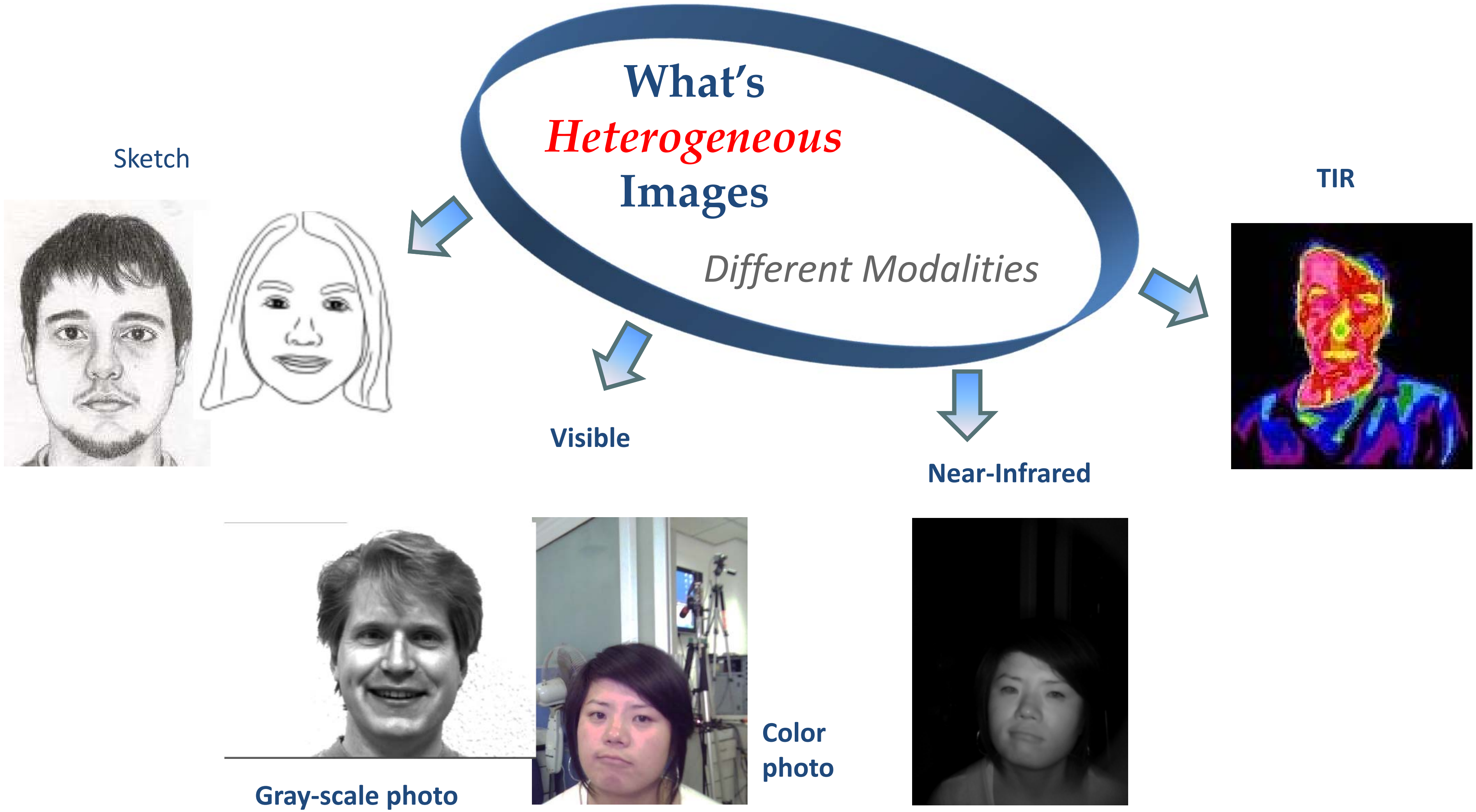
Heterogeneous Facial Image Synthesis

Nannan Wang, Dacheng Tao, Xinbo Gao, Xuelong Li, Jie Li

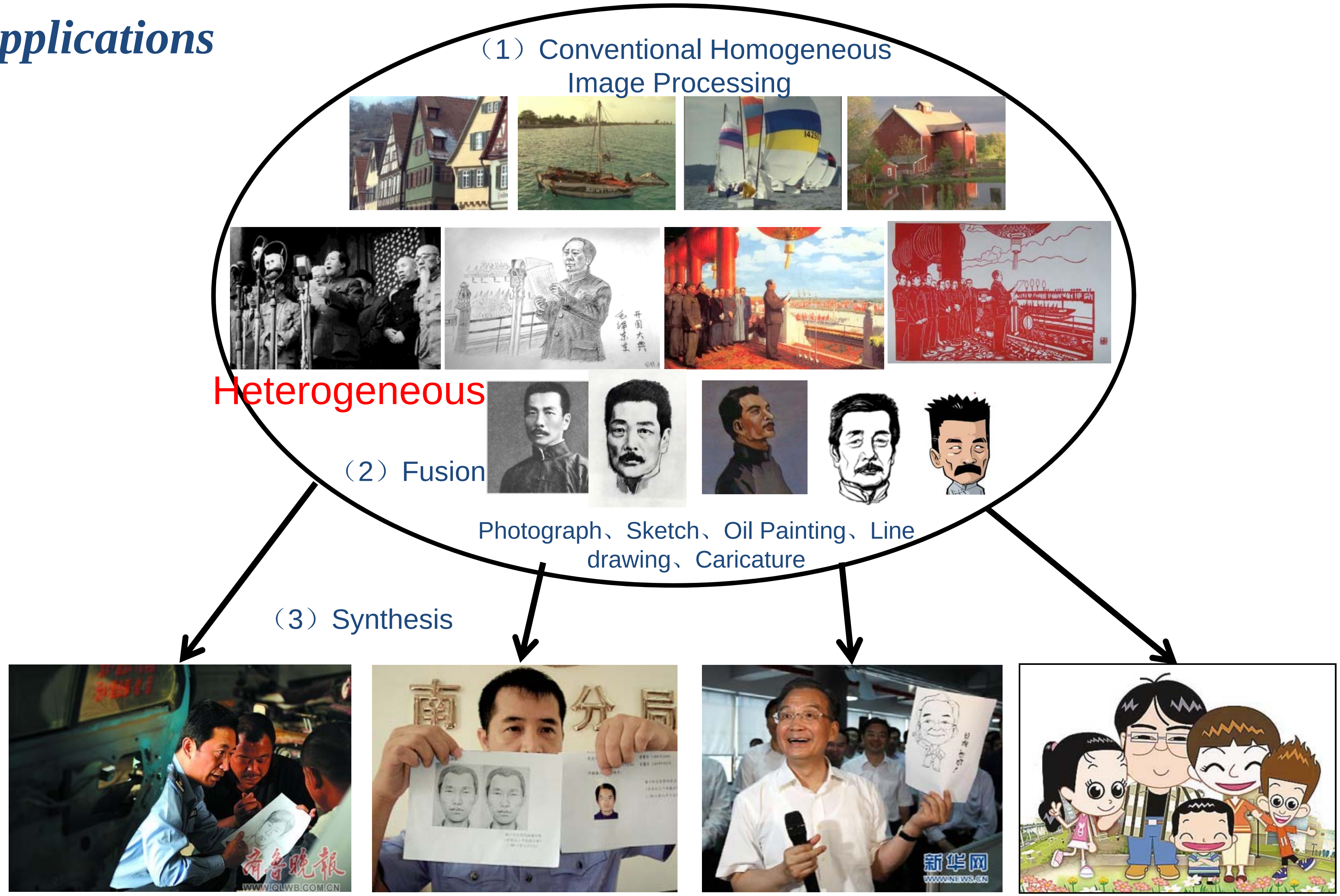

What's
Heterogeneous
Images
Different Modalities
Sketch
TIR
Visible
Near-Infrared
Color
photo
Gray-scale photo

Contents

- **Heterogeneous Facial Image *Synthesis***
  - ***Resolution:*** Low ←→ High (*Scaling*, *Super-Resolution*)
  - ***Color:*** Color ←→ gray-scale (*color2gray*, *pseudo-color*)
  - ***Modality:***
    - Near infrared image ←→ Visible image
    - Near infrared image ←→ Thermal infrared image
    - Thermal infrared image ←→ Visible image
    - CT image ←→ MRI
    - **Photo ←→ Sketch (The focus of this tutorial)**
    - ***Traditional Chinese Painting ←→ Oil Painting***

Photo vs. Sketch

Applications

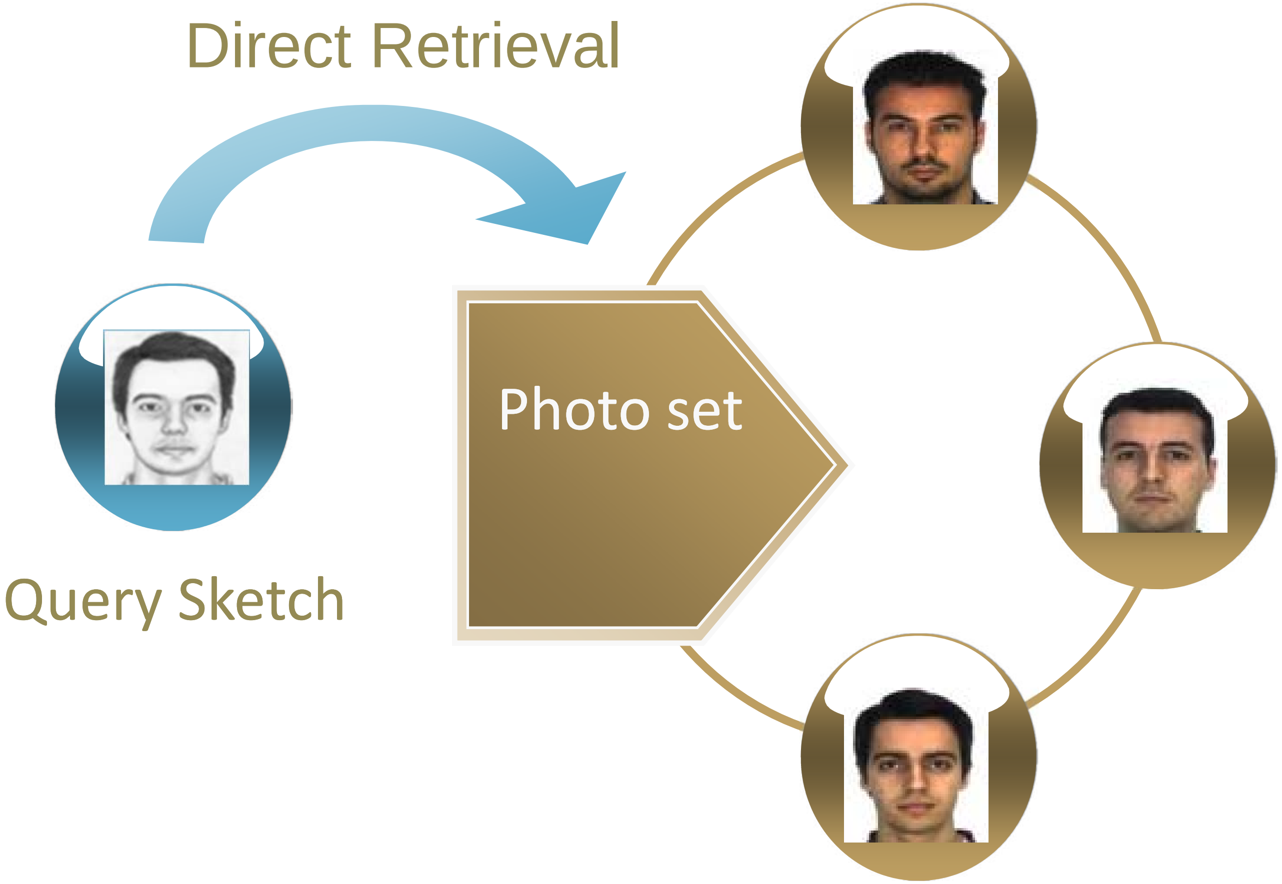


Search the criminal

Entertainment

OUTLINE

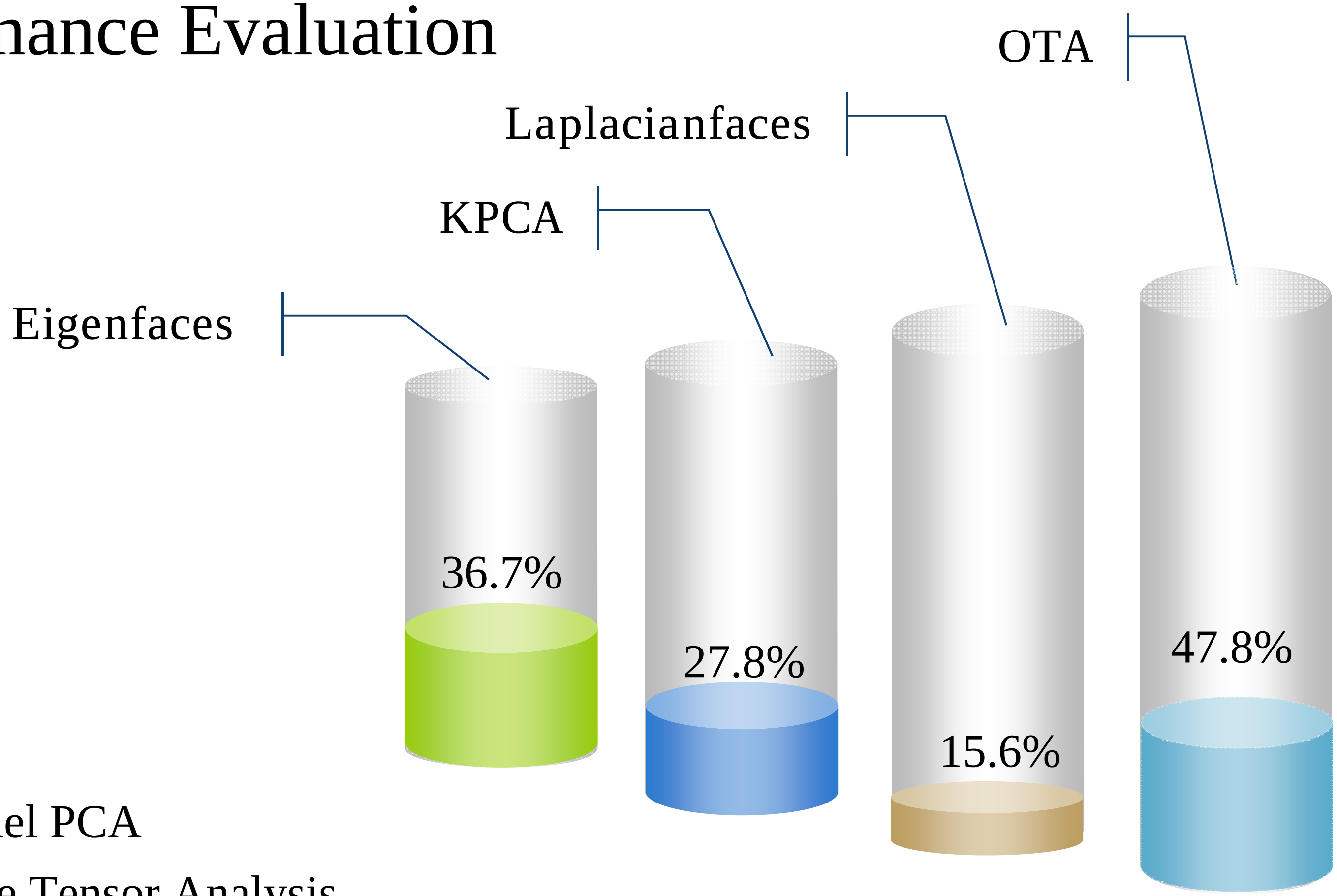
- Motivation & Introduction
- Subspace Learning-based Methods
- Sparse Representation-based Methods
- Bayesian Inference-based Methods
- Conclusions



Identify a person by a sketch

All ≤ 50 !

Performance Evaluation



KPCA: Kernel PCA

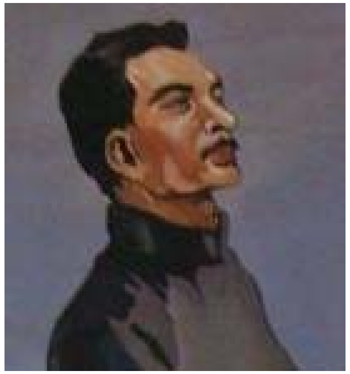
OTA: Offline Tensor Analysis

Why direct retrieval failed?

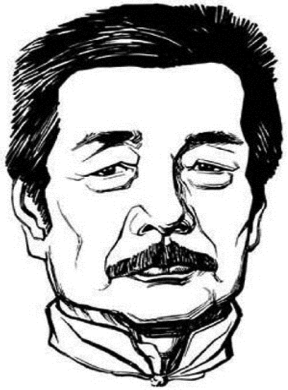
Why direct retrieval failed?

Visible image

Sketch Portrait



Oil Painting



Line Drawing

Caricature

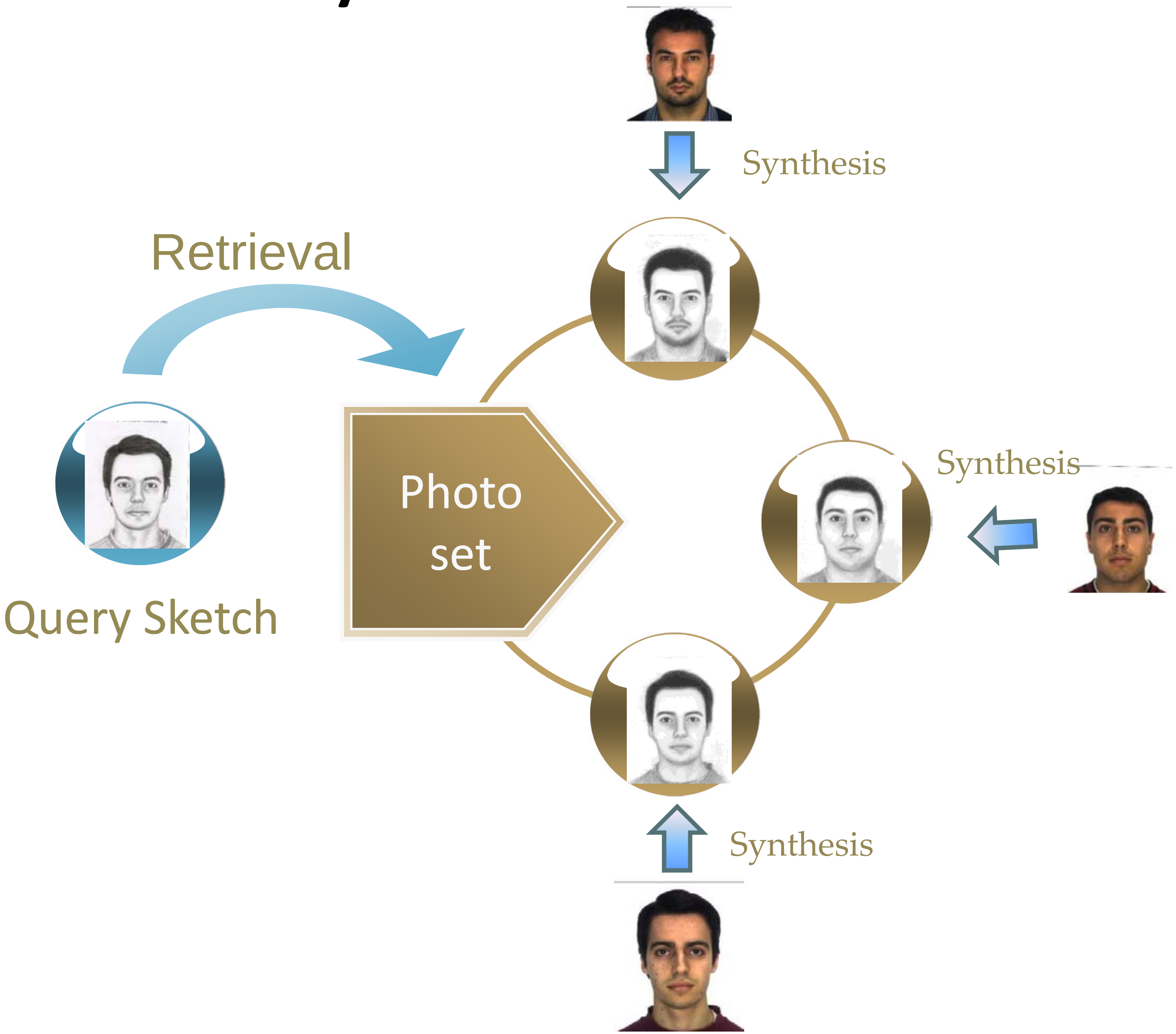
Characteristics:

- Differently expressed
- Highly Rendered
- Bold Exaggerated

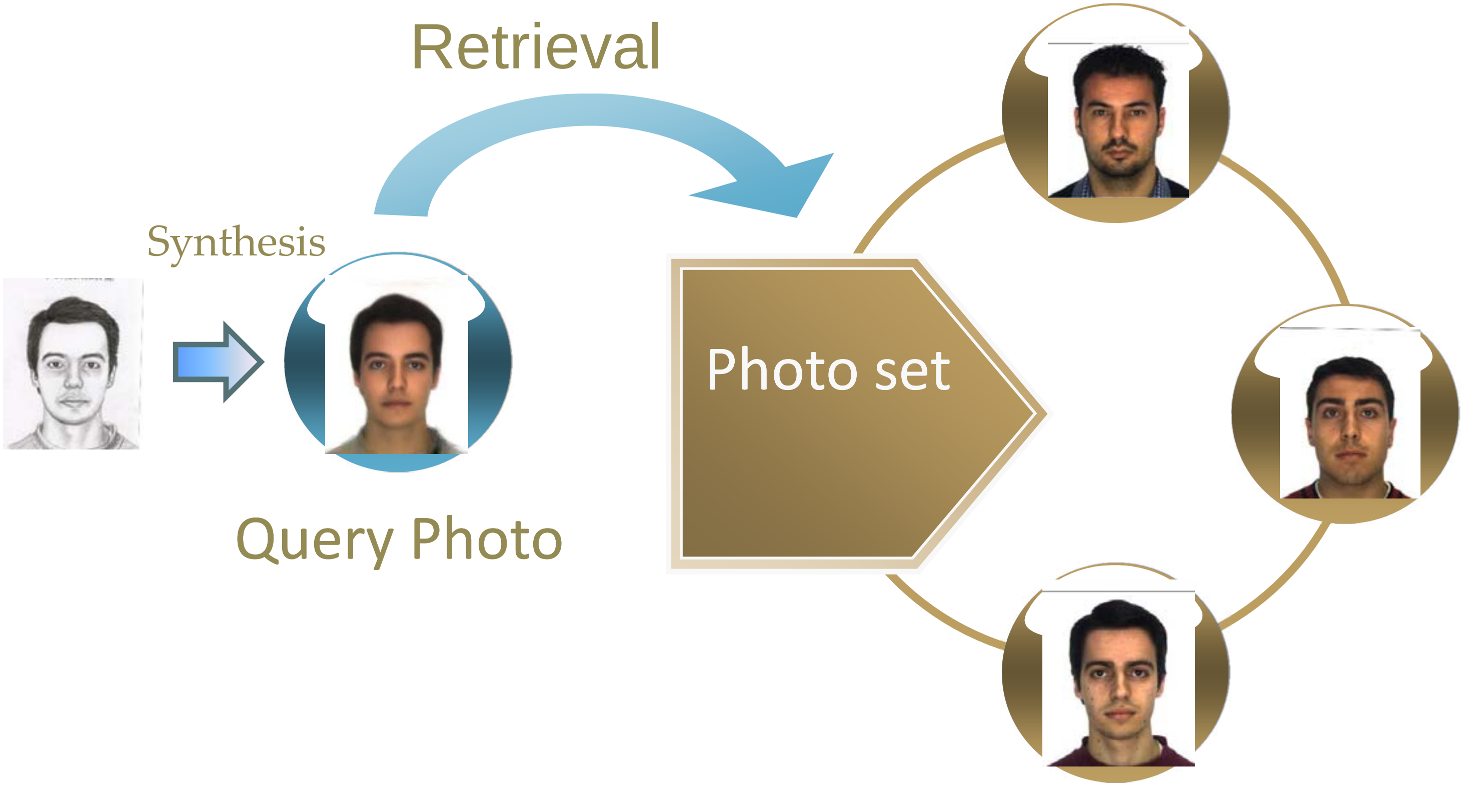
Challenges

- Complicated Mapping
- Diverse Quality Assessment Metrics
- Difficult to analyse their contents

