



(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)



1-0035297

(51)⁷ G10L 19/008; G10L 19/24

(13) B

(21) 1-2018-01905

(22) 07/10/2016

(86) PCT/EP2016/073970 07/10/2016

(87) WO 2017/060411 A1 13/04/2017

(30) 15306590.9 08/10/2015 EP; 62/361,809 13/07/2016 US

(45) 25/04/2023 421

(43) 25/07/2018 364A

(73) Dolby International AB (NL)

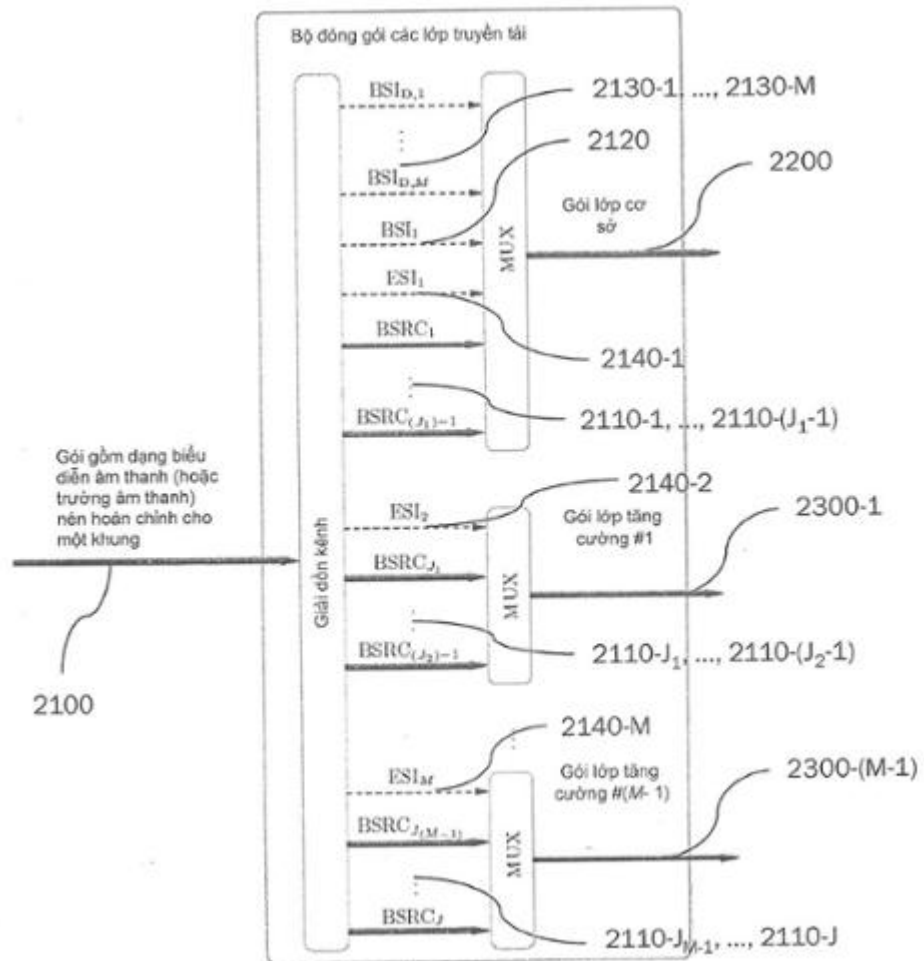
Apollo Building, 3E, Herikerbergweg 1-35, 1101 CN Amsterdam Zuidoost,
Netherlands

(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, VÀ PHƯƠNG TIỆN BẤT BIẾN ĐỌC ĐƯỢC BẰNG MÁY TÍNH

(57) Sáng chế đề cập đến phương pháp giải mã dạng biểu diễn âm thanh ambisonics bậc cao (Higher Order Ambisonics - HOA) nén của âm thanh hoặc trường âm thanh, trong đó dạng biểu diễn âm thanh nén được mã hóa trong nhiều lớp phân cấp bao gồm lớp cơ sở và một hoặc nhiều lớp tăng cường phân cấp, đồng thời đề cập đến thiết bị giải mã dạng biểu diễn ambisonics bậc cao (HOA) nén của âm thanh hoặc trường âm thanh. Ngoài ra, sáng chế cũng bộc lộ phương tiện bất biến đọc được bằng máy tính chứa các lệnh đọc được bằng máy tính mà, khi được thực thi bởi một hoặc nhiều bộ xử lý của thiết bị tính toán, sẽ khiến cho thiết bị tính toán thực hiện phương pháp nêu trên.



Lĩnh vực kỹ thuật được đề cập

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

Tình trạng kỹ thuật của sáng chế

Đối với việc phát trực tiếp dạng biểu diễn âm thanh (hoặc trường âm thanh) qua kênh truyền có các điều kiện thay đổi theo thời gian, mã hóa theo lớp là phương tiện để điều chỉnh chất lượng của dạng biểu diễn âm thanh thu được theo các điều kiện truyền, và cụ thể là để tránh không bị mất tín hiệu không mong muốn.

Để mã hóa theo lớp, dạng biểu diễn âm thanh (hoặc trường âm thanh) thường được chia nhỏ thành lớp cơ sở ưu tiên cao có kích thước tương đối nhỏ và các lớp tăng cường bổ sung có các mức ưu tiên giảm và kích thước tùy ý. Mỗi lớp tăng cường thường được tính toán để chứa thông tin gia tăng để bù cho tất cả các lớp thấp hơn để cải thiện chất lượng của dạng biểu diễn âm thanh (hoặc trường âm thanh). Mức bảo vệ chống lỗi cho cuộc truyền các lớp riêng được kiểm soát dựa trên mức ưu tiên của chúng. Cụ thể là, lớp cơ sở có mức bảo vệ chống lỗi cao là hợp lý và có thể chấp nhận được do kích thước nhỏ của nó.

Tuy nhiên, cần có sơ đồ mã hóa theo lớp cho (các phiên bản mở rộng của) các loại đặc biệt của các dạng biểu diễn nén của âm thanh hoặc trường âm thanh, chẳng hạn như các dạng biểu diễn âm thanh hoặc trường âm thanh HOA nén.

Sáng chế giải quyết các vấn đề trên. Cụ thể là, sáng chế đề cập đến các phương pháp và bộ mã hóa/bộ giải mã để mã hóa theo lớp các dạng biểu diễn âm thanh hoặc trường âm thanh được nén.

Bản chất kỹ thuật của sáng chế

Theo một khía cạnh, sáng chế đề xuất phương pháp mã hóa theo lớp dạng biểu diễn âm thanh nén của âm thanh hoặc trường âm thanh. Dạng biểu diễn âm thanh nén này có thể bao gồm dạng biểu diễn âm thanh nén cơ bản có nhiều thành phần. Các thành phần này có thể là các thành phần bù. Dạng biểu diễn âm thanh nén còn có thể bao gồm thông tin phụ trợ cơ bản để giải mã dạng biểu diễn âm thanh nén cơ bản thành dạng biểu diễn âm thanh tái tạo cơ bản của âm thanh hoặc trường âm thanh. Dạng biểu diễn âm thanh nén còn có thể bao gồm thông tin phụ trợ tăng cường bao gồm các thông số để cải thiện (ví dụ, tăng cường) dạng biểu diễn âm thanh tái tạo cơ bản. Phương pháp này có thể bao gồm bước chia nhỏ (ví dụ, phân nhóm) các thành phần thành nhiều nhóm thành phần. Phương pháp này còn có thể bao gồm bước gán (ví dụ, thêm) mỗi trong số nhiều nhóm cho một trong số nhiều lớp phân cấp tương ứng. Phép gán này có thể chỉ báo sự tương quan giữa các nhóm và lớp tương ứng. Các thành phần được gán cho lớp tương ứng có thể gọi là được bao gồm trong lớp đó. Số lượng nhóm có thể tương ứng với (ví dụ, bằng) số lượng lớp. Các lớp có thể bao gồm 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 này còn có thể bao gồm bước thêm thông tin phụ trợ cơ bản vào lớp cơ sở (ví dụ, bao gồm thông tin phụ trợ cơ bản trong lớp cơ sở, hoặc phân bổ thông tin phụ trợ cơ bản cho lớp cơ sở, ví dụ để truyền hoặc lưu trữ). Phương pháp này còn có thể bao gồm bước xác định các phần thông tin phụ trợ tăng cường từ thông tin phụ trợ tăng cường. Phương pháp này còn có thể bao gồm bước gán (ví dụ, thêm) mỗi trong số các phần thông tin phụ trợ tăng cường cho một lớp tương ứng trong số các lớp. Mỗi phần thông tin phụ trợ tăng cường có thể bao gồm các thông số để cải thiện dạng biểu diễn âm thanh tái tạo (ví dụ, giải nén) có thể thu được từ dữ liệu được bao gồm trong (ví dụ, được gán cho hoặc được thêm vào) lớp tương ứng và lớp bất kỳ thấp hơn lớp tương ứng này. Phương pháp mã hóa theo lớp có thể được thực hiện để truyền qua kênh truyền hoặc lưu trữ trong vật ghi thích hợp, chẳng hạn như đĩa CD, DVD hoặc Blu-ray Disc™.

Với cấu hình như vậy, phương pháp được đề xuất cho phép áp dụng kỹ thuật mã hóa theo lớp một cách hiệu quả cho các dạng biểu diễn âm thanh nén bao gồm các thành

phần cũng như thông tin phụ trợ thứ nhất và tăng cường (ví dụ, thông tin phụ trợ cơ bản độc lập và thông tin phụ trợ tăng cường) có các đặc điểm nêu trên. Cụ thể là, phương pháp được đề xuất đảm bảo rằng mỗi lớp bao gồm thông tin phụ trợ thích hợp để tái tạo dạng biểu diễn âm thanh tái tạo từ các thành phần được bao gồm trong lớp bất kỳ lên đến lớp đã nêu. Trong đó các lớp lên đến lớp đã nêu 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 đã nêu. Do đó, bất kể lớp dùng được cao nhất thực tế (ví dụ, lớp ở dưới lớp dưới cùng không được nhận một cách hợp lệ, sao cho tất cả các lớp bên dưới lớp dùng được cao nhất và chính lớp dùng được cao nhất đã được nhận một cách 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 có thể khác với dạng biểu diễn âm thanh hoàn chỉnh (ví dụ, đầy đủ). Cụ thể là, bất kể lớp dùng được cao nhất thực tế, bộ giải mã chỉ cần giải mã tải tin của thông tin phụ trợ tăng cường 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 có thể thu được trên cơ sở tất cả các thành phần chứa trong các lớp lên đến lớp dùng được cao nhất thực tế. Tức là, đối với mỗi khoảng thời gian (ví dụ, khung) thì chỉ có một tải tin duy nhất của thông tin phụ trợ tăng cường cần phải được 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 các phương án, các thành phần của dạng biểu diễn âm thanh nén cơ bản có thể tương ứng với tín hiệu nghe một tai (ví dụ, tín hiệu truyền hoặc tín hiệu truyền một tai). Tín hiệu nghe một tai có thể là tín hiệu âm thanh trội hoặc chuỗi hệ số của dạng biểu diễn HOA. Tín hiệu nghe một tai có thể được lượng tử hóa.

Theo các phương án, thông tin phụ trợ cơ bản có thể bao gồm thông tin định rõ việc giải mã (ví dụ, giải nén) một hoặc nhiều trong số nhiều thành phần riêng, độc lập với các thành phần khác. Ví dụ, thông tin phụ trợ cơ bản có thể biểu diễn thông tin phụ trợ liên quan đến tín hiệu nghe một tai riêng, độc lập với tín hiệu nghe một tai khác. Do đó, thông tin phụ trợ cơ bản có thể được gọi là thông tin phụ trợ cơ bản độc lập.

Theo các phương án, thông tin phụ trợ tăng cường có thể là thông tin phụ trợ tăng cường. Thông tin phụ trợ tăng cường có thể bao gồm các thông số dự đoán dành cho

dạng biểu diễn âm thanh nén cơ bản để cải thiện (ví dụ, tăng cường) dạng biểu diễn âm thanh tái tạo cơ bản mà có thể thu được từ dạng biểu diễn âm thanh nén cơ bản và thông tin phụ trợ cơ bản.

Theo các phương án, phương pháp còn có thể bao gồm bước tạo dòng truyền để truyền dữ liệu của các lớp (ví dụ, dữ liệu được gán hoặc được thêm vào các lớp tương ứng, hoặc nếu không thì được bao gồm trong các lớp tương ứng). Lớp cơ sở có thể có mức độ ưu tiên truyền cao nhất và các lớp tăng cường phân cấp có thể có mức độ ưu tiên truyền giảm. Tức là, mức độ ưu tiên truyền có thể giảm từ lớp cơ sở đến lớp tăng cường thứ nhất, từ lớp tăng cường thứ nhất đến lớp tăng cường thứ hai, v.v.. Mức bảo vệ chống lỗi khi truyền của dữ liệu của các lớp có thể được kiểm soát theo mức độ ưu tiên truyền tương ứng. Theo cách đó, có thể bảo đảm rằng ít nhất một số lớp thấp hơn chắc chắn được truyền, trong khi đó mặc 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, phương pháp còn có thể bao gồm, đối với mỗi trong số các lớp, bước tạo gói lớp truyền tải bao gồm dữ liệu của lớp tương ứng. Ví dụ, đối với mỗi khoảng thời gian (ví dụ, khung), gói lớp truyền tải tương ứng có thể được tạo cho mỗi lớp.

Theo các phương án, dạng biểu diễn âm thanh nén còn có thể bao gồm thông tin phụ trợ cơ bản bổ sung để giải mã dạng biểu diễn âm thanh nén cơ bản thành dạng biểu diễn âm thanh tái tạo cơ bản. Thông tin phụ trợ cơ bản bổ sung có thể bao gồm thông tin định rõ việc giải mã một hoặc nhiều trong số nhiều thành phần phụ thuộc vào các thành phần tương ứng khác. Phương pháp này còn có thể bao gồm bước phân tích thông tin phụ trợ cơ bản bổ sung thành các phần thông tin phụ trợ cơ bản bổ sung. Phương pháp này còn có thể bao gồm bước thêm các phần thông tin phụ trợ cơ bản bổ sung vào lớp cơ sở (ví dụ, bao gồm các phần thông tin phụ trợ cơ bản bổ sung trong lớp cơ sở, hoặc phân bổ các phần thông tin phụ trợ cơ bản bổ sung cho lớp cơ sở, ví dụ để truyền hoặc lưu trữ). Mỗi phần thông tin phụ trợ cơ bản bổ sung có thể tương ứng với lớp tương ứng và có thể bao gồm thông tin định rõ việc giải mã một hoặc nhiều thành phần được gán cho lớp tương ứng (chỉ) phụ thuộc vào các thành phần tương ứng khác được gán cho lớp tương ứng và lớp bất kỳ thấp hơn lớp tương ứng. Tức là, mỗi phần thông tin phụ

trợ cơ bản bổ sung định rõ các thành phần trong lớp tương ứng mà phần thông tin phụ trợ cơ bản bổ sung đó tương ứng với nó mà không tham chiếu đến các thành phần bất kỳ khác được gán cho các lớp cao hơn lớp tương ứng.

Với cấu hình như vậy, phương pháp được đề xuất tránh việc phân mảnh thông tin phụ trợ cơ bản bổ sung bằng cách thêm tất cả các phần vào lớp cơ sở. Nói cách khác, tất cả các phần thông tin phụ trợ cơ bản bổ sung được bao gồm trong lớp cơ sở. Việc phân tích thông tin phụ trợ cơ bản bổ sung đảm bảo rằng, đối với mỗi lớp, một phần thông tin phụ trợ cơ bản bổ sung đều khả dụng mà không cần biết đến các thành phần ở các lớp cao hơn. Do đó, bất kể lớp dùng được cao nhất thực tế, thì như thế cũng đủ để bộ giải mã giải mã thông tin phụ trợ cơ bản bổ sung chứa trong các lớp lên đến lớp dùng được cao nhất.

Theo các phương án, thông tin phụ trợ cơ bản bổ sung có thể bao gồm thông tin định rõ việc giải mã (ví dụ, giải nén) một hoặc nhiều trong số nhiều thành phần phụ thuộc vào các thành phần khác. Ví dụ, thông tin phụ trợ cơ bản bổ sung có thể biểu diễn thông tin phụ trợ liên quan đến tín hiệu nghe một tai riêng, độc lập với các tín hiệu nghe một tai khác. Do đó, thông tin phụ trợ cơ bản bổ sung có thể được gọi là thông tin phụ trợ cơ bản phụ thuộc.

