Attribute Pivots for Guiding Relevance Feedback in Image Search

Adriana Kovashka Kristen Grauman
The University of Texas at Austin

Problem

- Traditional relevance feedback methods focus on binary feedback.
- Attributes allow more precise semantic feedback.

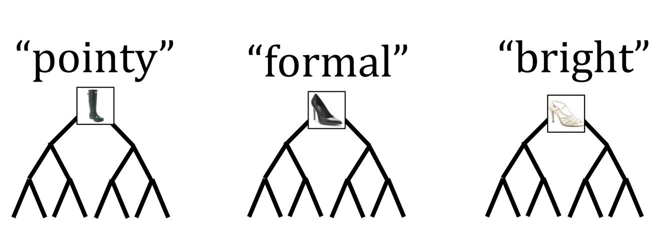
- But on which images would attribute feedback be *most informative*?

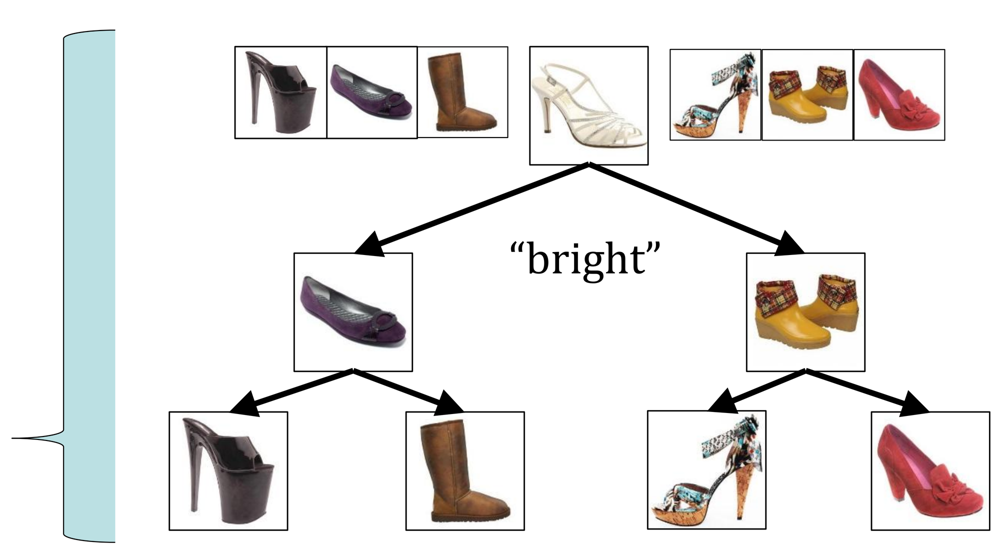
Our Idea

- Select series of most informative *visual comparisons* that user should make to help deduce target.
- Use binary search trees in attribute space for rapid active selection and to focus on useful comparisons.

Attribute Trees

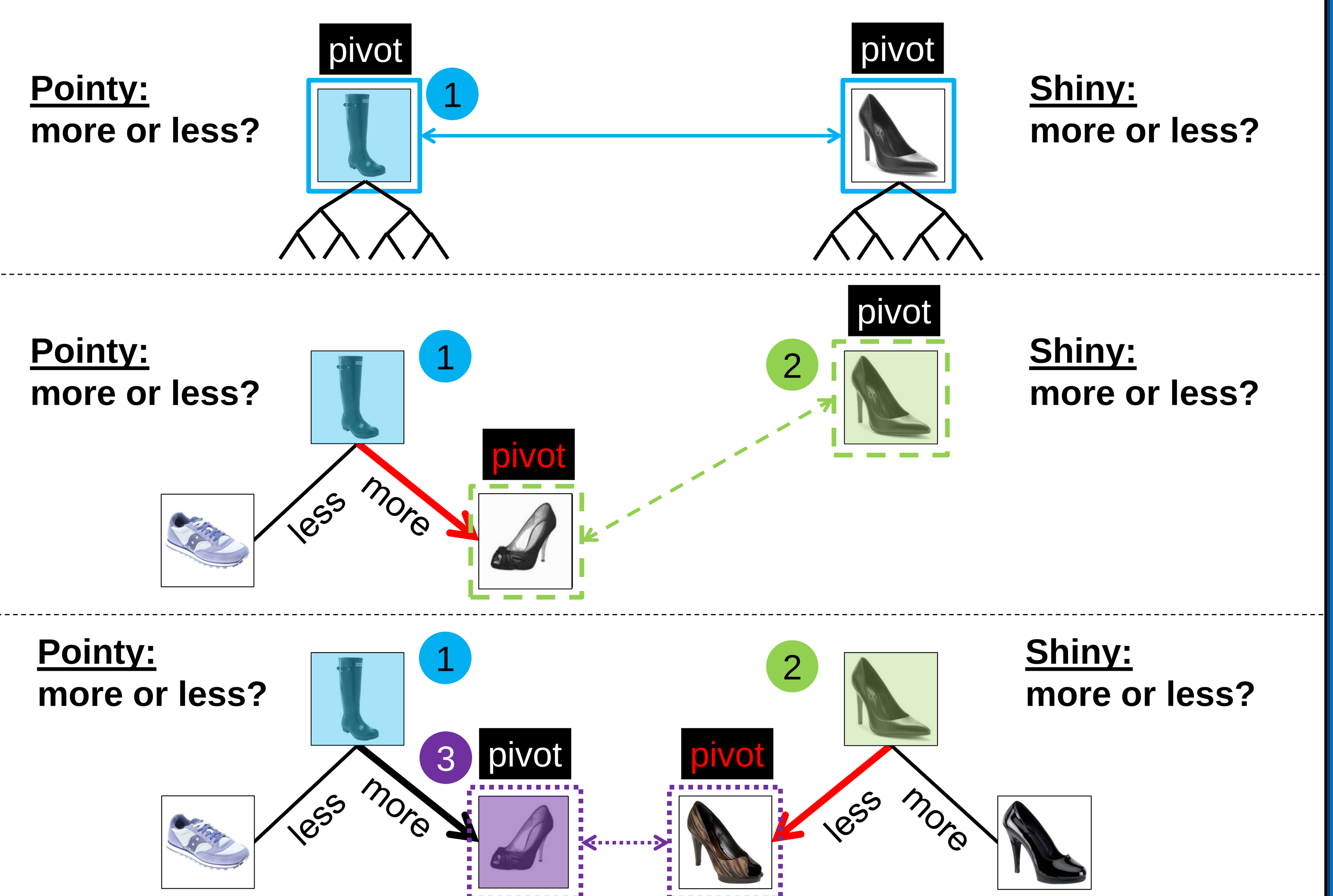
Relative attributes = learned ranking functions





