



(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)



1-0050253

(51)^{2021.01} G10L 19/008

(13) B

(21) 1-2022-03279

(22) 29/10/2020

(86) PCT/US2020/057885 29/10/2020

(87) WO 2021/087063 06/05/2021

(30) 62/927,790 30/10/2019 US; 63/086,465 01/10/2020 US

(45) 25/08/2025 449

(43) 25/07/2022 412A

(73) DOLBY LABORATORIES LICENSING CORPORATION (US)

1275 Market Street, San Francisco, California 94103 (US)

(72) MCGRATH, David S. (AU).

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

(54) PHƯƠNG PHÁP VÀ THIẾT BỊ XỬ LÝ DẠNG BIỂU DIỄN NÉN CỦA TÍN HIỆU
ÂM THANH KHÔNG GIAN, VÀ PHƯƠNG TIỆN LƯU TRỮ ĐỌC ĐƯỢC BẰNG
MÁY TÍNH

(21) 1-2022-03279

(57) Sáng chế đề cập đến các phương pháp xử lý tín hiệu âm thanh không gian để tạo ra dạng biểu diễn nén của tín hiệu âm thanh không gian. Các phương pháp bao gồm bước phân tích tín hiệu âm thanh không gian để xác định các hướng đến cho một hoặc nhiều phần tử âm thanh; đối với ít nhất một dải tần số con, xác định các chỉ báo tương ứng về công suất tín hiệu kết hợp với các hướng đến; tạo ra siêu dữ liệu chứa thông tin hướng mà bao gồm các chỉ báo về các hướng đến của các phần tử âm thanh, và thông tin năng lượng bao gồm các chỉ báo tương ứng về công suất tín hiệu; tạo ra tín hiệu âm thanh dựa trên kênh với số lượng kênh xác định trước dựa vào tín hiệu âm thanh không gian; và xuất ra, dưới dạng dạng biểu diễn nén, tín hiệu âm thanh dựa trên kênh và siêu dữ liệu. Sáng chế còn đề cập đến các phương pháp xử lý dạng biểu diễn nén của tín hiệu âm thanh không gian để tạo ra dạng biểu diễn tái tạo của tín hiệu âm thanh không gian, và đề cập đến thiết bị tương ứng và phương tiện lưu trữ đọc được bằng máy tính

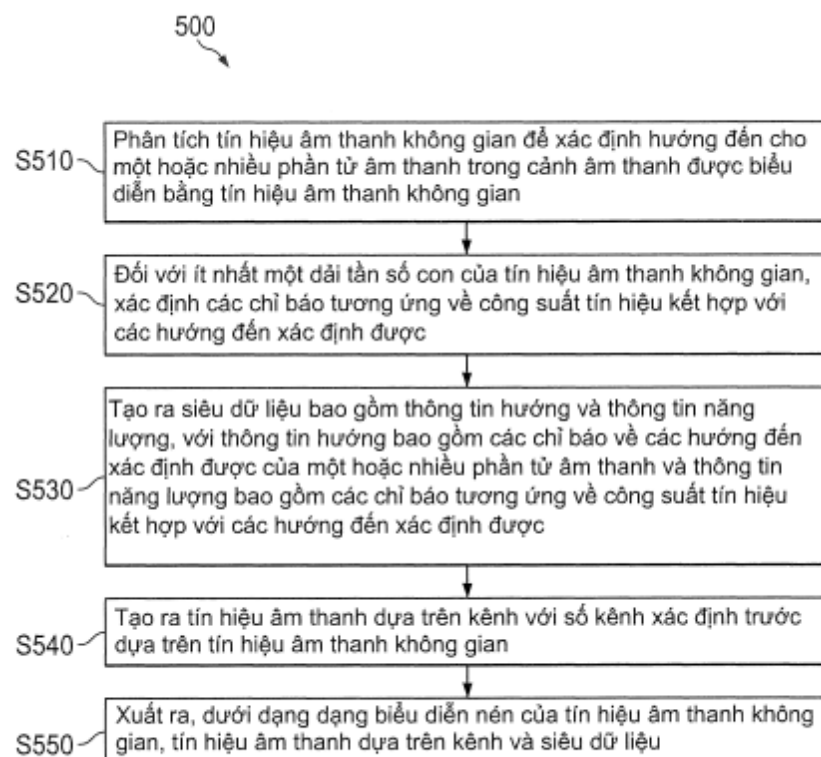


Fig.5

Lĩnh vực kỹ thuật được đề cập

Nói chung sáng chế đề cập đến xử lý tín hiệu âm thanh. Cụ thể, sáng chế đề cập đến các phương pháp xử lý tín hiệu âm thanh không gian (cảnh âm thanh không gian) để tạo ra dạng biểu diễn nén của tín hiệu âm thanh không gian và đến các phương pháp xử lý dạng biểu diễn nén của tín hiệu âm thanh không gian để tạo ra dạng biểu diễn tái tạo của tín hiệu âm thanh không gian.

Tình trạng kỹ thuật của sáng chế

Thính giác của con người cho phép người nghe cảm nhận môi trường của họ dưới dạng cảnh âm thanh không gian - theo đó thuật ngữ “cảnh âm thanh không gian” được dùng ở đây để chỉ môi trường âm thanh xung quanh người nghe, hoặc môi trường âm thanh cảm nhận được trong tâm trí của người nghe.

Trong khi trải nghiệm của con người gắn liền với các cảnh âm thanh không gian, kỹ thuật ghi lại và tái tạo âm thanh liên quan đến việc thu thập, thao tác, truyền và phát lại tín hiệu âm thanh, hoặc các kênh âm thanh. Thuật ngữ “dòng âm thanh” được dùng để chỉ tập hợp một hoặc nhiều tín hiệu âm thanh, đặc biệt trong đó dòng âm thanh được dùng để biểu diễn cảnh âm thanh không gian.

Dòng âm thanh có thể được phát lại cho người nghe, qua bộ chuyển đổi âm thanh điện tử hoặc bằng các phương tiện khác, để cung cấp cho một hoặc nhiều người nghe trải nghiệm nghe dưới dạng cảnh âm thanh không gian. Mục đích chung của những người thực hành ghi âm và nghệ sĩ âm thanh là tạo ra các dòng âm thanh nhằm cung cấp cho người nghe trải nghiệm về cảnh âm thanh không gian cụ thể.

Dòng âm thanh có thể đi kèm với dữ liệu liên quan, được gọi là siêu dữ liệu, hỗ trợ quy trình phát lại. Siêu dữ liệu đi kèm có thể bao gồm thông tin thay đổi theo thời gian có thể được dùng để tác động đến các sửa đổi trong quy trình xử lý được áp dụng trong quá trình phát lại.

Trong phần sau, thuật ngữ “trải nghiệm âm thanh thu được” có thể được dùng để chỉ dòng âm thanh cộng với siêu dữ liệu liên quan bất kỳ.

Trong một số ứng dụng, siêu dữ liệu chỉ bao gồm dữ liệu biểu thị cách bố trí loa dự kiến để phát lại. Thông thường, siêu dữ liệu này bị bỏ qua, với giả định rằng việc bố trí loa phát lại được chuẩn hóa. Trong trường hợp này, trải nghiệm âm thanh thu được chỉ bao gồm dòng âm thanh. Ví dụ về trải nghiệm âm thanh thu được như vậy là dòng âm thanh 2 kênh, được ghi trên đĩa compact, trong đó hệ thống phát lại dự kiến được giả định là ở dạng hai loa được bố trí trước mặt người nghe.

Ngoài ra, trải nghiệm âm thanh thu được dưới dạng tín hiệu âm thanh đa kênh dựa trên cảnh có thể được dùng để trình diễn cho người nghe bằng cách xử lý các tín hiệu âm thanh, qua ma trận trộn, để tạo ra tập hợp các tín hiệu loa, mỗi tín hiệu sau đó có thể được phát lại cho một loa tương ứng, trong đó các loa có thể được bố trí tùy ý theo không gian xung quanh người nghe. Trong ví dụ này, ma trận trộn có thể được tạo ra dựa trên kiến thức trước đó về định dạng dựa trên cảnh và cách bố trí loa phát lại.

Ví dụ về định dạng dựa trên cảnh là Ambisonics bậc cao hơn (Higher Order Ambisonics - HOA) và phương pháp làm ví dụ để tính toán các ma trận trộn phù hợp được đưa ra trong “Ambisonics”, Franz Zotter và Matthias Frank, ISBN: 978-3-030-17206-0, Chương 3, được đưa vào đây bằng cách viện dẫn.

Thông thường, các định dạng dựa trên cảnh như vậy bao gồm một số lượng lớn các kênh hoặc đối tượng âm thanh, dẫn đến các yêu cầu về băng thông hoặc lưu trữ tương đối cao khi truyền hoặc lưu trữ tín hiệu âm thanh không gian theo các định dạng này.

Do đó, cần có các dạng biểu diễn thu gọn của các tín hiệu âm thanh không gian biểu diễn các cảnh âm thanh không gian. Việc này áp dụng cho cả tín hiệu âm thanh không gian dựa trên kênh và dựa trên đối tượng.

Bản chất kỹ thuật của sáng chế

Sáng chế đề xuất các phương pháp xử lý tín hiệu âm thanh không gian để tạo ra dạng biểu diễn nén của tín hiệu âm thanh không gian, các phương pháp xử lý dạng biểu diễn nén của tín hiệu âm thanh không gian để tạo ra dạng biểu diễn tái tạo của tín hiệu

âm thanh không gian, các thiết bị, chương trình và phương tiện lưu trữ đọc được bằng máy tính tương ứng.

Một khía cạnh của sáng chế đề cập đến phương pháp xử lý tín hiệu âm thanh không gian để tạo ra dạng biểu diễn nén của tín hiệu âm thanh không gian. Ví dụ, tín hiệu âm thanh không gian có thể là tín hiệu đa kênh hoặc tín hiệu dựa trên đối tượng. Dạng biểu diễn nén có thể là dạng biểu diễn thu gọn hoặc rút gọn kích thước. Phương pháp có thể bao gồm bước phân tích tín hiệu âm thanh không gian để xác định hướng đến cho một hoặc nhiều phần tử âm thanh trong cảnh âm thanh (cảnh âm thanh không gian) được biểu diễn bằng tín hiệu âm thanh không gian. Các phần tử âm thanh có thể là các phần tử âm thanh trội. Ví dụ, các phần tử âm thanh (trội) có thể liên quan đến các đối tượng âm thanh (trội), nguồn âm thanh (trội) hoặc các thành phần âm thanh (trội) trong cảnh âm thanh. Một hoặc nhiều phần tử âm thanh có thể bao gồm từ một đến mười phần tử âm thanh, như bốn phần tử âm thanh, chẳng hạn. Các hướng đến có thể tương ứng với các vị trí trên hình cầu đơn vị biểu thị vị trí nhận biết được của các phần tử âm thanh. Phương pháp này còn có thể bao gồm bước, đối với ít nhất một dải tần số con (ví dụ, đối với tất cả các dải tần số con) của tín hiệu âm thanh không gian, xác định các chỉ báo tương ứng về công suất tín hiệu kết hợp với các hướng đến xác định được. Phương pháp này còn có thể bao gồm bước tạo ra siêu dữ liệu bao gồm thông tin hướng và thông tin năng lượng, với thông tin hướng bao gồm các chỉ báo về các hướng đến xác định được của một hoặc nhiều phần tử âm thanh và thông tin năng lượng bao gồm các chỉ báo tương ứng về công suất tín hiệu kết hợp với các hướng đến xác định được. Phương pháp còn có thể bao gồm bước tạo ra tín hiệu âm thanh dựa trên kênh với số lượng kênh được xác định trước dựa trên tín hiệu âm thanh không gian. Tín hiệu âm thanh dựa trên kênh có thể được gọi là tín hiệu hỗn hợp âm thanh hoặc dòng hỗn hợp âm thanh. Cần phải hiểu rằng số kênh của tín hiệu âm thanh dựa trên kênh có thể nhỏ hơn số kênh hoặc số đối tượng của tín hiệu âm thanh không gian. Phương pháp này có thể còn bao gồm bước xuất ra, dưới dạng dạng biểu diễn nén của tín hiệu âm thanh không gian, tín hiệu âm thanh dựa trên kênh và siêu dữ liệu. Siêu dữ liệu có thể liên quan đến dòng siêu dữ liệu.

Do đó, dạng biểu diễn nén của tín hiệu âm thanh không gian có thể được tạo ra chỉ bao gồm một số kênh giới hạn. Hơn nữa, bằng cách sử dụng thích hợp thông tin

hướng và thông tin năng lượng, bộ giải mã có thể tạo ra phiên bản tái tạo của tín hiệu âm thanh không gian ban đầu, gần đúng với tín hiệu âm thanh không gian ban đầu khi liên quan đến dạng biểu diễn của cảnh âm thanh không gian ban đầu.

Theo một số phương án, bước phân tích tín hiệu âm thanh không gian có thể được dựa trên nhiều dải tần số con của tín hiệu âm thanh không gian. Ví dụ, bước phân tích có thể được dựa trên toàn bộ khoảng tần số của tín hiệu âm thanh không gian (tức là, toàn bộ tín hiệu). Tức là, bước phân tích có thể được dựa trên tất cả các dải tần số con.

Theo một số phương án, bước phân tích tín hiệu âm thanh không gian có thể bao gồm bước áp dụng phân tích cảnh cho tín hiệu âm thanh không gian. Do đó, (hướng của) các phần tử âm thanh trội trong cảnh âm thanh có thể được xác định một cách tin cậy và hiệu quả.

Theo một số phương án, tín hiệu âm thanh không gian có thể là tín hiệu âm thanh đa kênh. Ngoài ra, tín hiệu âm thanh không gian có thể là tín hiệu âm thanh dựa trên đối tượng. Trong trường hợp này, phương pháp này có thể bao gồm thêm bước chuyển đổi tín hiệu âm thanh dựa trên đối tượng thành tín hiệu âm thanh đa kênh trước khi áp dụng phân tích cảnh. Việc này cho phép áp dụng một cách có ý nghĩa các công cụ phân tích cảnh cho tín hiệu âm thanh.

Theo một số phương án, chỉ báo về công suất tín hiệu kết hợp với hướng đến cho trước có thể liên quan đến một phần công suất tín hiệu trong dải tần số con đối với hướng đến cho trước liên quan đến tổng công suất tín hiệu trong dải tần số con.

Theo một số phương án, các chỉ báo về công suất tín hiệu có thể được xác định cho mỗi trong số nhiều dải tần số con. Trong trường hợp này, chúng có thể liên quan, đối với hướng đến cho trước và dải tần số con cho trước, đến một phần công suất tín hiệu trong dải tần số con cho trước đối với hướng đến cho trước liên quan đến tổng công suất tín hiệu trong dải tần số con cho trước. Đặc biệt, các chỉ báo về công suất tín hiệu có thể được xác định theo cách trên mỗi dải con, trong khi đó việc xác định các hướng đến (trội) có thể được thực hiện trên toàn bộ tín hiệu (tức là, dựa trên tất cả các dải tần số con).

Theo một số phương án, bước phân tích tín hiệu âm thanh không gian, bước xác định các chỉ báo tương ứng về công suất tín hiệu, và bước tạo ra tín hiệu âm thanh dựa

trên kênh có thể được thực hiện trên cơ sở từng đoạn thời gian. Do đó, dạng biểu diễn nén có thể được tạo ra và xuất ra cho mỗi trong số nhiều đoạn thời gian, với tín hiệu âm thanh giảm mức và siêu dữ liệu (khối siêu dữ liệu) cho từng đoạn thời gian. Ngoài ra hoặc cách khác, bước phân tích tín hiệu âm thanh không gian, bước xác định các chỉ báo tương ứng về công suất tín hiệu, và bước tạo ra tín hiệu âm thanh dựa trên kênh có thể được thực hiện dựa trên dạng biểu diễn thời gian-tần số của tín hiệu âm thanh không gian. Ví dụ, các bước nêu trên có thể được thực hiện dựa trên phép biến đổi Fourier rời rạc (như STFT chẳng hạn) của tín hiệu âm thanh không gian. Tức là, đối với mỗi đoạn thời gian (khối thời gian), các bước nêu trên có thể được thực hiện dựa trên các bin thời gian-tần số (các bin FFT) của tín hiệu âm thanh không gian, tức là, trên các hệ số Fourier của tín hiệu âm thanh không gian.

Theo một số phương án, tín hiệu âm thanh không gian có thể là tín hiệu âm thanh dựa trên đối tượng bao gồm nhiều đối tượng âm thanh và vector có hướng liên quan. Sau đó, phương pháp còn có thể bao gồm bước tạo ra tín hiệu âm thanh đa kênh bằng cách điều chỉnh vị trí (panning) các đối tượng âm thanh đến tập hợp các kênh âm thanh xác định trước. Trong đó, mỗi đối tượng âm thanh có thể được điều chỉnh vị trí đến tập hợp các kênh âm thanh xác định trước theo vector có hướng của nó. Hơn nữa, tín hiệu âm thanh dựa trên kênh có thể là tín hiệu giảm mức được tạo ra bằng cách áp dụng phép giảm mức cho tín hiệu âm thanh đa kênh. Ví dụ, tín hiệu âm thanh đa kênh có thể là tín hiệu Ambisonics bậc cao hơn.

Theo một số phương án, tín hiệu âm thanh không gian có thể là tín hiệu âm thanh đa kênh. Khi đó, tín hiệu âm thanh dựa trên kênh có thể là tín hiệu giảm mức được tạo ra bằng cách áp dụng phép giảm mức cho tín hiệu âm thanh đa kênh.

Một khía cạnh khác của sáng chế đề cập đến phương pháp xử lý dạng biểu diễn nén của tín hiệu âm thanh không gian để tạo ra dạng biểu diễn tái tạo lại của tín hiệu âm thanh không gian. Dạng biểu diễn nén có thể bao gồm tín hiệu âm thanh dựa trên kênh với số lượng kênh xác định trước và siêu dữ liệu. Siêu dữ liệu có thể bao gồm thông tin hướng và thông tin năng lượng. Thông tin hướng có thể bao gồm các chỉ báo về các hướng đến của một hoặc nhiều phần tử âm thanh trong một cảnh âm thanh (cảnh âm thanh không gian). Thông tin năng lượng có thể bao gồm, đối với ít nhất một dải tần số con, các chỉ báo tương ứng về công suất tín hiệu kết hợp với hướng đến. Phương pháp

có thể bao gồm bước tạo ra tín hiệu âm thanh của một hoặc nhiều phần tử âm thanh dựa trên tín hiệu âm thanh dựa trên kênh, thông tin hướng và thông tin năng lượng. Phương pháp còn có thể bao gồm bước tạo ra tín hiệu âm thanh dư mà một hoặc nhiều phần tử âm thanh về cơ bản không có ở đó, dựa trên tín hiệu âm thanh dựa trên kênh, thông tin hướng, và thông tin năng lượng. Tín hiệu dư có thể được biểu diễn theo định dạng âm thanh giống như tín hiệu âm thanh dựa trên kênh, ví dụ, có thể có cùng số kênh.

Theo một số phương án, chỉ báo về công suất tín hiệu kết hợp với hướng đến cho trước có thể liên quan đến một phần công suất tín hiệu trong dải tần số con đối với hướng đến cho trước liên quan đến tổng công suất tín hiệu trong dải tần số con.

Theo một số phương án, thông tin năng lượng có thể bao gồm các chỉ báo về công suất tín hiệu cho mỗi trong số nhiều dải tần số con. Sau đó, chỉ báo về công suất tín hiệu có thể liên quan, đối với hướng đến cho trước và dải tần số con cho trước, đến một phần công suất tín hiệu trong dải tần số con cho trước đối với hướng đến cho trước liên quan đến tổng công suất tín hiệu trong dải tần số con cho trước.

Theo một số phương án, phương pháp còn có thể bao gồm bước điều chỉnh vị trí tín hiệu âm thanh của một hoặc nhiều phần tử âm thanh đến tập hợp các kênh của định dạng âm thanh đầu ra. Phương pháp có thể còn bao gồm bước tạo ra tín hiệu âm thanh đa kênh tái tạo theo định dạng âm thanh đầu ra dựa trên một hoặc nhiều phần tử âm thanh được điều chỉnh vị trí và tín hiệu dư. Định dạng âm thanh đầu ra có thể liên quan đến dạng biểu diễn đầu ra, chẳng hạn như HOA hoặc định dạng đa kênh thích hợp khác bất kỳ. Bước tạo ra tín hiệu âm thanh đa kênh tái tạo có thể bao gồm bước tăng mức tín hiệu dư đến tập hợp các kênh của định dạng âm thanh đầu ra. Bước tạo ra tín hiệu âm thanh đa kênh tái tạo có thể còn bao gồm bước thêm một hoặc nhiều phần tử âm thanh đã điều chỉnh vị trí và tín hiệu dư đã tăng mức.

