Leveraging Geometry and Appearance Cues for Recognizing Family Photos

Xiaolong Wang， Guodong Guo*， Rohith MV and Chandra Kambhamettu

VIMs Lab, University of Delaware, Newark, USA

*West Virginia University, USA

Problem description

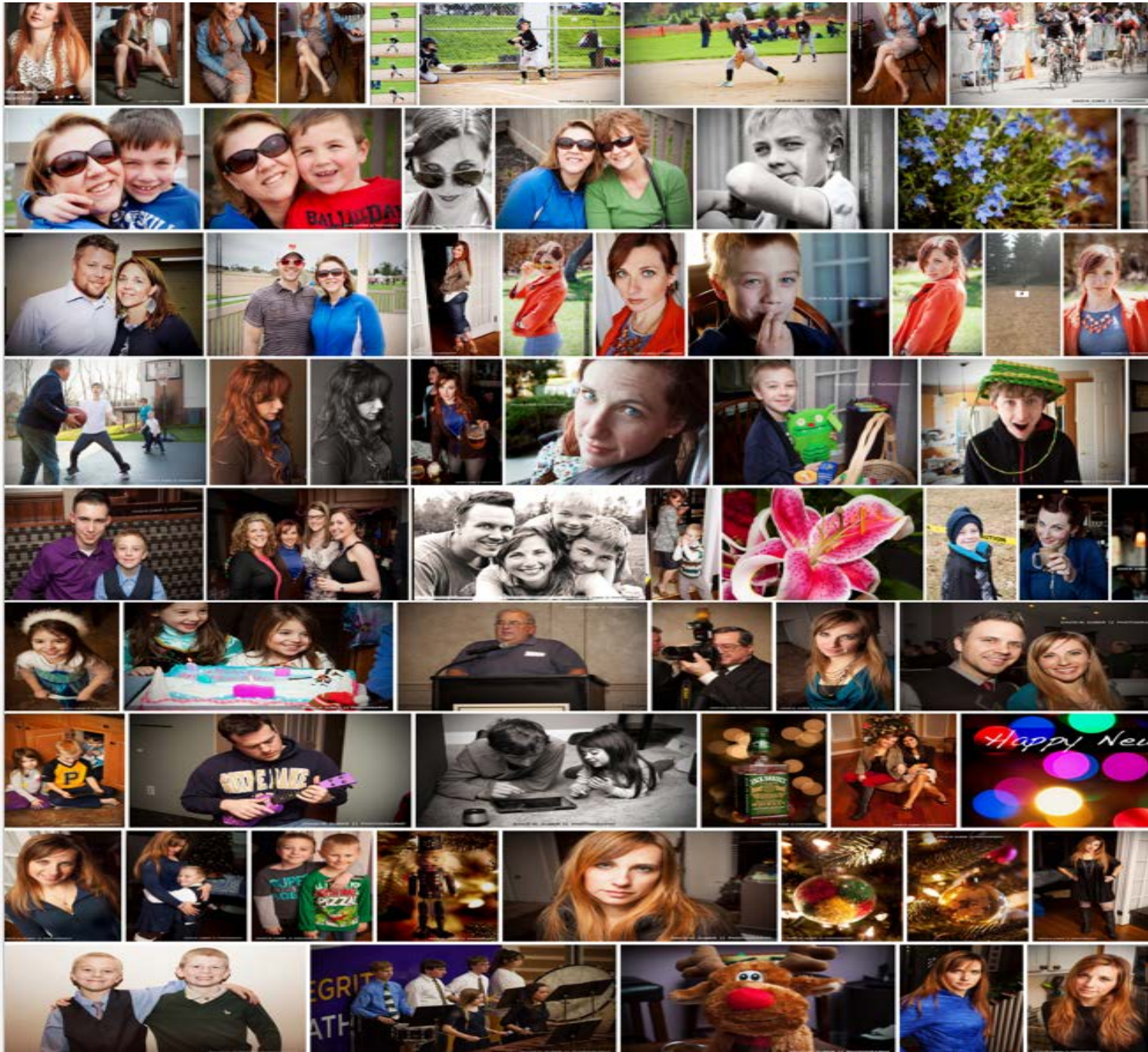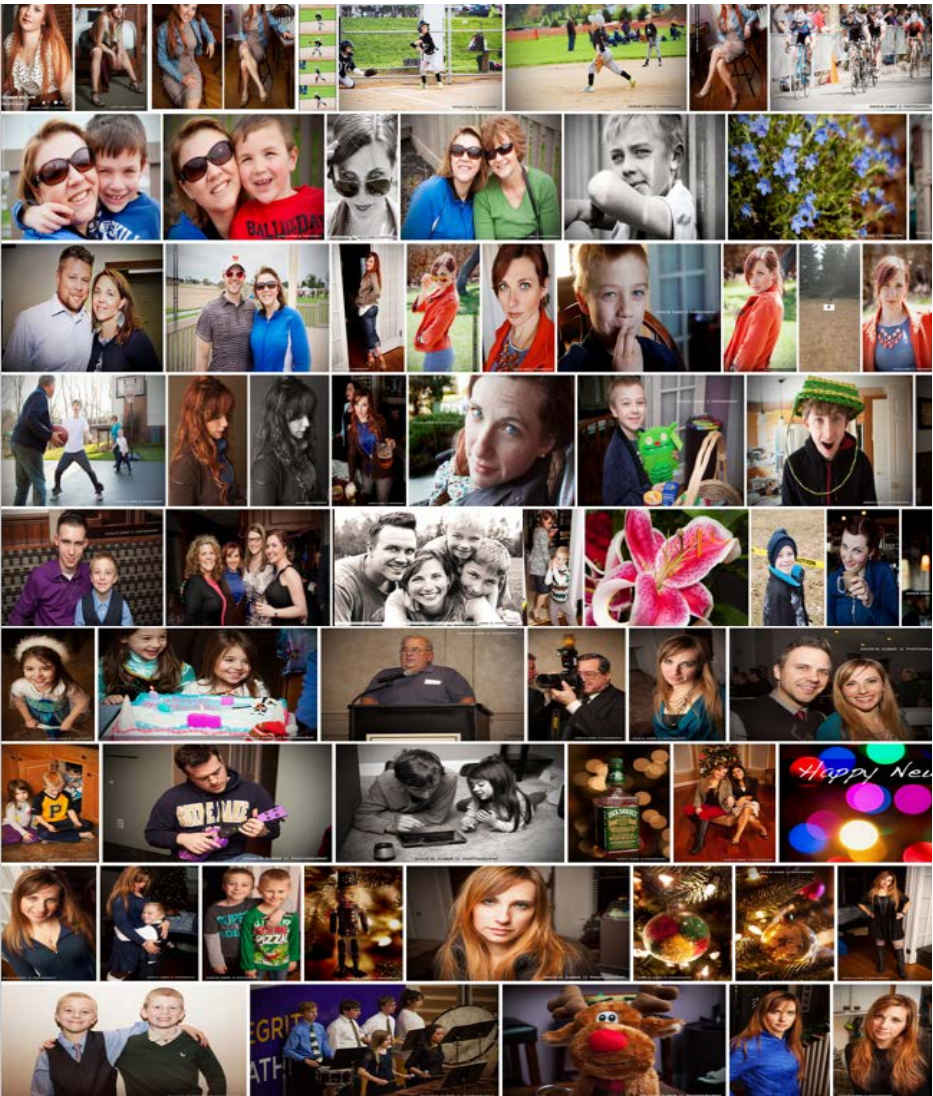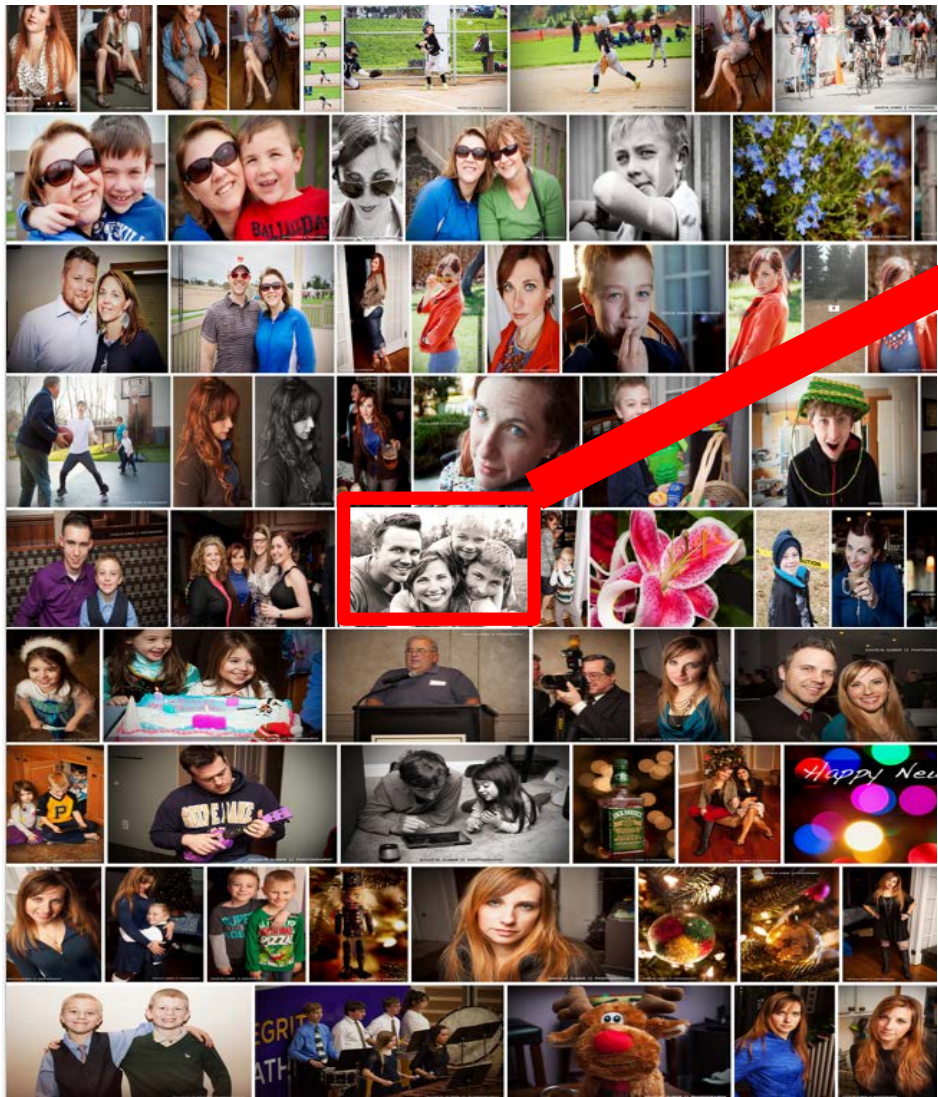# Problem description

