



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(51) Int. Cl.

H04S 5/00 (2006.01)

(45) 공고일자

2007년01월09일

(11) 등록번호

10-0666815

(24) 등록일자

2007년01월03일

(21) 출원번호	10-2005-7018212(분할)	(65) 공개번호	10-2005-0100012
(22) 출원일자	2005년09월27일	(43) 공개일자	2005년10월17일
심사청구일자	2005년10월11일		
번역문 제출일자	2005년09월27일		
(62) 원출원	특허10-2004-7000072	심사청구일자	2004년01월05일
	원출원일자 : 2004년01월03일	(87) 국제공개번호	WO 2003/007656
(86) 국제출원번호	PCT/SE2002/001372	국제공개일자	2003년01월23일
국제출원일자	2002년07월10일		

(30) 우선권주장	0102481-9	2001년07월10일	스웨덴(SE)
	0200796-1	2002년03월15일	스웨덴(SE)
	0202159-0	2002년07월09일	스웨덴(SE)

(73) 특허권자
 코딩 테크놀러지스 에이비
 스웨덴 스톡홀름 에스-113 52 도벨른스가탄 64

(72) 발명자
 헨 프레드릭
 스웨덴 브로나 에스-168 31 리타르바겐 14

크홀링 크리스토퍼
 스웨덴 솔나 에스-170 75 로스티겐 10

릴제리드 라스 구스타프
 스웨덴 솔나 에스-171 34 빈터바겐19

로텐 조나스
 스웨덴 솔나 에스-169 34 에릭 샌드버르그 가타 6

엔그데가아드 조나스
 스웨덴 스톡홀름 에스-115 43 벤스트롬스바겐 6

(74) 대리인
 양순석

심사관 : 선동국

전체 청구항 수 : 총 6 항

(54) 스테레오 신호 또는 멀티채널 신호의 채널을 발생하는 임향기 및 임향방법

(57) 요약

본 발명은 수신된 모노 신호의 후처리(post-processing)를 통하여 스테레오 착각(stereo-illusion)을 발생시키는 종래 기술의 오디오 코덱의 개선에 관한 것이다. 이러한 개선은, 인코더측에서 스테레오 이미지를 기술하는 파라미터의 추출에 의해서 달성되며, 이들 파라미터는 디코더측에 전송되어 스테레오 발생기의 제어에 이용된다. 또한, 본 발명은 새로운 형태의 파라미터 스테레오 코딩을 이용하여 간단한 의사 스테레오(pseudo-stereo) 방법 및 현재의 진정한(true) 스테레오 코딩 방법간의 캡을 이어준다. 더 진보된 스테레오 모드를 가능하게 하는 스테레오 파라미터가 도입되고, 또한 안내된 HFR (High Frequency Reconstruction)이 채용되는 시스템에서 특별히 이용되는 스펙트럼 포락선을 스테레오 코딩하는 새로운 방법의 기초를 형성한다. 특별한 경우로서, 스케일러블 HFR 기반 코덱(scalable HFR-based codec)에 이 스테레오 코딩 기술을 적용하는 것이 기술된다.

대표도

도 1

특허청구의 범위

청구항 1.

스테레오 오디오 신호 또는 멀티채널 오디오 신호의 제1채널 및 제2채널을 발생하는 잔향기에 있어서:

상기 오디오 신호에서 사운드 종료를 검출하는 사운드종료검출기와,

상기 사운드 종료의 검출에 반응하여 잔향 신호의 이득을 변경함으로써 잔향 꼬리를 감쇠하거나 제거하는 감쇠기를 포함하는 것을 특징으로 하는 잔향기.

청구항 2.

삭제

청구항 3.

제1항에 있어서,

변경가능한 이득을 이용하여 잔향 신호를 발생하는 발생 수단을 포함하고,

상기 감쇠기는 동작시에 상기 발생 수단의 이득을 변경하는 것을 특징으로 하는 잔향기.

청구항 4.

스테레오 오디오 신호 또는 멀티채널 오디오 신호의 제1채널 및 제2채널을 발생하기 위한 잔향 방법에 있어서:

상기 오디오 신호에서 사운드 종료를 검출하는 사운드종료검출단계와,

상기 사운드 종료의 검출에 반응하여, 잔향 꼬리를 감쇠하거나 제거하는 단계를 포함하는 것을 특징으로 하는 잔향 방법.

청구항 5.

스테레오 오디오 신호 또는 멀티채널 오디오 신호의 제1채널 및 제2채널을 발생하는 잔향기에 있어서:

상기 오디오 신호에서, 잔향기가 그에 대해 인공음을 발생하는 특정신호를 검출하는 특정신호검출기와,

상기 특정신호의 검출에 반응하여, 잔향 신호의 이득을 변경함으로써 모든 인공음을 감쇠하거나 제거하는 감쇠기를 포함하는 것을 특징으로 하는 잔향기.

청구항 6.

스테레오 오디오 신호 또는 멀티채널 오디오 신호의 제1채널 및 제2채널을 발생하기 위한 잔향 방법에 있어서:

상기 오디오 신호에서, 잔향기가 그에 대해 인공음을 발생하는 특정신호를 검출하는 특정신호검출단계와,

상기 특정신호의 검출에 반응하여, 상기 인공음을 감쇠하거나 제거하는 단계를 포함하는 것을 특징으로 하는 잔향 방법.

청구항 7.

제5항에 있어서,

상기 특정 신호는 과도 신호를 포함하고,

상기 특정신호검출기는 과도 신호 검출기인 것을 특징으로 하는 잔향기.

명세서

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 스테레오 신호 또는 멀티채널 신호의 채널을 발생하는 잔향기 및 잔향방법에 관한 것이다.

입력 신호의 스테레오 특성에 대한 다른 파라미터 표현들이 도입되고, 스펙트럼 포락선(spectral envelope)의 의사 스테레오 코딩(pseudo-stereo coding)으로부터 풀 스테레오 코딩(full stereo coding)에 걸쳐서 디코더측에서의 그 적용이 설명되는데, 여기서 후자는 특히 HFR(고주파재생:High Frequency Reconstruction) 기반 코덱(codec)에 적합하다.

오디오 소스 코딩 기술은 두 가지로 분류될 수 있는데, 자연음(natural audio) 코딩과 음성(speech) 코딩이 그것이다. 고비트레이트(high bitrate)를 위한 매체에서, 자연음 코딩은 통상 음성 및 음악 신호에 이용되고, 스테레오 전송 및 재생이 가능하다. 저 비트레이트만이 이용 가능한 애플리케이션, 예컨대 느린 전화 모뎀 접속을 갖는 사용자를 타깃으로 하는 인터넷 스트리밍 오디오에서, 또는 최근 생겨난 디지털 AM 방송 시스템에서는, 오디오 프로그램 자료의 모노 코딩(mono coding)이 불가피하다. 그러나, 특히 헤드폰으로 감상할 때는 스테레오 효과가 여전히 바람직하며, 어떤 경우에는 순수 모노 신호(pure mono signal)가 헤드 내에서 생긴 그대로 감지될 수 있으며, 이는 불쾌한 경험이 될 수 있다.

