



(19)대한민국특허청(KR)
(12) 공개특허공보(A)

(51) Int. Cl.

G10L 19/00 (2006.01)
H04S 3/00 (2006.01)
H04S 3/02 (2006.01)
H04S 5/00 (2006.01)

(11) 공개번호 10-2006-0132682
(43) 공개일자 2006년12월21일

(21) 출원번호 10-2006-7015754

(22) 출원일자 2006년08월03일

심사청구일자 없음

번역문 제출일자 2006년08월03일

(86) 국제출원번호 PCT/US2005/006359

(87) 국제공개번호 WO 2005/086139

국제출원일자 2005년02월28일

국제공개일자 2005년09월15일

(30) 우선권주장 60/549,368 2004년03월01일 미국(US)
60/579,974 2004년06월14일 미국(US)
60/588,256 2004년07월14일 미국(US)

(71) 출원인 돌비 레버러토리즈 라이선싱 코오폰레이션
미국 캘리포니아주 94103-4813 샌프란시스코 포트레로 애비뉴 100

(72) 발명자 데이비스, 마크, 프랭클린
미국, 캘리포니아 94103-4813, 샌 프란시스코, 포트레로 애비뉴100, 돌비 레버러토리즈 라이선싱 코오폰레이션

(74) 대리인 박경재

전체 청구항 수 : 총 62 항

(54) 멀티채널 오디오 코딩

(57) 요약

다중 채널(multiple channel)의 오디오는 단일 음의 합성 신호와 결합하거나, 다중 채널의 오디오가 재구성되는 관련된 보조 정보를 따르는 다중 채널의 오디오와 결합하고, 다중 오디오 채널로부터 혹은 단일 음의 오디오 채널로부터 유도된 다중 오디오 채널의 개선된 상관해제(decorrelation)와, 다중 오디오 채널로 혹은 단일 음의 오디오 신호로 다중 오디오 채널의 개선된 다운믹싱(downmixing)을 포함한다. 상기 기술된 발명의 특징은 오디오 인코더, 디코더, 인코드/디코드 시스템, 다운믹서, 업믹서(upmixer) 및 상관해제기(decorrelator)에서 이용가능하다.

대표도

도 1

특허청구의 범위

청구항 1.

적어도 두 개의 입력 오디오 채널을 수신하는 오디오 인코더에 있어서, 적어도 두 개의 입력 오디오 채널 중 한 세트의 공간 파라미터를 결정하는 단계를 포함하고, 제1 입력 채널의 스펙트럼 성분이 시간상으로 변하는 범위의 측정에 응답하고 다른 입력 채널의 것에 대한 상기 입력 채널의 상기 스펙트럼 성분 중 채널간 위상각의 유사성의 측정에 응답하는 제1 파라미터를 포함하는 상기 세트의 파라미터를 특징으로 하는, 오디오 인코딩 방법.

청구항 2.

제 1항에 있어서, 상기 제1 입력 채널의 스펙트럼 성분이 시간상으로 변하는 범위의 측정이 각각의 스펙트럼 성분의 진폭 혹은 에너지의 변화에 대한 것을 특징으로 하는, 오디오 인코딩 방법.

청구항 3.

제 1항 또는 제 2항에 있어서, 상기 다른 입력 채널의 스펙트럼 성분의 채널간 위상 각에 대해 상기 제1 입력 채널의 상기 스펙트럼 성분의 채널간 위상 각의 유사성의 측정이 상기 입력 채널과 다른 입력 채널 사이의 유령 이미지의 출현에 관련 되는 것을 특징으로 하는, 오디오 인코딩 방법.

청구항 4.

제1항 내지 제 3항 중의 어느 한 항에 있어서, 상기 세트의 파라미터들은 상기 다른 입력 채널에서의 스펙트럼 성분의 위상 각에 대한 상기 제1 입력 채널에서의 스펙트럼 성분의 위상 각에 응답하는 파라미터를 더 포함하는, 오디오 인코딩 방법.

청구항 5.

제 1항 내지 제 4항 중의 어느 한 항에 있어서, 상기 적어도 두 개의 입력 오디오 채널에서 유도된 단일 음의 오디오 신호를 발생하는 단계를 더 포함하는, 오디오 인코딩 방법.

청구항 6.

청구항 4항 기반의 제 5항에 있어서, 상기 단일 음의 오디오 신호는 상기 제1 파라미터와 상기 진일보 파라미터에 응답하여 상기 적어도 두 개 입력 오디오 채널 중 적어도 하나를 변경하는 과정에 의해 상기 적어도 두 개의 입력 오디오 채널에서 유도되는, 오디오 인코딩 방법.

청구항 7.

제 6항에 있어서, 상기 변경하는 단계는 적어도 두 개 입력 오디오 채널 중 상기 적어도 하나의 스펙트럼 성분의 위상 각을 변경하는, 오디오 인코딩 방법.

청구항 8.

제 5항, 제 6항 또는 제 7항에 있어서, 상기 단일 음의 오디오 신호와 상기 세트의 공간 파라미터를 나타내는 인코딩된 신호 혹은 신호들을 생성하는 단계를 더 포함하는, 오디오 인코딩 방법.

청구항 9.

제 1항 내지 4항 중의 어느 한 항에 있어서, 상기 적어도 두 개 입력 오디오 채널로부터 유도된 여러 오디오 신호들을 발생 하는 단계를 더 포함하는, 오디오 인코딩 방법.

청구항 10.

제 9항에 있어서, 상기 적어도 두 개의 입력 오디오 채널을 능동적으로나 수동적으로 매트릭스(matrix)하는 과정에 의해 상기 여러 오디오 신호가 상기 적어도 두 개의 입력 오디오 채널로부터 유도되는, 오디오 인코딩 방법.

청구항 11.

청구항 4항 기반의 제 9항 또는 제 10항에 있어서, 상기 제1 파라미터와 상기 진일보 파라미터에 응답하는 상기 적어도 두 개의 입력 오디오 채널 중 적어도 하나를 변경하는 과정에 의해 상기 여러 오디오 신호가 상기 적어도 두 개의 입력 오디오 채널에서 유도되는, 오디오 인코딩 방법.

청구항 12.

제 11항에 있어서, 상기 변경하는 단계는 상기 적어도 두 개의 입력 오디오 채널 중에 적어도 하나의 스펙트럼 성분의 위상 각을 변경하는, 오디오 인코딩 방법.

청구항 13.

제 10항, 제 11항 또는 제 12항에 있어서, 상기 세트의 공간 파라미터와 여러 오디오 신호를 나타내는 인코딩된 신호 혹은 신호들을 발생하는 단계를 더 포함하는, 오디오 인코딩 방법.

청구항 14.

제 1항 내지 제 13항 중의 어느 한 항에 있어서, 상기 세트의 파라미터는 상기 제1 입력 채널에서의 과도(transient)의 발생에 응답하는 파라미터를 포함하는, 오디오 인코딩 방법.

청구항 15.

제 1항 내지 제 14항 중의 어느 한 항에 있어서, 상기 세트의 파라미터는 상기 제1 입력 채널의 진폭이나 에너지에 응답하는 파라미터를 더 포함하는, 오디오 인코딩 방법.

청구항 16.

제 1항 내지 제 15항 중의 어느 한 항에 있어서, 상기 입력 채널의 스펙트럼 성분이 시간상으로 변하는 범위의 측정이 상기 제1 입력 채널의 주파수 대역에서 스펙트럼 성분에 대해 있고, 상기 다른 입력 채널의 상기 스펙트럼 성분의 채널간 위상 각에 대한 상기 제1 입력 채널의 상기 스펙트럼 성분의 채널간 위상 각의 유사성의 측정이, 상기 다른 입력 채널의 해당 주파수 대역에서 스펙트럼 성분에 대한 상기 제1 입력 채널의 상기 주파수 대역에서의 스펙트럼 성분에 대해 있는 것을 특징으로 하는, 오디오 인코딩 방법.

청구항 17.

적어도 두 개의 입력 오디오 채널을 수신하는 오디오 인코더에 있어서, 적어도 두 개의 입력 오디오 채널 중 한 세트의 공간 파라미터를 결정하는 단계를 포함하고, 상기 세트의 파라미터가 상기 제1 입력 채널에서 과도의 발생에 응답하는 제1 파라미터를 포함하는, 오디오 인코딩 방법.

청구항 18.

하나 이상의 다른 오디오 신호에 대한 오디오 신호를 상관 해제하는 방법에 있어서, 상기 오디오 신호는 복수의 주파수대역으로 분할되고, 각각의 대역은 하나 이상의 스펙트럼 성분을 포함하고,

제1 모드의 동작과 제2 모드의 동작에 따라 적어도 부분적으로 오디오 신호에서 스펙트럼 성분의 위상 각을 시프트(shift)하는 단계를 포함하는, 오디오 신호 상관해제 방법.

청구항 19.

제 18항에 있어서, 상기 제1 모드의 동작에 따른 오디오 신호에서의 스펙트럼 성분의 위상 각의 시프트는 제1 주파수 분석과 제1 시간 분석에 따른 오디오 신호에서의 스펙트럼 성분의 위상 각의 시프트를 포함하고, 제2 모드의 동작에 따른 오디오 신호에서의 스펙트럼 성분의 위상 각의 시프트는 제2 주파수 분석과 제2 시간 분석에 따른 오디오 신호에서의 스펙트럼 성분의 위상 각의 시프트를 포함하는, 오디오 신호 상관해제방법.

청구항 20.

제 19항에 있어서, 상기 제2 시간 분석이 제1 주파수 분석보다 더 양호한, 오디오 신호 상관해제 방법.

청구항 21.

제 19항에 있어서, 제2 주파수 분석은 제1 주파수 분석과 같거나 더 조잡하고, 제2 시간 분석은 제1 주파수 분석보다 더 양호한, 오디오 신호 상관해제 방법.

청구항 22.

제 18항 내지 제 21항 중의 어느 한 항에 있어서, 상기 제1 모드의 동작은 적어도 하나 이상의 복수의 주파수 대역에서 스펙트럼 성분의 위상 각을 시프트하는 단계를 포함하고, 각각의 스펙트럼 성분은 다른 각도로 시프트되고, 상기 각은 실제로 시간에 따라 변하고, 상기 제2 모드의 동작은 복수의 주파수 대역 중에 적어도 하나 이상의 대역에서 모든 스펙트럼 성분의 위상 각을 동일 각만큼 시프트하고, 위상 각이 시프트되는 각각의 주파수 대역에 다른 위상 각 시프트가 인가되는, 오디오 신호 상관해제 방법.

청구항 23.

제 22항에 있어서, 상기 제2 모드의 동작에서 주파수 대역 내에 스펙트럼 성분의 위상 각이, 주파수 대역 경계상에 스펙트럼 성분마다 위상 변화를 줄이도록 보간되는, 오디오 신호 상관해제 방법.

청구항 24.

제 18항에 있어서, 상기 제1 모드의 동작은 적어도 하나 이상의 복수의 주파수 대역에서 스펙트럼 성분의 위상 각을 시프트하는 단계를 포함하고, 각각의 스펙트럼 성분은 다른 각도로 시프트되고, 상기 각은 실제로 시간에 따라 변하고, 상기 제 2 모드의 동작은 스펙트럼 성분의 위상 각이 시프트되지 않는, 오디오 신호 상관해제 방법.

청구항 25.

제 18항 내지 24항 중의 어느 한 항에 있어서, 상기 시프트는 랜덤 화 된 시프트를 포함하는, 오디오 신호 상관해제 방법.

청구항 26.

제 18항 내지 25항 중의 어느 한 항에 있어서, 상기 랜덤 화 된 시프트의 양은 제어 가능한, 오디오 신호 상관해제 방법.

청구항 27.

제 18항 내지 26항 중의 어느 한 항에 있어서, 상기 모드의 작동은 상기 오디오 신호에 응답하는, 오디오 신호 상관해제 방법.

청구항 28.

제 27항에 있어서, 상기 모드의 동작은 상기 오디오 신호에서 과도의 출현에 응답하는, 오디오 신호 상관해제 방법.

청구항 29.

제 18항 내지 26항 중의 어느 한 항에 있어서, 상기 모드의 동작은 제어 신호에 응답하는, 오디오 신호 상관해제 방법.

청구항 30.

제 29항에 있어서, 상기 제어 신호는 오디오 신호에서 과도의 출현에 응답하는, 오디오 신호 상관해제 방법.

청구항 31.

제 18항 내지 30항 중의 어느 한 항에 있어서, 상기 오디오 신호에서 스펙트럼 성분의 크기를 시프트하는 단계를 더 포함하는, 오디오 신호 상관해제 방법.

청구항 32.

제 31항에 있어서, 상기 오디오 신호에서 스펙트럼 성분의 크기의 시프트는 제1 모드의 동작과 제2 모드의 동작에 따르는, 오디오 신호 상관해제 방법.

청구항 33.

제 32항에 있어서, 상기 모드의 동작은 상기 오디오 신호에 응답하는, 오디오 신호 상관해제 방법.

청구항 34.

제 33항에 있어서, 상기 모드의 동작은 상기 오디오 신호에서 과도 출현에 응답하는, 오디오 신호 상관해제 방법.

청구항 35.

제 14항에 있어서, 상기 모드의 동작은 제어 신호에 응답하는, 오디오 인코딩 방법.

청구항 36.

제35항에 있어서, 상기 제어 신호는 오디오 신호에서 과도의 출현에 응답하는, 오디오 인코딩 방법.

청구항 37.

제 30항 내지 36항 중의 어느 한 항에 있어서, 상기 크기의 시프트는 랜덤 화 된 시프트인, 오디오 인코딩 방법.

청구항 38.

제 37항에 있어서, 상기 크기의 시프트의 양은 제어 가능한, 오디오 인코딩 방법.

청구항 39.

M은 1 이상이고 N은 2 이상인 N 오디오 채널을 나타내는 M인코딩된 오디오 채널을 수신하고, N 오디오 채널에 관련된 한 세트의 공간 파라미터를 수신하는 오디오 디코더에 있어서,

상기 M 오디오 채널에서 N 오디오 채널을 유도하는 단계와, 각각의 오디오 채널에서 오디오 신호가 다수의 주파수 대역으로 분할되고, 각각의 대역이 하나 이상의 스펙트럼 성분을 포함하고, 및

상기 공간 파라미터 중 하나에 응답하여 상기 N 오디오 채널 중 적어도 하나의 오디오 신호에서의 스펙트럼 성분의 위상 각을 시프트하는 단계를 포함하고, 상기 시프트는 적어도 부분적으로 제1 모드의 동작과 제2 모드의 동작에 따르는, 오디오 디코딩 방법.

청구항 40.

제 39항에 있어서, 상기 M 오디오 채널을 능동적으로 혹은 수동적으로 디매트릭스하는 과정에 의해 상기 M 오디오 채널에서 상기 N 오디오 채널이 유도되는, 오디오 디코딩 방법.

청구항 41.

제 39항에 있어서, 상기 M은 2 이상이고, 상기 M 오디오 채널을 능동적으로 디매트릭스하는 과정에 의해 상기 M 오디오 채널에서 상기 N 오디오 채널이 유도되는, 오디오 디코딩 방법.

청구항 42.

제41항에 있어서, 상기 디매트릭스(dematrixing)는 상기 M 오디오 채널의 특성에 응답하여 적어도 부분적으로 동작하는, 오디오 디코딩 방법.

청구항 43.

제 41항 또는 제 42항에 있어서, 상기 디매트릭스는 상기 공간 파라미터 중 하나에 응답하여 적어도 부분적으로 동작하는, 오디오 디코딩 방법.

청구항 44.

제 39항에 있어서, 상기 제1 모드의 동작에 따른 오디오 신호에서의 스펙트럼 성분의 위상 각의 시프트는 제1 주파수 분석과 제1 시간 분석에 따른 오디오 신호에서의 스펙트럼 성분의 위상 각의 시프트를 포함하고, 제2 모드의 동작에 따른 오디오 신호에서의 스펙트럼 성분의 위상 각의 시프트는 제2 주파수 분석과 제2 시간 분석에 따른 오디오 신호에서의 스펙트럼 성분의 위상 각의 시프트를 포함하는, 오디오 디코딩 방법.

청구항 45.

제 44항에 있어서, 제2 시간 분석은 제1 시간 분석보다 더 양호한, 오디오 디코딩 방법.

청구항 46.

제 44항에 있어서, 상기 제2 주파수 분석은 제1 주파수 분석과 같거나 더 조잡하고, 제2 시간 분석은 제1 시간 분석보다 더 양호한 오디오 디코딩 방법.

청구항 47.

제 45항에 있어서, 제1 주파수 분석은 공간 파라미터의 주파수 분석보다 더 양호한, 오디오 디코딩 방법.

청구항 48.

제 46 또는 제 47항에 있어서, 상기 제2 시간 분석은 상기 공간 파라미터의 시간 분석보다 더 양호한, 오디오 디코딩 방법.

청구항 49.

제 39항 내지 제 48항 중의 어느 한 항에 있어서, 상기 제1 모드의 동작은 적어도 하나 이상의 복수의 주파수 대역에서 스펙트럼 성분의 위상 각을 시프트하는 단계를 포함하고, 각각의 스펙트럼 성분은 다른 각도로 시프트되고, 상기 각은 실제로 시간에 따라 변하고, 상기 제2 모드의 동작은 복수의 주파수 대역 중에 적어도 하나 이상의 대역에서 모든 스펙트럼 성분의 위상 각을 동일 각만큼 시프트하고, 위상 각이 시프트되는 각각의 주파수 대역에 다른 위상 각 시프트가 인가되는, 오디오 디코딩 방법.

청구항 50.

제 49항에 있어서, 상기 제2 모드의 동작에서 주파수 대역 내에 스펙트럼 성분의 위상 각이, 주파수 대역 경계상에 스펙트럼 성분마다 위상 변화를 줄이도록 보간되는, 오디오 디코딩 방법.

청구항 51.

제 39항에 있어서, 상기 제1 모드의 동작은 적어도 하나 이상의 복수의 주파수 대역에서 스펙트럼 성분의 위상 각을 시프트하는 단계를 포함하고, 각각의 스펙트럼 성분은 다른 각도로 시프트되고, 상기 각은 실제로 시간에 따라 변하고, 상기 제2 모드의 동작은 스펙트럼 성분의 위상 각이 시프트되지 않는, 오디오 디코딩 방법.

청구항 52.

제 39항 내지 51항 중의 어느 한 항에 있어서, 상기 시프트는 랜덤 화 된 시프트를 포함하는, 오디오 디코딩 방법.

청구항 53.

제 52항에 있어서, 상기 랜덤 화 된 시프트의 양은 제어 가능한, 오디오 디코딩 방법.

청구항 54.

제 39항 내지 53항 중의 어느 한 항에 있어서, 제1 모드의 동작과 제2 모드의 동작에 따르는 상기 공간 파라미터 중의 하나에 응답하여 상기 오디오 신호의 스펙트럼 성분 중의 크기를 시프트하는 단계를 더 포함하는, 오디오 디코딩 방법.

청구항 55.

제 54항에 있어서, 상기 크기의 시프트는 랜덤 랜덤 시프트를 포함하는, 오디오 디코딩 방법.

청구항 56.

제 54항 또는 제 55항에 있어서, 상기 크기의 시프트의 양은 제어 가능한, 오디오 디코딩 방법.

청구항 57.

M은 1 이상이고 N은 2 이상인 N 오디오 채널을 나타내는 M인코딩된 오디오 채널을 수신하고, N 오디오 채널에 관련된 한 세트의 공간 파라미터를 수신하는 오디오 디코더에 있어서, 상기 M 오디오 채널에서 N 오디오 채널을 유도하고, 상기 M 오디오 채널을 능동적으로 디매트릭스하는 과정에 의해 상기 N 오디오 채널이 상기 M 오디오 채널에서 유도되고, 상기 디매트릭스는 상기 공간 파라미터 중 하나에 응답하여 적어도 부분적으로 및 상기 M 오디오 채널의 특성에 응답하여 적어도 부분적으로 동작하는, 오디오 디코딩 방법.

청구항 58.

제 1항 내지 57항 중의 어느 한 항에 따른 방법을 실행하도록 적용된 장치.

청구항 59.

제 1항 내지 57항 중의 어느 한 항에 따른 방법을 컴퓨터가 실행하도록 하는 컴퓨터-판독 가능한 매체 상에 저장된 컴퓨터 프로그램.

청구항 60.

제 1항 내지 17항 중의 어느 한 항에 따른 방법에 의해 발생된 비트스트림.

청구항 61.

제 1항 내지 17항 중의 어느 한 항에 따른 방법을 실행하도록 적용된 장치에 의해 발생된 비트스트림.

청구항 62.

제 1항 내지 17항 및 제 39항 내지 57항 중의 어느 한 항에 따른 방법을 실행하는 인코딩/디코딩 시스템.

명세서**기술분야**

본 발상관 해제명은 일반적으로 오디오 신호 처리에 관한 것이다. 본 발명은 특히 낮은 비트율(bitrate)과 매우 낮은 비트율 오디오 신호 처리에서 유용하다. 더욱 특히, 본 발명의 특징은 인코더(혹은 인코딩 과정), 디코더(혹은 디코딩 과정), 및 복수의 오디오 채널이 합성 단일 음의("모노":mono) 오디오 채널과 보조("사이드체인":sidechain) 정보로 나타나는 오디오 신호에 대한 인코드/디코드 시스템(혹은 인코딩/디코딩 과정)에 관한 것이다. 대안으로, 복수의 오디오 채널은 복수의 오디오 채널과 사이드체인 정보로 나타난다. 또한, 본 발명의 특징은 멀티채널 대 합성 단일 음의 채널 다운믹서(downmixer :혹은 다운믹스 과정), 단일 음의 채널 대 멀티채널 업믹서(upmixer: 혹은 업믹스 과정) 및, 단일 음의 채널 대 멀티채널 상관해제기(decorrelator: 혹은 상관 해제 과정)에 관한 것이다. 본 발명의 다른 특징은, 멀티채널-대-멀티채널 다운믹서(혹은 다운믹스 과정), 멀티채널-대-멀티채널 업믹서(혹은 업믹스 과정), 및 상관 해제기(혹은 상관 해제 과정)에 관한 것이다.

배경기술

AC-3 디지털 오디오 인코딩 및 디코딩 시스템에 대해, 상기 시스템이 비트에 대해 갈망하게 될 때 채널은 고주파수에서 선택적으로 결합되거나 "연결"된다. AC-3 시스템의 세부 내용이 공지되어 있다 - 예를 들어, Advanced Television Systems Committee 발행의 2001년 8월 20일자 *ATSC 표준 A52/A: 디지털 오디오 압축 표준(AC-3), Revision A*,를 보라. 상기 A/52A 자료는 월드 와이드 웹 주소지: <http://www.atsc.org/standards.html>에서 이용할 수 있다. 여기서 상기 A/52A 자료는 그 전체가 참고로 기술된다.

AC-3 시스템이 요구가 있으면 채널을 결합하는 것 이상의 주파수는 "커플링"(coupling) 주파수로 불린다. 커플링 주파수 이상으로, 연결된 채널은 "커플링" 혹은 합성 채널과 결합한다. 인코더는 각 채널에서 커플링 주파수 이상으로 각 서브밴드(subband)에 대해 "커플링 좌표"(진폭 스케일 인자)를 생성한다. 커플링 좌표는 각각의 연결된 채널 서브밴드의 원래 에너지와 합성 채널에서 해당하는 서브밴드의 에너지의 비율을 나타낸다. 커플링 주파수 아래, 채널이 이산적으로 인코딩된다. 이상(異相) 신호 성분 소거를 줄이기 위해 채널이 하나 이상의 다른 연결된 채널과 결합하기 전에, 연결된 채널의 서브밴드의 위상 극성이 반전된다. 개별 서브밴드를 기반으로 커플링 좌표를 포함하는 사이드체인 정보에 따른 합성 채널과 채널의 위상이 반전되는지의 여부가 디코더로 보내진다. 실제, AC-3의 상업적인 실시예에 적용되었던 커플링 주파수는 약

10kHz ~ 약 3500Hz 범위를 갖는다. 미국 특허 5,583,962, 5,633,981, 5,727,119, 5,909,664 및 6,021,386호는 여러 오디오 채널을 하나의 합성 채널로의 결합과 보조 혹은 사이드체인 정보와 원래 멀티채널로의 근사의 복원에 관한 가르침을 포함한다. 상기 특허의 각각은 그 전체가 여기에 참고로 기술된다.

