



(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)



1-0036322

(51)⁷ G10L 19/008

(13) B

(21) 1-2018-01904

(22) 07/10/2016

(86) PCT/EP2016/073971 07/10/2016

(87) WO 2017/060412 A1 13/04/2017

(30) 15306591.7 08/10/2015 EP; 62/361,863 13/07/2016 US

(45) 25/07/2023 424

(43) 25/07/2018 364A

(73) Dolby International AB (NL)

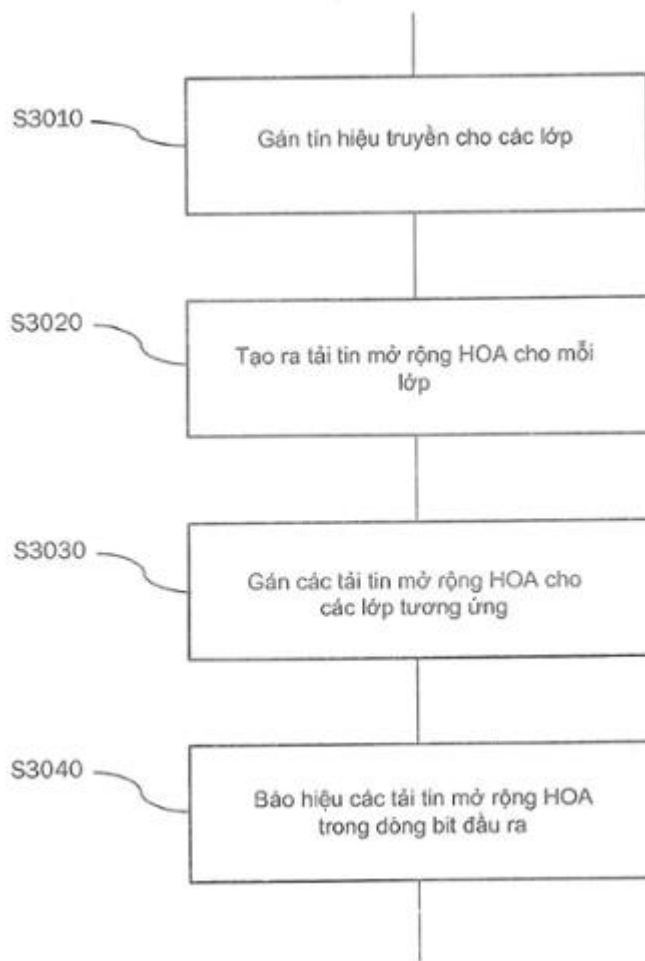
Apollo Building, 3E, Herikerbergweg 1-35, 1101 CN Amsterdam Zuidoost,
Netherlands

(72) KORDON, Sven (DE); KRUEGER, Alexander (DE).

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

(54) PHƯƠNG PHÁP VÀ THIẾT BỊ GIẢI MÃ DẠNG BIỂU DIỄN AMBISONICS BẬC CAO (HOA) NÉN CỦA ÂM THANH HOẶC TRƯỜNG ÂM THANH

(57) Sáng chế đề cập đến phương pháp mã hóa theo lớp đối với khung của dạng biểu diễn Ambisonics bậc cao, HOA, nén của âm thanh hoặc trường âm thanh. Dạng biểu diễn HOA nén bao gồm nhiều tín hiệu truyền. Phương pháp này bao gồm bước gán nhiều tín hiệu truyền cho nhiều lớp phân cấp, nhiều lớp bao gồm lớp cơ sở và một hoặc nhiều lớp tăng cường phân cấp, tạo ra, đối với mỗi lớp, tải tin mở rộng HOA tương ứng bao gồm thông tin phụ trợ để tăng cường tham số cho dạng biểu diễn HOA tái tạo thu được từ các tín hiệu truyền được gán cho lớp tương ứng và các lớp bất kỳ thấp hơn lớp tương ứng này, gán các tải tin mở rộng HOA đã tạo ra cho các lớp tương ứng của chúng, và báo hiệu các tải tin mở rộng HOA đã tạo ra trong dòng bit đầu ra. Sáng chế còn đề cập đến thiết bị và phương pháp giải mã dạng biểu diễn HOA nén của âm thanh hoặc trường âm thanh.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến các phương pháp và thiết bị để mã hóa âm thanh theo lớp. Cụ thể là sáng chế đề cập đến các phương pháp và thiết bị để mã hóa âm thanh theo lớp các khung của các dạng biểu diễn âm thanh (hoặc trường âm thanh) ambisonics bậc cao (Higher-Order Ambisonics - HOA) nén. Sáng chế còn đề cập đến cấu trúc dữ liệu (ví dụ, dòng bit) để biểu diễn các khung của các dạng biểu diễn âm thanh (hoặc trường âm thanh) ambisonics bậc cao (HOA) nén.

Tình trạng kỹ thuật của sáng chế

Theo định nghĩa hiện nay về mã hóa theo lớp HOA, thông tin phụ trợ cho các công cụ giải mã HOA như Dự báo tín hiệu không gian, Tổng hợp tín hiệu định hướng dải con và Bộ giải mã tái tạo môi trường tham số (Parametric Ambience Replication - PAR) được tạo ra để tăng cường dạng biểu diễn HOA đặc trưng. Đó là, theo định nghĩa hiện tại về mã hóa HOA theo lớp dữ liệu được cung cấp chỉ mở rộng thích đáng dạng biểu diễn HOA của lớp cao nhất (ví dụ, lớp tăng cường cao nhất). Đối với các lớp thấp hơn bao gồm lớp cơ sở các công cụ này không tăng cường dạng biểu diễn HOA được tái tạo một phần một cách thích đáng.

Các công cụ Tổng hợp tín hiệu định hướng dải con và Bộ giải mã tái tạo môi trường tham số được thiết kế đặc thù cho các tốc độ dữ liệu thấp, nơi chỉ có vài tín hiệu truyền là khả dụng. Tuy nhiên, trong mã hóa HOA theo lớp sự tăng cường thích đáng của các dạng biểu diễn HOA được tái tạo (một phần) là không thể đặc biệt là đối với các lớp có tốc độ bit thấp, như lớp cơ sở. Điều này rõ ràng là không mong muốn từ góc độ chất lượng âm thanh ở tốc độ bit thấp.

Ngoài ra, phát hiện ra rằng cách thông thường để xử lý các phần tử vector V được mã hóa đối với các tín hiệu dựa trên vector không dẫn đến việc giải mã thích hợp nếu `CodedVVecLength` bằng một được báo hiệu trong `HOADecoderConfig()` (tức là, chế độ mã hóa vector hoạt động). Theo chế độ mã hóa vector này các phần tử vector V không được truyền cho các chỉ số hệ số HOA mà được bao gồm trong tập `ContAddHoaCoeff`. Tập này bao gồm tất cả các chỉ số hệ số HOA `AmbCoeffIdx[i]` mà có

AmbCoeffTransitionState bằng không. Thông thường, không cần phải bổ sung tín hiệu vector V có trọng số vì chuỗi hệ số HOA ban đầu cho các chỉ số này được gửi (báo hiệu) rõ ràng. Vì vậy phần tử vector V được thiết lập về không cho các chỉ số này.

Tuy nhiên, trong chế độ mã hóa theo lớp tập chỉ số hệ số HOA liên tiếp phụ thuộc vào các kênh vận chuyển là một phần của lớp hoạt động hiện thời. Các chỉ số hệ số HOA bổ sung mà được gửi trong lớp cao hơn có thể thiếu trong lớp thấp hơn. Sau đó giả định rằng tín hiệu vector không đóng góp vào chuỗi hệ số HOA là sai đối với các chỉ số hệ số HOA mà thuộc chuỗi hệ số HOA bao gồm trong các lớp cao hơn.

Hệ quả là, vector V trong mã hóa HOA theo lớp có thể không thích hợp để giải mã các lớp bất kỳ dưới lớp cao nhất.

Do đó, có nhu cầu về các sơ đồ mã hóa và dòng bit mà được làm thích ứng để mã hóa theo lớp của các dạng biểu diễn HOA nén của âm thanh hoặc trường âm thanh.

Sáng chế giải quyết các vấn đề trên. Cụ thể, phương pháp và bộ mã hóa/bộ giải mã để mã hóa theo lớp các khung của dạng biểu diễn âm thanh hoặc trường âm thanh HOA nén cũng như cấu trúc dữ liệu để biểu diễn các khung của dạng biểu diễn âm thanh hoặc trường âm thanh HOA nén được mô tả.

Bản chất kỹ thuật của sáng chế

Theo một khía cạnh, phương pháp mã hóa theo lớp đối với khung của dạng biểu diễn Ambisonics bậc cao (HOA) nén của âm thanh hoặc trường âm thanh được mô tả. Dạng biểu diễn HOA nén phù hợp với chuẩn âm thanh 3D MPEG-H dự thảo và các chuẩn dự thảo hoặc được chấp nhận trong tương lai bất kỳ khác. Dạng biểu diễn HOA nén có thể bao gồm nhiều tín hiệu truyền. Các tín hiệu truyền có thể liên quan đến tín hiệu nghe một tai, ví dụ, biểu diễn tín hiệu âm thanh nổi trội hoặc chuỗi hệ số của dạng biểu diễn HOA. Phương pháp có thể bao gồm bước gán nhiều tín hiệu truyền cho nhiều lớp phân cấp. Ví dụ, các tín hiệu truyền có thể được phân bổ cho nhiều lớp. Các lớp có thể bao gồm một lớp cơ sở và một hoặc nhiều lớp tăng cường phân cấp. Nhiều lớp phân cấp có thể được sắp xếp, từ lớp cơ sở, đến lớp tăng cường thứ nhất, lớp tăng cường thứ hai, v.v., đến lớp tăng cường tổng thể cao nhất (lớp tổng thể cao nhất). Phương pháp có thể còn bao gồm bước tạo, đối với mỗi lớp, tải tin mở rộng HOA tương ứng bao gồm thông tin phụ trợ (ví dụ, thông tin phụ trợ tăng cường) để tăng cường tham số cho dạng biểu diễn HOA tái tạo thu được từ các tín hiệu truyền được gán cho lớp tương

ứng và các lớp bất kỳ thấp hơn lớp tương ứng này. Các dạng biểu diễn HOA tái tạo cho các lớp thấp hơn có thể được gọi là dạng biểu diễn HOA tái tạo một phần. Phương pháp có thể còn bao gồm bước gán các tải tin mở rộng HOA được tạo ra cho các lớp tương ứng của chúng. Phương pháp có thể còn bao gồm bước báo hiệu các tải tin mở rộng HOA được tạo ra trong dòng bit đầu ra. Các tải tin mở rộng HOA có thể được báo hiệu trong tải tin HOAEnhFrame(). Do đó, thông tin phụ trợ có thể được dịch chuyển từ HOAFrame() đến HOAEnhFrame().

Được tạo cấu hình như trên, phương pháp được đề xuất áp dụng mã hóa theo lớp cho (khung của) dạng biểu diễn HOA nén để kích hoạt giải mã chất lượng cao của chúng ngay cả ở tốc độ bit thấp. Cụ thể, phương pháp được đề xuất bảo đảm rằng mỗi lớp bao gồm tải tin mở rộng HOA thích hợp (ví dụ, thông tin phụ trợ tăng cường) để tăng cường dạng biểu diễn âm thanh tái tạo (một phần) thu được từ các tín hiệu truyền ở các lớp bất kỳ lên đến lớp hiện thời. Ở đó các lớp lên đến lớp hiện thời được hiểu là bao gồm, ví dụ, lớp cơ sở, lớp tăng cường thứ nhất, lớp tăng cường thứ hai, v.v., lên đến lớp hiện thời. Ở đó các lớp lên đến lớp hiện thời được hiểu bao gồm, ví dụ, lớp cơ sở, lớp tăng cường thứ nhất, lớp tăng cường thứ hai, v.v., lên đến lớp hiện thời. Ví dụ, bộ giải mã sẽ được kích hoạt để tăng cường dạng biểu diễn âm thanh tái tạo (một phần) thu được từ lớp cơ sở, liên quan đến tải tin mở rộng HOA gán cho lớp cơ sở. Theo phương thức thông thường, chỉ dạng biểu diễn HOA tái tạo của lớp tăng cường cao nhất có thể được tăng cường bởi tải tin mở rộng HOA. Do đó, bất kể lớp dùng được cao nhất thực tế (ví dụ, lớp ở dưới lớp dưới cùng không được thu hợp lệ, sao cho tất cả các lớp bên dưới lớp dùng được cao nhất và chính lớp dùng được cao nhất đã được thu hợp lệ), bộ giải mã cho phép cải thiện hoặc tăng cường dạng biểu diễn âm thanh tái tạo, ngay cả khi dạng biểu diễn âm thanh tái tạo (một phần) có thể khác với dạng biểu diễn âm thanh hoàn chỉnh (ví dụ, đầy đủ). Cụ thể, bất kể lớp dùng được cao nhất thực tế, bộ giải mã chỉ cần giải mã tải tin mở rộng HOA chỉ cho một lớp duy nhất (tức là, cho lớp dùng được cao nhất) để cải thiện hoặc tăng cường dạng biểu diễn âm thanh tái tạo (một phần) có thể thu được trên cơ sở các tín hiệu truyền có trong các lớp lên đến lớp dùng được cao nhất thực tế. Việc giải mã các tải tin mở rộng HOA của lớp cao hơn hoặc thấp hơn là không cần thiết. Mặt khác, phương pháp được đề xuất cho phép tận dụng hoàn toàn sự giảm băng thông cần thiết mà có thể đạt được khi áp dụng mã hóa theo lớp.

Theo các phương án, phương pháp có thể còn bao gồm bước truyền các tải tin dữ liệu cho nhiều lớp với các mức độ bảo vệ chống lỗi tương ứng. Các tải tin dữ liệu có thể bao gồm các tải tin mở rộng HOA tương ứng. Lớp cơ sở có thể có mức độ bảo vệ chống lỗi cao nhất và một hoặc nhiều lớp tăng cường có thể có mức độ bảo vệ chống lỗi giảm liên tiếp. Theo cách đó, có thể bảo đảm rằng ít nhất một số lớp thấp hơn chắc chắn được truyền, trong khi đó mặt khác làm giảm băng thông tổng thể cần thiết bằng cách không áp dụng sự bảo vệ chống lỗi quá mức lên các lớp cao hơn.

Theo các phương án, các tải tin mở rộng HOA có thể bao gồm các phần tử dòng bit cho công cụ giải mã dự báo tín hiệu không gian HOA. Ngoài ra hoặc theo cách khác, các tải tin mở rộng HOA có thể bao gồm các phần tử dòng bit cho công cụ giải mã tổng hợp tín hiệu định hướng con HOA. Ngoài ra hoặc theo cách khác, các tải tin mở rộng HOA có thể bao gồm các phần tử dòng bit cho công cụ giải mã tái tạo môi trường tham số HOA.

Theo các phương án, các tải tin mở rộng HOA có thể có `usacExtElementType` của `ID_EXT_ELE_HOA_ENH_LAYER`.

Theo các phương án, phương pháp có thể còn bao gồm bước tạo tải tin mở rộng cấu hình HOA bao gồm các phần tử dòng bit để tạo cấu hình công cụ giải mã dự báo tín hiệu không gian HOA, công cụ giải mã tổng hợp tín hiệu định hướng dải con HOA, và/hoặc công cụ giải mã tái tạo môi trường tham số HOA. Tải tin mở rộng cấu hình HOA có thể được bao gồm trong `HOADecoderEnhConfig()`. Phương pháp có thể còn bao gồm bước báo hiệu tải tin mở rộng cấu hình HOA trong dòng bit đầu ra.

Theo các phương án, phương pháp có thể còn bao gồm bước tạo tải tin cấu hình bộ giải mã HOA bao gồm thông tin chỉ báo việc gán các tải tin mở rộng HOA cho nhiều lớp. Phương pháp có thể còn bao gồm bước báo hiệu tải tin cấu hình bộ giải mã HOA trong dòng bit đầu ra.

Theo các phương án, phương pháp có thể còn bao gồm bước xác định liệu chế độ mã hóa vector có hoạt động không. Phương pháp có thể còn bao gồm, nếu chế độ mã hóa vector hoạt động, xác định, đối với mỗi lớp, tập chỉ số hệ số HOA liên tiếp trên cơ sở các tín hiệu truyền được gán cho lớp tương ứng. Các chỉ số hệ số HOA trong tập chỉ số hệ số HOA liên tiếp có thể là các chỉ số hệ số HOA bao gồm trong tập `ContAddHOACoeff`. Phương pháp có thể còn bao gồm bước tạo, đối với mỗi tín hiệu truyền, vector V trên cơ sở tập chỉ số hệ số HOA liên tiếp xác định được đối với lớp mà

tín hiệu truyền tương ứng được gán vào đó, sao cho vector V được tạo ra bao gồm các phần tử cho các tín hiệu truyền bất kỳ gán cho các lớp cao hơn lớp mà tín hiệu truyền tương ứng được gán vào đó. Phương pháp có thể còn bao gồm bước báo hiệu các vector V được tạo ra trong dòng bit đầu ra.

Theo một khía cạnh khác, phương pháp mã hóa theo lớp đối với khung của dạng biểu diễn Ambisonics bậc cao, HOA, nén của âm thanh hoặc trường âm thanh được mô tả. Dạng biểu diễn HOA nén có thể bao gồm nhiều tín hiệu truyền. Các tín hiệu truyền có thể liên quan đến tín hiệu nghe một tai, ví dụ, biểu diễn tín hiệu âm thanh nổi trội hoặc chuỗi hệ số của dạng biểu diễn HOA. Phương pháp có thể bao gồm bước gán nhiều tín hiệu truyền cho nhiều lớp phân cấp. Ví dụ, các tín hiệu truyền có thể được phân bổ cho nhiều lớp. Các lớp có thể bao gồm một lớp cơ sở và một hoặc nhiều lớp tăng cường phân cấp. Phương pháp có thể còn bao gồm bước xác định liệu chế độ mã hóa vector có hoạt động không. Phương pháp có thể còn bao gồm, nếu chế độ mã hóa vector hoạt động, xác định, đối với mỗi lớp, tập chỉ số hệ số HOA liên tiếp trên cơ sở các tín hiệu truyền được gán cho lớp tương ứng. Các chỉ số hệ số HOA trong tập chỉ số hệ số HOA liên tiếp có thể là các chỉ số hệ số HOA bao gồm trong tập `ContAddHOACoeff`. Phương pháp có thể còn bao gồm bước tạo, đối với mỗi tín hiệu truyền, vector V trên cơ sở tập chỉ số hệ số HOA liên tiếp xác định được đối với lớp mà tín hiệu truyền tương ứng được gán vào đó, sao cho vector V được tạo ra bao gồm các phần tử cho các tín hiệu truyền bất kỳ gán cho các lớp cao hơn lớp mà tín hiệu truyền tương ứng được gán vào đó. Phương pháp có thể còn bao gồm bước báo hiệu các vector V được tạo ra trong dòng bit đầu ra.

Với cấu hình như vậy, phương pháp được đề xuất bảo đảm rằng trong chế độ mã hóa vector vector V thích hợp là khả dụng cho mọi tín hiệu truyền thuộc các lớp lên đến lớp dùng được cao nhất. Cụ thể, phương pháp được đề xuất loại trừ trường hợp các phần tử của vector V tương ứng với các tín hiệu truyền ở các lớp cao hơn không được báo hiệu rõ ràng. Theo đó, thông tin bao gồm trong các lớp lên đến lớp dùng được cao nhất là đủ để giải mã các tín hiệu truyền bất kỳ thuộc các lớp lên đến lớp dùng được cao nhất. Theo cách đó, có sự giải nén thích hợp cho các dạng biểu diễn HOA tái tạo tương ứng đối với các lớp thấp hơn (các lớp có tốc độ bit thấp) ngay cả nếu các lớp cao hơn có thể không được thu hợp lệ bởi bộ giải mã. Mặt khác, phương pháp được đề xuất

cho phép tận dụng hoàn toàn sự giảm băng thông cần thiết mà có thể đạt được khi áp dụng mã hóa theo lớp.

