

음성 및 오디오 부호화기를 위한 저지연 윈도우 스위칭 modified discrete cosine transform

Low delay window switching modified discrete cosine transform for speech and audio coder

김영준,¹ 이인성[†]

(Young-Joon Kim¹ and In-Sung Lee^{1†})

¹충북대학교 정보통신공학부

(Received December 27, 2017; revised February 28, 2018; accepted March 29, 2018)

초 록: 본 논문에서는 음성/오디오 부호화기를 위한 저지연 윈도우 스위칭 MDCT(Modified Discrete Cosine Transform) 방법을 제안한다. 윈도우 스위칭 알고리즘을 사용하여 신호의 특성이 빨리 변하는 전이 구간에서 음질 저하를 개선하고, 저지연 TDAC(Time Domain Aliasing Cancellation)를 사용하여 알고리즘 지연을 1/2로 줄일 수 있는 MDCT 방법을 제안한다. 제안된 윈도우 스위칭 방법은 기존 윈도우 스위칭 알고리즘이 다른 길이의 중첩합(overlap-add)을 사용하는 것과 달리, 일정한 길이의 중첩합을 사용하여 알고리즘 지연을 1/2로 줄일 수 있었고, 신호의 특성에 따라 윈도우의 종류를 2개로 줄여 프레임 상태를 표현하는 정보 비트를 1 bit 감소시킬 수 있었다. 제안한 알고리즘을 MDCT 기반의 음성/오디오 부호화기인 ITU-T(International Telecommunication Union - Telecommunication) G.729.1 부호화기에 적용하여 성능을 확인하였으며, 알고리즘 지연을 절반으로 감소시키면서 동일한 음질을 유지할 수 있었다.

핵심어: 윈도우 스위칭, 저지연 알고리즘, 중첩합, MDCT (Modified Discrete Cosine Transform), TDAC (Time Domain Aliasing Cancellation)

ABSTRACT: In this paper, we propose a low delay window switching MDCT (Modified Discrete Cosine Transform) method for speech/audio coder. The window switching algorithm is used to reduce the degradation of sound quality in non-stationary transient duration and to reduce the algorithm delay by using the low delay TDAC (Time Domain Aliasing Cancellation). While the conventional window switching algorithms uses overlap-add with different lengths, the proposed method uses the fixed overlap add length. It results the reduction of algorithm delay by half and 1 bit reduction in frame indication information by using 2 window types. We apply the proposed algorithm to G.729.1 based on MDCT in order to evaluate the performance. The propose method shows the reduction of algorithm delay by half while speech quality of the proposed method maintains same as the conventional method.

Keywords: Window switching, Low-delay algorithm, Overlap-add, MDCT (Modified Discrete Cosine Transform), TDAC (Time Domain Aliasing Cancellation)

PACS numbers: 43.60.Ek, 43.60.Dh

1. 서 론

스마트폰이 대중화되면서 데이터 서비스 이용이

폭발적으로 증가했고, 이러한 데이터 사용량의 증가는 기존 3G 이동 통신 시스템의 부하를 가중시켰다. 이에 대응하기 위하여 국내외에서 고속의 데이터 서비스가 가능한 LTE(Long Term Evolution) 시스템이 상용화 되었다. LTE의 등장으로 VoIP(Voice over Internet Protocol)나 VoLTE(Voice over LTE)와 같은 패

[†]Corresponding author: In-Sung Lee (inslee@chungbuk.ac.kr)
Department of Radio Communication Engineering, Chungbuk National University, 1 Chungdae-ro, Seowon-Gu, Cheongju, Chungbuk 28644, Republic of Korea
(Tel: 82-43-261-3223)

킷망을 이용한 음성 서비스가 증가하고 있으며, 고음질의 음성코덱인 EVS(Enhanced Voice Service)도 패킷망에서의 통신을 지원하게 되었다.^[1] 이러한 패킷망을 통한 음성 서비스는 기존의 알고리즘 및 프로세싱 지연뿐만 아니라, 네트워크 상태에 따라 통신 지연이 추가적으로 발생되기 때문에 회선망을 이용한 음성 서비스보다 더 많은 지연이 발생하게 된다. 따라서 사용자가 불편함을 느끼지 못하는 200 ms 이하의 음성 서비스가 요구되고 있으며 패킷망을 이용한 음성 서비스가 증가함에 따라 알고리즘 지연은 코덱의 성능을 결정짓는 중요한 부분이 되었다.

