http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 14 (1): 16-23, 2015
ISSN 1812-5638

RESEARCH ARTICLE OPEN ACCESS

DOI: 10.3923/itj.2015.16.23

Feature Fusion Using Automatic Generated RBF Neural Network

[1]ZhiWen Wang and [2]GongKun Luo
[1]College of Computer Science and Communication Engineering,
[2]College of Electrical and Information Engineering, Guangxi University of Science and Technology, Liuzhou, China

ARTICLE INFO
Article History:
Received: April 13, 2014
Accepted: January 20, 2015

Corresponding Author:
ZhiWen Wang
College of Computer Science and Communication Engineering, Guangxi University of Science and Technology, 268 Dong Huan Road, Liuzhou, Guangxi, 545006, China

ABSTRACT
In this study, a strategy for feature fusion for team behaviors recognition using automatically generated RBF neural network is proposed, for various features need to be extracted in the course of team behaviors recognition and it is difficult to estimate the contribution of various features for identifying and decrypting team behaviors. The burden of high-level recognition algorithm is use eased by using the underlying features of moving target, such as the trajectory characteristics extracted by using the method of trajectory growth. KPCA algorithm is used to nonlinearly reduce dimensionality of extracted characteristics before feature fusion for extracted characteristics of team behaviors. Automatically generated RBF neural network is constructed and feature fusion is realized by using Dempster-Shafer combination rules and network learning. Parameters μ, α and γ are obtained through networks learning in the course of features fusion, s and p is decided by the decreased gradient of output error. The accuracy of behavior recognition is increased dramatically and the processing time is shorten significantly.

Key words: Feature fusion, automatic generated RBF neural network, nonlinear dimensionality reduction, recognizing of team behavior, KPCA algorithm

INTRODUCTION

Team behavior recognition using single feature will lead to unreliability for the results, as it is difficult for single feature to effectively characterize the team behavior (Lazebnik and Raginsky, 2009; Wang and Li, 2011). Firstly, characteristic dimension reduction for the extracted features will be carried on after the characteristics of clothing color of team, contour, trajectory are extracted in this study. Secondly, we merge these features using multi-feature fusion techniques. Finally, team behavior is described and team behavior recognition is carried on by using the fused characteristics. Selection of a learning algorithm for a particular application is critically dependent on its accuracy and speed. In practical online applications, sequential learning algorithms are generally preferred over batch learning algorithms as they do not require retraining whenever a new data is received. Compared with the batch learning algorithms, the sequential learning algorithms have the following distinguishing features:

- All the training observations are sequentially (one-by-one) presented to the learning system
- At any time, only one training observation is seen and learned
- The learning system has no prior knowledge as to how many total training observations will be presented
- A training observation is discarded as soon as the learning procedure for that particular observation is completed

Thus, if one strictly applies the above features of the sequential algorithms, many of the existing algorithms are not sequential. One major bottleneck seems to be that they need the entire training data ready for training before the training procedure starts and thus, they are not really sequential. This point is highlighted in a brief review of the existing algorithms given below.

The RBF networks which combines forward subset selection and zero-th-order regularization and achieves better generalization. Unlike other approaches involving several

preset parameters and thresholds (used for adding new centers and performing gradient descent) that must be tuned with each new problem. Thus, in this study, a strategy for feature fusion for team behaviors recognition using automatically generated RBF neural network is proposed, because it can overcome to the complexity of the network learning, to achieve the requirements of real-time image processing.

MATERIALS AND METHODS

In order to improve the accuracy of team behavior recognition and recognition efficiency, feature dimension reduction has been done for various extracted features and then use the RBF neural network is used to integrate these features.

Characteristics dimension reduction: In order to obtain compact description and efficient calculation of the team behaviors, KPCA algorithm is used to nonlinearly reduce the dimension of the extracted feature. Two aspects are mainly considered: (1) KPCA algorithm provides an effective learning method for subspace to discover the nonlinear structure of "behavioral space", (2) KPCA algorithm can be easily applied to any new data point and some nonlinear dimensionality reduction methods are still unclear how to describe the new data points, such as ISOMAP, LLE, etc.

