

G.723.1 보코더에서 잡음환경에 강인한 음성활동구간 검출기에 관한 연구

A Study on a Robust Voice Activity Detector Under the Noise Environment in the G.723.1 Vocoder

이 희 원*, 장 경 아**, 배 명 진***
(Hee Won Lee*, Kyung A Jang**, Myung Jin Bae***)

* 코오롱 정보통신 VI팀, ** 숭실대학교 대학원 정보통신공학과, *** 숭실대학교 정보통신공학과
(접수일자: 2001년 8월 3일; 수정일자: 2001년 12월 10일; 채택일자: 2002년 1월 28일)

일반적으로 음성활동 구간 검출기 (Voice Activity Detector)의 가장 큰 문제점은 어떠한 배경 잡음에 대해서도 음성 신호를 검출할 수 있어야 한다는 것이다. 이런 문제를 해결하기 위해서 G.723.1에서는 스펙트럼 특성과 입력 신호의 주기성을 이용하고 있지만 신호대 잡음비가 낮은 신호에 대해 정확한 판정을 한다는 것은 상당히 어렵다. 따라서 본 논문에서는 에너지, LSP (Line Spectrum Pair) 분산을 이용하는 새로운 방법을 제안한다. 제안한 알고리즘에 대한 처리시간 및 음질 평가를 측정한 결과 묵음구간이 정확하게 판정되므로 G.723.1에 비해 감소하였으며 주관적인 음질평가의 경우 기존의 G.723.1에 비해 차이가 거의 없었다. 마지막으로 전송률 측정을 위해 신호대 잡음비 5 dB에서 10 dB 사이의 음성 신호의 경우 최적의 전송률 감소 효과를 얻을 수 있었으며 처리시간의 비교 결과 평균 8% 정도의 처리시간 감소 효과를 얻을 수 있었다.

핵심용어: 음성활동 검출기, 에너지, LSP

투고분야: 음성처리 분야 (2.1)

Generally the one of serious problems in Voice Activity Detection (VAD) is speech region detection in noise environment. Therefore, this paper propose the new method using energy, lsp variation. As a result of processing time and speech quality of the proposed algorithm, the processing time is reduced due to the accurate detection of inactive period, and there is almost no difference in the subjective quality test. As a result of bit rate, proposed algorithm measures the number of VAD=1 and the result shows predominant reduction of bit rate as SNR of noisy speech is low (about 5 ~ 10 dB).

Keywords: Voice activity detector, Energy, LSP

ASK subject classification: Speech signal processing (2.1)

I. 서론

최근 십여 년간의 사회 동향을 볼 때 정보통신분야, 특히 멀티미디어와 이동 통신으로 대표되는 두 분야의 발전은 이루 헤아릴 수 없이 급속한 발전의 양상을 보이

고 있다. 이들 분야는 여러 가지 표현 수단을 가지고 최적의 조건하에서 정확한 정보의 전달을 목적으로 하고 있다. 그 중에서 음성 정보의 전달은 인간에게 있어서 가장 익숙한 정보 전달의 수단 중 하나이며 가장 오랫동안 정보의 전달에 사용되어 왔다.

특히 최근 디지털 이동통신 및 개인 통신망의 눈부신 발전으로 이동통신 사용자를 보다 많이 수용하고 양질의 서비스를 제공하기 위한 음성부호화 연구가 진행되고 있

책임저자: 이희원 (heewon0402@hotmail.com)
157-727 서울시 강서구 등촌동 628
코오롱 정보통신 VI팀
(전화: 02-3660-7769; 팩스: 02-3660-7439)

다. 가입자의 수용율을 증가시키는데 있어서 효과적인 방법 중의 하나는 부호화기의 전송율을 낮추는 방법이다. 이론적으로 디지털 셀룰라 망에서 보코더의 전송율이 1/2로 낮아질 경우 약 2배의 가입자를 수용할 수 있다고 알려져 있다[1]. 하지만 전송율 감소에 따른 음질저하의 발생은 필수적이므로 선진국에서는 다음과 같은 음성부호화기의 요구조건을 충족시키는 보코더 개발을 진행하고 있다[1]. 첫째, 가능한 원음의 음질을 유지하면서 낮은 전송율로 전송이 가능해야 하고 둘째, 간단한 알고리즘으로 하드웨어 구현시 복잡도가 낮아야 하고 저전력 소비로 단말기 이용시간을 증대해야 한다. 셋째, 부호기/복호기의 지연이 적어야 하고 에러에 강인해야 한다. 현재 우리나라의 경우 2세대 디지털 방식인 CDMA에서 음성 유무 검출 (Voice Activity Detection)을 이용한 QCELP (Quadric Code Excited Linear Prediction)가 현재 표준안으로 확정되어 이용되고 있으며 세계 각국에서 여러 음성 부호화 알고리즘을 자국의 실정에 맞게 표준화하여 사용하고 있다.

현재까지 발표된 음성 부호화기 중 가장 많은 연구가 이루어지고 있는 방식이 CELP (Code Excited Linear Prediction) 방식이다. 이 방식은 4.8 Kbps 내외의 전송율에서 양호한 음질을 얻을 수 있으며 ITU-T, TTA/EIA 등 여러 국제표준화 기구를 통해 다양한 응용분야에서 표준화가 이루어지고 있다. 특히 국내에서는 인터넷폰과 PCS 등의 출현으로 CELP 부호화기에 대해 많은 관심이 집중되고 있다[2].

