

# Error-Rate Based Biometrics Fusion

Kar-Ann Toh

Biometrics Engineering Research Center  
School of Electrical & Electronic Engineering  
Yonsei University, Seoul, Korea  
katoh@yonsei.ac.kr

**Abstract.** This paper addresses the face verification problem by fusing visual and infra-red face verification systems. Unlike the conventional least squares error minimization approach which involves fitting of a learning model to data density and then perform a threshold process for error counting, this work directly formulates the required target error count rate in terms of design model parameters. A simple power series model is adopted as the fusion classifier and our experiments show promising results.

## 1 Introduction

Existing 2D face recognition systems are much hindered by external factors such as illumination, pose, expression and etc (see e.g. [1]). Apart from adopting complicated compensation means for illumination variations, fusion of the visual (VS) and infra-red (IR) images has been experimented ([2,3,4,5,6]). Different from these works, this paper focuses on identity verification based on fusion of matching scores from visual and infra-red face images, paying particular attention to effective and efficient fusion classifier design under difficult situations when fusing a moderately accurate system with a weak system.

Essentially, the task of identity verification can be considered as a pattern classification task since a comparison between two face images results in either a match (genuine-user) or a non-match (imposter) output. Apart from the statistical means, learning from examples constitutes a major paradigm in pattern classification (see e.g. [7,8]). Under this paradigm, a classifier is a map between the input feature space and the output hypothesis space. Although one usually does not have full access to the expected input-output spaces, a subset of this input-output space, which we called training set, is often available. Our work here belongs to the training based approach (see e.g. [9,10,11]) working at match score level.

We propose to use the total error counting rate instead of the data density fitting objective for supervised classifier learning. The error rate is a percentage count of misclassified samples and this poses difficulty to analyze it directly without imposing strong assumptions regarding the data distribution. However, utilizing a smooth functional approximation to the error counting objective, we are able to formulate a single-step solution to solve this problem. The solution

is found to belong to a particular setting of the weighted least squares solution having a class specific sample normalization. A simple power series model is found to be suitable in this classification application. The power series was not seen in biometrics fusion literature, perhaps, due to its non-compactness related to functional approximation.

The face images used in our experiments were captured under various illumination, pose and expression conditions. This gives rise to the visual face data showing relatively weak verification performance as compared to the infrared face data. Hence, the data set used demonstrates a particular difficult situation of fusing a moderately accurate biometric with a weak biometric since the weak biometric may deteriorate the fusion outcome.

## 2 An Error Counting Power Series Model

### 2.1 Power Series Model

Linear parametric models have been widely used due to their tractability in optimization and related analysis. The embedment of nonlinearities such as kernels and other basis functions into linear regression models has even widen their scope of applications (see e.g. [12,13,14]). Common choice of the basis functions includes radial basis function, polynomials, and sigmoid function [15]. Here, we consider a simple power series model as the basis function for fusion of visual and infra-red face score data.

Consider a  $l$ -dimensional input  $\mathbf{x}$  and the following power series model:

$$g(\boldsymbol{\alpha}, \mathbf{x}) = \alpha_0 + \sum_{k=1}^r \sum_{j=1}^l \alpha_{k,j} x_j^k,$$

where  $x_j, j = 1, \dots, l$  are the biometric scores for fusion,  $\alpha_0$  and  $\alpha_{k,j}$  are the weighting coefficients to be estimated.  $l$  and  $r$  correspond to input-dimension and order of system respectively. The total number of terms in this model can be expressed as  $K = 1 + rl$ . This model is different from the polynomial model in [16] ( $K = 1 + r + l(2r - 1)$ ) since it has much smaller number of expansion terms.

The above power series model can be compactly written in vector form:

$$g(\boldsymbol{\alpha}, \mathbf{x}) = \sum_{k=0}^{K-1} \alpha_k p_k(\mathbf{x}) = \mathbf{p}(\mathbf{x}) \boldsymbol{\alpha}, \quad (1)$$

where each  $p_k(\mathbf{x})$  corresponds to a power basis expansion term, and collectively these power expansion terms form a row vector  $\mathbf{p}(\mathbf{x}) = [1, p_1(\mathbf{x}), \dots, p_{K-1}(\mathbf{x})]$ .  $\boldsymbol{\alpha} = [\alpha_0, \alpha_1, \dots, \alpha_{K-1}]^T$  denotes the column parameter vector to be estimated.