Theo một khía cạnh khác, phương pháp giải mã khung của dạng biểu diễn Ambisonics bậc cao, HOA, nén của âm thanh hoặc trường âm thanh, được mô tả. Dạng biểu diễn HOA nén có thể được mã hóa trong nhiều lớp phân cấp. Các lớp phân cấp có thể bao gồm một lớp cơ sở và một hoặc nhiều lớp tăng cường phân cấp. Phương pháp có thể bao gồm bước nhận dòng bit liên quan đến khung của dạng biểu diễn HOA nén. Phương pháp có thể còn bao gồm bước trích xuất các tải tin cho nhiều lớp. Mỗi tải tin có thể bao gồm các tín hiệu truyền gán cho lớp tương ứng. Phương pháp có thể còn bao gồm bước xác định lớp dùng được cao nhất trong số nhiều lớp để giải mã. Phương pháp có thể còn bao gồm trích xuất tải tin mở rộng HOA gán cho lớp dùng được cao nhất. Tải tin mở rộng HOA này có thể bao gồm thông tin phụ trợ để tăng cường tham số cho dạng biểu diễn HOA tái tạo (một phần) tương ứng với lớp dùng được cao nhất. Dạng biểu diễn HOA tái tạo (một phần) tương ứng với lớp dùng được cao nhất có thể thu được trên cơ sở các tín hiệu truyền gán cho lớp dùng được cao nhất và các lớp bất kỳ thấp hơn lớp dùng được cao nhất. Phương pháp có thể còn bao gồm việc tạo dạng biểu diễn HOA tái tạo (một phần) tương ứng với lớp dùng được cao nhất trên cơ sở các tín hiệu truyền gán cho lớp dùng được cao nhất và các lớp bất kỳ thấp hơn lớp dùng được cao nhất. Phương pháp có thể còn bao gồm bước tăng cường (ví dụ, tăng cường tham số) dạng biểu diễn HOA tái tạo (một phần) nhờ sử dụng thông tin phụ trợ bao gồm trong tải tin mở rộng HOA gán cho lớp dùng được cao nhất. Kết quả là, có thể thu được dạng biểu diễn HOA tái tạo tăng cường.

Với cấu hình như vậy, phương pháp được đề xuất bảo đảm rằng dạng biểu diễn HOA tái tạo (ví dụ, được tăng cường) cuối cùng có chất lượng tối ưu, nhờ sử dụng thông tin sẵn có (ví dụ, nhận được hợp lệ) ở mức độ tốt nhất có thể.

Theo các phương án, các tải tin mở rộng HOA có thể bao gồm các phần tử dòng bit cho công cụ giải mã dự báo tín hiệu không gian HOA. Ngoài ra hoặc theo cách khác, các tải tin mở rộng HOA có thể bao gồm các phần tử dòng bit cho công cụ giải mã tổng hợp tín hiệu định hướng dải con HOA. Ngoài ra hoặc theo cách khác, các tải tin mở rộng HOA có thể bao gồm các phần tử dòng bit cho công cụ giải mã tái tạo môi trường tham số HOA.

Theo các phương án, các tải tin mở rộng HOA có thể có usacExtElementType của ID_EXT_ELE_HOA_ENH_LAYER.

Theo các phương án, phương pháp có thể còn bao gồm bước trích xuất tải tin mở rộng cấu hình HOA bằng cách phân tích cú pháp dòng bit. Tải tin mở rộng cấu hình HOA có thể bao gồm các phần tử dòng bit để tạo cấu hình công cụ giải mã dự báo tín hiệu không gian HOA, công cụ giải mã tổng hợp tín hiệu định hướng dải con HOA, và/hoặc công cụ giải mã tái tạo môi trường tham số HOA.

Theo các phương án, phương pháp có thể còn bao gồm trích xuất các tải tin mở rộng HOA lần lượt được gán cho nhiều lớp. Mỗi tải tin mở rộng HOA có thể bao gồm thông tin phụ trợ để tăng cường tham số cho dạng biểu diễn HOA tái tạo (một phần) tương ứng với lớp được gán tương ứng của nó. Dạng biểu diễn HOA tái tạo (một phần) tương ứng với lớp được gán tương ứng của nó có thể thu được từ các tín hiệu truyền gán cho lớp đó và các lớp bất kỳ thấp hơn lớp đó. Việc gán các tải tin mở rộng HOA cho các lớp tương ứng có thể được biết từ thông tin cấu hình bao gồm trong dòng bit.

Theo các phương án, bước xác định lớp dùng được cao nhất có thể bao gồm bước xác định tập chỉ số lớp không hợp lệ chỉ báo các lớp mà không được thu hợp lệ. Nó có thể còn bao gồm bước xác định lớp dùng được cao nhất là lớp mà thấp hơn một lớp so với lớp được chỉ báo bởi chỉ số nhỏ nhất (thấp nhất) trong tập chỉ số lớp không hợp lệ. Lớp cơ sở có thể có chỉ số lớp thấp nhất (ví dụ, chỉ số lớp là 1), và các lớp tăng cường phân cấp có thể có các chỉ số lớp cao hơn liên tiếp. Theo cách đó, phương pháp được đề xuất bảo đảm rằng lớp dùng được cao nhất được chọn theo cách sao cho tất cả thông tin cần thiết để giải mã dạng biểu diễn HOA tái tạo (một phần) từ các lớp dùng được cao nhất và các lớp bất kỳ dưới lớp dùng được cao nhất là khả dụng.

Theo các phương án, bước xác định lớp dùng được cao nhất có thể bao gồm bước xác định tập chỉ số lớp không hợp lệ chỉ báo các lớp mà không được thu hợp lệ. Nó có thể còn bao gồm việc xác định lớp dùng được cao nhất của khung trước ở trước khung hiện thời. Nó có thể còn bao gồm việc xác định lớp dùng được cao nhất làm lớp thấp hơn của lớp dùng được cao nhất của khung trước và lớp mà thấp hơn một lớp so với lớp được chỉ báo bởi chỉ số nhỏ nhất trong tập chỉ số lớp không hợp lệ. Theo cách đó, lớp dùng được cao nhất đối với khung hiện thời được chọn theo cách sao cho tất cả thông tin cần thiết để giải mã dạng biểu diễn HOA tái tạo (một phần) từ lớp dùng được

cao nhất và các lớp bất kỳ dưới lớp dùng được cao nhất là khả dụng, cho dù khung hiện thời đã được mã hóa khác biệt so với khung trước.

Theo các phương án, phương pháp có thể còn bao gồm bước quyết định không thực hiện tăng cường tham số cho dạng biểu diễn HOA tái tạo (một phần) nhờ sử dụng thông tin phụ trợ bao gồm trong tải tin mở rộng HOA gán cho lớp dùng được cao nhất nếu lớp dùng được cao nhất của khung hiện thời là thấp hơn lớp dùng được cao nhất của khung trước và nếu khung hiện thời được mã hóa khác biệt so với khung trước. Theo cách đó, dạng biểu diễn HOA tái tạo có thể được giải mã không có lỗi trong các trường hợp trong đó khung hiện thời (bao gồm thông tin phụ trợ bao gồm trong tải tin mở rộng HOA gán cho lớp dùng được cao nhất) đã được mã hóa khác biệt so với khung trước.

Theo các phương án, tập chỉ số lớp không hợp lệ có thể được xác định bằng cách đánh giá cờ hợp lệ của các tải tin mở rộng HOA tương ứng. Chỉ số lớp của lớp cho trước có thể được bổ sung vào tập chỉ số lớp không hợp lệ nếu cờ hợp lệ cho tải tin mở rộng HOA được gán cho lớp tương ứng không được thiết lập. Theo cách đó, tập chỉ số lớp không hợp lệ có thể được xác định một cách hiệu quả.

Theo một khía cạnh khác, cấu trúc dữ liệu (ví dụ, dòng bit) biểu diễn khung của dạng biểu diễn Ambisonics bậc cao, HOA, nén của âm thanh hoặc trường âm thanh được mô tả. Dạng biểu diễn HOA nén có thể bao gồm nhiều tín hiệu truyền. Cấu trúc dữ liệu có thể bao gồm nhiều tải tin khung HOA tương ứng với các lớp tương ứng trong số nhiều lớp phân cấp. Tải tin khung HOA có thể bao gồm các tín hiệu truyền tương ứng. Nhiều tín hiệu truyền có thể được gán (ví dụ, phân bổ) cho nhiều lớp. Các lớp có thể bao gồm một lớp cơ sở và một hoặc nhiều lớp tăng cường phân cấp. Cấu trúc dữ liệu có thể còn bao gồm, đối với mỗi lớp, tải tin mở rộng HOA tương ứng bao gồm thông tin phụ trợ để tăng cường tham số cho dạng biểu diễn HOA tái tạo (một phần) thu được từ các tín hiệu truyền được gán cho lớp tương ứng và các lớp bất kỳ thấp hơn lớp tương ứng này.

Theo các phương án, các tải tin khung HOA và tải tin mở rộng HOA cho nhiều lớp có thể được tạo ra với các mức độ bảo vệ chống lỗi tương ứng. Lớp cơ sở có thể có mức độ bảo vệ chống lỗi cao nhất và một hoặc nhiều lớp tăng cường có thể có mức độ bảo vệ chống lỗi giảm liên tiếp.

Theo các phương án, các tải tin mở rộng HOA có thể bao gồm các phần tử dòng bit cho công cụ giải mã dự báo tín hiệu không gian HOA. Ngoài ra hoặc theo cách khác, các tải tin mở rộng HOA có thể bao gồm các phần tử dòng bit cho công cụ giải mã tổng hợp tín hiệu định hướng dải con HOA. Ngoài ra hoặc theo cách khác, các tải tin mở rộng HOA có thể bao gồm các phần tử dòng bit cho công cụ giải mã tái tạo môi trường tham số HOA.

Theo các phương án, các tải tin mở rộng HOA có thể có `usacExtElementType` của `ID_EXT_ELE_HOA_ENH_LAYER`.

Theo các phương án, cấu trúc dữ liệu có thể còn bao gồm tải tin mở rộng cấu hình HOA bao gồm các phần tử dòng bit để tạo cấu hình công cụ giải mã dự báo tín hiệu không gian HOA, công cụ giải mã tổng hợp tín hiệu định hướng dải con HOA, và/hoặc công cụ giải mã tái tạo môi trường tham số HOA.

Theo các phương án, cấu trúc dữ liệu có thể còn bao gồm tải tin cấu hình bộ giải mã HOA bao gồm thông tin chỉ báo việc gán các tải tin mở rộng HOA cho nhiều lớp.

Theo các phương án, sáng chế đề cập đến các phương pháp và thiết bị giải mã dạng biểu diễn Ambisonics bậc cao (HOA) nén của âm thanh hoặc trường âm thanh. Thiết bị này có thể được tạo cấu hình để hoặc phương pháp có thể bao gồm bước thu dòng bit chứa dạng biểu diễn HOA nén tương ứng với nhiều lớp phân cấp mà bao gồm lớp cơ sở và một hoặc nhiều lớp tăng cường phân cấp, trong đó nhiều lớp đã gán vào đó các thành phần của dạng biểu diễn âm thanh nén cơ bản của âm thanh hoặc trường âm thanh, các thành phần được gán cho các lớp tương ứng trong các nhóm thành phần tương ứng, xác định lớp dùng được cao nhất trong số nhiều lớp để giải mã; trích xuất tải tin mở rộng HOA gán cho lớp dùng được cao nhất, trong đó tải tin mở rộng HOA bao gồm thông tin phụ trợ để tăng cường tham số cho dạng biểu diễn HOA tái tạo tương ứng với lớp dùng được cao nhất, trong đó dạng biểu diễn HOA tái tạo tương ứng với lớp dùng được cao nhất là thu được trên cơ sở các tín hiệu truyền gán cho lớp dùng được cao nhất và các lớp bất kỳ thấp hơn lớp dùng được cao nhất; giải mã dạng biểu diễn HOA nén tương ứng với lớp dùng được cao nhất dựa trên thông tin lớp, các tín hiệu truyền gán cho lớp dùng được cao nhất và các lớp bất kỳ thấp hơn lớp dùng được cao nhất; và tăng cường tham số cho dạng biểu diễn HOA mã hóa nhờ sử dụng thông tin phụ trợ bao gồm trong tải tin mở rộng HOA gán cho lớp dùng được cao nhất.

Tải tin mở rộng HOA có thể bao gồm các phần tử dòng bit cho công cụ giải mã dự báo tín hiệu không gian HOA. Thông tin lớp có thể chỉ báo một số tín hiệu định hướng hoạt động trong khung hiện thời của lớp tăng cường.

Thông tin lớp có thể chỉ báo tổng số hệ số HOA xung quanh bổ sung cho lớp tăng cường. Thông tin lớp có thể bao gồm các chỉ số hệ số HOA cho mỗi hệ số HOA xung quanh bổ sung cho lớp tăng cường. Thông tin lớp có thể bao gồm thông tin tăng cường mà bao gồm ít nhất một trong số Dự báo tín hiệu không gian, Tổng hợp tín hiệu định hướng dải con và Bộ giải mã tái tạo môi trường tham số. Dạng biểu diễn HOA nén được làm thích ứng cho chế độ mã hóa theo lớp cho nội dung dựa trên HOA nếu `CodedVVecLength` bằng một được báo hiệu trong `HOADecoderConfig()`. Hơn nữa, các phần tử vector V có thể không được truyền đối với các chỉ số mà bằng các chỉ số của các hệ số HOA bổ sung bao gồm trong tập `ContAddHoaCoeff`. Tập `ContAddHoaCoeff` có thể được định nghĩa riêng biệt cho mỗi trong số nhiều lớp phân cấp. Thông tin lớp bao gồm các phần tử `NumLayers`, trong đó mỗi phần tử chỉ báo số lượng các tín hiệu truyền bao gồm trong tất cả các lớp lên đến lớp i . Thông tin lớp có thể bao gồm chỉ báo của tất cả các lớp được sử dụng thực tế cho khung thứ k . Thông tin lớp có thể còn chỉ báo rằng tất cả các hệ số cho vector nổi trội được chỉ rõ. Thông tin lớp có thể chỉ báo rằng các hệ số của các vector nổi trội tương ứng với số lượng lớn hơn `MinNumOfCoeffsForAmbHOA` được chỉ rõ. Thông tin lớp có thể chỉ báo rằng `MinNumOfCoeffsForAmbHOA` và tất cả các phần tử được xác định trong `ContAddHoaCoeff[lay]` không được truyền, trong đó `lay` là chỉ số của lớp chứa tín hiệu dựa trên vector tương ứng với vector này.

Theo một khía cạnh khác, bộ mã hóa để mã hóa theo lớp đối với khung của dạng biểu diễn Ambisonics bậc cao, HOA, nén của âm thanh hoặc trường âm thanh được mô tả. Dạng biểu diễn HOA nén có thể bao gồm nhiều tín hiệu truyền. Bộ mã hóa có thể bao gồm bộ xử lý được tạo cấu hình để thực hiện một số hoặc tất cả các bước phương pháp của các phương pháp theo khía cạnh được đề cập thứ nhất và khía cạnh được đề cập thứ hai nêu trên.

Theo một khía cạnh khác, bộ giải mã để giải mã khung của dạng biểu diễn Ambisonics bậc cao, HOA, nén của âm thanh hoặc trường âm thanh, được mô tả. Dạng biểu diễn HOA nén có thể được mã hóa trong nhiều lớp phân cấp mà bao gồm một lớp cơ sở và một hoặc nhiều lớp tăng cường phân cấp. Bộ giải mã có thể bao gồm bộ xử lý

được tạo cấu hình để thực hiện một số hoặc tất cả các bước phương pháp của các phương pháp theo theo khía cạnh được đề cập thứ ba nêu trên.

Theo một khía cạnh khác, sáng chế mô tả chương trình phần mềm. Chương trình phần mềm này có thể được làm thích ứng để thực thi trên bộ xử lý và để thực hiện một số hoặc tất cả các bước của phương pháp đã được nêu trong tài liệu này khi được tiến hành trên thiết bị máy tính.

Theo một khía cạnh khác nữa, sáng chế mô tả vật ghi. Vật ghi này có thể bao gồm chương trình phần mềm được làm thích ứng để thực thi trên bộ xử lý và để thực hiện một số hoặc tất cả các bước của phương pháp đã được nêu trong tài liệu này khi được tiến hành trên thiết bị máy tính.

Người có hiểu biết trung bình trong lĩnh vực hiểu rõ rằng các nội dung trình bày liên quan đến khía cạnh bất kỳ trong số các khía cạnh nêu trên hoặc các phương án của nó cũng áp dụng cho các khía cạnh khác tương ứng hoặc các phương án của chúng. Việc lặp lại các nội dung này cho mỗi và mọi khía cạnh hoặc phương án được lược bỏ để không dài dòng.

Cần lưu ý rằng các phương pháp và thiết bị bao gồm các phương án ưu tiên như được nêu trong tài liệu này có thể được sử dụng riêng rẽ hoặc kết hợp với các phương pháp và hệ thống khác được bộc lộ trong tài liệu này. Hơn thế nữa, tất cả các khía cạnh của phương pháp và thiết bị được nêu trong tài liệu này có thể được kết hợp tùy ý. Cụ thể, các dấu hiệu trong yêu cầu bảo hộ có thể được kết hợp với nhau theo cách bất kỳ.

Còn cần lưu ý rằng các bước của phương pháp và các dấu hiệu của thiết bị có thể thay thế cho nhau theo nhiều cách. Cụ thể, người có hiểu biết trung bình trong lĩnh vực hiểu rõ rằng các chi tiết của phương pháp được bộc lộ có thể được thực hiện dưới dạng thiết bị được làm thích ứng để thực hiện một số hoặc tất cả các bước của phương pháp, và ngược lại.

Mô tả vắn tắt các hình vẽ

Sáng chế được giải thích sau đây bằng cách minh họa và tham khảo hình vẽ đính kèm, trong đó:

Fig.1 là sơ đồ khối thể hiện gián lược việc gán các tải tin cho lớp cơ sở và $M - 1$ lớp tăng cường ở phía bộ mã hóa;

Fig.2 là sơ đồ khối thể hiện gián lược ví dụ về bộ thu và giai đoạn giải nén;

Fig.3 là lưu đồ thể hiện ví dụ về phương pháp mã hóa theo lớp đối với khung của dạng biểu diễn HOA nén theo các phương án của sáng chế;

Fig.4 là lưu đồ thể hiện ví dụ khác về phương pháp mã hóa theo lớp đối với khung của dạng biểu diễn HOA nén theo các phương án của sáng chế;

Fig.5 là lưu đồ thể hiện ví dụ về phương pháp giải mã khung của dạng biểu diễn HOA nén theo các phương án của sáng chế;

Fig.6 là sơ đồ khối thể hiện giản lược ví dụ về phương án thực hiện phần cứng của bộ mã hóa theo các phương án của sáng chế; và

Fig.7 là sơ đồ khối thể hiện giản lược ví dụ về phương án thực hiện phần cứng của bộ giải mã theo các phương án của sáng chế.

Mô tả chi tiết sáng chế

Đầu tiên, sáng chế mô tả dạng biểu diễn âm thanh (hoặc trường âm thanh) nén mà các phương pháp và bộ mã hóa/bộ giải mã theo sáng chế có thể áp dụng.

Để tạo dòng dạng biểu diễn âm thanh (hoặc trường âm thanh) nén qua kênh truyền với các điều kiện thay đổi về thời gian, mã hóa theo lớp là phương tiện để làm thích ứng chất lượng của dạng biểu diễn âm thanh thu được theo các điều kiện truyền, và cụ thể là để tránh các trường hợp mất tín hiệu không mong muốn.

Đối với mã hóa theo lớp, các dạng biểu diễn âm thanh (hoặc trường âm thanh) nén thường được chia nhỏ thành lớp cơ sở có mức độ ưu tiên cao có kích thước tương đối nhỏ và các lớp tăng cường bổ sung có mức độ ưu tiên giảm và kích thước tùy chọn. Mỗi lớp tăng cường thường được giả định chứa thêm nhiều thông tin để bù cho thông tin của tất cả các lớp thấp hơn nhằm cải thiện chất lượng của dạng biểu diễn âm thanh (hoặc trường âm thanh) nén. Sau đó, ý tưởng là kiểm soát mức độ bảo vệ chống lỗi cho quá trình truyền các lớp riêng lẻ theo mức độ ưu tiên của chúng. Cụ thể là, lớp cơ sở có độ bảo vệ chống lỗi cao, điều này cũng hợp lý và có thể chấp nhận được do lớp này có kích thước nhỏ.

Sau đây giả định rằng dạng biểu diễn âm thanh (hoặc trường âm thanh) nén hoàn chỉnh thường gồm có ba thành phần sau:

1. Dạng biểu diễn âm thanh (hoặc trường âm thanh) nén cơ bản gồm có một số thành phần bù, chiếm tỷ lệ phần trăm lớn nhất rõ ràng của dạng biểu diễn âm thanh (hoặc trường âm thanh) nén hoàn chỉnh.

