

# Gyroscope Fault Diagnosis Using Fuzzy SVM to Unbalanced Samples

Luo Qiufeng (罗秋凤)<sup>1\*</sup>, Zhang Rui (张锐)<sup>1</sup>, Li Yong (李勇)<sup>2</sup>,  
Yang Zhongqing (杨忠清)<sup>1</sup>

1. UAV Research Institute, Nanjing University of Aeronautics and Astronautics,  
Nanjing 210016, P. R. China;

2. Department of Automation, Nanjing University of Aeronautics and Astronautics,  
Nanjing 210016, P. R. China

(Received 13 November 2014; revised 11 January 2015; accepted 12 January 2015)

**Abstract:** A novel fuzzy support vector machine based on unbalanced samples (FSVM-US) is proposed to solve the high false positive rate problem since the gyroscope output is susceptible to unmanned aerial vehicle (UAV) airborne electromagnetic environment and the gyroscope abnormal signal sample is rather rare. Firstly, the standard deviation of samples projection to normal vector for SVM classifier hyper plane is analyzed. The imbalance feature expression reflecting the hyper plane shift for the number imbalance between samples and the dispersion imbalance within samples is derived. At the same time, the denoising factor is designed as the exponential decay function based on the Euclidean distance between each sample and the class center. Secondly, the imbalance feature expression and denoising factor are configured into the membership function. Each sample has its own weight denoted the importance to the classifier. Finally, the classification simulation experiments on the gyroscope fault diagnosis system are conducted and FSVM-US is compared with the standard SVM, FSVM, and the four typical class imbalance learning (CIL) methods. The results show that FSVM-US classifier accuracy is 12% higher than that of the standard SVM. Generally, FSVM-US is superior to the four CIL methods in total performance. Moreover, the FSVM-US noise tolerance is also 17% higher than that of the standard SVM.

**Key words:** fault diagnosis; gyroscope; fuzzy support vector machine (FSVM); unbalanced samples; membership function

**CLC number:** V240; O321

**Document code:** A

**Article ID:** 1005-1120(2015)01-0016-07

## 0 Introduction

Gyroscope is a key airborne sensor for unmanned aerial vehicle (UAV)<sup>[1]</sup>. It is reported that about 60% flight faults are associated with the gyroscope in an inertial navigation system. The gyroscope fault brings great risks into the flight safety of UAV. The main methods for sensor fault diagnosis are based on hardware redundancy, analytical redundancy, signal processing or artificial intelligence algorithmic techniques<sup>[2-6]</sup>. Support vector machines (SVM) is one of the artificial intelligence learning algorithms<sup>[7-8]</sup>, and it has been successfully applied to

many real-world classification problems from various domains. The gyroscope fault datasets are far less than the normal datasets. In addition, the airborne gyroscope output signals are blended with noise signals or outlier from the UAV airborne complex electromagnetic environment where radio signals, engine working radiation, radar signals and radio interference signals often exist synchronously. Although SVM often deals effectively with balanced datasets, it could produce suboptimal results when faced imbalanced datasets. The SVM classification hyper-plane would shift to the minority (positive) dataset when it is applied to noises and unbalanced datasets<sup>[9]</sup>. Its classifi-

\* **Corresponding author:** Luo Qiufeng, Researcher, E-mail: yqwl2002@126.com.

**How to cite this article:** Luo Qiufeng, Zhang Rui, Li Yong, et al. Gyroscope fault diagnosis using fuzzy SVM to unbalanced samples[J]. Trans. Nanjing U. Aero. Astro., 2015, 32(1):16-21.

<http://dx.doi.org/10.16356/j.1005-1120.2015.01.016>

cation accuracy greatly drops, and the majority (negative) dataset is over fitting.

Two main methods are addressed to solve the unbalanced dataset problem. One is based on the data preprocessing techniques such as over sampling, under sampling, and synthetic minority over sampling technique (SMOTE)<sup>[10-12]</sup>. The other is about the algorithmic modifications such as the kernel function improvement and different error costs (DEC)<sup>[13-15]</sup>. The SVM classification performance is enhanced to some degree when the number of the positive dataset and negative dataset is changed by resampling techniques, but the dataset structure and its integrity are damaged. Resampling techniques are adopted to tackle the imbalance problem between datasets, while it takes into few consideration the imbalance and noise problem within datasets. The fuzzy support vector machine (FSVM) algorithm is proposed in this paper to solve the number imbalance between datasets, dispersion imbalance within datasets, and noises from the UAV airborne electromagnetic environment when the imbalanced feature weights and denoising parameters are applied to the membership function. The gyroscope fault diagnosis system on fuzzy SVM for unbalanced samples is constructed.

