

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2014년 4월 10일 (10.04.2014)



(10) 국제공개번호
WO 2014/054918 A1

- (51) 국제특허분류:
G01L 19/00 (2006.01) *G11B 20/10* (2006.01)
- (21) 국제출원번호: PCT/KR2013/008905
- (22) 국제출원일: 2013년 10월 4일 (04.10.2013)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보:
10-2012-0110337 2012년 10월 4일 (04.10.2012) KR
- (71) 출원인: 부산대학교 산학협력단 (PUSAN NATIONAL UNIVERSITY INDUSTRY-UNIVERSITY COOPERATION FOUNDATION) [KR/KR]; 609-735 부산시 금정구 부산대학교로 63번길 2 (장전동, 부산대학교), Busan (KR).
- (72) 발명자: 박주성 (PARK, Jusung); 609-310 부산시 금정구 구서동 1049 롯데캐슬 607-701, Busan (KR). 이동훈 (LEE, Donghoo); 612-741 부산시 해운대구 우1동 경동아파트 105-1601, Busan (KR). 허경철 (XU, Jingzhe); 609-735 부산시 금정구 장전2동 부산대학교 기전관 502호, Busan (KR). 정승표 (JUNG, Seungpyo); 612-818 부산시 해운대구 우1동 148-6, Busan (KR).
- (74) 대리인: 오위환 (OH, Wihwan); 135-080 서울시 강남구 역삼동 601-18 은성빌딩 5층, Seoul (KR).

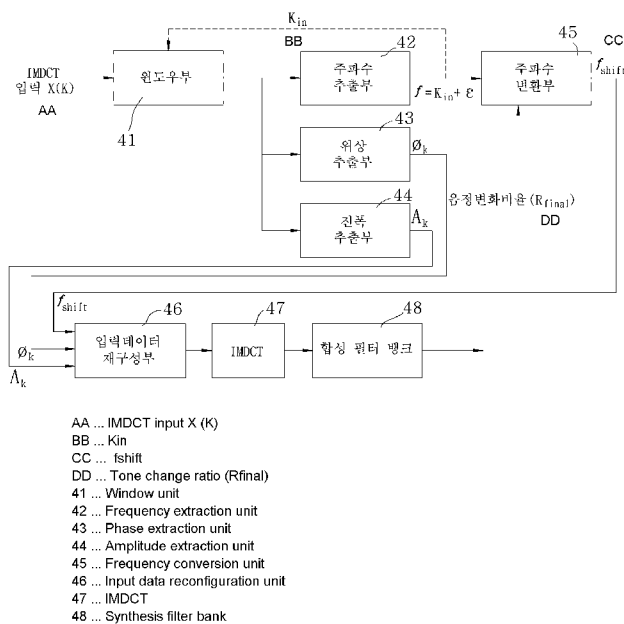
- (81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

공개:

— 국제조사보고서와 함께 (조약 제 21 조(3))

(54) Title: APPARATUS AND METHOD FOR VARYING TONE AND ACCELERATION OF AUDIO SIGNAL BY USING IMDCT INPUT SIGNAL

(54) 발명의 명칭: IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치 및 방법



(57) Abstract: The present invention relates to an apparatus and a method for varying the tone and the acceleration of an audio signal by using an inverse modified discrete cosine transform (IMDCT) to reduce the amount of computation and memory use by processing input data X(k) before converting into a time domain signal through the IMDCT, so that the tone and the acceleration of the audio signal can be changed, when varying the tone and the acceleration of the audio signal. The method comprises the steps of: determining the size of a window for determining the number of samples for extracting a frequency; extracting the frequency (k), phase, and amplitude of the input data X(k) needed in the IMDCT process from the window size that is selected; converting the extracted frequency, and reconfiguring the IMDCT input data by using the converted frequency and the extracted phase and amplitude; and outputting a coded audio signal by IMDCT processing and interpolating.

(57) 요약서: 본 발명은 오디오 신호의 음정 및 속도 가변시에 IMDCT(Inverse Modified Discrete Cosine Transform)를 통하여 시간의 영역의 신호로 변환하기 전에 입력 데이터 X(k)를 가공하여 음정과 속도를 변화시킬 수 있도록 하여 계산량 및 메모리의 사용을 줄일 수 있도록 한 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치 및

[다음 쪽 계속]



방법에 관한 것으로, 주파수 추출을 위한 샘플의 수를 결정하는 윈도우 크기 결정 단계; 선택된 윈도우 크기로 IMDCT 과정에 필요한 입력 데이터 $X(k)$ 의 주파수(k), 위상, 진폭을 추출하는 단계; 추출된 주파수를 변환하고, 변환된 주파수와 추출된 위상과 진폭을 이용하여 IMDCT 입력 데이터를 재구성하는 단계; IMDCT 처리 및 보간을 하여 코딩된 오디오 신호를 출력하는 단계;를 포함한다.

명세서

발명의 명칭: I M D C T 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치 및 방법

기술분야

- [1] 본 발명은 오디오 신호의 음정 및 속도가변에 관한 것으로, 구체적으로 IMDCT(Inverse Modified Discrete Cosine Transform)를 통하여 시간영역의 신호로 변환하기 전에 IMDCT 입력 데이터 $X(k)$ 를 가공하여 음정을 변화시킬 수 있도록 하여 계산량 및 메모리의 사용을 줄일 수 있도록 한 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치 및 방법에 관한 것이다.

배경기술

- [2] 일반적으로 많은 양의 오디오 데이터를 CD, 하드디스크, 이동저장매체에 저장하거나 유무선 방식으로 전송하기 위해서 오디오 데이터를 압축한다. 오디오 데이터 압축방법에는 시간영역에서 압축하는 방식과 주파수 영역에서 압축하는 방식이 있다.
- [3] 주파수 영역에서 압축하는 방식은 압축율이 높을 뿐만 아니라 음질도 좋으므로 시간영역의 오디오 신호를 주파수영역으로 변환하여 심리음향모델과 기타의 방식을 이용하여 압축한다.
- [4] MP3(MPEG 3)나 AAC(Advanced Audio Coding) 방식은 오디오 신호를 주파수 영역에서 압축하는 방식을 사용하고 있다. 사람이 오디오 신호를 청취하기 위해서는 압축된 데이터를 풀어서 압축되기 전 주파수영역의 신호로 복원하고 다시 시간영역의 신호로 변환해야 한다.
- [5] 시간영역으로 변환된 오디오 신호를 정상속도보다 빨리 재생하면 음정(tone)이 높아지고, 느리게 재생하면 음정이 낮아진다. 따라서 음정변화 없이 정상속도보다 빠르거나 느리게 재생하기 위하여 SOLA(Synchronous Overlap and Add)와 같은 방법을 이용하여 음정을 변화시킨다. 음정이 변환된 신호를 정상속도로 재생하면 변환된 음정이 그대로 재생된다. 음정이 변환된 오디오 신호를 재생속도에 따라 음정변화 없이 재생 시키거나, 음정과 재생속도를 동시에 가변시킬 수 있다.
- [6] 종래 기술의 경우 주파수영역에서 압축된 오디오 신호의 음정이나 속도를 가변시키고자 하는 경우에는 도 1에서와 같이, 시간영역의 신호로 일단 변환시킨 후 음정이나 속도를 가변 시킨다. 이러한 과정에서 시간영역에서 음정이나 속도를 가변 시키기 때문에 추가적인 계산이 요구되고 계산과정의 데이터를 저장하기 위하여 많은 메모리가 필요하게 된다.
- [7] MP3와 AAC방식에서 시간영역의 오디오 신호를 주파수영역의 신호로 압축하는 과정은 도 2에서와 같다.
- [8] 이 방식들에서 오디오 신호는 필터 뱅크(filter bank)(21)를 통하여 여러 대역의

주파수 성분으로 분해되고, 필터 뱅크(21)에서 분해된 성분은 MDCT 블록(22)에서 MDCT(Modified Discrete Cosine Transform)을 통하여 시간영역에서 주파수영역으로 변환된다.

- [9] 주파수영역으로 변환된 신호는 양자화부(23)에서 양자화(quantization)되고 코딩부(24)에서 코딩(coding)을 통하여 노이즈가 적고 손실이 적은 형태로 압축된다. 압축된 데이터는 비트스트림 인코딩부(25)에서 사이드 정보(side information)와 함께 비트스트림(bitstream)으로 만들어져 저장되거나 전송된다.
- [10] 주파수영역으로 압축된 신호를 시간영역으로 변환하는 일반적인 과정은 도 3에서와 같다.
- [11] 비트스트림 디코딩부(31)에서 부호화된 비트스트림으로부터 사이드 정보와 압축된 데이터를 분리한다.
- [12] 사이드 정보는 복호화 방법에 대한 정보를 포함하고 있으므로 주파수 영역으로 압축된 데이터를 IMDCT 블록(32)에서 IMDCT(Inverse Modified Discrete Cosine Transform)를 통하여 시간 영역으로 변환할 수 있다.
- [13] MP3나 AAC 방식으로 압축된 오디오 데이터는 여러 주파수 대역으로 나누어 압축하고 복원하므로 복원된 데이터를 합성 필터 뱅크(33)를 통하여 재구성하여 시간영역의 오디오 신호로 변환된다. 종래 기술의 경우에는 도 3에서와 같은 과정을 통하여 시간영역의 신호로 완전하게 변환시킨 후 오디오 신호의 음정 변화와 속도를 변화시키는 단계를 거치게 된다.
- [14] 시간영역의 오디오 신호 $x(n)$ 를 MDCT과정을 거쳐 주파수 정보 $X(k)$ 로 변환시키는 과정은 수학식 1과 같다. 수학식 1에서 $w(n)$ 은 윈도우 함수(window function)이고 수학식 2와 같이 표시되며, 수학식 1, 2에서 N 은 분석하는 윈도우 크기를 의미한다. (이하 모든 수학식에서 N 은 분석 윈도우의 크기를 의미한다.) 분석 윈도우 내에 있는 시간영역의 신호 $x(n)$ 에 MDCT을 하면 수학식 1과 같은 주파수 영역의 데이터 $X(k)$ 를 얻을 수 있다.
- [15] 주파수 영역으로 변환된 정보 $X(k)$ 를 시간 영역의 오디오 신호로 변환하는 IMDCT 과정은 수학식 3과 같다. 이러한 과정에서 주파수 정보 $X(k)$ 를 가공하여 오디오 신호의 주파수를 변화시켜 음정을 변화시키는 효과를 얻을 수 있다.

- [16] 수학식 1

$$X(k) = \sqrt{\frac{2}{N}} \sum_{n=0}^{2N-1} w(n)x(n) \cos \frac{(2n+1+N)(2k+1)\pi}{4N}$$

- [17] 수학식 2

$$w(n) = \sin \frac{(2n+1)\pi}{4N}, \quad n = 0, 1, \dots, 2N-1$$

- [18] 수학식 3

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) \cos \left\{ \frac{\pi}{N} \left(n + \frac{1}{2} + \frac{N}{2} \right) \left(k + \frac{1}{2} \right) \right\}$$

- [19] 종래기술에서는 주파수영역의 오디오 신호를 시간영역의 신호로 변환시킨 후 음정이나 속도를 변화시키기 때문에 많은 계산량이 요구되고 계산과정의 데이터를 저장하기 위하여 많은 메모리가 필요하게 된다.

