

블록 제한 트렐리스 부호화 양자화 기법을 이용한 협대역 음성 부호화기용 LPC 계수 양자화기 설계

정회원 전자경*, 준회원 박상국*, 종신회원 강상원*

Designing a Quantizer of LPC Parameters for the Narrowband Speech Coder using Block-Constrained Trellis Coded Quantization

Jakyoung Jun* *Regular Member*, Sangkuk Park* *Associate Member*
Sangwon Kang* *Lifelong Member*

요 약

본 논문에서는 기존의 트렐리스 부호화 양자화 기법을 이용, 변형하여 저 복잡도 블록 제한 격자 부호화 양자화 기법 (Block-Constrained Trellis Coded Quantization, 이하 BC-TCQ)을 제안하고, 이를 이용한 협대역 음성 부호화기용 예측 BC-TCQ를 설계하였다. 트렐리스 부호화 양자화 기법은 일종의 벡터 양자화 방식으로 부호화에 요구되는 벡터 코드북을 트렐리스 구조에 기반한 스칼라 코드북으로 구성함으로써 VQ와 비교 할 만한 성능을 보일 뿐 아니라 복잡도가 훨씬 작은 특성을 보인다. 본 논문에서 제안한 예측 BC-TCQ는 프레임당 26비트에서 IS-641 음성 부호화기보다 평균 SD가 0.4107dB 향상되었으며, 더하기 연산이 64.54%, 곱하기 연산이 76.93%, 비교 연산이 2.35% 감소하였다.

Key Words : LSF계수, BC-TCQ, 벡터 양자화, 음성 코딩

ABSTRACT

In this paper, low complexity block constrained trellis coded quantization (BC-TCQ) structures are introduced, and a predictive BC TCQ encoding method is developed for quantization of line spectrum frequencies (LSF) parameters for narrowband speech coding applications. Trellis-coded quantization(TCQ) is a form of VQ that builds the VQ codebook from interleaved constituent scalar quantization codebooks. The performance is compared to the other VQ, demonstrating reduction in spectral distortion and significant reduction in encoding complexity. The predictive BC-TCQ is about 0.47107 dB superior to the IS-641 split-VQ, 26bits/frame, in spectral distortion sense. The BC-TCQ is 64.54%, 76.93%, 2.35% of the IS-641 split-VQ, respectively, in the complexity of the additions, multiplies, comparisons.

I. 서 론

선형 예측은 보통 낮은 비트율 음성 코딩 시스템에서 사용된다. 선형 예측 계수는 효과적인 양자화와 코딩을 위해 LSF (Line Spectral Frequency) 계수^[1]로 표현된다. LSF 계수의 양자화 방법으로 크

게 스칼라 양자화 기법^[2]과 벡터 양자화 기법^{[3],[4]}을 바탕으로 변형된 여러 기법들이 소개 되었다. 스칼라 양자화 기법은 각각의 LSF 요소 값을 양자화하므로 성능에 비해 많은 비트(40비트)를 사용하는 단점이 있다. 벡터 양자화 기법은 스칼라 양자화 기법보다 훨씬 좋은 양자화 성능을 보이나, 복잡도가

* 한양대학교 전자전기 제어계측공학과 † 교신저자: 강상원

논문번호 : KICS2007-01-011, 접수일자 : 2007년 1월 10일, 최종논문접수일자 : 2007년 3월 14일

상당히 큰 문제점을 가지고 있다.

