(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2014-0008688
(43) 공개일자 2014년01월22일

(51) 국제특허분류(Int. Cl.)
G10L 19/00 (2006.01)
(21) 출원번호 10-2012-0075530
(22) 출원일자 2012년07월11일
심사청구일자 2012년07월11일

(71) 출원인
조선대학교산학협력단
광주광역시 동구 필문대로 309 (서석동)
(72) 발명자
이강현
광주광역시 남구 효사랑길 14 (봉선동, 포스코더샵아파트 108동 2005호)
(74) 대리인
정성종

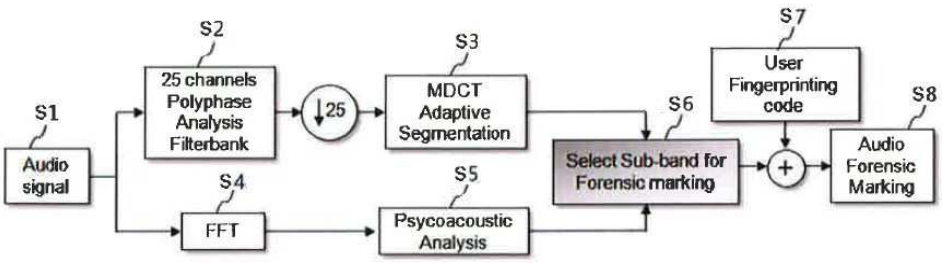
전체 청구항 수 : 총 6 항

(54) 발명의 명칭 **오디오 인코딩 방법**

(57) 요 약

본 발명에 따른 오디오 신호 인코딩 방법은 i) 아날로그 형태의 오디오 신호를 입력받아서 소정 개수의 주파수 대역으로 세분하고, 상기 세분된 신호에 대하여 변형이산여현변환(MDCT : Modified Discrete Cosine Transform)을 수행하는 단계; ii) 상기 아날로그 형태를 입력받아 FFT(Fast Fourier Transform)를 수행하고, 상기 FFT가 수행된 신호에 심리음향모델2를 적용하여 양자화를 진행하는 단계; iii) 상기 변형이산여현변환된 신호와 상기 심리음향모델2를 적용한 신호를 조합하여 MPEG 오디오 신호를 생성하는 단계; iv) 상기 MPEG 오디오 신호에서 포렌식 마킹코드를 삽입할 서브밴드를 선택하는 단계; 및 v) 상기 서브밴드에 각각의 사용자 고유의 포렌식 마킹코드를 삽입하여 포렌식 마킹된 신호를 생성하는 단계를 포함한다.

대 표 도 - 도2



특허청구의 범위

청구항 1

오디오 신호에 포렌식 마킹코드를 삽입함으로써 오디오 신호에 대한 공모공격 여부를 판단할 수 있는 오디오 신호 인코딩 방법에 있어서,

i) 아날로그 형태의 오디오 신호를 입력받아서 소정 개수의 주파수 대역으로 세분하고, 상기 세분된 신호에 대하여 변형이산여현변환(MDCT : Modified Discrete Cosine Transform)을 수행하는 단계;

ii) 상기 아날로그 형태를 입력받아 FFT(Fast Fourier Transform)를 수행하고, 상기 FFT가 수행된 신호에 심리음향모델2를 적용하여 양자화를 진행하는 단계;

iii) 상기 변형이산여현변환된 신호와 상기 심리음향모델2를 적용한 신호를 조합하여 MPEG 오디오 신호를 생성하는 단계;

iv) 상기 MPEG 오디오 신호에서 포렌식 마킹코드를 삽입할 서브밴드를 선택하는 단계; 및

v) 상기 서브밴드에 각각의 사용자 고유의 포렌식 마킹코드를 삽입하여 포렌식 마킹된 신호를 생성하는 단계;를 포함하는 것을 특징으로 하는 오디오 신호 인코딩 방법.

청구항 2

제1항에 있어서,

iv)단계는 MPEG 오디오 신호의 레벨이 최소가청한계 이하인 적어도 하나 이상의 대역을 서브밴드로 선택하는 것을 특징으로 하는 오디오 신호 인코딩 방법.

청구항 3

제1항에 있어서,

상기 단계 i)에서는 오디오 신호를 25개의 서브 밴드로 세분화하는 것을 특징으로 하는 오디오 신호 인코딩 방법.

청구항 4

제1항에 있어서,

상기 단계 ii)에서는 2,048 포인트 FFT를 수행하는 것을 특징으로 하는 오디오 신호 인코딩 방법.

