



(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)



1-0049180

(51)^{2023.01} H04S 3/00

(13) B

(21) 1-2023-06674

(22) 17/11/2016

(62) 1-2018-02535

(86) PCT/US2016/062497 17/11/2016

(87) WO 2017/087650 A1 26/05/2017

(30) 62/256,462 17/11/2015 US; 15199854.9 14/12/2015 EP

(45) 25/07/2025 448

(43) 25/03/2024 432

(73) 1. DOLBY LABORATORIES LICENSING CORPORATION (US)

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

2. DOLBY INTERNATIONAL AB (NL)

Apollo Building 3E Herikerbergweg 1-35 1101 CN Amsterdam Zuidoost (NL)

(72) BREEBAART, Dirk Jeroen (NL); COOPER, David Matthew (AU); DAVIS, Mark F. (US); MCGRATH, David S. (AU); KJOERLING, Kristofer (SE); MUNDT, Harald (DE); WILSON, Rhonda J. (GB).

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

(54) HỆ THỐNG ĐỂ MÃ HÓA VÀ GIẢI MÃ ÂM THANH ĐẦU VÀO DỰA TRÊN KÊNH HOẶC ĐỐI TƯỢNG VÀ PHƯƠNG TIỆN LƯU TRỮ BẤT BIẾN ĐỌC ĐƯỢC BẰNG MÁY TÍNH

(21) 1-2023-06674

(57) Sáng chế đề cập đến phương pháp mã hóa âm thanh đầu vào dựa trên kênh hoặc đối tượng để phát lại, phương pháp này bao gồm các bước: (a) ban đầu kết xuất âm thanh đầu vào dựa trên kênh hoặc đối tượng thành dạng trình diễn đầu ra ban đầu; (b) xác định giá trị ước tính của thành phần âm thanh nổi trội từ âm thanh đầu vào dựa trên kênh hoặc đối tượng và xác định một chuỗi hệ số gia trọng cho thành phần âm thanh nổi trội để ánh xạ dạng trình diễn đầu ra ban đầu thành thành phần âm thanh nổi trội; (c) xác định giá trị ước tính của hướng hoặc vị trí của thành phần âm thanh nổi trội; và (d) mã hóa dạng trình diễn đầu ra ban đầu, các hệ số gia trọng cho thành phần âm thanh nổi trội, hướng hoặc vị trí của thành phần âm thanh nổi trội dưới dạng tín hiệu mã hóa để phát lại. Sáng chế còn đề cập đến phương pháp giải mã tín hiệu âm thanh mã hóa, và giải mã và tái tạo dòng âm thanh, bộ máy và phương tiện lưu trữ đọc được bằng máy tính để thực hiện phương pháp này.

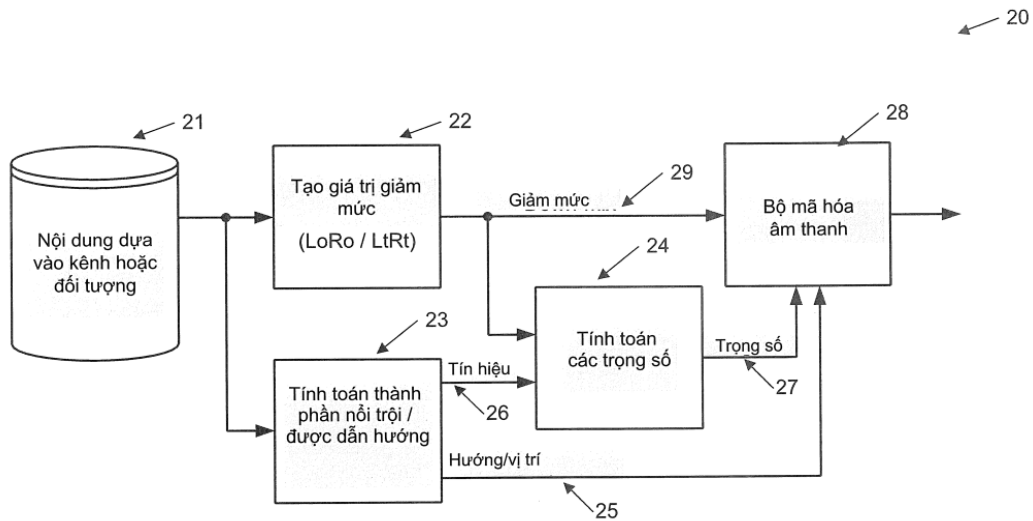


FIG. 2

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến hệ thống và phương pháp cho dạng đầu ra âm thanh lập thể hai tai tham số được cải thiện khi sử dụng tùy ý kỹ thuật theo dõi chuyển động của đầu.

Tình trạng kỹ thuật của sáng chế

Mọi sự thảo luận về giải pháp kỹ thuật trong phần tình trạng kỹ thuật trong toàn bộ bản mô tả này đều không được coi là thừa nhận rằng giải pháp kỹ thuật đó được biết đến rộng rãi hoặc tạo thành một phần của kiến thức chung thông thường trong lĩnh vực này.

Việc tạo nội dung, mã hóa, phân bố và tái tạo nội dung âm thanh thường được dựa vào kênh. Tức là, một hệ thống phát lại đích cụ thể được thiết kế cho nội dung trong toàn bộ hệ sinh thái nội dung. Ví dụ về các hệ thống phát lại đích như vậy là các hệ thống âm thanh đơn kênh (mono), âm thanh nổi (stereo), 5.1, 7.1, 7.1.4, và tương tự.

Nếu nội dung cần được tái tạo trên hệ hống phát lại khác với hệ thống được dự định, thì kỹ thuật giảm mức hoặc tăng mức có thể được áp dụng. Ví dụ, nội dung 5.1 có thể được tái tạo trên hệ thống phát lại âm thanh nổi bằng cách sử dụng các phương trình giảm mức cụ thể đã biết. Một ví dụ khác là phát lại nội dung âm thanh nổi trên hệ thống loa 7.1, mà có thể bao gồm quy trình gọi là quy trình tăng mức có thể được hoặc không được điều khiển bởi thông tin có trong tín hiệu âm thanh nổi như được dùng bởi các bộ mã hóa ma trận như Dolby Pro Logic. Để điều khiển quy trình tăng mức, thông tin về vị trí ban đầu của tín hiệu trước khi giảm mức có thể được báo hiệu ẩn bằng cách bao gồm các quan hệ pha riêng vào trong các phương trình giảm mức, hoặc nói cách khác là áp dụng các phương trình giảm mức có giá trị phức. Một ví dụ nổi tiếng về

phương pháp giảm mức như vậy sử dụng các hệ số giảm mức có giá trị phức cho nội dung với các loa được đặt ở hai hướng là LtRt (Vinton và cộng sự, 2015).

Tín hiệu giảm mức (âm thanh nổi) thu được có thể được tái tạo trên hệ thống loa âm thanh nổi, hoặc có thể được tăng mức với các hệ thống loa có các loa âm thanh vòm và/hoặc trên cao. Vị trí dự định của tín hiệu có thể được suy ra bởi bộ tăng mức từ các quan hệ pha giữa các kênh. Ví dụ, trong dạng biểu diễn âm thanh nổi LtRt, lý tưởng nếu tín hiệu ngoài pha (ví dụ, có hệ số tương quan chéo chuẩn hóa dạng sóng giữa các kênh gần với -1) nên được tái tạo bởi một hoặc nhiều loa âm thanh vòm, còn hệ số tương quan dương (gần với +1) chỉ báo rằng tín hiệu nên được tái tạo bởi các loa ở phía trước người nghe.

Đã phát triển nhiều thuật toán và chiến lược tăng mức mà khác nhau trong các chiến lược của chúng để tái tạo tín hiệu đa kênh từ quy trình giảm mức âm thanh nổi. Trong các bộ tăng mức tương đối đơn giản, hệ số tương quan chéo được chuẩn hóa của các tín hiệu dạng sóng âm thanh nổi được theo dõi dưới dạng hàm thời gian, trong khi (các) tín hiệu được dẫn hướng đến các loa phía trước hoặc phía sau tùy thuộc vào giá trị của hệ số tương quan chéo được chuẩn hóa. Phương pháp này hoạt động tốt đối với nội dung tương đối đơn giản mà trong đó chỉ một đối tượng có thể nghe thấy có mặt đồng thời. Các bộ tăng mức tiên tiến hơn được dựa vào thông tin thống kê mà được suy ra từ các vùng tần số riêng để kiểm soát dòng tín hiệu từ đầu vào âm thanh nổi đến đầu ra đa kênh (Gundry 2001, Vinton và cộng sự, 2015). Đặc biệt là, mô hình tín hiệu dựa trên thành phần được dẫn hướng hoặc thành phần nổi trội và tín hiệu dư (khuếch tán) âm thanh nổi có thể được sử dụng trong các ô thời gian/tần số riêng. Ngoài việc ước tính thành phần nổi trội và các tín hiệu dư, góc hướng (theo góc phương vị, có thể được tăng độ cao) cũng được ước tính, và sau đó tín hiệu thành phần nổi trội được dẫn hướng đến một hoặc nhiều loa để tái tạo vị trí (ước tính được) trong khi phát lại.

Việc sử dụng các bộ mã hóa và bộ giải mã/bộ tăng mức ma trận không bị giới hạn ở nội dung dựa vào kênh. Các phát triển gần đây trong ngành công nghiệp âm thanh được dựa vào đối tượng âm thanh chứ không phải vào kênh âm thanh, trong đó một hoặc nhiều đối tượng bao gồm tín hiệu âm thanh và siêu dữ liệu liên quan chỉ báo, trong

số những thứ khác, vị trí dự định của nó dưới dạng hàm thời gian. Đối với nội dung âm thanh dựa vào đối tượng như vậy, bộ mã hóa ma trận cũng có thể được dùng như được nêu trong tài liệu của Vinton và các cộng sự năm 2015. Trong hệ thống này, các tín hiệu đối tượng được giảm mức thành dạng biểu diễn tín hiệu âm thanh nổi với các hệ số giảm mức phụ thuộc vào siêu dữ liệu vị trí đối tượng.

Việc tăng mức và tái tạo nội dung được mã hóa ma trận không nhất thiết bị giới hạn ở việc phát lại trên các loa. Dạng biểu diễn của thành phần được dẫn hướng hoặc nổi trội bao gồm tín hiệu thành phần nổi trội và vị trí (dự định) cho phép tái tạo trên các tai nghe nhờ tích chập với các đáp ứng xung liên quan đến đầu (head-related impulse response - HRIR) (Wightman và cộng sự, 1989). Sơ đồ đơn giản 1 về hệ thống thực hiện phương pháp này được thể hiện trên Fig.1. Tín hiệu đầu vào 2, ở định dạng được mã hóa ma trận, trước tiên được phân tích ở bước 3 để xác định hướng và độ lớn của thành phần nổi trội. Tín hiệu thành phần nổi trội được tích chập ở bước 4, 5 thông qua cặp HRIR được suy ra từ bước tra cứu 6 dựa vào hướng thành phần nổi trội, để tính toán tín hiệu đầu ra để phát lại qua tai nghe 7 sao cho tín hiệu phát lại được hiểu là đến từ hướng được xác định bởi giai đoạn phân tích thành phần nổi trội 3. Sơ đồ này có thể được áp dụng trên các tín hiệu băng rộng cũng như trên các băng phụ riêng, và có thể được bổ sung với việc xử lý riêng các tín hiệu dư (hoặc khuếch tán) theo các cách khác nhau.

