論文93-30B-9-4

아날로그 음성 비화기의 비도및 음질 향상에 관한 연구

(A Study on the Improvements of Security and Quality for Analog Speech Scrambler)

孔 炳 球*, 趙 東 浩**

(Byung Goo Kong and Dong Ho Cho)

要 約

본 논문에서는 비화기의 높은 비도와 양호한 복원 음질 확보를 위한 새로운 알고리즘을 제안한다. 이 알고리즘은 fast fourier transform (FFT) 의 계수들을 재배치 하는 것을 기본으로 하고, 음성의 pre / post filtering 과 hamming window 및 적응 의사 스펙트럼 삽입을 부가한 것이다. 여기서 pre / post filter 는 음성 스펙트럼을 평활화 하고 대역외의 잡음 제거 효과를 위한 것이며, 적응 의사 스펙트럼 삽입은 묵음 구간과 음성 구간의 구별 정보를 제거하기 위함이다. 또한 hamming window 기법은 전화 라인의 동기 error 에 대한 내구성 (robustness) 유지를 위하여 적용 한다.

시뮬레이션 결과, 비화기의 비도와 음질이 주관적 및 객관적 성능 평가에서 기존의 알고리즘에 비하여 향상됨을 보이고 있고, 동시에 본 비화 알고리즘이 동기 error 에 대하여 내구성이 유지됨을 보인다.

Abstract

In this paper, a new algorithm for high level security and quality of speech is proposed. The algorithm is based on the rearrangement of the fast fourier transform (FFT) coefficents with pre and post filter process, hamming window and adaptive pseudo spectrum insertion. Then, the pre and post filters are used for the whitening of speech spectrum and the adaptive pseudo spectrum is inserted for the unclassification of silence/speech. Also, the hamming window technique is applied for the robustness to the syncronization error in the telephone line.

According to the simulation results, it can be seen that the security of scrambled signal and the quality of descrambled signal have been improved fairly in both subjective and objective performance test and the new FFT scrambler is robust to the synchronization error.

* 準會員, 三星電子 情報通信部門綜合研究所
(Key Technology Lab., Information System Business, Samsung Electronics Co.)
**正會員, 慶熙大學校 電子計算工學科
(Dept. of Computer Eng., Kyunghee Univ.)
接受日字: 1992年 7月 25日

Ⅰ. 서론

공중 전화망 (public switched telephone network) 이나 이동 통신망 (mobile telephone system) 에서의 음성 비화는 그 네트워크 특성에 따라 디지탈 또는 아날로그 방식으로 이루어진다. 음성의 디지탈 비화 방식은 저속도 전송의 vocoder 방식의 디지탈 정보열을 섞음으로써 간단히 구현된다.

음성을 압축하는 방식에는 waveform coding 과 source coding 방식이 있는데, 전자는 시간축상의 음성 신호를 그대로 압축하는 기법으로 pulse code modulation, adaptive differential pulse code modulation, adaptive delta modulation 등이 있다. 후자는 인간의 발성 기관의 발성 과정을 모델화하고 음성 신호를 그 모델에 모델링하여 특징을 추출하는 기법으로 선형 예측 부호화 (linear predictive coding, line spectrum pair, code exited linear prediction, vector sum exited linear prediction) 등이 이에 해당한다. 지금까지의 기술 수준에서 중속도 (16 Kbps) 이상에서는 waveform vocoder 를 사용하나, 저속도 (8 Kbps 이하) 에서는 source coding 기법이 음질이 훨씬 좋기 때문에 저속도 vocoder 로 사용된다. [10], [11]

그러나 아직까지 저속도 vocoder 의 음질 수준이 충분히 좋지는 않다.이에 비하여 실시간 처리를 위해서 다소 복잡한 하드웨어가 요구되지만 음질이 양호한 아날로그 비화 방식이 아직은 유효한 것으로 알려지고 있다. [1] 아날로그 음성 비화 기법중에서 FFT 비화 방식은 매우 효과 적이고, 음질도 상당히 좋으나 비도가 매우 낮아서 기밀 유지가 어렵다. 이러한 문제를 극복 하기 위하여 기본적인 FFT 비화 방식에 잡음에 해당하는 의사 스펙트럼을 삽입하거나 특정 대역의 스펙트럼 크기를 변형하는 등의 여러 비화 방식들이 제안되었다. 그러나 이러한 알고리즘은 비도는 어느 정도 향상되는 반면에 복원된 음성의 음질이 상당히 저하되는 결점을 보인다. [2], [3]

이에 본 논문에서는 먼저 기존의 비화방식을 구현하고 분석한다. 그리고 기존의 비화 방식이 채택한 부분적인 알고리즘과 대비되는 다른 알고리즘, 음성 특징에 적합하다고 판단되는 방식도 병행하여 살펴보고, 그 결과를 비교 분석한다. 이러한 분석 및 결과를 토대로 음성 정보의 기밀이 유지될수 있는 높은 수준의 비도가 확보되고 아울러 복원된 음성이 양호한 음질을 갖을수 있는 새로운 비화 알고리즘을 제안한다. 본 논문에서 사용하는 알고리즘은 기본적인 FFT 비화 방식을 근간으로 pre / post filter 와 hamming window 그리고 적응 의사 스펙트럼의 삽입기법을 활용한다.

