

# 융복합 시스템의 8kbps에 있어서 APC-MPC에 관한 연구

이시우  
상명대학교 정보통신공학과

## A Study on APC-MPC in 8kbps of Convergence System

See-Woo Lee

Dept. of Information and Telecommunication Engineering

**요 약** 유성음원과 무성음원을 사용하는 멀티펄스 음성부호화 방식(MPC)에 있어서, 유성음의 파형에서 일그러짐이 발생한다. 이러한 문제를 해결하기 위해, 재생파형의 일그러짐이 감소하도록 피치구간 마다 멀티펄스의 진폭과 위치를 보정하는 APC-MPC를 제안하였다. 또한 융복합 시스템의 8kbps 부호화 조건에서 APC-MPC의 SNRseg를 검토하고 부호화 시스템으로 구현하였다. APC-MPC의 SNRseg를 평가한 결과, APC-MPC의 남자음성에서 14.3dB, 여자음성에서 13.9dB 임을 확인할 수 있었다. 본 방법은 셀룰러폰이나 스마트폰과 같이 Low Bit Rate의 음원을 사용하여 음성신호를 부호화하는 방식에 활용할 수 있을 것으로 기대된다.

**주제어** : 융복합, 셀룰러폰, 스마트폰, 인터넷폰, 코딩시스템

**Abstract** In a MPC(Multi-Pulse Coding) using excitation source of voiced and unvoiced, it would be a distortion of voice waveform. This is caused by normalization of synthesis speech waveform of voiced in the process of restoration. To solve this problem, this paper present APC-MPC of amplitude-position compensation in a multi-pulses each pitch interval in order to reduce distortion of synthesis waveform. Also, I was implemented that the APC-MPC in coding system. And I evaluate the SNRseg of APC-MPC in 8kbps coding condition of convergence system. As a result, SNRseg of APC-MPC was 13.9dB for female voice and 14.3dB for male voice respectively. And so, I expect to be able to this method for cellular phone and smart phone using excitation source of low bit rate.

**Key Words** : Convergence, Cellular Phone, Smart Phone, Internet Phone, Coding System

### 1. 서론

스마트폰 사용 인구가 급증하는 추세 속에 스마트폰과 관련된 하드웨어와 소프트웨어 기술이 진보하고 있으

며, 하드웨어와 소프트웨어 기술의 핵심이라 할 수 있는 알고리즘과 디지털 신호처리 기술의 혁신은 스마트폰을 보다 다양한 형태로 진보하게 할 것이다. 다만, 스마트폰 사용자의 급증은 통신망 트래킹 과부하를 초래할 수 있

Received 21 April 2015, Revised 17 June 2015

Accepted 20 July 2015

Corresponding Author: See-Woo Lee  
(The Society of Digital Policy)

Email: swlee@smu.ac.kr

ISSN: 1738-1916

© The Society of Digital Policy & Management. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

기 때문에 이에 대한 대책이 필요하다. 특히 신호를 압축 복원하는 기술은 통신 트래킹 과부하를 저감시킬 수 있는 방법이 될 수 있다. 음성신호를 낮은 비트로 부호화하는 방식[1,2,3,4]은 일반적으로 음성신호를 주기적인 유성음과 비주기적인 무성음으로 구분하여 재생한다. 주기적인 유성음 재생에는 펄스성 음원을, 비주기적인 무성음 재생에는 백색잡음의 음원을 사용하는데, 특히 펄스성 음원은 피치정보를 이용하여 대표구간의 음원을 피치구간마다 재구성함으로써 Low Bit Rate을 실현한다[5,6]. 그러나 이와 같은 경우에 유성음을 재생하기 위하여 사용하는 구동음원은 대표구간이 갖는 펄스의 진폭과 위상으로 정규화될 수 밖에 없다. 따라서 신호대 잡음비가 낮아질 수 밖에 없다. 본 연구에서는 정규화된 펄스의 진폭과 위상을 보정하여 신호대 잡음비(SNR)를 개선하는 방법으로 APC-MPC(Amplitude and Phase Compensation-Multi Pulse Coding)을 제안하고자 한다.

