



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0075501
(43) 공개일자 2024년05월29일

(51) 국제특허분류(Int. Cl.)
B60Q 5/00 (2006.01) G10K 15/04 (2006.01)
H04R 1/22 (2006.01)
(52) CPC특허분류
B60Q 5/008 (2013.01)
G10K 15/04 (2013.01)
(21) 출원번호 10-2022-0157534
(22) 출원일자 2022년11월22일
심사청구일자 없음

(71) 출원인
현대모비스 주식회사
서울특별시 강남구 테헤란로 203 (역삼동)
(72) 발명자
이재영
경기도 이천시 증신로325번길 39(송정동, 이천 라온프라이빗) 103동 1101호
(74) 대리인
특허법인 플러스

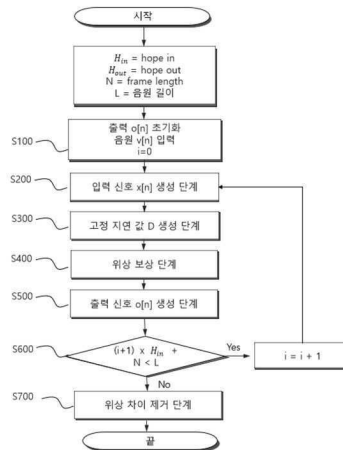
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 임베디드 시스템을 위한 가상 엔진음 생성 장치 및 그 제어 방법

(57) 요약

본 발명은 임베디드 시스템을 위한 가상 엔진음 생성 장치 및 그 제어 방법에 관한 것으로, 보다 상세하게는 소스 음원 정보를 입력 받는 입력부, 상기 소스 음원 정보의 전체 주파수에 대해 주파수 별 위상을 보상하여 보정된 음원 정보를 생성하는 신호 생성부 및 상기 보정된 음원 정보에 기초하여 최종 음원 정보를 제공하는 출력부를 포함하는 것을 특징으로 한다.

대표도 - 도2



(52) CPC특허분류

H04R 1/22 (2013.01)
B60Y 2200/91 (2013.01)
B60Y 2306/11 (2013.01)
G10K 2210/121 (2013.01)
G10K 2210/1282 (2013.01)
G10K 2210/51 (2013.01)
H04R 2499/13 (2013.01)
Y02T 10/64 (2020.08)

명세서

청구범위

청구항 1

소스 음원 정보를 입력 받는 입력부;

상기 소스 음원 정보에 기초하여 입력 신호를 생성하고, 상기 입력 신호의 전체 주파수에 대해 주파수 별 위상을 보상하여 보정된 음원 정보를 생성하는 신호 생성부; 및

상기 보정된 음원 정보에 기초하여 최종 음원 정보를 제공하는 출력부;를 포함하는 것

을 특징으로 하는 가상 엔진음 생성 장치.

청구항 2

제1항에 있어서,

상기 신호 생성부는,

상기 소스 음원 정보로부터 홑-인(hope-in) 간격으로 중복하여 추출한 프레임(frame)인 analysis frame을 입력 신호로 생성하는 것

을 특징으로 하는 가상 엔진음 생성 장치.

청구항 3

제2항에 있어서,

상기 신호 생성부는,

상기 analysis frame를 처리하여 합성 프레임인 synthesis frame을 생성하고, analysis frame과 상기 synthesis frame의 위치 차이에 기초하여 고정 지연 값을 생성하며,

상기 고정 지연 값에 기초하여 상기 입력 신호의 주파수 별 위상을 보상하는 것

을 특징으로 하는 가상 엔진음 생성 장치.

청구항 4

제3항에 있어서,

상기 신호 생성부는,

상기 입력 신호에 윈도우(window)를 곱하여 윈도우 신호를 생성하고, 상기 윈도우 신호에 FFT(Fast Fourier Transform)를 적용하여 주파수 영역 신호로 변환하며, 상기 주파수 영역 신호에 대해 상기 고정 지연 값을 주파수별로 적용하여 위상 보상 신호를 생성하고, 상기 위상 보상 신호에 IFFT(Inverse Fast Fourier Transform)를 적용하여 보정된 음원 정보를 출력하는 것

을 특징으로 하는 가상 엔진음 생성 장치.

청구항 5

제1항에 있어서,

상기 출력부는,

상기 보정된 음원 정보와 홉-아웃(hope-out) 간격에 기초하여 상기 최종 음원 정보를 제공하는 것을 특징으로 하는 가상 엔진음 생성 장치.

청구항 6

제3항에 있어서,

상기 홉-인(hope-in) 간격 및 상기 analysis frame의 길이 합과 상기 소스 음원 정보의 길이를 비교하는 비교부;를 더 포함하는 것

을 특징으로 하는 가상 엔진음 생성 장치.

청구항 7

제3항에 있어서,

저역 주파수 성분만을 통과시키는 LPF(Low pass filter);를 더 포함하고,

상기 LPF는,

상기 최종 음원 정보를 입력으로 하여 상기 synthesis frame과의 위상 차이를 제거하여 출력하는 것

을 특징으로 하는 가상 엔진음 생성 장치.

청구항 8

가상 엔진음 생성 장치의 제어 방법에 있어서,

(a) 소스 음원 정보를 입력 받는 단계;

(b) 상기 소스 음원 정보에 기초하여 입력 신호를 생성하고, 상기 입력 신호의 전체 주파수에 대해 주파수 별 위상을 보상하여 보정된 음원 정보를 생성하는 단계; 및

(c) 상기 보정된 음원 정보에 기초하여 최종 음원 정보를 제공하는 단계;를 포함하는 것

을 특징으로 하는 가상 엔진음 생성 장치의 제어 방법.

청구항 9

제8항에 있어서,

상기 단계 (a) 이후, 상기 단계 (b) 이전에,

(a-1) 상기 소스 음원 정보로부터 홉-인(hope-in) 간격으로 중복하여 추출한 프레임(frame)인 analysis frame을 입력 신호로 생성하는 단계;를 더 포함하는 것

을 특징으로 하는 가상 엔진음 생성 장치의 제어 방법.

