Image Ranking and Retrieval based on Multi-Attribute Queries

Behjat Siddiquie
behjat@cs.umd.edu

Rogerio S. Feris
rsferis@us.ibm.com

Larry S. Davis
lsd@cs.umd.edu

Attribute based Image Retrieval

"Young Asian woman wearing sunglasses"

Attribute based Image Retrieval

"Young Asian woman wearing sunglasses"

Attributes

Physical traits or characteristics of a person

Male
Asian
Middle-Aged

Female
White
Young

Attributes

Physical traits or characteristics of a person

Male
Asian
Middle-Aged

Female
White
Young

Object Properties

Round
White
Black

Attributes

Physical traits or characteristics of a person

Male
Asian
Middle-Aged

Female
White
Young

Object Properties

Round
White
Black

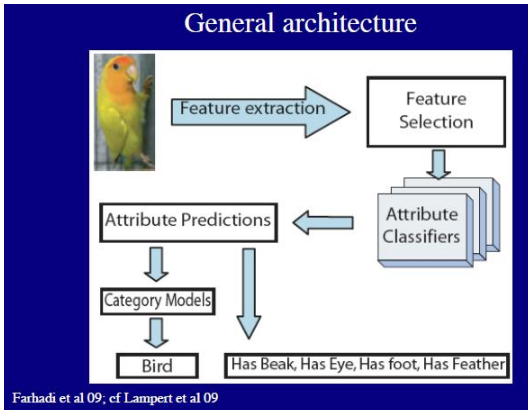
Object properties that span across object categories

Striped
Four-legged
Orange
Black

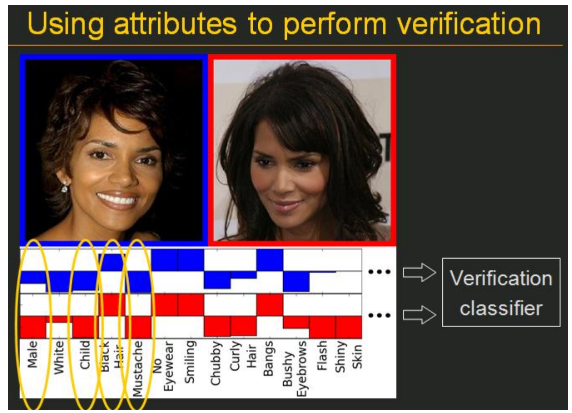
Striped
Four-legged
White
Black

Attribute based representation

Describing Images
Farhadi et al., CVPR 2009

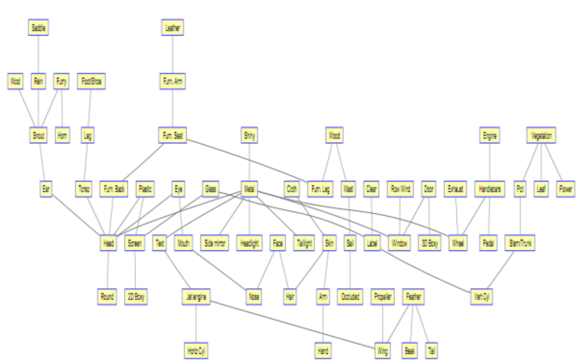


Face Verification
Kumar et al., ICCV 2009

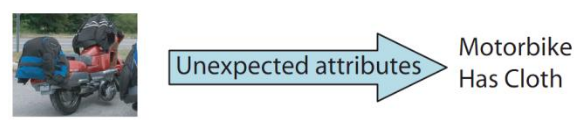


Transfer Learning
Lampert et al., CVPR 2009

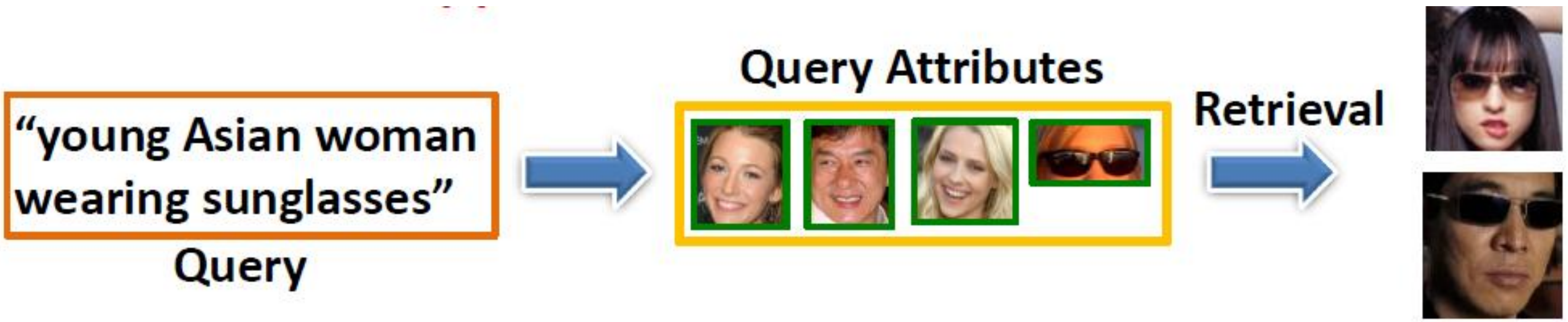
Object Recognition
Yang and Mori, ECCV 2010



Identifying Outliers
Farhadi et al., CVPR 2009



Multi-Attribute Image Retrieval

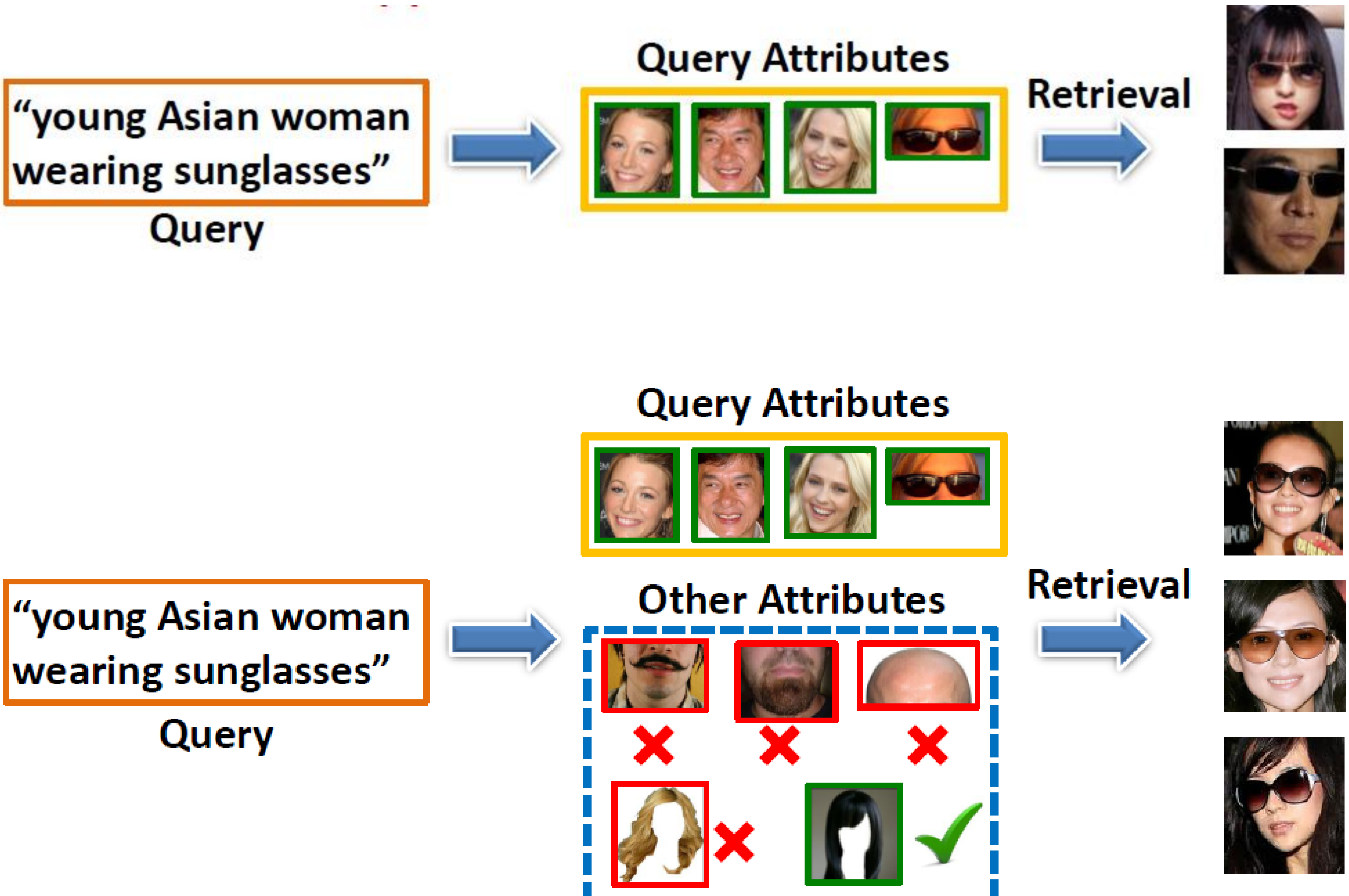


Number of possible queries is exponential

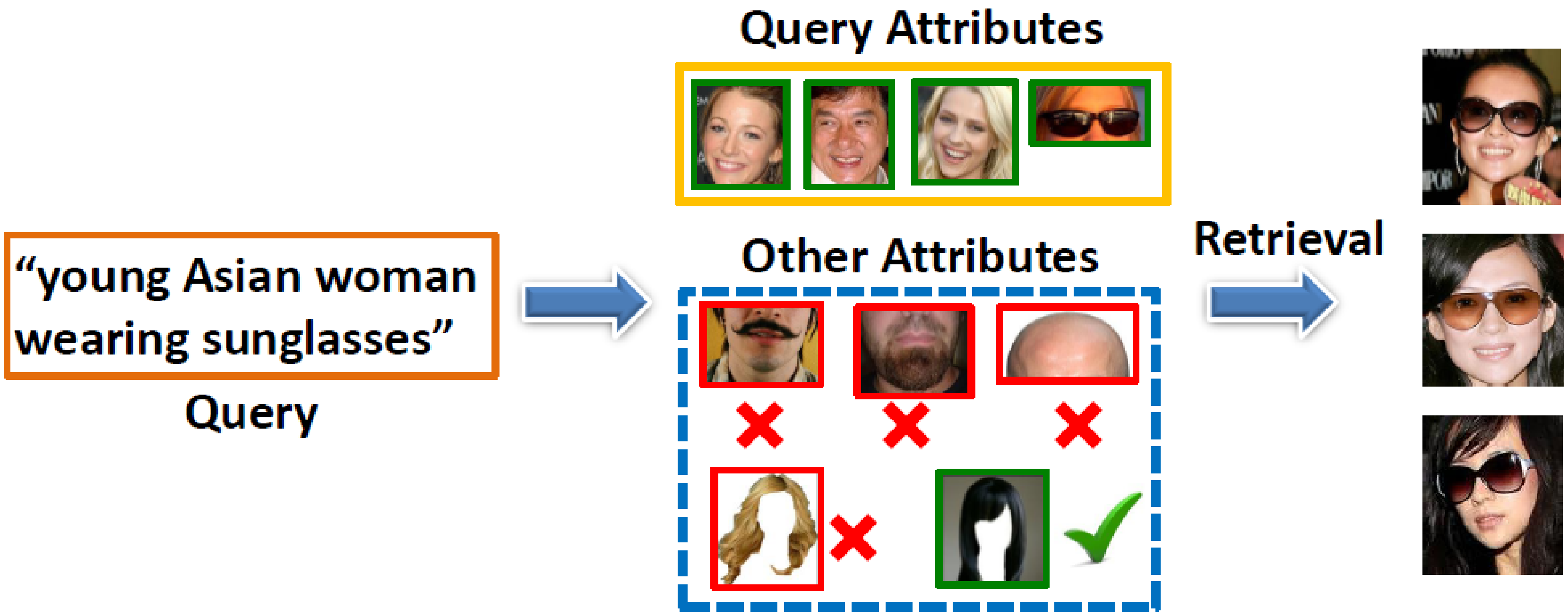
Existing Approaches

- Train independent classifiers for each attribute
- Sum up confidence scores

Multi-Attribute Image Retrieval



Multi-Attribute Image Retrieval



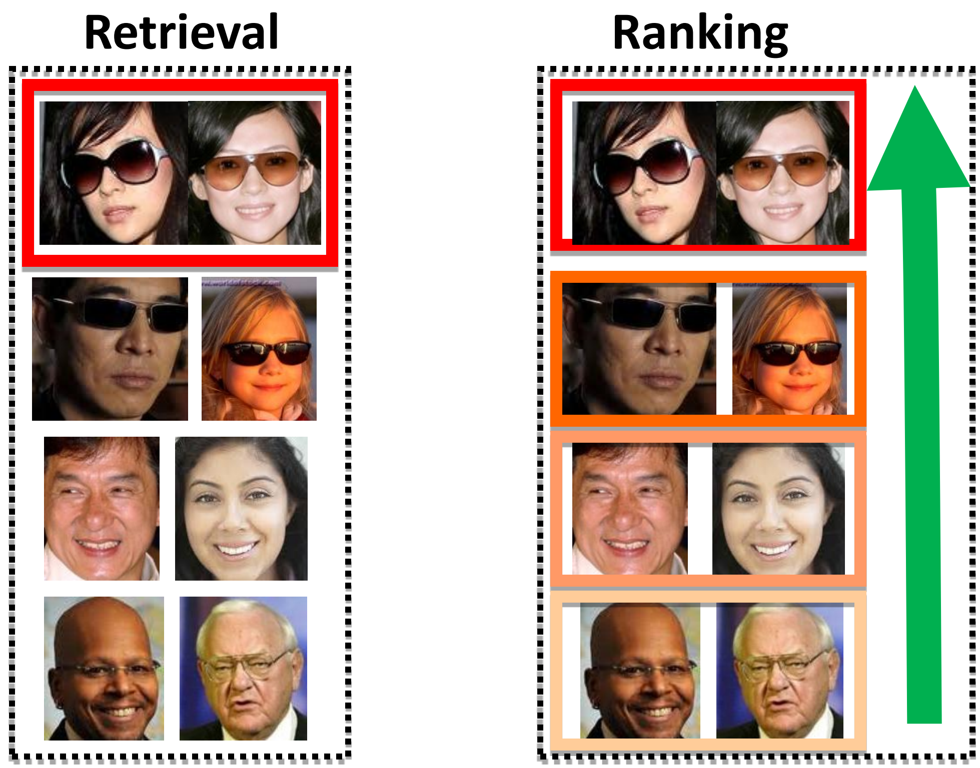
Model Correlations between Attributes

- Explicitly utilize information from non-query attributes

Multi-Attribute Image Retrieval

Joint Ranking and Retrieval Framework

- **Retrieval:** Set of images
- **Ranking:** Ordered set of images

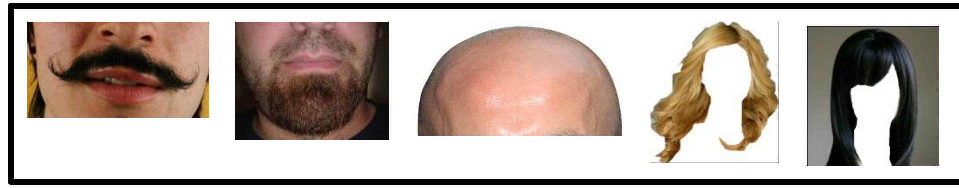


Multi-Attribute Image Ranking/Retrieval

We are given:

