공 학 석 사 학 위 논 문

Support Vector Machine을 이용한 냉간 압연기의 판 터짐 결함 검지

2010년 2월

부경대학교 대학원

기계설계공학과

양 승 욱

공 학 석 사 학 위 논 문

Support Vector Machine을 이용한 냉간 압연기의 판 터짐 결함 검지

지도교수 김 선 진

이 논문을 공학석사 학위논문으로 제출함

2010년 2월

부경대학교 대학원

기계설계공학과

양 승 욱

양승욱의 공학석사 학위논문을 인준함

2009 년 12 월 18 일

주 심 공학박사 김 병 탁 (인)

위 원 공학박사 김 선 진 (인)

위 원 공학박사 양 보 석 (인)

목 차

Strip Rupture Fault Detection of Cold Rolling Mill Using Support Vector Machine

Seung-Wook Yang

Department of Mechanical Design Engineering, Graduate School,
Pukyong National University

ABSTRACT

There are several types of fault occurred in cold rolling mill. In case of common faults such as strip breakage and chattering have been studied by some researchers. Furthermore, the study of strip rupture has been developing due to difficulties of signal analysis. The main reason is that fault usually happens in transient section.

In this paper, we propose a fault detection system to detect the strip rupture in six-high stand cold rolling mills based on transient current signal of an electrical motor. In this work, signal smoothing technique is used to highlight feature precisely between normal and fault condition. Subtracting the smoothed signal from the original signal gives the residuals that contain the information related to the normal or faulty condition. Then this signal is segmented to get more information from original signal. Using segmented residual signal, discrete wavelet transform (DWT) is performed and acquired the signal presenting fault feature well. Also, feature extraction and classification are employed by using principle component analysis (PCA), independent component analysis (ICA), kernel principle component analysis (KPCA), kernel independent component analysis (KICA), and support vector machine (SVM). The actual data were acquired from the domestic steel company for validating the proposed method.

1. 서 론

1.1 연구 배경

최근 세계적인 금융위기에 의한 수요 급감에 따른 치열한 시장경쟁으로 철강 제조업 분야는 생산성과 품질 향상을 통한 경쟁력 확보를 위한 많은 노력을 기울이고 있으며, 이를 위해 압연기 설비 또한 고속화 및 고성능화가 이루어지고 있는 추세이다. 하지만 이에 걸맞은 높은 수준의 생산설비의 유지 보수관리가 이루어져야 하지만 아직 부족한 부분이 많다는 것이 현실이다. 이에 따라 관련된 설비관리 및 운전 기술의 고도화를 위해 많은 연구개발이 이루어지고 있다. [1]

압연기(rolling mill)는 제품을 원하는 두께로 가공하기 위해 금속의 소성(塑性)을 이용하여 고온 또는 상온의 금속재료를 회전하는 2 개의 롤(roll) 사이로 통과시켜서 여러 가지 형태로 압착하여 만들어내는 공작기계이다. 그 특성상 제품의 품질에 가장 밀접한 영향을 끼치며, 압착하는 과정에서 판 파단(strip breakage), 판 터짐(strip rupture) 및 채터링(chattering)과 같은 다양한 결함들이 발생하고, 이는 제품의 생산성과 품질을 저해하는 주요 요인이 된다 [2, 3].

판 파단이나 채터링과 같은 결함은 90 년대 초부터 이를 방지하기 위한 연구가 지속적으로 수행되어 오고 있고, 여러 논문을 통해 그 검지 방법이 알려져 있으나, 판 터짐에 대한 연구는 매우 적다. 이에는 다양한 원인이 있으나, 주요 원인으로는 압연을 위해 속도가 증가되는 구간인 과도구간에서 발생되는 특성을 지니고 있는 점과, 초기 증속 구간에서 발생되는 만큼 결함 발생 시 나타나는 신호 또한 매우 미약한 관계로 감지가 어렵다는 것을 들 수 있다. 이 결함의 발생 정도는 한 달에 20~30 회 즉, 하루에 최대 1 회 정도로, 판 파

단이 한 달에 2 ~ 3 회 정도 발생하는 것에 비교하면 그 발생빈도가 매우 높다는 것을 알 수 있다.

대상은 다르나 유도전동기 결함 분석을 위해 과도 전류 신호를 웨이블릿 변환(wavelet transform)에 적용한 방법이 여러 논문을 통해 제시된 바 있다. [4, 5]

이 연구에서는 판 터짐 결함을 효과적으로 검지하기 위해 다양한 분석 기법을 이용하였고, 정상과 판 터짐 간의 구분을 위해 수행된 다양한 기법들을 비교하여, 결함을 검지하는데 있어 보다 유용한 방법을 제시하고자 한다. 여기서 사용된 데이터는 국내 제철소에서 취득된 실제 전류 신호를 이용하였다.

1.2 연구 내용

이 연구에서는 냉간 압연기에서 발생하는 결함 중 하나인 판 터짐을 검지하기 위해, 대상 설비로부터 취득된 신호로부터 웨이블릿 변환(wavelet transform)을 이용하여 다양한 특징(feature)들을 계산하고 이로부터 유용한 특징추출(feature extraction)을 통해 패턴분류 알고리즘의 하나인 Support Vector Machine (SVM)으로 결함을 분류하였다.

우선 이에 대한 배경지식으로 제 2장에서는 냉간 압연기의 개요와 이 설비에서 발생되는 결함에 대해 설명하였다. 냉간 압연기는 일반 강이나 스테인리스 강 등과 같은 제품을 재결정 온도보다 낮은 조건에서 회전하는 2개의 롤 사이에 통과시켜 단계적으로 두께를 얇게 하는 것과 동시에 길이를 늘리는 작업을 하는 설비를 말한다. 이 과정에서 여러 결함들이 발생되며, 이와 관련된 내용을 제시하였다.

제 3장에서는 냉간 압연기에서 발생되는 결함 중, 초기 과도구간에서 발생되는 결함인 판 터짐을 검지하기 위해 적용된 기법들의 이론적 설명을 제시하였다. 취득된 신호로부터 전처리(preprocessing)과정을 거친 후, 비정상신호(non-

stationary signal)를 분석하는데 우수한 신호처리기법인 웨이블릿 변환이 수행되었다. 여기서 얻어진 신호를 이용하여 특징 계산 및 추출을 수행하였다. 상태감시 및 결함진단에 있어 특징 파라미터는 매우 중요한 역할을 한다. 따라서 특징 파라미터의 선정에 있어 설비의 결함에 대해 가능한 많은 정보가 포함되어 있는 파라미터를 사용하여야 하며, 이에 적절한 신호처리 기법이 적용되어야 한다. 정상과 판 터짐 간의 분류율을 확인하기 위해 이진 분류에 있어 탁월한 성능을 보이고 있는 SVM이 사용되었다.

제 4장에서는 앞에서 제시된 이론을 바탕으로 정상신호와 실제 판 터짐이 발생한 신호를 이용하여 검지 여부 및 그 성능을 조사하였다.

제 5장에서는 본 연구의 주요 결론을 요약하였다.

2. 냉간 압연기의 개요 및 판 터짐

2.1 냉간 압연기 (Cold Rolling Mill)

2.1.1 압연

압연(rolling)이란 Fig. 2.1과 같이 금속재료를 회전하는 롤 사이에 넣은 후 가압을 통해 두께나 단면적을 감소시킴과 동시에 길이를 늘리는 가공을 말한다 [6, 7]. 금속재료에 적당한 힘을 가하면 파괴되지 않고 영구변형을 일으키는 성질인 소성을 이용한 것으로, 단조 등과 함께 금속 소성가공법의 일종이다. 후판, 박판, 봉강, 형강 및 이음매 없는 강관 등 강재를 비롯하여 각종 금속이나 합금의 박, 판, 봉 및 관 등이 이 방법으로 제조된다.

Fig. 2.1 Rolling process

압연 시 소재의 온도에 따라 크게 열간 압연(hot rolling)과 냉간 압연(cold rolling)으로 나누어 진다. 철강공장에서 주괴를 반제품인 강편으로 압연하는 분괴 압연(blooming)이 열간 압연의 대표적인 예이다. 열간 압연에서는 금속이 가공에 의한 경화를 일으키지 않는 재결정 온도 이상의 온도에서 압연이 행해지기 때문에 비교적 작은 롤 압력으로 큰 변형가공이 가능하다. 분괴 압연에

서는 주괴 속에 있는 기포와 조대 결정조직을 없애 균질 조직의 강편으로 만들 수 있다. 다만 열간 압연에서는 압연 중에 고온의 소재 표면이 대기중의 산소와 화합하여 산화 막을 형성하기 때문에, 제품의 표면이 거칠고 금속 광택이 없으며 치수 정밀도도 나쁘다. 냉간 압연은 실온에서 소재를 압연하는 방식으로, 제품의 표면이 평활하고 금속 광택이 있으며 치수 정밀도도 좋다. 박판 제조용의 냉간 스트립 밀(cold strip mill)이 대표적인 예이다. 냉간 압연과 열처리를 적절히 조합함으로써 압연 제품의 결정조직을 개선시켜 뛰어난 기계적, 물리적 성질을 갖게 할 수 있다. 그러나 냉간 압연에는 소요 동력이 크고 내부 응력이 커지며, 가공경화(work hardening)에 의한 취성(brittleness)을 동반하기 쉽다.

Fig. 2.2 Examples of hot rolling mill and cold rolling mill

2.1.2 압연기의 구조

압연기는 작업롤(work roll), 보강롤(backup roll)을 포함한 롤의 총수에 따라 2단식, 3단식, 4단식, 6단식 및 다단식 등으로 분류되며[6, 7], 그 일부를 Fig. 2.3에 나타내었다. 2단식 압연기는 가장 오래된 형식의 압연기로, 스트립 밀이 나오기 전에 압연기의 출구에서 나오는 압연판을 사람이 직접 입구 쪽에 있는 사람에게 건네 주어서 압연을 반복하는 방법(pull-over 방식)이다. 현재에는 2단식 가역 압연기가 3단식 압연기와 함께 분괴 압연기로 사용되며, 열간 압연기의 조압연기(粗壓延機)로도 사용된다.

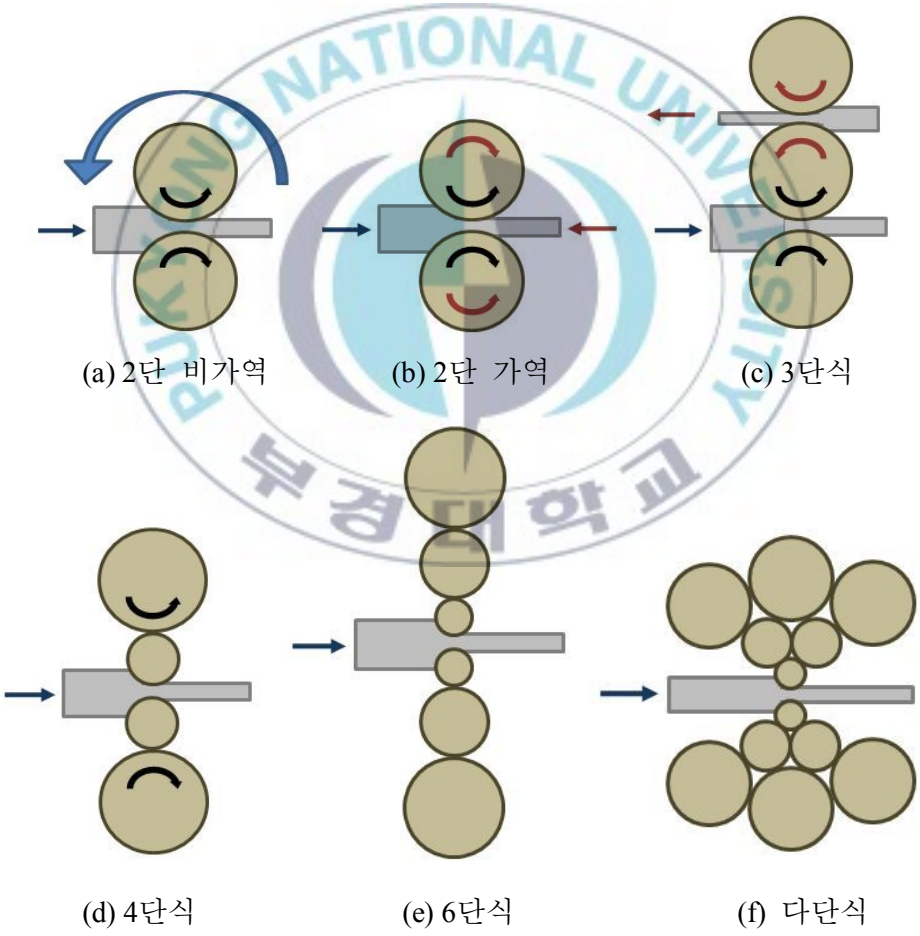


Fig. 2.3 Type of rolling mill

2단 가역 압연기는 Fig. 2.3 (a)와 비슷하나 롤을 역전시킬 수 있기 때문에 재압연 시 소재를 운반할 필요가 없다. 3단 압연기는 중간 롤이 구동되고, 상하의 롤은 마찰에 의하여 구동되는 방식으로 2단 압연기 2대의 기능을 한다. 4단 압연기는 소재와 접촉하는 롤의 지름이 작고, 외측에 보강 롤이 있다. 롤의 지름이 작으면 압연압력이 크게 되어 소성이 적어 압연이 어려운 소재의 압연에 사용된다. 주로 연강, 구리, 알루미늄 등의 광폭 판재의 압연에 적당하여 후판 압연기와 열간 및 냉간 압연기의 주요 압연기로 되어 있다. 6단식 이상의 다단 압연기는 스테인리스 강판, 규소 강판 등의 경질재료의 냉간 압연과 박에 이르기까지의 극박판의 압연용으로 고안된 것이다.

위에서 제시된 압연기 중 하나를 여러 개로 일렬로 나열한 것이 연속 압연기(tandem rolling mill)이다. 이 방식은 생산능률을 올릴 수 있으며, 압연 폭의 변화가 없는 조건이라면 소재가 적체되지 않도록 롤의 속도가 출구 측으로 갈수록 커져야 한다.

유성 압연기(planetary rolling mill)의 주요 부분의 롤 배치는 중앙의 큰 롤 주위에 20개 내외의 지름이 작은 롤이 있는데, 중앙의 지름이 큰 롤을 동력으로 구동시키면 지름이 작은 롤은 자전하면서 지름이 큰 롤 주위를 공전하여 차례로 재료에 닿아서 압연한다. 두꺼운 소재에서 1회 압연으로 1/10 내외 두께까지 압연하여 얇은 띠 강판재를 얻을 수 있는 것이 특징이다. 판재는 모두 긴 띠 강판 모양으로 압연하는 편이 경비도 절약할 수 있고, 또한 품질도 우수한 것을 얻을 수 있기 때문에 최근에는 마무리 압연을 거의 띠 강판 형식으로 하고 있다. 따라서 두꺼운 소재에서 띠 강판을 제조하는 압연기가 필요하며, 유성 압연기는 연속 압연기에 비해 설비비가 매우 적게 들기 때문에 생산량이 적은 경우의 띠 강판 압연설비로서 사용된다.

만능 압연기(universal rolling mill)는 상하 압연과 동시에 측면 압연도 할 수 있는 압연기로 레일(rail)과 같은 I 형재, H 형재, 채널(channel) 등의 압연에 사

용된다. 최근에는 대형 형강의 압연기로도 중요하게 사용되고 있다.

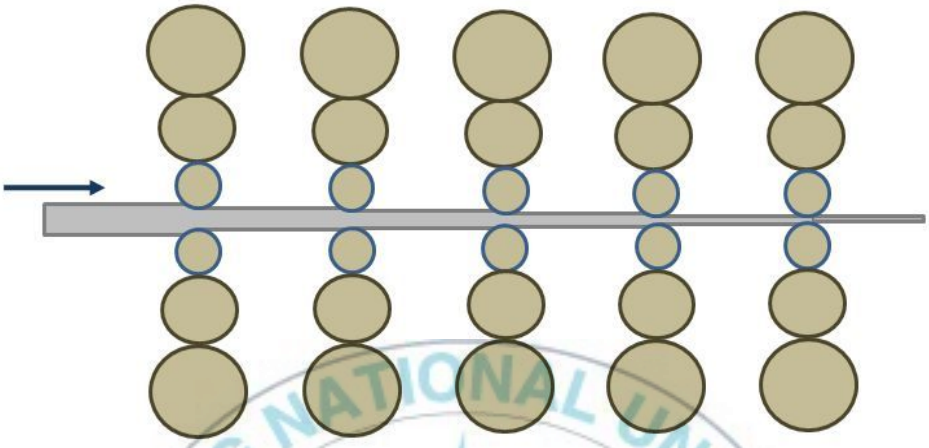


Fig. 2.4 Tandem rolling mill

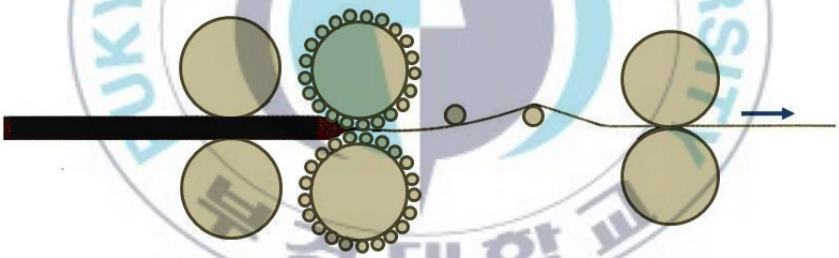


Fig. 2.5 Planetary rolling mill

2.1.3 냉연 강판의 제조공정

냉연강판의 제조공정은 Fig. 2.6과 같이 산세 → 냉간 압연 → 청정 → 소둔 → 조질 압연 → 정정 등의 복잡한 공정을 거쳐 만들어지며, 특히 공정간의 녹 발생이나 이물질 혼합에 의한 결함발생 가능성이 매우 높기 때문에 다른 공정과는 달리 세심한 주의가 필요하다[8, 9]. 따라서 공정 단축이나 연속화, 직결화 기술이 무엇보다 중요한 과제이다. 냉연 공정의 연속화 및 직결화는

일반 냉연 강판용 연속 소둔 기술(sheet-CAL)의 발전에 힘입어 일본을 중심으로 추진되었으며, 현재 산세에서 정정까지 전 공정을 통합하는 냉연공장 온라인화가 기술적으로 가능해졌다.

1) 산세(pickling)

산세 공정은 열연 코일을 스케일 브레이커 및 염산탱크에 통과시켜 최종 냉연 제품의 표면결함의 원인인 산화물 피막을 제거하는 공정이다.

2) 냉간 압연(cold rolling)

용도에 맞는 두께와 재질 확보를 위해 통상 40 ~ 90%의 압하율로 진행되며, 자동 두께 제어, 자동 형상 제어 등의 첨단 제어기기를 이용하여 균일한 두께 및 형상을 제어한다.

3) 전해 청정(electrolytic cleaning)

소둔에 앞서 냉연 코일을 알칼리 용액에 통과시켜 기계적, 화학적 반응을 통해 압연유와 오물을 제거하는 공정이다.

4) 소둔(annealing)

냉간 압연 시 경화된 강대의 재질을 연화시키기 위한 공정으로 급속 가열 및 급속 냉각을 통해 심가공용에서 고장력강까지 생산하는 생산성이 뛰어난 제조방법으로써 상자 소둔과 연속 소둔 방법을 이용한다.

