2010 IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP 2010

March 14-19, 2010 • Dallas, Texas, U.S.A.

IEEE Catalog Number: CFP10ICA-CDR
ISBN: 978-1-4244-4296-6 ISSN: 1520-6149

RADAR HRRP STATISTICAL RECOGNITION WITH LOCAL FACTOR ANALYSIS BY AUTOMATIC BAYESIAN YING YANG HARMONY LEARNING

Lei Shi[1], *Penghui Wang*[2], *Hongwei Liu*[2], *Lei Xu*[1*], *and Zheng Bao*[2]

[1] Department of Computer Sci. & Eng., The Chinese University of Hong Kong, Shatin, NT, Hong Kong
[2] National Laboratory of Radar Signal Processing, Xidian University, Xi'an, P. R. China

ABSTRACT

Radar high-resolution range profiles (HRRPs) are typical high-dimensional **non-Gaussian** and **inter-dimensional dependently** distributed data, the statistical modelling of which is a challenging task for HRRP based target recognition. Considering the inter-dimensional dependence, a recent work [1] applied Factor Analysis (FA) to model radar HRRP data and showed promising recognition results, which however still restricts to Gaussian distribution. This paper aims to simultaneously consider the inter-dimensional dependence and the non-Gaussian distribution, by using Local Factor Analysis (LFA) model. For not only learning parameters but also appropriately selecting the component number and local hidden dimensionalities, we adopt the automatic Bayesian Ying-Yang (BYY) harmony learning, in order to relieve the extensive computation and inaccurate evaluation encountered in the conventional two-phase implementation. Moreover, a heuristic aspect-frame partition is implemented based on the BYY harmony criterion rather than AIC or BIC in the previous work, to tackle the radar HRRP's target-aspect sensitivity. Experiments show improved recognition performances over [1] on the same measured HRRP dataset, i.e., for both equal interval and heuristic aspect-frame partitions, LFA automatically learned by BYY always outperforms FA selected by a two-phase procedure with either AIC or BIC.

***Index Terms*—** HRRP, sensitivity, non-Gaussian, inter-dimensional dependence, automatic model selection, BYY harmony learning, heuristic aspect-frame partition.

1. INTRODUCTION

Radar automatic target recognition (RATR) is to identify the unknown target from its radar-echoed signatures. A high-resolution range profile (HRRP) is the coherent summation amplitudes of the complex time returns from target scatterers in several range cells and along the radar line of sight (LOS), as illustrated in Fig. 1. The HRRP signal contains target structure information such as target size and scattering distribution, radar target recognition based on which has received intensive attention from the RATR community [2, 1, 3, 4, 5]. Bayesian classifier has been widely-used for this typical pattern recognition task, and thus the statistical modelling of the HRRP distribution becomes a main problem [1, 3]. Since nonparametric methods significantly suffer from the curse of dimensionality in this application [6, 7], parametric methods are usually preferred [1, 2], for which this paper follows still.

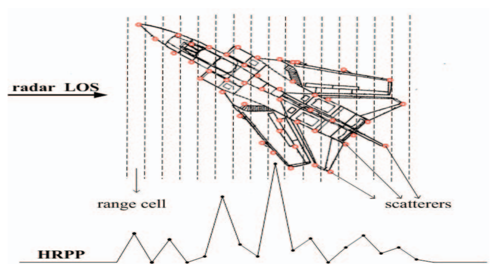


Fig. 1. Radar returns from the scatterers on the target are projected onto the LOS, resulting in an HRRP.

Radar HRRP data are typically *high-dimensional*, *non-Gaussian*, and *inter-dimensional dependently* distributed, the parametric modelling of which is a challenging task for HRRP based target recognition [2, 1, 3, 4]. In the literature, many efforts have been made on two important topics. The first is how to determine a family of parametric model that can appropriately describe HRRP data distribution. Then the second is how to appropriately learn the model based on a finite size of HRRP training samples.

On selecting an appropriate parametric model, several early efforts [4] assumed the range cells (dimensions) in HRRP are independently Gaussian distributed, while this assumption was later found inappropriate from physical arguments and results of empirical investigations [2, 1, 3]. There are mainly two types of efforts for improvement. On one hand, papers [2, 3] extended the distribution from Gaussian to non-Gaussian, while the independence is still assumed among different dimensions. On the other hand, Du et. al. [1] considered dependence among dimensions by Factor Analysis (FA) model, which is still in a Gaussian distribution. None of these efforts considers non-Gaussian and inter-dimensional dependence simultaneously. This paper is thus motivated to consider both by Local Factor Analysis (LFA) model [8, 9].

