

그림 1-1. 시간상에서의 음성신호 "아".
Fig.1-1. Speech signal "AH"
in the time domain.

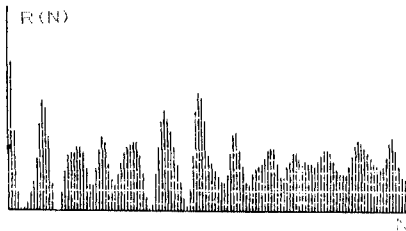


그림 1-2. "아"의 자기상관함수.
Fig.1-2. Autocorrelation function of "AH".

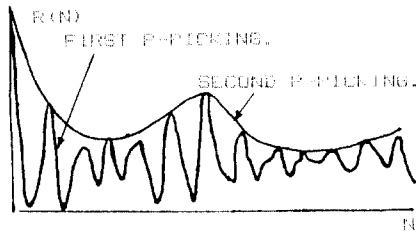


그림 1-3. "아"의 피크 피킹에 의한 피치검출.
Fig.1-3. Pitch detection of "AH" by
the peak-picking method.

2-3. 시간상에서 신호 압축

TDHS는 고조파를 이용하는 기술이며 신호 압축도 고조파 즉 피치와 관련이 상당히 크다. 신호압축을 위해서는 시간 축 상의 범위를 정해야하는데 이를 위해서 윈도우(window)를 채택하고 윈도우의 길이는 pitch와 관련이 있다. 또한 윈도우의 형태도 중요한 역할을 하여 시간상에서 신호압축은 삼각 윈도우(triangle window), $w(m)$ 와 그 보완도우(complement win-

dow), $w1(n)$ 를 채택한다.

윈도우의 형태를 수식적으로 표현하면 다음과 같다.

$$w(m) = 1 - m/p \quad m = 0, 1, \dots, p-1$$

$$w1(n) = 1 - w(m) = n/p \quad p = 0, 1, \dots, p-1$$

$$p = \text{pitch}$$

원래의 신호를 $s(m)$ 이라하면 압축된 신호 $sc(m)$ 는 다음과 같다.

$$\begin{aligned} sc(m) &= s(m) * w(m) + s(m+p) * w1(m) \\ &= s(m) * w(m) + s(m+p) * (1 - w(m+p)) \end{aligned}$$

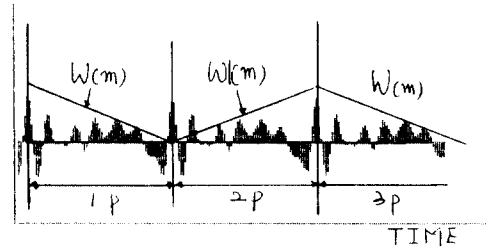


그림 2-1. "아"의 원래신호
Fig.2-1. Original signal of "AH".

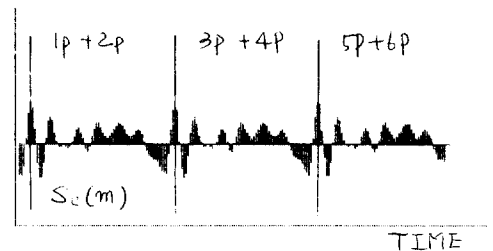


그림 2-2. "아"의 압축신호
Fig.2-2. Compressed signal of "AH".

그림 2-1.에서 제1 pitch는 윈도우 $w(m)$ 을 제2 pitch는 $w1(m)$ 을 택하고 제3 pitch는 $w(m)$ 를, 제4 pitch는 $w1(m)$ 을 계속 반복하여 택한다. 그림 2-2.에서는 $w(m)$ 으로 택하여진 제1 pitch에서 m 에 해당하는 데이터와 $w1(m)$ 으로 택하여진 제2 pitch의 $p+m$ 의 데이터가 더해진 것이다.

2-4. 시간상에서 신호 확장

확장에서의 윈도우는 압축에서의 윈도우의 경우와 마찬가지로 삼각윈도우를 택하는데 윈도우의 길이가 다르고 순서도 반대이다.

$$\bar{w}(m) = w_1(m) = 1 - w(m) = m/2p$$

$$\bar{w}_1(n) = w(n) = 1 - n/2p = 1 - \bar{w}(n+2p)$$

$$m, n = 0, 1, \dots, 2p-1$$

여기서 p 는 pitch이며 원도우 길이는 $2p$ 가 되는데 이유는 앞의 압축된 값들은 두 피치를 한 피치로 압축된 것이므로 확장할 때는 중복으로 계산 되어야 하며 두 피치를 주기로 중복계산 한다. 압축된 신호 $sc(m)$ 은 앞절에서 얻어졌고 확장된 $se(m)$ 은 다음과 같다.

$$\begin{aligned} se(m) &= sc(m) * w(m) + sc(m+2p) * \\ &w_1(m) \\ &= sc(m) * w(m) + sc(m+2p) \\ &* (1 - w(m)) \end{aligned}$$

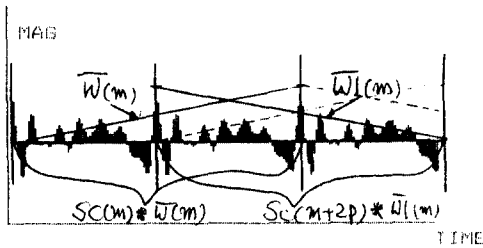


그림 2-3. "아"의 압축신호

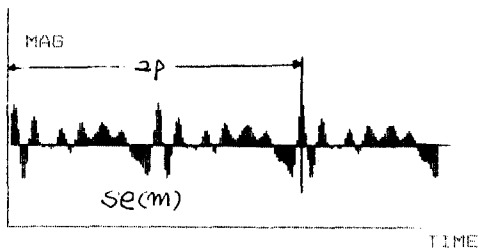


그림 2-4. "아"의 확장신호

Fig. 2-3. Compressed signal of "A".