When each element of  $\mathbf{x} \in \mathbb{R}^l$  has a known label  $y \in \{0, 1\}$  (e.g. ‘1’ indicates a genuine-user and ‘0’ indicates an imposter), we can form the learning data pairs  $(\mathbf{x}_i, y_i)$  from  $i = 1, 2, \dots, m$  observations. The learning problem can then

be supervised. Learning of the target labels (packed as  $\mathbf{y} = [y_1, y_2, \dots, y_m]^T$ ) is conventionally accomplished by minimizing a Sum of Squared-Errors criterion (SSE, the minimization process is also commonly called Least-Squares Error, LSE). To stabilize the solution for estimation, a weight decay regularization can be incorporated. The criterion function to be minimized thus becomes:

$$J(\boldsymbol{\alpha}) = \frac{1}{2} \|\mathbf{y} - \mathbf{P}\boldsymbol{\alpha}\|_2^2 + \frac{b}{2} \|\boldsymbol{\alpha}\|_2^2, \quad (2)$$

where  $b$  controls the weighting of regularization factor,  $\mathbf{P} \in \mathbb{R}^{m \times K}$  packs the training features in matrix form, and  $\|\cdot\|_2$  denotes the  $L_2$ -norm. The estimated training output is given by  $\hat{\mathbf{y}} = \mathbf{P}\hat{\boldsymbol{\alpha}}$  where the solution for  $\boldsymbol{\alpha}$  which minimizes  $J$  is

$$\text{LSE :} \quad \hat{\boldsymbol{\alpha}} = (\mathbf{P}^T \mathbf{P} + b\mathbf{I})^{-1} \mathbf{P}^T \mathbf{y}, \quad (3)$$

with  $b$  being chosen to be a small value for stability and not introducing much bias.  $\mathbf{I}$  is an identity matrix of similar dimension to  $\mathbf{P}^T \mathbf{P}$ .

For an unseen test data sample  $\mathbf{x}_t$ , a new power series vector  $\mathbf{p}(\mathbf{x}_t)$  (row vector) can be generated. Prediction of the class label for this test data can be performed using the above learned  $\hat{\boldsymbol{\alpha}}$  (i.e.  $\hat{g}(\mathbf{x}_t) = \mathbf{p}(\mathbf{x}_t)\hat{\boldsymbol{\alpha}}$ ) and a threshold ( $\tau$ ) process:

$$\text{cls}(\hat{g}(\mathbf{x}_t)) = \begin{cases} 1 & \text{if } \hat{g}(\mathbf{x}_t) \geq \tau \\ 0 & \text{if } \hat{g}(\mathbf{x}_t) < \tau \end{cases}. \quad (4)$$

For multiple samples of test data, the multiple row vectors of  $\mathbf{p}(\mathbf{x}_t)$  can be stacked to form a matrix  $\mathbf{P}_t$  and predictions can be made collectively using  $\mathbf{P}_t \hat{\boldsymbol{\alpha}}$ .

## 2.2 Classification Error Rates

The above LSE formulation is a simple and effective method for classifier learning. However, LSE fits the training data according to the distribution density instead of the desired classification error, and this frequently results in overfitting. In this section, we introduce a deterministic error counting formulation to overcome this problem. We shall discuss a conventional method using nonlinear function approximation prior to presentation of a novel formulation which can be solved in a single step.