In the space of $\Re^D$, a given training feature set $T_x = \{X_1, X_2,\ldots, X_M\}$ which has M elements, the goal of subspace learning is to find a embedded data set $E_y = \{Y_1, Y_2,\ldots, Y_M\}$ in low-dimensional space $\Re^D$(d<D). For the method of kernel principal component analysis, each vector X_i is firstly mapped to a non-linear Hilbert space H by $\phi:\Re^D\rightarrow H$. Then, in the space H, principal component analysis is applied to the mapping data $T_\phi = \{\phi(X_1), \phi(X_2),..., \phi(X_M)\}$. Due to the use of "core skills", the mapping process can be omitted. Let k is a positive semi-definite kernel function, nonlinear relationships of two feature vectors is defined by Eq. 1:

$$k(\vec{x}_i, \vec{x}_j) = (\phi(\vec{x}_i) \bullet \phi(\vec{x}_j)) \tag{1}$$

In the space H, the problem for searching the coefficients of main components can be attributed to diagonalization of kernel matrix κ:

$$\gamma\lambda\vec{e} = \kappa\vec{e} \tag{2}$$

Where:

$$\kappa_{ij} = k(\vec{x}_i, \vec{x}_j),\ \ \vec{e} = [e_1, e_2, \cdots, e_\gamma]^T$$

If:

$$Z = \sum_{i=1}^{\gamma} e_i \varphi(\vec{x}_i)$$

is used to represent spindle, a new point X is mapped to the j spindle Z^j can be expressed as:

$$(Z^j \cdot \phi(\vec{x})) = \sum_{i=1}^{\gamma} e_i^j (\phi(\vec{x}_i) \bullet \phi(\vec{x}_j)) = \sum_{i=1}^{\gamma} e_i^j k(\vec{x}_i, \vec{x}_j) \tag{3}$$

Gaussian kernel function is used in our experiments. After obtaining embedding space which includes first main component d, any video v can be mapped to a associated track $T_o = \{O_1, O_2,\ldots, O_T\}$ in a d dimensional feature space.

Feature fusion strategy for team behavior recognition: Three questions need to be considered in feature fusion of team behavior recognition: (1) Which features information need to be fused? Different characteristics information is reasonably selected to be fused, depending on the different application scenario, (2) At what level that feature integration is carried on? Feature fusion can choose to be implemented on the bottom level feature, mid-level keyword and senior-level semantic and (3) Which strategy is chosen in the course of features fusion? We need to select strategy for data normalization processing and fusion probability expression.

Selecting feature fusion strategy: Currently multi-feature fusion strategy have multiplicative integration, weighted fusion and discrete Karhunen-Loeve (K-L) transform fusion. For multiplicative integration, the joint distribution of multiple features is calculated by using feature weights Quadrature method which can effectively improve the accuracy of tracking moving targets and may amplify the noise. Weighted fusion adjusts weight coefficient of each feature according to the credibility of different characteristics and then calculates the total feature weights using weighted sum of each feature. Weighted feature fusion is not sensitive to noise, but can not raise the credibility of the fusion track (Wang and Yung, 2010). K-L transform fusion is a adaptive multi-feature fusion strategy which has the advantages of characteristics of remained entropy, energy, decorrelation, as well as energy re-allocation and concentration, etc., but the calculation is more complex and the learning process is lack. In this study, automatic generation of RBF neural network shown in Fig. 1 is used to integrate various features.

A wide class of multiple-input-single-output systems can be modeled by the RBF neural networks given by:

$$\hat{S} = RBF(X, \sigma, c, w) = \sum_{i=1}^{m} w_i \varphi_i(X, \sigma_i, c_i) \tag{4}$$

where, $\hat{S}$ denotes the network output, X is the input vector to the network, ϕ_i (X, σ_i, c_i) denotes the radial basis function (e.g., Gaussian basis function) of the i-th hidden node with the center $c_i\in\Re^n$ and the width $\sigma_i\in\Re^1$ and w_i is the linear output weight. The adjustable parameters in network (Eq. 4) are therefore, the center vector $c = (c_1, c_2,\ldots, c_m)$, the width vector $\sigma = (\sigma_1, \sigma_2,\ldots, \sigma_m)$ and the linear output weight vector $W = (w_1, w_2,\ldots, w_m)^T$. For the Gaussian radial basis function ϕ_i (X, σ_i, c_i)-exp(-$\|X-c_i\|^2/\sigma^2_i$, where, $\|\bullet\|$ denotes the Euclidean norm.

