

## PCRNN을 이용한 음향신호 식별에 관한 연구

안종구 · 최재하  
전자공학과

### 〈요 약〉

파워스펙트럼은 디지털신호처리 분야에서 널리 이용되어 왔으나 해석하고자 하는 신호의 위상에 관한 정보를 잃어 버리는 단점이 있어서, 위상에 관한 정보를 필요로 하는 경우에는 파워스펙트럼 해석법은 그 이용에 제약을 받는다. 바이스펙트럼은 그 계산에 시간이 많이 걸리는 단점은 있으나 반면에 위상에 관한 정보를 잃지 않는 장점이 있다. 본 논문에서는 여러가지 자동차들의 엔진 음향을 녹음하여 바이스펙트럼을 구한 후, 부분결합 회귀신경회로망을 이용하여 바이스펙트럼 패턴을 식별하고 결과적으로 신호원을 식별하는 방법에 관하여 연구하였다.

---

## A Study on the Identification of Acoustic Signals Using PCRNN

Chong Koo An · Jae Ha Choi

### 〈Abstract〉

Power spectrum has been widely utilized in the digital signal processing area. However, there is no information about the phase of a signal, the power spectral analysis technique can not be used to interpret the phase coherency of the signal produced by some nonlinear process. In this case, the third-order higher-order spectrum, the so called bispectrum, is very useful in analyzing such signals. In this paper, the bispectra of the measured acoustic signals of vehicles are computed and then the PCRNN (partially connected recurrent neural network) is used in order to identify the acoustic signal sources using the bispectrum patterns.

## I. 서 론

파워스펙트럼은 많은 응용 분야와 장점을 가졌음에도 불구하고 비선형 프로세스(nonlinear process)에 의해서 상호 작용하는 여러 주파수 성분을 해석하고자 할 때는 그 이용에 한계가 있다. 그 이유는 파워스펙트럼은 랜덤 신호(random signals)의 크기에 관한 정보는 갖고 있으나 랜덤 신호의 위상에 관한 정보는 갖고 있지 않기 때문이다<sup>[1]</sup>. 그러나 고차 스펙트럼 중 특히 3차 스펙트럼인 바이스펙트럼(bispectrum)을 이용하면 크기에 관한 정보는 물론, 위상에 관한 정보도 얻을 수 있다.

그 주파수가  $\omega_1$ ,  $\omega_2$  그리고  $\omega_3 (= \omega_1 + \omega_2)$ 인 3개의 주파수 성분을 갖는 신호가 있다고 가정한다. 이 때 세 주파수 성분이 랜덤하게 발생했다면 각 주파수 성분의 위상은 서로 상관관계가 없다고 할 수 있다. 그러나 그 주파수가  $\omega_3$ 인 주파수 성분이  $\omega_1$ 와  $\omega_2$ 의 상호관계(interaction 혹은 coupling)에 의해 생긴 것이라면 세 주파수 성분의 위상은 서로 상관관계 [phase coherency]를 갖게 될 것이다. 위상의 상관관계가 있거나 혹은 없거나 세 개의 주파수 성분으로 이루어진 신호의 파워스펙트럼은 서로 같다. 그러나 바이스펙트럼은 크기에 관한 정보 뿐만 아니라 세 주파수 성분의 상호간의 위상에 관한 정보도 갖고 있으므로 위상 상관관계가 있는 경우의 바이스펙트럼과 위상 상관관계가 없는 경우의 바이스펙트럼은 서로 다르다. 즉 위상 상관관계가 없는 경우는 그 바이스펙트럼이 영(zero)이지만, 위상 상관관계가 있는 경우는 그 바이스펙트럼이 영이 아닌 어떤 값을 갖게 된다<sup>[2]</sup>. 따라서 바이스펙트럼을 이용하면 어느 주파수 성분의 신호가 단순한 잡음인지 아니면 어떠한 의미를 갖는 신호인지 구별할 수 있다.

