

PCA-LDA 알고리즘을 이용한 고체절연물의 열화도 판별

박성희¹, 강성화², 임기조¹

충북대학교 전기전자컴퓨터공학부¹, 충청대학 소방안전과²

Evaluation on Degradation of Solid Insulator by PCA-LDA algorithm

Seong-hee Park¹ Seong-hwa Kang², Kee-joe Lim¹

Chungbuk National University¹, Chungcheong university²

Abstract - Electrical treeing occurrence is caused by some defect in solid insulator. Those are accompany the PD(partial discharge) occurrence. And lifetime of the insulator is affected by PD. So, detection of electrical treeing is important thing as this view. Especially, detection of the end treeing is more important and have meaning for industrial engineering because concerned with maintenance and replacement of equipment. In this paper, evaluation of treeing process were studied and PCA(principle component analysis)-LDA(linear discriminant analysis) as classification method were used. The result is present the good recognition.

1. 서 론

고체절연물은 여러 가지 절연물에서 가장 널리 사용되고 있어 전력기기내에서 중요한 역할을 차지하고 있다. 전기트리 현상은 전기기기의 실제적인 수명에 중요하게 영향을 주는 요소로서 국부 고전계 영역에서 시작되며 전기트리의 진전에는 부분방전을 수반하게 된다. 절연체 내부에 존재하는 오염물, 전극과 폴리머 경계면 사이의 갈라짐, 섬유로 보강된 절연체의 경우 반도체층의 돌출부, 그리고 절연체의 벌크내에서 가스로 채워진 보이드들과 같은 결함들은 국부적인 전계가 집중되어 전기트리가 생성될 수 있는 지점들이며 전기트리의 발생요인으로서는 침침단에 가해지는 반복되는 응력에 의한 기계적 피로, 침 침단에 존재하는 보이드, 기포등에 의한 미소 부분방전, 전하 주입 및 추출에 의한 전기트리의 발생등과 이론들이 제안되어 왔다. 이렇듯 전기트리는 다양한 형태로 진전이 될 수 있으며, 그것에 대한 검출은 상당히 난해한 점이 있는 것이 사실이다.

공학적으로 보았을 때, 전기트리의 검출은 그 말기진전시의 검출이 중요한 의미를 갖는데, 이는 말기시에는 전력기기자체의 교체 혹은 유지/보수등의 작업이 필요하기 때문이다. 결국 본 논문에서 연구의 목적은 이것에 있다고 할 수 있겠다. 즉, 전기트리의 각 성장별로 초기, 중기, 말기로 나누어 데이터를 획득한 후 각 시기별 패턴을 분류하여 말기시의 분류를 명확하게 하고자 하였다.

또한 본 논문에서 사용되는 데이터는 전기트리의 진전시에 수반되는 부분방전 데이터를 이용하였으며, 각 진전시기별로 데이터를 취득하였다. 각 진전시기별 분류는 PCA(principle component analysis-주성분분석기법)-LDA(linear discriminant analysis-선형판별분석기법)알고리즘을 적용하였다. PCA-LDA 알고리즘은 각 패턴에 대한 주성분이 특징벡터를 추출하여, 그것으로부터 분류를 하는 기법과 같은 클래스간의 분산과 다른 클래스 사이의 분산의 비가 최대가 되는 행렬을 구하여 적용하는 기법의 융합이다. PCA-LDA 기법에 대해서는 다

른 지면을 통해 보다 많은 설명을 할 것이다.

PCA-LDA에 입력데이터로서 트리진전시 발생하는 부분방전 데이터를 적용하여 진전시기별 분류를 하였다. 또한 전기트리 방전 데이터는 PD detector system을 이용하여 취득하였다.