또한 영상 신호 처리에 사용되는 DCT(Discrete Cosine Transform)를 음성이나 오디오 신호 처리에 사용할 때에는 몇몇 문제가 발생하게 된다.^[2] DCT나 FFT(Fast Fourier Transform)와 같은 주파수 변환방법을 사용하여 50% 중첩합을 하면 코덱에서 처리해야 할 프레임 수가 증가되는 문제점이 발생한다. 또한 심리음향 모델을 적용한 AAC(Advanced Audio Coding)나 MPEG-(Motion Picture image coding Experts Group)-Layer 3와 같은 손실압축 오디오 코덱에서 DCT를 사용할 경우 신호 손실된 부분에서 인공음이 만들어져 음질의 저하가 발생하게 된다.^[3,4] 이와 같은 DCT의 문제점을 보완하기 위해 MDCT(Modified Discrete Cosine Transform) 방법이 제안되었다.^[5] 하지만 MDCT는 DCT나 FFT와는 다르게 역MDCT를 통해 신호가 원래의 신호로 완벽하게 복원되지 않는다. 따라서 신호를 완벽 복원하기 위해서 TDAC(Time Domain Aliasing Cancellation) 알고리즘을 사용하며, 이 과정에서 프레임 길이 N 만큼의 지연이 발생하게 되는 단점을 가지고 있다.^[6]

따라서 본 논문에서는 고정된 윈도우 크기로 신호를 처리할 때 발생하는 음질 저하를 해결하기 위하여 신호의 종류에 따라 윈도우의 크기가 변하는 윈도우 스위칭 알고리즘과 지연 문제를 해결하기 위한 저지연 TDAC 알고리즘을 결합한 방법을 제안한다. 윈도우 스위칭 알고리즘을 사용하여 기존에 발생하는 음질 저하 구간을 억제하고, 저지연 TDAC 알고리즘을 사용함으로써 발생하는 지연을 줄이고자 한다.

II. 윈도우 스위칭 MDCT

2.1 윈도우 크기 선택 알고리즘

기준이 되는 샘플 길이 안에서 신호의 변화가 작은 정지적 신호라면 긴 윈도우 크기를 선택, 변화가 큰 비정지적 신호라면 변화하는 구간의 특성을 반영하기 위해 더 짧은 크기의 윈도우를 선택하는 것이 효율적이다.^[7,8] 신호의 특성에 맞는 적절한 크기의 윈도우를 선택하기 이전 프레임 구간의 에너지를 구하고 다음 윈도우를 적용할 인접 프레임의 에너지를 구하여 두 구간 사이의 에너지 차이를 계산하게 된다. 수식적으로는 다음과 같이 정의된다.

$$x_{DF}(N, m) = \frac{1}{N} \sum_{j=1}^N (x_N(j, m) - x_N(j, m-1))^2, \quad (1)$$

여기서 N 은 기준이 되는 가장 큰 프레임 크기이다. 그리고 j 는 기준이 되는 N 의 크기 안에서 j 번째 프레임, m 은 선택하는 후보 프레임의 길이이다. 가장 작은 에너지 차이를 나타낸 값의 크기를 다음 프레임에 적용할 윈도우 크기로 선택하게 된다. 수식적으로는 다음과 같이 정의된다.

$$x_{\min}(N) = \min_{m=1,2,\dots,M_{\max}} x_{DF}(N, m). \quad (2)$$

2.2 윈도우 스위칭 알고리즘

Fig. 1은 윈도우 스위칭의 알고리즘을 간략하게 나타낸 것이다. normal은 기준이 되는 윈도우의 길이를

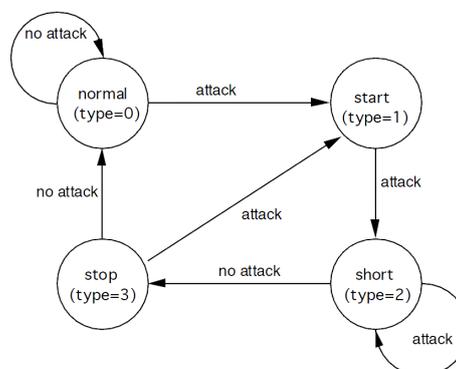


Fig. 1. Window switching algorithm.