청구항 5

제1항에 있어서,

상기 단계 v)의 각 사용자 고유의 포렌식 마킹코드는 7 비트로 구성된 것을 특징으로 하는 오디오 신호 인코딩 방법.

청구항 6

제5항에 있어서,

상기 7비트의 포렌식 마킹코드 중 4비트는 '1'이 삽입되는 것을 특징으로 하는 오디오 신호 인코딩 방법.

명 세 서

기 술 분 야

[0001] 본 발명은 오디오 인코딩 방법에 관한 것으로서, 보다 상세하게는 오디오 신호의 왜곡을 방지하면서도 사용자 고유의 포렌식 마킹코드를 삽입할 수 있는 오디오 인코딩 방법에 관한 것이다.

배 경 기 술

[0002] MPEG 오디오는 고품질 및 고능률의 스테레오 부호화를 위한 ISO/IEC의 표준 방식이다. 이러한 MPEG 오디오는 종래의 오디오 압축 부호화 방식에 비하여 음질이 뛰어난 장점으로 인하여, 최근 보급이 크게 확산되고 있는 다양한 종류의 모바일 기기에서의 널리 활용되고 있는 실정이다.

[0003] 도 1은 종래 MPEG 오디오 코딩 방법에 대한 개략도이다. 도 1에 도시된 바와 같이 입력된 아날로그 오디오 신호 s(n)는 32채널 필터뱅크에 의해 32개의 밴드로 세분화된 후 MDCT 단계로 입력된다. 다른 한편으로는 입력된 아날로그 오디오 신호에 대해 FFT를 수행한 후 심리음향분석을 진행한다. 이와 같이 MDCT가 수행된 신호와 심리음향분석 단계를 거친 신호를 이용해서 MP3 포맷의 신호를 생성하며, 생성된 MP3 포맷의 신호는 멀티플렉서를 거쳐 전송된다.

[0004] 한편, Apple의 스티브 잡스가 2007년에 DRM(Digital Right Management) Free 서비스를 제안하면서 DRM Free 서비스가 확산되고 있는 추세이지만 그에 비례하여 오디오 콘텐츠에 대한 저작권의 침해가 심각하게 대두되고 있는 실정이다.

[0005] 이와 같은 저작권 침해를 막기 위해 오디오 포렌식 마킹(Audio Forensic Marking)에 대한 연구가 활발히 진행되고 있다. 오디오 포렌식 마킹은 오디오 콘텐츠에 구매자 또는 사용자 고유의 코드를 삽입하여 콘텐츠가 불법으로 배포되었을 경우 그 배포자를 추적할 수 있도록 하는 기술이다.

[0006] 종래에 사용되는 오디오 포렌식 마킹 기술로는 확산 스펙트럼을 이용하는 방법, Phase 코딩 방법 및 Echo hiding 방법 등이 있으나, 이러한 방법들에서는 삽입되는 포렌식 마킹용 코드의 양이 많아서 원래의 오디오 신호에 왜곡이 발생되고 그 결과 오디오 신호의 음질이 저하되는 문제가 있었다.

발명의 내용

해결하려는 과제

[0007] 본 발명은 상기와 같은 문제를 해결하기 위한 것으로서, 오디오 신호의 에너지가 마스킹 스펙트럼 에너지보다 작은 서브밴드에 포렌식 마킹코드를 삽입함으로써 음질이 저하되는 것을 방지하면서 포렌식 마킹코드를 삽입할 수 있는 방법을 제공하는 것을 목적으로 한다.

과제의 해결 수단

[0008] 본 발명에 따른 오디오 신호 인코딩 방법은 i) 아날로그 형태의 오디오 신호를 입력받아서 소정 개수의 주파수 대역으로 세분하고, 상기 세분된 신호에 대하여 변형이산여현변환(MDCT : Modified Discrete Cosine Transform)을 수행하는 단계; ii) 상기 아날로그 형태를 입력받아 FFT(Fast Fourier Transform)를 수행하고, 상기 FFT가 수행된 신호에 심리음향모델2를 적용하여 양자화를 진행하는 단계; iii) 상기 변형이산여현변환된 신호와 상기 심리음향모델2를 적용한 신호를 조합하여 MPEG 오디오 신호를 생성하는 단계; iv) 상기 MPEG 오디오 신호에서 포렌식 마킹코드를 삽입할 서브밴드를 선택하는 단계; 및 v) 상기 서브밴드에 각각의 사용자 고유의 포렌식 마킹코드를 삽입하여 포렌식 마킹된 신호를 생성하는 단계;를 포함하는 것을 특징으로 한다.