발명의 상세한 설명

본 발명의 특징은 AC-3 인코딩 및 디코딩 시스템의 "커플링" 기술상의 개선으로 보여지고, 또한 오디오의 다중 채널이 단일 음의 합성 신호 혹은 관련된 보조 정보에 따른 오디오의 멀티채널과 결합되고 오디오의 멀티채널이 재구성되는 다른 기술상의 개선으로 보여진다.

본 발명의 특징은 N: 1: N 공간 오디오 코딩 기술(여기서 "N"은 오디오 채널의 수) 혹은 M: 1: N 공간 오디오 코딩 기술(여기서 "M"은 인코딩된 오디오 채널의 수이고, "N"은 디코딩된 오디오 채널의 수)에 적용되고, 이 기술은 개선된 위상 보상, 상관 해제(decorrelation) 메카니즘, 및 신호-의존의 가변 시정수(time constants)를 제공함으로써 채널 커플링을 개량한다. 본 발명의 특징은 또한 N: x: N 및 M: x: N 공간 오디오 코딩 기술("x"는 1 혹은 1 이상)에 적용된다. 디코더에서 위상 각과 상관 해제 정도를 복원함으로써 재생된 신호의 공간적인 차원을 개선하고, 다운믹싱(downmixing)하기 전에 상대적인 채널 간 위상을 조절함으로써 인코드 과정에서 커플링 소거 가공품의 감소를 목표로 포함한다. 본 발명의 특징은 실질적인 실시예로 구현될 때, 요구하는 대로의 채널 커플링보다는 계속적으로 채널 커플링을 허용해야 하고 예를 들어 AC-3 시스템에서보다는 더 낮은 커플링 주파수에서 허용되어야 하여서 요구되는 데이터 율(data rate)을 줄인다.

실시예

발명을 구현하는 가장 좋은 모드

기본적인 N: 1 인코더

도 1에 대해, 본 발명의 특징을 구현하는 N: 1 인코더 기능 혹은 장치가 도시된다. 도면은 본 발명의 특징을 구현하는 기본적인 인코더로서 작용하는 기능 혹은 구조의 예이다. 본 발명의 특징을 구현하는 다른 기능 혹은 구조적인 장치가, 아래 기술된 대안 및/또는 동일한 기능 혹은 구조적인 장치를 포함하여 적용된다.

두 개 이상의 오디오 입력 채널이 인코더에 인가된다. 주로, 본 발명의 특징이 아날로그, 디지털 혹은 하이브리드 아날로그/디지털 실시예로 구현되지만, 여기 기술된 예들은 디지털 실시예들이다. 이렇게 입력 신호는 시간샘플이고 이것은 아날로그 오디오 신호로부터 유도되어 왔다. 시간 샘플은 선형 펄스 코드 변조(PCM) 신호로 인코딩된다. 각각의 선형 PCM 오디오 입력 채널은 512-포인트 윈도우(window)된 선형 푸리에 변환(DFT)(고속 푸리에 변환(FFT)에 의해 구현되는)같은 동상(in-phase) 및 구상(quadrature) 출력을 갖는 필터 뱅크(filterbank) 기능 혹은 장치에 의해 처리된다. 필터뱅크는 시간-영역 대 주파수-영역 변환으로 여겨진다.

도 1은 필터뱅크 기능 혹은 장치, "필터뱅크"2에 인가된 제1 PCM 채널 입력(채널 "1")과 다른 필터뱅크 기능 혹은 장치, "필터뱅크"4에 각각 인가된 제2 PCM 채널 입력(채널 "n")을 도시한다. "n"개의 입력 채널이 있고, 여기서 "n"은 2 이상의 플러스 정수이다. 이렇게, 역시 "n" 개의 필터뱅크가 있고, 각각은 "n"개 입력 채널 중 유일한 하나를 수신한다. 간단히 표시하도록 도 1은 오직 두 개 입력 채널 "1"과 "n"을 보여준다.

필터뱅크가 FFT에 의해 구현될 때, 입력 시간-영역 신호는 연속 블록으로 분할되어 보통 중복된 블록에서 처리된다. FFT의 이산 주파수 출력(변환 계수)은 빈(bin: 디렉토리 이름)으로 언급되고, 각각은 동상과 구상 성분에 각각 해당하는 실수 및 허수 부분을 갖는 복소수 값을 갖는다. 인접하는 변환 빈이 사람 귀의 임계 대역폭에 근접한 서브밴드(subband)들로 그룹 지어지고, 앞으로 설명되듯이 인코더에 의해 생성된 대부분의 사이드체인 정보는 처리 자원을 최소화하고 비트율을 줄이기 위해 서브밴드 당 하나를 기반으로 계산되어 전송된다. 여러 개의 연속적인 시간-영역 블록이 프레임(frame)으로 그룹 지어지고, 이것은 각 블록 값이 평균이거나 그렇지 않으면 각 프레임 상에 결합되거나 누적되어 사이드체인 데이터 레이트를 최소화한다. 여기 기술된 예에서, 각각의 필터뱅크는 FFT에 의해 구현되고, 인접 변환 빈들은 서브밴드로 그룹 지어지고, 블록들은 프레임들로 그룹 지어지고, 사이드체인 데이터는 프레임당 한 번을 기본으로 보내진다. 대안으로, 사이드체인 데이터는 프레임당 한번 이상을 기본(예를 들어 한 블록당 한 번)으로 보내진다. 이제부터, 예를 들어 도 3과 그의 설명을 보라. 공지되듯이, 사이드체인 정보가 보내지는 주파수와 요구되는 비트율 사이에 교환(trade off)이 있다.

본 발명의 특징의 실질적인 적당한 구현은 48kHz 샘플링 레이트가 적용될 때 약 32밀리 초의 고정된 길이 프레임(length frame)을 채택하고, 각각의 프레임은 약 5.3 밀리 초의 간격으로 6 블록을 갖는다(예를 들어, 50% 중복으로 약 10.6밀리 초의 지속 시간을 갖는 블록을 채택). 그러나, 개별 프레임을 기본으로 보내지는 여기 기술된 정보가 대략 매 40밀리 초만 큼 자주 보내진다면, 그러한 타이밍도, 고정된 길이 프레임의 적용도, 고정된 수의 블록으로 그것들의 분할도 본 발명의 특징을 구현하는데 있어서 결정적이지 않다. 프레임은 임의 크기의 것으로 그것들의 크기는 동적으로 변한다. 가변 블록길이(Variable block lengths)는 상기의 AV-3 시스템에서처럼 적용된다. 여기 "프레임"과 "블록"이 기준이 된다는 것을 이해할 것이다.

실제로, 합성 모노(composite mono) 혹은 멀티채널 신호(들), 아니면 합성 모노 혹은 멀티채널 신호(들)과 이산-저-주파수 채널이 아래 설명되듯이 예를 든 것처럼 감지 코더(perceptual coder)에 의해 인코드 된다면, 감지 코더에 적용된 것처럼 동일 프레임과 블록 구조를 채택하는 것이 편리하다. 더욱이, 때때로 한 블록 길이에서 다른 것으로 스위칭이 있는 식으로 코더가 가변 블록 길이를 채택한다면, 그러한 블록 스위칭이 일어날 때 여기 기술된 하나 이상의 사이드체인 정보가 갱신된다면 그것은 바람직하다. 그러한 스위치가 발생할 때 사이드체인 정보의 갱신시에 데이터 오버헤드의 증가를 줄이기 위해, 갱신된 사이드체인 정보의 주파수 해상도가 감소한다.

도 3은 시간 축을 따르는(수평) 프레임과 블록과, 주파수 축을 따르는(수직) 빈(bin)과 서브밴드의 간략화된 개념적인 구조의 예를 도시한다. 빈이 임계 대역에 근접하는 서브밴드로 분할될 때, 가장 낮은 주파수 서브밴드는 가장 작은 빈(예를 들어, 하나)을 갖고 서브밴드 당 빈의 수는 증가하는 주파수에 따라 증가한다.

도 1로 돌아가서, 각각의 채널의 각 필터뱅크(이 예에서 필터뱅크 2와 4)에 의해 생성된 주파수-영역 버전의 각각의 n-시간-영역 입력 채널들은 가산 기능 혹은 장치 "가산기"에 의해 단일 음의("모노") 합성 오디오 신호와 함께 합산된다("다운 믹스").

다운 믹싱은 입력 오디오 신호의 전체 주파수 대역폭에 적용되거나, 선택적으로 주어진 "커플링" 주파수 위의 주파수로 제한되고, 다운 믹싱 과정의 가공품이 중간 내지 낮은 주파수에서 더 들을 수 있게 된다. 그러한 경우에, 채널은 커플링 주파수 아래 이산적으로 운반된다. 처리되는 가공품이 문제가 되지 아닐지라도 변환 빈을 임계-대역 같은 서브밴드(주파수에 비례하는 크기)로 그룹화하여 구성된 중/저 주파수 서브밴드가 저 주파수에서 작은 수의 변환 빈(매우 낮은 주파수에서 한 개 빈)을 갖기 쉽고, 사이드체인 정보를 갖는 다운 믹스된 모노 오디오 신호를 보낼 필요가 있기보다 작거나 더 작은 비트로 직접 코딩된다는 점에서 이러한 전략은 바람직하다. 4kHz, 2300Hz, 1000Hz 같이 낮은 커플링 혹은 변환 주파수 아니면 디코더에 인가된 오디오 신호의 주파수 대역의 최저치조차도 어떤 응용에 대해 수용 가능하고, 특히 그런 것에는 매우 낮은 비트율이 중요하다. 다른 주파수는 비트 절감과 청취자 수용 사이에 유용한 조화를 제공한다. 특정 커플링 주파수의 선택은 본 발명에 결정적이지 않다. 커플링 주파수는 가변으로, 가변이면 그것은 예를 들어 입력 신호 특성에 직접 혹은 간접으로 좌우된다.

다운믹싱하기 전에, 본 발명의 특징은 채널들이 결합될 때 이상(異相)신호 성분의 소거를 줄이기 위하여 서로 마주 대하는 채널들의 위상 각 정렬(phase angle alignments)을 개선하고, 개선된 모노 합성 채널을 제공하는 것이다. 이것은 채널들 중 하나에서 변환 빈 모두 혹은 일부의 "절대 각"(absolute angle)을 시간상으로 제어 가능하게 시프트(shift)함으로써 실행된다. 예를 들어, 커플링 주파수위의 오디오를 나타내는 모든 변환 빈이 흥미있는 주파수 대역을 정의하여, 각 채널에서 필요한 대로 시간상으로 제어 가능하게 시프트하거나, 한 개 채널이 기준으로 이용될 때 기준 채널 외의 모든 채널에서 시프트된다.

빈의 "절대 각"은 필터뱅크에 의해 생성된 각(angle) 복소수 값의 변환 빈의 크기-및-각 표시의 각도로서 취해진다. 한 채널에서 빈의 상기 절대 각의 제어 가능한 시프트는 각 회전 기능 혹은 장치("각 회전": Rotate Angle)에 의해 실행된다. 가산기(6)에 의해 제공된 다운-믹스 합산에 적용하기 전에 각 회전(8)은 필터뱅크(2)의 출력을 처리하는 동안에, 가산기(6)에 적용하기 전에 각 회전(10)은 필터뱅크(4)의 출력을 처리한다. 어떤 신호 조건 아래, 시간(여기 기술된 예에서, 프레임의 시간 지속시간)상으로 특정 변환 빈에 대해 각 회전이 필요 없다는 것을 알 수 있을 것이다. 커플링 주파수 아래, 채널 정보가 이산적으로 인코딩된다(도 1에 미 도시).

주로, 채널들의 서로에 대한 위상 각 정렬의 개선은, 흥미있는 주파수 대역을 통한 각각의 블록에서 각 변환 빈 혹은 서브밴드의 위상을 -(네거티브)의 절대 위상 각만큼 시프트함으로써 실행된다. 이것은 실제로 이상 신호 성분의 소거를 피하지만, 특히 나오는 모노 합성 신호가 고립된 것으로 들린다면 청취 가능한 가공품을 초래하기 쉽다. 이렇게, 디코더에 의해 재구성된 멀티 채널 신호의 공간 이미지 붕괴를 최소화하고 다운-믹스 과정에서 이상 소거를 최소화하도록 오직 필요한

만큼 채널에서 빈의 절대 각을 시프트 함으로서 "최소 처리"의 원리를 적용하는 것이 바람직하다. 그러한 각 시프트(angle shift)를 결정하는 기술이 아래 기술된다. 그러한 기술은 시간과 주파수 평활화(smoothing)와 신호 처리가 과도(transient)의 출현에 응답하는 방식을 포함한다.

또한 에너지 정규화(normalization)가 아래 기술되듯이, 고립된 빈들의 나머지 이상(異相) 소거를 좀더 줄이도록 인코더에서 빈 한 개당 기반으로 실행된다. 역시 아래 기술되듯이, 또한 에너지 정규화는 모노 합성 신호의 에너지가 기여된 채널들의 에너지의 합과 같도록 서브밴드 당 기반(디코더에서)으로 실행된다.

각각의 입력 채널은 다운믹스 합산기(6)에 인가되기 전에 채널에 인가된 각 회전의 양 혹은 정도를 제어하고 그 채널에 대해 사이드체인 정보를 발생하도록, 그와 연관된 오디오 분석기 기능 혹은 장치("오디오 분석기")를 구비한다. 채널 1과 n의 필터뱅크 출력이 오디오 분석기(12)와 오디오 분석기(14)에 각각 인가된다. 오디오 분석기(12)는 채널 1에 대한 사이드체인 정보와 채널 2에 대한 위상 각 회전의 양을 발생한다. 오디오 분석기(14)는 채널 n에 대한 사이드체인 정보와 채널 n에 대한 각 회전의 양을 발생한다. "각"에 대한 그러한 기준은 여기서 위상 각으로 언급된다는 것을 알 수 있을 것이다.

오디오 분석기에 의해 발생된 각 채널에 대한 사이드체인 정보는 다음을 포함한다:

진폭 스케일 인자("진폭 SF"),

각 제어 파라미터,

상관 해제 스케일 인자("상관 해제 SF")

과도 플래그(Flag), 및

선택적으로, 보간(Interpolation) 플래그.

그러한 사이드체인 정보는 채널의 공간 성질을 나타내고/거나 공간 처리에 관련된 신호 특성을 나타내는 "공간 파라미터"로 특징 지워진다. 그러한 경우에, 사이드체인 정보는 단일 서브밴드(과도 플래그와 보간 플래그를 제외하고, 이것의 각각이 한 채널 내에 모든 서브밴드에 인가)에 적용되고, 아래 기술된 예처럼 관련된 코더(coder)에서 블록 스위치의 발생시에 프레임당 한번 갱신된다. 여러 공간 파라미터의 더 상세한 내용이 아래 설명된다. 인코더에서 특정 채널에 대한 각 회전이, 사이드체인 정보의 일부를 형성하는 극성-반전의 각 제어 파라미터로 취해진다.

기준 채널이 채택되면, 그 채널은 오디오 분석기를 요하지 않거나, 대안으로 오직 진폭 스케일 인자 사이드체인 정보를 발생하는 오디오 분석기를 요한다. 스케일 인자가 다른, 비-기준 채널들의 진폭 스케일 인자들로부터 디코더에 의해 충분한 정확도로 추정될 수 있다면 진폭 스케일 인자를 보낼 필요가 없다. 인코더에서 에너지 정규화가 아래 기술되듯이, 어떤 서브밴드 내에 채널 너머 스케일 인자가 실제로 1과 일치하도록 하면, 디코더에서 기준 채널의 진폭 스케일 인자의 근사치를 추정하는 것이 가능하다. 추정된 근접의 기준 채널 진폭 스케일 인자 값이 재생된 멀티채널 오디오에서의 이미지 시프트가 되는 진폭 스케일 인자의 비교적 조잡한 양자화의 결과로서 에러를 갖는다. 그러나, 낮은 데이터 레이트 환경에서, 그러한 가공품은 기준 채널의 진폭 스케일 인자를 보내도록 비트를 이용하는 것보다 더욱 수용 가능하다. 그럼에도 불구하고, 어떤 경우에는 적어도 진폭 스케일 인자 사이드체인 정보를 발생하는 기준 채널에 대한 오디오 분석기를 적용하는 것이 바람직하다.

도 1은 PCM 시간 영역에 입력단에서 채널의 오디오 분석기까지 각각의 오디오 분석기로 선택적인 입력을 점선으로 도시한다. 상기 입력은 오디오 분석기에 의해 사용되어 시간 기간(여기 기술된 예에서, 블록 혹은 프레임의 기간)상의 과도를 검출하고 상기 과도에 응답하여 과도 표시기(예를 들어, 1비트의 "과도 플래그")를 생성한다. 대안으로, 아래의 도 4의 단계 408의 설명에 기술되듯이 과도는 주파수 영역에서 검출되고, 이런 경우에 오디오 분석기는 시간-영역 입력을 수신할 필요가 없다.

모든 채널(혹은 기준 채널을 제외한 모든 채널)에 대한 사이드체인 정보와 모노 합성 오디오 신호가 디코딩 과정 혹은 장치("디코더")에 저장, 전송되거나 저장되어 전송된다. 저장, 전송 혹은 저장과 전송 매체 혹은 미디어에 알맞은 하나 이상의 비트 스트림으로 패킹(pack)된다. 모노 합성 오디오는 예를 들어, 감지 인코더 같은 데이터-율 감소 인코딩 방법 혹은 장치에 인가되거나 저장, 전송 혹은 저장 및 전송에 감지 인코더 및 엔트로피 코더(entropy coder: 예를 들어, 산술 혹은 Huffman 코더)(때때로 "손실없는" 코더로 불림)가 인가된다. 또한, 위에 언급하였듯이, 모노 합성 오디오와 관련된 사이드체인 정보가 일정한 주파수 위의 오디오 주파수("커플링" 주파수)만으로 여러 입력 채널로부터 유도된다. 그런 경우에, 여

러 입력 채널의 각각에서 커플링 주파수 아래의 오디오 주파수가 이산 채널로서 저장되고, 전송되거나 저장되어 전송되고, 혹은 여기 기술된 것 이외의 다른 방식으로 결합되거나 처리된다. 그러한 이산적이거나 달리 결합된 채널은 예를 들어 감지 인코더 혹은 감지 인코더 및 엔트로피 인코더 같은 데이터 감소 인코딩 방법 혹은 장치에 역시 인가된다. 모노 합성 오디오 및 이산 멀티 채널 오디오는 모두 집적된 감지의 인코딩 혹은 감지 및 엔트로피 인코딩 과정 혹은 장치에 적용된다.

사이드체인 정보가 인코더 비트-스트림(bit-stream)으로 전송되는 특정 방식이 본 발명에는 결정적이지 않다. 필요하다면, 비트스트림이 유증 디코더(legacy decoder: 즉, 비트스트림이 나중에 양립)와 양립하는 식으로 사이드체인 정보가 운반된다. 그렇게 하기 위한 알맞은 다수의 기술이 공지되어 있다. 예를 들어, 많은 인코더는 디코더에 의해 무시되는 사용되지 않거나 무효 비트를 갖는 비트스트림을 생성한다. 그러한 장치의 예는 2004년 10월 19일 미국 특허 6,807,528 B1 트루먼 공동발명의 제목 "압축된 데이터 프레임으로의 데이터 추가"에 설명되고, 본 특허는 여기서 그 전체가 참고로 기술된다. 그러한 비트는 사이드체인 정보로 대체된다. 다른 예는 사이드체인 정보가 인코더의 비트 스트림으로 인코딩된다는 것이다. 대안으로, 유증 디코더와 양립하는 모노/스테레오 비트 스트림에 따른 그러한 정보의 전송 혹은 저장을 허용하는 어떤 기술에 의해 사이드체인 정보가 후방으로 양립하는 비트 스트림으로부터 개별적으로 저장 혹은 전송된다.

기본적인 1:N 및 1:M 디코더

도 2에 대해, 본 발명의 특징을 구현하는 디코더 기능 혹은 장치("디코더")가 도시된다. 도면은 본 발명의 특징을 구현하는 기본적인 디코더로서 작용하는 기능 혹은 구조의 예이다. 본 발명의 특징을 구현하는 다른 기능 혹은 장치가 아래 기술된 대안 및/또는 동일한 기능 혹은 구조적인 장치를 포함하여 적용된다.

디코더는 기준 채널을 제외하고는 모든 채널에 대한 모노 합성 오디오 신호와 사이드체인정보를 수신한다. 필요하다면, 합성 오디오 신호와 관련된 사이드체인 정보가 디멀티플렉스(demultiplex), 언팩(unpack) 및/또는 디코드된다. 디코딩이 테이블 검사를 채택한다. 여기 기술된 본 발명의 비트율-감소 기술을 가정하여, 목표는 도 1의 인코더에 인가된 오디오 채널의 각각의 것에 근접하는 복수의 개별 오디오 채널을 모노 합성 오디오 채널들로부터 유도하는 것이다.

물론, 인코더에 인가된 모든 채널을 복원하지 않도록 고르거나 단일 음의 합성 신호만을 사용하도록 후자는 선택한다. 대안으로, 인코더에 인가된 것에 더한 채널이 2002년 2월 7일 출원되고 2002년 8월 15일 공개된 국제 특허 출원 PCT/US 02/03619의 지정국이 미국으로 2003년 8월 5일 출원된 그의 미국 국내 단계 출원 S.N. 10/467,213과, 2003년 8월 6일 출원된 국제 특허 출원 PCT/US03/24570으로 지정국이 미국으로 2005년 2월 27일 출원된 그의 미국 국내 단계 출원 S.N. 10/522,515인 2001년 3월 4일 공개된 WO 2004/019656에 기술된 본 발명의 특징을 채택함으로써 본 발명의 특징에 따른 디코더의 출력으로부터 유래된다. 상기 출원은 여기서 그 전체가 참고로 기술된다. 본 발명의 특징을 구현하는 디코더에 의해 복원된 채널은, 복원된 채널이 유용한 채널간 진폭 관계뿐만 아니라 복원된 채널이 유용한 채널간 위상 관계를 갖는 점에서, 특히 인용된 출원의 채널 승산 기술과 연관되어 유용하다. 채널 승산에 대한 다른 대안은 부가적인 채널을 유도하는데 매트릭스 디코더(matrix decoder)를 채택하는 것이다. 본 발명의 채널 간 진폭 및 위상보존 특징은 본 발명의 특징을 구현하는 디코더의 출력 채널을 진폭-및 위상 감지 매트릭스 디코더에 인가하는데 특히 알맞도록 한다. 그러한 많은 매트릭스 디코더는, 그들에 인가된 신호가 신호 대역을 통틀어 스테레오일 때만이 적절하게 작동하는 광대역 제어 회로를 채용한다. 이렇게, 본 발명의 특징이 N이 2인 N: 1: N 시스템에서 구현된다면, 디코더에 의해 복원된 두 개 채널이 2: M 액티브 매트릭스 디코더에 인가된다. 적당한 많은 액티브 매트릭스 디코더가 공지되어 있고 예를 들어, "Pro Logic" 및 "Pro Logic II" 디코더("Pro Logic"은 돌비 연구소 라이선싱 코퍼레이션의 등록 상표)로 알려진 매트릭스 디코더를 포함한다. Pro Logic의 특징은 미국 특허 4,799,260과 4,941,177에 기술되어 있고, 이들 각각은 여기서 그 전체가 참고로 기술된다. Pro Logic II 디코더의 특징은 2001, 6,7에 WO 01/41504으로 공개되고 2000, 3, 22 출원된 "2 입력 오디오 신호에서 적어도 3 오디오 신호를 유도하는 방법"이란 제목의 Fostage의 진행중인 미국 특허 출원 S.N. 09/532,711과, 2004, 7, 1에 US 2004/0125960 A1으로 공개되고 2003, 2, 25 출원된 "오디오 매트릭스 디코딩을 위한 장치에 대한 방법"이란 제목의 Fostage 공동 발명의 진행중인 미국 특허 출원 S.N. 10/362,786에 기술되어 있다. 상기 출원의 각각은 그 전체가 여기에 참고로 기술된다. 돌비 Pro Logic과 Pro Logic II 디코더의 동작의 어떤 특징은 예를 들어, 돌비 연구소의 웹 사이트(www.dolby.com); Roger Rressher 저의 "돌비 씨라운드 Pro Logic 디코더 동작에서 이용 가능한 논문"에 설명된다. Jim Hilson 저의 "돌비 Pro Logic II 기술로의 믹싱" 적당한 다른 액티브 매트릭스 디코더는 하나 이상의 다음 미국 특허와 공개된 국제 특허 출원(각각 미국 지정)에 기술된 것을 포함하고, 그 각각은 그 전체가 여기에 참고로 기재된다: 5,046,098, 5,274,740, 5,400,433, 5,625,696, 5,644,640, 5,504,819, 5,428, 687, 5,172,415 및 WO 02/19768.