청구항 10

제9항에 있어서,

상기 단계 (b)는,

상기 analysis frame를 처리하여 합성 프레임인 synthesis frame을 생성하고, analysis frame과 상기

synthesis frame의 위치 차이에 기초하여 고정 지연 값을 생성하며,
상기 고정 지연 값에 기초하여 상기 입력 신호의 주파수 별 위상을 보상하는 것
을 특징으로 하는 가상 엔진음 생성 장치의 제어 방법.

청구항 11

제10항에 있어서,

상기 단계 (b)는,

(b-1) 상기 입력 신호에 윈도우(window)를 곱하여 윈도우 신호를 생성하는 단계;

(b-2) 상기 윈도우 신호에 FFT(Fast Fourier Transform)를 적용하여 주파수 영역 신호로 변환하는 단계;

(b-3) 상기 주파수 영역 신호에 대해 상기 고정 지연 값을 주파수별로 적용하여 위상 보상 신호를 생성하는 단계; 및

(b-4) 상기 위상 보상 신호에 IFFT(Inverse Fast Fourier Transform)를 적용하여 보정된 음원 정보를 출력하는 단계;를 포함하는 것

을 특징으로 하는 가상 엔진음 생성 장치의 제어 방법.

청구항 12

제8항에 있어서,

상기 단계 (c)는,

상기 보정된 음원 정보와 홉-아웃(hope-out) 간격에 기초하여 상기 최종 음원 정보를 제공하는 것

을 특징으로 하는 가상 엔진음 생성 장치의 제어 방법.

청구항 13

제10항에 있어서,

상기 단계 (c) 이후,

(c-1) 상기 홉-인(hope-in) 간격 및 상기 analysis frame의 길이 합과 상기 소스 음원 정보의 길이를 비교하는 단계;를 더 포함하되,

상기 단계 (c-1)은,

상기 소스 음원 정보의 길이가 더 클 경우 상기 단계 (a-1)로 되돌아가고,

상기 소스 음원 정보의 길이가 더 작을 경우 상기 최종 음원 정보를 제공하는 것

을 특징으로 하는 가상 엔진음 생성 장치의 제어 방법.

청구항 14

제13항에 있어서,

상기 단계 (c-1) 이후,

(c-2) LPF(Low pass filter)가 상기 최종 음원 정보를 입력으로 하여 상기 synthesis frame과의 위상 차이를 제거하여 출력하는 단계;를 더 포함하는 것

을 특징으로 하는 가상 엔진을 생성 장치의 제어 방법.

발명의 설명

기술 분야

[0001] 본 발명은 가상 엔진을 생성 장치 및 그 제어 방법에 관한 것으로, 보다 상세하게는 임베디드 시스템을 위한 가상 엔진을 생성 장치 및 그 제어 방법에 관한 것이다.

배경 기술

[0002] EASD(Electronic Active Sound Design)는 전기차 실내 스피커에서 가상 엔진음을 생성하여 출력하는 방법으로 운전자의 주행 몰입감을 향상시킬 수 있다. 최근 전기차 브랜드를 보다 잘 나타내기 위해 전통적인 엔진 소리에서 벗어나 미래 지향적이나 우주선 등의 소리를 엔진음으로 사용하면서, 음원을 사용하는 wavetable이나 granular synthesizer 방법이 사용되고 있다. 음원을 사용하여 소리를 만들 때, 음원의 주파수를 변조하여 동시에 출력할 경우 소리를 더 풍부하게 할 수 있다. 대표적으로, TSM(Time Scale Modification) 방법인 phase vocoder를 사용하여 길이를 변경한 후 샘플링 주파수를 변경함으로써 원래 음원과 동일한 길이를 갖지만 다른 대역의 주파수를 갖는 소리를 생성할 수 있다. 또한, 엔진의 order 성분과 같이 주파수가 변조된 음원 크기를 주행 상황에 따라 조절함으로써 운전자가 음색이 다른 실제 엔진음을 들을 수 있다.

[0003] 대표적인 TSM 기술인 phase vocoder는 음원의 길이 변화 결과가 자연스럽게 때문에 다수의 오디오 처리 분야에서 사용되고 있다. Phase vocoder는 최대값을 갖는 주파수 성분에 대해서만 이전 주파수 응답의 phase와 비교하여 정확한 주파수를 예측하고, 예측된 주파수를 사용하여 해당 주파수 성분의 위상을 보정하기 때문에 급격한 위상 변화 없이 연속성을 달성할 수 있다.

[0004] 하지만, 음원의 수직 위상 일관성, 즉 사인치의 위상 관계는 상실된다. 수직 위상 일관성은 과도 현상과 같은 파형 변화에 영향을 주기 때문에 phase vocoder 출력 신호에서 과도 현상이 사라질 때가 있다. 또한, 음색에 방향성이 추가되는 등의 인위 구조가 추가된다. Phase vocoder를 EASD 임베디드 제어기에서 사용하기 위해서는 제한된 연산량 및 메모리 용량에서 구현이 가능해야 하는 반면, 고품질 음원 확장에 사용되는 phase vocoder는 analysis frame과 synthesis frame의 간격이 좁기 때문에 전체 음원을 처리하기 위해서 다수의 frame을 처리해야 한다. 따라서 많은 연산량이 필요할 수밖에 없다. 또한, 두 frame의 주파수 응답을 사용하므로 메모리 사용량도 많으며, 최대 주파수를 검출하고, 검출된 주파수에 대해서 정확한 주파수를 예측하는 제어 과정이 필요하기 때문에 하드웨어 가속기로 구현하기 어려운 문제점이 있다.

[0005]

선행기술문헌

특허문헌

[0006] (특허문헌 0001) 한국등록특허 10-1895656호(공개일자: 2014.12.04)

발명의 내용

해결하려는 과제

[0007] 본 발명은 상기한 바와 같은 문제점을 해결하기 위해 안출된 것으로, 본 발명에 의한 임베디드 시스템을 위한 가상 엔진을 생성 장치 및 그 제어 방법의 목적은 임베디드 환경에서 음원의 길이를 변경하기 위하여 시간 지연 값을 사용한 위상 보상 장치 및 그 제어 방법을 제공함에 그 목적이 있다.