## 2. 음원 추출 및 보정

음성신호를 재생하기 위해서는 음성합성 필터와 합성 필터를 구동하는 음원이 필요하다. 이 음원으로 진폭과 위치를 달리하는 다수의 펄스를 탐색하고 추출하는 다양한 방법을 제안하였다[7]. 그러나 이러한 방법들은 펄스의 탐색과 추출에 자기상관법을 사용하고 있는데, 자기상관법은 음성신호의 시작부분과 끝부분, 프레임경계부분, 천이구간 등 음성신호의 에너지 변화가 있는 구간에서는 펄스의 진폭과 위치가 정규화되는 특징을 갖는다. 이러한 특징은 음질을 저하시키는 요인으로 작용한다. 이러한 요인을 극복하기 위하여 펄스의 진폭과 위치를 보상해주는 방법이 필요하나 그동안 멀티펄스 음원의 보상방법에 대해 다양한 연구가 진행되지 않았다. 이번 연구에서는 멀티펄스의 진폭과 위치를 보상함으로써 음질이 개선하는 음성부호화 방식인 APC-MPC방식(APC-MPC: Amplitude Position Compensation-Multi Pulse Coding)을 [Fig. 1]에 나타내었다.

음원은 원 음성신호 $x(n)$ 와 멀티펄스 $v(n)$ 에 의하여 재생된 음성신호  $\bar{x}(n)$ 로부터 식(1)이 최소가 되도록 펄스의 진폭( $g_i$ )과 위치( $m_i$ )를 결정하여 음원을 탐색하고 추출한다[8,9,10].

$$E = \sum_{n=0}^{N-1} [x(n) - \bar{x}(n) * w(n)]^2 \quad (1)$$

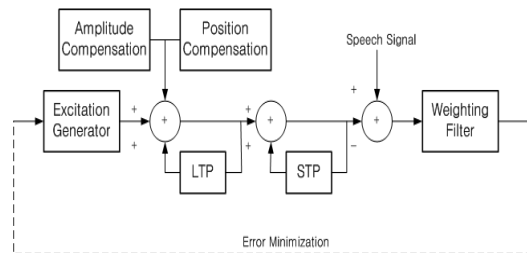
$$\bar{x}(n) = \sum_{i=1}^K g_i \cdot h(n - m_i)$$

여기에서  $K$ ,  $g_i$ ,  $m_i$ 는 멀티펄스의 수,  $i$ 번째의 펄스진폭과 위치를 나타낸다. [Fig. 1]에 있어서 피치재생필터(LTP)를  $h_p(n)$ , 음성재생필터(STP)를  $h(n)$ 이라 할 때, 음성재생 필터의 임펄스 응답 $h_s(n)$ 은 다음과 같이 나타낼 수 있다.

$$h(n) = h_p(n) * h_s(n) \quad (2)$$

또한, 식(1)의  $w(n)$ 은 Weighting 필터로서 다음 식과 같은 특성을 갖는다.

$$W(z) = (1 - \sum_{i=1}^p a_i z^{-i}) / (1 - \sum_{i=1}^p a_i r^i z^{-i}), \quad (0 < r < 1) \quad (3)$$



[Fig. 1] APC-MPC

그리고 APC-MPC의 전달함수 $H(z)$ 는 다음 식으로 나타낼 수 있다.

$$H(z) = 1 / (1 - \sum_{i=1}^Q b_i z^{-(T+i-1)}) \cdot (1 - \sum_{i=1}^P k_i z^{-i}) \quad (4)$$

