

정규화 코드북을 이용한 분할 벡터 구조의 ISF 적응적 양자화 기법

A Method of Adaptive ISF Split Vector Quantization Using Normalized Codebook

박 지 강, 임 종 하, 홍 기 봉, 이 인 성
(Zhigang Piao, Jongha Lim, Gibong Hong, Insung Lee)

충북대학교 전파통신공학과

(접수일자: 2011년 1월 28일; 수정일자: 2011년 4월 11일; 채택일자: 2011년 5월 31일)

본 논문에서는 ISF 계수의 순서화 성질을 이용하여 광대역 음성부화기의 분할구조 벡터양자화기의 단점을 보완함으로써 ISF 계수 양자화의 성능을 개선하는 알고리즘을 제안한다. 음성 부호화기의 ISF 벡터양자화기는 계산량과 메모리 수요량을 줄이기 위하여 벡터분할 구조를 사용한다. 이러한 벡터 분할구조의 양자화기는 분할된 벡터 사이의 상관도를 이용하지 못하였다. 제안하는 알고리즘은 ISF 계수의 순서화 특징을 이용하여 정규화 코드북을 만든다. 그리고 양자화 된 분할 벡터로 정규화 코드북의 분포범위를 적응적으로 변화시킴으로써 양자화 하여야 할 분할 벡터의 코드북을 효율적으로 만들어 준다. 제안된 알고리즘은 분할 벡터사이의 저하된 상관도를 순서화 특징을 통하여 다시 이용하는 방법으로 표준화 코덱인 AMR-WB의 ISF 양자화기에 적용하여 1.5 bit정도의 성능 개선을 얻었다.

핵심용어: LPC 양자화기, 순서화 성질, 정규화 코드북, AMR-WB

투고분야: 음성처리 분야 (2.2)

In most of the ISF (or LSF) based real time speech codec, SVQ (split vector quantization) method is used to decrease the quantizer complexity and memory size of codebook. However, it produces drawback that the level of correlation between code vectors can not be used during vector splits. This paper presents a new method of adaptive ISF vector quantization, which compensates the drawbacks of SVQ structured quantizer for wideband speech codec. In each different frame, the proposed method makes use of the correlation between splitted vectors by adaptively changing codebook distribution according to ordering property of ISF. The algorithm is evaluated in AMR-WB, and shows about 1.5 bit per frame improvement.

Keywords: LPC quantizer, Ordering property, Adaptive codebook, AMR-WB

ASK subject classification: Speech Signal Processing (2,2)

I. 서론

대부분의 음성 부호화기들은 음성의 다구간 스펙트럴 포락선 (spectral envelope) 정보를 전송하기 위하여 선형 예측 계수 (Linear Prediction Coefficient) [1]를 사용한다. LPC 계수들은 수학적으로 동등하며 양자화 및 보간에 좋은 성능을 갖는 LAR (Log Area Ratio) [2], Inverse Sine [3], LSP (Line Spectrum Pairs) [4] 혹은 ISP (Immitance Spectrum Pairs) [5]로 변환된다. 그 중

LSP와 ISP가 음성신호처리에 좋은 성능을 갖고 최근의 음성 코덱에서도 가장 많이 쓰인다. LPC 계수는 ISP와 LSP 형태로 변화되어 양자화 되는데 크게 스칼라 양자화와 벡터양자화 두 가지 기법으로 양자화된다. 벡터양자화기는 스칼라양자화기보다 더 우수한 성능을 보여주지만 계산량 많다는 단점이 있다. 그러나 최근 하드웨어의 연산능력이 향상됨에 따라 음성 양자화 코덱은 대다수가 벡터양자화 기법을 사용한다. 동시에 계산량과 메모리 수요량을 줄이기 위하여 벡터 분할 양자화 (SVQ) [10] 방법을 사용한다.

표준화 코덱인 G.729 [6], G.729.1 [7]은 10차 LPC을

LSP, LSF로 변환하였고 AMR-WB [8]는 16차 LPC을 ISP, ISF로 변환하여 MSVQ (Multistage-Split Vector Quantization) 방식으로 계산량과 메모리 수요량을 줄여 벡터 양자화 한다. 이 구조는 간단하고 계산량이 적은 특성을 갖고 있지만 벡터가 분할되면서 각 서브벡터 사이의 상관도가 사용되지 못하게 되는데, 이는 벡터양자화의 성능저하를 일으킨다.

최근 표준화된 G.718 [9] 음성 코덱에서는 LPC 계수를 분할하지 않고 Multistage 방식의 양자화기를 사용하여 양자화기의 성능을 높였지만 16차 ISF을 분할하지 않았기 때문에 벡터 차원이 크고 양자화 차수가 많아졌으며 양자화 복잡도가 커졌다. 따라서 계산량과 메모리 수요량은 기존 MSVQ 구조의 코덱보다 많다.