어떤 신호에 부가된 잡음은 그 확률분포

가 정규분포(Gaussian)인 경우가 많다. 가우시안 분포를 갖는 랜덤프로세스의 파워스펙트럼은 짝수차(즉 2차) 스펙트럼이므로 그 값이 영(zero)이 되지 않는다. 그러나 이러한 랜덤프로세스의 바이스펙트럼은 홀수차(즉 3차) 스펙트럼이므로 그 값이 영(zero)이 된다<sup>[3]</sup>. 따라서 비(非)가우시안(nonGaussian)인 신호에 가우시안 잡음이 부가된 경우의 바이스펙트럼을 구하면 가우시안 잡음에 의한 바이스펙트럼 성분은 영(zero)이 되므로 이론적으로는 잡음에 의한 성분은 모두 사라지고 신호에 의한 바이스펙트럼만 나타난다. 그러나 파워스펙트럼에서는 잡음에 의한 성분도 나타나게 된다.

바이스펙트럼은 최근에 비선형 현상과 관련된 데이터를 분석하는 과학과 엔지니어링의 여러분야에서 적용되고 있다. C.L. Nikias와 M.R. Raghuvier<sup>[1]</sup>에 의해서 많은 응용분야가 소개되었다. 바이스펙트럼을 이용한 위상추정(phase estimation)에 관한 연구<sup>[4]</sup>가 H. Bartelt, A.W. Lohmann, B. Wirniter에 의해서 이루어졌으며, 수중 음향 신호처리 분야에서는 바이스펙트럼을 이용한 시간 지연 추정(time delay estimation)에 관한 연구<sup>[5]</sup>가 이루어졌다. Walton과 Jouny<sup>[6]</sup>가 레이더 신호를 해석하기 위해서 바이스펙트럼 추정을 응용했으며, 비가우시안(nonGaussian) 신호를 식별하기 위해서도 바이스펙트럼 추정이 이용되었다<sup>[7]</sup>. 또한 생물학(biomedical) 분야에서도 바이스펙트럼 추정에 관한 연구가 이루어졌다<sup>[8]</sup>.

본 논문의 목적은 가우시안 잡음 환경하에서 바이스펙트럼을 이용하여 음향신호를 식별하여 궁극적으로는 신호원을 식별하는데 있다. 음향신호의 바이스펙트럼을 구한 후, 이 바이스펙트럼 패턴을 식별하기 위해서 부분결합 회귀 신경회로망(PCRNN)<sup>[9]</sup>을 이용하였는데 신경회로망의 학습 알고리즘으로는 Widrow-Hoff에 의해서 제안된 Delta Rule을 다층 신경망에 적용한 오류

역전파 알고리즘을 사용하였다. 본 연구에서 제안된 방법을 차량의 음향신호를 식별하는데 적용하였다. 2절에서는 바이스펙트럼 성질에 관해서 간략히 설명하고 3절에서는 PCRNN의 구성을 보인다. 4절에서는 측정된 차량의 음향을 이용하여 각 차량의 바이스펙트럼을 구한 후 이를 식별하기 위해서 PCRNN을 이용하여 각 차량의 바이스펙트럼에 대해서 출력노드의 활성화 값(activation value)을 구하고 인식율을 구한다. 끝으로 5절에서는 본 논문에 대한 검토와 결론을 맺는다.

## II. 바이스펙트럼의 성질

평균이 0(zero mean)이며 적어도 3차까지 정상인 랜덤프로세스  $x(t)$ 의 3차 상관함수는 다음과 같이 쓸 수 있다<sup>[7]</sup>.

$$E[x(t)x^*(t - \tau_1)x^*(t - \tau_2)] = R_{xxx}(\tau_1, \tau_2) \quad (1)$$

여기서  $E$ 는 기대치이고,  $R_{xxx}$ 는 3차 상관함수 (third-order correlation function)이며 이는 시간차(time difference)  $\tau_1$ 과  $\tau_2$ 의 함수이다.