Theo một số phương án, bước tạo ra các tín hiệu âm thanh của một hoặc nhiều phần tử âm thanh có thể bao gồm bước xác định các hệ số của ma trận trộn ngược M để ánh xạ tín hiệu âm thanh dựa trên kênh lên dạng biểu diễn trung gian bao gồm tín hiệu âm thanh dư và các tín hiệu âm thanh của một hoặc nhiều phần tử âm thanh, dựa vào thông tin hướng và thông tin năng lượng. Dạng biểu diễn trung gian cũng có thể được gọi là dạng biểu diễn tách riêng hoặc tách được, hoặc dạng biểu diễn lai.

Theo một số phương án, bước xác định các hệ số của ma trận trộn ngược M có thể bao gồm bước xác định, đối với mỗi trong số một hoặc nhiều phần tử âm thanh, vectơ điều chỉnh vị trí $Pan_{down}(dir)$ để điều chỉnh vị trí phần tử âm thanh đến các kênh của tín hiệu âm thanh dựa trên kênh, dựa vào hướng đến dir của phần tử âm thanh. Bước xác định các hệ số của ma trận trộn ngược M còn có thể bao gồm bước xác định ma trận trộn E mà được dùng để ánh xạ tín hiệu âm thanh dư và các tín hiệu âm thanh của một hoặc nhiều phần tử âm thanh lên các kênh của tín hiệu âm thanh dựa trên kênh, dựa vào các vectơ điều chỉnh vị trí xác định được. Bước xác định các hệ số của ma trận trộn ngược M còn có thể bao gồm bước xác định ma trận hiệp phương sai S của dạng biểu diễn trung gian dựa vào thông tin năng lượng. Bước xác định ma trận hiệp phương sai S có thể còn được dựa vào các vectơ điều chỉnh vị trí xác định được Pan_{down} . Bước xác định các hệ số của ma trận trộn ngược M còn có thể bao gồm thêm bước xác định các hệ số của ma trận trộn ngược M dựa vào ma trận trộn E và ma trận hiệp phương sai S .

Một số phương án, ma trận trộn E có thể được xác định theo $E = (I_N | Pan_{down}(dir_1) | \dots | Pan_{down}(dir_P))$. Ở đây, I_N có thể là $N \times N$ ma trận đơn vị, với N biểu thị số lượng kênh của tín hiệu dựa trên kênh $Pan_{down}(dir_p)$ có thể là vectơ điều chỉnh vị trí cho phần tử âm thanh thứ p có hướng đến liên quan dir_p mà sẽ điều chỉnh vị trí (chẳng hạn, ánh xạ) phần tử âm thanh thứ p đến N kênh của tín hiệu dựa trên kênh với $p = 1, \dots, P$ biểu thị phần tử âm thanh tương ứng trong số một hoặc nhiều phần tử âm thanh và P biểu thị tổng số của một hoặc nhiều phần tử âm thanh. Do đó, ma trận E có thể là ma trận $N \times P$. Ma trận E có thể được xác định cho mỗi trong số nhiều đoạn thời gian k . Trong trường hợp đó, ma trận E và các hướng đến dir_p có chỉ số k biểu thị đoạn thời gian, chẳng hạn, $E_k = (I_N | Pan_{down}(dir_{k,1}) | \dots | Pan_{down}(dir_{k,P}))$. Mặc dù phương pháp được đề xuất có thể hoạt động theo dải, ma trận E có thể giống nhau đối với tất cả các dải tần số con.

Theo một số phương án, ma trận hiệp phương sai S có thể được xác định là ma trận đường chéo theo $\{S\}_{n,n} = \text{rms}(Pan_{down})_n (1 - \sum_{p=1}^P e_p)$ với $1 \leq n \leq N$, và $\{S\}_{N+p,N+p} = e_p$ với $1 \leq p \leq P$. Ở đây, e_p có thể là công suất tín hiệu kết hợp với hướng đến của phần tử âm thanh thứ p . Ma trận S có thể được xác định cho mỗi trong số nhiều đoạn thời gian k , và/hoặc cho mỗi trong số nhiều dải tần số con b . Trong trường

hợp đó, ma trận S và công suất tín hiệu e_p có chỉ số k biểu thị đoạn thời gian và/hoặc chỉ số b biểu thị dải tần số con, chẳng hạn, $\{S_{k,b}\}_{n,n} = \text{rms}(Pan_{down})_n(1 - \sum_{p=1}^P e_{k,p,b})$ với $1 \leq n \leq N$, và $\{S_{k,b}\}_{N+p,N+p} = e_{k,p,b}$ với $1 \leq p \leq P$.

Theo một số phương án, bước xác định các hệ số của ma trận trộn ngược M dựa vào ma trận trộn E và ma trận hiệp phương sai S có thể bao gồm bước xác định giá trị giả ngược dựa vào ma trận trộn E và ma trận hiệp phương sai S .

Theo một số phương án, ma trận trộn ngược M có thể được xác định theo $M = S \times E^* \times (E \times S \times E^*)^{-1}$. Ở đây, “ $\times$ ” biểu thị phép nhân ma trận và “ $*$ ” biểu thị phép chuyển vị liên hợp của ma trận. Ma trận trộn ngược M có thể được xác định cho mỗi trong số nhiều đoạn thời gian k , và/hoặc mỗi trong số nhiều dải tần số con b . Trong trường hợp đó, các ma trận M và S có chỉ số k biểu thị đoạn thời gian và/hoặc chỉ số b biểu thị dải tần số con, và ma trận E có chỉ số k biểu thị đoạn thời gian, chẳng hạn, $M_{k,b} = S_{k,b} \times E_k^* \times (E_k \times S_{k,b} \times E_k^*)^{-1}$.

Theo một số phương án, tín hiệu âm thanh dựa trên kênh có thể là tín hiệu Ambisonics bậc nhất.

Khía cạnh khác đề cập đến thiết bị bao gồm bộ xử lý và bộ nhớ được ghép nối với bộ xử lý, trong đó bộ xử lý được làm thích ứng để thực hiện tất cả các bước của phương pháp theo bất kỳ một trong số các khía cạnh và phương án đề cập ở trên.

Một khía cạnh khác của sáng chế đề cập đến chương trình bao gồm các lệnh mà, khi được thực thi bởi bộ xử lý, khiến cho bộ xử lý thực hiện tất cả các bước của các phương pháp nêu trên.

Tuy nhiên, một khía cạnh khác của sáng chế đề cập đến phương tiện lưu trữ đọc được bằng máy tính lưu trữ chương trình nói trên.

Các phương án khác của sáng chế bao gồm phương pháp hiệu quả để biểu diễn cảnh âm thanh không gian dưới dạng dòng hỗn hợp âm thanh và dòng siêu dữ liệu có hướng, trong đó dòng siêu dữ liệu có hướng bao gồm dữ liệu chỉ báo về vị trí của các phần tử âm thanh có hướng trong cảnh âm thanh không gian và dữ liệu biểu thị công suất của từng phần tử âm thanh có hướng, trong một số dải con, so với tổng công suất của cảnh âm thanh không gian trong dải con đó. Tuy nhiên, các phương án khác liên

quan đến các phương pháp xác định dòng siêu dữ liệu có hướng từ cảnh âm thanh không gian đầu vào, và các phương pháp tạo ra cảnh âm thanh tái tạo từ dòng siêu dữ liệu có hướng và dòng hỗn hợp âm thanh kết hợp.

Theo một số phương án, phương pháp được dùng để biểu diễn cảnh âm thanh không gian ở dạng thu gọn hơn dưới dạng cảnh âm thanh không gian thu gọn bao gồm dòng hỗn hợp âm thanh và dòng siêu dữ liệu có hướng, trong đó dòng hỗn hợp âm thanh bao gồm một hoặc nhiều tín hiệu âm thanh, và trong đó dòng siêu dữ liệu có hướng bao gồm chuỗi thời gian của các khối siêu dữ liệu có hướng với mỗi trong số các khối siêu dữ liệu có hướng được kết hợp với đoạn thời gian tương ứng trong tín hiệu âm thanh, và trong đó cảnh âm thanh không gian bao gồm một hoặc nhiều phần tử âm thanh có hướng mà mỗi phần tử được kết hợp với hướng đến tương ứng, và trong đó mỗi trong số các khối siêu dữ liệu có hướng chứa:

- thông tin hướng biểu thị các hướng đến cho mỗi trong số các phần tử âm thanh có hướng, và
- Thông Tin Phần Dải Năng Lượng biểu thị năng lượng trong mỗi trong số các phần tử âm thanh có hướng, liên quan đến năng lượng trong đoạn thời gian tương ứng trong các tín hiệu âm thanh, cho mỗi trong số các phần tử âm thanh có hướng và cho mỗi dải con trong tập hợp hai hoặc nhiều dải con

Theo một số phương án, phương pháp được dùng để xử lý cảnh âm thanh không gian thu gọn bao gồm dòng hỗn hợp âm thanh và dòng siêu dữ liệu có hướng, để tạo ra dòng âm thanh không gian riêng bao gồm tập hợp một hoặc nhiều tín hiệu đối tượng âm thanh và dòng dư, trong đó dòng hỗn hợp âm thanh bao gồm một hoặc nhiều tín hiệu âm thanh, và trong đó dòng siêu dữ liệu có hướng bao gồm chuỗi thời gian của các khối siêu dữ liệu có hướng với mỗi trong số các khối siêu dữ liệu có hướng được kết hợp với đoạn thời gian tương ứng trong các tín hiệu âm thanh, trong đó đối với mỗi trong số nhiều dải con, phương pháp bao gồm các bước:

- xác định các hệ số của ma trận khử trộn (ma trận trộn ngược) từ thông tin hướng và Thông Tin Phần Dải Năng Lượng có trong dòng siêu dữ liệu có hướng, và
- trộn, bằng cách sử dụng ma trận khử trộn, các tín hiệu âm thanh để tạo ra dòng âm thanh không gian riêng biệt.

Theo một số phương án, phương pháp được dùng để xử lý cảnh âm thanh không gian để tạo ra cảnh âm thanh không gian thu gọn bao gồm dòng hỗn hợp âm thanh và dòng siêu dữ liệu có hướng, trong đó cảnh âm thanh không gian bao gồm một hoặc nhiều phần tử âm thanh có hướng mà mỗi phần tử được kết hợp với hướng đến tương ứng, và trong đó dòng siêu dữ liệu có hướng này bao gồm chuỗi thời gian của các khối siêu dữ liệu có hướng với mỗi trong số các khối siêu dữ liệu có hướng được kết hợp với đoạn thời gian tương ứng trong các tín hiệu âm thanh, phương pháp này bao gồm:

- bước xác định hướng đến cho một hoặc nhiều phần tử âm thanh có hướng, từ việc phân tích cảnh âm thanh không gian,
- bước xác định phần nào của tổng năng lượng trong cảnh không gian được đóng góp bởi năng lượng trong mỗi trong số các phần tử âm thanh có hướng, và
- bước xử lý cảnh âm thanh không gian để tạo ra dòng hỗn hợp âm thanh.

Cần phải hiểu rằng các bước có thể được thực hiện bằng các phương tiện hoặc đơn vị thích hợp, ví dụ như các bước này có thể được thực hiện bởi một hoặc nhiều bộ xử lý máy tính.

Cũng cần hiểu rằng các tính năng của thiết bị và các bước phương pháp có thể được thay thế cho nhau theo nhiều cách. Đặc biệt, người có hiểu biết trung bình trong lĩnh vực kỹ thuật này sẽ hiểu rằng các chi tiết của (các) phương pháp được bộc lộ có thể được thực hiện bởi thiết bị tương ứng và ngược lại. Hơn nữa, bất kỳ nội dung mô tả nào ở trên liên quan đến (các) phương pháp được hiểu là áp dụng tương tự cho thiết bị tương ứng và ngược lại.

Mô tả vắn tắt các hình vẽ

Ví dụ về các phương án của sáng chế được minh họa bằng ví dụ trong các hình vẽ kèm theo, trong đó các số tham chiếu giống nhau biểu thị các phần tử giống nhau hoặc tương tự và trong đó:

Fig.1 minh họa bằng sơ đồ một ví dụ về cách bố trí bộ mã hóa tạo ra dạng biểu diễn nén của cảnh âm thanh không gian và bộ giải mã tương ứng để tạo ra cảnh âm thanh tái tạo từ dạng biểu diễn nén, theo các phương án của sáng chế,

Fig.2 minh họa bằng sơ đồ một ví dụ khác về cách bố trí bộ mã hóa tạo ra dạng biểu diễn nén của cảnh âm thanh không gian và bộ giải mã tương ứng để tạo ra cảnh âm thanh tái tạo từ dạng biểu diễn nén, theo các phương án của sáng chế,

Fig.3 minh họa bằng sơ đồ một ví dụ về việc tạo ra dạng biểu diễn nén của cảnh âm thanh không gian, theo các phương án của sáng chế,

Fig.4 minh họa bằng sơ đồ một ví dụ về việc giải mã dạng biểu diễn nén của cảnh âm thanh không gian để tạo ra cảnh âm thanh tái tạo, theo các phương án của sáng chế,

Fig.5 và Fig.6 là các lưu đồ minh họa các ví dụ về các phương pháp xử lý cảnh âm thanh không gian để tạo ra dạng biểu diễn nén của cảnh âm thanh không gian, theo các phương án của sáng chế,

Các hình vẽ từ Fig.7 đến Fig.11 minh họa bằng sơ đồ các ví dụ về chi tiết của việc tạo ra dạng biểu diễn nén của cảnh âm thanh không gian, theo các phương án của sáng chế,

Fig.12 minh họa bằng sơ đồ một ví dụ về chi tiết của việc giải mã dạng biểu diễn nén của cảnh âm thanh không gian để tạo ra cảnh âm thanh tái tạo, theo các phương án của sáng chế,

Fig.13 là lưu đồ minh họa một ví dụ về phương pháp giải mã dạng biểu diễn nén của cảnh âm thanh không gian để tạo ra cảnh âm thanh tái tạo, theo các phương án của sáng chế,

Fig.14 là lưu đồ minh họa chi tiết của phương pháp trên Fig.13,

Fig.15 là lưu đồ minh họa một ví dụ khác về phương pháp giải mã dạng biểu diễn nén của cảnh âm thanh không gian để tạo ra cảnh âm thanh tái tạo, theo các phương án của sáng chế, và

Fig.16 minh họa bằng sơ đồ thiết bị tạo ra dạng biểu diễn nén của cảnh âm thanh không gian và/hoặc giải mã dạng biểu diễn nén của cảnh âm thanh không gian để tạo ra cảnh âm thanh tái tạo, theo các phương án của sáng chế.

Mô tả chi tiết sáng chế

Nói chung, sáng chế đề cập đến việc cho phép lưu trữ và/hoặc truyền cảnh âm thanh không gian bằng cách sử dụng lượng dữ liệu giảm.

Các khái niệm về xử lý âm thanh có thể được dùng trong ngữ cảnh của sáng chế sẽ được mô tả tiếp theo.

Các hàm điều chỉnh vị trí (panning)

Tín hiệu âm thanh đa kênh (hoặc dòng âm thanh) có thể được tạo ra bằng cách điều chỉnh vị trí các phần tử âm thanh riêng (hoặc các phần tử âm thanh, các đối tượng âm thanh) theo quy luật trộn tuyến tính. Ví dụ, nếu tập hợp R đối tượng âm thanh được biểu diễn bằng R tín hiệu, $\{o_r(t): 1 \leq r \leq R\}$, thì hỗn hợp được điều chỉnh vị trí đa kênh, $\{z_n(t): 1 \leq n \leq N\}$ có thể được tạo ra bằng

$$\begin{pmatrix} z_1(t) \\ z_2(t) \\ \vdots \\ z_N(t) \end{pmatrix} = \sum_{r=1}^R Pan(\theta_r) o_r(t) \quad (1)$$

Hàm điều chỉnh vị trí, $Pan(\theta_r)$, biểu diễn vector cột chứa N hệ số tỷ lệ (độ lợi điều chỉnh vị trí) biểu thị các độ lợi được dùng để trộn tín hiệu đối tượng, $o_r(t)$, để tạo ra đầu ra đa kênh, và trong đó θ_r chỉ báo vị trí của đối tượng tương ứng.

Một hàm điều chỉnh vị trí có thể có là bộ điều chỉnh vị trí Ambisonics bậc nhất (first-order Ambisonic - FOA). Ví dụ về hàm điều chỉnh vị trí FOA được xác định bởi

$$Pan_{FOA}(x, y, z) = \begin{pmatrix} 1 \\ y \\ z \\ x \end{pmatrix} \quad (2)$$

Hàm điều chỉnh vị trí khác là bộ điều chỉnh vị trí Ambisonics bậc ba (third-order Ambisonic - 3OA). Ví dụ về hàm điều chỉnh vị trí 3OA được xác định bởi

$$Pan_{30A}(x, y, z) = \begin{pmatrix} 1 \\ y \\ z \\ x \\ \sqrt{3}xy \\ \sqrt{3}yz \\ \frac{1}{2}(2z^2 - x^2 - y^2) \\ \sqrt{3}xz \\ \frac{\sqrt{3}}{2}(x^2 - y^2) \\ \frac{\sqrt{10}}{4}y(3x^2 - y^2) \\ \sqrt{15}xyz \\ \frac{\sqrt{6}}{4}y(4z^2 - x^2 - y^2) \\ \frac{1}{2}z(2z^2 - 3x^2 - 3y^2) \\ \frac{\sqrt{6}}{4}x(4z^2 - x^2 - y^2) \\ \frac{\sqrt{15}}{2}z(x^2 - y^2) \\ \frac{\sqrt{10}}{4}x(x^2 - 3y^2) \end{pmatrix} \quad (3)$$

Cần phải hiểu rằng sáng chế không giới hạn ở các hàm điều chỉnh vị trí FOA hoặc HOA, và người có hiểu biết trung bình trong lĩnh vực kỹ thuật này sẽ hiểu rằng việc sử dụng các hàm điều chỉnh vị trí khác có thể được xem xét

Biến đổi Fourier ngắn hạn

Ví dụ, dòng âm thanh, bao gồm một hoặc nhiều tín hiệu âm thanh, có thể được chuyển đổi thành dạng biến đổi Fourier ngắn hạn (short-term Fourier transform - STFT). Để đạt được mục đích này, phép biến đổi Fourier rời rạc có thể được áp dụng cho (được tạo cửa sổ tùy ý) các đoạn thời gian của tín hiệu âm thanh (ví dụ, các kênh, tín hiệu đối tượng âm thanh) của dòng âm thanh. Quy trình này, được áp dụng cho tín hiệu âm thanh $x(t)$, có thể được biểu diễn như sau

$$X_{c,k}(f) = STFT\{x_c(t)\} \quad (4)$$

Cần phải hiểu rằng STFT là một ví dụ về biến đổi tần số-thời gian và sáng chế sẽ không giới hạn ở các STFT.

Trong Phương trình (4), biến $X_{c,k}(f)$ biểu thị phép biến đổi Fourier ngắn hạn của kênh c ($1 \leq c \leq NumChans$), cho đoạn thời gian âm thanh k ($k \in \mathbb{Z}$), ở các bin tần số f ($1 \leq f \leq F$), trong đó F biểu thị số lượng bin tần số được tạo ra bởi phép biến đổi Fourier rời rạc. Cần phải hiểu rằng thuật ngữ được sử dụng ở đây là ví dụ, và chi tiết về phương án thực hiện cụ thể của một số phương pháp STFT (bao gồm một số hàm cửa sổ) có thể được biết đến trong lĩnh vực kỹ thuật này. Đoạn thời gian âm thanh k có thể được định nghĩa ví dụ như một loạt các mẫu âm thanh tập trung xung quanh $t = k \times stride + constant$, để đoạn thời gian được đặt cách đều theo thời gian, với khoảng cách bằng $stride$.

Các giá trị số của STFT (như $X_{c,k}(1)$, $X_{c,k}(2)$, ..., $X_{c,k}(F)$) có thể được gọi là các bin FFT.

Hơn nữa, dạng STFT có thể được chuyển đổi thành dòng âm thanh. Dòng âm thanh thu được có thể gần đúng với đầu vào ban đầu và có thể được xác định bởi

$$\begin{aligned} x'_c(t) &= STFT^{-1}\{X_{c,k}(f)\} \\ &\approx x_c(t) \end{aligned} \quad (5)$$

Phân tích tạo dải tần số

Dữ liệu đặc trưng có thể được tạo ra từ dòng âm thanh trong đó dữ liệu đặc trưng được kết hợp với một số dải tần số (dải tần số con), trong đó dải (dải con) được xác định bởi một vùng của khoảng tần số.