사용하는 것이고, **short**는 기준 길이의 윈도우보다 작은 길이의 윈도우를 사용하는 것이다. **start**, **stop**은 각각 길이가 달라지는 중첩합 구간을 이어주기 위해 **normal**과 **short**의 특성을 반반씩 갖는 윈도우이다. **attack**으로 표기된 부분은 윈도우의 길이의 변경을 요구하는 것이고, **no attack**은 윈도우 길이의 유지를 요구하는 것이다.

2.3 기존 사용된 윈도우 종류

기존 윈도우 스위칭에서 사용된 윈도우는 Fig. 2와 같이 **normal**, **start**, **stop**, **short**의 4 종류이다. (a)는 기준이 되는 가장 긴 길이의 윈도우이고 (c)는 3개의 짧은 중첩 윈도우로 이루어져 있다. (b)와 (d)는 긴 구간과 짧은 구간을 부드럽게 이어줄 수 있는 역할을 하는 윈도우이고 서로 반전 형태를 갖는다.

4가지 윈도우는 수식적으로 다음과 같이 표현된다. 먼저 **normal** 상태에서 사용하는 윈도우이다.

$$w(k) = \sin \frac{\pi}{2N} \left(k + \frac{1}{2} \right), \quad k = 0, 1, \dots, 2N-1, \quad (3)$$

여기서 N 은 기준이 되는 한 프레임의 길이이다. 다음은 **short** 상태에서 사용하는 윈도우이다.

$$w(k) = \sin \left(\frac{3\pi}{2N} \left(k + \frac{1}{2} \right) \right), \quad k = 0, 1, \dots, \frac{2N}{3}-1. \quad (4)$$

다음은 **start** 상태에서 사용하는 윈도우이다.

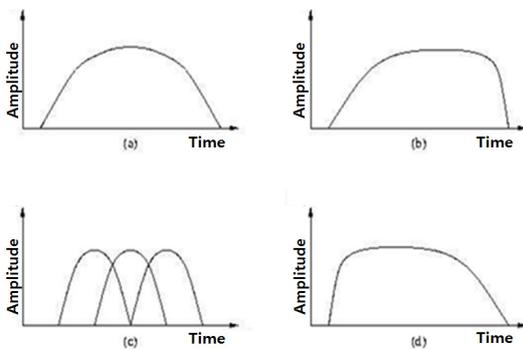


Fig. 2. Existing used windows - (a) normal window, (b) start window, (c) short window, (d) stop window.

$$w(k) = \begin{cases} \sin \frac{\pi}{2N} \left(k + \frac{1}{2} \right), & k = 0, 1, \dots, N-1 \\ 1, & k = N, \dots, \frac{4N}{3}-1 \\ \sin \left(\frac{3\pi}{2N} \left(k - N + \frac{1}{2} \right) \right), & k = \frac{4N}{3}, \dots, \frac{5N}{3}-1 \\ 0, & k = \frac{5N}{3}, \dots, 2N-1 \end{cases} \quad (5)$$

$$w(k) = \begin{cases} 0, & k = 0, \dots, N-1 \\ \sin \left(\frac{3\pi}{2N} \left(k - \frac{N}{3} + \frac{1}{2} \right) \right), & k = N, \dots, \frac{4N}{3}-1 \\ 1, & k = \frac{4N}{3}, \dots, \frac{5N}{3}-1 \\ \sin \frac{\pi}{2N} \left(k + \frac{1}{2} \right), & k = \frac{5N}{3}, \dots, 2N-1 \end{cases} \quad (6)$$

마지막으로 **stop** 상태에서 사용하는 윈도우이다.

$$w(k) = \begin{cases} 0, & k = 0, \dots, N-1 \\ \sin \left(\frac{3\pi}{2N} \left(k - \frac{N}{3} + \frac{1}{2} \right) \right), & k = N, \dots, \frac{4N}{3}-1 \end{cases} \quad (7)$$

$$w(k) = \begin{cases} 1, & k = \frac{4N}{3}, \dots, \frac{5N}{3}-1 \\ \sin \frac{\pi}{2N} \left(k + \frac{1}{2} \right), & k = \frac{5N}{3}, \dots, 2N-1 \end{cases} \quad (8)$$

2.4 MDCT / IMDCT(Inverse MDCT)

기존의 MDCT는 DCT-IV를 기반으로 만들어진 주파수 변환 방법으로 G.718, G.729.1, MPEG, AAC 코덱과 같은 다양한 음성 및 오디오 부호화기에서 사용한다. Fig. 3은 DCT를 사용하여 변환, 역변환하고 50% 중첩합을 사용하여 신호를 복원할 때 처리해야 할 프레임의 수가 2배로 증가하는 문제를 보여준다.^[9] 이러한 문제점을 해결하기 위한 MDCT는 $2N$ 개의 입력신호를 N 개의 주파수 계수로 표현하기 때문에 처리해야 할 데이터가 증가하지 않으며, 이는 Fig. 4에 표현되어 있다.