과제의 해결 수단

[0008] 상기한 바와 같은 문제점을 해결하기 위한 본 발명의 다양한 실시예에 의한 임베디드 시스템을 위한 가상 엔진을 생성 장치는 소스 음원 정보를 입력 받는 입력부, 상기 소스 음원 정보에 기초하여 입력 신호를 생성하고,

상기 입력 신호의 전체 주파수에 대해 주파수 별 위상을 보상하여 보정된 음원 정보를 생성하는 신호 생성부 및 상기 보정된 음원 정보에 기초하여 최종 음원 정보를 제공하는 출력부를 포함하는 것을 특징으로 한다.

[0009] 또한, 상기 신호 생성부는 상기 소스 음원 정보로부터 홉-인(hop-in) 간격으로 중복하여 추출한 프레임(frame)인 analysis frame을 입력 신호로 생성하는 것을 특징으로 한다.

[0010] 또한, 상기 신호 생성부는 상기 analysis frame를 처리하여 합성 프레임인 synthesis frame을 생성하고, analysis frame과 상기 synthesis frame의 위치 차이에 기초하여 고정 지연 값을 생성하며, 상기 고정 지연 값에 기초하여 상기 입력 신호의 주파수 별 위상을 보상하는 것을 특징으로 한다.

[0011] 또한, 상기 신호 생성부는 상기 입력 신호에 윈도우(window)를 곱하여 윈도우 신호를 생성하고, 상기 윈도우 신호에 FFT(Fast Fourier Transform)를 적용하여 주파수 영역 신호로 변환하며, 상기 주파수 영역 신호에 대해 상기 고정 지연 값을 주파수별로 적용하여 위상 보상 신호를 생성하고, 상기 위상 보상 신호에 IFFT(Inverse Fast Fourier Transform)를 적용하여 보정된 음원 정보를 출력하는 것을 특징으로 한다.

[0012] 또한, 상기 출력부는 상기 보정된 음원 정보와 홉-아웃(hop-out) 간격에 기초하여 상기 최종 음원 정보를 제공하는 것을 특징으로 한다.

[0013] 또한, 상기 홉-인(hop-in) 간격 및 상기 analysis frame의 길이 합과 상기 소스 음원 정보의 길이를 비교하는 비교부를 더 포함하는 것을 특징으로 한다.

[0014] 또한, 저역 주파수 성분만을 통과시키는 LPF(Low pass filter)를 더 포함하고, 상기 LPF는 상기 최종 음원 정보를 입력으로 하여 상기 synthesis frame과의 위상 차이를 제거하여 출력하는 것을 특징으로 한다.

[0015] 또한, 본 발명의 다양한 실시예에 의한 임베디드 시스템을 위한 가상 엔진을 생성 장치의 제어 방법은 가상 엔진을 생성 장치의 제어 방법에 있어서, (a) 소스 음원 정보를 입력 받는 단계, (b) 상기 소스 음원 정보에 기초하여 입력 신호를 생성하고, 상기 입력 신호의 전체 주파수에 대해 주파수 별 위상을 보상하여 보정된 음원 정보를 생성하는 단계 및 (c) 상기 보정된 음원 정보에 기초하여 최종 음원 정보를 제공하는 단계를 포함하는 것을 특징으로 한다.

[0016] 또한, 상기 단계 (a) 이후, 상기 단계 (b) 이전에, (a-1) 상기 소스 음원 정보로부터 홉-인(hop-in) 간격으로 중복하여 추출한 프레임(frame)인 analysis frame을 입력 신호로 생성하는 단계를 더 포함하는 것을 특징으로 한다.

[0017] 또한, 상기 단계 (b)는 상기 analysis frame를 처리하여 합성 프레임인 synthesis frame을 생성하고, analysis frame과 상기 synthesis frame의 위치 차이에 기초하여 고정 지연 값을 생성하며, 상기 고정 지연 값에 기초하여 상기 입력 신호의 주파수 별 위상을 보상하는 것을 특징으로 한다.

[0018] 또한, 상기 단계 (b)는 (b-1) 상기 입력 신호에 윈도우(window)를 곱하여 윈도우 신호를 생성하는 단계, (b-2) 상기 윈도우 신호에 FFT(Fast Fourier Transform)를 적용하여 주파수 영역 신호로 변환하는 단계, (b-3) 상기 주파수 영역 신호에 대해 상기 고정 지연 값을 주파수별로 적용하여 위상 보상 신호를 생성하는 단계 및 (b-4) 상기 위상 보상 신호에 IFFT(Inverse Fast Fourier Transform)를 적용하여 보정된 음원 정보를 출력하는 단계를 포함하는 것을 특징으로 한다.

[0019] 또한, 상기 단계 (c)는 상기 보정된 음원 정보와 홉-아웃(hop-out) 간격에 기초하여 상기 최종 음원 정보를 제공하는 것을 특징으로 한다.

[0020] 또한, 상기 단계 (c) 이후 (c-1) 상기 홉-인(hop-in) 간격 및 상기 analysis frame의 길이 합과 상기 소스 음원 정보의 길이를 비교하는 단계를 더 포함하되, 상기 단계 (c-1)은 상기 소스 음원 정보의 길이가 더 클 경우 상기 단계 (a-1)로 되돌아가고, 상기 소스 음원 정보의 길이가 더 작을 경우 상기 최종 음원 정보를 제공하는 것을 특징으로 한다.

[0021] 또한, 상기 단계 (c-1) 이후 (c-2) LPF(Low pass filter)가 상기 최종 음원 정보를 입력으로 하여 상기 synthesis frame과의 위상 차이를 제거하여 출력하는 단계를 더 포함하는 것을 특징으로 한다.

발명의 효과

[0022] 상기한 바와 같은 본 발명의 다양한 실시예에 의한 임베디드 시스템을 위한 가상 엔진을 생성 장치 및 그 제어 방법에 의하면 시간 지연 값을 사용하여 위상을 보상함으로써, 각 단계의 연산을 고정할 수 있는 효과가 있다.

[0023] 또한, 수직 위상 일관성을 유지할 수 있어 연산량을 크게 감소시킬 수 있고, 메모리 사용량도 절반가량 줄일 수 있어 임베디드 시스템에 구현할 수 있다.

[0024] 뿐만 아니라, 하드웨어 가속기로 구현이 가능하므로 프로세싱 유닛의 연산량 점유를 최소화할 수 있으며, domain controller(DCU)에 EASD 기능이 포함될 때 작은 연산량으로 구현이 가능하므로 보다 많은 기능을 DCU에 탑재할 수 있어 UX 향상 및 비용 절감이 가능한 효과가 있다.