- An attribute vocabulary

$$\mathcal{X} = \{x_1, x_2, \ldots, x_K\}$$



- A set of training images

$$\mathcal{Y} = \{y_1, y_2, \ldots, y_N\}$$

- Multi-label annotation for each image

Hat
Sunglasses
Female
Asian
Young

Multi-Attribute Image Ranking/Retrieval

We are given:

- An attribute vocabulary

$$\mathcal{X} = \{x_1, x_2, \ldots, x_K\}$$

- A set of training images

$$\mathcal{Y} = \{y_1, y_2, \ldots, y_N\}$$

- Multi-label annotation for each image

Hat
Sunglasses
Female
Asian
Young

Our goal is:

- For a multi-label query Q , where $Q \subset \mathcal{X}$

- Rank/Retrieve relevant images from a dataset

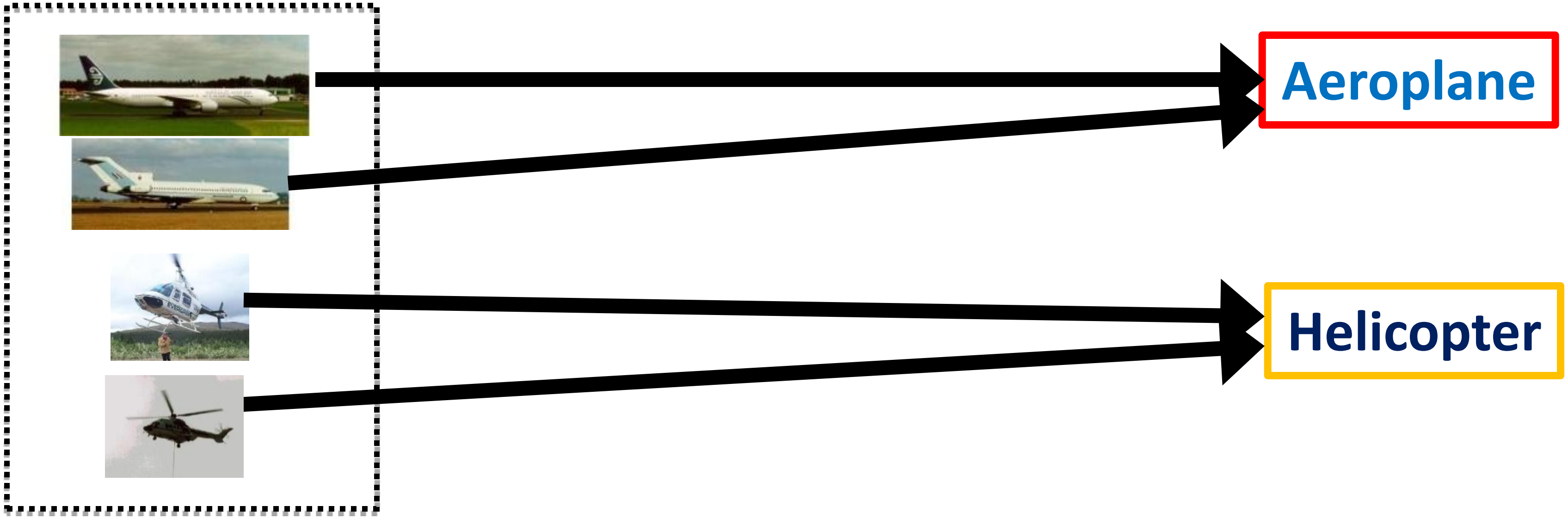
Male
Hat

Retrieval: Reverse Learning

- Reverse Multi-label Learning, Petterson and Caetano, NIPS 2010
- Given a label x_i such that $x_i \in \mathcal{X}$
- Predict the set of instances $y(\subset \mathcal{Y})$ that containing the label x_i

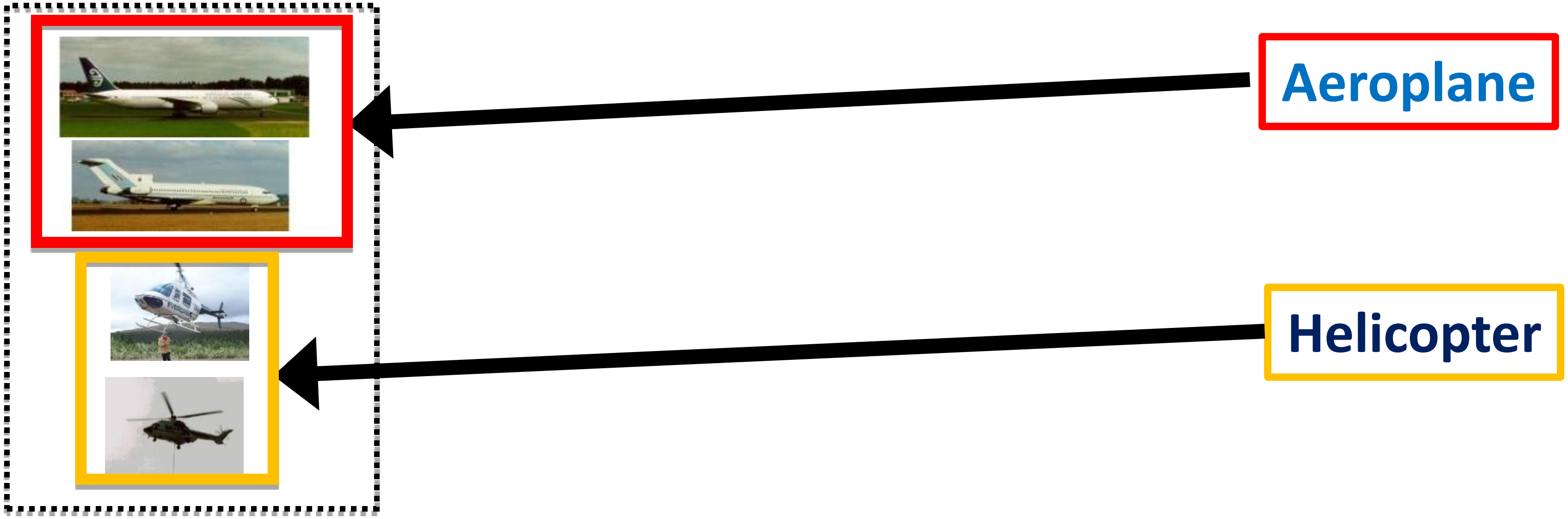
Retrieval: Conventional Learning

- **Conventional Learning**



Retrieval: Reverse Learning

- Given a label x_i such that $x_i \in \mathcal{X}$
- Predict the set of instances $y (\subset \mathcal{Y})$ that containing the label x_i



- Enables minimization of training loss based on a variety of metrics

Retrieval: Formulation

Given multi-attribute query $\mathcal{Q}$ **, output set of relevant images** y

Learn w **such that** $y^* = \arg\max_{y \subset \mathcal{Y}} w^T \psi(\mathcal{Q}, y)$

Retrieval: Formulation

Given multi-attribute query $\mathcal{Q}$ **, output set of relevant images** y

Learn w **such that** $y^* = \arg\max_{y \subset \mathcal{Y}} \boxed{w^T \psi(\mathcal{Q}, y)}$

where $\boxed{w^T \psi(\mathcal{Q}, y)} = \sum_{x_i \in \mathcal{Q}} w_i^a \Phi_a(x_i, y) + \sum_{x_i \in \mathcal{Q}} \sum_{x_j \in \mathcal{X}} w_{ij}^p \Phi_p(x_j, y)$

Retrieval: Formulation

Given multi-attribute query $\mathcal{Q}$, retrieve set of relevant images y

Learn w such that $$y^* = \arg\max_{y \subset \mathcal{Y}} w^T \psi(\mathcal{Q}, y)$$

where $$w^T \psi(\boxed{\mathcal{Q}}, y) = \sum_{x_i \in \mathcal{Q}} w_i^a \Phi_a(x_i, y) + \sum_{x_i \in \mathcal{Q}} \sum_{x_j \in \mathcal{X}} w_{ij}^p \Phi_p(x_j, y)$$

Asian woman wearing Sunglasses

Retrieval: Formulation

Given multi-attribute query $\mathcal{Q}$, retrieve set of relevant images y

Learn w such that $y^* = \arg\max_{y \subset \mathcal{Y}} w^T \psi(\mathcal{Q}, y)$

where $$w^T \psi(\mathcal{Q}, \boxed{y}) = \sum_{x_i \in \mathcal{Q}} w_i^a \Phi_a(x_i, y) + \sum_{x_i \in \mathcal{Q}} \sum_{x_j \in \mathcal{X}} w_{ij}^p \Phi_p(x_j, y)$$

Retrieval: Formulation

Given multi-attribute query $\mathcal{Q}$ **, retrieve set of relevant images** y

Learn w **such that** $y^* = \arg\max_{y \subset \mathcal{Y}} w^T \psi(\mathcal{Q}, y)$

where $\boxed{w^T \psi(\mathcal{Q}, y)} = \sum_{x_i \in \mathcal{Q}} w_i^a \Phi_a(x_i, y) + \sum_{x_i \in \mathcal{Q}} \sum_{x_j \in \mathcal{X}} w_{ij}^p \Phi_p(x_j, y)$

Retrieval: Formulation

Given multi-attribute query $\mathcal{Q}$, retrieve set of relevant images y

Learn w such that $y^* = \arg\max_{y \subset \mathcal{Y}} w^T \psi(\mathcal{Q}, y)$

where $$w^T \psi(\mathcal{Q}, y) = \sum_{x_i \in \mathcal{Q}} w_i^a \Phi_a(x_i, y) + \sum_{x_i \in \mathcal{Q}} \sum_{x_j \in \mathcal{X}} w_{ij}^p \Phi_p(x_j, y)$$

$$\Phi_a(x_i, y) = \sum_{y_k \in y} \boxed{\phi_a(x_i, y_k)}$$




Asian woman wearing Sunglasses

Retrieval: Formulation

Given multi-attribute query $\mathcal{Q}$, retrieve set of relevant images y

Learn w such that $y^* = \arg\max_{y \subset \mathcal{Y}} w^T \psi(\mathcal{Q}, y)$

where $w^T \psi(\mathcal{Q}, y) = \sum_{x_i \in \mathcal{Q}} w_i^a \boxed{\Phi_a(x_i, y)} + \sum_{x_i \in \mathcal{Q}} \sum_{x_j \in \mathcal{X}} w_{ij}^p \Phi_p(x_j, y)$

$$\boxed{\Phi_a(x_i, y)} = \sum_{y_k \in y} \phi_a(x_i, y_k)$$

Asian woman wearing Sunglasses

Retrieval: Formulation

Given multi-attribute query $\mathcal{Q}$, retrieve set of relevant images y

Learn w such that $y^* = \arg\max_{y \subset \mathcal{Y}} w^T \psi(\mathcal{Q}, y)$

where $$w^T \psi(\mathcal{Q}, y) = \boxed{\sum_{x_i \in \mathcal{Q}} w_i^a \Phi_a(x_i, y)} + \sum_{x_i \in \mathcal{Q}} \sum_{x_j \in \mathcal{X}} w_{ij}^p \Phi_p(x_j, y)$$

Asian + woman + wearing Sunglasses

Retrieval: Formulation

Given multi-attribute query $\mathcal{Q}$ **, retrieve set of relevant images** y

Learn w **such that** $y^* = \arg\max_{y \subset \mathcal{Y}} w^T \psi(\mathcal{Q}, y)$

where $w^T \psi(\mathcal{Q}, y) = \sum_{x_i \in \mathcal{Q}} w_i^a \Phi_a(x_i, y) + \sum_{x_i \in \mathcal{Q}} \boxed{\sum_{x_j \in \mathcal{X}} w_{ij}^p \Phi_p(x_j, y)}$

