

부대역 웨이팅 및 비트할당 알고리즘을 수정한 DSBC 음성 부호화기의 성능개선

성회원 김 선 영* 성회원 金 在 功**

Performance Improvement of DSBC Speech Coder by Subband Weighting and a Modified Bit Allocation Algorithm

Sun Young KIM*, Jae Kong KIM** *Regular Members*

要 約 DSBC 음성 부호화기의 성능 개선에 관한 두 방법을 제안하였다. 첫째는 계산량이 많은 종래의 비트할당을 수정함으로써 계산량을 줄일 수 있는 방법이고 둘째는 비전송 대역 재생시 백색잡음 주입으로 인한 히싱 문제를 제거하기 위한 부대역 웨이팅 방법이다. 시뮬레이션 결과 검토된 방법은 음성 출력의 성능 향상에 응용할 수 있음을 나타내었다.

ABSTRACT For the performance improvement in DSBC speech coder two possibilities are proposed. To reduce computational complexity the conventional dynamic bit allocation algorithm are modified. The subband weighting is also presented to avoid hissing noise effect when Gaussian noises are inserted in the regeneration of empty band. The simulation demonstrates that the discussed techniques may suitable for the performance enhancement at the speech output.

I. 서 론

SBC(Sub-Band Coding) 음성 부호화방식은 고음질 가청 신호의 전송 및 이동 통신용 음성 부호화기에 이용되고 있다^{1,2)}. 원래의 SBC 방법은 음성의 평균 스펙트럼에 근거하여 각 부대역

(Subband) 별로 고정 비트를 할당하기 때문에 비트할당은 상대적으로 저주파 부분에 치우쳐질 뿐만 아니라 전송 속도의 감축도 비교적 어려웠다. Ramstad(RA)는 종래의 SBC 방식에 비해 음질저하 없이 전송 속도를 낮출 수 있는 DSBC(Dynamic Bit Allocation SBC) 부호화 방법을 제안한 바 있으나^{3,4)} 이 방법에서는 원 음성신호를 몇개의 부대역으로 나눈 다음 이들 부대역의 에너지 크기 순으로 일부대역의 신호만을 전송하

*韓國電子通信研究所
ETRI

**東國大學校 電子工學科
Dept. of Electronics Engineering Dongguk University
論文番號 : 90-94(接受1990. 10. 6)

고 전송하지 않는 부대역 신호는 수신측에서 전송하지 않는 부대역의 RMS(Root Mean Square) 에너지에 해당하는 백색 잡음으로 대체함으로써 주파수 갭(Gap)을 메꾸어 신호를 재생한다. 이 방법은 낮은 전송속도에도 불구하고 고음질을 얻을 수 있다는 장점이 있는 반면 규정된 전체 비트가 각 부대역에 할당될 때까지는 계속적인 에너지 크기의 비교와 아울러 스케일이 이루어져야 하므로 비교적 많은 계산량을 필요로 하고 있다. 또 다른 문제점은 백색잡음 주입에 나타나는 히싱잡음(Hissing Noise) 영향이다³⁻⁶⁾. 특히 백색잡음 주입을 H/W로 구현할 때 백색잡음을 저장한 다음 음성재생 때 마다 매번 이용해야 하므로 히싱 현상이 심하게 된다.

이같은 단점을 개선코자 본 논문에서는 기존의 반복 비트할당 방법을 평균 비트할당으로 수정함으로써 종래의 DSBC 방식에 비해 계산량을

줄이고, 간단한 부대역 웨이팅만으로 히싱현상을 개선함으로써 음질 개선의 가능성을 제시하고자 한다. II 장에서는 DSBC 음성 부호화 방식을 간단히 살펴보고 III 장에서는 제안된 방식에 의거하여 성능개선의 원리를 설명하고 IV 장에서는 시뮬레이션을 그리고 결론을 V 장에 나타내었다.

II. DSBC 음성 부호화기

먼저 DSBC 음성 부호화기의 개요를 간략한다.