2. Thông tin phụ trợ cơ bản cần thiết để giải mã dạng biểu diễn âm thanh nén cơ bản, mà được giả định là có kích thước nhỏ hơn nhiều so với dạng biểu diễn âm thanh (hoặc trường âm thanh) nén cơ bản. Nó còn được giả định bao gồm trong phần lớn nhất của nó là hai thành phần sau, cả hai đều nêu rõ sự giải nén của chỉ một thành phần cụ thể của dạng biểu diễn âm thanh nén cơ bản:

a) Thành phần thứ nhất chứa thông tin phụ trợ mô tả từng thành phần bù riêng của dạng biểu diễn âm thanh (hoặc trường âm thanh) nén cơ bản *độc lập với các thành phần bù khác*.

b) Thành phần thứ hai (tùy ý) chứa thông tin phụ trợ mô tả từng thành phần bù riêng của dạng biểu diễn âm thanh (hoặc trường âm thanh) nén cơ bản *phụ thuộc vào các thành phần bù khác*. Cụ thể, sự phụ thuộc này có các thuộc tính sau:

- Thông tin phụ trợ phụ thuộc đối với mỗi thành phần bù riêng của dạng biểu diễn âm thanh (hoặc trường âm thanh) nén cơ bản đạt được mức độ cao nhất của nó trong trường hợp không chứa các thành phần bù nhất định khác trong dạng biểu diễn âm thanh (hoặc trường âm thanh) nén cơ bản.

- Trong trường hợp các thành phần bù nhất định khác được thêm vào dạng biểu diễn âm thanh (hoặc trường âm thanh) nén cơ bản, thông tin phụ trợ phụ thuộc đối với thành phần bù riêng được xem xét trở thành tập con của thông tin phụ trợ ban đầu, nhờ đó làm giảm kích thước của nó.

3. Thông tin phụ trợ tăng cường tùy ý để cải thiện dạng biểu diễn âm thanh (hoặc trường âm thanh) nén cơ bản. Kích thước của nó cũng được giả định là nhỏ hơn nhiều so với kích thước của dạng biểu diễn âm thanh (hoặc trường âm thanh) nén cơ bản.

Một ví dụ nổi bật về loại dạng biểu diễn âm thanh (hoặc trường âm thanh) nén hoàn chỉnh như vậy được đưa ra bởi dạng biểu diễn trường âm thanh HOA nén như được quy định bởi phiên bản sơ bộ của chuẩn âm thanh 3D MPEG-H.

1. Dạng biểu diễn trường âm thanh nén cơ bản của nó có thể được nhận dạng với một số tín hiệu nghe một tai lượng tử hóa, biểu diễn các tín hiệu âm thanh nổi bật hoặc chuỗi hệ số của thành phần trường âm thanh HOA xung

quanh.

2. Thông tin phụ trợ cơ bản mô tả, bên cạnh những cái khác, đối với mỗi tín hiệu nghe một tai này, sự đóng góp về không gian của nó như thế nào cho trường âm thanh. Thông tin này có thể còn được tách thành hai thành phần khác nhau như sau:

(a) Thông tin phụ trợ liên quan đến các tín hiệu nghe một tai riêng cụ thể, mà *độc lập* với sự tồn tại của các tín hiệu nghe một tai khác. Thông tin phụ trợ này có thể, chẳng hạn, mô tả tín hiệu nghe một tai để biểu diễn tín hiệu định hướng (nghĩa là sóng phẳng chung) với hướng tới nhất định. Theo cách khác, tín hiệu nghe một tai có thể được định rõ là chuỗi hệ số của dạng biểu diễn HOA ban đầu có chỉ số nhất định.

(b) Thông tin phụ trợ liên quan đến các tín hiệu nghe một tai riêng cụ thể, mà *phụ thuộc vào* sự tồn tại của các tín hiệu nghe một tai khác. Thông tin phụ trợ này xuất hiện ví dụ, nếu các tín hiệu nghe một tai được chỉ rõ là các tín hiệu dựa trên vector, nghĩa là chúng được phân bổ theo hướng trong trường âm thanh, trong đó sự phân bổ hướng được định rõ nhờ vector. Theo chế độ nhất định (tức là $\text{CodedVVecLength} = 1$), các thành phần cụ thể của vector này được ngầm thiết lập về không và không phải là một phần của dạng biểu diễn vector nén. Các thành phần này là các thành phần có các chỉ số bằng các chỉ số của chuỗi hệ số của dạng biểu diễn HOA ban đầu, mà là một phần của dạng biểu diễn trường âm thanh nén cơ bản. Nghĩa là nếu các thành phần riêng rẽ của vector được mã hóa, thì tổng số của chúng phụ thuộc vào dạng biểu diễn trường âm thanh nén cơ bản, cụ thể là phụ thuộc vào chuỗi hệ số nào của dạng biểu diễn HOA ban đầu mà nó chứa.

Nếu không có chuỗi hệ số của dạng biểu diễn HOA ban đầu được chứa trong dạng biểu diễn trường âm thanh nén cơ bản, thông tin phụ trợ cơ bản phụ thuộc cho mỗi tín hiệu dựa trên vector bao gồm tất cả các thành phần vector và có kích thước lớn nhất của nó. Trong trường hợp mà chuỗi hệ số của dạng biểu diễn HOA ban đầu có các chỉ số nhất định được thêm vào dạng biểu diễn trường âm thanh nén cơ bản, các thành

phần vector có các chỉ số đó được loại bỏ khỏi thông tin phụ trợ cho mỗi tín hiệu dựa trên vector, nhờ đó làm giảm kích thước của thông tin phụ trợ cơ bản phụ thuộc đối với các tín hiệu dựa trên vector.

3. Thông tin phụ trợ tăng cường bao gồm các thành phần sau:

- Các tham số liên quan đến dự báo không gian (dải rộng) để dự báo (tuyến tính) các phần thiếu của trường âm thanh từ các tín hiệu định hướng.

- Các tham số liên quan đến tổng hợp tín hiệu định hướng dải con và tái tạo môi trường tham số, mà là các công cụ nén cho phép sự dự báo tham số, phụ thuộc tần số cho các tín hiệu nghe một tai bổ sung cần phân bổ theo không gian để bù cho dạng biểu diễn HOA nén khuyết hoặc không hoàn chỉnh về không gian. Dự báo này là dựa trên chuỗi hệ số của dạng biểu diễn trường âm thanh nén cơ bản. Một khía cạnh quan trọng là sự đóng góp bù được đề cập vào trường âm thanh được biểu diễn trong dạng biểu diễn HOA nén không phải nhờ các tín hiệu lượng tử hóa bổ sung, mà là nhờ thông tin phụ trợ bổ sung có kích thước nhỏ hơn rất nhiều. Vì thế, hai công cụ mã hóa được đề cập đặc biệt thích hợp cho việc nén dạng biểu diễn HOA ở tốc độ dữ liệu thấp.

Ví dụ thứ hai về dạng biểu diễn nén của tín hiệu nghe một tai có cấu trúc nêu trên có thể gồm các thành phần sau:

1. Một số thông tin phổ được mã hóa cho các dải tần số rời lên đến tần số cao hơn nhất định, mà có thể được xem là dạng biểu diễn nén cơ bản.

2. Một số thông tin phụ trợ cơ bản nêu rõ thông tin phổ được mã hóa (bởi, ví dụ, số lượng và độ rộng của các dải tần số mã hóa).

3. Một số thông tin phụ trợ tăng cường gồm có các tham số của Tái tạo dải phổ (Spectral Band Replication - SBR), mô tả cách tái tạo tham số từ dạng biểu diễn nén cơ bản thông tin phổ cho các dải tần số cao hơn mà không được xem xét trong dạng biểu diễn nén cơ bản.

Tiếp theo, phương pháp mã hóa theo lớp dạng biểu diễn âm thanh (hoặc trường âm thanh) nén hoàn chỉnh có cấu trúc nêu trên sẽ được mô tả.

Giả định rằng sự nén này là dựa trên khung theo nghĩa là nó tạo ra các dạng biểu diễn nén (ví dụ, ở dạng gói dữ liệu hoặc các tải tin khung tương đương) trong các khoảng thời gian liên tiếp, ví dụ các khoảng thời gian có kích thước bằng nhau. Các gói dữ liệu này được giả định là chứa cờ hợp lệ, giá trị chỉ báo kích thước của chúng cũng như dữ liệu biểu diễn nén thực tế. Xuyên suốt phần mô tả sau sẽ tập trung hầu hết vào việc xử lý khung đơn, và vì vậy chỉ số khung sẽ được bỏ qua.

Mỗi tải tin khung của dạng biểu diễn âm thanh (hoặc trường âm thanh) nén hoàn chỉnh được xem xét 1100 được giả định chứa các gói dữ liệu J , mỗi gói cho một thành phần 1110-1, ..., 1110- J của dạng biểu diễn âm thanh (hoặc trường âm thanh) nén cơ bản, mà được biểu thị bằng $BSRC_j$, $j = 1, \dots, J$. Hơn nữa, tải tin này được giả định là chứa gói có thông tin phụ trợ cơ bản *độc lập* 1120 được biểu thị bằng cách BSI_1 chỉ rõ các thành phần cụ thể $BSRC_j$ của dạng biểu diễn âm thanh nén cơ bản độc lập với các thành phần khác. Tùy ý, ngoài ra tải tin này được còn được giả định là chứa gói có thông tin phụ trợ cơ bản *phụ thuộc* được biểu thị bằng cách BSI_D chỉ rõ các thành phần cụ thể $BSRC_j$ của dạng biểu diễn âm thanh nén cơ bản phụ thuộc vào các thành phần khác. Thông tin chứa trong hai gói dữ liệu BSI_1 và BSI_D có thể được nhóm tùy ý vào một gói dữ liệu đơn nhất BSI .

Cuối cùng, nó bao gồm tải tin thông tin phụ trợ tăng cường được biểu thị bằng ESI với sự mô tả cách để cải thiện âm thanh (hoặc trường âm thanh) tái tạo từ dạng biểu diễn nén cơ bản hoàn chỉnh.

Sơ đồ mô tả để mã hóa theo lớp chỉ ra các bước cần thiết để kích hoạt cả hai, phần nén bao gồm việc gói các gói dữ liệu để truyền cũng như bộ thu và phân giải nén. Mỗi phần sẽ được mô tả chi tiết trong phần sau.

Tiếp theo, việc nén và đóng gói để truyền sẽ được mô tả. Trong trường hợp mã hóa theo lớp (giả định tổng M lớp, tức là một lớp cơ sở và $M - 1$ lớp tăng cường) mỗi thành phần của dạng biểu diễn âm thanh (hoặc trường âm thanh) nén hoàn chỉnh 1100 được xử lý như sau:

- Dạng biểu diễn âm thanh (hoặc trường âm thanh) nén cơ bản được chia nhỏ thành các phần để được gán cho từng lớp riêng. Không làm mất tính tổng quát, việc tạo nhóm có thể được mô tả bằng $M + 1$ số J_m , $m = 0, \dots, M$ với $J_0 = 1$ và $J_M = J + 1$ sao cho $BSRC_j$ được gán vào lớp thứ m đối với $J_{m-1} \leq j < J_m$.

• Do kích thước nhỏ của nó, nên nó hợp lý để gán thông tin phụ trợ cơ bản hoàn chỉnh vào lớp cơ sở để tránh sự phân đoạn không cần thiết. Trong khi thông tin phụ trợ cơ bản độc lập BSI_I được giữ không thay đổi cho việc gán, thông tin phụ trợ cơ bản phụ thuộc phải được xử lý đặc biệt để mã hóa theo lớp, để cho phép, một mặt, giải mã chính xác ở phía bộ thu và mặt khác để làm giảm kích thước của thông tin phụ trợ phụ thuộc cần truyền. Đề xuất là phân tách nó thành M phần 1130-1, ..., 1130- M được biểu thị bằng $BSI_{D,m}$, $m = 1, \dots, M$, trong đó phần thứ m chứa thông tin phụ trợ phụ thuộc cho mỗi trong số các thành phần $BSRC_j$, $J_{m-1} \leq j < J_m$, của dạng biểu diễn âm thanh nén cơ bản gán cho lớp thứ m , nếu có tồn tại thông tin phụ trợ phụ thuộc tương ứng. Trong trường hợp thông tin phụ trợ phụ thuộc tương ứng không tồn tại, $BSI_{D,m}$ được giả định là rỗng. Thông tin phụ trợ $BSI_{D,m}$ phụ thuộc vào tất cả các thành phần $BSRC_j$, $1 \leq j < J_m$, chứa trong tất cả các lớp cho đến lớp thứ m .

• Trong trường hợp mã hóa theo lớp, quan trọng là hiểu rõ rằng thông tin phụ trợ tăng cường phải được tính toán đối với mỗi lớp thêm, vì dự định tăng cường âm thanh (hoặc trường âm thanh) giải nén sơ bộ, mà tuy nhiên phụ thuộc vào các lớp khả dụng để giải nén. Vì thế, việc nén phải tạo ra M gói dữ liệu thông tin phụ trợ tăng cường riêng 1140-1, ..., 1140- M , được biểu thị bằng ESI_m , $m = 1, \dots, M$, trong đó thông tin phụ trợ tăng cường trong gói dữ liệu thứ m ESI_m được tính toán để tăng cường dạng biểu diễn âm thanh (hoặc trường âm thanh) thu được từ tất cả dữ liệu chứa trong lớp cơ sở và các lớp tăng cường với các chỉ số thấp hơn m .

Tóm lại, ở giai đoạn nén, gói dữ liệu khung, được biểu thị bằng FRAME, phải được tạo ra có thành phần sau:

$$FRAME = [BSRC_1 \quad \dots \quad BSRC_J \quad BSI_I \quad BSI_{D,1} \quad \dots \quad BSI_{D,M} \quad ESI_1 \quad \dots \quad ESI_M]. \quad (1)$$

Hiểu rằng việc sắp thứ tự của các tải tin riêng rẽ với gói dữ liệu khung nhìn chung là tùy ý.

Việc gán đã mô tả của các tải tin riêng rẽ cho lớp cơ sở và lớp tăng cường được thực hiện bằng bộ đóng gói lớp truyền và được minh họa bằng sơ đồ trên Fig.1.

Tiếp theo, việc thu và giải nén sẽ được mô tả. Giải đoạn thu và giải nén tương ứng được minh họa trên Fig.2.

Thứ nhất, các gói lớp riêng rẽ 1200, 1300-1, ..., 1300-($M - 1$) được dồn kênh để tạo ra gói khung thu được

$$[BSI_1 \quad BSI_{D,1} \quad \dots \quad BSI_{D,M} \quad ESI_1 \quad BSRC_1 \quad \dots \quad BSRC_{(J_1)-1} \quad \dots \quad ESI_M \quad BSRC_{J(M-1)} \quad \dots \quad BSRC_J] \quad (2)$$

của dạng biểu diễn âm thanh (hoặc trường âm thanh) nén hoàn chỉnh, mà sau đó được chuyển qua bộ giải nén 2100. Giả định rằng nếu việc truyền lớp riêng rẽ không có lỗi, cờ hợp lệ của ít nhất tải tin thông tin phụ trợ tăng cường chứa trong đó được thiết lập là “đúng”. Trong trường hợp có lỗi do việc truyền lớp riêng rẽ, cờ hợp lệ trong ít nhất tải tin thông tin phụ trợ tăng cường trong lớp này được thiết lập là “sai”. Vì vậy, tính hợp lệ của gói lớp có thể được xác định từ tính hợp lệ của tải tin thông tin phụ trợ tăng cường chứa trong đó.

Ở bộ giải nén 2100, gói khung thu được được giải dồn kênh đầu tiên. Để đạt được mục đích này, thông tin về kích thước của mỗi tải tin có thể được khai thác để tránh việc phân tích cú pháp không cần thiết qua dữ liệu của các tải tin riêng rẽ.

Trong bước tiếp theo, số N_B của lớp cao nhất cần được sử dụng thực tế để giải nén cho dạng biểu diễn âm thanh cơ bản được chọn. Lớp *tăng cường* cao nhất cần được sử dụng thực tế để giải nén cho dạng biểu diễn âm thanh cơ bản được tạo ra bởi $N_B - 1$. Do mỗi lớp chứa chính xác một tải tin thông tin phụ trợ tăng cường, nên biết được qua mỗi tải tin thông tin phụ trợ tăng cường rằng lớp chứa có hợp lệ hay không. Vì vậy, việc lựa chọn có thể được thực hiện nhờ sử dụng tất cả tải tin thông tin phụ trợ tăng cường ESI_m , $m = 1, \dots, M$. Ngoài ra, chỉ số N_E của tải tin thông tin phụ trợ tăng cường cần sử dụng để giải nén được xác định, chỉ số luôn bằng N_B hoặc bằng không. Điều này nghĩa là sự tăng cường luôn được thực hiện theo dạng biểu diễn âm thanh cơ bản hoặc không. Dưới đây là phần mô tả chi tiết hơn về việc chọn này.

Lần lượt, các tải tin của các thành phần dạng biểu diễn âm thanh nén cơ bản $BSRC_1, SRcesC_J$ được chuyển cùng với tất cả các tải tin thông tin phụ trợ cơ bản (tức là BSI_1 và $BSI_{D,m}$, $m = 1, \dots, M$) và giá trị N_B đến bộ xử lý giải nén dạng biểu diễn cơ bản 2200, mà tái tạo dạng biểu diễn âm thanh (hoặc trường âm thanh) cơ bản nhờ sử dụng các thành phần dạng biểu diễn âm thanh nén cơ bản đó chứa trong các lớp N_B thấp nhất (tức là lớp cơ sở và $N_B - 1$ lớp tăng cường). Thông tin cần thiết về các thành phần nào của dạng biểu diễn âm thanh (hoặc trường âm thanh) nén cơ bản chứa trong các lớp riêng rẽ được giả định là sẽ được biết bởi bộ giải nén 2100 từ gói dữ liệu có

thông tin cấu hình, mà được giả định là được gửi và thu trước các gói dữ liệu khung. Việc giải mã thực tế mỗi tải tin thông tin phụ trợ cơ bản phụ thuộc riêng rẽ $BSI_{D,m}$, $m = 1, \dots, N_B$ có thể được tách thành hai phần như sau:

1. Giải mã sơ bộ cho mỗi tải tin $BSI_{D,m}$, $m = 1, \dots, N_B$, bằng cách khai thác sự phụ thuộc của nó vào $J_m - 1$ thành phần dạng biểu diễn âm thanh nén cơ bản thứ nhất $BSRC_1, \dots, BSRC_{(J_m)-1}$ chứa trong m lớp thứ nhất, mà được giả định ở bước mã hóa.

2. Hiệu chỉnh liên tiếp mỗi tải tin $BSI_{D,m}$, $m = 1, \dots, N_B$, bằng cách xem xét rằng thành phần âm thanh cơ sở cuối cùng được tái tạo từ $J_{N_B} - 1$ thành phần dạng biểu diễn âm thanh nén cơ bản thứ nhất $BSRC_1, \dots, BSRC_{(J_{N_B})-1}$ chứa trong $N_B > m$ lớp thứ nhất, mà nhiều thành phần hơn được giả định đối với giải mã sơ bộ. Vì vậy, việc hiệu chỉnh có thể được hoàn thành bằng cách loại bỏ thông tin lỗi thời, mà có thể được do thuộc tính giả định ban đầu của thông tin phụ trợ cơ bản phụ thuộc rằng nếu các thành phần bù nhất định được thêm vào dạng biểu diễn âm thanh (hoặc trường âm thanh) nén cơ bản, thông tin phụ trợ cơ bản phụ thuộc cho mỗi thành phần bù riêng rẽ trở thành tập con của thông tin phụ trợ ban đầu.

Cuối cùng, dạng biểu diễn âm thanh (hoặc trường âm thanh) cơ bản được tái tạo cùng với tất cả các tải tin thông tin phụ trợ tăng cường $ESI_1, \dots, ESI_M$, các tải tin thông tin phụ trợ cơ bản BSI_I và $BSI_{D,m}$, $m = 1, \dots, M$, và giá trị N_E được cung cấp cho bộ xử lý giải nén dạng biểu diễn tăng cường 2300, mà tính toán dạng biểu diễn âm thanh (hoặc trường âm thanh) tăng cường cuối cùng nhờ sử dụng chỉ tải tin thông tin phụ trợ tăng cường ESI_{N_E} và loại bỏ tất cả các tải tin thông tin phụ trợ tăng cường khác. Nếu giá trị của N_E là bằng không, tất cả các tải tin thông tin phụ trợ tăng cường được loại bỏ và dạng biểu diễn âm thanh (hoặc trường âm thanh) tăng cường cuối cùng được tái tạo là bằng dạng biểu diễn âm thanh (hoặc trường âm thanh) cơ bản được tái tạo.

Tiếp theo, việc chọn lớp sẽ được mô tả. Trong trường hợp tất cả các gói dữ liệu khung có thể được giải nén độc lập với nhau, cả số N_B của lớp cao nhất cần được sử dụng thực tế để giải nén cho dạng biểu diễn âm thanh cơ bản và chỉ số N_E của tải tin thông tin phụ trợ tăng cường cần được sử dụng để giải nén được thiết lập về số cao nhất L của tải tin thông tin phụ trợ tăng cường hợp lệ, mà chính nó có thể được xác định bằng cách đánh giá cờ hợp lệ trong các tải tin thông tin phụ trợ tăng cường. Nhờ

khai thác hiểu biết về kích thước của mỗi tải tin thông tin phụ trợ tăng cường, có thể tránh được việc phân tích cú pháp phức tạp qua dữ liệu thực tế của các tải tin để xác định tính hợp lệ của chúng.

