

Improving People Search Using Query Expansion: How Friends Help To Find People

Thomas Mensink & Jakob Verbeek

LEAR Team, INRIA Rhône-Alpes, Grenoble, France
Oral presentation at ECCV 2008, Marseille, France



Searching George Bush using Yahoo! news photo search



Tibet crisis won't dissuade **Bush** from attending Olympics ...
AFP/File via Yahoo! News - Mar 20 2:00 PM



Bush "cautiously optimistic" about missile deal with Moscow ...
AFP/Pool/File via Yahoo! News - Mar 20 10:24 AM



U.S. President **George W. Bush** shakes hands with Gov. Charlie Crist after arriving on Air Force One in Jacksonville ...
Reuters via Yahoo! News - Mar 18 9:00 AM



President **George W. Bush** receives a hug from his brother former Gov. Jeb **Bush** as ...
AP via Yahoo! News - Mar 18 5:42 PM



President **George W. Bush** speaks at the Blount Island Marine Terminal in Florida ...
Reuters via Yahoo! News - Mar 18 6:54 PM



A protestor dressed as U.S. President **George W. Bush** takes part in a demonstration marking the fifth anniversary of the U.S. invasion of Iraq in Washi ...
Reuters via Yahoo! News - Mar 19 7:48 AM



A protestor walks past a caricature depicting U.S. President **George W. Bush** at a demonstration marking the fifth anniversary of the U.S. invasion of I ...
Reuters via Yahoo! News - Mar 19 7:47 AM



President **George W. Bush** receives a hug from his brother former Gov. Jeb **Bush** as ...
AP via Yahoo! News - Mar 18 5:46 PM



U.S. President **George W. Bush** addresses the Economic Club of New York March 14, ...
Reuters via Yahoo! News - Mar 19 7:28 AM



Bush prays for US troops on Easter holiday ...
AFP/File via Yahoo! News - Mar 21 8:14 AM



President **George W. Bush** waves onstage at the Pentagon in Washington ...
Reuters via Yahoo! News - Mar 19 12:19 PM



White House back pedals on **Bush** comments on Iran bomb ...
AFP/File via Yahoo! News - Mar 21 8:48 AM



A protestor dressed as U.S. President **George W. Bush** takes part in a demonstration in Washington ...
Reuters via Yahoo! News - Mar 19 7:42 AM



President **George W. Bush** presents the President's Volunteer Service award to ...
AP via Yahoo! News - Mar 18 5:49 PM



U.S. President **George W. Bush** addresses the Economic Club of New York ...
Reuters via Yahoo! News - Mar 19 7:27 AM

Searching George Bush using Yahoo! news photo search



Tibet crisis won't dissuade **Bush** from attending Olympics ...
AFP/File via Yahoo! News - Mar 20 2:00 PM



Bush cautiously optimistic about missile deal with Moscow ...
AFP/Pool/File via Yahoo! News - Mar 20 10:24 AM



U.S. President **George W. Bush** shakes hands with Gov. Charlie Crist after arriving on Air Force One in Jacksonville ...
Reuters via Yahoo! News - Mar 18 9:00 AM



President **George W. Bush** receives a hug from his brother former Gov. Jeb **Bush** as ...
AP via Yahoo! News - Mar 18 5:42 PM



President **George W. Bush** speaks at the Blount Island Marine Terminal in Florida ...
Reuters via Yahoo! News - Mar 18 6:54 PM



A protestor dressed as U.S. President **George W. Bush** takes part in a demonstration marking the fifth anniversary of the U.S. invasion of Iraq in Washi ...
Reuters via Yahoo! News - Mar 19 7:48 AM



A protestor walks past a caricature depicting U.S. President **George W. Bush** at a demonstration marking the fifth anniversary of the U.S. invasion of I ...
Reuters via Yahoo! News - Mar 19 7:47 AM



President **George W. Bush** receives a hug from his brother former Gov. Jeb **Bush** as ...
AP via Yahoo! News - Mar 18 5:46 PM



U.S. President **George W. Bush** addresses the Economic Club of New York March 14, ...
Reuters via Yahoo! News - Mar 19 7:28 AM



Bush prays for US troops on Easter holiday ...
AFP/File via Yahoo! News - Mar 21 8:14 AM



President **George W. Bush** waves onstage at the Pentagon in Washington ...
Reuters via Yahoo! News - Mar 19 12:19 PM



White House back pedals on **Bush** comments on Iran bomb ...
AFP/File via Yahoo! News - Mar 21 8:48 AM



