

인지에 중요한 음향신호의 위상에 대해

On the Perceptually Important Phase Information in Acoustic Signal

김도석*
(Doh-Suk Kim*)

*삼성종합기술원

(접수일자: 1999년 11월 29일; 채택일자: 2000년 9월 6일)

음성 신호의 효과적인 양자화를 위해서 인간의 청각 인지 특성을 이용하는 방법이 널리 사용되고 있다. 그러나 아직까지는 음성 스펙트럼의 크기 정보에 대한 인지 특성만을 이용하고 있을 뿐, 음성의 위상 정보는 그 중요성에 비해 간과되어 온 것이 사실이다. 본 논문에서는, 음향 신호로부터 청각 인지 측면에서 관련 없는 위상 성분을 추정하는 PIPE (Perceptually Irrelevant Phase Elimination) 방법을 제안하였다. 제안된 방법은 청각 특성이 critical band 대역 내의 상대적인 위상 관계에 민감하다는 관측 결과에 근거하여, 일반적인 stationary 푸리에(Fourier) 신호와 고조파(harmonic) 신호에 대해 각각 인지 측면에서 관련 없는 위상을 판별하는 방법을 제공한다. 제안된 방법을 음성 분석-합성(analysis / synthesis) 시스템에 적용함으로써 일부 위상 정보만을 이용해 음성을 복원하였으며, 청취 실험 결과 제안된 방법의 효율성을 알 수 있었다.

핵심용어: 음향신호, 음성, 위상, 인지, 임계 위상 주파수

투고분야: 음악음향 및 음향심리 분야 (8.5), 음성처리 분야 (2.2, 2.6)

For efficient quantization of speech representation, it is common to incorporate perceptual characteristics of human hearing. However, the focus has been confined only to the magnitude information of speech, and little attention has been paid to phase information. This paper presents a novel approach, termed perceptually irrelevant phase elimination (PIPE), to find out irrelevant phase information of acoustic signals in terms of perception. The proposed method, which is based on the observation that the relative phase relationship within a critical band is perceptually important, is derived not only for stationary Fourier signal but also for harmonic signal. The proposed method is incorporated into the analysis/synthesis system based on harmonic representation of speech, and subjective test results demonstrate the effectiveness of proposed method.

Key words: Acoustic signal, Speech, Phase, Perception, Critical phase frequency

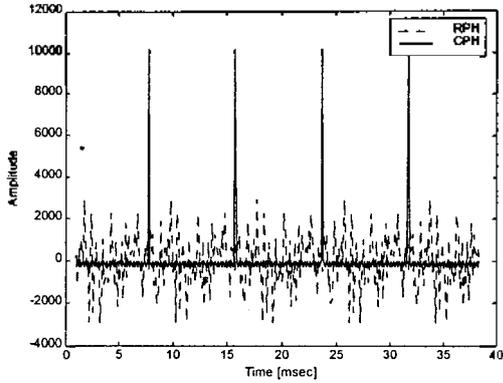
Subject classification: Musical acoustics and psychoacoustics (8.5), Speech signal processing (2.2, 2.6)

I. 서 론

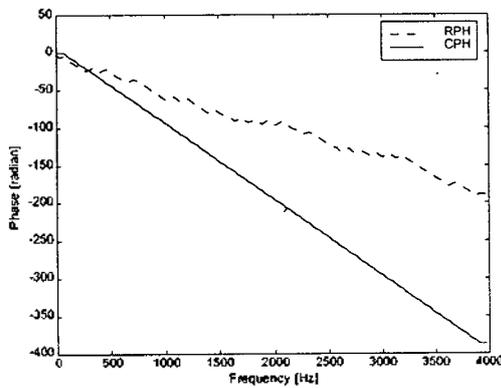
음성이나 음악 신호의 부호화에 있어 인간의 청각 인지적 특성을 이용하는 방법은 큰 성과를 거두고 있다. 이는 입력 신호에 따라 청각 시스템 내부에 특성이 다른 잡음 신호가 발생하며, 부호화 시스템에서 필연적으로 발생하는 양자화 잡음의 크기를 청각 내부 잡음보다 작도록 함으로써 양자화 잡음이 인지에 미치는 영향을 최소화시킨다는 기본적인 원리에 근거한다[1][2]. 그러나, 현재까지 부호화 방식에 적용되어 온 인지 특성은 신호의 크기 정보(magnitude information)에만 국한되어 왔으며, 신호의 위상 정보는 그 중요성에 비해 상대적으로 간과되어 왔다.

