

(19)대한민국특허청(KR)
(12) 등록특허공보(B1)

(51) 。 Int. Cl.
G11B 27/031 (2006.01)

(45) 공고일자 2006년09월13일
(11) 등록번호 10-0621424
(24) 등록일자 2006년08월31일

(21) 출원번호	10-2000-7004123	(65) 공개번호	10-2001-0024531
(22) 출원일자	2000년04월17일	(43) 공개일자	2001년03월26일
번역문 제출일자	2000년04월17일		
(86) 국제출원번호	PCT/US1998/021761	(87) 국제공개번호	WO 1999/21188
국제출원일자	1998년10월15일	국제공개일자	1999년04월29일

(81) 지정국 국내특허 : 오스트레일리아, 캐나다, 일본, 대한민국,

 EP 유럽특허 : 오스트리아, 벨기에, 스위스, 리히텐슈타인, 독일, 덴마크, 스페인, 프랑스, 영국, 그리스, 아일랜드, 이탈리아, 룩셈부르크, 모나코, 네덜란드, 포르투갈, 스웨덴, 핀란드,

(30) 우선권주장 08/953,618 1997년10월17일 미국(US)

(73) 특허권자 돌비 레버러토리즈 라이쎄싱 코오폰레이션
 미국 캘리포니아주 94103-4813 샌프란시스코 포트레로 애비뉴 100

(72) 발명자 필더,루이스,던
 미국,캘리포니아94103,샌프란시스코,포트레로애비뉴100

(74) 대리인 박경재

심사관 : 변성철

(54) 다이내믹 오디오 프레임 배열에 의해 비디오/오디오데이터 동기된 프레임 기반 오디오 코딩

요약

둘 또는 그 이상의 다른 정보 스트림들을 결합 편집하여 형성된 정보 스트림에 의해 재현되는 오디오의 질을 개선시키기 위하여 몇 가지 오디오신호 처리 기술들이 여러 가지 조합의 형태로 사용될 수 있다. 이 기술들은 오디오 정보를 비디오 정보에 결합시키는 적용분야에서 특히 유용하다. 이 중 하나의 기술에 의하면, 오디오 정보 스트림과 함께 전달되는 이득-제어 워드들이 결합부분을 가로질러 재생 음 레벨들을 보간하는데 사용된다. 다른 기술에 의하면, 특수한 필터뱅크들이나 특수한 형태의 TDAC 변환들이 결합부분의 어느 한 쪽에 있는 에일리어싱 아티팩트들을 억제하는데 사용된다. 또 다른 기술에 의하면, 특수한 필터뱅크들이나 크로스페이드 윈도우 함수들이 결합부분에서 생성되는 분광 스플래터의 감쇠를 최적화시키는데 사용된다. 또 다른 기술에 의하면, 오디오 샘플 레이트들이 예를들어, 오디오 정보가 비디오 정보와 함께 결합될 수 있도록 프레임 길이와 프레임 레이트에 따라서 변환된다. 또 다른 기술에 의하면, 결합부분을 가로질러 적절한 동기가 유지되도록 오디오 블록들이 다이내믹하게 정렬된다. NTSC 비디오와 결합되는 48 kHz 오디오의 실례가 설명된다.

대표도

도 4a

명세서

기술분야

본 발명은 오디오 정보 스트림들이 정보 프레임 형태로 배열되도록 한 오디오 신호처리에 관한 것이다. 보다 상세하게는, 본 발명은 프레임 기반 오디오 정보 스트림들을 결합시켜 형성된 오디오 정보 스트림들의 오디오 품질을 향상시키는 것과 관련된 것이다.

배경기술

오디오 또는 비디오 물의 편집 공정은 필수적으로 오디오 또는 비디오 물의 두 개의 세그먼트들을 함께 결합하거나 잘라내는 공정이다. 간단한 편집의 전형적인 예는 동화상(motion picture) 필름을 자르고 결합하는 공정이다. 이와 같이 결합시킬 두 개의 세그먼트들은 서로 다른 소스, 즉, 서로 다른 채널의 오디오 정보로부터 얻어지거나 또는 같은 소스로부터 얻어진다. 어느 경우든지, 그러한 결합은 사람이 인식할 수 있거나 또는 인식할 수 없는 오디오 또는 비디오 물 내의 불연속성을 만들어 내게 된다.

오디오 코딩

블록 처리

디지털 오디오의 사용 증가에 따라 가청 아티팩트(audible artifact)를 발생시키지 않고 오디오 물을 편집하는 것이 점점 더 어려워지고 있다. 이것은 디지털 오디오가 주로 블록으로서 처리되어야 하는 디지털 샘플 블록들로 빈번하게 처리 및 인코딩되기 때문에 발생한다. 많은 지각적(perceptual) 또는 심리음향적(psychoacoustic) 기반의 오디오 코딩 시스템들은, 원래의 신호를 그 복사물로서 복구하기 위하여 블록들로 역 변환되거나 합성 필터링되어야 하는 변환(transform) 계수들을 변경(convert)시키기 위해 또는 신호 샘플 블록들을 인코딩된 서브밴드(subband) 신호 샘플들로 바꾸기 위해, 변환기들을 이용하거나 또는 필터뱅크(filterbank)들을 이용한다. 처리된 오디오 신호의 편집은 최소한 블록 경계에서 이루어져야 한다. 최소한, 처리된 오디오 신호의 편집은 블록 경계에서 수행되어야 한다. 그렇지 않으면, 나머지 부분의 블록에 나타나는 오디오 정보는 적절하게 복구되지 못한다.

이하의 설명에서 "코딩"이나 "코더"와 같은 용어는 신호처리를 위한 여러 가지 방법과 장치들에 관련되며, "인코딩된"은 그와 같은 신호처리 결과에 관련된다. 이들 용어들의 어떤 것도 어떤 신호에서의 정보의 무관성(irrelevancy)이나 용장성(redundancy)을 감소시키는 것과 같은 신호처리의 특정형태를 의미하지는 않는다. 예를 들어, 코딩은 신호를 나타내기 위하여 펄스코드 변조(PCM) 샘플들을 발생시키는 것과 정보를 어떤 특정 사양에 따른 패킷이나 포맷으로 배열하는 것을 포함한다. 본 명세서에서 사용되는 "블록"이나 "프레임"과 같은 용어들은 정보의 그룹이나 그 간격과 관련되는데, 이 용어들은 종종 AES-3/EBU 디지털 오디오 표준으로 알려진, ANSI S4.40-1992 표준과 같이 다른 곳에서 사용되는 동일한 용어와는 다를 수 있다. 여기에서 사용되는 "필터"나 "필터뱅크"는 사각 미러 필터(quadrature mirror filter; QMF) 및 변환기들과 같은 재귀적(recursive) 및 비재귀적(non-recursive) 필터링의 형태를 필수적으로 포함하며, "필터링된" 정보는 그와 같은 필터들을 적용한 결과가 된다. 좀 더 상세한 설명은 변환기들에 의해 구현된 필터뱅크에서 이루어진다.

삭제

삭제

삭제

삭제

삭제

삭제

삭제

오디오 및 비디오 코딩

프레임 동기

오디오 및 비디오 정보를 함께 처리하는 편집에 있어서는, 적어도 두 가지의 이유로 인하여 더욱 큰 제한을 받게 된다. 한 가지 이유는 비디오 프레임 길이가 일반적으로 오디오 블록 길이와 같지 않다는 것이다. 두 번째 이유는 오디오 샘플 레이트(rate)의 정수배가 아닌 비디오 프레임 레이트를 갖는 NTSC와 같은 특정 비디오 표준에만 관련된다. 다음에 설명될 모든 실례들은 초당 48 k 샘플의 오디오 샘플 레이트를 갖는 것으로 가정한다. 대부분의 전문적인 장비는 이 샘플 레이트를 사용한다. 초당 44.1 k 샘플의 다른 샘플 레이트가 이와 유사하게 적용될 수도 있는 바, 이는 통상적으로 소비자의 장비에 사용되는 샘플 레이트이다.

몇 가지 비디오 및 오디오 코딩 표준에 의한 프레임 및 블록 길이가 테이블 I 및 테이블 II에 각각 나타나 있다. 테이블 내의 표제어들인 "MPEG II"와 "MPEG III"는 표준 ISO/IEC 13818-3 내의 국제 표준 기구의 동화상 기술자 그룹에 의해 설정된 MPEG-2 Layer II와 MPEG-2 Layer III 코딩기술에 관련된다. 표제어 "AC-3"은 Dolby Laboratories, Inc.에 의해 개발되고 표준 A-52 내의 Advanced Television Committee에 의해 설정된 코딩 기술에 관련된다. 48 kHz PCM에 대한 "블록길이"는 인접한 샘플들 사이의 시간 간격이다.

[표 1]

비디오 표준	프레임 길이
DTV (30Hz)	33.333 msec.
NTSC	33.367 msec.
PAL	40 msec.
필름	41.667 msec.
비디오 프레임 테이블 I	

[표 2]

오디오 표준	블록 길이
PCM	20.8 μ sec.
MPEG II	24 msec.
MPEG III	24 msec.
AC-3	32 msec.
오디오 프레임 테이블 II	

비디오 및 오디오 정보를 함께 결합시키는 적용분야에서, 오디오 블록들과 비디오 프레임들이 동기되는 것은 매우 드물다. 오디오/비디오 동기 발생 사이의 시간 간격이 테이블 III에 나타나 있다. 예를 들어, 이 테이블은, 초당 24 프레임으로 동화상 필름이 정확히 각 3 초 동안에 한 번씩 MPEG 오디오 블록 경계와 동기될 것이며, AC-3 오디오 블록과는 정확히 각 4 초 동안에 한 번씩 동기될 것임을 나타낸다.

[표 3]

오디오표준	DTV (30 Hz)	NTSC	PAL	필름
-------	-------------	------	-----	----

PCM	33.333 msec.	166.833 msec.	40 msec.	41.667 msec.
MPEG II	600 msec.	24.024 sec.	120 msec.	3 sec.
MPEG III	600 msec.	24.024 sec.	120 msec.	3 sec.
AC-3	800 msec.	32.032 sec.	160 msec.	4 sec.
오디오/비디오 동기 사이의 시간 간격 테이블 III				

이렇게 비디오 프레임들에 대한 오디오 블록들의 수로 표현되는 동기 발생 사이의 간격이 테이블 IV에 나타나 있다. 예를 들어, 5 개의 오디오 블록들과 4 개의 비디오 프레임들에 의한 간격 안에서 AC-3 블록들과 PAL 프레임들 사이에 동기가 정확하게 한 번 발생한다. 따라서, NTSC 비디오의 5 개의 프레임들이 PCM 오디오의 8,008 개의 샘플들과 동기되어야 한다. 이 관계가 아래에 기술된다

[표 4]

오디오표준	DTV (30 Hz)	NTSC	PAL	필름
PCM	1600 : 1	8008 : 5	1920 : 1	2000 : 1
MPEG II	25 : 18	1001 : 720	5 : 3	125 : 72
MPEG III	25 : 18	1001 : 720	5 : 3	125 : 72
AC-3	25 : 24	1001 : 960	5 : 4	125 : 96
: 오디오/비디오 동기 사이의 프레임 수 테이블 IV				

비디오 및 오디오 정보가 함께 결합될 때, 일반적으로 비디오 프레임 경계에서 편집이 이루어진다. 테이블 III 및 IV에서 나타난 정보로부터 이와 같은 편집이 오디오 프레임 경계에서는 드물게 발생한다는 것을 알 수 있다. 예를 들어, NTSC 비디오 및 AC-3 오디오에 있어서, 비디오 경계에서의 편집이 오디오 블록 경계에서도 발생할 수 있는 확률은 단지 1/960, 또는 약 0.1%에 불과하다. 물론, 결합체(splice)의 어느 한 쪽에서 발생하는 두 편집은 이와 같은 방식으로 동기되어야 하며, 그렇지 않으면, 몇몇 오디오 정보가 손실될 것이다. 따라서, 두 개의 임의 편집(random edits)을 위한 NTSC/AC-3 정보의 결합체는 오디오 블록 경계와는 다른 곳에서 발생하며, 이에 따라 하나 또는 두 개의 오디오 정보 블록이 손실될 것이 거의 확실하다. 그러나, AC-3가 TDAC 변환을 사용하기 때문에, 정보 블록들의 손실이 발생하지 않는 경우에도 상기의 이유로 인해 상쇄되지 않은 에일리어싱 변형이 발생하게 된다.

이 문제는 상기한 오디오 블록 처리 문제들과 유사하다. 종래의 기술에 따른 방법들과 장치들은 비디오/오디오 프레임 문제를 간과하거나, 또는 유사한 불만족스러운 해결책을 제시하고 있는데, 이것은 오디오 정보를 비디오 정보로부터 분리하고, 인코딩된 오디오 정보를 디코딩하며, 복구된 오디오 정보를 편집하고, 이 오디오 정보를 비디오 정보와 함께 재 인코딩 및 재 결합(re-bundling)시킴에 의해, 오디오를 "후처리" 하는 것이다. 이런 기술의 실시예가 제 US-A-4,903,148호에 개시된다.

데이터 동기

상기한 바와 같이, NTSC 비디오의 5 개의 프레임들이, 초당 48 k 개의 샘플 속도로서 PCM 오디오의 8008 개의 샘플들과 동기되어야 한다. 즉, NTSC 비디오 프레임들은 오디오 정보를 정수개의 샘플들로 구분하지 않는다. 각 NTSC 프레임은 1601.6 샘플들에 상응한다. 이와 유사하게, NTSC 프레임들은 인코딩된 오디오 정보를 정수개의 샘플들이나 계수들의 블록들로 구분하지 않는다. 이것은 오디오 샘플들을 예를 들어, 각각 1602, 1601, 1602, 1601 및 1602 개의 샘플들을 포함하는 오디오 프레임의 반복적 시퀀스로 배열함에 의해 조절될 수 있다. 그러나, 이것은, 편집이 5-프레임 시퀀스의 시작점(이하, 여기에서 "슈퍼프레임(superframe)"으로 언급됨)에서만 행해져야 하기 때문에, 편집 응용분야에 더욱 큰 제한을 갖게 된다. 불행하게도, 많은 응용분야에서 비디오 정보나 이 비디오와 묶여져 있는 오디오 정보 중 어느 것도 슈퍼프레임 경계의 표시를 수반하지 않는다.

슈퍼프레임 내에서의 가변길이 오디오 블록들은 많은 코딩 응용분야에서 또 다른 문제를 발생시킨다. 위에서 설명한 바와 같이, 많은 코딩 응용분야에서는 인코딩된 정보를 블록으로 처리한다. 만약, 신호가 어떤 형태의 동기신호를 수반하지 않

을 경우, 디코더는 각 슈퍼프레임에 대하여 어디에 경계가 있는지, 또는 편집이 슈퍼프레임의 부분을 제거했는지의 여부에 대하여 알 수가 없다. 즉, 디코더는 각 오디오 프레임이나 블록에 대하여 어디에 경계가 있는지를 알지 못한다. 상기 블록 경계에서의 불확실성을 하나의 샘플만큼 작게 줄이는 것이 가능할 수 있으나, 오디오 정보가 블록으로 처리될 때, 하나의 샘플이라도 충분히 복구된 오디오 정보의 복구를 방해할 수 있다.

1985년 10월 25일 공개된, 일본 특허 요약 공보 제 JP-A-60,212,874호는 오디오 샘플링 레이트가 비디오 프레임 레이트의 정수배가 아닐 때 오디오 및 비디오 정보를 기록 및 재생하도록 비디오 테이프 레코더(VTR)를 사용하기 위한 기술을 개시한다. 이런 기술에 따라, 의사(dummy) 샘플들은 변화하는-길이 필드 또는 오디오 샘플들의 블록들에 추가되어 고정-길이 블록들을 생성시킨다. 오디오 정보 및 의사 샘플들의 블록들은 시간-압축되어 비디오 정보로 기록된다. 재생동안, 블록들은 시간-확장되며, 의사 샘플들은 제거되고 연속적인 출력 오디오 신호는 나머지 오디오 정보로부터 발생된다. 불행하게도, 이 기술은 편집이 슈퍼프레임의 시작부분 이외의 어떤 다른 점에서라도 이루어진다면 비디오와 오디오 사이의 동기를 유지한다.