2. 실험 및 결과

2.1 실험

그림 1은 트리방전 발생을 위한 구성을 나타내고 있다. 전기트리방전 실험에 사용된 절연재료는 대한전선에서 생산되고 있는 22.9kV용 피복 전선의 XLPE 절연물 부분을 절개한 시편을 사용하였다. 절연물에 침을 삽입할 때 침과 절연물 사이의 경계면이 완전히 밀착되도록 하기 위해 XLPE에 120C의 온도의 hot press를 이용하여 가열·압축한 상태에서 시료를 제작하였다. 실험시에는 침단과 절연물 표면에서 발생하는 표면방전을 막기 위해 침전극 표면, 절연물과 침이 닿는 면 등에 절연재료인 에폭시를 도포하였다. 또한 코로나 방전과 표면방전의 발생을 방지하기 위하여 절연유에 함침하여 실험하였다.

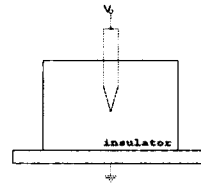


그림 1. 전기트리 방전 발생모델

그림 2는 모의 실험을 통한 결함의 부분방전 측정을 위한 실험장치의 구성을 나타내고 있다. 전기트리진전시의 측정은 아래의 구성도와 같이 PD free 변압기로 전압인가 후 PDASDA(partial discharge acquisition, storage and display system)에서 계측을 한다. 부분방전 펄스의 검출은 부분방전 검출기(Biddle Instrument사 제작, AVTM 662700Ja)로 검출한다. 검출된 방전펄스는 컴퓨터 프로그램을 통해 최대방전량, 평균방전전류, 발생빈도수 등을 연산하고 $\phi-q-n$ (위상-방전량-방전회수) 분포를 도출하게 된다. 특별히 전기트리 방전시에는 트리가 미세한 부분에서 발생하기 때문에 육안 관찰은 어렵다. 그런 이유로 트리의 성장속도를 측정하기 위해 시편에 광학 현미경(Olympus, SZ1145TR)을 설치하고 OCD(Hitron K-450) 카메라를 장착하여, 전기트리 성장을 측정하였다. 전기트리의 처음 시작은 상당히 빠른 시간에 진행이 되므로 검출에 대한 어려움이 있을 수 있다. 따라서 본 실험에서는 전기트리 발생 후 전체구간의 10% 성장시점을 트리 발생개시시점으로 선정하여, 개시 후 20 - 30%시점을 초기, 50% - 60% 성장시점을 중기,

80% 이상 성장시점을 말기로 구별하여 데이터를 취득하였다.

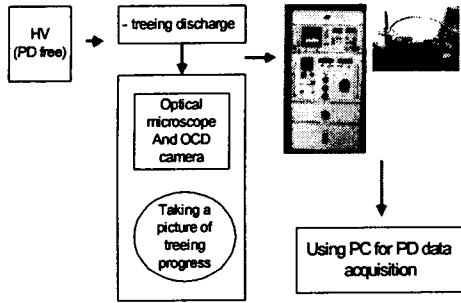


그림 2. 전기트리 및 부분방전 측정장치

그림 3은 부분방전 데이터의 측정 후, 통계적 분포의 도출, 알고리즘의 적용을 통한 패턴의 분류까지 일련의 과정을 나타내고 있다.

