



(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)



1-0043608

(51)^{2020.01} G10L 19/00; G10L 19/008

(13) B

(21) 1-2020-03429

(22) 16/11/2018

(86) PCT/EP2018/081620 16/11/2018

(87) WO 2019/097017 23/05/2019

(30) 17202393.9 17/11/2017 EP

(45) 25/02/2025 443

(43) 25/08/2020 389A

(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e. V. (DE)
Hansastrasse 27c, 80686 Muenchen, Germany

(72) FUCHS, Guillaume (FR); HERRE, Juergen (DE); KUECH, Fabian (DE); DOEHLA, Stefan (DE); MULTRUS, Markus (DE); THIERGART, Oliver (DE); WUEBBOLT, Oliver (DE); GHIDO, Florin (RO); BAYER, Stefan (AT); JAEGER, Wolfgang (DE).

(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS HANOI)

(54) THIẾT BỊ VÀ PHƯƠNG PHÁP MÃ HÓA HOẶC GIẢI MÃ CÁC THAM SỐ MÃ HÓA ÂM THANH CÓ HƯỚNG SỬ DỤNG CÁC ĐỘ PHÂN GIẢI THỜI GIAN/TẦN SỐ KHÁC NHAU

(21) 1-2020-03429

(57) Sáng chế đề cập đến thiết bị và phương pháp mã hóa hoặc giải mã các tham số mã hóa âm thanh có hướng sử dụng các độ phân giải thời gian/tần số khác nhau. Thiết bị mã hóa các tham số mã hóa âm thanh có hướng bao gồm các tham số độ khuếch tán và tham số hướng, bao gồm: bộ tính tham số (100) để tính toán các tham số độ khuếch tán với độ phân giải thời gian hoặc tần số thứ nhất và để tính toán các tham số hướng với độ phân giải thời gian hoặc tần số thứ hai; và bộ xử lý bộ lượng tử hóa và bộ mã hóa (200) để tạo sự biểu diễn được lượng tử hóa và được mã hóa của các tham số độ khuếch tán và các tham số hướng.

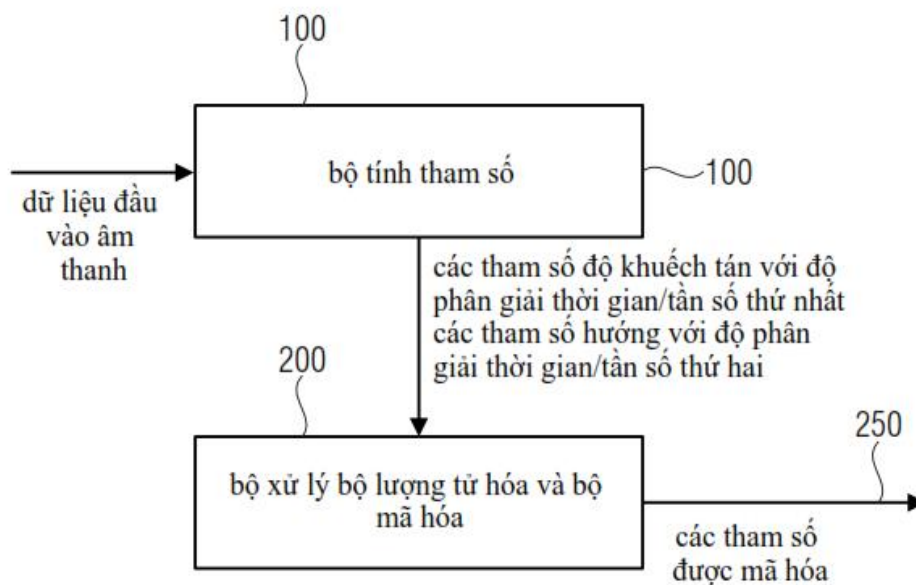


Fig. 2a

Lĩnh vực kỹ thuật được đề cập

Sáng chế hướng đến việc xử lý tín hiệu âm thanh và đặc biệt là các sơ đồ mã hóa hiệu quả của các tham số mã hóa âm thanh có hướng chẳng hạn như siêu dữ liệu DirAC.

Sáng chế nhằm mục đích đề xuất giải pháp mã hóa tốc độ bit thấp để mã hóa siêu dữ liệu không gian từ phép phân tích cảnh âm thanh 3D được thực hiện bởi mã hóa âm thanh có hướng (directional audio coding - DirAC), một kỹ thuật được thúc đẩy cảm giác để xử lý âm thanh không gian.

Tình trạng kỹ thuật của sáng chế

Việc truyền cảnh âm thanh theo ba chiều đòi hỏi phải xử lý nhiều kênh mà thường tạo ra một lượng lớn dữ liệu để truyền. Kỹ thuật mã hóa âm thanh có hướng (DirAC) [1] là một cách tiếp cận hiệu quả để phân tích cảnh âm thanh và biểu diễn nó theo tham số. DirAC sử dụng sự biểu diễn thúc đẩy cảm giác của trường âm thanh dựa trên hướng đến (direction of arrival - DOA) và độ khuếch tán đo được trên mỗi dải tần. Nó được xây dựng dựa trên giả định rằng tại một thời điểm tức thời và đối với một dải tới hạn, độ phân giải không gian của hệ thống thính giác bị giới hạn để giải mã một tín hiệu cho hướng và một tín hiệu khác cho sự kết hợp giữa các âm thanh nghe được. Sau đó, âm thanh không gian được tái tạo trong miền tần số bằng cách giảm dần âm hai luồng: luồng khuếch tán không có hướng và luồng không khuếch tán có hướng.

Sáng chế bộc lộ phương pháp mã hóa âm thanh 3D dựa trên sự biểu diễn và tái tạo âm thanh DirAC để đạt được việc truyền nội dung âm thanh đắm chìm ở các tốc độ bit thấp.

DirAC là kỹ thuật tái tạo âm thanh không gian thúc đẩy cảm giác. Người ta cho rằng tại một thời điểm tức thời và đối với một dải tới hạn, độ phân giải không gian của hệ thống thính giác bị giới hạn để giải mã một tín hiệu cho hướng và một tín hiệu khác cho sự kết hợp giữa các âm thanh nghe được.

Dựa trên các giả định này, DirAC biểu diễn âm thanh không gian trong một dải tần bằng cách giảm dần âm hai luồng: luồng khuếch tán không có hướng và luồng không khuếch tán có hướng. Quá trình xử lý DirAC được thực hiện theo hai giai đoạn: phân tích và tổng hợp như trên Fig.10a và Fig.10b.

Trong giai đoạn phân tích DirAC, micrô trùng khớp bậc 1 ở định dạng B được coi là đầu vào và độ khuếch tán và hướng đến của âm thanh được phân tích trong miền tần số.

Trong giai đoạn tổng hợp DirAC, âm thanh được chia thành hai luồng, luồng không khuếch tán và luồng khuếch tán. Luồng không khuếch tán được tái tạo dưới dạng các nguồn điểm bằng cách sử dụng quét biên độ, có thể được thực hiện bằng cách sử dụng quét biên độ cơ sở vectơ (vector base amplitude panning - VBAP) [2]. Luồng khuếch tán chịu trách nhiệm về cảm giác bao phủ và được tạo ra bằng cách truyền đến các loa phóng thanh các tín hiệu được tương quan qua lại.

Các tham số DirAC, còn được gọi là siêu dữ liệu không gian hoặc siêu dữ liệu DirAC sau đây, gồm các tập hợp độ khuếch tán và hướng. Hướng có thể được biểu diễn theo các tọa độ hình cầu bởi hai góc, góc phương vị và góc nâng, trong khi độ khuếch tán là hệ số vô hướng nằm trong khoảng giữa 0 và 1.

Fig.10a thể hiện giàn lọc 130 nhận tín hiệu đầu vào định dạng B. Phép phân tích năng lượng 132 và phân tích cường độ 134 được thực hiện. Phép tính trung bình theo thời gian cho các kết quả năng lượng được biểu thị ở 136 và phép tính trung bình theo thời gian cho các kết quả cường độ được biểu thị ở 138 được thực hiện, và từ dữ liệu trung bình, các giá trị độ khuếch tán cho các ngăn thời gian/tần số riêng được tính như biểu thị ở 110. Các giá trị hướng cho các ngăn thời gian/tần số được cung cấp bởi độ phân giải thời gian hoặc tần số của giàn lọc 130 được tính theo khối 120.

Trong sự tổng hợp DirAC được minh họa trên Fig.10b, một lần nữa giàn lọc phân tích 431 được sử dụng. Khối xử lý micrô ảo 421 được áp dụng, trong đó các micrô ảo ví dụ tương ứng với các vị trí loa phóng thanh ví dụ của thiết lập loa phóng thanh 5.1. Siêu dữ liệu độ khuếch tán được xử lý bởi các khối xử lý tương ứng 422 cho độ khuếch tán và bởi bảng độ khuếch đại VBAP (vectơ dựa trên việc quét biên độ) được biểu thị tại khối 423. Khối tính trung bình của loa phóng thanh 424 được tạo cầu

hình để thực hiện việc tính trung bình độ khuếch đại và khối chuẩn hóa tương ứng 425 được áp dụng để có các mức âm lượng được xác định tương ứng trong các tín hiệu loa phóng thanh cuối cùng. Việc bù micro được thực hiện trong khối 426.

Các tín hiệu thu được được sử dụng để tạo ra luồng khuếch tán 427 bao gồm giai đoạn giải tương quan trên một mặt và ngoài ra, luồng không khuếch tán 428 cũng được tạo ra. Cả hai luồng được thêm vào trong bộ cộng 429 cho băng con tương ứng và trong khối 431, việc bổ sung với các băng con khác, tức là chuyển đổi tần số thành thời gian được thực hiện. Do đó, khối 431 cũng có thể được coi là giàn lọc tổng hợp. Các thao tác xử lý tương tự được thực hiện cho các kênh khác từ thiết lập loa phóng thanh nhất định, trong đó, đối với kênh khác, việc thiết lập của micro ảo trong khối 421 sẽ khác.

Trong giai đoạn phân tích DirAC, micro trùng khớp bậc 1 ở định dạng B được coi là đầu vào và độ khuếch tán và hướng đến của âm thanh được phân tích trong miền tần số.

Trong giai đoạn tổng hợp DirAC, âm thanh được chia thành hai luồng, luồng không khuếch tán và luồng khuếch tán. Luồng không khuếch tán được tái tạo dưới dạng các nguồn điểm bằng cách sử dụng quét biên độ, có thể được thực hiện bằng cách sử dụng quét biên độ cơ sở vectơ (vector base amplitude panning - VBAP) [2]. Luồng khuếch tán chịu trách nhiệm về cảm giác bao phủ và được tạo ra bằng cách truyền đến các loa phóng thanh các tín hiệu được tương quan qua lại.

Các tham số DirAC, còn được gọi là siêu dữ liệu không gian hoặc siêu dữ liệu DirAC sau đây, gồm các tập hợp độ khuếch tán và hướng. Hướng có thể được biểu diễn theo các tọa độ hình cầu bởi hai góc, góc phương vị và góc nâng, trong khi độ khuếch tán là hệ số vô hướng nằm trong khoảng giữa 0 và 1.

Nếu STFT được coi là biến đổi thời gian-tần số với độ phân giải thời gian là 20ms mà thường được đề xuất trong một số bài báo và với sự chồng lấp 50% giữa các cửa sổ phân tích liên tiếp, phép phân tích DirAC sẽ tạo ra 288000 giá trị mỗi giây đối với đầu vào được lấy mẫu ở 48kHz, tương ứng nếu các góc được lượng tử hóa trên 8 bit với tổng tốc độ bit khoảng 2,3Mbit/s. Lượng dữ liệu không phù hợp để đạt được

phép mã hóa âm thanh không gian tốc độ bit thấp, và do đó cần có sơ đồ mã hóa hiệu quả của siêu dữ liệu DirAC.

Các công việc trước đây liên quan đến việc giảm siêu dữ liệu chủ yếu tập trung vào các trường hợp viển thông, trong đó khả năng của DirAC bị giảm đáng kể vì cho phép tốc độ dữ liệu tối thiểu của các tham số [4]. Thật vậy, nó được đề xuất để giới hạn việc phân tích có hướng theo góc phương vị trong mặt phẳng ngang để chỉ tái tạo cảnh âm thanh 2D. Hơn nữa, độ khuếch tán và góc phương vị chỉ được truyền tối đa 7kHz, hạn chế việc truyền thông đến tiếng nói dải rộng. Cuối cùng, độ khuếch tán được lượng tử hóa thô trên một hoặc hai bit, đôi khi chỉ bật hoặc tắt luồng khuếch tán trong giai đoạn tổng hợp, không đủ chung khi xem xét nhiều nguồn âm thanh và nhiều tiếng nói đơn lẻ hơn trên nhiều âm nền. Trong tài liệu [4], góc phương vị đã được lượng tử hóa trên 3 bit và người ta cho rằng nguồn, trong trường hợp đó là loa phóng thanh, có vị trí rất tĩnh. Do đó, các tham số chỉ được truyền với tần số cập nhật 50ms. Dựa trên các giả định mạnh mẽ này, nhu cầu về bit có thể giảm xuống còn khoảng 3kbit/s.

Tài liệu tham khảo

[1] V. Pulkki, M-V. Laitinen, J. Vilkamo, J. Ahonen, T. Lokki, and T. Pihlajamäki, “Directional audio coding - perception-based reproduction of spatial sound”, International Workshop on the Principles and Application on Spatial Hearing, Nov. 2009, Zao; Miyagi, Japan.

[2] V. Pulkki, “Virtual source positioning using vector base amplitude panning”, J. Audio Eng. Soc., 45(6):456-466, June 1997.

[3] J. Ahonen and V. Pulkki, “Diffuseness estimation using temporal variation of intensity vectors”, in Workshop on Applications of Signal Processing to Audio and Acoustics WASPAA, Mohonk Mountain House, New Paltz, 2009.

[4] T. Hirvonen, J. Ahonen, and V. Pulkki, “Perceptual compression methods for metadata in Directional Audio Coding applied to audiovisual teleconference”, AES 126th Convention, 2009, May 7–10, Munich, Germany.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất phương thức mã hóa âm thanh không gian được cải thiện.

Mục đích này đạt được bằng thiết bị mã hóa các tham số mã hóa âm thanh có hướng theo điểm 1, phương pháp mã hóa các tham số mã hóa âm thanh có hướng theo điểm 17, bộ giải mã để giải mã tín hiệu âm thanh được mã hóa theo điểm 18, phương pháp giải mã theo điểm 33 hoặc chương trình máy tính theo điểm 34.

Theo một khía cạnh, sáng chế dựa trên phát hiện rằng chất lượng được nâng cao trên một mặt và đồng thời, tốc độ bit giảm để mã hóa các tham số mã hóa âm thanh không gian trên mặt khác khi các tham số độ khuếch tán trên một mặt và các tham số hướng mặt khác được cung cấp với các độ phân giải khác nhau và các tham số khác nhau với các độ phân giải khác nhau được lượng tử hóa và mã hóa để thu được các tham số mã hóa âm thanh có hướng được mã hóa.

Theo phương án, độ phân giải thời gian hoặc tần số cho các tham số độ khuếch tán thấp hơn độ phân giải thời gian hoặc tần số của các tham số có hướng. Theo phương án khác, việc nhóm không chỉ theo tần số mà còn theo thời gian được thực hiện. Các tham số mã hóa âm thanh khuếch tán/có hướng ban đầu được tính toán với độ phân giải cao, ví dụ, đối với các ngăn thời gian/tần số độ phân giải cao, và việc nhóm và tốt hơn là việc nhóm với việc tính trung bình được thực hiện để tính toán tham số độ khuếch tán thu được với độ phân giải thời gian hoặc tần số thấp và để tính toán tham số có hướng thu được với độ phân giải thời gian hoặc tần số trung bình, nghĩa là với độ phân giải thời gian hoặc tần số nằm giữa độ phân giải thời gian hoặc tần số cho tham số độ khuếch tán và độ phân giải cao ban đầu, trong đó các tham số thô ban đầu đã được tính toán.

Theo các phương án, các độ phân giải thời gian thứ nhất và thứ hai là khác nhau và các độ phân giải tần số thứ nhất và thứ hai là giống nhau hoặc ngược lại, tức là các độ phân giải tần số thứ nhất và thứ hai là khác nhau nhưng các độ phân giải thời gian thứ nhất và lần thứ hai là giống nhau. Theo phương án khác, cả độ phân giải thời gian thứ nhất và thứ hai khác nhau và độ phân giải tần số thứ nhất và thứ hai cũng khác nhau. Do đó, độ phân giải thời gian hoặc tần số thứ nhất cũng có thể được coi là độ

phân giải thời gian-tần số thứ nhất và độ phân giải thời gian hoặc tần số thứ hai cũng có thể được coi là độ phân giải thời gian-tần số thứ hai.

Theo phương án khác, việc nhóm các tham số độ khuếch tán được thực hiện bằng phép cộng có trọng số, trong đó các hệ số trọng số của phép cộng có trọng số được xác định dựa trên công suất của tín hiệu âm thanh sao cho các ngăn thời gian/tần số có công suất cao hơn hoặc nói chung là số đo liên quan đến biên độ cao hơn cho tín hiệu âm thanh có ảnh hưởng cao hơn đến kết quả tham số độ khuếch tán cho ngăn thời gian/tần số, trong đó tín hiệu cần phân tích có công suất thấp hơn hoặc phép đo liên quan đến năng lượng thấp hơn.

Cũng ưu tiên thực hiện phép tính trung bình có trọng số gấp đôi để tính các tham số có hướng được nhóm. Phép tính trung bình có trọng số gấp đôi này được thực hiện theo cách các tham số có hướng từ các ngăn thời gian/tần số có ảnh hưởng cao hơn đến kết quả cuối cùng, khi công suất của tín hiệu gốc tương đối cao trong ngăn thời gian/tần số này. Đồng thời, giá trị độ khuếch tán của ngăn tương ứng cũng được tính đến để cuối cùng, tham số có hướng từ ngăn thời gian/tần số liên quan đến độ khuếch tán cao có tác động thấp hơn đến kết quả cuối cùng so với tham số có hướng có độ khuếch tán thấp, khi công suất giống nhau ở cả hai ngăn thời gian/tần số.

Nên thực hiện việc xử lý các tham số trong các khung, trong đó mỗi khung được tổ chức theo một số dải nhất định, trong đó mỗi dải bao gồm ít nhất hai ngăn tần số gốc trong đó các tham số đã được tính toán. Bảng thông của các dải, tức là, số lượng ngăn tần số gốc tăng theo số dải tăng sao cho các dải tần cao hơn rộng hơn các dải tần thấp hơn. Người ta đã phát hiện ra rằng, trong các phương án được ưu tiên, số tham số độ khuếch tán trên mỗi dải và khung bằng một trong khi số tham số có hướng trên mỗi khung và dải là hai hoặc thậm chí lớn hơn hai, chẳng hạn như bốn. Các tác giả sáng chế đã phát hiện ra rằng cùng một độ phân giải tần số, nhưng độ phân giải thời gian khác nhau, đối với tham số độ khuếch tán và tham số có hướng là hữu ích, tức là, số dải cho các tham số độ khuếch tán và các tham số có hướng trong khung bằng nhau. Các tham số được nhóm này sau đó được lượng tử hóa và mã hóa bởi bộ lượng hóa và bộ xử lý bộ mã hóa.

Theo khía cạnh thứ hai của sáng chế, mục đích đề xuất phương thức xử lý được cải thiện cho các tham số mã hóa âm thanh không gian đạt được bằng bộ lượng tử hóa tham số để lượng tử hóa các tham số độ khuếch tán và các tham số hướng và bộ mã hóa tham số được kết nối sau đó để mã hóa các tham số độ khuếch tán được lượng tử hóa và các tham số hướng được lượng tử hóa và giao diện đầu ra tương ứng để tạo sự biểu diễn tham số được mã hóa bao gồm thông tin về các tham số độ khuếch tán được mã hóa và các tham số hướng được mã hóa. Do đó, bằng cách lượng tử hóa và sau đó là mã hóa entropy, sẽ giảm được tốc độ dữ liệu đáng kể.

Các tham số độ khuếch tán và các tham số hướng được nhập vào bộ mã hóa có thể là các tham số độ khuếch tán/hướng có độ phân giải cao hoặc là các tham số mã hóa âm thanh có hướng độ phân giải thấp được nhóm hoặc không nhóm. Một tính năng của bộ lượng tử tham số được ưu tiên là độ chính xác lượng tử hóa cho tham số hướng lượng tử hóa được suy ra từ giá trị độ khuếch tán của tham số độ khuếch tán liên quan đến cùng vùng thời gian/tần số. Do đó, theo một tính năng của khía cạnh thứ hai, các tham số hướng liên quan đến các tham số độ khuếch tán có độ khuếch tán cao được lượng tử hóa ít chính xác hơn so với các tham số hướng liên quan đến các vùng thời gian/tần số có tham số độ khuếch tán biểu thị độ khuếch tán thấp.

Bản thân các tham số độ khuếch tán có thể được mã hóa entropy ở chế độ mã hóa thô, hoặc có thể được mã hóa ở chế độ mã hóa giá trị đơn khi các tham số độ khuếch tán cho các dải của khung có cùng giá trị trong toàn khung. Trong các phương án khác, các giá trị độ khuếch tán có thể được mã hóa trong quy trình chỉ hai giá trị liên tiếp.

Một tính năng khác của khía cạnh thứ hai là các tham số hướng được chuyển đổi thành sự biểu diễn góc phương vị/góc nâng. Trong tính năng này, giá trị góc nâng được sử dụng để xác định bảng chữ cái cho việc lượng tử hóa và mã hóa giá trị góc phương vị. Tốt hơn là, bảng chữ cái góc phương vị có số lượng các giá trị khác nhau lớn nhất khi góc nâng biểu thị góc không hoặc nói chung là góc xích đạo trên khối cầu đơn vị. Số lượng nhỏ nhất của các giá trị trong bảng chữ cái góc phương vị là khi góc nâng biểu thị cực bắc hoặc cực nam của khối cầu đơn vị. Do đó, giá trị bảng chữ cái giảm với giá trị tuyệt đối tăng của góc nâng được tính từ đường xích đạo.

Giá trị góc nâng này được lượng tử hóa với độ chính xác lượng tử hóa được xác định từ giá trị độ khuếch tán tương ứng, và bảng chữ cái lượng tử hóa trên một mặt và độ chính xác lượng tử hóa trên mặt khác xác định phép lượng tử hóa và mã hóa entropy điển hình của các giá trị góc phương vị tương ứng.

Do đó, quy trình xử lý hiệu quả và phù hợp với tham số được thực hiện nhằm loại bỏ càng nhiều sự không liên quan càng tốt, đồng thời, áp dụng độ phân giải cao hoặc độ chính xác cao cho các vùng có giá trị trong khi ở các vùng khác chẳng hạn như cực bắc hoặc cực nam của khối cầu đơn vị, độ chính xác không quá cao so với đường xích đạo của khối cầu đơn vị.

Phía bộ giải mã hoạt động theo khía cạnh thứ nhất thực hiện bất kỳ loại giải mã nào và thực hiện việc khử nhóm tương ứng với các tham số độ khuếch tán được mã hóa hoặc giải mã và các tham số hướng được mã hóa hoặc giải mã. Do đó, việc chuyển đổi độ phân giải tham số được thực hiện để nâng cao độ phân giải từ tham số mã hóa âm thanh có hướng được mã hóa hoặc giải mã thành độ phân giải được bộ kết xuất âm thanh sử dụng cuối cùng để thực hiện việc kết xuất cảnh âm thanh. Trong quá trình chuyển đổi độ phân giải này, việc chuyển đổi độ phân giải khác được thực hiện cho các tham số độ khuếch tán trên một mặt và cho các tham số hướng trên mặt khác.

Các tham số độ khuếch tán thường được mã hóa với độ phân giải thấp và do đó, một tham số độ khuếch tán phải được nhân hoặc sao chép một vài lần để thu được sự biểu diễn độ phân giải cao. Trên một mặt, tham số có hướng tương ứng phải được sao chép ít thường xuyên hơn hoặc được nhân ít thường xuyên hơn so với tham số độ khuếch tán, vì độ phân giải của các tham số hướng đã lớn hơn độ phân giải của các tham số độ khuếch tán trong tín hiệu âm thanh được mã hóa.

Theo phương án, các tham số mã hóa âm thanh có hướng được sao chép hoặc nhân được áp dụng khi chúng được xử lý như được làm mịn hoặc được lọc thông thấp để tránh các thành phần lạ gây ra bởi các tham số thay đổi mạnh theo tần số và/hoặc thời gian. Tuy nhiên, vì trong phương án được ưu tiên, việc áp dụng dữ liệu tham số được chuyển đổi độ phân giải được thực hiện trong miền phổ, việc chuyển đổi tần số-thời gian tương ứng của tín hiệu âm thanh được kết xuất từ miền tần số sang miền thời

gian thực hiện việc tính trung bình vốn có do ưu tiên quy trình chồng lấp và cộng được áp dụng là tính năng thường có trong các giàn lọc tổng hợp.

Ở phía bộ giải mã theo khía cạnh thứ hai, các quy trình cụ thể được thực hiện ở phía bộ mã hóa đối với phép mã hóa entropy trên một mặt và phép lượng tử hóa trên mặt khác được hoàn tác. Nên ưu tiên xác định độ chính xác khử lượng tử hóa ở phía bộ giải mã từ tham số độ khuếch tán được lượng tử hóa hoặc khử lượng tử hóa thông thường liên quan đến tham số hướng tương ứng.

Nên ưu tiên xác định bảng chữ cái cho tham số góc nâng từ giá trị độ khuếch tán tương ứng và độ chính xác khử lượng tử hóa liên quan của nó. Cũng nên ưu tiên cho khía cạnh thứ hai để thực hiện việc xác định bảng chữ cái khử lượng tử hóa cho tham số góc phương vị dựa trên giá trị của tham số góc nâng được lượng tử hóa hoặc tốt hơn là được khử lượng tử hóa.

Theo khía cạnh thứ hai, mặt khác, chế độ mã hóa thô trên một mặt hoặc chế độ mã hóa entropy trên mặt khác được thực hiện ở phía bộ mã hóa và chế độ dẫn đến số bit thấp hơn được chọn trong bộ mã hóa và được báo hiệu cho bộ giải mã thông qua một số thông tin phụ. Thông thường, chế độ mã hóa thô luôn được thực hiện cho các tham số hướng liên quan đến các giá trị độ khuếch tán cao trong khi chế độ mã hóa entropy được thử cho các tham số hướng liên quan đến các giá trị độ khuếch tán thấp hơn. Trong chế độ mã hóa entropy với mã hóa thô, các giá trị góc phương vị và góc nâng được gộp thành chỉ số hình cầu và chỉ số hình cầu sau đó được mã hóa bằng mã nhị phân hoặc mã thùng, và ở phía bộ giải mã mã hóa entropy này được hoàn tác tương ứng.

Trong chế độ mã hóa entropy với mô hình hóa, góc nâng trung bình và giá trị góc phương vị được tính cho khung, và các giá trị dư đối với các giá trị trung bình này thực sự được tính toán. Do đó, loại dự đoán được thực hiện và các giá trị dư dự đoán, tức là, khoảng cách cho góc nâng và góc phương vị được mã hóa entropy. Với mục đích này, nên thực hiện quy trình Golomb-Rice mở rộng dựa trên tham số Golomb-Rice được xác định ở phía bộ mã hóa và được mã hóa, ngoài các khoảng cách có dấu và các giá trị trung bình ưu tiên. Ở phía bộ giải mã ngay khi mã hóa entropy với mô hình hóa, tức là chế độ giải mã này, được báo hiệu và xác định bởi đánh giá thông tin

phụ trong bộ giải mã, việc giải mã với quy trình Golomb-Rice mở rộng được thực hiện bằng cách sử dụng các trung bình được mã hóa, các khoảng cách có dấu ưu tiên được mã hóa và các tham số Golomb-Rice tương ứng cho góc nâng và góc phương vị.

Mô tả vắn tắt các hình vẽ

Các phương án được ưu tiên của sáng chế sau đây sẽ được thảo luận liên quan đến các hình vẽ kèm theo, trong đó:

Fig.1a minh họa phương án được ưu tiên của phía bộ mã hóa của khía cạnh thứ nhất hoặc khía cạnh thứ hai;

Fig.1b minh họa phương án được ưu tiên của phía bộ giải mã của khía cạnh thứ nhất hoặc khía cạnh thứ hai;

Fig.2a minh họa phương án được ưu tiên của một thiết bị để mã hóa theo khía cạnh thứ nhất;

Fig.2b minh họa triển khai được ưu tiên của bộ tính tham số của Fig.2a;

Fig.2c minh họa triển khai khác cho việc tính tham số độ khuếch tán;

Fig.2d minh họa triển khai được ưu tiên khác của bộ tính tham số của 100 của Fig.2a;

Fig.3a minh họa sự biểu diễn thời gian/tần số thu được từ giàn lọc phân tích 130 của Hình Fig.1a hoặc 430 của Fig.1b với độ phân giải thời gian hoặc tần số cao;

Fig.3b minh họa triển khai của việc nhóm độ khuếch tán với độ phân giải thời gian hoặc tần số thấp, và đặc biệt là độ phân giải thời gian thấp cụ thể của một tham số độ khuếch tán duy nhất trên mỗi khung;

Fig.3c minh họa sự minh họa được ưu tiên của độ phân giải trung bình cho các tham số hướng có năm dải một mặt và bốn vùng thời gian trên mặt khác dẫn đến 20 vùng thời gian/tần số;

Fig.3d minh họa dòng bit đầu ra với các tham số độ khuếch tán được mã hóa và các tham số hướng được mã hóa;

Fig.4a minh họa thiết bị mã hóa các tham số mã hóa âm thanh có hướng theo khía cạnh thứ hai;

Fig.4b minh họa triển khai được ưu tiên của bộ lượng tử hóa tham số và bộ mã hóa tham số để tính toán các tham số độ khuếch tán được mã hóa;

Fig.4c minh họa triển khai được ưu tiên của bộ mã hóa theo Fig.4a liên quan đến sự kết hợp của các thành phần khác nhau;

Fig.4d minh họa độ bao phủ giả đều của khối cầu đơn vị được áp dụng cho mục đích lượng tử hóa trong phương án được ưu tiên;

Fig.5a minh họa tổng quan về hoạt động của bộ mã hóa tham số của Fig.4a hoạt động ở các chế độ mã hóa khác nhau;

Fig.5b minh họa quy trình tiền xử lý của các chỉ số hướng cho cả hai chế độ của Fig.5a;

Fig.5c minh họa chế độ mã hóa thứ nhất trong phương án được ưu tiên;

Fig.5d minh họa phương án được ưu tiên của chế độ mã hóa thứ hai;

Fig.5e minh họa triển khai được ưu tiên của mã hóa entropy của các khoảng cách có dấu và các giá trị trung bình tương ứng bằng cách sử dụng quy trình mã hóa GR;

Fig.5f minh họa triển khai được ưu tiên để xác định tham số Golomb-Rice tối ưu;

Fig.5g minh họa việc triển khai quy trình Golomb-Rice mở rộng để mã hóa các khoảng cách có dấu được sắp xếp lại như được biểu thị trong khối 279 của Fig.5e;

Fig.6a minh họa việc triển khai bộ lượng tử hóa tham số của Fig.4a;

Fig.6b minh họa triển khai được ưu tiên của các chức năng cho bộ khử lượng tử hóa tham số cũng được sử dụng trong các khía cạnh nhất định trong việc triển khai phía bộ mã hóa;

Fig.6c minh họa tổng quan về việc triển khai quy trình mã hóa hướng thô;

Fig.6d minh họa triển khai của việc tính toán và lượng tử hóa và khử lượng tử hóa cho hướng trung bình cho góc phương vị và góc nâng;

Fig.6e minh họa hình chiếu của dữ liệu góc nâng và góc phương vị trung bình;

Fig.6f minh họa việc tính toán các khoảng cách cho góc nâng và góc phương vị;

Fig.6g minh họa tổng quan về mã hóa hướng trung bình trong chế độ mã hóa entropy với mô hình hóa;

Fig.7a minh họa bộ giải mã để giải mã tín hiệu âm thanh được mã hóa theo khía cạnh thứ nhất,

Fig.7b minh họa triển khai được ưu tiên của bộ chuyển đổi độ phân giải tham số của Fig.7a và kết xuất âm thanh tiếp theo;

Fig.8a minh họa bộ giải mã để giải mã tín hiệu âm thanh được mã hóa theo khía cạnh thứ hai,

Fig.8b minh họa sự biểu diễn dòng bit dưới dạng sơ đồ cho các tham số độ khuếch tán được mã hóa theo phương án;

Fig.8c minh họa việc triển khai dòng bit khi chế độ mã hóa thô đã được chọn;

Fig.8d minh họa dòng bit dưới dạng sơ đồ khi chế độ mã hóa khác, tức là chế độ mã hóa entropy với mô hình hóa đã được chọn;

Fig.8e minh họa triển khai được ưu tiên của bộ giải mã tham số và bộ khử lượng tử hóa tham số, trong đó độ chính xác khử lượng tử hóa được xác định dựa trên độ khuếch tán cho vùng thời gian/tần số;

Fig.8f minh họa triển khai được ưu tiên của bộ giải mã tham số và bộ khử lượng tử hóa tham số, trong đó bảng chữ cái góc nâng được xác định từ độ chính xác khử lượng tử hóa và bảng chữ cái góc phương vị được xác định dựa trên độ chính xác khử lượng tử hóa và dữ liệu góc nâng cho vùng thời gian/tần số;

Fig.8g minh họa tổng quan về bộ giải mã tham số của Fig.8a minh họa hai chế độ giải mã khác nhau;

Fig.9a minh họa thao tác giải mã khi chế độ mã hóa thô được kích hoạt;

Fig.9b minh họa phép giải mã của hướng trung bình, khi chế độ giải mã entropy với mô hình hóa đang hoạt động;

Fig.9c minh họa việc khôi phục các góc nâng và các góc phương vị khi chế độ giải mã với mô hình hóa đang hoạt động, và sự khử lượng tử hóa tiếp theo;

Fig.10a minh họa bộ phân tích DirAC đã biết; và

Fig.10a minh họa bộ tổng hợp DirAC đã biết.

Mô tả chi tiết sáng chế

Sáng chế tổng quát hóa việc nén siêu dữ liệu DirAC cho bất kỳ trường hợp nào. Sáng chế được áp dụng trong hệ thống mã hóa không gian được minh họa trên Fig.1a và Fig.1b, trong đó bộ mã hóa và giải mã âm thanh không gian dựa trên DirAC được mô tả.

Bộ mã hóa thường phân tích cảnh âm thanh không gian ở định dạng B. Ngoài ra, phép phân tích DirAC có thể được điều chỉnh để phân tích các định dạng âm thanh khác nhau như các đối tượng âm thanh hoặc các tín hiệu đa kênh hoặc sự kết hợp của bất kỳ định dạng âm thanh không gian nào. Phép phân tích DirAC trích xuất sự biểu diễn theo tham số từ cảnh âm thanh đầu vào. Hướng đến (DOA) và độ khuếch tán được đo trên bộ phận thời gian-tần số tạo thành các tham số. Sau phép phân tích DirAC là bộ mã hóa siêu dữ liệu không gian mà lượng tử hóa và mã hóa các tham số DirAC để thu được sự biểu diễn theo tham số tốc độ bit thấp. Các mô đun sau là đối tượng của sáng chế này.

Cùng với các tham số, tín hiệu trộn giảm được suy ra từ các nguồn khác nhau hoặc các tín hiệu đầu vào âm thanh được mã hóa để truyền bởi bộ mã hóa lõi âm thanh thông thường. Trong phương án được ưu tiên, bộ mã hóa âm thanh EVS được ưu tiên để mã hóa tín hiệu trộn giảm, nhưng sáng chế không giới hạn ở bộ mã hóa lõi này và có thể được áp dụng cho bất kỳ bộ mã hóa lõi âm thanh nào. Tín hiệu trộn giảm bao gồm các kênh khác nhau, được gọi là các kênh truyền tải: tín hiệu có thể là ví dụ bốn tín hiệu hệ số tạo thành tín hiệu định dạng B, trộn giảm cặp âm thanh nổi hoặc đơn âm tùy theo tốc độ bit được nhắm tới. Các tham số không gian được mã hóa và dòng bit âm thanh được mã hóa được ghép kênh trước khi được truyền qua kênh truyền thông.

Trong bộ giải mã, các kênh truyền tải được giải mã bởi bộ giải mã lõi, trong khi siêu dữ liệu DirAC được giải mã trước tiên trước khi được chuyển với các kênh truyền tải được giải mã sang sự tổng hợp DirAC. Sự tổng hợp DirAC sử dụng siêu dữ liệu được giải mã để kiểm soát việc tái tạo luồng âm thanh thẳng và hỗn hợp của nó với

luồng âm thanh khuếch tán. Trường âm thanh tái tạo có thể được tái tạo trên bố trí loa phóng thanh tùy ý hoặc có thể được tạo ở định dạng Ambisonics (HOA/FOA) với thứ tự tùy ý.

Bộ mã hóa âm thanh để mã hóa tín hiệu âm thanh, chẳng hạn như tín hiệu đầu vào định dạng B được minh họa trên Fig.1a. Bộ mã hóa âm thanh bao gồm bộ phân tích DirAC 100. Bộ phân tích DirAC 100 có thể bao gồm giàn lọc phân tích 130, bộ ước lượng độ khuếch tán 110 và công cụ ước lượng hướng 120. Dữ liệu về độ khuếch tán và dữ liệu hướng được xuất ra bộ mã hóa siêu dữ liệu không gian 200 mà cuối cùng xuất ra siêu dữ liệu được mã hóa trên dòng 250. Tín hiệu định dạng B cũng có thể được chuyển đến bộ tạo chùm/bộ chọn tín hiệu 140 tạo ra, từ tín hiệu đầu vào, tín hiệu âm thanh truyền tải đơn âm hoặc âm thanh nổi mà sau đó được mã hóa trong bộ mã hóa âm thanh 150 mà tốt hơn là bộ mã hóa EVS (enhanced voice services: dịch vụ thoại nâng cao). Tín hiệu âm thanh được mã hóa được xuất ra tại 170. Các tham số mã hóa được mã hóa được biểu thị ở 250 được đưa vào bộ giải mã siêu dữ liệu không gian 300. Tín hiệu âm thanh được mã hóa 170 được đưa vào bộ giải mã âm thanh 340 được triển khai, theo phương án được ưu tiên và phù hợp với triển khai phía bộ mã hóa, dưới dạng bộ giải mã EVS.

Tín hiệu truyền tải được giải mã được đưa vào bộ tổng hợp DirAC 400 cùng với các tham số mã hóa âm thanh có hướng được giải mã. Theo phương án được minh họa trên Fig.1b, bộ tổng hợp DirAC bao gồm bộ tổng hợp đầu ra 420, giàn lọc phân tích 430 và giàn lọc tổng hợp 440. Ở đầu ra của giàn lọc tổng hợp 400, tín hiệu đa kênh được giải mã 450 thu được có thể được chuyển đến các loa phóng thanh hoặc thay vào đó có thể là tín hiệu âm thanh ở bất kỳ định dạng nào khác, chẳng hạn như định dạng Ambisonics bậc nhất (first order Ambisonics - FOA) hoặc định dạng Ambisonics bậc cao (high order Ambisonics - HOA). Đương nhiên, bất kỳ dữ liệu tham số khác như dữ liệu MPS (MPEG surround: âm thanh vòm MPEG) hoặc dữ liệu SAOC (spatial audio object coding: mã hóa đối tượng âm thanh không gian) có thể được tạo cùng với kênh truyền tải là kênh đơn âm hoặc kênh âm thanh nổi.

Nói chung, bộ tổng hợp đầu ra hoạt động bằng cách tính toán, cho mỗi ngăn thời gian-tần số được xác định bởi giàn lọc phân tích 430, một mặt là tín hiệu âm thanh

trực tiếp và mặt khác là tín hiệu âm thanh khuếch tán. Tín hiệu âm thanh thẳng được tính toán dựa trên các tham số hướng và mối quan hệ giữa tín hiệu âm thanh thẳng và tín hiệu âm thanh khuếch tán trong tín hiệu âm thanh cuối cùng cho ngắn thời gian/tần số này, được xác định dựa trên tham số độ khuếch tán để ngắn thời gian/tần số có tham số độ khuếch tán cao dẫn đến tín hiệu đầu ra có lượng tín hiệu khuếch tán cao và có lượng tín hiệu thẳng thấp, trong khi ngắn thời gian/tần số có độ khuếch tán thấp dẫn đến tín hiệu đầu ra có lượng tín hiệu thẳng cao và lượng tín hiệu khuếch tán thấp.

Fig.2a minh họa thiết bị mã hóa các tham số mã hóa âm thanh có hướng bao gồm các tham số độ khuếch tán và các tham số hướng theo khía cạnh thứ nhất. Thiết bị bao gồm bộ tính tham số 100 để tính toán các tham số độ khuếch tán với độ phân giải thời gian hoặc tần số thứ nhất và để tính toán các tham số hướng với độ phân giải thời gian hoặc tần số thứ hai. Thiết bị bao gồm bộ xử lý bộ lượng tử hóa và bộ mã hóa 200 để tạo sự biểu diễn được lượng tử hóa và được mã hóa của các tham số độ khuếch tán và các tham số hướng được minh họa ở 250. Bộ tính tham số 100 có thể chứa các phần tử 110, 120, 130 của Fig.1a, trong đó các tham số khác nhau đã được tính toán trong độ phân giải thời gian hoặc tần số thứ nhất hoặc thứ hai.

Ngoài ra, triển khai được ưu tiên được minh họa trên Fig.2b. Ở đây, bộ tính tham số và đặc biệt là các khối 110, 120 trên Fig.1a được tạo cấu hình như minh họa trong mục 130 của Fig.2b, tức là, chúng tính toán các tham số với độ phân giải thời gian hoặc tần số cao điển hình thứ ba hoặc thứ tư. Thao tác nhóm được thực hiện. Để tính toán các tham số độ khuếch tán, việc nhóm và tính trung bình được thực hiện như minh họa trong khối 141 để thu được sự biểu diễn tham số độ khuếch tán với độ phân giải thời gian hoặc tần số thứ nhất, và để tính toán các tham số hướng, việc nhóm (và tính trung bình) được thực hiện trong khối 142 để thu được sự biểu diễn tham số hướng trong độ phân giải thời gian hoặc tần số thứ hai.

