

Multi-class SVM 을 이용한 회전기계의 결함 진단 Fault diagnosis of rotating machinery using multi-class support vector machines

황원우*·양보석†

Won-Woo Hwang and Bo-Suk Yang

Key Words : Fault diagnosis(결함진단), Kernel Function(커널함수), Sequence Minimize Optimization(SMO), Support Vector Machines(SVM), Rotating Machinery(회전기계)

ABSTRACT

Condition monitoring and fault diagnosis of machines are gaining importance in the industry because of the need to increase reliability and to decrease possible loss of production due to machine breakdown. By comparing the vibration signals of a machine running in normal and faulty conditions, detection of faults like mass unbalance, shaft misalignment and bearing defects is possible. This paper presents a novel approach for applying the fault diagnosis of rotating machinery. To detect multiple faults in rotating machinery, a feature selection method and support vector machine (SVM) based multi-class classifier are constructed and used in the faults diagnosis. The results in experiments prove that fault types can be diagnosed by the above method.

기호설명

w	가중치 벡터(weight vector)
x	입력 벡터(input vector)
b	한계값(threshold value)
$(w \cdot x)$	벡터 w 와 x 의 내적
y_i	출력 값 (+1 또는 -1)
$\ w\ $	벡터 w 의 놈(norm)
α	Lagrange multipliers
L	Lagrangian
$\Phi(x)$	특징공간으로의 변환 함수
$K(x, y)$	Kernel 함수(= $\Phi(x) \cdot \Phi(y)$)
σ	RBF kernel 의 폭(width)
E	예측 오차

1. 서론

산업의 발전과 기술의 진보에 따라 각종 산업용 회전기계는 대용량, 고속화, 고 하중으로 되고, 이는 가혹한 운전조건으로 작용하여 많은 스트레스를 기계에 부가하게 된다. 제품의 생산 효율과 품질의 중요성과 아울러 안정적인 기능 유지, 보수 및 진단에 대한 인식이 높아지고 있으며 이미 커다란 경제적인 영역을 형성하고 있다.⁽¹⁾

회전기계는 발전플랜트 등 각종 산업 현장에서 광범위하게 사용되고 있으며 중요한 역할을 수행하기 때문에, 이상이 발생했을 때 즉각적인 고장

원인의 추정과 그에 대한 대응책이 요구된다.

회전기계에서 발생하는 이상이나 고장시의 진동 신호는 여러 상태변화의 결과이며, 그 관계가 매우 다양하고 복잡하기 때문에 원인을 추정하는 데에는 매우 전문적인 지식과 오랜 경험이 필요하다. 그러므로 시스템의 고성능화, 복잡화 및 전문가의 부족으로 인하여 진단의 자동화를 도모하고자 할 때에는 진동 원인과 발생하는 현상과의 관계를 잘 표현할 수 있는 기법이 요구되며 이에 대한 연구가 여러 방면에서 이루어지고 있다.

본 연구에서는 회전기계에서 발생할 수 있는 결함을 진단하기 위해 Support Vector Machine (SVM) 알고리즘을 이용한 회전기계의 결함진단 시스템을 제안하고, 인위적인 결함을 부여할 수 있는 실험 장치로부터 취득한 진동신호를 이용하여 제안된 알고리즘의 유용성을 확인하였다.

2. Support Vector Machines (SVMs)

2.1 SVM 알고리즘^(2,3)

SVM 의 이론은 1960 년대 후반에 통계학적인 학습이론을 기초로 하여 Vapnik 에 의해 처음 제안되었으나⁽⁴⁾, 당시 컴퓨터의 계산능력 부족으로 인해 1990 년 중반에서야 다양한 분야에 적용되기 시작하였으며, 또한 현재 알고리즘의 성능향상을 위하여 많은 연구가 이루어 지고 있다. 기본적인 SVM 은 이진분류문제에 널리 이용되며, hyperplane 을 중심으로 한쪽은 positive 클래스, 다른 한쪽은 negative 클래스로 나눈다. (x_i, y_i) ($i = 1, \dots, M$)를 훈

† 책임저자, 정회원, 부경대학교 기계공학부

E-mail : bsyang@pknu.ac.kr

Tel : (051) 620-1604, Fax : (051) 620-1405

* 부경대학교 대학원

런 데이터 집합 S , 데이터 $x_i \in R^N$ 가 클래스 $y_i \in \{-1, 1\}$ 에 속한다고 하면, 두 데이터 집합의 경계가 되는 hyperplane 은 식 (1)과 같이 정의된다.

$$\mathbf{w} \cdot \mathbf{x} + b = 0, \mathbf{w} \in R^N, b \in R \quad (1)$$

여기서, \mathbf{w} 는 두 데이터 집합의 경계가 되는 가중치 벡터, \mathbf{x} 는 N 차원의 입력 벡터, b 는 한계값 (threshold value)이다.