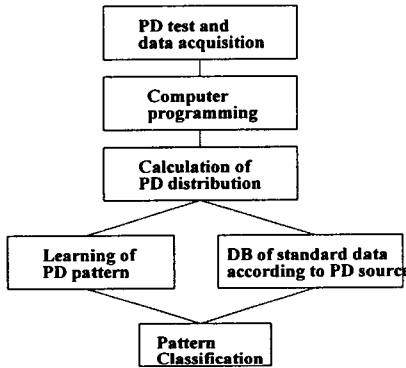


그림 3. Block diagram of pattern classification

2.2 PCA-LDA

2.2.1 PCA(principle component analysis)

주성분 분석은 서로 상관관계가 있는 벡터들이 존재할 때, 이러한 벡터들이 가지고 있는 정보들을 최대한 확보하는 적은 수의 새로운 벡터들을 생성하는 방법이다. 즉 서로 연관이 있는 벡터들에 대하여 주요한 관심사 중의 하나는 이 벡터들이 가지고 있는 변이를 측정하는 것이다. 여기서의 변이란 이 벡터들이 가지고 있는 정보의 양이라고 할 수 있으며, 주성분 분석에서는 원래의 벡터들이 가지고 있는 변이의 양을 가장 많이 확보하는 순서대로 변수들의 선형결합을 이용하여 새로운 변수를 구한다. 만일 P개의 a×b 사이즈를 가진 고차원 학습벡터를 N(a×b)으로서 1차원의 학습 벡터로 변환하여 $X = \{x^1, x^2, \dots, x^P\}$ 라고 가정하면 X에 대한 평균은 식 (1)과 같이 표현할 수 있다.

$$m = \frac{1}{P} \sum_{i=1}^P x^i \quad (1)$$

이 식에 1차원 벡터의 차를 구하면 식 (2)와 같다.

$$\bar{x}^i = x^i - m \quad (2)$$

그리고 식 (2)를 N×P 행렬로 이루어진 \bar{X} 를 식 (3)과 같이 표현하면 N×N 사이즈의 공분산 행렬을 식 (4)와 같이 표현할 수 있다. 다음으로 공분산 행렬 Ω 의 고유벡터를 구하여 \bar{X} 를 고유벡터 공간으로 사상시킬 수 있으며, 공분산의 고유벡터를 구하는 식은 식 (5)와 같이 나타낼 수 있다.

$$\bar{X} = [\bar{x}^1, \bar{x}^2, \dots, \bar{x}^P] \quad (3)$$

$$\Omega = \bar{X} \bar{X}^T \quad (4)$$

$$\bar{X} \bar{X}^T v_j = \lambda v_j \quad (5)$$

식 (5)에서 구해진 고유벡터 v_j 를 이용하여 학습신호 x 에 대한 고유벡터를 구하면 다음식과 같다. 즉, 입력신호에 대한 특징벡터는 공분산에 대한 고유벡터를 이용하여 PCA 변환된 공간으로 투영함으로써 얻을 수 있다.

$$z_i = v_j^T (x_i - m) \quad (6)$$

만일 공분산 행렬 Ω 의 사이즈 N×N가 클 경우 스냅샷 방법을 이용하여 아래와 같이 효율적으로 공분산 행렬의 고유벡터를 구할 수 있다. 먼저 식 (5)에 \bar{X}^T 를 곱해주고 v_j 를 V로 대체하면 식 (7)과 같이 된다.

$$\bar{X}^T \bar{X} (\bar{X}^T V) = \lambda (\bar{X}^T V) \quad (7)$$

식 (7)에서 우변에 식 (8)과 같이 V'를 대체하면 식 (7)은 식 (9)와 같이 나타낼 수 있다.

$$V' = \bar{X}^T V \quad (8)$$

$$\bar{X}^T \bar{X} (V) = \lambda (V) \quad (9)$$

식 (9)에서 V'는 초기 분산 대칭행렬 Ω 인 $\bar{X}^T \bar{X}$ 의 고유벡터를 나타낸다. 따라서 V'의 값은 (P×P) 사이즈를 갖는 $\bar{X}^T \bar{X}$ 의 행렬의 고유벡터를 구하면 되므로 N×N 사이즈를 갖는 원 분산행렬 Ω 인 $\bar{X} \bar{X}^T$ 의 고유벡터를 구하는 것보다 계산량을 줄일 수 있는 장점을 갖는다. 마지막으로 식 (9)에 의해 구해진 V'를 이용하여 최종적으로 알고자 하는 분산행렬 Ω 에 대한 고유벡터 V를 구하면 식 (10)과 같고,

$$\bar{X}^T = \bar{X}^{-1} \text{이므로 최종적으로 구하고자 하는 } V' \text{는}$$

식 (11)과 같다. 이러한 식 (11)을 이용하여 학습영상을 식 (6)에 의해 고유벡터 공간으로 투영하여 학습영상에 대한 특징벡터를 추출한다.