A protestor dressed as U.S. President **George Bush** takes part in a demonstration in Washington ...
Reuters via Yahoo! News - Mar 19 7:42 AM



President **George W. Bush** presents the President's Volunteer Service award to ...
AP via Yahoo! News - Mar 18 5:49 PM

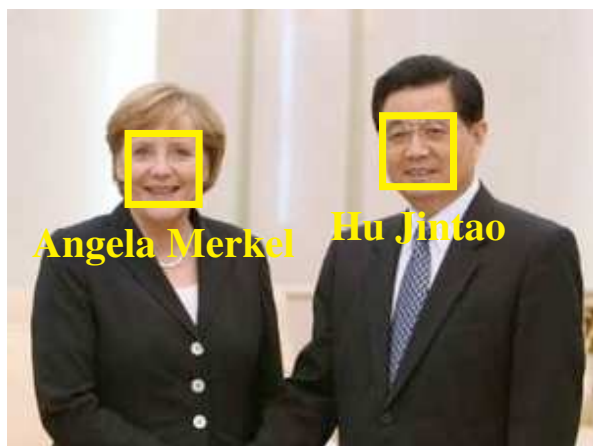


U.S. President **George W. Bush** addresses the Economic Club of New York ...
Reuters via Yahoo! News - Mar 19 7:27 AM

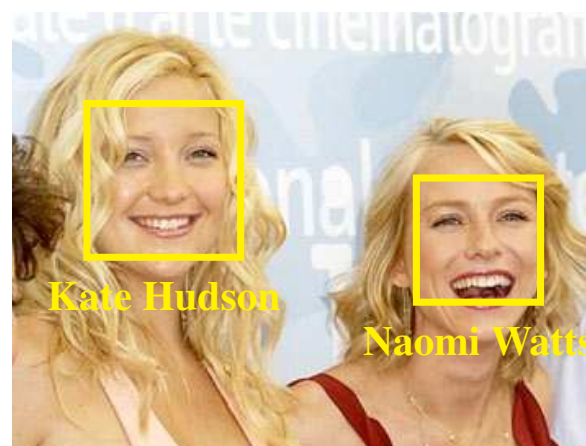
Presentation outline

- **Problem and challenges**
- **Related work and motivation of our work**
- **Query expansion implemented in two approaches**
 - ▶ **generative mixture model**
 - ▶ **linear discriminant model**
- **Conclusion**

Using captions as weak supervision to find people



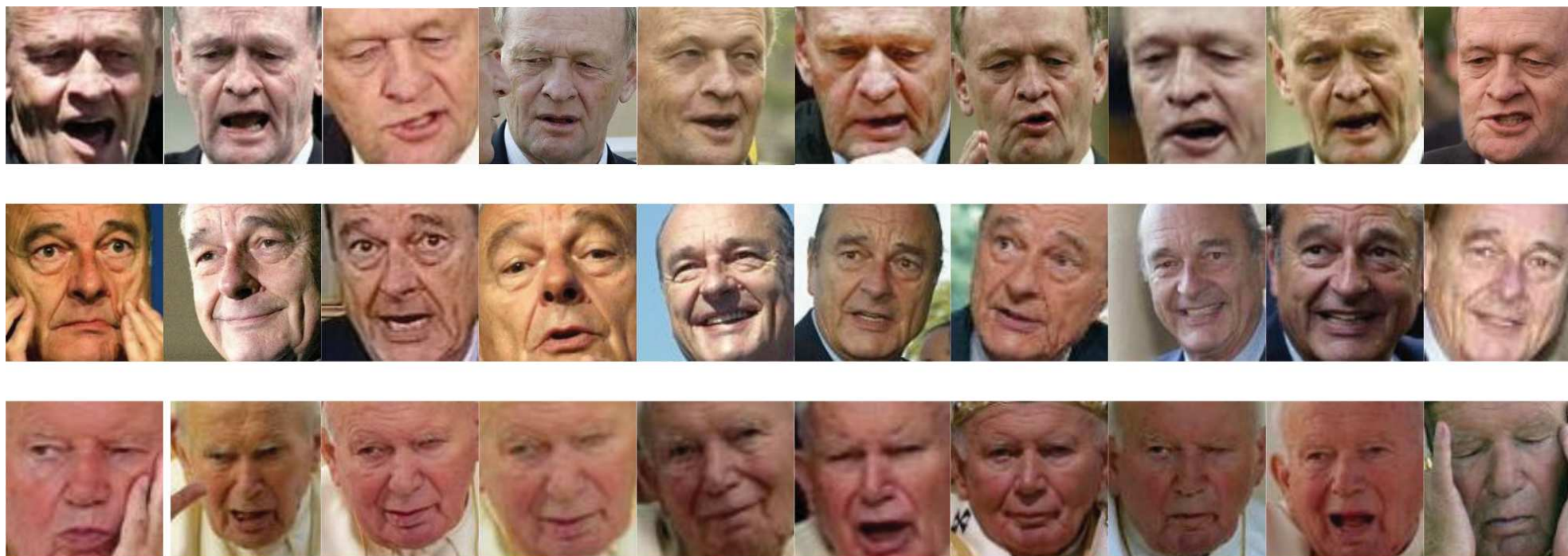
German Chancellor **Angela Merkel** shakes hands with Chinese President **Hu Jintao** (...)



