quantizing the differential vector into binary**

(2) Only low-order statistics

Local descriptors (SIFT, SURF) **Integration** Bag-of-Feature (BOF)



Drawbacks:

(1) Time-Consuming:
- **local descriptor extraction**
- **Enough number visual words**

(2) Information lost:
Approximation each local descriptors with the pre-defined visual words

(3) Only Low-order statistics

Proposed strategy

(1) Directly use the differential vector as local descriptor, called **micro-Texton**

3×3 patch

200	80	70
23	100	90
56	30	120

Subtraction

Difference

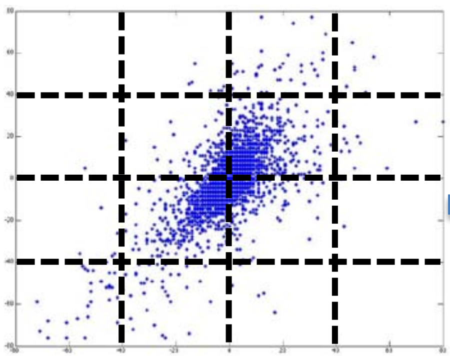
100	-20	-30
-77		-10
-44	-70	20

Differential vector

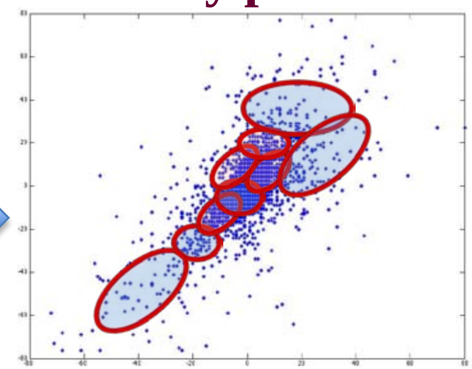
$$\begin{bmatrix} 100 \\ -20 \\ -30 \\ -10 \\ 20 \\ -70 \\ -44 \\ -77 \end{bmatrix}$$