다시 도 2에 대해, 복원된 다수의 오디오 채널이 유도되는 복수의 신호 통로에 수신된 모노 합성 오디오 채널이 인가된다. 각 채널-유도 통로는 차례로 진폭 조절 기능 혹은 장치("진폭 조절") 및 각 회전 기능 혹은 장치(각 회전: Rotate Angle)를 포함한다.

일정한 신호 조건 아래에, 출력 채널의 상대적인 출력 크기(혹은 에너지)가 인코더의 입력단에서 채널의 크기와 유사하도록, 진폭 조절은 모노 합성 신호에 이득 혹은 손실을 인가한다. 대안으로, "랜덤화 된(randomized)" 각 변화가 부가될 때 일정 신호 조건 아래, 다음에 설명되듯이 복원된 채널 중 다른 것에 대한 그의 상관 해제(decorrelation)를 개선하도록 "랜덤화 된" 진폭 변화의 제어 가능한 양이 역시 복원된 채널의 진폭에 부가된다.

일정한 신호 조건 아래, 모노 합성 신호로부터 유도된 출력 채널의 상대적인 위상 각이 인코더의 입력단에서의 채널의 것과 유사하도록 회전 각이 위상 회전을 인가한다. 바람직하게도, 일정 신호 조건 아래, 복원된 채널 중 다른 것에 대한 그의 상관 해제를 개선하도록 "랜덤화 된" 각 변동의 제어 가능한 양이 역시 복원된 채널의 각에 부가된다.

아래 논의되듯이, "랜덤화 된" 각 진폭 변동이 가짜-랜덤 및 진짜 랜덤 변동뿐 아니라 채널 사이에 크로스-상관 관계(cross-correlation)의 효과를 갖는 결정론적으로 발생된 변동을 포함한다. 이것은 아래의 도 5A의 단계 505의 설명에 더 논의된다.

개념적으로, 특정 채널에 대한 진폭 조절 및 각 회전이 채널에 대한 재구성된 변환 빈(bin) 값을 내도록 모노 합성 오디오 DFT 계수들을 비교한다.

각각의 채널에 대한 진폭 조절이 특정 채널에 대한 복원된 사이드체인 진폭 스케일 인자(Amplitude Scale Factor)에 의해서 적어도 제어되거나, 기준 채널의 경우에는 기준 채널에 대한 복원된 사이드체인 진폭 스케일로부터 혹은 다른, 비-기준, 채널의 복원된 사이드체인 진폭 스케일 인자로부터 추정된 진폭 스케일 인자로부터 제어된다. 대안으로, 복원된 채널의 상관 해제를 촉진하도록, 역시 진폭 조절이 특정 채널에 대한 복원된 사이드체인 과도 플래그(transient flag)와 특정 채널에 대한 복원된 사이드체인 상관 해제 스케일 인자로부터 유도된 랜덤 진폭 스케일 인자 파라미터에 의해 제어된다. 각각의 채널에 대한 각 회전은 적어도 복원된 사이드체인 각 제어 파라미터(Angle Control Parameter: 이 경우에, 디코더의 각 회전은 인코더의 각 회전에 의해 제공된 각 회전을 실제로 원상태로 해 놓는다)에 의해 제어된다. 복원된 채널의 상관 해제를 촉진하도록, 특정 채널에 대해 복원된 사이드체인 상관 해제 스케일 인자와 특정 채널에 대해 복원된 사이드체인 과도 플래그로부터 유도된 랜덤화 된 각 제어 파라미터에 의해 각 회전은 역시 제어된다. 채널에 대한 랜덤화 된 각 제어 파라미터와, 적용된다면, 채널에 대한 랜덤화 된 진폭 스케일 인자가 제어 가능한 상관 해제기 기능 혹은 장치("제어 가능한 상관 해제기")에 의해 복원된 과도 플래그와 채널에 대한 복원된 상관 해제 스케일 인자로부터 유도된다.

도 2의 예에 대해, 복원된 모노 합성 오디오는 제1 채널 오디오 복원 통로(22)에 인가되고, 이것은 채널 1 오디오를 유도하고, 제 2 채널 오디오 복원 통로(24)에 인가된 것은 채널 n 오디오를 유도한다. 오디오 통로(22)는 진폭 조절(26), 각 회전(28)을 포함하고, PCM 출력이 필요하다면, 역 필터뱅크 기능 혹은 장치("역 필터뱅크")(30)를 포함한다. 유사하게, 오디오 통로(24)는 진폭 조절(32), 각 회전(34)을 포함하고, PCM 출력이 필요하다면, 역 필터뱅크 기능 혹은 장치("역 필터뱅크")(36)를 포함한다. 도 1의 경우처럼, 간단히 나타나도록 오직 두 개 채널이 도시되고 2채널 이상이 있다는 것을 이해할 것이다.

제1 채널, 채널 1에 대해 복원된 사이드체인 정보가 진폭 스케일 인자, 각 제어 파라미터, 상관 해제 스케일 인자, 과도 플래그 및 선택적으로, 기본 인코더의 설명과 연관되어 위에 언급된 보간 플래그를 포함한다. 진폭 스케일 인자는 진폭 조절(26)에 인가된다. 선택적인 보간 플래그가 채택되면, 선택적인 주파수 보간기(interpolator) 혹은 보간기 기능("보간기")(27)이 채택되어 주파수 상의(예를 들어, 채널의 각 서브밴드에서 빈 상에) 각 제어 파라미터를 보간한다. 그러한 보간은 예를 들어, 각 서브밴드의 중심 사이에 빈 각도(bin angle)의 선형 보간이다. 1비트 보간 플래그의 상태는, 아래 설명되듯이 주파수 상의 보간이 채택될지의 여부를 선택한다. 과도 플래그와 상관 해제 스케일 인자는, 랜덤화 된 각 제어 파라미터를 발생하는 제어 가능한 상관 해제기(38)에 인가된다. 1비트 과도 플래그의 상태는 아래 더 설명되듯이, 랜덤화 된 각 상관 해제기의 두 개 다중 모드 중 하나를 선택한다. 보간 플래그와 보간기가 적용된다면 주파수 상에 보간된 각 제어 파라미터(Angle Control Parameter)와 랜덤화 된 각 제어 파라미터(Randomized Angle Control Parameter)가 가산기 혹은 결합 기능(40)에 의해 함께 합쳐져서 각 회전(28)에 대한 제어 신호를 제공한다. 대안으로, 제어 가능한 상관 해제기(38)는 역시 과도 플래그와 상관 해제 스케일 인자에 응답하여 랜덤화 된 진폭 스케일 인자를 발생하고, 랜덤화 된 각 제어 파라미터 발생을 부가한다. 가산기 혹은 결합하는 기능(미 도시)에 의해 진폭 스케일 인자는 그러한 랜덤화 된 진폭 스케일 인자와 함께 합산되어 진폭 조절(26)에 대한 제어 신호를 제공한다.

유사하게, 제2 채널에 대한 복원된 사이드체인 정보, 채널 n이 진폭 스케일 인자, 각 제어 파라미터, 상관 해제 스케일 인자, 과도 플래그 및 선택적으로, 기본 인코더의 설명과 연관되어 위에 언급된 보간 플래그를 포함한다. 진폭 스케일 인자는 진폭 조절(32)에 인가된다. 주파수 상에 각 제어 파라미터를 보간하기 위해 주파수 보간기 혹은 보간기 기능("보간기")(33)이 채택된다. 채널 1처럼, 1 비트 보간 플래그의 상태는 주파수 상의 보간의 적용 여부를 선택한다. 과도 플래그와 상관 해

제 스케일 인자는, 그에 응답하여 랜덤 화 된 각 제어 파라미터를 발생하는 제어 가능한 디코리레이터(42)에 인가된다. 채널 1에서처럼, 1비트 과도 플래그의 상태는 아래 더 설명되듯이, 랜덤 화 된 각 상관 해제의 두 개 다중 모드 중 하나를 선택한다. 각 제어 파라미터와 랜덤 화 된 각 제어 파라미터가 가산기 혹은 결합 기능(44)에 의해 함께 합해져서 회전 각(34)에 대한 제어 신호를 제공한다. 대안으로 채널 1과 연관지어 위에서 설명하였듯이, 제어 가능한 상관 해제기(42)는 역시 과도 플래그와 상관 해제 스케일 인자에 응답하여 랜덤 화 된 진폭 스케일 인자를 발생하고, 랜덤 화 된 각 제어 파라미터 발생을 부가한다.

가산기 혹은 결합하는 기능(미 도시)에 의해 진폭 스케일 인자 및 랜덤 화 된 진폭 스케일 인자는 함께 합산되어, 진폭 조절(32)에 대한 제어 신호를 제공한다.

방금 기술한 과정 혹은 위상 수학이 이해에 유용하지만, 반드시 동일하거나 유사한 결과를 얻는 다른 방법 혹은 위상 수학으로 동일 결과가 얻어진다. 예를 들어, 진폭 조절 26(32) 및 각 회전 28(34)의 순서가 반전되고/되거나 각 제어 파라미터에 응답하는 하나 이상의 각 회전이 있고, 랜덤 화 된 각 제어 파라미터에 응답하는 다른 것이 있다. 또한, 각 회전은 아래 기술된 도 5의 예에서처럼, 하나 혹은 두 개 기능 혹은 장치보다는 3개로 여겨진다. 랜덤 화 된 진폭 스케일 인자가 적용된다면, 하나 이상의 진폭 조절이 있다-진폭 스케일 인자에 응답하는 것과 랜덤 화 된 진폭 스케일 인자에 응답하는 것. 위상에 대한 진폭에 사람 귀의 더 큰 감도 때문에, 랜덤 화 된 진폭 스케일이 적용되면, 진폭 상의 그의 효과가 랜덤 화 된 각 제어 파라미터가 위상 각 상에 갖는 효과보다 작도록, 랜덤 화 된 각 제어 파라미터의 효과에 대한 그의 효과를 비교하는 것이 바람직하다. 다른 대안의 과정 혹은 위상 수학으로서, 상관 해제 스케일 인자가 랜덤 화 된 위상 각 대 기본적인 위상 각의 비율(기본적인 위상 각을 나타내는 파라미터에 랜덤 화 된 위상 각을 나타내는 파라미터를 가산하기보다는)을 제어하고, 또한 적용되면, 랜덤 화 된 진폭 시프트 대 기본적인 진폭 시프트의 비율을 제어한다(기본적인 진폭을 나타내는 스케일 인자에 랜덤 화 된 진폭을 나타내는 스케일 인자를 가산하기보다는) (즉, 각각의 경우에 가변 크로스페이드(crossfade)).

기준 채널이 적용되면 기본적인 인코더와 연관되어 위에 논의되었듯이, 각 회전, 제어 가능한 상관 해제기 및 기준 채널에 대한 사이드체인 정보로서 생략되는 채널에 대한 가산기가 진폭 스케일 인자만을 포함한다(그렇지 않으면, 대안으로 사이드체인 정보가 기준 채널에 대한 진폭 스케일 인자를 포함하지 않으면, 다른 채널들의 진폭 스케일 인자로부터 추정되고 이때는 인코더에서의 에너지 정규화가 서브 대역 내에 채널 상의 스케일 인자가 1에 일치하는 때이다). 진폭 조절이 기준 채널에 대해 제공되고, 그것은 기준 채널에 대해 수신되거나 유도된 진폭 스케일 인자에 의해 제어된다. 기준 채널의 진폭 스케일 인자가 사이드체인으로부터 유도되든지 아니면 디코더에서 추정되든지 간에, 복원된 기준 채널이 진폭-스케일 버전의 모노 합성 채널이다. 그것은 다른 채널의 회전에 대해 기준이기 때문에 각 회전을 요하지 않는다.

복원된 채널의 상대적인 진폭 조절이 적당한 정도의 상관 해제를 제공하지만, 진폭 조절만의 사용은 많은 신호 조건에 대해(예를 들어, "조잡한: 사운드필드) 공간화 혹은 영상화에서의 재생된 사운드 필드 부족을 초래할 것 같다. 진폭 조절은 귀에서 청각 사이 레벨 차이에 영향을 주고, 이것은 귀에 의해 적용된 사이코음향의 방향의 신호들 중 오직 하나이다. 이렇게, 본 발명의 특징에 따라, 일정한 각-조절 기술이 신호 조건에 따라 적용되어 부가적인 상관 해제를 제공한다. 본 발명의 특징에 따라 적용된 동작의 여러 각-조절 상관 해제 기술 혹은 모드를 이해하는데 유용한 생략된 해설을 제공하는 표 1이 참고된다. 도 8,9의 예와 관련지어 아래 기술된 다른 상관 해제 기술이 표 1의 기술 대신에 적용되거나 그에 더하여 적용된다.

실제, 각 회전과 크기 수정의 적용은 원형 컨벌루션(convolution)(역시 순환 혹은 주기적인 컨벌루션으로 알려짐)을 초래한다. 일반적으로, 원형 컨벌루션을 피하는 것이 바람직하지만, 원형 컨벌루션을 초래하는 바람직하지 않은 가청 가공품은 인코더와 디코더에서 보상 각 시프트에 의해 얼마간 감소된다. 게다가, 원형 컨벌루션의 효과는 본 발명의 특징의 저가 구현으로 취급되고, 특히 모노 혹은 여러 채널로의 다운믹싱이 예를 들어 1500Hz이상(이 경우에는 원형 컨벌루션의 가청 효과는 최소)같은 오디오 주파수 대역의 일부에서 오직 일어나는 경우이다. 대안으로, 원형 컨벌루션이 예를 들어, 제로 패딩(zero padding)의 적절한 이용을 포함하는 어떤 적당한 기술에 의해 피해지거나 최소화된다. 제로 패딩을 이용하는 한 가지 방법은, 제거된 주파수 영역 변동(각 회전과 진폭 스케일을 나타냄)을 시간 영역으로 변환하는 것으로, 그것을 윈도(window: 임의의 윈도로)하고, 제로로 패딩하여서 다시 주파수 영역으로 변환하여 주파수 영역 버전의 오디오로 곱한다(오디오는 윈도될 필요가 없다).

표1

각-조절 상관 해제 기술

	기술 1	기술 2	기술 3
--	------	------	------

신호의 유형(대표적인 예)	스펙트럼의 정적인 자원	복잡한 연속 신호	복잡한 총동적인 신호(과도)
상관 해제 상의 효과	저 주파수와 안정상태 신호 성분을 상관해제	비-총동적인 복잡한 신호 성분 상관 해제	총격의 고 주파수 신호 성분 상관 해제
프레임에 나타난 과도의 효과	단축된 시정수로 동작	동작 안함	동작
실행된 것	한 채널에서 빈 각을 천천히 시프트(프레임마다)	한 채널에서 개별 빈을 기반으로 기술1의 각을 시간-변하는 랜덤화 된 각에 가산	한 채널에서 개별 서브밴드를 기반으로, 기술 1의 각을 빨리 변하는 랜덤화 된 각에 부가
제어되거나 스케일되는 매체	기본적인 위상 각이 각 제어 파라미터에 의해 제어된다	랜덤화 된 각의 양은 상관 해제 SF에 의해 직접 스케일된다; 서브밴드 상에 같은 스케일링, 매 프레임 스케일링 갱신됨	랜덤화 된 각의 양은 상관 해제 SF에 의해 직접 스케일된다; 서브밴드 상에 같은 스케일링, 매 프레임 스케일링 갱신됨
각 시프트의 주파수 분석	서브밴드(각 서브밴드에 모든 빈에 인가된 동일 혹은 보간된 시프트 값)	빈(각각의 빈에 적용된 다른 랜덤화 된 시프트 값)	서브밴드(각 서브밴드에서 모든 빈에 인가된 동일 랜덤화 된 시프트 값; 채널에서 각 서브밴드에 인가된 다른 랜덤화 된 시프트 값)
시간 분석	프레임(시프트 값이 매 프레임 갱신됨)	랜덤화 된 시프트 값이 동일하게 남고 불변	블록(랜덤화 된 시프트 값이 매 블록 갱신됨)

예를 들어, 피치 파이프 노트 같은 스펙트럼으로 실제로 정적인 신호에 대해, 제1 기술("기술 1")은 인코더의 입력단에서 다른 채널에 대해 채널의 원래 각과 유사한 각으로, 다른 복원된 각각의 채널의 각에 대해 수신된 모노 합성 신호의 각을 복원한다. 귀가 오디오 신호의 개별 사이클을 따르는 약 1500Hz 이하의 저-주파수 신호 성분의 상관 해제를 제공하는데, 특히 위상 각 차이가 유용하다. 바람직하게, 기술 1이 모든 신호 조건 아래 동작하여서 기본적인 시프트를 제공한다.

약 1500Hz 이상의 고-주파수 신호 성분에 대해, 귀는 개별 사이클의 사운드를 따르지 않지만 대신에 파형 엔빌로프(envelopes)에 응답한다. 그리하여, 위상 각 차이보다는 신호 엔빌로프의 차이로 제공된 약 1500Hz 이상의 상관 해제가 더 좋다. 기술 1에 따른 위상 각 시프트만의 적용은 고-주파수 신호를 상관 해제하는데 충분히 신호의 엔빌로프를 변경하지 못한다. 제2 및 제3 기술(각각 "기술 2"와 "기술 3")은 제어 가능한 양의 랜덤화 된 각 변화를 일정 신호 조건 아래 기술 1에 의해 결정된 각도로 가산하여, 제어 가능한 양의 랜덤 엔빌로프 변화를 야기하고, 이것은 상관 해제를 촉진한다.

위상 각에서의 랜덤화 된 변화는 신호의 엔빌로프에서 랜덤화 된 변화를 초래하는 방법이다. 특정 엔빌로프는 서브밴드 내에 스펙트럼 성분의 진폭과 위상의 특별한 결합의 상호작용이 된다. 서브밴드 내에 스펙트럼 성분의 진폭의 변화가 엔빌로프를 변화시키지만, 엔빌로프에서 커다란 변화를 얻는데 큰 진폭 변화가 필요하고, 이것은 사람 귀가 스펙트럼 진폭의 변화에 민감하기 때문에 바람직하지 않다. 대비하여, 스펙트럼 성분의 위상 각 변화는 스펙트럼 성분의 진폭 변화보다 엔빌로프에 더 큰 영향을 미치고, 그래서 엔빌로프를 정의하는 보강과 감산이 상이한 시간에 일어나서 엔빌로프를 변화시킨다. 사람 귀가 어떤 엔빌로프 감도를 갖지만, 귀는 비교적 위상에는 귀머거리여서, 전체적인 사운드 품질은 실제로 유사하게 남는다. 그럼에도 불구하고, 어떤 신호 조건에 대해서, 그러한 진폭 랜덤화가 바람직하지 않은 가청 가공품을 야기하지 않는다면, 스펙트럼 성분의 위상의 랜덤화에 따른 스펙트럼 성분의 진폭의 어떤 랜덤화가 신호 엔빌로프의 촉진된 랜덤화를 제공한다.

바람직하게, 기술 2 혹은 기술 3의 제어 가능한 양이나 정도는 일정한 신호 조건 아래 기술 1에 따라 동작한다. 과도 플래그는 기술 2(과도 플래그가 프레임 혹은 블록으로 보내지는지에 따라, 과도가 프레임 혹은 블록에 나타나지 않는다) 혹은 기술 3(과도가 프레임 혹은 블록을 나타낸다)을 선택한다. 이렇게, 과도의 출현 여부에 따라 여러 모드의 동작이 있다. 대안으로, 일정한 신호 조건 아래 진폭 랜덤화의 제어 가능한 양 혹은 정도는, 역시 원래의 채널 진폭을 보존하도록 추구하는 진폭 스케일링에 따라 동작한다.

기술 2는 집중된 오케스트라 바이올린 같은 하모니가 충만한 복잡한 연속적인 신호에 적합하다. 기술 3은 박수갈채, 캐스터네즈 등과 같은 복잡한 총동적이거나 과도한 신호에 적합하다. 아래 더 설명되듯이, 가청 가공품을 최소화하기 위해, 기술 2와 3이 랜덤화 된 각 변동의 적용에 대한 다른 시간과 주파수 해결책을 갖는다-기술 2는 과도가 나타나지 않을 때 선택되고, 과도가 나타날 때 기술 3이 선택되는 것과 같다.

기술 1은 한 채널에서 빈 각(bin angle)을 천천히 시프트(프레임마다)한다. 이러한 기본적인 시프트의 양이나 정도는 각 제어 파라미터(파라미터가 0이면 시프트 불가)에 의해 제어된다. 아래에 더 설명되듯이, 같거나 보간된 파라미터가 각 서브밴드의 모든 빈에 적용되고 상기 파라미터는 프레임마다 갱신된다. 결과적으로, 각 채널의 각 서브밴드는 다른 채널에 대

해 위상 시프트(전이)를 하여서, 저-주파수(약 1500Hz)에서 한 등급의 상관 해제를 제공한다. 그러나, 기술 1, 그 자체는 그러한 박수갈채 같은 과도 신호에는 부적절하다. 그러한 조건에는, 재생된 채널이 불안정한 콤-필터(comb-filter) 효과를 낸다. 박수갈채의 경우에, 모든 채널이 프레임의 주기상에 동일 진폭을 갖기 쉽기 때문에 복원된 채널의 상대적인 진폭만을 조절함으로써 반드시 상관 해제가 제공되지 않는다.

