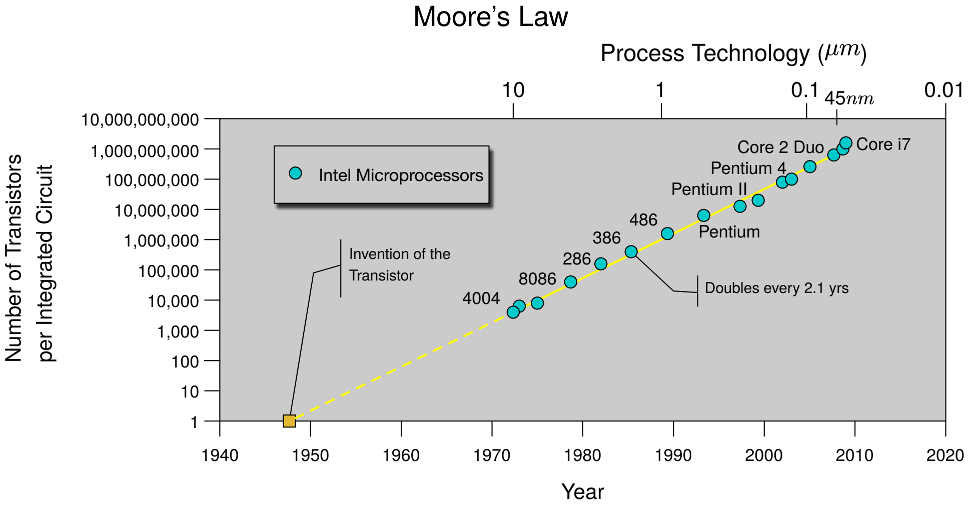
SRAF Insertion via Supervised Dictionary Learning

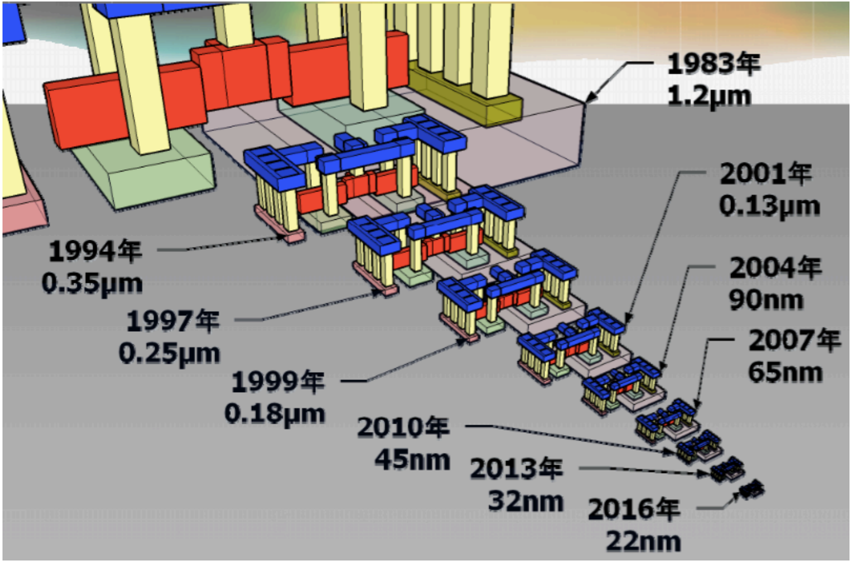
Hao Geng[1], Haoyu Yang[1], Yuzhe Ma[1], Joydeep Mitra[2], Bei Yu[1]

[1]The Chinese University of Hong Kong
[2]Cadence Inc.

Moore’s Law to Extreme Scaling

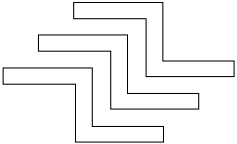


Nanometer Era of Manufacturing: An Inverter Example

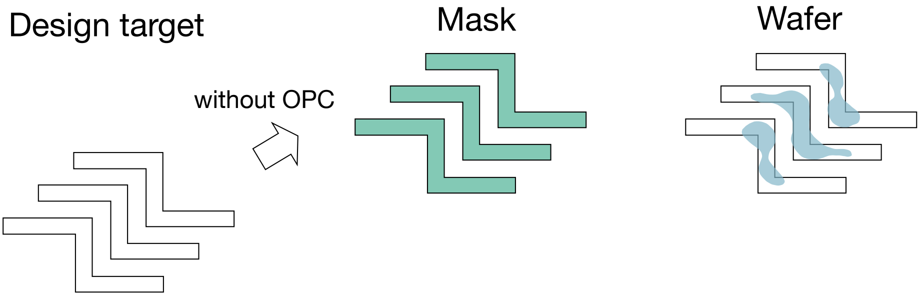


Optical Proximity Correction (OPC)