발명의 상세한 설명

본 발명의 목적은 비디오 및 오디오 정보의 프레임들 사이에 데이터 동기를 제공하므로써 두 개 또는 그 이상의 프레임 기반 오디오 정보 스트림들을 결합하여 형성된 오디오 정보 스트림에 의해 재현되는 오디오의 품질을 향상시키는 것이다. 본 발명의 한 특징에 의하면, 신호처리 방법 또는 장치는, 오디오 정보를 오디오 샘플 레이트로 재현하는 입력 샘플들을 포함하는 제 1 입력 신호를 수신하고; 정보를 입력 프레임 레이트로 반송하는 입력 프레임들을 포함하는 제 2 입력 신호를 수신하는데, 각각의 슈퍼 프레임은 상기 입력 프레임 레이트로 나누어진 상기 오디오 샘플 레이트가 정수가 아닌 상기 오디오 샘플 레이트의 곱이며 상기 입력 프레임 레이트로 나누어진 상기 제 1의 수는 실질적으로 정수와 등가이도록 제 1의 수와 등가인 다수의 상기 입력 프레임들을 포함하며; 제 1 입력 신호에 응답하여 오디오 프레임들의 시퀀스를 발생시키는데, 각각의 오디오 프레임은 해당 입력 프레임에 상응하고 상기 입력 샘플들의 시퀀스에 상응하는 인코딩된 오디오 정보를 포함하며, 상기 시퀀스는 상기 제 1 입력 신호로부터 취해지며 빠른 시작 샘플, 공칭 시작 샘플 및 뒤편의 정수 부분과 등가인 다수의 연속적인 샘플들을 포함하고, 상기 뒤편은 상기 입력 프레임 레이트로 나누어진 상기 오디오 샘플 레이트와 등가인 것으로서, 여기에서 상기 빠른 시작 샘플은 입력 샘플들의 상기 시퀀스 형태의 제 1 샘플이며 상기 공칭 시작 샘플은 상기 각각의 입력 프레임과 실질적으로 정렬되고; 출력 슈퍼프레임들로 그룹화된 출력 프레임들로 배열된 출력 신호를 발생시키는데, 각각의 출력 슈퍼프레임은 상기 제 1의 수와 등가인 다수의 상기 출력 프레임들을 포함하고, 각각의 출력 프레임은 각각의 오디오 프레임 및 상기 각각의 오디오 프레임에 대한 라벨을 포함하며, 상기 라벨은 각각의 출력 슈퍼프레임의 각각의 오디오 프레임에 대해 유일하다.

본 발명의 또 다른 특징에 의하면, 신호처리 방법 또는 장치는 완전 입력 슈퍼프레임 및 부분 입력 슈퍼프레임으로 그룹화된 입력 프레임들로 배열된 입력 신호를 수신하되, 각각의 완전 입력 슈퍼프레임은 하나 이상인 제 1의 수와 등가인 다수의 상기 입력 프레임들을 가지며 각각의 부분 입력 슈퍼프레임은 보다 적은 수의 상기 입력 프레임들을 가지며, 각각의 입력 프레임은 입력 프레임 레이트로 인코딩된 오디오 정보를 재현하는 오디오 프레임 및 상기 오디오 프레임과 결합된 라벨을 포함하는데, 여기에서, 상기 라벨은 각각의 완전 또는 부분 입력 슈퍼프레임의 해당 오디오 프레임에 대해 유일하고; 상기 오디오 프레임들로부터 샘플들의 시퀀스들을 유도하는 것으로, 여기에서 샘플들의 해당 시퀀스는 해당 오디오 프레임으로부터 유도되며 빠른 시작 샘플, 공칭 시작 샘플 및 제 2의 수와 등가인 다수의 연속적인 샘플들을 포함하고, 여기에서 상기 샘플들의 시퀀스는 오디오 샘플 레이트로 오디오 정보를 재현시키고 상기 제 2의 수는 뒤편의 정수부와 등가인데, 상기 뒤편은 상기 입력 프레임 레이트로 나누어진 상기 오디오 샘플 레이트와 등가이며; 샘플들의 각각의 시퀀스로부터 각각의 샘플들의 이어지는 연속물을 얻는데, 여기에서 각각의 샘플들의 시퀀스가 유도되는 오디오 프레임과 결합된 라벨에 응답하여, 상응하는 이어지는 연속물은 상기 각각의 시퀀스의 샘플들의 수 이하의 샘플들의 제 3의 수를 포함하며 빠른 시작 샘플, 공칭 시작 샘플 또는 상기 공칭 시작 샘플을 뒤따르는 샘플에서 시작하는데, 여기에서, 상기 제 3의 수는 제 2의 수 또는 상기 제 2의 수에 하나(1)를 더한 것과 등가이며; 각각의 연속물의 시작 및 바로 선행한 연속물의 시작이 상기 선행 연속물의 샘플들의 제 3의 수로 분리되는 연속물들의 배치로부터 출력 신호를 발생시킨다.

본 발명의 여러 가지 특징들과 그 바람직한 실시예들은 다음의 설명과 같은 부호가 같은 구성 요소들에게 부여된 첨부도면을 참조하여 더욱 잘 이해될 수 있다. 여러 가지 장치들을 도시한 도면들은 본 발명의 이해에 도움이 되는 주요 구성요소들을 나타내고 있다. 명확한 이해를 위하여, 이들 도면들은, 실제적인 실시예에서 중요할 수 있으나 본 발명의 개념들을 이해하는데 중요하지 않은 많은 다른 특징들을 생략하고 있다. 본 발명을 실시하는데 요구되는 신호처리는 마이크로프로세서, 디지털 신호 처리기, 로직 어레이 및 다른 형태의 계산회로들에 의해 수행되는 프로그램들을 포함하는 다양한 방법에 의해 달성될 수 있다. 신호 필터들은, 재귀적(recursive), 비-재귀적 및 래티스(lattice) 디지털 필터들을 포함하는 임의의 방법에 의해 달성될 수도 있다. 디지털 및 아날로그 기술이 응용분야의 필요 및 특성에 따라서 여러 가지 조합형태로 사용될 수 있다.

삭제

삭제

삭제

본 발명의 특징들은 보다 상세하게는 표제"다이나믹 오디오 프레임 배열"하의 단락에서 이 기술의 결말부에 기술되어 있다. 이 기술의 단락에 선행한 논의에서 본 발명의 특징을 보다 잘 이해하고 실효한 실시예들을 시행하기 위해 유용한 배경 물질을 제공한다. 좀더 상세한 설명은 오디오 및 비디오 정보 스트림들을 처리하도록 된 상태에서 이루어지게 되나, 본 발명의 특징들은 비디오 정보의 처리를 포함하지 않는 응용분야에서 실행될 수도 있다. 다음의 설명의 내용과 도면들은 오직 실예를 위한 것으로서, 본 발명의 범위를 제한하는 것으로 이해되지 않아야 한다.

도면의 간단한 설명

도 1a 및 1b는 블록들, 프레임들 및 슈퍼프레임들의 형태로 배열된 비디오 및 오디오 정보를 나타낸다.

도 2a 내지 2c는 윈도우 기능에 의해 변조된 중첩된 블록들과 그 결과로서 나타나는 상기 윈도우 블록들로 구성된 프레임들의 이득 프로파일을 나타낸다.

도 3은 에일리어싱 소거 변환에 의해 발생된 신호 및 에일리어싱 성분들을 나타낸다.

도 4a 내지 4c는 인코딩된 정보 스트림에서 이득-제어 워드들을 생성 및 변화시키고, 이 워드들에 응답하는 장치들의 기능 블록도이다.

도 5a 및 5b는 프레임 경계에서 에일리어싱 아티팩트들을 억제하기 위하여 대체 필터뱅크들을 적용하는 장치들의 기능 블록도이다.

도 6a 내지 6d는 프레임 경계에서 에일리어싱 아티팩트들을 억제하기 위하여 사용될 수 있는 윈도우 기능들을 나타낸다.

도 7은 프레임 경계에서 여러 가지 윈도우 기능들을 사용하여 얻어지는 주파수 응답 특성을 나타낸다.

도 8은 결합부분에서 분광 스플래터의 감쇠를 증가시키기 위하여 대체 필터뱅크들을 적용하는 장치의 기능 블록도이다.

도 9, 10a 및 11a는 도 8의 장치에 의해 수행되는 여러 가지 윈도우 기능들을 나타낸다.

도 10b 및 11b는 도 8의 장치에서의 여러 가지 윈도우 기능들을 사용하여 얻어지는 주파수 응답 특성을 나타낸다.

도 12a 및 12b는 오디오 샘플들과 비디오 프레임들 간의 동기를 이루기 위하여 샘플 레이트 변환(sample rate conversion)을 제공하는 장치들의 기능 블록도이다.

도 13a 및 13b는 결합부분과 교차하는 비디오 슈퍼프레임과 동기를 이루기 위하여 다이나믹 오디오 프레임 정렬을 제공하는, 본 발명에 따른 장치들의 기능 블록도이다.

도 14는 비디오 프레임 특성들과, 결합부분과 교차하는 다이나믹 오디오 프레임 정렬효과를 나타낸다.

실시예

본 발명을 수행하기 위한 모드들

신호들과 그 처리

신호 블록들과 프레임들

도 1a는 오디오 블록들(10-18)의 시퀀스 안에 배열된 인코딩된 오디오 정보와 비디오 프레임(1)과 같은 비디오 프레임들의 시퀀스 안에 배열된 비디오 정보의 스트림을 나타낸다. NTSC 비디오와 같은 몇몇 포맷에서, 각 비디오 프레임은 하나의 화상이나 영상을 공동으로 설정하는 두 개의 비디오 필드들로 구성된다. 오디오 블록들(11-17)은 비디오 프레임(1)과 함께 인코딩된 신호 프레임(21)으로 그룹지어진다.

전술한 그리고 테이블 IV에 나타난 바와 같이, 몇몇 응용분야에서의 비디오 프레임들은 인코딩된 오디오를 정수배의 샘플들이나 변환계수들 등으로 분리하지 않는다. 이것은 인코딩된 신호 프레임들의 그룹들을 해당 슈퍼프레임들로 배열함에 의해 조정될 수 있다. 슈퍼프레임(31)으로 분류된 5 개의 인코딩된 신호 프레임들(21-25)의 배열이 도 1b에 나타나 있다. 이와 같은 특정 배열은 NTSC 비디오와 48 k 샘플/초 PCM 오디오를 사용하는 응용분야에서 사용될 수 있다.

처리된 신호 블록들

인코딩된 오디오 정보 블록들의 시퀀스는 오디오 신호의 중첩 간격들을 나타낼 수 있다. 예를 들어, 몇몇 분할-밴드 지각적 코딩 시스템은, 블록 길이의 절반만큼 서로 중첩된 오디오 샘플들의 블록들을 처리한다. 통상적으로, 이 중첩된 블록들 내의 샘플들은 분석 윈도우 기능에 의해 변조된다.

도 2a는 중첩된 오디오 블록들의 시퀀스 내의 각 블록에 적용된 분석 윈도우 기능의 변조 엔벨로프(envelope)(61-67)를 나타낸다. 이 중첩의 길이는 블록 길이의 1.5배이다. 이 중첩 간격은 상기한 O-TDAC 변환과 프린센(Princen), 존슨(Johnson) 및 브래들리(Bradley)의, "Subband/Transform Coding Using Filter Bank Designs Based on Time Domain Aliasing Cancellation", ICASSP 1987 Conf. Proc. May 1987, pp.2161-64에 기술된 하나의 중첩된-블록 변환 같은 몇몇 신호 분석-합성 시스템에서 일반적으로 사용된다. 이런 변환은 홀수-적층된 임계적으로 샘플링된 단일-측대역 분석-합성 시스템의 시간-영역 등가물이며 본문에서 홀수-적층된 시간-영역 앨리어싱 소거(O-TDAC)로 칭해진다. 포워드 변환은 블록의 절반마다 서로 중첩되고 둘로 변환 계수들을 데서메이팅하여 임계 샘플링을 달성하는 샘플들의 블록들에 적용되며; 그러나, 이런 데서메이팅에 의한 정보 손실은 회복된 신호에 시간-영역 앨리어싱을 생성시킨다. 합성처리는 합성된 샘플들의 블록들을 발생시키기 위해 역변환을 변환계수들의 블록들에 적용하고, 합성된 샘플들의 블록들에 적절히 형성된 합성 윈도우 함수를 적용하며, 윈도우 블록들을 중첩하고 부가하므로써 이런 앨리어싱을 소거시킬 수 있다. 실제로, TDAC 코딩 시스템이 블록들 B_1 - B_2 의 시퀀스를 발생시킨다면, 블록 B_1 의 마지막 절반 및 블록 B_2 의 첫번째 절반의 앨리어싱 아티팩트는 서로 소거시킬 것이다.

도 2b는 인코딩된 신호 프레임에 대한 중첩 블록들의 시퀀스에 적용된 윈도우 기능의 변조 엔벨로프를 나타낸다. 도 2b에 도시된 바와 같이, 이 변조의 네트(net) 효과 또는 이득 프로파일(81)은 중첩 간격들 내의 인접 블록들에 대한 변조 엔벨로프들(71-77)의 합이 된다. 바람직하게는, 각 중첩부분과 교차하는 네트 효과는 단일 이득이 되어야 한다.

도 2c는 인접한 인코딩된 신호 프레임들과 교차하는 윈도우 기능 변조의 전체 효과를 나타낸다. 도시된 바와 같이, 이득 프로파일들(80-82)은 네트 효과가 단일 이득을 갖도록 중첩 및 추가된다.

분석 윈도우 기능만을 사용하는 시스템들에 있어서, 모든 윈도우 기능 변조의 효과는 분석 윈도우 기능만의 변조 효과들과 동등하다. 이상적인 이득 프로파일은, 분석 윈도우 기능의 변조 엔벨로프가 어떤 상수에 중첩 및 추가됨을 보증함으로써, 달성될 수 있다.

분석 및 합성 윈도우 기능들을 사용하는 시스템들에 있어서, 모든 윈도우 기능 변조의 네트 효과는 분석 윈도우 기능과 합성 윈도우 기능의 곱으로부터 형성된 "곱" 윈도우 기능의 네트 효과와 동등하다. 이와 같은 시스템들에 있어서, 이상적인 이득 프로파일은 곱 윈도우 기능의 변조 엔벨로프를 중첩 간격 내의 어떤 상수에 가산시킴에 의해 얻어질 수 있다.

이상에서와 같이, 분석 및 합성 윈도우 기능들을 함께 사용하는 코딩 시스템 및 방법들에 대하여 몇 가지 설명이 이루어졌다. 이와 관련해서, 중첩 분석 윈도우 기능들에 의한 이득 프로파일은 때때로 어떤 상수와 같다고 말하여 질 것이다. 마찬가지로, 중첩 합성 윈도우 기능들에 의한 이득 프로파일은 때때로 어떤 상수와 같다고 말하여 질 것이다. 그와 같은 설명은 시스템에서의 모든 윈도우 기능의 네트 변조 효과와 관련이 있음을 알아야 한다.

윈도우 기능

분석 윈도우 기능의 형태는 신호의 이득 프로파일에 영향을 미치는 물론 해당 필터뱅크의 주파수 응답 특성에도 영향을 미친다.

분광 스플래터(spectral splatter)

상기한 바와 같이, 많은 지각적 분할밴드 코딩 시스템(perceptual split-band coding system)들이 좀더 넓은 필터 통과대역과 대체된 필터 저지대역 내에서 주파수 감쇠를 증가시킴에 의해 지각적 코딩에 대한 최적의 주파수 응답 특성을 갖는 필터뱅크들을 사용하고 있다. 불행하게도, 결합 편집은 필터 저지대역으로 간주되는 주파수 범위 내가 아닌 주파수 범위에서 현저한 분광 아티팩트 또는 "분광 스플래터"를 생성하는 경향이 있다. 일반적인 지각적 코딩 성능을 최적화하도록 설계된 필터뱅크들은 결합 편집 시에 생성되는 이와 같은 분광 아티팩트들이 들리지 않도록 충분한 감쇠 기능을 발휘하지 않는다.

TDAC 변환 에일리어싱 소거

상기의 O-TDAC 변환에 대하여, 상기 분석 윈도우 기능은, 합성 변환의 적용 후에 적용되는 합성 윈도우 기능과 함께, 시간 영역 에일리어싱 아티팩트(aliasing artifact)들의 소거를 위하여 많은 제약 조건들을 만족시켜야만 한다.

합성 변환으로부터 복구된 신호는 원래의 신호와 상기 분석 변환에 의해 발생된 시간 영역 에일리어싱 성분들의 합으로서 개념지어진다. 도 3에서, 곡선 91, 93 및 95는 역 변환 또는 합성 변환으로부터 복구되고 분석 및 합성 윈도우 기능에 의해 변조된 입력 신호의 진폭 엔벨로프의 세그먼트들을 나타낸다. 곡선 92, 94 및 96은 역 변환 또는 합성 변환으로부터 복구되고 분석 및 합성 윈도우 기능에 의해 변조된 시간 영역 에일리어싱 성분들을 나타낸다. 도면에 도시되고 이후에 설명되는 바와 같이, 시간 영역 에일리어싱 성분들은 분석 및 합성 윈도우 기능에 의해 변조된 원래의 입력 신호의 복제물들로 나타난다.

분석 및 합성 O-TDAC 변환의 주요 기능들은 블록의 각각의 반쪽 내에서 윈도우 신호의 끝-대-끝(end-for-end)의 형태를 갖는 시간 영역 에일리어싱 성분들을 발생시키도록 설계된다. Princen에 개시되어 있듯이, O-TDAC 변환은 서로 다른 두 개의 영역에서 시간 영역 에일리어싱 성분들을 발생시킨다. 영역 2에서, 시간 영역 에일리어싱 성분은 해당 영역에서 원래의 신호의 끝-대-끝 윈도우 형태를 갖는다. 영역 1에서, 시간 영역 에일리어싱 성분은 해당 영역 내에서의 입력 신호의 끝-대-끝 윈도우 형태를 가지나, 그 진폭은 반전된다.

예를 들어, 에일리어싱 성분(94a)은 신호 성분(93a)의 끝-대-끝 윈도우 형태를 가지며, 에일리어싱 성분(92b)은 신호 성분(91b)의 끝-대-끝 윈도우 형태를 가지나 그 성분의 진폭은 반전된다.