도면의 간단한 설명

[0025] 도 1은 종래의 phase vocoder 기반 음원 길이 변경 방법을 도시한 도면이며,
 도 2는 본 발명의 일실시예에 의한 가상 엔진음 생성 장치의 제어 방법을 도시한 도면이고,
 도 3은 본 발명의 일실시예에 의한 프레임별 위상 보상 방법을 도시한 도면이며,
 도 4는 본 발명에 의한 가상 엔진음 생성 장치의 제어 방법에 따라 생성되는 신호를 도시한 도면이고,
 도 5는 종래의 phase vocoder 방식과 본 발명의 일실시예에 의한 연산량 및 메모리 사용량을 비교 도시한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0026] 본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시에 의하여 달성되는 목적을 설명하기 위하여 이하에서는 본 발명의 바람직한 실시예를 예시하고 이를 참조하여 살펴본다.

[0027] 먼저, 본 출원에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로서, 본 발명을 한정하려는 의도가 아니며, 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함할 수 있다. 또한 본 출원에서, "포함하다" 또는 "가지다" 등의 용어는 명세서 상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0028] 본 발명을 설명함에 있어서, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략한다.

[0029] 도 1은 종래의 phase vocoder 기반 음원 길이 변경 방법을 도시한 도면이다.

[0030] 본 발명을 설명하기에 앞서, 도 1을 통해 종래 phase vocoder 기반 음원 길이 변경 방법에 대해 자세히 설명하겠다.

[0031] 먼저, 음원 길이를 변경하기 위해 사용되는 TSM(Time Scale Modification) 방법은 주파수 특성 변화 없이 음원의 길이를 변경하기 위해 사용되는 방법이다. 여기서, 음원의 길이를 변경하기 위해 입력 신호를 프레임(frame) 단위로 나눌 수 있다. 이때, 각 frame을 analysis frame이라고 하고, 각 frame과 frame은 일반적으로 중복되며, 시작 간격을 analysis hopsize라고 한다. 또한, 상기 analysis frame을 처리하여 synthesis frame을 생성하고, 상기 입력 신호와 다른 위치에 배치할 수 있다. 이때, 각 frame의 시작 간격을 synthesis hopsize라고 한다. 따라서, analysis hopsize > synthesis hopsize인 경우 출력 신호의 길이는 줄어들며, analysis hopsize < synthesis hopsize 인 경우 출력 신호의 길이는 증가한다. 한편, 상기 analysis frame을 상기 synthesis frame의 위치로 옮길 때 frame이 위치한 시점이 변화하기 때문에 급격한 phase shift를 방지하기 위해서는 phase 보상이 필요하다. 구체적으로, 도 1에 도시된 상기 analysis frame의 시작 위치를 t_a 라고 하고, 상기 synthesis frame의 시작 위치를 t_s 라고 할 때, 상기 analysis frame의 시작 위치를 synthesis frame의 시작 위치인 t_s 로 옮길 때 하기 식(수식 1)을 이용하여 위상을 보상할 수 있다.

[0032] [수식 1]

$$x(t - (t_s - t_a)) \leftrightarrow e^{-j2\pi f(t_s - t_a)} X(f)$$

[0033]

[0034] 또한, phase vocoder는 frame 재배치에 따른 급격한 phase shift를 방지하기 위해서 상기 analysis frame를 처리하기 전에 hanning window를 사용하여 weighting하고, FFT(Fast Fourier Transform)를 사용하여 주파수 응답을 구할 수 있다. 이후, peak를 갖는 주 주파수 성분을 검출하여(Peak Detection), 해당 주파수의 이전 analysis frame과 현재 frame의 위상차이로 하기 식(수식 2)을 이용하여 FFT 해상도보다 높은 정확도를 갖는 주파수를 추정(Frequency Estimation)할 수 있다.

[0035] [수식 2]

$$Phase\ Difference = 2\pi f \frac{Analysis\ Hopsize}{Sampling\ Frequency} + 2\pi n, n \in N$$

[0036]

[0037] 이때, 해당 analysis frame의 위치를 synthesis frame의 위치로 이동시키기 전에 상기 기재된 수식 1을 이용하여 위상을 보정할 수 있으며(Phase Adjustment), 이후, IFFT(Inverse Fast Fourier Transform)를 사용하여 위상 보상된 신호의 길이가 변경된 신호를 생성할 수 있다. 보다 구체적으로, 입력 신호의 길이가 I일 때, 출력 신호의 길이 0는 하기 식(수식 3)과 같다.

[0038] [수식 3]

$$O = \left\lceil I - FFT\ Size / Analysis\ Hopsize \right\rceil \times Synthesis\ Hopsize + FFT\ Size$$

[0039]

[0040] 도 2는 본 발명의 일실시예에 의한 가상 엔진음 생성 장치의 제어 방법을 도시한 도면이고,

[0041] 도 3은 본 발명의 일실시예에 의한 프레임별 위상 보상 방법을 도시한 도면이며,

[0042] 도 4는 본 발명에 의한 가상 엔진음 생성 장치의 제어 방법에 따라 생성되는 신호를 도시한 도면이다.

[0043] 도 2 내지 도 4를 통해 본 발명의 일실시예에 의한 가상 엔진음 생성 장치의 제어 방법에 대해 자세히 설명하겠다.

[0044] 도 2에 도시된 바와 같이 본 발명의 일실시예에 의한 길이가 L인 소스 음원 정보 v[n]의 길이를 변경하기 위한 방법은 출력인 최종 음원 정보 o[n]을 초기화하고, 상기 v[n]을 입력 받는 단계(n은 0 이상 N 미만의 자연수, N은 프레임의 길이)(S100), 상기 v[n]으로부터 hope-in 간격인 H_in 간격으로 중복하여 추출한 프레임(frame)인 analysis frame을 입력 신호인 x[n]으로 생성하는 단계(S200), 상기 analysis frame을 처리하여 합성 프레임인 synthesis frame을 생성하고, 상기 analysis frame과 상기 synthesis frame의 위치 차이에 기초하여 고정 지연 값인 D를 생성하는 단계(S300), 상기 D에 기초하여 상기 x[n]의 전체 주파수 성분에 대해 프레임 별로 위상을 보상하여 보정된 음원 정보인 y[n]을 산출하는 단계(S400), 상기 y[n]과 hope-out 간격인 H_out에 기초하여 상기 출력인 최종 음원 정보 o[n]을 생성하는 단계(S500), 상기 H_in 및 상기 analysis frame의 길이인 N의 합과 상기 L을 비교하고, 상기 비교 결과에 기초하여 상기 최종 음원 정보 o[n]을 출력하는 단계(S600)를 포함할 수 있다.