(2) Model the micro-Texton using a parametric probability process: GMM

A simple example



Uniformly quantization

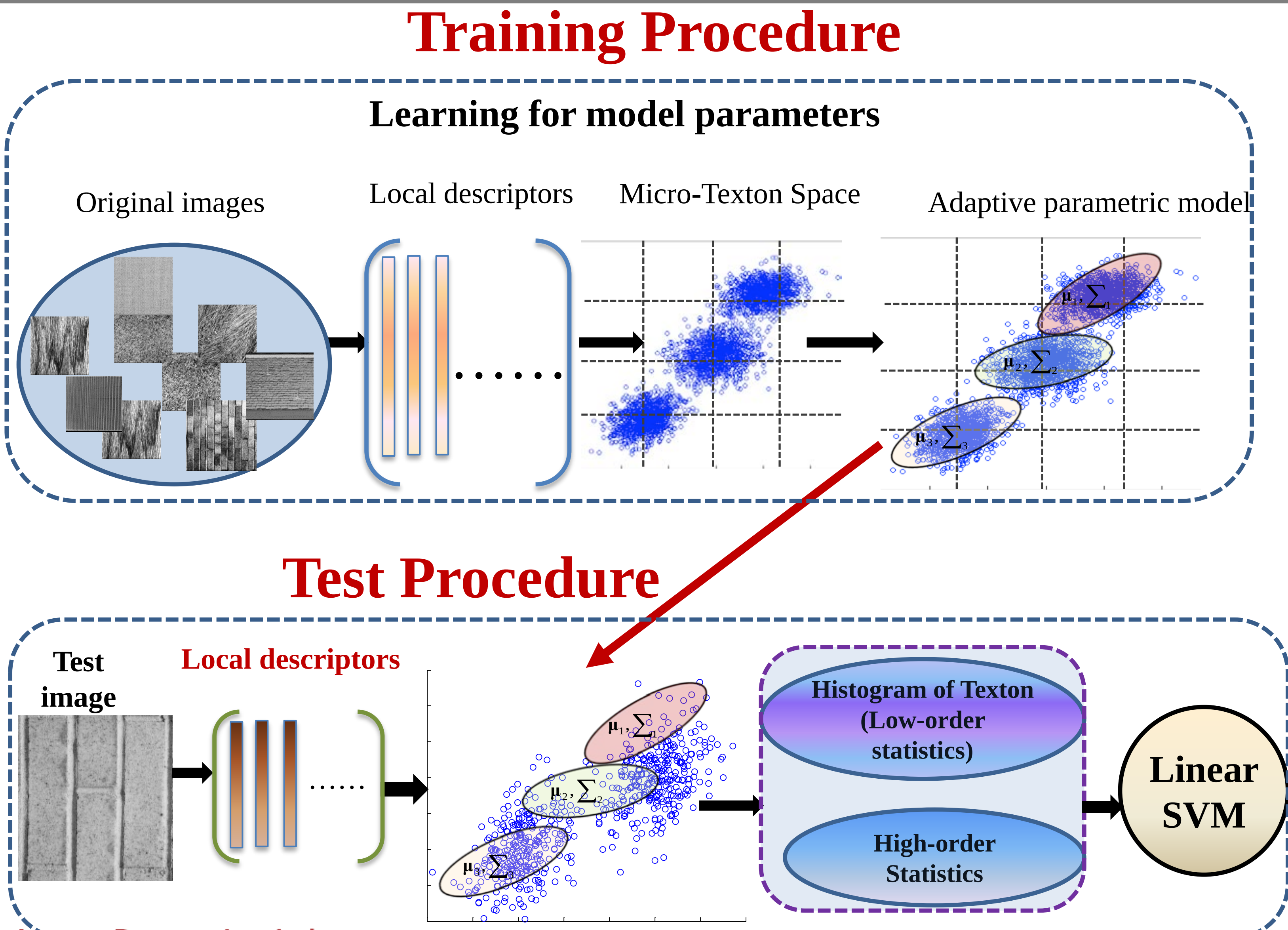


Adaptively modeling

Assign a local descriptor to several models (not only one like in BOF)

(3) Extract not only low-order statistics but also high-order statistics for image representation

Flowchart of the Proposed strategy



Proposed Strategy: Adaptively modeling micro-Texton

Given some local descriptors (micro-Texton) X extracted from Training images, **we model them using the parametric probability process: GMM**

$$P(\mathbf{X}/\lambda) = \sum_{k=1}^{K} w_k N(\mathbf{X}/\mu_k, \Sigma_k)$$

$$= \sum_{k=1}^{K} w_k \frac{1}{(2\pi)^{\frac{D}{2}} \mid \Sigma_k \mid^{\frac{1}{2}}} exp\{-\frac{1}{2}(\mathbf{X}-\mu_k)^T \Sigma_k^{-1}(\mathbf{X}-\mu_k)\}$$

Unknown parameters:
Need to learn using the micro-Texton ensemble X

Expectation maximization (EM) method: With initial paramters:

(1) Expectation Step: compute the responsibilities (posterior probability for each Texton samples):

$$\gamma_{t,k} = \frac{w_k P(k/\mathbf{x}_t, \lambda)}{\sum_{k=1}^{K} w_k P(k/\mathbf{x}_t, \lambda)}$$

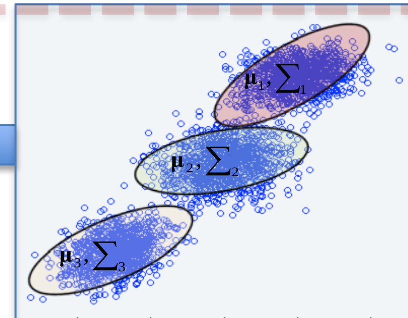
(2) **Maximization Step: compute the weighted** means, variances and mixture weights:

$$\mu_k = \frac{\sum_{t=1}^{T} \gamma_{t,k} \mathbf{x}_t}{\sum_{t=1}^{T} \gamma_{t,k}} \qquad \Sigma_k = \frac{\sum_{t=1}^{T} \gamma_{t,k} \mathbf{x}_t^2}{\sum_{t=1}^{T} \gamma_{t,k}} - \mu_k^2 \qquad w_k = \frac{1}{T}\sum_{t=1}^{T} \gamma_{t,k}$$

Proposed Strategy:
statistics (low and high orders)

Known model parameters

$$\{w_k, \mu_k, \boldsymbol{\Sigma}_k, k = 1, \ldots, K\}$$



Constructed model using training database

$$w_k = \frac{exp(\alpha_k)}{\sum_{j=1}^{K} exp(\alpha_j)} \qquad \sigma_k = diag(\boldsymbol{\Sigma}_k)$$

Only use diagonal elements

Given the local descriptors (micro-Textons) $X=[x_1,x_2,\ldots,x_T]$ extracted from a test image, our goal is to extract discriminant feature for image representation:

The deviation statistics of X to the learned parameters

Mathematical definition: $G_\lambda^{\mathbf{X}} = \nabla_\lambda \log P_\lambda(\mathbf{X})$