발명의 상세한 설명

기술적 과제

- [20] 본 발명은 종래의 기술에서 주파수 영역으로 변환된 오디오 신호의 음정과 음정 변화없이 재생속도를 가변시키는 과정에서 많은 계산량과 메모리가 요구되는 문제점을 해결하기 위한 것으로, IMDCT 과정에서 오디오 신호의 음정을 용이하게 변화시키기 위해서 IMDCT 입력 데이터 $X(k)$ 에서 다양한 주파수성분, 각 주파수의 진폭과 위상을 추출하여 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치 및 방법을 제공하는데 그 목적이 있다.
- [21] 본 발명은 주파수 영역의 신호를 시간영역의 신호로 변환하는 단계인 IMDCT(Inverse Modified Discrete Cosine Transform) 과정의 입력 데이터 $X(k)$ 를 가공하여 음정을 변화시킬 수 있도록 한 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치 및 방법을 제공하는데 그 목적이 있다.
- [22] 본 발명은 주파수 영역으로 변환된 오디오 신호의 음정 및 속도가변 방법에 있어, IMDCT 과정에서 주파수 변환이 가능하게 IMDCT 입력신호를 적절하게 변환시키는 전처리 장치 및 방법, IMDCT를 통하여 시간영역의 신호로 변환된 오디오의 재생속도를 IMDCT 입력신호 전처리 단계와 연동시켜 음정과 음정변화 없이 재생속도를 변화시킬 수 있게 보간(Interpolation)하여 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치 및 방법을 제공하는데 목적이 있다.
- [23] 본 발명은 MDCT를 통하여 주파수영역으로 변환된 오디오 신호의 주파수 영역 데이터를 IMDCT를 통하여 시간영역의 신호로 변환하기 전에 다양한 주파수 성분을 추출하고 각 주파수 성분의 진폭과 위상을 이용하여 주파수 변환이 가능하게 IMDCT 입력 데이터를 재생성하여 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치 및 방법을 제공하는데 그 목적이 있다.
- [24] 본 발명은 오디오 신호의 압축율을 높이기 위하여 주파수영역으로 변환된 신호의 주파수 정보를 활용하여 음정과 속도를 변화시키는 방법으로 시간영역에서 음정 변환시키는 과정을 생략할 수 있도록 한 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치 및 방법을 제공하는데 그 목적이 있다.
- [25] 본 발명의 목적들은 이상에서 언급한 목적들로 제한되지 않으며, 언급되지 않은 또 다른 목적들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제 해결 수단

- [26] 이와 같은 목적을 달성하기 위한 본 발명에 따른 IMDCT 입력신호를 이용한

오디오 신호의 음정 가변 장치는 처리할 샘플의 윈도우 크기를 결정하는 윈도우부; 상기 윈도우부에서 결정된 윈도우 크기로 IMDCT 입력 데이터 $X(k)$ 의 주파수와 위상을 추출하는 주파수 추출부 및 위상 추출부, 진폭을 추출하는 진폭 추출부; 상기 추출된 주파수를 변환하는 주파수 변환부; 상기 변환된 주파수, 추출된 위상과 진폭을 이용하여 IMDCT 입력 데이터를 재구성하는 입력 데이터 재구성부; IMDCT를 통하여 상기 재구성된 입력 데이터를 시간 영역으로 변환하는 IMDCT 블록; 상기 IMDCT를 통하여 시간영역으로 변환된 여러 서브밴드 성분을 합성하는 합성 필터뱅크;를 포함하는 것을 특징으로 한다.

[27] 여기서, 상기 윈도우부에서 IMDCT 입력 데이터인 $X(k)$ 를 분석하고 재구성하는 데 필요한 윈도우를 결정하기 위하여, 분석 대상이 되는 주파수 영역의 윈도우 크기를 다른 영역에 비하여 상대적으로 작게 하는 것을 특징으로 한다.

[28] 그리고 상기 윈도우부에서 IMDCT 입력 데이터인 $X(k)$ 를 분석하고 재구성하는데 필요한 윈도우를 결정하기 위하여, 서브밴드와 서브밴드, 프레임과 프레임 경계에서 분석윈도우를 중첩시켜 설정하고, 분석윈도우 내에서 스펙트럼을 계산을 통하여 정수주파수(k_{in})를 찾아 그 주파수를 중심으로 분석윈도우를 구성하는 것을 특징으로 한다.

[29] 다른 목적을 달성하기 위한 본 발명에 따른 IMDCT 입력신호를 이용한 오디오 신호의 음정 가변 방법은 IMDCT를 통하여 시간의 영역의 신호로 변환하기 전에 IMDCT에 입력되는 데이터($X(k)$)를 처리하여 음정을 변화시키기 위하여, 주파수 추출을 위한 샘플의 수를 결정하는 윈도우 크기 결정 단계; 선택된 윈도우 크기로 IMDCT 과정에 필요한 입력 데이터 $X(k)$ 의 주파수(k), 위상, 진폭을 추출하는 단계; 추출된 주파수를 변환하고, 변환된 주파수와 추출된 위상과 진폭을 이용하여 IMDCT 입력 데이터를 재구성하는 단계; 주파수 영역의 데이터를 시간영역으로 변환하는 IMDCT 단계; IMDCT 과정에서 만들어진 다양한 주파수의 시간영역 신호를 합성하는 합성필터뱅크 단계;를 포함하는 것을 특징으로 한다.

[30] 여기서, 상기 IMDCT 입력 데이터 $X(k)$ 의 주파수 추출 단계에서, IMDCT 입력 데이터 $X(k)$ 의 주파수를 정수부(k_{in})와 소수부(ϵ)로 나누어 $f = k_{in} + \epsilon$ 로 표시하고, 이웃하는 세 개의 주파수 성분 $X(k_{in}-1)$, $X(k_{in})$, $X(k_{in}+1)$ 을 이용하여 분석하는 윈도우 내에 존재하는 모든 스펙트럼 값을 구하여 그 중 가장 큰 스펙트럼 값(S_k)을 만드는 k_{in} 를 정수부 주파수 성분(k_{in})으로 하는 것을 특징으로 한다.

[31] 그리고 상기 주파수 성분의 소수부분 ϵ 을,

[32] $\alpha = -\frac{X(k_{in}-1)}{X(k_{in}+1)}$ 라고 두면,

[33] $\alpha \neq 1$ 인 경우에 대해서 $\epsilon_1 = \frac{3 + \alpha - \sqrt{\alpha^2 + 14\alpha + 1}}{2(1 - \alpha)}$

[34] $\beta = \frac{X(k_{in}-2)}{X(k_{in}+2)}$ 라고 두면,

- [35] $\beta \neq 1$ 인 경우에 대해서
$$\varepsilon_p = \frac{5 + 3\beta - \sqrt{\beta^2 + 62\beta + 1}}{2(1 - \beta)}$$
- [36] 2 종류를 구하고, α 와 β 중에서 어느 것을 사용한 것인지를 결정은, 윈도우 내의 가장 큰 주파수 성분 $X(k_{in})$ 의 절대값과 k_{in} 의 스펙트럼 값(S_k)의 비율을 $\lambda = \frac{|k_{in}|}{S_k}$ 로 정의하고, 그 비율이 특정 문턱값(threshold) λ_0 과 비교하여 작으면 α , 그 외의 경우엔 β 를 선택하여 주파수 성분의 소수부분인(ε)을 얻는 것을 특징으로 한다.
- [37] 그리고 IMDCT 데이터 $X(k)$ 의 위상 ϕ_k 를 추출하기 위해서 추출한 IMDCT 정수부 주파수 성분(k_{in})을 이용하여 계산하는 것을 특징으로 한다.
- [38] 그리고 상기 진폭을 추출하는 단계에서, 정수부 주파수 성분(K_{in})과 소수부(ε) 주파수 성분을 이용하여 IMDCT 입력 데이터의 진폭 A_k 를 구하는 것을 특징으로 한다.
- [39] 그리고 상기 IMDCT 입력 데이터를 재구성하는 단계에서, 상기 윈도우 선택과정, 주파수 추출과정, 위상 추출과정, 주파수 변환 과정으로부터 얻은 윈도우 크기(N), 주파수($f = k_{in} + \varepsilon$), 위상(ϕ_k), 변환된 주파수 $f_{shift} = f(1 + R_f)$ 를 이용하여 IMDCT 입력 $X'(k)$ 를 재구성하는 것을 특징으로 한다.
- [40] 또 다른 목적을 달성하기 위한 본 발명에 따른 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치는 처리할 샘플의 윈도우 크기를 결정하는 윈도우부; 상기 윈도우부에서 결정된 윈도우 크기로 IMDCT 입력 데이터 $X(k)$ 의 주파수와 위상을 추출하는 주파수 추출부 및 위상 추출부, 진폭을 추출하는 진폭 추출부; 상기 추출된 주파수를 변환하는 주파수 변환부; 상기 변환된 주파수, 추출된 위상과 진폭을 이용하여 IMDCT 입력 데이터를 재구성하는 입력 데이터 재구성부; IMDCT를 통하여 재구성된 입력 데이터를 시간 영역으로 변환하는 IMDCT 블록; IMDCT를 통하여 시간영역으로 변환된 여러 서브밴드 성분을 합성하는 합성 필터뱅크; 상기 합성 필터뱅크에서 출력되는 오디오 신호의 샘플링 간격을 조절하여 재생속도와 음정을 변화시키는 보간부(Interpolator);를 포함하는 것을 특징으로 한다.
- [41] 여기서, 상기 IMDCT 블록에서 오디오 신호의 음정을 변화시키기 위하여 IMDCT 입력 데이터 $X(k)$ 로부터 정현파 성분으로 분해하여 추출한 후, 원하는 만큼 주파수를 변환하여 IMDCT 입력을 재구성하는 것을 특징으로 한다.
- [42] 그리고 상기 윈도우부에서 IMDCT 입력 데이터인 $X(k)$ 를 분석하고 재구성하는데 필요한 윈도우를 결정하기 위하여, 분석 대상이 되는 주파수 영역의 윈도우 크기를 다른 영역에 비하여 상대적으로 작게 하는 것을 특징으로 한다.
- [43] 그리고 상기 윈도우부에서의 IMDCT 입력 데이터인 $X(k)$ 를 분석하고 재구성하는데 필요한 윈도우를 결정하는 데 있어 서브밴드와 서브밴드, 프레임과 프레임 경계에서 분석윈도우를 중첩시켜 설정하고, 분석윈도우 내에서 스펙트럼을 계산을 통하여 정수주파수(k_{in})를 찾아 그 주파수를 중심으로 분석윈도우를 구성하는 것을 특징으로 한다.

- [44] 또 다른 목적을 달성하기 위한 본 발명에 따른 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 방법은 IMDCT를 통하여 시간의 영역의 신호로 변환하기 전에 IMDCT에 입력되는 데이터($X(k)$)를 처리하여 음정을 변화시키기 위하여, 주파수 추출을 위한 샘플의 수를 결정하는 윈도우 크기 결정 단계;선택된 윈도우 크기로 IMDCT 과정에 필요한 입력 데이터 $X(k)$ 의 주파수(k), 위상, 진폭을 추출하는 단계;추출된 주파수를 변환하고, 변환된 주파수와 추출된 위상과 진폭을 이용하여 IMDCT 입력 데이터를 재구성하는 단계;IMDCT 처리 및 보간을 하여 코딩된 오디오 신호를 출력하는 단계;를 포함하는 것을 특징으로 한다.
- [45] 여기서, 상기 IMDCT 처리 과정에서 원하는 만큼 주파수를 변화시키기 위해서 IMDCT 입력 데이터 $X(k)$ 의 주파수, 위상, 진폭을 분리하여 방정식 형태로 표시하거나 그 방정식을 look-up 테이블 형태로 저장해두고 사용하는 과정을 포함하는 것을 특징으로 한다.
- [46] 그리고 IMDCT 입력 데이터 $X(k)$ 의 주파수 추출 단계에서, IMDCT 입력 데이터 $X(k)$ 의 주파수를 정수부(k_{in})와 소수부(ϵ)로 나누어 $f = k_{in} + \epsilon$ 로 표시하고, 이웃하는 세 개의 주파수 성분 $X(k_{in}-1)$, $X(k_{in})$, $X(k_{in}+1)$ 을 이용하여 분석하는 윈도우 내 존재하는 모든 스펙트럼 값을 구하여 그 중 가장 큰 스펙트럼 값(S_k)을 만드는 k_{in} 를 정수부 주파수 성분(k_{in})으로 하는 것을 특징으로 한다.
- [47] 그리고 상기 주파수 성분의 소수부분 ϵ 을,
- [48]
$$\alpha = -\frac{X(k_{in}-1)}{X(k_{in}+1)}$$
라고 두면,
- [49] $\alpha \neq 1$ 인 경우에 대해서
$$\epsilon_1 = \frac{3 + \alpha - \sqrt{\alpha^2 + 14\alpha + 1}}{2(1 - \alpha)}$$
- [50] $\beta = \frac{X(k_{in}-2)}{X(k_{in}+2)}$ 라고 두면,
- [51] $\beta \neq 1$ 인 경우에 대해서
$$\epsilon_2 = \frac{5 + 3\beta - \sqrt{\beta^2 + 62\beta + 1}}{2(1 - \beta)}$$
- [52] 2 종류를 구하고, α 와 β 중에서 어느 것을 사용한 것인지를 결정은, 윈도우 내의 가장 큰 주파수 성분 $X(k_{in})$ 의 절대값과 k_{in} 의 스펙트럼 값(S_k)의 비율을
$$\lambda = \frac{|X_{in}|}{S_k}$$
로 정의하고, 그 비율이 특정 문턱값(threshold) λ_0 과 비교하여 작으면 α , 그 외의 경우엔 β 를 선택하여 주파수 성분의 소수부분인(ϵ)을 얻는 것을 특징으로 한다.
- [53] 그리고 IMDCT 데이터 $X(k)$ 의 cosine 함수의 위상 ϕ_k 를 추출하기 위해서 추출한 IMDCT 정수부 주파수 성분(k_{in})을 이용하여 계산하는 것을 특징으로 한다.
- [54] 그리고 상기 진폭을 추출하는 단계에서,정수부 주파수 성분(K_{in})과 소수부(ϵ) 주파수 성분을 이용하여 IMDCT 입력 데이터의 진폭 A_k 를 구하는 것을 특징으로 한다.
- [55] 그리고 음정변화 없이 재생속도를 변화시키기 위하여, 상기 IMDCT 처리 및 보간을 하여 코딩된 오디오 신호를 출력하는 단계와 연계하여 원래 속도를 1로

할 때 가변속도, 원신호의 샘플링 간격(t_s), 새롭게 만들 신호의 샘플링 간격(t'_s), (원래속도/가변속도) = $t_s/t'_s = R_t$ 관계를 이용하여 R_t 를 구한 후 ($R_f \times R_t = 1$) 되게 R_f 를 결정한 다음, 상기 추출한 IMDCT 입력 데이터 $X(k)$ 의 주파수 성분(k)을 $f_{\text{shift}} = f(1+R_f)$ 변화시키는 것을 특징으로 한다.

