

G.723.1에서 Skipping Technique을 이용한 피치검색시간 단축에 관한 연구

김정진, 홍성훈, 박종찬, 배명진

승실대학교 정보통신공학과

On the Reduction of Pitch Search Time for G.723.1 Using the Skipping Technique

JcongJin KIM, SeongHoon HONG, JongChan PARK, MyungJin BAE

Dept. of Telecomm. Engr., Soongsil Univ. Seoul 156-743, Korea

mjbae@saint.soongsil.ac.kr

Abstract

G.723.1은 저 전송률 환경에서 고품질을 제공하여 주고 있으나 CELP형 부호화기가 갖는 합성에 의한 분석(analysis by synthesis) 방식의 구조로 인해 많은 처리 시간과 계산량을 요구하게 된다.

본 논문에서는 G.723.1에 대해 skipping 기법을 이용하여 피치 검색과정의 계산량을 줄여 부호화기의 전체 처리 시간을 감소시키는 방법을 제안하였다. 예측 피치를 찾기 위한 개화로 피치 예측(open loop pitch estimation) 과정에서 계산량을 줄이기 위해 skipping 기법을 사용하였다. 피치 예측 과정시 상관관계 곡형은 영과 음의 곡형이 교대로 나타나는 특징을 가지고 있기 때문에 계산시 음의 곡형을 생략하는 방법을 사용하였다.

실제 음성자료에 대해 제안한 피치 검색법을 적용하였을 때 부호화시 평균 처리시간은 약 10%정도 감소하였으며 기존 G.723.1과 제안한 방법을 적용한 G.723.1의 음질 비교를 위하여 MOS 평가를 했을 때 기존의 방법이 평균 3.76인데 비해 제안한 방법의 평균 MOS는 3.73으로 주관적인 음질 저하는 거의 나타나지 않았다.

1. 서론

음성 신호를 저장하거나 전송하기 위한 음성 부호화법은 크게 파형부호화법, 신호원부호화법, 혼성부호화법등 세 가지로 나눌 수 있다. 파형부호화법은 음성 신호의 성분 분

리 없이 곡형 자체의 잉여 성분만을 제거한 후 부호화 하여 전송하고 다시 합성하는 방식이다[1]. 파형부호화법은 고품질과 화자의 개성이 유지되는 반면에 전송 파형을 유지하기 위한 데이터 량이 많기 때문에 전송률이 높으며 내용량의 메모리가 필요하게 된다. 이에 비해, 신호원부호화법은 음성의 발생 모델에 근거하여 음성 신호의 여기 성분과 여파기 성분을 분석하여 각각을 독립적으로 분리시켜 부호화 하는 방법을 사용하기 때문에 전송 대역폭이 작고 메모리 사용량이 작다. 그러나, 분석사와 합성시의 오차가 누적되는 문제점을 갖기 때문에 합성 음질은 자연성과 명료성이 크게 떨어지게 된다[2][3].

혼성부호화법은 신호원부호화법의 메모리 효율성 및 저 전송률과 파형부호화법의 고품질 유지의 장점을 결합시킨 것으로 모던트 정보는 선형 예측 부호화법을 사용하게 되며, 그 나머지 잔여 신호를 어떻게 부호화 하느냐에 따라 RELP, VELP, MELP, CELP등이 제안되어져 있다[3].

이들 중에서 CELP 계열의 모코더들은 16kbps에서 2.4kbps까지의 저 전송률에서 고품질을 유지해 주기 때문에 이동 통신 및 멀티미디어 환경의 응용 분야에서 주로 사용되어져 있다. 이 계열의 부호화법들은 ITU-T에서 여러 기술들이 연구, 논의되어 왔으며 현재는 표준안으로 채택되어 있다. 이들 중에서 G.723.1은 멀티미디어 통신 환경하의 음성 전송 표준 모코더로 개발되었다.

