Laplacian Eigenmaps and Bayesian Clustering Based Layout Pattern Sampling and Its Applications to Hotspot Detection and OPC

Tetsuaki Matsunawa[1], Bei Yu[2] and David Z. Pan[3]

[1]Toshiba Corporation
[2]The Chinese University of Hong Kong
[3]The University of Texas at Austin

Outline

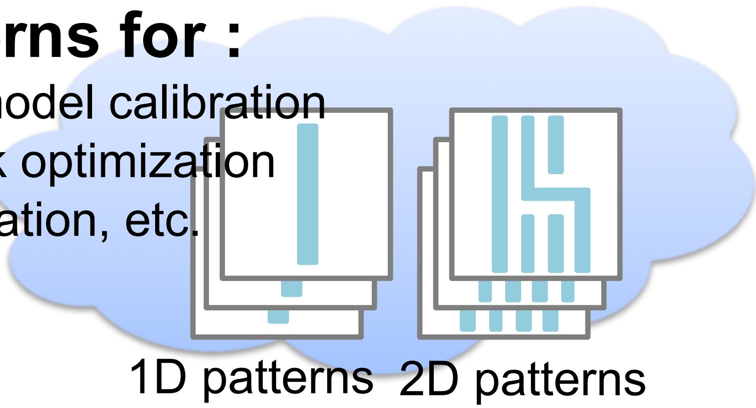
- **Background**
- **Pattern Sampling in Physical Verification**
- **Overall flow**
- **Laplacian Eigenmaps**
- **Bayesian Clustering**
- **Applications**
  - Lithography Hotspot Detection
  - OPC (Optical Proximity Correction)
- **Conclusion**

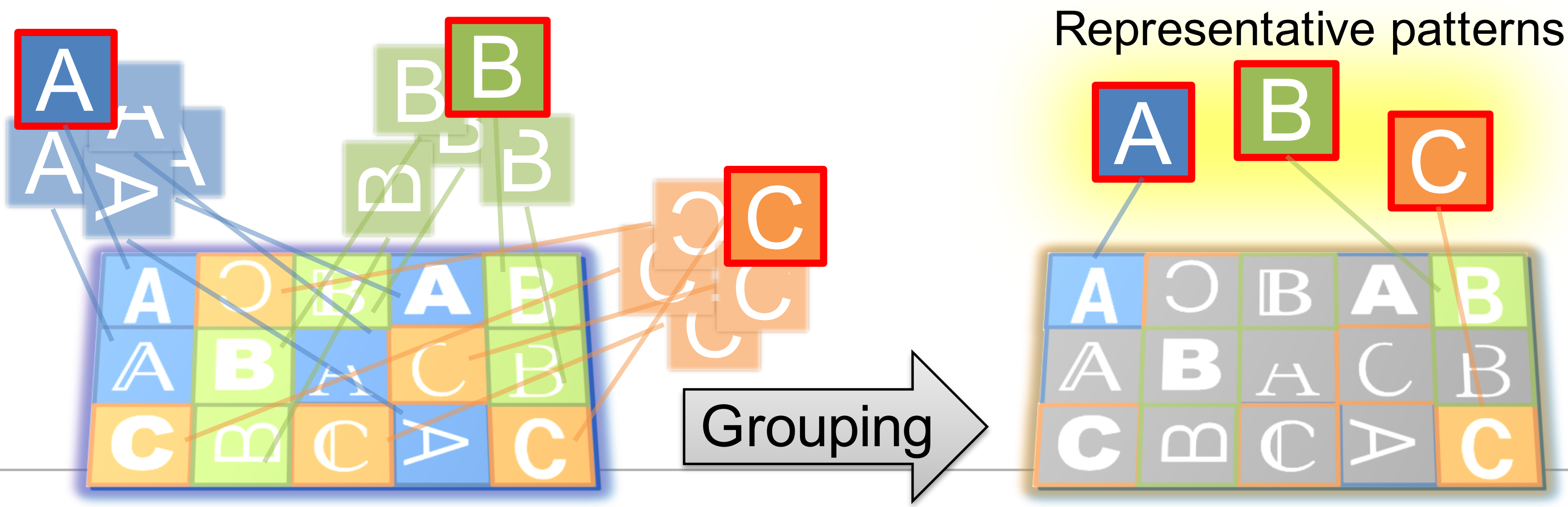
Background

- Issue: Systematic method for pattern sampling is not established
- Goal: Pattern sampling automation for process optimization

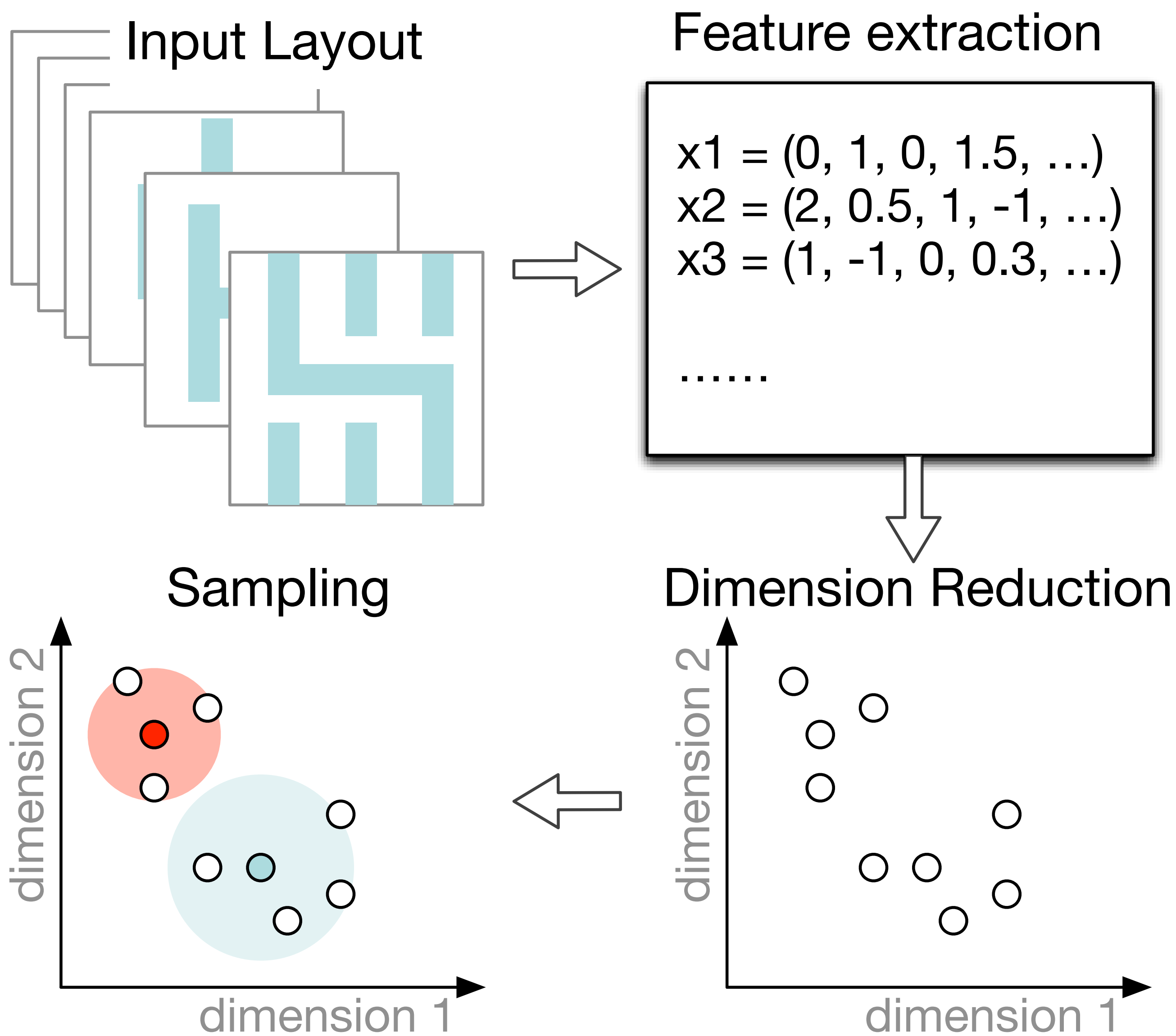
Test patterns for :