트렐리스 부호화 양자화(Trellis coded quantization)^[5]는 일종의 벡터 양자화 방식으로, 부호화에 사용되는 벡터 코드북내 각 요소값을 스칼라 코드북으로 구성하며, 길쌈 부호화기(convolutional coder)로 트렐리스 구조를 표현한다. Viterbi 알고리즘^[6]을 이용하여 최적 인코딩을 위한 트렐리스 경로를 찾으며, 트렐리스 부호화 양자화 기법은 복잡도가 unstructured VQ에 비하여 훨씬 작은 특성을 보인다. 전형적인 트렐리스 부호화 양자화 기법은 Viterbi 알고리즘 탐색 후 결정된 트렐리스 경로의 초기 스테이트를 추가 정보로서 전송하여야 한다. 이러한 부가 전송 정보는 차수가 큰 소스 벡터에 대해서는 큰 영향을 끼치지 않으나, 차수가 작은 소스에 대해서는 rate distortion 관점에서 상당히 중요한 결림돌로 작용한다. 하지만 Nikneshan과 Khandani가 “tail biting”^[7]방식을 소개함으로써 TCQ의 부가 정보 전송에 대한 단점은 해결되었다. 이들이 제안한 “tail-biting” TCQ는 차수가 작은 소스 벡터에 대하여 VQ와 비교할 만한 성능을 보인다.

본 논문에서는 tail-biting 개념을 변형하여, 10차 벡터 소스를 위한 Block-Constrained TCQ(BC-TCQ)를 개발하였으며, 이를 이용하여 LSF 계수 양자화를 위한 예측 BC-TCQ를 개발하였다. 예측 BC-TCQ의 성능은 협대역 음성을 이용해서 split-VQ^[8]와 비교 평가하였다.

본 논문의 구성은 서론에 이어, 2장에서는 BC-TCQ 개념을 소개하였고, 3장에서는 예측 BC-TCQ를 이용한 LSF 계수 양자화기에 대해 설명하였으며, 기존의 양자화기와 비교한 실험 결과를 4장에 나타내었다.

II. BC-TCQ

2.1 구조 및 복잡도

본 논문에서 사용하는 트렐리스 부호화 양자화기는 rate 1/2 길쌈 부호화기로 구성된 트렐리스 구조를 가지며, 피드백이 없는 인코더 구조를 가진다. 전체 트렐리스 구조의 스테이트 수는 $N=2^n$ 개이며, 각 트렐리스 스테이트로 들어오는 혹은 나가는 branch 갯수는 2개이다.

그림 1은 4 스테이트 트렐리스 구조와 길쌈부호화기(convolutional coder)를 나타낸 것이며, 각 branch에는 길쌈 부호화기의 출력 값이 할당된다. 모든 트렐리스 스테이트를 초기 스테이트로 선택하는 트렐리스 부호화 양자화 기법은 모든 초기 경로

메트릭을 '0' 혹은 같은 값으로 초기화 한 후, viterbi 인코딩을 마지막 스테이지까지 실행함으로써 최소 distortion을 갖는 트렐리스 경로를 결정할 수 있다. 이러한 인코딩으로 N 개의 트렐리스 스테이트

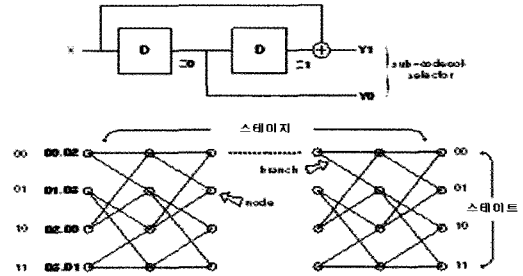


그림 1. 길쌈 부호화기 및 4 스테이트 트렐리스 구조

에서 시작하여, N 개의 트렐리스 스테이트로 끝나는 트렐리스 경로 중 최적의 경로가 결정된다. n 개의 샘플로 구성된 블록 소스를 이러한 트렐리스 부호화 양자화 기법을 이용하여 양자화하는 경우, 트렐리스 경로 정보를 전송하기 위하여 트렐리스 경로의 초기 스테이트를 전송하기 위한 v 비트와 전체 n 스테이지에 대한 천이 정보를 전송하기 위한 n 비트가 필요하므로, 전체 $n+v$ 비트가 요구된다.