실제로 인간의 청각 시스템은 신호의 위상 변화를 감지할 수 있다는 많은 실험 결과들이 발표되어 왔다 [3][4][5]. 예를 들어 크기 그림 1은 기본 주파수가 125Hz이며, 크기 스펙트럼은 같지만 위상 스펙트럼이 다른 두 고조파 신호를 나타낸 것이다. 그림에서 CPH는 모든 고조파의 시작 위상이 0인 cosine phase signal이며, RPH는 시작 위상이 랜덤한 값을 갖는 random phase signal이다. 이 두 신호는 크기 스펙트럼은 같지만, 서로 다른 위상 스펙트럼으로 인해 명백히 다른 음색 (timbre)을 지닌다. 간단한 인위적인 신호뿐 아니라 실제의 음성 부호화 시스템에서도 위상 스펙트럼에 따라 인지되는 음질에 차이가 있다는 사실도 알려져 있다[6]. 따라서 음성이나 음악의 효율적인 양자화를 위해서는 신호의 크기 스펙트럼에서와 마찬가지로, 인지에 중요하지 않은 위상 정보를 밝혀내고 이를 정량화 하는 일이 매우 중요하며, 최근 이 중요한 논제를 다루는 연구들이 주목 받고 있다[7][8]. 하지만 불행하게도 아직 위상 인지에 대해 완전한 이론이 없으며,

따라서 효과적인 위상 정보의 양자화와 전송이라는 문제는 아직도 숙제로 남아 있다.



(a) waveform in time domain



(b) phase spectrum

그림 1. 크기 스펙트럼은 같고 위상 스펙트럼이 다른 두 신호 CPH(cosine phase signal)와 RPH(random phase signal)
Fig. 1. Two signals with same magnitude spectrum but with different phase spectrum(CPH and RPH).

본 논문에서는 이 문제 해결을 위한 첫 시도로서, stationary 신호에 대해 인지와 무관한 음향신호의 위상을 찾아내는 PIPE (Perceptually Irrelevant Phase Elimination) 알고리즘을 제안한다. 제안된 방법을 심리음향학적 실험을 통해 검증하고, 음성 신호의 분석/합성(analysis/synthesis) 시스템에 이를 적용함으로써 향후 음성 부호화 시스템에의 적용 가능성을 제시하였다.

II. 인지에 무관한 위상 정보

2.1. 채널 내 위상 변화와 채널간 위상 변화

청각 시스템에서 음향 신호는 각 중심 주파수가 서로 다른 청각 필터군에 의해 먼저 분석된다. 각 청각 필터의 대역폭은 “임계 대역(Critical Bandwidth)”으로 결정되며 저주파에서는 좁은 대역폭을, 고주파에서는 넓은 대역폭을 지니고, 또한 이 필터들은 서로 중첩도가 매우 높다는 사실이 잘 알려져 있다. 여기서 중요한 사실은 신호가 필

터군에 의해 필터링되며 각 필터 출력의 포락선이 위상 인지와 관련된 중요한 요인이라는 점이다. 즉, 위상 스펙트럼의 변화는 각 필터 출력의 포락선 변화로 반영되며, 이로 인해 음색 변이가 발생한다고 알려져 있다. 이런 맥락에서 임계 대역내 신호 성분의 상대적인 관계는 매우 중요하며, 이 대역내 위상 관계를 변화 시키는 위상 변이를 “국소 위상 변화 (채널 내 위상 변화)” 라고 한다. 마찬가지로, 채널 내의 위상 관계는 유지시키면서 채널간 위상 관계를 변화 시키는 위상 변화를 “전역 위상 변화 (채널간 위상 변화)”라고 한다.

심리 음향학 분야에서는 채널 내 위상 변화는 인간의 인지 특성이 민감한 반면, 채널간 위상 변화는 느끼지 못한다고 알려져 있다[9]. 따라서, 임계 대역내 신호의 상대적 위상 관계를 유지하는 위상 변이는 임계 대역 신호의 포락선을 유지시키게 되고, 원래의 신호와 동일한 음색을 유지한다고 볼 수 있다.