이러한 문제에 접근하는 한가지 방법은 수신된 순수 모노 신호를 디코더측에서 스테레오 신호로 합성하는 것이다. 수년에 걸쳐서 몇 가지 다른 "의사 스테레오" 발생기들이 스테레오 신호 합성을 위하여 제안되어 왔다. 예컨대, USP 5,883,962호에서는, 신호의 지연/위상 시프트 버전을 미처리 신호에 부가하여 스테레오 착각(stereo-illusion)을 일으키는 모노 신호의 확대가 기재되어 있다. 이렇게 처리된 신호가 동일 레벨이지만 반대 부호를 갖는 두개의 출력 각각에 대한 원시 신호(original signal)에 부가되어, 두개의 채널이 나중에 신호 경로 상에서 더해져서 확대 신호가 상쇄되는 것을 보장한다. PCT WO98/57436호에서는, 유사하지만 상기 확대 신호와의 모노 호환성이 없는 시스템이 개시되어 있다. 종래 기술의 방법들은 순수 후처리(pure post-processing)로서 적용된다는 공통점을 갖는다. 환언하면, 스테레오 폭(stereo-width)의 크기에 관한 어떠한 정보도 스테레오 사운드 스테이지의 위치에서는 말할 것도 없이 디코더에서도 이용 가능하지 않다. 따라서, 의사 스테레오 신호는 원시 신호의 스테레오 특성의 유사성을 갖거나 갖지 않을 수도 있다. 종래 기술의 시스템이 부족한 특별한 상황은 원시 신호가 순수 모노 신호인 경우이며, 이는 종종 음성 기록의 경우이다. 이 모노 신호는 디코더에서 맹목적으로 합성 스테레오 신호로 변환되어, 음성의 경우에 종종 불쾌한 인공음을 초래하고, 선명함 및 음성 명료성을

감소시킬 수도 있다. 모노 신호가 "언제나" 모노 신호인 스피치(speech) 신호일 때, 그리고 의사 스테레오(pseudo-stereo) 발생기에 의해 처리될 때, 디코더측에서는 모노 임프레션이 파괴되는데, 이것은 모노 스피치 신호에서는 그러한 스테레오 임프레션이 절대로 존재하지 않았음에도 불구하고 디코더가 스테레오 임프레션을 생성하려고 하기 때문이다. 따라서, 스피치의 명확성과 명백성을 낮춘다는 점에서 문제가 있다.

저 비트레이트에서의 진정한 스테레오 전송을 목표로 하는 다른 종래 기술의 시스템들은 전형적으로 합(sum) 및 차(difference) 코딩 시스템을 채용한다. 따라서, 원시 좌측(L) 및 우측(R) 신호들이 합 신호 $Sum = (L + R)/2$ 및 차 신호 $Dif = (L - R)/2$ 로 변환되고, 이어서 인코딩되어 전송된다. 수신기는 Sum 및 Dif 신호를 디코딩하고, 그 결과 원시 L/R 신호가 연산 $L = Sum + Dif$ 및 $R = Sum - Dif$ 를 통하여 재구성된다. 이것의 이점은 종종 L 및 R 간의 리던던시(redundancy)가 가깝고, 따라서 인코딩될 Dif 내의 정보가 Sum 보다 더 적고 더 적은 비트를 필요로 한다. 분명하게, 극단적인 경우는 순수 모노 신호, 즉 L 및 R 이 동일한 경우이다. 전통적인 L/R 코덱은 이러한 모노 신호를 두 번 인코딩하는 한편, S/D 코덱이 이 리던던시를 검출하고, Dif 신호는 비트를 전혀 필요로 하지 않는다(이상적임). 다른 극단적인 경우는 "위상이 다른(out of phase)" 신호에 대응하는 $R = -L$ 인 상황에 의해 표현된다. 이제, Sum 신호가 제로이므로, Dif 신호는 L로 계산된다. 게다가, S/D 기술은 표준 L/R 코딩에 대한 분명한 이점을 갖는다. 그러나, 예컨대 스테레오 기록의 초기에는 흔하지 않았던, 진행 중에 $R = 0$ 인 상황을 고려한다. Sum 및 Dif 는 모두 $L/2$ 이고, S/D 기술은 어떤 이점도 제공하지 않는다. 반면에, L/R 코딩은 이것을 매우 잘 다룬다. R 신호는 어떤 비트도 필요로 하지 않는다. 이 때문에, 종래 기술의 코덱은 주어진 순간에 이용하기에 어느 방법이 가장 이로운가에 따라 이들 두 코딩 기술 사이에서 적합하게 스위칭하는 것을 채용한다. 위의 예들은 (음성 프로그램들에서는 흔한 듀얼 모노 경우를 제외하고는) 이론적인 것일 뿐이다. 따라서, 실제의 스테레오 프로그램 자료는 상당한 양의 스테레오 정보를 포함하고 있으며, 상술한 스위칭이 구현되더라도 결과적인 비트레이트는 많은 애플리케이션들에게는 여전히 너무 높다. 또한, 상술한 재합성 관계로부터 알 수 있는 바와 같이, 양자화 에러는 L 및 R 신호에서의 무시할 수 없는 레벨의 에러가 되기 때문에, 비트레이트를 더 저감하려는 시도에서 Dif 신호의 매우 거친 양자화는 적합하지 않다.

발명이 이루고자 하는 기술적 과제

본 발명의 목적은 오디오 신호를 파라메트릭 코딩 또는 디코딩하는 다양한 콘셉트를 제공하기 위한 것으로서, 본 발명의 하나의 목적은 잔향기를 제공하는 것이다. 본 발명의 다른 목적은 잔향 방법을 제공하는 것이다. 본 발명의 목적은 독립 청구항들에 따른 방법들과 장치들에 의해 성취된다. 이 독립 청구항들에는 벨런스 파라미터와 폭 파라미터가 인코딩 측에서 계산되어서 전송되거나 저장된다. 또 이 청구항들은 디코더 측에서도 사용되는데, 사운드스테이지에서 사운드 이미지를 위치시키고 분산시킬 수 있도록 하고, 원래 신호의 스테레오 임프레션을 미믹팅 할 때에 융통성을 제공한다. 본 발명은 코딩 및 전송 전에 신호 스테레오 특성의 검출을 채용한다. 가장 간단한 형태로, 검출기가 입력된 스테레오 신호에 존재하는 스테레오 특성의 양을 측정한다. 다음에 이 양은 원시 신호의 인코딩된 모노 합과 함께 스테레오 폭 파라미터로서 전송된다. 수신기는 상기 모노 신호를 디코딩하고, 상기 파라미터에 의해 제어되는 의사 스테레오 발생기를 이용하여 적당한 양의 스테레오 폭을 인가한다. 특별한 경우로서, 모노 입력 신호는 제로 스테레오 폭으로서 나타내어지고, 따라서 디코더에서는 어떠한 스테레오 합성도 가해지지 않는다. 본 발명에 따르면, 스테레오 폭의 유용한 측정이, 예컨대, 차 신호로부터 또는 원시 좌측 및 우측 채널로부터 유도될 수 있다. 이러한 계산들의 값은 적은 수의 상태들로 매핑되고, 적절한 시간에 또는 필요에 따라서 적절한 고정 레이트로 전송된다. 본 발명은 또한 통상적으로 저 비트레이트 코딩 신호와 연관되어 있는 코딩 인공음을 차단하지 못하는 위험을 감소시키기 위하여, 합성된 스테레오 성분들을 필터링하는 방법을 교시한다.

선택적으로, 전체적인 스테레오 벨런스 또는 스테레오 필드에서의 국부화가 인코더에서 검출된다. 이 정보는 인코딩된 모노 신호와 함께 그리고 선택적으로 상기 폭 파라미터와 함께 벨런스 파라미터로서 효율적으로 전송된다. 따라서, 두개의 출력 채널의 이득을 대응하여 바꿈으로써 사운드 스테이지의 어느 한쪽으로의 전치(displacements)가 디코더에서 재생될 수 있다. 본 발명에 따르면, 이 스테레오 벨런스 파라미터는 좌측 및 우측 신호 전력들의 비율로부터 유도된다. 양 타입의 파라미터들의 전송은 풀 스테레오 코딩에 비하여 매우 적은 비트를 필요로 하며, 이에 의해 총 비트레이트 요구가 낮게 유지된다. 보다 정밀한 파라미터 스테레오 표현을 제공하는 본 발명의 보다 복잡한 버전에서는, 몇 개의 벨런스 및 스테레오 파라미터들이 이용되는데, 이들 각각은 별개의 주파수 대역을 나타낸다.