Once a family of parametric models is chosen, the learning task consists of parameter learning for determining un-

*Corresponding author: Lei Xu. Email: lxu@cse.cuhk.edu.hk. The work described in this paper was supported by RGC Direct grant project 2050427.

known parameters given a model scale $\mathbf{k}$, and model selection for choosing one among a family of models with an appropriate scale $\mathbf{k}$ [10, 9]. For an LFA model, this $\mathbf{k}$ consists of the number of components and the local hidden dimensionalities. One widely used method for parameter estimation is the maximum-likelihood learning, usually implemented by the expectation-maximization (EM) algorithm [11]. Model selection is conventionally implemented via a two-phase procedure in help of a typical model selection criterion, such as Akaike's Information Criterion (AIC) and Bayesian Inference Criterion (BIC) [10, 1]. For high-dimensional data, this implementation inevitably suffers from two problems, namely, an extensive computation and unreliable evaluation to the criterion [9, 10], so that it is costly or even impractical for radar HRRP modeling. To tackle these problems, an automatic Bayesian Ying-Yang (BYY) harmony learning is thus adopted in this paper for LFA learning, which conducts model selection automatically during parameter learning [9].

In addition to the statistical modelling, there are three sensitivity problems in HRRP data, including translation, amplitude-scale and target-aspect sensitivities [3, 1]. The former two could be commonly tackled by translation alignment and amplitude-scale normalization, respectively. The last was tackled by the aspect-frame based modelling in [1], i.e., partitioning consecutive HRRP samples into several aspect-frames and then modelling different frames separately. In this paper, besides an equal interval partition, a heuristic aspect-frame partition mechanism is employed under the BYY harmony model selection criterion.

In help of automatic BYY harmony learning on the LFA model and aspect-frame partition based on BYY harmony criterion, the classification on the same dataset as [1] shows improved recognition performances. The rest of this paper is organized as follows. In Section 2, we introduce the LFA model and an automatic BYY harmony learning algorithm, together with a harmony criterion for aspect-frame partition. Section 3 reports the improved recognition performance on the same radar HRRP dataset as in [1]. Finally, some concluding remarks are drawn in Section 4.

2. LFA MODEL AND BYY HARMONY LEARNING

2.1. LFA Model

LFA, or also called Mixture of Factor Analyzers (MFA) [8, 9], performs clustering and local dimension reduction simultaneously. Considering a d-dimensional observable variable $\mathbf{x}$, LFA assumes that $\mathbf{x}$ is distributed according to a mixture of k underlying components, i.e., $p(\mathbf{x}) = \sum_{l=1}^{k} \alpha_l p(\mathbf{x}|l)$, with $\alpha_l \geq 0$ and $\sum_{l=1}^{k} \alpha_l = 1$. Each component $p(\mathbf{x}|l)$ is described by a single FA as follows:

$$p(\mathbf{x}|\mathbf{y}, l) = G(\mathbf{x}|\mathbf{U}_l\mathbf{y} + \boldsymbol{\mu}_l, \boldsymbol{\Psi}_l), \quad p(\mathbf{y}|l) = G(\mathbf{y}|\mathbf{0}, \boldsymbol{\Lambda}_l),$$
$$p(\mathbf{x}|l) = G(\mathbf{x}|\boldsymbol{\mu}_l, \mathbf{U}_l\boldsymbol{\Lambda}_l\mathbf{U}_l^T + \boldsymbol{\Psi}_l), \quad \text{s.t. } \mathbf{U}_l^T\mathbf{U}_l = \mathbf{I}_{m_l}. \qquad (1)$$

In each component l, α_l is the mixing weight, m_l is the hidden dimensionality, $\mathbf{U}_l$ is a $d \times m_l$ loading matrix constrained on the Stiefel manifold $\mathbf{U}_l^T\mathbf{U}_l = \mathbf{I}_{m_l}$, $\boldsymbol{\Lambda}_l$ is an $m_l \times m_l$-dimensional diagonal covariance, $\boldsymbol{\mu}_l$ is a d-dimensional mean vector, $\boldsymbol{\Psi}_l$ is a diagonal noise covariance. As a whole, LFA describes a *non-Gaussian* distribution by a mixture of Gaussian components, each of which is equipped with the inter-dimensional *dependence*. It should be noted that Eq. (1) is a parametrization different from the commonly used form [1, 8]. Although they are equivalent in maximum-likelihood learning, the formulation by Eq. (1) has shown advantages in the level of model selection, details of which are referred to another upcoming paper [12].