인접한 블록들을 중첩 및 가산함에 의해 원래의 신호가 복구되고 에일리어싱 성분들은 상쇄된다. 예를 들어, 신호 성분들(91b, 93a)은 윈도우 기능 변조 효과 없이 신호를 복구하도록 가산되며, 에일리어싱 성분들(92b, 94a)은 에일리어싱을 상쇄시키도록 가산된다. 마찬가지로, 신호 성분들(93b, 95a)은 신호를 복구하도록 가산되며, 에일리어싱 성분들(94b, 96a)은 에일리어싱을 상쇄시키도록 가산된다.

결합부분 바로 앞의 합성 오디오 샘플들의 반 블록 안에 있는 에일리어싱 아티팩트들이 상기 결합 경계의 바로 뒤의 합성 오디오 블록의 반 블록 안에 있는 반전된 에일리어싱 아티팩트들이 아니기 때문에 결합 경계의 어느 한 쪽에 있는 시간 영역 에일리어싱 아티팩트들은 일반적으로 상쇄되지 않을 것이다.

마찬가지로, Princen과 Bradley의 "시간 영역 에일리어싱 소거에 기반을 둔 분석/합성 필터뱅크의 설계", IEEE Trans. on Acoust., Speech, Signal Proc., vol. ASSP-34, 1986, pp. 1153-1161에 기재된 것과 같은, 다른 에일리어싱 소거 필터뱅크들에 상기의 사항들이 적용된다. 이 필터뱅크 시스템은 짝수적으로 적층되고 임계적으로 샘플링된 신호-사이드밴드 분석-합성 시스템의 시간 영역 등가물이며, 이것은 여기에서 짝수적으로 적층되는 시간 영역 에일리어싱 소거(E-TDAC)로 칭한다.

결합부분에서의 아티팩트들을 감쇠시키기 위한 이득제어

결합에 의해 생성된 아티팩트들의 가청성을 감소시키기 위하여 사용될 수 있는 하나의 기술은 재생신호의 진폭을 변경시키기 위하여 디코더나 재생 시스템에 명령하는 다수의 이득-제어 워드들을 인코딩된 오디오 신호로 결합시키는 것이다. 이와 같은 이득-제어 워드들을 사용하는 장치들의 간단한 실예들이 다음에 언급된다.

도 4a는 포맷터(111)가 경로 112 를 통해 비디오 정보, 다중 오디오 채널들을 나타내는 인코딩된 오디오 정보, 그리고 이득-제어 워드들로 구성된 프레임들의 형태로 배열된 출력신호를 발생시키도록 된 장치(100)의 기능 블록도이다. 포맷터(111)는 경로 108 을 통해 수신되는, 비디오 정보와 다중 오디오 채널들에 대한 인코딩된 오디오 정보를 수반하는 프레임들의 형태로 배열된 신호에 응하여, 그리고 경로 110 을 통해 수신되는, 이득-제어 워드들을 수반하는 신호에 응하여 상기 출력신호를 발생시킨다. 프로세서(109)는 각각 상기의 다중 오디오 채널들 중의 하나와 연관된 경로(103a, 103b)로부터의 다중 제어신호들을 수신하고, 이들 각 제어신호에 응하여 경로 110 을 통해 각 프레임 내에서의 시작 이득과 종료 이득을 나타내는 관련 오디오 채널에 대한 한 쌍의 이득-제어 워드들을 발생시킨다. 명확한 이해를 위하여, 도면에는 오직 두 개의 제어신호들(103)과 두 개의 관련 오디오 채널들(102)이 도시되어 있다. 이와 같은 이득제어 기술은 필요에 따라 두 개를 초과하는 채널들에 적용될 수 있다.

본 발명의 실시예에 있어서, 인코더(105)는 경로 102a 및 102b 로부터 수신되는 다중 오디오 채널 신호들에 응하여 다중 오디오 채널들에 대한 인코딩된 오디오 정보를 발생시키며, 프레이머(107)는 경로 101 로부터 수신되는 비디오 정보와 경로 106a 및 106b 로부터 수신되는 인코딩된 오디오 정보를 프레임들의 형태로 배열함에 의해 경로 108 을 통해 상기 신호를 발생시킨다.

이와 같은 이득제어 기술은 경로 108 을 통하는 상기 신호와 유사한 입력 신호들에 사용될 수 있다. 따라서, 인코더(105)나 프레이머(107)가 필요하지 않게 된다. 인코더(105)를 포함하는 실시예들에 있어서, 인코딩은 각 오디오 채널에 독립적으로 적용되거나, 또는 다중 오디오 채널들에 공동으로 적용될 수 있다. 예를 들어, AC-3 인코딩 기술은 채널들 사이의 중복성을 제거 또는 감소시킴에 의해 전체 대역폭 요구사항들을 낮추기 위하여 둘 또는 그 이상의 오디오 채널들에 공동으로 적용될 수 있다.

도 4c는 입력신호 내의 이득-제어 워드들에 따라서 다중 오디오 채널들을 재생하기 위하여 출력신호들을 발생시킨다. 디포맷터(142)는 경로 141 로부터 비디오 정보, 인코딩된 오디오 정보, 그리고 이득-제어 워드들로 구성된 프레임들의 형태로 배열된 입력신호를 수신한다. 이 디포맷터(142)는 입력신호의 각 프레임으로부터 다중 오디오 채널들을 나타내는 인코딩된 오디오 정보를 얻으며, 각 오디오 채널과 연관된 한 쌍의 이득-제어 워드들을 얻는다. 프로세서(148)는 경로 145 로부터 이득-제어 워드들을 수신하고, 이에 따라 경로들 149a 및 149b 를 통해 이득 제어 신호들을 발생시킨다. 디코더(146)는 경로들 144a 및 144b 로부터 인코딩된 오디오 정보의 다중 채널들을 수신하고, 이에 따라 각 출력 신호의 진폭 또는 레벨이 관련된 이득 제어 신호에 응하여 가변되도록 각 오디오 채널에 대한 출력신호를 발생시킨다.

한 쌍의 이득-제어 워드들은 특정 프레임 내에서 각 오디오 채널에 대한 시작 이득과 종료 이득을 나타낸다. 프로세서(148)는 각 이득-제어 워드 쌍의 보간(interpolation)을 나타내는 이득 제어 신호들을 발생시킨다. 이 보간은 직선, 직사각형, 대수 또는 지수함수와 같은 원하는 형태의 경로를 따를 수 있다. 예를 들어, 직선 보간에 있어서, 이득 제어 신호는 특정 프레임과 교차하여 직선적으로 변하는 이득을 나타낼 것이다.

디코딩은 각 오디오 채널에 독립적으로 적용되거나, 또는 다중 오디오 채널들에 공동으로 적용될 수 있다. 예를 들어, 디코딩은 채널들 사이의 중복성을 제거 또는 감소시키는 인코딩의 형태에 상보적(complementary)으로 될 수 있다. 합성 필터뱅크와 합성 윈도우 기능을 사용하는 분할밴드 코딩에 있어서, 출력 신호는 합성 필터뱅크의 적용 전에 인코딩된 오디오를 변형시키거나, 합성 윈도우 처리 이전에 합성 필터뱅크로부터 얻어지는 합성 오디오를 변형시키거나, 또는 합성 윈도우 기능의 적용에 의해 얻어진 오디오 정보를 변형시킴에 의해, 이득 제어 신호에 따라서 효과적으로 변조될 수 있다.

도 4b는 신호 내의 기존의 이득-제어 워드들을 변형시키는 장치(120)의 기능 블록도이다. 디포맷터(123)는 경로 121 로부터 비디오 정보, 다중 오디오 채널들을 나타내는 인코딩된 오디오 정보, 그리고 입력 이득-제어 워드들로 구성된 프레임들의 형태로 배열된 입력신호를 수신한다. 이 디포맷터(123)는 상기 입력신호로부터 상기 다중 오디오 채널들 중의 하나에 대한 인코딩된 오디오 정보와 관련된 하나 또는 그 이상의 입력 이득-제어 워드들을 얻으며, 이 입력 이득 제어 워드들을 경로 124a 및 124b 로 출력한다. 프로세서(126)는 경로 122 로부터 수신되는 제어 신호에 응하여 하나 또는 그 이상의 입력 이득-제어 워드들을 변형시킴에 의해 경로 127 을 통해 하나 또는 그 이상의 출력 이득-제어 워드들을 발생시킨다. 포맷터(128)는 경로 129 를 통해 비디오 정보, 다중 오디오 채널들에 대한 인코딩된 오디오 정보, 출력 이득-제어 워드들, 그리고 이 출력 이득-제어 워드들에 상응하지 않는 입력 이득-제어 워드들을 포함하는 프레임들의 형태로 배열된 출력신호를 발생시킨다.

편집에 있어서, 제어 신호(122)는 입력신호(121) 내의 하나의 결합부분을 나타낸다. 이에 응하여, 프로세서(126)는 상기 장치 140 과 같은 장치가 상기 결합 바로 이전에 재생신호를 감쇠시키고 상기 결합 바로 이후에는 상기 감쇠를 역전시키도록 하나 또는 그 이상의 출력 이득-제어 워드들을 발생시킨다. 이득의 변화는 여러 개의 프레임들을 가로질러 나타낼 수

있으나, 많은 적용 분야에서 이와 같은 변화는 상기 결합의 어느 한 쪽에 있는 하나의 프레임으로 제한된다. 이 이득변화 구간은 이득변화에 의해 발생된 변조 결과들의 가청도를 이득변화 자체의 가청도와 대조함에 의해 설정될 수 있다. 이 이득-제어 워드 기술은 편집 분야에만 한정되는 것은 아니다.

프레임 경계에서의 에일리어싱을 억제하기 위한 필터뱅크들

TDAC 변환에 의해 제공되는 것과 같은 에일리어싱 소거의 형태를 사용하는 코딩 시스템에 있어서, 결합 편집은 에일리어싱 아티팩트들이 상기한 이유로 인하여 결합부분의 어느 한 쪽에서 소거되는 것을 방지하게 된다. 이와 같이 소거되지 않은 에일리어싱 아티팩트들은 각 프레임의 시작 및 끝 부분에 있는 오디오 블록들에 교번되는 필터뱅크들을 인가함에 의해 피할 수 있다. 예를 들어, 도 1a의 프레임(21)을 참조하면, 제 1 필터뱅크가 블록 11에 인가되고, 제 2 필터뱅크가 블록 12 내지 16에 인가되며, 제 3 필터뱅크가 블록 17에 적용된다. 이와 같은 필터뱅크들의 특징은 각 프레임으로부터 복구된 오디오가 소거되지 않은 어떤 에일리어싱 아티팩트들도 포함하지 않는다는 것이다.

도 5a를 참조하면, 장치(200)는 오디오 정보 블록들을 수신하며, 경로 203 을 통해 하나의 오디오 블록이 프레임 내의 제 1 (즉, 시작) 블록인지, 마지막 (즉, 종료) 블록인지, 또는 중간 블록인지를 나타내는 제어신호를 발생시키는 버퍼(202)를 포함한다. 경로 203 으로부터 수신되는 제어신호에 응하여, 스위치(204)는 각 프레임 내의 제 1 (즉, 시작) 블록을 제 1 필터뱅크(205)에 연결하고, 각 프레임 내의 모든 중간 블록들을 제 2 필터뱅크(206)에 연결하며, 각 프레임 내의 마지막 (즉, 종료) 블록을 제 3 필터뱅크(207)에 연결한다. 포맷터(208)는 3 개의 필터뱅크 각각으로부터 수신된 필터링된 오디오 정보를 경로 209 를 통과하는 출력신호로 조립한다.

도 5b는 디포맷터(222)가 경로 221 로부터 입력신호를 수신하고, 이로부터 경로 224 를 통해 출력되는 인코딩된 오디오 정보를 얻으며, 경로 223 을 통해 상기 인코딩된 오디오 정보가 프레임 내의 제 1 (즉, 시작) 블록인지, 마지막 (즉, 종료) 블록인지, 또는 중간 블록인지를 나타내는 제어신호를 발생시키도록된 장치(220)를 도시한다. 이 경로 223 을 통해 수신되는 상기 제어신호에 응하여, 스위치(225)는 인코딩된 오디오 정보를 3 개의 필터뱅크들 중에서 어느 하나의 필터뱅크로 연결한다. 이 스위치(225)는 제 1 블록에 대한 인코딩된 오디오 정보를 제 1 합성 필터뱅크(226)로, 중간 블록들에 대한 인코딩된 오디오 정보를 제 2 합성 필터뱅크(227)로, 그리고 마지막 블록에 대한 인코딩된 오디오 정보를 제 3 합성 필터뱅크(228)로 연결한다. 버퍼(229)는 3 개의 필터뱅크들로부터 수신되는 합성 오디오 블록들에 응하여 경로 230 을 통해 출력신호를 발생시킨다.

제 2 필터뱅크

인코더의 한 실시예에 있어서, 상기 제 2 필터뱅크는 상기에서 인용된 Princen에 개시된 O-TDAC 변환에 따라 N-포인트 변형 DCT와 N-포인트 분석 윈도우 기능에 의해 구현된다. 상보적 디코더에 있어서, 상기 제 2 필터뱅크는 O-TDAC 변환에 따라 N-포인트 변형 역 DCT와 N-포인트 합성 윈도우 기능에 의해 구현된다. 이 정 및 역 O-TDAC 변환들은 각각 식 1 및 2에 나타나 있다.

수학식 1

$$X(k) = \sum_{n=0}^{M-1} x(n) \cos \left[\frac{2\pi}{M} \left(k + \frac{1}{2} \right) \left(n + \frac{m+1}{2} \right) \right] \quad (0 \leq k < M)$$

수학식 2

$$x(n) = \frac{1}{M} \sum_{k=0}^{M-1} X(k) \cos \left[\frac{2\pi}{M} \left(k + \frac{1}{2} \right) \left(n + \frac{m+1}{2} \right) \right] \quad (0 \leq n < M)$$

여기서, k = 주파수 지수,

n = 신호 샘플 번호,

M = 샘플 블록 길이,

m = O-TDAC에 대한 위상 항목,

x(n) = 윈도우 처리된 입력신호 샘플 n,

$X(k)$ = 변환 계수 k 이다.

제 2 필터뱅크들은 $M = N$ 의 길이를 가지며, 도 3에 도시된 바와 같이 블록의 중간 포인트에서의 두 영역사이의 하나의 경계와 함께 두 개의 에일리어싱 반사 영역들을 생성한다. 이와 같은 두 개의 영역을 생성하는데 필요한 TDAC 위상 항목은 $m = N/2$ 이다.

본 발명의 바람직한 실시예에 있어서, 분석 및 합성 윈도우 함수들은 다음에 기술된 기술에 따라 유도된다. 이 윈도우 함수들의 형태는 도 6a 내의 곡선(242)으로 도시된다. 설명을 쉽게 하기 위하여, 이 윈도우 함수들을 $W_2(n)$ 이라 칭한다.

제 1 필터뱅크

본 발명의 동일한 실시예에 있어서, 인코더 내의 제 1 필터뱅크들과 상보적 디코더는 상기한 변형 DCT와 변형된 형태의 윈도우 함수 $W_2(n)$ 에 의해 구현된다. 상기 정 및 역 변환들은 식 1 및 2에 각각 나타나 있다. 제 1 필터뱅크들은 $M = 3N/2$ 의 길이를 가지며, 하나의 에일리어싱 반사(reflection) 영역 1 을 생성한다. 에일리어싱 아티팩트들은 블록 내의 신호의 반전된 끝-대-끝 반사물(reflection)이 된다. 실제로, 반사 영역 2 는, "0"의 길이를 가지며, 두 영역 사이의 경계는 블록의 선단부나 우측단부에 있다. 이 하나의 영역을 생성하는데 필요한 TDAC 위상 항목은 $m = 0$ 이다.

제 1 필터뱅크에 대한 분석 및 합성 윈도우 함수들 $W_1(n)$ 은 동일하다. 이 윈도우 함수의 형태는 도 6b에서 곡선(241)으로 도시된다. 이 윈도우 함수는 3 개의 부분으로 이루어진다. 세그먼트 1 및 2로 지정된 제 1 및 제 2 부분들은 도 6a에 도시된 윈도우 함수 $W_2(x)$ 와 동일하다. 세그먼트 3으로 지정된 제 3 부분은 "0"과 같다.

제 1 분석 윈도우 함수 $W_1(n)$ 은 세그먼트 3 내의 신호가 "0"인 것을 보장한다. 결과적으로, 세그먼트 3에서 세그먼트 1로 반사되는 에일리어싱 아티팩트들도 "0"이 된다. 세그먼트 1에서 세그먼트 3으로 반사되는 에일리어싱 아티팩트들은 일반적으로 "0"이 되지 않을 것이다. 그러나 세그먼트 3으로 반사되는 아티팩트들은 제 1 합성 윈도우 함수 $W_1(n)$ 가 합성 오디오 블록에 적용될 때 제거될 것이다. 결과적으로, 에일리어싱 아티팩트들은 오직 세그먼트 2 내에 존재하게 된다.

