

---

# 음성 및 음성 관련 신호의 주파수 및 Quefrency 영역에서의 자기공분산 변화

김선일\*

\*거제대학

Variations of Autocovariances of Speech and its related Signals in time,  
frequency and quefrency domains

Seonil Kim\*

\*Koje College

E-mail : seonil@koje.ac.kr

## 요 약

자동차 엔진 소음과 같은 비음성신호군과 음성신호군을 구별하기 위해서는 시간영역, 주파수 영역 등에서 다양한 특징값들의 차이를 이용할 수 있는데 두 신호군을 구별하기에 적절한 명확한 차이를 가진 특징값들로서 무엇을 사용하느냐 하는 것은 중요한 관건이다. 두 신호군을 구별해내기 위해 시간, 주파수, quefrency 영역에서의 자기공분산을 제시하고 이 값들의 변화를 관찰하였다. 시간 영역에서는 단순한 공분산을, 주파수 및 quefrency 영역에서는 128개 데이터를 한 세그먼트로 하여 전체 데이터를 나눈 후 각 세그먼트에 대한 FFT 및 quefrency를 구하였다. 각 계수에 대해 세그먼트 사이의 공분산의 평균값을 구하여 각 음성신호군에 따른 공분산의 변화를 관찰하였고 주파수 영역에서 구한 공분산에서 각 신호군의 특징적인 변화를 발견할 수 있었다.

## ABSTRACT

To distinguish between a group of speech signals and nonspeech signals, you can use several features in domains like frequency, quefrency and time. It is very important to use features that differentiate two signal groups. As a feature to separate two signal groups, autocorrelation method was proposed and the variances between groups were studied. Autocovariances were just calculated for the time domain signal. Signals were divided into segments which consist of 128 data to be transformed to the frequency and quefrency domains. Autocovariances between each coefficient of segments in FFTs and quefrencies were found and they were averaged over wide spectrum. It is clear that the autocovariances in frequency domain show great differences between groups of signals.

## 키워드

공분산, FFT, quefrency, 음성신호, 비음성신호

## 1. 서 론

음성 신호에 섞여있는 잡음을 제거하는 방법은 여러 가지가 있지만 최근 모르는 둘 이상의 신호가 서로 섞여 있고 섞여있는 신호와 그 수가 같

거나 더 많은 마이크를 통해 입력된 신호를 분리해 내는 방법으로 ICA(Independent Component Analysis)[1]가 많이 이용된다. 이 방법은 다수의 신호를 분리해 낼 수는 있지만 분리된 신호가 어떤 신호인지 알려주지는 않는다. 본 연구에서는

자동차 주행 중 자동차 배기음 및 엔진 소음과 음성신호가 섞인 신호를 ICA로 분리해 낸 후 분리된 각 신호가 음성신호인지 아니면 자동차 배기음 및 엔진 소음과 같은 비음성신호인지 구별하기에 충분한 가능성이 있는 특징값을 찾아보고자 한다.

### II. 공분산

다음과 같은 공분산[2]을 이용하면 각 신호의 특성에 따른 변화를 알아낼 수 있다. 신호  $x(t)$ 에 대한 공분산  $c_x(t, \tau)$ 는

$$c_x(t-\tau) = E[x(t) - m_x(t)][x(t-\tau) - m_x(t-\tau)] \quad (1)$$

여기서  $m_x(t)$ 는  $x_t$ 의 평균값이고  $x(t-\tau)$ 는  $x_t$ 에서  $\tau$ 만큼 이동한 값이다.

### III. 시간 영역에서의 공분산

자동차 배기음 신호 2가지에 대해 시간 영역에서  $\tau = \pm 10$ 일 때  $x(t)$ 의 공분산  $c_x(t, \tau)$ 을 구하면 그림 1과 같다. 2개의 음성 신호에 대한 공분산은 그림 2에 나타나 있다.

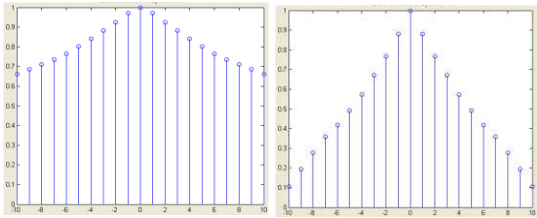


그림 1. 자동차 배기음의 시간 영역 공분산

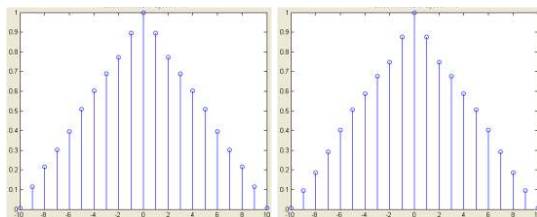


그림 2. 음성 신호의 시간 영역 공분산

자동차 배기음이나 음성 신호에서나 시간 영역에서의 공분산은  $\tau$ 가 증가함에 따라 선형적으로 감소하고 있다. 비록 기울기의 차이는 있지만  $\tau$ 의 증가에 따른 선형적인 감소는 동일하게 나타나고 있다. 따라서 공분산의 변화로 음성 신호인지 자동차 배기음인지 구별하기 어렵다. 시간 영역에서는 자동차 배기음 신호나 음성 신호나 다

같이 이웃하는 신호는 공분산 값이 크고 시간적으로 멀리 떨어진 신호는 공분산 값이 줄어들게 됨을 그림 1 및 그림 2를 통해 확인할 수 있다.

