Acta Technica 62 (2017), No. 6B, 645–656

Particle swarm optimization for rbf audio-visual merging feature classification for electromechanical system condition monitoring

Jiangwei Xu[1,2], Tiejun Li[1], Liang Zhang[1], Jinyue Liu[1]

Abstract. A classification method of audio-visual information fusion for optimized double-layer RBF of monitoring particle swarm on state of electromechanical system is proposed to improve monitoring algorithm precision of electromechanical system. Firstly, auralization handling of visual pattern for electromechanical system as well as normalization of audio-visual information and analysis process of principle component are described, and filtering of most background noise is realized; secondly, a double-layer RBF algorithm is introduced, and performance improvement of double-layer RBF algorithm is realized via using parameter adaptive particle swarm algorithm; finally, algorithm validity is verified via simulation experiment.

Key words. Feature of audio-visual information fusion Electromechanical system State monitoring.

1. Introduction

Technology of state monitoring and error diagnosis of mechanical system has always been hot fields noticed by researchers. Accident potential in electromechanical system can be effectively found, and equipment failures can be promptly eliminated, and vicious accident can be prevented, and personal injury as well as environmental pollution and great economical loss, etc caused by this can be avoided via applying technology of state monitoring and error diagnosis, thus effective state monitoring and reasonable error diagnosis is one of the key technologies to ensure safe operation of electromechanical system and realize scientific maintenance.

[1]School of Mechanical Engineering, Hebei University of Technology, Tianjin 300130, China
[2]Beijing C&W Eelectronics(Group)CO, .LTD, Beijing, 100015, China

Current monitoring methods on state of electromechanical system are mostly implemented via local monitoring of interior sensing, and method of error diagnosis is also based on expert-like reasoning of interior sensing information and control logic. For example, all kinds of systems of rotation, reciprocate mechanic monitoring and error diagnosis with objects as gear, bearing, motor, generator are established, and observer/filter method, parameter estimation method, analytical redundancy method, evidence theory method, model recognition method, etc are proposed via technologies such as oil analysis, sound emission analysis, vibration signal analysis, etc. However, current large-scale complex electromechanical equipment is equipped with features of complexity, openness, coupling, etc different from common small systems with development of technologies, defects of traditional state monitoring system such as structural discrete, big data size, low integration level, etc are gradually exposed. Monitoring methods based on machine vision have appeared in recent years, but there are many problems for medium to obtain equipment state information via visual sensing singly, where its main performance is sharp fall of system decision-making level when external environmental conditions are absent, such as light source restriction, object overlap, etc. ; above problems can be solved to a great extent via compensating audio information for visual information using diffraction feature of sound.

A kind of state monitoring system mixing audio-visual information is designed in this thesis to conduct real-time and effective intelligent monitoring of operation state for electromechanical equipment. Strengthening pretreatment on audio information and visual information needs to be respectively conducted before merging of audio and visual information, mainly aiming at filtering of collected signal to eliminate noise and redundant features, and in addition to conventional methods of strengthening treatment and pretreatment on audio-visual signal in system, such as binaryzation of image, grey processing, framing of audio signal, etc, auralization handling on visual image information is also conducted, and normalization processing as well as principle component analysis on audio-visual data are also conducted.

2. Handling methods of audio-visual features

2.1. Auralization handling of visual image

For visual information and video information belong to heterogeneous information, and they have different physical malleability and information dimensions: sound belongs to one-dimension information on time domain, while visual image belongs to two-dimension information. So audio-visual information fusion belongs to sensing information merging of different kinds, and these two types of information needs to be switched before merging. Commonly three ways are adopted to conduct identical switch of audio information and visual information to make them be disposed in unity in the same environment: (1) Converting visual image into sound information, such as voice blind assisting system and paint2sound mixmeister merging, etc; (2) Converting sound wave information into image information, such as acoustic imaging technique and ultrasonic technology of B type, etc; (3)

Representing these two different information via a third information.