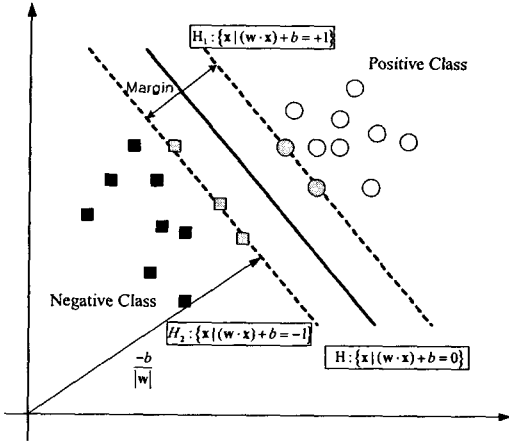


Fig. 1 Binary classification of dataset by SVM

Fig. 1 은 다른 두 데이터 집합이 경계가 되는 hyperplane H 에 의해 분리되어짐을 보이고 있다. SVM 은 경계와 각 클래스에 속한 가장 인접한 데이터와의 거리(margin)가 최대가 되게 한다. 그림에서 경계와 가장 인접한 회색의 원과 사각형 데이터 점을 support vector (SV) 라 하며, 이 SV 에는 두 클래스를 구분할 수 있는 모든 정보들이 포함되어 있기 때문에 이를 제외한 나머지 데이터는 필요 없게 되어 폐기된다. 따라서 데이터의 overfitting 과 테스트 시간이 상대적으로 모든 데이터를 사용하는 다른 알고리즘에 비하여 빠르게 된다. 새로운 데이터에 대한 결정함수는 식 (2)와 같다.

$$f(\mathbf{x}) = \text{sign}(\mathbf{w} \cdot \mathbf{x} + b) \quad (2)$$

두 hyperplane H_1 과 H_2 사이의 거리를 구하면 식 (3)과 같으며, 이 값이 최대가 되어야 한다.

$$\text{margin} = \frac{2}{\|\mathbf{w}\|} \quad (3)$$

식 (2)와 (3)을 정리하면 식 (4)와 같은 최적화문제에 귀착된다.

$$\begin{aligned} \text{Minimize } & \tau(\mathbf{w}) = 1/2 \|\mathbf{w}\|^2 \\ \text{subject to } & y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1, i = 1, 2, \dots, N \end{aligned} \quad (4)$$

식 (4)를 풀기 위하여 Lagrange multipliers $\alpha_i \geq 0$ ($i = 1, 2, \dots, N$) 의 개념을 도입하면, Lagrangian 은 식 (5)와 같이 된다.

$$L(\mathbf{w}, b, \boldsymbol{\alpha}) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^N \alpha_i y_i (\mathbf{w} \cdot \mathbf{x}_i - b) + \sum_{i=1}^N \alpha_i \quad (5)$$

여기서, $\partial L / \partial \mathbf{w} = 0, \partial L / \partial b = 0$ 이고, 이는 각각 다음과 같이 유도된다.

$$\mathbf{w} = \sum_{i=1}^N \alpha_i y_i \mathbf{x}_i, \quad \sum_{i=1}^N \alpha_i y_i = 0 \quad (6)$$

식 (6)을 식 (5)에 대입하여 정리하면,

$$\begin{aligned} \text{Maximize } & \\ L_D(\boldsymbol{\alpha}) = & \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j \end{aligned}$$

Subject to

$$\alpha_i \geq 0, i = 1, 2, \dots, N,$$

$$\sum_{i=1}^N \alpha_i y_i = 0 \quad (7)$$

식 (7)로부터 α_i 를 구할 수 있으며, 이 α_i 를 식 (6)에 대입하여 \mathbf{w} 를 구할 수 있다. 식 (6)을 사용하여 결정함수인 식 (2)를 재정의하면, 식 (8)과 같이 된다.

$$f(\mathbf{x}) = \text{sign} \left(\sum_{i=1}^N \alpha_i y_i (\mathbf{x}_i \cdot \mathbf{x}) + b \right) \quad (8)$$

따라서, 이제 한계값 b 를 구하게 되면, 새로운 입력 데이터 \mathbf{x} 를 분류할 수 있다. 여기서 b 는 SVM 을 위한 학습 알고리즘인 Sequence Minimize Optimization (SMO)으로부터 구해진다.