즉, MDCT는 $2N$ 개의 데이터를 이용하여 N 개의 주파수 계수로 표현하고, N 개의 MDCT 계수를 이용하여 역MDCT를 하면 다시 $2N$ 개의 데이터로 복원하게 된다. 이때 역MDCT 결과는 DCT나 DFT와는 다르게 완벽하게 복원되지 않고 Fig. 5와 같이 신호가 A, B 구간과 C, D 구간 사이에서 각각 대칭 특성을 가지게 된다. Fig. 5에서 아래첨자 R 은 역순을 의미한다. 이러한 문제를 해결하기 위하여 중첩하는 프레임 사이의

중첩합을 통하여 완벽 복원되는 성질을 가지는데, 이렇게 복원하는 과정을 TDAC라고 한다.

Fig. 6은 TDAC의 전체 과정을 자세하게 나타내고

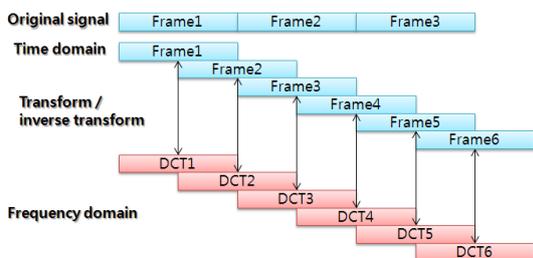


Fig. 3. The overlap-add problem of DCT.

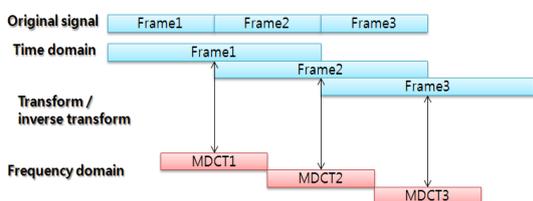


Fig. 4. Overlap-add using MDCT.

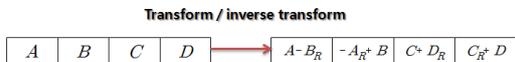


Fig. 5. Symmetrical characteristics of MDCT/IMDCT.

A	B	C	D	E	F
$A\omega_1$	$B\omega_2$	$C\omega_3$	$D\omega_4$	Windowed input	
$A\omega_1$	$-A\omega_{1R}$	$C\omega_3$	$C_R\omega_{3R}$	Output after MDCT/IMDCT	
$-B_R\omega_{2R}$	$+B\omega_2$	$+D_R\omega_{4R}$	$+D\omega_4$		
$A\omega_1\omega_1$	$-A\omega_{1R}\omega_2$	$C\omega_3\omega_3$	$C_R\omega_{3R}\omega_4$	Windowed Output	
$-B_R\omega_{2R}\omega_2$	$+B\omega_2\omega_2$	$+D_R\omega_{4R}\omega_3$	$+D\omega_4\omega_4$		
Windowed input	$C\omega_1$	$D\omega_2$	$E\omega_3$	$F\omega_4$	
Output after MDCT/IMDCT	$C\omega_1$	$-C\omega_{1R}$	$E\omega_3$	$E_R\omega_{3R}$	
	$-D_R\omega_{2R}$	$+D\omega_2$	$+F_R\omega_{4R}$	$+F\omega_4$	
Windowed Output	$C\omega_1\omega_1$	$-C\omega_{1R}\omega_2$	$E\omega_3\omega_3$	$E_R\omega_{3R}\omega_4$	
	$-D_R\omega_{2R}\omega_2$	$+D\omega_2\omega_2$	$+F_R\omega_{4R}\omega_3$	$+F\omega_4\omega_4$	
	C	D	Final overlap-add output		

Fig. 6. Overall operation of TDAC.

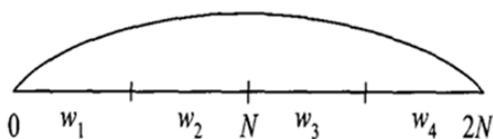


Fig. 7. The window used for MDCT divided into four section.

