Multi-Attribute Spaces: Calibration for Attribute Fusion and Similarity

Walter J. Scheirer[1], Neeraj Kumar[2], Peter N. Belhumeur[3], and Terrance E. Boult[1]
[1] Univ. of Colorado at Colorado Springs & Securics, Inc. | [2]Univ. of Washington | [3]Columbia University

Motivation

Visual attributes are a powerful approach for recognition, retrieval, image description and scene understanding. For example:

Binary: black/white stripes

Ferrari and Zisserman NIPS 2007

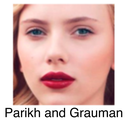
Has: Wheel, Helmet, Cloth

Farhadi et al. CVPR 2009

White: yes; Brown: no Stripes: no; Water: yes

Lampert et al. CVPR 2009

More white than Alex Rodriguez



Parikh and Grauman ICCV 2011

How do you calibrate multiple attribute applications for a single application?

Previous Approaches	Drawbacks
Use raw outputs	Not calibrated
Fit to a Gaussian	Most distributions are not Gaussian
Combine using SVM	Combinatorial constraints

Proposed Approach: Normalize scores using **Multi-Attribute Spaces**

Distance between raw attribute values does not correspond with visual similarity

Distances between points in a multi-attribute space correspond with perceptual visual similarity

Benefits

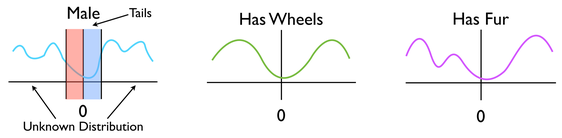
- Calibrated scores correspond to probability of human labels
- Greatly improves multi-attribute search queries
- Allows for "target attribute similarity searches"
- Does not assume any particular distribution over all scores
- Requires no ground truth labeling
- Fast and easy to compute
- Widely applicable to many types of attributes

Calibration of SVM Decision Scores

Goal: Map a set of raw binary SVM outputs to probabilities that humans would assign given labels to images.

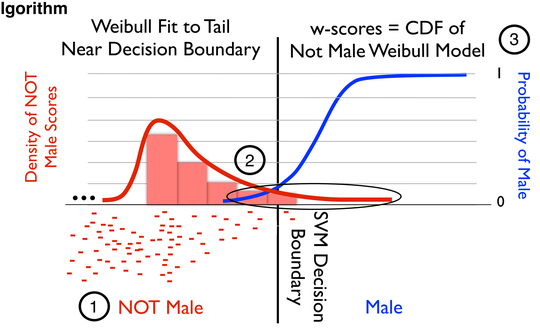
Problem: Outputs from SVM do not follow a standard distribution.

SVM decision score distributions for 3 classifiers:



- SVM scores that are far away from the decision boundary are not very informative
- The overall distribution can vary quite a bit depending on the data
- But values close to the decision boundary (around 0) are much more informative
- If scores are bounded from above and below, the Extreme Value Theory says that values in the tail **must** come from the Weibull distribution
- So given a set of SVM scores, a Weibull can be fit using the values in the tail
- The Weibull CDF gives the expected probability of a label being assigned an SVM value

Algorithm



1. Get tail of decision scores from the **opposite** class
2. Fit **Weibull** to values in the **tail:**

$$f(x;\lambda,k) = \begin{cases} \frac{k}{\lambda}(\frac{x}{\lambda})^{k-1} e^{-(x/\lambda)^k} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

3. Compute normalized scores using **CDF** of the Weibull:

$$F(x;k,\lambda) = 1 - e^{-(x/\lambda)^k}$$

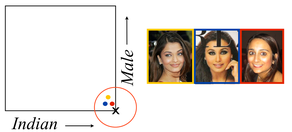
Combining Calibrated Attributes

1. maximize over I — $s^q = \| A_j(I) \|_1$
2. subject to — $A_j(I) = F(T(s_j(I)); W_j)$
3. for $\forall j \in J$ satisfying — $0 \leq \alpha_j \leq A_j(I) \leq \beta_j \leq 1$

The goal is to find images I that maximize the L_1 norm for each attribute j in the query set J

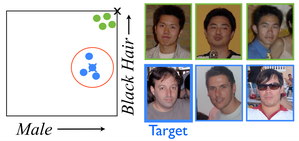
Multi-Attribute Search

"Indian Females"

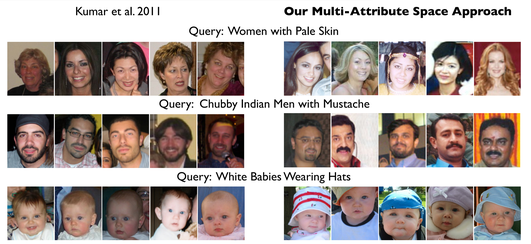


Target Attribute Similarity Search

"Male and Black Hair Like Target"



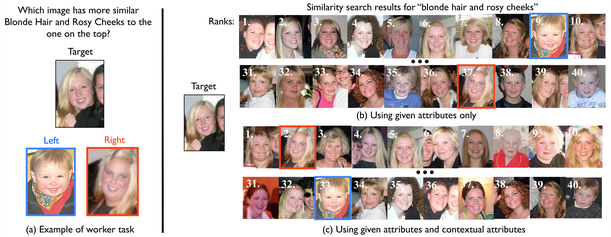
Multi-Attribute Queries



User Study: Which search results are more relevant for 900 trials?

Kumar et al.	13.1%
Proposed Approach	**86.9%**

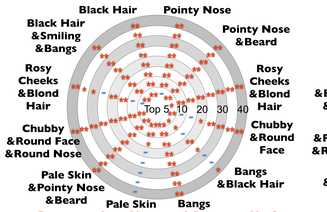
Target Attribute Similarity Searches



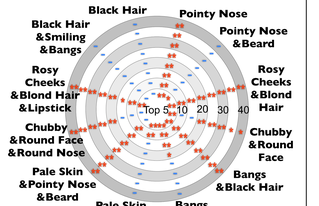
Quantitative Evaluation

Using 120,000 collected human rankings, we evaluated the statistical significance of improvements in target attribute similarity search results across 12 queries. Results shown in rings denoting top-5, top-10, etc.

Key:
** statistical significance at $p = 0.01$
* statistical significance at $p = 0.05$
- no statistical significance



Do query-only rankings match human ranking?



Are query+context rankings better than query-only?

Conclusions:

- Both algorithmic rankings match human rankings well
- Contextual attributes improve results for some queries but not others

Software & References

- The Search Engine: http://mughunt.securics.com
- The Attribute Service: http://afs.automaticfacesystems.com
- The Meta-Recognition Library: http://www.metarecognition.com

1. N. Kumar, A.C. Berg, P. N. Belhumeur and S.K. Nayar, "Describable Visual Attributes for Face Verification and Image Search," IEEE TPAMI, 33(10): 1962-1977, Oct. 2011.
2. W.J. Scheirer, A. Rocha, R. Micheals and T.E. Boult, "Robust Fusion: Extreme Value Theory for Recognition Score Normalization," *ECCV*, Sept. 2010.
3. W.J. Scheirer, A. Rocha, R. Micheals and T.E. Boult, "Meta-Recognition: The Theory and Practice of Recognition Score Analysis," *IEEE TPAMI*, 33(8): 1689-1695, Aug. 2011.
4. E. Gumbel, *Statistical Theory of Extreme Values and Some Practical Applications*. Number National Bureau of Standards Applied Mathematics in 33. U.S. GPO, Washington, D.C. (1954).

Acknowledgements

This work was supported by ONR SBIR Award N00014-11-C-0243 and ONR MURI Award N00014-08-1-0638.