[0009] 단계 iv)에서는 MPEG 오디오 신호의 레벨이 최소가청한계 이하인 적어도 하나 이상의 대역을 서브밴드로 선택하는 것이 바람직하며, 상기 단계 i)에서는 오디오 신호를 25개의 서브 밴드로 세분화되는 것이 바람직하다.

[0010] 상기 단계 ii)에서는 2,048 포인트 FFT를 수행하는 것을 특징으로 할 수 있으며, 상기 단계 v)의 각 사용자 고

유의 포렌식 마킹코드는 7 비트로 구성되며, 이 경우 7비트의 포렌식 마킹코드 중 BIBD 코드 기반으로 4개 비트의 '1'이 삽입되는 것이 바람직하다.

발명의 효과

[0011] 본 발명에 따르면 오디오 신호에서 마스킹 스펙트럼 에너지보다 에너지가 작은 주파수 대역을 선택하고 그 주파수 대역에 포렌식 마킹도르를 삽입함으로써 오디오 신호의 음질이 저하되지 않도록 하면서 포렌식 마킹코드를 삽입할 수 있게 된다.

도면의 간단한 설명

[0012] 도 1은 종래 MPEG 오디오 코딩 방법에 대한 개략도.

도 2는 본 발명에 따른 오디오 인코딩 방법을 설명하기 위한 블럭 다이어그램.

도 3은 입력 신호인 아날로그 형태의 오디오 신호의 예시도.

도 4는 아날로그 오디오 신호에 대해 FFT를 수행한 결과 스펙트럼.

도 5는 마스킹 스펙트럼에서의 임계대역 스펙트럼.

도 6은 아날로그 오디오 신호를 FFT 결과 및 마스킹 스펙트럼

도 7은 공모자를 추적하는 과정을 설명하기 위한 블럭 다이어그램.

도 8은 차분신호의 스펙트럼.

발명을 실시하기 위한 구체적인 내용

[0013] 이하 도면을 참조하여 본 발명에 대해 보다 상세히 설명하도록 한다.

[0014] 이하에서 설명할 실시예는 본 발명의 명확한 이해를 돕기 위한 것이며, 본 발명을 특정한 형태로 한정하려는 것은 아니다.

[0015] 또한, 이하에서 실시예를 설명함에 있어서 동일 구성에 대해서는 동일 명칭 및 동일 부호가 사용되며, 부가적인 설명의 반복은 생략하도록 한다.

[0016] 본 발명에 따른 오디오 인코딩 방법에서는, 아날로그 형태의 오디오 신호를 입력받아 디지털 신호로 변환시키고, 변환된 신호에 사용자 개개인별로 할당된 고유의 2진수 코드(이하 '포렌식 마킹코드'라 칭함)를 삽입하여 포렌식 마킹이 이루어진 오디오 신호를 생성함으로써 오디오 신호에 대한 공모공격(Collusion Attack) 여부 및 공모자를 확인하게 된다.

[0017] 도 2는 본 발명에 따른 오디오 인코딩 방법을 설명하기 위한 블럭 다이어그램이며, 도 3은 입력 신호인 아날로그 형태의 오디오 신호의 예시 및 그 오디오 신호의 150번째 프레임에 대한 부분 확대도이다.

[0018] 입력된 아날로그 형태의 오디오 신호(S1)는 폴리페이즈 필터뱅크(S2)로 전달된다. 폴리페이즈 필터뱅크(S2)는 오디오 신호의 통계적인 중복성을 제거하기 위해 입력된 아날로그 형태의 오디오 신호를 25개의 주파수 대역으로 세분화한다.

[0019] MDCT(S3)에서는 주파수 분해능(frequency resolution)을 증가시키기 위해 변형이산여현변환(MDCT : Modified Discrete Cosine Transform)이 수행된다.

[0020] 한편, 입력된 아날로그 오디오 신호(S1)에 대해서는 FFT(Fast Fourier Transform)도 수행된다(S4). 즉, 아날로그 오디오 신호(S1)는 FFT(S4)에 의해 주파수 대역의 스펙트럼으로 변환된 후 심리음향모델2에 따른 지각부호화가 진행된다(S5). 도 4는 아날로그 오디오 신호에 대해 FFT를 수행한 결과를 상기 25개의 주파수 대역에 대응되도록 도시한 것이다. 표 1은 25개 주파수 대역의 중심 주파수(Hz) 및 주파수 대역폭(Hz)을 나타낸 것이다.