본 논문에서 이용하는 3차 상관함수의 대칭성은 식(1)로부터 쉽게 얻을 수 있다.

$$\begin{aligned} R_{xxx}(\tau_1, \tau_2) &= R_{xxx}(\tau_2, \tau_1) \\ &= R_{xxx}(-\tau_2, \tau_1 - \tau_2) \\ &= R_{xxx}(\tau_2 - \tau_1, -\tau_1) \end{aligned} \quad (2)$$

한편 시간영역에서의 3차 상관함수  $R_{xxx}(\tau_1, \tau_2)$ 와 주파수영역에서의 3차 스펙트럼(third-order spectrum), 즉 바이스펙트럼(bispectrum),  $S_{xxx}(f_1, f_2)$ 와의 관계는 다음식으로 표현된다.

$$R_{xxx}(\tau_1, \tau_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} S_{xxx}(f_1, f_2) e^{j2\pi f_1 \tau_1 + f_2 \tau_2} df_1 df_2 \quad (3)$$

식(3)으로부터 랜덤프로세스  $x(t)$ 의 푸리에 변환(Fourier transform)과 바이스펙트럼과의 관계를 구하면 다음과 같다.

$$E[X(f_0)X^*(f_1)X^*(f_2)] = S_{xxx}(f_1, f_2) \delta(f_0 - f_1 - f_2) \quad (4)$$

여기서  $X(f_0)$ ,  $X(f_1)$  그리고  $X(f_2)$ 는  $x(t)$ 를 푸리에 변환하여 주파수 영역에서 나타낸 주파수 성분을 의미하고,  $S_{xxx}(f_1, f_2)$ 는 자기바이스펙트럼(autobispectrum)이다.

연속된 신호의 자기바이스펙트럼은 그림 1에서  $(f_1, f_2)$ 의 전영역에 걸쳐서 정의된다. 그러나 이산신호의 자기바이스펙트럼은  $f_s$ 를 표본화주파수(sampling frequency)라 할 때 표본화 이론으로부터 그림 1의  $f_1 + f_2 = \pm (f_s/2)$ ,  $f_1 = \pm (f_s/2)$ ,  $f_2 = \pm (f_s/2)$ 를 둘레로 하는 육각형의 내부에서 정의되며, 1의 영역에서 자기바이스펙트럼을 구하면 나머지 11개 영역의 자기바이스펙트럼은 식(2)를 주파수영역에서 표현한 대칭성을 이용하여 구할 수 있다. 그리고 이러한 12개 영역의 대칭성 때문에 자기바이스펙트럼을 그림 1의 12개의 전영역에 대하여 구한 후 등고선그림(contour plot)으로 표현하면 대칭적인 무늬(pattern)로 나타난다. 따라서 음향식별이라는 문제를 바이스펙트럼 패턴인식이라는 문제로 바꾸어 생각할 수 있다.

## III. 부분결합 회귀 신경회로망

부분결합 회귀 신경회로망(Partially Connected Recurrent Neural Network:

PCRNN)을 본 논문의 바이스펙트럼 패턴을 인식하기 위한 인식기로 사용하였다. 부분 결합 회귀 신경회로망은 한 층의 은닉층을 가진 2층구조의 퍼셉트론을 기본으로 하며 여기에 내부상태층(internal state layer)과 결정상태층(decision state layer)을 추가하였다. 그리고 은닉층으로부터 내부상태층에 이르는 결정상태 귀환루프가 결합되어 있고 결정상태층 내에는 1보다 작은 고정된 연결강도의 자기귀환루프가 있다(그림 2).

은닉층과 출력층의 활성화패턴은 시간스텝  $t-1$ 에서 내부상태층과 결정상태층으로 각각 복사되어 시간스텝  $t$ 에서 계통의 입력상태의 일부로 작용한다. 부분결합 회귀 신경회로망은 3종류의 사이클을 발생시킨다. 첫번째 사이클은 은닉층과 내부상태층간에 발생하며, 두번째는 결정상태층을 경유하여 은닉층과 출력층간에 발생한다. 그리고 마지막으로 결정상태층 내에서 발생한다.