Computational implementation:

$$\frac{\partial P(\mathbf{X}|\lambda)}{\partial \alpha_k} = \sum_{t=1}^{T} [\gamma_t(k) - w_k] \qquad \frac{\partial P(\mathbf{X}|\lambda)}{\partial \mu_k^d} = \sum_{t=1}^{T} \gamma_t(k) [\frac{x_t^d - \mu_k^d}{(\sigma_k^d)^2}]$$

$$\frac{\partial P(\mathbf{X}|\lambda)}{\partial \sigma_k^d} = \sum_{t=1}^{T} \gamma_t(k) [\frac{(x_t^d - \mu_k^d)^2}{(\sigma_k^d)^3} - \frac{1}{\sigma_k^d}]$$

Proposed Strategy: statistics (low and high orders)

Given the local descriptors (micro-Textons) $X=[x_1,x_2,\ldots,x_T]$ extracted from a test image, our goal is to extract discriminant feature for image representation:

The deviation statistics of X to the learned parameters

Computational implementation:

The deviation statistic to w_k: $\frac{\partial P(\mathbf{X}|\lambda)}{\partial \alpha_k} = \sum_{t=1}^{T}[\gamma_t(k) - w_k]$

Similar to histogram of texton in an image (like in BOF)
Called **0th order statistics** (low-order statistics)

The deviation statistic to $\mu_k, \mathbf{\Sigma}_k$

$$\frac{\partial P(\mathbf{X}|\lambda)}{\partial \mu_k^d} = \sum_{t=1}^{T} \gamma_t(k)[\frac{x_t^d - \mu_k^d}{(\sigma_k^d)^2}]$$

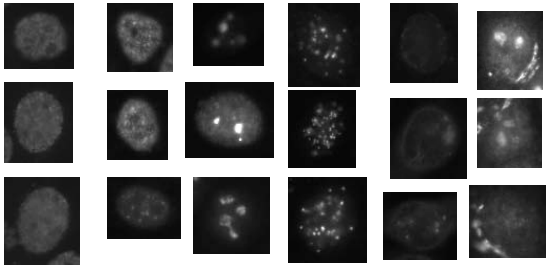
$$\frac{\partial P(\mathbf{X}|\lambda)}{\partial \sigma_k^d} = \sum_{t=1}^{T} \gamma_t(k)[\frac{(x_t^d - \mu_k^d)^2}{(\sigma_k^d)^3} - \frac{1}{\sigma_k^d}]$$

The **first and second order statistics**
(high-order statistics)

Experimental results

Three different Databases

(1) HEp-2 Cell database (ICIP2013 competition on cell classification by fluorescent image analysis):



1. **Two intensity Type: Positive and intermediate**
2. **More than 10000 images**
3. **6 cell patterns for each intensity type**

(2) Food image datasets (Pittsburgh fast-food image dataset: PFID): large-pose variance

(a) **61 categories**

(b) **3 instances (6 different pose images) for each category**

(3) Texture dataset (TIPS-2a): Large scale and pose variance

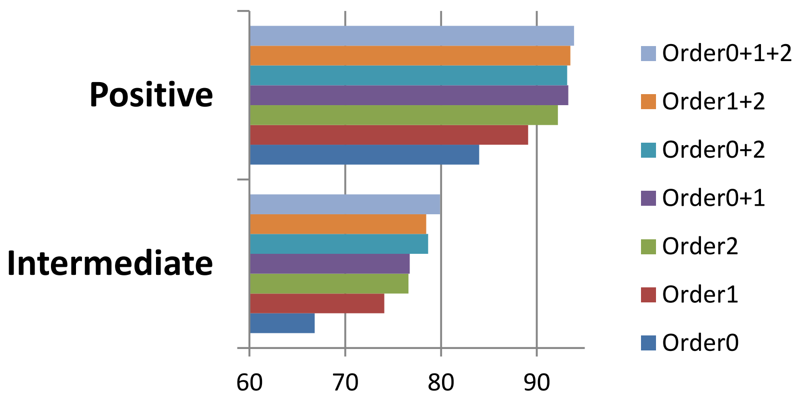
(a) **11 texture types (different materials);**
(b) **4 samples (with different rotated and scaled images) for each texture type**

Experimental results: Hep-2 Cell

(1) Classifier: Linear SVM; (2) Randomly select 600 images (about a half) from each pattern for training; the remainder for test . The average recognition rates of 10 runs:

