



(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)



1-0025861

(51)<sup>7</sup> G10L 19/008

(13) B

(21) 1-2016-00555

(22) 16/07/2014

(86) PCT/EP2014/065289 16/07/2014

(87) WO 2015/010998 A1 29/01/2015

(30) 13177378.0 22/07/2013 EP

(45) 26/10/2020 391

(43) 25/05/2016 338A

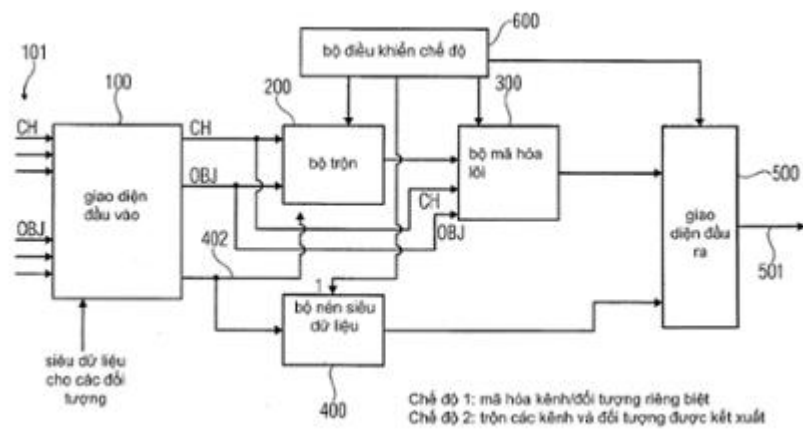
(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)  
Hansastrasse 27c, 80686 Muenchen, Germany

(72) ADAMI, Alexander (DE); BORSS, Christian (DE); DICK, Sascha (DE); ERTEL, Christian (DE); FUEG, Simone (DE); HERRE, Juergen (DE); HILPERT, Johannes (DE); HOELZER, Andreas (AT); KRATSCHMER, Michael (DE); KUECH, Fabian (DE); KUNTZ, Achim (DE); MURTAZA, Adrian (RO); PLOGSTIES, Jan (DE); SILZLE, Andreas (DE); STENZEL, Hanne (DE).

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

(54) BỘ GIẢI MÃ ÂM THANH, BỘ MÃ HÓA ÂM THANH, PHƯƠNG PHÁP MÃ HÓA DỮ LIỆU ĐẦU VÀO ÂM THANH, PHƯƠNG PHÁP GIẢI MÃ DỮ LIỆU ÂM THANH ĐƯỢC MÃ HÓA

(57) Sáng chế đề cập đến bộ mã hóa âm thanh, bộ giải mã âm thanh, phương pháp mã hóa dữ liệu đầu vào âm thanh, phương pháp giải mã dữ liệu âm thanh được mã hóa. Bộ mã hóa âm thanh để mã hóa dữ liệu đầu vào âm thanh (101) để thu được dữ liệu đầu ra âm thanh (501) bao gồm giao diện đầu vào (100) để nhận nhiều kênh âm thanh, nhiều đối tượng âm thanh và siêu dữ liệu liên quan đến một hoặc nhiều đối tượng âm thanh; bộ trộn (200) để trộn nhiều đối tượng và các kênh để thu được nhiều kênh được trộn trước, mỗi kênh được trộn trước bao gồm dữ liệu âm thanh của kênh và dữ liệu âm thanh của ít nhất một đối tượng; bộ mã hóa lõi (300) để mã hóa lõi dữ liệu đầu vào bộ mã hóa lõi; và bộ nén siêu dữ liệu (400) để nén siêu dữ liệu liên quan đến một hoặc nhiều đối tượng âm thanh trong số nhiều đối tượng âm thanh, trong đó bộ mã hóa âm thanh được cấu hình để hoạt động ở ít nhất một chế độ của nhóm gồm hai chế độ bao gồm chế độ thứ nhất, trong đó bộ mã hóa lõi được cấu hình để mã hóa nhiều kênh âm thanh và các đối tượng âm thanh được nhận bởi giao diện đầu vào như dữ liệu đầu vào bộ mã hóa lõi, và chế độ thứ hai, trong đó bộ mã hóa lõi (300) được cấu hình để nhận, như dữ liệu đầu vào bộ mã hóa lõi, nhiều kênh được trộn trước được tạo ra bởi bộ trộn (200).



### **Lĩnh vực kỹ thuật được đề cập**

Sáng chế đề cập đến việc mã hóa/giải mã âm thanh và, cụ thể là, sáng chế đề cập đến việc mã hóa âm thanh trong không gian và mã hóa đối tượng âm thanh trong không gian.

### **Tình trạng kỹ thuật của sáng chế**

Các công cụ mã hóa âm thanh trong không gian đã được biết đến trong lĩnh vực này và, ví dụ được chuẩn hóa theo tiêu chuẩn MPEG-bao quanh. Mã hóa âm thanh trong không gian bắt đầu từ các kênh đầu vào ban đầu như năm hoặc bảy kênh mà được nhận biết bởi vị trí của chúng trong việc thiết lập mô phỏng, tức là, kênh bên trái, kênh ở giữa, kênh bên phải, kênh bao quanh bên trái, kênh bao quanh bên phải và kênh tăng cường tần số thấp. Bộ mã hóa âm thanh trong không gian thường suy ra một hoặc nhiều kênh trộn giảm từ các kênh ban đầu và, ngoài ra, suy ra dữ liệu tham số liên quan đến các tín hiệu trong không gian như các chênh lệch mức liên kênh theo các trị liên kết kênh, các chênh lệch pha liên kênh, các chênh lệch thời gian liên kênh, v.v.. Một hoặc nhiều kênh trộn giảm được truyền cùng với thông tin phụ tham số biểu thị các tín hiệu trong không gian đến bộ giải mã âm thanh trong không gian mà giải mã kênh trộn giảm và dữ liệu tham số được kết hợp để cuối cùng thu được các kênh đầu ra mà là phiên bản gần đúng của các kênh đầu vào ban đầu. Cách bố trí của các kênh trong việc thiết lập đầu ra thường được cố định và là, ví dụ, định dạng 5.1, định dạng 7.1, v.v.

Ngoài ra, các công cụ mã hóa đối tượng âm thanh trong không gian đã được biết đến trong lĩnh vực kỹ thuật này và được chuẩn hóa theo tiêu chuẩn mã hóa đối tượng âm thanh trong không gian (SAOC = spatial audio object coding) MPEG. Trái với mã hóa âm thanh trong không gian bắt đầu từ các kênh ban đầu, việc mã hóa đối tượng âm thanh trong không gian bắt đầu từ các đối tượng âm thanh mà không chuyên

dụng theo cách tự động với thiết lập mô phỏng kết xuất nhất định. Thay vào đó, cách bố trí của các đối tượng âm thanh trong cảnh mô phỏng là linh hoạt và có thể được xác định bởi người dùng bằng cách nhập thông tin kết xuất nhất định vào trong bộ giải mã mã hóa đối tượng âm thanh trong không gian. Theo cách khác hoặc ngoài ra, thông tin kết xuất, tức là, thông tin mà tại đó vị trí trong mô phỏng thiết lập đối tượng âm thanh nhất định thường được đặt quá thời gian có thể được truyền như thông tin phụ bổ sung hoặc siêu dữ liệu. Để thu được việc nén dữ liệu nhất định, số lượng của các đối tượng âm thanh được mã hóa bằng bộ mã hóa SAOC mà tính toán, từ các đối tượng đầu vào, một hoặc nhiều kênh vận chuyển bằng cách trộn giảm các đối tượng phù hợp với thông tin trộn giảm nhất định. Hơn nữa, bộ mã hóa SAOC tính toán thông tin phụ tham số biểu diễn các tín hiệu liên đối tượng như các chênh lệch về mức đối tượng (object level difference - OLD), các trị liên kết đối tượng, v.v. Như trong mã hóa âm thanh trong không gian (Spatial Audio Coding - SAC), dữ liệu tham số liên đối tượng được tính toán cho các ô thời gian/tần số riêng biệt, tức là cho khung tín hiệu âm thanh nhất định bao gồm, ví dụ, 1024 hoặc 2048 mẫu, 24, 32, hoặc 64, v.v, các băng tần số được xét đến sao cho, cuối cùng, dữ liệu tham số tồn tại cho mỗi khung và mỗi băng tần số. Như ví dụ, khi đoạn âm thanh có 20 khung và khi mỗi khung được chia nhỏ thành 32 băng tần số, sau đó số lượng các ô tần số/ thời gian là 640.

Cho đến nay, chưa có kỹ thuật linh hoạt nào tồn tại một mặt là mã hóa kênh tổ hợp và mặt khác là mã hóa đối tượng sao cho thu được chất lượng âm thanh chấp nhận được ở tốc độ bit thấp.

### **Bản chất kỹ thuật của sáng chế**

Mục đích của sáng chế là đề xuất khái niệm được cải thiện để mã hóa âm thanh và giải mã âm thanh.

Mục đích này đạt được bởi bộ mã hóa âm thanh theo điểm 1, bộ giải mã âm thanh theo điểm 8, phương pháp mã hóa âm thanh theo điểm 22, phương pháp giải mã âm thanh theo điểm 23 hoặc chương trình máy tính theo điểm 24.

Sáng chế được dựa trên phát hiện ra rằng, đối với hệ thống tối ưu một mặt là tính linh hoạt và mặt khác là cung cấp hiệu quả nén tốt tại chất lượng âm thanh tốt đạt được bằng cách mã hóa âm thanh trong không gian tổ hợp, tức là, mã hóa âm thanh dựa trên kênh với mã hóa đối tượng âm thanh trong không gian, tức là, mã hóa dựa trên đối tượng. Cụ thể là, cung cấp bộ trộn để trộn các đối tượng và kênh đã ở phía bộ mã hóa cung cấp tính linh hoạt tốt, cụ thể là với các ứng dụng có tốc độ bit thấp, do sự truyền đối tượng bất kỳ sau đó có thể là không nhất thiết hoặc số đối tượng được truyền có thể bị giảm. Mặt khác, tính linh hoạt được yêu cầu sao cho bộ mã hóa âm thanh có thể được điều khiển ở hai chế độ khác nhau, tức là, ở chế độ mà các đối tượng được trộn với các kênh trước khi được mã hóa lỗi, trong khi ở chế độ còn lại, một mặt là dữ liệu đối tượng và mặt khác là dữ liệu kênh được mã hóa lỗi theo cách trực tiếp mà không có sự trộn bất kỳ nào ở giữa.

Việc này đảm bảo rằng người dùng có thể tách riêng các đối tượng được xử lý và các kênh ở phía bộ mã hóa sao cho tính linh hoạt hoàn toàn là sẵn có ở phía bộ giải mã nhưng, với điều kiện là tốc độ bit cao. Mặt khác, khi các yêu cầu về tốc độ bit nghiêm ngặt hơn, sáng chế đã cho phép thực hiện việc trộn /kết xuất trước ở phía bộ mã hóa, tức là, trộn /kết xuất trước một số hoặc tất cả đối tượng âm thanh đã được trộn với các kênh sao cho bộ mã hóa lỗi chỉ mã hóa dữ liệu kênh và các bit bất kỳ được yêu cầu để truyền dữ liệu đối tượng âm thanh ở dạng trộn giảm hoặc ở dạng dữ liệu liên đối tượng tham số là không được yêu cầu.

Trên phía bộ giải mã, người dùng cũng có sự linh hoạt cao một lần nữa do thực tế rằng cùng bộ giải mã âm thanh cho phép hoạt động ở hai chế độ khác nhau, tức là, chế độ thứ nhất mà việc mã hóa kênh và đối tượng riêng biệt hoặc tách biệt diễn ra và bộ giải mã có sự linh hoạt hoàn toàn để kết xuất các đối tượng và trộn với dữ liệu kênh. Mặt khác, khi việc trộn /kết xuất trước đã diễn ra trên phía bộ mã hóa, bộ giải mã được cấu hình để thực hiện việc xử lý sau mà không xử lý đối tượng trung gian bất kỳ. Mặt khác, việc xử lý sau cũng có thể được áp dụng với dữ liệu trong chế độ còn lại, tức là, khi việc kết xuất/trộn đối tượng diễn ra trên phía bộ giải mã. Do đó, sáng chế cho phép chương trình khung của các nhiệm vụ xử lý mà cho phép tái sử dụng các nguồn không chỉ trên phía bộ mã hóa mà còn trên phía bộ giải mã. Việc xử lý sau có

thể đề cập đến việc trộn giảm và lập thể hóa hoặc việc xử lý bất kỳ khác để thu được ngữ cảnh kênh cuối cùng như thiết kế mô phỏng được mong đợi.

Hơn nữa, trong trường hợp các yêu cầu tốc độ bit rất thấp, sáng chế cung cấp cho người dùng có đủ sự linh hoạt để phản ứng lại các yêu cầu tốc độ bit thấp, tức là, bằng cách kết xuất trước trên phía bộ mã hóa sao cho, đối với điều kiện một số tính linh hoạt, tuy nhiên thu được chất lượng âm thanh rất tốt trên phía bộ giải mã do thực tế rằng các bit mà đã được lưu bằng cách không cung cấp thêm dữ liệu đối tượng từ bộ mã hóa đến bộ giải mã có thể được sử dụng để mã hóa tốt hơn dữ liệu kênh như bằng cách lượng tử hóa mịn hơn dữ liệu kênh hoặc bằng các phương thức khác để cải thiện chất lượng hoặc để giảm việc mất mã hóa khi có đủ bit có thể dùng được.