[0021]

표 1

[0022]

서브밴드 No.	주파수 대역(Hz)	중심 주파수(Hz)
0	0 ~ 100	50
1	100 ~ 200	150
2	200 ~ 300	250
3	300 ~ 400	350
4	400 ~ 510	450
5	510 ~ 630	570
6	630 ~ 770	700
7	770 ~ 920	840
8	920 ~ 1080	1000
9	1080 ~ 1270	1170
10	1270 ~ 1480	1370
11	1480 ~ 1720	1600
12	1720 ~ 2000	1850
13	2000 ~ 2320	2150
14	2320 ~ 2700	2500
15	2700 ~ 3150	2900
16	3150 ~ 3700	3400
17	3700 ~ 4400	4000
18	4400 ~ 5300	4800
19	5300 ~ 6400	5800
20	6400 ~ 7700	7000
21	7700 ~ 9500	8500
22	9500 ~ 12000	10500
23	12000 ~ 15500	13500
24	15500 ~ 22050	19500

[0023] 이후, 변형이산여현변환된 신호와 심리음향모델2에 따른 지각부호화가 진행된 신호를 조합하여 MPEG 오디오 신호를 생성한다(S6). 이상에서의 MPEG 오디오 신호 생성과정은 해당 기술분야에서 자명한 사항이므로 MPEG 오디오 신호 생성과정과 관련된 보다 상세한 설명은 생략하도록 한다.

[0024] 본 발명에서는 상기의 MPEG 오디오 신호의 음질 손상을 방지하면서 MPEG 오디오 신호에 포렌식 마킹코드(S7)를 삽입하게 된다. 포렌식 마킹코드는 사용자별 MPEG 오디오 신호에는 변형이 발생되는 것이며 그 결과 음질의 손상을 초래할 수 있다. 그러나, 본 발명에서는 사람이 들을 수 없는 주파수 대역을 선택하여 포렌식 마킹코드를 삽입함으로써 음질의 저하를 방지할 수 있도록 한다.

[0025] 도 5는 마스킹 스펙트럼(Masking Spectrum)에서의 임계대역(Critical Band) 및 그 대역폭을 상기에서 설명한 25개 밴드에 대응되도록 도시한 것이다. 오디오 신호는 청취가 가능한 최소 레벨을 가지며, 이러한 최소 레벨은 주파수에 따라 달라진다. 한편, 청취가 가능한 최소 레벨은 함께 들리는 다른 음의 레벨에 영향을 받는다. 즉, 고요할 경우에는 청취가 가능한 레벨이라도 함께 들리는 음의 레벨이 상대적으로 높을 경우에는 들을 수 없게 되는데 이를 마스킹 효과라 하며, 이러한 마스킹 효과가 발생하는 주파수 대역을 임계대역이라 한다.

[0026] 만일 오디오 신호의 특정 주파수 성분의 레벨이 임계대역 내에서의 마스킹 스펙트럼의 레벨(최소가청한계)보다 낮을 경우, 사람은 오디오 신호 중 그 특정 주파수 성분을 들을 수 없게 된다. 따라서, 본 발명에서는 MPEG 오디오 신호 중 그 레벨이 마스킹 스펙트럼의 레벨보다 낮은 주파수 대역을 선택(이하 '서브밴드'라 칭함)하고, 선택된 서브밴드에 포렌식 마킹코드를 삽입하여 포렌식 마킹코드가 삽입된 오디오 신호를 생성한다(S8). 이처럼 상기의 서브밴드에 포렌식 마킹코드를 삽입할 경우, MPEG 오디오 신호를 구성하는 데이터에는 변형이 발생되지만, 그러한 변형은 변형 전의 MPEG 오디오 신호에서도 사람이 들을 수 없었던 주파수 대역에서만 이루어지므로 포렌식 마킹코드를 삽입한 이후에도 사람이 듣는 음질에는 변화가 없게 된다.

[0027] 도 6은 아날로그 오디오 신호를 FFT한 결과와 마스킹 스펙트럼을 상기에서 설명한 25개 주파수 대역에 대하여 도시한 것이다. 도 6에 도시된 바와 같이 본 실시예에 사용된 오디오 신호에서는 0번 대역, 6내지 9번 대역, 23내지 24번 대역에서 FFT된 오디오 신호의 레벨이 마스킹 스펙트럼의 레벨보다 낮은 것을 알 수 있다. 따라서,

위 3개의 영역이 포렌식 마킹코드를 삽입하기 위한 서브밴드로 선택된다(S6). 이처럼 서브밴드는 하나의 오디오 신호에서 복수개 존재할 수 있다.

[0028] 표 2는 7명의 사용자에게 할당된 포렌식 마킹코드를 예시한 것이다.