Simulation model calibration
Source mask optimization
Wafer verification, etc.





Pattern Sampling



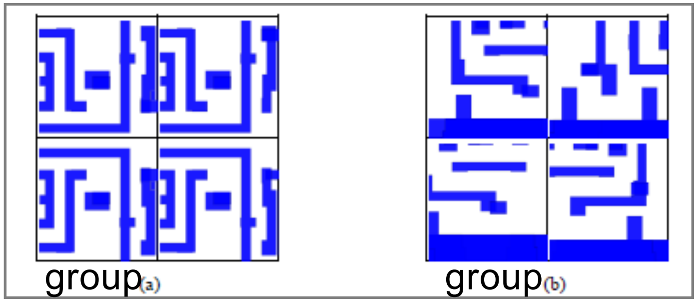
Pattern Sampling in Physical Verification

- Key techniques: Dimension reduction and Clustering

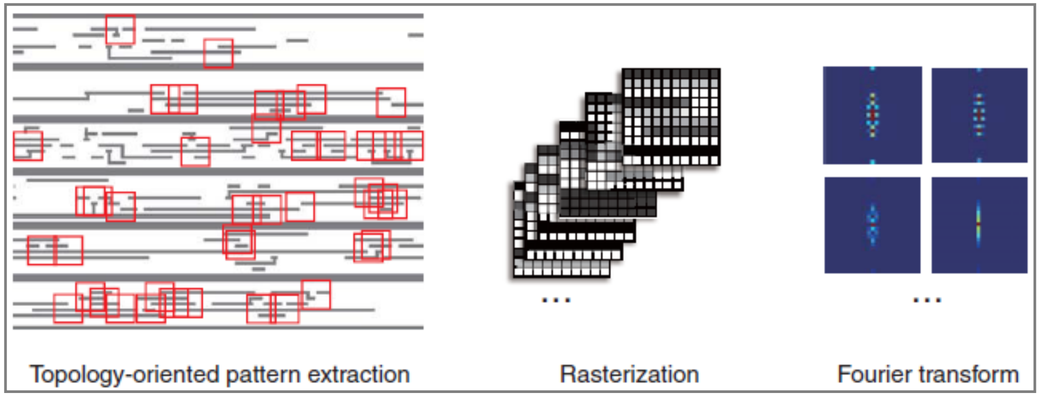
I. W. C. Tam, et al., “Systematic Defect Identification through Layout Snippet Clustering,” ITC, 2010

II. S. Shim, et al., “Synthesis of Lithography Test Patterns through Topology-Oriented Pattern Extraction and Classification,” SPIE, 2014

III. V. Dai, et al., “Systematic Physical Verification with Topological Patterns,” SPIE, 2014



Examples of clustering results[I] W. C. Tam



Classification flow [II] S. Shim

Open Questions

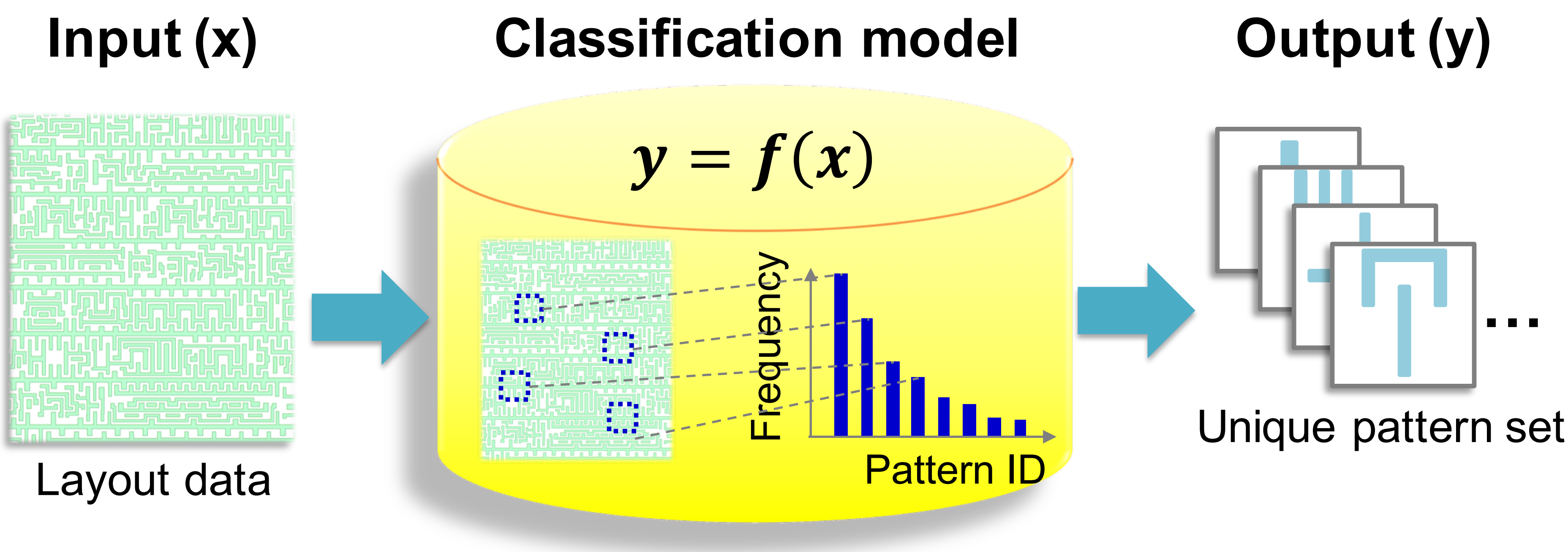
- Undefined similarity
  - A criterion for defining pattern similarity to evaluate essential characteristics in real layouts is unclear
- Manual parameter tuning
  - Most clustering algorithms require several preliminary experiments (total number of clusters)