Trong phương án được ưu tiên của sáng chế, bộ mã hóa còn bao gồm bộ mã hóa SAOC và còn cho phép không chỉ mã hóa các đối tượng đi vào bộ mã hóa mà còn mã hóa SAOC dữ liệu kênh để thu được chất lượng âm thanh tốt ở tốc độ bit yêu cầu thậm chí thấp hơn. Các phương án khác theo sáng chế cho phép chức năng xử lý sau mà bao gồm bộ kết xuất lập thể và/hoặc bộ chuyển đổi định dạng. Hơn nữa, được ưu tiên rằng toàn bộ quá trình xử lý trên phía bộ giải mã đã diễn ra đối với số lớn nhất định các loa phóng thanh như thiết lập loa phóng thanh kênh 22 hoặc 32. Tuy nhiên, sau đó bộ chuyển đổi định dạng, ví dụ, xác định rằng chỉ có đầu ra 5.1, tức là, đầu ra cho việc bố trí mô phỏng được yêu cầu mà có số lượng thấp hơn số lượng kênh tối đa, sau đó được ưu tiên rằng bộ chuyển đổi định dạng điều khiển bộ giải mã USAC hoặc bộ giải mã SAOC hoặc cả hai thiết bị để giới hạn hoạt động giải mã lỗi và hoạt động giải mã SAOC sao cho các kênh bất kỳ mà là, cuối cùng, tuy nhiên được trộn giảm thành sự chuyển đổi định dạng không được tạo ra trong giải mã. Thông thường, việc tạo ra các kênh trộn tăng yêu cầu việc xử lý giải tương quan và mỗi quá trình xử lý giải tương quan đưa vào một số mức giả âm. Do đó, bằng cách điều khiển bộ giải mã lỗi và/hoặc bộ giải mã SAOC bằng định dạng đầu ra được yêu cầu cuối cùng, mục đích to lớn của việc xử lý giải tương quan bổ sung được lưu trữ so với hoàn cảnh khi sự tương tác này không tồn tại mà không chỉ tạo ra chất lượng âm thanh được cải thiện mà còn dẫn đến độ phức tạp của bộ giải mã giảm và, cuối cùng, giảm sự tiêu thụ điện mà hữu dụng một cách đặc biệt cho các thiết bị di động chứa bộ mã hóa của sáng chế hoặc bộ giải

mã của sáng chế. Tuy nhiên, các bộ mã hóa/bộ giải mã của sáng chế không thể chỉ được đưa vào các thiết bị di động như điện thoại di động, điện thoại thông minh, máy tính cầm tay hoặc các thiết bị hàng hải mà còn có thể được sử dụng trong các máy tính để bàn đơn giản hoặc các thiết bị gia dụng không di động khác bất kỳ.

Việc thực hiện nêu trên, tức là không tạo ra một số kênh, có thể không tối ưu, do một số thông tin có thể bị mất (như chênh lệch mức giữa các kênh mà sẽ được trộn giảm). Thông tin chênh lệch mức này có thể không quan trọng, nhưng có thể dẫn đến tín hiệu đầu ra trộn giảm khác, nếu việc trộn giảm áp dụng các độ khuếch đại trộn giảm khác nhau đến các kênh trộn tăng. Giải pháp được cải thiện là chỉ bỏ sự giải tương quan trong trộn tăng, nhưng vẫn tạo ra các kênh trộn tăng với các chênh lệch mức chính xác (như được truyền tín hiệu bởi SAC tham số). Giải pháp thứ hai dẫn đến chất lượng âm thanh tốt hơn, nhưng giải pháp thứ nhất dẫn đến việc giảm độ phức tạp lớn hơn.

### **Mô tả vắn tắt các hình vẽ**

Các phương án được ưu tiên được thảo luận sau đây dựa trên các hình vẽ kèm theo, trong đó:

Fig.1 minh họa phương án thứ nhất của bộ mã hóa;

Fig.2 minh họa phương án thứ nhất của bộ giải mã;

Fig.3 minh họa phương án thứ hai của bộ mã hóa;

Fig.4 minh họa phương án thứ hai của bộ giải mã;

Fig.5 minh họa phương án thứ ba của bộ mã hóa;

Fig.6 minh họa phương án thứ ba của bộ giải mã;

Fig.7 minh họa bản đồ biểu thị các chế độ riêng biệt mà trong đó các bộ mã hóa/bộ giải mã phù hợp với các phương án có thể được vận hành;

Fig.8 minh họa sự thực hiện đặc biệt của bộ chuyển đổi định dạng;

Fig.9 minh họa sự thực hiện đặc biệt của bộ chuyển đổi lập thể;

Fig.10 minh họa phương án đặc biệt của bộ giải mã lỗi; và

Fig.11 minh họa phương án đặc biệt của bộ mã hóa để xử lý thành phần bốn kênh (quad channel element - QCE) và bộ giải mã QCE tương ứng.

### **Mô tả chi tiết sáng chế**

Fig.1 minh họa bộ mã hóa theo phương án của sáng chế. Bộ mã hóa được cấu hình để mã hóa dữ liệu đầu vào âm thanh 101 để thu được dữ liệu đầu ra âm thanh 501. Bộ mã hóa bao gồm giao diện đầu vào để nhận nhiều kênh âm thanh được biểu thị bởi CH và nhiều đối tượng âm thanh được biểu thị bởi OBJ. Hơn nữa, như được minh họa trong Fig.1, giao diện đầu vào 100 còn nhận siêu dữ liệu liên quan đến một hoặc nhiều đối tượng âm thanh trong số nhiều đối tượng âm thanh OBJ. Hơn nữa, bộ mã hóa bao gồm bộ trộn 200 để trộn nhiều đối tượng và nhiều kênh để thu được nhiều kênh được trộn trước, trong đó mỗi kênh được trộn trước bao gồm dữ liệu âm thanh của kênh và dữ liệu âm thanh của ít nhất một đối tượng.

Hơn nữa, bộ mã hóa bao gồm bộ mã hóa lõi 300 để mã hóa lõi dữ liệu đầu vào bộ mã hóa lõi, bộ nén siêu dữ liệu 400 để nén siêu dữ liệu liên quan đến một hoặc nhiều đối tượng âm thanh trong số nhiều đối tượng âm thanh. Hơn nữa, bộ mã hóa có thể bao gồm bộ điều khiển chế độ 600 để điều khiển bộ trộn, bộ mã hóa lõi và/hoặc giao diện đầu ra 500 trong một vài chế độ hoạt động, trong đó trong chế độ thứ nhất, bộ mã hóa lõi được cấu hình để mã hóa nhiều kênh âm thanh và nhiều đối tượng âm thanh được nhận bằng giao diện đầu vào 100 mà không có sự tương tác bất kỳ bởi bộ trộn, tức là, không có việc trộn bất kỳ bởi bộ trộn 200. Tuy nhiên, trong chế độ thứ hai mà trong đó bộ trộn 200 hoạt động, bộ mã hóa lõi mã hóa nhiều kênh được trộn, tức là, đầu ra được tạo ra bởi khối 200. Trong trường hợp sau, tốt hơn là không mã hóa thêm dữ liệu đối tượng bất kỳ. Thay vào đó, các vị trí biểu thị siêu dữ liệu của các đối tượng âm thanh đã được sử dụng bằng bộ trộn 200 để kết xuất các đối tượng lên các kênh như được biểu thị bởi siêu dữ liệu. Nói cách khác, bộ trộn 200 sử dụng siêu dữ liệu liên quan đến nhiều đối tượng âm thanh để kết xuất trước các đối tượng âm thanh và sau đó các đối tượng âm thanh được kết xuất trước được trộn với các kênh để thu được các kênh được trộn tại đầu ra của bộ trộn. Trong phương án này, các đối tượng bất kỳ có thể không nhất thiết được truyền và việc này cũng áp dụng cho siêu dữ liệu được nén như đầu ra bởi khối 400. Tuy nhiên, nếu không phải tất cả đối tượng đi vào giao diện 100 được trộn mà chỉ có lượng đối tượng nhất định được trộn, thì chỉ có các đối

tượng không được trộn còn lại và siêu dữ liệu được kết hợp tuy nhiên lần lượt được truyền đến bộ mã hóa lõi 300 hoặc bộ nén siêu dữ liệu 400.

Fig.3 minh họa phương án khác của bộ mã hóa mà, ngoài ra, còn bao gồm bộ mã hóa SAOC 800. Bộ mã hóa SAOC 800 được cấu hình để tạo ra một hoặc nhiều kênh vận chuyển và dữ liệu tham số từ dữ liệu đầu vào bộ mã hóa đối tượng âm thanh trong không gian. Như được minh họa trên Fig.3, dữ liệu đầu vào bộ mã hóa đối tượng âm thanh trong không gian là các đối tượng không được xử lý bởi bộ kết xuất trước/bộ trộn. Theo cách khác, miễn là bộ kết xuất trước/bộ trộn chuyển dòng đã bị bỏ qua như ở chế độ một mà việc mã hóa kênh/đối tượng riêng biệt hoạt động, tất cả đối tượng đi vào giao diện đầu vào 100 được mã hóa bởi bộ mã hóa SAOC 800.

Hơn nữa, như được minh họa trên Fig.3, bộ mã hóa lõi 300 tốt hơn là được thực hiện như bộ mã hóa USAC, tức là, bộ mã hóa như được xác định và tiêu chuẩn hóa theo tiêu chuẩn MPEG-USAC (USAC = unified speech and audio coding). Đầu ra của toàn bộ bộ mã hóa được minh họa trên Fig.3 là dòng dữ liệu MPEG 4 có cấu trúc như vật chứa đối với các loại dữ liệu riêng biệt. Hơn nữa, siêu dữ liệu được biểu thị là dữ liệu “OAM” và bộ nén siêu dữ liệu 400 trên Fig.1 tương ứng với bộ mã hóa OAM 400 để thu được dữ liệu OAM được nén được nhập vào bộ mã hóa USAC 300 mà, như có thể được thấy trên Fig.3, còn bao gồm giao diện đầu ra để thu được dòng dữ liệu đầu ra MP4 không chỉ có dữ liệu kênh/đối tượng được mã hóa mà còn có dữ liệu OAM được nén.

Fig.5 minh họa phương án khác của bộ mã hóa, mà trái với Fig.3, bộ mã hóa SAOC có thể được cấu hình để mã hóa, với thuật toán mã hóa SAOC, các kênh được cung cấp tại bộ kết xuất trước /bộ trộn 200 không hoạt động ở chế độ này hoặc, theo cách khác, để mã hóa SAOC các kênh được kết xuất trước cùng với các đối tượng. Do đó, trên Fig.5, bộ mã hóa SAOC 800 có thể hoạt động trên ba loại dữ liệu đầu vào khác nhau, tức là, các kênh không có các đối tượng được kết xuất trước bất kỳ, các kênh và các đối tượng được kết xuất trước hoặc riêng các đối tượng. Hơn nữa, tốt hơn là cung cấp bộ giải mã OAM 420 bổ sung trên Fig.5 sao cho bộ mã hóa SAOC 800 sử

dụng, cho việc xử lý của nó, cùng dữ liệu như trên phía bộ giải mã, tức là, dữ liệu thu được bởi sự giảm nén ngoại trừ hơn là dữ liệu OAM ban đầu.

Bộ mã hóa của Fig.5 có thể hoạt động trong một vài chế độ.

Ngoài các chế độ thứ nhất và thứ hai như được thảo luận trong ngữ cảnh của Fig.1, bộ mã hóa của Fig.5 còn có thể hoạt động trong chế độ thứ ba mà trong đó bộ mã hóa lỗi tạo ra một hoặc nhiều kênh vận chuyển từ các đối tượng riêng biệt khi bộ kết xuất trước /bộ trộn 200 không hoạt động. Theo cách khác hoặc ngoài ra, trong chế độ thứ ba này, bộ mã hóa SAOC 800 có thể tạo ra một hoặc nhiều kênh vận chuyển khác hoặc bổ sung từ các kênh ban đầu, tức là, một lần nữa khi bộ kết xuất trước/bộ trộn 200 tương ứng với bộ trộn 200 của Fig.1 không hoạt động.

Cuối cùng, bộ mã hóa SAOC 800 có thể mã hóa, khi bộ mã hóa được cấu hình trong chế độ thứ tư, các kênh cùng với các đối tượng được kết xuất trước như được tạo ra bởi bộ kết xuất trước/bộ trộn. Do đó, trong chế độ thứ tư, các ứng dụng có tốc độ bit thấp nhất sẽ cung cấp chất lượng tốt do thực tế là các kênh và các đối tượng đã được biến đổi hoàn toàn thành các kênh vận chuyển SAOC riêng biệt và thông tin phụ được kết hợp như được biểu thị trên các Fig.3 và 5 là “SAOC-SI” và, ngoài ra, siêu dữ liệu được nén bất kỳ không phải được truyền trong chế độ thứ tư này.

Fig.2 minh họa bộ giải mã theo phương án của sáng chế. Bộ giải mã nhận, như đầu vào, dữ liệu âm thanh được mã hóa, tức là, dữ liệu 501 của Fig.1.

Bộ giải mã bao gồm bộ giải nén siêu dữ liệu 1400, bộ giải mã lõi 1300, bộ xử lý đối tượng 1200, bộ điều khiển chế độ 1600 và bộ xử lý sau 1700.

Cụ thể là, bộ giải mã âm thanh được cấu hình để giải mã dữ liệu âm thanh được mã hóa và giao diện đầu vào được cấu hình để nhận dữ liệu âm thanh được mã hóa, dữ liệu âm thanh được mã hóa bao gồm các kênh được mã hóa và nhiều đối tượng được mã hóa và siêu dữ liệu được nén liên quan đến nhiều đối tượng trong chế độ nhất định.

Hơn nữa, bộ giải mã lõi 1300 được cấu hình để giải mã nhiều kênh được mã hóa và nhiều đối tượng được mã hóa và, ngoài ra, bộ giải nén siêu dữ liệu được cấu hình để giải nén siêu dữ liệu được nén.

Hơn nữa, bộ xử lý đối tượng 1200 được cấu hình để xử lý nhiều đối tượng được giải mã như được tạo ra bởi bộ giải mã lõi 1300 bằng cách sử dụng siêu dữ liệu được giải nén để thu được số lượng kênh đầu ra được xác định trước bao gồm dữ liệu đối tượng và các kênh được giải mã. Các kênh đầu ra này như được biểu thị ở 1205 sau đó được nhập vào bộ xử lý sau 1700. Bộ xử lý sau 1700 được cấu hình để chuyển đổi số lượng các kênh đầu ra 1205 thành định dạng đầu ra nhất định mà có thể là định dạng đầu ra lập thể hoặc định dạng đầu ra loa phóng thanh như định dạng đầu ra 5.1, 7.1, v.v.

