



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0029925

(51)⁷ G10L 19/00

(13) B

(21) 1-2017-03528

(22) 10/03/2016

(86) PCT/US2016/021666 10/03/2016

(87) WO 2016/149015 A1 22/09/2016

(30) 15159067.6 13/03/2015 EP; 62/133,800 16/03/2015 US

(45) 25/11/2021 404

(43) 25/12/2017 357A

(73) Dolby International AB (SE)

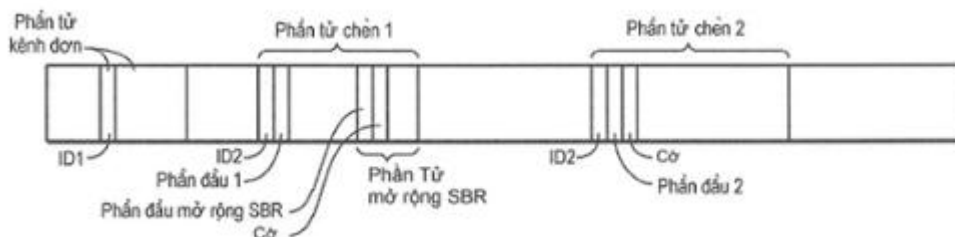
Apollo Building, 3E, Herikerbergweg 1-35, 1101 CN Amsterdam Zuidoost,
Netherlands

(72) VILLEMOES, Lars (SE); PURNHAGEN, Heiko (SE); EKSTRAND, Per (SE).

(74) Công ty TNHH Quốc tế D&N (D&N INTERNATIONAL CO.,LTD.)

(54) BỘ XỬ LÝ ÂM THANH VÀ PHƯƠNG PHÁP GIẢI MÃ DÒNG BIT ÂM THANH MÃ HÓA

(57) Các phương án đề cập đến bộ xử lý âm thanh bao gồm bộ đệm, bộ giải định dạng tải dữ liệu dòng bit, và hệ thống con giải mã. Bộ đệm lưu trữ ít nhất một khối của dòng bit âm thanh mã hóa. Khối này bao gồm phần tử chèn bắt đầu với ký hiệu nhận dạng và tiếp theo là dữ liệu chèn. Dữ liệu chèn bao gồm ít nhất một cờ nhận dạng xem việc xử lý sao chép dải phổ tăng cường (eSBR) có cần được thực hiện trên nội dung âm thanh của khối hay không. Sáng chế cũng đề xuất phương pháp tương ứng để giải mã dòng bit âm thanh mã hóa.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc xử lý tín hiệu âm thanh. Một số phương án đề cập đến việc mã hóa và giải mã các dòng bit âm thanh (ví dụ, các dòng bit có định dạng MPEG-4 AAC) bao gồm siêu dữ liệu để điều khiển việc sao chép dải phổ tăng cường (enhanced spectral band replication - eSBR). Các phương án khác đề cập đến việc giải mã các dòng bit này bằng các bộ giải mã kế thừa không được cấu hình để thực hiện xử lý eSBR và bỏ qua siêu dữ liệu này, hoặc đề cập đến việc giải mã dòng bit âm thanh không bao gồm siêu dữ liệu này bằng cách tạo dữ liệu điều khiển eSBR để đáp ứng lại dòng bit.

Tình trạng kỹ thuật của sáng chế

Dòng bit âm thanh thông thường bao gồm cả dữ liệu âm thanh (ví dụ, dữ liệu âm thanh được mã hóa) biểu thị một hoặc nhiều kênh nội dung âm thanh, và siêu dữ liệu biểu thị ít nhất một đặc trưng của dữ liệu âm thanh hoặc nội dung âm thanh. Một định dạng phổ biến để tạo dòng bit âm thanh mã hóa là định dạng MPEG-4 Mã hóa âm thanh tiên tiến (Advanced Audio Coding - AAC), được mô tả trong tiêu chuẩn MPEG ISO/IEC 14496-3:2009. Trong chuẩn MPEG-4, AAC có nghĩa là “mã hóa âm thanh tiên tiến” và HE-AAC có nghĩa là “mã hóa âm thanh tiên tiến hiệu quả cao”.

Chuẩn MPEG-4 AAC bao gồm một số biên dạng âm thanh, các biên dạng này xác định đối tượng và công cụ mã hóa nào có trong bộ mã hóa hoặc bộ giải mã tuân thủ. Ba trong số các biên dạng âm thanh này là (1) biên dạng AAC, (2) biên dạng HE-AAC, và (3) biên dạng HE-AAC v2. Biên dạng AAC bao gồm loại đối tượng AAC mức độ phức tạp thấp (hay “AAC-LC”). Đối tượng AAC-LC là bản sao của biên dạng MPEG-2 AAC có mức độ phức tạp thấp, với một số điều chỉnh, và không bao gồm loại đối tượng sao chép dải phổ (spectral band replication - “SBR”) và loại đối tượng âm thanh lập thể tham số (parametric stereo - “PS”). Biên dạng HE-AAC là tập lớn của biên dạng AAC và ngoài ra bao gồm loại đối tượng SBR. Biên dạng HE-AAC v2 là tập lớn của biên dạng HE-AAC và ngoài ra bao gồm loại đối tượng PS.

Loại đối tượng SBR chứa công cụ sao chép dải phổ, là một công cụ mã hóa quan trọng giúp cải thiện đáng kể hiệu quả nén của các codec âm thanh cảm nhận. SBR tái cấu trúc lại các thành phần tần số cao của tín hiệu âm thanh ở phía bộ phận thu (ví dụ, trong bộ giải mã). Theo đó, bộ mã hóa chỉ cần mã hóa và truyền các thành phần tần số thấp, cho phép tạo chất lượng âm thanh cao hơn nhiều với tốc độ dữ liệu thấp. SBR có nền tảng dựa trên sự sao chép các chuỗi sóng hài, đã bị cắt trước nhằm làm giảm tốc độ dữ liệu, từ tín hiệu có giới hạn băng thông và dữ liệu điều khiển có sẵn từ bộ mã hóa. Tỷ lệ giữa các thành phần âm và thành phần giống nhiễu âm được duy trì bằng cách lọc ngược thích nghi cũng như tùy ý bổ sung nhiễu âm và các hình sin. Trong chuẩn MPEG-4 AAC, công cụ SBR thực hiện chỉnh sửa phổ (spectral patching), trong đó số dải con của bộ lọc cầu phương (Quadrature Mirror Filter, QMF) liền kề được sao chép từ phần dải tần thấp của tín hiệu âm thanh được truyền đến phần dải tần cao của tín hiệu âm thanh được tạo trong bộ giải mã.

Việc chỉnh sửa phổ có thể không phải là tốt nhất đối với một số loại âm thanh, như nội dung âm nhạc có tần số cắt tương đối thấp. Vì thế, các kỹ thuật giúp cải thiện việc sao chép dải phổ là cần thiết.

Bản chất kỹ thuật của sáng chế

Lớp phương án thứ nhất đề cập đến các bộ xử lý âm thanh bao gồm bộ nhớ, bộ giải định dạng tải dữ liệu dòng bit, và hệ thống con giải mã. Bộ nhớ được cấu hình để lưu trữ ít nhất một khối của dòng bit âm thanh mã hóa (ví dụ, dòng bit MPEG-4 AAC). Bộ giải định dạng tải dữ liệu dòng bit được cấu hình để tách kênh khối âm thanh được mã hóa. Hệ thống con giải mã được cấu hình để giải mã nội dung âm thanh của khối âm thanh được mã hóa. Khối âm thanh được mã hóa bao gồm phần tử chèn với một ký hiệu nhận dạng biểu thị chỗ bắt đầu của phần tử chèn, và dữ liệu chèn sau ký hiệu nhận dạng. Dữ liệu chèn bao gồm ít nhất một cờ nhận dạng xem việc xử lý sao chép dải phổ tăng cường (eSBR) có cần được thực hiện trên nội dung âm thanh của khối âm thanh được mã hóa hay không.

Lớp phương án thứ hai đề cập đến các phương pháp giải mã dòng bit âm thanh mã hóa. Phương pháp bao gồm bước nhận ít nhất một khối của dòng bit âm thanh mã hóa, bước tách kênh ít nhất một số phần của ít nhất một khối của dòng bit âm thanh mã hóa, và bước giải mã ít nhất một số phần của ít nhất một khối của dòng bit âm thanh mã hóa. Ít

nhất một khối của dòng bit âm thanh mã hóa bao gồm phần tử chèn với một ký hiệu nhận dạng biểu thị chỗ bắt đầu của phần tử chèn và dữ liệu chèn sau ký hiệu nhận dạng. Dữ liệu chèn bao gồm ít nhất một cờ nhận dạng xem việc xử lý sao chép dải phổ tăng cường (eSBR) có cần được thực hiện trên nội dung âm thanh của ít nhất một khối của dòng bit âm thanh mã hóa hay không.

Các lớp phương án khác đề cập đến việc mã hóa và chuyển mã các dòng bit âm thanh chứa siêu dữ liệu nhận dạng xem việc xử lý sao chép dải phổ tăng cường (eSBR) có cần được thực hiện hay không.

Mô tả vắn tắt các hình vẽ

HÌNH 1 là biểu đồ khối của một phương án về hệ thống có thể được cấu hình để thực hiện một phương án của sáng chế.

HÌNH 2 là biểu đồ khối của bộ mã hóa, là một phương án về bộ xử lý âm thanh theo sáng chế.

HÌNH 3 là biểu đồ khối của hệ thống bao gồm bộ giải mã, là một phương án về bộ xử lý âm thanh theo sáng chế, và còn có thể có bộ hậu xử lý được ghép nối vào đó.

HÌNH 4 là biểu đồ khối của bộ giải mã, là một phương án về bộ xử lý âm thanh theo sáng chế.

HÌNH 5 là biểu đồ khối của bộ giải mã, là một phương án khác về bộ xử lý âm thanh theo sáng chế.

HÌNH 6 là biểu đồ khối của một phương án khác về bộ xử lý âm thanh theo sáng chế.

HÌNH 7 là sơ đồ của một khối của dòng bit MPEG-4 AAC, bao gồm các đoạn được phân chia của chúng.

Mô tả chi tiết sáng chế

Chú thích ký hiệu và thuật ngữ

Trong toàn bộ bản mô tả này, bao gồm trong các điểm yêu cầu bảo hộ, sự diễn đạt việc thực hiện một hoạt động “trên” tín hiệu hoặc dữ liệu (ví dụ, lọc, định tỷ lệ, biến đổi hoặc áp dụng độ lợi cho tín hiệu hoặc dữ liệu) được sử dụng theo nghĩa rộng để diễn đạt việc thực hiện hoạt động trực tiếp lên tín hiệu hoặc dữ liệu, hoặc lên

một phiên bản đã được xử lý của tín hiệu hoặc dữ liệu (ví dụ, lên một phiên bản của một tín hiệu đã được lọc sơ bộ hoặc đã được tiền xử lý trước khi thực hiện hoạt động trên đó).

Trong toàn bộ bản mô tả này, bao gồm trong các điểm yêu cầu bảo hộ, cụm từ “bộ xử lý âm thanh” được sử dụng theo nghĩa rộng, để chỉ một hệ thống, dụng cụ hoặc thiết bị được cấu hình để xử lý dữ liệu âm thanh. Các ví dụ về bộ xử lý âm thanh bao gồm, nhưng không giới hạn ở các bộ mã hóa (ví dụ, các bộ chuyển mã), các bộ giải mã, codec, các hệ thống tiền xử lý, các hệ thống hậu xử lý và các hệ thống xử lý dòng bit (thường được gọi là các công cụ xử lý dòng bit). Hầu như tất cả các sản phẩm điện tử tiêu dùng như điện thoại di động, ti vi, máy tính xách tay, máy tính bảng đều có chứa bộ xử lý âm thanh.

Trong toàn bộ bản mô tả này, bao gồm trong các điểm yêu cầu bảo hộ, thuật ngữ “ghép nối” hoặc “được ghép nối” được sử dụng theo nghĩa rộng để chỉ sự ghép nối trực tiếp hoặc gián tiếp. Theo đó, nếu thiết bị thứ nhất ghép nối với thiết bị thứ hai, kết nối này có thể là kết nối trực tiếp, hoặc là kết nối gián tiếp qua các thiết bị và kết nối khác. Ngoài ra, các thành phần được tích hợp vào hoặc tích hợp với các thành phần khác cũng được ghép nối với nhau.

Mô tả chi tiết các phương án của sáng chế

Chuẩn MPEG-4 AAC dự kiến rằng dòng bit MPEG-4 AAC được mã hóa bao gồm siêu dữ liệu biểu thị từng dạng xử lý SBR sẽ được áp dụng (nếu phải áp dụng một dạng bất kỳ) bởi một bộ giải mã để giải mã nội dung âm thanh của dòng bit, và/hoặc điều khiển việc xử lý SBR này, và/hoặc biểu thị ít nhất một đặc trưng hay tham số của ít nhất một công cụ SBR sẽ được sử dụng để giải mã nội dung âm thanh của dòng bit. Ở đây, chúng ta sử dụng cụm từ “siêu dữ liệu SBR” để biểu thị loại siêu dữ liệu này, mà đã được mô tả hoặc đề cập đến trong chuẩn MPEG-4 AAC.