이런 CELP 계열 부호화기 중 1996년도에 표준화가 이루어진 G.723.1과 G.729나 현재 국내의 디지털 셀룰라 서비스에 사용되고 있는 Qualcomm 사의 QCELP 등에서는 묵음 구간에서의 전송율을 감소시키기 위해 VAD (Voice Activity Detection)를 사용하고 있다. 이 중 G.723.1의 VAD는 현재 프레임의 음성신호 존재유무를 판정을 위해 에너지 파라미터를 사용하고 있으며, 연속적인 프레임의 피치값과 정현파 검출 여부를 이용하여 에너지 임계값을 설정한다[3].

일반적으로 VAD의 가장 큰 문제점은 어떠한 배경 잡음에 대해서도 음성 신호를 검출할 수 있어야 한다는 것이다. 이런 문제를 해결하기 위해서 G.723.1에서는 스펙트럼 특성과 입력신호의 주기성을 이용하고 있다. 하지만 이런 파라미터는 판정에 직접적으로 영향을 미치는 것이 아니기 때문에 신호대 잡음비가 낮은 신호에 대해 정확한 판정을 한다는 것은 상당히 어렵다.

따라서 본 논문에서는 잡음환경에서 강인한 음성활동 구간 검출하기 위한 다음과 같은 알고리즘을 제안한다. 처리시간의 불필요한 소모를 방지하기 위해 잡음의 영향을 배제한 에너지비가 미리 설정된 에너지 문턱값을 넘을 경우 음성 구간이라고 판정하는 간단한 결정 논리를 설명한다. 그리고 실험을 통하여 본 논문에서 제안한 알고리즘의 성능을 고찰하고 결론을 맺는다.

II. 잡음환경에 강인한 음성활동구간 검출기 알고리즘

2.1. 에너지 비, LSP 계수, 평균을 이용한 VAD 알고리즘

묵음 구간에서의 전송율을 낮추기 위한 G.723.1 VAD는 판정의 안정성과 연속성을 위해 2장에서 설명한 바와 같이 여러가지 파라미터를 사용하고 있다. 또한 신호대 잡음비가 낮은 신호에 대한 정확한 판정을 위해 스펙트럼 특성을 이용하고 있다. 하지만 위와 같은 방법은 실제 임계값이 현재의 에너지 레벨에 도달할 때까지 정확한 판정을 하는 것이 불가능하며 또한 신호대 잡음비가 낮은 신호가 입력으로 들어올 때 묵음이 존재하는 프레임에 대해서도 음성이 존재하는 프레임으로 판별함으로써 전송율을 낮추는데 기여하지 못하고 있다[3,12,13]. 따라서 본 논문에서는 기존의 방법이 가지는 장점은 해치지 않는 범위 내에서 묵음이 존재하는 프레임을 보다 정확히 검출함으로써 전송율을 낮추는 방법에 대해 제안한다.

본 논문에서 쓰이는 파라미터는 에너지비, LSP 계수, 평균이다. 먼저 입력 신호의 처음 세 프레임을 잡음이라고 생각하고 잡음 구간의 분산과 에너지를 이용하여 분산 문턱값을 만든다. 그 후 한 프레임의 에너지와 일정한 크기 이상의 신호의 에너지를 비를 조사한다. 이 값이 에너지 문턱값을 넘으면 음성 구간이라고 판단하고 그렇지 않으면 잡음과 음성 신호의 LSP 분산값을 이용한다. 이 값이 분산 문턱값을 넘으면 음성구간이라고 판단한다. 그렇지 않으면 백색 잡음 신호 크기의 평균이 영이라는 사실을 이용하여 신호 자체 크기의 평균이 주어진 평균 문턱값을 넘으면 음성이라고 판단한다. 평균 문턱값을 넘지 않는 경우가 일정 횟수보다 많이 발생하면 VAD를 0으로 설정한다. 다시 묵음구간으로 들어가게 되면 잡음 구간에 적응적으로 LSP변화도를 측정하기 위해 분산의 문턱값을 갱신하고 위의 과정을 반복한다.

2.2. 음성 구간 검출

2.2.1. 에너지 비가 문턱값을 넘는 경우

에너지 비는 다음과 같이 표현된다.

$$EneRatio = \frac{\sum_{n=0}^{N-1} s_L^2}{\sum_{n=0}^{N-1} s^2(n)},$$

$$s_L(n) = \begin{cases} s(n), & s(n) > EneTh \\ 0, & otherwise \end{cases} \quad (1)$$

여기서 $s(n)$ 은 음성 신호이고, $s_L(n)$ 은 다음과 같이 정의된다. $EneTh$ 는 에너지 문턱값을 의미한다. 식 (1)과 같은 에너지 비를 사용하면 에너지를 이용하여 음성 구간을 검출할 때 신호대 잡음비가 낮은 신호에서 잡음 구간을 음성 구간으로 판정하는 오류를 줄일 수 있다. 에너지 비가 에너지 문턱값을 넘는 경우 VAD=1로 판정한다.