회로망의 결합형태는 입력층과 은닉층간, 내부상태층과 은닉층간에는 임의 부분결합형태를 이루고 결정상태층과 은닉층간, 은닉층과 출력층간에는 완전결합형태를 이룬다. 여기서 은닉층의 처리요소는 입력층과 내부상태층의 처리요소로부터 무작위로 균일하게 선택된다. 이렇게 회로망이 부분결합형태를 가짐으로써 계통의 용장도(redundancy)가 감소되어 완전결합 회로망에 비해 빠르게 안정상태로 수렴할 수 있다.

부분결합 회귀 신경회로망의 학습에는 Rumelhart, Hinton, 그리고 Williams에 의해 개발된 오류역전파(error back propagation: EBP) 학습 알고리즘이 이용되었다<sup>[10]</sup>. 오류역전파 알고리즘은 비회귀구조를 가진 다층 전향 신경회로망의 학습알고리즘으로 개발되었으나 회귀망에도 이를 적용할 수 있다. 그러나 회귀망에서 역전파 과정이 적절하게 진행되기 위해서는 회귀루프를 가지는 처리요소가 이전 시간스텝에서의 활성을 보존해야 한다. 부분결합 회귀 신경회로망에서는 은닉노드와 출력노드의

활성이 내부상태노드와 결정상태노드에 각각 보존되므로써 역전파 과정이 정상적으로 진행되게 된다. 노드의 활성화 함수로는 0에서 1사이의 출력값을 갖는 시그모이드(sigmoid) 함수를 사용했으며 학습속도를 높이기 위해서 관성항(momentum)을 사용했다.

#### IV. 실험 및 결과

음향식별 알고리즘을 적용하기 위하여 세 종류의 음향을 수집하였다. 울산대학교 앞의 6차선 도로에서 오전 10시쯤 시동을 건 채로 정지해 있는 승용차, 45인승 버스, 2.5톤 소형 트럭의 엔진 소음을 녹음하였다. 이러한 음향신호의 파워스펙트럼(power spectrum)을 구해 본 결과 3가지 음향(엔진 소음)의 파워(power)가 대부분이 2KHz 이하에 분포해 있어서, 각 음향신호를 10KHz로 표본화(sampling)하여도 앨리어싱(aliasing)을 충분히 피할 수 있다고 생각되어 10KHz로 표본화 하였다. 따라서 각 음향신호의 총 데이터 양 (즉 data length)은 10KHz로 표본화 된 20,480 (약 2초간) 포인트이다.

실제로 수집된 각 차량의 음향은 도로상에서 수집되었기 때문에 각 차량의 음향 뿐만 아니라 주변의 소음까지도 같이 포함되어 있다. 따라서 주변의 소음(여기에서는 측정하고자 하는 자동차의 엔진에서 발생하는 소리를 제외한 노상의 잡음) 효과를 줄이기 위해서 원하는 차량의 엔진으로부터 1m 이내에 마이크로폰(microphone)을 설치하였다. 본 논문에서는 수집된 음향을 각 차량의 잡음이 없는 신호원(source signal)이라고 가정하였고, 부가 잡음은 IMSL을 이용하여 가우시안 잡음(Gaussian noise)을 발생시켰다.