있다. 먼저 A, B, C, D, E, F 의 각 신호 구간의 길이는 $N/2$ 으로 정의하고, $\omega_1, \omega_2, \omega_3, \omega_4$ 는 Fig. 7과 같이 $2N$ 길이의 윈도우를 $N/2$ 길이의 네 구간으로 나누는 것이고, 윈도우는 다음과 같은 조건을 만족하여야 한다.

$$\omega_1\omega_1 + \omega_3\omega_3 = \omega_2\omega_2 + \omega_4\omega_4 = 1. \tag{9}$$

$$\omega_1 = \omega_{4R}, \quad \omega_2 = \omega_{3R}. \tag{10}$$

이와 같은 방법으로 aliasing들이 제거되고, N 의 길이 만큼인 C, D 신호를 최종적으로 복원하게 된다.

III. 제안하는 윈도우 스위칭 알고리즘

3.1 저지연 윈도우 스위칭에 사용된 윈도우

제안하는 알고리즘에 사용할 윈도우는 기존에 사용하는 4가지가 아닌 Fig. 8의 long window, Fig. 9의 short window와 같이 두 가지 윈도우로 표현되고, 중첩합 구간의 길이를 동일하게 함으로써 start, stop과 같이 두 다른 길이의 윈도우를 부드럽게 이어줄 수 있는 역할을 하는 윈도우를 사용하지 않는다. 윈도우 크기 선택 알고리즘은 기존과 동일하게 사용한다. Fig. 8의 long window는 수식적으로는 다음과 같이 표현한다.

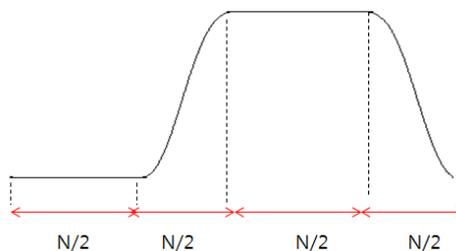


Fig. 8. Proposed long window.

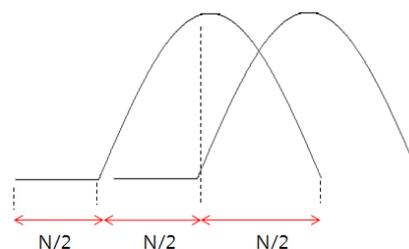


Fig. 9. Proposed short window.

$$w(k) = \begin{cases} 0, & k = 0, \dots, \frac{N}{2} - 1 \\ \sin \frac{\pi}{N} \left(k - \frac{N}{2} + \frac{1}{2} \right), & k = \frac{N}{2}, \dots, N-1 \\ 1, & k = N, \dots, \frac{3N}{2} - 1 \\ \sin \frac{\pi}{N} \left(k - N + \frac{1}{2} \right), & k = \frac{3N}{2}, \dots, 2N-1 \end{cases} \quad (11)$$

Fig. 9의 short window는 수식적으로 다음과 같이 표현하고 2개를 중첩하여 사용한다.

$$w(k) = \begin{cases} 0, & k = 0, \dots, \frac{N}{2} - 1 \\ \sin \frac{\pi}{N} \left(k - \frac{N}{2} + \frac{1}{2} \right), & k = \frac{N}{2}, \dots, N-1 \end{cases} \quad (12)$$

이러한 윈도우를 사용하게 되면 서로 다른 중첩합 길이를 조정해야 하는 기존의 윈도우 스위칭과 다르게 하나의 중첩합 길이로 고정하게 되면서 일관된 지연 감소효과를 얻게 된다. 또한 start window와 stop window를 사용하지 않아 4개의 상태 정보를 저장하여야 하는 윈도우 정보 bit가 2개의 상태만 필요하게 되면서 정보 1 bit의 감소효과를 얻게 된다.

3.2 상태별 과정

3제한하는 방법에서는 신호를 복원하는 상황을 3가지로 구분하였다. 첫째로 Fig. 10은 long window와 long window 사이의 중첩합 과정을 나타낸 것이다. 현재 프레임과 과거 프레임을 중첩합 하게 되면 0에서 $N/2$ 까지는 두 프레임의 중첩합으로 복원되고 $N/2$ 에서 N 까지는 상수 1의 값이 곱해져 있는 현재 프레임 값 그대로 복원되게 된다. 이 과정에서 N 에서 $3N/2$ 까지 $N/2$ 길이만큼의 look ahead, 즉 지연이 발생하게 된다. 이처럼 기존의 지연인 N 의 절반으로 줄어들게 된다.