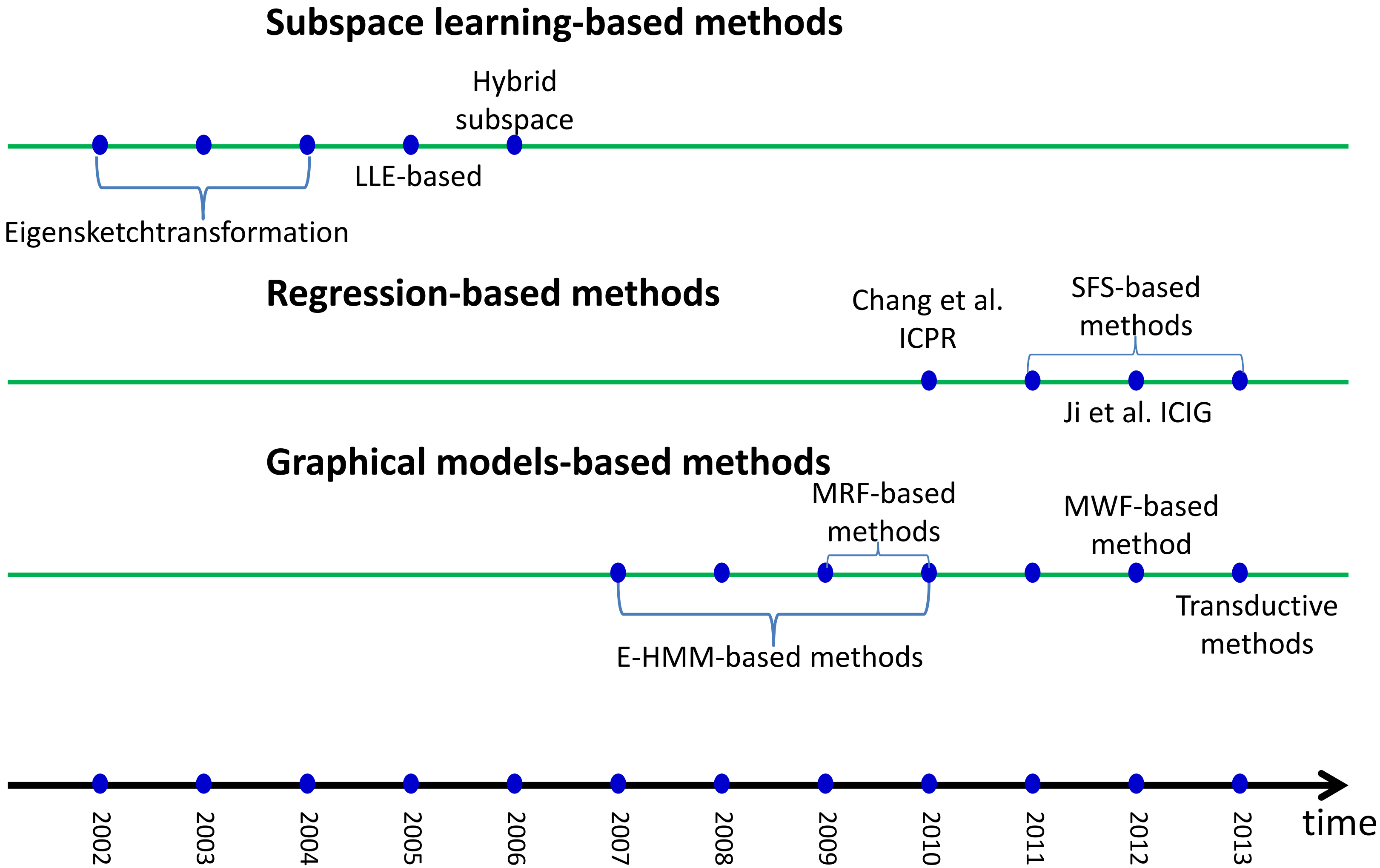
Solutions—Sketch Synthesis



Solutions—Photo Synthesis



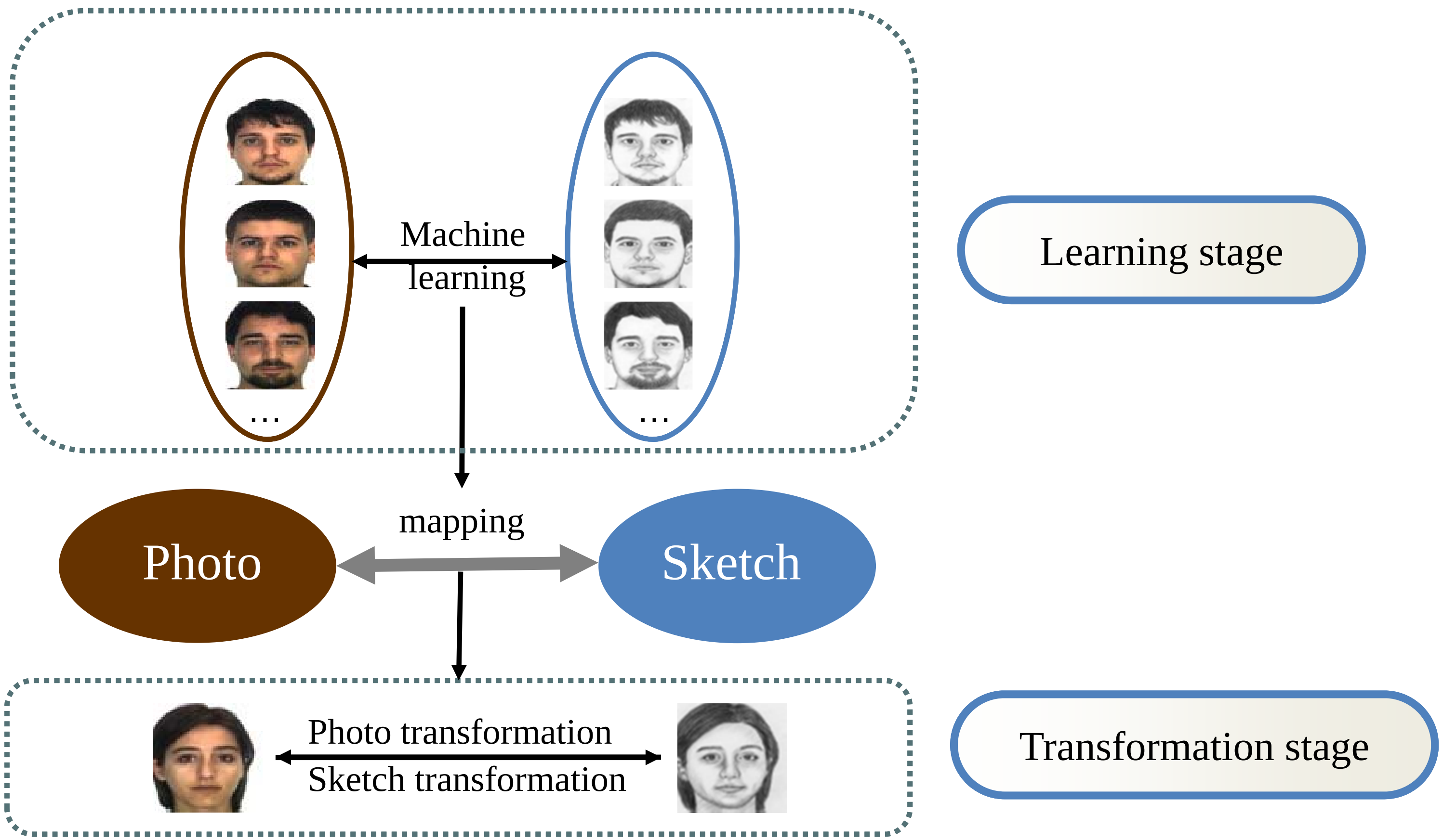
Development Timeline of Representative Methods



Notes:

E-HMM: Embedded Hidden Markov Models
LLE: Locally Linear Embedding
MRF: Markov Random Fields
MWF: Markov Weight Fields
SFS: Sparse Feature Selection

General Pipeline of Sketch-Photo Synthesis



OUTLINE

- Motivation & Introduction
- ➡ Subspace Learning-based Methods
- Sparse Representation-based Methods
- Bayesian Inference-based Methods
- Conclusions

Subspace Learning

Subspace learning refers to the technique of finding a subspace $\mathcal{R}^m$ embedded in a high dimensional space $\mathcal{R}^n (n > m)$.

- **Linear subspace learning** (e.g. principal component analysis):

it is mainly achieves by a projection matrix $U \in \mathcal{R}^{n \times m}$, which is learned from training examples. The matrix U can always be calculated by solving a standard eigenvalue decomposition problem or generalized eigenvalue decomposition problem:

$$\boldsymbol{A}\boldsymbol{u}_i = \lambda \boldsymbol{B}\boldsymbol{u}_i$$

Given an input image $\boldsymbol{f} \in \mathcal{R}^n$, we can find its projection on subspace $\mathcal{R}^m$ from $\boldsymbol{f}_{proc} = \boldsymbol{U}^T \boldsymbol{f}$.

- **Nonlinear subspace learning** (such as locally linear embedding)

It mainly refers to manifold learning. The concept of constructing a local neighborhood has been explored since the methods of such category have no explicit mapping function.

[N. -N. Wang, *et al.*, IJCV13]

OUTLINE

Motivation & Introduction

Subspace Learning-based Methods

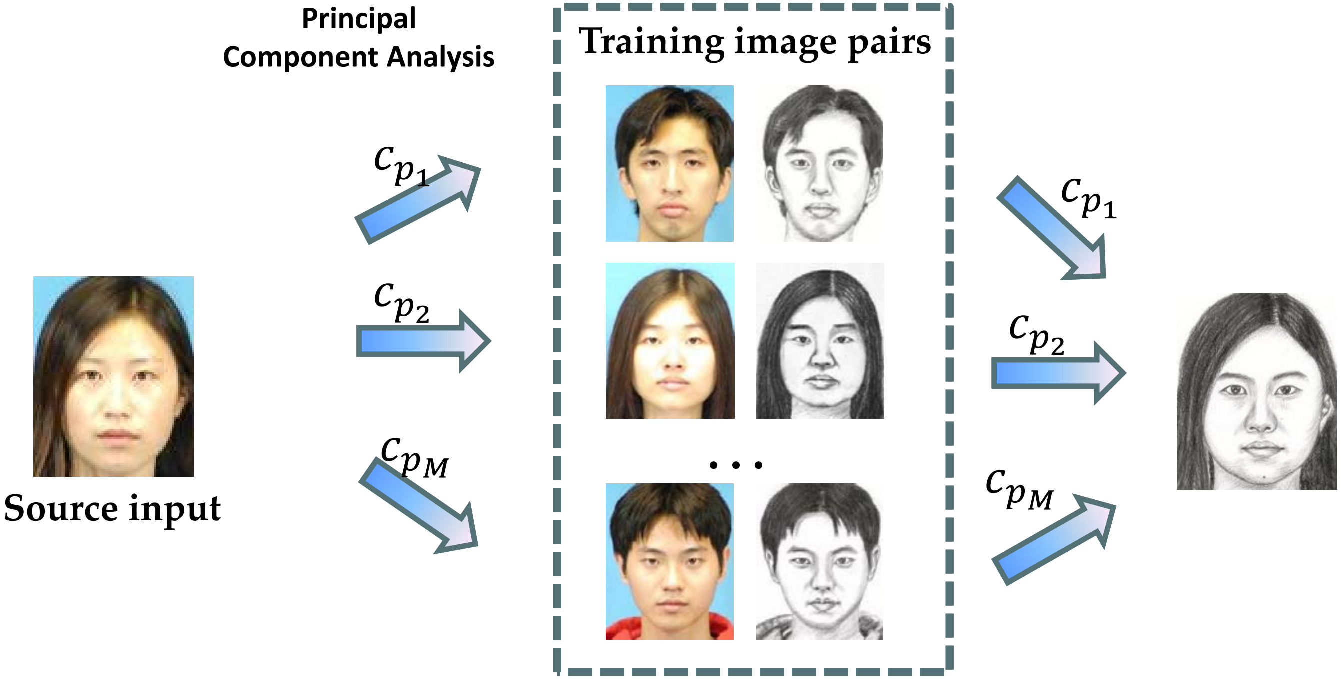
- Eigensketchtransformation
- LLE-based Method

Sparse Representation-based Methods

Bayesian Inference-based Methods

Conclusions

Eigensketchtransformation



$$\boldsymbol{P}_r = \boldsymbol{P}\boldsymbol{c}_p = \sum_{i=1}^{M} c_{p_i} \boldsymbol{P}_i \qquad \boldsymbol{S}_r = \boldsymbol{S}\boldsymbol{c}_p = \sum_{i=1}^{M} c_{p_i} \boldsymbol{S}_i$$

[X.-O. Tang *et al.*, ICIP 02, ICCV 03, CSVT 04]

OUTLINE

Motivation & Introduction

Subspace Learning-based Methods

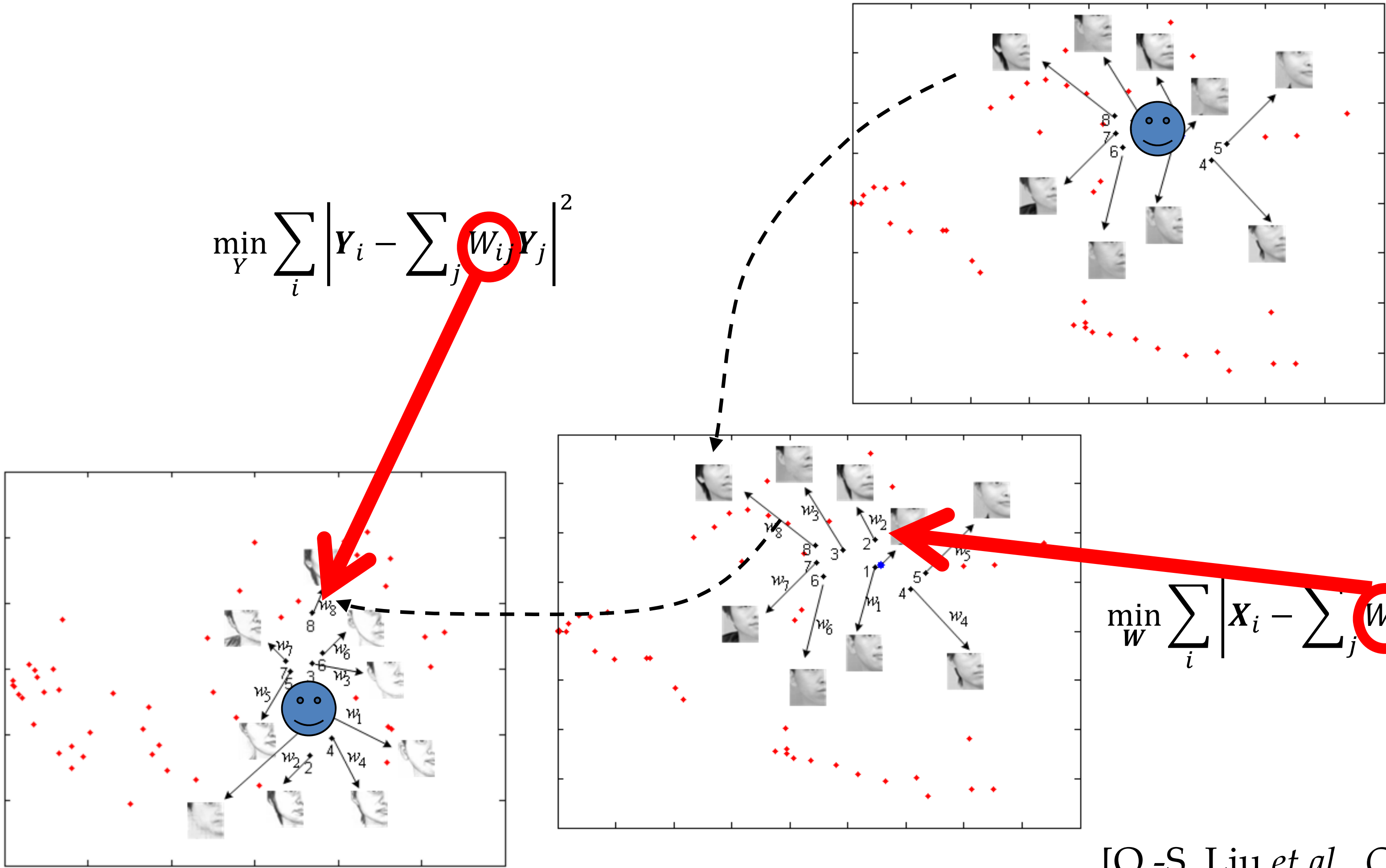
- Eigensketchtransformation
- LLE-based Method

Sparse Representation-based Methods

Bayesian Inference-based Methods

Conclusions

LLE-based Method



$$\min_{Y} \sum_i \left| \boldsymbol{Y}_i - \sum_j W_{ij} \boldsymbol{Y}_j \right|^2$$

$$\min_{\boldsymbol{W}} \sum_i \left| \boldsymbol{X}_i - \sum_j W_{ij} \boldsymbol{X}_j \right|^2$$

[Q.-S. Liu *et al.*, *CVPR* 05]

OUTLINE

- Motivation & Introduction
- Subspace Learning-based Methods
- ➡ Sparse Representation-based Methods
- Bayesian Inference-based Methods
- Conclusions

Sparse Representation

$$\min_{\boldsymbol{c}} \|\boldsymbol{c}\|_0 + \lambda \|\boldsymbol{A}\boldsymbol{c} - \boldsymbol{x}\|_2$$

Sparse coefficient vector

Overcompleted dictionary

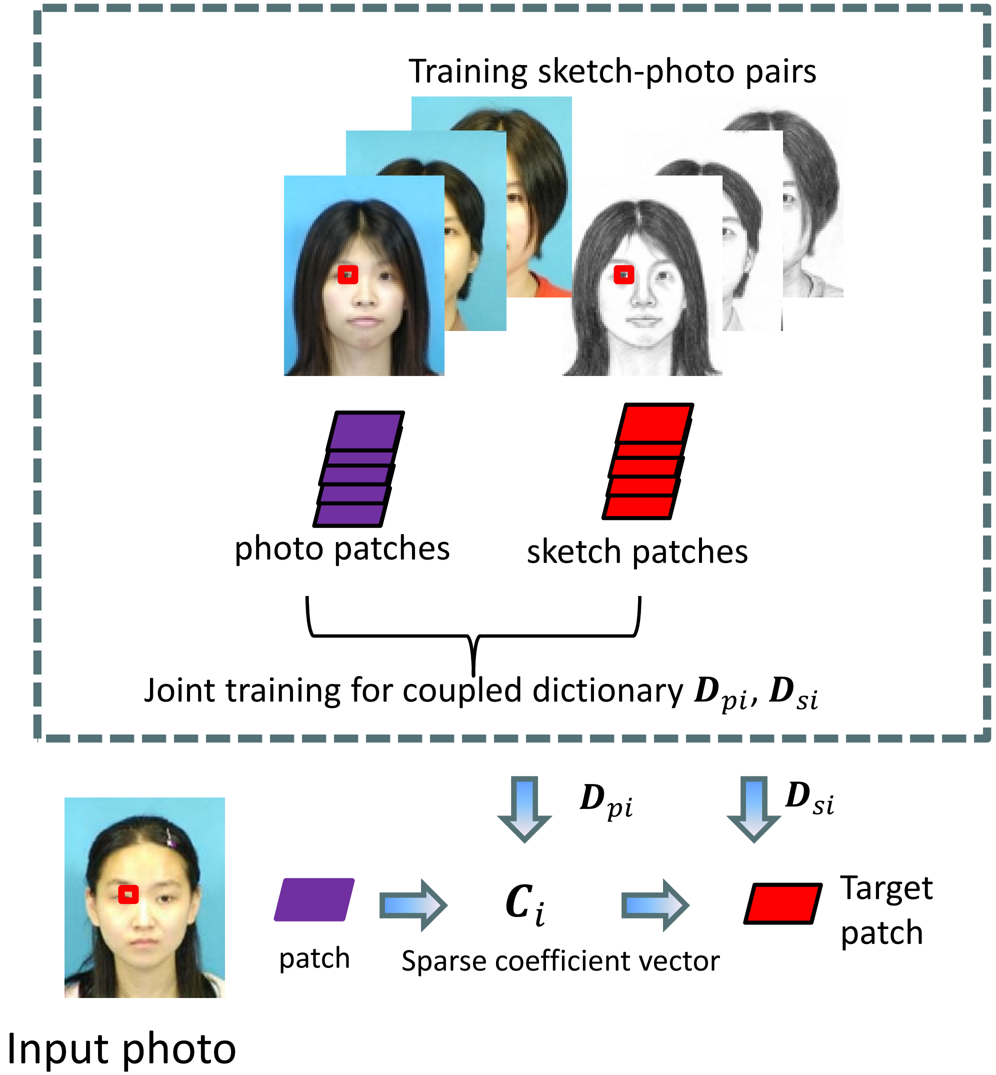
A signal

NP-hard!

$$\min_{\boldsymbol{c}} \|\boldsymbol{c}\|_1 + \lambda \|\boldsymbol{A}\boldsymbol{c} - \boldsymbol{x}\|_2$$

Convex

Sparse Representation-based Sketch Synthesis



Assuming sketch patch and corresponding photo patch have the same sparse representation!