Laplacian Eigenmaps and Bayesian Clustering

- **We develop**
  - An efficient feature comparison method
    - With nonlinear dimensionality reduction / kernel parameter optimization
  - An automated pattern sampling using Bayesian model based clustering
    - Without manual parameter tuning

Problem formulation: Layout Pattern Sampling

- **Problem:** Given layout data, a classification model is trained to extract representative patterns
- **Goal:** To classify the layout patterns into a set of classes minimizing the Bayes error



Bayes Error (BE)

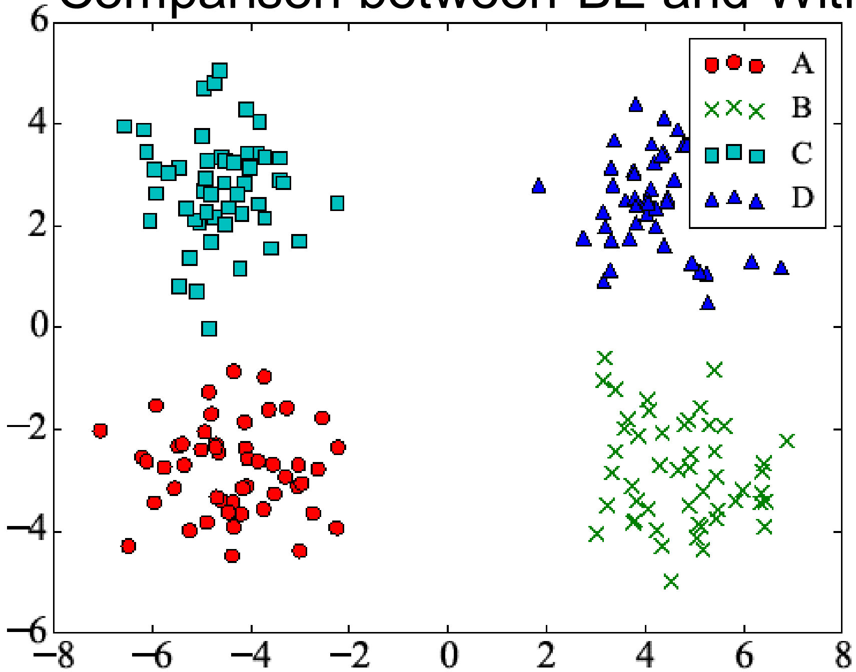
- **To quantify the clustering performance**
  - Define a quality of clustering distributions based on Bayes' theorem

$$BE = \int \min\{1 - p(\omega_k|x)\}p(x)dx$$

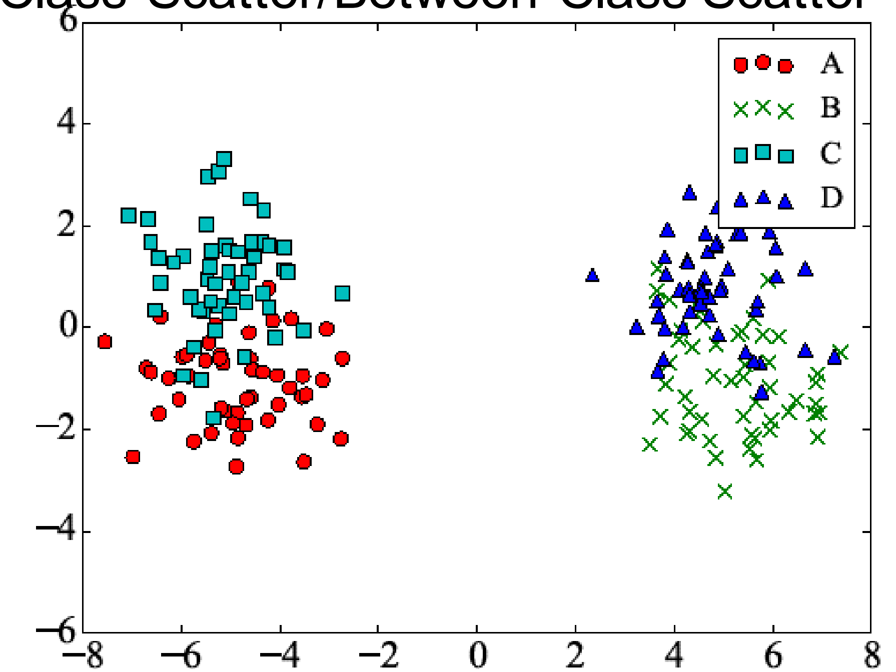
$P(\omega|x)$: conditional probability in class ω