Asian . Blonde hair

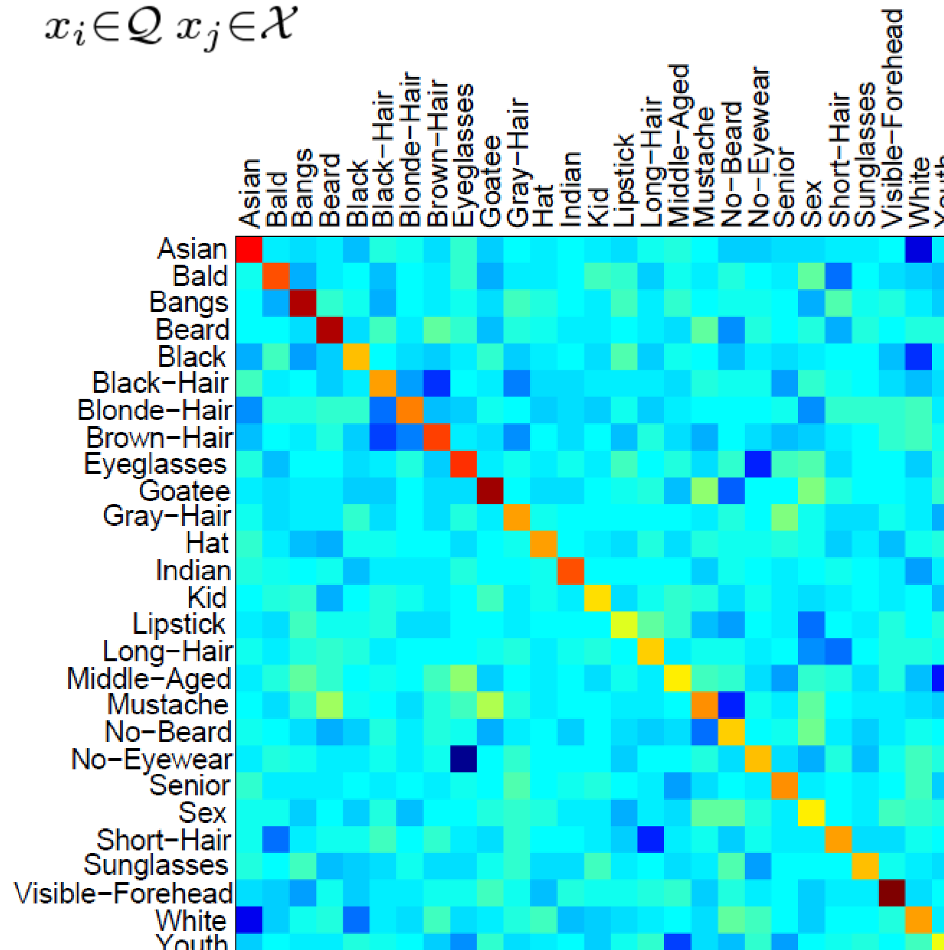
Asian . Gray hair

Asian . Black hair

Retrieval: Formulation

Weights Learnt

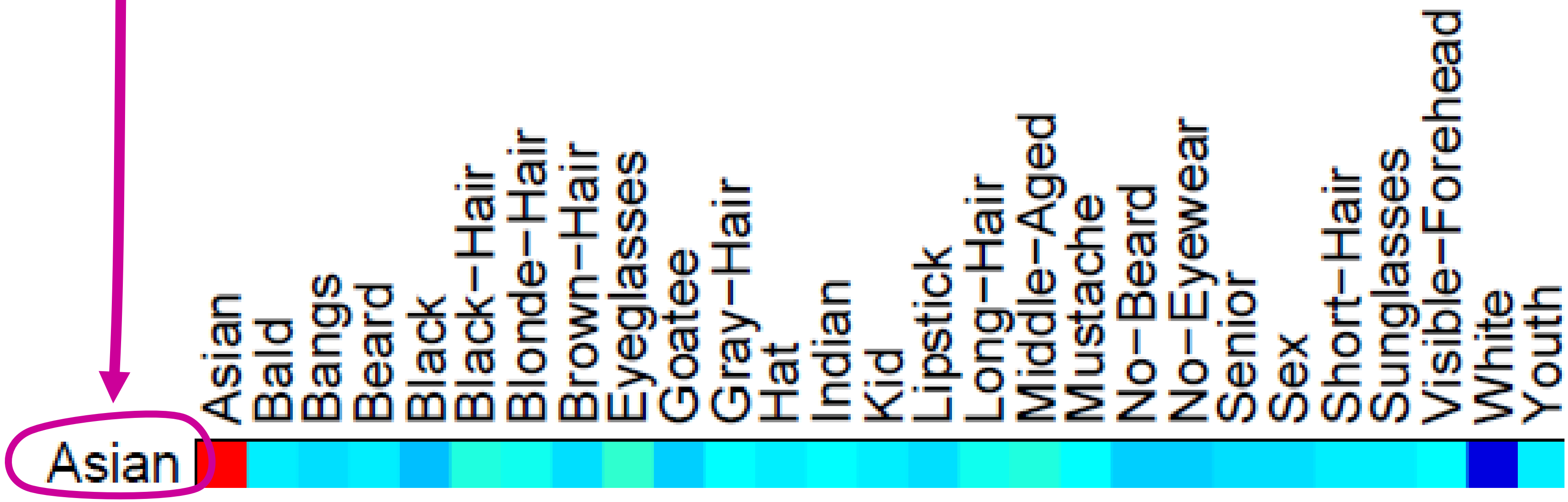
$$w^T \psi(\mathcal{Q}, y) = \sum_{x_i \in \mathcal{Q}} w_i^a \Phi_a(x_i, y) + \sum_{x_i \in \mathcal{Q}} \sum_{x_j \in \mathcal{X}} w_{ij}^p \Phi_p(x_j, y)$$



Retrieval: Formulation

Weights Learnt

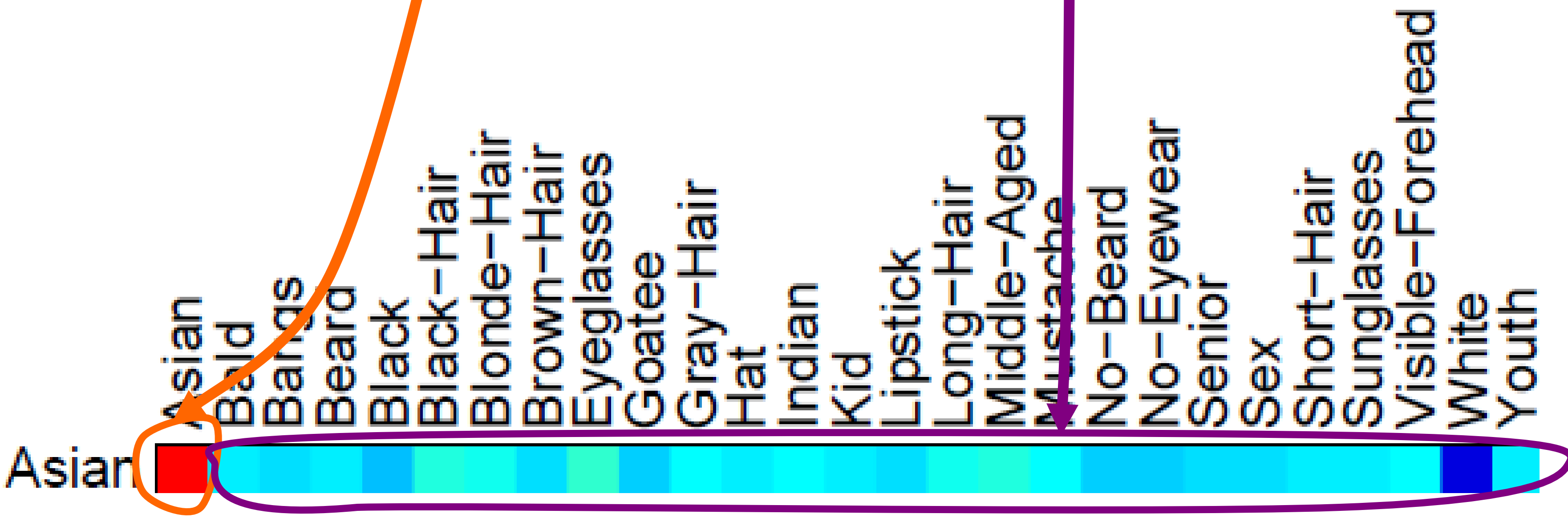
$$w^T\psi(\mathcal{Q}, y) = \sum_{x_i \in \mathcal{Q}} w_i^a \Phi_a(x_i, y) + \sum_{x_i \in \mathcal{Q}} \sum_{x_j \in \mathcal{X}} w_{ij}^p \Phi_p(x_j, y)$$



Retrieval: Formulation

Weights Learnt

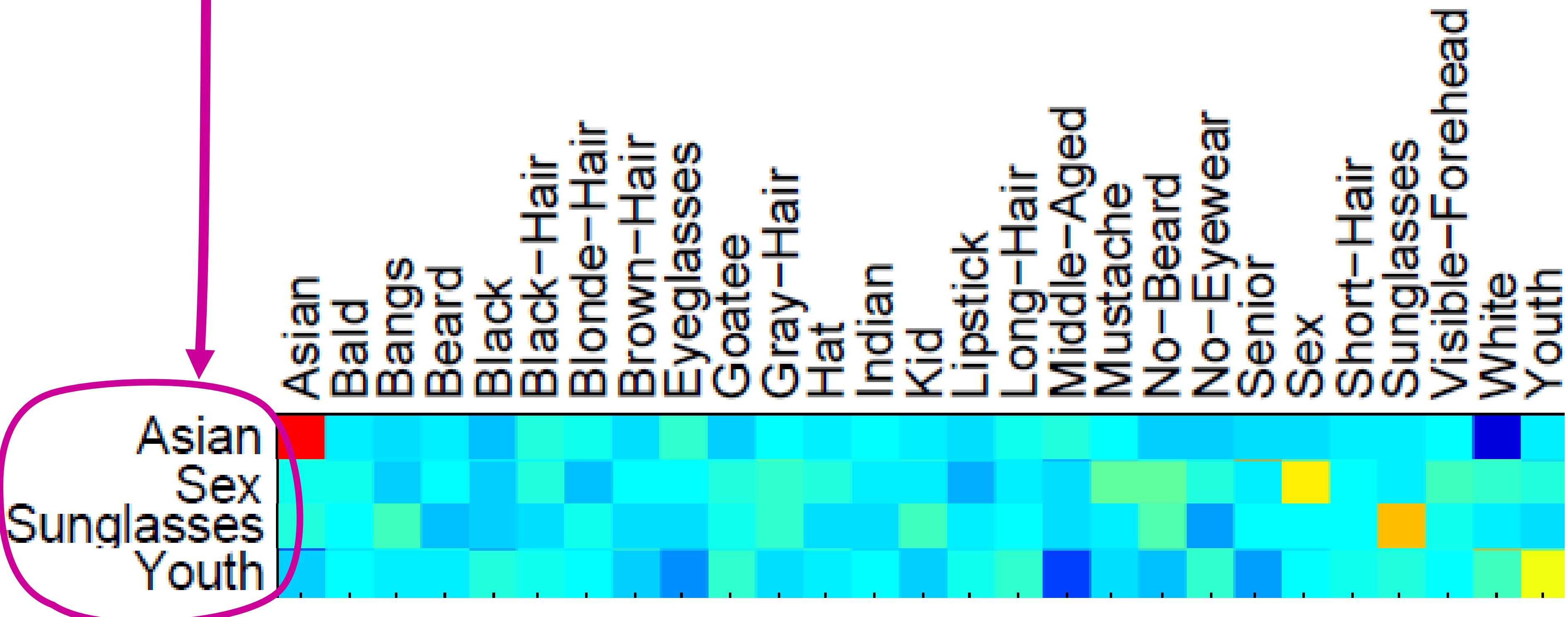
$$w^T \psi(\mathcal{Q}, y) = \sum_{x_i \in \mathcal{Q}} w_i^a \Phi_a(x_i, y) + \sum_{x_i \in \mathcal{Q}} \sum_{x_j \in \mathcal{X}} w_{ij}^p \Phi_p(x_j, y)$$



Retrieval: Formulation

Weights Learnt

$$w^T \psi(\mathcal{Q}, y) = \sum_{x_i \in \mathcal{Q}} w_i^a \Phi_a(x_i, y) + \sum_{x_i \in \mathcal{Q}} \sum_{x_j \in \mathcal{X}} w_{ij}^p \Phi_p(x_j, y)$$



Retrieval: Training

Training

$$\arg\min_{w,\xi} \quad w^T w + C \sum_t \xi_t$$

$$\forall \ t \quad w^T \psi(\mathcal{Q}_t, y_t^*) - w^T \psi(\mathcal{Q}_t, y_t) \geq \boxed{\Delta(y_t^*, y_t)} - \xi_t$$

Retrieval: Training

Training

$$\arg\min_{w,\xi} \quad w^T w + C \sum_t \xi_t$$

$$\forall \; t \quad w^T \psi(\mathcal{Q}_t, y_t^*) - w^T \psi(\mathcal{Q}_t, y_t) \geq \boxed{\Delta(y_t^*, y_t)} - \xi_t$$

loss function

$$\boxed{\Delta(y^*, y)} = \begin{cases} 1 - \frac{y \cap y^*}{y} & \text{precision} \\ 1 - \frac{y \cap y^*}{y^*} & \text{recall} \\ 1 - \frac{y \cap y^* + \bar{y} \cap \bar{y}^*}{\mathcal{Y}} & \text{hamming loss} \end{cases}$$