이 기본 가정에 근거하여, 인지와 무관한 위상 정보를 결정하는 판별법을 푸리에 신호와 고조파 신호와 같은 stationary 신호에 대해 전개하겠다.

2.2. Fourier 신호

식 (1)과 같은 신호 $s(t)$ 에서 $f_{k-1} < f_k < f_{k+1}$ 이고, $A_k \neq 0$ 이라고 하자.

$$s(t) = \sum_{k=1}^K A_k \cos(2\pi f_k t + \theta_k) \quad (1)$$

이제 그림 2에 나타낸 것과 같이, $C_L(f_k)$ 와 $C_H(f_k)$ 는 f_k 를 각각 상한과 하한으로 하는 임계 대역의 주파수 집합이라고 하겠다. 그러면 $C_L(f_k)$ 와 $C_H(f_k)$ 는

$$\begin{aligned} C_L(f_k) &= \{f \mid f_{k, LB} \leq f \leq f_k\} \\ C_H(f_k) &= \{f \mid f_k \leq f \leq f_{k, UB}\} \end{aligned} \quad (2)$$

와 같이 표현할 수 있다. 여기서 $f_{k, LB}$ 는 f_k 가 임계 대역의 상한 주파수일 때 그 대역의 하한 주파수를, $f_{k, UB}$ 는 f_k 가 임계 대역의 하한 주파수일 때 그 대역의 상한 주파수를 나타낸다.

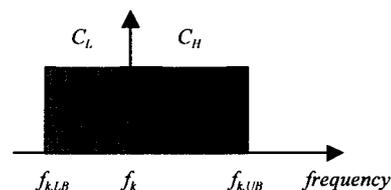


그림 2. f_k 를 상한과 하한으로 하는 두 임계 대역의 개념도
Fig. 2. Illustration of two critical bands with f_k as the upper and lower bound.

만약 청각 필터의 주파수 크기 응답 특성이 이상적이라 가정하면, 임계 대역폭은 ERB(Equivalent Rectangular Bandwidth)에 의해 표현되며, 주파수와 ERB 간의 변환식은

$$ERB = f/Q_{ear} + B_{min} \quad (3)$$

와 같이 근사화 할 수 있다[10]. 여기서 f 는 주파수[Hz], Q_{ear} 는 고주파에서의 청각 필터의 Q 를, B_{min} 은 저주파 채널의 최소 대역폭을 나타낸다. 이 관계식들로부터 $f_{k, LB}$ 와 $f_{k, UB}$ 는

$$f_{k, LB} = \frac{1}{2Q_{ear} + 1} [f_k(2Q_{ear} - 1) - 2Q_{ear}B_{min}]$$

$$f_{k, HB} = \frac{1}{2Q_{ear} - 1} [f_k(2Q_{ear} + 1) + 2Q_{ear}B_{min}] \quad (4)$$

와 같이 구할 수 있다.

이제 식 (5)를 만족하는 주파수 성분 f_k 가 존재한다고 가정하겠다.

$$f_{k-1} \in C_L(f_k) \text{ and } f_{k+1} \in C_H(f_k) \quad (5)$$

이 경우, f_k 를 원소를 하는 모든 임계 대역에는 f_k 이 외에 다른 신호 성분이 없으므로 f_k 의 위상 θ_k 를 임의의 값으로 대치하더라도 임계 대역 내의 상대적인 위상 관계는 변하지 않고 오직 임계 대역간의 위상 관계만 변하게 된다. 따라서 조건 식 (5)를 만족하는 주파수 성분의 위상은 인지 특성상 중요하지 않다고 할 수 있다.

2.3. 고조파 신호

고조파 신호의 경우 식 (1)은 식 (6)과 같이 기본 주파수와 그 정수배 주파수들로 구성된다.

$$S(t) = \sum_{k=1}^K A_k \cos(2\pi k f_0 t + \theta_k) \quad (6)$$

여기서 식 (7)을 만족하는 한 고조파 주파수 $f_x = x f_0$ 가 존재한다고 가정하자.