Nikneshan과 Khandani는 TCQ 방식에 비하여 좋은 성능을 보이는 tail-biting TCQ (TB-TCQ)방식을 제안하였다. TB-TCQ방식은 트렐리스 경로의 초기 스테이트와 마지막 스테이트가 같다는 제약 조건을 갖는다. TB-TCQ 인코딩 과정은 N 트렐리스 스테이트의 각 스테이트를 초기 스테이트로 하는 viterbi 알고리즘 인코딩이 요구됨으로 전체적으로 N 번의 viterbi 알고리즘 인코딩이 필요하며, 각 초기 스테이트에서 시작되는 N 개의 트렐리스 경로 중에서 최소 distortion을 나타내는 초기 스테이트와 트렐리스 경로가 결정된다. TB-TCQ에서 각 초기 스테이트에 따른 마지막 스테이트는 제약조건으로 정해져 있으므로 $n-v$ 스테이지에서 마지막 스테이지의 특정 스테이트로의 트렐리스 경로는 결정되므로, 마지막 v 스테이지에 대한 스테이트 천이 정보는 전송 할 필요가 없다. 이러한 TB-TCQ 인코딩 방식의 큰 단점은 viterbi 알고리즘 인코딩을 트렐리스 스테이트 수만큼 실행하여야 하므로 기본 TCQ에 비하여 N 배의 복잡도가 요구된다. Viterbi 알고리즘 탐색의 수가 트렐리스 스테이트의 수와 같으

므로, 이와 같은 인코딩을 본 논문에서는 “최대 계산량” 블록 제한 TCQ라 칭한다.

그림 2에 제시된 TB-TCQ에서 초기 스테이트와 마지막 스테이트가 같다는 제약조건 때문에 마지막 2개 스테이지의 스테이트 천이 정보는 유일하게 결정되므로 전송할 필요가 없다.

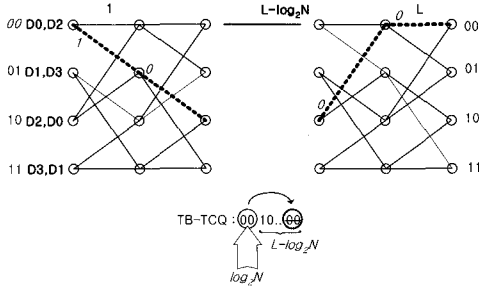


그림 2. TB-TCQ 인코딩 방식에서의 트렐리스 경로 정보

본 논문에서 제안한 block-constrained TCQ (BC-TCQ)는 N 트렐리스 스테이트 중 2^k , ($0 \leq k \leq v$) 개의 초기 스테이트를 가지며, 각각의 초기 스테이트마다 2^{v-k} 개의 마지막 스테이트를 가지도록 제약된 트렐리스 구조를 가진다. 다음 그림 3은 인코딩 과정을 나타낸 것이다.

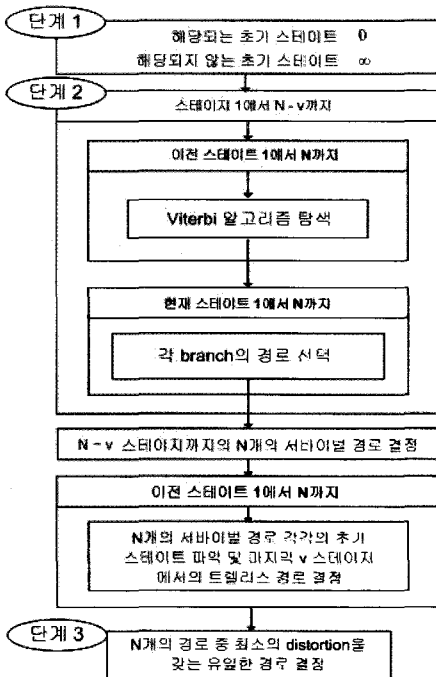


그림 3. BC-TCQ 인코딩 과정 플로우 차트

단계 1) 2^k 개의 초기 스테이트에 대한 경로 메트

릭은 '0'으로 초기화 시키며, 나머지 스테이트의 경로 메트릭은 '∞'로 초기화 시킨다.

단계 2) 한번의 viterbi 알고리즘 탐색이 $n-v$ 스테이지까지 진행되며, viterbi 알고리즘 탐색에 의해 결정된 트렐리스 경로의 초기 스테이트 정보에 따라 마지막 v 스테이지의 트렐리스 경로는 자동적으로 결정된다.