Consider the prediction and decision process in (4). For each operational setting of  $\tau$ , a *False Acceptance Rate* (FAR) and a *False Rejection Rate* (FRR) can be defined. Without loss of generality, consider the case where the scores of genuine-users are distributed at higher values than those of imposters. By varying the threshold  $\tau$  from  $-\infty$  to  $+\infty$  (or from 0 to 1 in a normalized case), the FRR shows an increasing trend while the FAR shows a decreasing trend, both with respect to this change of  $\tau$ . Along the variation of threshold  $\tau$ , there is a point (say, at  $\tau^*$ ) where the two curves (FAR and FRR) cross each other, and this point is called the *Equal Error Rate* (EER).

The *Total Error Rate* is defined as the sum of the False Acceptance and the False Rejection rates:  $TER = FAR + FRR$ . The EER mentioned above is frequently used as a performance index for biometric systems because at this particular operating point, the TER is frequently found to be near its minimum. This is particularly true when the genuine-user and imposter score distributions are normal. As such, the EER is frequently approximated by  $TER/2$  at  $\tau^*$  [17] and minimization of EER may be treated as minimization of the *minimum* TER. We shall minimize the empirical TER and observe its impact on the observed EER in this development.

Denote the genuine-user and imposter examples of variables  $(\mathbf{x}, m)$  by superscripts  $+$  and  $-$  respectively, it is not difficult to see that the FAR and FRR are merely the averaged counts of decision scores falling within the opposite pattern categories:

$$FAR = \frac{1}{m^-} \sum_{j=1}^{m^-} 1_{g(\mathbf{x}_j^-) \geq \tau} \quad , \quad (5)$$

$$FRR = \frac{1}{m^+} \sum_{i=1}^{m^+} 1_{g(\mathbf{x}_i^+) < \tau} \quad , \quad (6)$$

where the term  $1_{g(\mathbf{x}) \geq \tau}$  ( $1_{g(\mathbf{x}) < \tau}$ ) corresponds to a ‘1’ whenever  $g(\mathbf{x}) \geq \tau$  ( $g(\mathbf{x}) < \tau$ ), and ‘0’ otherwise.

### 2.3 Solving the Error Counting Problem

To solve the TER problem from (5)-(6), an approximation to the non-differentiable counting or step function is often adopted. A natural choice to approximate the above step function is the sigmoid function [18] where the minimization problem becomes  $\arg \min_{\alpha} TER(\alpha)$ :

$$TER(\alpha) = \frac{1}{m^-} \sum_{j=1}^{m^-} \sigma(\epsilon(\alpha, \mathbf{x}_j^-)) + \frac{1}{m^+} \sum_{i=1}^{m^+} \sigma(\epsilon(\alpha, \mathbf{x}_i^+)) \quad , \quad (7)$$

where

$$\sigma(x) = \frac{1}{1 + e^{-\gamma x}} \quad , \quad \gamma > 0 \quad , \quad (8)$$

and  $\epsilon(\alpha, \mathbf{x}_j^-) = g(\alpha, \mathbf{x}_j^-) - \tau$  for  $j = 1, 2, \dots, m^-$  and  $\epsilon(\alpha, \mathbf{x}_i^+) = \tau - g(\alpha, \mathbf{x}_i^+)$  for  $i = 1, 2, \dots, m^+$  are the error counts.

There are two problems associated with this approximation. The first problem is that the formulation is nonlinear with respect to the learning parameters. Although an iterative search can be employed for local solutions, different initializations may end up with different local solutions, hence incurring laborious trial and error efforts to select an appropriate setting. The second problem is that the objective function could be ill-conditioned due to the much local plateaus resulting from summing the flat regions of the sigmoid. A lot of search effort may be spent upon making little progress at locally flat regions.