[L. Chang *et al.*, ICPR10]

Refer to [S. –L. Wang *et al.*, CVPR12] for having different sparse representation

OUTLINE

Motivation & Introduction

Subspace Learning-based Methods

Sparse Representation-based Methods

- SFS-based Method
- Local Regression-based Method

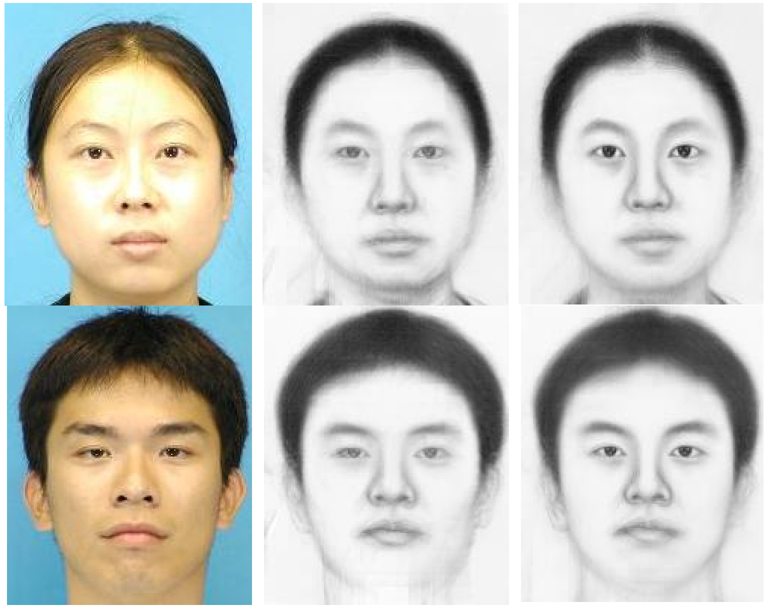
Bayesian Inference-based Methods

Conclusions

SFS-based Method

Motivation

(1) The defects of K-nearest neighbors-based synthesis algorithms: the number of nearest neighbors is **fixed but not adaptive** **(can be solved by sparse feature selection, SFS)**



Input Photo **K-NN** **SFS**

[X.-B. Gao *et al.*, ICIG2011, ICIP2011, CSVT12, PRL13]

SFS-based Method

Motivation

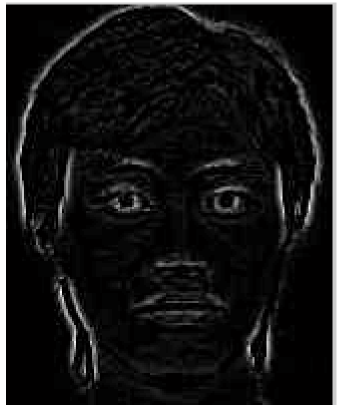
(2)

Sketch drawn
by the artist

−

Synthesized sketch
by K-NN-based method

=

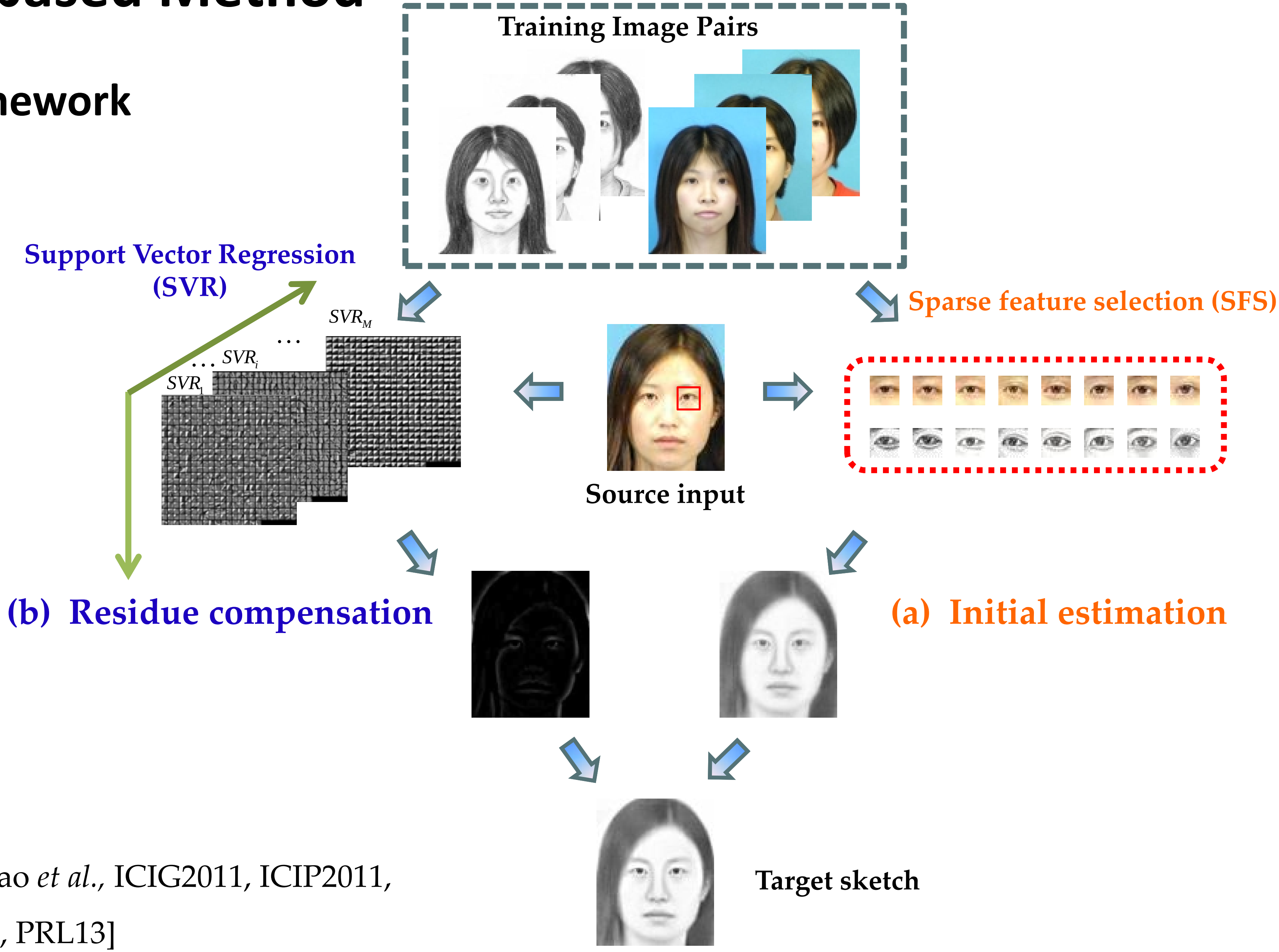


Residue
(can be compensated by
SVR-based hallucination)

[X.-B. Gao *et al.,* ICIG2011, ICIP2011, CSVT12, PRL13]

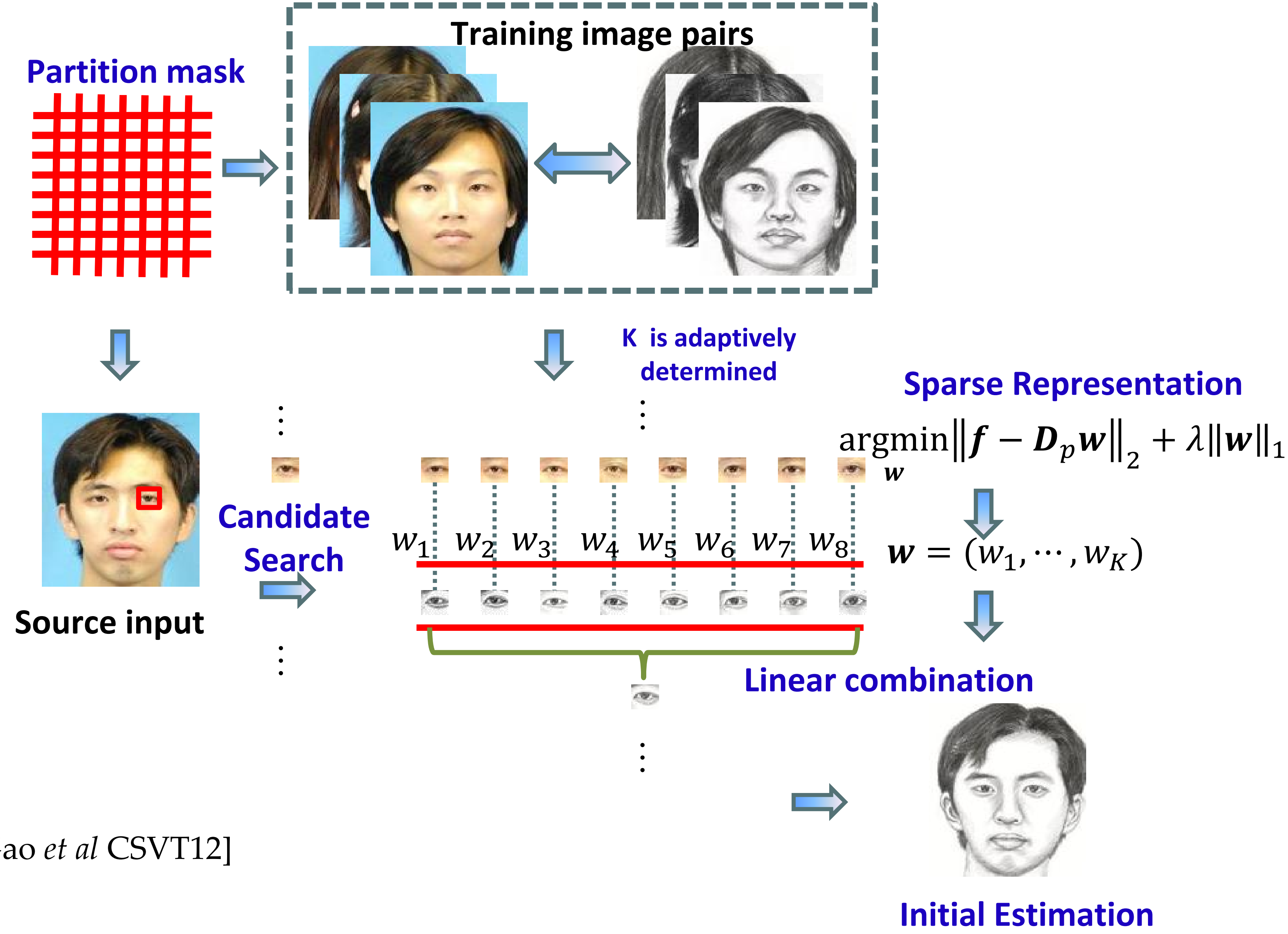
SFS-based Method

Framework



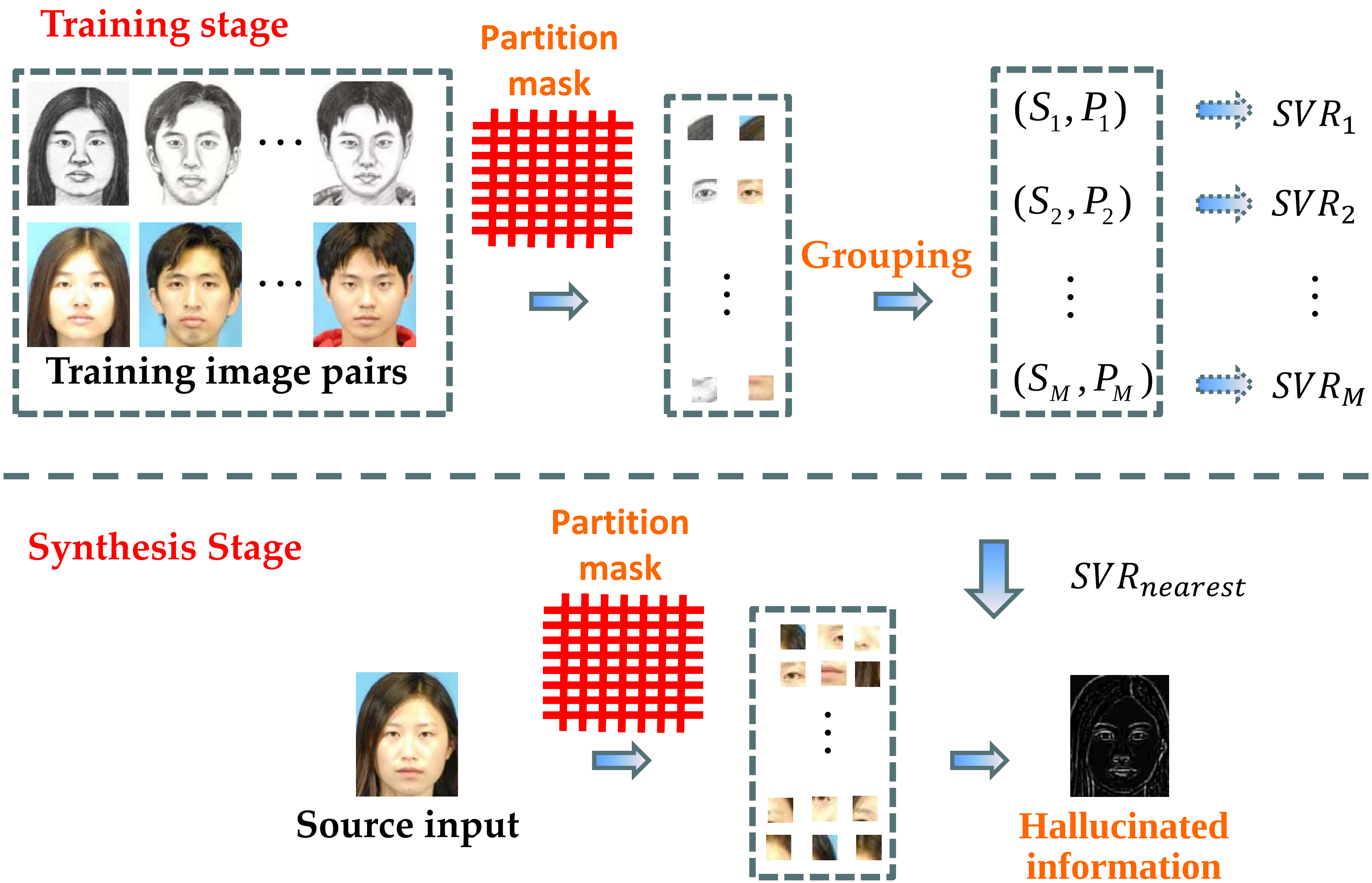
[X.-B. Gao *et al.*, ICIG2011, ICIP2011, CSVT12, PRL13]

SFS-based Method



[X.-B. Gao *et al* CSVT12]

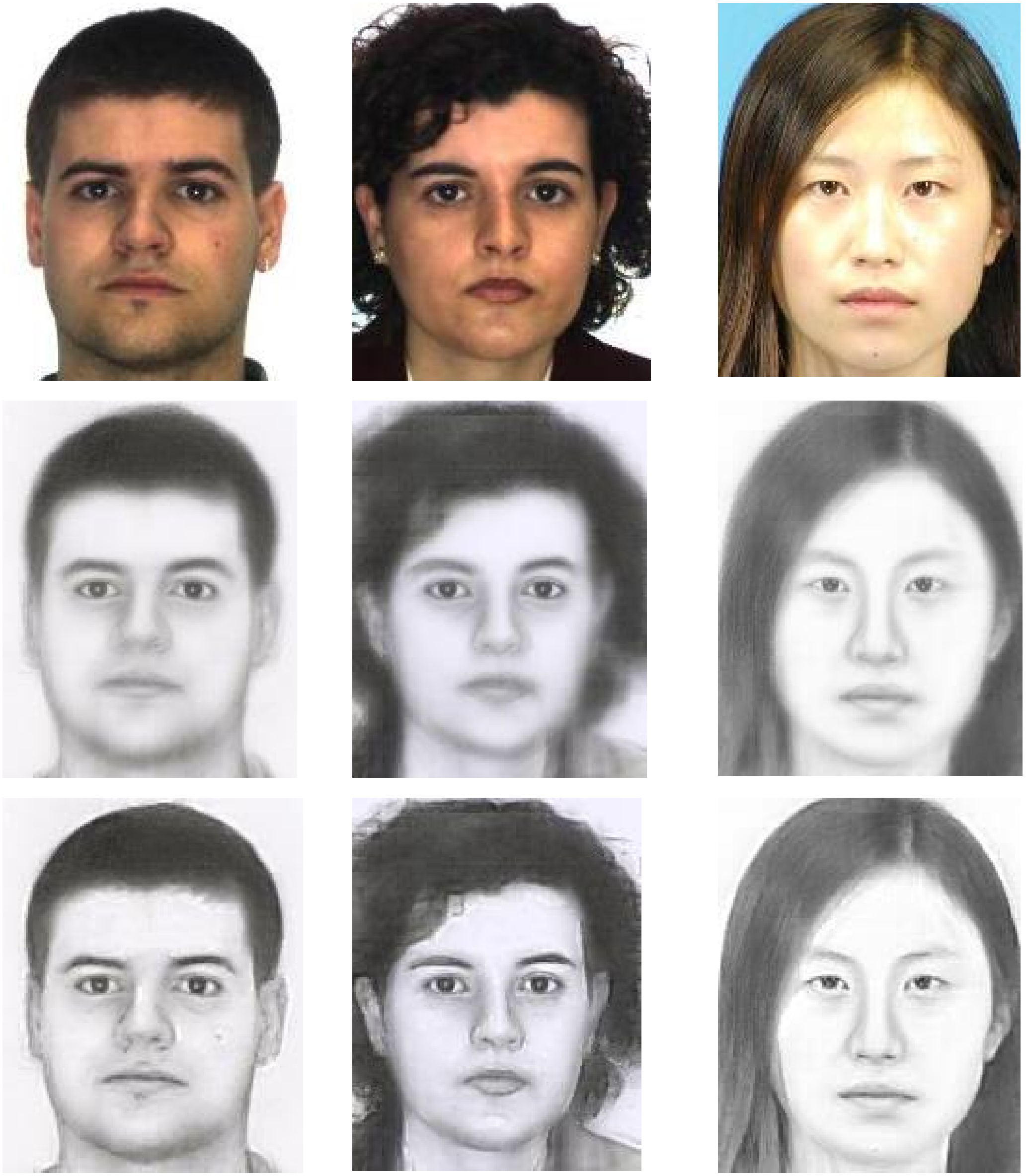
SFS-based Method



[N.-N. Wang, X. –B. Gao, *et al.*, PRL13]

SFS-based Method

Example Results-Synthesized Sketches



Input Photo

Results of SFS

Results of SFS-SVR

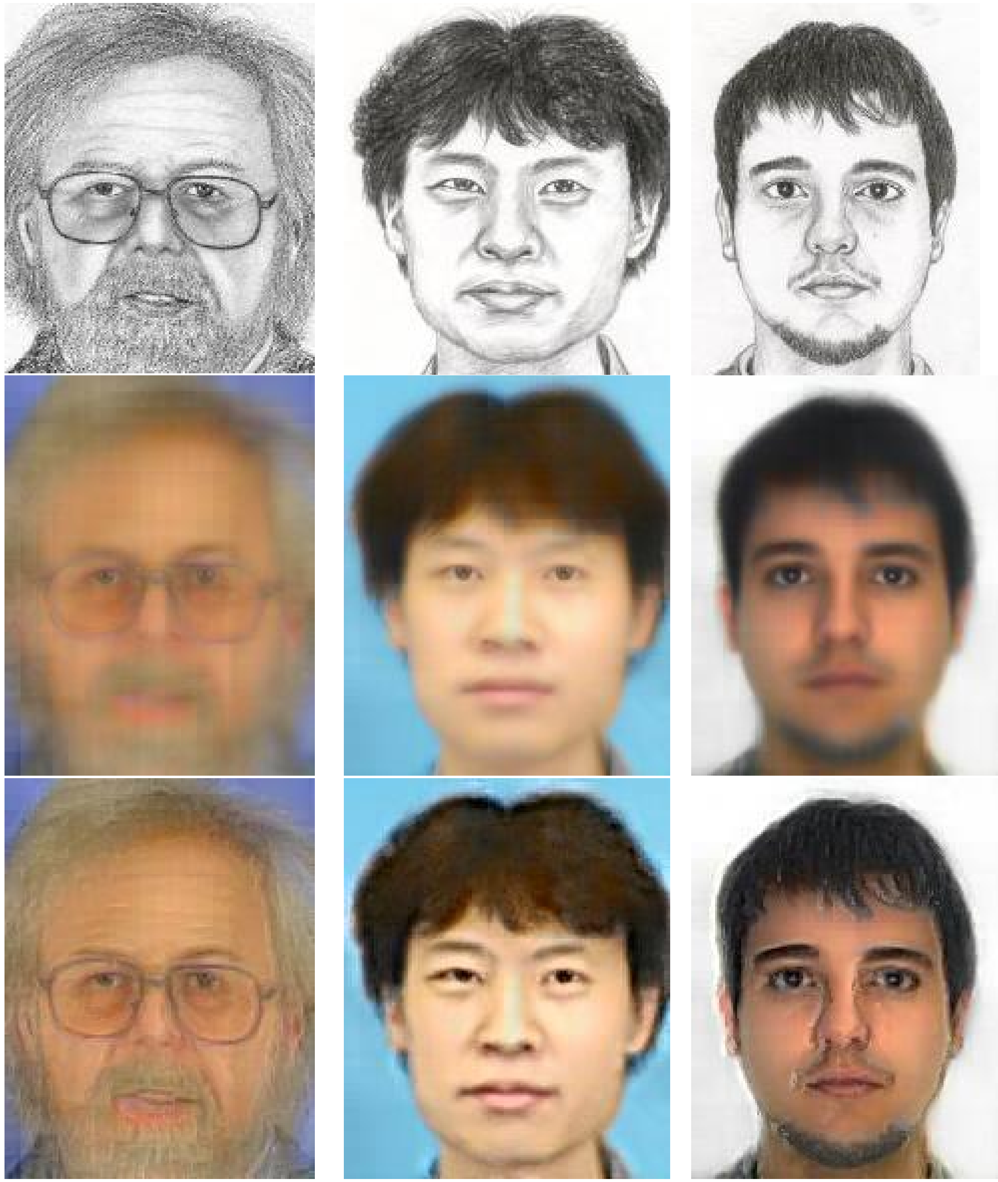
SFS-based Method

Example Results-Synthesized Photos

Input Sketch

Results of SFS

Results of SFS-SVR



SFS-based Method

Example Results-Comparisons with KNN based method

Input Photo

Results of LLE

Results of SFS-SVR

LLE: locally linear embedding-based method

SFS-based Method

Face Recognition

Method	LLE	E-HMM	SFS	SFS-SVR
Recognition Rate(%)	84	87	91	**93**

Image Quality Assessment (IQA)