2.2.2. 에너지 비가 문턱값을 넘지 않는 경우

에너지 비가 문턱값을 넘지 않는 경우는 두 가지 경우

로 나누어 생각한다. 프레임 처음 부분이거나 음성이 끝난 부분에 대해서는 분산의 문턱값을 갱신하고 음성이 진행 중인 경우에는 LSP 분산과 신호의 평균을 이용하여 음성 구간을 검출한다.

(1) 프레임의 처음 부분이거나 음성이 끝난 부분

잡음신호의 분산과 에너지를 이용하여 문턱값을 정하기 위해 먼저 백색잡음(White Noise)을 충분히 발생시켜 LSP의 평균 분포도를 구한다. 그림 1은 잡음신호에 대해 10차 선형예측분석을 통해 얻은 LSP 분포도이다. 10개의 LSP 분포가 일정한 간격을 유지하고 있는 것을 알 수 있다. 표 1은 10개의 LSP의 평균값을 나타낸다.

음성구간 검출을 위해서 먼저 구해진 평균 LSP를 이용하여 입력신호의 처음 3프레임 동안의 신호를 잡음으로 간주하고 LSP를 구하여 평균 LSP값을 갱신한다. 현재 입력신호의 LSP값과 평균 LSP값의 차이를 구하여 그 값을 평균 LSP값에 더한다. 이렇게 함으로써 현재 잡음신호의 특징을 반영할 수 있으며 식 (2)와 같다. 이것은 잡음의

표 1. 잡음 (White Noise) 10차 평균 LSP
Table 1. 10th LSP average of white noise.

LSP 차수	1	2	3	4	5	6	7	8	9	10
주파수 (Hz)	360	740	1080	1460	1820	2200	2540	2920	3280	3660

Occurrence

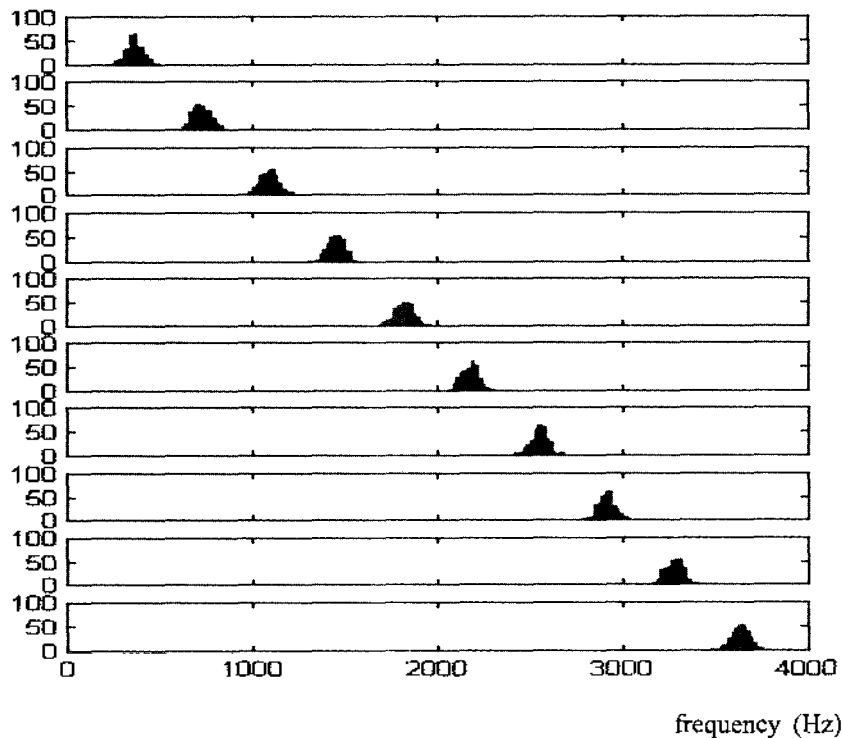


그림 1. 잡음 (White Noise)의 10차 LSP 분포도
Fig. 1. 10th LSP distribution of white noise.

갱신된 평균 LSP값이다.

$$T_{LSP}(i) = Ave_{LSP}(i) + \frac{1}{K} \sum_{k=0}^{K-1} (V_{LSP}(k, i) - Ave_{LSP}(i)) \quad (2)$$

V_{LSP} : 현재프레임의 LSP, $K=3$, $i=0,1,\dots, P-1$ (P :LPC order)

$$Ave_{LSP} = [360, 740, 1080, 1460, 1820, 2200, 2540, 2920, 3280, 3660]$$

var_{LSP} 는 분산의 문턱값과 음성구간 검출에 쓰이는 파라미터이다. 이런 경우 잡음의 분산을 나타낸다.

$$var_{LSP} = \frac{1}{P} \sum_{i=0}^{P-1} (T_{LSP}(i) - V_{LSP}(i))^2 \quad (3)$$

V_{LSP} : 현재프레임의 LSP

var_{LSP} 와 처음 세 프레임의 평균을 이용하여 분산 문턱값 $T var_{LSP}$ 를 구한다. 이 값은 음성이 끝나고 묵음 구간이 시작할 때마다 새롭게 갱신되어 잡음에 적응적인 판단을 가능하게 한다[5,6].