Notice that the power series model is linear with respect to its parameters  $\alpha$ . We then seek in this section a deterministic single-step solution from matching of the link-loss functional pair [19,20]. When we have all inputs normalized within  $[0, 1]$ , the step functional can be approximated by centering a quadratic functional at the origin. With this idea in mind, the following regularized quadratic TER approximation is proposed:

$$\text{TER} = \frac{b}{2} \|\alpha\|_2^2 + \frac{1}{2m^-} \sum_{j=1}^{m^-} [\epsilon(\alpha, \mathbf{x}_j^-) + \eta]^2 + \frac{1}{2m^+} \sum_{i=1}^{m^+} [\varepsilon(\alpha, \mathbf{x}_i^+) + \eta]^2. \quad (9)$$

where  $\eta > 0$  and  $\epsilon(\alpha, \mathbf{x}_j^-) = \mathbf{p}(\mathbf{x}_j^-)\alpha - \tau$ ,  $\varepsilon(\alpha, \mathbf{x}_i^+) = \tau - \mathbf{p}(\mathbf{x}_i^+)\alpha$ , for  $j = 1, 2, \dots, m^-$ ,  $i = 1, 2, \dots, m^+$ .

Abbreviating the row power series vectors  $\mathbf{p}_j = \mathbf{p}(\mathbf{x}_j^-) \in \mathbb{R}^K$  and  $\mathbf{p}_i = \mathbf{p}(\mathbf{x}_i^+) \in \mathbb{R}^K$ , the solution for  $\alpha$  which minimizes (9) can be written as

$$\alpha = \left[ b\mathbf{I} + \frac{1}{m^-} \sum_{j=1}^{m^-} \mathbf{p}_j^T \mathbf{p}_j + \frac{1}{m^+} \sum_{i=1}^{m^+} \mathbf{p}_i^T \mathbf{p}_i \right]^{-1} \left[ \frac{(\tau - \eta)}{m^-} \sum_{j=1}^{m^-} \mathbf{p}_j^T + \frac{(\tau + \eta)}{m^+} \sum_{i=1}^{m^+} \mathbf{p}_i^T \right], \quad (10)$$

where  $\mathbf{I}$  is an identity matrix of  $K \times K$  size. The learned  $\alpha$  can then be used to predict new test samples in a similar manner as the procedure described in section 2.1.

**Remark 1:** Here we see that (10) consists of class-specific normalization of the input covariates, and this result can be considered as a specialized setting of the more generic weight least squares. It is noted that the solution given by (10) is deterministic and single-step without the need of initialization. This is a clear advantage over those conventional formulations such as (7) which require an iterative search.  $\square$

### 3 Face Verification

#### 3.1 Visual Face Images

The visual face images used in this study were captured under various illumination and pose conditions using a Bumblebee CCD camera produced by Point Grey Research Inc. (see [6] for details). The resolution of the image used was  $320 \times 240$  pixels. The top row of Fig. 1 shows some visual image samples for an identity under various illumination and pose conditions. In this work, we adopted the holistic approach using PCA since it is among the most common choices due to its simplicity and effectiveness. To compare similarity between two face images, the Euclidean distance was used for the first 100 eigenvalues [6].



**Fig. 1.** **Top row:** visual face samples for an identity under different lighting, expression and pose conditions; **Bottom row:** infra-red face samples for the same identity under different lighting, expression and pose conditions

### 3.2 Infra-Red Face Images

The Infra-Red face images used in this study were captured using a ThermoVision S65 produced by FLIR Systems Inc. As in the visual face case, the images were captured under varying illumination and pose conditions with the resolution of the image being fixed at  $320 \times 240$  pixels. The bottom row of Fig. 1 shows some infra-red image samples for the same identity under various conditions. Similar to the visual face, we adopted the holistic approach using PCA for the infra-red face. To compare similarity between two face images, the Euclidean distance was used for the first 100 eigenvalues [6].

In the following experiments, each data set corresponding to visual face (Face-VS) and infra-red face (Face-IR) verification consists of 96 identities, wherein each identity contains 10 image samples. For training and test purposes, each of these biometric data sets are partitioned into two equal sets consisting of  $\mathcal{S}_{train}$  and  $\mathcal{S}_{test}$ , each with  $96 \times 5$  samples. The genuine-user and the imposter match-scores are generated from these two sets by intra-identity and inter-identity matching among the image samples for each biometric. A total of 960 sample match-scores are thus available for the genuine-user class in each training set and test set for each biometric. As for the imposter scores, there are 114000 sample match-scores for the 96 identities. Since both face systems captured multiple images from the same person for each identity, the fusion constitutes a true multi-modal fusion.