Tốt hơn là, bộ giải mã bao gồm bộ điều khiển chế độ 1600 được cấu hình để phân tích dữ liệu được mã hóa để phát hiện chỉ dẫn chế độ. Do đó, bộ điều khiển chế độ 1600 được kết nối với giao diện đầu vào 1100 trên Fig.2. Tuy nhiên, theo cách khác, bộ điều khiển chế độ không nhất thiết phải ở đó. Thay vào đó, bộ giải mã linh hoạt có thể được thiết lập trước bằng loại dữ liệu điều khiển khác bất kỳ như đầu vào người dùng hoặc điều khiển khác bất kỳ. Bộ giải mã âm thanh trên Fig.2 và, tốt hơn là được điều khiển bằng bộ điều khiển chế độ 1600, được cấu hình để bỏ qua bộ xử lý đối tượng và để nạp nhiều kênh được giải mã vào bộ xử lý sau 1700. Việc này hoạt động ở chế độ 2, tức là, trong đó chỉ các kênh được kết xuất trước được nhận, tức là, khi chế độ 2 đã được áp dụng trong bộ mã hóa của Fig.1. Theo cách khác, khi chế độ 1 đã được áp dụng trên bộ mã hóa, tức là, khi bộ mã hóa đã thực hiện việc mã hóa kênh/đối tượng riêng biệt, thì bộ xử lý đối tượng 1200 không bị bỏ qua, nhưng nhiều kênh được giải mã và nhiều đối tượng được giải mã được nạp vào bộ xử lý đối tượng 1200 cùng với siêu dữ liệu được giải nén được tạo ra bởi bộ giải nén siêu dữ liệu 1400.

Tốt hơn là, chỉ dẫn liệu chế độ 1 hoặc chế độ 2 được áp dụng là được bao gồm trong dữ liệu âm thanh được mã hóa và sau đó bộ điều khiển chế độ 1600 phân tích dữ liệu được mã hóa để phát hiện chỉ dẫn chế độ. Chế độ 1 được sử dụng khi chỉ dẫn chế độ biểu thị rằng dữ liệu âm thanh được mã hóa bao gồm các kênh được mã hóa và các

đối tượng được mã hóa và chế độ 2 được áp dụng khi chỉ dẫn chế độ biểu thị rằng dữ liệu âm thanh được mã hóa không chứa các đối tượng âm thanh bất kỳ, tức là, chỉ chứa các kênh được kết xuất trước thu được bởi chế độ 2 của bộ mã hóa trên Fig.1.

Fig.4 minh họa phương án được ưu tiên so với bộ giải mã của Fig.2 và phương án của Fig.4 tương ứng với bộ mã hóa của Fig.3. Ngoài việc thực hiện bộ giải mã của Fig.2, bộ giải mã trên Fig.4 bao gồm bộ giải mã SAOC 1800. Hơn nữa, bộ xử lý đối tượng 1200 của Fig.2 được thực hiện như bộ kết xuất đối tượng riêng biệt 1210 và bộ trộn 1220 trong khi, phụ thuộc vào chế độ, chức năng của bộ kết xuất đối tượng 1210 cũng có thể được thực hiện bằng bộ giải mã SAOC 1800.

Hơn nữa, bộ xử lý sau 1700 có thể được thực hiện là bộ kết xuất lập thể 1710 hoặc bộ chuyển đổi định dạng 1720. Theo cách khác, đầu ra trực tiếp của dữ liệu 1205 của Fig.2 cũng có thể được thực hiện như được minh họa bởi 1730. Do đó, tốt hơn là thực hiện việc xử lý ở bộ giải mã ở số lượng kênh cao nhất như 22.2 hoặc 32 để có tính linh hoạt và sau đó xử lý sau nếu định dạng nhỏ hơn được yêu cầu. Tuy nhiên, khi trở nên rõ ràng từ ngay lúc bắt đầu là chỉ định dạng nhỏ như định dạng 5.1 được yêu cầu, thì tốt hơn là, như được biểu thị bởi Fig.2 hoặc 6 bằng phím tắt 1727, mà sự điều khiển nhất định trên bộ giải mã SAOC và/hoặc bộ giải mã USAC có thể được áp dụng để tránh các hoạt động trộn tăng không cần thiết và các hoạt động trộn giảm sau đó.

Trong phương án được ưu tiên của sáng chế, bộ xử lý đối tượng 1200 bao gồm bộ giải mã SAOC 1800 và bộ giải mã SAOC được cấu hình để giải mã một hoặc nhiều kênh vận chuyển được xuất ra bởi bộ giải mã lõi và dữ liệu tham số được kết hợp và sử dụng siêu dữ liệu được giải nén để thu được nhiều đối tượng âm thanh được kết xuất. Để đạt được mục đích này, đầu ra OAM được kết nối với hộp 1800.

Hơn nữa, bộ xử lý đối tượng 1200 được cấu hình để kết xuất các đối tượng được giải mã được xuất ra bởi bộ giải mã lõi mà không được mã hóa trong các kênh vận chuyển SAOC nhưng được mã hóa riêng biệt ở các thành phần được tạo kênh đơn giản hình như được biểu thị bởi bộ kết xuất đối tượng 1210. Hơn nữa, bộ giải mã bao gồm giao diện đầu ra tương ứng với đầu ra 1730 để xuất ra đầu ra của bộ trộn đến các loa phóng thanh.

Theo phương án khác, bộ xử lý đối tượng 1200 bao gồm bộ giải mã mã hóa đối tượng âm thanh trong không gian 1800 để giải mã một hoặc nhiều kênh vận chuyển và thông tin phụ tham số được kết hợp biểu diễn các đối tượng âm thanh được mã hóa hoặc các kênh âm thanh được mã hóa, trong đó bộ giải mã mã hóa đối tượng âm thanh trong không gian được cấu hình để chuyển mã thông tin tham số được kết hợp và siêu dữ liệu được giải nén thành thông tin phụ tham số được chuyển mã có thể sử dụng để kết xuất trực tiếp định dạng đầu ra, như ví dụ được xác định ở phiên bản trước của SAOC. Bộ xử lý sau 1700 được cấu hình để tính toán các kênh âm thanh của định dạng đầu ra sử dụng các kênh vận chuyển được giải mã và thông tin phụ tham số được chuyển mã. Việc xử lý được thực hiện bởi bộ xử lý sau có thể tương tự với việc xử lý quanh MPEG hoặc có thể là việc xử lý khác bất kỳ như xử lý BCC hoặc tương tự.

Trong phương án khác, bộ xử lý đối tượng 1200 bao gồm bộ giải mã mã hóa đối tượng âm thanh trong không gian 1800 được cấu hình để trộn tăng và kết xuất trực tiếp các tín hiệu kênh cho định dạng đầu ra sử dụng các kênh vận chuyển được giải mã (bằng bộ giải mã lõi) và thông tin phụ tham số.

Hơn nữa, và quan trọng là, bộ xử lý đối tượng 1200 trên Fig.2 còn bao gồm bộ trộn 1220 mà nhận, như đầu vào, dữ liệu được xuất ra bởi bộ giải mã USAC 1300 trực tiếp khi các đối tượng được kết xuất trước được trộn với các kênh tồn tại, tức là, khi bộ trộn 200 trên Fig.1 hoạt động. Ngoài ra, bộ trộn 1220 nhận dữ liệu từ bộ kết xuất đối tượng thực hiện kết xuất đối tượng mà không giải mã SAOC. Hơn nữa, bộ trộn nhận dữ liệu đầu ra bộ giải mã SAOC, tức là, các đối tượng được kết xuất SAOC.

Bộ trộn 1220 được kết nối với giao diện đầu ra 1730, bộ kết xuất lập thể 1710 và bộ chuyển đổi định dạng 1720. Bộ kết xuất lập thể 1710 được cấu hình để kết xuất các kênh đầu ra thành hai kênh lập thể bằng cách sử dụng các hàm truyền liên quan đến vùng tiêu đề hoặc phản ứng xung phòng lập thể (binaural room impulse response - BRIR). Bộ chuyển đổi định dạng 1720 được cấu hình để chuyển đổi các kênh đầu vào định dạng đầu ra có số kênh thấp hơn so với các kênh đầu ra 1205 của bộ trộn và bộ chuyển đổi định dạng 1720 yêu cầu thông tin ở thiết kế bố trí như các loa 5.1 hoặc tương tự.

Bộ giải mã của Fig.6 khác với bộ giải mã của Fig.4 ở chỗ bộ giải mã SAOC không thể chỉ tạo ra các đối tượng được kết xuất mà còn các kênh được kết xuất và đây là trường hợp khi bộ mã hóa của Fig.5 đã được sử dụng và sự kết nối 900 giữa các kênh /các đối tượng được kết xuất trước và giao diện đầu vào của bộ mã hóa SAOC 800 hoạt động.

Hơn nữa, giai đoạn ghép nhẵn biên độ cơ sở vector (vector base amplitude panning-VBAP) 1810 được cấu hình mà nhận, từ bộ giải mã SAOC, thông tin ở thiết kế bố trí mô phỏng và xuất ra ma trận kết xuất đến bộ giải mã SAOC sao cho bộ giải mã SAOC có thể, cuối cùng, cung cấp các kênh được kết xuất mà không có hoạt động khác bất kỳ của bộ trộn ở định dạng kênh cao là 1205, tức là, 32 loa phóng thanh.

Khối VBAP tốt hơn là nhận dữ liệu OAM được giải mã để suy ra các ma trận kết xuất. Thông thường hơn, tốt hơn là yêu cầu thông tin hình học không chỉ của bố trí mô phỏng mà còn của các vị trí mà các tín hiệu đầu vào nên được kết xuất đến bố trí mô phỏng. Dữ liệu đầu vào hình học này có thể là dữ liệu OAM đối với các đối tượng hoặc thông tin vị trí kênh đối với các kênh được truyền bằng cách sử dụng SAOC.

Tuy nhiên, nếu chỉ giao diện đầu ra cụ thể được yêu cầu thì trạng thái VBAP 1810 có thể đã cung cấp ma trận kết xuất được yêu cầu cho, ví dụ, đầu ra 5.1. Sau đó, bộ giải mã SAOC 1800 thực hiện việc kết xuất trực tiếp từ các kênh vận chuyển SAOC, dữ liệu tham số có liên quan và siêu dữ liệu được giải nén, việc kết xuất trực tiếp vào định dạng đầu ra theo yêu cầu mà không có sự tương tác bất kỳ của bộ trộn 1220. Tuy nhiên, khi việc trộn nhất định giữa các chế độ được áp dụng, tức là, nếu một vài kênh được mã hóa SAOC nhưng không phải tất cả các kênh được mã hóa SAOC hoặc nếu một vài đối tượng được mã hóa SAOC nhưng không phải tất cả các đối tượng được mã hóa SAOC hoặc chỉ khi lượng đối tượng được kết xuất trước nhất định với các kênh được giải mã SAOC và các kênh còn lại được xử lý SAOC thì bộ trộn sẽ cùng đưa dữ liệu từ các phần đầu vào riêng biệt, tức là, trực tiếp từ bộ giải mã lõi 1300, từ bộ kết xuất đối tượng 1210 và từ bộ giải mã SAOC 1800.

Sau đó, Fig.7 được thảo luận để biểu thị các chế độ bộ mã hóa /bộ giải mã mà có thể được áp dụng bởi khái niệm bộ mã hóa/bộ giải mã âm thanh có chất lượng cao và linh hoạt cao.

Theo chế độ mã hóa thứ nhất, bộ trộn 200 trong bộ mã hóa của Fig.1 bị bỏ qua và, do đó, bộ xử lý đối tượng ở bộ giải mã của Fig.2 không được bỏ qua.

Trong chế độ thứ hai, bộ trộn 200 trên Fig.1 hoạt động và bộ xử lý đối tượng trên Fig.2 được bỏ qua.

Sau đó, trong chế độ mã hóa thứ ba, bộ mã hóa SAOC trên Fig.3 hoạt động nhưng chỉ có SAOC mã hóa các đối tượng ngoại trừ các kênh hoặc các kênh như được xuất ra bởi bộ trộn. Do đó, chế độ 3 yêu cầu, ở phía bộ giải mã được minh họa ở Fig.4, bộ giải mã SAOC chỉ hoạt động đối với các đối tượng và tạo ra các đối tượng được kết xuất.

Trong chế độ mã hóa thứ tư như được minh họa ở Fig.5, bộ mã hóa SAOC được cấu hình để mã hóa SAOC các kênh được kết xuất trước, tức là, bộ trộn hoạt động trong chế độ thứ hai. Trên phía bộ giải mã, việc giải mã SAOC được thực hiện đối với các đối tượng được kết xuất trước sao cho bộ xử lý đối tượng được bỏ qua như trong chế độ mã hóa thứ hai.

Hơn nữa, chế độ mã hóa thứ năm tồn tại có thể bằng hỗn hợp bất kỳ của các chế độ từ 1 đến 4. Cụ thể là, chế độ mã hóa hỗn hợp sẽ tồn tại khi bộ trộn 1220 trên Fig.6 nhận các kênh trực tiếp từ bộ giải mã USAC và, ngoài ra, nhận các kênh với các đối tượng được kết xuất trước từ bộ giải mã USAC. Hơn nữa, ở chế độ mã hóa hỗn hợp này, các đối tượng được mã hóa trực tiếp bằng cách sử dụng, tốt hơn là, một thành phần kênh của bộ giải mã USAC. Trong ngữ cảnh này, sau đó bộ kết xuất đối tượng 1210 sẽ kết xuất các đối tượng được giải mã này và chuyển chúng đến bộ trộn 1220. Hơn nữa, một vài đối tượng còn được mã hóa bằng bộ mã hóa SAOC sao cho bộ giải mã SAOC sẽ xuất các đối tượng được kết xuất đến bộ trộn và/hoặc các kênh được kết xuất khi một vài kênh được mã hóa bởi sự tồn tại của kỹ thuật.

