Heterogeneous Facial Image Synthesis

Nannan Wang, Dacheng Tao, Xinbo Gao, Xuelong Li, Jie Li
What's Heterogeneous Images

Different Modalities

Sketch

Visible

Near-Infrared

TIR

Gray-scale photo

Color photo

Color photo
Contents

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

(1) Conventional Homogeneous Image Processing

Heterogeneous

(2) Fusion

Photograph, Sketch, Oil Painting, Line drawing, Caricature

(3) Synthesis

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

Direct Retrieval

Query Sketch

Photo set
Performance Evaluation

- Eigenfaces: 36.7%
- Laplacianfaces: 27.8%
- KPCA: 15.6%
- OTA: 47.8%

KPCA: Kernel PCA
OTA: Offline Tensor Analysis

Why direct retrieval failed?

All ≤ 50!
Why direct retrieval failed?

**Characteristics:**
- Differently expressed
- Highly Rendered
- Bold Exaggerated

**Challenges:**
- Complicated Mapping
- Diverse Quality Assessment Metrics
- Difficult to analyse their contents
Solutions—Sketch Synthesis

Retrieval

Query Sketch

Photo set

Synthesis

Synthesis

Synthesis
Solutions—Photo Synthesis

Retrieval

Query Photo

Synthesis

Photo set
Development Timeline of Representative Methods

Subspace learning-based methods

- Eigensketch transformation
- LLE-based
- Hybrid subspace

Regression-based methods

- Chang et al. ICPR
- Ji et al. ICIG

Graphical models-based methods

- MRF-based methods
- SFS-based methods
- E-HMM-based methods
- MWF-based method
- Transductive methods

Notes:

LLE: Locally Linear Embedding  MWF: Markov Weight Fields
General Pipeline of Sketch-Photo Synthesis

Photo transformation
Sketch transformation

Machine learning
mapping

Photo
Sketch

Transformation stage
Learning stage
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 achieved 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:

$$ Au_i = \lambda Bu_i $$

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

◆ 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

Principal Component Analysis

Training image pairs

Source input

\[ P_r = P c_p = \sum_{i=1}^{M} c_{p_i} P_i \]

\[ S_r = S c_p = \sum_{i=1}^{M} c_{p_i} S_i \]

[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

\[
\min_Y \sum_i \left| Y_i - \sum_j W_{ij} Y_j \right|^2
\]

\[
\min_W \sum_i \left| X_i - \sum_j W_{ij} X_j \right|^2
\]

[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

\[
\min_c \|c\|_0 + \lambda \|Ac - x\|_2
\]

Sparse coefficient vector
Overcompleted dictionary
A signal

\[
\min_c \|c\|_1 + \lambda \|Ac - x\|_2
\]

NP-hard!
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)

[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

(a) Initial estimation

(b) Residue compensation

Support Vector Regression (SVR)

Sparse feature selection (SFS)

Source input

Target sketch

Training Image Pairs

[S.-B. Gao et al., ICIG2011, ICIP2011, CSVT12, PRL13]
SFS-based Method

Partition mask

Training image pairs

$K$ is adaptively determined

Sparse Representation

$\arg\min_w \|f - D_p w\|_2 + \lambda \|w\|_1$

$w = (w_1, \ldots, w_K)$

Linear combination

Initial Estimation

Source input

Candidate Search

$[X.-B. Gao et al CSVT12]$
SFS-based Method

Training stage

Partition mask

Training image pairs

Partition mask

Grouping

\((S_1, P_1) \rightarrow SVR_1\)

\((S_2, P_2) \rightarrow SVR_2\)

\(\vdots\)

\((S_M, P_M) \rightarrow SVR_M\)

Synthesis Stage

Partition mask

Source input

Partition mask

Hallucinated information

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

<table>
<thead>
<tr>
<th>Method</th>
<th>LLE</th>
<th>E-HMM</th>
<th>SFS</th>
<th>SFS-SVR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognition Rate(%)</td>
<td>84</td>
<td>87</td>
<td>91</td>
<td>93</td>
</tr>
</tbody>
</table>