[0045] 구체적으로, 상기 입력 신호인 x[n]은 다음과 같이 나타낼 수 있다.

[0046] [수식 4]

$$x[n] = v[n + i \times H_in]$$

[0048] 또한, 처음 i의 값은 0일 수 있다.

[0049] 구체적으로, 고정 지연 값인 상기 D는 다음과 같이 나타낼 수 있다.

[0050] [수식 5]

$$D = i \times (H_out - H_in)$$

[0052] 보다 구체적으로, 본 발명에 일실시예에 의한 프레임별 위상 보상 방법에 대해 자세히 설명하겠다.

[0053] 상기 analysis frame을 상기 synthesis frame의 위치에 배치할 때 구조적으로 위상을 맞추기 위하여 도 3에 도

시된 바와 같이, 주파수 별로 선형 위상 조정을 수행할 수 있다.

[0054] 다시 말하면, 위상 조정을 하기 위해 윈도우(windowing) 단계, FFT 단계, 위상 보상 단계 및 IFFT단계를 포함할 수 있다.

[0055] 보다 구체적으로, 윈도우 단계는 상기 $x[n]$ 에 윈도우(window)인 $w[n]$ 을 곱하여 웨이팅(weighting)하여 윈도우 신호인 $x_w[n]$ 을 생성할 수 있다. 윈도우는 FFT를 수행할 때 화면에 포착된 신호의 앞과 끝이 일치하지 않는 경우에, 불연속에 의해 발생하는 이슈를 줄이거나 제거할 때 적용될 수 있다. 이때, 상기 윈도우는 해닝 윈도우(hanning window)일 수 있다.

[0056] FFT 단계는 상기 $x_w[n]$ 을 FFT(Fast Fourier Transform)을 이용하여 주파수 영역의 신호인 $X[k]$ 로 변환할 수 있다.

[0057] 위상 보상 단계는 생성된 상기 D 를 상기 $X[k]$ 에 대해 주파수별로 위상을 보상할 수 있으며, 위상 보상 신호인 $X_{phase}[k]$ 를 생성할 수 있다.

[0058] IFFT 단계는 생성된 상기 $X_{phase}[k]$ 에 IFFT(Inverse Fast Fourier Transform)를 적용하여 시간영역의 보정된 음원 정보인 $y[n]$ 을 출력할 수 있다.

[0059] 다시 말하면, 상기 윈도우 신호인 $x_w[n]$ 은 다음과 같이 나타낼 수 있다.

[0060] [수식 6]

$$x_w[n] = w[n]x[n]$$

[0061]

[0062]

[0063] 상기 윈도우 단계를 통해 반복되는 영역에서 발생하는 급격한 변화를 없앨 수 있다.

[0064] 상기 주파수 영역의 신호인 $X[k]$ 는 다음과 같이 나타낼 수 있다.

[0065] [수식 7]

$$X[k] = \sum_{n=0}^{N-1} x_w[n] e^{-j2\pi \frac{k}{N}n}$$

[0066]

[0067] 상기 위상 보상 신호인 $X_{phase}[k]$ 는 다음과 같이 나타낼 수 있다.

[0068] [수식 8]

$$X_{phase}[k] = X[k] e^{j2\pi \frac{k}{N}D}$$

[0069]

[0070] 상기 보정된 음원 정보인 $y[n]$ 은 다음과 같이 나타낼 수 있다.

[0071] [수식 9]

$$y[n] = \text{real} \left(\frac{1}{N} \sum_{k=0}^N X_{phase}[k] e^{j2\pi \frac{n}{N}k} \right)$$

[0072]

[0073] 결론적으로, 상기 analysis frame을 상기 synthesis frame의 위치 차이에 의한 위상 변화를 보상하기 위하여 상기 상술된 단계를 통해 선형적으로 위상 보상을 수행할 수 있으며, 이를 통해 수직 위상 일관성이 유지될 수 있

다. 또한, Synchronous Overlap and Add나 Phase Vocoder처럼 몇 개의 주파수에 대하여 위상이 보상되는 것이 아닌 전체 주파수 성분에 대한 위상 보상이 가능하다. 따라서, 프레임의 위치 변화에 따른 고정 지연 값인 상기 D를 입력할 경우, 각 단계의 연산이 고정적이기 때문에 하드웨어 가속기로 구현이 가능한 효과가 있다.

[0074] 다시 도 2를 보면, 상기 $y[n]$ 과 hope-out 간격인 H_{out} 에 기초하여 상기 출력인 최종 음원 정보 $o[n]$ 을 생성하는 단계(S500)는 다음과 같이 나타낼 수 있다.

[0075] [수식 10]

[0076] $o[n + I \times H_{out}] = o[n + I \times H_{out}] + y[n]$

[0077] 또한, 상기 단계 S600은 다음과 같이 나타낼 수 있다.

[0078] $(i + 1) \times H_{in} + N < L$

[0079] 다시 말하면, 상기 H_{in} 및 상기 analysis frame의 길이인 N 의 합과 상기 L 을 비교하고, 상기 L 이 더 클 경우, 상기 단계 S200으로 되돌아가 다시 수행할 수 있다. 이때, i 는 $i + 1$ 로 대체될 수 있다.

[0080] 또한, 상기 단계 S500 이후에 상기 H_{in} 및 상기 프레임의 길이인 N 의 합과 상기 L 을 비교하고, 상기 L 이 더 작은 경우 상기 출력 $o[n + I \times H_{out}]$ 을 최종 출력 $o[n]$ 으로 출력할 수 있다. 구체적으로, 상기 단계 S600 이후에 저역 주파수 성분만을 통과시키는 LPF(Low pass filter)가 출력인 상기 출력 $o[n + I \times H_{out}]$ 을 입력으로 하여 상기 synthesis frame과의 위상 차이를 제거하여 출력하는 단계(S700)를 더 포함할 수 있다. 따라서, 도 4와 같은 출력이 생성될 수 있다.