Input layer of network is constituted of N neurons which use the same activation function ϕ, d is the distance calculated

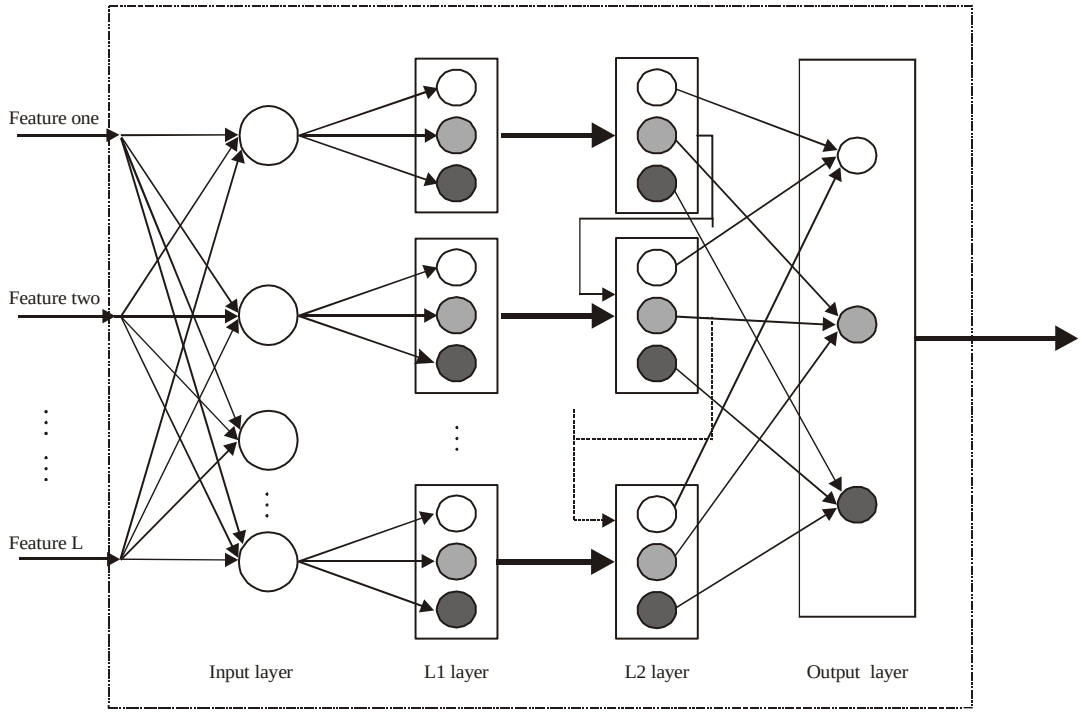

Fig. 1: RBF neural network for fusing multi-features

by using training data. $\alpha \in [0, 1]$ is weakening parameters of corresponding neuron i (Wang *et al*., 2010). There are:

$$\begin{cases} s_i = \alpha_i \phi(d_i) \\ \phi(d_i) = \exp(-\gamma_i (d_i)^2) \end{cases} \tag{5}$$

L1 layer is used to calculate the trust blocks m_i of i-th module connected to i-th neuron of previous level:

$$\begin{cases} m_i(\{w_q\}) = \alpha_i u_{q,i} \phi(d_i) \\ m_i(\Omega) = 1 - \alpha_i \phi(d_i) \end{cases} \tag{6}$$

where, $u_{q,i}$ is the member degree of each category of characteristic, q = {1, 2,…, M}.

L2 layer merges N different function blocks using the fusion composition rule of Dempster-Shafer in a single block, the combinations rules of shown in Eq. 7:

$$m(A) = (m_1 \oplus m_2 \oplus \cdots \oplus m_N) = \sum_{B_1 \cap B_2 \cap \cdots \cap B_N = A} \prod_{i=1}^{N} m_i(B_i) \tag{7}$$

Defining an excitation vector $\vec{u}_i$ of i-th module and $\vec{u}_i$ is obtained by Eq. 8 using recursive calculation:

$$\begin{cases} u_1 = m_1 \\ u_{i,j} = u_{i-1,j} m_{i,j} + u_{i-1,j} m_{i,M+1} + u_{i-1,M+1} m_{i,j} \\ u_{i,M} = u_{i-1,M} m_{i,M+1} \end{cases} \tag{8}$$