### Image Quality Assessment (IQA)

<table>
<thead>
<tr>
<th>Method</th>
<th>LLE</th>
<th>E-HMM</th>
<th>SFS</th>
<th>SFS-SVR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0921</td>
<td>0.0948</td>
<td>0.0956</td>
<td>0.0972</td>
</tr>
</tbody>
</table>

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

- **Training image pairs**
- **Source input**
- **Candidate Search**
- **Partition mask**
- **Local regression model**
- **Linear combination**
- **Final output**

\[ \mathbf{w} = (w_1, \ldots, w_K) \]

[N.-Y. Ji et al., ICIG11]
Local Regression-based Method

Regression methods

- K Nearest Neighbor (KNN)
  \[ w^{kNN} = \begin{cases} s(p, p_j), & \text{if } KNN \\ 0, & \text{otherwise} \end{cases} \]
  \( s(\cdot) \) is the similarity measurement metric

- Least Squares (LS)
  \[ w^{LS} = \arg\min_w \left| p - \sum w_i p_i \right|^2 \]

- Ridge Regression (RR)
  \[ w^{LS} = \arg\min_w \left| p - \sum w_i p_i \right|^2 + \lambda |w|^2 \]

- Lasso
  \[ w^{LS} = \arg\min_w \left| p - \sum w_i p_i \right|^2 + \lambda |w|^1 \]

[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:

$$I^*_{{out}} = \arg \max_{I_{out}} P(I_{out} | I_{in})$$

$$= \arg \max_{I_{out}} P(I_{in} | I_{out}) P(I_{out})$$

[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:

\[ \lambda = (\Pi, A, B, N, M) \]

<table>
<thead>
<tr>
<th>Param.</th>
<th>Name</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>State Number</td>
<td>Number of Jar: ( N = 3 )</td>
</tr>
<tr>
<td>( M )</td>
<td>Number of Observation Values</td>
<td>Color Number of Balls: ( M = 4 )</td>
</tr>
<tr>
<td>( A )</td>
<td>State Transition Probability Matrix</td>
<td>( P{\text{Jar } i \rightarrow \text{Jar } j}, i,j = 1,2,3 )</td>
</tr>
<tr>
<td>( B )</td>
<td>Observation Probability Matrix</td>
<td>( P{\text{Ball.color}_\text{Jar } i} )</td>
</tr>
<tr>
<td>( \Pi )</td>
<td>Initial State Distribution</td>
<td>( P{\text{Ball}_1 \text{ from Jar } i} )</td>
</tr>
</tbody>
</table>
E-HMM-based Methods

1D-HMM

- Forehead
- Eyes
- Nose
- Mouth
- Chin

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

2D-HMM

- High computational complexity

Symbols:
- $s_{0,0}$
- $s_{0,1}$
- $s_{1,0}$
- $s_{1,1}$

Graphical representation of 1D-HMM and 2D-HMM models.
E-HMM-based Methods

Pseudo 2D HMM

- Forehead
- Eyes
- Nose
- Mouth
- Chin

Flaw

- The signal is transformed to 1D observation sequence with a zigzag fashion.
E-HMM-based Methods

E-HMM Structure for Face Image

**Horizontal direction**

\[
\lambda = (\Pi_s, A_s, \Lambda_e, N_s)
\]

- Super states
- Embedded states

**Vertical direction**

- Parameters:
  - Initial distribution \( \Pi_s \)
  - State transition matrix \( A_s \)

**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

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

Foundation of the E-HMM Theory

(Q1) Computing the output probability: given the image observation sequence $O = (o_1, o_2, \ldots, o_T)$ and the E-HMM model $\lambda = \{\Pi_s, A_s, \Lambda_e\}$, how to get the output probability $P(O|\lambda)$

(Q2) Decoding the state sequence: with the image observation sequence $O = (o_1, o_2, \ldots, o_T)$ and the E-HMM model $\lambda = \{\Pi_s, A_s, \Lambda_e\}$, how to determine the optimal state sequence $Q = (q_1, q_2, \ldots, q_T)$ and the mixture indexes $M = (m_1, m_2, \ldots, m_T)$

(Q3) Estimating the model parameters: how to adjust the parameters of the E-HMM model $\lambda = \{\Pi_s, A_s, \Lambda_e\}$ to maximize $P(O|\lambda)$