Việc sử dụng các bộ mã hóa ma trận rất thích hợp để phân bố cho và tái tạo trên các bộ thu AV, nhưng có thể đem lại rắc rối cho các ứng dụng di động đòi hỏi tốc độ truyền dữ liệu thấp và tiêu hao điện năng thấp.

Bất kể sử dụng nội dung dựa trên kênh hay đối tượng, các bộ mã hóa và bộ giải mã ma trận đều phụ thuộc vào các quan hệ pha giữa các kênh tương đối chính xác của các tín hiệu được phân bố từ bộ mã hóa ma trận đến bộ giải mã ma trận. Nói cách khác, định dạng phân bố về cơ bản cần bảo toàn dạng sóng. Việc phụ thuộc như vậy vào sự bảo toàn dạng sóng có thể có vấn đề trong các điều kiện ràng buộc tốc độ bit, trong đó các bộ mã hóa-giải mã âm thanh sử dụng các phương pháp tham số thay vì các công cụ mã hóa dạng sóng để thu được chất lượng âm thanh tốt hơn. Ví dụ về các công

cụ tham số như vậy mà thường được biết là không bảo toàn dạng sóng thường được gọi là sao chép dải phổ, âm thanh nổi tham số, mã hóa âm thanh trong không gian, và tương tự như được thực hiện trong các bộ mã hóa-giải mã âm thanh MPEG-4 (ISO/IEC 14496-3:2009).

Như được đề cập ở phần trên, bộ tăng mức bao gồm thao tác phân tích và dẫn hướng (hoặc tích chập HRIR) các tín hiệu. Đối với các thiết bị chạy bằng điện, như bộ thu AV, thiết bị này thường không gây ra sự cố, nhưng đối với các thiết bị hoạt động bằng pin như điện thoại di động và máy tính bảng, độ phức tạp tính toán và các yêu cầu về bộ nhớ tương ứng liên quan đến các quy trình này thường không mong muốn vì chúng có ảnh hưởng xấu đến tuổi thọ pin.

Việc phân tích nêu trên thường cũng gây ra độ trễ âm thanh bổ sung. Độ trễ âm thanh như vậy là không mong muốn vì (1) nó đòi hỏi trễ video để duy trì sự nhép môi giữa âm thanh và video cần nhớ và công suất xử lý ở mức độ lớn, và (2) có thể gây ra sự không đồng bộ/độ trễ giữa chuyển động đầu và kết xuất âm thanh trong trường hợp theo dõi chuyển động đầu.

Kỹ thuật giảm mức mã hóa ma trận cũng có thể dường như không tối ưu âm thanh trên tai nghe hoặc loa âm thanh nổi, do có thể xuất hiện các thành phần tín hiệu mạnh ngoài pha.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất dạng đầu ra âm thanh lập thể hai tai tham số được cải thiện.

Theo khía cạnh thứ nhất, sáng chế đề xuất phương pháp mã hóa âm thanh đầu vào dựa trên kênh hoặc đối tượng để phát lại, phương pháp này bao gồm các bước: (a) ban đầu kết xuất âm thanh đầu vào dựa trên kênh hoặc đối tượng thành dạng trình diễn đầu ra ban đầu (ví dụ, sự biểu diễn đầu ra ban đầu); (b) xác định giá trị ước tính của thành phần âm thanh nổi trội từ âm thanh đầu vào dựa trên kênh hoặc đối tượng và xác

định chuỗi các hệ số gia trọng cho thành phần âm thanh nổi trội để ánh xạ sự trình diễn đầu ra ban đầu thành thành phần âm thanh nổi trội; (c) xác định giá trị ước tính của hướng hoặc vị trí của thành phần âm thanh nổi trội; và (d) mã hóa sự trình diễn đầu ra ban đầu, các hệ số gia trọng cho thành phần âm thanh nổi trội, hướng hoặc vị trí của thành phần âm thanh nổi trội dưới dạng tín hiệu mã hóa để phát lại. Việc cung cấp chuỗi các hệ số gia trọng cho thành phần âm thanh nổi trội để ánh xạ sự trình diễn đầu ra ban đầu thành thành phần âm thanh nổi trội có thể cho phép sử dụng các hệ số gia trọng cho thành phần âm thanh nổi trội và sự trình diễn đầu ra ban đầu để xác định giá trị ước tính của thành phần nổi trội.

Theo một số phương án, phương pháp còn bao gồm bước xác định giá trị ước tính của hỗn hợp dư là dạng trình diễn đầu ra ban đầu trừ đi phần kết xuất của thành phần âm thanh nổi trội hoặc giá trị ước tính của nó. Phương pháp có thể còn bao gồm bước tạo ra hỗn hợp âm thanh lập thể hai tai không tiếng vang của âm thanh đầu vào dựa trên kênh hoặc đối tượng, và xác định giá trị ước tính của hỗn hợp dư, trong đó giá trị ước tính của hỗn hợp dư có thể là hỗn hợp âm thanh lập thể hai tai không tiếng vang trừ đi phần kết xuất của thành phần âm thanh nổi trội hoặc giá trị ước tính của nó. Hơn nữa, phương pháp có thể bao gồm bước xác định chuỗi các hệ số ma trận dư để ánh xạ sự trình diễn đầu ra ban đầu thành giá trị ước tính của hỗn hợp dư.

Trình diễn đầu ra ban đầu có thể bao gồm trình diễn qua tai nghe hoặc qua loa. Âm thanh đầu vào dựa trên kênh hoặc đối tượng có thể được xếp lớp theo thời gian và tần số và bước mã hóa có thể được lặp lại đối với chuỗi các bước thời gian và chuỗi các băng tần số. Trình diễn đầu ra ban đầu có thể bao gồm hỗn hợp truyền qua loa âm thanh nổi.

Theo khía cạnh khác, sáng chế đề xuất phương pháp giải mã tín hiệu âm thanh mã hóa, tín hiệu âm thanh mã hóa này bao gồm: dạng trình diễn đầu ra thứ nhất (ví dụ, ban đầu) (ví dụ, biểu diễn đầu ra thứ nhất/ban đầu); -hướng của thành phần âm thanh nổi trội và các hệ số gia trọng cho thành phần âm thanh nổi trội; phương pháp này bao gồm các bước: (a) sử dụng các hệ số gia trọng cho thành phần âm thanh nổi trội và trình diễn đầu ra ban đầu để xác định thành phần nổi trội ước tính được; (b) kết xuất thành

phần nổi trội ước tính được với sự tạo âm thanh lập thể hai tai tại vị trí không gian tương ứng với người nghe đã định theo hướng của thành phần âm thanh nổi trội để tạo ra thành phần nổi trội ước tính được tạo âm thanh lập thể hai tai được kết xuất; (c) tái tạo giá trị ước tính của thành phần dư từ dạng trình diễn đầu ra thứ nhất (ví dụ, ban đầu); và (d) kết hợp thành phần nổi trội ước tính được tạo âm thanh lập thể hai tai được kết xuất và giá trị ước tính của thành phần dư để tạo ra tín hiệu âm thanh mã hóa theo không gian được xuất ra.

Tín hiệu âm thanh được mã hóa còn có thể bao gồm chuỗi hệ số ma trận dư thể hiện tín hiệu âm thanh dư và bước (c) còn có thể bao gồm bước (c1) áp dụng các hệ số ma trận dư cho sự trình diễn đầu ra thứ nhất (ví dụ, ban đầu) để tái tạo giá trị ước tính của thành phần dư.

Theo một số phương án, giá trị ước tính của thành phần dư có thể được tái tạo bằng cách lấy dạng trình diễn đầu ra thứ nhất (ví dụ, ban đầu) trừ đi thành phần nổi trội ước tính được tạo âm thanh lập thể hai tai được kết xuất. Bước (b) có thể bao gồm bước xoay ban đầu thành phần nổi trội ước tính theo tín hiệu theo dõi đầu đầu vào chỉ báo hướng đầu của người nghe đã định.

Theo khía cạnh khác, sáng chế đề xuất phương pháp giải mã và tái tạo dòng âm thanh cho người nghe sử dụng tai nghe, phương pháp này bao gồm các bước: (a) thu dòng dữ liệu chứa dạng biểu diễn âm thanh thứ nhất và dữ liệu biến đổi âm thanh bổ sung; (b) thu dữ liệu hướng đầu thể hiện hướng của người nghe; (c) tạo một hoặc nhiều tín hiệu phụ dựa vào dạng biểu diễn âm thanh thứ nhất và dữ liệu biến đổi thu được; (d) tạo ra dạng biểu diễn âm thanh thứ hai bao gồm tổ hợp của dạng biểu diễn âm thanh thứ nhất và (các) tín hiệu phụ, trong đó một hoặc nhiều trong số (các) tín hiệu phụ được sửa đổi để đáp lại dữ liệu hướng đầu; và (e) xuất ra dạng biểu diễn âm thanh thứ hai dưới dạng dòng âm thanh đầu ra.

Theo một số phương án, phương pháp này còn có thể bao gồm bước sửa đổi các tín hiệu phụ bao gồm mô phỏng đường âm thanh từ vị trí nguồn âm thanh đến tai người nghe. Dữ liệu biến đổi có thể bao gồm các hệ số ma trận và ít nhất một trong số:

vị trí nguồn âm thanh hoặc hướng nguồn âm thanh. Quy trình biến đổi có thể được áp dụng dưới dạng hàm thời gian hoặc tần số. Tín hiệu phụ có thể thể hiện ít nhất một thành phần nổi trội. Hướng hoặc vị trí nguồn âm thanh có thể được thu dưới dạng một phần của dữ liệu biến đổi và có thể được xoay để đáp lại dữ liệu hướng đầu. Theo một số phương án, lượng xoay tối đa bị giới hạn ở giá trị nhỏ hơn 360 độ trong góc phương vị hoặc độ cao. Dạng biểu diễn thứ hai có thể thu được từ dạng biểu diễn thứ nhất bằng cách tạo ma trận trong miền biến đổi hoặc miền giàn bộ lọc. Dữ liệu biến đổi còn có thể bao gồm các hệ số ma trận bổ sung, và bước (d) còn có thể bao gồm sửa đổi dạng trình diễn âm thanh thứ nhất để đáp lại các hệ số ma trận bổ sung trước khi kết hợp dạng trình diễn âm thanh thứ nhất và (các) tín hiệu âm thanh phụ.

