

극대 및 극소점에서서의 비균일 표본화에 의한 음성압축

Speech Compression by Non-uniform Sampling at the maxima and minima

임재열*, 백성준*, 안수길*, 김범훈**

(Jae Yeol Rheem*, SungJoon Baek*, Souguil Ann*, BumHoon Kim**)

요 약

리던던시(redundancy)가 존재하는 신호의 데이터 압축을 위한 한 방법으로 비균일 표본화 방법이 연구되어져 왔다. 그런데 많은 리던던시가 존재한다고 알려진 음성신호에 기존의 비균일 표본화 방법을 그대로 적용하며, 데이터량이 PCM과 같은 균일 표본화에 의한 것에 필적하게 되거나 오히려 많아지게 된다. 이 문제를 해결하기 위하여 본 논문에서는 음성의 인지 측면에서 리던던트한 샘플을 제거하는, 극대점과 극소점에서 비균일 표본화하는 방법을 제안한다. 그리고 본 논문에서는 제안된 방법의 성능을 분석하고자, 실험적인 방법으로 압축비 관계를 살펴봄, 아울러 묵음구간 처리에 의하여 압축비를 크게 향상시킬 수 있음을 보인다. 실험결과로써는 8kHz 8-bit PCM신호에 대하여 1.54배, 묵음처리된 한 경우에는 2.88배의 압축비를 얻었다.

ABSTRACT

To reduce the redundancy within samples that resulted from uniform sampling method, nonuniform sampling or nonredundant-sample coding methods can be considered. But it is well-known that when conventional nonuniform sampling methods are applied directly to speech signal, the amount of data required is comparable to or more than that required by uniform sampling method like PCM. To overcome this problem, we consider properties of speech signal in the sense of perception, and suggest a nonuniform sampling method at the maxima and minima of speech wave. To analyze the performance of the suggested method, compression ratio is considered. We show that compression ratio can be improved by silence detection, which can't be implemented by conventional methods based on uniform sampling. As experimental results, compression ratios of 1.54 without silence detection and 2.88 with silence detection for 8kHz 8-bit PCM signals are obtained.

I. 서 론

기존의 음성파형 표현 또는 코딩기법은 기본적으로 균일 표본화방법(uniform sampling method)에 관

기하여 음성신호에 내재하는 리던던시(redundancy)를 이용하는 방법에 따라 여러 양상화 방법이 분류되어 왔다[1]. 이러한 표현기법에 내재된 리던던시(inherent redundancy)는 균일 표본화에서 오는 샘플과 샘플간에 존재하는 상대적으로 높은 상관 관계(correlation)에서 기인하는 것으로 알려져 있다.[1]. 따라서 음성의 저장이나 전송을 위한 저장량이나 선

*서울대학교 전자공학과

**삼호전자

접수일자: 1992. 5. 30

송률을 줄이기 위해서는, 시간 축상에서 존재하는 리던던시를 줄이고자 샘플과 샘플간의 상관관계가 적은 방향으로 표본화를 하는 비균일 표본화방법(non-uniform sampling method, or nonredundant-sample coding)을 고려할 수 있다.[2]-[6] 그러나 비균일 표본화 방법으로 개발된 기존의 폴리노미알 예측기나 팬알고리즘과 같은 보간기(polynomial predictor or interpolator)같은 방법들은 음성신호에는 적합하지 않는다. 왜냐하면, 음성 파형은 상대적으로 그 변화가 급격하고 또한 지속적이므로, 기존의 방법에서와 같이 표본과 표본간의 크기나 기울기의 차를 이용하여 비균일 표본화를 수행할 경우, 요구되는 데이터량이 PCM과 같은 균일 표본화에 의한 것에 필적하거나 그 이상이 되기 때문이다.[6] 따라서 음성 신호를 비균일 표본화 방법으로 표현하여 데이터량을 줄이기 위해서는 음성의 특성에 맞는 새로운 방법을 개발해야 할 필요가 있다.

음성표현이나 코딩에 있어서의 주된 목적은 음성이 전달하는 정보중 특히 언어 정보의 전달이므로, 음성인지(perception of speech)측면에서 음성의 리던던시 특성을 살펴보자. 음성 파형을 극단적으로 왜곡시켜서 만든 파형을 가지고 인지실험을 한 것을 살펴보면, 음성의 인지를 유지하는 중요한 요소가 극대·극소점간의 간격(interval between maxima and minima, or local peak-to-peak interval)임을 알 수 있다.[7][8] Licklider와 Pollack의 인지실험의 경우, 원래의 음성 파형을 미분하고 극단적으로 클리핑시켜서 만든 음성파형의 인지도 값(intelligibility score)이 97%로써 원래의 음성파형의 인지도 99%에 비해 인지도의 저하가 거의 일어나지 않았다.[7] 원래의 파형을 미분하여 극단적으로 클리핑한 경우에는 그림1과 같이 일정한 크기를 가진 펄스열이 되며, 펄스열의 극성이나 모양이 음성의 인지도에 영향을 미치지 않는다[8].

