Niu, Gang and Han, Tian and Yang, Bo-Suk and Tan, Andy C. C. (2007) Multi-agent decision fusion for motor fault diagnosis. *Mechanical Systems and Signal Processing* 21(3):pp. 1285-1299.

Mechanical Systems and Signal Processing

Multi-agent decision fusion for motor fault diagnosis

Gang Niu[a], Tian Han[a], Bo-Suk Yang[a*], Andy Chit Chiow Tan[b]

[a] *Intelligent Condition Monitoring & Diagnosis Research Center, School of Mechanical Engineering, Pukyong National University, San 100, Yongdang-dong, Nam-gu, Busan 608-739, South Korea*

[b]*School of Mechanical, Manufacturing and Medical Engineering, Queensland University of Technology, GPO Box 2343, Brisbane, QLD 4001, Australia*

*** Corresponding Author:**

Bo-Suk Yang, Professor

School of Mechanical Engineering, Pukyang National University

San 100, Yongdang-dong, Nam-gu, Busan 608-739, South Korea

Phone: +82-51-620-1604

Fax: +82-51-620-1405

E-mail: bsyang@pknu.ac.kr

Total number of pages: 25

Total number of tables: 6

Total number of figures: 10

Multi-agent decision fusion for motor fault diagnosis

Gang Niu[a], Tian Han[a], Bo-Suk Yang[a*], Andy Chit Chiow Tan[b]

[a] *Intelligent Condition Monitoring & Diagnosis Research Center, School of Mechanical Engineering, Pukyong National University, San 100, Yongdang-dong, Nam-gu, Busan 608-739, South Korea*

[b] *School of Mechanical, Manufacturing and Medical Engineering, Queensland University of Technology, GPO Box 2343, Brisbane, QLD 4001, Australia*

Abstract

Improvement of recognition rate is the ultimate aim for fault diagnosis researchers using pattern recognition techniques. However, unique recognition method can only recognize a limited classification capability which is insufficient for application. An ongoing strategy is the decision fusion techniques. In order to avoid the shortage of single information source coupled with unique decision method, a new approach is required to generate better results. This paper proposes a decision fusion system for fault diagnosis, which integrates data sources of different types of sensors and decisions of multiple classifiers. First, non-commensurate sensors data sets are combined using an improved sensor fusion method at a decision-level by using relativity theory. The generated decision vectors are then selected based on correlation measure of classifiers in order to find an optimal sequence of classifiers fusion, which can lead to the best fusion performance. Finally, multi-agent classifiers fusion algorithm is employed as the core of the whole fault diagnosis system. The efficiency of the proposed system was demonstrated through fault diagnosis of induction motors. The experimental results show that this system can lead to super performance when compared with the best individual classifier with single source data.

Keywords: Fault diagnosis; Multiple-sensors data fusion; Correlation measure; Multi-agent algorithm; Classifiers fusion

1. Introduction

With the rapid advancement in industry, machinery systems are becoming more and more complex and required constant attention. In spite of their robustness and reliability, they do fail occasionally and result in unpredicted downtime that can be very costly [1]. This has lead to the development of numerous condition monitoring and fault diagnosis techniques and has received considerable attentions. The researches in this area, especially in early fault diagnosis, are progressing at a fast pace and many relevant research papers have been published during the past decades.

In fault diagnosis of machinery, many traditional approaches are still in used such as vibration analysis [2], frequency analysis [3] and other methods suitable only for special machinery. Nevertheless, it is hard to use these diagnosis methods because the engineers need to have special knowledge of the machine system and lack of experience in increasing sophistication of large-scale maintenance.

Recently, the development of artificial intelligence (AI) techniques has lead to their application in fault diagnosis area. Many intelligent diagnostic systems have been employed to assist condition monitoring tasks by correctly interpreting the fault data, such as expert systems, artificial neural networks (ANNs), support vector machines and fuzzy logic systems, and the results of these techniques are promising [4-7]. However, many researches have shown that when applying individual decision system with a single data source which can only acquire a limited classification capability and may be not enough for a particular application.

Therefore, the application of decision fusion system (DFS) has received a lot of interests in recent years, and researchers have achieved considerable successes from this approach to solve complex pattern recognition tasks. DFS is also called multiple classifiers fusion (MCF), combination of classifiers, multiple experts and hybrid method. Due to the integration of different decisions from multiple classifiers, the technique can boost the accuracy of recognition. The MCF can be categorized into two classes; the static and the dynamic methods. The static combination strategies are simple but only concentrate on the output of the classifiers, such as majority voting [8], minimum and maximum [9], and average [10]. In comparison, the dynamic methods are more elaborate which take into account the information from the training phase on the behavior of the classifiers which include the Bayesian method, behavior-knowledge space (BKS) and Dempster-Shafer theory [11-13].

In this paper, we proposed the dynamic fusion algorithm known multi-agent which was initially proposed by Kou and Zhang [14],To date we are unable to find more references in this area. This algorithm endues each classifier with the property of an agent. That is, each classifier can complete the task of pattern recognition by itself and at the same time it can exchange information with each other for the purpose of improving the recognition rate. The cooperation of all classifier agents finally produces an outcome which is better than the result of an individual agent. In structure, the method combines the Bayesian and the majority voting methods. One of the main characteristics of this method is the definition of co-decision matrix for information exchange between the classifier agents. Hence, the technique belongs to dynamic fusion approach. Based on the multi-agent fusion algorithm, a decision fusion system for fault diagnosis is developed, which contains of three parts: multiple sensors fusion (MSF), classifier selection and MCF. In the proposed system, raw data is initially collected from multiple sensors and feature parameters are calculated. The generated feature data are then grouped as the original input of the system to be sent into each classifier for recognition. Next, the decision vectors are selected in terms of correlation among the features data in order to obtain the best fusion performance with least classifiers. Finally, the optimal team obtained is applied in decision fusion using the multi-agent fusion method. The rest of this paper is organized in sequence as: background knowledge of decision fusion methods, multi-agent modal, proposed system, experiment and results, conclusion and future work.

2. Background knowledge of decision fusion methods

This section covers a brief introduction of some basic concepts and notation in decision fusion, classifier selection based on correlation measure and some common methods of MCF in brief.

2.1. Multiple sensors fusion (MSF)

In terms of different fusion phases of measured information, MSF can be categorized into three levels: data-level, feature-level and decision-level [15].

Data-level fusion: All sensor data from a measured object is combined directly and the features are then calculated from the fused data. At this stage, a pattern recognition process is performed. Fusion of data at this level contains the most information and can deliver good results. However, the sensors used in this level must be commensurate. That means the measurement has to be the same or has similar physical quantities or phenomena. As a consequence, data-level applications are limited in real environment where there are many physical quantities having to be measured for synthesis analysis.