그림 1을 보면 비트할당에 의한 DSBC 음성 부호화기로서 대역제한된 입력 $x(n)$ 으로 부터 재생된 신호 $x'(n)$ 을 최종 출력시킨다. 이 방식을 고정 비트할당 SBC와 비교하면 송신수단에서 QMF 필터는 동일하나 프레임링 에너지 크기에

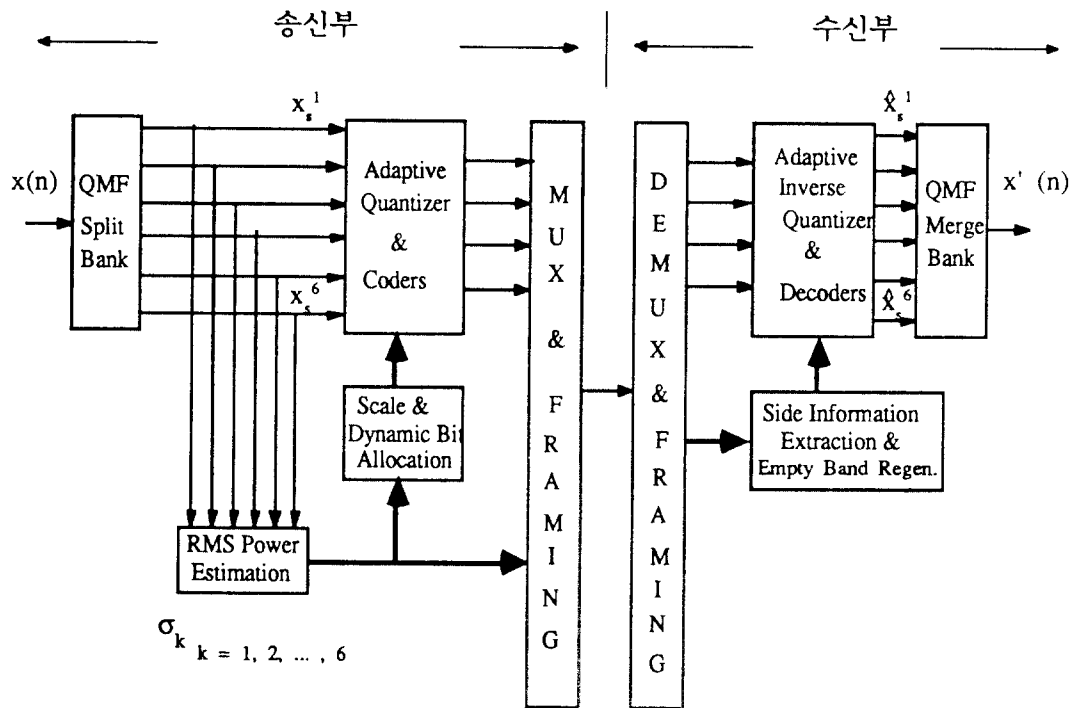


그림 1. DSBC 음성 부호화기 블럭도
Block Diagram of DSBC Speech Coder

따라 부대역 x_s^i , ($i=1, \dots, 6$)에 할당되는 비트 수가 가변이고 x_s^i 의 양자화에 $N(0, 1)$ 특성을 지닌 비선형 Max 양자화기를 이용하고, 보조정보로 RMS 에너지를 전송해야 하는 것들이 차이점이라 할 수 있다.

프레임 단위의 음성신호 $x(n)$ 은 필터링 및 데시메이션 된 후 각 부대역 출력 x_s^i 의 에너지의 순으로 비트할당되어 적응 양자화한 후 2진 부호화되어 전송된다. 여기서 비트할당을 위해 반복적 알고리즘을 이용함으로써 계산량이 많게 됨을 문제점으로 지적해 왔다. 양자화 간격은 신호 $x(n)$ 의 크기에 따라 매 16ms의 블록 단위로 비균일하게 양자화된다. 이 때 사용되는 APCM(Adaptive Pulse Code Modulation) 양자화기의 구성은 그림 2와 같이 양자화기 앞에 x_s^i 의 에너지 값에 반비례하는 스케일 인수 G 를 삽입하고 양자화기 뒤에 G 의 역 인수를 삽입함으로써 양자화기의 특성이 항상 일정하도록 한다¹⁾.

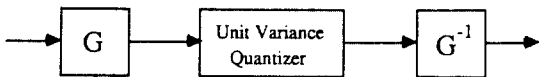


그림 2. APCM 양자화기 구성도
Schematic Diagram of APCM Quantizer

원래의 신호 $x(n)$ 은 전송된 부대역 데이터 및 보조정보 (각 부대역 에너지)를 이용하여 역과정으로 수신측에서 $x(n)$ 을 출력한다. 이 때 비전송 대역의 데이터 x_s^5, x_s^6 은 보조정보의 에너지에 해당하는 백색잡음열을 주입하여 재생한다. 그러나 이같은 방법은 잡음의 파형이 실제 비전송 대역 음성파형과 다르게 주입되므로 출력 음성에 이상 잡음이 나타나는 단점이 있다.