Sau đó, mỗi phần đầu vào của bộ trộn 1220 có thể, theo cách minh họa, có ít nhất khả năng nhận một số kênh như 32 như được biểu thị ở 1205. Do đó, về cơ bản là, bộ trộn có thể nhận 32 kênh từ bộ giải mã USAC và, ngoài ra, 32 kênh được kết xuất trước/được trộn từ bộ giải mã USAC và, ngoài ra, 32 “kênh” từ bộ kết xuất đối tượng và, ngoài ra, 32 “kênh” từ bộ giải mã SAOC, một mặt nếu mỗi “kênh” giữa các khối 1210 và 1218 và mặt khác khối 1220 có sự phân bố của các đối tượng tương ứng ở kênh loa phóng thanh tương ứng và sau đó bộ trộn 1220 trộn, tức là, bổ sung các phân bố riêng biệt cho mỗi kênh loa phóng thanh.

Theo phương án được ưu tiên của sáng chế, hệ thống mã hóa /giải mã dựa trên bộ mã hóa-giải mã USAC MPEG-D để mã hóa kênh và các tín hiệu đối tượng. Để làm tăng hiệu quả mã hóa lượng đối tượng lớn, kỹ thuật SAOC MPEG được thích ứng. Ba loại bộ kết xuất thực hiện nhiệm vụ kết xuất các đối tượng đến các kênh, kết xuất các kênh đến tai nghe hoặc kết xuất các kênh đến thiết lập loa phóng thanh khác. Khi các tín hiệu đối tượng được truyền rõ ràng hoặc được mã hóa tham số hóa bằng cách sử dụng SAOC, thông tin siêu dữ liệu đối tượng tương ứng được nén và dồn kênh đầu vào dữ liệu đầu ra được mã hóa.

Theo một phương án, bộ kết xuất trước/bộ trộn 200 được sử dụng để chuyển đổi kênh cộng với ngữ cảnh nhập đối tượng vào kênh trước khi mã hóa. Về chức năng, giống với bộ kết xuất đối tượng/tổ hợp bộ trộn ở phía bộ giải mã như được minh họa ở Fig.4 hoặc Fig.6 và như được biểu thị bởi bộ xử lý đối tượng 1200 của Fig.2. Việc kết xuất trước các đối tượng đảm bảo entropy tín hiệu tất định ở đầu vào bộ mã hóa mà về cơ bản độc lập với số tín hiệu đối tượng hoạt động đồng thời. Với việc kết xuất trước các đối tượng, không yêu cầu truyền siêu dữ liệu đối tượng. Các tín hiệu đối tượng rời rạc được kết xuất đến sự thiết bố trí kênh mà bộ mã hóa được cấu hình sử dụng. Các trọng lượng của các đối tượng với mỗi kênh thu được từ siêu dữ liệu đối tượng có liên quan OAM như được biểu thị bởi mũi tên 402.

Do lỗi /bộ mã hóa/bộ giải mã đối với các tín hiệu kênh loa phóng thanh, các tín hiệu đối tượng rời rạc, các tín hiệu trộn giảm đối tượng và các tín hiệu được kết xuất trước, kỹ thuật USAC được ưu tiên. Nó tiến hành mã hóa vô số tín hiệu bằng cách tạo

ra thông tin ánh xạ kênh và đối tượng (thông tin hình học và ngữ nghĩa của kênh đầu vào và chuyển đổi đối tượng). Thông tin ánh xạ này mô tả làm thế nào các kênh và các đối tượng đầu vào được ánh xạ đến các thành phần kênh USAC như được minh họa ở Fig.10, tức là, các thành phần cặp kênh (channel pair element - CPE), các thành phần kênh đơn (single channel element - SCE), các thành phần có bốn kênh (channel quad element - QCE) và thông tin tương ứng được truyền đến bộ giải mã lỗi từ bộ mã hóa lỗi. Tất cả trọng tải như dữ liệu SAOC hoặc siêu dữ liệu đối tượng đi qua các thành phần mở rộng và được xem xét ở sự điều khiển tốc độ của bộ mã hóa.

Việc mã hóa các đối tượng có thể theo các cách khác nhau, tùy thuộc vào các yêu cầu vào độ méo/tốc độ và các yêu cầu về tính tương tác đối với bộ kết xuất. Các biến thể mã hóa đối tượng sau đây có thể thực hiện được:

- Các đối tượng được kết xuất trước: các tín hiệu của đối tượng được kết xuất trước và trộn thành các tín hiệu kênh 22.2 trước khi mã hóa. Chuỗi mã hóa sau đó tạo ra các tín hiệu kênh 22.2.
- Các dạng sóng đối tượng rời rạc: các đối tượng được cung cấp là các dạng sóng đơn âm đến bộ mã hóa. Bộ mã hóa sử dụng các thành phần kênh đơn SCE để truyền các đối tượng ngoài các tín hiệu kênh. Các đối tượng được giải mã được kết xuất và trộn tại phía bộ nhận. Thông tin siêu dữ liệu đối tượng được nén được truyền đến dọc theo bộ nhận/bộ kết xuất.
- Các dạng sóng đối tượng tham số: các thuộc tính của đối tượng và tính tương quan của chúng với nhau được mô tả bằng các tham số SAOC. Việc trộn giảm của các tín hiệu đối tượng được mã hóa bằng USAC. Thông tin tham số được truyền dọc theo. Số kênh trộn giảm được chọn tùy thuộc vào số đối tượng và toàn bộ tốc độ dữ liệu. Thông tin siêu dữ liệu đối tượng được nén được truyền đến bộ kết xuất SAOC.

Bộ mã hóa và bộ giải mã SAOC cho các tín hiệu đối tượng dựa trên kỹ thuật SAOC MPEG. Hệ thống này có khả năng tái tạo, biến đổi và kết xuất số đối tượng âm thành dựa trên số lượng nhỏ hơn của các kênh được truyền và dữ liệu tham số bổ sung

(OLD, IOC (Sự kết dính liên đối tượng), DMG (các độ khuếch đại trộn giảm)). Dữ liệu tham số bổ sung có tốc độ dữ liệu thấp hơn đáng kể so với yêu cầu để truyền tất cả các đối tượng riêng biệt làm cho việc mã hóa rất hiệu quả.

Bộ mã hóa SAOC đóng vai trò là nhập các tín hiệu đối tượng /kênh làm các dạng sóng đơn âm và xuất ra thông tin tham số (được gói thành dòng bit âm thanh ba chiều) và các kênh vận chuyển SAOC (được mã hóa bằng cách sử dụng các thành phần kênh đơn và được truyền).

Bộ giải mã SAOC tái cấu trúc các tín hiệu đối tượng /kênh từ các kênh vận chuyển SAOC được giải mã và thông tin tham số, và tạo ra ngữ cảnh âm thanh đầu ra dựa trên bố trí mô phỏng, thông tin siêu dữ liệu đối tượng được giải nén và tùy ý ở thông tin tương tác người dùng.

Đối với mỗi đối tượng, siêu dữ liệu có liên quan cụ thể hóa vị trí hình học và thể tích của đối tượng trong không gian ba chiều được mã hóa hiệu quả bằng cách lượng tự hóa các thuộc tính đối tượng theo thời gian và không gian. Siêu dữ liệu đối tượng được nén cOAM được truyền đến bộ nhận là thông tin phụ. Thể tích của đối tượng có thể bao gồm thông tin ở phạm vi trong không gian và/hoặc thông tin của mức tín hiệu của tín hiệu âm thanh của đối tượng âm thanh này.

Bộ kết xuất đối tượng ứng dụng siêu dữ liệu đối tượng được nén để tạo ra các dạng sóng đối tượng theo định dạng mô phỏng được đưa ra. Mỗi đối tượng được kết xuất thành các kênh đầu ra nhất định theo siêu dữ liệu của nó. Đầu ra của khối này tạo ra từ tổng của các kết quả từng phần.

Nếu cả hai kênh dựa trên nội dung cũng như các đối tượng rời rạc/tham số được giải mã, các dạng sóng dựa trên kênh và các dạng sóng đối tượng được kết xuất được trộn trước khi xuất ra các dạng sóng thu được (hoặc trước khi dẫn chúng đến mô đun bộ xử lý sau như bộ kết xuất lập thể hoặc mô đun bộ kết xuất loa phóng thanh).

Mô đun bộ kết xuất lập thể tạo ra việc trộn giảm đơn âm của vật liệu âm thanh đa kênh, sao cho mỗi kênh đầu vào được biểu diễn bởi nguồn âm thanh ảo. Việc xử lý

này được tiến hành theo khung theo miền trong bộ lọc gương vuông góc (Quadrature Mirror Filterbank - QMF).

Sự lập thể hóa dựa trên các đáp ứng xung phòng lập thể đo được

Fig.8 minh họa phương án được ưu tiên của bộ chuyển đổi định dạng 1720. Bộ kết xuất loa phóng thanh hoặc bộ chuyển đổi định dạng chuyển đổi giữa cấu hình kênh bộ truyền và định dạng mô phỏng mong muốn. Bộ chuyển đổi định dạng này thực hiện các chuyển đổi với số lượng kênh đầu ra thấp hơn, tức là, nó tạo ra các trộn giảm. Cuối cùng, bộ trộn giảm 1722 tốt hơn là hoạt động ở miền QMF nhận tín hiệu đầu ra của bộ trộn 1205 và xuất ra các tín hiệu của loa phóng thanh. Tốt hơn là, bộ điều khiển 1724 để cấu hình bộ trộn giảm 1722 được tạo ra để nhận, làm đầu vào điều khiển, bố trí đầu ra của bộ trộn, tức là, bố trí mà dữ liệu 1205 được xác định và bố trí mô phỏng mong muốn thường được nhập thành khối chuyển đổi định dạng 1720 được minh họa ở Fig.6. Dựa trên thông tin này, bộ điều khiển 1724 tốt hơn là tự tạo ra các ma trận trộn giảm đối với sự tổ hợp tạo ra của các định dạng đầu vào và đầu ra và áp dụng các ma trận các ma trận này ở khối bộ trộn giảm 1722 trong quá trình trộn giảm. Bộ chuyển đổi định dạng cho phép các cấu hình loa phóng thanh tiêu chuẩn cũng như các cấu hình ngẫu nhiên có các vị trí loa phóng thanh không theo tiêu chuẩn.

Như được minh họa trong ngữ cảnh của Fig.6, bộ giải mã SAOC được thiết kế để kết xuất thiết kế bố trí kênh được xác định trước như 22.2 có sự chuyển đổi định dạng sau đó thành thiết kế bố trí mô phỏng đích. Theo cách khác, tuy nhiên, bộ giải mã SAOC được thực hiện để hỗ trợ chế độ “năng lượng thấp”, trong đó bộ giải mã SAOC được cấu hình để giải mã thành thiết kế bố trí mô phỏng trực tiếp mà không có sự chuyển đổi định dạng sau đó. Theo phương án này, bộ giải mã SAOC 1800 xuất ra trực tiếp tín hiệu loa phóng thanh như tín hiệu loa phóng thanh 5.1 và bộ giải mã SAOC 1800 yêu cầu thông tin bố trí mô phỏng và ma trận kết xuất sao cho việc đại diện độ cơ sở vectơ hoặc loại khác bất kỳ của bộ xử lý để tạo ra thông tin trộn giảm có thể vận hành.

Fig.9 minh họa phương án khác của bộ kết xuất lập thể 1710 của Fig.6. Cụ thể là, với các thiết bị di động, việc kết xuất lập thể được yêu cầu với tai nghe gắn vào các

thiết bị di động hoặc đối với loa phóng thanh được gắn vào các thiết bị di động nhỏ thông thường. Đối với các thiết bị di động này, các điều kiện bắt buộc có thể có để giới hạn bộ giải mã và kết xuất độ phức tạp. Ngoài việc bỏ qua băng tương quan trong các cảnh xử lý này, tốt hơn là lần trộn giảm thứ nhất sử dụng bộ trộn giảm 1712 đến trộn giảm trung gian, tức là, đến số lượng kênh đầu ra thấp hơn mà sau đó tạo số kênh đầu vào thấp hơn đối với bộ chuyển đổi lập thể 1714. Ví dụ, vật liệu kênh 22.2 được trộn giảm bằng bộ trộn giảm 1712 thành trộn giảm trung gian 5.1 hoặc, theo cách khác, việc trộn giảm trung gian được tính trực tiếp bằng bộ giải mã SAOC 1800 của Fig.6 ở loại chế độ “phím tắt”. Sau đó, việc kết xuất lập thể chỉ phải áp dụng mười hàm chuyển liên quan đến vùng tiêu đề (Head Related Transfer Functions – HRTF) hoặc các hàm BRIR để kết xuất năm kênh riêng biệt ở các vị trí khác nhau trái với áp dụng 44 HRTF cho các hàm BRIR nếu các kênh đầu vào 22.2 sẽ được kết xuất trực tiếp. Cụ thể là, các toán tử phức tạp cần thiết để kết xuất lập thể yêu cầu nhiều năng lượng xử lý và do đó làm giảm năng lượng xử lý này trong khi vẫn thu được chất lượng âm thanh chấp nhận được mà hữu dụng đặc biệt đối với các thiết bị di động.

Tốt hơn là, “phím tắt” như được minh họa bởi đường điều khiển 1727 bao gồm việc điều khiển bộ giải mã 1300 để giải mã số kênh thấp hơn, tức là, bỏ qua khối xử lý OTT hoàn thiện ở bộ giải mã hoặc định dạng chuyển đổi thành số kênh thấp hơn và, như được minh họa ở Fig.9, việc kết xuất lập thể được thực hiện đối với số lượng kênh thấp hơn. Cùng việc xử lý có thể được áp dụng không chỉ cho xử lý lập thể mà còn cho chuyển đổi định dạng như được minh họa bởi đường 1727 ở Fig.6.

Theo phương án khác, sự giao tiếp hiệu quả giữa các khối xử lý được yêu cầu. Cụ thể là ở Fig.6, đường dẫn tín hiệu âm thanh giữa các khối xử lý khác nhau được mô tả. Bộ kết xuất lập thể 1710, bộ chuyển đổi định dạng 1720, bộ giải mã SAOC 1800 và bộ giải mã USAC 1300, trong trường hợp SBR (sự tái tạo băng phổ) được áp dụng, tất cả hoạt động ở miền QMF hoặc QMF lai. Theo một phương án, tất cả các khối xử lý này cung cấp giao diện QMF hoặc QMF lai để cho phép các tín hiệu âm thanh đi qua giữa nhau ở miền QMF theo cách có hiệu quả. Ngoài ra, tốt hơn là thực hiện mô đun bộ trộn và mô đun bộ kết xuất đối tượng để cũng làm việc ở miền QMF hoặc QMF lai. Do đó, phân tích QMF hoặc QMF riêng biệt và các giải đoạn tổng hợp có thể tránh

được mà tạo ra các tiết kiệm về độ phức tạp đáng kể và sau đó chỉ giai đoạn tổng hợp QMF cuối cùng được yêu cầu để tạo ra các loa phóng thanh được biểu thị ở 1730 hoặc để tạo ra dữ liệu lập thể tại đầu ra của khối 1710 hoặc để tạo ra các tín hiệu loa bố trí mô phỏng tại đầu ra của khối 1720.