서론에 이어 Ⅱ장에서는 기본적인 FFT 비화 방식에 의사 잡음을 삽입하는 방식을 소개하고, Ⅲ장에서는 새롭게 제안한 비화 알고리즘을 자세히 설명한다. 또한 Ⅳ장에서는 기존의 알고리즘과 제안한 알고리즘의 성능 비교 평가를 실시하고 그 결과를 논한다. 계속해서 Ⅴ장에서 제안한 비화 알고리즘의 Hardware 설계를 검토하고,마지막으로 가장의 결론으로 본논문의 끝을 맺는다.

Ⅱ. 기존의 비화 방식

1. 기본적인 FFT 비화 방식

기본적인 FFT 비화 방식의 동작 과정은 다음과 같다. 전송측 즉 비화를 수행하는 측에서는 입력되는 음성 신호를 analog-to-digital converter (ADC) 를 통하여 디지탈 신호로 변환하고 이 신호를 FFT 기법을 이용하여 주파수 영역으로 전환한다. 이 변화된 신호를 스펙트럼 이라 부르는데, 스펙트럼을 나타내는 계수 (FFT coefficient) 들을 주어진 암호 key 에 의해 섞게된다. 이와같이 섞인 스펙트럼 계수들을 inverse fast fourier transform (IFFT) 를 통해서 시간 영역의 디지탈 신호로 전환한다. 계속해서 이 디지탈 신호를 digital-to-analog converter (DAC) 를 통하여 아날로그 신호로 변환한 후에, 상대편으로 전송한다. 수신하는 측에서는 비화 과정과 역순으로 진행되는데 스펙트럼의 계수들의 암호화 key 에의한 재정렬을 수행함으로써 복원된 음성 신호를 얻게된다. 이 기본적인 FFT 비화 방식을 그림 1에, 그 실험 결과를 원래의 음성 신호와 함께 그림 2에 보인다.

이 기본적인 FFT 비화 방식의 근간이 되는 개념은 시간 영역에서의 음성 신호를 섞는 것 보다는 주파수 영역에서의 스펙트럼 계수를 섞는 것이 비도 측면에서 상당히 효과적이라는 사실에 기인한 것이다 [1], [2].

따라서 지금까지 기본적인 FFT 비화 방식의 비도를 높이기 위하여 여러가지 알고리즘이 연구되어 왔고 제안되었는데, 어느정도의 비도 향상의 개선을 이룬 대표적인 알고리즘이 의사 잡음 삽입 알고리즘이다.

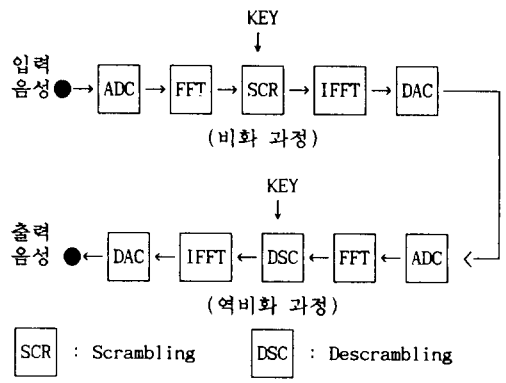
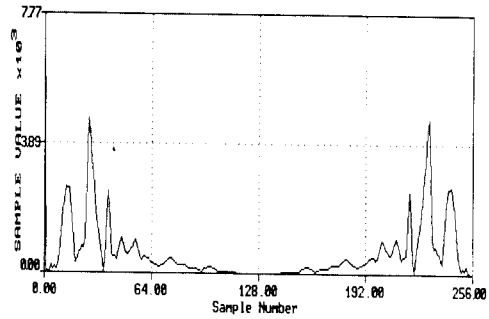
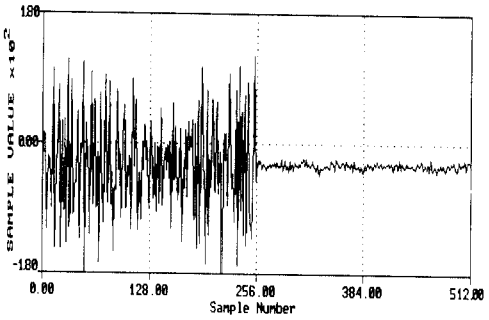


그림 1. 기본적인 FFT 비화 방식

Fig. 1. The Conventional Method of FFT Scrambling.



(a)



(b)

그림 2. 원래 음성 신호와 기본적인 FFT 비화 방식의 실험 결과
(a) 원래 음성 신호의 스펙트럼
(b) 기본적인 FFT 비화 방식으로 비화한 경우의 신호 파형

Fig. 2. The Spectrum of Natural Speech and of Scrambled Signal by Conventional Method.
(a) The Spectrum of natural Speech.
(b) The Scrambled Signal by Conventional Method.