- [56] 그리고 음정과 재생속도를 동시에 변화시키려는 경우에는 재생속도로부터 (원래속도/가변속도) = R_t 로부터 R_t 를 구하고, 변화시키고 싶은 반음의 수 n 에 따라 주파수 변화비율 $R_{\text{final}} = (1 \pm 0.06^n)$ 을 결정하고, $R_{\text{final}} = R_f \times R_t$ 관계로부터 IMCDT 전처리 단계의 주파수 변화율 R_f 를 결정하여 $f_{\text{shift}} = f(1+R_f)$ 을 이용하여 주파수를 변화시키는 것을 특징으로 한다.

- [57] 그리고 상기 IMDCT 입력 데이터를 재구성하는 단계에서, 상기 윈도우 선택과정, 주파수 추출과정, 위상 추출과정, 주파수 변환 과정으로부터 얻은 윈도우 크기(N), 주파수($f = k_{\text{in}} + \epsilon$), 위상(ϕ_k), 변환된 주파수 $f_{\text{shift}} = f(1+R_f)$, 을 이용하여 IMDCT 입력 $X'(k)$ 를 구하는 것을 특징으로 한다.

- [58] 그리고 가변속도에 따라 주파수 변환부의 주파수 변화량을 조절하고, 주파수 변화량에 따라 보간 단계의 샘플링 간격을 조절하는 것을 특징으로 한다.

- [59] 그리고 상기 샘플링 간격의 조절은, 원래속도/가변속도, 원 신호의 샘플링 간격(t_s), 보간에 의해 재생성되는 오디오 신호의 샘플링 간격(t'_s) 사이에 $R_t = (\text{원래속도/가변속도}) = t_s/t'_s$ 관계식이 성립하고, IMDCT 전 단계의 주파수 변환부의 주파수 변화량이 R_f 이라면 최종 음정이 가변속도와 $(R_f \times f) \times R_t$ 로 결정되는 것을 특징으로 한다.

발명의 효과

- [60] 이와 같은 본 발명에 따른 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치 및 방법은 다음과 같은 효과를 갖는다.

- [61] 첫째, 주파수 영역에서 압축된 오디오 신호를 IMDCT를 통하여 시간의 영역의 신호로 변환하기 전에 주파수 IMDCT 입력 데이터 $X(k)$ 를 가공하여 음정을 변화시킬 수 있다.

- [62] 둘째, 시간영역의 신호로 변환하기 전에 IMDCT 입력 데이터 $X(k)$ 를 가공하여 음정을 변화시킴으로써 계산량을 줄이게 되어 시스템의 CPU 부담을 줄여줄 수 있을 뿐만 아니라 소비전력을 줄일 수 있다.

- [63] 셋째, 시간영역의 음정변환 과정이 불필요하게 됨에 따라 데이터를 저장하는 메모리를 줄일 수 있어 하드웨어 시스템을 값싸게 구성할 수 있다.

- [64] 넷째, 입력 데이터 $X(k)$ 의 주파수 성분 및 위상과 진폭을 분리하여 방정식 형태로 표시하여 IMDCT 과정에서 주파수 변환이 용이하여 계산량을 줄일 수 있다.

- [65] 다섯째, 시간영역의 신호를 주파수 영역의 신호로 변환하는 과정에서 주파수 영역의 신호에 포함되는 정보를 활용하여 음정을 변화시키는 방법으로 시간영역에서 음정 변환하는 과정을 생략할 수 있다.

- [66] 여섯째, 주파수 추출 윈도우 크기를 변화시킴으로써 다양한 주파수를 세밀하게 추출할 수 있어 음정변환 음질을 높일 수 있다.
- [67] 일곱째, 스펙트럼성분이 큰 주파수를 중심으로 윈도우를 구성함으로써 불필요한 윈도우를 제거할 수 있어 계산량을 줄일 수 있다.
- [68] 여덟째, 서버밴드나 프레임 가장자리에서 윈도우를 중첩시킴으로써 그들 가장자리에 있는 주파수를 추출할 수 있어 음질을 개선할 수 있다.
- [69] 아홉째, IMDCT 앞 단계에서의 주파수 변환비율과 보간부의 샘플링 간격을 연동함으로써 오디오신호의 음정변화, 속도변화, 음정과 속도를 동시에 변화시킬 수 있다.

도면의 간단한 설명

- [70] 도 1은 종래 기술의 주파수영역에서 압축된 오디오 신호의 음정이나 속도를 가변 시키기 위한 구성도
- [71] 도 2는 종래 기술의 시간영역의 오디오 신호를 주파수 영역의 신호로 변환하기 위한 구성도
- [72] 도 3은 종래 기술의 주파수 영역으로 변환된 신호를 시간영역으로 변환하는 과정을 나타낸 구성도
- [73] 도 4는 본 발명에 따른 IMDCT를 이용한 오디오 신호의 음정 가변 장치의 구성도
- [74] 도 5는 본 발명에 따른 IMDCT를 이용한 오디오 신호의 음정 및 속도 가변 장치의 구성도
- [75] 도 6은 IMDCT 입력 데이터 $X(k)$ 의 한 프레임 내에서 주파수 대역에 따른 윈도우 구성 개념도
- [76] 도 7은 본 발명에 따른 윈도우에서 주파수 성분, 위상, 진폭 추출과정을 나타내는 플로우 차트
- [77] 도 8 내지 도 10은 보간(interpolation) 주파수에 따른 원신호의 음정변환 개념도

발명의 실시를 위한 최선의 형태

- [78] 이하, 본 발명에 따른 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치 및 방법의 바람직한 실시 예에 관하여 상세히 설명하면 다음과 같다.
- [79] 본 발명에 따른 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치 및 방법의 특징 및 이점들은 이하에서의 각 실시 예에 대한 상세한 설명을 통해 명백해질 것이다.
- [80] 도 4는 본 발명에 따른 IMDCT를 이용한 오디오 신호의 음정 가변 장치의 구성도이고, 도 5는 본 발명에 따른 IMDCT를 이용한 오디오 신호의 음정 및 속도 가변 장치의 구성도이다.
- [81] 본 발명은 IMDCT 단계에서 사용될 입력 신호를 가공하는 전처리 단계를 두어 음정과 속도를 가변시키는 것으로, 본 발명에 따른 IMDCT를 이용한 오디오 신호의 음정 가변 장치는 도 4에서와 같이, 처리할 샘플의 윈도우 크기(처리할

샘플의 개수)를 결정하는 윈도우부(41)와, 상기 윈도우부(41)에서 결정된 윈도우 크기로 IMDCT 입력 데이터 $X(k)$ 의 주파수와 위상을 추출하는 주파수 추출부(42) 및 위상 추출부(43), 진폭을 추출하는 진폭 추출부(44)와, 상기 추출된 주파수를 변환하는 주파수 변환부(45)와, 상기 변환된 주파수, 추출된 위상과 진폭을 이용하여 IMDCT 입력 데이터를 재구성하는 입력 데이터 재구성부(46)와, IMDCT를 통하여 재구성된 입력 데이터를 시간 영역으로 변환하는 IMDCT 블록(47)와, IMDCT를 통하여 시간영역으로 변환된 여러 서브밴드 성분을 합성하는 합성 필터뱅크(48)를 포함한다.

- [82] 그리고 본 발명에 따른 IMDCT를 이용한 오디오 신호의 음정 및 속도 가변 장치는 도 5에서와 같이, 처리할 샘플의 윈도우 크기(처리할 샘플의 개수)를 결정하는 윈도우부(41)와, 상기 윈도우부(41)에서 결정된 윈도우 크기로 IMDCT 입력 데이터 $X(k)$ 의 주파수와 위상을 추출하는 주파수 추출부(42) 및 위상 추출부(43), 진폭을 추출하는 진폭 추출부(44)와, 상기 추출된 주파수를 변환하는 주파수 변환부(45)와, 상기 변환된 주파수, 추출된 위상과 진폭을 이용하여 IMDCT 입력 데이터를 재구성하는 입력 데이터 재구성부(46)와, IMDCT를 통하여 재구성된 입력 데이터를 시간 영역으로 변환하는 IMDCT 블록(47)와, IMDCT를 통하여 시간영역으로 변환된 여러 서브밴드 성분을 합성하는 합성 필터뱅크(48)와, 오디오 신호의 샘플링 간격을 조절하여 재생속도를 변화시키는 보간부(Interpolator)(49)를 포함한다.
- [83] 이와 같은 본 발명에 따른 IMDCT 입력신호를 이용한 오디오 신호의 음정 가변 방법은 IMDCT 단계에서 사용될 입력 신호를 가공하는 전처리 단계를 두어 음정을 가변시키기 위하여, 주파수 성분을 세밀하게 추출하기 위하여 분석 윈도우의 크기를 조절하는 단계와, IMDCT 입력 데이터 $X(k)$ 의 주파수를 추출하는 단계와 위상을 추출하는 단계, 진폭을 추출하는 단계와, 음정변화 비율에 따라 추출된 주파수를 변환하는 단계와, 추출된 주파수, 위상, 진폭을 이용하여 IMDCT 입력을 재생성하는 단계와, 주파수 영역의 데이터를 시간영역으로 변환하는 IMDCT 단계와, IMDCT 과정에서 만들어진 다양한 주파수의 시간영역 신호를 합성하는 합성필터뱅크 단계를 포함한다.
- [84] 그리고 본 발명에 따른 IMDCT를 이용한 오디오 신호의 음정 및 속도 가변 방법은 IMDCT 단계에서 사용될 입력 신호를 가공하는 전처리 단계를 두어 음정과 속도를 가변시키기 위하여, 주파수 성분을 세밀하게 추출하기 위하여 분석 윈도우의 크기(처리할 샘플의 개수)를 조절하는 단계와, IMDCT 입력 데이터 $X(k)$ 의 주파수를 추출하는 단계와 위상을 추출하는 단계, 진폭을 추출하는 단계와, 음정변화 비율에 따라 추출된 주파수를 변환하는 단계와, 추출된 주파수, 위상, 진폭을 이용하여 IMDCT 입력을 재구성하는 단계와, 주파수 영역의 데이터를 시간영역으로 변환하는 IMDCT 단계와, IMDCT 과정에서 만들어진 다양한 주파수의 시간영역 신호를 합성하는 합성필터뱅크 단계와, 재생속도 변화비율에 따라 샘플링 간격을 조절하여 재생속도를

변화시키는 단계를 포함한다.