Model selection for LFA is to determine the component number and the local hidden dimensionalities appropriately based on a finite size of samples. Due to the problems of the two-phase implementation, an automatic BYY harmony learning is employed and will be introduced in the following.

2.2. Automatic BYY Harmony Learning on LFA

Firstly proposed in 1995 [13] and then systematically developed over a decade [9, 14, 15, 16], BYY harmony learning provides a general statistical learning framework for parameter learning and model selection. BYY harmony learning with typical structures leads to a set of new model selection criteria, new techniques for implementing regularization and a class of algorithms with automatic model selection ability during parameter learning.

Considered based on the observation $\mathbf{X}$ and its inner representation $\mathbf{R} = \{\mathbf{Y}, \boldsymbol{\Theta}\}$, where a parameter set $\boldsymbol{\Theta}$ collectively represents the underlying structure of $\mathbf{X}$, and $\mathbf{Y}$ correspondingly is the inner representation of $\mathbf{X}$, $p(\mathbf{X}, \mathbf{R}) = p(\mathbf{R}|\mathbf{X})p(\mathbf{X})$ and $q(\mathbf{X}, \mathbf{R}) = q(\mathbf{X}|\mathbf{R})q(\mathbf{R})$ form two types of decomposition called Yang machine and Ying machine, respectively. Such a Ying-Yang pair is called a BYY system. The harmony measure based on this system has been systematically studied in the general form restated in Eq. (2), which is different from the likelihood function. An important nature is that maximizing it leads to not only a best matching between the Ying-Yang pair, but also a compact model with a least complexity. Such an ability has been investigated from several perspectives [9, 14, 15, 16].

Applied on LFA model described in Eq. (1), observation $\mathbf{x}$ is assumed i.i.d. generated from the hidden representation $\mathbf{Y} = \{\mathbf{y}, l\}$, and thus the components in Ying machine $q(l)q(\mathbf{y}|l)q(\mathbf{x}|\mathbf{y}, l)$ are determined accordingly [9]. In Yang machine, a Parzen window smoothed density $p_h(\mathbf{x})$ is considered given a dataset $\{\mathbf{x}_t\}_{t=1}^{N}$, i.e. $p_h(\mathbf{x}) = \frac{1}{N}\sum_{t=1}^{N} G(\mathbf{x}|\mathbf{x}_t, h^2\mathbf{I}_d)$. According to the "uncertainty conversation principle", one choice of Yang-pathway consists of $p(l|\mathbf{x}) = \frac{\alpha_l q(\mathbf{x}|l)}{\sum_j \alpha_j q(\mathbf{x}|j)}$ and $p(\mathbf{y}|\mathbf{x}, l) = G(\mathbf{y}|\widetilde{\mathbf{W}}_l(\mathbf{x} - \boldsymbol{\mu}_l), \boldsymbol{\Gamma}_l)$, where $\boldsymbol{\Gamma}_l = [\frac{\partial^2 \ln[q(\mathbf{x}|\mathbf{y},l)q(\mathbf{y},l)]}{\partial\mathbf{y}\partial\mathbf{y}^T}]^{-1} = (\mathbf{U}_l^T\boldsymbol{\Psi}_l^{-1}\mathbf{U}_l + \boldsymbol{\Lambda}_l^{-1})^{-1}$, and $\widetilde{\mathbf{W}}_l$ is a Yang machine parameter. For parameter distributions, the prior is set as $q(\boldsymbol{\Theta}) = q(h)$ and $p(\boldsymbol{\Theta}|\mathbf{X})$ is considered as a free structure. After substituting each part into Eq. (2) and simplification, we arrive at the specific harmony

$$\text{General form:} \quad H(p||q) = \int p(\mathbf{R}|\mathbf{X})p(\mathbf{X}) \ln[q(\mathbf{X}|\mathbf{R})q(\mathbf{R})]d\mathbf{X}d\mathbf{R}. \tag{2}$$

$$\text{Typically on LFA:} \quad H(p||q,\mathbf{\Theta}) = H_f(p||q,\mathbf{\Theta}) - Z(h), \qquad Z(h) = -\ln q(h) = \ln[\frac{1}{N}\sum_{t=1}^{N}\sum_{\tau=1}^{N} G(\mathbf{x}_t|\mathbf{x}_\tau, h^2\mathbf{I}_d)],$$