$$T var_{LSP} = \frac{\frac{1}{K} \sum_{i=0}^{K-1} var_{LSP}(i)}{E} + B$$

$$E = R \times \frac{1}{N} \sum_{j=0}^{N-1} s(j)^2 \quad (4)$$

(2) 음성이 진행중인 경우

신호 자체 크기의 평균을 구하여 평균이 0에 근사하는

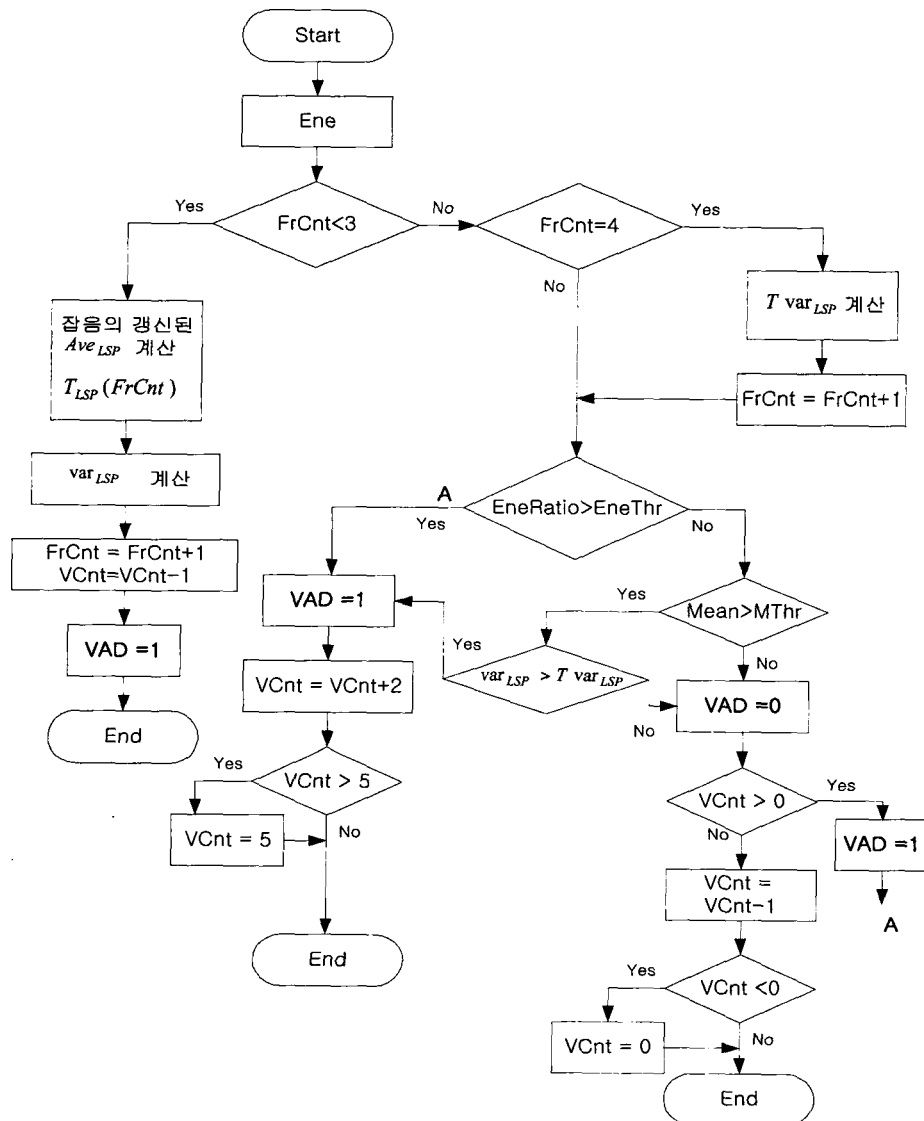


그림 2. 제안한 알고리즘의 순서도
Fig. 2. Block diagram of proposed algorithm.

문턱값 M_{Thr} 보다 작으면 묵음구간으로 결정하고 그렇지 않으면 식 (3)의 var_{LSP} 를 매 프레임마다 구하여 이미 정해진 $T var_{LSP}$ 와 비교하여 var_{LSP} 가 더 크면 음성 구간으로 판정한다. 음성 구간의 연속성을 유지하기 위해서 음성활동구간 검출기 카운터 (Counter)로 V_{Cnt} 라는 변수를 사용하였다.

$$VAD = \begin{cases} 1, & V_{Cnt} > 0 \\ 0, & V_{Cnt} = 0 \end{cases} \quad (5)$$

그림 2는 제안한 알고리즘의 순서도이다. 먼저 처음 세 프레임 동안은 잡음이라고 가정하고 LSP의 분산과 원 신호의 에너지를 이용하여 문턱값을 결정한다. 그 후 한 프레임 전체의 에너지와 일정한 값을 넘는 신호의 에너지 비를 사용하여 간단하게 음성 구간을 검출한다. 이 때 주어진 문턱값을 넘지 않으면 한 프레임의 평균값을 구하여 평균값이 영에 근사하면 잡음이라고 판정한다. 그렇지 않으면 잡음 구간의 LSP와 현재 신호의 LSP의 분산을 구하여 문턱값을 넘으면 음성 구간으로 그렇지 않으면 잡음 구간으로 판정한다.