$P(x)$: prior probability of data x

Comparison between BE and Within-Class Scatter/Between-Class Scatter



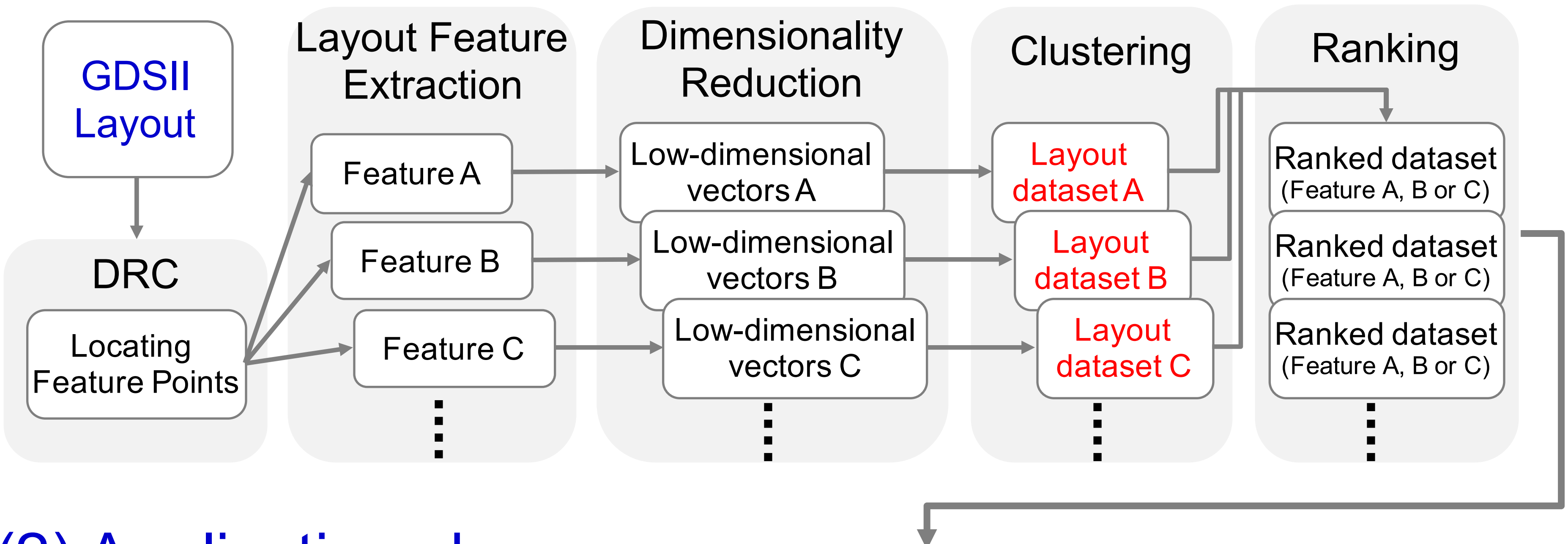
Bayes Error = 0.02, WCS/BCS = 0.07



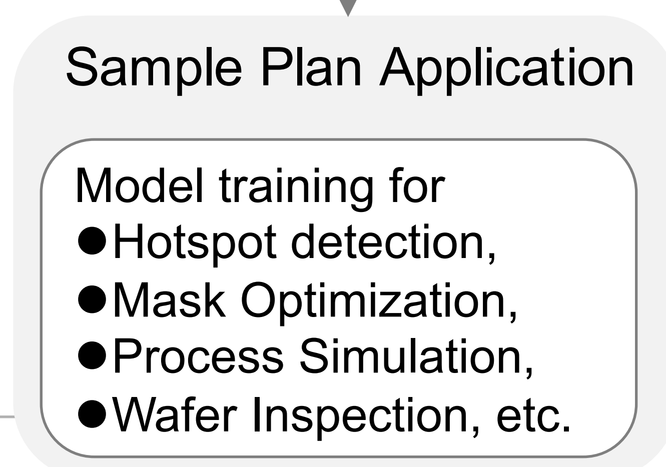
Bayes Error = 1.68, WCS/BCS = 0.07

Overall Flow

(1) Sampling phase



(2) Application phase



Feature Point Generation & Feature Extraction

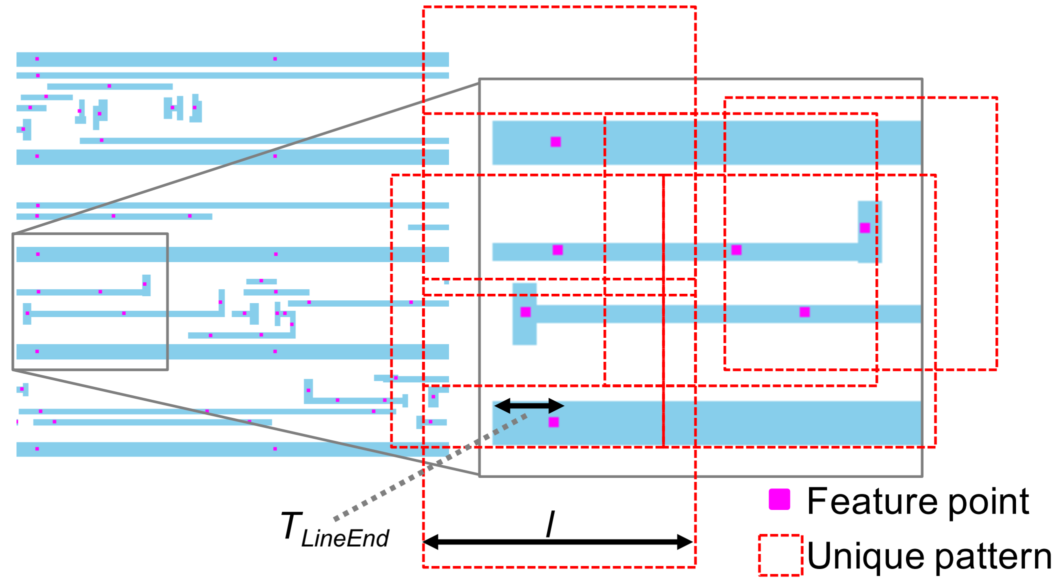
GDS

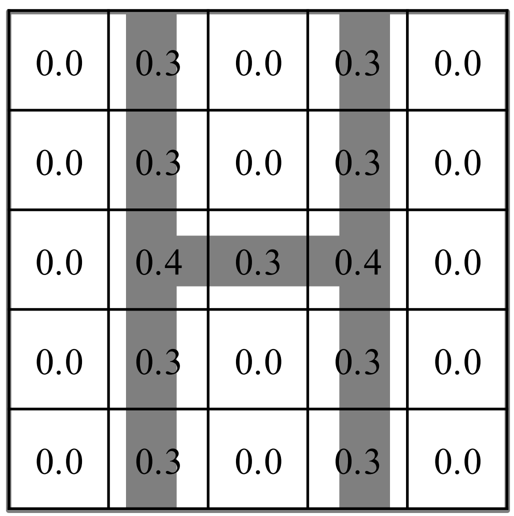
↓

Locating feature points

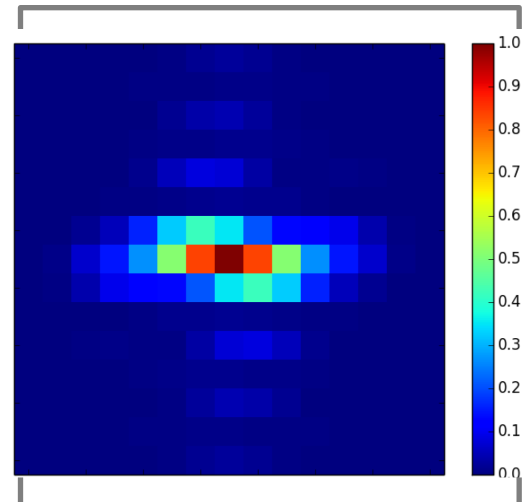
↓

Feature extraction





Density based encoding



Diffraction order distribution

Why dimension reduction and Bayesian clustering?

Required feature comparison for optimal feature selection

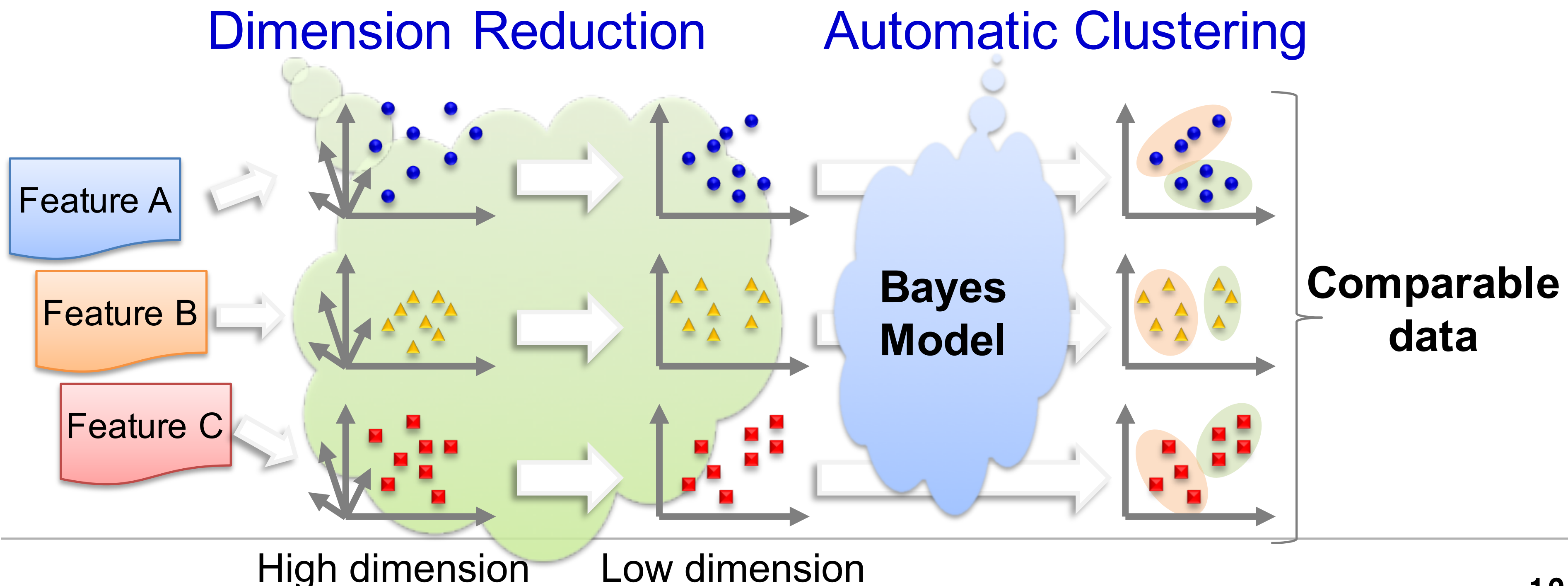
- The optimal characteristics for layout representation vary in different applications