Trong trường hợp sự giải nén khác biệt với sự phụ thuộc liên khung được sử dụng, quyết định từ khung trước đó phải được xem xét bổ sung. Với sự giải nén khác biệt, các gói dữ liệu khung độc lập được truyền ở các khoảng thời gian cách đều để cho phép bắt đầu giải nén từ các thời điểm này, trong đó việc xác định các giá trị N_B và N_E trở nên độc lập với khung và được tiến hành như mô tả ở trên.

Để giải thích chi tiết quyết định phụ thuộc khung, tác giả sáng chế trước hết là biểu thị cho khung thứ k

- số cao nhất của tải tin thông tin phụ trợ tăng cường hợp lệ bởi $L(k)$
- số lớp cao nhất cần được chọn và sử dụng để giải nén dạng biểu diễn âm thanh cơ bản bởi $N_B(k)$
- số tải tin thông tin phụ trợ tăng cường cần sử dụng để giải nén bởi $N_E(k)$.

Nhờ sử dụng cách ký hiệu này, số lớp cao nhất cần được sử dụng để giải nén dạng biểu diễn âm thanh cơ bản bởi $N_B(k)$ được tính theo

$$N_B(k) = \min(N_B(k-1), L(k)). \quad (3)$$

Bằng cách chọn $N_B(k)$ không lớn hơn $N_B(k-1)$ và $L(k)$ bảo đảm rằng tất cả thông tin cần thiết để giải nén khác biệt cho dạng biểu diễn âm thanh cơ bản là khả dụng.

Số $N_E(k)$ của tải tin thông tin phụ trợ tăng cường cần sử dụng để giải nén được xác định theo

$$N_E(k) = \begin{cases} N_B(k) & \text{nếu } N_B(k) = N_B(k-1) \\ 0 & \text{nếu không} \end{cases}. \quad (4)$$

Điều này cụ thể nghĩa là nếu số lượng lớp cao nhất $N_B(k)$ cần sử dụng để giải nén của dạng biểu diễn âm thanh cơ bản không thay đổi, thì cùng số lớp tăng cường tương ứng được chọn. Tuy nhiên, trong trường hợp có thay đổi $N_B(k)$, sự tăng cường bị vô hiệu hóa bằng cách thiết lập $N_E(k)$ về không. Do sự giải nén khác biệt giả định cho thông tin phụ trợ tăng cường, nên sự thay đổi của nó theo $N_B(k)$ là không thể vì nó đòi hỏi sự giải nén của lớp thông tin phụ trợ tăng cường tương ứng ở khung trước mà được giả định là chưa được tiến hành.

Theo cách khác, nếu ở bước giải nén tất cả các tải tin thông tin phụ trợ tăng cường với các số lên đến $N_E(k)$ được giải nén song song, quy tắc lựa chọn (4) có thể được thay thế bằng

$$N_E(k) = N_B(k). \quad (5)$$

Cuối cùng, cần lưu ý rằng đối với sự giải nén khác biệt, số lượng của lớp được sử dụng cao nhất có thể chỉ tăng ở các gói dữ liệu khung độc lập, trong khi đó sự giảm là có thể ở mọi khung.

Tiếp theo, các phương án của sáng chế liên quan đến mã hóa theo lớp cho khung của dạng biểu diễn âm thanh nén và đến cấu trúc dữ liệu (ví dụ, dòng bit) biểu diễn khung của âm thanh nén được mã hóa sẽ được mô tả đối với trường hợp dạng biểu diễn HOA nén. Cụ thể, các thay đổi được đề xuất đối với sơ đồ mã hóa theo lớp của dạng biểu diễn HOA nén sẽ được mô tả.

Theo sự hiệu chỉnh chế độ mã hóa theo lớp cho nội dung dựa trên HOA, `usacExtElementType` mới được xác định để làm thích ứng tốt hơn cấu hình và tải tin khung của các công cụ giải mã HOA như Dự báo tín hiệu không gian, Tổng hợp tín hiệu định hướng dải con và Bộ giải mã tái tạo môi trường tham số (Parametric Ambience Replication - PAR) với lớp tăng cường HOA tương ứng. Nếu Chế độ mã hóa theo lớp cho nội dung dựa trên HOA được kích hoạt, mà được báo hiệu bởi `SingleLayer==0`, đề xuất là di chuyển các phần tử dòng bit tương ứng của các công cụ này đến một tải tin mở rộng HOA bổ sung thuộc loại mới đối với mỗi lớp (bao gồm lớp cơ sở và một hoặc nhiều lớp tăng cường).

Sự mở rộng phải được thực hiện vì thông tin phụ trợ cho các công cụ này được tạo ra để tăng cường dạng biểu diễn HOA đặc trưng. Theo định nghĩa hiện tại về mã hóa HOA theo lớp dữ liệu được cung cấp chỉ mở rộng thích đáng dạng biểu diễn HOA của lớp cao nhất. Đối với các lớp thấp hơn, các công cụ này không tăng cường dạng biểu diễn HOA được tái tạo một phần một cách thích đáng.

Vì vậy, tốt hơn là cung cấp thông tin phụ trợ của các công cụ này đối với mỗi lớp để làm thích ứng tốt hơn chúng với dạng biểu diễn HOA tái tạo của lớp tương ứng.

Ngoài ra, các công cụ Tổng hợp tín hiệu định hướng dải con và Bộ giải mã tái tạo môi trường tham số được thiết kế đặc thù cho các tốc độ dữ liệu thấp, nơi chỉ có vài tín hiệu truyền là khả dụng. Do đó sự mở rộng đề xuất sẽ tạo ra khả năng làm thích ứng tối ưu thông tin phụ trợ của các công cụ này với số lượng tín hiệu truyền trong lớp.

Theo đó, chất lượng âm thanh của dạng biểu diễn HOA tái tạo cho các lớp tốc độ bit thấp, ví dụ, lớp cơ sở, có thể được tăng đáng kể so với phương thức theo lớp hiện có.

Hơn thế nữa, cú pháp dòng bit cho các phần tử vector V được mã hóa đối với các tín hiệu dựa trên vector phải được làm thích ứng cho việc mã hóa theo lớp HOA nếu `CodedVVecLength` bằng một được báo hiệu trong `HOADecoderConfig()`. Theo chế độ mã hóa vector này các phần tử vector V không được truyền cho các chỉ số hệ số HOA mà được bao gồm trong tập `ContAddHoaCoeff`. Tập này bao gồm tất cả các chỉ số hệ số HOA `AmbCoeffIdx[i]` mà có `AmbCoeffTransitionState` bằng không. Không cần phải bổ sung tín hiệu vector V có trọng số vì chuỗi hệ số HOA ban đầu cho các chỉ số này rõ ràng được gửi. Vì vậy phần tử vector V theo phương thức thông thường được thiết lập về không cho các chỉ số này.

Tuy nhiên, trong chế độ mã hóa theo lớp tập chỉ số hệ số HOA liên tiếp phụ thuộc vào các kênh truyền là một phần của lớp hoạt động hiện thời. Điều này nghĩa là các chỉ số hệ số HOA bổ sung được gửi trong lớp cao hơn là thiếu trong các lớp thấp hơn. Sau đó giả định rằng tín hiệu vector không đóng góp vào chuỗi hệ số HOA là sai đối với các chỉ số hệ số HOA mà thuộc chuỗi hệ số HOA được bao gồm trong các lớp cao hơn. Do đó, đề xuất báo hiệu (một cách rõ ràng) các phần tử vector V cho các chỉ số hệ số thiếu này.

Hệ quả là, đề xuất xác định tập của `ContAddHoaCoeff` đối với mỗi lớp và sử dụng tập của lớp khi tín hiệu vector V được bổ sung (tín hiệu truyền của tín hiệu vector V thuộc về) cho việc chọn các phần tử vector V hoạt động. Tuy nhiên, đề xuất là dữ liệu vector V ở trong `HOAFrame()` và không bị di chuyển đến `HOAEnhFrame()`.

Tiếp theo, sự tích hợp vào cú pháp dòng bit MPEG-H sẽ được mô tả. Phương pháp mã hóa tương ứng (ví dụ, phương pháp mã hóa theo lớp đối với khung của dạng biểu diễn HOA nén của âm thanh hoặc trường âm thanh) theo các phương án của sáng chế sẽ được mô tả tham chiếu đến Fig.3. Các thay đổi đề xuất đối với dòng bit MPEG-H 3D sẽ được mô tả dưới đây trong phụ lục.

Ở chế độ mã hóa theo lớp cơ `SingleLayer` trong `HOADecoderConfig()` là bất hoạt (`SingleLayer==0`) và số lượng lớp và số lượng tín hiệu truyền HOA được gán tương ứng của chúng được xác định. Nhìn chung, dạng biểu diễn HOA nén có thể bao gồm nhiều tín hiệu truyền.

Theo đó, ở S3010 trên Fig.3, nhiều tín hiệu truyền được gán cho nhiều lớp phân

cấp. Nói cách khác, các tín hiệu truyền được phân bổ cho nhiều lớp. Mỗi lớp có thể được cho là bao gồm các tín hiệu truyền tương ứng gán cho lớp đó. Mỗi lớp có thể có nhiều hơn một tín hiệu truyền được gán vào đó. Các lớp có thể bao gồm một lớp cơ sở và một hoặc nhiều lớp tăng cường phân cấp. Các lớp có thể được sắp xếp, từ lớp cơ sở, đến lớp tăng cường, đến lớp tăng cường tổng thể cao nhất (lớp tổng thể cao nhất).

Đề xuất bổ sung tải tin mở rộng cấu hình HOA bổ sung và tải tin mở rộng khung HOA với `usacExtElementType ID_EXT_ELE_HOA_ENH_LAYER` mới được xác định vào dòng bit MPEG-H để truyền một tải tin của Dự báo tín hiệu không gian, Tổng hợp tín hiệu định hướng dải con và dữ liệu bộ giải mã PAR cho mỗi lớp HOA tăng cường (bao gồm lớp cơ sở). Các tải tin bổ sung này sẽ trực tiếp theo sau tải tin thuộc loại `ID_EXT_ELE_HOA` trong `mpegh3daExtElementConfig()` và tương ứng trong `mpegh3daFrame()`.

Vì vậy, đề xuất di chuyển, trong trường hợp `SingleLayer==0`, các phần tử cấu hình cho dự báo tín hiệu không gian, tổng hợp tín hiệu định hướng dải con và bộ giải mã PAR từ `HOADecoderConfig()` đến `HOADecoderEnhConfig()` mới được xác định và lần lượt `HOAPredictionInfo()`, `HOADirectionalPredictionInfo()` và `HOAParInfo()` từ `HOAFrame()` đến `HOAEnhFrame()` mới được xác định.

Theo đó, ở S3020, tải tin mở rộng HOA tương ứng được tạo ra đối với mỗi lớp. Tải tin mở rộng HOA được tạo ra bao gồm thông tin phụ trợ để tăng cường tham số cho dạng biểu diễn HOA tái tạo có thể thu được từ các tín hiệu truyền được gán cho (ví dụ, bao gồm trong) lớp tương ứng và các lớp bất kỳ thấp hơn lớp tương ứng này. Như nêu trên đây, các tải tin mở rộng HOA có thể bao gồm các phần tử dòng bit cho một hoặc nhiều trong số công cụ giải mã dự báo tín hiệu không gian HOA, công cụ giải mã tổng hợp tín hiệu định hướng dải con HOA, và công cụ giải mã tái tạo môi trường tham số HOA. Hơn nữa, các tải tin mở rộng HOA có thể có `usacExtElementType` của `ID_EXT_ELE_HOA_ENH_LAYER`.

Ở S3030, các tải tin mở rộng HOA được tạo ra được gán cho các lớp tương ứng của chúng.

Hơn nữa (không được thể hiện trên Fig.3), tải tin mở rộng cấu hình HOA bao gồm các phần tử dòng bit để tạo cấu hình cho công cụ giải mã dự báo tín hiệu không gian HOA, công cụ giải mã tổng hợp tín hiệu định hướng dải con HOA, và/hoặc công cụ giải mã tái tạo môi trường tham số HOA có thể được tạo ra.

Hơn nữa (không được thể hiện trên Fig.3) tải tin cấu hình bộ giải mã HOA bao gồm thông tin chỉ báo việc gán các tải tin mở rộng HOA cho nhiều lớp có thể được tạo ra.

Tiếp theo, việc truyền dòng bit được tạo lớp (ví dụ, dòng bit MPEG-H) sẽ được mô tả. Do tất cả các tải tin mở rộng của dòng bit MPEG-H được căn chỉnh byte và kích thước của chúng được báo hiệu rõ ràng, việc cờ `elementLengthPresent` bằng một được giả định, nên bộ giải gói có thể phân tích cú pháp dòng bit MPEG-H và trích xuất các tải tin cho các lớp cao hơn một và truyền chúng một cách riêng rẽ qua các kênh truyền khác nhau. Lớp cơ sở chứa (ví dụ, bao gồm) dòng bit MPEG-H không chứa dữ liệu cho các lớp cao hơn. Các tải tin mở rộng thiếu được báo hiệu là trống hoặc bất hoạt. Đối với các tải tin thuộc loại `ID_USAC_SCE`, `ID_USAC_CPE` và `ID_USAC_LFE`, tải tin trống được báo hiệu bởi `elementLength` bằng không, trong đó `elementLengthPresent` cần được thiết lập về một. Tải tin trống thuộc loại `ID_USAC_EXT` có thể được báo hiệu bằng cách thiết lập cờ `usacExtElementPresent` về không (sai).

Theo đó, ở S3040, các tải tin mở rộng HOA được tạo ra được báo hiệu (ví dụ, được truyền, hoặc xuất ra) trong dòng bit đầu ra. Nhìn chung, nhiều lớp và các tải tin được gán vào đó được báo hiệu (ví dụ, được truyền hoặc xuất ra) trong dòng bit đầu ra. Hơn nữa, tải tin cấu hình bộ giải mã HOA và/hoặc tải tin mở rộng cấu hình HOA có thể được báo hiệu (ví dụ, được truyền hoặc xuất ra) trong dòng bit đầu ra.

Giả định rằng lớp cơ sở HOA (chỉ số lớp bằng một) được truyền với mức độ bảo vệ chống lỗi cao nhất và có tốc độ bit tương đối nhỏ. Mức độ bảo vệ chống lỗi cho các lớp sau (một hoặc nhiều lớp tăng cường HOA) được giảm đều theo sự tăng tốc độ bit của lớp tăng cường. Do các điều kiện truyền kém và mức độ bảo vệ chống lỗi thấp hơn, nên việc truyền các lớp cao hơn có thể thất bại và trong trường hợp xấu nhất chỉ lớp cơ sở được truyền chính xác. Giả định rằng mức độ bảo vệ chống lỗi kết hợp cho tất cả các tải tin của một lớp được áp dụng. Do đó nếu việc truyền của lớp thất bại, tất cả các tải tin của lớp tương ứng bị thiếu.

Nói cách khác, các tải tin dữ liệu cho nhiều lớp có thể được truyền với mức độ bảo vệ chống lỗi tương ứng, trong đó lớp cơ sở có mức độ bảo vệ chống lỗi cao nhất và một hoặc nhiều lớp tăng cường có mức độ bảo vệ chống lỗi giảm liên tục.

Trừ khi các bước đòi hỏi phải có các bước khác nhất định làm các điều kiện tiên quyết, các bước nêu trên có thể được thực hiện theo thứ tự bất kỳ và thứ tự ví dụ được

minh họa trên Fig.3 được hiểu là không làm giới hạn sáng chế.

Như nêu trên đây, cú pháp dòng bit cho các phần tử vector V được mã hóa đối với các tín hiệu dựa trên vectơ phải được làm thích ứng cho việc mã hóa theo lớp HOA nếu CodedVVecLength bằng một được báo hiệu trong HOADecoderConfig(). Phương pháp mã hóa tương ứng (ví dụ, phương pháp mã hóa theo lớp đối với khung của dạng biểu diễn HOA nén của âm thanh hoặc trường âm thanh) theo các phương án của sáng chế sẽ được mô tả tham chiếu đến Fig.4.

Ở S4010 trên Fig.4, nhiều tín hiệu truyền được gán cho nhiều lớp phân cấp. Bước này có thể được thực hiện theo cùng cách với S3010 mô tả ở trên.

Ở S4020, xác định liệu chế độ mã hóa vectơ có hoạt động hay không. Điều này bao gồm việc xác định liệu CodedVVecLength==1 hay không.

Như nêu trên đây, theo phương thức thông thường ở chế độ mã hóa vectơ, các phần tử vectơ V không được truyền cho các chỉ số hệ số HOA mà được bao gồm trong tập ContAddHoaCoeff. Tập này bao gồm tất cả các chỉ số hệ số HOA AmbCoeffIdx[i] mà có AmbCoeffTransitionState bằng không. Không cần phải bổ sung tín hiệu vectơ V có trọng số vì chuỗi hệ số HOA ban đầu cho các chỉ số này rõ ràng được gửi. Vì vậy phần tử vectơ V theo phương thức thông thường được thiết lập về không cho các chỉ số này.

Tuy nhiên, trong chế độ mã hóa theo lớp tập chỉ số hệ số HOA liên tiếp phụ thuộc vào các kênh truyền là một phần của lớp hoạt động hiện thời. Điều này nghĩa là các chỉ số hệ số HOA bổ sung được gửi trong lớp cao hơn là thiếu trong các lớp thấp hơn. Sau đó sự giả định rằng tín hiệu vectơ không đóng góp vào chuỗi hệ số HOA là sai đối với các chỉ số hệ số HOA mà thuộc chuỗi hệ số HOA bao gồm trong các lớp cao hơn.

Do đó, nếu chế độ mã hóa vectơ hoạt động, ở S4030 tập chỉ số hệ số HOA liên tiếp (ví dụ, ContAddHoaCoeff) được xác định (ví dụ, định rõ) đối với mỗi lớp trên cơ sở các tín hiệu truyền được gán cho lớp tương ứng.

Nếu chế độ mã hóa vectơ hoạt động, ở S4040, đối với mỗi tín hiệu truyền, vectơ V được tạo ra trên cơ sở tập chỉ số hệ số HOA liên tiếp xác định được đối với lớp mà tín hiệu truyền tương ứng được gán vào đó. Mỗi vectơ V được tạo ra có thể bao gồm các phần tử cho các tín hiệu truyền bất kỳ gán cho các lớp cao hơn lớp mà tín hiệu truyền tương ứng được gán vào đó. Bước này có thể bao gồm việc sử dụng tập chỉ số

hệ số HOA liên tiếp mà đã được xác định cho lớp mà tín hiệu vector V được thêm vào (lớp mà tín hiệu truyền của tín hiệu vector V thuộc về) cho việc lựa chọn các phần tử vector V hoạt động. Tuy nhiên, đề xuất là dữ liệu vector V ở trong HOAFrame() và không bị di chuyển đến HOAEnhFrame().

Sau đó, ở S4050 các vector V được tạo ra (các tín hiệu vector V) được báo hiệu trong dòng bit đầu ra. Điều này có thể bao gồm việc báo hiệu (một cách rõ ràng) các phần tử vector V đối với các chỉ số hệ số thiếu nêu trên.

Các bước S4020 đến S4050 trên Fig.4 cũng có thể được sử dụng trong ngữ cảnh của phương pháp mã hóa minh họa trên Fig.3, ví dụ, sau S3010. Trong trường hợp này, S3040 và S4050 có thể được kết hợp thành một bước báo hiệu đơn nhất.

Trừ khi các bước đòi hỏi phải có các bước nhất định khác làm điều kiện tiên quyết, các bước nêu trên có thể được thực hiện theo thứ tự bất kỳ và thứ tự ví dụ được minh họa trên Fig.4 được hiểu là không làm giới hạn sáng chế.

Ở phía bộ thu bộ đóng gói dòng bit MPEG-H có thể chèn lại các tải tin thu được một cách chính xác vào dòng bit MPEG-H của lớp cơ sở và chuyển nó đến bộ giải mã âm thanh MPEG-H 3D.

Tiếp theo, (cấu hình) khởi tạo giải mã HOA sẽ được mô tả. Các tải tin cấu hình HOA thuộc loại ID_EXT_ELE_HOA và ID_EXT_ELE_HOA_ENH_LAYER với các kích thước tương ứng của chúng ở đơn vị byte được nhập vào bộ giải mã HOA cho việc khởi tạo của nó. Các công cụ mã hóa HOA được tạo cấu hình theo các phần tử dòng bit được xác định trong HOAConfig(), mà được phân tích cú pháp từ tải tin thuộc loại ID_EXT_ELE_HOA. Hơn nữa, tải tin này chứa cách sử dụng của Chế độ mã hóa theo lớp, số lượng của các lớp và số lượng tín hiệu truyền tương ứng trên mỗi lớp. Sau đó, nếu mã hóa theo lớp được kích hoạt (SingleLayer==0), các HOAEnhConfig() được phân tích cú pháp từ các tải tin thuộc loại ID_EXT_ELE_HOA_ENH_LAYER để tạo cấu hình cho Dự báo tín hiệu không gian, Tổng hợp tín hiệu định hướng dải con và Bộ giải mã tái tạo môi trường tham số tương ứng của mỗi lớp.

