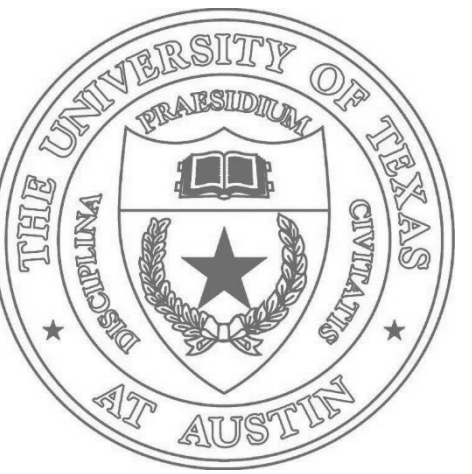


# Just Noticeable Differences in Visual Attributes

Aron Yu

Kristen Grauman

University of Texas at Austin



## Visual Comparisons



**Previous Assumption:** Relative attribute strength can and should always be *orderable* at test time.

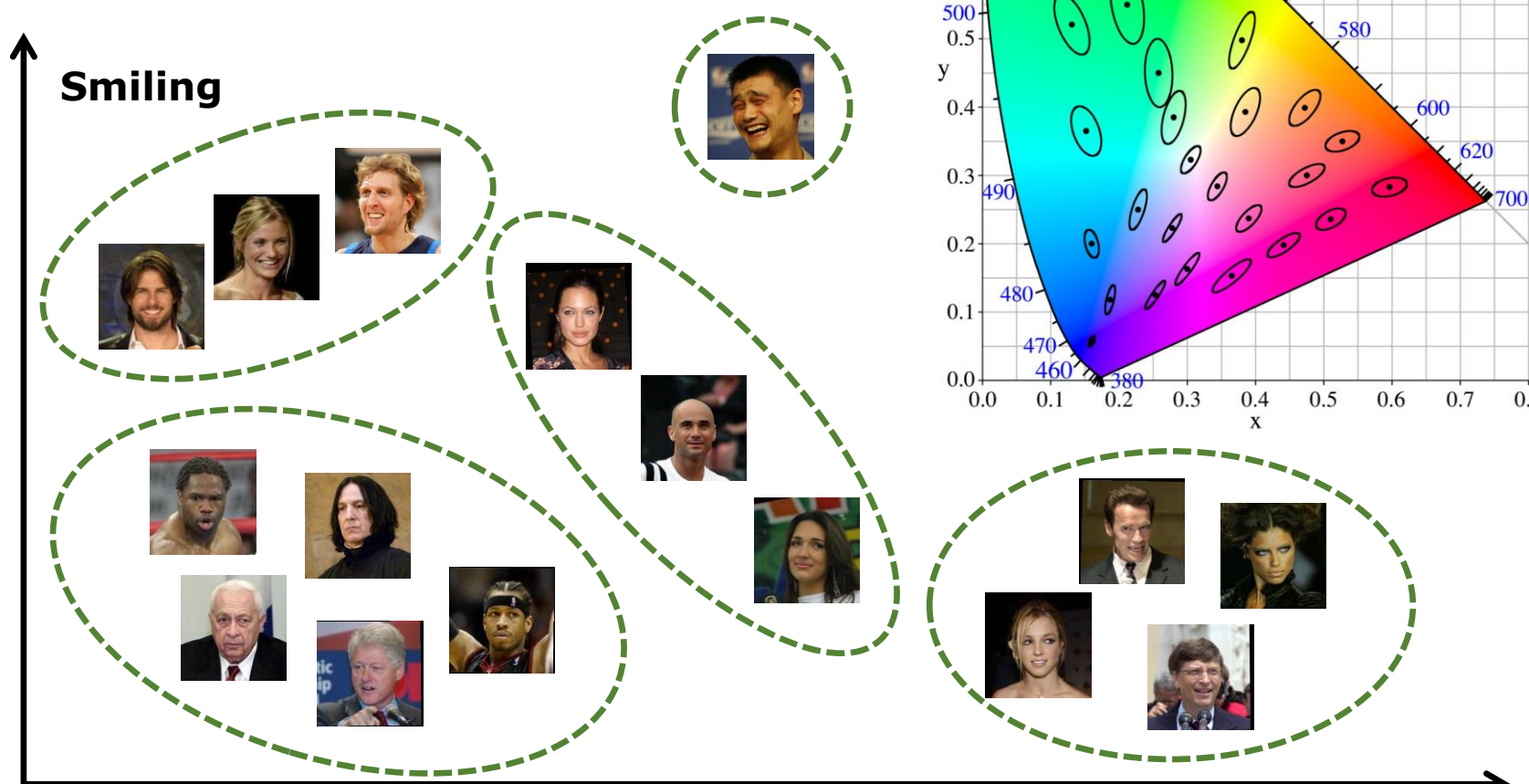
- lack of attention and evaluation on *equality pairs*
- for fine-grained domains, ~40% of human labels are equal!!