### 3.3 Fusion of Face-VS and Face-IR scores

Based on the biometrics data described above, we perform fusion of the visual and infra-red face verification scores in the following experiments. Our main consideration is that ground applications encounter large variation of illumination

conditions, and we believe that visual and infra-red images can complement each other. More importantly, these two modalities can be captured simultaneously and in a covert manner in practical scenarios.

In this experiment, we perform fusion experiments by combining the verification decisions from Face-VS and Face-IR where the images were captured simultaneously. Since the Face-VS has poor performance due to large variation of illumination conditions in the database, we shall observe the effects of fusing it with the much higher performed Face-IR. The following summarizes those fusion methods compared in our experiments:

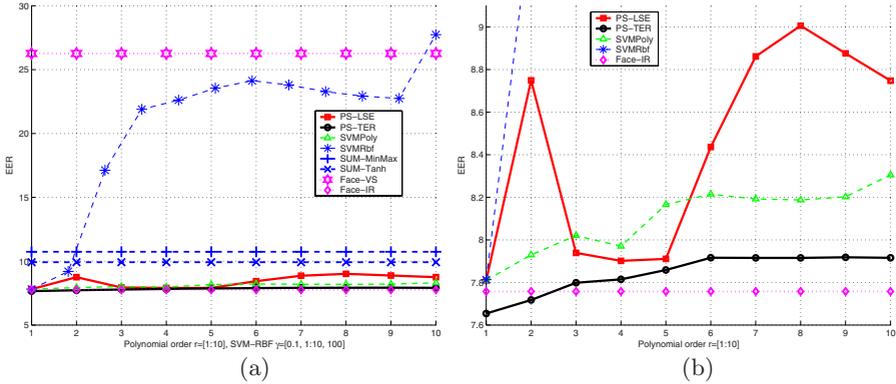
- i. Proposed PS-TER and PS-LSE: the proposed PS-TER (Power Series model trained by TER solution given by (10)) and PS-LSE (Power Series model trained by LSE solution given by (3)) are adopted for fusion of the above two face modalities. The power series are experimented with different model orders to observe the stability of test performance.
- ii. Two simple SUM methods: the simple SUM rule was reported to be among the best fusion methods for multi-biometrics when the inputs are normalized appropriately (see e.g. [21]). For benchmarking purpose, we include in our experiments two important normalization methods namely the Min-Max and the robust hyperbolic tangent (*tanh*) method (see e.g. [21]) to precede the simple SUM rule.
- iii. SVM adopting different kernels: in addition to the SUM rules, the Support Vector Machines adopting different kernels (polynomial and Radial Basis Function) are also included. This is to cater for nonlinear decision hyper-surfaces as well as to benchmark with the well-acclaimed SVM in classification and biometrics fusion (see e.g. [22,23,12]). The SVMs are experimented with different model settings (polynomial order and Gaussian width) to observe the operational performance.

*Performance evaluation:* the EER shall be adopted as the performance comparison measure in all the forth coming experiments following [24]. There are also two reasons behind this choice of performance criterion: (i) it is a single value index which has a clear indication of high and low performances and this can be advantages to the use of ROC or DET where the curves for different algorithms may cross each other, and (ii) it is somewhat related to our optimization objective (minimization of TER).

## Results:

Fig. 2 shows the average test EER performances for various model settings from two-fold validations. For SVM-Poly and PS model trained by LSE and TER, the model order varies for  $r \in [1, 10]$ . For SVM-Rbf, the kernel width (*gamma* parameter in [25]) varies in  $[0.1, 1, \dots, 10, 100]$ . The DET plots for best performed settings of each method from the first-fold are shown in Fig. 3.