최근에 많이 쓰이는 선형예측계수 양자화를 위한 변환방법인 LSP 변환과 ISP 변환은 포먼트 (Formant) 필터의 안정성을 보장하여 주는 순서화 특징을 가진다. 높은 차수 ISF 계수 값은 낮은 차수의 ISF 계수 값보다 커야한다 [3]. 순서화 특징을 이용하여 효율적인 양자화를 위한, 혹은 계산량을 줄이기 위한 많은 알고리즘들이 출시되고 있다. 1995년에 순서화 특징을 이용하여 설계된 코드북 중 순서화 특징에 가장 맞는 코드북을 선택하여주는 Linked Split VQ 기법이 제안되었다 [12]. 이러한 기법은 보다 우수한 성능을 제공할 수 있지만 선택할 수 있는 코드북이 제한적이고, 많은 메모리 수요량을 요구하는 벡터양자화기의 메모리 수요량을 더욱 증가시킨다. 2001년에는 G.723.1 코덱의 LPC 양자화기의 계산량을 감소하기 위하여 순서화특징을 이용한 보다 빠른 코드북검색 알고리즘이 제안되었다 [13]. 최근에 순서화 특징을 이용하여 양자화성능을 개선하는 기법이 제안되었다. 이러한 기법은 불필요한 코드북 인덱스를 검색하는 방법으로 음성의 통계적 특성에 따라 유효 인덱스를 추가 삽입 시켜주는 방법을 사용한다. 이러한 양자화기법은 서브벡터당 평균 15%의 코드북을 설계한다. 따라서 일반적인 벡터양자화기보다 많은 계산량을 갖는다 [14].

본 논문에서 제안하는 양자화 기법은 순서화 특징이 있는 ISF를 이용하여 LBG 알고리즘으로 적응적 양자화를 위한 정규화 코드북을 새롭게 만든다. 제안하는 양자화기는 ISF 계수의 순서화 특징을 이용하여 양자화하려는 서브벡터의 양자화 범위를 제한하여 정규화 코드북의 각 원소 분포범위를 적응적으로 확장/축소하면서 효율적인 양자화를 하게 된다. 즉 이러한 기법은 Linked Split VQ에 코드북 선택하는 기법과 비교하였을 때, 코드북을 정규화하고, 정규화 코드북을 이용하여 유효한 코드북을

프레임별로 만들어 준다고 볼 수 있다. 이러한 ISF 계수의 순서화 특성을 이용한 적응적 양자화 기법은 기존 표준화 코덱인 AMR-WB 코덱에 적용하여 실제 음성 데이터를 사용하여 실험하였고 AMR-WB와의 주파수 왜곡을 비교하였다 [8]. 제안된 양자화기법은 같은 순서화 특성을 갖는 LSP 양자화기에도 사용 가능하다.

본 논문의 내용은 다음과 같다. II장에서는 LPC 양자화 방법에 대하여 간략하게 서술하였고, 본 논문에서 사용한 ISF 파라미터의 특성에 대하여 기술하였다. III장에서는 정규화 코드북을 만드는 방법과 적응적 코드북에 응용하는 알고리즘을 설명하였고 IV장에서는 AMR-WB와의 성능을 비교하였다. 마지막으로 V장에서는 본 논문의 결론 및 향후 연구방향에 관하여 서술하였다.

II. ISF 계수의 정의 및 순서화 특성

대부분의 음성 부호화기는 음성의 다구간 스펙트랄 포락선 정보를 전송하기 위하여 선형 예측 계수 [1]를 사용하는데, 이는 수학적으로 동등하며 양자화 및 보간에 좋은 성능을 갖는 LSP, ISP 계수로 변환되어 양자화된다. AMR-WB 광대역 코덱은 16차의 선형예측계수를 같은 차수의 ISP로 변환한다. 변환된 16차의 ISP q_i 는 식 (1)을 통하여 코사인영역의 ISF로 변환하여 양자화한다 [5].

$$\begin{aligned} f_i &= \frac{f_s}{2\pi} \arccos(q_i), \quad i = 0, \dots, 14 \\ &= \frac{f_s}{4\pi} \arccos(q_i), \quad i = 15 \end{aligned} \quad (1)$$

식 (1)에서 f_i 는 ISF의 주파수영역 [0-6400 Hz]을 가리키고, $f_s=12800$ 는 샘플링 주파수이다. ISF 계수 벡터 원소 (Component)간의 상관성이 존재하는데 이는 주로 순서화 특성으로 표현된다. 순서화 특성이란 식 (2)와 같이 ISF 계수 벡터 원소들이 0-6400 Hz 사이의 주파수 영역에서 순서적으로 증가하는 것을 가리킨다.

$$0 < f_0 < f_1 < \dots < f_{p-2} < f_{p-1} < \pi, \quad \text{and} \quad -1 < f_{15} < 1 \quad (2)$$

p 는 16개의 ISF 계수 순서를 나타내고, f_{15} 는 ISF 계수의 이득 (Gain)이다. 만약 신호의 ISF 계수가 식 (2)의 순서화 특성을 만족한다면 다구간 합성 필터의 안정성은 보장된다. 본 논문에서는 이러한 순서화 특성을 이용하여 효율적인 ISF 계수 양자화를 하였다.