Sau đó, tham chiếu đến Fig.11 để giải thích các thành phần có bốn kênh (QCE). Trái với thành phần cặp kênh như được xác định theo tiêu chuẩn USAC-MPEG, thành phần có bốn kênh yêu cầu bốn kênh đầu vào 90 và xuất ra thành phần QCE được mã hóa 91. Theo một phương án, cây phân loại của hai hộp quanh MPEG ở chế độ 2-1-2 hoặc hai hộp hai thành một (Two To One - TTO) và kết nối bổ sung các công cụ mã hóa lập thể (ví dụ, lập thể MS) như được xác định ở USAC MPEG hoặc bao quanh MPEG được cung cấp và thành phần QCE không chỉ bao gồm hai kênh trộn giảm được mã hóa lập thể và tùy ý hai kênh dư được mã hóa lập thể kết hợp và, ngoài ra, dữ liệu tham số thu được từ, ví dụ, hai hộp TTO. Ở phía bộ giải mã, cấu trúc được áp dụng mà việc giải mã lập thể kết hợp của hai kênh trộn giảm và tùy ý hai kênh dư được áp dụng và ở giai đoạn thứ hai có hai hộp OTT mà sự trộn giảm và các kênh dư tùy ý được trộn tăng đến bốn kênh đầu ra. Tuy nhiên, các toán tử xử lý khác với một bộ mã hóa QCE có thể được áp dụng thay cho toán tử theo thứ bậc. Do đó, ngoài việc mã hóa kênh kết hợp của hai kênh, bộ mã hóa lõi/bộ giải mã còn sử dụng việc mã hóa kênh kết hợp của nhóm bốn kênh.

Hơn nữa, tốt hơn là thực hiện quy trình làm đầy tiếng ồn để cho phép mã hóa băng đầy đủ không bị hao mòn (18kHz) ở 1200kbps.

Bộ mã hóa được hoạt động theo phương thức ‘tốc độ không đổi có chứa bit’, sử dụng tối đa 6144 bit/kênh do đệm tốc độ đối với dữ liệu động.

Tất cả các trọng tải bổ sung như dữ liệu SAOC hoặc siêu dữ liệu đối tượng được đi qua các thành phần mở rộng và được xem xét trong điều khiển tốc độ của bộ mã hóa.

Để tận dụng các chức năng của SAOC cho nội dung âm thanh ba chiều, các phần mở rộng sau đây cho SAOC MPEG được thực hiện:

- Trộn giảm đến số lượng tùy ý của các kênh vận chuyển SAOC.
- Kết xuất tăng cường đến các cấu hình đầu ra có số loa phóng thanh cao (lên đến 22.2).

Mô đun bộ kết xuất lập thể tạo ra sự trộn giảm lập thể của vật liệu âm thanh đa kênh, sao cho mỗi kênh đầu vào (loại trừ các kênh LFE) được biểu diễn bởi nguồn âm thanh ảo. Việc xử lý này được tiến hành theo khung ở miền QMF.

Việc lập thể hóa dựa trên các đáp ứng xung phòng lập thể đo được. Âm thanh trực tiếp và các phản xạ sớm bị ảnh hưởng đến vật liệu âm thanh thông qua sự tiếp cận phức tạp ở miền giả FFT bằng cách sử dụng sự phức tạp nhanh ở trên miền QMF.

Mặc dù một số khía cạnh được mô tả ở ngữ cảnh của thiết bị, rõ ràng là các khía cạnh này cũng biểu diễn sự mô tả phương pháp tương ứng, trong đó khối hoặc thiết bị tương ứng với bước phương pháp hoặc dấu hiệu của bước phương pháp. Tương tự, các khía cạnh được mô tả trong ngữ cảnh của bước phương pháp cũng biểu diễn sự mô tả khối tương ứng hoặc mục hoặc dấu hiệu của thiết bị tương ứng. Một số hoặc tất cả các bước phương pháp có thể được thực hiện bằng (hoặc sử dụng) thiết bị phần cứng, như ví dụ, bộ vi xử lý, chương trình lập trình được hoặc mạch điện tử. Theo một số phương án, một hoặc một vài bước phương pháp quan trọng nhất được thực hiện bằng thiết bị này.

Tùy thuộc vào các yêu cầu thực hiện nhất định, các phương án theo sáng chế có thể được thực hiện trong phần cứng hoặc phần mềm. Việc thực hiện này có thể được tiến hành bằng cách sử dụng vật ghi lưu trữ cố định như vật ghi lưu trữ số, ví dụ, đĩa mềm, DVD, Blu-Ray, CD, ROM, PROM, và EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều khiển đọc được bằng điện tử được lưu trữ trên đó, mà kết hợp (hoặc có khả năng kết hợp) với hệ thống máy tính lập trình được sao cho phương pháp tương ứng được thực hiện. Do đó, vật ghi lưu trữ số có thể là bằng máy tính có thể đọc được.

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

Thông thường, các phương án theo sáng chế có thể được thực hiện như sản phẩm của chương trình máy tính có mã chương trình, mã chương trình hoạt động để thực hiện một trong số các phương pháp khi sản phẩm của chương trình máy tính chạy trên máy tính. Mã chương trình có thể, ví dụ, được lưu trữ trên vật mang đọc được bằng máy.

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

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

Do đó, phương án khác theo phương pháp của sáng chế là vật mang dữ liệu (hoặc vật ghi lưu trữ số, hoặc vật ghi đọc được bằng máy tính) bao gồm, được ghi trên đó, chương trình máy tính để thực hiện một trong số các phương pháp được mô tả ở đây. Vật mang dữ liệu, vật ghi lưu trữ số hoặc vật ghi được ghi lại thường là hữu hình và/hoặc không chuyển tiếp.

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

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

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

Phương án khác theo sáng chế bao gồm thiết bị hoặc hệ thống được cấu hình để truyền (ví dụ, bằng điện tử hoặc quang học) chương trình máy tính để thực hiện một trong số các phương pháp được mô tả ở đây đến bộ nhận. Bộ nhận này có thể, ví dụ, là máy tính, thiết bị di động, thiết bị nhớ hoặc thiết bị tương tự. Thiết bị hoặc hệ thống này có thể, ví dụ, bao gồm máy chủ tệp tin để truyền chương trình máy tính đến bộ nhận.

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

Các phương án nêu trên chỉ đơn thuần minh họa cho các nguyên tắc của sáng chế. Cần hiểu rằng các biến thể và biến đổi về cách bố trí và mô tả chi tiết ở đây sẽ rõ ràng đối với người có hiểu biết trung bình trong lĩnh vực này. Do đó, có ý định chỉ bị giới hạn bởi phạm vi của các yêu cầu bảo hộ đi kèm và không bị giới hạn bởi phần mô tả chi tiết được thể hiện bằng cách mô tả và giải thích của các phương án ở đây.

## YÊU CẦU BẢO HỘ

1. Bộ mã hóa âm thanh để mã hóa dữ liệu đầu vào âm thanh (101) để thu được dữ liệu đầu ra âm thanh (501) bao gồm:

giao diện đầu vào (100) được cấu hình để nhận nhiều kênh âm thanh, nhiều đối tượng âm thanh và siêu dữ liệu liên quan đến một hoặc nhiều đối tượng âm thanh trong số nhiều đối tượng âm thanh;

bộ trộn (200) được cấu hình để trộn nhiều đối tượng và nhiều kênh để thu được nhiều kênh được trộn trước, mỗi kênh được trộn trước bao gồm dữ liệu âm thanh của kênh và dữ liệu âm thanh của ít nhất một đối tượng;

bộ mã hóa lỗi (300) được cấu hình để mã hóa lỗi dữ liệu đầu vào bộ mã hóa lỗi; và

bộ nén siêu dữ liệu (400) được cấu hình để nén siêu dữ liệu liên quan đến một hoặc nhiều đối tượng âm thanh trong số nhiều đối tượng âm thanh,

trong đó, bộ mã hóa âm thanh được cấu hình để vận hành ở cả hai chế độ của nhóm của ít nhất hai chế độ bao gồm chế độ thứ nhất, trong đó bộ mã hóa lỗi được cấu hình để mã hóa nhiều kênh âm thanh và nhiều đối tượng âm thanh được nhận bằng giao diện đầu vào như dữ liệu đầu vào bộ mã hóa lỗi, và chế độ thứ hai, trong đó bộ mã hóa lỗi (300) được cấu hình để nhận, như dữ liệu đầu vào bộ mã hóa lỗi, nhiều kênh được trộn trước được tạo ra bởi bộ trộn (200) và để mã hóa nhiều kênh được trộn trước.

2. Bộ mã hóa âm thanh theo điểm 1, còn bao gồm:

bộ mã hóa đối tượng âm thanh trong không gian (800) để tạo ra một hoặc nhiều kênh vận chuyển và dữ liệu tham số từ dữ liệu đầu vào bộ mã hóa đối tượng âm thanh trong không gian,

trong đó, bộ mã hóa âm thanh được cấu hình để hoạt động thêm trong chế độ thứ ba, trong đó bộ mã hóa lỗi (300) mã hóa một hoặc nhiều kênh vận chuyển được

suy ra từ dữ liệu đầu vào bộ mã hóa đối tượng âm thanh trong không gian, dữ liệu đầu vào bộ mã hóa đối tượng âm thanh trong không gian bao gồm nhiều đối tượng âm thanh hoặc hai hoặc nhiều kênh âm thanh trong số nhiều kênh âm thanh.

3. Bộ mã hóa âm thanh theo điểm 1 hoặc 2, còn bao gồm:

bộ mã hóa đối tượng âm thanh trong không gian (800) để tạo ra một hoặc nhiều kênh vận chuyển và dữ liệu tham số từ dữ liệu đầu vào bộ mã hóa đối tượng âm thanh trong không gian,

trong đó, bộ mã hóa âm thanh được cấu hình để hoạt động thêm trong ngay cả chế độ khác, trong đó bộ mã hóa lõi mã hóa các kênh vận chuyển được suy ra bằng bộ mã hóa đối tượng âm thanh trong không gian (800) từ các kênh được trộn trước làm dữ liệu đầu vào bộ mã hóa đối tượng âm thanh trong không gian.

4. Bộ mã hóa âm thanh theo một trong số các điểm nêu trên, còn bao gồm bộ kết nối để kết nối đầu ra của giao diện đầu vào (100) với đầu vào bộ mã hóa lõi (300) trong chế độ thứ nhất và kết nối đầu ra của giao diện đầu vào (100) với đầu vào của bộ trộn (200) và kết nối đầu ra của bộ trộn (200) tới đầu vào của bộ mã hóa lõi (300) trong chế độ thứ hai, và

bộ điều khiển chế độ (600) để điều khiển bộ kết nối theo chỉ dẫn chế độ được nhận từ giao diện người dùng hoặc được tách từ dữ liệu đầu vào âm thanh (101).

5. Bộ mã hóa âm thanh theo điểm bất kỳ trong số các điểm nêu trên, còn bao gồm:

giao diện đầu ra (500) để cung cấp tín hiệu đầu ra như dữ liệu đầu ra âm thanh (501), tín hiệu đầu ra bao gồm, trong chế độ thứ nhất, đầu ra của bộ mã hóa lõi (300) và siêu dữ liệu được nén, và bao gồm, trong chế độ thứ hai, đầu ra bộ mã hóa lõi (300) mà không có siêu dữ liệu bất kỳ, và bao gồm, trong chế độ thứ ba, đầu ra của bộ mã hóa lõi (300), thông tin phụ SAOC và siêu dữ liệu được nén và bao gồm, ngay cả trong chế độ khác, đầu ra của bộ mã hóa lõi (300) và thông tin phụ SAOC.

6. Bộ mã hóa âm thanh theo một điểm bất kỳ trong số các điểm nêu trên,

trong đó, bộ trộn (200) được cấu hình để kết xuất trước nhiều đối tượng âm thanh bằng cách sử dụng siêu dữ liệu và chỉ dẫn về vị trí của mỗi kênh trong thiết lập tiếp âm, mà nhiều kênh được kết hợp ở đó,

trong đó, bộ trộn (200) được cấu hình để trộn đối tượng âm thanh với ít nhất hai kênh âm thanh và với việc này thì tổng số kênh âm thanh, khi đối tượng âm thanh được đặt giữa ít nhất hai kênh âm thanh ở thiết lập tiếp âm, như được xác định bằng siêu dữ liệu.

7. Bộ mã hóa âm thanh theo một trong số các điểm nêu trên,

còn bao gồm bộ giải nén siêu dữ liệu (420) để giải nén siêu dữ liệu được nén được xuất ra bởi bộ nén siêu dữ liệu (400), và

trong đó, bộ trộn (200) được cấu hình để trộn nhiều đối tượng theo siêu dữ liệu được giải nén, trong đó hoạt động nén được thực hiện bằng bộ nén siêu dữ liệu (400) là hoạt động mất nén bao gồm bước lượng tử hóa.