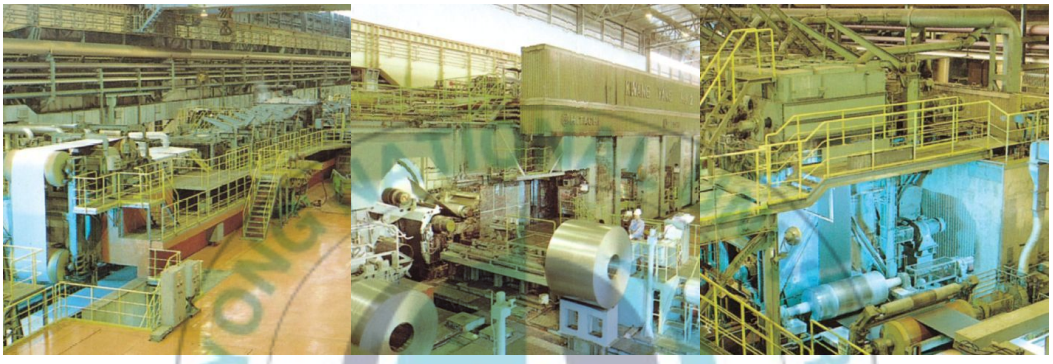
5) 조질 압연(temper rolling)

소둔을 거친 코일에 나타나기 쉬운 스트레처 스트레인(stretcher strain) 등의 결함을 판 위에 약 1% 정도의 압하를 가하여 제거하고 적당한 조도를 부여하

여 미려한 표면의 제품으로 생산한다.

6) 정정(finishing and inspection)

생산의 최종 공정으로 원하는 치수로 맞추어 제품의 결함검사 및 용도 적합 여부를 판단한다.



(a) Pickling (b) Cold rolling (c) Electrolytic cleaning

(d) Annealing (e) Temper rolling (f) Finishing and inspection

Fig. 2.6 Manufacturing process of cold rolled steel

2.2 판 터짐

제품을 압연하는 과정에서 다양한 결함이 발생할 수 있으며, 판 터짐, 판 파단 및 채터링 등이 대표적인 결함이다. 판 터짐과 판 파단은 각 스탠드(stand)의 롤과 롤 사이에서 각기 다른 속도 차로 인해 발생되는 결함이며, 채터링은 롤의 운동 형상에서의 문제로 인해 발생된다.

판 터짐과 판 파단의 차이는 발생되는 시점이 과도구간이냐 또는 정상구간이냐에 따라 구분된다. 판 터짐은 초기 과도구간에서 발생되는 결함으로 낮은 진동과 함께 제품이 끊어지는 현상으로 발생빈도가 매우 높다. 반면에 판 파단은 롤의 구동속도가 정상 궤도에 오른 정상구간에서 발생되는 만큼 큰 진동이 유발되고 제품이 찢어짐과 동시에 양쪽으로 감기는 현상으로 나타나며 발생빈도는 낮은 편이다.

2.2.1 판 터짐의 정의

판 터짐(strip rupture)이란 Fig. 2.7과 같이 코일을 압연하기 위해 초기에 롤을 구동하여 정상 압연 속도에 도달하기 위해 속도가 증가되는 과도구간에서 코일이 끊어지는 현상으로 대게 코일과 코일을 잇는 용접부분에서 발생된다.

발생 원인은 보통 제품이 각 스탠드를 통과할 때마다 두께가 얇아지게 되는데, 제품의 두께는 압하량만큼 그 길이가 늘어나게 되고, 길이에 비례하여 각 스탠드의 롤은 더 빠르게 구동되어야 한다. 그러나 그 속도가 충족되지 못하였을 때 판 터짐이 발생하게 되는 것이다. 이 결함은 판 파단과는 달리 조치를 취하는데 10 ~ 15분 정도로 그리 많은 시간이 소요되지 않고 큰 진동을 유발하지는 않으나, 한 달에 20 ~ 30회 정도 발생하는 결함으로 상대적으로 다른 종류의 결함에 비해 발생 빈도가 높다. Table 2.1에서는 압연공정에서 발생하는 주요 결함 중에서 한달 동안 발생된 판 파단과 판 터짐에 대한 발생 빈도를

나타내었다. 여기서 소재결함 판 파단은 정상구간에서 균열(crack)로 인한 장력으로 인해 발생하는 결함으로 진동이 거의 유발되지 않는다는 점에서 압연 이상으로 인한 판 파단과는 특성이 다르다.

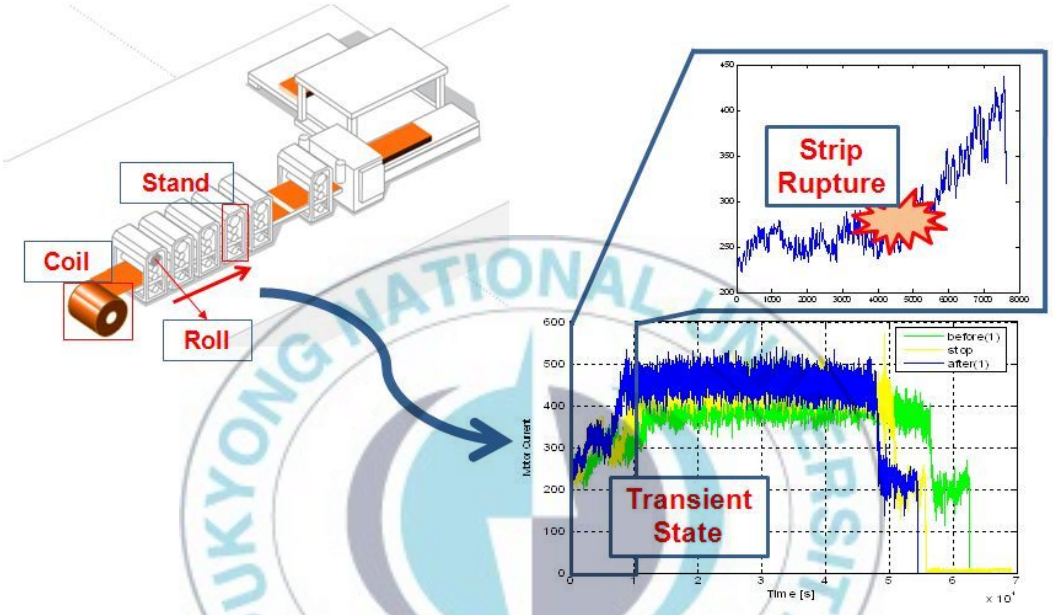


Fig. 2.7 Occurrence section of strip rupture

Table 2.1 Occurrence frequency of main faults

결함의 종류	발생 빈도(회/월)
판 터짐	24
판 파단: 소재 결함	3
판 파단: 압연 이상	3

2.2.2 데이터의 이력 정보

Table 2.2는 6일 동안 발생된 생산 장애를 나타낸다. 휴지시작 일시는 결함 조치를 취하기 위해 압연기를 휴지 상태로 한 시간을 의미하고, 휴지종료 일

시는 조치가 끝난 시간을 나타낸다. 여기서 판 터짐과 판 파단이 결함요소이며, 판 터짐이 판 파단보다 상대적으로 많이 발생함을 확인할 수 있다.

이 연구에서는 휴지시작 및 종료시간을 참고하여 결함이 발생하기 전과 후의 데이터를 정상데이터로, 결함이 발생한 데이터를 결함데이터로 구분하여, 이 두 상태가 어떠한 특징을 가지는지를 비교하였다. 그러나 아쉽게도 결함데이터 중 일부는 과도구간에 막 진입하였을 때 결함이 발생하여 압연기의 휴지로 인해 신호가 거의 취득되지 않았거나, 하나의 데이터에 하나의 압연공정 신호만 있어야 하는데, 그게 아닌 2개의 압연공정 신호가 취득되어 과도구간의 선정에서의 어려움으로 인해 사용될 수 없었다. 이는 현장 직원들의 빠른 시간 내로 결함 조치 후 제품 생산을 하고자 하는 인식이 강하다는 것과 데이터 관리에 대한 문제로 인한 것으로 판단된다. 그리하여 분석 가능한 12개의 결함 데이터와 동일한 수의 정상 데이터를 선정하였고 이를 분석에 이용하였다.

Table 2.2 Present status table about production blocking

No.	휴지시작	휴지종료	정지시간[h]	장애 명
1	08/12/12 05:35	08/12/12 05:42	0.12	부정기 교체
2	08/12/12 09:25	08/12/12 09:33	0.13	부정기 교체
3	08/12/12 15:45	08/12/12 15:51	0.1	판터짐
4	08/12/12 16:17	08/12/12 16:24	0.12	부정기 교체
5	08/12/13 01:54	08/12/13 02:01	0.12	W/R 2STD 교체
6	08/12/13 02:41	08/12/13 02:47	0.1	판터짐
7	08/12/13 03:12	08/12/13 03:28	0.27	판터짐
8	08/12/13 10:56	08/12/13 11:09	0.22	Welding Point
9	08/12/13 11:38	08/12/13 12:01	0.38	Welding Point
10	08/12/13 13:20	08/12/13 14:03	0.72	W/R 2STD 교체
11	08/12/14 00:56	08/12/14 01:03	0.12	부정기 교체
12	08/12/14 05:53	08/12/14 06:22	0.48	판터짐
13	08/12/14 10:28	08/12/14 10:43	0.25	부정기 교체
14	08/12/14 14:48	08/12/14 15:00	0.2	부정기 교체
15	08/12/14 21:02	08/12/14 21:10	0.13	부정기 교체
16	08/12/15 01:19	08/12/15 01:25	0.1	W/R 2STD 교체
17	08/12/15 07:29	08/12/15 07:38	0.15	판터짐
18	08/12/15 19:24	08/12/15 19:40	0.27	판터짐
19	08/12/15 20:04	08/12/15 20:17	0.22	Welding Point
20	08/12/15 23:10	08/12/15 23:16	0.1	부정기 교체
21	08/12/17 02:51	08/12/17 03:01	0.17	W/R 2STD 교체
22	08/12/17 03:43	08/12/17 03:52	0.15	판터짐
23	08/12/17 10:23	08/12/17 10:34	0.18	부정기 교체
24	08/12/17 11:21	08/12/17 11:30	0.15	부정기 교체
25	08/12/17 12:58	08/12/17 13:03	0.08	부정기 교체
26	08/12/17 13:20	08/12/17 13:30	0.17	부정기 교체
27	08/12/17 15:56	08/12/17 16:09	0.22	부정기 교체
28	08/12/18 04:54	08/12/18 07:49	2.92	판파단

3. 결함 검지 방법

3.1 웨이블릿 변환

전류 신호가 비정상(non-stationary) 혹은 과도(transient) 특성으로 나타날 때, 기존의 가장 널리 사용되는 푸리에 변환(Fourier transform)을 이용하는 것은 적절한 방법이 아니며 분석에 한계가 있다. 비정상신호는 STFT(short-time Fourier transform)나 웨이블릿 변환(wavelet transform)을 이용하여 분석할 수 있다[10-13]. 그러나 STFT는 길이가 짧은 창 함수(window function)를 시계열 데이터를 따라서 겹치지 않으면서 창에서 푸리에 변환을 수행하는 기법이다. 창 함수의 크기에 의해 주파수 분해능과 시간 분해능이 결정되고, 2개의 분해능 사이에는 불확정성(uncertainty principle) 관계가 존재한다. 따라서 고정된 창 함수를 사용함으로써 고 주파수 영역과 저 주파수 영역에서 동일한 해상도가 적용되는 단점이 존재하게 된다. 반면에 웨이블릿 변환은 Decomposition 과정을 통해 저주파에서는 주파수영역의 해상도를, 고주파에서는 시간영역의 해상도를 높게 하여 다양한 해상도의 해석을 수행한다. 따라서 다중 해상도 분석(multi-resolution analysis)이 가능하며 연산속도 또한 뛰어나다. 이를 Fig. 3.1에 나타내었다.

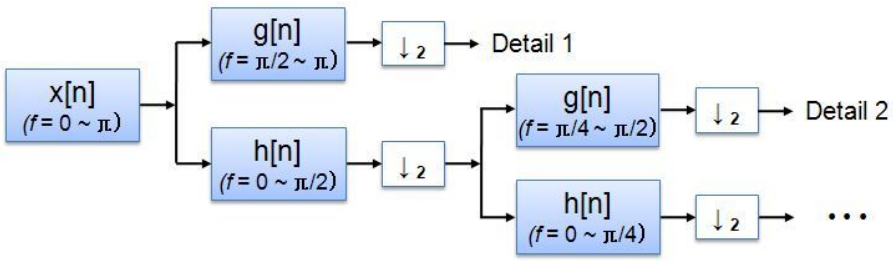


Fig. 3.1 Decomposition process

여기서, $x[n]$은 원 신호(raw signal), $g[n]$과 $h[n]$은 각각 Decomposition 수행 후의 고주파 및 저주파 대역의 신호, 그리고 f는 신호의 주파수 대역을 나타낸다.

웨이블릿 변환은 원 신호를 mother wavelet(wavelet basic function: 웨이블릿 기저함수)의 팽창(scaling)과 이동(translation)의 변화에 따라 여러 가지 신호성분으로 분해하는 것으로 시간에 따라 각각의 주파수성분의 변화추이를 알 수 있다. 이를 통해 신호의 시간 및 주파수 정보를 파악할 수 있다.

Fig. 3.2는 실제 많이 사용되고 있는 Daubechies 함수를 나타내었다.

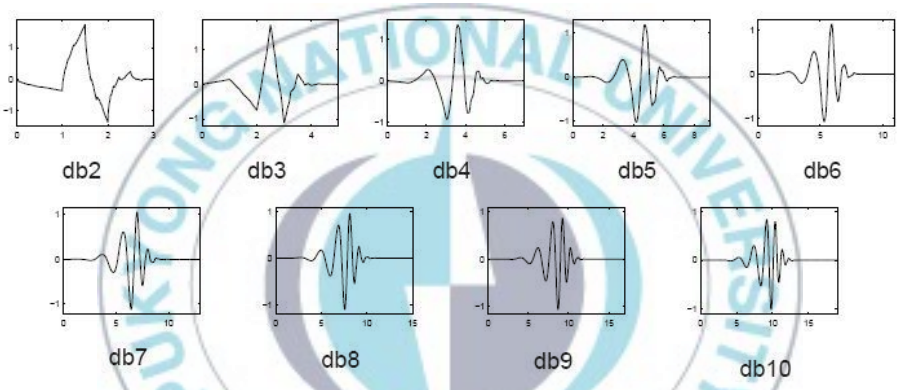


Fig. 3.2 Daubechies basis function

3.1.1 연속 웨이블릿 변환

연속 웨이블릿 변환(continuous wavelet transform, CWT)은 분석 대상신호 $f(t)$와 팽창 및 이동된 웨이블릿 기저함수 $\Psi_{a,b}$의 곱을 전 시간구간에서 적분한 것이다. 시간함수 $f(t)$의 연속 웨이블릿 변환은 식 (3.1) 및 (3.2)와 같다.

$$CWT(a,b) = \int_{-\infty}^{\infty} f(t)\Psi_{a,b}\,dt \tag{3.1}$$

$$\Psi_{a,b} = \frac{1}{\sqrt{a}}\Psi\left(\frac{t-b}{a}\right) \tag{3.2}$$

여기서 a는 팽창 파라미터, b는 이동 파라미터 그리고 Ψ는 웨이블릿 기저함수를 나타낸다. 웨이블릿 변환 결과는 특정한 이동에서의 신호 $f(t)$와 팽창 및 이동된 웨이블릿 기저함수와의 상관성을 의미한다. 웨이블릿 변환에 있어 중요한 개념인 이동은 앞섬과 지연을, 팽창은 확장과 압축을 의미한다.

3.1.2 이산 웨이블릿 변환

연속 웨이블릿 변환은 모든 팽창에서 웨이블릿 계수를 계산하므로 많은 시간이 요구되며 데이터 또한 많이 생성된다. 이러한 결점을 보완하기 위해 이산 웨이블릿 변환은 2의 누승이 되는 팽창에서 웨이블릿 변환을 수행함으로 정확성을 유지한 상태에서도 계산시간을 줄일 수 있으므로 보다 효율적인 변환이라고 할 수 있다. 이산 웨이블릿 변환(discrete wavelet transform, DWT)의 팽창함수 $\phi(t)$와 웨이블릿 기저함수 $\Psi(t)$는 식 (3.3) 및 (3.4)와 같이 표현된다.

$$\phi(t) = \sum_{k} c_k \phi(2t - k) \tag{3.3}$$

$$\Psi(t) = \sum_{k} (-1)^k c_k \phi(2t + k - N + 1) \tag{3.4}$$

여기서, N은 2의 누승인 데이터의 개수, c_k는 웨이블릿 계수이며, 다음의 조건을 만족해야 한다.

$$\sum_{k=0}^{N-1} c_k = 2, \qquad \sum_{k=0}^{N-1} c_k c_{k+2b} = 2\delta \tag{3.5}$$

여기서 δ는 Kronecker 델타를 의미한다.

3.2 특징 계산

설비에서 취득되는 진동 신호와 전동기의 전류 신호는 설비의 각 상태에 따른 일정한 특징을 지니고 있다. 그러므로 이를 분석하여 설비의 상태 파악 및 각종 결함 원인을 진단할 수 있다. 신호의 진폭 변화를 측정하여 적절한 신호처리기법을 통해 설비의 상태에 따른 특징을 계산할 수 있다. 이러한 과정을 특징 계산(feature calculation)이라 하며, 기본 원칙은 다음과 같다[10, 14].

1) 특징량을 정확하게 선택하여야 한다.
2) 동일한 특징량은 시간에 따라 변하기 않고 일정한 값이어야 한다.
3) 서로 다른 상태의 특징량 사이에는 뚜렷한 구별이 있어야 한다.
4) 특징량의 수는 각 상태를 충분히 표현할 수 있는 조건 하에서 가능한 적어야 한다.

특징 계산에 이용되는 파라미터는 다음과 같이 시간, 주파수 및 엔트로피 영역으로 구분하여 설명할 수 있다.

3.2.1 시간 영역

1) 평균(mean)

평균은 시계열 신호의 전체를 대표하는 값으로 다음 식 (3.6)과 같다.

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3.6}$$

이 값은 극단적인 관측 값의 영향에 민감하며, 다음과 같은 성질을 지닌다.

① 계열 신호의 각각의 값들의 편차의 합은 0이다.

② 평균에 대한 시계열 신호의 각각의 값들의 편차의 제곱 합은 다른 임의 값에 대한 각각의 값들의 제곱 합보다 작다.

2) 실효치(RMS)

진동의 심각한 정도를 나타내는 특성인 진동 진폭을 정량화하는 하나의 방법으로 평균치 정보나 분산(scattering) 정보가 포함되어 있고 시간에 대한 변화량을 고려하여 진동의 에너지량을 포함하므로 진동 크기의 표현에 적절하다.

$$x = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}} \tag{3.7}$$

3) 첨도(kurtosis)

파형의 4차 모멘트를 표준편차 σ^4 으로 나누어 규격화한 값으로 파형의 진폭 크기에 상관없이 파형의 형태에 의해 결정된다. 이는 시계열 신호의 확률밀도함수 분포가 갖는 첨예의 정도를 나타내는 측도로서 첨도의 적률계수의 정의는 다음과 같다.

$$\beta_2 = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i^4}{\sigma^4} \tag{3.8}$$

4) 파고율(crest factor)

진동의 피크 값과 Overall 값의 비를 나타내며, 측정값이 갑자기 커지거나 작아질 때 크게 나타난다.

$$C/F = x_p / x_s \tag{3.9}$$

5) 표준편차(standard deviation)

시계열 신호의 산포를 측정하는 척도로서 분산과 표준편차가 사용 가능하지만 표준편차가 보다 의미 있는 척도이다. 분산은 관측 값들의 제곱한 단위를 사용하지만 표준편차는 관측 값과 같은 단위를 사용한다.

