



(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)



1-0043605

(51)^{2020.01} G10L 19/008

(13) B

-
- (21) 1-2021-05137 (22) 21/01/2020
(86) PCT/EP2020/051396 21/01/2020 (87) WO2020/152154 30/07/2020
(30) 19152911.4 21/01/2019 EP
(45) 25/02/2025 443 (43) 25/11/2021 404A
(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)
Hansastrasse 27c, 80686 Muenchen, Germany
(72) KÜCH, Fabian (DE); THIERGART, Oliver (DE); FUCHS, Guillaume (FR);
DÖHLA, Stefan (DE); BOUTHÉON, Alexandre (FR); HERRE, Jürgen (DE);
BAYER, Stefan (AT).
(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS
HANOI)
-

- (54) THIẾT BỊ VÀ PHƯƠNG PHÁP MÃ HÓA SỰ BIỂU DIỄN ÂM THANH KHÔNG
GIAN, THIẾT BỊ VÀ PHƯƠNG PHÁP GIẢI MÃ TÍN HIỆU ÂM THANH ĐƯỢC
MÃ HÓA

(21) 1-2021-05137

(57) Sáng chế đề cập đến thiết bị và phương pháp để mã hóa sự biểu diễn âm thanh không gian, thiết bị và phương pháp để giải mã tín hiệu âm thanh được mã hóa. Thiết bị mã hóa sự biểu diễn âm thanh không gian biểu diễn ngữ cảnh âm thanh để thu được tín hiệu âm thanh được mã hóa, thiết bị bao gồm: bộ tạo sự biểu diễn vận chuyển (600) để tạo ra sự biểu diễn vận chuyển (611) từ sự biểu diễn âm thanh không gian, và để tạo ra siêu dữ liệu vận chuyển (610) về sự tạo ra sự biểu diễn vận chuyển (611) hoặc biểu thị một hoặc nhiều đặc điểm định hướng của sự biểu diễn vận chuyển (611); và giao diện đầu ra (640) để tạo ra tín hiệu âm thanh được mã hóa, tín hiệu âm thanh được mã hóa bao gồm thông tin về sự biểu diễn vận chuyển (611), và thông tin về siêu dữ liệu vận chuyển (610).

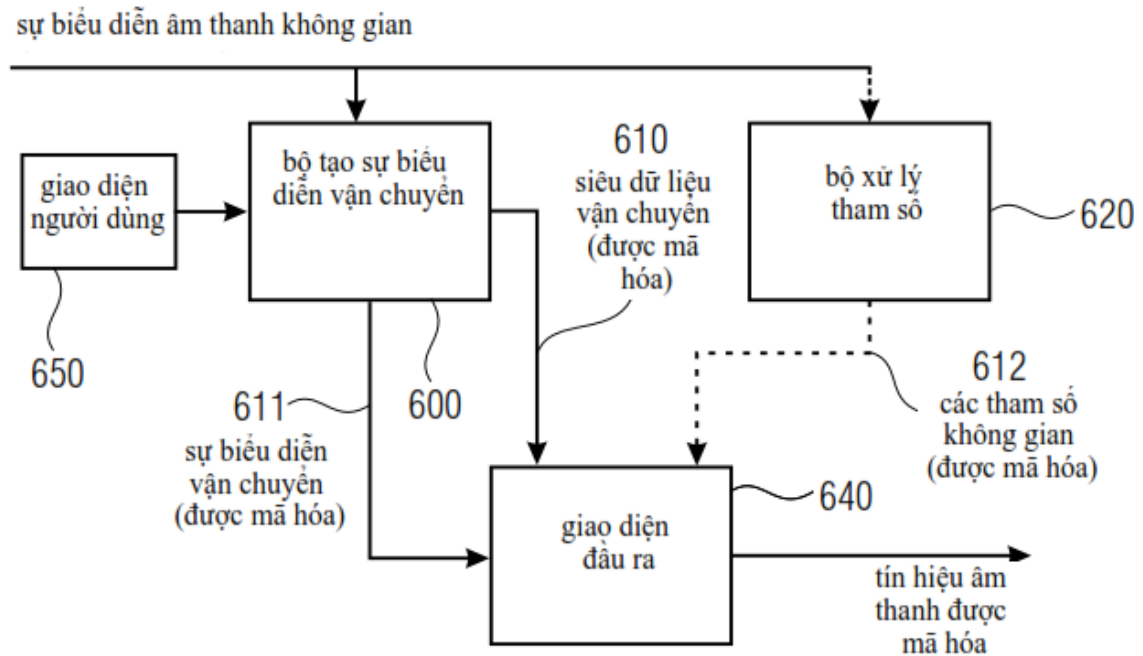


Fig. 6

Lĩnh vực kỹ thuật được đề cập

Các phương án theo sáng chế liên quan đến kênh vận chuyển hoặc sự tạo tín hiệu trộn giảm cho sự mã hóa âm thanh định hướng.

Tình trạng kỹ thuật của sáng chế

Công nghệ mã hóa âm thanh định hướng (Directional Audio Coding - DirAC) [Pulkki07] là phương pháp hiệu quả để phân tích và tái tạo âm không gian. DirAC sử dụng sự biểu diễn được kích thích cảm quan của trường âm trên cơ sở các tham số không gian, tức là, hướng tới (direction of arrival - DOA) và độ khuếch tán được đo mỗi băng tần. Được dựng dựa trên giả định rằng tại một nấc thời gian và tại một băng tới hạn, độ phân giải không gian của hệ thống thính giác được giới hạn để giải mã một tín hiệu điều khiển cho hướng và tín hiệu điều khiển khác cho sự đồng nhất thính giác với thính giác. Âm không gian khi đó được biểu diễn trong miền tần số bằng sự tăng giảm chéo hai dòng: dòng khuếch tán không định hướng và dòng không khuếch tán định hướng.

DirAC ban đầu được dự định cho âm định dạng B được ghi nhưng cũng có thể được mở rộng cho các tín hiệu micro phù hợp với thiết lập loa phát thanh cụ thể như 5.1 [2] hoặc cấu hình bất kỳ của các mảng micro [5]. Trong trường hợp mới nhất, độ linh hoạt hơn có thể đạt được bằng cách ghi các tín hiệu không dành cho thiết lập loa phát thanh cụ thể, mà thay vào đó ghi các tín hiệu của định dạng trung gian.

Định dạng trung gian như vậy, cũng được đưa ra trong thực tế, được biểu diễn bởi Ambisonics (thứ tự cao) [3]. Từ tín hiệu Ambisonics, người ta có thể tạo ra các tín hiệu cho mỗi thiết lập loa phát thanh mong muốn bao gồm các tín hiệu nhị phân cho sự tái tạo ống nghe. Điều này yêu cầu bộ kết xuất cụ thể mà được áp dụng cho tín hiệu Ambisonics, sử dụng bộ kết xuất Ambisonics tuyến tính [3] hoặc bộ kết xuất tham số như mã hóa âm thanh định hướng (Directional Audio Coding - DirAC).

Tín hiệu Ambisonics có thể được biểu diễn như tín hiệu đa kênh trong đó mỗi kênh (được đề cập đến như các thành phần Ambisonics) tương đương với hệ số của hàm được

gọi là hàm cơ bản không gian. Với tổng được gán trọng số của các hàm cơ bản không gian này (với các trọng số tương ứng với các hệ số), người ta có thể tái tạo trường âm ban đầu trong vị trí ghi [3]. Do đó, các hệ số hàm cơ bản không gian (tức là, các thành phần Ambisonics) biểu diễn sự mô tả thu gọn của trường âm trong vị trí ghi. Hiện có các loại hàm cơ bản không gian khác nhau, ví dụ, sóng hàm hình cầu (spherical harmonics - SH) [3] hoặc sóng hàm hình trụ (cylindrical harmonics - CH) [3]. Các CH có thể được sử dụng khi mô tả trường âm trong không gian 2D (ví dụ, sự tái tạo âm 2D) trong khi các SH có thể được sử dụng để mô tả trường âm trong không gian 2D và 3D (ví dụ, cho sự tái tạo âm 2D và 3D).

Như ví dụ, tín hiệu âm thanh $f(t)$ mà tới từ hướng nhất định (φ, θ) dẫn đến tín hiệu âm thanh $f(\varphi, \theta, t)$ có thể được biểu diễn ở định dạng Ambisonics bằng cách mở rộng các sóng hài hình cầu lên đến thứ tự cắt xén H :

$$f(\varphi, \theta, t) = \sum_{l=0}^H \sum_{m=-l}^{+l} Y_l^m(\varphi, \theta) \phi_{lm}(t)$$

trong đó $Y_l^m(\varphi, \theta)$ là sóng hài hình cầu có thứ tự l và chế độ m , và $\phi_{lm}(t)$ là các hệ số mở rộng. Với thứ tự cắt xén H tăng, sự mở rộng dẫn đến sự biểu diễn không gian chính xác hơn. Các sóng hài hình cầu lên đến thứ tự $H = 4$ với chỉ số lấy số kênh Ambisonics (Ambisonics Channel Numbering - ACN) được minh họa trên Fig.1a cho thứ tự n và chế độ m .

DirAC đã được mở rộng để phân phối các tín hiệu Ambisonics thứ tự cao hơn từ tín hiệu Ambisonics thứ tự thứ nhất (FOA được gọi là định dạng B) hoặc từ các mảng micro khác nhau [5]. Tài liệu này tập trung vào cách hiệu quả hơn để tổng hợp các tín hiệu Ambisonics thứ tự cao hơn từ các tham số DirAC và tín hiệu tham chiếu. Trong tài liệu này, tín hiệu tham chiếu, cũng được đề cập đến như tín hiệu trộn giảm, được xem xét như tập con của tín hiệu Ambisonics thứ tự cao hơn hoặc tổ hợp tuyến tính của tập con của các thành phần Ambisonics.

Trong sự phân tích DirAC, các tham số không gian của DirAC được ước lượng từ các tín hiệu đầu vào âm thanh. Ban đầu, DirAC đã được phát triển cho đầu vào Ambisonics thứ tự thứ nhất (first-order Ambisonics - FOA) mà có thể, ví dụ, thu được

từ các micro định dạng B, tuy nhiên, các tín hiệu đầu vào khác cũng khả thi. Trong sự tổng hợp DirAC, các tín hiệu đầu ra cho sự tái tạo không gian, ví dụ, các tín hiệu loa phát thanh, được tính toán từ các tham số DirAC và các tín hiệu âm thanh được kết hợp. Các giải pháp đã được mô tả để sử dụng tín hiệu âm thanh đẳng hướng chỉ cho sự tổng hợp hoặc sử dụng toàn bộ tín hiệu FOA [Pulkki07]. Cách khác, chỉ tập con của bốn thành phần tín hiệu FOA có thể được sử dụng cho sự tổng hợp.

Do sự biểu diễn hệ số của âm không gian DirAC cũng phù hợp như cơ bản cho các hệ thống mã hóa âm thanh không gian. Mục đích của hệ thống này là để mã hóa ngữ cảnh âm thanh không gian tại các tốc độ bit thấp và để tái tạo ngữ cảnh âm thanh ban đầu càng chính xác càng tốt sau khi truyền dẫn. Trong trường hợp này, sự phân tích DirAC theo sau bởi 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 tham số tốc độ bit thấp. Cùng với siêu dữ liệu, tín hiệu trộn giảm được suy ra từ các tín hiệu đầu vào âm thanh ban đầu được mã hóa cho sự truyền dẫn bởi bộ mã hóa lõi âm thanh ban đầu. Ví dụ, bộ mã hóa âm thanh trên cơ sở EVS có thể được thích ứng để mã hóa tín hiệu trộn giảm. Tín hiệu trộn giảm bao gồm các kênh khác nhau, được gọi là các kênh vận chuyển: Tín hiệu trộn giảm có thể là, ví dụ, bốn tín hiệu hệ số hợp thành tín hiệu định dạng B (tức là, FOA), cặp âm lập thể, hoặc sự trộn giảm đơn âm phụ thuộc vào tốc độ bit được đặt mục tiêu. Các tham số không gian được mã hóa và dòng bit âm thanh được mã hóa được dồn kênh trước sự truyền dẫn.

Ngữ cảnh: Tổng quan hệ thống về bộ mã hóa âm thanh không gian trên cơ sở DirAC

Sau đây, tổng quan về hệ thống mã hóa âm thanh không gian theo tình trạng kỹ thuật của sáng chế dựa trên DirAC được thiết kế cho các dịch vụ tiếng nói và âm thanh đắm chìm (Immersive Voice and Audio Services - IVAS) sẽ được trình bày. Mục đích của hệ thống này là để xử lý các định dạng âm thanh không gian khác nhau biểu diễn ngữ cảnh âm thanh và mã hóa chúng tại các tốc độ bit thấp và để tái tạo ngữ cảnh âm thanh ban đầu càng chính xác càng tốt sau khi truyền dẫn.

Hệ thống có thể chấp nhận các sự biểu diễn đầu vào khác nhau của ngữ cảnh âm thanh. Ngữ cảnh âm thanh đầu vào có thể được biểu diễn bởi các tín hiệu đa kênh nhằm

mục đích tái tạo tại các vị trí loa phát thanh khác nhau, các đối tượng thính giác cùng với siêu dữ liệu mô tả các vị trí của các đối tượng theo thời gian, hoặc định dạng Ambisonics thứ tự thứ nhất hoặc thứ tự cao hơn biểu diễn trường âm tại vị trí người nghe hoặc vị trí tham chiếu.

Tốt hơn là hệ thống dựa trên các dịch vụ tiếng nói nâng cao (Enhanced Voice Services - EVS) 3GPP vì giải pháp mong đợi vận hành với độ chờ thấp để cho phép các dịch vụ hội thoại trên mạng di động.

Phía bộ mã hóa của sự mã hóa âm thanh không gian trên cơ sở DirAC hỗ trợ các định dạng âm thanh khác nhau được minh họa trên Fig.1b. Đầu vào âm thanh/điện tử 1000 được nhập vào giao diện bộ mã hóa 1010, trong đó giao diện bộ mã hóa có chức năng cụ thể cho Ambisonics thứ tự thứ nhất (first-order Ambisonics - FOA) hoặc Ambisonics thứ tự cao (high order Ambisonics - HOA) được minh họa trên 1013. Hơn nữa, giao diện bộ mã hóa có chức năng cho dữ liệu đa kênh (multichannel - MC) như dữ liệu âm lập thể, dữ liệu 5.1 hoặc dữ liệu có nhiều hơn hai hoặc năm kênh. Hơn nữa, giao diện bộ mã hóa 1010 có chức năng để mã hóa đối tượng như, ví dụ, các đối tượng âm thanh được minh họa tại 1011. Bộ mã hóa IVAS bao gồm giai đoạn DirAC 1020 có khối phân tích DirAC 1021 và khối trộn giảm DMX 1022. Tín hiệu được phát ra bởi khối 1022 được mã hóa bởi bộ mã hóa lõi IVAS 1040 như bộ mã hóa AAC hoặc EVS, và siêu dữ liệu được tạo ra bởi khối 1021 được mã hóa sử dụng bộ mã hóa siêu dữ liệu DirAC 1030.

Fig.1b minh họa phía bộ mã hóa của sự mã hóa âm thanh không gian trên cơ sở DirAC hỗ trợ các định dạng âm thanh khác nhau. Như được thể hiện trên Fig.1b, bộ mã hóa (bộ mã hóa IVAS) có thể hỗ trợ các định dạng trình bày với hệ thống riêng biệt hoặc cùng lúc. Các tín hiệu âm thanh có thể là âm thanh trong tự nhiên, được chọn bởi micro, hoặc điện tử trong tự nhiên, mà cần được truyền dẫn đến các loa phát thanh. Các định dạng âm thanh được hỗ trợ có thể là các tín hiệu đa kênh (multi-channel signals - MC), Ambisonics thứ tự thứ nhất hoặc thứ tự cao hơn (FOA/HOA). Ngữ cảnh âm thanh phức tạp có thể cũng được mô tả bằng cách tổ hợp các định dạng đầu vào khác nhau. Tất cả các định dạng âm thanh khi đó được truyền dẫn đến sự phân tích DirAC, mà trích sự biểu diễn tham số của toàn bộ ngữ cảnh âm thanh. Hướng tới (direction-of-arrival - DOA) và

độ khuếch tán được đo mỗi đơn vị thời gian - tần số tạo ra các tham số không gian hoặc là phần của tập hợp các tham số lớn hơn. Sự phân tích DirAC theo sau bởi 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 tham số tốc độ bit thấp.

Ngoài các định dạng đầu vào trên cơ sở kênh, trên cơ sở HOA, và trên cơ sở đối tượng được mô tả trên đây, bộ mã hóa IVAS có thể nhận sự biểu diễn tham số của âm không gian được hợp thành từ siêu dữ liệu không gian và/hoặc định hướng và mã hóa tín hiệu đầu vào âm được kết hợp. Siêu dữ liệu có thể ví dụ, tương ứng với siêu dữ liệu DirAC, tức là, DOA và độ khuếch tán của âm. Siêu dữ liệu có thể cũng bao gồm các tham số không gian như nhiều DOA với các ước số năng lượng, giá trị khoảng cách hoặc vị trí, hoặc các ước số liên quan đến sự tương quan của trường âm. Các tín hiệu đầu vào âm thanh được hợp thành từ tín hiệu đơn âm, tín hiệu Ambisonics có thứ tự thứ nhất hoặc thứ tự cao hơn, tín hiệu âm lập thể X/Y, tín hiệu âm lập thể A/B, hoặc bất kỳ tổ hợp các tín hiệu dẫn đến từ các bản ghi với micro có các mẫu định hướng và/hoặc khoảng cách chung khác nhau.

Đối với đầu vào âm thanh không gian tham số, bộ mã hóa IVAS xác định tham số DirAC được sử dụng cho sự truyền dẫn dựa trên siêu dữ liệu không gian đầu vào.

Cùng với các tham số, tín hiệu trộn giảm (DMX) được suy ra từ các tín hiệu đầu vào âm thanh được mã hóa cho sự truyền dẫn bởi bộ mã hóa lõi âm thanh ban đầu. Trong trường hợp này, bộ mã hóa âm thanh trên cơ sở EVS được thích ứng để mã hóa tín hiệu trộn giảm. Tín hiệu trộn giảm bao gồm các kênh khác nhau, được gọi là các kênh vận chuyển: Tín hiệu có thể là, ví dụ, bốn tín hiệu hệ số hợp thành tín hiệu định dạng B hoặc tín hiệu Ambisonics thứ tự thứ nhất (FOA), cặp âm lập thể, hoặc sự trộn giảm đơn âm phụ thuộc vào tốc độ bit được đặt mục tiêu. Các tham số không gian được mã hóa và dòng bit âm thanh được mã hóa được dồn kênh trước khi được truyền dẫn qua kênh truyền thông.

Fig.2a minh họa phía bộ giải mã của sự mã hóa âm thanh không gian trên cơ sở DirAC hỗ trợ các định dạng âm thanh khác nhau. Trong bộ giải mã, được thể hiện trên Fig.2a, các kênh vận chuyển được giải mã bởi bộ giải mã lõi, trong khi siêu dữ liệu DirAC được giải mã đầu tiên trước khi được vận chuyển với các kênh vận chuyển được

giải mã đến sự tổng hợp DirAC. Tại giai đoạn này, các tùy chọn khác nhau có thể được xem xét. Có thể yêu cầu để phát ngữ cảnh âm thanh trực tiếp trên bất kỳ cấu hình loa phát thanh hoặc ống nghe như thường khả thi trong hệ thống DirAC thông thường (MC trên Fig.2a). Bộ giải mã có thể cũng phân phối các đối tượng riêng lẻ như chúng đã được trình bày tại phía bộ mã hóa (các đối tượng trên Fig.2a). Cách khác, có thể cũng được yêu cầu kết xuất ngữ cảnh cho định dạng Ambisonics (FOA/HOA trên Fig.2a) cho sự thao tác khác, như quay, phản chiếu, hoặc sự di chuyển của ngữ cảnh, hoặc để sử dụng bộ kết xuất bên ngoài không được định nghĩa trong hệ thống ban đầu.

Trong bộ giải mã, được thể hiện trên Fig.2a, các kênh vận chuyển được giải mã bởi bộ giải mã lõi, trong khi siêu dữ liệu DirAC được giải mã đầu tiên trước khi được vận chuyển với các kênh vận chuyển được giải mã đến sự tổng hợp DirAC. Tại giai đoạn này, các tùy chọn khác nhau có thể được xem xét. Có thể yêu cầu để phát ngữ cảnh âm thanh trực tiếp trên bất kỳ cấu hình loa phát thanh hoặc ống nghe như thường khả thi trong hệ thống DirAC thông thường (MC trên Fig.2a). Bộ giải mã có thể cũng phân phối các đối tượng riêng lẻ như chúng đã được trình bày tại phía bộ mã hóa (các đối tượng trên Fig.2a). Cách khác, có thể cũng được yêu cầu kết xuất ngữ cảnh cho định dạng Ambisonics cho các thao tác khác, như quay, phản chiếu, hoặc sự di chuyển của ngữ cảnh, hoặc để sử dụng bộ kết xuất bên ngoài không được định nghĩa trong hệ thống ban đầu.

Bộ giải mã của sự mã hóa âm thanh không gian DirAC phân phát các định dạng âm thanh khác nhau được minh họa trên Fig.2a và bao gồm bộ giải mã IVAS 1045 và giao diện bộ giải mã được nối về sau 1046. Bộ giải mã IVAS 1045 bao gồm bộ giải mã lõi IVAS 1060 mà được tạo cấu hình để thực hiện thuật toán giải mã nội dung được mã hóa bởi bộ mã hóa lõi 1040 của Fig.1b. Hơn nữa, bộ giải mã siêu dữ liệu DirAC 1050 được đề xuất mà phân phối chức năng giải mã để giải mã nội dung được mã hóa bởi bộ mã hóa siêu dữ liệu DirAC 1030. Bộ tổng hợp DirAC 1070 nhận dữ liệu từ khối 1050 và 1060 và sử dụng tương tác người dùng hoặc không, đầu vào được nhập vào giao diện bộ giải mã 1046 mà tạo ra dữ liệu FOA/HOA được minh họa tại 1083, dữ liệu đa kênh (dữ liệu MC) như được minh họa trong khối 1082, hoặc dữ liệu đối tượng như được minh họa trong khối 1080.

Sự tổng hợp HOA sử dụng hệ biến hóa DirAC được thể hiện trên Fig.2b. Tín hiệu đầu vào được gọi là tín hiệu trộn giảm được phân tích thời gian - tần số bởi giàn lọc tần số. Giàn lọc tần số 2000 có thể là giàn lọc giá trị phức như QMF giá trị phức hoặc sự biến đổi khối như STFT. Sự tổng hợp HOA tạo ra tại đầu vào tín hiệu Ambisonics có thứ tự H chứa $(H + 1)^2$ thành phần. Một cách tùy chọn, có thể cũng xuất ra tín hiệu Ambisonics được kết xuất trên lớp loa phát thanh cụ thể. Sau đây, chúng ta sẽ làm rõ cách thu được thành phần $(H + 1)^2$ từ tín hiệu trộn giảm được kết hợp trong một số trường hợp bằng các tham số không gian đầu vào.

Tín hiệu trộn giảm có thể là các tín hiệu micro ban đầu hoặc hỗn hợp của các tín hiệu ban đầu biểu thị ngữ cảnh âm thanh ban đầu. Ví dụ, nếu ngữ cảnh âm thanh được ghi bởi micro trường âm, tín hiệu trộn giảm có thể là thành phần đẳng hướng của ngữ cảnh (W), trộn giảm âm lập thể (L/R), hoặc tín hiệu Ambisonics thứ tự thứ nhất (FOA).

Đối với mỗi ô thời gian - tần số, hướng âm, cũng được gọi là hướng tới (DOA), và thừa số độ khuếch tán được ước lượng lần lượt bằng bộ ước lượng hướng 2020 và bằng bộ ước lượng độ khuếch tán 2010, nếu tín hiệu trộn giảm chứa thông tin đủ để xác định các tham số DirAC như vậy. Đó là trường hợp, ví dụ, nếu tín hiệu trộn giảm là tín hiệu Ambisonics thứ tự thứ nhất (FOA). Cách khác hoặc nếu tín hiệu trộn giảm không đủ để xác định các tham số này, các tham số có thể được chuyển trực tiếp đến sự tổng hợp DirAC qua dòng bit đầu vào chứa các tham số không gian. Dòng bit có thể chứa, ví dụ, các tham số được lượng tử hóa và mã hóa được nhận như thông tin phụ trong trường hợp ứng dụng truyền dẫn âm thanh. Trong trường hợp này, các tham số được suy ra bên ngoài mô đun tổng hợp DirAC từ các tín hiệu micro ban đầu hoặc các định dạng âm thanh đầu vào được đưa ra cho mô đun phân tích DirAC tại phía bộ mã hóa như được minh họa bởi chuyển mạch 2030 hoặc 2040.

Các hướng âm được sử dụng bởi bộ đánh giá độ khuếch đại định hướng 2050 để đánh giá, đối với mỗi ô thời gian - tần số trong số nhiều ô thời gian - tần số, một hoặc nhiều tập hợp gồm $(H + 1)^2$ độ khuếch đại định hướng $G_l^m(k, n)$, trong đó H là thứ tự của tín hiệu Ambisonics được tổng hợp.

Các độ khuếch đại định hướng có thể thu được bằng sự đánh giá của hàm cơ bản không gian cho mỗi hướng âm được ước lượng tại thứ tự mong muốn (mức) l và chế độ

m của tín hiệu Ambisonics để tổng hợp. Hướng âm có thể được thể hiện, ví dụ dưới dạng vectơ chuẩn đơn vị $n(k, n)$ hoặc dưới dạng góc phương vị $\varphi(k, n)$ và/hoặc góc nâng $\theta(k, n)$, mà liên quan đến ví dụ:

$$n(k, n) = \begin{bmatrix} \cos \varphi(k, n) \cos \theta(k, n) \\ \sin \varphi(k, n) \cos \theta(k, n) \\ \sin \theta(k, n) \end{bmatrix}$$

Sau khi ước lượng hoặc thu hướng âm, phản hồi hàm cơ bản không gian có chế độ mong muốn (mức) l và chế độ m có thể được xác định, ví dụ, bằng cách xem xét

các sóng hài hình cầu giá trị thực với sự chuẩn hóa SN3D như hàm cơ bản không gian:

$$Y_l^m(\varphi, \theta) = N_l^{|m|} P_l^{|m|} \sin \theta \begin{cases} \sin(|m|\varphi) & \text{if } m < 0 \\ \cos(|m|\varphi) & \text{if } m \geq 0 \end{cases}$$

với các phạm vi $0 \leq l \leq H$, và $-l \leq m \leq l$. $P_l^{|m|}$ là các hàm Legendre và $N_l^{|m|}$ là hạn chuẩn hóa cho cả hai hàm Legendre và hàm lượng giác mà có dạng sau đối với SN3D:

$$N_l^{|m|} = \sqrt{\frac{2 - \delta_m (l - |m|)!}{4\pi (l + |m|)!}}$$

trong đó delta Kronecker δ_m là 1 đối với $m = 0$ và nếu không thì bằng 0. Các độ khuếch đại định hướng sau đó được giảm trực tiếp cho mỗi ô thời gian - tần số có các chỉ số (k, n) như:

$$G_l^m(k, n) = Y_l^m(\varphi(k, n), \theta(k, n))$$

Các thành phần Ambisonics âm trực tiếp $P_{s,l}^m$ được tính toán bằng cách suy ra tín hiệu tham chiếu P_{ref} từ tín hiệu trộn giảm và được nhân với các độ khuếch đại định hướng và hàm thừa số có độ khuếch tán $\Psi(k, n)$:

$$P_{s,l}^m(k, n) = P_{ref}(k, n) \sqrt{1 - \Psi(k, n)} G_l^m(k, n)$$

Ví dụ, tín hiệu tham chiếu P_{ref} có thể là thành phần đẳng hướng của tín hiệu trộn giảm hoặc sự tổ hợp tuyến tính của K kênh của tín hiệu trộn giảm.

Thành phần Ambisonics âm khuếch tán có thể được tạo mô hình bằng cách sử dụng phản hồi hàm cơ bản không gian cho các âm tới từ tất cả các hướng khả thi. Một ví dụ là định nghĩa phản hồi trung bình D_l^m bằng cách xem xét sự tích hợp của cường độ được bình phương có hàm cơ bản không gian $Y_l^m(\varphi, \theta)$ trên tất cả các góc khả thi φ và θ :

$$D_l^m = \int_0^{2\pi} \int_0^\pi |Y_l^m(\varphi, \theta)|^2 \sin \theta d\theta d\varphi$$

Các thành phần Ambisonics âm khuếch tán $P_{d,l}^m$ được tính toán từ tín hiệu P_{diff} được nhân với phản hồi trung bình và hàm thừa số có độ khuếch tán $\Psi(k, n)$:

$$P_{d,l}^m(k, n) = P_{diff,l}^m(k, n) \sqrt{\Psi(k, n)} \sqrt{D_l^m}$$

Tín hiệu $P_{diff,l}^m$ có thể thu được bằng cách sử dụng các bộ giải tương quan được áp dụng cho tín hiệu tham chiếu P_{ref} .

Cuối cùng, thành phần Ambisonics âm trực tiếp và thành phần Ambisonics âm khuếch tán được tổ hợp 2060, ví dụ, qua toán tử cộng tổng, để thu được thành phần Ambisonics cuối cùng P_l^m có thứ tự (mức) mong muốn l và chế độ m cho ô thời gian - tần số (k, n) , tức là,

$$P_l^m(k, n) = P_{s,l}^m(k, n) + P_{diff,l}^m(k, n)$$

Các thành phần Ambisonics thu được có thể được biến đổi ngược về miền thời gian sử dụng giàn lọc nghịch đảo 2080 hoặc STFT nghịch đảo, được lưu trữ, truyền dẫn, hoặc được sử dụng, ví dụ, cho các ứng dụng tái tạo âm không gian. Cách khác, bộ kết xuất Ambisonics tuyến tính 2070 có thể được áp dụng cho mỗi băng tần để thu được các tín hiệu sẽ được phát trên lớp loa phát thanh cụ thể hoặc qua ống nghe trước khi biến đổi các tín hiệu loa phát thanh hoặc tín hiệu lập thể đến miền thời gian.

Nên lưu ý rằng [Thiergart17] cũng dẫn dắt khả năng rằng các thành phần âm khuếch tán $P_{diff,l}^m$ có thể cũng được tổng hợp lên đến thứ tự L , trong đó $L < H$. Điều này giảm độ phức tạp tính toán trong khi tránh các thành phần lạ tổng hợp do sự sử dụng dồn dập các bộ giải tương quan.

Mục đích của sáng chế là cung cấp khái niệm được cải thiện để tạo ra sự mô tả trường âm từ tín hiệu đầu vào.

Tình trạng kỹ thuật: Sự tổng hợp DirAC cho các tín hiệu đơn và các tín hiệu trộn giảm FOA

Sự tổng hợp DirAC thông thường, dựa trên dòng mã hóa âm thanh không gian dựa trên DirAC, được mô tả sau đây. Sự kết xuất được thực hiện bằng sự tổng hợp DirAC dựa trên các tín hiệu âm thanh trộn giảm được giải mã và siêu dữ liệu không gian được giải mã.

Tín hiệu trộn giảm là tín hiệu đầu vào của sự tổng hợp DirAC. Tín hiệu được biến đổi thành miền thời gian - tần số bằng giàn lọc. Giàn lọc có thể là giàn lọc giá trị phức như QMF giá trị phức hoặc sự biến đổi khối như STFT.