Ví dụ, công suất tín hiệu trong kênh c của dòng, trong dải tần số b (trong đó số dải là B và $1 \leq b \leq B$), trong đó dải b trải trên các bin FFT $f_{min} \leq f \leq f_{max}$, có thể được tính toán theo

$$power_{c,b,k} = \sum_{f=f_{min}}^{f_{max}} |X_{c,k}(f)|^2 \quad (6)$$

Theo ví dụ chung hơn, dải tần số b có thể được xác định bởi vector trọng số, $FR_b(f)$, mà gán các trọng số cho mỗi bin tần số, để quy trình tính toán khác của công suất trong dải được xác định bằng

$$power_{c,b,k} = \sum_{f=1}^F FR_b(f) |X_{c,k}(f)|^2 \quad (7)$$

Tổng quát hơn về phương trình (7), STFT của dòng bao gồm C tín hiệu âm thanh có thể được xử lý để tạo ra hiệp phương sai trong một số dải, trong đó hiệp phương sai, $R_{b,k}$ là ma trận $C \times C$, và trong đó phần tử $\{R_{b,k}\}_{i,j}$ được tính theo,

$$\{R_{b,k}\}_{i,j} = \sum_{f=1}^F FR_b(f) X_{i,k}(f) \overline{X_{j,k}(f)} \quad (8)$$

trong đó $\overline{X_{j,k}(f)}$ biểu diễn liên hợp phức của $X_{j,k}(f)$.

Trong một ví dụ khác, bộ lọc thông dải có thể được sử dụng để tạo ra các tín hiệu được lọc biểu diễn dòng âm thanh gốc trong các dải tần số theo các đáp ứng của bộ lọc thông dải. Ví dụ, tín hiệu âm thanh $x_c(t)$ có thể được lọc để tạo ra $x'_{c,b}(t)$, biểu diễn tín hiệu có năng lượng chủ yếu bắt nguồn từ dải b của $x_c(t)$, và do đó phương pháp khác để tính toán hiệp phương sai của dòng trong dải b cho khối thời gian k (tương ứng với các mẫu thời gian $t_{min} \leq t \leq t_{max}$) có thể được biểu diễn bằng

$$\{R_{b,k}\}_{i,j} = \sum_{t=t_{min}}^{t_{max}} x'_{i,b}(t) x'_{j,b}(t) \quad (9)$$

Trộn tạo dải tần số

Dòng âm thanh bao gồm N kênh có thể được xử lý để tạo ra dòng âm thanh bao gồm M kênh theo ma trận trộn tuyến tính $M \times N$, Q , sao cho

$$y_m(t) = \sum_{n=1}^N Q_{m,n} x_n(t) \quad (10)$$

có thể được viết dưới dạng ma trận là

$$\hat{y}(t) = Q \times \hat{x}(t) \quad (11)$$

trong đó $\hat{x}(t)$ được dùng để chỉ vector cột được tạo ra từ N phần tử: $x_1(t), x_2(t), \dots, x_N(t)$.

Hơn nữa, quy trình trộn khác có thể được thực hiện trong miền STFT, trong đó ma trận, Q , có thể nhận các giá trị khác nhau trong mỗi khối thời gian, k , và trong mỗi khối tần số, b . Trong trường hợp này, quy trình xử lý có thể được coi là được xác định bởi

$$Y_{m,k}(f) = \sum_{b=1}^B \sum_{n=1}^N FR_b(f) Q_{b,k,m,n} X_{n,k}(f) \quad (12)$$

hoặc, ở dạng ma trận

$$\hat{Y}_k(f) = \sum_{b=1}^B FR_b(f) (Q_{b,k} \times \hat{X}_k(f)) \quad (13)$$

Cần phải hiểu rằng các phương pháp thay thế có thể được dùng để tạo ra hoạt động tương đương với quá trình xử lý được mô tả trong phương trình (13).

Các phương án thực hiện làm ví dụ

Tiếp theo, các phương án thực hiện làm ví dụ của các phương pháp và thiết bị theo các phương án của sáng chế sẽ được mô tả chi tiết hơn.

Nói chung, các phương pháp theo các phương án của sáng chế đề cập đến cảnh âm thanh không gian dưới dạng dòng hỗn hợp âm thanh và dòng siêu dữ liệu có hướng, trong đó dòng siêu dữ liệu có hướng bao gồm dữ liệu biểu thị vị trí của các phần tử âm thanh có hướng trong cảnh âm thanh không gian và dữ liệu biểu thị công suất của từng phần tử âm thanh có hướng, trong một số dải con, so với tổng công suất của cảnh âm thanh không gian trong dải con đó. Các phương pháp khác theo các phương án của sáng chế đề cập đến việc xác định dòng siêu dữ liệu có hướng từ cảnh âm thanh không gian đầu vào, và để tạo ra cảnh âm thanh tái tạo (ví dụ, tái cấu trúc) từ dòng siêu dữ liệu có hướng và dòng hỗn hợp âm thanh kết hợp.

Ví dụ về các phương pháp theo các phương án của sáng chế là hiệu quả (ví dụ, về mặt giảm dữ liệu để lưu trữ hoặc truyền) trong quy trình biểu diễn cảnh âm thanh không gian. Cảnh âm thanh không gian có thể được biểu diễn bằng tín hiệu âm thanh không gian. Các phương pháp đã nêu có thể được thực hiện bằng cách xác định định dạng lưu trữ hoặc truyền (ví dụ, dòng âm thanh không gian thu gọn) bao gồm dòng hỗn hợp âm thanh và dòng siêu dữ liệu (ví dụ, dòng siêu dữ liệu có hướng).

Dòng hỗn hợp âm thanh bao gồm một số tín hiệu âm thanh vận chuyển dạng biểu diễn rút gọn của cảnh âm thanh không gian. Do đó, dòng hỗn hợp âm thanh có thể liên quan đến tín hiệu âm thanh dựa trên kênh với số kênh xác định trước. Cần phải hiểu rằng số kênh của tín hiệu âm thanh dựa trên kênh nhỏ hơn số kênh hoặc số đối tượng âm thanh của tín hiệu âm thanh không gian. Ví dụ, tín hiệu âm thanh dựa trên kênh có thể là tín hiệu âm thanh Ambisonics bậc nhất. Nói cách khác, dòng âm thanh không gian thu gọn có thể bao gồm dòng hỗn hợp âm thanh ở dạng dạng biểu diễn Ambisonics bậc nhất của trường âm.

Dòng siêu dữ liệu (có hướng) bao gồm siêu dữ liệu xác định các thuộc tính không gian của cảnh âm thanh không gian. Siêu dữ liệu có hướng có thể bao gồm chuỗi các khối siêu dữ liệu có hướng, trong đó mỗi khối siêu dữ liệu có hướng chứa siêu dữ liệu biểu thị các thuộc tính của cảnh âm thanh không gian trong đoạn thời gian tương ứng trong dòng hỗn hợp âm thanh.

Nói chung, siêu dữ liệu bao gồm thông tin hướng và thông tin năng lượng. Thông tin hướng bao gồm các chỉ báo về hướng đến của một hoặc nhiều phần tử âm thanh (trội) trong cảnh âm thanh. Thông tin năng lượng bao gồm, đối với mỗi hướng đến, chỉ báo

về công suất tín hiệu kết hợp với các hướng đến xác định được. Theo một số phương án thực hiện, các chỉ báo về công suất tín hiệu có thể được cung cấp cho một, một số hoặc mỗi trong số nhiều dải (dải tần số con). Hơn nữa, siêu dữ liệu có thể được cung cấp cho mỗi trong số nhiều đoạn thời gian liên tiếp, chẳng hạn như ở dạng khối siêu dữ liệu.

Trong một ví dụ, siêu dữ liệu (siêu dữ liệu có hướng) bao gồm siêu dữ liệu biểu thị các thuộc tính của cảnh âm thanh không gian trên một số dải tần, trong đó siêu dữ liệu xác định:

- Một hoặc nhiều hướng (ví dụ, hướng đến) biểu thị vị trí của các đối tượng âm thanh (phần tử âm thanh) trong cảnh âm thanh không gian, và
- phần năng lượng (hoặc công suất tín hiệu), trong mỗi dải tần, được quy cho đối tượng âm thanh tương ứng (ví dụ, được quy cho hướng tương ứng).

Chi tiết về việc xác định thông tin hướng và thông tin năng lượng sẽ được cung cấp bên dưới.

Fig.1 thể hiện bằng sơ đồ ví dụ về cách bố trí sử dụng các phương án của sáng chế. Cụ thể, hình vẽ thể hiện cách bố trí 100 trong đó cảnh âm thanh không gian 10 được đưa vào bộ mã hóa cảnh 200 để tạo ra dòng hỗn hợp âm thanh 30 và dòng siêu dữ liệu có hướng 20. Cảnh âm thanh không gian 10 có thể được biểu diễn bằng tín hiệu âm thanh không gian hoặc dòng âm thanh không gian được đưa vào bộ mã hóa cảnh 200. Dòng hỗn hợp âm thanh 30 và dòng siêu dữ liệu có hướng 20 cùng nhau tạo ra một ví dụ về cảnh âm thanh không gian thu gọn, tức là, dạng biểu diễn nén của cảnh âm thanh không gian 10 (hoặc của tín hiệu âm thanh không gian).

Dạng biểu diễn nén, tức là dòng âm thanh hỗn hợp 30 và dòng siêu dữ liệu có hướng 20 được đưa vào bộ giải mã cảnh 300 để tạo ra cảnh âm thanh tái tạo 50. Các phần tử âm thanh tồn tại trong cảnh âm thanh không gian 10 sẽ được biểu diễn trong dòng hỗn hợp âm thanh 30 theo hàm điều chỉnh vị trí hỗn hợp.

Fig.2 thể hiện bằng sơ đồ một ví dụ khác về cách bố trí sử dụng các phương án của sáng chế. Cụ thể, hình vẽ thể hiện cách bố trí thay thế 110, trong đó cảnh âm thanh không gian thu gọn, bao gồm dòng hỗn hợp âm thanh 30 và dòng siêu dữ liệu có hướng 20, được mã hóa thêm bằng cách cung cấp dòng hỗn hợp âm thanh 30 cho bộ mã hóa âm thanh 35 để tạo ra dòng âm thanh mã hóa có tốc độ bit giảm 37 và bằng cách cung cấp

cấp dòng siêu dữ liệu có hướng 20 cho bộ mã hóa siêu dữ liệu 25 để tạo ra dòng siêu dữ liệu mã hóa 27. Dòng âm thanh mã hóa có tốc độ bit giảm 37 và dòng siêu dữ liệu mã hóa 27 cùng tạo ra cảnh âm thanh không gian mã hóa (mã hóa có tốc độ bit giảm).

Cảnh âm thanh không gian được mã hóa có thể được phục hồi bằng cách trước tiên áp dụng dòng âm thanh mã hóa có tốc độ bit giảm 37 và dòng siêu dữ liệu mã hóa 27 cho các bộ giải mã tương ứng 36 và 26 để tạo ra dòng hỗn hợp âm thanh 38 và dòng siêu dữ liệu có hướng đã phục hồi 28. Các dòng phục hồi 38, 28 có thể giống hoặc gần bằng với các dòng 30, 20 tương ứng. Dòng hỗn hợp âm thanh đã phục hồi 38 và dòng siêu dữ liệu có hướng đã phục hồi 28 có thể được giải mã bởi bộ giải mã 300 để tạo ra cảnh âm thanh tái tạo 50.

Fig.3 minh họa bằng sơ đồ một ví dụ về cách bố trí để tạo ra dòng âm thanh mã hóa có tốc độ bit giảm và dòng siêu dữ liệu mã hóa từ cảnh âm thanh không gian đầu vào. Cụ thể, hình vẽ thể hiện cách bố trí 150 của bộ mã hóa cảnh 200 cung cấp dòng siêu dữ liệu có hướng 20 và dòng hỗn hợp âm thanh 30 cho các bộ mã hóa tương ứng 25, 35 để tạo ra cảnh âm thanh không gian mã hóa 40 bao gồm dòng âm thanh mã hóa có tốc độ bit giảm 37 và dòng siêu dữ liệu mã hóa 27. Tốt nhất là nên bố trí dòng âm thanh không gian mã hóa 40 để thích hợp cho việc lưu trữ và/hoặc truyền với yêu cầu dữ liệu giảm, so với dữ liệu cần thiết để lưu trữ/truyền cảnh âm thanh không gian ban đầu.

Fig.4 minh họa bằng sơ đồ một ví dụ về cách bố trí để tạo ra cảnh âm thanh không gian tái tạo từ dòng âm thanh mã hóa có tốc độ bit giảm và dòng siêu dữ liệu mã hóa. Cụ thể, hình vẽ thể hiện cách bố trí 160, trong đó dòng âm thanh không gian mã hóa 40, bao gồm dòng âm thanh mã hóa có tốc độ bit giảm 37 và dòng siêu dữ liệu mã hóa 27, được cung cấp làm đầu vào cho các bộ giải mã 36, 26 để lần lượt tạo ra dòng hỗn hợp âm thanh 38 và dòng siêu dữ liệu có hướng 28. Các dòng 38, 28 sau đó được xử lý bởi bộ giải mã cảnh 300 để tạo ra cảnh âm thanh tái tạo 50.

Chi tiết về việc tạo ra cảnh âm thanh không gian thu gọn, tức là dạng biểu diễn nén của cảnh âm thanh không gian (hoặc của tín hiệu âm thanh không gian/dòng âm thanh không gian) sẽ được mô tả tiếp theo.

Fig.5 là lưu đồ của một ví dụ về phương pháp 500 để xử lý tín hiệu âm thanh không gian để tạo ra dạng biểu diễn nén của tín hiệu âm thanh không gian. Phương pháp 500 bao gồm các bước từ S510 đến S550.

Ở bước S510, tín hiệu âm thanh không gian được phân tích để xác định hướng đến cho một hoặc nhiều phần tử âm thanh (ví dụ, phần tử âm thanh trội) trong cảnh âm thanh (cảnh âm thanh không gian) được biểu diễn bằng tín hiệu âm thanh không gian. Ví dụ, các phần tử âm thanh (trội) có thể liên quan đến các đối tượng âm thanh (trội), nguồn âm thanh (trội) hoặc các thành phần âm thanh (trội) trong cảnh âm thanh. Bước phân tích tín hiệu âm thanh không gian có thể bao gồm hoặc có thể liên quan đến việc áp dụng phân tích cảnh cho tín hiệu âm thanh không gian. Cần phải hiểu rằng một loạt các công cụ phân tích cảnh thích hợp được người có hiểu biết trung bình trong lĩnh vực kỹ thuật này biết rõ. Các hướng đến được xác định ở bước này có thể tương ứng với các vị trí trên hình cầu đơn vị biểu thị vị trí (cảm nhận được) của các phần tử âm thanh.

Theo phần mô tả ở trên về phân tích tạo dải tần số, bước phân tích tín hiệu âm thanh không gian ở bước S510 có thể dựa trên nhiều dải tần số con của tín hiệu âm thanh không gian. Ví dụ, bước phân tích có thể được dựa trên toàn bộ khoảng tần số của tín hiệu âm thanh không gian (tức là, toàn bộ tín hiệu). Tức là, bước phân tích có thể được dựa trên tất cả các dải tần số con.

Ở bước S520, các chỉ báo tương ứng về công suất tín hiệu kết hợp với các hướng đến xác định được được xác định cho ít nhất một dải tần số con của tín hiệu âm thanh không gian.

Ở bước S530, siêu dữ liệu bao gồm thông tin hướng và thông tin năng lượng được tạo ra. Thông tin hướng bao gồm các chỉ báo về các hướng đến xác định được của một hoặc nhiều phần tử âm thanh. Thông tin năng lượng bao gồm các chỉ báo tương ứng về công suất tín hiệu kết hợp với các hướng đến xác định được. Siêu dữ liệu được tạo ra ở bước này có thể liên quan đến dòng siêu dữ liệu.

Ở bước S540, tín hiệu âm thanh dựa trên kênh với số kênh xác định trước được tạo ra dựa trên tín hiệu âm thanh không gian.

Cuối cùng, ở bước S550, tín hiệu âm thanh dựa trên kênh và siêu dữ liệu được xuất ra dưới dạng dạng biểu diễn nén của tín hiệu âm thanh không gian.

Cần phải hiểu rằng các bước trên có thể được thực hiện theo thứ tự bất kỳ hoặc song song với nhau, miễn là thứ tự các bước đảm bảo rằng có sẵn thông tin đầu vào cần thiết cho mỗi bước.

Thông thường, cảnh không gian (hoặc tín hiệu âm thanh không gian) có thể được coi là bao gồm tổng của các tín hiệu âm thanh truyền tới người nghe từ tập hợp các hướng, liên quan đến vị trí nghe. Do đó cảnh âm thanh không gian có thể được mô hình hóa là tập hợp của R đối tượng âm thanh, trong đó đối tượng r ($1 \leq r \leq R$) được kết hợp với tín hiệu âm thanh $o_r(t)$ truyền tới vị trí nghe từ hướng đến được xác định bởi vectơ có hướng θ_r . Vectơ có hướng cũng có thể là vectơ có hướng thay đổi theo thời gian $\theta_r(t)$.

Do đó, theo một số phương án thực hiện, tín hiệu âm thanh không gian (dòng âm thanh không gian) có thể được xác định là tín hiệu âm thanh không gian dựa trên đối tượng (cảnh âm thanh không gian dựa trên đối tượng), dưới dạng tập hợp các tín hiệu âm thanh và vectơ có hướng liên quan

$$\text{Cảnh âm thanh không gian (dựa trên đối tượng)} = \{(o_r(t), \theta_r(t)): 1 \leq r \leq R\} \quad (14)$$

Hơn nữa, theo một số phương án thực hiện, tín hiệu âm thanh không gian (dòng âm thanh không gian) có thể được xác định theo tín hiệu biến đổi Fourier ngắn hạn, $O_{r,k}(f)$, theo phương trình (4) và vectơ có hướng có thể được định rõ theo chỉ số khối, k , sao cho:

$$\text{Cảnh âm thanh không gian (dựa trên đối tượng)} = \{(O_{r,k}(f), \theta_r(k)): 1 \leq r \leq R\} \quad (15)$$

Ngoài ra, tín hiệu âm thanh không gian (dòng âm thanh không gian) có thể được biểu diễn dưới dạng tín hiệu âm thanh không gian dựa trên kênh (cảnh âm thanh không gian dựa trên kênh). Dòng dựa trên kênh bao gồm tập hợp các tín hiệu âm thanh, trong đó mỗi đối tượng âm thanh từ cảnh âm thanh không gian được trộn vào các kênh theo hàm điều chỉnh vị trí ($Pan(\theta)$), theo phương trình (1). Theo một ví dụ, cảnh âm thanh không gian dựa trên kênh Q kênh, $\{C_{q,k}(f): 1 \leq q \leq Q\}$, có thể được tạo ra từ cảnh âm thanh không gian dựa trên đối tượng theo

$$\begin{aligned} \text{Cảnh âm thanh không gian (dựa trên đối tượng)} &= \{C_{q,k}(f): 1 \leq q \leq Q\} \\ \text{trong đó } \begin{pmatrix} C_{1,k}(f) \\ C_{2,k}(f) \\ \vdots \\ C_{Q,k}(f) \end{pmatrix} &= \sum_{r=1}^R Pan(\theta_r(k)) O_{r,k}(f) \end{aligned} \quad (16)$$

Cần phải hiểu rằng nhiều đặc điểm của cảnh âm thanh không gian dựa trên kênh được xác định bởi sự lựa chọn của hàm điều chỉnh vị trí, và đặc biệt là độ dài (Q) của vector cột được trả về bởi hàm điều chỉnh vị trí sẽ xác định số lượng kênh âm thanh chứa trong cảnh âm thanh không gian dựa trên kênh. Nói chung, dạng biểu diễn có chất lượng cao hơn của cảnh âm thanh không gian có thể được thực hiện bằng cảnh âm thanh không gian dựa trên kênh chứa số lượng kênh lớn hơn.

Ví dụ, tại bước S540 của phương pháp 500, tín hiệu âm thanh không gian (cảnh âm thanh không gian) có thể được xử lý để tạo ra tín hiệu âm thanh dựa trên kênh (dòng dựa trên kênh) theo phương trình (16). Hàm điều chỉnh vị trí có thể được chọn để tạo ra dạng biểu diễn có độ phân giải tương đối thấp của cảnh âm thanh không gian. Ví dụ, hàm điều chỉnh vị trí có thể được chọn làm hàm Ambisonics bậc nhất (FOA), chẳng hạn như được xác định trong phương trình (2). Do vậy, dạng biểu diễn nén có thể là dạng biểu diễn thu gọn hoặc rút gọn kích thước.