Forward-Backward Algorithm

Embedded Viterbi Algorithm

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

Sketch-photo synthesis based on E-HMM

Baum-Welch algorithm

Joint-training

Transform

$P_1$ → $S_1$

$P_2$ → State sequence → Mixture indices → Sketch synthesis

[J.Zhong, X.-B. Gao et al., ICASSP07]
E-HMM-based Methods

Joint-training Process

Training image A
Feature extraction

Training Image B
Feature extraction

Composition

Decomposition

E-HMM-A
E-HMM-B

EM algorithm

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

[J. Zhong, X.-B. Gao et al., ICASSP07]
Basic idea of joint-training

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.
Selective Ensemble Learning

Criteria of individual selection

- Individuals have good performance
- Individuals complement each other well

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

Sketch-photo synthesis based on E-HMM+SE

Photo set

Photo-sketch pairs with bigger similarity

Decode

Reconstruct

Pseudo-sketch

Normalization

Synthesized sketch

[ X. -B. Gao et al., TCSVT 08]
E-HMM-based Methods

Global $\rightarrow$ Local Strategy

- more specific information
- state estimation of EHMMs

This idea comes from the local linear embedding strategy.
E-HMM-based Methods

Sketch-photo synthesis based on E-HMM+SE + Local

Local strategy

Partitioning

E-HMM+SE

Assembling

Pseudo-sketch

[X.-B. Gao et al., Neurocomputing 08, Signal Processing 09]
E-HMM-based Methods

Averaging overlapping areas

Blurring effect
Blocking effect

Quilting overlapping areas
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 $(i, y_i)$

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

$$E^* = \arg \min_{E} \sum_{(i, y_i) \in E} |L_{ov}(i, y_i) - R_{ov}(i, y_i)|^2$$
E-HMM-based Methods
E-HMM-based Methods

Quilting strategy

Partitioning

E-HMM+SE

Pseudo-sketch

Quilting

Sketch-photo synthesis based on E-HMM+SE+Local+quilting

[B. Xiao, X.-B. Gao et al., Neurocomputing 2010]
E-HMM-based Methods

Graphical illustration of the model E-HMM. Here $x_1, \ldots, x_T$ and $y_1, \ldots, y_N$ denote the observations extracted from a photo-sketch pair respectively, i.e. $o_i = [x_i; y_i], i = 1, \ldots, T$. $z_1, \ldots, z_N$ are hidden variables.

$z^* = \underset{z}{\text{argmax}} P(O_{in}, z | \lambda_{p_i})$

$O_{out}^* = \underset{z}{\text{argmax}} P(O_{out} | z^*, \lambda_{S_i})$

Here $\lambda_{p_i}$ and $\lambda_{S_i}$ represent the joint trained photo and sketch model respectively.

$I_{out}^* = \underset{I_{out}, z}{\text{argmax}} P(I_{out}, z | I_{in})$

$= \underset{I_{out}, z}{\text{argmax}} P(I_{out}, z, I_{in})$

$= \underset{I_{out}, z}{\text{argmax}} P(z, I_{in})P(I_{out} | z, I_{in})$

$= \underset{I_{out}, z}{\text{argmax}} P(z, I_{in})P(I_{out} | z)$

[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
MRF-based Methods

Joint probability:

\[ P(I_{in}, I_{out}) = P(x_1, \ldots, x_N, y_1, \ldots, y_N) = \prod_{(j_1, j_2)} \Psi(x_k, y_k) \prod_j \Psi(x_j, y_j) \]

\[ P(I_{in}|I_{out}) \propto \prod_k \Phi(x_k, y_k) \]

\[ P(I_{out}) \propto \prod_{(j_1, j_2) \in \mathcal{E}} \Psi(x_{j_1}, x_{j_2}) \]
MRF-based Methods

Training sketch-photo pairs

Input photo patch $x_i$

photo patches

K Nearest Neighbors

sketch patches

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.

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.

