1L−9

On Imposter Models Used for Biometric Person Verification

Weijie Liu, Satoshi Haruyama and Hitoshi Sakano *
NTT DATA CORPORATION
{wjliu, haruyama, sakano@rd.nttdata.co.jp}

1 Introduction

Biometric person recognition systems are categorized into *identification* systems and *verification* systems. In verification systems, an individual is recognized through a one-to-one matching where his feature measurement is compared directly to the template he declared. Theoretically an imposter model is necessary for each individual to complete such a matching.

Though it seems that the topic of imposter models has not yet attracted much attention, several proposals do have been suggested in researches on speaker verification and face recognition. As a simple explanation about the topic, we introduce three main existing imposter models below.

Generally, imposter models are built based on a set of training samples. We assume that our set contains N feature vectors from each of M individuals, denoted by $\mathbf{X}_{ij}, i = 1 \sim M, j = 1 \sim N$.

a. Higgins Model[1]

The ith individual is described by his probability density function(*pdf*) $p_i(\cdot)$, which is estimated using the samples $\mathbf{X}_{ij}, j = 1 \sim N$. The likelihood (no longer a *pdf*) for the related imposters is then described by

$$q_i(\mathbf{X}) = \max_{r=1\sim M, r\neq i} p_r(\mathbf{X}), \qquad (1)$$

where $\mathbf{X}$ is an arbitrary measurement. The verification is implemented based on that if the ratio $\frac{p_i(\mathbf{X})}{q_i(\mathbf{X})}$ is greater than a threshold t.

b. Cohort Model[1]

Cohorts sometimes mean neighbors, but they may also be selected randomly. Suppose the C cohorts of the ith individual are the i_cth, $c = 1 \sim C$ individuals, then the *pdf* of the related imposters are obtained using $p_{i_c}(\cdot)$'s. For example, one solution is

$$q_i(\mathbf{X}) = \frac{1}{C}\sum_{c=1}^{C} p_{i_c}(\mathbf{X}). \qquad (2)$$

c. Global Model[2]

Global model builds the same imposter model for all individuals. Usually it tries to describe the background distribution of the whole populations. For example, the *pdf* of imposters can be defined by a Gaussian function

$$q_i(\mathbf{X}) = N(\mu, \Sigma), \qquad (3)$$

where, the mean μ and the covariance Σ are estimated using all $\mathbf{X}_{ij}$'s.

2 A New Proposal

By combining the original ideas of the existing models, we can propose a new imposter model.

At first, the *pdf*'s for all individuals, $p_i(\cdot), i = 1 \sim M$, are calculated. Usually they are assumed to be Gaussian, and their means and covariance matrices are estimated by

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} \mathbf{X}_{ij} \qquad (4)$$

and

$$\Sigma_i = \frac{1}{N-1}\sum_{j=1}^{N} (\mathbf{X}_{ij} - \mu_i)(\mathbf{X}_{ij} - \mu_i)^T, \qquad (5)$$

respectively.

According to the Bhattacharyya distance between two *pdf*'s[3], the K neighbors of the ith

*劉 偉傑、春山 智、坂野 鋭
㈱NTT データ マルチメディア技術センタ
〒104-0033 東京都中央区新川 1-21-2 茅場町タワー8 F

individual are obtained. They are denoted as $i_k, k = 1 \sim K$ below.

To obtain the *pdf* for related imposters, three parameters should be calculated. They are called *Average Imposter Radius(AIR)*, *Average Imposter Dispersal(AID)* and *Measurement-dependent Imposter Orientation(MIO)* respectively.

A. *AIR*

AIR represents the statistical distance between an individual and his neighbors. It can be defined by

$$AIR = \frac{1}{M}\frac{1}{K}\frac{1}{N}\sum_{i=1}^{M}\sum_{k=1}^{K}\sum_{j=1}^{N}|\mathrm{X}_{i_k j} - \mu_i| \qquad (6)$$

or

$$AIR = \frac{1}{M}\frac{1}{K}\sum_{i=1}^{M}\sum_{k=1}^{K} b_{ii_k}, \qquad (7)$$

where, b_{ii_k} is the Bhattacharyya distance between $p_i(\cdot)$ and $p_{i_k}(\cdot)$.

B. *AID*

AID represents the statistical covariance of all neighbors. It can be defined by

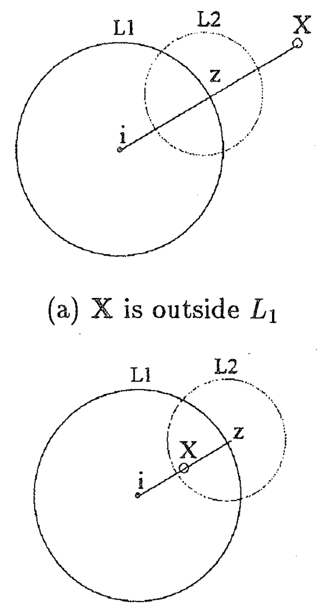
$$AID = \frac{1}{M}\frac{1}{K}\sum_{i=1}^{M}\sum_{k=1}^{K}\Sigma_{i_k}. \qquad (8)$$

C. *MIO*

As indicated in Figure 1, when i represents the ith individual, L_1 the circle with radius *AIR*, and X the measurement, the mean position of the related imposters is decided at the cross point z of line iX and circle L_1. Finally, the imposter model with respect to the ith individual and the measurement X is described by a *pdf* with mean z and covariance *AID*.

In practice, how to decide the number of neighbors K is a problem. A general consideration is to calculate the relation between K and the verification error based on the known samples. It is investigated in our experiments.

Otherwise, it should be noted that since *AID* are described as a diagonal matrix in many recognition algorithms, the calculation of imposter *pdf* then can be executed according to independent features.



(a) **X** is outside L_1

(b) **X** is inside L_1

Figure 1: An example of *MIO*

3 Experiments

The experiments are designed for face verification. Based on the concept of eigenface, the proposed imposter model is compared to the mentioned three existing models with respect to 100 individuals and about 90 face images per individual.

In nature the problem of imposter models concerns the effect of sample size. So it reflects some essential but very difficult aspects of biometric person verification[4].

References

[1] W. Liu, T. Isobe and N. Mukawa, "On optimum normalization method used for speaker verification," *Proceedings of ICSLP'98*, pp.165-168, Sydney, Dec. 1998.

[2] M. Golfarelli, D. Maio, and D. Maltoni, "On the error-reject trade-off in biometric verification systems," *IEEE Trans. on PAMI*, vol.19, no.7, pp.786-796, Jul.1997.

[3] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd Edition, Academic Press, 1990.

[4] H. Sakano, "The state of art of biometric person authentication," Technical Report of IEICE, PRMU99-29, Jun 1999.