## 1 SVM Algorithms

### 1.1 Standard SVM

SVM is a kind of classifier developed from statistical learning theory<sup>[7]</sup>. For a given dataset  $\{(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i), \dots\}$ , where  $x_i \in \mathbf{R}^n$  represents an  $n$ -dimensional data point,  $y_i \in \{-1, +1\}$  is a sample label. The goal of the SVM learning algorithm is to find a separating hyper plane that separates these datasets into two classes. According to the Mercer theorem, if the dataset is linearly separable, the optimal hyper plane can be transformed into a minimum target optimization problem as

$$\begin{aligned} \min & \left( \frac{1}{2} \boldsymbol{\omega} \cdot \boldsymbol{\omega} + C \sum_{i=1}^l \xi_i \right) \\ \text{s. t. } & y_i (\boldsymbol{\omega} \cdot \varphi(x_i) + b) \geq 1 - \xi_i \\ & \xi_i \geq 0, i = 1, 2, \dots, l \end{aligned} \quad (1)$$

where  $\varphi(x_i)$  is the mapping function from sample

attribute to high-dimensional feature space,  $\xi$  the slack variable, parameter  $C$  the trade-off between training accuracy and generalization,  $\boldsymbol{\omega}$  the weight vector that defines a direction perpendicular to the hyper plane of the decision function,  $b$  the bias that moves the hyper plane parallel to itself.

When an SVM classifier is trained using an imbalanced dataset, it can obtain a classification hyper plane bias toward the minority class. If the number between two class samples is roughly equivalent, the classification hyper plane will shift to the more scattered class. The classification hyper plane bias situation for imbalance dataset is shown in Figs. 1, 2, where  $H$  is the classifier hyper plane, and  $H_1, H_2$  are the support vector planes.

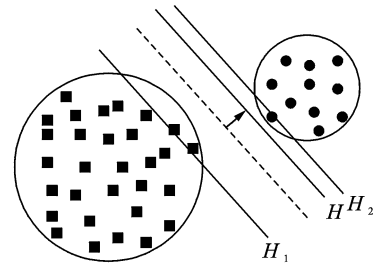


Fig. 1 Bias situation for the number imbalance

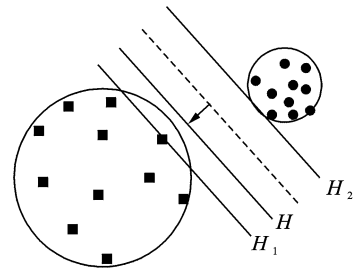


Fig. 2 Bias situation for the dispersion imbalance

### 1.2 Fuzzy SVM algorithm

Fuzzy technology is applied to SVM to reduce the effect of noise samples (or outliers) on SVM classification<sup>[16]</sup>. According to the effect of each sample during the SVM training, each sample is given a different membership value. Assuming that each sample  $(x_i, y_i)$  is assigned a membership value  $s_i$ , the parameter  $S = \{(x_1, y_1, s_1), (x_2, y_2, s_2), \dots, (x_i, y_i, s_i)\}$  forms a fuzzy training set. After a nonlinear transformation  $\varphi: \mathbf{R}^n \rightarrow H$  is defined, input space can be transformed into high dimensional Hilbert space  $H$ ,  $x \rightarrow \varphi(x)$ . The problem of the optimal hyper plane for FS-

VM model is deformed into the following optimization problems

$$\begin{cases} \min_{\omega, b, \xi} \left( \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^l s_i \xi_i \right) \\ \text{s. t. } y_i (\|\omega\| \varphi(x_i) + b) \geq 1 - \xi_i \\ \xi_i > 0, i = 1, 2, \dots, l \end{cases} \quad (2)$$

where the membership  $s_i$  ( $i = 1, 2, \dots, l$ ) represents the degree of training points  $(x_i, y_i, s_i)$  belonging to a certain class. The larger the value of  $s_i$  is, the less likely the corresponding samples are misclassified; and vice versa.