Ranking

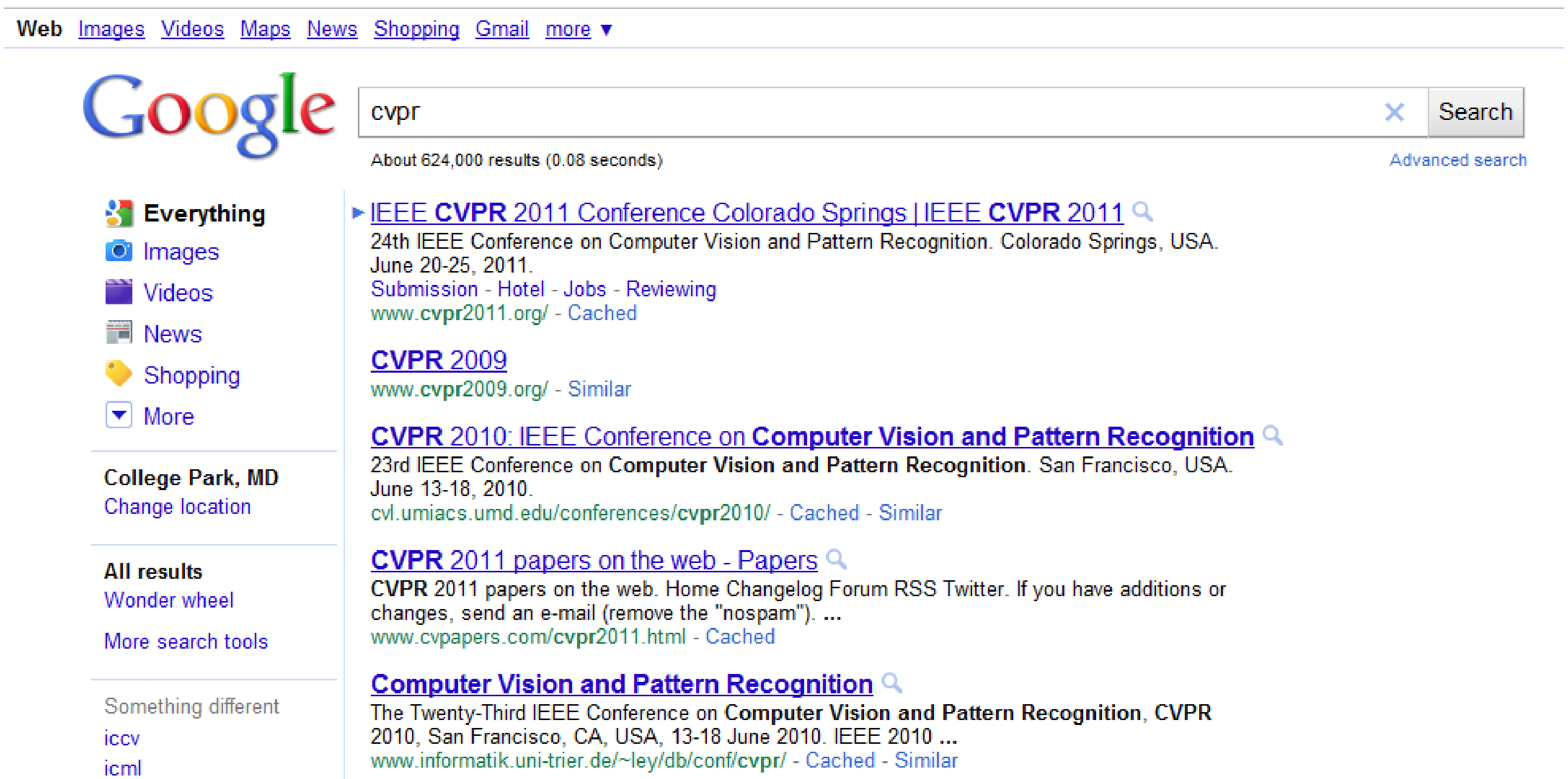
Given a query Q, rank documents in order of relevance

- **Output is an ordered set (permutation)**

Ranking

Given a query Q, rank documents in order of relevance

- **Output is an ordered set (permutation)**



Ranking

Given a query Q, rank documents in order of relevance

- Output is an ordered set (permutation)

Large Scale Datasets

- Ranking more important than retrieval

Ranking: Formulation

Given a multi-attribute query $\mathcal{Q}$, generate permutation of images z

Learn w such that $z^* = \arg\max_{z \in \pi(\mathcal{Y})} w^T \psi(\mathcal{Q}, z)$

where $w^T \psi(\mathcal{Q}, z) = \sum_{x_i \in \mathcal{Q}} w_i^a \hat{\Phi}_a(x_i, z) + \sum_{x_i \in \mathcal{Q}} \sum_{x_j \in \mathcal{X}} w_{ij}^p \hat{\Phi}_p(x_j, z)$

and

$$\hat{\Phi}_a(x_i, z) = \sum_{z_k \in z} A(r(z_k)) \phi_a(x_i, z_k)$$

$$\hat{\Phi}_p(x_j, z) = \sum_{z_k \in z} A(r(z_k)) \phi_p(x_j, z_k)$$

Ranking: Formulation

Given a multi-attribute query $\mathcal{Q}$, generate permutation of images z

Learn w such that $z^* = \arg\max_{z \in \pi(\mathcal{Y})} w^T \psi(\mathcal{Q}, z)$

where $$w^T \psi(\mathcal{Q}, z) = \sum_{x_i \in \mathcal{Q}} w_i^a \hat{\Phi}_a(x_i, z) + \sum_{x_i \in \mathcal{Q}} \sum_{x_j \in \mathcal{X}} w_{ij}^p \hat{\Phi}_p(x_j, z)$$

and $$\hat{\Phi}_a(x_i, z) = \sum_{z_k \in z} \boxed{A(r(z_k))} \phi_a(x_i, z_k)$$

$$\hat{\Phi}_p(x_j, z) = \sum_{z_k \in z} \boxed{A(r(z_k))} \phi_p(x_j, z_k)$$

Ranking: Training

Training $$\arg\min_{w,\xi} \quad w^T w + C \sum_t \xi_t$$

$$\forall \ t \quad w^T \psi(\mathcal{Q}_t, z_t^*) - w^T \psi(\mathcal{Q}_t, z_t) \geq \Delta(z_t^*, z_t) - \xi_t$$

Ranking: Training

Training $\arg\min_{w,\xi} \quad w^T w + C \sum_t \xi_t$

$$\forall \ t \quad w^T \psi(\mathcal{Q}_t, z_t^*) - w^T \psi(\mathcal{Q}_t, z_t) \geq \boxed{\Delta(z_t^*, z_t)} - \xi_t$$

loss function

$$\boxed{\Delta(z^*, z)} = 1 - \mathrm{NDCG}_k(\mathrm{z}^*, z)$$

Ranking: Training

Training $\arg\min_{w,\xi} \quad w^T w + C \sum_t \xi_t$

$$\forall \ t \quad w^T \psi(\mathcal{Q}_t, z_t^*) - w^T \psi(\mathcal{Q}_t, z_t) \geq \Delta(z_t^*, z_t) - \xi_t$$

loss function

$$\Delta(z^*, z) = 1 - \mathrm{NDCG}_k(\mathrm{z}^*, z)$$

where

$$\boxed{NDCG_k = \frac{1}{Z} \sum_{j=1}^{k} \frac{2^{\mathrm{rel}(j)} - 1}{\log(1 + j)}}$$

Ranking: Training

Training $\arg\min_{w,\xi} \quad w^T w + C \sum_t \xi_t$

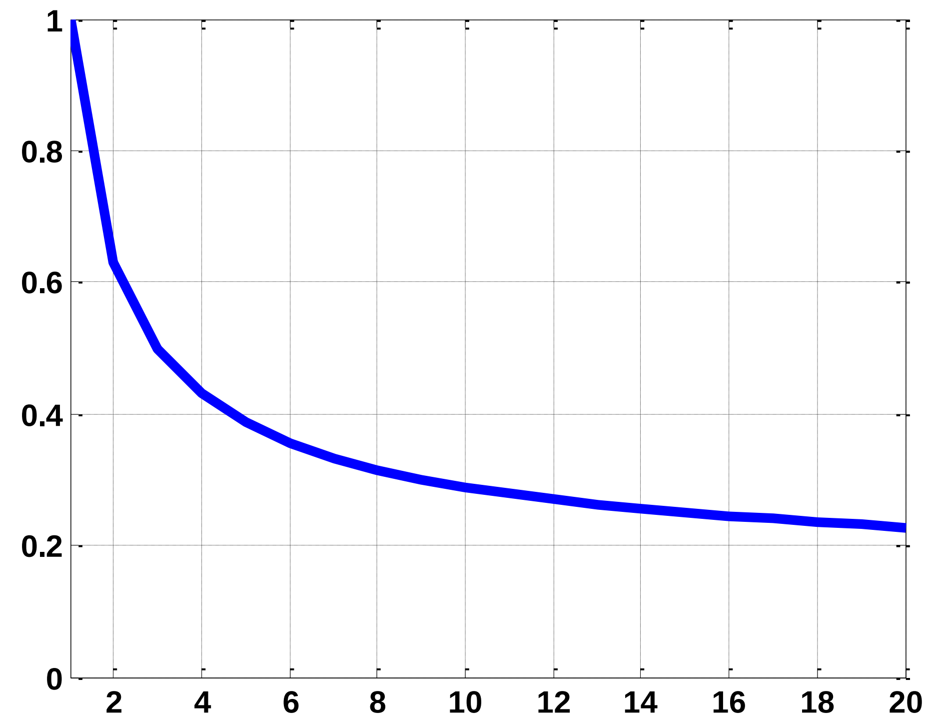
$$\forall \; t \quad w^T \psi(\mathcal{Q}_t, z_t^*) - w^T \psi(\mathcal{Q}_t, z_t) \geq \Delta(z_t^*, z_t) - \xi_t$$

loss function

$$\Delta(z^*, z) = 1 - \mathrm{NDCG}_k(\mathrm{z}^*, z)$$

where

$$NDCG_k = \frac{1}{Z} \sum_{j=1}^{k} \frac{2^{\mathrm{rel}(j)} - 1}{\boxed{\log(1 + j)}}$$



Ranking: Training

Training $\arg\min_{w,\xi} \quad w^T w + C \sum_t \xi_t$

$$\forall \ t \quad w^T \psi(\mathcal{Q}_t, z_t^*) - w^T \psi(\mathcal{Q}_t, z_t) \geq \Delta(z_t^*, z_t) - \xi_t$$

loss function

$$\Delta(z^*, z) = 1 - \mathrm{NDCG}_k(\mathrm{z}^*, z)$$

where

$$NDCG_k = \frac{1}{Z} \sum_{j=1}^{k} \frac{\boxed{2^{\mathrm{rel}(j)} - 1}}{\log(1+j)}$$

Asian + woman + wearing Sunglasses

Ranking: Training

Training $$\arg\min_{w,\xi} \quad w^T w + C \sum_t \xi_t$$

$$\forall \ t \quad w^T\psi(\mathcal{Q}_t, z_t^*) - w^T\psi(\mathcal{Q}_t, z_t) \geq \Delta(z_t^*, z_t) - \xi_t$$

loss function

$$\Delta(z^*, z) = 1 - \mathrm{NDCG}_k(\mathrm{z}^*, z)$$

3

where

$$NDCG_k = \frac{1}{Z}\sum_{j=1}^{k} \frac{\boxed{2^{\mathrm{rel}(j)} - 1}}{\log(1+j)}$$

Asian + woman + wearing Sunglasses

Ranking: Training

Training $\arg\min_{w,\xi} \quad w^T w + C\sum_t \xi_t$

$$\forall \ t \quad w^T\psi(\mathcal{Q}_t, z_t^*) - w^T\psi(\mathcal{Q}_t, z_t) \geq \Delta(z_t^*, z_t) - \xi_t$$

loss function

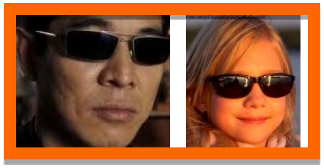
$$\Delta(z^*, z) = 1 - \mathrm{NDCG}_k(\mathrm{z}^*, z)$$

where

$$NDCG_k = \frac{1}{Z}\sum_{j=1}^{k}\frac{\boxed{2^{\mathrm{rel}(j)} - 1}}{\log(1+j)}$$

3

2



Asian + woman + wearing Sunglasses

Ranking: Training

Training $$\arg\min_{w,\xi} \quad w^T w + C \sum_t \xi_t$$

$$\forall \; t \quad w^T\psi(\mathcal{Q}_t, z_t^*) - w^T\psi(\mathcal{Q}_t, z_t) \geq \Delta(z_t^*, z_t) - \xi_t$$

loss function

$$\Delta(z^*, z) = 1 - \mathrm{NDCG}_k(\mathrm{z}^*, z)$$

where

$$NDCG_k = \frac{1}{Z}\sum_{j=1}^{k} \frac{\boxed{2^{\mathrm{rel}(j)} - 1}}{\log(1+j)}$$

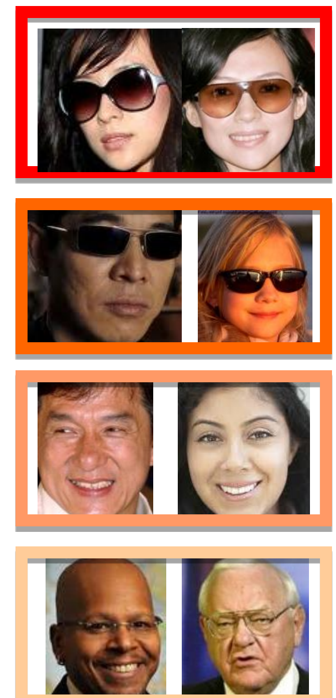
Asian + woman + wearing Sunglasses

3

2

1

0



Ranking: Training

Training $\arg\min_{w,\xi} \quad w^T w + C \sum_t \xi_t$

$$\forall \ t \quad w^T \psi(\mathcal{Q}_t, z_t^*) - w^T \psi(\mathcal{Q}_t, z_t) \geq \Delta(z_t^*, z_t) - \xi_t$$

loss function

$$\Delta(z^*, z) = 1 - \mathrm{NDCG}_k(\mathrm{z}^*, z)$$

where

$$NDCG_k = \frac{1}{Z} \sum_{j=1}^{k} \frac{\boxed{2^{\mathrm{rel}(j)} - 1}}{\log(1+j)}$$

Asian + woman + wearing Sunglasses

Strong Attributes
- Race
- Age
- Gender

Weak Attributes
- Hair Color
- Hair Style
- Facial Hair
- Eyewear

Ranking: Training

Training $$\arg\min_{w,\xi} \quad w^T w + C \textstyle\sum_t \xi_t$$

$$\forall \ t \quad w^T\psi(\mathcal{Q}_t, z_t^*) - w^T\psi(\mathcal{Q}_t, z_t) \geq \Delta(z_t^*, z_t) - \xi_t$$

loss function

$$\Delta(z^*, z) = 1 - \mathrm{NDCG}_k(\mathrm{z}^*, z)$$

where

$$NDCG_k = \frac{1}{\boxed{Z}} \sum_{j=1}^{k} \frac{2^{\mathrm{rel}(j)} - 1}{\log(1+j)}$$