- [85] 여기서, IMDCT 과정에서 원하는 만큼 주파수를 변화시키는데 있어, 계산량을 줄이고 주파수 변화량에 비례하게 IMDCT 전 단계에서 입력 데이터 $X(k)$ 의 주파수, 위상, 진폭을 분리하여 방정식 형태로 표시하거나 룩업(look-up) 테이블 형태로 저장하는 과정을 포함한다.
- [86] 그리고 IMDCT를 이용하여 압축된 데이터를 음정 변환 없이 단순히 복호화하는 경우에는 본 발명에서 제안하는 윈도우 과정, 주파수 추출, 진폭 추출, 입력 데이터 재구성, 주파수 변환 과정을 거치지 않고 부호화된 주파수 영역의 데이터 $X(k)$ 를 IMDCT 과정과 합성필터뱅크를 거쳐 복호화하면 된다.
- [87] 윈도우부(41)의 윈도우잉(windowing) 과정은 다음과 같다.
- [88] 입력신호의 음정을 변화시키기 위해서는 주파수 영역의 IMDCT 입력 데이터 $X(k)$ 의 주파수 성분을 가능한 세밀하게 추출하여 주파수 성분의 손실을 줄여야 한다. 추출의 정확도는 몇(N) 개의 $X(k)$ 샘플을 모아서 추출할 것인가를 결정하는 윈도우의 크기(N)에 의하여 결정된다.
- [89] 주파수 성분을 추출하는 방식은 N개의 샘플 윈도우 내에서 가장 큰 주파수 성분을 추출하는 방식이다. 추출하는 주파수가 중복되지 않은 범위 내에서 가능한 윈도우를 크게 잡는 것이 계산량을 줄이는 측면에서 좋다. 윈도우 크기를 작게 하면 주파수 성분을 세밀하게 추출할 수 있으나 계산량이 많아지는 문제점이 있다.
- [90] 오디오 신호의 주파수 대역은 20 KHz로 알려져 있으나, 오디오 신호의 주파수 특성을 분석해보면 저주파 영역의 주파수 성분이 고주파 영역의 주파수 성분보다 많다. 계산량도 줄이고 세밀하게 주파수 성분을 추출하기 위하여 도 6에서와 같이 IMDCT 입력 데이터 $X(k)$ 에서 주파수 추출할 때 저주파 영역의 윈도우 크기를 작게 하고 고주파 영역의 윈도우를 크게 한다.
- [91] 상기 IMDCT 단계에서 사용되는 입력 데이터 $X(k)$ 는 수학식 1과 같이 복잡한 형태이지만, 본 발명은
$$X(k) = \frac{A_k \sqrt{2N}}{2\pi} \frac{\sin(\pi f)}{(f-k-1)(f-k)} \cos\left(\frac{\pi}{2} k - \phi_k\right)$$
 을 사용하며, 여기서, f 와 ϕ_k 는 IMDCT 입력 데이터로부터 추출한 특정 주파수와 위상을 의미하며, k 는 MDCT에서 주파수 인덱스, A_k 는 주파수 인덱스의 진폭(amplitude)이다.
- [92] 그리고 상기 IMDCT 처리 과정에서 원하는 만큼 주파수를 변화시키기 위해서 IMDCT 입력 데이터 $X(k)$ 의 주파수 성분과 그 주파수의 진폭과 위상을 분리하여 간단한 방정식 형태로 표시하는 과정을 포함한다.
- [93] 그리고 상기 주파수 추출단계에서, IMDCT 입력 데이터의 주파수 성분을 $f = k_{in} + \epsilon$ 으로 하여 주파수 f 를 정수부(k_{in})와 소수부(ϵ)로 나누고, 윈도우 내에서 이웃하는 세 개의 주파수 성분 $X(k-1)$, $X(k)$, $X(k+1)$ 을 이용하여 스펙트럼 값을 수학식 4를 이용하여 구하여 윈도우 내에서 가장 큰 스펙트럼 값을 가지는 $X(k)$ 의 k 를 정수 주파수 성분 k_{in} 으로 한다.
- [94] 수학식 4

$$S_k = \sqrt{X^2(k) + \{X(k+1) - X(k-1)\}^2}$$

- [95] 보다 정확한 주파수 성분의 값을 알기 위해 소수부 주파수 성분을 정수부 주파수(k_{in}) 성분의 절대값과 스펙트럼 값의 비를 문턱 값(λ_0)과 비교하여 $X(k_{in}\pm 1)$ 이나 $X(k_{in}\pm 2)$ 를 사용하여 소수부를 계산한다.
- [96] 그리고 상기에서 추출한 주파수의 위상을 정수부 주파수(k_{in})와 정수부 주파수 성분 $X(k_{in})$ 과 이웃하는 주파수 성분 $X(k_{in}-1)$ 을 이용하여 구하고, 상기 단계에서 추출한 주파수($f = k_{in} + \epsilon$), $X(k_{in})$, $X(k_{in}-1)$, 소수부 주파수 (ϵ), 윈도우 크기(N)을 이용하여 추출한 주파수의 진폭을 구한다.
- [97] 그리고 상기와 같이 추출한 IMDCT 입력 데이터를 구성하는 각 주파수와 그 주파수의 위상과 진폭을 사용하여 음정을 가변하기 위해서 음정변화에 대응되게 주파수 f 를 변화시켜 변환주파수(f_{shift})를 $f_{shift} = f(1+R_f)$ 으로 표시하고, 여기서 R_f 은 주파수 변환비율이며 양의 값은 음정을 높이고 음의 값은 음정을 낮추는 경우이다.
- [98] 본 발명은 IMDCT 앞 단계에서의 주파수 변환을 통한 음정 변화와 보간부의 샘플링 간격을 연동시켜 속도가변, 음정가변, 음정과 재생속도 동시가변을 수행할 수 있다.
- [99] 그리고 윈도우를 구성함에 있어 서브밴드와 서브밴드, 프레임과 프레임 경계에 존재하는 정수주파수 성분을 추출하기 위해서 서브밴드와 프레임을 중첩시켜 윈도우를 설정한다.
- [100] 서브밴드나 프레임에서 수학식 4를 이용하여 스펙트럼을 계산하여 정수주파수 성분을 찾아 그 값이 큰 몇 개의 정수주파수(k_{in})를 중심으로 일정수의 샘플로 구성된 분석 윈도우를 구성한다.
- [101] 그리고 하나의 프레임이나 서브밴드에서 몇 개의 정수주파수 성분을 선택할 것인가 하는 문제는 프레임이나 서브밴드를 구성하는 샘플의 수에 따라 다르다.
- [102] MP3 방식과 같은 경우에는 하나의 서브밴드 내에서 5개 이내의 정수주파수 k_{in} 로 윈도우를 결정한다.
- [103] 시간영역의 오디오 신호를 MDCT를 통하여 주파수 영역의 신호로 변환하면 IMDCT 입력으로 사용될 $X(k)$ 는 수학식 1과 같이 윈도우 크기(N) 만큼의 항들의 합으로 표현된다. MP3나 AAC와 같은 방식과 같이 윈도우가 클 경우 수학식 5와 같은 단일 주파수(f)의 정현파를 MDCT 하면 수학식 6과 같이 $X(k)$ 는 하나의 항으로 근사화 할 수 있다.
- [104] 수학식 6를 분석해보면 IMDCT 입력은 MDCT 주파수 인덱스(k)와 그 주파수 인덱스의 위상(ϕ_k)과 진폭(A_k), MDCT의 입력으로 사용된 신호의 단일 주파수(f)로 표현됨을 알 수 있다. 시간영역의 복잡한 모양의 오디오 신호도 결국 여러 주파수의 정현파 합으로 생각할 수 있으므로 IMDCT 입력으로 사용될 $X(k)$ 의 각 주파수 성분 즉

인덱스(k)에 대한 정보(진폭, 주파수, 위상)를 추출하여 추출된 주파수를 변화시켜 IMDCT 단계를 거치면 음정을 변화시킬 수 있다.

[105] 수학식 5

$$x(n) = A \sin\left(\frac{\pi f}{N} n + \phi\right)$$

[106] 수학식 6

$$X(k) \cong \frac{A_k \sqrt{2N}}{2\pi} \frac{\sin(\pi f)}{(f - k - 1)(f - k)} \cos\left(\frac{\pi}{2} k - \phi_k\right)$$

[107] 그리고 주파수 추출부(42) 및 위상 추출부(43), 진폭 추출부(44)에서의 주파수, 위상, 진폭 추출과정은 도 7에서와 같다.

[108] IMDCT 입력 데이터 X(k)의 주파수를 정수부(k_{in})와 소수부(ϵ)로 나누어 $f = k_{in} + \epsilon$ 으로 표시할 수 있다. 이웃하는 세 개의 주파수 성분 $X(k_{in}-1)$, $X(k_{in})$, $X(k_{in}+1)$ 을 이용하여 분석할 윈도우 내 존재하는 모든 스펙트럼 값(61)을 구하여 그 중 가장 큰 스펙트럼 값(S_k)을 만드는 주파수 인덱스 k를 정수부 주파수 성분 k_{in} 으로 한다.(62)

[109] 소수부 주파수 성분을 정수부 주파수 성분 $X(k_{in})$ 의 바로 이웃하는 성분 $X(k_{in} \pm 1)$ 을 이용하여 구할 것인가, 아니면 그 다음 성분 $X(k_{in} \pm 2)$ 를 이용하여 구할 것인가를 정하기 위하여 정수부 주파수 인덱스 k_{in} 과 k_{in} 을 중심으로 한 스펙트럼 값의 비

$$\lambda = \frac{|k_{in}|}{S_k}$$

를 구한다.(63)

[110] 정수부 주파수 성분과 바로 인접한 성분 $X(k_{in} \pm 1)$ 을 이용하여 수학식 7을 이용하여 소수부 주파수 성분(ϵ_1)을 계산한다.(64, 65) 정수부 주파수 성분 $X(k_{in})$ 에서 2만큼 떨어진 주파수 성분 $X(k_{in} \pm 2)$ 를 이용하여 수학식 8을 이용하여 또 다른 소수부 주파수 성분(ϵ_2)를 구한다. (66, 67)

[111] 수학식 7

$$\alpha = -\frac{X(k_{in} - 1)}{X(k_{in} + 1)} \quad \text{라고 두면,}$$

$$\alpha \neq 1 \text{인 경우에 대해서} \quad \epsilon_1 = \frac{3 + \alpha - \sqrt{\alpha^2 + 14\alpha + 1}}{2(1 - \alpha)}$$

[112] 수학식 8

$$\beta = \frac{X(k_{in} - 2)}{X(k_{in} + 2)} \text{라고 두면,}$$

$$\beta \neq 1 \text{인 경우에 대해서} \quad \varepsilon_2 = \frac{5 + 3\beta - \sqrt{\beta^2 + 62\beta + 1}}{2(1 - \beta)}$$

[113] 정수부 주파수 k_{in} 과 k_{in} 을 중심으로 한 스펙트럼 값의 비

$$\lambda = \frac{|k_{in}|}{S_k}$$

가 특정 문턱 값보다 작으면 수학식 7을 이용하여 구한 소수부 주파수 성분(ε_1)을 선택하고, 클 경우는 수학식 8을 이용하여 구한 소수부(ε_2)를 선택한다. (68, 69)

[114] IMDCT 입력 데이터 $X(k)$ 의 추출된 정수부 주파수(k_{in})와 소수부 주파수(ε)을 이용하여 주파수(f)를 $f = k_{in} + \varepsilon$ 와 같이 구한다. (71)

[115] 정수부 IMDCT 입력 데이터 $X(k)$ 의 cosine 함수의 위상(ϕ_k)

은 정수부 주파수 성분 $X(k_{in})$ 과 바로 인접한 성분 $X(k_{in} \pm 1)$ 과 소수부 주파수 성분(ε)을 이용하여 수학식 9를 이용하여 구한다. (70)

[116] 수학식 9

$$\phi_k = \frac{\pi}{2} k_{in} - \tan^{-1} \left\{ \frac{(\varepsilon + 1)X(k_{in} - 1)}{(\varepsilon - 1)X(k_{in})} \right\}$$

[117] 그리고 본 발명에 따른 진폭 추출부에서의 진폭 추출 과정은 다음과 같다. (72)
수학식 6의 $X(k)$ 에 k_{in} 과 $k_{in}-1$ 을 각각 대입하여 수학식 10, 수학식 11을 얻는다.
수학식 10과 수학식 11을 이용하여 간단한 조작을 통하여 수학식 12와 같은 진폭(A_k)를 얻을 수 있다. 수학식 10, 11에서 f 는 추출된 주파수 정보 ($k_{in} + \varepsilon$)를 사용한다.