## 2 Fuzzy Membership Function Construction

The key of FSVM-US design is the construction of membership function. According to the working environment of the gyroscope, the design of membership function should solve two problems. One is to suppress the disturbance of noise and outlier in the class, and the other is to inhibit the imbalance of samples size and dispersion which affects the classification hyper plane shift. A membership function is represented by

$$\begin{cases} s_i^+ = m^+ f(x_i) \\ s_i^- = m^- f(x_i) \end{cases} \quad (3)$$

where  $s_i^+$  denotes the membership of the positive class (also called the fault samples),  $s_i^-$  the membership of the negative class (also called the normal samples).  $m^+$ ,  $m^-$  represent the uniform weight of positive class and negative class samples, respectively, which reflect the class sample size and dispersion imbalance characteristics.  $f(x_i) \in (0, 1)$  is the denoising fuzzy function that describes the importance of samples. Obviously, the membership function takes the interference of noise and outliers into account, as well as the classification hyper plane shift for the imbalance between or within class samples.

### 2.1 Imbalanced feature weight

Dispersion can be represented by the standard deviation of projection to  $\omega$  vector direction of classification hyper plane, the standard deviation of the positive class and the negative class samples point are

$$\sigma^+ = \sqrt{\omega^T \Sigma^+ \omega}, \quad \sigma^- = \sqrt{\omega^T \Sigma^- \omega}$$

$$\begin{aligned} \sigma^+ &= \sqrt{\frac{1}{l^+ - 1} \sum_{i=1}^{l^+} \left[ \sum_{i=1}^l \alpha_i y_i k(x_i, y_i) - \frac{1}{l^+} \sum_{i=1}^{l^+} \sum_{i=1}^l \alpha_i y_i k(x_i, y_i) \right]^2} \\ \sigma^- &= \sqrt{\frac{1}{l^- - 1} \sum_{i=1}^{l^-} \left[ \sum_{i=1}^l \alpha_i y_i k(x_i, y_i) - \frac{1}{l^-} \sum_{i=1}^{l^-} \sum_{i=1}^l \alpha_i y_i k(x_i, y_i) \right]^2} \end{aligned} \quad (4)$$

where  $\Sigma^+$ ,  $\Sigma^-$  are the covariance matrixes of the positive and negative classes in high-dimensional Hilbert space, respectively; and  $\varphi(x)$  is a kernel function to substitute the dot product of mapping function. After initial training, SVM can obtain the vector  $\omega$  of the classifier hyper plane.  $l^+$ ,  $l^-$  are the numbers of the positive class samples and negative class samples, respectively. To correct the offset caused by the different number of samples, the positive samples should be assigned higher weights than the negative ones. The weight is set inversely proportional to the number of samples.

$$\frac{m^+}{m^-} \propto \frac{l^-}{l^+} \quad (5)$$

Besides, due to the classification bias toward the larger dispersion class, the larger dispersion class should have a higher weight. The weights of the dispersion imbalance within class can be set as

$$\frac{m^+}{m^-} \propto \frac{\sigma^+}{\sigma^-} \quad (6)$$

Finally the imbalanced feature weight is designed as

$$m^+ = m, m^- = m \cdot \frac{l^+}{l^-} \cdot \frac{\sigma^-}{\sigma^+} \quad (7)$$

### 2.2 Denoising factor design

Denoising factor reflects the importance of sample within class. Assuming  $o_1$  and  $o_2$  are the centers of two class samples shown in Fig. 3, respectively, obviously the sample  $x_1$  and  $x_2$  are away from their respective class center for noise properties. To reduce this noise interference, FSVM algorithm assigns a less weight for samples away from their center.