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{3.10}$$

6) 왜도(skewness)

시계열 신호에 대한 관측 값들의 확률밀도함수 분포의 대칭 정도와 방향을 나타내는 값을 왜도라고 한다. 평균과 분산이 차원을 가지고 있지만 왜도는 무차원 수이다. 단봉형 분포에서 긴 왼쪽 꼬리를 가지면 음의 왜도(negative skewness)를 가지며 좌비대칭이라 하고, 오른쪽으로 긴 꼬리를 가지면 양의 왜도를 가지며 우비대칭이라 한다. 왜도 계수의 절대치가 커질수록 비대칭 정도가 커짐을 의미하고 우비대칭이면 양수, 좌비대칭이면 음수, 대칭이면 0의 값을 가진다. 왜도를 계산하는 방법은 식 (3.11)로 나타난다.

$$\beta_1 = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i^3}{\sigma^3} \tag{3.11}$$

7) 형상계수(shape factor)

진동의 Overall 값에 대한 진동 평균의 비를 나타낸다.

$$S/F = x_{rms}/\bar{x} \tag{3.12}$$

3.2.2 주파수 영역

1) 주파수 중심(frequency center)

이산 시계열 신호를 전 주파수 대역에서 보았을 때 스펙트럼 밀도의 중심을 나타낸다.

$$F/C = \frac{\sum_{i=2}^{n} \dot{x}_i x_i}{2\pi \sum_{i=1}^{n} x_i^2} \tag{3.13}$$

2) 평균제곱 주파수(mean square frequency)

$$MSF = \frac{\sum_{i=2}^{n} \dot{x}_i^2}{4\pi^2 \sum_{i=1}^{n} x_i^2} \tag{3.14}$$

3) RMS 주파수(root mean square frequency)

$$RMSF = \sqrt{MSF} \tag{3.15}$$

4) 분산 주파수(variance frequency)

분산 주파수는 전 스펙트럼에서 중심 주파수를 중심으로 다른 성분의 주파수들 사이의 분산을 나타내는 값으로, 이 값에 제곱근을 취한 값이다.

$$VF = \frac{\int_0^{\infty} (f - FC)^2 \square s(f) df}{\int_0^{\infty} s(f) df} \tag{3.16}$$

5) 제곱근 분산 주파수(root variance frequency)

$$RVF = \sqrt{\frac{\int_0^{\infty} (f - FC)^2 \cdot s(f) df}{\int_0^{\infty} s(f) df}} \tag{3.17}$$

3.2.3 엔트로피 영역

1) 엔트로피 추정(entropy estimation)

엔트로피는 불확실성의 측도로서 사용되며 식 (3.18)과 같이 정의된다.

$$H(x) = \int p(x) \cdot \ln p(x) dx \tag{3.18}$$

여기서, x는 추정오차, $p(x)$는 밀도함수를 나타낸다.

3.3 특징 추출

차원 축소(dimension reduction)는 고차원의 데이터 분석에서 중요한 전처리(preprocessing) 과정 중 하나이다. 차원 축소의 목적은 데이터에 포함되어 있는 원래의 정보를 보존하는 것과 동시에 고차원의 데이터를 낮은 차원의 공간으로 표현하는 것이다. 차원 축소를 이용하면 데이터를 간결하게 표현할 수 있으므로 시각화(visualization)나 분류(classification) 등의 분야에서 많이 사용된다 [15, 16].

패턴인식 분야의 연구에서는 데이터의 고차원 즉, 차원의 저주(curse of dimensionality)라고 불리는 어려움을 경험하고 있다. 일반적으로 차원이 확대되면 성능이 향상되나, 어느 시점을 초과하게 되면 오히려 성능이 떨어지는 현

상이 나타나게 된다. 이로 인해 데이터의 전 처리가 느려지고 많은 메모리와 시간이 요구된다. 고차원의 데이터가 가지는 다른 문제는 분류를 위한 데이터 학습에 있어 과대 적합(over-fitting)을 초래한다는 것이다. 따라서 이러한 점들이 학습 샘플의 일반화(generalization)에 나쁘게 작용된다.

따라서 특징 차원을 효율적으로 줄이기 위해 특징 추출(feature extraction)이 필요로 하게 된다. 원래의 특징 집합의 변환이나 조합에 기초하여 새로운 특징을 만드는 것을 특징 추출이라고 하며, 이를 이용하면 계산시간과 비용을 절약할 수 있다.

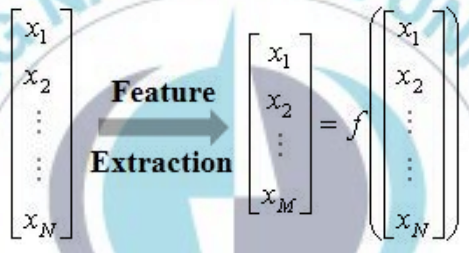


Fig. 3.3 Feature extraction

3.3.1 Principle Component Analysis (PCA)

PCA는 원래의 변수 집합을 더 적은 수를 가지는 무상관 변수(uncorrelated variable)의 집합으로 선형적으로 변환하는 통계적 기법으로, 이 변수 집합은 원래의 변수 집합이 가지고 있는 대부분의 정보를 나타내고 있다. 이것은 차원 축소를 위한 다변수의 통계적 분석의 고전적인 방법이다. 큰 무상관 변수 집합보다 작은 집합이 분석하는데 있어 적용하기가 더 쉽기 때문에, 클러스터 분석을 포함한 고차원 데이터의 시각화, 회귀, 데이터 비교 및 패턴 인식 등과 같은 데이터 비교 기술에서 광범위하게 이용된다.

주어진 중심 입력 벡터 집합 $\mathbf{x}_t$ ($(t = 1,...,l$ 및 $\sum \mathbf{x}_t = 0$)이고, 이 성분들은

m차원이며 $\mathbf{x}_t = [x_t(1), x_t(2), ..., x_t(m)]^T$, 보통 $m < l$ 이다. PCA는 식 (3.19)와 같이 벡터 $\mathbf{x}_t$ 에서 새로운 벡터인 $\mathbf{s}_t$ 로 선형 변환한다.

$$\mathbf{s}_t = \mathbf{U}^T \mathbf{x}_t \tag{3.19}$$

여기서 $\mathbf{U}$는 $m \times m$ 직교 행렬이며, $\mathbf{u}_i$ 는 공분산 행렬 $\mathbf{C}$의 고유벡터이다.

$$\mathbf{C} = \frac{1}{l} \sum_{t=1}^{l} \mathbf{x}_t \mathbf{x}_t^T \tag{3.20}$$

다시 말하면, PCA를 이용하여 고유치 문제를 해결한다.

$$\lambda_i \mathbf{u}_i = \mathbf{C} \mathbf{u}_i \qquad i = 1, ..., m \tag{3.21}$$

여기서 λ_i 는 $\mathbf{C}$의 고유치이고 $\mathbf{u}_i$ 는 이에 상응하는 고유벡터이다. $\mathbf{u}_i$ 를 기반으로 $\mathbf{s}_t$ 의 성분들은 $\mathbf{x}_t$ 의 직교 변환을 통해 계산된다.

$$\mathbf{s}_t(i) = u_i^T x_t \qquad i = 1, ..., m \tag{3.22}$$

새로운 성분을 주 성분(PC: principal component)이라 한다. 고유치를 크기 순으로 나열하여, 이들 중 가장 높은 관련된 고유벡터를 이용하여 $\mathbf{s}_t$ 의 PC의 수를 감소시킬 수 있다. 따라서 PCA는 차원을 감소시키는 특징을 가지고 있다.

3.3.2 Kernel Principle Component Analysis (KPCA)

KPCA는 비선형 접근 방법인 커널(kernel) 함수를 이용하여 일반화된 선형 PCA로 접근하는 방법이다. KPCA의 착안은 입력 벡터 $\mathbf{x}_j$ 를 고차원 특징 공간 $\mathbf{\Phi}(\mathbf{x}_j)$ 으로 사상시키고 $\mathbf{\Phi}(\mathbf{x}_j)$ 에서 PCA를 계산한다. $\mathbf{\Phi}(\mathbf{x}_j)$ 안에 선형 PCA는 $\mathbf{x}_j$ 안에 비선형 PCA와 관련이 있다. $\mathbf{x}_j$ 를 $\mathbf{\Phi}(\mathbf{x}_j)$ 에 사상(mapping)시킴으로써 차원은 학습 샘플의 수보다 높아지게 된다.

KPCA는 식 (3.23) 및 (3.24)와 같이 고유치로 풀 수 있다.

$$\lambda_j \mathbf{u}_j = \bar{\mathbf{C}}\mathbf{u}_j \qquad i = 1, ..., N \tag{3.23}$$

$$\bar{\mathbf{C}} = \frac{1}{l}\sum_{t=1}^{l} \mathbf{\Phi}(\mathbf{x}_t)\mathbf{\Phi}(\mathbf{x}_t)^T \tag{3.24}$$

여기서 $\bar{\mathbf{C}}$ 는 $\mathbf{\Phi}(\mathbf{x}_j)$ 의 상호 상관행렬이다. λ_j 는 $\bar{\mathbf{C}}$ 의 0이 아닌 고유치 중의 하나이다. $\mathbf{u}_j$ 는 고유벡터이다. 상호 상관은 비선형 공간의 입력 대신에 선형 특징 공간 F로 표현될 수 있는데, 식 (3.23)으로부터 얻어진 가장 큰 고유치 λ에 따른 고유벡터 $\mathbf{u}$는 특징공간 F 상에서 주 성분(PC)이 된다.

다양한 커널 함수를 Table 3.1에 제시하였다.

Table 3.1 Kernel function

Kernel	$K(\mathbf{x},\mathbf{x}_j)$
Linear	$\mathbf{x}^T \mathbf{x}_j$
Polynomial	$(\gamma \mathbf{x}^T \mathbf{x}_j + r)^d, \ \gamma > 0$
Gaussian RBF	$\exp(\frac{-\lVert \mathbf{x} - \mathbf{x}_j \rVert^2}{2\gamma^2})$
Sigmoid	$\tanh(\tau_0(x^T x_j) + \tau_1)$

3.3.3 Independent Component Analysis (ICA)

ICA는 임의의 다양한 신호를 통계적 판단 방법으로 상호독립적인 신호로 변환한다. 최근에 이 기술은 혼합된 소리나 진동신호로부터 독립적인 요소를 추출할 수 있는 것으로 설명되고 있다. 여기서 독립적이라는 것은 하나의 요소로부터 나타나는 정보는 다른 것으로 나타낼 수 없다는 것을 의미한다. 통계적으로 임의의 신호는 각 요소 확률의 내적(product)으로서 얻어진다. ICA 모델은 식 (3.25)와 같다.

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{3.25}$$

여기서 **A**는 혼합 행렬(mixing matrix), **S**는 독립요소 행렬(independent component matrix)이고, **x**는 측정된 데이터 행렬이다.

ICA의 문제는 **A**나 **s**의 정보 없이 측정된 데이터 **x**만으로 **A**와 **s**를 평가하는 것이다. 일반적으로 ICA는 몇 가지 방법을 통하여 독립 요소(independent component)를 구한다. ICA는 부 엔트로피(negentropy)와 최대 엔트로피(maximum entropy) 같은 이론을 통해 해를 찾아낸다. 먼저 받은 데이터를 상호 연관되지 않도록 pre-whiten 작업을 하여 $\tilde{\mathbf{x}}$를 구한다. 상호 상관행렬 $\mathbf{C} = E[\mathbf{x}\mathbf{x}^T]$의 특이 값 분해(SVD: singular value decomposition)는 식 (3.26)과 같다.

$$\mathbf{C} = \mathbf{\Psi}\sum\mathbf{\Psi}^{\mathbf{T}} \tag{3.26}$$

여기서 $\sum = diag(\sigma_1, \sigma_2, ..., \sigma_n)$는 특이 값의 대각행렬이고, $\mathbf{\Psi}$는 특이 벡터 행렬과 관련이 있다. 이렇게 해서 $\tilde{\mathbf{x}}$는 식 (3.27)과 같이 표현될 수 있다.

$$\tilde{\mathbf{x}} = \sum^{-1/2}\mathbf{\Psi}^{\mathbf{T}}\mathbf{x} = \mathbf{Q}\mathbf{A}\mathbf{s} = \mathbf{B}\mathbf{s} \tag{3.27}$$

여기서 **B**는 식 (3.28)에 의해 직교 행렬이 된다.

$$E[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T] = \mathbf{B}\,E[ss^T]\mathbf{B}^T = \mathbf{B}\mathbf{B}^T = \mathbf{I} \tag{3.28}$$

SVD 기반의 기술을 사용하는 장점은 주어진 레벨보다 적은 특이 값의 제거를 통해 노이즈를 줄일 수 있다. 다음으로는 고정 포인트 알고리즘을 사용한다. 취득한 데이터 벡터 **x**를 벡터 **y**로 변환하는 분리행렬 **W**를 정의하고 고정 포인트 알고리즘은 **y**의 첨도(kurtosis)의 절대값을 최대로 하는 분리행렬 **W**를 정의한다. 벡터 **y**는 독립요소가 요구된다. 그래서 식 (3.29)와 같이 쓸 수 있다.

$$\tilde{\mathbf{s}} = \mathbf{y} = \mathbf{W}\mathbf{x} \tag{3.29}$$

식 (3.27)로부터 우리는 $\tilde{\mathbf{s}}$를 다음과 같이 평가할 수 있다.

$$\tilde{\mathbf{s}} = \mathbf{B}^{\mathbf{T}}\tilde{\mathbf{x}} = \mathbf{B}^{\mathbf{T}}\mathbf{Q}\mathbf{x} \tag{3.30}$$

식 (3.29)와 (3.30)으로부터 **W**와 **B**의 관계는 다음과 같이 표현된다.

$$\mathbf{W} = \mathbf{B}^T\mathbf{Q} \tag{3.31}$$

B를 계산하기 위해 각 열의 b_i는 초기화되고 식 (3.32)에서 i번째 독립요소가 최대로 비가우시안(non-Gaussianty)을 가질 때까지 업데이트 된다.

$$\mathbf{s}_i = (\mathbf{b}_i)^T\tilde{\mathbf{x}} \tag{3.32}$$

3.3.4 Kernel Independent Component Analysis (KICA)

KICA는 KPCA를 사용한 중심화(centering) 및 백색화(whitening) 과정과 ICA를 이용한 반복하는 부분(iterative section)을 조합한 것이다. 우선 KPCA 방법을 이용하여 데이터를 취득한다.

$$\mathbf{S}_t^{\Phi} = \frac{1}{M}\sum_{j=1}^{M}\mathbf{\Phi}(\mathbf{x}_t)\mathbf{\Phi}(\mathbf{x}_t)^T \tag{3.33}$$

여기서 $\mathbf{S}_t^{\Phi}$는 커널 함수를 이용하여 계산된 고차원 행렬이다.

그리고 식 (3.34)와 같이 KPCA를 이용하여 **r**를 구하고, ICA를 이용하여 **r**로부터 독립 성분 **s**을 구한다.

$$\hat{\mathbf{s}} = \mathbf{W}\mathbf{x} = \mathbf{W}\mathbf{r} \tag{3.34}$$

요약하면, KICA를 사용한 비선형 특징 추출은 다음의 두 가지 단계를 수행한다.

1) Kernel PCA를 이용한 whitened 과정

2) Kernel PCA whitened 공간에서의 ICA 변환

3.4 Support Vector Machine (SVM)

SVM은 1995년에 Vapnik에 의해 제안되었고 인공 신경망이 가지는 몇 개의 단점을 극복하였으며, 명료한 이론적 근거에 기반을 두고 있다. 이는 데이터 마이닝 분야는 물론 얼굴인식과 같은 패턴인식 분야에도 널리 사용되고 있다 [10, 15, 17].

SVM은 패턴을 고차원 특징 공간으로 사상시킬 수 있다는 점과 대역적으로 최적의 식별이 가능한 특징을 가진다. 여기서 네트워크의 가중치(weight)는 선형 부등 조건을 가진 QP(quadratic programming) 문제를 해결함으로써 얻어진다. 또한 신경망을 포함하여 통계적 패턴인식 방법 등 전통적인 대부분의 패턴인식 기법들이 학습 데이터의 수행도를 최적화하기 위한 경험적인 위험 최소화(empirical risk minimization) 방법에 기초하고 있는데 반해, SVM은 고정되어 있지만 알려지지 않은 확률 분포를 갖는 데이터에 대해 잘못 분류하는 확률을 최소화하는 구조적인 위험 최소화(structural risk minimization) 방법에 기초하고 있다.

3.4.1 SVM의 기본 이론

SVM은 주로 이진 분류에 이용되며 Fig. 3.4와 같이 초평면(hyperplane)을 중심으로 한쪽은 Positive class, 다른 한쪽은 Negative class로 구분된다. 두 데이터 집합의 경계가 되는 초평면은 식 (3.35) 및 (3.36)과 같이 정의된다.

$$\mathrm{H}:(\mathbf{w}^T\mathbf{x})+b=\sum_{j=1}^{M}w_j x_j+b=0 \tag{3.35}$$

$$w\in R^N, b\in R \tag{3.36}$$

여기서 **w**는 두 데이터 집합의 경계가 되는 가중치 벡터, **x**는 N 차원의 입력벡터 그리고 b는 한계 값(threshold value)이다.

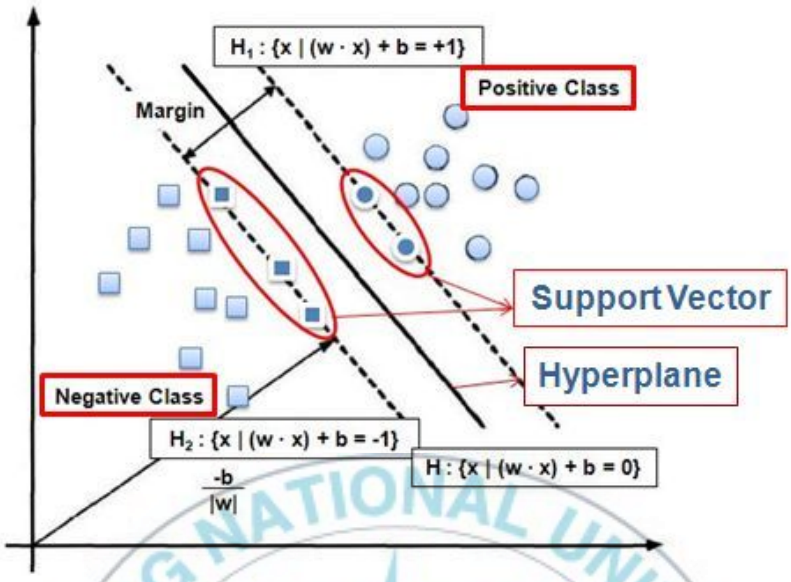


Fig. 3.4 Binary classification of dataset by SVM

Fig. 3.4는 다른 클래스의 데이터가 초평면에 의해 분류됨을 볼 수 있다. 초평면은 두 클래스의 거리(margin)가 최대가 되도록 선정되어야 하며, 이 때 최대 거리의 결정에 중요한 역할을 하는 데이터를 Support Vector(SV)라 한다. 즉, SV에는 두 클래스를 결정하는 정보를 가지고 있기 때문에, 이를 제외한 나머지 데이터들은 필요 없으므로 삭제된다. 그러므로 데이터의 과대 적합(over-fitting)과 테스트 시간이 모든 데이터를 사용하는 다른 알고리즘에 비해 빠르다. 새로운 데이터에 대한 결정함수(decision function)는 식 (3.37)과 같다.

$$f(\mathbf{x}) = sign((\mathbf{w}^T\mathbf{x}) + b) \tag{3.37}$$

두 초평면 H_1과 H_2 사이의 거리를 구하면 식 (3.38)과 같으며, 이 값이 최대가 되어야 한다.

$$\text{margin} = \frac{2}{\|\mathbf{w}\|} \tag{3.38}$$

식 (3.37)과 (3.38)을 정리하면, 식 (3.39) 및 (3.40)과 같다.

$$\textbf{minimize} \quad \frac{1}{2}\|\mathbf{w}\|^2 \tag{3.39}$$

$$\text{subject to } \mathrm{y_i}(w^T x_i + b) \geq 1, \quad y_i = \{+1 \text{ or } -1\} \tag{3.40}$$

위의 수식은 Kuhn-Tucker 조건을 가지는 문제를 등가 라그랑지안 쌍대 문제 (equivalent Lagrangian dual problem)로 변환하면 식 (3.41)과 같이 된다.

$$\textbf{minimize} \quad L(\mathbf{w}, b, \boldsymbol{\alpha}) = \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{N} \alpha_i y_i (\mathbf{w}\mathbf{x}_i + b) + \sum_{i=1}^{N} \alpha_i \tag{3.41}$$

0이 되기 위해 L의 도함수가 필요함과 동시에 $\mathbf{w}$와 b에 대해서 식 (3.41)을 최소화하는 것이 요구된다.

$$\frac{\partial L}{\partial \mathbf{w}} = 0, \quad \frac{\partial L}{\partial b} = 0 \tag{3.42}$$

이는 다음과 같이 각각 유도된다.

$$\mathbf{w} = \sum_{i=1}^{N} \alpha_i y_i \mathbf{x}_i, \quad \sum_{i=1}^{N} \alpha_i y_i = 0 \tag{3.43}$$

식 (3.43)을 식 (3.41)에 대입하여 정리하면,

$$\text{maximize } L(\alpha)=\sum_{i=1}^{N}\alpha_{\mathrm{i}}-\frac{1}{2}\sum_{i,j=0}^{N}\alpha_i\alpha_j y_i y_j \mathbf{x}_i\mathbf{x}_j \tag{3.44}$$

$$\textbf{subject to } \alpha_i \geq 0,\ i=1,2,...,N$$
$$\sum_{i=1}^{N}\alpha_{\mathrm{i}} y_i = 0 \tag{3.45}$$

위의 쌍대 최적화 문제를 풀면, 계수 α_i를 구할 수 있으며, 이 계수를 식 (3.43)에 대입하면 **w**를 구할 수 있다. 식 (3.43)을 이용하여 결정 함수인 식 (3.37)을 재정의하면 식 (3.46)과 같다.

$$f(\mathbf{x})=sign\left(\sum_{i,j=1}^{N}\alpha_i y_i(\mathbf{x}_i\mathbf{x}_j)+b\right) \tag{3.46}$$

따라서, 한계 값 *b*를 구하면 새로운 입력 데이터 **x**를 분류할 수 있게 된다. 여기서 *b*는 SVM의 학습 알고리즘인 SMO(sequence minimize optimization)에 의해 구해진다.