Các tham số DirAC có thể được vận chuyển trực tiếp đến sự tổng hợp DirAC qua dòng bit đầu vào chứa các tham số không gian. Dòng bit có thể chứa, ví dụ, các tham số được lượng tử hóa và mã hóa được nhận như thông tin phụ trong trường hợp ứng dụng truyền dẫn âm thanh.

Để xác định các tín hiệu kênh cho sự tái tạo kênh dựa trên loa phát thanh, mỗi tín hiệu loa phát thanh được xác định dựa trên các tín hiệu trộn giảm và các tham số DirAC. Tín hiệu của loa phát thanh thứ j $P_j(k, n)$ thu được như sự tổ hợp của thành phần âm trực tiếp và thành phần âm khuếch tán, tức là,

$$P_j(k, n) = P_{\text{dir},j}(k, n) + P_{\text{diff},j}(k, n)$$

Thành phần âm trực tiếp của kênh loa phát thanh thứ j $P_{\text{dir},j}(k, n)$ có thể thu được bằng cách định tỉ lệ cái được gọi là tín hiệu tham chiếu $P_{\text{ref},j}(k, n)$ với thừa số phụ thuộc vào tham số độ khuếch tán $\Psi(k, n)$ và thừa số khuếch đại định hướng $G_j(\mathbf{v}(k, n))$, trong đó thừa số khuếch đại phụ thuộc vào hướng tới (DOA) của âm và cũng có khả năng phụ thuộc vào vị trí của kênh loa phát thanh thứ j . Hướng tới của âm có thể được thể hiện, ví dụ dưới dạng vectơ chuẩn đơn vị $\mathbf{v}(k, n)$ hoặc dưới dạng góc phương vị $\varphi(k, n)$ và/hoặc góc nâng $\theta(k, n)$, mà liên quan đến ví dụ

$$\mathbf{v}(k, n) = \begin{bmatrix} \cos \varphi(k, n) \cos \theta(k, n) \\ \sin \varphi(k, n) \cos \theta(k, n) \\ \sin \theta(k, n) \end{bmatrix}$$

Thừa số khuếch đại định hướng $G_j(\mathbf{v}(k, n))$ có thể được tính toán sử dụng các phương pháp đã biết như sự quét biên độ góc véctor (VBAP) [Pulkki97].

Xem xét trên đây, thành phần âm trực tiếp có thể được biểu thị bằng cách

$$P_{\text{dir},j}(k, n) = P_{\text{ref},j}(k, n) \sqrt{1 - \Psi(k, n)} G_j(\mathbf{v}(k, n))$$

Các tham số mô tả DOA của âm và độ khuếch tán được ước lượng tại bộ giải mã từ các kênh vận chuyển hoặc được ước lượng từ siêu dữ liệu tham số nằm trong dòng bit.

Thành phần âm khuếch tán $P_{\text{diff},j}(k, n)$ có thể được xác định dựa trên tín hiệu tham chiếu và tham số độ khuếch tán:

$$P_{\text{diff},j}(k, n) = P_{\text{ref},j}(k, n) \sqrt{\Psi(k, n)} G_{\text{norm}}$$

Thừa số chuẩn hóa G_{norm} phụ thuộc vào cấu hình loa phát thanh phát lại. Thông thường, các thành phần âm khuếch tán được kết hợp với các kênh loa phát thanh khác nhau $P_{\text{diff},j}(k, n)$ còn được xử lý, tức là, chúng được giải tương quan qua lại. Điều này có thể cũng đạt được bằng cách giải tương quan tín hiệu tham chiếu cho mỗi kênh đầu ra, tức là,

$$P_{\text{diff},j}(k, n) = \tilde{P}_{\text{ref},j}(k, n) \sqrt{\Psi(k, n)} G_{\text{norm}},$$

trong đó $\tilde{P}_{\text{ref},j}(k, n)$ biểu thị phiên bản được giải tương quan của $P_{\text{ref},j}(k, n)$.

Tín hiệu tham chiếu cho kênh đầu ra thứ j thu được dựa trên các tín hiệu trộn giảm được truyền dẫn. Trong trường hợp đơn giản nhất, tín hiệu trộn giảm bao gồm tín hiệu đẳng hướng đơn âm (ví dụ, thành phần đẳng hướng $W(k, n)$ của tín hiệu FOA) và tín hiệu tham chiếu giống nhau cho tất cả các kênh đầu ra:

$$P_{\text{ref},j}(k, n) = W(k, n)$$

Nếu các kênh vận chuyển tương ứng với bốn thành phần của tín hiệu FOA, các tín hiệu tham chiếu có thể thu được bằng sự tổ hợp tuyến tính của các thành phần FOA.

Thông thường, các tín hiệu FOA được tổ hợp sao cho tín hiệu tham chiếu của kênh thứ j tương ứng với tín hiệu micro hai hướng ảo chỉ vào hướng của loa phát thanh thứ j [Pulkki07].

Sự tổng hợp DirAC thường cung cấp chất lượng tái tạo âm được cải thiện cho số lượng các kênh trộn giảm tăng, vì cả hai lượng được yêu cầu của sự giải tương quan tổng hợp, mức độ xử lý phi tuyến tính bởi các thừa số định hướng, hoặc sự xuyên âm giữa các kênh loa phát thanh khác nhau có thể được giảm và các thành phần lạ được kết hợp có thể được tránh hoặc làm giảm.

Nhìn chung, phương pháp trực tiếp để đưa nhiều tín hiệu vận chuyển khác nhau vào ngữ cảnh âm thanh được mã hóa một mặt là không linh hoạt và mặt khác tiêu thụ tốc độ bit. Thông thường, có thể không cần trong tất cả các trường hợp để đưa ra, ví dụ, tất cả bốn tín hiệu thành phần của tín hiệu Ambisonics thứ tự thứ nhất vào tín hiệu âm thanh được mã hóa, vì một hoặc nhiều thành phần không có sự đóng góp năng lượng quan trọng. Mặt khác, các yêu cầu tốc độ bit có thể chặt chẽ mà ngăn cản đưa vào nhiều hơn hai kênh vận chuyển vào tín hiệu âm thanh được mã hóa biểu diễn sự biểu diễn âm thanh không gian. Trong trường hợp yêu cầu tốc độ bit chặt chẽ như vậy, cần thiết cho bộ mã hóa và bộ giải mã đàm phán trước trước sự biểu diễn nhất định, và, dựa trên sự biểu diễn trước này, lượng nhất định của các tín hiệu vận chuyển được tạo ra dựa trên cách được đàm phán trước và, khi đó, bộ giải mã âm thanh có thể tổng hợp ngữ cảnh âm thanh từ tín hiệu âm thanh được mã hóa dựa trên kiến thức được đàm phán trước. Tuy nhiên, mặc dù điều này có ích đối với các yêu cầu tốc độ bit, nó không linh hoạt, và ngoài ra có thể tạo lượng trong chất lượng âm thanh được giảm đáng kể, vì quy trình đàm phán trước có thể không tối ưu đối với mảnh âm thanh nhất định hoặc không thể tối ưu hóa đối với tất cả các băng tần hoặc đối với tất cả các khung thời gian của mảnh âm thanh.

Do đó, quy trình theo tình trạng kỹ thuật là biểu diễn ngữ cảnh âm thanh là không tối ưu so với các yêu cầu tốc độ bit, là không linh hoạt, và, ngoài ra, có khả năng dẫn đến chất lượng âm thanh giảm đáng kể.

Tài liệu tham khảo

[Pulkki07] V. Pulkki, “ Spatial Sound Reproduction with Directional Audio Coding”, J. Audio Eng. Soc., Volume 55 Issue 6 pp. 503-516; June 2007.

[Pulkki97] V. Pulkki, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning" J. Audio Eng. Soc., Volume 45 Issue 6 pp. 456-466; June 1997

[Thiergart09] O. Thiergart, R. Schultz-Amling, G. Del Galdo, D. Mahne, F. Kuech, "Localization of Sound Sources in Reverberant Environments Based on Directional Audio Coding Parameters", AES Convention 127, Paper No. 7853, Oct. 2009

[Thiergart17] WO2017157803 A1, O. Thiergart et. al. "APPARATUS, METHOD OR COMPUTER PROGRAM FOR GENERATING A SOUND FIELD DESCRIPTION"

[Laitinen11] M. Laitinen, F. Kuech, V. Pulkki, "Using Spaced Microphones with Directional Audio Coding ", AES Convention 130, Paper No. 8433, May 2011

[Vilkamo13] J. Vilkamo, V. Pulkki, " Minimization of Decorrelator Artifacts in Directional Audio Coding by Covariance Domain Rendering", J. Audio Eng. Soc., Vol. 61, No. 9, 2013 September

[Veen88] B.D. Van Veen, K.M. Buckley, "Beamforming: a versatile approach to spatial filtering", IEEE ASSP Mag., vol. 5, no. 2, pp. 4-24, 1998

[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] M. V. Laitinen and V. Pulkki, "Converting 5.1 audio recordings to B-format for directional audio coding reproduction," 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, 2011, pp. 61-64

[3] R. K. Furness, "Ambisonics —An overview," in AES 8th International Conference, April 1990, pp. 181—189.

[4] C. Nachbar, F. Zotter, E. Deleflie, and A. Sontacchi, "AMBIX – A Suggested Ambisonics Format", Proceedings of the Ambisonics Symposium 2011

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là cung cấp giải pháp cải tiến để mã hóa sự biểu diễn âm thanh không gian hoặc giải mã tín hiệu âm thanh được mã hóa.

Mục đích này đạt được nhờ thiết bị mã hóa sự biểu diễn âm thanh không gian theo điểm 1, thiết bị giải mã tín hiệu âm thanh được mã hóa theo điểm 21, phương pháp mã hóa sự biểu diễn âm thanh không gian theo điểm 39, phương pháp giải mã tín hiệu âm thanh được mã hóa theo điểm 41, chương trình máy tính theo điểm 43, hoặc tín hiệu âm thanh được mã hóa theo điểm 44.

Sáng chế dựa trên việc tìm ra rằng sự cải thiện đáng kể đối với tốc độ bit, độ linh hoạt và chất lượng âm thanh thu được bằng cách sử dụng, ngoài sự biểu diễn vận chuyển được suy ra từ sự biểu diễn âm thanh không gian, siêu dữ liệu vận chuyển liên quan đến sự tạo ra sự biểu diễn vận chuyển hoặc biểu thị một hoặc nhiều đặc tính định hướng của sự biểu diễn vận chuyển. Thiết bị mã hóa sự biểu diễn âm thanh không gian biểu diễn ngữ cảnh âm thanh do đó tạo ra sự biểu diễn vận chuyển từ ngữ cảnh âm thanh, và, ngoài ra, siêu dữ liệu vận chuyển liên quan đến sự tạo ra sự biểu diễn vận chuyển hoặc biểu thị một hoặc nhiều đặc tính định hướng của sự biểu diễn vận chuyển hoặc liên quan đến sự tạo ra sự biểu diễn vận chuyển và biểu thị một hoặc nhiều đặc tính định hướng của sự biểu diễn vận chuyển. Hơn nữa, giao diện đầu ra tạo ra tín hiệu âm thanh được mã hóa bao gồm thông tin về sự biểu diễn vận chuyển, và thông tin về siêu dữ liệu vận chuyển.

Trên phía bộ giải mã, thiết bị giải mã tín hiệu âm thanh được mã hóa bao gồm giao diện để nhận tín hiệu âm thanh được mã hóa bao gồm thông tin về sự biểu diễn vận chuyển và thông tin về siêu dữ liệu vận chuyển và bộ tổng hợp âm thanh không gian khi đó tổng hợp sự biểu diễn âm thanh không gian sử dụng cả hai, thông tin về sự biểu diễn vận chuyển và thông tin về siêu dữ liệu vận chuyển.

Sự biểu thị rõ ràng về cách mà sự biểu diễn vận chuyển sao cho tín hiệu trộn giảm đã được tạo ra và/hoặc sự biểu thị rõ ràng về một hoặc nhiều đặc tính định hướng của sự biểu diễn vận chuyển bằng siêu dữ liệu vận chuyển bổ sung cho phép bộ mã hóa tạo ra ngữ cảnh âm thanh được mã hóa theo cách linh hoạt cao, mặt khác cung cấp chất lượng âm thanh tốt, và mặt khác, thỏa mãn các yêu cầu tốc độ bit. Ngoài ra, bằng siêu dữ liệu vận chuyển, thậm chí khả thi để bộ mã hóa tìm ra một mặt là sự cân bằng tối ưu

được yêu cầu giữa các yêu cầu tốc độ bit và mặt khác là chất lượng âm thanh được biểu diễn bởi tín hiệu âm thanh được mã hóa. Do đó, sự sử dụng siêu dữ liệu vận chuyển rõ ràng cho phép bộ mã hóa áp dụng nhiều cách khác nhau tạo ra sự biểu diễn vận chuyển và để thích ứng thêm sự tạo ra biểu diễn vận chuyển không chỉ từ mảnh âm thanh đến mảnh âm thanh, nhưng thậm chí từ một khung âm thanh đến khung âm thanh tiếp theo hoặc, trong một và cùng khung âm thanh từ băng tần đến băng tần khác. Một cách tự nhiên, độ linh hoạt thu được bằng cách tạo ra sự biểu diễn vận chuyển cho mỗi ô thời gian/tần số riêng biệt để, ví dụ, sự biểu diễn vận chuyển có thể được tạo ra cho tất cả các ngăn tần số trong khung thời gian hoặc, cách khác, sự biểu diễn vận chuyển tương tự có thể được tạo ra cho một và cùng băng tần trên nhiều khung thời gian âm thanh hoặc sự biểu diễn vận chuyển riêng có thể được tạo ra cho mỗi ngăn tần số của mỗi khung thời gian. Tất cả các thông tin này, tức là, cách tạo ra sự biểu diễn vận chuyển và liệu sự biểu diễn vận chuyển liên quan đến khung đầy đủ, hoặc chỉ đến ngăn thời gian/tần số hoặc băng tần nhất định theo nhiều khung thời gian cũng nằm trong siêu dữ liệu vận chuyển để bộ tổng hợp âm thanh không gian biết về những gì đã làm tại phía bộ mã hóa và khi đó có thể áp dụng quy trình tối ưu hóa tại phía bộ giải mã.

Tốt hơn là, các thay thế siêu dữ liệu vận chuyển nhất định là thông tin lựa chọn biểu thị thành phần nào trong số tập hợp nhất định gồm các thành phần biểu diễn ngữ cảnh âm thanh đã được lựa chọn. Thay thế siêu dữ liệu vận chuyển khác liên quan đến thông tin tổ hợp, tức là, liệu và/hoặc bằng cách nào các tín hiệu thành phần nhất định của sự biểu diễn âm thanh không gian đã được tổ hợp để tạo ra sự biểu diễn vận chuyển. Thông tin khác có ích như siêu dữ liệu vận chuyển liên quan đến thông tin phân/bán cầu biểu thị phần hoặc bán cầu nào tín hiệu vận chuyển nhất định hoặc kênh vận chuyển nhất định liên quan đến. Hơn nữa, siêu dữ liệu có ích trong ngữ cảnh theo sáng chế để thông tin hướng nhìn biểu thị hướng nhìn của tín hiệu âm thanh bao gồm như tín hiệu vận chuyển của, tốt hơn là, nhiều tín hiệu vận chuyển khác nhau trong sự biểu diễn vận chuyển. Thông tin hướng nhìn khác liên quan đến hướng nhìn micro, khi sự biểu diễn vận chuyển bao gồm một hoặc nhiều tín hiệu micro mà có thể, ví dụ, được ghi bằng micro vật lý trong mảng micro (được mở rộng về không gian) hoặc bằng micro trùng nhau hoặc, cách khác, các tín hiệu micro này có thể được tạo ra theo cách tổng hợp. Siêu dữ liệu vận chuyển khác liên quan đến dữ liệu tham số định hình biểu thị liệu tín hiệu

micro là tín hiệu đẳng hướng, hay có hình dạng khác như hình dạng hai hướng hoặc hình dạng hai cực. Siêu dữ liệu vận chuyển khác liên quan đến các vị trí của các micro trong trường hợp có nhiều hơn một tín hiệu micro trong sự biểu diễn vận chuyển. Siêu dữ liệu vận chuyển có ích khác liên quan đến dữ liệu định hướng của một hoặc nhiều micro, đến dữ liệu khoảng cách biểu thị khoảng cách giữa hai micro hoặc các mẫu định hướng của các micro. Hơn nữa, siêu dữ liệu vận chuyển bổ sung có thể liên quan đến sự mô tả hoặc sự nhận biết của mảng micro như mảng micro hình tròn hoặc tín hiệu micro nào từ mảng micro hình tròn như vậy đã được lựa chọn như sự biểu diễn vận chuyển.

Siêu dữ liệu vận chuyển khác có thể liên quan đến thông tin về tạo chùm, các trọng số tạo chùm tương ứng hoặc hướng tương ứng và, trong tình huống như vậy, sự biểu diễn vận chuyển bao gồm cụ thể tín hiệu được tạo ra tổng hợp được ưu tiên hơn có hướng chùm nhất định. Các thay thế siêu dữ liệu vận chuyển khác có thể tương ứng với thông tin thuần liệu các tín hiệu vận chuyển được bao gồm là các tín hiệu micro đẳng hướng hoặc các tín hiệu micro phi đẳng hướng như các tín hiệu hai hướng hoặc các tín hiệu hai cực.

Do đó, trở nên rõ ràng rằng các thay thế siêu dữ liệu vận chuyển khác nhau linh hoạt cao hơn và có thể được biểu diễn theo cách rút gọn cao hơn để siêu dữ liệu vận chuyển bổ sung thường không dẫn đến lượng tốc độ bit bổ sung đáng kể. Thay vào đó, các yêu cầu tốc độ bit cho siêu dữ liệu vận chuyển bổ sung có thể thường nhỏ như nhỏ hơn 1% hoặc thậm chí nhỏ hơn 1/1000 hoặc thậm chí nhỏ hơn lượng cho sự biểu diễn vận chuyển. Tuy nhiên, mặt khác, đây là lượng rất nhỏ siêu dữ liệu bổ sung dẫn đến độ linh hoạt cao hơn, và cùng lúc đó, sự tăng đáng kể chất lượng âm thanh do sự linh hoạt bổ sung và do tiềm năng có các sự biểu diễn vận chuyển thay đổi theo các mảnh âm thanh khác nhau hoặc, thậm chí trong một và cùng mảnh âm thanh theo các khung thời gian khác nhau và/hoặc các ngăn tần số.

Tốt hơn là, bộ mã hóa còn bao gồm bộ xử lý tham số để tạo ra các tham số không gian từ sự biểu diễn âm thanh không gian để, ngoài sự biểu diễn vận chuyển và siêu dữ liệu vận chuyển, các tham số không gian nằm trong tín hiệu âm thanh được mã hóa để nâng cao chất lượng âm thanh theo chất lượng chỉ có thể thu được bằng sự biểu diễn vận chuyển và siêu dữ liệu vận chuyển. Các tham số không gian này tốt hơn là dữ liệu hướng

tới (DoA) phụ thuộc thời gian và/hoặc tần số và/hoặc dữ liệu độ khuếch tán phụ thuộc tần số và/hoặc thời gian như, ví dụ, được biết đến từ sự mã hóa DirAC.

Trên phía bộ giải mã âm thanh, giao diện đầu vào nhận tín hiệu âm thanh được mã hóa bao gồm thông tin về sự biểu diễn vận chuyển, và thông tin về siêu dữ liệu vận chuyển. Hơn nữa, bộ tổ hợp âm thanh không gian được lắp trong thiết bị để giải mã tín hiệu âm thanh được mã hóa tổng hợp sự biểu diễn âm thanh không gian sử dụng cả thông tin về sự biểu diễn vận chuyển và thông tin về siêu dữ liệu vận chuyển. Trong phương án được ưu tiên, bộ giải mã còn sử dụng các tham số không gian được truyền dẫn một cách tùy chọn để tổng hợp sự biểu diễn âm thanh không gian không chỉ sử dụng thông tin về siêu dữ liệu vận chuyển và thông tin về sự biểu diễn vận chuyển, mà còn sử dụng các tham số không gian.

Thiết bị giải mã tín hiệu âm thanh được mã hóa nhận siêu dữ liệu vận chuyển, dịch hoặc phân tách siêu dữ liệu vận chuyển nhận được, và khi đó điều khiển bộ tổ hợp để tổ hợp các tín hiệu biểu diễn vận chuyển hoặc lựa chọn từ các tín hiệu biểu diễn vận chuyển hoặc tạo ra một hoặc một vài tín hiệu tham chiếu. Bộ tổ hợp/bộ lựa chọn/bộ tạo tín hiệu tham chiếu khi đó chuyển tiếp tín hiệu tham chiếu cho bộ tính toán tín hiệu thành phần mà tính toán các thành phần đầu ra được yêu cầu từ các tín hiệu tham chiếu được lựa chọn hoặc tạo ra một cách cụ thể. Trong các phương án được ưu tiên, không chỉ bộ tổ hợp/bộ lựa chọn/bộ tạo tín hiệu tham chiếu như trong bộ tổ hợp âm thanh không gian được điều khiển bởi siêu dữ liệu vận chuyển, mà bộ tính toán tín hiệu thành phần để, dựa trên dữ liệu vận chuyển nhận được, không chỉ sự tạo ra/lựa chọn tín hiệu tham chiếu được điều khiển, mà còn sự tính toán thành phần thực tế. Tuy nhiên, các phương án mà chỉ sự tính toán tín hiệu thành phần được điều khiển bởi siêu dữ liệu vận chuyển hoặc chỉ sự tạo ra hoặc sự lựa chọn tín hiệu tham chiếu chỉ được điều khiển bởi siêu dữ liệu vận chuyển cũng có ích và cung cấp sự linh hoạt được cải thiện theo các giải pháp có sẵn.

Các quy trình được ưu tiên có các thay thế lựa chọn tín hiệu khác nhau được lựa chọn một trong số nhiều tín hiệu trong sự biểu diễn vận chuyển như tín hiệu tham chiếu cho tập con thứ nhất gồm các tín hiệu thành phần và lựa chọn tín hiệu vận chuyển khác trong sự biểu diễn vận chuyển cho tập con thẳng góc khác gồm các tín hiệu thành phần

cho đầu ra đa kênh, đầu ra Ambisonics thứ tự thứ nhất hoặc thứ tự cao hơn, đầu ra đối tượng âm thanh âm thanh, hoặc đầu ra lập thể. Các quy trình khác phụ thuộc vào sự tính toán tín hiệu tham chiếu dựa trên sự tổ hợp tuyến tính của các tín hiệu riêng lẻ nằm trong sự biểu diễn vận chuyển. Phụ thuộc vào sự thực thi biểu diễn vận chuyển nhất định, siêu dữ liệu vận chuyển được sử dụng để xác định tín hiệu tham chiếu cho các kênh (ảo) từ các tín hiệu vận chuyển được truyền dẫn thực tế và xác định các thành phần bị thiếu dựa trên sự dự trữ, như thành phần tín hiệu đẳng hướng được truyền dẫn hoặc được tạo ra. Các quy trình này phụ thuộc vào sự tính toán các thành phần còn thiếu, tốt hơn là FOA hoặc HOA sử dụng phản hồi hàm cơ bản không gian liên quan đến chế độ và thứ tự nhất định của sự biểu diễn âm thanh không gian Ambisonics thứ tự thứ nhất hoặc thứ tự cao hơn.

Các phương án khác về siêu dữ liệu vận chuyển mô tả các tín hiệu micro nằm trong sự biểu diễn vận chuyển, và dựa trên tham số tạo hình được truyền dẫn và/hoặc hướng nhìn, sự xác định tín hiệu tham chiếu được thích ứng với siêu dữ liệu vận chuyển được nhận. Hơn nữa, sự tính toán các tín hiệu đẳng hướng hoặc các tín hiệu hai cực và sự tổng hợp bổ sung của các thành phần còn lại cũng được thực hiện dựa trên siêu dữ liệu vận chuyển biểu thị, ví dụ, rằng kênh vận chuyển thứ nhất là tín hiệu hai cực bên trái hoặc phía trước, và tín hiệu vận chuyển thứ hai là tín hiệu hai hướng bên phải hoặc phía sau.

Các quy trình khác liên quan đến sự xác định các tín hiệu tham chiếu dựa trên khoảng cách nhỏ nhất của loa nhất định đến vị trí micro nhất định hoặc sự lựa chọn, như tín hiệu tham chiếu, tín hiệu micro nằm trong sự biểu diễn vận chuyển với hướng nhìn gần nhất hoặc bộ tạo chùm gần nhất hoặc vị trí mảng gần nhất nhất định. Quy trình khác là sự lựa chọn tín hiệu vận chuyển tùy ý như tín hiệu tham chiếu cho tất cả các thành phần âm trực tiếp và sự sử dụng tất cả các tín hiệu vận chuyển có sẵn như các tín hiệu đẳng hướng được truyền dẫn từ các micro được tạo khoảng cách cho sự tạo ra tín hiệu tham chiếu âm khuếch tán và các thành phần tương ứng khi đó được tạo ra bằng cách thêm các thành phần trực tiếp và khuếch tán để thu được kênh cuối cùng hoặc thành phần Ambisonics hoặc tín hiệu đối tượng hoặc tín hiệu kênh lập thể. Các quy trình khác được thực thi cụ thể trong sự tính toán tín hiệu thành phần thực tế dựa trên tín hiệu tham chiếu nhất định liên quan trong thiết lập (tốt Hơn là hạn chế) lượng tương quan dựa trên khoảng cách micro nhất định.

Mô tả vắn tắt các hình vẽ

Các phương án ưu tiên của sáng chế được mô tả sau đây với việc đối chiếu với các hình vẽ đi kèm trong đó:

Fig.1a minh họa sóng hài hình cầu với số lượng kênh/thành phần Ambisonics;

Fig.1b minh họa phía bộ mã hóa của bộ xử lý mã hóa âm thanh không gian trên cơ sở DirAC;

Fig.2a minh họa phía bộ giải mã của bộ xử lý mã hóa âm thanh không gian trên cơ sở DirAC;

Fig.2b minh họa bộ xử lý tổng hợp Ambisonics thứ tự cao đã biết trong lĩnh vực kỹ thuật;

Fig.3 minh họa phía bộ mã hóa của sự mã hóa âm thanh không gian trên cơ sở DirAC hỗ trợ các định dạng âm thanh khác nhau;

Fig.4 minh họa phía bộ giải mã của sự mã hóa âm thanh không gian trên cơ sở DirAC hỗ trợ các định dạng âm thanh khác nhau;

Fig.5 minh họa phương án khác của thiết bị mã hóa sự biểu diễn âm thanh không gian;

Fig.6 minh họa phương án khác của thiết bị mã hóa sự biểu diễn âm thanh không gian;

Fig.7 minh họa phương án khác của thiết bị giải mã tín hiệu âm thanh được mã hóa;

Fig.8a minh họa tập hợp các sự thực thi cho bộ tạo sự biểu diễn vận chuyển có thể sử dụng riêng với nhau hoặc cùng với nhau;

Fig.8b minh họa bảng thể hiện các thay thế siêu dữ liệu vận chuyển khác nhau có thể sử dụng riêng với nhau hoặc cùng với nhau;

Fig.8c minh họa sự thực thi khác của bộ mã hóa siêu dữ liệu đối với siêu dữ liệu vận chuyển hoặc, nếu phù hợp, đối với các tham số không gian;

Fig.9a minh họa sự thực thi được ưu tiên của bộ tổng hợp âm thanh không gian của Fig.7;

Fig.9b minh họa tín hiệu âm thanh được mã hóa có sự biểu diễn vận chuyển với n tín hiệu vận chuyển, siêu dữ liệu vận chuyển và các tham số không gian tùy chọn;

Fig.9c minh họa bảng về chức năng bộ lựa chọn/bộ tạo tín hiệu tham chiếu phụ thuộc vào sự nhận biết loa và siêu dữ liệu vận chuyển;

Fig.9d minh họa phương án khác của bộ tổng hợp âm thanh không gian;

Fig.9e minh họa bảng khác thể hiện siêu dữ liệu vận chuyển khác;

Fig.9f minh họa sự thực thi khác của bộ tổng hợp âm thanh không gian;

Fig.9g minh họa phương án khác của bộ tổng hợp âm thanh không gian;

Fig.9h minh họa tập hợp khác gồm các thay thế sự thực thi cho bộ tổng hợp âm thanh không gian có thể sử dụng riêng với nhau hoặc cùng với nhau;

Fig.10 minh họa sự thực thi được ưu tiên minh họa để tính toán các thành phần trường âm thứ tự thấp hoặc giữa sử dụng tín hiệu trực tiếp và tín hiệu khuếch tán;

Fig.11 minh họa sự thực thi khác của sự tính toán các thành phần trường âm thứ tự cao hơn sử dụng thành phần trực tiếp mà không có thành phần khuếch tán; và

Fig.12 minh họa sự thực thi khác của sự tính toán các thành phần hoặc đối tượng tín hiệu loa phát thanh (ảo) sử dụng phần trực tiếp được tổ hợp với phần khuếch tán.

Mô tả chi tiết sáng chế

Fig.6 minh họa thiết bị mã hóa sự biểu diễn âm thanh không gian biểu diễn ngữ cảnh âm thanh. Thiết bị bao gồm bộ tạo sự biểu diễn vận chuyển 600 để tạo ra sự biểu diễn vận chuyển từ sự biểu diễn âm thanh không gian. Hơn nữa, bộ tạo sự biểu diễn vận chuyển 600 tạo ra siêu dữ liệu vận chuyển liên quan đến sự tạo ra sự biểu diễn vận chuyển hoặc biểu thị một hoặc nhiều đặc tính định hướng của sự biểu diễn vận chuyển. Thiết bị còn bao gồm giao diện đầu ra 640 để tạo ra tín hiệu âm thanh được mã hóa, trong đó tín hiệu âm thanh được mã hóa bao gồm thông tin về sự biểu diễn vận chuyển, và thông tin về siêu dữ liệu vận chuyển. Ngoài bộ tạo sự biểu diễn vận chuyển 600 và giao diện đầu ra 640, thiết bị tốt hơn là bao gồm giao diện người dùng 650 và bộ xử lý

tham số 620. Bộ xử lý tham số 620 được tạo cấu hình để suy ra các tham số không gian từ sự biểu diễn âm thanh không gian và tốt hơn là cung cấp tham số không gian (được mã hóa) 612. Hơn nữa, ngoài tham số không gian (được mã hóa) 612, siêu dữ liệu vận chuyển (được mã hóa) 610 và sự biểu diễn vận chuyển (được mã hóa) 611 được chuyển tiếp đến giao diện đầu ra 640 để tốt hơn là dồn kênh ba mục được mã hóa vào tín hiệu âm thanh được mã hóa.

Fig.7 minh họa sự thực thi được ưu tiên của thiết bị giải mã tín hiệu âm thanh được mã hóa. Tín hiệu âm thanh được mã hóa được nhập vào giao diện đầu vào 700 và giao diện đầu vào nhận, trong tín hiệu âm thanh được mã hóa, thông tin về sự biểu diễn vận chuyển và thông tin về siêu dữ liệu vận chuyển. Sự biểu diễn vận chuyển 711 được chuyển tiếp, từ giao diện đầu vào 700, đến bộ tổng hợp âm thanh không gian 750. Hơn nữa, bộ tổng hợp âm thanh không gian 750 nhận siêu dữ liệu 710 từ giao diện đầu vào và, nếu nằm trong tín hiệu âm thanh được mã hóa, tốt hơn là, ngoài tham số không gian 712. Bộ tổng hợp âm thanh không gian 750 sử dụng các mục 710, 711 và, tốt hơn là, ngoài mục 712 để tổng hợp sự biểu diễn âm thanh không gian.