Theo các phương án, dạng biểu diễn âm thanh nén có thể được xử lý trong các khoảng thời gian liên tiếp, ví dụ các khoảng thời gian có cùng kích thước. Các khoảng thời gian liên tiếp có thể là các khung. Do đó, phương pháp có thể hoạt động trên cơ sở mỗi khung, tức là, dạng biểu diễn âm thanh nén có thể được mã hóa theo từng khung. Dạng biểu diễn âm thanh nén có thể khả dụng trong mỗi khoảng thời gian liên tiếp (ví dụ, đối với mỗi khung). Tức là, hoạt động nén mà qua đó thu được dạng biểu diễn âm thanh nén có thể hoạt động trên cơ sở mỗi khung.

Theo các phương án, phương pháp còn có thể bao gồm bước tạo thông tin cấu hình mà chỉ báo, đối với mỗi lớp, các thành phần của dạng biểu diễn âm thanh nén cơ bản được gán cho lớp đó. Do đó, bộ giải mã có thể sẵn sàng truy cập thông tin cần cho quá trình giải mã mà không phải phân tích cú pháp không cần thiết qua các tải tin dữ liệu thu được.

Theo khía cạnh khác, sáng chế đề xuất phương pháp mã hóa theo lớp dạng biểu diễn âm thanh nén của âm thanh hoặc trường âm thanh. Dạng biểu diễn âm thanh nén có thể bao gồm dạng biểu diễn âm thanh nén cơ bản có nhiều thành phần. Các thành phần này có thể là các thành phần bù. Dạng biểu diễn âm thanh nén còn có thể bao gồm thông tin phụ trợ cơ bản (ví dụ, thông tin phụ trợ cơ bản độc lập) và thông tin thứ ba (ví dụ, thông tin phụ trợ cơ bản phụ thuộc) để giải mã dạng biểu diễn âm thanh nén cơ bản thành dạng biểu diễn âm thanh tái tạo cơ bản của âm thanh hoặc trường âm thanh. Thông tin phụ trợ cơ bản có thể bao gồm thông tin định rõ việc giải mã một hoặc nhiều trong số nhiều thành phần riêng, độc lập với các thành phần khác. Thông tin phụ trợ cơ bản bổ sung có thể bao gồm thông tin định rõ việc giải mã một hoặc nhiều trong số nhiều thành phần phụ thuộc vào các thành phần tương ứng khác. Phương pháp này có thể bao gồm bước chia nhỏ (ví dụ, phân nhóm) các thành phần thành nhiều nhóm thành phần. Phương pháp này còn có thể bao gồm bước gán (ví dụ, thêm) mỗi trong số nhiều nhóm cho một lớp phân cấp tương ứng. Phép gán này có thể chỉ báo sự tương quan giữa các nhóm và lớp tương ứng. Các thành phần được gán cho lớp tương ứng có thể gọi là được bao gồm trong lớp đó. Số lượng nhóm có thể tương ứng với (ví dụ, bằng) số lượng lớp. Các lớp có thể bao gồm một lớp cơ sở và một hoặc nhiều lớp tăng cường phân cấp. Phương pháp này còn có thể bao gồm bước thêm thông tin phụ trợ cơ bản vào lớp cơ sở (ví dụ, bao gồm thông tin phụ trợ cơ bản trong lớp cơ sở, hoặc phân bổ thông tin phụ trợ cơ bản cho lớp cơ sở, ví dụ để truyền hoặc lưu trữ). Phương pháp này còn có thể bao gồm bước phân tích thông tin phụ trợ cơ bản bổ sung thành các phần thông tin phụ trợ cơ bản bổ sung và thêm các phần thông tin phụ trợ cơ bản bổ sung vào lớp cơ sở (ví dụ, bao gồm các phần thông tin phụ trợ cơ bản bổ sung trong lớp cơ sở, hoặc phân bổ các phần thông tin phụ trợ cơ bản bổ sung cho lớp cơ sở, ví dụ để truyền hoặc lưu trữ). Mỗi phần thông tin phụ trợ cơ bản bổ sung có thể tương ứng với lớp tương ứng và bao gồm thông tin định rõ việc giải mã một hoặc nhiều thành phần được gán cho lớp tương ứng phụ thuộc vào các thành phần tương ứng khác được gán cho lớp tương ứng và lớp bất kỳ thấp hơn lớp tương ứng.

Với cấu hình như vậy, phương pháp được đề xuất đảm bảo rằng đối với mỗi lớp, thông tin phụ trợ cơ bản bổ sung phù hợp là khả dụng để giải mã các thành phần có trong lớp bất kỳ lên đến lớp tương ứng, mà không cần việc nhận hợp lệ hoặc giải mã (hoặc

nói chung là biết đến) các lớp cao hơn bất kỳ. Trong trường hợp của dạng biểu diễn HOA nén, phương pháp được đề xuất bảo đảm rằng trong chế độ mã hóa vector, vector V thích hợp là khả dụng đối với tất cả các thành phần thuộc các lớp lên đến lớp dùng được cao nhất. Cụ thể, phương pháp được đề xuất loại trừ trường hợp là các phần tử của vector V tương ứng với các thành phần 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 (ví dụ, giải nén) thành phần bất kỳ thuộc các lớp lên đến lớp dùng được cao nhất. Theo cách đó, việc giải nén thích hợp các dạng biểu diễn HOA tái tạo tương ứng đối với các lớp thấp hơn được đảm bảo ngay cả khi 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.

Các phương án của khía cạnh này có thể liên quan đến các phương án của khía cạnh nêu trên.

Theo khía cạnh khác, sáng chế đề xuất phương pháp giải mã dạng biểu diễn âm thanh nén của âm thanh hoặc trường âm thanh. Dạng biểu diễn âm thanh nén có thể đã được mã hóa trong các lớp phân cấp. Các lớp phân cấp có thể bao gồm một lớp cơ sở và một hoặc nhiều lớp tăng cường phân cấp. Các lớp có thể gán vào đó các thành phần của dạng biểu diễn âm thanh nén cơ bản của âm thanh hoặc trường âm thanh. Nói cách khác, các lớp có thể bao gồm các thành phần của thông tin phụ trợ cơ bản đã được nén. Các thành phần có thể được gán cho các lớp tương ứng trong các nhóm thành phần tương ứng. Các thành phần này có thể là các thành phần bù. Lớp cơ sở có thể bao gồm thông tin phụ trợ cơ bản để giải mã dạng biểu diễn âm thanh nén cơ bản. Mỗi lớp có thể bao gồm một phần thông tin phụ trợ tăng cường bao gồm các thông số để cải thiện dạng biểu diễn âm thanh tái tạo cơ bản có thể thu được từ dữ liệu có trong lớp tương ứng và lớp bất kỳ thấp hơn lớp tương ứng. Phương pháp này có thể bao gồm bước nhận các tải tin dữ liệu lần lượt tương ứng với các lớp phân cấp. Phương pháp này còn có thể bao gồm bước xác định chỉ số lớp thứ nhất chỉ báo lớp dùng được cao nhất trong số các lớp sẽ dùng để giải mã dạng biểu diễn âm thanh nén cơ bản thành dạng biểu diễn âm thanh tái tạo cơ bản của âm thanh hoặc trường âm thanh. Phương pháp này còn có thể bao gồm bước thu được dạng biểu diễn âm thanh tái tạo cơ bản từ các thành phần được gán cho lớp dùng được cao nhất và lớp bất kỳ thấp hơn lớp dùng được cao nhất, bằng cách sử

dụng thông tin phụ trợ cơ bản. Phương pháp này còn có thể bao gồm bước xác định chỉ số lớp thứ hai chỉ báo phần nào của thông tin phụ trợ tăng cường nên được dùng để cải thiện (ví dụ, nâng cao) dạng biểu diễn âm thanh tái tạo cơ bản. Phương pháp này còn có thể bao gồm bước thu được dạng biểu diễn âm thanh tái tạo của âm thanh hoặc trường âm thanh từ dạng biểu diễn âm thanh tái tạo cơ bản, tham chiếu đến chỉ số lớp thứ hai.

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

Theo các phương án, các thành phần của dạng biểu diễn âm thanh nén cơ bản có thể tương ứng với tín hiệu nghe một tai (ví dụ, tín hiệu truyền một tai). Tín hiệu nghe một tai có thể biểu diễn tín hiệu âm thanh trội hoặc chuỗi hệ số của dạng biểu diễn HOA. Tín hiệu nghe một tai có thể được lượng tử hóa.

Theo các phương án, thông tin phụ trợ cơ bản có thể bao gồm thông tin định rõ việc giải mã (ví dụ, giải nén) một hoặc nhiều trong số nhiều thành phần riêng, độc lập với các thành phần khác. Ví dụ, thông tin phụ trợ cơ bản có thể biểu diễn thông tin phụ trợ liên quan đến tín hiệu nghe một tai riêng, độc lập với tín hiệu nghe một tai khác. Do đó, thông tin phụ trợ cơ bản có thể được gọi là thông tin phụ trợ cơ bản độc lập.

Theo các phương án, thông tin phụ trợ tăng cường có thể biểu diễn thông tin phụ trợ tăng cường. Thông tin phụ trợ tăng cường có thể bao gồm các thông số dự đoán dành cho dạng biểu diễn âm thanh nén cơ bản để cải thiện (ví dụ, tăng cường) dạng biểu diễn âm thanh tái tạo cơ bản mà có thể thu được từ dạng biểu diễn âm thanh nén cơ bản và thông tin phụ trợ cơ bản.

Theo các phương án, phương pháp còn có thể bao gồm bước xác định, đối với mỗi lớp, xem lớp tương ứng đã được nhận một cách hợp lệ hay chưa. Phương pháp còn có thể bao gồm bước xác định chỉ số lớp thứ nhất là chỉ số lớp của lớp ngay dưới lớp dưới cùng không được nhận một cách hợp lệ.

Theo các phương án, việc xác định chỉ số lớp thứ hai có thể bao gồm xác định chỉ số lớp thứ hai bằng chỉ số lớp thứ nhất, hoặc xác định giá trị chỉ số là chỉ số lớp thứ hai mà chỉ báo không dùng thông tin phụ trợ tăng cường bất kỳ khi thu được dạng biểu

diễn âm thanh tái tạo. Trong trường hợp sau, dạng biểu diễn âm thanh tái tạo có thể bằng dạng biểu diễn âm thanh tái tạo cơ bản.

Theo các phương án, các tải tin dữ liệu có thể được thu và được xử lý trong các khoảng thời gian liên tiếp, ví dụ các khoảng thời gian có cùng kích thước. Các khoảng thời gian liên tiếp có thể là các khung. Do đó, phương pháp có thể thực hiện trên cơ sở từng khung. Phương pháp còn có thể bao gồm, nếu các dạng biểu diễn âm thanh nén trong các khoảng thời gian liên tiếp có thể được giải mã độc lập với nhau, bước xác định chỉ số lớp thứ hai bằng chỉ số lớp thứ nhất.

Theo các phương án, các tải tin dữ liệu có thể được thu và được xử lý trong các khoảng thời gian liên tiếp, ví dụ các khoảng thời gian có cùng kích thước. Các khoảng thời gian liên tiếp có thể là các khung. Do đó, phương pháp có thể thực hiện trên cơ sở từng khung. Phương pháp còn có thể bao gồm bước xác định, đối với mỗi lớp, xem lớp tương ứng đã được nhận một cách hợp lệ hay chưa, trong khoảng thời gian đã cho trong số các khoảng thời gian liên tiếp, nếu các dạng biểu diễn âm thanh nén cho các khoảng thời gian liên tiếp không thể được giải mã độc lập với nhau. Phương pháp còn có thể bao gồm bước xác định chỉ số lớp thứ nhất cho các khoảng thời gian đã cho là chỉ số nhỏ hơn trong số chỉ số lớp thứ nhất của khoảng thời gian ở trước khoảng thời gian đã cho và chỉ số lớp của lớp ngay dưới lớp dưới cùng không được nhận một cách hợp lệ.

Theo các phương án, phương pháp còn có thể bao gồm bước, với khoảng thời gian đã cho, nếu các dạng biểu diễn âm thanh nén trong các khoảng thời gian liên tiếp không thể được giải mã độc lập với nhau, xác định xem chỉ số lớp thứ nhất cho khoảng thời gian đã cho có bằng chỉ số lớp thứ nhất cho khoảng thời gian trước đó hay không. Phương pháp còn có thể bao gồm bước, nếu chỉ số lớp thứ nhất cho khoảng thời gian đã cho bằng chỉ số lớp thứ nhất cho khoảng thời gian trước đó, xác định chỉ số lớp thứ hai cho khoảng thời gian đã cho bằng chỉ số lớp thứ nhất cho khoảng thời gian đã cho. Phương pháp còn có thể bao gồm bước, nếu chỉ số lớp thứ nhất cho khoảng thời gian đã cho không bằng chỉ số lớp thứ nhất cho khoảng thời gian trước đó, xác định giá trị chỉ số là chỉ số lớp thứ hai chỉ báo không sử dụng thông tin phụ trợ tăng cường bất kỳ khi thu được dạng biểu diễn âm thanh tái tạo.

Theo các phương án, lớp cơ sở có thể bao gồm ít nhất một phần thông tin phụ trợ cơ bản bổ sung tương ứng với lớp tương ứng và bao gồm thông tin định rõ việc giải mã một hoặc nhiều thành phần trong số các thành phần được gán cho lớp tương ứng phụ thuộc vào các thành phần khác được gán cho lớp tương ứng và lớp bất kỳ thấp hơn lớp tương ứng. Phương pháp còn có thể bao gồm, đối với mỗi phần thông tin phụ trợ cơ bản bổ sung, bước giải mã phần thông tin phụ trợ cơ bản bổ sung bằng cách tham chiếu đến các thành phần được gán cho lớp tương ứng của nó và lớp bất kỳ thấp hơn lớp tương ứng. Phương pháp còn có thể bao gồm bước hiệu chỉnh phần thông tin phụ trợ cơ bản bổ sung bằng cách tham chiếu đến các thành phần được gán cho lớp dùng được cao nhất và lớp bất kỳ nằm giữa lớp dùng được cao nhất và lớp tương ứng. Dạng biểu diễn âm thanh tái tạo cơ bản có thể thu được từ các thành phần được gán cho lớp dùng được cao nhất và lớp bất kỳ thấp hơn lớp dùng được cao nhất, sử dụng thông tin phụ trợ cơ bản và các phần đã được hiệu chỉnh của thông tin phụ trợ cơ bản bổ sung thu được từ các phần thông tin phụ trợ cơ bản bổ sung tương ứng với các lớp lên đến lớp dùng được cao nhất.

Theo các phương án, thông tin phụ trợ cơ bản bổ sung có thể bao gồm thông tin định rõ việc giải mã (ví dụ, giải nén) một hoặc nhiều trong số nhiều thành phần phụ thuộc vào các thành phần khác. Ví dụ, thông tin phụ trợ cơ bản bổ sung có thể biểu diễn thông tin phụ trợ liên quan đến tín hiệu nghe một tai riêng, độc lập với tín hiệu nghe một tai khác. Do đó, thông tin phụ trợ cơ bản bổ sung có thể được gọi là thông tin phụ trợ cơ bản phụ thuộc.