	LLE	E-HMM	SFS	**SFS-SVR**
	0.0921	0.0948	0.0956	**0.0972**

E-HMM: embedded hidden Markov models-based method

OUTLINE

Motivation & Introduction

Subspace Learning-based Methods

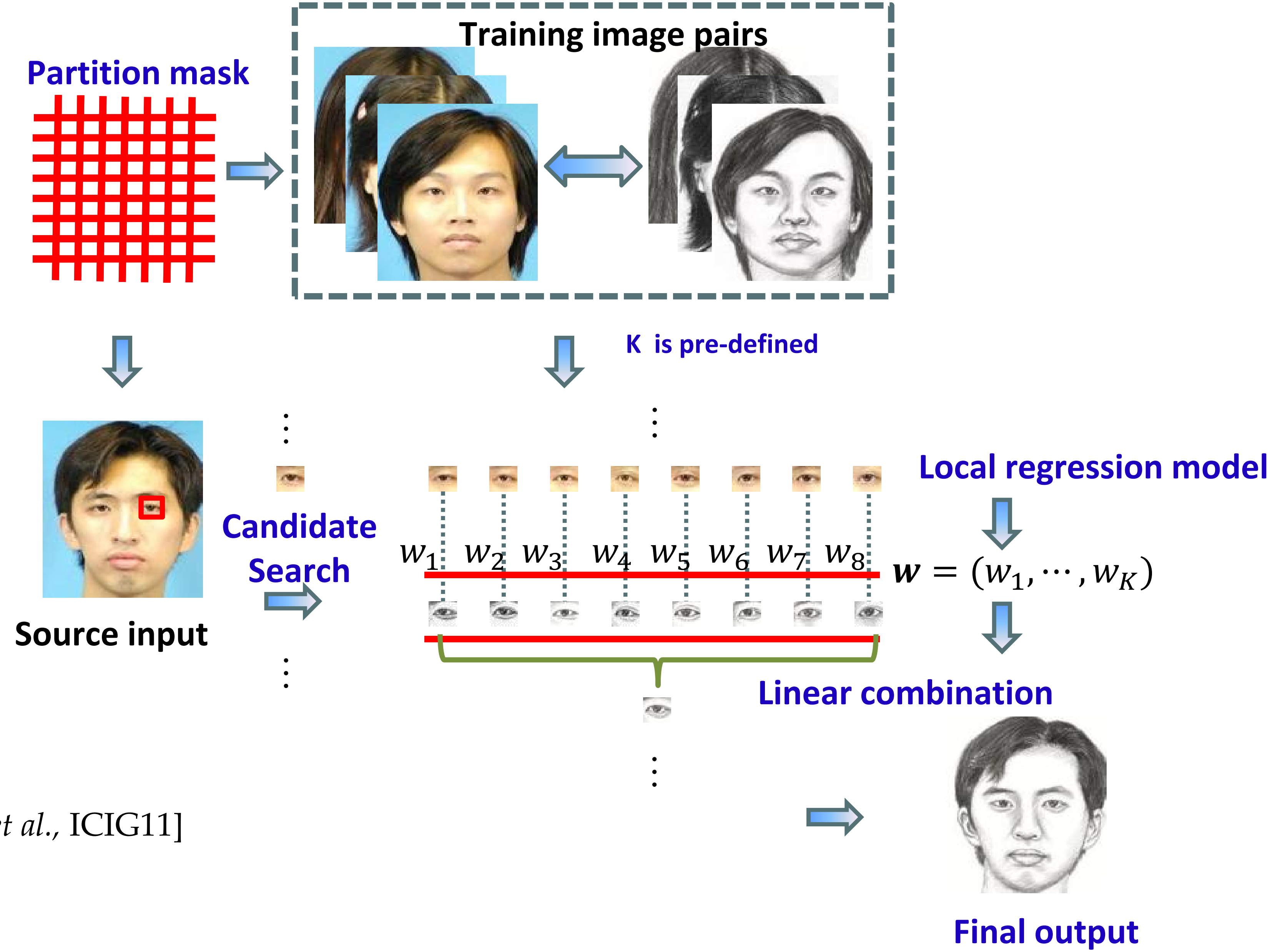
Sparse Representation-based Methods

- SFS-based Method
- Local regression-based Method

Bayesian Inference-based Methods

Conclusions

Local Regression-based Method



[N.-Y. Ji *et al.*, ICIG11]

Local Regression-based Method

Regression methods

- K Nearest Neighbor (KNN)

$$\boldsymbol{w}^{kNN} = \begin{cases} s(\mathbf{p}, \mathbf{p}_j), if\ KNN \\ 0, otherwise \end{cases}$$

$s(\cdot)$ is the similarity measurement metric

- Least Squares (LS)

$$\boldsymbol{w}^{LS} = \underset{\boldsymbol{w}}{\operatorname{argmin}} \left| \mathbf{p} - \sum \mathrm{w}_{\mathrm{i}} \mathbf{p}_{\mathrm{i}} \right|^2$$

- Ridge Regression (RR)

$$\boldsymbol{w}^{LS} = \underset{\boldsymbol{w}}{\operatorname{argmin}} \left| \mathbf{p} - \sum \mathrm{w}_{\mathrm{i}} \mathbf{p}_{\mathrm{i}} \right|^2 + \lambda |\boldsymbol{w}|_2$$

- Lasso

$$\boldsymbol{w}^{LS} = \underset{\boldsymbol{w}}{\operatorname{argmin}} \left| \mathbf{p} - \sum \mathrm{w}_{\mathrm{i}} \mathbf{p}_{\mathrm{i}} \right|^2 + \lambda |\boldsymbol{w}|_1$$

[N.-Y. Ji *et al.*, ICIG11]

OUTLINE

Motivation & Introduction

Subspace Learning-based Methods

Regression-based Methods

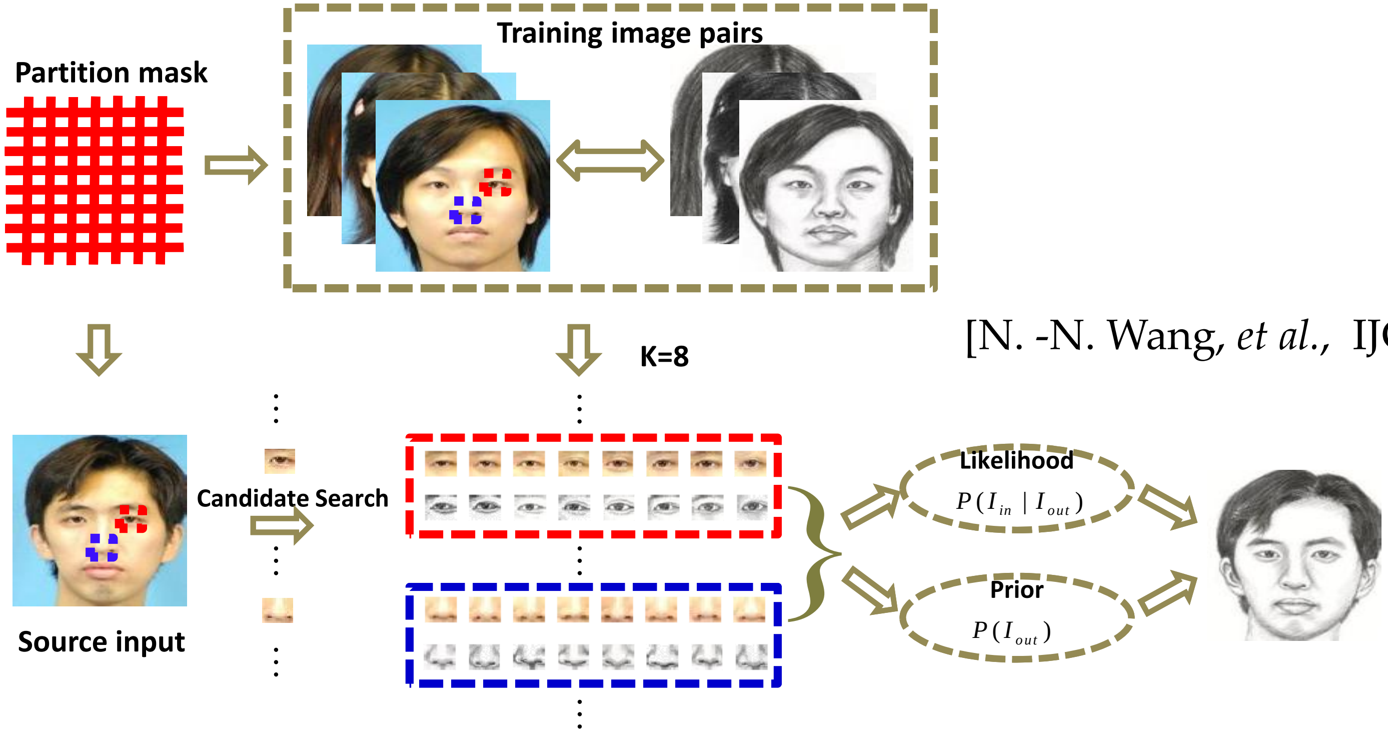
➡ Bayesian Inference-based Methods

Conclusions

Bayesian Inference-based Framework

Given that I_{in} and I_{out} denote the input (observation) and output image (to be estimated) for FH, respectively, the maximum a posteriori (MAP) decision rule in Bayesian statistics for FH is written as:

$$\boldsymbol{I}_{out}^{*} = \underset{\boldsymbol{I}_{out}}{\operatorname{argmax}} P(\boldsymbol{I}_{out}|\boldsymbol{I}_{in}) = \underset{\boldsymbol{I}_{out}}{\operatorname{argmax}} P(\boldsymbol{I}_{in}|\boldsymbol{I}_{out})P(\boldsymbol{I}_{out})$$



[N. -N. Wang, *et al.*, IJCV13]

OUTLINE

Motivation & Introduction

Subspace Learning-based Methods

Regression-based Methods

Bayesian Inference-based Methods

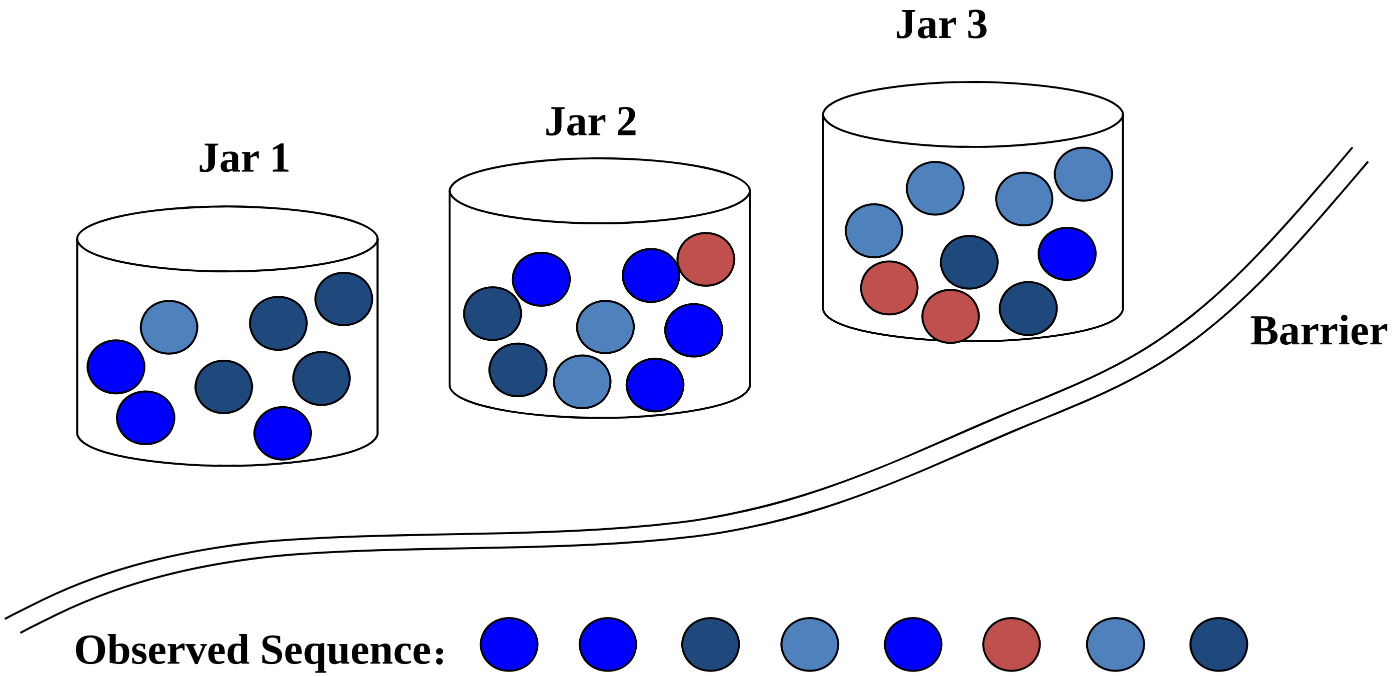
- E-HMM-based Methods
- MRF-based Method
- Transductive Method

Conclusions

E-HMM-based Methods

Sketch-photo Relationship Modeling:

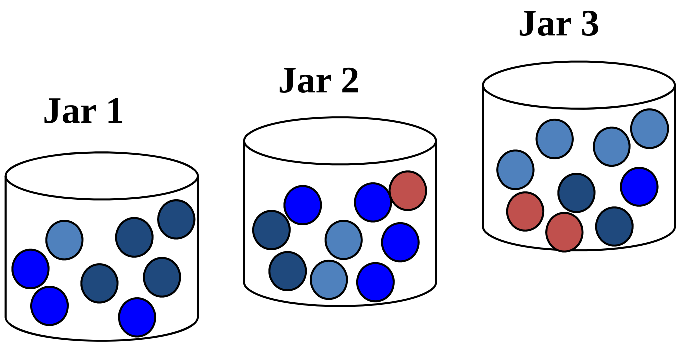
- Statistical Method → MRF: Markov Random Field
- HMM: Hidden Markov Model



E-HMM-based Methods

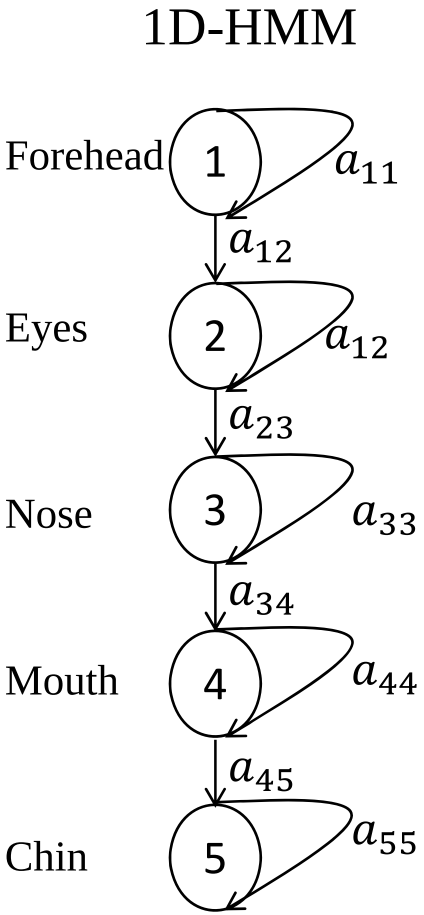
Hidden Markov Model:

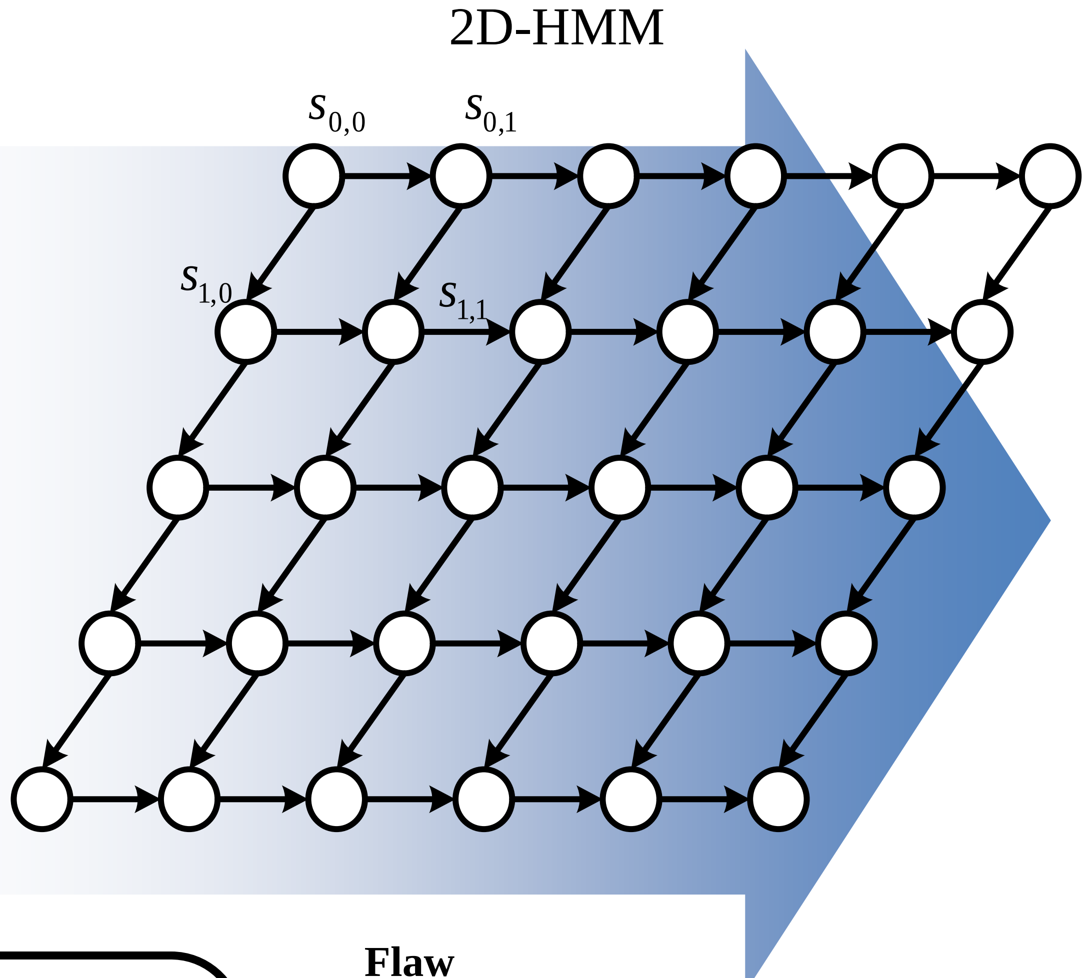
$$\boldsymbol{\lambda} = (\boldsymbol{\Pi}, \boldsymbol{A}, \boldsymbol{B}, N, M)$$



Param.	***Name***	***Example***
N	State Number	Number of Jar: $N = 3$
M	Number of Observation Values	Color Number of Balls: $M = 4$
$\boldsymbol{A}$	State Transition Probability Matrix	P{Jar i→Jar j}, $i,j = 1,2,3$
$\boldsymbol{B}$	Observation Probability Matrix	P{Ball.color_Jar i}
$\boldsymbol{\Pi}$	Initial State Distribution	P{Ball_1 from Jar i}

E-HMM-based Methods





Flaw

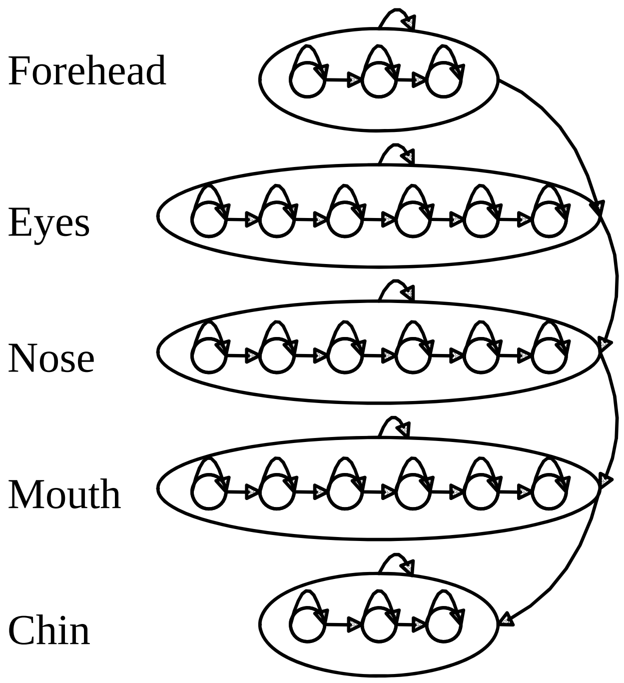
- Restricted to fixed-size face image
- The signal is transformed to 1D observation

Flaw

- High computational complexity

E-HMM-based Methods

Pseudo 2D HMM

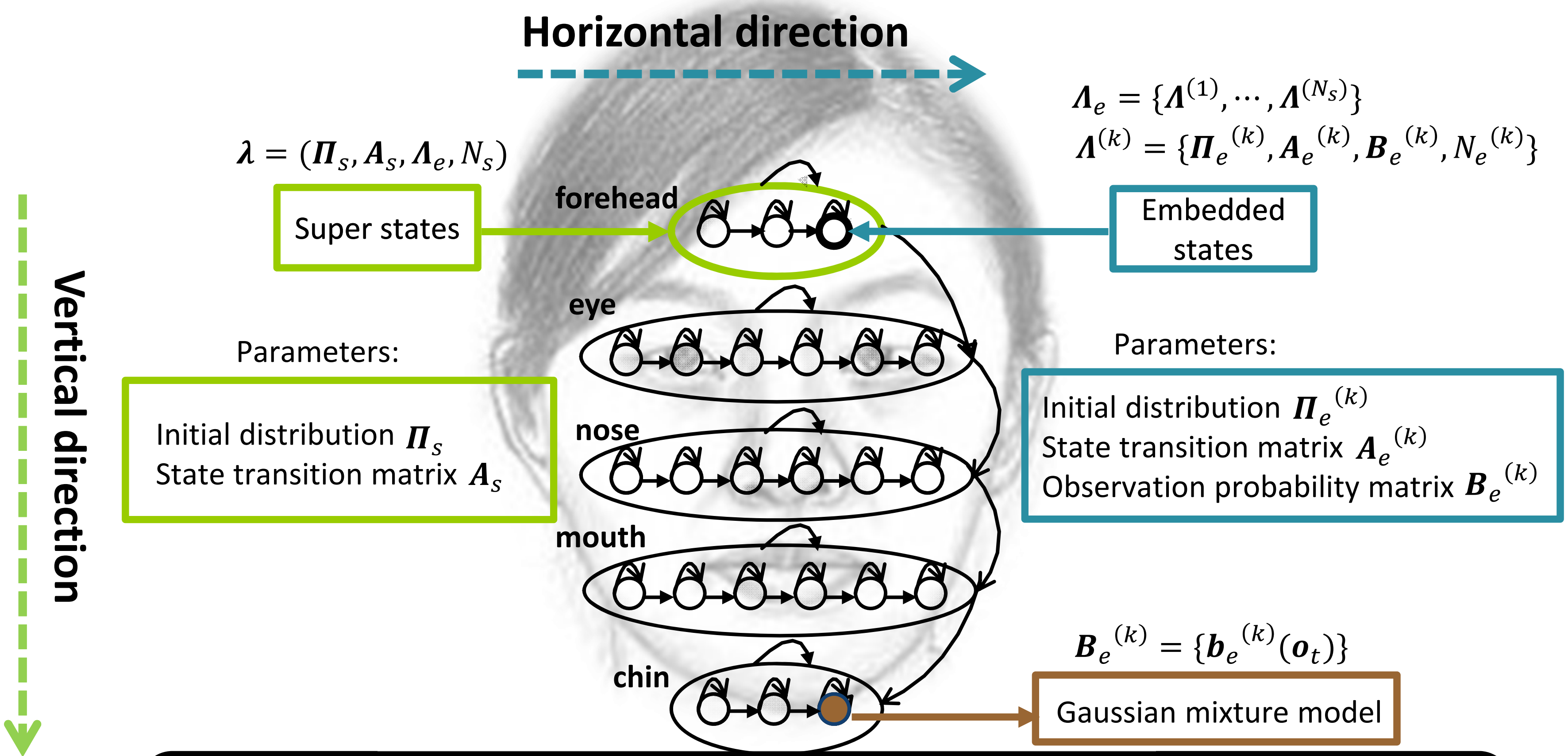


Flaw

- The signal is transformed to 1D observation sequence with a zigzag fashion

E-HMM-based Methods

E-HMM Structure for Face Image



Advantages of E-HMM:

- It can extract the main 2D facial features and has a moderate computational complexity
- It is robust to the change of pose and environment.