Feature-level fusion: At this level, the features are calculated from each sensor according to the type of raw data. Then, these non-commensurate sensors features are combined at the feature level. All features are combined in turn into a bigger single feature set, which are then used in a special classification modal such as neural network or cluster algorithm for decisions.

Decision-level fusion: In this structure, the processes of features calculation and pattern recognition are applied in sequence for single-source data obtained from each sensor. The decision vectors are then fused using decision-level fusion techniques such as voting strategy, Bayesian method, behavior knowledge space (BKS) and Dempster-Shafer theory. The flowchart for sensors fusion at decision level is shown in Fig. 1.

Fig. 1. Flowchart of the sensor fusion at decision level.

2.2. Correlation based classifier selection

It is essential for multiple classifiers fusion to have a proper method for classifier selection because the combination of different classifiers can affect fusion accuracy. When face with many classifiers and sensor data sets, the selection is often a problem before a final fusion strategy is employed. A proper classifier team should be robust and can generate the best fusion performance. It also should be optimal so that it can reduce the time for calculation and saving the data in the memory. Classifiers selection technique is an on-going active research area in recent years. Most of the selection methods are based on statistic theory such as: *Q* statistic, generalized diversity and agreement [16-18]. Among them, the degree of correlation is an interesting sub-direction belonging to agreement of classifiers. Many researchers have found the dependency between classifiers which can affect the fusion results. Goebel et al. [19] recommended an effective method for classifier selection based on calculating the correlation degree of n different classifiers and is shown in Eq. (1).

$$\rho_n = \frac{nN^f}{N - N^f - N^r + nN^f} \tag{1}$$

where, N^f means the number of samples which are misclassified by all classifiers, N^r means those samples which are classified correctly by all classifiers and N is the total number of experiments samples. Generally, smaller correlation degree ρ can lead to better performance of classifier fusion because the independent classifiers can give more effective information.

According to the correlation measurement principle, a team of classifiers needs to be selected and the steps of classifier selection can be summarized as:

Step 1: Select an appropriate performance measure as the initial evaluation criterion, such as accuracy rate which is the ratio of number of samples classified correctly to the total samples;

Step 2: Find the best performing classifier as the first classifier of the team;

Step 3: Calculate the correlation degree between the first classifier and the other classifiers respectively using Eq. (1);

Step 4: Select the classifier having the “low correlation” for fusion. A practical improvement in this paper is that when a similar low correlation degree appears for more than one classifier, the classifier that has the highest recognition rate is chosen;

Step 5: Repeat step 3 to step 4 between selected classifiers and the classifiers yet to be selected until all the classifiers are determined.

Finally, the optimal sequence of classifiers can be found.

2.3. Multiple classifiers fusion (MCF)

According to the characteristic of output information of the classifiers, classifier fusion methods can be divided into three styles [11]:

• The abstract style: a classifier C only generates a single class output with an input x;

• The rank style: a classifier C ranks all classes in a queue and chooses the top one;

• The measurement style: a classifier C evaluates each class using a probability value that the x subjects to the class.

Among the styles mentioned above, the required information for a classification increases in sequence and the abstract style contains the least information while the measurement style contains the most information. Accordingly, the classifier fusion algorithms of the measurement style can produce the best results. However, the classifiers being able to output each class’s probability are seldom available. As a result, the classifier fusion algorithms belonging to an abstract style are commonly used. The methods used in abstract style mainly consists of voting, Bayesian, BKS and Dempster-shafer theory et al.

In this section, we briefly introduce some popular methods of MCF used at abstract level: majority voting, Bayesian belief and behavior-knowledge space (BKS). A brief comment will be given for each method.

Majority voting method: Voting may be the easiest method in decision fusion [20]. There are various voting strategies such as unanimity, majority and Borda count. Among them, majority voting is the the most popular method. In this method the class voted by most of classifiers will be regarded as the result of fusion decision. If no class won more than half of the votes, the input is rejected. The method is simple and easy to realize. Nevertheless, it does not consider the characteristics of each classifier which are related with the performance of classifier fusion.

Bayesian belief method: To compare with voting method, Bayesian belief algorithm [21] offers a soft fusion strategy which is more dynamic. This method is based on the assumption of mutual independency of classifiers and considers the error of each classifier. For a multiple class recognition problem with classes 1 through *M*, the error for *k*th classifier can be represented by a two-dimensional confusion matrix shown in eg (2), where $k = 1, \ldots, K$,

$$PT_k = \begin{bmatrix} n_{11} & n_{12} & \cdots & n_{1M} \\ n_{21} & n_{22} & \cdots & n_{2M} \\ \cdots & \cdots & & \cdots \\ n_{M1} & n_{M2} & \cdots & n_{MM} \end{bmatrix} \quad (2)$$

where the rows stand for classes: $c_1, \ldots, c_M$ which consist of input sample x, and the columns indicate the classes which consist of the input sample assigned by the classifier e_k . The element n_{ij} illustrates the input samples from class c_i while assigned to class c_j by classifier e_k. On the basis of the confusion matrix, belief measure of recognition can be calculated for each classifier by the belief function:

$$Bel(x \in c_i / e_k(x)) = P(x \in c_i / e_k(x) = j_k) \quad (3)$$

where $i, j = 1, \ldots, M$ and

$$P(x \in c_i / e_k(x) = j_k) = n_{ij}^{(k)} \Big/ \sum_{i=1}^{M} n_{ij}^{(k)} \quad (4)$$

Combining the belief measures of all fusion classifiers will result in the final belief measure of the multiple classifier system and is shown as follows:

$$Bel(i) = P(x \in c_i) \frac{\prod_{k=1}^{K} P(x \in c_i / e_k(x) = j_k)}{\prod_{k=1}^{K} P(x \in c_i)} \quad (5)$$

The highest combined belief measure: $Bel(i)$ is chosen as the final classification decision. However, one of the significant limitations of Bayesian method is that it requires mutual independencies among multiple classifiers which doesn't usually hold in real application [22].

Behavior-knowledge space (BKS): When compared with the Bayesian method, this method does not emphasize independence of the decisions made by classifiers. BKS [23] is a *k*-dimensional space where each dimension corresponds to a single classifier. Each classifier could produce N+1 crisp decisions, N class labels and one rejection decision. The intersection of the decisions of every single classifier occupies one unit of the

BKS and each unit contains three elements: the total number of incoming samples, the best representative class and the number of incoming samples of each class. The unit, which is the intersection of the classifiers' decisions of the current input, is called the focal point. For an unknown test sample, the decision of the individual classifiers indexes a unit of BKS and the unknown sample is assigned to the class with the most training samples in the BKS unit. In the BKS method, large numbers of training data are required to build the BKS so that the lack of enough training data often is a problem.

3. Multi-agent modal

In this section, we intend to introduce multi-agent fusion algorithm in detail. The relationship between multi-agent with majority voting and Bayesian belief will also be discussed.