지금까지는 입력 공간이 선형적으로 분리가 가능한 경우만을 고려하였다. 그러나 실제로 많은 경우, 선형적으로 두 클래스를 구별할 수 없다. 하지만 N 차원 입력공간의 데이터를 고차원의 특징 공간(Q 차원)으로 변환함으로써 선형적으로 구별할 수 있으며(Fig. 2), 이를 위해 kernel 함수가 이용된다.

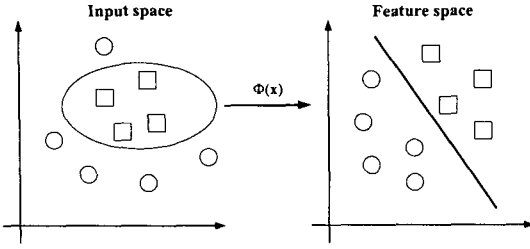


Fig. 2 Transformation to linear feature space from nonlinear input space

변환함수는 식 (9)와 같이 정의된다.

$$\mathbf{s} = \Phi(\mathbf{x}) \quad (9)$$

여기서, $\mathbf{x} \in R^N$, $\mathbf{s} \in R^Q$ 이다. 식 (9)를 결정함수 식 (8)에 대입하면 다음과 같다.

$$f(\mathbf{x}) = \text{sign} \left(\sum_{i=1}^N \alpha_i y_i (\Phi(\mathbf{x}) \cdot \Phi(\mathbf{x}_i)) + b \right) \quad (10)$$

결정함수를 kernel 함수 $K(\mathbf{x}, \mathbf{y}) = \Phi(\mathbf{x}) \cdot \Phi(\mathbf{y})$ 로 대체하면 식 (11)과 같이 된다.

$$f(\mathbf{x}) = \text{sign} \left(\sum_{i=1}^N \alpha_i y_i K(\mathbf{x}, \mathbf{x}_i) + b \right) \quad (11)$$

따라서, 식 (11)을 통하여 SVM 이 비선형적인 입력공간을 선형적인 특징공간으로 변환하여 분류할 수 있다. 일반적으로 사용되는 kernel 함수로는 RBF kernel, Polynomial kernel, Perceptron, Linear 등이 있다.

2.2 SMO 알고리즘^(5,6)

SOM 알고리즘은 SVM 을 학습시키기 위해 기존의 분해법과 비슷한 방법으로 가능한 가장 작은 크기의 Quadratic Programming (QP) 문제로 나누어서 전체 QP 최적화 문제를 풀게 된다. SMO 는 QP 문제를 해석적인 방법으로 풀기 때문에 내부 루프로서 시간 소비적인 수치적 QP 최적화를 행하지 않는다. 또한 필요한 메모리의 크기도 학습 데이터의 크기에 따라 선형적으로 달라진다. 따라서 매우 방대한 양의 학습 데이터도 다룰 수 있고, 또한 학습에 소요되는 시간은 SVM 의 결정함수를 평가하는데 지배되기 때문에 선형 SVM 과 학습 데이터 양이 적은 경우에 적합하다.

2.2.1 Lagrange multipliers α_i 의 최적화 과정

SMO 알고리즘은 목적함수를 최대화 하는 α_i 를 구하는데 있어 한번에 2 개의 α_i 값을 찾는다. 목적함수는 식 (12), 구속조건은 식 (13)과 같다.

$$L_D(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j \quad (12)$$

$$\sum_{i=1}^N \alpha_i y_i = 0 \quad (13)$$

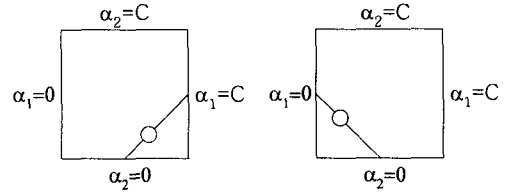
식 (13)의 조건을 이용하여 α_1, α_2 값을 제외한 나머지 α 값은 임의의 값으로 고정시킨다. 그러면 식 (13)을 다음과 같이 쓸 수 있다.

$$y_1 \alpha_1 + y_2 \alpha_2 = y_1 \alpha_1^{\text{old}} + y_2 \alpha_2^{\text{old}} = \text{const} \quad (14)$$

