

# FaceTracer:

## A Search Engine for Large Collections of Images with Faces

Authors:

Neeraj Kumar,

Peter Belhumeur, Shree Nayar

Columbia University

Presented by: Girish Malkarnenkar

19<sup>th</sup> October 2012

CS395T Visual Recognition

# Motivation

- Current search engines **use text annotations** to find images based on facial appearance.
- **Problems** with this approach:
  1. Manual labeling is time consuming
  2. Textual annotations can be misleading/incorrect
  3. Annotated images are only a small subset of all the images

# Google Images then...











smiling asian men with glasses - Google Image Search - Mozilla Firefox

file Edit View History Bookmarks Tools Help

http://images.google.com/images?um=1&tab=w&hl=en&q=smiling%20asian%20men%20with%20glasses

Images Showing: All image sizes Results 1 - 20 of about 26,400 for **smiling asian men with glasses** with SafeSearch on. (0.82 seconds)

View all web results for smiling asian men with glasses

 <p>Portrait of two elderly Asian men 167 x 250 - 22k - jpg www.jupiterimages.com</p>	 <p>asian man, asian men, asian people 167 x 250 - 22k - jpg www.jupiterimages.com [ More from images.jupiterimages.com ]</p>	 <p>Man with Glasses, Portrait 216 x 360 - 23k - jpg www.comstock.com</p>	 <p>Man with Glasses, Portrait 216 x 360 - 20k - jpg www.comstock.com [ More from images.comstock.com ]</p>	 <p>... to be less happy than men, ... 1000 x 664 - 273k - jpg www.menscience.com</p>
 <p>... to be less happy than men, ... 200 x 133 - 4k - jpg www.menscience.com</p>	 <p>Stock Photo: Smiling Arms Folded 253 x 380 - 45k www.istockphoto.com</p>	 <p>a handsome middle aged asian ... 166 x 300 - 11k www.photosearch.com</p>	 <p>... percent of Asian men suffer from ... 143 x 215 - 8k - png ahp.blogspot.com</p>	 <p>... The Eastbay Asian Youth Center 400 x 300 - 76k - jpg www.starfonex.com</p>
 <p>This is why the men are smiling 400 x 300 - 93k - jpg daitak.typepad.com</p>	 <p>Very Southeast Asian, very Thai 500 x 375 - 46k - jpg tenghouse.typepad.com</p>	 <p>... I came upon a monk smiling at 400 x 365 - 45k - jpg www.kachang.net</p>	 <p>In her novel China Men, Maxine Hong ... 350 x 276 - 17k - jpg www.kachang.net</p>	 <p>... Southeast Asian nations that men ... 364 x 245 - 30k - jpg content.cdbb.org</p>

Done

# Their method then..





# Their method now...



**Mug**Hunt2 smiling asian men with glasses Search

Asian ✕ Eyeglasses ✕ Male ✕ Smiling ✕ Strong Nose-Mouth Lines ✕

Searching 1.91 million images. Found 50 images in 0.0478 seconds.



# Google Images now...



## All results

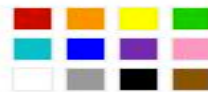
By subject  
Personal

## Any size

Large  
Medium  
Icon  
Larger than...  
Exactly...

## Any color

Full color  
Black and white



## Any type

Face  
Photo  
Clip art  
Line drawing

## Standard view

Show sizes

## PROBLEM STATEMENT:

1. Goal: A search engine based on **both Facial and Image appearance**
2. Since there are billions of images and hundreds of possible attributes, and we can only hope to get a few thousands of manual labels, the labeling of images needs to be done **automatically in a scalable manner**

# Database Creation: Downloading images

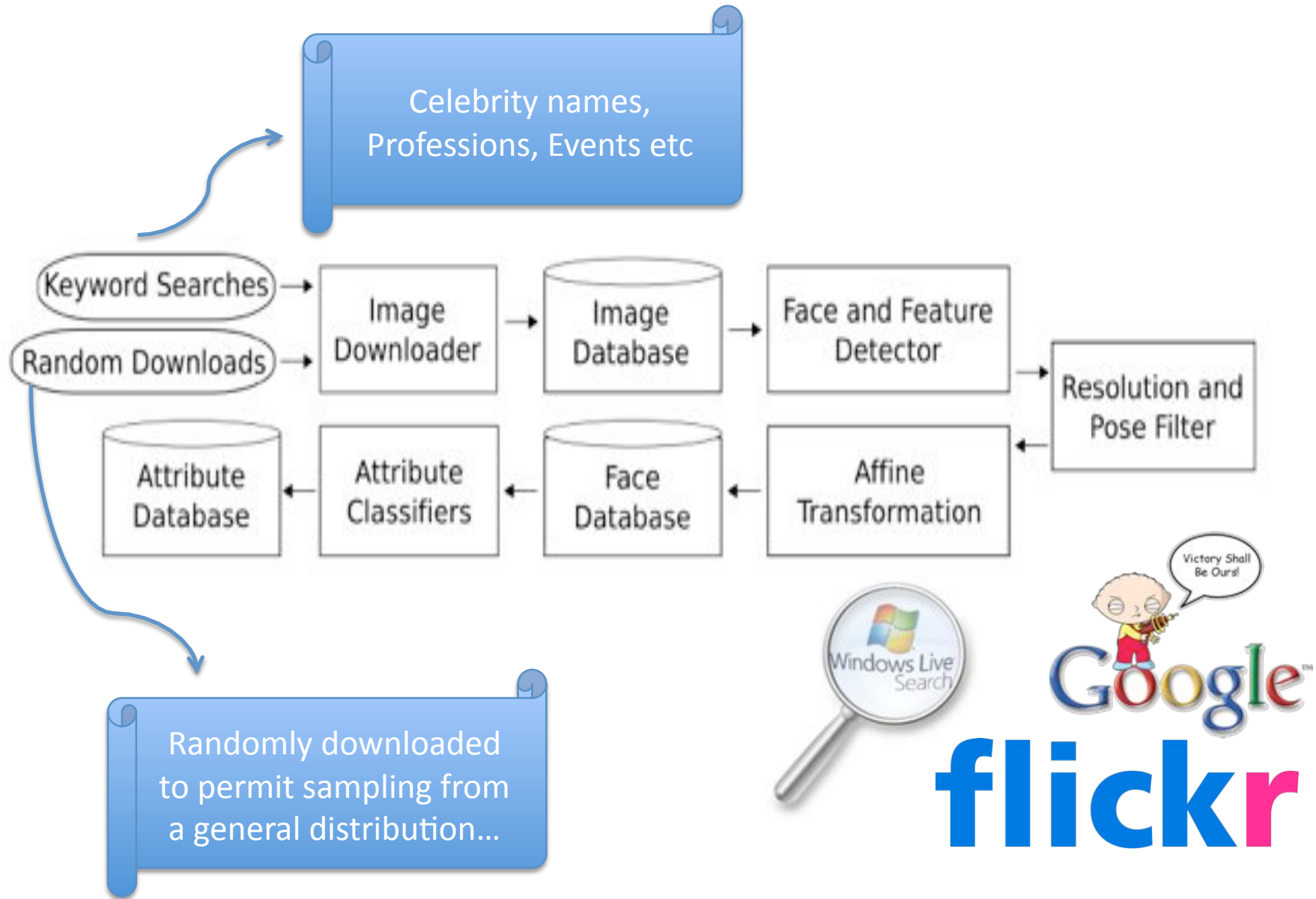


Image from: ECCV 2008 paper, Logos from: [link1](#), [link2](#), [link3](#)



# Database Creation: Face detection

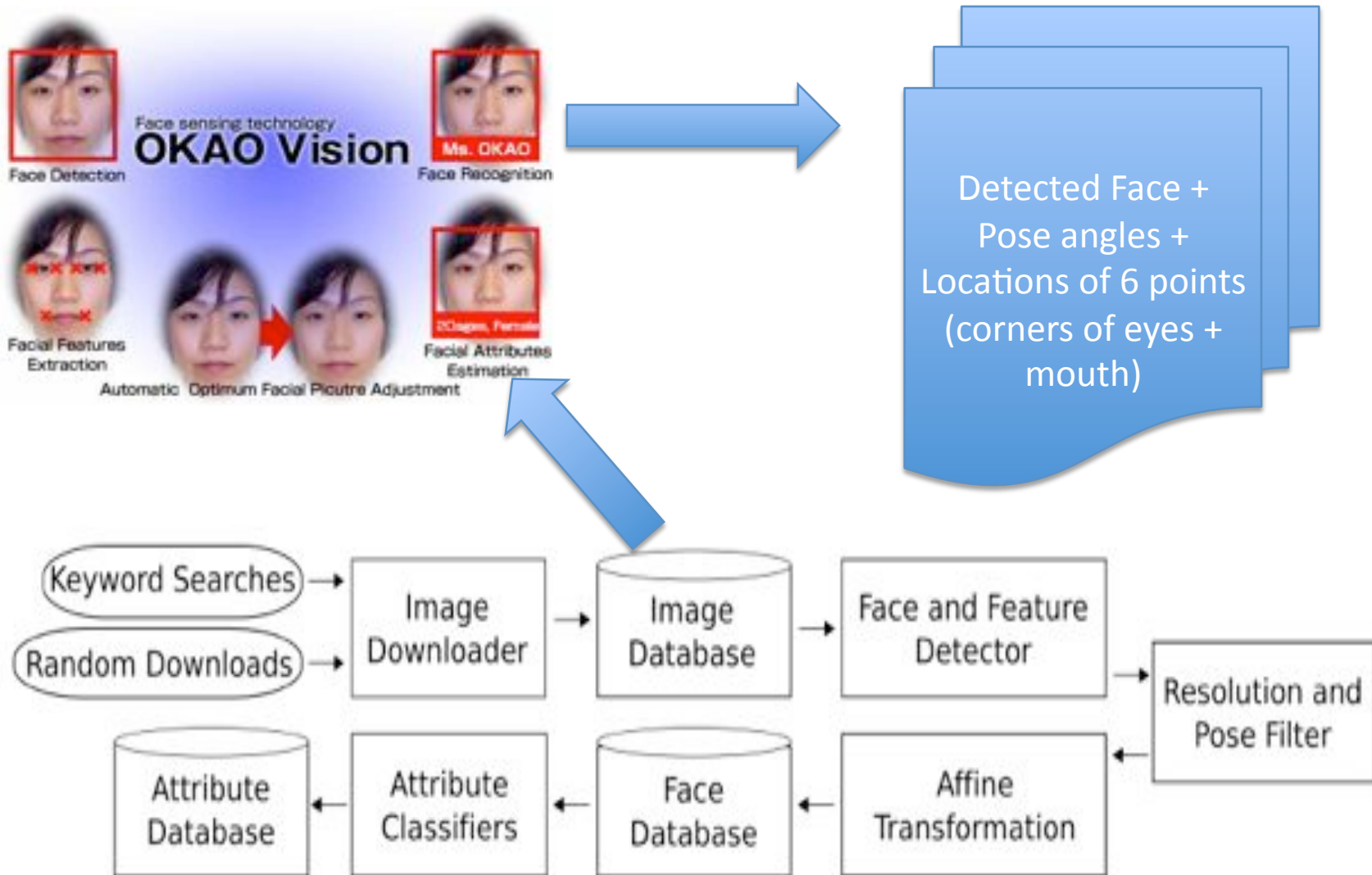
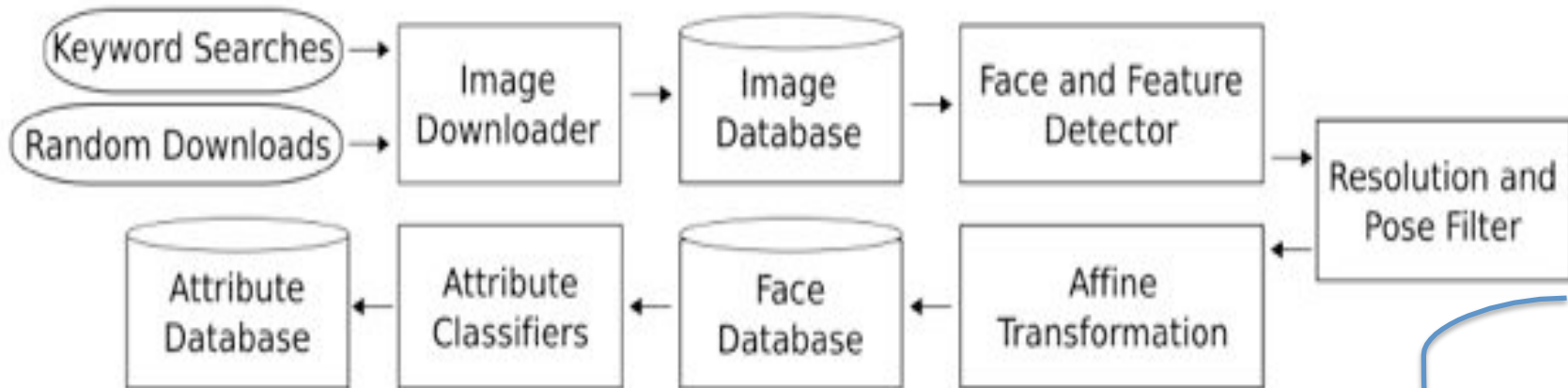


Image from: ECCV 2008 paper, OKAO logo: [link](#)

# Database Creation: Filter/Transformation



Affine transformation to canonical frontal pose using least squares on the 6 points w.r.t a template

Filter detected faces by pose (+/- 10 degrees from front/center)



Image from: ECCV 2008 paper, [Anakin](#), [Luke](#), [Affine](#), [Tamara Berg's paper](#)

# Image Database Statistics

Image Source	# Images	# Faces
Randomly Downloaded	4,248,194	2,124,472
Celebrities	105,568	109,748
Person Names	19,492	12,806
Face-Related Words	13,212	14,424
Event-Related Words	1,429	1,335
Professions	115,808	79,992
Series	7,551	8,585
Camera Defaults	2,153	879
Miscellaneous	10,855	16,201
<b>Total</b>	<b>4,539,886</b>	<b>2,373,533</b>

Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>

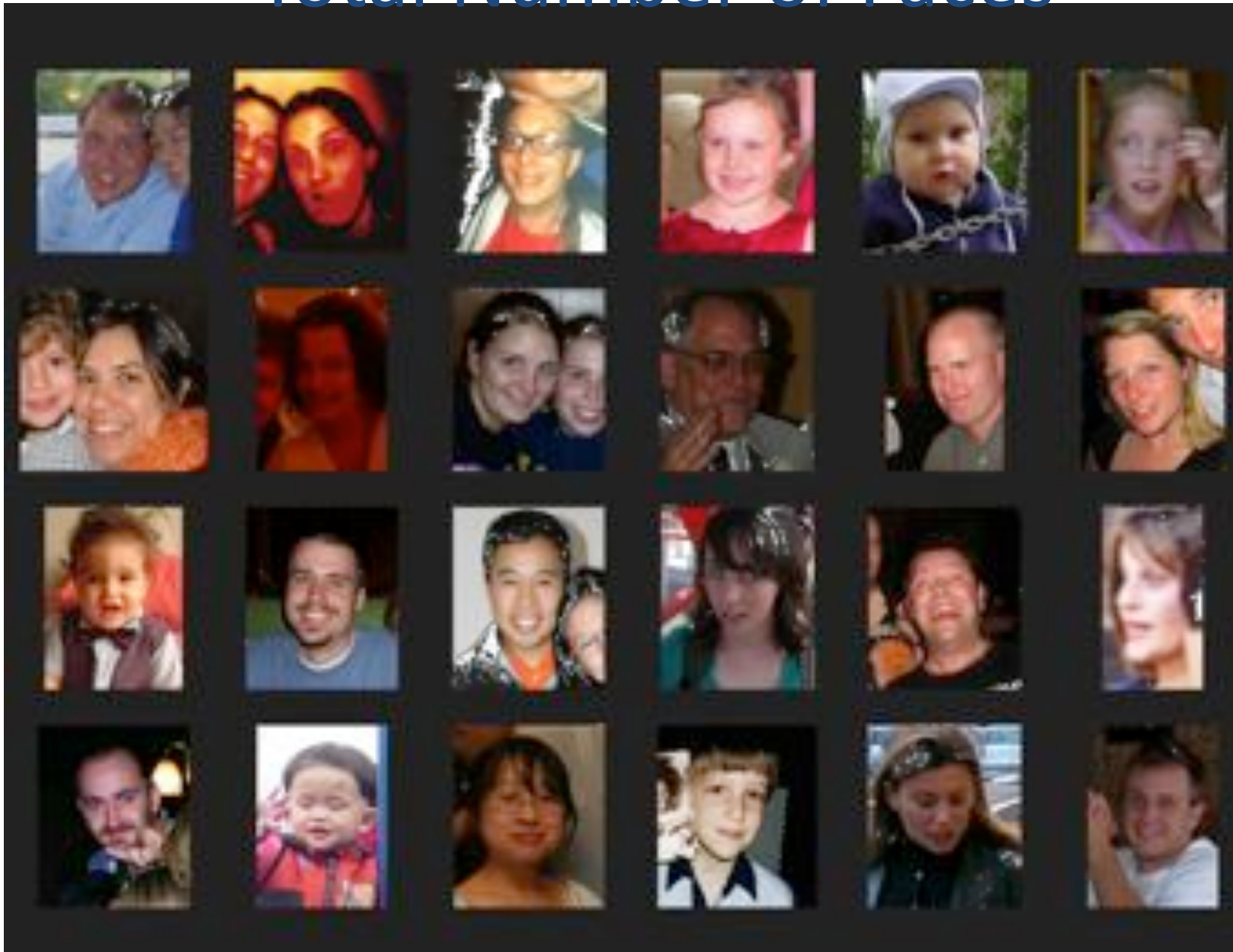
# Database Size Comparison

Database	# Face Images
MIT+CMU	130
Yale A	165
Yale B	5,760
FERET	14,051
CMU PIE	41,368
FRGC v2.0	50,000
<b>Proposed</b>	<b>2,373,533</b>

Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>



# Total Number of Faces



Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>

# Total Number of Faces



Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>



# Total Number of Faces



Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>

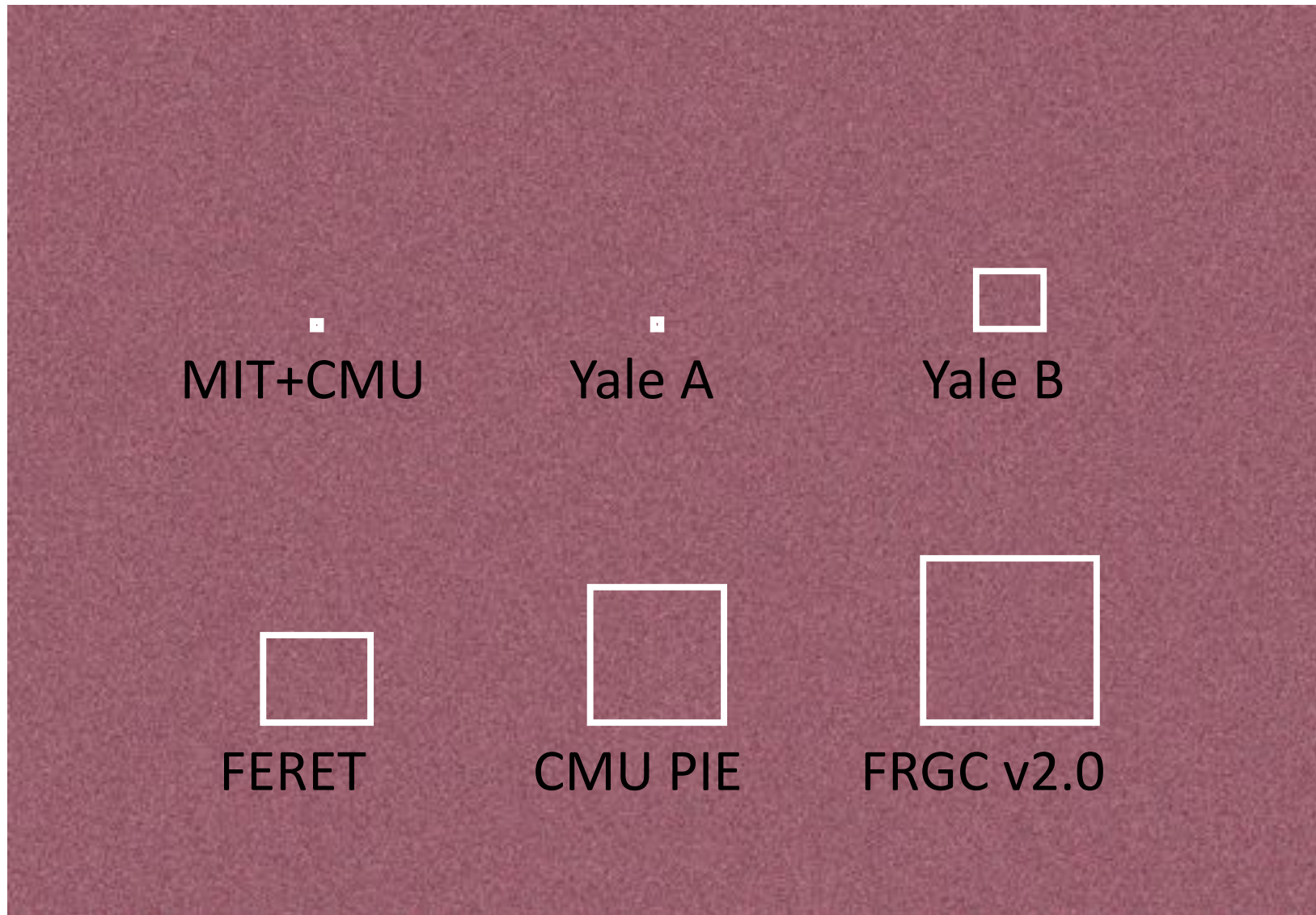
# Total Number of Faces



Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>

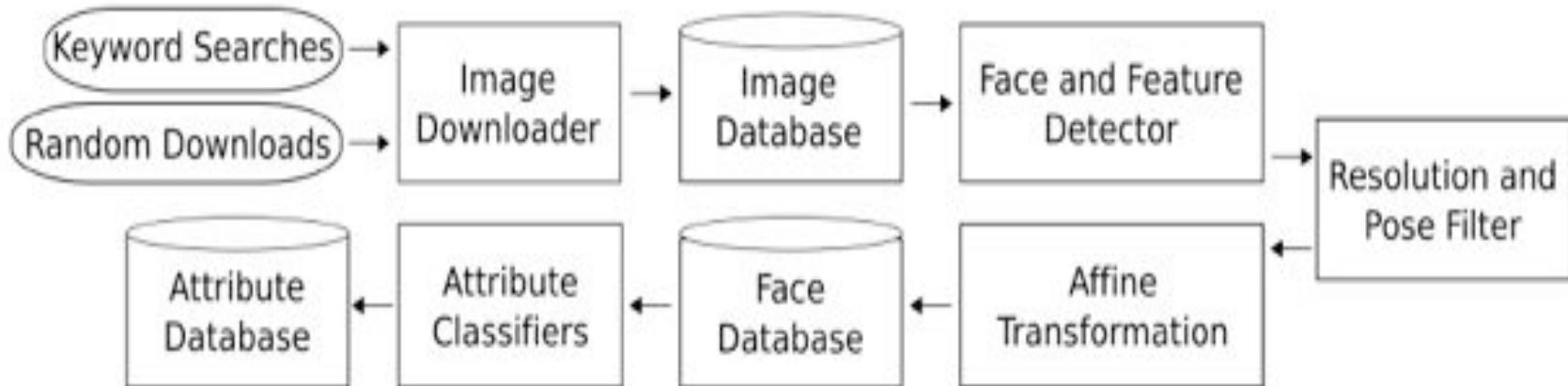


# Total Number of Faces



Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>

# Manual labeling of attributes...



So at this stage, we **have ~3.1 million images** (at the time of publication in 2008) and we need to train attribute classifier on them for 10 attributes

It is infeasible to manually label all the 3.1M images

**BUT**

we do need some labeled images for automatically labeling the remaining, so **we manually create ~17,000 attribute labeled images**

Image from: ECCV 2008 paper

# Labeled Attribute Statistics

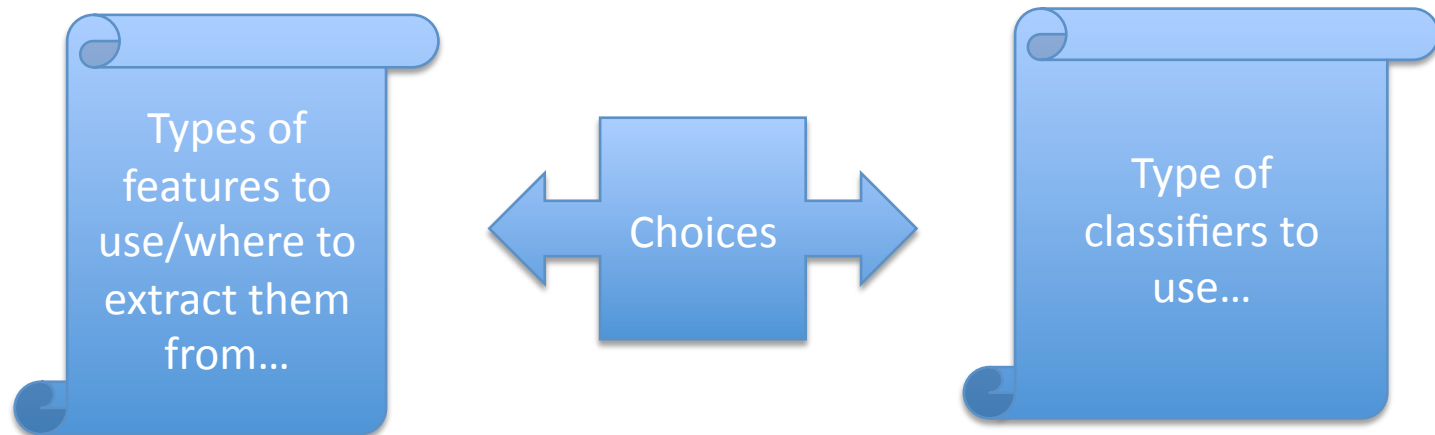
<b>Attribute</b>	<b># Labeled</b>	<b>Attribute</b>	<b># Labeled</b>
Gender	1,954	Mustache	1,947
Age	3,301	Smiling	1,571
Race	1,309	Blurry	1,763
Hair Color	1,033	Lighting	633
Eye Wear	2,360	Environment	1,583

**Total Number of Labels: 17,454**

Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/faceseach/#slides>

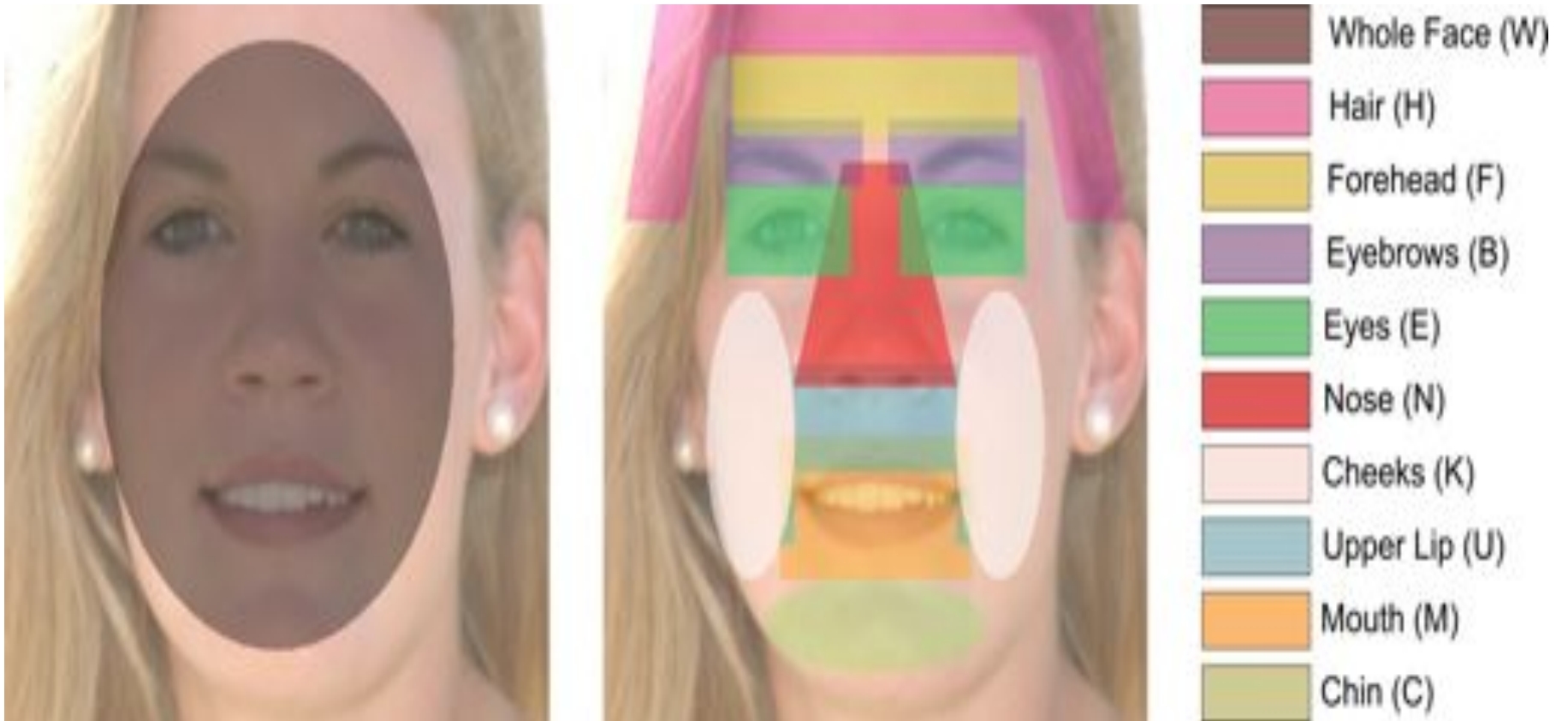
And this is where the fun starts...