[118] 수학식 10

$$(\varepsilon - 1)X(k_{in}) = \frac{-A_k \sqrt{2N}}{2\pi} \frac{\sin(\pi f)}{\varepsilon} \cos\left(\frac{\pi}{2} k_{in} - \phi_k\right)$$

[119] 수학식 11

$$(\varepsilon + 1)X(k_{in} - 1) = \frac{-A_k \sqrt{2N}}{2\pi} \frac{\sin(\pi f)}{\varepsilon} \cos\left(\frac{\pi}{2} k_{in} - \phi_k\right)$$

[120] 수학식 12

$$A_k = \frac{\sqrt{2\pi\epsilon}}{\sin(\pi f)} \sqrt{\frac{(\epsilon+1)^2 X(k_{in}-1)^2 + (\epsilon-1)^2 X(k_{in})^2}{N}}$$

- [121] 이상의 과정은 하나의 윈도우에서 그 윈도우를 대표하는 주파수, 위상, 진폭을 구하였다. MDCT 방식을 이용하는 MP3, AAC, AC-3 방식에 따라 하나의 프레임에는 위와 같은 분석이 필요한 윈도우가 여러 개 있다. 본 발명의 개념을 활용하여 방식에 따라 분석 윈도우의 크기를 적절하게 조절하여 하나의 프레임을 구성하는 다양한 주파수(f), 위상(ϕ_k), 진폭(A_k)을 분석할 수 있다.

- [122] 하나의 프레임에 대한 주파수 성분의 추출이 완료되면, 재생속도 변화 없이 음정을 변화시키는 경우 즉 원래 오디오 신호의 샘플링 간격과 보간부의 샘플링 간격이 동일한 경우에는 추출된 주파수를 $f_{\text{shift}} = f(1+R_f)$ 와 같이 변화시킨다. 여기서 R_f 는 주파수 변화율이다. 일반적으로 반음을 올리거나 내릴 경우 원래 주파수의 6% 만큼 주파수의 변화가 있으므로 n-반음정을 올리거나 내리기 위해서는 $f_{\text{shift}} = f\{1 \pm (0.06)^n\}$ 와 같이 주파수를 변화시키면 된다.

- [123] 음정과 재생속도를 동시에 가변 시키는 경우는 재생속도로부터 (원래속도/가변속도) = R_t 로 부터 R_t 를 구하고, 변화시키고 싶은 반음의 수 n에 따라 주파수 변화비율 $R_{\text{final}} = (1 \pm 0.06^n)$ 을 결정하고, $R_{\text{final}} = R_f \times R_t$ 로 부터 주파수 변화율 $R_f = (1 \pm 0.06^n) / R_t$ 결정하여 최종적으로 $f_{\text{shift}} = f(1+R_f)$ 을 구한다.

- [124] 그리고 입력 데이터 재구성부(46)의 입력 데이터 재구성(regenerating IMDCT input data) 과정은 다음과 같다.

- [125] 상기 과정을 거쳐 추출한 주파수 성분($f = k_{in} + \epsilon$), 음정변화에 따른 주파수 변화율(R_f), 위상(ϕ_k), 정수 주파수 성분(k_{in}), 진폭 정보(A_k)를 이용하여 수학식 13과 같이 IMDCT 입력 데이터 $X'(k)$ 를 다시 생성한다. 수학식 13에서 k는 IMDCT 영역의 주파수 성분이다.

- [126] 수학식 13
- $$X(k) \cong \frac{A_k \sqrt{2N}}{2\pi} \frac{\sin\{\pi f(1+R_f)\}}{\{f(1+R_f) - k - 1\}\{f(1+R_f) - k\}} \cos\left\{\frac{\pi}{2}(k - k_{in}) - \psi\right\}$$

- [127] 여기서

$$\psi = \frac{\pi}{2} k_{in} + \frac{\pi(N-1)}{2N} f(1+R_f) + \phi_k - \frac{\pi}{4}$$

- [128] 상기 수학식 13을 이용하여 여러 윈도우로 구성된 프레임을 각 각의 윈도우에 대하여 IMDCT 과정을 거치면 그 프레임에 대하여 음정이 변화된 시간영역의 오디오신호를 얻을 수 있다. 이러한 개념을 사용하여 시간영역의 오디오신호를 필요에 따라 적절하게 보간을 하면 음정가변, 속도 가변, 음정 및 속도 동시가변

효과를 얻을 수 있다.

- [129] 보간 단계에서 속도 및 음정변화 변화 개념을 도 8 내지 도 10을 이용하여 상세히 설명하면 다음과 같다.
- [130] 도 8을 음정 변화가 안된 원래의 신호라고 가정한다.
- [131] 도 9는 상기 IMDCT 과정을 통하여 주파수변환 즉 음정변환이 된 신호이다.
- [132] 도 8과 도 9에서 T_0/T_{sh} 를 주파수 변환비율 R_f 라고 한다. R_f 가 1보다 크게 되면 원래 오디오 신호보다 주파수가 올라가 음정이 높아진다.
- [133] 도 9와 도 10에서 신호를 샘플링 하는 간격 t_s 와, t'_s 의 비 t_s/t'_s 를 샘플링 간격의 비 R_t 라고 정의한다. 샘플링 간격의 비 R_t 는 재생속도의 비 (원래속도/가변속도)와 같이 생각해도 된다.
- [134] 여기서 재생속도가 느린 경우는 (원래속도/가변속도)가 1보다 큰 경우이다.
- [135] 샘플링 간격을 짧게 하면 주어진 시간에 많은 샘플 데이터를 얻을 수 있어, 음정을 변화시키지 않은 신호에 대하여 샘플링 간격의 비 R_t 를 1보다 크게하면 느리게 재생되고 음정이 낮아진다. 이러한 개념을 사용하면 샘플링 간격의 비를 조절함으로써 재생속도와 음정을 바꿀 수 있다.
- [136] 원래의 주파수(f)가 상기 IMDCT 전처리과정과 IMDCT를 통하여 R_f 비율로 주파수 변환이 일어났다면 보간단계 전의 주파수는 $(R_f \times f)$ 가 된다. 변환된 신호의 샘플링 간격과 보간 단계의 샘플링 간격의 비가 R_t 이라면 보간 단계를 거친 최종 신호의 주파수는 $(R_f \times f) \times R_t$ 가 된다.
- [137] 상기 IMDCT를 이용한 오디오 신호의 음정 및 속도 가변장치에서 음정변화 없이 재생속도를 변화시키려면 (원래속도/가변속도) = $t_s/t'_s = R_t$ 로 부터 R_t 를 구한 후 $(R_f \times R_t) = 1$ 되게 IMDCT 전처리 단계의 주파수 변화율 R_f 를 정하면 된다. 그리고 원신호의 샘플링 간격 t_s 는 이미 알고 있는 값이므로 음정변화 없이 재생속도를 변화시킬 수 있는 t'_s 를 구할 수 있다.
- [138] 상기 IMDCT를 이용한 오디오 신호의 음정 및 속도 가변장치에서 음정을 변화시키고자 하는 경우에는 음정변화비율 R_f 를 IMDCT 전처리 단계의 주파수 변화율 R_f 로 정하고 샘플링 간격 비율은 변화시키지 않고 재생하면 된다.
- [139] R_f 가 1보다 크면 음정이 높아지고 1보다 작으면 음정이 낮아진다. 일반적으로 반음은 주파수 측면에서 $\pm 6\%$ 변화를 가져오므로 변화시키려는 반음의 수(n)에 따라 주파수 변환율을 $R_f = (1 \pm 0.06^n)$ 형태로 결정하여 수학식 13을 이용하여 IMDCT 입력 $X'(k)$ 을 재생성한다.
- [140] 상기 IMDCT를 이용한 오디오 신호의 음정 및 속도 가변장치에서 음정과 재생속도를 동시에 변화시키려는 경우에는 재생속도로부터 (원래속도/가변속도) = R_t 로부터 R_t 를 구하고, 변화시키고 싶은 반음의 수 n 에 따라 주파수 변화비율 $R_{final} = (1 \pm 0.06^n)$ 을 결정하고 $R_{final} = R_f \times R_t$ 로부터 IMDCT 전처리 단계의 주파수 변화율 R_f 를 결정하여 수학식 13을 이용하여 IMDCT 입력 $X'(k)$ 을 재생성하고 재생성할 신호의 샘플링 간격(t'_s)은 (원래속도/가변속도) = $t_s/t'_s = R_t$ 을 이용하여 구한다.

- [141] 이와 같은 본 발명에 따른 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치 및 방법은 IMDCT(Inverse Modified Discrete Cosine Transform)를 통하여 시간의 영역의 신호로 변환하기 전에 입력되는 주파수 데이터 $X(k)$ 를 가공하여 음정을 변화시킬 수 있도록 한 것으로, IMDCT 과정에서 원하는 만큼 주파수를 용이하게 변화시키기 위해서는 입력 데이터 $X(k)$ 의 주파수 성분과 진폭을 분리하여 방정식 형태로 표시하는 과정을 포함하고, IMDCT 과정을 거쳐 시간영역으로 변환된 오디오 신호의 샘플링 간격을 조절하여 음정, 재생속도, 음정 및 재생속도 동시 가변이 가능하게 하는 보간 과정을 포함하여, 계산량 및 메모리의 사용을 줄일 수 있도록 한 것이다.
- [142] 상기 개념을 이용하면 음정변화뿐만 아니라 임의의 주파수와 임의의 재생속도를 얻을 수 있다.
- [143] 이상에서의 설명에서와 같이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 본 발명이 구현되어 있음을 이해할 수 있을 것이다.
- [144] 그러므로 명시된 실시 예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 하고, 본 발명의 범위는 전술한 설명이 아니라 특허청구 범위에 나타나 있으며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

산업상 이용가능성

- [145] 본 발명에 따른 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치 및 방법은 IMDCT(Inverse Modified Discrete Cosine Transform)를 통하여 시간의 영역의 신호로 변환하기 전에 입력되는 주파수 데이터 $X(k)$ 를 가공하여 음정을 변화시킬 수 있도록 한 것으로, IMDCT 과정을 거쳐 시간영역으로 변환된 오디오 신호의 샘플링 간격을 조절하여 음정, 재생속도, 음정 및 재생속도 동시 가변이 가능하다.

청구범위

[청구항 1]

처리할 샘플의 개수를 결정하는 윈도우부;
 상기 윈도우부에서 결정된 윈도우 크기로 IMDCT 입력 데이터 $X(k)$ 의 주파수와 위상을 추출하는 주파수 추출부 및 위상 추출부,
 진폭을 추출하는 진폭 추출부;
 상기 추출된 주파수를 변환하는 주파수 변환부;
 상기 변환된 주파수, 추출된 위상과 진폭을 이용하여 IMDCT 입력 데이터를 재구성하는 입력 데이터 재구성부;
 IMDCT(Inverse Modified Discrete Cosine Transform)를 통하여 상기 재구성된 입력 데이터를 시간 영역으로 변환하는 IMDCT 블록;
 상기 IMDCT를 통하여 시간영역으로 변환된 여러 서브밴드 성분을 합성하는 합성 필터뱅크;를 포함하는 것을 특징으로 하는 IMDCT 입력신호를 이용한 오디오 신호의 음성 가변 장치.

[청구항 2]

제 1 항에 있어서, 상기 윈도우부에서 IMDCT 입력 데이터인 $X(k)$ 를 분석하고 재구성하는 데 필요한 윈도우를 결정하기 위하여,
 분석 대상이 되는 주파수 영역의 윈도우 크기를 다른 영역에 비하여 상대적으로 작게 하는 것을 특징으로 하는 IMDCT 입력신호를 이용한 오디오 신호의 음성 가변 장치.

[청구항 3]

제 1 항에 있어서, 상기 윈도우부에서 IMDCT 입력 데이터인 $X(k)$ 를 분석하고 재구성 하는데 필요한 윈도우를 결정하기 위하여,
 서브밴드와 서브밴드, 프레임과 프레임 경계에서 분석윈도우를 중첩시켜 설정하고,
 분석윈도우 내에서 스펙트럼을 계산을 통하여 정수주파수(k_{in})를 찾아 그 주파수를 중심으로 분석 윈도우를 구성하는 것을 특징으로 하는 IMDCT 입력신호를 이용한 오디오 신호의 음성 가변 장치.

[청구항 4]

IMDCT를 통하여 시간의 영역의 신호로 변환하기 전에 IMDCT에 입력되는 데이터($X(k)$)를 처리하여 음정을 변화시키기 위하여,
 주파수 추출을 위한 샘플의 수를 결정하는 윈도우 크기 결정 단계;
 선택된 윈도우 크기로 IMDCT 과정에 필요한 입력 데이터 $X(k)$ 의 주파수(k), 위상, 진폭을 추출하는 단계;
 추출된 주파수를 변환하고, 변환된 주파수와 추출된 위상과 진폭을 이용하여 IMDCT 입력 데이터를 재구성하는 단계;
 주파수 영역의 데이터를 시간영역으로 변환하는 IMDCT 단계;
 IMDCT 과정에서 만들어진 다양한 주파수의 시간영역 신호를

합성하는 합성필터뱅크 단계;를 포함하는 것을 특징으로 하는
IMDCT 입력신호를 이용한 오디오 신호의 음정 가변 방법.