Fig. 11은 long window와 short window 사이의 중첩합 과정을 나타낸 것이다. 여기에서 현재 프레임과 과거 프레임의 중첩합 과정을 살펴보면 0에서 $N/2$ 사이의 구간은 과거 프레임과 현재 프레임안의 첫 번째 프레임과 중첩 합을 하여 복원하게 된다. $N/2$ 에서 N 까지의 구간은 현재 프레임안의 두 short 프레임의 중첩 합으로 복원이 되고 최종적으로 N 의 길이만큼의 신호가 복원된다. 이 과정에서도 절반이 감

소한 $N/2$ 의 지연이 발생한다. short window에서 long window로 넘어가는 구간도 동일하다.

마지막으로 Fig. 12는 start window와 short window 사이의 중첩합 과정을 나타낸 것이다. 현재 프레임과 과거 프레임의 중첩합 과정을 살펴보면 과거 프레임안의 2번째 프레임과 현재 프레임안의 1번째 프레임과 중첩 합 과정을 거쳐 0에서 $N/2$ 만큼의 신호가 복원된다. 그리고 현재 프레임안의 1번째 프레임과 2번째 프레임의 중첩 합을 통해 $N/2$ 에서 N 까지의 신호가 복원되고 최종적으로 N 의 길이가 복원된다. 여기에서도 $N/2$ 만큼의 지연이 발생한다.

이처럼 모든 중첩합 상황에서 절반의 지연이 감소한

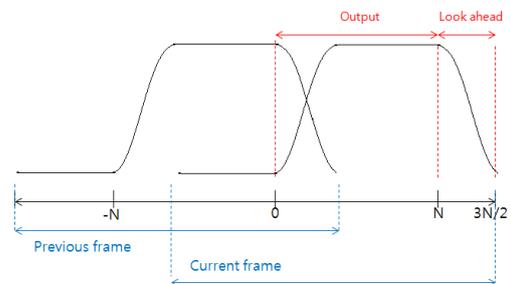


Fig. 10. Overlap-add between long window and long window.

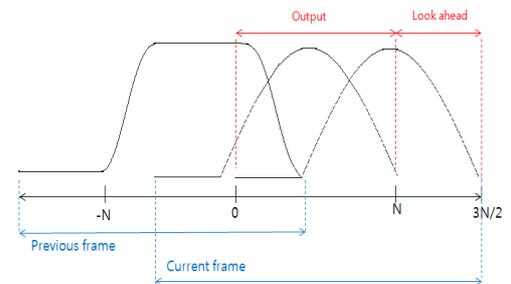


Fig. 11. Overlap-add between long window and short window.

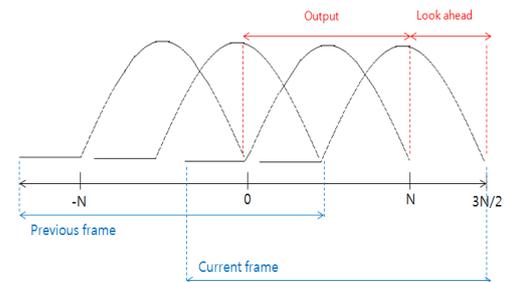


Fig. 12. Overlap-add between short window and short window.

일관된 $N/2$ 길이만큼의 지연이 발생하여 지연 감소 효과를 얻을 수 있고, 윈도우 스위칭을 사용하여 신호 구간의 특성별로 다른 윈도우를 사용하여 aliasing 발생을 억제하고 최적의 음질로 신호를 복원해낼 수 있다.

IV. 성능 평가

본 논문에서 제안하는 저지연 윈도우 스위칭 알고리즘의 성능을 평가하기 위하여 객관적인 방법을 이용한 성능 평가와 시각적인 성능 평가를 활용하였다. ITU-T(International Telecommunication Union-Telecommunication) 표준안으로 제정된 객관적인 음질 평가인 P.861 PESQ(Perceptual Evaluation of Speech Quality) 테스트를 이용하여 원 신호와 복호화된 신호의 음질을 비교하고, 시각적으로 스펙트럼을 분석하였다.