- Goal: Given the 17k attribute labels we now need to train attribute classifiers for all 10 attributes to automatically label the remaining images...





# Where to extract features from?



Face divided into 10 functional regions...

Image from: ECCV 2008 paper

# Feature Types

Pixel Value Type	Normalizations	Aggregation
RGB (r)	None (n)	None (n)
HSV (h)	Mean-Norm (m)	Histogram (h)
Image Intensity (i)	Energy-Norm (e)	Statistics (s)
Edge Magnitude (m)		
Edge Orientation (o)		

Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>

# Feature Types

Pixel Value Type	Normalizations	Aggregation
RGB (r)	None (n)	None (n)
HSV (h)	Mean-Norm (m)	Histogram (h)
Image Intensity (i)	Energy-Norm (e)	Statistics (s)
Edge Magnitude (m)		
Edge Orientation (o)		

RGB, Mean Norm., No Aggreg. (r.m.n)

Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/faceseach/#slides>

# Feature Types

Pixel Value Type	Normalizations	Aggregation
RGB (r)	None (n)	None (n)
HSV (h)	Mean-Norm (m)	Histogram (h)
Image Intensity (i)	Energy-Norm (e)	Statistics (s)
Edge Magnitude (m)		
Edge Orientation (o)		

Edge Orientations, No Norm, Histogram (o.n.h)

Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/faceseach/#slides>



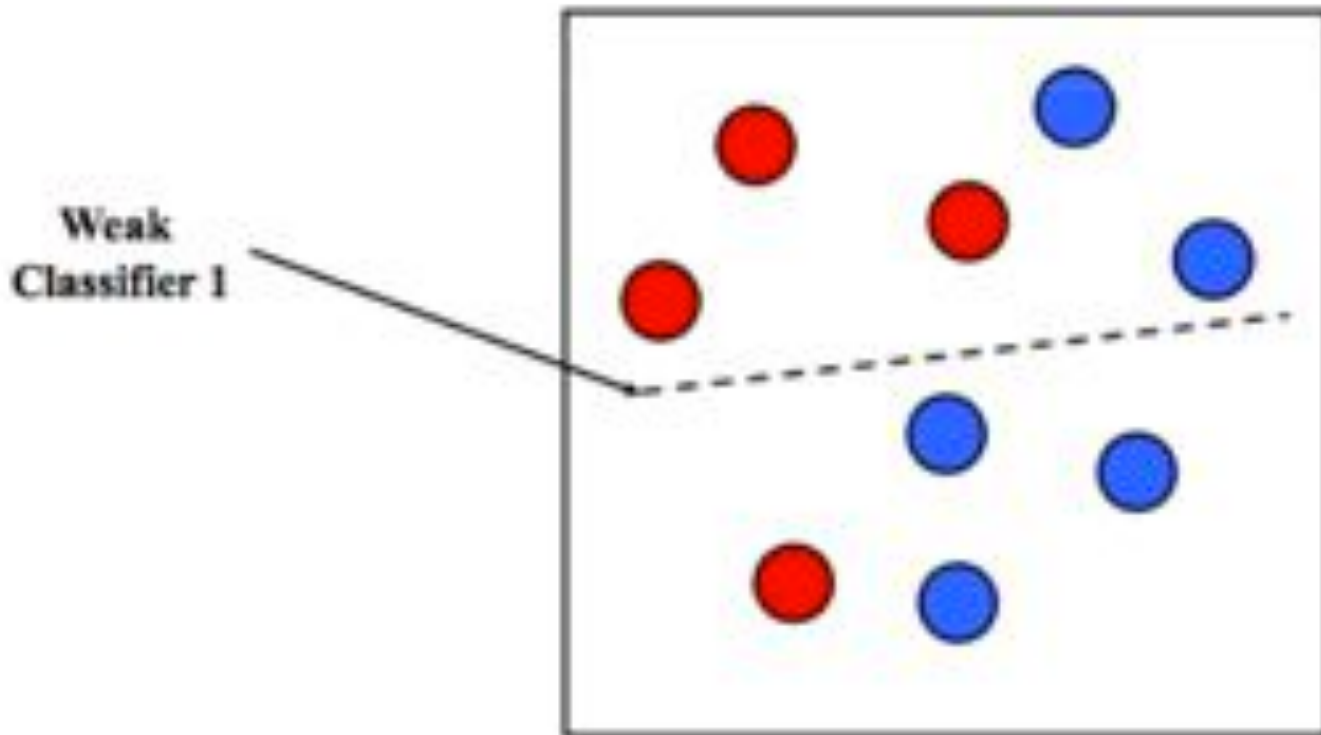
# Classifier architecture...

- Recent state of the art results in classification have mainly been achieved with SVMs
- The problem with SVMs is that **irrelevant features** can confuse/over-train the classifier...
- E.g. It might not make sense to use all facial pixels for training a classifier for just “is smiling”
- Given the large set of types of features/regions, we need a good way of selecting an optimal combination of features for each attribute...
- Enter Adaboost...

# Quick Review of Boosting...

## Boosting intuition

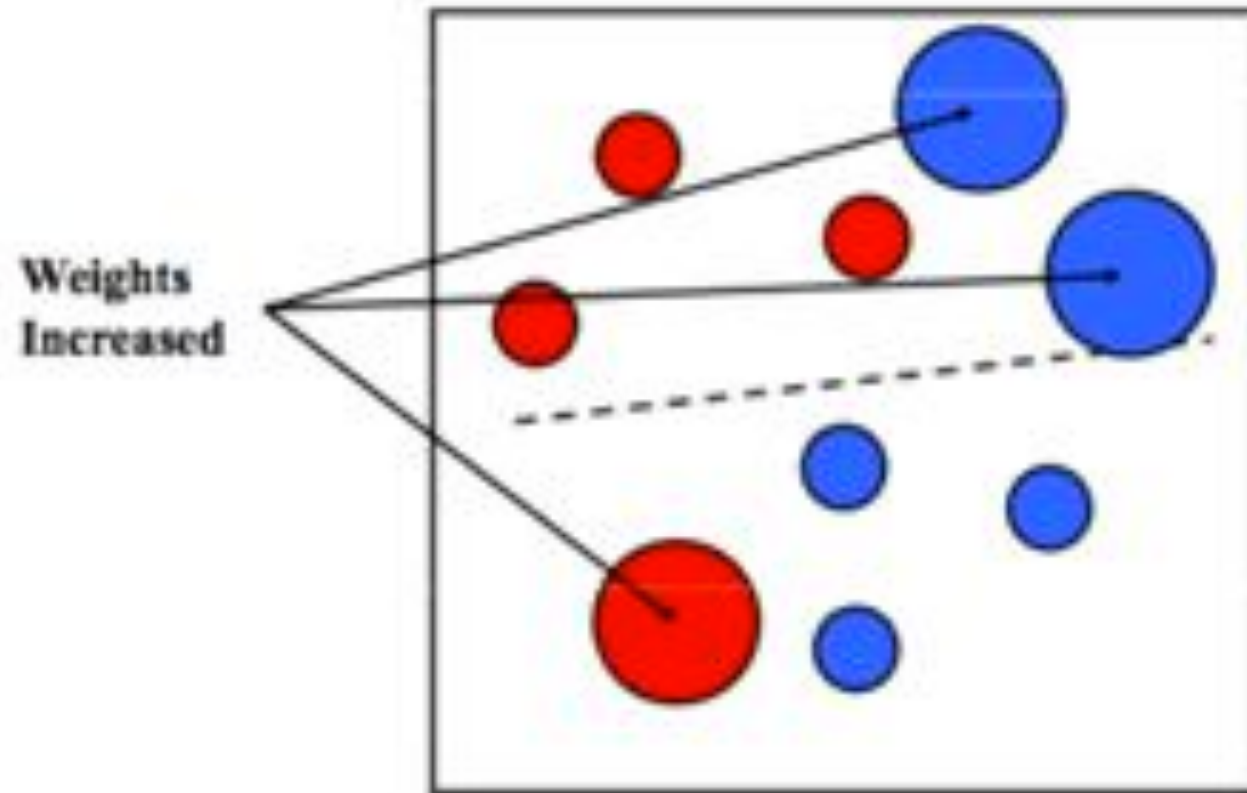
---



Source: [http://www.cs.utexas.edu/~cv-fall2012/slides/fall2012\\_04\\_categories\\_part1.pdf](http://www.cs.utexas.edu/~cv-fall2012/slides/fall2012_04_categories_part1.pdf)

# Boosting illustration

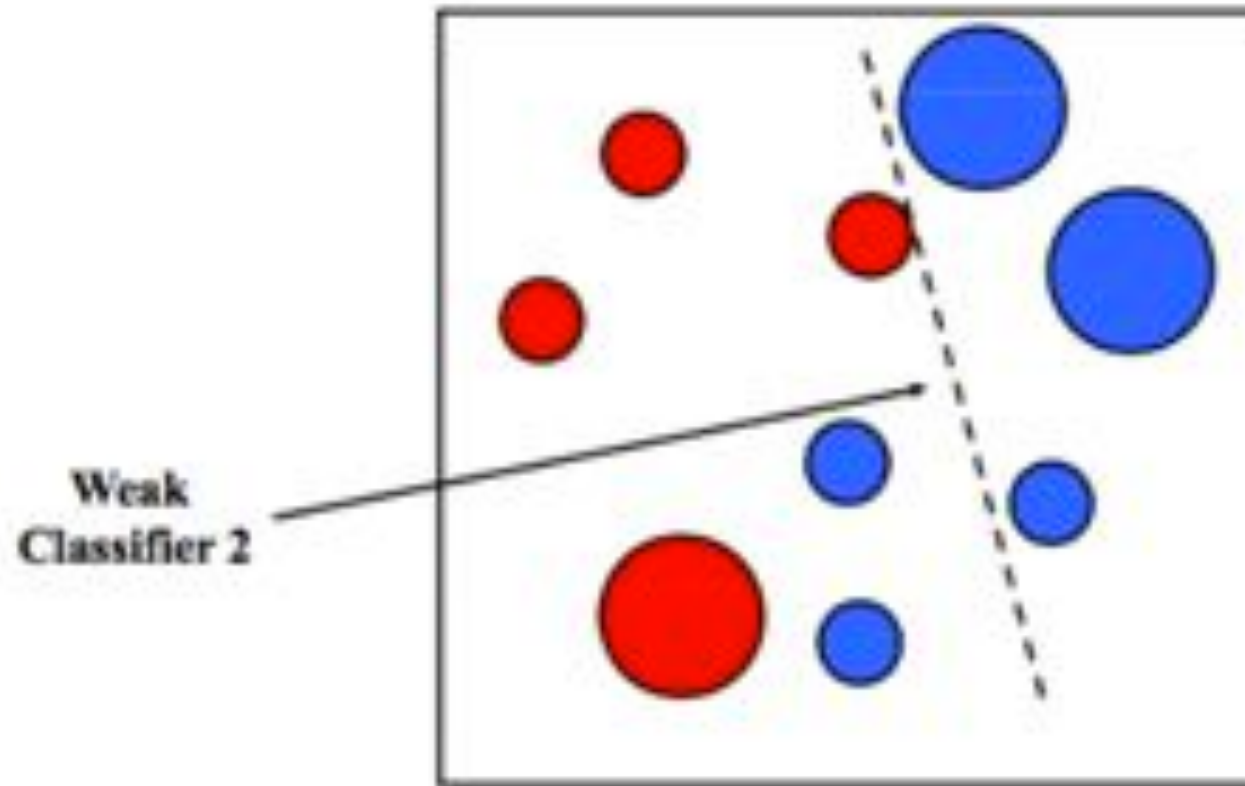
---



Source: [http://www.cs.utexas.edu/~cv-fall2012/slides/fall2012\\_04\\_categories\\_part1.pdf](http://www.cs.utexas.edu/~cv-fall2012/slides/fall2012_04_categories_part1.pdf)

# Boosting illustration

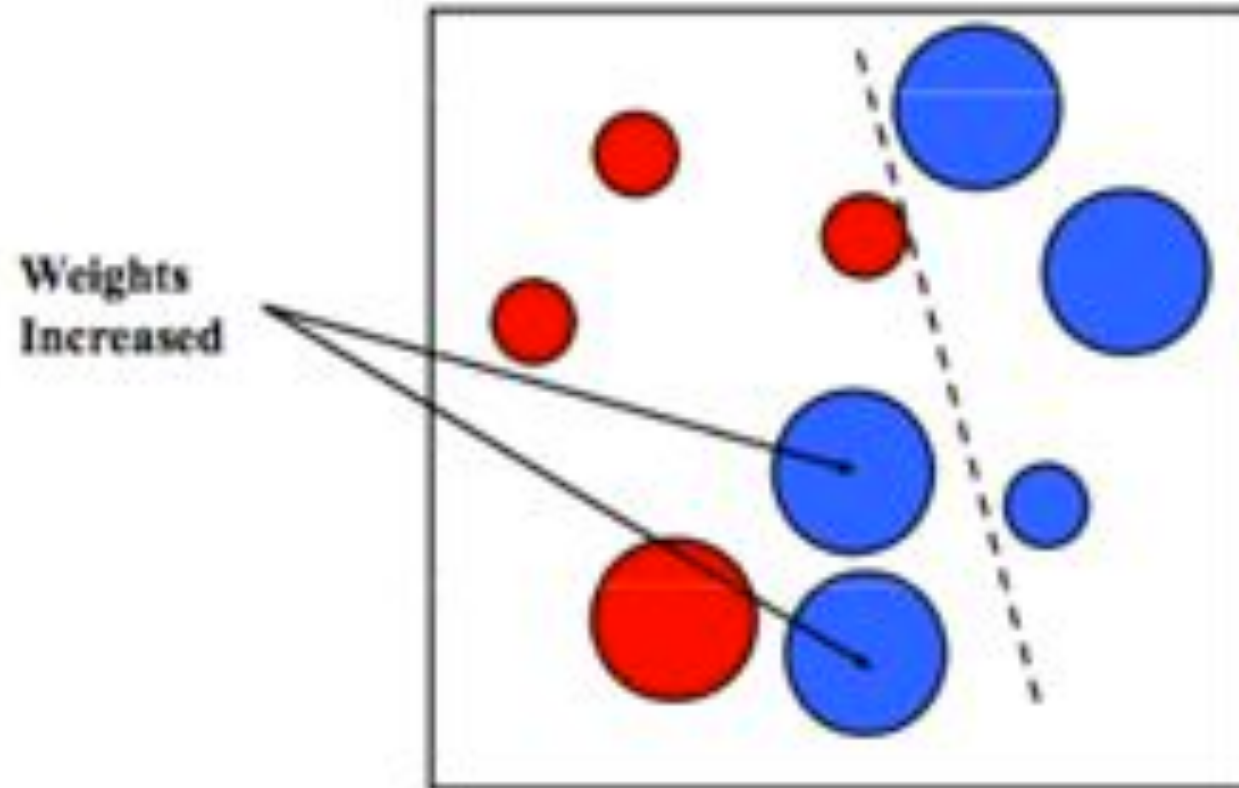
---



Source: [http://www.cs.utexas.edu/~cv-fall2012/slides/fall2012\\_04\\_categories\\_part1.pdf](http://www.cs.utexas.edu/~cv-fall2012/slides/fall2012_04_categories_part1.pdf)