본 논문에서는 바이스펙트럼을 하나의 패턴(pattern)화하여 전체적인 모양을 가지적

으로 보이기 위해서 그림 1에 보이는 바와 같이 ( $f_1, f_2$ )의 전체 평면에 대하여 바이스펙트럼을 구하였다. 그런데 세 가지 음향신호의 바이스펙트럼을 전체 평면에서 구한 결과 1kHz 밖에서는 바이스펙트럼의 값이 거의 0이어서 표본화주파수( $f_s$ )는 10KHz이지만 바이스펙트럼은 그림 1에서  $-1\text{KHz} \leq f_1 \leq 1\text{KHz}$ ,  $-1\text{KHz} \leq f_2 \leq 1\text{KHz}$ 의 정사각형 영역에 대하여 구하였다. 그림 3에서는 승용차, 버스, 트럭의 잡음이 없는 음향신호에 대한 바이스펙트럼 패턴을 보였고, 그림 4에서는 SNR=0dB인 경우(즉 신호의 파워와 부가잡음의 파워가 같은 경우)의 각 차량 음향신호에 대한 바이스펙트럼 패턴을 보였다. 본 논문에서는 512-point FFT를 사용하였으며 세그먼트를 40개로 나누어 평균을 취하였다.

PCRNN의 입력 패턴은 R(ROW)×C(COLUMN)으로 이루어진 바이스펙트럼의 크기 정보이다. 바이스펙트럼의 입력 패턴은 R과 C를 각각 99로 하였다. (99로 한 이유는 본 논문에 사용된 등고선 그림을 그리기 위한 소프트웨어가 100 X 100까지의 데이터를 처리할 수 있는데 그림에서는 직류성분을 고려해야 하므로 대칭성을 살리기 위해서 99 X 99로 등고선 그림을 그렸다). 바이스펙트럼 패턴을 위한 부분결합 회귀 신경회로망의 구성은 각각 표 1과 같다.

잡음이 없는 경우의 각 차량의 바이스펙트럼 패턴을 PCRNN의 학습 패턴으로 이용하였다. 가우시안 잡음을 IMSL을 이용하여 seed 값이 서로 다른 60개의 세그먼트(1 세그먼트 당 512 데이터 포인트)를 발생시킨 후 이를 각 차량의 음향신호에 대하여 잡음의 파워비(power ratio)가 0.2가 되도록 하여 승용차, 버스, 트럭의 음향신호에 부가하였다. 즉 승용차, 버스, 트럭의 3가지 음향신호에 대해 각 20개의 가우시안 잡음이 부가된 바이스펙트럼 패턴을 구할 수 있으며 이를 신경회로망의 입력 패턴으로 이용하였다. 음향신호 대 부가잡음의 파

워비가 0.4, 0.8, 1.0인 경우에 대하여 반복하였다. 그 결과 인식 기준을 0.9로 하였을 때 표 2에서 보인 바와 같은 인식율을 얻을 수 있었다.

## V. 결 론

본 논문에서는 가우시안 잡음 환경하에서 차량 음향신호의 바이스펙트럼을 구하여 부분결합 회귀 신경회로망을 이용하여 신호원을 식별시켜 보았다. 각 차량의 잡음이 없는 경우의 바이스펙트럼을 학습 패턴으로 이용하였으며, 인식 패턴은 컴퓨터를 이용하여 발생시킨 가우시안 잡음을 수집된 차량 음향신호에 파워 비 (power ratio)를 달리하여 부가한 후에 바이스펙트럼을 구하여 이용하였다. 신경회로망을 이용하여 신호원을 식별시켜본 결과 표2와 같은 결과를 얻었다.

이론적으로는 100%의 인식률을 얻어야 하나 실험에서는 잡음의 파워가 커질수록 인식률이 낮아졌다. 이와같이 인식률이 낮아지는 것은 다음의 두가지 이유가 있다고 생각된다.

첫째, 총 데이터 길이가 너무 짧아서 (2초간) 본 논문에서는 바이스펙트럼을 구하면서 평균을 40번 밖에 하지 못했다. 바이스펙트럼은 각 세그먼트의 바이스펙트럼을 구하여 이를 전 세그먼트에 대해 평균하는 과정에서 위상의 상관관계 (phase coherency)의 유무가 나타나므로 평균의 횟수가 결과에 큰 영향을 미치게 된다. 둘째, IMSL로 발생시킨 가우시안 잡음의 확률분포가 데이터의 길이가 짧으므로 어느 정도 가우시안에 가까운지 의문시 되며, 가우시안에 가깝다 하더라도 발생시킨 잡음을 세그먼트(segment)화 했을 때 각 세그먼트가 서로 독립적 (independent) 인지 단언하기 어려운 점이 있다. 각 세그먼트의 잡음이 서로 독립적이지 아니라면 위상의 상관관계가 있을 것이므로

잡음 성분의 바이스펙트럼은 영이 되지 않을 것이다.