[0081] 종합적으로, 종래의 phase vocoder의 경우, 이전 analysis frame과 현재 analysis frame의 위상 변화를 사용하여 synthesis frame의 위치를 예측하므로 analysis frame의 간격이 작을수록 정확한 예측이 가능하다. 따라서, 프레임 간격이 커서 구성하는 주파수 성분이 달라질 경우, 왜곡이 크게 발생하는 문제가 있다. 하지만 상기 상술된 방법은 입력 음원의 위상이 보전되기 때문에 프레임 간격을 보다 넓게 가져갈 수 있다.

[0082] 도 5는 본 발명에 의한 위상 보상 방법과 종래의 phase vocoder의 연산량 및 메모리 사용량을 비교한 표이다.

[0083] 도 5는 연산량 및 메모리 사용량을 비교하기 위하여 약 100만 샘플을 갖는 음원을 사용하여 음원 길이를 2배로 변경한 것이다.

[0084] 본 발명에 의한 위상 보상 방법은 종래의 phase vocoder보다 4배 더 프레임의 간격을 늘릴 수 있으므로 4배 적은 연산량으로 음원 길이 변경이 가능하다. 또한, 이전 프레임의 주파수 응답을 저장하지 않아도 되기 때문에 메모리 사용량도 절반 가량만 사용하여 구현할 수 있다. 따라서 본 발명에 의한 위상 보상 방법을 사용할 경우, 임베디드 환경에서도 음원 길이 변경이 가능하여, 주행 상황에 따라 보다 풍부한 가상 엔진음을 생성할 수 있는 것이다.

[0085] 또한, 소스 음원 정보인 $v[n]$ 의 길이인 L 을 변경하기 위한 가상 엔진음 생성 장치는 입력부, 신호 생성부, 출력부를 포함할 수 있다. (미도시)

[0086] 입력부는 상기 소스 음원 정보인 $v[n]$ 을 입력 받는다.

[0087] 신호 생성부는 상기 소스 음원 정보에 기초하여 입력 신호를 생성하고, 상기 입력 신호의 전체 주파수에 대해 주파수 별 위상을 보상하여 보정된 음원 정보를 생성한다.

[0088] 출력부는 상기 보정된 음원 정보에 기초하여 최종 음원 정보를 제공한다.

[0089] 구체적으로, 상기 신호 생성부는 상기 소스 음원 정보인 $v[n]$ 으로부터 hope-in 간격인 H_{in} 간격으로 중복하여 추출한 프레임(frame)인 analysis frame을 입력 신호인 $x[n]$ 으로 생성할 수 있다. 이때, n 은 0 이상 N 미만의 자연수이고, N 은 프레임의 길이를 의미한다.

[0090] 또한, 상기 신호 생성부는 상기 analysis frame을 처리하여 합성 프레임인 synthesis frame을 생성하고, 상기 analysis frame과 상기 synthesis frame의 위치 차이에 기초하여 고정 지연 값인 D 를 생성할 수 있다. 더불어, 상기 D 에 기초하여 상기 $x[n]$ 의 전체 주파수 성분에 대해 프레임 별로 위상을 보상하여 보정된 음원 정보인 $y[n]$ 을 산출할 수 있다.

[0091] 한편, 출력부는 상기 $y[n]$ 과 hope-out 간격인 H_{out} 에 기초하여 상기 출력인 최종 음원 정보 $o[n]$ 을 생성할 수 있다.