III. 성능 개선

앞절에서 언급된 비트할당에 필요한 계산량의 감소와 부대역 웨이팅에 의한 음질개선을 위해

다음과 같은 방법을 검토한다. 먼저 계산량 감소부터 언급한다.

1. 비트할당 알고리즘의 제안

앞에서 설명되었듯이 DSBC 방법은 비트가 블록 단위로 할당되기는 하나 허용된 전체 비트가 다 할당될 때까지 각 부대역 x_s^i 의 에너지 σ_k ($k=1, \dots, 6$)를 구하여, 그중 에너지가 가장 큰 대역에 1비트를 할당한 다음 이 대역이 에너지를 인수 r 로 스케일한다. 이 절차를 다시 반복하여 결국 각 대역에 0~5 비트를 할당하는 방법이기 때문에 반복적인 비트할당으로 인한 계산량은 엄청난 결과일 것이다. 그러나 프레임당 각 x_s^i 에 할당된비트의 평균을 구하여 이 평균값으로 각 x_s^i 의 할당비트를 미리 결정한다면 송수신시 각 대역에서 에너지 σ_k 의 크기비교는 블록당 한번으로 가능해질 수 있으므로 계산량의 축소를 예상할 수 있다. 따라서 종래의 비트할당 방법과 제안된 알고리즘을 표2와 같이 비교할 수 있으며

표 1. 가변 비트 할당 알고리즘
Dynamic Bit Allocation Algorithm

	종래 방법	제안된 방법
1.	각 부대역 에너지 σ_k 계산	각 부대역 에너지 σ_k 계산
2.	DO=1, NBIT Find max(σ_k)= σ_m k=1, NBAND m band에 1비트 할당 IF (# of bits in m band >=5) THEN $\sigma_m = \sigma_m \cdot r$ GOTO 1 ELSE $\sigma_m = \sigma_m / r$ 스케일 인수 END IF k bit=k bit+1 1 END DO	Find Indexing $\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6$
3.	비트할당	비트할당

¹⁾NBIT : 전체 할당비트, NBAND : 부대역 수

k bit : 각 부대역에 할당된 비트

연산의 복잡도가 종래의 방법이 $O(nbit \times nband)$ 임에 반해 제안된 방법에서는 $O(nband)$ 이므로 계산량 감소를 쉽게 예상할 수 있다.

그림 3은 그림 1의 수행절차를 나타낸다. 먼저

6개 부대의 필터에서 프레임당 $x_s, k=1, \dots, 6$ 에너지 $\sigma_k, k=1, \dots, 6$ 를 구하여 이를 보조정보로 전송하는 한편 σ_k 를 크기순으로 인덱싱 (Indexing)하고 각 대역별로 다음 인덱스 값에