Các tham số độ khuếch tán và tham số hướng được tính toán sao cho độ phân giải thời gian hoặc tần số thứ hai khác với độ phân giải thời gian hoặc tần số thứ nhất và độ phân giải thời gian thứ nhất thấp hơn độ phân giải thời gian thứ hai hoặc độ phân giải tần số thứ hai lớn hơn độ phân giải tần số thứ nhất hoặc, một lần nữa, thay vào đó,

độ phân giải thời gian thứ nhất thấp hơn độ phân giải thời gian thứ hai và độ phân giải tần số thứ nhất bằng độ phân giải tần số thứ hai.

Thông thường, các tham số độ khuếch tán và tham số hướng được tính cho tập hợp các dải tần, trong đó dải có tần số trung tâm thấp hơn thì hẹp hơn so với dải có tần số trung tâm cao hơn. Như đã thảo luận về Fig.2b, bộ tính tham số 100 được tạo cấu hình để thu được các tham số độ khuếch tán ban đầu có độ phân giải thời gian hoặc tần số thứ ba và bộ tính tham số 100 cũng được tạo cấu hình để thu được các tham số hướng ban đầu có độ phân giải thời gian hoặc tần số thứ tư trong đó, thông thường, độ phân giải thời gian hoặc tần số thứ ba và thứ tư bằng nhau.

Sau đó, bộ tính tham số được tạo cấu hình để nhóm và tính trung bình các tham số độ khuếch tán ban đầu để độ phân giải thời gian hoặc tần số thứ ba cao hơn độ phân giải thời gian hoặc tần số thứ nhất, tức là việc giảm độ phân giải được thực hiện. Bộ tính tham số cũng được tạo cấu hình để nhóm và tính trung bình các tham số hướng ban đầu để độ phân giải thời gian hoặc tần số thứ tư cao hơn độ phân giải thời gian hoặc tần số thứ hai, tức là, việc giảm độ phân giải được thực hiện. Tốt hơn là, độ phân giải thời gian hoặc tần số thứ ba là độ phân giải thời gian không đổi sao cho mỗi tham số độ khuếch tán ban đầu liên quan đến khe thời gian hoặc ngăn tần số có cùng kích thước. Độ phân giải thời gian hoặc tần số thứ tư cũng là độ phân giải tần số không đổi sao cho mỗi tham số hướng ban đầu liên quan đến khe thời gian hoặc ngăn tần số có cùng kích thước.

Bộ tính tham số được tạo cấu hình để tính trung bình trên nhiều tham số độ khuếch tán thứ nhất liên quan đến nhiều khe thời gian thứ nhất. Bộ tính tham số 100 cũng được tạo cấu hình để tính trung bình trên nhiều tham số độ khuếch tán thứ hai liên quan đến nhiều ngăn tần số thứ hai, và bộ tính tham số cũng được tạo cấu hình để tính trung bình trên nhiều tham số hướng thứ ba liên quan đến nhiều khe thời gian thứ ba hoặc bộ tính tham số cũng được tạo cấu hình để tính trung bình trên nhiều tham số hướng thứ tư liên quan đến nhiều ngăn tần số thứ tư.

Như sẽ được thảo luận liên quan đến Fig.2c và Fig.2d, bộ tính tham số 100 được tạo cấu hình để thực hiện phép tính trung bình có trọng số trong đó tham số độ khuếch tán hoặc tham số hướng được suy ra từ phần tín hiệu đầu vào có số đo liên

quan đến biên độ cao hơn được gán trọng số bằng cách sử dụng hệ số trọng số cao hơn so với tham số độ khuếch tán hoặc tham số hướng được suy ra từ phân tín hiệu đầu vào có số đo liên quan đến biên độ thấp hơn. Bộ tính tham số 100 được tạo cấu hình để tính toán 143 và số đo liên quan đến biên độ cho mỗi ngăn trong độ phân giải thời gian hoặc tần số thứ ba hoặc thứ tư như được minh họa trên Fig.2c, mục 143. Trong khối 144, các hệ số trọng số cho mỗi ngăn được tính toán, và trong khối 145, việc nhóm và tính trung bình được thực hiện bằng cách sử dụng sự kết hợp có trọng số chẳng hạn như phép cộng có trọng số trong đó các tham số độ khuếch tán cho các ngăn riêng lẻ được đưa vào khối 145. Ở đầu ra của khối 145, các tham số độ khuếch tán với độ phân giải thời gian hoặc tần số thứ nhất thu được mà sau đó có thể được chuẩn hóa trong khối 146 nhưng quy trình này chỉ là tùy chọn.

Fig.2d minh họa việc tính toán các tham số hướng với độ phân giải thứ hai. Trong khối 146, số đo liên quan đến biên độ được tính cho mỗi ngăn ở độ phân giải thứ ba hoặc thứ tư tương tự như mục 143 của Fig.2c. Trong khối 147, các hệ số trọng số được tính cho mỗi ngăn, nhưng không chỉ phụ thuộc vào số đo liên quan đến biên độ thu được từ khối 147 mà còn sử dụng tham số độ khuếch tán tương ứng trên mỗi ngăn như được minh họa trên Fig.2d. Do đó, đối với cùng một số đo liên quan đến biên độ, hệ số cao hơn thường được tính cho độ khuếch tán thấp hơn. Trong khối 148, việc nhóm và tính trung bình được thực hiện bằng cách sử dụng sự kết hợp có trọng số chẳng hạn như phép cộng và kết quả có thể được chuẩn hóa như được minh họa trong khối 146 tùy chọn. Do đó, ở đầu ra của khối 146, tham số hướng thu được dưới dạng vectơ đơn vị tương ứng với vùng hai chiều hoặc ba chiều chẳng hạn như vectơ Đề-các có thể dễ dàng chuyển đổi thành dạng cực có giá trị góc phương vị và giá trị góc nâng.

Fig.3a minh họa màn hình thời gian/tần số thu được từ phép phân tích giàn lọc 430 của Fig.1a và Fig.1b hoặc như được sự tổng hợp giàn lọc 440 của Fig.1b áp dụng. Theo phương án, toàn bộ dải tần được phân tách thành 60 dải tần và khung có thêm 16 khe thời gian. Độ phân giải thời gian hoặc tần số cao này tốt hơn là độ phân giải thời gian hoặc tần số cao thứ ba hoặc thứ tư. Do đó, bắt đầu từ 60 dải tần và 16 khe thời gian, thu được 960 ô hoặc ngăn thời gian/tần số trên mỗi khung.

Fig.3b minh họa việc giảm độ phân giải được thực hiện bởi bộ tính tham số và cụ thể là bởi khối 141 của Fig.2b để thu được sự biểu diễn độ phân giải thời gian hoặc tần số thứ nhất cho các giá trị độ khuếch tán. Theo phương án này, toàn bộ băng thông tần số được tách thành năm dải nhóm và chỉ có một khe thời gian duy nhất. Do đó, đối với một khung, cuối cùng, chỉ có năm tham số độ khuếch tán trên mỗi khung sau đó được lượng tử hóa và mã hóa thêm.

Fig.3c minh họa quy trình tương ứng được thực hiện bởi khối 142 của Fig.2b. Các tham số hướng có độ phân giải cao từ Fig.3a trong đó một tham số hướng được tính cho mỗi ngăn được nhóm lại và lấy trung bình thành sự biểu diễn độ phân giải trung bình trên Fig.3c, trong đó, mỗi khung có năm dải tần, nhưng ngược lại với Fig.3a, bây giờ có bốn khe thời gian. Do đó, cuối cùng, một khung nhận được 20 tham số hướng, tức là, 20 ngăn được nhóm trên mỗi khung cho các tham số hướng và chỉ có năm ngăn được nhóm trên mỗi khung cho các tham số độ khuếch tán của Fig.3b. Theo phương án được ưu tiên, các cạnh của dải tần là riêng biệt ở các cạnh trên của chúng, sao cho

Khi so sánh Fig.3b và Fig.3c, cần lưu ý rằng tham số độ khuếch tán cho dải thứ nhất, tức là, tham số độ khuếch tán 1 tương ứng với bốn tham số hướng cho dải thứ nhất hoặc liên quan đến chúng. Như sẽ được phác thảo sau này, độ chính xác lượng tử hóa cho tất cả các tham số hướng trong dải thứ nhất được xác định bởi tham số độ khuếch tán cho dải thứ nhất hoặc, chẳng hạn, độ chính xác lượng tử hóa cho các tham số hướng cho dải thứ năm, nghĩa là cho bốn tham số hướng tương ứng bao trùm dải thứ năm và bốn khe thời gian trong dải thứ năm được xác định bởi tham số độ khuếch tán đơn cho dải thứ năm.

Do đó, trong phương án này, trong đó chỉ có một tham số độ khuếch tán duy nhất có trong mỗi dải, tất cả các tham số hướng trong một dải có cùng độ chính xác lượng tử hóa/khử lượng tử hóa. Như sẽ được phác thảo sau này, bảng chữ cái để lượng tử hóa và mã hóa tham số góc phương vị phụ thuộc vào giá trị của tham số góc nâng ban đầu/được lượng tử hóa/được khử lượng tử hóa. Do đó, mặc dù mỗi tham số hướng cho mỗi dải có cùng tham số lượng tử hóa/khử lượng tử hóa, mỗi tham số góc phương vị

cho mỗi ngăn được nhóm hoặc vùng thời gian/tần số của Fig.3c có thể có bảng chữ cái khác nhau để lượng tử hóa và mã hóa.

Dòng bit kết quả được tạo bởi bộ xử lý bộ lượng tử hóa và bộ mã hóa 200 được minh họa ở 250 trên Fig.2a được minh họa chi tiết hơn trên Fig.3d. Dòng bit có thể chứa chỉ thị độ phân giải 260 biểu thị độ phân giải thứ nhất và độ phân giải thứ hai. Tuy nhiên, khi độ phân giải thứ nhất và độ phân giải thứ hai được thiết lập cố định bởi bộ mã hóa và bộ giải mã, thì chỉ thị độ phân giải này là không cần thiết. Các mục 261, 262 minh họa các tham số độ khuếch tán được mã hóa cho các dải tương ứng. Vì Fig.3d chỉ minh họa năm dải, chỉ có năm tham số độ khuếch tán được đưa vào luồng dữ liệu được mã hóa. Các mục 363, 364 minh họa các tham số hướng được mã hóa. Đối với dải thứ nhất, có bốn tham số hướng được mã hóa, trong đó chỉ số thứ nhất của tham số hướng biểu thị dải và tham số thứ hai biểu thị khe thời gian. Tham số hướng cho dải thứ năm và khe thời gian thứ tư, tức là, đối với ngăn tần số phía trên bên phải trên Fig.3c được biểu thị là DIR54.

Sau đây, triển khai được ưu tiên hơn nữa sẽ được thảo luận chi tiết.

Phân tách thời gian-tần số

Trong DirAC cả phép phân tích và sự tổng hợp được thực hiện trong miền tần số. Phép phân tích và tổng hợp thời gian-tần số có thể được thực hiện bằng cách sử dụng các biến đổi khối khác nhau, như biến đổi Fourier ngắn hạn (short-term Fourier Transform - STFT) hoặc các giàn lọc, như giàn lọc gương cầu phương được biến điệu phức tạp (quadrature mirror filterbank - QMF). Theo phương án được ưu tiên, các tác giả sáng chế mong muốn chia sẻ khung giữa quá trình xử lý DirAC và bộ mã hóa lỗi. Do bộ mã hóa lỗi tốt hơn là dựa trên bộ mã hóa-giải mã 3GPP EVS, nên cần có khung 20ms. Ngoài ra, các tiêu chí quan trọng như độ phân giải thời gian và tần số và độ mạnh mẽ khử răng cưa có liên quan đến xử lý thời gian-tần số rất tích cực trong DirAC. Vì hệ thống được thiết kế cho các truyền thông, độ trễ thuật toán là một khía cạnh quan trọng khác.

Vì tất cả những lý do này, giàn lọc độ trễ thấp được biến điệu phức tạp (complex modulated low-delay filterbank - CLDFB) là lựa chọn ưu tiên. CLDFB có độ phân giải thời gian là 1,25ms và chia khung 20ms thành 16 khe thời gian. Độ phân

giải tần số là 400Hz, có nghĩa là tín hiệu đầu vào được phân tách thành $(fs/2)/400$ dải tần. Hoạt động của giàn lọc được mô tả ở dạng chung theo công thức sau:

$$X_{CR}(t, k) = \sqrt{\frac{80}{L_c}} \cdot \sum_{n=0}^{n=10L_c-1} -w_c(10L_c - n) \cdot s_{HP}(10L_c - n + t \cdot L_c) \cos \left[\frac{\pi}{L_c} \left(n + \frac{1}{2} + \frac{L_c}{2} \right) \left(k + \frac{1}{2} \right) \right]$$

$$X_{CI}(t, k) = \sqrt{\frac{80}{L_c}} \cdot \sum_{n=0}^{n=10L_c-1} -w_c(10L_c - n) \cdot s_{HP}(10L_c - n + t \cdot L_c) \sin \left[\frac{\pi}{L_c} \left(n + \frac{1}{2} + \frac{L_c}{2} \right) \left(k + \frac{1}{2} \right) \right],$$

trong đó X_{CR} và X_{CI} tương ứng là giá trị của băng con thực và ảo, t là chỉ số thời gian của băng con với $0 \leq t \leq 15$ và k là chỉ số của băng con với $0 \leq k \leq L_c - 1$. Nguyên mẫu phân tích w_c là bộ lọc thông thấp không đối xứng với độ dài thích ứng tùy thuộc vào s_{HP} . Độ dài w_c được cho bởi $L_{w_c} = 10 \cdot L_c$ nghĩa là bộ lọc kéo dài qua 10 khối liên tiếp để chuyển đổi.

Chẳng hạn, CLDFB sẽ phân tách tín hiệu được lấy mẫu ở 48kHz thành $60 \times 16 = 960$ ô thời gian-tần số trên mỗi khung. Độ trễ sau khi phân tích và tổng hợp có thể được điều chỉnh bằng cách chọn các bộ lọc nguyên mẫu khác nhau. Các tác giả sáng chế đã thấy rằng độ trễ 5ms (phân tích và tổng hợp) là một sự thỏa hiệp tốt giữa chất lượng được phân phối và độ trễ được cam kết. Đối với mỗi ô thời gian-tần số, độ khuếch tán và hướng được tính toán.

Ước lượng tham số DirAC

Trong mỗi dải tần, hướng đến của âm thanh cùng với độ khuếch tán của âm thanh được ước lượng. Từ phép phân tích thời gian-tần số của thành phần định dạng B đầu vào $w^i(n), x^i(n), y^i(n), z^i(n)$, các vector áp suất và vận tốc có thể được xác định là:

$$P^i(n, k) = W^i(n, k)$$

$$U^i(n, k) = X^i(n, k) \mathbf{e}_x + Y^i(n, k) \mathbf{e}_y + Z^i(n, k) \mathbf{e}_z$$

trong đó i là chỉ số của đầu vào, n và k là các chỉ số thời gian và tần số của ô tần số thời gian, và $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$ biểu diễn các vector đơn vị Đề-các. $P(n, k)$ và $U(n, k)$ là cần thiết để tính toán các tham số DirAC, cụ thể là DOA và độ khuếch tán thông qua việc tính vector cường độ:

$$I(n, k) = \frac{1}{2} \Re \{ P(n, k) \cdot \overline{U(n, k)} \},$$

trong đó $\overline{(\cdot)}$ biểu thị liên hợp phức. Độ khuếch tán của trường âm thanh kết hợp được cho bởi:

$$\psi(n, k) = 1 - \frac{\|E\{I(n, k)\}\|}{cE\{E(n, k)\}}$$

trong đó $E\{\cdot\}$ biểu thị toán tử tính trung bình theo thời gian, c là tốc độ âm thanh và $E(n, k)$ là năng lượng trường âm thanh được cho bởi:

$$E(n, k) = \frac{\rho_0}{4} \|U(n, k)\|^2 + \frac{1}{\rho_0 c^2} |P(n, k)|^2$$

Độ khuếch tán của trường âm thanh được định nghĩa là tỷ lệ giữa cường độ âm thanh và mật độ năng lượng, có các giá trị nằm trong khoảng từ 0 đến 1.

Hướng đến (DOA) được biểu thị bằng phương tiện của vector đơn vị $direction(n, k)$, được xác định là

$$direction(n, k) = -\frac{I(n, k)}{\|I(n, k)\|}$$

Hướng đến được xác định bằng phép phân tích năng lượng của đầu vào định dạng B và có thể được xác định là hướng ngược lại của vector cường độ. Hướng được xác định theo các tọa độ Đề-các nhưng có thể dễ dàng chuyển đổi theo các tọa độ hình cầu được xác định bởi bán kính thông nhất, góc phương vị và góc nâng.

Tóm lại, nếu các giá trị tham số được chuyển đổi trực tiếp thành các bit, đối với mỗi ô thời gian-tần số, 3 giá trị phải được mã hóa: góc phương vị, góc nâng, và độ khuếch tán. Siêu dữ liệu sau đó có trong ví dụ về CLDFB có 2880 giá trị mỗi khung, tức là 144000 giá trị mỗi giây. Lượng dữ liệu khổng lồ này cần phải được giảm mạnh để đạt được việc mã hóa tốc độ bit thấp.

Nhóm và tính trung bình siêu dữ liệu DirAC

Để giảm số lượng tham số, các tham số được tính toán trong mỗi ô thời gian-tần số được nhóm và tính trung bình trước tiên dọc theo các dải tham số tần số và qua một vài khe thời gian. Việc nhóm được tách rời giữa độ khuếch tán và hướng, đây là một khía cạnh quan trọng của sáng chế. Thật vậy, việc tách rời dựa trên thực tế là độ

khuyến tán giữ lại đặc tính dài hạn của trường âm thanh hơn là hướng, đó là tín hiệu không gian phản ứng mạnh hơn.

Các dải tham số tạo thành sự phân tách không đồng nhất và không chồng lấp của các dải tần theo số lần gần bằng số lần của thang đo băng thông hình chữ nhật tương đương (equivalent rectangular bandwidth - ERB). Theo mặc định, thang đo ERB 9 lần được áp dụng cho tổng số 5 dải tham số cho băng thông âm thanh là 16kHz.

Độ khuếch tán được tính là:

$$\text{diff}[g, b] = \frac{\sum_{n=\text{slot_diff}[g]}^{\text{slot_diff}[g+1]-1} \sum_{k=\text{band_diff}[b]}^{\text{band_diff}[b+1]-1} \text{diffuseness}(n, k) \cdot \text{power}(n, k)^\alpha}{\sum_{n=\text{slot_diff}[g]}^{\text{slot_diff}[g+1]-1} \sum_{k=\text{band_diff}[b]}^{\text{band_diff}[b+1]-1} \text{power}(n, k)^\alpha},$$

trong đó $\text{power}(n, k)^\alpha$ là năng lượng của tín hiệu đầu vào được đo trong ô thời gian-tần số của các chỉ số (t, k) và được tăng lên lũy thừa α , và $\text{diffuseness}(n, k)$ là độ khuếch tán của tín hiệu đầu vào được đo trong ô thời gian-tần số của các chỉ số (n, k), và trong đó $\text{band_diff}[]$ xác định giới hạn của các dải tham số theo các chỉ số của dải tần số, và $\text{slot_diff}[]$ xác định giới hạn của việc nhóm theo thời gian trong các chỉ số của các khe thời gian. Ví dụ, các bảng có thể được xác định cho 5 dải tham số và 1 nhóm thời gian là:

$$\text{slot_diff} = [0, 16]$$

$$\text{band_diff} = [0, 1, 3, 7, 15, 60]$$

Vector hướng theo các tọa độ Đề-các được tính là:

$$\text{dv}[g, b] = \frac{\sum_{n=\text{slot_dv}[g]}^{\text{slot_dv}[g+1]-1} \sum_{k=\text{band_dv}[b]}^{\text{band_dv}[b+1]-1} \text{direction}(n, k) \cdot (1 - \text{diffuseness}(n, k)) \cdot \text{power}(n, k)^\alpha}{\sum_{n=\text{slot_dv}[g]}^{\text{slot_dv}[g+1]-1} \sum_{k=\text{band_dv}[b]}^{\text{band_dv}[b+1]-1} (1 - \text{diffuseness}(n, k)) \cdot \text{power}(n, k)^\alpha},$$

trong đó $\text{power}(n, k)^\alpha$ là năng lượng của tín hiệu đầu vào được đo trong ô thời gian-tần số của các chỉ số (t, k) và được tăng lên lũy thừa α , $\text{diffuseness}(n, k)$ là độ khuếch tán của tín hiệu đầu vào được đo trong ô thời gian-tần số của các chỉ số (n, k), và $\text{direction}(n, k)$ là hướng được đo trong ô thời gian-tần số của các chỉ số (n, k) theo các tọa độ Đề-các ba chiều, và trong đó $\text{band_dv}[]$ xác định giới hạn của các dải tham số theo các chỉ số của dải tần số, và $\text{slot_dv}[]$ xác định giới hạn của việc nhóm theo thời gian trong các chỉ

số của các khe thời gian. Ví dụ, các bảng có thể được xác định cho 5 dải tham số và 4 nhóm thời gian là:

$$\text{slot}_{dv} = [0, 4, 8, 12, 16]$$

$$\text{band}_{dv} = [0, 1, 3, 7, 15, 60]$$

Tham số α cho phép nén hoặc mở rộng các trọng số dựa trên công suất trong tổng trọng số được thực hiện để lấy trung bình các tham số. Theo chế độ được ưu tiên, $\alpha = 1$.

Nói chung, giá trị này có thể là số không âm thực, vì số mũ nhỏ hơn 1 cũng có thể hữu ích. Ví dụ số mũ 0,5 (căn bậc hai) vẫn sẽ tăng thêm trọng số cho các tín hiệu liên quan đến biên độ cao hơn, nhưng vừa phải hơn khi so sánh với số mũ bằng 1 hoặc lớn hơn 1.

Sau khi nhóm và tính trung bình, vector có hướng kết quả $dv[g, b]$ nói chung không còn là vector đơn vị. Phép chuẩn hóa là do đó là cần thiết:

$$dv[g, b] = \frac{dv[g, b]}{\|dv[g, b]\|}$$

Sau đây, phương án được ưu tiên của khía cạnh thứ hai của sáng chế được thảo luận. Fig.4a minh họa thiết bị mã hóa các tham số mã hóa âm thanh có hướng bao gồm các tham số độ khuếch tán và các tham số hướng theo khía cạnh thứ hai. Thiết bị bao gồm bộ lượng tử hóa tham số 210 nhận, ở đầu vào của nó, các tham số được nhóm như đã thảo luận liên quan đến khía cạnh thứ nhất hoặc các tham số chưa được nhóm hoặc đã được nhóm khác nhau.

Do đó, bộ lượng tử hóa tham số 210 và bộ mã hóa tham số 220 được kết nối sau đó để mã hóa các tham số độ khuếch tán được lượng tử hóa và các tham số hướng được lượng tử hóa được bao gồm cùng với giao diện đầu ra để tạo ra sự biểu diễn tham số được mã hóa bao gồm thông tin về các tham số độ khuếch tán được mã hóa và các tham số hướng được mã hóa ví dụ trong khối 200 của Fig.1a. Bộ xử lý bộ lượng tử hóa và bộ mã hóa 200 của Fig.2a có thể được triển khai như ví dụ được thảo luận sau đây liên quan đến bộ lượng tử hóa tham số 210 và bộ mã hóa tham số 220, nhưng bộ xử lý

bộ lượng hóa và bộ mã hóa 200 cũng có thể được triển khai theo bất kỳ cách nào khác đối với khía cạnh thứ nhất.

Tốt hơn là, bộ lượng tử hóa tham số 210 của Fig.4a được tạo cấu hình để lượng tử hóa tham số độ khuếch tán như được minh họa ở 231 trên Fig.4b bằng cách sử dụng bộ lượng tử hóa không đồng nhất để tạo ra các chỉ số độ khuếch tán. Bộ mã hóa tham số 220 của Fig.4a được tạo cấu hình như được minh họa trong mục 232, tức là, để mã hóa entropy các giá trị độ khuếch tán thu được cho khung bằng cách sử dụng ba chế độ khác nhau, mặc dù có thể sử dụng một chế độ duy nhất hoặc chỉ hai chế độ khác nhau. Một chế độ là chế độ thô được thực hiện theo cách mà mỗi giá trị độ khuếch tán riêng lẻ được mã hóa bằng cách sử dụng, ví dụ, mã nhị phân hoặc mã nhị phân thung. Ngoài ra, mã hóa vi sai có thể được thực hiện sao cho mỗi chênh lệch và giá trị tuyệt đối ban đầu được mã hóa bằng chế độ thô. Tuy nhiên, tình huống có thể là cùng một khung có cùng độ khuếch tán trên tất cả các dải tần và mã hóa chỉ có một giá trị có thể được sử dụng. Một lần nữa, thay vào đó, tình huống có thể là chỉ có các giá trị liên tiếp cho các độ khuếch tán, tức là, các chỉ số độ khuếch tán liên tiếp trong một khung, và sau đó chế độ mã hóa thứ ba có thể được áp dụng như được minh họa trong khối 232.

Fig.4c minh họa việc triển khai bộ lượng tử hóa tham số 210 của Fig.4a; Bộ lượng tử hóa tham số 210 của Fig.4a được tạo cấu hình để chuyển đổi tham số hướng thành dạng cực như được minh họa ở 233. Trong khối 234, độ chính xác lượng tử hóa cho ngăn được xác định. Ngăn này có thể là ngăn có độ phân giải cao ban đầu hoặc, thay vào đó, và tốt hơn là ngăn được nhóm có độ phân giải thấp.

Như đã thảo luận ở trên đối với Fig.3b và Fig.3c, mỗi dải có cùng giá trị độ khuếch tán nhưng có bốn giá trị hướng khác nhau. Độ chính xác lượng tử hóa tương tự được xác định cho toàn bộ dải, tức là, cho tất cả các tham số hướng trong một dải. Trong khối 235, tham số góc nâng được khối 233 xuất ra được lượng tử hóa bằng cách sử dụng độ chính xác lượng tử hóa. Bảng chữ cái lượng tử hóa để lượng tử hóa tham số góc nâng tốt hơn là cũng được lấy từ độ chính xác lượng tử hóa cho ngăn được xác định trong khối 234.

Với mục đích xử lý giá trị góc phương vị, bảng chữ cái góc phương vị được xác định 236 từ thông tin góc nâng cho ngăn thời gian/tần số (được nhóm) tương ứng.

Thông tin góc nâng có thể là giá trị góc nâng được lượng tử hóa, giá trị góc nâng ban đầu hoặc giá trị góc nâng được lượng tử hóa và một lần nữa được khử lượng tử hóa trong đó giá trị sau, nghĩa là giá trị góc nâng được lượng tử hóa và một lần nữa được khử lượng tử hóa được ưu tiên để có cùng tình huống ở phía bộ mã hóa và ở phía bộ giải mã. Trong khối 237, tham số góc phương vị được lượng tử hóa với bảng chữ cái cho ngăn thời gian/tần số này. Mặc dù có thể có cùng độ chính xác lượng tử hóa của dải như đã thảo luận trước đó đối với Fig.3b, vẫn có thể có các bảng chữ cái góc phương vị khác nhau cho mỗi ngăn thời gian/tần số được nhóm riêng lẻ liên quan đến tham số hướng.

Mã hóa siêu dữ liệu DirAC

Đối với mỗi ngăn, các tham số không gian DirAC được tính trên lưới chứa *nbands* dải tần, và đối với mỗi dải tần *b*, *num_slots* khe thời gian được nhóm thành số lượng *nblocks(b)* nhóm thời gian có kích thước bằng nhau. Tham số độ khuếch tán được gửi cho từng dải tần, và tham số hướng cho từng nhóm thời gian của từng dải tần.

Ví dụ, nếu *nbands* = 5 và *nblocks(b)* = 4, với *num_slots* = 16, điều này sẽ dẫn đến 5 tham số độ khuếch tán và 20 tham số hướng trên mỗi khung, mà sẽ được tiếp tục lượng tử hóa và được mã hóa entropy.

Lượng tử hóa các tham số độ khuếch tán

Mỗi tham số độ khuếch tán *diff(b)* được lượng tử hóa đến một trong số *diff_alpha* mức riêng biệt, bằng cách sử dụng bộ lượng tử hóa không đồng nhất tạo ra chỉ số độ khuếch tán *diff_idx(b)*. Ví dụ, bộ lượng tử hóa có thể được lấy từ bảng lượng tử hóa ICC được sử dụng trong tiêu chuẩn MPS, trong đó các ngưỡng và các mức khôi phục được tính toán bởi hàm *generate_diffuseness_quantizer*.

Tốt hơn là chỉ sử dụng các giá trị không âm từ bảng lượng tử hóa ICC, như *icc* = [1.0, 0.937, 0.84118, 0.60092, 0.36764, 0.0], chỉ chứa 6 mức trong số 8 mức ban đầu. Do ICC bằng 0.0 tương ứng với độ khuếch tán bằng 1.0 và ICC bằng 1.0 tương ứng với độ khuếch tán bằng 0.0, nên tập hợp các tọa độ *y* được tạo là *y* = 1.0 - *icc*, với tập hợp các tọa độ *x* tương ứng là *x* = [0.0, 0.2, 0.4, 0.6, 0.8, 1.0]. Phương pháp nội suy

bậc ba từng phần bảo toàn hình dạng, được biết đến là đa thức nội suy bậc ba từng phần (piecewise cubic hermite interpolating polynomial - PCHIP), được sử dụng để suy ra đường cong đi qua tập hợp các điểm được xác định bởi x và y . Số bước của bộ lượng tử hóa độ khuếch tán là `diff_alph`, mà trong triển khai được đề xuất là 8, nhưng nó không liên quan đến tổng số mức của bảng lượng tử hóa ICC, mà cũng là 8.

Tập hợp mới của `diff_alph` cách đều các tọa độ `x_interpiated` từ 0.0 đến 1.0 (hoặc gần nhưng nhỏ hơn 1.0, khi tránh trường hợp độ khuếch tán thuần túy bằng 1.0 do cân nhắc đến việc kết xuất âm thanh) được tạo ra, và các giá trị y tương ứng trên đường cong được sử dụng làm các giá trị khôi phục, các giá trị khôi phục này được đặt cách nhau không tuyến tính. Các điểm nửa vời giữa các giá trị `x_interpiated` liên tiếp cũng được tạo, và các giá trị y tương ứng của đường cong được sử dụng làm các giá trị ngưỡng để quyết định xem các giá trị nào ánh xạ tới chỉ số độ khuếch tán cụ thể và do đó là giá trị khôi phục. Đối với triển khai được đề xuất, các giá trị khôi phục và ngưỡng được tạo (được làm tròn đến 5 chữ số), được tính bởi hàm `generate_diffuseness_quantizer` là:

khôi phục = [0.0, 0.03955, 0.08960, 0.15894, 0.30835, 0.47388, 0.63232, 0.85010]

ngưỡng = [0.0, 0.01904, 0.06299, 0.11938, 0.22119, 0.39917, 0.54761, 0.73461, 2.0]

Giá trị ngưỡng lớn ngoài phạm vi phần giữ chỗ (2.0) được thêm vào cuối các ngưỡng để giúp tìm kiếm đơn giản hơn. Lấy ví dụ, nếu $\text{diff}(b) = 0.33$, đối với dải b cụ thể, thì $\text{thresholds}[4] \leq \text{diff}(b) < \text{thresholds}[5]$, do đó $\text{diff_idx}(b) = 4$, và giá trị khôi phục tương ứng là $\text{reconstructions}[4] = 0.30835$.

Quy trình trên chỉ là một lựa chọn có sẵn của bộ lượng tử hóa phi tuyến tính cho các giá trị độ khuếch tán.

Mã hóa entropy của các tham số độ khuếch tán

Hàm *EncodeQuasiUniform(value, alphabet_sz)* được sử dụng để mã hóa *value* với xác suất giả đều bằng cách sử dụng mã thùng. Với $value \in \{0, \dots, \text{alphabet_sz} - 1\}$, số các giá trị nhỏ nhất được mã hóa sử dụng $\lceil \log_2 \text{alphabet_sz} \rceil$ bit, và phần còn lại được mã hóa sử dụng $\lceil \log_2 \text{alphabet_sz} \rceil + 1$ bit. Nếu *alphabet_sz* là lũy thừa của hai, các kết quả mã hóa nhị phân.

Tùy thuộc vào các giá trị của chúng, các chỉ số độ khuếch tán được lượng tử hóa có thể được mã hóa bằng cách sử dụng một trong ba phương pháp có sẵn: mã hóa thô, chỉ một giá trị, và chỉ hai giá trị liên tiếp. Bit thứ nhất (*diff_use_raw_coding*) cho biết liệu phương pháp mã hóa thô có được sử dụng hay không. Đối với mã hóa thô, mỗi giá trị chỉ số độ khuếch tán được mã hóa bằng hàm *EncodeQuasiUniform*.

Nếu tất cả các giá trị chỉ số bằng nhau, phương pháp chỉ một giá trị được sử dụng. Bit thứ hai (*diff_have_unique_value*) được sử dụng để chỉ ra phương pháp này, sau đó giá trị đơn nhất được mã hóa bằng hàm *EncodeQuasiUniform*. Nếu tất cả các giá trị chỉ số chỉ chứa hai giá trị liên tiếp, phương pháp chỉ hai giá trị liên tiếp được sử dụng, được biểu thị bằng bit thứ hai ở trên. Giá trị nhỏ hơn trong hai giá trị liên tiếp được mã hóa bằng hàm *EncodeQuasiUniform*, có tính đến kích thước bảng chữ cái của nó được giảm xuống $diff_alph - 1$. Sau đó, với mỗi giá trị, chênh lệch giữa nó và giá trị tối thiểu được mã hóa bằng một bit.

Hàm *EncodeQuasiUniform* (giá trị, *alph_sz*) được ưu tiên triển khai cái được gọi là mã thùng. Nó có thể được xác định trong mã giả là:

```
bits = floor(log2(alphabet_sz))
thresh = 2 ^ (bits + 1) - alphabet_sz
if (value < thresh)
    write_bits(value, bits)
else
    write_bits(value + thresh, bits + 1)
```

Nếu *alphabet_sz* là lũy thừa của 2, thì $alphabet_sz = 2^{bit}$ và $thresh = 2^{bit}$, do đó nhánh khác không bao giờ được sử dụng, và kết quả mã hóa nhị phân. Mặt khác, các giá trị nhỏ nhất của ngưỡng thứ nhất được mã hóa bằng mã nhị phân có bit bit, và phần còn lại, bắt đầu bằng $value = thresh$, được mã hóa bằng mã nhị phân có bit + 1 bit. Mã nhị phân thứ nhất được mã hóa bằng bit + 1 bit có giá trị $value + thresh = thresh + thresh = thresh * 2$, do đó bộ giải mã có thể tìm ra, bằng cách chỉ đọc bit bit thứ nhất và so sánh giá trị của nó với *thresh*, nếu cần sẽ đọc thêm một bit nữa. Hàm giải mã, *DecodeQuasiUniform*(*alphabet_sz*) có thể được xác định trong mã giả là:

```
bits = floor(log2(alphabet_sz))
```

```

thresh = 2 ^ (bits + 1) - alphabet_sz
value = read_bits(bits)
if (value >= thresh)
    value = (value * 2 + read_bits(1)) - thresh
return value

```

Chuyển đổi các tham số hướng thành các tọa độ cực

Mỗi vector chỉ hướng 3 chiều dv được chuẩn hóa sao cho $dv[0]^2 + dv[1]^2 + dv[2]^2 = 1$ được chuyển đổi thành sự biểu diễn cực chứa góc nâng $el \in [-90, 90]$ và góc phương vị $az \in [0, 360]$, sử dụng hàm *DirectionVector2AzimuthElevation*. Việc chuyển đổi hướng ngược lại từ các tọa độ cực thành vector hướng được chuẩn hóa đạt được bằng cách sử dụng hàm *AzimuthElevation2DirectionVector*.

Lượng tử hóa các tham số hướng

Hướng được biểu diễn dưới dạng cặp góc nâng và góc phương vị được còn được lượng tử hóa. Đối với mỗi mức chỉ số độ khuếch tán được lượng tử hóa, độ chính xác góc yêu cầu được chọn từ *angle_spacing* vector cấu hình là $deg_req = angle_spacing(diff_idx(b))$ và được sử dụng để tạo ra tập hợp các điểm lượng tử hóa được phân bố đồng đều trên mặt cầu đơn vị.

Giá trị khoảng cách góc deg_req tốt hơn là không được tính từ độ khuếch tán $diff(b)$, mà từ chỉ số độ khuếch tán $diff_idx(b)$. Do đó, có $diff_alph$ possible deg_req giá trị, một giá trị cho mỗi chỉ số độ khuếch tán có thể có. Ở phía bộ giải mã, độ khuếch tán ban đầu $diff(b)$ không có sẵn, chỉ có chỉ số độ khuếch tán $diff_idx(b)$ mà có thể được sử dụng để chọn cùng một giá trị khoảng cách góc như trong bộ mã hóa. Trong triển khai được đề xuất, bảng khoảng cách góc là:

```
angle_spacing_table = [5.0, 5.0, 7.5, 10.0, 18.0, 30.0, 45.0, 90.0]
```

Các điểm phân bố gần như đồng đều trên mặt cầu đơn vị được tạo ra theo cách như vậy để đáp ứng một số tính chất mong muốn quan trọng. Các điểm nên được phân phối đối xứng với các trục X, Y, và Z. Việc lượng tử hóa hướng nhất định đến điểm gần nhất và ánh xạ tới chỉ số nguyên nên là thao tác thời gian không đổi. Cuối cùng,

việc tính điểm tương ứng trên mặt cầu từ chỉ số nguyên và việc khử lượng tử hóa theo hướng nên là thao tác thời gian không đổi hoặc logarit đối với tổng số điểm trên mặt cầu.

Có hai loại đối xứng đối với trục cho các điểm trên mặt phẳng nằm ngang: có hai điểm xuất hiện trong đó trục trục giao nhau với mặt cầu đơn vị trên mặt phẳng hiện tại, và không có bất kỳ điểm nào xuất hiện. Như ví dụ cho mặt phẳng ngang tùy ý, có ba trường hợp có thể xảy ra. Nếu số điểm là bội số của 4, như 8, thì có sự đối xứng với trục X (trái-phải) và hai điểm có mặt ở 90 và 270 độ trên trục Y, và đối xứng với trục Y (trước-sau) và hai điểm có mặt ở 0 và 180 độ trên trục X. Nếu số điểm chỉ là bội của 2, như 6, thì có sự đối xứng với trục X nhưng không có điểm nào tại 90 và 270 độ trên trục Y, và đối xứng với trục Y và có hai điểm tại 0 và 180 độ trên trục X. Cuối cùng, khi số điểm là một số nguyên tùy ý, như 5, có sự đối xứng với trục X nhưng không có điểm nào tại 90 và 270 độ trên trục Y, và không có sự đối xứng với trục Y.

Theo phương án được ưu tiên, có các điểm tại 0, 90, 180 và 270 độ trên tất cả các mặt phẳng ngang (tương ứng với tất cả các góc nâng được lượng tử hóa) được coi là hữu ích từ góc độ tâm thính học, chỉ ra rằng số điểm trên mỗi mặt phẳng ngang luôn là bội số của 4. Tuy nhiên, tùy thuộc vào ứng dụng cụ thể, điều kiện về số lượng điểm trên mỗi mặt phẳng ngang có thể được nói lỏng thành bội số của 2 hoặc một số nguyên tùy ý.

Ngoài ra, trong phương án được ưu tiên, đối với mỗi góc nâng, điểm phương vị "gốc" luôn tồn tại ở hướng trọi 0 độ (về phía trước). Tính chất này có thể được nói lỏng bằng cách chọn góc bù lượng tử hóa được tính toán trước cho từng góc nâng riêng biệt, với các điểm phương vị được phân phối tương ứng với nó thay vì hướng 0 độ. Nó có thể được triển khai dễ dàng bằng cách thêm phần bù trước khi lượng tử hóa và trừ nó sau khi khử lượng tử hóa.

Độ chính xác góc cần thiết là deg_req và nên là ước của 90 độ. Mặt khác, nó được tính toán lại trước khi sử dụng thực tế là $deg_req = 90 \div [90 \div deg_req]$. Ví dụ, danh sách các giá trị có thể có là $\{90, 45, 30, 22.5, 18, 15, 12.86, 11.25, 10, \dots, 5, \dots\}$. Góc nâng el được lượng tử hóa đồng đều, với kích thước bước deg_req , tạo ra

$el_idx = \text{round}(el + deg_req) + n_points$, một trong $el_alph = 2 \cdot n_points + 1$ chỉ số lượng tử hóa, trong đó $n_points = \lceil 90 + deg_req \rceil$. Chỉ số này tương ứng với góc nâng được lượng tử hóa $q_el = (el_idx - n_points) \cdot deg_req$. Tương đương, chỉ dựa trên kích thước bảng chữ cái, $el_idx = \text{round}\left(\frac{(el + 90) + 180}{deg_req} \cdot (el_alph - 1)\right)$ cho lượng tử hóa và $q_el = (el_idx + (el_alph - 1)) \cdot deg_req - 90$ cho sự khử lượng tử hóa.

Tại xích đạo, góc nâng az được lượng tử hóa đồng đều, với kích thước bước deg_req , tạo ra az_idx , một trong $4 \cdot n_points$ chỉ số lượng tử hóa. Đối với các góc nâng khác, khoảng cách góc ngang nhìn từ tâm của khối cầu đơn vị, tương ứng với độ dài hợp âm giữa hai điểm liên tiếp, có thể xấp xỉ bằng độ dài cung trên vòng tròn nằm ngang nằm ở q_el góc nâng. Do đó, số điểm tương ứng với 90 độ trên vòng tròn nằm ngang này bị giảm, so với vòng tròn xích đạo, tỷ lệ với bán kính của nó, do đó độ dài cung giữa hai điểm liên tiếp vẫn gần như nhau ở mọi nơi. Ở các cực, tổng số điểm trở thành một.

Có $az_alph = \max(4 \cdot \text{round}(\text{radius_len} \cdot n_points), 1)$ chỉ số lượng tử hóa, tương ứng với q_el góc nâng, trong đó $\text{radius_len} = \cos(q_el)$. Chỉ số lượng tử hóa tương ứng là $az_idx = \text{round}\left(\frac{az + 360}{deg_req} \cdot az_alph\right)$, trong đó giá trị tạo ra của az_alph được thay bằng 0. Chỉ số này tương ứng với góc phương vị được lượng tử hóa $q_az = az_idx \cdot deg_req$. Lưu ý, không bao gồm các cực trong đó $az_alph = 1$, các giá trị nhỏ nhất gần các cực là $az_alph = 4$ với $deg_req = 90$ và $deg_req = 45$, và $az_alph = 8$ với tất cả các phần còn lại.