The Euclidean distance between samples and the class center is written as

$$\bar{x} = \frac{1}{l^\pm} \sum_{i=1}^{l^\pm} x_i \quad (8)$$

$$d_i = \|x_i - \bar{x}\|^{\frac{1}{2}} \quad (9)$$

Then denoising factor can be calculated by the exponential decay function

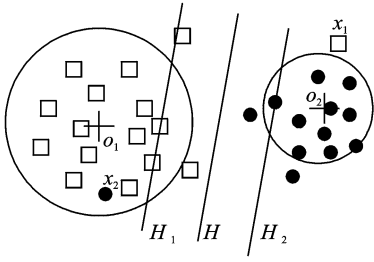


Fig. 3 SVM classifier of noise pollution dataset

$$f(x_i) = \frac{2}{1 + \exp(\beta d_i)} \quad (10)$$

## 3 Gyroscope Fault Diagnosis System

### 3.1 Gyroscope signal feature extraction

Pitch angle output sample data are from the gyroscope in a certain UAV real flight with the signal sampling period of 20 ms. The main fault types of gyroscope include complete fault, constant bias fault, drift fault and periodic interference fault. The three-layer db4 wavelet method is used to decompose signal samples, and the corresponding frequency band energy as a feature vector is obtained.

### 3.2 Imbalance classification algorithm evaluation criteria

Generally, the accuracy rate of minority class samples as fault event attracts more attention in practice. The classifier generalization performance evaluation criteria for imbalanced dataset mainly take geometric means (g-means) metric and F-measure. A slightly improved g-means criteria is adopted in this paper since g-means criteria is used for the binary classification events. The gyroscope fault detection classification belongs to multi-classifier events. To reflect the imbalanced data classification, the four fault abnormal samples are regarded as a positive class. The confusion matrix of the dataset is shown in Table 1.

Table 1 Confusion matrix of dataset

Predicted label	True	False
Positive class	TP	FP
Negative class	TN	FN

In Table 1, TP, FP denote the correctly

classified and misclassified numbers for positive samples, respectively. Likewise, TN, FN are defined for negative samples. The classification accuracy rate for positive samples is

$$acc_{pos} = \frac{TP}{(TP + FN)} \quad (11)$$

The classification accuracy rate for negative samples is defined as

$$acc_{neg} = \frac{TN}{(TN + FP)} \quad (12)$$

Moreover, the geometric average is

$$a_{g\text{-means}} = \sqrt{acc_{pos} \times acc_{neg}} \quad (13)$$

### 3.3 Simulation

Classification program of SVM is compiled on Libsvm-weights-3.17 toolbox developed by Dr. Lin. The dataset imbalanced ratio is 5 with 400 sets of normal samples to 80 sets of each kind gyroscope fault sample. The kernel function in the experiment is Gaussian radial basis function  $k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$ , and grid searching method and cross-validation method are synthesized to select the optimal kernel parameter  $\gamma$  and penalty parameter  $C$ . The two-parameter optimization process is shown in Fig. 4.

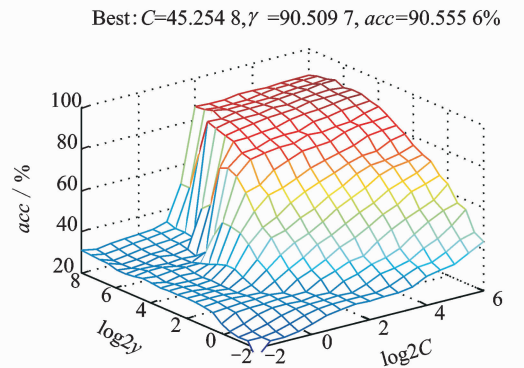


Fig. 4 Parameter optimization process of grid searching method and cross-validation method

#### 3.3.1 Standard SVM, FSVM and FSVM-US experiments

The classifier experiment of FSVM-US is conducted. Compared with standard SVM (SVC) and FSVM, the results are shown in Fig. 5.

From Fig. 5, the performance of FSVM-US classifier is generally better than those of traditional FSVM and SVM, whose fault diagnostic accuracy rates are 81.4% and 86.4%, respec-

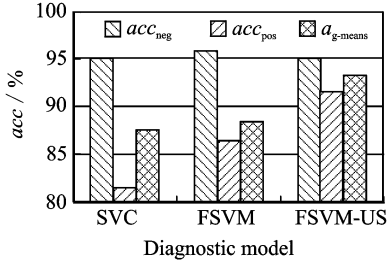


Fig. 5 Experimental results of three diagnostic machines on unbalanced samples

tively. Highly unbalanced samples result in classifier hyper plane bias toward the minority class, and the minority class has a high misclassification loss. The fault diagnostic accuracy rate of FSVM-US reaches 91.5%, 5% higher than that of the standard FSVM, and 10% higher than that of SVM approximately. The geometric average on FSVM-US is up to 93.2%, increases by approximately 5.6% and 4.7% relative to the standard SVM and FSVM, respectively. The experiment results indicate that the FSVM-US can increase the diagnosis accuracy and weaken the classifier shift influence resulted from unbalanced dataset effectively.