표 2

[0029]

사용자 번호	포렌식 마킹 코드
1	1 0 1 0 1 0 1
2	0 1 1 0 0 1 1
3	1 1 0 0 1 1 0
4	0 0 0 1 1 1 1
5	1 0 1 1 0 1 0
6	0 1 1 1 1 0 0
7	1 1 0 1 0 0 1

[0030] 본 발명에서는 7비트의 {v, k, λ} BIBD 코드로 마킹코드를 구성하되(v=7), 4비트에는 1을 할당하면서(k=4), 7명의 사용자 중 임의의 사용자 2명을 선택하여 1이 할당된 위치를 대비할 경우 1의 위치가 2군데(λ=2)에서 겹치도록 포렌식 마킹코드를 구성한다.

[0031] 도 7은 본 발명에 따른 포렌식 마킹코드를 삽입한 이후, 삽입된 포렌식 마킹코드를 이용하여 오디오 콘텐츠를 공모공격한 공모자를 추적하는 과정을 설명하기 위한 블럭 다이어그램이다.

[0032] 공모자를 추적하기 위해서는 우선 포렌식 마킹이 이루어진 신호, 즉 공모공격의 대상이 되었을 것으로 생각되는 신호(S11)에 대하여 FFT(S12)를 수행한다. 한편, 포렌식 마킹이 이루어지지 않은 원래의 오디오 신호(S13)에 대해서도 FFT(S14)를 수행한다. 이후 FFT(S12, S14)가 수행된 2개의 신호를 이용하여 차분신호를 구한다. 도 8은 차분신호를 도시한 예시이다.

[0033] 차분신호에서 마킹코드가 삽입되는 대역(예를 들어, 도 6에서의 3개의 영역)을 선택하고(S15), 선택된 영역에 대하여 차분신호와 포렌식 마킹코드간의 상관관계 계수를 구하여(S16, S17) 공모공격이 이루어졌는지를 판단하며(S18), 각 사용자 고유의 포렌식 마킹코드와 대비를 하면서 상관관계 계수를 살피게 되면 어느 사용자들이 공모공격에 관여를 했는지 살피게 된다.

[0034] 표 3은 2명 내지 7명이 공모공격에 관여한 경우에 대하여 공모자를 추적한 결과를 나타낸 것이다. 이때, A는 포렌식 마킹 세기를 의미하는 것으로서, 포렌식 마킹코드가 삽입된 신호의 강도를 기준값(1) 대비 0.1 내지 0.9 레벨로 변화를 시킨 것을 의미하며, B는 화이트 노이즈의 세기로서 화이트 노이즈의 세기 또한 0.1 내지 0.9 레벨로 변화를 시킨 조건에서 공모자의 추적결과를 나타낸 것이다. 본 본 실시예에서는 공모공격의 형태로는 평균화 공모공격을 사용하였다. 이러한 평균화 공모공격은 해당 기술분야에서의 모의실험에서 일반적으로 사용되는 방법인바 보다 상세한 설명은 생략하도록 한다.

표 3

[0035]

A	B	공모자의 수					
		2	3	4	5	6	7
0.1	0	1	1	1	1	1	0.14
	0.1	0.5	0.33	0	0.2	0.17	0.14
	0.3	0.5	0.33	0.25	0.2	0.17	0.14
	0.5	0.5	0.33	0.25	0.2	0.17	0.14
	0.7	0.5	0.33	0.25	0.2	0.17	0
	0.9	0.5	0.33	0.25	0.2	0.17	0