Mức cao nhất của dòng bit MPEG-4 AAC là một chuỗi các khối dữ liệu (các phần tử “raw_data_block”), mỗi khối là một đoạn dữ liệu (sau đây được gọi là “khối”) chứa dữ liệu âm thanh (thông thường cho một khoảng thời gian cho 1024 hoặc 960 mẫu), thông tin liên quan và/hoặc dữ liệu khác. Ở đây, chúng ta sử dụng thuật ngữ “khối” để chỉ một đoạn dòng bit MPEG-4 AAC bao gồm dữ liệu âm thanh (và siêu dữ liệu tương ứng và có thể còn có dữ liệu liên quan khác), dữ liệu này xác định hoặc biểu thị một (nhưng không nhiều hơn một) phần tử “raw_data_block”.

Mỗi khối của dòng bit MPEG-4 AAC có thể bao gồm một số các phần tử cú pháp (mỗi phần tử cũng được cụ thể hóa trong dòng bit dưới dạng đoạn dữ liệu). Bảy loại phần tử cú pháp này được chỉ rõ trong chuẩn MPEG-4 AAC. Mỗi phần tử cú pháp được xác định bằng một giá trị khác của phần tử dữ liệu “id_syn_ele”. Các ví dụ về các phần tử cú pháp này bao gồm “single_channel_element()”, “channel_pair_element()”, và “fill_element()”. Phần tử kênh đơn (single channel element) là một bộ chứa bao gồm dữ liệu âm thanh của một kênh âm thanh đơn (tín hiệu âm thanh đơn âm). Phần tử cặp kênh (channel pair element) bao gồm dữ liệu âm thanh của hai kênh âm thanh (đó là tín hiệu âm thanh lập thể).

Phần tử chèn (fill element) là một bộ chứa thông tin bao gồm ký hiệu nhận dạng (ví dụ, giá trị của phần tử “id_syn_ele” nêu trên), tiếp theo sau là dữ liệu, được gọi là “dữ liệu chèn”. Các phần tử chèn trước đây được sử dụng để điều chỉnh tốc độ bit tức thời của các dòng bit mà các dòng này sẽ được truyền qua kênh có tốc độ không đổi. Bằng cách thêm một lượng dữ liệu chèn phù hợp vào từng khối, có thể đạt được tốc độ dữ liệu không đổi.

Theo các phương án của sáng chế, dữ liệu chèn có thể bao gồm một hoặc nhiều tải dữ liệu mở rộng giúp mở rộng loại dữ liệu (ví dụ, siêu dữ liệu) có khả năng được truyền trong một dòng bit. Bộ giải mã nhận được các dòng bit có dữ liệu chèn chứa một loại dữ liệu mới có thể được tùy ý sử dụng bởi thiết bị thu nhận dòng bit (ví dụ, bộ giải mã) để mở rộng chức năng của thiết bị. Theo đó, người có hiểu biết trung bình trong lĩnh vực kỹ thuật có thể hiểu rằng các phần tử chèn là loại cấu trúc dữ liệu đặc biệt và khác với các cấu trúc dữ liệu thường được dùng để truyền dữ liệu âm thanh (ví dụ, tải dữ liệu âm thanh chứa dữ liệu kênh).

Trong một số phương án của sáng chế, ký hiệu nhận dạng được sử dụng để nhận dạng phần tử chèn có thể bao gồm số nguyên không dấu ba bit với bit có ý nghĩa nhất truyền trước (“uimsbf”) có giá trị 0x6. Trong một khối, có thể có một số thực thể có cùng loại với phần tử cú pháp (ví dụ, một số phần tử chèn).

Một tiêu chuẩn mã hóa dòng bit âm thanh khác là chuẩn MPEG Mã hóa âm thanh và tiếng nói hợp nhất (Unified Speech and Audio Coding - USAC)) (ISO/IEC 23003-3:2012). Chuẩn MPEG USAC mô tả sự mã hóa và giải mã nội dung âm thanh bằng cách sử dụng việc xử lý sao chép dải phổ (bao gồm xử lý SBR như được mô tả trong chuẩn MPEG-4 AAC, và đồng thời bao gồm các dạng xử lý sao chép dải phổ

tăng cường khác). Việc xử lý này sử dụng các công cụ sao chép dải phổ (trong bản mô tả này, thường được gọi là “công cụ SBR tăng cường” hoặc “công cụ eSBR”) là phiên bản mở rộng và tăng cường của bộ công cụ SBR được mô tả trong chuẩn MPEG-4 AAC. Theo đó, eSBR (như được định nghĩa trong chuẩn USAC) là một phiên bản được cải thiện của SBR (như được định nghĩa trong chuẩn MPEG-4 AAC).

Trong bản mô tả này, chúng ta sử dụng cụm từ “xử lý SBR tăng cường” (hay là “xử lý eSBR”) để chỉ việc xử lý sao chép dải phổ bằng cách sử dụng ít nhất một công cụ eSBR (ví dụ, ít nhất một công cụ eSBR được mô tả hoặc đề cập đến trong chuẩn MPEG USAC) mà không được mô tả hay đề cập đến trong chuẩn MPEG-4 AAC. Các ví dụ về các công cụ eSBR là chuyển vị sóng hài, tiền xử lý bổ sung chỉnh sửa QMF hay “làm phẳng trước” và tạo hình đường bao thời gian mẫu liên dải con hay “inter-TES”.

Dòng bit được tạo theo chuẩn MPEG USAC (còn được gọi là “dòng bit USAC” trong bản mô tả này) bao gồm nội dung âm thanh được mã hóa và thường bao gồm siêu dữ liệu biểu thị từng dạng xử lý sao chép dải phổ sẽ được áp dụng bởi một bộ giải mã để giải mã nội dung âm thanh của dòng bit USAC, và/hoặc siêu dữ liệu điều khiển việc xử lý sao chép dải phổ này và/hoặc biểu thị ít nhất một đặc trưng hay tham số của ít nhất một công cụ SBR và/hoặc eSBR sẽ được sử dụng để giải mã nội dung âm thanh của dòng bit USAC.

Trong bản mô tả này, chúng ta sử dụng cụm từ “siêu dữ liệu SBR tăng cường” (hay là “siêu dữ liệu eSBR”) để chỉ siêu dữ liệu biểu thị từng dạng xử lý sao chép dải phổ sẽ được áp dụng bởi một bộ giải mã để giải mã nội dung âm thanh của dòng bit âm thanh mã hóa (ví dụ dòng bit USAC), và/hoặc siêu dữ liệu điều khiển việc xử lý sao chép dải phổ này, và/hoặc biểu thị ít nhất một đặc trưng hay tham số của ít nhất một công cụ SBR và/hoặc eSBR sẽ được sử dụng để giải mã nội dung âm thanh này, nhưng không được mô tả hoặc đề cập đến trong chuẩn MPEG-4 AAC. Một ví dụ về siêu dữ liệu eSBR là siêu dữ liệu (biểu thị, hoặc để điều khiển việc xử lý sao chép dải phổ) được mô tả hoặc đề cập đến trong chuẩn MPEG USAC nhưng không được mô tả hoặc đề cập đến trong chuẩn MPEG-4 AAC. Theo đó, trong bản mô tả này, siêu dữ liệu eSBR không phải là siêu dữ liệu SBR, và siêu dữ liệu SBR không phải siêu dữ liệu eSBR.

Dòng bit USAC có thể bao gồm cả siêu dữ liệu SBR và siêu dữ liệu eSBR. Cụ thể hơn, dòng bit USAC có thể bao gồm siêu dữ liệu eSBR điều khiển hoạt động xử lý eSBR của bộ giải mã và siêu dữ liệu SBR điều khiển hoạt động xử lý SBR của bộ giải mã. Theo các phương án của sáng chế, siêu dữ liệu eSBR (ví dụ, dữ liệu cấu hình theo eSBR) có trong (theo sáng chế) dòng bit MPEG-4 AAC (ví dụ, trong bộ chứa `sbr_extension()` ở phía cuối tải dữ liệu SBR).

Việc thực hiện xử lý eSBR của bộ giải mã trong quá trình giải mã dòng bit mã hóa bằng bộ công cụ eSBR (bao gồm ít nhất một công cụ eSBR) sẽ tái tạo lại dải tần số cao của tín hiệu âm thanh, trên cơ sở sao chép các chuỗi sóng hài đã bị cắt trong quá trình mã hóa. Thông thường việc xử lý eSBR này sẽ điều chỉnh đường bao phổ của dải tần số cao được tạo, sẽ áp dụng lọc ngược và thêm các thành phần nhiễu âm và hình sin để tạo lại các đặc trưng phổ của tín hiệu âm thanh gốc.

Theo các phương án của sáng chế, siêu dữ liệu eSBR được bao gồm (ví dụ, một số lượng nhỏ các bit điều khiển là siêu dữ liệu eSBR được bao gồm) trong một hoặc nhiều đoạn siêu dữ liệu của dòng bit âm thanh mã hóa (ví dụ, dòng bit MPEG-4 AAC), các đoạn này cũng bao gồm dữ liệu âm thanh được mã hóa trong các đoạn khác (các đoạn dữ liệu âm thanh). Thông thường, ít nhất một đoạn siêu dữ liệu này của từng khối của dòng bit là (hoặc bao gồm) một phần tử chèn (bao gồm một ký hiệu nhận dạng chỉ báo chỗ bắt đầu của phần tử chèn), và siêu dữ liệu eSBR có trong phần tử chèn sau ký hiệu nhận dạng.

HÌNH 1 là biểu đồ khối của chuỗi xử lý âm thanh làm ví dụ (hệ thống xử lý dữ liệu âm thanh), trong đó một hoặc nhiều phần tử của hệ thống có thể được cấu hình theo một phương án của sáng chế. Hệ thống bao gồm các phần tử sau, ghép nối với nhau như được minh họa: bộ mã hóa 1, hệ thống con cung cấp 2, bộ giải mã 3, và bộ hậu xử lý 4. Trong các biến thể của hệ thống được minh họa, một hoặc nhiều phần tử có thể được bỏ qua, hoặc các bộ xử lý dữ liệu âm thanh bổ sung có thể được thêm vào.

Trong một số phương án thực hiện, bộ mã hóa 1 (tùy ý bao gồm bộ tiền xử lý) được cấu hình để thu nhận các mẫu PCM (miền thời gian) bao gồm nội dung âm thanh làm dữ liệu nhập, và để xuất dòng bit âm thanh mã hóa (có định dạng tuân thủ theo chuẩn MPEG-4 AAC) biểu thị nội dung âm thanh. Trong bản mô tả này, dữ liệu của dòng bit biểu thị nội dung âm thanh còn được gọi là “dữ liệu âm thanh” hoặc “dữ liệu

âm thanh được mã hóa”. Nếu bộ mã hóa được cấu hình theo một phương án của sáng chế, dòng bit âm thanh xuất từ bộ mã hóa sẽ bao gồm siêu dữ liệu eSBR (và thường cũng bao gồm siêu dữ liệu khác) và dữ liệu âm thanh.

Một hoặc nhiều dòng bit âm thanh mã hóa xuất từ bộ mã hóa 1 có thể được gửi đến hệ thống con cung cấp âm thanh được mã hóa 2. Hệ thống con 2 được cấu hình để lưu trữ và/hoặc cung cấp từng dòng bit mã hóa xuất từ bộ mã hóa 1. Dòng bit âm thanh mã hóa xuất từ bộ mã hóa 1 có thể được lưu trữ bởi hệ thống con 2 (ví dụ, dưới dạng DVD hoặc đĩa Blu-ray), hoặc được truyền bởi hệ thống con 2 (hệ thống con này có thể tạo ra mạng hoặc liên kết truyền) hoặc vừa được lưu trữ và được truyền bởi hệ thống con 2.

Bộ giải mã 3 được cấu hình để giải mã dòng bit âm thanh MPEG-4 AAC được mã hóa (được tạo bằng bộ mã hóa 1) nhận được từ hệ thống con 2. Trong một số phương án, bộ giải mã 3 được cấu hình để trích xuất siêu dữ liệu eSBR từ từng khối của dòng bit, và để giải mã dòng bit (bao gồm bằng cách thực hiện xử lý eSBR nhờ sử dụng siêu dữ liệu eSBR được trích xuất) để tạo dữ liệu âm thanh được giải mã (ví dụ, các dòng mẫu âm thanh PCM được giải mã). Trong một số phương án, bộ giải mã 3 được cấu hình để trích xuất siêu dữ liệu SBR từ dòng bit (nhưng bỏ qua siêu dữ liệu eSBR có trong dòng bit), và để giải mã dòng bit (bao gồm bằng cách thực hiện xử lý SBR nhờ sử dụng siêu dữ liệu SBR được trích xuất) để tạo dữ liệu âm thanh được giải mã (ví dụ, các dòng mẫu âm thanh PCM được giải mã). Thông thường, bộ giải mã 3 bao gồm một bộ đệm lưu trữ (ví dụ, theo cách thức không chuyên tiếp) các đoạn của dòng bit âm thanh mã hóa nhận được từ hệ thống con 2.