8. Bộ giải mã âm thanh để giải mã dữ liệu âm thanh được mã hóa, bao gồm:

giao diện đầu vào (1100) được cấu hình để nhận dữ liệu âm thanh được mã hóa, dữ liệu âm thanh được mã hóa bao gồm nhiều kênh được mã hóa hoặc nhiều đối tượng được mã hóa hoặc nén siêu dữ liệu liên quan đến nhiều đối tượng;

bộ giải mã lõi (1300) được cấu hình để giải mã nhiều kênh được mã hóa và các đối tượng được mã hóa;

bộ giải nén siêu dữ liệu (1400) được cấu hình để giải nén siêu dữ liệu được nén,

bộ xử lý đối tượng (1200) được cấu hình để xử lý nhiều đối tượng được giải mã bằng cách sử dụng siêu dữ liệu được giải nén để thu được số kênh đầu ra (1205) bao gồm dữ liệu âm thanh từ các đối tượng và các kênh được giải mã; và

bộ xử lý sau (1700) được cấu hình để chuyển đổi số kênh đầu ra (1205) thành định dạng đầu ra,

trong đó, bộ giải mã âm thanh được cấu hình để bỏ qua bộ xử lý đối tượng và để nạp nhiều kênh được giải mã vào bộ xử lý sau (1700), khi dữ liệu âm thanh được mã hóa không chứa các đối tượng âm thanh bất kỳ và nạp nhiều đối tượng được giải mã và nhiều kênh được giải mã vào bộ xử lý đối tượng (1200), khi dữ liệu âm thanh được mã hóa bao gồm các kênh được mã hóa và các đối tượng được mã hóa.

9. Bộ giải mã âm thanh theo điểm 8, trong đó bộ xử lý sau (1700) được cấu hình để chuyển đổi số kênh đầu ra (1205) để biểu diễn lập thể hoặc định dạng mô phỏng có số kênh nhỏ hơn so với số kênh đầu ra,

trong đó, bộ giải mã âm thanh được cấu hình để điều khiển bộ xử lý sau (1700) theo đầu vào điều khiển thu được từ giao diện người dùng hoặc được tách ra từ tín hiệu âm thanh được mã hóa.

10. Bộ giải mã âm thanh theo điểm 8 hoặc 9, trong đó bộ xử lý đối tượng bao gồm:

bộ kết xuất đối tượng để kết xuất các đối tượng được giải mã bằng cách sử dụng siêu dữ liệu được giải nén; và

bộ trộn (1220) để trộn các đối tượng được kết xuất và các kênh được giải mã để thu được số kênh đầu ra (1205).

11. Bộ giải mã âm thanh theo một trong số các điểm từ 8 đến 10, trong đó bộ xử lý đối tượng (1200) bao gồm:

bộ giải mã mã hoá đối tượng âm thanh trong không gian để giải mã một hoặc nhiều kênh vận chuyển và thông tin phụ tham số có liên quan biểu diễn các đối tượng âm thanh được mã hóa, trong đó bộ giải mã mã đối tượng âm thanh trong không gian được cấu hình để kết xuất các đối tượng âm thanh được giải mã theo thông tin kết xuất liên quan đến sự bố trí của các đối tượng âm thanh và điều khiển bộ xử lý đối tượng để trộn các đối tượng âm thanh được kết xuất và các kênh âm thanh được giải mã để thu được số kênh đầu ra (1205).

12. Bộ giải mã âm thanh theo một trong số các điểm từ 8 đến 10, trong đó bộ xử lý đối tượng (1200) bao gồm bộ giải mã mã đối tượng âm thanh trong không gian (1800)

để giải mã một hoặc nhiều kênh vận chuyển và thông tin phụ tham số có liên quan biểu diễn các đối tượng âm thanh được mã hóa và các kênh âm thanh được mã hóa,

trong đó, bộ mã hóa đối tượng âm thanh trong không gian được cấu hình để giải mã các đối tượng âm thanh được mã hóa và các kênh âm thanh được mã hóa bằng cách sử dụng một hoặc nhiều kênh vận chuyển và thông tin phụ tham số và trong đó bộ xử lý đối tượng được cấu hình để kết xuất các đối tượng âm thanh bằng cách sử dụng siêu dữ liệu được giải nén và giải nén các kênh và trộn chúng với các đối tượng được kết xuất để thu được số kênh đầu ra (1205).

13. Bộ giải mã âm thanh theo một trong số các điểm từ 8 đến 10, trong đó bộ xử lý đối tượng (1200) bao gồm bộ giải mã mã hóa đối tượng âm thanh trong không gian (1800) để giải mã một hoặc nhiều kênh vận chuyển và thông tin phụ tham số được kết hợp biểu diễn các đối tượng âm thanh được mã hóa hoặc các kênh âm thanh được mã hóa,

trong đó, bộ giải mã mã đối tượng âm thanh trong không gian được cấu hình để chuyển mã thông tin tham số được kết hợp và siêu dữ liệu được giải nén thành thông tin phụ tham số được chuyển mã có thể sử dụng được để kết xuất trực tiếp định dạng đầu ra, và trong đó bộ xử lý sau (1700) được cấu hình để tính toán các kênh âm thanh của định dạng đầu ra bằng cách sử dụng các kênh vận chuyển được giải mã và thông tin phụ tham số được chuyển mã, hoặc

trong đó, bộ giải mã mã hóa đối tượng âm thanh trong không gian được cấu hình để trộn tăng trực tiếp và kết xuất các tín hiệu kênh cho định dạng đầu ra bằng cách sử dụng các kênh vận chuyển được giải mã và thông tin phụ tham số.

14. Bộ giải mã âm thanh theo một trong số các điểm từ 8 đến 13,

trong đó, bộ xử lý đối tượng (1200) bao gồm bộ giải mã mã đối tượng âm thanh trong không gian để giải mã một hoặc nhiều kênh vận chuyển được xuất ra bằng bộ giải mã lõi (1300) và dữ liệu tham số được kết hợp và siêu dữ liệu được giải nén để thu được nhiều đối tượng âm thanh được kết xuất,

trong đó, bộ xử lý đối tượng (1200) còn được cấu hình để kết xuất các đối tượng được giải mã được xuất ra bởi bộ giải mã lõi (1300);

trong đó, bộ xử lý đối tượng (1200) còn được cấu hình để trộn các đối tượng được giải mã được kết xuất với các kênh được giải mã,

trong đó, bộ giải mã âm thanh còn bao gồm giao diện đầu ra (1730) để xuất đầu ra của bộ trộn (1220) đến các loa phóng thanh,

trong đó, bộ xử lý sau còn bao gồm:

bộ kết xuất lập thể để kết xuất các kênh đầu ra vào hai kênh lập thể bằng cách sử dụng các hàm truyền liên quan đến vùng khởi đầu hoặc các đáp ứng xung lập thể, và

bộ chuyển đổi định dạng (1720) để chuyển đổi các kênh đầu ra thành định dạng đầu ra có số kênh thấp hơn so với các kênh đầu ra của bộ trộn (1220) bằng cách sử dụng thông tin trên thiết kế bố trí mô phỏng.

15. Bộ giải mã âm thanh theo điểm bất kỳ trong số các điểm từ 8 đến 14,

trong đó, nhiều thành phần kênh được mã hóa hoặc nhiều đối tượng âm thanh được mã hóa được mã hóa làm các thành phần cặp kênh, các thành phần đơn kênh, các thành phần tần số thấp hoặc các thành phần có bốn kênh, trong đó thành phần có bốn kênh bao gồm bốn kênh hoặc đối tượng ban đầu, và

trong đó, bộ giải mã lõi (1300) được cấu hình để giải mã các thành phần cặp kênh, các thành phần đơn kênh, các thành phần có tần số thấp hoặc các thành phần có bốn kênh theo thông tin phụ được chứa trong dữ liệu âm thanh được mã hóa biểu thị thành phần cặp kênh, thành phần đơn kênh, thành phần có tần số thấp hoặc thành phần có bốn kênh.

16. Bộ giải mã âm thanh theo một trong số các điểm từ 8 đến 15,

trong đó, bộ giải mã lõi (1300) được cấu hình để áp dụng toán tử giải mã bằng đầy đủ bằng cách sử dụng toán tử điền đầy nhiều âm mà không có toán tử tái tạo bằng phổ.

17. Bộ giải mã âm thanh theo điểm 14, trong đó các thành phần bao gồm bộ kết xuất lập thể (1710), bộ chuyển đổi định dạng (1720), bộ trộn (1220), bộ giải mã SAOC (1800) và bộ giải mã lõi (1300) và bộ kết xuất đối tượng (1210) hoạt động ở miền giàn lọc gương cầu phương (QMF) và trong đó dữ liệu miền lọc gương cầu phương được truyền từ một trong số các thành phần đến một trong số các thành phần khác mà không cần phải có việc xử lý giàn lọc tổng hợp hoặc giàn lọc phân tích bất kỳ sau đó.

18. Bộ giải mã âm thanh theo điểm bất kỳ trong số các điểm từ 8 đến 17,

trong đó, bộ xử lý sau (1700) được cấu hình để trộn giảm các kênh được xuất ra bởi bộ xử lý đối tượng (1200) thành định dạng có ba hoặc nhiều kênh và có ít kênh hơn so với số kênh đầu ra (1205) của bộ xử lý đối tượng (1200) để thu được việc trộn giảm trung gian, và kết xuất lập thể (1210) các kênh của việc trộn giảm trung gian thành tín hiệu đầu ra lập thể có hai kênh.

19. Bộ giải mã âm thanh theo một trong số các điểm từ 8 đến 15, trong đó bộ xử lý sau (1700) bao gồm:

bộ trộn giảm được điều khiển (1722) để áp dụng ma trận trộn giảm; và

bộ điều khiển (1724) để xác định ma trận trộn giảm cụ thể bằng cách sử dụng thông tin về cấu hình kênh của đầu ra của bộ xử lý đối tượng (1200) và thông tin về thiết kế bố trí mô phỏng mong đợi.

20. Bộ giải mã âm thanh theo một trong số các điểm từ 8 đến 19,

trong đó, bộ giải mã lõi (1300) hoặc bộ xử lý đối tượng (1200) có thể điều khiển được, và

trong đó, bộ xử lý sau (1700) được cấu hình để điều khiển bộ giải mã lõi (1300) hoặc bộ xử lý đối tượng (1200) theo thông tin về định dạng đầu ra sao cho sự giải tương quan xảy ra kết xuất xử lý các đối tượng hoặc kênh không xảy ra là các kênh riêng biệt ở định dạng đầu ra bị giảm hoặc bị loại bỏ, hoặc sao cho đối với các đối tượng hoặc kênh xảy ra là các kênh riêng biệt ở định dạng đầu ra, hoạt động trộn tăng hoặc giải mã được thực hiện như nếu các đối tượng hoặc kênh sẽ xảy ra là các kênh

riêng biệt ở định dạng đầu ra, ngoại trừ việc xử lý giải tương quan bất kỳ đối với các đối tượng hoặc các kênh không xảy ra như các kênh riêng biệt ở định dạng đầu ra bị làm cho ngừng hoạt động.

21. Bộ giải mã âm thanh theo một trong số các điểm từ 8 đến 20,

trong đó, bộ giải mã lỗi (1300) được cấu hình để thực hiện việc giải mã biến đổi và giải mã tái tạo bằng phổ với thành phần đơn kênh, và thực hiện việc giải mã biến đổi, giải mã âm thanh nổi tham số với các thành phần cặp kênh và các thành phần bốn kênh.

22. Phương pháp mã hóa dữ liệu đầu vào âm thanh (101) để thu được dữ liệu đầu ra âm thanh (501), phương pháp này bao gồm các bước:

nhận (100) nhiều kênh âm thanh, các đối tượng âm thanh và siêu dữ liệu liên quan đến một hoặc nhiều đối tượng âm thanh trong số nhiều đối tượng âm thanh;

trộn (200) nhiều đối tượng và nhiều kênh để thu được nhiều kênh được trộn trước, mỗi kênh được trộn trước bao gồm dữ liệu âm thanh của kênh và dữ liệu âm thanh của ít nhất một đối tượng;

mã hóa lỗi (300) dữ liệu đầu vào mã hóa lỗi; và

nén (400) siêu dữ liệu liên quan đến một hoặc nhiều đối tượng âm thanh,

trong đó, phương pháp mã hóa âm thanh hoạt động ở hai chế độ của nhóm gồm hai hoặc nhiều chế độ bao gồm chế độ thứ nhất, trong đó việc mã hóa lỗi các kênh âm thanh và các đối tượng âm thanh được nhận làm dữ liệu đầu vào mã hóa lỗi, và chế độ thứ hai, trong đó việc mã hóa lỗi (300) nhận, như dữ liệu đầu vào mã hóa lỗi, nhiều kênh được trộn trước được tạo ra bằng việc trộn (200) và mã hóa lỗi nhiều kênh được trộn trước.