$$(x-1)f_0 \in C_L(xf_0) \text{ and } (x+1)f_0 \in C_H(xf_0) \quad (7)$$

즉, 식 (7)을 만족하는 x 에 대해서 $x f_0$ 와 $\alpha(x+1)f_0$ 는 같은 임계 대역 내에 존재하게 되지만, $K f_0$ 와 $(x-1)f_0$ 는 동일한 임계 대역 내에 존재하지 않는다. 각 고조파 주파수들은 주파수축 상에 균일하게 분포하며 임계 대역은 주파수가 증가함에 따라 그 폭이 넓어지게 되므로, $k \geq x$ 에 대해서 적어도 두 개의 인접한 주파수 성분 $k f_0$ 와 $(k+1)f_0$ 는 동일 임계 대역 내에 존재하게 된다. 반면, $k < x$ 일 경우에 어떠한 두 인접 주파수 성분

$k f_0$ 와 $(k+1)f_0$ 도 같은 임계 대역 내에 존재할 수 없게 된다. 여기서 $f_x = x f_0$ 를 “임계 위상 주파수(critical phase frequency)”라고 정의 하겠다.

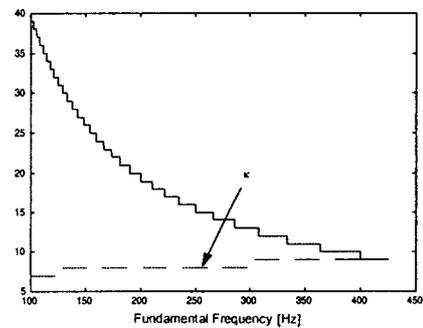
고조파 신호에 대해 이상의 특성을 고려하면, 임계 위상 주파수보다 높은 주파수의 위상 변화는 임계 대역 신호내의 상대적인 위상 관계를 변화 시킬 수 있으므로 국소 위상 변화에 해당하며, 이 결과는 신호의 음색 변이로 나타난다는 사실을 알 수 있다. 반면에 임계 위상 주파수보다 낮은 주파수 성분에 대해서는 두개 이상의 주파수 성분이 한 임계 대역내에 존재할 수 없으므로 위상을 변화 시키더라도 임계 대역간의 위상관계만 변화 시킬 뿐, 임계 대역내의 위상 관계는 변화 시키지 못하게 된다.

ERB 근사식과 고조파 신호를 구성하는 성분간의 주파수 간격을 고려하면 x 는 식 (8)과 같이 구해진다.

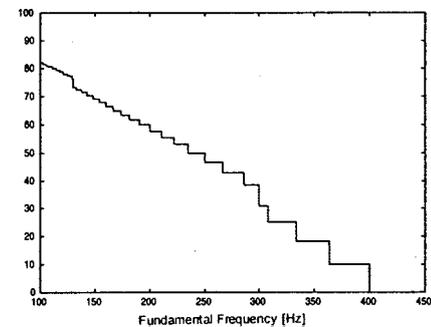
$$x = \lceil Q_{ear} \left(1 - \frac{B_{min}}{f_0} \right) - 0.5 \rceil \quad (8)$$

또한 고조파 신호의 기본주파수 f_0 가 식 (9)를 만족하면 f_x 는 $[f_0, \infty]$ 범위에서 존재한다는 사실을 쉽게 유도할 수 있다.

$$f_0 > Q_{ear} B_{min} / (Q_{ear} - 1.5) \quad (9)$$



(a) Number of perceptually important phase components



(b) Proportion of perceptually important phase components

그림 3. Nyquist 주파수가 4kHz 일 때 기본 주파수와 인지상 중요한 위상과의 관계
Fig. 3. Relation between perceptually important phase and fundamental frequency provided that the Nyquist frequency is 4kHz.

참고로, Glasberg 와 Moore의 제안에 따라 Q_{ear} 과 B_{min} 에 9.26449 와 24.7을 대입하면 f_0 가 100Hz 일 때 x 는 7이 된다. 따라서 $f_x = x/f_0$ 이상인 700Hz이상의 고조파 성분은 이웃하는 고조파 성분과 동일한 임계 대역 내에 존재하며, 서로 위상 변화에 영향을 받을 수 있다. 반면, 600Hz 이하의 고조파 성분들은 서로 동일한 임계 대역 내에 존재할 수 없으므로 위상 변화를 인지하지 못하게 된다.