Bộ hậu xử lý 4 trên Hình 1 được cấu hình để nhận dòng dữ liệu âm thanh được giải mã từ bộ giải mã 3 (ví dụ, các mẫu âm thanh PCM đã được giải mã) và để thực hiện hậu xử lý đối với chúng. Bộ hậu xử lý 4 cũng có thể được cấu hình để kết xuất nội dung âm thanh hậu xử lý (hoặc âm thanh được giải mã nhận được từ bộ giải mã 3) để được phát lại bởi một hoặc nhiều loa.

HÌNH 2 là biểu đồ khối của một bộ mã hóa (100), là một phương án về bộ xử lý âm thanh theo sáng chế. Thành phần hoặc phân tử bất kỳ của bộ mã hóa 100 có thể được triển khai dưới dạng một hoặc nhiều quy trình và/hoặc một hoặc nhiều mạch (ví dụ, ASIC (mạch tích hợp chuyên dụng), FPGA (mảng cổng lập trình được dạng trường), hoặc mạch tích hợp khác), trong phần cứng, phần mềm, hoặc trong

dạng kết hợp của phần cứng và phần mềm. Bộ mã hóa 100 bao gồm bộ mã hóa 105, bộ chèn/bộ định dạng trong bước 107, bước tạo siêu dữ liệu 106, bộ nhớ đệm 109, được ghép nối như trong hình. Thông thường, bộ mã hóa 100 cũng bao gồm các phần tử xử lý khác (không có trong hình). Bộ mã hóa 100 được cấu hình để chuyển đổi dòng bit âm thanh đầu vào thành dòng bit MPEG-4 AAC đầu ra được mã hóa.

Bộ tạo siêu dữ liệu 106 được ghép nối và được cấu hình để tạo (và/hoặc chuyển qua bước 107) siêu dữ liệu (bao gồm siêu dữ liệu eSBR và siêu dữ liệu SBR) sẽ được bao gồm trong bước 107 trong dòng bit mã hóa để được xuất từ bộ mã hóa 100.

Bộ mã hóa 105 được ghép nối và được cấu hình để mã hóa (ví dụ bằng cách thực hiện nén) dữ liệu âm thanh đầu vào, và sau đó gửi âm thanh đã được mã hóa đến bước 107 để được chứa trong dòng bit mã hóa để được xuất từ bước 107.

Bước 107 được cấu hình để ghép kênh âm thanh được mã hóa từ bộ mã hóa 105 và siêu dữ liệu (bao gồm siêu dữ liệu eSBR và siêu dữ liệu SBR) từ bộ tạo 106 để tạo dòng bit mã hóa để được xuất từ bước 107, tốt nhất là sao cho dòng bit mã hóa có định dạng như được xác định trong một trong các phương án của sáng chế.

Bộ nhớ đệm 109 được cấu hình để lưu trữ (ví dụ, theo cách thức không chuyển tiếp) ít nhất một khối của dòng bit âm thanh mã hóa xuất từ bước 107, và chuỗi các khối của dòng bit âm thanh mã hóa sau đó được gửi từ bộ nhớ đệm 109 dưới dạng dữ liệu xuất từ bộ mã hóa 100 đến hệ thống cung cấp.

HÌNH 3 là biểu đồ khối của một hệ thống bao gồm bộ giải mã (200), là một phương án về bộ xử lý âm thanh theo sáng chế, và còn có thể có bộ hậu xử lý (300) được ghép nối vào đó. Thành phần hoặc phần tử bất kỳ của bộ giải mã 200 và bộ hậu xử lý 300 có thể được triển khai dưới dạng một hoặc nhiều quy trình và/hoặc một hoặc nhiều mạch (ví dụ, ASIC, FPGA, hoặc mạch tích hợp khác), trong phần cứng, phần mềm, hoặc trong dạng kết hợp của phần cứng và phần mềm. Bộ giải mã 200 bao gồm bộ nhớ đệm 201, bộ giải định dạng (bộ phân tích cú pháp) tải dữ liệu dòng bit 205, hệ thống con giải mã âm thanh 202 (thường được gọi là bước giải mã “lỗi” hoặc hệ thống con giải mã “lỗi”), bước xử lý eSBR 203 và bước tạo bit điều khiển 204, được ghép nối như trong hình. Thông thường, bộ giải mã 200 còn bao gồm các phần tử xử lý khác (không có trong hình).

Bộ nhớ đệm (bộ đệm) 201 lưu trữ (ví dụ, theo cách thức không chuyển tiếp) ít nhất một khối của dòng bit âm thanh MPEG-4 AAC được mã hóa do bộ giải mã 200 nhận được. Trong quá trình hoạt động của bộ giải mã 200, một chuỗi khối của dòng bit được gửi từ bộ đệm 201 đến bộ giải định dạng 205.

Trong các biến thể của phương án trên Hình 3 (hoặc phương án trên Hình 4 sẽ được mô tả), một APU (bộ xử lý âm thanh) không phải là bộ giải mã (ví dụ APU 500 trên HÌNH 6) bao gồm bộ nhớ đệm (ví dụ, bộ nhớ đệm giống như bộ đệm 201) lưu trữ (ví dụ, theo cách thức không chuyển tiếp) ít nhất một khối của dòng bit âm thanh mã hóa (ví dụ, dòng bit âm thanh MPEG-4 AAC) cùng loại với dòng bit được nhận bởi bộ đệm 201 trên Hình 3 hoặc Hình 4 (tức là dòng bit âm thanh mã hóa bao gồm siêu dữ liệu eSBR).

Tham chiếu lại Hình 3, bộ giải định dạng 205 được ghép nối và được cấu hình để tách kênh từng khối của dòng bit để trích xuất siêu dữ liệu SBR (bao gồm dữ liệu đường bao được lượng tử hóa) và siêu dữ liệu eSBR (và thông thường còn có siêu dữ liệu khác) từ đó, để gửi ít nhất siêu dữ liệu eSBR và siêu dữ liệu SBR đến bước xử lý eSBR 203 và đồng thời cũng gửi siêu dữ liệu được trích xuất khác đến hệ thống con giải mã 202 (và cũng có thể đến bộ tạo bit điều khiển 204). Bộ giải định dạng 205 cũng được ghép nối và được cấu hình để trích xuất dữ liệu âm thanh từ từng khối của dòng bit, và để gửi dữ liệu âm thanh được trích xuất đến hệ thống con giải mã (bước giải mã) 202.

Hệ thống trên Hình 3 cũng tùy ý bao gồm bộ hậu xử lý 300. Bộ hậu xử lý 300 bao gồm bộ nhớ đệm (bộ đệm) 301 và các phần tử xử lý khác (không có trong hình) bao gồm ít nhất một phần tử xử lý được ghép nối với bộ đệm 301. Bộ đệm 301 lưu trữ (ví dụ, theo cách thức không chuyển tiếp) ít nhất một khối (hoặc khung) dữ liệu âm thanh được giải mã do bộ hậu xử lý 300 nhận được từ bộ giải mã 200. Các phần tử xử lý của bộ hậu xử lý 300 được ghép nối và được cấu hình để nhận và xử lý chuỗi khối (hoặc khung) âm thanh được mã hóa xuất từ bộ đệm 301 một cách phù hợp, sử dụng siêu dữ liệu xuất từ hệ thống con giải mã 202 (và/hoặc bộ giải định dạng 205) và/hoặc bit điều khiển xuất từ bước 204 của bộ giải mã 200.

Hệ thống con giải mã âm thanh 202 của bộ giải mã 200 được cấu hình để giải mã dữ liệu âm thanh được trích xuất bởi bộ phân tích cú pháp 205 (bước giải mã này có thể được gọi là hoạt động giải mã “lỗi”) để tạo dữ liệu âm thanh được giải mã, và để gửi

dữ liệu âm thanh được giải mã đến bước xử lý eSBR 203. Bước giải mã này được thực hiện trong miền tần số và thường bao gồm lượng tử hóa ngược sau đó là xử lý phổ. Thông thường, bước xử lý cuối cùng trong hệ thống con 202 áp dụng biến đổi miền tần số thành miền thời gian cho dữ liệu âm thanh miền tần số được giải mã, sao cho dữ liệu xuất của hệ thống con là miền thời gian, dữ liệu âm thanh được giải mã. Bước 203 được cấu hình để áp dụng các công cụ SBR và các công cụ eSBR được biểu thị bởi siêu dữ liệu SBR và siêu dữ liệu eSBR (được trích xuất bởi bộ phân tích cú pháp 205) cho dữ liệu âm thanh được giải mã (tức là, để thực hiện xử lý SBR và eSBR trên dữ liệu xuất của hệ thống con giải mã 202 bằng cách sử dụng siêu dữ liệu eSBR và SBR) để tạo dữ liệu âm thanh được giải mã hoàn toàn mà được xuất (ví dụ, đến bộ hậu xử lý 300) từ bộ giải mã 200. Thông thường, bộ giải mã 200 bao gồm bộ nhớ (có thể truy nhập bởi hệ thống con 202 và bước 203) lưu trữ dữ liệu âm thanh được giải định dạng và siêu dữ liệu được xuất từ bộ giải định dạng 205, và bước 203 được cấu hình để truy nhập dữ liệu âm thanh và siêu dữ liệu (bao gồm siêu dữ liệu eSBR và siêu dữ liệu SBR) khi cần trong quá trình xử lý SBR và eSBR. Xử lý SBR và xử lý eSBR trong bước 203 có thể được coi là hậu xử lý trên dữ liệu xuất của hệ thống con giải mã lõi 202. Tùy chọn, bộ giải mã 200 còn bao gồm hệ thống con trộn tăng cuối cùng (có thể áp dụng các công cụ âm lập thể tham số ("PS")) được xác định trong chuẩn MPEG-4 AAC, sử dụng siêu dữ liệu PS được trích xuất bởi bộ giải định dạng 205 và/hoặc bit điều khiển được tạo trong hệ thống con 204) được ghép nối và được cấu hình để thực hiện trộn tăng trên dữ liệu xuất của bước 203 để tạo âm thanh được trộn tăng, giải mã hoàn toàn, được xuất từ bộ giải mã 200. Theo cách khác, bộ hậu xử lý 300 được cấu hình để thực hiện trộn tăng trên dữ liệu xuất của bộ giải mã 200 (ví dụ, sử dụng siêu dữ liệu PS được trích xuất bởi bộ giải định dạng 205 và/hoặc bit điều khiển được tạo trong hệ thống con 204).

Đáp ứng lại việc siêu dữ liệu được trích xuất bởi bộ giải định dạng 205, bộ tạo bit điều khiển 204 có thể tạo dữ liệu điều khiển, và dữ liệu điều khiển có thể được sử dụng trong bộ giải mã 200 (ví dụ, trong hệ thống con trộn tăng cuối cùng) và/hoặc được gửi dưới dạng dữ liệu xuất của bộ giải mã 200 (ví dụ, đến bộ hậu xử lý 300 để sử dụng trong bước hậu xử lý). Đáp ứng lại việc siêu dữ liệu trích xuất từ dòng bit đầu vào (và cũng có thể đáp ứng lại dữ liệu điều khiển), bước 204 có thể tạo (và gửi đến bộ hậu xử lý 300) các bit điều khiển chỉ báo rằng dữ liệu âm thanh được giải mã được xuất từ bước xử lý eSBR 203 cần phải trải qua một bước hậu xử lý cụ thể.

Trong một số phương án thực hiện, bộ giải mã 200 được cấu hình để gửi siêu dữ liệu được trích xuất bởi bộ giải định dạng 205 từ dòng bit đầu vào đến bộ hậu xử lý 300, và bộ hậu xử lý 300 được cấu hình để thực hiện hậu xử lý trên dữ liệu âm thanh được giải mã xuất từ bộ giải mã 200 sử dụng siêu dữ liệu.

HÌNH 4 là biểu đồ khối của một bộ xử lý âm thanh (“APU”) (210), là một phương án khác về bộ xử lý âm thanh theo sáng chế. APU 210 là bộ giải mã kế thừa không được cấu hình để thực hiện xử lý eSBR. Thành phần hoặc phần tử bất kỳ của APU 210 có thể được triển khai dưới dạng một hoặc nhiều quy trình và/hoặc một hoặc nhiều mạch (ví dụ, ASIC, FPGA, hoặc mạch tích hợp khác), trong phần cứng, phần mềm, hoặc trong dạng kết hợp của phần cứng và phần mềm. APU 210 bao gồm bộ nhớ đệm 201, bộ giải định dạng (bộ phân tích cú pháp) tải dữ liệu dòng bit 215, hệ thống con giải mã âm thanh 202 (thường được gọi là bước giải mã “lỗi” hoặc hệ thống con giải mã “lỗi”) và bước xử lý SBR 213, được ghép nối như trong hình. Thông thường, APU 210 còn bao gồm các phần tử xử lý khác (không có trong hình).

Các phần tử 201 và 202 của APU 210 giống hệt với các phần tử được đánh số giống như vậy của bộ giải mã 200 (trên Hình 3) và phần mô tả trên đây về các phần tử này sẽ không được nhắc lại. Trong quá trình hoạt động của APU 210, chuỗi khối của dòng bit âm thanh mã hóa (dòng bit MPEG-4 AAC) nhận được bởi APU 210 được gửi từ bộ đệm 201 đến bộ giải định dạng 215.