# Boosting illustration

---

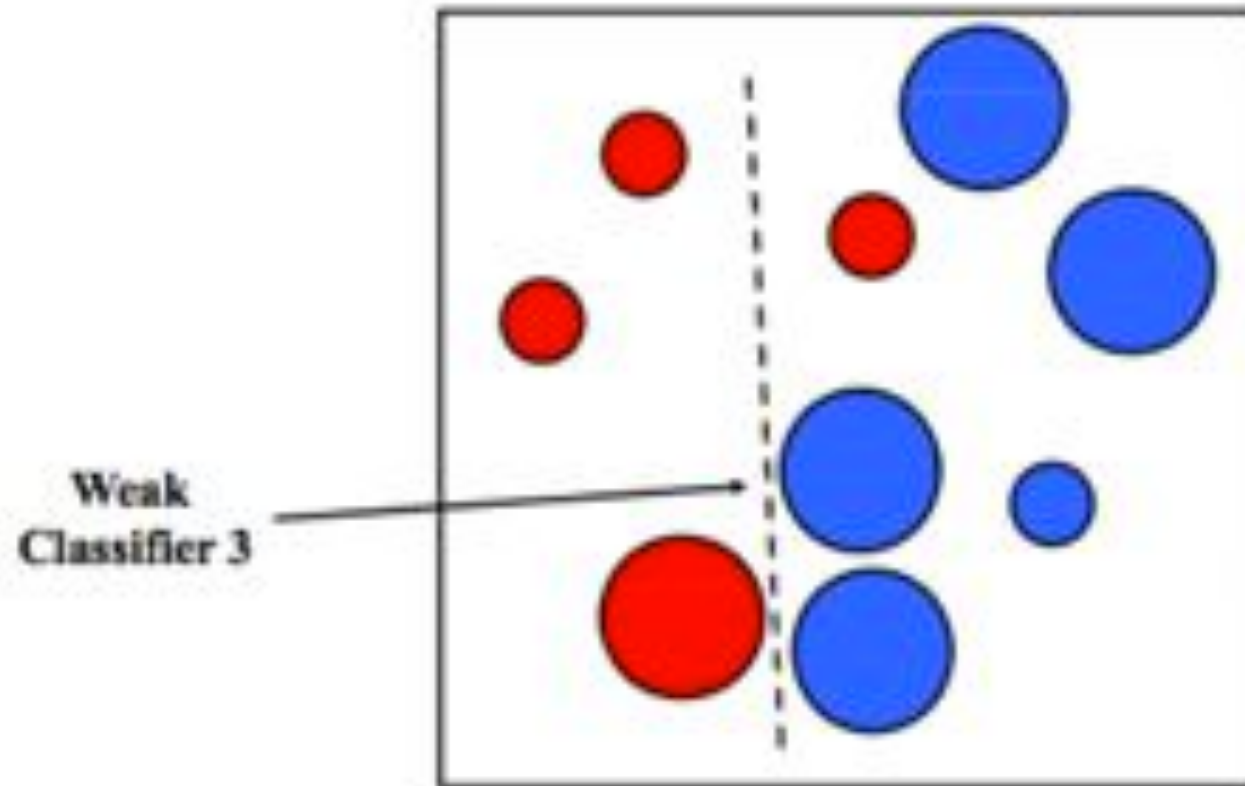


Source: [http://www.cs.utexas.edu/~cv-fall2012/slides/fall2012\\_04\\_categories\\_part1.pdf](http://www.cs.utexas.edu/~cv-fall2012/slides/fall2012_04_categories_part1.pdf)



# Boosting illustration

---

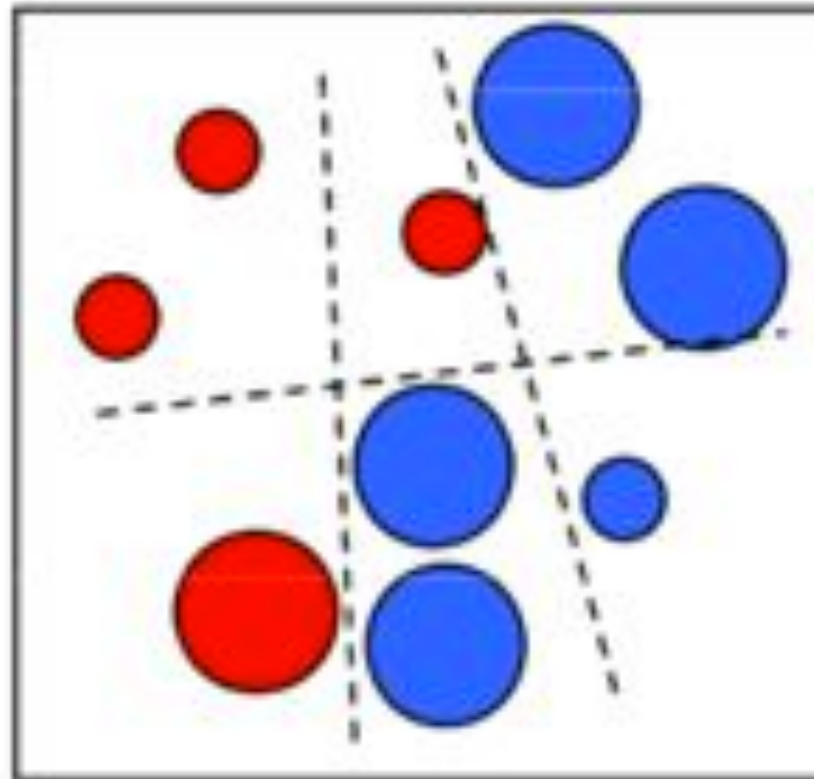


Source: [http://www.cs.utexas.edu/~cv-fall2012/slides/fall2012\\_04\\_categories\\_part1.pdf](http://www.cs.utexas.edu/~cv-fall2012/slides/fall2012_04_categories_part1.pdf)

# Boosting illustration

---

**Final classifier is  
a combination of weak  
classifiers**



# Combining Boosting with SVMs...

- The idea is to **construct a “local” SVM for every possible combination** of region, feature types and SVM parameters (LibSVM)
- And then to **use Adaboost** to create an optimal classifier using a **linear combination of these local SVMs**
- The usual Adaboost algorithm is modified so that no retraining is needed at the beginning of each round (since these SVMs are either powerful/useless classifiers depending on the relevance of the features used )

## Discussion...

- Boosting is meant to turn weak learners into strong learners. Does using boosting in this scenario where you have pre-trained SVMs make sense? Wouldn't using some feature selection approach be better?
- Performance degradation in boosting, (Wickramaratna, J. and Holden, S. and Buxton, B., Multiple Classifier Systems, **2001**) shows that boosting strong learners can cause performance degradation

## Discussion...

- While in this paper, they assumed that since SVMs were either powerful/useless learners, the normal retraining step in Adaboost wasn't needed, they have a related follow-up work (Attribute and Simile Classifiers for Face Verification, N. Kumar, A. Berg, P. Belhumeur, S. Nayar. ICCV 2009) where they use forward feature selection instead of Adaboost.
- While they don't get much better results, their system isn't restricted to only frontal poses.