III. 정규화 코드북을 이용한 ISF 적응적 분할 벡터양자화기법

3.1. 정규화 코드북을 사용한 벡터 양자화 기법

벡터 양자화는 양자화를 위한 코드북이 필요하다. 일반적인 벡터양자화의 코드북은 학습 데이터를 통하여 LBG 알고리즘 [11]으로 코드북을 설계한다. LBG 알고리즘은 학습데이터의 통계적인 벡터 분포특성을 통하여 설계되며 학습데이터에 대하여서는 최적화된 코드북이라고 할 수 있다. 본 논문에서 제안된 SVQ 구조의 양자화는 ISF 계수의 순서화 특성을 이용하여 양자화하려는 코드북의 양자화 범위내의 효율적인 양자화를 위한 정규화 코드북을 만들었다. AMR-WB와 같이 실시간 음성코덱에 사용되는 일반적인 ISF 양자화는 벡터를 분할기법이 쓰인다.

벡터 양자화 기법은 원소들을 묶어 하나로 벡터 양자화를 하게 된다. 그림 (1)과 같은 순서화 특성이 존재하는 6차 ISF 계수의 2차원 벡터 양자화를 예로 들어 설명하였다. f_1 부터 f_6 까지의 ISF 계수를 하나의 벡터로 만들어서 코드북을 설계하고 양자화하면 가장 좋은 성능 갖게 되는데 이는 각 벡터사이의 상관도를 최대한 이용했기 때문이다. 그러나 실제적인 선형예측계수 양자화에서는 계산량과 메모리 수효량 때문에 이러한 양자화기를 쓸 수 없다. 즉 벡터 분할기법을 통하여 10 bit의 코드북은 두 개의 5 bit 코드북으로 분할되면서 1024개의 코드 인덱스가 두 개의 32개 코드 인덱스로 감소된다. 광대역 코덱인 AMR-WB에 사용된 16차 ISF양자화기를 보면 46 bit를 사용하는데 이를 코드북으로 만들면 7×10^{13} 개의 코드북 인덱스가 생성된다. 따라서 실제적으로 응용되는 많은 코덱에서는 계산량을 줄이기 위하여 벡터 분할기법이나 멀티스테이지 양자화기법을 사용하지만 벡터가 분할되어 서로 독립적으로 양자화되기 때문에 서로간의 상관도를 사용 못하게 되어 양자화기의 성능은 저하될 수밖에 없다.

본 논문에서는 이러한 성능저하를 보상하여주기 위하여 ISF 원소의 순서화 특성을 사용하였다. 그림 (1)과 같은 6차 ISF 양자화기를 예로 들면 벡터가 $f_1 < f_2 < \dots < f_4 < f_6$ 이라고 할 때, 순서화 특성에 따르면 f_3 은 f_2 보다 크고, f_4 는 f_5 보다 작다. 즉, 특정한 프레임에서 f_2 와 f_5 가 주어졌을 때 제 2서브벡터 코드북 중에 f_3 원소가 f_2 원소보다 작은 인덱스와 f_4 원소가 f_5 원소보다 작은 인덱스들은 코드북상에 불필요한 부분이 된다. 본 논문

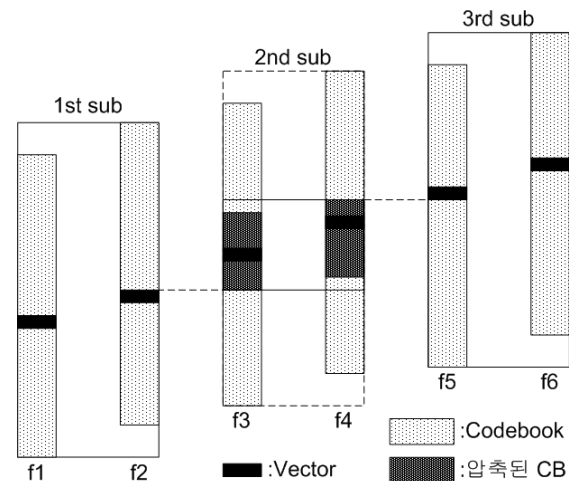


그림 1. 순서화 특성을 이용한 벡터범위 제한

Fig. 1. Distribution Restriction Using Ordering Property.

에서는 양자화 된 ISF 계수와 양자화 할 ISF 계수사이의 순서화 특성을 이용하여 0과 1 사이의 값으로 정규화 코드북을 만들고 정규화 코드북을 이용하여 ISF 벡터 분포 가능 범위만큼 적응적으로 변화시켜 코드북의 양자화 효율을 높인다. 그림 (2)는 코드북의 적응적 변화과정을 나타낸다.