Fig. 3 은 식 (14)를 도식적으로 나타낸 것이다. 이는 경계구속을 나타내는 것으로 α_2 의 경계값은 다음과 같이 된다.

$$\begin{cases} L = \max(0, \alpha_2^{\text{old}} - \alpha_1^{\text{old}}) \\ H = \min(C, C + \alpha_2^{\text{old}} - \alpha_1^{\text{old}}) \end{cases} \text{ if } y_1 \neq y_2 \quad (15)$$

$$\begin{cases} L = \max(0, \alpha_1^{\text{old}} + \alpha_2^{\text{old}} - C) \\ H = \min(C, \alpha_1^{\text{old}} + \alpha_2^{\text{old}}) \end{cases} \text{ if } y_1 = y_2$$



$$y_1 \neq y_2 \Rightarrow \alpha_1 - \alpha_2 = \gamma \quad y_1 = y_2 \Rightarrow \alpha_1 + \alpha_2 = \gamma$$

Fig. 3 $y_1 \alpha_1 + y_2 \alpha_2 = y_1 \alpha_1^{\text{old}} + y_2 \alpha_2^{\text{old}} = \text{const}$.

목적함수를 풀어 쓰면

$$\begin{aligned} L_D(\alpha) = & \alpha_1 + \alpha_2 + \text{const} \\ & - \frac{1}{2} (y_1 y_1 \mathbf{x}_1^T \mathbf{x}_1 \alpha_1^2 + y_2 y_2 \mathbf{x}_2^T \mathbf{x}_2 \alpha_2^2 \\ & + 2 y_1 y_2 \mathbf{x}_1^T \mathbf{x}_2 \alpha_1 \alpha_2 + 2 \left(\sum_{i=3}^N \alpha_i y_i \mathbf{x}_i^T \right) (y_1 \mathbf{x}_1 \alpha_1 + y_2 \mathbf{x}_2 \alpha_2) \\ & + \text{const}) \end{aligned} \quad (16)$$

식 (16)에서

$$v_j = \sum_{i=3}^N \alpha_i y_i \mathbf{x}_i^T \mathbf{x}_j = u_j^{\text{old}} + b^{\text{old}} - \alpha_1^{\text{old}} y_1 x_1^T x_j - \alpha_2^{\text{old}} y_2 x_2^T x_j$$

$$\alpha_2^{\text{new,clipped}} = \begin{cases} H & \text{if } \alpha_2^{\text{new}} \geq H; \\ \alpha_2^{\text{new}} & \text{if } L < \alpha_2^{\text{new}} < H; \\ L & \text{if } \alpha_2^{\text{new}} \leq L; \end{cases} \quad (22)$$

여기서, $u_j^{\text{old}} = \mathbf{x}_j^T \mathbf{w}^{\text{old}} - b^{\text{old}}$ 이며, \mathbf{x}_j 의 결과 값이다.

$K_{11} = \mathbf{x}_1^T \mathbf{x}_1, K_{22} = \mathbf{x}_2^T \mathbf{x}_2, K_{12} = \mathbf{x}_1^T \mathbf{x}_2$ 라 두면, 식 (16)은 다음과 같이 된다.

$$L_D(\alpha) = \alpha_1 + \alpha_2 - \frac{1}{2}(K_{11}\alpha_1^2 + K_{22}\alpha_2^2 + 2sK_{12}\alpha_1\alpha_2 + 2y_1v_1\alpha_1 + 2y_2v_2\alpha_2) + \text{const}$$

$$= \frac{1}{2}(2K_{12} - K_{11} - K_{22})\alpha_2^2 + (1-s + sK_{11}\gamma - sK_{12}\gamma + y_2v_1 - y_2v_2)\alpha_2 + \text{const} \quad (17)$$

$\eta = 2K_{12} - K_{11} - K_{22}$ 라 두고, α_2 의 계수는 다음과 같다.

$$1-s + sK_{11}\gamma - sK_{12}\gamma + y_2v_1 - y_2v_2 = y_2(E_1^{\text{old}} - E_2^{\text{old}}) - \eta\alpha_2^{\text{old}} \quad (18)$$

여기서 $E_i^{\text{old}} = u_i^{\text{old}} - y_i$ 로써 예측 오차이다.