Theo khía cạnh khác, sáng chế đề xuất phương pháp giải mã dạng biểu diễn âm thanh nén của âm thanh hoặc trường âm thanh. Dạng biểu diễn âm thanh nén có thể đã được mã hóa trong nhiều lớp phân cấp. Các lớp phân cấp có thể bao gồm lớp cơ sở và một hoặc nhiều lớp tăng cường phân cấp. Các lớp có thể gán vào đó các thành phần của dạng biểu diễn âm thanh nén cơ bản của âm thanh hoặc trường âm thanh. Nói cách khác, các lớp có thể bao gồm các thành phần của thông tin phụ trợ cơ bản đã được nén. Các thành phần có thể được gán cho các lớp tương ứng trong các nhóm thành phần tương ứng. Các thành phần này có thể là các thành phần bù. Lớp cơ sở có thể bao gồm thông tin phụ trợ cơ bản để giải mã dạng biểu diễn âm thanh nén cơ bản. Lớp cơ sở còn có thể bao gồm ít nhất một phần thông tin phụ trợ cơ bản bổ sung tương ứng với lớp tương ứng

và bao gồm thông tin định rõ việc giải mã một hoặc nhiều thành phần trong số các thành phần được gán cho lớp tương ứng phụ thuộc vào các thành phần khác được gán cho lớp tương ứng và lớp bất kỳ thấp hơn lớp tương ứng. Phương pháp này có thể bao gồm bước nhận các tải tin dữ liệu lần lượt tương ứng với các lớp phân cấp. Phương pháp này còn có thể bao gồm bước xác định chỉ số lớp thứ nhất chỉ báo lớp dùng được cao nhất trong số các lớp sẽ dùng để giải mã dạng biểu diễn âm thanh nén cơ bản thành dạng biểu diễn âm thanh tái tạo cơ bản của âm thanh hoặc trường âm thanh. Phương pháp này còn có thể bao gồm, đối với mỗi phần thông tin phụ trợ cơ bản bổ sung, bước giải mã phần thông tin phụ trợ cơ bản bổ sung bằng cách tham chiếu đến các thành phần được gán cho lớp tương ứng của nó và lớp bất kỳ thấp hơn lớp tương ứng. Phương pháp này còn có thể bao gồm, đối với mỗi phần thông tin phụ trợ cơ bản bổ sung, bước hiệu chỉnh phần thông tin phụ trợ cơ bản bổ sung bằng cách tham chiếu đến các thành phần được gán cho lớp dùng được cao nhất và các lớp bất kỳ nằm giữa lớp dùng được cao nhất và lớp tương ứng. Dạng biểu diễn âm thanh tái tạo cơ bản có thể thu được từ các thành phần được gán cho lớp dùng được cao nhất và lớp bất kỳ thấp hơn lớp dùng được cao nhất, sử dụng thông tin phụ trợ cơ bản và các phần đã được hiệu chỉnh của thông tin phụ trợ cơ bản bổ sung thu được từ các phần thông tin phụ trợ cơ bản bổ sung tương ứng với các lớp lên đến lớp dùng được cao nhất. Phương pháp này còn có thể bao gồm bước xác định chỉ số lớp thứ hai bằng chỉ số lớp thứ nhất hoặc chỉ báo sự bỏ sót thông tin phụ trợ tăng cường trong khi giải mã.

Với cấu hình như vậy, sáng chế đề xuất phương pháp đảm bảo rằng thông tin phụ trợ cơ bản bổ sung thực tế được dùng để giải mã dạng biểu diễn âm thanh nén cơ bản không bao gồm các phần tử dư thừa, qua đó tiến hành giải mã thực tế dạng biểu diễn âm thanh nén cơ bản một cách hiệu quả hơn.

Các phương án của khía cạnh này có thể liên quan đến các phương án của khía cạnh đã nêu trên.

Theo khía cạnh khác, sáng chế mô tả bộ mã hóa để mã hóa theo lớp dạng biểu diễn âm thanh nén của âm thanh hoặc trường âm thanh. Dạng biểu diễn âm thanh nén có thể bao gồm dạng biểu diễn âm thanh nén cơ bản có nhiều thành phần. Các thành phần này có thể là các thành phần bù. Dạng biểu diễn âm thanh nén còn có thể bao gồm

thông tin phụ trợ cơ bản để giải mã dạng biểu diễn âm thanh nén cơ bản thành dạng biểu diễn âm thanh tái tạo cơ bản của âm thanh hoặc trường âm thanh. Dạng biểu diễn âm thanh nén còn có thể bao gồm thông tin phụ trợ tăng cường bao gồm các thông số để cải thiện (ví dụ, tăng cường) dạng biểu diễn âm thanh tái tạo cơ bản. Bộ mã hóa có thể bao gồm bộ xử lý được tạo cấu hình để thực hiện một số hoặc tất cả các bước của các phương pháp theo khía cạnh được đề cập thứ nhất và khía cạnh được đề cập thứ hai nêu trên.

Theo khía cạnh khác, sáng chế mô tả bộ giải mã để giải mã dạng biểu diễn âm thanh nén của âm thanh hoặc trường âm thanh. Dạng biểu diễn âm thanh nén có thể đã 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. Các lớp có thể gán vào đó các thành phần của dạng biểu diễn âm thanh nén cơ bản của âm thanh hoặc trường âm thanh. Nói cách khác, các lớp có thể bao gồm các thành phần của thông tin phụ trợ cơ bản đã được nén. Các thành phần có thể được gán cho các lớp tương ứng trong các nhóm thành phần tương ứng. Các thành phần này có thể là các thành phần bù. Lớp cơ sở có thể bao gồm thông tin phụ trợ cơ bản để giải mã dạng biểu diễn âm thanh nén cơ bản. Mỗi lớp có thể bao gồm một phần thông tin phụ trợ tăng cường bao gồm các thông số để cải thiện (ví dụ, tăng cường) dạng biểu diễn âm thanh tái tạo cơ bản có thể thu được từ dữ liệu có trong lớp tương ứng và lớp bất kỳ thấp hơn lớp tương ứng. Bộ giải mã có thể bao gồm bộ xử lý được tạo cấu hình để thực hiện một số hoặc tất cả các bước của các phương pháp theo khía cạnh được đề cập thứ ba và khía cạnh được đề cập thứ tư nêu trên.

Theo các khía cạnh khác, các phương pháp, thiết bị và hệ thống hướng đến giải mã dạng biểu diễn âm thanh Ambisonics bậc cao (HOA) của âm thanh hoặc trường âm thanh. Thiết bị có thể có bộ thu được tạo cấu hình để thu hoặc phương pháp có thể thu dòng bit chứa dạng biểu diễn HOA nén tương ứng với các lớp phân cấp bao gồm lớp cơ sở và một hoặc nhiều lớp tăng cường phân cấp. Các 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. Thiết bị có thể có bộ giải mã được tạo cấu hình để giải mã hoặc phương pháp có thể giải mã dạng biểu diễn HOA nén dựa trên thông tin phụ trợ cơ bản được kết hợp với lớp cơ sở và dựa trên thông tin phụ trợ tăng cường được kết hợp với một hoặc nhiều lớp tăng cường phân

cấp. Thông tin phụ trợ cơ bản có thể bao gồm thông tin phụ trợ cơ bản độc lập liên quan đến tín hiệu nghe một tai riêng thứ nhất sẽ được giải mã độc lập với tín hiệu nghe một tai khác. Mỗi lớp trong số một hoặc nhiều lớp tăng cường phân cấp có thể bao gồm một phần thông tin phụ trợ tăng cường bao gồm các thông số cải thiện dạng biểu diễn âm thanh tái tạo cơ bản có thể thu được từ dữ liệu chứa trong các lớp tương ứng và lớp bất kỳ thấp hơn lớp tương ứng.

Thông tin phụ trợ cơ bản độc lập có thể chỉ báo rằng tín hiệu nghe một tai riêng thứ nhất biểu diễn tín hiệu định hướng với một hướng tới. Thông tin phụ trợ cơ bản còn có thể bao gồm thông tin phụ trợ cơ bản phụ thuộc liên quan đến tín hiệu nghe một tai riêng thứ hai sẽ được giải mã một cách phụ thuộc vào tín hiệu nghe một tai khác. Thông tin phụ trợ cơ bản phụ thuộc có thể bao gồm tín hiệu dựa trên vector được phân bổ theo hướng trong trường âm thanh, trong đó sự phân bổ theo hướng này có thể được định rõ bằng vector. Các thành phần của vector này được thiết lập về không và không phải là một phần của dạng biểu diễn vector nén.

Các thành phần của dạng biểu diễn âm thanh nén cơ bản có thể tương ứng với tín hiệu nghe một tai mà biểu diễn tín hiệu âm thanh trội hoặc chuỗi hệ số của dạng biểu diễn HOA. Dòng bit bao gồm các tải tin dữ liệu lần lượt tương ứng với các lớp phân cấp. Thông tin phụ trợ tăng cường có thể bao gồm các thông số liên quan đến ít nhất một trong số: dự đoán theo không gian, tổng hợp tín hiệu định hướng dải con, và sao chép thông số xung quanh. Thông tin phụ trợ tăng cường có thể bao gồm thông tin cho phép dự đoán việc bỏ lỡ các phần của âm thanh hoặc trường âm thanh từ tín hiệu định hướng. Còn có thể xác định được, đối với mỗi lớp, liệu lớp tương ứng đã được nhận một cách hợp lệ hay chưa và chỉ số lớp của lớp ngay dưới lớp thấp nhất chưa được nhận hợp lệ.

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.

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 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.

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 thi 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 lưu đồ thể hiện ví dụ về phương pháp mã hóa theo lớp theo các phương án của sáng chế;

Fig.2 là sơ đồ khối thể hiện giản lược ví dụ về giai đoạn bộ mã hóa theo các phương án của sáng chế;

Fig.3 là lưu đồ thể hiện ví dụ về phương pháp giải mã dạng biểu diễn âm thanh nén của âm thanh hoặc trường âm thanh đã được mã hóa thành các lớp phân cấp, theo các phương án của sáng chế;

Fig.4A và Fig.4B là các sơ đồ khối thể hiện giản lược các ví dụ về giai đoạn bộ giải mã theo các phương án của sáng chế;

Fig.5 là sơ đồ khối thể hiện ví dụ về cài đặt phần cứng của bộ mã hóa theo các phương án của sáng chế; và

Fig.6 là sơ đồ khối thể hiện ví dụ về cài đặt 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 (sau đây gọi ngắn gọn là dạng biểu diễn âm thanh nén) mà các phương pháp và bộ mã hóa/bộ giải mã theo sáng chế có thể áp dụng. Nhìn chung, dạng biểu diễn âm thanh (hoặc trường âm thanh) nén hoàn chỉnh (sau đây gọi ngắn gọn là dạng biểu diễn âm thanh nén hoàn chỉnh) có thể bao gồm (ví dụ, gồm có) ba thành phần sau đây: dạng biểu diễn âm thanh (trường âm thanh) nén cơ bản (sau đây gọi ngắn gọn là dạng biểu diễn âm thanh nén cơ bản), thông tin phụ trợ cơ bản, và thông tin phụ trợ tăng cường.

Dạng biểu diễn âm thanh nén cơ bản bao gồm (ví dụ, gồm có) một số thành phần (ví dụ, các thành phần bù). Dạng biểu diễn âm thanh nén cơ bản có thể chiếm tỷ lệ phần trăm lớn nhất rõ ràng của dạng biểu diễn âm thanh nén hoàn chỉnh. Dạng biểu diễn âm thanh nén cơ bản có thể bao gồm tín hiệu truyền một tai biểu diễn tín hiệu âm thanh trội hoặc các chuỗi hệ số của dạng biểu diễn HOA ban đầu.

Thông tin phụ trợ cơ bản là cần thiết để giải mã dạng biểu diễn âm thanh nén cơ bản và có thể được giả định là có kích thước nhỏ hơn nhiều so với dạng biểu diễn âm thanh nén cơ bản. Nó có thể chiếm phần lớn nhất trong số các phần không liên nhau, mỗi phần định rõ việc giải nén chỉ một thành phần cụ thể của dạng biểu diễn âm thanh nén cơ bản. Thông tin phụ trợ cơ bản có thể bao gồm phần thứ nhất là thông tin phụ trợ cơ bản độc lập và phần thứ hai là thông tin phụ trợ cơ bản bổ sung.

Cả phần thứ nhất và phần thứ hai, thông tin phụ trợ cơ bản độc lập và thông tin phụ trợ cơ bản bổ sung, có thể định rõ việc giải nén các thành phần cụ thể của dạng biểu diễn âm thanh nén cơ bản. Phần thứ hai là tùy ý và có thể lược bỏ. Trong trường hợp này, dạng biểu diễn âm thanh nén có thể được xem là bao gồm phần thứ nhất (ví dụ, thông tin phụ trợ cơ bản).

Phần thứ nhất (ví dụ, thông tin phụ trợ cơ bản) có thể chứa thông tin phụ trợ mô tả các thành phần (bù) riêng của dạng biểu diễn âm thanh nén cơ bản độc lập với các thành phần (bù) khác. Cụ thể là, phần thứ nhất (ví dụ, thông tin phụ trợ cơ bản) có thể định rõ việc giải mã một hoặc nhiều trong số nhiều thành phần riêng, độc lập với các thành phần khác. Do đó, phần thứ nhất có thể được gọi là thông tin phụ trợ cơ bản độc lập.

Phần thứ hai (tùy ý) có thể chứa thông tin phụ trợ, còn gọi là thông tin phụ trợ cơ bản bổ sung, có thể mô tả các thành phần (bù) riêng của dạng biểu diễn âm thanh nén cơ bản phụ thuộc vào các thành phần (bù) khác. Phần thứ hai này còn có thể được gọi là thông tin phụ trợ cơ bản phụ thuộc. Cụ thể, sự phụ thuộc này có thể có các thuộc tính sau:

- Thông tin phụ trợ cơ bản phụ thuộc dành cho mỗi thành phần (bù) riêng của dạng biểu diễn âm thanh nén cơ bản có thể đạt được mức độ cao nhất của nó khi không có các thành phần (bù) nhất định khác chứa trong dạng biểu diễn âm thanh nén cơ bản.
- Trong trường hợp các thành phần (bù) nhất định khác được thêm vào dạng biểu diễn âm thanh nén cơ bản, thông tin phụ trợ cơ bản phụ thuộc dành cho thành phần (bù) riêng được xem xét có thể trở thành tập con của thông tin phụ trợ cơ bản phụ thuộc ban đầu, nhờ đó làm giảm kích thước của nó.

Thông tin phụ trợ tăng cường cũng là tùy ý. Nó có thể được dùng để cải thiện hoặc tăng cường (ví dụ, cải thiện hoặc tăng cường bằng thông số) dạng biểu diễn âm thanh nén cơ bản. Kích thước của nó còn có thể được giả định là nhỏ hơn nhiều so với kích thước của dạng biểu diễn âm thanh nén.

Do đó, trong các phương án dạng biểu diễn âm thanh nén có thể bao gồm dạng biểu diễn âm thanh nén cơ bản bao gồm các thành phần, thông tin phụ trợ cơ bản để giải mã (ví dụ, giải nén) dạng biểu diễn âm thanh nén cơ bản thành dạng biểu diễn âm thanh tái tạo cơ bản của âm thanh hoặc trường âm thanh, và thông tin phụ trợ tăng cường bao gồm các thông số để cải thiện hoặc tăng cường (ví dụ, cải thiện hoặc tăng cường bằng thông số) dạng biểu diễn âm thanh tái tạo cơ bản. Dạng biểu diễn âm thanh nén còn có thể bao gồm thông tin phụ trợ cơ bản bổ sung để giải mã (ví dụ, giải nén) dạng biểu diễn

âm thanh nén cơ bản thành dạng biểu diễn âm thanh tái tạo cơ bản, mà có thể bao gồm thông tin định rõ việc giải mã một hoặc nhiều trong số nhiều thành phần phụ thuộc vào các thành phần tương ứng khác.

Một ví dụ về loại dạng biểu diễn âm thanh nén hoàn chỉnh như vậy là dạng biểu diễn trường âm thanh Ambisonics bậc cao (HOA) nén như được mô tả rõ trong phiên bản sơ bộ của chuẩn âm thanh MPEG-H 3D (Đối chứng 1), chương 12 và phụ lục C. 5. Tức là, dạng biểu diễn âm thanh nén có thể tương ứng với dạng biểu diễn âm thanh (hoặc trường âm thanh) HOA nén của âm thanh hoặc trường âm thanh.

Trong ví dụ này, dạng biểu diễn trường âm thanh nén cơ bản (dạng biểu diễn âm thanh nén cơ bản) có thể bao gồm (ví dụ, có thể được nhận dạng bằng) một số thành phần. Các thành phần này có thể là (ví dụ, tương ứng với) tín hiệu nghe một tai. Tín hiệu nghe một tai có thể là tín hiệu nghe một tai được lượng tử hóa. Tín hiệu nghe một tai có thể biểu diễn tín hiệu âm thanh trội hoặc chuỗi hệ số của thành phần trường âm thanh HOA xung quanh.

Thông tin phụ trợ cơ bản có thể 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 vào trường âm thanh. Ví dụ, thông tin phụ trợ cơ bản có thể định rõ tín hiệu âm thanh trội là tín hiệu định hướng thuần, nghĩa là sóng phẳng chung với hướng tới nhất định. Theo cách khác, thông tin phụ trợ cơ bản có thể định rõ tín hiệu nghe một tai là chuỗi hệ số của dạng biểu diễn HOA ban đầu có chỉ số nhất định. Thông tin phụ trợ cơ bản còn có thể được tách thành phần thứ nhất và phần thứ hai, như nêu trên.