How to compare diverse layout feature types?

- #of dimensions differs with different types of features

Hard to achieve completely automatic clustering

- Hypothetical parameters are required for typical clustering task



Laplacian Eigenmaps

To reduce dimensions while preserving complicated structure

Solve an eigenvalue problem: $L\psi = \gamma D\psi$

Laplacian matrix

$$L = D - W$$

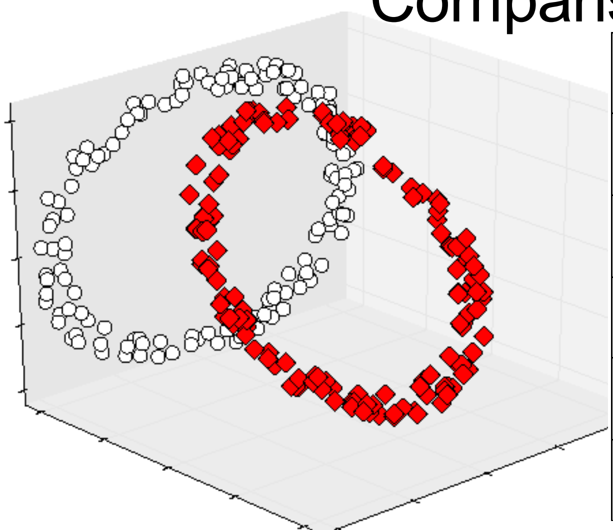
Diagonal matrix

$$D = \mathrm{diag}\left(\sum_{i'=1}^{n} W_{i,i'}\right)$$

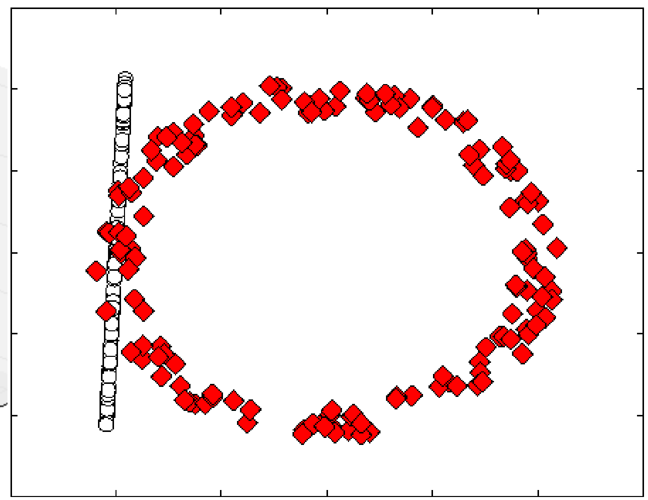
Kernel : k-nearest neighbors

$$W_{i,i'} = \begin{cases} 1 & \text{if } x_i \in kNN(x_{i'}) \\ & \quad or\ x_{i'} \in kNN(x_i) \\ 0 & \text{otherwise} \end{cases}$$

Comparison with linear/nonlinear algorithm

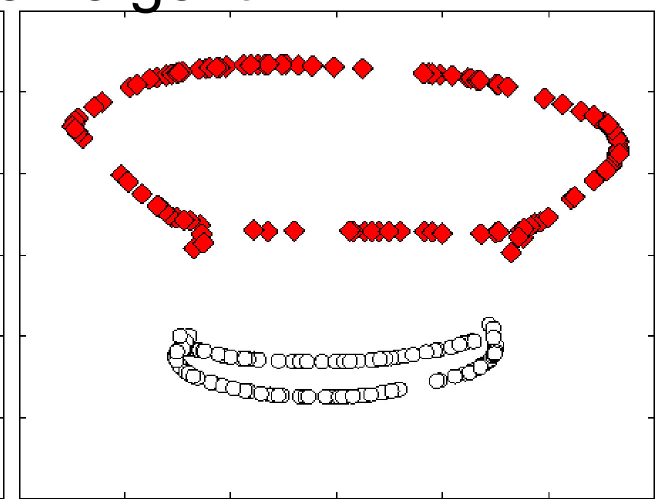


Original data (3D)



Linear(2D)

Principal Component Analysis



Nonlinear(2D)

Laplacian Eigenmaps

Kernel Parameter Optimization

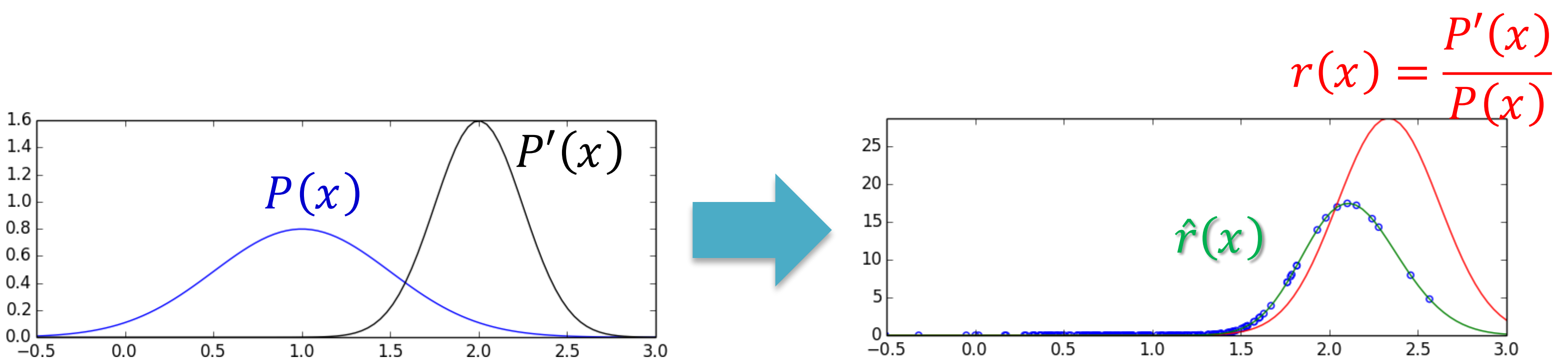
- Optimization through estimating density-ratio $\hat{r}(\mathbf{x}) = \mathbf{w}\boldsymbol{\Phi}(\mathbf{x})$ between given feature vectors $P(x)$ and embedded feature vectors $P'(x)$

$$\max_{w} \sum_{i=1}^{n'} \log\left(w^{\mathrm{T}}\phi(x_i')\right)$$

$$\text{Subject to } \sum_{i=1}^{n} w^{\mathrm{T}}\phi(x_i) = n \text{ and } w \geq 0$$

n': #of test samples
n : #of training samples

This is convex optimization, so repeating gradient ascent and constraint satisfaction converges to global solution