Interactive Selection

Find series of useful comparisons, a la *relative* 20 questions game.



Probabilistic Model of Relevance

Relevance score for an image: $P(y_i = 1 | I_i, \mathcal{F}) = \sum_{k=1}^{T} \log \underbrace{P(S_{k,i} = 1 | I_i)}_{k\text{-th constraint satisfied?}}$

Probability a constraint is satisfied:

$$P(S_{k,i} = 1 | I_i) = \begin{cases} P(A_m(I_i) > A_m(I_p)) & \text{if } r = \text{"more"} \\ P(A_m(I_i) < A_m(I_p)) & \text{if } r = \text{"less"} \\ P(A_m(I_i) = A_m(I_p)) & \text{if } r = \text{"equally"} \end{cases}$$

True attribute comparison — User response

Selecting the Next Comparison

Entropy of system given current feedback:

$$H(\mathcal{F}) = -\sum_{i=1}^{N} \sum_{\ell} P(y_i = \ell | I_i, \mathcal{F}) \log P(y_i = \ell | I_i, \mathcal{F})$$

Choose pivot comparison that minimizes expected entropy:

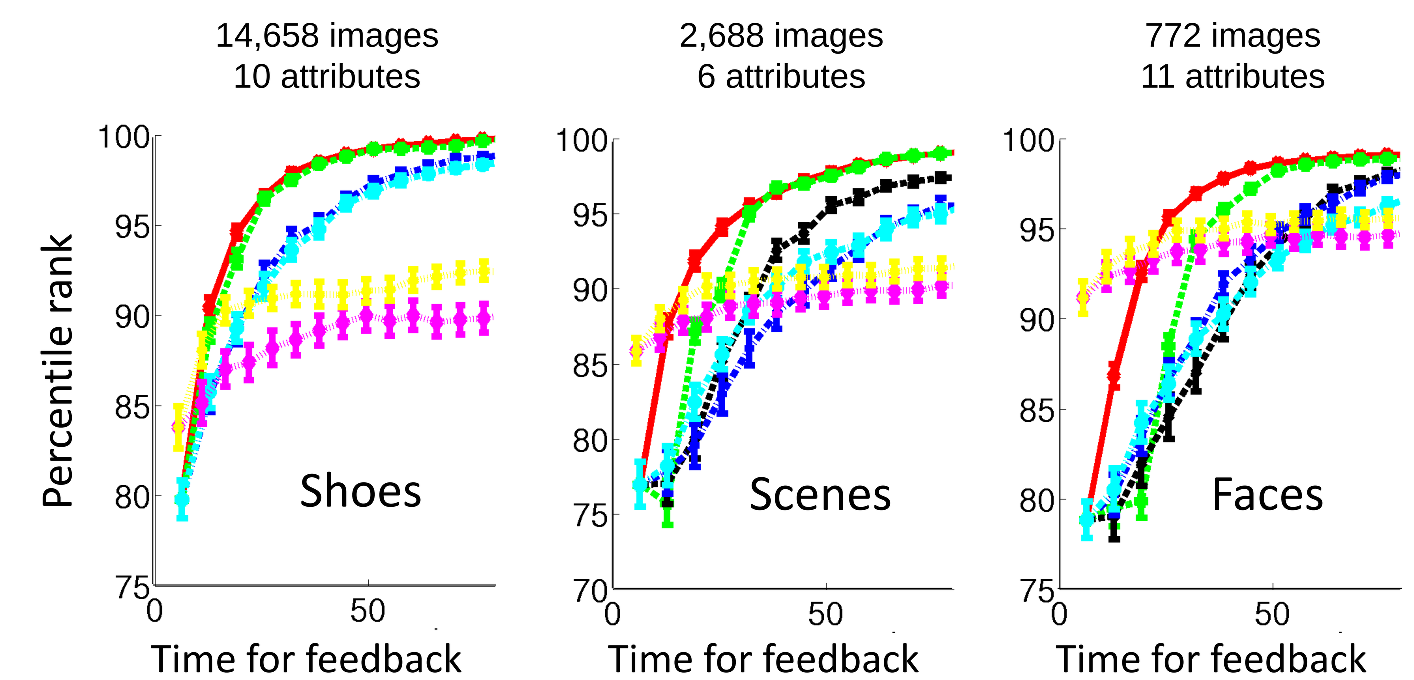
$$\underbrace{I_p^*}_{\text{Most informative pivot}} = \arg\min_{I_{pm} \in \mathcal{P}} \sum_{r} \underbrace{P(R = r | I_{pm}, \mathcal{F})}_{\text{Likelihood that user responds with } r} \overbrace{H(\mathcal{F} \cup (I_{pm}, r))}^{\text{Entropy given added feedback}}$$