**Question:** At what point is the strength of an attribute *indistinguishable* between a pair of images?

## Just Noticeable Differences (JND)

**Definition:** The amount a *stimulus* has to be changed in order for it to be *detectable* by human observers.

- adapted from psychophysics
- analogous to MacAdam ellipses in the CIE xy color space



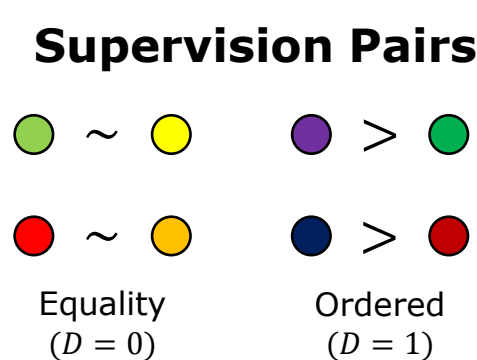
### Challenges

- detecting subtle, yet still perceptible differences
- stringent annotation requirements for precision
- non-uniformity of attribute space => no global threshold

## Local Bayesian Model

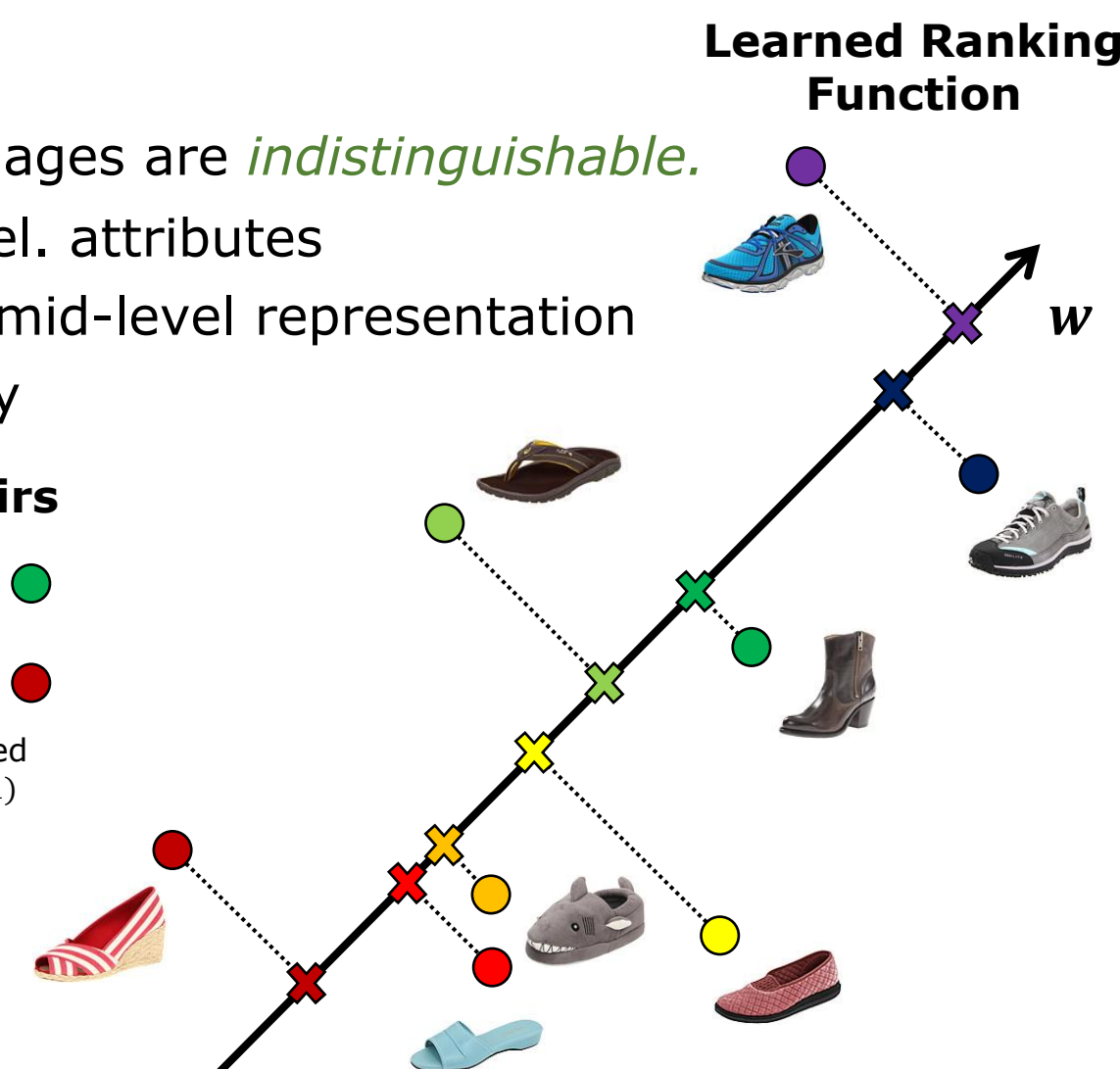
**Goal:** Given an attribute, infer when images are *indistinguishable*.