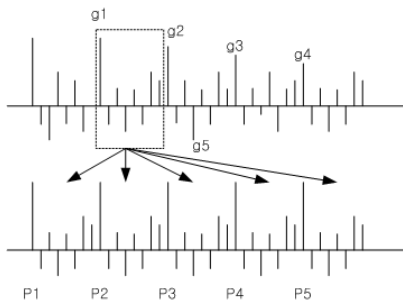
여기에서  $b_i$ ,  $T$ ,  $a_i$ 는 자기상관함수와 상호상관함수로부터 구하여진 피치계수, 피치주기, 예측계수를 나타내며, 식(1)을 최소화하는 멀티펄스의 진폭과 위치는 다음 식을 만족한다.

$$\sum_{j=1}^K g_j \cdot R[m_i - m_j] = \Phi(m_i), \quad (1 \leq i \leq K) \quad (5)$$

윗 식에서  $R$ 은 자기상관함수,  $\Phi$ 는 상호상관함수를 나

타낸다. 위 식에서 표현하는 멀티펄스는 음성에너지의 증감에 따라서 멀티펄스의 진폭과 위치도 함께 증감하는 펄스로 구성된다. 그러나 정보 전송률을 낮추기 위하여 대표 피치구간의 멀티펄스를 사용하여 구동음원을 재구성함으로써 멀티펄스의 진폭과 위상은 정규화 된다.

이러한 경우에 나타나는 재생음성의 음질열화 현상을 억제하기 위하여 셀룰러폰과 스마트폰에 적용된 CELP 방식에서는 코드북의 구동음원을 재구성할 때 발생하는 에러신호를 이용하여 음원이득을 제어하여 음질을 개선하고 있다[11].



[Fig. 2] Exciting Source of Multi Pulses

APC-MPC에서는 음질을 개선하기 위하여 멀티펄스 음원의 진폭과 위치를 보정하여 음질의 객관적 평가 척도인  $SNR_{seg}$  을 개선하는 방법을 살펴보고자 한다. 우선, 유성음을 다음식과 같은 n개의 파형으로 나타낸다.

$$s(t) = \sum_{k=1}^M (A_V(k) e^{(-\pi B_V(k)t}) \cdot \cos(2\pi F_V(k)t + \theta_V(k)) \quad (6)$$

여기에서  $F_V(k)$ 는 공진주파수,  $B_V(k)$ 는 대역폭,  $A_V(k)$ 와  $\theta_V(k)$ 는 파형의 진폭과 위상을 나타낸다. 그러나 [Fig. 2]에서와 같이 알 수 있듯이 피치위치에 있는  $g_2, g_3, g_4$ 의 진폭은 멀티펄스의 재구성시에  $g_1$ 의 진폭으로 정규화 된다. 그러나 실제 유성음의 진폭과 위상은 시간에 따라 변화되고, 특히 유성음의 음성파형은 진폭레벨이 증가하거나 감소하는 패턴을 갖는다. 이러한 진폭레벨 증감 패턴을 고려하여 멀티펄스의 진폭과 위치를 식(7)과 식(8)로 나타낼 수 있다.

$$A_V(k) = \gamma \cdot \sum_{k=1}^M \alpha_k \quad (7)$$

$$\theta_V(k) = \sum_{k=1}^M g_k \cdot \delta(n - m_k) \quad (8)$$

여기에서  $\gamma$ 는 각 프레임의 진폭레벨의 증감패턴,  $\alpha_k$ 는 피치위치에 있는 멀티펄스의 진폭 값을 나타낸다. 아울러  $\alpha_k$ 는 피치위치의 멀티펄스 진폭 값  $g_k$ 와  $g_{k+1}$ 의 차이를 식(9)와 같이 계산하여 수신측에 전송한다.