G.723.1은 5.3/6.3kbps의 이중 전송률을 갖는 구조로 현재 별정 통신으로 상용화되는 인터넷과 그 외의 이동 통신용 모코더로 사용되어지고 있으며 낮은 전송률에 비해서 우수한 음질을 제공하고 있다. 더불어 최적의 전송 환경을 위하

여 두 개의 전송률을 사용하기 때문에 다른 코덱 표준안들에 비해서 더욱 응용성이 높다. 그러나 G.723.1 역시 음성 신호를 성분 분리하여 합성하는 방식인 CELP 코덱 계열의 합성에 의한 분석방법을 사용하기 때문에 많은 계산량으로 인한 처리 시간의 소모를 피할 수 없다는 문제점을 갖고 있다[2]-[8].

본 논문에서는 음성부호화기를 실시간으로 구현하는 경우에 발생하는 문제점들을 극복하여 음질 향상과 저 전송률 환경하에서 계산량을 감소시키는 기법을 제안하고자 하는 것이다. 이를 위하여 고음질 저전송 코덱인 G.723.1에 관한 연구를 수행하여 이후 저전송 환경하에서도 음질 열하 없이 처리가 가능한 음성 부호화기를 개발할 수 있는 기반 기술을 확보하고자 한다. 논문에서는 G.723.1의 피치 검색과정에서 skipping 기법을 적용시킨 후 기존의 G.723.1에 비해 피치 검색시간을 감소시켜 전체 처리 시간의 감소를 유도하는 방법을 제안하였다.

II. G.723.1 Dual Rate 음성 부호화기

G.723.1 부호화기는 5.3kb/s와 6.3kb/s의 두 가지 전송률을 가지며 6.3kbps인 경우가 더 좋은 음질을 제공한다. 부호화기는 선형 예측 합성에 의한 분석 부호화의 원칙에 기본을 두고 있고 인지 가중화된 신호의 최소 오차를 얻으며 240표본의 프레임단위로 처리되어 8kHz 표본화에서 30ms가 된다. 분석을 위해 각 프레임은 60표본씩 4개의 부프레임으로 나누어지고 모든 부프레임에서 10차의 선형 예측 부호화기 필터는 입력신호를 사용하여 계산되어진다. 마지막 부프레임을 위한 선형 예측 부호화기 필터는 예측 분할 벡터 양자화기(predictive split vector quantizer)를 사용하여 양자화된다. 양자화 되지 않은 선형 예측 부호화기의 계수는 인지 가중화된 음성 신호를 얻기 위한 단구간 인지 가중화 필터를 구축한다[3].

두 개의 부프레임(120표본)을 처리하기 위한 개회로 피치 주기 예측은 인지 가중화된 음성 신호를 사용하여 계산되어진다. 피치 주기는 18에서 142표본의 구간에서 찾으며 피치 예측은 120표본의 블록 단위로 실행되어진다. 이전에 계산되어 예측된 피치 주기를 사용하여 하모닉 노이즈 셰이핑 필터(harmonic noise shaping filter)가 만들어지며 LPC 합성 필터(LPC synthesis filter), 포먼트 인지 가중화 필터(formant perceptual weighting filter), 하모닉 노이즈 셰이핑 필터의 결합은 임펄스 응답을 생성하기 위해 사용된다. 피치 주기 예측값과 임펄스 응답을 이용하여 폐회로 피치 예측기(closed loop pitch predictor)값이 계산되어진다. 5차 피치 예측기가 사용되는데 피치주기는 개회로 피치 예측값 주위의 차이값이 작은 구간에서 계산되어진다. 피치 예측기는 초기 목적벡터(initial target vector)로부터 값을 뺀 때

이용되는데 피치 주기와 차이값 둘 다 복호화기로 전송되어진다. 마지막으로 여기신호의 비주기 요소가 근사화 되어진다. 부호화기의 블록도는 그림 2-1에서 나타내었다[12].

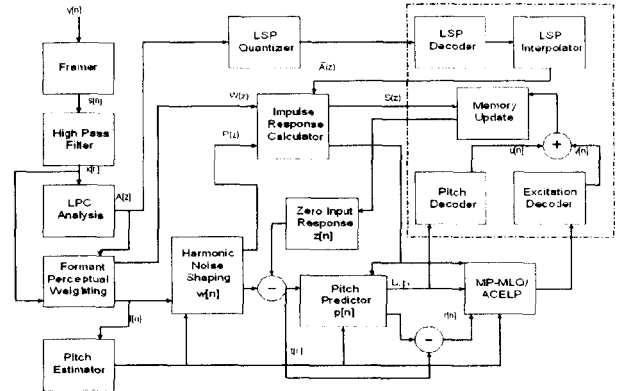


그림 2-1. G.723.1 음성 부호화기 블록도.