[N. -N. Wang, TNNLS13]
Transductive Method

Training Image Pairs

Input Photo

Partition mask

Training Image Patches

Test Image Patches

Candidate Search

Boundary min-cut

Yes

Convergence?

weights of patches in training set
\[ W_1 \cdots W_i \cdots W_{tr} \]

weights of patches on test image
\[ W_1 \cdots W_i \cdots W_{te} \]

Framework
Transductive Method

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

\[
\]

\[
= P(Y | X, W) P(X | W) P(W)
\]

\[
= P(Y | W) P(X | W) P(W)
\]

\[
P(I_{in} | I_{out}) \propto P(Y | W) P(X | W)
\]

\[
P(I_{out}) \propto P(W)
\]

\[
P(W) \propto \prod_{(i,j) \in E} \exp \left\{ -\frac{\| \sum_{k \in N(i)} W_{ik} Y_k^{(i,j)} - \sum_{l \in N(j)} W_{il} Y_l^{(j,i)} \|^2}{2\sigma_i^2} \right\} \text{ s.t. } \sum_{k \in V} W_{ik} = 0, W_{ik} > 0 \forall i \in V
\]

\[
P(X | W) \propto \prod_{i \in V} \exp \left\{ -\frac{\| X_i - \sum_{j \in N(i)} W_{ij} X_j \|^2}{2\sigma_{dp}^2} \right\}
\]

\[
P(Y | W) \propto \prod_{i \in V} \exp \left\{ -\frac{\| Y_i - \sum_{j \in N(i)} W_{ij} Y_j \|^2}{2\sigma_{ds}^2} \right\}
\]

\[
P(Y, X, W) \propto \min_{W, Y_1, \ldots, Y_M} tr(Y^T MY) + \alpha tr(X^T MX) + \beta \sum_{(i,j) \in E} \left\| \sum_{k \in N(i)} W_{ik} Y_k^{(i,j)} - \sum_{l \in N(j)} W_{il} Y_l^{(j,i)} \right\|^2
\]

\[
\text{s.t. } \sum_{k \in V} W_{ik} = 0, W_{ik} > 0 \forall i \in V
\]
Transductive Method

Optimization

- Fixing $W$, update $y_1, \ldots, y_M$ by solving:
  $$\min_{y_1, \ldots, y_M} \operatorname{tr}(Y^TMY) \quad (1)$$

- Then fixing $y_1, \ldots, y_M$ to be the above obtained value, and update $W$ by solving
  $$\min_W \|U - WU\|^2 + \beta \sum_{(i,j) \in E} \left\| \sum_{k \in \mathcal{N}(i)} w_{ik}Y_k^{(i,j)} - \sum_{l \in \mathcal{N}(j)} w_{il}Y_l^{(i,i)} \right\|^2$$
  $\quad $ s.t. $\sum_{k \in \mathcal{V}} w_{ik} = 0, w_{ik} > 0 \ \forall i \in \mathcal{V}$

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

Effect of neighborhood size $K$
Transductive Method

Synthesized Sketches

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

Synthesized Photos

(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

Input Photos

LLE-based method

MWF-based method

Transductive method

LLE: Locally Linear Embedding
MWF: Markov Weight Fields
## Transductive Method

## Face Recognition on CUHK Sketch Database

<table>
<thead>
<tr>
<th>Method</th>
<th>Eigensketch</th>
<th>MRF-sketch</th>
<th>MRF-photo</th>
<th>T-Sketch</th>
<th>T-Photo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fisherface</td>
<td>79.7</td>
<td>89.3</td>
<td>93.3</td>
<td>91.3</td>
<td>96.3</td>
</tr>
<tr>
<td>NLDA</td>
<td>84.0</td>
<td>90.7</td>
<td>94.7</td>
<td>93.7</td>
<td>96.3</td>
</tr>
<tr>
<td>RS-LDA</td>
<td>90.0</td>
<td>93.3</td>
<td>96.3</td>
<td>95.7</td>
<td>97.7</td>
</tr>
</tbody>
</table>

NLDA: null space linear discriminant analysis  
RS-LDA: random sampling linear discriminant analysis  
Eigensketch: 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

- Quality assessment of synthesized images
- Multi-scale local strategy
- Better model for mapping sketch-photo pair
- Compressed Sensing & Sparse Representation
- Heterogeneous images fusion, synthesis and recognition
- Sketch-photo Transformation
- Further research
- Sketch-photo Caricature Transformation & Recognition
- Sketch-photo Recognition
References


References

- 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.