식 (8)로부터, 기본 주파수가 증가하게 되면 임계 위상 주파수도 따라서 증가하게 됨을 알 수 있다. 이 특성을 그림 3에 나타내었다. 그림 3(a)에서 실선은 Nyquist 주파수가 4KHz 라고 가정했을 때 총 위상 성분 수이며, 점선은 x 를 나타낸다. 인지 특성상 중요한 위상 성분에 해당하는 것은 실선과 점선 사이의 정보이며, 점선보다 아래에 존재하는 위상 정보는 중요하지 않다. 그림 3(b)는 전체 고조파 성분수에 대해 인지 특성상 중요한 위상의 비율이 기본 주파수에 따라 어떤 특성을 지니는지를 나타낸 것이다. 기본 주파수가 높을수록 중요한 위상 정보의 비율은 점점 낮아지는 것을 알 수 있다.

III. 실험 결과

3.1. 임계 위상 주파수

제안된 이론의 타당성을 검증하기 위해 길이 512msec인 톤에 대해 청취 실험을 수행하였다. 실험에 사용된 고조파 신호는 크기 스펙트럼이 전 주파수 대역에 대해 같은 값을 가지며, 모든 고조파의 시작 위상이 0인 CPH(cosine phase)신호이다 (A). 각 고조파에 대한 위상 JND(Just-noticeable Difference)를 측정하기 위해 CPH 신호에서 하나의 주파수의 위상 값을 변화 시킨 신호를 또한 발생시킨다 (B). 이제 청취자는 2AFC(Two-Alternative Forced Choice) 실험 방식에 의해, AKG 250 헤드폰의 한쪽으론 세 신호 ABX를 듣고 (1) X가 A와 같은지 또는 (2) X가 B와 같은지 둘 중에서 판단하게 된다. 여기서 X는 A와 B 중에서 랜덤하게 고른 것이며, A와 B의 순서도 랜덤하게 변화 시킨다. 실험은 UDTR (up-down transform) 방법[11] 에 의해 변환된 심리 응답 곡선 상에서 79.4% 정답점을 찾도록 설계되었다. 그림 4는 한 고조파 성분의 위상 JND를 측정하는 UDTR 방법의 과정을 보인 것이다

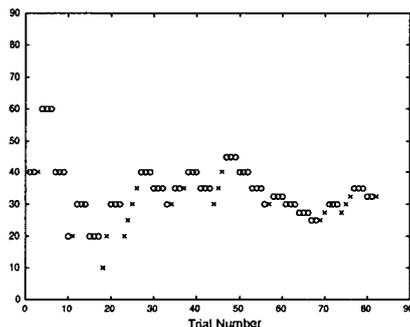


그림 4. 위상 JND를 측정하기 위한 UDTR 실험의 한 예
Fig. 4. An example of UDTR method to measure JND of phase.

원은 응답자가 맞는 답을 선택한 것을, X는 틀린 답을 선택한 것을 나타낸 것이다. 그림에서 알 수 있듯이, 응답자의 반응에 따라 adaptive하게 위상 변이값을 조절함으로써 JND 값으로 수렴하도록 되어 있다. 이와 같이 모든 고조파 성분에 대해 JND를 측정하였으며, 그 결과 낮은 고조파 성분에 대해서는 여러 번의 시도에도 불구하고 수렴하는 JND값을 측정할 수 없었다 -- 임의의 위상 값을 부여하더라도 청취자는 그 차이를 전혀 느낄 수 없었다. 임계 위상 주파수는 위상 JND를 갖는 고조파 중 가장 낮은 값으로 추정하였으며, 한 청취자 (정상청력을 지닌 20대 여성) 에 대한 결과를 표 1에 나타내었다.

표 1. 임계 위상 주파수
Table 1. Critical phase frequency.

FO	100 Hz	200 Hz	350 Hz
k , theoretic	7	8	9
k , experimental	6	7	7

실험 결과 이론적으로 유도된 결과와 실험 결과가 매우 유사함을 알 수 있다. 그러나 정확히 일치하지는 않으며, 그 이유 중 하나는 제안된 위상 판별 방법이 청각 필터의 주파수 응답 특성이 이상적이라는 가정하에 전개된 것이라는 점에 있다. 즉, 청각 필터의 실제 특성은 이상적이 아니며 $C_L(f)$ 이나 $C_H(f)$ 에 속하지 않는 주파수 성분이 f_0 를 포함하는 임계 대역 내의 위상 관계에 어느 정도 영향을 미칠 수 있다(그림 2 참조). 그럼에도 불구하고, 그 효과는 $C_L(f)$ 이나 $C_H(f)$ 에 속하는 주파수 성분에 비해 무시될 수 있다.