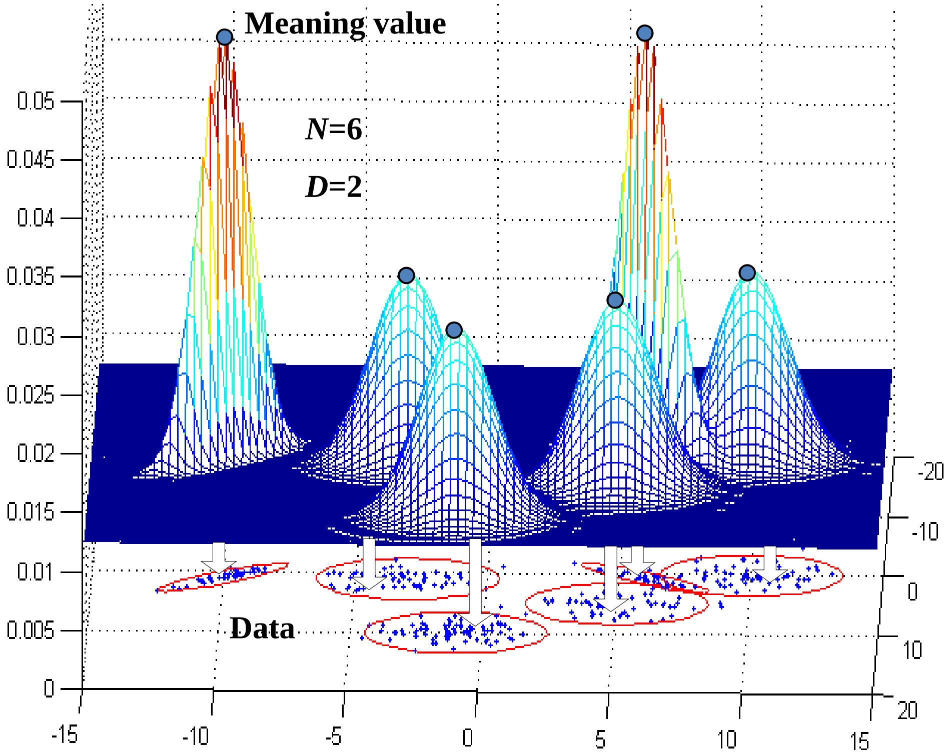
[X.-B. Gao *et al.*, ICASSP07, CSVT08]

E-HMM-based Methods

Gaussian mixture model

$$b_i^{(k)}(o_t) = \sum_{m=1}^{N_i^{(k)}} c_{im}^{(k)} N(o_t, \mu_{im}^{(k)}, \Sigma_{im}^{(k)}) = \sum_{m=1}^{N_i^{(k)}} c_{im}^{(k)} \frac{1}{\sqrt{(2\pi)^D \left|\Sigma_{im}^{(k)}\right|}} \exp\left(-\frac{1}{2}\left(o_t - \mu_{im}^{(k)}\right)^T \left(\Sigma_{im}^{(k)}\right)\left(o_t - \mu_{im}^{(k)}\right)\right)$$

Observation Probability



E-HMM-based Methods

Foundation of the E-HMM Theory

(Q1) Computing the output probability: given the image observation sequence $\boldsymbol{O} = (\boldsymbol{o}_1, \boldsymbol{o}_2, \cdots, \boldsymbol{o}_T)$ and the E-HMM model $\boldsymbol{\lambda} = \{\boldsymbol{\Pi}_S, \boldsymbol{A}_S, \boldsymbol{\Lambda}_{\boldsymbol{e}}\}$, how to get the output probability $P(\boldsymbol{O}|\boldsymbol{\lambda})$

Forward-Backward Algorithm

(Q2) Decoding the state sequence: with the image observation sequence $\boldsymbol{O} = (\boldsymbol{o}_1, \boldsymbol{o}_2, \cdots, \boldsymbol{o}_T)$ and the E-HMM model $\boldsymbol{\lambda} = \{\boldsymbol{\Pi}_S, \boldsymbol{A}_S, \boldsymbol{\Lambda}_{\boldsymbol{e}}\}$, how to determine the optimal state sequence $\boldsymbol{Q} = (q_1, q_2, \cdots, q_T)$ and the mixture indexes $\boldsymbol{M} = (m_1, m_2, \cdots, m_T)$

Embedded Viterbi Algorithm

(Q3) Estimating the model parameters: how to adjust the parameters of the E-HMM model $\boldsymbol{\lambda} = \{\boldsymbol{\Pi}_S, \boldsymbol{A}_S, \boldsymbol{\Lambda}_{\boldsymbol{e}}\}$ to maximize $P(\boldsymbol{O}|\boldsymbol{\lambda})$

Baum–Welch Algorithm

E-HMM-based Methods

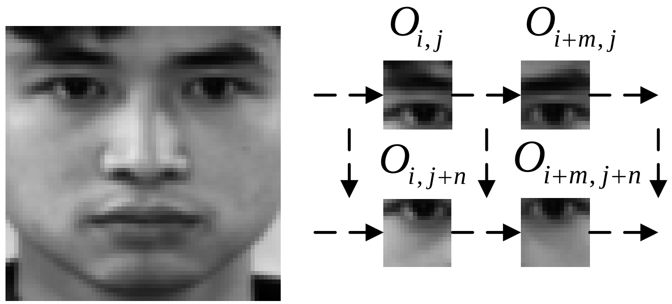
Face Representation Ability of E-HMM

(a) Original image

(b) Reconstructed image

(a) Modeling by Baum-Welch algorithm (Q3)
(b) Reconstructed face by Viterbi algorithm (Q2)

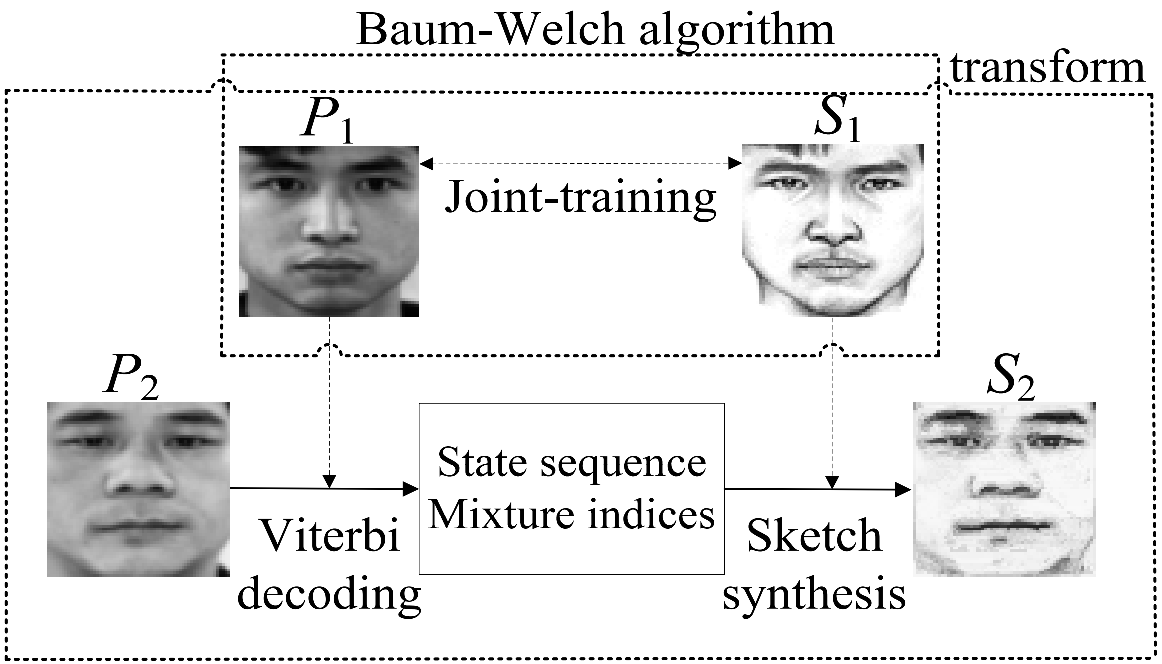
It is shown that E-HMM is able to represent face well.



E-HMM-based Methods

①

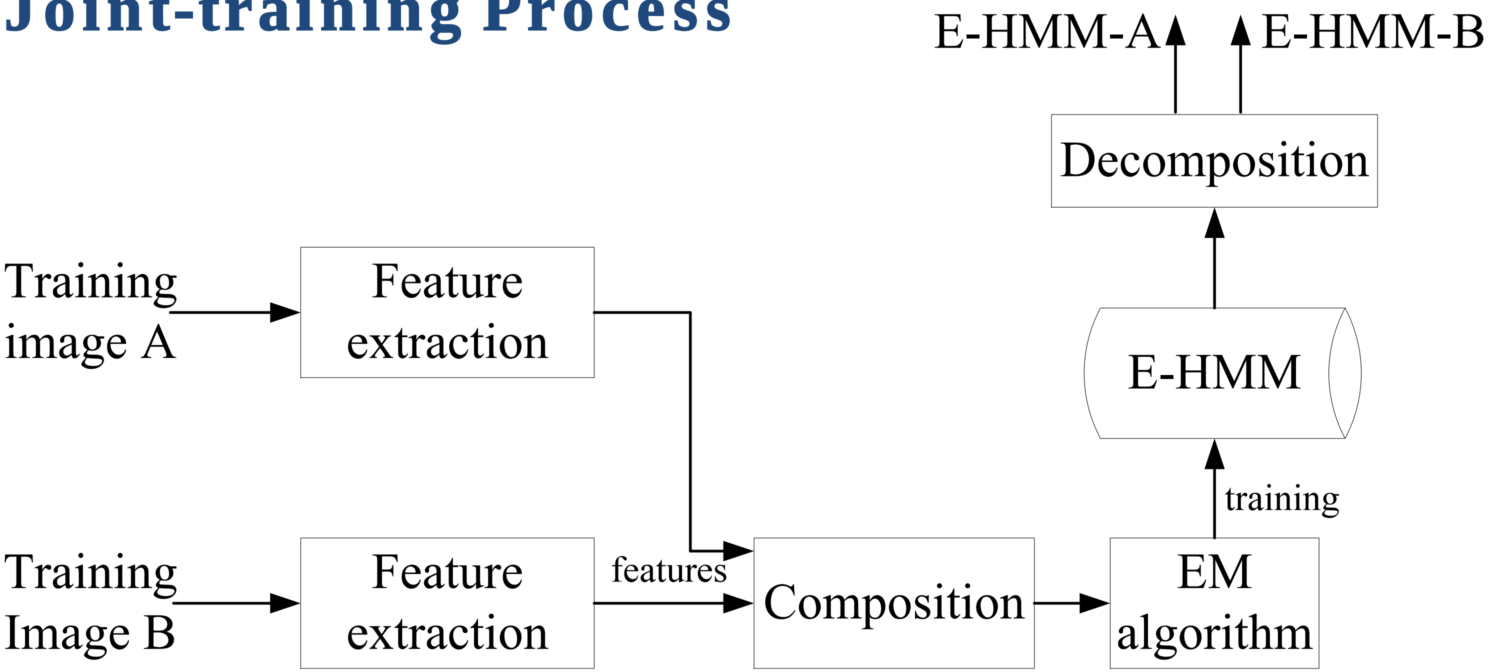
Sketch-photo synthesis based on E-HMM



[J.Zhong, X.-B. Gao *et al.*, ICASSP07]

E-HMM-based Methods

Joint-training Process

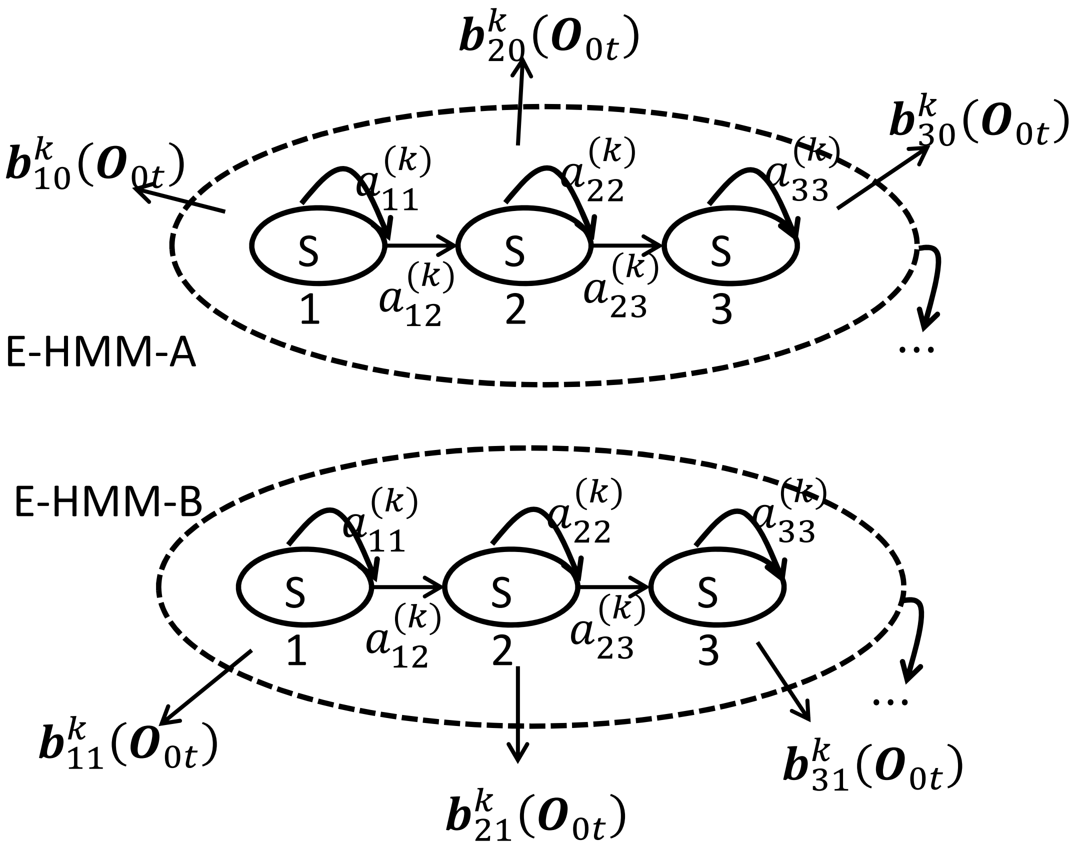


[J. Zhong, X.-B. Gao *et al.*, ICASSP07]

Observation vector=[gray value, Gaussian operator, Laplacian operator, horizontal derivative and vertical derivative]

E-HMM-based Methods

Basic idea of joint-training



$$\boldsymbol{b}_j^k(\overline{\boldsymbol{O}_t}) = \sum_{m=1}^{N} c_{jm}^{(k)} N(\overline{\boldsymbol{O}_t}; \mu_{jm}^{(k)}, \Sigma_{jm}^{(k)})$$

$$= \sum_{m=1}^{N} c_{jm}^{(k)} \prod_{l=0}^{1} \{N(\boldsymbol{O}_{lt}; \boldsymbol{\mu}_{jml}^{(k)}, \boldsymbol{\Sigma}_{jml}^{(k)})\}$$

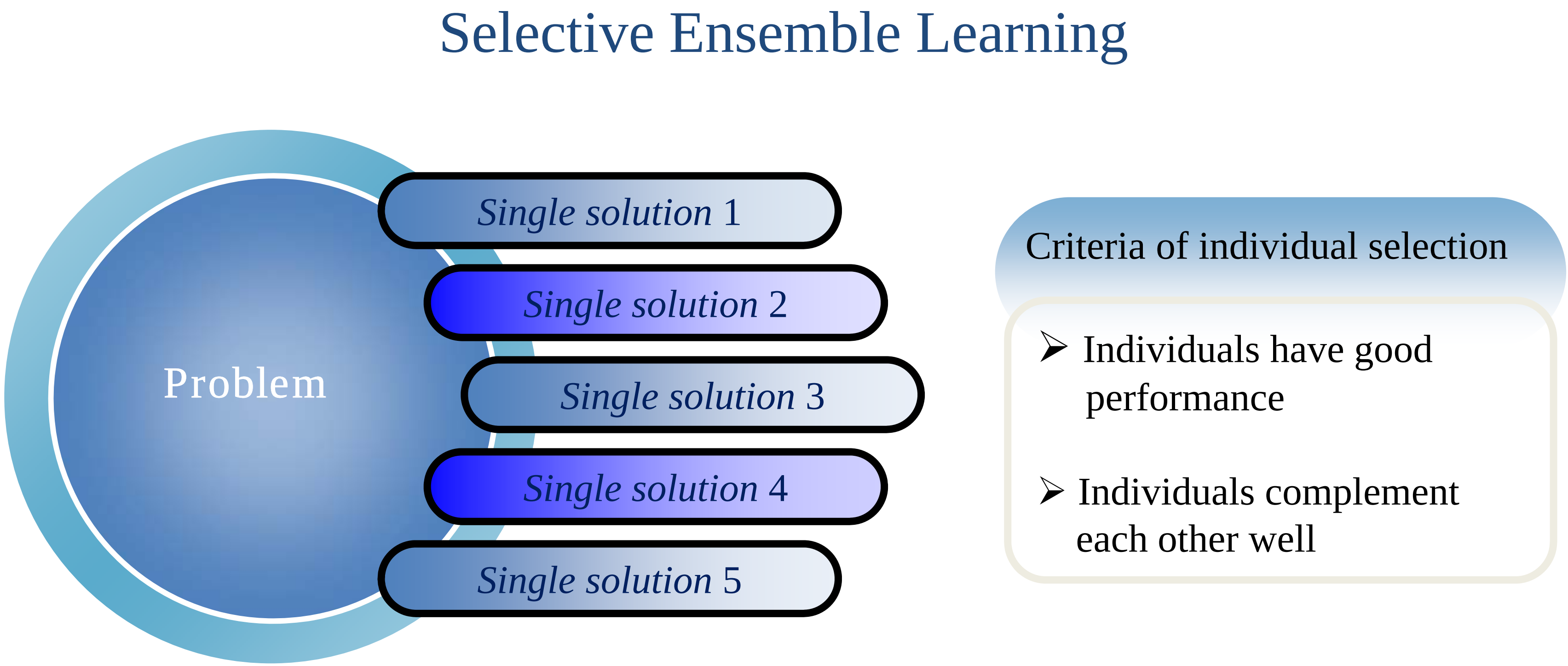
$$\boldsymbol{O}_t = [\boldsymbol{O}_{0t}^T, \boldsymbol{O}_{1t}^T]^T$$

$$\boldsymbol{\mu}_{jm}^{(k)} = \left[{\boldsymbol{\mu}_{jm0}^{(k)}}^T, {\boldsymbol{\mu}_{jm1}^{(k)}}^T\right]^T$$

$$\boldsymbol{\Sigma}_{jm}^{(k)} = \begin{bmatrix} \boldsymbol{\Sigma}_{jm0}^{(k)} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}_{jm1}^{(k)} \end{bmatrix}$$

The E-HMMs of photo and sketch share the same state and same state transition probability matrix.
The mean values and the covariance matrixes in the same state are different.

E-HMM-based Methods



Given many solutions to a problem, we usually choose the best one as the final decision; while in selective ensemble, the optimal decision is given by combining several complementary solutions with weights.