$$\alpha_k = |g_k - g_{k+1}| \quad (k=1,2,3,\dots,M) \quad (9)$$

수신측에 전송된 대표구간의 멀티펄스는 피치구간마다 멀티펄스의 진폭과 위치를 식(7)과 식(8)로 보정하여 멀티펄스를 재구성한다. 여기에서  $\alpha_k$ 를 전송하는데 사용하는 비트를 줄이기 위해  $g_k$ 와 피치위치에 있는 멀티펄스의 최대 진폭 값( $g_{max}$ )으로부터 다음 식과 같은 진폭 보정계수( $\beta_k$ )를 구하여 사용한다. 즉, 전송된  $\beta_k$ 와  $g_{max}$ 의 정보는 수신측에서  $\alpha_k$ 를 구하여 적용할 수 있고,  $\alpha_k$ 에 비해  $\beta_k$ 가 작은 값이기 때문에 보다 작은 비트를 할당할 수 있다.

$$\beta_k = |\alpha_k / g_{max}| \quad (10)$$

수신측에 전송한  $g_{max}$ 와  $\beta_k$ 로 피치위치에 있는 멀티펄스의 진폭을 보정하여 생성된 멀티펄스 음원을 다음 식으로 나타내었다.

$$v(n) = \sum_{k=1}^M (g_k + \gamma / \beta_i \cdot g_{max}) \cdot \delta(n - m_k) \quad (11)$$

위 식으로 표현된 멀티펄스 음원을 피치간격마다 재구성하여 음성재생필터를 구동하여 음성신호를 재생한다.

### 3. 부호화조건

APC-MPC에 있어서 재생신호를 얻기 위하여 사용되는 파라미터에 할당하는 비트를 <Table 1>에 제시하였다. 일반적인 이동통신시스템에서는 3.4kHz LPF로 주파수 대역을 제안하고 있으므로 본 연구에서도 음성신호를 3.4kHz LPF로 주파수 대역을 제안하고, 10kHz, 12bit로

표본화 및 양자화 하였으며, 프레임 길이는 25.6ms로 하였다. 한편, APC-MPC와 MPC를 객관적인 척도인  $SNR_{seg}$  로 평가하는 것도 주요 연구대상이나, 부호화 방식을 실제 이동통신 시스템에 적용하기 위해서는 프레임 당 전송율이 약 8kbps이하가 되어야 한다. 따라서 이번 연구에서는 APC-MPC가 이동통신 시스템에 적용하기에 적합한 전송율인 8kbps이하가 되도록 하였다.

<Table 1> Coding Condition

Parameter[bit]	MPC	APC-MPC
V/UV(1)	2	2
Coefficient of Synthesis Filter	7,6,5,5,4 3,3,3,3,3	7,6,5,5,4 3,3,3,3,3
$g_{max}$	10	10
$g_k$	8	6
$m_k$	8	6
$\beta_k$		3
$\gamma$		2
Number of Multi-Pulse	10	10
Average Pitch	8	8
kbps	8	8

APC-MPC와 MPC 모두 유성음원(V)과 무성음원(UV)을 사용하므로[12,13,14,15], V/UV의 선택정보에 각각 2bit를 할당하였으며, V의 경우에는 평균 피치정보에 8bit, 멀티펄스 진폭 및 위치에 할당된 비트는 MPC의 경우에 8bit, APC-MPC의 경우에는 6bit를 할당하였다. 아울러 멀티펄스의 최대진폭에 10bit, 진폭레벨 패턴에 2bit, 진폭보정계수에 3bit를 할당하였다. 그리고 음성재생필터의 차수는 10차를 사용하였으며, 음성재생필터 계수의 변화가 스펙트럼의 변화에 미치는 영향은 낮은 차수의 계수일수록 영향이 크기 때문에 낮은 차수일수록 많은 bit를 할당하였다[16].