단계 3) 단계 2에서 결정된 트렐리스 경로 정보는 k 비트의 초기 스테이트 정보와 $n-v$ 스테이지까지의 스테이트 천이 정보를 나타내기 위한 $n-v$ 비트와 $n-v$ 스테이지에서 마지막 스테이트의 2^{v-k} 개 스테이트 중 선택된 스테이트를 나타내기 위한 $v-k$ 비트로 구성된다. 이러한 인코딩 방식은 k 값에 관계없이 한번의 Viterbi 알고리즘 탐색으로 이루어진다.

2.2 인코딩 성능

예측 구조를 이용한 LSF 계수 양자화 시, 예측 오차들은 대략 라플라시안 분포를 이룬다^[9]. 따라서, 개발된 BC-TCQ의 성능을 비메모리 라플라시안 소스에 대하여 평가 하였다. Marcellin과 Fischer^[5]는 비메모리 라플라시안 소스를 이용한 fixed-rate TCQ 인코딩에서 2^{2+R} 크기의 “quadrupled” 코드북을 사용함으로써 향상된 양자화 성능을 얻었다. 이 결과를 바탕으로, 본 논문에서는 4개의 코드워드를 갖는 8개의 부 코드북으로 구성되는 샘플당 3 비트의 16 스테이트 BC-TCQ 인코더를 설계하였다. 트렐리스의 부 코드북 할당은^[5]에 제시된 방법과 동일하게 적용하였으며, 4개의 부 코드북은 전체 32 개의 트렐리스 branch 중 28개에 할당하고, 나머지 4개 부 코드북은 나머지 4개의 branch에 각각 하나씩 할당하였다. 코드북은 훈련용 샘플을 이용해서, 트렐리스 코드 설계를 위한 일반적인 Lloyd 알고리즘을 기반으로 최적화하였다.

표 1. 라플라시안 소스에 대한 초기 스테이트 수에 따른 BC-TCQ의 SNR 성능

| 초기 스테이트 수 | Block length | | | |
|-----------|--------------|---------|---------|---------|
| | 8 | 16 | 32 | 64 |
| 1 | 13.6287 | 14.4819 | 15.1030 | 15.5636 |
| 2 | 14.8567 | 15.2100 | 15.5808 | 15.8499 |
| 4 | 14.9591 | 15.4942 | 15.7731 | 15.9887 |
| 8 | 13.4285 | 14.5864 | 15.3346 | 15.7704 |
| 16 | 11.6558 | 13.2499 | 14.4951 | 15.2912 |

표 1은 복잡도를 고려한 초기 및 마지막 스테이트

의 다양한 조합들의 SNR 성능을 나타내었다. 성능이 가장 좋은 설계는 초기 스테이트 s 가 4개, 마지막 스테이트 f 가 4개로 $s = 4i, f = 4i + j, i = 0, 1, 2, 3; j = 0, 1, 2, 3$ 로 구성된 조합이다. 이 설계 값(본 장 1절에서 사용되는 파라미터 k 값이 2에 해당)을 본 논문에서 언급한 모든 BC-TCQ에 사용하였다.

III. 예측 BC-TCQ를 이용한 LSF 계수 양자화기

LSF 계수는 프레임간 및 프레임 내 상관도가 높은 특징을 가지고 있으며, 이러한 성질을 이용하기 위해 본 논문에서는 예측 BC-TCQ 양자화 방식을 제안하였다. LSF 계수 양자화를 위한 예측 BC-TCQ의 설계 과정 및 최적화 과정을 살펴보고, 인코딩 복잡도를 분석하였다.