제 3 필터뱅크

본 발명의 동일한 실시예에 있어서, 인코더 내의 제 3 필터뱅크와 상보적 디코더는 상기한 변형 DCT와 변형된 형태의 윈도우 함수 $W_2(n)$ 에 의해 구현된다. 상기 정 변환 및 역 변환들은 식 1 및 2에 각각 나타나 있다. 제 3 필터뱅크들은 $M = 3N/2$ 의 길이를 가지며, 하나의 에일리어싱 반사 영역 2 를 생성한다. 에일리어싱 아티팩트들은 블록 내의 신호의 끝-대-끝 반사물이 된다. 실제로, 반사 영역 1 은 "0"의 길이를 가지며, 두 영역 사이의 경계는 블록의 후단부나 좌측단부에 있다. 이 단일 영역을 생성하는데 필요한 TDAC 위상 항목은 $m = 3N/2$ 이다.

제 3 필터뱅크에 대한 분석 및 합성 윈도우 함수들 $W_3(n)$ 은 동일하다. 적절한 윈도우 함수의 형태가 도 6c에서 곡선(243)으로 도시되어 있다. 이 윈도우 함수는 3 개의 부분으로 이루어진다. 세그먼트 1로 지정된 제 1 부분은 "0"이고, 세그먼트 2 및 3으로 지정된 제 2 및 제 3 부분들은 도 6a에 도시된 윈도우 함수 $W_2(x)$ 와 동일하다.

제 3 분석 윈도우 함수 $W_3(n)$ 은 세그먼트 1 내의 신호가 "0"인 것을 보장한다. 결과적으로, 세그먼트 1에서 세그먼트 3로 반사되는 에일리어싱 아티팩트들도 "0"이 된다. 세그먼트 3에서 세그먼트 1로 반사되는 에일리어싱 아티팩트들은 일반적으로 "0"이 되지 않을 것이다. 그러나, 세그먼트 1로 반사되는 아티팩트들은 제 3 합성 윈도우 함수 $W_3(n)$ 가 합성 오디오 블록에 인가될 때 제거될 것이다. 결과적으로, 에일리어싱 아티팩트들은 오직 세그먼트 2 내에 존재하게 된다.

도 6d는 어떻게 윈도우 함수들 $W_1(n)$, $W_2(n)$ 및 $W_3(n)$ (241-243)이 서로 중첩되는지를 도시하고 있다. 이득 프로파일(240)은 TDAC에 있어서 해당 분석 및 합성 윈도우 함수들로부터 형성된 중첩된 곱 윈도우 함수들의 시퀀스인 끝-대-끝 윈도우 처리 결과의 네트 효과를 나타낸다. 분석-합성 윈도우 함수들 $W_1(n)$ 에 의해 가중된(weighted) 블록(11)의 세그먼트 2 내의 에일리어싱 아티팩트들은 분석-합성 윈도우 함수들 $W_2(n)$ 에 의해 가중된 블록(12)의 앞쪽 절반 안의 에일리어싱 아티팩트들에 의해 소거된다. 분석-합성 윈도우 함수들 $W_3(n)$ 에 의해 가중된 블록(17)의 세그먼트 2 내의 에일리어싱

아티팩트들은 분석-합성 윈도우 함수들 $W_2(n)$ 에 의해 가중된 블록(16)의 뒷쪽 절반 안의 에일리어싱 아티팩트들에 의해 소거된다. 블록들 12 및 13 또는 블록들 15 및 16 과 같은 중간 블록 쌍들 안의 신호 복구 및 에일리어싱 소거는 종래의 TDAC에 따라 이루어진다.

이 기술을 사용하여, 어느 프레임 경계에서든 결합 편집이 이루어지며, 어떤 에일리어싱 아티팩트들도 비소거 상태로는 남아있지 않게 된다.

윈도우 함수들의 유도

윈도우 함수 $W_2(n)$ 는 다음에서 설명하는 기술을 이용한 기본 윈도우 함수로부터 유도될 수 있다. 적절한 중첩-가산 특성을 갖는 어떠한 윈도우 기능도 기본 윈도우 함수로 사용될 수 있지만, 본 발명의 바람직한 실시예에서 사용된 기본 윈도우 함수들은 Kaiser-Bessel 윈도우 함수이다.

$$W_{KB}(n) = \frac{I_0 \left[\pi \alpha \sqrt{1 - \left(\frac{n}{N/2} \right)^2} \right]}{I_0[\pi \alpha]} \quad (0 \leq n < N) \quad \text{수학식 3}$$

여기서, α = Kaiser-Bessel 윈도우 함수의 알파 인자,

n = 윈도우 샘플 번호,

N = 샘플들 내의 윈도우 길이,

$$I_0[x] = \sum_{k=0}^{\infty} \frac{(x/2)^k}{k!}$$

이다.

상기 유도식은, Kaiser-Bessel 윈도우 함수 $W_{KB}(n)$ 를, 블록 길이 N 에서 중첩 간격 v 를 뺀 값과 같은 길이를 갖는 직사각형 윈도우 함수 $s(k)$ 를 결합하여, 분석-합성 곱 윈도우 함수 $W_p(n)$ 를 발생시킨다. 즉,

$$WP(n) = \frac{\sum_{k=0}^{N-1} s(k) W_{KB}(n-k)}{\sum_{k=0}^v W_{KB}(k)} \quad (0 \leq n < N)$$

상기 식은 다음과 같이 간단하게 될 수 있다.

$$WP(n) = \frac{\sum_{k=0}^{N-v-1} W_{KB}(n-k)}{\sum_{k=0}^v W_{KB}(k)} \quad (0 \leq n < N)$$

여기서, n = 곱-윈도우 샘플 번호,

v = 윈도우 중첩 구간 내의 샘플 수,

N = 곱-윈도우의 소망의 길이,

$W_{KB}(n) = v+1$ 길이의 기본 윈도우 함수,

$WP(n)$ = 길이 N 의 유도된 곱-윈도우,

$$s(k)=\begin{cases} 1 & (0 \leq k < N-v) \\ 0 & (\text{기타}) \end{cases} \quad \text{이다.}$$

O-TDAC 변환에 있어서, 중첩 간격 $v=N/2$ 이고, 분석 윈도우 함수 및 합성 윈도우 함수들은 동일하다. 따라서, 어느 윈도우 함수도 다음으로부터 얻어질 수 있다.

$$W_2(n)=\sqrt{\frac{\sum_{k=0}^{N/2-1} W_{KB}(n-k)}{\sum_{k=0}^{N/2} W_{KB}(k)}} \quad (0 \leq n < N) \quad \text{수학식 4}$$

이와 같은 방식으로 유도되는 분석 및 합성 윈도우 함수들은 여기서 Kaiser-Bessel-Derived(KBD) 윈도우 함수라 칭한다. 상기 곱 윈도우 함수는 KBD 곱 윈도우 함수라 칭한다. 기본 Kaiser-Bessel 윈도우 함수의 알파 인자는 코딩 성능을 최적화시키기 위하여 선택될 수 있다. 많은 적용분야에서 코딩을 위한 최적의 알파 인자는 2 ~ 6의 범위에 있다.

프레임을 통하여 소거되지 않은 에일리어싱 아티팩트들이 없다는 것은 필수적으로 어떤 윈도우 함수도 결합부분에서 사용될 수 있다는 것을 의미한다. 일반적으로, 이와 같은 윈도우 함수들은 중첩 간격을 가로질러 일정한 이득 프로파일을 유지하는 형태를 취하게 된다. 결합부분에서 많은 프레임들을 가로질러 중첩 간격이 존재할 수 있다. 그러나, 많은 적용분야에서는 5 ~ 30 msec 범위에 있는 "결합-중첩 간격"을 사용할 것이다. 다음에 설명하는 이유 때문에 결합부분을 가로지르는 중첩 간격이 증가될 수 있음을 알 수 있다.

결합부분에서의 분광 스플래터를 감소시키기 위한 필터뱅크들

상기한 범위 내에서의 알파 인자는 지각적 코딩을 최적화한다는 의미에서 많은 코딩 적용분야에 대하여 최적화된다. 즉, 일반적으로, 코딩은 좀더 넓은 필터 통과 대역과 대체되는 필터 저지대역에서의 주파수 감쇠를 증가시킴에 의해 최적화된다. 지각적 코딩에 대하여 최적화된 필터에 있어서, 통상의 주파수 응답의 실예가 도 7의 곡선 342로 나타나 있다. 이 곡선은 $\alpha=6$ 인 KBD 윈도우 함수들을 사용하며 256 개의 샘플들과 동일한 프레임 중첩 간격을 갖는 O-TDAC 분석-합성 시스템의 프레임 이득 프로파일의 주파수 응답을 나타낸다. 통과대역과 저지대역 사이의 경계가 명확하게 정해지지 않더라도, 이 실예에서 통과 대역은 약 200 Hz까지의 주파수들을 포함하며, 저지대역은 약 1 kHz까지의 주파수들을 포함한다. 전이영역은 이들 두 개의 대역 사이에 존재하게 된다.

256-샘플 블록들에 적용되는 변환을 사용하는 적용분야에 있어서, 결합 편집은 약 200 Hz 내지 1 kHz 내의 의사 분광 성분들 또는 "분광 스플래터"를 발생시키게 된다. 다른 길이의 블록들을 사용하는 적용분야에서는, 이 주파수 범위가 블록 길이에 의해 구분된 두 개의 상수들의 항목으로 표현될 수 있다. 따라서, 분광 스플래터는 각 블록 길이로 구분되며 약 50,000 내지 256,000 Hz로 표현되는 주파수 범위 내에서 발생하게 된다.

도 7에 도시된 실예에서, 이들 주파수들은 필터 저지대역으로 간주되는 범위의 바깥쪽에 존재하게 된다. 지각적 코딩 성능을 최적화하도록 설계된 필터뱅크들은 결합 편집 부분에서 생성되는 분광 스플래터를 충분히 감쇠시키지 못한다. 이들 아티팩트들은 일반적으로 신호에 의해 차폐되기에는 너무 크기 때문에 흔히 청취 가능하게 된다.

도 7의 곡선 341 과 곡선 343 은 저지대역에서 더욱 적은 감쇠기능을 제공하나 결합부분에서 생성되는 분광 스플래터에 의해 영향을 받는 주파수 범위에서는 더욱 큰 감쇠기능을 제공하는 두 개의 다른 분석-합성 시스템들의 주파수 응답을 도시하고 있다. 지각적 코딩의 성능으로서, 분광 스플래터의 감쇠를 증가시키는 것이 있다. 바람직하게는, 주파수 응답에 의해 256-샘플 블록들을 필터링하는 시스템에서 200 Hz와 600 Hz를 포함하는 주파수 범위 내에서, 또는 각각 블록 길이에 의해 구분되는 약 50,000 및 150,000의 주파수 범위 내에서 분광 에너지의 감쇠가 최적화된다.

가끔 일반적인 코딩과 결합부분에서의 프레임들의 크로스페이딩에 대한 주파수 응답의 요구사항들을 만족시키기 위한 절충이 이루어질 수 있다. 이와 같은 절충이 이루어질 수 없는 적용분야에서는 결합부분이 검출되고, 분석-합성 시스템의 주파수 응답이 변하게 된다. 이와 같은 변화는 분석 필터뱅크가 일반적으로 결합 동작을 고려할 수 없기 때문에 합성 필터링과 관련하여 이루어져야 한다.

도 8은 분석-합성 시스템의 끝-대-끝 주파수 응답을 변경(alter)시킴에 의해 결합부분에서의 분광 스플래터를 감소시키는 데 사용될 수 있는 장치(320)를 도시한다. 이 장치에 있어서, 디포맷터(322)는 경로 321로부터 입력신호를 수신하고, 이로부터 경로 324를 통해 출력되는 인코딩된 오디오 정보를 얻으며, 결합이 프레임의 시작 또는 끝 부분 중 어느 부분에서 발생하는지를 나타내는 제어신호를 경로 323을 통해 발생시킨다. 이와 같은 결합부분의 발생은 입력신호에 수반되거나, 또는 입력신호에 수반되는 다른 정보로부터 추정될 수 있다.

예를 들어, AES-3/EBU 표준에 의하면, 오디오 정보의 순차적인 블록들은 0에서 255까지 증가한 후 다시 0으로 감소하는 블록 번호들을 포함한다. 순차적이지 않은 두 개의 인접한 블록 번호들은 결합부분(splice)을 나타낼 수 있으나, 이와 같은 테스트는 AES/EBU 데이터 스트림을 처리하는 몇몇 장치들이 해당 번호들을 증가시키지 않기 때문에 신뢰성이 없다. 오디오 스트림이 인코딩될 때, 그 인코딩 기능이 순차적인 번호부여나 다른 형태의 예보적인 정보를 제공할 수 있다. 만약, 정보가 예상하는 것과 일치하지 않을 경우, 결합부분의 존재를 나타내기 위하여 신호가 발생할 수 있다.

경로 323 으로부터 수신된 제어신호에 의하여, 스위치(325)는 인코딩된 오디오 정보를 3 개의 합성 필터뱅크들 중에서 어느 하나의 합성 필터뱅크로 연결한다. 이 스위치(325)는 결합부분의 다음에 오는 프레임 내의 제 1 블록에 대한 인코딩된 오디오 정보를 제 1 합성 필터뱅크(326)에 연결하고, 결합부분의 앞쪽에 있는 프레임 내의 마지막 블록에 대한 인코딩된 오디오 정보를 제 3 합성 필터뱅크(328)에 연결하며, 다른 블록들에 대한 인코딩된 오디오 정보를 제 2 합성 필터뱅크(327)에 연결한다. 선택적으로, 이들 다른 블록들에 대한 인코딩된 오디오 정보는 도 5b와 관련하여 상술한 기술에 따라 3 개의 필터뱅크들 중 어느 하나의 필터뱅크에 연결될 수 있다. 버퍼(329)는 3 개의 합성 필터뱅크들로부터 수신된 합성 오디오 블록들에 의하여 출력신호를 경로 330을 통해 발생시킨다.

제 1 및 제 3 필터뱅크들은 어떤 분석 필터뱅크와 관련하여 소망의 주파수 응답을 이루도록 설계된다. 많은 적용분야에서 이 분석 필터뱅크는 제 2 합성 필터뱅크와 함께 일반적인 코딩 성능을 최적화하도록 설계된다. 제 1 및 제 3 합성 필터뱅크들은 필수적으로 소망의 전체 주파수 응답을 제공하는 방법이면 어떠한 방법으로든 구현될 수 있다. 일반적으로, 두 개의 필터뱅크들은 동일한 주파수 응답을 가지나, 서로 시간-반전된(time-reversed) 복사물들인 임펄스 응답을 갖는다. 변환과 윈도우 함수들을 사용하는 필터뱅크들을 구현하는 적용분야에 있어서, 결합부분의 어느 한 쪽에 있는 인접한 프레임들 사이의 중첩 간격을 증가시키는 합성 윈도우 함수들을 사용함에 의해 적절한 필터뱅크들이 구현될 수 있다.

합성 오디오의 변조

합성 오디오의 변조는 여러 가지 방법에 의해 이루어질 수 있다. 한 방법은 결합부분의 어느 한 쪽에 있는 프레임들이 서로 크로스플레이드(crossfade)되도록 합성 필터뱅크로부터 복구된 합성 오디오 신호를 변조하는 것이다. 이것은 도 4c에 도시된 장치(140)와 같은 장치에 의해 이루어진다. 디코더(146)는 소망의 결합-중첩 간격을 가로지르는 결합부분의 앞쪽에 있는 프레임 내의 합성 신호의 진폭을 감소시킨다. 실제로, 결합부분의 앞쪽에 있는 프레임의 이득 프로파일은 상기 간격을 가로질러 특정 단위 레벨로부터 좀더 낮은 레벨로 감소한다. 디코더(146)는 또한, 소망의 결합-중첩 간격을 가로지르는 결합부분의 뒤쪽의 프레임 내의 합성 신호의 진폭을 증가시킨다. 실제로, 결합부분의 뒤쪽의 프레임의 이득 프로파일은 이 간격을 가로질러 상기의 낮은 레벨로부터 특정 단위 레벨로 증가하게 된다. 이득 프로파일들 내에서의 실제적인 변화들이 분석-합성 윈도우 기능의 변조 효과들로 될 경우, 중첩된 프레임들의 전체 이득은 보존될 수 있다.

이득 프로파일들 내에서의 실제적인 변화는 선형적일 수 있다. 도 7의 곡선 343은 약 5 msec 기간의 선형적으로 점점 가늘어지는 프레임 이득 프로파일의 주파수 응답 특성을 나타낸다. 초당 48 k 샘플의 샘플 레이트에서, 이 간격은 약 256 샘플들에 상응한다. 많은 코딩 적용분야에서, 256 샘플들을 갖는 샘플블록들에 변환이 인가되며, 이에 따라 특정 적용분야에서 256 샘플의 경사지거나 선형적으로 점점 가늘어지는 이득 프로파일이 프레임 경계에서의 "종료" 블록을 가로지르거나 이 종료블록과 중첩되는 인접한 블록의 부분을 가로질러 뻗어있게 된다. 이것은 하나의 필터뱅크를 종료블록에 인가하고, 다른 하나의 필터뱅크를 바로 인접한 블록에 인가하며, 또 다른 하나의 필터뱅크를 프레임의 내부에 있는 다른 블록들에 인가하는 것과 동등하다. 도 8에 도시된 장치(320)를 참조하면, 두 개의 부가 합성 필터뱅크들은 상기 "종료" 블록들과 인접하고 중첩되는 블록들을 처리하는데 필요할 것이다.