Nếu điều kiện về số lượng điểm trên mỗi mặt phẳng nằm ngang chỉ là bội số của 2 thì bảng chữ cái góc phương vị trở thành $az_alph = \max(2 \cdot \text{round}(\text{radius_len} \cdot (2 \cdot n_points)), 1)$, bởi vì có $2 \cdot n_points$ tương ứng với 180 độ trên mặt phẳng xích đạo. Nếu điều kiện về số điểm được nói lỏng thành một số nguyên tùy ý, bảng chữ cái góc phương vị trở thành $az_alph = \max(\text{round}(\text{radius_len} \cdot (4 \cdot n_points)), 1)$, bởi vì có $4 \cdot n_points$ tương ứng với 360 độ trên mặt phẳng xích đạo. Trong cả hai trường hợp, trên mặt phẳng xích đạo, số điểm luôn là bội số của 4, bởi vì $\text{radius_len} = 1$ và n_points là số nguyên.

Quá trình lượng tử hóa và khử lượng tử hóa được mô tả ở trên đạt được bằng cách sử dụng các hàm *QuantizeAzimuthElevation* và *DequantizeAzimuthElevation*.

Tốt hơn là, hàm $\text{round}(x)$ làm tròn x đến số nguyên gần nhất, thường được thực hiện theo điểm cố định là $\text{round}(x) = \text{floor}(x + 0.5)$. Việc làm tròn cho các số thập phân giá trị chính xác nửa vởi giữa các số nguyên, như 1,5, có thể được thực hiện theo một vài cách. Định nghĩa trên làm tròn các số thập phân hướng tới + vô cực (1,5 được làm tròn thành 2, 2,5 được làm tròn thành 3). Việc triển khai dấu phẩy động thường có chức năng làm tròn hàm số nguyên, làm tròn các số thập phân thành các số nguyên chẵn (1,5 được làm tròn thành 2, 2,5 được làm tròn thành 2).

Fig.4d được biểu thị là "độ bao phủ gần như đồng nhất của mặt cầu đơn vị" minh họa ví dụ về độ bao phủ gần như đồng đều của mặt cầu đơn vị sử dụng độ chính xác góc 15 độ, thể hiện các hướng được lượng tử hóa. Hình chiếu 3D là từ phía trên, chỉ có bán cầu trên được vẽ để hình dung tốt hơn, và đường xoắn ốc được nối bằng các chấm chỉ để hình dung trực quan dễ dàng hơn các điểm từ cùng một đường tròn hoặc mặt phẳng ngang.

Sau đây, triển khai được ưu tiên của bộ mã hóa tham số 220 của Fig.4a với mục đích mã hóa các tham số hướng được lượng tử hóa, tức là, các chỉ số góc nâng được lượng tử hóa và các chỉ số góc phương vị được lượng tử hóa được minh họa. Như được minh họa trên Fig.5a, bộ mã hóa được tạo cấu hình để phân loại 240 từng khung liên quan đến các giá trị độ khuếch tán trong khung. Khối 240 nhận các giá trị độ khuếch tán, theo phương án của Fig.3b, chỉ có năm giá trị độ khuếch tán cho khung. Nếu khung chỉ bao gồm các giá trị độ khuếch tán thấp, chế độ mã hóa độ khuếch tán thấp 241 được áp dụng. Khi năm giá trị độ khuếch tán trong khung chỉ là các giá trị độ khuếch tán cao, thì chế độ mã hóa độ khuếch tán cao 242 được áp dụng. Khi xác định rằng các giá trị độ khuếch tán trong khung ở cả dưới và trên ngưỡng độ khuếch tán ec_max , thì chế độ mã hóa độ khuếch tán hỗn hợp 243 được áp dụng. Trong cả chế độ mã hóa độ khuếch tán thấp 241 và chế độ mã hóa độ khuếch tán cao 242, và cả đối với các dải độ khuếch tán thấp, đối với khung độ khuếch tán hỗn hợp, việc mã hóa mã hóa thô trên một mặt và mã hóa entropy trên mặt khác được thử, tức là, được thực hiện như được biểu thị ở 244a, 244b và 244c. Tuy nhiên, đối với các dải độ khuếch tán cao

trong khung độ khuếch tán hỗn hợp, chế độ mã hóa thô luôn được sử dụng như được biểu thị ở 244d.

Trong trường hợp sử dụng các chế độ mã hóa khác nhau, tức là, chế độ mã hóa thô và chế độ mã hóa entropy (có mô hình hóa) được sử dụng, kết quả được chọn bởi bộ điều khiển bộ mã hóa chọn chế độ dẫn đến số bit nhỏ hơn để mã hóa các chỉ số lượng tử hóa. Điều này được biểu thị ở 245a, 245b và 245c.

Mặt khác, người ta chỉ có thể sử dụng chế độ mã hóa thô cho tất cả các khung và dải hoặc chỉ chế độ mã hóa entropy với mô hình hóa cho tất cả các dải hoặc bất kỳ chế độ mã hóa nào khác để mã hóa các chỉ số chẳng hạn như chế độ mã hóa Huffman hoặc chế độ mã hóa số học thích ứng hoặc không thích ứng bối cảnh.

Tùy thuộc vào kết quả của quy trình được chọn trong các khối 245a, 245b và 245c, thông tin phụ được thiết lập cho toàn bộ khung như được minh họa trong các khối 246a, 246b hoặc chỉ được thiết lập cho các dải tương ứng, nghĩa là các dải có độ khuếch tán thấp trong khối 246c. Ngoài ra, thông tin phụ cũng có thể được thiết lập cho toàn bộ khung trong trường hợp mục 246c. Trong trường hợp này, việc xác định các dải có độ khuếch tán cao có thể được thực hiện chỉ trong bộ giải mã để mặc dù thông tin phụ được thiết lập cho toàn bộ khung, thì bộ giải mã xác định rằng có khung khuếch tán hỗn hợp và các tham số hướng cho các dải có giá trị độ khuếch tán cao trong khung độ khuếch tán hỗn hợp này được mã hóa bằng chế độ mã hóa thô mặc dù thông tin phụ cho khung cho biết chế độ mã hóa entropy với mô hình hóa.

Theo phương án được ưu tiên, $\text{diff_alph} = 8$. Sau đó, giá trị ngưỡng ec_max được chọn là 5, bằng cách giảm thiểu kích thước nén trung bình trên lượng thử nghiệm lớn. Giá trị ngưỡng ec_max này được sử dụng trong chế độ sau, tùy thuộc vào phạm vi các giá trị cho các chỉ số độ khuếch tán của khung hiện thời:

- đối với các khung có độ khuếch tán thấp đến trung bình, trong đó $\text{diff_idx} (b) \leq \text{ec_max}$, cho tất cả các dải b , tất cả các hướng được mã hóa bằng cả mã hóa thô và entropy và tốt nhất được chọn và được biểu thị bởi một bit làm thông tin phụ (được xác định ở trên là $\text{dir_use_raw_coding}$);

- đối với các khung độ khuếch tán hỗn hợp, trong đó $\text{diff_idx}(b) \leq \text{ec_max}$, đối với một số dải b , các hướng tương ứng với các dải đó được mã hóa chính xác như trong trường hợp thứ nhất; tuy nhiên, đối với các dải có độ khuếch tán cao b khác, trong đó $\text{diff_idx}(b) > \text{ec_max}$, các hướng tương ứng với các dải khác này luôn được mã hóa thành dạng thô (để tránh trộn lẫn các thống kê mã hóa entropy của các hướng có độ khuếch tán thấp đến trung bình với các độ khuếch tán cao mà cũng được lượng tử hóa rất thô);

- đối với các khung có độ khuếch tán cao, trong đó $\text{diff_idx}(b) > \text{ec_max}$, đối với tất cả các dải b , ngưỡng ec_max được thiết lập trước thành $\text{ec_max} = \text{diff_alph}$ cho khung hiện thời (vì các chỉ số độ khuếch tán được mã hóa trước các hướng, thiết lập này có thể được thực hiện trước khi nhận dạng ở phía bộ giải mã), do đó trường hợp này trở thành trường hợp thứ nhất.

Fig.5b minh họa quá trình tiền xử lý được ưu tiên nhưng theo cách tùy chọn của các chỉ số hướng cho cả hai chế độ. Đối với cả hai chế độ, các chỉ số hướng được lượng tử hóa, tức là các chỉ số góc phương vị được lượng tử hóa và các chỉ số góc nâng được lượng tử hóa được xử lý trong khối 247 thành sự chuyển đổi các chỉ số góc nâng/góc phương vị tạo ra các giá trị có dấu, trong đó chỉ số không tương ứng với góc nâng hoặc góc phương vị không. Việc chuyển đổi 248 tiếp theo thành các giá trị không dấu bao gồm việc xen kẽ các giá trị dương/âm được thực hiện để có sự biểu diễn nhỏ gọn hơn của các chỉ số góc phương vị/góc nâng không dấu được sắp xếp lại.

Fig.5c minh họa triển khai được ưu tiên của chế độ mã hóa thứ nhất 260, tức là chế độ mã hóa thô mà không cần mô hình hóa. Các chỉ số góc phương vị/góc nâng được xử lý trước được nhập vào khối 261 để gộp cả hai chỉ số thành một chỉ số hình cầu duy nhất. Dựa trên độ chính xác lượng tử hóa có được từ chỉ số độ khuếch tán có liên quan, tức là deg_req , phép mã hóa có chức năng mã hóa như EncodeQuasiUniform hoặc mã nhị phân (thùng) được thực hiện 262. Do đó, thu được các chỉ số hình cầu được mã hóa cho các dải hoặc cho toàn bộ khung. Các chỉ số hình cầu được mã hóa cho toàn bộ khung thu được trong trường hợp khung chỉ có độ khuếch tán thấp trong đó mã hóa thô đã được chọn hoặc trong khung chỉ có độ khuếch tán cao, trong đó một lần nữa mã hóa thô được chọn hoặc các chỉ số hình cầu được mã

hóa chỉ cho các dải độ khuếch tán cao của khung thu được trong trường hợp khung độ khuếch tán hỗn hợp được biểu thị ở 243 trên Fig.5a, trong đó, đối với các dải khác có độ khuếch tán thấp hoặc trung bình, chế độ mã hóa thứ hai chẳng hạn như mã hóa entropy với mô hình hóa đã được chọn.

Fig.5d minh họa chế độ mã hóa thứ hai này mà ví dụ có thể là chế độ mã hóa entropy với mô hình hóa. Các chỉ số được xử lý trước mà ví dụ được phân loại cho khung độ khuếch tán hỗn hợp như minh họa trên Fig.5a tại 240 được nhập vào khối 266 mà thu thập dữ liệu lượng tử hóa tương ứng như các chỉ số góc nâng, các bảng chữ cái góc nâng, các chỉ số góc phương vị, các bảng chữ cái góc phương vị, và dữ liệu này được thu thập vào các vector riêng cho khung. Trong khối 267, các giá trị trung bình được tính cho góc nâng và góc phương vị rõ ràng dựa trên thông tin có được từ sự khử lượng tử hóa và chuyển đổi vector tương ứng như được thảo luận ở phần sau. Các giá trị trung bình này được lượng tử hóa với độ chính xác góc cao nhất được sử dụng trong khung được biểu thị ở khối 268. Các chỉ số góc nâng và góc phương vị được dự đoán được tạo ra từ các giá trị trung bình như được minh họa trong khối 269, và khoảng cách có dấu cho góc nâng và góc phương vị từ các chỉ số ban đầu và liên quan đến các chỉ số góc nâng và góc phương vị dự đoán được tính toán và giảm xuống tùy ý tới khoảng giá trị nhỏ hơn khác.

Như được minh họa trên Fig.5e, dữ liệu được tạo bởi thao tác mô hình hóa bằng cách sử dụng thao tác chiếu để suy ra các giá trị dự đoán được minh họa trên Fig.5d được mã hóa entropy. Thao tác mã hóa này được minh họa trên Fig.5e cuối cùng tạo ra các bit mã hóa từ dữ liệu tương ứng. Trong khối 271, các giá trị trung bình cho góc phương vị và góc nâng được chuyển đổi thành các giá trị có dấu, và việc sắp xếp lại 272 nhất định được thực hiện để có sự biểu diễn nhỏ gọn hơn, và các giá trị trung bình đó được mã hóa 273 bằng mã nhị phân hoặc mã nhị phân thung để tạo ra các bit trung bình góc nâng 274 và các bit trung bình góc phương vị. Trong khối 275, tham số Golomb-Rice được xác định như được minh họa trên Fig.5f, và tham số này sau đó cũng được mã hóa bằng mã nhị phân (thung) được minh họa ở khối 276 để có tham số Golomb-Rice cho góc nâng và một tham số Golomb-Rice khác cho góc phương vị được minh họa ở 277. Trong khối 278, các khoảng cách có dấu (giảm) được tính bởi khối 270 được sắp xếp lại và sau đó được mã hóa bằng phương pháp Golomb-Rice mở

rộng được minh họa ở 279 để có các khoảng cách góc nâng được mã hóa và các khoảng cách góc phương vị được biểu thị ở 280.

Fig.5f minh họa triển khai được ưu tiên để xác định tham số Golomb-Rice trong khối 275 mà được thực hiện cả để xác định tham số Golomb-Rice góc nâng hoặc để xác định tham số Golomb-Rice góc phương vị. Trong khối 281, khoảng được xác định cho tham số Golomb-Rice tương ứng. Trong khối 282, tổng số bit cho tất cả các khoảng cách có dấu giảm được tính, cho mỗi giá trị ứng cử, và trong khối 283, giá trị ứng cử dẫn đến số bit nhỏ nhất được chọn làm tham số Golomb-Rice cho việc xử lý góc phương vị hoặc góc nâng.

Tiếp theo, Fig.5g được thảo luận để minh họa thêm cho quy trình trong khối 279 của Fig.5e, tức là, phương pháp Golomb-Rice mở rộng. Dựa trên tham số Golomb-Rice đã chọn p , chỉ số khoảng cách hoặc cho góc nâng hoặc cho góc phương vị được phân tách trong phần có nghĩa nhất MSP và phần ít có nghĩa hơn LSP được minh họa ở bên phải của khối 284. Trong khối 285, bit không kết thúc của phần MSP bị loại bỏ, trong trường hợp khi MSP là giá trị tối đa có thể, và trong khối 286, kết quả được mã hóa bằng mã nhị phân (thùng).

Phần LSP cũng được mã hóa bằng mã nhị phân (thùng) được minh họa ở 287. Do đó, trên các đường 288 và 289, các bit được mã hóa cho phần có nghĩa nhất MSP và các bit được mã hóa cho phần ít có nghĩa hơn LSP thu được mà cùng biểu diễn cho các khoảng cách có dấu giảm được mã hóa tương ứng cho góc nâng hoặc góc phương vị.

Fig.8d minh họa ví dụ cho hướng được mã hóa. Bit chế độ 806 biểu thị, ví dụ, chế độ mã hóa entropy với mô hình hóa. Mục 808a minh họa các bit trung bình góc phương vị và mục 808b minh họa các bit trung bình góc nâng như đã thảo luận trước đó đối với mục 274 của Fig.5e. Tham số góc phương vị Golomb-Rice 808c và tham số góc nâng Golomb-Rice 808d cũng có trong dạng được mã hóa trong dòng bit của Fig.8d tương ứng với những gì đã thảo luận trước đó đối với mục 277. Khoảng cách góc nâng được mã hóa và khoảng cách góc phương vị được mã hóa 808e và 808f có trong dòng bit như thu được ở 288 và 289 hoặc như đã thảo luận trước đó đối với mục 280 trên Fig.5e và Fig.5g. Mục 808g minh họa các bit tải tiếp theo cho các khoảng

cách góc nâng/góc phương vị tiếp theo. Các giá trị trung bình cho góc nâng và góc phương vị và các tham số Golomb-Rice cho góc nâng và góc phương vị chỉ được yêu cầu một lần duy nhất cho mỗi khung, nhưng nếu cần cũng có thể được tính hai lần cho khung hoặc hơn, nếu khung khá dài hoặc các thông kê tín hiệu thay đổi mạnh mẽ trong khung.

Fig.8c minh họa dòng bit khi bit chế độ biểu thị mã hóa thô như được xác định trên Fig.5c, khối 260. Bit chế độ 806 biểu thị chế độ mã hóa thô và mục 808 biểu thị các bit tải cho các chỉ số hình cầu, tức là, kết quả của khối 262 của Fig.5c.

Mã hóa entropy của các tham số hướng

Khi mã hóa hướng được lượng tử hóa, chỉ số góc nâng *el_idx* luôn được mã hóa trước, trước chỉ số góc phương vị *az_idx*. Nếu cấu hình hiện tại chỉ tính đến mặt phẳng xích đạo nằm ngang, thì không có gì được mã hóa cho góc nâng và nó được coi là không ở mọi nơi.

Trước khi mã hóa, các giá trị có dấu được ánh xạ tới các giá trị không dấu bằng cách sử dụng phép chuyển đổi sắp xếp lại chung, xen kẽ các số dương và âm thành các số không dấu là $u_val = 2 \cdot |s_val| - (s_val < 0)$, được triển khai bởi hàm *ReorderGeneric*. Biểu thức (*condition*) đánh giá thành 1 nếu *condition* là đúng, và đánh giá thành 0 nếu *condition* là sai.

Bởi vì số giá trị không dấu nhỏ hơn được mã hóa hiệu quả hơn, với ít bit hơn một chút, bằng cách sử dụng hàm *EncodeQuasiUniform*, cả chỉ số góc nâng và chỉ số góc phương vị mà đã được bỏ dấu được chuyển đổi thành có dấu sao cho giá trị chỉ số có dấu bằng không tương ứng với góc nâng hoặc góc phương vị bằng không, và chỉ sau đó hàm *ReorderGeneric* được áp dụng. Bằng cách trước tiên chuyển đổi thành có dấu, giá trị không nằm ở giữa khoảng giá trị có dấu có thể có, và sau khi áp dụng hàm *ReorderGeneric*, giá trị chỉ số góc nâng được sắp xếp lại không dấu tạo ra là $el_idx_r = \text{ReorderGeneric}(el_idx - |el_alph + 2|)$, và giá trị chỉ số góc phương vị được sắp xếp lại không dấu tạo ra là $az_idx_r = \text{ReorderGeneric}(az_idx - az_alph \cdot (az_idx \geq az_alph + 2))$.

Đối với mã hóa thô không có mô hình hóa, hai chỉ số được sắp xếp lại không dấu được gộp thành chỉ số hình cầu không dấu duy nhất $sphere_idx = sphere_offsets(deg_req, el_idx_r) + az_idx_r$, trong đó hàm $sphere_offsets$ tính tổng của tất cả các bảng chữ cái góc phương vị az_alph tương ứng với các chỉ số góc nâng được sắp xếp lại không dấu nhỏ hơn el_idx_r . Ví dụ: khi $deg_req = 90$, trong đó $el_idx_r = 0$ (góc nâng 0 độ) có $az_alph = 4$, $el_idx_r = 1$ (góc nâng -90 độ) có $az_alph = 1$, và $el_idx_r = 2$ (góc nâng 90 độ) có $az_alph = 1$, $sphere_offsets(90,2)$ sẽ lấy giá trị $4 + 1$. Nếu cầu hình hiện tại chỉ tính đến mặt phẳng xích đạo nằm ngang, thì el_idx_r luôn luôn bằng 0 và chỉ số hình cầu không dấu đơn giản hóa thành $sphere_idx = az_idx_r$. Nói chung, tổng số điểm trên hình cầu, hoặc số điểm đếm trên hình cầu là $sphere_alph = sphere_offsets(deg_req, el_alph + 1)$.

Chỉ số hình cầu không dấu $sphere_idx$ được mã hóa bằng hàm *EncodeQuasiUniform*. Đối với mã hóa entropy có mô hình hóa, các hướng được lượng tử hóa được nhóm thành hai loại. Loại thứ nhất chứa các hướng được lượng tử hóa cho các chỉ số độ khuếch tán $diff_idx(b) \leq ec_max$ được mã hóa entropy, và loại thứ hai chứa các hướng được lượng tử hóa cho các chỉ số độ khuếch tán $diff_idx(b) > ec_max$ được mã hóa thô, trong đó ec_max là ngưỡng được chọn tối ưu tùy thuộc vào $diff_alph$. Cách tiếp cận này hoàn toàn loại trừ khỏi việc mã hóa entropy các dải tần có độ khuếch tán cao, khi các dải tần có độ khuếch tán thấp đến trung bình cũng có mặt trong khung, để tránh trộn lẫn các số liệu thống kê của các phần dư. Đối với khung độ khuếch tán hỗn hợp, mã hóa thô luôn được sử dụng cho các dải tần có độ khuếch tán cao. Tuy nhiên, nếu tất cả các dải tần có độ khuếch tán cao, $diff_idx(b) > ec_max$, ngưỡng được đặt trước thành $ec_max = diff_alph$ để cho phép mã hóa entropy cho tất cả các dải tần.

Đối với loại hướng được lượng tử hóa thứ nhất được mã hóa entropy, các chỉ số góc nâng tương ứng el_idx , các bảng chữ cái góc nâng el_alph , các chỉ số góc phương vị az_idx , và các bảng chữ cái góc phương vị az_alph được thu thập thành các vector riêng để xử lý thêm.

Vector hướng trung bình được suy ra, bằng cách chuyển đổi từng hướng được lượng tử hóa được mã hóa entropy trở lại thành vector hướng, tính toán giá trị trung bình, trung tuyến, hoặc chế độ của các vector hướng chứa phép tái chuẩn hóa, và chuyển đổi vector hướng trung bình thành góc nâng trung bình el_avg và góc phương vị az_avg . Hai giá trị này được lượng tử hóa bằng cách sử dụng độ chính xác góc tốt nhất deg_req được sử dụng bởi các hướng được lượng tử hóa mà được mã hóa entropy, ký hiệu là deg_req_avg , thường là độ chính xác góc cần thiết tương ứng với chỉ số độ khuếch tán nhỏ nhất $\min(diff_idx(b))$, for $b \in \{0, \dots, nbands - 1\}$ and $diff_idx(b) \leq ec_max$.

Sử dụng giá trị n_points_avg tương ứng được suy ra từ deg_req_avg , el_avg được lượng tử hóa thường tạo ra el_avg_idx và el_avg_alph , tuy nhiên, az_avg được lượng tử hóa sử dụng độ chính xác tại đường xích đạo, tạo ra az_avg_idx và $az_avg_alph = 4 \cdot n_points_avg$.

Đối với mỗi hướng được mã hóa entropy, góc nâng trung bình q_el_avg và góc phương vị q_az_avg được chiếu sử dụng độ chính xác của hướng đó, để thu được chỉ số góc nâng và góc phương vị dự đoán. Đối với chỉ số góc nâng el_idx , độ chính xác của nó mà có thể được suy ra từ el_alph được sử dụng để tính toán chỉ số góc nâng trung bình được chiếu $el_avg_idx_p$. Đối với chỉ số góc phương vị tương ứng az_idx , độ chính xác của nó trên đường tròn nằm ngang nằm ở góc nâng q_el mà có thể được suy ra từ az_alph được sử dụng để tính chỉ số góc phương vị trung bình được chiếu $az_avg_idx_p$.

Phép chiếu để thu được các chỉ số góc nâng và góc phương vị dự đoán có thể được tính theo một số cách tương đương. Đối với góc nâng $el_avg_idx_p = \text{round}\left(\frac{q_el_avg + 90}{180} \cdot (el_alph - 1)\right)$ mà có thể dễ dàng được rút gọn thành $el_avg_idx_p = \text{round}\left(\frac{el_avg_idx}{el_avg_alph - 1} \cdot (el_alph - 1)\right)$. Để tạo điều kiện cho thao tác khớp bit, công thức trước đó có thể được viết lại bằng cách sử dụng toán học chỉ trên số nguyên, chứa phép chia, như

$el_avg_idx_p = (2 \cdot el_avg_idx \cdot (el_alph - 1) + (el_avg_alph - 1)) \text{ div } (2 \cdot (el_avg_alph - 1))$.
 . Đối với góc phương vị, $az_avg_idx_p = \text{round}\left(\frac{q_az_avg}{360} \cdot az_alph\right) \bmod az_alph$, mà có thể dễ dàng được rút gọn thành $az_avg_idx_p = \text{round}\left(\frac{az_avg_idx}{az_avg_alph} \cdot az_alph\right) \bmod az_alph$.
 Để tạo điều kiện cho thao tác khớp bit, công thức trước đó có thể được viết lại bằng cách sử dụng toán học chỉ trên số nguyên, chứa phép chia, như $az_avg_idx_p = ((2 \cdot az_avg_idx \cdot az_alph + az_avg_alph) \text{ div } (2 \cdot az_avg_alph)) \bmod az_alph$.
 . Ở các cực, trong đó $az_alph = 1$, ta luôn có $az_idx = 0$ và đặt $az_avg_idx_p = 0$ trực tiếp.

Khoảng cách có dấu el_idx_dist được tính là chênh lệch giữa từng chỉ số góc nâng el_idx và $el_avg_idx_p$ tương ứng của nó. Ngoài ra, do sự chênh lệch tạo ra các giá trị trong khoảng $\{-el_alph + 1, \dots, el_alph - 1\}$, chúng được giảm xuống khoảng $\{-|el_alph \div 2|, \dots, |el_alph \div 2|\}$ bằng cách cộng el_alph cho các giá trị quá nhỏ và trừ el_alph cho các giá trị quá lớn, như trong số học môđun. Nếu khoảng cách giảm này liên quan đến $el_avg_idx_p$ được giải thích bằng cách sử dụng việc bao quanh, nó có thể tạo ra tất cả các giá trị từ đến bằng chữ cái không dấu có chứa el_alph giá trị.

Tương tự, khoảng cách có dấu az_idx_dist được tính là chênh lệch giữa từng chỉ số góc phương vị az_idx và $az_avg_idx_p$ tương ứng của nó. Thao tác chênh lệch tạo ra các giá trị trong khoảng $\{-az_alph + 1, \dots, az_alph - 1\}$ được giảm xuống khoảng $\{-az_alph \div 2, \dots, az_alph \div 2 - 1\}$ bằng cách cộng az_alph cho các giá trị quá nhỏ và trừ az_alph cho các giá trị quá lớn. Khi $az_alph = 1$, chỉ số góc phương vị luôn là $az_idx = 0$ và không có gì cần phải mã hóa.

Tùy thuộc vào giá trị của chúng, các chỉ số góc nâng và góc phương vị được lượng tử hóa có thể được mã hóa bằng một trong hai phương pháp có sẵn: mã hóa thô hoặc mã hóa entropy. Bit thứ nhất (*diff_use_raw_coding*) cho biết liệu phương pháp mã hóa thô có được sử dụng hay không. Đối với mã hóa thô, các chỉ số hình cầu không dấu đơn *sphere_index* được gộp được mã hóa trực tiếp bằng hàm *EncodeQuasiUniform*.

Mã hóa Entropy bao gồm một số phần. Tất cả các chỉ số góc nâng và góc phương vị được lượng tử hóa tương ứng với các chỉ số độ khuếch tán $diff_idx(b) > ec_max$ được mã hóa giống như mã hóa thô. Sau đó, đối với các phần khác, phần góc nâng được mã hóa entropy trước tiên, tiếp theo là phần góc phương vị.

Phần góc nâng chứa ba thành phần: chỉ số góc nâng trung bình, tham số Golomb-Rice, và các khoảng cách góc nâng có dấu giảm. Chỉ số góc nâng trung bình el_avg_idx được chuyển đổi thành có dấu, sao cho giá trị không nằm ở giữa khoảng có dấu của các giá trị có thể có, hàm *ReorderGeneric* được áp dụng, và kết quả được mã hóa bằng hàm *EncodeQuasiUniform*. Tham số Golomb-Rice, có kích thước bằng chữ cái tùy thuộc vào giá trị tối đa của các kích thước bằng chữ cái của các chỉ số góc nâng, được mã hóa bằng hàm *EncodeQuasiUniform*. Cuối cùng, với mỗi khoảng cách góc nâng có dấu giảm el_idx_dist , hàm *ReorderGeneric* được áp dụng để tạo $el_idx_dist_r$, và kết quả được mã hóa bằng phương pháp Golomb-Rice mở rộng với tham số được chỉ ra ở trên.

Ví dụ, nếu độ chính xác góc tốt nhất deg_req_min được sử dụng là 5 độ, thì giá trị tối đa của kích thước bằng chữ cái góc nâng el_alph sẽ là $el_alph_max = 2 \cdot \lceil 90 \div deg_req_min \rceil + 1 = 37$. Trong trường hợp này, các giá trị tham số Golomb-Rice (ký hiệu là p trong mô tả của phương pháp Golomb-Rice bên dưới) bị giới hạn trong khoảng $\{0,1,2,3,4\}$. Nói chung, giá trị hữu ích thế vị lớn nhất của tham số Golomb-Rice là $\lceil \log_2 el_alph_max \rceil - 1$ tạo ra các từ mã nhị phân của chiều dài bằng hoặc dài hơn một chút so với độ dài được tạo bởi mã hóa thô bằng hàm *EncodeQuasiUniform*. Giá trị tối ưu của tham số Golomb-Rice el_gr_param được chọn bằng cách tính toán hiệu quả mà không cần mã hóa, cho từng giá trị trong khoảng trên, tổng kích thước tính theo bit cho cả $el_idx_dist_r$ giá trị được mã hóa bằng phương pháp Golomb-Rice mở rộng, và chọn một giá trị cung cấp kích thước bit nhỏ nhất.

Phần góc phương vị cũng chứa ba thành phần: chỉ số góc phương vị trung bình, tham số Golomb-Rice, và các khoảng cách góc phương vị có dấu giảm. Chỉ số góc phương vị trung bình az_avg_idx được chuyển đổi thành có dấu, sao cho giá trị không

nằm ở giữa khoảng có dấu của các giá trị có thể có, hàm *ReorderGeneric* được áp dụng, và kết quả được mã hóa bằng hàm *EncodeQuasiUniform*. Tham số Golomb-Rice, có kích thước bảng chữ cái tùy thuộc vào giá trị tối đa của các kích thước bảng chữ cái của các chỉ số góc phương vị, được mã hóa bằng hàm *EncodeQuasiUniform*. Cuối cùng, với mỗi khoảng cách góc phương vị có dấu giảm *az_idx_dist*, hàm *ReorderGeneric* được áp dụng để tạo *az_idx_dist_r*, và kết quả được mã hóa bằng phương pháp Golomb-Rice mở rộng với tham số được chỉ ra ở trên.

Ví dụ, nếu độ chính xác góc tốt nhất *deg_req_min* được sử dụng là 5 độ, thì giá trị tối đa của kích thước bảng chữ cái góc phương vị *az_alph* sẽ là $az_alph_max = 4 \cdot [90 \div deg_req_min] = 72$. Trong trường hợp này, các giá trị tham số Golomb-Rice (ký hiệu là p trong mô tả phương pháp Golomb-Rice bên dưới) được giới hạn trong khoảng $\{0,1,2,3,4,5\}$. Giá trị tối ưu của tham số Golomb-Rice *az_gr_param* được chọn bằng cách tính toán hiệu quả, cho từng giá trị trong khoảng trên, tổng kích thước tính theo bit cho cả *az_idx_dist_r* giá trị được mã hóa bằng phương pháp Golomb-Rice mở rộng, và chọn một giá trị cung cấp kích thước bit nhỏ nhất.

Một tính chất quan trọng cần tính đến cho mã hóa entropy hiệu quả là mỗi khoảng cách góc nâng giảm được sắp xếp lại *el_idx_dist_r* có thể có kích thước bảng chữ cái khác nhau, chính xác là *el_alph* của giá trị chỉ số góc nâng ban đầu *el_idx*, và phụ thuộc vào chỉ số độ khuếch tán tương ứng *diff_idx(b)*. Ngoài ra, mỗi khoảng cách góc phương vị giảm được sắp xếp lại *az_idx_dist_r* có thể có kích thước bảng chữ cái khác nhau, chính xác là *az_alph* của giá trị chỉ số góc phương vị ban đầu *az_idx*, và phụ thuộc cả vào q_{el} tương ứng của đường tròn ngang và chỉ số độ khuếch tán *diff_idx(b)* của nó.

Phương pháp mã hóa entropy Golomb-Rice hiện có với tham số nguyên $p \geq 0$ được sử dụng để mã hóa số nguyên không dấu u . Trước tiên, u được chia thành phần ít có nghĩa nhất với p bit $u_lsp = u \bmod 2^p$, và phần có nghĩa nhất $u_msp = \lfloor u \div 2^p \rfloor$. Phần

có nghĩa nhất được mã hóa đơn phân, sử dụng một bit u_msp và bit không kết thúc, và phần ít có nghĩa nhất nhất được mã hóa nhị phân.

Do các số nguyên lớn tùy ý có thể được mã hóa, một số hiệu quả mã hóa có thể bị mất khi các giá trị thực được mã hóa có kích thước bằng chữ cái đã biết và tương đối nhỏ. Một nhược điểm khác là khả năng giải mã giá trị ngoài phạm vi hoặc không hợp lệ hoặc khả năng đọc một số lượng bit rất lớn, trong trường hợp có lỗi truyền hoặc dòng bit không hợp lệ được tạo ra có chủ ý.

Phương pháp Golomb-Rice mở rộng kết hợp ba cải tiến so với phương pháp Golomb-Rice hiện có, để mã hóa vector của các giá trị, mỗi giá trị có kích thước bằng chữ cái đã biết và khác nhau về thế vị u_alph . Trước tiên, kích thước bằng chữ cái của phần có nghĩa nhất có thể được tính là $u_msp_alph = \lceil u_alph + 2^p \rceil$. Nếu giá trị tối đa có thể có của phần có nghĩa nhất được mã hóa, $u_msp_alph - 1$, bit không kết thúc có thể được loại bỏ, bởi vì điều kiện này có thể được phát hiện ngầm ở phía bộ giải mã, phép biến đổi là phương pháp Golomb-Rice có giới hạn hiện có. Ngoài ra, trong trường hợp tương tự khi $u_msp = u_msp_alph - 1$, kích thước bằng chữ cái của phần ít có nghĩa nhất u_lsp có thể được tính là $u_alph - (u_msp_alph - 1) \cdot 2^p$, có thể nhỏ hơn 2^p , cho phép sử dụng hàm *EncodeQuasiUniform* thay vì mã hóa nhị phân bằng p bit. Điều này cũng hữu ích khi giá trị cụ thể u có bằng chữ cái u_alph nhỏ hơn 2^p . Cuối cùng, khi $u_msp_alph \leq 3$ phương pháp Golomb-Rice có giới hạn tạo ra các mã chỉ có một độ dài là p hoặc $p + 1$ bit, hoặc các mã chỉ có hai độ dài là $p + 1$ và $p + 2$ bit. Hàm *EncodeQuasiUniform* là tối ưu cho tối đa hai độ dài, do đó, nó được sử dụng thay thế.

Ngưỡng 3 là giá trị ưu tiên đặc biệt, bởi vì khi $u_msp_alph = 3$, các từ mã của Golomb-Rice có giới hạn đối với phần có nghĩa nhất là 0, 10, 11; do đó, tổng độ dài của mã là $1+p$, $2+p$, và $2+p$, trong đó p là số bit cho phần ít có nghĩa nhất; bởi vì mã thùng luôn luôn tối ưu cho tối đa hai độ dài, nên nó được sử dụng thay thế, thay thế cả phần có nghĩa nhất và ít có nghĩa nhất.

Hơn nữa, cần phải phác thảo rằng hàm *EncodeQuasiUniform* chính xác là mã thùng, nó hoàn toàn trở thành mã nhị phân khi kích thước bằng chữ cái là lũy thừa của

hai. Nói chung, mã thùng là tối ưu và được xác định duy nhất với kích thước bảng chữ cái; nó tạo ra các mã chỉ có một hoặc hai độ dài; đối với 3 độ dài mã liên tiếp trở lên, các mã có thể có không còn đồng nhất nữa và có nhiều lựa chọn khác nhau cho số lượng mã có thể có của mỗi độ dài.

Sáng chế này không giới hạn ở mô tả chính xác ở trên. Ngoài ra, sáng chế có thể dễ dàng được mở rộng dưới dạng sơ đồ mã hóa dự đoán liên khung, trong đó với mỗi dải tham số, vectơ hướng trung bình được tính bằng cách sử dụng các vectơ hướng trước đó theo thời gian, từ khung hiện thời và cũng tùy ý từ các khung trước đó, hơn là tính toán một vectơ hướng trung bình duy nhất cho toàn bộ khung hiện thời và lượng tử hóa và mã hóa nó thành thông tin phụ. Giải pháp này sẽ có lợi thế là hiệu quả hơn trong mã hóa nhưng cũng chống lại mạnh mẽ hơn việc mất gói có thể có.

Các hình vẽ từ Fig.6a đến Fig.6g minh họa các quy trình tiếp theo được thực hiện trong bộ mã hóa như đã thảo luận trước đây. Fig.6a minh họa tổng quan chung về bộ lượng tử tham số 210 bao gồm hàm góc nâng lượng tử hóa 210a, hàm góc phương vị lượng tử hóa 210b và hàm góc nâng khử lượng tử hóa 210c. Phương án được ưu tiên theo Fig.6a minh họa bộ lượng tử hóa tham số có hàm góc phương vị 210c dựa trên giá trị góc nâng được lượng tử hóa và một lần nữa được khử lượng tử hóa q_{el} .

Fig.6c minh họa bộ khử lượng tử hóa tương ứng để khử lượng tử hóa góc nâng như đã được thảo luận trước đây đối với Fig.6a cho bộ mã hóa. Tuy nhiên, phương án của Fig.6b cũng hữu ích cho bộ khử lượng tử hóa được minh họa trong mục 840 của Fig.8a. Dựa trên độ chính xác khử lượng tử hóa deg_{req} , chỉ số góc nâng trên một mặt và chỉ số góc phương vị trên mặt khác được khử lượng tử hóa để cuối cùng thu được giá trị góc nâng đã khử lượng tử hóa q_{el} và giá trị góc phương vị được khử lượng tử hóa q_{az} . Fig.6c minh họa chế độ mã hóa thứ nhất, tức là chế độ mã hóa thô như được thảo luận liên quan đến các mục từ 260 đến 262 trên Fig.5c. Fig.6c cũng minh họa thêm quá trình tiền xử lý được thảo luận trên Fig.5b thể hiện việc chuyển đổi dữ liệu góc nâng thành các giá trị có dấu ở 247a và chuyển đổi tương ứng dữ liệu góc phương vị thành các giá trị có dấu ở 247b. Việc sắp xếp lại được thực hiện cho góc nâng như được biểu thị ở 248a và cho góc phương vị như được biểu thị ở 248b. Quy trình đếm điểm hình cầu được thực hiện trong khối 248c để tính toán bảng chữ cái hình cầu dựa

trên độ chính xác lượng tử hóa hoặc khử lượng tử hóa. Trong khối 261, việc gộp cả hai chỉ số thành một chỉ số hình cầu duy nhất được thực hiện và, phép mã hóa trong khối 262 được thực hiện bằng mã nhị phân hoặc mã nhị phân thung, trong đó, ngoài chỉ số hình cầu này, còn có bảng chữ cái hình cầu cho độ chính xác khử lượng tử hóa tương ứng được suy ra như cũng được minh họa trên Fig.5c.

Fig.6d minh họa quy trình được thực hiện cho chế độ mã hóa entropy với mô hình hóa. Trong mục 267a, việc khử lượng tử hóa dữ liệu góc phương vị và dữ liệu góc nâng được thực hiện dựa trên các chỉ số tương ứng và độ chính xác khử lượng tử hóa. Các giá trị được khử lượng tử hóa được đưa vào khối 267b để tính toán vector hướng từ các giá trị được khử lượng tử hóa. Trong khối 267c, việc tính trung bình được thực hiện cho các vector có chỉ số độ khuếch tán liên quan dưới ngưỡng tương ứng để thu được vector trung bình. Trong khối 267d, vector hướng trung bình hướng một lần nữa được chuyển đổi thành trung bình góc nâng và trung bình góc phương vị, và các giá trị này sau đó được lượng tử hóa bằng cách sử dụng độ chính xác cao nhất được xác định bởi khối 268e. Phép lượng tử hóa này được minh họa ở 268a, 268b và các kết quả lượng tử hóa trong các chỉ số lượng tử hóa và các bảng chữ cái lượng tử hóa tương ứng trong đó các bảng chữ cái được xác định bằng độ chính xác lượng tử hóa cho giá trị trung bình. Trong các khối 268c và 268d, một lần nữa phép khử lượng tử hóa được thực hiện để thu được giá trị trung bình được khử lượng tử hóa cho góc nâng và góc phương vị.

Trên Fig.6e, trung bình góc nâng được chiếu được tính trong khối 269a và trung bình góc phương vị được chiếu được tính trong khối 269b, tức là, Fig.6e minh họa triển khai được ưu tiên của khối 269 của Fig.5d. Như được minh họa trên Fig.6e, các khối 269a, 269b tốt hơn là nhận các giá trị trung bình được lượng tử hóa và một lần nữa được khử lượng tử hóa cho góc nâng và góc phương vị. Ngoài ra, phép chiếu cũng có thể được thực hiện trực tiếp trên đầu ra của khối 267d, mặc dù quy trình với lượng tử hóa và một lần nữa khử lượng tử hóa được ưu tiên cho độ chính xác cao hơn và khả năng tương thích cao hơn với các trạng thái ở phía bộ mã hóa và ở phía bộ giải mã.

Trên Fig.6f, quy trình được minh họa tương ứng với khối 270 của Fig.5d trong phương án được ưu tiên. Trong các khối 278a, 278b, các chênh lệch tương ứng hay

"các khoảng cách" như được gọi trong khối 270 của Fig.5d được tính giữa các chỉ số ban đầu và các chỉ số được chiếu. Việc giảm khoảng tương ứng được thực hiện trong các khối 270c cho góc nâng và 270d cho dữ liệu góc phương vị. Sau đó sắp xếp lại trong khối 270e, 270f, thu được dữ liệu tuân theo mã hóa Golomb-Rice mở rộng như đã thảo luận trước đó đối với các hình vẽ từ Fig.5e đến Fig.5g.