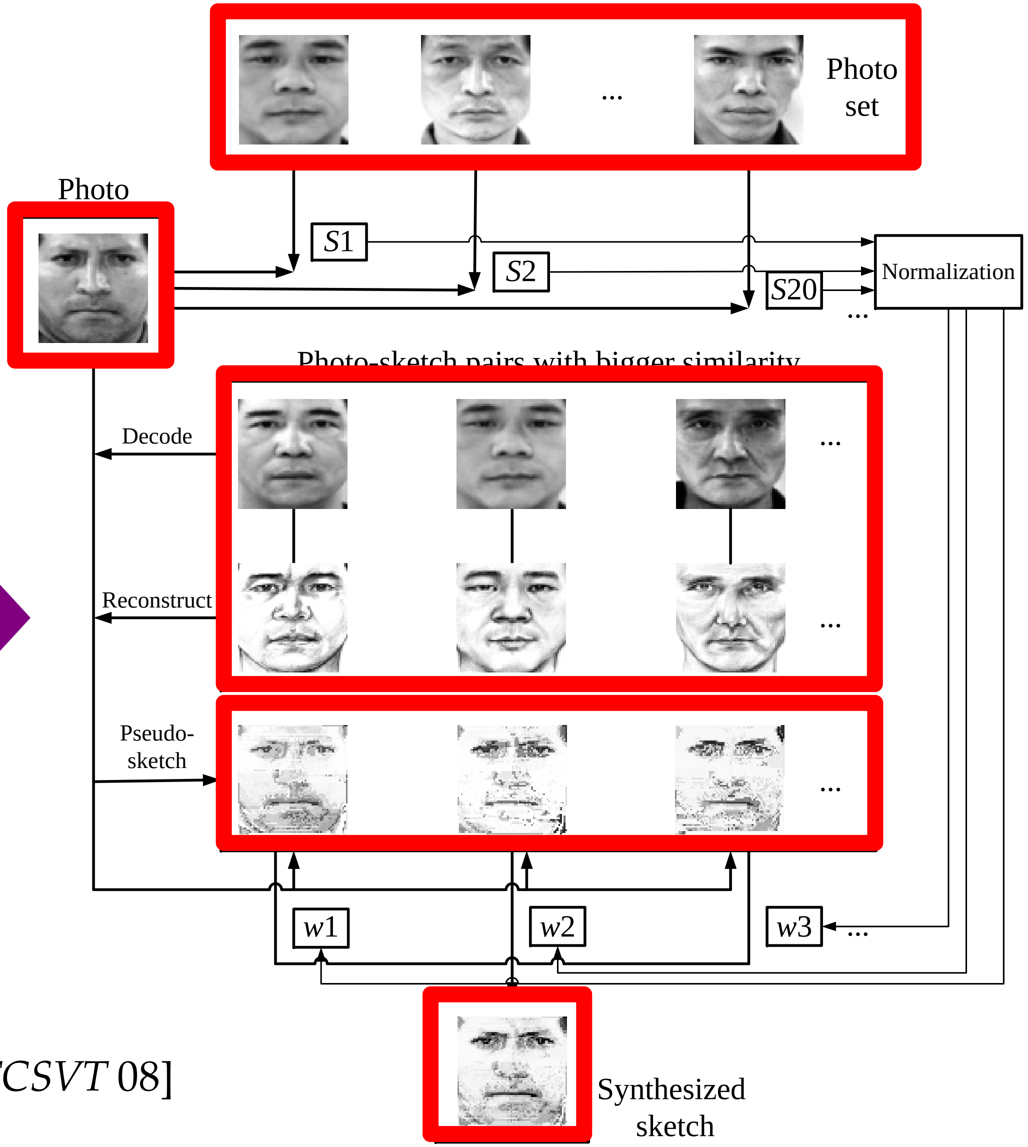
[Z.-H. Zhou *et al.*, AIJ02]

E-HMM-based Methods

Selective Ensemble

②

Sketch-photo synthesis based on E-HMM+SE



[X. -B. Gao *et al.*, *TCSVT* 08]

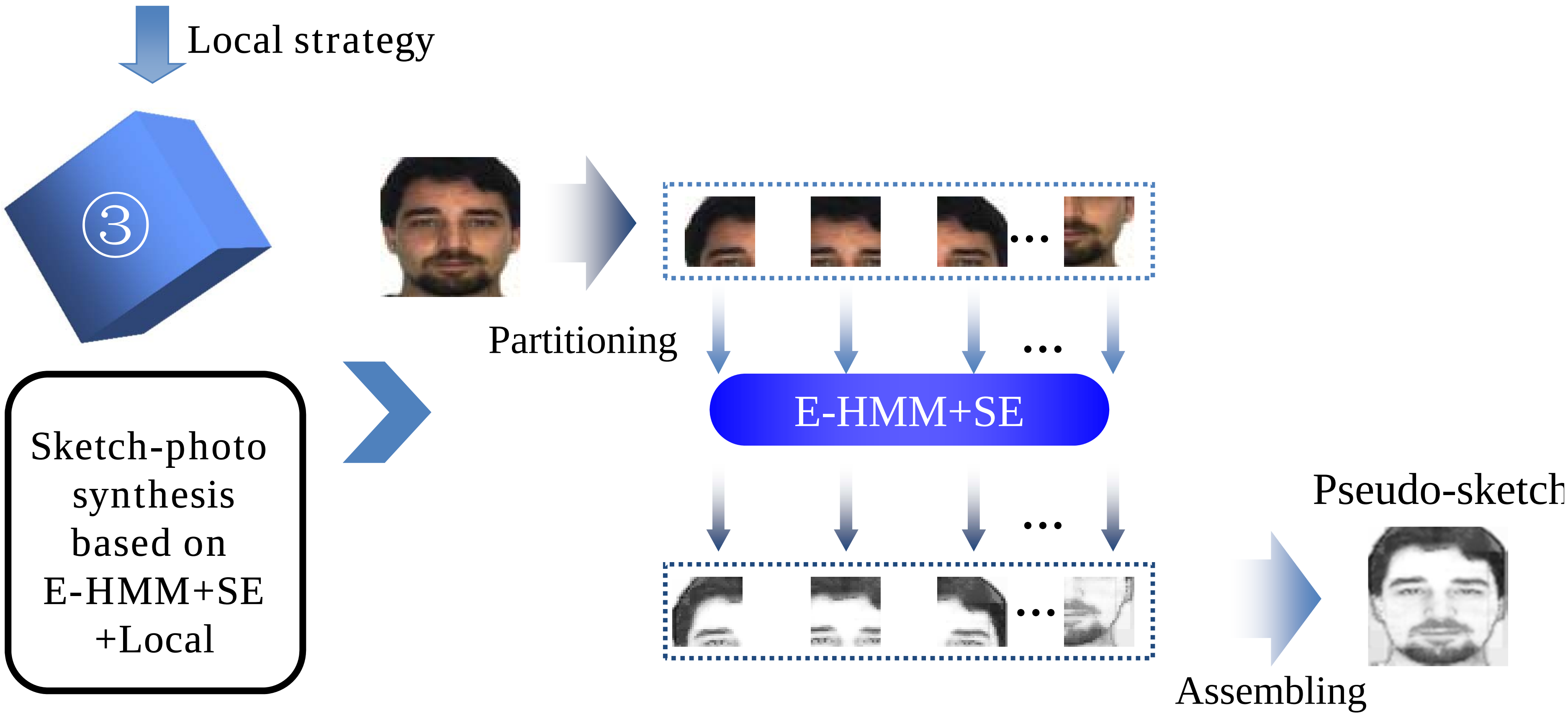
E-HMM-based Methods

Global → Local Strategy

- more specific information
- state estimation of EHMMs

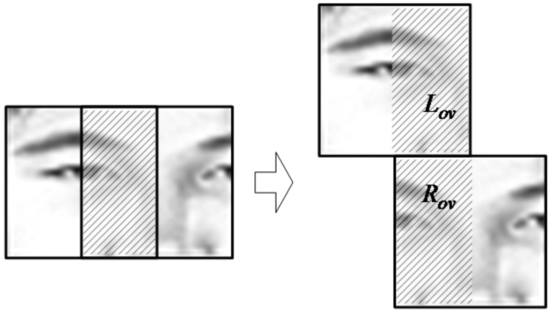
This idea comes from the local linear embedding strategy.

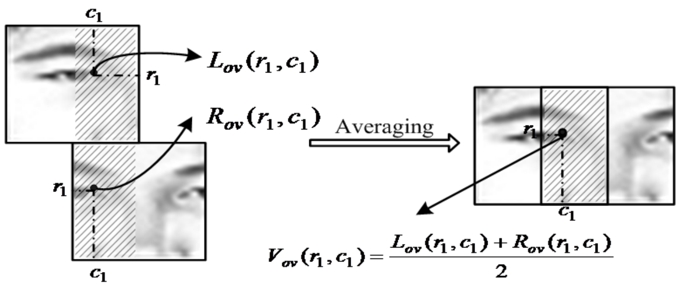
E-HMM-based Methods



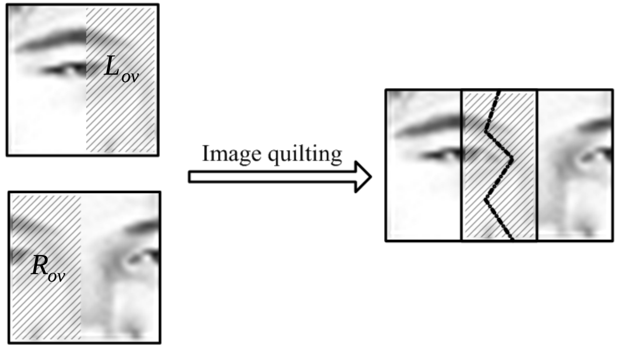
[X. -B. Gao *et al.*, Neurocomputing 08, Signal Processing09]

E-HMM-based Methods





Averaging overlapping areas

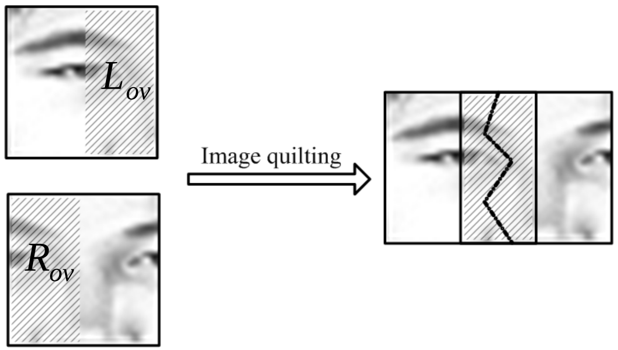


Quilting overlapping areas

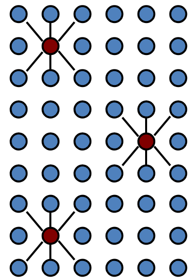
Blurring effect
Blocking effect

E-HMM-based Methods

Image Quilting



Quilting overlapping areas



Graph of overlapping area

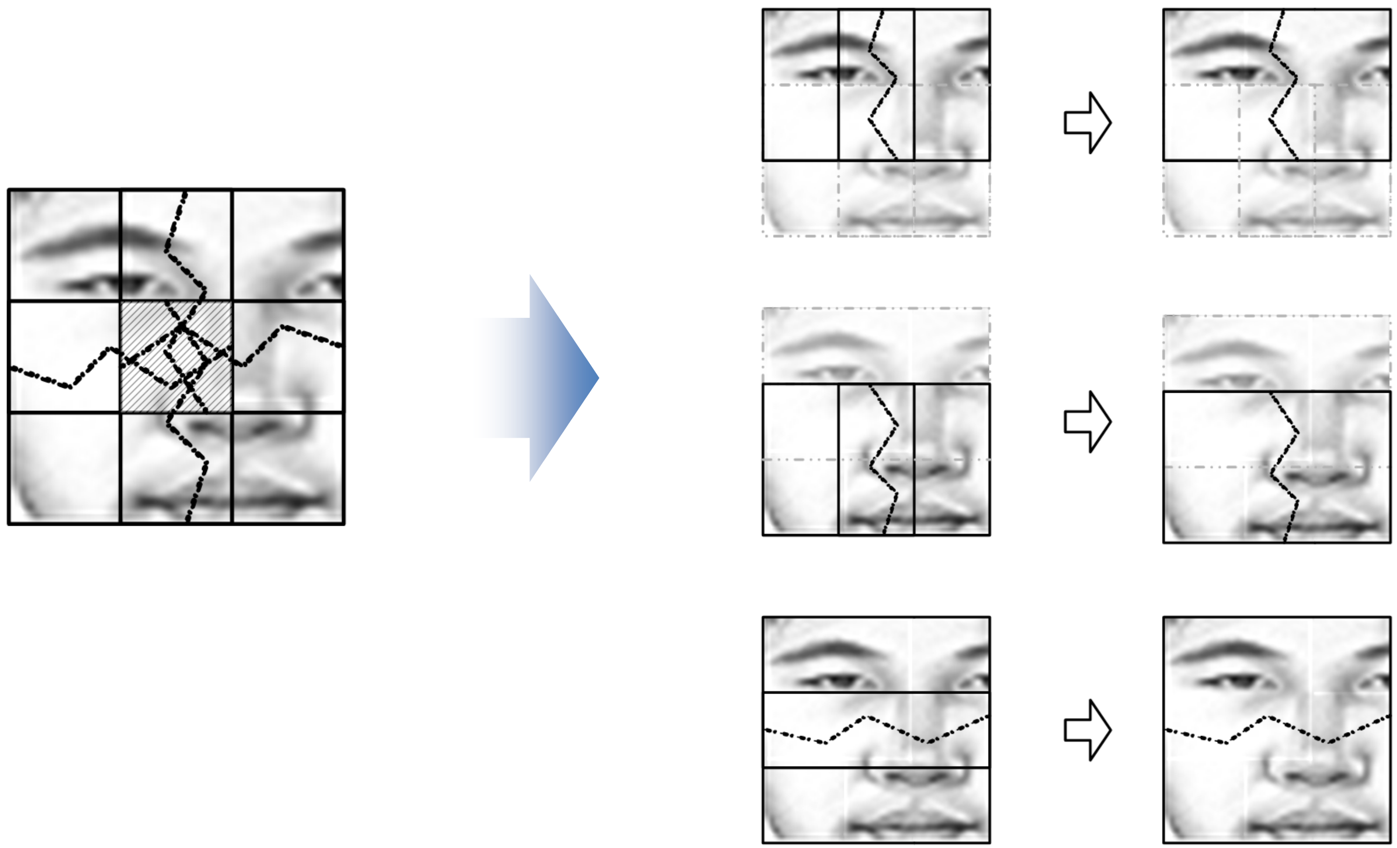
The difference between $L_{ov}(i, y_i)$ and $R_{ov}(i, y_i)$ → The cost of traversing (i, y_i)

Dijkstra algorithm

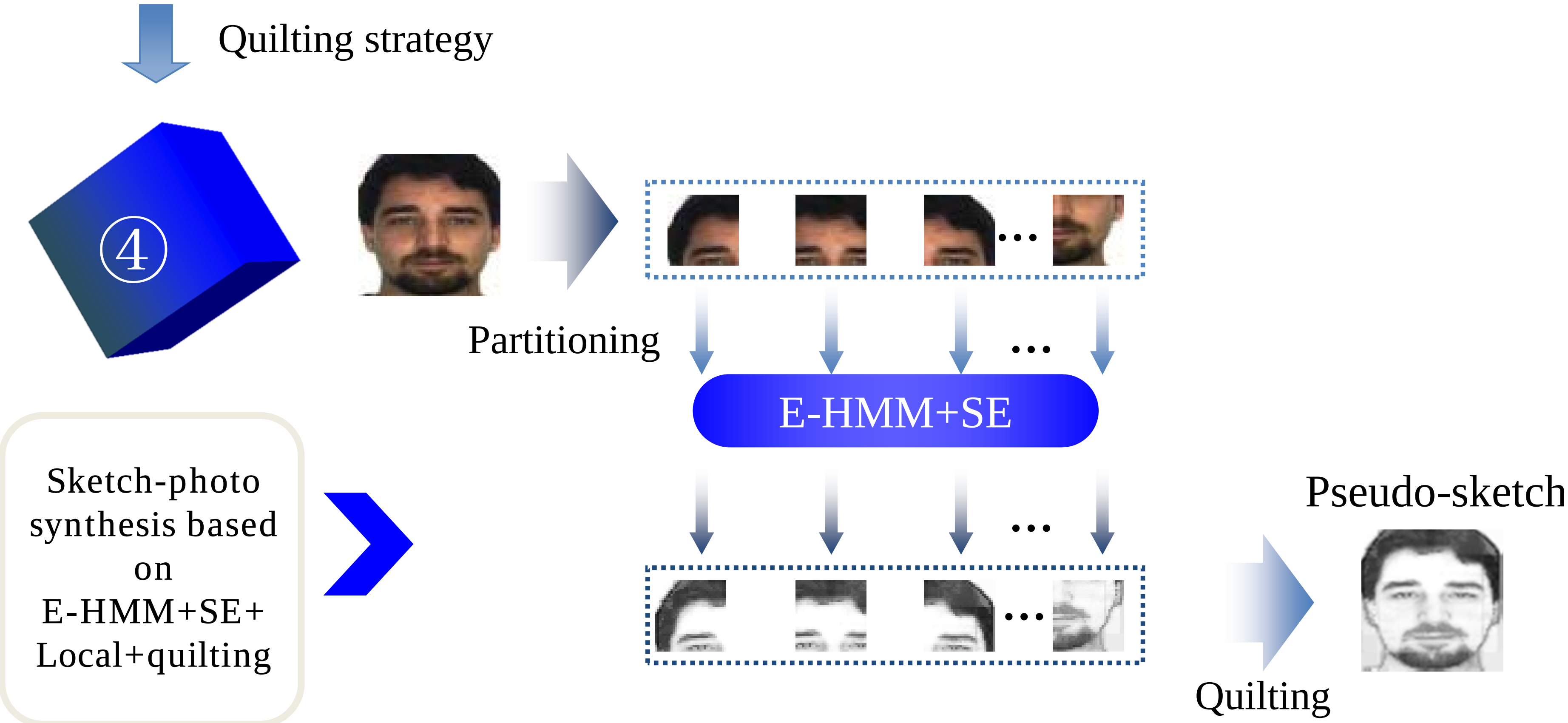
The edge is determined by searching for the minimum cost path $\boldsymbol{E}^* = \{(1, y_1), \cdots, (r, y_r)\}$

$$\boldsymbol{E}^* = \underset{\boldsymbol{E}}{\operatorname{argmin}} \sum\nolimits_{(i,y_i)\in\boldsymbol{E}} |L_{ov}(i, y_i) - R_{ov}(i, y_i)|^2$$

E-HMM-based Methods

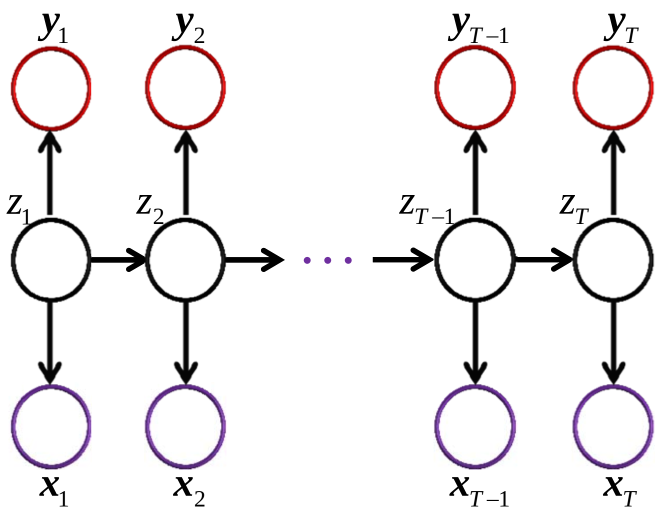


E-HMM-based Methods



[B. Xiao, X. -B. Gao *et al.*, Neurocomputing 2010]

E-HMM-based Methods



Graphical illustration of the model E-HMM. Here $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_T$ and $\boldsymbol{y}_1, \cdots, \boldsymbol{y}_N$ denote the observations extracted from a photo-sketch pair respectively, i.e. $\boldsymbol{o}_i = [\boldsymbol{x}_i; \boldsymbol{y}_i]$, $i = 1, \cdots, T$. $\boldsymbol{z}_1, \cdots, \boldsymbol{z}_N$ are hidden variables

$$\boldsymbol{I}_{out}^* = \underset{\boldsymbol{I}_{out}, \boldsymbol{z}}{\operatorname{argmax}} P(\boldsymbol{I}_{out}, \boldsymbol{z} | \boldsymbol{I}_{in})$$
$$= \underset{\boldsymbol{I}_{out}, \boldsymbol{z}}{\operatorname{argmax}} P(\boldsymbol{I}_{out}, \boldsymbol{z}, \boldsymbol{I}_{in})$$
$$= \underset{\boldsymbol{I}_{out}, \boldsymbol{z}}{\operatorname{argmax}} P(\boldsymbol{z}, \boldsymbol{I}_{in}) P(\boldsymbol{I}_{out} | \boldsymbol{z}, \boldsymbol{I}_{in})$$
$$= \underset{\boldsymbol{I}_{out}, \boldsymbol{z}}{\operatorname{argmax}} P(\boldsymbol{z}, \boldsymbol{I}_{in}) P(\boldsymbol{I}_{out} | \boldsymbol{z})$$

[N. -N. Wang, *et al.*, IJCV13]

$$\boxed{\begin{aligned} \boldsymbol{z}^* &= \underset{\boldsymbol{z}}{\operatorname{argmax}} P(\boldsymbol{O}_{in}, \boldsymbol{z} | \boldsymbol{\lambda}_{\boldsymbol{P}_i}) \\ \boldsymbol{O}_{out}^* &= \underset{\boldsymbol{z}}{\operatorname{argmax}} P(\boldsymbol{O}_{out}, | \boldsymbol{z}^*, \boldsymbol{\lambda}_{S_i}) \end{aligned}}$$

Here $\boldsymbol{\lambda}_{\boldsymbol{P}_i}$ and $\boldsymbol{\lambda}_{S_i}$ represent the joint trained photo and sketch model respectively.

