

Biometrics on Visual Preferences: a "Pump and Distill" Regression Approach



Cristina Segalin¹, Alessandro Perina², Marco Cristani¹ ¹Computer Science Department, University of Verona, Verona, Italy ²Istituto Italiano di Tecnologia (IIT), Genova, Italy



Motivations and Goals

Behavioral biometric traits encode a characteristic linked to the behavior of a person [1]. The so-called HCI-based behavioral biometrics [5] are based on the idea that every person has a unique way to interact with a personal computer.

We present a statistical behavioral biometric approach for recognizing people by their **aesthetic preferences** using color images, on a dataset of 40000 Flickr images and 200 users.

The "Pump and Distill" Regression Approach

PHASE 0 - Features extraction:

Each image x is composed by the concatenation of features ranging from color statistics, image aesthetics cues, objects detections.

PHASE 1 – Pumping:

- Dataset divided in *gallery* and *probe* set (100 preferred images/user both)
- Cluster gallery set into K cluster \rightarrow thematic exemplar μ_k , k = 1, ..., K
- Augment the gallery set by **bagging** [3] \rightarrow For each user v, v = 1, ..., V, bagging generates G new training sets, the bags $B_q^{(\nu)}$, g = 1, ..., G.

Results: Recognition

- Given the classifier template $w^{(v)}$ of the user v, the matching scores is aimed at measuring how likely the set $\{x^{(u)}\}$ of the user u contains images which are in accord with the surrogates $\{z^{(v)}\}$ by the user v.
- Compute for every image $x_m^{(u)}$ in the testing pool of user u the regression score $\beta_m^{(u,v)} = \boldsymbol{w}^{(v)T} \boldsymbol{x}_m^{(u)}$
- Final matching score of the whole pool for each classifier $\{w^{(v)}\}$ is determined as the averaged regression scores of the images belonging to it: $\beta^{(u,v)} = \frac{1}{M_{TE}} \sum_{m=1}^{M_{TE}} \beta_m^{(u,v)}$
- Build a **CMC curve**: given a probe set of images coming from a single user and the matching score, the curve tells the rate at which the correct user is found within the first r matches, with all possible rspanned on the x-axis (in our case, r = 1, ..., 200).

PHASE 2 – Distill:

Generate a set of surrogates $\{z\}$ for each bag $B_a^{(v)}$, a user-customized representation of a cluster:

- Assign each image $\mathbf{x} \in B_q^{(v)}$ to the nearest thematic exemplar μ_k
- Create a surrogate z by averaging all images associated to a given cluster

PHASE 3 – Regression:

Use surrogates of all the users to learn a per-user classifier performing LASSO [4] Training:

- Assign to all the training surrogates $\{z_n\}, n = 1, ..., N^+$ of a user the $y_n = +1$ label, and -1 to all the other N^- surrogates
- Regress a label assuming it as a linear combination of the image features: $y_n = w^{(v)T} z_n$ where $w^{(v)}$ is the linear classifier for the user v obtained by minimizing the error function with the standard least squared estimate: $E(w^{(v)}) = \sum_{n=1}^{M} (y_n - w^{(v)T} z_n)^2$

Testing: Match a pool of probe images of user *u* with the gallery biometrical traits of user v, represented by his positive surrogates

Results: Authentication



CMC curve using M = 5 images per bag, and TE = 20 test images, against the [2] approach. The highest difference in terms of recognition rate (the probability of having the correct match in the first r ranked positions) is localized in the first ranks, where is expected to find the correct match.

$M (= ^{(v)}B_g)$	nAUC CMC			
	$M_{\rm TE}=1$	$M_{\rm TE}=5$	$M_{\rm TE}=20$	$M_{\rm TE}=100$
5	0.69	0.83	0.92	0.96
10	0.69	0.83	0.92	0.96
20	0.68	0.82	0.91	0.96
50	0.68	0.80	0.90	0.95
100	0.66	0.79	0.89	0.95

For every user u, take the probe set of user u as *client* images, and all the other probe sets as *impostor* images.

Given an *"authentication threshold"*, i.e. a value over which the subject is authenticated, sensitivity (TPR) and specificity (TNR) are computed using the averaged matching scores derived from the set of probe images. By varying this threshold, the **ROC curve** is obtained.

$M (= ^{(v)}B_g)$	nAUC ROC			
	$M_{\rm TE}=1$	$M_{\rm TE}=5$	$M_{\rm TE}=20$	$M_{\rm TE}$ =100
5	0.70	0.83	0.91	0.95
10	0.67	0.83	0.91	0.95
20	0.67	0.82	0.90	0.94
50	0.67	0.80	0.90	0.94
100	0.66	0.79	0.88	0.94
[2]	0.65	0.76	0.82	0.87

nAUC values of ROC curves varying num. of images per bag (M), and the M_{TE}

0.66 0.77 0.84 0.88 [2]

nAUC values of the CMC curves varying the num. of images per bag (M), and the num. of test images (M_{TE}). We can think of [2] as having bags formed by one image. This suggests that pooling together images into surrogates produces more discriminative information.

References

[1] A.K. Jain, P. Flynn, and A.A. Ross, Handbook of Biometrics, Springer, 2008.

[2] P. Lovato, A. Perina, N. Sebe, O. Zandonà, A. Montagnini, M. Bicego, M. Cristani: "Tell Me What You Like and I'll Tell You What You Are: Discriminating Visual Preferences on Flickr Data", ACCV 2012, pp. 45-56

[3] Leo Breiman, "Bagging predictors," Machine learning, vol. 24, no. 2, pp. 123–140, 1996. [4] R. Tibshirani, "Regression shrinkage and selection via the lasso," J. of the Royal Statistical Society, Series B, vol. 58, pp. 267–288, 1994.

[5] R.V. Yampolskiy and V. Govindaraju, "Behavioural biometrics: a survey and classification," Int. J. Biometrics, vol. 1, no. 1, pp. 81–113, 2008.

IEEE International Conference on Image Processing - ICIP 2014