식 (3)은 적응적으로 코드북을 변화시키는 수식이다. 여기서 X_{norm} 은 정규화 코드북, X_{real} 은 ISF 벡터 양자화를 위한 적응적 코드북이며 R_{max} 코드벡터 분포가능범위의 최대치 R_{min} 은 분포가능범위의 최소치를 나타낸다. n 은 코드벡터 및 코드북의 차원을 나타낸다. 우선 제 2서브벡터의 적응적양자화를 위하여 제 1서브벡터와 제 3서브벡터를 양자화한다. 따라서 f_2 와 f_5 을 알게 되는데 이는 각각 제2서브벡터의 분포범위, R_{min} 과 R_{max} 에 해당된다. 분포범위를 알면 정규화 코드북 X_{norm} 을 통하여 그림 (2)에서처럼 제 2서브벡터의 적응적 코드북 X_{real} 을 설계할 수 있다. 만들어진 적응적 코드북으로 제 2서브벡터가 양자화 되는데 원소 f_3 와 f_4 을 조합하여 가장 작은 MSE (Mean Square Error)를 갖는 인덱스가 선택된다. 이러한 기법은 분할된 벡터의 코드북 원소 분포범위를 제한시켜 코드북 심볼 분포가 보다 밀집하게 만들어 준다. 광대역 음성신호에서 16차 ISF를 추출했을 때 4차원 서브벡터가 순서화 특징을 이용하여 범위를 제한하면 평균적으로 15%의 무효 코드북 인덱스가 검색된다. 따라서 R_{min} 과 R_{max} 두 방향으로 범위를 제한하면 평균 30%의 무효 코드북 인덱스가 검색된다. 코드북 원소가 균등하게 분포되었다고 가정할 때, 제안한 방법은 유효한 70%의 코드북 인덱스에 평균적으로 30%의 무효한 코드북

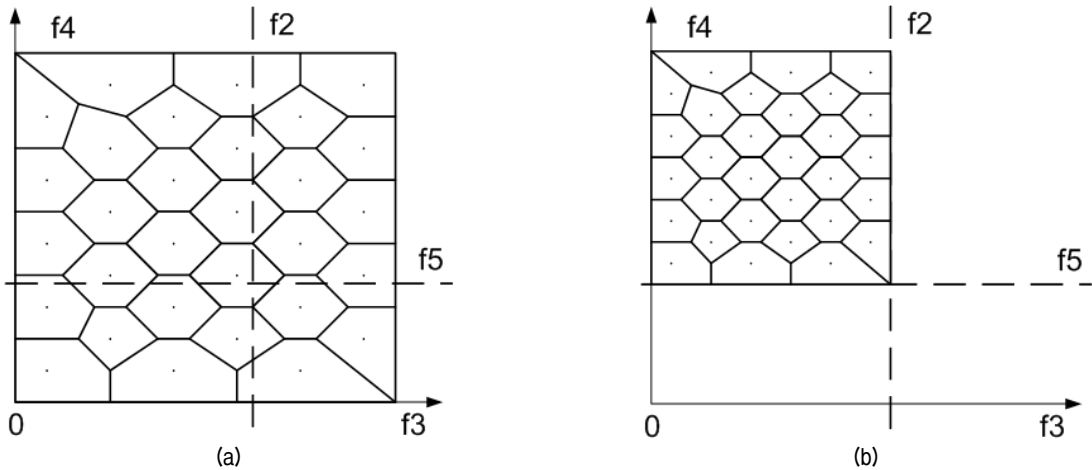


그림 2. 코드북의 적응적 확장
 (a) 일반적인 벡터양자화에 사용된 코드북 기법, (b)정규화 코드북 기법
 Fig. 2. Codebook adaptively Expansion.
 (a) Codebook of conventional VQ, (b) Normalized Codebook

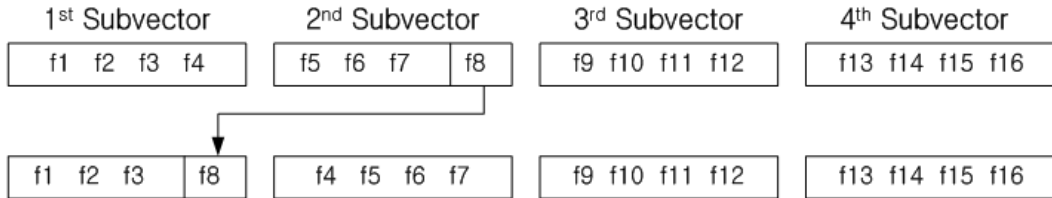


그림 3. 적응적 양자화 기법을 위한 16차 ISF 벡터분할
 Fig. 3. Vector Split for Adaptive quantization method.