OUTLINE

Motivation & Introduction

Subspace Learning-based Methods

Regression-based Methods

Bayesian Inference-based Methods

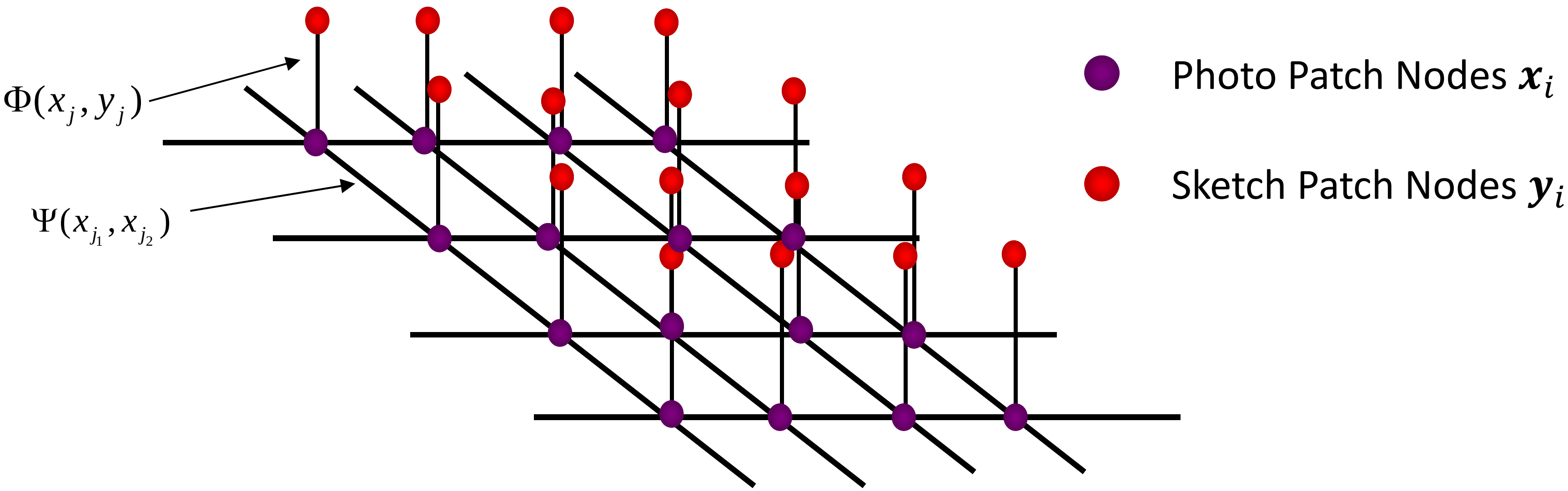
- E-HMM-based Methods
- MRF-based Method
- Transductive Method

Conclusions

MRF-based Methods

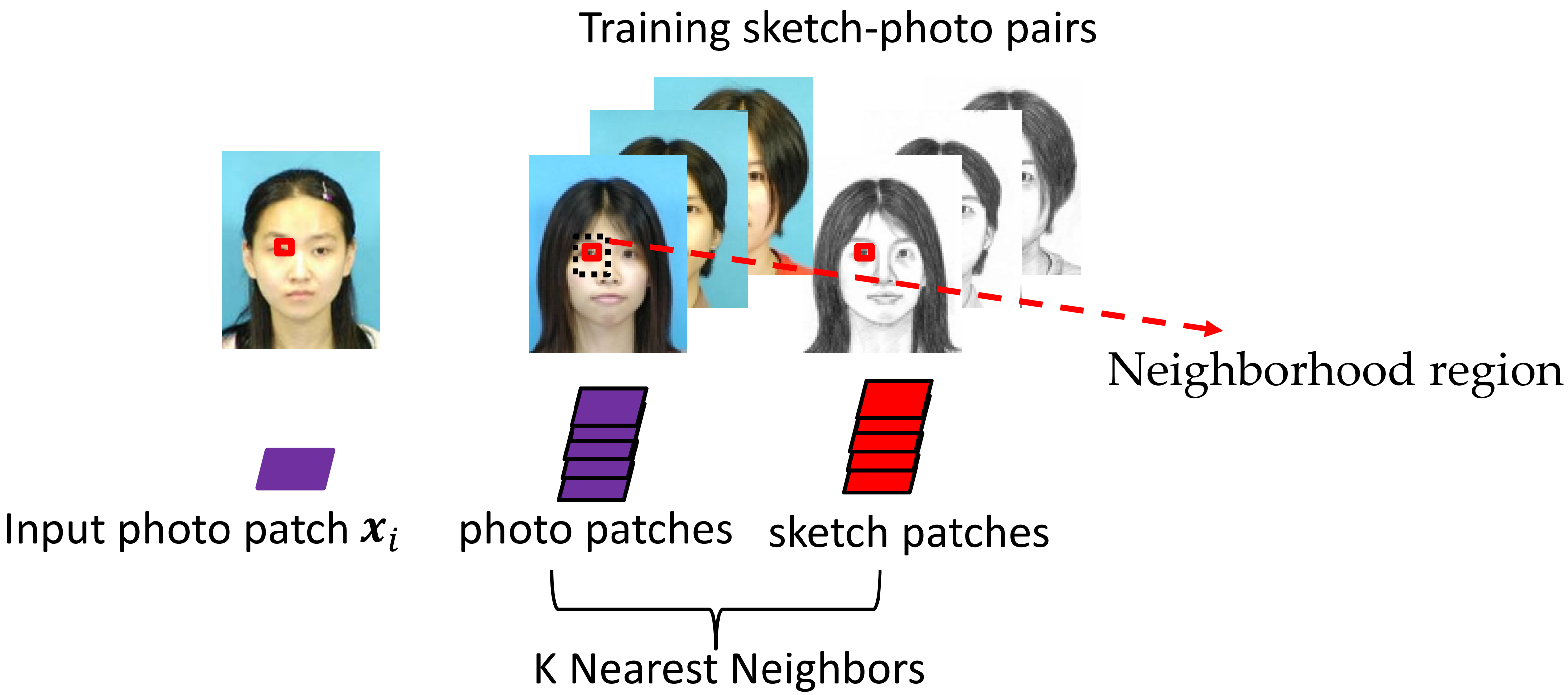
joint probability:

$$P(\boldsymbol{I}_{in}, \boldsymbol{I}_{out}) = P(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N, \boldsymbol{y}_1, \cdots, \boldsymbol{y}_N) = \prod_{(j_1, j_2)} \Psi(\boldsymbol{x}_k, \boldsymbol{y}_k) \prod_{j} \Psi(\boldsymbol{x}_j, \boldsymbol{y}_j)$$



$$P(\boldsymbol{I}_{in} | \boldsymbol{I}_{out}) \propto \prod_{k} \Phi(\boldsymbol{x}_k, \boldsymbol{y}_k) \qquad P(\boldsymbol{I}_{out}) \propto \prod_{(j_1, j_2) \in \mathcal{E}} \Psi(\boldsymbol{x}_{j_1}, \boldsymbol{x}_{j_2})$$

MRF-based Methods

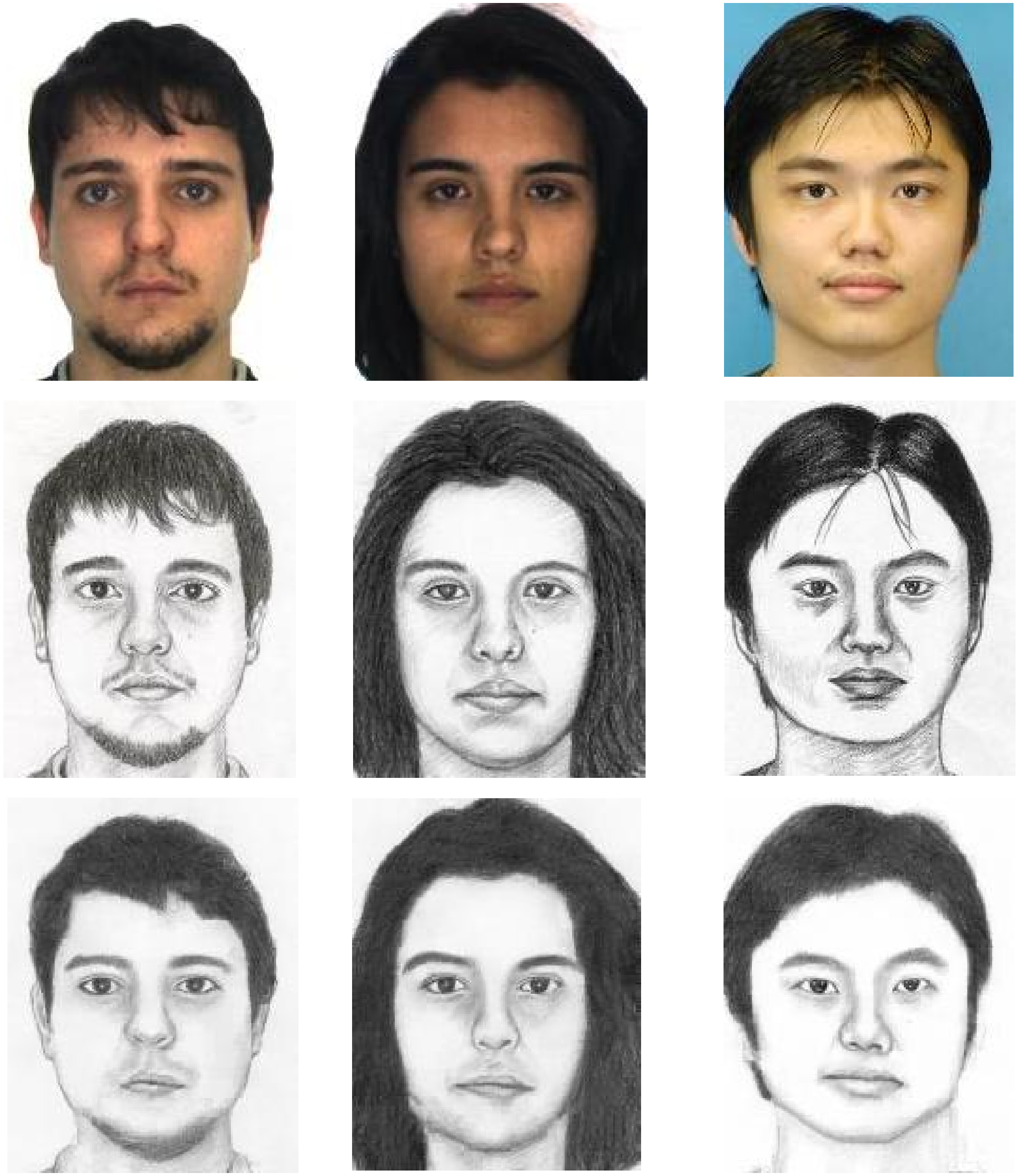


Neighborhood selection

[X. -G. Wang, X. –O. Tang, PAMI 09]

MRF-based Method

Example Results-Synthesized Sketches



Input Photo

Ground Truth

Synthesis Results

MRF-based Method

Example Results-Synthesized Photos

Input Sketch

Ground Truth

Synthesis Results

OUTLINE

Motivation & Introduction

Subspace Learning-based Methods

Regression-based Methods

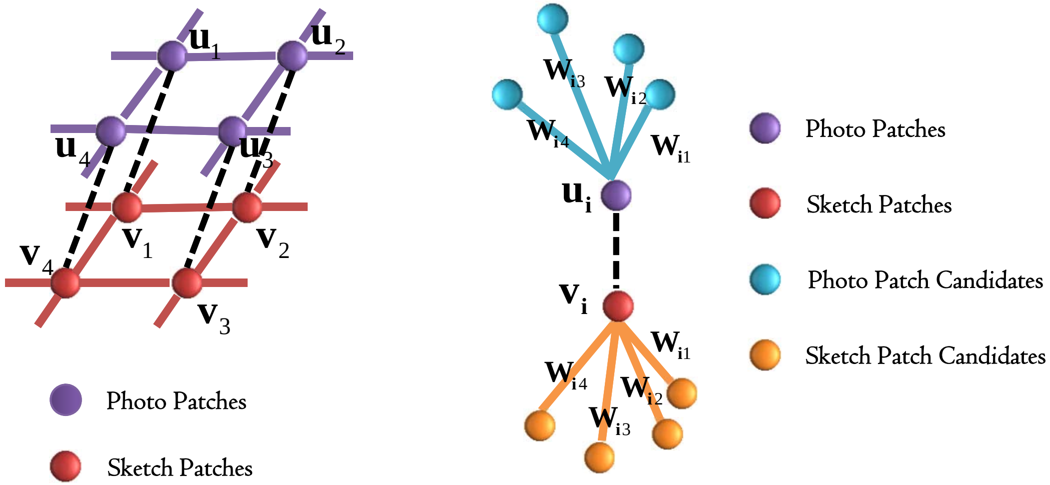
Bayesian Inference-based Methods

- E-HMM-based Methods
- MRF-based Method
- Transductive Method

Conclusions

Transductive Method

All above methods are based on **inductive learning**, which result in **high losses for test samples**. This is because inductive learning minimizes the empirical loss for training samples. **The transductive method** could incorporate the given test samples into the learning process and optimize the performance on these test samples.

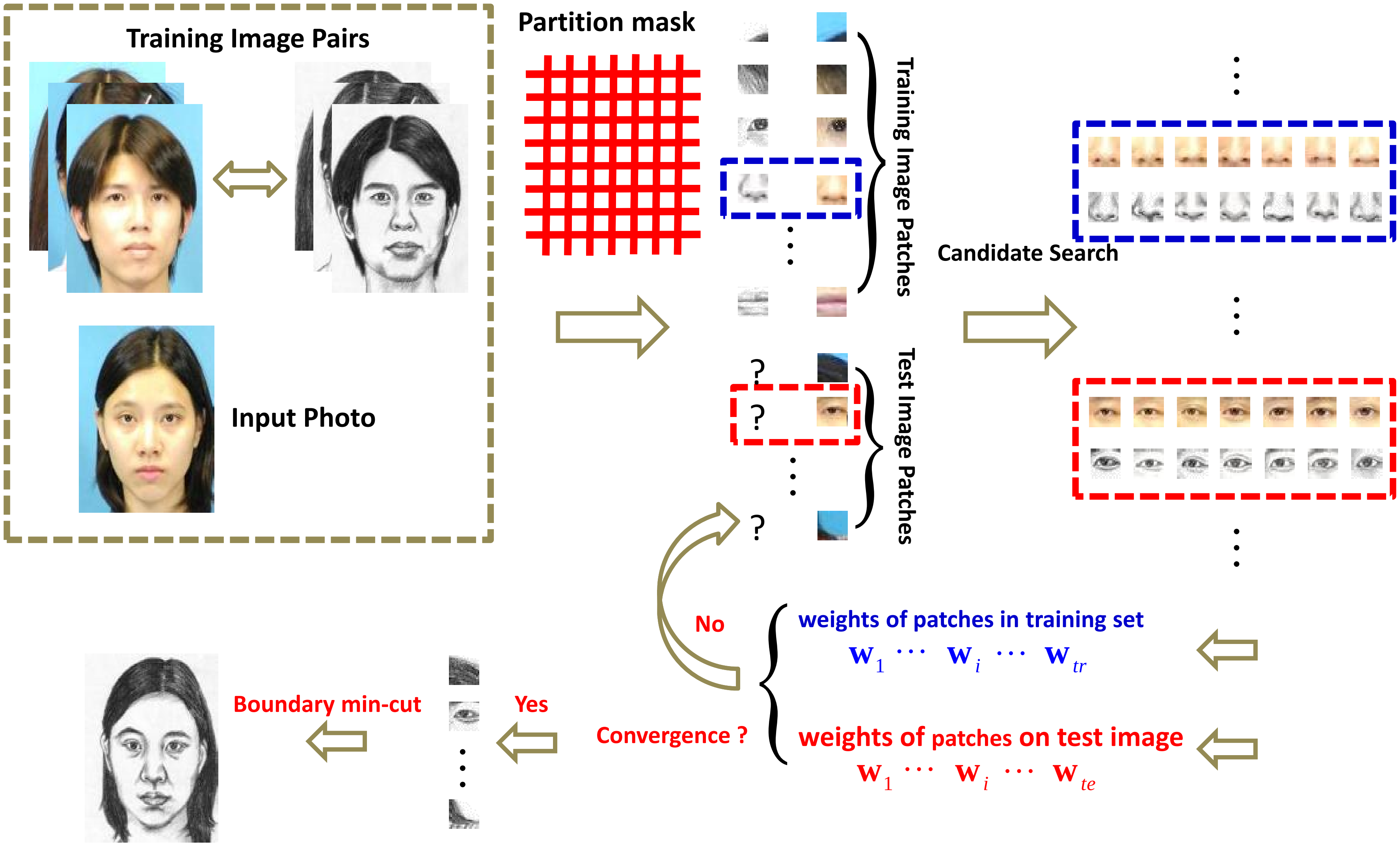


[N. -N. Wang, TNNLS13]

(a) (b)

Illustration of the constructed graph. (a) Graph G = (V, E,W). Photo patches (or sketch patches) can represent either training photo patches (training sketch patches) or test photo patches (target sketch patches) because we will construct the model from the perspective of transductive learning. (b) Illustration of the candidate selection criterion. The number of nearest neighbors is K = 4. Weights on edges illustrate the similarity between a patch and its candidates.

Transductive Method



Framework

Transductive Method

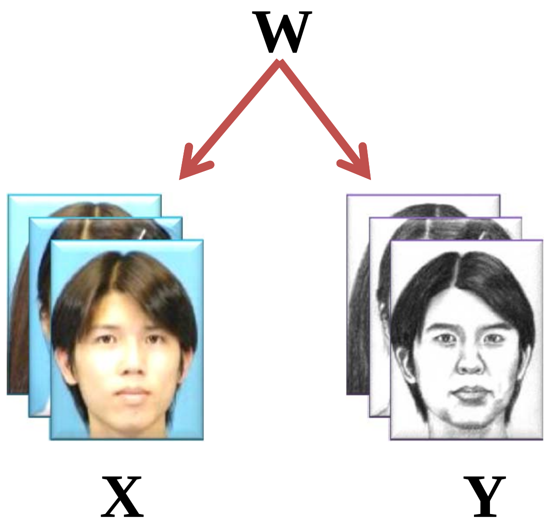


Illustration of the generative process of photo patches and sketch patches from common hidden parameters.

$$P(\boldsymbol{Y},\boldsymbol{X},\boldsymbol{W}) = P(\boldsymbol{Y},\boldsymbol{X}|\boldsymbol{W})P(\boldsymbol{W})$$
$$= P(\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{W})P(\boldsymbol{X}|\boldsymbol{W})P(\boldsymbol{W})$$
$$= P(\boldsymbol{Y}|\boldsymbol{W})P(\boldsymbol{X}|\boldsymbol{W})P(\boldsymbol{W})$$

$$P(\boldsymbol{I}_{in}|\boldsymbol{I}_{out}) \propto P(\boldsymbol{Y}|\boldsymbol{W})P(\boldsymbol{X}|\boldsymbol{W})$$
$$P(\boldsymbol{I}_{out}) \propto P(\boldsymbol{W})$$

$$P(W) \propto \prod_{(i,j)\in\mathcal{E}} \exp\left\{-\frac{\left\|\sum_{k\in\mathcal{N}(i)} \boldsymbol{W}_{ik} {Y_k}^{(i,j)} - \sum_{l\in\mathcal{N}(j)} \boldsymbol{W}_{il} {\boldsymbol{Y}_l}^{(j,i)}\right\|^2}{2\sigma_r^2}\right\} \quad \text{s.t.} \sum_{k\in\mathcal{V}} \boldsymbol{W}_{ik} = 0, \boldsymbol{W}_{ik} > 0 \ \ \forall i \in \mathcal{V}$$

$$P(X|W) \propto \prod_{i\in\mathcal{V}} \exp\left\{-\frac{\left\|\boldsymbol{X}_{i\cdot} - \sum_{j\in\mathcal{N}(i)} \boldsymbol{W}_{ij}\boldsymbol{X}_{j\cdot}\right\|^2}{2\sigma_{dp}^2}\right\}, \qquad P(Y|W) \propto \prod_{i\in\mathcal{V}} \exp\left\{-\frac{\left\|\boldsymbol{Y}_{i\cdot} - \sum_{j\in\mathcal{N}(i)} \boldsymbol{W}_{ij}\boldsymbol{Y}_{j\cdot}\right\|^2}{2\sigma_{ds}^2}\right\}$$

$$P(Y,X,W) \propto \min_{\boldsymbol{W},\boldsymbol{y}_1,\cdots,\boldsymbol{y}_M} tr(Y^T M Y) + \alpha tr(X^T M X) + \beta \sum_{(i,j)\in\mathcal{E}} \left\|\sum_{k\in\mathcal{N}(i)} \boldsymbol{W}_{ik} {Y_k}^{(i,j)} - \sum_{l\in\mathcal{N}(j)} \boldsymbol{W}_{il} {\boldsymbol{Y}_l}^{(j,i)}\right\|^2$$
$$\text{s.t.} \sum_{k\in\mathcal{V}} \boldsymbol{W}_{ik} = 0, \boldsymbol{W}_{ik} > 0 \ \ \forall i \in \mathcal{V}$$

Transductive Method

Optimization

- Fixing $\boldsymbol{W}$, update $\boldsymbol{y}_1, \cdots, \boldsymbol{y}_M$ by solving:

$$\min_{\boldsymbol{y}_1,\cdots,\boldsymbol{y}_M} tr(\boldsymbol{Y}^T \boldsymbol{M}\boldsymbol{Y}) \qquad (1)$$

- Then fixing $\boldsymbol{y}_1, \cdots, \boldsymbol{y}_M$ to be the above obtained value, and update $\boldsymbol{W}$ by solving

$$\min_{\boldsymbol{W}} \|\boldsymbol{U} - \boldsymbol{W}\boldsymbol{U}\|^2 + \beta \sum_{(i,j)\in\mathcal{E}} \left\| \sum\nolimits_{k\in\mathcal{N}(i)} \boldsymbol{W}_{ik} Y_k{}^{(i,j)} - \sum\nolimits_{l\in\mathcal{N}(j)} \boldsymbol{W}_{il} \boldsymbol{Y}_l{}^{(j,i)} \right\|^2$$

$$\text{s.t.} \quad \sum\nolimits_{k\in\mathcal{V}} \boldsymbol{W}_{ik} = 0, \boldsymbol{W}_{ik} > 0 \ \ \forall i \in \mathcal{V} \qquad \text{where } U = [\boldsymbol{X}\ \sqrt{\alpha}\boldsymbol{Y}] \qquad (2)$$

The optimization method for solving (2) can refer to [H. Zhou *et al.* CVPR12]

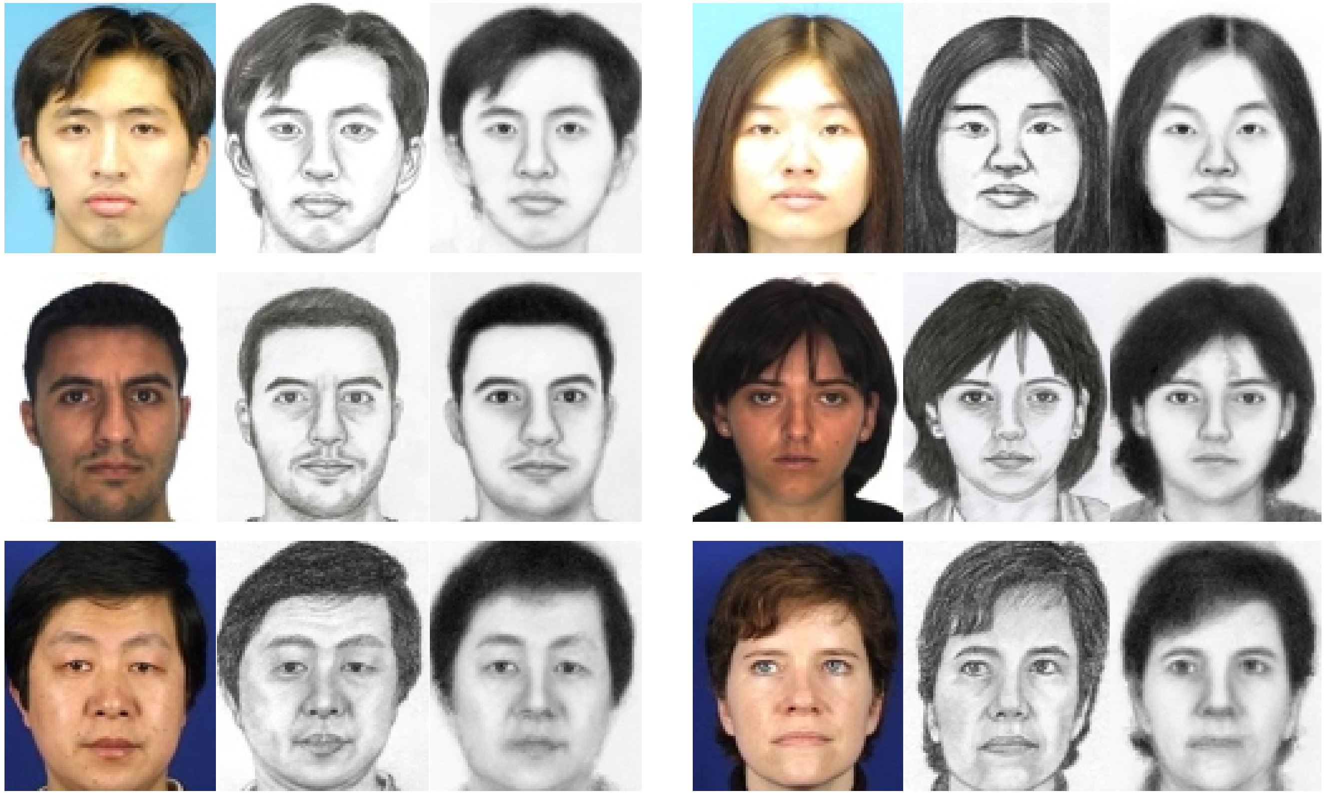
Transductive Method

Effect of neighborhood size K

$K = 1$ $K = 3$ $K = 5$ $K = 10$ $K = 20$ $K = 30$

Transductive Method

Synthesized Sketches

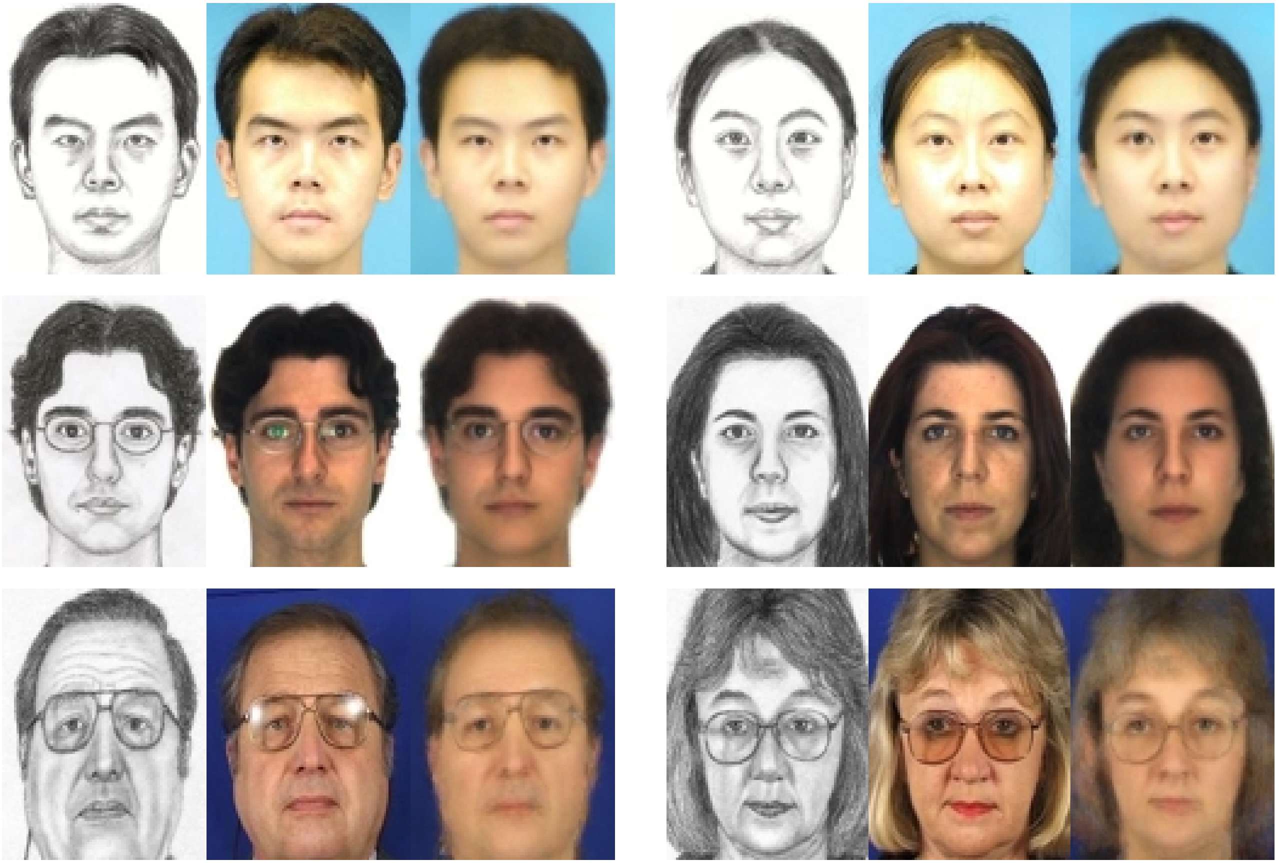


(a) (b) (c) (d) (e) (f)