Mô tả vắn tắt các hình vẽ

Sau đây, các phương án của sáng chế sẽ được mô tả, chỉ bằng ví dụ, có tham chiếu đến các hình vẽ kèm theo, trong đó:

Fig.1 minh họa dưới dạng biểu đồ bộ giải mã tai nghe sử dụng nội dung được được mã hóa theo ma trận;

Fig.2 minh họa dưới dạng biểu đồ bộ mã hóa theo một phương án;

Fig.3 là sơ đồ khối giản lược của bộ giải mã;

Fig.4 là sự thể hiện chi tiết của bộ mã hóa; và

Fig.5 minh họa chi tiết hơn một dạng của bộ giải mã.

Mô tả chi tiết sáng chế

Theo các phương án, sáng chế đề xuất hệ thống và phương pháp biểu diễn nội dung âm thanh dựa vào đối tượng hoặc kênh mà (1) tương thích với phát lại âm thanh nổi, (2) cho phép phát lại âm thanh lập thể hai tai bao gồm theo dõi chuyển động đầu,

(3) có độ phức tạp giải mã thấp, và (4) không phụ thuộc vào nhưng vẫn tương thích với mã hóa ma trận.

Điều này đạt được bằng cách kết hợp việc phân tích ở phía bộ mã hóa một hoặc nhiều thành phần nổi trội (hoặc đối tượng nổi trội hoặc sự kết hợp của chúng) bao gồm các trọng số để dự báo các thành phần nổi trội này từ giá trị giảm mức, kết hợp với các tham số bổ sung mà giảm đến mức tối thiểu lỗi giữa kết xuất âm thanh lập thể hai tai dựa vào các thành phần nổi trội hoặc được dẫn hướng riêng lẻ, và dạng trình diễn âm thanh lập thể hai tai mong muốn của nội dung hoàn chỉnh.

Theo một phương án, phép phân tích thành phần nổi trội (hoặc nhiều thành phần nổi trội) được thực hiện trong bộ mã hóa chứ không phải trong bộ giải mã/bộ kết xuất. Sau đó dòng âm thanh được tăng cường với siêu dữ liệu chỉ báo hướng của thành phần nổi trội, và thông tin về cách thức (các) thành phần nổi trội có thể thu được từ tín hiệu giảm mức liên quan.

Fig.2 minh họa một dạng của bộ mã hóa 20 theo một phương án được ưu tiên. Nội dung dựa vào đối tượng hoặc kênh 21 được phân tích ở bước 23 để xác định (các) thành phần nổi trội. Bước phân tích này có thể diễn ra dưới dạng hàm thời gian và tần số (giả sử nội dung âm thanh được chia nhỏ thành các ô thời gian và các ô phụ tần số). Kết quả của quá trình này là tín hiệu thành phần nổi trội 26 (hoặc nhiều tín hiệu thành phần nổi trội), và thông tin về (các) vị trí hoặc (các) hướng liên quan 25. Sau đó, các trọng số được ước tính ở bước 24 và xuất ra ở bước 27 để cho phép tái tạo (các) tín hiệu thành phần nổi trội từ giá trị giảm mức được truyền. Bộ tạo giá trị giảm mức 22 không nhất thiết phải tuân theo quy tắc giảm mức LtRt, nhưng có thể là giá trị giảm mức ITU (LoRo) chuẩn sử dụng các hệ số giảm mức không âm, có giá trị thực. Cuối cùng, tín hiệu giảm mức được xuất ra 29, các trọng số 27, và dữ liệu vị trí 25 được đóng gói bằng bộ mã hóa âm thanh 28 và chuẩn bị cho việc phân bố.

Fig.3 minh họa bộ giải mã tương ứng 30 theo một phương án được ưu tiên. Bộ giải mã âm thanh tái tạo tín hiệu giảm mức. Tín hiệu được nhập vào ở bước 31 và được giải nén bởi bộ giải mã âm thanh 32 thành tín hiệu giảm mức, các trọng số và hướng

của các thành phần nổi trội. Sau đó, các trọng số ước tính thành phần nổi trội được dùng để tái tạo ở bước 34 (các) thành phần được dẫn hướng, các thành phần này được kết xuất ở bước 36 bằng cách sử dụng dữ liệu vị trí hoặc hướng được truyền. Dữ liệu vị trí có thể tùy ý được sửa đổi ở bước 33 phụ thuộc vào thông tin xoay hoặc dịch chuyển đầu 38. Ngoài ra, (các) thành phần nổi trội được tái tạo có thể bị trừ đi ở bước 35 từ giá trị giảm mức. Một cách tùy ý, có sự trừ đi (các) thành phần nổi trội trong đường giảm mức, nhưng theo cách khác, phép trừ này cũng có thể xảy ra ở bộ mã hóa, như được mô tả dưới đây.

Để cải thiện việc loại bỏ hoặc khử thành phần nổi trội được tái tạo trong bộ trừ ở bước 35, trước tiên đầu ra của thành phần nổi trội có thể được kết xuất bằng cách sử dụng dữ liệu vị trí hoặc hướng được truyền trước khi trừ. Giai đoạn kết xuất tùy ý ở bước 39 này được thể hiện trên Fig.3.

Sau đây sẽ mô tả đầu tiên bộ mã hóa chi tiết hơn, Fig.4 thể hiện một dạng của bộ mã hóa 40 để xử lý nội dung âm thanh dựa vào đối tượng (ví dụ Dolby Atmos). Các đối tượng âm thanh được lưu trữ ban đầu dưới dạng các đối tượng Atmos ở bước 41 và ban đầu được chia thành các ô thời gian và tần số bằng cách sử dụng dàn bộ lọc gương cầu phương giá trị phức lai (hybrid complex-valued quadrature mirror filter - HCQMF) ở bước 42. Các tín hiệu đối tượng đầu vào có thể được biểu thị bởi $x_i[n]$ khi các tác giả sáng chế lược bỏ các chỉ số thời gian và tần số tương ứng; vị trí tương ứng trong khung hiện thời được xác định bằng vector đơn vị $\vec{p}_i$, và chỉ số i thể hiện số lượng đối tượng, và chỉ số n thể hiện thời gian (ví dụ, chỉ số mẫu băng phụ). Các tín hiệu đối tượng đầu vào $x_i[n]$ là ví dụ của âm thanh đầu vào dựa trên kênh hoặc đối tượng.

Hỗn hợp âm thanh lập thể hai tai, băng phụ, không tiếng vang Y (y_l, y_r) được tạo ra ở bước 43 bằng cách sử dụng các vô hướng có giá trị phức $H_{l,i}, H_{r,i}$ (ví dụ, các HRTF một chạm 48) mà biểu diễn dạng biểu diễn băng phụ của các HRIR tương ứng với vị trí $\vec{p}_i$:

$$y_l[n] = \sum_i H_{l,i} x_i[n]$$

$$y_r[n] = \sum_i H_{r,i} x_i[n]$$

Theo cách khác, hỗn hợp âm thanh lập thể hai tai $Y(y_l, y_r)$ có thể được tạo ra bằng phép tích chập sử dụng các đáp ứng xung liên quan tới đầu (head-related impulse responses – HRIR). Ngoài ra, giá trị giảm mức âm thanh nổi z_l, z_r (ví dụ bao gồm dạng trình diễn đầu ra ban đầu) được tạo ra ở bước 44 bằng cách sử dụng các hệ số độ lợi quét theo biên độ $g_{l,i}, g_{r,i}$:

$$z_l[n] = \sum_i g_{l,i} x_i[n]$$

$$z_r[n] = \sum_i g_{r,i} x_i[n]$$

Vectơ chỉ hướng của thành phần nổi trội $\vec{p}_D$ (ví dụ bao gồm hướng hoặc vị trí của thành phần âm thanh nổi trội) có thể được ước tính bằng cách tính toán thành phần nổi trội ở bước 45 khi ban đầu tính toán tổng được gán trọng số của các vectơ chỉ hướng đơn vị cho mỗi đối tượng:

$$\vec{p}_D = \frac{\sum_i \sigma_i^2 \vec{p}_i}{\sum_i \sigma_i^2}$$

với σ_i^2 năng lượng của tín hiệu $x_i[n]$:

$$\sigma_i^2 = \sum_n x_i[n] x_i^*[n]$$

và với $(.)^*$ là toán tử tiếp hợp phức.

Tín hiệu nổi trội/được dẫn hướng, $d[n]$ (ví dụ bao gồm thành phần âm thanh nổi trội) sau đó được xác định bởi:

$$d[n] = \sum_i x_i[n] \mathcal{F}(\vec{p}_D, \vec{p}_i)$$

với $\mathcal{F}(\vec{p}_1, \vec{p}_2)$ một hàm tạo ra độ lợi giảm khi tăng khoảng cách giữa các vector đơn vị $\vec{p}_1, \vec{p}_2$. Ví dụ, để tạo ra micrô ảo có mẫu định hướng dựa vào hàm điều hòa cầu bậc cao hơn, một phương án thực hiện sẽ tương ứng với:

$$\mathcal{F}(\vec{p}_1, \vec{p}_2) = (a + b\vec{p}_1^T \cdot \vec{p}_2)^c$$

với $\vec{p}_i$ là vector hướng đơn vị trong hệ tọa độ hai hoặc ba chiều, $(.)$ là toán tử tích vô hướng của hai vector, và với các tham số làm ví dụ a, b, c (ví dụ $a=b=0,5; c=1$).

Các trọng số hoặc các hệ số dự báo $w_{l,d}, w_{r,d}$ được tính toán ở bước 46 và được dùng để tính toán tín hiệu dẫn hướng ước tính được $\hat{d}[n]$ ở bước 47:

$$\hat{d}[n] = w_{l,d}z_l + w_{r,d}z_r$$

với các trọng số $w_{l,d}, w_{r,d}$ cực tiểu hóa sai số bình phương trung bình giữa $d[n]$ và $\hat{d}[n]$ cho các tín hiệu giảm mức z_l, z_r . Các trọng số $w_{l,d}, w_{r,d}$ là ví dụ cho các hệ số gia trọng cho thành phần âm thanh nổi trội để ánh xạ sự trình diễn đầu ra ban đầu (ví dụ, z_l, z_r) thành thành phần âm thanh nổi trội (ví dụ, $\hat{d}[n]$). Một phương pháp đã biết để suy ra các trọng số này là áp dụng bộ dự báo sai số bình phương tối thiểu (minimum mean-square error - MMSE):

$$\begin{bmatrix} w_{l,d} \\ w_{r,d} \end{bmatrix} = (R_{zz} + \epsilon I)^{-1} R_{zd}$$

với R_{ab} là ma trận hiệp phương sai giữa các tín hiệu của các tín hiệu a và các tín hiệu b , và ϵ là tham số điều chuẩn.