주파수 대역당 동작에 대하여 일반화된 벨런스 파라미터는, 좌측 및 우측 신호 전력들의 합으로서 계산된 레벨 파라미터의 대응하는 대역당 동작과 함께 스테레오 신호의 전력 스펙트럼 밀도에 대한 새롭고 임의적인 상세한 표현을 가능하게 한다. 이러한 표현의 특별한 이익은, S/D 시스템도 이용하는 스테레오 리던던시로부터의 이익 이외에도, 스테레오 스펙트럼 포락선으로 역변환될 때 양자화 에러는 레벨상의 에러라기 보다는 "공간상의 에러", 즉 스테레오 파노라마에 있어서의 감지된 국부화를 초래하기 때문에, 벨런스 신호가 동일 레벨보다 낮은 정밀도로 양자화될 수 있다는 것이다. 전통적인 스위칭 L/R 및 S/D 시스템과 유사하게, 레벨/밸런스 기술은, 전체 신호가 어느 한 채널쪽으로 심하게 오프셋될 때 레벨L/레벨R 신호를 위하여, 적응적으로 스위칭 오프될 수 있다. 상기 스펙트럼 포락선 코딩 기술은 전력 스펙트럼 포락선의 효율적인 코

당이 필요할 때마다 이용될 수 있고, 새로운 스테레오 소스 코덱에 툴(tool)로서 병합될 수 있다. 특히 흥미로운 애플리케이션은 원시 신호 고대역 포락선에 대한 정보에 의해 안내되는 HFR 시스템에 있다. 이러한 시스템에서, 저대역은 임의의 코덱에 의해 코딩 및 디코딩되고, 고대역은 디코딩된 저대역 신호 및 전송된 고대역 포락선 정보를 이용하여 재생된다 (PCT WO98/57436). 또한 스케일러블 HFR 기반 스테레오 코덱을 구성할 가능성은 포락선 코딩을 레벨/밸런스 동작에 한정함으로써 제공된다. 이에 의해, 구현 예에 따라서 통상적으로 모노 신호로 디코딩되는 1차 비트스트림으로 레벨 값들이 공급된다. 밸런스 값들은, 1차 비트스트림 이외에도 예로서 IBOC(In-Band On-channel) 디지털 AM 방송 시스템을 취하는 송신기에 가까운 수신기에 이용 가능한 2차 비트스트림으로 공급된다. 두개의 비트스트림이 결합되면, 디코더는 스테레오 출력 신호를 생성한다. 1차 비트스트림은 레벨 값들 이외에도 스테레오 파라미터, 예컨대 폭 파라미터를 포함할 수 있다. 따라서, 이 비트스트림만의 디코딩은 이미 스테레오 출력을 야기하며, 양측 비트스트림이 모두 이용 가능한 경우에 향상된다.

발명의 구성

이하의 첨부 도면을 참조하여 본 발명의 범위 및 사상을 제한하지 않는 예시적인 실시 예를 통하여 본 발명을 설명한다.

(바람직한 실시 예의 설명)

후술하는 실시 예들은 단지 본 발명의 원리의 예시를 위한 것일 뿐이다. 여기에 기재된 세부 사항들 및 배치들의 여러 변경 예 및 변형 예들은 이 분야에서 통상의 지식을 가진 자들에게는 명백함을 이해할 수 있을 것이다. 따라서, 실시 예들의 기재 및 설명을 통하여 제시되는 특정한 세부 사항들이 아니라, 첨부한 특허청구범위의 범위에 의해서만 본 발명이 제한되어야 한다. 명백함을 위하여, 이하의 모든 예들은 2 채널 시스템을 가정하지만, 이 분야에서 통상의 지식을 가진 자들에게는 그 방법들이 5.1 채널과 같은 다중 채널 시스템에 적용될 수 있음이 명백할 것이다.

도 1은 모노럴(monaural) 모드에서 인코더 및 디코더가 작동하는, 인코더(107) 및 디코더(115)를 구비하는 임의의 소스 코딩 시스템이 본 발명에 따른 파라미터 스테레오 코딩에 의해 얼마나 향상될 수 있는지를 보여준다. L 및 R은 AD 변환기 (101)에 공급되는 좌측 및 우측 아날로그 입력 신호를 나타낸다고 한다. AD 변환기로부터의 출력은 모노 신호로 변환되고 (105), 이 모노 신호는 인코딩된다(107). 또한, 스테레오 신호는 후술하는 하나 또는 수개의 스테레오 파라미터를 계산하는 파라미터 스테레오 인코더(103)로 보내진다. 이들 파라미터는 다중화기(multiplexer)(109)기에 의해 상기 인코딩된 모노 신호와 결합되어 비트스트림을 형성한다(111). 이 비트스트림은 저장되거나 전송되고, 이어서 디코더측에서 역다중화기(demultiplexer)(113)에 의해 추출된다. 이 모노 신호는 디코딩되고(115), 스테레오 파라미터(117)를 제어 신호로서 이용하는 파라미터 스테레오 디코더(119)에 의해 스테레오 신호로 변환된다. 마지막으로, 스테레오 신호는 아날로그 출력(L' 및 R')을 공급하는 DA 변환기(121)로 보내진다. 도 1에 따른 토플로지(topology)는 덜 복잡한 버전으로 시작되는 이하에 상세하게 기재되는 파라미터 스테레오 코딩 방법의 집합과 공통이다.

본 발명에 따른 스테레오 특성을 파라미터화하는 한 방법은 인코더측에서 원시 신호 스테레오 폭을 결정하는 것이다. 대략적으로 말해서 L과 R간의 고도의 유사성이 Dif의 작은 값으로 계산되고 그 역도 마찬가지이므로, 이 스테레오 폭의 첫 번째 근사는 차 신호 $Dif = L - R$ 이다. 특별한 경우는 $L = R$ 이고 따라서 $Dif = 0$ 인 듀얼 모노이다. 따라서, 이러한 간단한 알고리즘이라도 통상 뉴스 방송과 연관된 모노 입력 신호의 타입을 검출할 수 있으며, 이 경우에 의사 스테레오는 바람직하지 않다. 그러나, 서로 다른 레벨에서 L 및 R로 공급되는 모노 신호는 감지된 폭이 제로인 경우에도 제로 Dif 신호를 야기하지는 않는다. 따라서, 예컨대 상관법을 이용하는 실제로 복잡한 검출기가 필요할 수도 있다. 레벨 독립적인 검출기를 달성하기 위하여, 좌우 차이 또는 상관 관계를 어떤 식으로 나타내는 값이 총 신호 레벨로 정규화 된다는 것을 확인해야 한다. 상술한 검출기가 갖는 문제는 모노 음성이 훨씬 약한 스테레오 신호, 예컨대 음성-음악/음악-음성 전이 중에 스테레오 노이즈 또는 배경 음악과 혼합되는 경우이다. 음성 정지 시에 검출기는 넓은 스테레오 신호를 나타낸다. 이것은 이전의 총 에너지의 정보를 포함하는 신호, 예컨대 총 에너지의 피크 쇠퇴(peak decay) 신호로 상기 스테레오 폭 값을 정규화 함으로써 해결된다. 또한, 스테레오 폭 검출기가 고주파 노이즈 또는 채널 차 고주파 왜곡에 의해 트리거 되는 것을 방지하기 위하여, 검출기 신호들은, 통상 음성의 제2 포르만트(formant) 위의 어딘가의 차단 주파수를 갖는 로우패스(저역 통과) 필터에 의해 또는 하이패스(고역 통과) 필터에 의해 필터링 되어 언밸런스 신호 오프셋이나 잡음을 회피해야 한다. 검출기 탑재에 관계없이, 계산된 스테레오 폭은 모노로부터 넓은 스테레오에 걸친 전 범위를 커버하는 유한 집합의 값들로 매핑된다.

도 2a는 도 1에 소개된 파라미터 스테레오 디코더의 상세한 예를 보여준다. 파라미터 B에 의해 제어되는 "밸런스"로 표시된 블록(211)은 나중에 설명되며, 여기서 "B"는 밸런스 파라미터를 의미한다. "폭"으로 표시된 블록(205)은 모노 입력 신호를 취하며, 스테레오 폭의 흔적을 합성적으로 재생하는데, 여기서 폭의 양은 파라미터 W에 의해 제어된다. 선택적인 롤 오프 필터 파라미터 S 및 지연 파라미터 D는 나중에 설명한다. 본 발명에 따르면, 저주파수 범위를 단단히 변하지 않게 유지하기 위하여, 종종 로우패스 필터(203) 및 하이패스 필터(201)로 구성되는 크로스오버(crossover) 필터를 합체함으로

써 본질적으로 더 나은 음질을 얻을 수 있다. 이로써, 하이패스 필터로부터의 출력만이 폭 블록으로 보내진다. 폭 블록으로 부터의 스테레오 출력은 207 및 209에 의해 로우패스 필터로부터의 모노 출력에 부가되어, 스테레오 출력 신호를 형성한다.