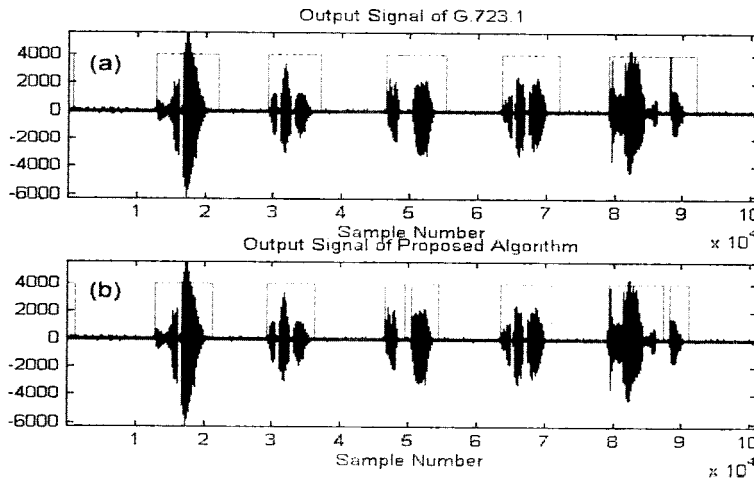
III. 실험 및 결과

제안한 알고리즘을 실험하기 위한 장비는 IBM-PC 586 (333 MHz)에 상용화된 AD/DA 컨버터를 인터페이스한 시스템을 사용하였다. 입력신호는 G.723.1의 입력신호

와 같이 8 kHz로 샘플링하고 16 bit로 양자화한 음성을 입력으로 하며 각 음성시료에 대해 한 프레임의 길이를 240샘플, 부프레임의 길이를 60샘플로 하여 처리하였다. 처리결과의 성능을 측정하기 위해 다음의 음성 시료에 대해 화이트 가우시안 잡음 AWGN과 실제 실험실 환경의 잡음을 첨가하여 사용하였다. 본 논문에서는 음성시료의 신호대 잡음비가 각각 -5, 0, 5, 10, 20 dB가 되도록 화이트 가우시안 잡음 AWGN을 잡음이 없는 환경에서 녹음된 음성에 첨가한 음성과 실제 실험실 환경에서 접할 수 있는 잡음을 마이크를 통하여 입력받은 다음 잡음이 없는 환경에서 녹음된 음성 Clean Speech에 더한 음성시료를 사용하였다. 또한 묵음구간의 검출 결과를 알아보기 위해 각 어절마다 묵음구간을 길게 삽입하여 묵음의 검출여부를 관찰하였으며 대표적인 문장을 연령층이 다른 남녀 화자가 발성한 음성을 음성시료로 사용하였다.

- 발성1: /인수네 꼬마는 천재소년을 좋아한다./
- 발성2: /예수님께서 천지창조의 교훈을 말씀하셨다./
- 발성3: /창공을 헤쳐 나가는 인간의 도전은 끝이 없다./
- 발성4: /숭실대학교 정보통신과 음성통신 연구팀이다./
- 발성5: /공일이삼사오육칠팔구/

제안한 알고리즘의 시뮬레이션은 C-언어로 구현하여 수행하였다. 성능 비교는 G.723.1 Annex A를 통과한 음성과 제안한 알고리즘을 통과한 음성을 비교하였다. 전송률 감소량을 측정하기 위해 전체 프레임 중에서 $VAD=1$ 로 판정한 프레임 수를 비교하였으며 처리시간 감소량

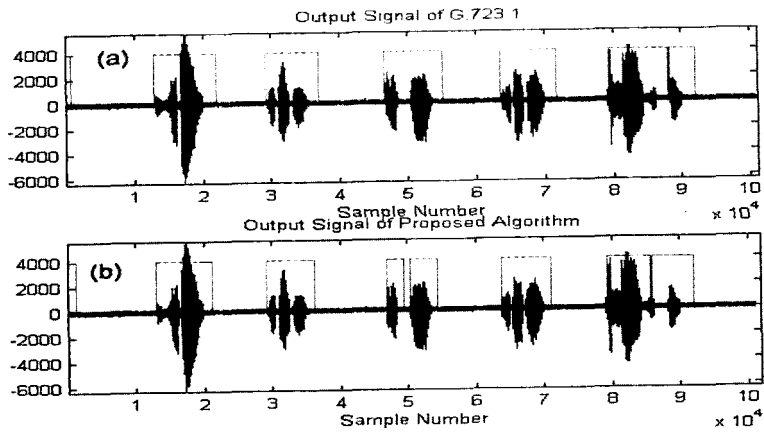


/인수네 꼬마는 천재 소년을 좋아한다/

(a) G.723.1 알고리즘 (b) 제안한 알고리즘
(a) G.723.1 algorithm (b) Proposed algorithm

그림 3. Clean Speech

Fig. 3. Clean speech (male speaker) / Insune komaneun cheonjae sonyuneul joahanda.