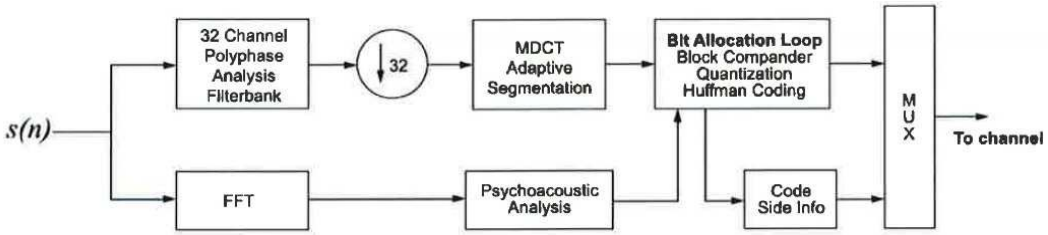
0.3	0	1	1	1	1	1	0.14
	0.1	1	0.33	0.25	0.2	0.17	0.14
	0.3	0.5	0.33	0	0.2	0.17	0.14
	0.5	0.5	0.33	0.25	0.2	0.17	0.14
	0.7	0.5	0.33	0.25	0.2	0.17	0.14
	0.9	0.5	0.33	0.25	0.2	0.17	0.14
0.5	0	1	1	1	1	1	0.14
	0.1	1	1	0.5	0.4	0.17	0.14
	0.3	0.5	0.33	0.5	0.2	0.17	0.14
	0.5	0.5	0.33	0	0.2	0.17	0.14
	0.7	0.5	0.33	0.25	0.2	0.17	0.14
	0.9	0.5	0.33	0.25	0.2	0.17	0.14
0.7	0	1	1	1	1	1	0.28
	0.1	1	1	0.75	0.4	0.33	0.14
	0.3	1	0.33	0.5	0.2	0.17	0.14
	0.5	0.5	0.33	0	0.2	0.17	0.14
	0.7	0.5	0.33	0	0.2	0.17	0.14
	0.9	0.5	0.33	0.25	0.2	0.17	0.14
0.9	0	1	1	1	1	1	0.43
	0.1	1	1	1	0.6	0.33	0.28
	0.3	1	0.67	0.5	0.4	0.17	0.14
	0.5	0.5	0.33	0.25	0.2	0.17	0.14
	0.7	0.5	0.33	0	0.2	0.17	0.14
	0.9	0.5	0.33	0	0.2	0.17	0.14

[0036] 표 3에 나타난 바와 같이, 화이트 노이즈의 세기가 0(zero)인 경우 2명 내지 6명이 공모에 참여한 경우 상관관계 계수가 모두 '1'로 산출되었는바, 2명 내지 6명이 공모에 참여한 경우 해당 공모자가 정확히 추적됨을 알 수 있다.

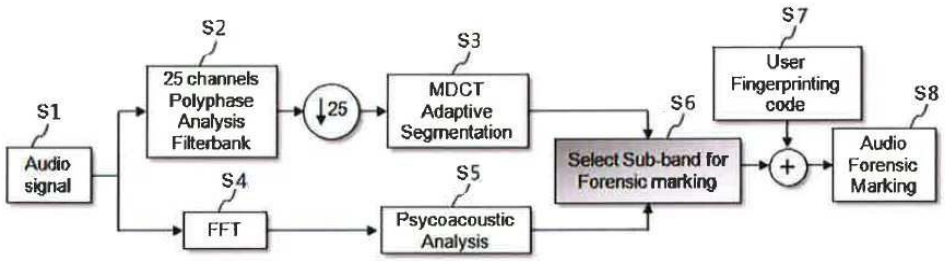
[0037] 이상에서 살펴본 바와 같이, 본 발명에 따른 오디오 인코딩 방법에서는, 청취가 불가능한 주파수 대역에 포렌식 마킹코드를 삽입함으로써 음질의 저하를 방지할 수 있음과 동시에, 원래(original)의 신호와 포렌식 마킹코드와의 차분신호를 구한 후, 이 차분신호와 포렌식 마킹코드와의 상관관계를 분석함으로써 오디오 신호에 대한 공모 공격여부 및 공모자를 명백히 파악할 수 있게 된다.