3.2. 분석/합성 음질 테스트

제안된 방법이 음성 부호화에 효과적으로 적용될 수 있다는 것을 보이기 위해 임계 위상 주파수 개념을 SMX(Spectrally Mixed Excitation) 부호화기[12]에 근간을 둔 음성 분석/합성 시스템에 적용시켰다. 여기서 음성 신호는 여기 신호 $e(n)$ 을 입력으로 하는 all pole 모델 $h(n)$ 으로 표현되며, 부호화기에서 사용되는 양자화는 수행하지 않았다. All pole 모델은 20ms마다 얻어지는 10차 LPC 계수로 표현되며, 여기 신호는 10msec마다 얻어지는 고조파 신호로 모델링 되었다. 여기 신호는 추정된 기본 주파수와 각 고조파 위치에서 스펙트럼의 크기와 위상에 의해 결정된다. 추정된 기본 주파수로부터 음성의 매 프레임마다 임계 위상 주파수를 식 (8)에 의해 구하게 되며, 임계 위상 주파수 이상의 위상만 원래 위상 정보를 이용하고 그 이하의 위상은 과거 프레임의 위상, 현재와 과거 프레임의 기본 주파수를 이용해 이차 보간법으로 진행시킨다[13].

임계 위상 주파수 개념이 도입된 음성 분석/합성 시스템에 대한 평가를 위해 주관적 AB 음질 선호도 평가 실험을 수행하였다. 실험에 사용된 음성은 8명의 화자(남자4, 여자4)가 발성한 16 문장이며, 5명이 두 시스템 - 모든

위상 정보를 사용한 시스템(all phase)과 임계 위상 주파수 이상의 주파수 위상만 사용한 시스템(PIPE) - 에 의해 처리된 같은 음성을 듣고 선호도를 평가한 결과를 표 2에 나타내었다. 성능 비교를 위해 기존의 sinusoidal 부호화기에서 많이 사용되는 zero phase의 결과도 함께 나타내었다. Zero phase는 원래 신호의 위상 정보 없이 모든 위상이 cosine phase (zero phase)라고 가정하고 위상을 진행시키는 방법이다. 실험 결과 PIPE에 의한 음질이 모든 위상 정보를 사용한 all phase의 음질에 비해 약간 저하됨을 알 수 있으며, 이는 제안된 이론에 부합하지 않는 결과이다. (제안된 이론에 의하면 음질 차이가 존재하지 않아야 한다.) 이 불일치는 제안된 이론이 stationary 신호에 대해 전개 되었으며, 음성 신호는 시변 신호(time-varying signal)라는 점에 기인한다. 즉, 음성 신호의 인접한 두 프레임의 기본 주파수가 다를 경우 프레임 경계 부분에서 위상 변이가 발생할 수 있으며, 이는 이 논문에서 제안하고 있는 이론의 범위를 넘어서는 것이다. 그럼에도 불구하고 all phase와 zero phase의 비교 실험(표 2-b)에 비해 PIPE의 경우 "no preference" 비율이 매우 높다는 사실은 제안된 방법의 효과와 우수성을 입증하는 것이다. 더욱이, 실험에 사용된 음성에서 남자의 경우 원래 위상 정보량의 27.3%, 여성 음성의 경우 48.7%가 사용되지 않은 상태에서 이와 같은 결과를 얻을 수 있었다는 것은 매우 고무적이라 할 수 있다.

표 2. 음질 평가 결과(%)
Table 2. Preference test results(%).

	Male speech	Female speech	Total
All Phase	25.0	18.7	21.9
PIPE	10.0	16.3	13.1
No preference	65.0	65.0	65.0

(a) all phase vs. PIPE

	Male speech	Female speech	Total
All Phase	66.3	46.3	56.3
Zero phase	12.5	20.0	16.2
No preference	21.2	33.7	27.5

(b) all phase vs. zero phase

IV. 결 론

본 논문에서는 청각 필터의 임계 대역 내의 상대적인 위상 정보가 중요하다는 데에 착안하여, stationary 신호에서 인지 특성상 무관한 위상 정보를 판별하는 이론과 방법을 제안하였다. 이 이론에 의해 고조파 신호에서는 "임계 위상 주파수"가 정의되어 이 주파수 이상에 해당하는 위상 정보만 인지 특성상 중요하게 된다.