3.1 예측 BC TCQ의 설계 및 최적화

양자화기에 inter 및 intra 프레임 예측 구조를 도입함으로써 상당히 큰 성능 이득을 얻을 수 있다. Moving-average(MA)필터는 auto-regressive(AR)필터에 비하여 무선 통신과 같이 전달 오류가 많이 발생하는 환경에 강하다. 따라서 LSF 코딩 시스템의 inter 프레임 선형 예측에는 4차(보통 3차에서 5차가 좋은 성능을 얻는다.) MA 필터를 사용하고, MA inter예측 에러 벡터 내 요소들에 대한 intra 프레임 예측은 1차 AR 필터를 사용하였다. 그림 4는 4차 MA inter예측과 1차 AR intra예측 BC-TCQ 양자화기의 구조를 나타낸다.

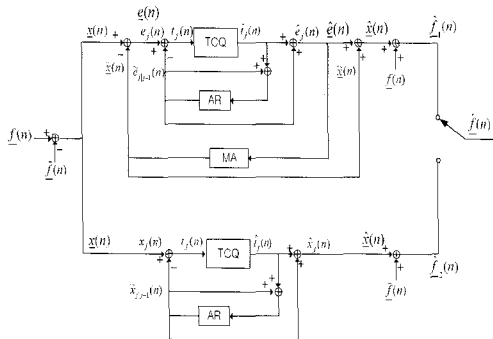


그림 4. Safety net 구조의 예측 BC-TCQ

MA 필터 예측 양자화기의 최적 예측 계수 선택과 선택된 예측 계수에 대한 최적 코드북 설계를 위해 ‘공동 최적화’ 설계 방식^[10]을 이용 하였으며

그 과정은 다음과 같다. 예측 계수와 코드북의 설계 과정에서 예측 계수와 코드북을 구성 하는 예측 에러 벡터의 스케일이 DC제거 LSF 계수 값에 관계 없이 교환되는 현상이 발생한다. 이러한 현상을 막기 위하여 예측 계수에 대해 다음 식 (1)과 같은 제약 조건을 사용하였다.

$$\sum_{i=0}^Q \mu_i = 1 \quad (1)$$

식 (1)의 조건을 만족하는 모든 경우의 예측 계수에 대해 최적의 BC-TCQ 코드북 설계가 이루어졌다. 선택된 예측 계수에 최적으로 설계된 코드북을 4차 MA inter예측과 1차 AR intra예측 BC-TCQ로 양자화하여 전체 훈련용 샘플에 대한 예측 에러의 에너지 값이 최소가 되는 예측 계수와 최적 코드북을 설계한다. 예측 TCQ^[5]에서와 같이 intra예측 BC-TCQ에서는 트렐리스 구조에 intra 요소간 예측이 포함된다. Intra예측 TCQ 인코딩 구조에서 각 스테이지의 각 트렐리스 노드에서 다음 입력 샘플에 대한 예측이 이루어지며, 이때 예측은 해당 트렐리스 노드까지 결정지어진 서바이벌 경로상의 출력 값을 바탕으로 이루어진다.

급격히 변화 하는 프레임에 대해서는 적절한 예측이 불가능하여 성능이 떨어질 수 밖에 없으며 “outliers”발생의 원인이 된다. 따라서 예측이 적절히 되지 않는 outlier 음성 프레임들의 신호 특성에 최적화된 양자화기를 병렬로 구성하여 전체 성능을 향상 시키도록 설계된 구조를 safety net구조라 하고 그림 4과 같다. 이러한 outlier 프레임들도 프레임내 상관도는 존재함으로 이를 이용하고자 intra예측 코딩 방식을 이용하였다.

Safety net구조를 이루는 예측 BC-TCQ 양자화기의 최적화 설계 과정은 다음과 같다. 먼저, 메모리 기반 양자화기 및 비 메모리 기반 양자화기의 초기 intra예측 계수 및 코드북은 훈련용 샘플을 이용하여 독립구조에서 최적화된 파라미터들을 사용한다. 두 번째 과정은 전체 훈련용 샘플은 safety net 구조의 양자화 과정을 통해 두 개 부분의 훈련용 샘플로 분류되는데, 두 개 양자화기 구조 중 적은 MSE의 결과를 가지는 구조로 분류된다. 다음으로 각 양자화기 구조의 코드북과 intra예측 계수는 분류된 훈련용 샘플을 이용하여 새로 디자인되어 업데이트된다. 네 번째 과정은 새로 업데이트된 코드북과 intra 예측 계수 조건하에서 평균 distortion, $d(k)$

(여기서 k 는 현재 iteration을 나타낸다)와 distortion의 변화율, $Thr = |d(k-1) - d(k)| / |d(k-1)|$ 을 구하여 종료 조건을 만족하면 설계를 멈추고 아니면 두 번째 과정부터 반복한다.