인덱스를 유효로 만들어 주기 때문에 유효 인덱스 밀집 정도는 일반적인 코드북에 비하여 평균 1.43배로 향상된다. 따라서 양자화 오차도 줄어든다.

$$X_{real}[n] = X_{norm}[n] \times (R_{max} - R_{min}) + R_{min}, \quad (3)$$

3.2. 정규화 코드북 설계

본 논문에서 제안하는 적응적 벡터 양자화기에 사용되는 정규화 코드북을 만드는 방법은 다음과 같다.

$$X_{norm}[n] = \frac{X_{real}[n] - R_{min}}{R_{max} - R_{min}}. \quad (4)$$

여기서 n은 벡터 차원을 나타낸다. 식 (4)은 식 (3)으로 구할 수 있다. 실제적인 음성데이터에서 LPC를 추출하고 ISF로 변환하여 충분한 학습데이터를 만든다. 그림 (1)에서의 제 1서브벡터와 제 3서브벡터와 같은 일반적인 벡터 양자화에 사용되는 코드북은 학습데이터를 이용하여 바로 LBG 알고리즘으로 벡터 양자화 코드북을 설계한다. 그림 (1)의 제 2서브벡터와 같이 적응적 양자화기에 사용

하는 정규화 코드북은 같은 프레임에서 양자화 된 R_{max} 와 R_{min} 을 이용하여 식 (4)을 통하여 사용될 학습데이터를 0부터 1사의 원소 값으로 적응적 양자화 코드북의 분포특징에 맞게 정규화하고 LBG 알고리즘을 통하여 정규화 코드북을 만든다.

3.3. 적응적 ISF 벡터 양자화기 설계

일반적인 벡터양자화는 MSVQ를 이용하여 코드북 크기를 줄이고 계산량을 감소한다. 제안하는 광대역 적응적 ISF 양자화기는 벡터의 순서화 특성을 최대한 이용하기 위하여 16차 ISF 벡터를 그림 (3)과 같이 4개의 4차원 서브벡터로 분할하여 제 2서브벡터와 제 3서브벡터를 적응적 양자화 기법으로 양자화 한다.

우선 제 1서브벡터와 제 4서브벡터를 벡터 양자화 하는데 각각 4차 벡터의 네 개 원소가 조합하여 가장 작은 MSE를 갖는 인덱스가 선택된다. 이에 따라 양자화 된 f_2, f_7 과 f_{12} 를 알게 되는데 f_2 와 f_7 은 그림 (4)와 같이 제2서브벡터 f_3, f_4, f_5, f_6 의 분포범위 즉 R_{min} 과 R_{max} 에 해당하기 때문에 정규화 코드북 X_{norm} 을 통하여 제 2서브벡터의 적응적 코드북 X_{real} 을 만들 수 있다. 또한

제 3서브벡터는 f_8, f_9, f_{10}, f_{11} 인데 f_7 과 f_{12} 에 의하여 분포범위가 제한되며 제 3서브벡터에 대응하는 정규화 코드북으로 실제 코드북이 만들어진다. 제안하는 적응적 양자화 기법은 실시간 음성코덱에서 프레임 별 이러한 방식으로 제 2서브벡터와 제 3서브벡터를 적응적 벡터 양자화한다.

제안하는 양자화 기법은 코드북을 프레임마다 변화시켜주기 때문에 매우 많은 계산량을 요구할 거라고 추측한

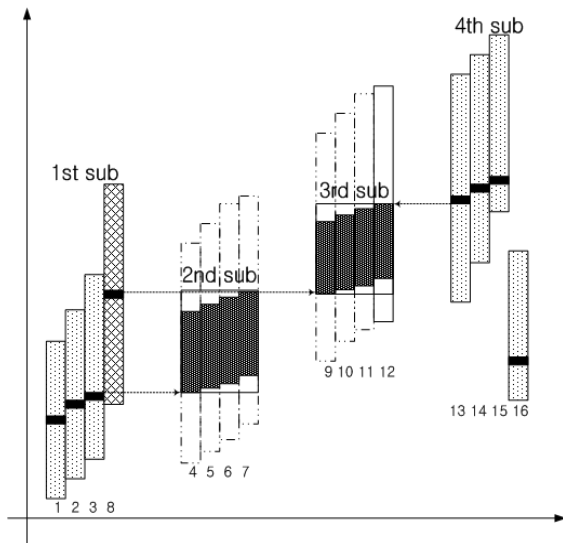


그림 4. 16차 ISF양자화기의 벡터 분포 제한
Fig. 4. Distribution Constriction for 16th order ISF VQ.

다. 그러나 실제 양자화기에서는 코드북을 변화시키는 대신 코드벡터를 R_{min} 과 R_{max} 를 이용하여 정규화 코드북 범위로 변화하여 양자화한다. 이러한 방법으로 하나의 서브벡터 양자화에서 수많은 데이터를 가진 코드북을 변화시킬 대신 단 한번만 양자화벡터를 정규화하기 때문에 일반적인 벡터양자화와 똑같은 계산량을 갖게 된다. 또한 제안하는 ISF 양자화기는 AMR-WB와 같은 비트전송률 46 bit에서의 성능을 평가하기위하여 멀티스테이지 양자화 기법을 사용하여 양자화 오차를 다시 한번 양자화하였다. 프레임 예측기법을 추가하여 G, 718 코덱에 사용한 Safety-net [9] 구조로 예측이 없는 부분과 예측이 있는 두 부분으로 나뉘어 각각 정규화 코드북을 사용하였다. 제안하는 양자화기는 그림 (5)와 같다.