Phần tử LayerIdx từ HOAEnhConfig() cùng với thứ bậc của các tải tin cấu hình lớp tăng cường HOA trong mpeg3daExtElementConfig() chỉ báo thứ bậc của lớp tăng cường HOA. Thứ bậc của các tải tin khung lớp tăng cường HOA thuộc loại ID_EXT_ELE_HOA_ENH_LAYER trong mpeg3daFrame() là giống với thứ bậc của các tải tin cấu hình trong mpeg3daExtElementConfig() để gán một cách rõ ràng các

tải tin khung vào các lớp tương ứng.

Trong trường hợp `SingleLayer==1` (mã hóa một lớp) các tải tin thuộc loại `ID_EXT_ELE_HOA_ENH_LAYER` được bỏ qua và Dự báo tín hiệu không gian, Tổng hợp tín hiệu định hướng dải con và Bộ giải mã tái tạo môi trường tham số sử dụng dữ liệu tương ứng từ `HOADecoderConfig()` cho cấu hình của chúng.

Tiếp theo, việc giải mã khung HOA ở chế độ theo lớp sẽ được mô tả. Phương pháp giải mã tương ứng (ví dụ, phương pháp giải mã khung của dạng biểu diễn HOA nén của âm thanh hoặc trường âm thanh) theo các phương án của sáng chế sẽ được mô tả tham chiếu đến Fig.5. Hiểu rằng dạng biểu diễn HOA nén (ví dụ, đầu ra của các phương pháp trên Fig.3 hoặc Fig.4 nêu trên) đã được mã hóa trong nhiều lớp phân cấp bao gồm lớp cơ sở và một hoặc nhiều lớp tăng cường.

Ở S5010 trên Fig.5, dòng bit liên quan đến khung của dạng biểu diễn HOA nén được thu.

Bộ giải mã lõi âm thanh 3D giải mã các tín hiệu truyền HOA được truyền chính xác và tạo ra các tín hiệu truyền với tất cả các mẫu bằng không đối với các tải tin không hợp lệ tương ứng. Các tín hiệu truyền được giải mã cùng với các cờ `usacExtElementPresent`, dữ liệu và kích thước của các tải tin HOA thuộc loại `ID_EXT_ELE_HOA` và `ID_EXT_ELE_HOA_ENH_LAYER` được nhập vào bộ giải mã HOA. Các tải tin mở rộng từ loại `ID_USAC_EXT` với cờ `usacExtElementPresent` được thiết lập là sai phải được báo hiệu dưới dạng các tải tin thiếu cho bộ giải mã HOA để bảo đảm việc gán các tải tin vào các lớp tương ứng.

Ở S5020, các tải tin cho nhiều lớp được trích xuất. Mỗi tải tin có thể bao gồm các tín hiệu truyền gán cho lớp tương ứng.

Ở bước này, bộ giải mã HOA có thể phân tích cú pháp `HOAFrame()` từ tải tin thuộc loại `ID_EXT_ELE_HOA`.

Sau đó, các tải tin hợp lệ thuộc loại `ID_EXT_ELE_HOA_ENH_LAYER` và các tải tin không hợp lệ thuộc loại `ID_EXT_ELE_HOA_ENH_LAYER` được xác định bằng cách đánh giá cờ `usacExtElementPresent` tương ứng của các tải tin, trong đó tải tin không hợp lệ được chỉ báo bởi cờ `usacExtElementPresent` bằng sai và việc gán các tải tin tăng cường HOA vào các chỉ số lớp tăng cường được biết từ cấu hình bộ giải mã HOA.

Ở S5030, lớp dùng được cao nhất trong số nhiều lớp để giải mã được xác định.

Do các lớp là phụ thuộc vào nhau về các tín hiệu truyền, bộ giải mã HOA có thể chỉ giải mã một lớp khi tất cả các lớp có chỉ số thấp hơn được thu một cách chính xác. Lớp dùng được cao nhất có thể được chọn ở bước này để cho tất cả các lớp lên đến lớp dùng được cao nhất được thu một cách chính xác. Chi tiết của các bước này sẽ được mô tả dưới đây.

Ở S5040, tải tin mở rộng HOA gán cho lớp dùng được cao nhất được trích xuất. Như nêu trên đây, tải tin mở rộng HOA có thể bao gồm thông tin phụ trợ để tăng cường tham số cho dạng biểu diễn HOA tái tạo tương ứng với lớp dùng được cao nhất. Ở đây, dạng biểu diễn HOA tái tạo tương ứng với lớp dùng được cao nhất có thể là thu được trên cơ sở các tín hiệu truyền gán cho lớp dùng được cao nhất và các lớp bất kỳ thấp hơn lớp dùng được cao nhất.

Ngoài ra, các tải tin mở rộng HOA lần lượt được gán cho các tải tin còn lại của nhiều lớp có thể được trích xuất. Mỗi tải tin mở rộng HOA có thể bao gồm thông tin phụ trợ để tăng cường tham số cho dạng biểu diễn HOA tái tạo tương ứng với lớp được gán tương ứng của nó. Dạng biểu diễn HOA tái tạo tương ứng với lớp được gán tương ứng của nó có thể thu được từ các tín hiệu truyền gán cho lớp đó và các lớp bất kỳ thấp hơn lớp đó.

Hơn nữa (không được thể hiện trên Fig.5), phương pháp giải mã có thể bao gồm bước trích xuất tải tin mở rộng cấu hình HOA. Điều này có thể được thực hiện bằng cách phân tích cú pháp dòng bit. Tải tin mở rộng cấu hình HOA có thể bao gồm các phần tử dòng bit để tạo cấu hình cho công cụ giải mã dự báo tín hiệu không gian HOA, công cụ giải mã tổng hợp tín hiệu định hướng dải con HOA, và/hoặc công cụ giải mã tái tạo môi trường tham số HOA.

Ở S5050, dạng biểu diễn HOA tái tạo (một phần) tương ứng với lớp dùng được cao nhất được tạo ra trên cơ sở các tín hiệu truyền gán cho lớp dùng được cao nhất và các lớp bất kỳ thấp hơn lớp dùng được cao nhất.

Số lượng các tín hiệu truyền được sử dụng thực tế $I_{\text{ADD,LAY}}(k)$ được thiết lập theo (chỉ số $M_{\text{LAY}}(k)$ của) lớp dùng được cao nhất và dạng biểu diễn HOA sơ bộ thứ nhất được giải mã từ $\text{HOAFrame}()$ và từ các tín hiệu truyền tương ứng của lớp này và các lớp thấp hơn bất kỳ.

Sau đó, ở S5060 dạng biểu diễn HOA tái tạo được tăng cường (ví dụ, được tăng cường tham số) nhờ sử dụng thông tin phụ trợ bao gồm trong tải tin mở rộng HOA gán cho lớp dùng được cao nhất.

Tức là, dạng biểu diễn HOA thu được ở bước S5050 sau đó được tăng cường bởi Dự báo tín hiệu không gian, Tổng hợp tín hiệu định hướng dải con và Bộ giải mã tái tạo môi trường tham số nhờ sử dụng dữ liệu HOAEnhFrame() được phân tích cú pháp từ tải tin mở rộng lớp tăng cường HOA thuộc loại ID_EXT_ELE_HOA_ENH_LAYER của lớp hoạt động hiện thời $M_{LAY}(k)$, tức là, lớp dùng được cao nhất.

Thông tin được sử dụng ở các bước S5020-S5060 có thể được gọi là thông tin lớp.

Trừ khi các bước đòi hỏi phải có các bước nhất định khác làm các điều kiện tiên quyết, các bước nêu trên có thể được thực hiện theo thứ tự bất kỳ và thứ tự ví dụ được minh họa trên Fig.5 được hiểu là không làm giới hạn sáng chế.

Tiếp theo, chi tiết về việc xác định (ví dụ, chọn) lớp dùng được cao nhất ở bước S5030 sẽ được mô tả.

Như nêu trên đây, bộ giải mã HOA chỉ có thể giải mã một lớp khi tất cả các lớp có chỉ số thấp hơn được thu một cách chính xác, do các lớp phụ thuộc vào nhau về các tín hiệu truyền.

Để chọn lớp giải mã được cao nhất, bộ giải mã HOA có thể tạo tập chỉ số lớp không hợp lệ, trong đó chỉ số nhỏ nhất từ tập này trừ một tạo ra chỉ số M_{LAY} của lớp tăng cường giải mã được cao nhất. Tập chỉ số lớp không hợp lệ có thể được xác định bằng cách đánh giá cờ hợp lệ của các tải tin mở rộng HOA tương ứng.

Nói cách khác, bước xác định lớp dùng được cao nhất có thể bao gồm việc xác định tập chỉ số lớp không hợp lệ chỉ báo các lớp mà không được thu hợp lệ. Nó có thể còn bao gồm xác định lớp dùng được cao nhất là lớp mà thấp hơn một lớp so với lớp được chỉ báo bởi chỉ số nhỏ nhất trong tập chỉ số lớp không hợp lệ. Theo cách đó, bảo đảm rằng tất cả các lớp dưới lớp dùng được cao nhất đã được thu hợp lệ.

Trong trường hợp mã hóa khác biệt các khung, chỉ số của lớp dùng được cao nhất của khung trước (ví dụ, ngay trước) sẽ phải được tính đến. Thứ nhất, tình huống sẽ được mô tả trong đó chỉ số của lớp dùng được cao nhất của khung trước (ví dụ, đứng trước) được giữ.

Nếu chỉ số của lớp dùng được cao nhất (ví dụ, lớp giải mã được cao nhất) đối với khung hiện thời là bằng chỉ số lớp của khung trước $M_{LAY}(k-1)$, chỉ số lớp của khung hiện thời $M_{LAY}(k)$ được thiết lập là $M_{LAY}(k-1)$.

Sau đó, số lượng các tín hiệu truyền được sử dụng thực tế $I_{ADD,LAY}(k)$ được thiết lập theo $M_{LAY}(k)$ và dạng biểu diễn HOA sơ bộ thứ nhất được giải mã từ HOAFrame() và từ các tín hiệu truyền tương ứng của lớp này và các lớp thấp hơn bất kỳ, như chỉ ra ở trên. Dạng biểu diễn HOA này sau đó được tăng cường bởi Dự báo tín hiệu không gian, Tổng hợp tín hiệu định hướng dải con và Bộ giải mã tái tạo môi trường tham số nhờ sử dụng dữ liệu HOAEnhFrame() được phân tích cú pháp từ tải tin mở rộng lớp tăng cường HOA thuộc loại ID_EXT_ELE_HOA_ENH_LAYER của lớp hoạt động hiện thời $M_{LAY}(k)$, như được chỉ ra ở trên.

Tiếp theo, tình huống sẽ được mô tả trong đó đã được chuyển sang chỉ số thấp hơn chỉ số của lớp dùng được cao nhất của khung trước (ví dụ, đứng trước). Đó là, trong trường hợp mà chỉ số của lớp giải mã được cao nhất đối với khung hiện thời là nhỏ hơn chỉ số của lớp của khung trước $M_{LAY}(k-1)$, bộ giải mã HOA thiết lập $M_{LAY}(k)$ về chỉ số của lớp giải mã được cao nhất đối với khung hiện thời. Sự giải mã của các tải tin cho Dự báo tín hiệu không gian, Tổng hợp tín hiệu định hướng dải con và Bộ giải mã tái tạo môi trường tham số cho các lớp mới có thể chỉ bắt đầu ở khung HOA tiếp theo với hoaIndependencyFlag bằng một. Cho đến khi HOAFrame() như vậy được thu, dạng biểu diễn HOA của lớp của chỉ số $M_{LAY}(k)$ được tái tạo mà không thực hiện Dự báo tín hiệu không gian, Tổng hợp tín hiệu định hướng dải con và Bộ giải mã tái tạo môi trường tham số. Điều này nghĩa là số lượng các tín hiệu truyền được sử dụng thực tế $I_{ADD,LAY}(k)$ được thiết lập theo $M_{LAY}(k)$ và chỉ dạng biểu diễn HOA sơ bộ thứ nhất được giải mã từ HOAFrame() và từ các tín hiệu truyền tương ứng của lớp này và các lớp thấp hơn bất kỳ. Sau đó, nếu HOAFrame() với hoaIndependencyFlag bằng một được thu, các tải tin cho Dự báo tín hiệu không gian, Tổng hợp tín hiệu định hướng dải con và Bộ giải mã tái tạo môi trường tham số được phân tích cú pháp và giải mã để tăng cường dạng biểu diễn HOA sơ bộ, do đó chất lượng trọn vẹn của lớp hiện thời được cung cấp cho khung này.

Do đó, phương pháp đề xuất có thể bao gồm (không được thể hiện trên Fig.5) bước quyết định không thực hiện tăng cường tham số cho dạng biểu diễn HOA tái tạo nhờ sử dụng thông tin phụ trợ bao gồm trong tải tin mở rộng HOA gán cho lớp dùng

được cao nhất nếu lớp dùng được cao nhất của khung hiện thời là thấp hơn lớp dùng được cao nhất của khung trước (nếu khung hiện thời được mã hóa khác biệt so với khung trước).

Nhìn chung, bước xác định lớp dùng được cao nhất cho khung hiện thời có thể bao gồm việc xác định tập chỉ số lớp không hợp lệ chỉ báo các lớp mà không được thu hợp lệ cho khung hiện thời. Nó có thể còn bao gồm bước xác định lớp dùng được cao nhất của khung trước ở trước khung hiện thời. Nó có thể còn bao gồm bước xác định lớp dùng được cao nhất là lớp thấp hơn của lớp dùng được cao nhất của khung trước và lớp mà thấp hơn một lớp so với lớp được chỉ báo bởi chỉ số nhỏ nhất trong tập chỉ số lớp không hợp lệ (nếu khung hiện thời được mã hóa khác biệt so với khung trước).

Giải pháp thay thế có thể luôn phân tích cú pháp tất cả các tải tin lớp tăng cường hợp lệ (ví dụ, các tải tin mở rộng HOA) song song cho dù chúng hiện thời bất hoạt. Điều này sẽ kích hoạt việc chuyển đổi trực tiếp sang lớp có chỉ số thấp hơn với chất lượng trọn vẹn, trong đó Dự báo tín hiệu không gian, Tổng hợp tín hiệu định hướng dải con và Bộ giải mã Tái tạo môi trường tham số (Parametric Ambience Replication - PAR) có thể được áp dụng trực tiếp ở khung được chuyển đổi.

Tiếp theo, tình huống sẽ được mô tả trong đó đã được chuyển đổi sang chỉ số cao hơn chỉ số của lớp dùng được cao nhất của khung trước (ví dụ, đứng trước). Việc chuyển sang lớp có chỉ số cao hơn này chỉ có thể được áp dụng nếu `mpegh3daFrame()` có `usacIndependencyFlag` bằng một (ví dụ, nếu khung là khung độc lập) vì các tải tin tương ứng hoặc các trạng thái giải mã của khung trước là thiếu. Do đó, bộ giải mã HOA giữ chỉ số lớp $HOA_{LAY}(k)$ bằng $M_{LAY}(k-1)$ cho đến khi thu được `anmpegh3daFrame()` với `usacIndependencyFlag` bằng một (ví dụ, khung độc lập) mà chứa dữ liệu hợp lệ cho lớp giải mã được cao hơn. Sau đó $M_{LAY}(k)$ được thiết lập là chỉ số lớp giải mã được cao nhất đối với khung hiện thời và theo đó, số lượng tín hiệu truyền $I_{ADD,LAY}(k)$ được sử dụng thực tế được xác định. Dạng biểu diễn HOA sơ bộ của lớp đó được giải mã từ `HOAFrame()` và các tín hiệu truyền tương ứng và được tăng cường bởi Dự báo tín hiệu không gian, Tổng hợp tín hiệu định hướng dải con và Bộ giải mã tái tạo môi trường tham số nhờ sử dụng `HOAEnhFrame()` được phân tích cú pháp từ tải tin mở rộng lớp tăng cường HOA thuộc loại `ID_EXT_ELE_HOA_ENH_LAYER` của lớp hoạt động hiện thời $M_{LAY}(k)$.

Hiểu rằng phương pháp được đề xuất để mã hóa theo lớp dạng biểu diễn âm thanh nén có thể được thực hiện bằng bộ mã hóa để mã hóa theo lớp dạng biểu diễn âm thanh nén. Bộ mã hóa như vậy có thể bao gồm các đơn vị tương ứng được làm thích ứng để tiến hành các bước tương ứng mô tả ở trên. Ví dụ về bộ mã hóa 6000 này được minh họa bằng sơ đồ trên Fig.6. Chẳng hạn, bộ mã hóa 6000 có thể bao gồm đơn vị gán tín hiệu truyền 6010 được làm thích ứng để thực hiện bước S3010 nêu trên, đơn vị tạo tải tin lớp mở rộng HOA 6020 được làm thích ứng để thực hiện S3020 nêu trên, đơn vị gán tải tin mở rộng HOA 6030 được làm thích ứng để thực hiện S3030 nêu trên, và đơn vị báo hiệu hoặc đơn vị đầu ra 6040 được làm thích ứng để thực hiện S3040 nêu trên. Hiểu thêm rằng các đơn vị tương ứng của bộ mã hóa này có thể được thực thi bởi bộ xử lý 6100 của thiết bị máy tính mà được làm thích ứng để thực hiện việc xử lý tiến hành bởi mỗi trong các đơn vị tương ứng này, tức là được làm thích ứng để tiến hành một số hoặc tất cả các bước nêu trên của phương pháp mã hóa đề xuất được minh họa bằng sơ đồ trên Fig.3. Ngoài ra hoặc theo cách khác, bộ xử lý 6100 có thể được làm thích ứng để tiến hành từng bước của phương pháp mã hóa được minh họa bằng sơ đồ trên Fig.4. Với mục đích này, bộ xử lý 6100 có thể được làm thích ứng để thực thi các đơn vị tương ứng của bộ mã hóa. Bộ mã hóa hoặc thiết bị tính toán có thể còn bao gồm bộ nhớ 6200 mà có thể truy cập được bởi bộ xử lý 6100.

Hiểu thêm rằng phương pháp được đề xuất để giải mã dạng biểu diễn âm thanh nén mà được mã hóa trong nhiều lớp phân cấp có thể được thực hiện bởi bộ giải mã để giải mã dạng biểu diễn âm thanh nén mà được giải mã trong nhiều lớp phân cấp. Bộ giải mã này có thể bao gồm các đơn vị tương ứng được làm thích ứng để tiến hành các bước tương ứng mô tả ở trên. Ví dụ về bộ giải mã 7000 như vậy được minh họa bằng sơ đồ trên Fig.7. Chẳng hạn, bộ giải mã 7000 như vậy có thể bao gồm bộ thu 7010 được làm thích ứng để thực hiện bước S5010 nêu trên, bộ trích xuất tải tin 7020 được làm thích ứng để thực hiện bước S5020 nêu trên, bộ xác định lớp dùng được cao nhất 7030 được làm thích ứng để thực hiện bước S5030 nêu trên, bộ trích xuất tải tin mở rộng HOA 7040 được làm thích ứng để thực hiện bước S5040 nêu trên, bộ tạo dạng biểu diễn HOA tái tạo 7050 được làm thích ứng để thực hiện bước S5050 nêu trên, và bộ tăng cường 7060 được làm thích ứng để thực hiện bước S5060 nêu trên. Hiểu thêm rằng các đơn vị tương ứng của bộ giải mã này có thể được thực thi bởi bộ xử lý 7100 của thiết bị tính toán mà được làm thích ứng để thực hiện việc xử lý tiến hành bởi từng

đơn vị tương ứng, tức là được làm thích ứng để tiến hành một số hoặc tất cả các bước nêu trên của phương pháp giải mã đề xuất. Bộ giải mã hoặc thiết bị tính toán có thể còn bao gồm bộ nhớ 7200 mà truy cập được bởi bộ xử lý 7100.

Tiếp theo, cấu trúc dữ liệu (ví dụ, dòng bit) để cung cấp (ví dụ, biểu diễn) dạng biểu diễn HOA nén theo chế độ mã hóa theo lớp sẽ được mô tả. Cấu trúc dữ liệu này có thể sinh ra từ việc sử dụng phương pháp mã hóa đề xuất và có thể được giải mã (ví dụ, giải nén) bằng cách sử dụng phương pháp giải mã đề xuất.