이 선형적으로 점점 가늘어지는 경사부의 주파수 응답은 다른 주파수 응답의 값을 구하기 위한 기반 응답을 나타낸다. 일반적으로, 이 기반 응답에 대하여 분광 에너지의 감쇠를 최적화하는 필터뱅크들은 결합부분에서 생성되는 분광 스플래터를 감소시키는데 효과적이다.

변형된 합성 윈도우 함수

분석-합성 시스템의 전체 주파수 응답 특성을 변경시키는 다른 방법은 분석-합성 윈도우 기능의 네트 효과로서 소망의 응답을 얻을 수 있도록 합성 윈도우 함수를 변형시키는 것이다. 실제로, 전체 주파수 응답은 결과적인 분석-합성 곱 윈도우 함수에 따라 변화된다.

도 7의 곡선 341 은 곡선 343 으로 표시된 5 msec 의 선형적으로 점점 가늘어지는 이득 프로파일의 주파수 응답보다 더욱 넓게 결합부분에서의 분광 스플래터를 감쇠시키는 주파수 응답을 나타낸다. 곡선 341 의 응답은 256-포인트 변환들과 $\alpha=1$ 인 KBD 윈도우 함수들을 이용하는 O-TDAC 분석-합성 시스템에 의해 얻어진다. 상기한 바와 같이, 곡선 342 는 $\alpha=6$ 인 KBD 윈도우 함수들에 상응한다.

이들 분석-합성 시스템들의 끝-대-끝 주파수 응답은 분석 윈도우 함수와 합성 윈도우 함수의 곱으로부터 형성된 윈도우의 주파수 응답과 동등하다. 이것은 대수적으로 다음과 같이 표현된다.

수학식 5a

$$WP_6(n) = WA_6(n)WS_6(n)$$

수학식 5b

$$WP_1(n) = WA_1(n)WS_1(n)$$

여기서, $WA_6(n) = \alpha=6$ 인 분석 KBD 윈도우 함수,

$WS_6(n) = \alpha=6$ 인 합성 KBD 윈도우 함수,

$WP_6(n) = \alpha=6$ 인 KBD 곱 윈도우 함수,

$WA_1(n) = \alpha=1$ 인 분석 KBD 윈도우 함수,

$WS_1(n) = \alpha=1$ 인 합성 KBD 윈도우 함수,

$WP_1(n) = \alpha=1$ 인 KBD 곱 윈도우 함수이다.

만약, 합성 윈도우 함수가 끝-대-끝 주파수 응답을 어떤 다른 소망의 응답으로 변환시키도록 변형될 경우, 그것은 그것 자체와 분석 윈도우 함수의 곱이 소망의 응답을 갖는 곱 윈도우와 같아지도록 변형되어야 한다. 만약, WP_1 에 상응하는 주파수 응답이 요구되며, 분석 윈도우 함수 WA_6 이 신호 분석에 사용될 경우, 이 관계는 대수적으로 다음과 같이 표현된다.

수학식 5c

$$WP_1(n) = WA_6(n)WX(n)$$

여기서, $WX(n)$ = 주파수 응답을 변환(convert)하는데 필요한 합성 윈도우 함수이다. 이것은 다음과 같이 표현될 수 있다.

수학식 5d

$$WX(n) = \frac{WP_1(n)}{WA_6(n)}$$

만약, 결합-중첩 간격이 프레임 내의 "종료"블록에 중첩되는 인접한 오디오 블록까지 뻗어있을 경우, 윈도우 함수 WX 의 실제 형태는 식 5d 로 표시된 것보다 더욱 복잡해진다. 이것은 다음에 더욱 상세히 설명된다. 어떤 경우이건 식 5d 는 프레임 내에서 어떤 다른 블록하고도 중첩되지 않는 종료블록의 해당 부분에서 윈도우 함수 WX 에서 무엇이 요구되는 지를 정확하게 나타낸다. O-TDAC를 사용하는 시스템들에서, 상기 부분은 블록 길이의 절반, 또는 $0 \leq n < N/2$ 가 된다.

만약, 합성 윈도우 함수 WX가 끝-대-끝 주파수 응답을 좀더 높은 알파 프로파일로부터 좀더 낮은 알파 프로파일로 변환하는데 사용될 경우, 그것은 프레임 경계 주변에서 매우 큰 값들을 가져야한다. 도 9에서, 곡선 351은 $\alpha=1$ 인 KBD 분석 또는 합성 윈도우 함수를, 곡선 352는 $\alpha=1$ 인 KBD 곱 윈도우, 곡선 356은 $\alpha=6$ 인 KBD 분석 또는 합성 윈도우 함수를, 그리고 곡선 359는 식 5d에 따른 합성 윈도우 함수를 실제로 나타낸다. 곡선 356이 프레임 경계에 접근하면서, 곡선 352보다 매우 작아지게 되며, 이에 따라 곡선 359는 매우 커지게 된다. 불행하게도, 윈도우 함수 WX의 끝단에서 큰 증가를 갖는 곡선 359와 같은 형태를 갖는 합성 윈도우 함수는 매우 빈약한 주파수 응답 특성을 가지며, 복구된 신호의 음질을 떨어뜨릴 것이다. 이 문제점을 해결할 수 있는 두 가지 기술이 다음에 설명된다.

샘플들의 버리기(discarding samples)

합성 윈도우 함수를 변형시키기 위한 첫 번째 기술은 분석 윈도우 함수가 가장 작은 값들을 갖는 프레임 경계에서 몇 개의 샘플들을 버림에 의해 윈도우 함수 WX에서의 큰 증가를 피하는 것이다. 이렇게 버려지는 샘플들의 수를 가변시킴에 의해, 프레임 중첩 간격 내에서 샘플들을 전달하는데 필요한 대역폭이 디코더 내에서의 불충분한 주파수 응답 특성에 의한 시스템의 코딩 성능 감소와 교환될 수 있다.

예를 들어, 만약, 프레임 내의 처음 3개의 블록들에 대한 합성 윈도우 함수가 곱 윈도우 함수 WP_1 에 상응하는 소망의 주파수 응답을 얻기 위하여 변형되고, 신호 분석을 위해 사용된 윈도우 함수가 WA_6 일 경우, 요구되는 변형 합성 윈도우 함수들은 다음과 같다.

$$\text{수학식 6a} \\ WX1(n) = \begin{cases} 0 & (0 \leq n < x) \\ \frac{WP_1(n-x)}{WA_6(n)} & (x \leq n < \frac{N}{2}) \\ WP_1(n-x) WA_6(n) & (\frac{N}{2} \leq n < N) \end{cases}$$

$$\text{수학식 6b} \\ WX2(n) = \begin{cases} WP_1(n-x+\frac{N}{2}) WA_6(N) & (0 \leq n < \frac{N}{2}+x) \\ WA_6(n) & (\frac{N}{2}+x \leq n < N) \end{cases}$$

$$\text{수학식 6c} \\ WX3(n) = \begin{cases} WP_1(n-x+N) WA_6(n) & (0 \leq n < x) \\ WA_6(n) & (x \leq n < N) \end{cases}$$

여기서, $WX1(n)$ = 제 1 블록에 대한 변형 합성 윈도우 함수,

$WX2(n)$ = 제 2 블록에 대한 변형 합성 윈도우 함수,

$WX3(n)$ = 제 3 블록에 대한 변형 합성 윈도우 함수,

x = 프레임 경계에서 폐기된 샘플 수이다.

도 10a는, x 의 몇가지 값들에 대하여, $\alpha=6$ 인 KBD 분석 윈도우 함수를 사용하는 256-포인트 O-TDAC 분석-합성 시스템을, 256 샘플들과 동일한 프레임 중첩 간격과 함께 $\alpha=1$ 인 KBD 분석 및 합성 윈도우 함수들을 사용하는 시스템의 주파수 응답과 동등한 주파수 응답을 갖는 분석-합성 시스템으로, 변환하기 위하여 필요한 변형 합성 윈도우 함수의 형태를 도시한다. 곡선들 361, 362, 363 및 364는 각각, $x=8, 16, 24$ 및 32에서의 변형 합성 윈도우 함수들이다.

이들 변형 윈도우 함수들을 사용하는 합성 필터뱅크들의 주파수 응답들이 도 10b에 나타나있다. 곡선들 372, 373 및 374는 각각, $x=8, 16$ 및 24에서의 주파수 응답들이다. 곡선 371은 $\alpha=1$ 인 KBD 윈도우 함수를 사용하는 합성 필터뱅크의 주파수 응답이다. 도면에 도시된 바와 같이, $x=16$ 인 변형 합성 윈도우 함수는, 약 200 Hz를 초과하는 주파수를, $\alpha=1$ 인

KBD 윈도우 함수들을 사용하는 합성 필터뱅크에 의해 얻어지는 것과 같은 주파수로 감쇠시킨다. 즉, $x = 16$ 샘플들을 버리도록 된 합성 필터뱅크는, 분석 필터뱅크와 $\alpha = 6$ 인 분석 윈도우 함수와 관련하여 사용될 때, $\alpha = 1$ 인 분석 및 합성 윈도우 함수들을 사용하는 시스템의 끝-대-끝 주파수 응답과 동등한 끝-대-끝 분석-합성 시스템 주파수 응답을 이룰 수 있으며, 동시에, 약 200 Hz를 초과하는 주파수들을 $\alpha = 1$ 인 합성 윈도우 함수를 사용하는 합성 필터뱅크의 주파수 정도로 감쇠시키는 합성 필터뱅크 주파수 응답을 제공한다.

정상적 코딩에 대하여 좀더 낮은 알파 값들을 갖는 KBD 윈도우 함수들을 사용하는 시스템들은 일반적으로 합성 윈도우 함수에 대하여 좀더 작은 변환을 필요로 하며, 좀더 적은 수효의, 프레임의 단부에서 버려지는 샘플들을 필요로 한다. 프레임의 단부에서 요구되는 변형 합성 윈도우 함수들은 시간 반전을 제외하고는 식 6a 내지 6c에 나타나 있는 윈도우 함수들과 유사하다.

프레임 이득 프로파일의 변조

합성 윈도우 함수를 변형하기 위한 두 번째 기술은 결합부분의 바로 어느 한쪽에서 프레임 이득 프로파일이 이상적인 레벨에서 조금 벗어나도록 함에 의해 윈도우 함수 WX에서의 큰 증가를 피하는 것이다. 이렇게 이득 프로파일에서의 이탈을 가변시킴에 의해, 이 이탈의 가청도가 분광 스플래터의 가청도와 교환될 수 있다.

이 기술은 프레임 경계나 그 근처에서 윈도우 함수가 작은 값들을 갖도록 변형 합성 윈도우 함수를 평활화시킨다. 이와 같은 평활화가 적절하게 이루어질 때, 그 결과로서의 합성 윈도우 함수는 수용 가능한 주파수 응답을 가지게 되며, 이득이 상대적으로 낮은 프레임 경계나 그 근처에서 프레임 이득 프로파일이 이상적인 KBD 곱 윈도우 함수로부터 벗어나게 된다. 분광 스플래터의 감쇠는 이상적인 크로스플레이드(crossfade) 이득 형태에 의해 제공되는 것에 비교하여 볼 때 단지 약간만 질을 떨어뜨릴 것이다.

예를 들어, 만약, 프레임 내의 처음 3 개의 블록들에 대한 합성 윈도우 함수가 소망의 주파수 응답을 얻기 위하여 변형되어야 할 경우, 상기 제 2 및 제 3 블록들에 요구되는 변형 합성 윈도우 함수들 WX는 일반적으로 식 6b 및 6c 에서 $x = 0$ 에 대하여 상기 나타낸 바와 같게 된다. 식 6a 에서 보인 변형 합성 윈도우 함수 WX1은 평활화 윈도우 함수의 길이의 앞쪽 절반을 넘도록 평활화 윈도우 함수로 점-대-점(point-by-point) 방식으로 곱함에 의해 평활화된다. 그 결과로서 나타나는 첫 번째 블록에 대한 변형 합성 윈도우 함수는 다음과 같다.

$$\text{수학식 7}$$

$$WX1(n) = \begin{cases} \frac{WP_1(n)WM(n)}{WA_6(n)} & (0 \leq n < \frac{p}{2}) \\ \frac{WP_1(n)}{WA_6(n)} & (\frac{p}{2} \leq n < \frac{N}{2}) \\ WP_1(n)WA_6(N) & (\frac{N}{2} \leq n < N) \end{cases}$$

여기서, $WM(n)$ = 평활화 윈도우 함수,

$p = N$ 보다 작은 것으로 추정되는, 평활화 윈도우 함수의 길이이다.

프레임의 단부에서 요구되는 변형 합성 윈도우 함수는 시간 반전을 제외하고는 이 윈도우 함수와 동일하다.

평활화 윈도우 함수 WM은 필수적으로 어떠한 윈도우 함수에도 근거할 수 있지만, KBD 평활화 윈도우 함수가 잘 동작하는 것 같다. 이 실예에서, 평활화 윈도우 함수는 $\alpha = 6$ 이고 길이가 128인 KBD 윈도우 함수이다. 도 11a에서, 곡선 381 은 평활화되지 않은 변형 합성 윈도우 함수의 형태를 나타내며, 곡선 382 은 평활화된 변형 합성 윈도우 함수의 형태를 나타낸다.

평활화된 변형 윈도우 함수를 사용하는 분석-합성 시스템의 주파수 응답이 도 11b에 도시되어 있다. 곡선 391 은 평활화된 변형 윈도우 함수를 사용함에 의해 나타나는 주파수 응답을 나타낸다. 곡선 341 은 $\alpha = 1$ 인 KBD 윈도우 함수들을 사용하는 분석-합성 시스템의 주파수 응답을 나타내며, 곡선 393 은, 상기 기술된 앞서 곡선 343 으로 도시된 바와 같은, 기간이 약 5 msec 인 전형적으로 점점 가늘어지는 프레임 크로스플레이드 윈도우 함수들을 사용함에 의해 나타나는 주파수 응답의 피크 엔벨로프를 나타낸다. 도시한 바와 같이, 평활화된 변형 합성 윈도우 함수는 $\alpha = 1$ 인 KBD 윈도우 함수들을 사용하는 분석-합성 시스템에 의해 얻어지는 주파수 응답과 유사한 주파수 응답을 얻는다.

하이브리드 분석-합성 윈도우 함수 변형

상기한 기술들에 있어서, 프레임 이득 프로파일에 대한 모든 변화는 신호 합성 처리과정에서 이루어진다. 선택적으로, 분석 처리는 프레임 경계들에서의 블록들에 대한 하나의 주파수 응답을 갖는 필터뱅크들을 사용할 수 있으며, 또한 내부 블록들에 대한 다른 필터뱅크를 사용할 수 있다. 프레임 경계들에서의 블록들에 대하여 사용되는 필터뱅크들은 결합부분에서의 분광 스플래터를 충분히 감쇠하기 위하여 합성 처리과정에서 요구되는 변형의 정도를 감소시키도록 설계될 수 있다.

데이터 동기

비디오 및 오디오 정보를 함께 처리하는 적용분야에서, 비디오 프레임 길이는 일반적으로 오디오 블록 길이와 같지 않다. 테이블 III과 IV에서 보인 표준에서와 같이, 비디오 프레임들과 오디오 블록들이 동기되는 것은 드물다. 즉, 비디오 프레임 경계에서의 비디오/오디오 정보의 편집은 대개는 오디오 블록 경계에서 이루어지지 않는다. 결과적으로, 블록 코딩 시스템들에 있어서, 나머지 부분적인 블록에 의해 재현되는 오디오 정보는 적절하게 복구되지 못한다. 이 문제를 해결하기 위한 두 가지의 기술이 다음에 설명된다. 첫번째 기술에 대한 논의는 본 발명에 따라 두번째 기술과 관련되는 다양한 특징들에 대해 도입으로서 제공된다.

오디오 샘플 레이트 변환(conversion)

첫 번째 기술은 외부 레이트(rate)로 수신된 입력 오디오 신호를 코딩 시스템의 내부 처리에서 사용되는 다른 레이트로 변환하는 것이다. 이 내부 레이트는 내부 신호에 대하여 충분한 대역폭을 제공하고 적절한 수의 샘플들이 각 비디오 프레임으로 분류되도록 선택된다. 디코딩 또는 재생 시에, 출력신호는 내부 레이트로부터 외부 레이트로 변환되는데, 이 외부 레이트는 원래의 입력 오디오 신호의 외부 레이트와 같은 필요는 없다.

아래의 테이블 V는, 몇몇 비디오 표준에 대해, 비디오 프레임 길이, 그러한 비디오 프레임 길이와 같은 초당 48 k 샘플 레이트에서의 오디오 샘플들의 수, 이들 오디오 샘플들을 타겟 샘플 수로 변환하는데 필요한 내부 레이트, 그리고 샘플들로 이루어지는 내부 오디오 프레임 길이 (이하에서 언급됨) 를 나타낸다. 각 비디오 표준에서 괄호 안에 표시된 숫자는 Hz 단위의 비디오 프레임 레이트이다. 30 Hz를 초과하는 비디오 프레임 레이트에 있어서, 타겟 샘플 수는 896이다. 30 Hz 이하의 비디오 프레임 레이트에 있어서, 타겟 샘플 수는 1792이다. 이들 타겟 길이는 설명을 위한 것이나, 이들 길이가 128 샘플 만큼 서로 중첩되는 256-샘플 블록들의 정수배로 구분될 수 있기 때문에 많은 코딩 적용분야에서 적절한 길이가 된다.