Sau đó có thể lấy hỗn hợp âm thanh lập thể hai tai không tiếng vang y_l, y_r trừ đi giá trị ước tính kết xuất được của tín hiệu thành phần nổi trội $\hat{d}[n]$ ở bước 49 để tạo

ra hỗn hợp âm thanh lập thể hai tai dư $\tilde{y}_l, \tilde{y}_r$ bằng cách sử dụng HRTF (HRIR) $H_{l,D}, H_{r,D}$ liên quan đến hướng/vị trí $\vec{p}_D$ của tín hiệu thành phần nổi trội $\hat{d}$:

$$\tilde{y}_l[n] = y_l[n] - H_{l,D}\hat{d}[n]$$

$$\tilde{y}_r[n] = y_r[n] - H_{r,D}\hat{d}[n]$$

Cuối cùng, một tập hợp hệ số dự báo khác hoặc các trọng số $w_{i,j}$ được ước tính ở bước 51 cho phép tái tạo hỗn hợp âm thanh lập thể hai tai dư $\tilde{y}_l, \tilde{y}_r$ từ hỗn hợp âm thanh nổi z_l, z_r bằng cách sử dụng các ước tính sai số bình phương trung bình tối thiểu:

$$\begin{bmatrix} w_{1,1} & w_{1,2} \\ w_{2,1} & w_{2,2} \end{bmatrix} = (R_{zz} + \epsilon I)^{-1} R_{z\tilde{y}}$$

với R_{ab} là ma trận hiệp phương sai giữa các tín hiệu của dạng biểu diễn a và dạng biểu diễn b , và ϵ là tham số điều chuẩn. Các hệ số dự báo hoặc các trọng số $w_{i,j}$ là ví dụ của các hệ số ma trận dư để ánh xạ dạng trình diễn đầu ra ban đầu (ví dụ, z_l, z_r) thành giá trị ước tính của hỗn hợp âm thanh lập thể hai tai dư $\tilde{y}_l, \tilde{y}_r$. Biểu thức trên có thể bị hạn chế mức bổ sung để khắc phục mọi sự suy hao dự báo. Bộ mã hóa xuất ra thông tin sau:

Hỗn hợp âm thanh nổi z_l, z_r (ví dụ bao gồm dạng trình diễn đầu ra ban đầu);

Các hệ số để ước tính thành phần nổi trội $w_{l,d}, w_{r,d}$ (ví dụ bao gồm các hệ số gia trọng cho thành phần âm thanh nổi trội);

Vị trí hoặc hướng của thành phần nổi trội $\vec{p}_D$;

Và tùy ý, các trọng số dư $w_{i,j}$ (ví dụ bao gồm các hệ số ma trận dư).

Mặc dù phần mô tả trên đây đề cập đến việc kết xuất dựa vào một thành phần nổi trội, nhưng theo một số phương án bộ mã hóa có thể được làm thích ứng để phát hiện nhiều thành phần nổi trội, xác định các trọng số và các hướng cho mỗi trong số nhiều thành phần nổi trội, kết xuất và trừu tượng hóa mỗi trong số nhiều thành phần nổi trội từ hỗn

hợp âm thanh lập thể hai tai không tiếng vang Y , và sau đó xác định các trọng số dư sau khi mỗi trong số nhiều thành phần nổi trội đã được trừ đi từ hỗn hợp âm thanh lập thể hai tai không tiếng vang Y .

Bộ giải mã/bộ kết xuất

Fig.5 minh họa chi tiết hơn một dạng của bộ giải mã/bộ kết xuất 60. Bộ giải mã/bộ kết xuất 60 áp dụng quy trình nhằm tái tạo hỗn hợp âm thanh lập thể hai tai y_l, y_r để xuất ra cho người nghe 71 từ thông tin đầu vào được mở gói z_l, z_r ; $w_{l,d}, w_{r,d}$; $\vec{p}_D$; $w_{i,j}$. Ở đây, hỗn hợp âm thanh nổi z_l, z_r là một ví dụ về dạng biểu diễn âm thanh thứ nhất, và các hệ số dự báo hoặc các trọng số $w_{i,j}$ và/hoặc hướng/vị trí $\vec{p}_D$ của tín hiệu thành phần nổi trội ã là các ví dụ về dữ liệu biến đổi âm thanh bổ sung.

Ban đầu, giá trị giảm mức âm thanh nổi được tách thành các ô thời gian/tần số bằng cách sử dụng dàn bộ lọc hoặc phép biến đổi phù hợp ở bước 61, chẳng hạn như dàn bộ lọc phân tích HCQMF 61. Các phép biến đổi khác như biến đổi Fourier rời rạc, biến đổi cosin hoặc sin (đã điều chế), dàn bộ lọc miền thời gian, hoặc các biến đổi sóng con cũng có thể được áp dụng như nhau. Sau đó, tín hiệu thành phần nổi trội ước tính được $\hat{d}[n]$ được tính toán ở bước 63 bằng cách sử dụng các trọng số hệ số dự báo $w_{l,d}, w_{r,d}$:

$$\hat{d}[n] = w_{l,d}z_l + w_{r,d}z_r$$

Tín hiệu thành phần nổi trội ước tính được $\hat{d}[n]$ là một ví dụ về tín hiệu phụ. Do đó, bước này có thể được xem là tương ứng với bước tạo một hoặc nhiều tín hiệu phụ dựa vào dạng biểu diễn âm thanh thứ nhất và dữ liệu biến đổi thu được.

Tín hiệu thành phần nổi trội này sau đó được kết xuất ở bước 65 và được sửa đổi ở bước 68 với các HRTF ở bước 69 dựa vào dữ liệu vị trí/hướng được truyền $\vec{p}_D$, có thể được sửa đổi (được xoay) dựa vào thông tin thu được từ bộ theo dõi chuyển động đầu 62. Cuối cùng, toàn bộ đầu ra âm thanh lập thể hai tai không tiếng vang bao gồm

tín hiệu thành phần nổi trội kết xuất được được cộng ở bước 66 với các giá trị dư tái tạo được $\tilde{y}_l, \tilde{y}_r$ dựa vào các trọng số hệ số dự báo $w_{i,j}$:

$$\begin{bmatrix} \tilde{y}_l \\ \tilde{y}_r \end{bmatrix} = \begin{bmatrix} w_{1,1} & w_{1,2} \\ w_{2,1} & w_{2,2} \end{bmatrix} \begin{bmatrix} z_l \\ z_r \end{bmatrix}$$

$$\begin{bmatrix} \hat{y}_l \\ \hat{y}_r \end{bmatrix} = \left(\begin{bmatrix} w_{1,1} & w_{1,2} \\ w_{2,1} & w_{2,2} \end{bmatrix} + \begin{bmatrix} H_{l,D} \\ H_{r,D} \end{bmatrix} \begin{bmatrix} w_{l,d} & w_{r,d} \end{bmatrix} \right) \begin{bmatrix} z_l \\ z_r \end{bmatrix}$$

Toàn bộ đầu ra âm thanh lập thể hai tai không tiếng vang là một ví dụ của dạng biểu diễn âm thanh thứ hai. Do đó, bước này có thể được xem là tương ứng với bước tạo ra dạng biểu diễn âm thanh thứ hai bao gồm tổ hợp của dạng biểu diễn âm thanh thứ nhất và (các) tín hiệu phụ, trong đó một hoặc nhiều trong số (nhiều) tín hiệu phụ đã được sửa đổi để đáp lại dữ liệu hướng đầu.

Cần lưu ý thêm rằng nếu nhận được thông tin trên nhiều hơn một tín hiệu nổi trội, thì mỗi tín hiệu nổi trội có thể được kết xuất và được thêm vào tín hiệu dư tái tạo được.

Nếu không áp dụng xoay hoặc dịch chuyển đầu, thì các tín hiệu đầu ra $\hat{y}_l, \hat{y}_r$ rất gần (về sai số bình phương trung bình gốc) với các tín hiệu âm thanh lập thể hai tai tham chiếu y_l, y_r miễn là

$$\hat{d}[n] \approx d[n]$$

Các đặc tính cơ bản

Như có thể thấy từ công thức phương trình trên, hoạt động hiệu quả để xây dựng dạng sự trình diễn âm thanh lập thể hai tai không tiếng vang từ dạng trình diễn âm thanh nổi bao gồm ma trận 2×2 ở bước 70, trong đó các hệ số ma trận là phụ thuộc vào thông tin được truyền $w_{l,d}, w_{r,d}; \vec{p}_D; w_{i,j}$ và thông tin chuyển động xoay và/hoặc dịch chuyển của máy theo dõi chuyển động đầu. Điều này chỉ báo rằng độ phức tạp của quy

trình này là tương đối thấp vì việc phân tích các thành phần nổi trội được áp dụng ở bộ mã hóa thay vì bộ giải mã.

Nếu thành phần nổi trội không được ước tính (ví dụ, $w_{l,d}, w_{r,d} = 0$), thì giải pháp được mô tả là tương đương với phương pháp lập thể theo tham số.

Trong trường hợp muốn loại trừ một số đối tượng nhất định ra khỏi thông tin theo dõi xoay đầu/thông tin theo dõi đầu, các đối tượng này có thể bị loại khỏi (1) quy trình phân tích hướng thành phần nổi trội, và (2) quy trình dự báo tín hiệu thành phần nổi trội. Kết quả là, các đối tượng này sẽ được chuyển đổi từ âm thanh nổi sang âm thanh lập thể hai tai qua các hệ số $w_{i,j}$ và do đó không bị ảnh hưởng bởi sự xoay hoặc dịch chuyển của đầu.

Với ý tưởng tương tự, các đối tượng có thể được thiết lập ở chế độ ‘đi qua’, điều này có nghĩa là ở dạng trình diễn âm thanh lập thể hai tai, các đối tượng sẽ được quét biên độ chứ không phải là tích chập HRIR. Điều này có thể đạt được bằng cách sử dụng đơn giản các độ lợi quét biên độ cho các hệ số $H_{.,i}$ thay vì HRTF một chạm hoặc quy trình xử lý âm thanh lập thể hai tai thích hợp khác bất kỳ.

Mở rộng

Các phương án không bị giới hạn ở việc sử dụng các giá trị giảm mức âm thanh nổi vì số lượng kênh khác cũng có thể được sử dụng.

Bộ giải mã 60 được mô tả có tham chiếu đến Fig.5 có tín hiệu đầu ra bao gồm hướng thành phần nổi trội kết xuất được cộng tín hiệu đầu vào được tạo ma trận bằng các hệ số ma trận $w_{i,j}$. Các hệ số ma trận có thể được suy ra theo các cách khác, ví dụ:

1. Các hệ số $w_{i,j}$ có thể được xác định trong bộ mã hóa bằng cách tái tạo tham số các tín hiệu $\tilde{y}_l, \tilde{y}_r$. Nói cách khác, theo phương án thực hiện này, các hệ số $w_{i,j}$ tập trung vào việc tái tạo chính xác các tín hiệu âm thanh lập thể hai tai y_l, y_r mà sẽ thu

được khi kết xuất các đối tượng/kênh đầu vào gốc theo cách âm thanh lập thể hai tai; nói cách khác, các hệ số $w_{i,j}$ được điều chỉnh nội dung.