배경 기술 부분에서 언급한 것들 또는 쉬레더(Schroeder) 타입의 조기 반사 시뮬레이션 유닛(멀티탭 지연)이나 잔향기(reverberator)와 같은 임의의 종래 기술의 의사 스테레오 발생기가 상기 폭 블록에 이용될 수 있다.

도 2b는 모노 신호 M에 의해 공급된 의사 스테레오 발생기의 예를 제시한다. 스테레오 폭의 양은 215의 이득에 의해 결정되며, 이 이득은 스테레오 폭 파라미터 W의 함수이다. 이득이 더 높을수록, 스테레오 흔적은 더 넓어지며, 제로 이득은 순수 모노 재생에 대응한다. 215로부터의 출력은 지연되고(221), 반대 부호를 이용하여 두개의 다이렉트 신호 인스턴스(direct signal instance)에 부가된다(223 및 225). 스테레오 폭을 변경할 때 전체적인 재생 레벨을 유효하게 바꾸지 않기 위해서, 다이렉트 신호의 보상 감쇠가 병합될 수 있다(213). 예컨대, 지연 신호의 이득이 G라면, 다이렉트 신호의 이득은 $\sqrt{1-G^2}$ 로서 선택될 수 있다. 본 발명에 따르면, 고주파 롤오프(roll-off)가 지연 신호 경로(217)내에 합체되어, 코딩 인공음을 차단하지 못하게 하는 의사 스테레오의 회피를 도울 수 있다. 선택적으로, 크로스오버 필터, 롤오프 필터 및 지연 파라미터들이 비트스트림 내에 보내져서, 도 2a 및 도 2b에 신호 X, S 및 D로서 나타낸 바와 같이 원시 신호의 스테레오 특성을 모방할 더 많은 가능성을 제공할 수 있다. 여기서 S는 롤오프 필터 파라미터이고, D는 지연 파라미터이다. X는 크로스오버 필터 파라미터 인데, 도 2a에서 파라미터 X가 하이패스 필터(201)와 로우패스 필터(203)을 제어한다는 것을 보여주고 있다. 물론, 이것은 이 크로스오버 필터 파라미터가 로우패스 필터(203)의 컷오프 주파수와 하이패스 필터(201)의 컷오프 주파수를 결정한다는 의미이다. 따라서, X는 더 높은 컷오프 주파수로 상승될 때, 로우패스 필터(203)에 의한 신호 출력은 더 높은 주파수 요소들을 포함한다. 그리고 하이패스 필터(201)의 신호 출력은 더 높은 주파수 요소들과 함께 출발한다는 것이 명백해진다. 이것은 크로스오버 필터 파라미터 X가 더 작았을 경우와 비교하여 볼 때 그러하다는 것이다. 잔향기(reverberator)가 스테레오 신호를 생성하는데 이용된다면, 때때로 사운드의 끝 직후에는 잔향 쇠퇴가 원해지지 않을 수도 있다. 그러나, 이들 원치 않는 잔향 꼬리(reverb-tails)는 단지 잔향 신호의 이득을 바꿈으로써 쉽게 감쇠되거나 완전히 제거될 수 있다. 사운드 끝을 찾아내기 위하여 설계된 검출기가 이러한 목적에 이용될 수 있다. 잔향기가 일부 특정 신호에서, 예컨대 과도 신호에서 인공음을 발생한다면, 이들 신호에 대한 검출기 또한 그것을 감쇠 시키는데 이용될 수 있다.

본 발명에 따른 스테레오 특성을 검출하는 선택적인 방법을 이하에 기술한다. 또, L 및 R이 좌측 및 우측 입력 신호를 나타내고, P_L 은 왼쪽 채널 L의 신호전력, P_R 은 오른쪽 채널 R의 신호전력을 나타낸다. 그러면 대응하는 신호 전력들은 $P_L \sim L^2$ 및 $P_R \sim R^2$ 에 의해 주어진다. 이제, 스테레오 벨런스의 치수가 두 신호 전력의 비율, 더 상세하게는 $B = (P_L + e)/(P_R + e)$ 로서 계산될 수 있으며, 여기서 e는 0으로의 나눗셈을 소거하는 임의의 매우 작은 수이다. 벨런스 파라미터 B는 관계식 $B_{dB} = 10\log_{10}(B)$ 에 의해 주어진 dB로 표현될 수 있다. 예로서, 세 가지 경우 $P_L = 10P_R$, $P_L = P_R$ 및 $P_L = 0.1P_R$ 는 $+10dB$, $0dB$ 및 $-10dB$ 에 각각 대응한다. 명백하게, 이들 값은 "좌", "중앙" 및 "우"의 위치들에 매핑된다. 실험에 따르면, 벨런스 파라미터의 범위는 예컨대 $+/-40dB$ 로 제한될 수 있는데, 이들 극한값들은 이미 사운드가 두개의 라우드스피커(loudspeaker)나 헤드폰 드라이버 중 하나로부터 완전히 생기는 것처럼 감지되기 때문이다. 이러한 제한은 전송 시에 커버할 신호 공간을 감소시켜서 비트레이트 감소를 제공한다. 또한, 진보적인 양자화 기술이 이용될 수 있으며, 이에 의해 제로 주변에서는 보다 작은 양자화 스텝들이 이용되고 외측 한계값들로 갈수록 보다 큰 스텝들이 이용되어, 비트레이트를 더 감소시킨다. 연장된 진행을 위한 시간에 걸쳐서 벨런스는 일정한 경우가 많다. 따라서, 필요한 평균 비트 수를 상당히 감소시키는 최종 조치가 다음과 같이 취해질 수 있다. 초기 벨런스 값의 전송 후에, 연속하는 벨런스 값들간의 차이만이 전송되고, 이에 의해 엔트로피 코딩이 채용된다. 아주 흔히, 이 차이는 제로이며, 따라서 최단 가능 코드어로 나타내어진다. 명백하게, 비트 에러가 가능한 애플리케이션에서, 제어되지 않은 에러 전파를 제거하기 위하여 이 델타 코딩(delta coding)은 적절한 시간 간격으로 리셋 되어야 한다.

상기 벨런스 파라미터의 가장 기본적인 디코더 사용법은, 도 2c에 블록 227 및 229로 나타낸 바와 같이 상기 모노 신호를 양측 출력 모두에 공급하고 제어 신호에 대응하여 그 이득을 조절함으로써, 단순하게 두개의 재생 채널 중 어느 하나를 향해서 상기 모노 신호를 오프셋 하는 것이다. 이것은 "파노라마" 마디를 혼합 테스크상으로 돌리는 것과 유사하며, 두개의 스테레오 스피커 사이에서 모노 신호를 합성적으로 "이동시키는" 것이다.

상술한 폭 파라미터에 더하여 벨런스 파라미터도 전송되어, 사운드 스테이지 내에 사운드 이미지를 제어된 방식으로 위치 결정하고 확산시킬 가능성을 제공하고, 원시 스테레오 흔적을 모방할 때 유연성을 제공한다. 이전의 섹션에서 언급한 바와 같은 의사 스테레오 발생 및 파라미터 제어 벨런스를 결합함에 있어서의 한가지 문제점은 중앙 위치에서 멀리 떨어진 벨런

스 위치에서 의사 스테레오 발생기로부터의 원치 않는 신호 기여이다. 이것은, 스테레오 폭 값에 대하여 모노 촉진 기능을 인가함으로써 해결되며, 이는 극단적 위치에서 밸런스 위치들에서의 스테레오 폭 값의 보다 큰 감쇠 및 중앙 위치에서 가까운 밸런스 위치에서의 더 적은 또는 전혀 없는 감쇠를 초래한다.

지금까지 설명한 방법들은 매우 낮은 비트레이트 애플리케이션용으로 의도된 것이다. 보다 높은 비트레이트가 이용 가능한 애플리케이션에서는, 상기 폭 및 밸런스 방법의 보다 복잡한 버전을 이용할 수 있다. 스테레오 폭 검출은 몇몇의 주파수 대역에서 수행되어, 각 주파수 대역에 대한 개별적인 스테레오 폭 값들을 초래한다. 마찬가지로, 밸런스 계산은 다중 대역 방식으로 동작할 수 있으며, 이는 모노 신호에 의해 공급되는 두개의 채널에 다른 필터 곡선을 적용하는 것과 동일하다.