# Train Classifiers



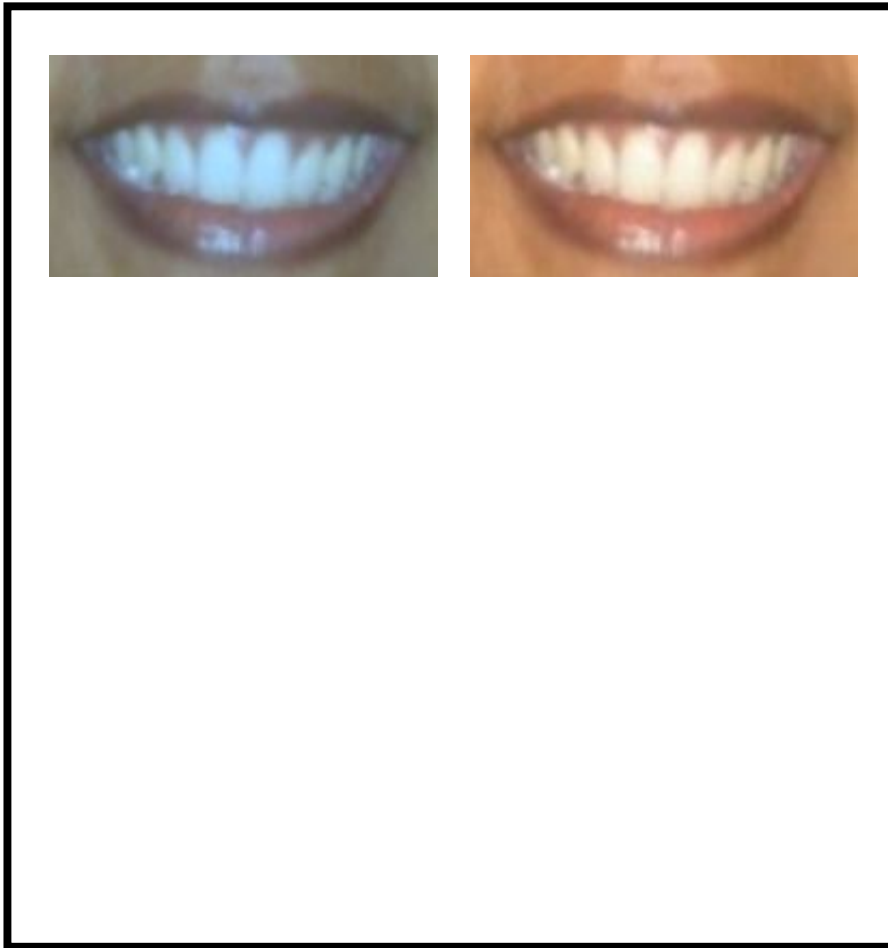
Mouth  
Raw RGB

Pool of Classifiers

Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/faceseach/#slides>



# Train Classifiers

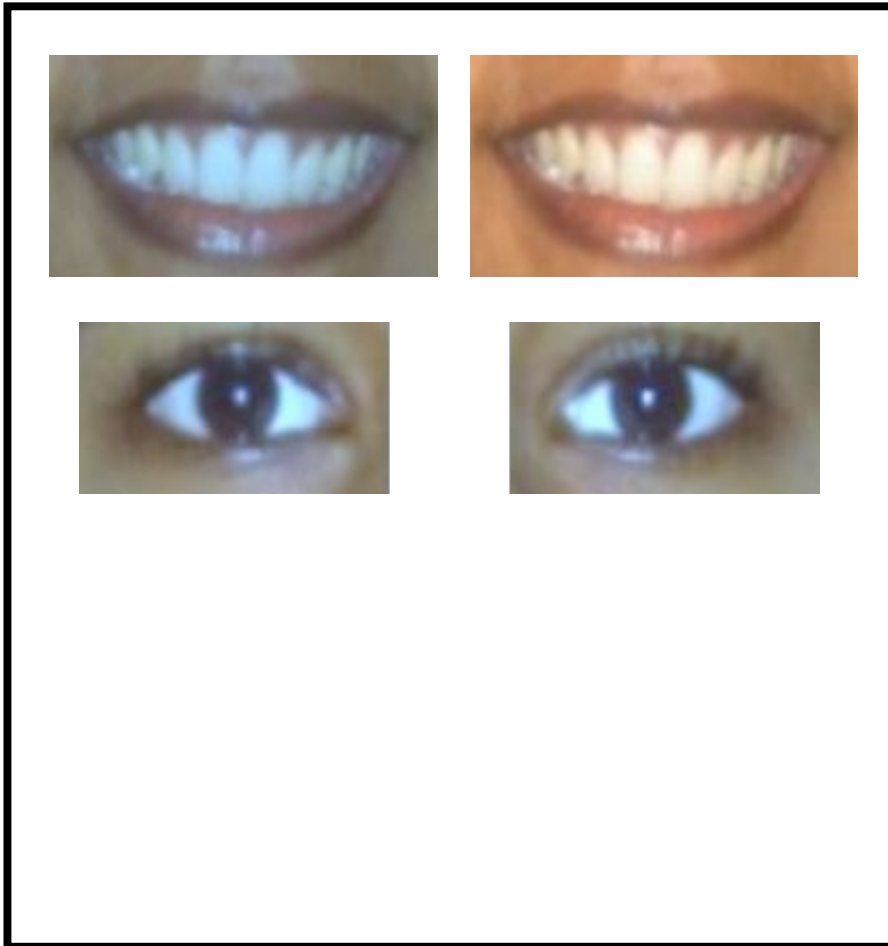


Eyes  
Mean-Normalized RGB

## Pool of Classifiers

Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/faceseach/#slides>

# Train Classifiers

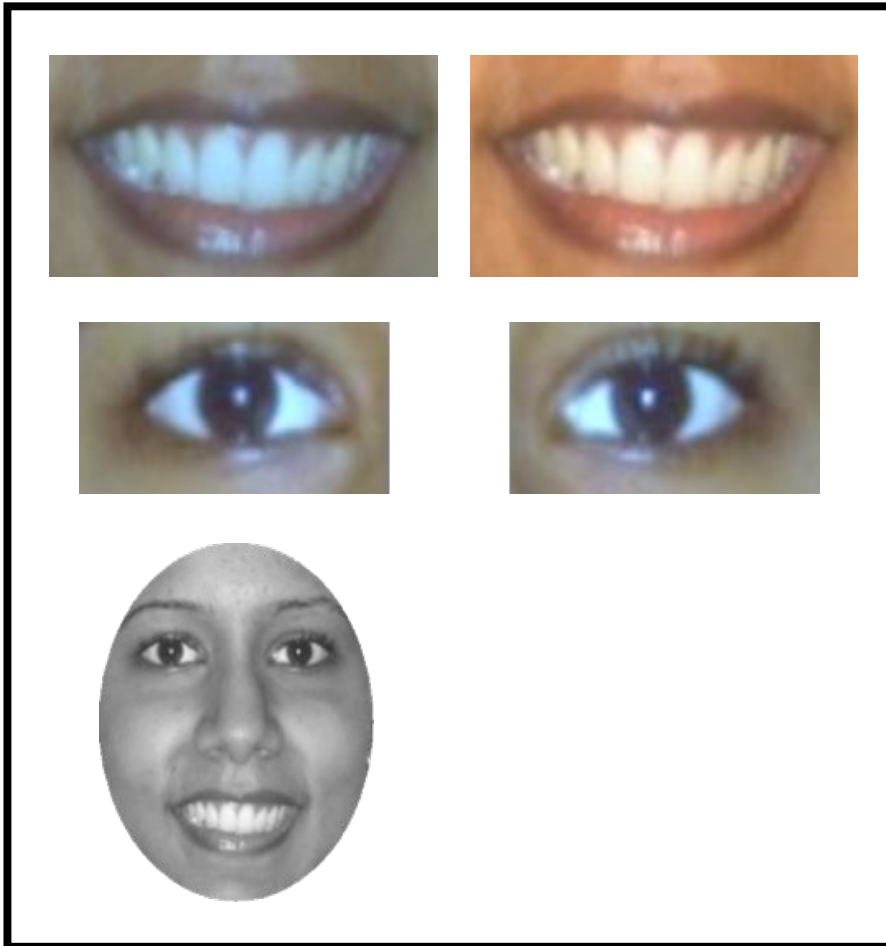


Whole Face  
Raw Intensity

## Pool of Classifiers

Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/faceseach/#slides>

# Train Classifiers



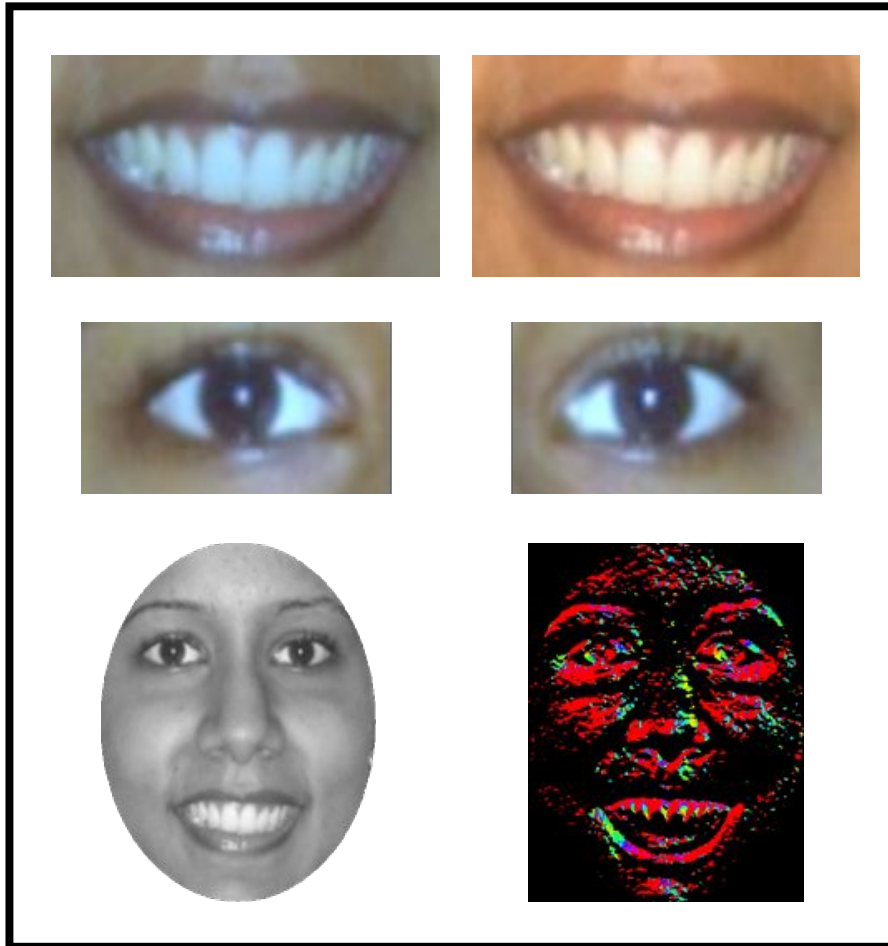
Pool of Classifiers



Whole Face  
Gradient Directions

Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/faceseach/#slides>

# Select Classifiers

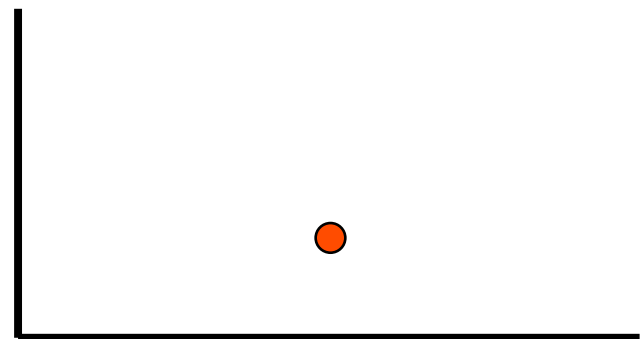


Pool of Classifiers



Selected Classifiers

Error Rate



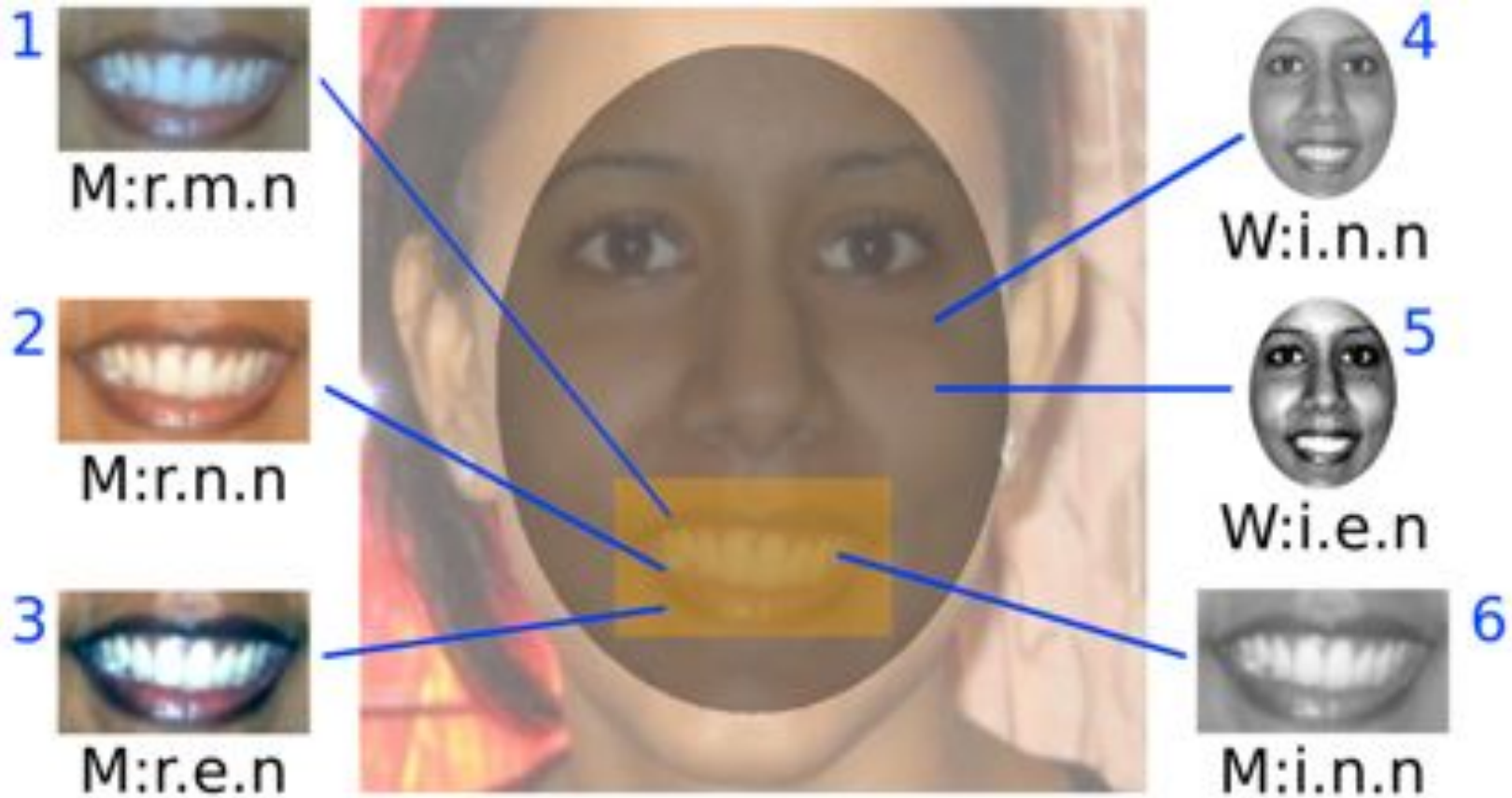
Iteration

# Feature Selection: Smiling

1. Mouth: RGB, Mean Norm., No Aggreg. (M:r.m.n)
2. Mouth: RGB, No Norm., No Aggreg. (M:r.n.n)
3. Mouth: RGB, Energy Norm., No Aggreg. (M:r.e.n)
4. Whole Face: Intensity, No Norm., No Aggreg. (W:i.n.n)
5. ...