일반적으로 바이스펙트럼의 값을 구하는 것이 파워 스펙트럼의 값을 구하는 것보다 상대적으로 많은 시간이 걸린다. 그러나 이러한 단점은 최근에 처리 능력이 빠른 컴퓨터의 등장으로 많이 해소되었다고 할 수 있다. 본 논문에서는 386 혹은 486 PC를 이용하였기 때문에 하나의 바이스펙트럼 패턴을 구하는 데 수 분 정도의 시간이 걸렸다.

실제로는 음향신호에 대한 가우시안 잡음의 파워비가 2.0, 4.0, 8.0인 경우도 실험을 행하였다. 바이스펙트럼의 이론상 높은 인식율을 얻을 수 있을 것이라는 기대를 갖고 있었으나 얻어진 인식율은 그러하지 못하였다. 잡음이 큰 환경 하에서도 신호원을 식별해 내는 방법을 찾는 것이 앞으로의 과제라 생각한다.

### 참고문헌

- [1] C.L.Nikias and M.R.Raghuveer, "Bispectrum Estimation: A Digital Signal Processing Framework," *Proc. IEEE*, Vol.75, No.7, pp. 869-891, July 1987.
- [2] M.R.Raghuveer and C.L.Nikias, "Bispectrum estimation via AR Modeling," *Signal Processing (Special Issue on Modern Trends in Spectral Analysis)*, Vol.9, No. 1, pp35-48, Jan.1986.
- [3] J.S.Bendat and A.G.Piesol, *Random Data: Analysis and Measurement Procedures*, 2nd ed., Wiley-Interscience, New York, 1986.
- [4] H.Bartelt, A.W.Lohmann and B. Wirniter, "Phase and Amplitude Recovery from Bispectra," *Applied Optics*, Vol.23, No.18, pp.3121-3129, Sept.1984.
- [5] J.K.Tugnait, "On Time Delay Estimation with Unknown Spatially Correlated Gaussian Noise Using Fourth-Order Cumulants and Cross Cumulants," *IEEE Trans. Signal Processing*, Vol.39, No6., pp.1258-1267, June 1991.
- [6] E.K.Walton and I.Jouny, "Application of Bispectral Techniques to Radar Signature Analysis," *Workshop on Higher-Order Spectral Analysis*, Vail, Colorado, pp.56-61, June 28-30, 1989.
- [7] C.K.An, D.M.Lee and T.H.Lee, "Identification of Acoustic Signals of Vehicles Using Bispectrum," *Jour. Acoustical Society of Korea*, Vol.11, No.1, pp.5-13, Feb.1992.
- [8] E.S.Manolakos, H.M.Stellakis and D.H.Brooks, "Parallel Processing for Biomedical Signal Processing-Higher-Order Spectral Analysis," *IEEE Computer Magazine*, Vol.24, No.3, pp.33-44, March 1991.
- [9] 김성석, 순차구조 처리를 위한 부분결합 회귀 신경회로망, 공학박사 학위논문, 울산대학교, 1990년 8월.
- [10] D.E.Rumelhart, G.E.Hinton, and R.J.Williams, "Learning representations by back propagation errors," *Nature* 323, pp.533-536, Oct.1986.

표 1. 바이스펙트럼 패턴을 위한 부분결합 회귀 신경회로망의 구성  
 Table 1. Configuration of the PCRNN for bispectrum pattern

입력층	은닉층	출력층	내부상태층	결정상태층	전체연결수
9801/297	33/x	3/3	33/5	3/3	10494

※ 1. A/B: A는 각 층에 있는 처리 요소의 수를 나타내고, B는 은닉 노드에 결합된 각 층의 처리요소의 수를 나타낸다.