The first way is adopted in this thesis, namely converting visual information into sound information. The following description can be roughly conducted on the process of this converting method: pixel value of image and position information of pixel vale are selected as image mapping features, and amplitude and frequency of sound are selected as dimension of sound. Mapping relation of image and sound is:

$$s_i = \sum_{i=1}^{n} g_{i,j} \sin 2\pi f_i t\,. \tag{1}$$

In the formula, $i = 1, 2, \cdots, n$, and acoustic pattern in ith line of the image is presented by s_i; pixel value of image is presented by $g_{i,j}$; frequency of pixel in jth column is presented by f_i: play time of sound is decided by t.

$$f_j = \text{Pixel frequency/sample frequency}\,. \tag{2}$$

$$s_{i,j} = g_{i,j} \sin 2\pi f_j t \tag{3}$$

In the formula, $i = 1, 2, \cdots, n$, $j = 1, 2, \cdots, n$. For sound which can be heard and common sound information in error diagnosis of electromechanical system are both concentrated in scope of a certain frequency domain, so above frequency domain shall be avoided as possible and vacant frequency domain shall be selected to avoid mutual interference of sound information converted by image information and sound wave information collected by microphone when image information is converted into frequency parameter of sound.

2.2. Normalization and principle component analysis of audio-visual information

Audio-visual information is converted into two types of data form by above procedures: grey value data of amplitude scope as 0∼255 and frequency-domain analysis amplitude of sound, and normalization processing shall be conducted via linear function conversion to unify two types of data. Linear transfer function:

$$\text{Value upon normalization} = \frac{(\text{value before normalization} - \ \text{Min}V)}{(\text{Max}V - \text{Min}V)}\,. \tag{4}$$

In the formula, Min$V$ and Min$V$are respectively maximum and minimum value of sample. Normalization is to unify judge results in different scales in a scale to facilitate comparison and calculation. Data features of the features themselves can be removed via mathematical methods to create the same competitive ability of all feature values, thus some more universal solving methods can be used to free view emphasis from local data, and weight selection can be also avoided before superposition of two types of data.

Information data upon normalization still has relatively high dimension, while high-dimension measuring volume cannot effectively reflect nature of measured object and cannot be classified or recognized. To facilitate calculation and designing classifier and proposing recognition method, audio information and image information need to be transferred to feature space of greatly reduced dimension from measuring volume. For size of variance is considered as standard of measuring information amount for principle component analysis method, and it is a multivariate statistical method where many indexes can be converted into several comprehensive indexes in premise of losing little information, and it is easy to calculate and has optimal linear error, better uncovering distribution of high-dimension data set with linear structure in the whole situation. So principle component analysis method is adopted as main algorithm of feature extraction. Specific steps of this algorithm are as follows:

(1) Suppose there are 200 samples in the training set, composed of data from several steps hereinbefore, and every sample size is N, and vector x_i is N-dimension column vector of ith sample.

(2) Calculating of average vector:

$$\psi = \frac{1}{200}\sum_{i=1}^{200} x_i \,. \tag{5}$$

(3) Calculating difference value of every data vector and average data vector:

$$d_i = x_i - \psi, i = 1, 2, \cdots, 200 \,. \tag{6}$$

(4) Building covariance matrix:

$$C = \frac{1}{200}\sum_{i=1}^{200} d_i d_i^T = \frac{1}{200} AA^T \,. \tag{7}$$

In the formula, $A = (d_1, d_2, \cdots, d_{200})$.

(5) Getting eigenvalue and eigenvector of covariance matrix and building feature space.

Dimension of covariance matrix is $N \times N$, so eigenvalue and eigenvector of AA^T are obtained adopting singular value decomposition theorem considering its relatively high dimensions and complex calculation. The first p maximum feature vectors and their corresponding featuring vectors are selected according to contribution rate of feature value, and contribution rate refers to ratio between sum of selected feature value and that of all feature value, namely:

$$\varphi = \sum_{i=1}^{P} \lambda_i \Bigg/ \sum_{i=1}^{i=200} \lambda_i \geq a \,. \tag{8}$$

Generally a >99%, namely making 99% energy is concentrated on projection of the first p feature vector set in training sample. Feature vector of original covariance matrix can be obtained:

$$u_1 = \frac{1}{\sqrt{\lambda_i}} A v_i, (i = 1, 2, \cdots, p) . \tag{9}$$

Then feature space of merging data is:

$$w = (u_1, u_2, \cdots, u_P) . \tag{10}$$

(6) Difference value vector can be projected on feature vector space of every vector and average vector before merging, namely:

$$\Omega_i = w^T d_i, (i = 1, 2, \cdots, 200) . \tag{11}$$

Dimensionality reduction operation can be finished after association of two types of information via above process, meanwhile interested signal feature is extracted, and difference value is obtained via calculation of every data vector and average data vector in this process, namely filtering most background noise.