Bộ giải định dạng 215 được ghép nối và được cấu hình để tách kênh từng khối của dòng bit để trích xuất siêu dữ liệu SBR (bao gồm dữ liệu đường bao được lượng tử hóa) và thông thường còn cả siêu dữ liệu khác từ đó, nhưng bỏ qua siêu dữ liệu eSBR có thể có trong dòng bit theo phương án bất kỳ của sáng chế. Bộ giải định dạng 215 được cấu hình để gửi ít nhất siêu dữ liệu SBR đến bước xử lý SBR 213. Bộ giải định dạng 215 cũng được ghép nối và được cấu hình để trích xuất dữ liệu âm thanh từ từng khối của dòng bit, và để gửi dữ liệu âm thanh được trích xuất đến hệ thống con giải mã (bước giải mã) 202.

Hệ thống con giải mã âm thanh 202 của bộ giải mã 200 được cấu hình để giải mã dữ liệu âm thanh được trích xuất bởi bộ giải định dạng 215 (bước giải mã này có thể được coi là hoạt động giải mã “lỗi”) để tạo dữ liệu âm thanh được giải mã, và để gửi dữ liệu âm thanh được giải mã đến bước xử lý SBR 213. Bước giải mã được thực hiện trong miền tần số. Thông thường, bước xử lý cuối cùng trong hệ thống con 202 áp dụng biến đổi

miền tần số thành miền thời gian cho dữ liệu âm thanh miền tần số được giải mã, sao cho dữ liệu xuất của hệ thống con là miền thời gian, dữ liệu âm thanh được giải mã. Bước 213 được cấu hình để áp dụng các công cụ SBR (nhưng không phải các công cụ eSBR) được biểu thị bằng siêu dữ liệu SBR (được trích xuất bởi bộ giải định dạng 215) cho dữ liệu âm thanh được giải mã (tức là, để thực hiện xử lý SBR trên dữ liệu xuất của hệ thống con giải mã 202 bằng cách sử dụng siêu dữ liệu SBR) để tạo dữ liệu âm thanh được giải mã hoàn toàn mà được xuất (ví dụ, đến bộ hậu xử lý 300) từ APU 210. Thông thường, APU 210 bao gồm bộ nhớ (có thể truy nhập bởi hệ thống con 202 và bước 213) lưu trữ dữ liệu âm thanh được giải định dạng và siêu dữ liệu được xuất từ bộ giải định dạng 215, và bước 213 được cấu hình để truy nhập dữ liệu âm thanh và siêu dữ liệu (bao gồm siêu dữ liệu SBR) khi cần trong quá trình xử lý SBR. Xử lý SBR trong bước 213 có thể được coi là hậu xử lý trên dữ liệu xuất của hệ thống con giải mã lỗi 202. Tùy chọn, APU 210 cũng có thể bao gồm hệ thống con trộn tăng cuối cùng (có thể áp dụng các công cụ âm lập thể tham số ("PS")) được định nghĩa trong chuẩn MPEG-4 AAC, sử dụng siêu dữ liệu PS được trích xuất bởi bộ giải định dạng 215) được ghép nối và được cấu hình để thực hiện trộn tăng trên dữ liệu xuất của bước 213 để tạo âm thanh được trộn tăng, giải mã hoàn toàn, được xuất từ APU 210. Theo cách khác, bộ hậu xử lý được cấu hình để thực hiện trộn tăng trên dữ liệu xuất của APU 210 (ví dụ, sử dụng siêu dữ liệu PS được trích xuất bởi bộ giải định dạng 215 và/hoặc bit điều khiển được tạo trong APU 210).

Các phương án thực hiện khác nhau của bộ mã hóa 100, bộ giải mã 200 và APU 210 được cấu hình để thực hiện các phương án khác nhau theo phương pháp của sáng chế.

Theo một số phương án, siêu dữ liệu eSBR được bao gồm (ví dụ, một số lượng nhỏ các bit điều khiển là siêu dữ liệu eSBR được bao gồm) trong dòng bit âm thanh mã hóa (ví dụ, dòng bit MPEG-4 AAC), sao cho các bộ giải mã kế thừa (không được cấu hình để phân tích cú pháp siêu dữ liệu eSBR, hoặc để sử dụng công cụ eSBR bất kỳ mà siêu dữ liệu eSBR có liên quan) có thể bỏ qua siêu dữ liệu eSBR nhưng vẫn giải mã dòng bit ở mức cao nhất có thể mà không sử dụng đến siêu dữ liệu eSBR hoặc công cụ eSBR bất kỳ mà siêu dữ liệu eSBR có liên quan, mà thông thường không gặp một hậu quả đáng kể nào trong chất lượng âm thanh được giải mã. Tuy nhiên, các bộ giải mã eSBR được cấu hình để phân tích cú pháp dòng bit để nhận dạng siêu dữ liệu eSBR và để sử dụng ít nhất một công cụ eSBR để đáp ứng siêu dữ liệu eSBR sẽ được hưởng các

lợi ích của việc sử dụng ít nhất một công cụ eSBR này. Vì vậy, các phương án theo sáng chế sẽ cung cấp phương tiện truyền siêu dữ liệu hoặc dữ liệu điều khiển sao chép dải phổ tăng cường (eSBR) bằng cách tương thích ngược một cách hiệu quả.

Thông thường, siêu dữ liệu eSBR trong dòng bit biểu thị (ví dụ, biểu thị ít nhất một đặc trưng hay tham số của) một hoặc nhiều trong số các công cụ eSBR sau đây (được mô tả trong chuẩn MPEG USAC, và có thể được áp dụng hoặc có thể không được áp dụng bởi bộ mã hóa trong quá trình tạo dòng bit):

- Chuyển vị sóng hài;
- Tiền xử lý bổ sung chỉnh sửa QMF (làm phẳng trước); và
- Tạo hình đường bao thời gian mẫu liên dải con hay “inter-TES”.

Ví dụ, siêu dữ liệu eSBR có trong dòng bit có thể biểu thị các giá trị của các tham số (được mô tả trong chuẩn MPEG USAC và trong bản mô tả này): `harmonicSBR[ch]`, `sbrPatchingMode[ch]`, `sbrOversamplingFlag[ch]`, `sbrPitchInBins[ch]`, `bs_interTes`, `bs_temp_shape[ch][env]`, `bs_inter_temp_shape_mode[ch][env]`, và `bs_sbr_preprocessing`.

Ở đây, ký hiệu `X[ch]`, trong đó `X` là tham số, có nghĩa là tham số liên quan đến kênh (channel - “ch”) nội dung âm thanh của dòng bit mã hóa cần được giải mã. Để đơn giản hóa, chúng ta đôi khi bỏ qua cụm `[ch]`, và giả định rằng tham số liên quan gắn liền với kênh nội dung âm thanh.

Ở đây, ký hiệu `X[ch][env]`, trong đó `X` là tham số, có nghĩa là tham số liên quan đến đường bao SBR (“env”) của kênh (channel, “ch”) nội dung âm thanh của dòng bit mã hóa cần được giải mã. Để đơn giản hóa, chúng ta đôi khi bỏ qua cụm `[env]` và `[ch]`, và giả định rằng tham số liên quan gắn liền với đường bao SBR của kênh nội dung âm thanh.

Như đã nêu, chuẩn MPEG USAC dự kiến rằng dòng bit USAC bao gồm siêu dữ liệu eSBR điều khiển hoạt động xử lý eSBR của bộ giải mã. Siêu dữ liệu eSBR bao gồm các tham số siêu dữ liệu một bit sau đây: `harmonicSBR`; `bs_interTES`; và `bs_pvc`.

Tham số “`harmonicSBR`” biểu thị việc sử dụng kỹ thuật chỉnh sửa sóng hài (chuyển vị sóng hài) cho SBR. Cụ thể, `harmonicSBR = 0` biểu thị chỉnh sửa phổ, không phải sóng hài như được mô tả trong mục 4.6.18.6.3 của chuẩn MPEG-4 AAC; và `harmonicSBR = 1` biểu thị sự chỉnh sửa SBR sóng hài (thuộc loại được sử dụng

trong eSBR, như được mô tả trong mục 7.5.3 hoặc 7.5.4 của chuẩn MPEG USAC). Chính sửa SBR sóng hài không được sử dụng theo sao chép dải phổ không phải eSBR (nghĩa là, SBR mà không phải là eSBR). Trong bản mô tả này, chỉnh sửa phổ được gọi là sao chép dải phổ dạng cơ bản, trong khi đó chuyển vị sóng hài được gọi là dạng sao chép dải phổ dạng tăng cường.

Giá trị của tham số “bs_interTES” biểu thị việc sử dụng công cụ inter-TES của eSBR.

Giá trị của tham số “bs_pvc” biểu thị việc sử dụng công cụ PVC của eSBR.

Trong quá trình giải mã dòng bit đã được mã hóa, việc thực hiện chuyển vị sóng hài trong bước xử lý eSBR của quá trình giải mã (đối với mỗi kênh, “ch”, nội dung âm thanh được biểu thị bởi dòng bit) được điều khiển bởi các tham số siêu dữ liệu eSBR sau đây: sbrPatchingMode[ch]; sbrOversamplingFlag[ch]; sbrPitchInBinsFlag[ch]; và sbrPitchInBins[ch].

Giá trị “sbrPatchingMode[ch]” biểu thị loại chuyển vị sử dụng trong eSBR: sbrPatchingMode[ch] = 1 biểu thị chỉnh sửa không phải sóng hài như được mô tả trong mục 4.6.18.6.3 của chuẩn MPEG-4 AAC; sbrPatchingMode[ch] = 0 biểu thị sự chỉnh sửa SBR sóng hài như được mô tả trong mục 7.5.3 hoặc 7.5.4 của chuẩn MPEG USAC.

Giá trị “sbrOversamplingFlag[ch]” biểu thị việc sử dụng việc tăng tần số lấy mẫu miền tần số thích ứng với tín hiệu trong eSBR kết hợp với chỉnh sửa SBR sóng hài trên cơ sở DFT (biến đổi Fourier rời rạc) như được mô tả trong mục 7.5.3 của chuẩn MPEG USAC. Cờ này điều khiển kích cỡ của các DFT được sử dụng trong chuyển vị: 1 biểu thị việc tăng tần số lấy mẫu miền tần số thích ứng với tín hiệu được kích hoạt như được mô tả trong mục 7.5.3.1 của chuẩn MPEG USAC; 0 biểu thị việc tăng tần số lấy mẫu miền tần số thích ứng với tín hiệu bị vô hiệu hóa như được mô tả trong mục 7.5.3.1 của chuẩn MPEG USAC.

Giá trị “sbrPitchInBinsFlag[ch]” điều khiển sự diễn dịch tham số sbrPitchInBins[ch]: 1 biểu thị rằng giá trị trong sbrPitchInBins[ch] là hợp lệ và lớn hơn không; 0 biểu thị rằng giá trị của sbrPitchInBins[ch] được thiết lập bằng không;

Giá trị “sbrPitchInBins[ch]” điều khiển sự bổ sung các toán hạng tích có hướng trong bộ chuyển vị SBR sóng hài. Giá trị sbrPitchInBins[ch] là giá trị nguyên trong khoảng [0,127] và thể hiện khoảng cách đo được trong các bin tần số

(frequency bin) cho hoạt động DFT 1536 dòng trên tần số lấy mẫu của bộ mã hóa lõi.

Trong trường hợp dòng bit MPEG-4 AAC biểu thị cặp kênh SBR mà các kênh của nó không được ghép nối (thay vì kênh SBR đơn), dòng bit biểu thị hai thực thể của cú pháp nói trên (để chuyển vị sóng hài hoặc không phải sóng hài), mỗi thực thể cho từng kênh của `sbr_channel_pair_element()`.

Chuyển vị sóng hài của công cụ eSBR thường cải thiện chất lượng của các tín hiệu nhạc được giải mã ở các tần số cắt tương đối thấp. Chuyển vị sóng hài nên được thực hiện trong bộ giải mã bằng chuyển vị sóng hài trên cơ sở DTF hoặc QMF. Chuyển vị không phải sóng hài (nghĩa là, chỉnh sửa hoặc sao chép phổ kế thừa) thường cải thiện các tín hiệu tiếng nói. Vì thế, điểm xuất phát trong quyết định ưu tiên loại chuyển vị nào hơn để mã hóa nội dung âm thanh cụ thể chính là chọn phương pháp chuyển vị tùy vào việc nhận dạng tiếng nói/âm nhạc, trong đó chuyển vị sóng hài được sử dụng cho nội dung âm nhạc và chỉnh sửa phổ cho nội dung tiếng nói.

Việc làm phẳng trước trong khi xử lý eSBR được điều khiển bởi giá trị của tham số siêu dữ liệu eSBR một bit là “`bs_sbr_preprocessing`”, có nghĩa là việc thực hiện hay không thực hiện làm phẳng trước tùy thuộc vào giá trị của bit đơn này. Khi thuật toán chỉnh sửa QMF đối với SBR, như đã mô tả trong mục 4.6.18.6.3 của chuẩn MPEG-4 AAC, được sử dụng, có thể thực hiện bước làm phẳng trước (khi được chỉ báo bởi tham số “`bs_sbr_preprocessing`”) để có thể tránh sự đứt đoạn trong hình dạng đường bao phổ của tín hiệu cao tần được nhập vào bộ điều chỉnh đường bao tiếp theo (bộ điều chỉnh đường bao thực hiện bước xử lý eSBR khác). Việc làm phẳng trước thường cải thiện hoạt động của bước điều chỉnh đường bao tiếp theo, tạo ra tín hiệu dải tần cao được nhận thấy là ổn định hơn.