신호를 미분하고 극단적으로 클리핑시켰을 때 남아 있는 유일한 정보는 극대점과 극소점의 발생시각이므로 극대·극소점 사이의 샘플들은 인지 측면에서 리던던트한 샘플임을 알 수 있다. 이러한 사실은 극대·극소점간의 간격이 주파수 영역에서의 주파수와 밀접한 관련이 있다는 사실과 우리의 인지작용이 각각의 주파수 처리에 의존한다는 사실과 잘 부합된다. 그리고 우리는 극대·극소점간의 간격이, 주파수 영역에서의 주파수 성분을 구성하는 시간영역에서의 최소 구성요소이므로, 시간축상에서 음의 인지를 결정하는 음의 기본요소(sound element)임을 알

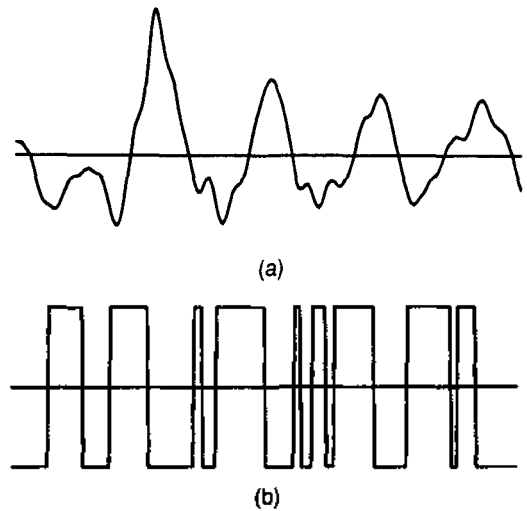


그림 1. 왜곡된 파형
(a)원래의 파형 (b)미분및 클리핑된 파형.
Fig.1. Amplitude distorted waveform
(a)original waveform
(b)differentiated and clipped waveform.

수 있다.

본 논문에서는, 음성의 인지가 기본적으로 음성 파형의 극대점과 극소점간의 간격에 의존하고 있다는 사실에 근거하여, 음성신호를 극대점및 극소점에서 비균일 표본화하여 표현하는 음성 표현 방법을 제안한다. 그리고 음성신호에서의 극대·극소점간 간격의 분포를 구하고, 표본화 주파수에 따른 압축비를 살펴본다. 제안된 방법은 비균일 표본화에 의해서 기본적으로 비동기적인(asynchronous) 성격을 지니게 된다.

II. 극대·극소점에서의 비균일 표본화 방법

극대·극소점의 간격 정보만으로도 음성의 인지가 가능하므로, 균일 표본화 시에 나타나는 극대·극소점 사이에 나타나는 샘플들은 음성인지 측면에서는 리던던트한 샘플로 간주할 수 있다. 따라서 음성인지의 측면에서 음성 파형을 나타낼 경우 극대·극소점에서 비균일 표본화하여 나타낼 수 있다. 이때 표본 간격이 일정하지 않으므로, 표본시의 샘플값및 앞 샘플과의 간격을 시간정보로써 저장해 주어야 한다. 재구성 시에는 샘플과 샘플을 직선으로 연결해 주는 선형 보간(linear interpolation)을 하면 된다.

본 논문에서는 균일 표본화된 신호와 비교하기 위하여 PCM신호에서의 비균일 표본화를 고려하기로

한다.

2.1)PCM 신호에서의 극대-극소점의 고찰

PCM 신호 $x(n)$ 에서 극대-극소점을 검출하는데는 표본화 및 양자화에 따른 영향으로 발생하게되는 기울기 0인 샘플들을 고려해야 한다. 그림 2에서 (a)의 경우에는 전형적인 극대점으로 쉽게 결정지을 수 있으나, (b), (c)인 경우에는 기울기가 0인 점들의 고려가 필요하다. 이러한 기울기가 0인 점들은 표본화 및 양자화의 영향으로 극대-극소점이 분해되지 않은 경우로 볼 수 있으며, 특히 (c)의 경우, 실제환경에서는 A/D변환시 과부하(overload)에 의한 클리핑(clipping)현상의 결과일 수도 있다. 원래의 연속신호 $x_c(t)$ 에서는 (b)의 경우에는 극대, 극소점이, (c)의 경우에는 한개의 극대점이 존재했을 것이다.

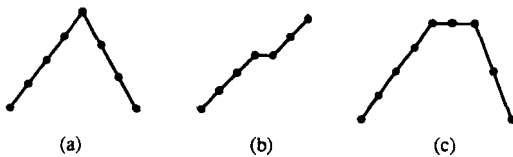


그림 2. PCM 신호에서의 극대점 및 극소점의 고려

- (a) 정상적인 극대점
 - (b) 양자화 오류로 분해되지 않은 극대, 극소점
 - (c) 양자화 오류 또는 과부하에 의해서 분해되지 않은 극대점
- Fig.2. Considerations on maxima and minima of PCM signal
- (a) normal maxima
 - (b) unresolved maxima and minima by quantization error
 - (c) unresolved maxima by quantization error or overload clipping.