$$H_f(p||q,\mathbf{\Theta}) = \frac{1}{2N}\sum_{t=1}^{N}\sum_{l=1}^{k} p(l|\mathbf{x}_t)\{2\ln\alpha_l - (d+m_l)\ln(2\pi) - m_l - \ln|\mathbf{\Lambda}_l| - \ln|\mathbf{\Psi}_l|$$
$$-Tr[[\widetilde{\mathbf{W}}_l^T\mathbf{\Lambda}_l^{-1}\widetilde{\mathbf{W}}_l + (\mathbf{I}_d - \mathbf{U}_l\widetilde{\mathbf{W}}_l)^T\mathbf{\Psi}_l^{-1}(\mathbf{I}_d - \mathbf{U}_l\widetilde{\mathbf{W}}_l)][(\mathbf{x}_t-\boldsymbol{\mu}_l)(\mathbf{x}_t-\boldsymbol{\mu}_l)^T + h^2\mathbf{I}_d]]\}. \tag{3}$$

measure $H(p||q,\mathbf{\Theta})$ for LFA in Eq. (3), where the regularization term $Z(h)$ helps adjust h to an appropriate value during learning, coming from an Induced Bias Cancelation (IBC) improper prior on h [9].

To maximize $H(p||q,\mathbf{\Theta})$ in Eq. (3) w.r.t. parameters $\mathbf{\Theta} = \{\{\alpha_l, \boldsymbol{\mu}_l, \mathbf{\Psi}_l, \mathbf{U}_l, \mathbf{\Lambda}_l, \widetilde{\mathbf{W}}_l\}_{l=1}^k, h\}$, an adaptive learning algorithm is referred to Figure 8 in [9] and Section 2.3 in [16]. Maximizing the harmony measure in Eq. (3) will provide an intrinsic force to push $\alpha_l \to 0$ if a component l is extra and thus can be discarded. Also, the variance $\mathbf{\Lambda}_l^{(j)}$ of a component l will be forced to approach zero if the corresponding local dimension $\mathbf{y}_l^{(j)}$ is extra and thus discarded. As long as the component number k and the local hidden dimensions $\{m_l\}_{l=1}^k$ are initialized at large enough values, model selection will be conducted automatically during parameter learning. More algorithms are referred to Algorithm III in [15], where a unified procedure is given for a number of different settings.

2.3. Harmony Criterion based Aspect-Frame Partition

One remaining but important problem for radar HRRP recognition is the target-aspect sensitivity, i.e., the distribution of HRRPs varies with the target aspect angle. Two kinds of methods are usually used to tackle this problem. The first uses a mixture model to describe the distribution of the total samples based on all aspect angles, which however is usually so complex that its learning suffers from great computational complexity [1]. The second method adopts the divide-and-conquer policy and models different sectors separately [3, 1], with each sector defined as an aspect-frame. Specifically, after HRRPs of a target c are divided into K_c aspect-frames, a parametric model is used to describe the distribution $p_c(\mathbf{x}|j)$ in each frame j. Then during classification, the class-conditional probability $p_c(\mathbf{x}_t)$ for class c is calculated as $p_c(\mathbf{x}_t) \triangleq \max_j[p_c(\mathbf{x}_t|j)p_c(j)]$, where $p_c(j)$ is a prior of the j-th aspect-frame in target c, which could be assigned proportionally to its sample size in this frame. Due to its efficiency and effectiveness, this method has been investigated in the literature [3, 1] and is adopted in this paper.

Therein, how to partition the consecutive training samples into aspect-frames becomes a key problem. As a typical combinatorial optimization problem, both the frame number and the frame partitions need to be determined. We adopt the heuristic partition mechanism proposed in Section III.B of [1] for comparison consistency, whose implementation consists of two nested levels. The inner level determines the model in a frame, and the outer level determines frames sequentially in a two-phase procedure.

In addition to the nature of automatic model selection by $\max_{\mathbf{\Theta}} H(p||q,\mathbf{\Theta})$, Eq. (2) in a two stage implementation also consists of seeking the minimum of the following BYY harmony criterion [9]:

$$J_{BYY}(\mathbf{k}) = -H(p||q,\mathbf{\Theta}) + \frac{1}{2N}D(\mathbf{k}), \tag{4}$$

where $D(\mathbf{k}) = 2kd + k - 1 + \sum_{l=1}^{k} m_l(d - \frac{m_l-1}{2})$ is the number of free parameters in an LFA model with a scale $\mathbf{k} = \{k, \{m_l\}_{l=1}^k\}$. Although based on a mechanism similar to that in [1], our partition procedure has two key differences. In the inner level, an LFA model is automatically learned by BYY, instead of an FA model selected by a two-phase procedure. In the outer level, the above BYY harmony criterion $J_{BYY}(\mathbf{k})$ replaces either AIC or BIC in [1].