Is it possible?

Photo labeling

Photo labeling

Photo labeling

Photo labeling

Photo labeling

Photo labeling

Related work

- Chen et al. Discovering Informative Social Subgraphs and Predicting Pairwise Relationships from Group Photo, MM, 2012.

Related work

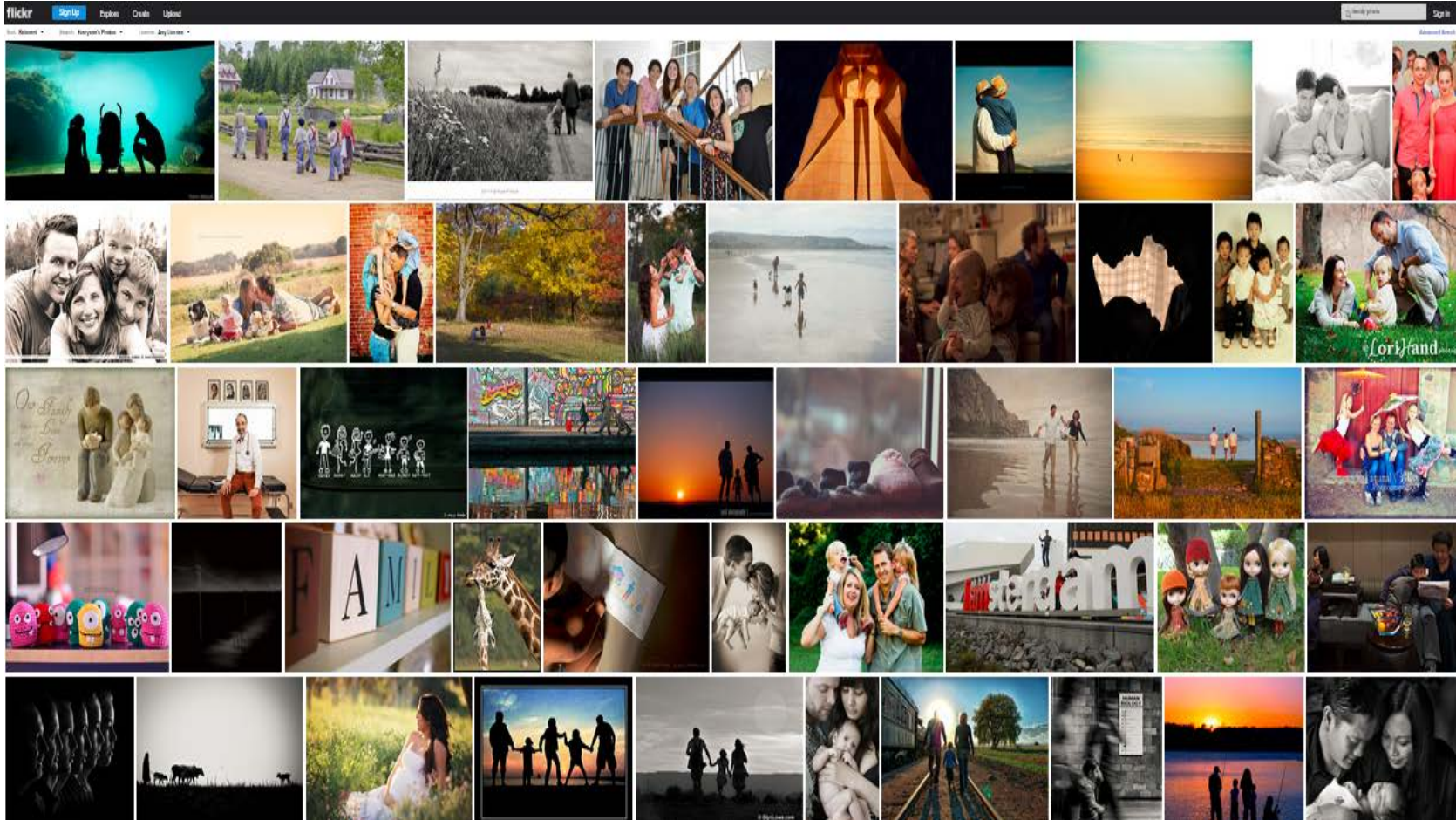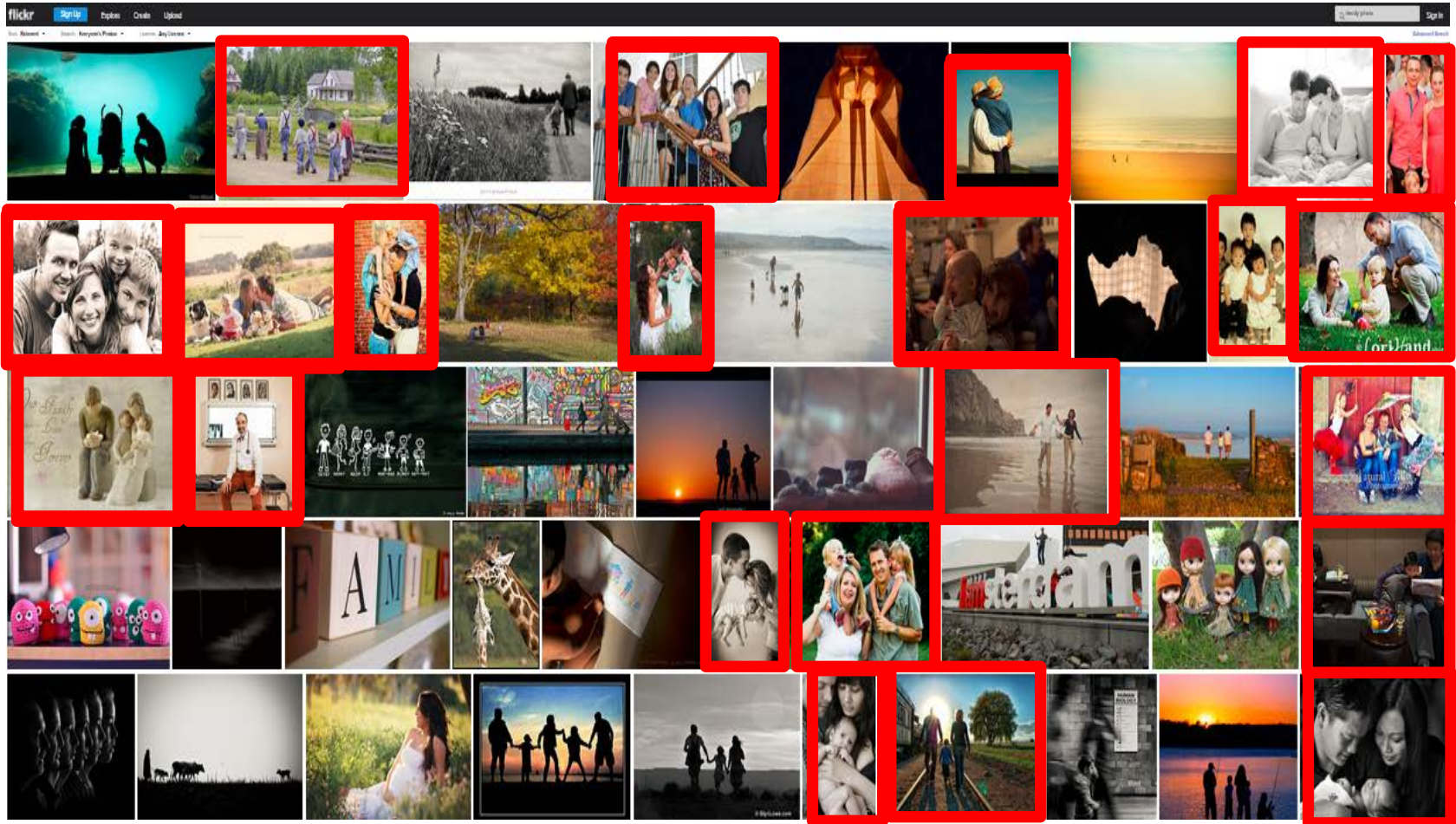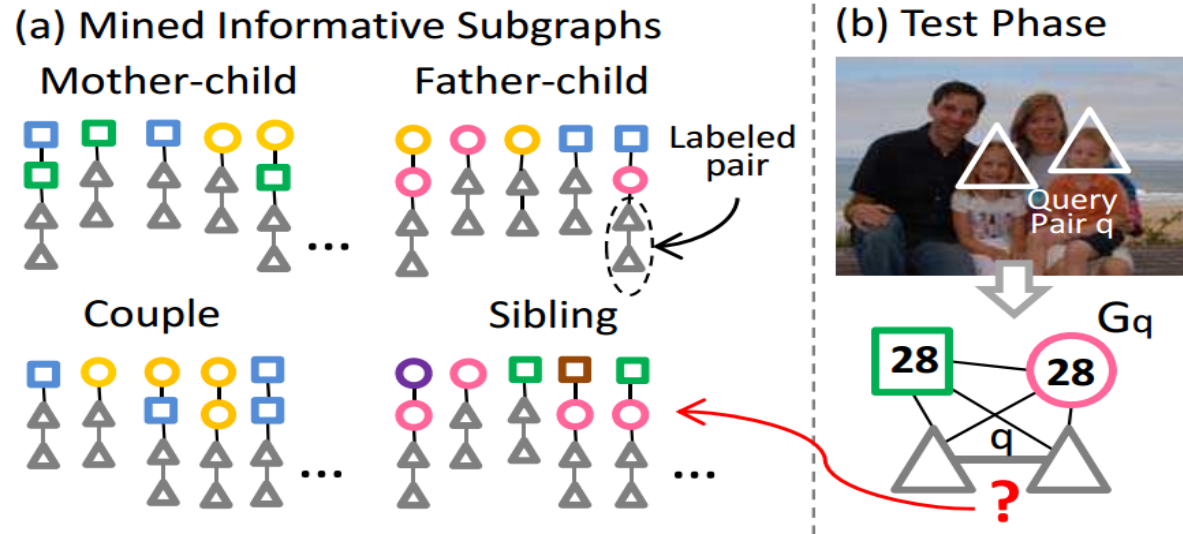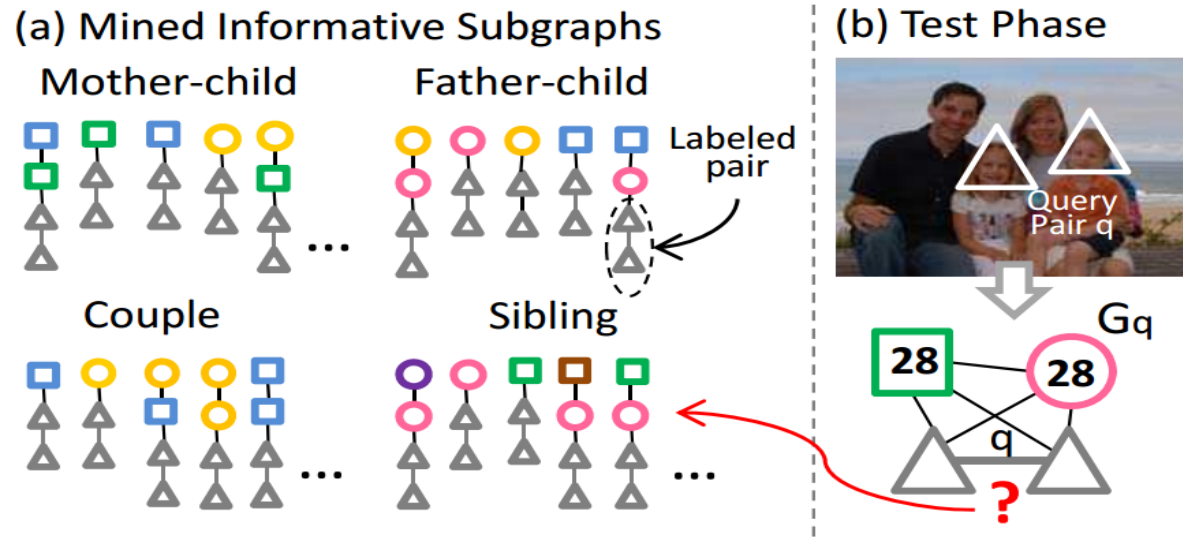- Chen et al. Discovering Informative Social Subgraphs and Predicting Pairwise Relationships from Group Photo, MM, 2012.