### IV. 주파수 영역에서의 공분산

자동차 배기음이나 음성 신호를 128개의 데이터를 갖는 세그먼트(segment)로 나눈 후 각 세그먼트를 FFT 변환한다. 이 중에서 첫 번째 값은 DC 값이므로 제외하고 음성 신호는 저주파 영역에 성도에 관한 정보를 많이 갖고 있으므로 저주파 영역의 10개 주파수에 대해 공분산을 구한 다음 이를 평균하였다. 각 세그먼트의 FFT 값을 열벡터로 나타내면 한 신호 전체에 대한 값은 식(2)와 같은 행렬로 나타낼 수 있다.

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,128} \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots & a_{2,128} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \dots & a_{n,128} \end{bmatrix} \quad (2)$$

식 (2)에서 두 번째 행벡터의 각 계수 사이의 공분산을 식(3)과 같이 행벡터로 나타내면

$$c_1 = \begin{bmatrix} c_1(-\tau) \\ c_1(-\tau+1) \\ \vdots \\ c_1(-1) \\ c_1(0) \\ c_1(1) \\ \vdots \\ c_1(\tau-1) \\ c_1(\tau) \end{bmatrix} \quad (3)$$

세 번째 및 i 번째 행벡터의 공분산은 식(4)와 같은 행벡터로 나타낼 수 있고

$$c_2 = \begin{bmatrix} c_2(-\tau) \\ c_2(-\tau+1) \\ \vdots \\ c_2(-1) \\ c_2(0) \\ c_2(1) \\ \vdots \\ c_2(\tau-1) \\ c_2(\tau) \end{bmatrix} \quad c_i = \begin{bmatrix} c_i(-\tau) \\ c_i(-\tau+1) \\ \vdots \\ c_i(-1) \\ c_i(0) \\ c_i(1) \\ \vdots \\ c_i(\tau-1) \\ c_i(\tau) \end{bmatrix} \quad (4)$$

각 공분산 행벡터 내의 값들을 식(5)와 같이 평균하여  $\tau = \pm 10$  일 때 그래프를 그리면 그림 3

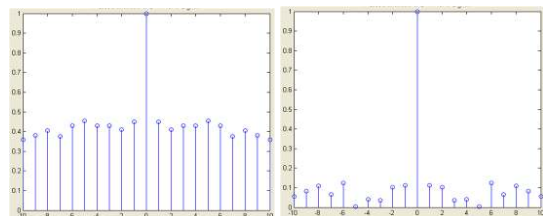


그림 3. 자동차 배기음의 주파수 영역 공분산

및 그림 4와 같은 결과를 얻게 된다.

$$c = \frac{1}{p} \begin{bmatrix} \sum_{i=1}^p c_i(-\tau) \\ \sum_{i=1}^p c_i(-\tau+1) \\ \vdots \\ \sum_{i=1}^p c_i(-1) \\ \sum_{i=1}^p c_i(0) \\ \sum_{i=1}^p c_i(1) \\ \vdots \\ \sum_{i=1}^p c_i(\tau-1) \\ \sum_{i=1}^p c_i(\tau) \end{bmatrix} \quad (5)$$

이 그래프를 살펴보면 음성 신호에 대해서는  $\tau$ 의 증가에 따른 공분산의 변화가 대략 선형적으로 감소하지만 자동차 배기음 신호에 대해서는  $\tau$ 의 변화와 무관하며  $\tau$ 가 0인 경우와 아닌 경우에 상당한 차이를 나타내고 있다.