III. 피치 검색

피치 예측은 매 프레임당 두 번씩 계산되며, 하나는 처음 두 부프레임에서 다른 하나는 마지막 두 부프레임에서 계산된다. 개회로 피치 주기 예측 L_{ol} 은 인지 가중화된 음성신호를 이용하여 계산된다. 최대 상관관계 $C_{ol}(j)$ 는 피치 주기를 결정하는데 사용되며 다음과 같이 표현된다.

$$C_{ol}(j) = \frac{\left(\sum_{n=0}^{119} f(n) \cdot f(n-j) \right)^2}{\sum_{n=0}^{119} f(n-j) \cdot f(n-j)} \quad 18 \leq j \leq 142 \quad (3.2)$$

상관관계 $C_{ol}(j)$ 가 최대가 되는 인덱스 j 는 두 부프레임에 적당한 개회로 피치 예측값에서 선택된다. 최적의 인덱스를 찾는 동안 얻어진 선택 값들로부터 피치 매수를 찾게 되는 것을 피하기 위해 피치 주기보다 작도록 주어진다. $C_{ol}(j)$ 의 최대값은 j 가 18부터 찾기 시작한다. 모든 최대값 $C_{ol}(j)$ 을 찾으며 이들 값들은 최적 상태이면서 이전의 최대값 $C_{ol}(j')$ 과 비교된다. 만약 인덱스 j 와 j' 의 차이가 18보다 작고 $C_{ol}(j) > C_{ol}(j')$ 이면 새로운 최대값이 선택된다. 만약 인덱스간의 차이가 18보다 크거나 같고 $C_{ol}(j)$ 이 $C_{ol}(j')$ 보다 1.25dB 클 때는 새로운 최대값이 선택된다.

피치 예측기의 피치 예측 기여(contribution)는 기존의 적용 코덱북 기여로 처리한다. 피치 예측기는 5차의 피치 예측기를 사용하며 첫 번째와 세 번째 부프레임에 대해서는 폐회로 피치 지연이 +1의 범위를 갖는 적당한 개회로 피치 지연값으로부터 결정된다. 두 번째와 네 번째 부프레임에

내해서는 이전 부프레임 지연에서 -1, 0, +1, +2의 차이 값을 갖는 피치 지연값을 사용한다. 이전 부프레임 지연값에서의 차이값들이 양자화되고 부호화된 피치 지연값은 이 점으로부터 L_1 로 간주된다. 피치 예측기 이득은 170 엔트리를 갖는 코드북을 사용하여 벡터 양자화를 한다.

IV. 제안한 방법

기존의 G.723.1에서 피치검색은 시간영역에서 open loop 알고리즘을 이용하여 18부터 142 사이에서 피치값을 선택하므로 피치 검색시 많은 시간이 소요된다. 이러한 G.723.1의 복잡한 피치검색방법은 음질의 향상을 가져오지만 피치검색시간이 많이 소요되는 단점을 갖고 있다. 따라서 우리는 기존의 skipping 기법을 G.723.1의 피치검색시에 적용시키는 방법을 제안하였다.