Fig.6 là lưu đồ cung cấp một cách trình bày khác của phương pháp 600 để tạo ra dạng biểu diễn thu gọn của cảnh âm thanh không gian. Phương pháp 600 được cung cấp dòng đầu vào, dưới dạng cảnh âm thanh không gian hoặc dòng dựa trên cảnh, và tạo ra cảnh âm thanh không gian thu gọn dưới dạng dạng biểu diễn thu gọn. Để đạt được mục đích này, phương pháp 600 bao gồm các bước từ S610 đến S660. Trong đó, bước S610 có thể được xem là tương ứng với bước S510, bước 620 có thể được xem là tương ứng với bước S520, bước S630 có thể được xem là tương ứng với bước S540, bước S650 có thể được xem là tương ứng với bước S530 và bước S660 có thể được xem như tương ứng với bước S550.

Ở bước S610, dòng đầu vào được phân tích để xác định các hướng đến trội.

Ở bước S620, đối với mỗi dải (dải tần số con), một phần năng lượng phân bổ cho mỗi hướng được xác định, so với tổng năng lượng trong dòng của dải đó.

Ở bước S630, dòng giảm mức được tạo ra, chứa một số kênh âm thanh biểu diễn cảnh âm thanh không gian.

Ở bước S640, dòng giảm mức được mã hóa để tạo ra dạng biểu diễn nén của dòng.

Ở bước S650, thông tin hướng và thông tin phần năng lượng được mã hóa để tạo ra siêu dữ liệu mã hóa.

Cuối cùng, ở bước S660 dòng giảm mức mã hóa được kết hợp với siêu dữ liệu mã hóa để tạo ra cảnh âm thanh không gian thu gọn.

Cần phải hiểu rằng các bước trên có thể được thực hiện theo thứ tự bất kỳ hoặc song song với nhau, miễn là thứ tự các bước đảm bảo rằng có sẵn thông tin đầu vào cần thiết cho mỗi bước.

Các hình vẽ từ Fig.7 đến Fig.11 minh họa bằng sơ đồ các ví dụ về chi tiết của việc tạo ra dạng biểu diễn nén của cảnh âm thanh không gian, theo các phương án của sáng chế. Cần phải hiểu rằng các chi tiết về, ví dụ, phân tích tín hiệu âm thanh không gian để xác định hướng đến, xác định chỉ báo về công suất tín hiệu kết hợp với các hướng đến xác định được, tạo ra siêu dữ liệu bao gồm thông tin hướng và thông tin năng lượng, và/hoặc tạo ra kênh tín hiệu âm thanh dựa trên kênh với số lượng kênh xác định trước như được mô tả dưới đây có thể độc lập với bố cục hệ thống cụ thể và có thể áp dụng cho, ví dụ, bất kỳ trong số các cách bố trí được thể hiện trên các hình vẽ từ Fig.7 đến Fig.11, hoặc cách bố trí khác thích hợp bất kỳ.

Fig.7 minh họa bằng sơ đồ ví dụ thứ nhất về các chi tiết của việc tạo ra dạng biểu diễn nén của cảnh âm thanh không gian. Cụ thể, Fig.7 thể hiện bộ mã hóa cảnh 200 trong đó cảnh âm thanh không gian 10 được xử lý bằng hàm giảm mức 203 để tạo ra dòng hỗn hợp âm thanh N kênh 30, theo, ví dụ, các bước S540 và S630. Theo một số phương án, hàm giảm mức 203 có thể bao gồm quy trình điều chỉnh vị trí theo phương trình (1) hoặc phương trình (16), trong đó hàm điều chỉnh vị trí giảm mức được chọn: $Pan(\theta) = Pan_{down}(\theta)$. Ví dụ, bộ điều chỉnh vị trí Ambisonics bậc nhất có thể được chọn làm hàm điều chỉnh vị trí giảm mức: $Pan_{down}(\theta) = Pan_{FOA}(\theta)$ và do đó $N = 4$.

Đối với mỗi đoạn thời gian âm thanh, phân tích cảnh 202 lấy cảnh âm thanh không gian làm đầu vào, và xác định hướng đến của tối đa P thành phần âm thanh trội trong cảnh âm thanh không gian, ví dụ, theo các bước S510 và S610. Các giá trị điển hình của P ở giữa 1 và 10, và giá trị ưu tiên của P là $P \approx 4$. Do đó, một hoặc nhiều phần tử âm thanh được xác định ở bước S510 có thể bao gồm từ một đến mười phần tử âm thanh, như bốn phần tử âm thanh, chẳng hạn.

Phân tích cảnh 202 tạo ra dòng siêu dữ liệu 20 bao gồm thông tin hướng 21 và thông tin đoạn dải năng lượng 22 (thông tin năng lượng). Tùy ý, phân tích cảnh 202 cũng có thể cung cấp hệ số 207 cho hàm giảm mức 203 để cho phép thay đổi việc giảm mức.

Không có giới hạn dự kiến, việc phân tích tín hiệu âm thanh không gian (ví dụ, ở bước S510), xác định các chỉ báo tương ứng về công suất tín hiệu (ví dụ, ở bước S520), và tạo ra tín hiệu âm thanh dựa trên kênh (ví dụ, ở bước S540) có thể được thực hiện trên cơ sở từng đoạn thời gian, ví dụ, theo mô tả ở trên về STFT. Việc này ngụ ý rằng dạng biểu diễn nén sẽ được tạo ra và xuất cho mỗi trong số nhiều đoạn thời gian, với tín hiệu âm thanh giảm mức và siêu dữ liệu (khối siêu dữ liệu) cho từng đoạn thời gian.

Đối với mỗi đoạn thời gian, k , thông tin hướng 21 (ví dụ, thể hiện bằng hướng đến của một hoặc nhiều phần tử âm thanh) có thể có dạng P vectơ có hướng, $\{dir_{k,p} : 1 \leq p \leq P\}$. Vectơ có hướng p biểu thị hướng kết hợp với chỉ số đối tượng trội p , và có thể được biểu diễn theo các vectơ đơn vị,

$$\begin{aligned} dir_{k,p} &= (x_{k,p}, y_{k,p}, z_{k,p}) \\ \text{trong đó: } x_{k,p}^2 + y_{k,p}^2 + z_{k,p}^2 &= 1 \end{aligned} \tag{17}$$

hoặc theo tọa độ cầu,

$$\begin{aligned} dir_{k,p} &= (az_{k,p}, el_{k,p}) \\ \text{trong đó: } -180 \leq az_{k,p} \leq 180 \text{ và } -90 \leq el_{k,p} \leq 90 \end{aligned} \tag{18}$$

Theo một số phương án, các chỉ báo tương ứng về công suất tín hiệu được xác định ở bước S520 có dạng một phần công suất tín hiệu. Tức là, chỉ báo về công suất tín hiệu kết hợp với hướng đến cho trước trong dải tần số con liên quan đến một phần công

suất tín hiệu trong dải tần số con đối với hướng đến cho trước liên quan đến tổng công suất tín hiệu trong dải tần số con.

Hơn nữa, theo một số phương án, các chỉ báo về công suất tín hiệu được xác định cho mỗi trong số nhiều dải tần số con (chẳng hạn, theo cách trên mỗi dải con). Sau đó, chúng liên quan, đối với hướng đến cho trước và dải tần số con cho trước, đến một phần công suất tín hiệu trong dải tần số con cho trước đối với hướng đến cho trước liên quan đến tổng công suất tín hiệu trong dải tần số con cho trước. Đặc biệt, mặc dù các chỉ báo về công suất tín hiệu có thể được xác định theo cách trên mỗi dải con, nhưng việc xác định hướng đến (trội) vẫn có thể được thực hiện trên toàn bộ tín hiệu (tức là, dựa trên tất cả các dải tần số con).

Tuy nhiên, theo một số phương án, việc phân tích tín hiệu âm thanh không gian (ví dụ, ở bước S510), xác định các chỉ báo tương ứng về công suất tín hiệu (ví dụ, ở bước S520), và tạo ra tín hiệu âm thanh dựa trên kênh (ví dụ, ở bước S540) được thực hiện dựa trên dạng biểu diễn thời gian-tần số của tín hiệu âm thanh không gian. Ví dụ, các bước nêu trên và các bước khác nếu thích hợp có thể được thực hiện dựa trên phép biến đổi Fourier rời rạc (như STFT chẳng hạn) của tín hiệu âm thanh không gian. Ví dụ, đối với mỗi đoạn thời gian (khối thời gian), các bước nêu trên có thể được thực hiện dựa trên các bin thời gian-tần số (các bin FFT) của tín hiệu âm thanh không gian, tức là dựa trên các hệ số Fourier của tín hiệu âm thanh không gian.

Như đã nêu ở trên, đối với mỗi đoạn thời gian, k , và đối với mỗi chỉ số đối tượng trội p ($1 \leq p \leq P$), thông tin phân dải năng lượng 22 có thể bao gồm giá trị phân số $e_{k,p,b}$ cho mỗi dải b trong tập hợp các dải ($1 \leq b \leq B$). Giá trị phân số $e_{k,p,b}$ được xác định cho đoạn thời gian k theo:

$$e_{k,p,b} = \frac{\text{Năng lượng theo hướng } dir_{k,p} \text{ cho dải } b}{\text{Tổng năng lượng trong cảnh cho dải } b} \quad (19)$$

Giá trị phân số $e_{k,p,b}$ có thể biểu diễn một phần năng lượng trong vùng không gian xung quanh hướng $dir_{k,p}$, sao cho năng lượng của nhiều đối tượng âm thanh trong cảnh âm thanh không gian ban đầu có thể được kết hợp để biểu diễn một thành phần âm thanh trội duy nhất gán cho hướng $dir_{k,p}$. Theo một số phương án, năng lượng của tất

cả các đối tượng âm thanh trong cảnh có thể được gán trọng số, bằng cách sử dụng hàm trọng số theo chênh lệch góc $w(\theta)$ biểu diễn việc gán trọng số lớn hơn cho hướng, θ , tức là gần với $dir_{k,p}$, và việc gán trọng số nhỏ hơn cho hướng, θ , tức là xa $dir_{k,p}$. Chênh lệch hướng có thể được coi là gần đối với chênh lệch góc nhỏ hơn, ví dụ, 10° và là xa đối với chênh lệch góc lớn hơn, ví dụ, 45° . Trong các phương án thay thế, hàm trọng số có thể được chọn dựa trên các lựa chọn thay thế của sự khác biệt góc gần/góc xa.

Nói chung, tín hiệu âm thanh không gian đầu vào mà dạng biểu diễn nén được tạo ra có thể là tín hiệu âm thanh đa kênh hoặc tín hiệu âm thanh dựa trên đối tượng, ví dụ. Trong trường hợp thứ hai, phương pháp tạo ra dạng biểu diễn nén của tín hiệu âm thanh không gian sẽ bao gồm thêm bước chuyển đổi tín hiệu âm thanh dựa trên đối tượng thành tín hiệu âm thanh đa kênh trước khi áp dụng phân tích cảnh (ví dụ, trước bước S510).

Trong ví dụ trên Fig.7, tín hiệu âm thanh không gian đầu vào có thể là tín hiệu âm thanh đa kênh. Sau đó, tín hiệu âm thanh dựa trên kênh được tạo ra ở bước S540 sẽ là tín hiệu giảm mức được tạo ra bằng cách áp dụng phép giảm mức cho tín hiệu âm thanh đa kênh.

Fig.8 minh họa bằng sơ đồ một ví dụ khác về các chi tiết tạo ra dạng biểu diễn nén của cảnh âm thanh không gian. Tín hiệu âm thanh không gian đầu vào trong trường hợp này có thể là tín hiệu âm thanh dựa trên đối tượng bao gồm nhiều đối tượng âm thanh và vectơ có hướng liên quan. Trong trường hợp này, phương pháp tạo ra dạng biểu diễn nén của tín hiệu âm thanh không gian bao gồm bước tạo ra tín hiệu âm thanh đa kênh, dưới dạng dạng biểu diễn trung gian hoặc cảnh trung gian, bằng cách điều chỉnh vị trí các đối tượng âm thanh đến tập hợp các kênh âm thanh xác định trước, trong đó mỗi đối tượng âm thanh được điều chỉnh vị trí đến tập hợp các kênh âm thanh xác định trước theo vectơ có hướng của nó. Do đó, Fig.8 thể hiện phương án thay thế của bộ mã hóa cảnh 200, trong đó cảnh âm thanh không gian 10 được đưa vào bộ chuyển đổi 201 tạo ra cảnh trung gian 11 (ví dụ, được thể hiện bằng tín hiệu đa kênh). Cảnh trung gian 11 có thể được tạo ra theo phương trình (1) trong đó hàm điều chỉnh vị trí được chọn để tích vô hướng của vectơ độ lợi điều chỉnh vị trí $Pan(\theta_1)$ và $Pan(\theta_2)$ biểu diễn gần đúng hàm gán trọng số theo chênh lệch góc, như được mô tả ở trên.

Theo một số phương án, hàm điều chỉnh vị trí được dùng trong bộ chuyển đổi 201 hàm điều chỉnh vị trí Ambisonics bậc ba, $Pan_{30A}(\theta)$, như thể hiện trong Phương trình (3). Do đó, tín hiệu âm thanh đa kênh có thể là tín hiệu Ambisonics bậc cao, chẳng hạn.

Sau đó cảnh trung gian 11 được xuất ra cho phân tích cảnh 202. Phân tích cảnh 202 có thể xác định các hướng, $dir_{k,p}$, của các đối tượng âm thanh trội trong cảnh âm thanh không gian từ việc phân tích cảnh trung gian 11. Bước xác định các hướng trội có thể được thực hiện bằng cách ước tính năng lượng theo tập hợp các hướng, với năng lượng ước tính lớn nhất biểu diễn hướng trội.

Thông tin phần dải năng lượng 22 của đoạn thời gian k có thể bao gồm giá trị phân số $e_{k,p,b}$ đối với mỗi dải b mà bắt nguồn từ năng lượng trong dải b của cảnh trung gian theo mỗi hướng $dir_{k,p}$, so với tổng năng lượng trong dải b của cảnh trung gian 11 trong đoạn thời gian k .

Dòng hỗn hợp âm thanh 30 (ví dụ, tín hiệu âm thanh dựa trên kênh) của cảnh âm thanh không gian thu gọn (ví dụ, dạng biểu diễn thu gọn) trong trường hợp này là tín hiệu giảm mức được tạo ra bằng cách áp dụng hàm giảm mức 203 (phép giảm mức) cho cảnh âm thanh không gian.

Fig.10 thể hiện cách bố trí thay thế của bộ mã hóa cảnh bao gồm bộ chuyển đổi 201 để chuyển đổi cảnh âm thanh không gian 10 thành định dạng trung gian dựa trên cảnh 11. Định dạng trung gian 11 là đầu vào cho phân tích cảnh 202 và cho hàm giảm mức 203. Theo một số phương án, hàm giảm mức 203 có thể bao gồm bộ trộn ma trận với các hệ số được làm thích ứng để chuyển đổi định dạng trung gian 11 thành dòng hỗn hợp âm thanh 30. Tức là, dòng hỗn hợp âm thanh 30 (ví dụ, tín hiệu âm thanh dựa trên kênh) của cảnh âm thanh không gian thu gọn (ví dụ, dạng biểu diễn thu gọn) trong trường hợp này có thể là tín hiệu giảm mức được tạo ra bằng cách áp dụng hàm giảm mức 203 (phép giảm mức) cho cảnh trung gian (ví dụ, tín hiệu âm thanh đa kênh).

Trong một phương án thay thế, được thể hiện trên Fig.11, bộ mã hóa không gian 200 có thể nhận đầu vào ở dạng đầu vào dựa trên cảnh 11, trong đó các đối tượng âm thanh được biểu diễn theo quy tắc điều chỉnh vị trí, $Pan(\theta)$. Theo một số phương án, hàm điều chỉnh vị trí có thể là hàm điều chỉnh vị trí Ambisonics bậc cao hơn. Trong một phương án ví dụ, hàm điều chỉnh vị trí là hàm điều chỉnh vị trí Ambisonics bậc ba.

Trong một phương án thay thế khác, được minh họa trên Fig.9, cảnh âm thanh không gian 10 được chuyển đổi bằng bộ chuyển đổi 201 trong bộ mã hóa không gian 200 để tạo ra cảnh trung gian 11 mà được nhập vào hàm giảm mức 203. Phân tích cảnh 202 được cung cấp đầu vào từ cảnh âm thanh không gian 10.

Fig.12 minh họa bằng sơ đồ một ví dụ về chi tiết của việc giải mã dạng biểu diễn nén của cảnh âm thanh không gian để tạo ra cảnh âm thanh tái tạo, theo các phương án của sáng chế. Cụ thể, hình vẽ thể hiện một bộ giải mã cảnh 300 bao gồm bộ khử trộn 302 lấy dòng hỗn hợp âm thanh 30 và tạo ra dòng âm thanh không gian riêng 70. Dòng âm thanh không gian riêng 70 bao gồm P tín hiệu đối tượng trội 90 và dòng dư 80. Bộ giải mã dư 81 lấy đầu vào từ dòng dư 80 và tạo ra dòng dư giải mã 82. Bộ điều chỉnh vị trí đối tượng 91 lấy đầu vào từ các tín hiệu đối tượng trội 90 và tạo ra dòng đối tượng được điều chỉnh vị trí 92. Dòng dư đã giải mã 82 và dòng đối tượng đã điều chỉnh vị trí 92 được cộng lại 75 để tạo ra cảnh âm thanh tái tạo 50.

Hơn nữa, Fig.12 thể hiện thông tin hướng 21 và thông tin phân dải năng lượng 22 được nhập vào máy tính ma trận khử trộn 301 để xác định ma trận khử trộn 60 (ma trận trộn ngược) được dùng bởi bộ khử trộn 302.

Chi tiết về việc xử lý cảnh âm thanh không gian thu gọn (ví dụ, dạng biểu diễn nén của tín hiệu âm thanh không gian) để tạo ra dạng biểu diễn tái tạo của tín hiệu âm thanh không gian sẽ được mô tả tiếp theo.

Fig.13 là lưu đồ của một ví dụ của phương pháp 1300 để xử lý dạng biểu diễn nén của tín hiệu âm thanh không gian để tạo ra dạng biểu diễn tái tạo của tín hiệu âm thanh không gian. Cần phải hiểu rằng dạng biểu diễn nén bao gồm tín hiệu âm thanh dựa trên kênh (ví dụ, được thể hiện bởi dòng hỗn hợp âm thanh 30) với số lượng kênh xác định trước và siêu dữ liệu, siêu dữ liệu bao gồm thông tin hướng (ví dụ, được thể hiện bằng thông tin hướng 21) và thông tin năng lượng (ví dụ, được thể hiện bằng thông tin đoạn dải năng lượng 22), với thông tin hướng bao gồm các chỉ báo về hướng đến của một hoặc nhiều phần tử âm thanh trong cảnh âm thanh và thông tin năng lượng bao gồm, đối với ít nhất một dải tần số con, các chỉ báo tương ứng về công suất tín hiệu kết hợp với các hướng đến. Ví dụ, tín hiệu âm thanh dựa trên kênh có thể là tín hiệu Ambisonics bậc nhất. Phương pháp 1300 bao gồm các bước S1310 và S1320, và tùy ý, các bước

S1330 và S1340. Ví dụ, cần hiểu rằng các bước này có thể được thực hiện bởi bộ giải mã cảnh 300 trên Fig.12.

Ở bước S1310, tín hiệu âm thanh của một hoặc nhiều phần tử âm thanh được tạo ra dựa trên tín hiệu âm thanh dựa trên kênh, thông tin hướng và thông tin năng lượng.

Ở bước S1320, tạo ra tín hiệu âm thanh dư mà một hoặc nhiều phần tử âm thanh về cơ bản không có ở đó, dựa vào tín hiệu âm thanh dựa trên kênh, thông tin hướng và thông tin năng lượng. Ở đây, tín hiệu dư có thể được biểu diễn theo định dạng âm thanh giống như tín hiệu âm thanh dựa trên kênh, ví dụ, có thể có cùng số kênh với tín hiệu âm thanh dựa trên kênh.

Ở bước tùy chọn S1330, tín hiệu âm thanh của một hoặc nhiều phần tử âm thanh được điều chỉnh vị trí đến tập hợp các kênh có định dạng âm thanh đầu ra. Ở đây, định dạng âm thanh đầu ra có thể liên quan đến dạng biểu diễn đầu ra, chẳng hạn như HOA hoặc định dạng đa kênh thích hợp khác bất kỳ.