과도가 나타나지 않을 때 기술 2가 동작한다. 기술 2는 기술 1의 각 시프트에 한 채널에서 개별 빈을 기반으로(각 빈은 다른 랜덤 화 된 시프트를 갖는다) 시간에 따라 변하지 않는 랜덤 화 된 각 시프트를 가산하여, 채널의 엔빌로프가 서로 다르도록 하고, 그리하여 채널 사이에 복잡한 신호의 상관 해제를 제공한다. 랜덤 화 된 위상 각 값을 시간상으로 일정하게 유지하는 것은, 빈 위상 각의 블록-대-블록 혹은 프레임-대-프레임 변경을 초래하는 블록 혹은 프레임 가공품을 피한다. 과도가 나타나지 않을 때 이러한 기술이 매우 유용한 상관 해제 툴인 반면에, 그것은 과도를 임시로 손상한다(이것은 종종 "이전-잡음"으로 언급되고-이후-과도 손상은 과도에 의해 차폐된다). 기술 2에 의해 제공된 추가적인 시프트의 양이나 정도는 상관 해제 스케일 인자(스케일 인자가 0이면)에 의해 직접 스케일된다. 이상적으로, 기술 2에 따른 베이스 각 시프트(기술 1의)에 추가된 랜덤 화 된 위상 각의 양은 가청 신호 가공품을 최소화하는 식으로 상관 해제 스케일 인자에 의해 제어된다. 상관 해제 스케일 인자가 유도되는 방식에서 신호 지저귀는 가공품의 그러한 최소화가 아래 기술되듯이 생긴다. 다른 추가적인 랜덤 화 된 각 시프트 값이 각각의 빈에 인가되어 그 시프트 값이 변하지 않지만, 같은 스케일링이 서브밴드 상에 인가되어 스케일링이 매 프레임 갱신된다.

기술 3은 과도 플래그가 보내지는 비율에 의존하여 프레임 혹은 블록에서 과도의 출현에 동작한다. 그것은 유일한 랜덤 화 된 각도 값을 갖고 블록마다 한 채널에서 각각의 서브밴드에서 모든 빈을 시프트하고, 서브밴드에서 모든 빈에 공통이고, 엔빌로프뿐 아니라 한 채널에서 신호의 진폭과 위상이 블록마다 다른 채널에 대해 변화하도록 한다. 각 랜덤 화의 시간과 주파수 분석의 이러한 변화는 채널 사이에 안정-상태 신호 유사성을 줄이고 "이전-잡음" 가공품을 일으키지 않고 실제로 채널의 상관 해제를 제공한다. 각 랜덤 화의 주파수 분석의 변화는 기술 2에서의 매우 양호(한 채널에서 모든 빈이 다름)에서 기술 3에서의 조잡(서브밴드 내 모든 빈이 동일, 각 서브밴드 다름)까지 "이전-잡음" 가공품을 최소화하는데 특히 유용하다. 귀가 고주파수에서 직접 순수한 각 변화에 응답하지 않지만, 두 개 이상의 채널이 확성기에서 청취자까지 도중에 음향적으로 혼합된 때, 위상 차이는 들을 수 있고 이의 있는 진폭 변화(콤-필터 효과)를 유발하고, 이것은 기술 3에 의해 분쇄된다. 신호의 충격적인 특성은 달리 발생하는 블록 율 가공품을 최소화한다. 이렇게, 기술 3은 한 채널에서 개별 서브밴드를 기반으로 기술 1의 위상 시프트에 빨리 변하는(개별 블록마다) 랜덤 화 된 각 시프트를 추가한다. 아래 기술되듯이, 추가적인 시프트의 양이나 정도는 상관 해제 스케일 인자에 의해 간접으로 스케일된다(scale: 스케일 인자가 0이면 추가적인 시프트가 없다). 동일 스케일링이 서브밴드 상에 인가되어 스케일링이 매 프레임 갱신된다.

각-조정 기술이 3개 기술로 특징지어졌지만, 이것은 의미론의 문제로 역시 두 개 기술로 특징지어진다: (1) 기술 1과 기술 2의 가변 정도의 조합, 이것이 0이다. (2) 기술 1과 기술 3의 가변 정도의 조합, 이것은 0이다. 간단히 나타내도록 기술은 3개 기술로 취급된다.

여러 모드 상관 해제 기술의 특징과 그들의 변경은, 그러한 오디오 채널이 본 발명의 특징에 따라 인코더로부터 유도되지 않을 때조차도 업믹싱(upmixing)에 의하듯이 하나 이상의 오디오 채널로부터 유도된 오디오 신호의 상관 해제 제공시 적용된다. 모노 오디오 채널에 적용될 때, 그러한 장치는 때때로 "의사-스테레오"장치와 기능으로 언급된다. 어떤 적당한 장치 혹은 기능("업믹서")이 모노 오디오 채널 혹은 여러 오디오 채널들로부터 여러 신호를 유도하는데 적용된다. 그러한 여러 오디오 채널이 업-믹서에 의해 유도되면, 여기 기술된 여러 모드 상관 해제 기술을 적용함으로써 그들 중 하나 이상이 다른 유도된 오디오 신호 중 하나 이상에 대해 상관 해제된다. 그러한 응용에서, 상관 해제 기술이 적용되는 각각의 유도된 오디오 채널이 유도된 오디오 채널 자체에서 과도를 검출함으로써 한 모드의 동작에서 다른 것으로 스위칭 된다. 대안으로, 과도-출현 기술의 동작(기술 3)이 간략화되어 과도가 출현할 때 스펙트럼 성분의 위상 각의 시프트를 제공하지 않는다.

사이드체인 정보

위에 언급하였듯이, 사이드체인 정보는 다음을 포함한다.: 진폭 스케일 인자, 각 제어 파라미터, 상관 해제 스케일 인자, 과도 플래그 및 선택적으로 보간 플래그. 본 발명의 특징의 실질적인 실시예에 대한 그러한 사이드 체인 정보는 다음 표 2로 요약된다. 통상적으로, 사이드체인 정보는 프레임당 한번 갱신된다.

표 2

한 채널에 대한 사이드체인 정보 특징

사이드체인 정보	값 범위	표시("-의 측정이다")	양자화 레벨	주요 목적
서브밴드 각 제어 파라미터	$0 \rightarrow + 2\pi$	한 채널에 대한 서브 밴드에서 각 빈의 가과 기준 채널의 서브밴드에서 해당하는 빈의 각 사이에 차이의 평탄하게 된 시간 평균	6비트(64레벨)	채널에서 각각의 빈에 대해 기본적인 각 회전을 제공
서브밴드 상관 해제 스케일 인자	$0 \rightarrow 1$ 스펙트럼 안정 인자와 채널간 각 일치 인자가 낮다면 서브밴드 상관 해제 스케일 인자는 높다	한 채널의 서브밴드에서 시간상 신호 특징의 스펙트럼 안정(스펙트럼 안정 인자)과 기준 채널의 해당 빈에 대해 빈 각의 동일 서브밴드에서 일치(채널간 각 일치 인자)	3비트(8 레벨)	기본적인 각 회전에 추가된 스케일 랜덤 각 시프트와, 적용된다면 역시 기본적 진폭 스케일 인자에 추가된 스케일 랜덤 진폭 스케일 인자 및, 선택적으로 반향의 스케일 정도
서브밴드 진폭 스케일 인자	$0 \sim 31$ (전체 정수) 0은 최고조의 진폭이고, 31은 최저의 진폭이다	모든 채널상의 동일 서브밴드를 위한 에너지나 진폭에 대해 한 채널의 서브밴드의 에너지나 진폭	5비트(32레벨) 알갱이 끝은 1.5dB이고, 그래서 범위는 $31 \times 1.5 = 46.5 \text{ dB}$ + 마지막 값 = 오프.	한 채널에서 서브밴드의 빈의 스케일 진폭
과도 플래그	1, 0 (진/가) (극성은 임의)	프레임에서 혹은 블록에서 과도의 출현	1 비트(2레벨)	랜덤화 된 시프트 혹은 양쪽 각 시프트 및 진폭 시프트를 가산하는데 어떤 기술이 적용되는 지를 결정
보간 플래그	1, 0 (진/가) (극성은 임의)	한 채널 내 위상 각 혹은 서브밴드 경계 가까이 스펙트럼 피크가 선형 진행 구비	1 비트(2 레벨)	기본적인 각 회전이 주파수 상에 보간되는지의 결정

각각의 경우에, 한 채널의 사이드체인 정보는 단일 서브밴드에 인가되고(과도 플래그와 보간 플래그를 제외하고, 각각이 한 채널에서 모든 서브밴드에 적용), 한 프레임마다 한번 갱신된다. 나타난 시간 분석(한 프레임당 한번), 주파수 분석(서브밴드), 값 범위 및 양자화 레벨이 낮은 비트율과 성능 사이에 유용한 성능과 유용한 절충안을 제공하는 것으로 알려져 왔지만, 이러한 시간과 주파수 분석, 값 범위 및 양자화 레벨은 결정적이지 않고, 다른 분석, 범위 및 레벨이 본 발명의 특징을 구현하는데 적용된다는 것을 알 것이다. 예를 들어, 과도 플래그 및/또는 보간 플래그가 적용된다면, 사이드체인 데이터 오버헤드에서 오직 최소 증가로 한 블록마다 한번 갱신된다. 과도 플래그의 경우에, 그렇게 하는 것은 기술 2에서 기술 3으로의 스위칭이 더 정확하다는 이점을 갖는다. 게다가, 위에 언급하였듯이, 관련된 코더의 블록 스위치의 발생시에 사이드체인 정보는 갱신된다.

동일한 서브밴드 상관 해제 스케일 인자가 서브밴드의 모든 빈에 적용될지라도, 위에 기술된(역시 표1 참고) 기술 2는 서브밴드 주파수 분석(즉, 다른 의사 랜덤 위상 각 시프트가 각각의 서브밴드보다는 각 빈에 적용된다)보다는 빈 주파수 분석을 제공한다는 것을 알 것이다. 동일한 서브밴드 상관 해제 스케일 인자가 서브밴드의 모든 빈에 적용될지라도, 위에 기술된(역시 표 1 참고) 기술 3이 블록 주파수 분석(즉, 다른 랜덤 화 된 위상 각 시프트가 각 프레임보다는 각 블록에 적용된다)을 제공한다는 것을 알 것이다. 사이드체인 정보의 분석보다 더 큰 그러한 분석은 랜덤 화 된 위상 각 시프트가 디코더에서 발생하여 인코더에서 알려질 필요가 없기에 가능하다(이것은 인코더가 역시 인코딩된 모노 합성 신호에 랜덤 화 된 위상 각 시프트를 적용할지라도 같은 경우이고, 대안은 아래 기술된다). 환언하면, 상관 해제 기술이 그러한 알갱이 풀을 채택할지라도 빈 혹은 블록 알갱이 풀을 갖는 사이드체인 정보를 보낼 필요가 없다. 예를 들어, 디코더는 하나 이상의 검사표의 랜덤 화 된 빈 위상 각을 채택한다. 사이드체인 정보 레이트보다 더 큰 상관 해제를 위한 시간 및/또는 주파수 분석의 획득은 본 발명의 특징들 사이에 존재한다. 이렇게, 랜덤 화 된 위상을 통한 상관 해제가 시간(기술 2), 혹은 조잡한 주파수 분석(밴드마다)((및 양호한 시간 분석(블록 율))(기술 3)에 따라 변하지 않는 양호한 주파수 분석(빈 마다)으로 실행된다.

랜덤 화 된 위상 시프트의 증가하는 정도가 복원된 채널의 위상 각에 부가됨에 따라, 복원된 채널의 절대 위상 각이 상기 채널의 원래 절대 위상 각과 더욱 달라진다는 것을 역시 알 수 있을 것이다. 랜덤 화 된 위상 시프트가 본 발명의 특징에 따라 추가되는 식으로 신호 조건이 될 때, 복원된 채널의 나오는 절대 위상 각이 원래 채널의 것과 매칭될 필요가 없는 것을 이해하는 것이 본 발명의 한 특징이다. 예를 들어, 상관 해제 스케일 인자가 랜덤 화 된 위상 시프트의 최고치를 초래할 때

의 극한의 경우에, 기술 2 혹은 기술 3에 의해 야기된 위상 시프트는 기술 1에 의해 초래된 기본적인 위상 시프트를 압도한다. 그럼에도 불구하고, 랜덤 화 된 위상 시프트의 어느 정도의 추가를 야기하는 상관 해제 스케일 인자를 일으키는 원래 신호의 다른 랜덤 위상과, 들을 수 있게 동일한 랜덤 화 된 위상 시프트에는 관심이 없다.

위에 언급되었듯이, 랜덤 화 된 진폭 시프트는 랜덤 화 된 위상 시프트에 부가하여 적용된다. 예를 들어, 진폭 조절은 역시 특정 채널에 대한 복원된 사이드체인 과도 플래그 특정 채널에 대한 복원된 사이드체인 상관해제 스케일 인자로부터 유도된 랜덤 화 된 진폭 스케일 인자 파라미터에 의해 제어된다. 그러한 랜덤 화 된 진폭 시프트는 랜덤 위상 시프트의 적용과 유사한 방식으로 두 가지 모드로 동작한다. 예를 들어, 과도 부재시에 시간에 따라 변하지 않는 랜덤 화 된 시프트는 개별 빈을 기반으로(빈마다 다른) 추가되고, 과도의 출현시에(프레임 혹은 블록에서), 랜덤 화 된 진폭 시프트는 개별 블록을 기반으로(블록마다 다른) 변하고, 서브밴드마다 변한다(서브밴드에서 모든 빈에 대해 동일 시프트; 서브밴드마다 다른). 랜덤 화 된 진폭 시프트가 추가되는 양이나 정도가 상관 해제 스케일 인자에 의해 제어되지만, 가청 가공품을 피하기 위해 특정 스케일 인자 값이, 동일한 스케일 인자 값을 내는 해당 랜덤 화 된 위상 시프트보다 더 작은 진폭 시프트를 초래해야 한다고 믿어진다.

과도 플래그(Transient Flag)가 프레임에 인가될 때, 프레임 윌 혹은 블록 윌조차 보다 더 양호한 임시의 분석을 제공하기 위해, 과도 플래그가 기술 2 혹은 기술 3 중에 어느 것을 선택하는 지의 시간 분석은 디코더에서 보충적인 과도 검출기를 제공함으로써 촉진된다. 그러한 보충적인 과도 검출기는 디코더에 의해 수신된 모노 혹은 멀티채널 합성 오디오 신호에서의 과도의 발생을 검출하여, 그러한 검출 정보는 각각의 제어 가능한 상관해제기 (도 2의 38, 42처럼)로 보내진다. 다음에, 그 채널에 대한 과도 플래그의 수신시에, 제어 가능한 상관 해제기는 디코더의 로컬 과도 검출 표시의 수신시에 기술 2에서 기술 3까지 스위칭한다. 이렇게, 임시의 분석에서 사이드체인 비트레이트 증가 없이 실질적인 개선이 가능하다(인코더는 그의 다운-믹싱 전에 각각의 입력 채널에 과도를 검출하고, 이처럼 디코더에서의 검출이 다운-믹싱 후에 이루어진다).

개별 프레임 기반으로 사이드체인 정보를 보내는 대안으로서, 사이드체인 정보는 적어도 아주 동적인 신호에 대해 매 블록 갱신된다. 위에 언급되었듯이, 매 블록 과도 플래그 및/또는 보간 플래그의 갱신은 사이드체인 데이터 오버헤드에서 오직 소량의 증가를 초래한다. 사이드체인 데이터 레이트를 실제로 증가시키지 않고 다른 사이드체인 정보에 대해 임의 분석에서 그러한 증가를 실행하기 위해, 블록-플로팅-포인트(block-floating-point) 미분 코딩 장치가 사용된다. 예를 들어, 연속 변화 블록은 프레임 상에 6그룹으로 모인다. 전체 사이드체인 정보가 제1 블록에서 각 서브밴드-채널에 대해 보내진다. 5개의 다음의 블록에서, 오직 미분 값이 보내지고, 현재 블록 진폭과 각 사이의 차이 각각과 이전 블록으로부터 동일한 값이 보내진다. 이것은 피치 파이프 노트(pitch pipe note) 같은 정적(static) 신호에 대해 매우 낮은 데이터 레이트가 된다. 더 동적인 신호에 대해, 더 큰 범위의 차이 값이 요구되지만 정밀도가 떨어진다. 그래서, 각 그룹의 5개 미분 값에 대해, 지수가 먼저 예를 들어 3비트를 이용하여 보내지고, 다음에 미분 값이 예를 들어 2-비트 정확도로 양자화된다. 이런 장치는 평균 사이드체인 데이터 레이트를 약 2의 인자만큼 줄인다. 위에 언급했듯이, 예를 들어 산술 코딩을 이용하여 기준 채널(이것은 다른 채널로부터 유도될 수 있기 때문에)에 대해 사이드체인 데이터를 생략함으로써 그 이상의 감소가 얻어진다. 대안으로 혹은 게다가, 예를 들어 서브밴드 각 혹은 진폭에서의 차이를 보내므로서 주파수 상의 미분 코딩이 적용된다.

사이드체인 정보가 개별 프레임 기반으로 보내질지 더욱 자주 보내지든 간에, 한 프레임에서 블록 상에 사이드체인 값을 보간하는 것이 유용하다. 시간상으로 선형 보간이, 아래 기술되듯이 주파수 상에 선형 보간의 방식으로 적용된다.

본 발명의 특징 중의 한 가지 적당한 구현은, 다음에 설명되듯이 기능적으로 관련되고 각각의 처리 단계를 구현하는 처리 단계 혹은 장치를 적용한다. 아래 기재된 인코딩과 디코딩 단계 각각이 아래 기재된 단계의 순서로 동작하는 컴퓨터 소프트웨어 명령 시퀀스(sequence)에 의해 실행되지만, 어떤 양은 더 빠른 것에서 유도된다는 것을 고려하여, 동일하거나 유사한 결과가 다른 방식으로 지시된 단계에 의해 얻어지기도 한다는 것을 이해할 것이다. 예를 들어, 일정한 순서의 단계가 병렬로 실행되도록 멀티-스레드(multi-thread) 컴퓨터 소프트웨어 명령 순서가 채택된다. 대안으로, 상기 기술된 단계는 상기 기술된 기능을 실행하는 장치로서 구현되고, 여러 장치는 이후에 기술되듯이 기능과 기능적인 상호관계를 갖는다.

인코딩

인코더가 사이드체인 정보를 유도하고 프레임의 오디오 채널을 하나의 단일 음(모노) 오디오 채널(위에 설명된 도 1의 예의 방식으로), 혹은 여러 오디오 채널(아래 설명된 도 6의 예의 방식으로)로 다운-믹스 하기 전에, 인코더 혹은 인코딩 기능은 프레임의 가치있는 데이터를 모은다. 그렇게 함으로서, 사이드체인 정보는 먼저 디코더로 보내져서, 디코더가 모노 혹은 여러 채널 오디오 정보의 수신시에 즉시 디코딩을 시작하도록 한다. 인코딩 과정의 단계("인코딩 단계")는 다음과 같이 기술된다. 인코딩 단계에 대해, 도 4를 참고하여 이것은 흐름도와 기능 블록도를 합한 성질의 것이다. 단계 419를 통해, 도 4는 하나의 채널에 대한 인코딩 단계를 도시한다. 도 6의 예와 연관되어 아래에 설명되듯이 여러 채널을 제공하도록 함께 매트릭스 되거나, 합성 모노 신호 출력을 제공하도록 결합되는 모든 여러 채널에 단계 420, 421가 적용된다.

단계 401. 과도 검출

- a. 입력 오디오 채널에서 PCM 값의 과도 검출 실행
- b. 과도가 그 채널에 대해 한 프레임의 어떤 블록에 출현한다면, 1-비트 과도 플래그를 참(True)으로 설정.

단계 401에 대한 설명:

과도 플래그는 사이드체인 정보의 일부를 형성하고, 아래 기술되듯이 역시 단계 411에서 사용된다. 디코더에서 블록을 보다 더 양호한 과도 분석이 디코더 성능을 개선한다. 위에 논의했듯이, 프레임-율(frame-rate) 과도 플래그 보다는 블록-율이 비트율의 적당히 증가된 사이드체인 정보의 일부를 형성하지만, 디코더에서 수신된 모노 합성 신호에서 과도의 발생을 검출함으로써 사이드체인 비트율을 증가시키지 않고 감소된 공간적인 정확도를 갖는 유사한 결과가 실행된다.

프레임당 채널당 하나의 과도 플래그가 있는데, 이것이 시간 영역에서 유도되기에 그 채널 내에 모든 서브밴드에 반드시 인가된다. 길고 짧은 길이 오디오 블록 사이에 언제 스위칭할 지의 결정을 제어하기 위해, AC-3 인코더에 적용된 것과 유사한 방식으로 과도 검출이 실행되지만, 블록에 대한 과도 플래그가 참인(AC-3 인코더는 블록 기반으로 과도를 검출한다) 어떤 프레임에 대해 더 높은 감도와 과도 플래그 참(True)을 갖는다. 특히, 상기 인용된 A/52A 서류의 섹션 8.2.2를 보라. 감도 인자 F를 거기에 설정된 식에 추가함으로써 섹션 8.2.2에 기술된 과도 검출의 감도가 증가한다. A/52A 문서의 섹션 8.2.2가 아래 설명되는데, 감도 인자가 부가된다(저주파 필터가 공개된 A/52A 문서에서와 같이 "Form I"보다는 연속의 4차 직접 Form II IIR 필터인 것을 가르치도록 아래 재생된 섹션 8.2.2가 정정된다; 섹션 8.2.2가 더 빠른 A/52A 문서에서 정정되었다). 그것이 결정적이지는 않지만, 0.2의 감도 인자가 본 발명의 특징의 실질적인 실시예에서 적당한 값으로 보여져 왔다.

대안으로, 미국 특허 5,394,473에 기술된 유사한 과도 검출 기술이 채택된다. '473특허는 A/52A 문서의 특징인 과도 검출기를 더 상세히 기술한다. 상기 A/52A 문서와 상기 '473 특허는 그 전체가 여기에 참고로 기술된다.

다른 대안으로서, 과도가 시간 영역보다는 주파수 영역에서 검출된다(단계 408의 주석을 보라). 그런 경우에, 단계(401)는 생략되고, 다른 단계가 아래 기술되듯이 주파수 영역에 적용된다.

단계 402. 원도 및 DFT.

PCM 시간 샘플의 중복 블록을 시간 원도로 곱하고, 그것들을 FFT에 의해 구현된 DFT를 통해 복소수 주파수 값으로 변환하라.

단계 403. 복소수 값을 크기와 각으로 변환

각각의 주파수-영역 복소수 변환 빈(bin) 값(a + jb)을 표준 복소수 조작의 이용으로 크기와 각 표시로 변환하라.

- a. 크기 = 스퀘어_루트($a^2 + b^2$)
- b. 각 + arctan(b/a)

단계 403에 대한 설명:

다음 단계의 일부는 대안으로서 상기 자승 크기로 정위된(즉, 에너지 + $(a^2 + b^2)$), 빈의 에너지를 이용하거나 할 수도 있다.

단계 404. 서브밴드 에너지 계산

- a. 각 서브밴드 내에 빈 에너지값을 가산함으로써 블록당 서브밴드 에너지를 계산하라(주파수 상에 합산).

b. 한 프레임에 모든 블록에서 에너지를 평균하거나 축적함으로써 프레임당 서브밴드 에너지를 계산.(시간상으로 평균/누적)

c 인코더의 커플링 주파수가 약 1000Hz 이하이면, 서브밴드 프레임-평균이거나 프레임-누적된 에너지를 상기 주파수 이하와 커플링 주파수 이상의 모든 서브밴드에서 동작하는 시간 평활기(time smoother)에 인가하라.