Ranking: Training

Training $$\arg\min_{w,\xi} \quad w^T w + C \sum_t \xi_t$$

$$\forall \ \ t \quad w^T \psi(\mathcal{Q}_t, z_t^*) - w^T \psi(\mathcal{Q}_t, z_t) \geq \Delta(z_t^*, z_t) - \xi_t$$

loss function

$$\Delta(z^*, z) = 1 - \mathrm{NDCG}_k(\mathrm{z}^*, z)$$

where

$$NDCG_k = \frac{1}{Z} \sum_{j=1}^{k} \frac{2^{\mathrm{rel}(j)} - 1}{\log(1+j)}$$

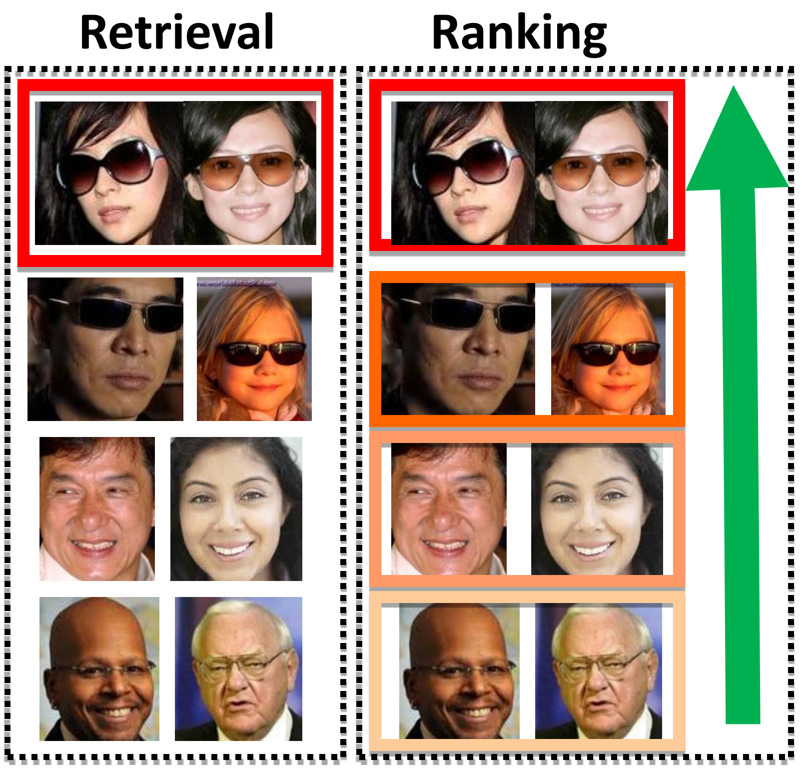
Multi-Attribute Ranking and Retrieval

Ranking and Retrieval

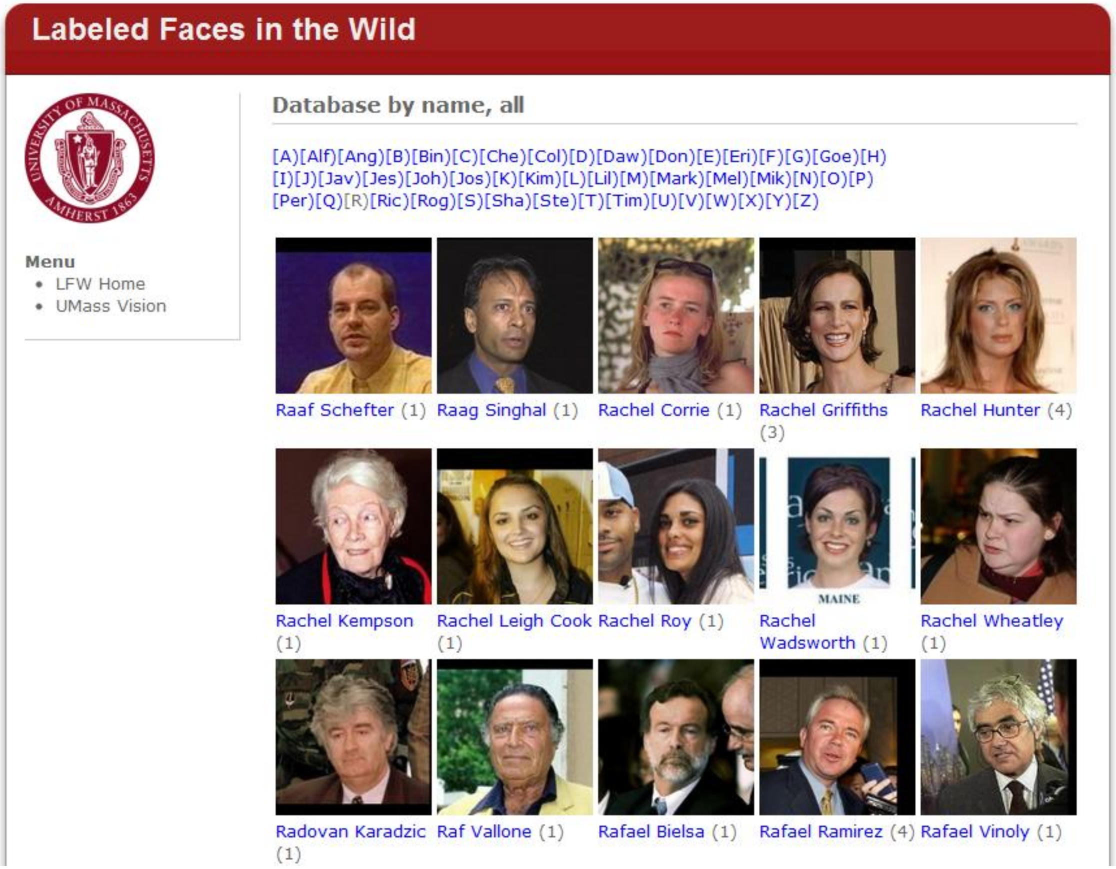
- Typically treated as separate problems

Structured Formulation

- Optimize the same model according to different performance measures



Labeled Faces in the Wild(LFW) Dataset



Attribute Annotation

- **9992 images**
- **27 attributes**

LFW Dataset: Attributes

Race
- Asian
- Black
- White
- Indian

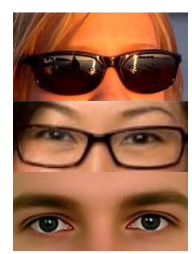
Hair Color
- Black Hair
- Blonde Hair
- Brown Hair
- Gray Hair

Facial Hair
- Mustache
- Beard
- Goatee
- No Beard

Age
- Kid
- Youth
- Middle-Aged
- Senior

Hairstyle
- Long Hair
- Short Hair
- Bangs
- Bald

Eyewear
- Sunglasses
- Eyeglasses
- No Eyewear



Gender
- Sex

Other
- Hat
- Lipstick
- Visible Forehead

LFW Dataset: Feature Extraction

Features

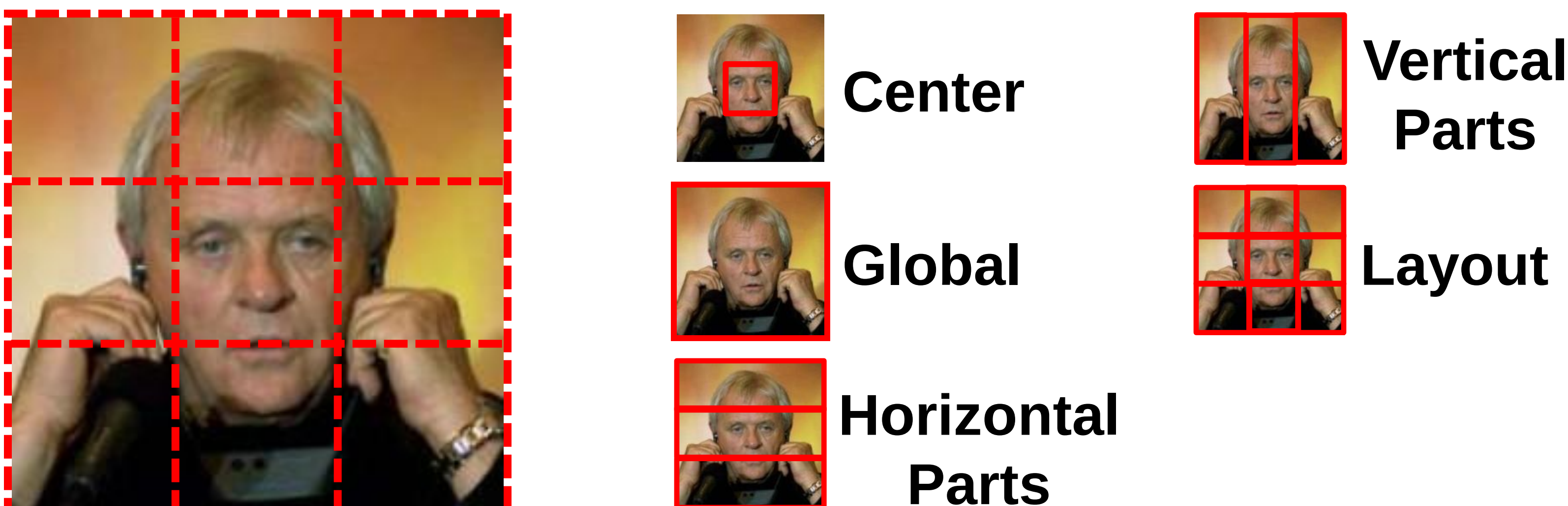
Color
- Color Histograms
- Color Correlograms
- Color Moments
- Color Wavelet

Texture
- Wavelet Texture
- LBP Histogram
- LBP PCA

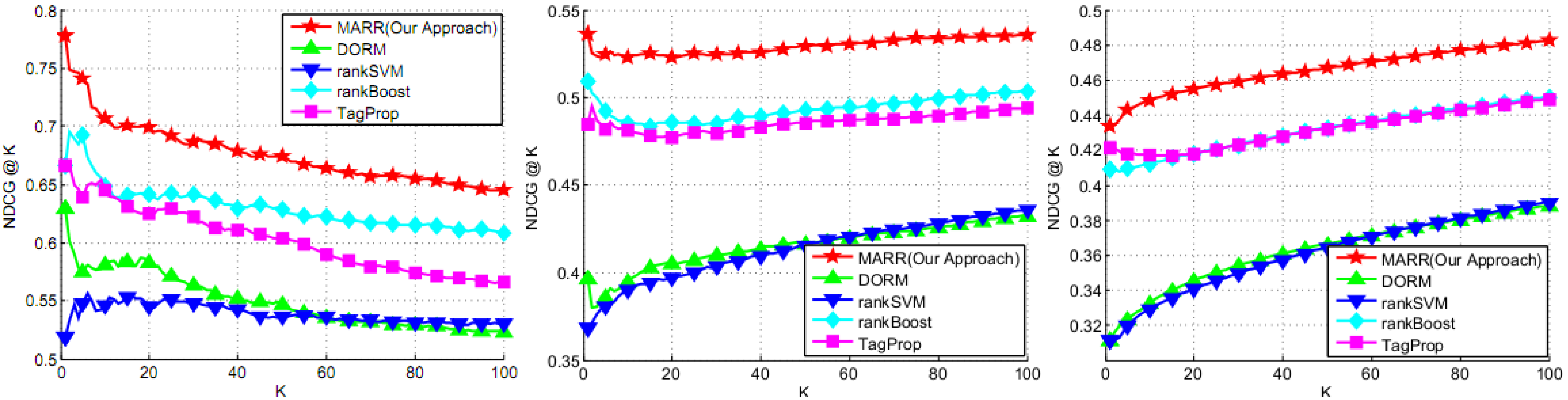
Skin Information
- Skin Bitmap
- Skin Color
- Spatial Skin

Shape
- Edge Histogram
- Shape Moments
- SIFTogram

Features: Spatial Configurations



LFW Dataset: Quantitative Results

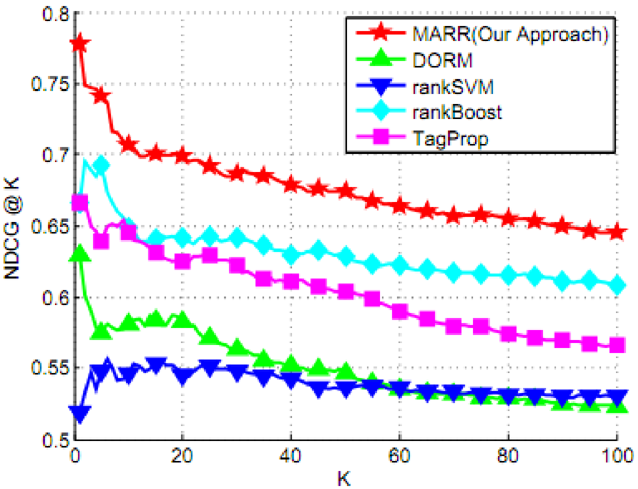


(a) Single Attribute Queries (b) Double Attribute Queries (c) Triple Attribute Queries

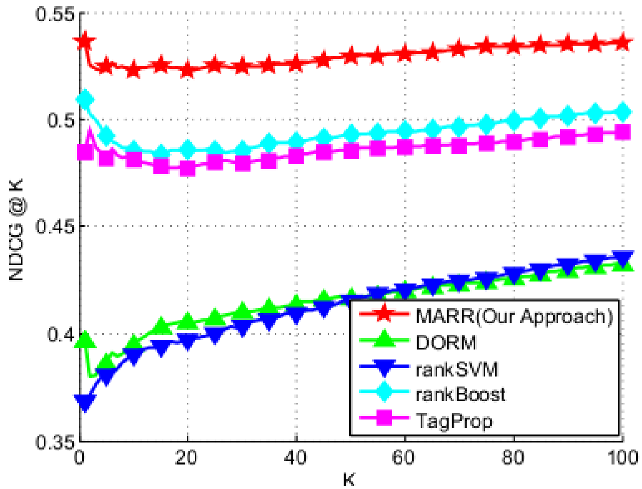
Ranking

- **Baselines**
  - **RankSVM** (T. Joachims, KDD 2002)
  - **RankBoost** (Y. Freund, I. Iyer, R. Schapire, Y. Singer, JMLR 2003)
  - **DORM** (Q. Li, A. Smola, NIPS workshop 2008)
  - **TagProp** (M. Guillaumin, T. Mensink, J. Verbeek, C. Schmid, ICCV 2009)
- **NDCG@K**