Kate Hudson and **Naomi Watts**, *Le Divorce*, Venice Film Festival - 8/31/2003.

- **Task:** Find all face images of a particular person
 - ▶ Manual construction of labeled training sets costly
 - ▶ Continued labeling effort needed for online system with new people arriving
- **Using caption alone does not work:** only 44% of faces are correct
 - ▶ Averaged over our set of 23 people with ground truth annotation
- **Better approach:** combine information in caption with visual analysis

Challenges in the data



- **Appearance variations:** illumination, expression, pose, scale, occlusion, ...

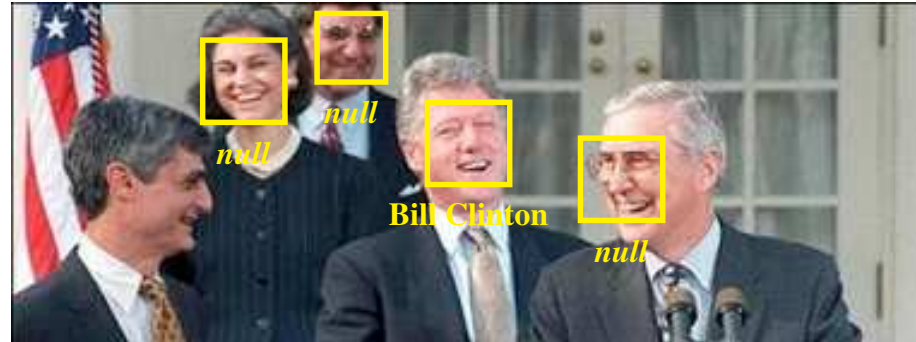
Challenges in the data



- **Appearance variations:** illumination, expression, pose, scale, occlusion, ...
- **Naming variations:** *Bush, George W. Bush, US president, ...*
- **Imperfect detectors:** both for names & faces

Work on related problems

- **Matching all names and faces** in captioned news images: many possible matches (Berg et al. CVPR '04)



Lloyd Bentsen is pictured here announcing his retirement in 1994 at the White House with former US President **Bill Clinton**, Chief of Staff **Leon Panetta**, **Robert Rubin** and **Judy Rubin** (...)

- **Naming characters in TV series** combining tracking and video-script alignment (Everingham et al. BMVC '06)



- **Labeling personal photo collections** exploiting social networks, e.g. FaceBook to predict co-occurrence (Stone et al. CVPR '08)



Previous work on our problem: find all instances of X

- **Approach in previous work on same problem:**

(Ozkan & Duygulu CVPR'06), (Guillaumin, Mensink, Verbeek & Schmid CVPR'08)

- ▶ Given query name X
- ▶ Select all images with X in caption
- ▶ Analyze faces in those images to rank or classify them