평균 distortion은 전체 반복 과정에서 단조 감소하지 않는다. 하지만 대부분의 iteration에서 평균 distortion은 크게 감소하며, 설계 과정은 종료 조건에 의해 멈춘다. 본 논문에서는 $Thr < 0.0001$ 인 경우 설계 과정이 멈추도록 하였다. 전체 iteration 과정에서 최소 distortion을 갖는 코드북 및 intra예측 계수를 최적의 코드북 및 계수로 선택하였다. 모든 가능한 MA 예측 계수에 대하여 최적의 코드북 및 intra예측 계수를 얻기 위하여 위의 과정이 수행되었으며, 그 중 최소 distortion을 보이는 코드북과 계수를 선택하였다.

제안된 BC-TCQ 양자화 방식은 제약된 트렐리스 경로상에서 단 한번의 viterbi 알고리즘 탐색만을 하게 된다. BC-TCQ 복잡도는 전체 트렐리스 스테이트, 입력 벡터의 차원, 허용된 초기 스테이트의 수 및 각 초기 스테이트에 할당된 마지막 스테이트의 수 등에 좌우된다.

IV. 실험 및 결과

본 개발 양자화기의 예측 BC-TCQ는 quadruple 출력 레벨을 가지는 16 스테이트 트렐리스 구조를 사용한다. LSF 계수 양자화기 설계를 위하여 10분 분량의 8kHz 샘플링 음성 샘플을 사용하였으며 성능 테스트를 위하여 10분 분량의 음성 샘플을 사용하였다. 훈련용 및 테스트용 음성 샘플은 영어 여자 및 남자, 한국어 여자 및 남자의 음성 순으로 반복 구성되어 있다. LSF 양자화기 성능을 평가하기 위해 spectral distortion (SD)을 사용하였다. 본 논문에서는 distance measure로서 weighted Euclidean distance^[8]를 이용하였다. 예측 BC-TCQ 양자화기 성능은 협대역 음성 부호화기인 IS-641의 LPC 계수 양자화기와 비교 평가하였다. 예측 BC-TCQ 양자화기 성능을 IS-641 음성 부호화기의 LPC 계수 양자화기와 균등한 조건에서 비교하기 위하여 LPC 계수 분석의 전처리 과정을 IS-641 음성 부호화기와 똑같이 적용하였다. 예측 BC-TCQ 양자화기의 10차 LSF 계수에 대한 양자화기의 비트 할당은 프레임당 26비트로 표 2와 같다.

표 2. 20ms 프레임의 협대역 음성에 대한 예측 BC-TCQ 양자화기의 비트 할당

| 파라미터 | 비트 (bits) |
|---------------|---|
| 패스 정보 | 2+6+2 |
| 코드워드 정보 | 2 x 5 (stage 1 to 5) 1 x 5 (stage 6 to 10) |
| Safety net 정보 | 1 |
| 총 계 | 26 |

예측 BC-TCQ와 IS 641 음성 부호화기에서 사용되는 split-VQ 양자화기와의 성능 비교를 통해 표 3에서의 같이 본 논문에서 제안하는 양자화기의 SD 성능이 우수함을 알 수 있다. 또한, 비트 별 예측 BC-TCQ의 SD 성능은 표 4와 같고 프레임 당 대략 22비트에서 transparent한 품질이 만족됨을 알 수 있다. 비트가 줄어들에 따른 비트 할당은 표 2에서 높은 주파수의 LSF 계수에 대한 비트 할당을 줄임으로써 수행된다.