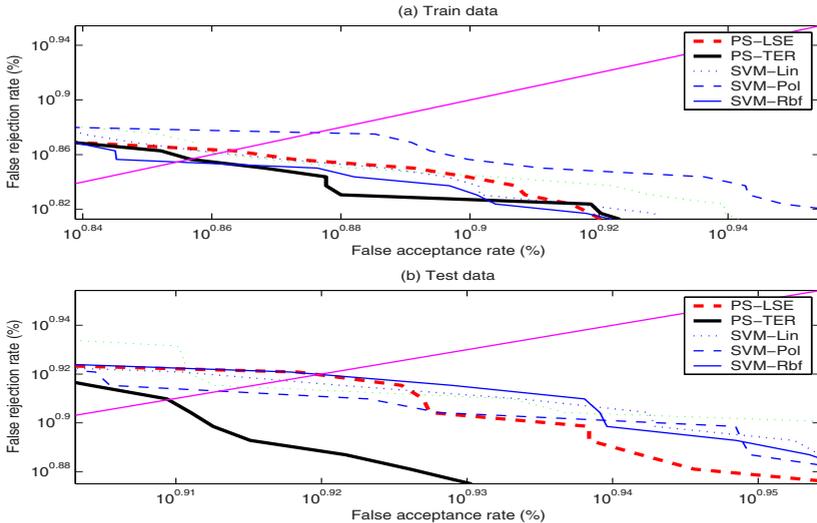
The results from Fig. 2(a) show that SVM-Rbf's performance deteriorates with small kernel widths (large *gamma* values). Both the fusion outputs from the SUM methods generated poorer performance than that from unimodal infra-red face.



**Fig. 2.** (a) Average test EER values plotted over different classifier settings, (b) a zoom-in view

A zoom-in view of Fig. 2(a) in Fig. 2(b) shows superiority of PS-TER over PS-LSE for all power series order settings experimented. Two settings ( $r = 1, 2$ ) of PS-TER show gain in fusion performance comparing to unimodal Face-IR (maximum improvement of 1.34% at  $r = 1$ , this can mean additional 1523 imposters being rejected based on the size of the database). For SVM-Poly and SVM-Rbf, no gain in performance is observed for all model settings.

The above results suggest that fusion of a moderate biometric with a weak one can have deteriorated performance when the fusion classifier is not well tuned. For the experimented data set, the proposed TER method shows good



**Fig. 3.** Detection Error Tradeoff (DET) curves

generalization property with respect to different model tunings as compared to LSE and SVM. This shows that useful information gain can still be acquired from the weaker biometric with appropriate fusion classifier design.

## 4 Conclusion

In this paper, we proposed to minimize an error rate formulation for face recognition. Unlike conventional learning and threshold processes to arrive at a decision error count rate, the proposed method directly optimized the error count rate with respect to a power series model design. The identity verification was performed using both visual and infra-red face image cues which constitutes a difficult scenario of fusing a moderately accurate biometric to a weak one. Our experiments show stability of fusion results thereby illustrating useful information gain can still be acquired from a weak biometric in fusion.

## Acknowledgements

The author would like to acknowledge Prof Anil. K. Jain (Michigan State University) for useful discussions regarding the research topic. The face data sets prepared by Mr Sangki Kim and Prof Sangyoun Lee are also acknowledged.

This work was supported by the Korea Science and Engineering Foundation (KOSEF) through the Biometrics Engineering Research Center (BERC) at Yonsei University.

## References

1. Li, S.Z., Jain, A.K. (eds.): Handbook of Face Recognition. Springer, New York (2004)
2. Chen, X., Flynn, P.J., Bowyer, K.W.: IR and visible light face recognition. *Computer Vision and Image Understanding* 99(3), 332–358 (2005)
3. Kong, S., Heo, J., Abidi, B., Paik, J., Abidi, M.: Recent advances in visual and infrared face recognition—a review. *Computer Vision and Image Understanding* 97(1), 103–135 (2005)
4. Wilder, J., Phillips, P.J., Jiang, C., Wiener, S.: Comparison of visible and infrared imagery for face recognition. In: Proc. 2nd International Conference on Automatic Face and Gesture Recognition, pp. 182–187 (1996)
5. Socolinsky, D.A., Selinger, A., Neuheisel, J.D.: Face recognition with visible and thermal infrared imagery. *Vision and Image Understanding* 91(1-2), 72–114 (2003)
6. Kim, S.-K., Lee, H., Yu, S., Lee, S.: Robust face recognition by fusion of visual and infrared cues. In: Proceedings of the First IEEE Conference on Industrial Electronics and Applications, Singapore, May 2006, pp. 804–808 (2006)
7. Duda, R.O., Hart, P.E., Stork, D.G.: *Pattern Classification*, 2nd edn. John Wiley & Sons, Inc. New York (2001)
8. Schürmann, J.: *Pattern Classification: A Unified View of Statistical and Neural Approaches*. John Wiley & Sons, Inc. New York (1996)