식 (18)을 식 (17)에 대입하여 정리하면

$$L_D(\alpha) = \frac{1}{2}\eta\alpha_2^2 + (y_2(E_1^{\text{old}} - E_2^{\text{old}}) - \eta\alpha_2^{\text{old}}) + \text{const} \quad (19)$$

식(19)를 α_2 에 대해 1계 및 2계 미분을 구하면

$$\frac{dL_D}{d\alpha_2} = \eta\alpha_2 + (y_2(E_1^{\text{old}} - E_2^{\text{old}}) - \eta\alpha_2^{\text{old}}) \quad (20)$$

$$\frac{d^2L_D}{d\alpha_2^2} = \eta$$

여기서 목적함수가 최대값을 가지기 위해서는 1계 미분의 값이 0이어야 한다. 따라서,

$$\alpha_2^{\text{new}} = \alpha_2^{\text{old}} + \frac{y_2(E_1^{\text{old}} - E_2^{\text{old}})}{\eta} \quad (21)$$

이 된다. 식 (15)의 경계값을 고려하면

그리고,

$$\alpha_1^{\text{new}} = \alpha_1^{\text{old}} + s(\alpha_2^{\text{old}} - \alpha_2^{\text{new}}) \quad (23)$$

2.2.2 한계값 b 의 계산

예측오차 $E(\mathbf{x}, y)$ 와 예측오차의 변화량 $\Delta E(\mathbf{x}, y)$ 는 다음과 같이 정의된다.

$$E(\mathbf{x}, y) = \sum_{i=1}^N \alpha_i y_i \mathbf{x}_i^T \mathbf{x} - b - y$$

$$\Delta E(\mathbf{x}, y) = \Delta \alpha_1 y_1 \mathbf{x}_1^T \mathbf{x} + \Delta \alpha_2 y_2 \mathbf{x}_2^T \mathbf{x} - \Delta b$$

한계값 b 는 $E(\mathbf{x}, y)^{\text{new}} = 0$ 일 때이므로, 식(23)에 의해

$$E(\mathbf{x}, y)^{\text{new}} = E(\mathbf{x}, y)^{\text{old}} + \Delta E(\mathbf{x}, y)$$

따라서,

$$\Delta b = E(\mathbf{x}, y)^{\text{old}} + \Delta \alpha_1 y_1 \mathbf{x}_1^T \mathbf{x} + \Delta \alpha_2 y_2 \mathbf{x}_2^T \mathbf{x}$$

가 된다.

3. Multi-class SVM의 전략

기본적으로 SVM은 이진 분류를 수행하며, 여러 클래스를 가지는 실제 응용분야에의 적용에는 한계가 있다. 그래서 이를 극복하기 위한 방안으로 이진 분류의 조합에 대한 전략들이 제시되었다. 대표적인 전략으로는 one-against-all⁽⁷⁾, one-against-one^(8,9), directed acyclic graph⁽¹⁰⁾ 및 binary tree⁽¹¹⁾ 등과 같은 것들이 있다. 여기에서는 one-against-all, one-against-one 기법을 적용하였으며 두 방법에 대해 분류성능을 비교, 검토하였다.

3.1 One against all 기법⁽⁷⁾

이 방법은 초기 multi-class 분류화를 수행하기 위해 사용된 모델이다. k 개의 상태를 분류하기 위해 $k-1$ 개의 SVM이 필요하며, 학습데이터는 $k-1$ 개 모두 동일하나 상태의 소속을 나타내는 label은 수행되는 SVM에 따라 달라진다. 가령, i 번째 SVM은 소속 i 와 그 나머지를 구별하는 일을 수

행하며 소속 i 에 해당되는 학습데이터는 +1 값을 가지며 나머지 즉, 다른 소속에 속하는 학습데이터는 -1 값을 가진다. 그래서 구성된 학습데이터를 이용하여 학습하고 미지의 새로운 데이터에 대하여 순차적으로 SVM 의 수행 후 그 결과값이 +1 값을 도출한 SVM 에서 그 상태를 판별하게 된다.