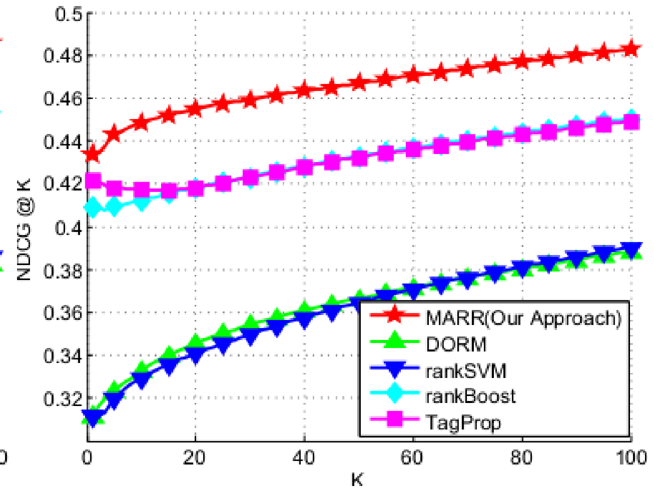
LFW Dataset: Quantitative Results



(a) Single Attribute Queries



(b) Double Attribute Queries

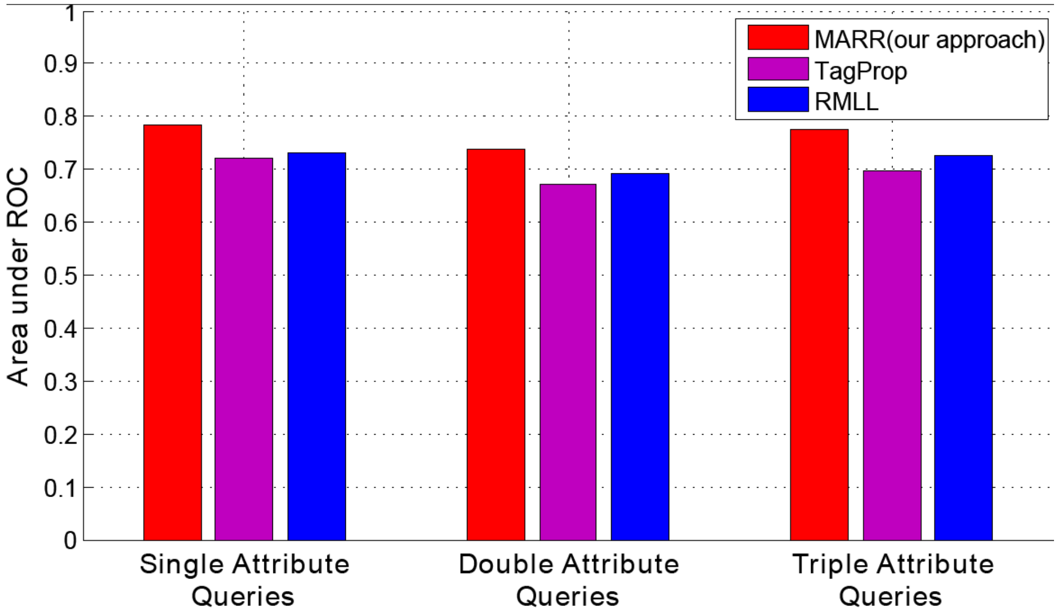


(c) Triple Attribute Queries

Results

- **rankBoost is the 2nd best**
- **Performance gain**
  - **Single Attribute Queries: 8.9% improvement in NDCG@10**
  - **Double Attribute Queries: 7.7% improvement in NDCG@10**
  - **Triple Attribute Queries: 8.8% improvement in NDCG@10**

LFW Dataset: Quantitative Results

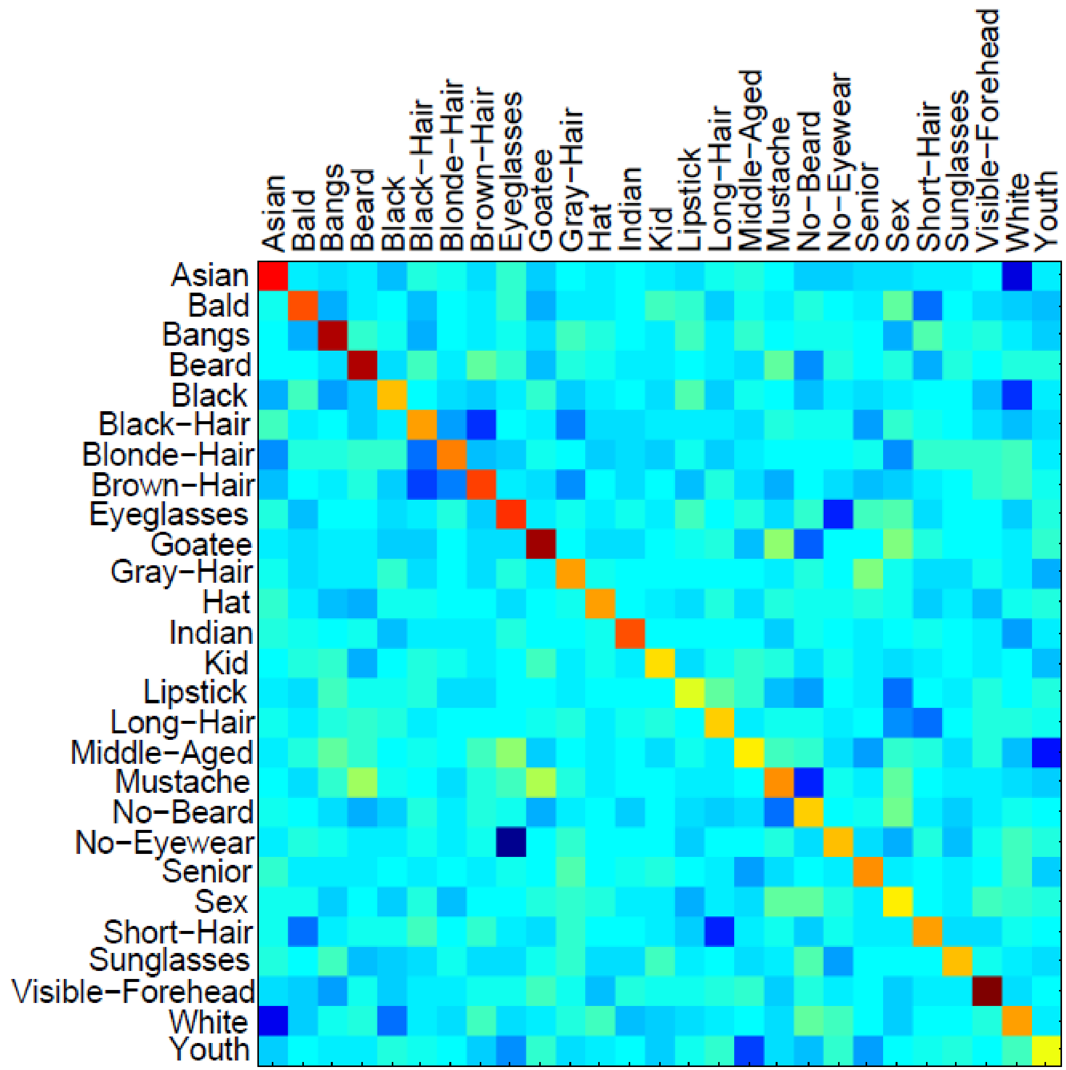


Retrieval

- **Baselines**
  - **Reverse Multi-Label Learning, (J. Petterson and T. Caetano, NIPS 2010)**
  - **TagProp, (M. Guillaumin, T. Mensink, J. Verbeek, C. Schmid, ICCV 2009)**
- **mean Area under ROC**
- **Performance gain**
  - **~5% w.r.t. RMLL**
  - **~7% w.r.t. TagProp**

LFW Dataset: Analysis

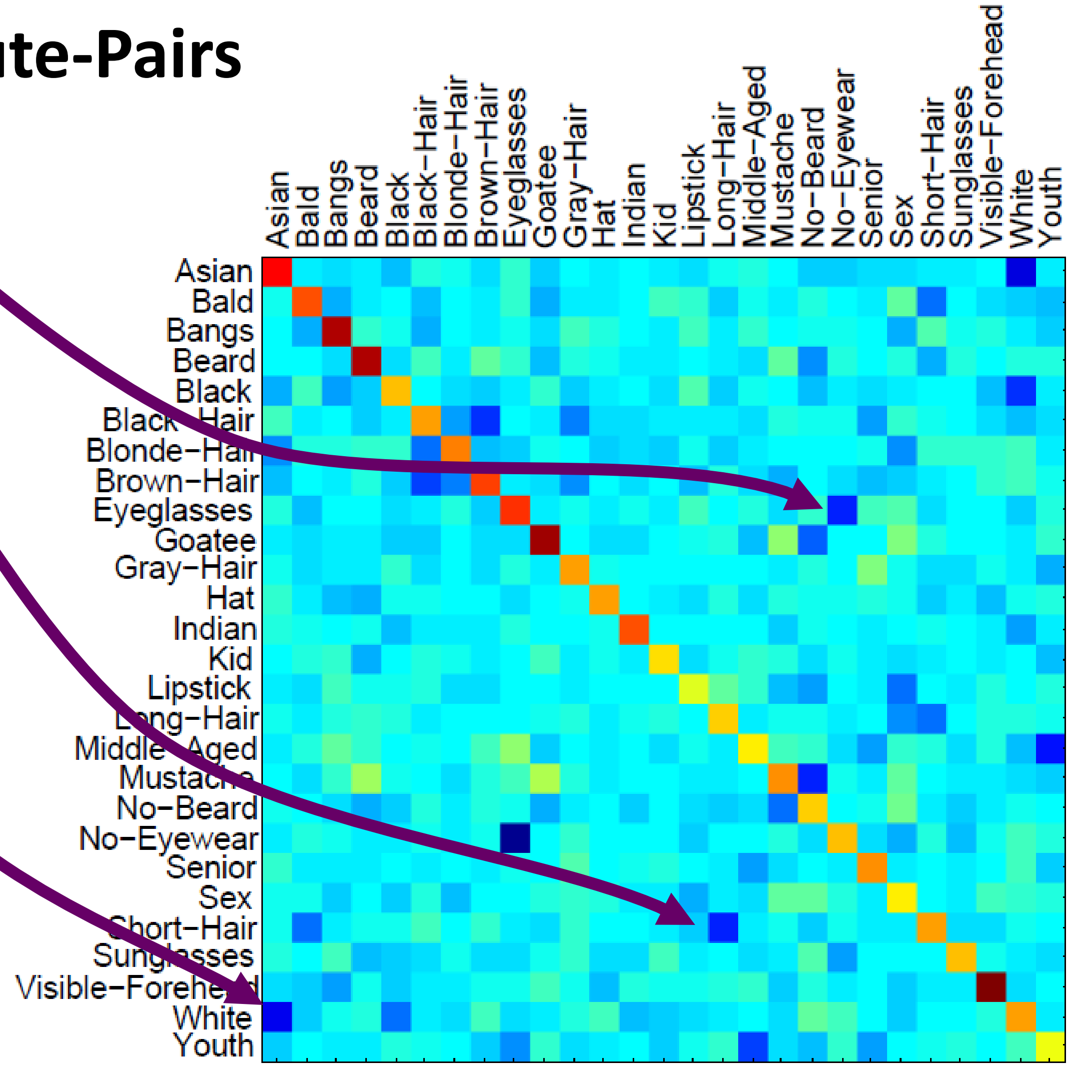
Weights Learnt



LFW Dataset: Analysis

Mutually Exclusive Attribute-Pairs

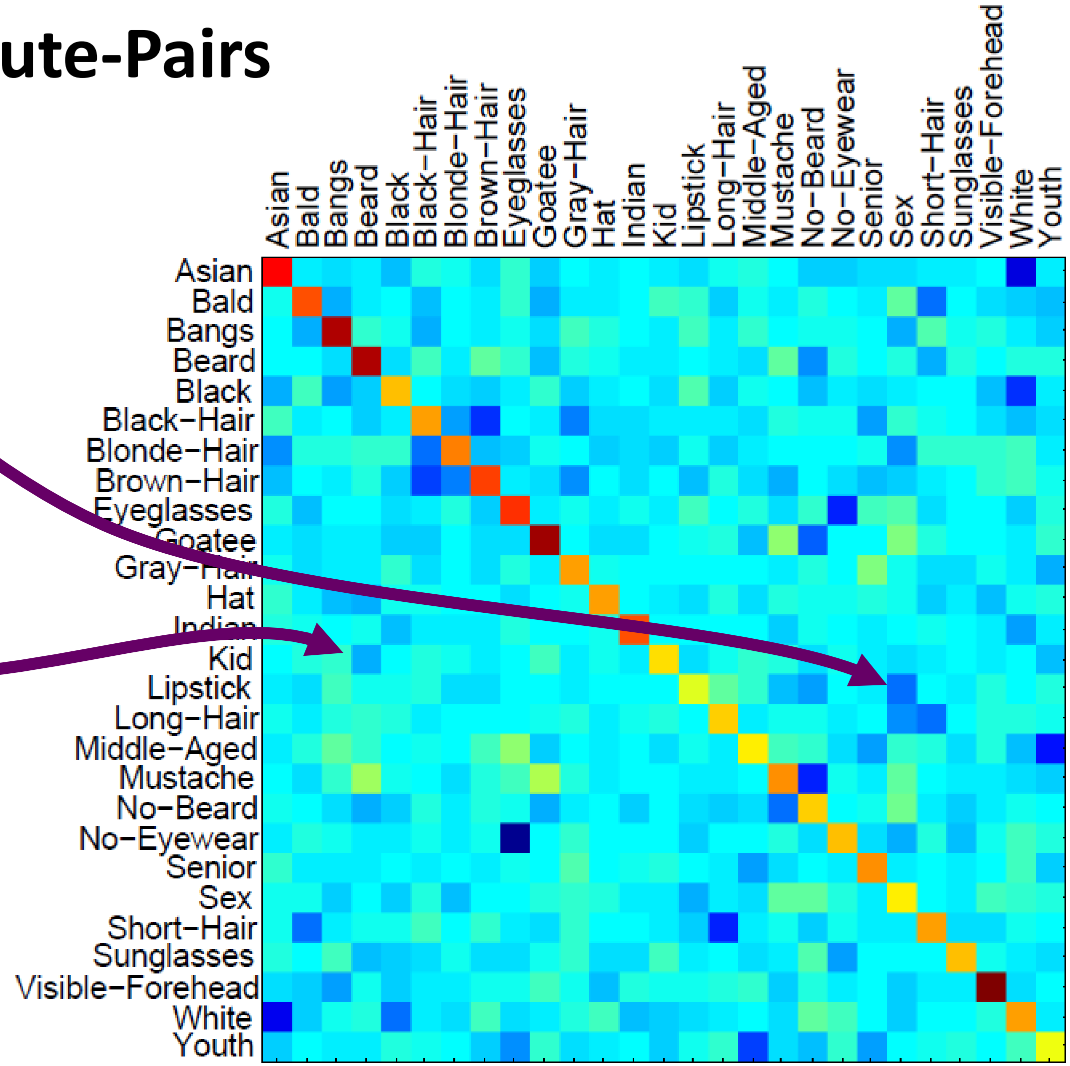
- (White, Asian)
- (Eyeglasses, No-Eyewear)
- (Short-Hair, Long-Hair)