단계 404c에 대한 설명:

저-주파수 서브밴드에서 프레임-간 평활기를 제공하기 위한 시간 평활화가 유용하다. 서브밴드 경계에서 빈 값 사이에 가공품 -유발하는 불연속을 피하기 위해, 거의 들을 수 있지만 시간 평활화 효과가 측정 가능하지만 들을 수 없는 더 높은 주파수 서브밴드를 통해 커플링 주파수(평활화가 커다란 효과를 가짐)와 둘러싸는 가장 낮은 주파수 서브밴드로부터 점차로 줄어드는 시간 평활화를 적용하는 것이 유용하다. 가장 낮은 주파수 범위 서브밴드(서브밴드가 임계 밴드라면 서브밴드는 단일 빈이다)에 대한 적당한 시정수가 예를 들어, 50~ 100 밀리 초의 범위에 있다. 점차로 감소하는 시간 평활화가 약 1000Hz를 포함하는 서브밴드를 통해 계속되고 여기서 시정수는 예를 들어, 약 10 밀리 초이다. 제 1차 평활기가 적당하지만, 평활기는 과도에 응답하여 시간을 감소하고 그의 공격을 줄이는 가변 시정수를 갖는 두 단계의 스무더이다(그러한 두 단계의 평활기는 미국 특허 3,846,719와 4,922,535에 기술된 아날로그 두-단계 평활기의 디지털 등가물이고, 그 각각은 여기에 그 전체가 기술된다). 환언하면, 안정된 상태 시정수는 주파수에 따라 스케일(scale)되고, 역시 과도에 응답하여 가변적이다. 대안으로, 그러한 평활화는 단계 412에 적용된다.

단계 405. 빈 크기의 합 계산

a. 각 서브밴드의 빈 크기(단계 403)의 블록당 합 계산(주파수 상의 합산).

b. 프레임에서 블록 상에 단계 405의 크기를 평균하거나 누적(시간상에 평균/누적)함으로써 각 서브밴드의 빈 크기의 프레임당 합을 계산, 상기 합은 아래 단계 410에서 채널 간 각 일치 인자를 계산하는데 이용된다.

c. 인코더의 커플링 주파수가 약 1000 Hz이하라면, 커플링 주파수 이상이고 상기 주파수 이하의 모든 서브밴드에서 동작하는 시간 평활기에 서브밴드 프레임-평균 혹은 프레임-누적된 크기를 인가하라.

단계 405c에 대한 설명:

단계 405c의 경우에 시간 평활화는 단계 410의 부분으로 실행된다는 것을 제외하고 단계 404c에 대한 해설을 보라.

단계 406 상대적인 채널 간 빈 위상 각 계산.

단계 403의 빈 각에서 기준 채널(예를 들어, 제1 채널)의 해당 빈 각을 빼서, 각 블록이 각 변환 빈의 상대적인 채널 간 위상 각을 계산하라.

다른 각 가산 혹은 여기서는 감산으로 그 결과는 $-\pi$ 내지 $+\pi$ 의 필요한 범위 내일때까지 2π 를 더하거나 빼서 modulo($\pi, -\pi$) 래디안으로 취해진다.

단계 407. 채널간 서브밴드 위상 각 계산.

각 채널에 대해, 다음처럼 각 서브밴드에 대한 프레임-울 진폭-측정된 채널 간 위상 각 계산:

a. 각 빈에 대해, 단계 406의 상대적인 채널간 빈 위상 각과 단계 403의 크기로부터 복소수 구성

b. 각 서브밴드 상에 단계 407a의 구축된 복소수 가산(주파수 상에 합산)

단계 407b에 대한 설명:

예를 들어, 서브밴드가 2번을 갖고 빈 중에 하나가 $1 + j1$ 의 복소수 값을 갖고 다른 빈이 $2 + j2$ 의 복소수 값을 갖으면, 그 두 복소수 합은 $3 + j3$ 이다.

- c. 각 프레임의 블록 상에 단계 407b의 각 서브밴드에 대해 블록당 복소수 합의 평균 혹은 누적(시간상에 평균 혹은 누적).
- d. 인코더의 커플링 주파수가 약 1000Hz이하 라면, 서브밴드 프레임 평균 혹은 프레임-누적의 복소수를 상기 주파수 이하와 커플링 주파수 이상의 모든 서브밴드에서 동작하는 시간 평활기에 적용하라.

단계 407d에 대한 설명:

시간 평활화가 단계 407e 혹은 410의 일부로서 실행되는 단계 407d의 경우를 제외하고 단계 404c에 대한 해설을 보라.

- e. 단계 403당 단계 407d의 복잡한 결과의 크기 계산.

단계 407e에 대한 설명: 이 크기는 아래 단계 410a에서 이용된다.

단계 407b로 주어진 간단한 예에서, $3 + j3$ 의 크기는 $\sqrt{9 + 9} = 4.24$ 이다.

- f. 단계 403당 복잡한 결과의 각도 계산.

단계 407f에 대한 설명: 단계 407b에서 주어진 간단한 예에서, $3 + j3$ 의 각은 $\arctan(3/3) = 45^\circ = \pi/4$ 라디안이다. 상기 서브밴드 각은 신호-의존의 시간-스무드되고(단계 41을 보라) 양자화 되어(단계 414를 보라), 아래 기술되듯이 서브밴드 각 제어 파라미터 사이드체인 정보를 생성한다.

단계 408. 빈(bin) 스펙트럼 안정 인자 계산

각 빈에 대해, 다음처럼 0 ~1의 범위에서 빈 스펙트럼-안정 인자를 계산하라:

- a. x_m = 단계 403에서 계산된 현재 블록의 빈 크기로 놓는다.

- b. y_m = 이전 블록의 해당 빈 크기로 놓는다.

c. $x_m > y_m$ 라면, 빈 동적인 진폭 인자 = $(y_m/x_m)^2$;

d. 그외 $y_m > x_m$ 라면, 빈 동적인 진폭 인자 = $(x_m/y_m)^2$,

- e. 그외 $y_m = x_m$ 라면, 빈 스펙트럼-안정 인자 = 1.

단계 408에 대한 설명:

"스펙트럼 안정"은 스펙트럼 성분(예를 들어, 스펙트럼 계수 혹은 빈 값)이 시간상으로 변하는 범위의 측정이다. 1의 빈 스펙트럼 안정 인자는 주어진 시간상에 불변을 가르친다.

또한 스펙트럼 안정은 과도가 출현할 지의 표시기로서 취해진다. 과도는 하나 이상의 블록의 시간 주기상에 스펙트럼(빈) 진폭에서 갑작스런 상승 및 하락을 일으키고, 블록과 그들의 경계에 대한 위치에 좌우된다. 결과적으로, 작은 수의 블록 상에 높은 값에서 낮은 값으로 빈 스펙트럼-안정 인자에서의 변화는, 더 낮은 값을 갖는 블록 혹은 블록들에서 과도의 출현의 표시로 취해진다. 과도 출현의 확인 또는 빈 스펙트럼 안정 인자를 채택하는 대안은 블록 내에 빈의 위상 각을 관찰하는 것이다(예를 들어, 단계 403의 위상 각 출력에서). 과도가 블록 내에 단일 임시 위치를 점유할 것 같고 블록에서 현저한 에너지를 갖기에, 과도의 존재와 위치는 블록에서 빈마다 위상의 일정한 지연으로 나타난다-즉, 주파수의 함수로서 실제로 선형 기술기의 위상 각. 아직도 상기 확인 혹은 대안은 작은 수의 블로 상에 빈 진폭을 관찰하는 것이다(예를 들어, 단계 403의 크기 출력에서)-즉 스펙트럼 레벨의 갑작스런 상승과 하락에 대해 직접 봄으로서.

대안으로, 단계 408은 한 블록 대신에 3개의 연속 블록을 본다. 인코더의 커플링 주파수가 약 1000Hz 이하라면, 단계 408은 3개 이상의 연속 블록을 본다. 서브밴드 주파수 범위가 감소하는 대로 그 수가 점차로 증가하는 식으로, 연속 블록의 수는 주파수에 따라 변하는 것으로 고려된다. 빈 스펙트럼 안정 인자가 하나 이상의 블록에서 얻어진다면, 방금 기술하듯이 과도를 검출하는데 유용한 블록의 수에 오직 응답하는 개별 단계에 의해 과도의 검출이 결정된다.

다른 대안으로서, 빈(bin) 에너지가 빈 크기 대신에 이용된다.

또 다른 대안으로서, 다음 단계 409 해설에서 아래 기술되듯이, 단계 408이 "이벤트 결정" 검출 기술을 채택한다.

단계 409. 서브밴드 스펙트럼-안정 인자 계산.

다음과 같이 프레임에서 블록 상에 각 서브밴드 내에 빈 스펙트럼 안정 인자의 크기-측량된 평균을 형성함으로써 0~1의 스케일 상에서 프레임-율 서브밴드 스펙트럼 안정 인자를 계산.

- a. 각각의 빈에 대해, 단계 403의 빈 크기와 단계 408의 빈 스펙트럼 안정 인자의 적(積) 계산.
- b. 각각의 서브밴드 내에 적 계산(주파수 상에 합산).
- c. 프레임에서 모든 블록에 단계 409b의 합산을 평균 혹은 누적(시간상에 평균/누적).
- d. 인코더의 커플링 주파수가 약 1000Hz 이하이면, 커플링 주파수 이상 및 그 주파수 이하의 모든 서브밴드에서 작동하는 시간 평활기에 서브밴드 프레임-평균의 혹은 프레임-누적된 합산을 적용.

단계 409d에 대해 설명:

단계 409d의 경우에 시간 평활화가 대안으로 실행되는 이후의 적당한 단계가 없는 것을 제외하고 단계 404c에 대한 주석을 보라.

- e. 단계 409c 혹은 단계 409d의 결과를 서브밴드 내에 빈 크기의 합(단계 403)으로 나누라.

단계 409e에 대한 설명:

단계 409a에서 크기에 의한 승산과 단계 409e에서 크기의 합에 의한 제산은 진폭 측정을 제공한다. 단계 408의 출력은 절대 크기에 독립적이고, 진폭 측정이 되지 않으면, 출력 혹은 단계 409가 매우 작은 진폭으로 제어될 수 있고 이것은 바람직하지 않다.

- f. {0.5..1} 내지 {0...1} 범위를 매핑함으로써 서브밴드 스펙트럼-안정 인자를 얻는 결과를 스케일. 이것은 2에 의한 결과를 승산, 1을 감산, 0의 값으로 0 보다작은 결과를 제한함으로써 이루어진다.

단계 409f에 대한 설명: 채널의 잡음이 0의 서브밴드 스펙트럼-안정 인자가 되는데있어 단계 409f가 유용하다.

단계 408과 409에 대한 설명:

단계 408, 409의 목표는 스펙트럼-안정을 측정하는 것이다-한 채널의 서브밴드에서 시간상으로 스펙트럼 성분의 변화. 대안으로, 국제 특허 공개 번호 WO 02/097792 A1(미국 지정)에 기술된 것 같은 "이벤트 결정" 감지의 특징은 단계 408, 409와 연관되어 방금 기술된 접근 대신에 스펙트럼-안정을 측정하도록 채택된다. 2003년 11월 20일 출원된 미국 특허 출원 S.N. 10/478,538호는 공개된 PCT 출원 WO 02/097792 A1의 미국 국내 출원이다. 공개된 PCT 출원과 미국 출원이 여기에 전체가 참고로 기술된다. 참고된 출원에 따르면, 각각의 빈의 복소수 FFT 계수의 크기가 계산되어 정규화된다(예를 들어, 가장 큰 크기는 1의 값으로 설정된다). 다음에 연속 블록의 해당 빈의 크기(dB로), 빈 사이의 차이가 합산되고, 합산이 임계치를 초과하면 블록 경계는 청각의 이벤트 경계선으로 여겨진다. 대안으로, 블록마다의 진폭의 변화는 역시 스펙트럼 크기 변화를 따라 여겨진다(필요한 정규화의 양을 봄으로서).

기술된 이벤트-감지 출원의 특징이 스펙트럼-안정을 측정하도록 채택된다면, 정규화는 필요하지 않고 스펙트럼 크기의 변화(정규화가 생략되면 크기 변화가 측정되지 않는다)는 바람직하게 서브밴드 기반으로 고려된다. 위에 나타난 단계 408 실행 대신에, 해당하는 빈 사이의 스펙트럼 크기의 데시벨 차이가 상기 출원의 가르침에 따라 합산된다. 다음에, 블록마다 스펙트럼 변화의 정도를 나타내는 그런 합산의 각각은 스케일되어 그 결과가 0~1의 범위를 갖는 스펙트럼-안정 인자이고, 여기서 1의 값은 최고조의 안정, 주어진 빈에 대한 블록당 0dB의 변화를 나타낸다. 가장 낮은 안정을 나타내는 0의 값은 예를 들어, 12dB 같은 적당한 양이거나 더 큰 데시벨 변화로 할당된다. 이러한 결과인 빈 스펙트럼-안정 인자는, 단계 409가 상기의 단계 408의 결과를 이용하는 동일한 방식으로 단계 409에 의해 이용된다. 방금 기술한 다른 이벤트 결정 감지 기술을 채택함으로써 얻어진 빈 스펙트럼-안정 인자를 단계 409가 수신할 때, 단계 409의 서브밴드 스펙트럼-안정 인자는 역시 과도의 표시기로서 사용된다. 예를 들어, 단계 409에 의해 나온 값의 범위가 0~1이라면, 서브밴드 스펙트럼-안정 인자가 실질적인 스펙트럼-불안정을 나타내는, 예를 들어 0.1같이 작은 값일 때 과도는 출현하는 것으로 여겨진다.

단계 408에 대한 방금 기술한 대안에 의해 및 단계 408에 의해 나온 빈 스펙트럼-안정 인자 각각이, 그것들이 블록마다 상대적인 변화에 근거하는 점에서 일정한 정도로 가변 임계치를 제공한다. 선택적으로, 예를 들어 프레임에서 여러 과도 혹은 더 작은 과도 사이에 큰 과도(예를 들어, 중간에서 낮은 레벨로의 박수갈채로의 커다란 과도)에 응답하여 임계치에서 시프트를 제공함으로써 그러한 유산을 보충하는 것이 유용하다. 나중의 예의 경우에, 이벤트 검출기는 초기에 이벤트로서 각각의 찰칵 소리를 식별하지만, 큰 과도(예를 들어, 드럼치는 소리)는 드럼치는 소리만이 이벤트로서 식별되도록 임계치를 시프트하는 것이 바람직하게 만든다.

대안으로, 시간상으로 스펙트럼-안정의 측정 대신에 랜덤 매트릭스가 채택된다(예를 들어, 미국 특허 Re 36,714에 기술되듯이, 여기에 그 전체가 참고로 기술된다).

단계 410. 채널 간 각 일치 인자 계산.

하나 이상의 빈을 갖는 각 서브밴드에 대해, 다음처럼 프레임-레이트 채널 간 각 일치 인자를 계산:

- a. 단계 407e의 복소수 합의 크기를 단계 405의 크기의 합으로 나눔. 나오는 "원래" 각 일치 인자는 0~1 범위의 수이다.
- b. 정정 인자 계산: n = 상기 단계에서 2개 양에 기여하는 서브밴드 상에 값의 수로 놓는다(환언하면, " n "은 서브밴드에서 빈의 개수이다). n 이 2보다 작다면 각 일치 인자는 1로 놓고 단계 411, 413으로 진행한다.
- c. r = 예측된 랜덤 변동 = $1/n$ 으로 놓는다. 단계 410b의 결과에서 r 을 뺀.
- d. $(1-r)$ 로 나누어서 단계 410c의 결과를 정규화. 그 결과는 1의 최대값을 갖는다. 필요한 데로 최소값을 0으로 제한.

단계 410에 대한 설명:

채널 간 위상 각이 프레임 주기상에 서브밴드 내에 있는 것이 얼마나 유사한지의 측정이 채널 간 각 일치이다. 서브밴드의 모든 빈 채널 간 각이 같다면, 채널 간 각 일치 인자가 1.0이다; 이처럼, 채널 간 각이 임의로 퍼지면, 값이 0에 접근한다.

서브밴드 각 일치 인자가 채널 사이에 유령 이미지가 있는지를 가르친다. 일치이 낮으면, 채널을 상관 해제하는 것이 바람직하다. 높은 값은 융합된 이미지를 나타낸다. 이미지 융합은 다른 신호 특성과 무관하다.

비록 각 파라미터지만, 서브밴드 각 일치 인자가 두 개 크기로부터 간접으로 결정된다는 것을 알 수 있다. 채널 간 각이 모두 같지만, 복소수 값을 가산하여 크기를 갖는 것이 모든 크기를 취하여 그들을 가산하는 것과 동일한 결과를 낳아서, 몫이 1이다. 채널 간 각이 퍼지면, 복소수 값 가산(다른 값을 갖는 벡터 가산 값은) 적어도 부분적인 소거가 되고, 그래서 합의 크기가 크기의 합보다 더 작고, 몫은 1보다 작다.

다음은 두 개 빈을 갖는 서브밴드의 간단한 예이다:

두 개 복소수 빈 값이 $(3 + j4)$ 와 $(6 + j8)$ 인 것으로 가정. (동일 각 각각의 경우: 각 = $\arctan(\text{허수}/\text{실수})$, 그래서 각 1 = $\arctan(4/3)$ 와 각 2 = $\arctan(8/6) = \arctan(4/3)$). 복소수 값 가산하면, 합 = $(9 + j12)$, 이것의 크기는 자승_루트($81 + 144$) = 15이다.

크기의 합은 $(3 + j4)$ 의 크기 + $(6 + j8)$ 의 크기 = $5 + 10 = 15$ 이다. 그러므로, 몫은 $15/15 = 1 =$ 일치($1/n$ 정규화전에, 정규화 후에 역시 1이다)(정규화된 일치 = $(1 - 0.5)/(1 - 0.5) = 1.0$).

상기 빈 중에 하나가 다른 각도를 갖는다면, 제 2의 것은 복소수 값 $(6 - j8)$ 을 갖는다고 말하고, 이것은 동일 크기 10을 갖는다. 복소수 합은 이제 $(9 - j4)$ 로, 이것은 자승_루트 $(81 + 16) = 9.85$ 의 크기를 갖고, 그래서 몫은 $9.85/15 = 0.66 =$ 일치(정규화 전에)이다. 정규화하도록, $1/n = 1/2$ 감산하여 $(1 - 1/n)$ (정규화된 일치 = $(0.66 - 0.5)/(1 - 0.5) = 0.32$)로 계산.

서브밴드 각 일치 인자를 결정하는 상기의 기술이 유용하다고 보여 지지만, 그의 이용은 결정적이지 않다. 다른 적당한 기술이 채택되기도 한다. 예를 들어, 누군가 표준 공식을 이용하여 각의 표준 편차를 계산할 수 있었다. 어떤 경우에, 계산된 일치 값 상에서 작은 신호의 효과를 최소화하도록 진폭 측정을 채택하는 것이 바람직하다.

게다가, 서브밴드 각 일치 인자의 다른 편차가 크기 대신에 에너지(크기의 자승)를 사용한다. 이것은 단계 405, 407에 적용되기 전에 단계 403에서 크기를 제공함으로써 실행된다.

단계 411. 서브밴드 상관 해제 스케일 인자 유도.

다음과 같이 각 서브밴드에 대한 프레임-울 상관 해제 스케일 인자를 유도:

- 단계 409f의 $x =$ 프레임-울 스펙트럼-안정 인자로 놓는다.
- 단계 410e의 $y =$ 프레임-울 각 일치 인자로 놓는다.
- 다음에 프레임-울 서브밴드 상관 해제 스케일 인자 = $(1 - x) * (1 - y)$, 0과 1 사이의 수.

단계 411에 대한 설명:

서브밴드 상관 해제 스케일 인자는 기준 채널의 대응하는 빈에 대한 한 채널의 빈 각의 같은 서브밴드에서의 일치와(채널 간 각 일치 인자), 한 채널의 서브밴드에서 시간상으로 신호 특성의 스펙트럼-안정의 함수(스펙트럼-안정 인자)이다. 스펙트럼-안정 인자와 채널 간 각 일치 인자가 낮다면 서브밴드 상관 해제 스케일 인자가 높다.

위에 설명되듯이, 상관 해제 인자는 디코더에서 제공된 엔빌로프 상관 해제의 정도를 제어한다. 시간상으로 스펙트럼-안정을 나타내는 신호는 다른 채널에 무엇이 발생하는 지에 관계없이 바람직하게 그들의 엔빌로프를 수정함으로써 디코리레이트 되지 말아야 하고, 그것은 들을 수 있는 가공품, 즉 신호의 지저귐 혹은 흔들림이 된다.

단계 412. 서브밴드 크기 스케일 인자 유도

모든 채널의 서브밴드 프레임 에너지값으로부터 및 단계 404의 서브밴드 프레임 에너지 값으로부터(이것은 단계 404에 해당하는 단계에 의해 얻어지거나), 다음과 같이 프레임-울 서브밴드 크기 스케일 인자를 유도하라:

- 각 서브밴드에 대해, 모든 입력 채널 상에 프레임당 에너지값을 합산.
- 0 ~ 1의 범위에서 값을 내도록 프레임당 각 서브밴드 에너지값을 모든 입력 채널(단계 412a로부터) 상에 에너지값의 합으로 나눈다(단계 404로부터).
- $-\infty \sim 0$ 의 범위에서 각 비율을 dB로 변환.
- 스케일 인자 알갱이 꼴로 나누고, 이것은 1.5dB로 설정되고, 예를 들어 비-마이너스 값이 되도록 신호를 변경하여 최대치로 제한하고, 이것은 예를 들어 31(즉 5-비트 정밀도)이고, 양자화된 값을 내도록 가장 가까운 정수로 라운딩한다. 이런 값은 프레임-울 서브밴드 크기 스케일 인자로서 사이드체인 정보의 일부로서 운반된다.
- 인코더의 커플링 주파수가 약 1000Hz라면, 서브밴드 프레임-평균이거나 프레임-누적된 크기를, 커플링 주파수 이상과 그 주파수 이하의 모든 서브밴드에서 동작하는 시간 평활기에 적용.

단계 412e에 관한 설명: 시간 평활화가 대안으로 실행되는 적절한 후속 단계가 없는 단계 412e의 경우를 제외하고 단계 404c에 대한 해설을 보라.

단계 412에 대한 설명:

여기 나타난 알갱이 풀(분석)과 양자화 정밀도가 유용하다고 보여 지지만, 그것은 결정적이지 않고 다른 값은 수용 가능한 결과를 제공한다.

대안으로, 서브밴드 크기 스케일 인자를 발생하도록 에너지 대신에 진폭을 이용한다. 진폭을 이용한다면, 에너지를 이용할 지라도 $dB = 20 * \log(\text{진폭 비율})$ 을 이용하고, dB를 $dB = 10 * \log(\text{에너지 비율})$ 로 변환하고, 여기서 진폭 비율 = 제곱 루트(에너지 비율).

단계 413. 신호-의존의 시간 평활 채널 간 서브밴드 위상 각.