Fig.3 minh họa sự thực thi được ưu tiên của thiết bị mã hóa sự biểu diễn âm thanh không gian được biểu thị như tín hiệu âm thanh không gian trên Fig.3. Cụ thể, tín hiệu âm thanh không gian được nhập vào khối tạo ra trộn giảm 610 và vào khối phân tích âm thanh không gian 621. Các tham số không gian 615 được suy ra từ khối phân tích âm thanh không gian 621 từ tín hiệu âm thanh không gian được nhập vào bộ mã hóa siêu dữ liệu 622. Hơn nữa, các tham số trộn giảm 630 được tạo ra bằng khối tạo ra trộn giảm 601 cũng được nhập vào bộ mã hóa siêu dữ liệu 603. Cả hai bộ mã hóa siêu dữ liệu 621 và bộ mã hóa siêu dữ liệu 603 được biểu thị như khối đơn trên Fig.3 nhưng cũng có thể được thực thi như các khối riêng. Tín hiệu âm thanh trộn giảm 640 được nhập vào bộ mã hóa lõi 603 và sự biểu diễn được mã hóa lõi 611 được nhập vào bộ tạo dòng bit 641 mà còn nhận các tham số trộn giảm được mã hóa 610 và các tham số không gian được mã hóa 612. Do đó, bộ tạo sự biểu diễn vận chuyển 600 được minh họa trên Fig.6 bao gồm, trong phương án của Fig.3, khối tạo ra trộn giảm 601 và khối bộ mã hóa lõi 603. Hơn nữa, bộ xử lý tham số 620 được minh họa trên Fig.6 bao gồm khối phân tích âm thanh không gian 621 và khối bộ mã hóa siêu dữ liệu 622 cho tham số không gian 615. Hơn nữa, bộ tạo sự biểu diễn vận chuyển 600 của Fig.6 còn bao gồm khối bộ mã hóa

siêu dữ liệu 603 cho siêu dữ liệu vận chuyển 630 mà được xuất ra như siêu dữ liệu vận chuyển được mã hóa 610 bằng bộ mã hóa siêu dữ liệu 603. Giao diện đầu ra 40 là, trong phương án của Fig.3, được thực thi như bộ tạo dòng bit 641.

Fig.4 minh họa sự thực thi được ưu tiên của thiết bị giải mã tín hiệu âm thanh được mã hóa. Cụ thể, thiết bị bao gồm bộ giải mã siêu dữ liệu 752 và bộ giải mã lõi 751. Bộ giải mã siêu dữ liệu 752 nhận, như đầu vào, siêu dữ liệu vận chuyển được mã hóa 710 và bộ giải mã lõi 751 nhận sự biểu diễn vận chuyển được mã hóa 711. Hơn nữa, bộ giải mã siêu dữ liệu 752 tốt hơn là nhận, khi có sẵn, các tham số không gian được mã hóa 712. Bộ giải mã siêu dữ liệu giải mã siêu dữ liệu vận chuyển 710 để nhận tham số trộn giảm 720, và bộ giải mã siêu dữ liệu 752 ưu tiên giải mã các tham số không gian được mã hóa 712 để thu được tham số không gian được giải mã 722. Sự biểu diễn vận chuyển được giải mã hoặc sự biểu diễn âm thanh trộn giảm 721 cùng với siêu dữ liệu 720 được nhập vào khối tổng hợp âm thanh không gian 753 và, ngoài ra, khối tổng hợp âm thanh không gian 753 có thể nhận tham số không gian 722 để sử dụng hai thành phần 721 và 720 hoặc tất cả ba thành phần 721, 720 và 722 để tạo ra sự biểu diễn âm thanh không gian bao gồm sự biểu diễn thứ tự thứ nhất hoặc thứ tự cao hơn (first order or higher order - FOA/HOA) 754 hoặc bao gồm sự biểu diễn đa kênh (multichannel - MC) 755 hoặc bao gồm sự biểu diễn đối tượng (các đối tượng) 756 như được minh họa trên Fig.4. Do đó, thiết bị giải mã tín hiệu âm thanh được mã hóa được minh họa trên Fig.7 bao gồm, trong bộ tổng hợp âm thanh không gian 750, khối 752, 751 và 753 của Fig.4, và sự biểu diễn âm thanh không gian có thể bao gồm một trong số các thay thế 754, 755 và 756 của Fig.4.

Fig.5 minh họa sự thực thi khác của thiết bị mã hóa sự biểu diễn âm thanh không gian biểu diễn ngữ cảnh âm thanh. Ở đây, sự biểu diễn âm thanh không gian biểu diễn ngữ cảnh âm thanh được cung cấp như các tín hiệu micro và, tốt hơn là, các tham số không gian được kết hợp với các tín hiệu micro. Do đó, sự biểu diễn vận chuyển 600 được thảo luận viện dẫn đến Fig.6 bao gồm, trong phương án của Fig.5, khối tạo ra trộn giảm 601, bộ mã hóa siêu dữ liệu 603 cho các tham số trộn giảm 613 và bộ mã hóa lõi 602 cho sự biểu diễn âm thanh trộn giảm. Ngược với phương án của Fig.3, khối bộ phân tích âm thanh không gian 621 không nằm trong thiết bị mã hóa, vì đầu vào micro đã có,

tốt hơn trong dạng được phân tách, một mặt là các tín hiệu micro và mặt khác là các tham số không gian.

Trong phương án được thảo luận viện dẫn đến Fig.3 đến Fig.5, âm thanh trộn giảm 614 biểu diễn sự biểu diễn vận chuyển, và các tham số trộn giảm 613 biểu diễn sự thay thế của siêu dữ liệu vận chuyển mà liên quan đến sự tạo ra sự biểu diễn vận chuyển hoặc, sẽ được nêu ra phân sau đây, biểu thị một hoặc nhiều đặc tính định hướng của sự biểu diễn vận chuyển.

Các phương án được ưu tiên của sáng chế: Tín hiệu trộn giảm cho cấu hình kênh vận chuyển linh hoạt

Trong một số ứng dụng, không thể truyền dẫn tất cả bốn thành phần của tín hiệu FOA như các kênh vận chuyển do giới hạn bit, nhưng chỉ tín hiệu trộn giảm với số lượng giảm các thành phần tín hiệu hoặc kênh. Để đạt được chất lượng tái tạo được cải thiện tại bộ giải mã, sự tạo ra các tín hiệu trộn giảm được truyền dẫn có thể được thực hiện theo cách biến thiên thời gian và có thể được thích ứng cho tín hiệu đầu vào âm thanh không gian. Nếu hệ thống mã hóa âm thanh không gian cho phép bao gồm các tín hiệu trộn giảm linh hoạt, quan trọng không chỉ truyền dẫn các kênh vận chuyển này mà ngoài ra bao gồm siêu dữ liệu mà chỉ rõ các đặc điểm không gian quan trọng của các tín hiệu trộn giảm. Sự tổng hợp DirAC được định vị tại bộ giải mã của hệ thống mã hóa âm thanh không gian khi đó có thể thích ứng quy trình kết xuất theo cách tối ưu xem xét các đặc điểm không gian của các tín hiệu trộn giảm. Sáng chế này do đó đề xuất bao gồm siêu dữ liệu liên quan đến trộn giảm trong dòng mã hóa âm thanh không gian theo tham số mà được sử dụng để chỉ rõ hoặc mô tả các đặc điểm không gian quan trọng của các kênh vận chuyển trộn giảm để cải thiện chất lượng kết xuất tại bộ giải mã âm thanh không gian.

Sau đây, các ví dụ minh họa cho các cấu hình tín hiệu trộn giảm thực tế được mô tả.

Nếu tín hiệu âm thanh không gian đầu vào chủ yếu bao gồm năng lượng âm trong mặt phẳng nằm ngang, chỉ ba thành phần tín hiệu âm thanh không gian đầu tiên của tín hiệu FOA tương ứng với tín hiệu đẳng hướng, tín hiệu hai cực được căn chỉnh với trục

x và tín hiệu hai cực được căn chỉnh với trục y của hệ tọa độ Cartesian nằm trong tín hiệu âm trộn giảm, trong đó tín hiệu hai cực được căn chỉnh với trục z được loại trừ.

Theo ví dụ khác, chỉ hai tín hiệu trộn giảm có thể được truyền dẫn để giảm thêm tốc độ bit được yêu cầu cho các kênh vận chuyển. Ví dụ, nếu năng lượng âm chủ yếu bắt nguồn từ bán cầu bên trái, có lợi để tạo ra kênh trộn giảm mà bao gồm năng lượng âm chủ yếu từ hướng bên trái và kênh trộn giảm bổ sung bao gồm âm bắt nguồn chủ yếu từ hướng đối diện, tức là, bán cầu bên phải trong ví dụ này. Điều này có thể đạt được bằng sự tổ hợp tuyến tính của các thành phần tín hiệu FOA sao cho các tín hiệu thu được tương ứng với các tín hiệu micro định hướng với các mẫu định hướng hai hướng lần lượt chỉ về bên trái và bên phải. Tương tự, các tín hiệu trộn giảm tương ứng với các mẫu định hướng thứ tự thứ nhất lần lượt chỉ về hướng phía trước và phía sau, hoặc bất kỳ mẫu định hướng nào có thể được tạo ra bằng cách tổ hợp phù hợp các tín hiệu đầu vào FOA.

Trong giai đoạn tổng hợp DirAC, sự tính toán các kênh đầu ra loa phát thanh dựa trên siêu dữ liệu không gian được truyền dẫn (ví dụ, DOA của âm và độ khuếch tán) và các kênh vận chuyển âm thanh phải được thích ứng đến cấu hình trộn giảm được sử dụng thực sự. Cụ thể hơn, lựa chọn phù hợp nhất cho tín hiệu tham chiếu của loa phát thanh thứ j $P_{\text{ref},j}(k, n)$ phụ thuộc vào đặc điểm định hướng của các tín hiệu trộn giảm và vị trí của loa phát thanh thứ j .

Ví dụ, nếu các tín hiệu trộn giảm tương ứng với hai tín hiệu micro hai hướng lần lượt chỉ vào bên trái và bên phải, tín hiệu tham chiếu của loa phát thanh được đặt vào bán cầu bên trái nên chỉ sử dụng tín hiệu hai hướng chỉ vào bên trái như tín hiệu tham chiếu $P_{\text{ref},j}(k, n)$. Loa phát thanh được đặt ở giữa có thể sử dụng tổ hợp tuyến tính của cả hai tín hiệu trộn giảm thay vào đó.

Mặt khác, nếu các tín hiệu trộn giảm tương ứng với hai tín hiệu micro hai hướng lần lượt chỉ vào phía trước và phía sau, tín hiệu tham chiếu của loa phát thanh được đặt trong bán cầu phía trước nên chỉ sử dụng tín hiệu hai hướng chỉ vào phía trước như tín hiệu tham chiếu $P_{\text{ref},j}(k, n)$.

Điều quan trọng là lưu ý rằng sự suy giảm đáng kể chất lượng âm thanh không gian phải được mong đợi nếu sự tổng hợp DirAC sử dụng tín hiệu trộn giảm sai như tín hiệu tham chiếu cho sự kết xuất. Ví dụ, nếu tín hiệu trộn giảm tương ứng với micro hai hướng chỉ vào bên trái được sử dụng cho sự tạo ra tín hiệu kênh đầu ra cho loa phát thanh được đặt trong bán cầu bên phải, các thành phần tín hiệu bắt nguồn từ bán cầu bên trái của trường âm đầu vào sẽ được hướng chủ yếu đến bán cầu bên phải của hệ thống tái tạo dẫn đến ảnh không gian không chính xác của đầu ra.

Do đó, ưu tiên hơn bao gồm thông tin tham số trong dòng mã hóa âm thanh không gian mà chỉ rõ các đặc điểm không gian của các tín hiệu trộn giảm như các mẫu định hướng của các tín hiệu micro định hướng tương ứng. Sự tổng hợp DirAC được định vị tại bộ giải mã của hệ thống mã hóa âm thanh không gian khi đó có thể thích ứng quy trình kết xuất theo cách tối ưu xem xét các đặc điểm không gian của các tín hiệu trộn giảm như được mô tả trong siêu dữ liệu về trộn giảm.

Sự trộn giảm linh hoạt cho đầu vào âm thanh FOA và HOA sử dụng sự lựa chọn thành phần Ambisonics.

Trong phương án, tín hiệu âm thanh không gian, tức là, tín hiệu đầu vào âm thanh đến bộ mã hóa, tương ứng với tín hiệu âm thanh Ambisonics thứ tự thứ nhất (first-order Ambisonics - FOA) hoặc Ambisonics thứ tự cao hơn (higher-order Ambisonics - HOA). Sơ đồ khối tương ứng của bộ mã hóa được biểu thị trên Fig.3. Đầu vào cho bộ mã hóa là tín hiệu âm thanh không gian, ví dụ, tín hiệu FOA hoặc HOA. Trong khối “phân tích âm thanh không gian”, các tham số DirAC, tức là, các tham số không gian (ví dụ, DOA và độ khuếch tán), được ước lượng như giải thích trước đó. Các tín hiệu trộn giảm của sự trộn giảm linh hoạt được đề xuất được tạo ra trong khối “tạo ra trộn giảm”, mà được giải thích cụ thể hơn dưới đây. Các tín hiệu trộn giảm được tạo ra được đề cập đến như $D_m(k, n)$, trong đó m là chỉ số của kênh trộn giảm. Tín hiệu trộn giảm được tạo ra khi đó được mã hóa trong khối “bộ mã hóa lõi”, ví dụ, sử dụng bộ mã hóa âm trên cơ sở EVS như được giải thích trước đó. Các tham số trộn giảm, tức là, các tham số mô tả thông tin liên quan về cách sự trộn giảm được tạo ra hoặc các đặc điểm định hướng khác của tín hiệu trộn giảm, được mã hóa trong bộ mã hóa siêu dữ liệu cùng với các tham số

không gian. Cuối cùng, siêu dữ liệu được mã hóa và các tín hiệu trộn giảm được biến đổi vào dòng bit, mà có thể được gửi đến bộ giải mã.

Sau đây, khối “tạo ra trộn giảm” và các tham số trộn giảm được giải thích cụ thể hơn. Nếu, ví dụ, tín hiệu âm thanh không gian đầu vào chủ yếu bao gồm năng lượng âm trong mặt phẳng nằm ngang, chỉ ba thành phần tín hiệu âm thanh không gian của tín hiệu FOA/HOA tương ứng với tín hiệu đẳng hướng $W(k, n)$, tín hiệu hai cực $X(k, n)$ được căn chỉnh với trục x, và tín hiệu hai cực $Y(k, n)$ được căn chỉnh với trục y của hệ tọa độ Cartesian nằm trong tín hiệu âm trộn giảm, trong đó tín hiệu hai cực $Z(k, n)$ được căn chỉnh với trục z (và tất cả các thành phần thứ tự cao hơn khác, nếu có) được loại trừ. Điều này có nghĩa là, các tín hiệu trộn giảm được cho bởi

$$D_1(k, n) = W(k, n), D_2(k, n) = X(k, n), D_3(k, n) = Y(k, n).$$

Cách khác, nếu, ví dụ, tín hiệu âm thanh không gian đầu vào chủ yếu bao gồm năng lượng âm trong mặt phẳng x-z, các tín hiệu trộn giảm bao gồm tín hiệu hai cực $Z(k, n)$ thay vì $Y(k, n)$.

Trong phương án này, các tham số trộn giảm, được mô tả trên Fig.3, chứa thông tin mà các thành phần FOA/HOA đã nằm trong các tín hiệu trộn giảm. Thông tin này có thể, ví dụ, là tập hợp các số nguyên tương ứng với các chỉ số của các thành phần FOA được lựa chọn, ví dụ $\{1, 2, 4\}$ nếu các thành phần $W(k, n)$, $X(k, n)$, và $Z(k, n)$ được bao gồm.

Lưu ý rằng sự lựa chọn các thành phần FOA/HOA cho tín hiệu trộn giảm có thể được thực hiện, ví dụ, dựa trên đầu vào người dùng thủ công hoặc tự động. Ví dụ, khi tín hiệu đầu vào âm thanh không gian được ghi tại đường băng sân bay, có thể giả định rằng hầu hết năng lượng âm nằm trong mặt phẳng Cartesian thẳng đứng cụ thể. Trong trường hợp này, ví dụ, các thành phần $W(k, n)$, $X(k, n)$ và $Z(k, n)$ được lựa chọn. Ngược lại, nếu bản ghi được thực hiện tại ngã tư đường, có thể giả định rằng hầu hết năng lượng âm nằm trong mặt phẳng Cartesian nằm ngang. Trong trường hợp này, ví dụ, các thành phần $W(k, n)$, $X(k, n)$ và $Y(k, n)$ được lựa chọn. Cách khác, nếu, ví dụ, máy ghi video được sử dụng cùng với sự ghi âm, thuật toán nhận biết khuôn mặt có thể được sử dụng để phát hiện mà mặt phẳng Cartesian nào, người nói được định vị và do đó, các thành phần FOA tương ứng với mặt phẳng này có thể được lựa chọn cho sự trộn

giảm. Cách khác, người ta có thể xác định mặt phẳng của hệ tọa độ Cartesian với năng lượng cao nhất bằng cách sử dụng thuật toán định vị nguồn âm thanh theo tình trạng kỹ thuật.

Cũng lưu ý rằng sự lựa chọn thành phần FOA/HOA và siêu dữ liệu trộn giảm tương ứng có thể phụ thuộc vào thời gian và tần số, ví dụ, tập hợp khác gồm các thành phần và chỉ số, lần lượt, có thể được lựa chọn tự động cho mỗi băng tần và nấc thời gian (ví dụ, bằng cách xác định tự động mặt phẳng Cartesian với năng lượng cao nhất cho mỗi điểm thời gian - tần số). Sự cục bộ hóa năng lượng âm trực tiếp có thể được thực hiện ví dụ, bằng cách khai thác thông tin nằm trong các tham số không gian phụ thuộc vào thời gian - tần số [Thiergart09].

Sơ đồ khối bộ giải mã tương ứng với phương án được biểu thị trên Fig.4. Việc nhập vào bộ giải mã là dòng bit chứa siêu dữ liệu được mã hóa và các tín hiệu âm thanh trộn giảm được mã hóa. Các tín hiệu âm thanh trộn giảm được giải mã trong “bộ giải mã lõi” và siêu dữ liệu được giải mã trong “bộ giải mã siêu dữ liệu”. Siêu dữ liệu được giải mã bao gồm các tham số không gian (ví dụ, DOA và độ khuếch tán) và các tham số trộn giảm. Các tín hiệu âm thanh trộn giảm được giải mã và các tham số không gian được sử dụng trong khối “tổng hợp âm thanh không gian” để tạo ra các tín hiệu đầu ra âm thanh không gian mong muốn mà có thể, ví dụ, là các tín hiệu FOA/HOA, các tín hiệu đa kênh (multi-channel - MC) (ví dụ, các tín hiệu âm thanh loa phát thanh), các đối tượng âm thanh hoặc đầu ra âm lập thể cho sự phát lại ống nghe. Sự tổng hợp âm thanh không gian ngoài ra được điều khiển bởi các tham số trộn giảm, như được giải thích sau đây.

Sự tổng hợp âm thanh không gian (sự tổng hợp DirAC) được mô tả trước đó yêu cầu tín hiệu tham chiếu phù hợp $P_{\text{ref},j}(k, n)$ cho mỗi kênh đầu ra j . Trong sáng chế này, được đề xuất tính toán $P_{\text{ref},j}(k, n)$ từ các tín hiệu trộn giảm $D_m(k, n)$ sử dụng siêu dữ liệu trộn giảm bổ sung. Trong phương án này, các tín hiệu trộn giảm $D_m(k, n)$ bao gồm các thành phần được lựa chọn cụ thể của tín hiệu FOA hoặc HOA, và siêu dữ liệu trộn giảm mô tả các thành phần FOA/HOA đã được truyền dẫn đến bộ giải mã.

Khi kết xuất đến các loa phát thanh (tức là, đầu ra MC của bộ giải mã), đầu ra chất lượng cao có thể đạt được khi tính toán cho mỗi kênh loa phát thanh tín hiệu micro ảo,

mà được đưa trực tiếp về phía loa phát thanh tương ứng, như được giải thích trong [Pulkki07]. Thông thường, sự tính toán các tín hiệu micro ảo yêu cầu tất cả các thành phần FOA/HOA có sẵn trong sự tổng hợp DirAC. Tuy nhiên, trong phương án này, chỉ tập con của các thành phần FOA/HOA ban đầu có sẵn tại bộ giải mã. Trong trường hợp này, các tín hiệu micro ảo có thể được tính chỉ cho mặt phẳng Cartesian, mà các thành phần FOA/HOA có sẵn, như được biểu thị bởi siêu dữ liệu trộn giảm. Ví dụ, nếu siêu dữ liệu trộn giảm biểu thị rằng thành phần $W(k, n)$, $X(k, n)$, và $Y(k, n)$ đã được truyền dẫn, các tác giả sáng chế có thể tính toán các tín hiệu micro ảo cho tất cả các loa phát thanh trong mặt phẳng x-y (mặt phẳng nằm ngang), trong đó sự tính toán có thể được thực hiện như được mô tả trong [Pulkki07]. Đối với các loa phát thanh được nâng bên ngoài mặt phẳng nằm ngang, các tác giả sáng chế có thể sử dụng giải pháp dự trữ cho tín hiệu tham chiếu $P_{\text{ref},j}(k, n)$, ví dụ, các tác giả sáng chế có thể sử dụng thành phần đẳng hướng $W(k, n)$.

Lưu ý rằng khái niệm tương tự có thể được sử dụng khi kết xuất đầu ra âm lập thể, ví dụ, cho sự phát lại ống nghe. Trong trường hợp này, hai micro ảo cho hai kênh đầu ra được dẫn hướng về phía các loa phát thanh âm lập thể ảo, trong đó vị trí của các loa phát thanh phụ thuộc vào hướng đầu của người nghe. Nếu các loa phát thanh ảo được đặt trong mặt phẳng Cartesian, mà các thành phần FOA/HO đã được truyền dẫn như được biểu thị bởi siêu dữ liệu trộn giảm, các tác giả sáng chế có thể tính toán các tín hiệu micro ảo tương ứng. Nếu không, giải pháp dự trữ được sử dụng cho tín hiệu tham chiếu $P_{\text{ref},j}(k, n)$, ví dụ, thành phần đẳng hướng $W(k, n)$.

Khi kết xuất đến FOA/HOA (đầu ra FOA/HOA của bộ giải mã trên Fig.4), siêu dữ liệu trộn giảm được sử dụng như sau: Siêu dữ liệu trộn giảm biểu thị các thành phần FOA/HOA đã được truyền dẫn. Các thành phần này không cần phải được tính toán trong sự tổng hợp âm thanh không gian, vì các thành phần được truyền dẫn có thể được sử dụng trực tiếp tại đầu vào bộ giải mã. Tất cả các thành phần FOA/HOA được tính toán trong sự tổng hợp âm không gian, ví dụ, bằng cách sử dụng thành phần đẳng hướng $W(k, n)$ như tín hiệu tham chiếu $P_{\text{ref},j}(k, n)$. Sự tổng hợp các thành phần FOA/HOA từ thành phần đẳng hướng $W(k, n)$ sử dụng siêu dữ liệu không gian được mô tả, ví dụ trong [Thiergart17].

Sự trộn giảm linh hoạt cho đầu vào âm thanh FOA và HOA sử dụng các thành phần Ambisonics được tổ hợp

Trong phương án, tín hiệu âm thanh không gian, tức là, tín hiệu đầu vào âm thanh đến bộ mã hóa, tương ứng với tín hiệu âm thanh Ambisonics thứ tự thứ nhất (first-order Ambisonics - FOA) hoặc Ambisonics thứ tự cao hơn (higher-order Ambisonics - HOA). Sơ đồ khối tương ứng của bộ mã hóa và được biểu thị lần lượt trên Fig.3 và Fig.4. Trong phương án này, chỉ hai tín hiệu trộn giảm có thể được truyền dẫn từ bộ mã hóa đến bộ giải mã để giảm thêm tốc độ bit được yêu cầu cho các kênh vận chuyển. Ví dụ, nếu năng lượng âm chủ yếu bắt nguồn từ bán cầu bên trái, có lợi để tạo ra kênh trộn giảm mà bao gồm năng lượng âm chủ yếu từ bán cầu bên trái và kênh trộn giảm bổ sung bao gồm âm bắt nguồn chủ yếu từ hướng đối diện, tức là, bán cầu bên phải trong ví dụ này. Điều này có thể đạt được bằng sự tổ hợp tuyến tính của các thành phần tín hiệu đầu vào âm thanh FOA hoặc HOA sao cho các tín hiệu thu được tương ứng với, ví dụ, các tín hiệu micro định hướng với các mẫu định hướng hai hướng lần lượt chỉ về bên trái và bên phải. Tương tự, các tín hiệu trộn giảm tương ứng với các mẫu định hướng thứ tự thứ nhất (hoặc thứ tự cao hơn) lần lượt chỉ về hướng phía trước và phía sau, hoặc bất kỳ mẫu định hướng nào có thể được tạo ra bằng cách tổ hợp phù hợp lần lượt các tín hiệu đầu vào FOA hoặc HOA.

Các tín hiệu trộn giảm được tạo ra trong bộ mã hóa trong khối “tạo ra trộn giảm” trên Fig.3. Các tín hiệu trộn giảm thu được từ tổ hợp tuyến tính của các thành phần tín hiệu FOA hoặc HOA. Ví dụ, trong trường hợp của các tín hiệu đầu vào âm thanh FOA, bốn thành phần tín hiệu FOA tương ứng với tín hiệu đẳng hướng $W(k, n)$ và ba tín hiệu hai cực $X(k, n)$, $Y(k, n)$, và $Z(k, n)$ với các mẫu định hướng được căn chỉnh với trục x, y, z của hệ tọa độ Cartesian. Bốn tín hiệu này thường được đề cập đến như các tín hiệu định dạng B. Các mẫu định hướng thu được, mà có thể thu được bằng sự tổ hợp tuyến tính của các thành phần định dạng B, thường được đề cập đến như các mẫu định hướng thứ tự thứ nhất. Mẫu định hướng thứ tự thứ nhất hoặc các tín hiệu tương ứng có thể được biểu thị bằng các cách khác nhau. Ví dụ, tín hiệu trộn giảm thứ m $D_m(k, n)$ có thể được biểu thị bằng sự tổ hợp tuyến tính của các tín hiệu định dạng B với các trọng số được kết hợp, tức là,

$$D_m(k, n) = a_{m,W}W(k, n) + a_{m,X}X(k, n) + a_{m,Y}Y(k, n) + a_{m,Z}Z(k, n).$$

Lưu ý rằng trong trường hợp các tín hiệu đầu vào âm thanh HOA, tổ hợp tuyến tính có thể được thực hiện tương tự sử dụng các hệ số HOA có sẵn. Các trọng số cho tổ hợp tuyến tính, tức là, các trọng số $a_{m,W}$, $a_{m,X}$, $a_{m,Y}$, và $a_{m,Z}$ trong ví dụ này, xác định mẫu định hướng của tín hiệu micro định hướng thu được, tức là, của tín hiệu trộn giảm thứ m $D_m(k, n)$. Trong trường hợp các tín hiệu đầu vào âm thanh FOA, các trọng số cho sự tổng hợp tuyến tính có thể được tính như

$$a_{m,W} = c_m$$

$$\begin{bmatrix} a_{m,X} & a_{m,Y} & a_{m,Z} \end{bmatrix}^T = (1 - c_m) \mathbf{w}_m$$

Trong đó

$$\mathbf{w}_m = \begin{bmatrix} \cos \Phi_m \cos \Theta_m \\ \sin \Phi_m \cos \Theta_m \\ \sin \Theta_m \end{bmatrix}.$$

Ở đây, c_m là tham số thứ tự thứ nhất hoặc tham số hình dạng và Φ_m và Θ_m là góc phương vị mong muốn và góc nâng của hướng nhìn của tín hiệu micro định hướng thứ m được tạo ra. Ví dụ, đối với $c_m = 0.5$, đạt được micro định hướng với sự định hướng hai hướng, $c_m = 1$ tương ứng với đặc điểm đẳng hướng $c_m = 0$ tương ứng với đặc điểm hai cực. Nói cách khác, tham số c_m mô tả hình dạng chung của mẫu định hướng thứ tự thứ nhất.

Các trọng số cho sự tổ hợp tuyến tính, ví dụ, $a_{m,W}$, $a_{m,X}$, $a_{m,Y}$, và $a_{m,Z}$, hoặc các tham số tương ứng c_m , Φ_m , và Θ_m , mô tả mẫu định hướng của các tín hiệu micro định hướng tương ứng. Thông tin này được biểu diễn bởi các tham số trộn giảm trong bộ mã hóa trên Fig.3 và được truyền dẫn đến bộ giải mã như một phần của siêu dữ liệu.

Các chiến thuật mã hóa khác nhau có thể được sử dụng để biểu diễn một cách hiệu quả các tham số trộn giảm trong dòng bit bao gồm sự lượng tử hóa của thông tin định hướng hoặc đề cập đến mục nhập bảng bởi chỉ số, trong đó bảng bao gồm tất cả các tham số liên quan.

Trong một số phương án, đã đủ hoặc hiệu quả hơn để sử dụng chỉ số lượng giới hạn các thiết lập trước cho các hướng nhìn Φ_m và Θ_m cũng như cho tham số hình dạng

c_m . Điều này hiển nhiên tương ứng với việc sử dụng số lượng giới hạn các thiết lập trước cũng cho các trọng số $a_{m,W}$, $a_{m,X}$, $a_{m,Y}$, và $a_{m,Z}$. Ví dụ, các tham số hình dạng có thể được giới hạn ở biểu diễn chỉ ba mẫu định hướng định hướng khác nhau: đặc điểm đẳng hướng, hai hướng và hai cực. Số lượng các hướng nhìn khả thi Φ_m và Θ_m có thể được giới hạn sao cho chúng chỉ biểu diễn các trường hợp bên trái, bên phải, phía trước, phía sau, phía trên, và phía dưới.

Trong phương án đơn giản hơn khác, tham số hình dạng được giữ cố định và luôn luôn tương ứng với mẫu hai hướng hoặc tham số hình dạng không được định nghĩa. Các tham số trộn giảm được kết hợp với hướng nhìn được sử dụng để tạo tín hiệu liệu cặp các kênh trộn giảm tương ứng với cấu hình cặp kênh trái/phải hoặc trước/sau sao cho quy trình kết xuất tại bộ giải mã có thể sử dụng kênh trộn giảm tối ưu như tín hiệu tham chiếu để kết xuất kênh loa phát thanh nhất định được đặt trong bán cầu bên trái, bên phải, hoặc phía trước.

Trong ứng dụng thực tế, tham số c_m có thể được định nghĩa, ví dụ, một cách thủ công (thường là $c_m = 0.5$). Các hướng nhìn Φ_m và Θ_m có thể được thiết lập tự động (ví dụ, bằng cách cục bộ hóa các nguồn âm hoạt động sử dụng phương pháp cục bộ hóa nguồn âm theo tình trạng kỹ thuật và định hướng tín hiệu trộn giảm thứ nhất về phía nguồn được cục bộ hóa và tín hiệu trộn giảm thứ hai về phía hướng đối diện).

Lưu ý rằng tương tự như trong phương án trước đó, các tham số trộn giảm có thể phụ thuộc vào thời gian - tần số, tức là, cấu hình trộn giảm khác nhau có thể được sử dụng cho mỗi thời gian và tần số (ví dụ, khi định hướng các tín hiệu trộn giảm phụ thuộc vào hướng nguồn hoạt động được cục bộ hóa riêng biệt trong mỗi băng tần). Sự cục bộ hóa có thể được thực hiện ví dụ, bằng cách khai thác thông tin nằm trong các tham số không gian phụ thuộc vào thời gian - tần số [Thiergart09].