표 3. 예측 BC-TCQ와 IS-641의 split-VQ 양자화기의 SD 성능 비교 (프레임당 26비트)

| | Ave SD (dB) | SD Outliers (%) | |
|-----------------|-------------|-----------------|--------|
| | | 2dB ~ 4 dB | > 4 dB |
| | | 예측 BC TCQ | 0.8035 |
| IS 641 split VQ | 1.2142 | 4.0875 | 0 |

표 4. 10차 LSF 계수 예측 BC-TCQ의 비트별 SD 성능

| Bits | Ave SD (dB) | SD Outliers (%) | |
|------|-------------|-----------------|--------|
| | | 2dB ~ 4 dB | > 4 dB |
| | | 20 | 1.1705 |
| 22 | 1.0345 | 2.7730 | 0 |
| 24 | 0.9162 | 1.4475 | 0 |
| 26 | 0.8035 | 0.6278 | 0 |

표 5는 예측 BC-TCQ의 각 스테이지에 비트가 할당된 경우 10차 LSF 계수 양자화에 필요한 연산량을 IS-641의 split-VQ 양자화기와 비교한 표이다 (프레임당 26비트). 예측 BC-TCQ 양자화기는 IS-641보다 더하기 연산이 64.54%, 곱하기 연산이 76.93%, 비교 연산이 2.35% 감소하였다.

표 6 및 표 7은 채널 에러 환경하의 성능 분석을 위해 bit error rate(BER)에 따른 개발 “예측

표 5. 예측 BC-TCQ와 IS-641의 spit-VQ 양자화기와의 복잡도 비교 (프레임당 26비트)

| | Operation | 예측 | IS 641 |
|-------------|-----------|--------|----------|
| | | BC TCQ | split VQ |
| 복잡도 | 가산 연산 | 2632 | 7424 |
| | 승산 연산 | 2008 | 8704 |
| | 비교 연산 | 1247 | 1277 |
| ROM (words) | | 480 | 4352 |

BC-TCQ”의 SD 성능을 IS-641 split-VQ 양자화기의 BER 성능과 비교한 표로서 예측 BC-TCQ가 IS-641 split-VQ보다 채널 에러에도 강함을 알 수 있다.

표 6. BER에 따른 협대역용 예측 BC-TCQ (26비트)의 SD 성능

| BER (%) | SD | | |
|---------|-------------|--------------|--------|
| | Ave SD (dB) | Outliers (%) | |
| | | 2dB ~ 4dB | > 4dB |
| 0 | 0.8035 | 0.6278 | 0 |
| 0.01 | 0.8104 | 0.7938 | 0.0026 |
| 0.1 | 0.8719 | 2.1167 | 0.4047 |
| 0.5 | 1.0996 | 7.2503 | 1.6627 |
| 1 | 1.3732 | 13.2555 | 3.7432 |
| 2 | 1.8601 | 24.4747 | 7.9715 |

표 7. BER에 따른 IS-641 LPC 계수 양자화기 (26비트)의 SD 성능

| BER (%) | SD | | |
|---------|-------------|--------------|---------|
| | Ave SD (dB) | Outliers (%) | |
| | | 2dB ~ 4dB | > 4dB |
| 0 | 1.2142 | 4.0875 | 0 |
| 0.01 | 1.2266 | 4.3231 | 0.1662 |
| 0.1 | 1.2993 | 5.9859 | 0.9838 |
| 0.5 | 1.5792 | 11.6669 | 4.6556 |
| 1 | 1.9560 | 18.5257 | 9.6439 |
| 2 | 2.5880 | 27.3105 | 19.0106 |

V. 결론

Tail-biting TCQ (TB-TCQ)는 각 블록의 첫 번째 스테이지의 스테이트와 마지막 스테이지의 스테이트가 같다는 제약조건을 도입함으로써 초기 스테이트 정보 비트를 절약하여 낮은 dimension의 소스 벡터에 대한 양자화에도 VQ 와 비교할 만한 성능을 제공한다. 본 논문에서는 TB-TCQ의 복잡도를 크게 개선하고 TB-TCQ에 비해 성능 감쇄가