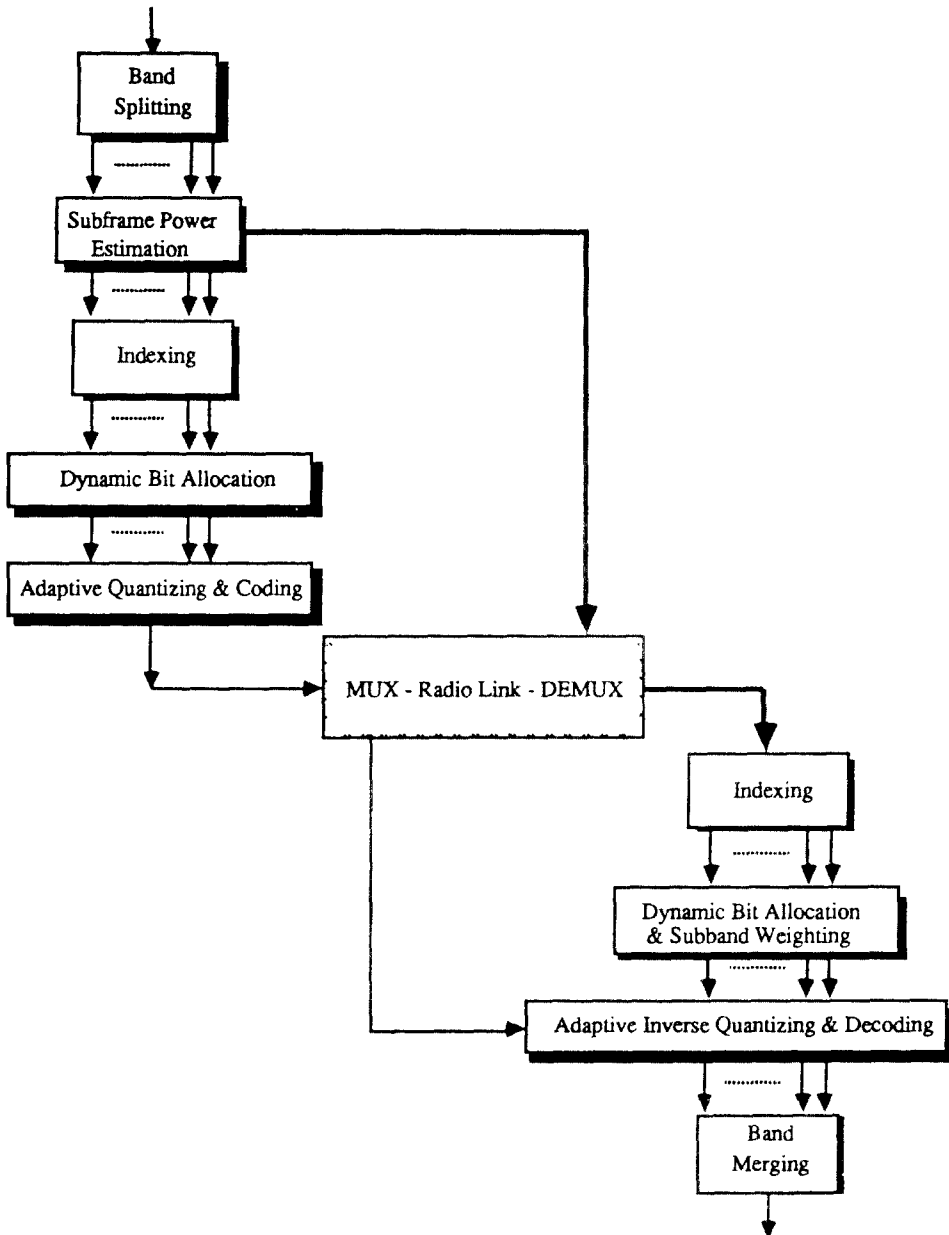


그림 3. DSBC 시스템 블록도 및 순서
Flow Diagram of DSBC System

따라 정해진 평균 비트를 할당하여 데이터를 전송하지만 그러나 0비트가 할당되는 부대역 데이터는 전송하지 않는다. 이 때 전체 전송속도를 12kbps로 목표한다면 각 부대역에 할당된 총 비트수는 보조정보 전송을 위한 2kbps를 제외하면 10kbps가 될 것이다. 따라서 이를 6개 부대역 중 4개 부대역에 에너지 크기순으로 3, 3, 2, 2, 또는 4, 3, 2, 1 비트를 할당하여 전송한다면 나머지 2개 대역 이 데이터는 전송하지 않아도 된다. 이같은 비트 할당 방법의 한 예를 표 3과 같이 나타낼 수 있다.

표 3. 가변 비트할당의 한 예
An Example of Dynamic Bit Allocation

대역	주파수(HZ)	Power	Index	비트 할당	비트 할당	Bit Rate	비트 할당	Bit Rate
1	0-500	67.0	4	2	2K	2	2K	2K
2	500-1000	9.69	5	3	3K	3	3K	3K
3	1000-1500	10.24	6	4	4K	3	3K	3K
4	1500-2000	4.26	3	1	1K	2	2K	2K
5	2000-2500	3.21	2	0	0	0	0	0
6	2500-3000	1.50	1	0	0	0	0	0
Total					10K			10K

표 3에서 볼 수 있듯이 임의의 한 프레임에서 $x_s^1, x_s^2, \dots, x_s^6$ 의 에너지 $\sigma_1, \sigma_2, \dots, \sigma_6$ 크기순으로 인덱싱하면 σ_k 가 가장 작은 부대역 x_s^6 의 에너지 $\sigma_6=1.50$ 이 1, ..., 가장 큰 x_s^3 의 $\sigma_3=10.24$ 가 6으로 인덱싱된다. 이와 같이 매 프레임당 비트를 인덱싱 값에 따라 할당할 수가 있다.