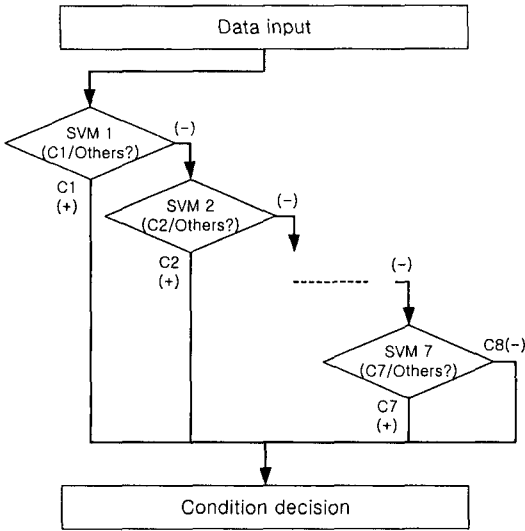


Fig. 4 One against all approach

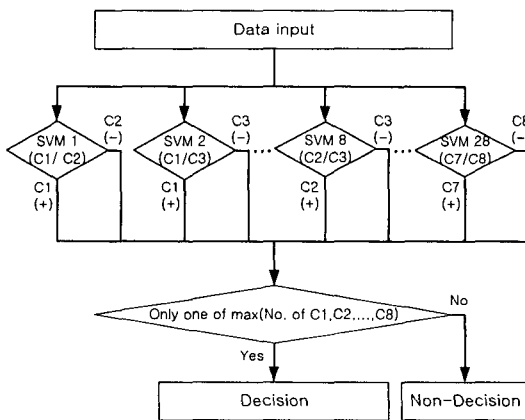


Fig. 5 One against one approach

3.2 One against one 기법^(8,9)

이 방법은 one-against-all 방법과는 달리 $k(k-1)/2$ 개의 SVM 으로 구성되며 각각의 학습데이터는 두 개의 소속을 나타내는 데이터로만 구성된다. 즉, 첫번째 SVM 은 class 1 과 class 2 만으로 구성된 학습데이터를 가지며 미지에 데이터에 대해서도 class 1 과 class 2 에 대해서만 분류하게 된다. 그래서 다수의 $k(k-1)/2$ 개)SVM 이 필요하다. 미지의 테

이터에 대한 판별은 모든 SVM 을 수행한 후 가장 많은 투표 값을 가진 class 로 그 소속을 판별하게 된다.

4. 결함 진단 시스템의 구성

4.1 실험 장치 및 데이터 취득

아래의 실험장치는 인위적으로 시스템에 결함을 인가할 수 있는 장치로서 3 상 유도 전동기로 구동되며 커플링, 베어링, 디스크, 축으로 구성되어 있다. 결함은 커플링 부분의 편심(parallel) 및 편각(angular) 정렬불량, 베어링 결함(내륜, 외륜, 볼, 및 복합 결함), 디스크의 불평형, 정상상태의 총 8 가지의 상태에 대하여 각각 연속적으로 20 개씩 총 160 개의 시간신호(time signal)를 취득하였다.

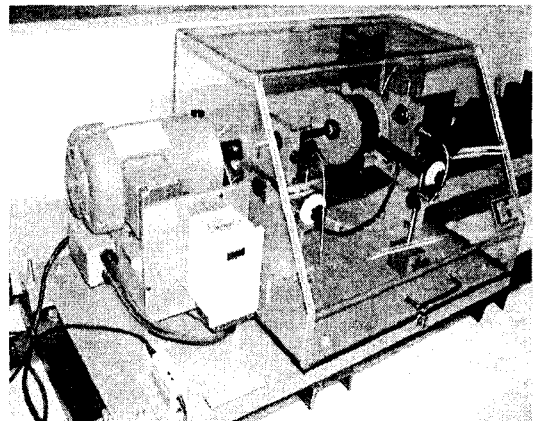


Fig. 6 Machinery fault simulator

Table 1 Description of each fault condition

Fault type		Description
Normal		No fault
Bearing	Outer race	Spalling
	Inner race	Spalling
	Ball	Spalling
	Complex	Outer, Inner, Ball
Misalignment	Angular	0.7 °
	Parallel	2mm
Unbalance		578 g·mm

4.2 특징 추출

센서로부터 취득된 시계열 신호는 시간상 연속적으로 취득된 것이 아니라 통계적인 표본 값이므로 시스템의 상태변화를 정확하게 표현하기 어렵다. 따라서 신호에 대하여 통계적인 처리기법을 이용하여 처리 분석하고 일정한 통계적 규칙을 갖

고 있는 통계량으로 시계열 신호의 특징을 표현하여야 한다.⁽¹²⁾ 시계열 신호의 특징을 표현할 수 있는 파라미터는 여러 가지가 있으나, 본 연구에서는 원 신호(raw signal)에 대한 평균(mean), 실효치(RMS), 형상계수(shape factor), 왜도(skewness), 첨도(kurtosis), crest factor, entropy estimation 및 entropy error 와 같이 8 가지의 통계량을 추출하였다. 또한 시간신호의 Daubechies-10 (db-10) wavelet 변환을 level 3 까지 수행하고 각 level 에 대한 위의 8 가지 통계값을 추출하여 총 32 개를 기본 파라미터로 선정하였다. 그러나 일반적으로 많은 파라미터의 사용은 SVM 의 성능을 저하시킬 수 있으며 학습 및 테스트 시간을 길게 하므로 총 32 개의 파라미터를 대상으로 Fig. 7 과 같이 유효성 평가를 실시하여 유효성 계수가 높은 5~13 개 정도의 파라미터를 SVM 의 입력 파라미터로 선택하였다.⁽¹²⁾