This modal, proposed by Kou and Zhang [14], absorbs the properties of multi agent system (MAS) into the algorithm of classifiers fusion. It integrates Bayesian belief at the starting phase and majority voting at the final phase. A co-decision matrix is set up for information exchange between the classifier agents so that Bayesian belief matrix can be modified dynamically until a predetermined criterion is satisfied. Finally, a combination decision is made. The flowchart of multi-agent algorithm is shown in Fig. 2:

Fig. 2. Flowchart of multi-agents fusion algorithm.

Confusion matrix is first created as a training parameter, which accumulates the errors of each classifier. Then an initial belief matrix can be calculated easily for each test sample based on the training parameter. In the initial belief matrix, the rows indicate kth classifier, where $k = 1, \ldots, K$, and columns stand for class $c_1, \ldots, c_M$. The elements in kth row show the probabilities of an input sample x belonging to different classes estimated by kth classifier using Eq. (5). The processes to calculate the confusion matrix and initial belief matrix are based on Bayesian belief method.

After calculating the two matrixes, a five-dimensional co-decision matrix is required as the last training parameter. Each cell in the co-decision matrix stands for decision correlation between two classifiers, which is calculated through following equation:

$$
\begin{aligned}
d_{j_1,j_2,i,k_1,k_2} &= P\left(E = i \middle| e_{k_1} = j_1, e_{k_2} = j_2\right) \\
&= \frac{\left|\left\{x \middle| E(x) = i, e_{k_1}(x) = j_1, e_{k_2}(x) = j_2, \forall x \in U_2\right\}\right|}{\sqrt{\left|\left\{x \middle| E(x) = i, e_{k_1}(x) = j_1, \forall x \in U_2\right\}\right|} \cdot \sqrt{\left|\left\{x \middle| E(x) = i, e_{k_2}(x) = j_2, \forall x \in U_2\right\}\right|}}
\end{aligned}
\tag{6}
$$

where $E(x) = i$ is the expectation of input sample x, that is, the real class of x range from c_1 to c_M; j_1 and j_2 respectively stands for the decision of classifiers k_1 and k_2 where $k_1 \neq k_2$ and U_2 are the training samples set of the fusion modal. Each element in the matrix shows the probability of classifier k_1 classifying x as j_1 class and classifier k_2 assigning x as j_2 class.

After obtaining the necessary statistical parameters, the confusion matrix and co-decision matrix, the initial vote rates for input sample x can be calculated. The column class corresponding to the maximum of kth row of belief matrix is regarded as kth classifier's decision. By doing this, the belief matrix can be transformed into a decision label vector. Then the voting strategy can be employed. The original vote rate of each class is calculated for input x.

Next, an accordance criterion is set to compare with the maximum vote rate. Higher accordance criterion is set to allow for less different decisions. If the maximum vote rate is less than the threshold, a repeating modification scheme is fired and the original belief degrees have to be modified dynamically. The exchange of information of the two classifiers based on the co-decision matrix is added to the vote rates using following equation:

$$b_{ki} = b_{ki} + \frac{1}{K} \sum_{k_n=1, k_n \neq k}^{K} d_{j, j_n, i, k, k_n} \cdot \sqrt{b_{ki} \cdot b_{k_n, i}} \tag{7}$$

where the original belief matrix b is acquired by the confusion matrix based on Eq. (5); K is the number of total fusion classifiers; b_{ki} represents belief probability of classifier k to class i and d_{j, j_n, i, k, k_n} which is the exchange of information between kth classifier and k_nth classifier.

After the modified belief degree, a normalization process is required to bring the summation of each row probabilities of new belief matrix equals to one. Then the new belief matrix b can be transformed into a decision vector, so the new vote rates can be acquired. If the maximum vote rate is still less than the predetermined criterion, the repeating modification process will continue until the maximum vote rate reaches the threshold. Finally, an improved majority voting method is utilized for the output of fusion decision, which only chooses the class gaining the most votes as the fusion decision and needs not beyond half of votes as original voting strategy.

4. Proposed system

Traditionally, multiple classifier fusion contains two main steps. First, each classifier assigns a single source of data set respectively. Then the acquired decision vectors group is sent into the fusion modal for combination and a fused decision is generated. In this paper, the significant characteristics of the proposed system include three parts: non-commensurate sensors data fusion at decision-level, classifiers selection based on correlation measure and multi-agents based classifier fusion algorithm. For the purpose of clarity, we applied this system to fault diagnosis of rotating machinery. The system can be extended to other diagnosis areas. The diagram of the proposed fusion decision system is shown in Fig. 3.

Fig. 3. Diagram of the proposed fusion decision system.

The procedure of this system can be summarized as follows:

Step 1: Sensors data fusion. For fault diagnosis of rotating machinery such as electric motor and generator, stator current signal analysis is equally important as vibration analysis. So, data fusion of the two types of sensors is expected to provide more accurate information to multi-classifiers system. Usually, the classes of a vector are diverse for different classifiers with a same data set. Utilizing the relativity theory, we noticed that the outputs could also be changed for different data sets which are classified by a same classifier. For instance, two data sets classified by a same classifier separately can be seen as one data set assigned by two different classifiers. In this case, k classifiers with i data sets can be regarded as $i \times k$ classifiers with one data set. Here, v_i and c_i indicate decision vectors of vibration and current data respectively assigned by classifier i, where $i = 1 \ldots k$.

Step 2: Classifier selection. A classifier selection process is executed for the $i \times k$ classifiers (a collection of decisions vectors) by correlation measure method. As a result, an optimal team of classifiers containing classes of information from both the vibration and the current signals is formed to improve classification accuracy.

Step 3: Decision fusion. After classifier selection, multi-agent classifiers fusion algorithm is employed in the decision fusion system.

5. Experiments and results

In order to demonstrate the effectiveness of the proposed system, experiments were carried out using a self-designed test rig which consists of a motor, pulleys, belt, shaft and fan with changeable blade pitch angle, as shown in Fig. 4.

Fig. 4. Experiment apparatus.

The test specimens consist of six 0.5kW, 60Hz, 4-pole induction motors to create the data needed. This motor was set to operate at full-load conditions. [24] One of the motors is normal (healthy), which is used as a benchmark for comparing with faulty motors. The others are faulty motors with broken rotor bar, bowed rotor, bearing outer race fault, rotor unbalance, adjustable eccentricity motor (misalignment) and phase unbalance, as shown in Fig. 5. The conditions of faulty induction motors are described in Table 1. The load of the motors can be changed by adjusting the blade pitch angle or the number of the blades.

Fig. 5. Faults on the induction motors.