도 3은 도 2c에 나타낸 바와 같이 블록 309, 319 및 329로 표시된 다중 대역 밸런스 조절과 결합된, 블록 307, 317 및 327로 표시된 도 2b에 따른 N개의 의사 스테레오 발생기를 이용하는 파라미터 스테레오 디코더의 일례를 나타낸다. 개별적인 통과 대역들은 대역 통과 필터들(305, 315 및 325)의 집합에 상기 모노 입력 신호(M)를 공급함으로써 얻어진다. 상기 밸런스 조절기들로부터의 대역 통과 스테레오 출력들이 더해져서(311, 321, 313, 323), 스테레오 출력 신호(L 및 R)를 형성한다. 이전의 스칼라 폭 및 밸런스 파라미터들은 이제 어레이들(W(k) 및 B(k))로 치환된다. 도 3에서, 모든 의사 스테레오 발생기 및 밸런스 조절기는 고유한 스테레오 파라미터를 갖는다. 그러나, 전송되거나 저장되는 총 데이터량을 감소시키기 위하여, 몇 개의 주파수 대역들로부터의 파라미터들은 인코더에서 그룹단위로 평균될 수 있고, 이 보다 적은 파라미터 수는 디코더에서의 폭 및 밸런스 블록들의 대응하는 그룹으로 매핑된다. 명백하게도, 다른 그룹핑(grouping) 기술들 및 길이들이 어레이들(W(k) 및 B(k))에 이용될 수 있다. S(k)는 폭 블록들에서의 지역 신호 경로의 이득을 나타내고, D(k)는 지역 파라미터들을 나타낸다. 게다가, S(k) 및 D(k)는 비트스트림에서는 선택적이다.

상기 파라미터 밸런스 코딩 방법은, 특히 보다 낮은 주파수 대역들에 대하여, 주파수 해상도의 부족에 기인하거나 동일시간이지만 다른 밸런스 위치들에서의 하나의 주파수 대역에서 발생하는 너무 많은 사운드 이벤트들에 기인하여 다수 불안정한 행위를 제공할 수 있다. 이들 밸런스 결함(balance-glitches)은 보통 짧은 시간동안의 비정상적인 밸런스 값, 통상적으로는 개신 레이트에 따라 계산된 하나 또는 수개의 연속하는 값들을 특징으로 한다. 혼란스러운 밸런스 결함들을 회피하기 위하여, 안정화 처리가 밸런스 데이터에 가해질 수 있다. 이 처리는 현재 시간 위치의 전후에 많은 밸런스 값들을 이용하여, 그들의 중간값을 계산할 수도 있다. 중간값은 이어서 현재 밸런스 값에 대한 한계값으로서 이용될 수 있다. 즉, 현재 밸런스 값은 중간값을 넘어 가도록 허용되지 않아야 한다. 현재 값은 최종 값과 중간값 사이의 범위에 의해 한정된다. 선택적으로, 현재 밸런스 값은 소정의 오버슛 요인(certain overshoot factor)에 의해 한계값을 통과하도록 허용될 수 있다. 또한, 중간값을 계산하는데 이용된 밸런스 값들의 수뿐만 아니라 오버슛 요인도 주파수 종속적인 특성으로서 보여져야 하고, 따라서 각 주파수 대역에 대하여 개별적이어야 한다.

밸런스 정보의 낮은 개신 속도에서, 시간 해상도의 부족은 스테레오 이미지의 모션들과 실제 사운드 이벤트들간의 동기화 시에 실패를 초래할 수 있다. 동기화에 의한 이러한 활동을 향상시키기 위하여, 사운드 이벤트들을 식별하는 것에 기초한 보간 기술이 이용될 수 있다. 여기에서 보간은 두개의 시간적으로 연속적인 밸런스 값들간의 보간을 말한다. 수신기 측에서의 모노 신호를 연구함으로써, 서로 다른 사운드 이벤트들의 시작 및 종료에 대한 정보를 얻을 수 있다. 한가지 방법은 특정한 주파수 대역에서의 신호 에너지의 갑작스런 증가 및 감소를 검출하는 것이다. 이 보간은 적절한 에너지 포락선으로 부터의 안내 후에 밸런스 위치의 변경이 작은 신호 에너지를 포함하는 시간 세그먼트동안에 바람직하게 수행되어야 함을 확실히 하여야 한다. 인간의 귀는 사운드의 꼬리 부분보다는 시작부분에 더 민감하기 때문에, 상기 보간 기술은 예컨대 피크 홀드(peak-hold)를 에너지에 인가함으로써 사운드의 시작을 찾아내는 것으로부터 이익을 얻게 되고, 상기 밸런스 값 증분들은 피크 홀드 에너지의 합수로 하며, 여기서 작은 에너지 값이 큰 증분을 부여하고, 그 역도 성립한다. 시간적으로 균일하게 분포된 에너지를 포함하는 시간 세그먼트에 대하여, 즉 어떤 고정 신호들에 대하여, 이 보간 방법은 두개의 밸런스 값들 사이의 선형 보간과 동일하다. 밸런스 값들이 좌측 및 우측 에너지의 비율이라면, 좌우 대칭 요인들에 대하여는 대수적인 밸런스 값들이 바람직하다. 대수적인 도메인에서 전체 보간 알고리즘을 인가하는 또 다른 이점은 인간의 귀가 대수적인 스케일에 대하여 레벨을 관련시키려는 경향이다.

또한, 스테레오 폭 이득 값들의 낮은 개신 비율에 대하여, 보간이 동일하게 적용될 수 있다. 간단한 방법은 두개의 시간적으로 연속하는 스테레오 폭 값들 사이를 선형적으로 보간하는 것이다. 스테레오 폭의 보다 안정한 동작은 몇몇의 스테레오 폭 파라미터를 포함하고 있는 보다 긴 시간 세그먼트에 걸쳐서 스테레오 폭 이득 값들을 평활하게 함으로써 달성될 수 있다. 서로 다른 개시 및 해제 시간 상수에 의한 평활화를 활용함으로써, 혼합되거나 삽입된 음성 및 음악을 포함하는 프로그램 자료에 매우 적합한 시스템이 얻어진다. 그러한 평활 필터의 적절한 설계가 짧은 개시 시간 상수를 이용하여 행해져서, 짧은 상승 시간을 얻고, 따라서 스테레오 내의 음악 엔트리에 대한 즉각적인 응답을 얻고, 긴 해제 시간을 이용하여 긴 하강 시간을 얻는다. 갑작스런 음성 엔트리에 대하여 바람직 할 수 있는 넓은 스테레오 모드로부터 모노 모드로 빠르게 스위칭 할 수 있기 위하여, 이 이벤트를 나타냄으로써 평활 필터를 우회하거나 리셋 할 가능성이 있다. 또한, 개시 시간 상수들, 해제 시간 상수들 및 다른 평활 필터 특성들도 인코더에 의해 신호로 나타내어 질 수 있다.

심리 음향 코덱으로부터의 차단된 왜곡을 포함하고 있는 신호들에 대하여, 코딩된 모노 신호에 기반하고 있는 스테레오를 도입함에 따른 한가지 공통적인 문제는 상기 왜곡의 미차단 효과이다. 보통 "스테레오 언마스킹(stereo unmasking)"이라 불리는 이러한 현상은 마스킹(차단) 기준을 이행하지 않는 비중심적인 사운드들의 결과이다. 스테레오 언마스킹에 따른 문제는 디코더측에서 그러한 상황을 목표로 하는 검출기를 도입함으로써 해결되거나 부분적으로 해결될 수 있다. 신호대 마스크 비율을 측정하는 공지된 기술을 이용하여 잠재적인 스테레오 언마스킹을 검출할 수 있다. 일단 검출되면, 분명하게 나타내어지거나 스테레오 파라미터들은 단순히 감소될 수 있다.