3. RBF algorithm based on particle swarm algorithm

3.1. Parameter adaptive particle swarm algorithm

Velocity pattern of particle swarm mainly consists [4, 5] of three parts:

$$v_{id}^{k+1} = v_{id}^k + c_r r_1 (p_{id}^k - x_{id}^k) + c_2 r_2 (p_{gd}^k - x_{id}^k) . \tag{12}$$

$$\rho = x \cos\theta + y \sin\theta . \tag{13}$$

$y = f(x)$ is called part of "history" in related literature, and its function is mainly to present inertia trend of population individual of previous generation, and it is particle flying speed of previous generation and basic speed of flying particle, and it plays the role of inertia.

$c_1 r_1 \left(p_{id}^k - x_{id}^k\right)$ is called part of "cognition", which mainly decides behavior trend in next step by observation of particles on current evolutionary process, and particles will approach to vicinity of self optimal value of the whole situation when this value increases.

$c_2 r_2 (p_{gd}^k - x_{id}^k)$ is called part of "social", for current particles are compared with neighboring particles, and good aspects of neighboring particles in forage process can be taken for example, and behavior trend of population particle can be inherited and carried forward.

It is indicated that population individual tends to regard itself as center, which does good to keeping diversity of population when particles fly towards "recognition" via simple logic analysis; while group particles can accelerate converging to optimal value of the whole situation when particles fly towards the part of "social". So two different variation content of "social" and "cognition" in the algorithm plays the role

[6, 7] of guiding particle behavior and balancing ability of “depth development” and “global exploration” for PSO algorithm.

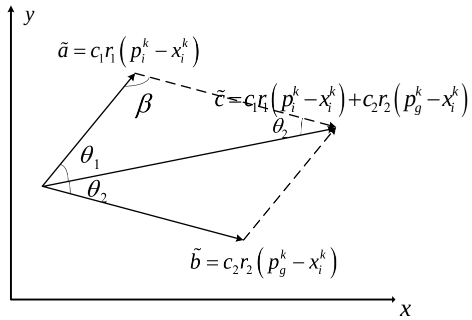

Fig. 1. Behavior vector figure of two-dimension particle

Take two-dimension vector as an example, for instance, vector superposition form of $c_1r_1(p_{id}^k - x_{id}^k) + c_2r_2(p_{gd}^k - x_{id}^k)$ is shown in Fig 1. It can be directly found by vector figure that vector angle β and $\theta_1 + \theta_2$ are fixed when $r_1(p_{id}^k - x_{id}^k)$ and $r_2(p_{gd}^k - x_{id}^k)$ are determined, then whether particles fly towards the part of “cognition” or the part of “social” will be affected by relative change of θ_1, θ_2, namely behavior trend of “development” and “exploration” for PSO algorithm can be changed via adjusting relative size of θ_1, θ_2.

Good algorithm should focus on “exploration” in initial period of searching and keep population diversity as possible to prevent premature convergence of algorithm, while in later period of searching, “development” should be emphasized, and search speed should be accelerated. Algorithm thought is implanted on every individual particle by us, and it is made to gradually ten to “development” via adjusting relative value of $c_1?c_2$ when individual tends to “exploration” excessively, vice versa, and finally balanced state of particle individual can be reached. Parameter adjustment way of $c_1?c_2$ is designed as follows in combination of individual adaptive value:

$$\begin{cases} c_1 = a - prob_i\,, \\ c_2 = b + prob_i\,, \\ prob_i = fit_i / \max fit\,. \end{cases} \tag{14}$$

In it, $\max fit$ is optimal value of current PSO population, and fit_i is adaptive value of ith particle for current PSO population, so value range of $prob_i$ is between $[0, 1]$ (aiming at optimal problem of getting maximum extremum), and a, b are constant value, then it can be achieved that $c_1 \in [a-1, a]$ and $c_2 \in [b, b+1]$. Algorithm step of PSA-PSO is as Fig. 2:

In the figure, VTR is set function value of termination objective, and $k_{\max}$ is set maximum iteration number.