$$V = (\bar{X}^T)^{-1} V' \quad (10)$$

$$V = (\bar{X} V') \quad (11)$$

표1. PCA의 특징추출과정

Step	Process
1	Transition from measured PD signal for learning to one dimension column vector, then calculate the mean of total signal from following formula $m = \frac{1}{P} \sum_{i=1}^P x_i$
2	Make the covariance by below formula $\Omega = \bar{X} \bar{X}^T, \bar{X} [x_1, x_2, \dots, x_P] \bar{x}_i = x_i - m$
3	Obtain eigen value λ and eigen vector v_j from covariance Ω by following formula $\Omega v_j = \lambda v_j$
4	Obtain the feature vector as presented transit PCA $z_i = v_j^T (x_i - m)$
5	Extraction feature from input PD data by PCA $z_i = v_j^T (x_i - m)$ Where x_i is input PD data
6	Calculation distance between feature vector z_i and input vector z_i'

2.2.2 LDA(linear discriminant analysis)

선형판별분석은 클래스 내의 분산을 나타내는 행렬 (Within-Scatter Matrix) S_w 와 클래스간 분산을 나타

내는 행렬(Between-Scatter Matrix) S_B 의 비율이 최대가 되는 행렬을 구하여 이 행렬을 이용하여 선형 변환을 하는 기법으로 두 행렬을 식으로 정의하면 각각 식 (12) 및 식 (13)과 같다.

$$S_B = \sum_{i=1}^c n_i(m_i - m)(m_i - m)^T \quad (12)$$

$$S_W = \sum_{i=1}^c \sum_{x \in C_i} (x - m_i)(x - m_i)^T \quad (13)$$

식 (12)에서 n_i 는 i 번째 클래스 C_i 에서의 데이터 수이며, m_i 는 i 번째 클래스 C_i 내의 평균값을 나타낸다. 또한 식 (13)에서 m 은 전체 클래스의 평균값이 된다. 선형판별분석을 이용하여 구하고자 하는 최적의 투영행렬 W 는 식 (14)와 같이 투영된 샘플의 WCS행렬에 대한 BCS 행렬의 비를 최대화하는 정규직교 행들을 가진 행렬로서 선택되어진다.

$$W = \arg \max_W \frac{|W^T S_B W|}{|W^T S_W W|} = [w_1, w_2, \dots, w_m] \quad (14)$$

식 (14)에서 w_i 는 식 (15)와 같이 S_B 와 S_W 에 대한 일반화된 고유벡터의 집합이다.

$$S_B w_i = \lambda_i S_W w_i, \quad i = 1, 2, \dots, m \quad (15)$$

식 (15)에서 S_B 의 rank가 C-1개의 (C : 클래스의 수)의 영이 아닌 일반화된 고유값을 가지므로, m은 C-1개 까지 취할 수 있다. 또한 S_W 의 rank도 P-C개(P : 학습입력의 수)를 나타낸다. 일반적으로 학습 신호의 수 (P)가 신호의 길이(N)보다 적을 경우 WCS 행렬 S_W 는 항상 비정칙(singular)이 된다. 이것은 투영 행렬이 적당히 선택되지 않는다면 투영된 WCS 행렬은 영이 될 수 있음을 의미한다. 이러한 문제점은 PCA를 이용하여 입력신호를 저차원공간으로 투영함으로써 해결될 수 있다. 최종적으로 어떤 입력신호 x_i 에 대한 특징벡터 $T = (T_1, T_2, \dots, T_N)$ 는 식 (16)에 의해 얻어진다.

$$T_i = W^T z_i = W^T v_i^T (x_i - m) \quad (16)$$

표 2. LDA를 이용한 특징추출 과정

Step	Process
1	From formula (12) and (13) obtain WCS matrix S_W and BCS matrix S_B
2	Make the optimize presentation matrix W from formula (14)
3	From formula (16), obtain the feature vector T_i
4	Extract feature by LDA $T_i' = W^T z_i' = W^T v_i'^T (x_i' - m)$
5	Calculation distance between feature vector T_i' and input vector T_i'

