

채널 간 위상차 파라미터 생략 기법을 이용한 오디오 부호화

*김현휘, **김인철

서울시립대학교 전자전기컴퓨터공학부

*msplkhh@gmail.com, **rin@uos.ac.kr

An Audio Coding Technique Employing the Inter-channel Phase Difference Skip

*Hyun-Hwi Kim **Rin-Chul Kim

Dept. of Electrical & Computer Eng., University of Seoul

요 약

본 논문에서는 공간 오디오 부호화 기법인 MPEG 서라운드에서 공간 파라미터 전송 시 위상 파라미터를 생략하는 기법에 대해 다룬다. 기존 방법에서는 한 프레임이 모두 적은 위상차를 가지는 경우에도 정상적으로 처리하여 전송한다. 이러한 경우 위상차 파라미터를 생략하여 비트 효율을 향상시킬 수 있다. 스테레오 복원 과정에서 발생하는 채널 간 시간차에 기반해 설계된 양자화기를 생략 기법에 적용하면 기존에 비해 평균적으로 40 ~ 50% 정도의 위상 파라미터 절감 효과를 얻을 수 있다.

1. 서론

MPEG 서라운드(MPS : MPEG Surround) [1]는 공간 오디오 부호화(SAC : spatial audio coding) 기법의 한 종류로 스테레오 채널을 모노 채널로 다운믹스하는 과정에서 공간 파라미터를 추출하는데 여기는 채널 간 레벨차(CLD : inter-channel level difference), 채널 간 상관도(ICC : inter-channel correlation), 채널 간 위상차(IPD : inter-channel phase difference)가 있다. IPD는 채널 간 시간차(ITD : inter-channel time difference)로 나타낼 수 있는데 ITD는 SAC의 한 종류인 양이 단서 부호화(BCC : binaural cue coding) [2]에서 적용되던 파라미터로 이와 관련해 시간차로 인한 방향 인지의 청취 민감도를 추정하는 연구가 활발히 진행되어 왔다.

본 논문에서는 MPS에서의 스테레오 복원과 복원 시 실제 발생하는 시간차에 대해 알아보고 이를 통하여 청취 민감도를 고려한 양자화기를 설계한다. 제안된 양자화기를 제안된 생략 기법에 적용하고 비트 효율과 주관적 음질평가를 통해서 그 성능을 평가한다.

2. 스테레오 복원 시 발생하는 시간차

다운믹스 과정에서 추출된 공간 파라미터는 모노로 다운믹스된 신호와 함께 복호화기에서 스테레오로 복원되는데 CLD와 ICC에 의해 레벨과 상관도가 복원된 후, IPD와 계산된 OPD에 의해 위상 정정되는 과정을 거친다. 앞서 언급한대로 채널 간 위상차인 IPD는 ITD로 나타낼 수 있다. b 를 파라미터 밴드의 인덱스, $f_c[b]$ 를 해당 파라미터 밴드의 중심 주파수라고 하면 IPD와 ITD의 관계는 식 (1)과 같이 정의된다.

$$IPD[b] = 2\pi f_c[b]ITD[b] \quad (1)$$

IPD는 위상이 청취자의 방향 인지에 지대한 영향을 미치는 약 1.5kHz 미만의 파라미터 밴드에서 추출되는데, 연속된 파라미터 밴드 간의 위상을 청취자가 완전히 분리하여 듣지 못하므로 실제로는 단일 파라미터 밴드의 IPD에서 얻을 수 있는 ITD에 의한 시간차보다 작은 시간차를 인지하게 된다. 채널 간 레벨차

가 없다고 가정하고 $PB(i)$ 를 i 번째 파라미터 밴드, $PB(i+1)$ 를 그 다음 파라미터 밴드라고 할 때 단일 파라미터 밴드가 가지는 시간차를 그림으로 나타내면 그림 1(a)과 같다. 연속된 파라미터 밴드가 특정 채널을 지속적으로 선행시키면 그림 1(b), 지속적으로 선행시키지 않으면 그림 1(c)로 나타낼 수 있다.