Fig.6g minh họa chi tiết thêm về quy trình được thực hiện để tạo các bit được mã hóa cho trung bình góc nâng và trung bình góc phương vị. Khối 271a và 271b minh họa việc chuyển đổi dữ liệu trung bình góc nâng và góc phương vị thành dữ liệu có dấu và hàm ReorderGeneric sau đó được minh họa đối với cả hai loại dữ liệu trong khối 272a và 272b. Các mục 273a và 273b minh họa việc mã hóa dữ liệu này bằng cách sử dụng mã nhị phân (thùng), chẳng hạn như hàm giả đều mã hóa được thảo luận ở trên.

Fig.7a minh họa bộ giải mã theo khía cạnh thứ nhất để giải mã tín hiệu âm thanh được mã hóa bao gồm các tham số mã hóa âm thanh có hướng được mã hóa, các tham số mã hóa âm thanh có hướng được mã hóa bao gồm các tham số độ khuếch tán được mã hóa và các tham số hướng được mã hóa. Thiết bị bao gồm bộ xử lý tham số 300 để giải mã các tham số mã hóa âm thanh có hướng được mã hóa để thu được các tham số độ khuếch tán được giải mã với độ phân giải thời gian hoặc tần số thứ nhất và các tham số hướng được giải mã với độ phân giải thời gian hoặc tần số thứ hai. Bộ xử lý tham số 300 được kết nối với bộ chuyển đổi độ phân giải tham số 710 để chuyển đổi các tham số độ khuếch tán được giải mã hoặc các tham số hướng được giải mã thành các tham số độ khuếch tán được chuyển đổi hoặc các tham số hướng được chuyển đổi. Ngoài ra, như được minh họa bởi đường nét đứt, bộ chuyển đổi độ phân giải tham số 710 có thể đã thực hiện xử lý độ phân giải tham số với dữ liệu theo tham số được mã hóa và các tham số được mã hóa được chuyển đổi được gửi từ bộ chuyển đổi độ phân giải tham số 710 đến bộ xử lý tham số 300. Trong trường hợp ở sau, bộ xử lý tham số 300 sau đó cung cấp các tham số được xử lý, tức là được giải mã trực tiếp vào bộ kết xuất âm thanh 420. Tuy nhiên, nên thực hiện chuyển đổi độ phân giải tham số với các tham số độ khuếch tán được giải mã và các tham số hướng được giải mã.

Các tham số hướng và độ khuếch tán được giải mã thường có độ phân giải tần số hoặc thời gian thứ ba hoặc thứ tư khi chúng được cung cấp cho bộ kết xuất âm thanh 420, trong đó độ phân giải thứ ba hoặc thứ tư lớn hơn độ phân giải vốn có của các tham số này khi chúng được xuất ra bởi bộ xử lý tham số 300.

Bộ chuyển đổi độ phân giải tham số 710 được tạo cấu hình để thực hiện chuyển đổi độ phân giải tham số khác với các tham số độ khuếch tán được giải mã và các tham số hướng được giải mã, do độ phân giải thời gian hoặc tần số vốn có của các tham số độ khuếch tán được giải mã và các tham số hướng được giải mã khác nhau, và thông thường, các tham số độ khuếch tán được giải mã có độ phân giải thời gian hoặc tần số thấp hơn so với các tham số hướng được giải mã. Như đã thảo luận trước đó về các hình vẽ từ Fig.3a đến Fig.3c, độ phân giải cao nhất được sử dụng bởi bộ kết xuất âm thanh 420 là độ phân giải được minh họa trên Fig.3b và độ phân giải trung gian như được minh họa trên Fig.3c là độ phân giải vốn có của các tham số hướng được giải mã và độ phân giải thấp vốn có của các tham số độ khuếch tán được giải mã là độ phân giải được minh họa trên Fig.3b.

Các hình vẽ từ Fig.3a đến 3c chỉ là các ví dụ minh họa ba độ phân giải thời gian hoặc tần số rất cụ thể. Bất kỳ độ phân giải thời gian hoặc tần số nào khác có cùng xu hướng là có độ phân giải thời gian hoặc tần số cao, độ phân giải trung bình và độ phân giải thấp cũng có thể được áp dụng theo sáng chế. Độ phân giải thời gian hoặc tần số thấp hơn một độ phân giải thời gian hoặc tần số khác khi cả hai độ phân giải này có cùng độ phân giải tần số nhưng độ phân giải thời gian khác nhau, hoặc ngược lại, như đã được minh họa trong ví dụ của Fig.3b và Fig.3c. Trong ví dụ này, độ phân giải tần số giống như trên Fig.3b và Fig.3c, nhưng độ phân giải thời gian cao hơn trên Fig.3c sao cho Fig.3c minh họa độ phân giải trung bình trong khi Fig.3b minh họa độ phân giải thấp.

Kết quả của bộ kết xuất âm thanh 420 hoạt động ở độ phân giải thời gian hoặc tần số cao thứ ba hoặc thứ tư sau đó được chuyển đến bộ chuyển đổi phổ/thời gian 440 mà sau đó tạo ra tín hiệu âm thanh đa kênh miền thời gian 450 như đã thảo luận trước đó về Fig.1b. Bộ chuyển đổi phổ/thời gian 440 chuyển đổi dữ liệu từ miền phổ được tạo bởi bộ kết xuất âm thanh 420 thành miền thời gian trên đường 450. Miền phổ mà

trong đó bộ kết xuất âm thanh 420 hoạt động bao gồm, đối với khung, số khe thời gian thứ nhất và số dải tần thứ hai. Khung bao gồm số ngăn thời gian/tần số bằng với kết quả phép nhân của số thứ nhất và số thứ hai, trong đó số thứ nhất và số thứ hai xác định độ phân giải thời gian hoặc tần số thứ ba, tức là độ phân giải thời gian hoặc tần số cao.

Bộ chuyển đổi độ phân giải 710 được tạo cấu hình để tạo, từ tham số độ khuếch tán liên quan đến độ phân giải thời gian hoặc tần số thứ nhất, một số ít nhất bốn tham số độ khuếch tán, trong đó hai trong số các tham số độ khuếch tán này dành cho các ngăn thời gian/tần số liền kề nhau theo thời gian và hai tham số còn lại trong số ít nhất bốn tham số độ khuếch tán này dành cho các ngăn thời gian/tần số liền kề nhau theo tần số.

Do độ phân giải thời gian hoặc tần số cho các tham số độ khuếch tán thấp hơn so với tham số hướng, bộ chuyển đổi độ phân giải tham số được tạo cấu hình để tạo, cho tham số độ khuếch tán được giải mã, tập hợp các tham số độ khuếch tán được chuyển đổi và cho tham số hướng được giải mã, tập hợp các tham số hướng được chuyển đổi thứ hai, trong đó tập hợp thứ hai cao hơn tập hợp thứ nhất.

Fig.7b minh họa quy trình được ưu tiên được thực hiện bởi bộ chuyển đổi độ phân giải tham số. Trong khối 721, bộ chuyển đổi độ phân giải tham số 710 thu được các tham số độ khuếch tán/hướng cho khung. Trong khối 722, phép nhân các tham số độ khuếch tán hoặc thao tác sao chép vào ít nhất bốn ngăn thời gian/tần số có độ phân giải cao được thực hiện. Trong khối 723, quá trình xử lý tùy chọn như làm mịn hoặc lọc thông thấp được thực hiện cho các tham số được nhân trong sự biểu diễn độ phân giải cao. Trong khối 724, các tham số độ phân giải cao được áp dụng cho dữ liệu âm thanh tương ứng trong các ngăn thời gian/tần số độ phân giải cao tương ứng.

Fig.8a minh họa việc triển khai được ưu tiên của bộ giải mã để giải mã tín hiệu âm thanh được mã hóa bao gồm các tham số mã hóa âm thanh có hướng được mã hóa gồm các tham số độ khuếch tán được mã hóa và các tham số hướng được mã hóa theo khía cạnh thứ nhất. Tín hiệu âm thanh được mã hóa được đưa vào giao diện đầu vào. Giao diện đầu vào 800 nhận tín hiệu âm thanh được mã hóa và tách ra, từ tín hiệu âm thanh được mã hóa, các tham số độ khuếch tán được mã hóa và các tham số hướng

được mã hóa, thường ở trong khung theo cách đóng khung. Dữ liệu này được nhập vào bộ giải mã tham số 820 mà tạo ra, từ các tham số được mã hóa, các tham số độ khuếch tán được lượng tử hóa và các tham số hướng được lượng tử hóa trong đó các tham số hướng được lượng tử hóa ví dụ là các chỉ số góc phương vị và các chỉ số góc nâng. Dữ liệu này được nhập vào bộ khử lượng tử hóa tham số 840 để xác định, từ các tham số độ khuếch tán được lượng tử hóa và các tham số hướng được lượng tử hóa, các tham số độ khuếch tán được khử lượng tử hóa và các tham số hướng được khử lượng tử hóa. Dữ liệu này sau đó có thể được sử dụng để chuyển đổi một định dạng âm thanh thành một định dạng âm thanh khác hoặc có thể được sử dụng để kết xuất tín hiệu âm thanh thành tín hiệu đa kênh hoặc trong bất kỳ sự biểu diễn nào khác chẳng hạn như sự biểu diễn Ambisonics, sự biểu diễn MPS hoặc sự biểu diễn SAOC.

Đầu ra các tham số được khử lượng tử hóa theo khối 840 có thể được nhập vào bộ chuyển đổi độ phân giải tham số tùy chọn như đã thảo luận trước đó đối với Fig.7a ở khối 710. Các tham số được chuyển đổi hoặc không được chuyển đổi có thể được nhập vào bộ kết xuất âm thanh 420, 440 được minh họa trên Fig.8a. Khi tín hiệu âm thanh được mã hóa bao gồm thêm tín hiệu truyền được mã hóa, giao diện đầu vào 800 được tạo cấu hình để tách tín hiệu truyền được mã hóa ra khỏi tín hiệu âm thanh được mã hóa và đưa dữ liệu này vào bộ giải mã tín hiệu truyền âm thanh 340 đã được thảo luận trước đó đối với Fig.8b. Kết quả là đầu vào vào bộ chuyển đổi thời gian-phổ 430 cấp cho bộ kết xuất âm thanh 420. Khi bộ kết xuất âm thanh 420 được triển khai như được minh họa trên Fig.1b, việc chuyển đổi thành miền thời gian được thực hiện bằng cách sử dụng giàn lọc tổng hợp 440 của Fig.1b.

Fig.8b minh họa một phần của tín hiệu âm thanh được mã hóa thường được tổ chức theo dòng bit liên quan đến các tham số độ khuếch tán được mã hóa. Các tham số độ khuếch tán có liên quan đến tốt hơn là hai bit chế độ 802 để biểu thị ba chế độ khác nhau được minh họa trên Fig.8b và đã thảo luận trước đó. Dữ liệu được mã hóa cho các tham số độ khuếch tán bao gồm dữ liệu tải trọng 804.

Các phần dòng bit cho các tham số hướng được minh họa trên Fig.8c và Fig.8d như đã thảo luận trước đó, trong đó Fig.8c minh họa tình huống khi chế độ mã hóa thô

đã được chọn và Fig.8d minh họa tình huống trong đó chế độ giải mã entropy với mô hình hóa đã được chọn/chỉ định bởi bit chế độ hoặc cờ chế độ 806.

Bộ giải mã tham số 820 của Fig.8a được tạo cấu hình để giải mã dữ liệu tải trọng khuếch tán cho vùng thời gian/tần số như được biểu thị trong khối 850 và vùng thời gian/tần số là vùng thời gian/tần số với độ phân giải thấp trong phương án được ưu tiên. Trong khối 851, độ chính xác khử lượng tử hóa cho vùng thời gian/tần số được xác định. Dựa trên độ chính xác khử lượng tử hóa này, khối 852 của Fig.8e minh họa phép giải mã và/hoặc khử lượng tử hóa của các tham số hướng sử dụng độ chính xác khử lượng tử hóa tương tự cho vùng thời gian/tần số mà tham số độ khuếch tán có liên quan. Đầu ra của Fig.8e là tập hợp các tham số hướng được giải mã cho vùng thời gian/tần số, chẳng hạn như cho một dải của Fig.3c, tức là, trong ví dụ minh họa, bốn tham số hướng cho một dải trong khung.

Fig.8f minh họa tính năng khác nữa của bộ giải mã, và đặc biệt là bộ giải mã tham số 820 và bộ khử lượng tử hóa tham số 840 của Fig.8a. Bất kể liệu độ chính xác khử lượng tử hóa được xác định dựa trên tham số độ khuếch tán hoặc được báo hiệu rõ ràng hoặc được xác định ở bất kỳ đâu, khối 852a biểu thị việc xác định bằng chữ cái góc nâng từ độ chính xác khử lượng tử hóa được báo hiệu cho vùng thời gian/tần số. Trong khối 852b, dữ liệu góc nâng được giải mã và khử lượng tử hóa một cách tùy chọn bằng cách sử dụng bảng chữ cái góc nâng cho vùng thời gian/tần số để thu được, ở đầu ra của khối 852b, các tham số góc nâng được khử lượng tử hóa. Trong khối 852c, bảng chữ cái góc phương vị cho vùng thời gian/tần số được xác định không chỉ từ độ chính xác khử lượng tử hóa từ khối 851, mà còn từ dữ liệu góc nâng được lượng tử hóa hoặc được khử lượng tử hóa cũng như để phản ánh tình huống đã được thảo luận trước đó đối với độ bao phủ giả đều của khối cầu đơn vị trên Fig.4d. Trong khối 852d, việc giải mã và khử lượng tử hóa một cách tùy chọn dữ liệu góc phương vị với bảng chữ cái góc phương vị được thực hiện cho vùng thời gian/tần số.

Sáng chế theo khía cạnh thứ hai tốt hơn là kết hợp hai tính năng đó, nhưng hai tính năng, tức là, một tính năng theo Fig.8a hoặc một tính năng khác theo Fig.8f cũng có thể được áp dụng riêng biệt với nhau.

Fig.8g minh họa tổng quan về phép giải mã tham số tùy thuộc vào việc chọn chế độ giải mã thô hay chế độ giải mã với mô hình hóa như được chỉ định bởi bit chế độ 806 được thảo luận trên Fig.8c và Fig.8d. Khi áp dụng phép giải mã thô, thì các chỉ số hình cầu cho dải được giải mã như được biểu thị ở 862 và các tham số góc phương vị/góc nâng được lượng tử hóa cho dải được tính từ các chỉ số hình cầu được giải mã như được biểu thị ở khối 864.

Khi giải mã với mô hình hóa được chỉ định bởi bit chế độ 806, thì các giá trị trung bình của dữ liệu góc phương vị/góc nâng trong dải/khung được giải mã như được biểu thị bởi khối 866. Trong khối 868, các khoảng cách cho thông tin góc phương vị/góc nâng trong dải được giải mã, và trong khối 870, các tham số góc nâng và góc phương vị được lượng tử hóa tương ứng được tính thường bằng cách sử dụng thao tác bổ sung.

Không phụ thuộc vào việc chế độ giải mã thô hay chế độ giải mã với mô hình hóa đã được áp dụng, các chỉ số góc phương vị/góc nâng đã giải mã được khử lượng tử hóa 872 như cũng được minh họa ở 840 trên Fig.8a, và trong khối 874, và kết quả có thể được chuyển đổi thành các tọa độ Đề-các cho dải. Ngoài ra, khi dữ liệu góc phương vị và góc nâng có thể được sử dụng trực tiếp trong bộ kết xuất âm thanh, thì bất kỳ chuyển đổi nào như vậy trong khối 874 là không cần thiết. Bất kỳ chuyển đổi độ phân giải tham số có khả năng được sử dụng có thể được áp dụng trước hoặc sau khi chuyển đổi nếu việc chuyển đổi thành các tọa độ Đề-các được thực hiện bằng mọi cách.

Sau đây, tham chiếu đến các hình vẽ từ Fig.9a đến Fig.9c liên quan đến các triển khai được ưu tiên bổ sung của bộ giải mã. Fig.9a minh họa thao tác giải mã được minh họa trong khối 862. Tùy thuộc vào độ chính xác khử lượng tử hóa được xác định bởi khối 851 trên Fig.8e hoặc Fig.8f, việc đếm điểm hình cầu chức năng của khối 248c được thực hiện để xác định bảng chữ cái hình cầu thực tế cũng được áp dụng trong quá trình mã hóa. Các bit cho chỉ số hình cầu được giải mã trong khối 862 và việc phân tách thành hai chỉ số được thực hiện như được minh họa ở 864a và được nêu chi tiết hơn trên Fig.9a. Các hàm sắp xếp lại 864b, 864c và các hàm chuyển đổi tương ứng trong khối 864d và 864e được thực hiện để cuối cùng thu được các chỉ số góc nâng,

các chỉ số góc phương vị và các bảng chữ cái tương ứng cho việc khử lượng tử hóa tiếp theo trong khối 872 của Fig.8g.

Fig.9b minh họa các quy trình tương ứng cho chế độ giải mã khác, tức là chế độ giải mã với mô hình hóa. Trong khối 866a, độ chính xác khử lượng tử hóa cho các giá trị trung bình được tính toán phù hợp với những gì đã được thảo luận trước đó đối với phía bộ mã hóa. Các bảng chữ cái được tính trong khối 866b và trong các khối 866c và 866d, các bit tương ứng 808a, 808b của Fig.8d được giải mã. Các hàm sắp xếp lại 866e, 866f được thực hiện trong các thao tác chuyển đổi tiếp theo 866g, 866h để hoàn tác hoặc bắt chước các thao tác tương ứng được thực hiện ở phía bộ mã hóa.

Fig.9c minh họa thêm thao tác khử lượng tử hóa hoàn toàn 840 trong phương án được ưu tiên. Các khối 852a xác định bảng chữ cái góc nâng như đã được thảo luận liên quan đến Fig.8f và việc tính toán tương ứng của bảng chữ cái góc phương vị cũng được thực hiện trong khối 852c. Thao tác tính toán phép chiếu 820a, 820e cũng được thực hiện cho góc nâng và góc phương vị. Các quy trình sắp xếp lại cho góc nâng 820b và góc phương vị 820f cũng được thực hiện và các thao tác bổ sung tương ứng 820c, 820g cũng được thực hiện. Việc giảm khoảng tương ứng trong các khối 820d đối với góc nâng và 820h đối với góc phương vị cũng được thực hiện, và việc khử lượng tử hóa góc nâng được thực hiện trong khối 840a và 840b. Fig.9c thể hiện quy trình này có thứ tự nhất định, tức là dữ liệu góc nâng được xử lý trước tiên và dựa trên dữ liệu góc nâng được khử lượng tử hóa, việc giải mã và khử lượng tử hóa dữ liệu góc phương vị được thực hiện theo phương án được ưu tiên của sáng chế.

Sau đây, các lợi ích và lợi thế của các phương án được ưu tiên được tóm tắt:

- Mã hóa hiệu quả của siêu dữ liệu không gian được tạo bởi DirAC mà không ảnh hưởng đến tính tổng quát của mô hình. Nó là công cụ hỗ trợ chính để tích hợp DirAC vào sơ đồ mã hóa tốc độ bit thấp.
- Nhóm và tính trung bình các tham số hướng và độ khuếch tán với các độ phân giải thời gian (hoặc tần số tùy chọn) khác nhau: độ khuếch tán được tính trung bình trong thời gian dài hơn hướng, vì độ khuếch tán giữ đặc tính dài hạn của trường âm thanh hơn là hướng, mà là tín hiệu không gian phản ứng lại hơn.

- Độ bao phủ động giả đều của hình cầu 3D, đối xứng hoàn toàn với các trục tọa độ X, Y, và Z, và bất kỳ độ phân giải góc mong muốn nào đều có thể có.
- Các thao tác lượng tử hóa và khử lượng tử hóa là độ phức tạp không đổi (không cần tìm kiếm vectơ mã gần nhất).
- Mã hóa và giải mã một chỉ số điểm được lượng tử hóa có độ phức tạp không đổi hoặc nhiều nhất là logarit đối với tổng số điểm được lượng tử hóa trên khối cầu.
- Kích thước mã hóa entropy của trường hợp không mong muốn nhất của toàn bộ siêu dữ liệu không gian DirAC cho một khung hình luôn bị giới hạn chỉ nhiều hơn 2 bit so với mã hóa thô.
- Phương pháp mã hóa Golomb-Rice mở rộng là tối ưu để mã hóa vectơ các ký hiệu với các kích thước bảng chữ cái có khả năng khác nhau.
- Sử dụng hướng trung bình để mã hóa entropy hiệu quả của các hướng, ánh xạ hướng trung bình được lượng tử hóa từ độ phân giải cao nhất tới độ phân giải của từng góc phương vị và góc nâng.
- Luôn sử dụng mã hóa thô cho các hướng có độ khuếch tán cao, trên ngưỡng được xác định trước, cho các khung độ khuếch tán hỗn hợp.
- Sử dụng độ phân giải góc cho mỗi hướng như là một hàm của độ khuếch tán tương ứng của nó.

Khía cạnh thứ nhất của sáng chế hướng đến việc xử lý các tham số độ khuếch tán và các tham số hướng với các độ phân giải thời gian hoặc tần số thứ nhất và thứ hai và sau đó lượng tử hóa và mã hóa các giá trị đó. Khía cạnh thứ nhất này cũng đề cập đến việc nhóm các tham số với các độ phân giải thời gian/tần số khác nhau. Khía cạnh khác nữa có liên quan đến việc thực hiện gán trọng số liên quan đến đo biên độ trong việc nhóm và khía cạnh bổ sung khác liên quan đến việc gán trọng số cho việc tính trung bình và việc nhóm các tham số hướng sử dụng các tham số độ khuếch tán tương ứng làm cơ sở cho các trọng số tương ứng. Các khía cạnh trên cũng được mô tả và xây dựng trong tập hợp yêu cầu bảo hộ thứ nhất.

Khía cạnh thứ hai của sáng chế sau đây được xây dựng nhiều hơn trong tập hợp các ví dụ kèm theo là hướng đến thực hiện lượng tử hóa và mã hóa. Khía cạnh này có thể được thực hiện mà không có các tính năng được nêu trong khía cạnh thứ nhất hoặc

có thể được sử dụng cùng với các tính năng tương ứng được xây dựng ở khía cạnh thứ nhất.

Do đó, tất cả các khía cạnh khác nhau như được xây dựng trong các điểm yêu cầu bảo hộ và tập hợp các ví dụ và như được xây dựng trong các điểm yêu cầu bảo hộ phụ thuộc khác nhau của các yêu cầu và các ví dụ có thể được sử dụng độc lập với nhau hoặc có thể được sử dụng cùng nhau và đặc biệt ưu tiên cho phương án được ưu tiên nhất là tất cả các khía cạnh của tập hợp các điểm yêu cầu bảo hộ được sử dụng cùng với tất cả các khía cạnh của tập hợp các ví dụ.

Tập hợp các ví dụ bao gồm các ví dụ sau:

1. Thiết bị mã hóa các tham số mã hóa âm thanh có hướng bao gồm các tham số độ khuếch tán và các tham số hướng, bao gồm:

bộ lượng tử hóa tham số (210) để lượng tử hóa các tham số độ khuếch tán và các tham số hướng;

bộ mã hóa tham số (220) để mã hóa các tham số độ khuếch tán được lượng tử hóa và các tham số hướng được lượng tử hóa; và

giao diện đầu ra (230) để tạo sự biểu diễn tham số được mã hóa bao gồm thông tin về các tham số độ khuếch tán được mã hóa và các tham số hướng được mã hóa.

2. Thiết bị theo ví dụ 1,

trong đó bộ lượng tử hóa tham số (210) được tạo cấu hình để lượng tử hóa các tham số độ khuếch tán bằng cách sử dụng bộ lượng tử hóa không đồng nhất để tạo ra các chỉ số độ khuếch tán.

3. Thiết bị theo ví dụ 2,

trong đó bộ lượng tử hóa tham số (210) được tạo cấu hình để suy ra bộ lượng tử hóa không đồng nhất bằng cách sử dụng bảng lượng tử hóa kết hợp liên kênh để thu được các ngưỡng và các mức khôi phục của bộ lượng tử hóa không đồng nhất.

4. Thiết bị theo một trong số các ví dụ từ 1 đến 3,

trong đó bộ mã hóa tham số (220) được tạo cấu hình:

để mã hóa các tham số độ khuếch tán được lượng tử hóa trong chế độ mã hóa thô bằng cách sử dụng mã nhị phân nếu bảng chữ cái mã hóa có kích thước bằng lũy thừa của hai, hoặc

để mã hóa các tham số độ khuếch tán được lượng tử hóa trong chế độ mã hóa thô bằng cách sử dụng mã thùng, nếu bảng chữ cái mã hóa khác lũy thừa của hai, hoặc

để mã hóa các tham số độ khuếch tán được lượng tử hóa trong chế độ chỉ một giá trị bằng cách sử dụng số chỉ cụ thể thứ nhất và từ mã cho một giá trị từ chế độ mã hóa thô, hoặc

để mã hóa các tham số độ khuếch tán được lượng tử hóa trong chế độ chỉ hai giá trị liên tiếp bằng cách sử dụng số chỉ cụ thể thứ hai, mã cho giá trị nhỏ hơn trong số hai giá trị liên tiếp và bit cho chênh lệch giữa một hoặc mỗi giá trị thực và giá trị nhỏ hơn trong số hai giá trị liên tiếp.

5. Thiết bị theo ví dụ 4,

trong đó bộ mã hóa tham số (220) được tạo cấu hình để xác định, đối với tất cả các giá trị độ khuếch tán liên quan đến phần thời gian hoặc phần tần số, dù chế độ mã hóa là chế độ mã hóa thô, chế độ chỉ một giá trị hoặc chế độ chỉ hai giá trị liên tiếp,

trong đó chế độ thô được báo hiệu bằng cách sử dụng một trong hai bit, trong đó chế độ chỉ một giá trị được báo hiệu bằng cách sử dụng một bit khác trong số hai bit có giá trị thứ nhất, và trong đó chế độ chỉ hai giá trị liên tiếp được báo hiệu bằng cách sử dụng một bit khác trong số hai bit có giá trị thứ hai.

6. Thiết bị theo một trong các ví dụ trên, trong đó bộ lượng tử hóa tham số (210) được tạo cấu hình

để nhận, đối với mỗi tham số hướng, vectơ Đề-các có hai hoặc ba thành phần, và

để chuyển đổi vectơ Đề-các thành sự biểu diễn có giá trị góc phương vị và giá trị góc nâng.

7. Thiết bị theo một trong các ví dụ trên,

trong đó bộ lượng tử hóa tham số (210) được tạo cấu hình để xác định, đối với phép lượng tử hóa của tham số hướng, độ chính xác lượng tử hóa, độ chính xác lượng tử hóa phụ thuộc vào tham số độ khuếch tán liên quan đến tham số hướng, sao cho tham số hướng liên quan đến tham số độ khuếch tán thấp hơn được lượng tử hóa chính xác hơn tham số hướng liên quan đến tham số độ khuếch tán cao hơn.

8. Thiết bị theo ví dụ 7,

trong đó bộ lượng tử hóa tham số (210) được tạo cấu hình để xác định độ chính xác lượng tử hóa

sao cho các điểm được lượng tử hóa được phân bố giả đều trên mặt cầu đơn vị, hoặc

sao cho các điểm được lượng tử hóa được phân bố đối xứng với trục x, trục y hoặc trục z, hoặc

sự lượng tử hóa của hướng nhất định đến điểm lượng tử hóa gần nhất hoặc một trong một số điểm lượng tử hóa gần nhất bằng cách ánh xạ tới chỉ số nguyên là thao tác thời gian không đổi, hoặc

sao cho việc tính điểm tương ứng trên mặt cầu từ chỉ số nguyên và việc khử lượng tử hóa đến hướng là thao tác thời gian không đổi hoặc logarit đối với tổng số điểm trên mặt cầu.

9. Thiết bị theo một trong số các ví dụ 6, 7 hoặc 8,

trong đó bộ lượng tử hóa tham số (210) được tạo cấu hình để lượng tử hóa góc nâng có các giá trị âm và dương cho tập hợp các chỉ số lượng tử hóa không dấu, trong đó nhóm các chỉ số lượng tử hóa thứ nhất biểu thị các góc nâng âm và nhóm các chỉ số lượng tử hóa thứ hai biểu thị các góc nâng dương.

10. Thiết bị theo một trong các ví dụ trên,

trong đó bộ lượng tử hóa tham số (210) được tạo cấu hình để lượng tử hóa góc phương vị bằng cách sử dụng số các chỉ số lượng tử hóa có thể có, trong đó số các chỉ số lượng tử hóa giảm từ các góc nâng thấp hơn tới các góc nâng cao hơn sao cho số các chỉ số lượng tử hóa có thể có thứ nhất cho góc nâng thứ nhất có độ lớn thứ nhất

cao hơn số các chỉ số lượng tử hóa có thể có thứ hai cho góc nâng thứ hai có độ lớn thứ hai, độ lớn thứ hai có giá trị tuyệt đối lớn hơn độ lớn thứ nhất.

11. Thiết bị theo ví dụ 10, trong đó bộ lượng tử hóa tham số (210) được tạo cấu hình

để xác định từ giá trị độ khuếch tán liên quan đến góc phương vị độ chính xác cần thiết,

để lượng tử hóa góc nâng liên quan đến góc phương vị bằng cách sử dụng độ chính xác cần thiết, và

để lượng tử hóa góc phương vị bằng cách sử dụng góc nâng được lượng tử hóa.

12. Thiết bị theo một trong các ví dụ trên,

trong đó tham số hướng được lượng tử hóa có góc nâng được lượng tử hóa và góc phương vị được lượng tử hóa, và trong đó bộ mã hóa tham số (220) được tạo cấu hình để trước tiên mã hóa góc nâng được lượng tử hóa và sau đó mã hóa góc phương vị được lượng tử hóa.

13. Thiết bị theo một trong các ví dụ trên,

trong đó các tham số hướng được lượng tử hóa bao gồm các chỉ số không dấu cho cặp góc phương vị và góc nâng,

trong đó bộ mã hóa tham số (220) được tạo cấu hình để chuyển đổi các chỉ số không dấu thành các chỉ số có dấu, sao cho chỉ số cho biết góc không nằm ở giữa khoảng có dấu của các giá trị có thể có và,

trong đó bộ mã hóa tham số (220) được tạo cấu hình để thực hiện chuyển đổi sắp xếp lại thành các chỉ số có dấu để xen kẽ các số dương và âm thành các số không dấu.

14. Thiết bị theo một trong các ví dụ trên,

trong đó các tham số hướng được lượng tử hóa bao gồm các chỉ số góc phương vị và góc nâng không dấu không được sắp xếp lại, và

trong đó bộ mã hóa tham số (220) được tạo cấu hình để gộp các chỉ số của cặp thành chỉ số hình cầu, và

để thực hiện mã hóa thô chỉ số hình cầu.

15. Thiết bị theo ví dụ 14,

trong đó bộ mã hóa tham số (220) được tạo cấu hình để suy ra chỉ số hình cầu từ phần bù hình cầu và chỉ số góc phương vị được sắp xếp lại hoặc không được sắp xếp lại hiện thời, và

trong đó phần bù hình cầu được suy ra từ tổng của các bảng chữ cái góc phương vị tương ứng với các chỉ số góc nâng được sắp xếp lại hoặc không được sắp xếp lại nhỏ hơn chỉ số góc nâng được sắp xếp lại hoặc không được sắp xếp lại hiện thời.

16. Thiết bị của một trong các ví dụ trên, trong đó bộ mã hóa tham số (220) được tạo cấu hình để thực hiện mã hóa entropy cho các tham số hướng được lượng tử hóa liên quan đến các giá trị độ khuếch tán thấp hơn hoặc bằng ngưỡng và để thực hiện mã hóa thô cho các tham số hướng được lượng tử hóa liên quan đến các giá trị độ khuếch tán lớn hơn ngưỡng.

17. Thiết bị theo ví dụ 16,

trong đó bộ mã hóa tham số (220) được tạo cấu hình để xác định ngưỡng động bằng cách sử dụng bảng chữ cái lượng tử hóa và sự lượng tử hóa các tham số độ khuếch tán, hoặc trong đó bộ mã hóa tham số (220) được tạo cấu hình để xác định ngưỡng dựa trên bảng chữ cái lượng tử hóa của các tham số độ khuếch tán.

18. Thiết bị theo một trong các ví dụ trên,

trong đó bộ lượng tử hóa tham số (210) được tạo cấu hình để xác định, dưới dạng các tham số hướng được lượng tử hóa, các chỉ số góc nâng, các bảng chữ cái góc nâng liên quan đến các chỉ số góc nâng, các chỉ số góc phương vị và các bảng chữ cái góc phương vị liên quan đến các chỉ số góc phương vị,

trong đó bộ mã hóa tham số (220) được tạo cấu hình để

suy ra vector hướng trung bình từ các vector hướng được lượng tử hóa cho phần thời gian hoặc phần tần số của tín hiệu đầu vào,

định lượng vector hướng trung bình bằng cách sử dụng độ chính xác góc tốt nhất của các vector cho phần thời gian hoặc phần tần số, và

mã hóa vector hướng trung bình được lượng tử hóa, hoặc

trong đó giao diện đầu ra (230) được tạo cấu hình để nhập vector hướng trung bình được mã hóa vào sự biểu diễn tham số được mã hóa dưới dạng thông tin phụ bổ sung.

19. Thiết bị theo ví dụ 18,

trong đó bộ mã hóa tham số (220) được tạo cấu hình

để tính toán các chỉ số góc nâng dự đoán và các chỉ số góc phương vị dự đoán bằng cách sử dụng vector hướng trung bình, và

để tính các khoảng cách có dấu giữa các chỉ số góc nâng và các chỉ số góc nâng dự đoán, và giữa các chỉ số góc phương vị và các chỉ số góc phương vị dự đoán.

20. Thiết bị theo ví dụ 19,

trong đó bộ mã hóa tham số (220) được tạo cấu hình để chuyển đổi các khoảng cách có dấu thành khoảng giảm bằng cách cộng giá trị cho các giá trị nhỏ và trừ đi giá trị cho các giá trị lớn.

21. Thiết bị theo một trong các ví dụ trên,

trong đó bộ mã hóa tham số (220) được tạo cấu hình để quyết định, liệu các tham số hướng được lượng tử hóa được mã hóa bởi chế độ mã hóa thô hay chế độ mã hóa entropy, và trong đó giao diện đầu ra (230) được tạo cấu hình để đưa chỉ báo tương ứng vào sự biểu diễn tham số được mã hóa.

22. Thiết bị theo một trong các ví dụ trên,

trong đó bộ mã hóa tham số (220) được tạo cấu hình để thực hiện mã hóa entropy bằng cách sử dụng phương pháp Golomb-Rice hoặc sự cải biên của phương pháp này.

23. Thiết bị theo một trong các ví dụ từ 18 đến 22, trong đó bộ mã hóa tham số (220) được tạo cấu hình để

chuyển đổi các thành phần của vector hướng trung bình thành sự biểu diễn có dấu sao cho giá trị không tương ứng nằm ở giữa khoảng có dấu của các giá trị có thể có,

thực hiện phép biến đổi sắp xếp lại các giá trị có dấu để xen kẽ các số dương và âm thành các số không dấu,

mã hóa kết quả bằng cách sử dụng hàm mã hóa để thu được các thành phần được mã hóa của vector hướng trung bình; và

mã hóa tham số Golomb-Rice bằng cách sử dụng kích thước bảng chữ cái tùy thuộc vào mức tối đa của các kích thước bảng chữ cái cho thành phần tương ứng của vector hướng.

24. Thiết bị theo một trong số các ví dụ từ 19 đến 23,

trong đó bộ mã hóa tham số (220) được tạo cấu hình để thực hiện phép biến đổi sắp xếp lại các khoảng cách có dấu hoặc các khoảng cách có dấu giảm tới các số dương và âm xen kẽ thành các số không dấu.

trong đó bộ mã hóa tham số (220) được tạo cấu hình để mã hóa các khoảng cách có dấu được sắp xếp lại hoặc các khoảng cách có dấu giảm được sắp xếp lại bằng cách sử dụng phương pháp Golomb-Rice hoặc sự cải biên của phương pháp này.

25. Thiết bị theo ví dụ 24, trong đó bộ mã hóa tham số (220) được tạo cấu hình để áp dụng phương pháp Golomb-Rice hoặc sự cải biên của phương pháp này bằng cách sử dụng

việc xác định phần có nghĩa nhất và phần ít có nghĩa nhất của giá trị sẽ được mã hóa;

việc tính toán bảng chữ cái cho phần có nghĩa nhất;

việc tính toán bảng chữ cái cho phần ít có nghĩa nhất; và

việc mã hóa phần có nghĩa nhất theo kiểu đơn phân bằng cách sử dụng bảng chữ cái cho phần có nghĩa nhất và mã hóa phần ít có nghĩa nhất theo kiểu nhị phân bằng cách sử dụng bảng chữ cái cho phần ít có nghĩa nhất.

26. Thiết bị theo một trong các ví dụ trên,

trong đó bộ mã hóa tham số (220) được tạo cấu hình để áp dụng phương pháp Golomb-Rice hoặc sự cải biên của phương pháp này, sử dụng việc xác định phần có

nghĩa nhất và phần ít có nghĩa nhất của giá trị sẽ được mã hóa; và việc tính toán bảng chữ cái cho phần có nghĩa nhất,

trong đó khi bảng chữ cái của phần có nghĩa nhất nhỏ hơn hoặc bằng giá trị được xác định trước, chẳng hạn như 3, phương pháp EncodeQuasiUniform được sử dụng để mã hóa toàn bộ giá trị, trong đó phương pháp EncodeQuasiUniform mẫu, như mã thùng, tạo ra các mã chỉ có một độ dài chỉ hoặc các mã chỉ có hai độ dài, hoặc

để mã hóa phần có nghĩa nhất trong chế độ mã hóa thô bằng cách sử dụng mã nhị phân nếu bảng chữ cái mã hóa có kích thước bằng lũy thừa của hai, hoặc mã hóa phần ít có nghĩa nhất trong chế độ mã hóa thô bằng cách sử dụng mã thùng, nếu bảng chữ cái mã hóa khác lũy thừa của hai.

27. Thiết bị theo một trong các ví dụ trên, còn bao gồm bộ tính tham số để tính toán các tham số độ khuếch tán với độ phân giải thời gian hoặc tần số thứ nhất và để tính toán các tham số hướng với độ phân giải thời gian hoặc tần số thứ hai như được xác định theo điểm bất kỳ trong số các điểm kèm theo từ 1 đến 15.

28. Phương pháp mã hóa các tham số mã hóa âm thanh có hướng bao gồm các tham số độ khuếch tán và các tham số hướng, phương pháp bao gồm các bước:

lượng tử hóa các tham số độ khuếch tán và các tham số hướng;

mã hóa các tham số độ khuếch tán được lượng tử hóa và các tham số hướng được lượng tử hóa; và

tạo sự biểu diễn tham số được mã hóa bao gồm thông tin về các tham số độ khuếch tán được mã hóa và các tham số hướng được mã hóa.

29. Bộ giải mã để giải mã tín hiệu âm thanh được mã hóa bao gồm các tham số mã hóa âm thanh có hướng được mã hóa bao gồm các tham số độ khuếch tán được mã hóa và các tham số hướng được mã hóa, bao gồm;

giao diện đầu vào (800) để nhận tín hiệu âm thanh được mã hóa và tách ra, từ tín hiệu âm thanh được mã hóa, các tham số độ khuếch tán được mã hóa và các tham số hướng được mã hóa;

bộ giải mã tham số (820) để giải mã các tham số độ khuếch tán được mã hóa và các tham số hướng được mã hóa để thu được các tham số độ khuếch tán được lượng tử hóa và các tham số hướng được lượng tử hóa; và

bộ khử lượng tử hóa tham số (840) để xác định, từ các tham số độ khuếch tán được lượng tử hóa và các tham số hướng được lượng tử hóa, các tham số độ khuếch tán được khử lượng tử hóa và các tham số hướng được khử lượng tử hóa.

30. Bộ giải mã theo ví dụ 29,

trong đó giao diện đầu vào (800) được tạo cấu hình để xác định, từ chỉ báo chế độ mã hóa (806) có trong tín hiệu âm thanh được mã hóa, liệu bộ giải mã tham số (820) sẽ sử dụng chế độ giải mã thứ nhất là chế độ giải mã thô hay chế độ giải mã thứ hai là chế độ giải mã với mô hình hóa và khác với chế độ giải mã thứ nhất, để giải mã các tham số hướng được mã hóa.

31. Bộ giải mã theo điểm 29 hoặc điểm 30,

trong đó bộ giải mã tham số (820) được tạo cấu hình để giải mã tham số độ khuếch tán được mã hóa (804) cho khung của tín hiệu âm thanh được mã hóa để thu được tham số độ khuếch tán được lượng tử hóa cho khung,

trong đó bộ khử lượng tử hóa (840) được tạo cấu hình để xác định độ chính xác khử lượng tử hóa cho sự khử lượng tử hóa ít nhất một tham số hướng cho khung bằng cách sử dụng tham số độ khuếch tán được lượng tử hóa hoặc được khử lượng tử hóa, và

trong đó bộ khử lượng tử hóa tham số (840) được tạo cấu hình để khử lượng tử hóa tham số hướng được lượng tử hóa bằng cách sử dụng độ chính xác khử lượng tử hóa.

32. Bộ giải mã theo ví dụ 29, 30 hoặc 31,

trong đó bộ giải mã tham số (820) được tạo cấu hình để xác định, từ độ chính xác khử lượng tử hóa, bảng chữ cái giải mã để giải mã tham số hướng được mã hóa cho khung, và

trong đó bộ giải mã tham số (820) được tạo cấu hình để giải mã tham số hướng được mã hóa bằng cách sử dụng bảng chữ cái giải mã để thu được tham số hướng được lượng tử hóa.

33. Bộ giải mã theo một trong số các ví dụ từ 29 đến 32,

trong đó bộ giải mã tham số (820) được tạo cấu hình để suy ra chỉ số hình cầu được lượng tử hóa từ tham số hướng được mã hóa, và để phân tách chỉ số hình cầu được lượng tử hóa thành chỉ số góc nâng được lượng tử hóa và chỉ số góc phương vị được lượng tử hóa.

34. Bộ giải mã theo một trong các ví dụ từ 29 đến 33, trong đó bộ mã hóa tham số (820) được tạo cấu hình

để xác định, từ độ chính xác khử lượng tử hóa, bảng chữ cái góc nâng hoặc

để xác định, từ tham số góc nâng được lượng tử hóa hoặc tham số góc nâng được khử lượng tử hóa, bảng chữ cái góc phương vị.