Trong giai đoạn “tổng hợp âm thanh không gian” trong bộ giải mã trên “Fig.4”, sự tính toán các tín hiệu đầu ra bộ giải mã (đầu ra FOA/HOA, đầu ra MC, hoặc đầu ra đối tượng), mà sử dụng các tham số không gian được truyền dẫn (ví dụ, DOA của âm và độ khuếch tán) và các kênh âm thanh trộn giảm $D_m(k, n)$ như được mô tả trước đó, phải được thích ứng đến cấu hình trộn giảm được sử dụng thực sự, mà được chỉ rõ bởi siêu dữ liệu trộn giảm.

Ví dụ, khi tạo ra các kênh đầu ra loa phát thanh (đầu ra MC), sự tính toán các tín hiệu tham chiếu $P_{\text{ref},j}(k, n)$ phải được thích ứng với cấu hình trộn giảm được sử dụng thực tế. Cụ thể hơn, lựa chọn phù hợp nhất cho tín hiệu tham chiếu $P_{\text{ref},j}(k, n)$ của loa phát thanh thứ j phụ thuộc vào đặc điểm định hướng của các tín hiệu trộn giảm (ví dụ, hướng nhìn của nó) và vị trí của loa phát thanh thứ j . Ví dụ, nếu siêu dữ liệu trộn giảm biểu thị rằng các tín hiệu trộn giảm tương ứng với hai tín hiệu micro hai hướng lần lượt chỉ vào bên trái và bên phải, tín hiệu tham chiếu của loa phát thanh được đặt vào bán cầu bên trái nên chủ yếu hoặc chỉ sử dụng tín hiệu trộn giảm hai hướng chỉ vào bên trái như tín hiệu tham chiếu $P_{\text{ref},j}(k, n)$. Loa phát thanh được đặt ở giữa có thể sử dụng tổ hợp tuyến tính của cả hai tín hiệu trộn giảm thay vào đó (ví dụ, tổng của hai tín hiệu trộn giảm). Mặt khác, nếu các tín hiệu trộn giảm tương ứng với hai tín hiệu micro hai hướng lần lượt chỉ vào phía trước và phía sau, tín hiệu tham chiếu của loa phát thanh được đặt trong bán cầu phía trước nên chủ yếu hoặc chỉ sử dụng tín hiệu hai hướng chỉ vào phía trước như tín hiệu tham chiếu $P_{\text{ref},j}(k, n)$.

Khi tạo ra đầu ra FOA hoặc HOA trong bộ giải mã trên Fig.4, sự tính toán tín hiệu tham chiếu $P_{\text{ref},j}(k, n)$ cũng phải được thích ứng với cấu hình trộn giảm được sử dụng thực sự, mà được mô tả bởi siêu dữ liệu trộn giảm. Ví dụ, nếu siêu dữ liệu trộn giảm biểu thị rằng các tín hiệu trộn giảm tương ứng với hai tín hiệu micro hai hướng lần lượt chỉ vào bên trái và bên phải, tín hiệu tham chiếu $P_{\text{ref},1}(k, n)$ để tổng hợp thành phần FOA thứ nhất (thành phần đẳng hướng) có thể được tính toán như tổng của hai tín hiệu trộn giảm hai hướng, tức là,

$$P_{\text{ref},1}(k, n) = D_1(k, n) + D_2(k, n).$$

Thực tế, đã được biết đến rằng tổng của hai tín hiệu hai hướng với hướng nhìn đối diện dẫn đến tín hiệu đẳng hướng. Trong trường hợp này, $P_{\text{ref},1}(k, n)$ trực tiếp dẫn đến thành phần thứ nhất của tín hiệu đầu ra FOA hoặc HOA mong muốn, tức là, không có sự tổng hợp âm không gian khác được yêu cầu cho thành phần này. Tương tự, thành phần FOA thứ ba (thành phần hai cực theo hướng y) có thể được tính toán như sự chênh lệch của hai tín hiệu trộn giảm hai hướng, tức là,

$$P_{\text{ref},3}(k, n) = D_1(k, n) - D_2(k, n).$$

Thực tế, đã được biết đến rằng sự chênh lệch của hai tín hiệu hai hướng với hướng nhìn đối diện dẫn đến tín hiệu hai cực. Trong trường hợp này, $P_{\text{ref},3}(k, n)$ trực tiếp dẫn đến thành phần thứ ba của tín hiệu đầu ra FOA hoặc HOA mong muốn, tức là, không có sự tổng hợp âm không gian khác được yêu cầu cho thành phần này. Tất cả các thành phần FOA hoặc HOA có thể được tổng hợp từ tín hiệu tham chiếu đẳng hướng, mà chứa thông tin âm thanh từ tất cả các hướng. Điều này có nghĩa là, trong ví dụ này,, tổng của hai tín hiệu trộn giảm được sử dụng cho sự tổng hợp của các thành phần FOA hoặc HOA còn lại. Nếu siêu dữ liệu trộn giảm biểu thị rằng sự định hướng khác nhau của hai tín hiệu trộn giảm âm thanh, sự tính toán của các tín hiệu tham chiếu $P_{\text{ref},j}(k, n)$ có thể được điều chỉnh theo đó. Ví dụ, nếu hai tín hiệu trộn giảm âm thanh hai hướng được dẫn hướng về phía trước và phía sau (thay vì bên trái và bên phải), sự khác biệt của hai tín hiệu trộn giảm có thể được sử dụng để tạo ra thành phần FOA thứ hai (thành phần hai cực theo hướng x) thay vì thành phần FOA thứ ba. Nhìn chung, như được thể hiện bởi các ví dụ trên đây, tín hiệu tham chiếu tối ưu $P_{\text{ref},j}(k, n)$ có thể được tìm thấy bởi tổ hợp tuyến tính của các tín hiệu âm thanh trộn giảm thu được, tức là,

$$P_{\text{ref},j}(k, n) = A_{1,j}D_1(k, n) + A_{2,j}D_2(k, n)$$

trong đó các trọng số $A_{1,j}$ và $A_{2,j}$ của sự tổ hợp tuyến tính phụ thuộc vào siêu dữ liệu trộn giảm, tức là, trên cấu hình kênh vận chuyển và tín hiệu tham chiếu thứ j được xem xét (ví dụ, khi kết xuất đến loa phát thanh thứ j).

Lưu ý rằng sự tổng hợp các thành phần FOA hoặc HOA từ thành phần đẳng hướng sử dụng siêu dữ liệu không gian được mô tả, ví dụ trong [Thiergart17].

Nhìn chung, điều quan trọng là lưu ý rằng sự suy giảm đáng kể chất lượng âm thanh không gian phải được mong đợi nếu sự tổng hợp âm thanh không gian sử dụng tín hiệu trộn giảm sai như tín hiệu tham chiếu cho sự kết xuất. Ví dụ, nếu tín hiệu trộn giảm tương ứng với micro hai hướng chỉ vào bên trái được sử dụng cho sự tạo ra tín hiệu kênh đầu ra cho loa phát thanh được đặt trong bán cầu bên phải, các thành phần tín hiệu bắt nguồn từ bán cầu bên trái của trường âm đầu vào sẽ được hướng chủ yếu đến bán cầu bên phải của hệ thống tái tạo dẫn đến ảnh không gian không chính xác của đầu ra.

Sự trộn giảm linh hoạt cho đầu vào âm thanh không gian tham số

Trong phương án này, đầu vào đến bộ mã hóa tương ứng với tín hiệu đầu vào âm không gian theo tham số, mà bao gồm các tín hiệu âm thanh có cấu hình mảng tùy ý bao gồm hai hoặc nhiều hơn hai micro cùng với các tham số không gian của âm không gian (ví dụ, DOA và độ khuếch tán).

Bộ mã hóa cho phương án này được thể hiện trên Fig.5. Các tín hiệu mảng micro được sử dụng để tạo ra một hoặc nhiều tín hiệu trộn giảm âm thanh trong khối “tạo ra trộn giảm”. Các tham số trộn giảm, mà mô tả cấu hình kênh vận chuyển (ví dụ, bằng cách nào tín hiệu trộn giảm được tính toán hoặc một số đặc tính của chúng), cùng với các tham số không gian biểu diễn siêu dữ liệu bộ mã hóa, mà được mã hóa trong khối “bộ mã hóa siêu dữ liệu”. Lưu ý rằng không có bước phân tích âm thanh không gian được yêu cầu cho đầu vào âm thanh không gian theo tham số (ngược với các phương án trước đó), vì các tham số không gian đã được cung cấp như được nhập vào bộ mã hóa. Tuy nhiên, lưu ý rằng các tham số không gian của tín hiệu đầu vào âm thanh không gian và các tham số không gian nằm trong dòng bit cho sự truyền dẫn được tạo ra bởi bộ mã hóa âm thanh không gian không phải giống nhau. Trong trường hợp này, sự chuyển mã hoặc ánh xạ của các tham số không gian đầu vào và các tham số được sử dụng cho sự truyền dẫn phải được thực hiện tại bộ mã hóa. Các tín hiệu âm thanh trộn giảm được mã hóa trong khối “bộ mã hóa lõi”, ví dụ, sử dụng bộ mã hóa-giải mã âm thanh trên cơ sở EVS. Các tín hiệu trộn giảm âm thanh được mã hóa và siêu dữ liệu được mã hóa tạo ra dòng bit mà được truyền dẫn đến bộ giải mã. Đối với bộ giải mã, sơ đồ khối tương tự trên Fig.4 áp dụng cho các phương án trước đó.

Sau đây, mô tả cách các tín hiệu trộn giảm âm thanh và siêu dữ liệu trộn giảm tương ứng có thể được tạo ra.

Trong ví dụ thứ nhất, các tín hiệu trộn giảm âm thanh được tạo ra bằng cách lựa chọn tập con của các tín hiệu micro đầu vào có sẵn. Sự lựa chọn có thể được thực hiện thủ công (ví dụ, dựa trên các thiết lập trước) hoặc tự động. Ví dụ, nếu các tín hiệu micro của mảng hình tròn đồng đều với các micro đẳng hướng được tạo không gian M được sử dụng như đầu vào cho bộ mã hóa âm thanh không gian và hai kênh vận chuyển trộn giảm âm thanh được sử dụng cho sự truyền dẫn, sự lựa chọn thủ công có thể bao gồm,

ví dụ, lựa chọn cặp tín hiệu tương ứng với micro tại phía trước và phía sau của mảng, hoặc cặp tín hiệu tương ứng với các micro tại bên trái và bên phải của mảng. Việc lựa chọn micro phía trước và phía sau như các tín hiệu trộn giảm cho phép sự phân biệt tốt giữa các âm phía trước và các âm từ phía sau khi tổng hợp âm không gian tại bộ giải mã. Tương tự, lựa chọn micro bên trái và bên phải sẽ cho phép sự phân biệt tốt các âm không gian dọc theo trục y khi kết xuất âm không gian tại phía bộ giải mã. Ví dụ, nếu nguồn âm được ghi được định vị tại phía bên trái của mảng micro, có sự chênh lệch về thời gian tới của tín hiệu nguồn lần lượt tại micro bên trái và bên phải. Nói cách khác, tín hiệu tới micro bên trái đầu tiên, và sau đó micro bên phải. Tại quy trình kết xuất tại bộ giải mã, do đó quan trọng để sử dụng tín hiệu trộn giảm được kết hợp với tín hiệu micro bên trái để kết xuất đến các loa phát thanh được đặt trong bán cầu bên trái và tương tự để sử dụng tín hiệu trộn giảm được kết hợp với tín hiệu micro bên phải để kết xuất với các loa phát thanh được đặt trong bán cầu bên phải. Nếu không, sự chênh lệch thời gian lần lượt nằm trong các tín hiệu trộn giảm bên trái và bên phải, sẽ được hướng đến các loa phát thanh theo cách không đúng và dấu hiệu cảm quan thu được bởi các tín hiệu loa phát thanh là không đúng, tức là, ảnh âm thanh không gian được nhận biết bởi người nghe cũng sẽ không chính xác. Tương tự, sẽ quan trọng để có thể tại bộ giải mã phân biệt giữa các kênh trộn giảm tương ứng với phía trước và phía sau hoặc phía trên và phía dưới để đạt được chất lượng kênh tối ưu.

Sự lựa chọn các tín hiệu micro phù hợp có thể được thực hiện bằng cách xem xét mặt phẳng Cartesian mà chứa hầu hết năng lượng âm thanh, hoặc điều được mong đợi để chứa hầu hết năng lượng âm liên quan. Để thực hiện sự lựa chọn tự động, người ta có thể thực hiện, ví dụ, sự cục bộ hóa nguồn âm thanh theo tình trạng kỹ thuật, và khi đó lựa chọn hai micro gần nhất với trục tương ứng với hướng nguồn. Khái niệm tương tự có thể được áp dụng, ví dụ, nếu mảng micro bao gồm M micro định hướng trùng nhau (ví dụ, hai hướng) thay vì các micro đẳng hướng được tạo không gian. Trong trường hợp này, người ta có thể lựa chọn hai micro định hướng mà được định hướng theo hướng và theo hướng đối diện của trục Cartesian mà chứa (hoặc hi vọng chứa) hầu hết năng lượng âm thanh.

Trong ví dụ thứ nhất này, siêu dữ liệu trộn giảm chứa thông tin liên quan về các micro được lựa chọn. Thông tin này có thể chứa, ví dụ, các vị trí micro của các micro

được lựa chọn (ví dụ, về các tọa độ tuyệt đối hoặc tương đối trong hệ tọa độ Cartesian) và hoặc các khoảng cách liên micro và/hoặc hướng (ví dụ, về các tọa độ trong hệ tọa độ có cực, tức là, về góc phương vị và góc nâng Φ_m và Θ_m). Ngoài ra, siêu dữ liệu trộn giảm có thể bao gồm thông tin về mẫu định hướng của các micro được lựa chọn, ví dụ, bằng cách sử dụng tham số thứ tự thứ nhất c_m được mô tả trước đó.

Trên phía bộ giải mã (Fig.4), siêu dữ liệu trộn giảm được sử dụng trong khối “hệ thống âm thanh không gian” để thu được chất lượng kết xuất tối ưu. Ví dụ, đối với đầu ra loa phát thanh (đầu ra MC), khi siêu dữ liệu trộn giảm biểu thị rằng hai micro đẳng hướng tại hai vị trí cụ thể được truyền dẫn như các tín hiệu trộn giảm, tín hiệu tham chiếu $P_{ref,j}(k,n)$, từ đó tín hiệu loa phát thanh được tạo ra như giải thích trước đó, có thể được lựa chọn để tương ứng với các tín hiệu trộn giảm mà có khoảng cách nhỏ nhất đến vị trí loa phát thanh thứ j. Tương tự, nếu siêu dữ liệu trộn giảm biểu thị rằng hai micro định hướng với hướng nhìn $\{\Phi_m, \Theta_m\}$ đã được truyền dẫn, $P_{ref,j}(k,n)$ có thể được lựa chọn để tương ứng với tín hiệu trộn giảm với hướng nhìn gần nhất về phía vị trí loa phát thanh. Cách khác, sự tổ hợp tuyến tính của các tín hiệu trộn giảm định hướng trùng nhau được truyền dẫn có thể được thực hiện, như được giải thích trong phương án thứ hai.

Khi tạo ra đầu ra FOA/HOA tại bộ giải mã, tín hiệu trộn giảm đơn có thể được lựa chọn (theo ý muốn) để tạo ra âm trực tiếp cho tất cả các thành phần FOA/HOA nếu siêu dữ liệu trộn giảm biểu thị rằng các micro đẳng hướng được tạo khoảng cách đã được truyền dẫn. Thực tế, mỗi micro đẳng hướng chứa thông tin tương tự trên âm trực tiếp sẽ được tái tạo do các đặc điểm đẳng hướng. Tuy nhiên, đối với sự tạo ra các tín hiệu tham chiếu âm khuếch tán $\tilde{P}_{ref,j}$, người ta có thể xem xét tất cả các tín hiệu trộn giảm đẳng hướng được truyền dẫn. Thực tế, nếu trường âm là khuếch tán, các tín hiệu trộn giảm đẳng hướng được tạo không gian sẽ được giải tương quan riêng phần sao cho ít sự giải tương quan được yêu cầu để tạo ra các tín hiệu tham chiếu không tương quan $\tilde{P}_{ref,j}$. Các tín hiệu tham chiếu không tương quan lẫn nhau có thể được tạo ra từ các tín hiệu âm trộn giảm được truyền dẫn bằng cách sử dụng, ví dụ, phương pháp kết xuất dựa trên hiệp phương sai được đề xuất trong [Vilkamo13].

Được biết đến rộng rãi rằng sự tương quan giữa các tín hiệu của hai micro trong trường âm khuếch tán phụ thuộc vào khoảng cách giữa các micro: khoảng cách các micro càng lớn, càng ít các tín hiệu ghi được trong trường âm khuếch tán được tương quan [Laitinen11]. Thông tin liên quan đến khoảng cách micro nằm trong các tham số trộn giảm có thể được sử dụng tại bộ giải mã để xác định bao nhiêu kênh trộn giảm phải được giải tương quan tổng hợp để phù hợp cho sự kết xuất các thành phần âm khuếch tán. Trong trường hợp các tín hiệu trộn giảm đã được giải tương quan đủ do các khoảng cách micro lớn, sự giải tương quan nhân tạo có thể thậm chí bị loại bỏ và bất kỳ thành phần lạ liên quan đến sự giải tương quan có thể được tránh.

Khi siêu dữ liệu trộn giảm biểu thị rằng, ví dụ, các tín hiệu micro định hướng trùng nhau đã được truyền dẫn như các tín hiệu trộn giảm, khi đó các tín hiệu tham chiếu $P_{\text{ref},j}(k, n)$ cho đầu ra FOA/HOA có thể được tạo ra như được giải thích trong phương án thứ hai.

Lưu ý rằng thay vì lựa chọn tập hợp gồm các micro như các tín hiệu âm thanh âm thanh trộn giảm trong bộ mã hóa, người ta có thể lựa chọn tất cả tín hiệu âm thanh đầu vào micro có sẵn (ví dụ, hai hoặc nhiều hơn hai) như tín hiệu âm thanh trộn giảm. Trong trường hợp này, siêu dữ liệu trộn giảm mô tả toàn bộ cấu hình mảng micro, ví dụ, về các vị trí micro Cartesian, các hướng nhìn micro Φ_m và Θ_m trong các tọa độ có cực, hoặc các hoạt động hướng micro về các tham số thứ tự thứ nhất c_m .

Trong ví dụ thứ hai, các tín hiệu âm thanh trộn giảm được tạo ra trong bộ mã hóa trong khối “tạo ra trộn giảm” sử dụng sự tổ hợp tuyến tính của các tín hiệu micro đầu vào, ví dụ, sự lọc không gian (tạo chùm). Trong trường hợp này, các tín hiệu trộn giảm $D_m(k, n)$ có thể được tính toán như

$$D_m(k, n) = \mathbf{w}_m^H \mathbf{x}(k, n)$$

Ở đây, $\mathbf{x}(k, n)$ là vectơ chứa tất cả các tín hiệu micro đầu vào và $\mathbf{w}_m^H$ là các trọng số cho sự tổ hợp tuyến tính, tức là, các trọng số của bộ lọc không gian hoặc bộ tạo chùm, cho tín hiệu trộn giảm âm thanh thứ m . Có các cách khác nhau để tính toán các bộ lọc không gian hoặc các bộ tạo chùm theo cách tối ưu [Veen88]. Trong nhiều trường hợp, hướng nhìn $\{\Phi_m, \Theta_m\}$ được định nghĩa, về phía bộ tạo chùm được định hướng. Các

trọng số bộ tạo chùm khi đó có thể được tính toán, ví dụ, như bộ tạo chùm làm trễ và cộng tổng hoặc bộ tạo chùm MVDR [Veen88]. Trong phương án này, hướng nhìn bộ tạo chùm $\{\Phi_m, \Theta_m\}$ được định nghĩa cho mỗi tín hiệu trộn giảm âm thanh. Điều này có thể được thực hiện thủ công (ví dụ, dựa trên các thiết lập trước) hoặc tự động theo cách tương tự như được mô tả trong phương án thứ hai. Hướng nhìn $\{\Phi_m, \Theta_m\}$ của các tín hiệu bộ tạo chùm, mà biểu diễn các tín hiệu trộn giảm âm thanh khác nhau, khi đó có thể biểu diễn siêu dữ liệu trộn giảm mà được truyền dẫn đến bộ giải mã trên Fig.4.

Ví dụ khác phù hợp đặc biệt khi sử dụng đầu ra loa phát thanh tại bộ giải mã (đầu ra MC). Trong trường hợp này, tín hiệu trộn giảm $D_m(k, n)$ được sử dụng như $P_{\text{ref},j}(k, n)$ mà hướng nhìn bộ tạo chùm gần nhất với hướng loa phát thanh. Hướng nhìn bộ tạo chùm được yêu cầu được mô tả bởi siêu dữ liệu trộn giảm.

Lưu ý rằng trong tất cả các ví dụ, cấu hình kênh vận chuyển, tức là, các tham số trộn giảm, có thể phụ thuộc thời gian - tần số được điều chỉnh, ví dụ, dựa trên các tham số không gian, tương tự trong các phương án trước đó.

Tiếp theo, các phương án khác của sáng chế này hoặc các phương án đã được mô tả trước đã được thảo luận đối với các khía cạnh tương tự hoặc bổ sung hoặc khác.

Tốt hơn là, bộ tạo sự biểu diễn vận chuyển 600 của Fig.6 bao gồm một hoặc một số đặc điểm được minh họa trên Fig.8a. Cụ thể, bộ xác định vị trí năng lượng 606 được cung cấp mà điều khiển khối 602. Khối 602 có thể bao gồm bộ lựa chọn để lựa chọn từ các tín hiệu hệ số Ambisonics khi đầu vào là tín hiệu FOA hoặc HOA. Cách khác, hoặc ngoài ra, bộ tổ hợp vị trí năng lượng 606 điều khiển bộ tổ hợp để tổ hợp tín hiệu hệ số Ambisonics. Ngoài ra, hoặc cách khác, sự lựa chọn từ sự biểu diễn đa kênh hoặc từ các tín hiệu micro được thực hiện. Trong trường hợp này, đầu vào có các tín hiệu micro hoặc sự biểu diễn đa kênh hơn là dữ liệu FOA hoặc HOA. Ngoài ra hoặc cách khác, sự tổ hợp kênh hoặc sự tổ hợp các tín hiệu micro được thực hiện như được biểu thị tại 602 trên Fig.8a. Đối với hai thay thế thấp hơn, sự biểu diễn đa kênh hoặc các tín hiệu micro được nhập vào.

Dữ liệu vận chuyển được tạo ra bởi một hoặc một vài khối trong số các khối 602 được nhập vào bộ tạo siêu dữ liệu vận chuyển 05 nằm trong bộ tạo sự biểu diễn vận chuyển 60 của Fig.6 để tạo ra siêu dữ liệu vận chuyển (được mã hóa) 610.

Một khối bất kỳ trong số các khối 602 tạo ra sự biểu diễn vận chuyển không được mã hóa 614 mà khi đó còn được mã hóa bởi bộ mã hóa lõi 603 sao cho được minh họa trên Fig.3 hoặc Fig.5.

Lưu ý rằng sự thực thi thực tế của bộ tạo sự biểu diễn vận chuyển 600 có thể bao gồm chỉ một khối đơn trong số các khối 602 trên Fig.8a hoặc hai hoặc nhiều hơn hai khối trong số các khối được minh họa trên Fig.8a. Ở trường hợp sau, bộ tạo siêu dữ liệu vận chuyển 605 được tạo cấu hình để còn bao gồm mục siêu dữ liệu vận chuyển vào siêu dữ liệu vận chuyển 610 mà biểu thị phần (thời gian và/hoặc tần số) của sự biểu diễn âm thanh không gian nào, bất kỳ một trong số các thay thế được biểu thị tại mục 602 đã được lấy. Do đó, Fig.8a minh họa tình huống trong đó chỉ một trong số các thay thế 602 là hoạt động hoặc trong đó hai hoặc nhiều hơn hai thay thế đang hoạt động và sự chuyển mạch phụ thuộc tín hiệu có thể được thực hiện trong số các thay thế khác nhau cho sự tạo ra sự biểu diễn vận chuyển hoặc trộn giảm và siêu dữ liệu vận chuyển tương ứng.

Fig.8b minh họa bảng các thay thế siêu dữ liệu vận chuyển khác nhau mà có thể được tạo ra bởi bộ tạo sự biểu diễn vận chuyển 600 của Fig.6 và có thể được sử dụng bởi bộ tổng hợp âm thanh không gian của Fig.7. Các thay thế siêu dữ liệu vận chuyển bao gồm thông tin lựa chọn cho siêu dữ liệu biểu thị tập con nào của tập hợp các thành phần dữ liệu đầu vào âm thanh đã được lựa chọn như sự biểu diễn vận chuyển. Ví dụ, một ví dụ là chỉ hai hoặc ba thành phần trong số, ví dụ, bốn thành phần FOA đã được lựa chọn. Cách khác, thông tin lựa chọn có thể biểu thị các tín hiệu micro nào của mảng tín hiệu micro đã được lựa chọn. Thay thế khác của Fig.8b là tổ hợp thông tin biểu thị cách mà thành phần đầu vào biểu diễn âm thanh hoặc các tín hiệu nhất định đã được tổ hợp. Thông tin tổ hợp nhất định có thể viện dẫn đến các trọng số cho sự tổ hợp tuyến tính hoặc đến các kênh nào đã được tổ hợp, ví dụ, các trọng số bằng nhau hoặc được định nghĩa trước. Thông tin khác đề cập đến thông tin phản hoặc bán cầu được kết hợp với tín hiệu vận chuyển nhất định. Phần của thông tin bán cầu có thể đề cập đến phần bên trái hoặc phần bên phải hoặc phần phía trước hoặc phần phía sau so với vị trí nghe hoặc, cách khác, phần nhỏ hơn phần 180°.

Các phương án khác liên quan đến siêu dữ liệu vận chuyển biểu thị tham số hình dạng đề cập đến hình dạng của, ví dụ, sự định hướng micro vật lý hoặc ảo nhất định tạo

ra tín hiệu biểu diễn vận chuyển tương ứng. Tham số hình dạng có thể biểu thị hình dạng tín hiệu micro đẳng hướng hoặc hình dạng tín hiệu micro hai hướng hoặc hình dạng tín hiệu micro hai cực hoặc bất kỳ hình dạng liên quan khác. Các thay thế siêu dữ liệu vận chuyển liên quan đến các vị trí micro, các hướng micro, khoảng cách giữa các micro hoặc mẫu định hướng của micro mà có, ví dụ, được tạo ra hoặc được ghi các tín hiệu biểu diễn vận chuyển nằm trong sự biểu diễn vận chuyển (được mã hóa) 614. Các phương án khác về hướng nhìn hoặc nhiều hướng nhìn của tín hiệu nằm trong sự biểu diễn vận chuyển hoặc thông tin về các trọng số tạo chùm hoặc các hướng bộ tạo chùm hoặc, ngoài ra hoặc cách khác, liên quan đến việc liệu các tín hiệu micro được bao gồm là các tín hiệu micro đẳng hướng hoặc các tín hiệu micro đẳng hướng hoặc các tín hiệu micro hai hướng hoặc các tín hiệu khác. Rất ít thông tin phụ siêu dữ liệu vận chuyển (đối với tốc độ bit) có thể được tạo ra bằng cách đơn giản bao gồm cờ hiệu đơn biểu thị liệu các tín hiệu vận chuyển là các tín hiệu micro từ micro đẳng hướng hoặc từ micro bất kỳ khác với micro đẳng hướng.

Fig.8c minh họa sự thực thi được ưu tiên của bộ tạo siêu dữ liệu vận chuyển 605. Cụ thể, đối với siêu dữ liệu vận chuyển số học, bộ tạo siêu dữ liệu vận chuyển bao gồm bộ lượng tử hóa siêu dữ liệu vận chuyển 605a hoặc 622 và bộ mã hóa entropy siêu dữ liệu vận chuyển được kết nối tiếp theo 605b. Các quy trình được minh họa trên Fig.8c có thể cũng được áp dụng cho siêu dữ liệu tham số và, cụ thể, cho các tham số không gian.

Fig.9a minh họa sự thực thi được ưu tiên của bộ tổng hợp âm thanh 750 trên Fig.7. Bộ tổng hợp âm thanh không gian 750 bao gồm bộ phân tách siêu dữ liệu vận chuyển để phiên dịch siêu dữ liệu vận chuyển (được giải mã) 710. Dữ liệu đầu ra từ khối 752 được đưa vào bộ tổ hợp/bộ lựa chọn/bộ tạo tín hiệu tham chiếu 760 mà, ngoài ra, nhận tín hiệu vận chuyển 711 như nằm trong sự biểu diễn vận chuyển thu được từ giao diện đầu vào 700 của Fig.7. Dựa trên siêu dữ liệu vận chuyển, bộ tổ hợp/bộ lựa chọn/bộ tạo tín hiệu tham chiếu tạo ra một hoặc nhiều tín hiệu tham chiếu và chuyển tiếp các tín hiệu tham chiếu này đến bộ tính toán tín hiệu thành phần 770 mà tính toán các thành phần của sự biểu diễn âm thanh không gian được tổng hợp như các thành phần chung cho đầu ra đa kênh, các thành phần Ambisonics cho đầu ra FOA hoặc HOA, các kênh bên trái

và bên phải cho sự biểu diễn âm lập thể hoặc các thành phần đối tượng âm thanh, trong đó thành phần đối tượng âm thanh là tín hiệu đối tượng đơn âm hoặc âm lập thể.

Fig.9b minh họa tín hiệu âm thanh được mã hóa bao gồm, ví dụ, n tín hiệu vận chuyển T1, T2, T_n được biểu thị tại mục 611 và, ngoài ra, bao gồm siêu dữ liệu vận chuyển 610 và các tham số không gian tùy chọn 612. Thứ tự của các khối dữ liệu khác nhau và kích thước của khối dữ liệu nhất định đối với khối dữ liệu khác chỉ được minh họa dạng giản lược trên Fig.9b.

Fig.9c minh họa bảng tổng quan cho quy trình của bộ tổ hợp/bộ lựa chọn/bộ tạo tín hiệu tham chiếu 760 đối với siêu dữ liệu vận chuyển nhất định, sự biểu diễn vận chuyển nhất định và thiết lập loa nhất định. Cụ thể, trong phương án trên Fig.9c, sự biểu diễn vận chuyển bao gồm tín hiệu vận chuyển bên trái (hoặc tín hiệu vận chuyển phía trước hoặc tín hiệu đẳng hướng hoặc hai hướng) và sự biểu diễn vận chuyển còn bao gồm tín hiệu vận chuyển thứ hai T2 ví dụ, là tín hiệu vận chuyển bên phải (hoặc tín hiệu vận chuyển phía sau, tín hiệu vận chuyển đẳng hướng hoặc tín hiệu vận chuyển hai hướng). Trong trường hợp bên trái/bên phải, tín hiệu tham chiếu cho loa bên trái A được lựa chọn là tín hiệu vận chuyển thứ nhất T1 và tín hiệu tham chiếu cho loa bên phải được lựa chọn như tín hiệu vận chuyển T2. Đối với sự bao quanh bên trái và sự bao quanh bên phải, các tín hiệu bên trái và bên phải được lựa chọn như được nêu trong bảng 771 cho các kênh tương ứng. Đối với kênh ở giữa, tổng của tín hiệu vận chuyển bên trái và bên phải T1 và T2 được lựa chọn như tín hiệu tham chiếu cho thành phần kênh ở giữa của sự biểu diễn âm thanh không gian được tổng hợp.