2.3 분류결과

표 1은 PCA-LDA를 이용한 각 시기의 분류결과를 나타내고 있다. 서론에서도 언급하였듯이 트리는 말기시의 검출이 공학적으로 중요한 의미를 지닌다고 언급한바 있다. 따라서 각 시기별로 어떤 형태가 보다 더 정확한 패턴분류결과를 보이는지 확인할 필요가 있다. 따라서 본 논문에서도 초기-중기-말기, 초기-말기와 중기-말기시의

분류에 대한 결과를 나타내었다. 결과를 살펴보면, 결국엔 초기-말기 두 단계로 구별했을 경우가 100%의 인식률을 나타내고 있음을 알 수 있으며, 중기의 진전시기에는 초기의 진전시기의 데이터와 말기시의 데이터가 상당히 중복되어 나타나고 있음을 확인할 수 있었다.

표 3. 분류결과

	진전 시기별					
	초기-중기-말기		초기-말기		중기-말기	
	입력 데이터	분류된 데이터	입력 데이터	분류된 데이터	입력 데이터	분류된 데이터
초기	20	18	20	20	-	-
중기	20	20	-	-	20	16
말기	20	18	20	20	20	19
인식률	93%		100%		88%	

3. 결 론

지금까지 우리는 고체절연물내에서 발생하는 전기트리를 단계별로 구별하여 그것으로부터 수반되는 부분방전 데이터를 취득하여 각 단계를 구별하고자 하였다. 각 단계는 성장단계별로 초기, 중기, 말기로 구분을 하였으며, 각 단계별로 부분방전 데이터를 취득하였다. 결과적으로 말하자면 시기별 패턴에 대한 분류는 전체적으로 잘 이루어졌다. 하지만 서론에서도 밝혔듯이 공학적으로 의미를 갖는 것은 유지/보수 및 기기의 교체와 관련되는 말기시의 분류이므로, 본 논문에서는 각각을 보다 세밀하게 분류를 하였다. 즉, 초기-중기-말기, 초기-말기, 중기-말기의 단계로 구별하여 인식률을 나타내었다. 그 결과 초기-말기의 구별에서의 패턴 분류가 100%의 분류를 나타내었다. 반면에 중기시의 패턴이 입력으로 사용되면 초기와 말기 데이터와의 중복으로 인해 인식률이 저하되는 결과를 나타내었다. 결국 말기시의 분류에 있어서는 초기진전단계와 말기의 진전단계만 구별한다면 충분한 의미가 있다고 하겠다.

[참 고 문 헌]

- [1] F. H. Kreuger, Partial Discharge Detection in High-Voltage Equipment, Butterworth and Co. Ltd, 1989
- [2] Qian Du, Chin-I Chang, A linear constrained distance-based discriminant analysis for hyperspectral image classification, *The Journal of the pattern recognition* 34, pp. 361-373, 2001
- [3] R.O. Duda, P.E.Hart, Pattern Classification and Scene Analysis, Wiley, NewYork, 1973.
- [4] Chengjun Liu, A Shape- and Texture-Based Enhanced Fisher Classifier for Face Recognition, *IEEE Transactions on Image Processing*, Vol. 10, No. 4, pp 598-600, April 2001
- [5] Huang Weigang, Study on Conductor Configuratio of 500-kV Chang-Fang Compact Line. *IEEE Transactions on Power Delivery*, vol 18, no 3, pp.1002-1008, 2003.
- [6] Wang Liming, Liang Xidong, Guan Zhicheng, Que Weiguo. Research on 500kV Phase to Phase Composite Spacer for Compact Lines. *Proceedings of the 6th International Conference on Properties and Applications of Dielectric Materials*, Xi'an, China, pp. 346-349, 2000.