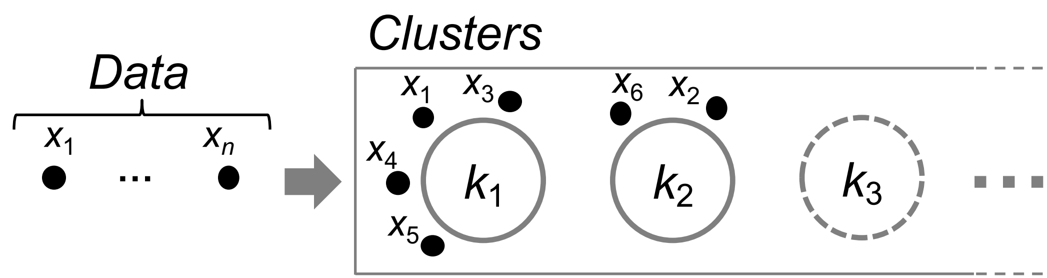
Bayesian Clustering

- Clustering automation without arbitrary parameter tuning
- Bayesian based method: express a parameter distribution as an infinite dimensional distribution

mixture ratio — Gaussian distribution

$$p\big(\mathbf{x}|\alpha, p(\boldsymbol{\theta})\big) = \sum_{k=1}^{\infty} \pi_k \mathcal{N}(\mu_k, \sigma_k)$$

Centroid — Similarity



Prior probability : $\dfrac{4}{\alpha+n-1}$ $\dfrac{2}{\alpha+n-1}$ $\dfrac{\alpha}{\alpha+n-1}$

$$p(z_n = k|x_n, z_1, \dots, z_{n-1}) \propto \begin{cases} p(x_n|k)\dfrac{n_k}{\alpha+n-1} & (k = 1 \cdots K) \\ p(x_n|k^{new})\dfrac{\alpha}{\alpha+n-1} & (k = K+1) \end{cases}$$

Experiments

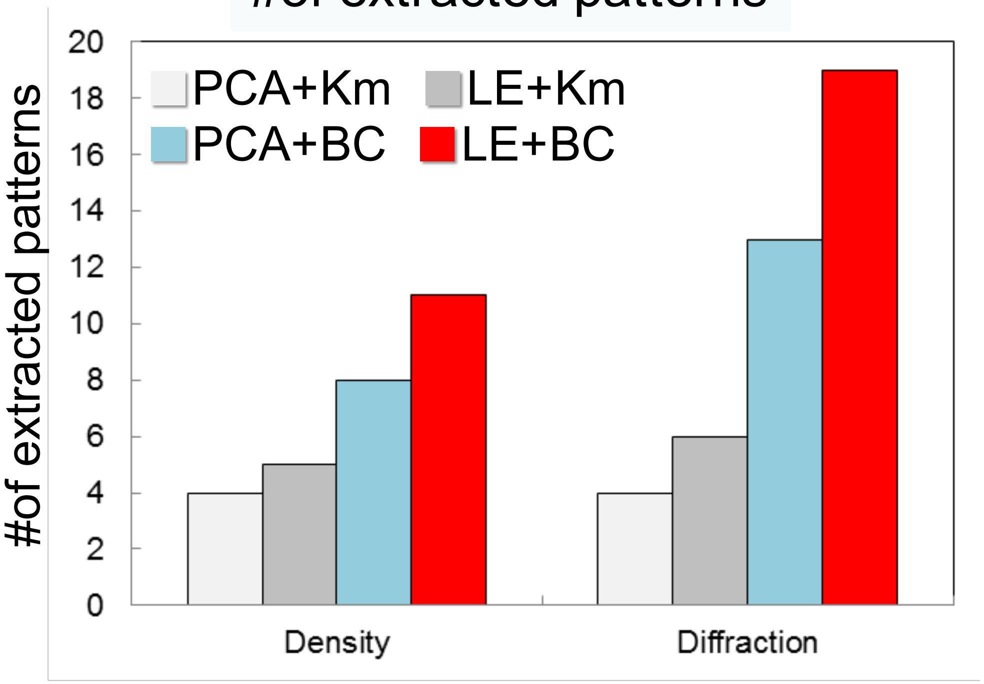
- **Pattern sampling**
  - Comparison of conventional methods
    - Dimensionality reduction
      - Principal Component Analysis (PCA) vs. Laplacian Eigenmaps (LE)
    - Clustering
      - K-means (Km) vs. Bayesian clustering (BC)