표 2. 차량 음향신호의 바이스펙트럼 패턴에 대한 인식률  
 Table 2. Recognizing rate of bispectrum pattern of acoustic signals of vehicles

구 분 \ 파워비	0.2	0.4	0.8	1.0
	승 용 차	100	100	100
버 스	100	100	85	85
트 렉	100	100	90	80

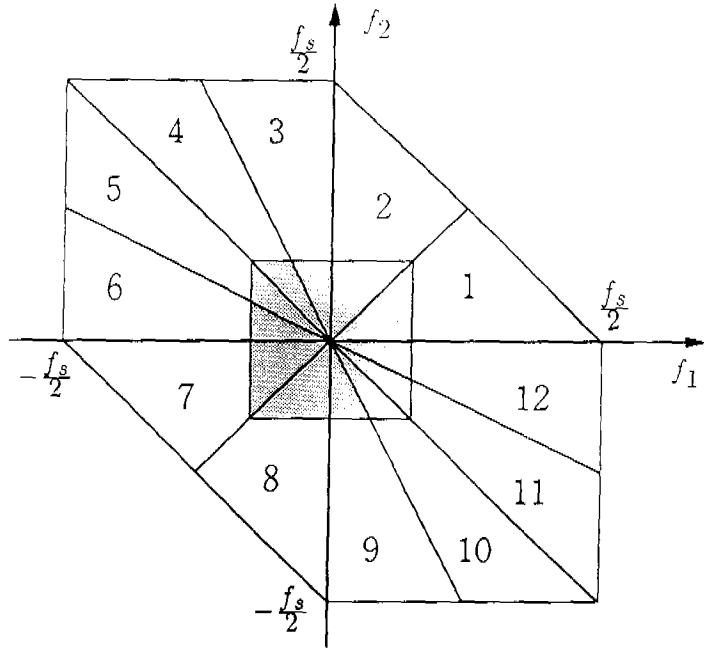


그림 1. 바이스펙트럼의 영역 및 대칭성

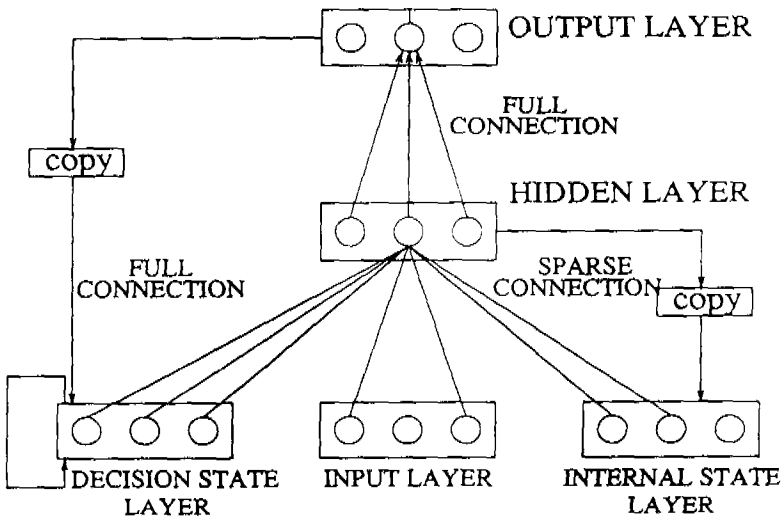


그림 2. 부분 결합 회귀 신경회로망의 구조



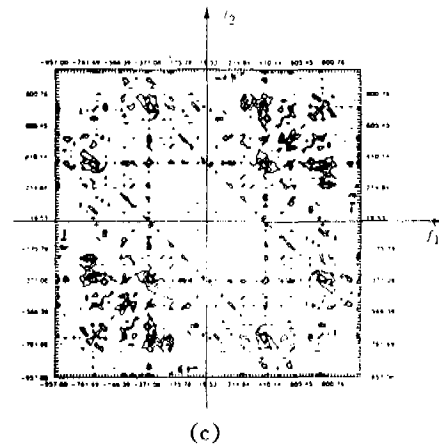