2. Các hệ số $w_{i,j}$ có thể được gửi từ bộ mã hóa đến bộ giải mã để biểu diễn HRTF cho các vị trí không gian cố định, ví dụ tại các góc phương vị $\pm 45^\circ$. Nói cách khác, tín hiệu dư được xử lý để mô phỏng sự tái tạo trên hai loa ảo ở một số vị trí. Do các hệ số thể hiện HRTF được truyền từ bộ mã hóa đến bộ giải mã, nên các vị trí của các loa ảo có thể thay đổi theo thời gian và tần số. Nếu phương pháp này được thực hiện bằng cách sử dụng loa ảo tĩnh để biểu diễn tín hiệu dư, thì các hệ số $w_{i,j}$ không cần truyền từ bộ mã hóa đến bộ giải mã, và thay vào đó có thể được nối cứng trong bộ giải mã. Một biến thể của phương pháp này bao gồm một tập hợp vị trí tĩnh giới hạn mà có sẵn trong bộ giải mã, cùng với các hệ số tương ứng của chúng $w_{i,j}$, và việc chọn vị trí tĩnh nào được dùng để xử lý tín hiệu dư được báo hiệu từ bộ mã hóa đến bộ giải mã.

Các tín hiệu $\tilde{y}_l, \tilde{y}_r$ có thể phụ thuộc vào thiết bị được gọi là bộ tăng mức, tái tạo nhiều hơn 2 tín hiệu nhờ phân tích thống kê các tín hiệu này tại bộ giải mã, sau đó là kết xuất âm thanh lập thể hai tai các tín hiệu tăng mức thu được.

Các phương pháp được mô tả còn có thể được áp dụng trong hệ thống trong đó tín hiệu được truyền Z là tín hiệu âm thanh lập thể hai tai. Trong trường hợp cụ thể đó, bộ giải mã 60 trên Fig.5 vẫn như nguyên, còn khối có nhãn ‘Tạo ra hỗn hợp âm thanh nổi (LoRo)’ 44 trên Fig.4 cần được thay thế bằng ‘Tạo ra hỗn hợp âm thanh lập thể hai tai không tiếng vang’ 43 (Fig.4) mà giống như khối tạo ra cặp tín hiệu Y. Ngoài ra, các dạng trộn khác có thể được tạo ra theo yêu cầu.

Phương pháp này có thể được mở rộng với các phương pháp tái tạo một hoặc nhiều tín hiệu đầu vào FDN từ hỗn hợp âm thanh nổi được truyền mà bao gồm một tập hợp con các đối tượng hoặc kênh cụ thể.

Phương pháp này có thể được mở rộng với nhiều thành phần nổi trội được dự báo từ hỗn hợp âm thanh nổi được truyền, và được kết xuất ở phía bộ giải mã. Không có giới hạn cơ bản nào về việc chỉ dự báo một thành phần nổi trội cho mỗi ô thời gian/tần

số. Cụ thể là, số lượng các thành phần nổi trội có thể khác nhau theo mỗi ô thời gian/tần số.

Giải thích

Sự tham chiếu trong toàn bộ bản mô tả đến “một phương án”, “một số phương án” hoặc “phương án” có nghĩa là một dấu hiệu, cấu trúc hoặc đặc điểm cụ thể được mô tả liên quan đến phương án này được bao gồm trong ít nhất một phương án theo sáng chế. Do đó, sự xuất hiện của cụm từ “theo một phương án”, “theo một số phương án” hoặc “theo phương án” ở nhiều chỗ khác nhau trong toàn bộ bản mô tả không nhất thiết là đề cập đến cùng phương án, nhưng có thể là đề cập đến cùng. Hơn thế nữa, các dấu hiệu, cấu trúc hoặc đặc điểm 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 qua phân bố cục này, theo một hoặc nhiều phương án.

Như được sử dụng ở đây, trừ phi được quy định khác, việc sử dụng các tính từ chỉ thứ tự "thứ nhất", "thứ hai", "thứ ba", v.v., để mô tả một đối tượng chung, đơn thuần chỉ ra rằng các trường hợp khác nhau của các đối tượng giống nhau đang được đề cập đến, và không nhằm để ngụ ý rằng các đối tượng được mô tả như vậy phải theo một trình tự cho trước, hoặc về thời gian, không gian, trong việc xếp hạng hoặc theo cách thức khác bất kỳ.

Trong các điểm yêu cầu bảo hộ dưới đây và phần mô tả này, thuật ngữ bất kỳ trong số các thuật ngữ bao gồm (comprising), chứa (comprised of) hoặc mà bao gồm (which comprises) là thuật ngữ mở có nghĩa là bao gồm ít nhất nhiều phần tử/dấu hiệu theo sau, chứ không phải là loại trừ các phần tử/dấu hiệu khác. Do đó, thuật ngữ bao gồm, khi được sử dụng trong yêu cầu bảo hộ, không được hiểu là giới hạn ở các phương tiện hoặc phần tử hoặc bước được liệt kê sau nó. Ví dụ, phạm vi của cụm từ thiết bị bao gồm A và B không bị giới hạn ở các thiết bị chỉ gồm có các thành phần A và B. Một thuật ngữ bất kỳ trong số các thuật ngữ bao gồm (including) hoặc mà bao gồm (which includes) hoặc mà bao gồm (that includes) như được sử dụng ở đây cũng là một thuật ngữ mở mà cũng có nghĩa là bao gồm ít nhất các phần tử /dấu hiệu theo sau thuật ngữ

này, chứ không loại trừ các thành phần/dấu hiệu khác. Do đó, bao gồm (including) là từ đồng nghĩa với và có nghĩa là bao gồm (comprising).

Như được sử dụng ở đây, thuật ngữ "làm ví dụ" được sử dụng theo nghĩa đưa ra các ví dụ, không thể hiện phẩm chất. Tức là, "phương án làm ví dụ" là một phương án được đưa ra như là một ví dụ, chứ không phải là một phương án có phẩm chất ví dụ.

Cần phải hiểu rằng trong phần mô tả trên đây về các phương án làm ví dụ theo sáng chế, các dấu hiệu khác nhau của sáng chế đôi khi được nhóm cùng nhau trong một phương án, hình vẽ hoặc phần mô tả của chúng nhằm mục đích hợp lý hóa sáng chế và giúp hiểu một hoặc nhiều trong số các khía cạnh của sáng chế. Tuy nhiên, phương pháp bộc lộ này không được hiểu là thể hiện ý định rằng sáng chế được yêu cầu bảo hộ cần nhiều dấu hiệu hơn được thể hiện trong mỗi điểm yêu cầu bảo hộ. Đúng hơn là, như các điểm yêu cầu bảo hộ sau đây thể hiện, các khía cạnh của sáng chế nằm trong ít hơn tất cả các dấu hiệu của một phương án được bộc lộ trên đây. Do vậy, các điểm yêu cầu bảo hộ sau phần mô tả chi tiết được đưa rõ ràng vào phần mô tả chi tiết này, với mỗi điểm yêu cầu bảo hộ là một phương án riêng của sáng chế.

Hơn thế nữa, mặc dù một số phương án mô tả trong sáng chế bao gồm một số dấu hiệu nhưng không bao gồm các dấu hiệu khác nằm trong các phương án khác, nhưng các tổ hợp của các dấu hiệu của các phương án khác nhau được hiểu là nằm trong phạm vi của sáng chế và tạo thành các phương án khác nhau, như được hiểu bởi người có hiểu biết trung bình trong lĩnh vực này. Ví dụ, trong các điểm yêu cầu bảo hộ dưới đây, phương án bất kỳ trong số các phương án được yêu cầu bảo hộ có thể được sử dụng trong tổ hợp bất kỳ.

Hơn thế nữa, một số trong số các phương án được mô tả ở đây là phương pháp hoặc tổ hợp của các phần tử của phương pháp mà có thể thực hiện được bởi bộ xử lý của hệ thống máy tính hoặc bởi phương tiện khác để thực hiện chức năng. Do vậy, bộ xử lý với các lệnh cần thiết để thực hiện phương pháp như vậy hoặc một phần tử của phương pháp tạo ra một phương tiện để thực hiện phương pháp hoặc một phần tử của phương pháp. Ngoài ra, một phần tử được mô tả ở đây của một phương án thiết bị là

một ví dụ về phương tiện thực hiện chức năng được thực hiện bởi phần tử nhằm mục đích thực hiện sáng chế.

Trong phần mô tả được ra trong sáng chế, nhiều chi tiết cụ thể được nêu. Tuy nhiên, cần hiểu rằng các phương án 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 được biết rộng rãi không được thể hiện một cách chi tiết để không làm khó hiểu phần mô tả này.

Tương tự, cần lưu ý rằng thuật ngữ được ghép nối, khi được sử dụng trong phần yêu cầu bảo hộ, không nên được hiểu là chỉ giới hạn ở các kết nối trực tiếp. Thuật ngữ “được ghép nối” và “được kết nối,” cùng với các dạng khác của chúng có thể được sử dụng. Cần hiểu rằng các thuật ngữ này không được dùng làm các từ đồng nghĩa với nhau. Do vậy, phạm vi của cụm từ thiết bị A được ghép nối với thiết bị B không nên bị giới hạn ở các thiết bị hoặc hệ thống trong đó đầu ra của thiết bị A được nối trực tiếp với đầu vào của thiết bị B. Điều đó có nghĩa là có đường dẫn giữa đầu ra của A và đầu vào của B mà có thể là đường dẫn gồm các thiết bị hoặc phương tiện khác. “Ghép nối” có thể có nghĩa là hai hoặc nhiều phần tử tiếp xúc vật lý hoặc điện trực tiếp, hoặc hai hoặc nhiều phần tử không tiếp xúc trực tiếp với nhau nhưng vẫn phối hợp hoặc tương tác với nhau.

Do đó, mặc dù các phương án của sáng chế được mô tả, nhưng người có hiểu biết trung bình trong lĩnh vực kỹ thuật này nhận ra rằng các cải biến khác và cải biến tiếp có thể được thực hiện đối với các phương án đó mà không nằm ngoài mục đích của sáng chế, và dự tính yêu cầu bảo hộ tất cả các thay đổi và cải biến như vậy đều nằm trong phạm vi của sáng chế. Ví dụ, công thức bất kỳ được đưa ra trên đây chỉ đại diện cho các quy trình có thể được sử dụng. Chức năng có thể được bổ sung hoặc loại ra 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 xóa ra khỏi phương pháp được mô tả trong phạm vi của sáng chế.