35. Bộ giải mã theo một trong các ví dụ từ 29 đến 34, trong đó bộ mã hóa tham số (820) được tạo cấu hình

để giải mã, từ các tham số hướng được mã hóa, tham số góc nâng được lượng tử hóa, và để giải mã, từ các tham số hướng được mã hóa, tham số góc phương vị được lượng tử hóa,

trong đó bộ khử lượng tử hóa tham số (840) được tạo cấu hình để xác định, từ tham số góc nâng được lượng tử hóa hoặc tham số góc nâng được khử lượng tử hóa, bảng chữ cái góc phương vị, trong đó kích thước của bảng chữ cái góc phương vị lớn hơn đối với dữ liệu góc nâng biểu thị góc nâng của góc nâng tuyệt đối thứ nhất so với dữ liệu góc nâng biểu thị góc nâng của góc nâng tuyệt đối thứ hai, góc nâng tuyệt đối thứ hai lớn hơn góc nâng tuyệt đối thứ nhất, và

trong đó bộ giải mã tham số (820) được tạo cấu hình để sử dụng bảng chữ cái góc phương vị để tạo tham số góc phương vị được lượng tử hóa, hoặc trong đó bộ khử lượng tử hóa tham số được tạo cấu hình để sử dụng, để khử lượng tử hóa tham số góc phương vị được lượng tử hóa, bảng chữ cái góc phương vị.

36. Bộ giải mã theo một trong số các ví dụ từ 29 đến 35,

trong đó giao diện đầu vào (800) được tạo cấu hình để xác định chế độ giải mã với mô hình hóa từ chỉ báo chế độ giải mã (806) trong tín hiệu âm thanh được mã hóa,

trong đó bộ giải mã tham số (820) được tạo cấu hình để thu được chỉ số góc nâng trung bình hoặc chỉ số góc phương vị trung bình.

37. Bộ giải mã theo ví dụ 36, trong đó bộ giải mã tham số (820) được tạo cấu hình để xác định, từ chỉ số độ khuếch tán được lượng tử hóa cho khung, độ chính xác khử lượng tử hóa cho khung (851),

để xác định (852a) từ độ chính xác khử lượng tử hóa cho khung, bảng chữ cái trung bình góc nâng hoặc bảng chữ cái trung bình góc phương vị, và

để tính chỉ số góc nâng trung bình sử dụng các bit (808b) trong tín hiệu âm thanh được mã hóa và bảng chữ cái trung bình góc nâng, hoặc để tính chỉ số góc phương vị trung bình sử dụng các bit (808a) trong tín hiệu âm thanh được mã hóa và bảng chữ cái trung bình góc phương vị.

38. Bộ giải mã theo một trong số các ví dụ 36 hoặc 37,

trong đó bộ giải mã tham số (820) được tạo cấu hình để giải mã các bit (808c) nhất định trong tín hiệu âm thanh được mã hóa để thu được tham số Golomb-Rice được giải mã và giải mã các bit (808c) tiếp theo trong tín hiệu âm thanh được mã hóa để thu được các khoảng cách góc nâng được giải mã, hoặc

trong đó bộ giải mã tham số (820) được tạo cấu hình để giải mã các bit nhất định (808a) trong tín hiệu âm thanh được mã hóa để thu được tham số góc phương vị Golomb-Rice được giải mã, và để giải mã các bit tiếp theo (808f) trong tín hiệu âm thanh được mã hóa để thu được các khoảng cách góc phương vị được giải mã,

trong đó bộ giải mã tham số (820) được tạo cấu hình để tính toán các tham số góc nâng được lượng tử hóa từ tham số Golomb-Rice góc nâng và các khoảng cách góc nâng được giải mã và chỉ số trung bình góc nâng, hoặc để tính toán các tham số góc phương vị được lượng tử hóa từ tham số Golomb-Rice góc phương vị và các khoảng cách góc phương vị được giải mã và chỉ số trung bình góc phương vị.

39. Bộ giải mã theo một trong số các ví dụ từ 29 đến 38,

trong đó bộ giải mã tham số (820) được tạo cấu hình để giải mã tham số độ khuếch tán cho phần thời gian và tần số từ tín hiệu âm thanh được mã hóa để thu được tham số độ khuếch tán được lượng tử hóa (850),

trong đó bộ khử lượng tử hóa tham số (840) được tạo cấu hình để xác định độ chính xác khử lượng tử hóa từ tham số độ khuếch tán được lượng tử hóa hoặc được khử lượng tử hóa (851),

trong đó bộ giải mã tham số (820) được tạo cấu hình để suy ra bảng chữ cái góc nâng từ độ chính xác khử lượng tử hóa (852a) và để sử dụng bảng chữ cái góc nâng để thu được tham số góc nâng được lượng tử hóa cho phần thời gian và tần số của khung, và

trong đó bộ khử lượng tử hóa được tạo cấu hình để khử lượng tử hóa tham số góc nâng được lượng tử hóa bằng cách sử dụng bảng chữ cái góc nâng để thu được tham số góc nâng được khử lượng tử hóa cho phần thời gian và tần số của khung.

40. Bộ giải mã theo một trong số các ví dụ từ 29 đến 39,

trong đó bộ giải mã tham số (820) được tạo cấu hình để giải mã tham số hướng được mã hóa để thu được tham số góc nâng được lượng tử hóa,

trong đó bộ khử lượng tử hóa tham số (840) được tạo cấu hình để xác định bảng chữ cái góc phương vị từ tham số góc nâng được lượng tử hóa hoặc tham số góc nâng được khử lượng tử hóa (852c), và

trong đó bộ giải mã tham số (820) được tạo cấu hình để tính toán tham số góc phương vị được lượng tử hóa bằng cách sử dụng bảng chữ cái góc phương vị (852d), hoặc trong đó bộ khử lượng tử hóa tham số (840) được tạo cấu hình để khử lượng tử hóa tham số góc phương vị được lượng tử hóa bằng cách sử dụng bảng chữ cái góc phương vị.

41. Bộ giải mã theo một trong các ví dụ từ 29 đến 40, trong đó bộ khử lượng tử hóa tham số (840) được tạo cấu hình

để xác định bảng chữ cái góc nâng bằng cách sử dụng độ chính xác khử lượng tử hóa (852a), và

để xác định bảng chữ cái góc phương vị (852c) bằng cách sử dụng độ chính xác khử lượng tử hóa và tham số góc nâng được lượng tử hóa hoặc được khử lượng tử hóa được tạo ra bằng cách sử dụng bảng chữ cái góc nâng, và

trong đó bộ giải mã tham số (820) được tạo cấu hình để sử dụng bảng chữ cái góc nâng để giải mã tham số hướng được mã hóa để thu được tham số góc nâng được lượng tử hóa và để sử dụng bảng chữ cái góc phương vị để giải mã tham số hướng được mã hóa để thu được tham số góc phương vị được lượng tử hóa, hoặc trong đó bộ khử lượng tử hóa tham số (840) được tạo cấu hình để khử lượng tử hóa tham số góc nâng được lượng tử hóa bằng cách sử dụng bảng chữ cái góc nâng và để khử lượng tử hóa tham số góc phương vị được lượng tử hóa bằng cách sử dụng bảng chữ cái góc phương vị.

42. Bộ giải mã theo ví dụ 33, trong đó bộ giải mã tham số (820) được tạo cấu hình

để tính toán chỉ số góc nâng dự đoán hoặc chỉ số góc phương vị dự đoán bằng cách sử dụng chỉ số góc nâng trung bình hoặc chỉ số góc phương vị trung bình, và

để thực hiện thao tác giải mã Golomb-Rice, hoặc sự cải biên của nó, để thu được khoảng cách cho tham số góc phương vị hoặc góc nâng, và

để cộng khoảng cách cho tham số góc phương vị hoặc góc nâng vào chỉ số góc nâng trung bình hoặc chỉ số góc phương vị trung bình để thu được chỉ số góc nâng được lượng tử hóa hoặc chỉ số góc phương vị được lượng tử hóa.

43. Bộ giải mã theo một trong số các ví dụ từ 29 đến 42, còn bao gồm:

bộ chuyển đổi độ phân giải tham số (710) để chuyển đổi độ phân giải thời gian/tần số của tham số độ khuếch tán được khử lượng tử hóa hoặc độ phân giải thời gian hoặc tần số của góc phương vị hoặc tham số góc nâng hoặc sự biểu diễn tham số được suy ra từ tham số góc phương vị được khử lượng tử hóa hoặc tham số góc nâng được khử lượng tử hóa thành độ phân giải thời gian hoặc tần số đích, và

bộ kết xuất âm thanh (420) để áp dụng các tham số độ khuếch tán và các tham số hướng trong độ phân giải thời gian hoặc tần số đích cho tín hiệu âm thanh để thu được tín hiệu âm thanh đa kênh được giải mã.

44. Bộ giải mã theo ví dụ 43, bao gồm:

bộ chuyển đổi phổ/thời gian (440) để chuyển đổi tín hiệu âm thanh đa kênh tạo sự biểu diễn miền phổ thành sự biểu diễn miền thời gian có độ phân giải thời gian cao hơn độ phân giải thời gian của độ phân giải thời gian hoặc tần số đích.

45. Bộ giải mã theo một trong số các ví dụ từ 29 đến 44,

trong đó tín hiệu âm thanh được mã hóa bao gồm tín hiệu truyền được mã hóa, trong đó giao diện đầu vào (800) được tạo cấu hình để trích xuất tín hiệu truyền được mã hóa,

trong đó bộ giải mã bao gồm bộ giải mã âm thanh tín hiệu truyền (340) để giải mã tín hiệu truyền được mã hóa,

trong đó bộ giải mã còn bao gồm bộ chuyển đổi thời gian/phổ (430) để chuyển đổi tín hiệu truyền được giải mã thành sự biểu diễn phổ, và

trong đó bộ giải mã bao gồm bộ kết xuất âm thanh (420, 440) để kết xuất tín hiệu âm thanh đa kênh bằng cách sử dụng các tham số độ khuếch tán được khử lượng tử hóa và các tham số hướng được khử lượng tử hóa, và

trong đó bộ giải mã còn bao gồm bộ chuyển đổi phổ/thời gian (440) để chuyển đổi tín hiệu âm thanh được kết xuất thành sự biểu diễn miền thời gian.

46. Phương pháp giải mã tín hiệu âm thanh được mã hóa bao gồm các tham số mã hóa âm thanh có hướng được mã hóa bao gồm các tham số độ khuếch tán được mã hóa và các tham số hướng được mã hóa, phương pháp bao gồm các bước:

nhận (800) tín hiệu âm thanh được mã hóa và tách ra, từ tín hiệu âm thanh được mã hóa, các tham số độ khuếch tán được mã hóa và các tham số hướng được mã hóa;

giải mã (820) các tham số độ khuếch tán được mã hóa và các tham số hướng được mã hóa để thu được các tham số độ khuếch tán được lượng tử hóa và các tham số hướng được lượng tử hóa; và

xác định (840), từ các tham số độ khuếch tán được lượng tử hóa và các tham số hướng được lượng tử hóa, các tham số độ khuếch tán được khử lượng tử hóa và các tham số hướng được khử lượng tử hóa.

47. Chương trình máy tính để thực hiện, khi chạy trên máy tính hoặc bộ xử lý, phương pháp theo ví dụ 28 hoặc 46.

Các tín hiệu âm thanh được mã hóa theo sáng chế bao gồm sự biểu diễn tham số có thể được lưu trữ trên phương tiện lưu trữ kỹ thuật số hoặc phương tiện lưu trữ không tạm thời hoặc có thể được truyền trên phương tiện truyền phát như phương tiện truyền không dây hoặc phương tiện truyền có dây chẳng hạn như Internet.

Mặc dù một vài khía cạnh đã được mô tả trong ngữ cảnh của thiết bị, rõ ràng rằng các khía cạnh này cũng thể hiện mô tả của phương pháp tương ứng, trong đó khối hoặc thiết bị tương ứng với bước xử lý của phương pháp hoặc đặc điểm của bước xử lý của phương pháp. Tương tự, các khía cạnh được mô tả trong ngữ cảnh của bước xử lý của phương pháp cũng biểu diễn sự mô tả của khối hoặc mục hoặc dấu hiệu tương ứng của thiết bị tương ứng.

Tùy thuộc vào các yêu cầu triển khai nhất định, các phương án của sáng chế có thể được thực hiện trong phần cứng hoặc trong phần mềm. Việc triển khai có thể được thực hiện sử dụng vật lưu trữ số, ví dụ, đĩa mềm, DVD, CD, ROM, PROM, EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều khiển có thể đọc được bằng điện được lưu trữ trên đó, mà kết hợp (hoặc có thể kết hợp) với hệ thống máy tính khả trình sao cho phương pháp tương ứng được thực hiện.

Một số phương án theo sáng chế bao gồm vật mang dữ liệu có các tín hiệu điều khiển có thể đọc được bằng điện tử, mà có khả năng kết hợp với hệ thống máy tính có thể lập trình, sao cho một trong số các phương pháp được mô tả ở đây được thực hiện.

Nói chung, các phương án của sáng chế có thể được thực hiện như sản phẩm chương trình máy tính với mã chương trình, mã chương trình có tác dụng để thực hiện một trong số các phương pháp khi sản phẩm chương trình máy tính chạy trên máy tính. Mã chương trình có thể, ví dụ, được lưu trữ trên vật mang có thể đọc được bằng máy.

Các phương án khác bao gồm chương trình máy tính để thực hiện một trong số các phương pháp được mô tả ở đây, được lưu trữ trên vật mang có thể đọc được bằng máy hoặc vật ghi lưu trữ không tạm thời.

Do đó, nói cách khác, phương án của phương pháp theo sáng chế là, chương trình máy tính có mã chương trình để thực hiện một trong số các phương pháp được mô tả ở đây, khi chương trình máy tính chạy trên máy tính.

Do đó, phương án khác của các phương pháp theo sáng chế là vật mang dữ liệu (hoặc vật ghi lưu trữ số, hoặc vật ghi có thể đọc được bằng máy tính) bao gồm chương trình máy tính đã được ghi lại trên đó để thực hiện một trong số các phương pháp được mô tả ở đây.

Do đó, phương án nữa của các phương pháp theo sáng chế là dòng dữ liệu hoặc chuỗi tín hiệu biểu diễn chương trình máy tính để thực hiện một trong số các phương pháp đã được mô tả ở đây. Ví dụ, có thể tạo cấu hình dòng dữ liệu hoặc chuỗi tín hiệu để được truyền thông qua sự kết nối truyền thông dữ liệu, ví dụ thông qua Internet.

Phương án khác bao gồm phương tiện xử lý, ví dụ máy tính, hoặc thiết bị logic lập trình được, được tạo cấu hình để hoặc được làm thích ứng để thực hiện một trong số các phương pháp được mô tả ở đây.

Phương án khác bao gồm máy tính có chương trình máy tính được cài đặt trên đó để thực hiện một trong số các phương pháp được mô tả ở đây.

Theo một số phương án, thiết bị logic lập trình được (ví dụ, mảng cổng lập trình được dạng trường) có thể được sử dụng để thực hiện một số hoặc tất cả các chức năng của các phương pháp được mô tả ở đây. Trong một số phương án, mảng cổng lập trình được dạng trường có thể kết hợp với bộ vi xử lý để thực hiện một trong số các phương pháp được mô tả ở đây. Thông thường, các phương pháp tốt hơn là được thực hiện bởi thiết bị phân cứng bất kỳ.

Các phương án được mô tả ở trên chỉ mang tính minh họa cho các nguyên lý của sáng chế. Cần hiểu rằng các biến thể và biến đổi của các phương án và các chi tiết được mô tả ở đây sẽ là rõ ràng đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật tương ứng. Do đó, mục đích của sáng chế chỉ bị giới hạn bởi phạm vi của các

điểm yêu cầu bảo hộ sắp đưa ra dưới đây và không bị giới hạn bởi các chi tiết cụ thể được biểu diễn bằng cách mô tả và giải thích của các phương án ở đây.

YÊU CẦU BẢO HỘ

1. Thiết bị mã hóa các tham số mã hóa âm thanh có hướng bao gồm các tham số độ khuếch tán và các tham số hướng, bao gồm:

bộ tính tham số để tính toán các tham số độ khuếch tán với độ phân giải thời gian hoặc tần số thứ nhất và để tính toán các tham số hướng với độ phân giải thời gian hoặc tần số thứ hai, trong đó độ phân giải thời gian hoặc tần số thứ hai khác với độ phân giải thời gian hoặc tần số thứ nhất; và

bộ xử lý bộ lượng tử hóa và bộ mã hóa để tạo sự biểu diễn được lượng tử hóa và được mã hóa của các tham số độ khuếch tán và các tham số hướng.

2. Thiết bị theo điểm 1, trong đó bộ tính tham số được tạo cấu hình để tính toán các tham số độ khuếch tán và các tham số hướng sao cho độ phân giải thời gian thứ nhất thấp hơn độ phân giải thời gian thứ hai, hoặc độ phân giải tần số thứ hai lớn hơn độ phân giải tần số thứ nhất, hoặc độ phân giải thời gian thứ nhất thấp hơn độ phân giải thời gian thứ hai và độ phân giải tần số thứ nhất bằng độ phân giải tần số thứ hai.

3. Thiết bị theo một điểm bất kỳ trong số các điểm trên,

trong đó bộ tính tham số được tạo cấu hình để tính các tham số độ khuếch tán và các tham số hướng cho tập hợp các dải tần, trong đó dải có tần số trung tâm thấp hơn thì hẹp hơn so với dải có tần số trung tâm cao hơn.

4. Thiết bị theo một điểm bất kỳ trong số các điểm trên,

trong đó bộ tính tham số được tạo cấu hình để thu được các tham số độ khuếch tán ban đầu có độ phân giải thời gian hoặc tần số thứ ba và để thu được các tham số hướng ban đầu có độ phân giải thời gian hoặc tần số thứ tư, và

trong đó bộ tính tham số được tạo cấu hình để nhóm và tính trung bình các tham số độ khuếch tán ban đầu sao cho độ phân giải thời gian hoặc tần số thứ ba cao hơn độ phân giải thời gian hoặc tần số thứ nhất, hoặc

trong đó bộ tính tham số được tạo cấu hình để nhóm và tính trung bình các tham số hướng ban đầu sao cho độ phân giải thời gian hoặc tần số thứ tư cao hơn độ phân giải thời gian hoặc tần số thứ hai.

5. Thiết bị theo điểm 4,

trong đó độ phân giải thời gian hoặc tần số thứ ba và độ phân giải thời gian hoặc tần số thứ tư bằng nhau.

6. Thiết bị theo điểm 5,

trong đó độ phân giải thời gian hoặc tần số thứ ba là độ phân giải thời gian hoặc tần số không đổi, sao cho mỗi tham số độ khuếch tán ban đầu liên quan đến khe thời gian hoặc ngăn tần số có cùng kích thước, hoặc

trong đó độ phân giải thời gian hoặc tần số thứ tư là độ phân giải thời gian hoặc tần số không đổi, sao cho mỗi tham số hướng ban đầu liên quan đến khe thời gian hoặc ngăn tần số có cùng kích thước, và

trong đó bộ tính tham số được tạo cấu hình để tính trung bình trên nhiều tham số độ khuếch tán thứ nhất liên quan đến nhiều khe thời gian thứ nhất, hoặc

trong đó bộ tính tham số được tạo cấu hình để tính trung bình trên nhiều tham số độ khuếch tán thứ hai liên quan đến nhiều ngăn tần số thứ hai, hoặc

trong đó bộ tính tham số được tạo cấu hình để tính trung bình trên nhiều tham số hướng thứ ba liên quan đến nhiều khe thời gian thứ ba, hoặc

trong đó bộ tính tham số được tạo cấu hình để tính trung bình trên nhiều tham số hướng thứ tư liên quan đến nhiều ngăn tần số thứ tư.

7. Thiết bị theo một trong số các điểm từ 4 đến 6,

trong đó bộ tính tham số được tạo cấu hình để tính trung bình sử dụng phép tính trung bình có trọng số, trong đó tham số độ khuếch tán hoặc tham số hướng được suy ra từ phần tín hiệu đầu vào có số đo liên quan đến biên độ cao hơn được gán trọng số bằng cách sử dụng hệ số trọng số cao hơn so với tham số độ khuếch tán hoặc tham số hướng được suy ra từ phần tín hiệu đầu vào có số đo liên quan đến biên độ thấp hơn.

8. Thiết bị theo điểm 7,

trong đó số đo liên quan đến biên độ là công suất hoặc năng lượng trong phần thời gian hoặc phần tần số hoặc công suất hoặc năng lượng được nâng lên lũy thừa bởi số thực không âm bằng hoặc khác 1 trong phần thời gian hoặc phần tần số.

9. Thiết bị theo một trong số các điểm từ 4 đến 8,

trong đó bộ tính tham số được tạo cấu hình để thực hiện tính trung bình sao cho tham số độ khuếch tán hoặc tham số hướng được chuẩn hóa theo số đo liên quan đến biên độ được suy ra từ phần thời gian của tín hiệu đầu vào tương ứng với độ phân giải thời gian hoặc tần số thứ nhất hoặc thứ hai.

10. Thiết bị theo một trong số các điểm từ 4 đến 8,

trong đó bộ tính tham số được tạo cấu hình để nhóm và tính trung bình các tham số hướng ban đầu bằng cách sử dụng phép tính trung bình có trọng số, trong đó tham số hướng thứ nhất liên quan đến phần thời gian thứ nhất có tham số độ khuếch tán thứ nhất biểu thị độ khuếch tán thấp hơn được gán trọng số mạnh hơn tham số hướng thứ hai liên quan đến phần thời gian thứ hai có tham số độ khuếch tán thứ hai biểu thị độ khuếch tán cao hơn.

11. Thiết bị theo một điểm bất kỳ trong số các điểm trên,

trong đó bộ tính tham số được tạo cấu hình để tính toán các tham số hướng ban đầu sao cho các tham số hướng ban đầu bao gồm vector Đề-các có thành phần cho mỗi hướng trong số hai hoặc ba hướng, và trong đó bộ tính tham số được tạo cấu hình để thực hiện phép tính trung bình cho từng thành phần riêng của vector Đề-các một cách riêng rẽ, hoặc trong đó các thành phần được chuẩn hóa sao cho tổng các thành phần bình phương của vector Đề-các cho tham số hướng bằng với phần tử đơn vị.

12. Thiết bị theo một điểm bất kỳ trong số các điểm trên còn bao gồm:

bộ phân tách thời gian-tần số để phân tách tín hiệu đầu vào có nhiều kênh đầu vào thành sự biểu diễn thời gian-tần số cho mỗi kênh đầu vào, hoặc

trong đó bộ giải mã thời gian-tần số được tạo cấu hình để phân tách tín hiệu đầu vào có nhiều kênh đầu vào thành sự biểu diễn thời gian-tần số cho mỗi kênh đầu vào có độ phân giải thời gian hoặc tần số thứ ba hoặc độ phân giải thời gian hoặc tần số thứ tư.

13. Thiết bị theo một trong số các điểm từ 5 đến 12,

trong đó bộ phân tách thời gian-tần số bao gồm giàn lọc được biến điệu tạo ra các giá trị phức tạp cho từng tín hiệu băng con, trong đó mỗi tín hiệu băng con có nhiều khe thời gian trên mỗi khung và dải tần.

14. Thiết bị theo một điểm bất kỳ trong số các điểm trên,

trong đó thiết bị được tạo cấu hình để liên kết chỉ báo của độ phân giải thời gian hoặc tần số thứ nhất hoặc thứ hai với sự biểu diễn được lượng tử hóa và được mã hóa để truyền đến bộ giải mã hoặc để lưu trữ.

15. Thiết bị theo một trong số các điểm trên, trong đó bộ xử lý bộ lượng tử hóa và bộ mã hóa để tạo sự biểu diễn được lượng tử hóa và được mã hóa của các tham số độ khuếch tán và các tham số hướng bao gồm bộ lượng tử hóa tham số để lượng tử hóa các tham số độ khuếch tán và tham số hướng và bộ mã hóa tham số để mã hóa các tham số độ khuếch tán được lượng tử hóa và các tham số hướng được lượng tử hóa như được xác định theo ví dụ bất kỳ trong số các ví dụ từ 1 đến 26 ở trên.

16. Phương pháp mã hóa các tham số mã hóa âm thanh có hướng bao gồm các tham số độ khuếch tán và các tham số hướng, phương pháp bao gồm các bước:

tính toán các tham số độ khuếch tán với độ phân giải thời gian hoặc tần số thứ nhất và để tính toán các tham số hướng với độ phân giải thời gian hoặc tần số thứ hai, trong đó độ phân giải thời gian hoặc tần số thứ hai khác với độ phân giải thời gian hoặc tần số thứ nhất; và

tạo sự biểu diễn được lượng tử hóa và được mã hóa của các tham số độ khuếch tán và các tham số hướng.

17. Bộ giải mã để giải mã tín hiệu âm thanh được mã hóa bao gồm các tham số mã hóa âm thanh có hướng chứa các tham số độ khuếch tán được mã hóa và các tham số hướng được mã hóa, bộ giải mã bao gồm:

bộ xử lý tham số để giải mã các tham số mã hóa âm thanh có hướng được mã hóa để thu được các tham số độ khuếch tán được giải mã với độ phân giải thời gian hoặc tần số thứ nhất và các tham số hướng được giải mã với độ phân giải thời gian hoặc tần số thứ hai; và

bộ chuyển đổi độ phân giải tham số để chuyển đổi các tham số độ khuếch tán được mã hóa hoặc được giải mã hoặc các tham số hướng được mã hóa hoặc được giải mã thành các tham số độ khuếch tán được chuyển đổi hoặc các tham số hướng được chuyển đổi có độ phân giải thời gian hoặc tần số thứ ba, độ phân giải thời gian hoặc tần số thứ ba khác với độ phân giải thời gian hoặc tần số thứ nhất hoặc độ phân giải thời gian hoặc tần số thứ hai hoặc khác với độ phân giải thời gian hoặc tần số thứ nhất và độ phân giải thời gian hoặc tần số thứ hai.

18. Bộ giải mã theo điểm 17, còn bao gồm bộ kết xuất âm thanh hoạt động trong miền phổ, miền phổ bao gồm, đối với khung, số khe thời gian thứ nhất và số dải tần thứ hai, sao cho khung bao gồm số ngăn thời gian/tần số bằng với kết quả nhân của số thứ nhất và số thứ hai, trong đó số thứ nhất và số thứ hai xác định độ phân giải thời gian hoặc tần số thứ ba.

19. Bộ giải mã theo điểm 17 hoặc 18, còn bao gồm bộ kết xuất âm thanh hoạt động trong miền phổ, miền phổ bao gồm, đối với khung, số khe thời gian thứ nhất và số dải tần thứ hai, sao cho khung bao gồm số ngăn thời gian/tần số bằng với kết quả nhân của số thứ nhất và số thứ hai, trong đó số thứ nhất và số thứ hai xác định độ phân giải thời gian hoặc tần số thứ tư, trong đó độ phân giải thời gian hoặc tần số thứ tư bằng hoặc cao hơn độ phân giải thời gian hoặc tần số thứ ba.

20. Bộ giải mã theo một điểm bất kỳ trong số các điểm từ 17 đến 19.

trong đó độ phân giải thời gian hoặc tần số thứ nhất thấp hơn độ phân giải thời gian hoặc tần số thứ hai, và

trong đó bộ chuyển đổi độ phân giải tham số được tạo cấu hình để tạo, từ tham số độ khuếch tán được giải mã, tập hợp thứ nhất của các tham số độ khuếch tán được chuyển đổi và để tạo, từ tham số hướng được giải mã, tập hợp thứ hai của các tham số hướng được chuyển đổi, trong đó tập hợp thứ hai lớn hơn tập hợp thứ nhất.

21. Bộ giải mã theo một trong số các điểm từ 17 đến 20,

trong đó tín hiệu âm thanh được mã hóa bao gồm chuỗi các khung, trong đó mỗi khung được tổ chức theo các dải tần, trong đó mỗi khung chỉ bao gồm một tham

số độ khuếch tán được mã hóa trên mỗi dải tần và ít nhất hai tham số hướng tuần tự theo thời gian trên mỗi dải tần, và

trong đó bộ chuyển đổi độ phân giải tham số được tạo cấu hình để liên kết tham số độ khuếch tán được giải mã với tất cả các ngăn thời gian trong dải tần hoặc với từng ngăn thời gian/tần số có trong dải tần trong khung, và

để liên kết một trong số ít nhất hai tham số hướng của dải tần với nhóm các ngăn thời gian thứ nhất và với từng ngăn thời gian/tần số có trong dải tần, và để liên kết tham số hướng được giải mã thứ hai trong số ít nhất hai tham số hướng với nhóm các ngăn thời gian thứ hai và với mỗi ngăn thời gian/tần số có trong dải tần, trong đó nhóm thứ hai không chứa bất kỳ ngăn thời gian nào trong nhóm thứ nhất.

22. Bộ giải mã theo một trong số các điểm từ 17 đến 21, trong đó tín hiệu âm thanh được mã hóa bao gồm tín hiệu truyền âm thanh được mã hóa, trong đó bộ giải mã bao gồm:

bộ giải mã âm thanh để giải mã tín hiệu âm thanh truyền được mã hóa để thu được tín hiệu âm thanh được giải mã, và

bộ chuyển đổi thời gian/tần số để chuyển đổi tín hiệu âm thanh được giải mã thành sự biểu diễn tần số có độ phân giải thời gian hoặc tần số thứ ba.

23. Bộ giải mã theo một trong số các điểm từ 17 đến 22 bao gồm:

bộ kết xuất âm thanh để áp dụng các tham số độ khuếch tán được chuyển đổi và các tham số hướng được chuyển đổi thành sự biểu diễn phổ của tín hiệu âm thanh ở độ phân giải thời gian hoặc tần số thứ ba để thu được sự biểu diễn phổ tổng hợp; và

bộ chuyển đổi phổ/thời gian để chuyển đổi sự biểu diễn phổ tổng hợp ở độ phân giải thời gian hoặc tần số thứ ba hoặc thứ tư để thu được tín hiệu âm thanh không gian miền thời gian tổng hợp có độ phân giải thời gian cao hơn độ phân giải thời gian hoặc tần số thứ ba.

24. Bộ giải mã theo một trong số các điểm từ 17 đến 23,

trong đó bộ chuyển đổi độ phân giải tham số được tạo cấu hình để nhân tham số hướng được giải mã bằng cách sử dụng thao tác sao chép hoặc để nhân tham số độ

khuếch tán được giải mã bằng cách sử dụng thao tác sao chép hoặc để làm mịn hoặc lọc thông thấp tập hợp các tham số hướng đã nhân hoặc tập hợp các tham số độ khuếch tán đã nhân.

25. Bộ giải mã theo một trong số các điểm từ 17 đến 24,

trong đó độ phân giải thời gian hoặc tần số thứ hai khác độ phân giải thời gian hoặc tần số thứ nhất.

26. Bộ giải mã theo một trong số các điểm từ 17 đến 25,

trong đó độ phân giải thời gian thứ nhất thấp hơn độ phân giải thời gian thứ hai, hoặc độ phân giải tần số thứ hai lớn hơn độ phân giải tần số thứ nhất, hoặc độ phân giải thời gian thứ nhất thấp hơn độ phân giải thời gian thứ hai và độ phân giải tần số thứ nhất bằng độ phân giải tần số thứ hai.

27. Bộ giải mã theo một trong số các điểm từ 17 đến 26,

trong đó bộ chuyển đổi độ phân giải tham số được tạo cấu hình để nhân các tham số độ khuếch tán được giải mã và các tham số hướng được giải mã thành số các tham số chuyển đổi liên kế tần số tương ứng cho tập hợp các dải, trong đó dải có tần số trung tâm thấp hơn nhận được ít tham số được nhân hơn so với dải có tần số trung tâm cao hơn.

28. Bộ giải mã theo một trong số các điểm từ 17 đến 27,

trong đó bộ xử lý tham số được tạo cấu hình để giải mã tham số độ khuếch tán được mã hóa cho khung tín hiệu âm thanh được mã hóa để thu được tham số độ khuếch tán được lượng tử hóa cho khung, và trong đó bộ xử lý tham số được tạo cấu hình để xác định độ chính xác khử lượng tử hóa cho sự lượng tử hóa của ít nhất một tham số hướng cho khung bằng cách sử dụng tham số độ khuếch tán được lượng tử hóa hoặc được khử lượng tử hóa, và

trong đó bộ xử lý tham số được tạo cấu hình để khử lượng tử hóa tham số hướng được lượng tử hóa bằng cách sử dụng độ chính xác khử lượng tử hóa.

29. Bộ giải mã theo một trong số các điểm từ 17 đến 28,

trong đó bộ xử lý tham số được tạo cấu hình để xác định, từ độ chính xác khử lượng tử hóa, sẽ được sử dụng bởi bộ xử lý tham số để khử lượng tử hóa, bảng chữ cái giải mã để giải mã tham số hướng được mã hóa cho khung, và

trong đó bộ xử lý tham số được tạo cấu hình để giải mã tham số hướng được mã hóa bằng cách sử dụng bảng chữ cái giải mã được xác định và để xác định tham số hướng được khử lượng tử hóa.

30. Bộ giải mã theo một trong số các điểm từ 17 đến 29,

trong đó bộ xử lý tham số được tạo cấu hình để xác định, từ độ chính xác khử lượng tử hóa sẽ được sử dụng bởi bộ xử lý tham số (300) để khử lượng tử hóa tham số hướng, bảng chữ cái góc nâng để xử lý tham số góc nâng được mã hóa và để xác định, từ chỉ số góc nâng thu được bằng cách sử dụng bảng chữ cái góc nâng, bảng chữ cái góc phương vị, và

trong đó bộ xử lý tham số được tạo cấu hình để khử lượng tử hóa tham số góc phương vị được mã hóa bằng cách sử dụng bảng chữ cái góc phương vị.

31. Phương pháp giải mã tín hiệu âm thanh được mã hóa bao gồm các tham số mã hóa âm thanh có hướng chứa các tham số độ khuếch tán được mã hóa và các tham số hướng được mã hóa, phương pháp bao gồm các bước:

giải mã các tham số mã hóa âm thanh có hướng được mã hóa để thu được các tham số độ khuếch tán được giải mã với độ phân giải thời gian hoặc tần số thứ nhất và các tham số hướng được giải mã với độ phân giải thời gian hoặc tần số thứ hai; và

chuyển đổi các tham số độ khuếch tán được mã hóa hoặc được giải mã hoặc các tham số hướng được mã hóa hoặc được giải mã thành các tham số độ khuếch tán được chuyển đổi hoặc các tham số hướng được chuyển đổi có độ phân giải thời gian hoặc tần số thứ ba, độ phân giải thời gian hoặc tần số thứ ba khác với độ phân giải thời gian hoặc tần số thứ nhất hoặc độ phân giải thời gian hoặc tần số thứ hai hoặc khác với độ phân giải thời gian hoặc tần số thứ nhất và độ phân giải thời gian hoặc tần số thứ hai.

32. Vật ghi đọc được bằng máy tính chứa chương trình máy tính để thực hiện phương pháp theo điểm 16 hoặc điểm 31 khi chạy trên máy tính hoặc bộ xử lý.