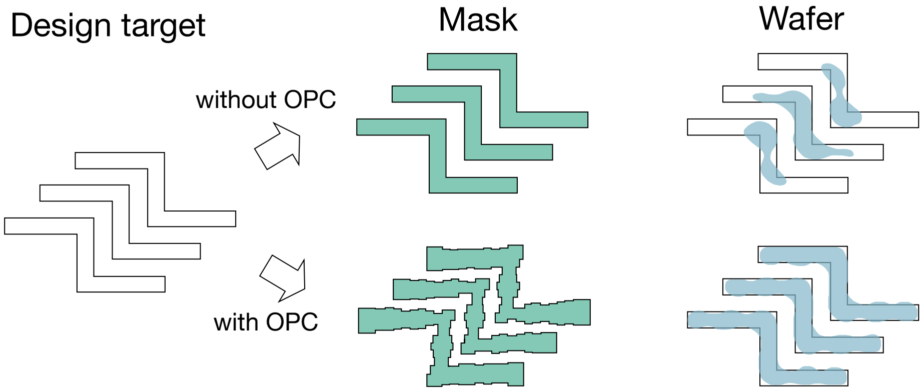
Design target



Optical Proximity Correction (OPC)

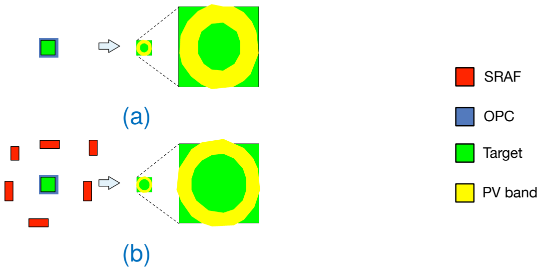


Optical Proximity Correction (OPC)



What is SRAF?

- Patterns deliver light to target features without printing themselves
- Make isolated features more dense
- Improve the robustness of the target patterns
- Rule-based [Jun+,SPIE'15], Model-based [Shang+,Mentor'05], Machine learning model-based [Xu+,ISPD'16]



(a) Printing with OPC only (2688 nm^2 PV band area); (b) Printing with both OPC and SRAF (2318 nm^2 PV band area).

Outline

Outline

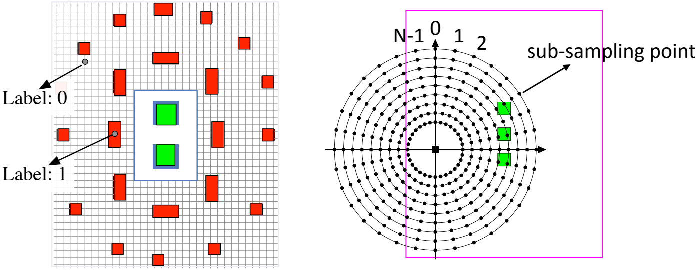
Supervised Feature Revision

SRAF Insertion

Experimental Results

Concentric Circle Area Sampling

- Initial feature extraction method in SRAF generation



(a) (b)

(a) SRAF label; (b) CCAS feature extraction method in machine learning model-based SRAF generation.

Introduction to Dictionary Learning

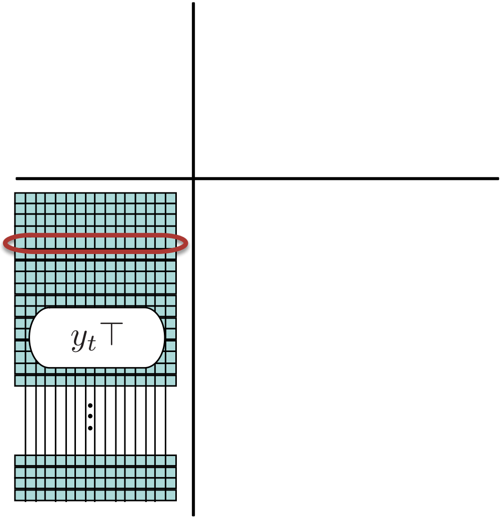
Overview

Originally, the dictionary learning model is composed of two parts. One is **sparse coding** and the other is **dictionary constructing**. The joint objective function with respect to $\mathbf{D}$ and $\mathbf{x}$ is below.