The output of output layer of RBF neural network are very sensitive to the number of original features, small changes in the number of features may cause large change of output fused characteristics (Laptev, 2005; Wang and Li, 2012; Boiman and Irani, 2007), thus excitation vector of original characteristics are considered to calculate output to reduce the impact caused by number change of features in the course of block calculation. Equation 9 is specific expression of calculation:

$$\begin{cases} O_j = \dfrac{\sum_{i=1}^{N} u_{i,j}}{\sum_{i=1}^{N} \sum_{j=1}^{M+1} u_{i,j}} \\ p_q = O_q + O_{M+1} \end{cases} \tag{9}$$

Estimation of fusion parameter: The RBF has only three preset parameter, the basis function width. The computation cost (number of floating point operations) for parameters adjustment at each learning cycle can be up to $O(n^3)$, where, n is the number of training data which is usually very large. u, α, γ are the importance factor of each feature in the process of characteristic integration which we can get through network learning.

RESULTS

The result of characteristic reduction using KPCA algorithm for two-dimension origin features is shown in Fig. 2. Where, gray× expresses input characteristic vectors and red+ is the reduced feature vectors. Figure 3 shows mapping

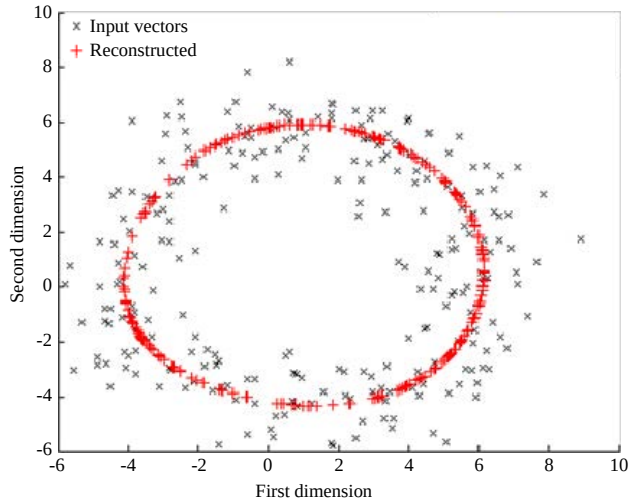

Fig. 2: Result of feature reduction for two-dimension characteristics

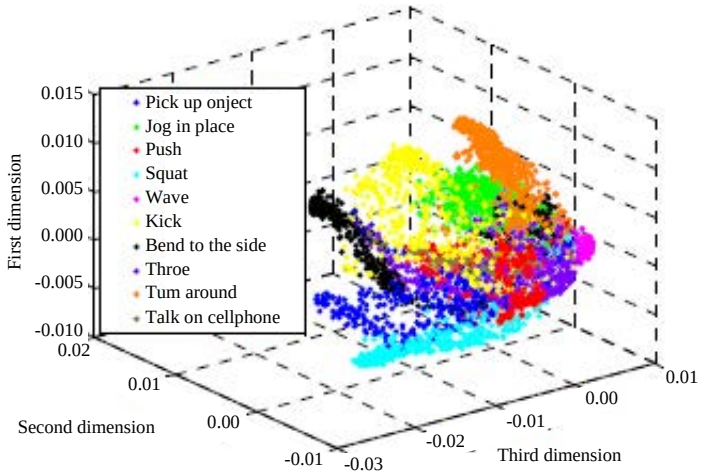

Fig. 3: Three-dimensional view of mapping track for behavior in KPCA derived subspace

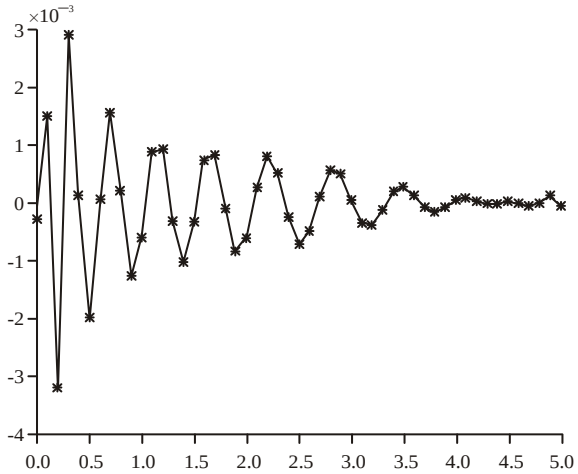

Fig. 4: Learning process of RBF neural network

trajectories (PTM) of behavior of the dataset in literature (Veeraraghavan *et al*., 2005) and the time sequence of frame is marked unclearly. Learning of RBF neural network is shown as Fig. 4. The s and p are determined by the gradient descent of the output error.