Skipping 기법은 음성음의 경우 실제 피치 지연은 항상 상관관계 함수의 과형의 양의 봉우리에 위치한다는 사실을 이용하여 피치검색시 양의 봉우리에 대해서만 수행하게 되면 불필요한 피치 검색시간을 줄일 수 있게되는 아주 유용한 방법이다[9]. 이것은 다음에 나타나는 음성신호의 상관관계 함수의 특성들에 의해 가능하다: (1)음성음에서 과형은 천천히 변화하기 때문에, 단구간에서 음성신호는 매우 높은 상관관계를 갖는다. (2)양의 봉우리 구간과 음의 봉우리 구간은 번갈아 나타난다. (3)음성음의 제1포먼트의 영향이 지배적이기 때문에 각 봉우리 구간의 폭은 거의 변화하지 않는다. 이들 특성을 이용하여 음의 봉우리 구간은 전의 양의 봉우리 구간의 폭을 계산함으로써 쉽게 제외할 수 있다. 즉, 양의 봉우리 구간의 폭을 계산하고 음의 봉우리가 시작하면 양의 봉우리 구간만큼을 제외하고 상관관계 함수를 수행함으로써 계산량을 감소시킬 수 있다. 그림 4-1에 원래 음성파형과 자기상관 함수파형 그리고 제안한 방법에 의한 처리파형을 나타내었다.

이렇게 하여 상관관계 파형의 봉우리가 양의 값을 나타내는 구간에 대해서 완전한 피치검색을 수행하게 되면, 음양의 봉우리가 교대로 나타날 확률이 50%이기 때문에 검색시간을 약 50% 정도로 절약할 수 있게 된다. 더 필요하다면 양의 봉우리 정점을 기준으로 전후 각각 3표본 정도에 대해서만 수행하게 되면, 양의 봉우리가 평균 6.4개 정도 나타나기 때문에 완전검색에 비해 30% 정도의 피치 검색시간만 필요로 하게 된다. 그러나 본 논문에서는 음의 봉우리검색 과정을 건너뛰는 방법만을 적용하였다.

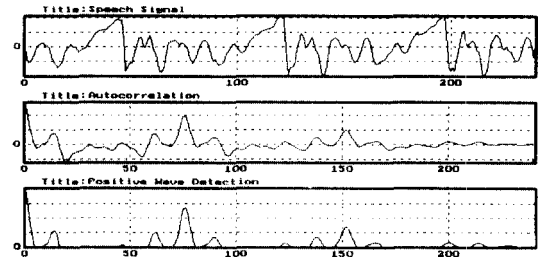


그림 4-1. 제안한 방법의 결과예시.
(a) 음성파형, (b) 기존의 상관관계파형, (c) 제안한 방법의 상관관계파형

V. 실험 및 결과

컴퓨터 시뮬레이션에 이용한 장비는 IBM-PC 586(166MHz)에 상용화된 AD/DA 컨버터를 인터페이스한 시스템이다. 8kHz로 음성을 표본화하였으며 각 사료에 대해 한 프레임의 길이를 240표본으로 하여 부프레임을 60표본 단위로 처리하였다. 처리결과와 성능을 측정하기 위해 다음의 대표적인 문장을 연립송이 다양한 남녀 5명의 화자가 각 5번씩 발생하여 시료로 사용하였다:

- 발성1: /인수네 꼬마는 전재소년을 좋아한다./
- 발성2: /예수님께서 진지장로의 교훈을 말씀하셨나./
- 발성3: /송일대 정보통신과 음성통신 연구팀이다./
- 발성4: /창장을 해쳐나가는 인간의 도전은 끝이 없다./
- 발성5: /May I Help You?/

피치검색은 G.723.1의 피치 검색과정을 C-언어로 구현하여 수행하였다. 성능비교는 기존의 G.723.1의 피치검색과 제안한 방법을 이용한 피치검색을 이용하여 수행하였다. 본 논문에서는 피치검색의 정확성을 기하기 위해 양의 상관관계 파형에 대해서만 완전한 피치검색을 수행하였다. 표5-1에서 볼 수 있듯이 전체 음성처리 시간이 평균 9.6%의 향상을 보였다.

MOS결과는 표5-2에서 볼 수 있듯이 기존의 G.723.1과 제안한 방법의 MOS값이 거의 같음을 볼 수 있다.