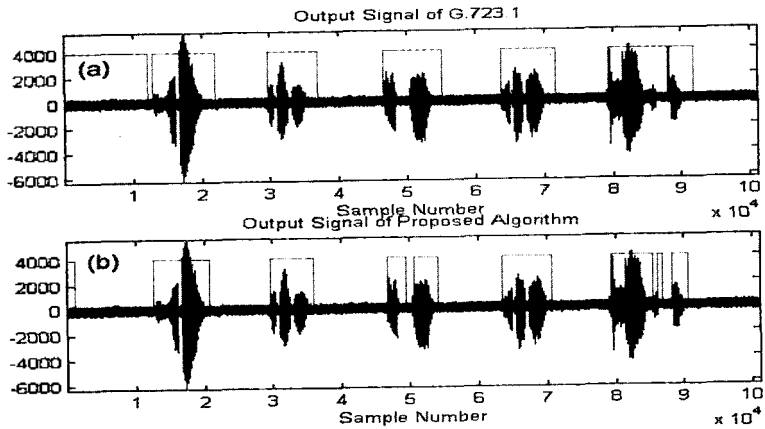


/인수네 꼬마는 천재 소년을 좋아한다/

(a) G.723.1 알고리즘 (b) 제안한 알고리즘
 (a) G.723.1 algorithm (b) Proposed algorithm

그림 4. Clean Speech + AWGN (20dB)

Fig. 4. Clean speech + AWGN (SNR 20dB) / Insune komaneun cheonjae sonyuneul joahanda.

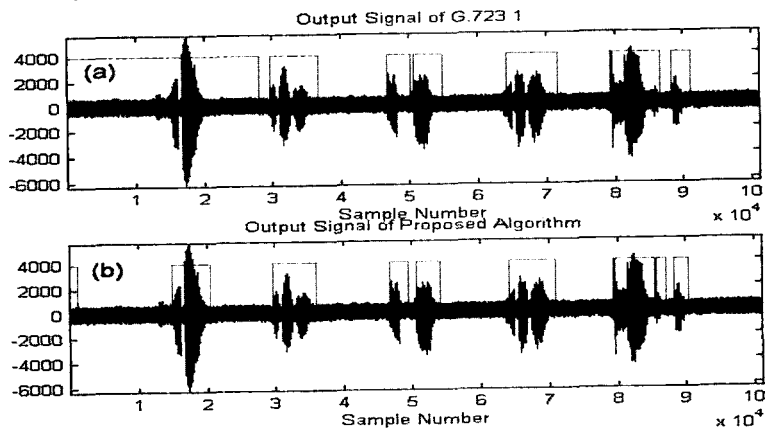


/인수네 꼬마는 천재 소년을 좋아한다/

(a) G.723.1 알고리즘 (b) 제안한 알고리즘
 (a) G.723.1 algorithm (b) Proposed algorithm

그림 5. Clean Speech + AWGN (10dB)

Fig. 5. Clean speech + AWGN (SNR 10dB) / Insune komaneun cheonjae sonyuneul joahanda.



/인수네 꼬마는 천재 소년을 좋아한다/

(a) G.723.1 알고리즘 (b) 제안한 알고리즘
 (a) G.723.1 algorithm (b) Proposed algorithm

그림 6. Clean Speech + AWGN (5dB)

Fig. 6. Clean speech + AWGN (SNR 5dB) / Insune komaneun cheonjae sonyuneul joahanda.

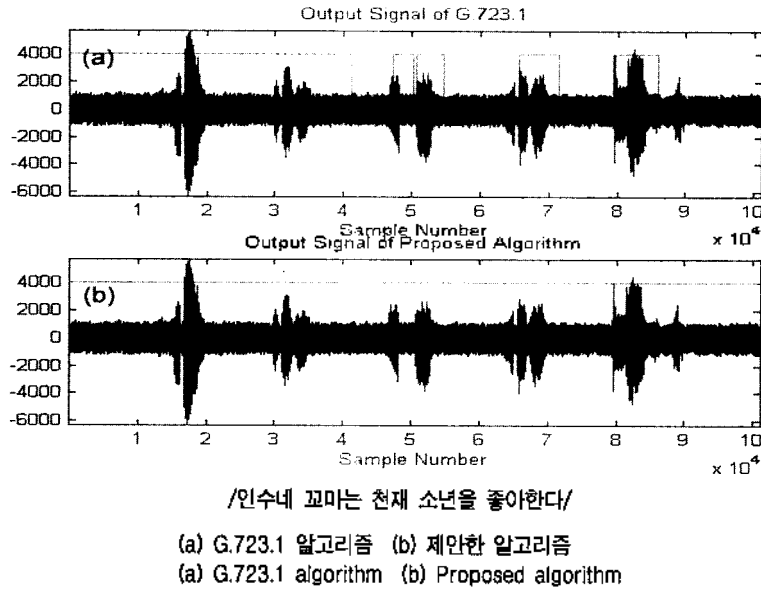


그림 7. Clean Speech + AWGN (0dB)
 Fig. 7. Clean speech + AWGN (SNR 0dB) / Insune komaneun cheonjae sonyuneul joahanda.