DISCUSSION

Radial Basis Function (RBF) neural networks have been widely used in many areas, such as data mining, pattern recognition, signal processing, time series prediction and nonlinear system modeling and control, because of the simple topological structure and universal approximation ability (Park and Sandberg, 1991). This is also stated in previous studies (Adams and Payandeh, 1996; Chen and Billings, 1992; Er *et al*., 2005; Gonzalez *et al*., 2003; Hong *et al*., 2003; Peng *et al*., 2006; Li and Wang, 2014; Li *et al*., 2004, 2006;

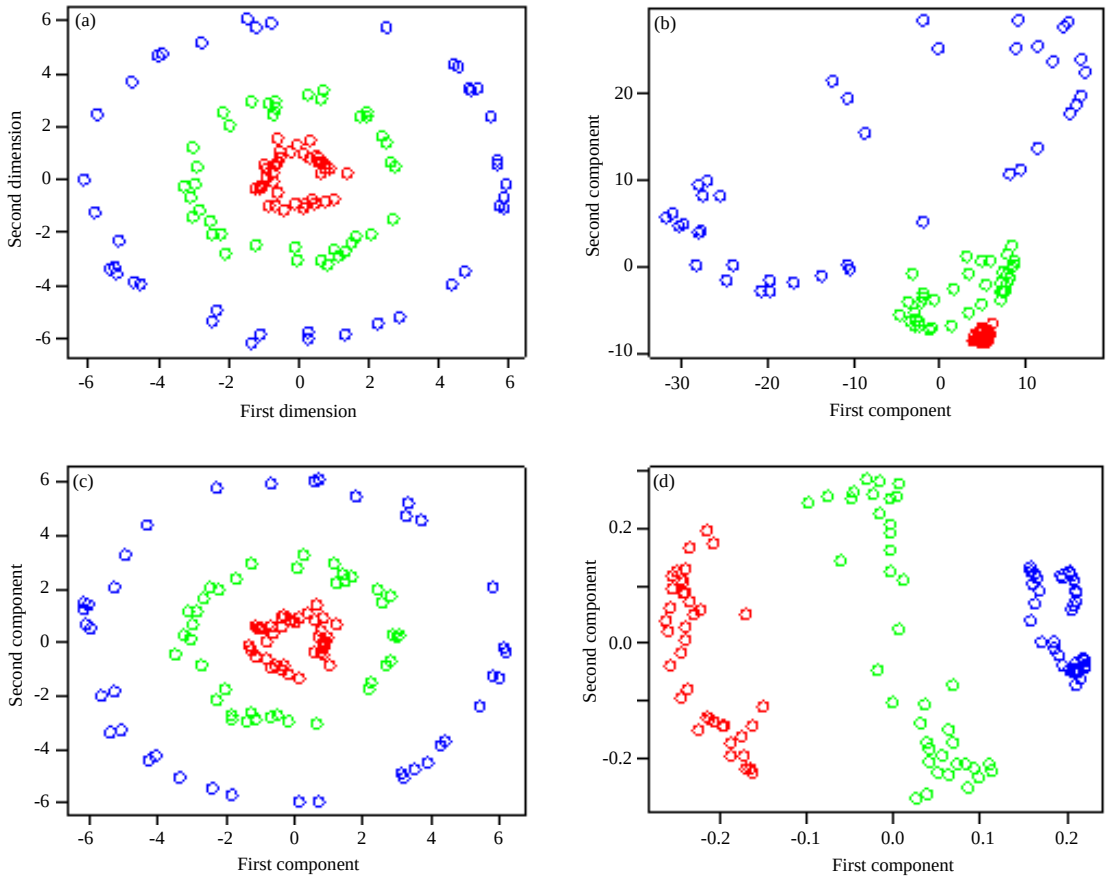

Fig. 5(a-d): Comparing of PCA and KPCA, (a) Origin, (b) Kernel PCA (Poly), (c) Traditional PCA and (d) Kernel PCA (Gaussian)

Oyang *et al*., 2005; Park and Sandberg, 1991; Polycarpou, 2001; Xie and Leung, 2005; Zhang *et al*., 2004; Zhu and Billings, 1996). Gaussian radial basis functions have also been widely used in support vector machines, an important class of machine learning algorithms. One of the most important issues in the RBF neural network applications is the network learning, i.e., to optimize the adjustable parameters which include the center vectors, the variances (or the widths of the basis functions) and the linear output weights connecting the RBF hidden nodes to the output nodes. Another important issue is to determine the network structure or the number of RBF nodes based on the parsimonious principle (Huang *et al*., 2005; Leung *et al*., 2003).