Cấu trúc dữ liệu có thể bao gồm nhiều tải tin khung HOA tương ứng với các lớp tương ứng trong số nhiều lớp phân cấp. Nhiều tín hiệu truyền có thể được gán (ví dụ, có thể thuộc về) các lớp tương ứng trong số nhiều lớp. Cấu trúc dữ liệu có thể bao gồm tải tin mở rộng HOA tương ứng bao gồm thông tin phụ trợ để tăng cường tham số cho dạng biểu diễn HOA tái tạo thu được từ các tín hiệu truyền được gán cho lớp tương ứng và các lớp bất kỳ thấp hơn lớp tương ứng này. Các tải tin khung HOA và tải tin mở rộng HOA cho nhiều lớp có thể được cung cấp các mức độ bảo vệ chống lỗi tương ứng, như nêu trên đây. Hơn nữa, các tải tin mở rộng HOA có thể bao gồm các phần tử dòng bit chỉ ra ở trên và có thể có `usacExtElementType` của `ID_EXT_ELE_HOA_ENH_LAYER`. Cấu trúc dữ liệu có thể còn bao gồm tải tin mở rộng cấu hình HOA và/hoặc tải tin cấu hình bộ giải mã HOA bao gồm các phần tử dòng bit chỉ ra ở trên.

Cần lưu ý rằng phần mô tả và hình vẽ chỉ minh họa các nguyên lý của phương pháp và thiết bị được đề xuất. Do đó người có hiểu biết trong lĩnh vực này sẽ được hiểu là có thể nghĩ ra các cải tiến khác nhau mà, mặc dù không được mô tả và thể hiện rõ ràng ở đây, thực thi các nguyên lý của sáng chế và được bao gồm trong phạm vi của sáng chế. Hơn thế nữa, tất cả các ví dụ được đề cập về nguyên tắc được dự tính rõ ràng chỉ cho các mục đích sư phạm để giúp người đọc hiểu rõ các nguyên lý của các phương pháp và thiết bị được đề xuất và các ý đồ được đóng góp bởi các tác giả sáng chế để cải tiến lĩnh vực này, và cần được hiểu là không giới hạn ở các ví dụ và điều kiện được đề cập cụ thể này. Ngoài ra, tất cả các nội dung trình bày ở đây đề cập đến các nguyên lý, khía cạnh và phương án của sáng chế, cũng như các ví dụ cụ thể của chúng, được dự tính bao gồm các dạng biến đổi tương đương của chúng.

Phương pháp và thiết bị mô tả trong sáng chế có thể được thực thi dưới dạng phần mềm, phần sụn và/hoặc phần cứng. Một số thành phần nhất định có thể, ví dụ,

được thực thi dưới dạng phần mềm chạy trên bộ xử lý hoặc bộ vi xử lý tín hiệu số. Các thành phần khác có thể, ví dụ, được thực thi dưới dạng phần cứng và hoặc dạng mạch tích hợp chuyên dụng. Các tín hiệu gặp phải trong phương pháp và thiết bị được mô tả có thể được lưu trữ trên phương tiện như bộ nhớ truy cập ngẫu nhiên hoặc phương tiện nhớ quang. Chúng có thể được chuyển qua các loại mạng, như mạng vô tuyến, mạng vệ tinh, mạng không dây hoặc mạng có dây, ví dụ mạng Internet.

Phụ lục:

Các thay đổi dòng bit MPEG-H 3D được đề xuất

Các thay đổi được đánh dấu bằng cách làm nổi bật bằng màu xám

>

Bảng 1 — Cú pháp của mpeg3daExtElementConfig()

Cú pháp	Số bit	lượng Mã nhớ	gọi
<pre> mpeg3daExtElementConfig() { usacExtElementType = escapedValue(4, 8, 16); usacExtElementConfigLength = escapedValue(4, 8, 16); if (usacExtElementDefaultLengthPresent) { usacExtElementDefaultLength = escapedValue(8, 16, 0) + 1; } else { usacExtElementDefaultLength = 0; } usacExtElementPayloadFrag; switch (usacExtElementType) { case ID_EXT_ELE_FILL: /* No configuration element */ break; case ID_EXT_ELE_MPEGS: SpatialSpecificConfig(); break; case ID_EXT_ELE_SAOC: SAOCSpecificConfig(); break; case ID_EXT_ELE_AUDIOPREROLL: </pre>	1	uimsbf	
<pre> usacExtElementPayloadFrag; </pre>	1	uimsbf	

```

    /* No configuration element */
    break;
case ID_EXT_ELE_UNI_DRC:
    mpeg3daUniDrcConfig();
    break;
case ID_EXT_ELE_OBJ_METADATA:
    ObjectMetadataConfig();
    break;
case ID_EXT_ELE_SAOC_3D:
    SAOC3DSpecificConfig();
    break;
case ID_EXT_ELE_HOA:
    HOAConfig();
    break;
case ID_EXT_ELE_HOA_ENH_LAYER:
    HOAEnhConfig();
    break;
case ID_EXT_ELE_FMT_CNRTR
    /* No configuration element */
    break;
default:
    while (usacExtElementConfigLength--) {
        tmp;
    }
    break;
}
}

```

LƯU Ý

8

uimbsf

LƯU Ý: Mục nhập mặc định cho `usacExtElementType` được sử dụng cho các `extElementType` chưa biết để các bộ giải mã kế thừa có thể đối phó với các phần mở rộng thêm trong tương lai.

Bảng 2 - Giá trị của usacExtElementType

usacExtElementType	Giá trị
ID_EXT_ELE_FILL	0
ID_EXT_ELE_MPEGS	1
ID_EXT_ELE_SAOC	2
ID_EXT_ELE_AUDIOPREROLL	3
ID_EXT_ELE_UNI_DRC	4
ID_EXT_ELE_OBJ_METADATA	5
ID_EXT_ELE_SAOC_3D	6
ID_EXT_ELE_HOA	7
ID_EXT_ELE_FMT_CNVTR	8
ID_EXT_ELE_HOA_ENH_LAYER	9
/* reserved for ISO use */	10-127
/* reserved for use outside of ISO scope */	128 and higher
LUU Ý: Các giá trị usacExtElementType chuyên dụng bị bắt buộc trong không gian dành riêng để sử dụng bên ngoài phạm vi ISO. Chúng bị bỏ qua bởi bộ giải mã do cấu trúc tối thiểu được đòi hỏi bởi bộ giải mã bỏ qua các phần mở rộng này.	

Bảng 3 — Diễn giải các khối dữ liệu để giải mã tải tin mở rộng

usacExtElementType	usacExtElementSegmentData ràng buộc thể hiện:
ID_EXT_ELE_FILL	Chuỗi fill_byte
ID_EXT_ELE_MPEGS	SpatialFrame() như được định nghĩa trong ISO/IEC 23003-1
ID_EXT_ELE_SAOC	SAOCFrame() như được định nghĩa trong ISO/IEC 23003-2
ID_EXT_ELE_AUDIOPREROLL	AudioPreRoll()
ID_EXT_ELE_UNI_DRC	uniDrcGain() như được định nghĩa trong ISO/IEC 23003-4

ID_EXT_ELE_OBJ_METADATA	object_metadata()
ID_EXT_ELE_SAOC_3D	Saoc3DFrame()
ID_EXT_ELE_HOA	HOAFrame()
ID_EXT_ELE_HOA_ENH_LAYER	HOAEnhFrame()
ID_EXT_ELE_FMT_CNVRTR	FormatConverterFrame()
unknown	dữ liệu không biết. Khối dữ liệu sẽ được loại bỏ.

Bảng 4 - Cú pháp ChannelSideInfoData(i)

Cú pháp	Số lượng bit	Mã nhớ	gọi
HOADecoderConfig(numHOATransportChannels)			
{			
MinAmbHoaOrder = escapedValue(3,5,0) - 1;	3,8	uimsbf	
MinNumOfCoeffsForAmbHOA = (MinAmbHoaOrder + 1)^2;			
NumOfAdditionalCoders = numHOATransportChannels - MinNumOfCoeffsForAmbHOA;			
if(SingleLayer == 0){	1	bslbf	
HOALayerChBits = ceil(log2(NumOfAdditionalCoders));			
NumHOAChannelsLayer[0] = codedLayerCh +	HOALayer ChBits	uimsb f	
MinNumOfCoeffsForAmbHOA;			
remainingCh = numHOATransportChannels - NumHOAChannelsLayer[0];			
NumLayers = 1;			
while (remainingCh > 1) {			
HOALayerChBits = ceil(log2(remainingCh));			
NumHOAChannelsLayer[NumLayers] =	HOALayer	uimsb	
NumHOAChannelsLayer[NumLayers-1] +	ChBits	f	
codedLayerCh + 1;			
remainingCh = remainingCh -			
NumHOAChannelsLayer[NumLayers];			
NumLayers++;			
}			

if (remainingCh) {		
NumHOAChannelsLayer[NumLayers] =		
NumOfAdditionalCoders;		
NumLayers++;		
}		
}		
MaxNoOfDirSigsForPrediction =	2	uimsbf
MaxNoOfDirSigsForPrediction + 1;		
NoOfBitsPerScalefactor = NoOfBitsPerScalefactor + 1;	4	uimsbf
CodedSpatialInterpolationTime;	3	uimsbf
SpatialInterpolationMethod;	1	bslbf
CodedVVecLength;	2	uimsbf
MaxGainCorrAmpExp;	3	uimsbf
HOAFrameLengthIndicator;	2	uimsbf
MaxHOAOrderToBeTransmitted = escapedValue(2,5,0) +		
MinAmbHoaOrder;		
MaxNumOfCoeffsToBeTransmitted =		
(MaxHOAOrderToBeTransmitted + 1)^2;		
MaxNumAddActiveAmbCoeffs =		
MaxNumOfCoeffsToBeTransmitted		
- MinNumOfCoeffsForAmbHOA;		
VqConfBits = ceil(log2(ceil(log2(NumOfHoaCoeffs))))		
NumVVecVqElementsBits++;	VqConf	uimsbf
	Bits	
if(MinAmbHoaOrder == 1) {		
UsePhaseShiftDecorr;	1	bslbf
}		
if(SingleLayer==1) {		
HOADecoderEnhConfig();	2	uimsbf
}		

```

    AmbAssignmBits = ceil( log2(
MaxNumAddActiveAmbCoeffs ) );
    ActivePredIdsBits = ceil( log2( NumOfHoaCoeffs ) );
    i = 1;
    while( i * ActivePredIdsBits + ceil( log2( i ) ) <
NumOfHoaCoeffs ){
        i++;
    }
    NumActivePredIdsBits = ceil( log2( max( 1, i - 1 ) ) );
    GainCorrPrevAmpExpBits = ceil( log2( ceil( log2(
        1.5 * NumOfHoaCoeffs ) )
        + MaxGainCorrAmpExp + 1 ) );
    for (i=0; i<NumOfAdditionalCoders; ++i){
        AmbCoeffTransitionState[i] = 3;
    }
}

```

LUU Ý: MinAmbHoaOrder = 30 ... 37 được dự trữ. HOAFrameLengthIndicator = 3 được dự trữ.

Bảng mới ? — Cú pháp của HOAEnhConfig()

Cú pháp	Số lượng bit	Mã nhớ	gợi
HOAEnhConfig() { LayerIdx HOALayerChBits HOADecoderEnhConfig(); }			

Bảng mới ? — Cú pháp của HOADecoderEnhConfig()

[illegible]

```

    MaxNumOfPredDirsPerBand = escapedValue(3,2,5) +
1;
    NumOfBitsPerDirIdx =                                2          uimsbf
        NumOfBitsPerDirIdxTable[DirGridTableIdx];
}
if( ParSubbandTableIdx < 3 ) {                          2          uimsbf
    NumOfParSubbands =
        NumOfParSubbandsTable[ParSubbandTableIdx];
    ParSubbandWidths =
        ParSubbandWidthTable[ParSubbandTableIdx];
}
else {
    CodedNumberOfSubbands                                5          uimsbf
    NumOfParSubbands = CodedNumberOfSubbands+1;
    ParSubbandWidths =
        getSubbandWidths(NumOfParSubbands);
}
if( NumOfParSubbands > 0 ) {
    LastFirstOrderSubbandIdxBits =
        ceil( log2(NumOfParSubbands + 1) );
    LastFirstOrderSubbandIdx;                            LastFir  uimsbf
                                                         stOrder
                                                         Subban
                                                         dIdxBit
                                                         s

    for ( idx = 0; idx < NumOfParSubbands; idx++) {
        UseRealCoeffsPerParSubband[idx];                1          bslbf
    }
    for ( idx = 0; idx < LastFirstOrderSubBandIdx; idx++)
{
    UpmixHoaOrderPerParSubband[idx] = 1;

```

```
        MaxNumOfDecoSigs[idx] =  
            (UpmixHoaOrderPerParSubband[idx] + 1)^2;  
    }  
    for ( idx = LastFirstOrderSubBandIdx;  
        idx < NumOfParSubbands; idx++) {  
        UpmixHoaOrderPerParSubband[idx] = 2;  
        MaxNumOfDecoSigs[idx] =  
            (UpmixHoaOrderPerParSubband[idx] + 1)^2;  
    }  
}
```

Bảng 5— Cú pháp của HOAFrame

Cú pháp	Số bit	lượng Mã nhớ	gọi
<pre> HOAFrame() { NumOfDirSigs = 0; for(lay=0; (lay< NumLayers) & !SingleLayer; ++lay){ NumOfDirSigsPerLayer[lay] = 0; NumOfAddHoaChansPerLayer[lay] = 0; NumOfContAddHoaChans[lay] = 0; } NumOfVecSigs = 0; NumOfAddHoaChans = 0; NumOfAddVVecValCoeffIdx = 0; hoaIndependencyFlag; for(i=0; i< NumOfAdditionalCoders; ++i){ ChannelSideInfoData(i); HOAGainCorrectionData(i); switch ChannelType[i] { case 0: DirSigChannelIds[NumOfDirSigs] = i + 1; NumOfDirSigs++; for(lay=0; (lay< NumLayers) & !SingleLayer; ++lay){ if((MinNumOfCoeffsForAmbHOA + i) < NumHOAChannelsLayer[lay]){ NumOfDirSigsPerLayer[lay]++; } } break; case 1: </pre>	1	bslbf	

```

VecSigChannelIds[NumOfVecSigs] = i + 1;
VecSigLayerIdx[NumOfVecSigs] = 0;
if (SingleLayer == 0) {
    lay = 0;
    while( (MinNumOfCoeffsForAmbHOA + i )
           ≥ NumHOAChannelsLayer[lay]) {
        lay ++;
    }
    VecSigLayerIdx[NumOfVecSigs] = lay;
}
NumOfVecSigs++;
break;
case 2:
    if (AmbCoeffTransitionState[i] == 0) {
        for(lay=0; (lay< NumLayers); ++lay){
            if( (MinNumOfCoeffsForAmbHOA + i ) <
                NumHOAChannelsLayer[lay]){
                ContAddHoaCoeff[lay]
                [NumOfContAddHoaChans[lay]] =
                    = AmbCoeffIdx[i];
                NumOfContAddHoaChans[lay]++;
            }
        }
        AddHoaCoeff[NumOfAddHoaChans] =
        AmbCoeffIdx[i];
        for(lay=0; (lay< NumLayers) & (SingleLayer ==
0); ++lay){
            if( (MinNumOfCoeffsForAmbHOA + i ) <
                NumHOAChannelsLayer[lay]){

```

```

AddHoaCoeffPerLayer[lay][NumOfAddHoaChans]
    = AmbCocffIdx[i];
    NumOfAddHoaChansPerLayer[lay]++;
}
}
NumOfAddHoaChans++;
break;
}
}

for ( i= NumOfAdditionalCoders;
      i<NumHOATransportChannels; ++i){
    HOAGainCorrectionData(i);
}

for(i=0; i<NumOfVecSigs; ++i){
    VVectorData ( VecSigChannelIds(i) );
}

if(SingleLayer==1) {
    HOAEnhFrame();
}
}

```

LƯU Ý: bộ mã hóa sẽ thiết lập hoaIndependencyFlag là 1 nếu
 usacIndependencyFlag (see mpegH3daFrame()) được thiết lập là 1.
 LƯU Ý: Nếu SingleLayer == 1 thiết lập NumLayers = 1.

NumOfDirSigsPerLayer[lay] Phần tử này xác định số lượng tín hiệu định hướng
 hoạt động trong HOAFrame() hiện thời được sử dụng
 thực tế trong lớp tăng cường HOA *lay*.
 AddHoaCoeffPerLayer[lay] Mảng này chứa chỉ số hệ số HOA cho mỗi hệ số HOA

xung quanh bổ sung được sử dụng thực tế trong lớp tăng cường HOA *lay*.

NumOfAddHoaChansPerLayer[lay] Phần tử này báo hiệu tổng số lượng hệ số HOA xung quanh bổ sung được sử dụng thực tế trong lớp tăng cường HOA *lay*.

Thêm bảng này

Bảng mới ?— Cú pháp của HOAEnhFrame

Cú pháp	Số lượng Mã gọi bit nhớ
<pre> HOAEnhFrame() { if(((SingleLayer==1) & (NumOfDirSigs > 0)) ((SingleLayer==0) & (NumOfDirSigsPerLayer[lay] > 0))){ HOAPredictionInfo() } if(NumOfPredSubbands > 0) { HOADirectionalPredictionInfo(); } if(NumOfParSubbands > 0) { HOAParInfo(); } } </pre>	
Lưu ý: lay là chỉ số của lớp tăng cường HOA hoạt động hiện thời	

Cập nhật bảng này:

Bảng 6- Cú pháp ChannelSideInfoData(i)

Cú pháp	Số lượng bit	Mã nhó	gợi
<pre> VVectorData(i) { if (CodedVVecLength == 1) { VVecLengthUsed = VVecLength[i]; VVecCoeffIdUsed = VVecCoeffId[i]; } else { VVecLengthUsed = VVecLength; VVecCoeffIdUsed = VVecCoeffId; } if (NbbitsQ(k)[i] == 4){ if (NumVvecIndices(k)[i] == 1) { VecIdx[0] = VvecIdx + 1; WeightVal[0] = ((SgnVal*2)-1); } else { WeightIdx; nbbitsIdx = ceil(log2(NumOfHocCoeffs)); for (j=0; j<NumVvecIndices(k)[i]; ++j) { VvecIdx[j] = VvecIdx + 1; WeightVal[j] = ((SgnVal*2)-1) * } } } WeightValCdbk[CodebkIdx(k)[i]][WeightIdx][j]; } else if (NbbitsQ(k)[i] == 5) { for (m=0; m< VVecLengthUsed; ++m){ aVal[i][m] = (VecVal / 128.0) -- 1.0; } } </pre>	10 1 8 nbbitsIdx 1	uimbsf uimbsf uimbsf uimbsf uimbsf	

```

    }
    else if(NbitsQ(k)[i] >= 6) {
        for (m=0; m< VVecLengthUsed; ++m){
            huffIdx = huffSelect(VVecCoeffIdUsed[m],
PFlag[i],
                                CbFlag[i]);
            cid = huffDecode(NbitsQ[i], huffIdx, huffVal);    động    huffDeco
                                                                de

            aVal[i][m] = 0.0;
            if ( cid > 0 ) {
                aVal[i][m] = sgn = (sgnVal * 2) - 1;          1        bsbf
                if (cid > 1) {
                    aVal[i][m] = sgn * (2.0^(cid -1 ) +        cid-1    uimsbf
intAddVal);
                }
            }
        }
    }
}

```

LƯU Ý: Xem 0 để biết thêm về tính toán VVecLength

Bảng 7 — Cú pháp của HOAPredictionInfo(DirSigChannelIds, NumOfDirSigs)

Cú pháp	Số bit	lượng Mã nhớ	gọi
<pre> HOAPredictionInfo() { if(SingleLayer==1){ PredIdsBits = ceil(log2(NumOfDirSigs + 1)); } else{ PredIdsBits = ceil(log2(NumOfDirSigsPerLayer[lay] + 1)); } if(PSPredictionActive){ NumActivePred = 0; if(KindOfCodedPredIds){ NumActivePred = NumActivePredIds + 1; i=0; while(i < NumActivePred){ PredIds[i] = PredIds[i] + 1; i++; } } else{ for (i=0; i<(HoaOrder +1)^2; i++) { if(ActivePred[i]) { NumActivePred ++; } } } } </pre>	1 1 NumActivePredI dsBits ActivePredIdsBit s 1	bslbf bslbf uimsbf uimsbf	

```

NumOfGains=0;
for (i=0; i<NumActivePred *MaxNoOfDirSigsForPrediction;
i++) {
    if (PredDirSigIds[i]> 0) {
        PredIdsBits      uimsbf
        PredDirSigIds[i]=
            DirSigChannelIds[PredDirSigIds[i]
- 1 ];
        NumOfGains++;
    }
}
n=0;
for (i=0; i< NumOfGains; i++) {
    if (PredDirSigIds[i]>0) {
        PredGains[n];
        NoOfBitsPerScal  bslbf
        efactor

        n++;
    }
}
}
}