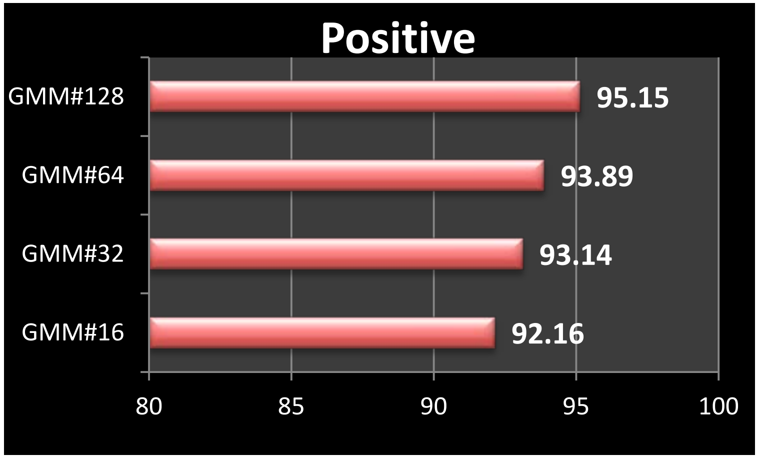
	Positive	Intermediate
LBP	58.87±1.2	37.49±1.78
Our	**93.89**±0.42	**79.89**±0.43

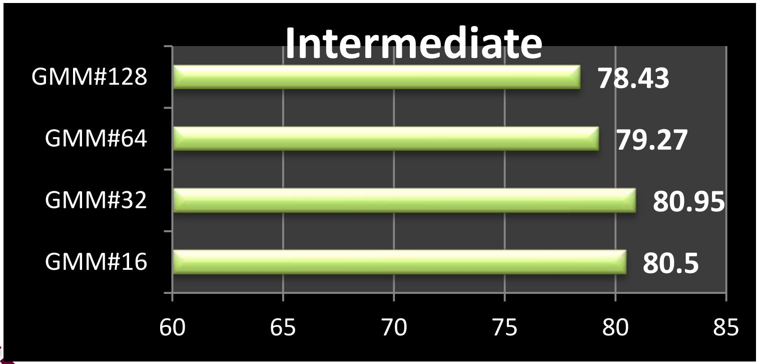
Compared with LBP histogram (64 Gaussian components: all order statistics)