인코더측에서, 본 발명에 의해 교시되는 바와 같이 한가지 옵션은 입력 신호에 대하여 힐버트 변환기(Hilbert transformer)를 채용하는 것이다. 즉, 두개의 채널간의 90도 위상 시프트가 도입된다. 이어서 두개의 신호의 더함에 의하여 모노 신호를 형성하는 경우에, 힐버트 변환은 중심 정보에 대하여 3dB 감쇠를 도입하기 때문에, 중심이 팬된(center-panned) 모노 신호와 "진정한(true)" 스테레오 신호들간의 보다 나은 밸런스가 달성된다. 실제로, 이것은 예컨대, 하나의 모노 소스를 이용하여 리드 보컬과 베이스 기타를 기록하는 당 시대의 팝음악의 모노 코딩을 향상시킨다.

다중 대역 밸런스 파라미터 방법은 도 1에 기재된 애플리케이션 타입에 한정되는 것은 아니다. 목적이 스테레오 신호의 전력 스펙트럼 포락선을 효율적으로 인코딩하는 것인 경우에는 언제나 유리하게 이용될 수 있다. 따라서, 상기 방법은 스테레오 스펙트럼 포락선에 이외에도 대응하는 스테레오 나머지가 코딩되는 스테레오 코덱에서 툴로서 이용될 수 있다. 레벨 파라미터인 총전력 P 가 $P = P_L + P_R$ 로 정의되며, 여기서 P_L 및 P_R 은 상술한 바와 같은 신호 전력들이라고 한다. 이 정의는 좌측 내지 우측 위상 관계를 고려하지 않음에 유의해야 한다(예컨대, 동일하지만 반대 부호의 좌측 및 우측 신호들은 제로 총전력을 초래하지 않는다). B 와 유사하게, P 는 $P_{dB} = 10\log_{10}(P/P_{ref})$ dB로 표현될 수 있으며, 여기서 P_{ref} 는 임의의 기준 전력이고, 델타 값들은 엔트로피 코딩된다. 밸런스 경우와는 반대로, P 에 대하여는 어떠한 진보적인 양자화도 채용되지 않는다. 스테레오 신호의 스펙트럼 포락선을 표현하기 위하여, P 및 B 가 인간의 청력의 임계 대역에 관련되어 있는 대역폭을 갖는 주파수 대역들의 집합에 대하여 계산되지만, 통상적으로 반드시 그런 것은 아니다. 예컨대, 이들 대역은 인정한 대역 폭 필터 뱅크에서 채널을 그룹핑함으로써 형성될 수 있으며, 여기서 P_L 및 P_R 은 각각의 대역 및 주기에 대응하는 하위대역 샘플들의 제곱의 시간 및 주파수 평균들로서 계산된다. 집합들 $P_0, P_1, P_2, \dots, P_{N-1}$ 및 $B_0, B_1, B_2, \dots, B_{N-1}$ 은 델타 및 호프만(Huffman) 코딩되고, 전송되거나 저장되고, 마지막으로 인코더에서 계산된 양자화된 값들로 디코딩되는데, 여기서 첨자는 N 대역 표현에서의 주파수 대역을 나타낸다. 최종 단계는 P 및 B 를 P_L 및 P_R 로 역변환하는 것이다. P 및 B 의 정의로부터 쉽게 알 수 있는 바와 같이, 역관계들은 (B 의 정의에서 e 를 무시한 경우) $P_L = BP/(B+1)$ 및 $P_R = P/(B+1)$ 이다.

상기 포락선 코딩 방법의 한가지 특히 흥미 있는 애플리케이션은 HFR 기반 코덱에 대한 고대역 스펙트럼 포락선의 코딩이다. 이 경우에, 고대역 잔류 신호는 전송되지 않는다. 대신에 이 잔류는 저대역으로부터 유도된다. 따라서, 잔류 및 포락선 표현간에는 엄격한 관계가 없으며, 포락선 양자화가 더 중요하다. 양자화의 효과를 연구하기 위하여, Pq 및 Bq 가 P 및 B 의 양자화된 값들을 나타낸다고 한다. Pq 및 Bq 는 다음에 상기 관계식에 삽입되고, 그 합은 다음과 같이 얻어진다.

$$P_Lq + P_Rq = BqPq/(Bq+1) + Pq/(Bq+1) = Pq(Bq+1)/(Bq+1) = Pq$$

여기서 흥미 있는 특징은 Bq 가 소거되고 총전력에서의 P 에서의 양자화 에러에 의해 단독으로 결정된다는 것이다. 이것은, P 의 양자화에 충분한 정밀도가 이용된다고 가정하면, B 가 크게 양자화 되더라도 감지된 레벨이 올바르다는 것을 내포하고 있다. 바꿔 말하면, B 에서의 왜곡이 레벨 상이라기 보다는 공간상에서의 왜곡에 매핑된다. 음원이 시간적으로 공간에 걸쳐서 고정되어 있는 한, 스테레오 감지에 있어서의 이러한 왜곡도 또한 고정적이며, 주의하기 어려운 것이다. 전술한 바와 같이, 중심선에 대한 각도가 큰 경우에 인간의 청력 특성으로 인하여 주어진 dB 에러가 감지된 각도에서의 더 작은 에러에 대응하기 때문에 스테레오 밸런스의 양자화도 외측 극단을 향할수록 더 거칠어질(coarser) 수 있다.

주파수 종속적인 데이터, 예컨대 다중 대역 스테레오 폭 이득 값들이나 다중 대역 밸런스 값들을 양자화할 때, 양자화 방법의 해상도 및 범위는 감지할 수 있는 스케일의 특성들을 일치시키도록 유리하게 선택될 수 있다. 이러한 스케일이 주파수 종속적으로 되면, 다른 양자화 방법들이나 소위 양자화 등급들이 다른 주파수 대역들에 대하여 선택될 수 있다. 다른 주파수 대역들을 나타내는 인코딩된 파라미터 값들은 어떤 경우에는 동일한 값을 가지더라도 다른 방식으로 해석, 즉 다른 값들로 디코딩되어야 한다.

스위칭된 L/R 내지 S/D 코딩 기술과 유사하게, 극단적인 신호들에 더 잘 대처하기 위하여 P 및 B 신호들이 P_L 및 P_R 로 적합하게 치환될 수 있다. PCT/SE00/00158에 의해 교시된 바와 같이, 포락선 샘플들의 델타 코딩은, 어느 방향이 특정한 모

엔트에서의 비트수에 의하여 가장 효율적인지에 따라서, 델타-인-타임(delta-in-time)으로부터 델타-인-주파수(delta-in-frequency)로 스위칭될 수 있다. 밸런스 파라미터도 이 기술을 이용할 수 있다. 즉, 예컨대 시간에 대하여 스테레오 필드에서 이동하는 소스를 고려하는 것이다. 명백하게도, 이것은 시간에 대한 밸런스 값들의 연속적인 변화에 대응하며, 이는 파라미터들의 갱신 속도에 대한 소스의 속도에 따라 엔트로피 코딩을 채용하는 경우의 큰 코드어에 대응하는 큰 델타-인-타임 값들에 대응할 수도 있다. 그러나, 소스가 주파수에 대한 균일한 사운드 방사를 갖는다고 하면, 밸런스 파라미터의 델타-인-주파수 값들이 작은 코드어에 또 대응하는 모든 시점에서 제로이다. 따라서, 이 경우에 주파수 델타 코딩 방향을 이용할 때 보다 낮은 비트레이트가 달성된다. 또 다른 예는 실내에서 고정되어 있지만 균일하지 않은 방사를 갖는 소스이다. 이제 델타-인-주파수 값들은 크므로, 델타-인-타임이 바람직한 선택이다.

P/B 코딩 기술은 도 4에 나타낸 바와 같이 스케일러블 HFR 코덱을 구성할 가능성을 제공한다. 스케일러블 코덱은 둘 이상의 부분들로 분할되는 것을 특징으로 하고 있으며, 여기서 더 고차 부분의 수신 및 디코딩은 선택적이다. 이 예는, 이하에서 1차(419) 및 2차(417)라고 불리는 두 개의 비트스트림 부분을 가정하지만, 고 많은 부분들로의 확장도 분명히 가능하다. 도 4a에서 인코더측은 스테레오 입력 신호(IN)에 대하여 작용하는 임의의 스테레오 저대역 코어 인코더(403) (AD 및 DA 변환 각각의 세부적인 단계들은 도시되어 있지 않다), 고대역 스펙트럼 포락선 및 상기 스테레오 입력 신호에 작용하는 선택적으로 추가되는 스테레오 파라미터들(401)을 평가하는 파라미터 스테레오 인코더, 및 1차 및 2차 비트스트림 각각을 위한 두 개의 다중화기(415, 413)를 구비하고 있다. 이 애플리케이션에서, 고대역 포락선 코딩은 P/B 연산에 한정되고, P 신호(407)가 415에 의해 1차 비트스트림으로 보내지는 한편, B 신호(405)는 413에 의해 2차 비트스트림으로 보내진다.