【표 5】

비디오표준	프레임 길이 (msec.)	오디오 길이 (샘플)	내부 샘플 레이트 (kHz)	내부 오디오 프레임 길 이
DTV (60)	16.667	800	53.76	1024
NTSC (59.94)	16.683	800.8	53.706	1024
PAL (50)	20	960	44.8	1024
DTV (30)	33.333	1600	53.76	1920
NTSC (29.97)	33.367	1601.6	53.706	1920
PAL (25)	40	1920	44.8	1920
Film (24)	41.667	2000	43	1920
DTV (23.976)	41.7	2002	42.965	1920

비디오 및 오디오 샘플 레이트
테이블 V

예를 들어, 초당 48 k 샘플로 입력 오디오 신호를, 그리고 초당 25 프레임으로 PAL 비디오 신호를 처리하는 적용분야에서는, 입력 오디오 신호를 초당 44.8 k 샘플 레이트를 갖는 내부 신호로 변환시킬 수 있다. 이 내부 신호 샘플들은 처리를 위해 내부 오디오 프레임들로 배열될 수 있다. 테이블 V에서 보인 실예에 있어서, 내부 오디오 프레임의 길이는 1920 샘플들이다. 이들 실예에서, 내부 오디오 프레임의 길이는 비디오 프레임의 길이와 같지 않다. 이와 같은 불일치는 하나의 프레임 내의 오디오 샘플들이 다른 프레임 내의 오디오 샘플들과 중첩되는 샘플 수에 기인한다.

도 2c에 도시된 실예를 참조하면, 각 프레임은 얼마간의 샘플 수 만큼 서로 중첩된다. 이 샘플 수는 프레임 중첩 간격을 형성한다. 많은 적용분야에서, 프레임 중첩 간격은 각 프레임 안에서의 인접한 오디오 블록들 사이의 중첩 간격과 동일하다. 비디오 프레임 길이와 동일한 샘플 수는 한 프레임의 시작부분에서 다음의 프레임의 시작부분까지의 간격에 걸쳐있는 샘플 수이다. 이것은 내부 오디오 프레임의 길이와 같고, 프레임 중첩 간격 내의 샘플 수보다 적다.

테이블 V에 보인 실예들에서, 비디오 프레임의 길이와 같은 샘플 수는 비디오 프레임 레이트에 따라 1792 또는 896이 된다. 프레임 중첩 간격은 128 샘플들이다. 30 Hz가 넘는 비디오 프레임 레이트에 있어서, 각 내부 오디오 프레임은 1024 (896+ 128) 샘플들을 포함하는데, 이 샘플들은 128 샘플들 만큼 서로 중첩되는, 256 샘플들로 이루어진 7 개의 블록들로 배열될 수 있다. 좀더 낮은 비디오 프레임 레이트에 있어서, 각 내부 오디오 프레임은 1920 (1792+ 128) 샘플들을 포함하며, 이 샘플들은 128 샘플들 만큼 서로 중첩되는, 256 샘플들로 이루어진 14 개의 블록들로 배열될 수 있다.

만약, 프레임 경계들에서 에일리어싱 아티팩트들을 발생시키지 않는 필터뱅크들이 사용될 경우, 프레임 중첩 간격은 바람직하게는 256 샘플까지 증가하는데, 이것은 30 Hz 이상의 비디오 프레임 레이트에 대하여 내부 프레임의 길이를 1152 (896+ 256)까지 증가시키며, 좀더 낮은 비디오 프레임 레이트에 대하여는 2048 (1792+ 256)까지 증가시키게 된다.

오디오 신호를 소망의 비디오 프레임 레이트와 동기시키는데 필요한 내부 샘플 레이트는 그 비디오 프레임 레이트 및 비디오 프레임 길이와 같은 샘플 수의 곱과 같다. 이것은 다음과 동일하다.

$$\text{수학식 8}$$

$$R_I = R_V * (L_A - L_O)$$

여기서, R_I = 내부 샘플 레이트,

R_V = 비디오 프레임 레이트,

L_A = 내부 오디오 프레임 길이,

L_O = 프레임 중첩 간격이다.

도 12a는 컨버터(403)가 경로 402로부터 외부 샘플 레이트를 갖는 입력 오디오 신호를 수신하고, 이 입력 오디오 신호를 내부 샘플 레이트를 갖고 내부 오디오 프레임 길이를 갖는 내부 오디오 프레임들로 배열되는 내부 신호로 변환하는 장치(400)의 기능 블록도이다. 내부 신호는 인코더(404)에 전달된다. 상기 내부 신호에 응하여, 인코더(404)는 경로 405를 통해 인코딩된 오디오 프레임들의 형태로 배열된, 인코딩된 신호를 발생시킨다. 포맷터(406)는 경로 401로부터 프레임들의 형태로 배열된 비디오 정보를 수신하고, 각 비디오 프레임으로 인코딩된 오디오 프레임을 구성하여 경로 407을 통해 출력 신호를 발생시킨다.

도 12b는 디포맷터(412)가 경로 411로부터 비디오 정보와 인코딩된 오디오 정보로 구성된 프레임들의 형태로 배열된 인코딩된 입력신호를 수신하도록 된 장치(410)의 기능 블록도이다. 디포맷터(412)는 인코딩된 입력신호로부터 경로 413을 통과하는 비디오 정보를 얻으며, 또한 인코딩된 입력신호로부터 경로 414를 통과하는 인코딩된 오디오 프레임들의 형태로 배열된, 인코딩된 오디오 정보를 얻는다. 디코더(415)는 인코딩된 오디오 정보를 디코딩하여 내부 샘플 레이트를 가지며 내부 오디오 프레임 길이를 갖는 내부 오디오 프레임들의 형태로 배열되는 내부 신호를 발생시킨다. 내부 신호는 컨버터(416)에 전달된다. 컨버터(416)는 내부 신호를 외부 샘플 레이트를 갖는 출력신호로 변환한다.

삭제

근본적으로, 샘플 레이트 변환을 위한 어떠한 기술도 사용될 수 있다. 이 샘플 레이트 변환을 위한 여러 가지 고려 사항 및 구현 방법이 Adams와 Kwan의 "비동기 샘플 레이트 변환을 위한 이론 및 VLSI 구조," J. of Audio Engr. Soc., July 1993, vol. 41, no. 7/8, pp. 539-555에 개시되어 있다.

다이나믹 오디오 프레임 정렬

만약, 샘플 레이트 변환이 사용되지 않을 경우, 오디오 프레임 레이트는 비디오 프레임 레이트와 함께 가변되어야 한다. 내부 오디오 프레임의 길이는 변환을 사용하는 분할밴드 코딩과 같은 블록 처리가 용이하도록 적절한 길이, 즉, 이들 둘의 상당히 큰 파워의 정수배로 설정되어야 한다. 이후에, 프레임 중첩 간격은, 내부 오디오 프레임의 길이 및 비디오 프레임에 정확하게 걸쳐있는 샘플 개수 사이의 차이와 동일하게 설정된다. 이것은 다음과 같이 주어진다.

$$\text{수학식 9}$$

$$L_O = L_A - L_V$$

여기서, L_V = 오디오 샘플 수로 표현된 비디오 프레임의 길이이다.

불행하게도, 테이블 V에서 보는 바와 같이, 이 기술은 NTSC 비디오 프레임 레이트가 오디오 샘플 레이트의 정수배가 아니기 때문에 NTSC 비디오를 처리하는 적용분야에 대하여 더욱 복잡해진다. 결과적으로, NTSC 프레임 길이는 오디오 샘플 수의 정수배와 같지 않다. 테이블 IV에서 보는 바와 같이, 5 개의 NTSC 비디오 프레임들은 초당 48 k 샘플 레이트에서 8008 개의 오디오 샘플들과 동기되는 것이 필요하다. 이 5 개 프레임의 그룹은 이하에서 슈퍼프레임으로 언급된다.

슈퍼프레임 내의 각 비디오 프레임과 상응하는 오디오 샘플 수는 일정하지 않고 가변된다. 많은 배열이 가능하지만, 29.97 Hz NTSC 비디오의 바람직한 배열은 각각 1602, 1601, 1602, 1601, 1602 샘플들에 상응하는 5 개 프레임들의 시퀀스이다. 59.94 Hz의 NTSC 비디오에 있어서, 801-샘플 블록 쌍이 각 1602 블록과 대체되고, 801/800-샘플 블록 쌍이 각 1601 블록과 대체되도록 된, 유사한 시퀀스가 사용될 수 있다. 다음의 설명에서는 29.97 Hz의 비디오 프레임들을 처리하는 적용분야에 대한 해결책이 주어진다. 이 개념들은 다른 비디오 프레임 레이트에도 적용될 수 있다.

식 9 에서와 같이, 디코더는 중첩 간격의 길이를 정확하게 정할 수 있도록 비디오 프레임 길이 L_V 를 설정할 수 있어야 한다. 만약, 디코더가 프레임 경계에서의 결합 편집과 직면하게 될 경우, 결합부분의 다음에 오는 프레임은 5 개의 가능한 슈퍼프레임 정렬들 중에서 어느 하나를 나타낸다. 상기 디코더는 만약 블록들이 슈퍼프레임 정렬과 같지 않을 경우, 결합부분의 다음에 오는 상기 블록들에 의해 재현되는 오디오가 복구되지 않을 것이다. 이것은 다음에 설명하는 다이내믹 오디오 프레임 정렬 기술에 의해 이루어질 수 있다.

이 기술에 의하면, 도 13a에 도시된 것과 같은 장치(420)에 있어서, 인코더(423)는 경로 422로부터 오디오 정보를 수신하고, 각 슈퍼프레임 내의 프레임들이 각 프레임에 대하여 유일하게 주어지는 라벨(label)에 의해 확인되도록 된 슈퍼프레임들의 형태로 배열된, 인코딩된 오디오 정보를 발생시킨다. 상기의 인코딩된 오디오 정보의 슈퍼프레임들은 경로 425를 통해 출력되고, 프레임 라벨들은 경로 424를 통해 출력된다. 포맷터(426)는 경로 421로부터 비디오 정보 프레임들을 수신하고, 이 비디오 정보, 즉, 인코딩된 오디오 정보 프레임들과 이에 상응하는 라벨들을 경로 427를 통해 출력될 출력신호로 조합하게 된다.

도 13b에 도시된 장치(430)에 있어서, 디포맷터(432)는 경로 431로부터 입력신호를 수신하고, 경로 433을 통해 출력되는 비디오 정보 프레임들을 얻으며, 경로 435를 통해 출력되는 인코딩된 오디오 정보의 슈퍼프레임 시퀀스들을 얻고, 경로 434를 통해 출력되는 인코딩된 오디오 정보의 각 프레임에 대한 라벨들을 얻는다. 프로세서(436)는 상기 지표에 응하여 인코딩된 오디오 정보의 각 프레임에 대한 시작 샘플과 프레임 길이를 설정하고, 디코더(438)는 프로세서(436)에 의해 설정된 시작 샘플과 프레임 길이에 따라 인코딩된 오디오 정보 프레임들을 디코딩함에 의해 경로 439를 통해 출력되는 출력신호를 발생시킨다.

본 발명의 바람직한 실시예에 있어서, 각 슈퍼프레임 내의 프레임들은 0, 1, 2, 3 및 4로 라벨이 붙여진다. 프레임 0에서의 시작 샘플은 비디오 신호의 프레임 경계와 정확하게 동기되는 것으로 추정된다. 슈퍼프레임 내의 각 프레임은 "빠른(early)샘플", "공칭(normal)시작 샘플", 그리고 1603 개의 전체 샘플들 중에서 다른 1601 개의 샘플들을 구비한, 동일한 구조로 발생된다. 본 발명의 바람직한 실시예에 있어서, 샘플들은 0에서 1602까지의 번호가 부여되고, 여기서 "공칭 시작 샘플"은 샘플 번호 1로 됨으로서 비디오 프레임 길이는 1603으로 된다. 상기한 바와 같이, 내부 오디오 프레임의 길이는 프레임 중첩 간격에 의해 더욱 커질 수 있다. 적절한 내부 오디오 프레임의 길이는 1792 샘플들이다. 프레임 이득 프로파일은 1603의 비디오 프레임 길이에 따라 설정된다. 바로 언급한 실예에서, 프레임 중첩 간격은 189 (1792-1603) 샘플들이다.

상기 장치(430)는 소망의 슈퍼프레임 정렬을 가지며, 비디오 정보 내에서 적절한 동기가 이루어지도록 각 오디오 프레임의 정렬을 변경시킨다. 이 정렬은 각 프레임에 대한 시작 샘플과 길이를 선택함에 의해 변경된다. 상기한 바와 같이, 길이는 슈퍼프레임 안의 5-프레임 패턴에 따라 1601과 1602 샘플들 사이에서 가변된다. 이 다이내믹 정렬의 효과는 동반된 비디오 정보와의 동기를 보존하는, 결합부분의 다음에 오는 적절한 정렬을 즉시 이룰 수 있다는 것이다.

상기한 실시예에 있어서, 시작 샘플 번호와 비디오 프레임 길이는 다음의 키에 따른 테이블로부터 얻을 수 있다.

$$\text{수학식 10} \\ K = (F_E - F_D) \text{ modulo } 5$$

여기서, K = 정렬 테이블 접근 키,

F_E = 인코더 프레임 라벨,

F_D = 디코더 프레임 라벨이다.

디코더는 인코딩된 신호로부터 인코더 프레임 라벨을 얻는다. 디코더 프레임 라벨은 디코더에 의해 추정되는 슈퍼프레임 정렬에 따라 0에서 4까지의 반복적인 시퀀스의 형태로 디코더에서 발생된다.

디코더는 테이블 접근 키로서 K를 사용하여 테이블 VI(표6 및 표7)로부터 적절한 프레임 시작 샘플 번호와 비디오 프레임 길이를 얻는다.

[표 6]

접근 키	인코딩 프레임 라벨	시작 샘플	비디오 프레임 길이
0	0	1	1602
0	1	1	1601
0	2	1	1602
0	3	1	1601
0	4	1	1602
2	0	1	1601
2	1	0	1601
2	2	1	1602
2	3	1	1601
2	4	1	1602
4	0	1	1601
4	1	0	1601
4	2	1	1601
4	3	0	1601
4	4	1	1602

[표 7]

접근 키	인코딩 프레임 라벨	시작 샘플	비디오 프레임 길이
1	0	1	1602
1	1	1	1602
1	2	2	1602
1	3	1	1602
1	4	2	1602
3	0	1	1602
3	1	1	1601
3	2	1	1602
3	3	1	1602
3	4	2	1602
다이나믹 오디오 프레임 정렬 테이블 VI			

다이나믹 정렬의 실례가 도 14에 도시되어 있다. 이 실례에서, 슈퍼프레임은 프레임 453 으로 시작되고, 프레임 455 다음에 오는 결합부분에 의해 중단된다. 이 슈퍼프레임 내의 마지막 프레임(456)은 결합부분의 다음에 오게 되며, 프레임 457로 시작되는 새로운 슈퍼프레임이 그 다음에 오게 된다. 각 프레임에서의 오디오 정보의 이상적인 길이는 행 450의 상자 안에 나타나 있다. 라벨 0은 각 슈퍼프레임 내의 제 1 프레임에 해당된다. 이 실례에서, 디코더에 의해 추정되는 디코더 라벨 F_D 는 행 462에 나타나 있다. 이들 두 라벨들 간의 차이는 식 10에 따라 계산되며, 행 463에 나타나 있는 정렬 테이블 접근 키 K를 설정하게 된다. 상기 정렬 테이블로부터 설정되는 시작 및 종료 샘플 수는 행 464에 나타나 있다. 예를 들어, 부호 0-1601은 샘플 0에서 시작하고 샘플 1601에서 끝나는 1602-샘플 프레임을 나타낸다.

프레임 451에서, 디코더는 1602 샘플들 만큼 긴 블록을 처리한다. 이 프레임은 인코더 슈퍼프레임 정렬에 따라 "이상적인" 프레임 길이 보다 한 샘플 만큼 길다. 따라서, 프레임 452는 한 샘플 늦게 시작되며, 이에 의해 이상적인 길이 보다 한 샘플 짧게 된다. 이로 인해, 프레임 453은 슈퍼프레임의 제 1 프레임과 정확하게 동기된 상태에서 샘플 번호 1에서 시작하게 된다. 프레임들 454 및 455의 정렬은 이상적인 정렬과 같다.

결합부분 바로 뒤에서, 프레임 456의 정렬은 이상적 정렬과 같다. 프레임 457은 다음의 슈퍼프레임의 시작과 정확하게 동기된 상태에서 샘플 번호 1에서 시작된다. 프레임 457의 길이는 이상적인 길이 보다 한 샘플 작으나, 프레임 458이 한 샘플 일찍 시작되고, 이상적인 프레임 길이 보다 한 샘플 큰 길이를 갖는다. 프레임 459의 시작은 이상적인 것과 같으나, 이상적인 프레임 보다 한 샘플 짧다. 따라서, 프레임 460은 한 샘플 일찍 시작되고 이상적인 프레임보다 한 샘플 길다.