SVM은 커널 함수를 이용하여 비선형 분류로도 사용된다. 실제 응용문제에서는 대부분의 입력 벡터들은 입력 공간 내에 비선형으로 분포하게 된다. 이를 위해 분류되어야 할 데이터는 고차원 특징 공간에 사상되는데, 이는 선형 분류만이 가능하다. *n* 차원의 입력 벡터 **x**를 1차원 특징 공간으로의 사상을 위해 선형 벡터 함수 $\mathbf{\Phi(x)}=(\phi_1(\mathbf{x}),...,\phi_l(\mathbf{x}))$을 사용하고자 할 때, 선형 결정함수는 다음과 같이 주어진다.

$$f(\mathbf{x}) = sign\left(\sum_{i,j=1}^{N} \alpha_i y_i (\mathbf{\Phi}^T(\mathbf{x}_i)\mathbf{\Phi}(\mathbf{x}_j)) + b \right) \tag{3.47}$$

고차원 특징 공간에서 작업하는 것은 복잡한 함수의 표현이 가능하나, 이 또한 문제를 만들어 낸다. 계산을 요구하는 문제가 넓은 벡터로 인해 발생되고 고차원으로 인해 over-fitting 또한 존재하게 된다. 후자의 문제는 커널 함수를 사용하여 해결할 수 있다. 커널은 $K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{\Phi}^T(\mathbf{x}_i)\mathbf{\Phi}_j(\mathbf{x}_j))$로 대체함으로써 원래의 데이터 포인트의 특징 공간 사상의 내적(dot product)을 보상해주는 함수이다. 커널 함수를 적용할 때, 특징 공간에서의 학습은 $\mathbf{\Phi}$의 양함수 평가를 요구하지 않으며 결정함수는 다음과 같다.

$$f(\mathbf{x}) = sign\left(\sum_{i,j=1}^{N} \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}_j) + b \right) \tag{3.48}$$

Mercer의 이론을 만족하는 함수는 특징 공간에서의 내적을 계산하기 위해 커널 함수로 사용될 수 있다. SVM에 사용되는 커널 함수는 linear, polynomial, Sigmoid 및 Gaussian RBF 등과 같이 다양하게 있다. 커널은 학습 데이터를 분류하는 특징 공간을 정의하기 때문에, 적절한 커널 함수를 선택하는 것이 매우 중요하다. 다양한 커널 함수를 Table 3.2에 나타내었으며, 이 중에서 Gaussian RBF 함수가 가장 많이 사용되고 있다.

Table 3.2 Formulation of kernel function

Kernel	$K(\mathbf{x}, \mathbf{x}_j)$
Linear	$\mathbf{x}^T \cdot \mathbf{x}_j$
Polynomial	$(\gamma \mathbf{x}^T \cdot \mathbf{x}_j + r)^d, \ \gamma > 0$
Sigmoid	$\tanh(a \times x \cdot y / b + 2)$
Laplacian RBF	$\lvert \mathbf{x} - \mathbf{y} \rvert$
Gaussian RBF	$\exp(-\lVert \mathbf{x} - \mathbf{x}_j \rVert^2 / 2\gamma^2)$

3.4.2 Sequential Minimum Optimization (SMO)

SMO는 SVM을 학습시키기 위해, 기존의 분해법과 유사한 방법으로 가능한 가장 작은 크기의 Quadratic Programming (QP) 문제로 나누어서 전체 QP 최적화 문제를 풀게 된다. SMO는 QP문제를 해석적인 방법으로 풀기 때문에 내부 루프로서 시간 소비적인 수치적 QP 최적화를 수행하지 않는다. 또한 필요한 메모리의 크기도 학습 데이터(training data)의 크기에 따라 선형적으로 달라진다. 따라서 매우 방대한 양의 학습 데이터도 다룰 수 있을 뿐만 아니라 학습에 소요되는 시간은 SVM의 결정 함수를 평가하는데 지배되기 때문에 선형 SVM과 학습 데이터 양이 적은 경우에 적합하다.

1) Quadratic Programming (QP)

Vapnik는 Chunking으로 알려져 있는 SVM-QP 문제의 해결을 위한 Projected Conjugate Gradient 알고리즘을 사용한 방법이라고 하였다. Chunking 알고리즘은 0 라그랑지 승수와 일치하는 행렬의 행과 열을 제거했을 시 2차 형식(quadratic form)의 값이 같다는 사실을 이용한다. 그러므로 Chunking은 0이 아닌 라그랑지 승수의 수만큼 제곱된 학습 데이터의 수로부터 행렬의 크기를 많이 줄일 수 있다. 그러나 줄어든 행렬이 메모리와 일치하지 않으므로, Chunking은 여전

히 넓은 범위의 학습 문제는 다룰 수가 없다. Osuna 등은 SVM을 위한 QP 알고리즘의 전체 새로운 집합을 제안하는 원리를 증명하였다. 이 원리는 큰 QP 문제는 더 작은 QP의 하위 문제의 연속으로 분석된다는 것을 증명하였다.

2) Sequential Minimum Optimization (SMO)

Platt에 의해 제안된 SMO은 추가적 행렬 저장이나 수치적 QP 최적화 과정 없이도 SVM-QP 문제를 풀 수 있는 간단한 알고리즘이다. 이 방법은 수렴을 위해 Osuna의 원리를 이용하여 전체 QP 문제를 QP 하위 문제로 분해한다.

두 개의 라그랑지 승수 α_1, α_2를 구하기 위해, 첫 번째로 SMO은 이들 승수에 대한 제약조건을 계산하고 구속되는 최소값을 구한다. 이들 승수의 새로운 값들은 $0 \le \alpha_1, \alpha_2 \le C$로 정의된 (α_1, α_2) 공간 내에 한 줄로 있어야 한다.

$$\alpha_1 y_1 + \alpha_2 y_2 = \alpha_1^{\text{old}} y_1 + \alpha_2^{\text{old}} y_2 = \text{constant} \tag{3.49}$$

식 (3.49)를 Fig. 3.5로 표현할 수 있다.

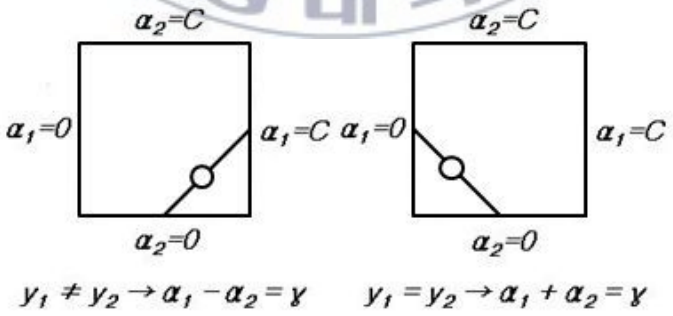


Fig. 3.5 $\alpha_1 y_1 + \alpha_2 y_2 = \alpha_1^{\text{old}} y_1 + \alpha_2^{\text{old}} y_2 = \text{constant}$

이 알고리즘은 첫째로 두 번째 라그랑지 승수 α_2^{new}를 계산하고, 이를 이용하여 α_1^{new}을 얻는다. α_2의 경계 값은 다음과 같다.

$$\begin{aligned} &\text{If } y_1 \neq y_2;\ L = \mathbf{max}(0, \alpha_2^{\text{old}} - \alpha_1^{\text{old}}) \\ &\qquad\qquad H = \mathbf{min}(C, C + \alpha_2^{\text{old}} - \alpha_1^{\text{old}}) \end{aligned} \tag{3.50}$$

$$\begin{aligned} &\text{If } y_1 \neq y_2;\ L = \mathbf{max}(0, \alpha_2^{\text{old}} + \alpha_1^{\text{old}} - C) \\ &\qquad\qquad H = \mathbf{min}(C, C + \alpha_2^{\text{old}} + \alpha_1^{\text{old}}) \end{aligned} \tag{3.51}$$

대각선을 따라 목적함수의 2차 도함수를 다음과 같이 나타낼 수 있다.

$$\eta = K(\mathbf{x}_1, \mathbf{x}_1) + K(\mathbf{x}_2, \mathbf{x}_2) - 2K(\mathbf{x}_1, \mathbf{x}_1) \tag{3.52}$$

일반적인 경우에서, 목적함수는 궁극적으로 명확해질 것이고 선형 대등 제약조건의 지시에 따라 최소값일 것이며, η는 0보다 클 것이다. 이 경우에서, SMO는 제약조건의 지시에 따라 최소값을 계산한다.

$$\alpha_2^{\text{new}} = \alpha_2^{\text{old}} + \frac{y_2(E_1^{\text{old}} - E_2^{\text{old}})}{\eta} \tag{3.53}$$

여기서 E_i는 i번째 학습 데이터의 예측 오차이다. 다음 과정에서는 구속 최소값이 선분(line segment)의 끝에서 비구속 최소값을 잘라 냄으로써 구해진다.

$$\alpha_2^{\text{new, clipped}} = \begin{cases} H & \text{if } \alpha_2^{\text{new}} \geq H; \\ \alpha_2^{\text{new}} & \text{if } L < \alpha_2^{\text{new}} < H; \\ L & \text{if } \alpha_2^{\text{new}} \leq L; \end{cases} \tag{3.54}$$

$s = y_1 y_2$ 라고 하자. α_1^{new} 의 값은 새로운 값인 α_2^{new} 으로부터 구해진다.

$$\alpha_1^{\text{new}} = \alpha_1^{\text{old}} + s(\alpha_2^{\text{old}} - \alpha_2^{\text{new}}) \tag{3.55}$$

라그랑지 승수에 관한 식 (3.44)를 푼다는 것만으로 SVM의 한계값 b를 구할 수 없다. 따라서 한계값 b는 개별적으로 계산되어야 한다. 한계값 b_1, b_2 는 새로운 α_1, α_2 가 각각의 영역에 있지 않을 때에 유효하다.

$$b_1 = E_1 + y_1(\alpha_1^{\text{new}} - \alpha_1^{\text{old}})K(\mathbf{x}_1, \mathbf{x}_1) + y_2(\alpha_2^{\text{new, clipped}} - \alpha_2^{\text{old}})K(\mathbf{x}_1, \mathbf{x}_2) + b^{\text{old}} \tag{3.56}$$

$$b_2 = E_2 + y_1(\alpha_1^{\text{new}} - \alpha_1^{\text{old}})K(\mathbf{x}_1, \mathbf{x}_2) + y_2(\alpha_2^{\text{new, clipped}} - \alpha_2^{\text{old}})K(\mathbf{x}_2, \mathbf{x}_2) + b^{\text{old}} \tag{3.57}$$

b_1와 b_2이 유효할 때, 이 값들은 동일하다. 두 개의 새로운 라그랑지 승수가 영역에 있고 L이 H와 같지 않을 때, b_1와 b_2간의 간격이 한계 값이 된다. 이것은 QP 문제의 최적 포인트를 위한 필요 충분조건이다. 이 경우에서 SMO는 b_1와 b_2의 중간에서 한계 값을 선택하게 된다.

4. 판 터짐 검지 적용 예

Fig. 4.1은 판 터짐을 검지하기 위한 과정을 도식화하였다. 작업롤(work roll)을 구동하는 유도 전동기의 전류 신호를 취득하였고, 이를 전 처리 후 웨이블릿 변환을 수행하여 각 레벨 별로 Detail 신호를 얻었다. 이를 이용하여 특징 계산 및 추출을 수행하였고 SVM으로 분류를 수행하여 정상과 결함 간의 통계적 수치를 나타내었다. 이 과정의 일환으로 정상과 결함 데이터 각 12개 중 7개가 학습 데이터로 선택되고 이를 SMO 알고리즘에 입력하여 학습시킨 후, 각 상태 간의 구분을 위해 필요로 하는 데이터는 SV로 선정되고 나머지 데이터는 필요 없으므로 폐기된다. 이는 새로운 미지의 데이터가 입력되었을 때 그 상태를 판별하는 데이터베이스(DB)가 된다. 또한 필요에 의해 추가적인 학습 데이터가 입력될 수 있고 이 데이터가 상태 구분에 있어 명확한 성질을 지니게 되면 SV로 선정되어 업데이트된다.

실제 압연기에서 취득된 데이터를 이용하여 위에서 제시된 방법으로 각 과정의 내용과 적용을 설명한다.

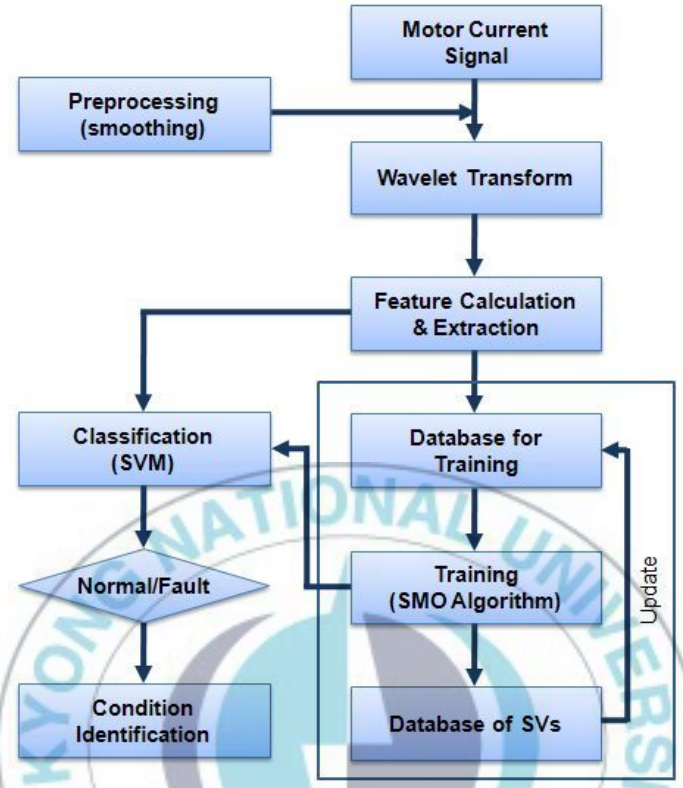


Fig. 4.1 Process for strip rupture detection

4.1 적용된 냉간 압연기

냉간 압연기는 주로 4단 냉간 압연기가 사용되나, 이 연구에서는 Fig. 4.2와 같이 5단계로 연속 압연하는 6단 냉간 압연기가 사용되었다. 연속 압연기는 한 번에 단계적으로 압연을 함으로써 생산효율이 증대되는 장점을 지니며, 6단이 4단보다 압하력이 강하여 코일의 두께를 보다 얇게 할 수 있다.

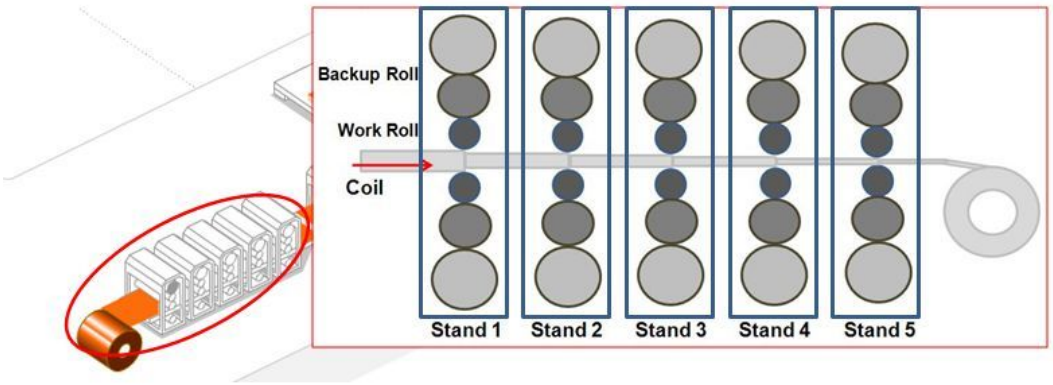


Fig. 4.2 6-high cold rolling mill

Fig. 4.2에서 보여지는 바와 같이 적용된 냉간 압연기는 5개의 스탠드로 구성되어 있으며, 각 스탠드는 2개의 보강 롤(backup roll)과 1개의 작업 롤(work roll)로 구성되어 있다. 스탠드의 수는 연속압연과 관련 있는 것으로 몇 번의 단계를 거쳐 재료를 압연할 지에 따라 결정된다. 보강롤은 작업롤과 접촉 회전하면서 압연에 필요로 하는 압하력을 작업롤에 전달해 주는 역할을 하며, 작업롤은 코일과 접촉하여 이를 압연 및 이송하는 역할을 한다.

Fig. 4.3은 냉간압연기의 측면도이다. 구동을 위해 유도전동기가 사용되고 두 개의 작업롤에 동력이나 회전 전달 및 정확한 각속도비로 전달하기 위해 기어 상자(gear box)가 요구된다. 여기서 발생된 힘을 작업롤에 효율적으로 전달하기 위해 유니버설 조인트(universal joint)가 사용된다.