상단은 예측기법을 추가하지 않은 양자화기이고 하단은 1차 AR (Auto-Regressive) 예측 기법을 추가한 양자화기이다. 예측기법을 추가한 양자화기는 양자화 성능이 우수하지만 ISF 계수의 순서화 특징이 사라진다. 적응적 양자화기법을 적용하기 위하여 적응적 양자화기법을 사용하지 않은 제1서브벡터와 제4서브벡터만 예측구조로 양자화 하였다. 또한 예측기법을 이용한 벡터양자화기는 채널 에러에 민감하여 한 번의 채널에러가 발생하면 에러가 누적되어 LPC 복원에 엄청난 영향을 미치게 된다. 이러한 현상에 대비하여 제안한 양자화기는 예측기법을 사용한 양자화기와 예측기법을 사용하지 않은 양자화기로

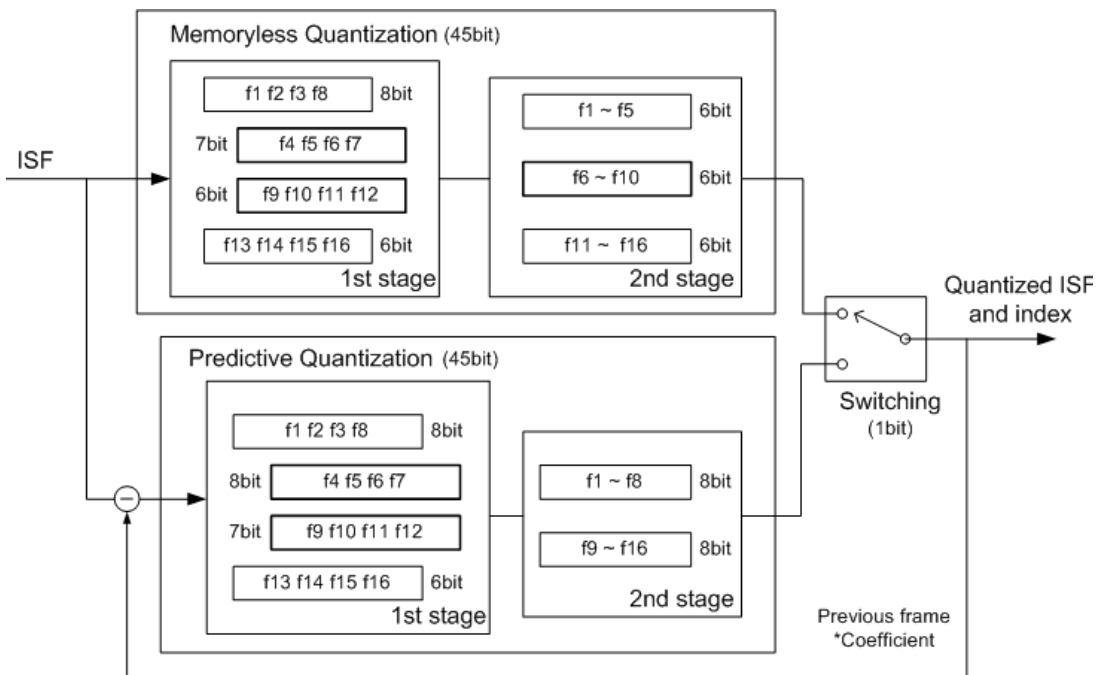


그림 5. 프레임 예측기법을 사용한 ISF 벡터 양자화기
Fig. 5. ISF quantizer with prediction.

두 번 양자화 하는 방법을 사용하였으며, 따라서 계산량은 일반적인 벡터양자화의 두 배정도 된다.

IV. 제안된 적응적 ISF 벡터 양자화기의 성능평가

본 장에서는 제안된 알고리즘의 성능을 III장에서 설계된 양자화기 AMR-WB와 성능을 비교하였다. 프레임 성능 평가의 정량적 지표로는 주파수 왜곡 (Spectral Distortion) [3]을 사용한다. 주파수 왜곡은 다음 식 (5)와 같이 정의된다.

$$SD(dB) = \frac{1}{NF} \sum_{n=1}^{NF} \left(\frac{1}{\pi} \int_0^{\pi} [10 \log_{10} |A_n(e^{j\omega})|^2 - 10 \log_{10} |\hat{A}_n(e^{j\omega})|^2]^2 d\omega \right)^{\frac{1}{2}} \quad (5)$$

여기서 NF 는 프레임 수이고 $A_n(e^{j\omega})$ 및 $\hat{A}_n(e^{j\omega})$ 는 양자화 하기전의 LPC 계수와 양자화 된 LPC 계수의 스펙트럴 값이다. 설계된 광대역 LPC 양자화기에서 투명한 성능이란 평균 SD가 1 dB 이하이고 2-4 dB의 스펙트럴 왜곡을 갖는 프레임이 2% 이하이며 4 dB 이상의 스펙트럴 왜곡을 갖는 프레임이 없어야 된다.

성능측면에서 본 논문에서는 실험을 위하여 여성 화자 60명 남성 화자 58명의 5분 43초의 NTT 음성데이터를

사용하였다. 표 (1)은 III장에서 기술된 구조로 테스트한 실험결과이다.