Both the issues to determine the network size and to adjust the parameters on the continuous parameter space are closely coupled. It is a mixed integer hard problem if the two issues are considered simultaneously. Evolutionary algorithms have been used to address this problem (Gonzalez *et al*., 2003; Leung *et al*., 2003), however, they are computationally very expensive to implement and it is also well known that these algorithms suffer the slow and premature convergence problems. Despite that no analytic method is available to efficiently and effectively address this integrated problem, the two separate issues have been studied extensively in the literature.

With respect to the RBF neural network learning, conventional approach takes a two stage procedure, i.e., unsupervised learning of both the centers and widths for the RBF nodes and supervised learning of the linear output weights. With respect to the center location, clustering techniques have been proposed. For the width learning, if the input samples are uniformly distributed, an identical width can be set for all the basis functions, otherwise a particular width has to be set for each individual basis function to reflect the input distribution (Park and Sandberg, 1991). Once the centers and the widths are determined, the linear output weights can be obtainable using Cholesky factorization, orthogonal least squares, or singular value decomposition.

In contrast to the conventional two stage learning procedure, supervised learning methods aim to optimize all the network parameters (Neruda and Kudova, 2005). To improve the convergence, various techniques have been introduced. For example, hybrid algorithms combine the gradient-based search for the nonlinear parameters (the widths and centers) of the

RBF nodes and the least squares estimation of the linear output weights (Panchapakesan *et al*., 2002; Peng *et al*., 2003). Second-order algorithms have also been proposed which use an additional adaptive momentum term to the Levenberg-Marquardt algorithm in order to maintain the contumacy between successive minimization directions, resulting in good convergence for some well-known hard problems. Neruda and Kudova (2005), the performances of three different RBF learning methods are compared a gradient-based algorithm (gradient descent with a momentum term), a three-step hybrid learning algorithm and a genetic algorithm. Generally speaking, although supervised learning is thought to be superior to conventional two stage approaches, it can be computationally more demanding.

Nevertheless, these above learning methods are only applicable to RBF networks of fixed structure. If the network size also has to be determined, one of the simplest ways is to repeat these above learning procedures with different network size until the optimal one is acquired based on some network selection criterion such as the Akaike Information Criterion (AIC). This method is, however, computationally too demanding. With respect to the determination of the RBF neural network structure, a popular approach is to formulate it as a linear-in-the parameters problem, where all the training patterns (or samples) are usually used as the candidate RBF centers and the RBF widths are chosen a priori. A parsimonious network is then determined from these candidates using an efficient forward subset selection method, such as the Orthogonal Least Squares (OLS) algorithms (Zhu and Billings, 1996) or the Forward Recursive Algorithm (FRA) (Li *et al*., 2005). To improve the network generalization, the Regularized Forward Selection (RFS) algorithm has been proposed (Orr, 1995) which combines subset selection with zero-order regularization. Backward selection methods have also been used in RBF center selection (Li *et al*., 2004, 2006). However, forward selection algorithms are thought to be superior to backward methods in terms of computational efficiency, since they do not need to solve the equations explicitly with the full set of initial candidate centers.

Generally speaking, existing (forward and backward) subset selection methods have several major disadvantages. First, since the RBF widths are set a priori and the centers of the RBF nodes are selected from a set of training samples with limited size, the optimal values on the continuous parameter space for the center and width parameters of RBF nodes can be easily missed out on. This means that the stepwise forward or backward procedures can easily miss out on a good RBF neural network. Second, in order to increase the chance of obtaining a satisfactory RBF network, one has to use a very large set of candidate RBF nodes of different centers and widths. This is, however, sometimes computationally too expensive or impossible to implement, since all the candidate RBF nodes have to be stored for batch operations and the number of all candidates will increase exponentially as the search space dimension increases. Part of this is usually referred to as the curse of dimensionality problem in the literature.