- **Applications to**
  - Lithography Hotspot Detection
  - OPC

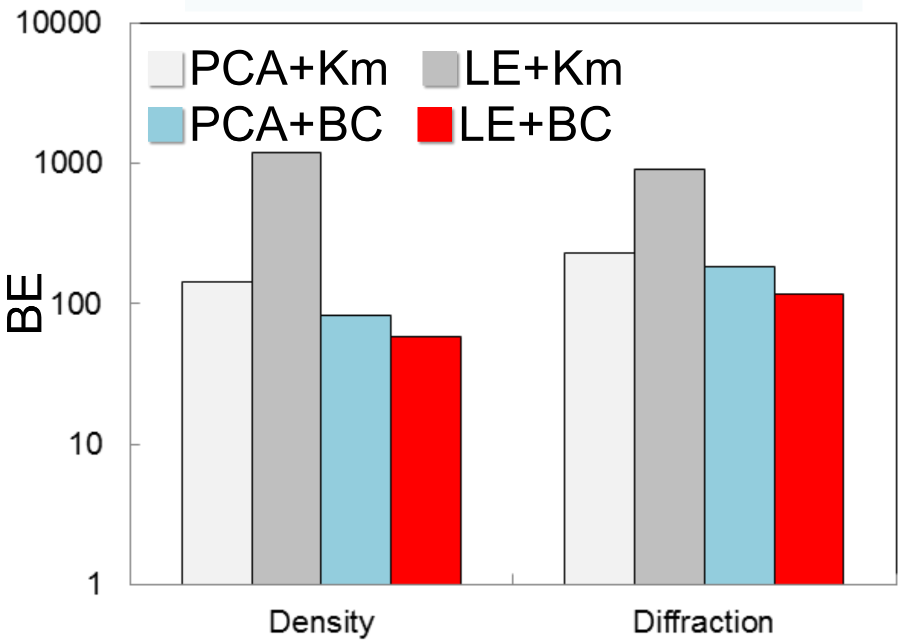
Effectiveness of Pattern Sampling

- Representative patterns could be automatically selected

Clustering results:
#of extracted patterns



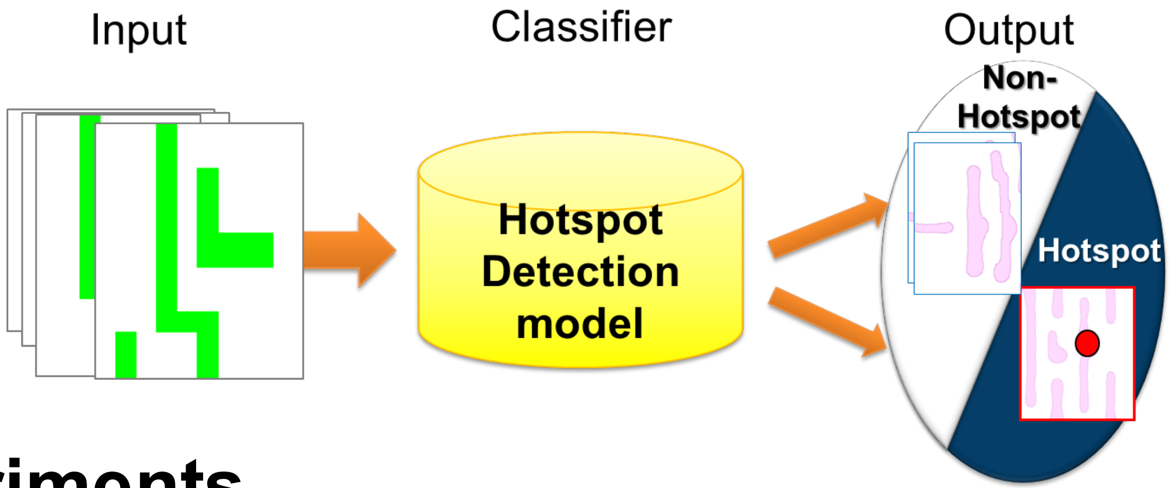
Misclassification error rate:
Bayes Error



Ratio: PCA+Km: 1.0 LE+Km: 5.6
PCA+BC: 0.7 LE+BC: 0.5

Application to Lithography Hotspot Detection

- **To detect hotspot in short runtime**

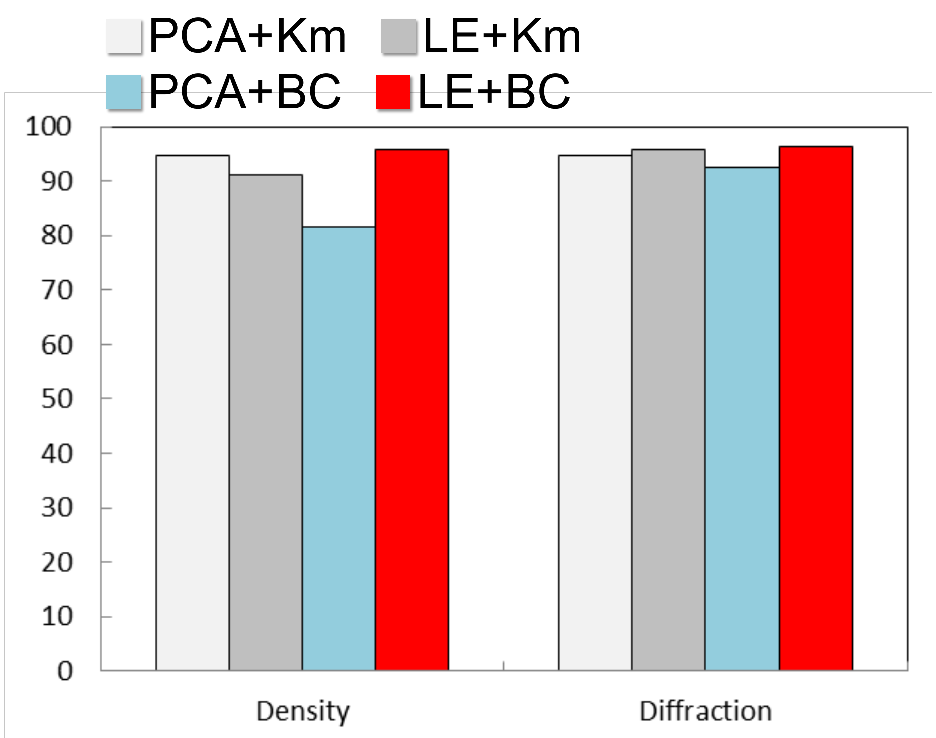


- **Experiments**
  - Detection model training with different patterns
    - PCA+Km, LE+Km, PCA+BC, LE+BC
    - Learning algorithm is fixed to Adaptive Boosting (AdaBoost)
  - Metrics: detection accuracy and false alarm

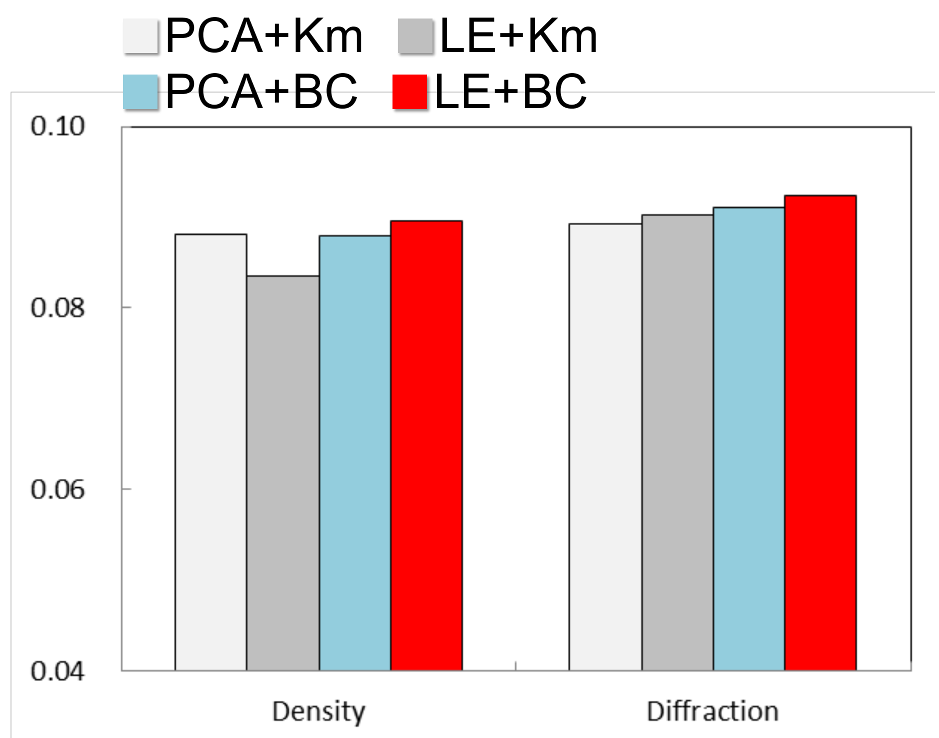
Effectiveness of Hotspot Detection

- Comparison with conventional clustering method
- Result: Proposed framework achieved the best false-alarm