2. 부대역 웨이팅에 의한 음질개선

음질개선 방법의 하나로 부대역 웨이팅을 생각해 볼 수 있다. RA방법에서는 비트가 할당되면 먼저 부대역 x_s^1 을 x_s^1 의 에너지 σ_k 로 나누어 분산을 1로 정규화한다. 그 다음 양자화기를 통과시키고 2진 부호화하여 전송한다. 수신측에서는 역으로 프레임마다 다시 인덱싱하여 비트를 할당하고 빠진 대역의 신호를 재생해야 하는 바 비전

송 부대역의 σ_k 에 해당하는 $N(0, \sigma_k)$ 백색잡음을 주입하는 방법이 이용되었다. 그러나 이 방법은 재생음성에 듣기 거북한 히싱잡음이 나타난다^(3,4). 실제 구현시 백색잡음 데이터를 저장하였다가 매번 이용해야 하는 과정에서 같은 형태의 데이터가 더해지게 되므로 히싱 현상은 더욱 심하게 나타난다. 그러나 전송된 최하위 대역 즉 4번째 대역의 데이터로 비전송 대역의 에너지와 웨이팅 한다면 5, 6번째의 비전송 부대역 신호의 재생이 가능해질 수 있다.

그림 4는 음성의 어느 한 부분에서 얻어진 각 부대역 신호 x_s^i 의 크기순 나열이며 비전송 x_s^5 및 x_s^6 의 에너지가 전송된 x_s^4 의 에너지와 거의 같음을 볼 수 있다. 따라서 이중 에너지가 가장 작은 비전송 x_s^5 및 x_s^6 에 해당하는 데이터 \hat{x}_s^5 및 \hat{x}_s^6 는 다음식에 의해 재생이 가능해진다.

$$\hat{x}_s^5(i) = \sigma_5 x_s^4$$

$$\hat{x}_s^6(i) = \sigma_6 x_s^4, i=1, \dots, 16$$

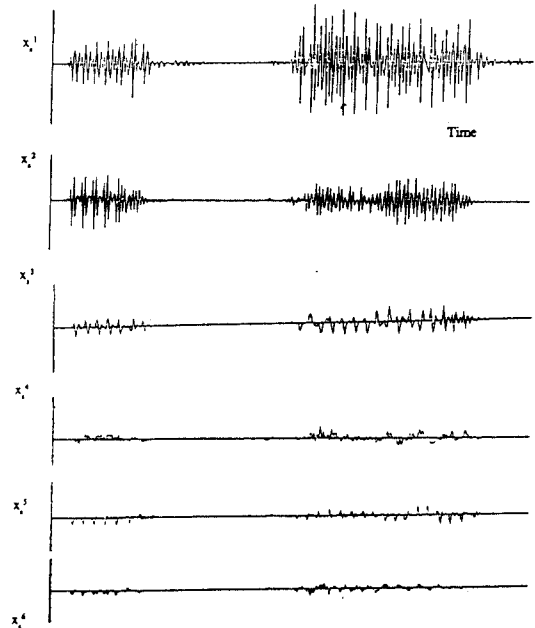


그림 4. 부대역의 출력
Subband Output

여기서 \hat{x}_s^k 와 σ_k 는 수신측에서 k번째 대역의 신호 및 RMS 에너지이다. 따라서 이 방법을 매프레임마다 이용하면 주입되는 신호의 에너지 σ_s 와 부대역 에너지 크기가 같으면서도 입력 신호 x_s^k 의 패턴에 따라 매프레임마다 형태를 달리한 잡음 주입이 가능해질 수 있다.

IV. 결과 및 검토

제안된 방법의 타당성을 검토하기 위하여 다음과 같이 시뮬레이션하였다. 문헌⁴⁾에 근거하여 샘플링 주파수 8KHz, 프레임 길이 16ms, 전체 전송속도 12kbps, 부대역수 8, 32차의 QMF (Quadrature Mirror Filter) 필터뱅크, Max의 최적 양자화기와 남녀 각 4초간의 음성용을 이용하였다. 표 3은 RA의 방법과 제안된 방법에 의한 각 대역별 평균 비트수이다.