Ở bước tùy chọn S1340, tín hiệu âm thanh đa kênh tái tạo theo định dạng âm thanh đầu ra được tạo ra dựa trên một hoặc nhiều phần tử âm thanh được điều chỉnh vị trí và tín hiệu dư. Bước tạo ra tín hiệu âm thanh đa kênh tái tạo có thể bao gồm bước tăng mức tín hiệu dư đến tập hợp các kênh của định dạng âm thanh đầu ra. Bước tạo ra tín hiệu âm thanh đa kênh tái tạo có thể còn bao gồm bước thêm một hoặc nhiều phần tử âm thanh đã điều chỉnh vị trí và tín hiệu dư đã tăng mức.

Cần phải hiểu rằng các bước trên có thể được thực hiện theo thứ tự bất kỳ hoặc song song với nhau, miễn là thứ tự các bước đảm bảo rằng có sẵn thông tin đầu vào cần thiết cho mỗi bước.

Theo phần mô tả ở trên về các phương pháp xử lý cảnh âm thanh không gian để tạo ra dạng biểu diễn nén của cảnh âm thanh không gian, chỉ báo về công suất tín hiệu kết hợp với hướng đến cho trước có thể liên quan đến một phần công suất tín hiệu trong dải tần số con đối với hướng đến cho trước liên quan đến tổng công suất tín hiệu trong dải tần số con.

Hơn nữa, theo một số phương án, thông tin năng lượng có thể bao gồm các chỉ báo về công suất tín hiệu cho mỗi trong số nhiều dải tần số con. Sau đó, chỉ báo về công suất tín hiệu có thể liên quan, đối với hướng đến cho trước và dải tần số con cho trước,

đến một phần công suất tín hiệu trong dải tần số con cho trước đối với hướng đến cho trước liên quan đến tổng công suất tín hiệu trong dải tần số con cho trước.

Bước tạo ra tín hiệu âm thanh của một hoặc nhiều phần tử âm thanh ở bước S1310 có thể bao gồm bước xác định các hệ số của ma trận trộn ngược M để ánh xạ tín hiệu âm thanh dựa trên kênh lên dạng biểu diễn trung gian bao gồm tín hiệu âm thanh dư và tín hiệu âm thanh của một hoặc nhiều phần tử âm thanh, dựa trên thông tin hướng và thông tin năng lượng. Dạng dạng biểu diễn trung gian cũng có thể được gọi là dạng biểu diễn riêng hoặc tách rời được, hoặc dạng biểu diễn kết hợp.

Chi tiết về bước xác định các hệ số của ma trận trộn ngược M sẽ được mô tả tiếp theo dựa vào lưu đồ trên Fig.14. Phương pháp 1400 được minh họa bởi lưu đồ này bao gồm các bước từ S1410 đến S1440.

Ở bước S1410, đối với mỗi trong số một hoặc nhiều phần tử âm thanh, vector điều chỉnh vị trí $Pan_{down}(dir)$ để điều chỉnh vị trí phần tử âm thanh đến các kênh của tín hiệu âm thanh dựa trên kênh được xác định, dựa vào hướng đến dir của phần tử âm thanh.

Ở bước S1420, ma trận trộn E sẽ được dùng để ánh xạ tín hiệu âm thanh dư và các tín hiệu âm thanh của một hoặc nhiều phần tử âm thanh lên các kênh của tín hiệu âm thanh dựa trên kênh được xác định, dựa vào các vector điều chỉnh vị trí xác định được.

Ở bước S1430, ma trận hiệp phương sai S cho dạng biểu diễn trung gian được xác định dựa vào thông tin năng lượng. Bước xác định ma trận hiệp phương sai S có thể còn được dựa vào các vector điều chỉnh vị trí xác định được Pan_{down} .

Cuối cùng, ở bước S1440, các hệ số của ma trận trộn ngược M được xác định dựa vào ma trận trộn E và ma trận hiệp phương sai S .

Cần phải hiểu rằng các bước trên có thể được thực hiện theo thứ tự bất kỳ hoặc song song với nhau, miễn là thứ tự các bước đảm bảo rằng có sẵn thông tin đầu vào cần thiết cho mỗi bước.

Trở lại hình vẽ trên Fig.12, máy tính ma trận khử trộn 301 tính toán ma trận khử trộn 60 (ma trận trộn ngược), $M_{k,b}$ theo quy trình bao gồm các bước sau:

1. Các đầu vào máy tính ma trận khử trộn, cho đoạn thời gian k , là thông tin hướng, $dir_{k,p}$ ($1 \leq p \leq P$), và thông tin phân dải năng lượng, $e_{k,p,b}$ ($1 \leq p \leq P$ và $1 \leq b \leq B$). P biểu diễn số lượng thành phần âm thanh trội và B biểu thị số lượng dải tần số.
2. Đối với mỗi dải, b , ma trận khử trộn $M_{k,b}$ được tính toán theo:

$$M = S \times E^* \times (E \times S \times E^*)^{-1} \quad (20)$$

trong đó “ $\times$ ” biểu thị phép nhân ma trận và “ $*$ ” biểu thị phép chuyển vị liên hợp của ma trận. Ví dụ việc tính toán theo Phương trình (20) có thể tương ứng với bước S1440.

Ma trận khử trộn M có thể được xác định cho mỗi trong số nhiều đoạn thời gian k , và/hoặc mỗi trong số nhiều dải tần số con b . Trong trường hợp đó, các ma trận M và S có chỉ số k biểu thị đoạn thời gian và/hoặc chỉ số b biểu thị dải tần số con, và ma trận E có chỉ số k biểu thị đoạn thời gian, chẳng hạn,

$$M_{k,b} = S_{k,b} \times E_k^* \times (E_k \times S_{k,b} \times E_k^*)^{-1} \quad (20a)$$

Nói chung, bước xác định các hệ số của ma trận trộn ngược M dựa vào ma trận trộn E và ma trận hiệp phương sai S có thể bao gồm bước xác định giá trị giả ngược dựa vào ma trận trộn E và ma trận hiệp phương sai S . Một ví dụ về giá trị giả ngược như vậy được xác định trong các phương trình (20) và (20a).

Trong Phương trình (20), ma trận E_k (ma trận trộn) được tạo ra bằng cách xếp chồng lên nhau ma trận đơn vị $N \times N$ (I_N) và P cột được tạo ra bằng hàm điều chỉnh vị trí được áp dụng cho các hướng của mỗi trong số P thành phần âm thanh trội:

$$E = (I_N | Pan_{down}(dir_1) | \dots | Pan_{down}(dir_P)) \quad (21)$$

trong Phương trình (21), I_N là $N \times N$ ma trận đơn vị, với N biểu thị số lượng kênh của tín hiệu dựa trên kênh. $Pan_{down}(dir_p)$ là vectơ điều chỉnh vị trí cho phần tử âm thanh thứ p có hướng đến liên quan dir_p mà điều chỉnh vị trí phần tử âm thanh thứ

N cho N kênh của tín hiệu dựa trên kênh, với $p = 1, \dots, P$ biểu thị phần tử âm thanh tương ứng trong số một hoặc nhiều phần tử âm thanh và P biểu thị tổng số của một hoặc nhiều phần tử âm thanh. Các dấu số thẳng trong Phương trình (21) biểu thị phép toán tăng ma trận. Do đó, ma trận E là ma trận $N \times P$.

Hơn nữa, ma trận E có thể được xác định cho mỗi trong số nhiều đoạn thời gian k . Trong trường hợp đó, ví dụ, ma trận E và các hướng đến dir_p có chỉ số k biểu thị đoạn thời gian, ví dụ,

$$E_k = (I_N | Pan_{down}(dir_{k,1}) | \dots | Pan_{down}(dir_{k,P})) \quad (21a)$$

Nếu phương pháp được đề xuất hoạt động theo dải, thì ma trận E có thể là như nhau đối với tất cả các dải tần số con.

Theo bước S1420, ma trận E_k là ma trận trộn mà được dùng để ánh xạ tín hiệu âm thanh dư và các tín hiệu âm thanh của một hoặc nhiều phần tử âm thanh lên các kênh của tín hiệu âm thanh dựa trên kênh. Như có thể thấy từ các phương trình (21) và (21a), ma trận E_k được dựa vào các vector điều chỉnh vị trí $Pan_{down}(dir)$ được xác định ở bước S1410.

Trong Phương trình (20), ma trận S là ma trận đường chéo $(N + P) \times (N + P)$. Nó có thể được xem như ma trận hiệp phương sai của dạng biểu diễn trung gian. Các hệ số của nó có thể được tính toán dựa vào thông tin năng lượng, theo bước S1430. N phần tử đường chéo thứ nhất được xác định bởi

$$\{S\}_{n,n} = \text{rms}(Pan_{down})_n \left(1 - \sum_{p=1}^P e_p \right) \quad (22)$$

với $1 \leq n \leq N$, và P phần tử đường chéo còn lại được xác định bởi

$$\{S\}_{N+p,N+p} = e_p \quad (23)$$

với $1 \leq p \leq P$, trong đó e_p là công suất tín hiệu kết hợp với hướng đến của phần tử âm thanh thứ p .

Ma trận hiệp phương sai S có thể được xác định cho mỗi trong số nhiều đoạn thời gian k , và/hoặc cho mỗi trong số nhiều dải tần số con b . Trong trường hợp đó, ma trận hiệp phương sai S và công suất tín hiệu e_p có chỉ số k biểu thị đoạn thời gian và/hoặc chỉ số b biểu thị dải tần số con. N phần tử đường chéo thứ nhất sẽ được xác định bởi

$$\{S_{k,b}\}_{n,n} = rms(Pan)_{down} \left(1 - \sum_{p=1}^P e_{k,p,b} \right) \quad (1 \leq n \leq N) \quad (22a)$$

và P phần tử đường chéo còn lại sẽ được xác định bởi

$$\{S_{k,b}\}_{N+p,N+p} = e_{k,p,b} \quad (1 \leq p \leq P) \quad (23a)$$

Theo phương án ưu tiên, ma trận khử trộn $M_{k,b}$ được áp dụng, bằng bộ khử trộn 302, để tạo ra dòng âm thanh không gian riêng 70 (là một ví dụ của dạng biểu diễn trung gian), theo phương án thực hiện mô tả ở trên ở bước S1310, trong đó N kênh thứ nhất là các dòng dư 80 và P kênh còn lại biểu diễn các thành phần âm thanh trội.

Dòng không gian riêng biệt $N + P$ kênh 70, $Y_k(f)$, các tín hiệu đối tượng trội P kênh 90 (là các ví dụ của các tín hiệu âm thanh của một hoặc nhiều phần tử âm thanh được tạo ra ở bước S1310), $O_k(f)$, và dòng dư N kênh 80 (là ví dụ của tín hiệu âm thanh dư được tạo ra ở bước S1320), $R_k(f)$, được tính toán từ hỗn hợp âm thanh N kênh 30, $X_k(f)$, theo:

$$\begin{aligned} Y_k(f) &= \sum_{b=1}^B FR_b(f) M_{k,b} \times X_k \\ R_k(f) &= \{Y_k(f)\}_{1..N} \\ O_k(f) &= \{Y_k(f)\}_{N+1..N+P} \end{aligned} \quad (24)$$

trong đó các tín hiệu được biểu diễn dưới dạng STFT, biểu thức $\{Y_k(f)\}_{1..N}$ biểu thị tín hiệu N kênh được tạo ra từ các kênh $1..N$ của $Y_k(f)$, và $\{Y_k(f)\}_{N+1..N+P}$ biểu thị tín hiệu P kênh từ các kênh $N+1..N+P$ của $Y_k(f)$. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật này sẽ hiểu rằng việc áp dụng ma trận $M_{k,b}$ có thể được thực hiện theo

các phương pháp thay thế đã biết đến trong lĩnh vực, tạo ra hàm gần đúng tương đương với hàm của phương trình (24).

Ngoài phần mô tả ở trên, theo một số phương án, số lượng thành phần âm thanh trội P có thể được làm thích ứng để nhận giá trị khác nhau cho mỗi đoạn thời gian, sao cho P_k có thể phụ thuộc vào chỉ số đoạn thời gian, k . Ví dụ, phân tích cảnh 202 trong bộ mã hóa cảnh 200 có thể xác định giá trị P_k cho mỗi đoạn thời gian. Nói chung, số lượng thành phần âm thanh trội P có thể phụ thuộc vào thời gian. Lựa chọn P (hoặc P_k) có thể bao gồm sự cân bằng giữa tốc độ dữ liệu siêu dữ liệu và chất lượng của cảnh âm thanh được tái tạo.

Trở lại Fig.12, bộ giải mã không gian 300 tạo ra cảnh âm thanh tái tạo M kênh 50 trong đó dòng M kênh được kết hợp với bộ điều chỉnh vị trí đầu ra $Pan_{out}(\theta)$. Quy trình này có thể được thực hiện theo bước S1340 mô tả ở trên. Các ví dụ về bộ điều chỉnh vị trí đầu ra bao gồm các hàm điều chỉnh vị trí âm thanh nổi, các hàm điều chỉnh vị trí biên độ dựa trên vector như đã biết trong lĩnh vực này, và các hàm điều chỉnh vị trí Ambisonics bậc cao hơn, như đã biết trong lĩnh vực kỹ thuật này.

Ví dụ, bộ điều chỉnh vị trí đối tượng 91 trên Fig.12 có thể được làm thích ứng để tạo ra dòng đối tượng được điều chỉnh vị trí M kênh 92, Z_p , theo

$$Z_p(f) = \sum_{p=1}^P Pan_{out}(\theta_p) O_{k,p}(f) \quad (25)$$

Fig.15 là lưu đồ cung cấp một cách trình bày khác của phương pháp 1500 để giải mã cảnh âm thanh không gian thu gọn để tạo ra cảnh âm thanh tái tạo. Phương pháp 1500 bao gồm các bước từ S1510 đến S1580.

Ở bước S1510, cảnh âm thanh không gian thu gọn được nhận và dòng giảm mức mã hóa và dòng siêu dữ liệu mã hóa được trích ra.

Ở bước S1520, dòng giảm mức mã hóa được giải mã để tạo ra dòng giảm mức.

Ở bước S1530, dòng siêu dữ liệu mã hóa được giải mã để tạo ra thông tin hướng và thông tin phần năng lượng.

Ở bước S1540, ma trận khử trộn trên mỗi dải được tạo ra từ thông tin hướng và thông tin phân năng lượng.

Ở bước S1550, dòng giảm mức được xử lý theo ma trận khử trộn để tạo ra dòng riêng.

Ở bước S1560, các tín hiệu đối tượng được trích từ dòng riêng và được điều chỉnh vị trí để tạo ra các tín hiệu đối tượng được điều chỉnh vị trí theo thông tin hướng và định dạng đầu ra mong muốn.

Ở bước S1570, các tín hiệu dư được trích từ dòng riêng và được xử lý để tạo ra các tín hiệu dư đã giải mã theo định dạng đầu ra mong muốn.

Cuối cùng, ở bước S1580, các tín hiệu đối tượng đã điều chỉnh vị trí và các tín hiệu dư đã giải mã được kết hợp để tạo ra cảnh âm thanh tái tạo.

Cần phải hiểu rằng các bước trên có thể được thực hiện theo thứ tự bất kỳ hoặc song song với nhau, miễn là thứ tự các bước đảm bảo rằng có sẵn thông tin đầu vào cần thiết cho mỗi bước.

Các phương pháp xử lý tín hiệu âm thanh không gian để tạo ra dạng biểu diễn nén của tín hiệu âm thanh không gian, cũng như các phương pháp xử lý dạng biểu diễn nén của tín hiệu âm thanh không gian để tạo ra dạng biểu diễn tái tạo của tín hiệu âm thanh không gian đã được mô tả ở trên. Ngoài ra, sáng chế cũng đề cập đến thiết bị để thực hiện các phương pháp này. Một ví dụ về thiết bị 1600 như vậy được minh họa bằng sơ đồ trên Fig.16. Thiết bị 1600 có thể bao gồm bộ xử lý 1610 (ví dụ, bộ xử lý trung tâm (central processing unit - CPU), bộ xử lý đồ họa (graphics processing unit - GPU), bộ xử lý tín hiệu số (digital signal processor - DSP), một hoặc nhiều mạch tích hợp chuyên dụng (application specific integrated circuit - ASIC), một hoặc nhiều mạch tích hợp tần số vô tuyến (radio-frequency integrated circuit - RFIC), hoặc kết hợp bất kỳ của chúng) và bộ nhớ 1620 được ghép nối với bộ xử lý 1610. Bộ xử lý có thể được làm thích ứng để thực hiện một số hoặc tất cả các bước của các phương pháp được mô tả trong sáng chế. Nếu thiết bị 1600 hoạt động như bộ mã hóa (ví dụ, bộ mã hóa cảnh), thì nó có thể nhận, dưới dạng đầu vào 1630, tín hiệu âm thanh không gian (ví dụ, cảnh âm thanh không gian), chẳng hạn. Sau đó, thiết bị 1600 có thể tạo ra, dưới dạng đầu ra 1640, dạng biểu diễn nén của tín hiệu âm thanh không gian. Nếu thiết bị 1600 hoạt động như bộ giải

mã (ví dụ, bộ giải mã cảnh), thì nó có thể nhận, dưới dạng đầu vào 1630, dạng biểu diễn nén. Sau đó, thiết bị có thể tạo ra, dưới dạng đầu ra 1640, cảnh âm thanh đã tái tạo.

Thiết bị 1600 có thể là máy chủ, máy khách, máy tính cá nhân (personal computer - PC), máy tính bảng, đầu thu giải mã tín hiệu (set-top box - STB), thiết bị kỹ thuật số hỗ trợ cá nhân (personal digital assistant - PDA), điện thoại di động, điện thoại thông minh, thiết bị web, bộ định tuyến mạng, bộ chuyển mạch hoặc cầu nối, hoặc máy bất kỳ có khả năng thực thi các lệnh (tuần tự hoặc theo cách khác) xác định các hành động mà thiết bị đó thực hiện. Hơn nữa, mặc dù chỉ một thiết bị 1600 được minh họa trên Fig.16, sáng chế đề cập đến tập hợp thiết bị bất kỳ thực thi riêng hoặc chung các lệnh để thực hiện một hoặc nhiều phương pháp bất kỳ được mô tả ở đây.

Sáng chế còn đề cập đến chương trình (ví dụ, chương trình máy tính) bao gồm các lệnh mà, khi được thực thi bởi bộ xử lý, khiến cho bộ xử lý thực hiện một số hoặc tất cả các bước của các phương pháp được mô tả ở đây.

Ngoài ra, sáng chế đề cập đến phương tiện lưu trữ đọc được bằng máy tính (hoặc được bằng máy) lưu trữ chương trình nói trên. Ở đây, thuật ngữ “phương tiện lưu trữ đọc được bằng máy tính” bao gồm, nhưng không giới hạn ở, kho dữ liệu ở dạng bộ nhớ trạng thái rắn, phương tiện quang học và phương tiện từ tính, chẳng hạn.

Xem xét cấu hình bổ sung

Trừ khi có quy định cụ thể khác, rõ ràng từ các phần mô tả sau đây, sẽ hiểu rằng xuyên suốt các nội dung trong bản mô tả này việc sử dụng các thuật ngữ như “xử lý,” “tính toán,” “tính,” “xác định,” “phân tích” hoặc thuật ngữ tương tự chỉ hoạt động và/hoặc các quy trình của máy tính hoặc hệ thống máy tính, hoặc các thiết bị máy tính điện tử tương tự, mà thao tác và/hoặc biến đổi dữ liệu được biểu diễn dưới dạng các đại lượng vật lý, như điện tử thành dữ liệu khác được biểu diễn tương tự dưới dạng các đại lượng vật lý.

Theo cách tương tự, thuật ngữ “bộ xử lý” có thể chỉ thiết bị hoặc phần bất kỳ của thiết bị mà xử lý dữ liệu điện tử, ví dụ, các thanh ghi và/hoặc bộ nhớ để biến đổi dữ liệu điện tử thành dữ liệu điện tử khác mà, ví dụ, có thể được lưu trữ trong các thanh ghi và/hoặc bộ nhớ. “Máy tính” hoặc “máy tính toán” hoặc “nền tảng tính toán” có thể bao gồm một hoặc nhiều bộ xử lý.