Table 1 Description of faulty induction motors

Three AC current probes and three accelerometers were used to measure the stator current of three phase power supply and vibration signals in horizontal, vertical and axial directions for evaluating the fault diagnosis system, respectively. The maximum frequency of the signals was 3 kHz, the number of sampled data was 16384 and measured time was 2.1333 seconds. The collected signal waveforms of vibration and current are shown in Fig. 6. For each condition 40 samples were measured, 20 of them were used for training parameters of the classifiers, 10 for training parameters of the multi-agent fusion modal and the other 10 for the test.

After data acquisition, a process of features calculation was exerted. 21 feature parameters of the time domain (10 parameters), frequency domain (3 parameters) and regression estimation (8 parameters) are acquired from each sensor using the collected three-direction vibration and three-phase currents signals shown in Table 2. A total of 63 features (21 parameters × 3 signals) are calculated respectively for vibration and current signals. Table 2 shows the feature parameters

Fig. 6. Vibration and current signals of each condition.

Table 2 Feature parameters

Next, six classifiers were utilized to classify the calculated features of vibration and current. The utilized classifiers are described as follows:

- *Support Vector Machine (SVM):* The SVM [25] is a machine learning algorithm based on the statistical learning theory and compared with other classifier, this technique can lead to good recognition rate with a few training samples. Kernel function is an important parameter for SVM classifier which contains linear, polynomial, Gausian RBF and sigmoid.

- *Linear Discriminant Analysis (LDA):* LDA [26] is popular for features drop-dimension and also can be used for classification. It projects features from parametric space onto feature space through a linear transformation matrix. This classifier can be efficiently computed in the linear case even with large data sets.
- *k Nearest Neighbors (k-NN):* *k*-NN [27] is an easy and effective classifier, the aim is to find the nearest neighbors of an unidentified test pattern within a hyper-sphere of pre-defined radius in order to determine its true class. It can detect a single or multiple number of nearest neighbors.
- *Improved Iterative Scaling (IIS):* IIS [28] is a generating weight value of maximum entropy model classifier. It finds the Conditional Exponential Model weights that define the maximum entropy classifier for a given feature set and training corpus.
- *Gaussian Mixture Model (GMM):* GMM [29] is a classifier based on Gaussian component functions. The linear combination of the Gaussian functions is capable of representing a large class of sample distributions. In principle, it is a compromise between performance and complexity. The Gaussian mixture has remarkable capability to model the irregular data.
- *Learning Vector Quantization (LVQ):* LVQ [30] is a typical classifier in neural network (NN) proposed by Kohonen. It is a simple and intuitive, though very successful prototype-based clustering algorithm. It combines the simplicity of self-organizing learning with the accuracy of supervised training algorithms.

The relevant parameters setup for these classifiers can be found in Table 3.

Table 3 The parameters of individual classifier

Fig. 7 shows the training accuracy of the six classifiers. In our experiment, the classification accuracy is evaluated using a ratio of number of the samples classified correctly to the total samples. It can be seen that the classification accuracy of vibration signal is far better than the ones obtained from stator current signal. The best classification accuracy using vibration data is 0.8667 by SVM and *k*-NN classifiers while 0.6778 when using current signal by SVM. As far as performance of the six classifiers is concerned, SVM and *k*-NN produced superior results and followed by LDA and IIS. GMM and LVQ are not sutiable in this work for vibration or current signals. This implies that the scatter of training samples does not fit for the two types classifiers. In this experiment, we did not avoid using bad classifications because in practice, the classifiers that are available are often predetermined and the measured singals are generally different. It is almost impossible that all these classifiers will achieve the best performance at the same time. In addition, M. Petrakos et. al [18] also emphasis that the fusion of a group of pure good or bad classifiers to achieve accuracy rates may not necessary improve the results. So GMM and LVQ classifiers still require further work. In this experiment, SVM was selected as the best individual classifier and regarded as the top classifier among the optimal team of classifiers.

Fig. 7. Comparison of single classification accuracy.

5.1. Sensors data fusion

In the experiment, the measured information of current probes and accelerometers are non-commensurate, therefore, the fusion of which was laid on decision-level. The six classifiers were used to classify the features data of vibration samples and the generated decision vectors were named as vectors 1 to 6 in sequence. The process was then repeated for the current data with the result vectors named from 7 to 12. Next, the 12 decision vectors were sent for classifiers selection in order to find the best sequence of classifier fusion. Finally, multi-agent algorithm was utilized for decision fusion. The comparison of classification accuracy between the fusion data and single data using the vibration or current signal is shown in Fig. 8. It can be seen that the highest accuracy of single source data, after classifiers selection and combination, is 0.922 for vibration data and 0.722 for current data. Table 4 shows the fusion performance of vibration and current data. The fusion sequence of classifiers is acquired by the selection step for the 12 decision vectors. According to the selected sequence, the decision vectors of different numbers of classifiers are fused using step of classifiers fusion. The number 1 to 12 means the sequence of classifiers to be fused and the corresponding location of number i, (i =1,…,12) shows the fusion accuracy using the decision vectors of number 1 to number i. For example, the fusion accuracy of 0.956 in Table 3 is the result of fusing SVM to vibration data, LVQ to current data, *k*-NN to current data, LVQ to vibration data and *k*-NN to vibration data. Table 4 shows that the best accuracy rate after fusion of vibration and current data is 0.989.

Fig. 8. Comparison of classification fusion accuracy.

Table 4 Fusion performances of vibration and current data

5.2. Selection of classifiers

This section illustrates the use of correlation measure method mentioned above for sequence selection of classifier fusion. The result is shown in Fig. 9. The highest classification accuracy using majority voting method is 0.9111 with classifiers selection and 0.8778 if the selection process is neglected. A bad team of classifiers may even lead to worse performance than best single classifier due to the fusion of error decisions. The selection results for different numbers of classifiers are shown in Table 5.

Fig. 9. Effect of classifiers selection (majority voting).

Table 5 Sequence of classifier fusion

5.3. Multi-agent method

This section describes the use of multi-agent method. We applied the same process of multiple sensor fusion (MSF) and classifier selection. The final outputs of a group of decision vectors were sent to three different classifiers fusion algorithms for comparison, namely, majority voting, Bayesian belief and multi-agent. The results are shown in Fig. 10. In multi-agent method, accordance criterion is a sensitive parameter and can affect the final accuracy. The larger the data set, the longer executing time it takes but normally produces better accuracy. In order to find an appropriate accordance criterion, we tried the values from 0.5 to 1 in step of 0.05 and the results are shown in Table 6. When the value was increased to 0.65, best performance was achieved at the cost of fusion of 11 classifiers. With further increase of accordance criterion, the number of fusion classifiers required for highest accuracy gradually reduced. At the accordance criterion of 0.9, the number of required classifiers is only 7.