Cues

Geometry

Appearance

Cues

Geometry

+

Appearance

Geometry Model

1. Convex Hull is used as the basis to represent the geometry of people position.

$$\{\sum_{i=1}^{n} a_i s_i \mid (\forall i : a_i \geq 0) \wedge \sum_{i=1}^{n} a_i = 1\}$$

Geometry Model

1. Convex Hull is used as the basis to represent the geometry of people position.

$$\{\sum_{i=1}^{n} a_i s_i \mid (\forall i : a_i \geq 0) \wedge \sum_{i=1}^{n} a_i = 1\}$$
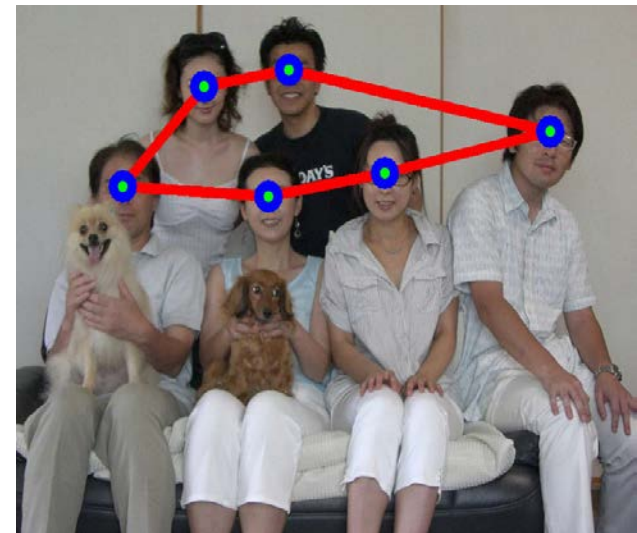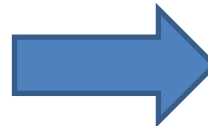
Geometry Model

2. Incorporating all the vertices

Geometry Model

2. Incorporating all the vertices
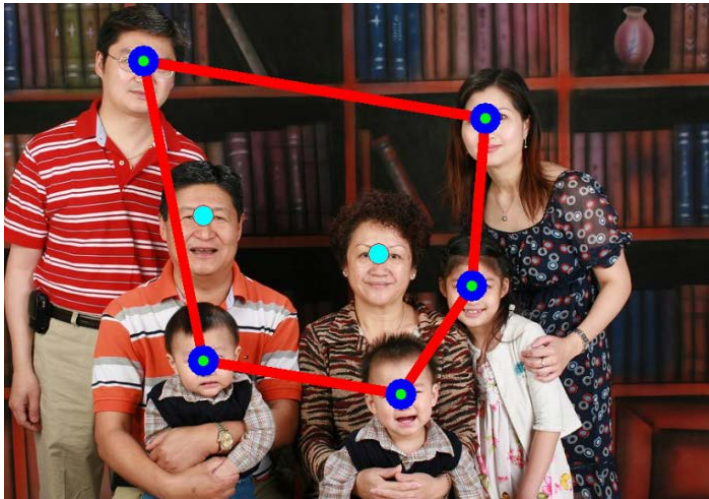
$$d = \sqrt{(u_i - x_{ai})^2 + (v_i - y_{ai})^2}.$$

Geometry Model

2. Incorporating all the vertices

$$d = \sqrt{(u_i - x_{ai})^2 + (v_i - y_{ai})^2}.$$

Geometry Model

Geometry Model

Geometry Model

Geometry Model

$\mathsf{L}(\theta_j)$

θ_j
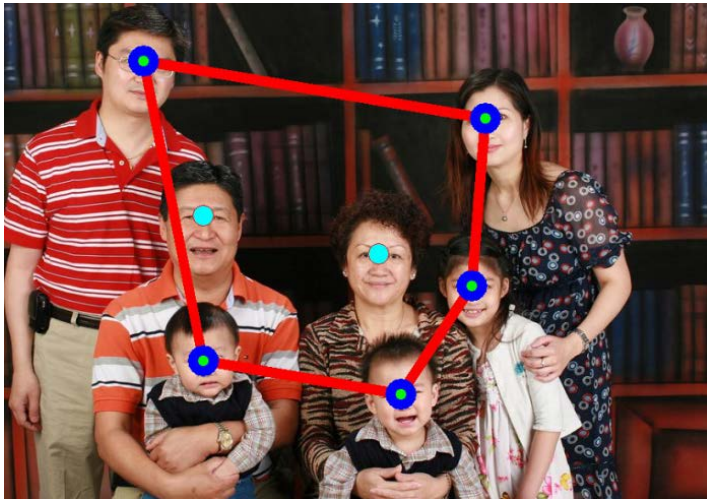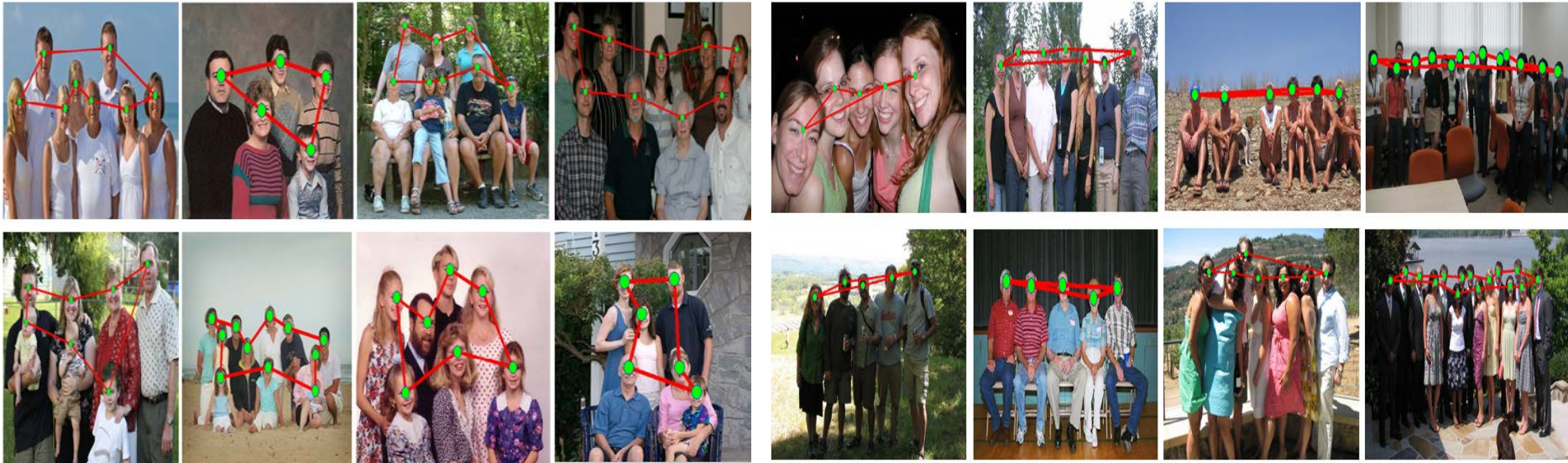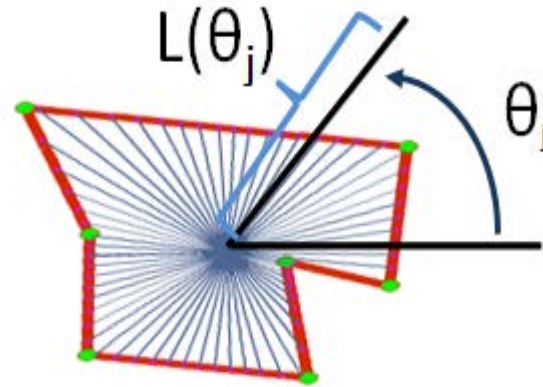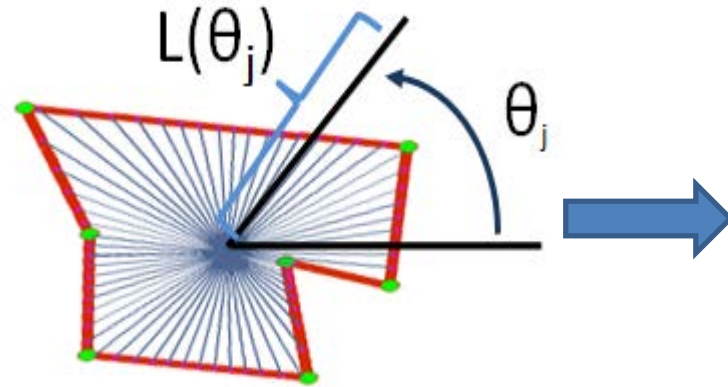
FFT of $\mathsf{L}(\theta_j)$