Trên Fig.9c, sự lựa chọn khác được minh họa khi tín hiệu vận chuyển thứ nhất T1 là tín hiệu vận chuyển phía trước và tín hiệu vận chuyển thứ hai T2 là tín hiệu vận chuyển bên phải. Khi đó, tín hiệu vận chuyển thứ nhất T1 được lựa chọn cho tín hiệu vận chuyển bên trái, bên phải, ở giữa, và tín hiệu vận chuyển thứ hai T2 được lựa chọn cho sự bao quanh bên trái và sự bao quanh bên phải.

Fig.9d minh họa sự thực thi được ưu tiên khác của bộ tổng hợp âm thanh không gian của Fig.7. Trong khối 910, dữ liệu vận chuyển hoặc trộn giảm được tính toán về sự lựa chọn Ambisonics thứ tự thứ nhất hoặc Ambisonics thứ tự cao hơn nhất định. Bốn thay thế lựa chọn khác nhau là, ví dụ, được minh họa trên Fig.9d trong đó, trong thay

thể thứ tư, chỉ hai tín hiệu vận chuyển T1, T2 được lựa chọn hơn là thành phần thứ ba mà, trong các thay thế khác, các thành phần đẳng hướng.

Tín hiệu tham chiếu cho các kênh (ảo) được xác định dựa trên dữ liệu trộn giảm vận chuyển và quy trình dự trữ được sử dụng cho thành phần còn thiếu, tức là, cho thành phần thứ tư đối với các ví dụ trên Fig.9d hoặc đối với hai thành phần bị thiếu trong trường hợp ví dụ thứ tư. Khi đó, tại khối 912, các tín hiệu kênh được tạo ra sử dụng các tham số định hướng được nhận và suy ra từ dữ liệu vận chuyển. Do đó, các tham số định hướng hoặc không gian có thể được nhận thêm như được minh họa tại 712 trên Fig.7 hoặc có thể được suy ra từ sự biểu diễn vận chuyển bởi sự phân tích tín hiệu của các tín hiệu biểu diễn vận chuyển.

Trong sự thực thi thay thế, sự lựa chọn thành phần như thành phần FOA được thực hiện như được biểu thị trong khối 913 và sự tính toán của thành phần còn thiếu được thực hiện sử dụng phản hồi hàm cơ bản không gian như được minh họa tại mục 914 trên Fig.9d. Quy trình nhất định sử dụng phản hồi chức năng cơ sở không gian được minh họa trên Fig.10 tại khối 410 trong đó, trên Fig.10, khối 826 cung cấp phản hồi trung bình cho phần khuếch tán trong khi khối 410 trên Fig.10 cung cấp phản hồi cụ thể cho mỗi chế độ m và thứ tự l cho phần tín hiệu trực tiếp.

Fig.9e minh họa bảng khác biểu thị siêu dữ liệu vận chuyển nhất định cụ thể bao gồm tham số hình dạng hoặc hướng nhìn thêm vào tham số hình dạng hoặc thay thế tham số hình dạng. Tham số hình dạng có thể bao gồm thừa số hình dạng c_m bằng 1, 0,5 hoặc 0. Thừa số $c_m = 1$ biểu thị hình dạng đẳng hướng của micro ghi đặc điểm, trong khi thừa số 0,5 biểu thị hình dạng hai hướng và giá trị 0 biểu thị hình dạng hai cực.

Hơn nữa, các hướng nhìn khác nhau có thể bao gồm bên trái, bên phải, phía trước, phía sau, phía trên, phía dưới, hướng tới cụ thể bao gồm góc phương vị ϕ và góc nâng θ hoặc, cách khác, siêu dữ liệu ngắn bao gồm sự biểu thị rằng cặp tín hiệu trong sự biểu diễn vận chuyển bao gồm cặp bên trái/bên phải hoặc cặp phía trước/phía sau.

Trên Fig.9f, sự thực thi khác của bộ tổng hợp âm thanh không gian được minh họa trong đó, trong khối 910, siêu dữ liệu vận chuyển được đọc, ví dụ, như được thực hiện bởi giao diện đầu vào 700 của Fig.7 và cổng đầu vào của bộ tổng hợp âm thanh không gian 750. Trong khối 950, sự xác định tín hiệu tham chiếu được thích ứng với siêu dữ

liệu vận chuyển được đọc như được thực hiện, ví dụ, bởi khối 760. Khi đó, trong khối 916, đầu ra đa kênh, FOA/HOA, đối tượng hoặc lập thể và, cụ thể các thành phần cụ thể cho các loại đầu ra dữ liệu này được tính toán sử dụng tín hiệu tham chiếu thu được qua khối 915 và dữ liệu tham số được truyền dẫn tùy chọn 712 nếu có sẵn.

Fig.9g minh họa sự thực thi khác của bộ tổ hợp/bộ lựa chọn/bộ tạo tín hiệu tham chiếu 760. Khi siêu dữ liệu vận chuyển minh họa, ví dụ, rằng tín hiệu vận chuyển thứ nhất T11 là tín hiệu hai hướng bên trái và tín hiệu vận chuyển thứ hai T2 là tín hiệu hai hướng bên phải, khi đó, trong khối 920, tín hiệu đẳng hướng được tính toán bằng cách cộng T1 và T2. Như được liệt kê bởi khối 921, tín hiệu hai cực Y được tính toán bằng cách thu được sự chênh lệch giữa T1 và T2 hoặc sự chênh lệch giữa T2 và T1. Khi đó, trong khối 922, các thành phần còn lại được tổng hợp sử dụng tín hiệu đẳng hướng như sự tham chiếu. Tín hiệu đẳng hướng được sử dụng như sự tham chiếu trong khối 822 tốt hơn là đầu ra của khối 920. Ngoài ra, như được liệt kê tại mục 712, các tham số không gian tùy chọn có thể được sử dụng cũng cho sự tổng hợp các thành phần còn lại như các thành phần FoA hoặc HOA.

Fig.9h minh họa sự thực thi khác của các thay thế khác nhau cho quy trình mà có thể được thực hiện bởi bộ tổng hợp âm thanh không gian hoặc bộ tổ hợp/bộ lựa chọn/bộ tạo ra tín hiệu tham chiếu 760 khi, như được nêu ra trong khối 930, hai hoặc nhiều hơn hai tín hiệu micro được nhận biết như sự biểu diễn vận chuyển và siêu dữ liệu vận chuyển được kết hợp cũng được nhận. Như được nêu ra trong khối 931, sự lựa chọn có thể được thực hiện như tín hiệu tham chiếu cho thành phần tín hiệu nhất định, của tín hiệu vận chuyển với khoảng cách nhỏ nhất đến vị trí loa phát thanh, ví dụ, nhất định. Sự thay thế khác được minh họa trên khối 932 bao gồm sự lựa chọn tín hiệu micro với hướng nhìn gần nhất như tín hiệu tham chiếu cho loa nhất định hoặc với bộ tạo chùm gần nhất hoặc vị trí lỗi so với loa phát thanh nhất định hoặc nguồn âm ảo ví dụ, như bên trái/bên phải trong sự biểu diễn lập thể. Sự thay thế khác được minh họa trong khối 933 là sự lựa chọn tín hiệu vận chuyển tùy ý như tín hiệu tham chiếu cho tất cả các thành phần âm trực tiếp như đối với sự tính toán các thành phần FOA hoặc HOA hoặc cho sự tính toán tín hiệu loa phát thanh. Thay thế khác được minh họa tại 934 đề cập đến sự sử dụng tất cả các tín hiệu vận chuyển có sẵn như các tín hiệu đẳng hướng để tính toán các tín hiệu tham chiếu âm khuếch tán. Các thay thế khác về thiết lập hoặc hạn chế lượng tương quan cho

sự tính toán của tín hiệu thành phần dựa trên khoảng cách micro nằm trong siêu dữ liệu vận chuyển.

Đối với mục đích thực hiện một hoặc một vài thay thế 931 đến 935, một số siêu dữ liệu vận chuyển được kết hợp có ích mà được biểu thị ở phía bên phải của Fig.9h như bao gồm các vị trí micro của các micro chọn lọc, khoảng cách liên micro, các hướng micro hoặc các mẫu định hướng như c_m , sự mô tả mảng, các thừa số tạo chùm w_m hoặc hướng tới thực tế hoặc hướng âm với góc phương vị φ và góc nâng θ , ví dụ, đối với mỗi kênh vận chuyển.

Fig.10 minh họa sự thực thi được ưu tiên của bộ tạo các thành phần thứ tự giữa cho quy trình trực tiếp/khuếch tán. Cụ thể, bộ tạo các thành phần thứ tự thấp hoặc giữa bao gồm bộ tạo tín hiệu tham chiếu 821 mà nhận tín hiệu đầu vào và tạo ra tín hiệu tham chiếu bằng cách sao chép hoặc lấy như khi tín hiệu đầu vào là tín hiệu đơn hoặc bằng cách suy ra tín hiệu tham chiếu từ tín hiệu đầu vào bằng sự tính toán như được thảo luận trước hoặc như được thảo luận trong WO 2017/157803 A1 được đi kèm trong sáng chế như tài liệu tham khảo với toàn bộ sự dẫn dắt và tốt hơn là được điều khiển bởi siêu dữ liệu vận chuyển.

Hơn nữa, Fig.10 minh họa bộ tính toán độ khuếch đại định hướng 410 mà được tạo cấu hình để tính toán, từ thông tin DOA nhất định (Φ , θ) và từ chế độ nhất định số m và thứ tự nhất định số l và độ khuếch đại định hướng G_l^m . Trong phương án được ưu tiên, trong đó sự xử lý được thực hiện trong miền thời gian/tần số cho mỗi ô riêng lẻ được tham chiếu bởi k , n , độ khuếch đại định hướng được tính toán cho mỗi ô thời gian/tần số như vậy. Bộ gán trọng số 820 nhận tín hiệu tham chiếu và dữ liệu độ khuếch tán cho ô thời gian/tần số nhất định và kết quả của bộ gán trọng số 820 là phần trực tiếp. Phần khuếch tán được tạo ra bằng sự xử lý được thực hiện bởi bộ lọc giải tương quan 823 và bộ gán trọng số tiếp theo 824 nhận giá trị độ khuếch tán $\square$ cho khung thời gian nhất định và ngân tần số và, cụ thể, nhận phản hồi trung bình cho chế độ m và thứ tự l nhất định được biểu thị bởi D_l được tạo ra bởi bộ cung cấp phản hồi trung bình 826 mà nhận, như đầu vào, chế độ m được yêu cầu và thứ tự l được yêu cầu.

Kết quả của bộ gán trọng số 824 là phần khuếch tán và phần khuếch tán được thêm vào phần trực tiếp bằng bộ cộng 825 để thu được thành phần trường âm thứ tự giữa cho

chế độ nhất định m và thứ tự nhất định l . Tốt hơn là áp dụng độ khuếch đại bù khuếch tán được thảo luận đối với Fig.6 chỉ cho phần khuếch tán được tạo ra bởi khối 823. Điều này có thể được thực hiện một cách có lợi trong quy trình được thực hiện bởi bộ gán trọng số (khuếch tán). Do đó, chỉ phần khuếch tán trong tín hiệu được nâng cao để bù sự tổn hao năng lượng khuếch tán phải chịu bởi các thành phần cao hơn mà không nhận sự tổng hợp đầy đủ như được minh họa trên Fig.10.

Phần trực tiếp chỉ tạo ra như được minh họa trên Fig.11 cho bộ tạo các thành phần thứ tự cao hơn. Cụ thể, bộ tạo các thành phần thứ tự cao được thực thi theo cách tương tự như bộ tạo các thành phần thứ tự thấp hoặc giữa đối với nhánh trực tiếp nhưng không bao gồm các khối 823, 824, 825 và 826. Do đó, bộ tạo thành phần thứ tự cao hơn chỉ bao gồm bộ gán trọng số (trực tiếp) 822 nhận dữ liệu đầu vào từ bộ tính toán độ khuếch đại định hướng 410 và nhận tín hiệu tham chiếu từ bộ tạo tín hiệu tham chiếu 821. Tốt hơn là, chỉ tín hiệu tham chiếu đơn cho bộ tạo các thành phần thứ tự cao và bộ tạo các thành phần thứ tự thấp và giữa được tạo ra. Tuy nhiên, cả hai khối cũng có thể có các bộ tạo tín hiệu tham chiếu riêng lẻ như trường hợp có thể. Tuy nhiên, tốt hơn là chỉ có bộ tạo tín hiệu tham chiếu đơn. Do đó, sự xử lý được thực hiện bởi bộ tạo các thành phần thứ tự cao là vô cùng hiệu quả, vì chỉ hướng gán trọng số đơn với độ khuếch đại định hướng nhất định G_l^m với thông tin độ khuếch tán nhất định $\square$ cho ô thời gian/tần số sẽ được thực hiện. Do đó, các thành phần trường âm thứ tự cao có thể được tạo ra vô cùng hiệu quả và nhanh chóng và bất kỳ lỗi do sự không tạo ra các thành phần khuếch tán hoặc sự không sử dụng các thành phần khuếch tán trong tín hiệu đầu ra được bù một cách dễ dàng bằng cách nâng cao các thành phần trường âm thứ tự thấp hoặc tốt hơn là chỉ phần khuếch tán của các thành phần trường âm thứ tự giữa. Quy trình được minh họa trên Fig.11 cũng có thể được sử dụng cho sự tạo ra thành phần thứ tự thấp hoặc giữa.

Fig.10, do đó, minh họa sự tạo ra các thành phần trường âm thứ tự thấp hoặc giữa mà có phần khuếch tán, trong khi Fig.11 minh họa quy trình tính toán các thành phần trường âm thứ tự cao hoặc, nhìn chung, các thành phần không yêu cầu hoặc không nhận bất kỳ phần khuếch tán nào.

Tuy nhiên, trong sự tạo ra các thành phần trường âm, cụ thể đối với sự biểu diễn FoA hoặc HOA, quy trình của Fig.10 với phần khuếch tán hoặc quy trình của Fig.11 không có phần khuếch tán có thể được áp dụng. Bộ tạo tín hiệu tham chiếu 821, 760 được điều khiển trong cả hai quy trình trên Fig.10 và Fig.11 bằng siêu dữ liệu vận chuyển. Hơn nữa, bộ gán trọng số 822 được điều khiển không chỉ bằng phản hồi hàm cơ bản không gian G_1^n mà tốt hơn là cũng bằng các tham số không gian như các tham số độ khuếch tán 712, 722. Hơn nữa, trong phương án được ưu tiên, bộ gán trọng số 824 cho phần khuếch tán cũng được điều khiển bởi siêu dữ liệu vận chuyển và, cụ thể, bởi khoảng cách micro. Mỗi quan hệ nhất định giữa khoảng cách micro D và thừa số gán trọng số W được minh họa trong bản vẽ giản lược trên Fig.10. Khoảng cách cao D dẫn đến thừa số gán trọng số nhỏ và khoảng cách nhỏ dẫn đến thừa số gán trọng số cao. Do đó, khi có hai tín hiệu micro nằm trong sự biểu diễn tín hiệu vận chuyển mà có khoảng cách đến với nhau cao, người ta có thể giả định rằng cả hai tín hiệu micro đã hơi giải tương quan và, do đó, đầu ra của bộ lọc giải tương quan có thể được gán trọng số với thừa số gán trọng số gần 0 để, cuối cùng, tín hiệu được nhập vào bộ cộng 825 rất nhỏ so với tín hiệu được nhập vào bộ cộng từ bộ gán trọng số trực tiếp 822. Trong trường hợp cực hạn, nhánh tương quan thậm chí có thể được tắt mà có thể, ví dụ, đạt được bằng cách thiết lập trọng số $W = 0$. Một cách tự nhiên, có các cách khác để tắt nhánh khuếch tán bằng cách sử dụng sự chuyển mạch được tính toán bởi thuật toán ngưỡng hoặc như vậy.

Một cách tự nhiên, sự tạo ra thành phần được minh họa trên Fig.10 có thể được thực hiện chỉ bằng cách điều khiển bộ tạo tín hiệu tham chiếu 821, 760 bằng siêu dữ liệu vận chuyển mà không điều khiển bộ gán trọng số 804 hoặc, cách khác, chỉ bằng cách điều khiển bộ gán trọng số 804 mà không có sự tạo ra tín hiệu tham chiếu bất kỳ điều khiển khối 821, 760.

Fig.11 minh họa tình huống trong đó nhánh khuếch tán bị thiếu và trong đó, do đó, bất kỳ sự điều khiển bộ gán trọng số khuếch tán 824 của Fig.10 cũng không được thực hiện.

Fig.10 và Fig.12 minh họa bộ tạo tín hiệu khuếch tán nhất định 830 bao gồm bộ lọc giải tương quan 823 và bộ gán trọng số 824. Một cách tự nhiên, thứ tự trong sự xử

lý tín hiệu giữa bộ gán trọng số 824 và bộ lọc giải tương quan 823 có thể được trao đổi để sự gán trọng số của tín hiệu tham chiếu được tạo ra hoặc được xuất ra bằng bộ tạo tín hiệu tham chiếu 821, 760 được thực hiện trước tín hiệu được nhập vào bộ giải tương quan 823.

Trong khi Fig.10 minh họa sự tạo ra các thành phần trường âm thứ tự thấp hoặc giữa của sự biểu diễn thành phần trường âm như FOA hoặc HOA, tức là, sự biểu diễn với các tín hiệu thành phần hình cầu hoặc hình trụ, Fig.12 minh họa sự thực thi thay thế hoặc chung cho sự tính toán các tín hiệu thành phần loa phát thanh hoặc các đối tượng. Cụ thể, đối với sự tạo ra và tính toán các tín hiệu/đối tượng loa phát thanh, bộ tạo tín hiệu tham chiếu 821, 760 được cung cấp mà tương ứng với khối 760 của Fig.9a. Hơn nữa, bộ tính toán tín hiệu thành phần 770 được minh họa trên Fig.9a bao gồm, đối với nhánh trực tiếp, bộ gán trọng số 822, và, đối với nhánh khuếch tán, bộ tạo tín hiệu khuếch tán 830 bao gồm bộ lọc giải tương quan 823 và bộ gán trọng số 824. Hơn nữa, bộ tính toán tín hiệu thành phần 770 của Fig.9a còn bao gồm bộ cộng 825 mà thực hiện phép cộng tín hiệu trực tiếp P_{dir} và tín hiệu khuếch tán P_{diff} . Đầu ra của bộ cộng là tín hiệu loa phát thanh (ảo) hoặc tín hiệu đối tượng hoặc tín hiệu lập thể như được biểu thị bởi các số tham chiếu minh họa 755, 756. Cụ thể, bộ tính toán tín hiệu tham chiếu 821, 760 được điều khiển bởi siêu dữ liệu vận chuyển 710 và bộ gán trọng số khuếch tán 824 có thể cũng được điều khiển bởi siêu dữ liệu vận chuyển 710. Nhìn chung, bộ tính toán tín hiệu thành phần tính toán phần trực tiếp, ví dụ, sử dụng các độ khuếch đại quét như độ khuếch đại quét biên độ cơ sở ảo (virtual base amplitude panning - VBAP). Các độ khuếch đại được suy ra từ thông tin hướng tới, tốt hơn là được cho bởi góc phương vị φ và góc nâng θ . Điều này dẫn đến phần trực tiếp P_{dir} .

Hơn nữa, tín hiệu tham chiếu được tạo ra bởi bộ tính toán tín hiệu tham chiếu P_{ref} được nhập vào bởi bộ lọc giải tương quan 823 để thu được tín hiệu tham chiếu được giải tương quan và khi đó tín hiệu được gán trọng số, tốt hơn là sử dụng tham chiếu độ khuếch tán và cũng tốt hơn là sử dụng khoảng cách micro thu được từ siêu dữ liệu vận chuyển 710. Đầu ra của bộ gán trọng số P_{diff} và bộ cộng 825 cộng thành phần trực tiếp và thành phần khuếch tán để thu được tín hiệu loa phát thanh nhất định hoặc tín hiệu đối tượng hoặc kênh lập thể cho sự biểu diễn tương ứng. Cụ thể, khi các tín hiệu loa phát thanh ảo được tính toán bằng quy trình được thực hiện bởi bộ tạo tín hiệu tham chiếu

821, 760 phản hồi siêu dữ liệu vận chuyển có thể được thực hiện như được minh họa trên Fig.9c, cách khác, các tín hiệu tham chiếu có thể được tạo ra như các kênh chỉ ra từ vị trí nghe được định nghĩa đến loa cụ thể, và sự tính toán này của tín hiệu tham chiếu có thể được thực hiện sử dụng sự tổ hợp tuyến tính của các tín hiệu nằm trong sự biểu diễn vận chuyển.

Các phương án được ưu tiên của sáng chế như danh sách

Đầu vào trên cơ sở FOA

- Bộ mã hóa ngữ cảnh âm thanh không gian
 - Nhận các tín hiệu đầu vào âm thanh không gian biểu diễn ngữ cảnh âm thanh không gian (ví dụ, các thành phần FOA)
 - Tạo ra hoặc nhận các tham số âm thanh không gian bao gồm ít nhất một tham số hướng
 - Tạo ra tín hiệu âm thanh âm thanh trộn giảm dựa trên các tín hiệu đầu vào âm thanh thu được (Tùy chọn: sử dụng cũng các tham số âm thanh không gian cho sự tạo ra trộn giảm thích ứng).
 - Tạo ra các tham số trộn giảm mô tả các đặc tính định hướng của các tín hiệu trộn giảm (ví dụ, các hệ số trộn giảm hoặc các mẫu định hướng).
 - Mã hóa các tín hiệu trộn giảm, các tham số âm thanh không gian và các tham số trộn giảm.
- Bộ giải mã ngữ cảnh âm thanh không gian
 - Nhận ngữ cảnh âm thanh không gian được mã hóa bao gồm tín hiệu âm thanh trộn giảm, các tham số âm thanh không gian và các tham số trộn giảm
 - Giải mã các tín hiệu âm thanh trộn giảm, các tham số âm thanh không gian và các tham số kênh trộn giảm/vận chuyển
 - Bộ kết xuất âm thanh không gian để kết xuất về không gian sự biểu diễn được giải mã dựa trên các tín hiệu âm thanh trộn giảm, các tham số âm thanh không gian và các tham số (vị trí) trộn giảm

Đầu vào dựa trên các bản ghi micro được tạo khoảng cách và siêu dữ liệu không gian được kết hợp (đầu vào âm không gian theo tham số):

- Bộ mã hóa ngữ cảnh âm thanh không gian
 - Tạo ra hoặc nhận ít nhất hai tín hiệu đầu vào âm thanh không gian được tạo ra từ các tín hiệu micro được ghi
 - Tạo ra hoặc nhận các tham số âm thanh không gian bao gồm ít nhất một tham số hướng
 - Tạo ra hoặc nhận các tham số vị trí mô tả các đặc điểm địa lý hoặc vị trí của các tín hiệu đầu vào âm thanh không gian được tạo ra từ các tín hiệu micro được ghi (ví dụ, vị trí tương đối hoặc các không gian liên micro).
 - Mã hóa các tín hiệu đầu vào âm thanh không gian hoặc các tín hiệu trộn giảm được suy ra từ các tín hiệu đầu vào âm thanh không gian, các tham số âm thanh không gian và các tham số vị trí.
- Bộ giải mã ngữ cảnh âm thanh không gian
 - Nhận ngữ cảnh âm thanh không gian được mã hóa bao gồm ít nhất hai tín hiệu âm thanh, các tham số âm thanh không gian và các tham số vị trí (liên quan đến các đặc tính vị trí của các tín hiệu âm thanh).
 - Giải mã các tín hiệu âm thanh, các tham số âm thanh không gian và các tham số vị trí
 - Bộ kết xuất âm thanh không gian để kết xuất về không gian sự biểu diễn được giải mã dựa trên các tín hiệu âm thanh, các tham số âm thanh không gian và các tham số vị trí.

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 phạm vi của bước phương pháp cũng biểu diễn sự mô tả của khối hoặc mục hoặc dấu hiệu tương ứng của thiết bị tương ứng.

Phụ thuộc vào các yêu cầu thực thi nhất định, các phương án của sáng chế có thể được thực thi trong phần cứng hoặc trong phần mềm. Phương án 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 khả 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 thi như sản phẩm chương trình máy tính với mã chương trình, mã chương trình hoạt động để thực hiện một trong 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 chuyển tiếp.

Nói cách khác, do đó, 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.

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

Phương án khác của các phương pháp theo sáng chế là, dòng dữ liệu hoặc chuỗi tín hiệu thể hiện chương trình máy tính để thực hiện một trong các phương pháp đã được mô tả ở đây. Ví dụ, có thể 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 Liên mạng.

Phương án khác gồm có phương tiện xử lý, ví dụ máy tính, hoặc thiết bị logic lập trình được, được 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 máy tính 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 phần 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 cải biên và biến đổi của các phương án và các chi tiết được mô tả ở đây sẽ là rõ ràng đối với người có trình độ trong lĩnh vực kỹ thuật tương ứng. Do đó, mục đích của sáng chế là chỉ được 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 bản mô tả và giải thích của các phương án ở đây.

YÊU CẦU BẢO HỘ

1. Thiết bị mã hóa sự biểu diễn âm thanh không gian biểu diễn ngữ cảnh âm thanh để thu được tín hiệu âm thanh được mã hóa, thiết bị này bao gồm:

bộ tạo sự biểu diễn vận chuyển để tạo ra sự biểu diễn vận chuyển từ sự biểu diễn âm thanh không gian, và để tạo ra siêu dữ liệu vận chuyển về sự tạo ra sự biểu diễn vận chuyển; và

giao diện đầu ra để tạo ra tín hiệu âm thanh được mã hóa, tín hiệu âm thanh được mã hóa bao gồm thông tin về sự biểu diễn vận chuyển, và thông tin về siêu dữ liệu vận chuyển,

trong đó sự biểu diễn âm thanh không gian là sự biểu diễn Ambisonics thứ tự thứ nhất hoặc Ambisonics thứ tự cao hơn bao gồm vô số tín hiệu hệ số, hoặc sự biểu diễn đa kênh bao gồm nhiều kênh âm thanh,

trong đó bộ tạo sự biểu diễn vận chuyển được tạo cấu hình để tổ hợp hai hoặc nhiều hơn hai tín hiệu hệ số từ sự biểu diễn Ambisonics thứ tự cao hơn hoặc sự biểu diễn Ambisonics thứ nhất, hoặc để tổ hợp hai hoặc nhiều hơn hai kênh âm thanh từ sự biểu diễn đa kênh, và

trong đó bộ tạo sự biểu diễn vận chuyển được tạo cấu hình để tạo ra, như siêu dữ liệu vận chuyển, thông tin về cách hai hoặc nhiều hơn hai tín hiệu hệ số từ sự biểu diễn Ambisonics thứ tự thứ nhất hoặc sự biểu diễn Ambisonics thứ tự cao hoặc hai hoặc nhiều hơn hai kênh âm thanh từ sự biểu diễn đa kênh đã được tổ hợp, hoặc tín hiệu nào trong số hai hoặc nhiều hơn hai tín hiệu hệ số từ sự biểu diễn Ambisonics thứ tự thứ nhất hoặc sự biểu diễn Ambisonics thứ tự cao hơn hoặc kênh nào trong số hai hoặc nhiều hơn hai kênh âm thanh từ sự biểu diễn đa kênh đã được tổ hợp.

2. Thiết bị theo điểm 1, thiết bị còn bao gồm bộ xử lý tham số để suy ra các tham số không gian từ sự biểu diễn âm thanh không gian,

trong đó tín hiệu âm thanh được mã hóa còn bao gồm thông tin về tham số không gian.

3. Thiết bị theo điểm 1 hoặc 2,

trong đó bộ tạo sự biểu diễn vận chuyển được tạo cấu hình để xác định liệu năng lượng âm trội bắt nguồn từ phần hoặc bán cầu cụ thể như bán cầu bên trái hoặc bán cầu bên phải hoặc bán cầu phía trước hoặc bán cầu phía sau,

trong đó bộ tạo sự biểu diễn vận chuyển được tạo cấu hình để tạo ra tín hiệu vận chuyển thứ nhất từ phần hoặc bán cầu cụ thể, nơi mà năng lượng âm trội bắt nguồn, và tín hiệu vận chuyển thứ hai từ phần hoặc bán cầu khác nhau như phần hoặc bán cầu có hướng ngược so với vị trí tham chiếu và so với phần hoặc bán cầu cụ thể, và

trong đó bộ tạo sự biểu diễn vận chuyển được tạo cấu hình để xác định siêu dữ liệu vận chuyển để siêu dữ liệu vận chuyển bao gồm thông tin nhận biết phần hoặc bán cầu cụ thể, hoặc nhận biết phần hoặc bán cầu khác nhau.

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

trong đó bộ tạo sự biểu diễn vận chuyển được tạo cấu hình để tổ hợp các tín hiệu hệ số của sự biểu diễn âm thanh không gian để tín hiệu thu được thứ nhất là tín hiệu vận chuyển thứ nhất tương ứng với tín hiệu micro định hướng được hướng đến phần hoặc bán cầu cụ thể, và tín hiệu thu được thứ hai là tín hiệu vận chuyển thứ hai tương ứng với tín hiệu micro định hướng được hướng đến phần hoặc bán cầu khác nhau.

5. Thiết bị theo điểm bất kỳ trong số các điểm nêu trên,

trong đó bộ tạo sự biểu diễn vận chuyển được tạo cấu hình để tạo ra sự biểu diễn vận chuyển và siêu dữ liệu vận chuyển theo cách biến thiên thời gian hoặc phụ thuộc tần số, để sự biểu diễn vận chuyển và siêu dữ liệu vận chuyển cho khung thứ nhất khác với sự biểu diễn vận chuyển và siêu dữ liệu vận chuyển cho khung thứ hai, hoặc để sự biểu diễn vận chuyển và siêu dữ liệu vận chuyển cho băng tần thứ nhất khác với sự biểu diễn vận chuyển và siêu dữ liệu vận chuyển cho băng tần khác thứ hai.

6. Thiết bị theo điểm bất kỳ trong số các điểm nêu trên,

trong đó bộ tạo sự biểu diễn vận chuyển được tạo cấu hình để tạo ra một hoặc hai tín hiệu vận chuyển bằng sự tổ hợp được gán trọng số gồm hai hoặc nhiều hơn hai tín hiệu hệ số của sự biểu diễn âm thanh không gian, và

trong đó bộ tạo sự biểu diễn vận chuyển được tạo cấu hình để tính toán siêu dữ liệu vận chuyển để siêu dữ liệu vận chuyển bao gồm thông tin về trọng số được sử dụng

trong tổ hợp được gán trọng số, hoặc thông tin về góc phương vị và/hoặc góc nâng như hướng nhìn của tín hiệu micro định hướng được tạo ra, hoặc thông tin về tham số biểu thị đặc điểm định hướng của tín hiệu micro định hướng.

7. Thiết bị theo điểm bất kỳ trong số các điểm nêu trên,

trong đó bộ tạo sự biểu diễn vận chuyển được tạo cấu hình để tạo ra siêu dữ liệu vận chuyển định lượng, để lượng tử hóa siêu dữ liệu vận chuyển định lượng để thu được siêu dữ liệu vận chuyển, và mã hóa entropy siêu dữ liệu vận chuyển được lượng tử hóa, và trong đó giao diện đầu ra được tạo cấu hình để bao gồm siêu dữ liệu vận chuyển được mã hóa thành tín hiệu âm thanh được mã hóa.

8. Thiết bị theo điểm bất kỳ trong số các điểm nêu trên,

trong đó bộ tạo sự biểu diễn vận chuyển được tạo cấu hình để vận chuyển siêu dữ liệu vận chuyển vào chỉ số bảng hoặc tham số được thiết lập trước, và

trong đó giao diện đầu ra được tạo cấu hình để bao gồm chỉ số bảng hoặc tham số được thiết lập trước vào tín hiệu âm thanh được mã hóa.