저대역 코덱에 대해서 다른 가능성들이 존재한다. 즉, S/D 모드에서 일정하게 작동할 수도 있으며, S 및 D 신호들은 각각 1차 및 2차 비트스트림으로 보내질 수 있다. 이 경우에, 1차 비트스트림의 디코딩은 전체 대역 모노 신호를 초래한다. 물론, 이 모노 신호는 본 발명에 따른 파라미터 스테레오 방법들에 의해 확대될 수 있으며, 이 경우에 스테레오 파라미터(들)도 1차 비트스트림에 위치해야 한다. 또 다른 가능성은 선택적으로 고대역 폭 및 밸런스 파라미터와 함께 스테레오 코딩된 저대역 신호를 1차 비트스트림으로 공급하는 것이다. 이제 1차 비트스트림의 디코딩은 저대역에 대해서는 진정한 스테레오 및 고대역에 대해서는 매우 사실적인 의사 스테레오를 초래하는데, 이는 저대역의 스테레오 특성이 고주파 재생에 반영되기 때문이다. 또 다른 방식에 대하여 말하자면, 이용 가능한 고대역 포락선 표현이나 스펙트럼 거친 구조가 모노이더라도, 합성된 고대역 잔류 또는 스펙트럼 미세 구조는 아니다. 이러한 타입의 구현 예에서는, 2차 비트스트림은 더 저대역 정보를 포함할 수도 있으며, 이는 1차 비트스트림과 결합되는 경우에 고품질의 저대역 재생을 초래한다. 415 및 417에 각각 접속된 1차 및 2차 저대역 코어 인코더 출력 신호(411, 409)가 상술한 신호 타입들의 어느 하나를 포함할 수 있기 때문에, 도 4의 토플로지는 양측 경우를 모두 나타내고 있다.

비트스트림들이 전송되거나 저장되고, 419만이나 419와 417 모두의 어느 한쪽이 도 4b의 디코더에 공급된다. 1차 비트스트림은 423에 의해 저대역 코어 디코더 1차 신호(429) 및 P 신호(431)로 역다중화된다. 마찬가지로, 2차 비트스트림은 421에 의해 저대역 코어 디코더 2차 신호(427) 및 B 신호(425)로 역다중화된다. 저대역 신호(들)가 출력(435)을 생성하는 저대역 코어 디코더(433)로 보내지며, 이 출력은 다시 1차 비트스트림만을 디코딩하는 경우에 상술한 타입 중 어느 하나이다(모노 또는 스테레오). 신호(435)는 HFR 유닛(437)에 공급되는데, 여기에서 합성 고대역이 생성되고 역시 HFR 유닛에 접속된 P에 따라 조절된다. 디코딩된 저대역은 HFR 유닛에서 고대역과 결합되며, 저대역 및/또는 고대역은, 최종적으로 시스템 출력으로 공급되기 전에 의사 스테레오 발생기(HFR 유닛에 설치되어 있음)에 의해 선택적으로 확대되어, 출력 신호(OUT)를 형성한다. 2차 비트스트림(417)이 존재하는 경우에, HFR 유닛도 또한 입력 신호(425)로서 B 신호를 갖고 435는 스테레오이며, 이에 의해 시스템은 풀 스테레오 출력 신호를 생성하고, 의사 스테레오 발생기는 있다면 우회된다.

다시 말해서, 입력신호의 스테레오 특성에 대한 코딩 방법은, 인코더측에서 상기 입력신호의 스테레오-폭을 나타내는 폭-파라미터를 계산하는 단계와, 디코더측에서 스테레오 출력신호를 발생시키고 상기 폭-파라미터를 이용하여 상기 출력신호의 스테레오-폭을 제어하는 단계를 포함한다. 본 방법은 상기 인코더에서 상기 입력신호로부터 모노신호를 형성하는 단계를 더 포함하는데, 여기에서 상기 디코더에서의 상기 발생은 상기 모노신호에서 의사-스테레오 방식 동작을 포함한다. 본 방법은 상기 폭-파라미터에 의해 제어된 레벨(들)에서, 상기 모노신호를 두 개의 신호로 분할하고 상기 모노신호의 지연 버전을 상기 두 개의 신호에 부가하는 단계를 더 포함한다. 본 방법은 상기 지연 버전이 상기 두 개의 신호에 부가되기 전에 하이패스 필터링되고 더 고주파수에서 점진적으로 감쇠되는 것을 더 포함한다. 본 방법은 상기 폭-파라미터는 벡터이며, 상기 벡터의 성분들은 별개의 주파수 대역들에 대응하는 것을 더 포함한다. 이 방법은 상기 입력 신호가 듀얼 모노 타입인 경우에, 상기 출력 신호도 또한 듀얼 모노 타입인 것을 더 포함한다.

입력 신호의 스테레오 특성을 코딩하는 방법으로서, 인코더측에서, 상기 입력 신호의 스테레오-밸런스를 나타내는 밸런스-파라미터를 계산하고, 디코더측에서, 스테레오 출력 신호를 발생시키고 상기 밸런스 파라미터를 이용하여 상기 출력 신호의 스테레오 밸런스를 제어하는 것을 포함한다.

본 방법에서, 상기 인코더측에서, 상기 입력 신호로부터 모노 신호가 형성되고, 상기 디코더측에서, 상기 발생은 상기 모노 신호를 두개의 신호로 분할하는 것을 포함하고, 상기 제어는 상기 두개의 신호의 레벨을 조정하는 것을 포함한다. 본 방법은 상기 입력 신호의 각 채널에 대한 전력이 계산되고, 상기 밸런스-파라미터가 상기 전력들간의 비율로부터 계산되는 것을 더 포함한다. 본 방법은 상기 전력들 및 상기 밸런스-파라미터는 모든 성분이 특정 주파수 대역에 대응하는 벡터들인 것을 더 포함한다. 본 방법은 상기 디코더측에서, 상기 모노 신호의 대응 전력의 순간 값이 순간 보간을 얼마나 가파르게 행하여져야 하는지를 제어하는 방식으로, 상기 밸런스 파라미터들의 두개의 시간적으로 연속하는 값들 사이에서 보간이 행해지는 것을 더 포함한다. 본 방법은 상기 보간 방법은 대수 값으로서 표현되는 밸런스 값들에 대하여 수행되는 것을 더 포함한다. 본 방법은 상기 밸런스 파라미터 값들은 이전의 밸런스 값 및 중앙 필터나 다른 필터 처리에 의해 다른 밸런스 값들로부터 추출된 밸런스 값 사이의 범위로 제한되고, 소정의 요인에 의해 상기 범위의 경계를 이동함으로써 상기 범위가 더 확장될 수 있는 것을 더 포함한다. 본 방법은 밸런스 값들에 대한 제한 경계를 추출하는 상기 방법은 다중 대역 시스템의 경우에 주파수 종속적인 것을 더 포함한다. 본 방법은 추가적인 레벨 파라미터가 상기 전력들의 벡터 합으로서 계산되고, 상기 디코더로 보내져서, 상기 디코더에 상기 입력 신호의 스펙트럼 포락선 표현을 제공하는 것을 더 포함한다. 본 방법은 상기 레벨-파라미터 및 상기 밸런스-파라미터는 상기 전력들에 의해 적합하게 치환되는 것을 더 포함한다. 본 방법은 상기 스펙트럼 포락선이 디코더에서의 HFR 처리를 제어하는데 이용되는 것을 더 포함한다. 본 방법은 상기 레벨 파라미터는 스케일러블 HFR 기반 스테레오 코덱의 1차 비트스트림으로 공급되고, 상기 밸런스 파라미터는 상기 코덱의 2차 비트스트림으로 공급되는 것을 더 포함한다. 상기 모노 신호 및 상기 폭 파라미터는 상기 1차 비트스트림에 공급된다. 또한, 상기 폭 파라미터는, 중앙 위치로부터 멀리 있는 밸런스 위치에 대하여는 작은 밸런스 값을 부여하는 기능에 의해 처리된다. 상기 밸런스 파라미터의 양자화는, 중앙 위치 주변에서는 더 작은 양자화 스텝을 채용하고 외측 위치들에 대하여는 더 큰 양자화 스텝을 채용하는 것을 더 포함한다. 본 방법은 상기 폭 파라미터들과 상기 밸런스 파라미터들은, 다중 대역 시스템의 경우에 주파수 종속적인 해상도 및 범위에 의한 양자화 방법을 이용하여 양자화되는 것을 더 포함한다. 본 방법은 상기 밸런스 파라미터는 엘타-인-타임 또는 엘타-인-주파수로 적합하게 엘타 코딩되는 것을 더 포함한다. 본 방법은 상기 입력 신호는 상기 모노 신호를 형성하기 전에 힐버트 변환기(Hilbert transformer)를 통과하는 것을 더 포함한다.