In order to optimize the RBF center and width parameters along with the network structure determination process, a Sparse Incremental Regression (SIR) modeling method was proposed in Chen *et al*. (2005). This method appends repressor in an incremental modeling process. For each repressor to be appended, the nonlinear parameters are tuned using a boosting search base d on a correlation criterion. In this way, the network structure and the associated nonlinear parameters are determined simultaneously. However, the search for the optimal values of the nonlinear parameters (RBF centers and widths) is a continuous optimization problem. The boosting approach in SIR which employs a stochastic search process, tends to be slower in convergence than calculus-based optimization techniques. In addition, all the nonlinear parameters are treated equally in SIR and the difference between the center and width parameters of a RBF node is ignored; this is again another factor that slows down the search process. Finally, the boosting search in SIR has three to five parameters that need to be tuned empirically. All the above potential problems with SIR are illustrated in the simulation examples at the end of this study.

Different from existing methods in RBF neural network construction, this study proposes a novel Hybrid Forward Algorithm (HFA) which performs simultaneous network growing and parameter optimization within an integrated analytic framework, leading to two main technical advantages. First, the network performance can be significantly improved through the optimization of the nonlinear RBF parameters on the continuous parameter space. Second, conventional forward selection algorithms tend to use all training samples to produce a very large set of candidate RBF nodes from which the final RBF network is selected (Adams and Payandeh, 1996; Gonzalez *et al*., 2003). The currently proposed method, however, only uses a very small number of training samples just for the initialization of the RBF centers, aiming to speed up the continuous optimization procedure.

We use KPCA to reduce the dimension of input features in this study. As a result, the memory requirement is significantly reduced. Figure 5 shows the results of two-dimensional features reduction using the methods of traditional PCA and KPCA (using kernel function poly and Gaussian) which are shown in Fig. 5c and 5d, respectively. Figure 5b shows the original features.

Automatically generated RBF neural network is constructed and extracted features of team behavior are merged by using combination rule of Dempster-Shafer and network learning. The accuracy of behavior recognition is increased dramatically and the processing time is shorten significantly in this study.

CONCLUSION

Strategy of feature fusion for team behavior recognition using automatically generated RBF neural network is presented in this study. Firstly, the underlying characteristics are extracted to reduce the burden of high-level recognition algorithm, such as trajectory feature of moving targets is extracted by using the growth method of trace. Secondly, nonlinear dimensionality reduction for extracted characteristics is implemented using KPCA algorithm. Finally, automatically generated RBF neural network is constructed and extracted features of team behavior are merged by using combination rule of Dempster-Shafer and network learning. Parameters u, α, γ are obtained through network learning in the integration process, s and p are calculated by the gradient descent algorithm of output error.

ACKNOWLEDGMENT

This study was supported in part by National Natural Science Foundation of China (61462008) and a grant from Natural Science Foundation of Guangxi (2013GXNSFAA019336) and the Foundation of Doctor of Guangxi University of Science and Technology (12z14).

REFERENCES

Adams, J. and S. Payandeh, 1996. Methods for low velocity friction compensation: Theory and experimental study. J. Robot. Syst., 13: 391-404.

Boiman, O. and M. Irani, 2007. Detecting irregularities in images and in video. Int. J. Comput. Vision, 74: 17-31.

Chen, S. and S.A. Billings, 1992. Neural network for nonlinear dynamic system modelling and identification. Int. J. Control, 56: 319-346.

Chen, S., X.X. Wang and D.J. Brown, 2005. Sparse incremental regression modeling using correlation criterion with boosting search. IEEE Signal Process. Lett., 12: 198-201.

Er, M.J., W. Chen and S. Wu, 2005. High-speed face recognition based on discrete cosine transform and RBF neural networks. IEEE Trans. Neural Networks, 16: 679-691.

Gonzalez, J., I. Rojas, J. Ortega, H. Pomares, F.J. Fernandez and A.F. Diaz, 2003. Multiobjective evolutionary optimization of the size, shape and position parameters of radial basis function networks for function approximation. IEEE Trans. Neural Networks, 14: 1478-1495.

Hong, X., C.J. Harris, S. Chen and P.M. Sharkey, 2003. Robust nonlinear model identification methods using forward regression. IEEE Trans. Syst. Man Cybern. Part A: Syst. Hum., 33: 514-523.

Huang, G.B., P. Saratchandran and N. Sundararajan, 2005. A generalized growing and pruning RBF (GGAP-RBF) neural network for function approximation. IEEE Trans. Neural Networks, 16: 57-67.