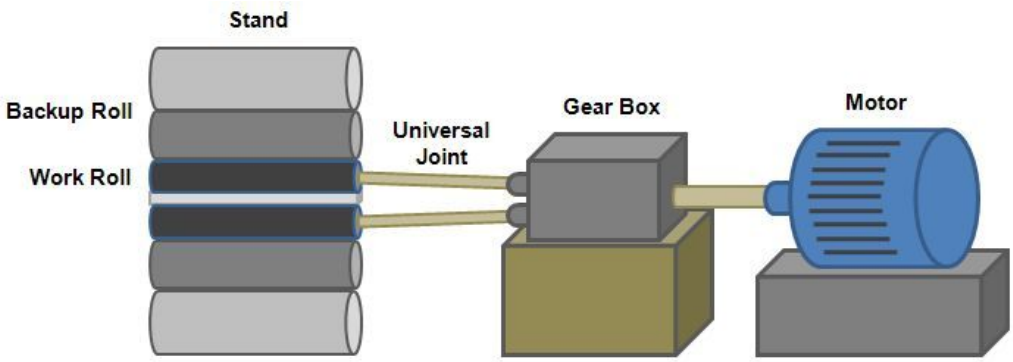
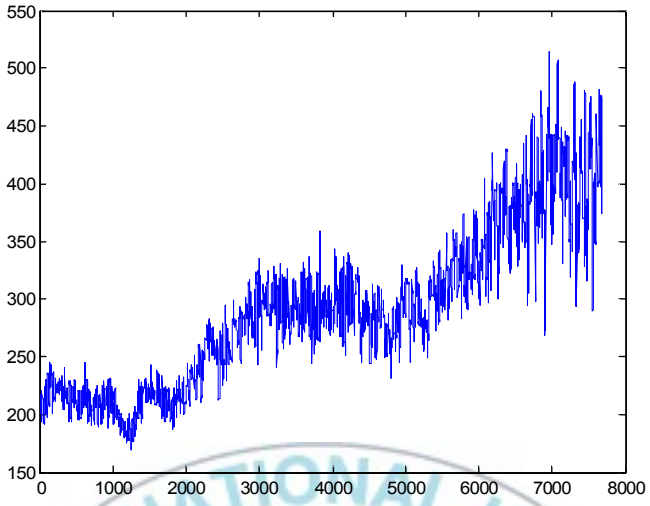


Fig. 4.3 Side view of cold rolling mill

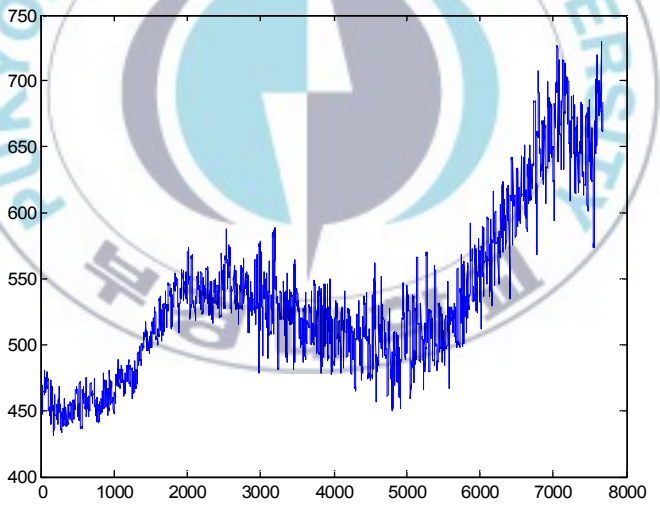
4.2 데이터 취득

Fig. 4.4와 4.5는 작업롤을 구동하는 전동기에서 취득된 정상 및 판 터짐 전류의 원 신호(raw signal)를 나타낸다. 신호는 결함이 발생되는 구간인 과도구간만을 선정하였으며, 비교를 위해 정상신호 또한 과도구간을 선정하였다.

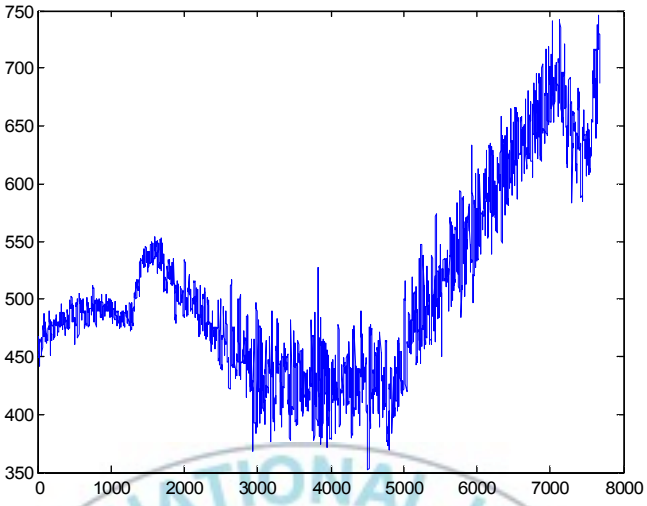
선정된 데이터의 시간 즉, 과도구간에 도달되는 시간은 제어기기나 조작자의 의지 혹은 제품의 종류에 따라 1~3분으로 매우 임의적이나, 대게 1분 정도이다. 또한 그림을 통해 후반부 스탠드로 갈수록 전류 값이 상승함을 알 수 있는데, 이는 코일이 각 스탠드를 거치며 압연될 때마다 두께가 점점 얇아지게 되고, 얇아진 만큼 길이가 늘어나게 된다. 즉, 늘어난 길이만큼 전동기 속도의 빨라짐이 요구되므로 전류 값도 그만큼 상승하게 되는 것이다. 그러나 전류 값이 선형적으로 상승하는 것이 아니라, 이 또한 위에 제시된 원인에 따라 그 특성을 달리한다.



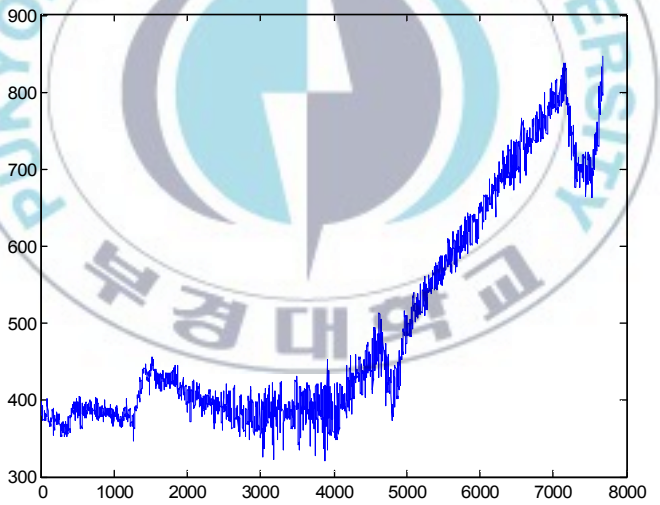
(a) Stand 1



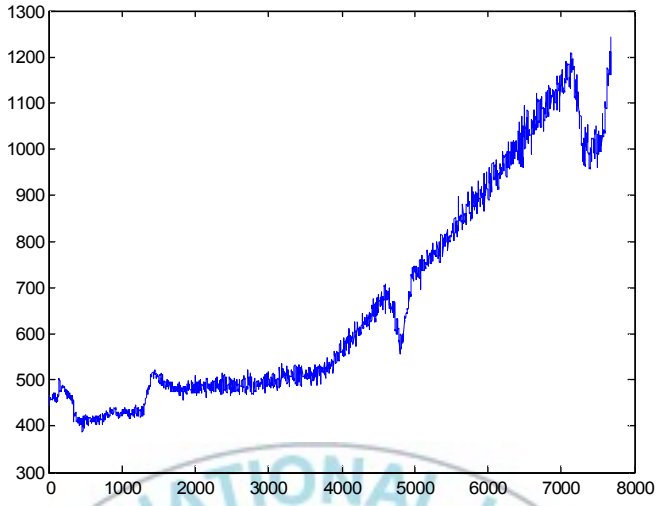
(b) Stand 2



(c) Stand 3

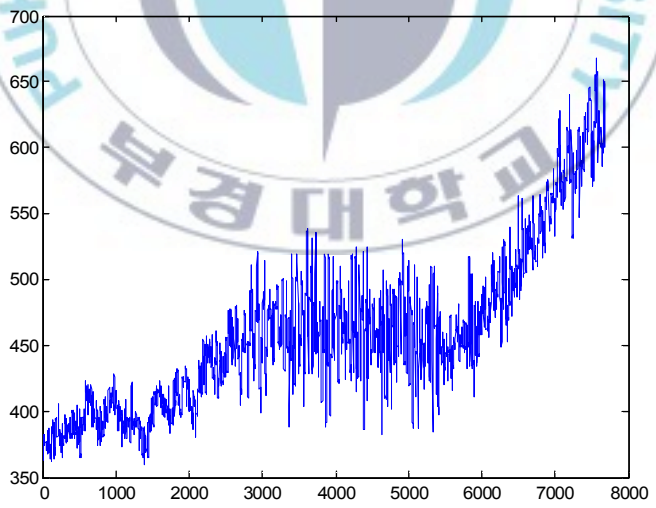


(d) Stand 4

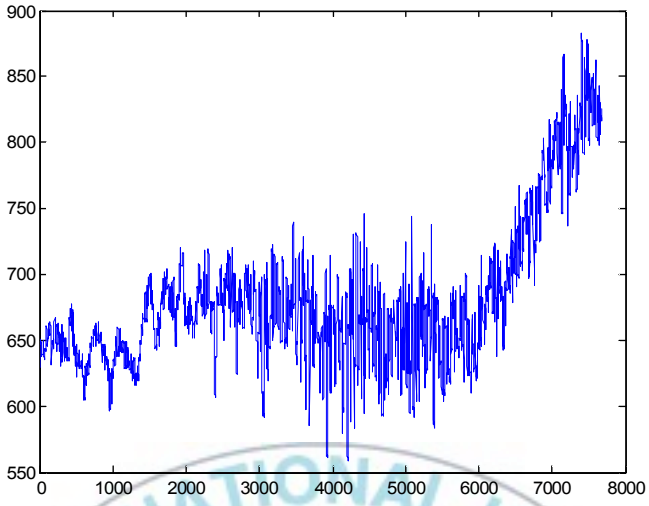


(e) Stand 5

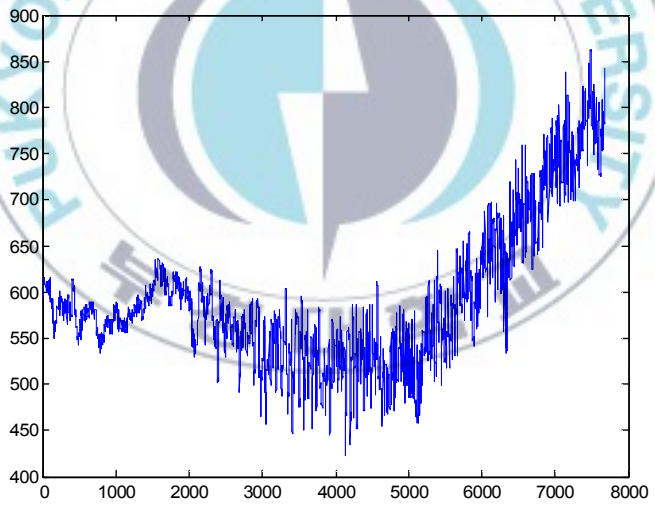
Fig. 4.4 Current raw signal of normal condition of stand 1~5



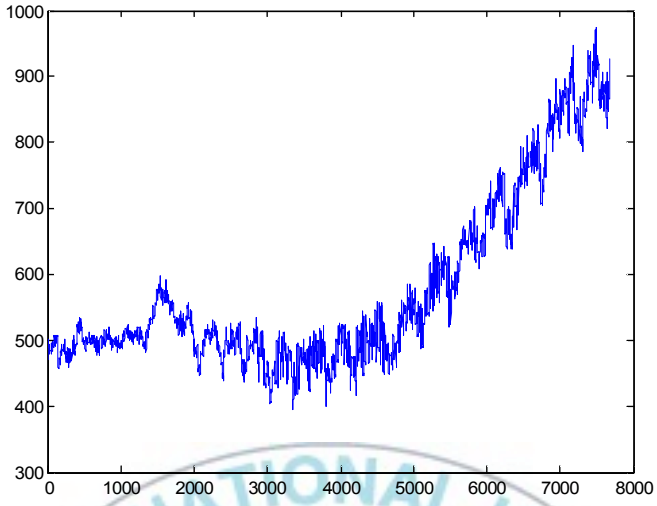
(a) Stand 1



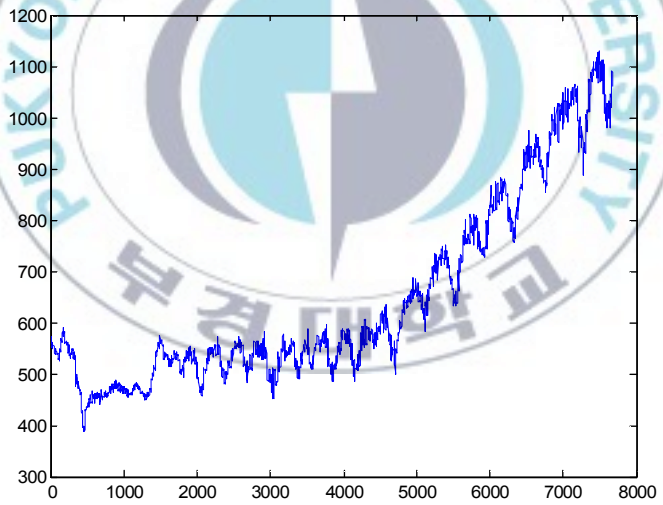
(b) Stand 2



(c) Stand 3



(d) Stand 4



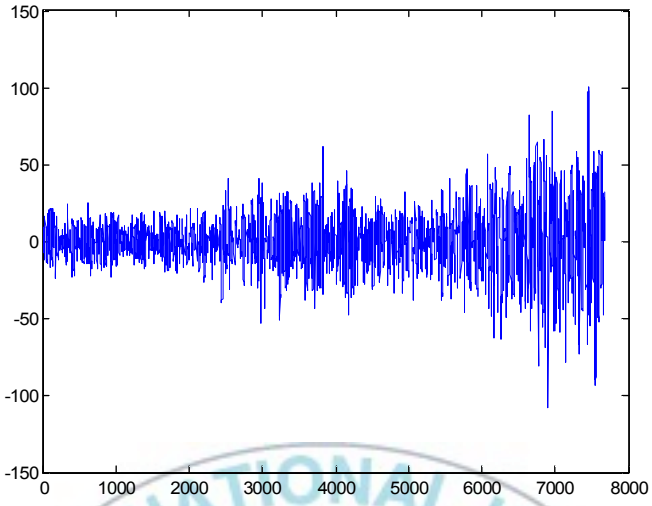
(e) Stand 5

Fig. 4.5 Current raw signal of fault condition of stand 1~5

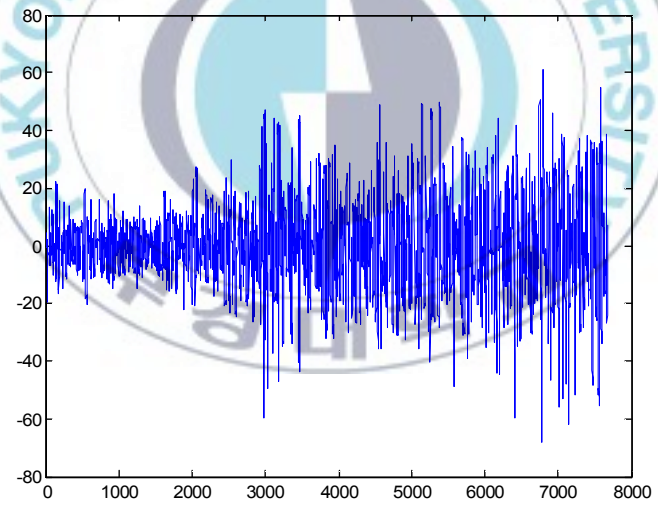
4.3 전처리

전류 신호에서는 전원 주파수(line frequency)의 영향을 없애거나 줄이는 절차가 요구되는데, 이를 위해 평활화 처리(smoothing process)가 수행되어야 한다. 이 과정을 통해 얻어진 잔여 신호(residual signal)를 Fig. 4.6과 Fig. 4.7에 나타내었다. 이 신호는 제품의 품질에 영향을 주는 장비의 상태를 알려주는 정보를 지니고 있다 [13]. 또한 변환된 신호는 다양한 크기를 갖는 정현 파형과 같이 나타나게 되어 과도 형태의 신호를 정상 형태의 신호로 변환되는 효과를 지니게 된다. 그러나 제시된 신호뿐만 아니라 전반적으로 신호 전체에서 두 상태를 구분할 수 있는 명확한 특징을 육안으로는 구별할 수 없었으므로 추가적인 기법이 요구된다.

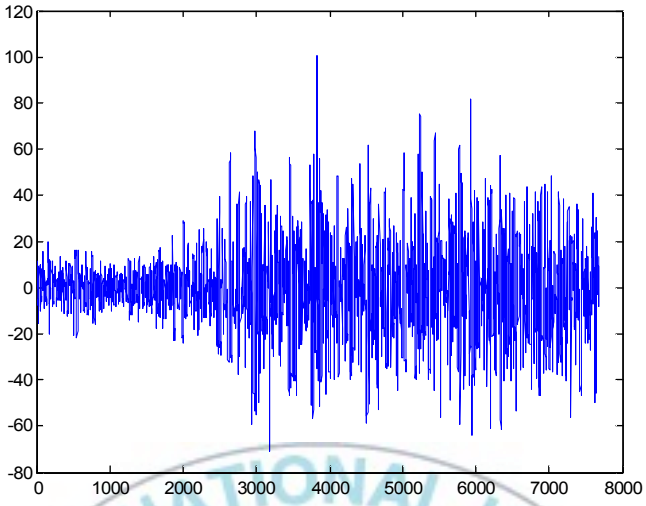
Fig. 4.8과 같이 신호의 정보를 보다 정확하게 얻기 위해 신호를 세분화(segment) 하였고 이를 5등분으로 나누었다. 기존의 신호를 이용하여 특징을 표현하게 되면 신호에서 표현하고 있는 모든 값들이 하나의 특징으로 표현되는데 반해, 세분화를 실시하게 되면 부분적으로 특징을 표현할 수 있으므로, 보다 구체적이고 세부적으로 그 신호의 정보를 표현할 수 있는 장점이 있다.