2. 의사 잡음 삽입 알고리즘

스펙트럼의 계수들의 재배치 처리 이전에 잡음에 해당하는 스펙트럼을 삽입하는 방식으로 재배치의 효과를 높힘으로써 높은 수준의 비도를 확보하는 알고리즘이다. 이때 삽입하는 위치와 삽입 형태에 따라 여러가지 알고리즘이 제안되어 왔다. 또한 스펙트럼 외에 시간축상의 음성 신호에 임의의 잡음을 삽입하는 알고리즘도 제안되었다. 일반적으로 삽입하는 위치는 묵음구간에 해당하는 주파수 대역 또는 시간축상의 대역으로 정해지고, 역비화 과정에서 삽입 위치를 검출하여 제거하는 것이 기본적인 원리이다. 이러한 잡음 삽입 알고리즘중에서 비교적 효과적이고 간단한 의사 잡음 삽입 알고리즘을 소개한다.[3] [6]

이 의사 잡음 삽입 방식은 스펙트럼 계수들을 몇개의 블록으로 나누고, 각 블록의 에너지가 전체 에너지 보다 적은 블록을 삽입 위치로 정하여 특정 형태의 블록을 기존 블록과 대치 삽입하는 알고리즘으로 그림 3에 그 구성도를 보이고, 그림 4에 그 실험 결과를 보인다.

그림 4에서 보는 바와 같이 비도가 어느 정도 향상되었으나, 아직 음성 구간의 구별 정보와 억양 패턴 등의 정보가 남아있고, 복원된 음성 신호의 음질은 기본적 FFT 비화 방식에 비해 현저히 저하된다.

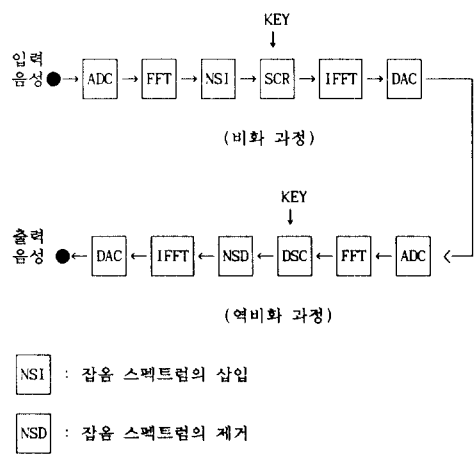
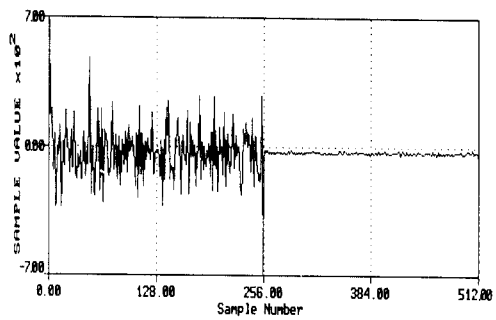
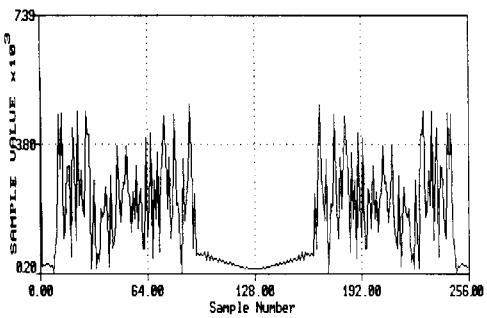


그림 3. 의사 잡음 삽입 방식

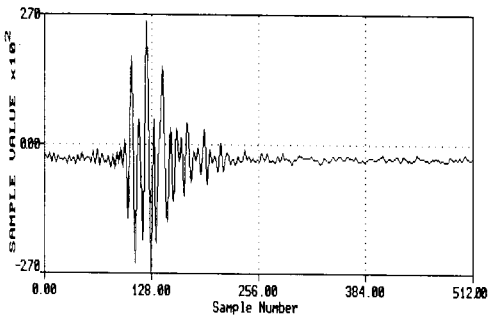
Fig. 3. The Method of Pseudo-noise Insertion.



(a)



(b)



(c)

그림 4. 의사 잡음 삽입 방식의 실험 결과
(a) 의사 잡음 삽입 방식으로 비화한 신호의 파형
(b) 의사 잡음 삽입 방식으로 비화한 스펙트럼
(c) 의사 잡음 삽입 방식으로 복원한 신호의 파형

Fig. 4. The Results of Pseudo-Noise Insertion Method.
(a) The waveform of Scrambled Signal by Pseudo-noise Insertion Method.
(b)The Spectrum of Scrambled Signal by Pseudo-noise Insertion Method.
(c)The waveform of Reconstructed Signalby Pseudo-noise Insertion Method.

Ⅲ. 새로운 비화 알고리즘

의사 잡음 삽입 방식을 부가한 기존의 FFT 비화 방식의 성능을 향상시키기 위하여 기본적인 FFT 비화 방식에 pre / post filtering 과 hamming window, 적응 의사 스펙트럼의 삽입을 추가한 새로운 알고리즘을 제안한다. 이 제안한 알고리즘의 전체 블록도는 그림 5 와 같으며 그 실험 결과를 그림 6에 보인다.