Geometry Model
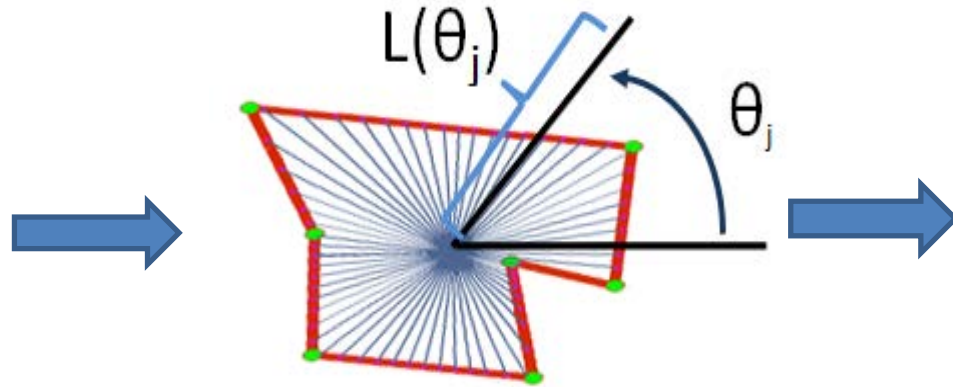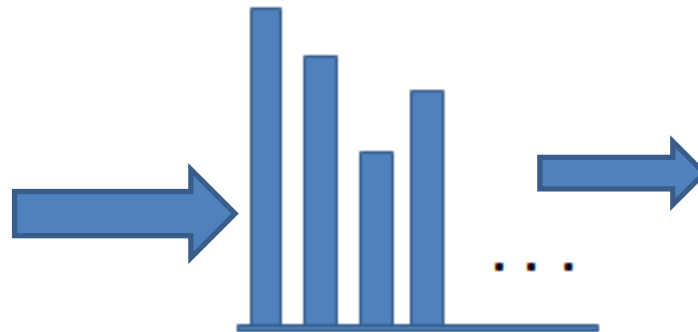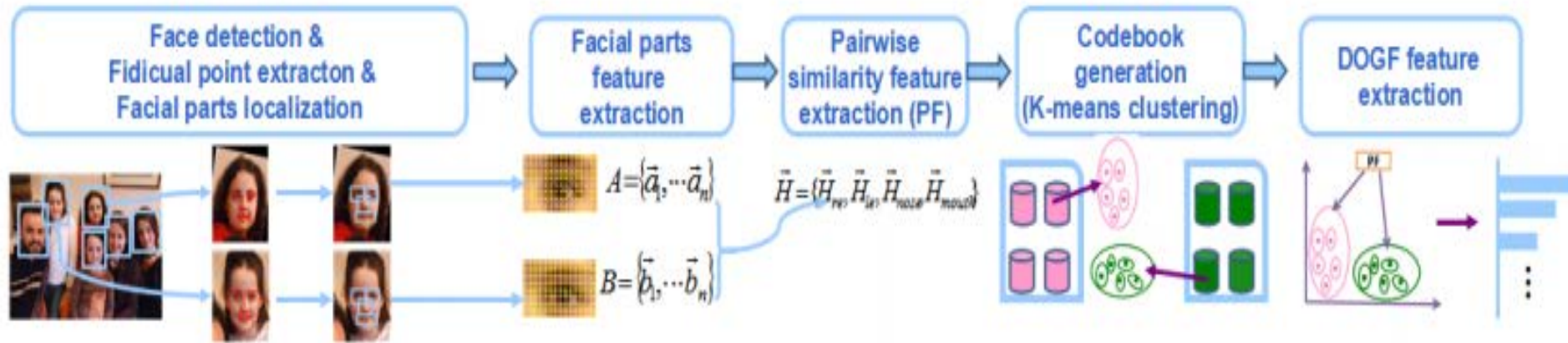
FFT of $L(\theta_j)$ $\rightarrow$ $\rightarrow$ **Classification**

Degree Of Group similarity Feature (DOGF) feature extraction

Degree Of Group similarity Feature (DOGF) feature extraction
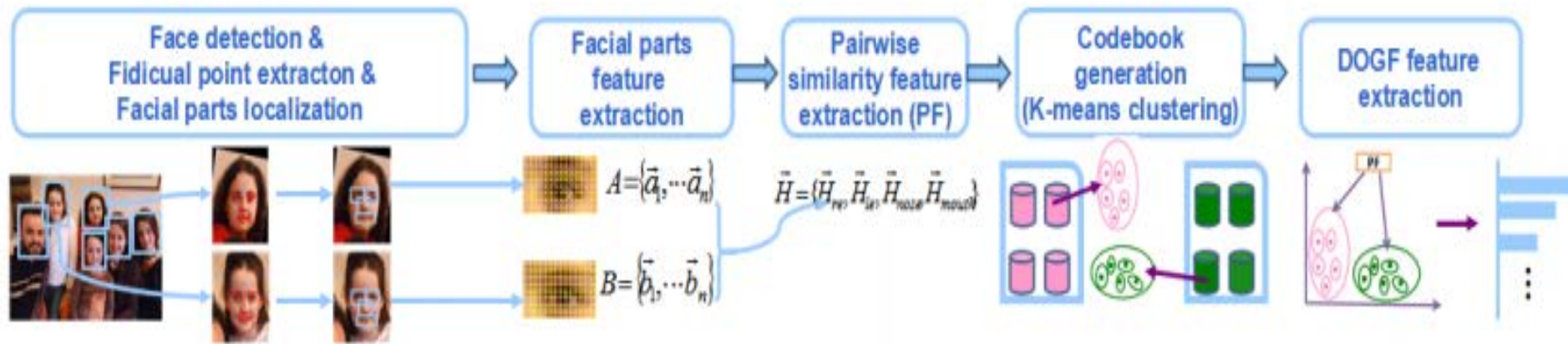
- Pairwise Feature (PF) Extraction