이 실례가 나타내는 바와 같이, 디코더는 결합부분에 의해 생성되는 불연속성과 무관하게 각 슈퍼프레임의 시작에 정확하게 동기시키게 된다.

상기 장치(430)는 식 6a 내지 6c와 관련하여 상술한 바와 유사한 방식으로 적절한 끝-대-끝(end-to-end) 프레임 이득 프로파일을 이루기 위하여 변형 합성 윈도우 함수를 사용한다. 각 프레임의 시작점에서 상기 변형 합성 윈도우 함수는, 프레임 경계에서 "버려지는" 샘플들의 수 x 가 빠른 시작 샘플에 대하여 상대적으로 오프셋되는 프레임 시작 정렬과 동일하도록 된 식 6a에 따라 설정된다. 예를 들어, 샘플 2에서 시작되는 프레임에서, $x = 2$ 이다. 각 프레임의 끝에서의 변형 합성 윈도우 함수 또한 시간-반전 방식을 제외하고는 식 6a에 따라 설정된다.

(57) 청구의 범위

청구항 1.

신호처리 방법에 있어서,

오디오 샘플 레이트로 오디오 정보를 재현하는 입력 샘플들을 포함하는 제 1 입력 신호를 수신하는 단계;

슈퍼프레임들로 그룹화된 입력 프레임 레이트로 정보를 반송하는 입력 프레임들을 포함하는 제 2 입력 신호를 수신하는 단계로서, 각각의 슈퍼프레임은 상기 입력 프레임들의 제 1의 수를 포함하여 상기 입력 프레임 레이트로 나누어진 상기 오디오 샘플 레이트는 정수가 아니지만, 상기 입력 프레임 레이트로 나누어진 상기 오디오 샘플 레이트와 상기 제 1의 수의 곱은 실질적으로 정수와 등가인, 제 2 입력 신호를 수신하는 단계;

상기 제 1 입력 신호에 응답하여 오디오 프레임들의 시퀀스를 발생시키는 단계로서, 각 오디오 프레임은 각각의 입력 프레임에 상응하며 상기 입력 샘플들의 시퀀스에 상응하는 인코딩된 오디오 정보를 포함하고, 상기 시퀀스는 상기 제 1 입력 신호에서 취해지며 상기 입력 샘플들의 시퀀스에서 제 1 샘플인 빠른(early) 시작 샘플, 상기 해당 입력 프레임과 실질적으로 정렬되는 공칭(normal) 시작 샘플, 및 상기 입력 프레임 레이트로 나누어진 상기 오디오 샘플 레이트와 등가인 몫의 정수부분과 등가의 다수의 연속적인 샘플들을 포함하는, 오디오 프레임들의 시퀀스를 발생시키는 단계; 및

출력 슈퍼프레임들로 그룹화된 출력 프레임들로 배열되는 출력 신호를 발생시키는 단계로서, 각 출력 슈퍼프레임이 상기 제 1의 수와 등가인 다수의 상기 출력 프레임들을 포함하며, 각각의 출력 슈퍼프레임은 각각의 오디오 프레임 및 상기 각각의 오디오 프레임에 대한 라벨을 포함하되, 상기 라벨은 각각의 출력 슈퍼프레임의 각 오디오 프레임에 대해 유일한, 출력 신호를 발생시키는 단계;

를 포함하는 것을 특징으로 하는 신호처리 방법.

청구항 2.

제 1 항에 있어서, 상기 오디오 샘플 레이트는 48kHz이며, 상기 입력 프레임 레이트는 29.97Hz와 실질적으로 등가이고, 상기 제 1의 수는 5이며, 상기 연속적인 샘플들의 수는 1601인 것을 특징으로 하는 신호처리 방법.

청구항 3.

제 1 항에 있어서, 상기 오디오 샘플 레이트는 48kHz이며, 상기 입력 프레임 레이트는 59.94Hz와 실질적으로 등가이고, 상기 제 1의 수는 5이며, 상기 연속적인 샘플들의 수는 800인 것을 특징으로 하는 신호처리 방법.

청구항 4.

제 1 항에 있어서, 상기 오디오 프레임들의 발생은 상기 오디오 정보에 필터뱅크를 적용시키는 것을 포함하는 것을 특징으로 하는 신호처리 방법.

청구항 5.

신호 처리 방법에 있어서,

완전한 입력 슈퍼프레임들과 부분 입력 슈퍼프레임들로 그룹화된 입력 프레임들로 배열되는 입력 신호를 수신하는 단계로서, 각 완전한 입력 슈퍼프레임은 상기 입력 프레임의 제 1의 수를 가지되 상기 제 1의 수는 1 보다 크고, 각 부분 입력 슈퍼프레임은 상기 입력 프레임의 제 4의 수를 가지되 상기 제 4의 수는 상기 제 1의 수보다 작고, 상기 입력 프레임은 입력 프레임 레이트로 인코딩된 오디오 정보를 재현하는 오디오 프레임과 상기 오디오 프레임과 관련한 라벨을 포함하되, 상기 라벨이 각각의 완전한 또는 부분 입력 슈퍼프레임의 각 오디오 프레임에 대해 유일한, 입력 신호를 수신하는 단계;

상기 오디오 프레임들로부터 샘플들의 시퀀스를 유도하기 위한 단계로서, 각각의 샘플들의 시퀀스는 각각의 오디오 프레임으로부터 유도되고 빠른 시작 샘플, 공칭 시작 샘플 및 연속적인 샘플들의 제 2의 수를 포함하며, 상기 샘플들의 시퀀스는 오디오 정보를 오디오 샘플 레이트로 재현하고 상기 제 2의 수는 몫의 정수 부분과 등가이며, 상기 몫은 상기 입력 프레임 레이트로 나누어진 상기 오디오 샘플 레이트와 등가인, 샘플들의 시퀀스를 유도하는 단계;

각 샘플들의 시퀀스로부터 샘플들의 각각의 이어지는 연속물(subsequence)을 얻는 단계로서, 각각의 샘플들의 시퀀스가 유도되는 오디오 프레임과 관련된 라벨에 응답하여 상응하는 이어지는 연속물이 상기 각각의 시퀀스의 샘플들의 수 이하의 샘플들의 제 3의 수를 포함하며 상기 빠른 시작 샘플, 상기 공칭 시작 샘플 또는 상기 공칭 시작 샘플들에 이어지는 샘플에서 시작하며, 상기 제 3의 수는 제 2의 수 또는 상기 제 2의 수에 하나(1)를 더한 것과 등가인, 연속물을 얻는 단계; 및

각각의 이어지는 연속물의 시작 및 바로 전의 연속물의 시작은 상기 바로 전의 연속물의 샘플들의 상기 제3의 수로 분리되는 연속물들의 배치로부터 출력 신호를 발생시키는 단계;

를 포함하는 것을 특징으로 하는 신호처리 방법.

청구항 6.

제 5 항에 있어서, 상기 오디오 샘플 레이트는 48kHz이며, 상기 입력 프레임 레이트는 29.97Hz와 실질적으로 등가이고, 상기 제 1의 수는 5이며, 상기 제 2의 수는 1601인 것을 특징으로 하는 신호처리 방법.

청구항 7.

제 5 항에 있어서, 상기 오디오 샘플 레이트는 48kHz이며, 상기 입력 프레임 레이트는 59.94Hz와 실질적으로 등가이고, 상기 제 1의 수는 5이며, 상기 제 2의 수는 800인 것을 특징으로 하는 신호처리 방법.

청구항 8.

제 5 항에 있어서, 상기 각각의 샘플들의 시퀀스를 유도(derivation)하는 단계는 합성 필터뱅크를 상기 각각의 오디오 프레임의 인코딩된 오디오 정보에 적용시키는 것을 포함하는 것을 특징으로 하는 신호처리 방법.

청구항 9.

신호처리 장치에 있어서,

오디오 샘플 레이트로 오디오 정보를 재현하는 입력 샘플들을 포함하는 제 1 입력 신호를 수신하는 수단;

슈퍼프레임들로 그룹화된 입력 프레임 레이트로 정보를 반송하는 입력 프레임들을 포함하는 제 2 입력 신호를 수신하는 수단으로서, 각각의 슈퍼프레임은 상기 입력 프레임들의 제 1의 수를 포함하여 상기 입력 프레임 레이트로 나누어진 상기 오디오 샘플 레이트는 정수가 아니지만, 상기 입력 프레임 레이트로 나누어진 상기 제 1의 수와 상기 오디오 샘플 레이트의 곱은 실질적으로 정수와 등가인, 제 2 입력 신호를 수신하는 수단;

상기 제 1 입력 신호에 응답하여 오디오 프레임들의 시퀀스를 발생시키는 수단으로서, 각 오디오 프레임은 각각의 입력 프레임에 상응하며 상기 입력 샘플들의 시퀀스에 상응하는 인코딩된 오디오 정보를 포함하고, 상기 시퀀스는 상기 제 1 입력 신호에서 취해지며 입력 샘플들의 상기 시퀀스에서 제 1 샘플인 빠른(early) 시작 샘플, 상기 해당 입력 프레임과 실질적으로 정렬되는 공칭(normal) 시작 샘플, 및 상기 입력 프레임 레이트로 나누어진 상기 오디오 샘플 레이트와 등가인 몫의 정수부분과 등가의 다수의 연속적인 샘플들을 포함하는, 오디오 프레임들의 시퀀스를 발생시키는 수단; 및

출력 슈퍼프레임들로 그룹화된 출력 프레임으로 배열되는 출력 신호를 발생시키는 수단으로서, 각 출력 슈퍼프레임은 상기 제 1의 수에 등가인 다수의 출력 프레임들을 포함하며, 각각의 출력 프레임은 각각의 오디오 프레임 및 상기 각각의 오디오 프레임에 대한 라벨을 포함하되, 상기 라벨은 각각의 출력 슈퍼프레임의 각 오디오 프레임에 대해 유일한, 출력 신호를 발생시키는 수단;

을 포함하는 것을 특징으로 하는 신호처리 장치.

청구항 10.

제 9 항에 있어서, 상기 오디오 샘플 레이트는 48kHz이며, 상기 입력 프레임 레이트는 29.97Hz와 실질적으로 동가이고, 상기 제 1의 수는 5이며, 상기 연속적인 샘플들의 수는 1601인 것을 특징으로 하는 신호처리 장치.

청구항 11.

제 9 항에 있어서, 상기 오디오 샘플 레이트는 48kHz이며, 상기 입력 프레임 레이트는 59.94Hz와 실질적으로 동가이고, 상기 제 1의 수는 5이며, 상기 연속적인 샘플들의 수는 800인 것을 특징으로 하는 신호처리 장치.

청구항 12.

제 9 항에 있어서, 상기 오디오 프레임들의 발생은 상기 오디오 정보에 필터뱅크를 적용하는 수단을 포함하는 것을 특징으로 하는 신호처리 장치.

청구항 13.

신호 처리 장치에 있어서,

완전한 입력 슈퍼프레임들과 부분 입력 슈퍼프레임들로 그룹화된 입력 프레임들로 배열되는 입력 신호들을 수신하는 수단으로서, 각 완전한 입력 슈퍼프레임은 상기 입력 프레임들의 제 1의 수를 가지되 상기 제 1의 수는 1보다 크고, 각 부분 입력 슈퍼프레임은 상기 입력 프레임의 제 4의 수를 가지되 상기 제 4의 수는 상기 제 1의 수보다 작고, 각 입력 프레임은 입력 프레임 레이트로 인코딩된 오디오 정보를 재현하는 오디오 프레임과 상기 오디오 프레임과 관련된 라벨을 포함하되, 상기 라벨이 각각의 완전한 또는 부분 입력 슈퍼프레임에 대해 유일한, 입력 신호를 수신하는 수단;

상기 오디오 프레임들로부터 샘플들의 시퀀스를 유도하기 위한 수단으로서, 각각의 샘플들의 시퀀스는 각각의 오디오 프레임으로부터 유도되고 빠른 시작 샘플, 공칭 시작 샘플 및 연속적인 샘플들의 제 2의 수를 포함하며, 상기 샘플들의 시퀀스는 오디오 정보를 오디오 샘플 레이트로 재현하고 상기 제 2의 수는 몫의 정수 부분과 동가이며, 상기 몫은 상기 입력 프레임 레이트로 나누어진 상기 오디오 샘플 레이트와 동가인, 샘플들의 시퀀스를 유도하는 수단; 및

각 샘플들의 시퀀스로부터 샘플들의 각각의 이어지는 연속물(subsequence)을 얻는 수단으로서, 각각의 샘플들의 시퀀스가 유도되는 오디오 프레임과 관련된 라벨에 응답하여 상응하는 이어지는 연속물이 상기 각각의 시퀀스의 샘플들의 수 이하의 샘플들의 제 3의 수를 포함하며 상기 빠른 시작 샘플, 상기 공칭 시작 샘플 또는 상기 공칭 시작 샘플들에 이어지는 샘플에서 시작하며, 상기 제 3의 수는 제 2의 수 또는 상기 제 2의 수에 하나(1)를 더한 것과 동가인, 연속물을 얻는 수단; 및

각각의 이어지는 연속물의 시작 및 바로 전의 연속물의 시작은 상기 바로 전의 연속물의 샘플들의 상기 제3의 수로 분리되는 연속물들의 배치로부터 출력 신호를 발생시키는 수단;

을 포함하는 것을 특징으로 하는 신호처리 장치.

청구항 14.

제 13 항에 있어서, 상기 오디오 샘플 레이트는 48kHz이며, 상기 입력 프레임 레이트는 29.97Hz와 실질적으로 동가이고, 상기 제 1의 수는 5이며, 상기 제 2의 수는 1601인 것을 특징으로 하는 신호처리 장치.

청구항 15.

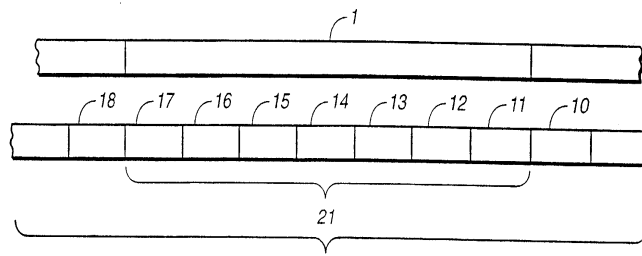
제 13 항에 있어서, 상기 오디오 샘플 레이트는 48kHz이며, 상기 입력 프레임 레이트는 59.94Hz와 실질적으로 등가이고, 상기 제 1의 수는 5이며, 상기 제 2의 수는 800인 것을 특징으로 하는 신호처리 장치.

청구항 16.

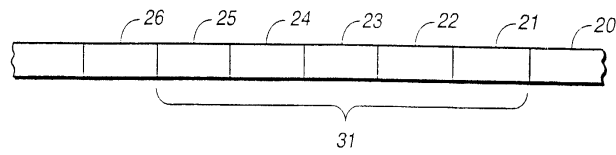
제 13 항에 있어서, 상기 샘플들의 시퀀스를 유도하는 수단은 합성 필터뱅크를 상기 각각의 오디오 프레임의 인코딩된 오디오 정보에 적용하는 수단을 포함하는 것을 특징으로 하는 신호처리 장치.

