Auto Arrangement and Threshold to Database for Face Recognition

Mu Chien Hsu[†] Waseda University [†]
Jui Chun Shyur[‡] Compal Electronics[‡]
Hiroshi Watanabe [†] Waseda University [†]

Abstract: We are given a real face database for analysis and process to improve the performance of 1:1 recognition (P1) and 1:N identification (P2) face recognition problems[1]. We identified that Euclidian distances among face embeddings are normal distributions. Base on the statistical result, we can simply use normal distribution model to replace clusters of photos, and meanwhile determine a statistically meaningful threshold according to different requirement. We now use normal distribution model to auto-arrange database and auto-determine threshold according to different requirements. Results show that we can achieve equivalent recognition accuracies (P1 AUC = 0.999, P2 top-1 accuracy = 98.88%) with 90 times faster, and the prediction error of P1 true positive rate (TPR) is only 1.14% and P2 false positive rate (FPR) only 0.05% on average in the meaningful threshold range.

1 Introduction

The face set ($\mathscr{F}$) is from a real unconstrained environment [2,3] with 1,903 persons, 190,033 images, each person $\boldsymbol{p}$ has 20 ~ 250 faces in a cluster C_p. For convenience, we use face and face embedding interchangeably. In this section, we will show how normal distributions are observed as insights for auto-arrangement and threshold to take effect. For C_p, we define:

- f_{pc} : *center face*, the projection from the mean of $f_{p,i}$ to the hyper sphere surface
- $D_{p,intra}$= $\{\|f_{pc} - f_{p,i}\|_2\}$: the set of *intra-personal distance*
- $D_{p,extra} = \{\ \|f_{pc} - f_{q,i}\|_2\}$ for all $f_{q,i}$ *of* C_q, $q \neq p$: the set of *extra-personal distance*
- $\epsilon_{p,intra} = \overline{\|f_{pc} - f_{p,\iota}\|_2}$, $\epsilon_{p,extra} = \overline{\|f_{pc} - f_{q,\iota}\|_2}$: the mean of $D_{p,intra}$ and $D_{p,extra}$
- $\sigma_{p,intra}$, $\sigma_{p,extra}$: the standard deviation associated with $\epsilon_{p,intra}$ and $\epsilon_{p,extra}$

Note that $D_{p,intra}$ or $D_{p,extra}$ of different persons have different scale of mean and std, they are normalized to the same base and scale for comparison, i.e., the mean is offset by ϵ_p to be 0, and the std divided by σ_p. A few examples of the normalized distance distribution obtained are shown in Fig.1.

We then unite all these normalized $D_{p,intra}$ and $D_{p,extra}$ for all persons to obtain global set D_{intra} and set D_{extra}. By displaying their distribution in the same figure (Fig.2), we are able to see two perfect normal distributions:

F1: D_{intra} on the left, is a normal distribution where $(\epsilon_i, \sigma_i) = (0.18, 1.17)$, ϵ_i, σ_i are mean and std associated with D_{intra}

F2: D_{extra} on the right, is a normal distribution where $(\epsilon_e, \sigma_e) = (10.04, 3.05)$, ϵ_e, σ_e are mean and std associated with D_{extra}

From Fig.2, the problem of effective feature representations to reduce intra-personal variations while enlarging inter-personal differences [4] becomes an easier one by finding ways to push D_{intra} to the left, and D_{extra} to the right as far as possible, while keeping the normalization distributions. This can be optimized by our auto-arrangement in the next section. Meanwhile, n_1 in Fig.2 will predict P1 TPR and P2 FPR and form the basis of our auto-threshold process, explained in the next section.

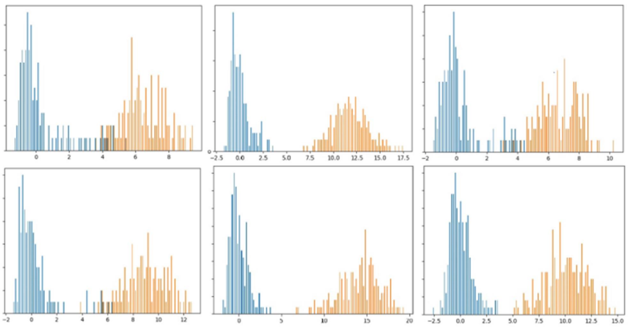

Fig.1 Six normalized ($\epsilon_p = 0$) personal $D_{p,intra}$ (left) & $D_{p,extra}$ (right) examples

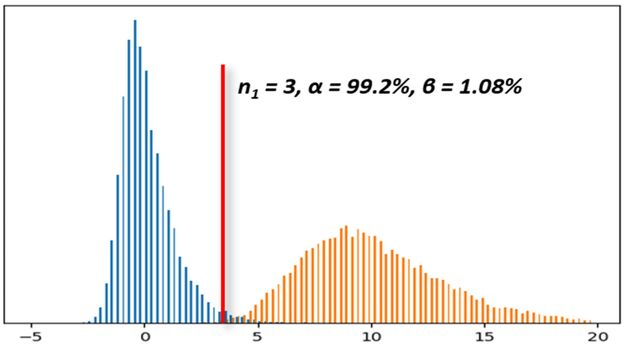

Fig.2 Normalized D_{intra} (left) & D_{extra} (right) for prediction.

2 Our Approach

We separate the face set into a test set $\mathscr{F}_t$ by randomly removed 7 ~ 10 faces from each C_p to get C'_p and use the remaining $\mathscr{F}$- $\mathscr{F}_t$ for auto-arrangement and auto-threshold. Test set $\mathscr{F}_t$ is used for experimental results.