9. Thiết bị theo điểm bất kỳ trong số các điểm nêu trên,

trong đó sự biểu diễn âm thanh không gian bao gồm ít nhất hai tín hiệu âm thanh và các tham số không gian,

trong đó bộ xử lý tham số được tạo cấu hình để suy ra các tham số không gian từ sự biểu diễn âm thanh không gian bằng cách trích các tham số không gian từ sự biểu diễn âm thanh không gian,

trong đó giao diện đầu ra được tạo cấu hình để bao gồm thông tin về các tham số không gian vào tín hiệu âm thanh được mã hóa hoặc để bao gồm thông tin về các tham số không gian được xử lý được suy ra từ các tham số không gian thành tín hiệu âm thanh được mã hóa, hoặc

trong đó bộ tạo sự biểu diễn vận chuyển được tạo cấu hình để tổ hợp ít nhất hai tín hiệu âm thanh hoặc tập con gồm ít nhất hai tín hiệu âm thanh và tính toán siêu dữ liệu vận chuyển sao cho siêu dữ liệu vận chuyển bao gồm thông tin về sự tổ hợp của các tín

hiệu âm thanh được thực hiện để tính toán sự biểu diễn vận chuyển của sự biểu diễn âm thanh không gian.

10. Thiết bị theo điểm bất kỳ trong số các điểm nêu trên,

trong đó bộ tạo sự biểu diễn vận chuyển được tạo cấu hình để tổ hợp các tín hiệu âm thanh nằm trong sự biểu diễn âm thanh không gian sử dụng sự lọc không gian hoặc sự tạo chùm, và

tạo sự biểu diễn vận chuyển được tạo cấu hình để bao gồm thông tin về hướng nhìn của sự biểu diễn vận chuyển hoặc thông tin về các trọng số tạo chùm được sử dụng để tính toán sự biểu diễn vận chuyển vào siêu dữ liệu vận chuyển.

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

trong đó sự biểu diễn âm thanh không gian là sự mô tả trường âm về vị trí tham chiếu, và

trong đó bộ xử lý tham số được tạo cấu hình để suy ra các tham số không gian từ sự biểu diễn âm thanh không gian, trong đó các tham số không gian định nghĩa các tham số biến thời gian hoặc phụ thuộc tần số về hướng tới của âm tại vị trí tham chiếu hoặc các tham số biến thời gian hoặc phụ thuộc tần số về độ khuếch tán của trường âm tại vị trí tham chiếu, hoặc

trong đó bộ tạo sự biểu diễn vận chuyển bao gồm bộ trộn giảm để tạo ra, như sự biểu diễn vận chuyển, sự biểu diễn trộn giảm có số lượng thứ hai các tín hiệu riêng nhỏ hơn số lượng thứ nhất các tín hiệu riêng nằm trong sự biểu diễn âm thanh không gian, trong đó bộ trộn giảm được tạo cấu hình để tổ hợp các tín hiệu riêng nằm trong sự biểu diễn âm thanh không gian để giảm số lượng thứ nhất các tín hiệu đến số lượng thứ hai các tín hiệu.

12. Thiết bị theo điểm bất kỳ trong số các điểm nêu trên,

trong đó bộ xử lý tham số bao gồm bộ phân tích âm thanh không gian để suy ra các tham số không gian từ sự biểu diễn âm thanh không gian bằng cách thực hiện sự phân tích tín hiệu âm thanh, và

trong đó bộ tạo sự biểu diễn vận chuyển được tạo cấu hình để tạo ra sự biểu diễn vận chuyển dựa trên kết quả của bộ phân tích âm thanh không gian, hoặc

trong đó sự biểu diễn vận chuyển bao gồm bộ mã hóa lõi để mã hóa lõi một hoặc nhiều tín hiệu âm thanh của các tín hiệu vận chuyển của sự biểu diễn vận chuyển, hoặc

trong đó bộ xử lý tham số được tạo cấu hình để lượng tử hóa và mã hóa entropy các tham số không gian, và

trong đó giao diện đầu ra được tạo cấu hình để bao gồm sự biểu diễn vận chuyển được mã hóa lõi như thông tin về sự biểu diễn vận chuyển vào tín hiệu âm thanh được mã hóa hoặc để bao gồm các tham số không gian được mã hóa entropy như thông tin về các tham số không gian vào tín hiệu âm thanh được mã hóa.

13. Thiết bị giải mã tín hiệu âm thanh được mã hóa, gồm có:

giao diện đầu vào để nhận tín hiệu âm thanh được mã hóa bao gồm thông tin về sự biểu diễn vận chuyển và thông tin về siêu dữ liệu vận chuyển; và

bộ tổ hợp âm thanh không gian để tổng hợp sự biểu diễn âm thanh không gian sử dụng thông tin về sự biểu diễn vận chuyển bao gồm các tín hiệu vận chuyển và thông tin về siêu dữ liệu vận chuyển,

trong đó siêu dữ liệu vận chuyển biểu thị tín hiệu vận chuyển như đề cập đến phần hoặc bán cầu thứ nhất về vị trí tham chiếu của sự biểu diễn âm thanh không gian và tín hiệu vận chuyển thứ hai như đề cập đến phần hoặc bán cầu khác thứ hai về vị trí tham chiếu của sự biểu diễn âm thanh không gian, trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để

tạo ra tín hiệu thành phần của sự biểu diễn âm thanh không gian được kết hợp với phần hoặc bán cầu thứ nhất sử dụng tín hiệu vận chuyển thứ nhất và không sử dụng tín hiệu vận chuyển thứ hai, hoặc

tạo ra tín hiệu thành phần khác của sự biểu diễn âm thanh không gian được kết hợp với phần hoặc bán cầu thứ hai sử dụng tín hiệu vận chuyển thứ hai và không sử dụng tín hiệu vận chuyển thứ nhất, hoặc

tạo ra tín hiệu thành phần khác của sự biểu diễn âm thanh không gian được kết hợp với phần hoặc bán cầu thứ nhất sử dụng tổ hợp của tín hiệu vận chuyển thứ nhất và thứ hai, hoặc

tạo ra tín hiệu thành phần khác của sự biểu diễn âm thanh không gian được kết hợp với phần hoặc bán cầu thứ hai khác sử dụng sự tổ hợp thứ hai gồm các tín hiệu vận chuyển thứ nhất và thứ hai, trong đó sự tổ hợp thứ nhất ảnh hưởng bởi tín hiệu vận chuyển thứ nhất mạnh hơn sự tổ hợp thứ hai, hoặc trong đó sự tổ hợp thứ hai ảnh hưởng bởi tín hiệu vận chuyển thứ hai mạnh hơn sự tổ hợp thứ nhất,

hoặc

trong đó siêu dữ liệu vận chuyển biểu thị thông tin về các hệ số gán trọng số hoặc các hướng nhìn của các tín hiệu vận chuyển của sự biểu diễn vận chuyển, và trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để tính toán các thành phần Ambisonics thứ tự thứ nhất khác của sự biểu diễn âm thanh không gian sử dụng thông tin về các hướng nhìn hoặc các hệ số gán trọng số, và sử dụng các tín hiệu vận chuyển.

14. Thiết bị theo điểm 13, trong đó giao diện đầu vào được tạo cấu hình để nhận tín hiệu âm thanh được mã hóa còn bao gồm thông tin về các tham số không gian, và

trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để tổng hợp sự biểu diễn âm thanh không gian còn sử dụng thông tin về các tham số không gian.

15. Thiết bị theo điểm 13 hoặc 14, trong đó bộ tổng hợp âm thanh không gian bao gồm:

bộ giải mã lỗi để giải mã lỗi hai hoặc nhiều hơn hai tín hiệu vận chuyển được mã hóa biểu diễn thông tin về sự biểu diễn vận chuyển để thu được hai hoặc nhiều hơn hai tín hiệu vận chuyển được giải mã, hoặc

trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để tính toán sự biểu diễn Ambisonics thứ tự thứ nhất hoặc Ambisonics thứ tự cao hơn hoặc sự biểu diễn đối tượng hoặc sự biểu diễn lập thể của sự biểu diễn âm thanh không gian, hoặc

trong đó bộ tổng hợp âm thanh không gian bao gồm bộ giải mã siêu dữ liệu để giải mã thông tin về siêu dữ liệu vận chuyển để suy ra siêu dữ liệu vận chuyển được giải mã

hoặc để giải mã thông tin về tham số không gian để thu được các tham số không gian được giải mã.

16. Thiết bị theo điểm 13, 14 hoặc 15,

trong đó sự biểu diễn âm thanh không gian bao gồm nhiều tín hiệu thành phần,

trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để xác định, đối với tín hiệu thành phần của sự biểu diễn âm thanh không gian, tín hiệu tham chiếu sử dụng thông tin về sự biểu diễn vận chuyển và thông tin về siêu dữ liệu vận chuyển, và

tính toán tín hiệu thành phần của sự biểu diễn âm thanh không gian sử dụng tín hiệu tham chiếu và thông tin về các tham số không gian, hoặc để tính toán tín hiệu thành phần của sự biểu diễn âm thanh không gian sử dụng tín hiệu tham chiếu.

17. Thiết bị theo điểm bất kỳ trong số các điểm từ 13 đến 16,

trong đó siêu dữ liệu vận chuyển bao gồm thông tin về đặc tính định hướng được kết hợp với các tín hiệu vận chuyển của sự biểu diễn vận chuyển,

trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để tính toán các tín hiệu micro ảo sử dụng các tín hiệu Ambisonics thứ tự thứ nhất hoặc Ambisonics thứ tự cao hơn, các vị trí loa phát thanh và siêu dữ liệu vận chuyển, hoặc

trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để xác định đặc điểm định hướng của các tín hiệu vận chuyển sử dụng siêu dữ liệu vận chuyển và để xác định thành phần Ambisonics thứ tự thứ nhất hoặc Ambisonics thứ tự cao hơn từ các tín hiệu vận chuyển cùng dòng với các đặc điểm định hướng được định trước của các tín hiệu vận chuyển, và

để xác định thành phần Ambisonics thứ tự thứ nhất hoặc Ambisonics thứ tự cao hơn khác không được kết hợp với các đặc điểm định hướng của các tín hiệu vận chuyển phù hợp với quy trình rút lui.

18. Thiết bị theo điểm bất kỳ trong số các điểm từ 13 đến 17,

trong đó siêu dữ liệu vận chuyển bao gồm thông tin về hướng nhìn thứ nhất được kết hợp với tín hiệu vận chuyển thứ nhất, và thông tin về hướng nhìn thứ hai được kết hợp với tín hiệu vận chuyển thứ hai,

trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để lựa chọn tín hiệu tham chiếu cho sự tính toán tín hiệu thành phần hoặc tín hiệu thành phần khác của sự biểu diễn âm thanh không gian dựa trên siêu dữ liệu vận chuyển và vị trí của loa phát thanh được kết hợp với tín hiệu thành phần của sự biểu diễn âm thanh không gian.

19. Thiết bị theo điểm 18,

trong đó hướng nhìn thứ nhất biểu thị bán cầu bên trái hoặc phía trước, trong đó hướng nhìn thứ hai biểu thị bán cầu bên phải hoặc phía sau,

trong đó, đối với sự tính toán tín hiệu thành phần hoặc tín hiệu thành phần khác cho loa phát thanh trong bán cầu bên trái, tín hiệu vận chuyển thứ nhất và không phải tín hiệu vận chuyển thứ hai được sử dụng, hoặc trong đó đối với sự tính toán tín hiệu thành phần hoặc tín hiệu thành phần khác cho tín hiệu loa phát thanh trong bán cầu bên phải, tín hiệu vận chuyển thứ hai và không phải tín hiệu vận chuyển thứ nhất được sử dụng, hoặc

trong đó đối với sự tính toán của tín hiệu thành phần hoặc tín hiệu thành phần khác cho loa phát thanh ở bán cầu phía trước, tín hiệu vận chuyển thứ nhất và không phải tín hiệu vận chuyển thứ hai được sử dụng, hoặc trong đó đối với sự tính toán của loa phát thanh ở bán cầu phía sau, tín hiệu vận chuyển thứ hai và không phải tín hiệu vận chuyển thứ nhất được sử dụng, hoặc

trong đó đối với sự tính toán tín hiệu thành phần hoặc tín hiệu thành phần khác cho loa phát thanh trong vùng trung tâm, tổ hợp gồm tín hiệu vận chuyển bên trái và tín hiệu vận chuyển bên phải được sử dụng, hoặc trong đó đối với sự tính toán tín hiệu thành phần hoặc tín hiệu thành phần khác cho tín hiệu loa phát thanh được kết hợp với loa phát thanh trong vùng giữa bán cầu phía trước và bán cầu phía sau, tổ hợp gồm tín hiệu vận chuyển thứ nhất và tín hiệu vận chuyển thứ hai được sử dụng.

20. Thiết bị theo điểm bất kỳ trong số các điểm từ 13 đến 19,

trong đó thông tin về siêu dữ liệu vận chuyển biểu thị, như hướng nhìn thứ nhất, hướng bên trái đối với tín hiệu vận chuyển thứ nhất và biểu thị, như hướng nhìn thứ hai, hướng nhìn bên phải cho tín hiệu vận chuyển thứ hai,

trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để tính toán thành phần Ambisonics thứ nhất bằng cách cộng tín hiệu vận chuyển thứ nhất và tín hiệu vận chuyển thứ hai, hoặc tính toán thành phần Ambisonics thứ hai bằng cách trừ tín hiệu vận chuyển thứ nhất và tín hiệu vận chuyển thứ hai, hoặc trong đó thành phần Ambisonics khác được tính toán sử dụng tổng của tín hiệu vận chuyển thứ nhất và tín hiệu vận chuyển thứ hai.

21. Thiết bị theo điểm bất kỳ trong số các điểm từ 13 đến 17,

trong đó siêu dữ liệu vận chuyển biểu thị, đối với tín hiệu vận chuyển thứ nhất, hướng nhìn phía trước và biểu thị, đối với tín hiệu vận chuyển thứ hai, hướng nhìn phía sau,

trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để tính toán thành phần Ambisonics thứ tự thứ nhất cho hướng x bằng cách thực hiện sự tính toán của độ chênh lệch giữa các tín hiệu vận chuyển thứ nhất và tín hiệu vận chuyển thứ hai, và để tính toán thành phần Ambisonics thứ tự thứ nhất đẳng hướng sử dụng phép cộng tín hiệu vận chuyển thứ nhất và tín hiệu vận chuyển thứ hai, và

tính toán thành phần Ambisonics thứ tự thứ nhất khác sử dụng tổng tín hiệu vận chuyển thứ nhất và tín hiệu vận chuyển thứ hai.

22. Thiết bị theo điểm 13,

trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để tính toán các thành phần Ambisonics thứ tự thứ nhất của sự biểu diễn âm thanh không gian còn sử dụng các tham số không gian.

23. Thiết bị theo điểm bất kỳ trong số các điểm từ 13 đến 22,

trong đó siêu dữ liệu vận chuyển bao gồm thông tin về các tín hiệu vận chuyển được suy ra từ các tín hiệu micro tại hai vị trí khác nhau hoặc với các hướng nhìn khác nhau,

trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để lựa chọn tín hiệu tham chiếu mà có vị trí gần nhất với vị trí loa phát thanh, hoặc lựa chọn tín hiệu tham chiếu mà có hướng nhìn gần nhất so với hướng từ vị trí tham chiếu của sự biểu diễn âm thanh không gian và vị trí loa phát thanh, hoặc để thực hiện sự tổ hợp tuyến tính với các

tín hiệu vận chuyển để xác định tín hiệu tham chiếu cho loa phát thanh được đặt giữa hai hướng nhìn được biểu thị bởi siêu dữ liệu vận chuyển.

24. Thiết bị theo điểm bất kỳ trong số các điểm từ 13 đến 23,

trong đó siêu dữ liệu vận chuyển bao gồm thông tin về khoảng cách giữa các vị trí micro được kết hợp với các tín hiệu vận chuyển,

trong đó bộ tổng hợp âm thanh không gian bao gồm bộ tạo tín hiệu khuếch tán, và trong đó bộ tạo tín hiệu khuếch tán được tạo cấu hình để điều khiển lượng tín hiệu giải tương quan trong tín hiệu khuếch tán được tạo ra bởi bộ tạo tín hiệu khuếch tán sử dụng thông tin về khoảng cách, để, đối với khoảng cách thứ nhất, lượng cao hơn của tín hiệu giải tương quan nằm trong tín hiệu khuếch tán so với lượng tín hiệu giải tương quan đối với khoảng cách thứ hai, trong đó khoảng cách thứ nhất thấp hơn khoảng cách thứ hai, hoặc

trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để tính toán, đối với khoảng cách thứ nhất giữa các vị trí micro, tín hiệu thành phần cho sự biểu diễn âm thanh không gian sử dụng tín hiệu đầu vào của bộ lọc giải tương quan D được tạo cấu hình để giải tương quan tín hiệu tham chiếu hoặc tín hiệu tham chiếu được định tỉ lệ và tín hiệu tham chiếu được gán trọng số sử dụng độ khuếch đại được suy ra từ thông tin hướng tới của âm và tính toán, đối với khoảng cách thứ hai giữa các vị trí micro, tín hiệu thành phần cho sự biểu diễn âm thanh không gian sử dụng tín hiệu tham chiếu được gán trọng số sử dụng độ khuếch đại được suy ra từ thông tin hướng tới của âm mà không có sự xử lý giải tương quan nào, khoảng cách thứ hai lớn hơn khoảng cách thứ nhất hoặc lớn hơn ngưỡng khoảng cách.

25. Thiết bị theo điểm bất kỳ trong số các điểm từ 13 đến 24,

trong đó siêu dữ liệu vận chuyển bao gồm thông tin về sự tạo chùm hoặc sự lọc không gian được kết hợp với các tín hiệu vận chuyển của sự biểu diễn vận chuyển, và

trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để tạo ra tín hiệu loa phát thanh cho loa phát thanh sử dụng tín hiệu vận chuyển có hướng nhìn gần nhất hướng nhìn từ vị trí tham chiếu của sự biểu diễn âm thanh không gian đến loa phát thanh.

26. Thiết bị theo điểm bất kỳ trong số các điểm từ 13 đến 25,

trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để xác định các tín hiệu thành phần của sự biểu diễn âm thanh không gian của thành phần âm thanh trực tiếp và thành phần âm khuếch tán, trong đó thành phần âm trực tiếp thu được bằng cách định tỉ lệ tín hiệu tham chiếu với thừa số phụ thuộc vào tham số khuếch tán hoặc tham số định hướng, trong đó tham số định hướng phụ thuộc vào hướng tới của âm, trong đó sự xác định của tín hiệu tham chiếu được thực hiện dựa trên thông tin về siêu dữ liệu vận chuyển, và trong đó thành phần âm khuếch tán được xác định sử dụng tín hiệu tham chiếu và tham số khuếch tán giống nhau.

27. Thiết bị theo điểm bất kỳ trong số các điểm từ 13 đến 26,

trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để xác định các tín hiệu thành phần của sự biểu diễn âm thanh không gian của thành phần âm thanh trực tiếp và thành phần âm khuếch tán, trong đó thành phần âm trực tiếp thu được bằng cách định tỉ lệ tín hiệu tham chiếu với thừa số phụ thuộc vào tham số khuếch tán hoặc tham số định hướng, trong đó tham số định hướng phụ thuộc vào hướng tới của âm, trong đó sự xác định của tín hiệu tham chiếu được thực hiện dựa trên thông tin về siêu dữ liệu vận chuyển, và trong đó thành phần âm khuếch tán được xác định sử dụng tín hiệu tham chiếu và tham số khuếch tán giống nhau.

28. Thiết bị một trong số các điểm từ 13 đến 27, trong đó sự biểu diễn vận chuyển bao gồm ít nhất hai tín hiệu micro khác nhau,

trong đó siêu dữ liệu vận chuyển bao gồm thông tin biểu thị, liệu ít nhất hai tín hiệu micro khác nhau là ít nhất một trong số các tín hiệu đẳng hướng, tín hiệu lưỡng cực hoặc tín hiệu hai hướng, và

trong đó bộ tổng hợp âm thanh không gian được tạo cấu hình để thích ứng sự xác định tín hiệu tham chiếu cho siêu dữ liệu vận chuyển để xác định, đối với các thành phần của sự biểu diễn âm thanh không gian, các tín hiệu tham chiếu riêng và tính toán thành phần tương ứng sử dụng tín hiệu tham chiếu riêng được xác định cho thành phần tương ứng.

29. Phương pháp mã hóa sự biểu diễn âm thanh không gian biểu diễn ngữ cảnh âm thanh để thu được tín hiệu âm thanh được mã hóa, phương pháp này bao gồm các bước:

tạo ra sự biểu diễn vận chuyển từ sự biểu diễn âm thanh không gian;

tạo ra siêu dữ liệu vận chuyển liên quan đến sự tạo ra sự biểu diễn vận chuyển; và

tạo ra tín hiệu âm thanh được mã hóa, tín hiệu âm thanh được mã hóa bao gồm thông tin về sự biểu diễn vận chuyển, và thông tin về siêu dữ liệu vận chuyển

trong đó sự biểu diễn âm thanh không gian là sự biểu diễn Ambisonics thứ tự thứ nhất hoặc Ambisonics thứ tự cao hơn bao gồm vô số tín hiệu hệ số, hoặc sự biểu diễn đa kênh bao gồm nhiều kênh âm thanh,

trong đó sự tạo ra sự vận chuyển bao gồm tổ hợp hai hoặc nhiều hơn hai tín hiệu hệ số từ sự biểu diễn Ambisonics thứ tự cao hơn hoặc sự biểu diễn Ambisonics thứ nhất, hoặc để tổ hợp hai hoặc nhiều hơn hai kênh âm thanh từ sự biểu diễn đa kênh, và

trong đó sự tạo ra siêu dữ liệu vận chuyển bao gồm tạo ra, như siêu dữ liệu vận chuyển, thông tin về cách hai hoặc nhiều hơn hai tín hiệu hệ số từ sự biểu diễn Ambisonics thứ tự thứ nhất hoặc sự biểu diễn Ambisonics thứ tự cao hoặc hai hoặc nhiều hơn hai kênh âm thanh từ sự biểu diễn đa kênh đã được tổ hợp, hoặc tín hiệu nào trong số hai hoặc nhiều hơn hai tín hiệu hệ số từ sự biểu diễn Ambisonics thứ tự thứ nhất hoặc sự biểu diễn Ambisonics thứ tự cao hơn hoặc kênh nào trong số hai hoặc nhiều hơn hai kênh âm thanh từ sự biểu diễn đa kênh đã được tổ hợp.

30. Phương pháp giải mã tín hiệu âm thanh được mã hóa, phương pháp này bao gồm các bước:

nhận tín hiệu âm thanh được mã hóa bao gồm thông tin về sự biểu diễn vận chuyển và thông tin về siêu dữ liệu vận chuyển; và

tổng hợp sự biểu diễn âm thanh không gian sử dụng thông tin về sự biểu diễn vận chuyển và thông tin về siêu dữ liệu vận chuyển,

trong đó siêu dữ liệu vận chuyển biểu thị tín hiệu vận chuyển như đề cập đến phần hoặc bản cầu thứ nhất về vị trí tham chiếu của sự biểu diễn âm thanh không gian và tín hiệu vận chuyển thứ hai như đề cập đến phần hoặc bản cầu khác thứ hai về vị trí tham chiếu của sự biểu diễn âm thanh không gian, trong đó sự tổng hợp bao gồm

tạo ra tín hiệu thành phần của sự biểu diễn âm thanh không gian được kết hợp với phần hoặc bán cầu thứ nhất sử dụng tín hiệu vận chuyển thứ nhất và không sử dụng tín hiệu vận chuyển thứ hai, hoặc

tạo ra tín hiệu thành phần khác của sự biểu diễn âm thanh không gian được kết hợp với phần hoặc bán cầu thứ hai sử dụng tín hiệu vận chuyển thứ hai và không sử dụng tín hiệu vận chuyển thứ nhất, hoặc

tạo ra tín hiệu thành phần khác của sự biểu diễn âm thanh không gian được kết hợp với phần hoặc bán cầu thứ nhất sử dụng tổ hợp của tín hiệu vận chuyển thứ nhất và thứ hai, hoặc

tạo ra tín hiệu thành phần khác của sự biểu diễn âm thanh không gian được kết hợp với phần hoặc bán cầu thứ hai khác sử dụng sự tổ hợp thứ hai gồm các tín hiệu vận chuyển thứ nhất và thứ hai, trong đó sự tổ hợp thứ nhất ảnh hưởng bởi tín hiệu vận chuyển thứ nhất mạnh hơn sự tổ hợp thứ hai, hoặc trong đó sự tổ hợp thứ hai ảnh hưởng bởi tín hiệu vận chuyển thứ hai mạnh hơn sự tổ hợp thứ nhất,

hoặc

trong đó siêu dữ liệu vận chuyển biểu thị thông tin về các hệ số gán trọng số hoặc các hướng nhìn của các tín hiệu vận chuyển của sự biểu diễn vận chuyển, và trong đó sự tổng hợp bao gồm tính toán các thành phần Ambisonics thứ tự thứ nhất khác của sự biểu diễn âm thanh không gian sử dụng thông tin về các hướng nhìn hoặc các hệ số gán trọng số, và sử dụng các tín hiệu vận chuyển.

31. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính bao gồm các lệnh mà, khi chương trình máy tính được thực thi bởi máy tính hoặc bộ xử lý, khiến máy tính hoặc bộ xử lý thực hiện phương pháp theo điểm 29 hoặc 30.

32. Vật ghi lưu trữ số chứa trên đó tín hiệu âm thanh được mã hóa bao gồm:

thông tin về sự biểu diễn vận chuyển của sự biểu diễn âm thanh không gian; và

thông tin về siêu dữ liệu vận chuyển, trong đó siêu dữ liệu vận chuyển biểu thị tín hiệu vận chuyển như đề cập đến phần hoặc bán cầu thứ nhất về vị trí tham chiếu của sự biểu diễn âm thanh không gian và tín hiệu vận chuyển thứ hai như đề cập đến phần hoặc bán cầu khác thứ hai về vị trí tham chiếu của sự biểu diễn âm thanh không gian, trong

đó siêu dữ liệu vận chuyển biểu thị thông tin về các hệ số gắn trọng số hoặc các hướng nhìn của các tín hiệu vận chuyển của sự biểu diễn âm thanh không gian.