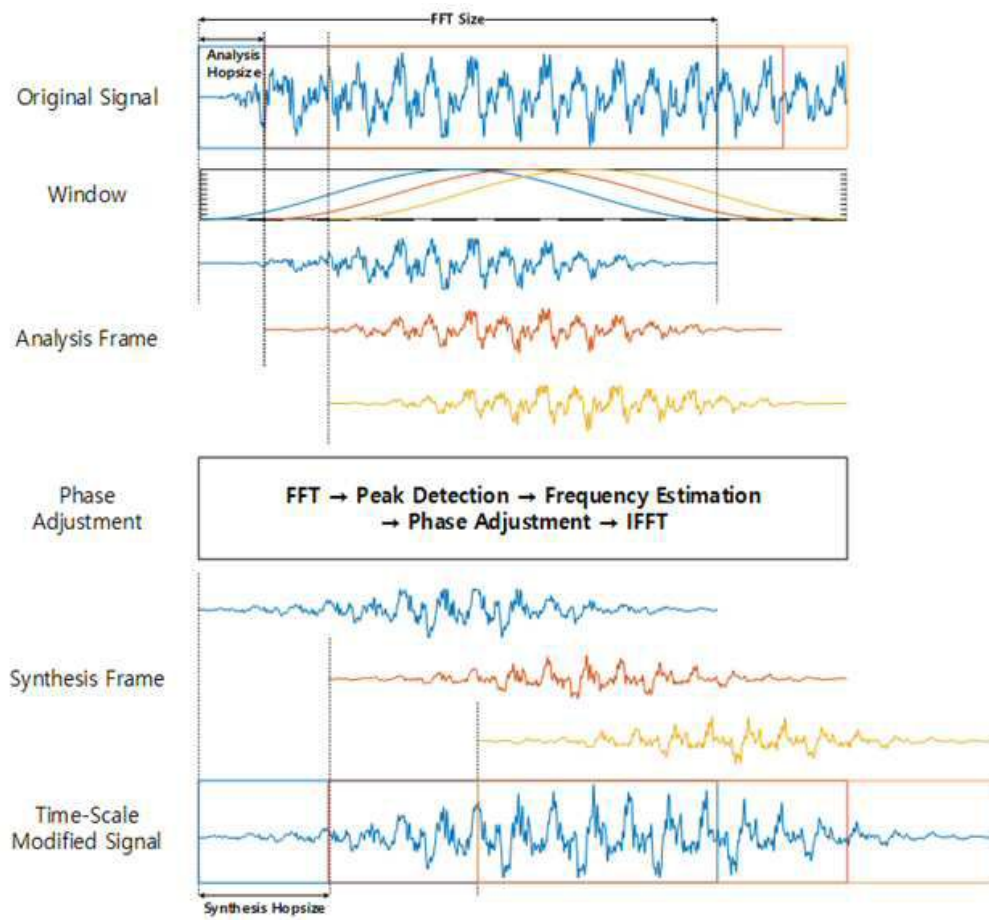
- [0092] 또한, 가상 엔진을 생성 장치는 비교부를 더 포함할 수 있다. (미도시)
- [0093] 비교부는 상기 H_{in} 및 상기 analysis frame의 길이인 N 의 합과 상기 L 을 비교하고, 상기 비교 결과에 기초하여 상기 $o[n]$ 을 출력할 수 있다.
- [0094] 구체적으로, 상기 신호 생성부는 상기 $x[n]$ 에 윈도우인 $w[n]$ 을 곱하고 웨이팅 하여 윈도우 신호인 $x_w[n]$ 을 생성할 수 있다. 또한 생성된 상기 $x_w[n]$ 을 FFT를 이용하여 주파수 영역 신호인 $X[k]$ 로 변환할 수 있다. 이후, 상기 신호 처리부에서 생성된 상기 D 를 상기 $X[k]$ 에 대해 주파수별로 위상을 보상하여 위상 보상 신호인 $X_{phase}[k]$ 를 생성할 수 있다. 또한, 생성된 상기 $X_{phase}[k]$ 에 IFFT를 적용하여 시간 영역의 보정된 음원 정보인 $y[n]$ 을 출력할 수 있다.
- [0095] 보다 구체적으로, 상기 윈도우는 해닝 윈도우일 수 있다.
- [0096] 또한, 본 발명은 저역 주파수 성분만을 통과시키는 LPF를 더 포함할 수 있다. 상기 LPF는 상기 비교부에서 출력되는 상기 $o[n]$ 을 입력으로 하여, 상기 synthesis frame과의 위상 차이를 제거하여 출력할 수 있다.
- [0097] 결론적으로, 상기 analysis frame을 상기 synthesis frame의 위치 차이에 의한 위상 변화를 보상하기 위하여 상기 상술된 장치를 통해 선형적으로 위상 보상을 수행할 수 있으며, 이를 통해 수직 위상 일관성이 유지될 수 있다. 또한, Synchronous Overlap and Add나 Phase Vocoder처럼 몇 개의 주파수에 대하여 위상이 보상되는 것이 아닌 전체 주파수 성분에 대한 위상 보상이 가능하다. 따라서, 프레임의 위치 변화에 따른 고정 지연 값인 상기 D 를 입력할 경우, 각 단계의 연산이 고정적이기 때문에 하드웨어 가속기로 구현이 가능한 효과가 있다.
- [0098] 또한, 종합적으로, 종래의 phase vocoder의 경우, 이전 analysis frame과 현재 analysis frame의 위상 변화를 사용하여 synthesis frame의 위치를 예측하므로 analysis frame의 간격이 작을수록 정확한 예측이 가능하다. 따라서, 프레임 간격이 커서 구성하는 주파수 성분이 달라질 경우, 왜곡이 크게 발생하는 문제가 있다. 하지만 상기 상술된 시스템은 입력 음원의 위상이 보전되기 때문에 프레임 간격을 보다 넓게 가져갈 수 있다.
- [0099] 이상에서 본 발명의 바람직한 실시 예에 대하여 설명하였으나, 본 발명에 개시된 실시예들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것일 뿐이다. 따라서, 본 발명의 기술 사상은 개시된 각각의 실시예 뿐 아니라, 개시된 실시예들의 조합을 포함하고, 나아가, 이러한 실시예에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 또한, 본 발명이 속하는 기술분야에서 통상의 지식을 가지는 자라면 첨부된 특허 청구범위의 사상 및 범주를 일탈함이 없이 본 발명에 대한 다수의 변경 및 수정이 가능하며, 그러한 모든 적절한 변경 및 수정은 균등물들로 본 발명의 범위에 속하는 것으로 간주되어야 할 것이다.

부호의 설명

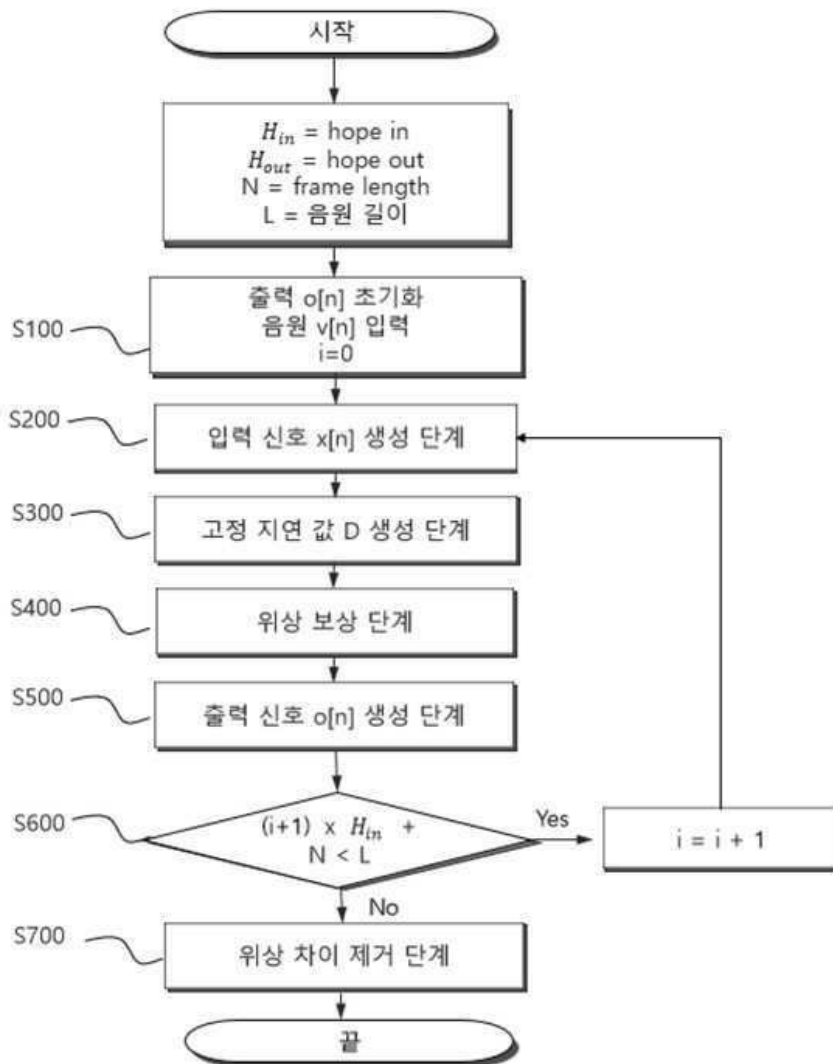
- [0100] S100 : 출력 $o[n]$ 초기화, 음원(소스 음원 정보) $v[n]$ 입력 단계
- S200 : 입력 신호 $x[n]$ 생성 단계
- S300 : 고정 지연 값 D 생성 단계
- S400 : 위상 보상 단계
- S500 : 출력 신호(최종 음원 정보) $o[n]$ 생성 단계
- S600 : 비교 단계
- S700 : 위상 차이 제거 단계