Laptev, I., 2005. On space-time interest points. Int. J. Comput. Vision, 64: 107-123.

Lazebnik, S. and M. Raginsky, 2009. Supervised learning of quantizer codebooks by information loss minimization. IEEE Trans. Pattern Anal. Mach. Intell., 31: 1294-1309.

Leung, F.H.F., H.K. Lam, S.H. Ling and P.K.S. Tam, 2003. Tuning of the structure and parameters of a neural network using an improved genetic algorithm. IEEE Trans. Neural Networks, 14: 79-88.

Li, Y., S. Qiang, X. Zhuang and O. Kaynak, 2004. Robust and adaptive backstepping control for nonlinear systems using RBF neural networks. IEEE Trans. Neural Networks, 15: 693-701.

Li, K., J.X. Peng and G.W. Irwin, 2005. A fast nonlinear model identification method. IEEE Trans. Autom. Control, 50: 1211-1216.

Li, K., J.X. Peng and E.W. Bai, 2006. A two-stage algorithm for identification of nonlinear dynamic systems. Automatica, 42: 1189-1197.

Li, S.Z. and Z.W. Wang, 2014. Adaptive fractal-wavelet image denoising based on multivariate statistical model. Chin. J. Comput., 27: 1380-1389.

Neruda, R. and P. Kudova, 2005. Learning methods for radial basis function networks. Future Gen. Comput. Syst., 21: 1131-1142.

Orr, M.J.L., 1995. Regularization in the selection of radial basis function centers. Neural Comput., 7: 606-623.

Oyang, Y.J., S.C. Hwang, Y.Y. Ou, C.Y. Chen and Z.W. Chen, 2005. Data classification with radial basis function networks based on a novel kernel density estimation algorithm. IEEE Trans. Neural Network, 16: 225-236.

Panchapakesan, C., M. Palaniswami, D. Ralph and C. Manzie, 2002. Effects of moving the center's in an RBF network. IEEE Trans. Neural Networks, 13: 1299-1307.

Park, J. and I.W. Sandberg, 1991. Universal approximation using radial-basis-function networks. Neural Comput., 3: 246-257.

Peng, H., T. Ozaki, V. Haggan-Ozaki and Y. Toyoda, 2003. A parameter optimization method for radial basis function type models. IEEE Trans. Neural Networks, 14: 432-438.

Peng, J.X., K. Li and D.S. Huang, 2006. A hybrid forward algorithm for RBF neural network construction. IEEE Trans. Neural Networks, 17: 1439-1451.

Polycarpou, M.M., 2001. Fault accommodation of a class of multivariable nonlinear dynamical systems using a learning approach. IEEE Trans. Autom. Control, 46: 736-742.

Veeraraghavan, A., A.K. Roy-Chowdhury and R. Chellappa, 2005. Matching shape sequences in video with applications in human movement analysis. IEEE Trans. Pattern Anal. Mach. Intell., 27: 1896-1909.

Wang, L. and N.H. Yung, 2010. Extraction of moving objects from their background based on multiple adaptive thresholds and boundary evaluation. IEEE Trans. Intell. Transp. Syst., 11: 40-51.

Wang, Z., S. Li, Y. Lv and K. Yang, 2010. Remote sensing image enhancement based on orthogonal wavelet transformation analysis and pseudo-color processing. Int. J. Comput. Intell. Syst., 3: 745-753.

Wang, Z. and S. Li, 2011. Face recognition using skin color segmentation and template matching algorithms. Inform. Technol. J., 10: 2308-2314.

Wang, Z. and S. Li, 2012. A secure and high-efficient dynamic key management scheme for group communication using optimized GDH. ICIC Express Lett., 6: 1815-1820.

Xie, N. and H. Leung, 2005. Blind Equalization using a predictive radial basis function neural network. IEEE Trans. Neural Networks, 16: 709-720.

Zhang, X., T. Parisini and M.M. Polycarpou, 2004. Adaptive fault-tolerant control of nonlinear uncertain systems: An information-based diagnostic approach. IEEE Trans. Autom. Control, 49: 1259-1274.

Zhu, Q.M. and S.A. Billings, 1996. Fast orthogonal identification of nonlinear stochastic models and radial basis function neural networks. Int. J. Control, 64: 871-886.