Các phương pháp được mô tả ở đây, theo một phương án ví dụ, thực hiện được bởi một hoặc nhiều bộ xử lý mà chấp nhận mã đọc được bằng máy tính (còn được gọi là đọc được bằng máy) chứa tập lệnh mà khi được thực thi bởi một hoặc nhiều trong số các bộ xử lý thực hiện ít nhất một trong các phương pháp mô tả ở đây. Bộ xử lý bất kỳ có khả năng thực thi tập lệnh (theo tuần tự hoặc theo cách khác) mà chỉ ra các hành động cần được thực hiện cũng được bao gồm. Do đó, một ví dụ là hệ thống xử lý điển hình mà bao gồm một hoặc nhiều bộ xử lý. Mỗi bộ xử lý có thể bao gồm một hoặc nhiều trong số CPU, đơn vị xử lý đồ họa, và đơn vị DSP lập trình được. Hệ thống xử lý còn có thể bao gồm hệ thống con bộ nhớ bao gồm RAM chính và/hoặc RAM tĩnh, và/hoặc ROM. Hệ thống con bus có thể được bao gồm để truyền thông giữa các thành phần này. Hệ thống xử lý còn có thể là hệ thống xử lý phân tán với các bộ xử lý được ghép nối bởi mạng. Nếu hệ thống xử lý cần có màn hiển thị, thì màn hiển thị như vậy có thể được bao gồm, ví dụ, màn hình tinh thể lỏng (liquid crystal display - LCD) hoặc màn hình ống tia catot (cathode ray tube - CRT). Nếu có yêu cầu nhập dữ liệu thủ công, hệ thống xử lý còn bao gồm thiết bị đầu vào như một hoặc nhiều đơn vị nhập chữ số như bàn phím, thiết bị điều khiển con trỏ như chuột, và v.v.. Hệ thống xử lý có thể còn bao gồm hệ thống lưu trữ như đơn vị ổ đĩa. Hệ thống xử lý theo một số cấu hình có thể bao gồm thiết bị đầu ra âm thanh, và thiết bị giao diện mạng. Do đó hệ thống con bộ nhớ bao gồm phương tiện mang đọc được bằng máy tính mà mang mã đọc được bằng máy tính (ví dụ, phần mềm) bao gồm tập lệnh để thực hiện, khi được thực thi bởi một hoặc nhiều bộ xử lý, một hoặc nhiều trong số các phương pháp được mô tả ở đây. Lưu ý rằng khi phương pháp bao gồm một vài phần tử, ví dụ, một vài bước, thì không có sự ngụ ý thứ tự các phần tử đó, trừ khi có quy định cụ thể. Phần mềm có thể nằm trong đĩa cứng, hoặc cũng có thể nằm, hoàn toàn hoặc ít nhất một phần, trong RAM và/hoặc trong bộ xử lý trong quá trình thực thi của nó bởi hệ thống máy tính. Do đó, bộ nhớ và bộ xử lý còn cấu thành phương tiện mang đọc được bằng máy tính mang mã đọc được bằng máy tính. Hơn nữa, phương tiện mang đọc được bằng máy tính có thể tạo ra, hoặc được bao gồm trong sản phẩm chương trình máy tính.

Theo các phương án ví dụ khác, một hoặc nhiều bộ xử lý hoạt động dưới dạng thiết bị độc lập hoặc có thể được kết nối, ví dụ, nối mạng với (các) bộ xử lý khác, theo phương án triển khai nối mạng, một hoặc nhiều bộ xử lý có thể hoạt động với khả năng

của máy chủ hoặc máy người dùng trong môi trường mạng máy chủ-người dùng, hoặc như máy ngang hàng trong môi trường mạng ngang hàng hoặc mạng phân tán. Một hoặc nhiều bộ xử lý có thể tạo ra máy tính cá nhân (personal computer - PC), máy tính bảng, thiết bị hỗ trợ số cá nhân (Personal Digital Assistant - PDA), điện thoại di động, thiết bị web, bộ định tuyến mạng, bộ chuyển mạch hoặc cầu nối, hoặc máy bất kỳ có khả năng thực thi tập lệnh (theo tuần tự hoặc theo cách khác) mà chỉ rõ các hành động phải thực hiện bởi máy móc đó.

Lưu ý rằng thuật ngữ “máy” cũng bao gồm tập hợp máy móc bất kỳ mà thực thi một cách riêng lẻ hoặc kết hợp tập (hoặc nhiều tập) lệnh để thực hiện một hoặc nhiều phương pháp bất kỳ trong số các phương pháp được mô tả ở đây.

Do đó, một phương án ví dụ của mỗi trong số các phương pháp được mô tả ở đây là ở dạng phương tiện mang đọc được bằng máy tính mang tập lệnh, ví dụ, chương trình máy tính để thực thi trên một hoặc nhiều bộ xử lý, ví dụ, một hoặc nhiều bộ xử lý mà là một phần của bố cục máy chủ web. Do đó, người có hiểu biết trung bình trong lĩnh vực này sẽ hiểu rằng, các phương án ví dụ của sáng chế có thể được thể hiện dưới dạng phương pháp, thiết bị như thiết bị chuyên dụng, thiết bị như hệ thống xử lý dữ liệu, hoặc phương tiện mang đọc được bằng máy tính, ví dụ, sản phẩm chương trình máy tính. Phương tiện mang đọc được bằng máy tính mang mã đọc được bằng máy tính bao gồm tập lệnh mà khi được thực thi trên một hoặc nhiều bộ xử lý khiến cho bộ xử lý hoặc các bộ xử lý thực hiện một phương pháp. Do đó, các khía cạnh của sáng chế có thể có dạng phương pháp, phương án ví dụ phần cứng hoàn toàn, phương án ví dụ phần mềm hoàn toàn hoặc phương án ví dụ kết hợp các khía cạnh phần cứng và phần mềm. Hơn nữa, sáng chế có thể có dạng phương tiện mang (ví dụ, sản phẩm chương trình máy tính trên phương tiện lưu trữ đọc được bằng máy tính) mang mã chương trình đọc được bằng máy tính được bao gồm trong phương tiện.

Phần mềm có thể còn được truyền hoặc nhận qua mạng qua thiết bị giao diện mạng. Mặc dù phương tiện mang là một phương tiện đơn theo phương án ví dụ, thuật ngữ “phương tiện mang” nên được hiểu là bao gồm một phương tiện hoặc nhiều phương tiện (ví dụ, cơ sở dữ liệu tập trung hoặc phân tán, và/hoặc bộ nhớ cache và máy chủ kết hợp) mà lưu trữ một hoặc nhiều tập lệnh. Thuật ngữ “phương tiện mang” sẽ được hiểu là bao gồm phương tiện bất kỳ mà có khả năng lưu trữ, mã hóa hoặc mang tập lệnh để

thực thi bởi một hoặc nhiều bộ xử lý và khiến cho một hoặc nhiều bộ xử lý này thực hiện một hoặc nhiều phương pháp bất kỳ trong số các phương pháp của sáng chế. Phương tiện mang có thể có nhiều dạng, bao gồm nhưng không chỉ giới hạn ở, phương tiện bất biến, phương tiện khả biến, và phương tiện truyền. Phương tiện bất biến bao gồm, chẳng hạn, đĩa quang, đĩa từ, và đĩa từ-quang. Phương tiện khả biến bao gồm bộ nhớ động, như bộ nhớ chính. Phương tiện truyền bao gồm cáp đồng trục, dây đồng và cáp quang, bao gồm các dây dẫn chứa hệ thống con bus. Phương tiện truyền cũng có thể có dạng của các sóng âm hoặc ánh sáng, chẳng hạn như các sóng sinh ra trong quá trình trao đổi dữ liệu hồng ngoại và sóng vô tuyến. Chẳng hạn, thuật ngữ “phương tiện mang” do đó sẽ được hiểu là bao gồm, nhưng không chỉ giới hạn ở, các bộ nhớ trạng thái rắn, sản phẩm máy tính được bao gồm trong phương tiện quang và từ; phương tiện mang tín hiệu truyền phát hiện được bởi ít nhất một bộ xử lý hoặc một hoặc nhiều bộ xử lý và biểu diễn tập lệnh mà, khi được thực thi, sẽ thực hiện phương pháp; và phương tiện truyền trong mạng mang tín hiệu truyền phát hiện được bởi ít nhất một bộ xử lý của một hoặc nhiều bộ xử lý và biểu diễn tập lệnh.

Sẽ hiểu rằng các bước của các phương pháp được bàn luận được thực hiện trong một phương án ví dụ bởi bộ xử lý (hoặc các bộ xử lý) thích hợp của hệ thống xử lý (ví dụ, máy tính) thực thi các lệnh (mã đọc được bằng máy tính) lưu trữ trong bộ nhớ. Cũng sẽ hiểu rằng sáng chế không bị giới hạn ở phương án thực hiện hoặc kỹ thuật lập trình cụ thể và rằng sáng chế có thể được thực hiện nhờ sử dụng các kỹ thuật thích hợp bất kỳ để thực hiện chức năng được mô tả ở đây. Sáng chế không bị giới hạn ở ngôn ngữ lập trình hoặc hệ điều hành cụ thể bất kỳ.

Việc nói đến “một phương án ví dụ”, “một số phương án ví dụ” hoặc “phương án ví dụ” trên khắp bản mô tả nghĩa là một dấu hiệu, cấu trúc hoặc đặc tính cụ thể được mô tả liên quan đến phương án ví dụ được bao gồm trong ít nhất một phương án ví dụ của sáng chế. Vì thế, các cụm từ “theo một phương án ví dụ”, “theo một số phương án ví dụ” hoặc “theo phương án ví dụ” ở khắp bản mô tả không nhất thiết đều nói đến cùng một phương án ví dụ. Hơn nữa, các dấu hiệu, cấu trúc hoặc đặc tính cụ thể này có thể được kết hợp theo cách thức phù hợp bất kỳ, là hiển nhiên đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này từ bản mô tả này, theo một hoặc nhiều phương án ví dụ.

Như được sử dụng trong bản mô tả này, trừ khi có quy định theo cách khác, việc sử dụng các tính từ thông thường “thứ nhất”, “thứ hai”, “thứ ba”, v.v., để mô tả một đối tượng chung, chỉ cho biết là các trường hợp khác của các đối tượng tương tự đang được nói đến và không được dự định để ngụ ý rằng các đối tượng được mô tả phải theo trình tự cho trước, theo thời gian, không gian, theo thứ hạng hoặc theo cách thức bất kỳ khác.

Trong yêu cầu bảo hộ dưới đây và phần mô tả này, thuật ngữ bất kỳ trong số các thuật ngữ gồm, gồm có hoặc bao gồm là thuật ngữ mở có nghĩa là bao gồm ít nhất một thành phần/dấu hiệu theo sau, nhưng không loại trừ các thành phần/dấu hiệu khác. Do đó, thuật ngữ bao gồm, khi được dùng trong phần yêu cầu bảo hộ, không được hiểu là giới hạn ở các phương tiện hoặc thành phần hoặc bước được liệt kê sau đó. Ví dụ, phạm vi biểu diễn thiết bị bao gồm A và B không bị giới hạn ở các thiết bị chỉ bao gồm các phần tử A và B. Một thuật ngữ bất kỳ trong số các thuật ngữ chứa hoặc bao gồm hoặc gồm như được dùng ở đây cũng là thuật ngữ mở cũng có nghĩa là gồm ít nhất các thành phần/dấu hiệu theo sau thuật ngữ này, nhưng không loại trừ các thành phần/dấu hiệu khác. Do đó, bao gồm là từ đồng nghĩa với và có nghĩa là gồm.

Cần phải hiểu rằng trong phần mô tả trên về các phương án làm ví dụ theo sáng chế, các dấu hiệu khác nhau của sáng chế đôi khi được nhóm lại với nhau trong một phương án, hình vẽ, hoặc phần mô tả làm ví dụ của nó nhằm mục đích đơn giản hóa phần mô tả và giúp hiểu rõ một hoặc nhiều khía cạnh sáng tạo khác nhau. Tuy nhiên, cách thể hiện này không được hiểu là phản ánh dự định rằng các yêu cầu bảo hộ nhiều dấu hiệu hơn là các dấu hiệu được đề cập rõ ràng trong mỗi điểm yêu cầu bảo hộ. Đúng hơn là, như các điểm yêu cầu bảo hộ sau đây phản ánh, các khía cạnh của sáng chế nằm trong ít hơn tất cả các dấu hiệu của một phương án ví dụ được mô tả trên đây. Do vậy, yêu cầu bảo hộ sau phần Mô tả sáng chế được đưa vào phần mô tả này, với mỗi điểm yêu cầu bảo hộ là một phương án riêng biệt của sáng chế.

Hơn nữa, mặc dù một số phương án làm ví dụ được mô tả trong sáng chế gồm một số dấu hiệu này mà không phải là các dấu hiệu khác trong các phương án làm ví dụ khác, tổ hợp các dấu hiệu của các phương án làm ví dụ khác nhau được hiểu là nằm trong phạm vi của sáng chế, và tạo ra các phương án ví dụ khác nhau, như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực này. Ví dụ, trong phần yêu cầu bảo hộ

dưới đây, phương án bất kỳ trong số các phương án làm ví dụ được yêu cầu bảo hộ có thể được dùng trong tổ hợp bất kỳ.

Trong phần mô tả này, nhiều chi tiết cụ thể được nêu. Tuy nhiên, cần hiểu rằng các phương án ví dụ theo sáng chế có thể được thực hiện mà không cần tới các chi tiết cụ thể này. Trong các trường hợp khác, các phương pháp, cấu trúc và kỹ thuật đã biết không được thể hiện một cách chi tiết để không làm rối rắm phần mô tả này.

Do đó, mặc dù các phương án được cho là tốt nhất của sáng chế đã được mô tả, người có hiểu biết trung bình trong lĩnh vực này sẽ hiểu rằng các cải biến khác và cải biến thêm có thể được thực hiện mà không vượt ra ngoài phạm vi của sáng chế, và yêu cầu bảo hộ dự tính bao hàm tất cả các thay đổi và cải biến trong phạm vi của sáng chế này. Ví dụ, công thức bất kỳ được đưa ra trên đây chỉ là đại diện của các thủ tục mà có thể được dùng. Chức năng có thể được bổ sung hoặc loại bỏ khỏi sơ đồ khối và các hoạt động có thể thay thế cho nhau giữa các khối chức năng. Các bước có thể được bổ sung hoặc loại bỏ ra khỏi các phương pháp được mô tả nằm trong phạm vi của sáng chế

Các khía cạnh, phương án và phương án thực hiện làm ví dụ khác của sáng chế sẽ trở nên rõ ràng từ các phương án làm ví dụ được liệt kê (enumerated example embodiment - EEE) được liệt kê bên dưới.

EEE 1 đề cập đến phương pháp biểu diễn cảnh âm thanh không gian dưới dạng cảnh âm thanh không gian thu gọn bao gồm dòng hỗn hợp âm thanh và dòng siêu dữ liệu có hướng, trong đó dòng hỗn hợp âm thanh bao gồm một hoặc nhiều tín hiệu âm thanh, và trong đó dòng siêu dữ liệu có hướng bao gồm chuỗi thời gian của các khối siêu dữ liệu có hướng với mỗi trong số các khối siêu dữ liệu có hướng kết hợp với đoạn thời gian tương ứng trong các tín hiệu âm thanh, và trong đó cảnh âm thanh không gian bao gồm một hoặc nhiều phần tử âm thanh có hướng mà mỗi phần tử âm thanh có hướng được kết hợp với hướng đến tương ứng, và trong đó mỗi trong số các khối siêu dữ liệu có hướng chứa: (a) thông tin hướng biểu thị các hướng đến cho mỗi trong số các phần tử âm thanh có hướng, và (b) thông tin về dải năng lượng biểu thị năng lượng trong mỗi trong số các phần tử âm thanh có hướng, liên quan đến năng lượng trong đoạn thời gian tương ứng trong các tín hiệu âm thanh, cho mỗi trong số các phần tử âm thanh có hướng và cho mỗi dải con trong tập hợp hai hoặc nhiều dải con.

EEE 2 đề cập đến phương pháp theo EEE 1, trong đó (a) Thông tin phần dải năng lượng biểu thị các thuộc tính của cảnh âm thanh không gian trong mỗi trong số các dải con, và (b) đối với ít nhất một hướng đến, dữ liệu được bao gồm trong Thông tin hướng biểu thị các thuộc tính của cảnh âm thanh không gian trong một cụm gồm hai hoặc nhiều dải con.

EEE 3 đề cập đến phương pháp xử lý cảnh âm thanh không gian thu gọn bao gồm dòng hỗn hợp âm thanh và dòng siêu dữ liệu có hướng, để tạo ra dòng âm thanh không gian riêng bao gồm tập hợp một hoặc nhiều tín hiệu đối tượng âm thanh và dòng dư, trong đó dòng hỗn hợp âm thanh bao gồm một hoặc nhiều tín hiệu âm thanh và trong đó dòng siêu dữ liệu có hướng bao gồm chuỗi thời gian của các khối siêu dữ liệu có hướng với mỗi trong số các khối siêu dữ liệu có hướng được kết hợp với đoạn thời gian tương ứng trong các tín hiệu âm thanh, trong đó đối với mỗi trong số nhiều dải con, phương pháp bao gồm: (a) xác định các hệ số của ma trận khử trộn từ Thông tin hướng và thông tin Phần Dải Năng Lượng có trong dòng siêu dữ liệu có hướng, và (b) trộn, sử dụng ma trận khử trộn, dòng hỗn hợp âm thanh để tạo ra dòng âm thanh không gian riêng.

EEE 4 đề cập đến phương pháp theo EEE 3, trong đó mỗi trong số các khối siêu dữ liệu có hướng chứa: (a) thông tin hướng chỉ báo hướng đến cho mỗi trong số các phần tử âm thanh có hướng, và (b) Thông tin Phần Dải Năng Lượng chỉ báo năng lượng trong mỗi trong số các phần tử âm thanh có hướng, liên quan đến năng lượng trong đoạn thời gian tương ứng trong các tín hiệu âm thanh, cho mỗi trong số các phần tử âm thanh có hướng và cho mỗi dải con trong tập hợp hai hoặc nhiều dải con.

EEE 5 đề cập đến phương pháp theo EEE 3, trong đó (a) đối với mỗi trong số các khối siêu dữ liệu có hướng, thông tin hướng và thông tin phần dải năng lượng được dùng để tạo ra ma trận, S , biểu diễn hiệp phương sai gần đúng của dòng âm thanh không gian riêng, và (a) thông tin phần dải năng lượng được dùng để tạo ra ma trận, E , biểu diễn ma trận trộn lại mà xác định việc chuyển đổi dòng âm thanh không gian riêng thành dòng hỗn hợp âm thanh, và (b) ma trận khử trộn, U , được tính toán theo $U = S \times E^* \times (E \times S \times E^*)^{-1}$.

EEE 6 đề cập đến phương pháp theo EEE 5, trong đó ma trận, S , là ma trận đường chéo.

EEE 7 đề cập đến phương pháp theo EEE 3, trong đó (a) dòng dư được xử lý để tạo ra dòng dư tái tạo, (b) mỗi trong số các tín hiệu đối tượng âm thanh được xử lý để tạo ra dòng đối tượng tái tạo tương ứng, và (c) dòng dư tái tạo và mỗi trong số các dòng đối tượng tái tạo được kết hợp để tạo ra Tín hiệu Âm thanh Tái tạo, trong đó Tín hiệu Âm thanh Tái tạo chứa các phần tử âm thanh có hướng theo cảnh âm thanh không gian thu gọn.

EEE 8 đề cập đến phương pháp theo EEE 7, trong đó Tín hiệu Âm thanh Tái tạo bao gồm hai tín hiệu để biểu diễn cho người nghe qua bộ chuyển đổi ở hoặc gần mỗi tai để cung cấp trải nghiệm hai tai về cảnh âm thanh không gian gồm các phần tử âm thanh có hướng theo cảnh âm thanh không gian thu gọn.

EEE 9 đề cập đến phương pháp theo EEE 7, trong đó Tín hiệu Âm thanh Tái tạo bao gồm một số tín hiệu biểu diễn cảnh âm thanh không gian dưới dạng các hàm điều chỉnh vị trí điều hòa cầu.

EEE 10 đề cập đến phương pháp xử lý cảnh âm thanh không gian để tạo ra cảnh âm thanh không gian thu gọn bao gồm dòng hỗn hợp âm thanh và dòng siêu dữ liệu có hướng, trong đó cảnh âm thanh không gian bao gồm một hoặc nhiều phần tử âm thanh có hướng mà mỗi phần tử âm thanh có hướng được kết hợp với hướng đến tương ứng, và trong đó dòng siêu dữ liệu có hướng bao gồm chuỗi thời gian của các khối siêu dữ liệu có hướng với mỗi trong số các khối siêu dữ liệu có hướng được kết hợp với đoạn thời gian tương ứng trong các tín hiệu âm thanh, phương pháp này bao gồm: (a) phương tiện xác định hướng đến cho một hoặc nhiều phần tử âm thanh có hướng, từ việc phân tích cảnh âm thanh không gian, (b) phương tiện xác định phần nào của tổng năng lượng trong cảnh không gian được đóng góp bởi năng lượng trong mỗi trong số các phần tử âm thanh có hướng, và (c) phương tiện xử lý cảnh âm thanh không gian để tạo ra dòng hỗn hợp âm thanh.