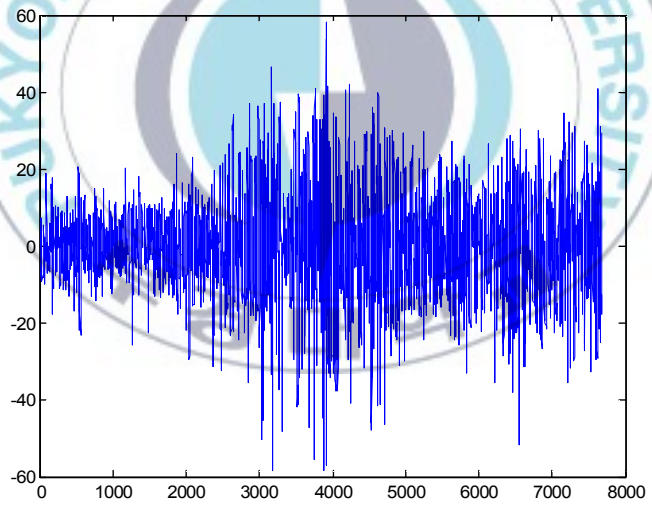
(a) Stand 1



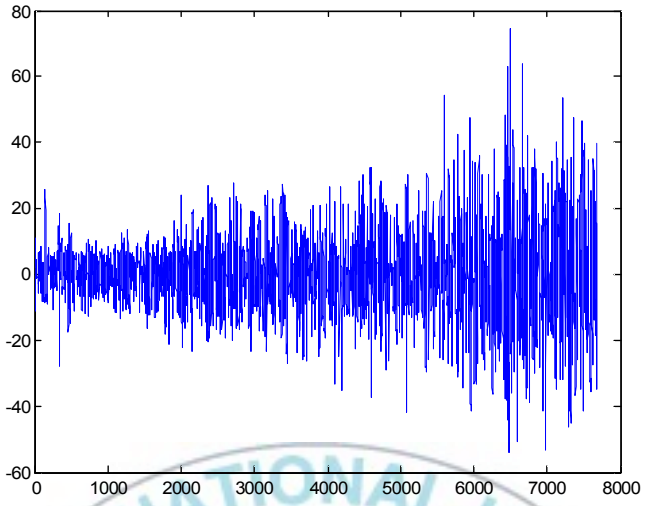
(b) Stand 2



(c) Stand 3

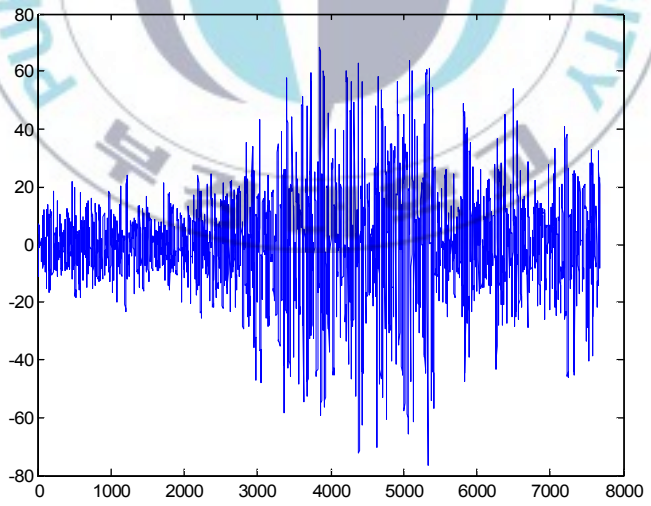


(d) Stand 4

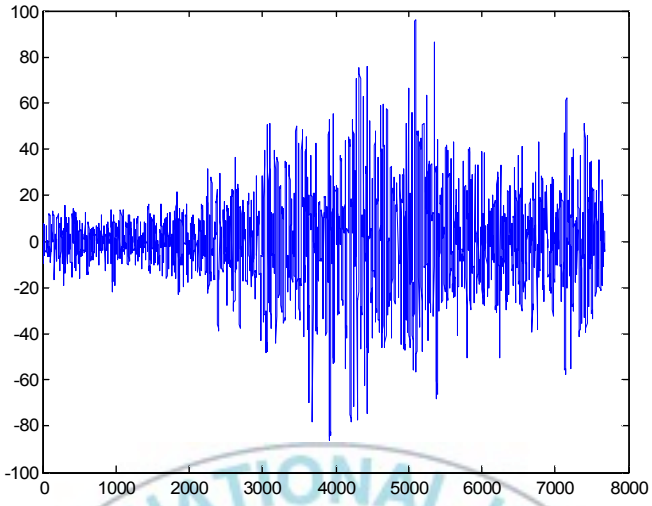


(e) Stand 5

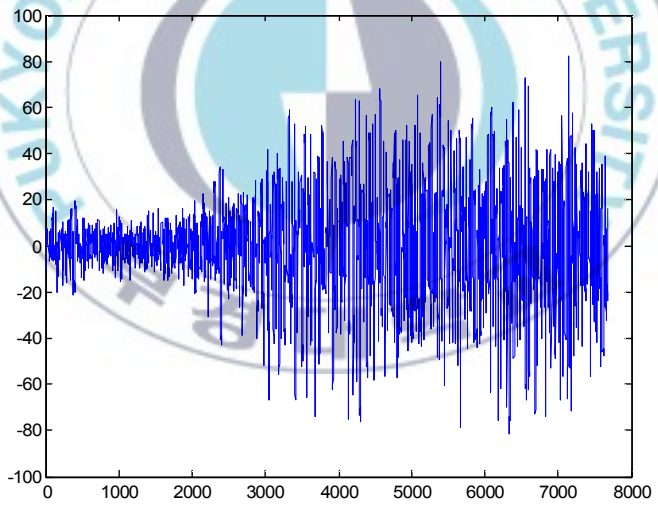
Fig. 4.6 Residual signal of normal condition of stand 1~5



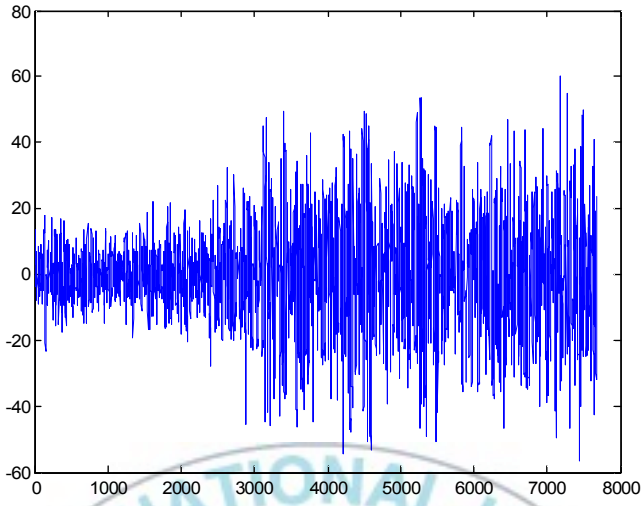
(a) Stand 1



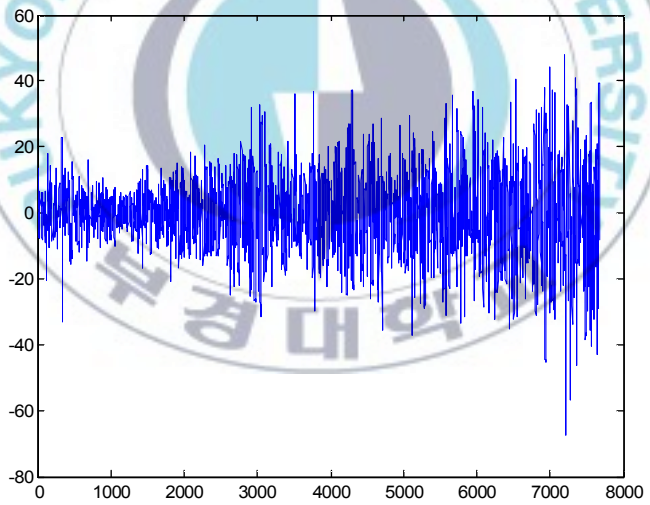
(b) Stand 2



(c) Stand 3



(d) Stand 4



(e) Stand 5

Fig. 4.7 Residual signal of fault condition of stand 1~5

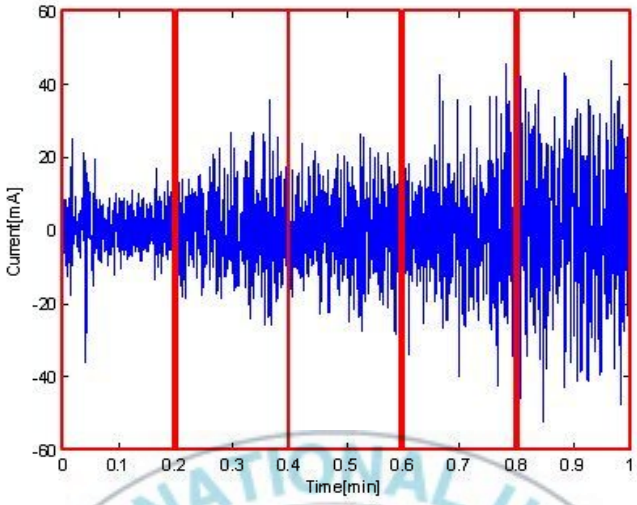
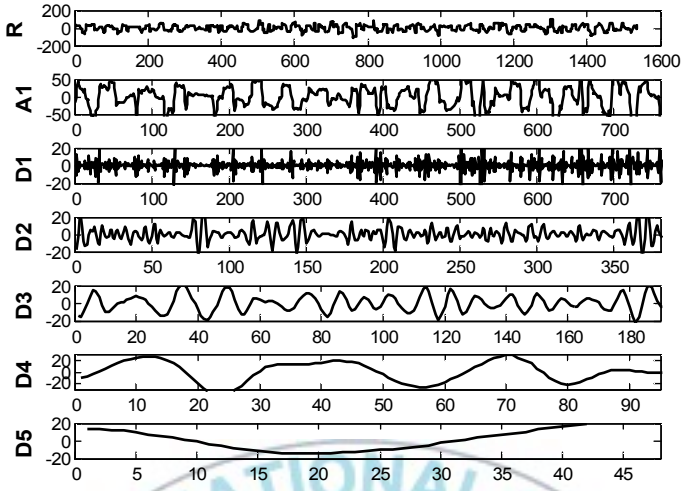


Fig. 4.8 Segmented residual signal in 5 divisions

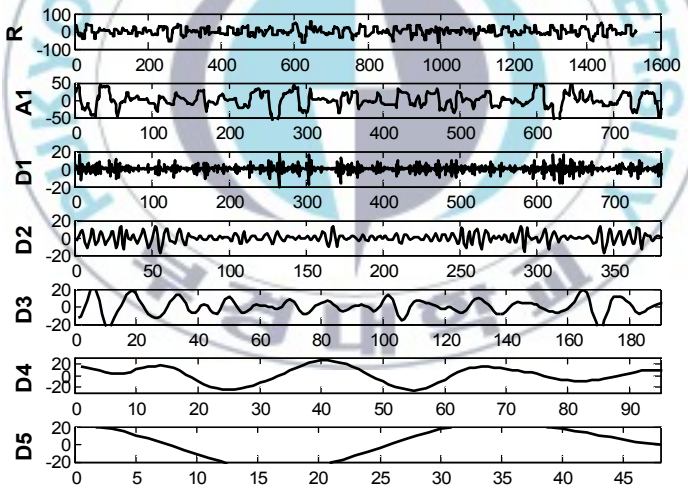
4.4 웨이블릿 변환

위의 과정에서 얻은 잔여 신호를 이용하여 웨이블릿 변환을 수행하였다. 기저함수는 현재 가장 많이 사용되며 좋은 성능을 보이고 있는 Daubechies 기저함수를 이용하였고 레벨은 실험을 통해 좋은 성능을 보인 db8로 설정하였다.

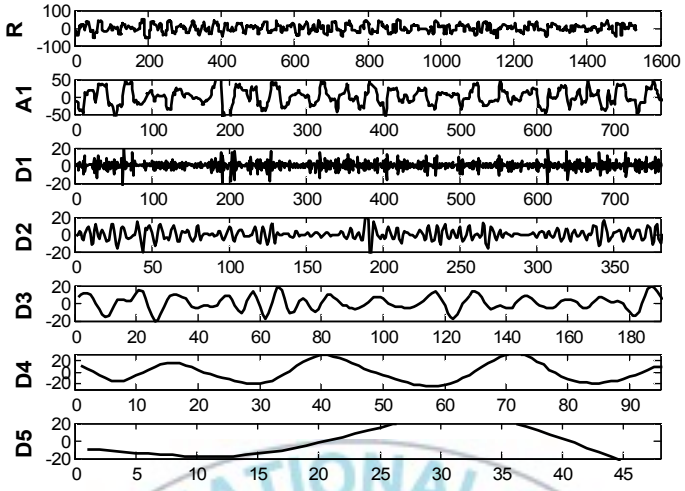
Fig. 4.9와 Fig. 4.10은 세분화된 정상 및 결함 데이터의 웨이블릿 변환 결과를 나타낸다. D1~5(Detail 1~5)은 각각의 주파수 영역에서 나타내고 있는 신호를 의미한다. 그러나 정상과 결함 상태에 대한 각각의 특성을 육안으로 구별하기가 매우 어려웠다. 즉, 각 상태의 명확한 구분을 위해 추가적인 방법이 요구되며, 취득된 각 Detail 신호를 이용하여 특징 계산을 수행하였다.



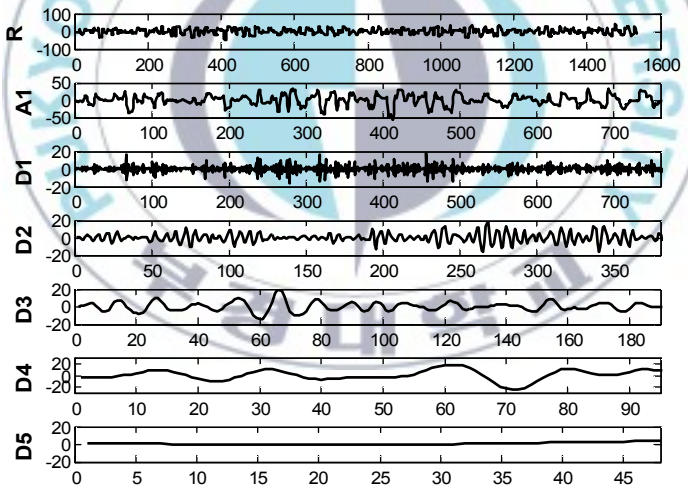
(a) Stand 1



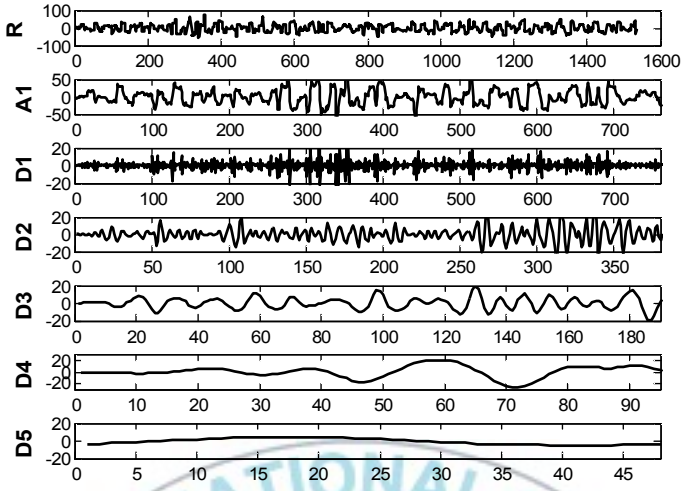
(b) Stand 2



(c) Stand 3

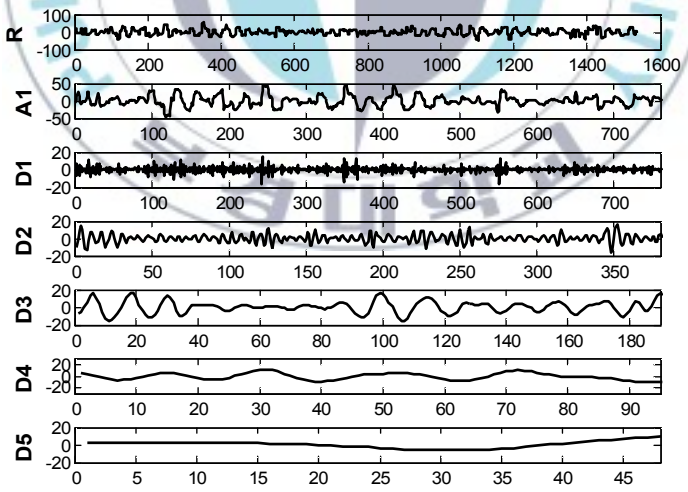


(d) Stand 4

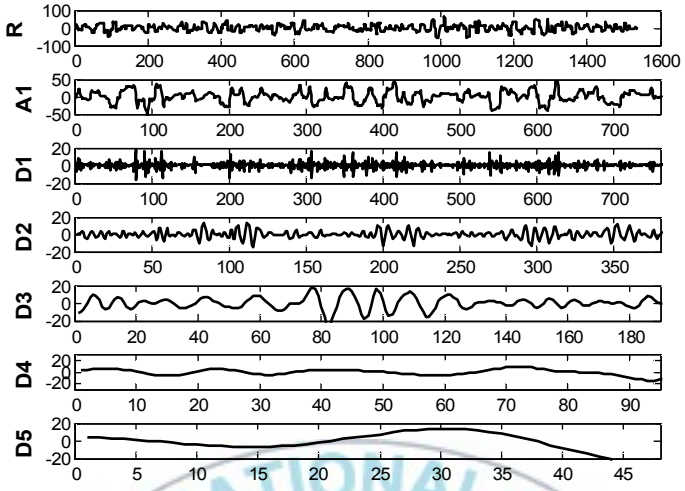


(e) Stand 5

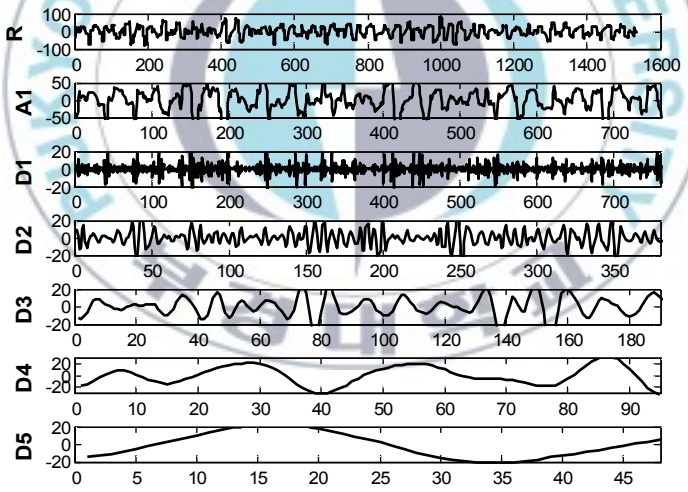
Fig. 4.9 Wavelet transform using normal signal of stand 1~5



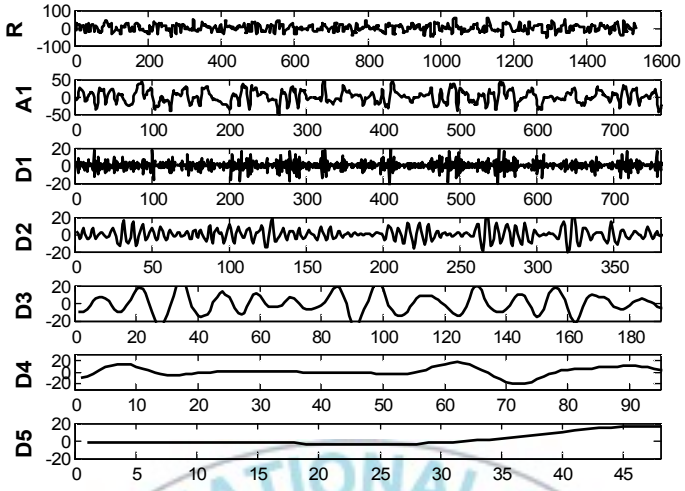
(a) Stand 1



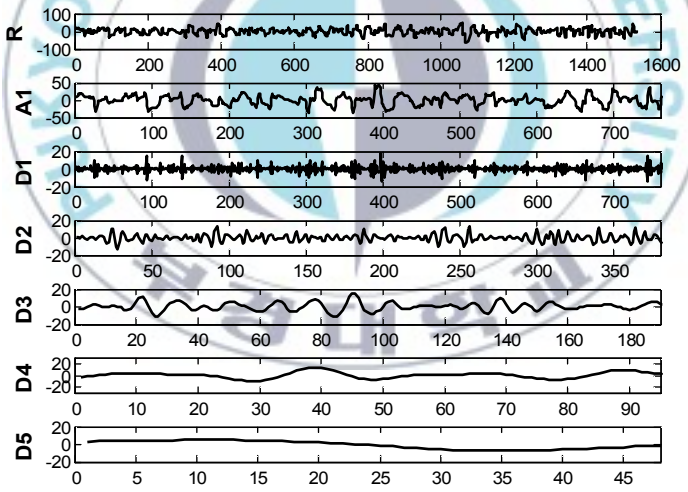
(b) Stand 2



(c) Stand 3



(d) Stand 4



(e) Stand 5

Fig. 4.10 Wavelet transform using fault signal of stand 1~5

4.5 특징 계산

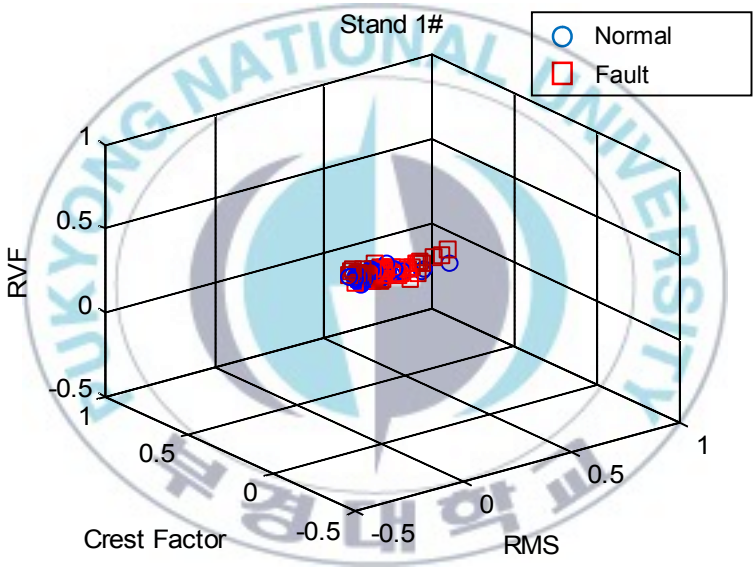
각 상태를 평가하기 위해 웨이블릿 변환을 통해 얻어진 각각의 Detail 신호를 3.2장에서 제시된 수식을 이용하여 하나의 특징값으로 표현하였다. 총 21개의 특징이 사용되었으며, 그 특징들을 Table 4.1에 제시하였다.