```

Lưu ý: lay là chỉ số của lớp tăng cường HOA hoạt động hiện thời

Bảng AMD1.2 - Cú pháp HOADirectionalPredictionInfo()

Cú pháp	Số lượng bit	Mã nhó	gợi
HOADirectionalPredictionInfo ()			
{			
if(UseDirectionalPrediction) {	1	bslbf	
if (!hoaIndependencyFlag) {			
KeepPreviousGlobalPredDirsFlag;	1	bslbf	
}			
else{			
KeepPreviousGlobalPredDirsFlag= 0;			
}			
if(!KeepPreviousGlobalPredDirsFlag) {			
NumOfGlobalPredDirs= NumOfGlobalPredDirs	MaxNum	bslbf	
+ 1;	OfPredDir		
	sLog2		
NumBitsForRelDirGridIdx = ceil(			
log2(NumOfGlobalPredDirs));			
for (idx=0; idx < NumOfGlobalPredDirs; idx++)			
{			
GlobalPredDirsIds[idx];	NumOfBi	uimsbf	
	tsPerDirId		
	x		
}			
}			
else{			
/* Keep values from previous			
HOADirectionalPredictionInfo			
payload for NumOfGlobalPredDirs and			
GlobalPredDirsIds. */			
}			

```

    if(SingleLayer==1){
        SortedAddHoaCoeff = sort(AddHoaCoeff,
                                'ascend');

        NumOfAddHoaChansUsed =
NumOfAddHoaChans;
    }
    else{
        SortedAddHoaCoeff =
sort(AddHoaCoeffPerLayer[lay],
    'ascend');

        NumOfAddHoaChansUsed =
            NumOfAddHoaChansPerLayer[lay];
    }
    for ( band = 0; band < NumOfPredSubbands; band++
) {
        for ( dir = 0; dir < MaxNumOfPredDirsPerBand;
dir++) {
            for ( hoaIdx = 0;
                hoaIdx < MinNumOfCoeffsForAmbHOA;
                hoaIdx++ ) {
                DecodedMagDiff[band][dir][hoaIdx] = 0;
                DecodedAngleDiff[band][dir][hoaIdx] = 0;
            }
        }
    }

    for ( band = 0; band < NumOfPredSubbands; band++
) {
        if (!hoaIndependencyFlag) {
            KeepPreviousDirPredMatrixFlag[band];      1      bs1bf

```

```

    }
    else{
        KeepPreviousDirPredMatrixFlag[band]= 0;
    }
    if (!KeepPreviousDirPredMatrixFlag[band]) {
        UseHuffmanCodingDiffMag;           1          bslbf
        if( band < FirstSBRSubbandIdx ) {
            UseHuffmanCodingDiffAngle;      1          bslbf
            for ( dir = 0; dir <
MaxNumOfPredDirsPerBand; dir++) {
                if ( DirIsActive[band][dir] ) {           1          bslbf
                    RelDirGridIdx;                        NumBitsF   uimsbf
                                                            orRelDir
                                                            GridIdx

                    PredDirGridIdx[band][dir] =

GlobalPredDirsIds[RelDirGridIdx];
                    for ( hoIdx = 0;
                        hoIdx <
MinNumOfCoeffsForAmbHOA;
                            hoIdx++ ) {
                                readDirPredDiffValues (band, dir,
hoIdx,

                                UseHuffmanCodingDiffAbs,
                                UseHuffmanCodingDiffAngle,
                                FirstSBRSubbandIdx);
                            }
                    for ( idx = 0;
                        idx < NumOfAddHoaChansUsed;
                            idx++ ) {

```

```
readDirPredDiffValues (band, dir,  
    SortedAddHoaCoeff[idx] -1,  
    UseHuffmanCodingDiffAbs,  
    UseHuffmanCodingDiffAngle,  
    FirstSBRSubbandIdx);  
    }  
    }  
    }  
    }  
    }  
    }  
    }
```

Lưu ý: lay là chỉ số của lớp tăng cường HOA hoạt động hiện thời

Bảng 8 — SingleLayer definition

Giá trị	Ý nghĩa
0	Tín hiệu HOA được tạo ra trong nhiều lớp; kích hoạt việc báo hiệu sự phân bổ của các kênh truyền HOA vào các lớp khác nhau
1	Tín hiệu HOA được tạo ra trong một lớp

codedLayerCh	Phần tử này chỉ báo cho lớp thứ nhất (tức là cơ sở) số lượng tín hiệu truyền được bao gồm, mà được cho bởi codedLayerCh + MinNumOfCocffsForAmbHOA. Đối với các lớp cao hơn (tức là tăng cường), phần tử này chỉ báo số lượng tín hiệu bổ sung bao gồm vào lớp tăng cường so với lớp thấp hơn tiếp theo, mà được cho bởi codedLayerCh + 1.
HOALayerChBits	Phần tử này chỉ báo số lượng bit để đọc codedLayerCh .
NumLayers	Phần tử này chỉ báo (sau khi đọc HOADecoderConfig()) tổng số lượng lớp trong dòng bit.
NumHOAChannelsLayer	Phần tử này là mảng gồm có các phần tử NumLayers, có phần tử thứ <i>i</i> chỉ báo số lượng tín hiệu truyền bao gồm trong tất cả các lớp lên đến lớp thứ <i>i</i> .

12.4.1.x Khung và các tham số phụ thuộc người dùng

$M_{LAY}(k)$	Số lượng tất cả các lớp được sử dụng thực tế cho khung thứ <i>k</i> (cần được nêu rõ) ở phía bộ giải mã. Trong trường hợp mã hóa theo lớp (được chỉ báo bởi SingleLayer==0) số lượng này phải thấp hơn hoặc bằng tổng số lượng lớp có mặt trong dòng bit, tức là $M_{LAY} \leq \text{NumLayers}$. Trong trường hợp mã hóa đơn lớp (được chỉ báo bởi SingleLayer==1) M_{LAY} được thiết lập là một.
--------------	---

Tùy thuộc vào sự lựa chọn $M_{LAY}(k)$ số lượng $I_{ADD,LAY}(k)$ kênh truyền bổ sung được sử dụng thực tế để giải mã HOA không gian (tức là thêm vào các kênh O_{MIN} mà luôn được ngầm sử dụng) được tính như sau:

```

if(SingleLayer | (!SingleLayer & MLAY(k) == NumLayers))
{
    IADD,LAY(k) = NumOfAdditionalCoders;
}
else
{
    IADD,LAY(k) = NumHOACannelsLayer[MLAY(k) - 1] -
    MinNumOfCoeffsForAmbHOA;
}

```

VVecLength và VVecCoeffId

Từ codedVVecLength chỉ báo:

- 0) Độ dài vector hoàn chỉnh (các phần tử NumOfHoaCoeffs). Chỉ báo rằng tất cả các hệ số cho vector nổi trội (NumOfHoaCoeffs) được nêu rõ.
- 1) Các phần tử vector 1 cho MinNumOfCoeffsForAmbHOA và tất cả các phần tử được xác định trong ContAddHoaCoeff[lay] của lớp hoạt động hiện thời của chỉ số lay=0...NumLayers-1 không được truyền. Đối với chế độ một lớp SingleLayer==1 biến số NumLayers phải được thiết lập bằng một. Chỉ báo rằng chỉ các hệ số của các vector nổi trội tương ứng với số lượng lớn hơn MinNumOfCoeffsForAmbHOA được nêu rõ. Hơn nữa, các hệ số NumOfContAddAmbHoaChan[lay] này được nhận ra trong ContAddAmbHoaChan[lay] bị trừ đi. Danh sách ContAddAmbHoaChan[lay] nêu rõ các kênh bổ sung tương ứng với thứ tự vượt quá thứ tự của MinAmbHoaOrder.
- 2) Phần tử vector 1 cho MinNumOfCoeffsForAmbHOA không được truyền. Chỉ báo rằng các hệ số của các vector nổi trội tương ứng với số lượng lớn hơn MinNumOfCoeffsForAmbHOA được nêu rõ.

Trong trường hợp codedVVecLength==1 cả mảng VVecLength[i] cũng như mảng VVecCoeffId[i][m] 2D đều hợp lệ cho VVector có chỉ số i, trong các trường hợp khác, cả phần tử VVecLength cũng như mảng VVecCoeffId[m] đều hợp lệ cho tất cả VVector trong HOAFrame. Đối với thuật toán gán bên dưới, chức năng giúp đỡ được xác định như sau.

```

switch CodedVVecLength{
    case 0:
        VVecLength = NumOfHoaCoeffs;
        for (m=0; m<VVecLength; ++m) {
            VVecCoeffId[m] = m;
        }
        break;
    case 1:
        for (i=0; i < NumOfVecSigs; ++i) {
            lay = VecSigLayerIdx[i];
            VVecLength[i] = NumOfHoaCoeffs
                -.MinNumOfCoeffsForAmbHOA
                - NumOfContAddHoaChans[lay];
            CoeffIdx = m+1;
            for (m=0; m<VVecLength[i]; ++m) {
                bIsInArray = isMemberOf(CoeffIdx,
                    ContAddHoaCoeff[lay],
                    NumOfContAddHoaChans[lay]);
                while (bIsInArray) {
                    CoeffIdx++;
                    bIsInArray = isMemberOf(CoeffIdx,
                        ContAddHoaCoeff[lay],
                        NumOfContAddHoaChans[lay]);
                }
                VVecCoeffId[i][m] = CoeffIdx-1;
            }
        }
        break;
    case 2:
        VVecLength = NumOfHoaCoeffs – MinNumOfCoeffsForAmbHOA;
        for (m=0; m< VVecLength; ++m) {
            VVecCoeffId[m] = m + MinNumOfCoeffsForAmbHOA;

```

```

    }
}

```

Lệnh chuyển mạch thứ nhất có ba trường hợp (trường hợp từ 0 đến 2) đưa ra cách để xác định độ dài vector nội trội dưới dạng số (VVecLength) và các chỉ số hệ số (VVecCoeffId).

12.4.1.X Chuyển đổi thành phần tử VVec

Loại khử lượng tử hóa của vector V-được báo hiệu bởi từ NbbitsQ. Giá trị NbbitsQ bằng 4 chỉ báo lượng tử hóa vector. Khi NbbitsQ bằng 5, bước khử lượng tử hóa vô hướng 8 bit đồng nhất được thực hiện. Ngược lại, giá trị NbbitsQ lớn hơn hoặc bằng 6 chỉ báo ứng dụng giải mã Huffman cho vector V được lượng tử hóa vô hướng. Chế độ dự báo được ký hiệu là PFlag, trong khi CbFlag biểu thị cho Bảng thông tin bit Huffman.

```

if (CodedVVecLength == 1) {
    VVecLengthUsed = VVecLength[i];
    VVecCoeffIdUsed = VVecCoeffId[i];
} else {
    VVecLengthUsed = VVecLength;
    VVecCoeffIdUsed = VVecCoeffId;
}
if (NbbitsQ(k)[i] == 4) {
    if (NumVVecIndices == 1) {
        for (m=0; m< VVecLengthUsed; ++m) {
            idx = VVecCoeffIdUsed[m];
             $v_{idx}^{(i)}(k) = \text{WeightVal}[0] * \text{VecDict}[900].[\text{VvecIdx}[0]][idx];$ 
        }
    } else {
        cdbLen = 0;
        if (N==4) {
            cdbLen = 32;
        }
        for (m=0; m<O; ++m) {
            TmpVVec[m] = 0;
            for (j=0; j<NumVVecIndices; ++j) {

```

```

        TmpVVec[m] += WeightVal[j] * VecDict[cdbLen].[VvecIdx[j]][m];
    }
}
FNorm = 0.0;
for (m=0; m<O; ++m) {
    FNorm += TmpVVec[m] * TmpVVec[m];
}
FNorm = (N+1)/sqrt(FNorm);
for (m=0; m< VVecLengthUsed; ++m) {
    idx = VVecCoeffldUsed[m];
     $v_{idx}^{(i)}(k) = \text{TmpVVec}[idx] * \text{FNorm};$ 
}
}
}
elseif (NbitsQ(k)[i] == 5) {
    for (m=0; m< VVecLengthUsed; ++m) {
         $v_{VVecCoeffldUsed[m]}^{(i)}(k) = (N+1)*aVal[i][m];$ 
    }
}
elseif (NbitsQ(k)[i] >= 6) {
    for (m=0; m< VVecLengthUsed; ++m) {
         $v_{VVecCoeffldUsed[m]}^{(i)}(k) = (N+1) * (2^{(16 - NbitsQ(k)[i])} * aVal[i][m]) / 2^{15};$ 
        if (PFlag(k)[i] == 1) {
             $v_{VVecCoeffldUsed[m]}^{(i)}(k) += v_{VVecCoeffldUsed[m]}^{(i)}(k - 1);$ 
        }
    }
}
}
}

```

YÊU CẦU BẢO HỘ

1. Phương pháp giải mã dạng biểu diễn Ambisonics bậc cao (Higher Order Ambisonics - HOA) nén của âm thanh hoặc trường âm thanh, phương pháp này bao gồm các bước:

thu dòng bit chứa dạng biểu diễn HOA nén tương ứng với nhiều lớp phân cấp mà bao gồm lớp cơ sở và một hoặc nhiều lớp tăng cường phân cấp, trong đó nhiều lớp đã gán vào đó các thành phần của dạng biểu diễn âm thanh nén cơ bản của âm thanh hoặc trường âm thanh, các thành phần được gán cho các lớp tương ứng trong các nhóm thành phần tương ứng,

xác định lớp dùng được cao nhất trong số nhiều lớp để giải mã;

trích xuất tải tin mở rộng HOA gán cho lớp dùng được cao nhất, trong đó tải tin mở rộng HOA bao gồm thông tin phụ trợ để tăng cường tham số cho dạng biểu diễn HOA tái tạo tương ứng với lớp dùng được cao nhất, trong đó dạng biểu diễn HOA tái tạo tương ứng với lớp dùng được cao nhất có được trên cơ sở các tín hiệu truyền gán cho lớp dùng được cao nhất và các lớp bất kỳ thấp hơn lớp dùng được cao nhất;

giải mã dạng biểu diễn HOA nén tương ứng với lớp dùng được cao nhất dựa trên thông tin lớp, các tín hiệu truyền gán cho lớp dùng được cao nhất và các lớp bất kỳ thấp hơn lớp dùng được cao nhất; và

tăng cường tham số cho dạng biểu diễn HOA được giải mã nhờ sử dụng thông tin phụ trợ bao gồm trong tải tin mở rộng HOA gán cho lớp dùng được cao nhất.

2. Phương pháp theo điểm 1, trong đó thông tin lớp chỉ báo tổng số hệ số HOA xung quanh bổ sung cho lớp tăng cường.

3. Phương pháp theo điểm 1, trong đó thông tin lớp bao gồm các chỉ số hệ số HOA cho mỗi hệ số HOA xung quanh bổ sung cho lớp tăng cường.

4. Phương pháp theo điểm 1, trong đó thông tin lớp bao gồm thông tin tăng cường mà bao gồm ít nhất một trong số Dự báo tín hiệu không gian, Tổng hợp tín hiệu định hướng dải con và bộ giải mã tái tạo môi trường tham số.

5. Phương pháp theo điểm 1, phương pháp này còn bao gồm các phần tử vector V mà không được truyền đổi với các chỉ số mà bằng các chỉ số của các hệ số HOA bổ sung bao gồm trong tập `ContAddHoaCoeff`.
6. Phương pháp theo điểm 1, trong đó thông tin lớp bao gồm các phần tử `NumLayers`, trong đó mỗi phần tử chỉ báo số lượng các tín hiệu truyền bao gồm trong tất cả các lớp lên đến lớp thứ i .
7. Phương pháp theo điểm 1, trong đó thông tin lớp bao gồm chỉ báo của tất cả các lớp được sử dụng thực tế cho khung thứ k .
8. Phương pháp theo điểm 1, trong đó thông tin lớp chỉ báo rằng tất cả các hệ số đối với các vector nổi trội được nêu rõ.
9. Phương pháp theo điểm 1, trong đó thông tin lớp chỉ báo rằng các hệ số của các vector nổi trội tương ứng với số lượng lớn hơn `MinNumOfCoeffsForAmbHOA` được nêu rõ.
10. Phương pháp theo điểm 1, trong đó thông tin lớp chỉ báo `MinNumOfCoeffsForAmbHOA` và tất cả các phần tử được xác định trong `ContAddHoaCoeff[lay]` không được truyền, trong đó lay là chỉ số của lớp chứa tín hiệu dựa trên vector tương ứng với vector này.
11. Thiết bị giải mã dạng biểu diễn Ambisonics bậc cao (HOA) nén của âm thanh hoặc trường âm thanh, thiết bị này bao gồm:
 - bộ thu được tạo cấu hình để thu dòng bit chứa dạng biểu diễn HOA nén tương ứng với nhiều lớp phân cấp mà bao gồm lớp cơ sở và một hoặc nhiều lớp tăng cường phân cấp, trong đó nhiều lớp đã gán vào đó các thành phần của dạng biểu diễn âm thanh nén cơ bản của âm thanh hoặc trường âm thanh, các thành phần được gán cho các lớp tương ứng trong các nhóm thành phần tương ứng,
 - bộ giải mã được tạo cấu hình để:
 - xác định lớp dùng được cao nhất trong số nhiều lớp để giải mã;

trích xuất tải tin mở rộng HOA gán cho lớp dùng được cao nhất, trong đó tải tin mở rộng HOA bao gồm thông tin phụ trợ để tăng cường tham số cho dạng biểu diễn HOA tái tạo tương ứng với lớp dùng được cao nhất, trong đó dạng biểu diễn HOA tái tạo tương ứng với lớp dùng được cao nhất có được trên cơ sở các tín hiệu truyền gán cho lớp dùng được cao nhất và các lớp bất kỳ thấp hơn lớp dùng được cao nhất;

giải mã dạng biểu diễn HOA nén tương ứng với lớp dùng được cao nhất dựa trên thông tin lớp, các tín hiệu truyền gán cho lớp dùng được cao nhất và các lớp bất kỳ thấp hơn lớp dùng được cao nhất; và

tăng cường tham số cho dạng biểu diễn HOA được giải mã nhờ sử dụng thông tin phụ trợ bao gồm trong tải tin mở rộng HOA gán cho lớp dùng được cao nhất.