YÊU CẦU BẢO HỘ

1. Phương pháp (1300) xử lý dạng biểu diễn nén của tín hiệu âm thanh không gian để tạo ra dạng biểu diễn tái tạo của tín hiệu âm thanh không gian, trong đó dạng biểu diễn nén bao gồm tín hiệu âm thanh dựa trên kênh với số lượng kênh xác định trước và siêu dữ liệu, siêu dữ liệu bao gồm thông tin hướng và thông tin năng lượng, với thông tin hướng bao gồm các chỉ báo về các hướng đến của một hoặc nhiều phần tử âm thanh trong cảnh âm thanh và thông tin năng lượng bao gồm, đối với ít nhất một dải tần số con, các chỉ báo tương ứng về công suất tín hiệu kết hợp với các hướng đến, phương pháp bao gồm các bước:

tạo ra (S1310) tín hiệu âm thanh của một hoặc nhiều phần tử âm thanh dựa vào tín hiệu âm thanh dựa trên kênh, thông tin hướng, và thông tin năng lượng; và

tạo ra (S1320) tín hiệu âm thanh dư mà một hoặc nhiều phần tử âm thanh về cơ bản là không có ở đó, dựa vào tín hiệu âm thanh dựa trên kênh, thông tin hướng, và thông tin năng lượng.

trong đó bước tạo ra các tín hiệu âm thanh gồm một hoặc nhiều phần tử âm thanh bao gồm:

xác định (1400) các hệ số của ma trận trộn ngược M để ánh xạ tín hiệu âm thanh dựa trên kênh lên dạng biểu diễn trung gian bao gồm tín hiệu âm thanh dư và các tín hiệu âm thanh của một hoặc nhiều phần tử âm thanh, dựa vào thông tin hướng và thông tin năng lượng.

trong đó bước xác định các hệ số của ma trận trộn ngược M bao gồm:

xác định (S1410), đối với mỗi trong số một hoặc nhiều phần tử âm thanh, vector điều chỉnh vị trí $Pan_{down}(dir)$ để điều chỉnh vị trí phần tử âm thanh đến các kênh của tín hiệu âm thanh dựa trên kênh, dựa vào hướng đến dir của phần tử âm thanh;

xác định (S1420) ma trận trộn E mà được dùng để ánh xạ tín hiệu âm thanh dư và các tín hiệu âm thanh của một hoặc nhiều phần tử âm thanh lên các kênh của tín hiệu âm thanh dựa trên kênh, dựa vào các vector điều chỉnh vị trí xác định được;

xác định (S1430) ma trận hiệp phương sai S cho dạng biểu diễn trung gian dựa vào thông tin năng lượng; và

xác định (S1440) các hệ số của ma trận trộn ngược M dựa vào ma trận trộn E và ma trận hiệp phương sai S .

2. Phương pháp theo điểm 1, trong đó chỉ báo về công suất tín hiệu liên quan đến, đối với hướng đến cho trước và dải tần số con cho trước, một phần công suất tín hiệu trong dải tần số con cho trước đối với hướng đến cho trước liên quan đến tổng công suất tín hiệu trong dải tần số con cho trước.

3. Phương pháp theo điểm 1 hoặc 2, trong đó thông tin năng lượng bao gồm các chỉ báo về công suất tín hiệu cho mỗi trong số nhiều dải tần số con.

4. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 3, phương pháp này còn bao gồm các bước:

điều chỉnh (S1330) vị trí các tín hiệu âm thanh của một hoặc nhiều phần tử âm thanh đến tập hợp các kênh của định dạng âm thanh đầu ra; và

tạo ra (S1340) tín hiệu âm thanh đa kênh tái tạo trong định dạng âm thanh đầu ra dựa vào một hoặc nhiều phần tử âm thanh được điều chỉnh vị trí và tín hiệu dư.

5. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 4, trong đó ma trận trộn E được xác định theo

$$E = (I_N | Pan_{down}(dir_1) | \dots | Pan_{down}(dir_P))$$

trong đó I_N là ma trận đơn vị $N \times N$, với N biểu thị số lượng kênh của tín hiệu dựa trên kênh, $Pan_{down}(dir_p)$ là vectơ điều chỉnh vị trí cho phần tử âm thanh thứ p có hướng đến kết hợp dir_p mà sẽ điều chỉnh vị trí phần tử âm thanh thứ p đến N kênh của tín hiệu dựa trên kênh với $p = 1, \dots, P$ biểu thị phần tử âm thanh tương ứng trong số một hoặc nhiều phần tử âm thanh và P biểu thị tổng số của một hoặc nhiều phần tử âm thanh.

6. Phương pháp theo điểm 5, trong đó ma trận hiệp phương sai S được xác định là ma trận đường chéo theo

$$\{S\}_{n,n} = \text{rms}(Pan_{down})_n \left(1 - \sum_{p=1}^P e_p \right)$$

với $1 \leq n \leq N$, và

$$\{S\}_{N+p,N+p} = e_p$$

với $1 \leq p \leq P$,

trong đó e_p là công suất tín hiệu kết hợp với hướng đến của phần tử âm thanh thứ p .

7. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 6, trong đó bước xác định các hệ số của ma trận trộn ngược dựa vào ma trận trộn và ma trận hiệp phương sai bao gồm bước xác định giá trị giả ngược dựa vào ma trận trộn và ma trận hiệp phương sai.

8. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 7, trong đó ma trận trộn ngược M được xác định theo

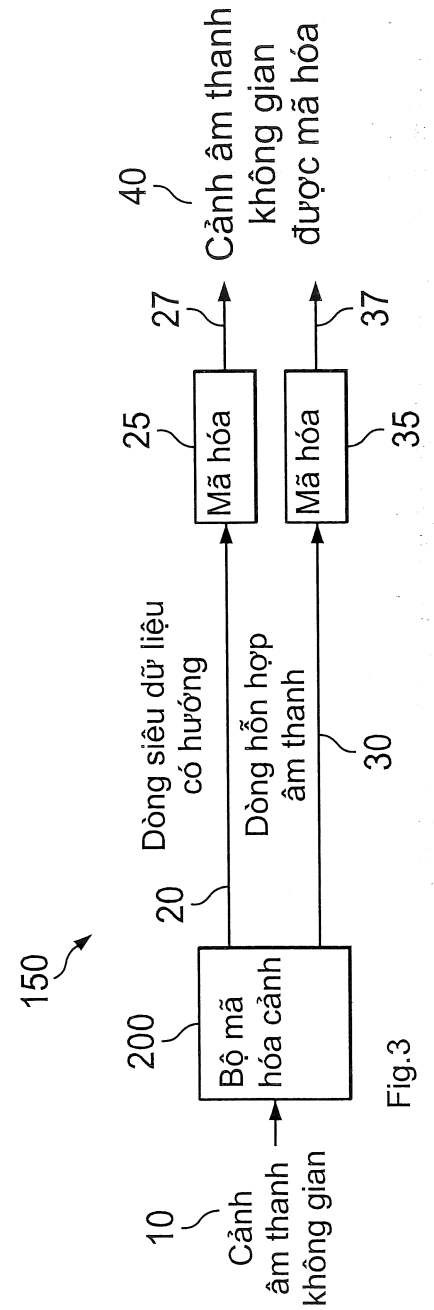
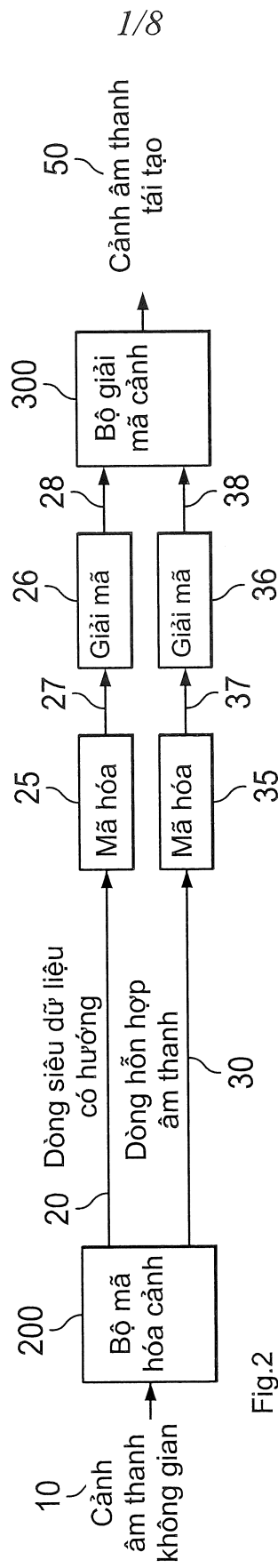
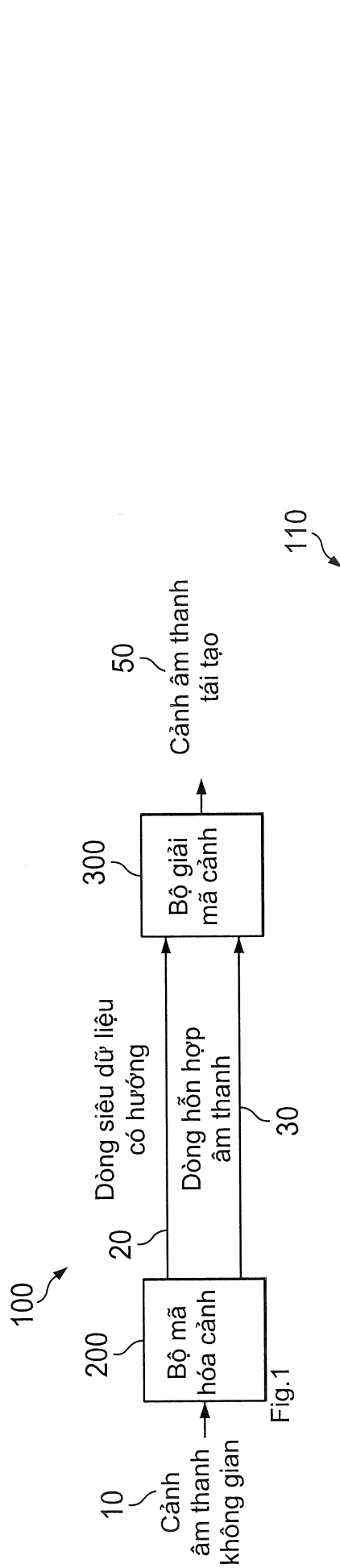
$$M = S \times E^* \times (E \times S \times E^*)^{-1}$$

trong đó “ $\times$ ” biểu thị phép nhân ma trận và “ $*$ ” biểu thị phép chuyển vị liên hợp của ma trận.

9. Phương pháp theo điểm bất kỳ trong số các điểm đã nêu, trong đó tín hiệu âm thanh dựa trên kênh là tín hiệu Ambisonics bậc nhất.

10. Phương tiện lưu trữ đọc được bằng máy tính lưu trữ chương trình bao gồm các lệnh mà, khi thực thi bằng bộ xử lý, khiến cho bộ xử lý thực hiện tất cả các bước của phương pháp theo điểm bất kỳ trong số các điểm trước đó.

11. Thiết bị xử lý dạng biểu diễn nén của tín hiệu âm thanh không gian, bao gồm bộ xử lý và bộ nhớ được ghép nối với bộ xử lý, trong đó bộ xử lý được làm thích ứng để thực hiện tất cả các bước của phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 9.



2/8

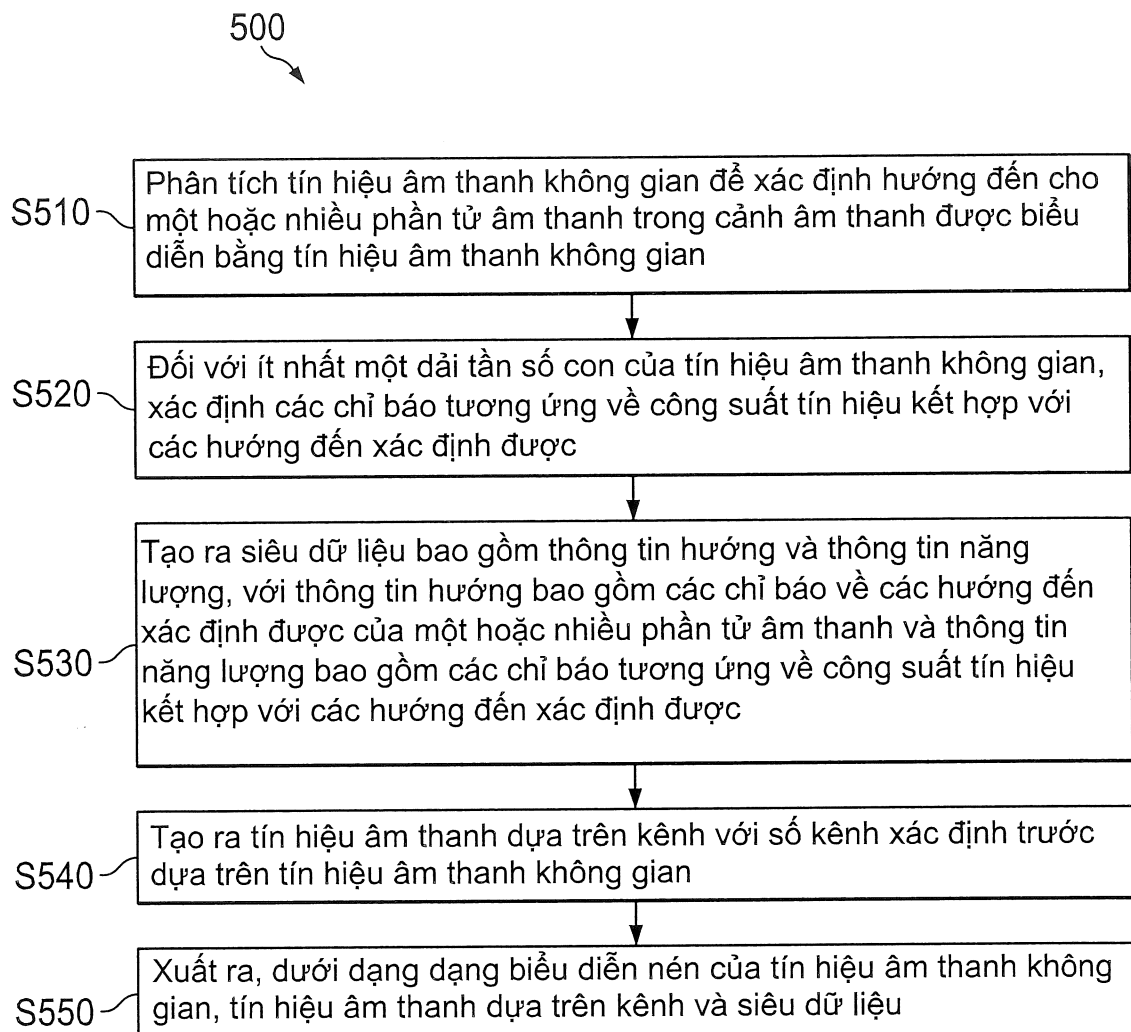
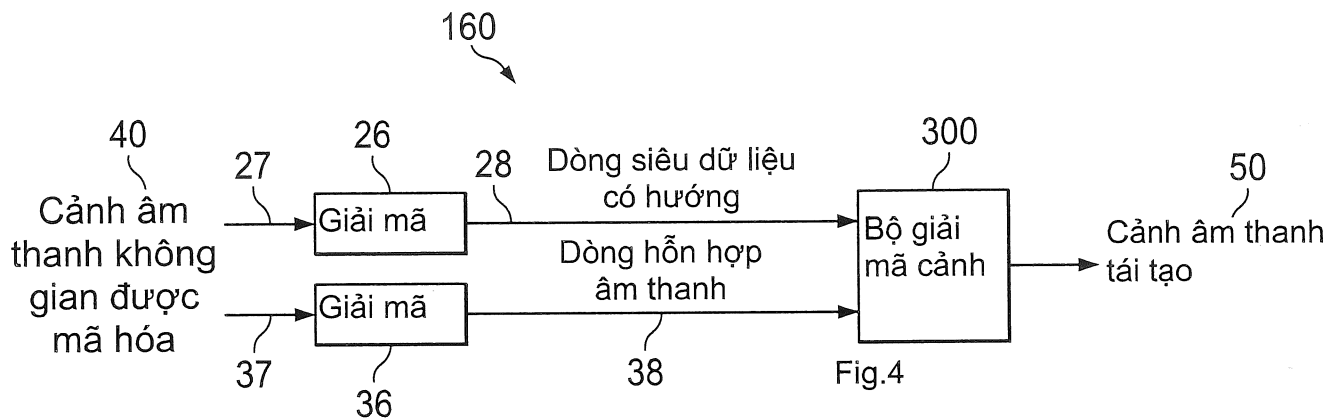