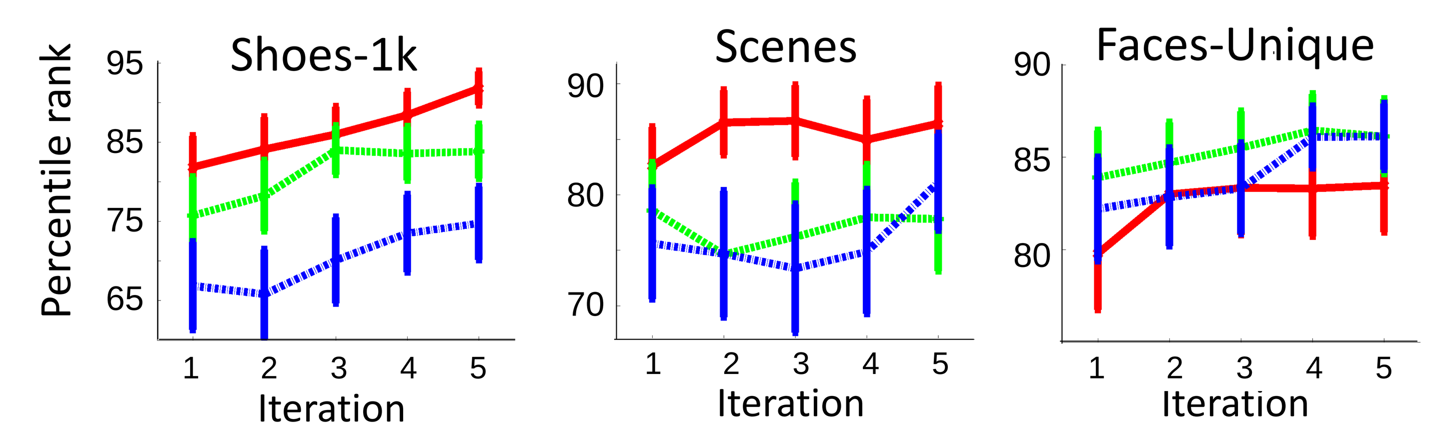
more / less / equally

Examples of Real User Searches

Quantitative Results

- Active attribute pivots (Ours)
- Attribute pivots (Ours)
- Active attribute exhaustive
- Top
- Passive
- Active binary feedback
- Passive binary feedback





Our method correctly places the target near the top of the results page, with less user effort than both passive and active methods.

Computational Efficiency

Our method is much faster than the traditional exhaustive active approach that scans all images.

Method / Dataset	Shoes	Scenes	Faces
Active attribute pivots (Ours) – *O(MN)*	0.05 s	0.01 s	0.01 s
Active attribute exhaustive – *O(MN²)*	656.27 s	28.20 s	3.42 s
Number of images (*N*)	14,658	2,688	772
Number of attributes (*M*)	10	6	11

Conclusion

- Our method takes up to 11 fewer iterations per query, and saves the user 70 seconds.
- This retrieval speed-up is achieved at a low computational cost.