Degree Of Group similarity Feature (DOGF) feature extraction
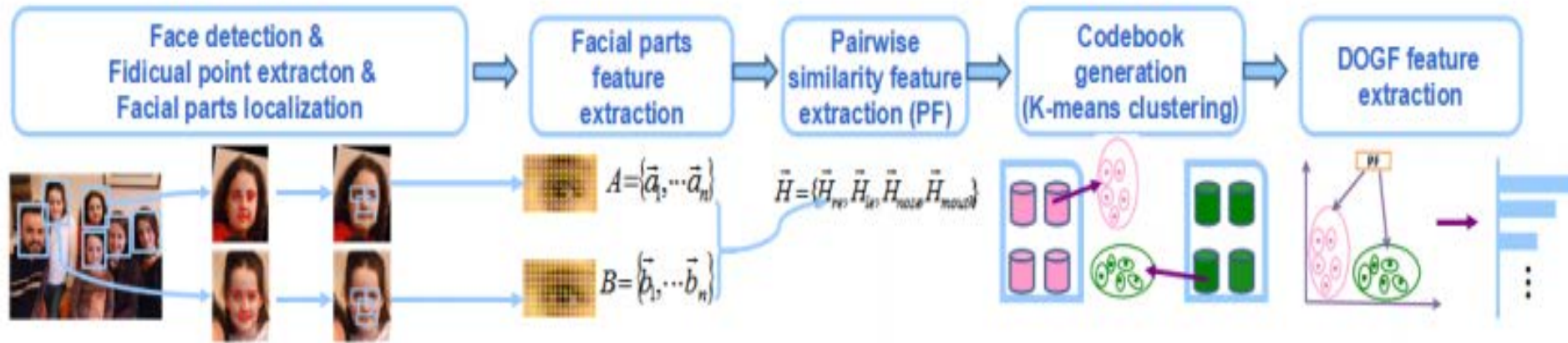
- Pairwise Feature (PF) Extraction
- Facial Codebook Construction

Degree Of Group similarity Feature (DOGF) feature extraction

- Pairwise Feature (PF) Extraction
- Facial Codebook Construction
- Formulation of DOGF Feature

Pairwise Feature (PF) Extraction

Dense SIFT

$$A = \{\vec{a}_1, \cdots \vec{a}_n\}$$

$$B = \{\vec{b}_1, \cdots \vec{b}_n\}$$

Modified Hausdorff Distance (MHD)

$$\vec{H}_{re}$$

$$PF = \{\vec{H}_{re}, \vec{H}_{le}, \vec{H}_{nose}, \vec{H}_{mouth}\}$$
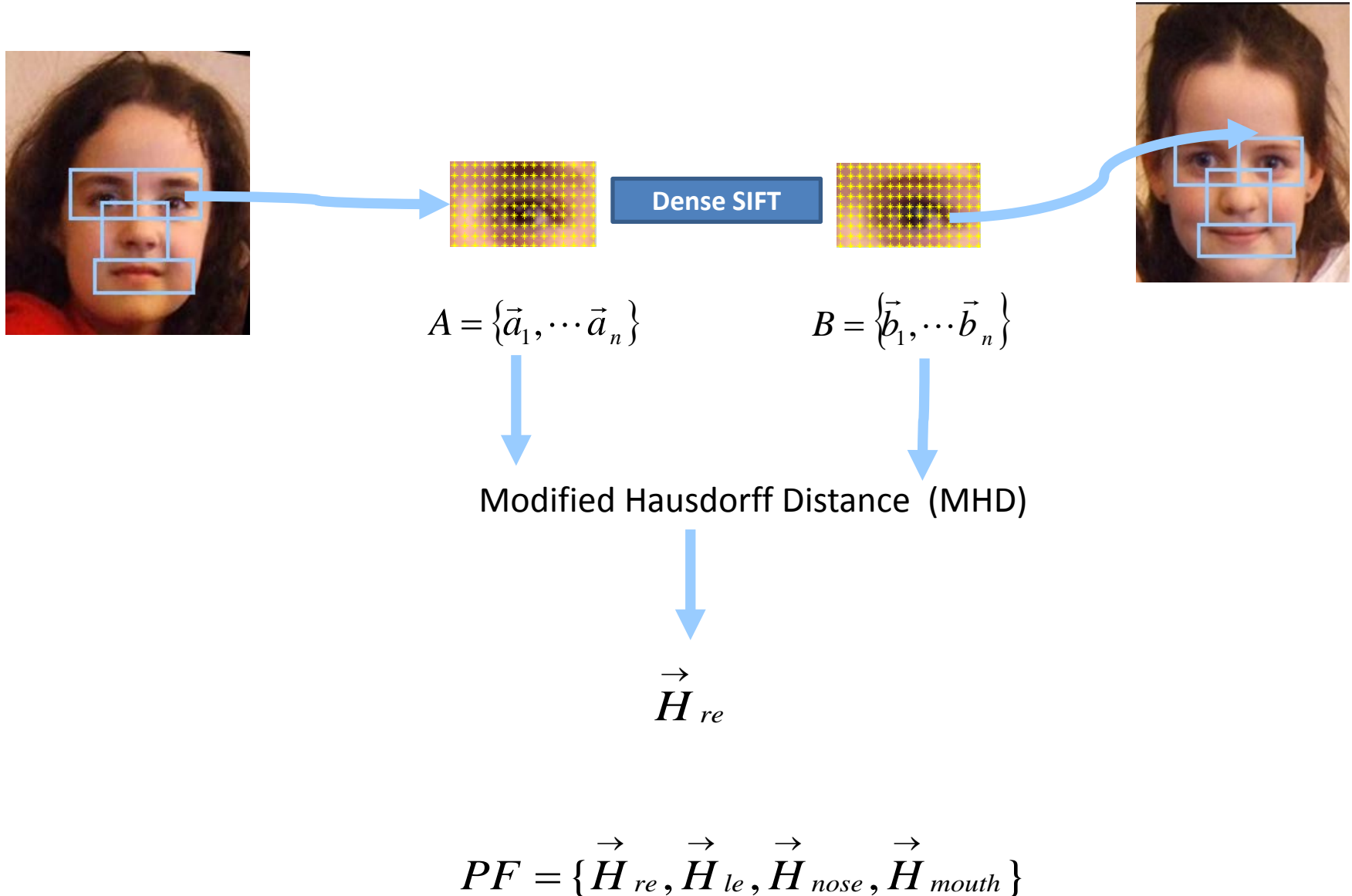
Facial Codebook Construction
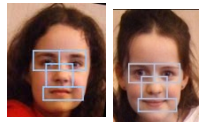
Age Gap (p) :   (a) P < 10;   (b)10<P<20; (c) 20 < P<40; (d) P > 40

PF
Age gap (p)

Age gap: (a)   (b)   (c)   (d)