- **Underlying principles:**

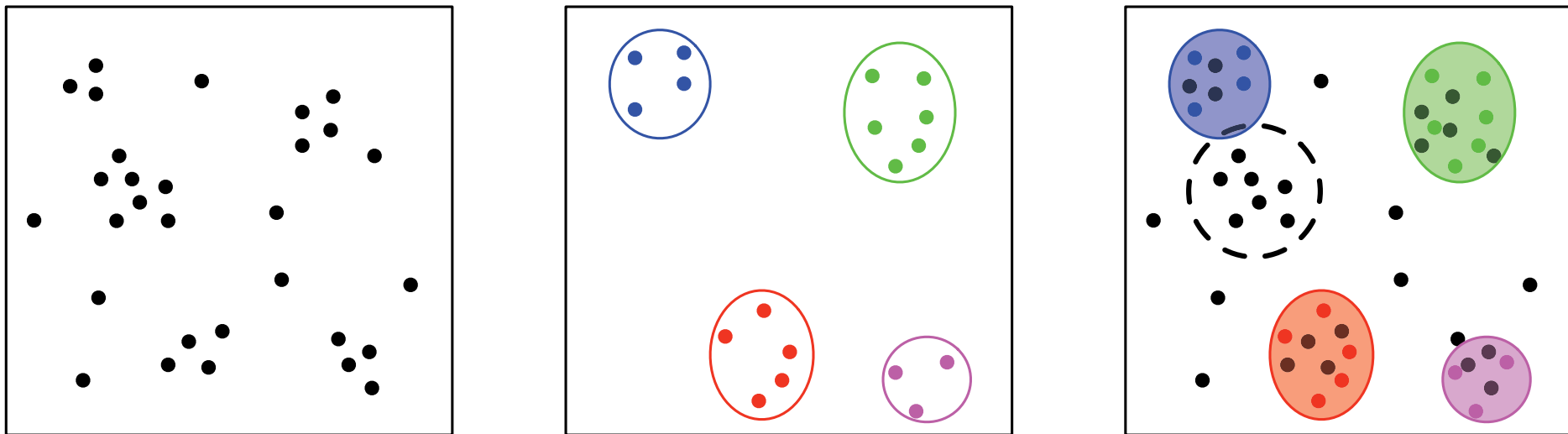
- ▶ Text filtering makes queried person the most frequent
- ▶ Task is reduced to finding the big mode among clutter

- **Failure case:**

- ▶ If text-filtering yields a precision $< 40\%$
- ▶ Mode finding might return wrong person

Improving people search using query expansions

- **Motivation:** avoid confusion with co-occurring people
- **Query Expansion:** use more images than just those with X in caption
 - ▶ Find names co-occurring with the queried person: “friends”
 - ▶ Query database for images with friends in caption, *but not* X
 - ▶ Adds “negative” examples, different from typical query expansion in retrieval
- **Example:** search for “Bush”, expand with “Powell”, “Rumsfeld”, and “Rice”



Initial situation (left), models based on queries for friends (middle), simplified person identification (right).

Query expansion example: Berlusconi



Faces 1 - 25



Faces 1 - 25

<-- 1 2 3 4 5 6 7 -->

Show All

Show Friends | Hide Friends | Show Console output | Hide Console

Friend 1 bush



Friend 2 schroeder



Friend 3 saddam hussein

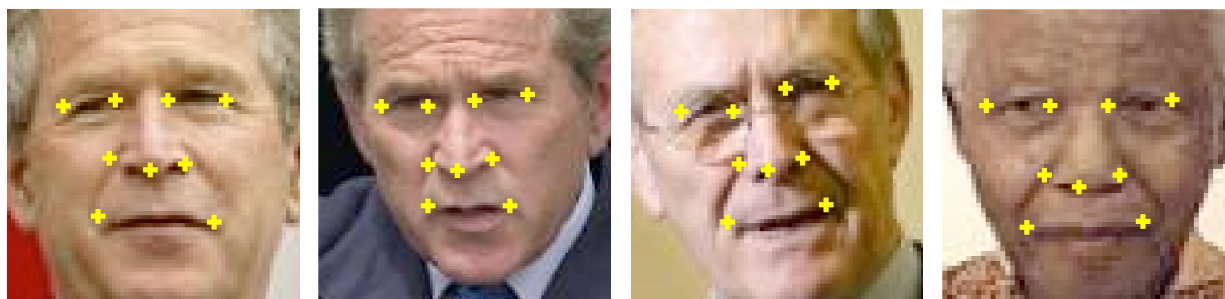


Friend 4 tony blair

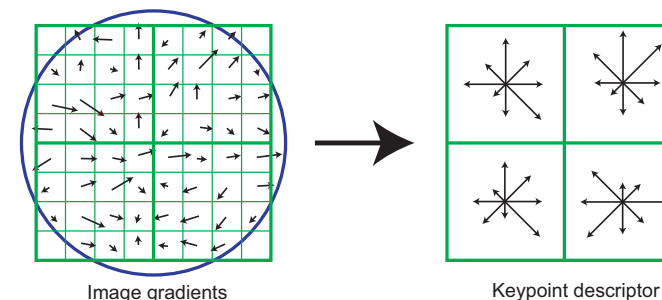


Data and pre-processing pipeline

- **Data set:** 15.000 captioned images from *Yahoo! News* (Collected by Tamara Berg)
 - ▶ Hand labeling of all faces in images with one of the 23 query names in caption
- **Name detection:** off-the-shelf detector (Deschacht & Moens, WOLP'06)
- **Face detection:** off-the-shelf detector (Mikolajczyk, Schmid & Zisserman, ECCV'04)
- **Face representation:** based on local features
 - ▶ Detector of facial features: mouth, nose, eyes, ... supervised training (Everingham et al. BMVC '06)
 - ▶ Concatenate SIFT descriptors of all facial feature