Phần thứ nhất là thông tin phụ trợ (ví dụ, thông tin phụ trợ cơ bản độc lập) liên quan đến tín hiệu nghe một tai riêng đặc biệt. Thông tin phụ trợ cơ bản độc lập này độc lập với sự tồn tại của 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 (ví dụ, 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. Phần thứ nhất có thể được gọi là thông tin phụ trợ cơ bản độc lập. Nhìn chung, phần thứ nhất (ví dụ, thông tin phụ trợ cơ bản) có thể định rõ việc giải mã một hoặc nhiều trong số nhiều tín hiệu nghe một tai riêng, độc lập với tín hiệu nghe một tai khác.

Phần thứ hai là thông tin phụ trợ (ví dụ, thông tin phụ trợ cơ bản bổ sung) liên quan đến tín hiệu nghe một tai riêng đặc biệt. Thông tin phụ trợ này phụ thuộc vào sự tồn tại của tín hiệu nghe một tai khác. Thông tin phụ trợ như thế có thể được sử dụng, ví dụ, nếu tín hiệu nghe một tai được định rõ là tín hiệu dựa trên vector (xem, ví dụ, Đối chứng 1, phần 12.4.2.4.4). Tín hiệu này được phân bổ theo hướng trong trường âm thanh, trong đó sự phân bổ theo hướng này có thể được định rõ bằng vector. Theo chế độ nhất định (ví dụ, `CodedVVecLength = 1`), các thành phần cụ thể của vector này được thiết lập ngầm định 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 và một phần của dạng biểu diễn âm thanh nén cơ bản. Điều này có nghĩa là nếu các thành phần riêng của vector được mã hóa, thì tổng số của chúng có thể phụ thuộc vào dạng biểu diễn âm thanh nén cơ bản. Cụ thể là tổng số này có thể phụ thuộc vào chuỗi hệ số nào mà dạng biểu diễn HOA ban đầu có chứa.

Nếu không có chuỗi hệ số của dạng biểu diễn HOA ban đầu chứa trong dạng biểu diễn 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 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 cho tín hiệu dựa trên vector.

Thông tin phụ trợ tăng cường (ví dụ, thông tin phụ trợ tăng cường) có thể bao gồm các thông số liên quan đến dự đoán theo không gian (dải rộng) (xem Đối chứng 1, phần 12.4.2.4.3) và/hoặc các thông số liên quan đến tổng hợp tín hiệu định hướng dải con và sao chép thông số xung quanh.

Các thông số liên quan đến dự báo không gian (dải rộng) có thể được dùng để dự báo (tuyến tính) các phân thiếu của trường âm thanh từ tín hiệu hướng.

Tổng hợp tín hiệu định hướng dải con và sao chép thông số xung quanh là các công cụ nén mới được đưa vào chuẩn âm thanh MPEG-H 3D sửa đổi [xem Đối chứng 2, phần 1]. Hai công cụ này cho phép dự đoán thông số phụ thuộc vào tần số của tín hiệu nghe một tai bổ sung sẽ được phân bổ theo không gian để bù cho dạng biểu diễn HOA

nén chưa hoàn chỉnh hoặc thiếu hụt về mặt không gian. Phép dự đoán này có thể dựa trên chuỗi hệ số của dạng biểu diễn âm thanh nén cơ bản.

Một điểm quan trọng cần lưu ý là phần đóng góp bù được đề cập trên đây vào trường âm thanh được dạng biểu diễn trong dạng biểu diễn HOA nên không phải nhờ 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 đã được đề cập đặc biệt thích hợp để nén dạng biểu diễn HOA ở tốc độ dữ liệu thấp.

Ví dụ thứ hai về dạng biểu diễn nén của một hoặc nhiều tín hiệu nghe một tai có cấu trúc nêu trên có thể bao gồm thông tin phổ được mã hóa dành cho các dải tần số ngắt quãng lên đến tần số cao hơn nhất định, đây có thể coi là dạng biểu diễn nén cơ bản; thông tin phụ trợ cơ bản chỉ rõ thông tin phổ được mã hóa (ví dụ, bằng số lượng và độ rộng của các dải tần số được mã hóa); và thông tin phụ trợ tăng cường bao gồm (ví dụ, gồm có) các thông số của phép sao chép dải phổ (Spectral Band Replication - SBR), mô tả cách tái tạo bằng thông số thông tin phổ từ dạng biểu diễn nén cơ bản đối với các dải tần cao không được xem xét trong dạng biểu diễn nén cơ bản.

Sáng chế đề xuất phương pháp mã hóa theo lớp dạng biểu diễn âm thanh (hoặc trường âm thanh) nén hoàn chỉnh có cấu trúc nêu trên.

Việc nén có thể dựa trên khung nghĩa là nó cung cấp các dạng biểu diễn nén (dưới dạng gói dữ liệu hoặc các tải tin khung tương đương) trong các khoảng thời gian liên tiếp. Các khoảng thời gian có thể có kích thước bằng hoặc khác nhau. Các gói dữ liệu này có thể được giả định là chứa cờ hợp lệ, giá trị chỉ báo kích thước của chúng cũng như dữ liệu dạng biểu diễn nén thực tế. Sau đây, không nhằm mục đích giới hạn sáng chế, giả định rằng việc nén là dựa trên khung. Ngoài ra, trừ khi có chỉ dẫn khác và không nhằm mục đích giới hạn sáng chế, phần mô tả sau đây sẽ tập trung 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 đang xem xét được giả định là chứa J gói dữ liệu (hoặc tải tin khung), mỗi gói cho một thành phần của dạng biểu diễn âm thanh nén cơ bản, mà được biểu thị bằng $BSRC_j$, $j = 1, \dots, J$. Hơn nữa, tải tin được giả định là chứa gói có thông tin phụ trợ cơ bản *độc*

lập (thông tin phụ trợ cơ bản) được biểu thị bằng BSI_I định 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 còn được giả định là chứa gói có thông tin phụ trợ cơ bản *phụ thuộc* (thông tin phụ trợ cơ bản bổ sung) được biểu thị bằng BSI_D định 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_I và BSI_D có thể được nhóm tùy ý vào một gói dữ liệu đơn nhất BSI của thông tin phụ trợ cơ bản. Một gói dữ liệu BSI có thể được xem là chứa, trong số những thứ khác, J phần, mỗi phần định rõ một thành phần cụ thể $BSRC_j$ của dạng biểu diễn âm thanh nén cơ bản. Mỗi phần này lần lượt có thể được xem là chứa một phần thông tin phụ trợ độc lập và, tùy ý, một phần thông tin phụ trợ phụ thuộc.

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

Sáng chế đề xuất giải pháp mã hóa theo lớp chỉ ra các bước cần thiết để kích hoạt cả phần nén bao gồm việc gói các gói dữ liệu để truyền cũng như bộ thu và phần giải nén. Sau đây, mỗi phần sẽ được mô tả chi tiết.

Đầu tiên, việc nén và đóng gói (ví dụ, để truyền) sẽ được mô tả. Cụ thể là, các thành phần và phần tử của dạng biểu diễn âm thanh (hoặc trường âm thanh) nén hoàn chỉnh trong trường hợp mã hóa theo lớp sẽ được mô tả.

Fig.1 thể hiện giản lược về lưu đồ của ví dụ về phương pháp nén và đóng gói (ví dụ, phương pháp mã hóa, hoặc phương pháp mã hóa theo lớp dạng biểu diễn âm thanh nén của âm thanh hoặc trường âm thanh). Phép gán (ví dụ, phân bổ) các tải tin riêng cho lớp cơ sở và $(M - 1)$ lớp tăng cường có thể được thực hiện bởi bộ đóng gói lớp truyền tải. Fig.2 thể hiện giản lược sơ đồ khối của ví dụ về việc gán/phân bổ các tải tin riêng.

Như đã nêu trên đây, dạng biểu diễn âm thanh nén hoàn chỉnh 2100 có thể liên quan đến, ví dụ, dạng biểu diễn HOA nén bao gồm dạng biểu diễn âm thanh nén cơ bản. Dạng biểu diễn âm thanh nén hoàn chỉnh 2100 có thể bao gồm các thành phần (ví dụ,

tín hiệu nghe một tai) 2110-1, ... 2110- J , thông tin phụ trợ cơ bản độc lập (thông tin phụ trợ cơ bản) 2120, thông tin phụ trợ tăng cường tùy ý (thông tin phụ trợ tăng cường) 2140, và thông tin phụ trợ cơ bản phụ thuộc tùy ý (thông tin phụ trợ cơ bản bổ sung) 2130. Thông tin phụ trợ cơ bản 2120 có thể là thông tin để giải mã dạng biểu diễn âm thanh nén cơ bản thành dạng biểu diễn âm thanh tái tạo cơ bản của âm thanh hoặc trường âm thanh. Thông tin phụ trợ cơ bản 2120 có thể bao gồm thông tin định rõ việc giải mã một hoặc nhiều thành phần (ví dụ, tín hiệu nghe một tai) riêng, độc lập với các thành phần khác. Thông tin phụ trợ tăng cường 2140 có thể bao gồm các thông số để cải thiện (ví dụ, tăng cường) dạng biểu diễn âm thanh tái tạo cơ bản. Thông tin phụ trợ cơ bản bổ sung 2130 có thể là thông tin (thêm) để giải mã dạng biểu diễn âm thanh nén cơ bản thành dạng biểu diễn âm thanh tái tạo cơ bản, và có thể bao gồm thông tin định rõ việc giải mã một hoặc nhiều trong số nhiều thành phần phụ thuộc vào các thành phần tương ứng khác.

Fig.2 thể hiện giả định cơ sở trong đó có các lớp phân cấp, bao gồm một lớp cơ sở (lớp cơ bản) và một hoặc nhiều lớp tăng cường (phân cấp). Ví dụ, có thể có tổng M lớp, tức là một lớp cơ sở và $M - 1$ các lớp tăng cường. Các lớp phân cấp có chỉ số lớp tăng nối tiếp. Giá trị thấp nhất của chỉ số lớp (ví dụ, chỉ số lớp 1) tương ứng với lớp cơ sở. Còn có thể hiểu rằng các lớp được sắp xếp, từ lớp cơ sở, đến lớp tăng cường, đến lớp tăng cường tổng thể cao nhất (lớp tổng thể cao nhất).

Phương pháp được đề xuất có thể được thực hiện trên cơ sở khung (tức là, theo từng khung). Cụ thể là, dạng biểu diễn âm thanh nén 2100 có thể được nén trong các khoảng thời gian liên tiếp, ví dụ các khoảng thời gian có cùng kích thước. Mỗi khoảng thời gian có thể tương ứng với một khung. Các bước được mô tả dưới đây có thể được thực hiện đối với mỗi khoảng thời gian (ví dụ, khung) liên tiếp.

Ở bước S1010 trên Fig.1, nhiều thành phần 2110 được chia nhỏ thành nhiều nhóm thành phần. Mỗi trong số nhiều nhóm sau đó được gán (ví dụ, thêm hoặc phân bổ) cho một trong số nhiều lớp phân cấp tương ứng. Trong đó, số lượng nhóm tương ứng với số lớp. Ví dụ, số nhóm có thể bằng số lớp, nên sẽ có một nhóm thành phần dành cho mỗi lớp. Như nêu trên, các lớp có thể bao gồm một lớp cơ sở và một hoặc nhiều (ví dụ, $M - 1$) lớp tăng cường phân cấp.

Nói cách khác, dạng biểu diễn 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 các thành phần BSRC_j được gán cho lớp thứ m với $J_{m-1} \leq j < J_m$.

Tại bước S1020, các nhóm thành phần được gán cho các lớp tương ứng của chúng. Tại bước S1030, thông tin phụ trợ cơ bản 2120 được thêm (ví dụ, được phân bổ) vào lớp cơ sở (tức là, lớp thấp nhất trong số các lớp phân cấp).

Tức là, do nó có kích thước nhỏ nên sáng chế đề xuất bao gồm thông tin phụ trợ cơ bản hoàn chỉnh (thông tin phụ trợ cơ bản và thông tin phụ trợ cơ bản bổ sung tùy ý) vào lớp cơ sở để tránh sự phân mảnh không cần thiết.

Nếu dạng biểu diễn âm thanh nén đang xem xét bao gồm thông tin phụ trợ cơ bản phụ thuộc (thông tin phụ trợ cơ bản bổ sung), phương pháp còn có thể bao gồm (không được thể hiện trên Fig.1) bước phân tích thông tin phụ trợ cơ bản bổ sung thành các phần 2130-1, ..., 2130- M của thông tin phụ trợ cơ bản bổ sung. Khi đó các phần thông tin phụ trợ cơ bản bổ sung có thể được thêm (ví dụ, được phân bổ) vào lớp cơ sở. Nói cách khác, các phần thông tin phụ trợ cơ bản bổ sung có thể được chứa trong lớp cơ sở. Mỗi phần thông tin phụ trợ cơ bản bổ sung có thể tương ứng với lớp tương ứng và có thể bao gồm thông tin định rõ việc giải mã một hoặc nhiều thành phần được gán cho lớp tương ứng phụ thuộc vào các thành phần khác được gán cho lớp tương ứng và lớp bất kỳ thấp hơn lớp tương ứng.

Do đó, trong khi thông tin phụ trợ cơ bản độc lập BSI₁ (thông tin phụ trợ cơ bản) 2120 được giữ không đổi cho việc gán, thông tin phụ trợ cơ bản phụ thuộc phải được xử lý đặc biệt cho việc 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 cơ bản phụ thuộc cần truyền. Đề xuất là phân tích thông tin phụ trợ cơ bản phụ thuộc thành M phần được biểu thị là BSI_{D,m}, $m = 1, \dots, M$, trong đó phần thứ m chứa thông tin phụ trợ phụ thuộc cho mỗi thành phần BSRC_j, $J_{m-1} \leq j < J_m$, của dạng biểu diễn âm thanh nén cơ bản được gán cho lớp thứ m , giả định rằng có tồn tại thông tin phụ trợ cơ bản phụ thuộc tùy ý cho dạng biểu diễn âm thanh nén đang xem xét. Trong trường hợp thông tin phụ trợ phụ thuộc tương ứng không tồn tại, thì các phần dạng biểu diễn âm thanh nén BSI_{D,m} có thể

được giả định là rỗng. Mỗi phần thông tin phụ trợ cơ bản phụ thuộc $BSI_{D,m}$ có thể phụ thuộc vào tất cả các thành phần $BSRC_j$, $1 \leq j \leq J_m$, chứa trong tất cả các lớp lên đến lớp thứ m , (tức là, chứa trong tất cả các lớp $j = 1, \dots, m$).

Nếu gói thông tin phụ trợ cơ bản độc lập BSI_I có kích thước quá nhỏ thì giữ nguyên và thêm (gán) nó vào lớp cơ sở mới là hợp lý. Tùy ý, việc phân tích tương tự đối với thông tin phụ trợ cơ bản phụ thuộc còn có thể được thực hiện với cả thông tin phụ trợ cơ bản độc lập, cung cấp các gói $BSI_{I,m}$, $m = 1, \dots, M$. Điều này là có lợi khi giảm kích thước của lớp cơ sở bằng cách thêm (gán) các phần thông tin phụ trợ cơ bản độc lập vào các lớp có các thành phần tương ứng của dạng biểu diễn âm thanh nén cơ bản.

Tại bước S1040, các phần 2140-1, ..., 2140- M của thông tin phụ trợ tăng cường có thể được xác định. Mỗi phần thông tin phụ trợ tăng cường có thể bao gồm các thông số để cải thiện (ví dụ, tăng cường) dạng biểu diễn âm thanh tái tạo có thể thu được từ dữ liệu chứa trong lớp tương ứng và lớp bất kỳ thấp hơn lớp tương ứng.

Lý do để thực hiện bước này là 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 bổ sung, vì dự tính là 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. Cụ thể là, âm thanh (hoặc trường âm thanh) nén sơ bộ dành cho lớp giải mã được cao nhất cho trước (lớp dùng được cao nhất) phụ thuộc vào các thành phần có trong lớp giải mã được cao nhất và lớp bất kỳ bên dưới lớp giải mã được cao nhất. Vì thế, việc nén phải tạo ra M gói dữ liệu thông tin phụ trợ tăng cường riêng (các phần thông tin phụ trợ tăng cường), được biểu thị là 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 (ví dụ, tất cả dữ liệu chứa trong lớp thứ m và lớp bất kỳ bên dưới lớp thứ m).