4.1 G.729.1 코덱

제안하는 알고리즘의 성능을 평가하기 위해 VoIP 환경에서 많이 사용하는 MDCT 기반의 음성 코덱인 G.729.1^[10]에 제안하는 방법을 적용하여 성능을 평가하였다. G.729.1은 CELP(Code Excited Linear Prediction), TDBWE(Time Domain BandWidth Extension), TDAC 세 개의 모듈로 구성되어 있고, 8 kHz와 16 kHz의 입력 신호를 사용하며, 입력된 신호는 QMF(Quadrature Mirror Filter)를 통과하며 저주파수 신호와 고주파수 신호로 나뉘게 된다. 저주파수 신호는 CELP모듈을 통해 부호화 되고, CELP의 잔여신호는 MDCT를 통해 MDCT 계수를 부호화기로 전송하게 된다. 고주파수 신호는 TDBWE 모듈과 MDCT를 이용하여 부호화를 한다. 이와 같은 부호화과정에서 프레임 길이 20ms, MDCT look-ahead 20ms, CELP의 LPC(Linear Predictive Coding) 분석의 look-ahead 5ms, QMF 3.9375ms의 알고리즘 지연이 발생하여 총 48.9375ms의 지연이 발생한다. 이때 MDCT look-ahead에서 발생하는 지연을 감소시키며 최적의 음질로 복원하기 위해 제안하는 알고리즘을 적용하여 실험을 하였다.

4.2 저지연 윈도우 스위칭 알고리즘 성능 평가

본 논문에서 제안한 알고리즘을 평가하기 위하여

테스트 신호로는 성인 남성과 여성의 한국어 문장과 영어 문장을 사용하였다. 16kHz로 샘플링 된 입력 신호와 16 kbps의 전송률 환경에서 실험을 하였다. Table 1은 지연 변화에 따른 원 신호와 복호화된 신호의 PESQ 음질 테스트 결과를 나타낸다. 처음 20ms의 지연은 기존의 MDCT를 사용한 결과이고, 나머지는 제안하는 알고리즘을 사용하여 지연을 10ms까지 감소시킨 실험 결과이다. 제안하는 알고리즘은 기존의 MDCT보다 짧은 영역의 중첩합을 통해서 감소된 지연으로 완벽 복원이 가능하다. 따라서 Table 1에서 보듯이 지연이 감소함에도 불구하고 객관적인 음질 테스트인 PESQ 값의 변화가 없는 것을 볼 수 있었고,

Table 1. PESQ experiment results.

Test signals	Delay (20 ms)	Delay (10 ms)
English female 1	3.7744	3.7744
English female 2	3.7987	3.7987
English male 1	3.9142	3.9144
English male 2	3.8642	3.8641
Korean female 1	3.6948	3.6949
Korean female 2	3.7478	3.7477
Korean male 1	3.9991	3.9991
Korean male 2	3.9927	3.9928

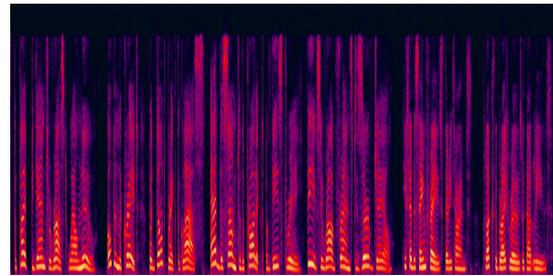


Fig. 13. Results of using existing MDCT with 20 ms delay.

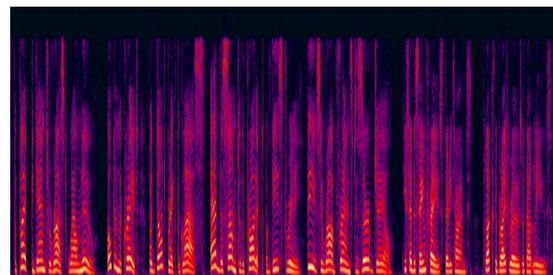


Fig. 14. Results of using low-delay window switching algorithm with 10 ms delay.

표준 편차 역시 0.10604와 0.10596으로 거의 차이가 없다는 것을 확인하였다.

Figs. 13과 14는 기존의 MDCT의 결과와 제안하는 알고리즘을 적용한 결과의 스펙트럼을 비교한 그림이다. Figs. 13과 14에서 보듯이 기존 MDCT를 사용한 복원신호와 제안한 알고리즘을 적용하여 지연을 감소시킨 복원신호의 스펙트럼간의 차이를 발생시키지 않으면서 완벽하게 복원된 결과를 확인할 수 있었다.