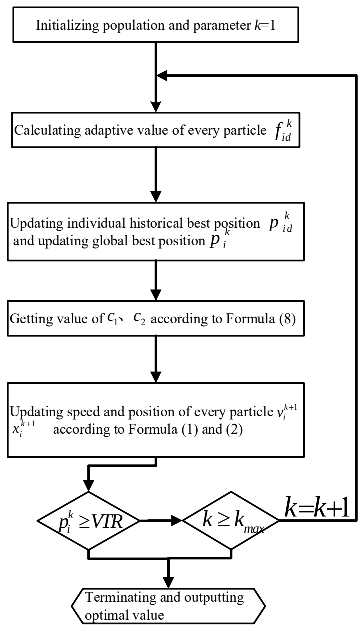

Fig. 2. Algorithm step of parameter self-adaptive particle swarm

3.2. Double-layer RBF network

RBF is a three-layer network [13], and its connecting weight from input layer to implicit strata is a fixed value, equaling to one, and connecting weight from implicit strata to output layer is variable. Suppose sample training set of RBF network is $\{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^{l}$, and $\mathbf{x}_i \in \mathbf{R}^n$ is input data of fixed RBF network, and $\mathbf{y}_i \in \mathbf{R}^n$ is output data of network, then value form of three-layer forward RBF neural network is:

$$\mathbf{y}_j = \sum_{i=1}^{N} \omega_i g_i(\mathbf{x}) + \mathbf{b}_j \,. \tag{15}$$

In the formula, $g_i(x)$ is activation function of implicit strata for RBF network, and it can be normally presented in the following formula:

$$g_i(\mathbf{x}) = \exp\left(-\|\mathbf{x} - \mathbf{c}_i\|^2 \big/ \sigma_i\right) \,. \tag{16}$$

In the formula, c_i is radial basis function center of RBF network, and σ_i is width of this function. All input sample can be radial basis function center of RBF network, and fixed radial basis function width is selected, and suppose:

$$\begin{cases} \mathbf{W} = (\omega_1, \omega_2, \cdots, \omega_l)^T \,, \\ \mathbf{G}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \cdots, g_l(\mathbf{x}))^T \,. \end{cases} \tag{17}$$

Via above definition, RBF network form can be defined as:

$$\mathbf{y} = \mathbf{W}^T \mathbf{G}(\mathbf{x}) + \mathbf{b}. \tag{18}$$

It can be found that $\mathbf{G}(\mathbf{x})$ in the formula plays the role of mapping function $\phi(\mathbf{x})$ similar to kernel function in last chapter via analysis of Formula (18). The difference lies in that there is no need to know mapping dimension for mapping function $\phi(\mathbf{x})$, while mapping dimension of $G(\mathbf{x})$ is known, being input sample number l of RBF network.

It can be found that canonical variable $\mathbf{u}$ and $\mathbf{v}$ related to analysis of kernel canonical can be obtained based on double-layer RBF network when all RBF network input samples are used as radial basis function center via analysis of Formula (18) and (16). Solving form of double-layer RBF network is shown in Fig. 3.

Its form of objective function can be designed as based on Lagrange algorithm:

$$J = E(\mathbf{uv}) + \frac{1}{2}\lambda_1\left(1 - \|\mathbf{u}\|^2\right) + \frac{1}{2}\lambda_2\left(1 - \|\mathbf{v}\|^2\right). \tag{19}$$

Training rule form is obtained as follows via gradient descent algorithm on Formula (19):

$$\begin{cases} \Delta\omega_{1j} = \eta_1 \phi_j(\mathbf{x})(\mathbf{v} - \lambda_1 \mathbf{u}). \\ \Delta\lambda_1 = \eta_2(1 - \|\mathbf{u}\|^2). \\ \Delta\omega_{2j} = \eta_1 \phi_j(\mathbf{y})(\mathbf{u} - \lambda_1 \mathbf{v}), \\ \Delta\lambda_2 = \eta_2(1 - \|\mathbf{v}\|^2). \end{cases} \tag{20}$$