도면

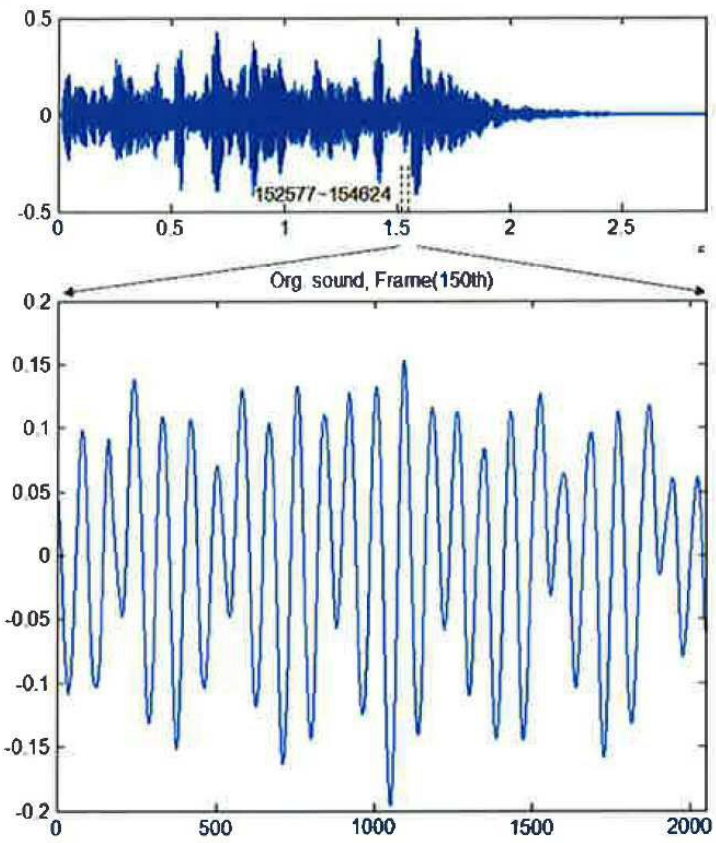
도면1



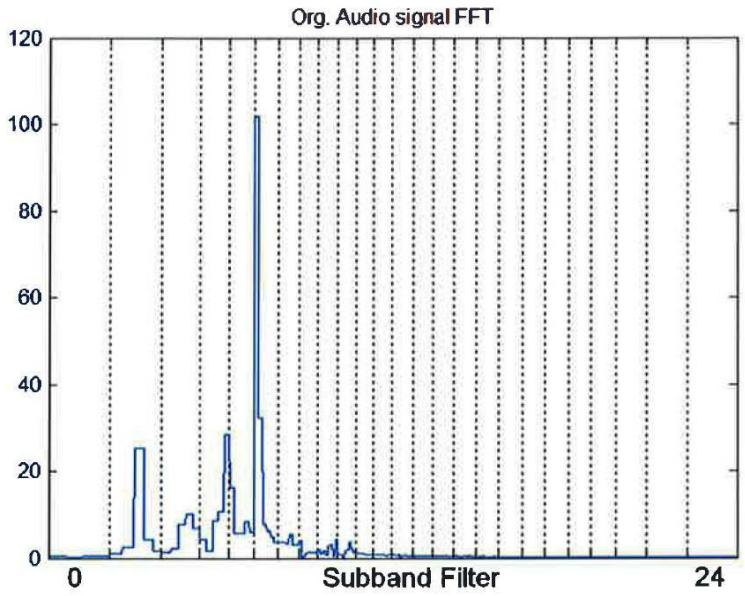
도면2


S1
Audio signal
S2
25 channels Polyphase Analysis Filterbank
↓25
S3
MDCT Adaptive Segmentation
S4
FFT
S5
Psycoacoustic Analysis
S6
Select Sub-band for Forensic marking
S7
User Fingerprinting code
+
S8
Audio Forensic Marking

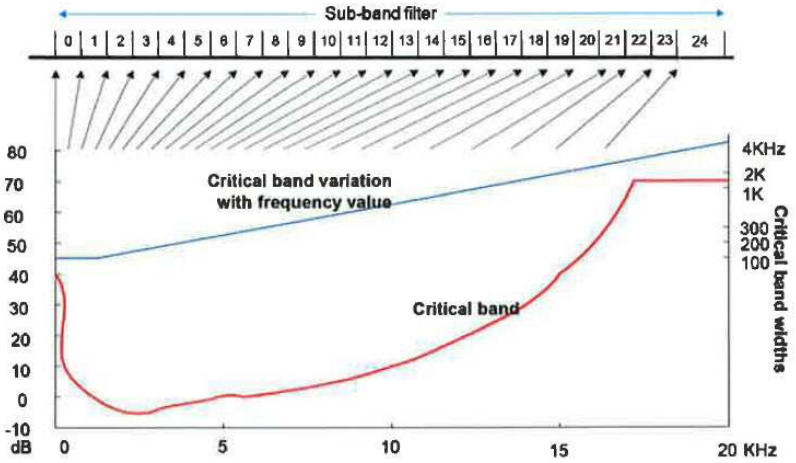
도면3


152577~154624
Org sound, Frame(150th)