(a)(d): Input Photos; (c)(f): Output Sketches; (b)(e): Ground Truth

Transductive Method

Synthesized Photos

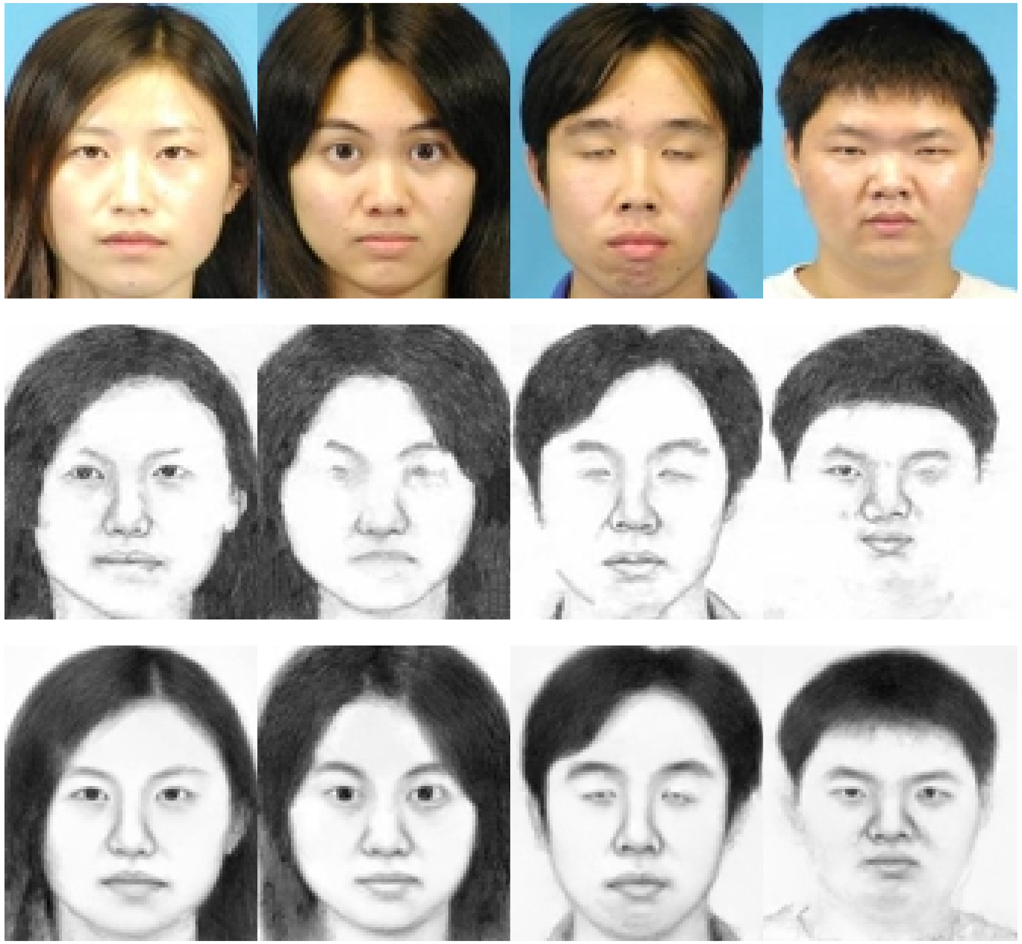


(a) (b) (c) (d) (e) (f)

(a)(d): Input Sketches; (c)(f): Output photos; (b)(e): Ground Truth

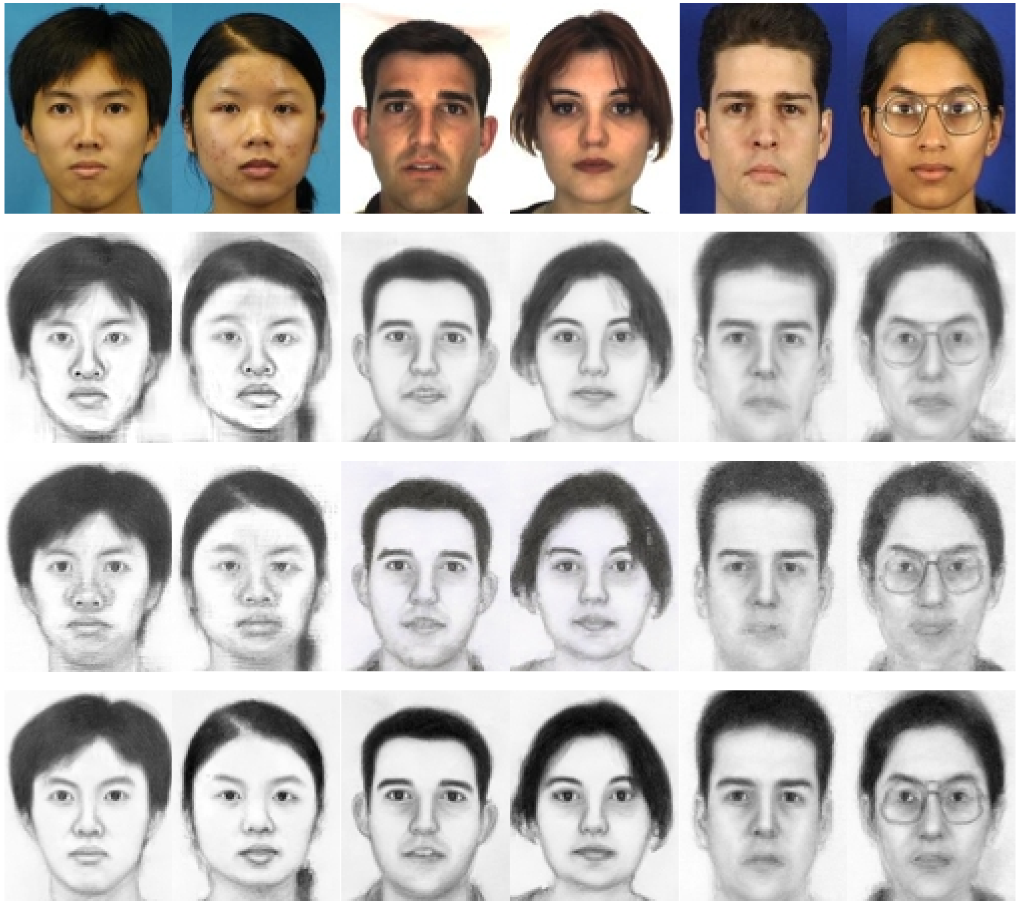
Transductive Method

Comparison with MRF-based Methods



Transductive Method

Comparison with MRF-based Methods



Input Photos

LLE-based method

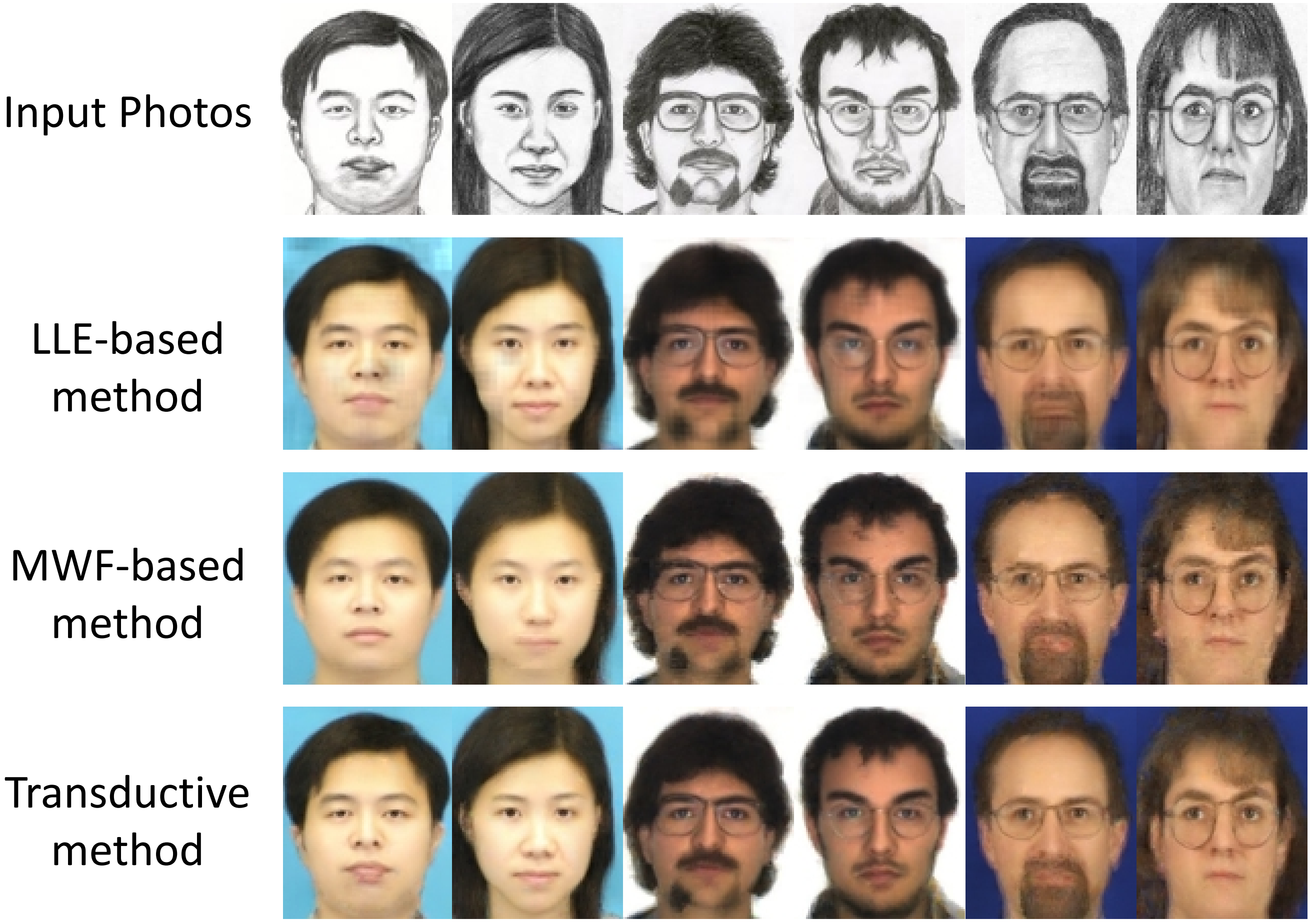
MWF-based method

Transductive method

LLE: Locally Linear Embedding
MWF: Markov Weight Fields

Transductive Method

Comparison with MRF-based Methods



LLE: Locally Linear Embedding
MWF: Markov Weight Fields

Transductive Method

Face Recognition on CUHK Sketch Database

	Eigensketch	MRF-sketch	MRF-photo	T-Sketch	T-Photo
Fisherface	79.7	89.3	93.3	91.3	96.3
NLDA	84.0	90.7	94.7	93.7	96.3
RS-LDA	90.0	93.3	96.3	95.7	97.7

NLDA: null space linear discriminant analysis

RS-LDA: random sampling linear discriminant analysis

Eigensektch: eigensketchtransformation method

MRF-sketch: MRF-based sketch synthesis

MRF-photo: MRF-based photo synthesis

T-sketch: Transductive sketch synthesis

T-photo: Transductive photo synthesis

OUTLINE

Motivation & Introduction

Subspace Learning-based Methods

Regression-based Methods

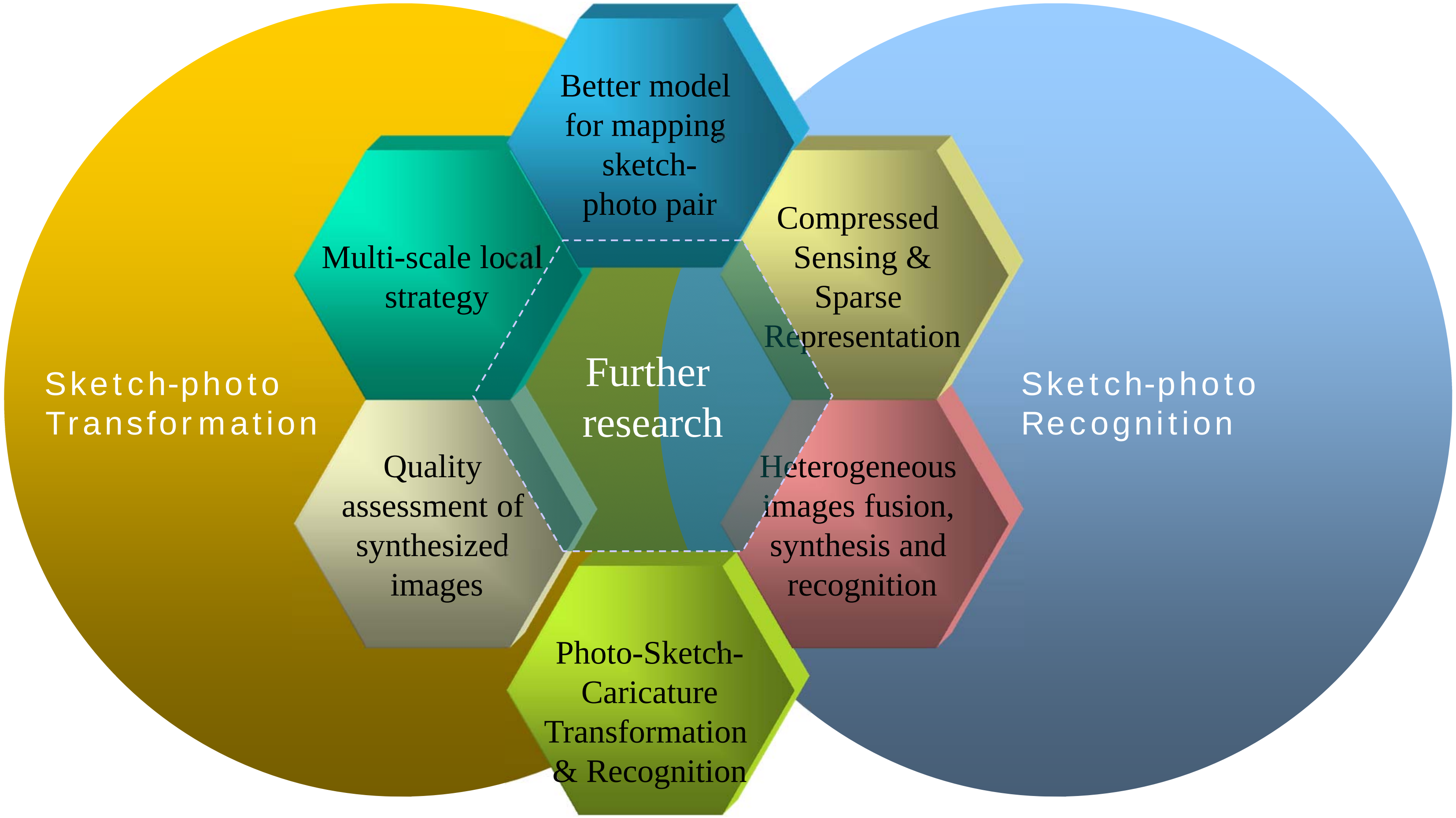
Bayesian Inference-based Methods

➡ Conclusions

Conclusions

- Subspace learning-based and sparse representation-based methods synthesize each image patch independently neglecting the neighboring relation. This results in the incompatibility between neighboring patches. Bayesian inference based approaches construct the model from the neighborhood relation and thus have promising results;
- Linear subspace learning-based methods synthesize a whole image which may lead to some critical local details lost;
- E-HMM-based method are most time-consuming due to the iterative Vieterbi decoding algorithm; However, E-HMM-based methods need most less examples to synthesize images.
- In all methods which explore kNN, the kNN process is the most time consuming part.

Open problems



References

- L. Chang, M. Zhou, Y. Han, and X. Deng, “Face Sketch Synthesis via Sparse Representation,” in *International Conference on Pattern Recognition*, 2010, pp. 2146-2149.
- X. Gao, N. Wang, D. Tao, and X. Li, “Face Sketch-Photo Synthesis and Retrieval Using Sparse Representation ,” *IEEE Transactions on Circuits and Systems for Video Technology*, to appear, 2012
- X. Gao, J. Zhong, J. Li, and C. Tian, “Face Sketch Synthesis Algorithm Based on E-HMM and Selective Ensemble,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 4, pp. 487-496, 2008.
- X. Gao, J. Zhong, D. Tao, and X. Li, “Local Face Sketch Synthesis Learning,” *Neurocomputing*, vol. 71, no. 10-12, pp. 1921-1930, 2008.
- N. Ji, X. Chai, S. Shan, and X. Chen, “Local Regression Model for Automatic Face Sketch Generation,” in *International Conference on Image and Graphics*, 2011, pp. 412-417.
- Y. Li, M. Savvides, and V. Bhagavatula, “Illumination Tolerant Face Recognition Using a Novel Face from Sketch Synthesis Approach and Advanced Correlation Filters,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2006, pp. 357-360.
- Q. Liu, X. Tang, H. Jin, H. Lu, and S. Ma, “A Nonlinear Approach for Face Sketch Synthesis and Recognition,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2005, pp. 1005-1010.
- X. Tang and X. Wang, “Face Sketch Recognition,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 1-7, 2004.
- X. Tang and X. Wang, “Face Sketch Synthesis and Recognition,” *IEEE International Conference on Computer Vision*, 2003, pp. 687-694.
- X. Tang and X. Wang, “Face Photo Recognition Using Sketch,” in *IEEE International Conference on Image Processing*, 2002, pp. 257-260.

References

- S. Wang, L. Zhang, Y. Liang, and Q. Pan, "Semi-Coupled Dictionary Learning with Applications to Image Super-Resolution and Photo-Sketch Synthesis," in *IEEE Conference on Computer Vision and Pattern Recognition*, to appear, 2012.
- N. Wang, X. Gao, D. Tao, and X. Li, "Face Sketch-Photo Synthesis under Multi-dictionary Sparse Representation Framework," in *International Conference on Image and Graphics*, 2011, pp. 82-87.
- N. Wang, D. Tao, X. Gao, X. Li, and J. Li, "Transductive Face Sketch-Photo Synthesis," IEEE Transactions on Neural Networks and Learning Systems, vol. 24, no. 9, 2013.
- N. Wang, D. Tao, X. Gao, X. Li, and J. Li, "A Comprehensive Survey to Face Hallucination," International Journal of Computer Vision, in press, 2013.
- X. Wang and X. Tang, "Face Photo-Sketch Synthesis and Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 11, pp. 1955-1967, 2009.
- B. Xiao, X. Gao, D. Tao, Y. Yuan, and J. Li, "Photo-Sketch Synthesis and Recognition Based on Subspace Learning," *Neurocomputing*, vol. 73, no. 4-6, pp. 840-852, 2010.
- B. Xiao, X. Gao, D. Tao, and X. Li, "A New Approach for Face Recognition by Sketches in Photos," *Signal Processing*, vol. 89, no. 8, pp. 1576-1588, 2009.
- J. Zhang, N. Wang, X. Gao, D. Tao, and X. Li, "Face Sketch-Photo Synthesis Based on Support Vector Regression," in *IEEE International Conference on Image Processing*, 2011, pp. 1149-1152.
- W. Zhang, X. Wang, and X. Tang, "Lighting and Pose Robust Face Sketch Synthesis," in *Proceedings of European Conference on Computer Vision*, 2010, pp. 420-423.
- J. Zhong, X. Gao, and C. Tian, "Face Sketch Synthesis Algorithm Based on E-HMM and Selective Ensemble," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2007, pp. 485-488.
- H. Zhou, Z. Kuang, and K. Wong, "Markov Weight Fields for Face Sketch Synthesis," in *IEEE Conference on Computer Vision and Pattern Recognition*, to appear, 2012.