In Formula (20), ω_{1j} is weight vector (the jth) corresponding to $\mathbf{u}$; in a similar way, $\Delta\omega_{2j}$ is weight vector (the jth) corresponding to $\mathbf{v}$, and $\phi_j(\mathbf{x})$ as well as $\phi_j(\mathbf{y})$ are input and output corresponding to implicit strata j, and η_1 as well as η_2 are study parameters. Optimal weight can be obtained with minimum training objective, then canonical vectors $\mathbf{u}$ and $\mathbf{v}$ are achieved.

4. Experimental analysis

Above method was tested in August of 2014 in Beijing General Research Institute of Mining and Metallurgy. Detection objective was boxing manipulator of manganese. Position of this platform manipulator is composition of air cylinder motion for X-axis and Z-axis, mingling and placing of manganese are finished by pneumatic finger. The experiment aims at monitoring core action procedure of functional actions in process of moving manganese-additid with manipulator.

For the experiment aims at monitoring core action procedure of functional actions in process of moving manganese-additid with manipulator, while its core action is mainly to pick out manipulator for finishing the series of cyclic motion such as grabbing manganese-additid with pneumatic finger, ascent of Z electric cylinder, left shift of X Na electric cylinder, descent of Z uranium electric cylinder and manganese additid release with pneumatic finger, for position motion of motion shift is relatively

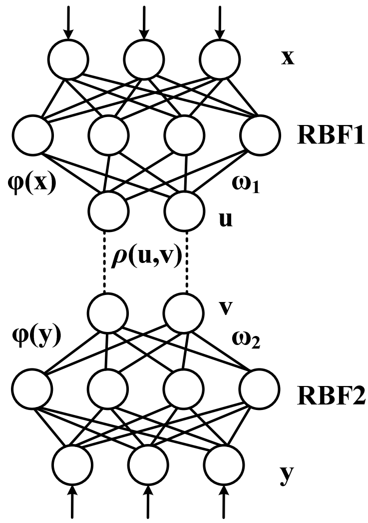

Fig. 3. Double-layer RBF network

complex and unstability condition is most likely to occur in this series of motions, thus leading to abnormal sound in these point positions easily, so monitoring is conducted via selecting proceeding state of motion shift position in its motion cycle for manipulator.

For this, four related monitoring position points are designed in this experiment, and X_1Z_2 is position where manipulator grabs workpieces, and X_1Z_1 is zero position of manipulator, and X_2Z_1 is transition position of manipulator to transport workpieces, and X_2Z_2 is objective position of manipulator to transport workpieces. There are four key points of experimental monitoring, and it is described with eight states, and coding result on these eight states is shown in Table 1 according to way of binary coding.

Table 1. Coding table of state information

Code	State
0000	Position normality of X_1Z_1
0001	Position abnormality of X_1Z_1
0010	Position normality of X_1Z_2
0011	Position abnormality of X_1Z_2
0100	Position normality of X_2Z_1
0101	Position abnormality of X_2Z_1
0110	Position normality of X_2Z_2
0111	Position abnormality of X_2Z_2

Electromagnetic environment of industrial field is very complex, so there are high-frequency noise and harmonic interference in many forms inevitably. These disturbances are of relatively strong randomness and uncertainty, and quantitative disturbance information cannot be achieved in reality, so method of software simulation is adopted to conduct tests, namely Gaussian noise with overlaid average value as zero and variance as 0.2 after normalization of audio and visual information.

60 experiments are conducted aiming at the following four situations:

(1) State monitoring of single vision modal is conducted in environment without noise; (2) State monitoring of merging audio and visual information is conducted in environment without noise; (3) State monitoring of single vision modal is conducted on condition of visual noise; (4) State monitoring of merging audio and visual information is conducted on condition of visual noise. Standard RBF and BP algorithms are selected for contrast algorithm, and correction rate of state recognition of manipulator is shown in Table 2.