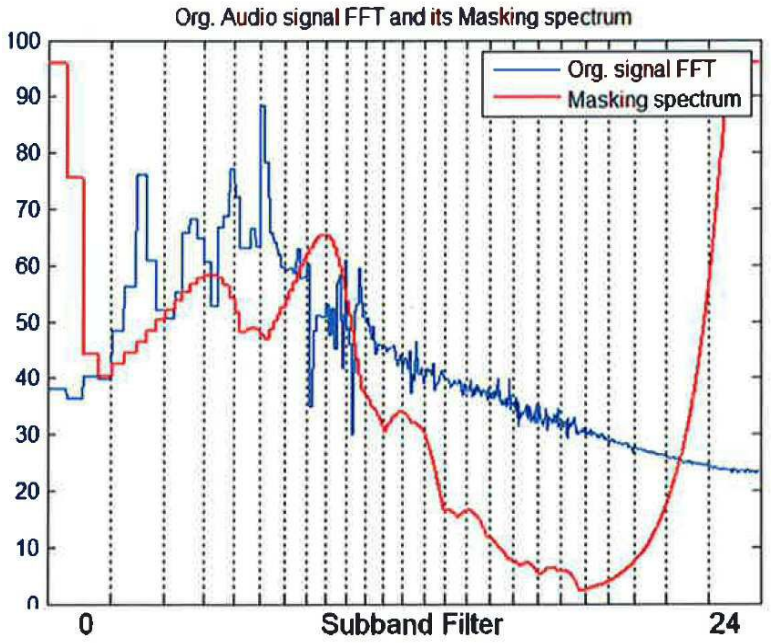
도면4


Org. Audio signal FFT
120
100
80
60
40
20
0
0
Subband Filter
24

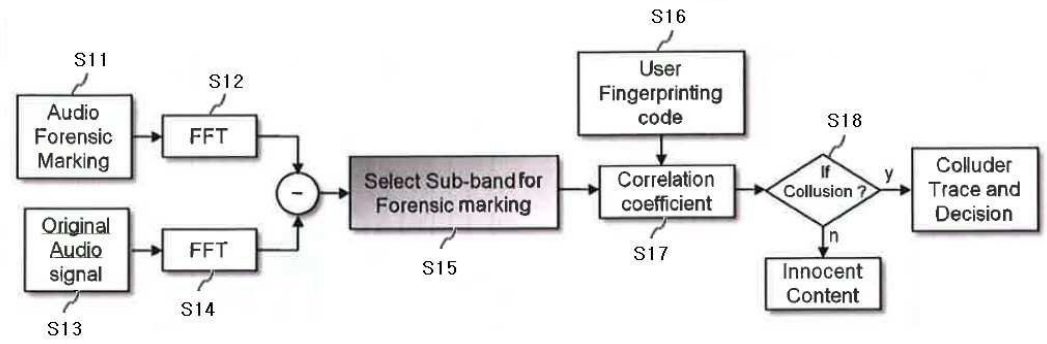
도면5


Sub-band filter
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
80
70
60
50
40
30
20
10
0
-10
dB
Critical band variation
with frequency value
Critical band
4KHz
2K
1K
300
200
100
Critical band widths
0
5
10
15
20 KHz

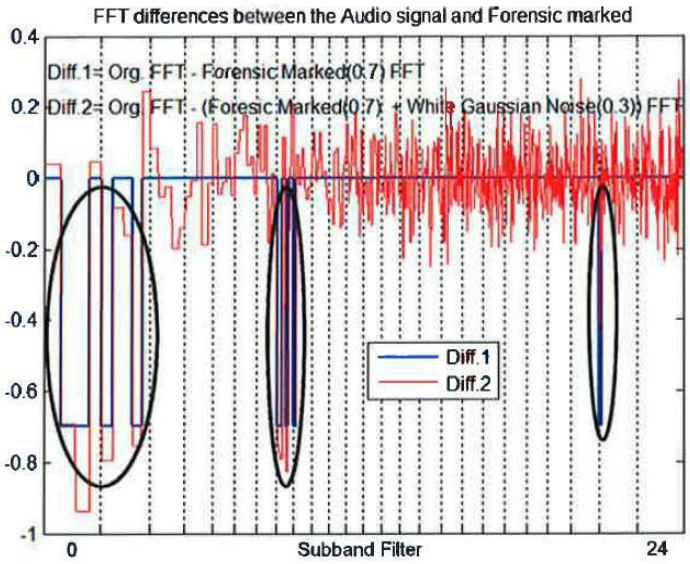
도면6


Org. Audio signal FFT and its Masking spectrum
Org. signal FFT
Masking spectrum
100
90
80
70
60
50
40
30
20
10
0
0
Subband Filter
24

도면7


S11
Audio Forensic Marking
S12
FFT
Original Audio signal
S13
FFT
S14
Select Sub-band for Forensic marking
S15
S16
User Fingerprinting code
Correlation coefficient
S17
S18
If Collusion ?
y
n
Colluder Trace and Decision
Innocent Content

도면8


FFT differences between the Audio signal and Forensic marked
Diff.1= Org. FFT - Forensic Marked(0.7) FFT
Diff.2= Org. FFT - (Foresic Marked(0.7) + White Gaussian Noise(0.3)) FFT
Diff.1
Diff.2
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0
Subband Filter
24