신호-의존의 임시 평활화를 단계 407f에서 유도된 서브밴드 프레임-을 채널 간 각에 적용:

- a. v = 단계 409d의 서브밴드 스펙트럼-안정 인자로 놓는다.
- b. w = 단계 410e의 해당하는 각 일치 인자로 놓는다.
- c. $x = (1-v)*w$ 로 놓는다. 이것은 0과 1 사이의 값이고 이것은 스펙트럼-안정 인자가 낮고 각 일치 인자가 높으면 높다.
- d. $y = 1-x$ 로 놓고, 스펙트럼-안정 인자가 높고, 각 일치 인자가 낮다면 y 는 높다.
- e. $z = y^{\text{exp}}$ 로 놓고, 여기서 exp 는 상수이고, 이것은 0.1이다. z 는 역시 0 ~1의 범위에 있지만 1을 향해 비스듬하고 늦은 시정수에 해당한다.
- f. 채널에 대한 과도 플래그(단계 401)가 설정된다면, $z = 0$ 으로 설정, 과도의 출현에 빠른 시정수에 해당.
- g. lim 계산, z 의 최대 허용값 $\text{lim} = 1-(0.1*w)$. 이것은 각 일치 인자가 1.0으로 높으면 이것은 0.9로부터의 범위이다.
- h. 필요한 데로 lim 으로 z 제한: ($z > \text{lim}$)라면 $z = \text{lim}$ 이 됨.
- i. 각 서브밴드에 대해 유지된 각의 평활하게 된 값과 z 의 값을 이용하여 단계 407f의 서브밴드 각 평활화. 단계 407f의 각 = A , $\text{RSA} =$ 이전 블록의 구동되는 평활하게 된 각 값이고, 새로운 RSA 가 구동되는 평활하게 된 각의 새로운 값이라면: 새로운 $\text{RSA} = \text{RSA}*z + A*(1-z)$. 다음 블록을 처리하기 전에 RSA 의 값은 새로운 RSA 와 같게 설정된다. 새로운 RSA 는 단계 413의 단일-의존의 시간-평활하게 된 각 출력이다.

단계 413에 대한 설명:

과도가 검출될 때, 서브밴드 각 갱신 시간 상수가 0으로 설정되어 빠른 서브밴드 각 변화를 허용한다. 이것은 빨리-변하는 신호가 빠른 시정수로 처리되는 반면, 정적이거나 준-정적 신호 동안에 이미지를 최소화하는 비교적 늦은 시정수의 범위를 사용하도록 정상적인 각 갱신 메커니즘을 허용하기에 이것은 바람직하다.

다른 평활화 기술과 파라미터가 이용 가능하지만, 단계 413을 구현하는 제 1차 평활기가 적당하다고 여겨져 왔다. 제 1차 평활기/저역 통과 필터로 구현된다면, 변수 " z "는 선행-공급 계수(때때로 "ff0"로 표시)에 해당하는 반면에, " $(1-z)$ "는 궤환 계수(때때로 "ff1"로 표시)에 해당한다.

단계 414. 평활하게 된 채널간 서브밴드 위상 각의 양자화.

단계 413에서 유도된 시간-평활하게 된 서브밴드 채널간 각을 양자화하여 서브밴드 각 제어 파라미터를 얻음:

- a. 값이 0보다 작다면, 양자화될 모든 각의 값이 0 내지 2π 의 범위가 되도록 2π 가산.
- b. 각 알갱이 꼴로 나누고(분해), 이것은 $2\pi/64$ 라디안이고 정수로 라운딩. 최대값은 6-비트 양자화에 해당하는 63으로 설정된다.

단계 414에 대한 설명:

양자화된 값이 비-마이너스 정수로 처리되어, 각을 양자화하는 쉬운 방법은 비-마이너스 플로팅 포인트 수로 매핑하는 것으로(0보다 작으면 2π 를 가산하여, 0 내지 2π 의 범위를 만들고), 알갱이 꼴로 스케일링하여 정수로 라운딩한다. 유사하게, 그 정수의 역양자화(이것은 달리 간단한 테이블 검사로 이루어짐)는, 각 그래뉴러티 인자의 역을 스케일링하고, 비-마이너스 정수를 비-마이너스 플로팅 포인트 각(다시, 0 ~ 2π 의 범위)으로 변환하여서 이루어진다. 서브밴드 각 제어 파라미터의 그러한 양자화가 유용하다고 여겨졌지만, 그런 양자화는 결정적이지 않아서 다른 양자화가 수용 가능한 결과를 제공한다.

단계 415. 서브밴드 상관 해제 스케일 인자의 양자화

단계 411로 발생된 서브밴드 상관 해제 스케일 인자를 7.49를 곱하고 가장 근접한 정수로 라운딩하여 예를 들어, 8레벨(3비트)로 양자화한다. 이렇게 양자화된 값은 사이드체인 정보의 일부이다.

단계 415에 대한 설명:

서브밴드 상관 해제 스케일 인자의 그런 양자화가 유용하다고 여겨졌지만, 그런 예를 이용한 양자화는 결정적이지 않아서 다른 양자화가 수용 가능한 결과를 제공한다.

단계 416. 서브밴드 각 제어 파라미터의 역양자화.

다운믹싱 하기 전에 서브밴드 각 제어 파라미터(단계 414를 보라)를 역양자화한다.

단계 416에 대한 설명:

인코더와 디코더 사이의 조화를 유지하도록 돕는 인코더의 양자화된 값의 이용.

단계 417. 블록 상에 프레임-울 역양자화 된 서브밴드 각 제어 파라미터의 분배.

다운믹싱에 대한 준비로, 프레임 내에 각 블록의 서브밴드에 시간상 단계 416의 프레임당 한번 역양자화된 서브밴드 각 제어 파라미터를 분배.

단계 417에 대한 설명:

동일 프레임 값이 프레임의 각각의 블록에 할당된다. 대안으로, 프레임에서 블록 상에 서브밴드 각 제어 파라미터 값을 보간하는 것이 유용하다. 시간상의 선형 보간은 아래 기술되듯이, 주파수 상의 선형 보간의 방식으로 적용된다.

단계 418. 블록 서브밴드 각 제어 파라미터를 빈으로 보간

아래 기술된 선형 보간을 이용하여, 주파수 상에 각각의 채널에 대한 단계 417의 블록 서브밴드 각 제어 파라미터를 빈으로 분배.

단계 418에 대한 설명:

주파수 상의 선형 보간이 적용되면, 단계 418은 서브밴드 상에 빈마다 위상 각 변화를 최소화한다. 그런 선형 보간은 예를 들어, 단계 422의 다음 설명을 따라서 가능하게 된다. 서브밴드 각은 서로에 무관하게 계산되어, 각각이 서브밴드 상에 평균을 나타낸다. 이렇게, 서브밴드마다 커다란 차이가 있다. 서브밴드에 대한 전체 각도 값이 서브밴드에서 모든 빈에 적용되면("장방향" 서브밴드 분배), 한 서브밴드에서 인접 서브밴드까지의 전체 위상 변화가 두 개 빈 사이에 일어난다. 거기에

강한 신호 성분이 있다면, 심한, 들을 수 있는 앨리어싱이 있다. 예를 들어, 각 서브밴드의 중심 사이에 선형 보간은 서브밴드에서 모든 빈 상에 걸쳐 위상 각 변화를 뿌려서, 빈의 어느 쌍 사이의 변화를 최소화한다. 환언하면, 장방향 서브밴드 분배 대신에, 서브밴드 각 분배가 사다리꼴 모양이다.

예를 들어, 가장 낮은 연결된 서브밴드가 한 개 빈과 20도의 서브밴드 각도를 갖는다면, 다음 서브밴드는 3개 빈과 40도의 서브밴드 각을 갖고 제 3 서브밴드는 5개 빈과 100도의 서브밴드 각을 갖는다. 보간 없이, 제1 빈(한 서브밴드)이 20도의 각도로 시프트 되면 다음 3개 빈은(다른 서브밴드) 40도의 각도로 시프트되고 다음 5개 빈은(또 다른 서브밴드) 100도의 각도로 시프트된다. 상기 예에서, 빈 4에서 빈 5까지 60도의 최대 변화가 있다. 선형 보간으로, 여전히 제1 빈이 20도 만큼 시프트되고, 다음 3개 빈이 약 30, 40, 50도 만큼 시프트되고, 다음 5개 빈이 약 67, 83, 100, 117 및 133도 만큼 시프트된다. 평균 서브밴드 각 시프트가 같지만, 최대 빈-대-빈 변화가 17도로 줄어든다.

선택적으로, 여기에 기술된 단계와 연관된 단계 417 같은 서브밴드 마다의 크기 변화가 역시 유사한 보간의 방식으로 취급된다. 그러나, 한 서브밴드에서 다음까지 더욱 자연스러운 연속성 때문에 그렇게 할 필요가 없다.

단계 419. 위상 각 회전을 채널에 대한 빈 변환 값에 적용.

다음처럼 각각의 빈 변환 값에 위상 각 회전 적용.

- a. x = 단계 418에서 계산된 빈에 대한 빈 각도를 놓는다.
- b. $y = -x$ 로 놓는다;
- c. z , 각 y 를 갖는 1-크기의 복소수 위상 회전 스케일 인자 계산, $z = \cos(y) + j \sin(y)$.
- d. 빈 값($a + jb$)과 z 를 곱함.

단계 419에 대한 설명:

인코더에 적용된 위상 각 회전은 서브밴드 각 제어 파라미터로부터 유도된 각도의 역이다.

여기 기술되듯이, 인코더 혹은 인코딩 과정에서 다운믹싱 (단계 420)하기 전에 위상 각 조절은 여러 장점을 갖는다: (1) 여러 채널에 매트릭스 되거나 모노 합성 신호로 합산되는 채널의 소거를 최소화하고, (2) 에너지 정규화(단계 421) 상의 의존도를 최소화, (3) 디코더 역 위상 각 회전을 미리 보상하여, 앨리어싱(aliasing)을 줄인다.

서브밴드에 각각의 변환 빈 값의 각에서 각각의 서브밴드 위상 정정 값을 감산하여 위상 정정 인자가 인코더에 적용될 수 있다. 이것은 위상 정정 인자의 네가티브와 같은 각과 1.0의 크기의 복소수로 각 복소수 빈 값을 곱한 것이 같다. 크기 1의 복소수, 각도 A 가 $\cos(A) + j \sin(A)$ 와 같다는 것을 알라. 후자의 양은 $A = -$ 상기 서브밴드에 대한 위상 정정으로 각 채널의 각 서브밴드에 대해 한번 계산되어, 각각의 빈 복소수 신호 값으로 곱해서 위상 전이된 빈 값이 된다.

상기 위상 전이는 원형으로, 원형 컨벌루션 (상기 언급)이 된다. 원형 컨벌루션이 어떤 연속적인 신호에 대해 양성인 동안, 일정한 연속적인 복잡한 신호(피치 파이프 같은)에 대한 스펙트럼 성분을 만들거나, 다른 위상 각이 다른 서브밴드에 이용된다면 과도의 퍼짐을 유발한다. 결과적으로, 원형 컨벌루션을 피하는 적당한 기술이 적용되거나 과도 플래그가 참일 때 예를 들어, 각 계산 결과가 중복되는 식으로 과도 플래그가 적용되고, 채널의 모든 서브밴드가 제로 혹은 랜덤 값 같은 동일 위상 정정 인자를 사용한다.

단계 420. 다운믹스.

아래 기술되듯이, 도 6의 예의 방식으로 입력 채널을 매트릭스 하여 여러 채널로 다운믹스하거나, 모노 합성 채널을 발생하도록 채널 상에 해당 복소수 변환 빈을 부가하여 모노로 다운믹스.

단계 420에 대한 설명:

인코더에서, 모든 채널의 변환 빈이 위상 전이되면, 채널이 빈마다 합산되어 모노 합성 오디오 신호를 만든다. 대안으로, 도 1의 N: 1 인코딩처럼 한 채널 혹은 여러 채널로 간단한 합산을 제공하는 수동 혹은 능동 매트릭스에 채널이 인가된다. 매트릭스 계수는 실수 혹은 복소수(실수 및 허수)이다.

단계 421. 정규화하다.

동상(in-phase) 신호의 오버-엠퍼시스(over-emphasis)와 고립된 빈의 소거를 피하기 위해, 기여되는 에너지의 합과 실제로 같은 에너지를 갖도록 다음처럼 모노 합성 채널의 각각의 빈의 크기를 정규화:

- a. x = 빈 에너지의 채널 상의 합(즉, 단계 403에서 계산된 빈 크기의 제곱)으로 놓는다.
- b. y = 단계 403에 따라 계산된, 모노 합성 채널의 해당하는 빈의 에너지.
- c. z = 스케일 인자 = 제곱_루트(x/y)로 놓는다. $x = 0$ 이면, $y = 0$ 이고 z 는 1로 설정.
- d. 예를 들어 100의 최대치로 z 를 제한, z 가 초기에 100보다 크면(다운믹싱 으로부터 강한 소거를 포함), 임의 값, 예를 들어, 모노 합성 빈의 실수와 허수부에 $0.01 * \text{제곱_루트}(x)$ 가산, 이것은 다음 단계에 의해 정규화되도록 충분히 크다는 것을 확신할 것이다.
- e. 복소수 모노 합성 빈 값과 z 를 승산.

단계 421에 대한 설명:

인코딩과 디코딩 양쪽에 같은 위상 인자를 이용하는 것이 일반적으로 바람직하지만, 서브밴드 위상 정정 값의 적절한 선택은 서브밴드 내에 하나 이상의 가청 스펙트럼 성분이 인코드 다운믹스 과정 동안에 취소되게끔 하는 것은 단계 419의 위상 전이가 빈 기반보다는 서브밴드 기반으로 실행되기 때문이다. 이런 경우에, 그런 빈의 합산 에너지가 그 주파수에서 개별 채널 빈의 에너지 합보다 작다고 검출되면, 인코더에서 고립된 빈에 대한 다른 위상 인자가 사용된다. 일반적으로 그러한 고립된 정정 인자를 디코더에 적용하는 것이 필요하지 않고, 그러기에 보통 고립된 빈은 전체의 이미지 품질에 거의 영향을 미치지 않는다. 모노 채널보다는 다중 채널이 적용된다면 유사한 정규화가 채택되기도 한다.

단계 422. 비트스트림으로 어셈블 및 패키징.

공통 모노 합성 오디오 혹은 매트릭스 된 다중 채널에 따르는 각각의 채널에 대한 진폭 스케일 인자, 각 제어 파라미터, 상관 해제 스케일 인자, 및 과도 플래그 사이드 채널 정보가 요구되는 만큼 멀티플렉싱되어, 저장, 전송 혹은 저장 및 전송 매체 혹은 미디어에 적당한 하나 이상의 비트스트림으로 패키징된다.

단계 422에 대한 설명:

모노 합성 오디오 혹은 다중 채널 오디오가 패키징되기 전에 예를 들어, 감지 인코더 같은 데이터-율 감소 인코딩 과정 혹은 장치, 아니면 감지 인코더 및 엔트로피 코더 (예를 들어, 산술 혹은 후프만 코더)에 인가된다. 또한, 상술하였듯이, 모노 합성 오디오(혹은 다중 채널 오디오)와 관련된 사이드체인 정보는 일정 주파수 이상의 오디오 주파수("커플링 주파수")에 대한 다중 입력 채널로부터 유도된다. 그런 경우에, 다중 입력 채널의 각각에서 커플링 주파수 이하의 오디오 주파수가 저장되고, 전송 혹은 이산 채널로서 저장 및 전송되거나, 여기 기술된 외의 방식으로 결합되거나 처리된다. 이산 혹은 달리 결합된 채널이 예를 들어 감지 인코더 혹은 감지 인코더 및 엔트로피 인코더 같은 데이터 감소의 인코딩 과정 혹은 장치에 역시 인가된다. 모노 합성 오디오(혹은 다중 채널 오디오) 및 이산 멀티채널 오디오가 패키징하기 전에 집적된 감지 인코딩 혹은 감지 및 엔트로피 인코딩 과정 혹은 장치에 모두 인가된다.

선택적인 보간 플래그(도 4에 미 도시)

서브밴드 각 제어 파라미터에 의해 제공된 기본적인 위상 각 시프트의 주파수 상의 보간은 인코더(단계 418) 및/또는 디코더(단계 505,아래)에서 가능하게 된다. 선택적인 보간 플래그 사이드체인 파라미터가 적용되어 디코더에서 보간이 가능해진다. 보간 플래그와 이와 유사한 구현 플래그가 인코더에 사용된다. 인코더가 빈 레벨에서 데이터에 액세스하기에 디코더 보다는 다른 보간 값을 갖고, 이것은 사이드체인 정보에서의 서브밴드 각 제어 파라미터를 보간한다는 것을 인지하라.

인코더 혹은 디코더에서 주파수 상의 그러한 보간의 이용은, 다음 두 가지 조건을 만족하면 가능해진다:

조건 1. 실제로 다른 위상 회전 각 할당을 갖는 두 개 서브밴드의 경계선 근처 혹은 경계선에 강하고 고립된 스펙트럼 피크가 위치하는지 이다.

이유: 보간 없이, 경계선에서 큰 위상 변화가 고립된 스펙트럼 성분에서의 지지권을 유도한다. 빈 내에 빈 값 상에 대역-대역 위상 각을 뿌리도록 하는 보간을 이용함으로써, 서브밴드 경계선에서의 변화의 양이 줄어든다. 이러한 조건을 만족하는 스펙트럼 피크 강도에 대한 임계치, 경계선에의 근접, 및 서브밴드마다 위상 회전의 차이가 경험적으로 조절된다.

조건 2. 과도 출현에 따라, 채널 간 위상 각(무 과도) 혹은 채널 내 절대 위상 각(과도) 양쪽이 선형 처리에 잘 맞는지 이다.

이유: 데이터 재구성 시의 보간 이용은 원래 데이터에 더 잘 맞는다. 각 데이터가 여전히 서브밴드 기반으로 디코더에 운반되기에 선형 처리의 기울기가 각 서브밴드 내에 이외의 모든 주파수 상에 일정할 필요가 없고, 이것은 보간기로의 입력을 형성한다. 이러한 조건을 만족하도록 양호하게 맞는 그 정도는 역시 경험적으로 결정된다.

경험적으로 결정된 다른 조건은 주파수상에 보간으로부터 얻어진다. 이제 방금 설명한 두 가지 조건의 존재는 다음으로 결정된다:

조건 1. 실제로 다른 위상 회전 각 할당을 갖는 두 개 서브밴드의 경계선 근처 혹은 경계선에 강하고 고립된 스펙트럼 피크가 위치하는지 이다.

디코더에 의해 사용되도록 보간 플래그에 대해, 서브밴드 각 제어 파라미터(단계 414의 출력), 및 인코더 내에 단계(418)의 구동을 위해, 양자화전에 단계 (4130의 출력이 서브밴드마다 회전 각을 결정한다.

보간 플래그를 위해 및 인코더 내에 단계 403의 크기 출력을 구동하기 위해, 현재의 DFT 크기는 서브밴드 경계에서 고립된 피크를 찾는다.

조건 2. 과도의 출현에 따라, 채널 간 위상 각(무 과도) 혹은 채널 내 절대 위상 각(과도) 양쪽이 선형 처리에 잘 맞는지 이다.

과도 플래그가 참이 아니면(무 과도), 단계(406)에서 선형 결정까지 상대적인 채널간 빈 위상 각을 이용, 및

과도 플래그가 참이면(과도), 단계(403)에서 채널의 절대 위상 각을 이용.

디코딩

디코딩 과정("디코딩 단계")의 단계가 다음과 같이 기술된다. 디코딩 단계에 대해, 도 5가 참고되고, 이것은 흐름도와 기능 블록도를 혼합으로 이용한다. 간단히 나타나도록, 도면은 한 채널에 대한 사이드체인 정보 성분의 이탈을 도시하고, 설명되듯이 채널이 그러한 성분에 대한 기준 채널이 아니면 사이드체인 정보 성분이 각 채널에 대해 얻어져야 한다는 것을 이해할 것이다.

단계 501. 사이드체인 정보 언패킹 및 디코드.

필요한 대로, 각 채널(도 5에 도시된 한 채널)의 각각의 프레임에 대한 사이드체인 데이터 성분(진폭 스케일 인자, 각 제어 파라미터, 상관 해제 스케일 인자, 및 과도 플래그)을 언패킹 하여 디코딩(역양자화 포함). 테이블 검사는 진폭 스케일 인자, 각 제어 파라미터, 및 상관 해제 스케일 인자를 디코드한다.

단계 501에 대한 설명:

위에 설명되듯이, 기준 채널이 적용되면, 기준 채널에 대한 사이드체인 데이터는 각 제어 파라미터, 상관 해제 스케일 인자, 및 과도 플래그를 포함하지 않는다.

단계 502. 모노 합성 혹은 멀티채널 오디오 신호 인패킹 및 디코딩

필요한 대로 모노 합성 혹은 멀티채널 오디오 신호 정보를 인패크하여 디코드하여 모노 합성 혹은 멀티채널 오디오 신호의 각 변환 빈에 대한 DFT 계수를 제공.

단계 502에 대한 설명:

단계 501과 502는 단일의 인패킹 및 디코딩 단계의 일부로 여겨진다. 단계(502)는 수동 혹은 능동 매트릭스를 포함한다.

단계 503. 블록 상에 각 파라미터 값 분배

블록 서브밴드 각 제어 파라미터 값은 역-양자화 프레임 서브밴드 각 제어 파라미터 값에서 유도된다.

단계 503에 대한 설명:

단계 503은 프레임에서 각 블록에 동일 파라미터를 분배함으로써 구현된다.

단계 504. 블록 상에 서브밴드 상관 해제 스케일 인자 분배

블록 서브밴드 상관 해제 스케일 인자 값은 디쿼타이징 프레임 서브밴드 상관 해제 스케일 인자 값에서 유도된다.

단계 504에 대한 설명:

단계 504는 동일 스케일 인자 값을 프레임의 매 블록에 분배하여 구현된다.

단계 505. 주파수 상에 선형으로 보간

선택적으로, 디코더 단계(418)와 연관지어 위에 기술된 주파수 상에 선형 보간에 의해 디코더 단계(503)의 블록 서브밴드 각에서 빈 각을 유도. 단계 505에서의 선형 보간은 보간 플래그가 사용되어 참일 때 구동된다.

단계 506. 랜덤 화 된 위상 각 오프셋 부가(기술 3)

상기의 기술 3에 따라 위에 기술된대로, 과도 플래그가 과도를 가르칠 때 단계 503에 의해 제공된 블록 서브밴드 각 제어 파라미터를 부가하고, 이것은 단계 505에 의해 주파수상에 선형 보간된다(스케일링은 본 단계에서 설명된대로 이차적이다).

- a. y = 블록 서브밴드 상관 해제 스케일 인자로 놓는다.
- b. $z = y^{\text{exp}}$ 로 놓고, 여기서 exp는 상수로 예를 들어 5이다. z 는 역시 0 내지 1의 범위지만 0을 향해 기울어서 상관 해제 스케일 인자값이 높지 않으면 랜덤 변동의 낮은 레벨을 향한 바이어스를 반사한다.
- c. $x = + 1.0$ 과 -1.0 사이의 랜덤 화 된 숫자로, 각 블록의 각각의 서브밴드에 대해 개별적으로 선택된다.
- d. 다음에, 블록 서브밴드 각 제어 파라미터에 부가된 값이 기술 3에 따른 랜덤 화 된 각 오프셋 값에 부가되어 $x \cdot \pi \cdot z$ 이다.

단계 506에 대한 설명:

당 분야의 기술자에게 인지되듯이, 상관 해제 스케일 인자에 의한 스케일링에 대한 "랜덤 화 된" 각(혹은 진폭이 역시 스케일 된다면 "랜덤 화 된" 진폭)이 가짜-랜덤 및 진짜 랜덤 변동뿐 아니라 결정론적으로-발생된 변동을 포함하고, 위상각 혹은 위상 각과 진폭에 인가될 때 채널 사이에 크로스-상관관계(cross-correlation)를 줄이는 효과를 갖는다.

그러한 "랜덤 화 된" 변동은 여러 방식으로 얻어진다. 예를 들어, 여러 종자 값을 갖는 의사-난수 발생기가 채택된다. 대안으로, 진짜로 난수(random number)가 하드웨어 난수 발생기를 이용하여 생성된다. 그리하여 약 1도의 랜덤 화 된 각 분석이 충분하고, 두 개나 세 개의 십진 자릿수(예로, 0.84 혹은 0.844)를 갖는 난수표가 적용된다. 바람직하게, 임의의 값(단계 505c를 참고로, -1.0과 +1.0 사이의)이 각각의 채널 상에 통계적으로 균일하게 분배된다.