표 3. 각 대역별 평균 비트 할당율
Average Bit Allocation Rate of Each Subband

비트할당 방 법	각 대역별 평균 비트 할당					
	부대역1	부대역2	부대역3	부대역4	부대역5	부대역6
Proposed 4, 3, 2, 1	남 3.57 여 3.69	2.77 3.04	1.71 1.54	0.78 0.77	0.57 0.36	0.59 0.60
Proposed 3, 3, 2, 2	남 2.83 여 2.95	2.67 2.83	1.86 1.77	1.03 1.17	0.81 0.49	0.75 0.78
Ramstad 0~5	남 3.71 여 3.86	2.65 2.96	1.60 1.31	0.77 0.68	0.68 0.48	0.58 0.71

두 방법을 서로 비교해 볼 때 RA 방법에서 4, 3, 2, 1의 비트할당이 3, 3, 2, 2의 경우보다 양호함을 나타내고 있다. 환언하면 4, 3, 2, 1 경우는 3, 3, 2, 2에 비해 비트할당이 에너지 분포에 알맞도록 이루어졌음을 의미한다. 표 4는 RA의 방법에서 비트할당 및 부대역 재생 방법을 달리했을 때의 SEGSR (Segmental Signal to Noise Ratio)이다. 비트할당만 달리한 경우 SEGSR은 남녀 공히 1~2dB 정도 개선되

었고 부대역 재생 방법을 달리할 경우 SEGSR은 0.1~0.2 dB 정도의 개선을 나타내나 이는 하성 잡음의 제거를 의미하므로 음질 개선 효과는 오히려 부대역 재생방법에서 더 크다고 할 수 있다. 또한 제안된 평균 비트할당 방법 가운데서 4, 3, 2, 1의 경우는 3, 3, 2, 2의 경우보다 약 1.3dB 개선되었으나 실제 청취결과 거의 구분하기 어려웠다. 그러나 3, 3, 2, 2 바드로 할당하면 실제 구현시 4, 1 비트 양자화가 필요없게 될 것이므로 이에 관련된 500워드 정도의 메모리 감축이 가능해질 수 있다.

표 4. 시뮬레이션 결과
Simulation Results

비트할당 방 법	SEGSR (dB)	
	백색잡음 주입방법	부대역 웨이팅방법
Proposed 4, 3, 2, 1	남 12.20 여 14.07	12.31 14.16
Proposed 3, 3, 2, 2	남 10.88 여 12.85	10.93 12.92
Ramstad 0~5	남 9.95 여 11.95	10.04 12.24

부호화기 출력파형의 관찰은 부호화기 성능평가의 또다른 방법일 수 있다.그림 5는 부호화기 출력으로서 그림에서 (a)는 원음, (b)는 RA방법, 그리고 (c)는 제안된 방법에 의한 출력이다. 이들을 서로 비교해 볼 때 그림 (c)는 (b)에 비해 원음에 더욱 가까움을 얻을 수 있다.

다음으로 제안된 웨이팅 효과를 스펙트럼을 통해 알아본다. 그림 6은 그림 5의 한 프레임만을 FFT한 것으로 (a)는 원래 음성을, (b)는 백색잡음 주입 방식을, 그리고 (c)는 부대역 웨이팅을 나타냈는데 그림에서 보건데 DSBC 백색잡음 주입의 경우 스펙트럼은 대역전반에 걸쳐 비교적 평탄하게 매프레임 같은 형태로 나타났으므로 신호 에너지가 낮다면 잡음으로 들릴 우려가 있다. 반면 제안된 방법에서는 신호가, 원음의 스펙트럼 형태를 거의 비슷하게 따라가므로 이들

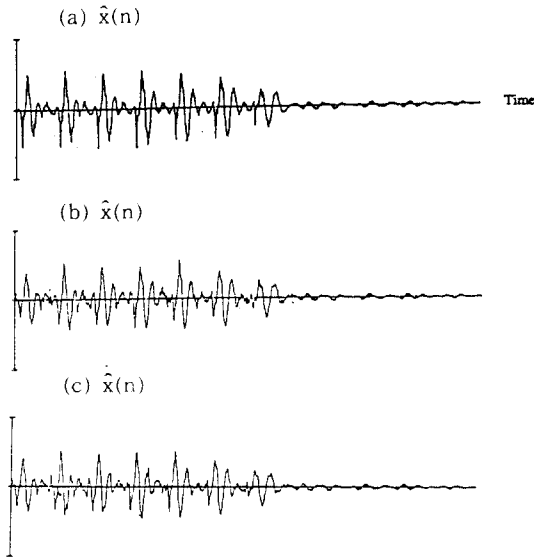


그림 5. DSBC 부호화기 출력파형
a) 원음 b) Ramstad c) 제안된 방법
Output Waveform of DSBC Speech Coder
a) Original Waveform b) Ramstad c) Proposed Method

프레임당 성분을 서로 더하면 평탄한 잡음 스펙트럼을 변형시키는 효과를 나타내게 할 수 있을 것이며 결국 청각적으로 인지 효과(Perceptual Enhancement)가 나타날 것이다.