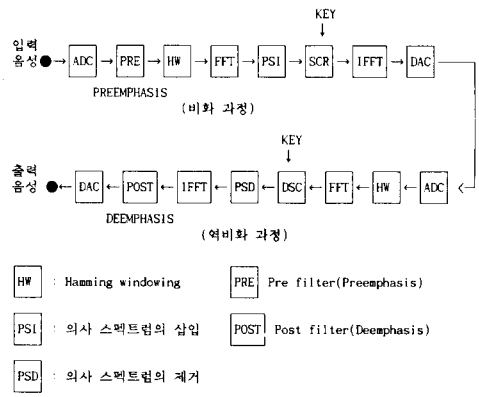
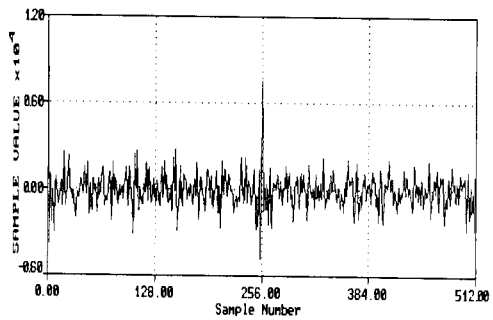
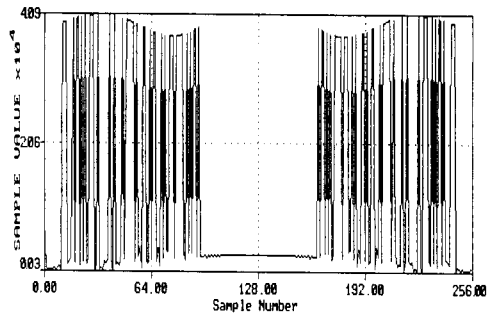


그림 5. 제안한 비화기의 블록도

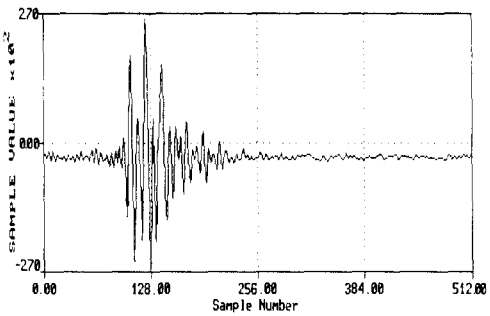
Fig. 5. The Block Diagram of Proposed Algorithm.



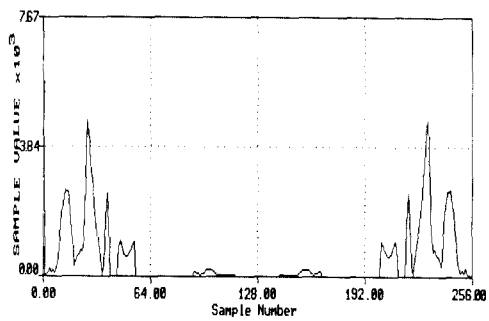
(a)



(b)



(c)



(d)

그림 6. 제안한 비화 방식의 실험 결과
(a) 제안한 비화 방식으로 비화한 신호 파형
(b) 제안한 비화 방식으로 비화한 스펙트럼
(c) 제안한 비화 방식으로 복원한 신호 파형
(d) 제안한 비화 방식으로 복원한 신호 파형

Fig. 6. The Results of Proposed Method).
(a) The Waveform of Scrambled Signal by Proposed Method.
(b) The Spectrum of Scrambled Signal by Proposed Method.
(c) The Waveform of Reconstructed Signal by
(d) The Spectrum of Signal by Proposed Method.

1. Pre / Post filtering (preemphasis 와 deemphasis)

유성음 신호의 전형적인 스펙트럼이 그림 2의 (b)에 나타나 있는데, 유성음은 저주파수 대역에 대부분의 에너지가 몰려있다. 반면에 묵음과 무성음 신호는 비교적 주파수 전대역에 걸쳐 고르게 에너지가 분포하게된다. 이러한 현상은 유성음의 경우 입술의 방사와 성문 모양의 효과에 기인하는 것으로 주파수가 증가함에 따라 약 6dB 의 기울기를 갖고 에너지가 감소하는 특성을 지닌다. 이러한 특징은 비화 알고리즘의 섞는 효과를 반감시켜 비도를 저하시키는 한 요인이된다. 따라서 비도를 향상시키기 위해서는 스펙트럼의 형태를 섞기전에 평활화 하는 것이 필요하다. 이러한 평활화는 preemphasis filter 를 통해서 구현이 가능하고 그 원리는 식 (1) 과 같이 시간 축상의 함수로 간단히 나타낼수 있다.

$$\hat{s}[n] = s[n] - 0.95 \times s[n-1] \tag{1}$$

여기서 $\hat{s}[n]$은 filtering 된 음성 신호이고, s [n] 은 원래의 음성 신호이다. 이렇게 filtering 된 음성 신호는 수신측에서 역비화 과정의 맨끝 부분 즉 아날로그 신호로의 변환 직전에 deemphasis filtering 을 거침으로써 preemphasis filter 효과가 제거된다. 이때 Deemphasis filtering 에 관한 수식은 아래 식 (2) 와 같이 표시된다.