Tại bước S1050, các phần 2140-1, ..., 2140- M của thông tin phụ trợ tăng cường được gán (ví dụ, thêm hoặc phân bổ) cho các lớp. Mỗi trong số nhiều phần thông tin phụ trợ tăng cường được gán cho một trong số nhiều lớp tương ứng. Ví dụ, mỗi trong số nhiều lớp bao gồm phần tương ứng của thông tin phụ trợ tăng cường.

Việc gán thông tin phụ trợ cơ bản và/hoặc tăng cường cho các lớp tương ứng có thể được chỉ báo trong thông tin cấu hình được tạo ra bởi phương pháp mã hóa. Nói cách khác, sự tương quan giữa thông tin phụ trợ cơ bản và/hoặc tăng cường với các lớp tương ứng có thể được chỉ báo trong thông tin cấu hình. Ngoài ra, thông tin cấu hình có thể chỉ báo, đối với mỗi lớp, các thành phần của dạng biểu diễn âm thanh nén cơ bản được gán cho (ví dụ, được bao gồm trong) lớp đó. Các phần thông tin phụ trợ cơ bản bổ sung được bao gồm trong lớp cơ sở, còn có thể tương ứng với các lớp khác với lớp cơ sở.

Tóm lại, ở giai đoạn nén, gói dữ liệu khung, được biểu thị bằng FRAME, được đề xuất là có thành phần sau:

$$\text{FRAME} = [\text{BSRC}_1 \quad \dots \quad \text{BSRC}_J \quad \text{BSI}_I \quad \text{BSI}_{D,1} \quad \dots \quad \text{BSI}_{D,M} \quad \text{ESI}_1 \quad \dots \quad \text{ESI}_M] \quad (1)$$

Ngoài ra, các gói BSI_I và $\text{BSI}_{D,m}$ với $m = 1, \dots, M$ có thể được kết hợp thành một gói BSI , trong trường hợp mà gói dữ liệu khung, được biểu thị bằng FRAME sẽ có thành phần sau:

$$\text{FRAME} = [\text{BSRC}_1 \quad \text{BSRC}_2 \quad \dots \quad \text{BSRC}_J \quad \text{BSI} \quad \text{ESI}_1 \quad \text{ESI}_2 \quad \dots \quad \text{ESI}_M] \quad (2)$$

Việc sắp xếp các tải tin riêng với gói dữ liệu khung nhìn chung là có thể tùy ý.

Các gói dữ liệu riêng khi đó có thể được nhóm trong các tải tin, được xác định là các gói dữ liệu đặc biệt chứa cờ hợp lệ, giá trị chỉ báo kích thước của chúng cũng như dữ liệu biểu diễn nén thực. Việc sử dụng các tải tin cho phép giải dồn kênh đơn giản ở phía bộ thu, đem lại một lợi ích là có thể loại bỏ các tải tin lỗi thời mà không cần phải phân tích cú pháp chúng. Một phân nhóm có thể thực hiện là

- gán (ví dụ, phân bổ) mỗi BSRC_j gói, $j = 1, \dots, J$, cho một tải tin riêng được biểu thị là $\overline{BP}_j$.
- gán (ví dụ, phân bổ) gói dữ liệu thông tin phụ trợ tăng cường thứ m ESI_m và gói dữ liệu thông tin phụ trợ phụ thuộc thứ m $\text{BSI}_{D,m}$ cho một tải tin tăng cường được biểu thị bằng $\overline{EP}_m$, $m = 1, \dots, M$.
- gán gói thông tin phụ trợ cơ bản độc lập BSI_I cho tải tin thông tin phụ trợ riêng được biểu thị bằng $\overline{BSIP}$.

Tùy ý, nếu thông tin phụ trợ cơ bản độc lập có kích thước lớn, mỗi thành phần thứ m trong số các thành phần, $BSI_{I,m}$, $m = 1, \dots, M$, có thể được gán (ví dụ, phân bổ) cho tải tin tăng cường $\overline{EP}_m$. Trong trường hợp này, tải tin thông tin phụ trợ $\overline{BSIP}$ là rỗng và có thể bỏ qua.

Lựa chọn khác là gán tất cả các gói dữ liệu thông tin phụ trợ cơ bản phụ thuộc $BSI_{D,m}$ vào tải tin thông tin phụ trợ $\overline{BSIP}$, phương án này hợp lý khi thông tin phụ trợ cơ bản phụ thuộc có kích thước nhỏ.

Cuối cùng, gói dữ liệu khung, được biểu thị bằng $FRAME$, có thể được đề xuất có thành phần sau

$$FRAME = [\overline{BP}_1 \quad \dots \quad \overline{BP}_J \quad \overline{BSIP} \quad \overline{EP}_1 \quad \dots \quad \overline{EP}_M] \quad (3)$$

Việc sắp xếp các tải tin riêng với gói dữ liệu khung nhìn chung là có thể tùy ý.

Phương pháp còn có thể bao gồm (không được thể hiện trên Fig.1) bước tạo, đối với mỗi trong số nhiều lớp, gói lớp truyền tải (ví dụ, lớp cơ sở gói 2200 và M-1 gói chứa lớp tăng cường 2300-1, ..., 2300-($M - 1$)) bao gồm dữ liệu của lớp tương ứng (ví dụ, các thành phần, thông tin phụ trợ cơ bản và thông tin phụ trợ tăng cường cho lớp cơ sở, hoặc các thành phần và thông tin phụ trợ tăng cường cho một hoặc nhiều lớp tăng cường).

Các gói lớp truyền tải cho các lớp khác nhau có thể có mức độ ưu tiên truyền khác nhau. Do đó, phương pháp còn có thể bao gồm (không được thể hiện trên Fig.1), bước tạo dòng truyền tải để truyền dữ liệu của các lớp, trong đó lớp cơ sở có mức độ ưu tiên truyền cao nhất và các lớp tăng cường phân cấp có mức độ ưu tiên truyền giảm. Trong đó, mức độ ưu tiên truyền cao hơn có thể tương ứng với mức bảo vệ chống lỗi lớn hơn, và ngược lại.

Trừ khi các bước đòi hỏi 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.1 được hiểu là không làm giới hạn sáng chế.

Fig.3 thể hiện phương pháp giải mã dạng biểu diễn âm thanh nén của âm thanh hoặc trường âm thanh) để giải mã hoặc giải nén (mở gói). Ví dụ về bộ thu tương ứng và giai đoạn giải nén được thể hiện giản lược trong các sơ đồ khối trên Fig.4A và Fig.4B.

Theo phần mô tả nêu trên, dạng biểu diễn âm thanh nén có thể được mã hóa trong nhiều lớp phân cấp. Các lớp đã có thể gắn vào đó (ví dụ, có thể bao gồm) các thành phần của dạng biểu diễn âm thanh nén cơ bản, các thành phần được gắn cho các lớp tương ứng trong các nhóm thành phần tương ứng. Lớp cơ sở có thể bao gồm thông tin phụ trợ cơ bản để giải mã dạng biểu diễn âm thanh nén cơ bản. Mỗi lớp có thể bao gồm một trong số các phần nêu trên của thông tin phụ trợ tăng cường bao gồm các thông số để cải thiện dạng biểu diễn âm thanh tái tạo cơ bản có thể thu được từ dữ liệu có trong lớp tương ứng và lớp bất kỳ thấp hơn lớp tương ứng.

Phương pháp được đề xuất có thể được thực hiện trên cơ sở khung (tức là, theo từng khung). Cụ thể là, dạng biểu diễn âm thanh hoặc trường âm thanh được khôi phục có thể được tạo ra cho các khoảng thời gian liên tiếp, ví dụ các khoảng thời gian có cùng kích thước. Các khoảng thời gian có thể là các khung, ví dụ. Các bước được mô tả dưới đây có thể được thực hiện đối với mỗi khoảng thời gian (ví dụ, khung) liên tiếp.

Tại bước S3010, thu được các tải tin dữ liệu (ví dụ, các gói lớp truyền tải) tương ứng với các lớp. Các tải tin dữ liệu có thể được thu dưới dạng một phần dòng bit chứa dạng biểu diễn HOA nén của âm thanh hoặc trường âm thanh, dạng biểu diễn này tương ứng với các lớp phân cấp. Các lớp phân cấp bao gồm một lớp cơ sở và một hoặc nhiều lớp tăng cường phân cấp. Các lớp gắn 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 gắn cho các lớp tương ứng trong các nhóm thành phần tương ứng.

Các gói lớp riêng có thể có thể được dồn kênh để cung cấp gói khung thu được của dạng biểu diễn âm thanh nén hoàn chỉnh. Gói khung thu được có thể được thể hiện bởi

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

(4)

Trong trường hợp khác, các gói BSI_I và $BSI_{D,m}$ với $m = 1, \dots, M$ được kết hợp thành một gói BSI , các gói lớp riêng có thể được dồn kênh để tạo gói khung thu được của dạng biểu diễn âm thanh nén hoàn chỉnh được chỉ báo bởi

$$[BSI \quad ESI_1 \quad BSRC_1 \quad \dots \quad BSRC_{(J_1)-1} \quad \dots \quad ESI_M \quad BSRC_{J(M-1)} \quad \dots \quad BSRC_J] \quad (5)$$

Về phân tải tin, gói khung thu được có thể được thể hiện bởi

$$FRAME = [\overline{BP}_1 \quad \dots \quad \overline{BP}_J \quad \overline{BSIP} \quad \overline{EP}_1 \quad \dots \quad \overline{EP}_M] \quad (6)$$

Gói khung thu được khi đó có thể được chuyển đến bộ giải mã hoặc bộ giải nén 4100. Nếu việc truyền lớp riêng không có lỗi, thì cờ hợp lệ của ít nhất là một phần tải tin thông tin phụ trợ tăng cường $\overline{EP}_m$ chứa trong đó (ví dụ, tương ứng với một phần thông tin phụ trợ tăng cường) được thiết lập là “đúng”. Trong trường hợp có lỗi do việc truyền lớp riêng thì 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 (ví dụ, từ cờ hợp lệ của nó).

Ở bộ giải nén 4100, gói khung thu được có thể được giải dồn kênh. 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.

Tại bước S3020, chỉ số lớp thứ nhất chỉ báo lớp cao nhất (ví dụ, lớp dùng được cao nhất hoặc lớp giải mã được cao nhất) được xác định trong số các lớp sẽ dùng để giải mã dạng biểu diễn âm thanh nén cơ bản thành dạng biểu diễn âm thanh tái tạo cơ bản của âm thanh hoặc trường âm thanh.

Hơn nữa, tại bước S3020, có thể lựa chọn giá trị (ví dụ, chỉ số lớp) N_B của lớp cao nhất (lớp dùng được cao nhất) mà sẽ được dùng để giải nén dạng biểu diễn âm thanh cơ bản. Lớp *tăng cường* cao nhất thực sự sẽ được dùng để giải nén dạng biểu diễn âm thanh cơ bản được thể hiện bởi $N_B - 1$. Do mỗi lớp chứa chính xác một tải tin thông tin phụ trợ tăng cường (phần thông tin phụ trợ tăng cường), nên có thể được xác định dựa trên tải tin thông tin phụ trợ tăng cường xem liệu lớp chứa có hợp lệ hay không (ví dụ, đã được nhận một cách hợp lệ hay chưa). 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 , $m = 1, \dots, M$ (hoặc tương ứng $\overline{EP}_m$, $m = 1, \dots, M$).

Tại bước S3030, thu được dạng biểu diễn âm thanh tái tạo cơ bản. Dạng biểu diễn âm thanh tái tạo cơ bản có thể thu được từ các thành phần được gán cho lớp dùng được cao nhất được chỉ báo bởi chỉ số lớp thứ nhất và lớp bất kỳ thấp hơn lớp dùng được cao nhất này, bằng cách sử dụng thông tin phụ trợ cơ bản (hoặc nói chung là sử dụng thông tin phụ trợ cơ bản).

Các tải tin của các thành phần dạng biểu diễn âm thanh nén cơ bản $BSRC_1, \dots, BSRC_J$ có thể được đưa đến bộ xử lý giải nén dạng biểu diễn cơ bản 4200, cùng với (tất cả) các tải tin thông tin phụ trợ cơ bản (ví dụ, BSI hoặc BSI_I và $BSI_{D,m}$, $m = 1, \dots, M$) và giá trị N_B . Bộ xử lý giải nén dạng biểu diễn cơ bản 4200 (được thể hiện trên Fig.4A và Fig.4B) tái tạo dạng biểu diễn âm thanh (hoặc trường âm thanh) cơ bản bằng cách chỉ sử dụng các thành phần dạng biểu diễn âm thanh nén cơ bản chứa trong các lớp N_B thấp nhất, tức là lớp cơ sở và $N_B - 1$ lớp tăng cường (tức là, các lớp lên đến lớp được chỉ báo bởi chỉ số lớp thứ nhất). Theo cách khác, chỉ các tải tin của các thành phần dạng biểu diễn âm thanh nén cơ bản chứa trong các lớp N_B thấp nhất cùng với các tải tin thông tin phụ trợ cơ bản tương ứng có thể được đưa đến bộ xử lý giải nén dạng biểu diễn cơ bản 4200.

Thông tin cần thiết về các thành phần nào của dạng biểu diễn âm thanh (hoặc trường âm thanh) nén cơ bản được chứa trong các lớp riêng rẽ được giả định là sẽ được biết bởi bộ giải nén 4100 từ gói dữ liệu có thông tin cấu hình, mà được giả định là được gửi và thu trước các gói dữ liệu khung.

Nhằm cung cấp các gói dữ liệu thông tin phụ trợ phụ thuộc $BSI_{D,m}$, $m = 1, \dots, N_B$ và gói dữ liệu thông tin phụ trợ tăng cường ESI_{N_E} , tất cả các tải tin tăng cường có thể được đưa vào bộ phân tích cú pháp một phần 4400 (xem Fig.4B) của bộ giải nén 4100 cùng với giá trị N_E và giá trị N_B . Bộ phân tích cú pháp có thể loại trừ tất cả các tải tin và các gói dữ liệu không dùng cho việc giải nén thực. Nếu giá trị N_E bằng không, tất cả các gói dữ liệu thông tin phụ trợ tăng cường có thể được giả định là rỗng.

Nếu lớp cơ sở bao gồm ít nhất một tải tin thông tin phụ trợ cơ bản phụ thuộc (phần thông tin phụ trợ cơ bản bổ sung) tương ứng với lớp tương ứng, việc giải mã mỗi tải tin thông tin phụ trợ cơ bản phụ thuộc riêng (ví dụ, $BSI_{D,m}$, $m = 1, \dots, N_B$ (phần thông tin phụ trợ cơ bản bổ sung)) có thể bao gồm các bước (i) giải mã phần thông tin phụ trợ cơ bản bổ sung bằng cách tham chiếu đến các thành phần được gán cho lớp tương ứng của nó và lớp bất kỳ thấp hơn lớp tương ứng (giải mã sơ bộ), và (ii) hiệu chỉnh phần thông tin phụ trợ cơ bản bổ sung bằng cách tham chiếu đến các thành phần được gán cho lớp dùng được cao nhất và lớp bất kỳ giữa lớp dùng được cao nhất và lớp tương ứng (hiệu chỉnh). Trong đó, thông tin phụ trợ cơ bản bổ sung tương ứng với lớp tương ứng bao gồm thông tin định rõ việc giải mã một hoặc nhiều thành phần trong số các thành phần được gán cho lớp tương ứng phụ thuộc vào các thành phần khác được gán cho lớp tương ứng và lớp bất kỳ thấp hơn lớp tương ứng.

Sau đó, dạng biểu diễn âm thanh tái tạo cơ bản có thể được thu (ví dụ, được tạo ra) từ các thành phần được gán cho lớp dùng được cao nhất và lớp bất kỳ thấp hơn lớp dùng được cao nhất, sử dụng thông tin phụ trợ cơ bản và các phần đã được hiệu chỉnh của thông tin phụ trợ cơ bản bổ sung thu được từ các phần thông tin phụ trợ cơ bản bổ sung tương ứng với các lớp lên đến lớp dùng được cao nhất.

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

Hiệu chỉnh liên tiếp mỗi tải tin $BSI_{D,m}$, $m = 1, \dots, N_B$, có thể bao gồm bước coi 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 cho giải mã sơ bộ. Vì vậy, việc hiệu chỉnh có thể được thực hiện bằng cách loại bỏ thông tin lỗi thời, điều này là có thể 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 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.