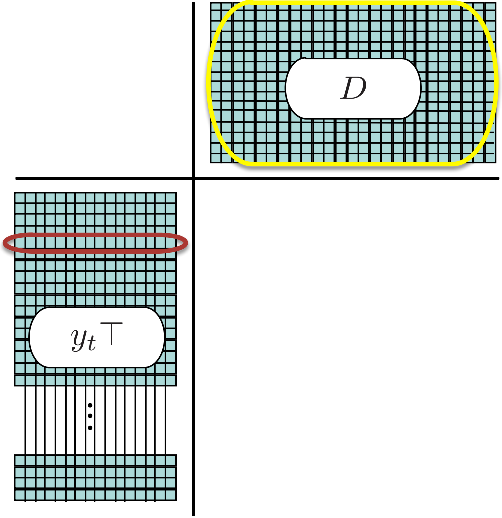
$$\min_{\mathbf{x},\mathbf{D}} \frac{1}{N} \sum_{t=1}^{N} \{ \frac{1}{2} \|\mathbf{y}_t - \mathbf{D}\mathbf{x}_t\|_2^2 + \lambda \|\mathbf{x}_t\|_p \}, \tag{1}$$

- $\mathbf{y}_t \in \mathbb{R}^{(n)}$: the t-th input data vector
- $D = \{\mathbf{d}_j\}_{j=1}^{s}, \mathbf{d}_j \in \mathbb{R}^{(n)}$: the dictionary where every column is called an atom.
- $\mathbf{x}_t \in \mathbb{R}^{(s)}$: the sparse code
- λ: hyper-parameter
- p: the norm type of penalty term, e.g. l_1 norm

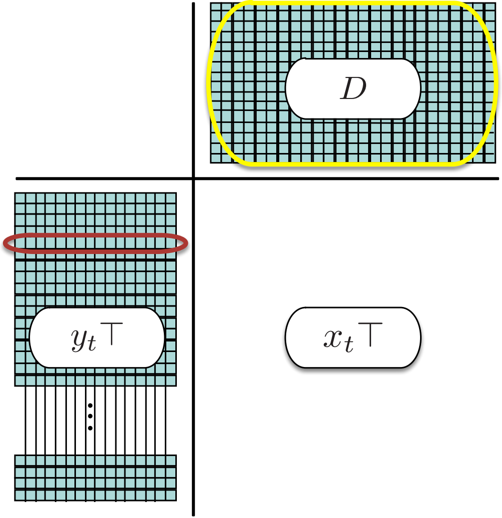
The Illustration for Dictionary Learning



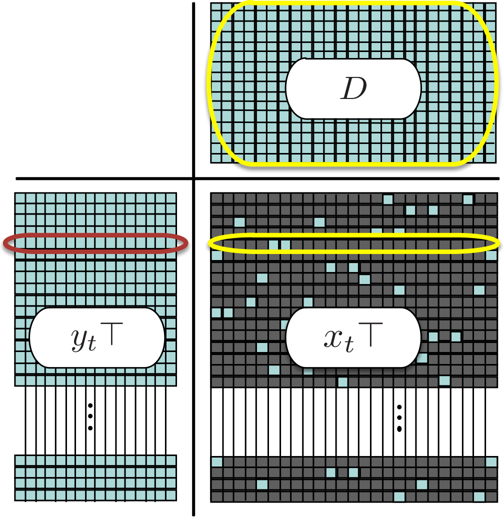
The Illustration for Dictionary Learning



The Illustration for Dictionary Learning



The Illustration for Dictionary Learning



Online Learning Framework

Sparse Coding

The subproblem with $\mathbf{D}$ fixed is convex. The objective function for sparse coding of i-th training data vector in memory is

$$\mathbf{x}_t \triangleq \arg\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y}_t - \mathbf{D}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_p. \tag{2}$$

Solver Details

- $p = 0$: l_0 norm and NP-hard [Mallat+,TIP'93], [Pati+,ACSSC'93]
- $p = 1$: LASSO problem [Friedman+,JSS'10], [Beck+,SIIMS'09]

Online Learning Framework

Dictionary Constructing

The subproblem with $\mathbf{x}$ fixed is also convex. The objective function for dictionary constructing is

$$\mathbf{D} \triangleq \arg\min_{\mathbf{D}} \frac{1}{N} \sum_{t=1}^{N} \frac{1}{2} \|\mathbf{y}_t - \mathbf{D}\mathbf{x}_t\|_2^2 + \lambda \|\mathbf{x}_t\|_p. \tag{3}$$