$$s[n] = \hat{s}[n] - 0.95 \times s[n-1] \tag{2}$$

이 pre / post filter 인 preemphasis 와 deemphasis 의 처리 효과가 대역 필터 역할을 거의 수행하기 때문에 복원 음성의 음질 향상을 위해 별도의 대역 필터를 필요로 하지 않는 장점이 있다. 즉 pre / post filter 는 복원 음성의 음질 향상에 상당히 기여함을 알수있다. [5]

2. Hamming window

기존의 FFT 비화기는 일정 구간의 단위마다의 처리를 위해 구형 (rectangular) window 기법을 채택하는데 이는 복원 음질에 어느정도 음질 저하 요인이 된다. 여기서 음질 저하가 발생함에도 불구하고 다른 기법을 채택하지 않는 것은 계산량이 비교적 적게되어 실시간 구현이 가능하기 때문이다. 그러나 급격한 마이크로 회로의 발달로 다른 window 기법의 처리도 간단히 실시간으로 구현할수 있기 때문에, 본 논문에서는 구형 window 기법 대신에 hamming window 기법을 채택한다.

이 hamming window 는 구형 window 기법에 비하여 계산량은 증가하나 edge effect 가 감소하고, 실제의 전송 라인상의 error 에 대하여 내구성을 향상 시킨다. 즉 인접하는 단위간의 연결이 hamming window 의 중첩 효과에의해 좋아지고, 가중치 효과에 따라 국부적인 error 를 window 전구간으로

smoothing 하여 내구성을 향상시킨다. 여기에서는 한 frame 의 길이를 10msec 로 하고, window 의 길이를 30msec, 즉 3개의 frame 으로 하며 한 frame 씩 움직이는 sliding windowing 기법을 채택하였다. 여기서 해당 주기의 시간축상의 데이타 수와 FFT 변환 기법에 의한 데이타 수가 달라지게 되는데 이는 overlap addition method 로 해결하였다. [7]

3. 적응 의사 스펙트럼 삽입

의사 스펙트럼 삽입의 목적은 스펙트럼의 계수들을 섞기전에 원래의 음성 스펙트럼의 형태를 변화시켜 비음성 형태를 갖게 함으로써 섞는 효과를 향상시키고, 해독자가 해독시 이미 알고있는 음성 스펙트럼의 형태 정보를 무력화하는 것이다. 의사 스펙트럼의 삽입에는 그 삽입 위치와 삽입하는 형태에 따라 기법이 다양하고, 그 효과가 상당히 달라지게 된다. 여기서 삽입위치와 삽입 형태는 다음과 같이 결정하였다.

1) 삽입 위치 결정 알고리즘

먼저 삽입 위치는 한단위를 연속되는 블록으로 나누고, 각 블록을 의사 스펙트럼으로 대치 삽입할 것인가를 판단하여 삽입이 결정된 블록의 위치로 정한다. 이때의 한 블록의 길이는 5 ~ 6 개의 스펙트럼 계수를 갖도록 하였다. 삽입을 결정하는 방식은 각각의 블록의 에너지를 계산하고, 각 블록의 에너지를 frame 의 블록 중에서 최대 에너지를 갖는 블록의 에너지와 비교한다. 이때 에너지의 비가 일정 수준이하가 되는 블록을 대체하여 삽입할 블록으로 정하고, 그위치를 삽입 위치로 한다. 에너지 비교치인 Er 을 구하는 식은 아래 식 (3) 과 같다.

$$E_r = 20 \times \log_{10}\left(\frac{segment\ energy}{\max.\ segment\ energy}\right) \qquad (3)$$

이러한 기준을 음성 신호 전구간에 대하여 적용하면 음성이 존재하는 구간과는 달리 묵음의 구간에서는 삽입이 음성 구간과는 뚜렷이 구분되게 결정되어 음성 구간을 판별할수 있는 정보가 남게된다. 이는 해독자에게 음절적 분석 정보를 제공함으로 비도가 떨어지게 된다. 따라서 묵음 구간에서의 삽입 위치 결정 알고리즘을 별도로 설정하여 비화 후의 음성 신호에서 음절적 구별 정보를 제거한다. 즉 먼저 음성 구간과 묵음 구간을 각 블록의 절대 에너지로 판단하고, 묵음 구간으로 판단되는 블록은 에너지 비교치와는 상관 없이 대체 삽입함으로써 음성 구간과 구별할수없도록 하는 것이다.

여기에서는 음성 신호를 12 bit 크기의 8kHz sampling 데이타 80개 (10msec) 음성 신호를 256 개의 FFT 계수로 변환하여 한 frame 으로하고, 그 frame 을 연속되는 5개의 계수들의 길이를 갖도록 52개의 블록을 만든다. 이러한 환경에서 반복적인 실험을 거쳐 묵음 구간과 음성 구간에 있어서의 삽입 위치 결정에 필요한 임계값을 표 1 에 나타낸다. 표에 나타낸 묵음 구간에 대한 임계값은 최대값을 1 로 정규화한 경우의 절대값으로 임계값 이하인 에너지는 묵음 구간으로 분류한다.