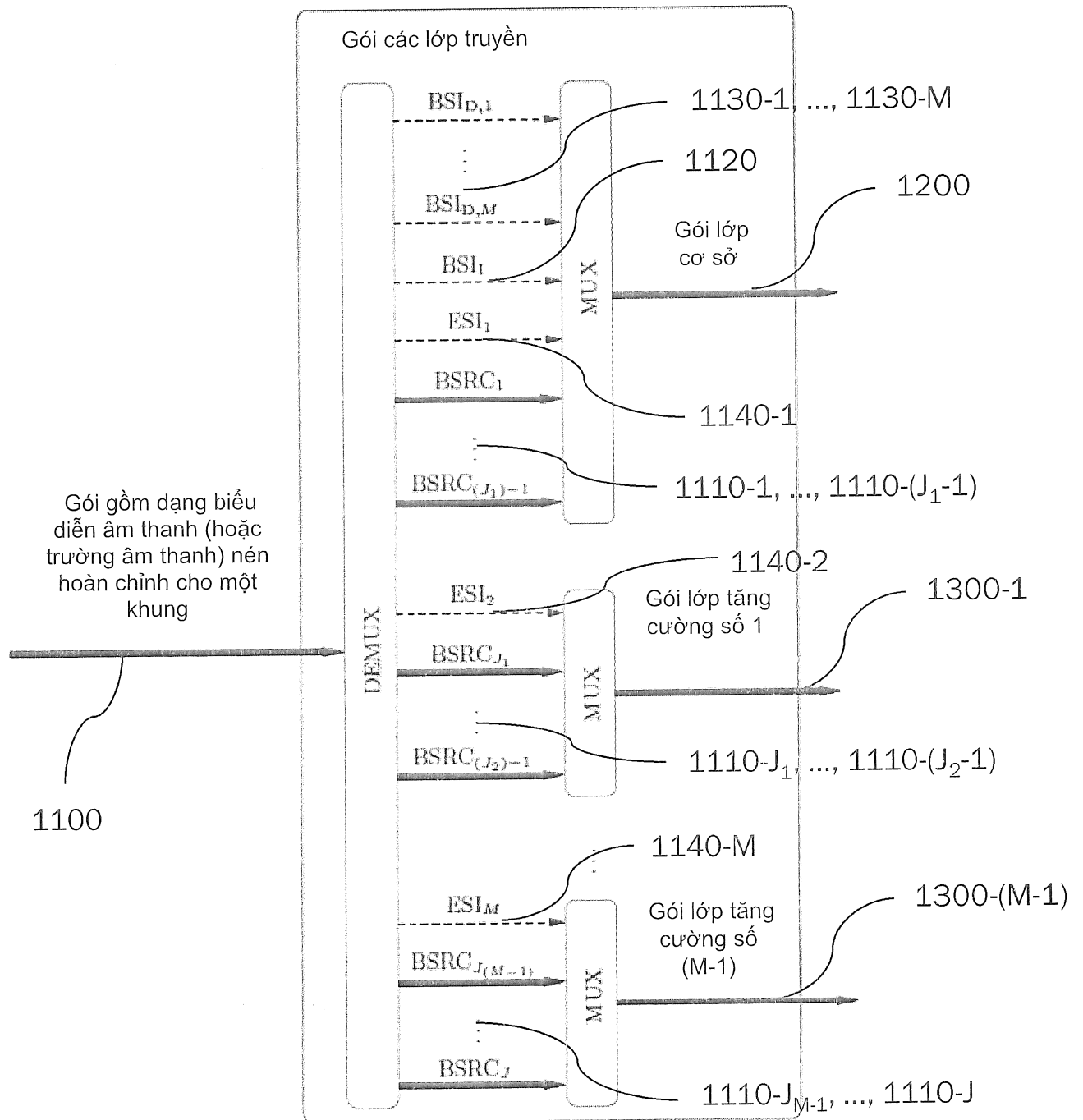


Fig. 1

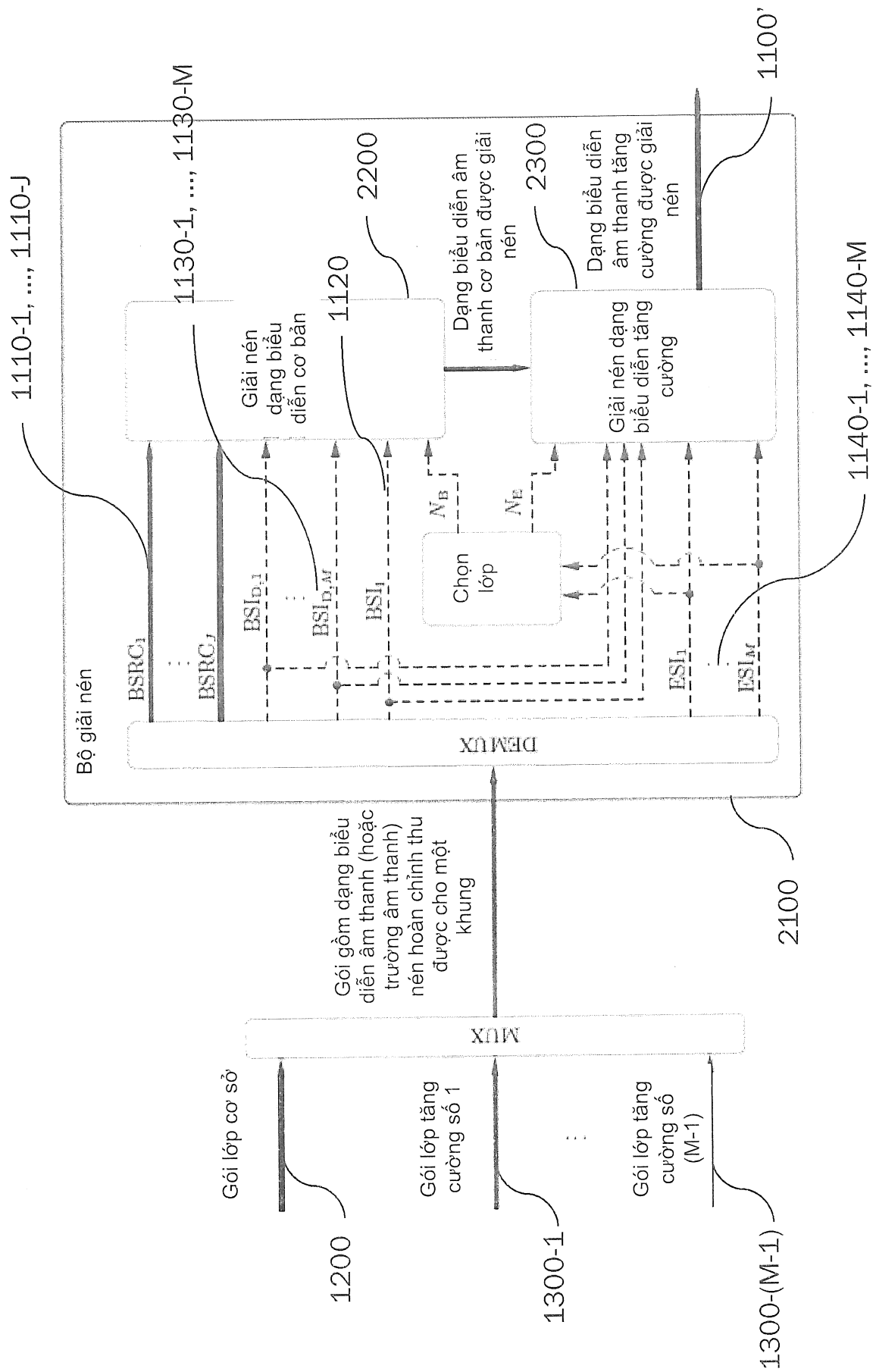


Fig. 2

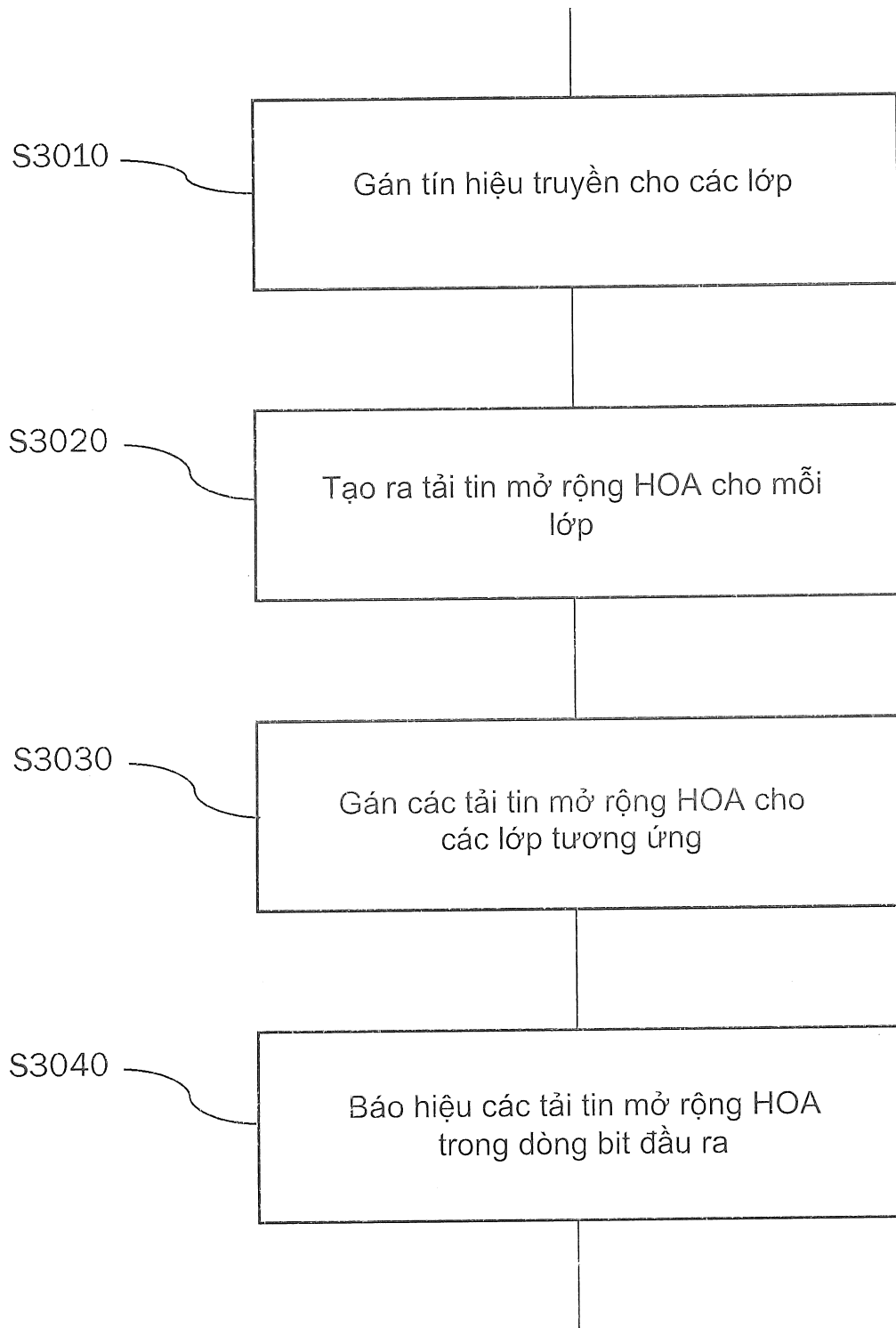


Fig. 3

4/7

S4010

Gán tín hiệu truyền cho các lớp

S4020

Xác định xem chế độ mã hóa vector
có hoạt động không

S4030

Xác định tập hệ số HOA liên tục cho
mỗi lớp

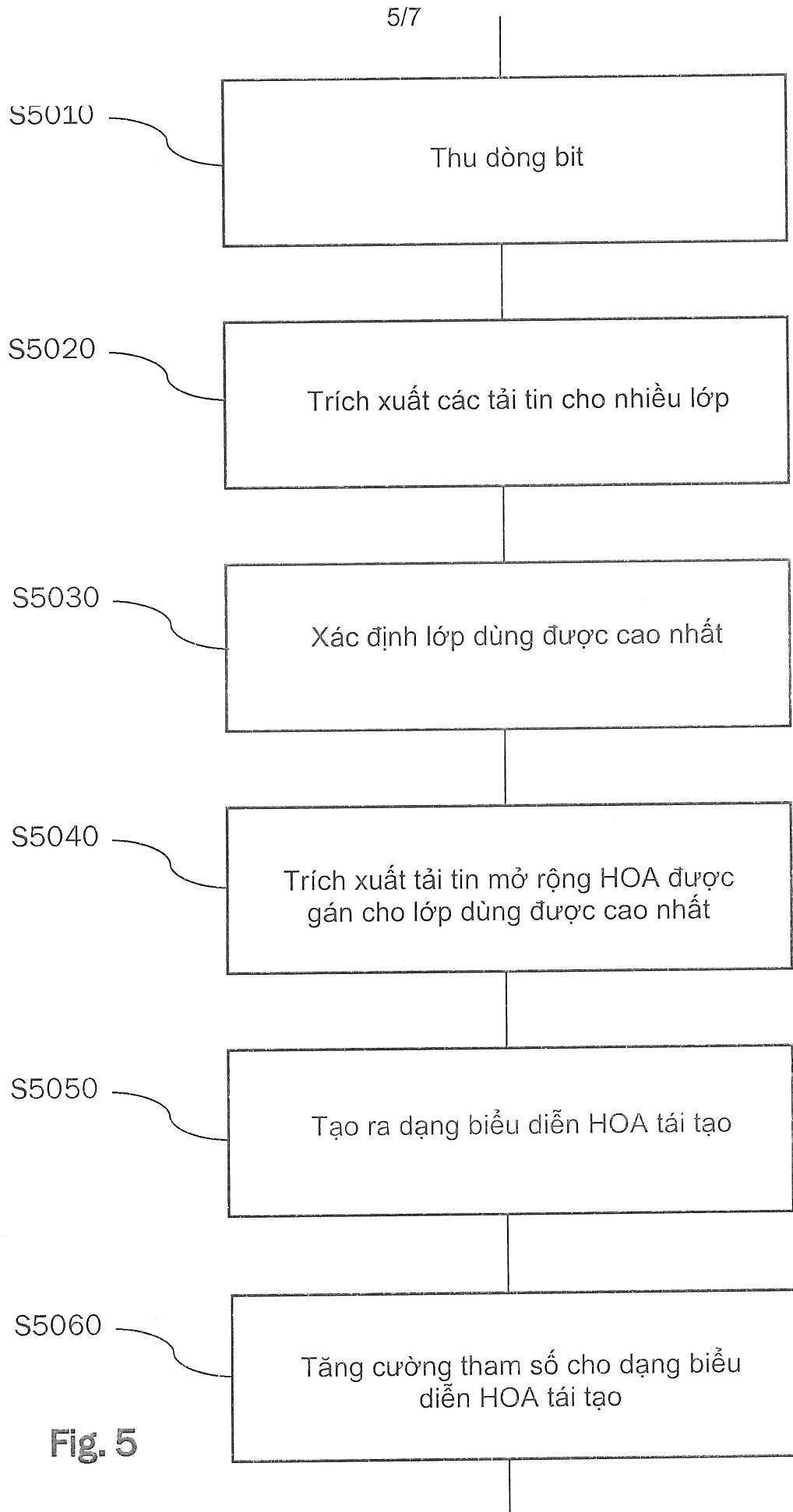
S4040

Tạo ra vector V cho mỗi tín hiệu
truyền

S4050

Báo hiệu vector V trong dòng bit đầu
ra

Fig. 4



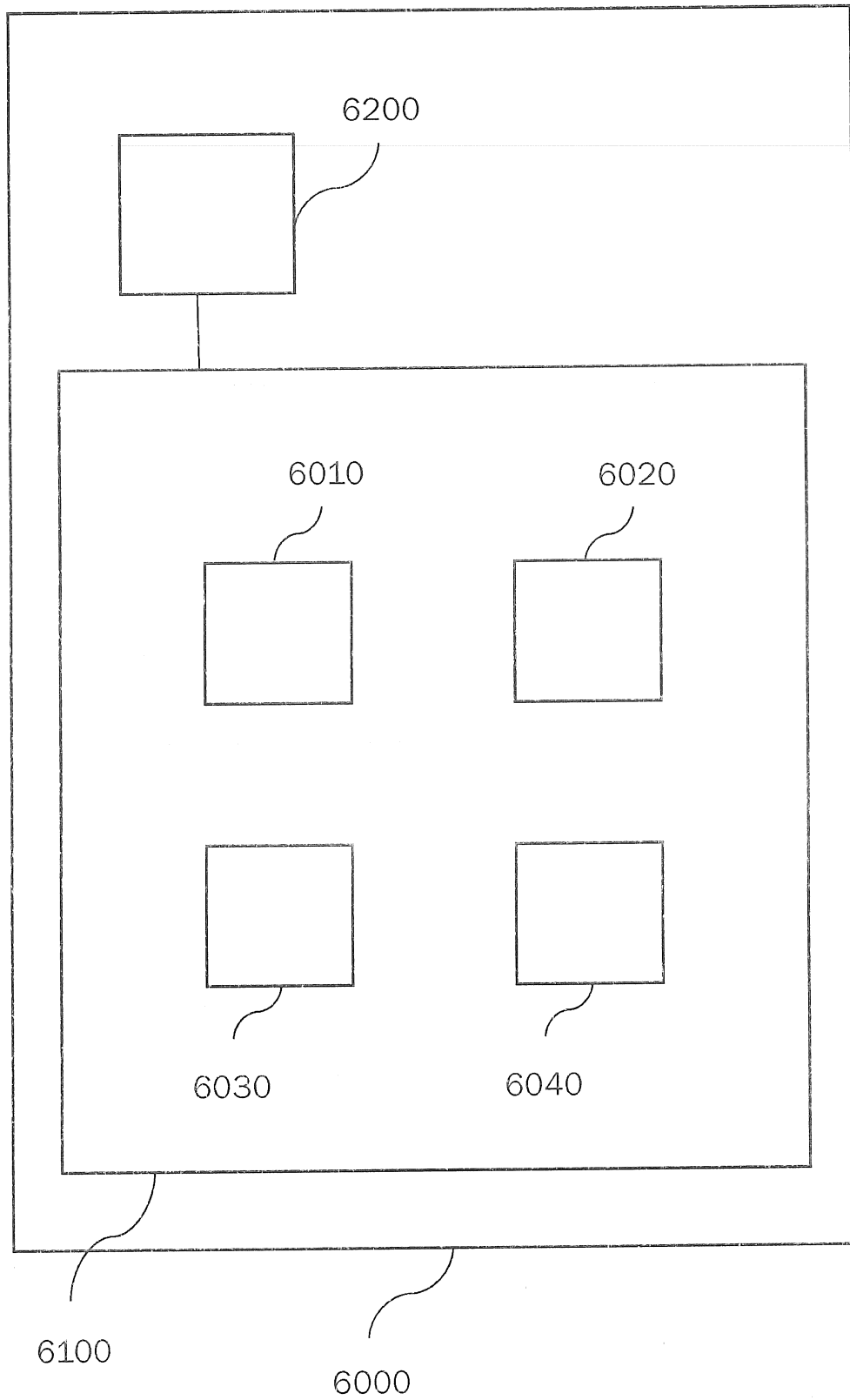


Fig. 6

7/7

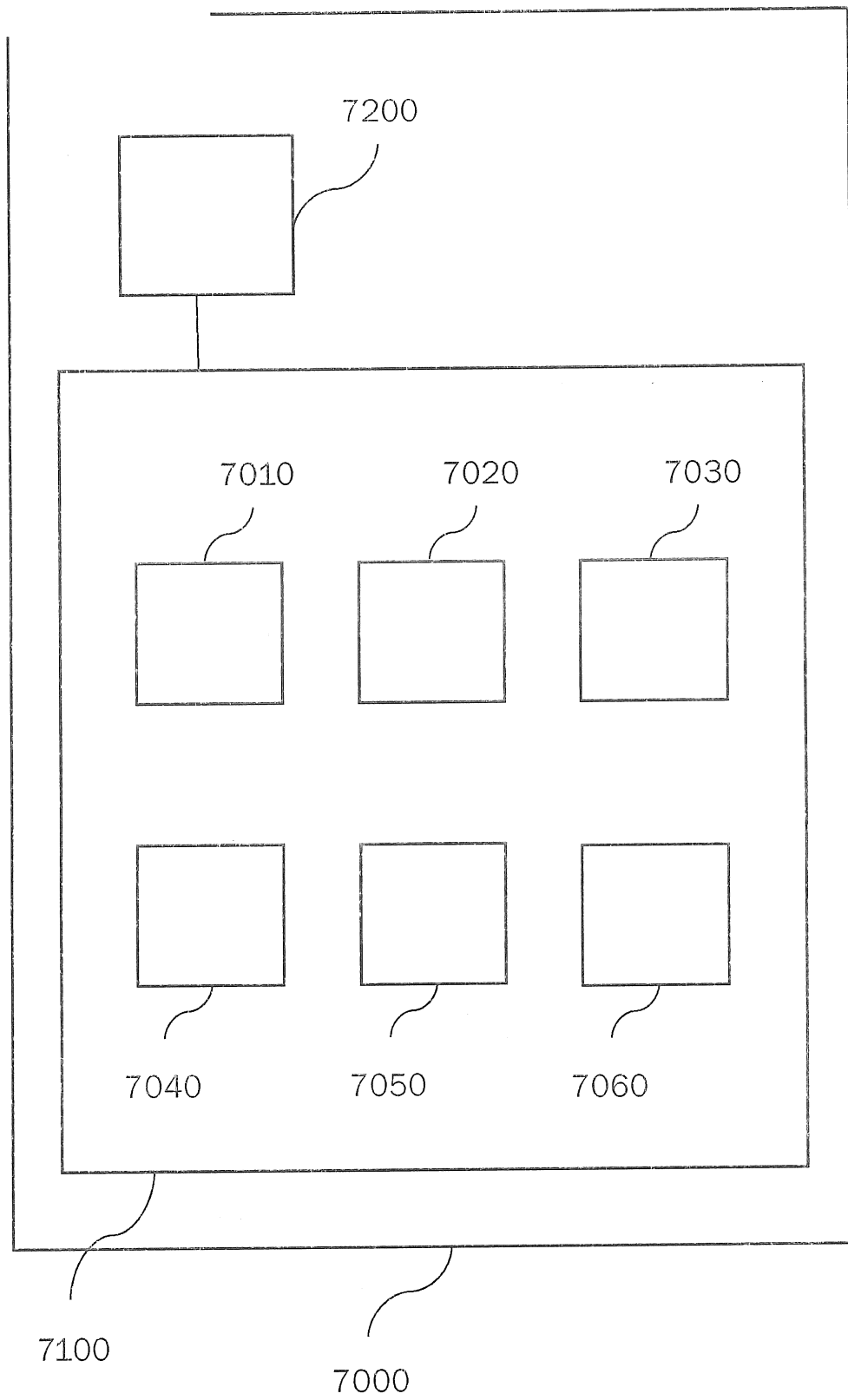


Fig. 7