9. Kittler, J., Messer, K.: Fusion of multiple experts in multimodal biometric personal identity verification systems. In: Proceedings of the 2002 12th IEEE Workshop on Neural Networks for Signal Processing, pp. 3–12 (2002)
10. Kuncheva, L.I., Bezdek, J.C., Duin, R.: Decision templates for multiple classifier design: An experimental comparison. *Pattern Recognition* 34(2), 299–314 (2001)
11. Huang, Y.S., Suen, C.Y.: A method of combining multiple experts for the recognition of unconstrained handwritten numerals. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 17(1), 90–94 (1995)
12. Boser, B.E., Guyon, I.M., Vapnik, V.N.: A training algorithm for optimal margin classifier. In: Proceedings of the 5th ACM Workshop on Computational Learning Theory, Pittsburgh, PA, pp. 144–152 (1992)
13. Poggio, T., Girosi, F.: Networks for approximation and learning. *Proceedings of the IEEE* 78(9), 1481–1497 (1990)
14. Schölkopf, B., Smola, A.J.: *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, Cambridge, MA (2002)
15. Toh, K.-A.: Training a  $\phi$ -machine classifier using feature scaling-space. In: Proceedings of the 4th International IEEE Conference on Industrial Informatics, Singapore, August 2006, pp. 1334–1339 (2006)
16. Toh, K.-A., Tran, Q.-L., Srinivasan, D.: Benchmarking a reduced multivariate polynomial pattern classifier. *IEEE Trans. Pattern Analysis and Machine Intelligence* 26(6), 740–755 (2004)
17. Poh, N., Bengio, S.: How do correlation and variance of base-experts affect fusion in biometric authentication tasks? *IEEE Trans. on Signal Processing* 53(11), 4384–4396 (2005)
18. Toh, K.-A.: Learning from target knowledge approximation. In: Proceedings of the First IEEE Conference on Industrial Electronics and Applications, Singapore, May 2006, pp. 815–822 (2006)
19. Gordon, G.J.: Generalized<sup>2</sup> Linear<sup>2</sup> Models. In: *Advances in Neural Information Processing Systems (NIPS 2002)*, Vancouver, British Columbia, Canada, December 2002, pp. 577–584 (2002)
20. McCullagh, P., Nelder, J.A.: *Generalized Linear Models*, 2nd edn. Chapman and Hall, London (1989)
21. Jain, A.K., Nandakumar, K., Ross, A.: Score normalization in multimodal biometric systems. *Pattern Recognition* 38, 2270–2285 (2005)
22. Ben-Yacoub, S., Abdeljaoued, Y., Mayoraz, E.: Fusion of face and speech data for person identity verification. *IEEE Trans. on Neural Networks* 10(5), 1065–1074 (1999)
23. Vapnik, V.N.: *Statistical Learning Theory*. Wiley-Interscience Pub. Chichester (1998)
24. Maio, D., Maltoni, D., Cappelli, R., Wayman, J.L., Jain, A.K.: FVC2000: Fingerprint verification competition. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 24(3), 402–412 (2002)
25. Ma, J., Zhao, Y., Ahalt, S.: OSU SVM classifier Matlab toolbox (ver 3.00), the Ohio State University (2002), [[http://eewww.eng.ohio-state.edu/~maj/osu\\_svm/](http://eewww.eng.ohio-state.edu/~maj/osu_svm/)]