그림 2-3.은 "아"의 압축신호이고 그림 2-4.는 확장신호를 보여 주는데 계산방법은 압축방법과 비슷한대원도우의 길이는 $2pitch$ 이며 중복계산하는 것이 다르다.

3. 실험 및 검토

연속 유성음 아-우(AH-WOO)를 대상으로 약 640개의 데이터를 가지고 처리하였다. 그림4에서 보듯이 압축된 데이터의 주파수 스펙트럼(frequency spectrum)을 보면 원래 신호와는 고주파성분에서 큰 차이를 갖는다. 다시 말하면 압축된 신호는 고주파 성분이 증가된다. 이유는 원래신호의 앞의 피치 동안의 신호 그 다음 피치동안의 신호변화가 완만함에 반하여 압축된 신호의 피치간의 변화는 급격하기때문에 고주파 성분이 증가된다. 확장된 신호는 원래의 신호와 시간상에서나 주파수상에서 거의 같게 나타난다. 결국 압축된 신호의 데이터량은 반으로 줄어들었지만 원래의 신호를 찾는다는 어려움이 없다.

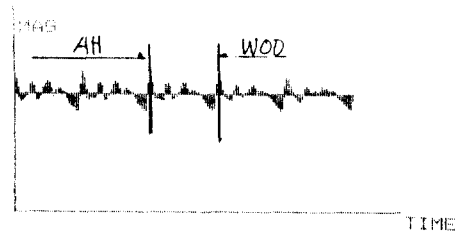


그림 3-1. "아-우"의 원래신호

Fig. 3-1. Original signal of "AH-WOO".

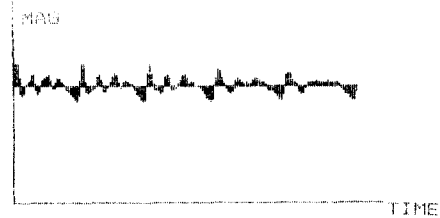


그림 3-2. "아-우"의 압축신호

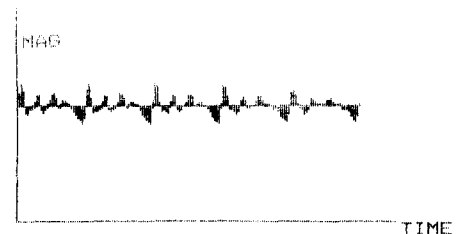


그림 3-3. "아-우"의 확장신호

Fig. 3-3. Expanded signal of "AH-WOO"

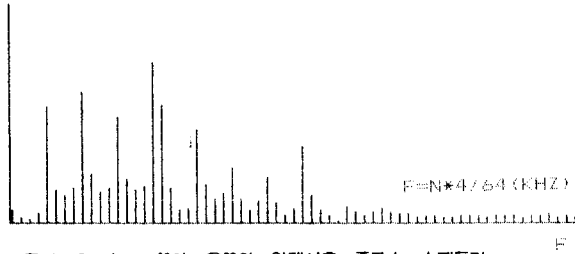


그림 3-4. "아-우"의 원래신호 주파수 스펙트럼
Fig.3-4. Freq.spectrum of original signal for "AH-WOO".

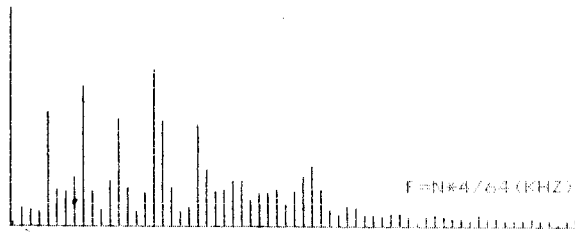


그림 3-5. "아-우"의 압축신호 주파수 스펙트럼
Fig.3-5. Freq.spectrum of compressed signal for "AH-WOO".

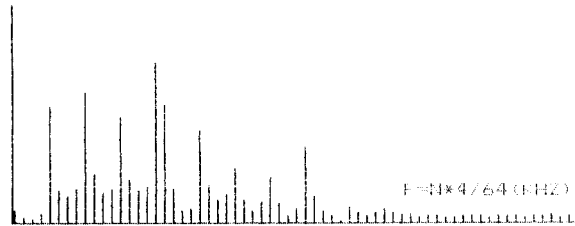


그림 3-6. "아-우"의 확장신호 주파수 스펙트럼
Fig.3-6. Freq.spectrum of expanded signal for "AH-WOO".

4. 결 론

많은 양의 데이터를 적은 오차로 줄인다는 것은 경제적인 면에서 큰 잇점을 갖는다. 그러한 면에서 연구는 계속 진행되고 본 논문에서 제시한 시간상에서 연속 음성용의 압축, 확장은 데이터를 반으로 줄일 수 있다는 것을 보여준다. 음성 뿐만 아니라 주기적이거나 준주기적인 어떠한 신호든지 압축이 가능하다.

5. 참고문헌

1. Digital Processing of Speech Signal : Rabiner/Schfer.
2. Speech Analysis, Synthesis and Perception : Flanagan
3. Computer Speech Processing : Fallside/Woods
4. Digital signal processing : Oppenheim/Schafer
5. Real Time Implementation of Time Domain Harmonic Scaling of speech for Rate Modification and Coding : Cox /Crochiere/Johnson
- IEEE.Trans on ASSP V.31 NO.1 FAB.1983
6. Performance of Transform and Subband Coding system combined with Harmonic Scaling of Speech : Malah/Crochiere/Cox
- IEEE.Trans on ASSP V.29 NO.2 APR.1981.