Tại bước S3040, chỉ số lớp thứ hai có thể được xác định. Chỉ số lớp thứ hai có thể chỉ báo (các) phần thông tin phụ trợ tăng cường nên dùng để cải thiện (ví dụ, nâng cao) dạng biểu diễn âm thanh tái tạo cơ bản.

Ngoài chỉ số lớp thứ nhất, có thể xác định chỉ số (chỉ số lớp thứ hai) N_E của tải tin thông tin phụ trợ tăng cường (một phần thông tin tăng cường thứ hai) sẽ dùng để giải nén. Chỉ số lớp thứ hai N_E có thể luôn luôn bằng chỉ số lớp thứ nhất N_B hoặc bằng không. Việc tăng cường có thể luôn luôn được thực hiện dựa theo dạng biểu diễn âm thanh cơ bản thu được từ lớp dùng được cao nhất, hoặc hoàn toàn không theo dạng biểu diễn âm thanh cơ bản này.

Tại bước S3050, dạng biểu diễn âm thanh tái tạo của âm thanh hoặc trường âm thanh được thu (ví dụ, được tạo ra) từ dạng biểu diễn âm thanh tái tạo cơ bản, tham chiếu đến chỉ số lớp thứ hai.

Tức là, dạng biểu diễn âm thanh tái tạo được thu bằng cách cải thiện hoặc tăng cường (bằng thông số) dạng biểu diễn âm thanh tái tạo cơ bản, chẳng hạn như bằng cách sử dụng thông tin phụ trợ tăng cường (phần thông tin phụ trợ tăng cường) được chỉ báo bởi chỉ số lớp thứ hai. Như được đề cập thêm dưới đây, chỉ số lớp thứ hai có thể chỉ báo hoàn toàn không dùng thông tin phụ trợ tăng cường bất kỳ ở giai đoạn này. Khi đó, dạng biểu diễn âm thanh tái tạo sẽ tương ứng với dạng biểu diễn âm thanh tái tạo cơ bản.

Với mục đích này, dạng biểu diễn âm thanh cơ bản tái tạo cùng với tất cả các tải tin thông tin phụ trợ tăng cường $ESI_1, \dots, ESI_M$, tải tin thông tin phụ trợ cơ bản (ví dụ, BSI hoặc BSI_I và $BSI_{D,m}$, $m = 1, \dots, M$), và giá trị N_E được đưa đến bộ xử lý giải nén dạng biểu diễn tăng cường 4300 (được thể hiện trên Fig.4A và Fig.4B), bộ xử lý này tính toán dạng biểu diễn âm thanh (hoặc trường âm thanh) tăng cường cuối cùng 2100' bằng cách chỉ sử dụng tải tin thông tin phụ trợ tăng cường ESI_{N_E} và loại trừ tất cả các tải tin thông tin phụ trợ tăng cường khác. Theo cách khác, chỉ có tải tin thông tin phụ trợ tăng cường ESI_{N_E} , chứ không phải là tất cả các tải tin thông tin phụ trợ tăng cường, có thể được đưa đến bộ xử lý giải nén dạng biểu diễn tăng cường 4300. Nếu giá trị N_E bằng không, tất cả các tải tin thông tin phụ trợ tăng cường sẽ bị loại trừ (hoặc thay vào đó, không có tải tin thông tin phụ trợ tăng cường được cung cấp) và dạng biểu diễn âm thanh tăng cường cuối cùng được tái tạo 2100' là bằng dạng biểu diễn âm thanh cơ bản

tái tạo. Tải tin thông tin phụ trợ tăng cường ESI_{NE} có thể đã được thu bởi bộ phân tích cú pháp một phần 4400.

Nhìn chung, Fig.3 còn thể hiện việc giải mã dạng biểu diễn HOA nén dựa trên thông tin phụ trợ cơ bản được kết hợp với lớp cơ sở và dựa trên thông tin phụ trợ tăng cường được kết hợp với một hoặc nhiều lớp tăng cường phân cấp.

Trừ khi các bước đòi hỏi 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.3 được hiểu là không làm giới hạn sáng chế.

Sau đây, chi tiết về việc lựa chọn lớp để giải nén (chọn các chỉ số lớp thứ nhất và thứ hai) tại các bước S3020 và S3040 sẽ được mô tả.

Việc xác định chỉ số lớp thứ nhất có thể bao gồm bước xác định, đối với mỗi lớp, liệu lớp tương ứng đã được nhận một cách hợp lệ hay chưa. Việc xác định chỉ số lớp thứ nhất còn có thể bao gồm bước xác định chỉ số lớp thứ nhất là chỉ số lớp của lớp ngay dưới lớp dưới cùng không được nhận một cách hợp lệ. Việc lớp đã được nhận một cách hợp lệ hay chưa có thể được xác định bằng cách đánh giá liệu tải tin thông tin phụ trợ tăng cường của lớp đó đã được nhận một cách hợp lệ hay chưa. Việc này sau đó có thể được thực hiện bằng cách đánh giá các cờ hợp lệ trong các tải tin thông tin phụ trợ tăng cường.

Bước xác định chỉ số lớp thứ hai có thể thường đòi hỏi xác định chỉ số lớp thứ hai là bằng chỉ số lớp thứ nhất, hoặc xác định một giá trị chỉ số là chỉ số lớp thứ hai (ví dụ, giá trị chỉ số 0) mà chỉ báo không dùng thông tin phụ trợ tăng cường bất kỳ khi thu được dạng biểu diễn âm thanh tái tạo.

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 (lớp dùng được cao nhất) cần dùng thực tế để giải nén dạng biểu diễn âm thanh cơ bản và chỉ số N_E của tải tin thông tin phụ trợ tăng cường cần dùng để giải nén có thể 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.

Tức là, chỉ số lớp thứ hai có thể được xác định là bằng chỉ số lớp thứ nhất nếu các dạng biểu diễn âm thanh nén trong các khoảng thời gian liên tiếp có thể được giải mã độc lập. Trong trường hợp này, dạng biểu diễn âm thanh cơ bản tái tạo có thể được tăng cường dựa trên tải tin thông tin phụ trợ tăng cường của lớp dùng được cao nhất.

Trong trường hợp sử dụng giải nén vi sai với sự phụ thuộc nội khung, quyết định từ khung trước đó phải được xem xét bổ sung. Lưu ý rằng với giải nén vi sai, thì thường 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 sẽ độc lập theo khung và được tiến hành như mô tả ở trên.

Để giải thích cho quyết định phụ thuộc khung đề xuất, số cao nhất (ví dụ, chỉ số lớp) của tải tin thông tin phụ trợ tăng cường hợp lệ cho khung thứ k được biểu thị bởi $L(k)$, số lớp cao nhất (ví dụ, chỉ số lớp) cần được chọn và dùng để giải nén dạng biểu diễn âm thanh cơ bản là $N_B(k)$, và số (ví dụ, chỉ số lớp) tải tin thông tin phụ trợ tăng cường cần dùng để giải nén là $N_E(k)$.

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

$$N_B(k) = \min(N_B(k-1), L(k)). \quad (7)$$

Bằng việc 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 vi sai cho dạng biểu diễn âm thanh cơ bản là khả dụng.

Tức là, nếu các dạng biểu diễn âm thanh nén trong các khoảng thời gian (ví dụ, các khung) liên tiếp không thể được giải mã độc lập với nhau, việc xác định chỉ số lớp thứ nhất có thể bao gồm bước xác định, đối với mỗi lớp, liệu lớp tương ứng đã được nhận một cách hợp lệ hay chưa, và xác định chỉ số lớp thứ nhất cho khoảng thời gian đã cho là chỉ số nhỏ hơn trong số chỉ số lớp thứ nhất của khoảng thời gian ở trước khoảng thời gian đã cho và chỉ số lớp của lớp ngay dưới lớp dưới cùng không được nhận một cách hợp lệ.

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

$$N_E(k) = \begin{cases} N_B(k) & \text{nếu } N_B(k) = N_B(k-1) \\ 0 & \text{nếu không} \end{cases}. \quad (8)$$

Trong đó, việc lựa chọn 0 cho $N_E(k)$ chỉ báo rằng dạng biểu diễn âm thanh cơ bản tái tạo không được cải thiện hoặc tăng cường bằng cách sử dụng thông tin phụ trợ tăng cường.

Cụ thể điều này nghĩa là miễn là số lượng lớp cao nhất $N_B(k)$ cần dùng để giải nén dạng biểu diễn âm thanh cơ bản không thay đổi, cùng số lớp tăng cường tương ứng được chọn. Tuy nhiên, trong trường hợp thay đổi $N_B(k)$, sự tăng cường bị vô hiệu hóa bằng cách thiết lập $N_E(k)$ về không. Do việc giải nén vi sai được giả định cho thông tin phụ trợ tăng cường, sự thay đổi của nó theo $N_B(k)$ là không thể do nó đòi hỏi giải nén lớp thông tin phụ trợ tăng cường tương ứng ở khung trước đó mà được giả định là chưa được tiến hành.

Tức là, nếu các dạng biểu diễn âm thanh nén trong các khoảng thời gian (ví dụ, các khung) liên tiếp không thể được giải mã độc lập với nhau, việc xác định chỉ số lớp thứ hai có thể bao gồm bước xác định liệu chỉ số lớp thứ nhất cho khoảng thời gian đã cho có bằng chỉ số lớp thứ nhất cho khoảng thời gian trước đó hay không. Nếu chỉ số lớp thứ nhất cho khoảng thời gian đã cho bằng chỉ số lớp thứ nhất cho khoảng thời gian trước đó, thì chỉ số lớp thứ hai cho khoảng thời gian đã cho có thể được xác định (ví dụ, được chọn) là bằng chỉ số lớp thứ nhất cho khoảng thời gian đã cho. Mặt khác, nếu chỉ số lớp thứ nhất cho khoảng thời gian đã cho không bằng chỉ số lớp thứ nhất cho khoảng thời gian trước đó, giá trị chỉ số có thể được xác định (ví dụ, lựa chọn) là chỉ số lớp thứ hai chỉ báo không sử dụng thông tin phụ trợ tăng cường bất kỳ khi thu được dạng biểu diễn âm thanh tái tạo.

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 trong phương trình (4) có thể được thay thế bằng

$$N_E(k) = N_B(k). \quad (9)$$

Cuối cùng, cần lưu ý rằng đối với giải nén vi sai, số lượng của lớp dùng được cao nhất N_B có thể chỉ tăng ở các gói dữ liệu khung độc lập, trong khi đó có thể giảm ở mọi khung.

Cần 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 mô tả ở trên. Ví dụ về bộ mã hóa 5000 này được minh họa bằng sơ đồ trên Fig.5. Ví dụ, bộ mã hóa 5000 này có thể bao gồm bộ chia nhỏ thành phần 5010 được làm thích ứng để thực hiện bước S1010 nêu trên, bộ gán thành phần 5020 được làm thích ứng để thực hiện bước S1020 nêu trên, bộ gán thông tin phụ trợ cơ bản 5030 được làm thích ứng để thực hiện bước S1030 nêu trên, bộ phân chia thông tin phụ trợ tăng cường 5040 được làm thích ứng để thực hiện S1040 nêu trên, và bộ gán thông tin phụ trợ tăng cường 5050 được làm thích ứng để thực hiện bước S1050 nêu trên. Cần hiểu thêm rằng các đơn vị tương ứng của bộ mã hóa này có thể được thực thi bởi bộ xử lý 5100 của thiết bị tính toán mà được làm thích ứng để thực hiện việc xử lý được 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, cũng như các bước bất kỳ khác của phương pháp mã hóa được đề xuất. Bộ mã hóa hoặc thiết bị tính toán còn có thể bao gồm bộ nhớ 5200 mà có thể truy cập được bởi bộ xử lý 5100.

Cần 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 mã hóa 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 6000 được minh họa bằng sơ đồ trên Fig.6. Ví dụ, bộ giải mã 6000 này có thể bao gồm bộ thu 6010 được làm thích ứng để thực hiện bước S3010 nêu trên, bộ xác định chỉ số lớp thứ nhất 6020 được làm thích ứng để thực hiện bước S3020 nêu trên, bộ tái tạo cơ bản 6030 được làm thích ứng để thực hiện bước S3030 nêu trên, bộ xác định chỉ số lớp thứ hai 6040 được làm thích ứng để thực hiện S3040 nêu trên, và bộ tái tạo tăng cường 6050 được làm thích ứng để thực hiện bước S3050 nêu trên. Cần hiểu thêm rằng các đơn vị tương ứng của bộ giải mã này có thể được thực thi bởi bộ xử lý 6100 của thiết bị tính toán mà được làm thích ứng để

thực hiện việc xử lý được 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, cũng như các bước bất kỳ khác của phương pháp giải mã được đề xuất. Bộ giải mã hoặc thiết bị tính toán còn có thể bao gồm bộ nhớ 6200 mà truy cập được bởi bộ xử lý 6100.

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 hiểu rằng có thể nghĩ ra các cách sắp xếp khác nhau mà, mặc dù không được mô tả và thể hiện rõ ràng ở đây, cụ thể hóa các nguyên lý của sáng chế và được bao gồm trong ý tưởng và phạm vi của sáng chế. Hơn thế nữa, tất cả các ví dụ được nêu ở đây 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 ở đây đề cập đến các nguyên lý, khía cạnh và phương án của sáng chế, cũng như các ví dụ của chúng, được dự tính bao gồm các dạng tương đương của chúng.

Phương pháp và thiết bị mô tả trong sáng chế có thể được ứng dụng là phần mềm, phần sụn và/hoặc phần cứng. Một số thành phần ví dụ có thể được ứng dụng là phần mềm chạy trên bộ xử lý hoặc bộ vi xử lý tín hiệu số. Các thành phần khác ví dụ có thể được ứng dụng là phần cứng và hoặc dạng mạch tích hợp chuyên dụng. 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 truyền qua các loại mạng, như mạng vô tuyến, mạng vệ tinh, mạng không dây hoặc mạng có dây, ví dụ mạng Internet.

Đối chứng 1: ISO/IEC JTC1/SC29/WG11 23008-3:2015(E). Công nghệ thông tin – Mã hóa hiệu suất cao và phân phối phương tiện trong các môi trường không đồng nhất – Phần 3: Âm thanh 3D, tháng 2/2015.

Đối chứng 2: ISO/IEC JTC1/SC29/WG11 23008-3:2015/PDAM3. Công nghệ thông tin – Mã hóa hiệu suất cao và phân phối phương tiện trong các môi trường không đồng nhất – Phần 3: Âm thanh 3D, SỬA ĐỔI LẦN 3: Âm thanh MPEG-H 3D pha 2, tháng 7/2015.

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 hai hoặc nhiều lớp tăng cường phân cấp, và chứa thông tin phụ trợ cơ bản được kết hợp với lớp cơ sở và thông tin phụ trợ tăng cường được kết hợp với hai 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,

trong đó hai hoặc nhiều lớp tăng cường phân cấp bao gồm lớp tăng cường phân cấp dùng được cao nhất, và

trong đó mỗi lớp trong số hai hoặc nhiều lớp tăng cường phân cấp bao gồm một phần thông tin phụ trợ tăng cường bao gồm các thông số để cải thiện dạng biểu diễn âm thanh tái tạo cơ bản có thể thu được từ dữ liệu chứa trong lớp tương ứng và lớp bất kỳ thấp hơn lớp tương ứng; và

giải mã dạng biểu diễn HOA nén dựa trên thông tin phụ trợ cơ bản được kết hợp với lớp cơ sở, dựa trên phần thông tin phụ trợ tăng cường được kết hợp với lớp tăng cường phân cấp dùng được cao nhất, và không dựa trên phần thông tin phụ trợ tăng cường được kết hợp với lớp khác bất kỳ trong số hai hoặc nhiều lớp tăng cường phân cấp.

2. Phương pháp theo điểm 1, trong đó thông tin phụ trợ tăng cường bao gồm các thông số liên quan đến ít nhất một trong số: dự đoán theo không gian, tổng hợp tín hiệu định hướng dải con, và sao chép thông số xung quanh.

3. Phương pháp theo điểm 1, trong đó thông tin phụ trợ tăng cường bao gồm thông tin cho phép dự đoán các phần bỏ sót của âm thanh hoặc trường âm thanh từ tín hiệu định hướng.

4. Phương pháp theo điểm 1, phương pháp này còn bao gồm các bước:

xác định, đối với mỗi lớp, liệu lớp tương ứng đã được nhận một cách hợp lệ hay chưa; và

xác định chỉ số lớp của lớp ngay dưới lớp thấp nhất chưa được nhận hợp lệ.