1/19

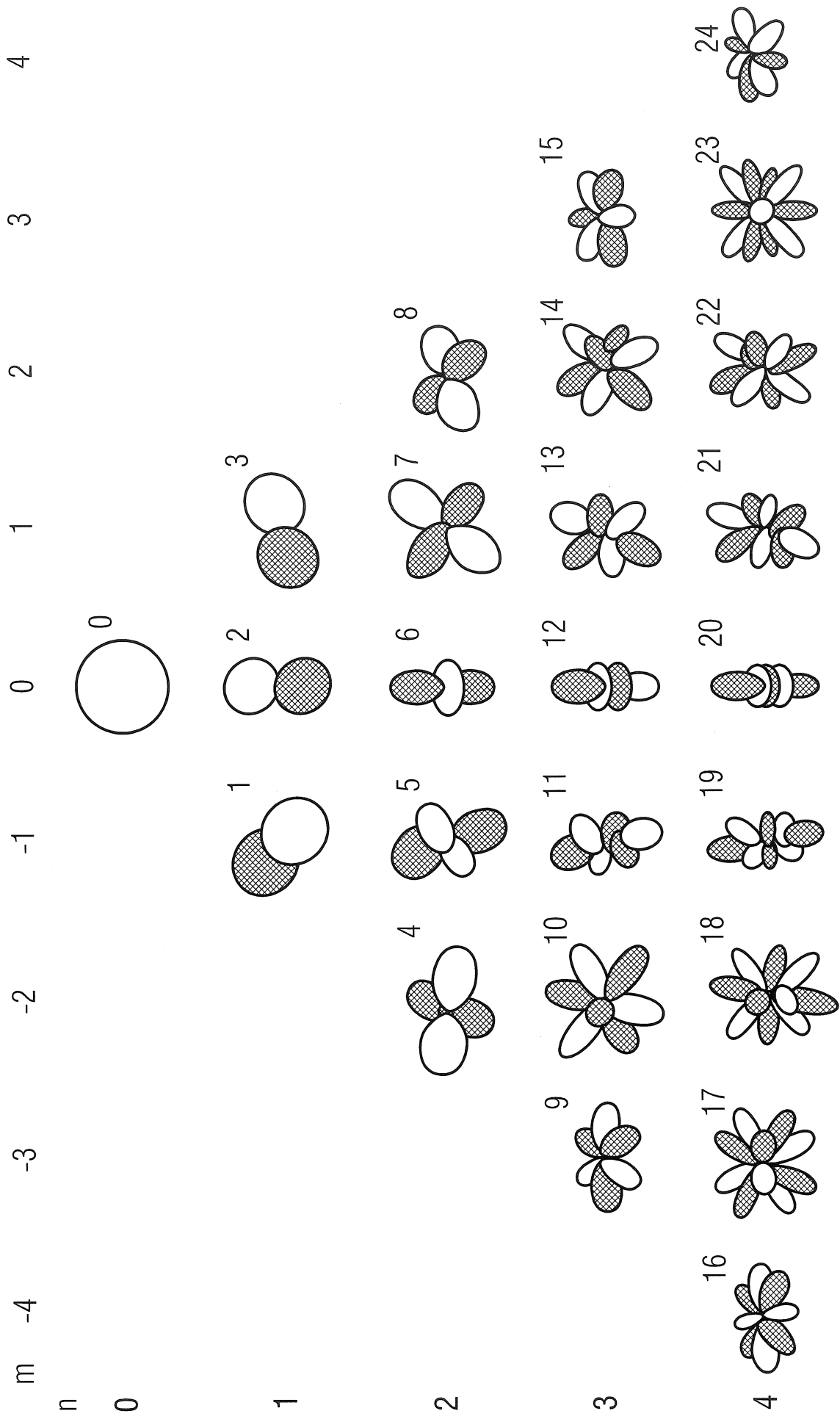


Fig. 1a

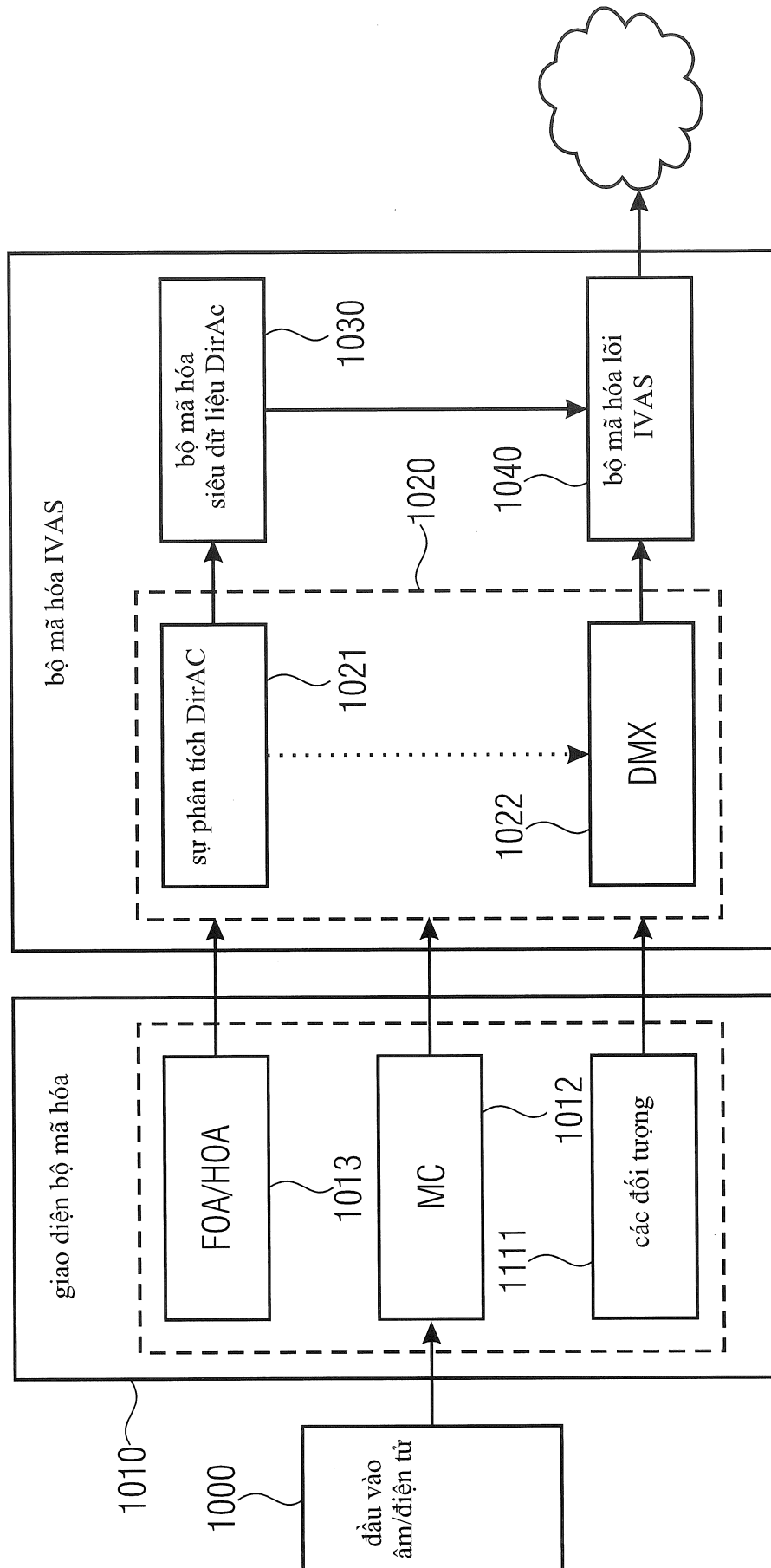


Fig. 1b

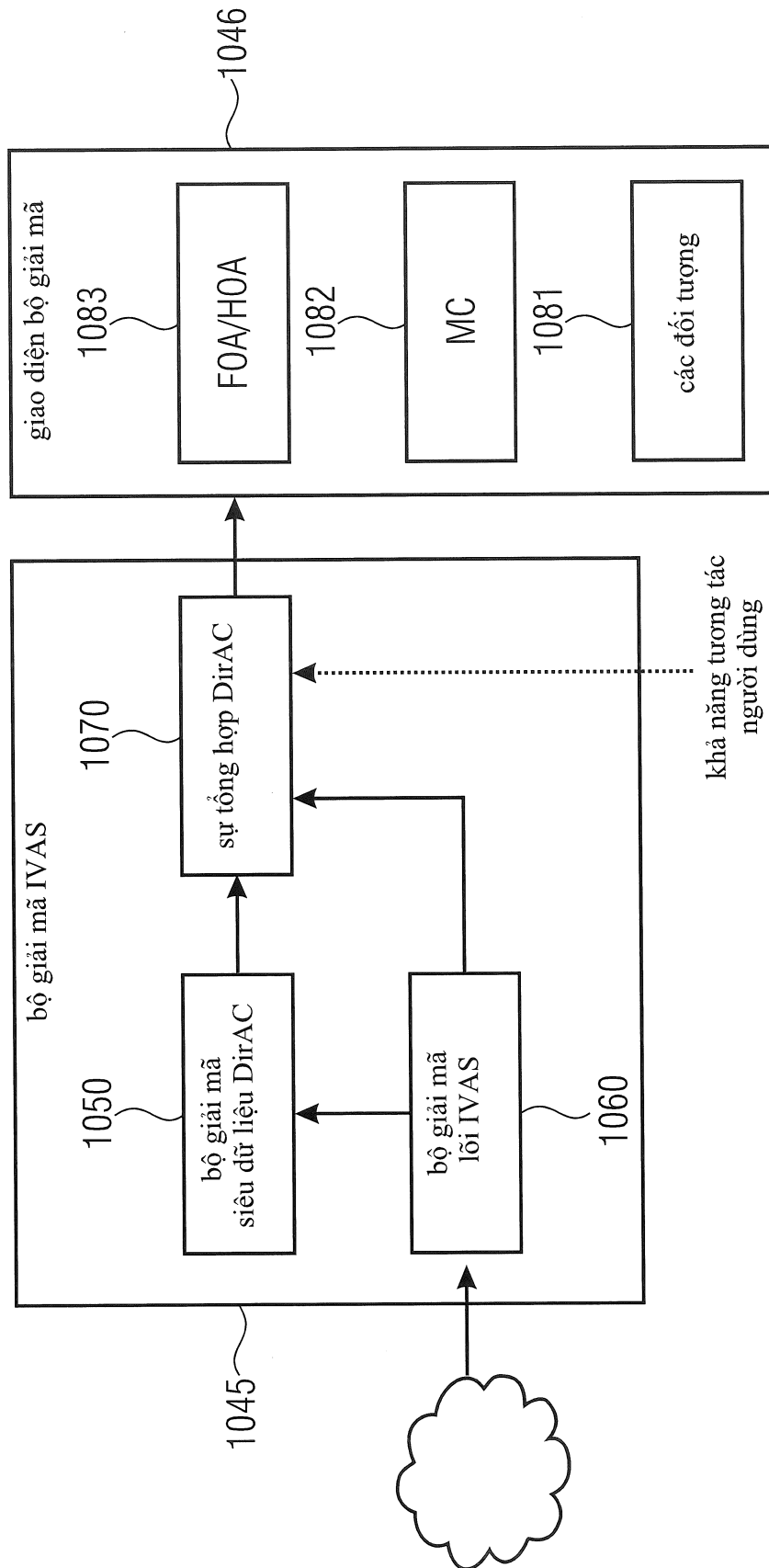


Fig. 2a

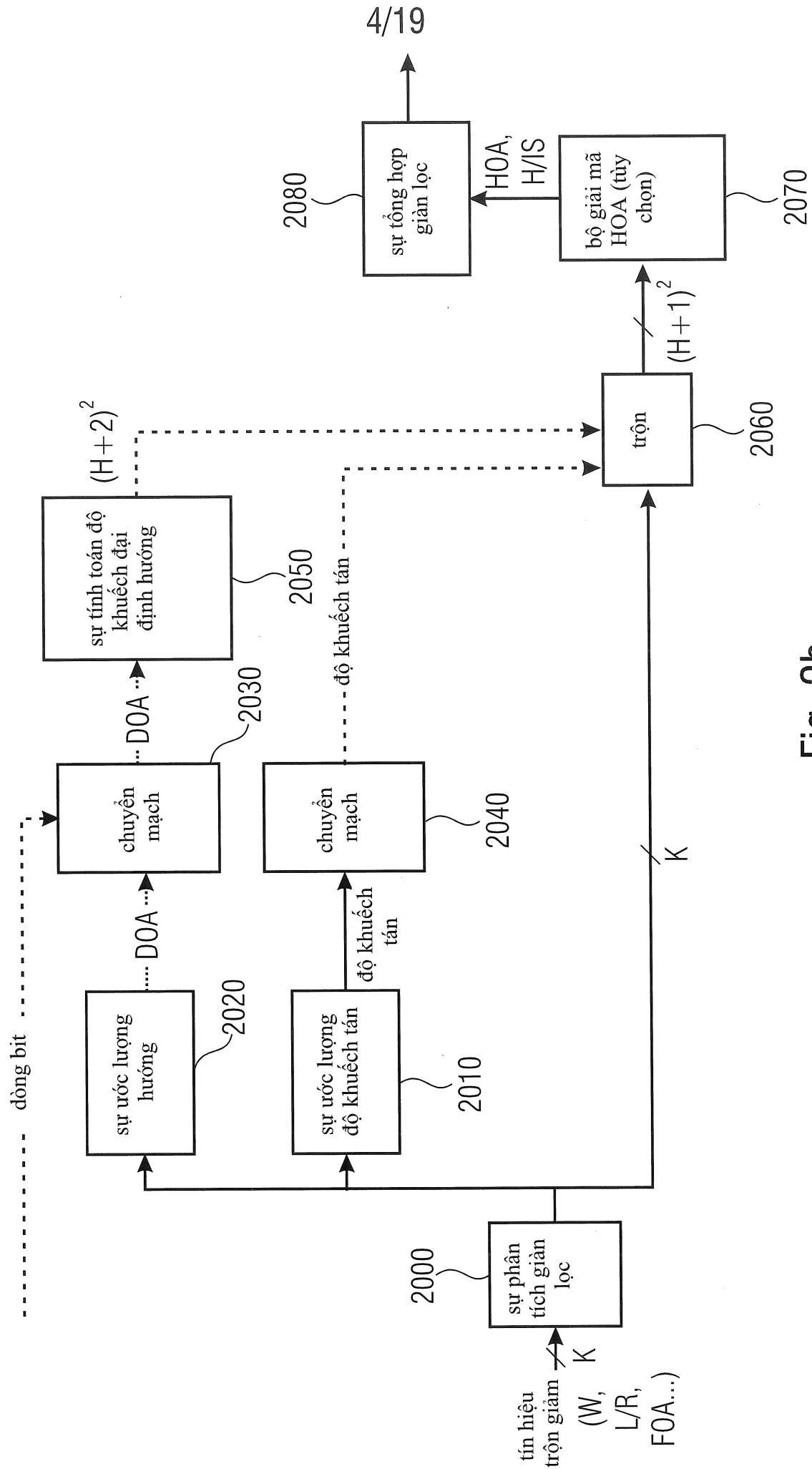


Fig. 2b

5/19

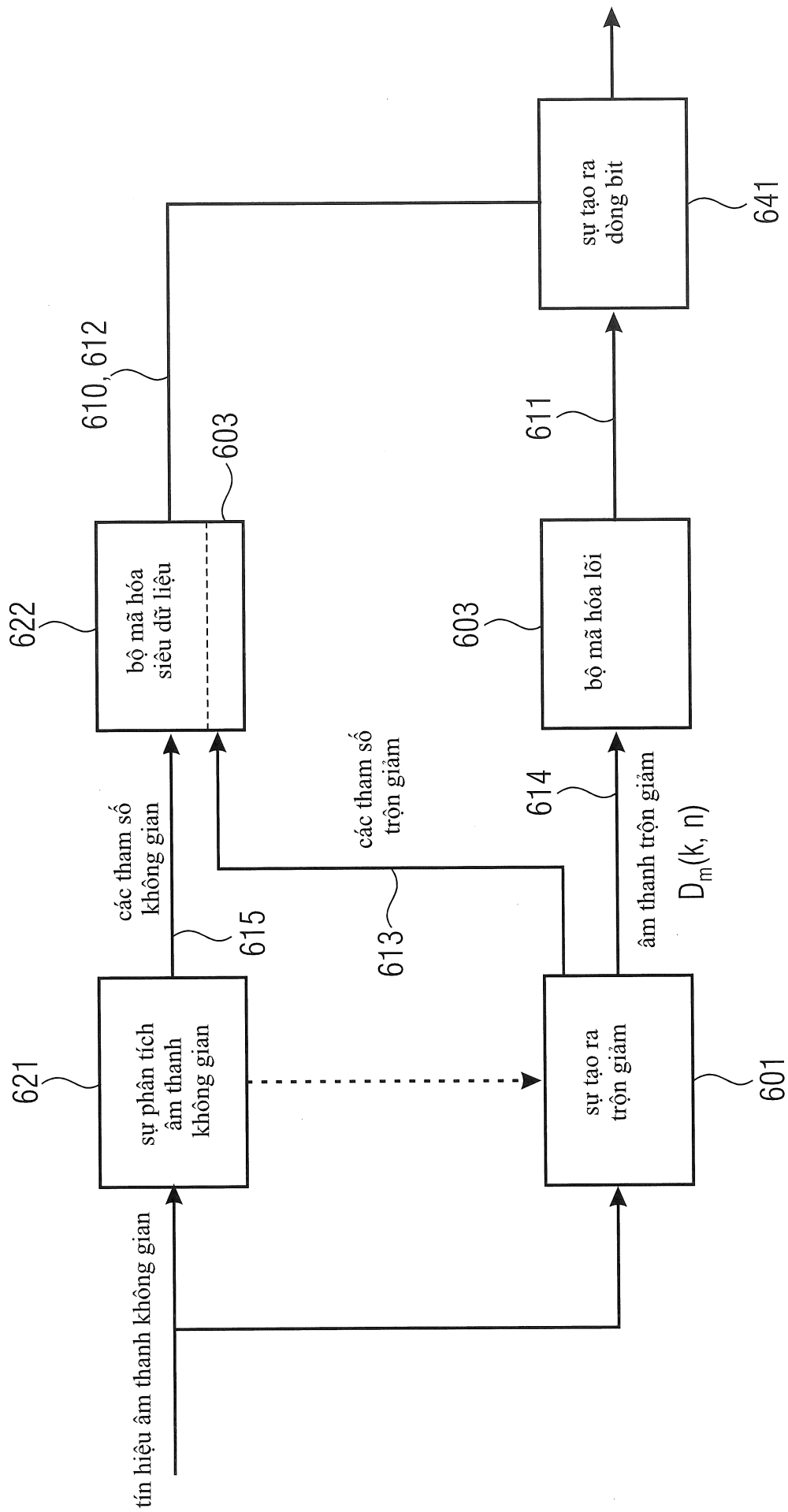


Fig. 3

6/19

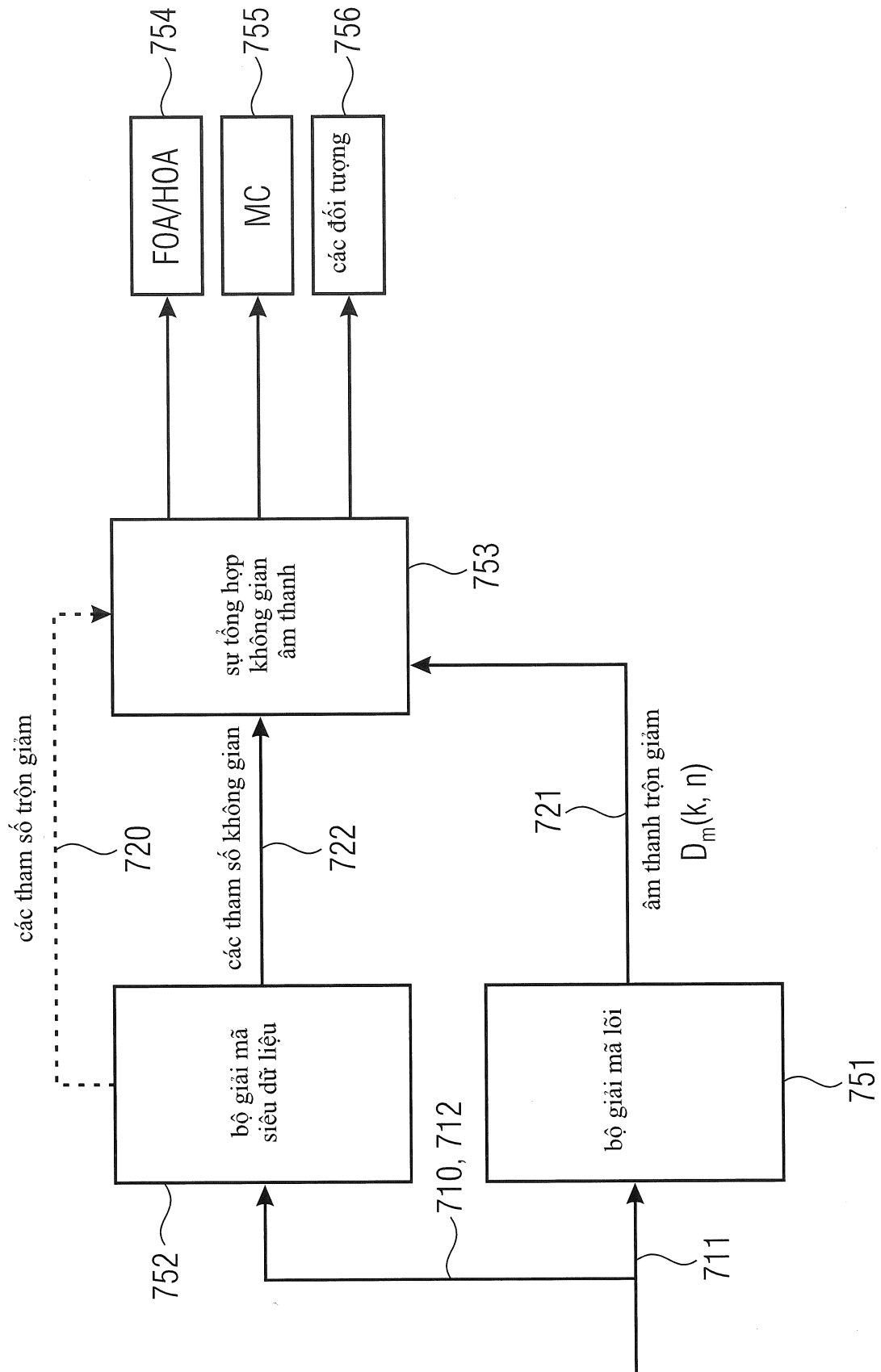


Fig. 4

7/19

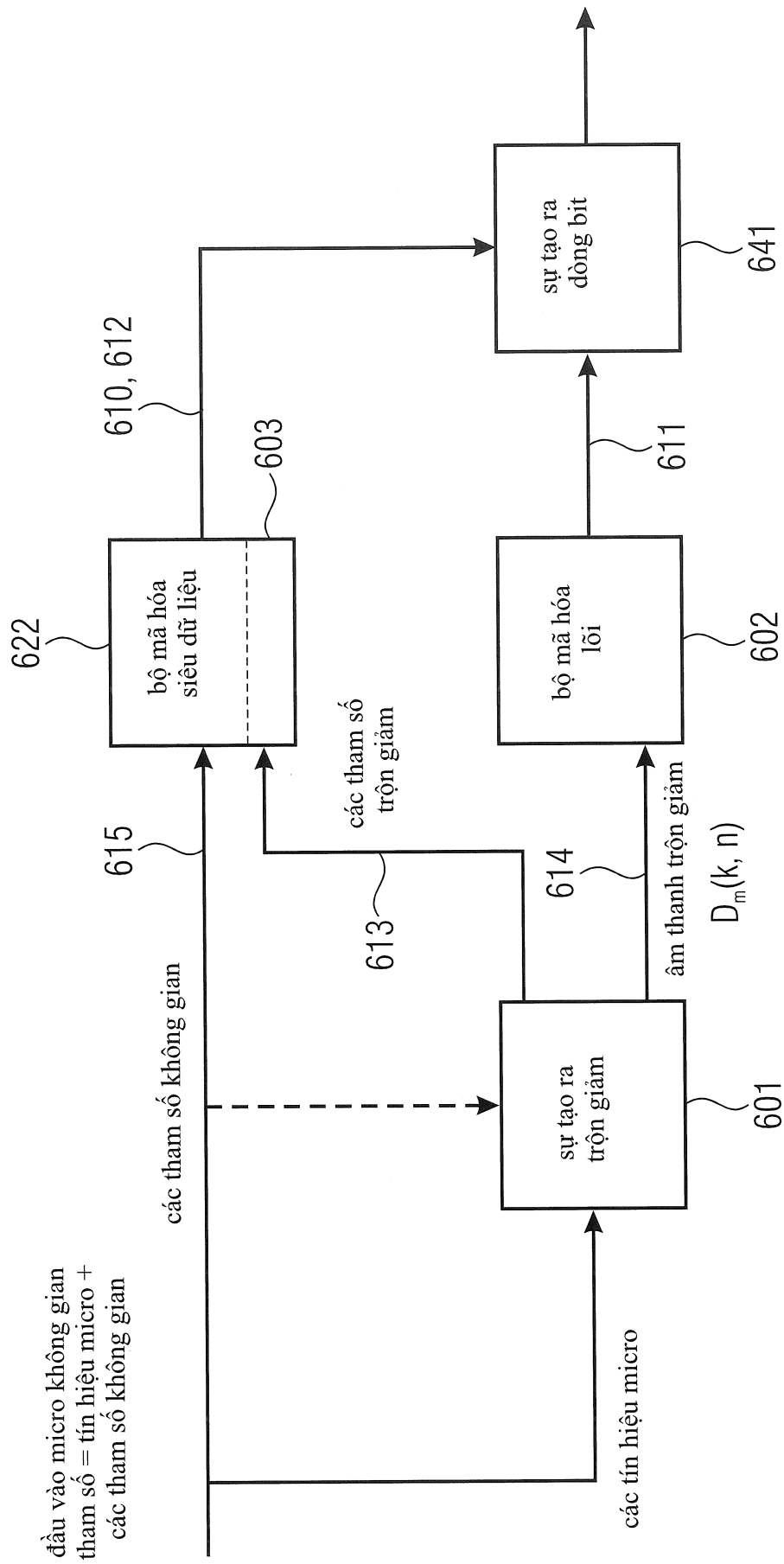


Fig. 5

8/19

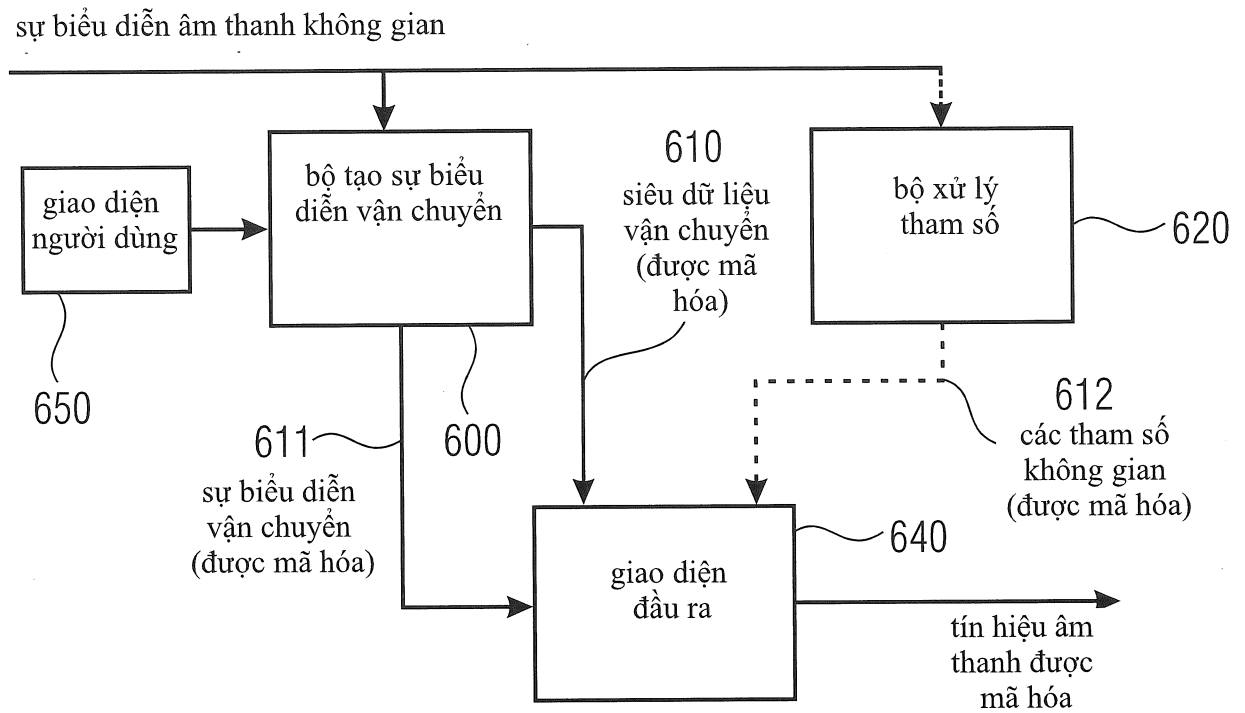


Fig. 6

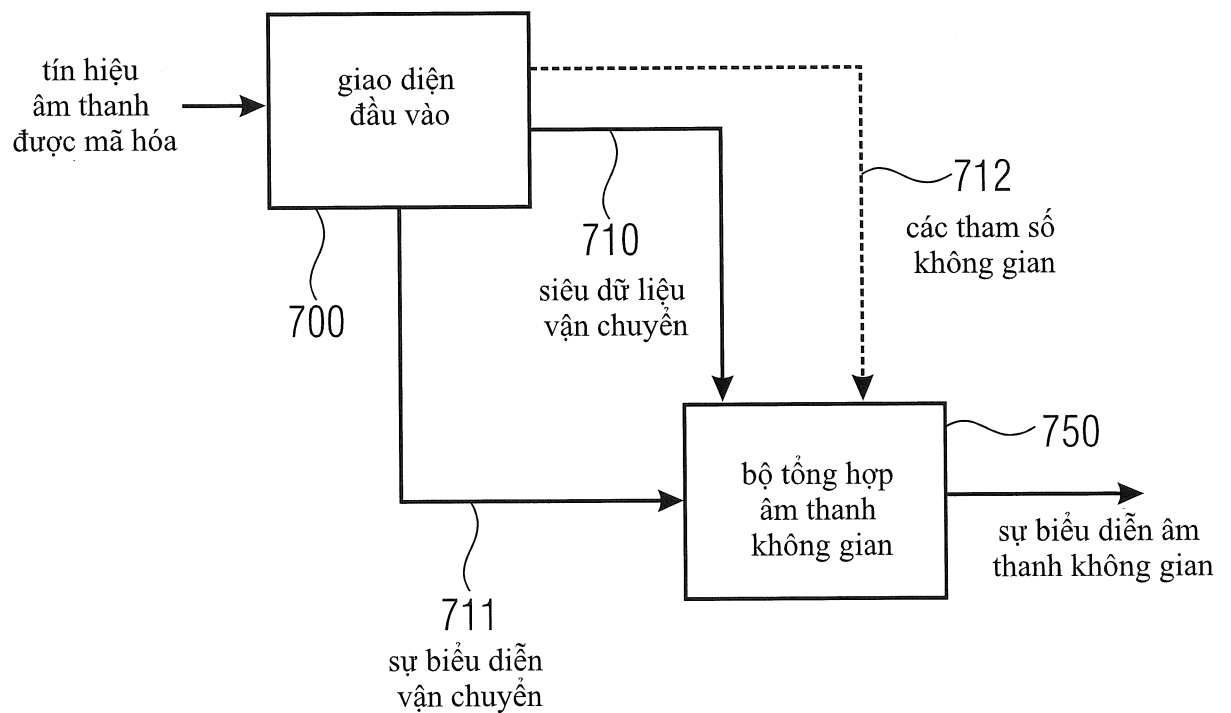


Fig. 7

9/19

ví dụ, FOA/HOA

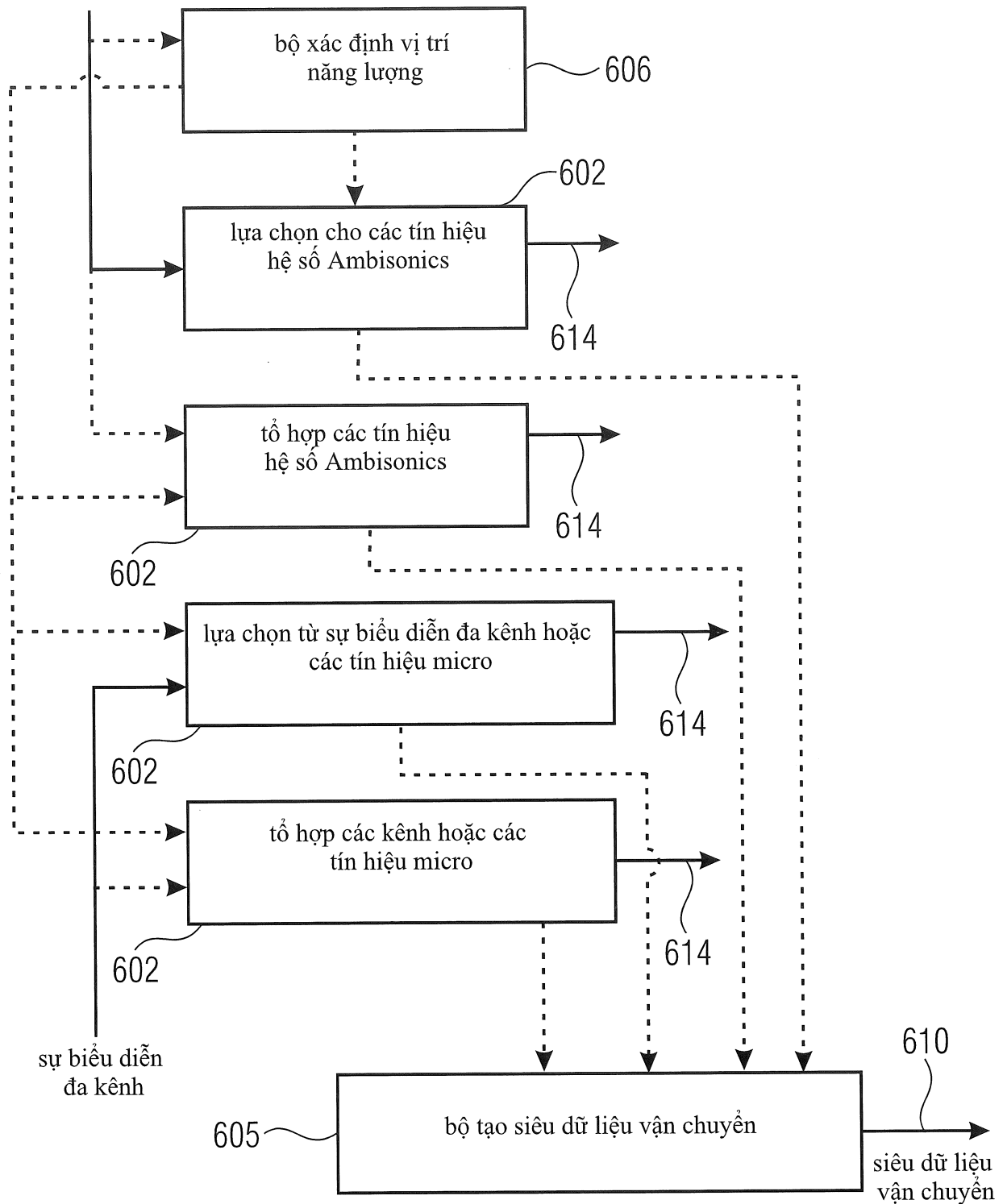


Fig. 8a

10/19

các thay thế cho siêu dữ liệu vận chuyển

thông tin lựa chọn
thông tin tổ hợp
thông tin phân/bán cầu
thông tin hướng nhìn
tham số định hình
các địa điểm, hướng, khoảng cách, mẫu định hướng (các) micro
sự định danh/mô tả mảng micro
(các) hướng nhìn của sự biểu diễn vận chuyển
thông tin về tạo chùm các trọng số/hướng
các tín hiệu micro đẳng hướng hoặc các tín hiệu micro hai hướng