A1: Auto-arrangement

We want to remove those faces either too far from center, or repeatedly sampled ones, thus to have a distribution moving more to the left side in Fig.2, with smaller std. To achieve this, for each C'_p, we repeatedly calculate f_{pc}, ϵ_p, σ_p and prune those $f_{p,i}$ with $\|f_{pc} - f_{p,i}\|_2 > \epsilon_p + 2\sigma_p$. The arranged cluster is then used to obtain the normalized $D_{p,intra}$ and $D_{p,extra}$ which completes the auto-arrangement process.

A2: Auto-threshold and prediction for **P1** TPR and **P2** FPR

The procedure "**A1**" will also results in the global sets D_{intra} and D_{extra} as mentioned in Introduction. Referring to Fig.2 which is used to determine P1 personal thresholds and predict P1 TPR by α and P2 FPR by β, with a given n_1 (number of std), we have $T_p = \epsilon_p + n_1 \times \sigma_p$ as the personal threshold for every C_p. Also noted that $\epsilon_i + n_1 \times \sigma_i = \epsilon_e - n_2 \times \sigma_e$, n_2 for **P2** can be calculated from n_1. According to the characteristic of normal distributions, n_1 and n_2 will one-to-one map to α and β. Thus by choosing n_1, say in a range from 2.0 to 3.0, we can have T_p for all persons, and derive the prediction of α and β. This completes the auto-threshold and α β prediction process.

3 Experiment Result

The tests for **P1** and **P2** for p are performed according to the following criteria:

1) **P1 :** positive if $\|f_{pc} - f_t\|_2 < t_p$ for $f_t \in C_p \cap \mathcal{F}_t$
2) **P2 :** negative: if $\|f_{pc} - f_t\|_2 < t_p$, for $f_t \in \mathcal{F}_t - C_p$.
3) **P2:** positive, if $\min_q \|f_{pc} - f_t\|_2 < t_p$, $q = p$, that is, the top-1 accuracy, for all $f_t \in \mathcal{F}_t$, f_t corresponds to q's embedding.

We feed $\mathcal{F}_t$ for tests of different n_1 as in Table.1 A good result for P1 is achieved by $n1$ = 4.1, with TPR = 99.21%, FPR = 0.84%; and for P2, a good one is by $n1$ = 6.9, with TPR = 98.88%, FPR = 1.12%. Our prediction α and β are accurately matching the test results as shown in Fig.3. The P1 AUC is 0.9996 to show it is an excellent classifier. As we only compare the test face against the center face, we only need 1 comparison and reduce 90% other comparisons.

4 Conclusion and Future Work

We fully utilize the normal distribution characteristic to simplify the face recognition problems with promising results in accuracy, computation, and accuracy prediction. We use the global sets D_{intra} and D_{extra} to help determine personal thresholds for individual persons. Though α and β can accurately predict P1 TPR and P2 FPR, they are limited in certain ranges. We observed these can be caused by: 1) arranged model has a smaller σ than it should have, 2) the $d_{p,extra}$ is not arranged and is less normal distribution like, 3) the face quality in $\mathcal{F}_t$ is not good enough, such as blurred and misaligned are observed and lower down P1 TPR. In the future, we will enhance on 1) face image preprocessing, and 2) Use personalized n_p instead of identical for all, to increase **P1** true positive rate (TPR) and decrease **P2** false positive rate (FPR), and to have a better prediction α and β.

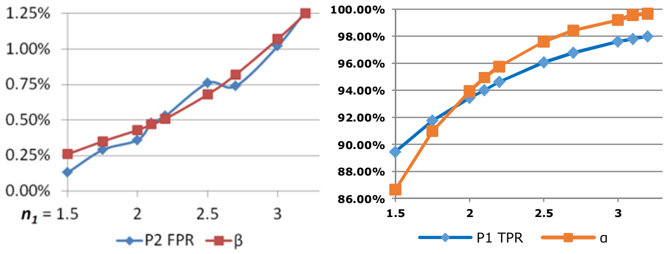

Fig. 3 Prediction accuracy of α and β

Table 1 Test Result with Predictions α and β

n_1	P2 TPR	P2 FPR	β	P1 TPR	α	P1 FPR
2.0	93.36%	0.36%	0.43%	93.45%	93.94%	0.01%
2.5	96.53%	0.76%	0.68%	96.07%	97.61%	0.03%
3.0	97.91%	1.02%	1.07%	97.62%	99.20%	0.10%
3.5	98.14%	1.25%	1.61%	98.47%	99.85%	0.29%
4.1	98.51%	1.36%	2.62%	99.21%	99.96%	0.84%
6.9	98.88%	1.12%	15.39%	100.00%	100.00%	17.29%

5 References

[1] NIST FRVT(https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt-ongoing)

[2] O. M. Karkhi, A. Vedaldi, and Z. Zieerman: "Deep face recognition," Proc. British Machine Vision, pp.41.1-41.12 Sep. 2015.

[3] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf: "Deepface: closing the gap to human-level performance in face verification," Proc. Computer Vision and Pattern Recognition (CVPR 2014), pp.1701-1708 June 2014.

[4] Y. Sun, Y. Chen, X. Wand, and X. Tang: "Deep learning face representation by joint identification-verification," Proc. NIPS, Dec. 2014.

[5] F. Schroff, D. Kalenichenko, and J. Philbin: "Facenet: A unifiedembedding for face recognition and clustering," Proc. Computer Vision and Pattern Recognition (CVPR 2015), pp.815-823 June 2015.

Dept. of.Communications and Computer Eng., School of Fundamental Science and Eng.
Shillman Hall 401, 3-14-9 Okubo, Shinjuku, Tokyo 169-0072
Phone: 03-5286-2509
michael.hsu@asagi.waseda.jp