Table 2. Recognition result

Noise	Modal	Recognition rate		
		BP	RBF	Algorithm in this thesis
No noise	Single vision	72.1	83.2	85.2
No noise	Merging of audio and visual information	95.3	96.2	98.3
Visual noise	Single vision	65.2	75.1	89.4
Visual noise	Merging of audio and visual information	85.2	89.3	94.2

It is indicated in result of the experiment that in environment of industrial field, recognition rate of state monitoring based on single vision sensing is only 65.2%, failing to satisfy monitoring requirement of operation state for equipment. While correction recognition rate in noise environment of state monitoring system can be steadily maintained above 89.4% via abnormal sound in abnormal operation state of equipment based on image auralization technology merging audio-visual information, and it is indicated in result of the experiment that recognition rate of the algorithm in this thesis is higher than that of standard RBF and BP algorithms, manifesting validity of algorithm.

5. Conclusion

A classification method of audio-visual information fusion for optimized double-layer RBF of monitoring particle swarm on state of electromechanical system is proposed in this thesis, and double-layer RBF algorithm performance improvement is realized and state monitoring audio-visual merging classification is conducted via parameter adaptive particle swarm algorithm, and algorithm validity is verified by simulation result. There are mainly the following research directions: (1) algorithm performance improvement of particle swarm; (2) structural improvement of RBF algorithm; (3) complexity optimization of algorithm calculation.

6. Acknowledgement

The National Natural Science Foundation of China under Grant No.51175145.

References

[1] Xu L, Qian F, Li Y, et al.: (2016) *Resource allocation based on quantum particle swarm optimization and RBF neural network for overlay cognitive OFDM System*[J]. Neurocomputing, 173(P3):1250-1256.

[2] Qasem S N, Shamsuddin S M: (2009) *Hybrid Learning Enhancement of RBF Network Based on Particle Swarm Optimization*[M]// Advances in Neural Networks – ISNN 2009. Springer Berlin Heidelberg, 2009:19-29.

[3] Niros A D, Tsekouras G E, Tsolakis D, et al.: (2015) *Hierarchical Fuzzy Clustering in Conjunction with Particle Swarm Optimization to Efficiently Design RBF Neural Networks*[J]. Journal of Intelligent & Robotic Systems, 78(1):1-21.

[4] Oh S K, Kim W D, Pedrycz W, et al.: (2011) *Polynomial-based radial basis function neural networks (P-RBF NNs) realized with the aid of particle swarm optimization*[J]. Fuzzy Sets & Systems, 163(1):54-77.

[5] Kitayama S, Yasuda K, Yamazaki K: (2005) *The Integrative Optimization by RBF Network and Particle Swarm Optimization*[C]// Hybrid Intelligent Systems, 2005. HIS '05. Fifth International Conference on. IEEE, 2005:233-238.

[6] Qasem S N, Shamsuddin S M: (2009) *Hybrid Learning Enhancement of RBF Network Based on Particle Swarm Optimization*[J]. Studies in Computational Intelligence, 201:381-397.

[7] Zhang Y, Wang S, Ji G, et al.: (2013) *An MR Brain Images Classifier System via Particle Swarm Optimization and Kernel Support Vector Machine*[J]. Scientific World Journal, 2013(20):130134.

[8] Qasem S N, Shamsuddin S M: (2013) *Radial basis function network based on time variant multi-objective particle swarm optimization for medical diseases diagnosis*[J]. Applied Soft Computing, 11(1):1427-1438.

[9] Zhong B, Wang D, Li T: (2009) *Application of optimized neural network based on particle swarm optimization algorithm in fault diagnosis*[C]// IEEE International Conference on Cognitive Informatics, Icci 2009, June 15-17, 2009, Hong Kong, China. DBLP, 2009:476-480.

[10] Satpute K, Agrawal S, Agrawal J, et al.: (2013) *A Survey on Anomaly Detection in Network Intrusion Detection System Using Particle Swarm Optimization Based Machine Learning Techniques*[M]// Proceedings of the International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA). Springer Berlin Heidelberg, 2013:6 pp.-1320.

[11] Yassi M, Moattar M H: (2015) *Optimal SVM parameters estimation using chaotic accelerated particle swarm optimization for genetic data classification*[C]// International Congress on Technology, Communication and Knowledge. IEEE, 2015:1-9.

Received May 7, 2017