Việc thực hiện tạo hình đường bao thời gian mẫu liên dải con (công cụ “`inter-TES`”), trong quá trình xử lý eSBR trong bộ giải mã, được điều khiển bởi các tham số siêu dữ liệu eSBR sau đây cho mỗi đường bao (“`env`”) SBR của mỗi kênh (“`ch`”) nội dung âm thanh của dòng bit USAC đang được giải mã: `bs_temp_shape[ch][env]`; và `bs_inter_temp_shape_mode[ch][env]`.

Công cụ `inter-TES` xử lý các mẫu dải con QMF tiếp theo sau bộ điều chỉnh đường bao. Bước xử lý này tạo hình đường bao thời gian của dải tần số cao hơn có

độ mịn thời gian cao hơn độ mịn thời gian của bộ điều chỉnh đường bao. Bằng cách áp dụng hệ số độ lợi cho mỗi mẫu dải con QMF trong đường bao SBR, công cụ inter-TES tạo hình đường bao thời gian trong các mẫu dải con QMF.

Tham số “bs_temp_shape[ch][env]” là cờ báo hiệu việc sử dụng công cụ inter-TES. Tham số “bs_inter_temp_shape_mode[ch][env]” biểu thị (như đã xác định trong chuẩn MPEG USAC) các giá trị của tham số γ trong inter-TES.

Tốc độ bit tổng thể được yêu cầu để đưa vào siêu dữ liệu eSBR dòng bit MPEG-4 AAC biểu thị các công cụ eSBR nêu trên (chuyển vị sóng hài, làm phẳng trước và inter_TES) dự kiến bằng khoảng vài trăm bit mỗi giây vì chỉ có dữ liệu điều khiển vi sai cần thiết để thực hiện xử lý eSRB là được truyền theo một số phương án của sáng chế. Các bộ giải mã kế thừa có thể bỏ qua thông tin này vì nó được đưa vào bằng cách tương thích ngược (sẽ được giải thích sau). Vì vậy, ảnh hưởng bất lợi đối với tốc độ bit liên quan đến việc bao gồm siêu dữ liệu eSBR là không đáng kể, vì một số lý do, bao gồm lý do sau:

- Sự giảm tốc độ bit (do bao gồm siêu dữ liệu eSBR) chiếm tỷ lệ rất nhỏ trong tốc độ bit tổng thể vì chỉ có dữ liệu điều khiển vi sai cần thiết để thực hiện xử lý eSBR là được truyền (và không truyền dữ liệu điều khiển SBR đồng thời theo nhiều kênh);
- Sự điều chỉnh thông tin điều khiển liên quan đến SBR thường không phụ thuộc vào các chi tiết chuyển vị; và
- Công cụ inter-TES (được sử dụng trong khi xử lý eSBR) thực hiện hậu xử lý tín hiệu được chuyển vị một ngõ.

Theo đó, các phương án của sáng chế cung cấp phương tiện truyền hiệu quả siêu dữ liệu hoặc dữ liệu điều khiển sao chép dải phổ tăng cường (eSBR) bằng cách tương thích ngược. Việc truyền dữ liệu điều khiển eSBR hiệu quả làm giảm yêu cầu về bộ nhớ trong các bộ giải mã, bộ mã hóa và bộ chuyển mã bằng cách sử dụng các phương án của sáng chế mà vẫn không gây ảnh hưởng bất lợi đáng kể nào đối với tốc độ bit. Ngoài ra, các yêu cầu về độ phức tạp và xử lý liên quan đến việc thực hiện eSBR theo các phương án của sáng chế cũng giảm do dữ liệu SBR chỉ cần được xử lý một lần và không phải truyền đồng thời trên nhiều kênh, đây là trường hợp nếu eSBR được xử lý

như một loại đối tượng riêng biệt hoàn toàn trong MPEG-4 AAC thay vì được tích hợp vào codec MPEG-4 AAC bằng cách tương thích ngược.

Tiếp theo, theo HÌNH 7, chúng ta mô tả các phần tử của khối (“raw_data_block”) của dòng bit MPEG-4 AAC mà trong đó siêu dữ liệu eSBR được bao gồm theo một số phương án của sáng chế. HÌNH 7 là sơ đồ của khối (“raw_data_block”) của dòng bit MPEG-4 AAC, minh họa một số đoạn của khối.

Khối của dòng bit MPEG-4 AAC có thể bao gồm ít nhất một “single_channel_element()” (ví dụ, phần tử kênh đơn được minh họa trong hình 7), và/hoặc ít nhất một “channel_pair_element()” (không được minh họa cụ thể trong Hình 7 mặc dù nó có thể hiện diện), bao gồm dữ liệu âm thanh cho chương trình âm thanh. Khối này cũng có thể bao gồm một số “fill_element” (ví dụ, phần tử chèn 1 và/hoặc phần tử chèn 2 trên Hình 7) bao gồm dữ liệu (ví dụ, siêu dữ liệu) liên quan đến chương trình. Mỗi “single_channel_element()” bao gồm ký hiệu nhận dạng (ví dụ “ID1” trên Hình 7) chỉ báo chỗ bắt đầu của phần tử kênh đơn, và có thể bao gồm dữ liệu âm thanh biểu thị kênh khác của chương trình âm thanh đa kênh. Mỗi “channel_pair_element” bao gồm ký hiệu nhận dạng (không có trong Hình 7) chỉ báo chỗ bắt đầu của phần tử hai kênh, và có thể bao gồm dữ liệu âm thanh biểu thị hai kênh của chương trình.

fill_element (ở đây được gọi là phần tử chèn) của dòng bit MPEG-4 AAC bao gồm ký hiệu nhận dạng (“ID2” trên Hình 7) chỉ báo chỗ bắt đầu của phần tử chèn, và dữ liệu chèn sau ký hiệu nhận dạng. Ký hiệu nhận dạng ID2 có thể bao gồm số nguyên không dấu 3 bit với bit quan trọng nhất truyền trước (“uimsbf”) có giá trị 0x6. Dữ liệu chèn có thể bao gồm phần tử extension_payload() (còn được gọi là tải dữ liệu mở rộng) có cú pháp được trình bày trong Bảng 4.57 của chuẩn MPEG-4 AAC. Có một số loại tải dữ liệu mở rộng và được nhận dạng thông qua tham số “extension_type”, đây là số nguyên không dấu 4 bit với bit quan trọng nhất truyền trước (“uimsbf”).

Dữ liệu chèn (ví dụ, tải dữ liệu mở rộng của nó) có thể bao gồm phần đầu hoặc ký hiệu nhận dạng (ví dụ, “phần đầu 1” trên Hình 7) chỉ báo đoạn dữ liệu chèn biểu thị đối tượng SBR (tức là, phần đầu khởi tạo loại “đối tượng SBR”, được gọi là sbr_extension_data() theo chuẩn MPEG-4 AAC). Ví dụ, tải dữ liệu mở rộng sao chép dải phổ (SBR) được nhận dạng với giá trị ‘1101’ hoặc ‘1110’ cho trường extension_type trong phần đầu, với ký hiệu nhận dạng ‘1101’ nhận dạng tải dữ liệu

mở rộng với dữ liệu SBR và ‘1110’ nhận dạng tải dữ liệu mở rộng với dữ liệu SBR bằng phương pháp Kiểm tra độ dư vòng (CRC) để xác minh độ chính xác của dữ liệu SBR.

Khi phân đầu (ví dụ, trường `extension_type`) khởi tạo loại đối tượng SBR, siêu dữ liệu SBR (thường được gọi là “dữ liệu sao chép dải phổ”, và được gọi là `sbr_data()` theo chuẩn MPEG-4 AAC) theo sau phân đầu, và ít nhất một phần tử mở rộng sao chép dải phổ (ví dụ, “phần tử mở rộng SBR” của phần tử chèn 1 trên Hình 7) có thể theo sau siêu dữ liệu SBR. Phần tử mở rộng sao chép dải phổ này (đoạn dòng bit) được gọi là bộ chứa “`sbr_extension()`” theo chuẩn MPEG-4 AAC. Phần tử mở rộng sao chép dải phổ tùy ý bao gồm phân đầu (ví dụ “phần đầu mở rộng SBR” của phần tử chèn 1 trên Hình 7).

Chuẩn MPEG-4 AAC dự kiến rằng phần tử mở rộng sao chép dải phổ có thể bao gồm dữ liệu PS (âm lập thể tham số) cho dữ liệu âm thanh của chương trình. Chuẩn MPEG-4 AAC dự kiến rằng khi phân đầu của phần tử chèn (ví dụ, tải dữ liệu mở rộng của nó) khởi tạo loại đối tượng SBR (như “phần đầu 1” trên Hình 7 đã làm) và phần tử mở rộng sao chép dải phổ của phần tử chèn bao gồm dữ liệu PS, phần tử chèn (ví dụ tải dữ liệu mở rộng của nó) bao gồm dữ liệu sao chép dải phổ, và tham số “`bs_extension_id`” có giá trị (tức là, `bs_extension_id = 2`) chỉ báo là dữ liệu PS có trong phần tử mở rộng sao chép dải phổ của phần tử chèn.

Theo một số phương án của sáng chế, siêu dữ liệu eSBR (ví dụ, cờ biểu thị sự xử lý sao chép dải phổ tăng cường (eSBR) có cần được thực hiện trên nội dung âm thanh của khối hay không) có trong phần tử mở rộng sao chép dải phổ của phần tử chèn. Ví dụ, cờ này được biểu thị trong phần tử chèn 1 trên Hình 7, trong đó cờ xuất hiện sau phân đầu (“phần đầu mở rộng SBR” của phần tử chèn 1) của “phần tử mở rộng SBR” của phần tử chèn 1. Tùy chọn, cờ này và siêu dữ liệu eSBR bổ sung có trong phần tử mở rộng sao chép dải phổ sau phân đầu của phần tử mở rộng sao chép dải phổ (ví dụ, trong phần tử mở rộng SBR của phần tử chèn 1 trong Hình 7, sau phần đầu mở rộng SBR). Theo một số phương án của sáng chế, phần tử chèn bao gồm siêu dữ liệu eSBR cũng bao gồm tham số “`bs_extension_id`” có giá trị (ví dụ, `bs_extension_id = 3`) chỉ báo siêu dữ liệu eSBR có trong phần tử chèn và việc xử lý eSBR cần được thực hiện trên nội dung âm thanh của khối liên quan.

Theo một số phương án của sáng chế, siêu dữ liệu eSBR có trong phần tử chèn (ví dụ, phần tử chèn 2 trên Hình 7) của dòng bit MPEG-4 AAC ngoài phần tử mở rộng

sao chép dải phổ (phần tử mở rộng SBR) của phần tử chèn. Đó là vì các phần tử chèn chứa `extension_payload()` với dữ liệu SBR hoặc dữ liệu SBR với CRC không chứa bất kỳ tải dữ liệu mở rộng nào khác của các kiểu mở rộng khác. Vì vậy, trong các phương án mà siêu dữ liệu eSBR được lưu trữ, tải dữ liệu mở rộng của riêng nó, phần tử chèn riêng biệt được sử dụng để lưu trữ siêu dữ liệu eSBR. Phần tử chèn này bao gồm ký hiệu nhận dạng (ví dụ, “ID2” trên Hình 7) chỉ báo phần bắt đầu của phần tử chèn, và dữ liệu chèn sau ký hiệu nhận dạng. Dữ liệu chèn có thể bao gồm phần tử `extension_payload()` (còn được gọi là tải dữ liệu mở rộng) có cú pháp được trình bày trong Bảng 4.57 của chuẩn MPEG-4 AAC. Dữ liệu chèn (ví dụ, tải dữ liệu mở rộng của nó) bao gồm phần đầu (ví dụ, “phần đầu 2” của phần tử chèn 2 trên Hình 7) biểu thị đối tượng eSBR (tức là, phần đầu khởi tạo loại đối tượng sao chép dải phổ tăng cường (eSBR)), và dữ liệu chèn (ví dụ, tải dữ liệu mở rộng của nó) bao gồm siêu dữ liệu eSBR sau phần đầu. Ví dụ, phần tử chèn 2 trên Hình 7 bao gồm phần đầu này (“phần đầu 2”) và cũng bao gồm, sau phần đầu, siêu dữ liệu eSBR (tức là, “cờ” trong phần tử chèn 2, biểu thị việc xử lý sao chép dải phổ tăng cường (eSBR) có cần được thực hiện trên nội dung âm thanh của khối hay không). Tùy chọn, siêu dữ liệu eSBR bổ sung cũng có trong dữ liệu chèn của phần tử chèn 2 trên Hình 7, sau phần đầu 2. Trong các phương án đang được mô tả trong đoạn này, phần đầu (ví dụ, phần đầu 2 trên Hình 7) có giá trị nhận dạng không phải là một trong các giá trị quy ước được xác định trong Bảng 4.57 của chuẩn MPEG-4 AAC, và thay vào đó biểu thị tải dữ liệu mở rộng eSBR (sao cho trường `extension_type` của phần đầu chỉ báo dữ liệu chèn bao gồm siêu dữ liệu eSBR).