1/38

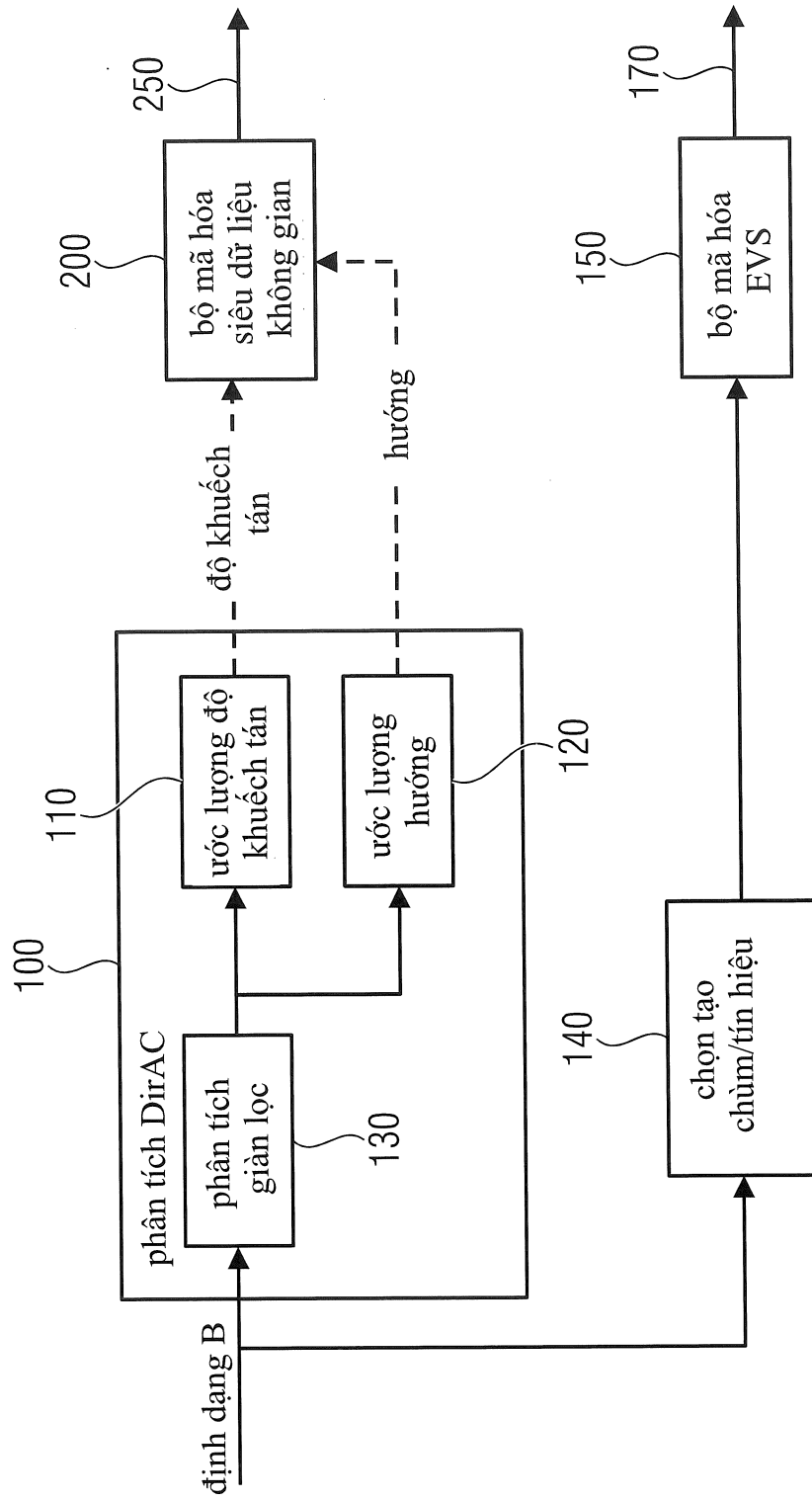


Fig. 1a

2/38

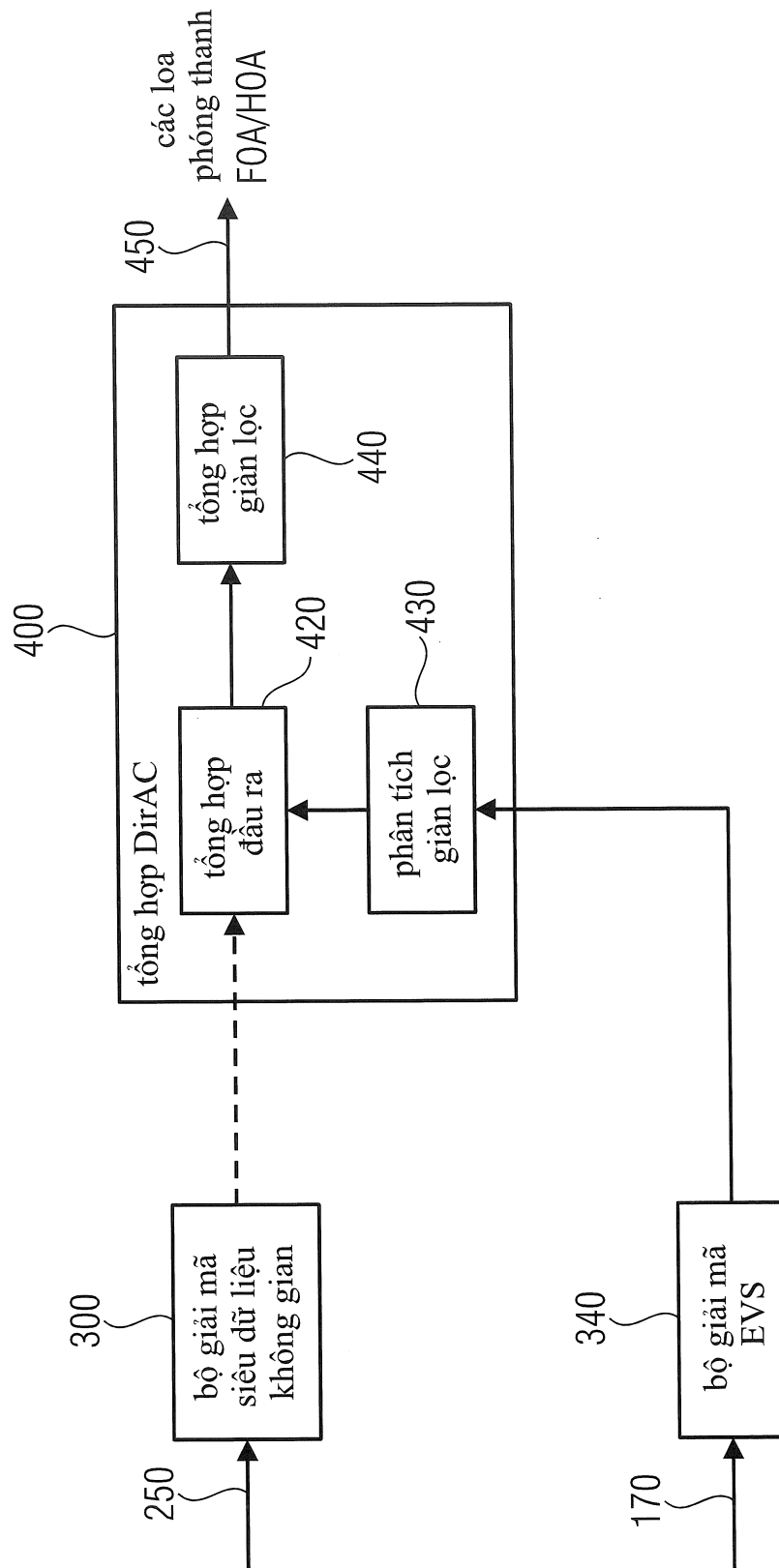


Fig. 1b

3/38

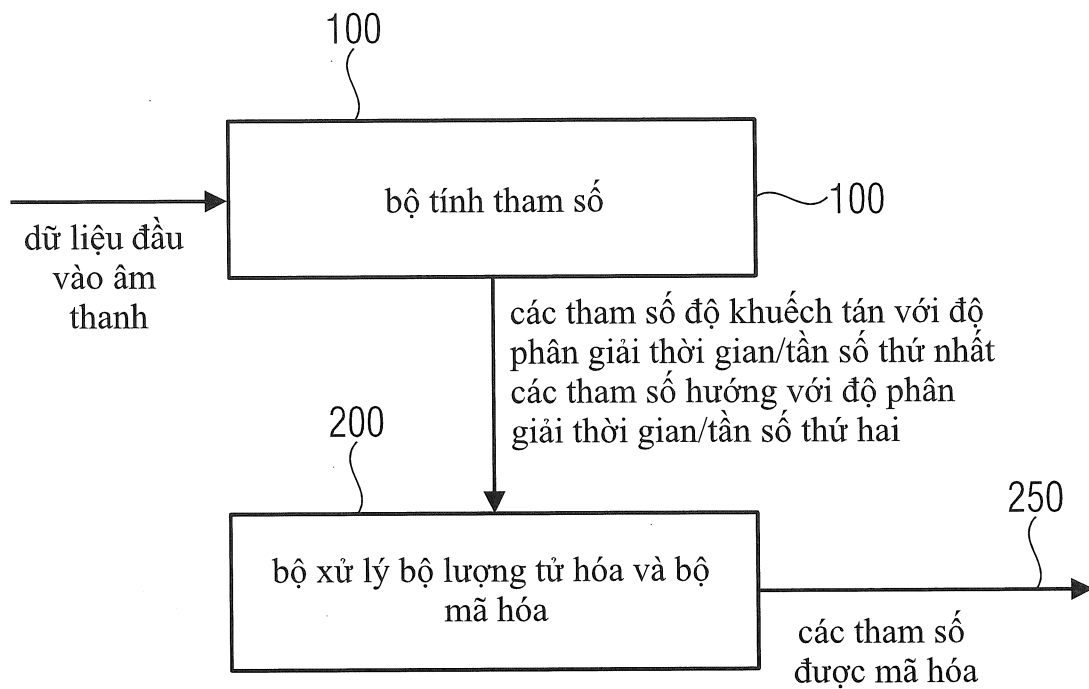


Fig. 2a

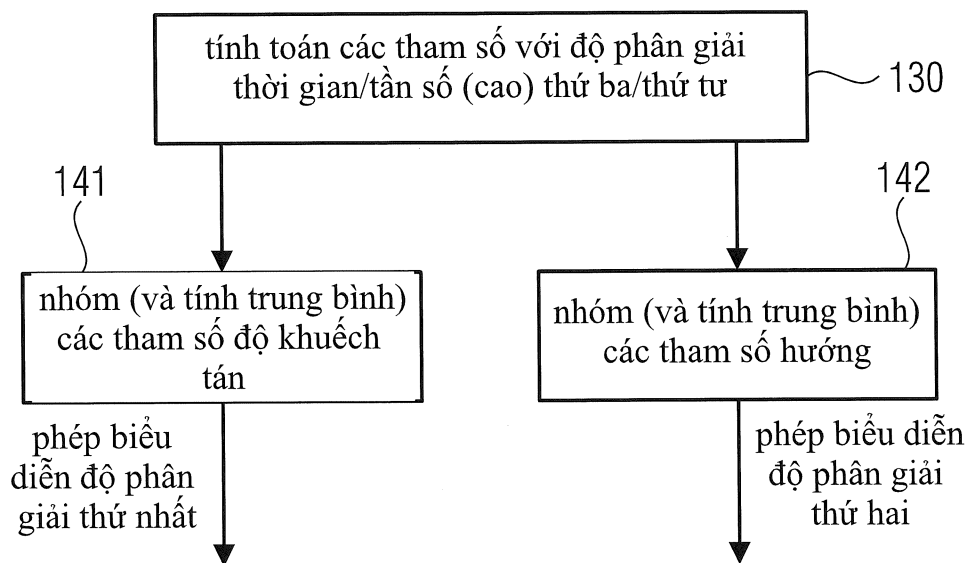


Fig. 2b

4/38

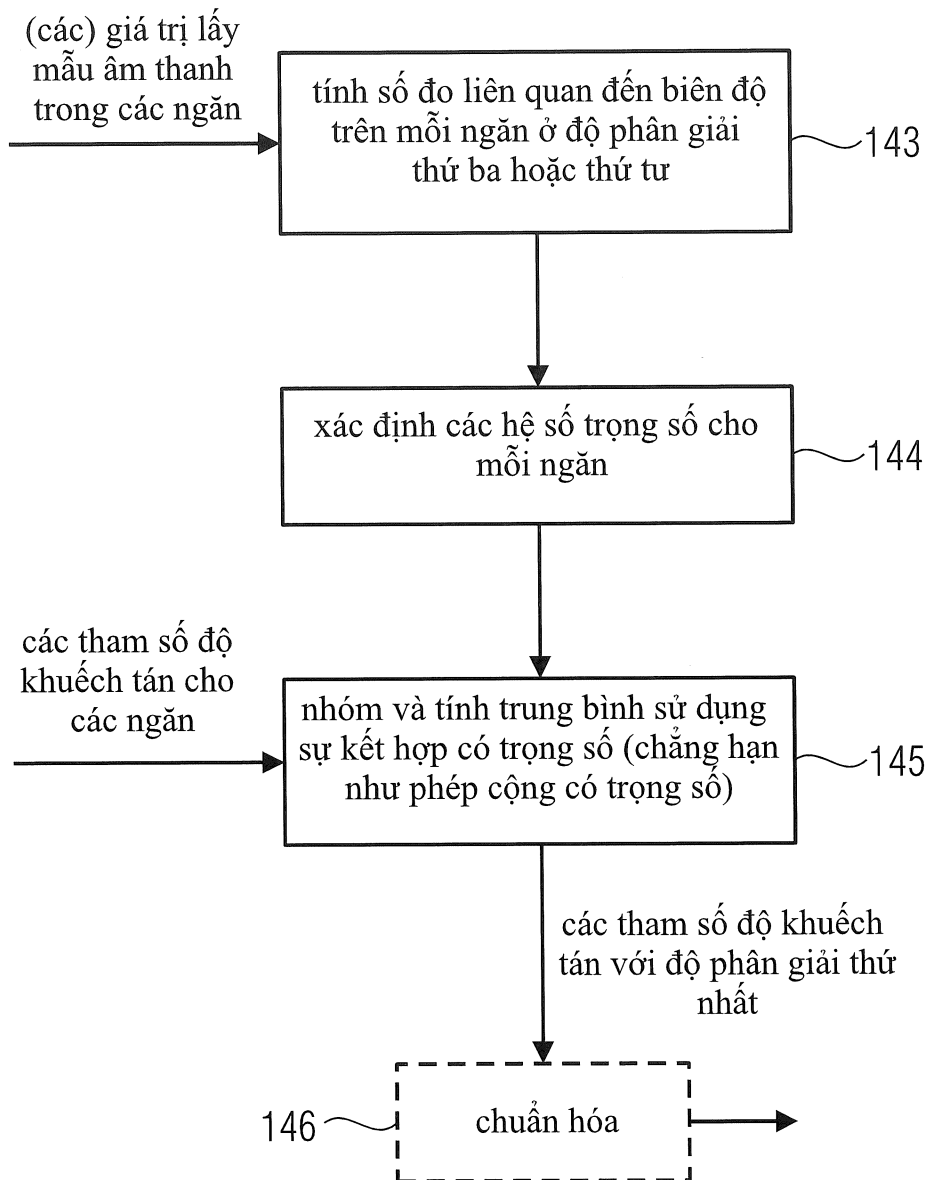


Fig. 2c

5/38

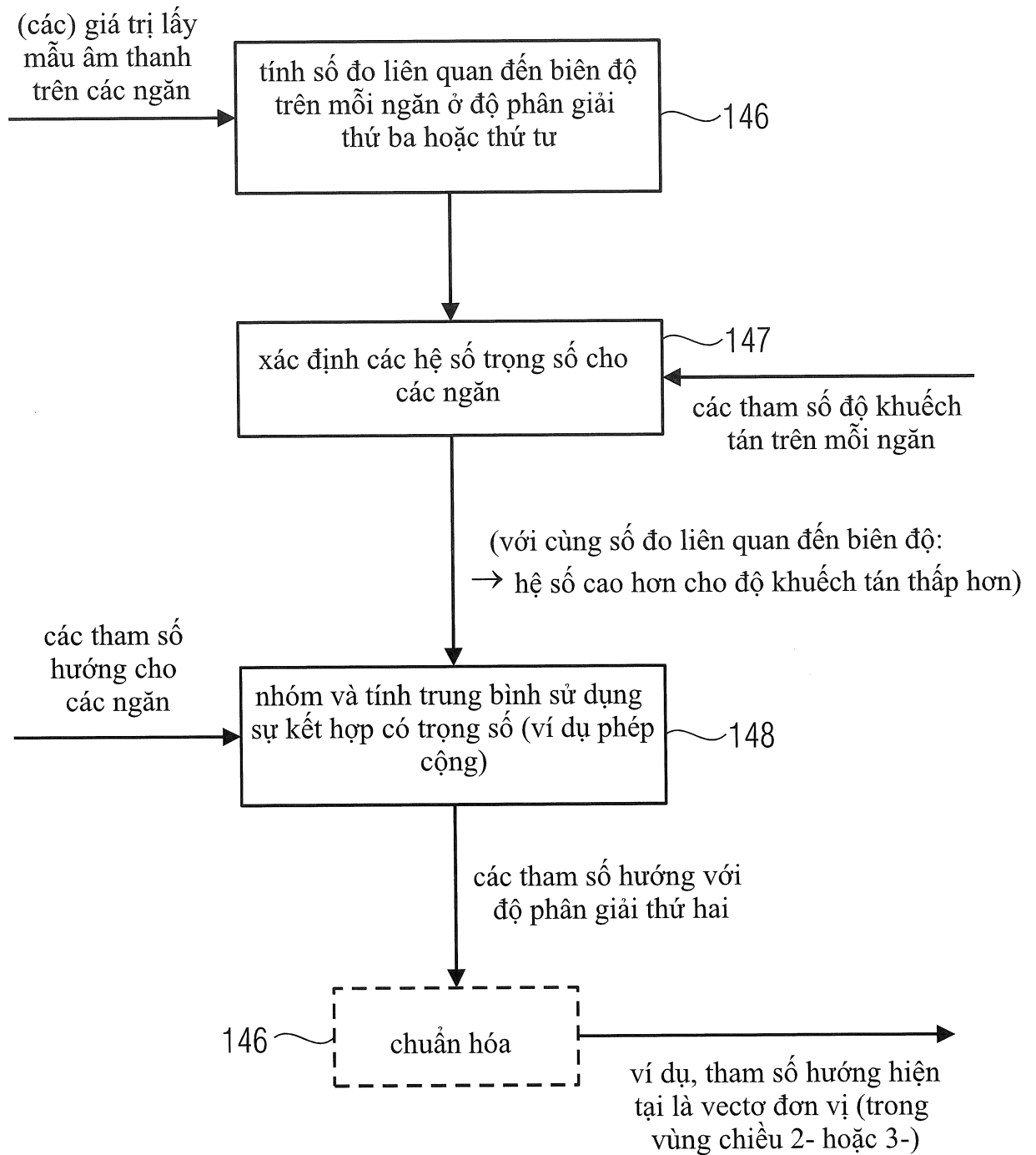


Fig. 2d

6/38

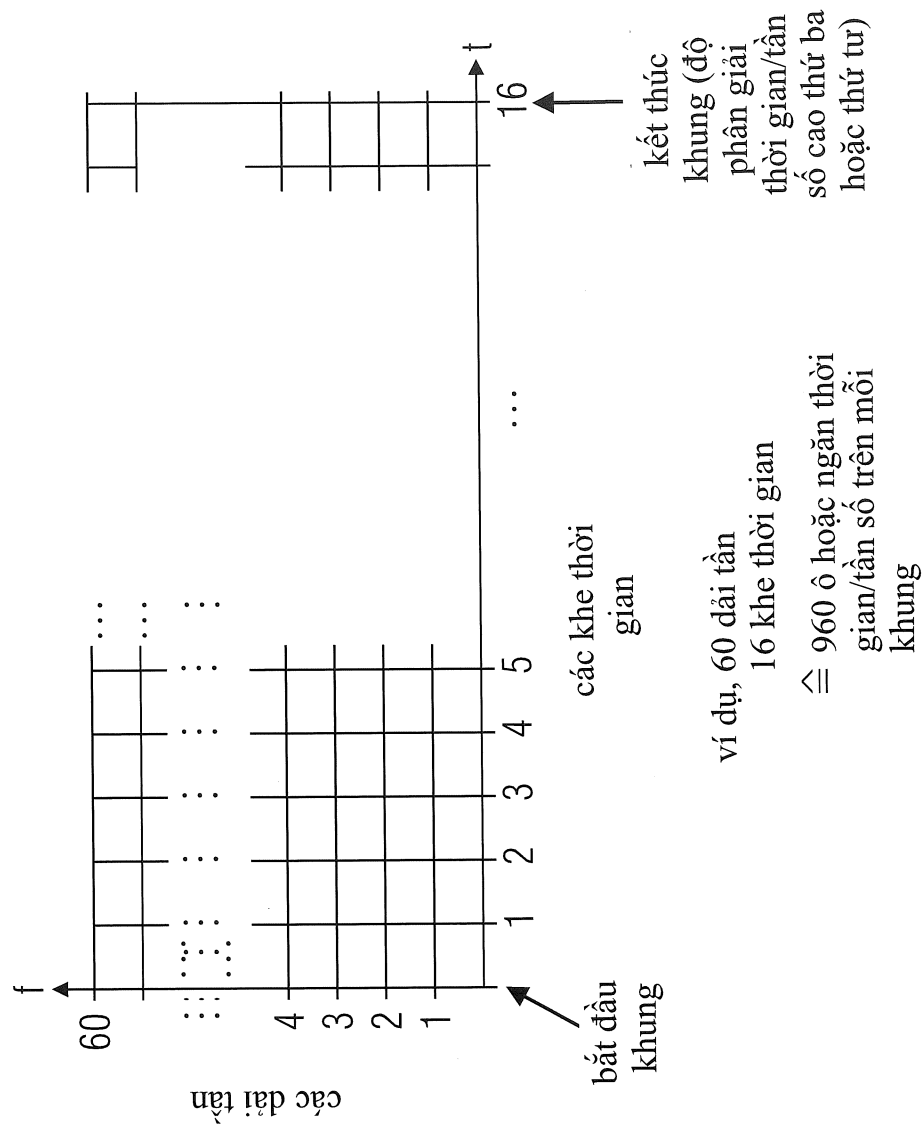
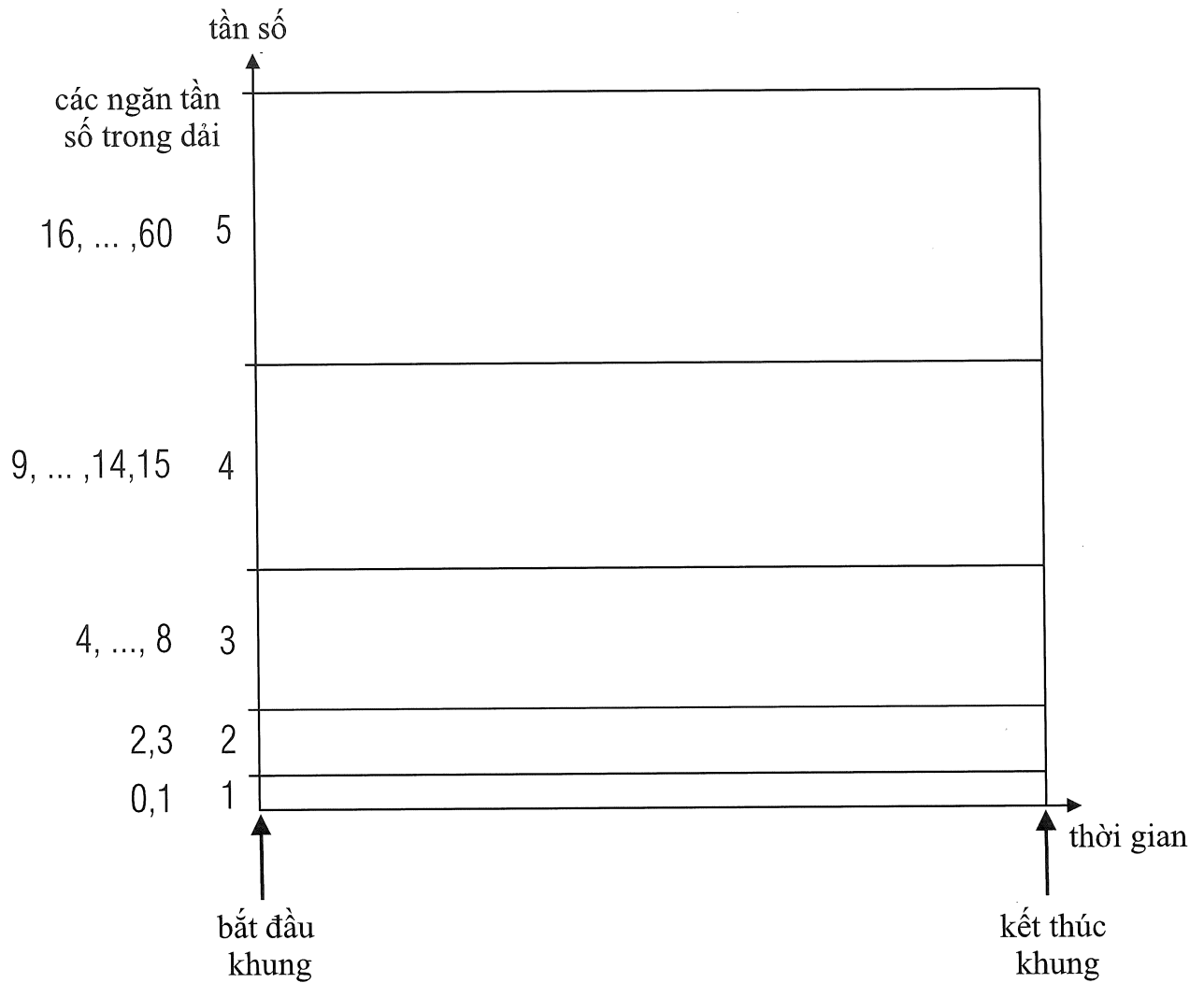


Fig. 3a
(độ phân giải cao)

7/38



nhóm độ khuếch tán

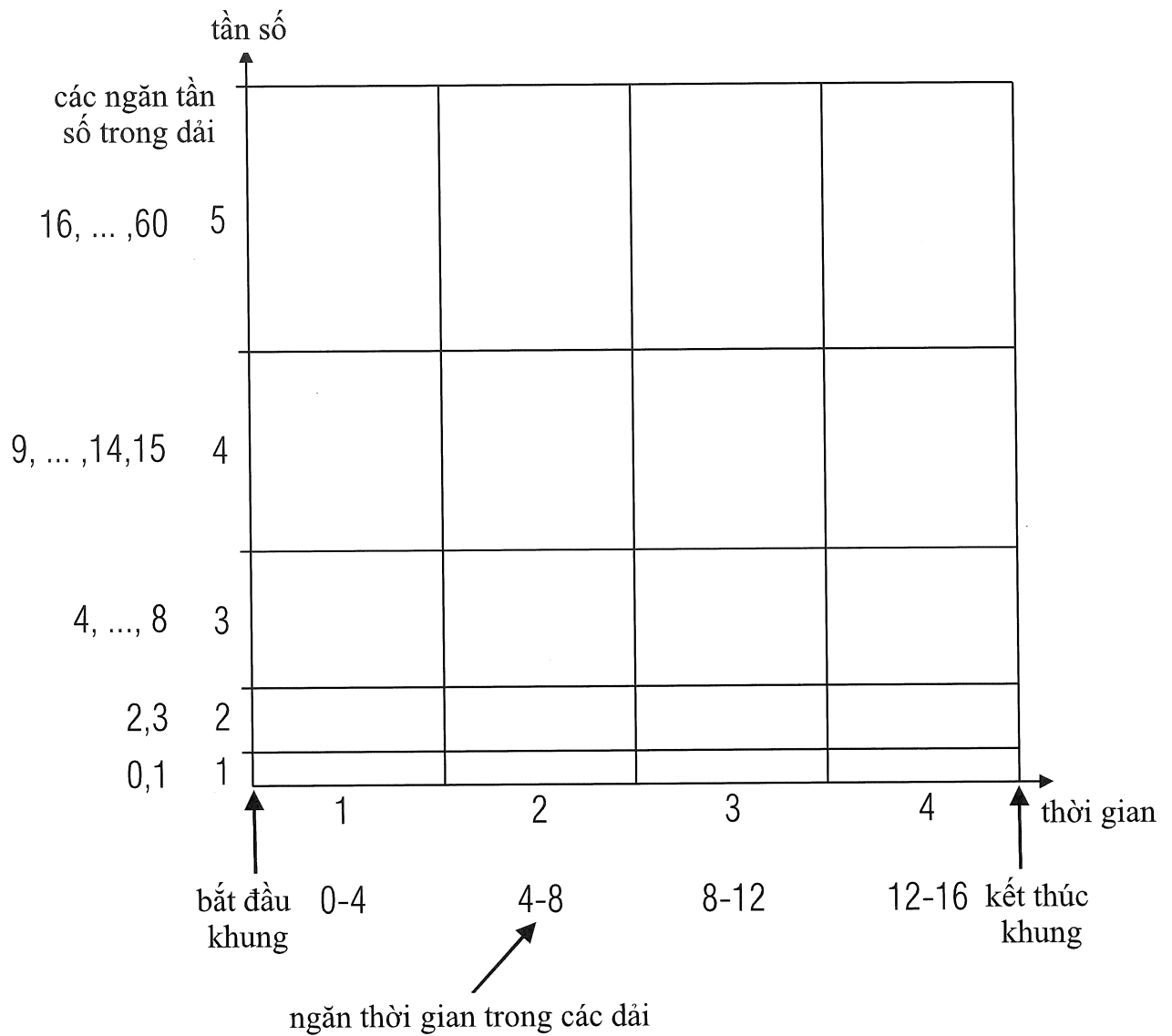
độ phân giải thời gian/tần số thứ nhất

ví dụ, 1 khe thời gian

5 dải (băng thông tăng từ
tần số thấp đến cao) $\hat{=}$ 5 ngăn được nhóm trên mỗi khung

Fig. 3b
(độ phân giải thấp)

8/38

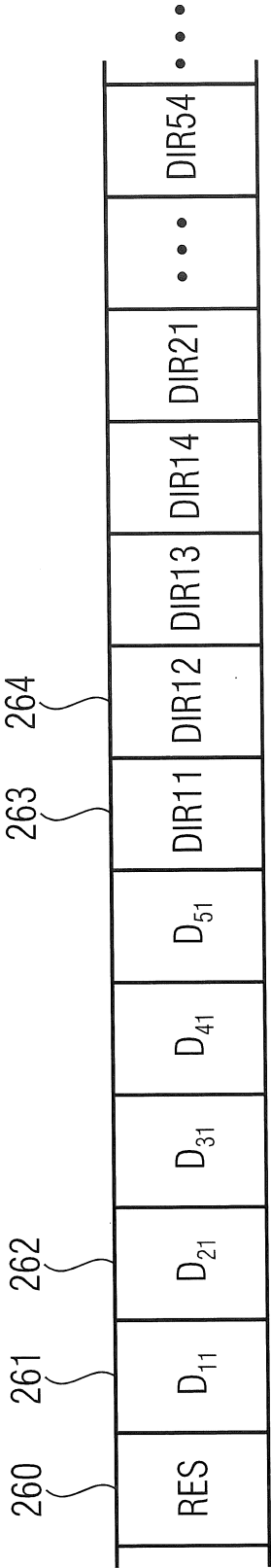


nhóm tham số hướng
độ phân giải thời gian/tần số thứ hai

ví dụ, 4 khe thời gian (mỗi khe có 4 ngăn thời gian)
5 dải (bảng thông tăng từ thấp đến cao)

$\cong$ 20 ngăn được nhóm trên mỗi khung

Fig. 3c
(độ phân giải trung bình)



D₁₁ : tham số độ khuếch tán được mã hóa cho dài thứ nhất và khe thời gian thứ nhất

D₂₁ : tham số độ khuếch tán được mã hóa cho dài thứ hai và khe thời gian thứ nhất

•
•
•

— — — — —

DIR₁₁ : tham số hướng được mã hóa cho dài thứ nhất và khe thời gian thứ nhất

•
•
•

DIR₅₄ : tham số hướng được mã hóa cho dài thứ năm và khe thời gian thứ tư

RES : chỉ báo độ phân giải cho độ phân giải thứ nhất và/hoặc thứ hai (không cần thiết, khi thiết lập cố định bởi bộ mã hóa và bộ giải mã)

Fig. 3d

10/38

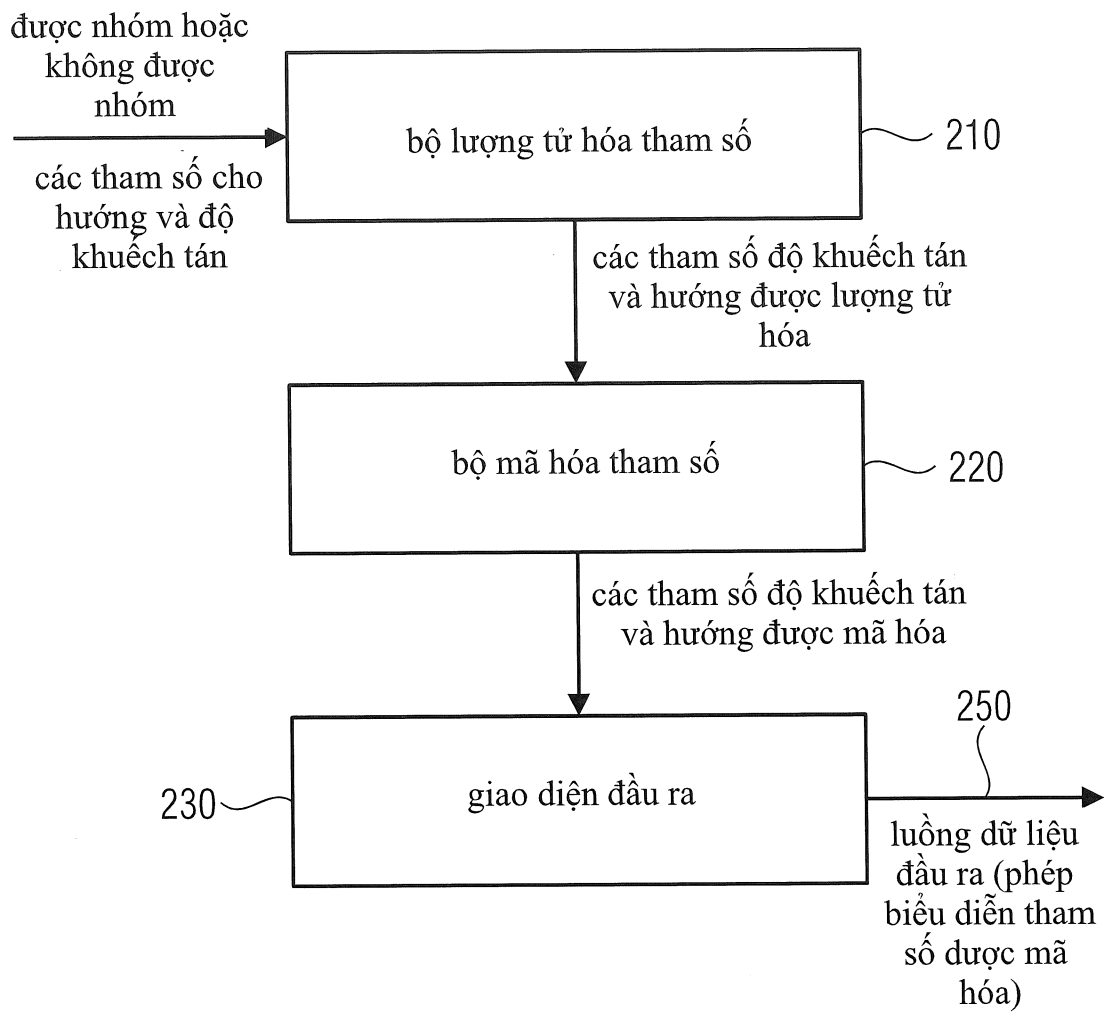


Fig. 4a

11/38

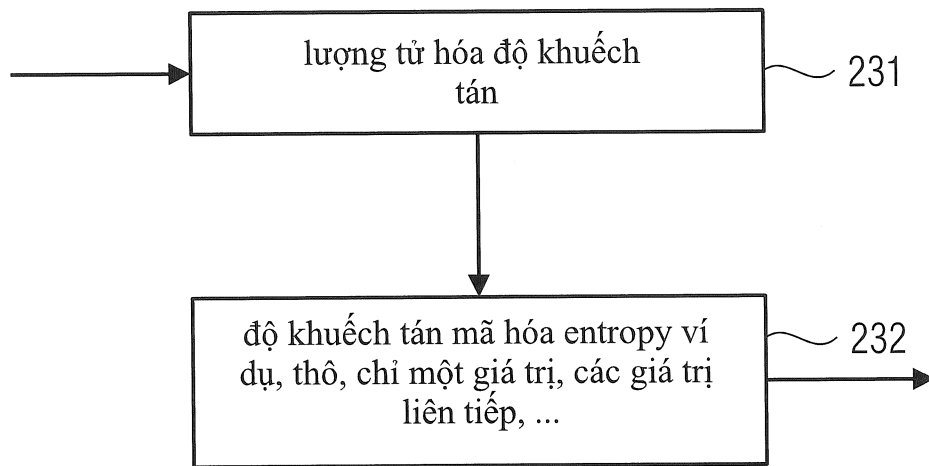


Fig. 4b

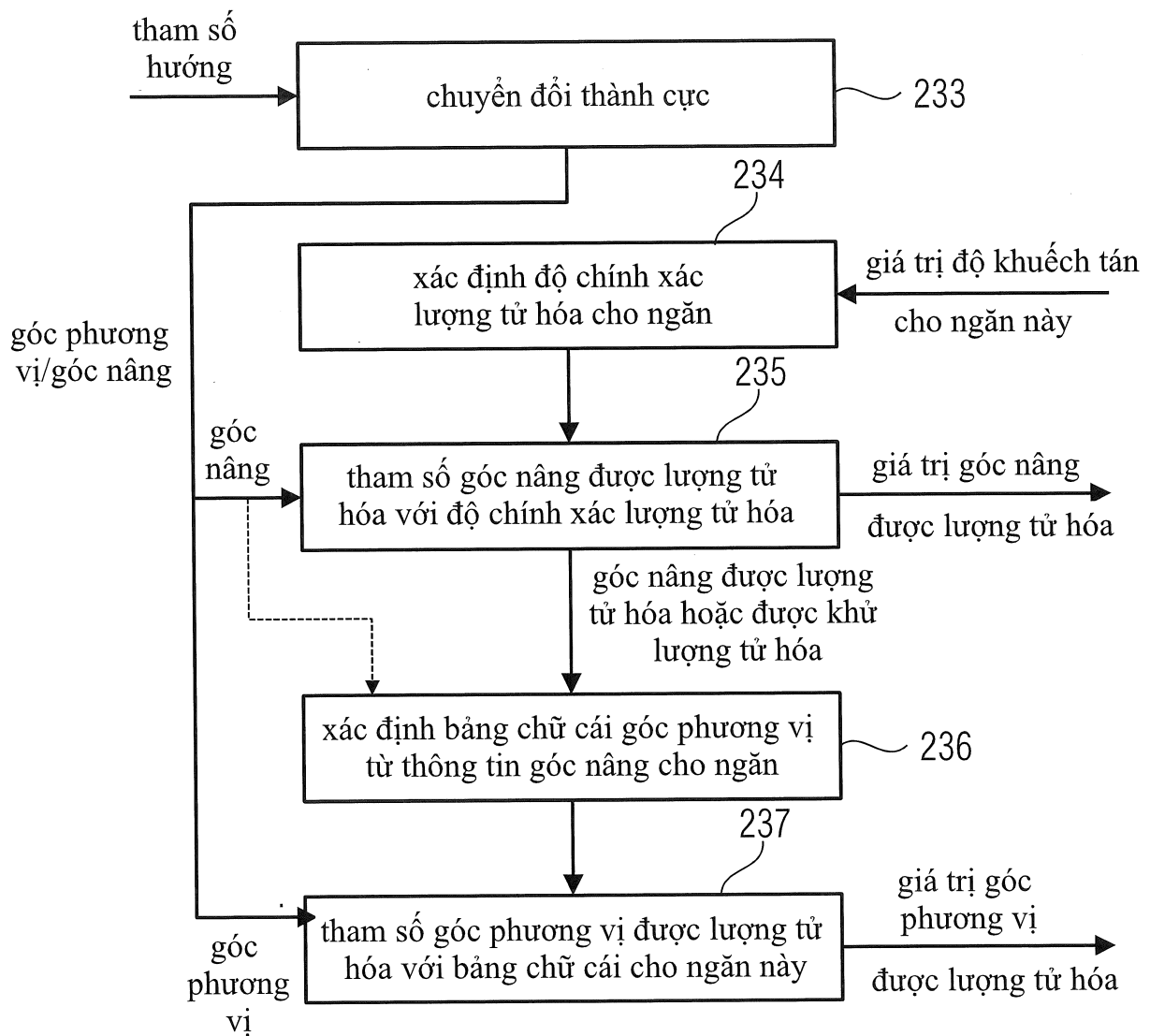
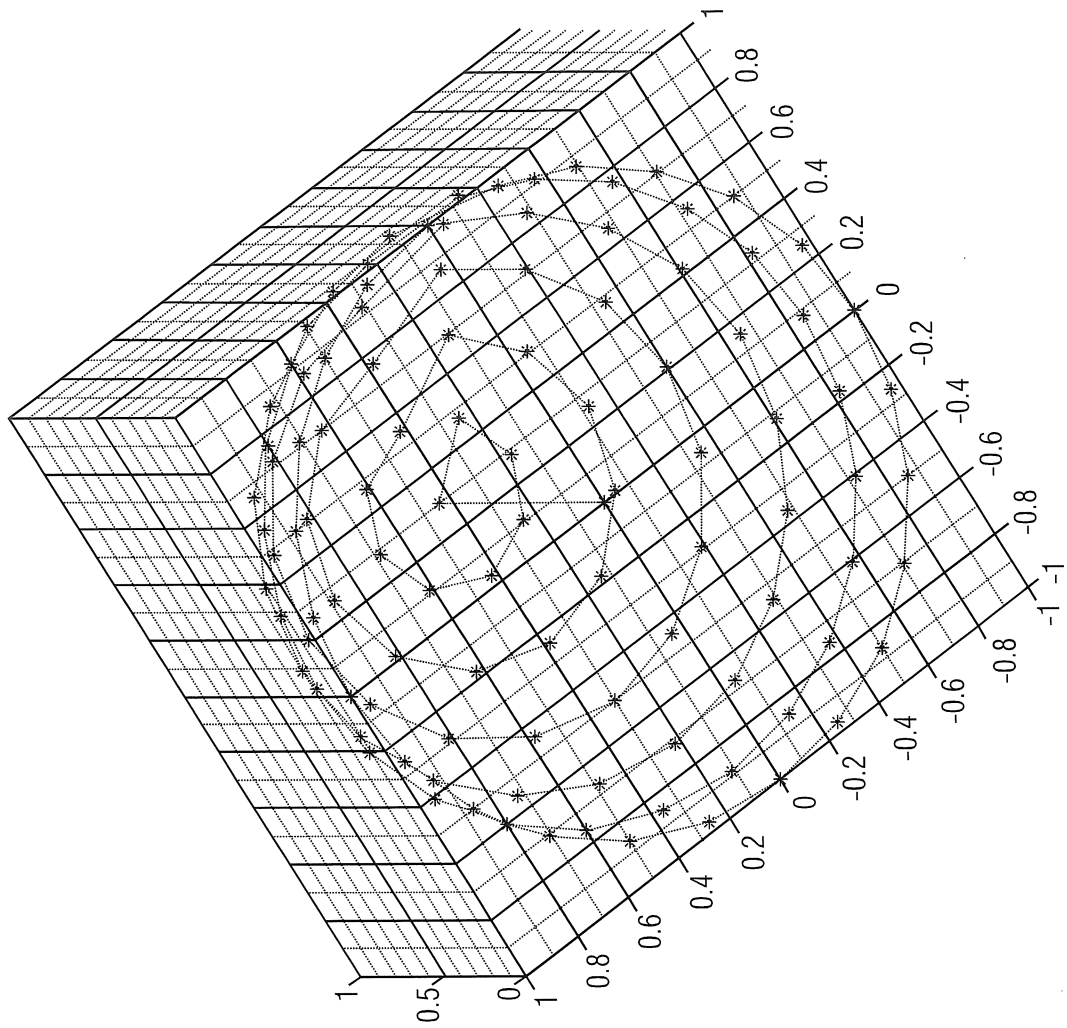


Fig. 4c

12/38

**Fig. 4d**

sự bao phủ giả đều của khối cầu đơn vị

13/38

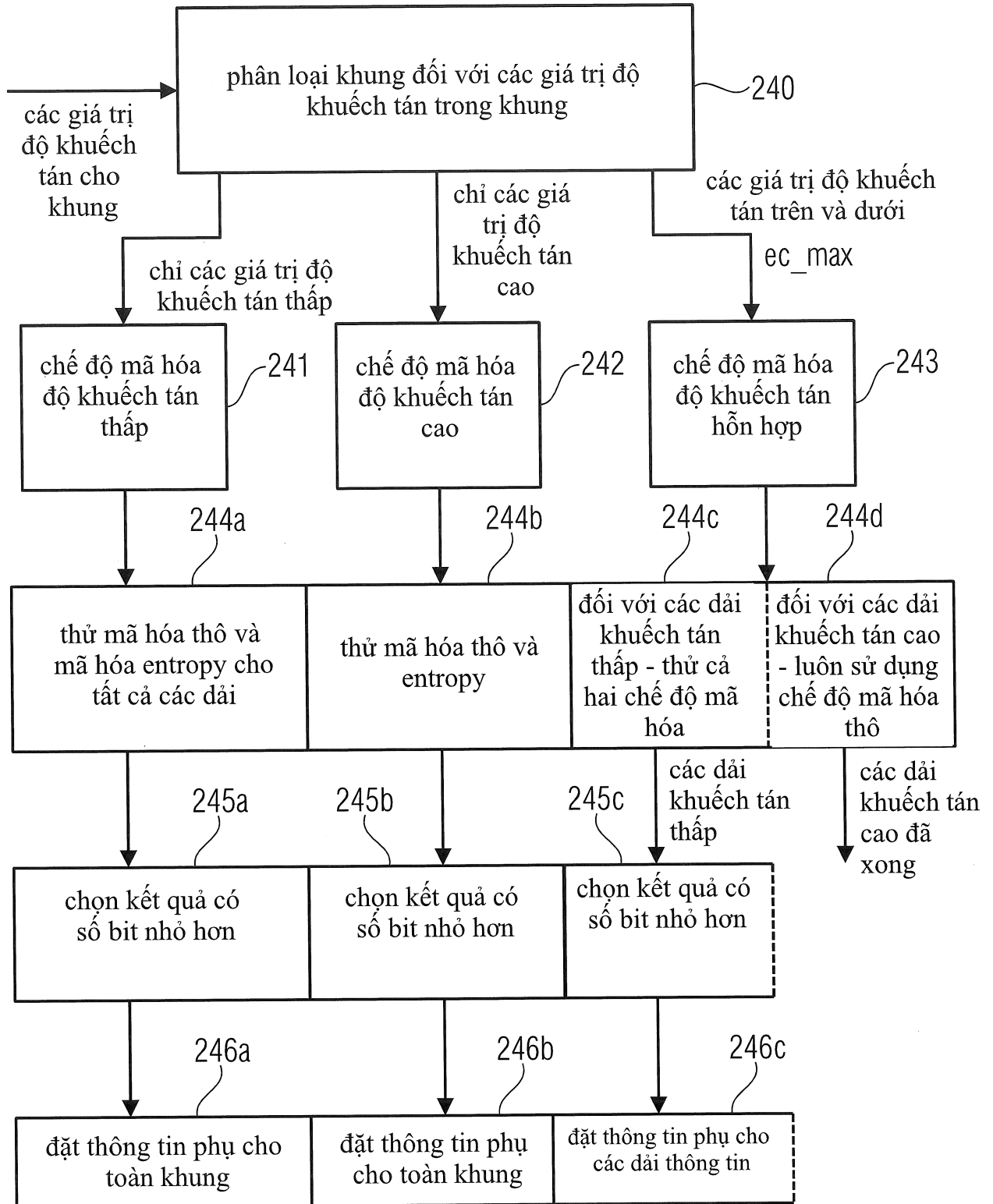
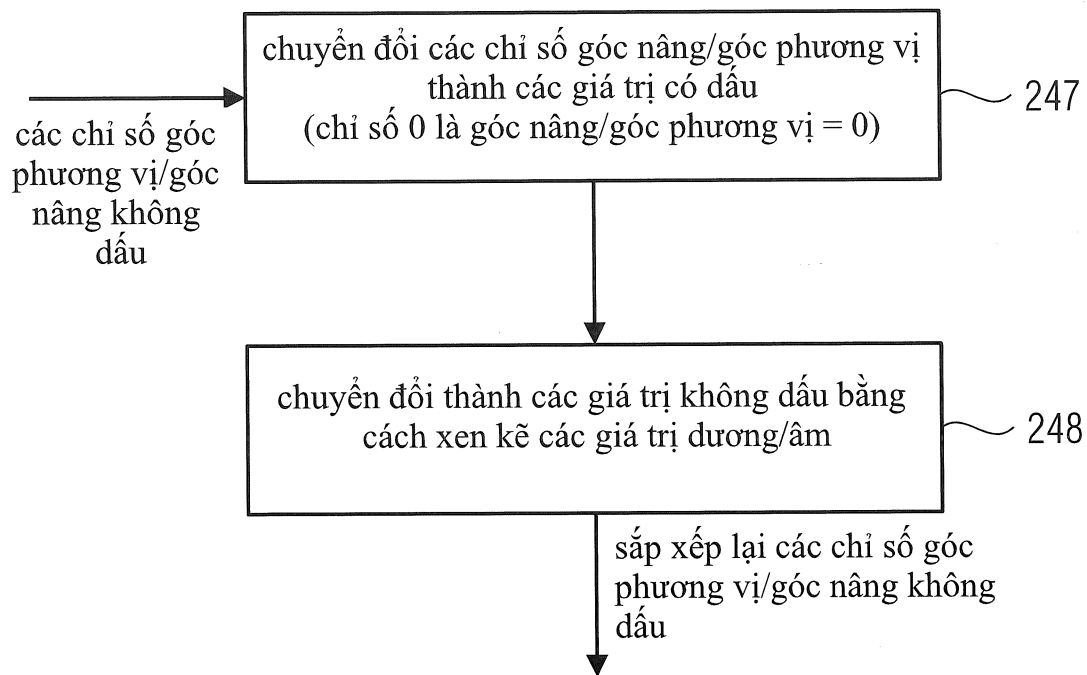