Examples of facial feature detection



SIFT descriptor

Presentation outline

- Problem and challenges
- Related work and motivation of our work
- **Query expansion implemented in two approaches**
 - ▶ generative mixture model
 - ▶ linear discriminant model
- **Conclusion**

Approach 1: Gaussian mixture model

- **Goal:** which, if any, of the F faces in this image is X ?
 - ▶ Coded in assignment variable $\gamma \in \{0, 1, \dots, F\}$
- **Mixture model** over set of feature vectors \mathcal{F}
 - ▶ **Data not i.i.d. !**
 - ▶ A-priori over γ : equal for $\gamma \neq 0$
 - ▶ Gaussian density for faces of X
 - ▶ generic “background model” for other faces

Approach 1: Gaussian mixture model

- **Goal:** which, if any, of the F faces in this image is X ?
 - ▶ Coded in assignment variable $\gamma \in \{0, 1, \dots, F\}$
- **Mixture model** over set of feature vectors \mathcal{F}
 - ▶ **Data not i.i.d. !**
 - ▶ A-priori over γ : equal for $\gamma \neq 0$
 - ▶ Gaussian density for faces of X
 - ▶ generic “background model” for other faces

$$p(\mathcal{F}) = \sum_{\gamma=0}^F p(\gamma)p(\mathcal{F}|\gamma), \quad p(\mathcal{F}|\gamma) = \prod_{i=1}^F p(f_i|\gamma),$$

$$p(f_i|\gamma) = \begin{cases} p_{\text{BG}}(f_i) = \mathcal{N}(f_i; \mu_{\text{BG}}, \Sigma_{\text{BG}}) & \text{if } \gamma \neq i \\ p_{\text{FG}}(f_i) = \mathcal{N}(f_i; \mu_{\text{FG}}, \Sigma_{\text{FG}}) & \text{if } \gamma = i \end{cases}$$

Approach 1: Gaussian mixture model

- **Goal:** which, if any, of the F faces in this image is X ?
 - ▶ Coded in assignment variable $\gamma \in \{0, 1, \dots, F\}$
- **Mixture model** over set of feature vectors \mathcal{F}
 - ▶ **Data not i.i.d. !**
 - ▶ A-priori over γ : equal for $\gamma \neq 0$
 - ▶ Gaussian density for faces of X
 - ▶ generic “background model” for other faces

$$p(\mathcal{F}) = \sum_{\gamma=0}^F p(\gamma)p(\mathcal{F}|\gamma), \quad p(\mathcal{F}|\gamma) = \prod_{i=1}^F p(f_i|\gamma),$$

$$p(f_i|\gamma) = \begin{cases} p_{\text{BG}}(f_i) = \mathcal{N}(f_i; \mu_{\text{BG}}, \Sigma_{\text{BG}}) & \text{if } \gamma \neq i \\ p_{\text{FG}}(f_i) = \mathcal{N}(f_i; \mu_{\text{FG}}, \Sigma_{\text{FG}}) & \text{if } \gamma = i \end{cases}$$

- **EM algorithm** to find face model and assignments
 - ▶ Background model fixed, only foreground Gaussian and prior updated
 - ▶ After convergence evaluate $p(\gamma|\mathcal{F})$

Query expansion in the Gaussian mixture model

- **Learn a Gaussian for each friend using standard 2-component model**
- **Use images with friend in the caption but without X**
 - ▶ At most 15 friends, at least 5 images per friend
- **Define new background model:** mixture of N friends + generic model

$$p_{\text{BG}}(f) = \frac{1}{N+1} \sum_{n=0}^N \mathcal{N}(f; \mu_n, \Sigma_n)$$