표 1. 임계값
Table 1. Threshholds.

종 류	값
THD (음성)	-15dB
THD (묵음)	0.05

2) 삽입 스펙트럼 형태

의사 스펙트럼의 형태를 결정하는 것은 비도를 향상 시키는것 뿐만이 아니라, 수신측의 삽입 위치 검출에 있어서도 매우 중요하다. 먼저 비도를 높이기 위해서는 비화된 신호가 비음성 형태이며 평활화가 잘되어야 한다. 이는 삽입하는 스펙트럼은 음성의 스펙트럼의 형태를 갖는 의사 스펙트럼이 되어야 하고, 또한 단위내에서 최대 에너지를 갖는 블록의 에너지를 갖어야 함을 의미한다. 따라서 음성 신호의 특정 부분 (유성 모음의 최대 에너지를 함유하는 부분) 의 스펙트럼의 형태를 삽입 형태로 정하고, 그 크기는 단위마다의 최대 에너지와 같도록 적응한 형태의 스펙트럼을 삽입하는 것이다.

그러나 묵음 구간에서는 원래 신호의 에너지가 상대적으로 매우 작기 때문에 해당 단위에 적응하는 크기를 적용하면 삽입되는 스펙트럼도 작게된다. 이것은 음성 구간과 묵음 구간의 변별적 정보가 되므로 묵음 구간에서는 자체 크기에 적응시키지 않고 묵음 구간의 앞에오는 음성 구간의 크기에 적용한다. 또한 의사 스펙트럼이 음성 스펙트럼과 유사하기 때문에 수신측에서 삽입 위치를 검출하기가 매우 어려워 송신측에서 비화된 음성 신호와 함께 삽입 위치 정보를 전송해야 한다. 그러나 이 경우 삽입 위치 정보의 누출 위험이 따를 뿐만이 아니라 별도의 전송 채널등을 필요로 하기 때문에 복잡도가 증가한다.

이러한 정보 누출과 복잡도의 증가를 피하기위하여 삽입 패턴을 특정화 한다. 즉 수신측에서 삽입 위치 정보의 수신없이도 용이하게 삽입 위치를 검출할수 있도록 하는 것이다. 음성 종류에 따라 적응되는 특

성을 살리기 위하여 크기는 위에서 언급한 적응 형태로 유지하고 위상을 나타내는 데이타로 특정화를 수행한다. 이때 특정화는 연속되는 계수들이 일정한 부호의 위상값을 갖도록 하는 것이다.

일반적인 음성 신호의 위상은 부호의 측면에서 전체적으로 균형을 이루고 있기 때문에 삽입하는 의사 스펙트럼의 위상이 항상 음 또는 양의 부호만 갖게되면 삽입 블록의 갯수가 추정될수 있다. 이러한 삽입 정보가 비화 신호에 남지 않도록 삽입 되는 스펙트럼의 위상부호를 음과 양으로 번갈아 삽입한다. 이는 수신측에서 역비화를 수행한후에 스펙트럼상에서 쉽게 삽입 의사 스펙트럼을 검출하여 삽입 의사 스펙트럼의 제거를 용이하게 하며 해독자에게는, 역비화 되지 않은 스펙트럼 상에서 삽입 위치 추정을 불가능하게 한다.

Ⅳ. 시뮬레이션 결과와 검토

기본적인 FFT 비화기와 dummy 스펙트럼 삽입을 부가한 FFT 비화기를, 제안한 비화기 와 성능비교하고 그 결과를 표 2 에 나타낸다. 시뮬레이션은 위에서 언급한 비화기를 일반 PC 상에서 별도 hardware unit 의 구성없이 구현하였다. 각각의 비화기에 입력으로 사용한 음성 신호는 일반적인 A/D 보드를 통하여 미리 음성 화일에 저장 하였다. 또한 각각의 비화기의 출력은 출력 음성 화일로 저장하여, 원래의 입력 음성 신호 화일과의 정량적 분석 (SNR)에 이용하고, 일반적인 D/A 를 통하여 청취도 평가를 수행하였다. 이때 실험실 환경에서는 음성 화일을 마이크를 통해 입력하였고, 전화 라인 환경은 마이크 대신에 실제 전화기를 통해서 입력하였다.

입력 FFT 계수들을 단순히 섞는 기본적인 FFT 비화기는 복원 음질이 매우 뛰어나지만 실제 청취시에는 edge effect 에 의한 주기적인 스펙트럼 왜곡이 있어 그에 기인하는 잡음이 내재된다. 또한 비화된 신호에 원래 음성 신호의 음절적 정보, 언어적 정보, 운율적 정보등이 완전히 제거되지 않고 남아있어 그 비도는 상당히 낮다.