2.2)제한된 비균일 표본화 방법

그림 3에 제안된 비균일 표본화 방법의 알고리즘이 있다. 극대-극소점의 결정은 연속한 두 샘플의 기울기의 값을 이용하여 결정한다. 연속한 두 샘플의 기울기의 곱 값이 0보다 크면 상승, 또는 하강 중임을 의미하므로 다음 샘플을 읽어들이며, 그 값이 0보다 작으면 정상적인 극대 또는 극소점을 의미하므로 그때의 샘플값 자체와 이전 샘플값과의 간격을 차례로 저장한다. 기울기의 곱이 0인 경우에는 기울기 0인 점들을 나타내게 되므로 그림 2의 (b), (c)의 상태를 고려하여야 하며, 이를 위하여 상태 변수(status)를 도입

입하였다. 정상적인 극대 또는 극소점(그림 2의 (a))인 경우에는 상태변수에 0값을, 기울기가 0인 직선에 이르면 직전 기울기를 저장하여 기울기 0인 점들의 상태를 판단하게 된다. 그리고 전체 시간정보를 유지하기 위하여 처음 샘플과 마지막 샘플을 극대-극소값에 관계없이 저장하기로 한다.

본 알고리즘에서 고려한 사항은 다음과 같다.

- (1)그림 2의 (b)의 경우에서 기울기 0인 점들의 존재여부는 표본화 주파수 ω_s 와 양자화 레벨수 L 에 종속되지만, 음성파형을 관찰하여 보면 그 간격이 상대적으로 작으므로 제안된 알고리즘에서는 무시하기로 한다. 실제로 고려했을 경우와 무시한 경우를 비교해 보면, 1% 미만의 데이터량 차이를 보이며 재구성된 신호에서 음질의 차이를 크게 느낄수 없다.
- (2)그림 2의 (c)의 경우에는 기울기 0인 점들 중 가운데것을 극대점으로 선택하였다. 기울기 0인 점들 중 처음 점이나 마지막 점도 극대점으로 가능하며, 마지막 점이 구현하기 쉬우나, 각 경우에 대하여 실험을 해보아도 인지측면에서는 큰 차이를 보이지 않기 때문에, 제안된 알고리즘에서는 기존과형을 최대한 유지한다는 측면에서 가능한 한 가운데 점을 택하기로 한다.

2.3)재구성(reconstruction) 방법

재구성은 각각의 비균일 표본화 점들 값의 간격을 고려하여 비균일 표본화의 일반적인 재구성 방법인 선형 보간을 해주면 된다¹⁾. 비균일 표본화된 k번째 샘플의 크기와 간격을 각각 $a(k)$ 와 $i(k)$ 라 하면, 이전 샘플 크기 $a(k-1)$ 이후에서 $a(k)$ 까지 선형 보간에 의해서 재구성된 신호 $y_k(n)$ 은

$$y_k(n) = \left[- \frac{a(k) - a(k-1)}{i(k)} \cdot n + a(k-1) \right]_L, 1 \leq n \leq i(k) \tag{1}$$

으로 나타낼 수 있다. 여기서 $[\]_L$ 은 L 레벨의 양자화 함수이다.

선형보간법을 이용하여 비균일 표본점들 사이를 직선으로 연결하여 재구성하게되면, 원래의 파형이 식별 이해도와 음색이 유지되나 원래의 파형과는 차이를 보이게 된다. 자음과 묵음의 경우에는 직선으로 나타내어도 원래의 파형과는 큰 차이를 보이지 않겠지만 음성을 특히 비음과 같이 사인파형에 가까운 파형의 경우에는 파형상의 차이가 매우 크게 된다. 따