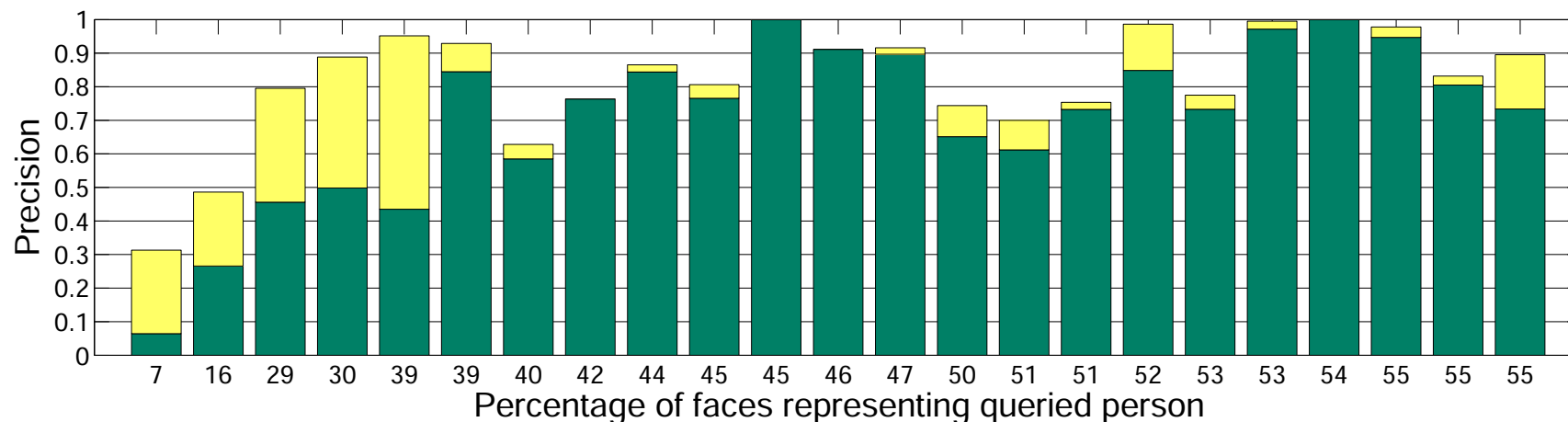
Query expansion in the Gaussian mixture model

- **Learn a Gaussian for each friend using standard 2-component model**
- **Use images with friend in the caption but without X**
 - ▶ At most 15 friends, at least 5 images per friend
- **Define new background model:** mixture of N friends + generic model

$$p_{\text{BG}}(f) = \frac{1}{N+1} \sum_{n=0}^N \mathcal{N}(f; \mu_n, \Sigma_n)$$

- **Run EM on standard 2-component model using mixture background**
- **Possibly errors in friend models, but trained on images without X**