⋮

Fig. 8b

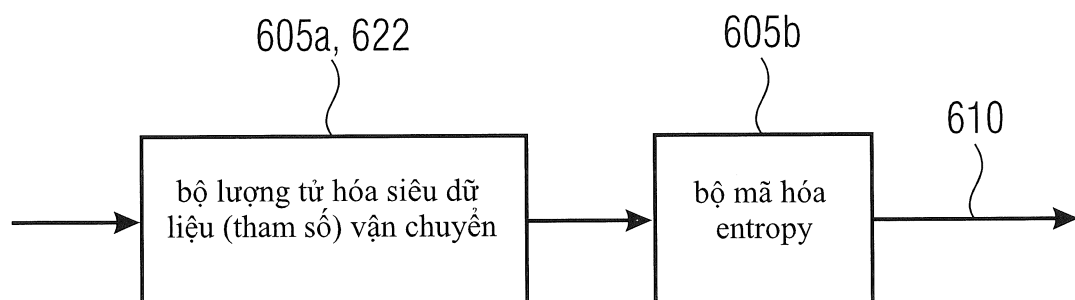


Fig. 8c

11/19

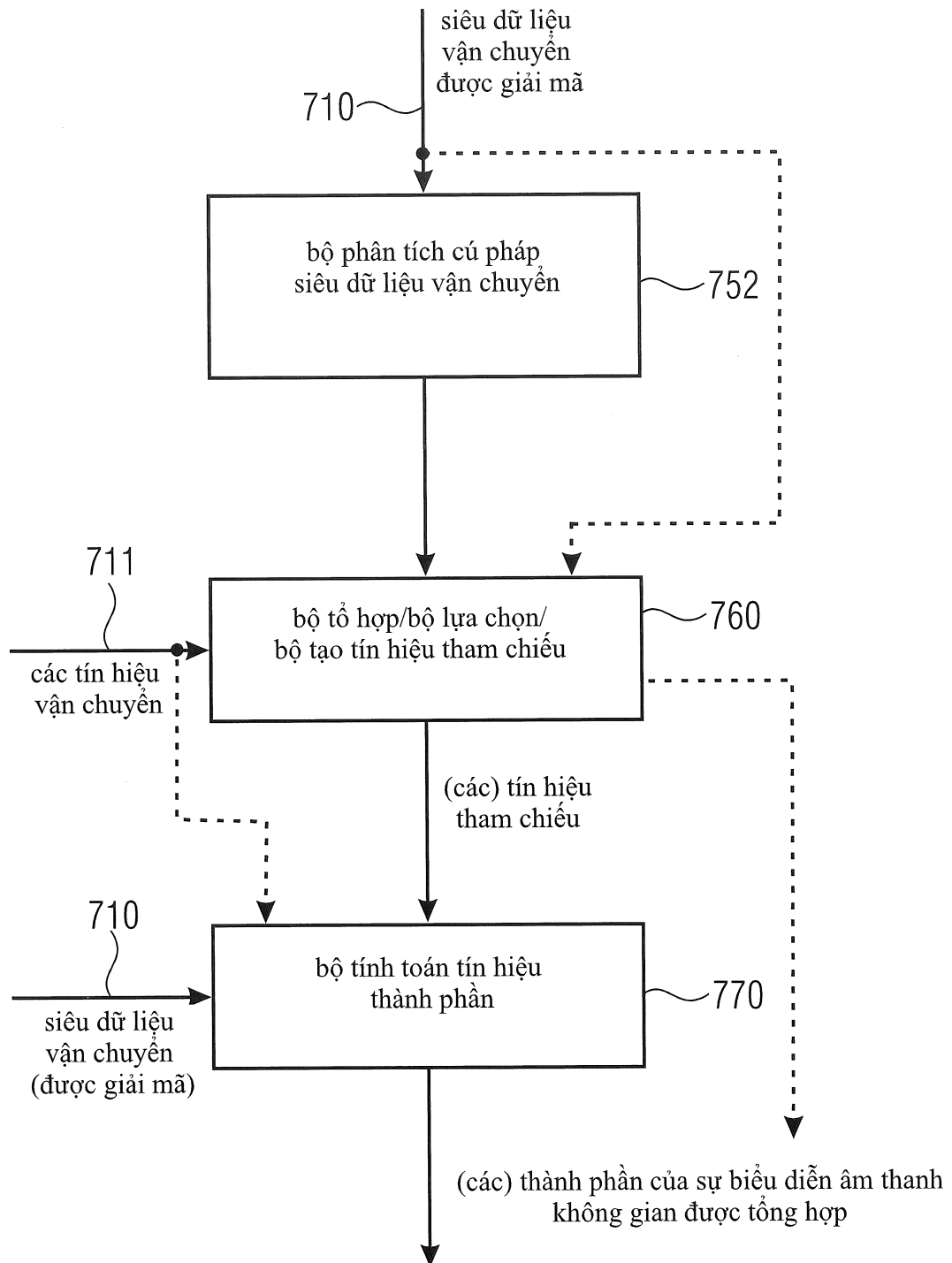


Fig. 9a

12/19

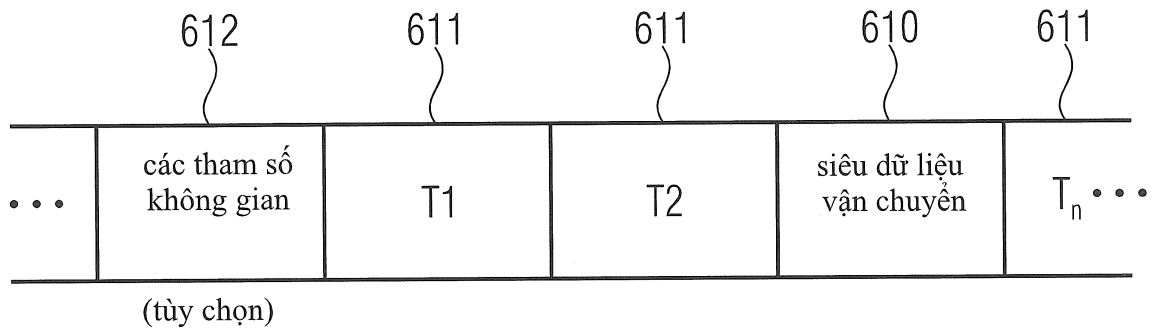


Fig. 9b

loa	bộ lựa chọn/bộ tạo tín hiệu tham chiếu	siêu dữ liệu T
L	T1/T1	T1: bên trái/phía trước (đẳng hướng hoặc hai hướng) T2: bên phải/phía sau (đẳng hướng hoặc hai hướng)
R	T2/T1	
L _s	T1/T2	
R _s	T2/T2	
C	T1+T2/T1	

Fig. 9c

13/19

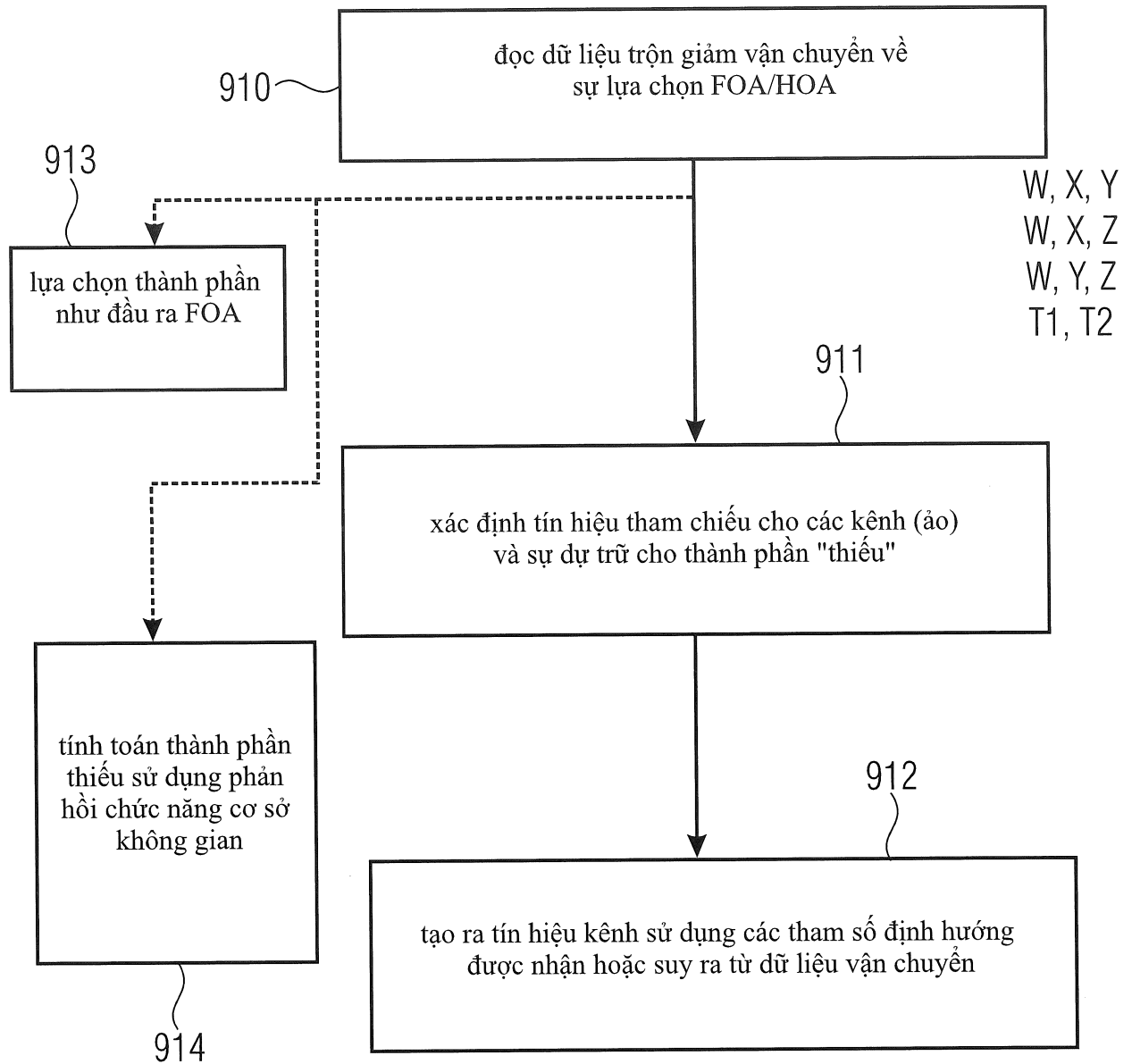


Fig. 9d

14/19

	tham số hình dạng	hướng nhìn
$C_M = 1$	đẳng hướng	trái
		phải
$C_M = 0.5$	hai hướng	trước
		sau
$C_M = 0$	lưỡng cực	trên
		dưới
		(được lượng tử hóa) ϕ, ϑ
		cặp trái/phải
		cặp trước/sau

Fig. 9e

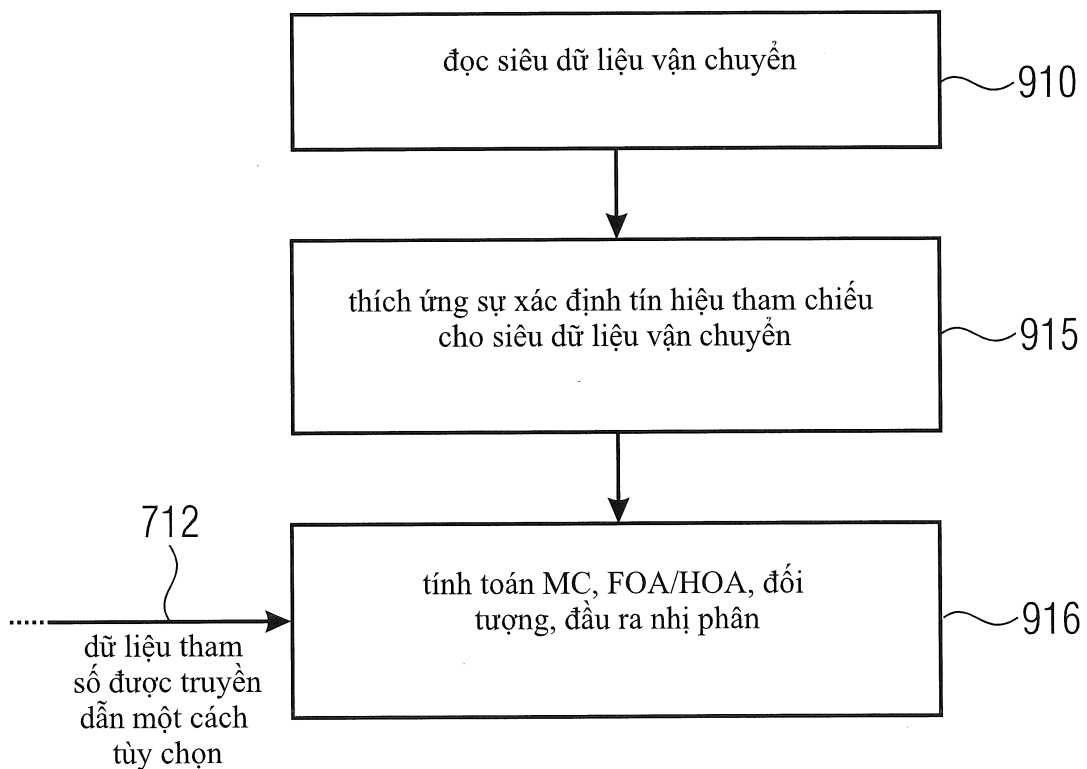


Fig. 9f

15/19

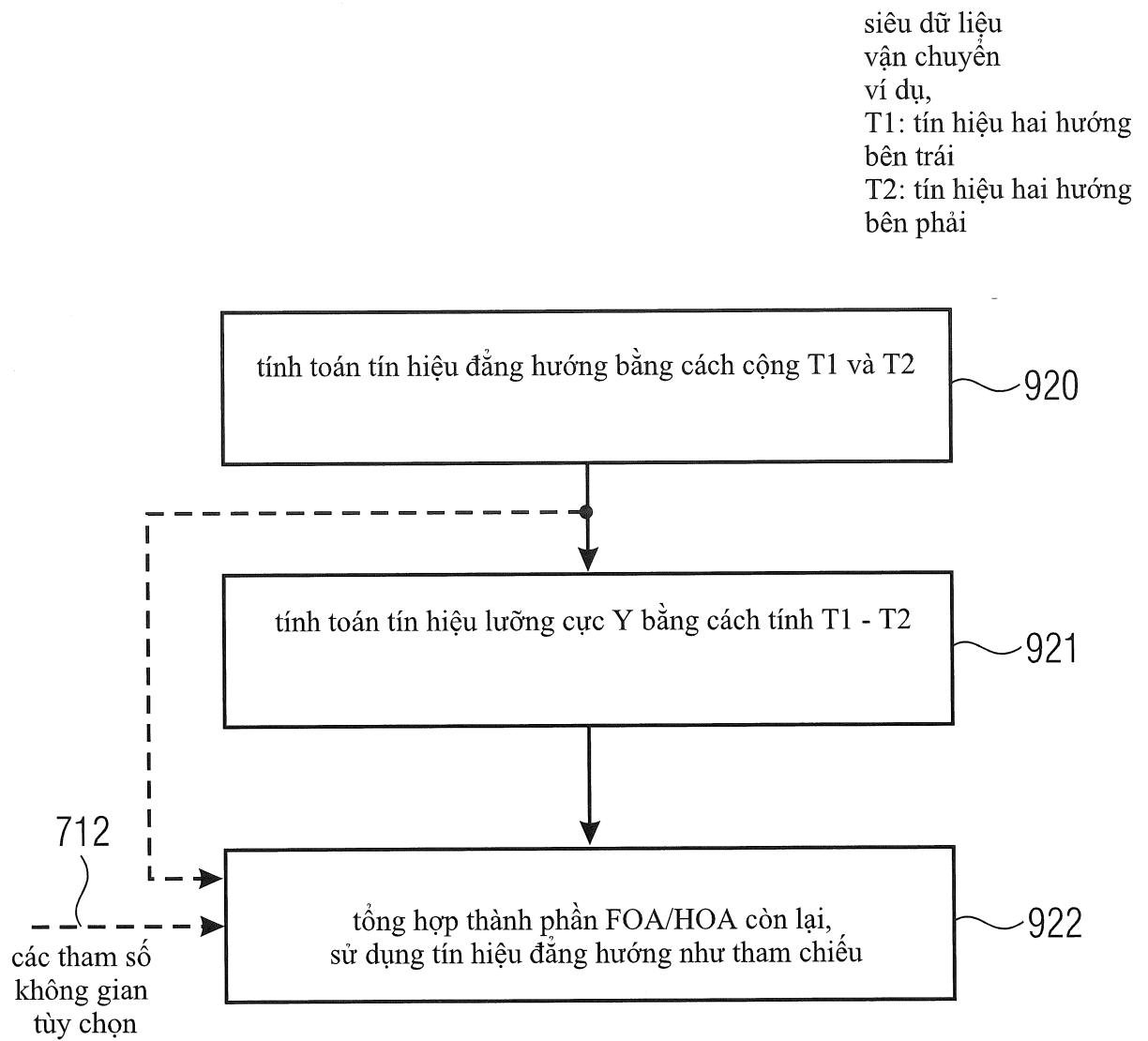


Fig. 9g

16/19

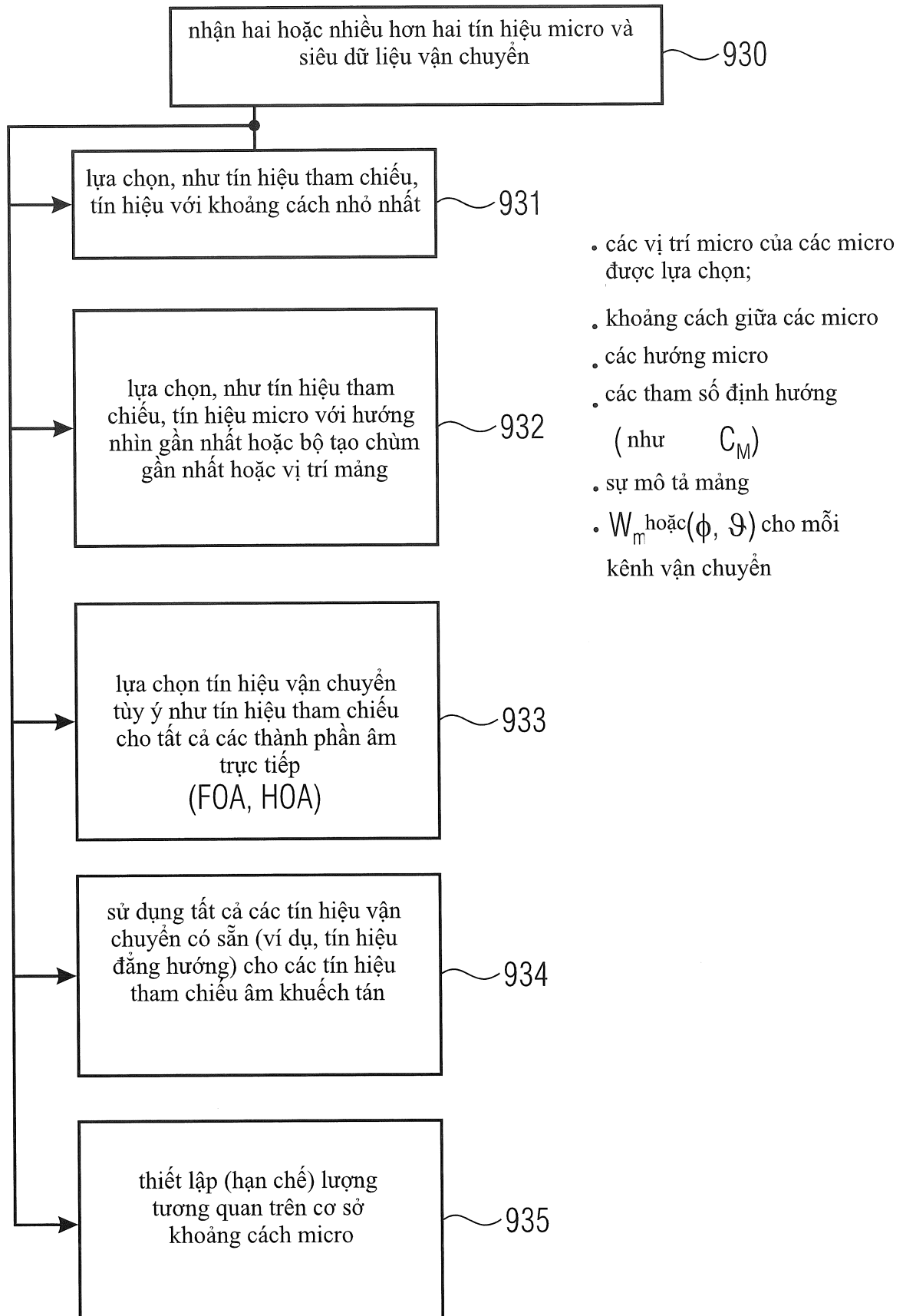


Fig. 9h

17/19

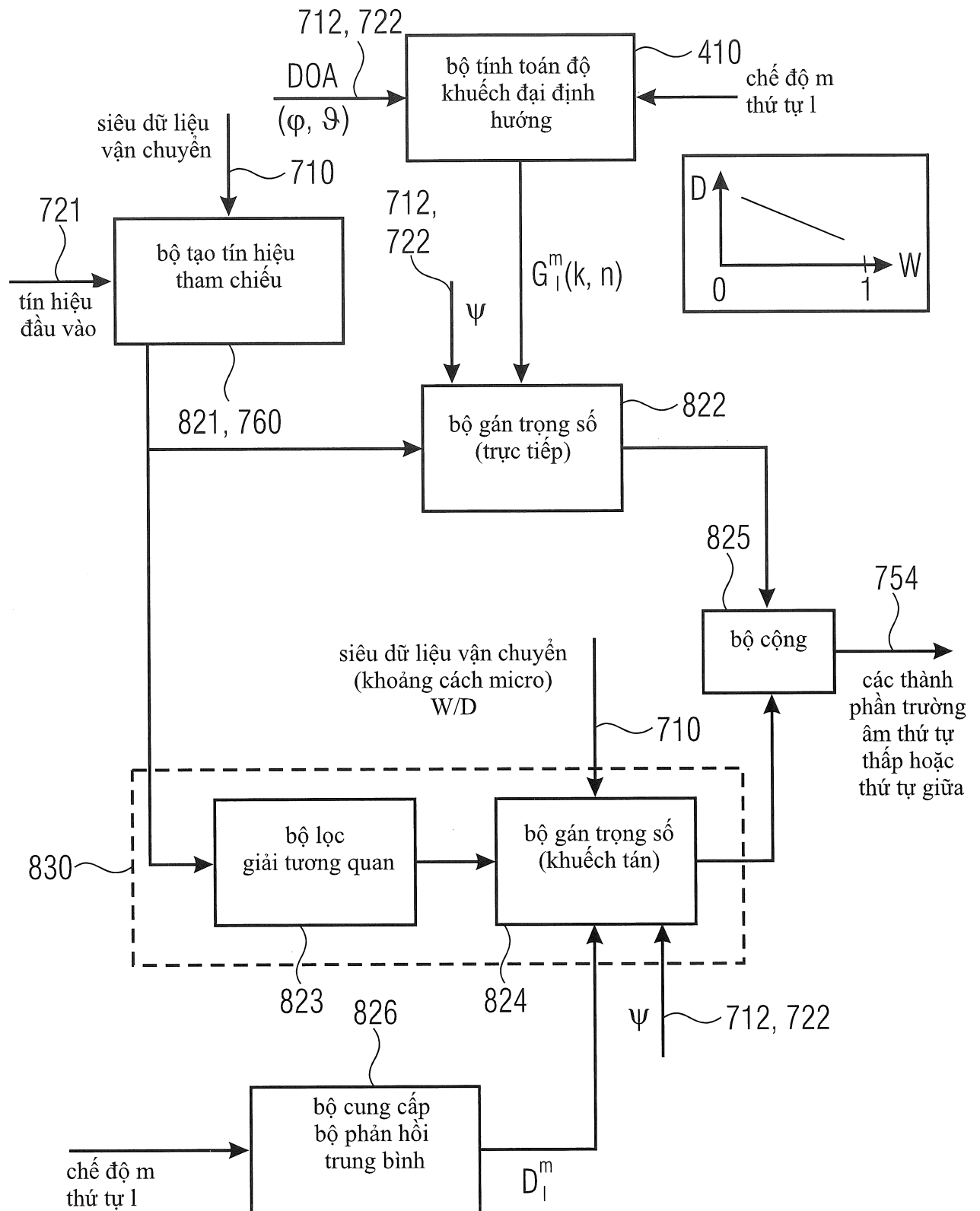


Fig. 10

18/19

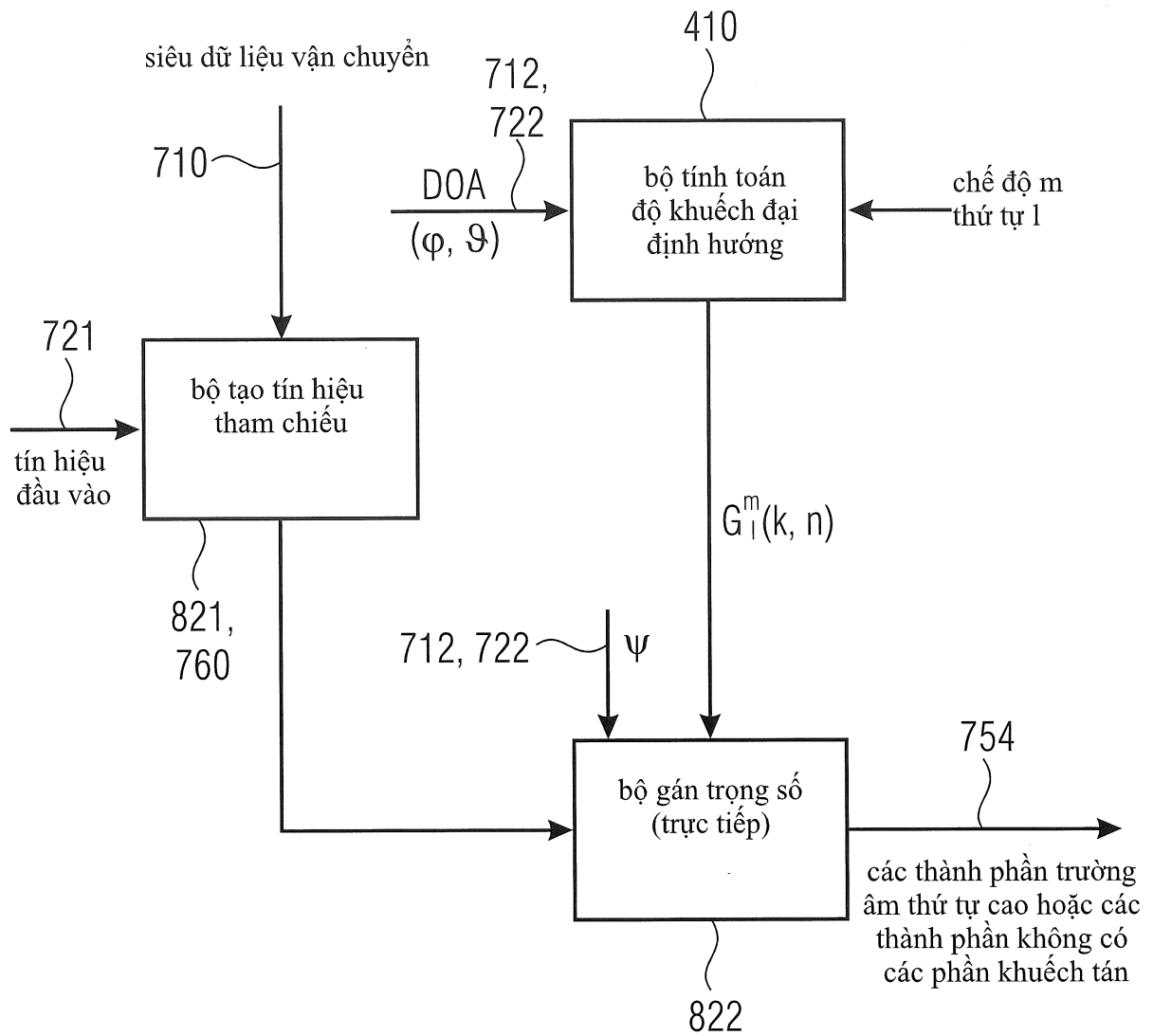


Fig. 11

19/19

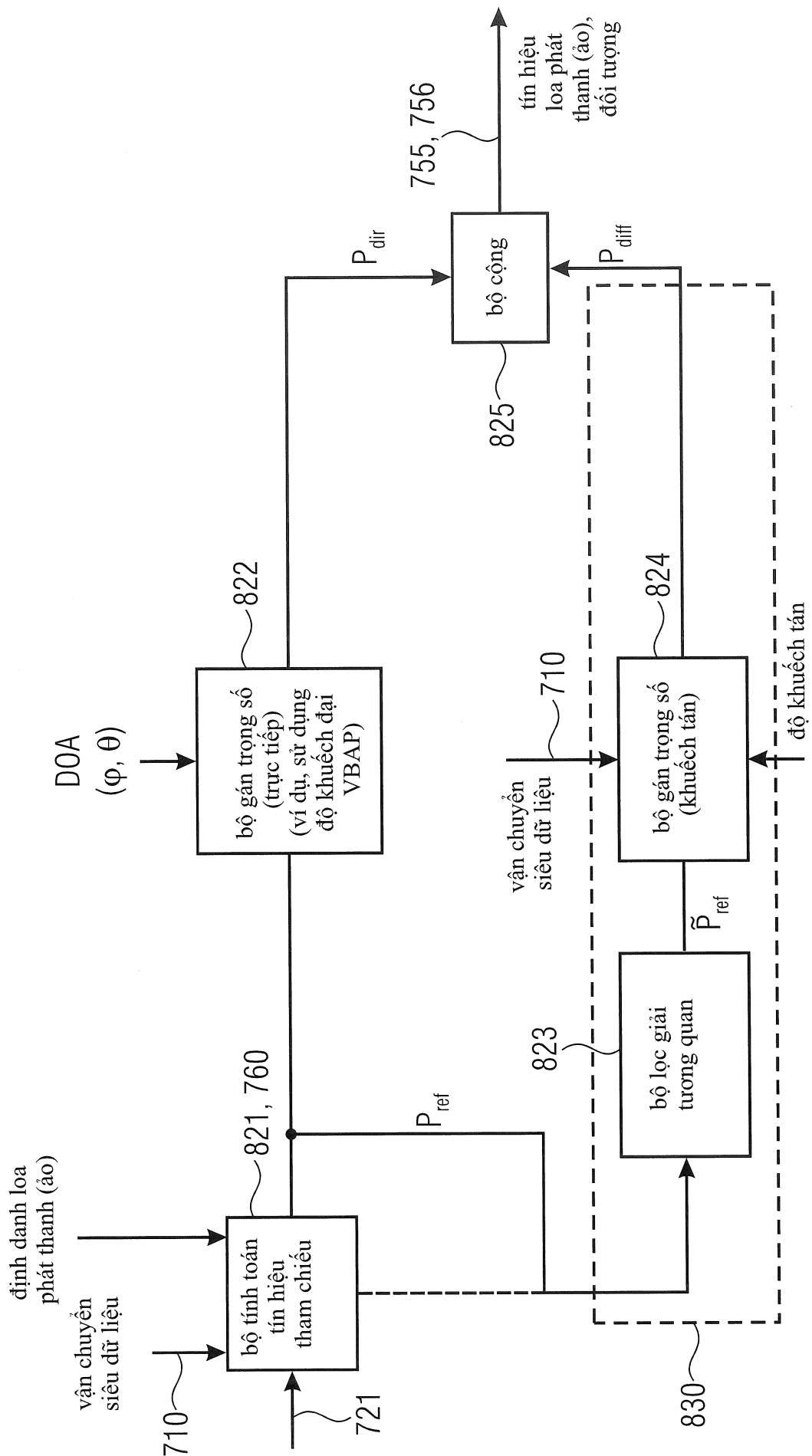


Fig. 12