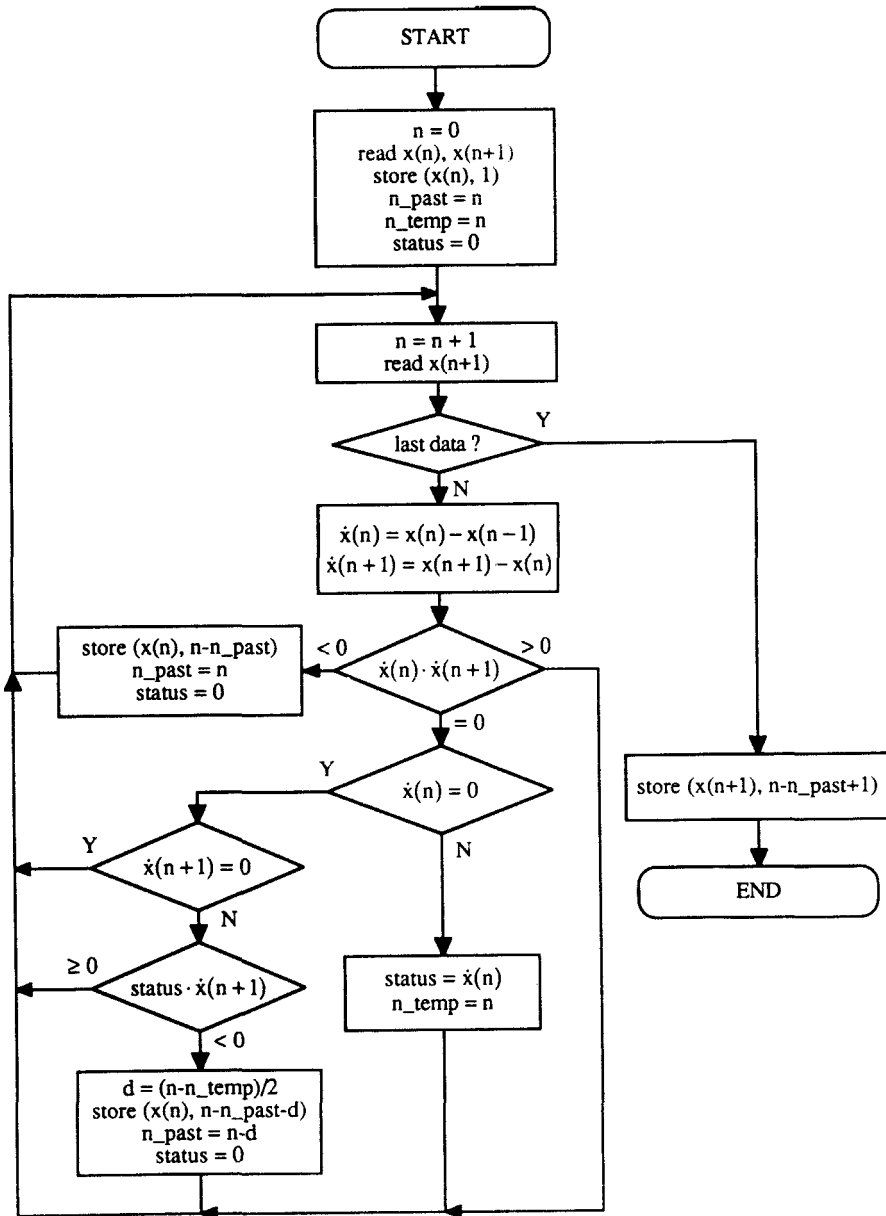


그림 3. 비균일 표본화 알고리즘
Fig.3. Nonuniform sampling algorithm

라서 이를 보완하기 위하여, 코사인 파형의 반주기를 이용하여 근사시킴으로써, 유성음의 경우 원래의 파형에 가깝도록 재구성을 할 수 있다. 이때 재구성된 신호 $y_k(n)$ 은

$$y_k(n) = \left[\frac{a(k-1) - a(k)}{2} \cdot \cos\left(\frac{\pi}{i(k)} \cdot n\right) + \frac{a(k-1) + a(k)}{2} \right]_L, \quad 1 \leq n \leq i(k) \quad (2)$$

으로 표현된다. 이때 유성음과 비유성음을 구별할 필요는 없다. 유성음이 아닌 경우에는 그 간격이 대체로 작기 때문에 크기와 관계없이 선형보간사와 큰 차이가 없기 때문이다.

III. 압축비

제안된 방법의 성능을 분석하기 위해서 압축비(compression ration)를 계산해보자. 압축비를 해석적으로 구하기 위해서는 극대-극소점간의 간격의 확률분포를 알아야 한다. 대상 신호가 정상(stationary) 신호이고, 균일 표본화하였을 경우에 발생하는 샘플과 샘플간의 발생확률분포 및 그 조건확률이 주어진다면 해석적으로 간격의 분포를 구하여 압축비를 계산할 수 있다^[6].

그런데 음성신호의 경우에는 샘플간의 발생 조건확률 모델이 정확하게 알려져 있지 않으며^[10], 실사 10차 이상의 마코프 과정으로 음성 신호가 모델링된다고 할지라도, 계산량 및 기억용량측면에서 그 조건확률을 구하는 것이 쉽지 않다^[6]. 그리고 음성신호는 기본적으로 비정상적인(nonstationary) 신호이다. 따라서 본 논문에서는 음성신호의 간격에 대한 확률분포를 상대적 빈도수(relative frequency)에 의하여 실험적으로 구하고 그것으로부터 평균간격을 구하여 압축비를 구하기로 한다.