The maximum fusion accuracy for multi-agent method was 0.989, while the accuracy using Bayesian belief and majority voting methods were 0.989 and 0.911 respectively. In Fig. 10, it can be seen that the multi-agent and Bayesian belief methods are far superior than the majority voting strategy, because the former methods involve dynamic fusion. Multi-agent method is slightly better than the Bayesian strategy when the classifiers fused are insufficient. However, to compare the results of the two fusion methods, the highest accuracy rates are same that could be due to the close correlation of decision vectors. Another possible reason might be that when the fusion accuracy using Bayesian method is high enough, any further improvement by multi-agent algorithm is hard to achieve.

Fig. 10. Comparison of multi-agent, Bayesian and majority voting method.

Table 6 Relationship of accordance criterion, number of classifiers fused and accuracy

6. Conclusions and future work

In this paper, a decision fusion system for fault diagnosis is proposed which has the following characters: non-commensurate sensor data fusion at decision-level in terms of theory of relativity, correlation measure based classifiers selection and classifier fusion based on multi-agent algorithm. The effectiveness of the proposed methodology was tested with examples of motor fault diagnosis. The core of this fusion diagnosis system, the multi-agent classifiers fusion algorithm, is tested and proposed because of its capability of information exchange and combination of Bayesian belief and majority voting methods.

Based on the decision fusion framework, our future work will concentrate at the following three parts:

- Extend scope of data fusion containing of not only the signals of commensurate and non-commensurate sensors but also the ones of transient and static states of operating machinery. Integration techniques of different levels will be researched.
- Investigate other methods of classifier selection and the effectiveness of the methods will be evaluated.
- Compare multi-agent algorithm with other methods of classifier fusion.

Acknowledgment

This work is supported by the Center for Advanced Environmentally Friendly Energy Systems, Pukyong National University, Korea (Project number: R12-2005-001-01006-0) and partially supported by the Brain Korea 21 Project in 2005.

References

[1] B.S. Yang, S.K. Jeong, Y.M. Oh, A.C.C. Tan, Case-based reasoning system with Petri nets for induction motor fault diagnosis, Expert Systems with Applications 27 (2) (2004) 301-311.

[2] W.R Finley, M.M. Hodowanec, W.G. Holter, An analytical approach to solving motor vibration problems, IEEE Trans. on Industry Applications 36 (5) (2000) 1467-1480.

[3] S. K. Lee and P. R. White, Higher-order time–frequency analysis and its application to fault detection in rotating machinery, Mechanical Systems and Signal Processing 11 (4) (1997) 637-650.

[4] Y. Qian, X. Li, Y. Jiang, Y. Wen, An expert system for real-time fault diagnosis of complex chemical processes, Expert Systems with Applications 24 (4) (2003) 425-432.

[5] B. Samanta and K.R. Al-Balushi, Artificial neural network based fault diagnostics of rolling element bearings using time-domain features, Mechanical System and Signal Processing 17 (2) (2003) 317-328.

[6] M. Ge, R. Du, G. Zhang, Y. Xu, Fault diagnosis using support vector machine with an application in sheet metal stamping operations, Mechanical System and Signal Processing 12 (1) (2004) 143-159.

[7] S.O.T. Ogaji, L. Marinai, S. Sampath, R. Singh, S.D. Prober, Gas-turbine fault diagnostics: a fuzzy-logic approach, Applied Energy 82 (1) (2005) 81-89.
[8] D. Ruta and B. Gabrys, An overview of classifier fusion methods, IEEE Computing and Information System (2000) 1-10.
[9] L. Kuncheva, A theoretical study on six classifier fusion strategies, IEEE Transactions on Systems 24 (2) (2002) 281-286.
[10] M. Taniguchi and V. Tresp, Averaging regularized estimators, Neural Computation 9 (1997) 1163-1178.
[11] L. Xu, A. Krzyzak, C. Suen, Methods of combining multiple classifiers and their applications to handwriting recognition, IEEE Transactions on Systems, Man, and Cybernetics 22(3) (1992) 418-435.
[12] Y.S. Hang, C.Y. Suen, The behavior-knowledge space method for combination of multiple classifiers, IEEE (1993) 347-352.
[13] B.S. Yang, K.J. Kim, Application of Dempster-Shafer theory in fault diagnosis of induction motors using vibration and current signals, Mechanical Systems and Signal Processing 20(2) (2006) 403-420.
[14] J.B. Kou and C.S. Zhang, Multi-agent based classifier combination, Chinese Journal of Computers (2003) 1-5. (in Chinese)
[15] D.L. Hall, J. Llinas, An introduction to multisensor data fusion, IEEE Digital Object Identifier 85 (1) (1997) 6-23.
[16] L.I. Kuncheva, C.J. Whitaker, C.A. Shipp, R.P.W. Duin, Limits on the majority vote accuracy in classifier fusion, Pattern Analysis and Applications 6 (2003) 22-31.
[17] D. Partridge, W. Krzanowski, Software diversity: practical statistics for its measurement arid exploitation, Information and Software Technology 39 (10) (1997) 707-717.
[18] M. Petrakos and J.A. Benediktsson, The effect of classifier agreement on the accuracy of the combined classifier in decision level fusion, Transaction on Geoscience and Remote Sensing 39 (11) (2001) 2569-2546.
[19] K. Goebel, W.Z. Yan, W. Cheetham, A method to calculate classifier correlation for decision fusion. Proceedings of Decision and Control (2002) 135-140.
[20] R. Battiti and A. Colla, Democracy in neural nets: Voting schemes for classification, Neural Networks, 7(4) (1994) 691-707.
[21] L. Lam and C. Suen, Optimal combination of pattern classifiers, Pattern Recognition Letters, 16 (1995) 945-954.
[22] E. Kim, W. Kim, Y. Lee, Combination of multiple classifiers for the customer's purchase behavior prediction, Decision Support System 34 (2002) 167-175.
[23] Y. Huang, K. Liu, C. Suen, The combination of multiple classifiers by a neural network approach, Journal of Pattern Recognition and Artificial Intelligence 9 (1995) 579-597.
[24] T. Han, B.S. Yang, J.M. Lee, A new condition monitoring and fault diagnosis system of induction motors using artificial intelligent algorithms, Proceedings of IEEE Int. Electric Machines and Drives Conference (2005) 1-8.
[25] B.S. Yang, T. Han, W.W. Hwang, Fault diagnosis of rotating machinery based on multi-class support vector machines, Journal of Mechanical Science and Technology 19 (3) (2005) 846-859.
[26] T. Han, Development of a Feature Based Fault Diagnostics System and Its Application to Induction Motors, Ph.D thesis, Pukyong National University, South Korea(2005).
[27] S. Singh, J. Haddon, M. Markou, Nearest-neighbor classifiers in natural scene analysis, Pattern Recognition 34 (2001) 1601-1612.
[28] A. Berger, The improved iterative scaling algorithm: A gentle introduction (1997),

http://www.cs.cmu.edu/afs/~aberger/www/ps/scaling.ps.
[39] Q.Y. Hong and S. Kwong, A genetic classification method for speaker recognition, Engineering Applications of Artificial Intelligence 18(1) (2005) 13-19.
[30] Th. Villmann, F. Schleif, B. Hammer, Comparison of relevance learning vector quantization with other metric adaptive classification methods, Neural Networks (2005) (in press) 1-13.