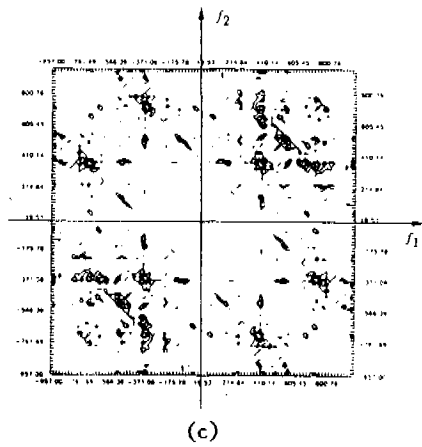
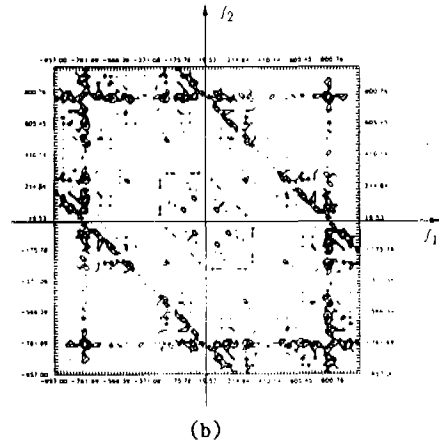
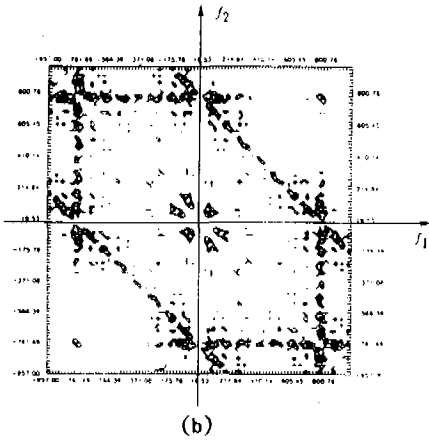
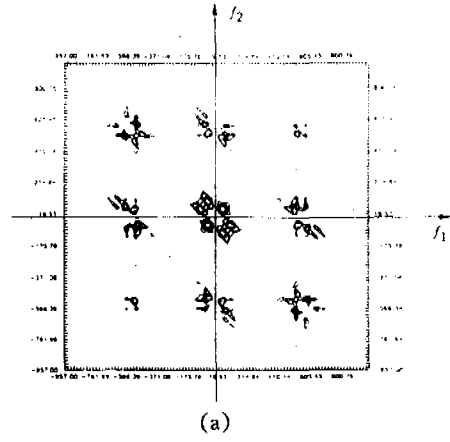
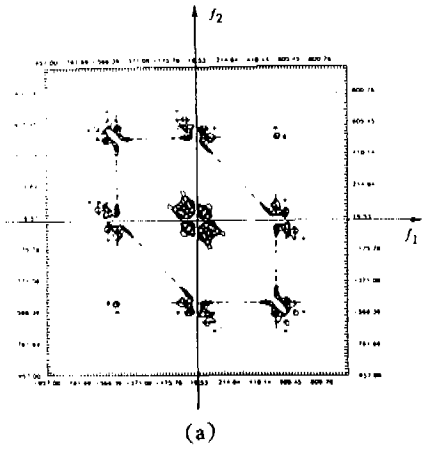


그림 3. 음향의 바이스펙트럼

Fig. 3. Bispectra of the acoustic signals no noise: (a) car, (b) bus, (c) truck

그림 4. 음향의 바이스펙트럼

Fig. 4. Bispectra of the acoustic signals SNR=0dB: (a) car, (b) bus, (c) truck