도면

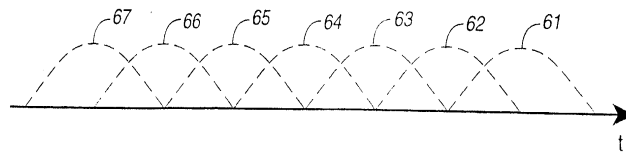
도면1a



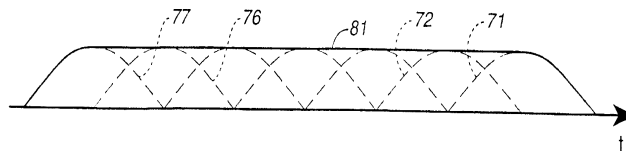
도면1b



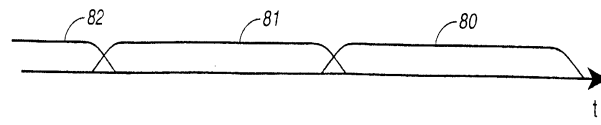
도면2a



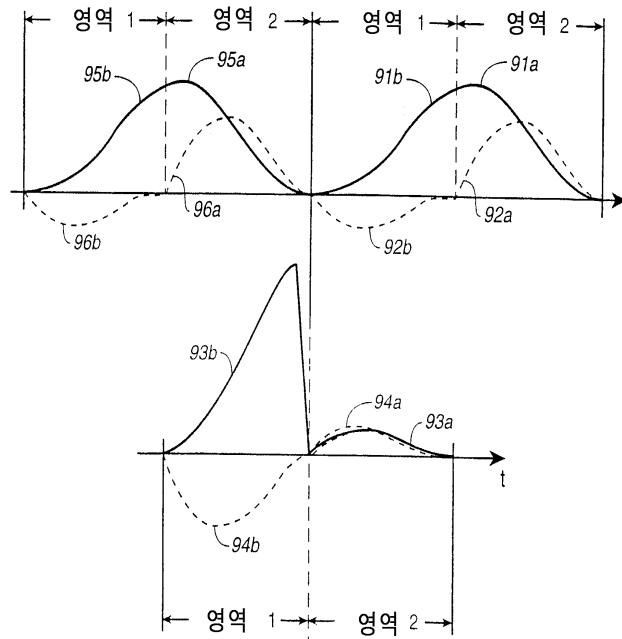
도면2b



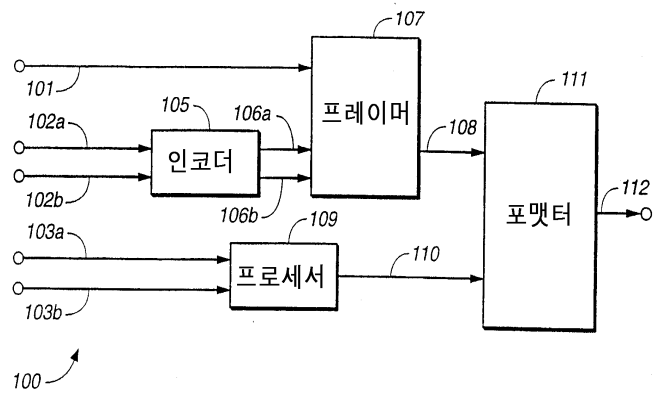
도면2c



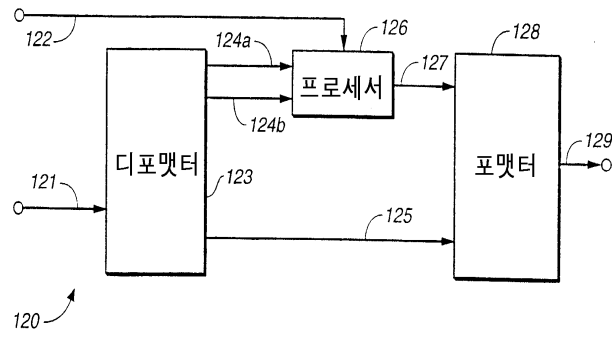
도면3



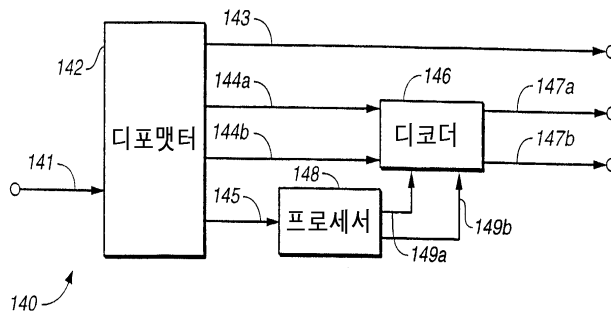
도면4a



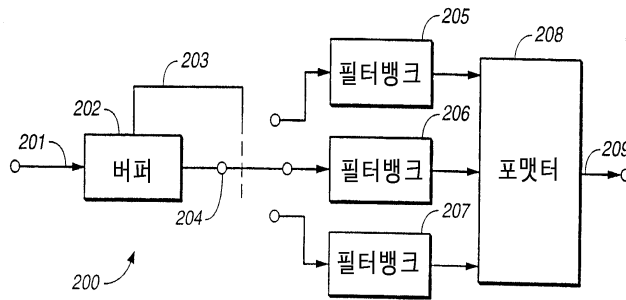
도면4b



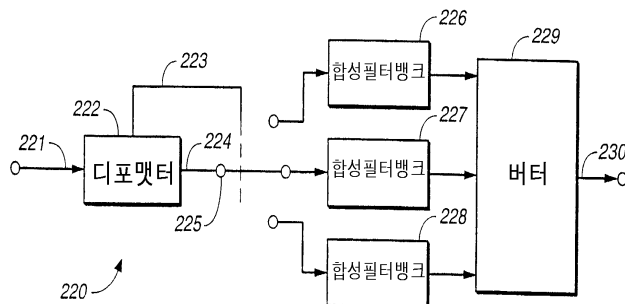
도면4c



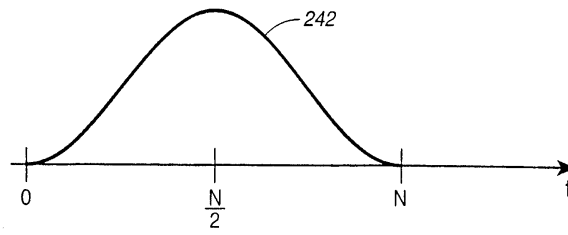
도면5a



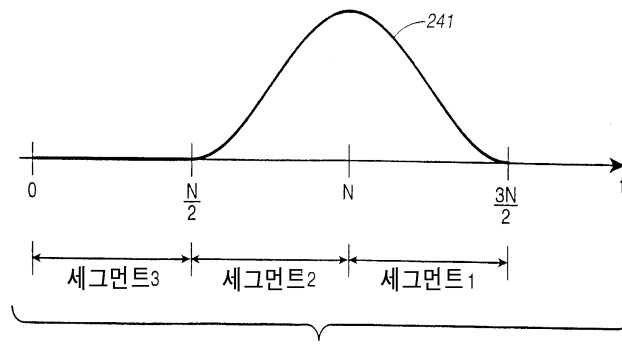
도면5b



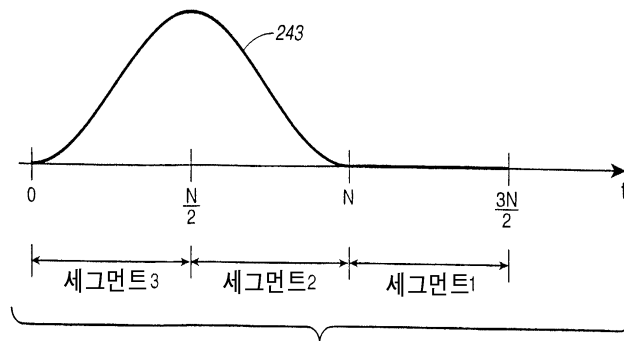
도면6a



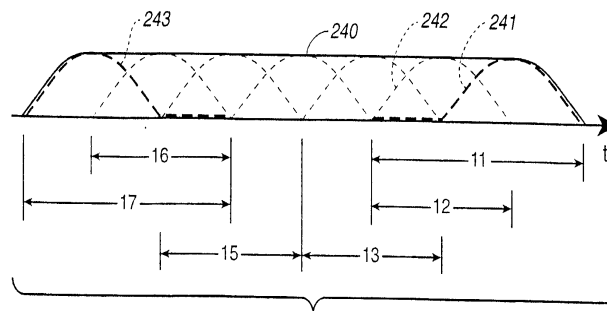
도면6b



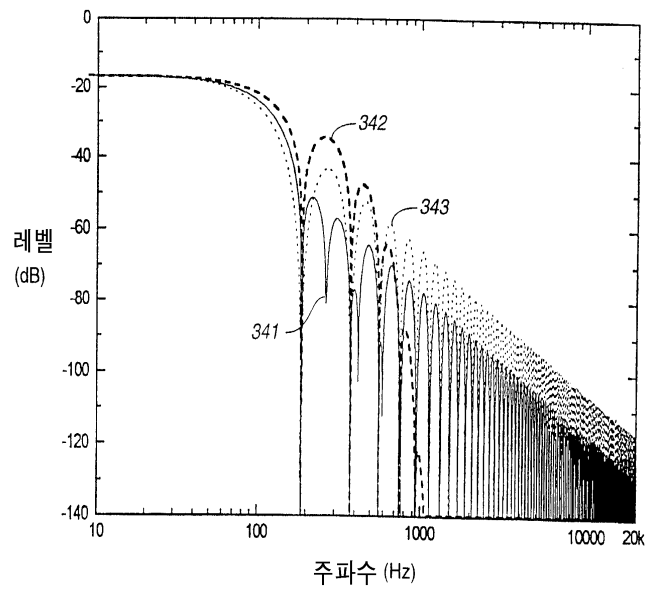
도면6c



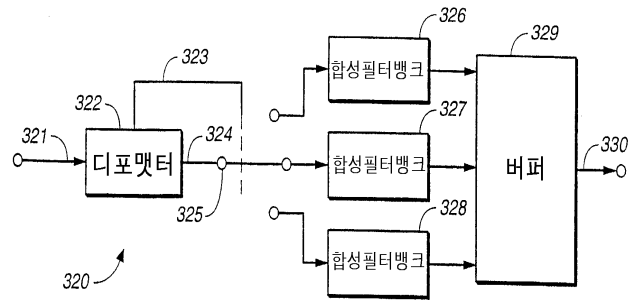
도면6d



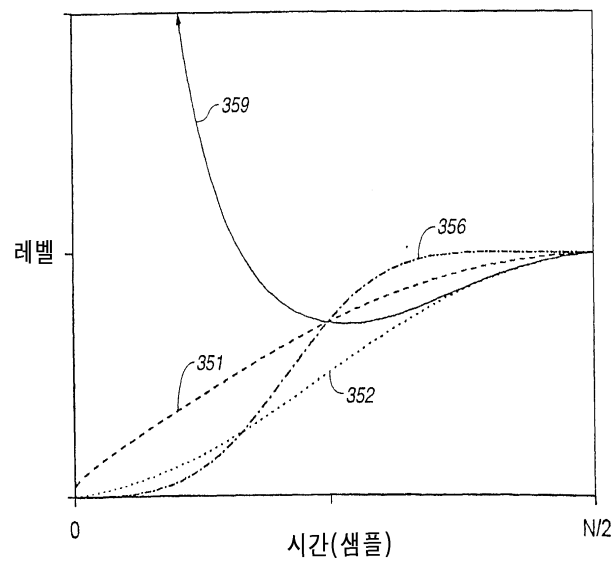
도면7



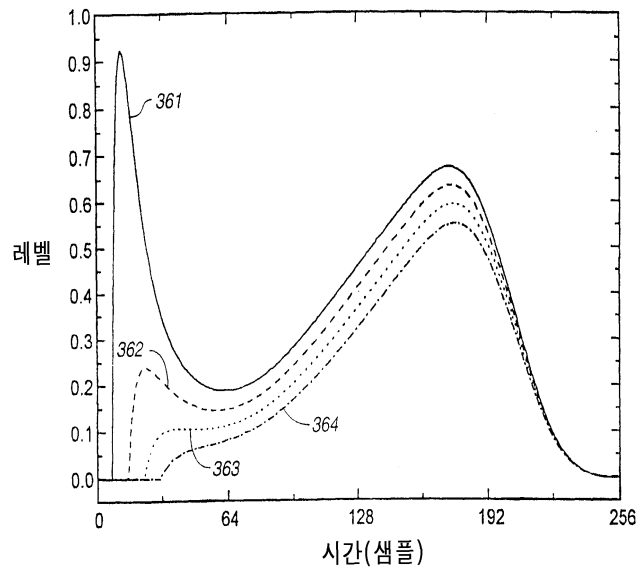
도면8



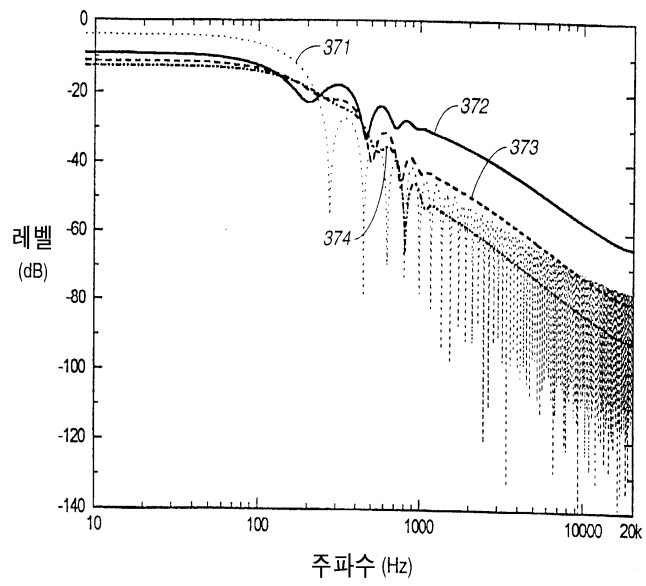
도면9



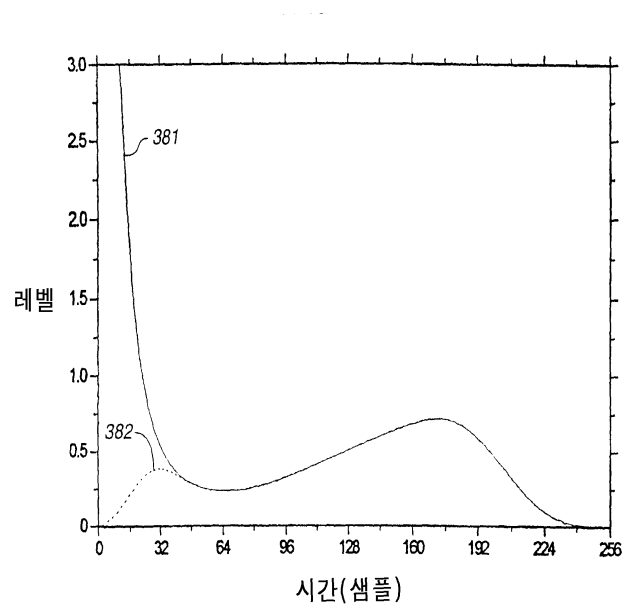
도면10a



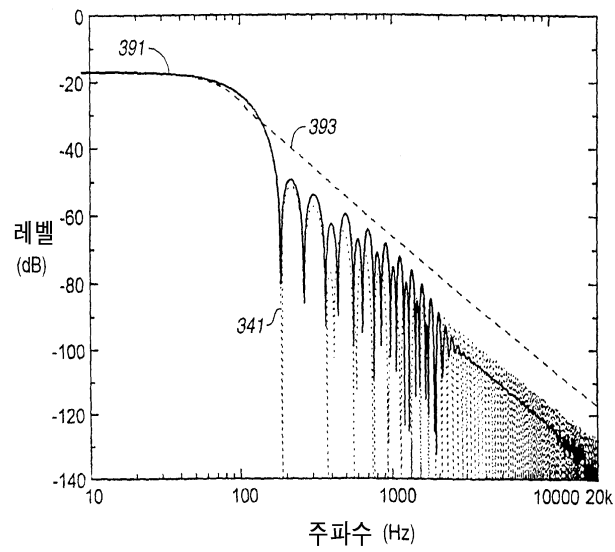
도면10b



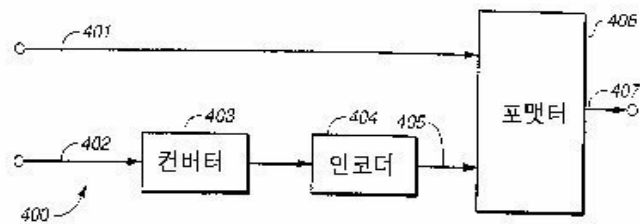
도면11a



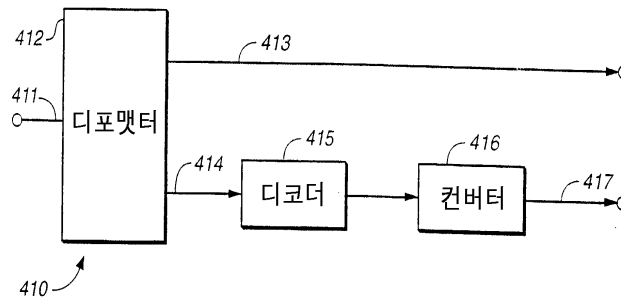
도면11b



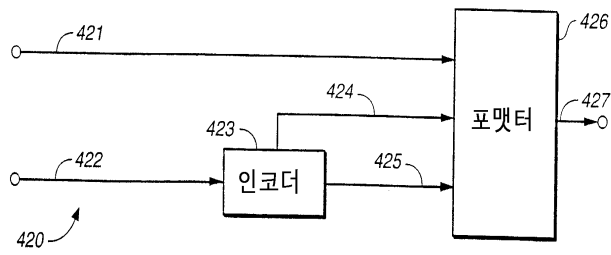
도면12a



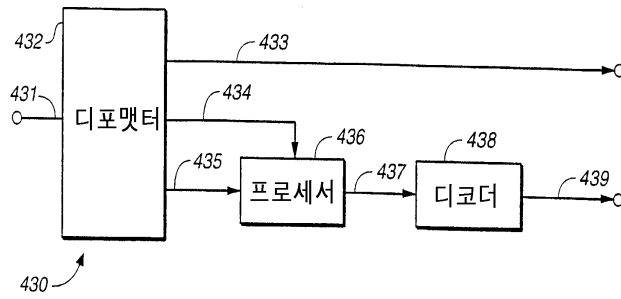
도면12b



도면13a



도면13b



도면14

451 ↓	1601	1602	1602	1601	1602	1602	1602	1601	1602	1601
452 ↓										
453 ↓										
454 ↓										
455 ↓										
456 ↓										
457 ↓										
458 ↓										
459 ↓										
460 ↓										
450 →	3	4	0	1	2	4	0	1	2	3
461 →	0	1	2	3	4	0	1	2	3	4
462 →	3	3	3	3	4	4	4	4	4	4
463 →	1-1602	2-1602	1-1602	1-1601	1-1602	1-1602	1-1601	0-1601	1-1601	0-1601
464 →										