파라미터 스테레오 코딩용 장치는 인코더측에서는, 입력 신호의 스테레오 폭을 나타내는 폭 파라미터를 계산하는 수단, 및 상기 입력 신호로부터 모노 신호를 형성하는 수단을 포함하고, 디코더측에서는, 상기 모노 신호로부터 스테레오 출력 신호를 발생시키고 상기 폭 파라미터를 이용하여 상기 출력 신호의 스테레오 폭을 제어하는 수단을 포함한다.

발명의 효과

실시 예들은 단지 본 발명의 원리의 예시를 위한 것일 뿐이다. 여기에 기재된 세부 사항들 및 배치들의 여러 변경 예 및 변형 예들은 이 분야에서 통상의 지식을 가진 자들에게는 명백함을 이해할 수 있을 것이다. 따라서, 실시 예들의 기재 및 설명을 통하여 제시되는 특정한 세부 사항들이 아니라, 첨부한 특허청구범위의 범위에 의해서만 본 발명이 제한되어야 한다. 명백함을 위하여, 이하의 모든 예들은 2 채널 시스템을 가정하지만, 이 분야에서 통상의 지식을 가진 자들에게는 그 방법들이 5.1 채널과 같은 다중 채널 시스템에 적용될 수 있음이 명백할 것이다.

본 발명은 잔향 신호의 이득을 바꿔서 모든 잔향 꼬리를 감쇠하거나 완전히 제거함으로써 그로인한 문제가 해결되게 된다.

도면의 간단한 설명

도 1은 파라미터 스테레오 인코더 모듈에 의해 향상된 인코더 및 파라미터 스테레오 디코더 모듈에 의해 향상된 디코더를 포함하는 소스 코딩 시스템을 나타낸다.

도 2a는 파라미터 스테레오 디코더 모듈의 개략적인 블록도이다.

도 2b는 제어 파라미터 입력을 갖는 의사 스테레오 발생기의 개략적인 블록도이다.

도 2c는 제어 파라미터 입력을 갖는 밸런스 조절기의 개략 블록도이다.

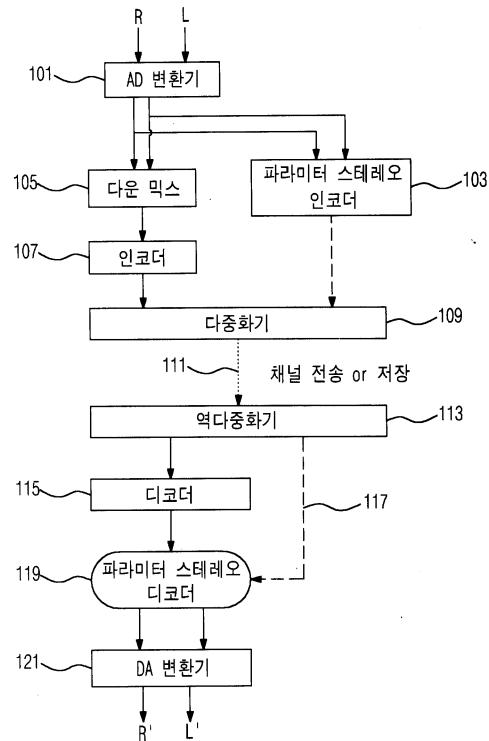
도 3은 다중 대역 밸런스 조절(multiband balance adjustment)과 결합된 다중 대역 의사 스테레오 발생(multiband pseudo-stereo generation)을 이용하는 파라미터 스테레오 디코더 모듈의 개략 블록도이다.

도 4a는 스펙트럼 포락선의 레벨/밸런스 코딩을 채용한 스케일러블 HFR 기반 스테레오 코덱의 인코더측의 개략 블록도이다.

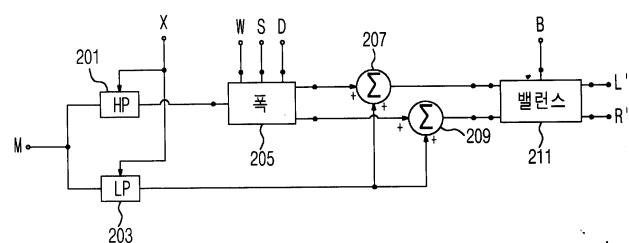
도 4b는 대응하는 디코더측의 개략 블록도이다.

도면

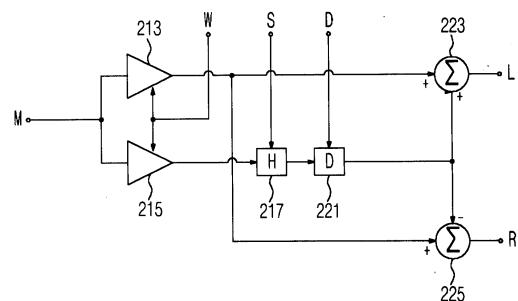
도면1



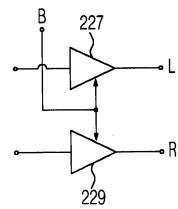
도면2a



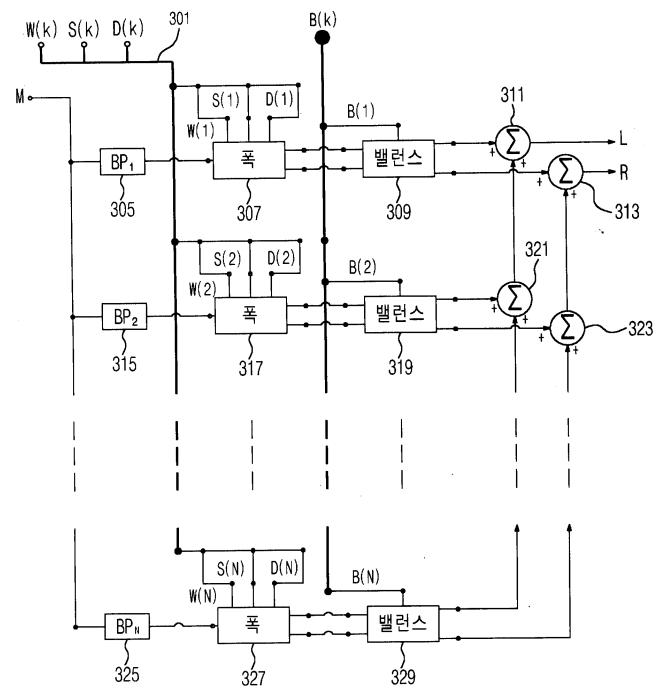
도면2b



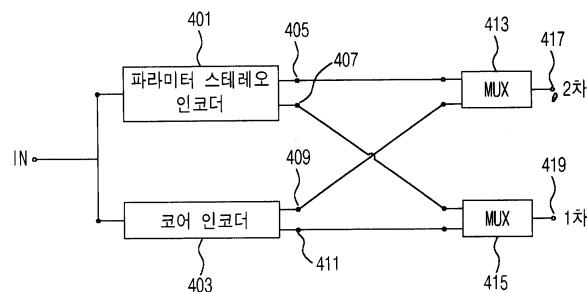
도면2c



도면3



도면4a



도면4b