V. 결 론

본 논문에서는 윈도우 스위칭 알고리즘과 저지연 알고리즘을 결합한 알고리즘을 제안하였다. 윈도우 스위칭 알고리즘을 사용하여 기존에 발생하는 음질 저하 구간을 억제하고, 저 지연 알고리즘을 통해 발생하는 지연을 줄이고자 하였다.

제안하는 저지연 윈도우 스위칭 알고리즘의 성능을 평가하기 위해 객관적인 방법을 이용한 성능 평가와 시각적인 성능 평가를 활용 하였다. 객관적인 음질평가인 PESQ 테스트를 이용하여 원 신호와 복호화된 신호의 음질을 비교하여 제안하는 알고리즘의 성능을 평가하고 기존에 사용하는 MDCT의 결과와 제안한 알고리즘의 결과 비교를 통해 시각적으로 평가를 하였다. PESQ 평가 결과와 시각적인 성능평가에서 지연을 감소시켰음에도 불구하고 기존 방법과 음질의 차이가 없음을 확인할 수 있었다. 따라서 제안하는 알고리즘들을 다양한 MDCT 기반의 음성 및 오디오 코덱에 사용하여 알고리즘 지연을 효율적으로 감소시킬 수 있다.

감사의 글

본 연구는 교육과학기술부가 후원하는 한국연구재단 기초과학연구프로그램(NRF-2017R1D1A3B03027961) 지원으로 수행되었음(This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology).

References

1. M. S. Lee, D. Y. Kim, and B. S. Lee, "Trends of codec technology for 4 G mobile enhanced voice service" (in Korean), *Electronics and Telecommunications Trends*, **25**, 29-37 (2010).
2. L. Yaroslavsky and Y. Wang, "DFT, DCT, MDCT, DST and signal Fourier spectrum analysis," *Proc. 10th EUSIPCO*, 1065-1068 (2000).
3. ISO/IEC 13818-3: "Information technology - generic coding of moving pictures and associated audio information - Part 3: Audio," 1998.
4. ISO/IEC 14496-3: "2005/FPDAM9, Enhanced low delay AAC," 2007.
5. H. S. Malvar, "Lapped transforms for efficient transform/subband codin," *IEEE Trans. Acoustics, Speech, Signal Process*, **38**, 969-978 (1990).
6. J. Princen and A. Bradley, "Analysis/Synthesis filter bank design based on time domain aliasing cancellation," *IEEE Trans. Acoustics, Speech, Signal Process*, **34**, 1153-1161 (1986).
7. R. Geiger, J. Herre, M. Jander, M. Multrus, M. Schmidt, M. Schell, and G. Schuller, "Enhanced MPEG-4 low delay AAC - low bitrate high quality communication:," *Audio Engineering Society Convention 122* (2007).
8. G. Fuchs, C. R. Helmrich, G. Markovic, M. Neusinger, E. Ravelli, and T. Moriya, "Low delay LPC and MDCT-based audio coding in the EVS codec," *IEEE Trans. Acoustics, Speech, Signal Process*, 5723-5727 (2015).
9. Y. Wang and M. Vilermo, "The modified discrete cosine transform: Its implications for audio coding and error concealment," *J. Audio Eng. Soc.* **51**, 1-10 (2012).
10. ITU-T Rec. G.729.1, *An 8-32 Kbit/S Scalable Wideband Coder Bitstream Interoperable with G.729*, 2006.

저자 약력

▶ 김 영 준 (Young-Joon Kim)



2004년 2월: 충북대학교 전기전자 및 컴퓨터 공학부 학사
2008년 2월: 충북대학교 전파공학과 석사
2009년 2월: 충북BIT연구중심대학육성사업단 전임 연구원
2009년 3월 ~ 현재: 충북대학교 전파통신 공학과 박사 수료

▶ 이 인 성 (In-Sung Lee)



1983년 2월: 연세대학교 전자공학과 학사
1985년 2월: 연세대학교 전자공학과 석사
1992년 12월: Texas A&M University 전기공학
학과 박사
1986년 5월 ~ 1987년 7월: 한국통신 연구
개발단 전임 연구원
1993년 2월 ~ 1995년 9월: 한국전자통신
연구원 이동 통신기술연구단 선임 연
구원
1995년 9월 ~ 현재: 충북대학교 전자정보
대학 교수