V. 결 론

DSBC 음성 부호화기에서 성능개선의 두 방법을 제안하였고 이론의 타당성을 시뮬레이션으로 확인하였다. 평균 비트할당의 개념에 의거 음질저하 없이도 종래의 가변 비트할당 방법보다 비트할당에 필요한 계산량을 감축시킬 수 있었으며 간단한 부대역 웨이팅 방법에 의해 종래 백색잡음 주입방법의 단점인 히싱현상을 제거할 수 있는 청각적 음질 개선의 효과를 얻을 수 있다.

제안된 방법은 DSBC의 음질 개선에 적합하다고 믿어지며 잡음 음색 효과측면에서 후단 필터링(Post-Filtering) 효과와의 비교는 차후검토의 대상이 될 수 있다고 생각된다.

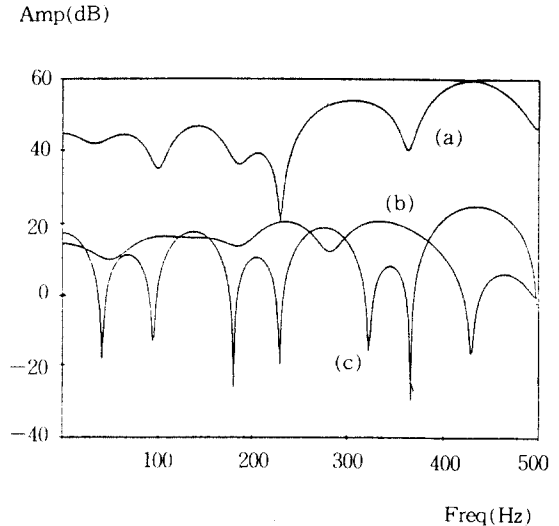
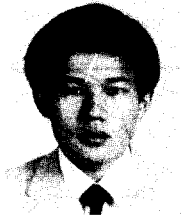


그림 6. 부대역 웨이팅 방법과 백색잡음 주입방법의 스펙트럼 비교
Spectral Comparison Subband Weighting with White Noise

參 考 文 獻

1. X. Maitre, "7 khz audio coding within 64 kbit/s", IEEE J. on SAC, vol. 6, no. 2, pp. 283-298, Feb, 1988.
2. M. J. McLaughlin, "Speech and channel coding for digital land mobile radio", IEEE J. on SAC, vol. 6, no. 2, pp. 332-345, Feb. 1988.
3. R. V. Cox, "New directions in subband coding", IEEE J. on SAC, vol. 6, no. 2, pp. 391-409, Feb. 1988.
4. T. A. Ramstad, "Subband coder with a simple adaptive bit allocation algorithm", in Proc. ICASSP, pp. 203-2207, 1982.
5. R. E. Crochiere, "On the design of subband coders for low bit rate speech communication", BSTJ, vol. 56, no. 5, pp. 747-770, May, 1977.
6. Jayant and Noll, Digital Coding of Waveforms, Prentice Hall, 1984.
7. Sun Y. Kim and Jae K. Kim, "A study on the implementation of DSBC speech codec for digital mobile communications", KIEE J, vol. 27, no. 8, August 1990.



김 선 영(Sun Young KIM) 正會員
1982년 : 동국대학교 전자공학과(석사)
1984년 : 동국대학교 전자공학과(석사)
1988년 : 동국대학교 전자공학과(박사수료)
현재 : 한국전자통신연구원 이동통신연구실 근무



金 在 功(Jae Kong KIM) 正會員
1938年生
1961年 : 漢陽大學校 工科學 電氣工學科 卒業
1964年 : 同 大學院 卒業
1966年 : 日本國 早稻田大學 電氣通信科 研究
1980年 : 英國 Loughborough 大學校 電子電氣科 研究員(10年)
1970年 ~ 現在 : 東國大學校 工科學 電子工學科 教授(工博)
※關心分野 : 信號 傳送