23. Phương pháp giải mã dữ liệu âm thanh được mã hóa, phương pháp này bao gồm các bước:

nhận (1100) dữ liệu âm thanh được mã hóa, dữ liệu âm thanh được mã hóa bao gồm nhiều kênh được mã hóa hoặc các đối tượng được mã hóa hoặc siêu dữ liệu được nén liên quan đến các đối tượng;

giải mã lỗi (1300) các kênh được mã hóa và các đối tượng được mã hóa;

giải nén (1400) siêu dữ liệu được nén,

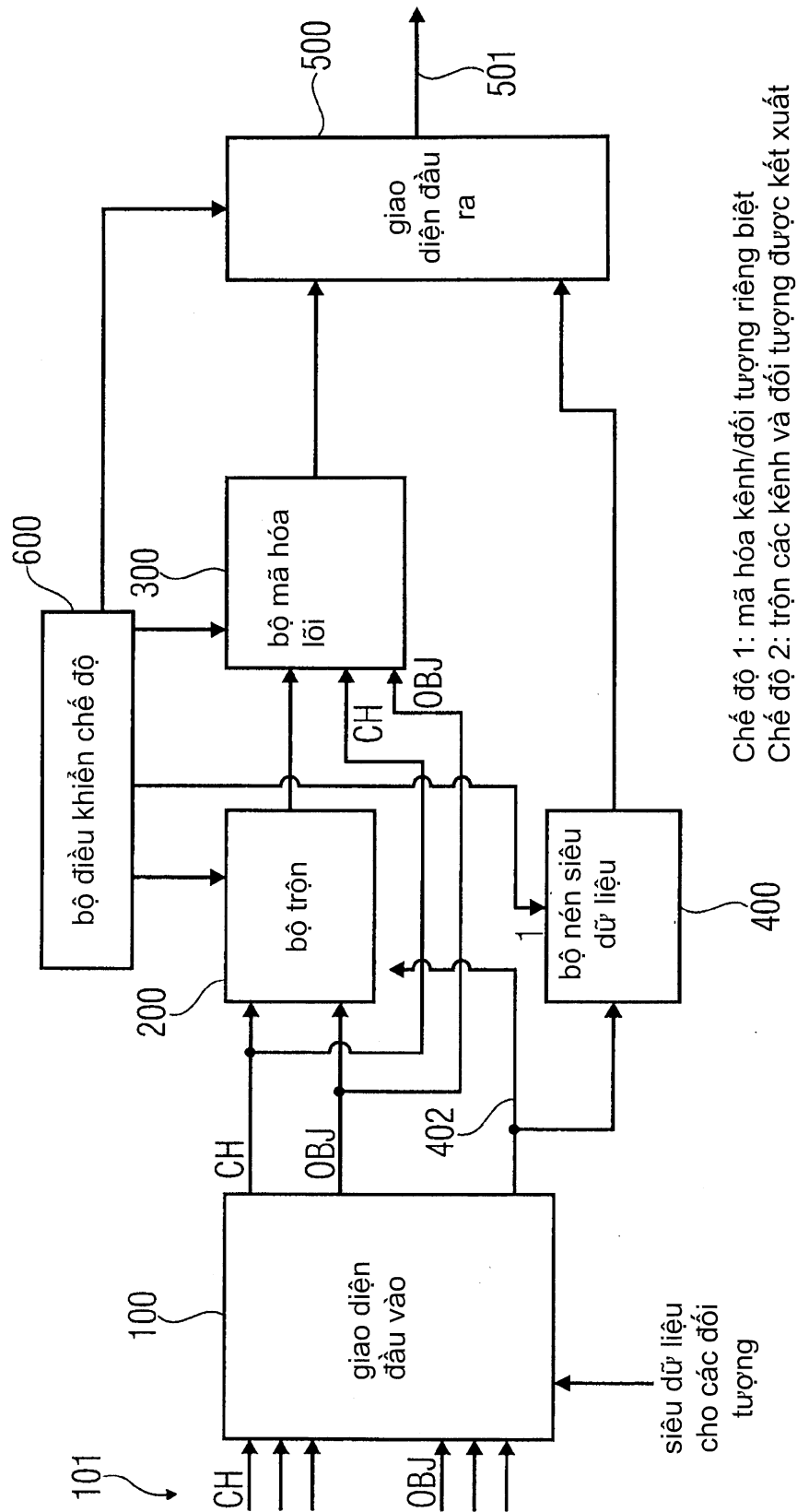
xử lý (1200) nhiều đối tượng được giải mã bằng cách sử dụng siêu dữ liệu được giải nén để thu được số kênh đầu ra (1205) bao gồm dữ liệu âm thanh từ các đối tượng và các kênh được giải mã; và

chuyển đổi (1700) số kênh đầu ra (1205) thành định dạng đầu ra,

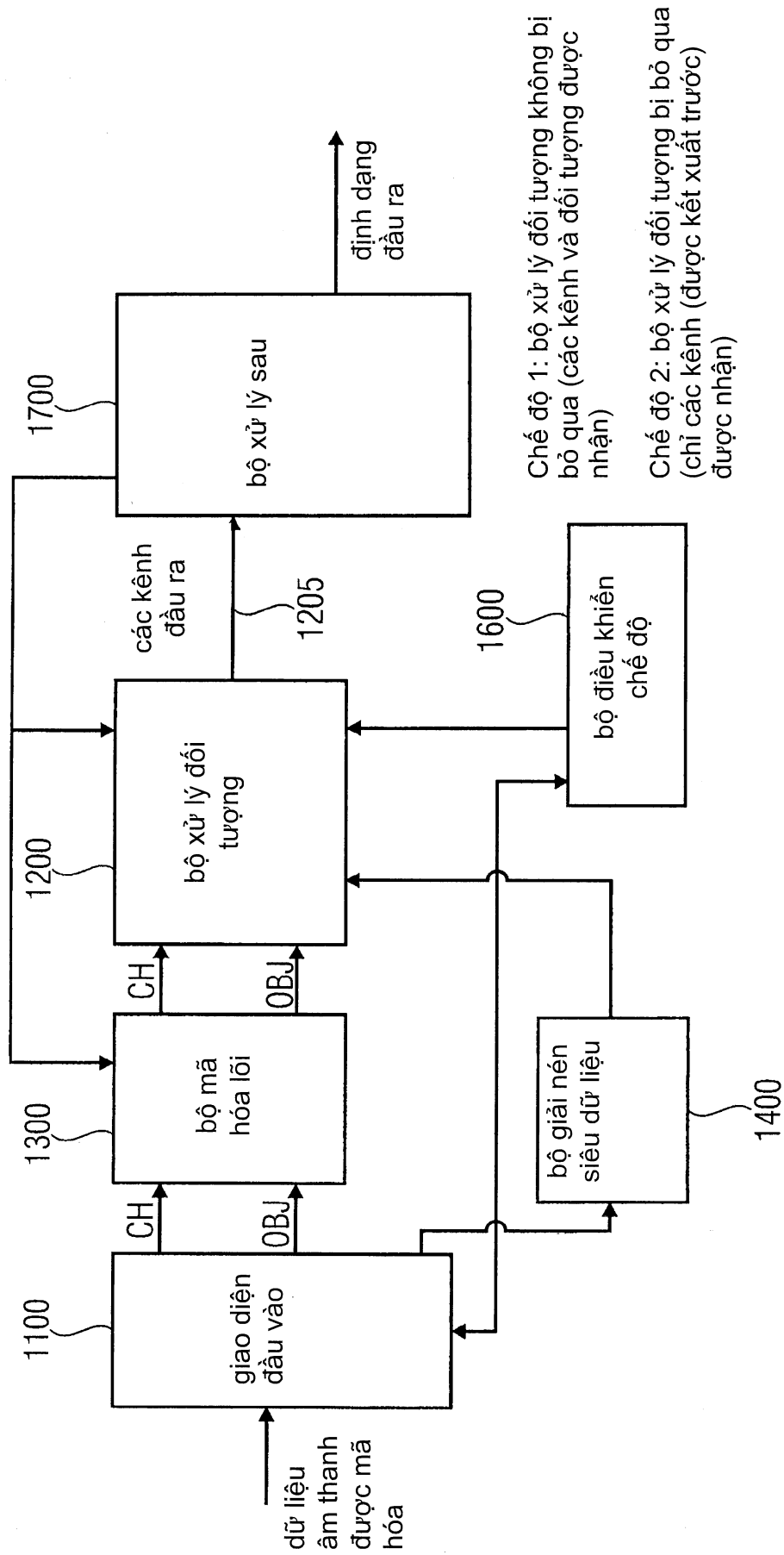
trong đó, trong phương pháp giải mã âm thanh, việc xử lý (1200) nhiều đối tượng được giải mã bị bỏ qua và nhiều kênh được giải mã được nạp vào việc xử lý sau (1700), khi dữ liệu âm thanh được mã hóa không chứa các đối tượng âm thanh bất kỳ và nhiều đối tượng được giải mã và các kênh được giải mã được nạp vào việc xử lý (1200) nhiều đối tượng được giải mã, khi dữ liệu âm thanh được mã hóa bao gồm các kênh được mã hóa và các đối tượng được mã hóa.

24. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính để thực hiện phương pháp theo điểm 22 khi chạy trên máy tính hoặc bộ xử lý.

25. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính để thực hiện phương pháp theo điểm 23 khi chạy trên máy tính hoặc bộ xử lý.



**FIG 1**  
(BỘ MÃ HÓA)



**FIG 2**  
(BỘ GIẢI MÃ)

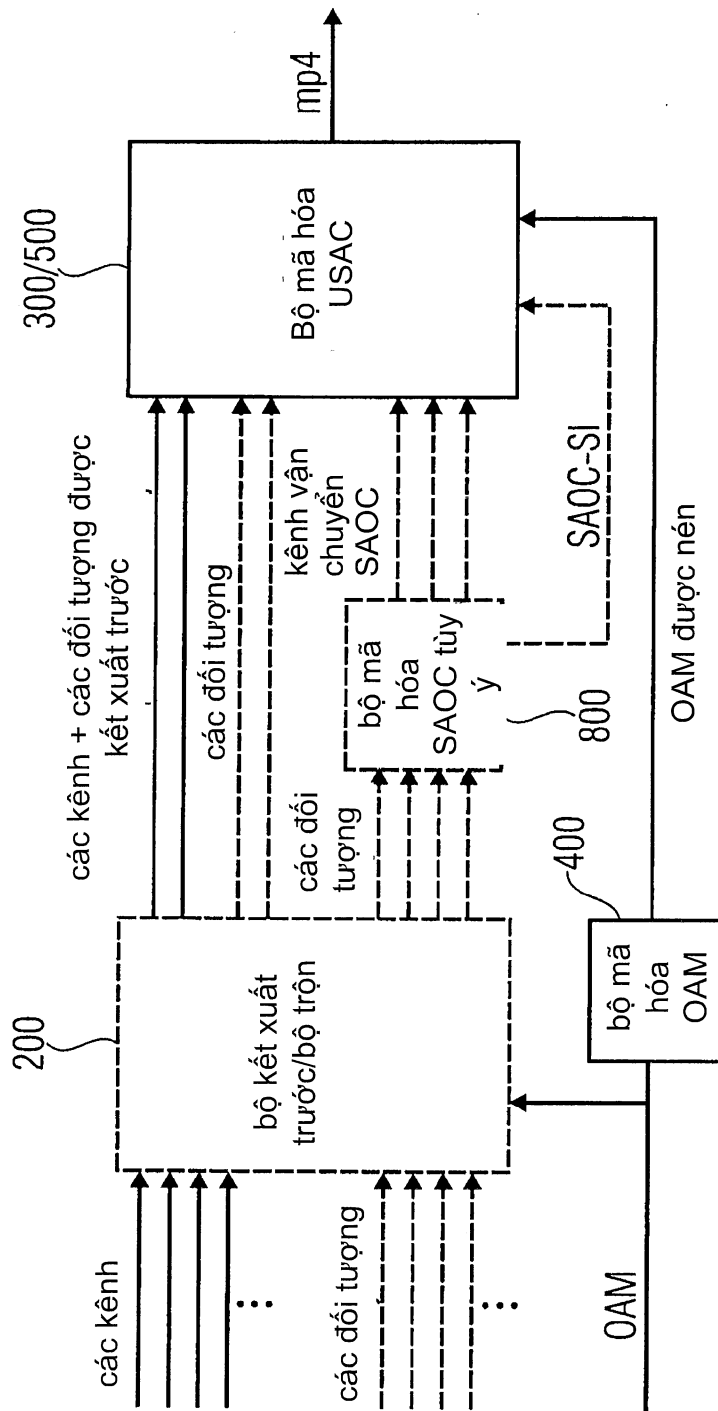
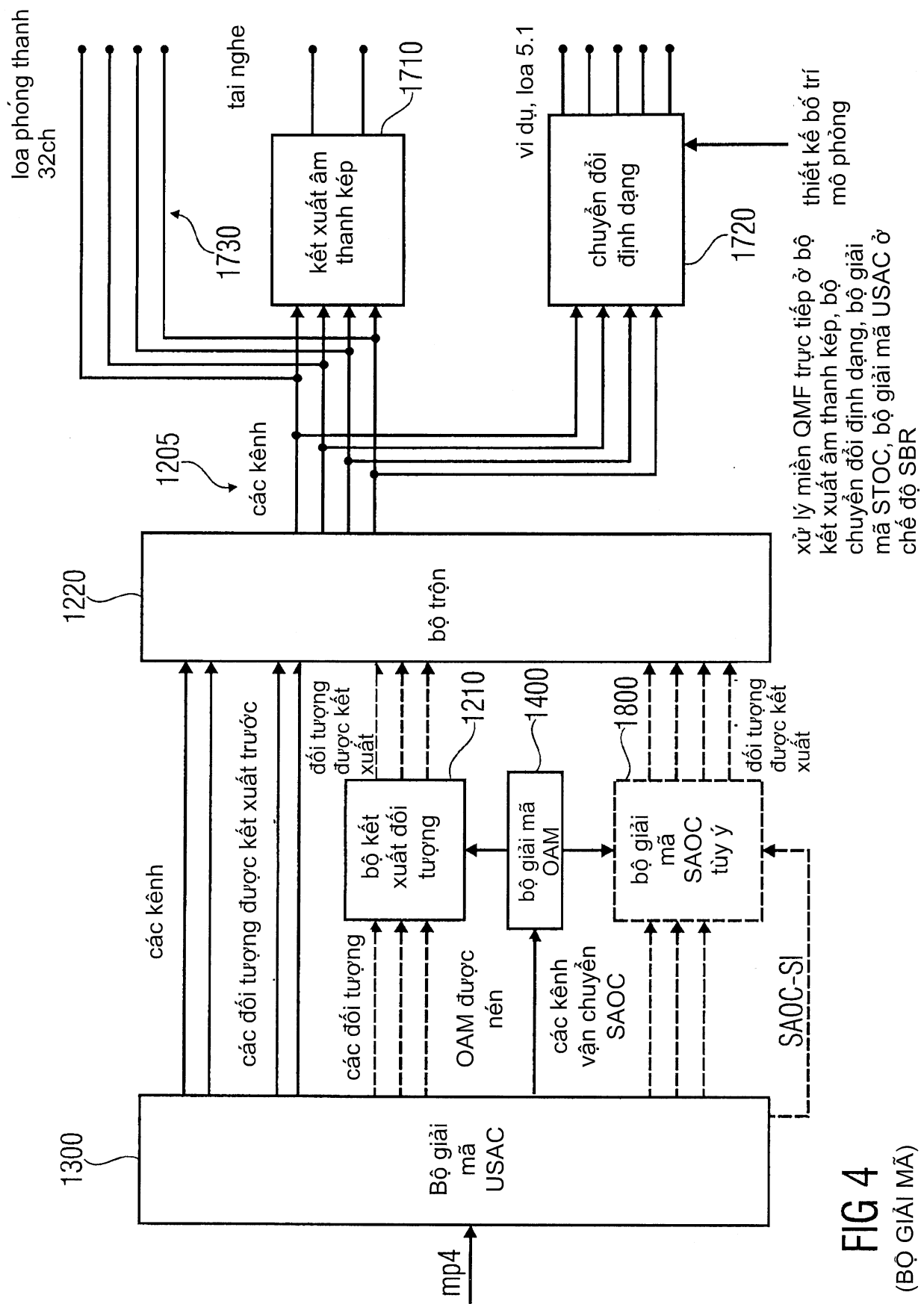
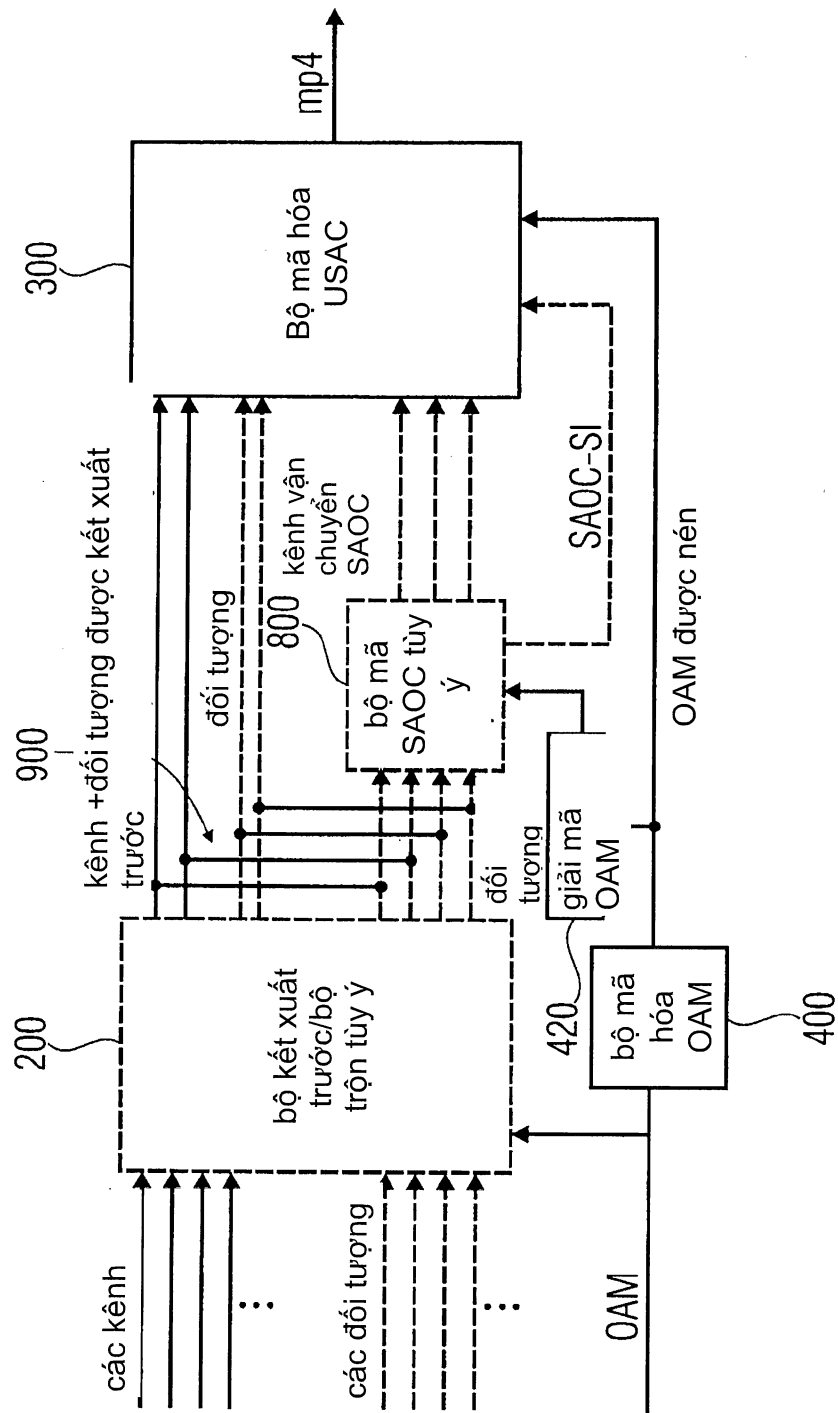