Detection accuracy:
#correctly detected hotspots / #total hotspots

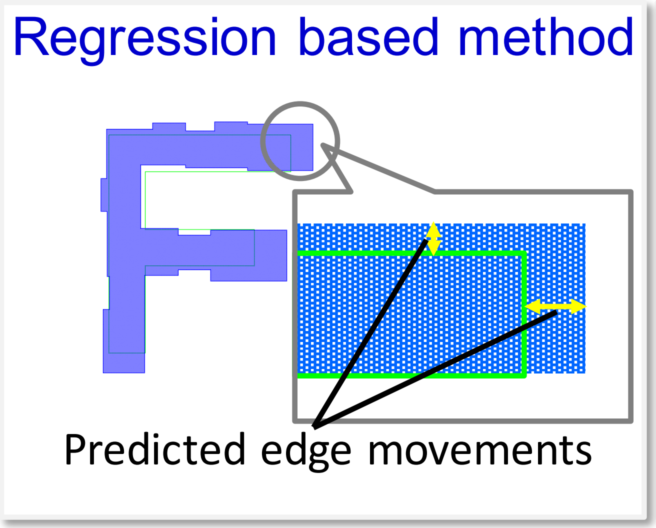


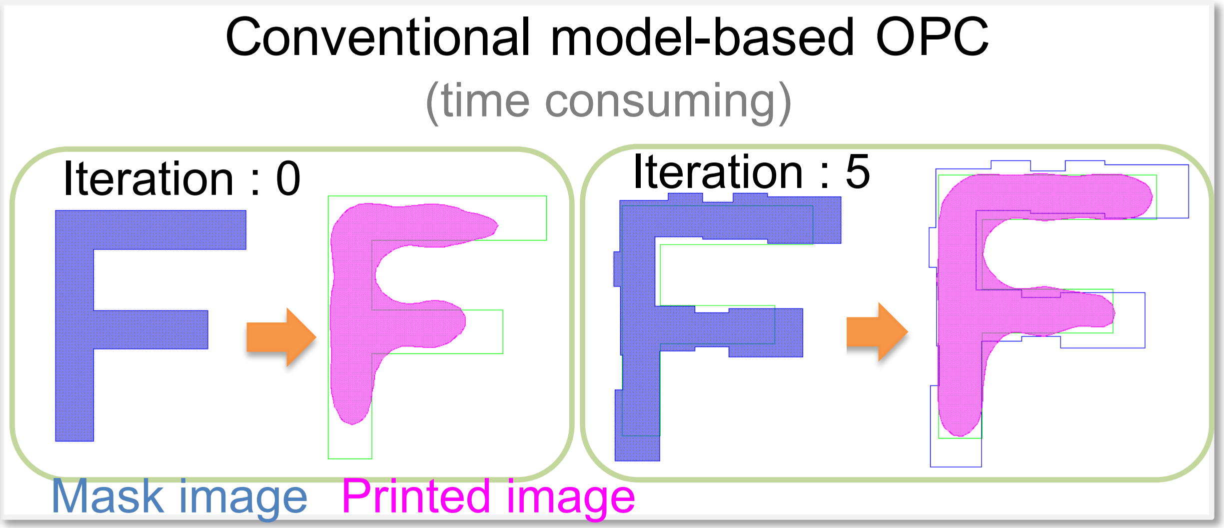
False alarm:
#correctly detected hotspots / #falsely detected hotspots



Application to Regression-based OPC

- **To predict edge movements in short runtime**



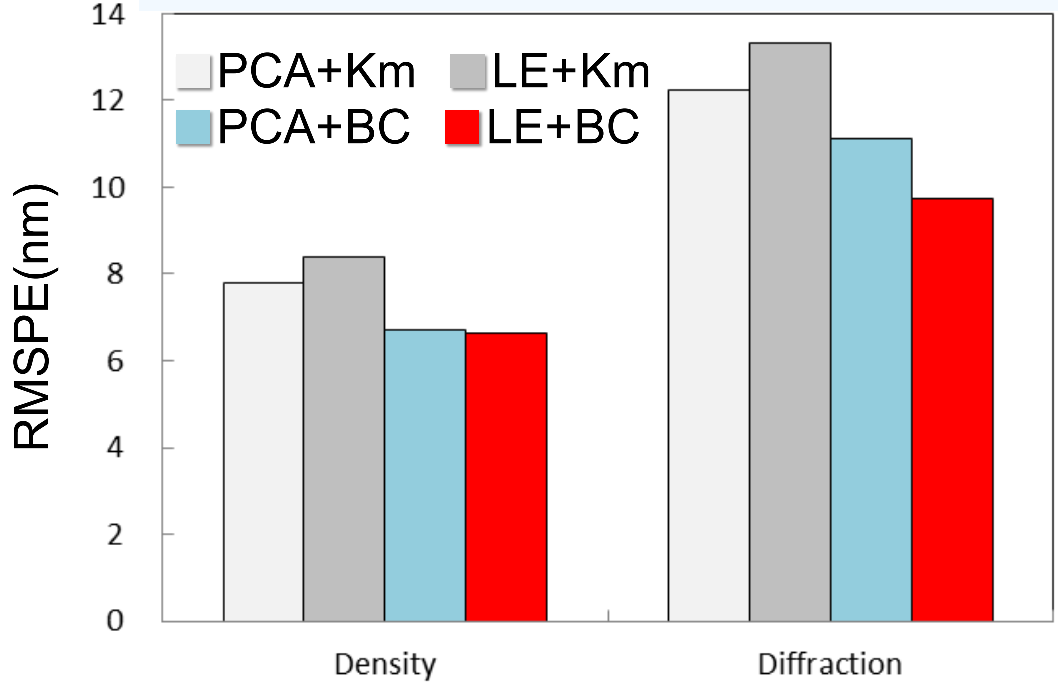


- **Experiments**
  - Prediction model training with different patterns
    - PCA+Km, LE+Km, PCA+BC, LE+BC
    - Learning algorithm is fixed to Linear regression
  - Metric: RMSPE (Root Mean Square Prediction Error)

Effectiveness of OPC regression

- Proposed framework achieved the best prediction accuracy

Prediction accuracy:
RMSPE: Root mean square prediction error



Ratio: PCA+Km: 1.0 LE+Km: 1.1
PCA+BC: 0.9 LE+BC: 0.8

Conclusion

- We have introduced a new method to sample unique patterns.
  - By applying our dimension reduction technique, dimensionality- and type-independent layout feature can be used in accordance with applications.
  - The Bayesian clustering is able to classify layout data without manual parameter tuning.
  - The experimental results show that our proposed method can effectively sample layout patterns that represent characteristics of whole chip layout.