Table 4.1 Applied features

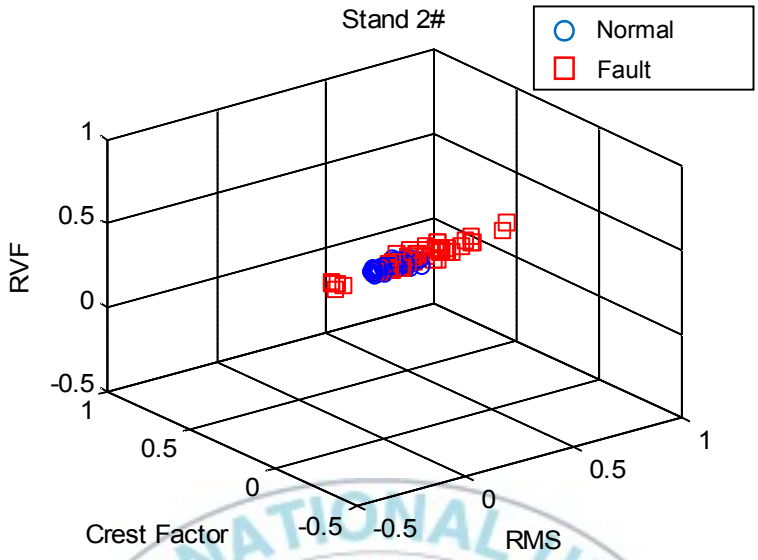
Domain		
Time	Frequency	Auto regression
Mean	Frequency center	AR coefficient
RMS	Root variance frequency	(a_1~a_8)
Kurtosis	Root mean square frequency	
Crest factor		
Shape factor		
Skewness		
Entropy estimation		
Entropy estimation error		
Lower histogram		
Upper histogram		

Fig. 4.11에 스탠드 1~5에 대한 정상 및 결함상태 간의 군집(clustering)을 나타내었다. 이 결과는 D3 데이터를 이용한 것으로 나머지 데이터를 이용하였을 때는 어느 스탠드에서도 군집의 형성이 이루어지지 않았다. 그림에서 기호 ○는 정상상태, □는 결함상태를 나타낸다. 적용된 특징 중에서 정상과 결함 간의 군집이 잘 이루어지는 가장 적절한 특징 3개를 이용하여 3차원 공간에 특징 데이터를 표시하였을 때, 스탠드 1에서 4까지의 결과는 스탠드 5의 결과에 비해 군집특성이 좋지 않은 것을 확인할 수 있다.

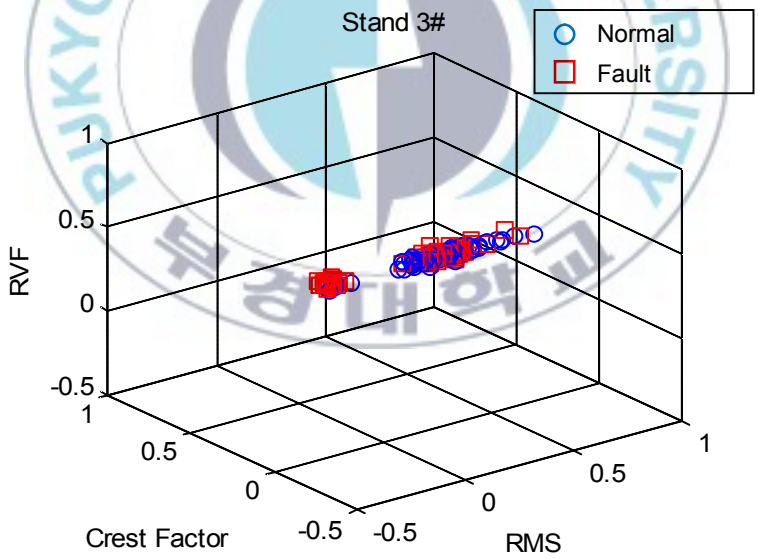
이는 스탠드 1~4에서 취득된 신호의 특징데이터는 판터짐 시에 발생하는 특징을 잘 표현하지 못함을 나타내고, 이로 인해 각 상태의 구분을 위한 사용에는 적절치 못하다는 것을 알 수 있다. Fig. 4.12 (e)의 스탠드 5에 대한 그림 중에 일부의 정상데이터가 결함데이터로 분류되고 있으나, RMS, Crest Factor 및 Root Variance Frequency(RVF)를 특징으로 하였을 때 상대적으로 스탠드 5에서 가장 군집이 잘 이루어짐을 확인할 수 있었다.



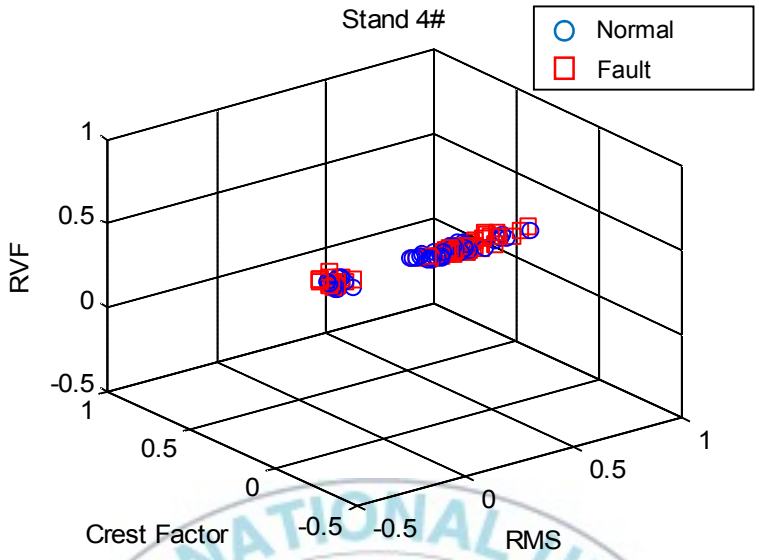
(a) Stand 1



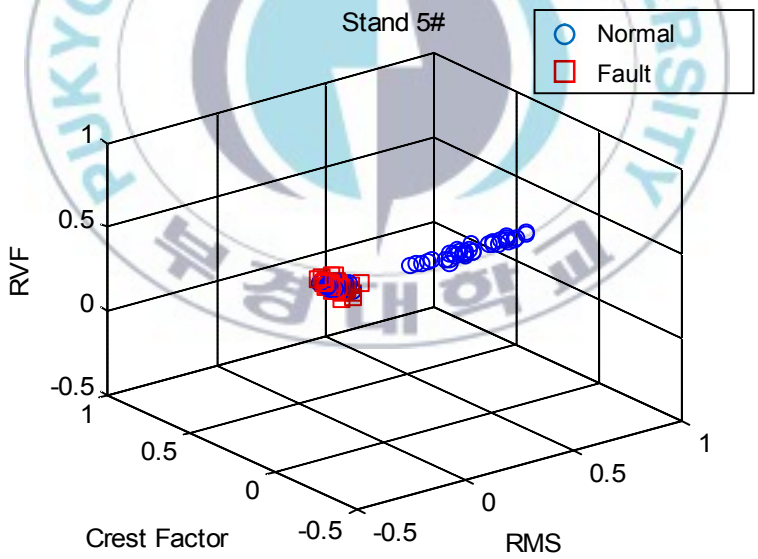
(b) Stand 2



(c) Stand 3



(d) Stand 4



(e) Stand 5

Fig. 4.11 Feature calculation of stand 1~5

4.6 특징 추출

특징 계산에서 얻어진 원래의 특징은 한 쪽 영역에서 혼합되는 형태를 가지고 있었으며, 이는 분류기의 성능을 떨어뜨리는 결과를 야기한다. 즉, 공분산 행렬(covariance matrix)의 고유치를 이용하여 특징 추출 및 차원 축소를 하는 성분 분석(component analysis)이 요구된다.

Fig. 4.12는 성분 분석에서의 특징 감소를 위한 공분산 행렬의 고유치를 나타낸다. 이 행렬의 고유치에서 99%가 되는 지점을 선정하여 차원 축소가 이루어지며, 여기서는 21개의 특징 중에서 10개의 성분 분석이 선정되었다[18].

21개의 특징 계산을 통해 얻어진 데이터를 바탕으로 PCA, ICA, KPCA 및 KICA를 수행하였다. 5개의 스탠드 중에서 이전의 특징 계산을 통해 스탠드 5에서 각 상태의 구분이 상대적으로 명확하게 이루어짐을 확인할 수 있었고, 이 데이터를 이용하여 특징 추출한 결과를 Fig. 4.13에 나타내었다. 그러나 특징 계산에서의 결과와 크게 다르게 나타나지 않았다. 아무래도 속도가 정상에 이르기 전인 초기 속도 상태에서 판이 터진다는 것과 결함 특성상 약하게 터지는 현상으로 인해, 그 특성이 신호에 나타난다 하더라도 매우 작기 때문에 정상과 결함간의 구분이 있어 어려움이 따른다는 것을 알 수 있었다. 그리하여 이런 약점을 극복하기 위해서는 보다 성능을 향상시킬 수 있는 전처리 과정이 수행되어야 할 것이 요구된다.

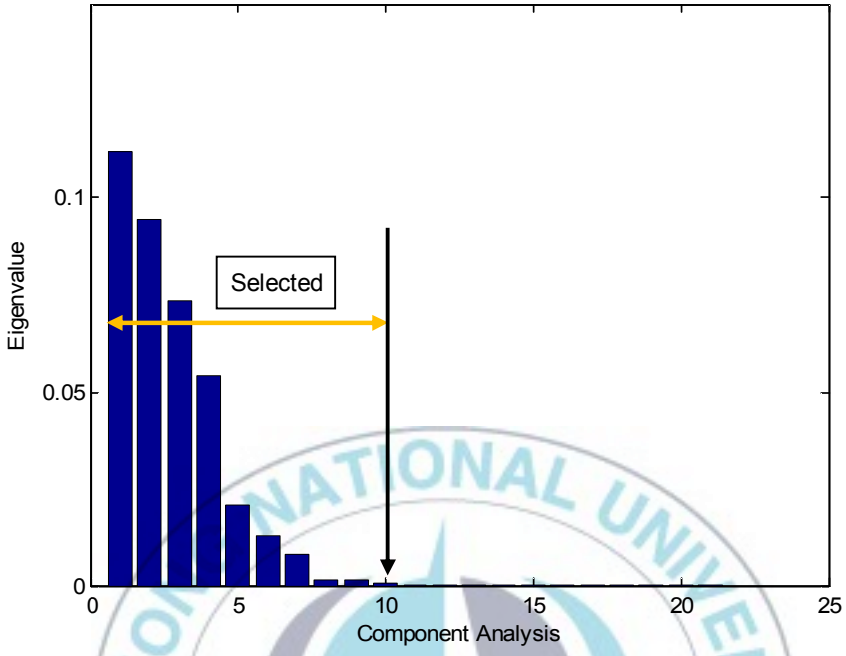
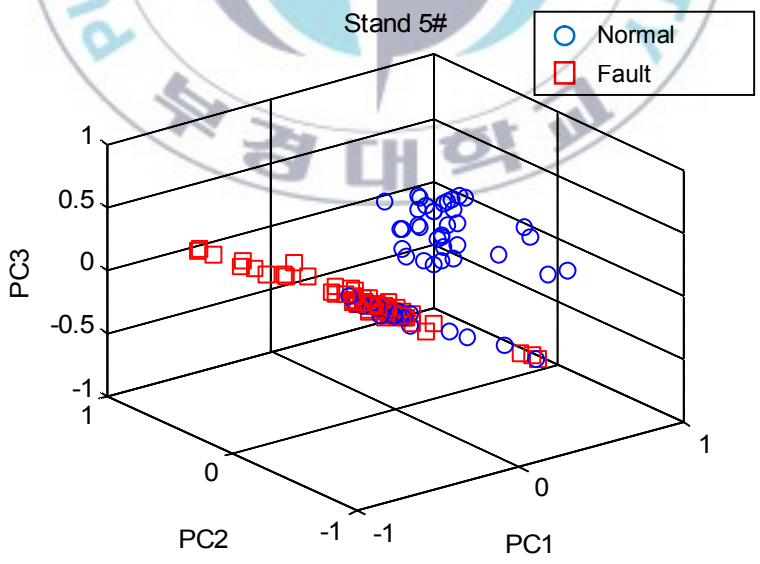
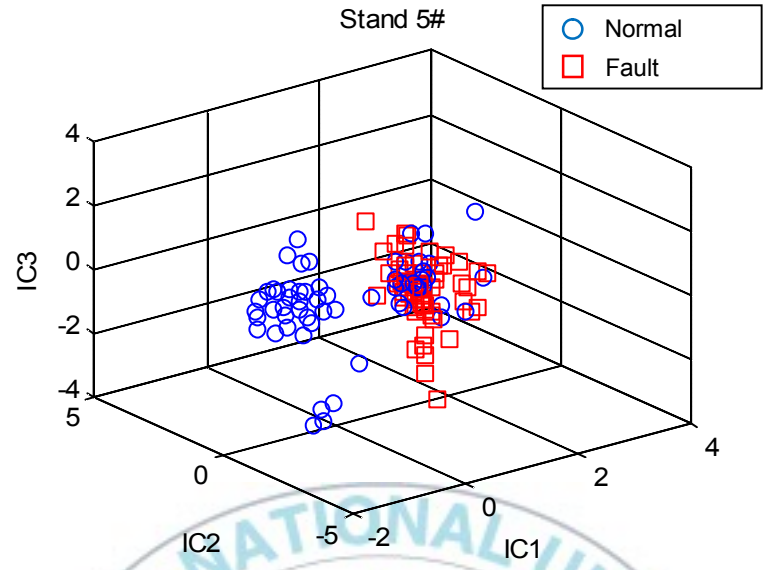


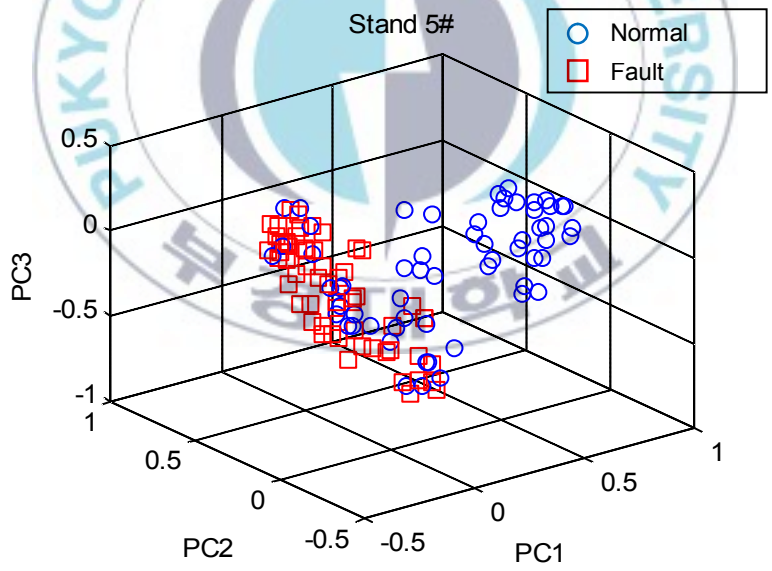
Fig. 4.12 Eigenvalue of covariance matrix for feature reduction



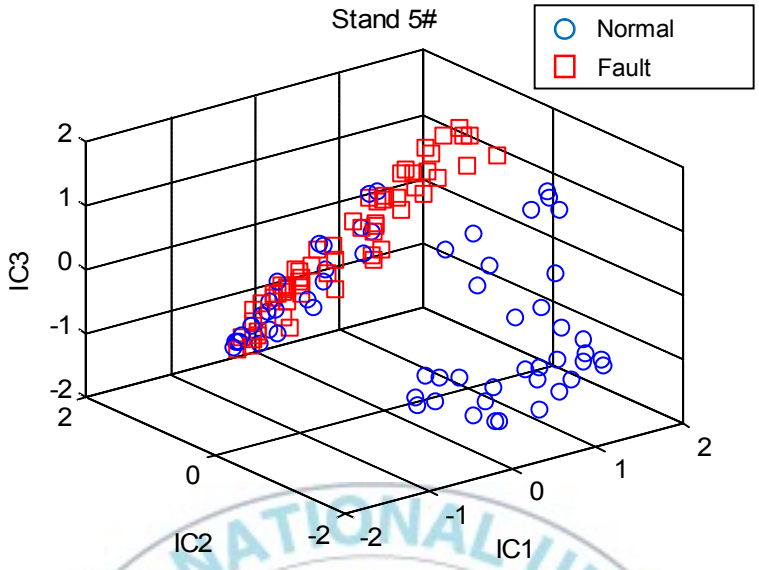
(a) PCA



(b) ICA



(c) KPCA



(d) KICA

Fig. 4.13 Feature extraction of stand 5 using PCA, ICA, KPCA and KICA

4.7 결함 분류

결함검지 성능을 향상시키기 위해 분류 알고리즘으로 SVM을 이용하여 정상과 판 터짐을 분류하였다. SVM에 사용된 커널 함수는 SVM의 기본 커널 함수로써 좋은 성능을 보이고 있는 RBF 커널 함수를 사용하였다. RBF 커널 함수에서 SVM 학습 과정에 이용되는 두 개의 파라미터가 있는데, C는 페널티 조건(penalty term)에서의 상계(upper bound)를 말하고, γ는 RBF 커널 파라미터이다. 이 2개의 파라미터를 이용하여 분류가 수행되므로, 좋은 분류율을 얻기 위해서는 파라미터 선정이 매우 중요하다. 따라서 이에 대한 방안으로 n-fold cross-validation을 적용하여 설정된 C와 γ 값을 n번만큼 반복 수행하여, 분류를 하는데 있어 발생하는 에러를 최소화하였다. 본 연구에서는 5-fold cross-validation으로 수행되었고 각 파라미터의 값의 범위는 $C = \{2^5, 2^6, 2^7, 2^8, 2^9\}$, γ

= $\{2^{-2}, 2^{-1}, 1, 2, 2^2, 2^3, 2^4\}$으로 선택되었다.