주어진 신호가 특정 표본화 주파수로 균일 표본화되어 1에서 L까지의 값으로 균일 양자화되어 있고, 여기에 제안된 방법을 이용하여 비균일 표본화 하여 얻은 간격이 1에서 T의 범위의 값을 가진다고 가정하자. N개의 음성 샘플로부터 간격 $n(1 \leq n \leq T)$ 에 대한 상대적 빈도수에 의한 확률이 $p_N(n)$ 으로 구해진다면, $N \rightarrow \infty$ 일 경우 일반성을 잃지 않고 $p_N(n) \rightarrow p(n)$ 이 될 것이다. 이때 평균간격 \bar{n} 은

$$\bar{n} = \sum_{n=1}^T n \cdot p(n) \quad (3)$$

으로 표현된다. 압축비는 균일 표본화시의 평균비트수 대 비균일 표본화시 비트수로 정의되므로, 비균일 표본화시의 평균간격에 대한 균일 표본화시의 비트수 대 비균일 표본화로 얻은 한 샘플을 나타내는데 필요한 비트수로 나타낼 수 있다. 따라서 제안된 방법의 압축비는

$$C = \frac{\bar{n} \cdot \log_2 L}{\log_2 T + \log_2 L} \quad (4)$$

IV. 묵음구간 처리에 의한 압축향상

음성신호에 대한 극대-극소점간의 간격의 확률분포를 구하게 되면 간격이 짧은 구간에서 확률이 큼을 알 수 있다. 이것은 음성신호중 많은 부분을 차지하는 묵음 구간과 무성자음 구간의 불규칙성으로 설명될 수 있다. 그런데 묵음 구간에는 음성의 언어적 정보가 없으므로 묵음 구간을 검출하여 하나의 간격으로 표시하게 되면, 상대적으로 많은 압축 효과를 볼 수 있으며, 이러한 압축방법은 기존의 균일 표본화 방법에서는 전송측면에서 별도의 도움이 없이 직접적으로 구현될 수 있는 것은 아니다.

묵음 구간의 검출은 일정한 폭을 가지는 윈도우를 이용하여, 윈도우내에서 간격 및 샘플값의 쌍들의 통계적 성질을 이용하여 결정할 수 있다. 이때 주의해야 할 점은 무성자음 구간이 포함되지 않도록 하기 위하여 순수 묵음 구간만을 검출해야 하는 것이다.

4.1) 묵음 구간 검출에 의한 비균일 표본화

묵음 구간 검출에 의한 비균일 표본화 알고리즘은 다음과 같다. 그림 5에서와 같이 길이 W의 윈도우를 잡고, 윈도우내의 극대점 및 극소점을 그림 3의 알고리즘에 의해서 결정하여, 간격과 샘플 값의 쌍으로 비균일 표본화하여 나타낸다. 이 순서쌍의 통계적 성질을 조사하여 묵음으로 판정이 난 경우에는, 사이에 있는 샘플을 무시하고, 윈도우 내의 처음과 마지막 샘플을 직선으로 연결하여 나타낸다(그림 4의 w-th 윈도우). 묵음 구간이 아닌 것으로 판정이 난 경우에는 윈도우 내의 모든 극대점과 극소점의 간격과 샘플 값을 저장한다(그림 4의 (w+1)-th window).

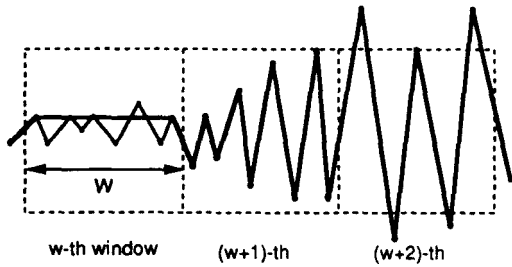


그림 4. 묵음 구간 검출에 의한 비균일 표본화
Fig.4. Nonuniform sampling by silence interval detection.

V. 실험 및 결과

간격의 빈도수를 측정하기 위하여 음성신호를 3.8kHz 저역필터링한 후 균일표본화(8, 10, 14, 20, 21, 40, 그리고 42kHz)와 양자화를 거쳐(8-bit PCM) 얻은 신호에 대하여 제안된 비균일 표본화 알고리즘을 수행했다. 이렇게 하여 얻어진 크기와 간격의 쌍을 저장한 후 간격의 발생빈도수를 추정했다. 통계작업에 사용된 음성시료는 3명의 20대 남성화자(JDI, SBL, NKM)와 1명의 20대 여성화자(JYM)가 책을 읽은 것을 3분씩 녹음하여 사용하였다. 각 화자의 책을 읽는 속도는 JDI, SBL, NKM, JYM이 각각 5.18, 5.49, 5.89, 4.78 음절 / 초(syllable / sec)이다.

그림 5에는 원래의 파형과 제안된 비균일 표본화 방법에 의해서 재구성된 파형을 보여준다. 선형보간에 의해서 재구성된 신호는 톱니파형과 유사하게 되어 유성음 부분에서 원래의 파형과 차이를 보이나, 코사인 파형으로 재구성한것은 원래의 파형에 가깝게 근사됨을 알 수 있고, 묵음 구간의 경우에도 큰 차이가 없음을 알 수 있다. 그러나 재구성된 신호를 청취 테스트하여 보면, 이해도의 차이는 전혀 발생하지 않으며, 음질의 차이도 거칠어 졌다는 느낌을 확실히 받으나, 그 차이가 크게 느껴지지는 않는다. 그리고 선형보간법에 의한 것이 상대적으로 조금 더 거칠게 느껴진다.