- unifies the notion of "equality" into rel. attributes
- use rel. attribute ranks as imperfect mid-level representation
- relies on local statistics of orderability

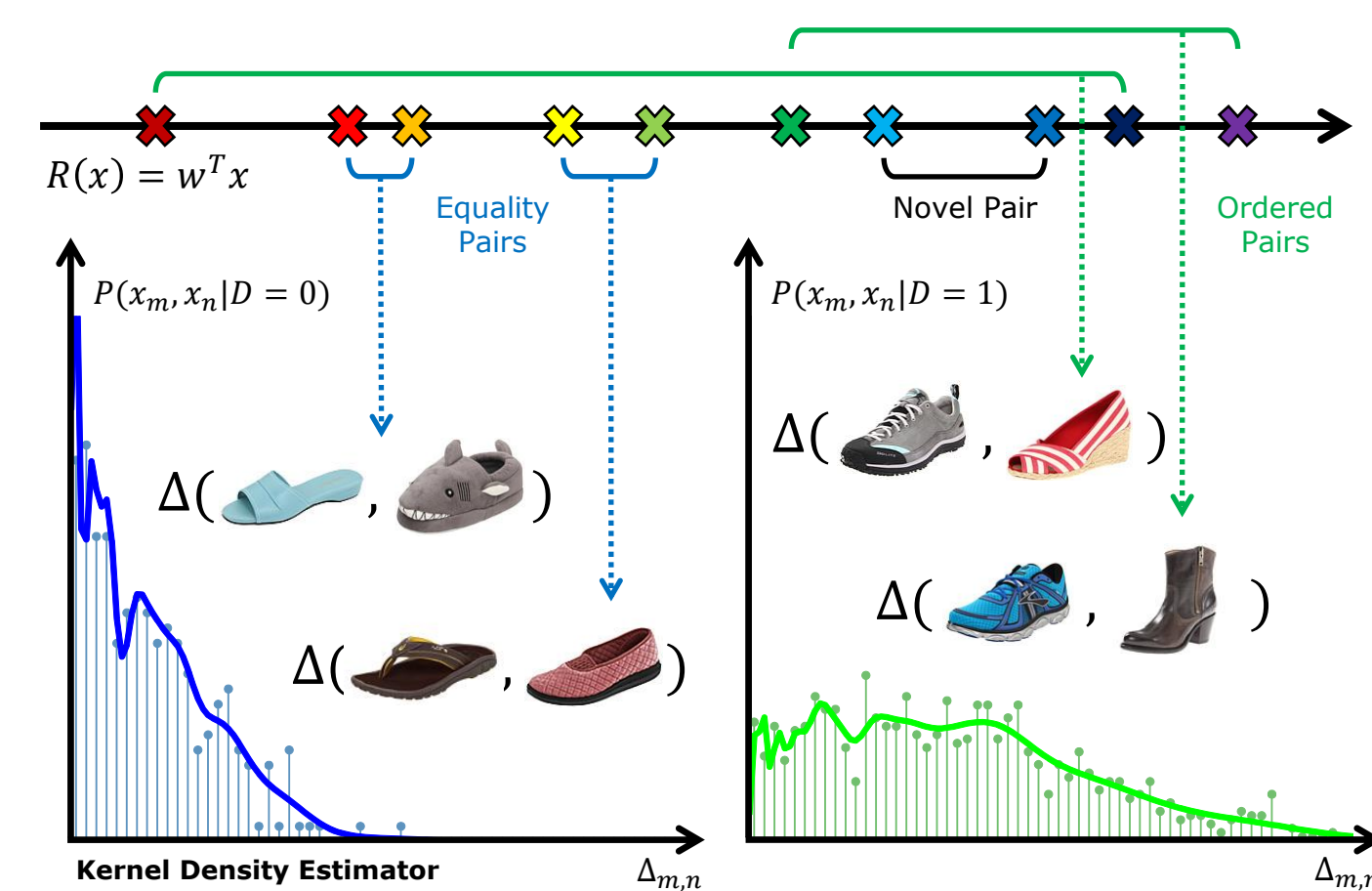


$$P(D|x_m, x_n) \propto P(x_m, x_n|D)P(D)$$

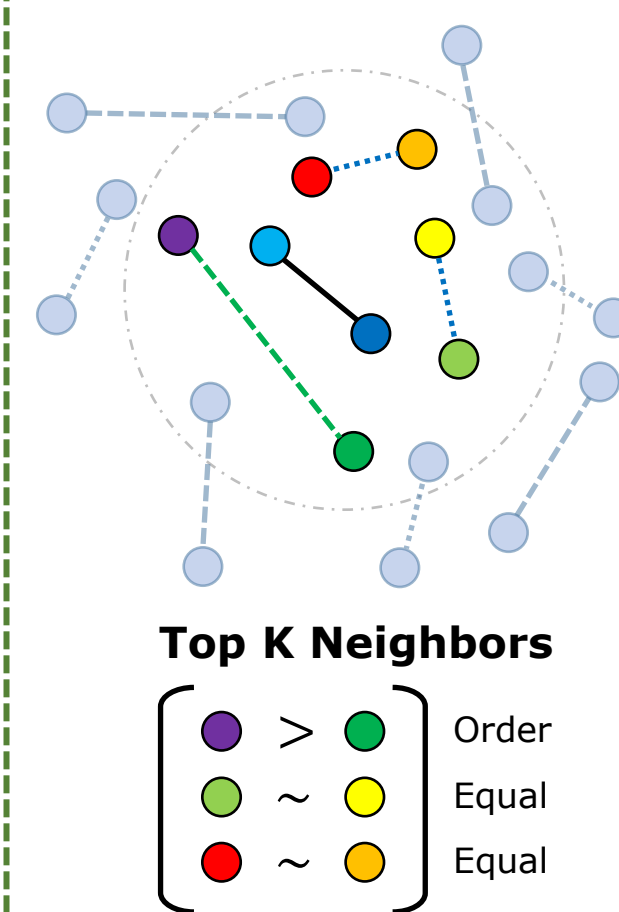
$$d^* = \operatorname{argmax}_d P(D = d|x_m, x_n)$$



### Likelihood Model ( $P(x_m, x_n|D)$ )



### Prior Model ( $P(D)$ )



## Experimental Setup

### Datasets

- instance-level relative supervision
  - ordered and equality labels
- (1) **UT-Zap50K** [Yu & Grauman 14]
    - two-stage crowdsourcing w/ confidence
    - 50,025 shoe images, 4 attributes
    - 4,778 pairs, 800 ordered, 350 equality
  - (2) **LFW-10** [Sandeep et al. 14]
    - 2,000 face images, 10 attributes
    - 5,543 pairs, 230 ordered, 320 equality
    - 8,300-dim part-based features