[청구항 5]

제 4 항에 있어서, 상기 IMDCT 입력 데이터 $X(k)$ 의 주파수 추출
단계에서,

IMDCT 입력 데이터 $X(k)$ 의 주파수를 정수부(k_{in})와 소수부(ϵ)로
나누어 $f = k_{in} + \epsilon$ 로 표시하고, 이웃하는 세 개의 주파수 성분 $X(k_{in} - 1)$, $X(k_{in})$, $X(k_{in} + 1)$ 을 이용하여 분석하는 윈도우 내 존재하는 모든
스펙트럼 값을 구하여 그 중 가장 큰 스펙트럼 값(S_k)을 만드는 k_{in}
를 정수부 주파수 성분(k_{in})으로 하는 것을 특징으로 하는 IMDCT
입력신호를 이용한 오디오 신호의 음정 가변 방법.

[청구항 6]

제 5 항에 있어서, 상기 주파수 성분의 소수부분 ϵ 을,

$$\alpha = -\frac{X(k_{in} - 1)}{X(k_{in} + 1)} \text{라고 두면,}$$

$$\alpha \neq 1 \text{인 경우에 대해서 } \epsilon_1 = \frac{3 + \alpha - \sqrt{\alpha^2 + 14\alpha + 1}}{2(1 - \alpha)}$$

$$\beta = \frac{X(k_{in} - 2)}{X(k_{in} + 2)} \text{라고 두면,}$$

$$\beta \neq 1 \text{인 경우에 대해서 } \epsilon_2 = \frac{5 + 3\beta - \sqrt{\beta^2 + 62\beta + 1}}{2(1 - \beta)}$$

2 종류를 구하고, α 와 β 중에서 어느 것을 사용한 것인지의 결정은,
윈도우 내의 가장 큰 주파수 성분 $X(k_{in})$ 의 절대값과 k_{in} 의 스펙트럼
값(S_k)의 비율을 $\lambda = \frac{|X(k_{in})|}{S_k}$ 로 정의하고, 그 비율이 특정

문턱값(threshold) λ_0 과 비교하여 작으면 α , 그 외의 경우엔 β 를
선택하여 주파수 성분의 소수부분인(ϵ)을 얻는 특징으로 하는
IMDCT 입력신호를 이용한 오디오 신호의 음정 가변 방법.

[청구항 7]

제 4 항에 있어서, IMDCT 데이터 $X(k)$ 의 위상 ϕ_k 를 추출하기
위해서 추출한 IMDCT 정수부 주파수 성분(k_{in})을 이용하여
계산하는 것을 특징으로 하는 IMDCT 입력신호를 이용한 오디오
신호의 음정 가변 방법.

[청구항 8]

제 4 항에 있어서, 상기 진폭을 추출하는 단계에서,
정수부 주파수 성분(K_{in})과 소수부(ϵ) 주파수 성분을 이용하여
IMDCT 입력 데이터의 진폭 A_k 를 구하는 것을 특징으로 IMDCT
입력신호를 이용한 오디오 신호의 음정 가변 방법.

[청구항 9]

제 4 항에 있어서, 상기 IMDCT 입력 데이터를 재구성하는
단계에서,

상기 윈도우 선택과정, 주파수 추출과정, 위상 추출과정, 주파수
변환 과정으로부터 얻은 윈도우 크기(N), 주파수($f = k_{in} + \epsilon$), 위상(ϕ_k
), 변환된 주파수 $f_{shift} = f(1 + R_f)$, 을 이용하여 IMDCT 입력 $X'(k)$ 를

- 구하는 것을 특징으로 하는 IMDCT 입력신호를 이용한 오디오 신호의 음정 가변 방법.
- [청구항 10] 처리할 샘플의 윈도우 크기를 결정하는 윈도우부;
 상기 윈도우부에서 결정된 윈도우 크기로 IMDCT 입력 데이터 $X(k)$ 의 주파수와 위상을 추출하는 주파수 추출부 및 위상 추출부, 진폭을 추출하는 진폭 추출부;
 상기 추출된 주파수를 변환하는 주파수 변환부;
 상기 변환된 주파수, 추출된 위상과 진폭을 이용하여 IMDCT 입력 데이터를 재구성하는 입력 데이터 재구성부;
 IMDCT를 통하여 재구성된 입력 데이터를 시간 영역으로 변환하는 IMDCT 블록;
 IMDCT를 통하여 시간영역으로 변환된 여러 서브밴드 성분을 합성하는 합성 필터뱅크;
 상기 합성 필터뱅크에서 출력되는 오디오 신호의 샘플링 간격을 조절하여 재생속도를 변화시키는 보간부(Interpolator);를 포함하는 것을 특징으로 하는 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치.
- [청구항 11] 제 10 항에 있어서, 상기 IMDCT 블록에서 오디오 신호의 음정을 변화시키기 위하여 IMDCT 입력 데이터 $X(k)$ 로 부터 정현파 성분으로 분해하여 추출한 후, 원하는 만큼 주파수를 변환하여 IMDCT 입력을 재구성하는 것을 특징으로 하는 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치.
- [청구항 12] 제 10 항에 있어서, 상기 윈도우부에서 IMDCT 입력 데이터인 $X(k)$ 를 분석하고 재구성 하는데 필요한 윈도우를 결정하기 위하여,
 분석 대상이 되는 주파수 영역의 윈도우 크기를 다른 영역에 비하여 상대적으로 작게 하는 것을 특징으로 하는 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치.
- [청구항 13] 제 10 항에 있어서, 상기 윈도우부에서의 IMDCT 입력 데이터인 $X(k)$ 를 분석하고 재구성 하는데 필요한 윈도우를 결정하는 데 있어 서브밴드와 서브밴드, 프레임과 프레임 경계에서 분석윈도우를 중첩시켜 설정하고,
 분석윈도우 내에서 스펙트럼을 계산을 통하여 정수주파수(k_{in})를 찾아 그 주파수를 중심으로 분석 윈도우를 구성하는 것을 특징으로 하는 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치.
- [청구항 14] IMDCT를 통하여 시간의 영역의 신호로 변환하기 전에 IMDCT에 입력되는 데이터($X(k)$)를 처리하여 음정을 변화시키기 위하여,

주파수 추출을 위한 샘플의 수를 결정하는 윈도우 크기 결정 단계;
 선택된 윈도우 크기로 IMDCT 과정에 필요한 입력 데이터 $X(k)$ 의
 주파수(k), 위상, 진폭을 추출하는 단계;
 추출된 주파수를 변환하고, 변환된 주파수와 추출된 위상과
 진폭을 이용하여 IMDCT 입력 데이터를 재구성하는 단계;
 IMDCT 처리 및 보간을 하여 코딩된 오디오 신호를 출력하는
 단계;를 포함하는 것을 특징으로 하는 IMDCT 입력신호를 이용한
 오디오 신호의 음정 및 속도 가변 방법.

[청구항 15]

제 14 항에 있어서, 상기 IMDCT 처리 과정에서 원하는 만큼
 주파수를 변화시키기 위해서 IMDCT 입력 데이터 $X(k)$ 의 주파수,
 위상, 진폭을 분리하여 방정식 형태로 표시하거나 그 방정식을
 look-up 테이블 형태로 저장해두고 사용하는 과정을 포함하는
 것을 특징으로 하는 IMDCT 입력신호를 이용한 오디오 신호의
 음정 및 속도 가변 방법.

[청구항 16]

제 14 항에 있어서, IMDCT 입력 데이터 $X(k)$ 의 주파수 추출
 단계에서,
 IMDCT 입력 데이터 $X(k)$ 의 주파수를 정수부(k_{in})와 소수부(ϵ)로
 나누어 $f = k_{in} + \epsilon$ 로 표시하고, 이웃하는 세 개의 주파수 성분 $X(k_{in}$
 $-1)$, $X(k_{in})$, $X(k_{in}+1)$ 을 이용하여 분석하는 윈도우 내 존재하는 모든
 스펙트럼 값을 구하여 그 중 가장 큰 스펙트럼 값(S_k)을 만드는 k_{in}
 를 정수부 주파수 성분(k_{in})으로 하는 것을 특징으로 하는 IMDCT
 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 방법.

[청구항 17]

제 16 항에 있어서, 상기 주파수 성분의 소수부분 ϵ 을,

$$\alpha = -\frac{X(k_{in}-1)}{X(k_{in}+1)} \text{라고 두면,}$$

$$\alpha \neq 1 \text{인 경우에 대해서 } \epsilon_1 = \frac{3 + \alpha - \sqrt{\alpha^2 + 14\alpha + 1}}{2(1 - \alpha)}$$

$$\beta = \frac{X(k_{in}-2)}{X(k_{in}+2)} \text{라고 두면,}$$

$$\beta \neq 1 \text{인 경우에 대해서 } \epsilon_2 = \frac{5 + 3\beta - \sqrt{\beta^2 + 62\beta + 1}}{2(1 - \beta)}$$

2 종류를 구하고, α 와 β 중에서 어느 것을 사용한 것인지의 결정은,
 윈도우 내의 가장 큰 주파수 성분 $X(k_{in})$ 의 절대값과 k_{in} 의 스펙트럼
 값의 비율을 $\lambda = \frac{|k_{in}|}{S_k}$ 로 정의하고, 그 비율이 특정 문턱값(threshold) λ_0

과 비교하여 작으면 α , 그 외의 경우엔 β 를 선택하여 주파수
 성분의 소수부분인(ϵ)을 얻는 것을 특징으로 하는 IMDCT
 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 방법.

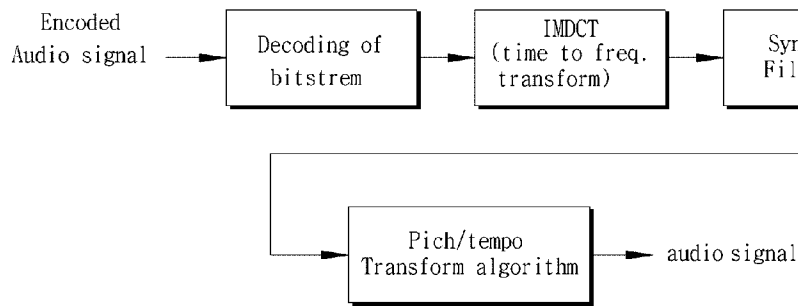
[청구항 18]