Trong lớp phương án thứ nhất, sáng chế là bộ xử lý âm thanh (ví dụ, bộ giải mã), bao gồm:

bộ nhớ (ví dụ, bộ đệm 201 trên Hình 3 hoặc 4) được cấu hình để lưu trữ ít nhất một khối của dòng bit âm thanh mã hóa (ví dụ, ít nhất một khối của dòng bit MPEG-4 AAC);

bộ giải định dạng tải dữ liệu dòng bit (ví dụ, phần tử 205 trên Hình 3 hoặc phần tử 215 trên Hình 4) được ghép nối với bộ nhớ và được cấu hình để tách kênh ít nhất một phần của khối của dòng bit nêu trên; và

hệ thống con giải mã (ví dụ, các phần tử 202 và 203 trên Hình 3, hoặc các phần tử 202 và 213 trên Hình 4), được ghép nối và được cấu hình để giải mã ít nhất một phần của nội dung âm thanh của khối của dòng bit đã nêu, trong đó khối bao gồm:

phần tử chèn, bao gồm ký hiệu nhận dạng chỉ báo phần bắt đầu của phần tử chèn (ví dụ, ký hiệu nhận dạng “id_syn_ele” có giá trị 0x6, của Bảng 4.85 của chuẩn MPEG-4 AAC), và dữ liệu chèn sau ký hiệu nhận dạng, trong đó dữ liệu chèn bao gồm:

ít nhất một cờ nhận dạng xem việc xử lý sao chép dải phổ tăng cường (eSBR) có cần được thực hiện trên nội dung âm thanh của khối hay không (ví dụ, sử dụng dữ liệu sao chép dải phổ và siêu dữ liệu eSBR có trong khối).

Cờ là siêu dữ liệu eSBR, và một ví dụ về cờ là cờ sbrPatchingMode. Ví dụ khác về cờ là cờ harmonicSBR. Cả hai cờ này chỉ báo xem sao chép dải phổ dạng cơ bản hay sao chép dải phổ dạng tăng cường sẽ được thực hiện trên dữ liệu âm thanh của khối. Sao chép dải phổ dạng cơ bản là sự chỉnh sửa phổ, và sao chép dải phổ dạng tăng cường là sự chuyển vị sóng hài.

Trong một số phương án, dữ liệu chèn cũng bao gồm siêu dữ liệu eSBR bổ sung (tức là, siêu dữ liệu eSBR không phải là cờ).

Bộ nhớ có thể là bộ nhớ đệm (ví dụ bộ đệm 201 trên Hình 4) lưu trữ (ví dụ, theo cách thức không chuyển tiếp) ít nhất một khối của dòng bit âm thanh mã hóa.

Ước tính rằng độ phức tạp khi bộ giải mã eSBR thực hiện xử lý eSBR (bằng cách chuyển vị sóng hài eSBR, làm phẳng trước, và các công cụ inter_TES) trong quá trình giải mã dòng bit MPEG-4 AAC bao gồm siêu dữ liệu eSBR (biểu thị các công cụ eSBR này) sẽ như sau (cho quy trình giải mã thông thường với các tham số đã nêu):

- Chuyển vị sóng hài (16 kbps, 14400/28800 Hz)
 - Trên cơ sở DFT: 3,68 WMOPS (một triệu phép toán có trọng số trong một giây);
 - Trên cơ sở QMF: 0,98 WMOPS;
- Tiền xử lý chỉnh sửa QMF (làm phẳng trước): 0,1 WMOPS; và
- Tạo hình đường bao thời gian mẫu liên dải con (inter-TES): Tối đa 0,16 WMOPS.

Chúng ta biết rằng sự chuyển vị dựa trên DFT thường được thực hiện tốt hơn sự chuyển vị dựa trên QMF cho nốt đệm.

Theo một số phương án của sáng chế, phần tử chèn (của dòng bit âm thanh mã hóa) bao gồm siêu dữ liệu eSBR cũng bao gồm tham số (ví dụ, tham số “bs_extension_id”) có giá trị (ví dụ, bs_extension_id = 3) chỉ báo rằng siêu dữ liệu

eSBR có trong phần tử chèn và việc xử lý eSBR cần được thực hiện trên nội dung âm thanh của khối liên quan, và/hoặc tham số (ví dụ, cùng tham số “bs_extension_id”) có giá trị (ví dụ, bs_extension_id = 2) chỉ báo rằng bộ chứa sbr_extension() của phần tử chèn bao gồm dữ liệu PS. Ví dụ, như được nêu trong Bảng 1 bên dưới, tham số này có giá trị bs_extension_id = 2 có thể chỉ báo rằng bộ chứa sbr_extension() của phần tử chèn bao gồm dữ liệu PS, và tham số này có giá trị bs_extension_id = 3 có thể chỉ báo rằng bộ chứa sbr_extension() của phần tử chèn bao gồm siêu dữ liệu eSBR:

Bảng 1

bs_extension_id	Ý nghĩa	
0	Dự trữ	
1	Dự trữ	
2	EXTENSION_ID_PS	
3	EXTENSION_ID_ESBR	

Theo một số phương án của sáng chế, cú pháp của mỗi phần tử mở rộng sao chép dải phổ bao gồm siêu dữ liệu eSBR và/hoặc dữ liệu PS như được nêu trong Bảng 2 bên dưới (trong đó “sbr_extension()” biểu thị bộ chứa là phần tử mở rộng sao chép dải phổ, “bs_extension_id” là như đã mô tả trong Bảng 1 bên trên, “ps_data” biểu thị dữ liệu PS, và “esbr_data” biểu thị siêu dữ liệu eSBR):

Bảng 2

sbr_extension(bs_extension_id, num_bits_left)		
{		
switch (bs_extension_id) {		
case EXTENSION_ID_PS:		
num_bits_left -= ps_data();		Ghi chú 1

break;		
case EXTENSION_ID_ESBR:		
num_bits_left -= esbr_data();		Ghi chú 2
break;		
default:		
bs_fill_bits;	Ghi chú 3	
num_bits_left = 0;		
break;		
}		
}		
Ghi chú 1: ps_data() trả về số bit được đọc.		
Ghi chú 2: esbr_data() trả về số bit được đọc.		
Ghi chú 3: tham số bs_fill_bits bao gồm N bit, trong đó N = num_bits_left.		

Trong phương án ví dụ, esbr_data() được đề cập trong Bảng 2 bên trên biểu thị các giá trị của các tham số siêu dữ liệu sau đây:

1. mỗi trong các tham số siêu dữ liệu 1 bit được mô tả bên trên: “harmonicSBR”; “bs_interTES”; và “bs_sbr_preprocessing”;
2. với mỗi kênh (“ch”) nội dung âm thanh của dòng bit mã hóa cần được giải mã, mỗi trong các tham số được mô tả bên trên: “sbrPatchingMode[ch]”; “sbrOversamplingFlag[ch]”; “sbrPitchInBinsFlag[ch]”; và “sbrPitchInBins[ch]”; và
3. với mỗi đường bao SBR (“env”) của mỗi kênh (“ch”) nội dung âm thanh của dòng bit mã hóa cần được giải mã, mỗi trong các tham số được mô tả bên trên: “bs_temp_shape[ch][env]”; và “bs_inter_temp_shape_mode[ch][env]”.

Ví dụ, trong một số phương án, esbr_data() có thể có cú pháp được nêu trong Bảng 3, để biểu thị các tham số siêu dữ liệu này:

Bảng 3

esbr_data()		
{		
harmonicSBR;	1	
bs_interTes;	1	
bs_sbr_preprocessing;	1	
if (harmonicSBR) {		
if (sbrPatchingMode[0] == 0) {	1	
sbrOversamplingFlag[0];	1	
if (sbrPitchInBinsFlag[0])	1	
sbrPitchInBins[0];	7	
Else		
sbrPitchInBins[0] = 0;		
} else {		
sbrOversamplingFlag[0] = 0;		
sbrPitchInBins[0] = 0;		
}		
}		
if (bs_interTes) {		
/* vòng lặp qua ch và env được thực hiện */		

bs_temp_shape[ch][env];	1	
if (bs_temp_shape[ch][env]) {		
bs_inter_temp_shape_mode[ch][env];	2	
}		
}		
}		

Trong Bảng 3, chữ số ở cột giữa chỉ số bit của tham số tương ứng ở cột bên trái.

Cú pháp trên cho phép thực hiện hiệu quả sao chép dải phổ dạng tăng cường, như sự chuyển vị sóng hài, làm phần mở rộng cho bộ giải mã kế thừa. Cụ thể, dữ liệu eSBR của Bảng 3 chỉ bao gồm các tham số cần thiết để thực hiện sao chép dải phổ dạng tăng cường mà không được hỗ trợ trong dòng bit hoặc không được dẫn xuất trực tiếp từ các tham số đã được hỗ trợ trong dòng bit. Tất cả các tham số và dữ liệu xử lý khác cần thiết để thực hiện sao chép dải phổ dạng tăng cường được trích xuất từ các tham số có từ trước tại những vị trí đã được xác định trong dòng bit. Điều này trái ngược với phương án thực hiện khác (và ít hiệu quả hơn) là chỉ truyền tất cả siêu dữ liệu xử lý dùng để sao chép dải phổ tăng cường.

Ví dụ, bộ giải mã theo chuẩn MPEG-4 HE-AAC hoặc HE-AAC v2 có thể được mở rộng để bao gồm sao chép dải phổ dạng tăng cường, như chuyển vị sóng hài. Sao chép dải phổ dạng tăng cường này cùng với sao chép dải phổ dạng cơ bản đã được bộ giải mã hỗ trợ. Trong ngữ cảnh của bộ giải mã theo chuẩn MPEG-4 HE-AAC hoặc HE-AAC v2, sao chép dải phổ dạng cơ bản này là công cụ SBR chỉnh sửa phổ QMF theo định nghĩa trong mục 4.6.18 của chuẩn MPEG-4 AAC.

Khi thực hiện sao chép dải phổ dạng tăng cường, bộ giải mã HE-AAC mở rộng có thể sử dụng lại nhiều tham số dòng bit đã có trong tải dữ liệu mở rộng SBR của dòng bit. Các tham số cụ thể có thể được sử dụng lại bao gồm, ví dụ, các tham số khác nhau xác định băng dải tần số chủ. Các tham số này bao gồm bs_start_freq (tham số xác định chỗ bắt đầu của tham số băng tần số chủ), bs_stop_freq (tham số xác định chỗ dừng băng tần số chủ), bs_freq_scale (tham số xác định số dải tần số cho mỗi bát độ), và

bs_alter_scale (tham số thay đổi thang đo dải tần số). Các tham số có thể được sử dụng lại cũng bao gồm các tham số xác định băng dải nhiễu âm (bs_noise_bands) và các tham số băng dải hạn chế (bs_limiter_bands).

Ngoài các tham số, các phần tử dữ liệu khác có thể cũng có thể được sử dụng lại bởi bộ giải mã HE-AAC mở rộng khi thực hiện sao chép dải phổ dạng tăng cường theo các phương án của sáng chế. Ví dụ, dữ liệu đường bao và dữ liệu sản nhiễu cũng có thể được trích xuất từ dữ liệu bs_data_env và bs_noise_env và được sử dụng trong sao chép dải phổ dạng tăng cường.

Về bản chất, các phương án này khai thác các tham số cấu hình và dữ liệu đường bao đã được hỗ trợ bởi bộ giải mã HE-AAC hoặc HE-AAC v2 kế thừa trong tải dữ liệu mở rộng SBR để cho phép sao chép dải phổ dạng tăng cường yêu cầu càng ít dữ liệu truyền bổ sung càng tốt. Theo đó, các bộ giải mã mở rộng hỗ trợ sao chép dải phổ dạng tăng cường có thể được tạo ra một cách rất hiệu quả bằng cách dựa vào các phần tử dòng bit đã được xác định (ví dụ, các phần tử trong tải dữ liệu mở rộng SBR) và chỉ thêm các tham số cần thiết để hỗ trợ sao chép dải phổ dạng tăng cường (trong tải dữ liệu mở rộng phần tử chèn). Đặc tính giảm dữ liệu này kết hợp với việc sắp xếp các tham số bổ sung mới trong trường dữ liệu dự trữ, như bộ chứa mở rộng, đã làm giảm đáng kể những rào cản đối với việc tạo bộ giải mã hỗ trợ sao chép dải phổ dạng tăng cường bằng cách đảm bảo dòng bit tương thích ngược với bộ giải mã kế thừa không hỗ trợ sao chép dải phổ dạng tăng cường.

Trong một số phương án, sáng chế đề cập đến phương pháp bao gồm bước mã hóa dữ liệu âm thanh để tạo dòng bit mã hóa (ví dụ, dòng bit MPEG-4 AAC), bao gồm bằng cách đưa siêu dữ liệu eSBR vào trong ít nhất một đoạn của ít nhất một khối của dòng bit mã hóa và dữ liệu âm thanh ở ít nhất một đoạn khác của khối. Trong các phương án điển hình, phương pháp bao gồm bước ghép kênh dữ liệu âm thanh có siêu dữ liệu eSBR trong mỗi khối của dòng bit mã hóa. Trong bước giải mã điển hình đối với dòng bit mã hóa trong bộ giải mã eSBR, bộ giải mã trích xuất siêu dữ liệu eSBR từ dòng bit (bao gồm bước phân tích cú pháp và bước tách kênh siêu dữ liệu eSBR và dữ liệu âm thanh) và sử dụng siêu dữ liệu eSBR để xử lý dữ liệu âm thanh để tạo dòng dữ liệu âm thanh được giải mã.