없는 block-constrained TCQ (BC-TCQ)를 개발하였다. 개발된 BC-TCQ를 음성 부호화기용 LPC 계수 양자화기에 적용하기 위해 예측 에러를 BC-TCQ로 양자화하는 예측 BC-TCQ 인코딩 방식을 설계하였다. 예측 BC-TCQ 방식의 성능 평가를 위해 협대역 표준 음성 부호화기 IS-641의 LPC 계수 양자화기와 비교하였고, 스펙트럼 왜곡(Spectral Distortion) 및 부호화 복잡도에서 향상된 성능을 보였다. 예측 BC-TCQ 양자화기는 협대역 표준 음성 부호화기에서 프레임당 대략 22비트에서 transparent한 성능을 나타내고 프레임당 26 비트에서 협대역 표준 음성 부호화기 IS-641의 20% 복잡도만을 가진다. 또한, 예측 BC-TCQ 양자화기는 IS-641보다 채널 에러에 강하다.

감사의 글

본 연구는 2006년도 한국과학재단의 특정기초연구사업 (과제번호: R01-2006-000-10885-0)의 지원으로 수행되었습니다.

참고 문헌

- [1] F. Itakura, "Line spectral representation of linear predictive coefficients of speech signals," J. Acoust. Soc. Amer., vol. 57, Supplement no. 1, S35, 1975.
- [2] N. Sugamura and N. Farvardin, "Quantizer design in LSP speech analysis synthesis," IEEE J. Sel. Areas Commun., SAC 6, pp. 432-440, 1988.
- [3] K. K. Paliwal and B. S. Atal, "Efficient vector quantization of LPC parameter at 24 bits/frame," IEEE Trans. Speech Audio Processing, vol. 1, no. 1, pp. 3-14, Jan. 1993.
- [4] A. Gersho and R. M. Gray, Vector Quantization and Signal Compression. Norwell, MA: Kluwer, 1992.
- [5] M. W. Marcellin and T. R. Fischer, "Trellis coded quantization of memoryless and GaussMarkov sources," IEEE Trans. Communications, vol. 38, issue 1, pp. 82-93, Jan. 1990.
- [6] G. D. Forney, Jr., "The Viterbi algorithm,"

Proc. IEEE (Invited Paper), vol. 61, pp. 268
278, Mar. 1973.

- [7] S. Nikneshan and A. K. Khandani, "Soft decision decoding of fixed rate entropy coded trellis coded quantizer over a noisy channel," IEEE Trans. Vehicular Tech., to appear.
- [8] TIA/EIA/IS641, "TDMA cellular/PCS radio interface enhanced full rate speech codec," May 1996.
- [9] S. Chi, S. Kang, and C. Lee, "Safety net pyramid VQ of LSF parameters for wideband speech codecs," Electronics Letters, vol. 37, no. 11, pp. 724-726, May 2001.
- [10] Y. Shoham, "Coding the line spectral frequencies by jointly optimized MA prediction and vector quantization," IEEE Worksho. Speech Coding Proceedings, pp. 46-48, 1999.

전 자 경 (Jakyoung Jun)

정회원



2000년~2004년 : 한양대학교 전
자컴퓨터공학부 (공학사)
2004년~2006년 : 한양대학교 전
자전기제어계측공학과(공학석
사)
2007년~현재 : 삼성전자 차세대
단말팀

박 상 국 (Sangkuk Park)

준회원



1996년~2001년 : 한양대학교 전
자컴퓨터공학부 (공학사)
2005년~2007년 : 한양대학교 전
자전기제어계측공학과(공학석
사)

강 상 원 (Sangwon Kang)

종신회원



1976년~1980년 : 한양대학교 전
자공학과 (공학사)
1980년~1982년 : 서울대학교 전
자공학과 (공학석사)
1985년~1990년 : Texas A&M 대
학교 전기공학과 (공학박사)
1982년~1994년2월 : 한국전자통

신연구원 신호처리 연구실

1994년 3월~현재: 한양대학교 전자컴퓨터공학부, 교수
<관심분야> 음성/오디오 부호화, 신호처리, 이동통신