# Selected Features

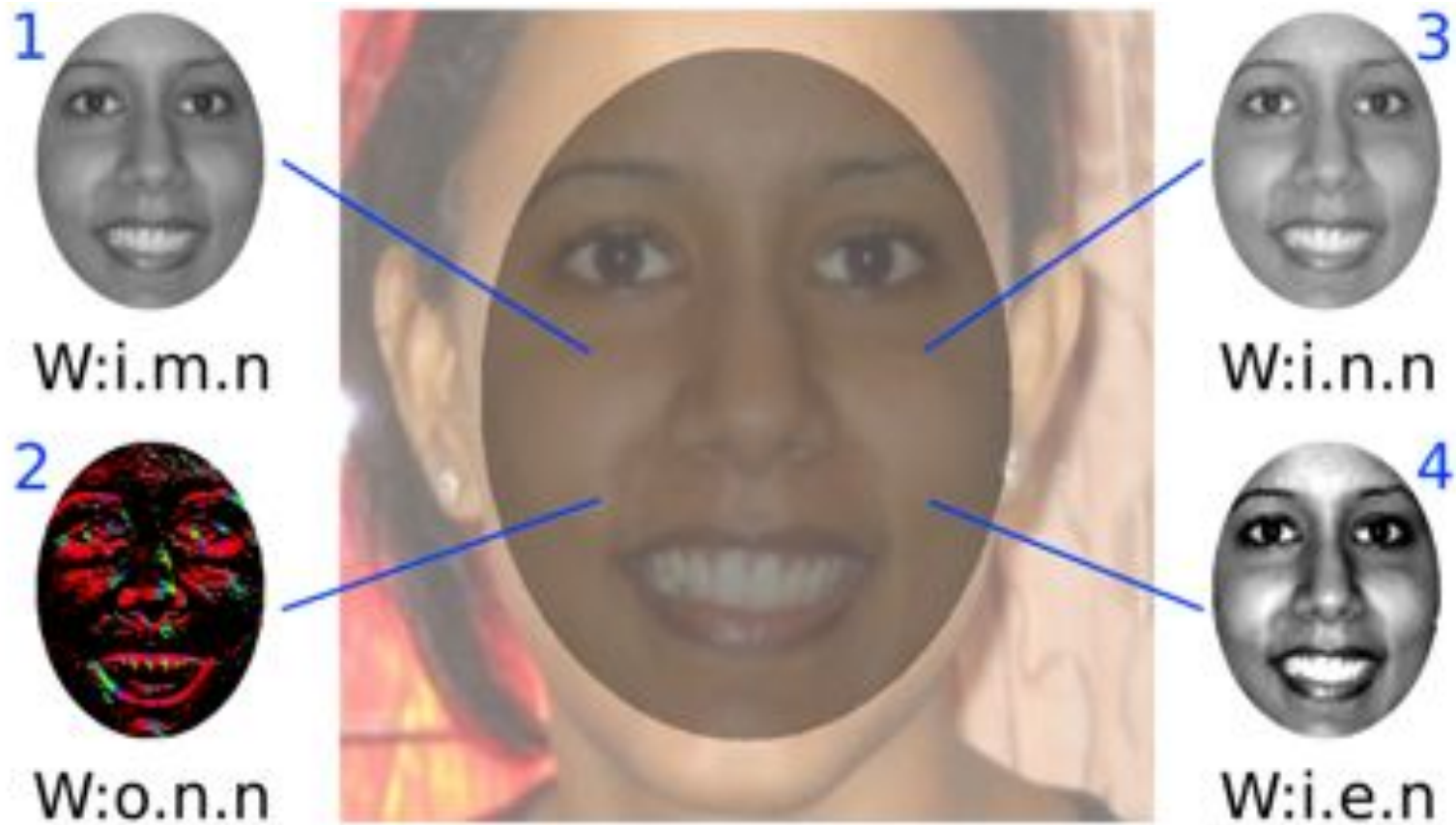


## Smiling

Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>



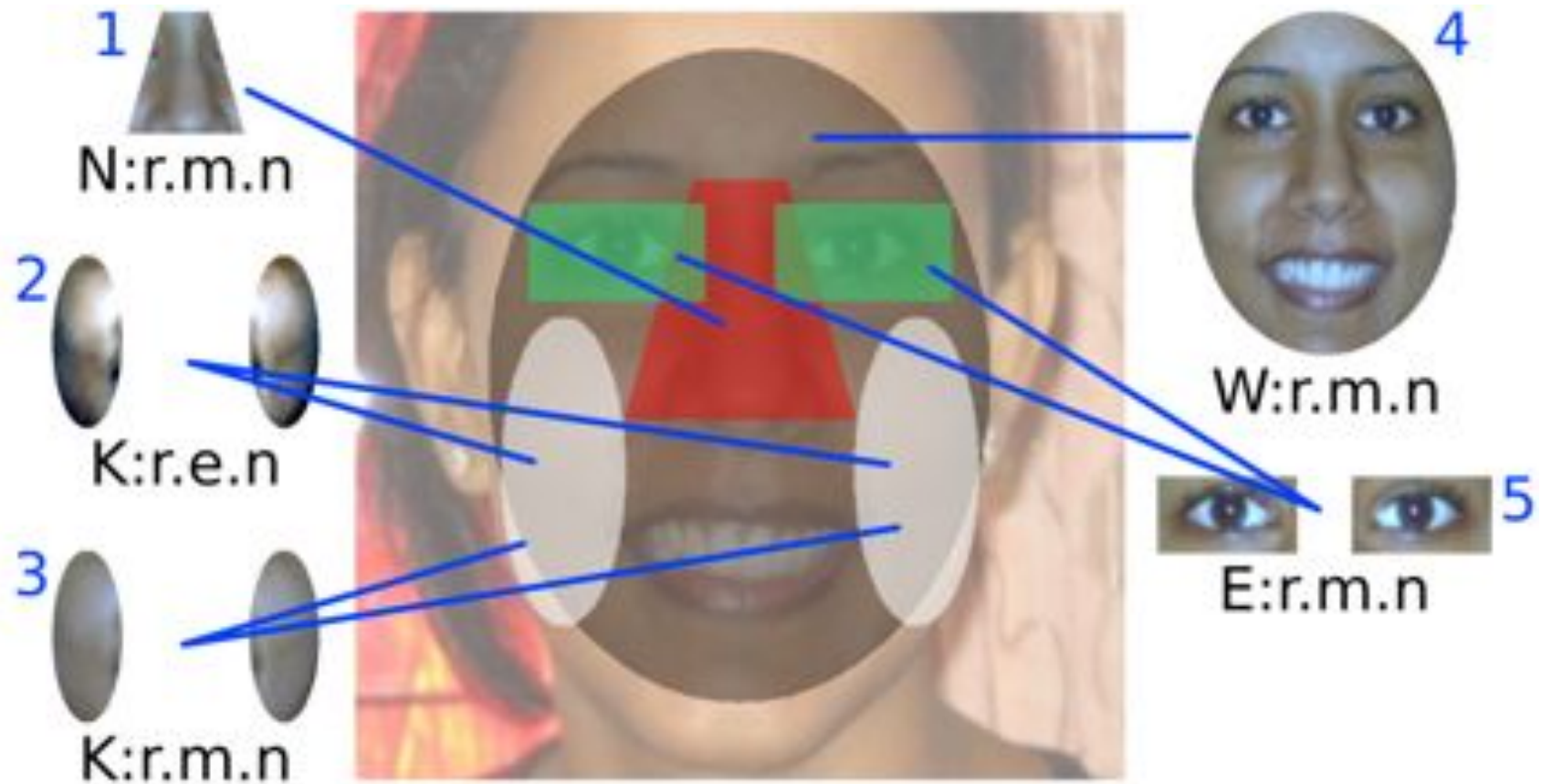
# Selected Features



## Gender

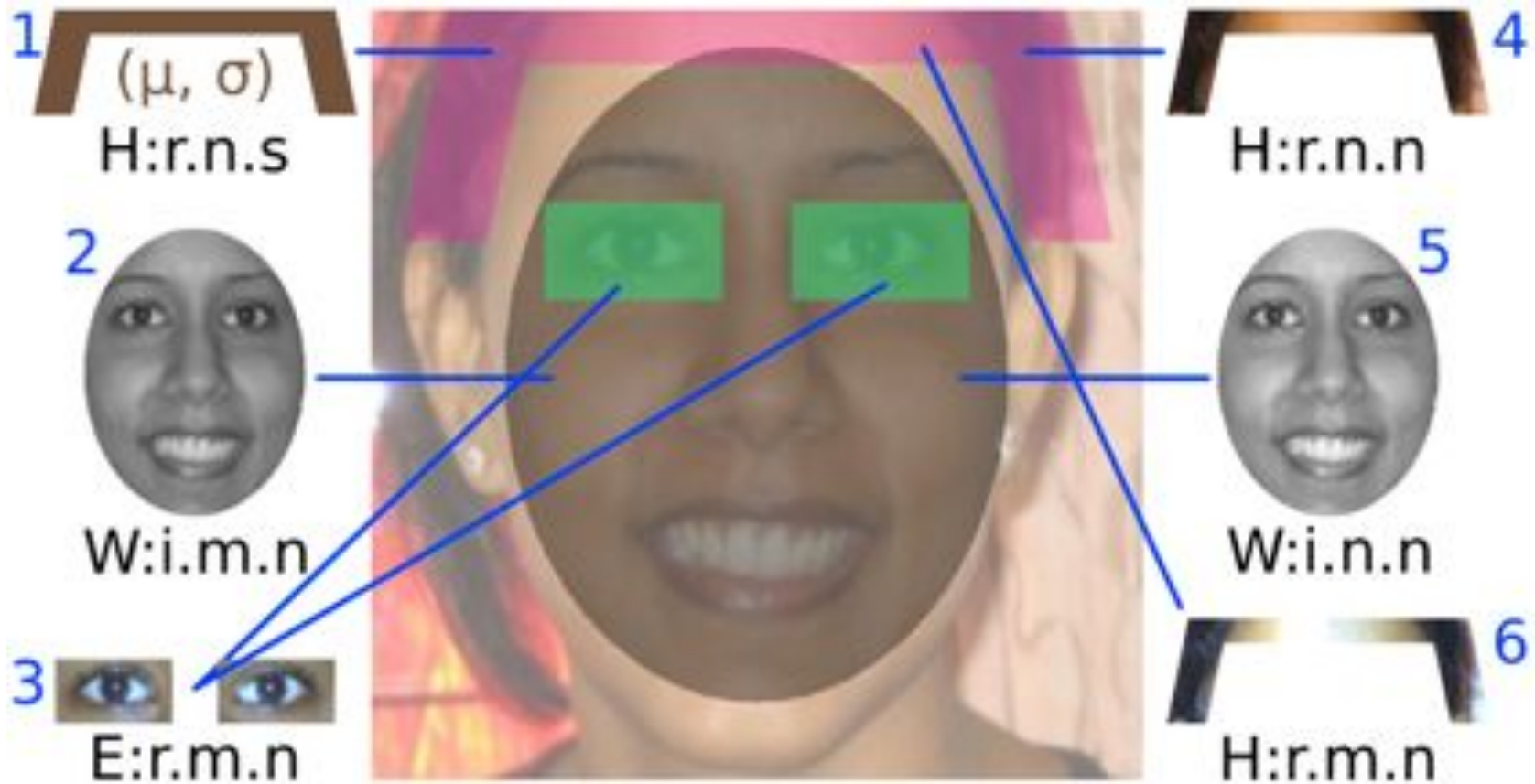
Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/faceseach/#slides>

# Selected Features



## Indoor/Outdoor

# Selected Features



## Hair Color

# One global SVM to rule them all...

- The **drawback** of such a local SVM based architecture is that it requires **keeping a large number of SVMs in memory and evaluating all** of them for every new input image...
- This is solved by training **one global SVM** on the **union** of the features from the top N highest weighted (by Adaboost) SVMs
- The next slide shows the comparative error rates for the local v/s global SVM approach...

# Discussion...

- While concatenating the local feature sets (for creating the global SVM from the local top ranked SVMs) they do not seem to use the Adaboost weights/scores for those feature-region sets. Could this help?



# Classification Accuracy

Attribute	Error Rates for Attribute-Tuned Local SVMs	Error Rates for Attribute-Tuned Global SVM	Top Feature Combinations in Ranked Order Each combination is represented as Region:pixtype.norm.aggreg
Gender	9.42%	8.62%	W:i.m.n   W:o.n.n   W:i.n.n   W:i.e.n
Age	17.34%	16.65%	W:i.m.n   W:i.n.n   H:r.e.n   E:r.m.n   H:r.e.s   W:o.n.n
Race	7.75%	6.49%	W:i.m.n   E:r.e.n   C:o.n.n   M:r.m.n   W:o.n.n
Hair Color	7.85%	5.54%	H:r.n.s   W:i.m.n   E:r.m.n   H:r.n.n   W:i.n.n   H:r.m.n
Eye Wear	6.22%	5.14%	W:m.n.n   W:i.n.n   K:o.n.h   W:m.m.n   N:r.n.n
Mustache	6.42%	4.61%	U:r.e.n   M:r.m.n
Smiling	4.60%	4.60%	M:r.m.n   M:r.n.n   M:r.e.n   W:i.n.n   W:i.e.n   M:i.n.n
Blurry	3.94%	3.41%	W:m.m.n   H:m.n.n   W:m.n.n   H:m.m.n   M:m.m.n
Lighting	2.82%	1.61%	W:i.n.n   W:i.e.n   K:r.n.n   C:o.n.n   E:o.n.n
Environment	12.25%	12.15%	N:r.m.n   K:r.e.n   K:r.m.n   W:r.m.n   E:r.m.n

Source of slide: The corresponding ECCV 2008 paper

# Comparison to State-of-the-Art

Method	Gender Error Rate	Smiling Error Rate
Proposed	8.62%	4.60%
Adaboost (pixel comparison feats) Baluja & Rowley, IJCV 2007	13.13%	7.41%
Adaboost (Haar-like features) Shakhnarovich et al., ICAFGR 2002	12.88%	6.40%
Full face SVM Moghaddam & Yang, TPAMI 2002	9.52%	13.54%

Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>

# Discussion...

- This approach seems to boil down to feature selection. Could their improved results be because of their rich set of features rather than the Boosting+SVMs approach?

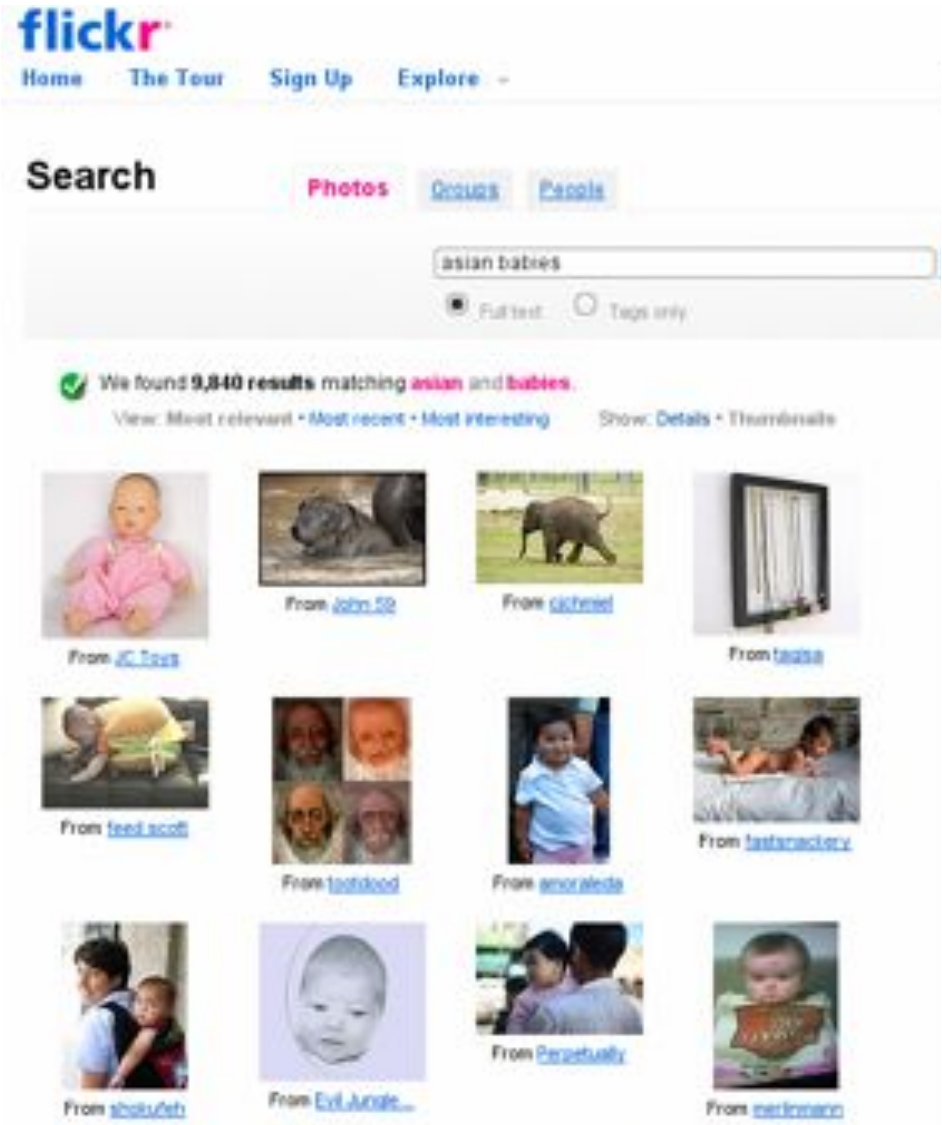
# FaceTracer Engine...