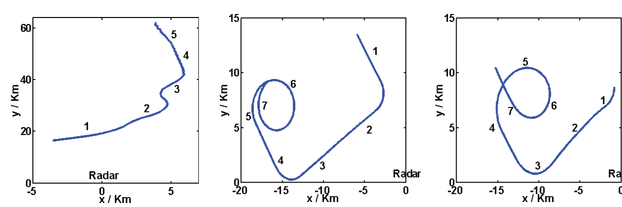
3. EXPERIMENTAL RESULTS

Experimental investigation is conducted on the 3-class measured HRRP dataset same as in [1], including *Yak-42*, *Cessna Citation S/II*, and *An-26*. The radar signal is of 400MHz bandwidth, and the dimensionality of the HRRP data, i.e., the number of range cells, is $d = 256$, with the plane targets' parameters listed in Table 1. The projections of target trajectories onto ground plane are shown in Fig. 2, where the measured data are segmented into training and testing data [1]. Particularly, the 2nd and the 5th segments of *Yak-42*, the 6th and the 7th segments of *Cessna Citation S/II*, the 5th and the 6th segments of *An-26* are taken as the training samples, with the remaining left for testing.

During dividing aspect-frames, we implement both equal interval partition and heuristic partition approaches. For the equal interval partition, same as in [1], we set 35 aspect-frames for *Yak-42*, 50 for *Cessna*, and 50 for *An-26*, with

Table 1. Description of the plane parameters.

plane	length (m)	width (m)	height (m)
Yak-42	36.38	34.88	9.83
Cessna Citation S/II	14.40	15.90	4.57
An-26	23.80	29.20	9.83

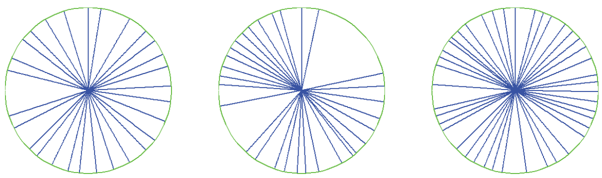


(a) *Yak-42* (b) *Cessna* (c) *An-26*

Fig. 2. Projections of target trajectories onto ground plane.

Table 2. The confusion matrices and average correct recognition rates ($ACRR$) by LFA classifier learned with BYY. For short, Y, C, and A stand for the three planes, respectively. Results labeled by (*) are the $ACRR$ collected from [1] for comparison.

	Equal Interval Partition						Heuristic Partition					
	no smoothing			smoothing			no smoothing			smoothing		
Plane	Y	C	A	Y	C	A	Y	C	A	Y	C	A
Y	100	0.80	3.30	100	0.45	2.70	100	0.15	1.85	100	0.10	1.75
C	0	97.15	0.45	0	98.00	0.45	0	99.20	0.35	0	99.35	0.25
A	0	1.95	96.25	0	1.55	96.75	0	0.65	97.80	0	0.55	98.00
ACRR	97.80			**98.25**			99.00			**99.12**		
(*)FA-AIC	94.50						98.53					
(*)FA-BIC	96.78						98.73					



(a) *Yak-42* (with $K_Y = 28$) (b) *Cessna* (with $K_C = 30$) (c) *An-26* (with $K_A = 42$)

Fig. 3. The heuristically partitioned aspect-frames. Each separated sector refers to an aspect-frame, with the angular size proportional to its sample size. The notation K_i represents the number of resulted aspect-frames in class i.

1024 HRRP samples in each frame. Based on samples in each aspect-frame of each plane, LFA learning by BYY is implemented, with and without data smoothing. There are thus four different implementations in total. A classifier is composed once all aspect-frames have been determined and each frame gets a learned LFA model. The recognition accuracy on testing samples by each implementation is reported in Table 2. The results by FA selected with AIC and BIC are collected from [1] for comparison. Fig. 3 shows the sample size proportions of aspect-frames obtained by the heuristic partition.