Solver Details

- Block coordinate descent method with warm start
- Introducing two auxiliary variables $\mathbf{B}$ and $\mathbf{C}$ to speed up convergence rate
- Sequentially updating atoms in a dictionary $\mathbf{D}$

$$\vec{B}_t \leftarrow \frac{t-1}{t} \vec{B}_{t-1} + \frac{1}{t} \vec{y}_t \vec{x}_t^\top, \tag{4}$$

$$\vec{C}_t \leftarrow \frac{t-1}{t} \vec{C}_{t-1} + \frac{1}{t} \vec{x}_t \vec{x}_t^\top. \tag{5}$$

Further Exploration: Supervised Dictionary Learning

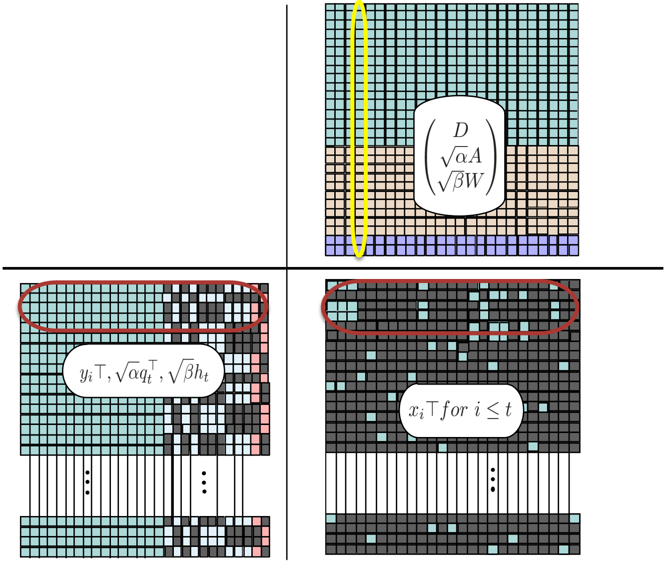
Exploring Latent Label Information

$$\min_{\mathbf{x},\mathbf{D},\mathbf{A}} \frac{1}{N}\sum_{t=1}^{N}\{\frac{1}{2}\left\|\left(\mathbf{y}_t^\top, \sqrt{\alpha}\mathbf{q}_t^\top\right)^\top - \begin{pmatrix}\mathbf{D}\\ \sqrt{\alpha}\mathbf{A}\end{pmatrix}\mathbf{x}_t\right\|_2^2 + \lambda\|\mathbf{x}_t\|_p\}. \tag{6}$$

Exploiting both Latent and Direct Label Information

$$\min_{\mathbf{x},\mathbf{D},\mathbf{A},\mathbf{W}} \frac{1}{N}\sum_{t=1}^{N}\{\frac{1}{2}\left\|\left(\mathbf{y}_t^\top, \sqrt{\alpha}\mathbf{q}_t^\top, \sqrt{\beta}h_t\right)^\top - \begin{pmatrix}\mathbf{D}\\ \sqrt{\alpha}\mathbf{A}\\ \sqrt{\beta}\mathbf{W}\end{pmatrix}\mathbf{x}_t\right\|_2^2 + \lambda\|\mathbf{x}_t\|_p\}. \tag{7}$$

The Illustration for Supervised Online Dictionary Learning



Outline

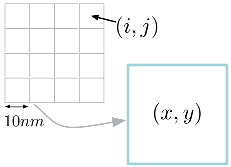
SRAF Insertion

Preliminary Work

- SRAF probability learning for each grid: Logistic regression
- SRAF grid model construction: Merging

$$c(x,y) = \begin{cases} \sum_{(i,j)\in(x,y)} p(i,j), & \text{if } \exists\ p(i,j) \geq \text{threshold}, \\ -1, & \text{if all } p(i,j) < \text{threshold}. \end{cases} \tag{8}$$

- $p(i,j)$: the probability of a grid with index (i,j)
- $c(x,y)$: the summed probability value of merged grid with index (x,y)



SRAF grid model construction.

SRAF Insertion via ILP

$$\max_{a(x,y)} \sum_{x,y} c(x,y) \cdot a(x,y) \tag{9a}$$

$$\text{s.t.} \quad a(x,y) + a(x-1,y-1) \le 1, \quad \forall(x,y), \tag{9b}$$

$$a(x,y) + a(x-1,y+1) \le 1, \quad \forall(x,y), \tag{9c}$$

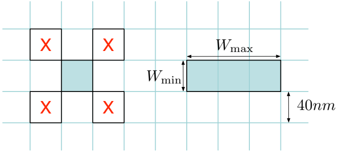
$$a(x,y) + a(x+1,y-1) \le 1, \quad \forall(x,y), \tag{9d}$$