제안된 방법에 의하면 여성 음성처럼 높은 기본주파수 신호보다 남성 음성과 같은 낮은 기본주파수의 신호에 인지 특성상 중요한 위상 정보의 비율이 더 높게 된다. 이 특성은 여성 음성의 경우 음성의 크기 스펙트럼 정보에 많은 bit을 할당하는 Sinusoidal coder가 성능을 잘 발휘하고, 남성 음성의 경우 위상 정보에 많은 bit을 할당하는 CELP 계열 부호화기가 음질이 더 좋은 사실과 일치한다.

비록 주관적 음질 평가 결과가 제안된 방법의 우수성을 입증해 주고는 있지만, 제안된 방법을 음성이나 음악 부호화에 효과적으로 적용시키기 위해서는 프레임 경계에서의 위상 변이를 감소시킬 수 있는 보다 향상된 방법의 제안이 앞으로 필요하다.

참 고 문 헌

1. B. S. Atal and M. R. Schroeder, "Predictive coding of speech signals and subjective error criteria," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 27, no. 3, pp. 247-254, 1979.
2. J. Johnston, "Transform coding of audio signals using perceptual noise criteria," *IEEE J. Sel. Areas in Comm.*, pp. 314-323, 1988.
3. M. R. Schroeder, "New results concerning monaural phase sensitivity," *J. Acoust. Soc. America*, 1579, 1959.
4. F. A. Bilson, "On the influence of the number and phase of harmonics on the perceptibility of the pitch of complex signals," *Acoustica*, vol. 28, pp. 60-65, 1973.
5. R. Plomp and H. J.M. Steeneken, "Effect of phase on the timbre of complex tones," *J. Acoust. Soc. America*, vol. 46, pp. 409-421, 1969.
6. A. E. Rosenberg, "Effect of glottal pulse shape on the quality of natural vowels," *J. Acoust. Soc. America*, vol. 49, no. 2, pp. 583-590, 1971.
7. J. Sköglund, W. B. Kleijn, and P. Hedlin, "Audibility of pitch-synchronously modulated noise," in *IEEE Speech Coding Workshop*, pp. 51-52, 1997.
8. H. Poblath and W. B. Kleijn, "On phase perception in speech," in *Proc. Int. Conf. on Acoust., Speech, Signal Processing*, (Phoenix, Arizona), 1999.
9. R. D. Patterson, "A pulse ribbon model of monaural phase perception," *J. Acoust. Soc. America*, vol. 82, no. 5, pp. 1560-1586, 1987.
10. B. R. Glasberg and B. C. J. Moore, "Derivation of auditory filter shapes from notched-noise data," *Hearing Res.*, vol. 47, pp. 103-108, 1990.
11. H. Levitt, "Transformed up-down method in psychoacoustics," *J. Acoust. Soc. America*, vol. 49, pp. 467-477, 1971.
12. Y. D. Cho, M. Y. Kim, and S. R. Kim, "A spectrally mixed excitation (SMX) vocoder with robust parameters determination," in *Proc. Int. Conf. On Acoust., Speech, Signal Processing*, (Seattle, WA, USA), pp. 601-604, 1998.
13. R. J. McAulay and T. F. Quatieri, "Sinusoidal coding," in *Speech coding and synthesis* (W. B. Kleijn and K. K. Paliwal, eds.), pp. 121-173, Elsevier, 1995.

▲ 김도석(Doh-Suk Kim)

1987년 3월 ~ 1991년 2월 : 한양대학교 전자공학과(공학사)

1991년 3월 ~ 1993년 2월 : 한국과학기술원 전기및전자공학
학과(공학석사)

1993년 3월 ~ 1997년 2월 : 한국과학기술원 전기및전자공
학과(공학박사)

1993년 1월 ~ 1996년 6월 : 시스템 공학연구소 별정직
연구원

1997년 3월 ~ 1997년 10월 : 한국과학기술원 정보전자
연구소

1997년 11월 ~ 1998년 10월 : Bell Laboratories, Lucent
Technologies

1998년 11월 ~ 현재 : 삼성종합기술원 HCI LAB

※ 주관심분야: 심리음향, 청각 모델링, 객관적 음질 평
가, 음성 부호화, 음성 인식