Một số khía cạnh của sáng chế có thể được hiểu rõ qua các phương án làm ví dụ được liệt kê dưới đây (enumerated example embodiments - EEES):

EEE 1. Phương pháp mã hóa âm thanh đầu vào dựa trên kênh hoặc đối tượng để phát lại, phương pháp này bao gồm các bước:

(a) ban đầu kết xuất âm thanh đầu vào dựa trên kênh hoặc đối tượng thành dạng trình diễn đầu ra ban đầu;

(b) xác định giá trị ước tính của thành phần âm thanh nổi trội từ âm thanh đầu vào dựa trên kênh hoặc đối tượng và xác định một chuỗi hệ số gia trọng cho thành phần âm thanh nổi trội để ánh xạ dạng trình diễn đầu ra ban đầu thành thành phần âm thanh nổi trội;

(c) xác định giá trị ước tính của hướng hoặc vị trí của thành phần âm thanh nổi trội; và

(d) mã hóa dạng trình diễn đầu ra ban đầu, các hệ số gia trọng cho thành phần âm thanh nổi trội, hướng hoặc vị trí của thành phần âm thanh nổi trội dưới dạng tín hiệu mã hóa để phát lại.

EEE 2. Phương pháp theo EEE 1, phương pháp này còn bao gồm bước xác định giá trị ước tính của hỗn hợp dư là dạng trình diễn đầu ra ban đầu trừ đi phần kết xuất của thành phần âm thanh nổi trội hoặc giá trị ước tính của nó.

EEE 3. Phương pháp theo EEE 1, phương pháp này còn bao gồm bước tạo hỗn hợp âm thanh lập thể hai tai không tiếng vang của âm thanh đầu vào dựa trên kênh hoặc đối tượng, và xác định giá trị ước tính của hỗn hợp dư, trong đó giá trị ước tính của hỗn hợp dư là hỗn hợp âm thanh lập thể hai tai không tiếng vang trừ đi phần kết xuất của thành phần âm thanh nổi trội hoặc giá trị ước tính của nó.

EEE 4. Phương pháp theo EEE 2 hoặc 3, phương pháp này còn bao gồm bước xác định một chuỗi hệ số ma trận dư để ánh xạ dạng trình diễn đầu ra ban đầu thành giá trị ước tính của hỗn hợp dư.

EEE 5. Phương pháp theo EEE nêu trên bất kỳ trong đó dạng trình diễn đầu ra ban đầu bao gồm dạng trình diễn qua tai nghe hoặc loa.

EEE 6. Phương pháp theo EEE nêu trên bất kỳ trong đó âm thanh đầu vào dựa trên kênh hoặc đối tượng được tạo ô thời gian và tần số và bước mã hóa được lặp lại cho một loạt các bước thời gian và một loạt băng tần số.

EEE 7. Phương pháp theo EEE nêu trên bất kỳ trong đó dạng trình diễn đầu ra ban đầu bao gồm hỗn hợp loa âm thanh nổi.

EEE 8. Phương pháp giải mã tín hiệu âm thanh được mã hóa, tín hiệu âm thanh được mã hóa này bao gồm:

- dạng trình diễn đầu ra thứ nhất;

- hướng của thành phần âm thanh nổi trội và các hệ số gia trọng cho thành phần âm thanh nổi trội;

phương pháp này bao gồm các bước:

- (a) sử dụng các hệ số gia trọng cho thành phần âm thanh nổi trội và dạng trình diễn đầu ra ban đầu để xác định thành phần nổi trội ước tính được;

- (b) kết xuất thành phần nổi trội ước tính được với sự tạo âm thanh lập thể hai tai ở vị trí không gian tương ứng với người nghe đã định theo hướng của thành phần âm thanh nổi trội để tạo ra thành phần nổi trội ước tính được tạo âm thanh lập thể hai tai được kết xuất;

- (c) tái tạo giá trị ước tính của thành phần dư từ dạng trình diễn đầu ra thứ nhất; và

- (d) kết hợp thành phần nổi trội ước tính được tạo âm thanh lập thể hai tai được kết xuất và giá trị ước tính của thành phần dư để tạo ra tín hiệu âm thanh mã hóa theo không gian được xuất ra.

EEE 9. Phương pháp theo EEE 8 trong đó tín hiệu âm thanh mã hóa còn bao gồm một chuỗi các hệ số ma trận dư biểu diễn tín hiệu âm thanh dư và bước (c) còn bao gồm:

(c1) áp dụng các hệ số ma trận dư cho dạng trình diễn đầu ra thứ nhất để tái tạo giá trị ước tính của thành phần dư.

EEE 10. Phương pháp theo EEE 8, trong đó giá trị ước tính của thành phần dư được tái tạo bằng cách lấy dạng trình diễn đầu ra thứ nhất trừ đi thành phần nhiễu ước tính được tạo âm thanh lập thể hai tai được kết xuất.

EEE 11. Phương pháp theo EEE 8 trong đó bước (b) bao gồm bước xoay ban đầu của thành phần nhiễu ước tính được theo tín hiệu đầu vào theo dõi chuyển động đầu chỉ báo hướng đầu của người nghe đã định.

EEE 12. Phương pháp giải mã và tái tạo dòng âm thanh cho người nghe bằng cách sử dụng tai nghe, phương pháp này bao gồm các bước:

(a) thu dòng dữ liệu chứa dạng biểu diễn âm thanh thứ nhất và dữ liệu biến đổi âm thanh bổ sung;

(b) thu dữ liệu hướng đầu thể hiện hướng của người nghe;

(c) tạo ra một hoặc nhiều tín hiệu phụ dựa vào dạng biểu diễn âm thanh thứ nhất và dữ liệu biến đổi thu được;

(d) tạo ra dạng biểu diễn âm thanh thứ hai chứa tổ hợp của dạng biểu diễn âm thanh thứ nhất và (các) tín hiệu phụ, trong đó một hoặc nhiều trong số (các) tín hiệu phụ được sửa đổi để đáp lại dữ liệu hướng đầu; và

(e) xuất ra dạng biểu diễn âm thanh thứ hai dưới dạng dòng âm thanh đầu ra.

EEE 13. Phương pháp theo EEE 12, trong đó bước sửa đổi các tín hiệu phụ bao gồm mô phỏng đường âm thanh từ vị trí nguồn âm thanh đến tai người nghe.

EEE 14. Phương pháp theo EEE 12 hoặc 13, trong đó dữ liệu biến đổi bao gồm các hệ số ma trận và ít nhất một trong số: vị trí nguồn âm thanh hoặc hướng nguồn âm thanh.

EEE 15. Phương pháp theo EEE bất kỳ trong số các EEE từ 12 đến 14, trong đó quy trình biến đổi được áp dụng dưới dạng hàm thời gian hoặc tần số.

EEE 16. Phương pháp theo EEE bất kỳ trong số các EEE từ 12 đến 15, trong đó các tín hiệu phụ thể hiện ít nhất một thành phần nổi trội.

EEE 17. Phương pháp theo EEE bất kỳ trong số các EEE từ 12 đến 16, trong đó hướng hoặc vị trí nguồn âm thanh thu được là một phần của dữ liệu biến đổi được xoay đáp lại dữ liệu hướng đầu.

EEE 18. Phương pháp theo EEE 17, trong đó mức độ xoay tối đa bị hạn chế ở giá trị nhỏ hơn 360 độ theo góc phương vị hoặc độ cao.

EEE 19. Phương pháp theo EEE bất kỳ trong số các EEE từ 12 đến 18, trong đó dạng biểu diễn thứ hai thu được từ dạng biểu diễn thứ nhất bằng cách tạo ma trận trong miền biến đổi hoặc miền khối lọc.

EEE 20. Phương pháp theo EEE bất kỳ trong số các EEE từ 12 đến 19, trong đó dữ liệu biến đổi còn bao gồm các hệ số ma trận bổ sung, và bước (d) còn bao gồm sửa đổi dạng trình diễn âm thanh thứ nhất đáp lại các hệ số ma trận bổ sung trước khi kết hợp dạng trình diễn âm thanh thứ nhất và (các) tín hiệu âm thanh phụ.

EEE 21. Bộ máy, bao gồm một hoặc nhiều thiết bị, được tạo cấu hình để thực hiện phương pháp theo EEE bất kỳ trong số các EEE từ 1 đến 20.

EEE 22. Phương tiện lưu trữ đọc được bằng máy tính bao gồm chương trình chứa các lệnh mà, khi được thực thi bởi một hoặc nhiều bộ xử lý, khiến cho một hoặc nhiều thiết bị thực hiện phương pháp theo EEE bất kỳ trong số các EEE từ 1 đến 20.

Tài liệu tham khảo

Gundry, K., "A New Matrix Decoder for Surround Sound," AES 19th International Conf., Schloss Elmau, Germany, 2001.

Vinton, M., McGrath, D., Robinson, C., Brown, P., "Next generation surround decoding and up-mixing for consumer and professional applications", AES 57th International Conf, Hollywood, CA, USA, 2015.

Wightman, F. L., và Kistler, D. J. (1989). "Headphone simulation of free-field listening. I. Stimulus synthesis," J. Acoust. Soc. Am. 85, 858–867.

ISO/IEC 14496-3:2009 - Information technology -- Coding of audio-visual objects -- Part 3: Audio, 2009.

Mania, Katerina và cộng sự. "Perceptual sensitivity to head tracking latency in virtual environments with varying degrees of scene complexity." Proceedings of the 1st Symposium on Applied perception in graphics and visualization. ACM, 2004.

Allison, R. S., Harris, L. R., Jenkin, M., Jasiobedzka, U., & Zacher, J. E. (3/2001). Tolerance of temporal delay in virtual environments. In Virtual Reality, 2001. Proceedings. IEEE (pp. 247-254). IEEE.

Van de Par, Steven, và Armin Kohlrausch. "Sensitivity to auditory-visual asynchrony and to jitter in auditory-visual timing." Electronic Imaging. International Society for Optics và Photonics, 2000.

Yêu cầu bảo hộ

1. Hệ thống được tạo cấu hình để mã hóa âm thanh đầu vào dựa trên kênh hoặc đối tượng để phát lại, hệ thống này bao gồm:

một hoặc nhiều bộ xử lý; và

phương tiện đọc được bằng máy tính lưu trữ các lệnh mà, khi được thực thi bằng một hoặc nhiều bộ xử lý, khiến cho một hoặc nhiều bộ xử lý thực hiện các hoạt động bao gồm:

kết xuất âm thanh đầu vào dựa trên kênh hoặc đối tượng thành dạng trình diễn đầu ra ban đầu;

xác định giá trị ước tính của thành phần âm thanh nổi trội từ âm thanh đầu vào dựa trên kênh hoặc đối tượng, bước xác định này bao gồm:

xác định một chuỗi hệ số gia trọng cho thành phần âm thanh nổi trội để ánh xạ dạng trình diễn đầu ra ban đầu thành thành phần âm thanh nổi trội; và

xác định giá trị ước tính của thành phần âm thanh nổi trội dựa vào các hệ số gia trọng cho thành phần âm thanh nổi trội và dạng biểu diễn đầu ra ban đầu;

xác định giá trị ước tính của hướng hoặc vị trí của thành phần âm thanh nổi trội; và

mã hóa dạng trình diễn đầu ra ban đầu, các hệ số gia trọng cho thành phần âm thanh nổi trội, và ít nhất một trong số hướng hoặc vị trí của thành phần âm thanh nổi trội dưới dạng tín hiệu mã hóa để phát lại.