#### 4. 실험결과

APC-MPC와 MPC의 비교하기 위하여 음질의 객관적 척도인  $SNR_{seg}$  사용하고자 하였다.  $SNR_{seg}$  측정하기 위해서는 원 음성신호와 음성재생 필터의 출력 음성신호가 필요하다. 제 2장에서와 같이 추출한 멀티펄스 음원에서 피치구간에 근거한 대표구간의 멀티펄스를 송신측에 전송한다. 수신측에서는 대표구간의 멀티펄스를 피치간격

마다 되풀이 하여 멀티펄스 음원을 얻는다. 이 음원으로 음성재생 필터를 구동함으로써 음성신호를 재생하고, 재생된 음성신호와 원 음성신호와의 차로부터  $SNR_{seg}$  를 얻는다. 이 APC-MPC와 MPC의  $SNR_{seg}$ 로부터 음질의 개선정도를 측정할 수 있다. 이번 실험에서는 남녀 4명이 발성한 <Table 2>의 음성샘플을 사용하여 APC-MPC와 MPC의  $SNR_{seg}$  를 측정하였다.

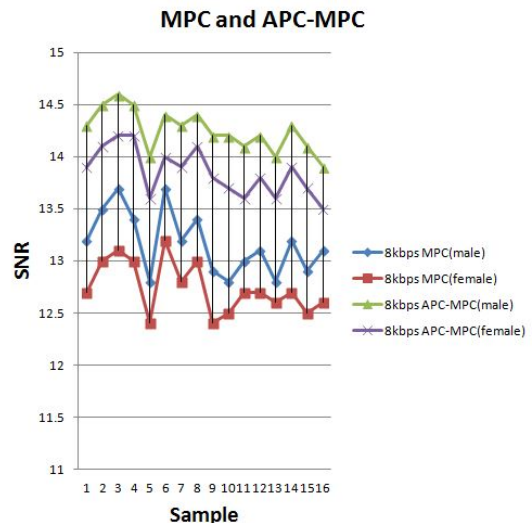
<Table 2> Speech Sample

	male	female
participants	4	4
Time	54.4 sec	54.4 sec
sentences	16	16
Vowel	145	145
Unvoiced Coseonant	34	34

$SNR_{seg}$  를 측정한 결과, <Table 3>에 나타낸바와 같이 8kbps에 있어서 APC-MPC는 남자음성에서 14.3dB, 여자음성에서 13.9dB를 얻었다. 기존의 MPC에 비하여 남자음성에서 1.1dB, 여자음성에서 1.2dB 정도 개선되었다.

<Table 3>  $SNR_{seg}$  of MPC and APC-MPC

Method [dB]	kbps	male	female
MPC	8	13.2	12.7
APC-MPC	8	14.3	13.9



[Fig. 3] SNR of MPC and APC-MPC

<Table 2>의 표본에 따른 APC-MPC와 MPC의  $SNR_{seg}$ 를 [Fig. 3]에 나타내었다. 이러한 결과로 부터 대표구간의 멀티펄스에 의하여 재생된 음성신호에 비하여 멀티펄스의 진폭과 위치를 보정함으로써 원 음성신호에 보다 가까운 재생신호를 얻을 수 있음을 알 수 있다.

## 5. 결론

본 연구에서는 멀티펄스 음원의 진폭과 위치를 보정하는 APC-MPC를 제안하였다. 실험결과, 기존의 MPC에 비하여 남자음성에서 1.1dB, 여자음성에서 1.2dB 정도 개선된 것을 알 수 있었다. 다만, APC-MPC에 있어서 진폭과 위치 보정이 피치정보를 갖는 음성신호의 프레임에는 효과적이거나, 피치정보를 갖지 않는 음성신호의 프레임에는 비효과적이라 할 수 있다. 따라서 향후 본 연구에서는 APC-MPC에 있어서, 피치정보를 갖지 않는 음성신호의 프레임을 보완하는 방법과 8kbps 이하에서도 같은 음질을 유지하거나 향상된 음질을 갖는 부호화 방식을 연구하고자 한다.