단계 506의 비-선형 간접 스케일링이 유용하다고 보여져 왔지만, 그것은 결정적이지 않아서 다른 적당한 스케일링이 적용된다-특히, 지수에 대한 다른 값이 적용되어 유사한 결과를 얻는다.

서브밴드 상관 해제 스케일 인자 값이 1일 때, $-\pi$ 에서 $+\pi$ 로의 전체 범위의 랜덤 각이 가산된다(이 경우에 단계 503에 의해 발생된 블록 서브밴드 각 제어 파라미터 값은 무관하다). 서브밴드 상관 해제 스케일 인자 값이 0으로 감소될 때, 역시 랜덤 화 된 각 오프셋이 0으로 감소되고, 이것은 단계 506의 출력이 단계 503에 의해 발생된 서브밴드 각 제어 파라미터 값을 향해 이동하도록 한다.

필요하다면, 상기의 인코더는 역시 기술 3에 따른 스케일 된 랜덤 오프셋을 다운-믹싱 하기 전에 채널에 인가된 각 시프트에 가산한다. 그렇게 하는 것은 디코더에서 앨리어스(alias) 소거를 개선한다. 그것은 또한 인코더와 디코더의 동기화를 개선하는데 유리하다.

단계 507. 랜덤 화 된 위상 각 오프셋 가산(기술 2).

위에 설명된대로 기술 2에 따라서, 과도 플래그가 각각의 빈에 대해 과도를 나타내지 않을 때, 단계(503)에 의해 제공된 프레임에서 모든 블록 서브밴드 각 제어 파라미터에 상관 해제 스케일 인자로 스케일 된 다른 랜덤 오프셋 값을 부가하라(스케일링은 이 단계에서 설명된 대로 직접이다):

- a. $y =$ 블록 서브밴드 상관해제기 스케일 인자로 놓는다.
- b. $x = x +$ 1.0과 1.0 사이의 랜덤 화 된 숫자로, 각 프레임의 각각의 빈에 대해 개별적으로 선택된다.
- c. 다음에, 블록 빈 각 제어 파라미터에 부가된 값이 기술 3에 따른 랜덤 화 된 각 오프셋 값에 부가되어 $x*\pi*y$ 이다.

단계 507에 대한 설명:

랜덤 화 된 각 오프셋에 대한 단계 505에 관한 상기 주석을 보라.

단계 507의 직접 스케일링이 유용하다고 보였지만, 그것은 결정적이지 않고, 다른 적당한 스케일링이 적용된다.

임시 불연속을 최소화하도록, 각 채널의 각 빈에 대한 독특한 랜덤 각 값은 바람직하게 시간에 따라 변하지 않는다. 한 서브밴드의 모든 빈의 랜덤 각 값은 동일 서브밴드 상관해제기 스케일 인자에 의해 스케일되고, 이것은 프레임 율로 갱신된다. 이렇게, 서브밴드 상관 해제 스케일 인자 값이 1일 때, $-\pi$ 내지 $+\pi$ 범위의 랜덤 각이 부가된다(이 경우에 역양자화 프레임 서브밴드 각 값으로부터 유도된 블록 서브밴드 각 값이 무관). 서브밴드 상관 해제 스케일 인자 값이 0으로 감소될 때, 랜덤 화 된 각 오프셋은 역시 0으로 감소된다. 단계 504와는 달리, 단계 507의 스케일링은 서브밴드 상관 해제 스케일 인자 값의 직접 기능이다. 예를 들어, 0.5의 서브밴드 상관 해제 스케일 인자 값은 0.5만큼 매 랜덤 각 변동으로 비례적으로 감소한다.

다음에 스케일 된 무자위 각 값이 디코더 단계 506에서 빈 각(bin angle)으로 가산된다. 상기 상관 해제 스케일 인자 값이 프레임당 한번 갱신된다. 상기 프레임에 대한 과도 플래그의 출현시에, 이러한 단계는 건너뛰어서 과도 이전-잡음 인공물을 피한다.

필요하다면, 상기의 인코더는 역시 기술 2에 따른 스케일 된 랜덤 오프셋을 다운-믹싱 하기 전에 인가된 각 시프트에 가산한다. 그렇게 하는 것은 디코더에서 앨리어스 소거를 개선한다. 그것은 또한 인코더와 디코더의 동기화를 개선하는데 유리하다.

단계 508. 진폭 스케일 인자 정규화.

합의 제곱이 1이 되도록 채널 상의 진폭 스케일 인자 정규화.

단계 508에 대한 설명:

예를 들어, 두 채널이 -3.0 dB의 스케일 인자를 역-양자화했다면(= $2 * 1.5\text{dB}$ 의 알갱이 폭)(.70795), 그 제곱의 합이 1.002이다. 각각을 1.002의 제곱 루트로 나누면 .7072(-3.01 dB)의 두 값이 된다.

단계 509. 서브밴드 스케일 인자 레벨 증폭(선택적인).

선택적으로, 과도 플래그(Transient Flag)가 미-과도를 가르칠 때, 서브밴드 상관 해제 스케일 인자 레벨에 따라 서브밴드 스케일 인자 레벨을 부가적으로 약간 올려라: 각각의 정규화된 서브밴드 진폭 스케일 인자에 작은 인자(예를 들어, $1 + 0.2 * \text{서브밴드 상관 해제 스케일 인자}$)를 곱하라. 상기 과도 플래그가 참일 때, 본 단계를 건너뛰어라.

단계 509에 대한 설명:

상기 디코더 상관 해제 단계 507이 마지막 역 필터뱅크 과정에서 약간 감소된 레벨이 되기 때문에 본 단계는 유용하다.

단계 510. 빈 상에 서브밴드 진폭 값 분배.

동일 서브밴드 진폭 스케일 인자 값을 서브밴드에서 각 빈에 분배함으로써 단계 510이 구현된다.

단계 510a. 랜덤 화 된 진폭 오프셋을 가산(선택적인)

선택적으로, 서브밴드 상관 해제 스케일 인자 레벨과 과도 플래그에 따라 정규화된 서브밴드 진폭 스케일 인자에 랜덤 화 된 변화를 인가하라. 고도의 부재시에, 개별 빈을 기반으로(빈 마다 다른) 시간에 따라 변하지 않는 랜덤 화 된 진폭 스케일 인자를 부가하고, 과도 출현시에(프레임 혹은 블록에서), 개별 블록 기반으로 변하고(블록마다 다른) 서브밴드마다 변하는(하나의 서브밴드에서 모든 빈에 대해 동일한 시프트; 서브밴드마다 다른) 랜덤 화 된 진폭 스케일 인자를 부가하라. 단계 510a는 도면에 도시되지 않는다.

단계 510a에 대한 설명:

랜덤 화 된 진폭 시프트가 부가되는 정도가 상관 해제 스케일 인자에 의해 제어되지만, 가청 인공물을 피하기 위해 동일 스케일 인자 값이 되는 해당 랜덤 위상 시프트보다는 더 작은 진폭 시프트를 특정 스케일 인자 값이 초래해야 하는 것으로 믿어진다.

단계 511. 업믹스.

a. 각각의 출력 채널의 각각의 빈에 대해, 디코더 단계 508의 진폭과 디코더 단계 507의 빈 각으로부터 복잡한 업믹스 스케일 인자를 구성하라: $(\text{진폭} * (\cos(\text{각}) + j \sin(\text{각})))$.

b. 각각의 출력 채널에 대해, 상기 채널의 각각의 빈의 업믹스 된 복소수 출력 빈 값을 발생하도록 복소수 빈 값과 복소수 업믹스 스케일 인자를 곱하라.

단계 512. 역 DFT 실행(선택적인).

선택적으로, 멀티채널 출력 PCM 값을 내도록 각각의 출력 채널의 빈에서 역 DFT 변환을 실행하라. 잘 알려졌듯이, 그러한 역 DFT 변환과 연관지어 개별 블록의 시간 샘플은 윈도우(window)되고, 마지막으로 연속적인 시간 출력 PCM 오디오 신호를 재구성하기 위해 인접 블록들은 중복되어 함께 가산된다.

단계 512에 대한 설명:

본 발명에 따른 디코더는 PCM 출력을 제공하지 않는다. 디코더 과정이 주어진 커플링 주파수 이상으로 오직 적용되고 이산 MDCT 계수가 상기 주파수 이하로 각각의 채널에 대해 보내지는 경우에, 디코더 업믹싱 단계 511a와 511b에 의해 유

도된 DFT 계수를 MDCT 계수로 변환하는 것이 바람직하여서, 역변환이 실행되는 외부 기기로의 적용을 위한 표준 AC-3 SP/DIF 비트스트림같은, 대량의 설치된 사용자를 갖는 인코딩 시스템과 양립할 수 있는 비트스트림을 예를 들어 제공하도록 더 낮은 주파수 이산 MDCT 계수와 결합되어 재양자화 될 수 있다. 역 DFT 변환은 출력 채널들 중의 하나에 적용되어 PCM 출력을 제공한다.

A/52A 문서의 섹션 8.2.2

감도 인자 "F"가 부가된 채로

8.2.2. 과도 검출

이전-에코 성능을 개선하도록 길이 오디오 블록을 짧게 하는 데는 언제 스위칭할지를 결정하기 위해 과도가 완전-대역폭 채널에서 검출된다. 고역 필터링 버전의 신호가 하나의 서브-블록 시간-세그먼트에서 다음 까지 에너지의 증가 동안 검사된다. 서브-블록은 다른 시간 스케일에서 검사된다. 과도가 한 채널에서 오디오 블록의 제2 반에서 검출된다면 채널은 짧은 블록으로 스위칭한다. 블록-스위칭된 채널이 D45 지수 전략을 이용한다[즉, 임시 분석에서 증가가 되는 데이터 오버헤드를 줄이도록 데이터가 더 조잡한 주파수 분석을 한다].

과도 검출은 긴 변환 블록(길이 512)에서 짧은 블록(길이 256)까지 언제 스위칭하는지를 결정하는데 이용된다. 이것은 두 개 관문으로 이루어지고, 각 관문은 256 샘플을 처리한다. 과도 검출이 4단계로 분석된다: 1)고역 필터링, 2) 블록의 여러 개로 세분화, 3) 각 서브-블록 세그먼트 내에 피크 크기 검출, 및 4) 임계치 비교. 과도 검출기는 각각의 완전 대역폭 채널에 대한 플래그 blksw[n]을 출력하고, 이것은 해당 채널에 대한 512 길이 입력 블록의 제2의 반 쪽에서 과도의 출현을 언제 "1"로 설정되는지가 나타낸다.

- 1) 고역-통과 필터링: 고역-통과 필터는 8kHz의 차단 주파수를 갖는 연속된 4차 직접 형식 II III 필터로 구현된다.
- 2) 블록 세그먼트: 256 고역-통과 필터링 샘플의 블록은 계급 트리의 레벨로 분절되는데, 레벨 1은 256 길이 블록을 나타내고, 레벨 2는 길이 128의 두 개 세그먼트이고, 레벨 3은 길이 64의 4개 세그먼트이다.
- 3) 피크 검출: 가장 큰 크기의 샘플은 계급 트리의 각 레벨에서 각각의 세그먼트에 대해 식별된다. 단일 레벨에 대한 피크는 다음처럼 도시된다:

$$P[j][k] = \max(x(n))$$

여기서 $n = (512 \times (k-1)/2^j), (512 \times (k-1)/2^j) + 1, \dots, (512 \times k/2^j) - 1$

및 $k = 1, \dots, 2^{(j-1)}$;

여기서: $x(n)$ = 256 길이 블록에서 n번째 샘플

$j = 1, 2, 3$ 은 계급 구조 레벨의 번호

$k =$ 레벨 j 내에 세그먼트 번호

$P[j][0]$, (즉, $k=0$)은 현재 트리 이전에 바로 계산된 트리의 레벨 j 상의 마지막 세그먼트의 피크로 정의된다. 예를 들어, 선행 트리의 $P[3][4]$ 는 현재 트리에서 $P[3][0]$ 이다.

- 4) 임계치 비교: 제1 단계의 임계치 비교기는 현재 블록에 중대한 신호 레벨이 있는지를 체크한다. 이것은 "침묵 임계치"와 현재 블록의 전체 피크 값 $P[1][1]$ 을 비교함으로써 이루어진다. $P[1][1]$ 가 상기 임계치 이하라면, 긴 블록이 힘을 받는다. 침묵 임계치는 100/32768이다. 다음 단계의 비교기는 계급 구조의 트리의 각 레벨 상에서 인접 세그먼트의 상대적인 피크 레벨을 체크한다. 특정 레벨 상의 두 개의 인접한 세그먼트의 피크 비율이 그 레벨에 대한 소정의 임계치를 초과하면, 플래그는 현재의 256-길이 블록에서 과도의 출현을 나타내도록 설정된다. 상기 비율은 다음으로 비교된다:

$$\text{mag}(P[j][k]) \times T[j] > (F * \text{mag}(P[j][(k-1)])) \text{ ["F"는 감도 인자]}$$

여기서: $T[j]$ 는 레벨 j 에 대한 이전-정의된 임계치이고, 다음으로 정의:

$$T[1] = .1$$

$$T[2] = .075$$

$$T[3] = .05$$

이러한 부등식이 어떤 레벨에서 두 개의 세그먼트 피크에 대해 참(true)이라면, 과도는 512 길이 입력 블록의 제1의 반 쪽에 대해 나타난다. 이러한 과정을 통한 제 2의 관문은 512 길이 입력 블록의 제2의 반 쪽에서 과도의 출현을 결정한다.

N:M 인코딩

본 발명의 특징은 도 1과 연관지어 기술된 N: 1 인코딩에 국한되지 않는다. 더욱 일반적으로는, 본 발명의 특징은 도 6의 방식으로(즉, N: M 인코딩) 어떤 수의 입력 채널(n 입력 채널)을 어떤 수의 출력 채널(m 출력 채널)로 변환 가능하게 한다. 많은 공통의 응용에서, 입력 채널의 수 n 이 출력 채널의 수 m 보다 더 크기에, 도 6의 N: M 인코딩 장치는 간략한 표현으로 "다운믹싱(downmixing)"으로 일컬어진다.

도 6의 세부 내용에 대해, 도 1의 장치에서처럼 가산 결합기(6)에서 회전 각(8)과 회전 각(10)의 출력을 합산하는 대신에, 상기 출력은 다운믹스 매트릭스 장치 혹은 기능 6'("다운믹스 매트릭스")에 인가된다. 다운믹스 매트릭스 6'는 도 1의 N: 1 인코딩에서처럼 한 채널, 혹은 여러 채널에 간단한 합산을 제공하는 능동 혹은 수동 매트릭스이다. 매트릭스 계수는 실수 혹은 복소수(실수 및 허수)이다. 도 6의 다른 장치와 기능이 도 1 장치에서와 동일하고, 동일 참조번호를 쓴다.

예를 들어, 주파수 범위 f_2 내지 f_3 에서 $m_{f_2-f_3}$ 채널과 주파수범위 f_1 내지 f_2 에서 $m_{f_1-f_2}$ 채널을 제공하는 식으로 다운믹스 매트릭스 6'이 하이브리드 주파수-기반의 기능을 제공한다. 예를 들어, 1000Hz의 커플링 주파수 이하로 다운믹스 매트릭스 6'는 두 채널을 제공하고, 커플링 주파수 이상으로 다운믹스 매트릭스 6'는 한 채널을 제공한다. 커플링 주파수 이하로 두 채널을 채택함으로써, 더 좋은 공간 충실도가 얻어지고, 특히 두 채널이 수평 방향을 나타낼 때이다(사람 귀의 수평 상태가 매치).

도 6이 도 1 장치에서처럼 각각의 채널에 대해 동일한 사이드체인 정보의 발생을 보여준다면, 하나 이상의 채널이 다운믹스 매트릭스 6'의 출력에 의해 제공될 때 사이드체인 정보의 일정 부분의 생략이 가능하다. 어떤 경우에는, 진폭 스케일 인자 사이드체인 정보만이 도 6의 장치에 의해 제공될 때 수용 가능한 결과가 얻어지기도 한다. 사이드체인 옵션에 대한 더 상세한 부분은 도 7, 8, 9의 설명과 연관지어 아래 논의된다.

위에 언급되었듯이, 다운믹스 매트릭스 6'에 의해 발생된 여러 채널은 입력 채널의 수 n 보다 더 작을 필요가 없다. 도 6에서와 같은 인코더의 목적이 전송 혹은 저장용 비트의 수를 줄이는 것일 때, 다운믹스 매트릭스 6'에 의해 발생된 채널의 수는 입력 채널의 수 n 보다 더 작을 것 같다. 그러나, 도 6의 장치가 역시 "업믹서"(upmixer)로도 사용된다. 그런 경우에, 다운믹스 매트릭스 6'로 발생된 채널 수 m 이 입력 채널 수 n 보다 더 많은 어플리케이션이 존재한다.

그러한 디코더에 의해 디코딩될 때 오디오 정보와 사이드체인 정보가 합당한 결과를 제공하는지를 결정하기 위해, 도 2, 5 및 6의 예와 연관지어 설명된 인코더가 역시 그들 자체의 로컬 디코더 혹은 디코딩 기능을 포함한다. 그러한 결정의 결과는 예를 들어, 반복적인 과정을 적용함으로써 파라미터를 개선하는데 이용된다. 블록 인코딩과 디코딩 시스템에서, 한 블록의 오디오 정보와 그에 연관된 공간 파라미터의 전송시의 지연을 최소화하도록 다음 블록이 종료되기 전에 반복적인 계산이 매 블록에서 실행될 수 있다.

인코더가 또한 그 자체의 디코더 혹은 디코딩 기능을 포함하는 장치가 역시 채택될 수 있는 때는 공간 파라미터가 어떤 블록에만 저장되지 않거나 보내질 때이다. 불합리한 디코딩이 공간-파라미터 사이드체인 정보를 보내지 않는다고 하면, 그러한 정보는 특정 블록에 대해 보내진다. 이런 경우에, 디코더는 그것이 커플링 주파수 이상의 주파수에 대해 입력 스트림으로부터 공간-파라미터 사이드체인 정보를 복원할 능력과, 커플링 주파수 이하에서 스테레오 정보로부터 시뮬레이트된 공간-파라미터 사이드체인 정보를 발생시킬 수 있는 능력을 갖는 점에서, 도 2, 5 혹은 6의 디코더 혹은 디코딩 기능의 변경안이다.

로컬 디코더 혹은 디코더 기능을 갖기보다는 그러한 로컬-디코더-조합의 인코더 예의 간략화된 대안에서, 인코더는 커플링 주파수 이하로 어떤 신호 내용이 있는지를 단순히 체크하여(주파수 범위를 통해 주파수 빈에 에너지의 합을 합당한 방식으로 결정), 없으면 그것은 에너지가 임계치 이상이라면 공간-파라미터 사이드체인 정보를 보내거나 저장한다. 인코딩 체계에 따라서, 커플링 주파수 이하의 낮은 신호 정보가 역시 사이드체인 정보를 보내는데 유용한 더 많은 비트를 야기한다.

M:N 디코딩

도 2의 장치의 더 일반적인 형태가 도 7에 도시되고, 여기서 업믹스 매트릭스 기능 혹은 장치("업믹스 매트릭스") 20가 도 6의 장치에 의해 발생된 1 내지 m 채널을 수신한다. 업믹스 매트릭스(20)는 수동 매트릭스이다. 그것은 도 6 장치의 다운믹스 매트릭스 6'의 치환(즉, 보충)이다. 달리, 업믹스 매트릭스(20)는 능동 매트릭스-가변 매트릭스를 갖는 조합의 가변 매트릭스 혹은 수동 매트릭스이다. 능동 매트릭스 디코더가 적용되면, 완화되거나 조용한 상태에서 그것은 다운믹스의 복잡한 결합이거나 독립된 다운믹스 매트릭스이다. 사이드체인 정보가 도 7에 도시되듯이 인가되어 진폭 조절, 회전 각, 및 (선택) 보간기 기능 또는 장치를 제어한다. 그런 경우에, 능동 매트릭스라면, 업믹스 매트릭스는 사이드체인 정보에 관계없이 동작하여 그에 인가된 채널에만 응답한다. 대안으로, 일부 혹은 모든 사이드체인 정보가 능동 매트릭스에 인가되어 그의 동작을 지원한다. 그런 경우에, 진폭 조절, 각 회전, 및 보간기 기능 또는 장치의 일부 또는 전부가 생략된다. 도 7의 디코더 예가 또한 도 2, 5와 연관되어 위에 기술된대로 일정 신호 조건 아래 랜덤 화 된 진폭 변동의 정도를 적용하는 대안을 채택한다.

업믹스 매트릭스(20)가 능동 매트릭스일 때, 도 7의 장치는 "하이브리드 매트릭스 디코더"로 특징지어져서 "하이브리드 매트릭스 인코더/디코더 시스템"으로 동작한다. 상기 '하이브리드'는 디코더가 그의 입력 오디오 신호(즉, 능동 매트릭스는 그에 인가된 채널에 인코딩된 공간 정보에 응답한다)에서 어떤 측정의 제어 정보와 공간-파라미터 사이드체인 정보로부터 측정 제어정보를 유도한다. 도 7의 다른 소자는 도 2의 장치에서처럼 동일 부재 번호를 쓴다.

하이브리드 매트릭스 디코더에 사용되는 적당한 능동 매트릭스 디코더는 상기의 것과 같은 능동 매트릭스 디코더를 포함하고 예를 들어, "Pro Logic"과 Pro Logic II" 디코더로 알려진 매트릭스 디코더를 포함한다.

선택적인 상관 해제

도 8, 9는 도 7의 일반화된 디코더 상의 변화를 도시한다. 특히, 도 8과 도 9의 장치는 도 2, 7의 상관 해제 기술의 대안을 보여준다. 도 8에서, 각각의 상관 해제 기능 혹은 장치("상관해제기")(46, 48)는 시간 영역에 있고, 각각은 그들의 채널에서 각각의 역 필터뱅크(30, 36)를 따른다. 도 9에서, 각각의 상관 해제 기능 혹은 장치("상관해제기")(50, 52)는 주파수 영역에 있고, 각각은 그들의 채널에서 역 필터뱅크(30, 36)를 선행한다. 도 8, 9의 장치에서, 상관해제기(46, 48, 50, 52) 각각은 그 출력이 서로에 대해 상호 상관해제되도록 독특한 특성을 갖는다. 상관 해제 스케일 인자는 예를 들어, 각각의 채널에 제공된 비-상관 관계와 상관 해제의 비율을 제어한다. 선택적으로, 아래 설명되듯이 과도 플래그는 역시 디코리레이터의 동작 모드를 시프트한다. 도 8, 9의 장치에서, 각각의 상관 해제기는 그 자체의 독특한 필터 특성을 갖는 쉬로더-형(Schroeder-type) 반향기이고, 여기서 반향 정도의 양은 상관 해제 스케일 인자(예를 들어, 상관 해제기 출력이 상관 해제기 입력과 출력의 선형 조합의 일부를 형성하는 정도를 제어함으로써 구현되는)에 의해 제어된다. 대안으로, 다른 제어 가능한 상관 해제 기술은 쉬로더-형 반향기 자체에 혹은 서로가 조합한 조합체에 적용된다. 상기 쉬로더-형 반향기는 공지되어 있고, 두 개의 저널 논문에 원문이 실려 있다: M.R. Schroeder와 B.F.Logan 공저의 1961년 판 *IRE Transactions on Audio*, vol. AU-9, pp. 209-214 제목 "'무색' 인공 반향" 및 M.R. Schroeder저의 1962년 7월 판 *Journal A.E.S.*, vol. 10, no2, pp. 219-223에 제목 "자연적인 사운드 인공적인 반향".