제 14 항에 있어서, IMDCT 데이터 $X(k)$ 의 위상 ϕ_k 를 추출하기

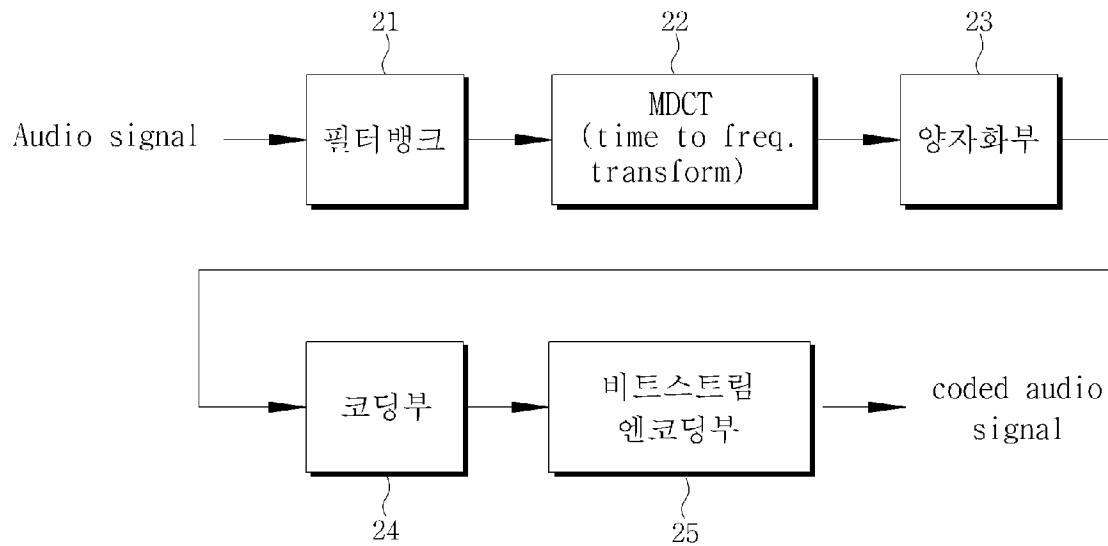
- 위해서 추출한 IMDCT 정수부 주파수 성분(k_{in})을 이용하여 계산하는 것을 특징으로 하는 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 방법.
- [청구항 19] 제 14 항에 있어서, 상기 진폭을 추출하는 단계에서, 정수부 주파수 성분(K_{in})과 소수부(ϵ) 주파수 성분을 이용하여 IMDCT 입력 데이터의 진폭 A_k 를 구하는 것을 특징으로 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 방법.
- [청구항 20] 제 14 항에 있어서, 가변속도와 주파수 변화량에 따라 주파수 변환부와 보간부를 연동하여, 주파수 변화량과 샘플링 간격을 조절하는 것을 특징으로 하는 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 방법.
- [청구항 21] 제 20 항에 있어서, 음정변화 없이 재생속도를 변화시키기 위하여, 상기 IMDCT 처리 및 보간을 하여 코딩된 오디오 신호를 출력하는 단계와 연계하여 원래 속도를 1로 할 때 가변속도, 원신호의 샘플링 간격(t_s), 새롭게 만들 신호의 샘플링 간격(t'_s), $(\text{원래속도/가변속도}) = t_s/t'_s = R_t$ 관계를 이용하여 R_t 를 구한 후 $(R_f \times R_t) = 1$ 되게 R_f 를 결정한 다음, 상기 추출한 IMDCT 입력 데이터 $X(k)$ 의 주파수 성분(k)을 $f_{\text{shift}} = f(1+R_f)$ 변화시키는 것을 특징으로 하는 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 방법.
- [청구항 22] 제 20 항에 있어서, 음정과 재생속도를 동시에 변화시키려는 경우에는 재생속도로부터 $(\text{원래속도/가변속도}) = R_t$ 로부터 R_t 를 구하고, 변화시키고 싶은 반음의 수 n 에 따라 주파수 변화비율 $R_{\text{final}} = (1 \pm 0.06^n)$ 을 결정하고, $R_{\text{final}} = R_f \times R_t$ 관계로부터 IMCDT 전처리 단계의 주파수 변화율 R_f 를 결정하여 $f_{\text{shift}} = f(1+R_f)$ 을 이용하여 주파수를 변화시키는 것을 특징으로 하는 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 장치 및 방법.
- [청구항 23] 제 20 항에 있어서, 상기 IMDCT 입력 데이터를 재구성하는 단계에서, 상기 윈도우 선택과정, 주파수 추출과정, 위상 추출과정, 주파수 변환 과정으로부터 얻은 윈도우 크기(N), 주파수($f = k_{in} + \epsilon$), 위상(ϕ_k), 변환된 주파수 $f_{\text{shift}} = f(1+R_f)$, 을 이용하여 IMDCT 입력 $X'(k)$ 를 구하는 것을 특징으로 하는 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 방법.
- [청구항 24] 제 20 항에 있어서, 상기 샘플링 간격의 조절은, 원래속도/가변속도, 원 신호의 샘플링 간격(t_s), 보간에 의해 재생성되는 오디오 신호의 샘플링 간격(t'_s) 사이에 R_t

$=(\text{원래속도}/\text{가변속도})=t_s/t'_s$ 관계식이 성립하고, IMDCT 전 단계의 주파수 변환부의 주파수 변화량이 R_f 이라면 최종 음정이 가변속도와 $(R_f \times f) \times R_i$ 로 결정되는 것을 특징으로 하는 IMDCT 입력신호를 이용한 오디오 신호의 음정 및 속도 가변 방법.

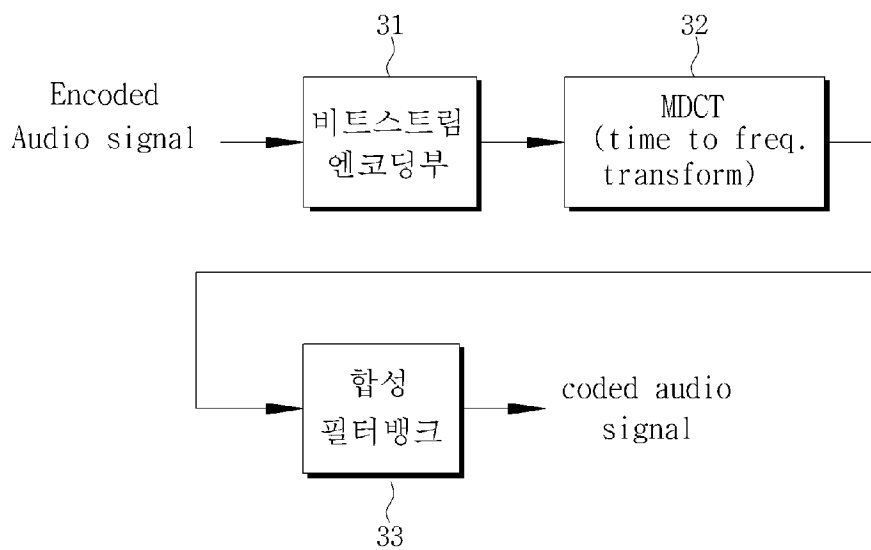
[Fig. 1]



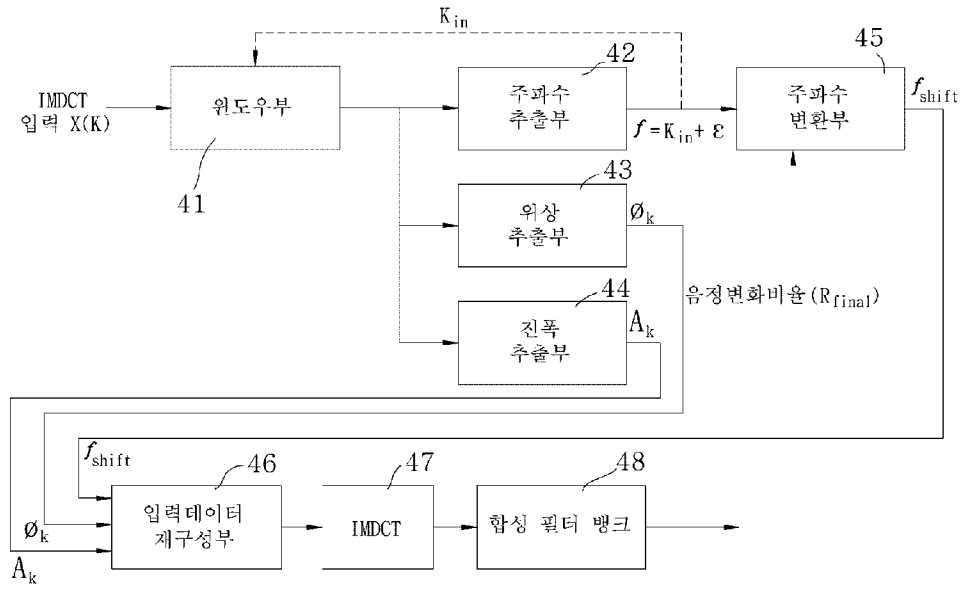
[Fig. 2]



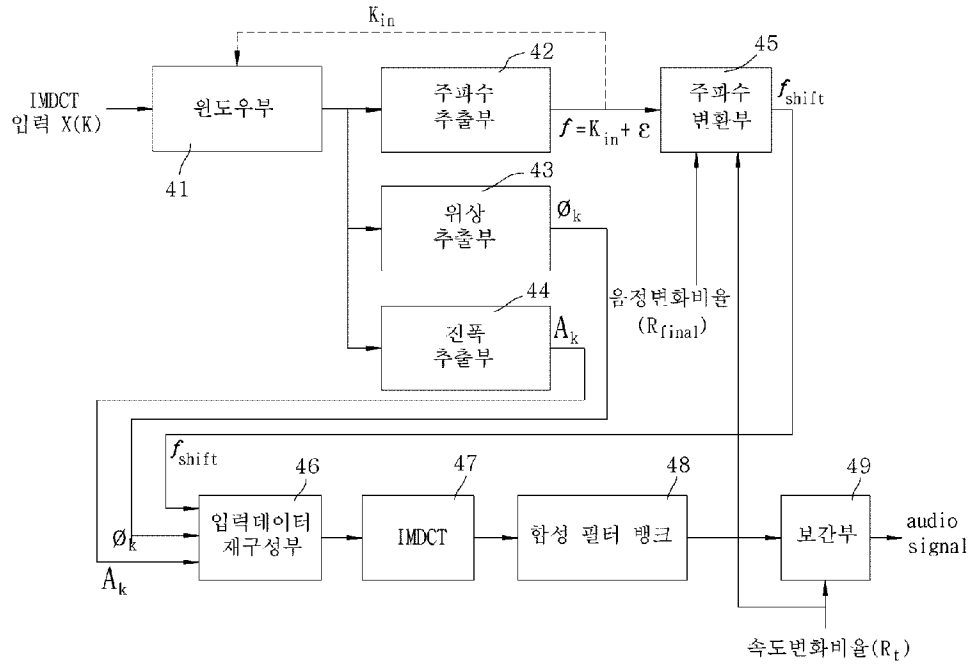
[Fig. 3]



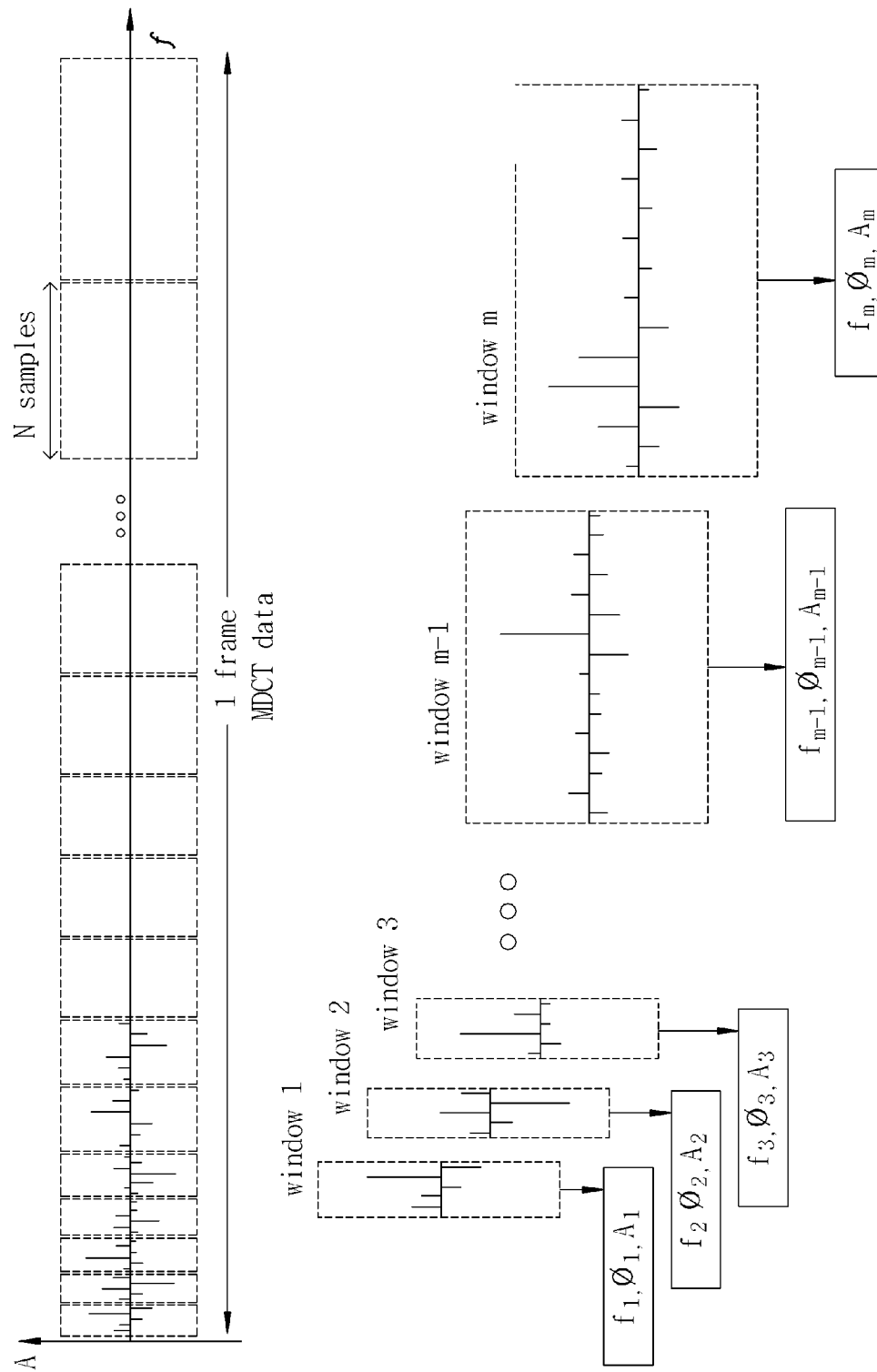
[Fig. 4]