특징계산만을 통해 얻은 특징(Original)과 선형 기법인 PCA와 ICA, 비선형 기법인 KPCA와 KICA를 이용한 특징추출 결과, 총 5개를 이용하여 분류를 수행한 결과를 Table 4.2 ~ 4.6에 나타내었다. Table 4.3 ~ 4.6에서는 성분 분석의 수에 따른 분류율의 향상 정도를 확인하기 위해 그 수를 3개에서 10개까지 다양화하여 분류를 수행하였다. 스탠드 5에서 취득된 정상과 결함 데이터 각각 12개 중 7개를 학습 데이터로, 나머지 5개를 테스트 데이터로 사용하였다.

각 Table에서 훈련 정확도(training accuracy)는 학습 데이터 중 일부를 학습 데이터와 비교한 수치를 나타내고, 시험 정확도(testing accuracy)는 학습 데이터와 시험 데이터 간을 비교한 수치를 나타낸다. Number of SVs는 각 상태를 구분 짓기 위해 필요로 하는 SV의 개수를 나타내며, CPU Time은 분류하는데 요구되는 시간을 의미한다. 여기서는 정상과 판터짐의 구별 정도를 나타내는 시험 정확도가 가장 중요한 의미를 갖는다.

Table 4.2에 제시된 SVM과 Original 특징을 이용하였을 때는 78%의 분류 결과가 나타났으며, 요구되는 SV의 개수는 32개였다. RBF을 위한 적절한 파라미터는 C와 γ에서 각각 64와 4로 나타났다.

Table 4.3은 SVM과 PCA를 이용하였을 때의 분류 결과를 나타낸다. 학습 및 시험에 대한 정확성은 각각 82.1%와 82%이다. 대게 PC의 수가 증가할수록 학습 정확성이 향상됨을 확인되는데, 이는 특징이 많을수록 더 좋은 수행능력을 보여줌을 의미한다. 여기서는 9개의 PC를 사용하였을 때 가장 좋은 성능을 나타냈으며, 이 때 SV는 33개, $C = 32$ 그리고 $\gamma = 8$이었다.

SVM과 ICA를 이용한 분류 결과를 Table 4.4에 나타냈으며, IC의 수가 7일 때 가장 좋은 성능을 나타내었다. 학습 및 시험에 대한 정확도는 85.7%와 82%였으며, SV = 24, $C = 128$ 그리고 $\gamma = 16$이었다.

Table 4.5에서는 SVM과 KPCA를 이용한 분류 결과를 나타낸다. 학습 및 시

험에 대한 정확도는 83.9%와 84%이었으며, 이 때 PC의 수는 5였다. 적절한 SV의 수는 32개였으며, 파라미터 C는 32, γ는 1이였다.

SVM과 ICA를 이용한 분류 결과를 나타낸 Table 4.6을 보게 되면, 학습과 시험에 대한 정확도는 IC의 수가 9일 때 각각 80.4%와 84%로 가장 좋게 나타났다. 그리고 SV = 24, C = 128 그리고 γ = 16이었다.

Table 4.2 Classification result using SVM and Original feature

Accuracy (%)		Number of SVs	CPU time (s)	RBF kernel parameters	
Training	Testing			C	γ
80.4	78.0	32	0.130	64	4

Table 4.3 Classification results using SVM and PCA

PCs	Accuracy (%)		Number of SVs	CPU time (s)	RBF kernel parameters	
	Training	Testing			C	γ
3	78.6	80.0	33	0.142	32	16
4	82.1	80.0	29	0.108	256	1
5	76.8	80.0	36	0.022	32	2
6	80.4	80.0	31	0.090	128	1
7	80.4	78.0	31	0.016	512	4
8	82.1	80.0	27	0.118	128	4
9	**82.1**	**82.0**	**33**	**0.017**	**32**	**8**
10	83.9	80.0	28	0.329	128	16

Table 4.4 Classification results using SVM and ICA

ICs	Accuracy (%)		Number of SVs	CPU time (s)	RBF kernel parameters	
	Training	Testing			C	γ
3	80.4	80.0	26	1.097	64	8
4	78.6	82.0	24	0.027	128	4
5	84.0	80.0	23	0.129	32	8
6	83.9	78.0	25	0.133	32	8
7	**85.7**	**82.0**	**24**	**0.105**	**128**	**16**
8	87.5	76.0	19	0.049	512	4
9	89.3	76.0	24	0.334	128	16
10	89.3	76.0	27	0.059	32	16

Table 4.5 Classification results using SVM and KPCA

PCs	Accuracy (%)		Number of SVs	CPU time (s)	RBF kernel parameters	
	Training	Testing			C	γ
3	78.6	80.0	35	0.071	64	2
4	78.6	82.0	34	0.035	64	16
5	**83.9**	**84.0**	**32**	**0.162**	**32**	**1**
6	80.4	80.0	34	0.213	64	4
7	83.9	78.0	31	0.071	256	1
8	76.8	82.0	32	0.038	64	16
9	78.6	80.0	36	0.187	256	2
10	82.1	80.0	31	0.045	64	1

Table 4.6 Classification results using SVM and KICA

ICs	Accuracy (%)		Number of SVs	CPU time (s)	RBF kernel parameters	
	Training	Testing			C	γ
3	78.6	80.0	29	0.052	512	4
4	78.6	80.0	30	0.051	128	4
5	85.7	80.0	19	0.210	64	16
6	78.6	80.0	33	0.053	32	16
7	78.6	74.0	32	0.049	128	4
8	78.6	80.0	31	0.098	32	16
9	**80.4**	**84.0**	**37**	**0.044**	**32**	**8**
10	75.0	80.0	36	0.029	128	1

Fig. 4.14에서 SVM을 이용한 각 기법에 따른 분류 결과를 정리하였다. 전반적으로 각 기법 별로 유사한 결과를 나타냈으나, 특징 계산만을 수행하였을 때보다 특징 추출까지 수행하였을 때 더 좋은 분류 결과로 나타났으며, 커널 함수를 사용하였을 때가 그렇지 않았을 때보다 분류율이 더 좋았다. 학습 정확성까지 고려하였을 때, KPCA를 이용한 결과가 가장 좋은 것으로 나타났다.

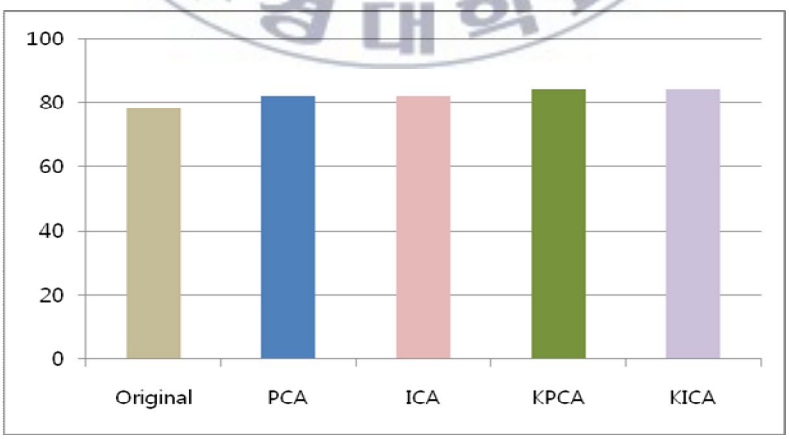


Fig. 4.14 Classification results using SVM (current signal)

Fig. 4.15와 Fig. 4.16에 각 스탠드의 상부와 하부 작업롤의 진동신호를 이용한 분류 결과를 나타내었다. 이 신호 역시 전류신호와 동일한 전처리 과정으로 수행되었다. 상부 작업롤의 진동신호는 D5, 하부 작업롤의 진동신호는 D3에서 정상과 판 터짐 간의 구분이 상대적으로 명확하게 이루어졌으며, 이 신호를 이용하여 분류가 수행되었다. 상부 작업롤의 진동신호의 경우, 전류신호를 이용하였을 때의 분류 결과와 유사하였으나, 훈련 정확도에서 Original 62.5%, PCA 69.6%, ICA 73.2%, KPCA 71.4% 그리고 KICA 71.4%로 전류신호에서와 비교하였을 때 낮게 나타났다. 이는 그 만큼 훈련 데이터의 신뢰도가 떨어진다는 것을 의미한다. 하부 작업롤의 진동신호의 경우는 훈련 및 시험 정확도 모두 전류신호보다 좋지 않았다. 훈련 정확도는 Original 67.9%, PCA 66.1%, ICA 80.4%, KPCA 67.9% 및 KICA 83.9%로 나타났다. 이를 통해 진동신호보다는 전류신호를 이용하였을 때 더 좋은 분류 결과가 나타난다는 것을 알 수 있다.

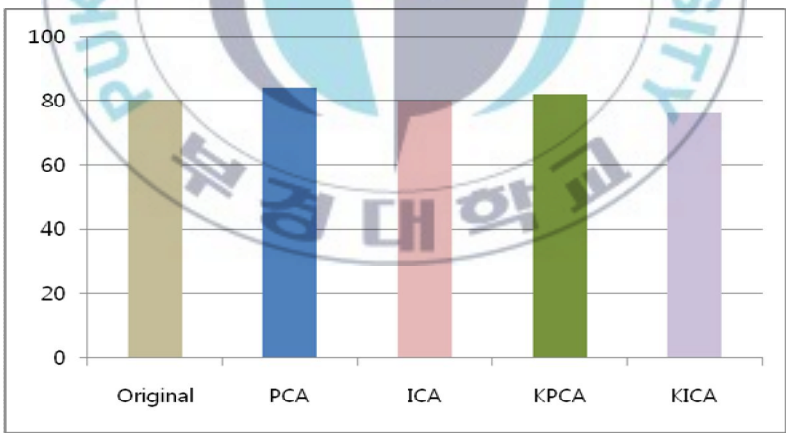


Fig. 4.15 Classification results using SVM (vibration signal – up)

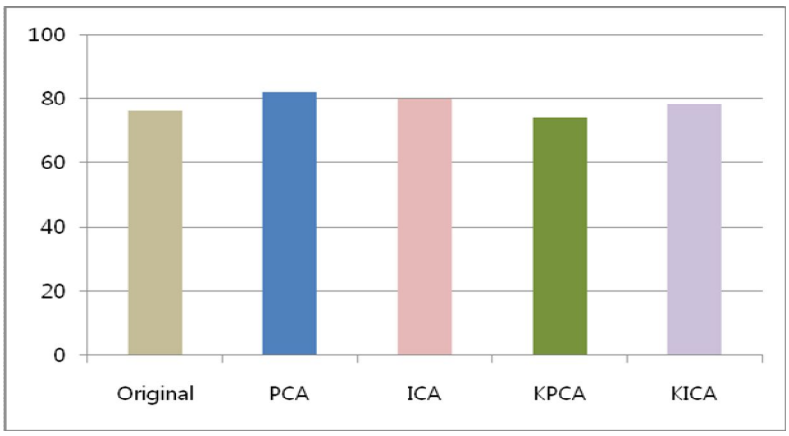


Fig. 4.16 Classification results using SVM (vibration signal – down)

5. 결 론

이 연구에서는 냉간 압연기에서 발생하는 판 터짐을 검지하기 위해 결함이 발생되는 초기 과도 구간의 전류 신호를 이용, 분석하여 검지하는 결함 검지 시스템을 제안하였다. 과도 구간은 정상 구간의 신호에 비해 노이즈가 덜 포함되어있어 더 많은 정보를 얻을 수 있는 것으로 알려져 있으나, 비정상적인 기본파 성분(non-stationary fundamental)을 가지고 있어 이를 제거하는 전 처리 과정에 많은 어려움이 있는 것이 사실이다. 그리하여 이 연구에서는 이를 해결하기 위해 전 처리로서 평활화(smoothing), segment 및 웨이블릿 변환을 이용하였다.

웨이블릿 변환에 이용된 기저함수는 현재 많이 사용되면서 좋은 성능을 보이고 있는 Daubechies 함수를 사용하였고 레벨은 db8로 선정하였다. 이를 통해 특징 계산을 수행한 결과, D3를 이용하였을 때 스탠드 5에서 두 상태의 구분이 이루어짐을 확인할 수 있었다. 이 때 사용된 특징은 RMS, Crest Factor 그리고 Root Variance Frequency였다.

분류 향상과 차원 감소를 위해 특징 추출을 수행하였고 공분산 행렬의 고유치 확인을 통해 원래의 특징으로부터 10개의 성분 분석이 선정되었다. 또한 어떤 개수에서 분류 성능이 좋은지 확인하기 위해 그 개수를 3개부터 10개까지 그 수를 달리하였으며, 추출 기법 별로 그 결과가 달리 나타난 것으로 확인되었다. 대게 수가 많을수록 더 좋은 분류 결과로 나타나는 게 일반적이나, 충분치 못한 학습 데이터의 수와 입력 데이터의 좋지 않은 품질이 분류율을 저해하고 있음을 확인할 수 있었다. 특징 추출 기법으로는 선형 기법인 PCA 및 ICA와 비선형 기법인 Kernel 함수가 적용된 KPCA와 KICA가 이용되었다.

분류를 위해 이진 분류 문제에 주로 사용되는 알고리즘인 SVM을 이용하여

정상과 결함 간의 구분을 수행하였다. 구분은 모든 스탠드에서 이루어지지 않았고 스탠드 5에서만 이루어졌으며, 이는 마지막 스탠드인 스탠드 5에서 부하가 가장 많이 걸리는 것과 관계가 있을 거라 사료된다. 또한 KPCA를 이용하여 분류를 수행하였을 때 가장 좋은 결과가 나타났다. 그리고 상·하부 작업롤의 진동신호를 이용하여 분류를 수행하였으나, 결과는 전류신호보다 좋지 못함을 확인하였다. 따라서 위에서 제시된 기법 및 데이터를 이용하여 스탠드 5를 중점적으로 감시하면 결함을 조기에 검지할 수 있을 것으로 생각된다.

참고 문헌

[1] 이원호, 신남호, 2002, 냉간 압연 공정에서의 판 파단 예지 알고리즘, 기계 관련 산학연 연합심포지엄 강연 및 논문 초록집, pp. 1079~1084

[2] 신남호, 강명구, 임은섭, 1998, 냉간 압연기에서 채터 마크 예방에 관한 연구, 한국정밀공학회 학술발표대회 논문집, pp. 335~338

[3] J. Mackel, 1999, Condition Monitoring and Diagnostic Engineering for Rolling Mills, International Congress of COMADEM. pp. 1~12

[4] H. Douglas, P. Pillay and A. Ziarani, 2004, A New Algorithm for Transient Motor Current Signature Analysis Using Wavelet, IEEE Trans. Industry Applications, Vol. 40, No. 5, pp. 1361~1368

[5] H. Douglas, P. Pillay and A. Ziarani, 2005, The Impact of Wavelet Selection on Transient Motor Current Signature Analysis, IEEE ICEMD, pp. 80~85

[6] Yahoo 백과사전, http://kr.dictionary.search.yahoo.com/search/dictionaryp?pk=15887100&p=%BE%D0%BF%AC&field=id&type=enc&subtype=enc

[7] 서남섭, 최신 기계공작법, http://sns.chonbuk.ac.kr/manufacturing/mclass-3-8.htm

[8] Naver blog, http://blog.naver.com/yunbh1004?Redirect=Log&logNo=80027115235

[9] POSCO homepage, http://www.posco.co.kr/homepage/docs/kor/html/product/ info/s91e6000010c.html

[10] 황원우, 2004, Support Vector Machine을 이용한 회전기계의 상태 분류 및 결함 진단, 부경대 공학석사 학위논문, pp. 8~18, pp. 29~39

[11] 김병욱, 1999, Wavelet 변환을 이용한 회전기계의 이상진동진단, 부경대 공학석사 학위논문, pp. 3~9

[12] T. Han, 2005, Development of a Feature-Based Fault Diagnostics System and Its

Application to Induction Motor, 부경대 공학박사 학위논문, pp.12~15

[13] A. Widodo, B. S. Yang, 2008, Wavelet Support Vector Machine for Induction Machine Fault Diagnosis Based on Transient Current Signal, Expert System with Applications, Vol. 35, No. 1-2, pp. 307~316

[14] 안경룡, 2002, 인공신경망을 이용한 회전기계의 고장진단, 부경대 공학박사 학위논문, pp. 15

[15] A. Widodo, 2007, Support Vector Machine for Machine Fault Diagnosis and Prognosis, 부경대 공학박사 학위논문, pp.50~57, pp.58~65

[16] B.S. Yang, A. Widodo, 2010, Intelligent Machine Fault Diagnosis and Prognosis, Nova Science Publishers, New York, pp. 93~116 (출판 예정)

[17] 한학용, 2005, 패턴인식 개론: MATLAB 실습을 통한 입체적 학습, 한빛미디어, pp.518~524

[18] A. Widodo, B.S. Yang, 2009, Intelligent Fault Diagnosis System of Induction Motor Based on Transient Current Signal, Machatronis, Vol. 19, Issue 5, pp. 680~689

감사의 글

이 논문이 나오기까지 정성 어린 지도와 관심을 주신 지도교수 김선진 교수님께 깊은 감사를 드립니다. 아울러 이 분야의 관심을 석사과정으로 이끌어 주신 양보석 교수님의 많은 조언으로 인해 이렇게 논문을 완성할 수 있게 된 것에 대해 감사 드립니다.

회사 업무를 통해 진동 분야에 관심을 갖고 자습하며 보낸 시간이 실험실 생활로 이어지게 되었습니다. 그 시작이 작년 7월. 비록 남들에 비해 늦은 시작과 부족한 지식이지만, 보다 더 진취적인 사고방식으로 앎을 위해 지속적인 관심과 노력으로 이를 채워나가도록 하겠습니다.

실험실 생활 시작부터 양질의 조언을 아끼지 않고 힘들 땐 삶의 쉼터가 되어준 성실의 대명사인 종덕 형, 형으로 인해 지적으로 많은 깨달음을 얻었습니다. 과제 및 각종 업무로 인해 매우 바쁜 머리 좋은 민찬 형, 연구에 관한 궁금증을 풀어주는 Widodo박사님, 가끔씩 좋은 지식을 나누어 주시는 진대 형, 그리고 파트타임으로 매주 학교 오는 부지런한 상윤 형, 대우조선에 근무 중이며 많은 격려와 관심 주신 영모 형님, 또한 좋은 정보 제공해 주신 이원호 박사님 모두 감사합니다.

나의 동기들: 이번 연구에 많이 도와준 재미있는 친구 Caesar, 마음 여리고 순진한 Younus, 말 길을 잘 못 알아듣지만 똑똑한 Thom 그리고 같이 졸업 못해 아쉽고 이상하게 빨리 친해진 김곤 부장님, 모두 저의 동기인 것이 자랑스럽습니다.

논문 때문에 고민하는 술꾼 원정이, 코드를 잘 짜는 드라마 매니아 정민, 지금은 여기에 없지만 각자 삶을 살아가고 있는 유부남 성도, 운동이 필요한 술꾼 성원, 파트타임이 된 애희 그리고 해외에서 연구하고 있는 학과장 Tung교수님, 착한 Niu Gang박사님, 실험실 온지 얼마 안된 부지런한 선용, 막내 훈민, 모두들 감사합니다.

평생 함께할 영원한 벗, 민식, 상우, 인호, 영호, 준호에게도 감사를 표합니다.

저를 위해 기도해 주신 임용환 목사님, 민형, 정호 형, 남희 누나를 비롯한 다른 청년부원들 모두의 기도가 대학원 삶에 있어 큰 힘이 되었습니다. 감사 드리며 앞으로도 많은 기도 부탁 드립니다.

끝으로 저의 삶에 있어 주인공이 되시는 부모님, 표현을 잘 못해도 항상 감사함을 느끼고 있습니다. 자식 걱정 안 하시도록 멋진 삶으로 보답해 드리도록 하겠습니다.