2. Hệ thống theo điểm 1, trong đó các hoạt động còn bao gồm bước xác định giá trị ước tính của hỗn hợp dư là dạng trình diễn đầu ra ban đầu trừ đi phần kết xuất của thành phần âm thanh nổi trội hoặc giá trị ước tính của nó.

3. Hệ thống theo điểm 1, trong đó các hoạt động còn bao gồm bước tạo hỗn hợp âm thanh lập thể hai tai không tiếng vang của âm thanh đầu vào dựa trên kênh hoặc đối tượng, và xác định giá trị ước tính của hỗn hợp dư, trong đó giá trị ước tính của hỗn hợp

đư là hỗn hợp âm thanh lập thể hai tai không tiếng vang trừ đi phần kết xuất của thành phần âm thanh nổi trội hoặc giá trị ước tính của nó.

4. Hệ thống theo điểm 2, trong đó các hoạt động còn bao gồm bước xác định chuỗi hệ số ma trận đư để ánh xạ dạng trình diễn đầu ra ban đầu thành giá trị ước tính của hỗn hợp đư.

5. Hệ thống theo điểm 1, trong đó dạng trình diễn đầu ra ban đầu bao gồm dạng trình diễn qua tai nghe hoặc dạng trình diễn qua loa.

6. Hệ thống theo điểm 1, trong đó âm thanh đầu vào dựa trên kênh hoặc đối tượng được chia ô thời gian và tần số và bước mã hóa được lặp lại cho một loạt bước nhảy thời gian và một loạt băng tần số.

7. Hệ thống theo điểm 1, trong đó dạng trình diễn đầu ra ban đầu bao gồm hỗn hợp qua loa âm thanh nổi.

8. Hệ thống được tạo cấu hình để giải mã tín hiệu âm thanh, hệ thống này bao gồm:

một hoặc nhiều bộ xử lý; và

phương tiện đọc được bằng máy tính lưu trữ các lệnh mà, khi được thực thi bằng một hoặc nhiều bộ xử lý, khiến cho một hoặc nhiều bộ xử lý thực hiện các hoạt động bao gồm:

nhận tín hiệu âm thanh đã mã hóa, tín hiệu âm thanh đã mã hóa này bao gồm:

dạng trình diễn đầu ra ban đầu bao gồm dạng giảm mức âm thanh nổi;

hướng của thành phần âm thanh nổi trội; và

các hệ số gia trọng cho thành phần âm thanh nổi trội;

xác định thành phần âm thanh nổi trội dựa vào các hệ số gia trọng cho thành phần âm thanh nổi trội và dạng trình diễn đầu ra ban đầu;

tạo thành phần nổi trội ước tính được tạo âm thanh lập thể hai tai được kết xuất bao gồm kết xuất thành phần nổi trội ước tính được với sự tạo âm thanh lập thể hai tai ở vị trí không gian tương ứng với người nghe dự tính theo hướng của thành phần âm thanh nổi trội;

tái tạo giá trị ước tính của thành phần dư từ dạng trình diễn đầu ra ban đầu; và

tạo tín hiệu âm thanh đầu ra không gian hóa bằng cách kết hợp thành phần nổi trội ước tính được tạo âm thanh lập thể hai tai được kết xuất với giá trị ước tính của thành phần dư.

9. Hệ thống theo điểm 8, trong đó tín hiệu âm thanh mã hóa còn bao gồm một chuỗi hệ số ma trận dư thể hiện tín hiệu âm thanh dư và bước tái tạo giá trị ước tính của thành phần dư còn bao gồm bước: áp dụng các hệ số ma trận dư cho dạng trình diễn đầu ra ban đầu để tái tạo giá trị ước tính của thành phần dư.

10. Hệ thống theo điểm 8, trong đó giá trị ước tính của thành phần dư được tái tạo bằng cách lấy dạng trình diễn đầu ra ban đầu trừ đi thành phần nổi trội ước tính được tạo âm thanh lập thể hai tai được kết xuất.

11. Hệ thống theo điểm 8, trong đó bước tạo thành phần nổi trội ước tính được tạo âm thanh lập thể hai tai được kết xuất bao gồm bước xoay ban đầu thành phần nổi trội ước tính được theo tín hiệu đầu vào theo dõi chuyển động đầu chỉ báo hướng đầu của người nghe dự tính.

12. Hệ thống theo điểm 8, trong đó giá trị ước tính của thành phần dư được tái tạo bằng cách lấy dạng trình diễn đầu ra ban đầu trừ đi thành phần nổi trội ước tính được tạo âm thanh lập thể hai tai được kết xuất và trong đó bước tạo thành phần nổi trội ước tính được tạo âm thanh lập thể hai tai được kết xuất bao gồm bước xoay ban đầu thành phần nổi trội ước tính được theo tín hiệu đầu vào theo dõi chuyển động đầu chỉ báo hướng đầu của người nghe dự tính.

13. Phương tiện lưu trữ bất biến đọc được bằng máy tính chứa các lệnh mà, khi được thực thi bởi một hoặc nhiều bộ xử lý, khiến cho một hoặc nhiều bộ xử lý thực hiện các hoạt động bao gồm:

kết xuất âm thanh đầu vào dựa trên kênh hoặc đối tượng thành dạng trình diễn đầu ra ban đầu;

xác định giá trị ước tính của thành phần âm thanh nổi trội từ âm thanh đầu vào dựa trên kênh hoặc đối tượng, bước xác định này bao gồm:

xác định một chuỗi hệ số gia trọng cho thành phần âm thanh nổi trội để ánh xạ dạng trình diễn đầu ra ban đầu thành thành phần âm thanh nổi trội; và

xác định ước lượng thành phần âm thanh nổi trội dựa vào các hệ số gia trọng cho thành phần âm thanh nổi trội và dạng biểu diễn đầu ra ban đầu;

xác định giá trị ước tính của hướng hoặc vị trí của thành phần âm thanh nổi trội; và

mã hóa dạng trình diễn đầu ra ban đầu, các hệ số gia trọng cho thành phần âm thanh nổi trội, và ít nhất một trong số hướng hoặc vị trí của thành phần âm thanh nổi trội dưới dạng tín hiệu mã hóa để phát lại.