제안하는 기법의 성능을 평가하기 위하여 제안된 양자화기와 같은 분할벡터구조를 가진 일반적인 양자화를 설계하였다. 일반적인 양자화기는 프레임 당 46 bit에서 0.9 dB의 평균 스펙트럴 왜곡이 일어났고 2-4 dB의 왜곡을 갖는 프레임은 1.58% 정도였다. 제안하는 정규화 코드북기법을 추가하여 평균 스펙트럴 왜곡은 0.09 dB 개선되었고 2-4 dB 왜곡을 갖는 프레임도 반으로 줄었다. 또한 프레임예측기법을 추가하여 그림 (5)와 같이 Safety-net 구조로 양자화기를 설계하여 성능평가 하였다. 표준 코덱인 AMR-WB는 46 bit 프레임당 전송률에서 0.77 dB의 평균 스펙트럴 왜곡을 가졌고 2-4 dB 왜곡을 갖는 프레임이 0.28% 존재하였다. 제안하는 방법은 같은 비트 전송률에서 0.73 dB의 평균 스펙트럴 왜곡으로 0.04 dB의 성능향상을 가져왔는데 프레임 당 약 1.5 bit정도가 절약된다. PESQ 성능상 제안하는 46 bit 적응적양자화기를 AMR-WB에 적용하였을 때 4.0710의 PESQ 성능이 4.0851로 향상되는 것을 보여주었으며 같은 기법으로 44 bit의 적응적 양자화기를 적용하였을 때 4.0699의 PESQ 성능으로 AMR-WB의 성능에 접근하였다.

계산량 측면에서는 제안하는 기법은 III장 2절에서 서술한 코드북 대신 순서화 특징을 이용하여 코드벡터를

표 1. 적응적 양자화기의 성능 비교
Table 1. Performance of adaptive quantization method.

	스펙트럴 왜곡 Spectrum Distortion	평균 왜곡 (dB)	Outlier (in %)		PESQ
			2-4 dB	>4 dB	
프레임 예측이 없음	일반적인 벡터양자화기	0.8996	1.582	0.000	4.0327
	정규화 코드북을 사용한 양자화기 (46 bit)	0.8099	0.838	0.000	4.0603
프레임 예측이 있음	AMR-WB	0.7699	0.279	0.000	4.0710
	정규화 코드북을 사용한 양자화기 (46 bit)	0.7290	0.248	0.000	4.0851
	정규화 코드북을 사용한 양자화기 (45 bit)	0.7483	0.279	0.000	4.0769
	정규화 코드북을 사용한 양자화기 (44 bit)	0.7796	0.51	0.000	4.0699

표 2. 적응적 양자화기의 계산량 비교표
Table 2. Complexity of adaptive quantization method.

	양자화기의 계산량 Complexity of quantizer	프레임 당 코드북 원소 검색회수		
		일반구조	예측구조	전체
프레임 예측이 없음	일반적인 벡터양자화기	6656	/	6656
	정규화 코드북을 사용한 양자화기 (46 bit)	3456	/	3456
프레임 예측이 있음	AMR-WB	/	5376	5376
	정규화 코드북을 사용한 양자화기 (46 bit)	3072	6912	9984
	정규화 코드북을 사용한 양자화기 (45 bit)	3072	6656	9728
	정규화 코드북을 사용한 양자화기 (44 bit)	2912	6656	9568

0과 1사이의 값으로 정규화하여 양자화하는 방법을 적용하였다. 따라서 제안하는 양자화기법의 계산량은 일반적인 벡터양자화기법과 유사하며 코드북 검색차수와 밀접한 관계를 갖는다. 제안하는 Safety-net 구조의 양자화기는 적응적 양자화기와 AR예측구조 양자화기, 크게 두 개의 블록으로 나뉘게 되며, 따라서 전체적인 계산량은 일반적인 벡터양자화기의 2배 정도가 된다. 계산량 비교 관점에서 제안하는 양자화기 및 AMR-WB에 사용된 양자화기의 프레임 당 코드북 원소 검색수를 계산하였다. 계산량 감소기법이 추가되지 않은 일반적인 벡터 양자화기에서 양자화에 사용되는 계산량은 코드북 인덱스 검색차수와 밀접한 관계를 가진다. 코드북 인덱스가 많고 차원이 높으면, 즉 코드북 원소가 많으면 사용되는 계산량이 많고, 반대로 코드북 원소가 적으면 계산량이 적다. 표 (2)에서 제안하는 양자화기와 AMR-WB 양자화기의 각 서브벡터 코드북 원소 검색차수 총 합이다.