List of tables

List of figures

Table 1

Description of faulty induction motors

Fault condition	Fault description	Others
Broken rotor bar	Number of broken bar: 12 ea	Total number of 34 bars
Bowed rotor	Maximum bowed shaft deflection: 0.075 mm	Air-gap: 0.25 mm
Faulty bearing	A spalling on the outer raceway	#6203
Rotor unbalance	Unbalance mass (8.4 g) on the rotor	
Eccentricity	Parallel and angular misalignments	Adjusting the bearing pedestal
Phase unbalance	Add resistance to one phase	8.4%

Table 2

Feature parameters

Signals	Position	Feature parameters		
		Time domain	Frequency domain	Auto regression
Vibration	Vertical Horizontal Axial	• Mean • RMS • Shape factor • Skewness • Kurtosis • Crest factor • Entropy error • Entropy estimation • Histogram lower • Histogram upper	• Root mean square frequency • Frequency center • Root variance frequency	• AR coefficients ($a_1 \sim a_8$)
Current	Phase A Phase B Phase C			

Table 3

Parameters of individual classifier

Classifier	SVM	*k*-NN	LVQ	IIS
Parameters setup	Linear kernel function, Euclidean distance type	k = 3	Number of neurons = 20, epochs = 50	Number of iterations = 50

Table 4

Fusion performance of vibration and current data

Classifier	Signal	SVM	LVQ	*k*-NN	LDA	IIS	GMM
Fusion sequence of classifiers	Vibration	1	4	5	7	10	11
	Current	6	2	3	8	9	12
Fusion accuracy	Vibration	0.867	0.944	0.956	0.989	0.989	0.989
	Current	0.967	0.867	0.944	0.989	0.989	0.989

Table 5

Results of optimal sequence of classifier fused

Numbers of classifiers selected	Serial number of classifiers											
1	1											
2	1	12										
3	1	12	9									
4	1	12	9	6								
5	1	12	9	6	3							
6	1	12	9	6	3	7						
7	1	12	9	6	3	7	2					
8	1	12	9	6	3	7	2	8				
9	1	12	9	6	3	7	2	8	10			
10	1	12	9	6	3	7	2	8	10	4		
11	1	12	9	6	3	7	2	8	10	4	5	
12	1	12	9	6	3	7	2	8	10	4	5	11

Table 6

Relationship of accordance criterion, number of classifiers fused and accuracy

Accordance criterion ρ	Number of classifiers fused											
	1	2	3	4	5	6	7	8	9	10	11	12
	Accuracy											
0.50	0.867	0.867	0.900	0.944	0.956	0.967	0.978	0.978	0.967	0.956	0.956	0.978
0.55	0.867	0.867	0.900	0.944	0.956	0.967	0.978	0.978	0.967	0.956	0.978	0.978
0.60	0.867	0.867	0.900	0.944	0.956	0.967	0.978	0.978	0.978	0.978	0.978	0.978
0.65	0.867	0.867	0.900	0.944	0.956	0.967	0.978	0.978	0.978	0.978	0.989	0.978

0.70	0.867	0.867	0.944	0.944	0.956	0.967	0.978	0.978	0.978	0.989	0.989	0.978
0.75	0.867	0.867	0.944	0.944	0.956	0.967	0.978	0.989	0.978	0.989	0.989	0.989
0.80	0.867	0.867	0.944	0.944	0.956	0.967	0.978	0.989	0.989	0.989	0.989	0.989
0.85	0.867	0.867	0.944	0.944	0.956	0.978	0.978	0.989	0.989	0.989	0.989	0.989
0.90	0.867	0.867	0.944	0.944	0.956	0.978	0.989	0.989	0.989	0.989	0.989	0.989
0.95	0.867	0.867	0.944	0.944	0.956	0.978	0.989	0.989	0.989	0.989	0.989	0.989

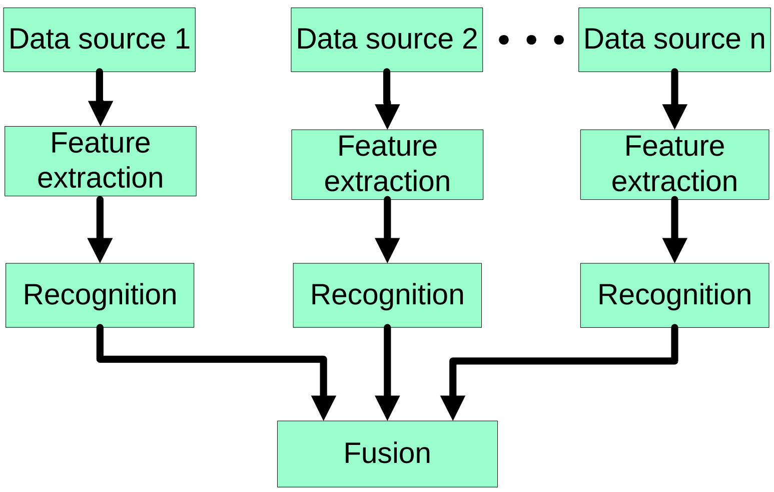


Fig. 1. Diagram of the sensors fusion at decision level.

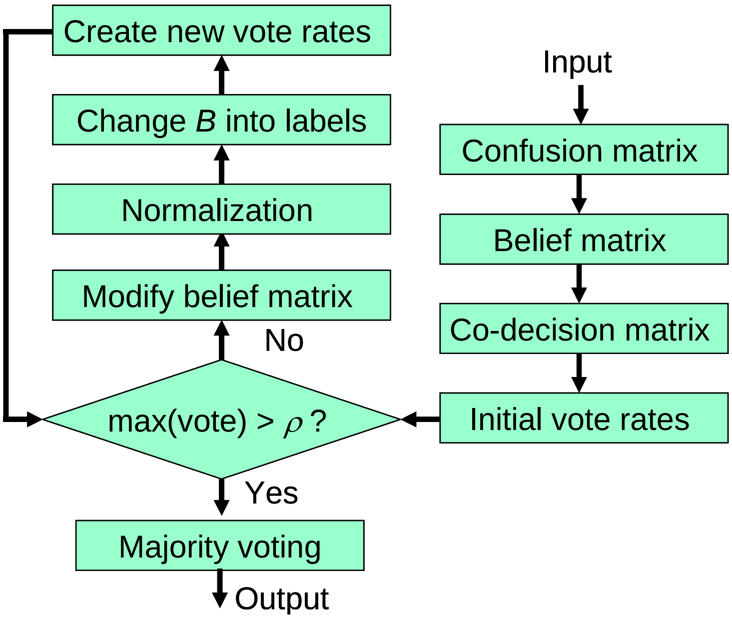


Fig. 2. Flowchart of multi-agents fusion algorithm.