- Step1: [Offline] Training attribute-tuned global SVM classifiers for each of the 10 attributes
- Step2: [Offline] All the images in the database are sent through these classifiers to be labeled
- These labels are stored for fast **online** search.
- The search interface accepts simple text based queries and maps these onto the labels by using a dictionary of terms.
- It returns the results in the order of decreasing confidence.
- Applications: law enforcements, social networking sites, personal snap collection management

## Discussion...

- They mention that for multiple-attribute query terms, they convert the classifier confidences into probabilities and then use the product of these probabilities for scoring/ranking. Is this the right approach? Don't these different attribute classification scores need to be calibrated properly?
- Multi-Attribute Spaces: Calibration for Attribute Fusion and Similarity Search. W. Scheirer, N. Kumar, P. Belhumeur, T. Boult. CVPR 2012

# “Asian Babies”



Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>



# “Adults Outside”



Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>

# “Middle-Aged White Men”

phizard  
face search

middle-aged white men


Search Faces

Example searches: women with no glasses, non adults

Show Attributes

Parsed query as: age is middle aged, race is white, race is not asian, gender is m  
Found 1000 results in 1.053 secs. Displaying results

Aligned Faces Images



(c) 2008 Automatic Face Systems, Inc., New York, NY

Web Images Maps News Shopping Gmail more

Google

middle-aged white men

Search Images

Search the Web

Advanced Image Search Preferences

Moderate SafeSearch is on

New! Google Image Labels

Images Showing: All image sizes

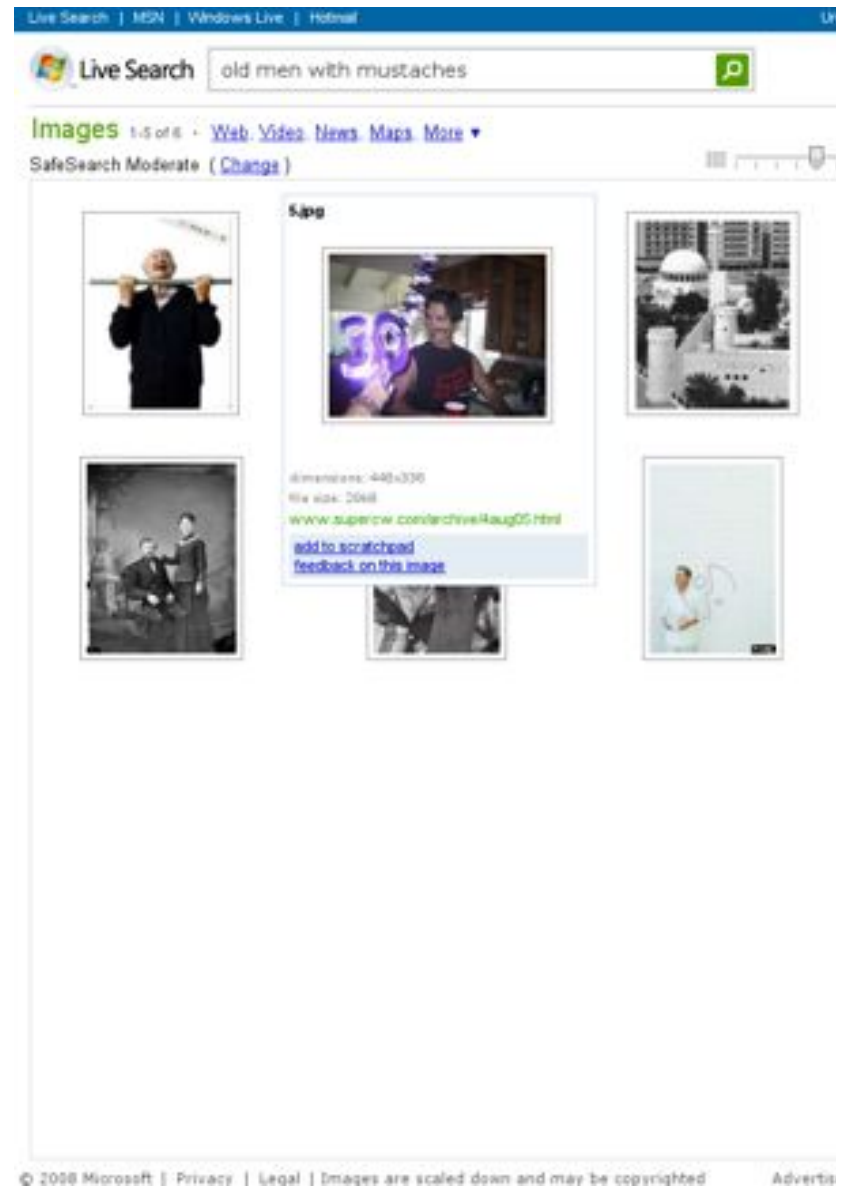
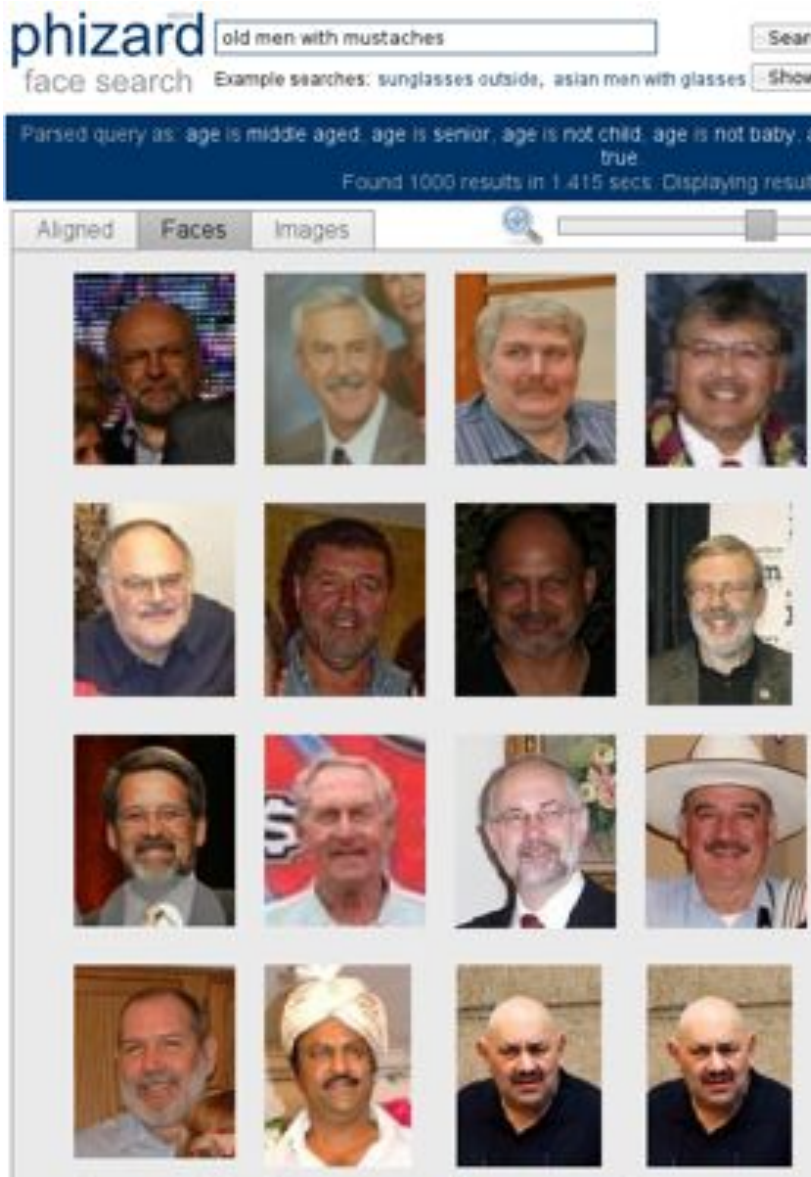
Results 1 - 20 of about 169,000 for middle-aged white men (0.04 seconds)



Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>

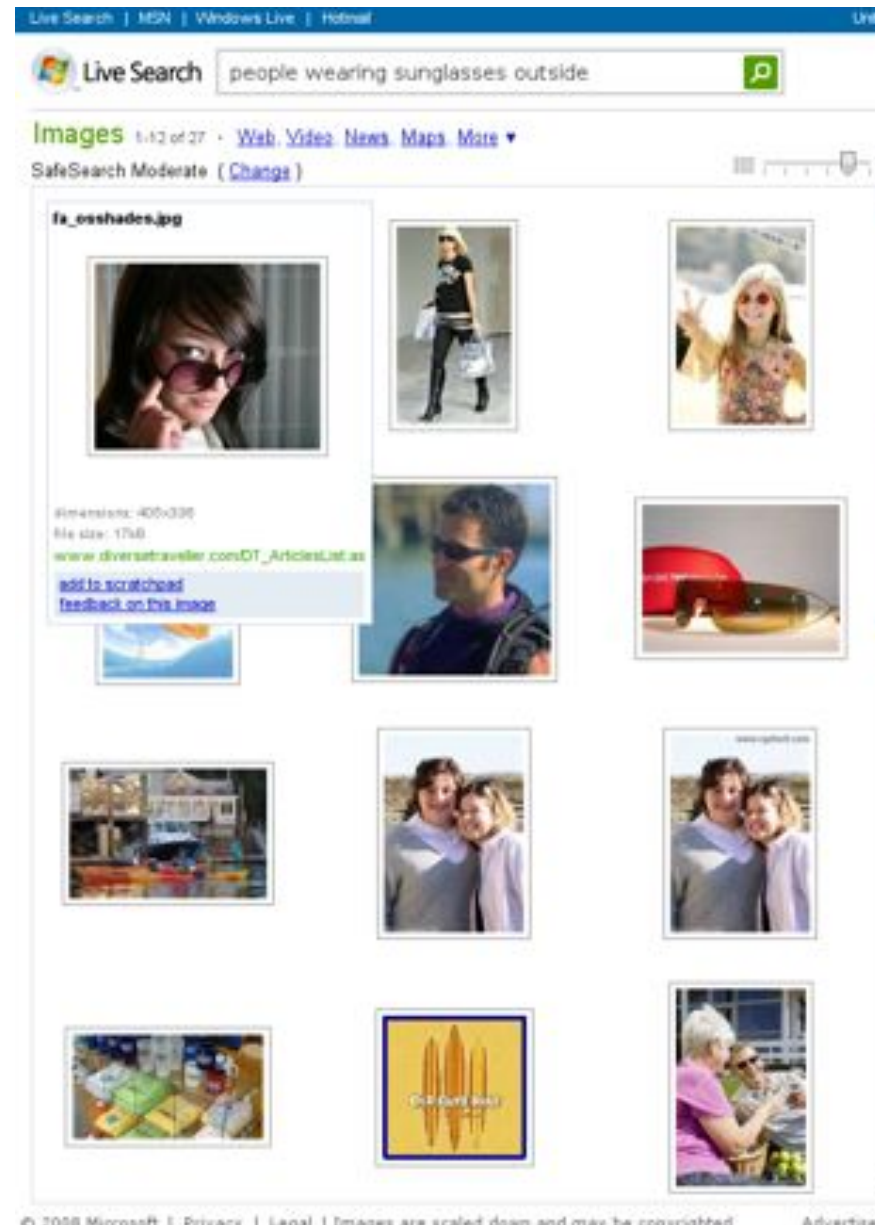


# “Old Men With Mustaches”



Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>

# “People Wearing Sunglasses Outside”



Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>

# “Kids Indoors Not Smiling”

phizard  
face search



kids indoors not smiling

Search Faces Find Similar

Example searches: asian men with glasses, men with dark hair Show Attribute List

Parsed query as: age is child, age is not middle aged, environment is indoor, expression is not smiling  
Found 1000 results in 0.521 secs. Displaying results 1 to 48

Aligned Faces Images



[Open Original Image](#)  
[Open Source Page](#)

phizard  
face search

kids indoors not smiling

Search

Example searches: asian men with glasses, men with dark hair Show Attribute List

Parsed query as: age is child, age is not middle aged, environment is indoor, expression is not smiling  
Found 1000 results in 0.521 secs. Displaying results 1 to 48

Aligned Faces Images



flickr

Home The Tour Sign Up Explore

Search

Photos

Groups

People

kids indoors not smiling

SEARCH

[Advanced Search](#)  
[Search by Camera](#)

Full text  Tags only

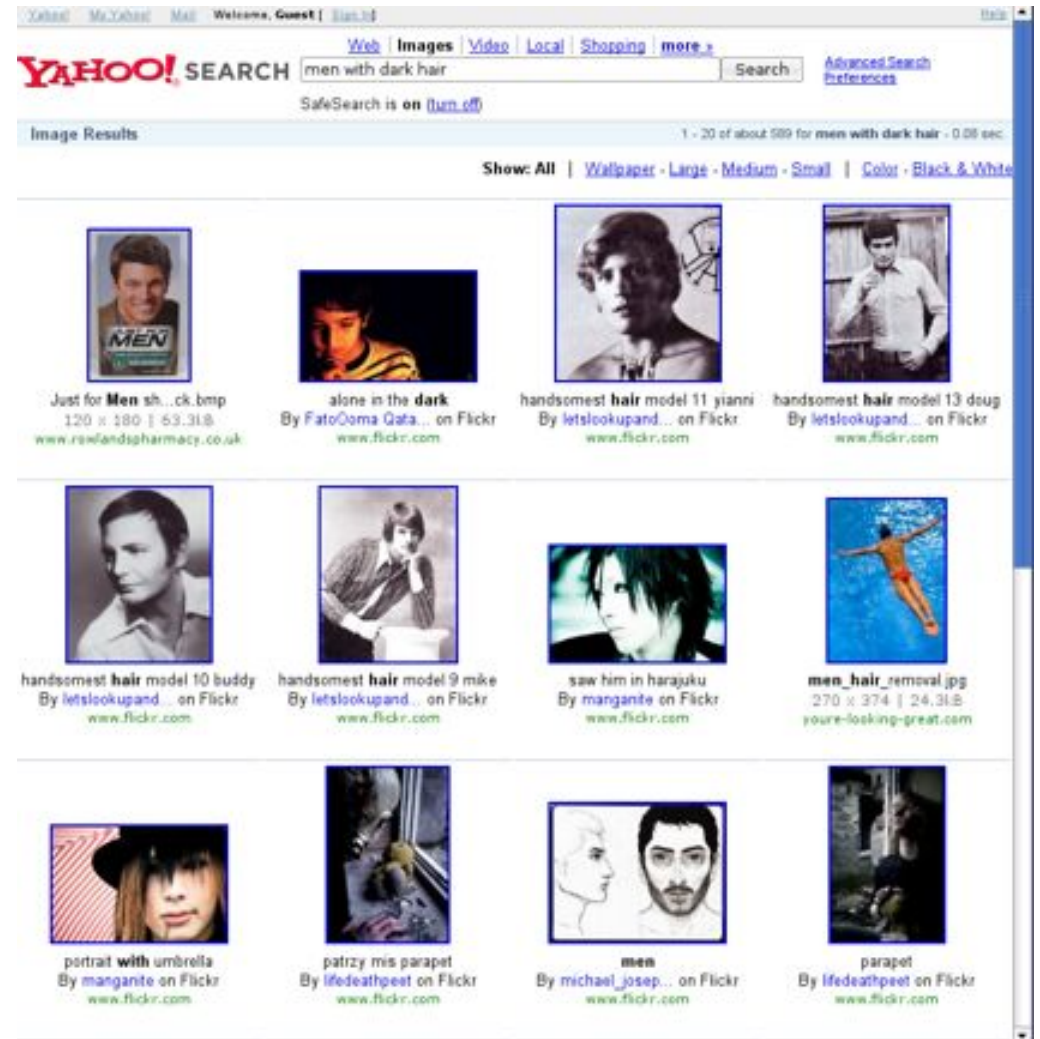
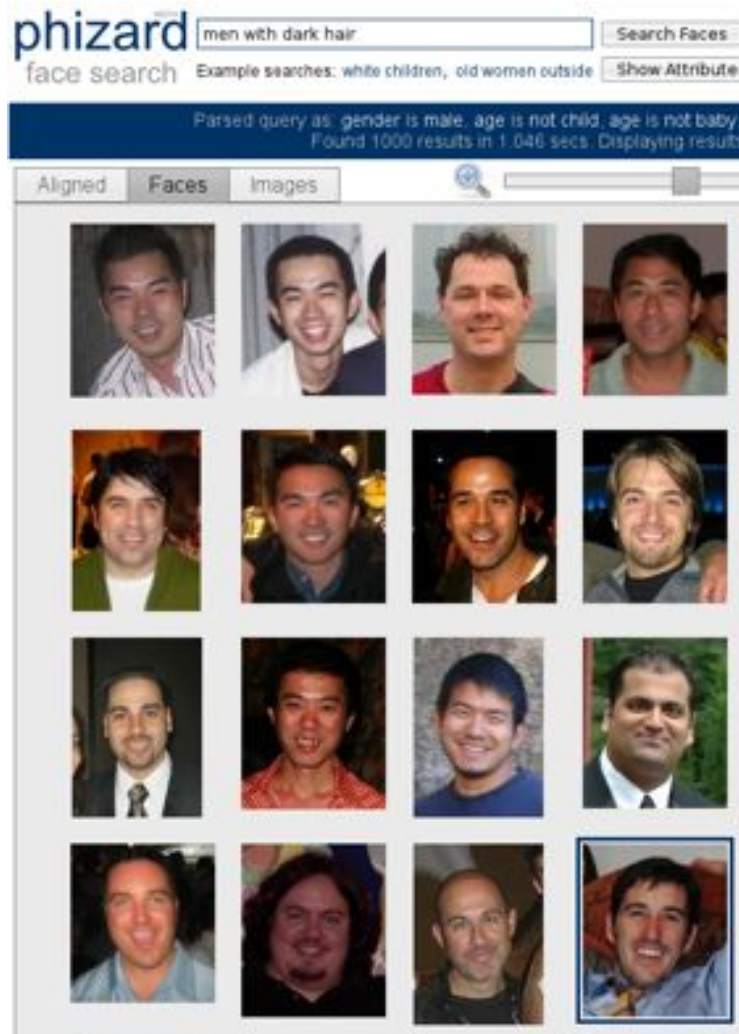
**!** We couldn't find any results matching "kids indoors" and not smiling.