Results using Gaussian mixture model

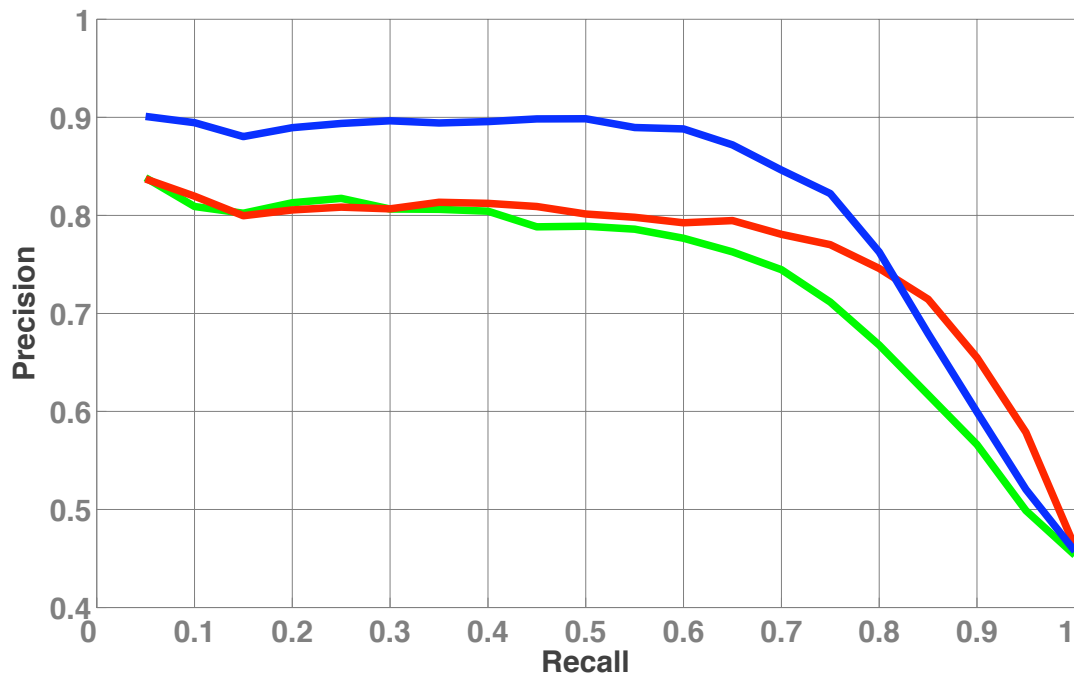


Comparing mixture model without (green), and with (yellow) query expansion

- **Failure case of previous work:** low text-based precision ($<40\%$)
- **Progress mainly in those cases:** 20%-50% increase in precision

Results using Gaussian mixture model (2)

- **Green:** 1 background Gaussian: fitted to all faces with X in caption
- **Red:** 1 background Gaussian: fitted to all faces in expansion
- **Blue:** Mixture background: composed of Gaussian for friends + expansion



Precision averaged over the 23 queries at different levels of recall

Approach 2: logistic discriminant model

- **Motivation:** diagonal Gaussian model rather limited
 - ▶ Too little data to allow learning of richer model
- **Logistic discriminant:** same nr. of parameters put to use for separation
 - ▶ Laplace prior for sparsity in the weight vector

$$p(y = 1|f) = \frac{1}{1 + \exp(w^\top f)}$$

Approach 2: logistic discriminant model

- **Motivation:** diagonal Gaussian model rather limited
 - ▶ Too little data to allow learning of richer model
- **Logistic discriminant:** same nr. of parameters put to use for separation
 - ▶ Laplace prior for sparsity in the weight vector

$$p(y = 1|f) = \frac{1}{1 + \exp(w^\top f)}$$

- **Positive examples:** all faces in images with X in caption
- **Negative examples:**
 - ▶ random set of faces without X in caption
 - ▶ faces in query expansion

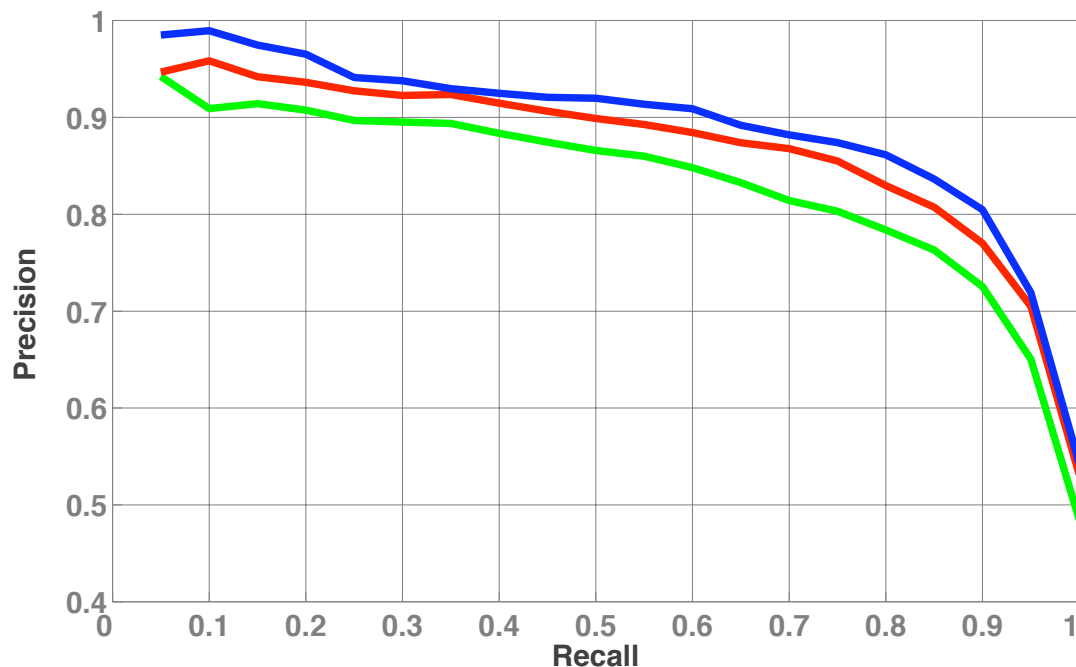
Iterative re-labeling of noisy positive data