도면

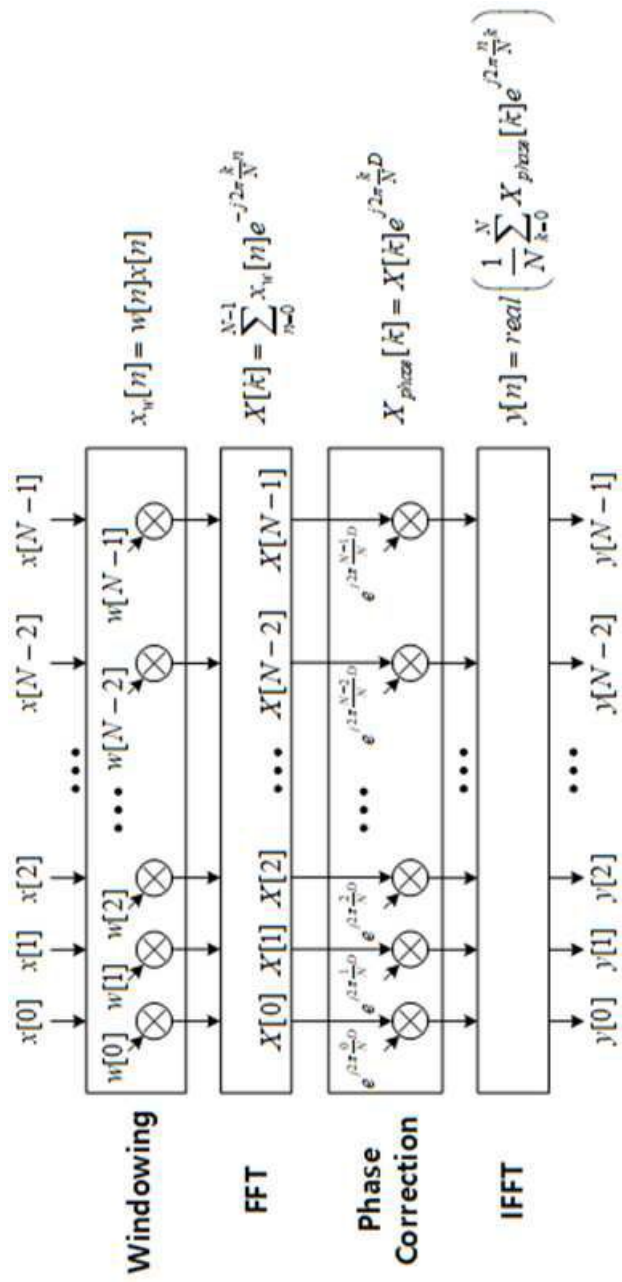
도면1



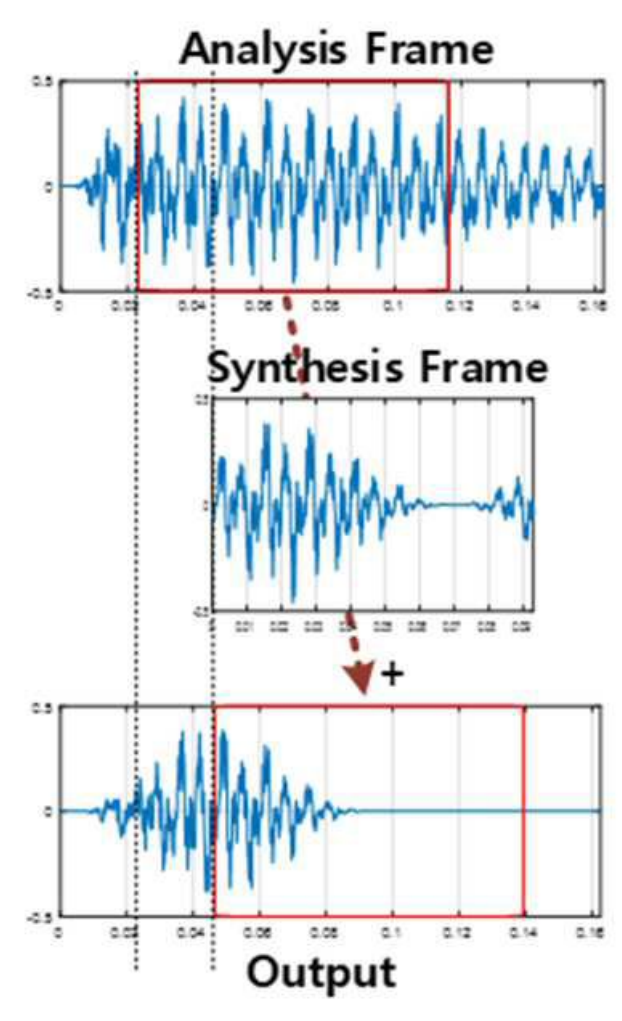
도면2



도면3



도면4



도면5

Item	Phase Vocoder (Hin = 256, Hout = 512)	Proposed Method (Hin = 1024, Hout = 2048)
#frames	3,779	944
#FFT	7,558	1,888
Memory(Byte)	32,768	16,384