### 3.3.2 Experiment on five-class imbalanced learning algorithms

The above experimental results show that the FSVM-US is superior to the standard SVM and traditional FSVM. With respect to the algorithms specially dealing with the unbalanced problem in Introduction, the advantage of FSVM-US is still not sure. Hence, several typical unbalanced learning algorithms, such as under sampling, over sampling, SMOTE and DEC, are compared with the algorithm proposed in this paper. The comparison results are shown in Fig. 6.

In Fig. 6, the five-class imbalanced learning algorithms obviously increase the classifier accuracy rate for fault samples compared with the standard SVM. The geometric average on FSVM-US algorithms is still 2% higher than the best one, SMOTE methods, and much higher than other class imbalanced learning algorithms.

### 3.3.3 FSVM-US noise tolerance capability experiment

To verify noise tolerance capability of

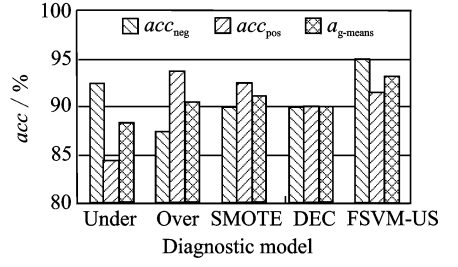


Fig. 6 Experimental results of five-class imbalanced learning

FSVM-US, one adds Gaussian white noise whose amplitude mean is 0, variance is 0.5 to signal samples of gyroscope. Different attenuation coefficient  $\beta$  in denoising fuzzy function is given as  $\{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$ . The results are shown in Fig. 7.

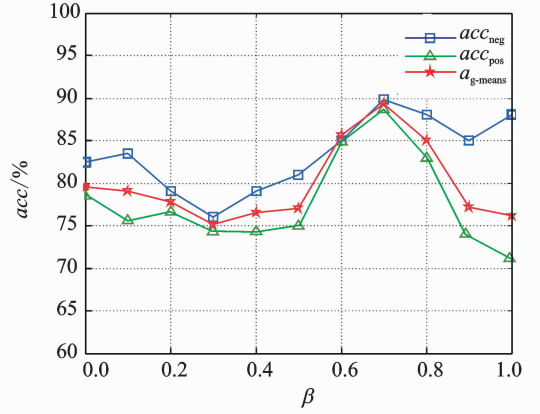


Fig. 7 FSVM-US diagnosis results with different  $\beta$

Fig. 7 demonstrates that different attenuation coefficient has great influences on the diagnostic accuracy of the algorithm. When  $\beta$  is equivalent to 0.7, the FSVM-US performance is optimum. The gyroscope fault diagnosis experiment is performed on SVM, FSVM and FSVM-US under the condition of  $\beta=0.7$ .

From Fig. 8, the diagnostic accuracy of standard SVM is the worst, i. e. 72.3% when noises are add-

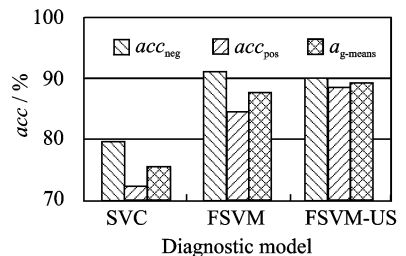


Fig. 8 Results of three diagnosis models under noise condition

ed to the imbalanced samples, it is reduced by 10% relative to the one without noise. Luckily, the accuracies of the traditional FSVM and FSVM-US are 87.5% and 89.2%, only reduced by 1%—4%, compared with those in the situation without noise.