$\vec{G}_{family}$
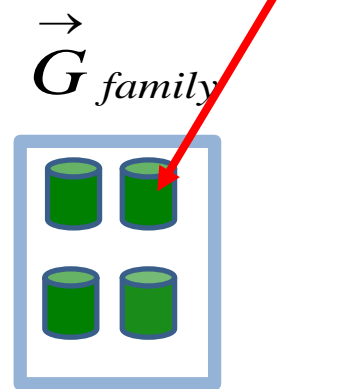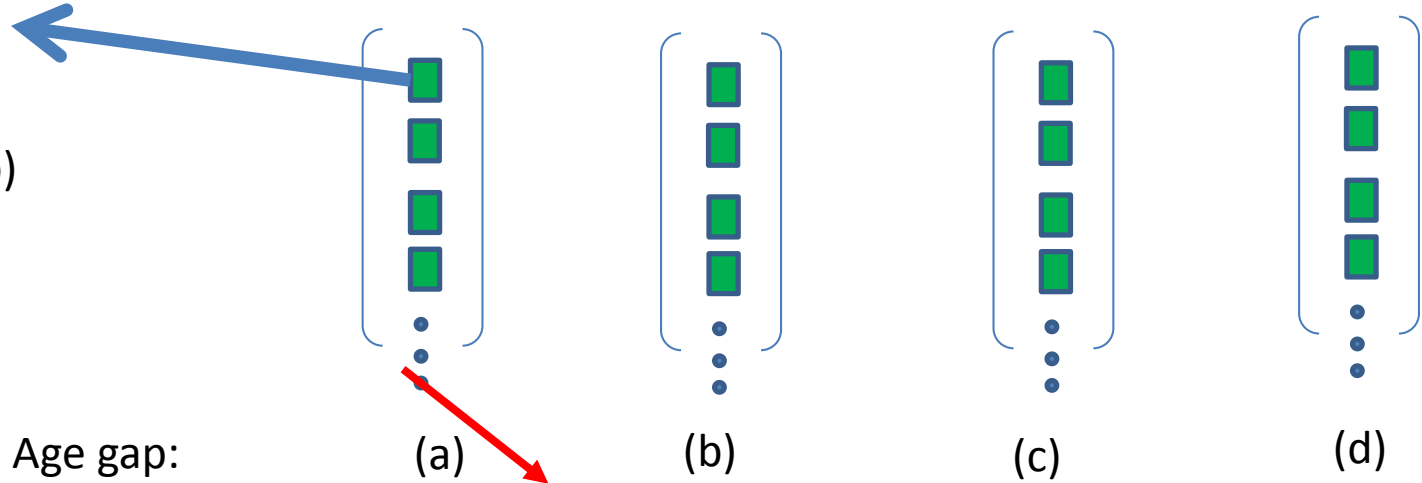
DOGF feature extraction
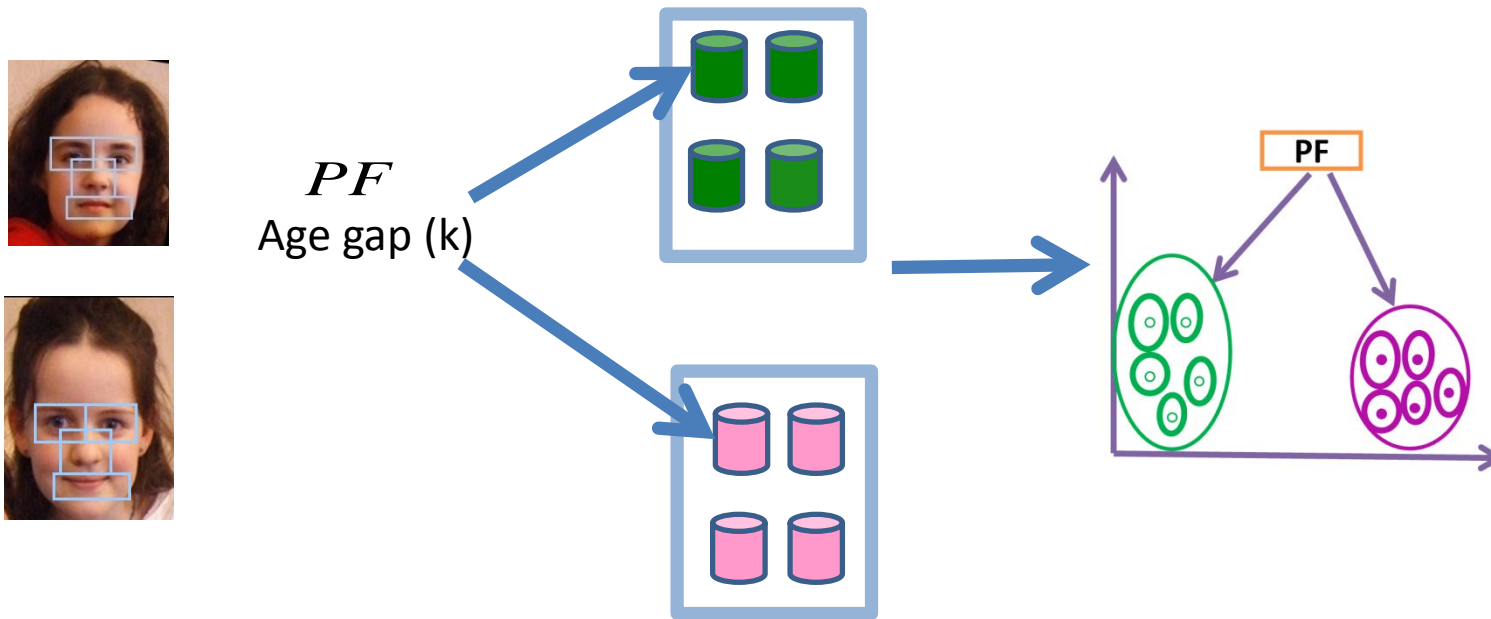
$$d_1 = d\,(\vec{H}_j, \vec{G}_{mk})$$

$$d_2 = d(\vec{H}_j, \vec{G'}_{mk})$$

$$\vec{d}_j = [d_1, d_2]$$

$$\vec{F} = \frac{1}{J}\sum_{j=1}^{J}\vec{d}_j$$

J : Number of facial pairs

Model Fusion

Score level fusion

$$R_c(x, y) = \sum_m w_m e^{-\frac{||x-y||^2}{2\sigma_m^2}}$$

W_m can be set a priori or via cross-validation.

W_m is set using cross-validation.

Experiments

Dataset:

	Family photos	Non-Family photos
Dataset One [Chen et al. MM 2012]	1167	1263
Dataset Two (New)	1420	1420

Note:

1. Dataset One is used in [Chen et al. MM 2012].
2. 16818 individuals are included in Dataset Two.
3. All photos used in our experiment contain at least three individuals.
4. Five-fold cross- validation.

Experiments

Correctly-classified samples using geometry model only

Family Photo

Non-Family Photo

Experiments

Misclassified samples using geometry model

Family Photo

Non-Family Photo

Experiments

Correctly-classified samples using appearance model

Family Photo

Non-Family Photo

Experiments

Misclassified samples using geometry model and correctly classified after fusion

Family

Experiments

| Method	Dataset One	Dataset Two
Chen et al [MM 12]	90.03%	-----
Geometry (Ours)	86.4%	87.3%
Appearance (Ours)	89.4%	89.0%
Fusion (Ours)	93.9%	93.4%