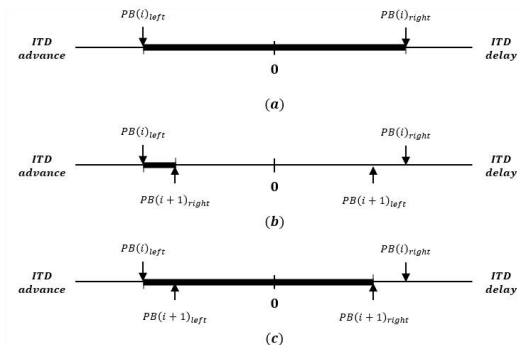


그림 1. 실제 청취 환경에서 인지하게 되는 시간차

시간차의 민감도는 ICC가 작을수록 둔감해지게 된다. 중심 주파수 인근 대역에 잡음이 분포한 환경에서 중심 주파수가 225Hz이고 ICC가 0.85일 땐 125us, 중심 주파수가 715Hz이고 ICC가 0.8일 땐 50us 이상의 ITD 문턱 값을 가진다[3]. 실험에 이용된 샘플들의 경우 ICC가 이보다 낮고 다양한 주파수 성분이 포함되어 있으므로 ITD 문턱 값은 이보다 높다.

연속된 파라미터 밴드가 특정 채널을 지속적으로 선행시킬 경우 선행 여부가 명확한데 반해 지속적으로 선행시키지 않을 경우 방향 인지에 어려움을 겪는다. IPD 양자화기의 $-\pi/2$ 에서 $\pi/2$ 까지 표현 레벨의 조합은 특정 채널을 지속적으로 선행시키지 않을 경우 위에서 살펴본 ITD 문턱 값 이하의 시간차를 발생시키기 때문에 0으로 처리가 가능하다.

3. 제안된 생략 기법

기존 방법에서는 한 프레임 전체의 IPD가 0일 경우에도 부

호화기에서 이를 모두 처리하는데 제안하는 생략 기법을 사용하면 생략 플래그를 이용해 복호화기에서의 1 비트를 복호화기 한 프레임 전체의 IPD로 처리할 수 있다.

생략의 빈도를 높이기 위해서 새로운 양자화기와 역양자화기를 제안한다. $-\pi/2$ 에서 $\pi/2$ 까지의 위상을 제외한 데드존을 가지는 구조로 표 1과 같다.

Representation level							
0	1.570796327f	2.094395102f	2.617993878f	3.141592654f	3.665191429f	4.188790205f	4.712388960f
0	$\frac{6\pi}{12}$ 12 (24)	$\frac{8\pi}{12}$ 12 (16) 12 (24)	$\frac{10\pi}{12}$ 12 (20) 12 (24)	$\frac{12\pi}{12}$ 12 (24) 12 (24)	$\frac{14\pi}{12}$ 12 (28) 12 (24)	$\frac{16\pi}{12}$ 12 (32) 12 (24)	$\frac{18\pi}{12}$ 12 (36) 12 (24)
Decision level							
1.308996939f	1.832595715f	2.356194490f	2.879793266f	3.403392041f	3.926990817f	4.450589593f	4.974188368f
10 π 24	14 π 24	18 π 24	22 π 24	26 π 24	30 π 24	34 π 24	38 π 24

표 1. 제안된 양자화, 역양자화기 구조

생략 기법의 적용을 위해 비트 스트림 구조를 살펴보면 그림 2와 같다. 그림 2(a)는 성공적으로 생략되었을 때의 위상 파라미터 비트 스트림 구조이며 그림 2(b)는 생략하지 못했을 경우이다. 기존의 비트 스트림 구조는 그림 2(c)와 같다.

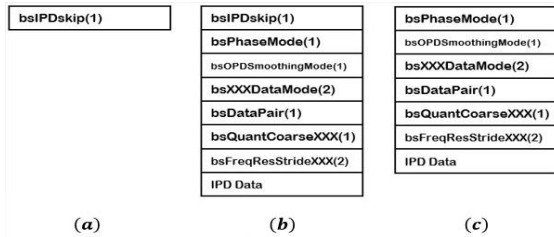


그림 2. 제안된 비트 스트림 구조

생략 기법을 사용하기 위해 생략 플래그를 추가한다. 한 프레임 전체 파라미터 밴드의 IPD가 0이라면 생략 플래그를 1로 설정하고 전송한다. 복호화기에서는 생략 플래그가 1이라면 해당 프레임 전체 파라미터 밴드의 IPD를 0으로 설정하고 처리한다. (b)는 (c)에 비해 비트 소모가 1 비트 증가하므로 생략하지 못했을 경우 비트 소모가 늘어나는데 생략된 경우 1 비트로만 표현되므로 평균적인 비트 소모는 낮을 것으로 예측할 수 있다.

4. 모의 실험

채널 간 위상차 파라미터 생략 기법의 효율을 여러 실험 샘플을 통해 확인한 결과는 표 2와 같다.