Phương án khác của sáng chế là bộ giải mã eSBR được cấu hình để thực hiện xử lý eSBR (ví dụ, sử dụng ít nhất một trong các công cụ eSBR là chuyển vị sóng

hải, làm phẳng trước, hoặc inter_TES) trong quá trình giải mã dòng bit âm thanh mã hóa (ví dụ, dòng bit MPEG-4 AAC) không bao gồm siêu dữ liệu eSBR. Ví dụ về bộ giải mã này sẽ được mô tả theo Hình 5.

Bộ giải mã eSBR (400) trên Hình 5 bao gồm bộ nhớ đệm 201 (giống với bộ nhớ 201 trên Hình 3 và 4), bộ giải định dạng tải dữ liệu dòng bit 215 (giống với bộ giải định dạng 215 trên Hình 4), hệ thống con giải mã âm thanh 202 (thường được gọi là giai đoạn giải mã “lỗi” hoặc hệ thống con giải mã “lỗi”, và giống với hệ thống con giải mã lỗi 202 trên Hình 3), hệ thống con tạo dữ liệu điều khiển eSBR 401, và bước xử lý eSBR 203 (giống với bước 203 trên Hình 3), được ghép nối như trong hình. Ngoài ra, thông thường bộ giải mã 400 còn bao gồm các phần tử xử lý khác (không được minh họa).

Trong quá trình hoạt động của bộ giải mã 400, chuỗi khối của dòng bit âm thanh mã hóa (dòng bit MPEG-4 AAC) nhận được bởi bộ giải mã 400 được gửi từ bộ đệm 201 đến bộ giải định dạng 215.

Bộ giải định dạng 215 được ghép nối và được cấu hình để tách kênh mỗi khối của dòng bit để trích xuất siêu dữ liệu SBR (bao gồm dữ liệu đường bao được lượng tử hóa) và thông thường còn cả siêu dữ liệu khác từ đó. Bộ giải định dạng 215 được cấu hình để gửi ít nhất siêu dữ liệu SBR đến bước xử lý eSBR 203. Bộ giải định dạng 215 cũng được ghép nối và được cấu hình để trích xuất dữ liệu âm thanh từ mỗi khối của dòng bit, và để gửi dữ liệu âm thanh được trích xuất đến hệ thống con giải mã (bước giải mã) 202.

Hệ thống con giải mã âm thanh 202 của bộ giải mã 400 được cấu hình để giải mã dữ liệu âm thanh được trích xuất bởi bộ giải định dạng 215 (bước giải mã này có thể được coi là hoạt động giải mã “lỗi”) để tạo dữ liệu âm thanh được giải mã, và gửi dữ liệu âm thanh được giải mã đến bước xử lý eSBR 203. Bước giải mã được thực hiện trong miền tần số. Thông thường, bước xử lý cuối cùng trong hệ thống con 202 áp dụng biến đổi miền tần số thành miền thời gian cho dữ liệu âm thanh miền tần số được giải mã, sao cho dữ liệu xuất của hệ thống con là miền thời gian, dữ liệu âm thanh được giải mã. Bước 203 được cấu hình để áp dụng các công cụ SBR (và các công cụ eSBR) được biểu thị bằng siêu dữ liệu SBR (được trích xuất bởi bộ giải định dạng 215) và bằng siêu dữ liệu eSBR được tạo trong hệ thống con 401, cho dữ liệu âm thanh được giải mã (tức là, để thực hiện xử lý SBR và eSBR trên dữ liệu

xuất của hệ thống con giải mã 202 bằng cách sử dụng siêu dữ liệu eSBR và SBR) để tạo dữ liệu âm thanh được giải mã hoàn toàn mà được xuất từ bộ giải mã 400. Thông thường, bộ giải mã 400 bao gồm bộ nhớ (có thể truy nhập bởi hệ thống con 202 và bước 203) lưu trữ dữ liệu âm thanh được giải định dạng và siêu dữ liệu được xuất từ bộ giải định dạng 215 (và cũng có thể từ hệ thống con 401), và bước 203 được cấu hình để truy nhập siêu dữ liệu và dữ liệu âm thanh khi cần trong quá trình xử lý SBR và eSBR. Xử lý SBR trong bước 203 có thể được coi là hậu xử lý trên dữ liệu xuất của hệ thống con giải mã lõi 202. Tùy chọn, bộ giải mã 400 cũng bao gồm hệ thống con trộn tăng cuối cùng (có thể áp dụng các công cụ âm lập thể tham số (“PS”) được xác định trong chuẩn MPEG-4 AAC, sử dụng siêu dữ liệu PS được trích xuất bởi bộ giải định dạng 215) được ghép nối và được cấu hình để thực hiện trộn tăng trên dữ liệu xuất của bước 203 để tạo âm thanh được trộn tăng, giải mã hoàn toàn, được xuất từ APU 210.

Hệ thống con tạo dữ liệu điều khiển 401 trên Hình 5 được ghép nối và được cấu hình để phát hiện ít nhất một đặc tính của dòng bit âm thanh mã hóa cần được giải mã, và để tạo dữ liệu điều khiển eSBR (có thể là hoặc bao gồm siêu dữ liệu eSBR thuộc bất kỳ loại nào có trong dòng bit âm thanh mã hóa theo các phương án khác của sáng chế) để đáp ứng lại ít nhất một kết quả của bước phát hiện. Dữ liệu điều khiển eSBR được gửi đến bước 203 để kích hoạt việc áp dụng từng công cụ eSBR riêng lẻ hoặc tổ hợp công cụ eSBR ngay khi phát hiện một đặc tính cụ thể (hoặc tổ hợp đặc tính) của dòng bit, và/hoặc để điều khiển việc áp dụng các công cụ eSBR này. Ví dụ, để điều khiển việc thực hiện xử lý eSBR bằng cách sử dụng sự chuyển vị sóng hài, một số phương án về hệ thống con tạo dữ liệu điều khiển 401 sẽ bao gồm: bộ dò nhạc (ví dụ, phiên bản tối giản của bộ dò nhạc thông thường) để thiết lập tham số `sbrPatchingMode[ch]` (và gửi tham số được thiết lập đến bước 203) để đáp ứng việc phát hiện dòng bit biểu thị hoặc không biểu thị nhạc; bộ dò nốt đệm để thiết lập tham số `sbrOversamplingFlag[ch]` (và gửi tham số được thiết lập đến bước 203) để đáp ứng sự phát hiện việc có hay không có các nốt đệm trong nội dung âm thanh được biểu thị bởi dòng bit; và/hoặc bộ dò âm vực để thiết lập các tham số `sbrPitchInBinsFlag[ch]` và `sbrPitchInBins[ch]` (và gửi các tham số được thiết lập đến bước 203) để đáp ứng sự phát hiện âm vực của nội dung âm thanh được biểu thị bởi dòng bit. Các phương án khác của sáng chế là các phương pháp giải mã dòng bit âm thanh được thực hiện theo

phương án bất kỳ về bộ giải mã của sáng chế được mô tả trong đoạn này và đoạn trước.

Các khía cạnh của sáng chế bao gồm phương pháp mã hóa hoặc giải mã mà phương án bất kỳ về APU, hệ thống hoặc thiết bị của sáng chế được cấu hình (ví dụ, được lập trình) để thực hiện. Các phương án khác của sáng chế bao gồm hệ thống hoặc thiết bị được cấu hình (ví dụ, được lập trình) để thực hiện phương án bất kỳ về phương pháp của sáng chế, và vật ghi đọc được bằng máy tính (ví dụ, đĩa) lưu trữ mã (ví dụ, theo cách thức không chuyển tiếp) để thực hiện phương án bất kỳ về phương pháp hoặc các bước của sáng chế. Ví dụ, hệ thống của sáng chế có thể là hoặc bao gồm bộ xử lý đa dụng lập trình được, bộ xử lý tín hiệu số, hoặc bộ vi xử lý, được lập trình với phần mềm hoặc phần sụn và/hoặc được cấu hình để thực hiện hoạt động bất kỳ trong số rất nhiều hoạt động trên dữ liệu, bao gồm phương án về phương pháp của sáng chế hoặc các bước của phương pháp. Bộ xử lý đa dụng này có thể là hoặc bao gồm hệ thống máy tính gồm thiết bị đầu vào, bộ nhớ, và hệ mạch xử lý được lập trình (và/hoặc được cấu hình) để thực hiện phương án về phương pháp của sáng chế (hoặc các bước của phương pháp) để đáp ứng dữ liệu được gửi đến.

Các phương án của sáng chế có thể được thực hiện trong phần cứng, phần sụn, hoặc phần mềm, hoặc kết hợp cả hai (ví dụ, mảng logic lập trình được). Trừ khi được quy định khác, các thuật toán hoặc các quy trình được bao gồm như một phần của sáng chế vốn không liên quan đến bất kỳ máy tính cụ thể nào hoặc các thiết bị khác. Cụ thể, các máy đa dụng khác nhau có thể được sử dụng với những chương trình được viết theo các hướng dẫn trong tài liệu này, hoặc có thể sẽ thuận tiện hơn khi thiết lập các thiết bị chuyên dụng hơn (ví dụ, các mạch tích hợp) để thực hiện các bước của phương pháp theo yêu cầu. Theo đó, sáng chế có thể được thực hiện trong một hoặc nhiều chương trình máy tính thực thi trên một hoặc nhiều hệ thống máy tính lập trình được (ví dụ, hệ thống bao gồm phần tử bất kỳ trên Hình 1, hoặc bộ mã hóa 100 trên Hình 2 (hoặc phần tử của bộ mã hóa), hoặc bộ giải mã 200 trên Hình 3 (hoặc phần tử của bộ giải mã), hoặc bộ giải mã 210 trên Hình 4 (hoặc phần tử của bộ giải mã), hoặc bộ giải mã 400 trên Hình 5 (hoặc phần tử của bộ giải mã)), mỗi hệ thống bao gồm ít nhất một bộ xử lý, ít nhất một hệ thống lưu trữ dữ liệu (bao gồm bộ nhớ khả biến và bất biến và/hoặc các phần tử lưu trữ), ít nhất một cổng hoặc thiết bị đầu vào, và ít nhất một cổng hoặc thiết bị

đầu ra. Mã chương trình được áp dụng cho dữ liệu nhập để thực hiện các chức năng được mô tả trong bản mô tả này và tạo thông tin đầu ra. Thông tin đầu ra được áp dụng cho một hoặc nhiều thiết bị đầu ra theo cách đã biết.

Mỗi chương trình này có thể được thực hiện bằng bất kỳ ngôn ngữ máy tính mong muốn nào (bao gồm ngôn ngữ lập trình định hướng máy, hợp ngữ, hoặc thủ tục bậc cao, logic, hoặc đối tượng) để giao tiếp với hệ thống máy tính. Trong mọi trường hợp, ngôn ngữ có thể là ngôn ngữ biên dịch hoặc diễn dịch.

Ví dụ, khi được thực hiện bởi các chuỗi chỉ lệnh phần mềm máy tính, các chức năng và các bước khác nhau của các phương án theo sáng chế có thể được thực hiện bởi chuỗi chỉ lệnh phần mềm đa luồng chạy trong phần cứng xử lý tín hiệu số thích hợp, trong trường hợp này các thiết bị, các bước và chức năng khác nhau của các phương án có thể tương ứng với các phần chỉ lệnh phần mềm.

Mỗi chương trình máy tính này tốt hơn là được lưu trữ vào hoặc tải về thiết bị hoặc phương tiện lưu trữ (ví dụ, phương tiện hoặc bộ nhớ mạch rắn, hoặc phương tiện từ tính hoặc quang học) đọc được bằng máy tính được lập trình đa dụng hoặc chuyên dụng, để cấu hình và vận hành máy tính khi thiết bị hoặc phương tiện lưu trữ được đọc bởi hệ thống máy tính để thực hiện các quy trình được mô tả trong tài liệu này. Hệ thống theo sáng chế cũng có thể được triển khai dưới dạng vật ghi đọc được bằng máy tính, được cấu hình với (tức là, lưu trữ) chương trình máy tính, trong đó vật ghi được cấu hình sao cho hệ thống máy tính hoạt động một cách cụ thể và được xác định trước để thực hiện các chức năng được mô tả trong tài liệu này.

Một số phương án của sáng chế đã được mô tả. Tuy nhiên, cần hiểu rằng có thể thực hiện các phương án cải biến khác nhau mà không tách rời nội dung và phạm vi của sáng chế. Có thể có các phương án cải biến và biến đổi đối với sáng chế theo các hướng dẫn nêu trên. Cần hiểu rằng trong phạm vi của các điểm yêu cầu bảo hộ kèm theo, sáng chế có thể được thực hiện theo cách khác ngoài cách được mô tả cụ thể trong tài liệu này. Các số tham chiếu có trong các điểm yêu cầu bảo hộ sau đây chỉ nhằm mục đích minh họa và không được sử dụng để diễn dịch hoặc giới hạn các điểm yêu cầu bảo hộ dưới mọi hình thức.