- **Positive data is very noisy**
 - ▶ on average only 44% correct
- **People appear once per image**
 - ▶ most of the time
- **Iterative re-labeling of noisy positive examples**
 - ▶ Learn initial classifier from all faces after text search
 - ▶ Re-label most suspicious faces as negative
 - ▶ Re-train classifier using new labels
 - ▶ Repeat until one face per positive image is left



Results logistic discriminant model

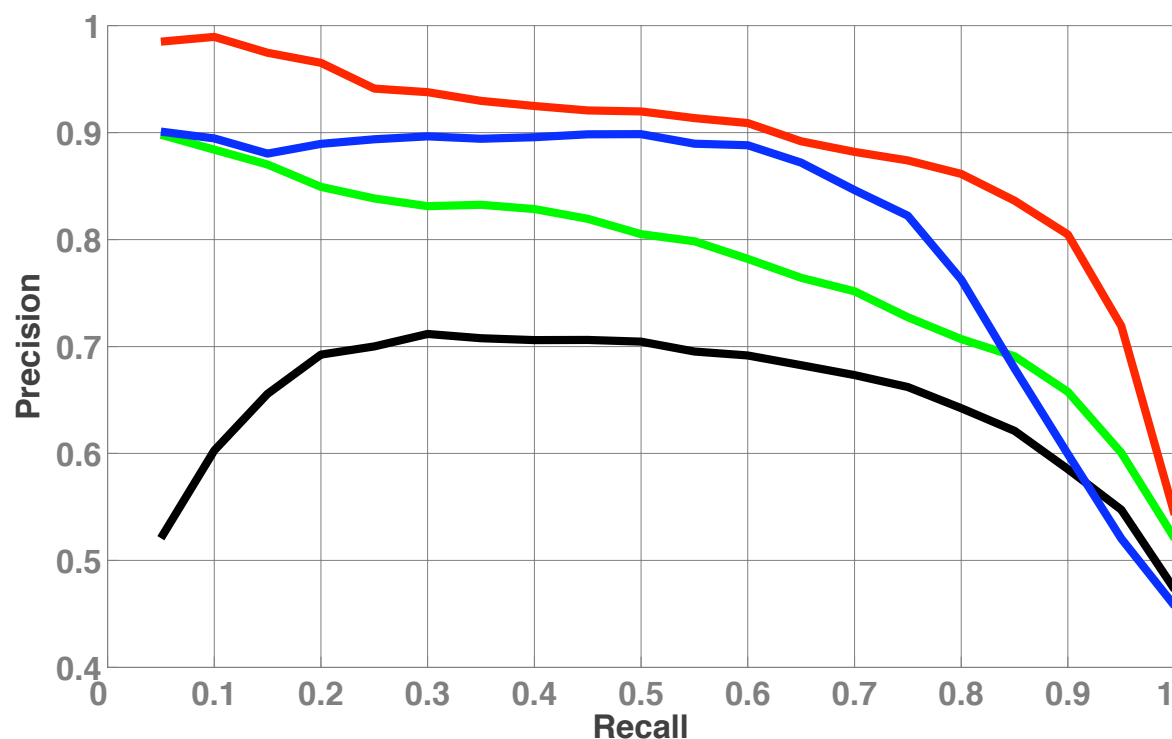
- **Green:** Discriminate noisy positives from a set of random faces
- **Red:** Iteratively re-labeling of noisy positive set
- **Blue:** Idem, but use query expansion as negative example set



Precision averaged over the 23 queries at different levels of recall

Comparison of results with state-of-the-art

- **Red:** Discriminative model, re-labeling, query expansion (this paper)
- **Blue:** Gaussian mixture, query expansion (this paper)
- **Green:** Similarity-based method (our CVPR '08)
- **Black:** Similarity-based method (Ozkan & Duygulu, CVPR '06)



Precision averaged over the 23 queries at different levels of recall

- More than 10% increase in precision for recall levels up to 90%

Performance in absence of captions

- Classifiers learned from caption based supervision
- Test on “Labeled Faces in the Wild” data set
 - ▶ public data set, 13.000 hand labeled faces, no captions



In each row: top 10 ranked faces for one person

Conclusions

- **Query expansion improves people search**
 - ▶ Generative model benefits most from expansion
 - ▶ Discriminative model yields best performance
- **Significant progress when text-based precision is low**
 - ▶ These remain the most difficult cases
- **Our methods using query expansion improves earlier work**
 - ▶ +10% precision compared to our CVPR'08 work
 - ▶ +20% precision compared to Ozkan & Duygulu CVPR'06

Questions?