Fig.5

3/8

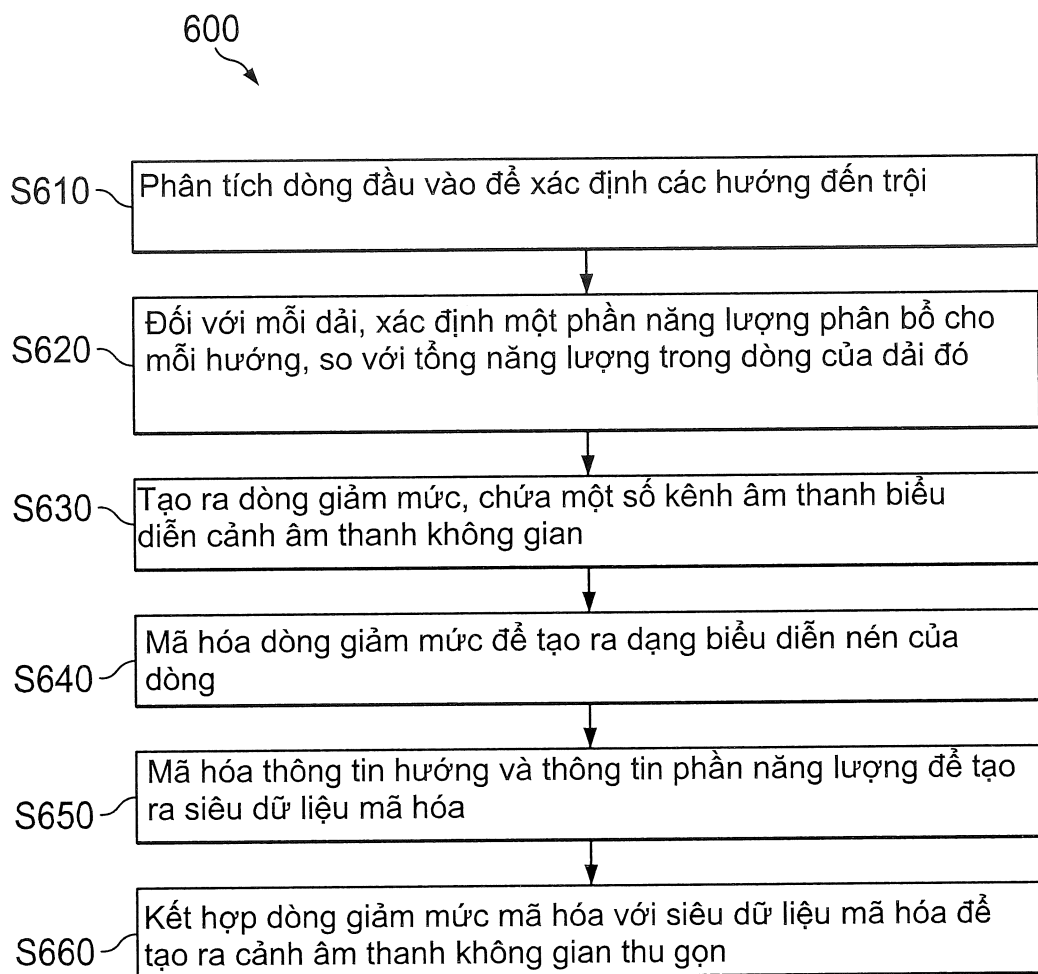
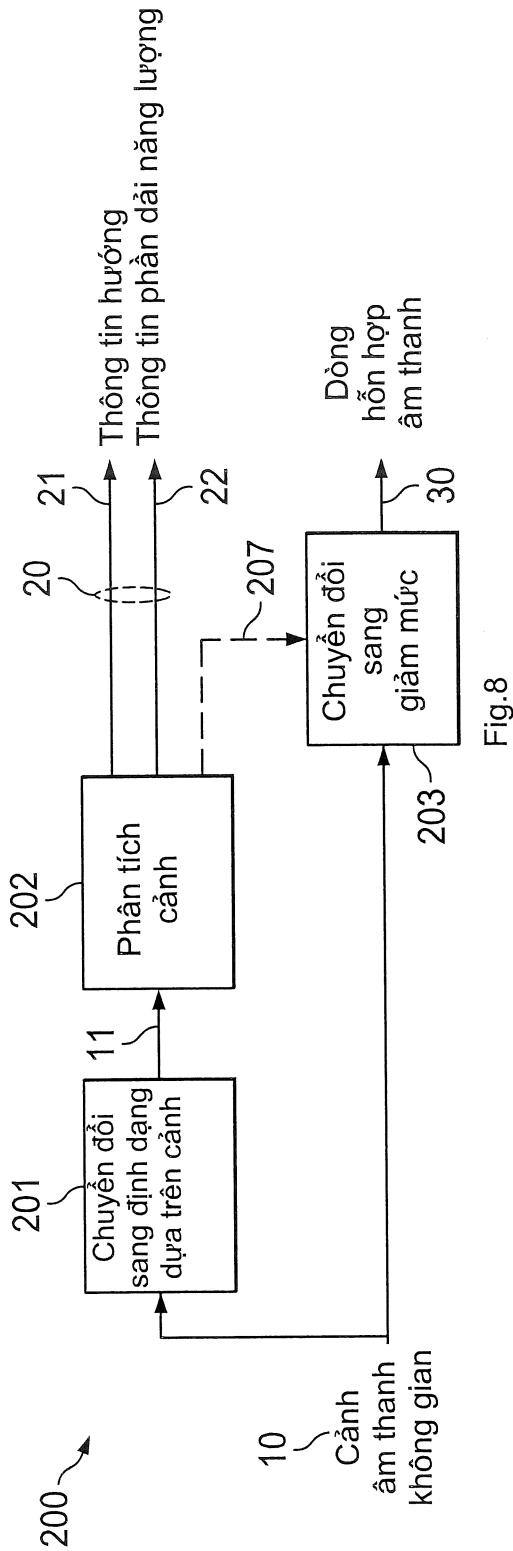
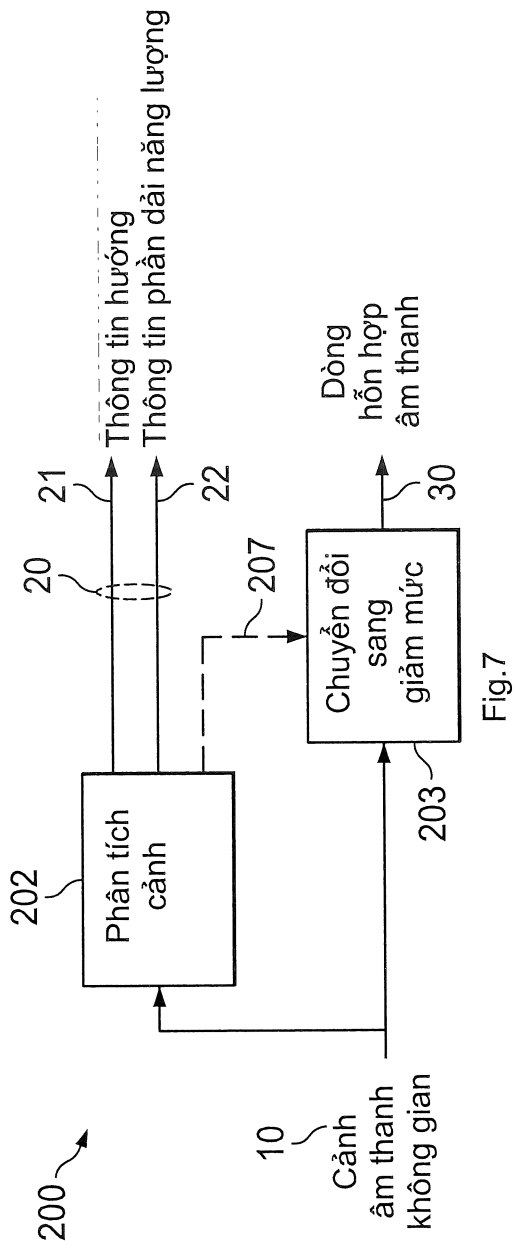


Fig.6



5/8

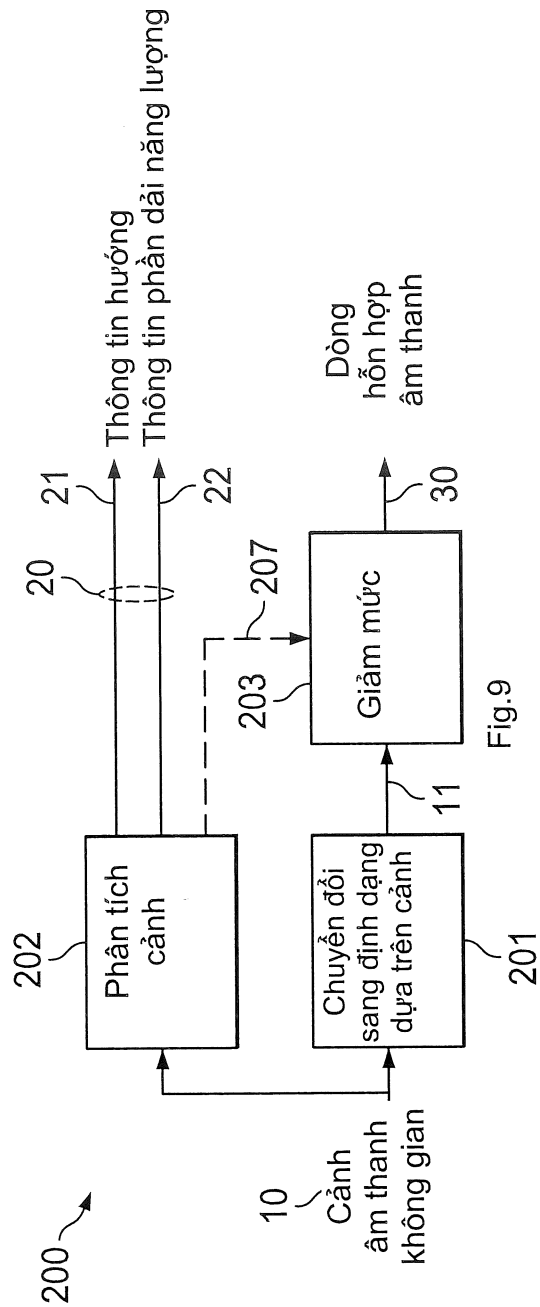


Fig. 9

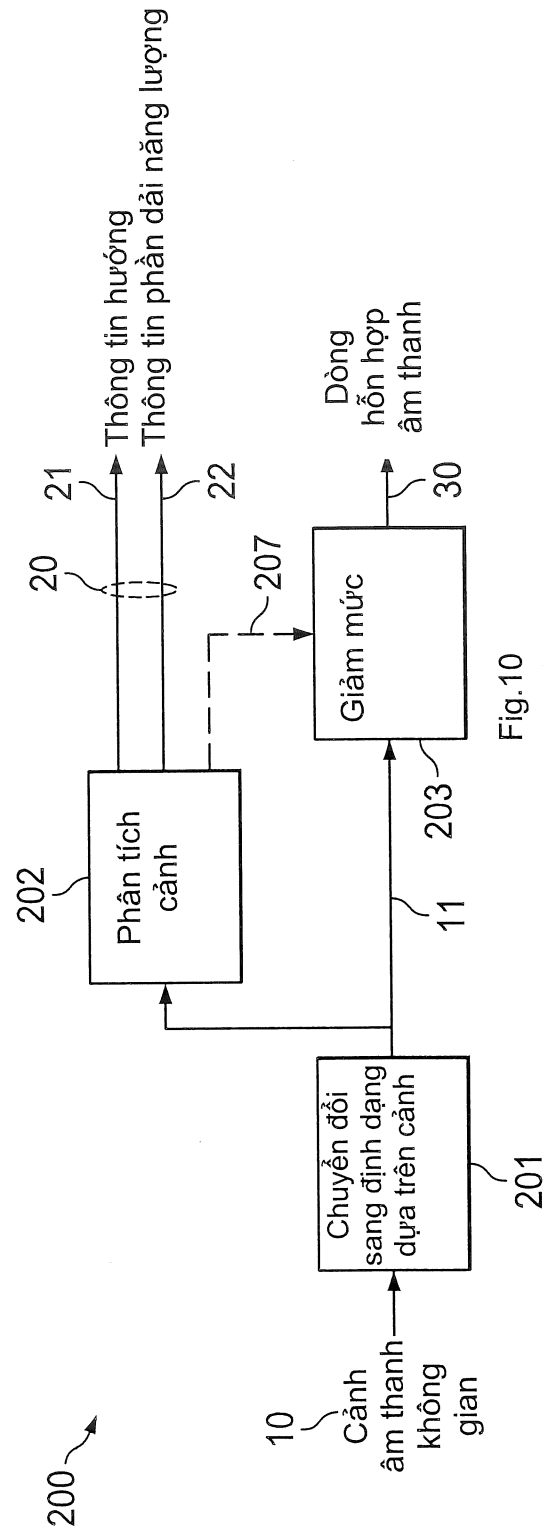
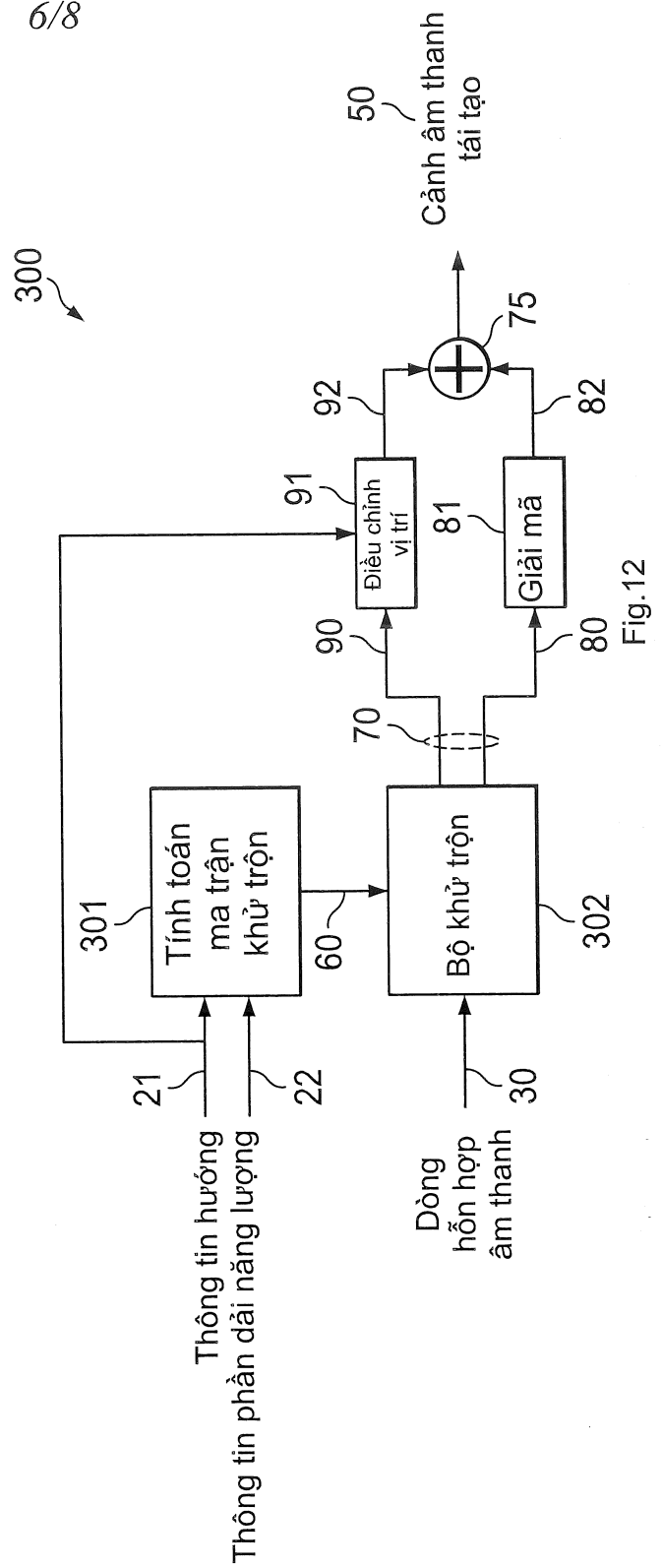
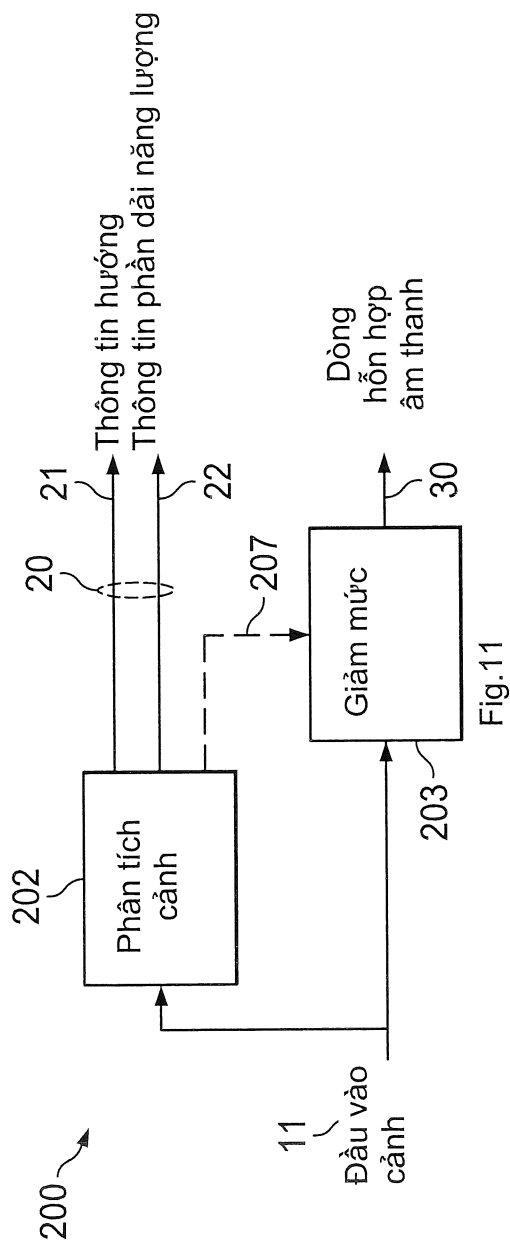
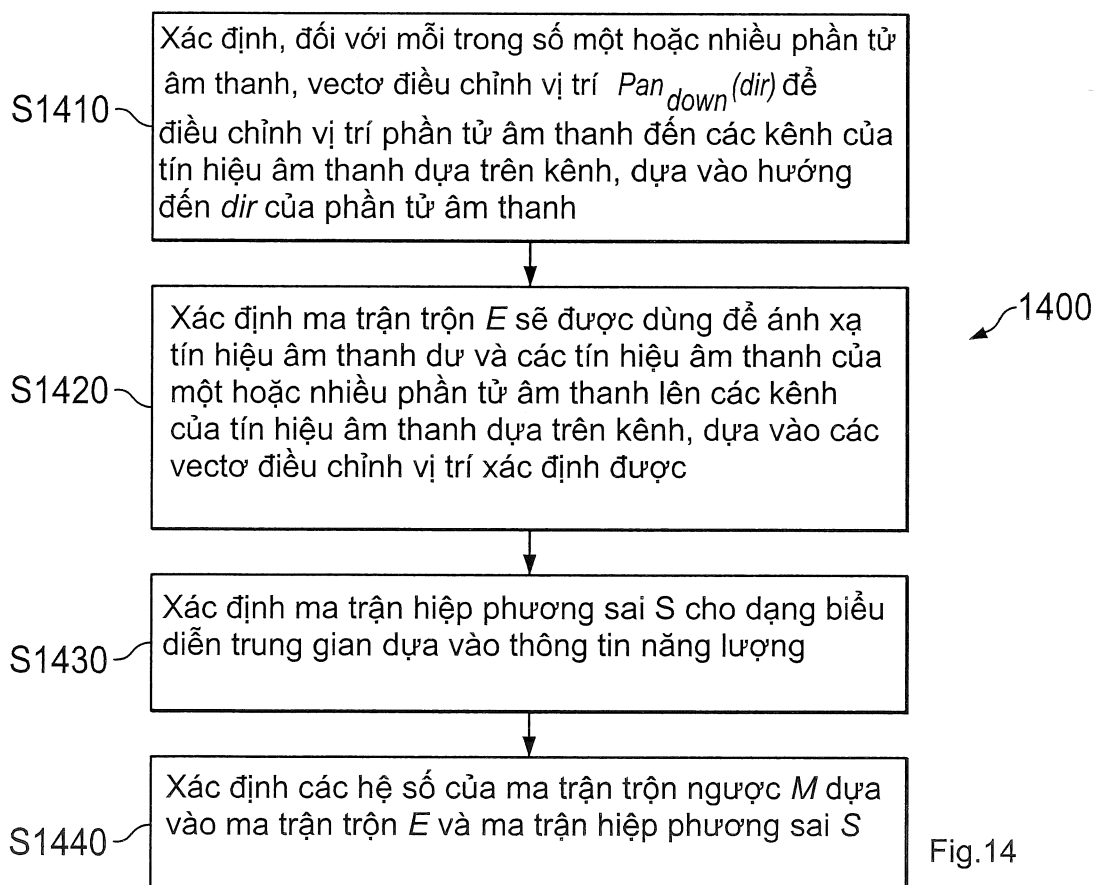
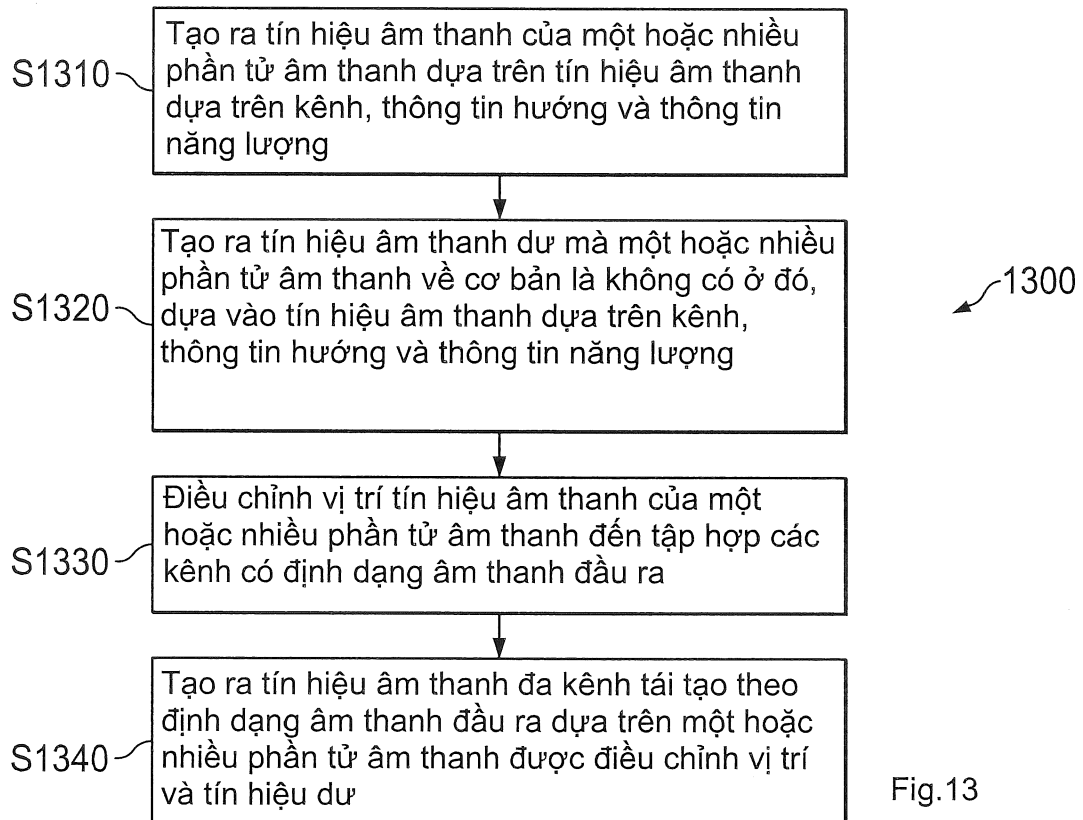


Fig. 10

6/8



7/8



8/8