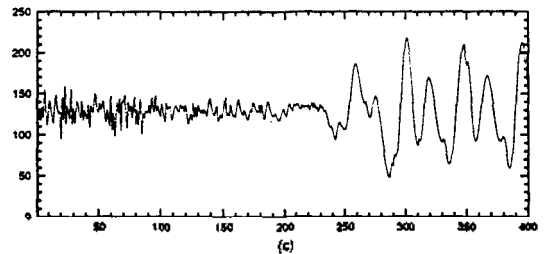
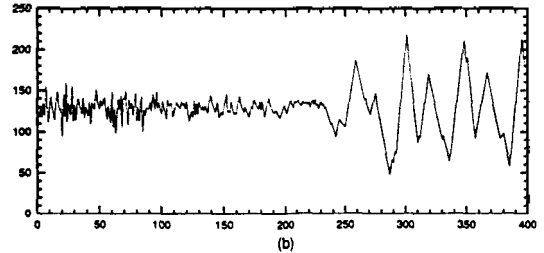
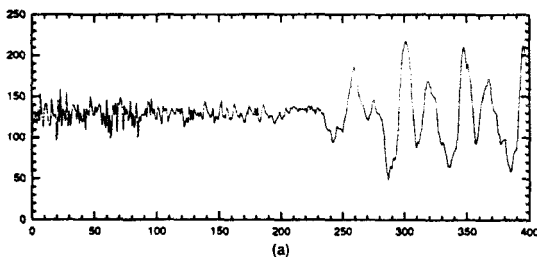
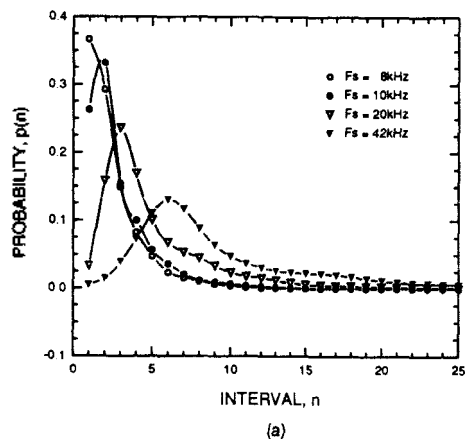


그림 5. 재구성된 신호
(a)원래의 PCM 파형 (b)선형보간법에 의한 파형
(c)사인함수에 의한 파형.

Fig.5. Reconstructed signal
(a)original PCM data
(b)reconstructed signal by linear interpolation
(c)reconstructed signal by sinusoidal interpolation.

그림 6에는 제안된 그림 3의 알고리즘에 의한 확률분포 및 압축비 계산결과를 보여준다. 확률분포의 경우 표본화 주파수에 따른 간격의 분포 결과가 잘 나타나고 있다. 주목할 만한 것은 40kHz 표본화시의 압축비가 약 5.4로써 8kHz 균일 표본화 시보다 적은 데이터양과 명료한 음질로 데이터를 저장할 수 있다는 것을 알 수 있다. 8kHz 표본화시의 압축비는 약 1.54로써 약 30%의 압축효과가 나타난다.



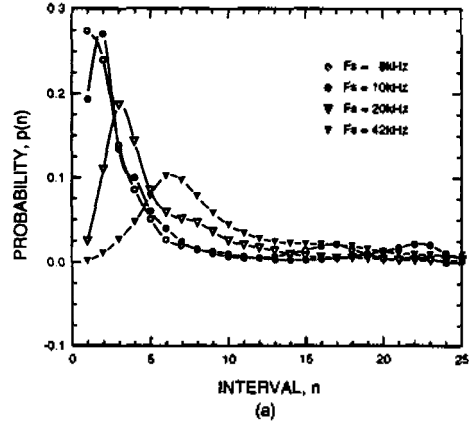
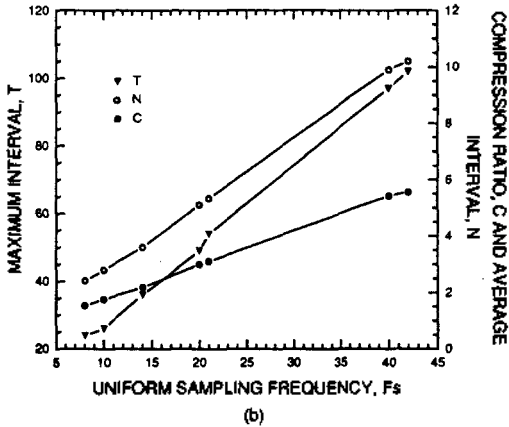


그림 6. 제안된 알고리즘의 결과
 (a)간격의 확률분포
 (b)최대간격T, 평균간격N, 압축비 C의 관계.
 Fig.6. Results of proposed algorithm
 (a)probability distribution of interval
 (b)relation among T, N, and C.