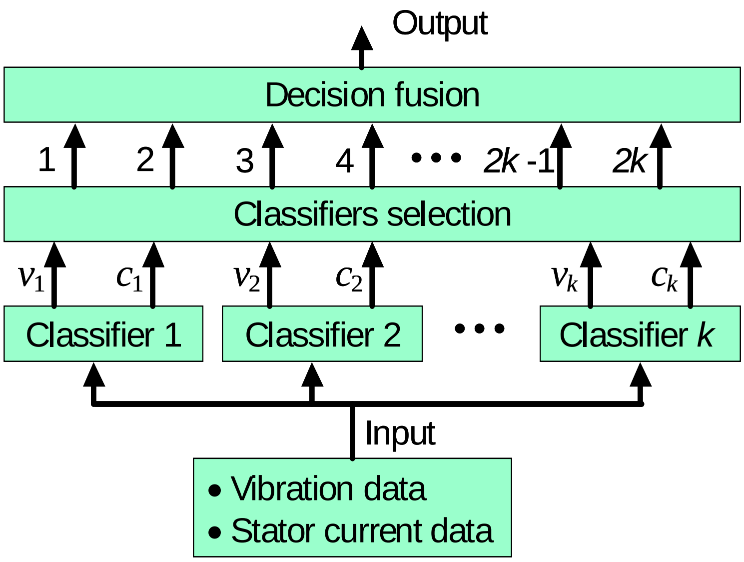


Fig. 3. Diagram of the proposed fusion decision system.

Fig. 4. Experiment apparatus.

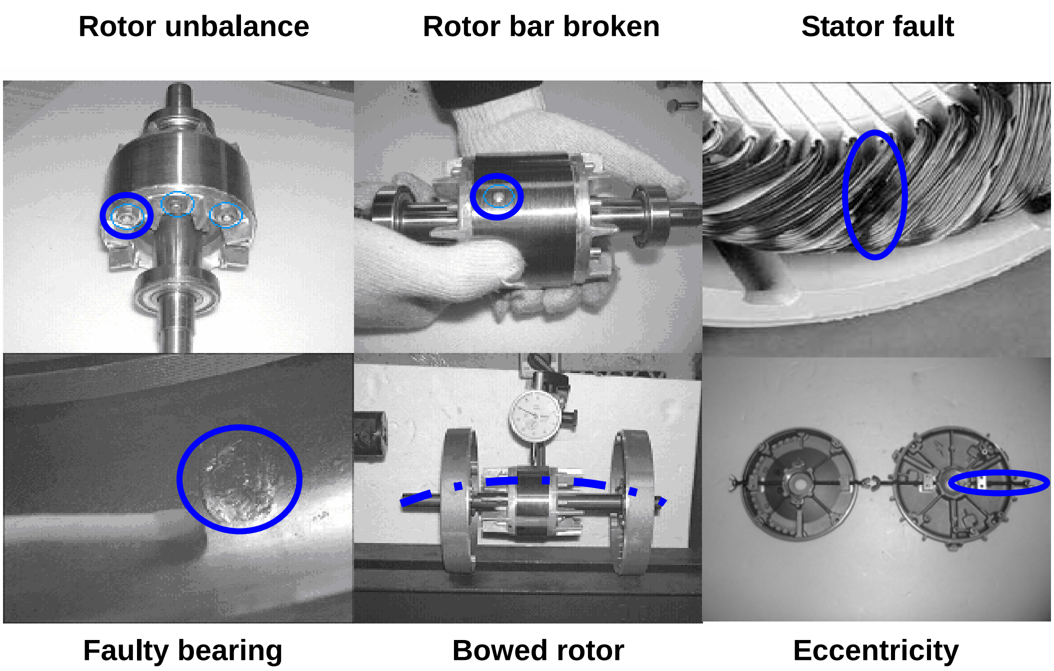
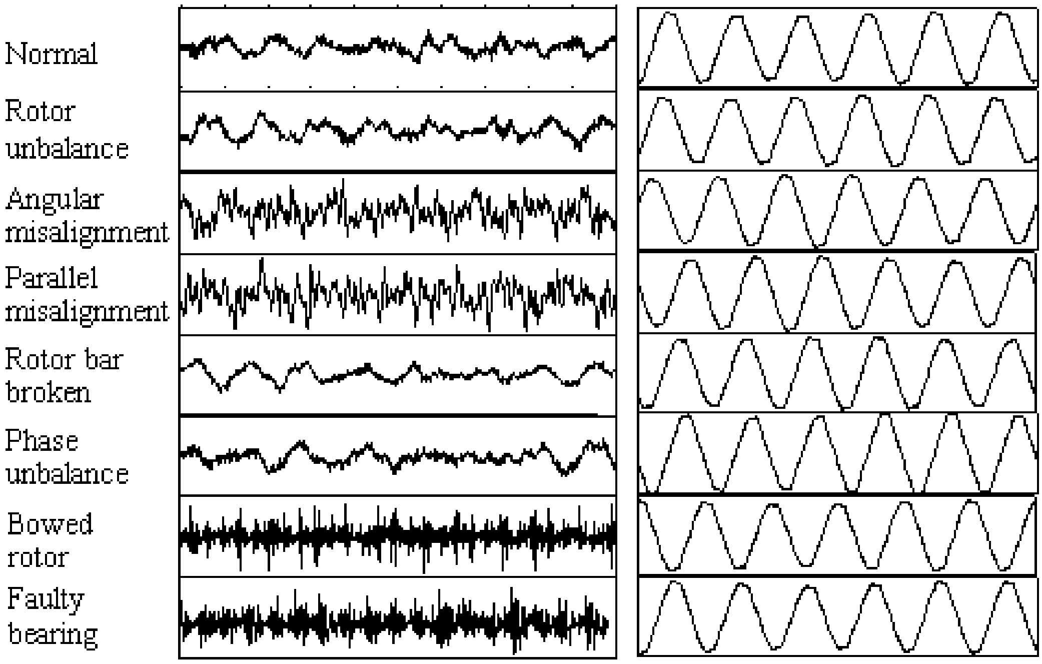


Fig. 5. Faults on the induction motors.



(a) Vibration signal (b) Current signal

Fig. 6. Vibration and current signals of each condition.