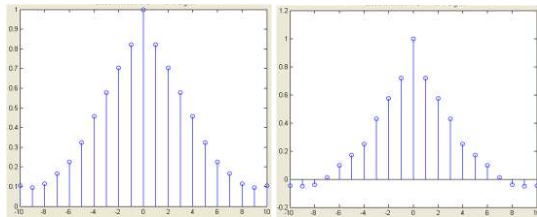


그림 4. 음성 신호의 주파수 영역 공분산

그림 3과 그림 4의 그래프들은 각 블록의 FFT 값 사이의 공분산을 보여주고 있다. 각 블록의 FFT 값은 일종의 SFFT(Shorttime FFT)에 해당되는데 음성 신호의 주파수 변화는 특히 저주파 영역에서는 성도의 변화와 거의 일치한다고 볼 수 있다. 성도는 짧은 구간에서 급격히 변화하지 않으므로 그림 4와 같이 공분산의 변화가 거의 선형적으로 나타나리라고 기대할 수 있다. 즉 시간적으로 가까운 영역에서는 공분산의 값이 크고 멀어질수록 값이 작아지게 되고 그림 4에서 이를 확인할 수 있다. 하지만 자동차 배기음의 경우에는 소리의 발생 원인이 달라서 성도와 같은 음성 기관에서 나오는 소리가 아니므로 각 블록 사이의 주파수 변화는 서로 독립적이라고 볼 수 있다. 따라서 그림 3과 같이 블록 사이의 시간적 거리에 따른 공분산의 변화를 기대하기 어렵다.

### V. Quefreny 영역에서의 공분산

FFT로 얻은 계수의 절대값의 로그를 IFFT 하면 Cepstrum[3]이 계산되는데 이 때 계산된 계수가 quefreny에 해당된다. 이 계수에 대해서 FFT에서와 같은 방식으로 공분산을 계산하여 그래프로 그리면 그림 5 및 그림 6과 같은 결과를 얻을 수 있다. 신호를 블록으로 나누어서 FFT, log, IFFT를 거친 후 각 블록 사이의 공분산을 구한 것이므로 대체로 주파수 영역에서 구한 것과 유사한 형태를 보인다. 하지만 대체적으로 낮은 값을 나타내고 있다.

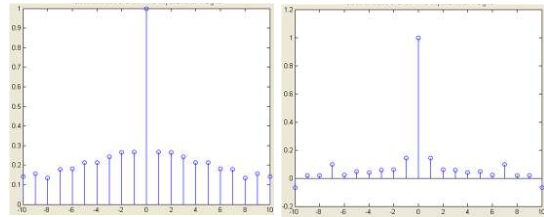


그림 5. 자동차 배기음의 Quefreny 영역 공분산

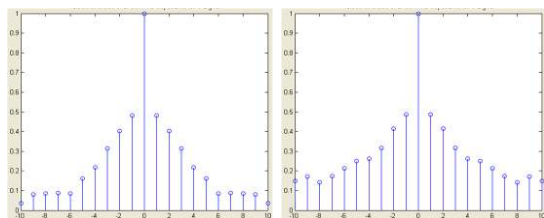


그림 6. 음성 신호의 Quefreny 영역 공분산

### VI. 결 론

시간 영역, 주파수 영역, quefreny 영역에서 음성 신호 및 자동차 배기음의 공분산 변화를 관찰하였다. 자동차 배기음은 일상 생활에서 자동차를 많이 이용하는 현대인에게 익숙한 소음이고 텔레메틱스 등 사람의 음성이 중요한 역할을 하는 시스템에서 사람의 음성과 배경 소음을 분리해 내는 것은 상당히 중요한 기술 중의 하나이다. 설사 분리해 내었다라도 어느 것이 사람의 음성이고 어느것이 아닌지 구분해 내는 것도 매우 중요하다. 구분해 내야 할 음성 신호의 음원을 이용해서 이 음원과 대상 신호와의 상관관계를 이용해[4] 좋은 결과를 얻기도 하였다. 하지만 이 경우에 음성 신호의 음원을 얼마나 적절히 확보하느냐에 따라 다른 결과를 얻을 수 있다.

본 연구에서는 시간 영역에서의 공분산이 두 형태의 신호를 구분하는데 도움이 되지 못하지만 주파수 영역에서 각 블록 사이의 공분산은 의미 있는 구분을 할 수 있는 근거를 제공하고 있음을 제시하고 있다. 물론 quefreny에서도 어느 정도 구분이 가능하지만 더 많은 계산을 해야함에도

불구하고 주파수 영역에서 보여주는 것과 같은 명확성을 보여주지 못하고 있다. 따라서 주파수 영역에서의 블록 사이의 공분산을 이용하면 ICA와 같은 기술을 사용하여 분리된 신호들 중에서 음성 신호를 가려내는데 상당히 유용하게 사용될 수 있을 것이며 이를 이용한 후속 연구를 기대할 수 있다.

#### 참고문헌

- [1] A. Hyvarinen, "Survey on Independent Component Analysis," Neural Computing Surveys 2, pp. 94-128, 1999
- [2] Sophocles J. Orfanidis, Introduction to Signal Processing, Prentice Hall Inc., 1996
- [3] A. V. Oppenheim and R. W. Schaffer, "From Frequency to Quefrency: A History of the Cepstrum," IEEE Signal Processing Magazine, pp. 95,-106, Sep., 2004
- [4] 김선일, "ICA로 분리한 신호의 분류," 대한 전자공학회논문지, 제 47권 IE편, 제 4호, 2010. 12