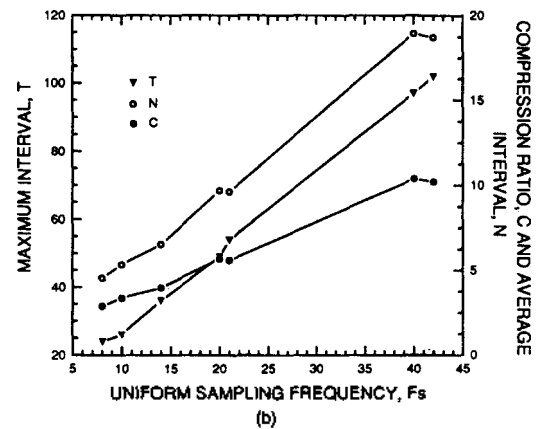


그림 7에는 묵음구간 처리를 이용한 결과를 보여준다. 실험시 사용된 윈도우의 크기는 $W=T$ 이며, 묵음구간의 결정에 사용된 파라미터는 윈도우 내의 비균일 표본화 샘플의 개수 및 평균크기로서, 비균일 표본화 개수가 6개 이상이고 평균크기가 5이하인 경우에는 묵음구간으로 판정했다. 이것은 균일표본화시에 많이 사용되는 영교차율(zero-crossing rate) 및 평균크기(average magnitude)에 해당하는 파라미터로 볼 수 있다. 확률분포의 경우 묵음처리에 의한 간격 분포의 변화가 명확히 나타난다. 그 결과 압축비가, 묵음구간을 처리하지 않았을 때에 비하여 평균 1.85배 향상되었으며, 8kHz 표본화시에는 PCM에 비하여 약 65%의 데이터량 감소가 발생되었다. 묵음구간 처리를 사람이 수작업으로 처리할 경우에는 8kHz 표본화시에 약 75%까지도 데이터량의 감소가 가능하게 되는데, 이것은 묵음구간의 결정알고리즘을 향상시킬 수 있으면 압축비를 더욱 향상시킬 수 있다는 것을 의미한다. 묵음구간만을 정확히 검출하여 처리한 경우에는 묵음구간을 처리하지 않았을 때에 비하여 재구성된 음성파형의 음질저하가 일어나지 않는다. 묵음구간의 검출시 사용된 문턱값을 높이면 결과적으로 압축비가 높아지며, 어느 한계까지는 음질의 저하가 발생하나 이해도가 떨어지는 않는다. 이것은 윈도우의 크기를 높이는 경우에도 마찬가지이다.

그림 7. 묵음처리의 결과
 (a)간격의 확률분포
 (b)최대간격T, 평균간격N, 압축비 C의 관계.
 Fig.7. Results of silence detection
 (a)probability distribution of interval
 (b)relation among T, N, and C.

표1에 묵음구간 처리에 따른 압축비 실험 결과가 정리되어 있다.

표 1. 묵음구간처리 전후에 따른 압축비 관계

표본화 주파수	간격 최대값	묵음구간 처리전		묵음구간 처리후	
		평균간격	압축비	평균간격	압축비
8	24	2.43	1.54	4.52	2.88
10	26	2.80	1.76	5.34	3.36
14	36	3.60	2.20	6.52	3.96
20	49	5.11	3.00	9.68	5.69
21	54	5.34	3.11	9.61	5.59
40	97	9.90	5.43	18.94	10.38
42	102	10.20	5.56	18.71	10.20

객관적으로 음질을 평가하기 위하여, 화자 JDI의 데이터에 대하여 원래의 PCM 신호와 선형보간 및 코사인 파형 보간으로 재구성한 신호에 대한 세그멘탈 SNR^[11]을 구했다. 그 결과 각각 11.4dB, 11.3dB를 얻어 주관적인 음질 평가와는 큰 차이를 보여 줄 수 있다. 원래의 PCM 신호와 재구성된 신호는 오직 비균일 표본화 값에서만 일치하고 리던던트한 구간에서는 일치하지 않으므로 상대적으로 낮은 SNR 값이 얻어진다.

임의의 음성을 비균일 표본화하여 저장할 경우에도 위의 실험 결과와 잘 부합된다. 실제의 경우에는 녹음 환경과 화자의 발생속도, 발생음량, 음질에 따라 개인차가 심하게 나타나게 되나, 그 평균값은 사람에 따라 일정하며, 특히 복음을 처리한 경우에는 압축비가 화자에 관계없이 거의 일정하게 유지됨을 관찰할 수 있다.

VI. 결 과

본 논문에서는 음성의 인지 측면에서 음성과정의 극대값과 극소값 사이의 샘플들이 리던던트하다는 사실에 근거하여, 극대 및 극소점에서의 비균일 표본화 방법을 제안했다. 제안된 방법은, 기존의 비균일 표본화 방법을 그대로 음성에 적용하였을 경우에 요구되는 데이터 양이 PCM과 같은 균일 표본화 방법에 근거한 것에 필적하거나 오히려 증가하는 단점을 극복하여, 음성의 경우에도 비균일 표본화 방법으로 적은 데이터량으로 저장 또는 전송이 가능함을 보였다. 데이터량의 감소의 결과로 약간의 음질 저하가 발생하나 음성의 인지도에는 영향을 주지 않는다. 그리고 음성의 상당한 부분을 차지하는 복음구간을 처리할 경우에는 압축비를 표본화 주파수에 따라 약 1배까지 향상 시킬 수 있으며, 이러한 복음구간의 처리는 균일표본화 방법으로는 직접적으로 구현할 수가 없다.