These results indicate the following observations. First, compared to FA-AIC and FA-BIC in [1], the results by LFA show improved recognition performances. For the equal interval partition, LFA's best result relatively improves FA-AIC and FA-BIC by 3.97% and 1.52%, respectively. For the heuristic partition, LFA's best result relatively outperforms them by 0.60% and 0.40%, respectively. Additionally, the heuristic partition always outperforms the equal interval partition, which reconfirms the observations in [1]. Second, the results with data smoothing outperform those without data smoothing by 1.04% in average. Third, interestingly, there are $(28, 30, 42)$ heuristically partitioned frames for the three planes respectively, which are all smaller than $(32, 39, 48)$ by FA-AIC and $(31, 35, 47)$ by FA-BIC in [1]. It should be noted that, these observations come from a combination of several effects, including the description ability of LFA model, the learning performance by BYY-A, and the aspect-frame partition appropriateness, etc.

4. CONCLUDING REMARKS

As a further investigation of a recent work [1] on radar HRRP target recognition, this paper considers LFA for describing not only the inter-dimensional dependence but also the non-Gaussian distribution of HRRPs. To determine an appropriate model scale for LFA, this work adopts the automatic BYY harmony learning. Furthermore, the BYY harmony criterion is employed for the heuristic aspect-frame partition to tackle the target-aspect sensitivity. Recognition results on the same HRRP dataset as [1] show promising improvement.

5. REFERENCES

[1] L. Du, H.-W. Liu, and Z. Bao, "Radar HRRP Statistical Recognition: Parametric Model and Model Selection," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1931–1944, 2008.

[2] K. Copsey and A. R. Webb, "Bayesian Gamma Mixture Model Approach to Radar Target Recognition," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1201–1217, October 2003.

[3] L. Du, H.-W. Liu, Z. Bao, and J.-Y. Zhang, "A two-distribution compounded statistical model for radar HRRP target recognition," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2226–2238, June 2006.

[4] H.-J. Li and S.-H. Yang, "Using range profiles as feature vectors to identify aerospace objects," *IEEE Trans. Antennas Propag.*, vol. 41, no. 3, pp. 261–268, 1993.

[5] A. Zyweck and R. E. Bogner, "Radar target classification of commercial aircraft," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 3, pp. 598–606, 1996.

[6] E. Parzen, "On the estimation of a probability density function and mode," *Annals of Math. Stat.*, vol. 33, pp. 1065–1076, 1962.

[7] A. R. Webb, *Statistical Pattern Recognition, 2nd Edition*, John Wiley & Sons, October 2002.

[8] Geoffrey E. Hinton, Michael Revow, and Peter Dayan, "Recognizing handwritten digits using mixtures of linear models.," in *Advances in NIPS*, 1994, pp. 1015–1022.

[9] L. Xu, "Bayesian Ying Yang Learning," in *Scholarpedia 2(3):1809, http://scholarpedia.org/article/Bayesian_Ying_ Yang_Learning*, 2007.

[10] Kenneth P. Burnham and David Anderson, *Model Selection and Multi-Model Inference*, Springer, July 2002.

[11] R. Redner and H. Walker, "Mixture densities, maximum likelihood and the EM algorithm," *SIAM Review*, vol. 26, no. 2, pp. 195–239, 1984.

[12] S. Tu and L. Xu, "On the two parameterizations of factor analysis: which one is better?," *(in preparation)*, 2009.

[13] L. Xu, "Bayesian-Kullback coupled Ying-Yang machines: Unified learnings and new results on vector quantization," in *Proc. of ICONIP'95*, Beijing, China, 1995, pp. 977–988.

[14] L. Xu, "Trends on Regularization and Model Selection in Statistical Learning: A Perspective from Bayesian Ying Yang Learning," in *Studies in Computational Intelligence*, W. Duch, J. Mandziuk, and J.M. Zurada, Eds., pp. 365–406. Springer-Verlag, 2007.

[15] L. Xu, "Learning Algorithms for RBF Functions and Subspace Based Functions," in *Ch. 3 in Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods and Techniques*, Olivas, Guerrero, Sober, Benedito, and Lopez, Eds., pp. 60–94. IGI Global publiation, 2009.

[16] L. Xu, "Bayesian Ying-Yang System, Best Harmony Learning, and Five Action Circling," in *an invited special issue on Emerging Themes on Information Theory and Bayesian Approach, Frontiers of Electrical and Electronic Engineering in China, a journal jointly published by Higher Education Press of China and Springer (to appear)*. 2010.