### Baselines

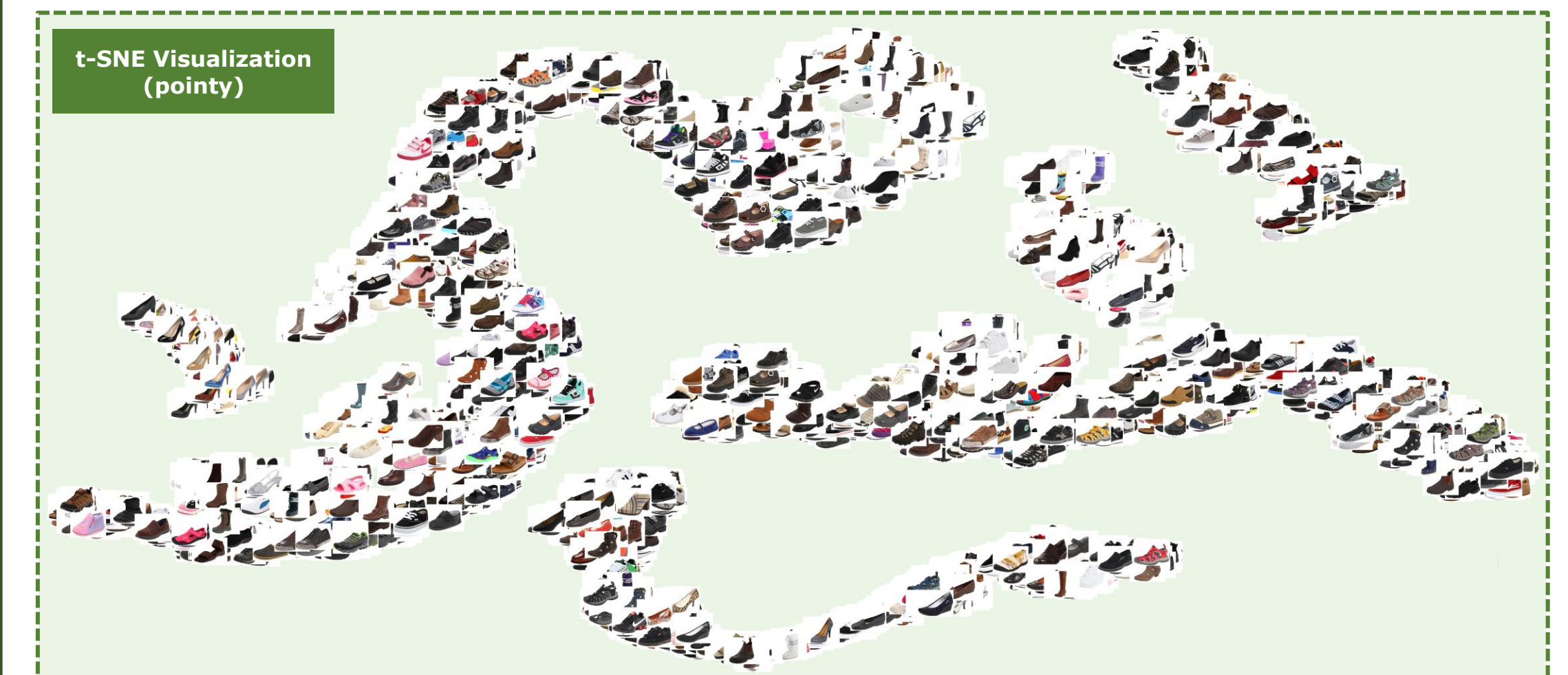
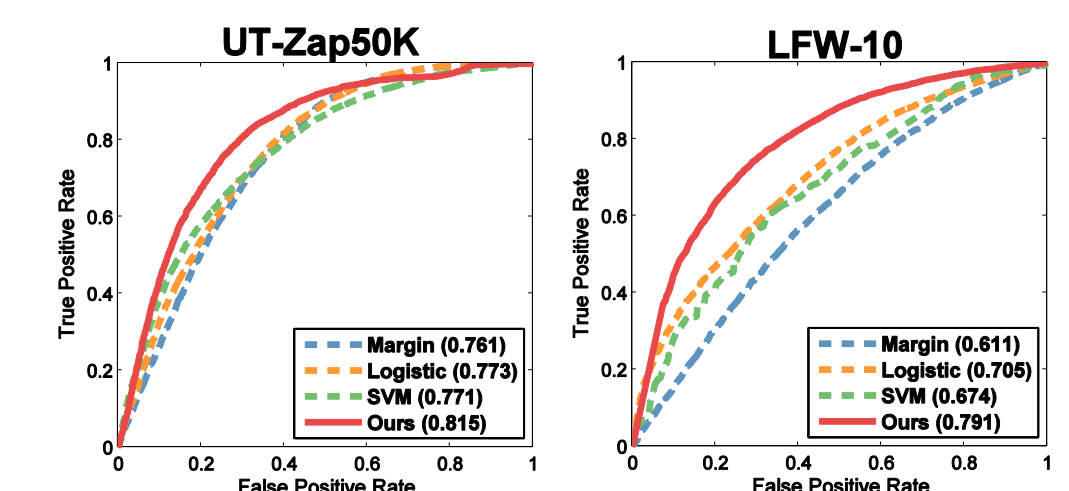
- **Rank Margin:** treat  $\Delta_{m,n}$  as confidence for distinguishability and apply global threshold
- **Logistic Classifier** [Kovashka & Grauman 13]: logistic regression classifier where pairs are represented by  $\Delta_{m,n}$
- **SVM Classifier:** nonlinear SVM classifier with RBF kernel; encode pairs of images down to single points
- **Mean Shift:** mean shift clustering on  $R(x)$  for all training images; images in the same cluster are indistinguishable

## Experimental Results

### JND Detection

	Indistinguishable			Distinguishable		
Sporty						
Smiling						
Error Cases						

**Observation:** Qualitative prediction results. Pairs may look very different overall yet still be indistinguishable *in the context of a specific attribute*. Those that are distinguishable may only exhibit subtle differences.



### Image Search Application



**Observation:** We enhance existing relative attribute search technique called WhittleSearch [Kovashka et al. 12] with our JND model.