## 4 Conclusions

The classification accuracy of the traditional SVM is decreased because it is susceptible to the influence of noise pollution and unbalanced samples. FSVM-US algorithm is similar to most existing algorithms in considering of the sensitivity of classifier hyper plane to the quantity imbalance between negative and positive class samples. Moreover, it also considers the sensitivity to the dispersion imbalance with the same class samples, as well as the one to the noise pollution in trained samples. The experiment results show that the FSVM-US classification accuracy is improved obviously, and the proposed algorithm has strong robustness. Its fault accuracy performance has many advantages in general.

Theoretically, the gyroscope fault diagnosis system based on FSVM-US can sharply reduce the risk of UAV flight safety. The work on FSVM-US algorithm will be transplanted to hardware platform for the further investigation on its engineering practice.

## Acknowledgement

This work was supported by the Fundamental Research Fund for the Central Universities (No. 56XZA12017).

## References:

- [1] Guo Xiuzhong. Theory of gyro inertial navigation system [M]. Beijing: National Defence Industry Press, 1996. (in Chinese)
- [2] Wang Xiaomeng, Zhang Ren. A sensor fault diagnosis method research based on wavelet transform and Hilbert-Huang transform [C]// Fifth Conference on Measuring Technology and Mechatronics Automation. [S. l.]: IEEE, 2013: 81-84.
- [3] Mirzaee A, Salahshoor K. Fault diagnosis and accommodation of nonlinear systems based on multiple model adaptive unscented Kalman filter and switched MPC and H-infinity loop-shaping controller [J]. Journal of Process Control, 2012(22): 626-634.
- [4] Zhu Daqi, Liu Yongan. Information fusion method for fault diagnosis [J]. Control and Decision, 2007, 22(12): 1321-1328.
- [5] Xiao Lizhi, Sun Dexiang, Liu Yuwei, et al. A combined method based on expert system and BP neural network for UAV systems fault diagnosis [C]// 2010 International Conference on Artificial Intelligence and Computational Intelligence. [S. l.]: IEEE, 2010: 3-6.
- [6] Zhang Yalin, Wei Min, Chen Pengzhang, et al. Overview on sensor fault diagnosis technology [J]. Transducer and Microsystem Technologies, 2009, 28(1): 4-12. (in Chinese)
- [7] Vapnik V. The nature of statistical learning theory [M]. New York: Springer Verlag, 2000.
- [8] Yu Jun, Lou Peihuang, Wu Xing, et al. Automated guided vehicle cross path recognition based on rough set and hierarchical support vector machine [J]. Journal of Nanjing University of Aeronautics & Astronautics, 2013, 45(1): 62-69. (in Chinese)
- [9] Akbani R, Kwek S, Japkowicz N. Applying support vector machines to imbalanced datasets [C]// 15th European Conference on Machine Learning. Berlin: Springer Verlag, 2004: 39-50.
- [10] He Haibo, Garcia E A. Learning from imbalanced data [J]. IEEE Transactions on Knowledge and Data Engineering, 2009, 21(9): 1262-1284.
- [11] Chawla N, Bowyer K, Kegelmeyer P. SMOTE: Synthetic minority oversampling technique [J]. Journal of Artificial Intelligence Research, 2002(16): 321-357.
- [12] Tao Xinmin, Liu Furong, Tong Zhijing, et al. Novel fault detection method based on SVM with unbalanced datasets [J]. Journal of Vibration and Shock, 2010, 29(10): 8-12. (in Chinese)
- [13] Veropoulos K, Campbell C, Cristianini N. Controlling the sensitivity of support vector machines [C]// Proc Int Joint Conf Artif Intell. Stockholm, Sweden: [s. n.], 1999: 55-60.
- [14] Li Zhongguo, Hou Jie, Wang Kai, et al. Application of fuzzy support vector machine on road type recognition [J]. Journal of Data Acquisition and Processing, 2014, 29(1): 146-151. (in Chinese)
- [15] Li P, Chan K, Fang W. Hybrid kernel machine ensemble for imbalanced datasets [C]// Proceedings of the 18th International Conference on Pattern Recognition. [S. l.]: IEEE Computer Society, 2006: 1108-1111.
- [16] Lin C F, Wang S D. Fuzzy support vector machines [J]. IEEE Transactions on Neural Network, 2002, 13(3): 466-471.