		Current coarse IPD	Proposed coarse IPD		Current coarse IPD	Proposed coarse IPD	
es01	Skip frame	0 (54)	179	Average	Skip frame percentage	0 %	48.32 %
	Bit cost	9044	2480		Bitcost	9044	4912
	Percentage	100%	27.42%		Bitcost percentage	100 %	54.31 %
twinkle_ff51	Skip frame	0 (0)	164				
	Bit cost	9044	3050				
	Percentage	100%	33.72%				
te15	Skip frame	0 (0)	2				
	Bit cost	9044	9206				
	Percentage	100%	101.79%				

표 2. 제안된 생략 기법을 적용하였을 때의 비트 효율

괄호 안의 수는 기존 양자화기에 제안된 생략 기법만 적용하였을 때 생략이 가능한 프레임 수이다. es01의 경우 IPD가 기존 양자화기 데드존 안쪽으로 밀집되어 있어 제안된 생략 기법만 적용하여도 비트 효율을 증가시킬 수 있으나 제안된 양자화기를 함께 적용하여 더 많은 프레임을 생략할 수 있었다. twinkle_ff51의 경우 기존 양자화기로는 생략 가능한 프레임이

없지만 제안된 양자화기를 도입하면 많은 프레임을 생략할 수 있어 비트 효율이 증가한다. te15의 경우 한 프레임에 다양한 위상 성분을 포함하고 있어 제안된 양자화기를 적용하여도 생략이 거의 불가능하다. 결과를 통해 전체적으로 약 45% 정도의 위상 파라미터를 감소시킬 수 있음을 확인하였다.

청취 평가에는 MUSHRA 실험 방법 [4]을 이용하였으며, 13명의 피실험자가 참여하였다. IPD 파라미터를 사용하지 않은 경우까지 3가지를 비교하였으며, 3개의 실험 아이টে에 대해 Sennheiser HD640 헤드폰으로 실험한 결과는 그림 3과 같다. es01의 경우 작은 IPD가 주를 이루어 많은 프레임이 생략되거나 음질 하락이 없다. twinkle_ff51의 경우 기존 양자화기로는 생략이 불가능하나 제안된 양자화기를 통해 음질 하락 없이 많은 프레임이 생략 가능하다. te15의 경우 생략이 거의 불가능하여 비트 효율이 기존과 흡사한데 음질에는 차이가 없다. es01을 제외한 샘플들은 위상 성분을 다양하게 가지고 있어 전체 프레임의 IPD를 0으로 처리하였을 때 음질 하락이 발생한다.

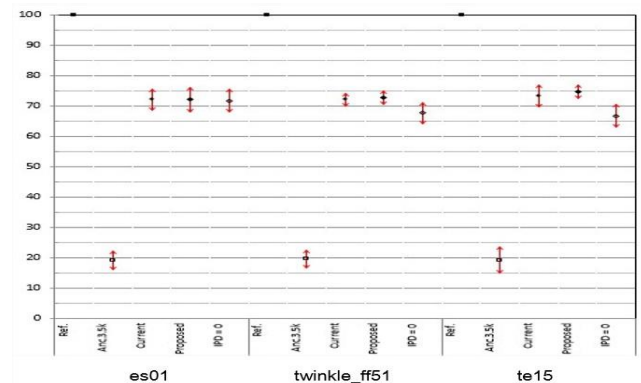


그림 3. MUSHRA 청취 실험 결과

5. 결론

본 논문에서는 채널 간 위상차 파라미터를 생략하기 위한 기법에 대해 다루었다. 이를 위해 복원 시 실제 발생하는 위상 차에 따른 시간차를 청취 민감도와 비교하였다. 이를 통해 새로운 양자화기를 설계하고 양자화기를 적용하였을 때의 비트 효율과 MUSHRA 청취 실험을 통해 결과를 확인하였다. 제안된 생략 기법을 사용하면 음질 하락 없이 비트 소모 효율이 증가할 수 있음을 확인하였다.

참고 논문

[1] ISO/IEC 23003-1 : Information technology – MPEG audio technologies – Part 1 : MPEG Surround, 2006.
 [2] C. Faller and F. Baumgarte, “Binaural cue coding – part I : psychoacoustic fundamentals and design principles,” IEEE Trans. On Speech and Audio Proc., vol. 11, no. 6, pp. 509–519, 2003.
 [3] B. Rakerd and W. Hartmann, “Localization of sound in room. V. Binaural coherence and human sensitivity to interaural time differences in noise,” J. Acoust. Soc. Am. Vol. 128, no. 5, pp. 3052–3063, 2010.
 [4] ITU-R BS.1534-1, Method for the subjective assessment of intermediate sound quality (MUSHRA), 2011.