LFW Dataset: Analysis

Rarely Co-occurring Attribute-Pairs

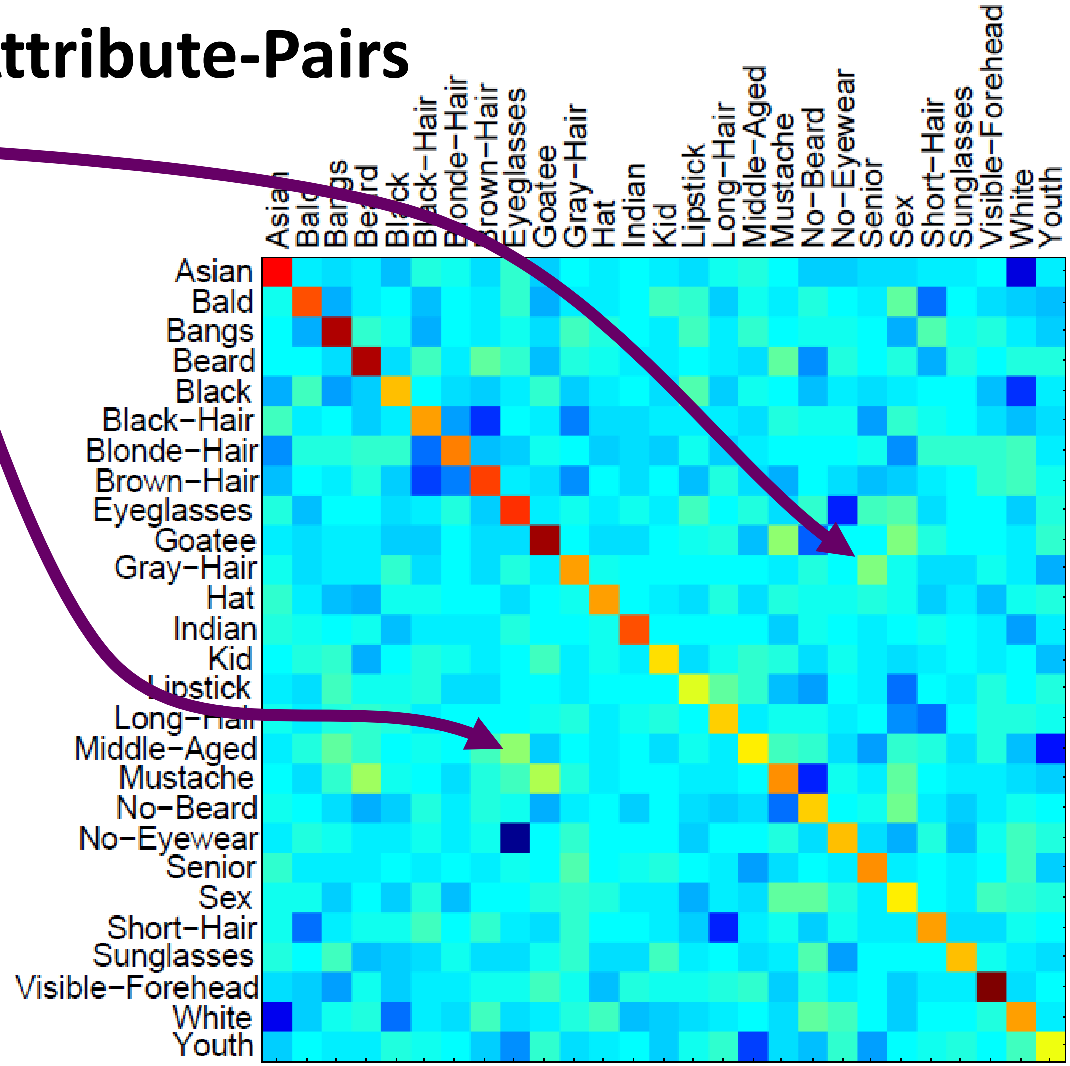
- (Lipstick, Male)
- (Kid, Beard)



LFW Dataset: Analysis

Commonly Co-occurring Attribute-Pairs

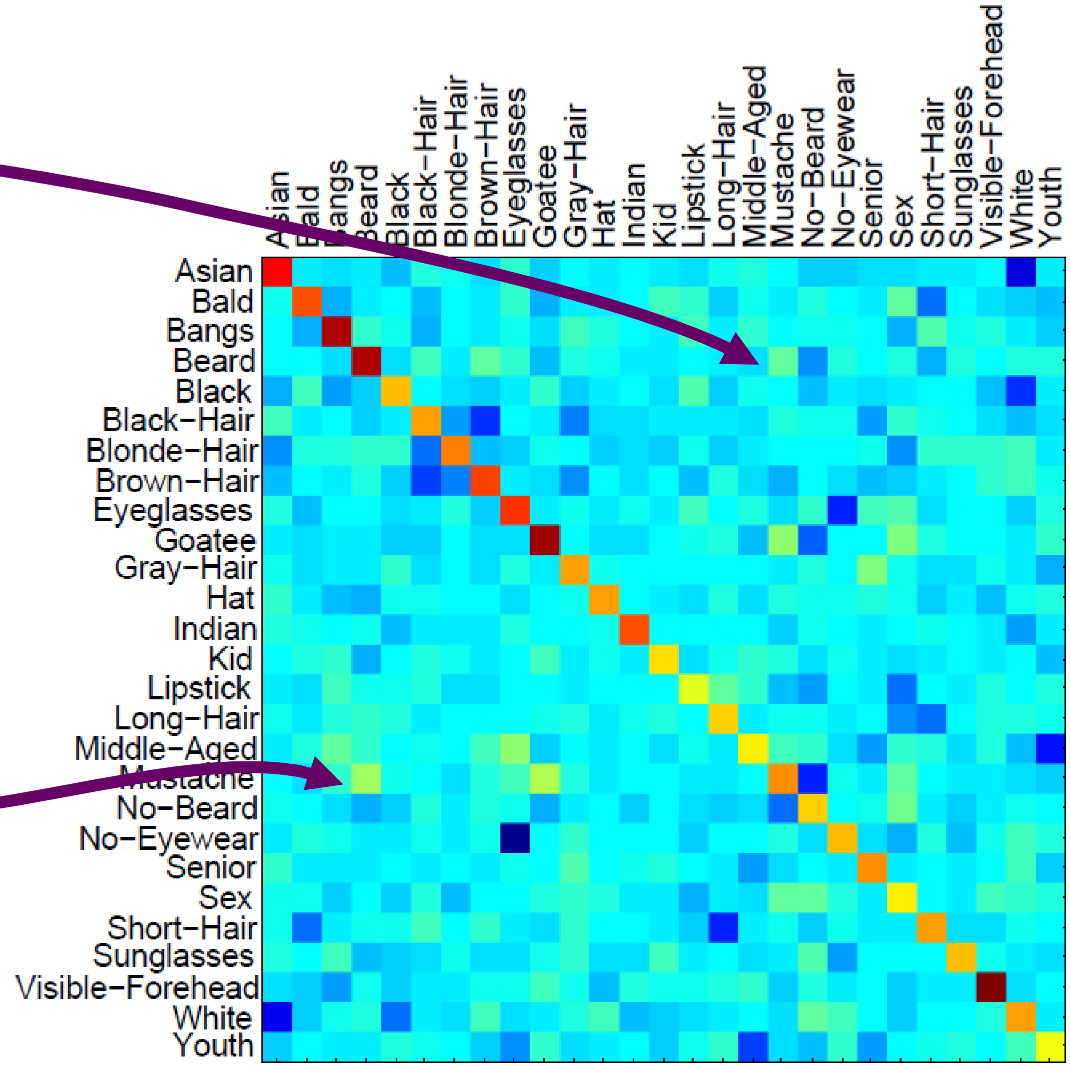
- (Gray-Hair, Senior)
- (Middle-aged, Eyeglasses)