Fig. 5a

14/38

**Fig. 5b**

(tiền xử lý tùy ý các chỉ số hướng cho cả hai chế độ)

15/38

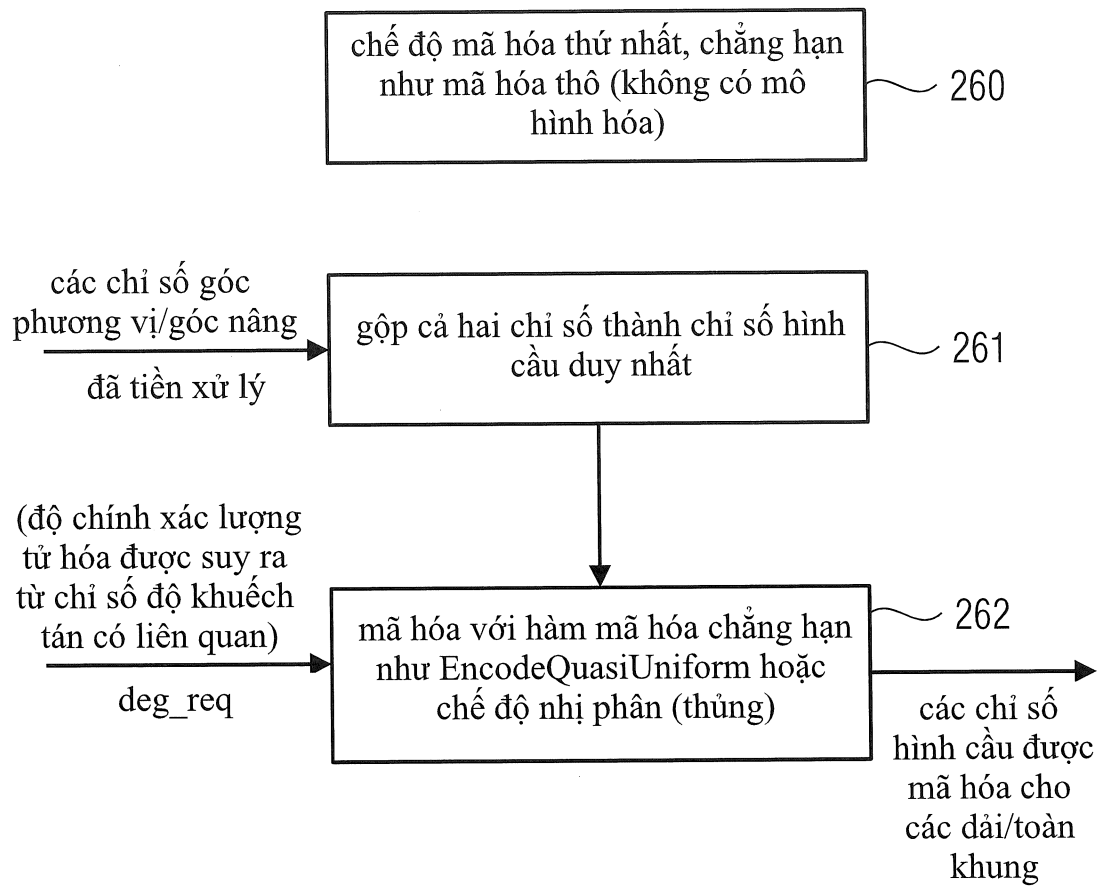


Fig. 5c

16/38

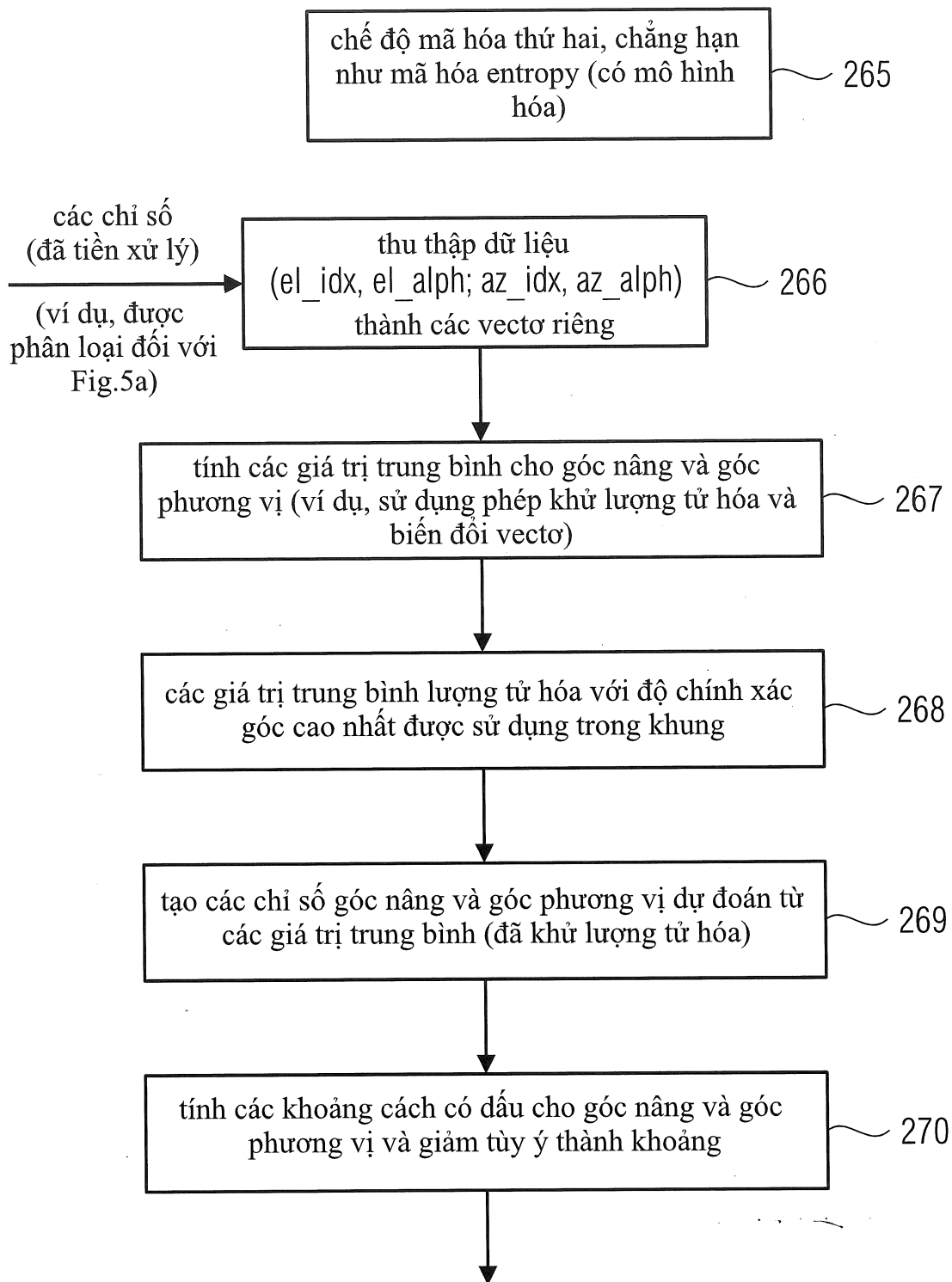


Fig. 5d

(mô hình hóa với phép chiếu cho sự dự đoán)

17/38

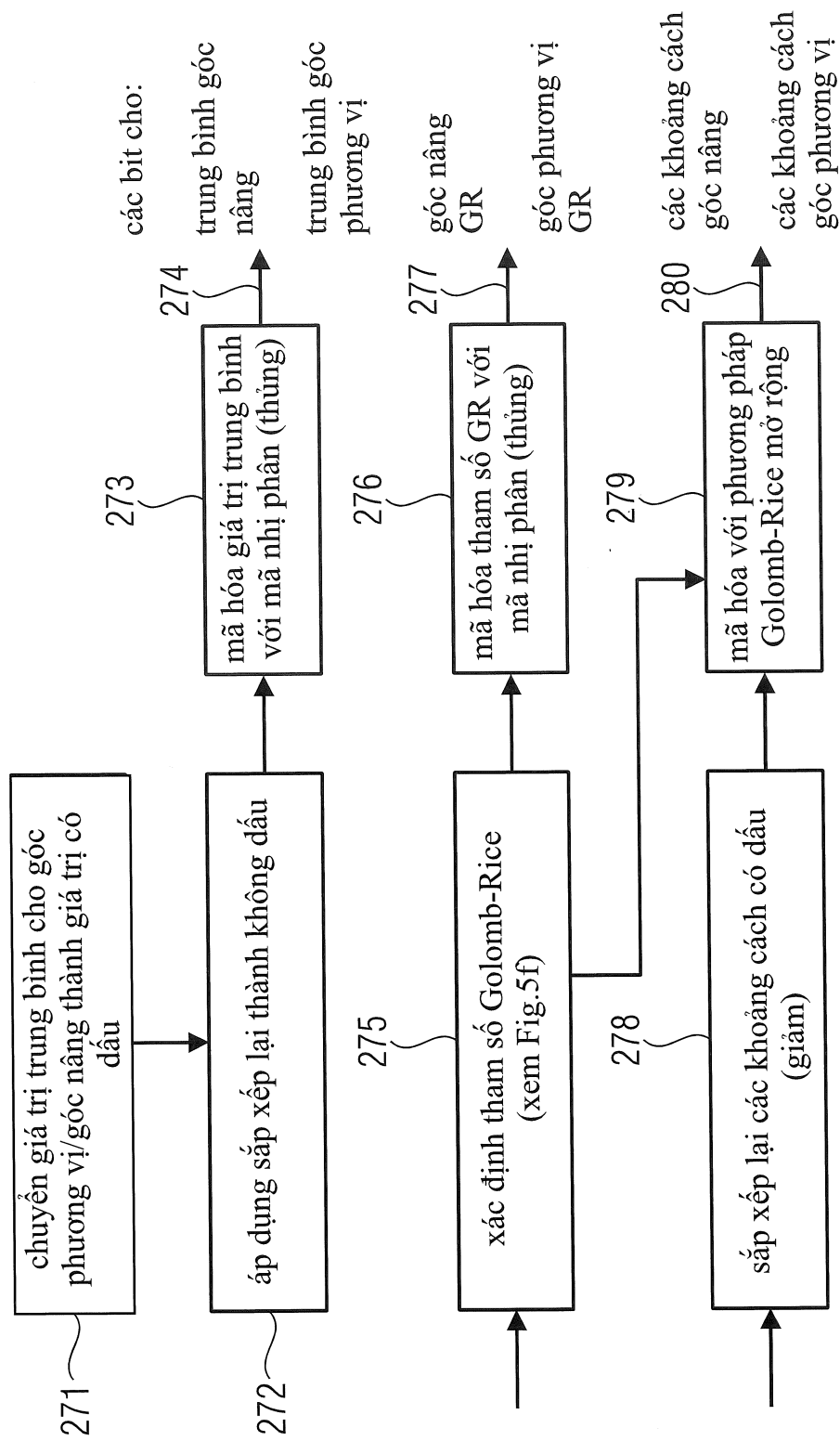


Fig. 5e

mã hóa, tức là chuyển đổi thành các bit cho việc lưu trữ phép truyền)

18/38

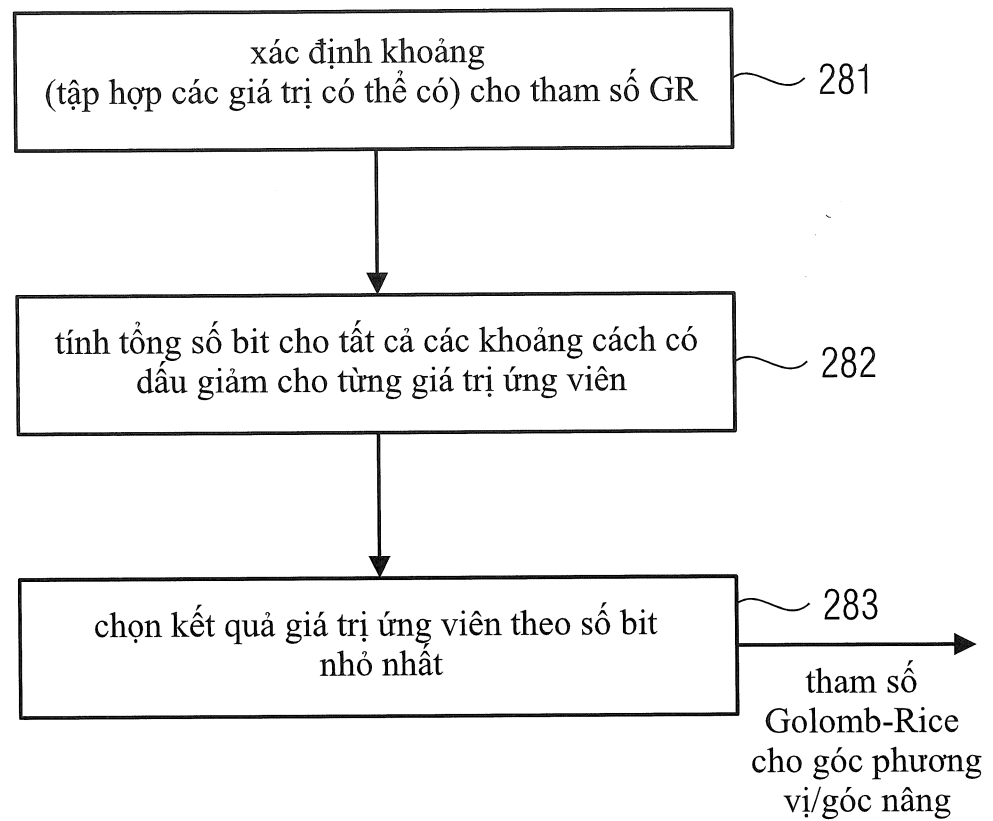


Fig. 5f

19/38

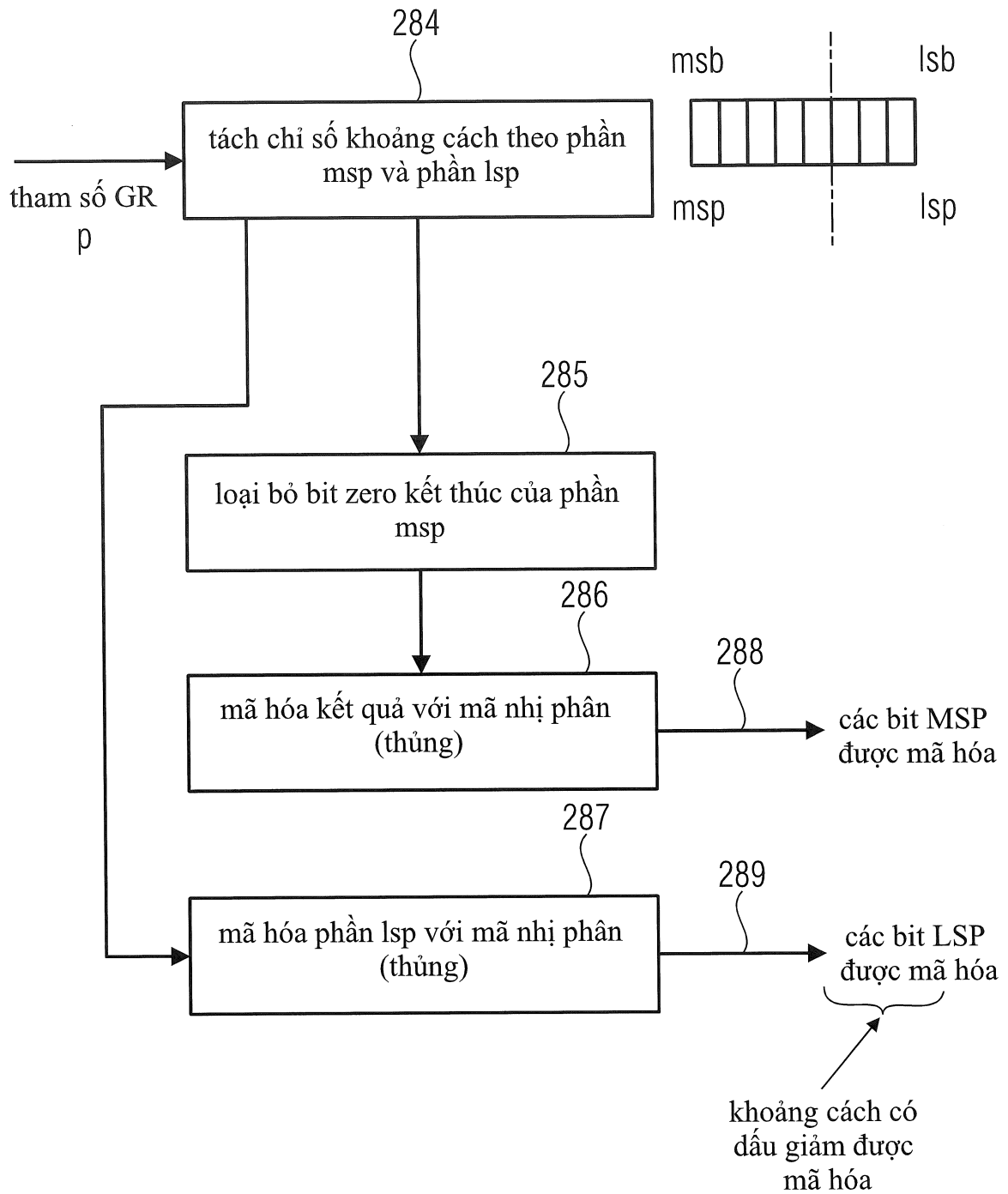


Fig. 5g

20/38

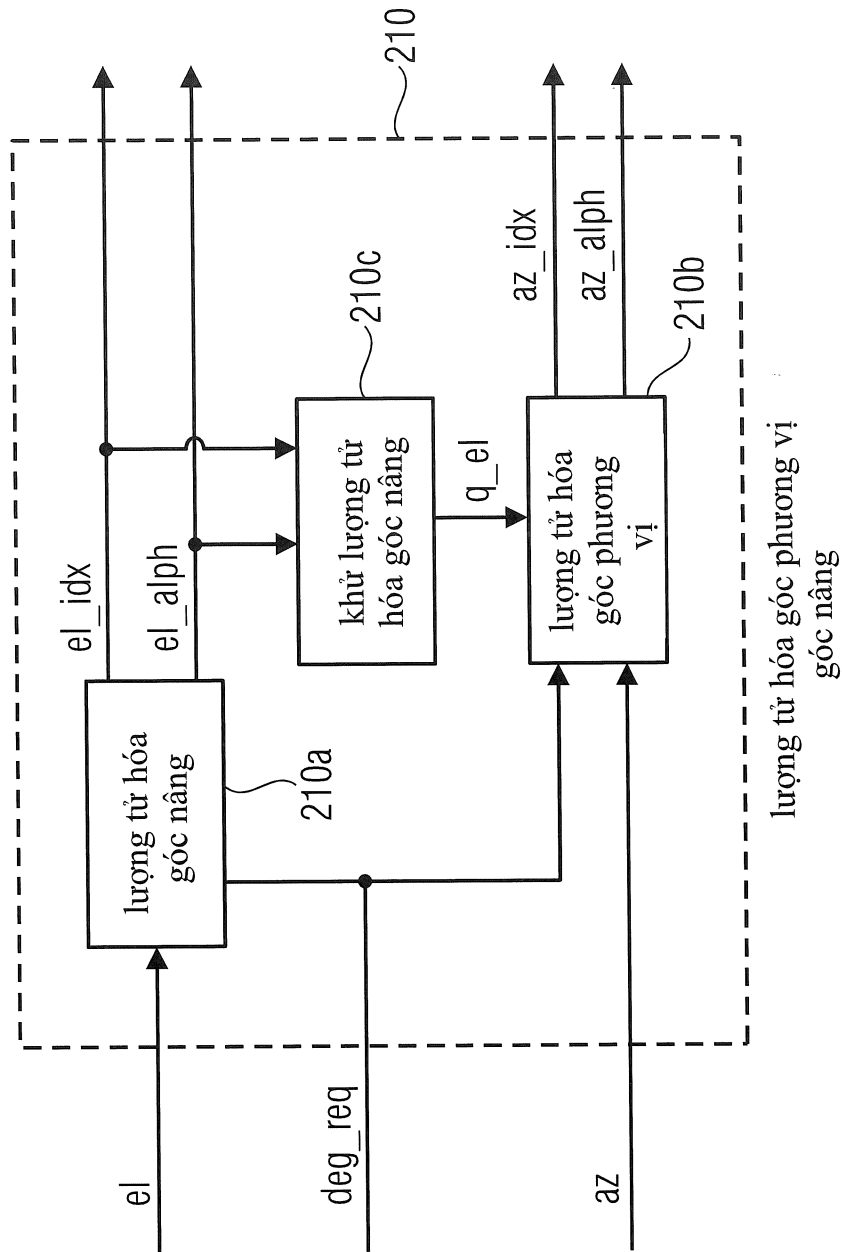


Fig. 6a
Bộ mã hóa

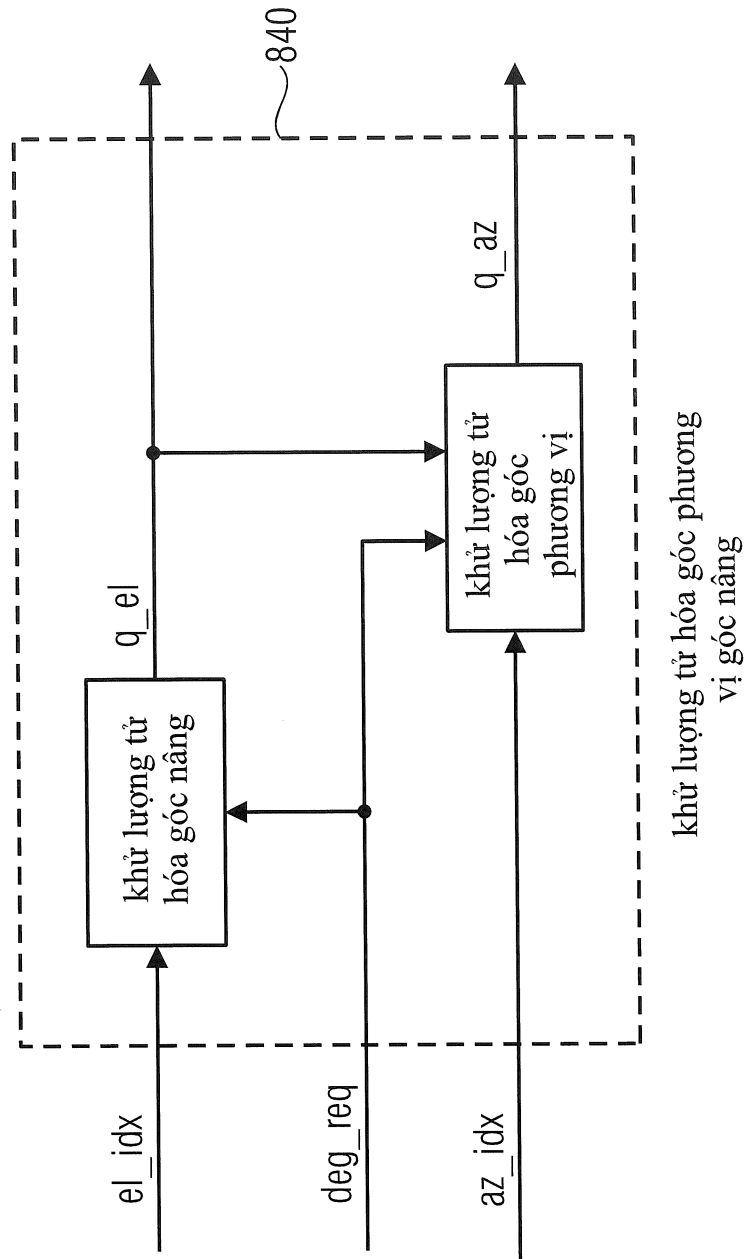


Fig. 6b
Bộ mã hóa/Bộ giải mã

22/38

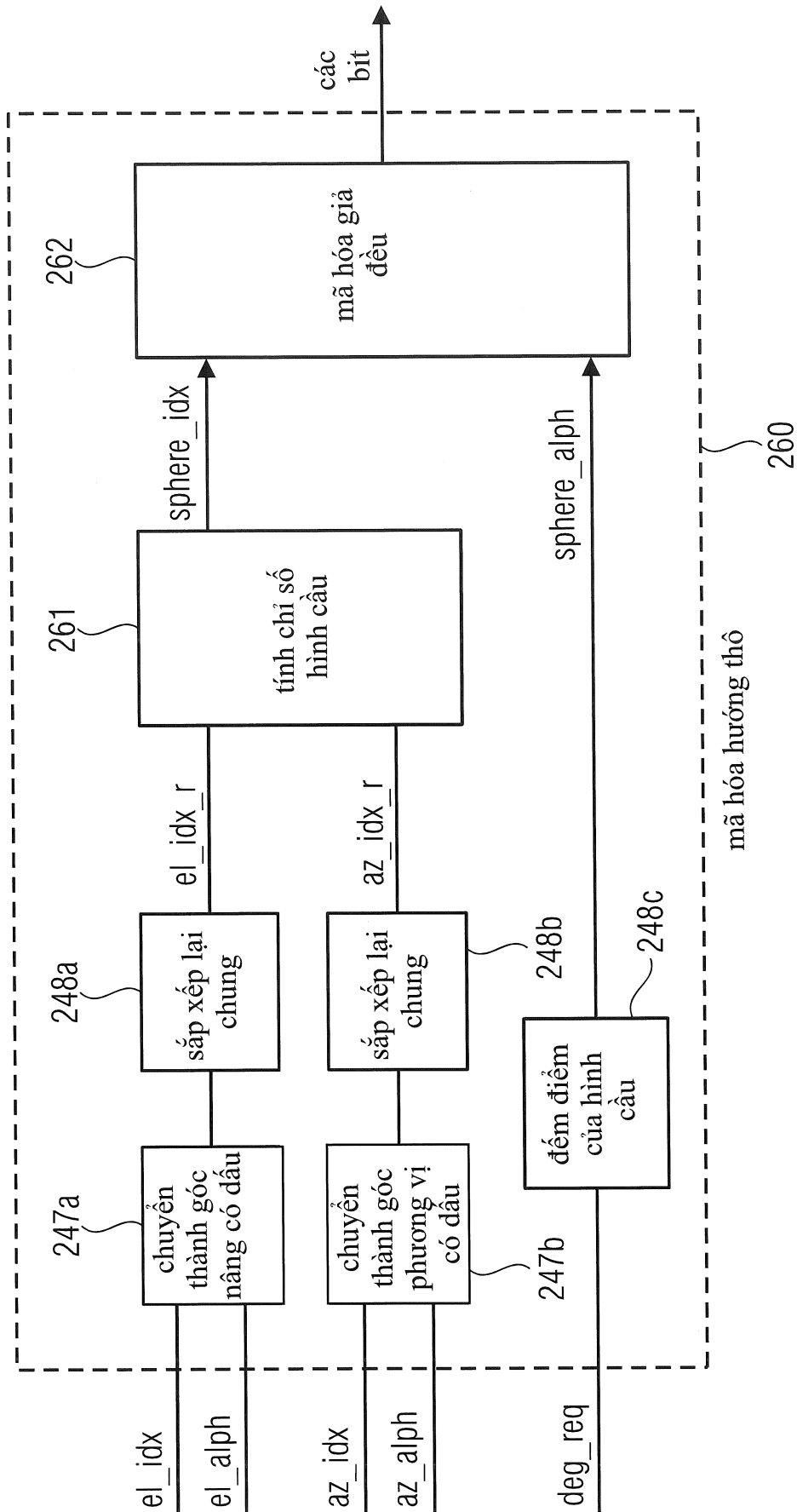
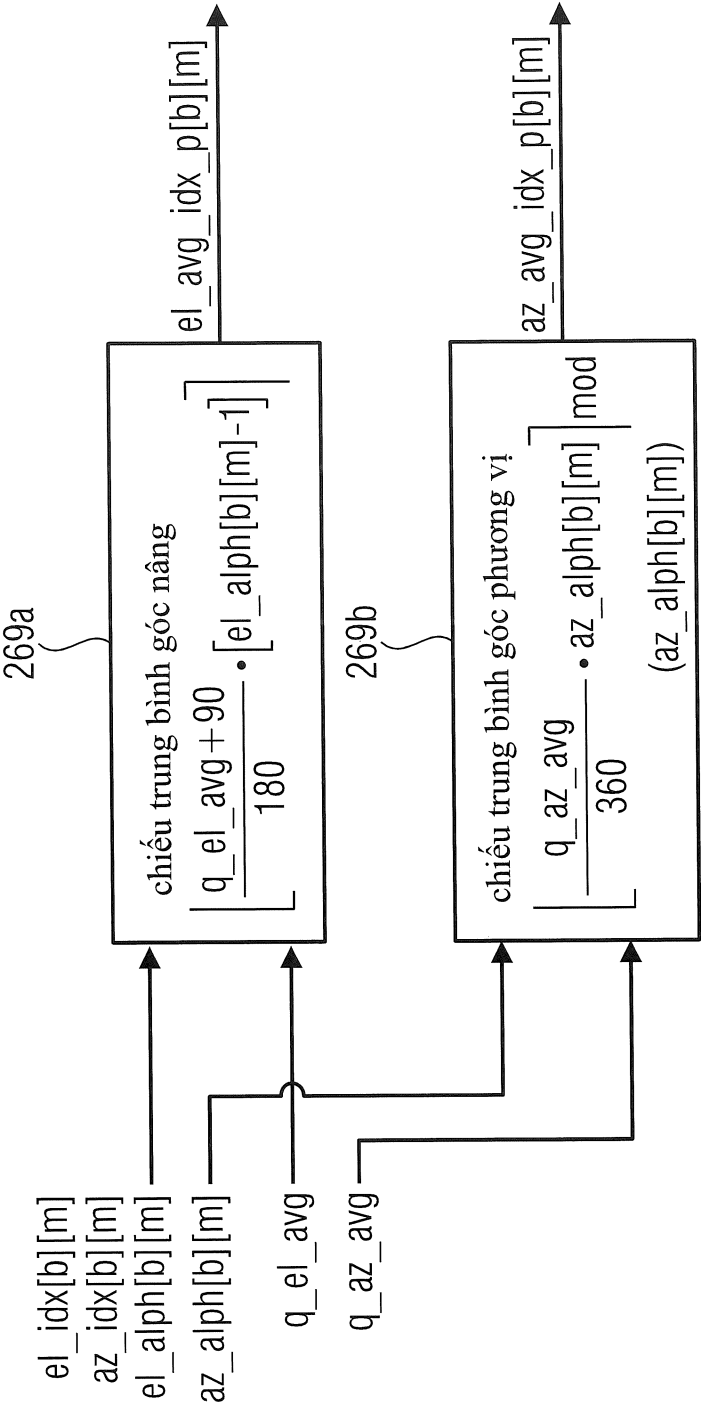


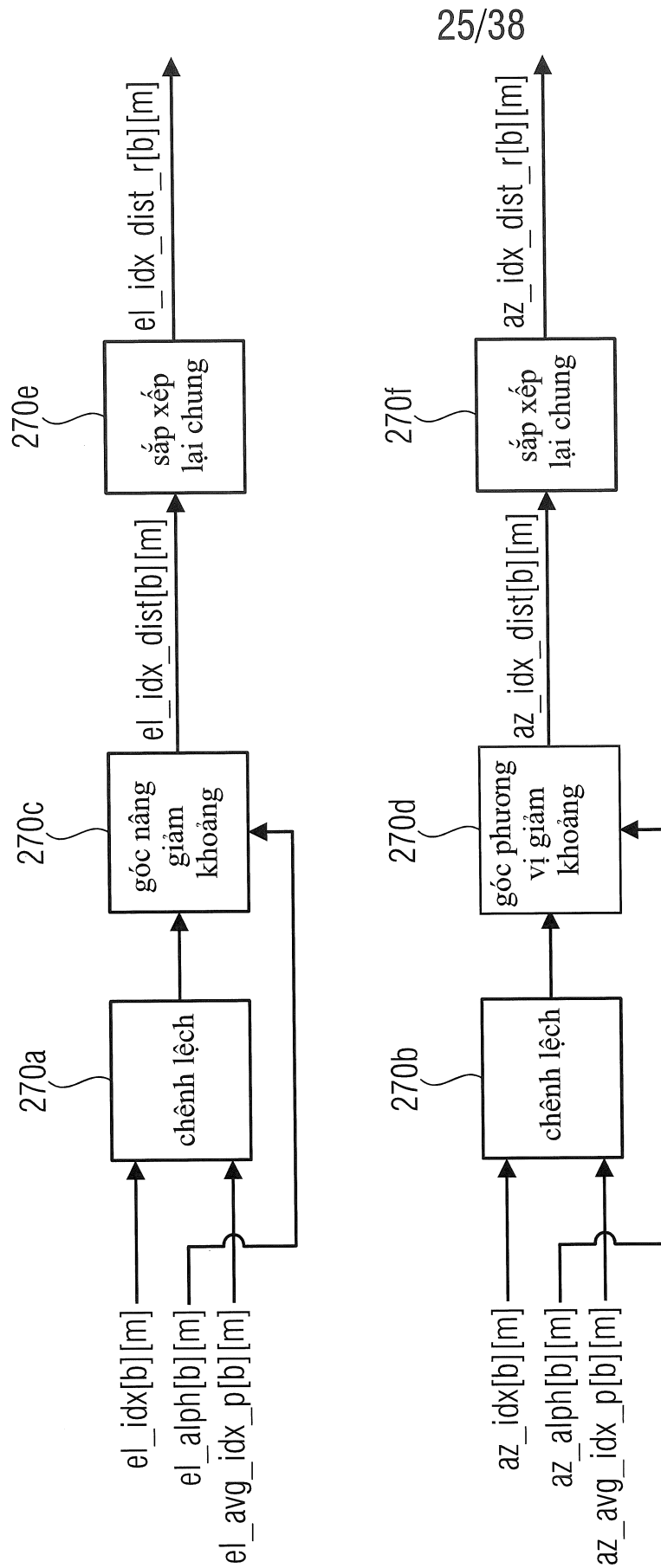
Fig. 6c
Bộ mã hóa



chiều trung bình góc nâng và góc phương vị
(chỉ đối với các hướng với $diff_idx[b] \leq ec_max$)

$$\lfloor \cdot \rfloor = \text{round}(\cdot)$$

Fig. 6e
Bộ mã hóa



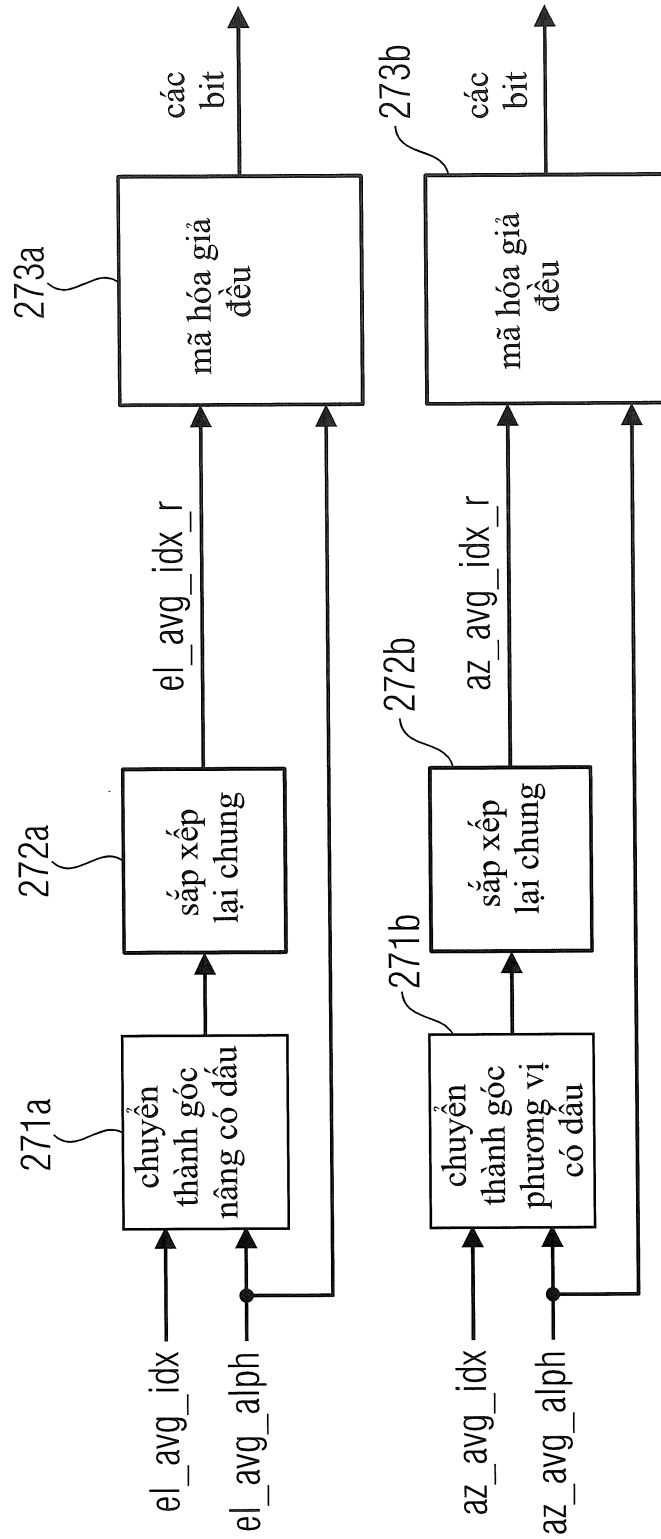
tính các giá trị mà sẽ được mã hóa với Golomb-Rice mở rộng, riêng biệt
cho học nâng và góc phương vị
(chỉ đối với các hướng với $diff_idx[b] \leq ec_max$)