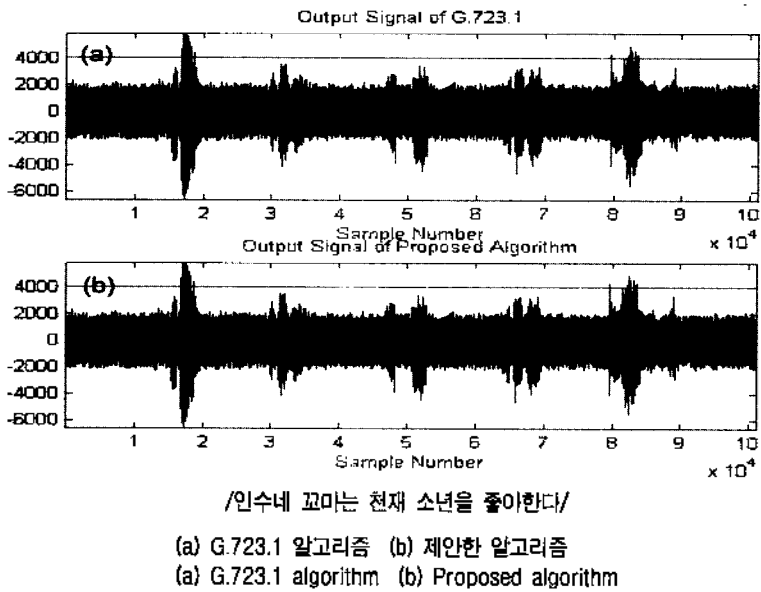


그림 8. Clean Speech + AWGN (-5dB)
 Fig. 8. Clean speech + AWGN (SNR -5dB) / Insune komaneun cheonjae sonyuneul joahanda.

측정은 C 알고리즘이 제공하는 clock 함수로 측정하였다. 실험 환경은 MP-MLQ 모드에서 실험실 환경의 잡음이 첨가된 음성을 사용하여 5번 측정된 다음 평균값을 사용하였다. 음질 측면에서는 MOS 테스트를 사용하였다.

그림 3-8은 음성 시료 1을 여성화자가 발성한 것으로서, 0dB음성시료를 사용한 경우 제안한 알고리즘이 전송률은 증가하나, 기존의 알고리즘보다 정확한 판정을 수행하는 것을 알 수 있다. 그리고 잡음이 없는 환경에서 녹음된 음성인 clean speech, 20 dB, 10 dB, 5 dB음성에서는 전송률의 뚜렷한 감소가 나타나는 것을 알 수 있다.

표 2와 3은 전송률 감소량을 측정하기 위해 VAD=1로

표 2. VAD=1로 판정한 프레임의 수 (음성 시료 2, 남성화자)
 (전체 367 프레임, 눈으로 확인한 음성구간 152 프레임)
 Table 2. The number of active speech frame (Speech data 2, male speaker) (Total 367frames, 152 speech frames confirmed with eyes).

	Clean	-5dB	0dB	5dB	10dB	20dB	실험실 잡음
G.723.1 VAD 알고리즘	170	358	253	235	213	173	232
제안한 VAD 알고리즘	166	358	358	137	154	152	150
감소율 (%)	2.35	0	-41.5	41.7	27.7	12.1	35.3

표 3. VAD=1로 판정한 프레임의 수 (음성 시료 3, 여성화자)
(전체 424 프레임, 눈으로 확인한 음성구간 202 프레임)
Table 3. The number of active speech frame (Speech data 3, female speaker) (Total 424frames, 202 speech frames confirmed with eyes).

	Clean	-5dB	0dB	5dB	10dB	20dB	실험실 잡음
G.723.1 VAD 알고리즘	231	263	252	239	243	252	296
제한한 VAD 알고리즘	235	207	167	187	197	225	193
감소율 (%)	-1.73	21.3	33.7	17.3	18.9	10.7	34.8

표 4. 처리시간의 비교 (sec, 실험실 환경의 잡음을 첨가한 경우)
Table 4. Comparison of processing time (sec, in case of adding a background noise of laboratory).

	발성 1	발성 2	발성 3	발성 4	발성 5
G.723.1 VAD 알고리즘	8.92	10.36	10.57	10.32	11.93
제한한 VAD 알고리즘	7.89	9.69	9.68	9.37	10.79
감소율 (%)	6.89	6.46	8.58	9.21	9.55

표 5. 주관적 음질의 비교 (MOS, 실험실 잡음을 첨가한 경우)
Table 5. Comparison of subjective quality (MOS, in case of adding a background noise of laboratory).

	발성 1	발성 2	발성 3	발성 4	발성 5	평균
G.723.1	3.7	3.8	3.7	3.9	3.8	3.78
제한한 알고리즘	3.7	3.7	3.7	3.9	3.8	3.76

판정한 프레임 수를 비교한 것으로 신호대 잡음비 5 dB 이상의 음성신호에서 잡음의 에너지가 클수록 전송률 감소효과가 큰 것을 알 수 있다. 표 3은 처리시간을 비교한 것으로 VAD 알고리즘에 의해 묵음 구간으로 판정한 경우 파라미터 추출 과정을 생략함으로써 부가적으로 처리시간 감소효과를 얻을 수 있으므로 제한한 알고리즘이 기존의 알고리즘에 비해 처리시간이 감소하는 것을 알 수 있다. 표 5는 음질을 비교한 것으로 기존의 알고리즘에 비해 전송률은 감소하는 반면 음질에는 거의 차이가 없음을 알 수 있다.