5. Phương pháp theo điểm 4, trong đó phương pháp này còn bao gồm bước xác định chỉ số lớp tiếp theo bằng chỉ số lớp hoặc chỉ báo sự bỏ sót thông tin phụ trợ tăng cường trong khi giải mã.

6. Phương pháp theo điểm 1, trong đó lớp cơ sở bao gồm ít nhất một phần thông tin phụ trợ cơ bản bổ sung tương ứng với lớp tương ứng và bao gồm thông tin định rõ việc giải mã một hoặc nhiều thành phần trong số các thành phần được gán cho lớp tương ứng phụ thuộc vào các thành phần khác được gán cho lớp tương ứng và lớp bất kỳ thấp hơn lớp tương ứng, phương pháp này bao gồm, đối với mỗi phần thông tin phụ trợ cơ bản bổ sung, các bước:

giải mã phần thông tin phụ trợ cơ bản bổ sung bằng cách tham chiếu đến các thành phần được gán cho lớp tương ứng của nó và lớp bất kỳ thấp hơn lớp tương ứng; và

hiệu chỉnh phần thông tin phụ trợ cơ bản bổ sung bằng cách tham chiếu đến các thành phần được gán cho lớp tăng cường phân cấp dùng được cao nhất và lớp bất kỳ giữa lớp tăng cường phân cấp dùng được cao nhất và lớp tương ứng,

trong đó dạng biểu diễn âm thanh tái tạo cơ bản thu được từ các thành phần được gán cho lớp tăng cường phân cấp dùng được cao nhất và lớp bất kỳ thấp hơn lớp tăng cường phân cấp dùng được cao nhất này, sử dụng thông tin phụ trợ cơ bản và các phần đã được hiệu chỉnh của thông tin phụ trợ cơ bản bổ sung thu được từ các phần thông tin phụ trợ cơ bản bổ sung tương ứng với các lớp lên đến lớp tăng cường phân cấp dùng được cao nhất.

7. 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 để thu dòng bit chứa dạng biểu diễn HOA nén tương ứng với các lớp phân cấp bao gồm lớp cơ sở và hai hoặc nhiều lớp tăng cường phân cấp, và chứa thông tin

phụ trợ cơ bản được kết hợp với lớp cơ sở và thông tin phụ trợ tăng cường được kết hợp với hai hoặc nhiều lớp tăng cường phân cấp,

trong đó nhiều lớp đã gắn 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 gắn cho các lớp tương ứng trong các nhóm thành phần tương ứng,

trong đó hai hoặc nhiều lớp tăng cường phân cấp bao gồm lớp tăng cường phân cấp dùng được cao nhất, và

trong đó mỗi lớp trong số hai hoặc nhiều lớp tăng cường phân cấp bao gồm một phần thông tin phụ trợ tăng cường bao gồm các thông số để cải thiện dạng biểu diễn âm thanh tái tạo cơ bản có thể thu được từ dữ liệu chứa trong lớp tương ứng và lớp bất kỳ thấp hơn lớp tương ứng; và

bộ giải mã để giải mã dạng biểu diễn HOA nén dựa trên thông tin phụ trợ cơ bản được kết hợp với lớp cơ sở, dựa trên phần thông tin phụ trợ tăng cường được kết hợp với lớp tăng cường phân cấp dùng được cao nhất, và không dựa trên phần thông tin phụ trợ tăng cường được kết hợp với lớp khác bất kỳ trong số hai hoặc nhiều lớp tăng cường phân cấp.

8. Thiết bị theo điểm 7, trong đó thông tin phụ trợ tăng cường bao gồm các thông số liên quan đến ít nhất một trong số: dự đoán theo không gian, tổng hợp tín hiệu định hướng dải con, và sao chép thông số xung quanh.

9. Thiết bị theo điểm 7, trong đó thông tin phụ trợ tăng cường bao gồm thông tin cho phép dự đoán các phân bố sóng của âm thanh hoặc trường âm thanh từ tín hiệu định hướng.

10. Thiết bị theo điểm 7, thiết bị này được tạo cấu hình để:

xác định, đối với mỗi lớp, liệu lớp tương ứng đã nhận được hợp lệ hay chưa; và xác định chỉ số lớp của lớp ngay dưới lớp thấp nhất chưa được nhận hợp lệ.

11. Thiết bị theo điểm 10, thiết bị này còn được tạo cấu hình để xác định chỉ số lớp tiếp theo bằng chỉ số lớp hoặc chỉ báo sự bỏ sót thông tin phụ trợ tăng cường trong khi giải mã.

12. Thiết bị theo điểm 7,

trong đó lớp cơ sở bao gồm ít nhất một phần thông tin phụ trợ cơ bản bổ sung tương ứng với lớp tương ứng và bao gồm thông tin định rõ việc giải mã một hoặc nhiều thành phần trong số các thành phần được gán cho lớp tương ứng phụ thuộc vào các thành phần khác được gán cho lớp tương ứng và lớp bất kỳ thấp hơn lớp tương ứng, và

trong đó, đối với mỗi phần thông tin phụ trợ cơ bản bổ sung, thiết bị được tạo cấu hình để:

giải mã phần thông tin phụ trợ cơ bản bổ sung bằng cách tham chiếu đến các thành phần được gán cho lớp tương ứng của nó và lớp bất kỳ thấp hơn lớp tương ứng; và

hiệu chỉnh phần thông tin phụ trợ cơ bản bổ sung bằng cách tham chiếu đến các thành phần được gán cho lớp tăng cường phân cấp dùng được cao nhất và lớp bất kỳ giữa lớp tăng cường phân cấp dùng được cao nhất và lớp tương ứng,

trong đó dạng biểu diễn âm thanh tái tạo cơ bản thu được từ các thành phần được gán cho lớp tăng cường phân cấp dùng được cao nhất và lớp bất kỳ thấp hơn lớp tăng cường phân cấp dùng được cao nhất này, sử dụng thông tin phụ trợ cơ bản và các phần đã được hiệu chỉnh của thông tin phụ trợ cơ bản bổ sung thu được từ các phần thông tin phụ trợ cơ bản bổ sung tương ứng với các lớp lên đến lớp tăng cường phân cấp dùng được cao nhất.

13. Phương tiện bất biến đọc được bằng máy tính bao gồm các lệnh đọc được bằng máy tính mà, khi được thực thi bởi một hoặc nhiều bộ xử lý của thiết bị tính toán, sẽ khiến cho thiết bị tính toán này thực hiện phương pháp theo điểm 1.