# Experiments

ROC illustration of Dataset One and Two
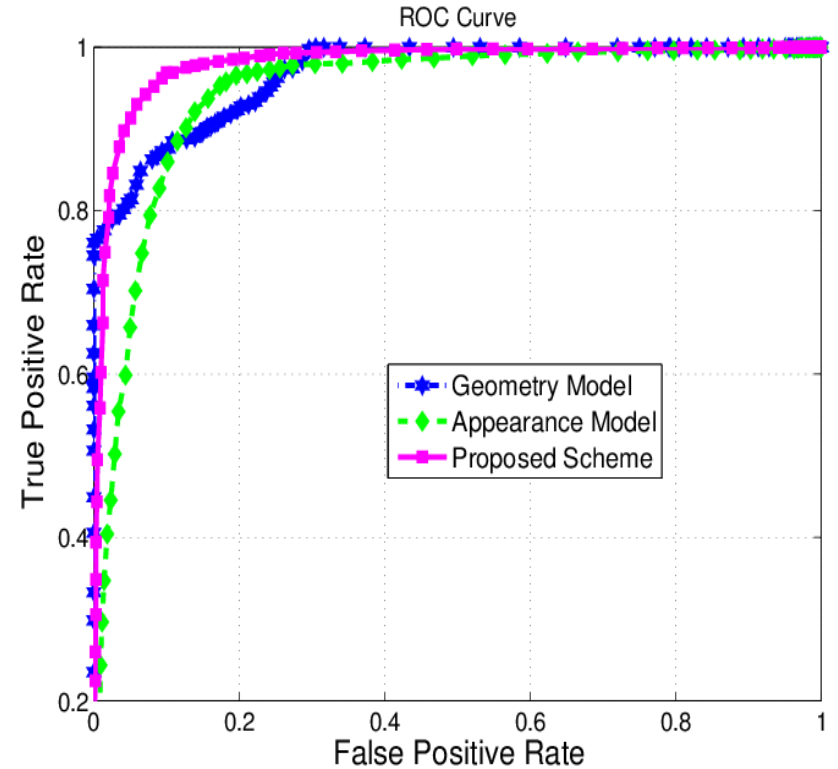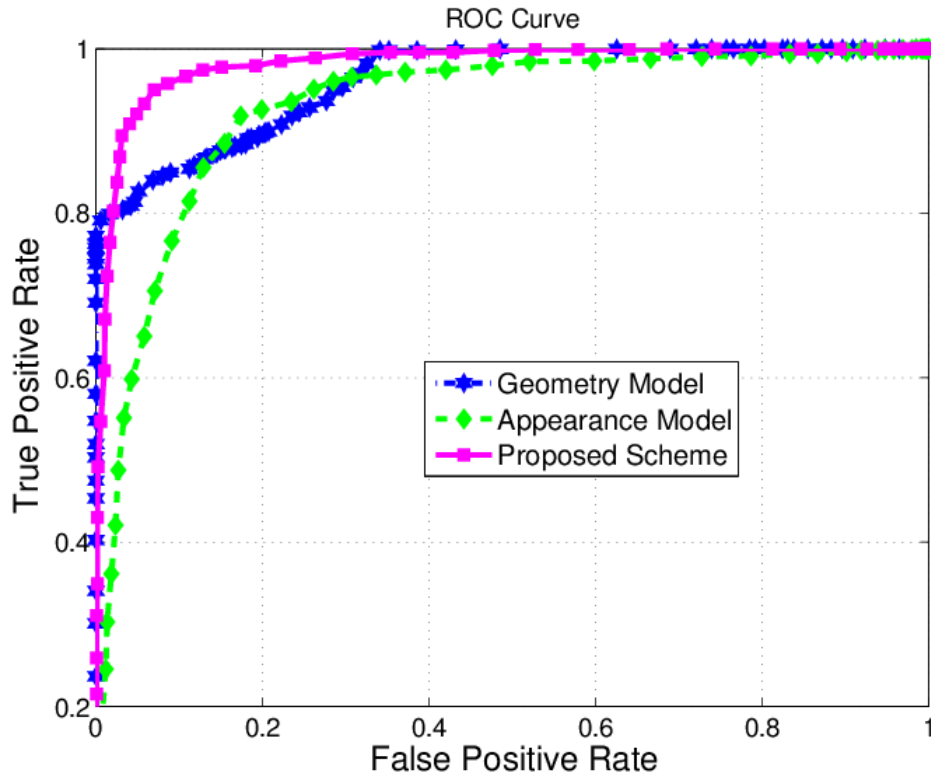
Experiments

Experimental Setting:

1. Geometry model baseline: the distances calculated from the center to all vertices are used as the feature.

2. Appearance model:

   a. Modified Hausdorff Distance(MHD) vs Hausdorff Distance(HD) .

   b. Using age information vs no-age information.

Experiments

Bar illustrations of different schemes
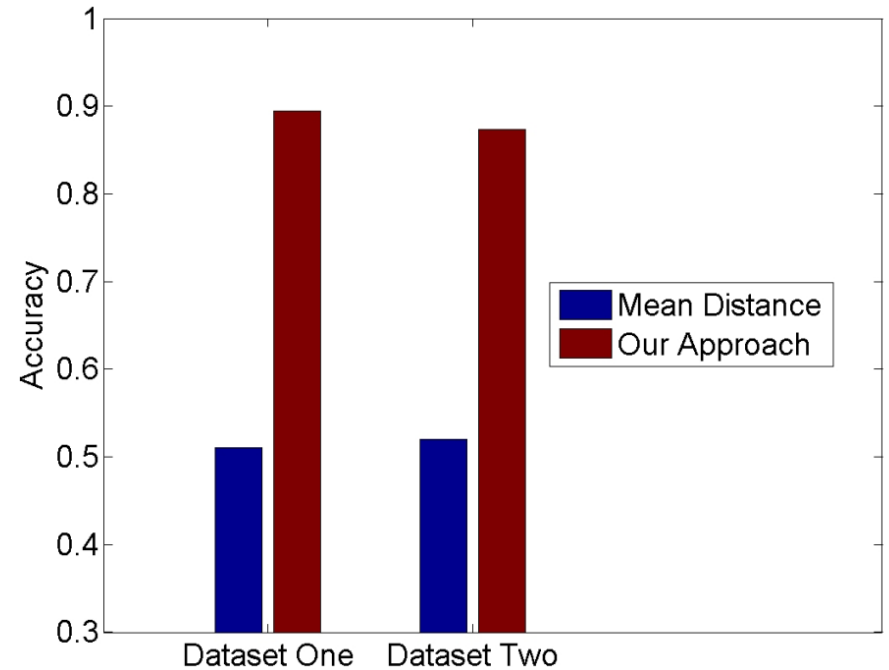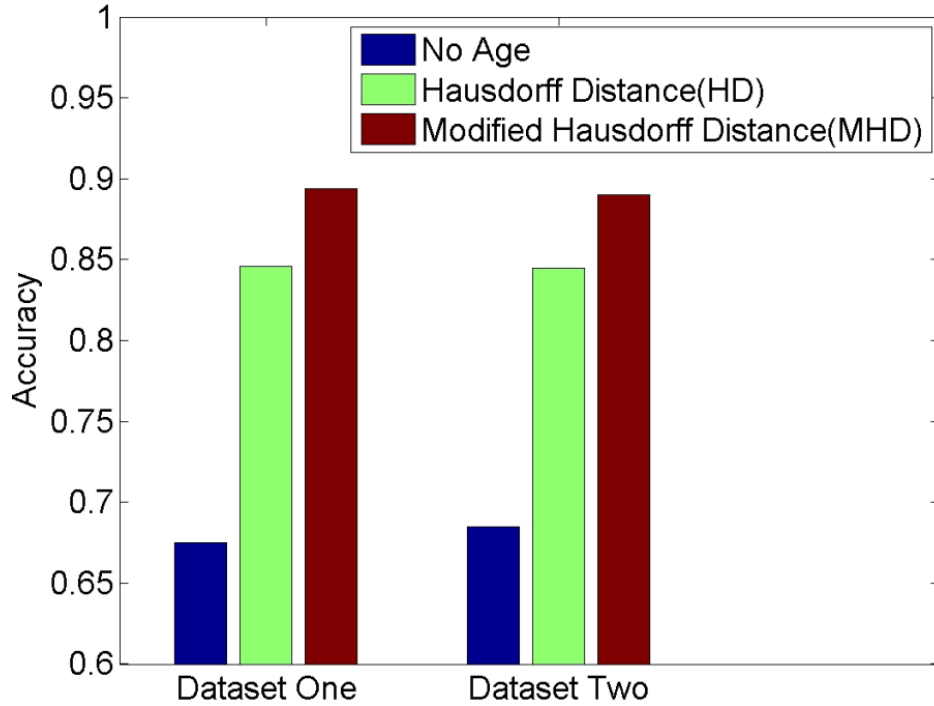
Conclusions

1. A new global geometry feature is proposed to represent the standing pattern of people in a group photo and we have obtained very promising results.

2. Our method demonstrated that fusing geometry with appearance feature is an efficient scheme for recognizing family photos.

3. In our future research, we like to explore automatically labeling photos for other groups (such as work colleagues, students, etc)

Thank you !