IV. 결론

CELP 계열 음성 부호화기 중에서 인터넷 폰 및 화상통신을 위해 개발된 G.723.1 6.3/5.3 Kbps 이중 전송율 음성 코덱은 묵음 구간에서의 전송율을 감소시키기 위해 음성활동

검출기를 사용하고 있으며, 음성활동 유무판정에 의해 묵음으로 간주된 프레임에 대해서는 신호합성에 필요한 최소한의 파라미터만을 전송하여 전송을 감소와 실시간 구현을 위한 처리시간 감소 효과도 더불어 얻을 수 있다. 본 논문에서는 처리시간의 불필요한 소모를 방지하기 위해 잡음의 영향을 배제한 에너지비가 미리 설정된 에너지 문턱값을 넘을 경우 음성 구간이라고 판정하는 간단한 결정 논리를 사용하였다. 에너지 비가 에너지 문턱값보다 작을 경우 묵음 구간에서 구해 놓은 잡음의 평균 LSP와 음성 신호의 LSP 간의 분산을 구하여 분산 문턱값을 넘으면 음성구간이라고 판단한다. 그렇지 않으면 백색 잡음의 평균이 0이라는 사실을 이용하여 신호의 평균이 주어진 평균 문턱값을 넘으면 음성이라고 판단한다. 하지만 다시 묵음구간으로 들어가면 잡음 구간에 적응적으로 LSP변화도를 측정하기 위해 분산의 문턱값을 갱신하고 위의 과정을 반복한다. 위의 실험 결과 신호대 잡음비 5 dB에서 10 dB 사이의 음성 신호의 경우 최적의 전송률 감소 효과를 얻을 수 있었으며 처리시간의 비교 결과 평균 8% 정도의 처리시간 감소 효과를 얻을 수 있었다. 주관적 음질 평가의 결과 음질의 열하는 거의 발생하지 않았다.

참고 문헌

1. A. M. Kondoz, "Digital Speech-Coding for Low Bit Rate Communication System," 1994.
2. W. B. Kleijn and K. K. Paliwal, "Speech Coding and Synthesis," ELSEVIER SCIENCE B, V., pp. 6-10, 1995.
3. ITU-T Recommendation G.723.1, March, 1996.
4. ITU-T Recommendation G.729, March, 1996.
5. 나덕수, 강은영, 배명진, "LSF를 이용한 음성신호의 골점검출 방법," 한국음향학회, 음성통신 및 신호처리 워크샵 논문집, 2000.
6. 민병준, 강병준, "EVRC 패킷에서 LSP거리를 이용한 음성 골점 검출," 한국음향학회지, 제18권 6호, 8, 1999.
7. L. R. Rabiner, R. W. Schafer, "Digital Processing of Speech Signal," Prentice Hall, 1978.
8. 이태진, 권혁제, 최형기, 신유석, 김종교, "잡음환경에서 음성구간 검출방법에 관한 연구," 신호처리합동학술대회논문집, pp. 257-260, 1997.
9. 이희원, 나덕수, 정찬중, 배명진, "LSP 파라미터를 이용한 음성신호의 성분분리에 관한 연구," 대한전자공학회, 하계종합학술대회 논문집, 1999.
10. 배명진, "디지털 음성부호화," 동영출판사, 1996.
11. 배명진, "디지털 음성분석," 동영출판사, 1998.
12. 김정진, 장경아, 배명진, 홍성훈, 심민규, 성규나, "LSP 정보와 에너지 파라미터를 이용한 G.723.1 음성 활동 검출 장치," 신호처리합동학술대회논문집, 10, 1999.
13. 이희원, 나덕수, 배명진, "LSP변화도를 이용한 G.723.1 보코더의 VAD 성능향상에 관한 연구," 한국음향학회, 2000년도 한국음향학회추계학술대회논문집, vol. 19, no. 2(s), pp. 19-22.

- 2000.
14. 이희원, 배명진, "LSP분포 특성을 이용한 G.723.1 보코더의 계산량 감소," 제13회 신호처리합동학술대회 논문집, vol. 13, no. 1, pp. 127-130, 2000.
 15. 이희원, 나덕수, 배명진, "시간축 스케일링 윈도우를 이용한 스펙트럼누설 감소," 한국음향학회, 2000년도 한국음향학회 하계 학술대회 논문집, vol. 19, no. 1(s), pp. 107-110, 2000.

저자 약력

● 이 희 원 (Hee Won Lee)

1999년 2월: 숭실대학교 정보통신공학과 (공학사)
 2001년 2월: 숭실대학교 정보통신공학과 (공학석사)
 2001년 2월~현재: 코오롱 정보통신
 * 주관심분야: 음성인식

● 장 경 아 (Kyung A Jang)

한국음향학회지 제20권 제3호 참조

● 배 명 진 (Myung Jin Bae)

한국음향학회지 제14권 제2호 참조