Golomb-Rice mở rộng: xem ví dụ Fig.5a đến 5g

Fig. 6f

Bộ mã hóa

26/38



mã hóa hướng trung bình

Fig. 6g

Bộ mã hóa

27/38

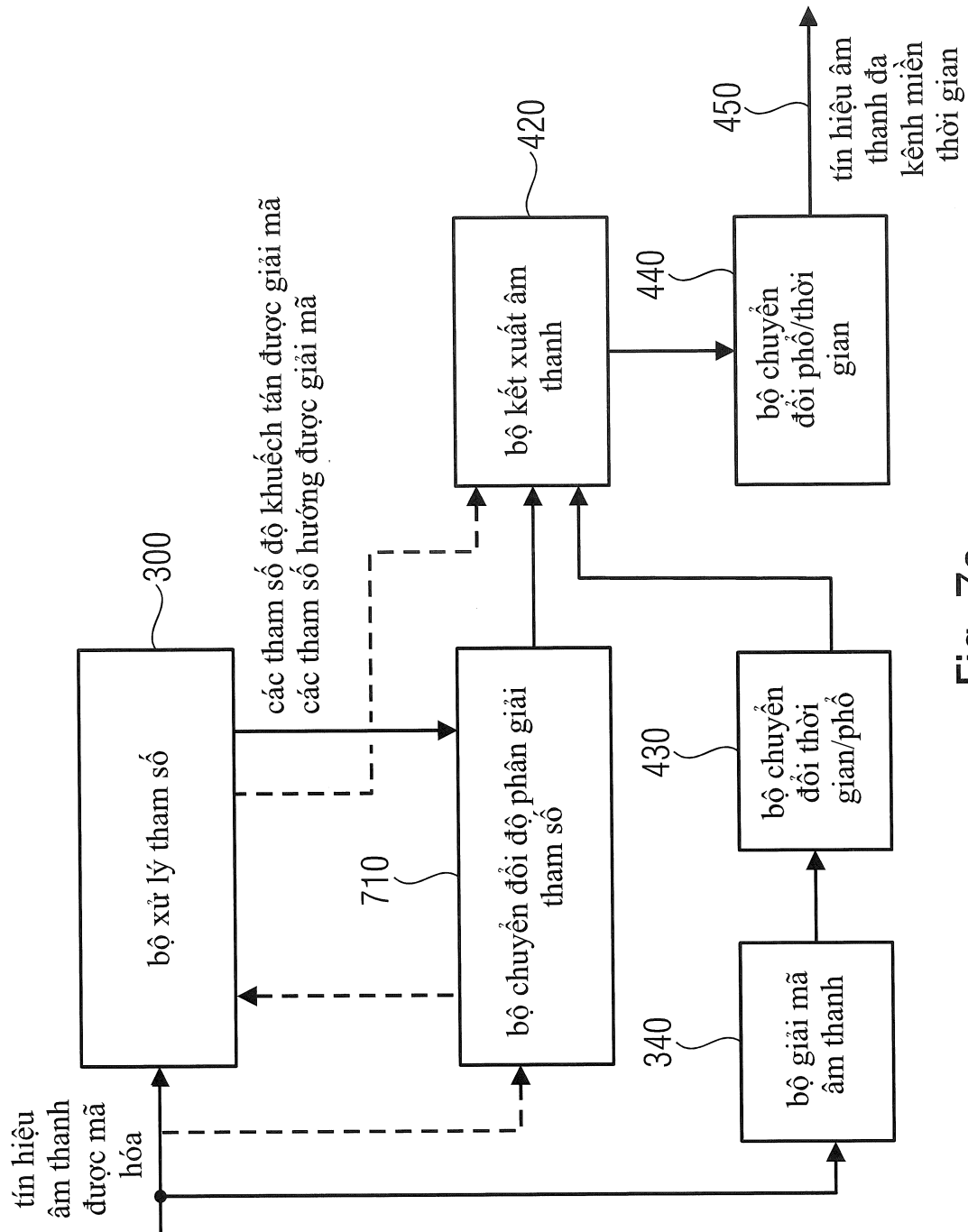


Fig. 7a

28/38

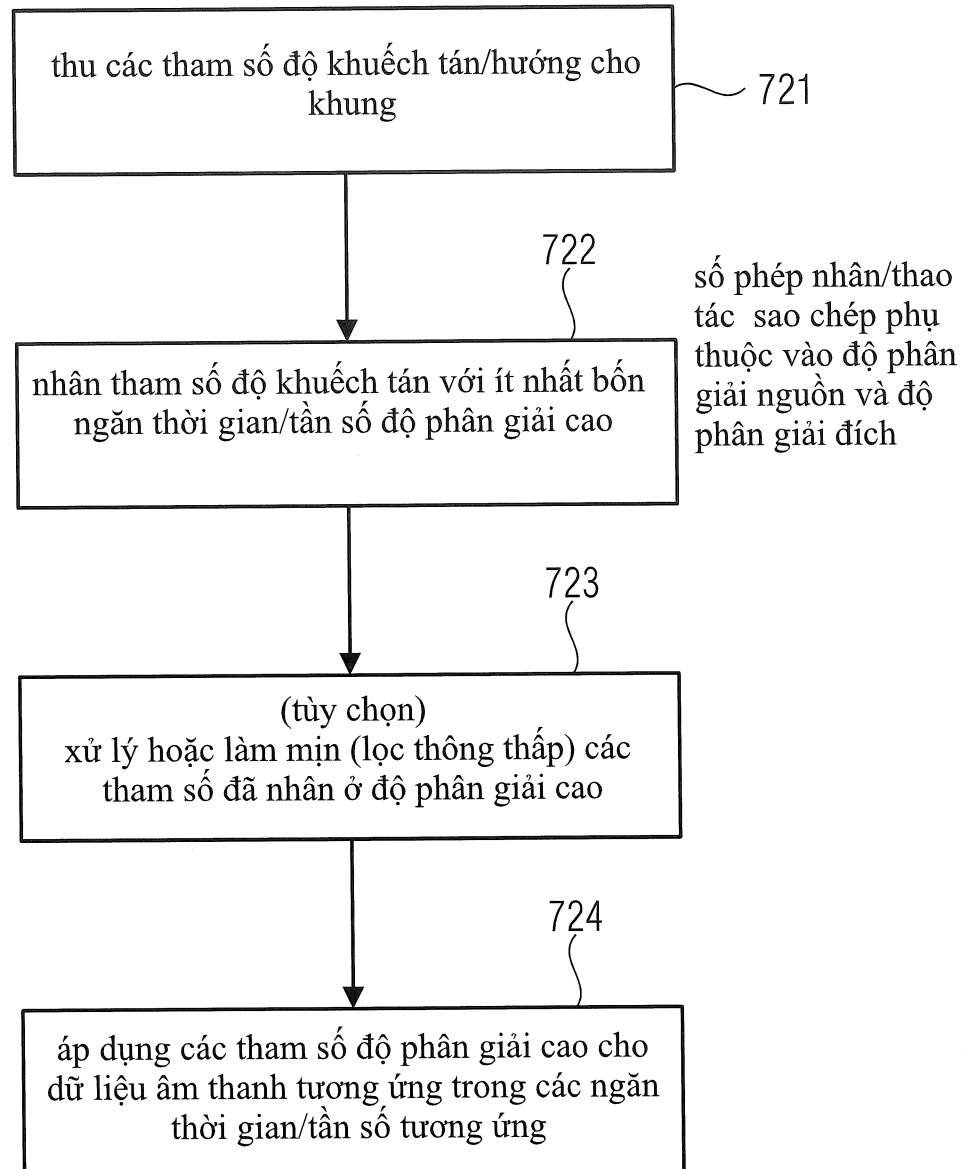


Fig. 7b

29/38

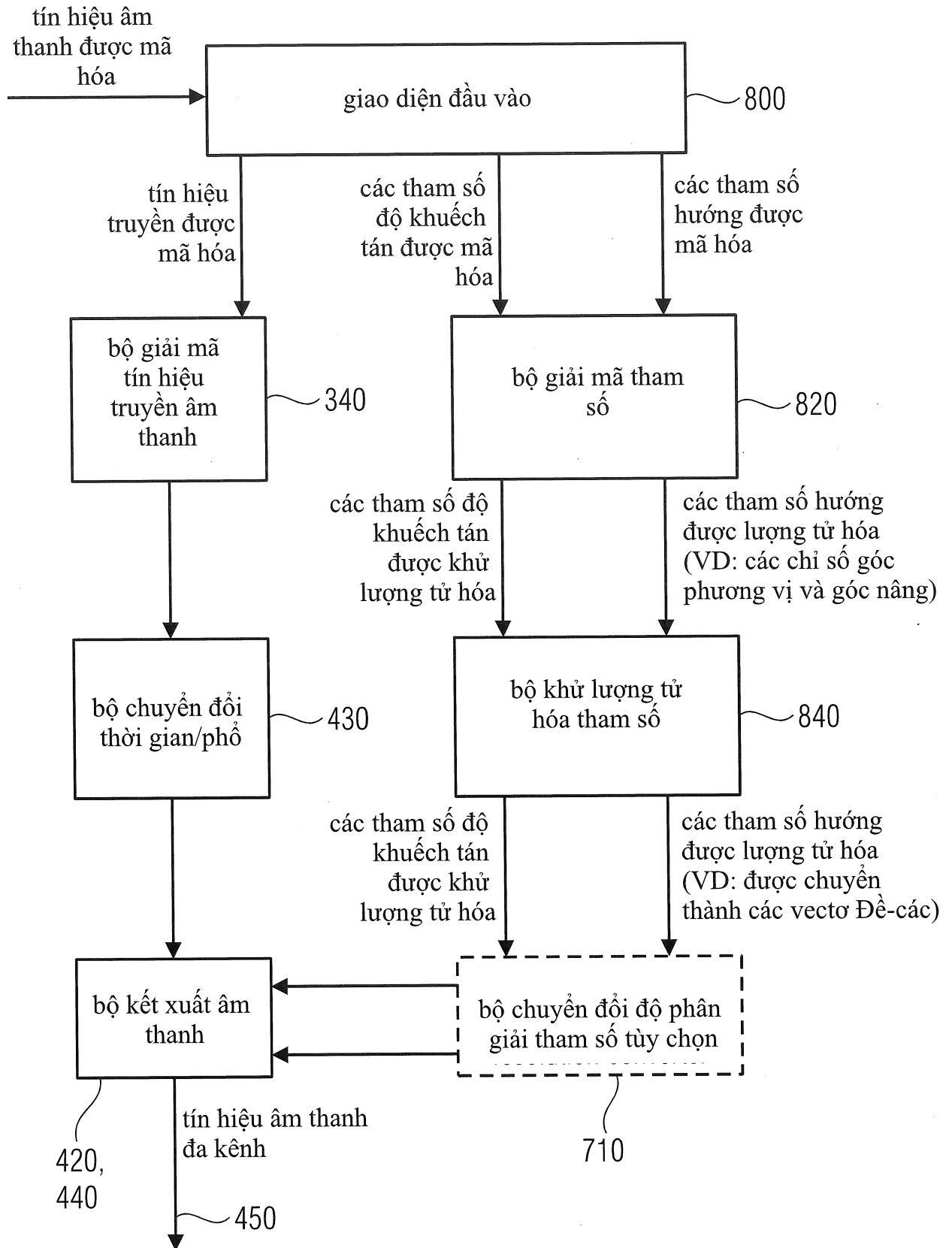


Fig. 8a

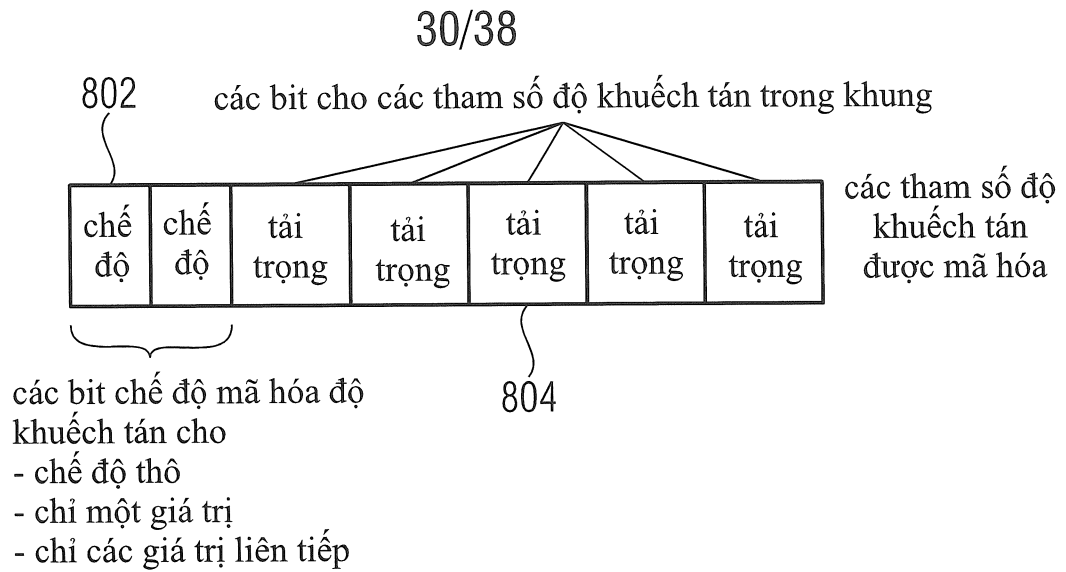


Fig. 8b

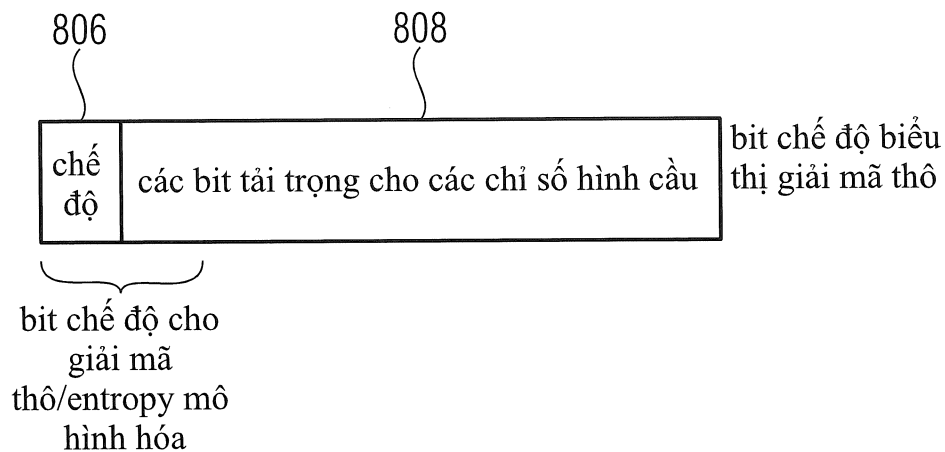


Fig. 8c

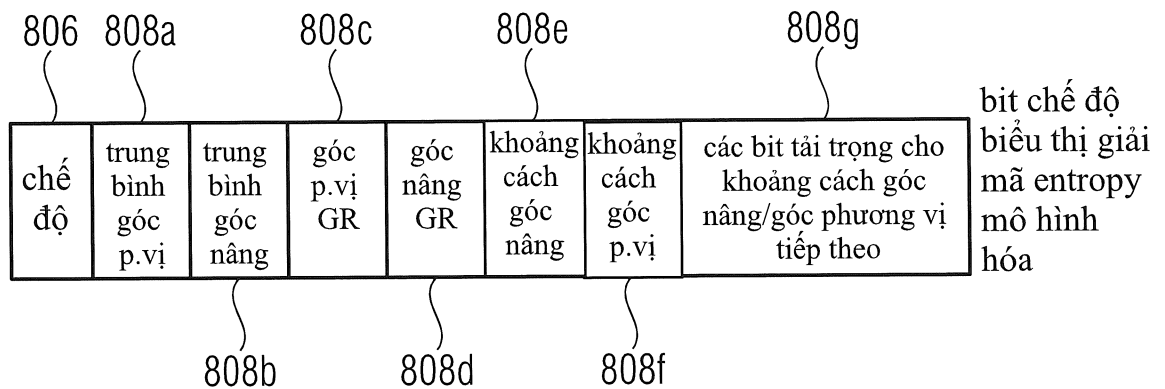


Fig. 8d

31/38

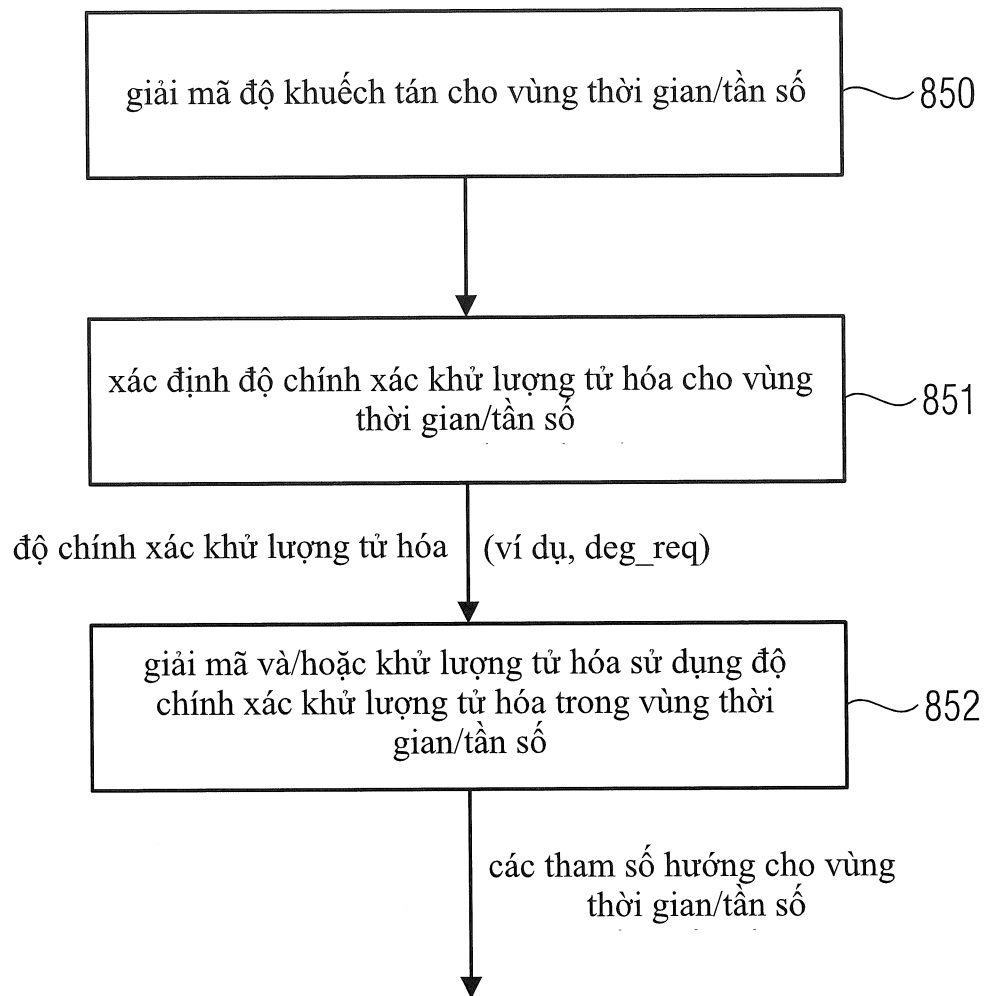


Fig. 8e

32/38

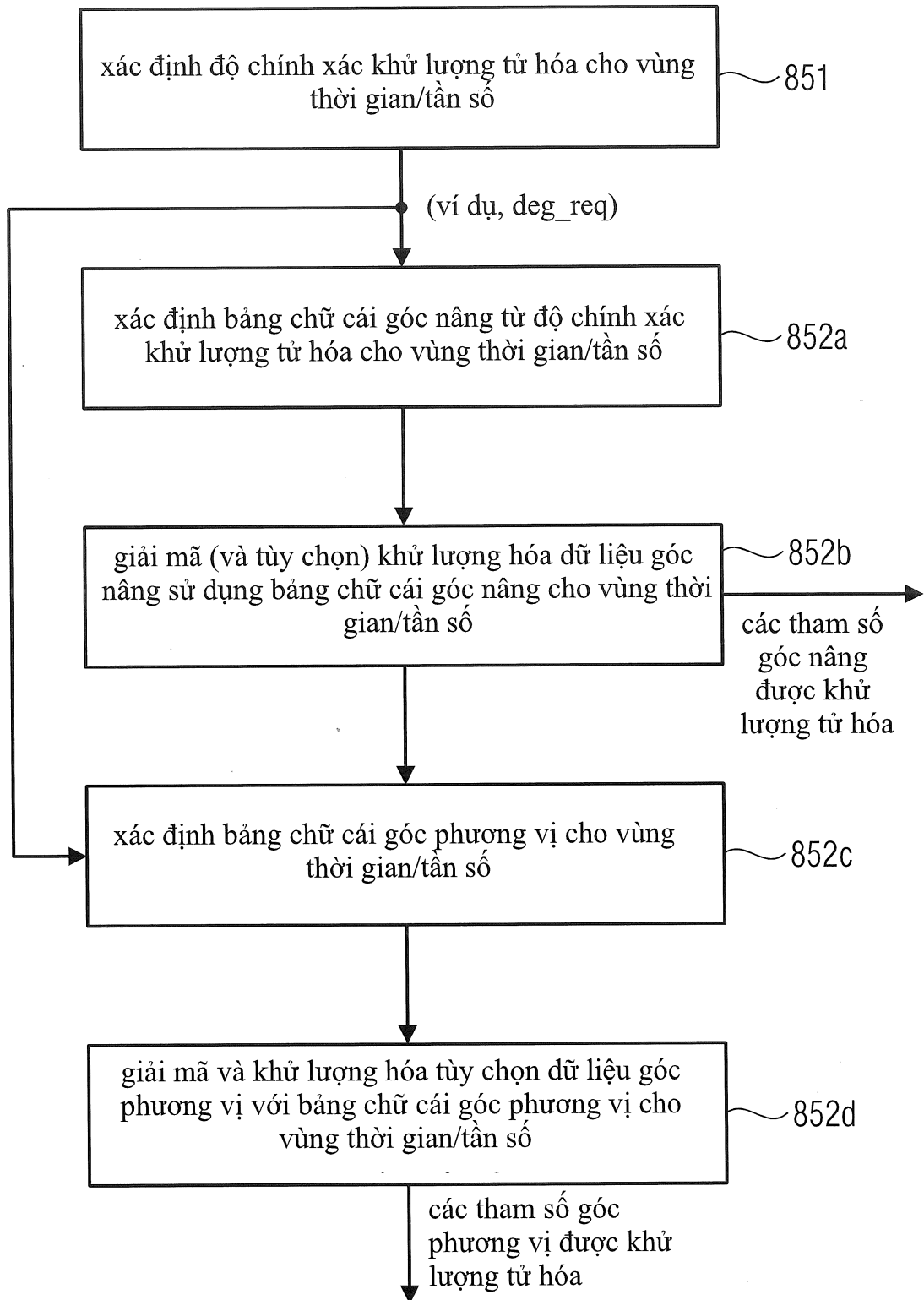


Fig. 8f

33/38

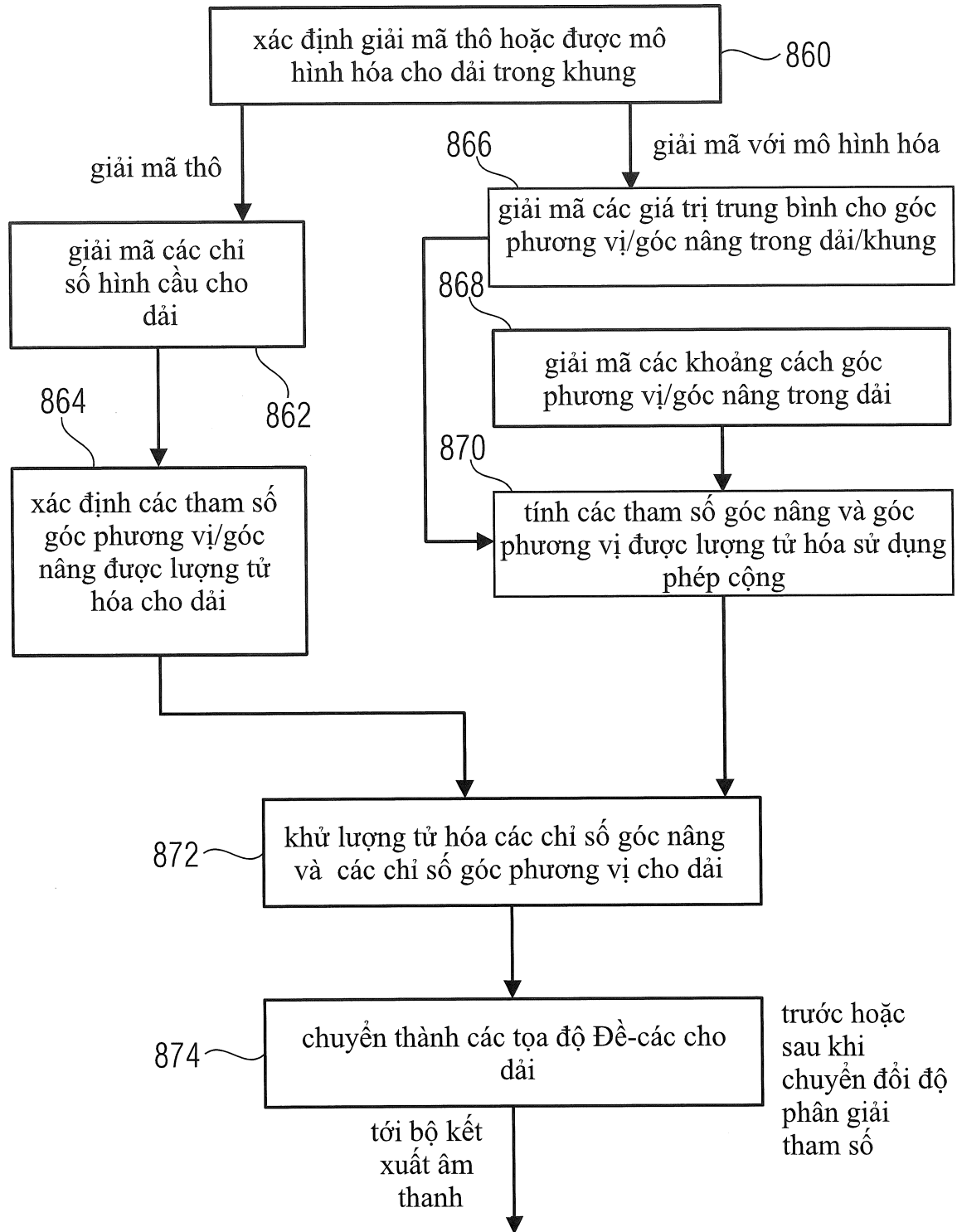
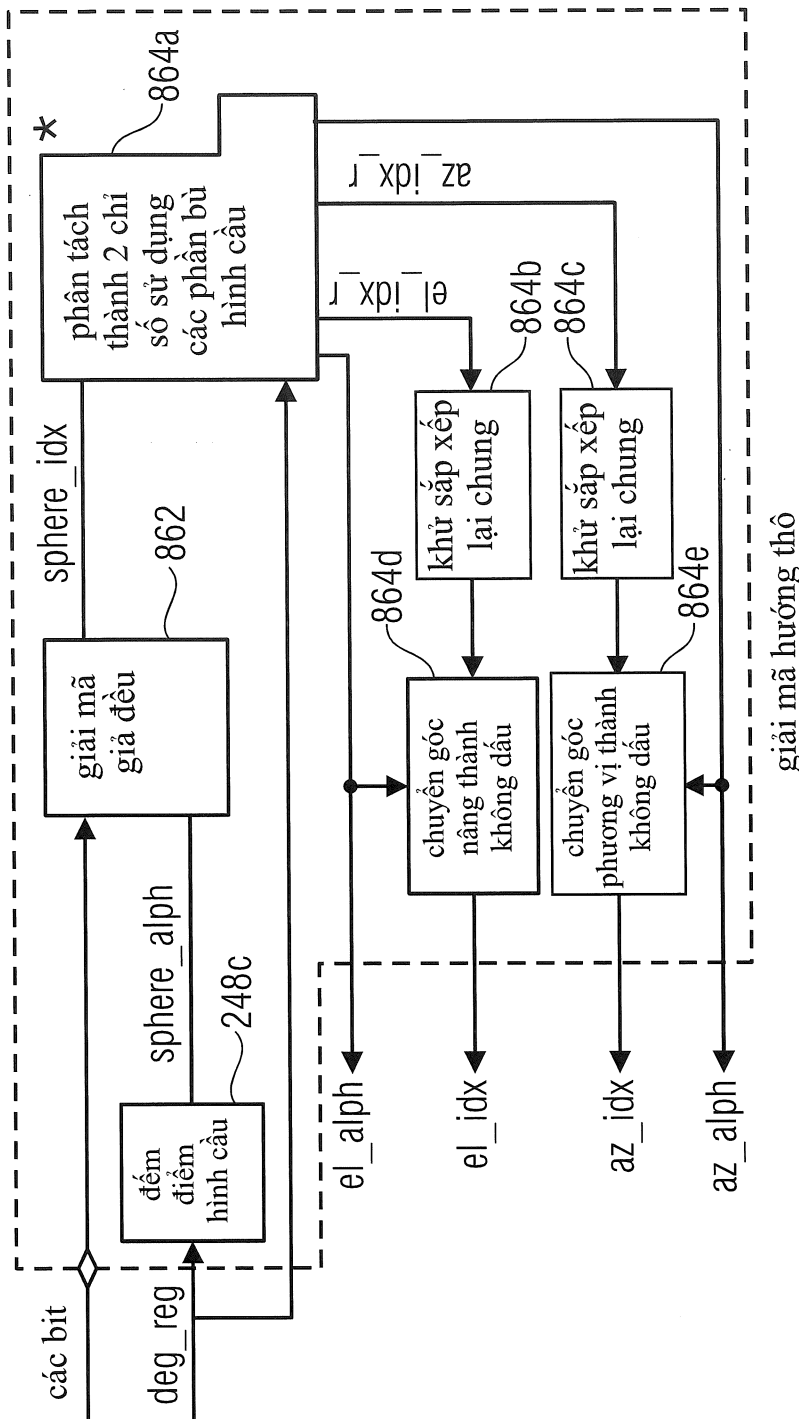


Fig. 8g

34/38



★ $n_points = 90 : \lceil 90 : deg_req \rceil$

$el_alph = 2 \cdot n_points + 1$

el_idx_r được tìm thấy là thỏa mãn mối quan hệ

$A = sphere_offsets(deg_req, el_idx_r) \leq sphere_idx < sphere_offsets(deg_req, el_idx_r + 1) = B$

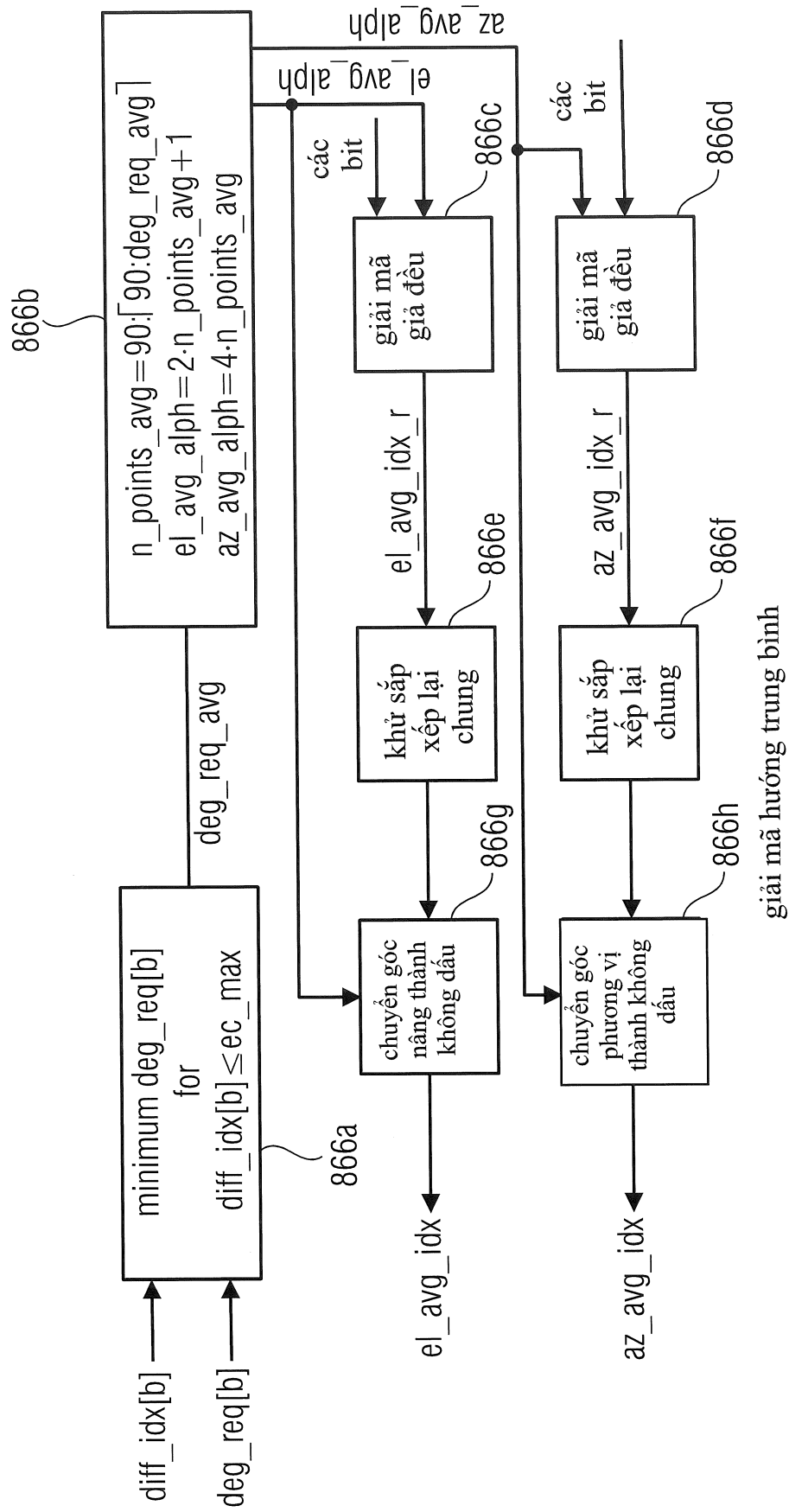
$az_alph = B - A$

$az_idx_r = sphere_idx - A$

Fig. 9a

Bộ giải mã

35/38



36/38

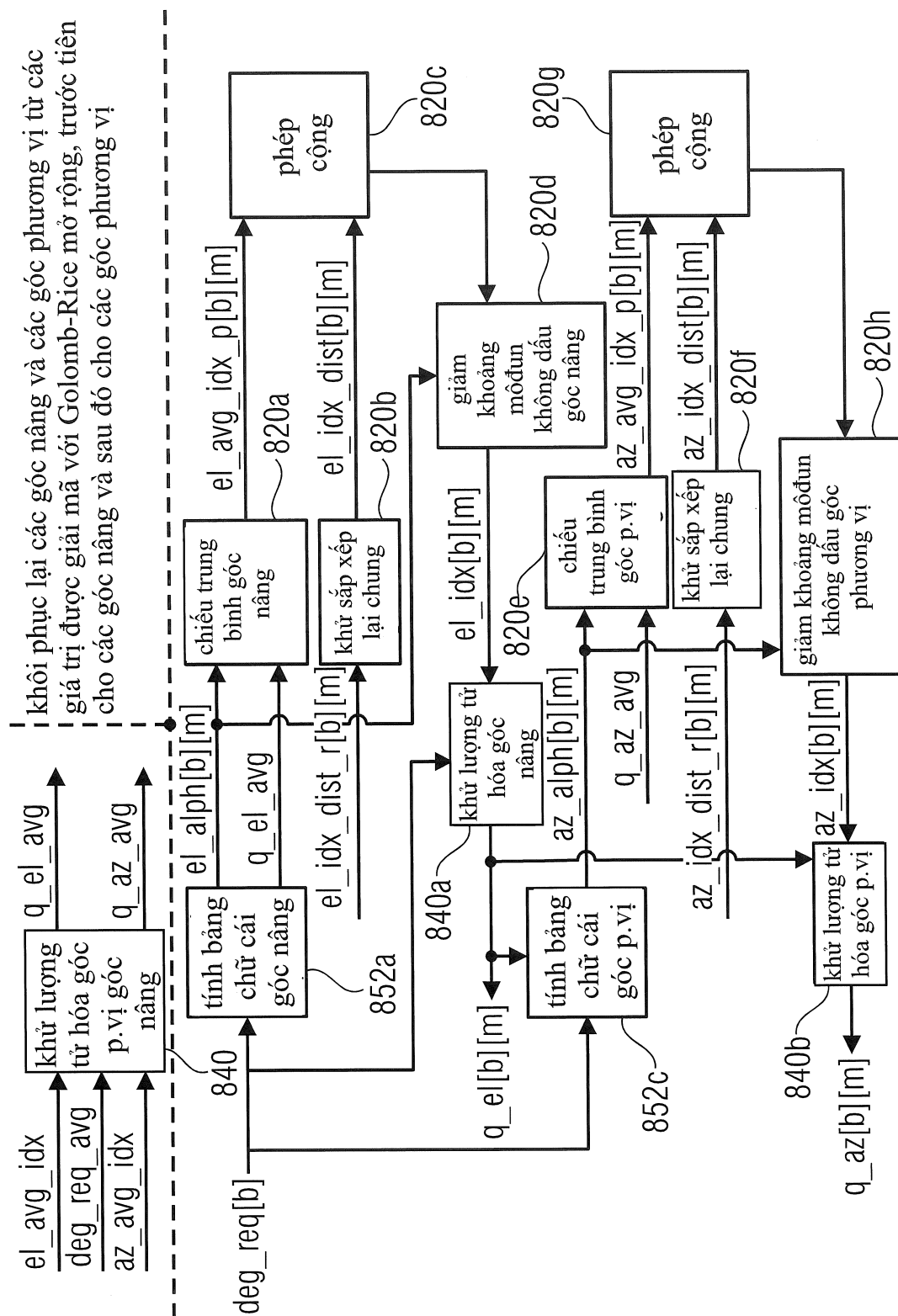


Fig. 9c

Bộ giải mã

37/38

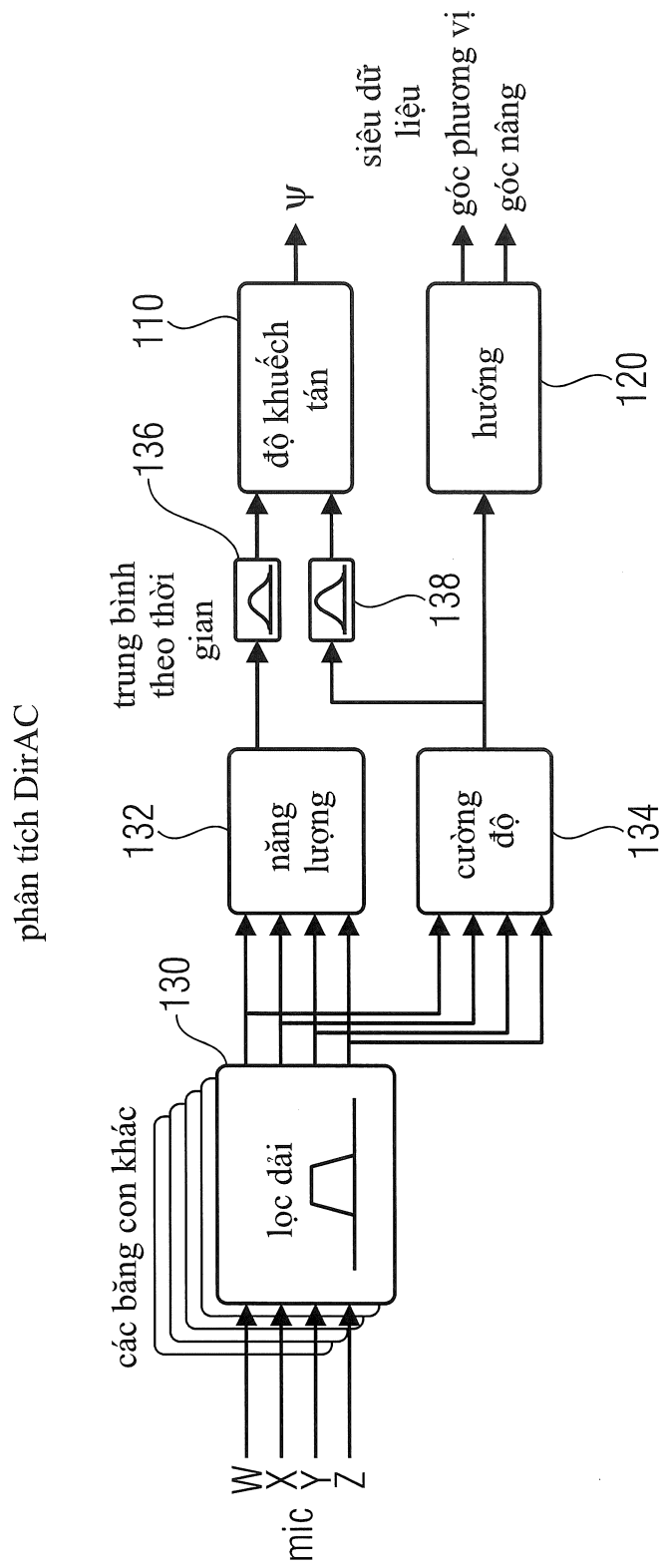


Fig. 10a
(KỸ THUẬT TRƯỚC)

38/38

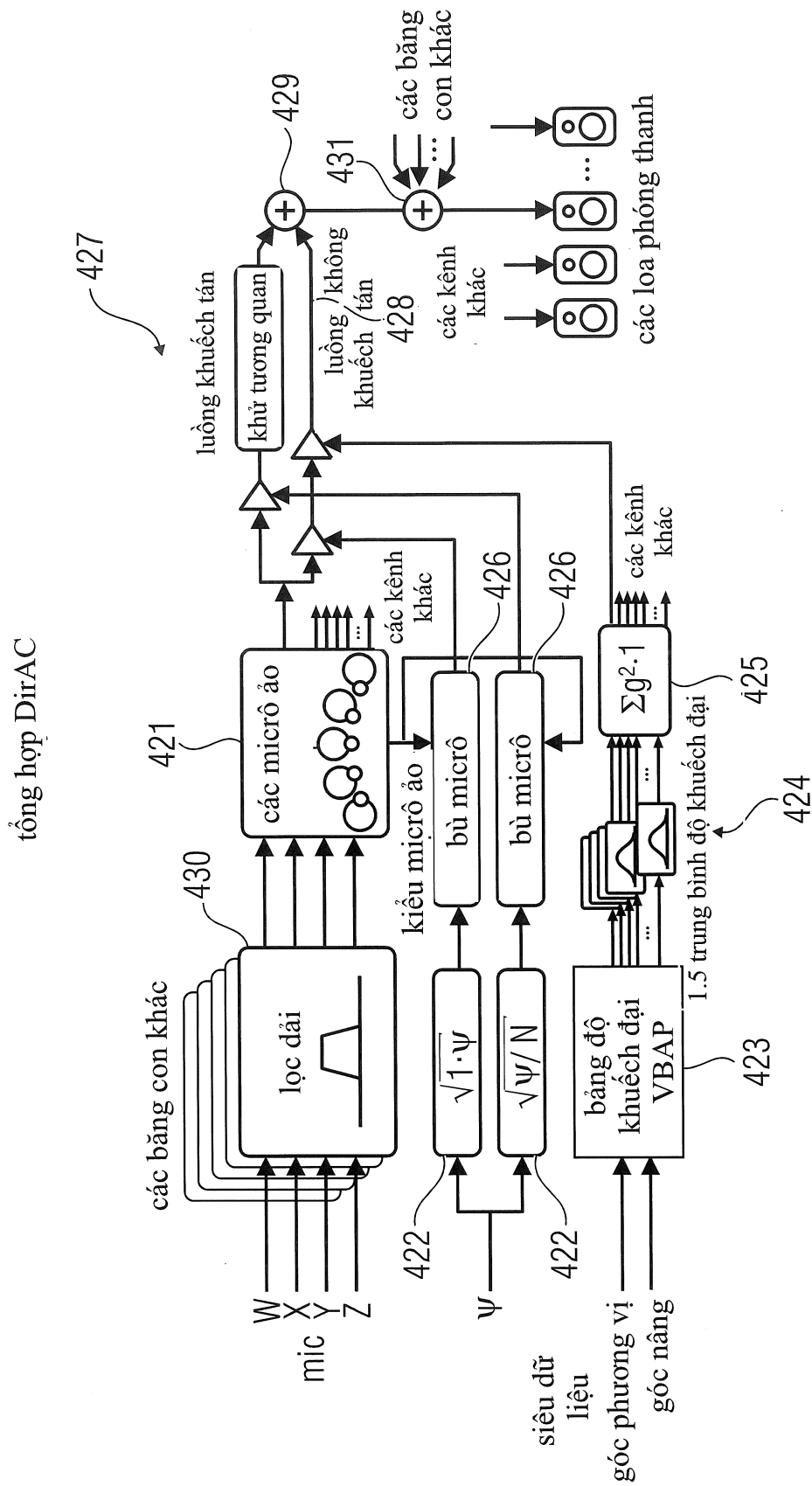


Fig. 10b

(KỸ THUẬT TRƯỚC)