제안된 방법의 장점은 간단한 방법으로 데이터량의 감소를 얻을 수 있다는 것과, 개인용 컴퓨터에서도 특별한 장치없이 실시간 처리가 가능하다는 점들을 들 수 있다. 특히 제안된 방법으로 얻어지는 극대 및 극소점간의 간격은 음성 신호처리에 있어서 시간영역에서의 파라미터로서의 가능성이 있다. 극대 및 극소점간의 간격을 살펴보면, 음성의 크기는 발음시간에 따라 그 변화가 심한데비하여, 간격은 발음에 따라 그 차이가 심하지 않고 일정하게 유지되는 것을 관찰할 수 있으며, 표본화 주파수가 높을수록 분명하

게 나타나게 된다. 이것은 극대 및 극소점간의 간격이 주파수영역에서 주파수 성분을 구성하는 시간영역의 기본요소라는 것과 같은 발음일 경우에는 발생시간에 따른 시간영역에서의 변화보다 주파수 영역에서의 변화가 작다는 사실로써 알 수 있다. 따라서, 음성 신호의 시간영역처리에서 주로 사용되어진 에너지, 영교차율과 같은 파라미터가 주로 사전처리과정에서 사용되어졌던 것에 비하여, 극대 및 극소점간의 간격은 본격적인 분석 및 인식과정에서도 사용가능한 것으로 보인다.

제안된 방법의 단점은 비균일 표본화 방법에 근본적으로 존재하는 비동기적인 데이터가 얻어진다는 점과 요구되는 데이터 양이 개인의 음질이나 발생환경에 종속된다는 점이다. 비동기적인 변은 전송시의 버퍼링과 리던던트 샘플(nonredundant sample)의 시간코딩(TNRS coding) 방법등을 통하여 극복할 수 있다 할 지라도, 발생화자에 따른 데이터량의 종속성은 전송측면에서 앞으로 연구되어야 할 것이다. 저장측면에서는 비동기적인 측면과 데이터량의 화자 종속성이 큰 문제가 되지 않으므로 개인용 컴퓨터를 이용한 음성정보의 저장 및 전송에는 제안된 알고리즘이 매우 유용하리라고 본다.

참 고 문 헌

1. N.S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, Prentice-Hall, Inc., 1984.
2. T.J. Lynch, *Data Compression: Techniques and Applications*, Lifetime Learning Pub., 1985.
3. J.W. Mark and T.D. Todd, "A Nonuniform Sampling Approach to Data Compression," *IEEE Trans. on Com.*, Vol.COM-29, No. 1, pp.24-32, Jan. 1981.
4. L. Ehrman, "Analysis of Some Redundancy Removal Bandwidth Compression Techniques," *Proc. IEEE*, Vol.55, No.3, March 1967.
5. T.J. Lynch, "The Probability of a straight-line Sequence from a Uniform Independent Sample Source," *IEEE Trans. on Info. Theory*, Vol.IT-14, No. 5, pp 773-774, Sept. 1968.
6. L.D. Davisson, "Data Compression Using Straight Line Interpolation," *IEEE Trans. on Info. Theory*, Vol.IT-14, No. 3, pp.300-304, May 1968.
7. J.C.R. Licklider and I. Pollack, "Effects of Differentiation, Integration, and Infinite Peak Clipping upon the Intelligibility of Speech," *J. Acoust. Soc. Am.*, Vol. 31, 1490, 1948.

8. W.A. Ainsworth, "Relative Intelligibility of Different Transforms of Clipped Speech," *J. Acoust. Soc. Am.*, Vol.41, pp.1272-1276, 1967.

9. R.A.Cole, A.I.Rudnicky, V.W.Zue and D.R.Reddy, "Speech as Patterns on Paper," *Perception and Production of Fluent Speech*, LEA, 1980.

10. W.B. Davenport, Jr, "An Experimental Study of Speech-wave Probability Distributions," *J. Acoust. Soc. Am.*, Vol. 24, No.4, pp.390-399, July 1952.

11. S. Furui and M.M. Sondhi, *Advances in Speech Signal Processing*, Marcel Dekker, Inc., 1992.

▲임 재 열

1986년 2월 : 서울대 전자공학과 졸업(공학사)
 1988년 2월 : 서울대 대학원 전자공학과 졸업(공학석사)
 현 재 : 서울대 대학원 전자공학과 박사과정

▲백 성 준



1989년 2월 : 서울대 전자공학과 졸업(공학사)
 1991년 2월 : 서울대 대학원 전자공학과 졸업(공학석사)
 현 재 : 서울대 대학원 전자공학과 박사과정

▲안 수 길

현 서울대학교 전자공학과 교수(9권 1호 참조)

▲김 범 훈

삼호전자 상무이사