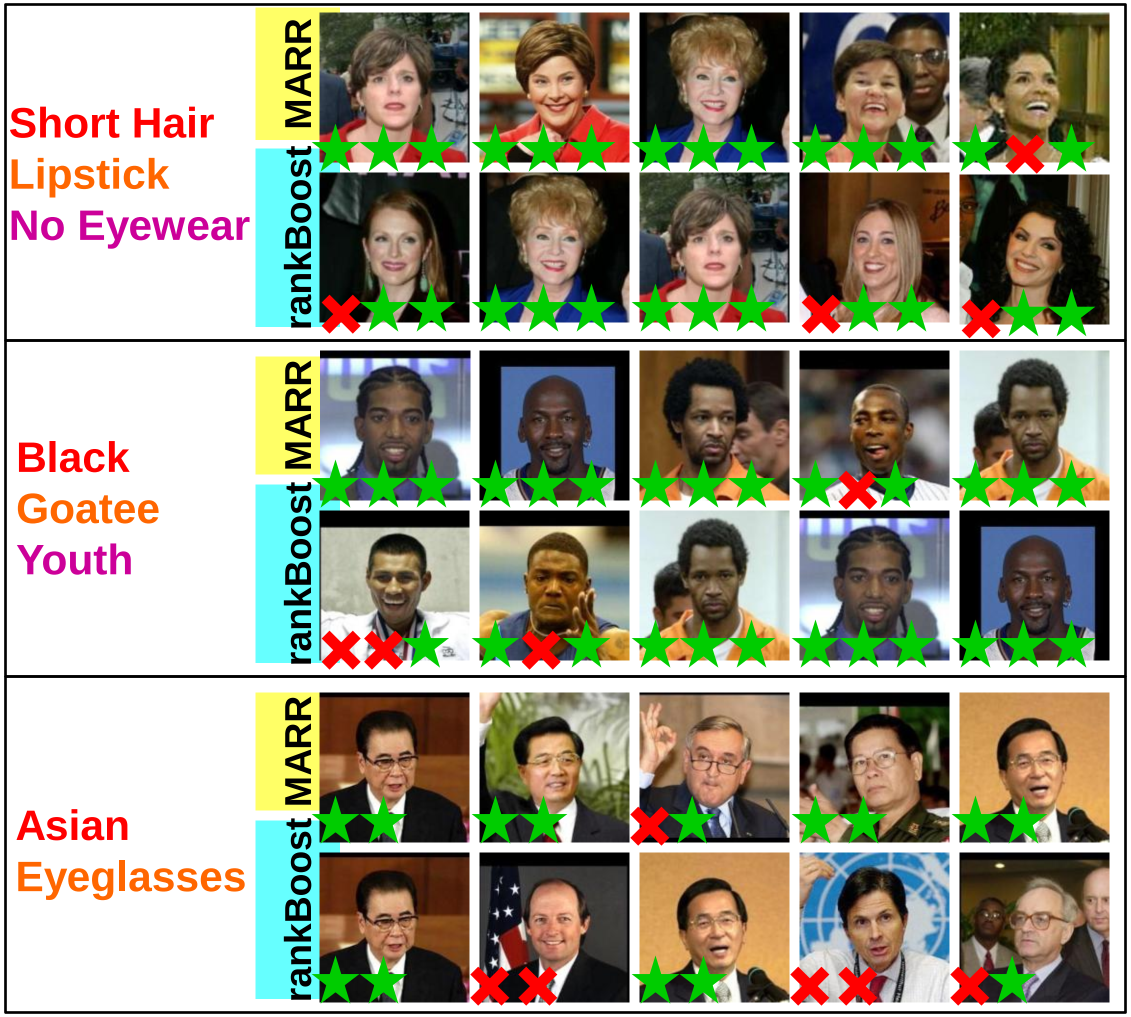
LFW Dataset: Analysis

Asymmetric weights

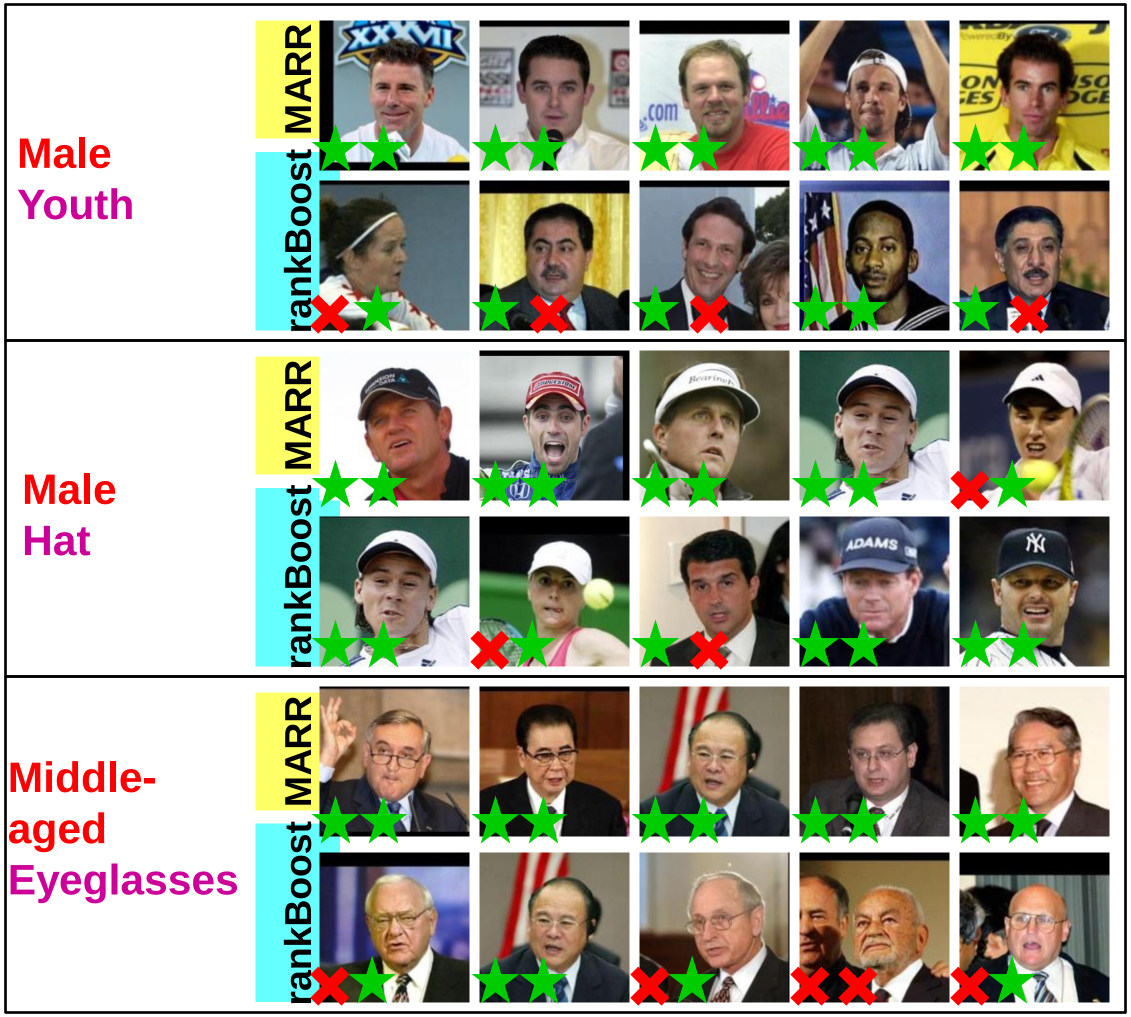
- (Beard, Mustache)
- (Mustache, Beard)



LFW Dataset: Qualitative Results



LFW Dataset: Qualitative Results



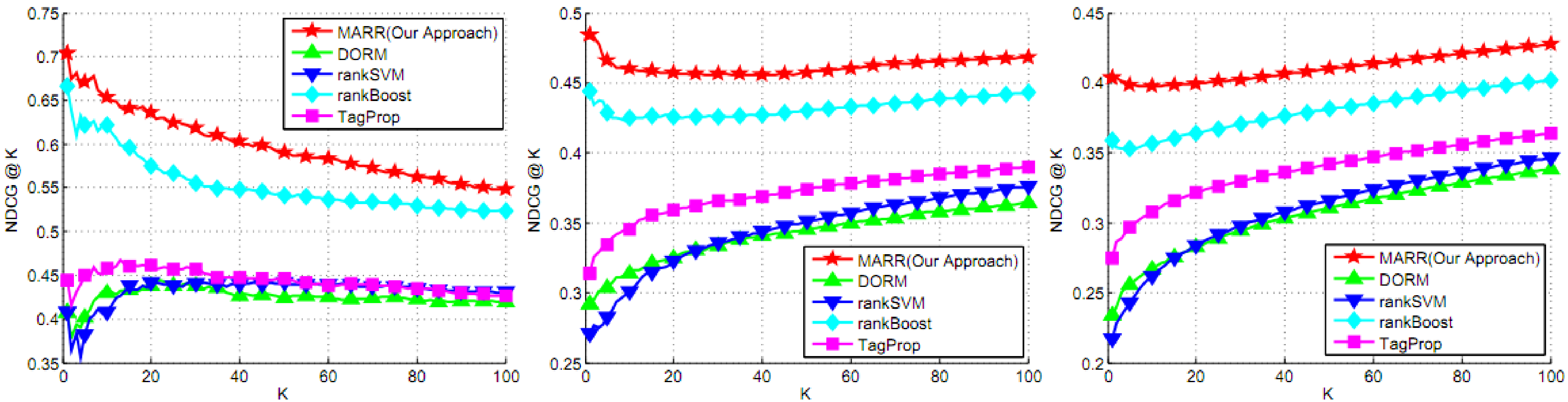
FaceTracer Dataset

(Kumar et. al, ECCV 2008)

Attribute Annotation

- 3000 images
- 27 attributes

FaceTracer Dataset: Quantitative Results



(a) Single Attribute Queries (b) Double Attribute Queries (c) Triple Attribute Queries

Results

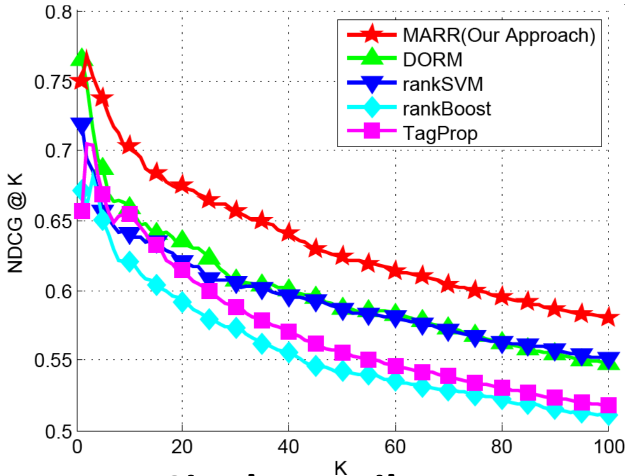
- **rankBoost is the 2nd best**
- **Performance gain**
  - **Single Attribute Queries: 5.0% improvement in NDCG@10**
  - **Double Attribute Queries: 8.1% improvement in NDCG@10**
  - **Triple Attribute Queries: 11.6% improvement in NDCG@10**

Pascal Dataset

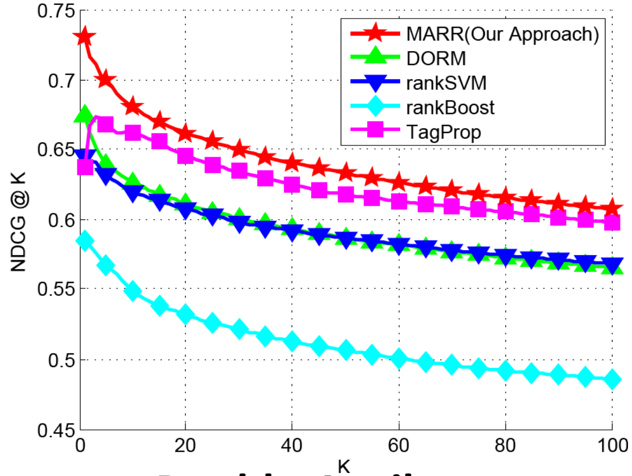
Dataset Statistics

- ➢ **12695 images (6340 train, 6355 test)**
- ➢ **20 classes**
  - **Airplane, bicycle, bus, horse, person, …**
- ➢ **64 Attributes and Parts**
  - **Attributes: 2D Boxy, Round, Vertical Cylinder, Horizontal Cylinder, …**
  - **Parts: Window, Headlight, Text, Leg, …**

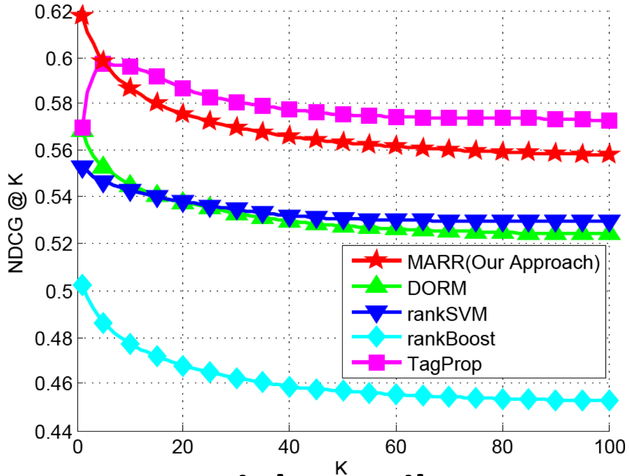
Pascal Dataset: Quantitative Results



Single-Attribute Queries



Double-Attribute Queries



Triple-Attribute Queries

Ranking

Retrieval

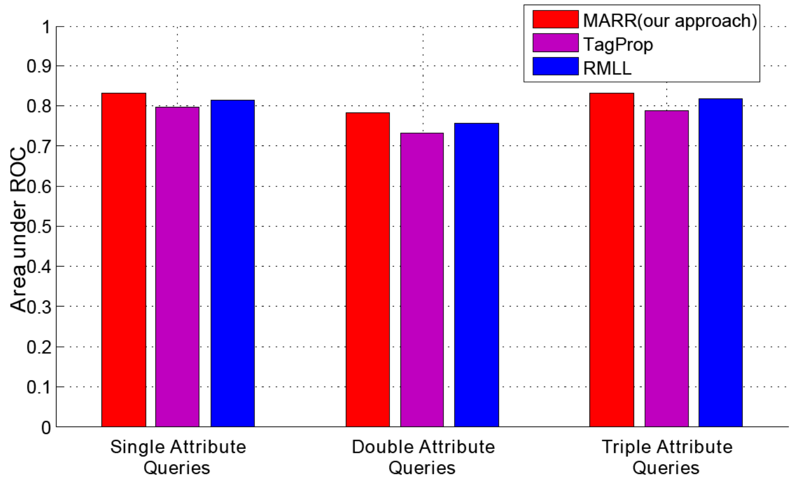
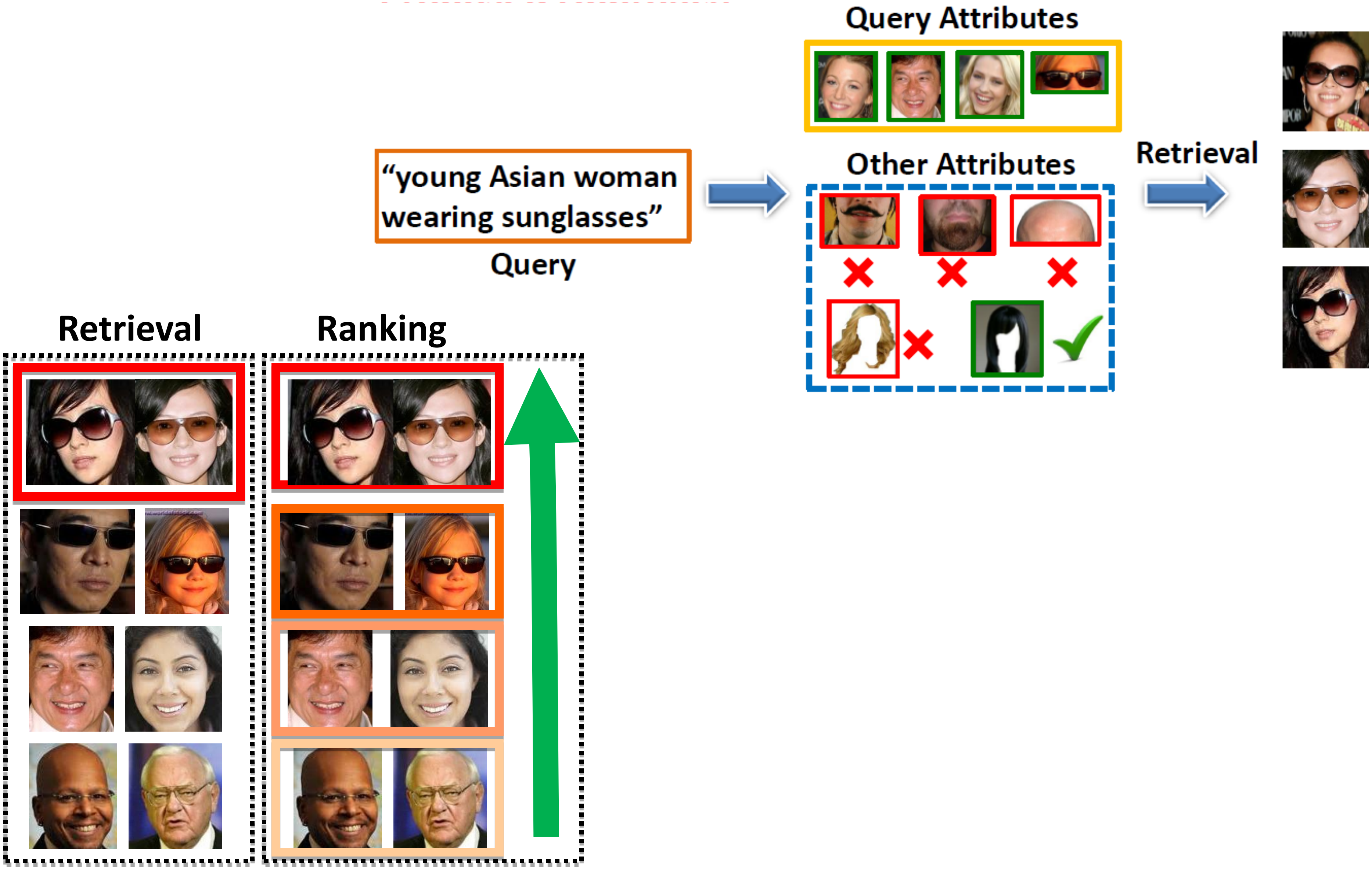


Image Ranking and Retrieval based on Multi-Attribute Queries



Questions?