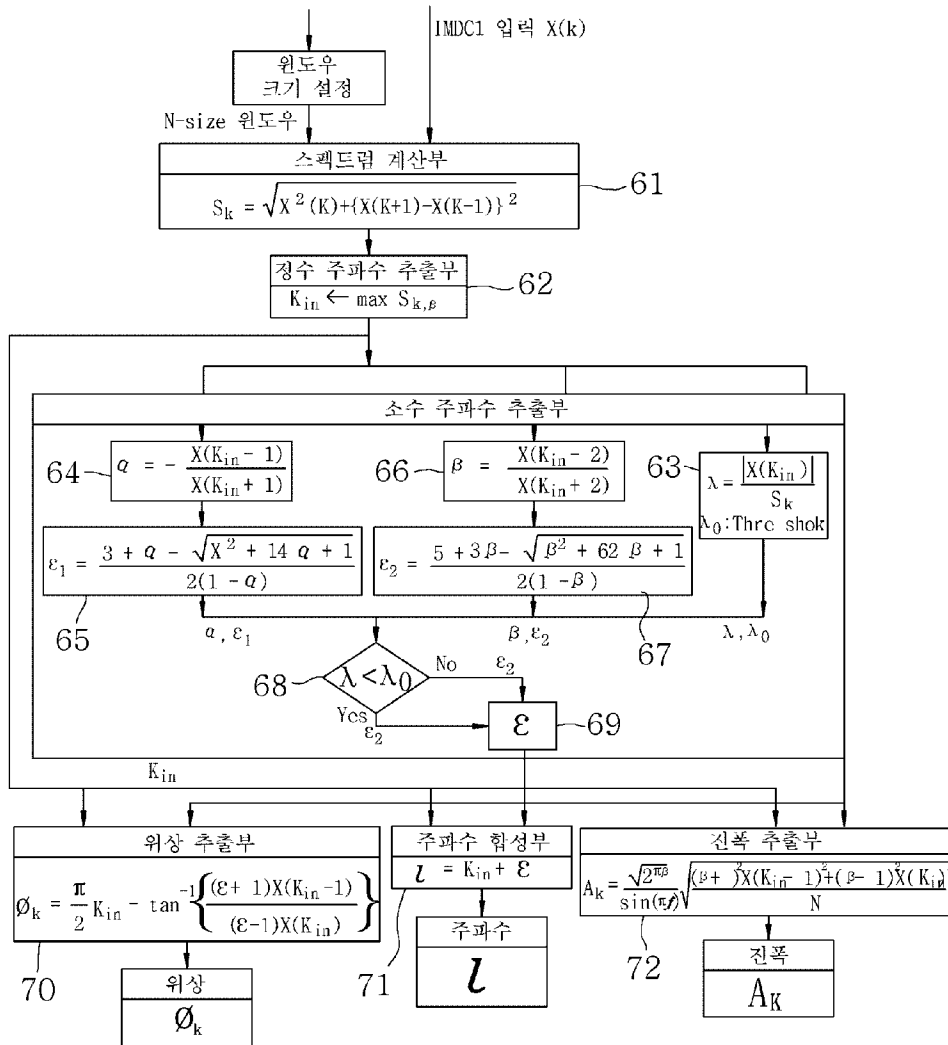
[Fig. 5]



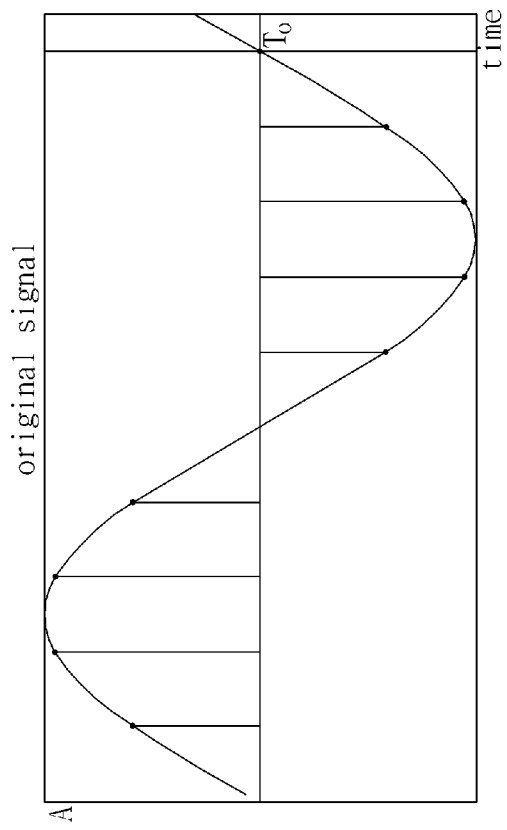
[Fig. 6]



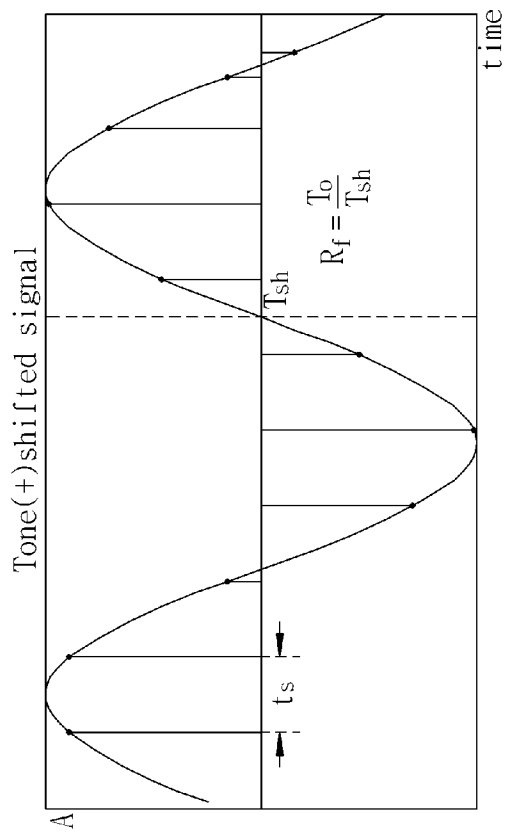
[Fig. 7]



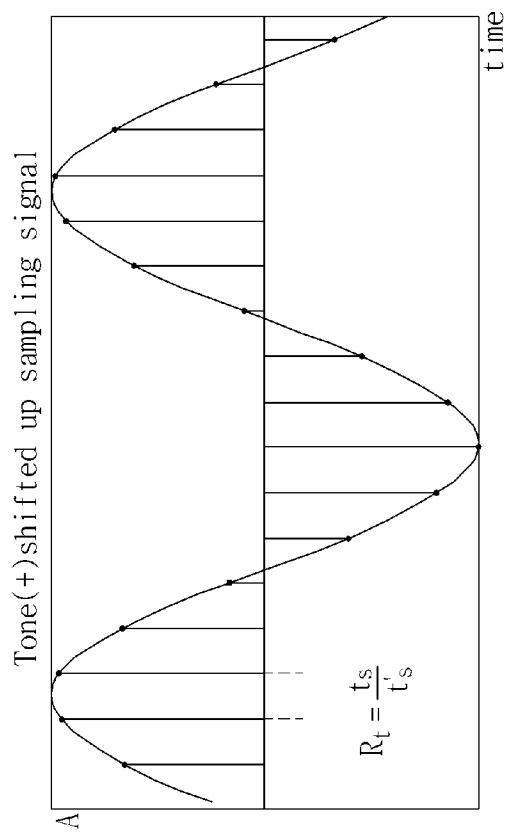
[Fig. 8]



[Fig. 9]



[Fig. 10]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2013/008905**A. CLASSIFICATION OF SUBJECT MATTER****G10L 19/00(2006.01)i, G11B 20/10(2006.01)i**

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L 19/00; H03M 7/30; G10L 19/08; G10L 13/06; G11B 20/10

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean Utility models and applications for Utility models: IPC as above

Japanese Utility models and applications for Utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & Keywords: audio, IMDCT, frequency, phase, amplitude, tune, speed, audio, IMDCT, frequency, hase, pulse width, tone, speed

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 10-1997-0050862 A (NIPPON VICTOR K.K.) 29 July 1997 See abstract, claim 1	1-24
A	JP 2009-501353 A (KONINKLIJKE PHILIPS NV) 15 January 2009 See abstract, claim 2	1-24
A	JP 2001-102932 A (SANYO ELECTRIC CO LTD) 13 April 2001 See abstract, figure 1	1-24



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

06 NOVEMBER 2013 (06.11.2013)

Date of mailing of the international search report

11 NOVEMBER 2013 (11.11.2013)

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex-Daejeon, 189 Seonsa-ro, Daejeon 302-701,
Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2013/008905

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-1997-0050862 A	29/07/1997	CN1135531 C0	21/01/2004
		CN1164084 A0	05/11/1997
		JP 03-265962B2	18/03/2002
		JP 09-185392A	15/07/1997
		TW418384 B	11/01/2001
		US 05862232 A	19/01/1999
JP 2009-501353 A	15/01/2009	AT443318 T	15/10/2009
		CN101223581 A0	16/07/2008
		DE602006009271 D1	29/10/2009
		EP 1905009 B1	16/09/2009
		ES2332108 T3	26/01/2010
		JP 2009-501353 T	15/01/2009
		RU2008105555 A	20/08/2009
		US 2010-0131276 A1	27/05/2010
		WO 2007-007253 A1	18/01/2007
JP 2001-102932 A	13/04/2001	JP 3778739 B2	24/05/2006

A. 발명이 속하는 기술분류(국제특허분류(IPC)) G10L 19/00(2006.01)i, G11B 20/10(2006.01)i		
B. 조사된 분야 조사된 최소문헌(국제특허분류를 기재) G10L 19/00; H03M 7/30; G10L 19/08; G10L 13/06; G11B 20/10 조사된 기술분야에 속하는 최소문헌 이외의 문헌 한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우)) eKOMPASS(특허청 내부 검색시스템) & 키워드: 오디오, IMDCT, 주파수, 위상, 진폭, 음정, 속도, audio, IMDCT, frequency, hase, pulse width, tone, speed		
C. 관련 문헌		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
A	KR 10-1997-0050862 A (닛폰비쿠다가부시키가이샤) 1997.07.29 요약, 청구항 제1항	1-24
A	JP 2009-501353 A (필립스) 2009.01.15 요약, 청구항 제2항	1-24
A	JP 2001-102932 A (SANYO ELECTRIC CO LTD) 2001.04.13 요약, 도면 1	1-24
☐ 추가 문헌이 C(계속)에 기재되어 있습니다. ☒ 대응특허에 관한 별지를 참조하십시오.		
* 인용된 문헌의 특별 카테고리: “A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌 “E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌 “L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌 “O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌 “P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌 “T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌 “X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다. “Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다. “&” 동일한 대응특허문헌에 속하는 문헌		
국제조사의 실제 완료일 2013년 11월 06일 (06.11.2013)		국제조사보고서 발송일 2013년 11월 11일 (11.11.2013)
ISA/KR의 명칭 및 우편주소  대한민국 특허청 (302-701) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사) 팩스 번호 +82-42-472-7140		심사관 이수철 전화번호 +82-42-481-8120 

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-1997-0050862 A	1997/07/29	CN1135531 C0 CN1164084 A0 JP 03-265962B2 JP 09-185392A TW418384 B US 05862232 A	2004/01/21 1997/11/05 2002/03/18 1997/07/15 2001/01/11 1999/01/19
JP 2009-501353 A	2009/01/15	AT443318 T CN101223581 A0 DE602006009271 D1 EP 1905009 B1 ES2332108 T3 JP 2009-501353 T RU2008105555 A US 2010-0131276 A1 WO 2007-007253 A1	2009/10/15 2008/07/16 2009/10/29 2009/09/16 2010/01/26 2009/01/15 2009/08/20 2010/05/27 2007/01/18
JP 2001-102932 A	2001/04/13	JP 3778739 B2	2006/05/24