YÊU CẦU BẢO HỘ

1. Bộ xử lý âm thanh bao gồm:

bộ đệm được cấu hình để lưu trữ ít nhất một khối của dòng bit âm thanh mã hóa;

bộ giải định dạng tải dữ liệu dòng bit được ghép nối với bộ đệm và được cấu hình để tách kênh ít nhất một phần của ít nhất một khối của dòng bit âm thanh mã hóa; và

hệ thống con giải mã được ghép nối với bộ giải định dạng tải dữ liệu dòng bit và được cấu hình để giải mã ít nhất một phần của ít nhất một khối của dòng bit âm thanh mã hóa, trong đó ít nhất một khối của dòng bit âm thanh mã hóa bao gồm:

phần tử chèn có ký hiệu nhận dạng chỉ báo chỗ bắt đầu của phần tử chèn và dữ liệu chèn sau ký hiệu nhận dạng, trong đó dữ liệu chèn bao gồm:

ít nhất một cờ nhận dạng xem sao chép dải phổ dạng cơ bản hay sao chép dải phổ dạng tăng cường cần được thực hiện trên nội dung âm thanh của ít nhất một khối của dòng bit âm thanh mã hóa, trong đó sao chép dải phổ dạng cơ bản bao gồm sự chỉnh sửa phổ, sao chép dải phổ dạng tăng cường bao gồm sự chuyển vị sóng hài, một giá trị của cờ chỉ báo rằng sao chép dải phổ dạng tăng cường này phải được thực hiện trên nội dung âm thanh, và giá trị khác của cờ chỉ báo rằng sao chép dải phổ dạng cơ bản chứ không phải chuyển vị sóng hài phải được thực hiện trên nội dung âm thanh.

2. Bộ xử lý âm thanh theo điểm 1, trong đó dữ liệu chèn còn bao gồm siêu dữ liệu sao chép dải phổ tăng cường.

3. Bộ xử lý âm thanh theo điểm 2, trong đó siêu dữ liệu sao chép dải phổ tăng cường được chứa trong tải dữ liệu mở rộng của phần tử chèn.

4. Bộ xử lý âm thanh theo điểm bất kỳ trong số các điểm từ 2 đến 3, trong đó siêu dữ liệu sao chép dải phổ tăng cường bao gồm một hoặc nhiều tham số xác định băng dải tần số chủ.

5. Bộ xử lý âm thanh theo điểm bất kỳ trong số các điểm từ 2 đến 3, trong đó siêu dữ liệu sao chép dải phổ tăng cường bao gồm các hệ số tỷ lệ đường bao hoặc hệ số tỷ lệ sản nhiễu.

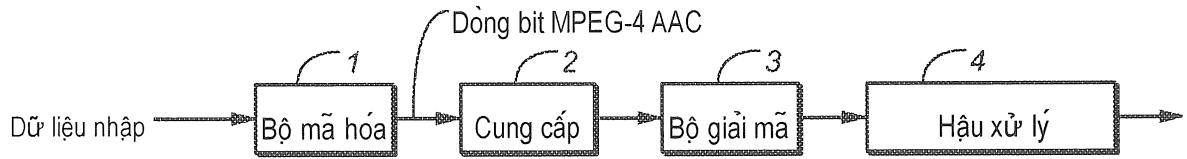
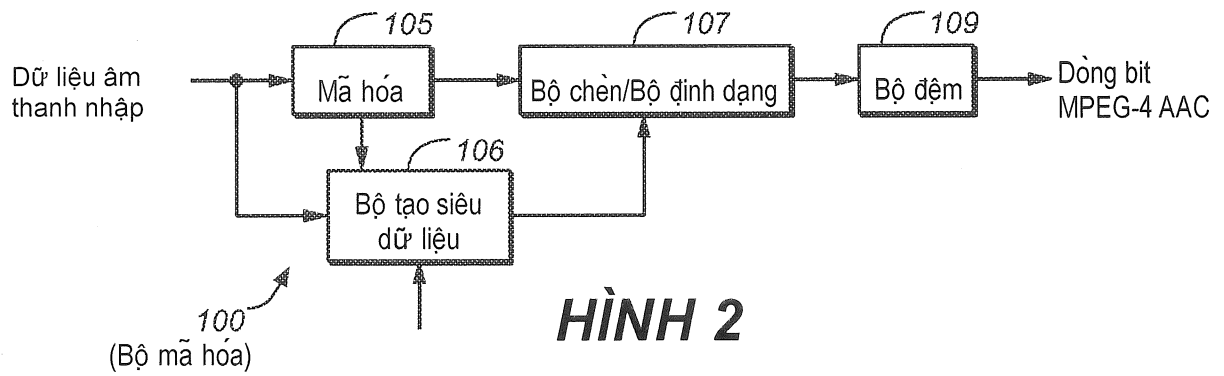
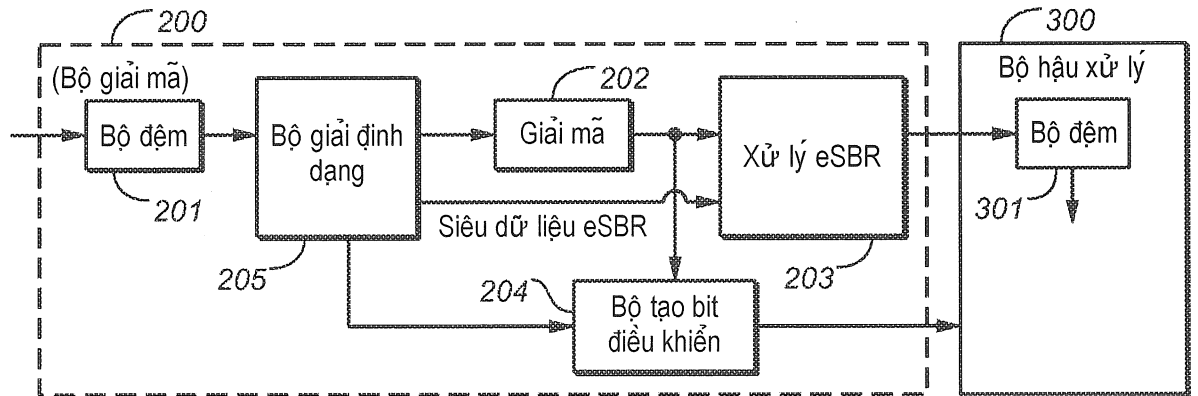
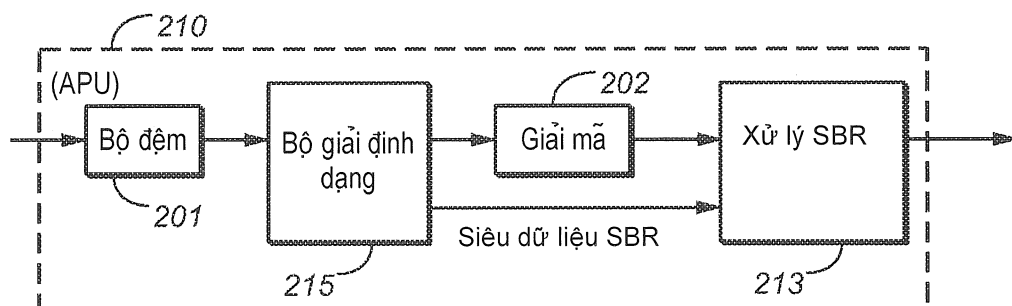
6. Phương pháp giải mã dòng bit âm thanh mã hóa, phương pháp này bao gồm các bước:

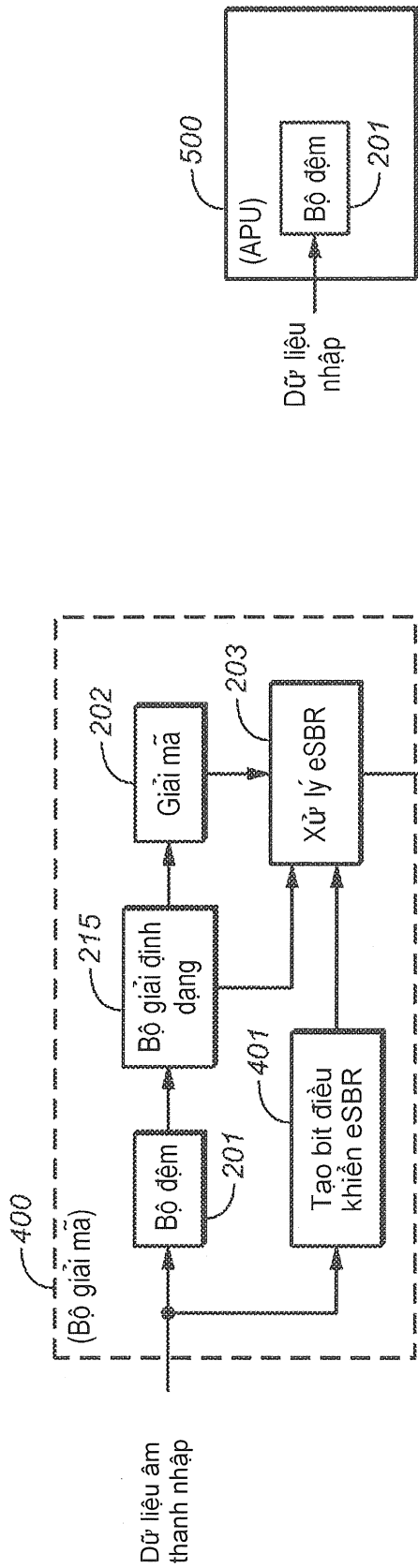
nhận ít nhất một khối của dòng bit âm thanh mã hóa;
tách kênh ít nhất một phần của ít nhất một khối của dòng bit âm thanh mã hóa; và
giải mã ít nhất một phần của ít nhất một khối của dòng bit âm thanh mã hóa, trong đó ít nhất một khối của dòng bit âm thanh mã hóa bao gồm:
phần tử chèn có ký hiệu nhận dạng chỉ báo chỗ bắt đầu của phần tử chèn và dữ liệu chèn sau ký hiệu nhận dạng, trong đó dữ liệu chèn bao gồm:

ít nhất một cờ nhận dạng xem sao chép dải phổ dạng cơ bản hay sao chép dải phổ dạng tăng cường cần được thực hiện trên nội dung âm thanh của ít nhất một khối của dòng bit âm thanh mã hóa, trong đó sao chép dải phổ dạng cơ bản bao gồm sự chỉnh sửa phổ, sao chép dải phổ dạng tăng cường bao gồm sự chuyển vị sóng hài, một giá trị của cờ chỉ báo rằng sao chép dải phổ dạng tăng cường này phải được thực hiện trên nội dung âm thanh, và giá trị khác của cờ này chỉ báo rằng sao chép dải phổ dạng cơ bản chứ không phải chuyển vị sóng hài phải được thực hiện trên nội dung âm thanh.

7. Phương pháp theo điểm 6, trong đó ký hiệu nhận dạng là số nguyên không dấu ba bit với bit có ý nghĩa nhất truyền trước và có giá trị 0x6.

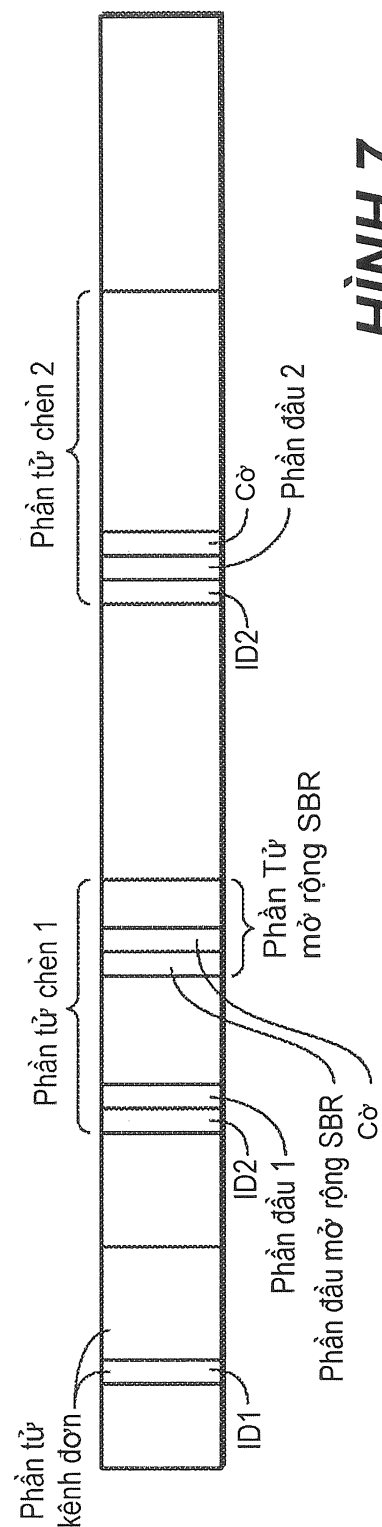
8. Phương pháp theo điểm 6 hoặc 7, trong đó dữ liệu chèn còn bao gồm siêu dữ liệu sao chép dải phổ tăng cường.

**HÌNH 1****HÌNH 2****HÌNH 3****HÌNH 4**



HÌNH 6

HÌNH 5 Dữ liệu xuất



HÌNH 7