V. 결론

본 논문에서는 ISF 계수의 순서화 특성을 이용하여 선형예측계수 양자화를 위한 ISF 분할구조를 가진 양자화기의 성능개선 알고리즘에 관하여 연구하였다. 제안하는 알고리즘은 ISF 계수의 순서화 특성을 이용하여 정규화 코드북을 만든다. 그리고 양자화 된 분할 벡터로 정규화 코드북의 분포범위를 적응적으로 확장시켜줌으로 양자화 하여야 할 분할 벡터의 코드북을 효율적으로 만들어 준다. 제안한 기법은 AMR-WB 부호화기에 적용하여 46 bit 벡터양자화기에서 일반적인 벡터양자화기보다 0.9 dB 정도의 성능향상을 얻었다. 또 프레임 예측기법이 사용된 AMR-WB와의 성능 비교를 위하여 프레임 예측양자화기법을 적용하였고 프레임 손실에 어느 정도 강인하게하기 위하여 Safety-net 기법을 사용하였다. 제안된 알고리즘은 표준화코덱인 AMR-WB의 ISF 양자화와 같은 계산량에서 0.04 dB, 약 프레임 당 1.5 bit 정도의 성능 개선을 얻었다.

본 논문에서 제안한 알고리즘은 AMR-WB와 같은 간

단한 구조의 실시간 음성통신환경의 ISF 및 LSF 벡터양자화기의 성능개선에 활용 될 수 있으리라 사료된다.

감사의 글

이 논문은 2010년도 충북대학교 학술연구지원사업의 연구비 지원에 의하여 연구되었습니다.

참고 문헌

- O'Shaughnessy, D., "Linear predictive coding", *IEEE Potentials* vol. 7, pp. 29-32, Feb. 1988.
- Makhoul J. (1975) Linear Prediction: A Tutorial Review, *Proceedings of the IEEE*, vol. 63, pp. 561-579.
- A.M.Kondoz, "Digital Speech", pp. 106-107.
- F. Itakura, "Line spectrum representation of linear predictive coefficients of speech signals," *J.Acoust. Soc. Am*, vol. 57, no. 535(A), 1975.
- Y. Bistriz and S. Peller, "Immittance spectral pairs (ISP) for speech encoding", in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 1993.
- ITU-T Rec. G.729, "Coding of Speech at 8 kbit/s using Conjugate Structure Algebraic Code Excited Linear Prediction (CS-ACELP)," March 1996.
- ITU-T Rec. G.729.1, "An 8-32 kbit/s scalable wideband coder bitstream interoperable with G.729," May 2006.
- B. Bessette, R. Salami, R. Lefebvre, M. Jelinek, J. Rotola-Pukkila, J. Vainio, H. Mikkola, K. Jarvinen, "The adaptive multirate wideband speech codec (AMR-WB)," *IEEE Trans. Speech Audio Processing*, vol. 10, no. 8, Nov. 2002.
- ITU-T Rec. G.718 "Frame error robust narrowband and wideband embedded variable bit-rate coding of speech and audio from 8-32 kb/s," Jun. 2008.
- Paliwal K.K and Atal B.S., "Efficient vector quantization of LPC parameters at 24 bit/frame", *IEEE Trans.* vol. 1, pp. 3-14, Jan. 1993.
- Linder, Y, Buzo, A, and Gray, R., "An algorithm for vector quantizer design", *IEEE Trans.* vol. 28, pp. 84-95, Jan. 1980.
- Moo Young Kim, Nam Kyu Ha, and Sang Ryong Ki, "Linked split-vector quantizer of LPC parameters," *Proc. ICASSP* 1996.
- Sangwon Kang, Changyong Son, Hosang Sung, "A fast-search algorithm for the predictive split VQ of LPC parameters", *IEEE Communications Letters*, vol. 5, no. 5, pp. 212-214, May 2001.
- Jongha Lim, Gyu-Hyeok Jeong, Gibong Hong and Insung Lee, "적응적인 확장된 코드북을 이용한 분할 벡터 양자화기 구조의 ISF 양자화기 개선" *한국음향학회지* 30권, 1호.

저자 약력

●박 지 강 (Zhigang Piao)



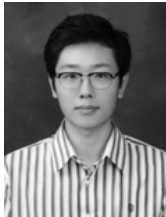
2009년: 중국 길림성 연변대학교 전자정보공학과 (학사)
2009년 ~ 현재: 충북대학교 전파통신공학과 (석사)
※ 주관심 분야: 음성 / 오디오 신호처리

●임 중 하 (Jongha Lim)



2009년: 충북대학교 전파통신공학과 (학사)
2011년: 충북대학교 전파통신공학과 (석사)
2011년 ~ 현재: 삼성전자
※ 주관심 분야: 음성 / 오디오 신호처리

●홍 기 봉 (Gibong Hong)



2010년: 충북대학교 정보통신학과 (학사)
2010년 ~ 현재: 충북대학교 전파통신공학과 (석사)
※ 주관심 분야: 음성 / 오디오 신호처리

●이 인 성 (Insung Lee)



1983년: 연세대학교 전자공학과 (학사)
1985년: 연세대학교 전자공학과 (석사)
1992년: Texas A&M University 전기공학과 (박사)
1993년 ~ 1995년: 한국전자 통신연구원 이동통신
기술연구단 선임연구원
1995년 ~ 현재: 충북대학교 전파통신공학과 교수
※ 주관심 분야: 음성/영상 신호압축, 이동통신, 적
응적 필터