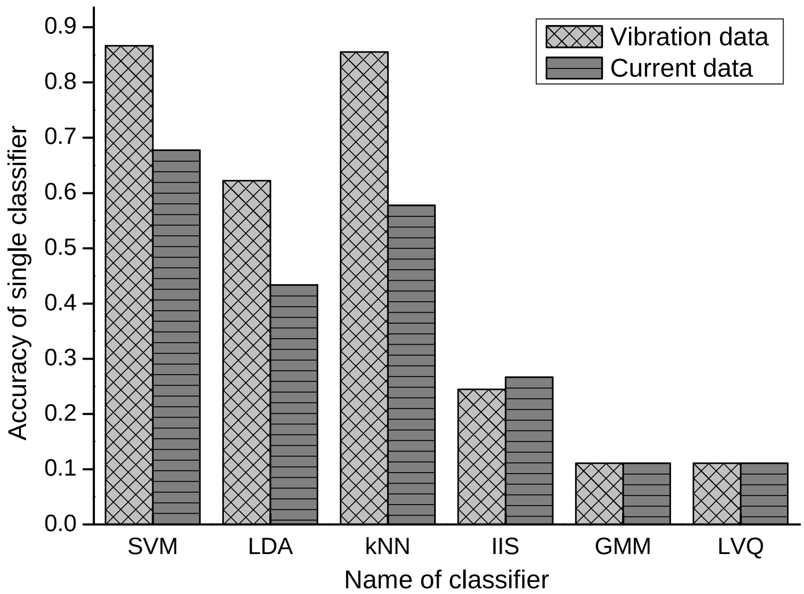


Fig. 7. Comparison of single classification accuracy.

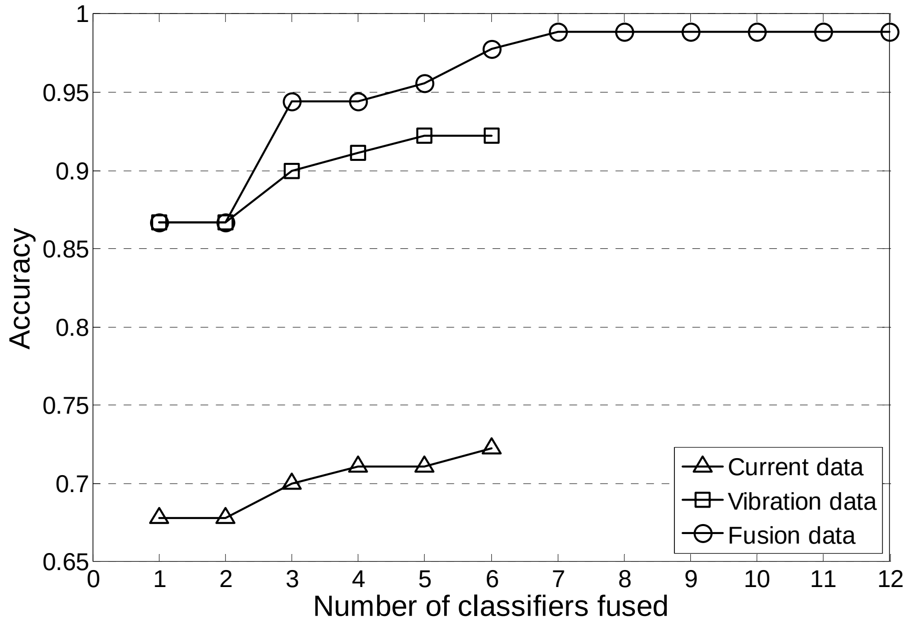


Fig. 8. Comparison of classification accuracy of fusion data and single data.

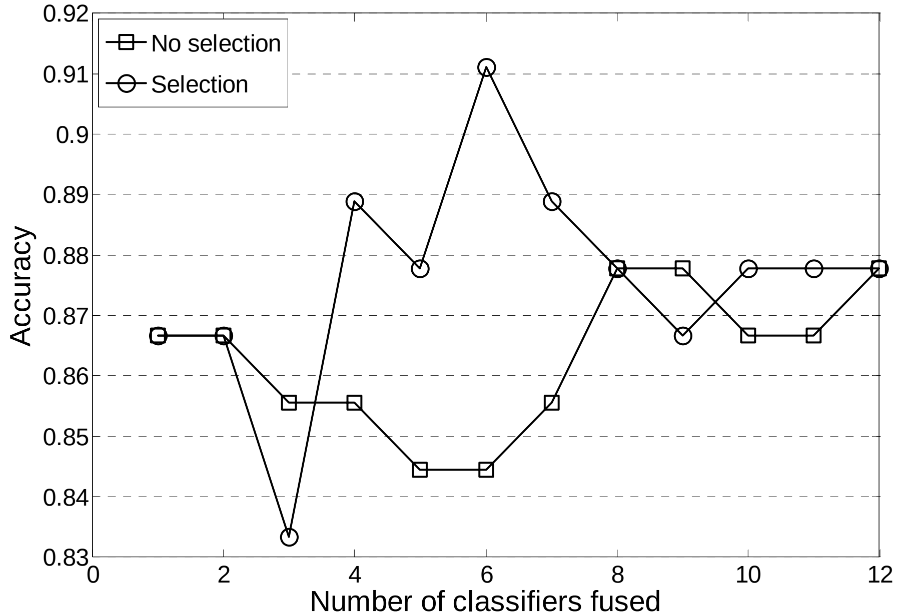


Fig. 9. Effect of classifiers selection (majority voting).

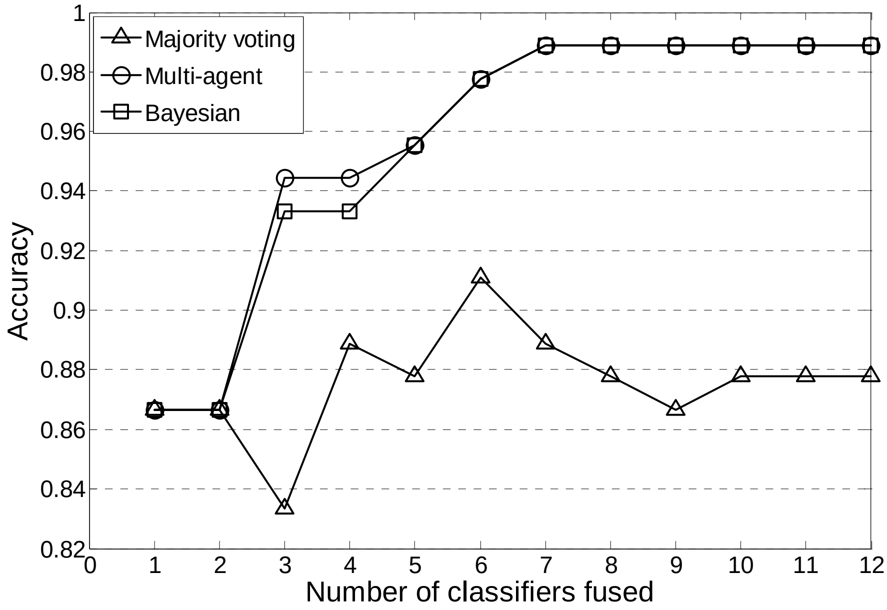


Fig. 10. Comparison of multi-agent, Bayesian and majority voting method.