Would you like to try a search for [smile](#), [happy girl portrait](#) or [woman](#) instead?

Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>



# “Men With Dark Hair”



Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>





# “Smiling Asian Men With Glasses”

phizard

face search Example searches: men with dark hair, flash photos of kids

Parsed query as: expression is smiling, race is asian, gender is male, age is not child, age is not baby, eye wear is eyeglasses.  
Found 1000 results in 0.017 secs. Displaying results 1 to 48

Aligned Faces Images




[Open Original Image](#)  
[Open Source Page](#)

Yahoo! SEARCH   [Advanced Search](#) [Preferences](#)


SafeSearch is on [Turn off](#)

Image Results 1 - 3 of about 3 for smiling asian men with glasses - 0.05 sec


Show: All | Wallpaper | Large | Medium | Small | Color | Black & White



istockphoto\_488...an.jpg  
253 x 380 | 23.4kB  
[www.istockphoto.com](http://www.istockphoto.com)



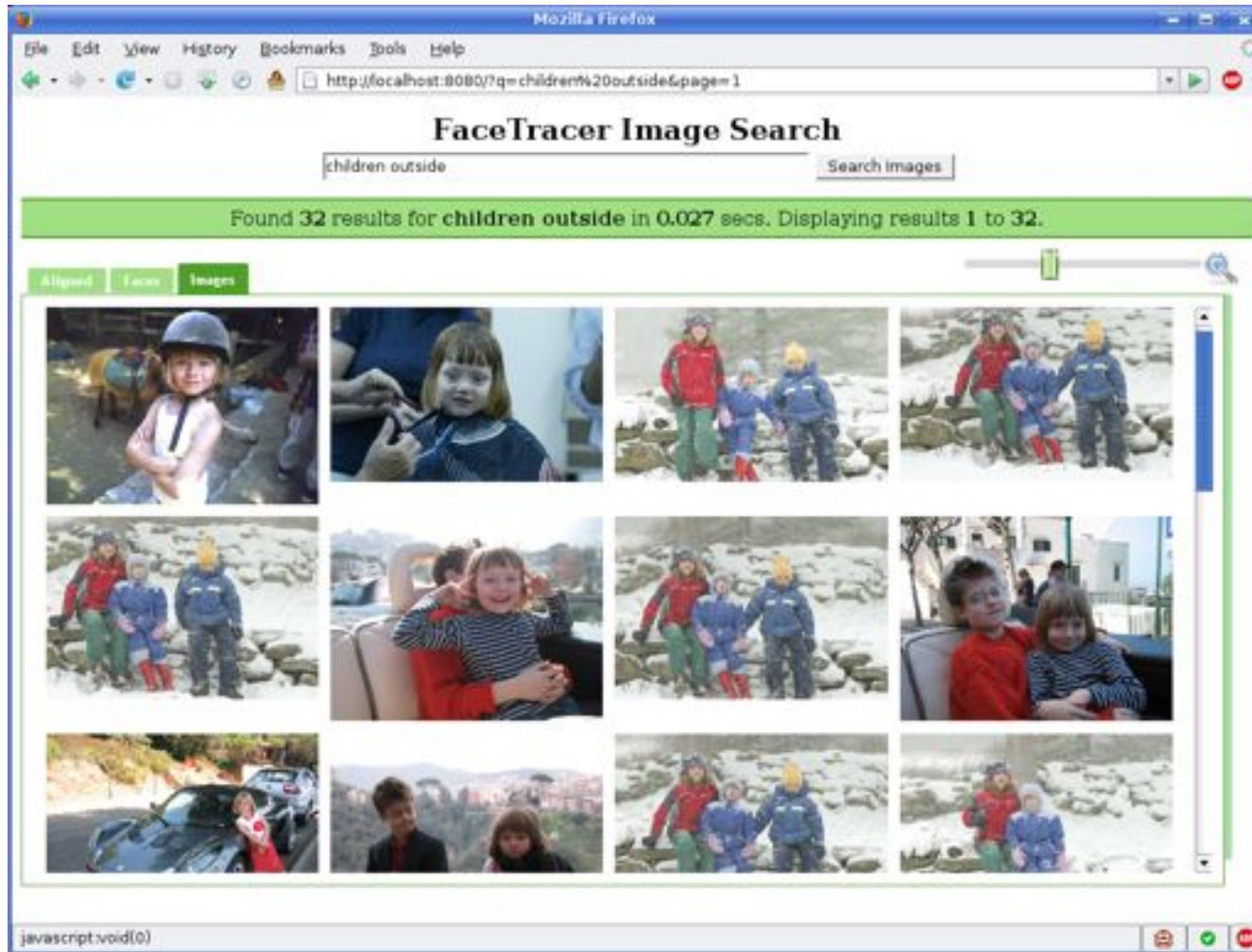
dr029.jpg  
400 x 291 | 30.9kB  
[www.inmagine.com](http://www.inmagine.com)



spo033.jpg  
400 x 294 | 37.6kB  
[www.inmagine.com](http://www.inmagine.com)

Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/facesearch/#slides>

# Personal FaceTracer Search



“Children outside”

Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/faceseach/#slides>

# Discussion...

- While this method works very well for descriptive attribute based search on facial images, is **it scalable to the general type of queries** that most image search engines use?

# Major contribution of this paper

The idea of allowing people to search for faces with descriptive terms by learning nameable semantic attributes for facial images.

# Strengths

- The idea of combining Boosting with SVMs
- This helps on two fronts:
  1. SVMs are powerful classifiers unlike the usual “weak” classifiers that Boosting is used on. This can help where Boosting usually fails.
  2. Boosting helps in selecting the optimal set of features from the variety of feature/region choices available.
- Combining Boosting with SVMs seems to obtain better results than the state-of-the-art approaches that use solely Boosting or SVMs.

# Strengths

- Finds an optimal set of relevant features/ regions for training for each attribute
- Approach implemented on the largest collection of images of “real-world” faces.
- Easily extensible to new attributes
- Handles both facial attributes & image attributes



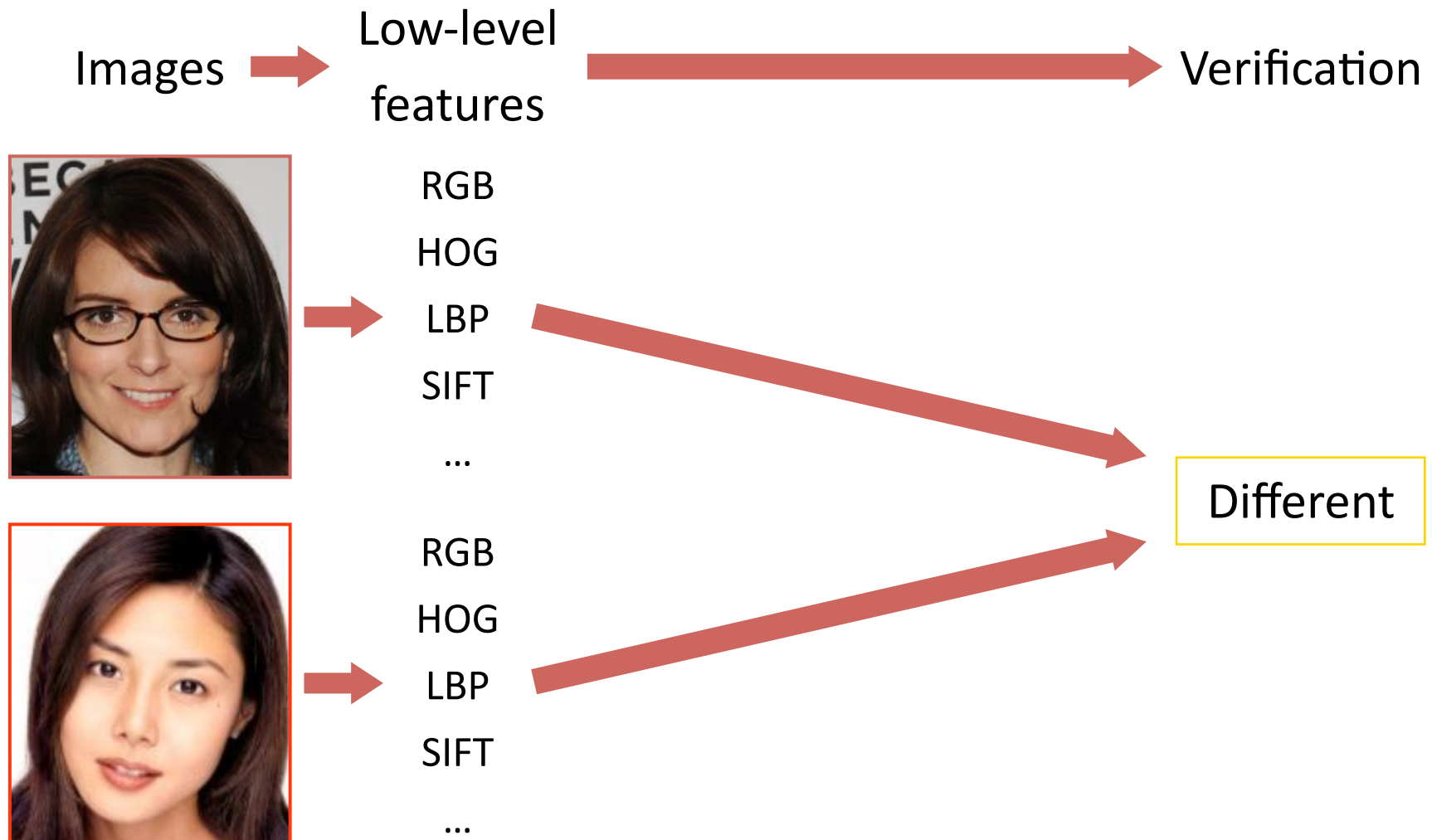
# Weaknesses

- Limited to frontal poses only
- The methods they compare with do not use the same/similar set of features and instead use a relatively impoverished set of features. As such, they do not seem to be fair baselines

# Related work

- "Attribute and Simile Classifiers for Face Verification," (oral presentation)
- [Neeraj Kumar](#), [Alexander C. Berg](#), [Peter N. Belhumeur](#), [Shree K. Nayar](#), Proceedings of the 12th IEEE International Conference on Computer Vision (ICCV),
- October 2009.

# Prior approaches



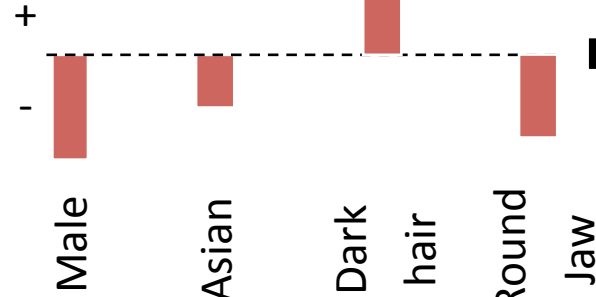
Source of slide: <http://homes.cs.washington.edu/~neeraj/projects/faceverification/>

# Their approach: attributes

Images → Low-level features → Attributes → Verification



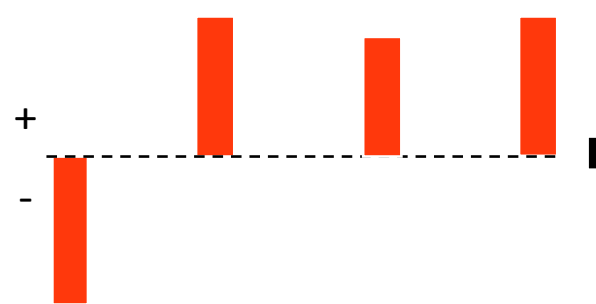
RGB  
HOG  
LBP  
SIFT  
...



Different



RGB  
HOG  
LBP  
SIFT  
...



# References/Resources

- <http://homes.cs.washington.edu/~neeraj/projects/facesearch/>
- Attribute and Simile Classifiers for Face Verification, N. Kumar, A. Berg, P. Belhumeur, S. Nayar. ICCV 2009
- Multi-Attribute Spaces: Calibration for Attribute Fusion and Similarity Search. W. Scheirer, N. Kumar, P. Belhumeur, T. Boult. CVPR 2012
- FaceTracer: A Search Engine for Large Collections of Images with Faces. N. Kumar, P. Belhumeur, and S. Nayar. ECCV 2008