도 8의 장치에서처럼, 상관해제기(46, 48)가 시간 영역에서 동작할 때 단일(즉, 광대역) 상관 해제 스케일 인자가 필요하다. 이것은 여러 방법으로 얻어진다. 예를 들어, 오직 단일 상관 해제 스케일 인자가 도 1 혹은 도 7의 인코더에서 발생한다. 대안으로, 도 1 혹은 도 7의 인코더가 서브밴드 기반으로 상관 해제 스케일 인자를 발생하면, 서브밴드 상관 해제 스케일 인자는 도 8의 디코더에서 혹은 도 1, 7의 인코더에서 크기나 파워가 합산된다.

상관해제기(50, 52)가 도 9의 장치에서처럼 주파수 영역에서 동작할 때, 그들은 각 서브밴드 혹은 서브밴드의 그룹에 대해 상관 해제 스케일 인자를 수신하고, 부수적으로 그러한 서브밴드 혹은 그룹의 서브밴드에 대한 균형잡힌 정도의 상관 해제를 제공한다.

도 8의 상관해제기(46, 48)와 도 9의 상관해제기(50, 52)는 선택적으로 과도 플래그를 수신한다. 도 8의 시간-영역 상관해제기에서, 과도 플래그가 각각의 상관해제기의 동작의 모드를 시프트하도록 적용된다. 예를 들어, 상관해제기는 과도 플래그의 부재시에 쉬로더-형 반향기로 동작하지만, 플래그의 수신시에 짧은 시간동안 즉, 1 내지 10 밀리 초 동안 고정된 지연기로 동작한다. 각각의 채널은 소정의 고정된 지연기를 갖거나 그 지연기는 복수의 과도에 응답하여 짧은 시간 내에 변한다. 도 9의 주파수-영역 상관해제기에서, 과도 플래그는 역시 각 상관해제기의 동작 모드를 시프트하도록 적용된다. 그러나, 이 경우에, 과도 플래그의 수신은 예를 들어, 플래그가 발생한 채널에서 진폭의 짧은(수 밀리 초) 증가를 트리거한다.

도 8과 9의 장치에서, 선택적인 과도 플래그에 의해 제어되는 보간기 27 (33)는 상기 방식으로 회전 각 28(33) 밖의 위상각의 주파수 상에 보간을 제공한다.

위에 기술되듯이, 두 개 이상의 채널이 사이드체인 정보에 부가하여 보내질 때, 사이드체인 파라미터의 수를 줄이는 것은 수용 가능하다. 예를 들어, 진폭 스케일 인자만을 보내는 것이 수용 가능하고, 이런 경우에 디코더에서의 상관 해제과 각도 장치 혹은 기능이 생략된다.

대안으로, 오직 진폭 스케일 인자, 상관 해제상관 해제기, 및 선택적으로, 과도 플래그가 보내진다. 그런 경우에, 도 7, 8, 9의 장치가 적용된다(그들의 각각에서 회전 각(28, 34) 생략).

다른 대안으로, 오직 진폭 스케일 인자와 각 제어 변수가 보내진다. 그런 경우에, 도 7, 8, 9의 장치가 적용된다(도 7의 상관 해제기 (38, 42)와, 도 8, 9의 46, 48, 50, 52 생략).

도 1과 2에서처럼, 도 6-9의 장치는 간략히 나타내도록 오직 2 채널이 도시되었지만, 입력 및 출력 채널이 어떤 개수도 무방하다.

본 발명과 이의 여러 특징의 다른 변경과 수정의 구현이 당 분야의 기술자에게는 명백하고, 본 발명이 여기 기술된 특정 실시예에만 국한되지 않는다는 것을 이해해야 한다. 그러므로, 여기에 개시된 본 발명의 기본적인 원리와 진정한 정신을 벗어나지 않는 범위 내에서 본 발명의 모든 변경, 수정 등은 본 발명에 의해 수용되는 것은 물론이다.

도면의 간단한 설명

도 1은 본 발명의 특징을 구현하는 N: 1 인코딩 장치의 주요 기능 혹은 장치를 도시한 이상적인 블록도이다.

도 2는 본 발명의 특징을 구현하는 1: N 디코딩 장치의 주요 기능 혹은 장치를 도시한 이상적인 블록도이다.

도 3은 시간 축을 따른(수평) 프레임과 주파수 축을 따른(수직) 서브밴드와 빈의 간략화된 개념적인 구조의 예를 도시한다.

도 4는 본 발명의 특징을 구현하는 인코딩 장치의 기능을 실행하는 인코딩 단계 혹은 장치를 도시한, 흐름도와 기능 블록도를 혼합한 성질의 도면이다.

도 5는 본 발명의 특징을 구현하는 디코딩 장치의 기능을 실행하는 디코딩 단계 혹은 장치를 도시한, 흐름도와 기능 블록도를 혼합한 성질의 도면이다.

도 6은 본 발명의 특징을 구현하는 제 1의 N: x 인코딩 장치의 주요 기능 혹은 장치를 도시한 이상적인 블록도이다.

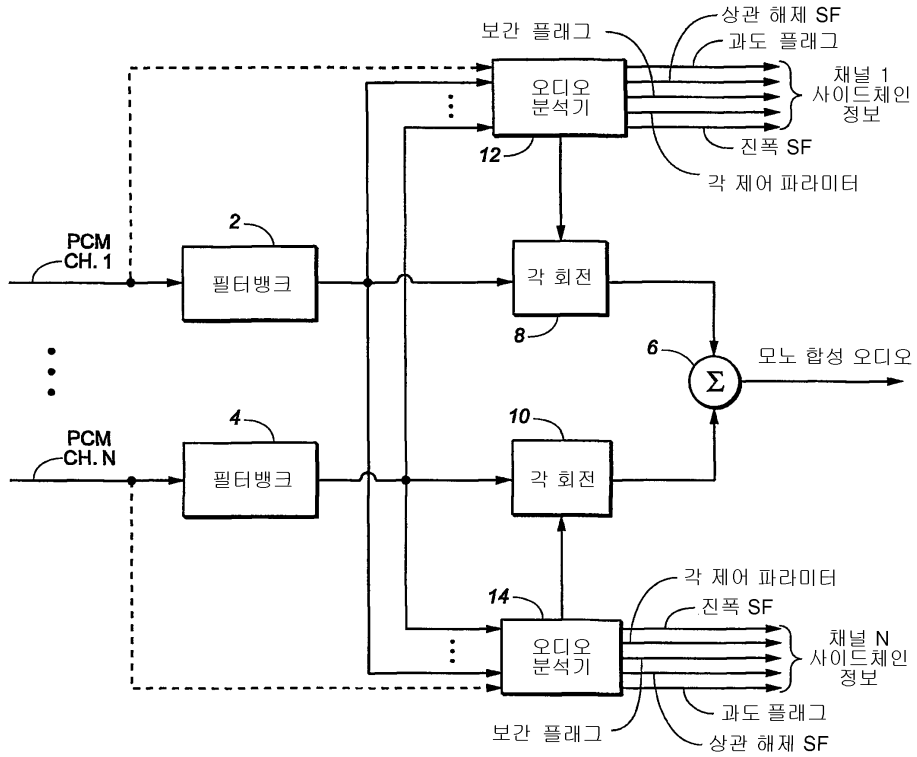
도 7은 본 발명의 특징을 구현하는 x: M 디코딩 장치의 주요 기능 혹은 장치를 도시한 이상적인 블록도이다.

도 8은 본 발명의 특징을 구현하는 제 1의 다른 x: M 디코딩 장치의 주요 기능 혹은 장치를 도시한 이상적인 블록도이다.

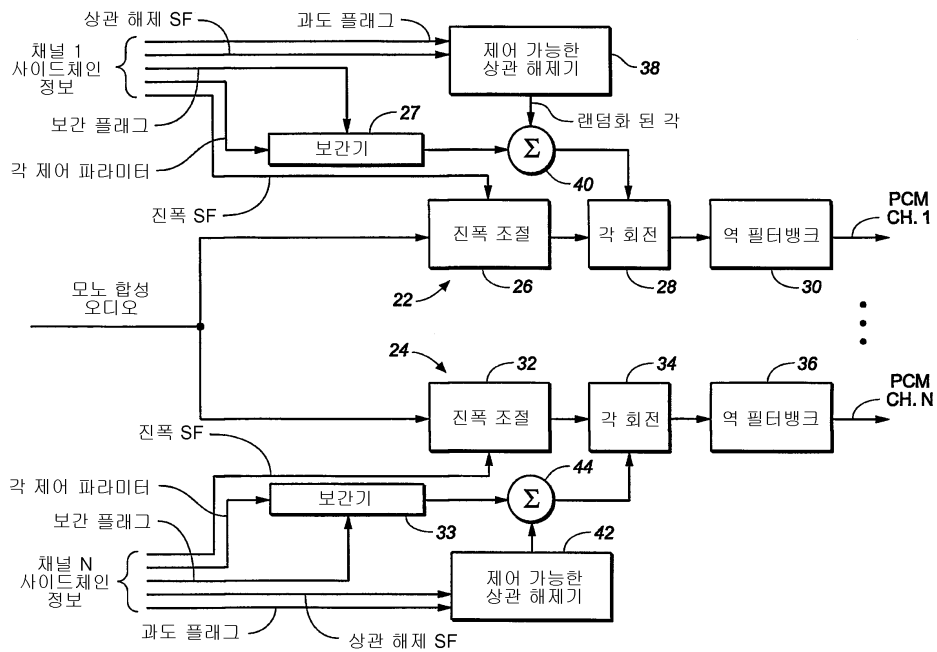
도 9는 본 발명의 특징을 구현하는 제 2의 다른 x: M 디코딩 장치의 주요 기능 혹은 장치를 도시한 이상적인 블록도이다.

도면

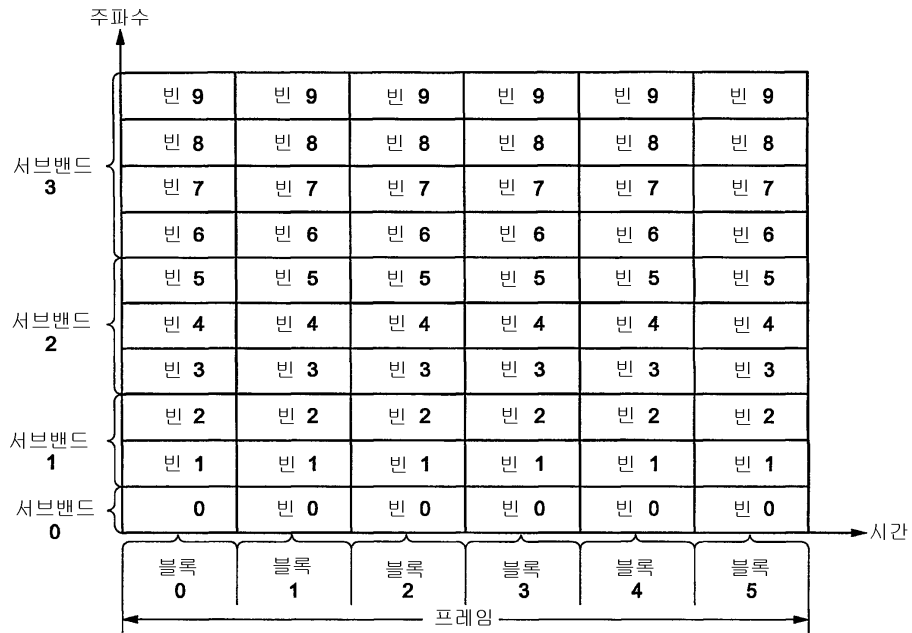
도면1



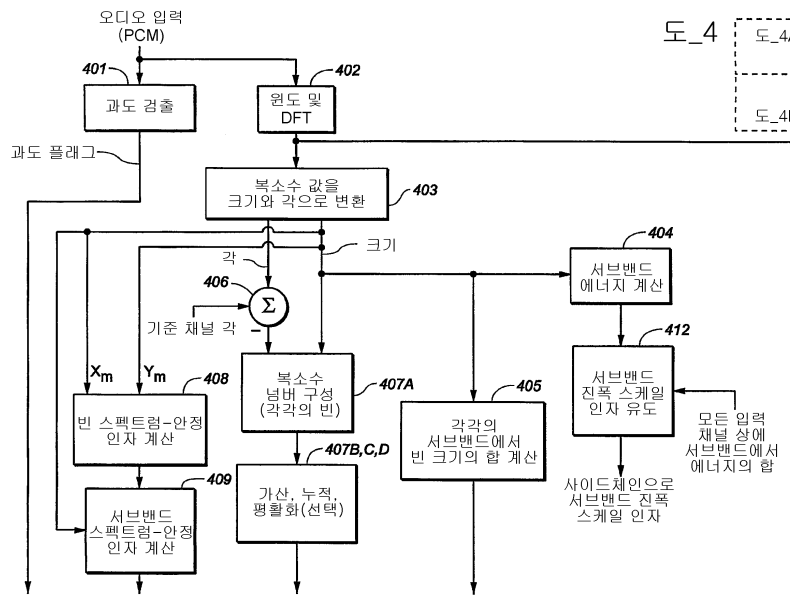
도면2



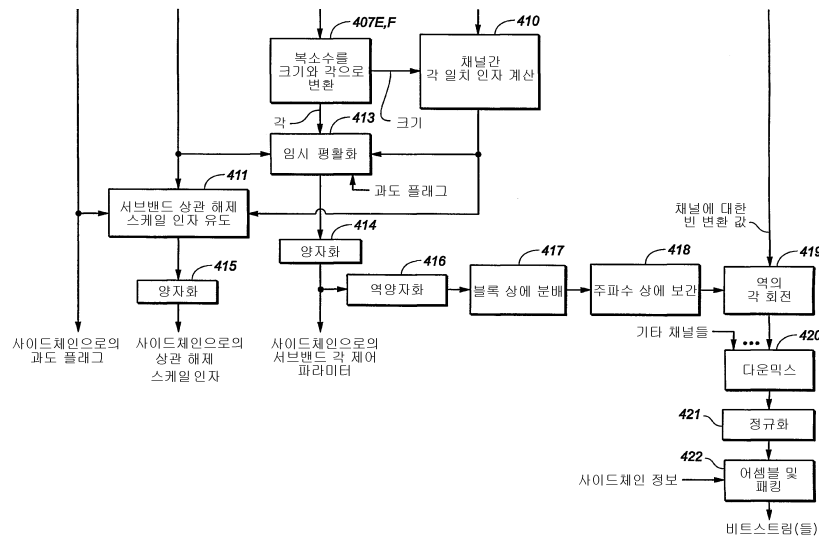
도면3



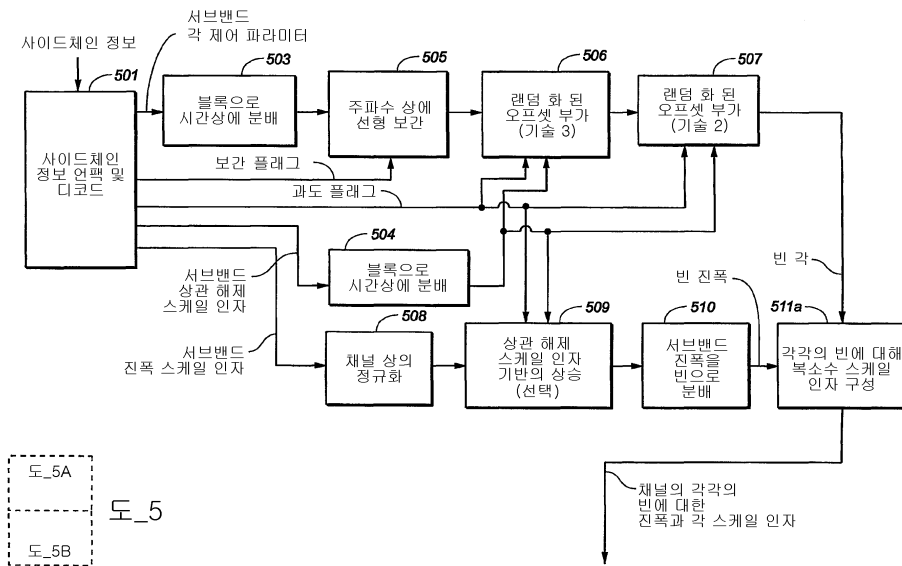
도면4A



도면4B

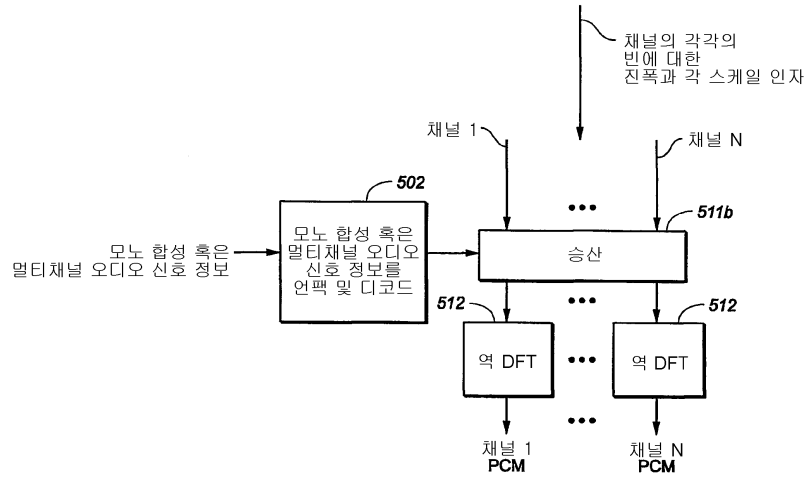


도면5A

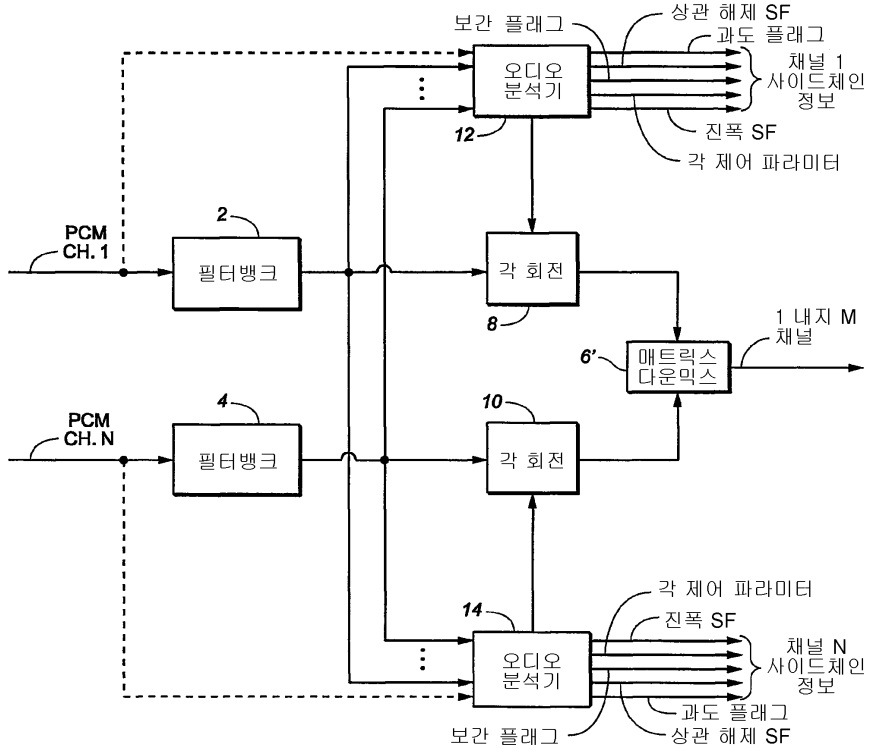


도_5A
도_5
도_5B

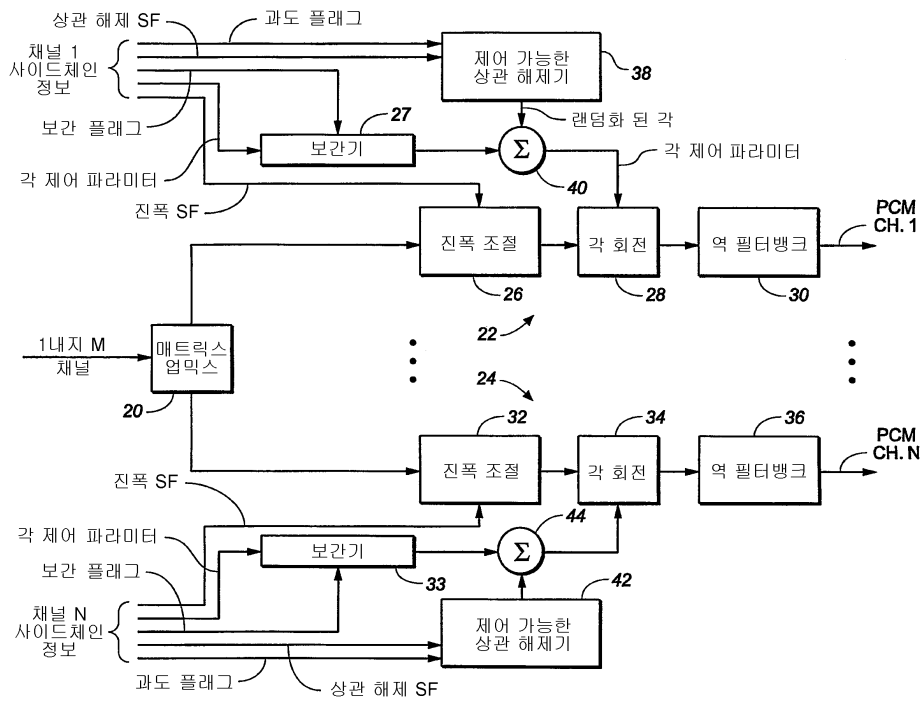
도면5B



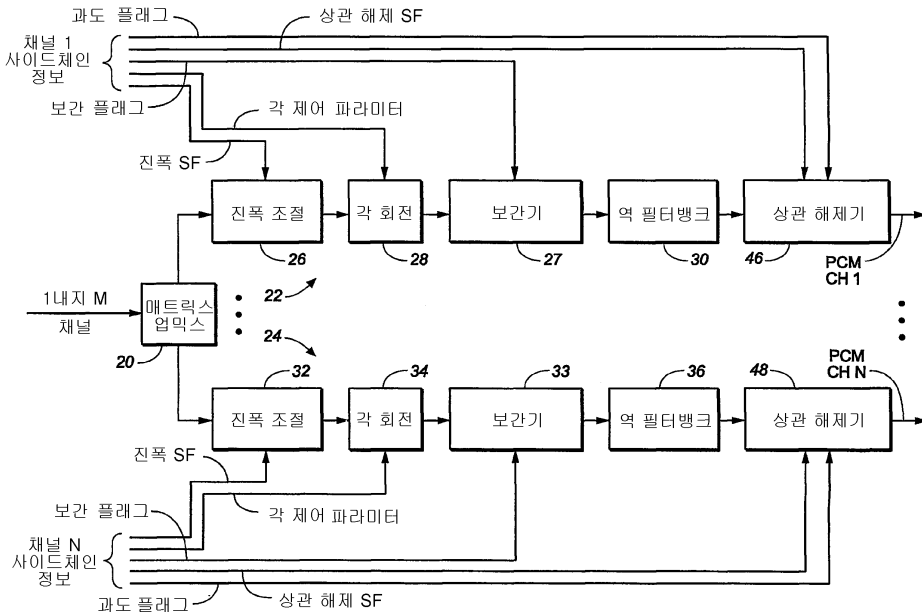
도면6



도면7



도면8



도면9