$$a(x,y) + a(x+1,y+1) \le 1, \quad \forall(x,y), \tag{9e}$$

$$a(x,y) + a(x,y+1) + x(x,y+2) + a(x,y+3) \le 3, \quad \forall(x,y), \tag{9f}$$

$$a(x,y) + a(x+1,y) + x(x+2,y) + a(x+3,y) \le 3, \quad \forall(x,y), \tag{9g}$$

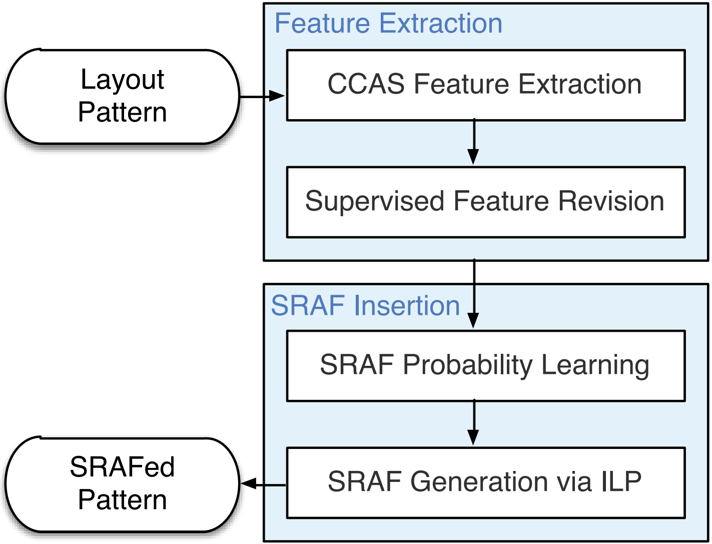
$$a(x,y) \in \{0,1\}, \quad \forall(x,y). \tag{9h}$$



SRAF insertion design rule under the grid model.

Outline

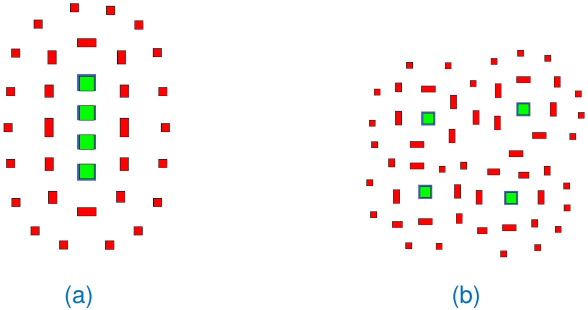
The Overall Flow



Experimental Bed

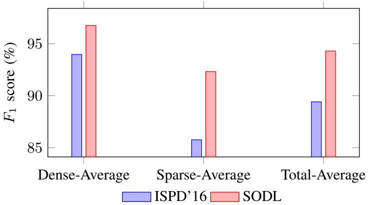
Benchmark Set

- The same benchmark set as applied in [Xu+,ISPD'16]
- 8 dense layouts and 10 sparse layouts with contacts sized 70nm
- 70nm spacing for dense and $\geq$ 70nm spacing for sparse layouts

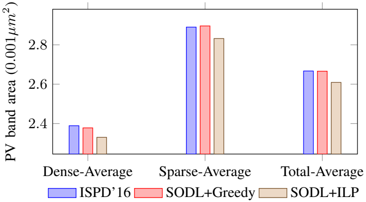


(a) Dense layout with golden SRAFs; (b) Sparse layout with golden SRAFs.

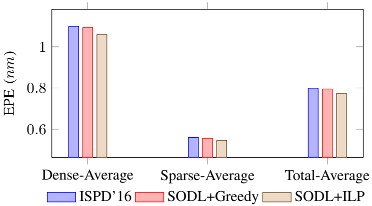
Results



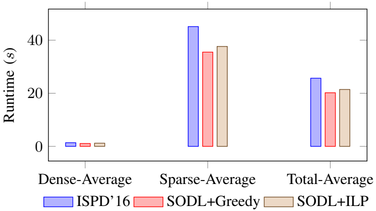
(a)



(b)



(c)



(d)

Lithographic performance comparisons with a state-of-the-art machine learning based SRAF insertion tool.

Conclusion

Summary:

- First introduced the concept of dictionary learning into the layout feature extraction stage
  and further proposed a supervised online dictionary learning algorithm.
- ILP for SRAF generation in a global view.
- Boost F_1 score and enhance lithographic performance with less time overhead.

Future Work:

- Speed up SRAF insertion process
- Consider more SRAF design rules into ILP
- ...