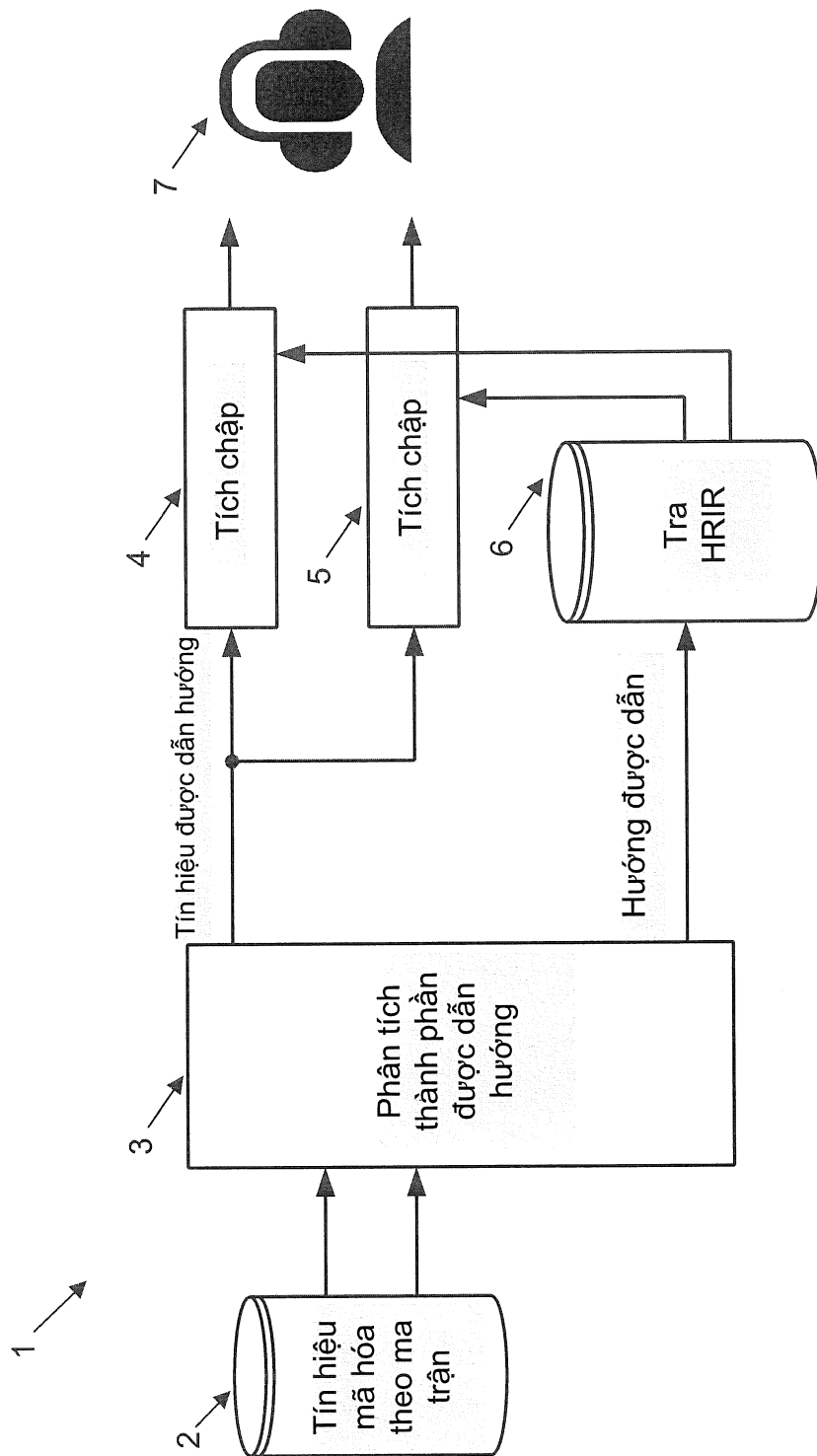


FIG. 1

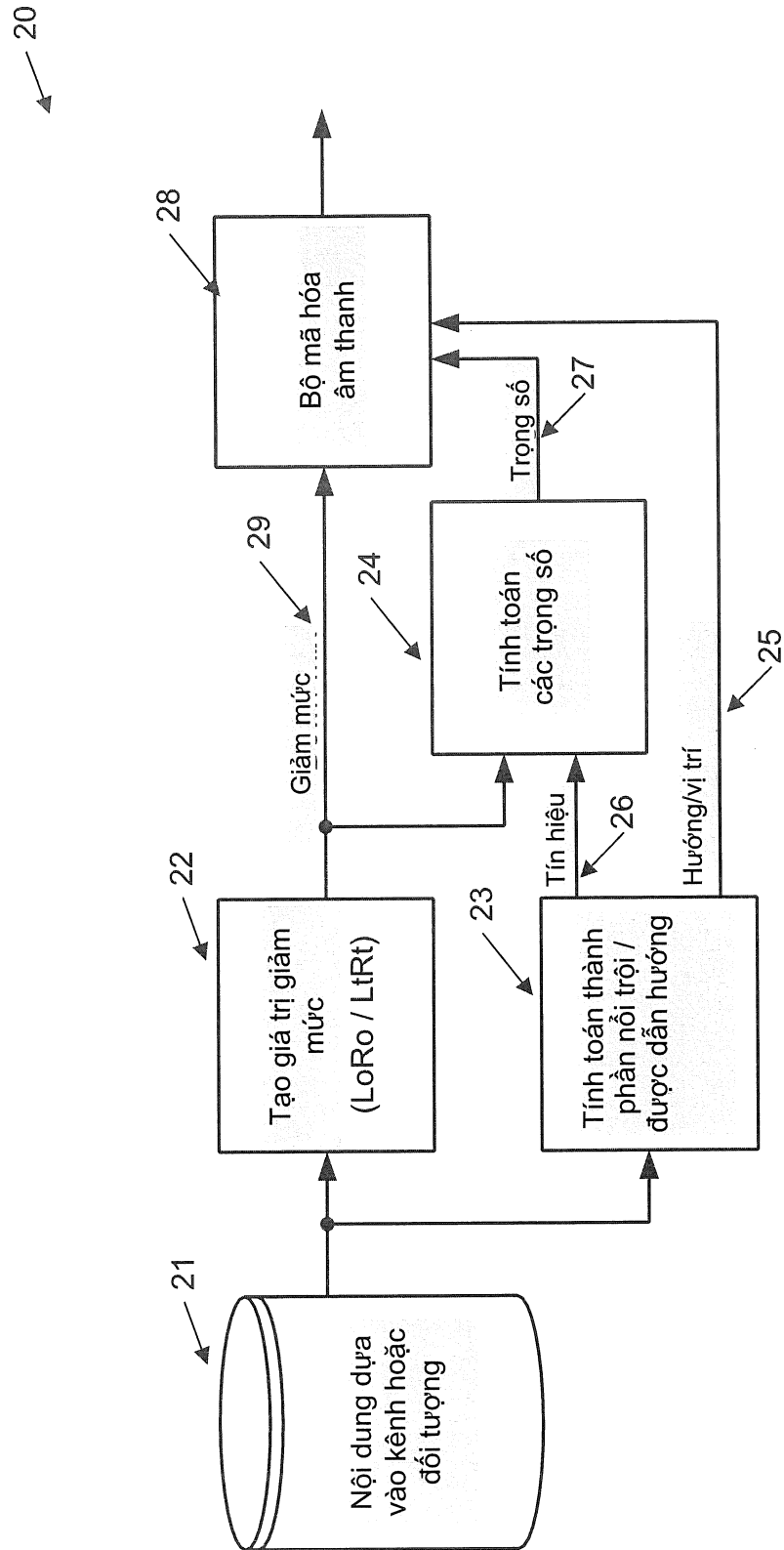


FIG. 2

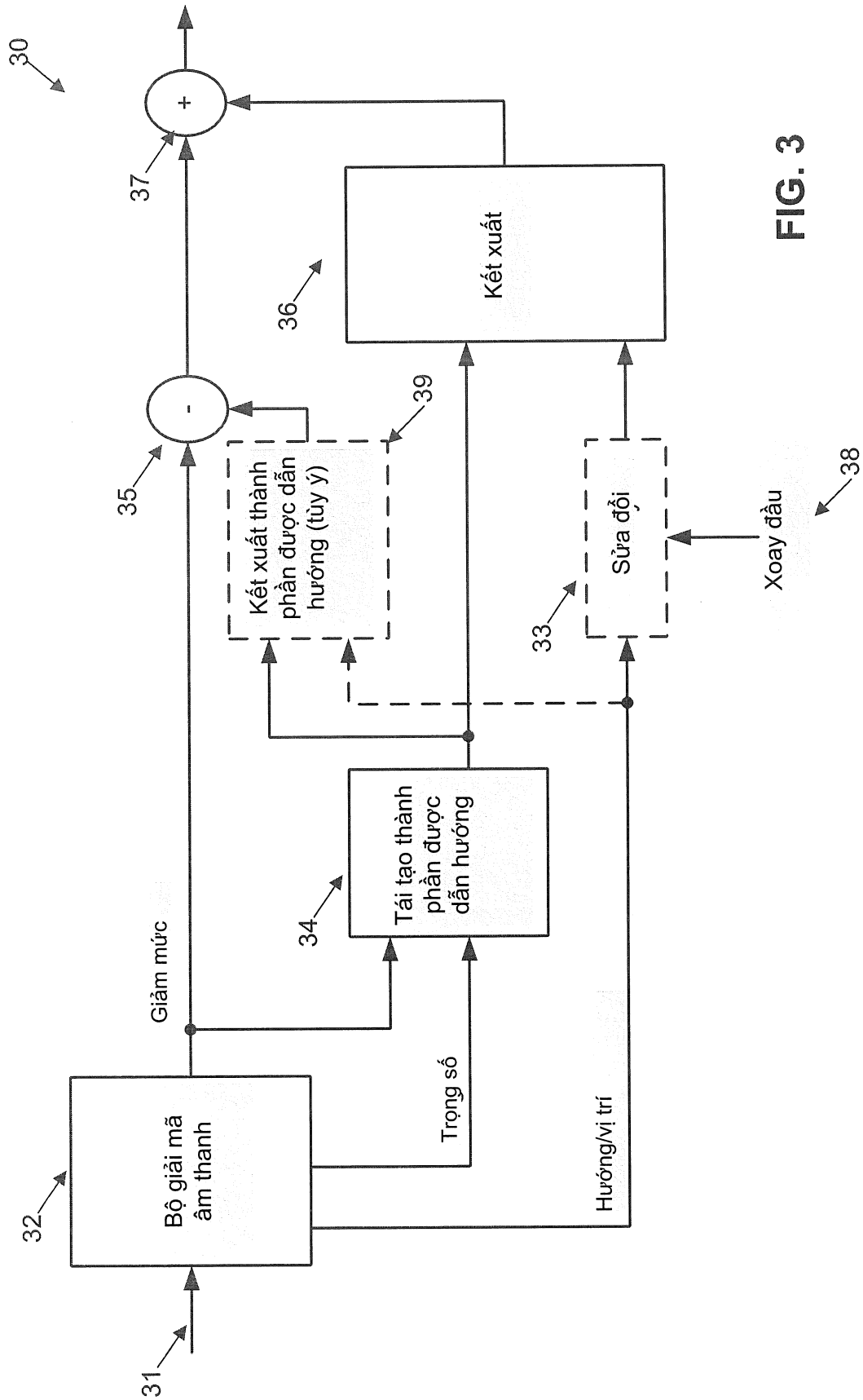


FIG. 3

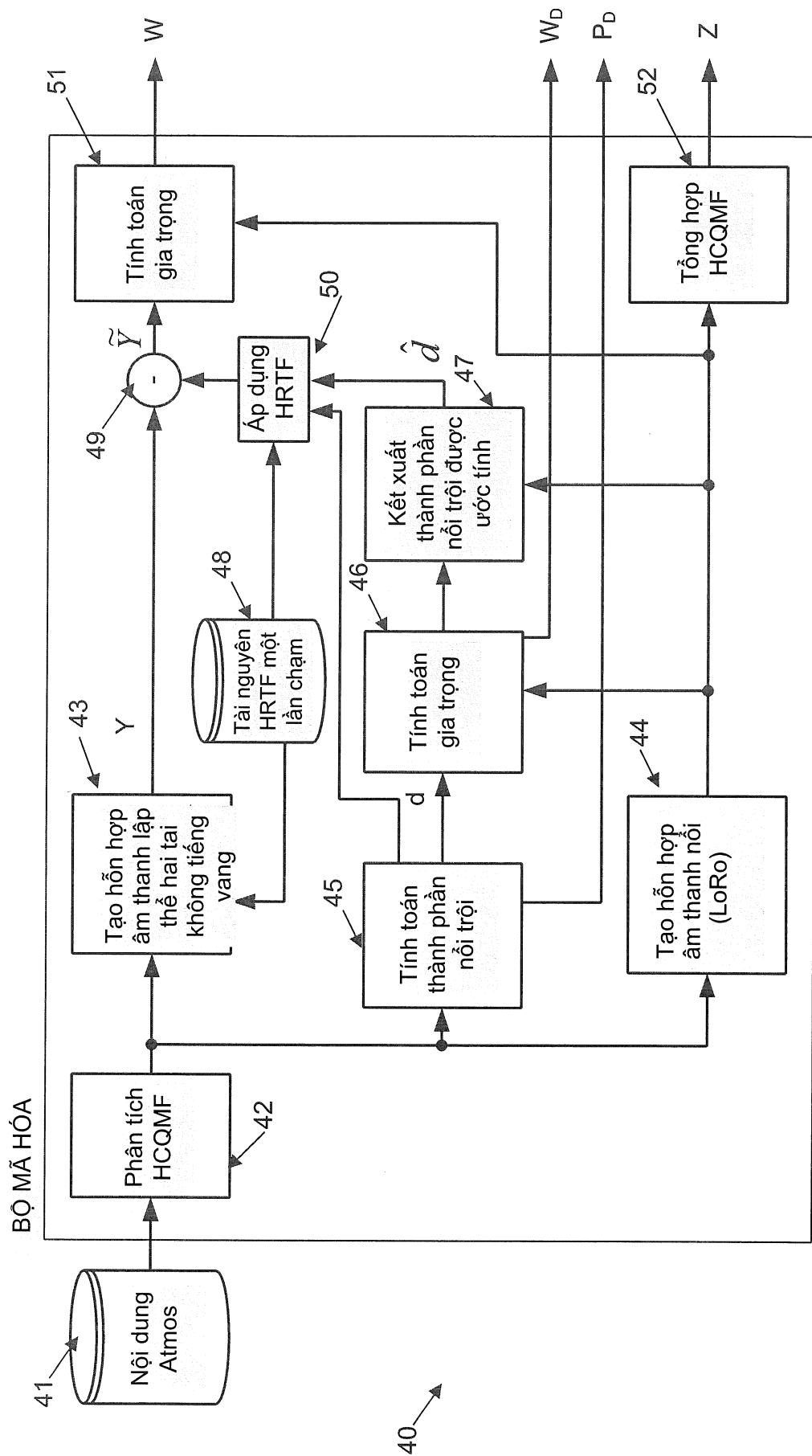


FIG. 4