## REFERENCES

- [1] Selma Ozaydm, Buyurman Baykal: "Matrix quantization and mixed excitation based linear predictive speech coding at very low bit rates", *Speech Communication* 41, p381-392, 2003
- [2] Ghaemmaghami, S., Sridharan, S.: "Very low rate speech coding using temporal decomposition". *IEEE Electron. Lett.* 35(6), p456-457.1999
- [3] McCree, A.V, Barnwell, T.P., "A mixed excitation LPC vocoder model for low bit rate speech coding", *IEEE Trans. Speech Audio Process*, p242-250,1995
- [4] Phu Chien Nguyen, Masato Akagi, Binh Phu Nguyen: "Limited error based event localizing temporal decomposition and its application to variable-rate speech coding", *Speech Communication* 49, p292-304, 2007
- [5] LeBlanc, W.P., Bhattacharya, B., Mahmoud, S.A.: "Efficient search and design procedures for robust multi stage vector quantization of LPC parameters for 4kbps speech coding". *IEEE Trans. Speech Audio Process*. p373-385.1993
- [6] David A. Krubsack and Russell J. Niederjohn: "An Autocorrelation Pitch Detector and Voicing Decision with Confidence Measures Developed for Noise-Corrupted Speech", *IEEE, Transactions of Signal Processing*, Vol.39, No.2, 1991
- [7] Kazunori Ozawa, Shigeru Ono and Takashi Araseki: "A study on pulse search algorithm for multipulse excited speech coder realization", *IEEE, Journal on Selected areas in Communications*, Vol. SAC-4, No.1, 1986
- [8] B.S. Atal and J.R. Remdo: "A New Model of LPC Excitation for Producing Natural Sounding Speech at Low Bit Rates", *IEEE, ICASSP*, p614-617, 1982
- [9] Z.A. Putnins, G.A. Wilson, J. Kumar and R.D. Trupp: "A Multi-Pulse LPC Synthesizer for Telecommunications use", *IEEE, ICASSP*, Mar, 1985
- [10] Kazunori OZAWA, Takashi ARASEKI: "Multi-Pulse Excited Speech Coding Utilizing Pitch Information at Rates Between 9.6 and 4.8 kbit/s", *IEICE, Vol. J72-D-2*, No.8, 1989
- [11] K. Krishna, V.L.N. Murty, R. Ramakrishnan: "Vector quantization of excitation gains in speech coding", *Signal Processing* 81, p203-209, 2001
- [12] Widrow B. and Hoff M. E.: "Adaptive Switching Circuit", *IRE WESCON Conv. Rec.*, June 2000
- [13] Campbell, J.P., Tremain, T.E.: "Voiced/unvoiced classification of speech with applications to the U.S. Government LPC-10e algorithm", *Proc. IEEE Int. Conf. on Acoustics, Speech, Signal Processing*, p473-476.1986
- [14] LEAH, J. SIEGE and ALAN C. BESSEY: "Voiced/Unvoiced/Mixed Excitation Classification of Speech", *IEEE, Vol. ASSP-30*, No.3, 1982
- [15] HIDEFUMI KOBATAKE: "Optimization of Voiced/Unvoiced Decisions in Nonstationary Noise Environments", *IEEE, Vol. ASSP-35*, No.1, 1987
- [16] Nobuhiko KITAWAKI, Fumitada ITAKURA and Shuzo SAITO: "Optimum Coding of Transmission Parameters in PARCOR Speech Analysis Synthesis System", *IEICE, Vol. J61-A No.2*, 1978

이 시 우(Lee. See Woo)



- 1990년 3월 : 日本大學(Nihon Univ)  
전자공학과 (공학석사)
- 1994년 3월 : 日本大學(Nihon Univ)  
전자공학과 (공학박사)
- 1994년 3월 ~ 1998년 2월 : (주)삼  
성전자 통신연구소/멀티미디어 연  
구소
- 1998년 3월 ~ 현재 : 상명대학교 정보통신공학과 교수
- 관심분야 : 음성신호처리, 음주관독시스템, 유무선통신
- E-Mail : swlee@smu.ac.kr