Comparison with different order statistics

Compared results with different component numbers





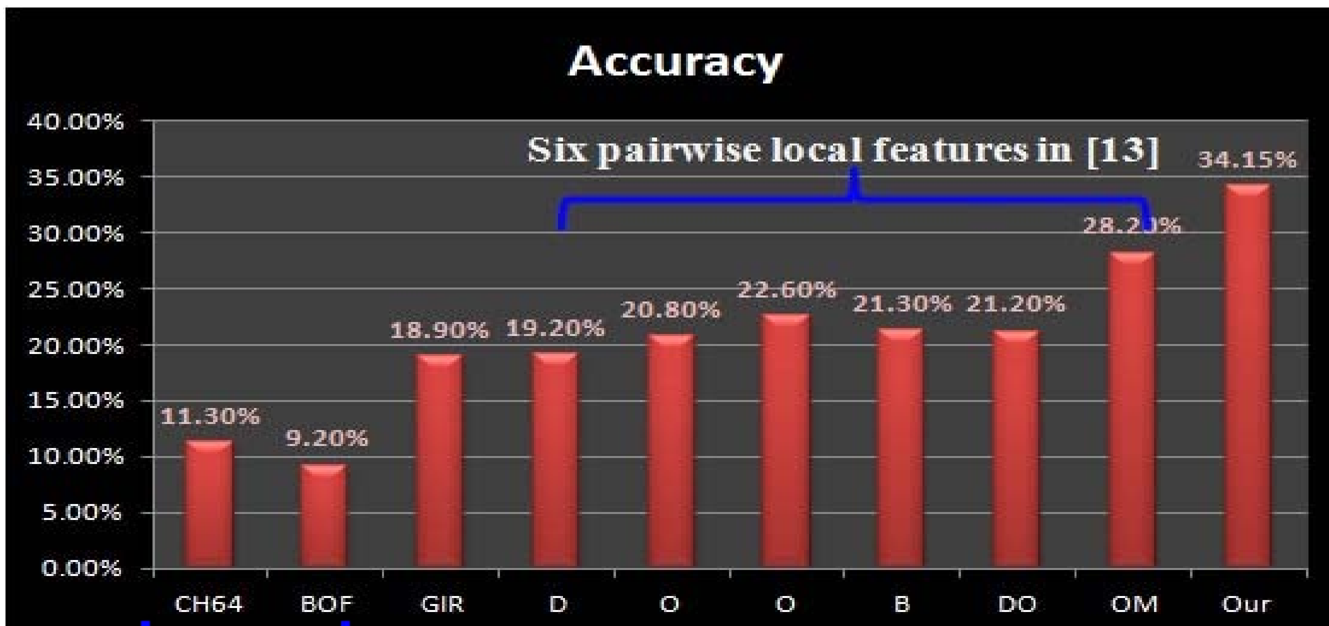
Experimental results: Food images

(1) **Classifier: linear SVM**

(2) **Select the images of two instances for training, and the remainder one for test (same to the experimental setting in the state-of-the-art methods)**

Average recognition rates of 3 runs

The recognition performance compared to the state-the-of-art methods



Baseline evaluations

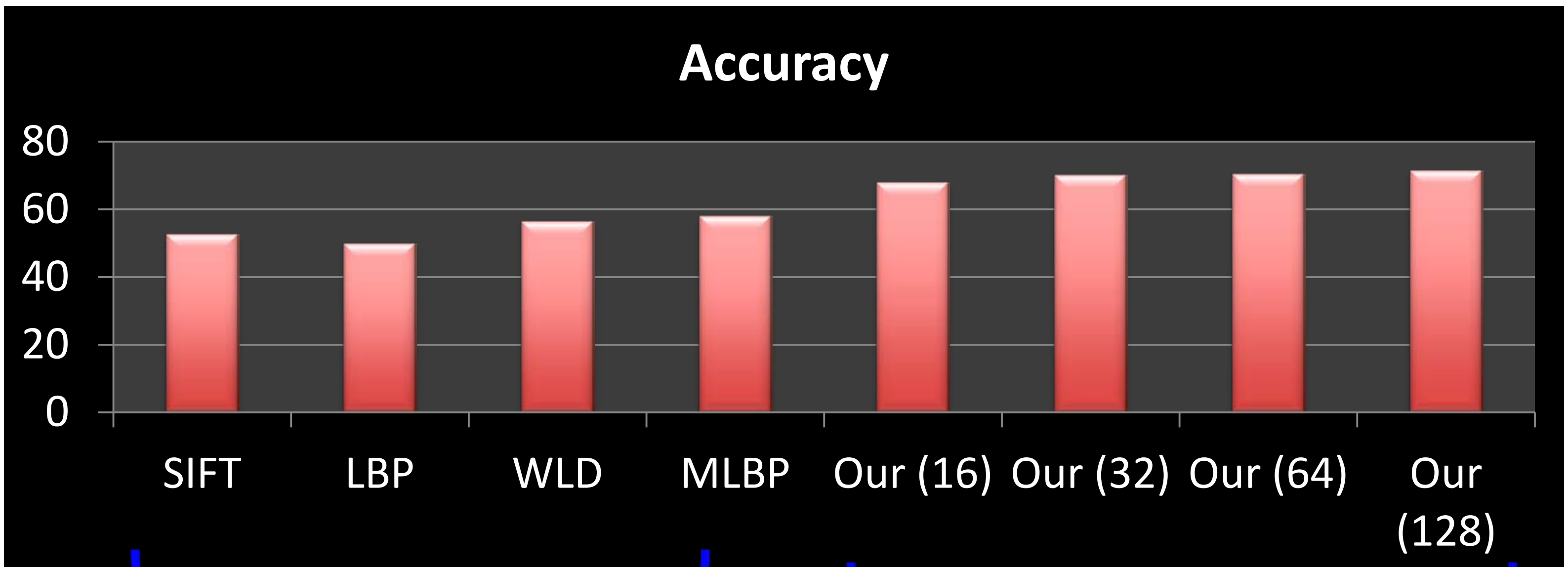
CVPR2010 Paper

Experimental results: Texture database

(1) **Classifier: linear SVM**

(2) **Select the images of three instances for training, and the remainder one for test (same to the experimental setting in the state-of-the-art methods)**

Average recognition rates of 4 runs



PAMI2009: **Low-order statistics (Histogram of the local descriptors)**

High-order statistics (our strategy)

Conclusions

- **Proposed a simple, yet powerful local descriptor for local structure representation.**

- **Adaptively modeling the local descriptor space using a parametric probability process instead of quantization.**

- **Extract not only low-order statistics but also high-order statistics for image representation.**

- **Apply the proposed strategy to three image classification application, and prove the possible promising performance..**

Thank you for your attention!