VI. 결론

G.723.1 모코더에서 피치검색시간은 저가형 정수 DSP를 사용할 때 총 부호화 처리시간의 약 40% 정도

를 차지한다. 따라서 우리는 G.723.1 보코더의 피치 검색시간을 개선하는 검색법을 새로이 제안하였다. 먼저, 피치검색시에 자기 상관관계값 중에 양의 봉우리만큼 음의 봉우리를 건너뛰는 skipping 기법을 이용하여 피치를 구하였다. 즉, 간단한 상관관계 함수를 통해 얻어진 상관관계값 중에서 양의 봉우리를 이루는 구간을 예비피치 구간으로 선택하였다. 그런 다음에 예비피치 구간에 대해서만 피치검색을 수행하여 피치 검색시간을 절약하는 새로운 방법을 제안하였다.

실제음성에 대해 제안한 피치검색을 사용하여 G.723.1 부호화를 수행하였을 때, 기존의 G.723.1 피치 검색법 비해 처리시간이 평균 9.6% 감소함을 보였다. 또한 음질의 경우는 MOS 테스트 결과 거의 같음을 보였다.

표 5-1. 기존의 G.723.1과 제안한 G.723.1의 처리시간 비교 (sec).

	발성1	발성2	발성3	발성4	발성5	평균
기존 G.723.1	5.15	5.31	5.14	5.37	4.43	5.08
제안된 G.723.1	4.72	4.78	4.64	4.83	3.98	4.59

표 5-2. 기존의 G.723.1과 제안된 G.723.1의 MOS SCORE 비교.

	발성1	발성2	발성3	발성4	발성5	평균
기존 G.723.1	3.7	3.63	3.87	3.83	3.75	3.76
제안된 G.723.1	3.7	3.61	3.87	3.8	3.68	3.73

VII. 참고문헌

[1] N. S. Jayant and P. Noll, Digital Coding of Waveform-Principles and Applications to Speech and Video, pp.220-221, Prentice-Hall, 1978.
 [2] B. Atal. "Efficient coding of LPC parameter by temporal decomposition", Proc. of ICASSP, pp.81-85, 1983.
 [3] A. M. Kondoz, Digital Speech, John Willy & Sons, 1994.
 [4] J. P. Campbell, Jr., V. C. Welch and T. E. Treiman, "Error-protected 4800 bps CELP

Coder(U. S. Fedral Standard 4800 bps Voice Coder)", Proc. Int. Conf. on Acoustics, Speech and Signal Processing, 1989.
 [5] W. B. Kleijn et al., "Fast Methods for the CELP Speech Coding Algorithm", IEEE Trans., Acoustics, Speech and Signal Processing, Vol.38, No.8, pp.1330-1341, August 1990.
 [6] J. H. LEE, H. Y. JEON, M. BAE, and S. ANN, "A Fast Pitch Searching Algorithm using Correlation Characteristics in the CELP Vocoder", IEEE Communication Society, Proceeding of MILCOM '94, pp.699-702, Oct.2-5 1994.
 [7] D. KIM, M. BAE, J. KIM, K. BYUN, K. HAN, H.YOO, "On a Reduction of Pitch Searching Time by Preprocessing in the CELP Vocoder", J. Acoust. Soc. Korea, Vol.13, No.3, pp.33-40, June 1994.
 [8] D. KIM, M. BAE, J. KIM, K. BYUN, K. HAN, H.YOO, "On a Reduction of Pitch Searching Time by Preliminary Pitch in the CELP Vocoder", J. Acoust. Soc. Korea, Vol.13, No.2E, pp.51-57, July 1994.
 [9] J.H. LEE, H.Y. JEON, M. BAE, and S. ANN, "A Fast Pitch Searching Algorithm using Correlation Characteristics in CELP Vocoder," IEEE Communication Society, Proceeding of MILCOM'94, pp.699-702, Oct.2-5 1994.
 [10] J.H.Lee, M.J.Bae, S.G.Ann, H.Y.Yoo, "The Skipping Technique: A Simple and Fast Algorithm to Find the Pitch in CELP Vocoder", IEEE Communication Society, Proceeding of MILCOM'95, pp1263-1266, Nov.6-8 1995.
 [11] J.H.Lee, M.J.Bae, H.Y.Yoo, "A New Fast Pitch Search Algorithm Using The Abbreviated Correlation Function in CELP Vocoder", IEEE Communication Society, Proceeding of MILCOM '96, PP.653-657, Oct.21-24 1996.