기본적인 FFT 비화기에 dummy 스펙트럼 삽입 기법을 부가한 비화기는 원래의 음성 신호 정보를 어느정도 제거하여 비도가 향상된다. 그러나 묵음 구간과 음성 구간이 뚜렷이 구분되고 억양의 전체 패턴이 완전히 제거되지않아 비록 비화된 신호의 청취는 상당히 어렵지만 해독 정보를 남겨두기 때문에 해독을 가능케 한다. 또한 기본적인 FFT 비화기와 마찬가지로 edge effect 로 인한 잡음 문제가 그대로 남아있고 dummy 스펙트럼 삽입에 따른 복원 음성 신호의 음질 저하가 현저히 발생한다.

표 2. 기존의 비화기와 제안한 비화기의 성능 비교

Table 2. Performance Comparison.

방식	실험실 환경		전화 라인 환경		청취시험 (MOS 5.0)
	비화신호의 SEGSNG	복원신호의 SEGSNG	비화신호의 SEGSNG	복원신호의 SEGSNG	
기본적인 FFT 비화기	4.1dB	35.2dB	-3.1dB	31.3dB	3.1
기본적인 FFT 비화기에 의사 스펙트럼 삽입 부가	-5.8dB	19.8dB	-7.5dB	10.6dB	2.4
제안한 비화기	-25.6dB	22.3dB	-34.6dB	12.7dB	3.8

(Mean Opinion Score)

본 논문에서 제안한 비화 방식은 원래의 음성 신호의 스펙트럼 형태를 preemphasis 하여 스펙트럼을 더욱 평활화하고, 묵음구간도 음성구간과 같이 비화되도록 함으로써 비화된 음성신호에서 원래의 억양 정보 및 음절적 정보를 거의 완전하게 제거한다. 이는 청취 불능은 물론 비화된 신호 자체에 해독 정보를 남기지 않아 비도를 현저히 향상 시킨다. 또한 비도 향상을 위한 알고리즘이 초래하는 스펙트럼의 왜곡 및 edge effect에 기인한 음질 저하를 개선하여 좋은 통화 품질을 유지한다. 여기서 비도 향상을 위해서는 Ⅲ장에서 언급한 pre / post filtering 기법과 적응 의사 스펙트럼 삽입 기법을 적용하고, 음질 저하를 개선하기 위해서 hamming window 기법을 채택 하였다. 또한 pre / post filtering 은 대역 필터의 효과를 발휘하여 edge effect 에 기인하는 잡음 문제를 별도의 대역 필터의 사용 없이 해결하였다.

이러한 시뮬레이션 결과는 새로이 제안한 알고리즘이 비도의 측면 뿐만아니라, 복원 음질 수준에 있어서도 다른 기존의 알고리즘보다 뛰어남을 보이고 있다. 또한 실험실 환경이 아닌 실제 전화 라인상에서의 성능 평가 결과는 이 알고리즘이 선로 error 및 잡음 환경하에서 내구성이 있음을 보이고 있다. 그리고 그림 2와 그림 4 그리고 그림 6에서 보는 바와같이 실제 비화된 신호의 스펙트럼의 평활화가 매우 잘 되어 있으며, 복원 음성 신호의 파형도 상당히 개선됨을 알수있다. 여기에서 평가한 주관적 평가 항목인 MOS 는 성인 4명(남자 2인, 여자 2인)이 발성한 6

개의 문장을 대상으로 하였고, 20명의 평가 대상자를 선정하여 표에 정의한 점수를 기록하게 하였다. 이와 같은 평가를 각각 3회 반복하고 각 알고리즘 별로 평균을 구한것이다. 이때 청취에 사용한 음성 시료는 구현된 각각의 알고리즘을 적용한 것이며, 전화 라인을 통한 실제 환경 시험은 실시간 구현 대신에 OFF-LINE 으로 행하였다.

Ⅴ. Hardware 설계

Hardware 의 블록도를 그림 7에 보인다. 여기서 제안한 비화기는 그 구성이 비화부와 역비화부 ,동기 제어부, 아날로그부로 구성되는데, 비화및 역비화에 사용된 알고리즘과 동기를 제어하기 위해 사용하는 TONE 신호 처리 알고리즘 모두가 10 MIPS 이하가 되어 10 MIPS 이상이 되는 DIGITAL SIGNAL PROCEESOR(DSP) 를 선정하면 하나의 DSP chip 과 전화기및 전화 라인의 접속부만으로 hardware 구성이 가능해진다. 제안한 비화기에 필요한 알고리즘중 가장 큰 계산량을 요구하는 부분은 FFT 수행부분으로 비화및 역비화 과정에서 4번의 계산을 요구한다. 한번의 FFT (256 포인트) 계산에 약 10000 개의 명령어가 수행되므로 한 블록의 시간 10 ms 동안 40000 번의 명령어의 수행이 요구된다. 이는 약 4MIPS를 나타내며 전체 알고리즘의 60%를 차지하므로, 비화기 전체는 7MIPS 정도로 가능하다. 여기에 전송 에러에 대한 복구 작업등을 위하여 3MIPS의 여유를 갖도록하여도 10 MIPS 의 성능이면 충분하다.