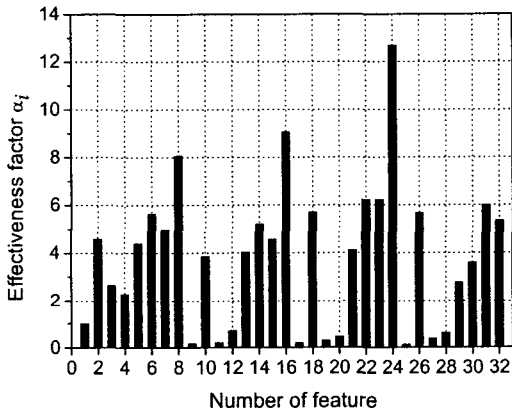


Fig. 7 Effectiveness factor of features

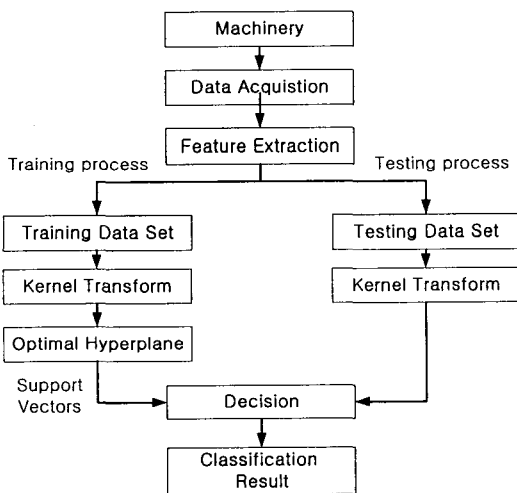


Fig. 8 Block diagram of SVM classifier system

4.3 진단시스템의 구성

그림 8 은 제안된 진단 시스템의 전체 구조를 개략적으로 나타낸다.

5. Simulation 결과

5.1 학습 및 테스트 데이터의 구성

취득한 총 160 개의 데이터 중에서 학습데이터로 각 결함에 대해 10 개씩 총 80 개를 사용하였으며 나머지 80 개를 테스트 데이터로 사용하였다.

5.2 학습 및 테스트 결과

본 연구에서 사용한 kernel 함수와 그에 대한 정의를 Table 2 에 나타내었다.

Table 3 은 유효성 계수가 높은 특징 파라미터 8 개를 선택하여 kernel 함수 및 전략에 따른 학습 및 테스트 데이터의 분류율을 나타내고 있다. 학습 데이터의 분류율은 kernel 함수와 multi-class 전략에 관계없이 거의 모두 100%의 양호한 분류율을 보이지만, 테스트 데이터의 분류율은 Gaussian RBF kernel 이 다른 kernel 비해 좀 더 나은 분류율을 보인다. 그리고 전반적으로 one against one 방법이 one against all 방법보다 나은 결과를 보이고 있다. 일반적으로 kernel 함수의 선정은 사전에 결정하는 것이 불가능하므로 적절한 함수를 선정하는 것이 양호한 분류율을 얻기 위해서는 매우 중요함을 알 수 있다. 또한 Fig. 9 는 선택한 특징 파라미터수에 따른 분류율과 학습시간을 나타내고 있다.

Table 2 Formulation for used kernel functions

Kernel	$K(x, y)$
Linear	$\mathbf{x} \cdot \mathbf{y}$
Polynomial	$(\mathbf{x} \cdot \mathbf{y} + 1)^d$
Gaussian RBF	$\exp\{-\frac{\ \mathbf{x} - \mathbf{y}\ ^2}{2\sigma^2}\}$

Table 3 Classification results due to kernel and multi-class classification strategy

Kernel	Multi-class approach	Classification rate (%)	
		Training	Testing
Linear	One vs. one	100	93.75
	One vs. all	98.125	90.00
Polynomial ($d = 1$)	One vs. one	100	93.75
	One vs. all	98.125	90.00
Polynomial ($d = 2$)	One vs. one	100	92.5
	One vs. all	100	90.00
Polynomial ($d = 3$)	One vs. one	100	93.75
	One vs. all	100	88.75
Polynomial ($d = 4$)	One vs. one	100	93.75
	One vs. all	100	91.25
RBF ($\sigma = 0.168$)	One vs. one	100	98.75
	One vs. all	100	92.50