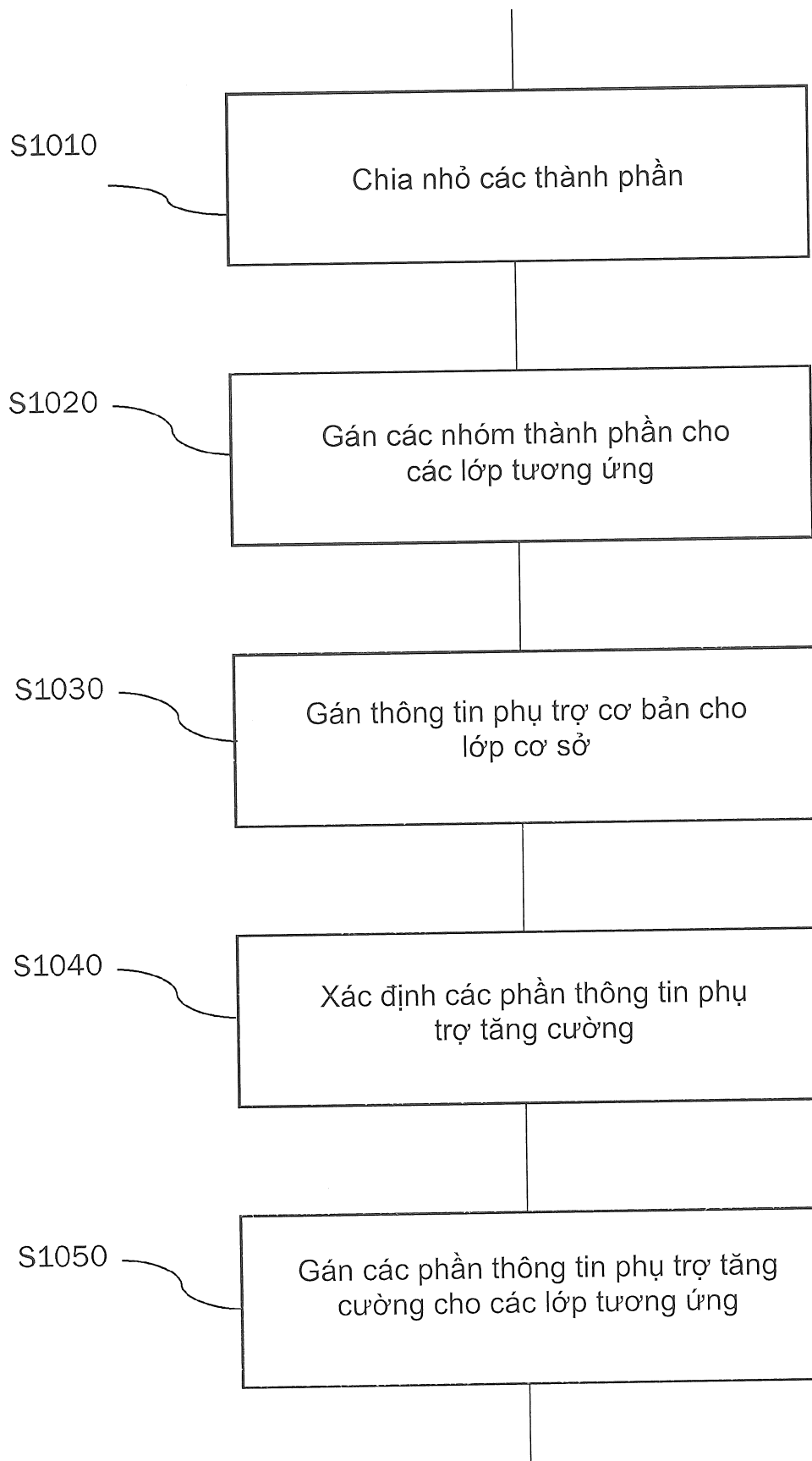


Fig.1

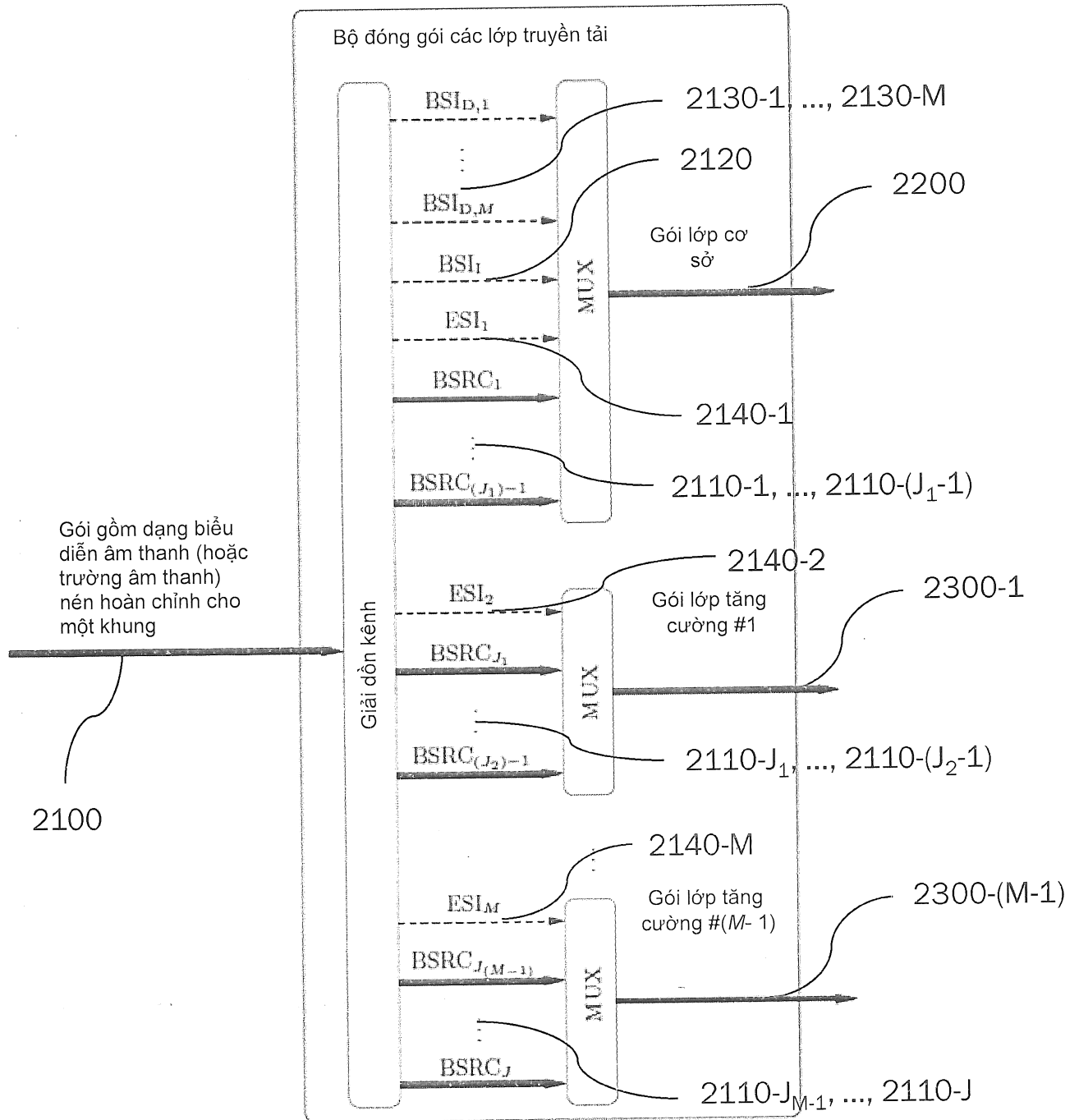


Fig.2

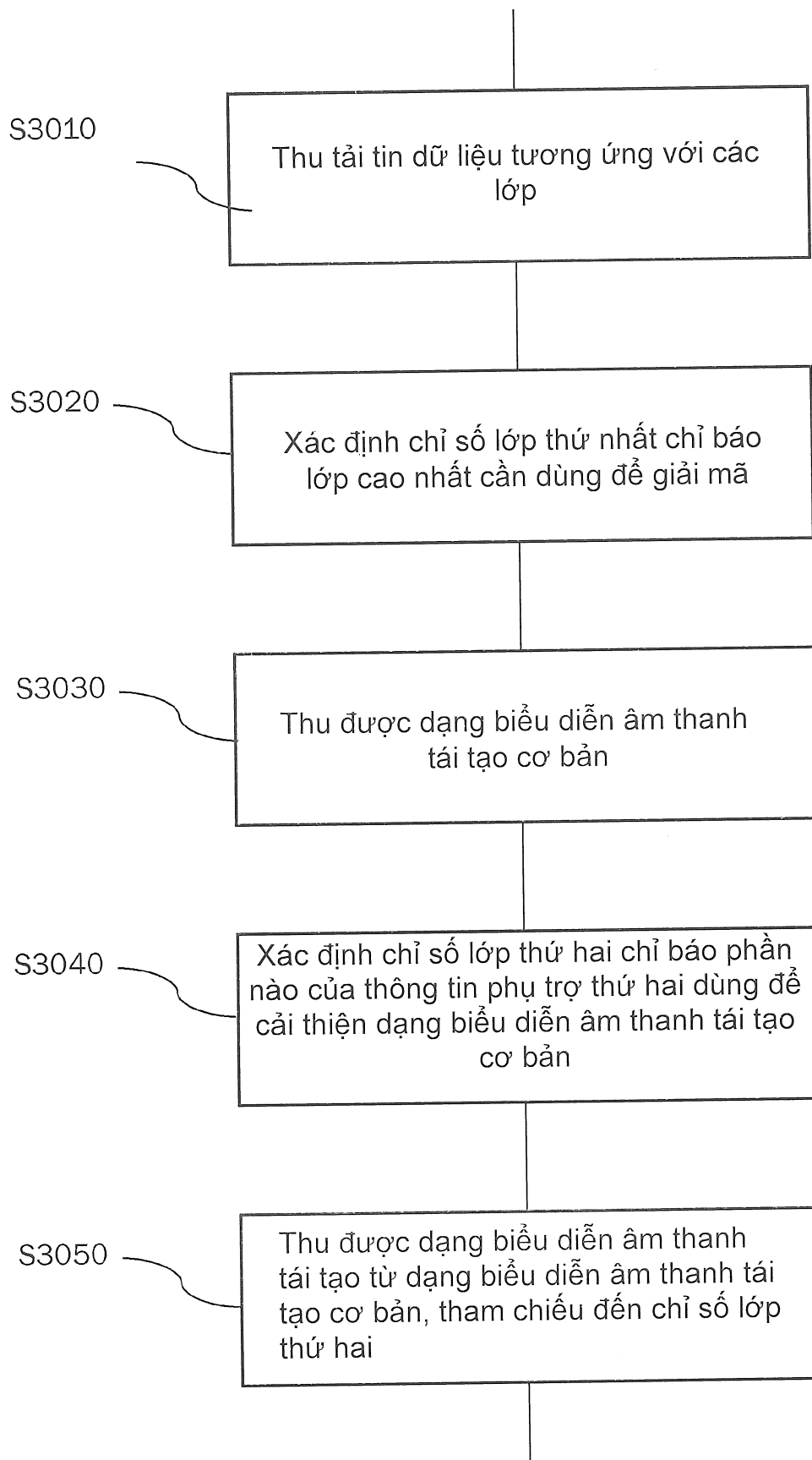


Fig.3

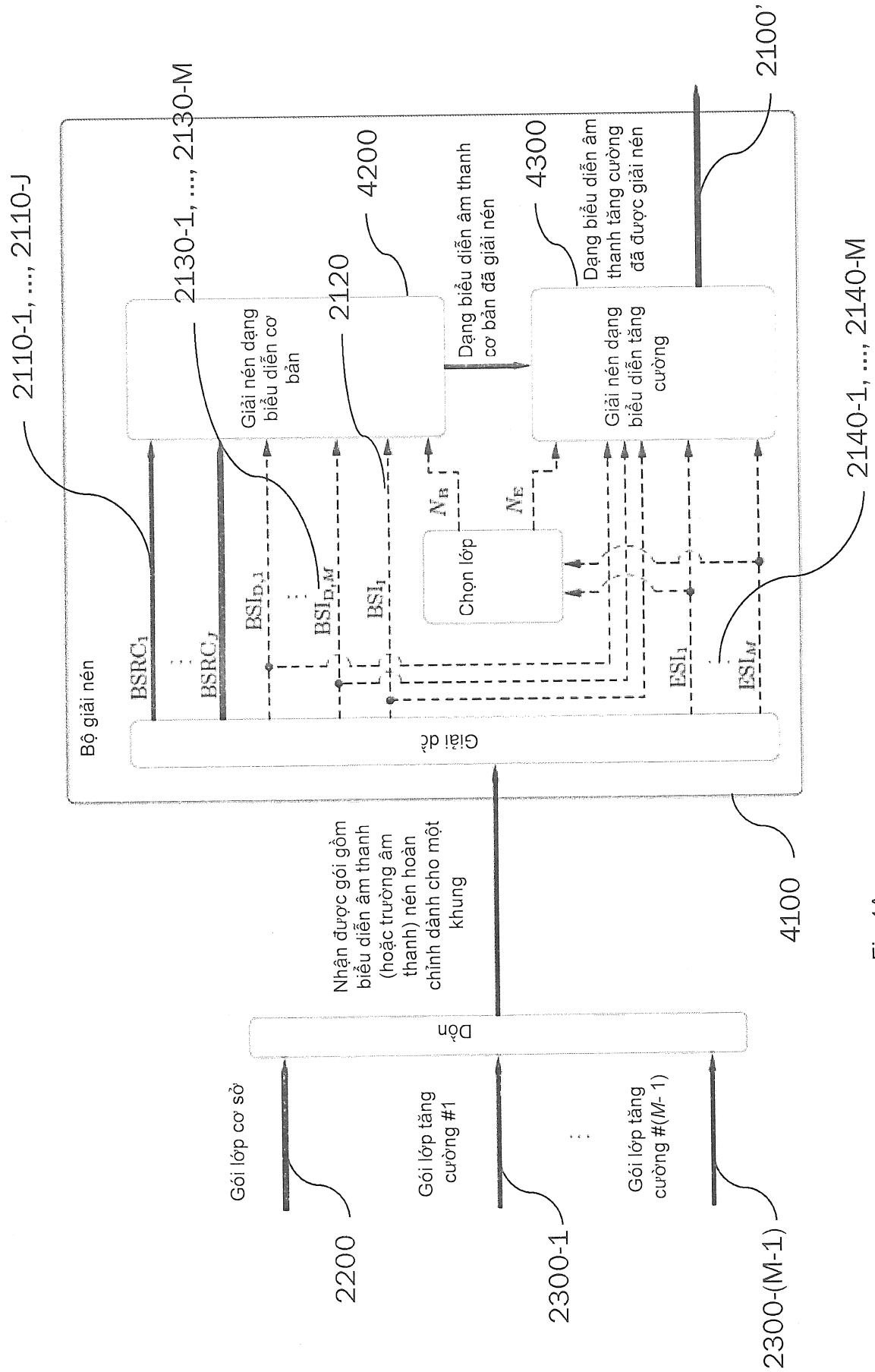


Fig. 4A

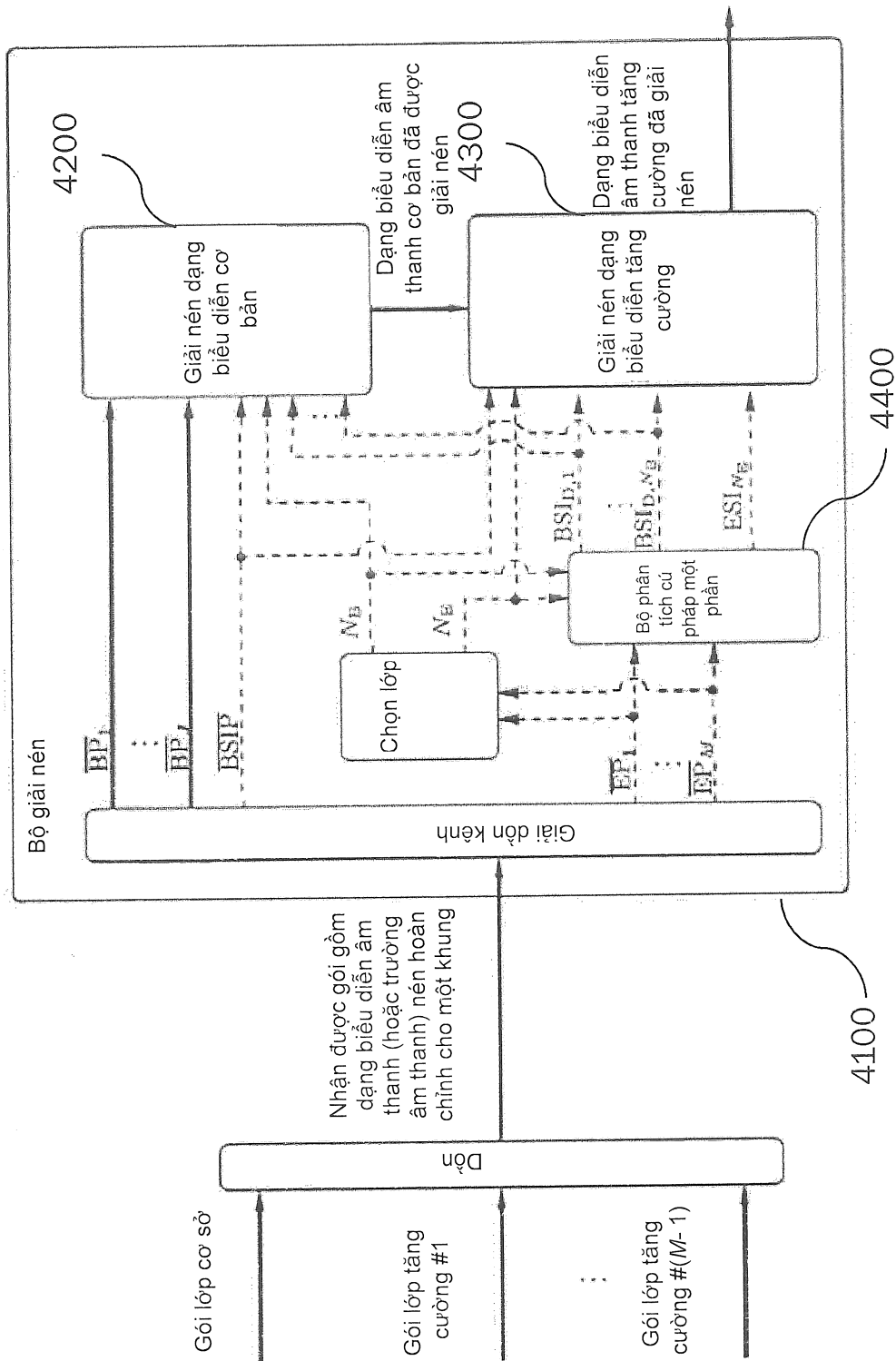


Fig.4B

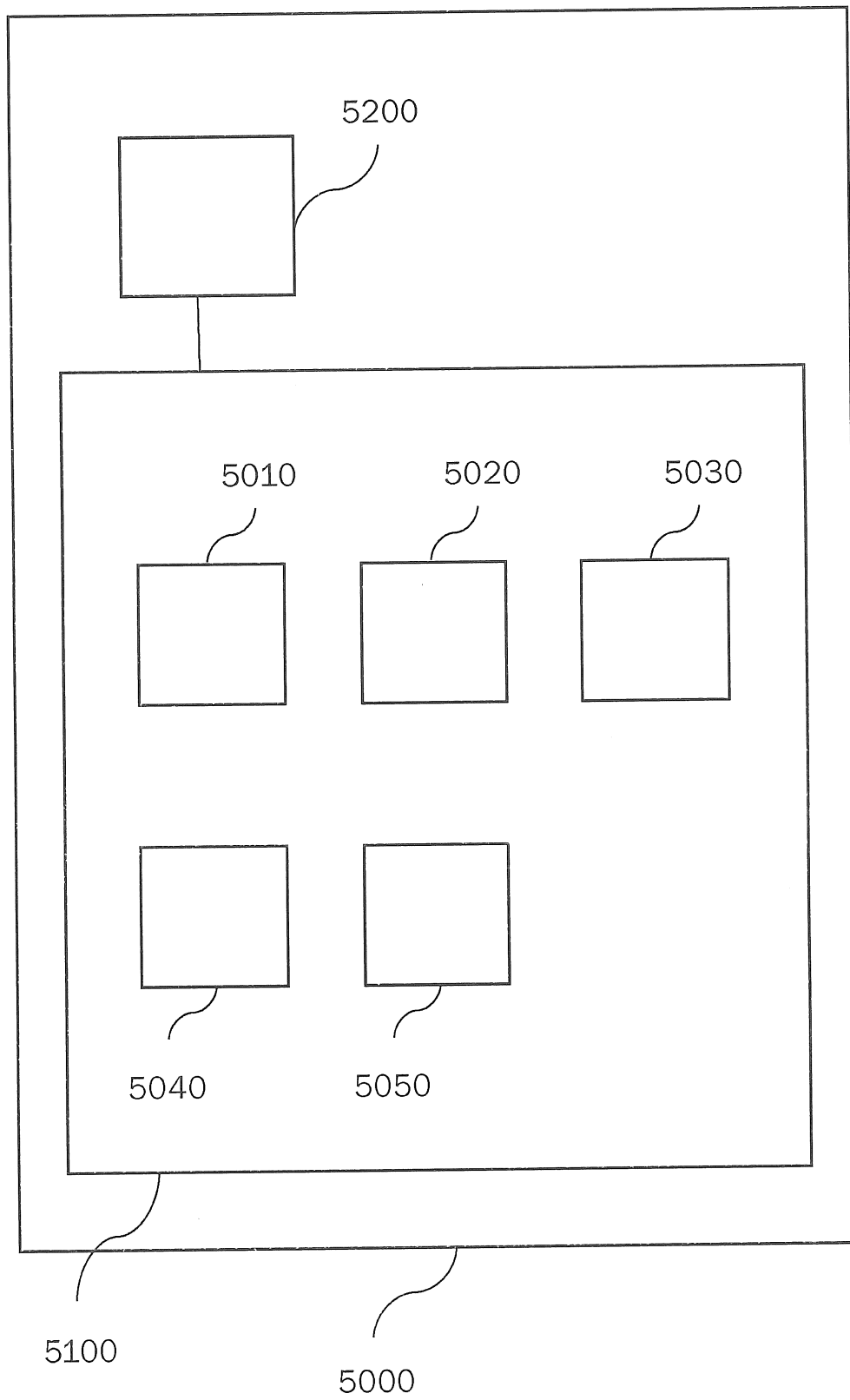


Fig.5

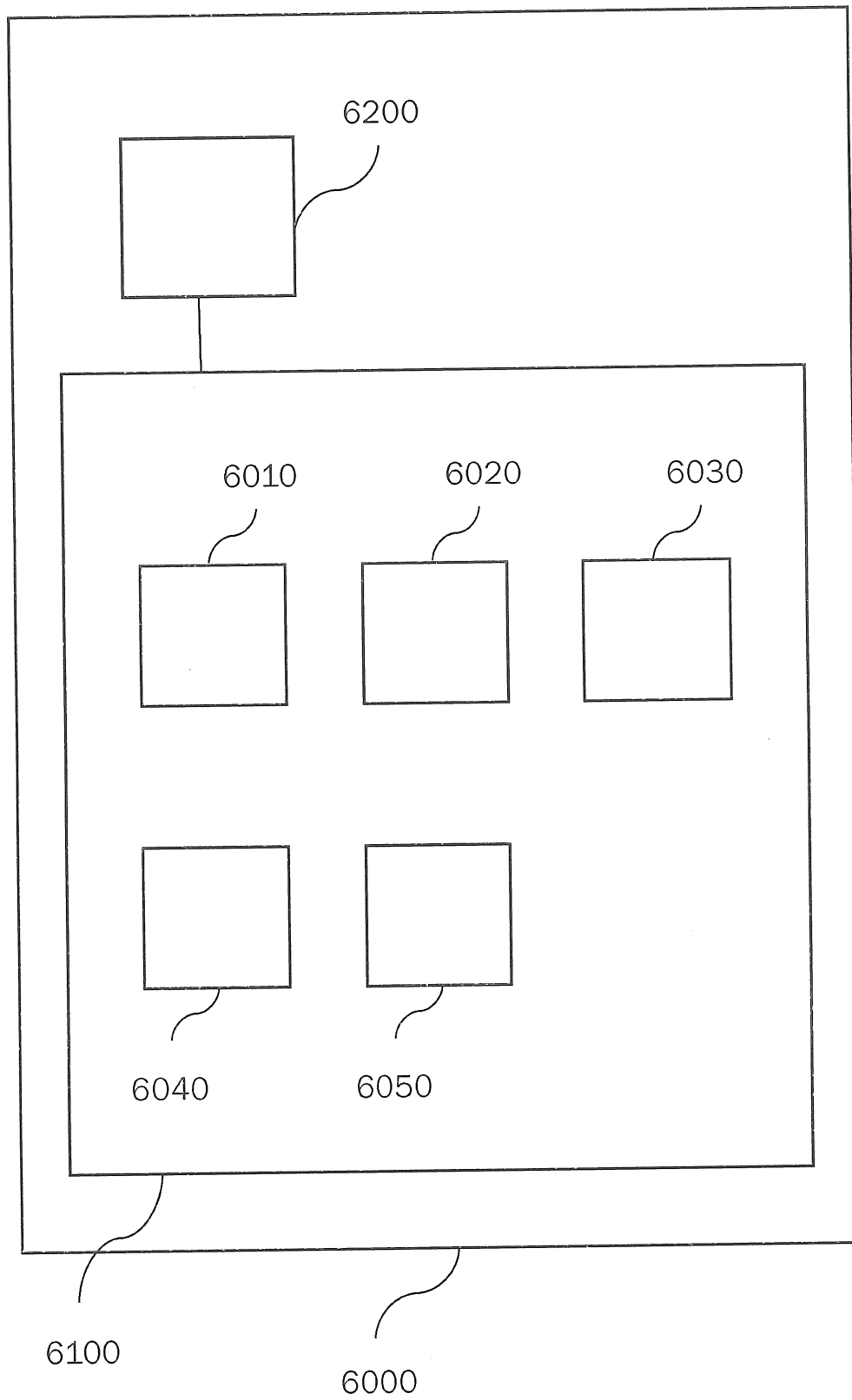


Fig.6