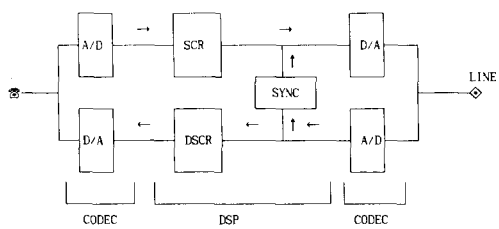


그림 7. 비화기의 HARDWARE 블록도
Fig. 7. The Block Diagram of Scrambler.

Ⅵ. 결론

지금까지 음성의 아날로그 비화 방식에 있어서 비화 신호의 비도를 높이고 복원 신호의 음질을 향상시키기 위한 많은 연구들이 수행되고 그 결과 여러가지 알고리즘들이 제안 되어왔다. 그러나 비도 향상을 위한 알고리즘은 복원 음질의 저하를 수반하고 반대로 복원 신호의 음질 향상 기법은 비화 신호의 비도 수준을 떨어뜨리고 있다.

본 논문에서는 비화 신호의 비도를 향상시키는 동시에 그에 따른 음질 저하를 최소화하여 복원 음성 신호의 통화 품질 수준을 유지하는 알고리즘을 제안하였다. 그리고 컴퓨터 시뮬레이션을 통한 실제 통화 라인에서의 청취 시험을 통하여 기존의 비화 방식들과의 성능 비교 평가를 행한 결과 그우수성이 입증되었다. 먼저 pre / post filtering 기법과 적응 의사 스펙트럼 삽입 기법의 적용은 비화 신호의 비도를 상당히 향상시켰으며 pre / post filtering 과 hamming window 기법의 채택으로 복원 음성 신호의 음질을 향상시킴은 물론 실제의 통화 라인상에서 통화 품질을 유지하는 내구성을 보였다. 또한 여기서 제안한 알고리즘은 ADSP-2101 DSP 하나의 칩을 이용하여 간단히 실시간 아날로그 음성 비화기를 구현할수 있음을 알 수 있었다. [4]

參 考 文 獻

[1] N. S. Jayant, "Analog scrambler for speech privacy. in Computer & Security", New York: North-Holland, pp. 275-289, 1982

[2] L. S. Lee and et al., "A simple sample value scrambler using FFT algorithms for secure voice communications," in Proc. NTC '80, pp. 49.4.1-49.4.5, 1980

[3] A. Matsunaga, K. Koga, and M. Ohkawa, "An analog speech scrambler using FFT technique with high-level security," *in Proc.* ICC'88, Philadelpia, PA, pp. 49.7.1-49.7.7, 1988

[4] ADSP 2100 Family User's Guide, Analog Device Incorporated, 1990.

[5] D. H. Cho, U. K. Un and J. W. Kim, "On the use of pre- and post-filters in speech waveform coding", *Journal of Acoust.*, Soc. of Korea, 4(21), 33-41, 1985

[6] Hasegawa T., Suzuki N Hakura Y., "A Proposal of Analog with Speech-Pseudonoise Signal for Higher Security against Eavesdropping", Technical

report of IEICE, vol.SSTA89-40, pp.33-37, Nov., 1989

[7] R. L. Rabiner and R. W. Schafer, "Digital Processing of Speech Signals", Prentice-Hall, Englewood Cliffs:NJ, 1987

[8] SAMSUNG Linear IC vol.3 (Telecom) DATA BOOK, Samsung Electronics Co.,1991.

[9] A. Shamir, "Identity-based cryptosystem and signature scheme," in Advances in Cryptology:Proceedings of Crypto'84. Berlin, West Germany:Springer-Verlag, 1985, pp. 47-53.

[10] Jon E. Natvig, "Evaluation of six Medium Bit-Rate Coders for the Pan-European Digital Mobile Radio System," in IEEE Journal on Selected Areas in Communications, vol. 6, no. 2, February, 1988, pp. 324-331

[11] A. Fukasawa et al., "An advanced 32 kbit/s ADPCM coding to transmit speech and high-speed voiceband data," in Proc. *IEEE ICASSP Conf.*, April 1986, pp. 821-824

著者紹介

孔 炳 球(正會員)

1961年 3月 9日生. 1984年 2月 한양대학교 화학공학과(학사). 1992年 9月 경희대학교 전자계산기공학과(석사). 1984年~1993年 2月 삼성전자 정보통신부분 종합연구소. 1993年 2月~현재 삼성종합기술원 선임연구원 . 주관심분야는 음성합성, 음성인식, 음성압축 등임.

趙 東 浩(正會員)

1956年 4月 3日生. 1979年 2月 서울대학교 공과대학 전자공학과 (학사). 1981年 2月 한국과학기술원 전기 및 전자공학과 (석사). 1985年 2月 한국과학기술원 전기 및 전자공학과 (박사). 1985년 3月~1987年 2月 한국과학기술원 통신공학연구실 선임연구원. 1987年 3月~현재 경희대학교 전자계산공학과 부교수. 1989年 9月~현재 경희대학교 전자계산소장. 주관심분야는 멀티미디어 통신, 무선 데이타 통신, 통신 소프트웨어 등임.