참고문헌

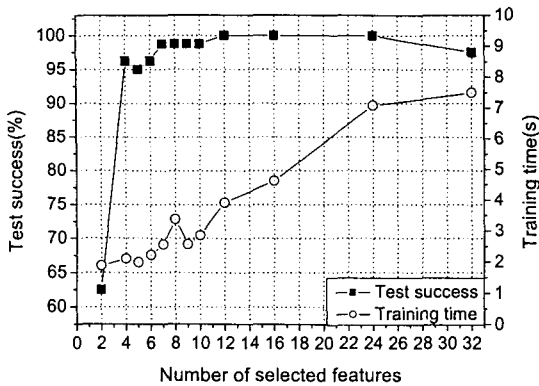


Fig. 9 Performance of SVMs for different number of selected features

6. 결론

본 연구에서는 이진 분류에 탁월한 분류성능을 가진 SVM 알고리즘을 조합한 Multi-class SVM 을 이용한 회전기계의 결함진단 시스템을 제안하였다. 진동 가속도신호의 정상 특징을 통계적 모멘트를 이용하여 특징을 추출하였으며, 이들 특징 중에서 유효성계수가 큰 특징을 선정하고 이를 입력데이터로 하여 분류기를 훈련(학습)하여 회전기계의 다양한 결함 상태를 분류하였다. 제안한 회전기계의 결함진단시스템을 적용한 결과 one against one 방법과 Gaussian RBF kernel 을 사용하였을 때 거의 100%에 가까운 양호한 결과를 얻었다. 따라서 제안된 시스템은 회전기계의 자동 상태감시 및 진단 시스템과 실시간 실행을 가능하도록 함으로서 심각한 고장에 이르기 전에 문제를 사전에 탐지하고 즉시 보수작업의 스케줄을 설정하는데 크게 기여할 수 있을 것으로 기대된다.

- (1) Williams, Davies and Drakes, "Condition-based Maintenance and Machine Diagnostics," Chapman & Hall, 1994.
- (2) M.A. Hearst, B. Scholkopf, S. Dumais, E. Osuna, J. Platt, 1998, "Trends and Controversies-Support Vector Machines," IEEE Intelligent System, Vol. 13, No. 4, pp. 18-28
- (3) N. Cristianini, J.S. Taylor, 2000, "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods," Cambridge University Press.
- (4) V.N. Vapnik, 1982, Estimation of Dependences Based on Empirical Data, Springer-Verlag.
- (5) J. Platt, 1999, "Fast Training of Support Vector Machines using Sequential Minimal Optimization," In B. Scholkopf, C.J.C. Burges, A.J. Smola, "Advances in Kernel Methods-Support Vector Learning," MIT Press, pp. 336-342
- (6) S.S. Keerthi, S.K. Shevade, 2002, "SMO Algorithm for Least Squares SVM Formulations," Control Division Technical Report CD-02-8.
- (7) C.W. Hsu, C.J. Lin, A comparison of methods for multiclass support vector machines, IEEE Transaction on Neural Networks 13 (2) 415-425, 2002.
- (8) S. Knerr, L. Personnaz, G. Dreyfus, Single-layer learning revisited: A stepwise procedure for building and training a neural network, in Neurocomputing: Algorithms, Achitutures and Applications, J. Fogelman, Ed. Springer-Verlag, New York, 1990.
- (9) U. Kreßel, Pairwise classification and support vector machines, in Advances in Kernel Methods- Support Vector Learning, B. Scholkopf, C.J.C. Burges, A.J. Smola, Eds. MIT Press, Cambridge, 255-268, 1999.
- (10) J.C. Platt, N. Cristianini, J. Shawe-Taylor, Large margin DAG's for multiclass classification, Advances in Neural Information Processing Systems 12 (2000) 547-553.
- (11) F. Schwenker, Hierarchical support vector machines for multi-class pattern recognition, Proceeding of 4th International Conference on Knowledge-Based Intelligent Engineering Systems & Allied Technologies, pp. 561-565, 2000.
- (12) B.S. Yang, T. Han, J.L. An, 2003 "ART-Kohonen neural network for fault diagnosis of rotating machinery," Mechanical Systems and Signal Processing (in press).