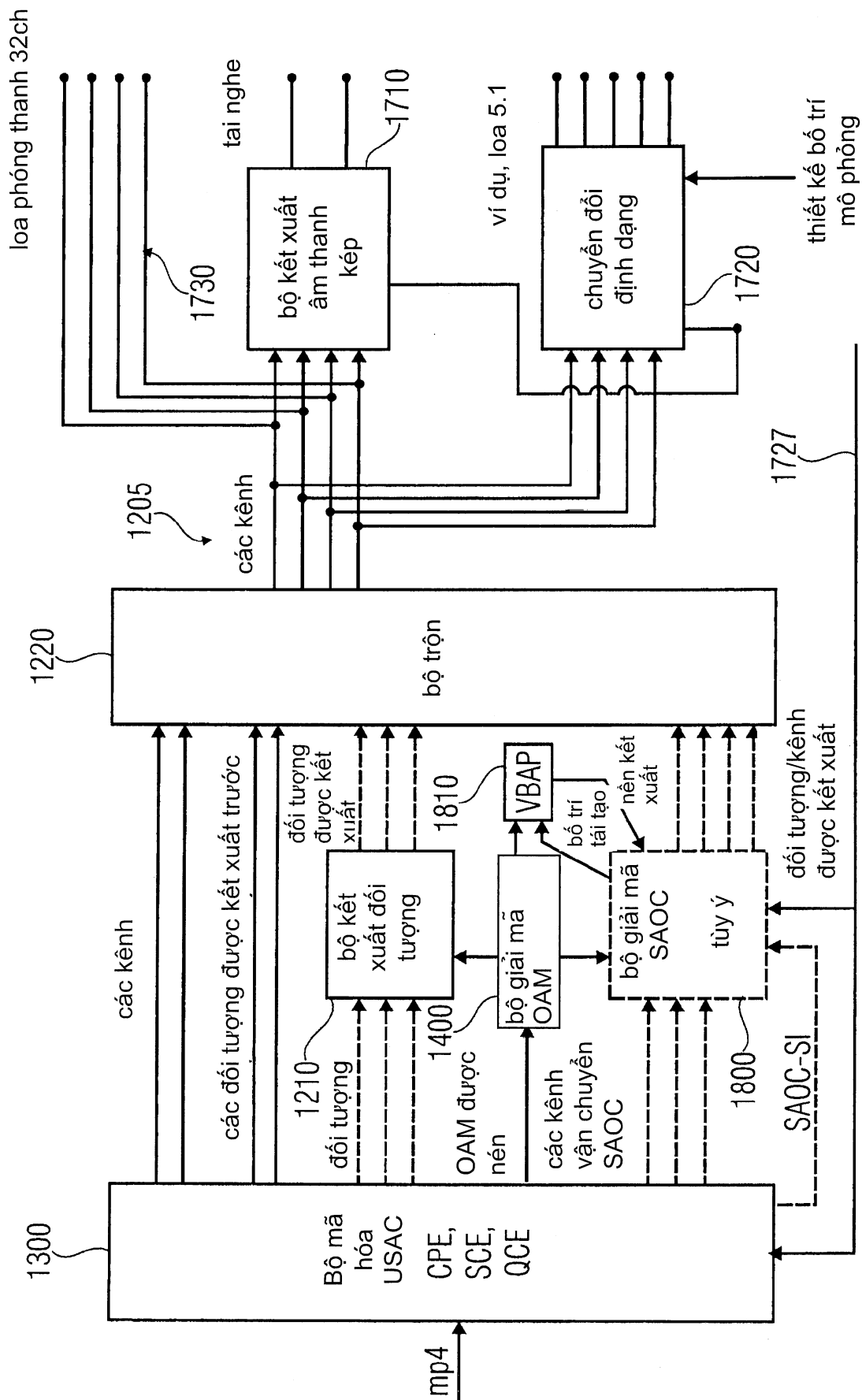
FIG 3

(BỘ MÃ HÓA)





**FIG 5**  
(BỘ MÃ HÓA)



**FIG 6**  
(BỘ GIẢI MÃ)

| chế độ | bộ mã hóa                                                      | bộ giải mã                                                                 |
|--------|----------------------------------------------------------------|----------------------------------------------------------------------------|
| 1      | bộ trộn bị bỏ qua                                              | bộ xử lý đối tượng không bị bỏ qua                                         |
| 2      | bộ trộn hoạt động                                              | bộ xử lý đối tượng bằng cách bỏ qua                                        |
| 3      | mã hóa SAOC chỉ với đối tượng                                  | giải mã SAOC chỉ với các đối tượng                                         |
| 4      | mã hóa SAOC với các kênh được kết xuất trước/bộ trộn hoạt động | giải mã SAOC cho đối tượng được kết xuất trước (xử lý đối tượng bị bỏ qua) |
| 5      | trộn bất kỳ trong các chế độ 1 đến 4                           | trộn bất kỳ trong các chế độ 1 đến 4                                       |

FIG 7

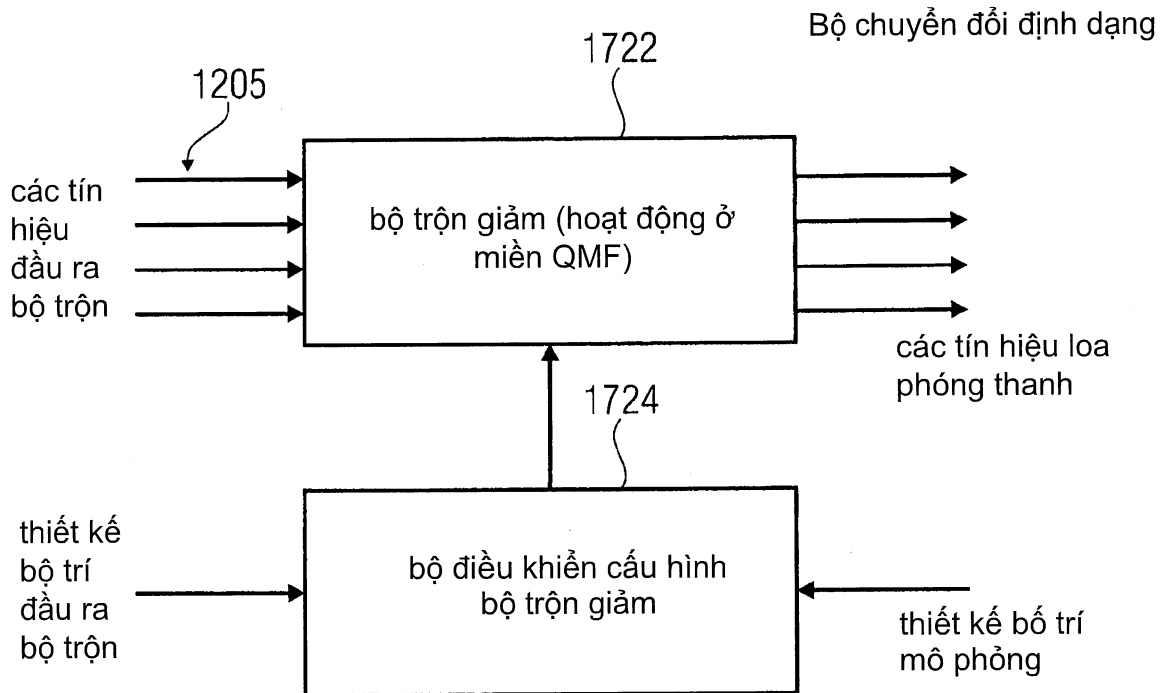


FIG 8

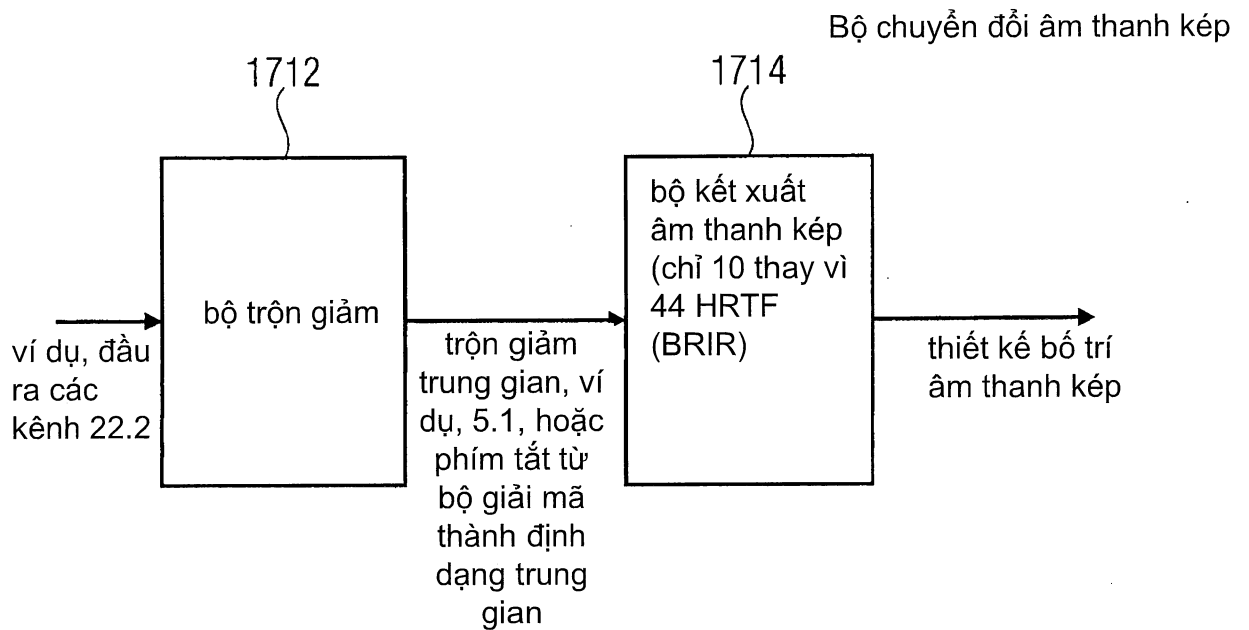


FIG 9

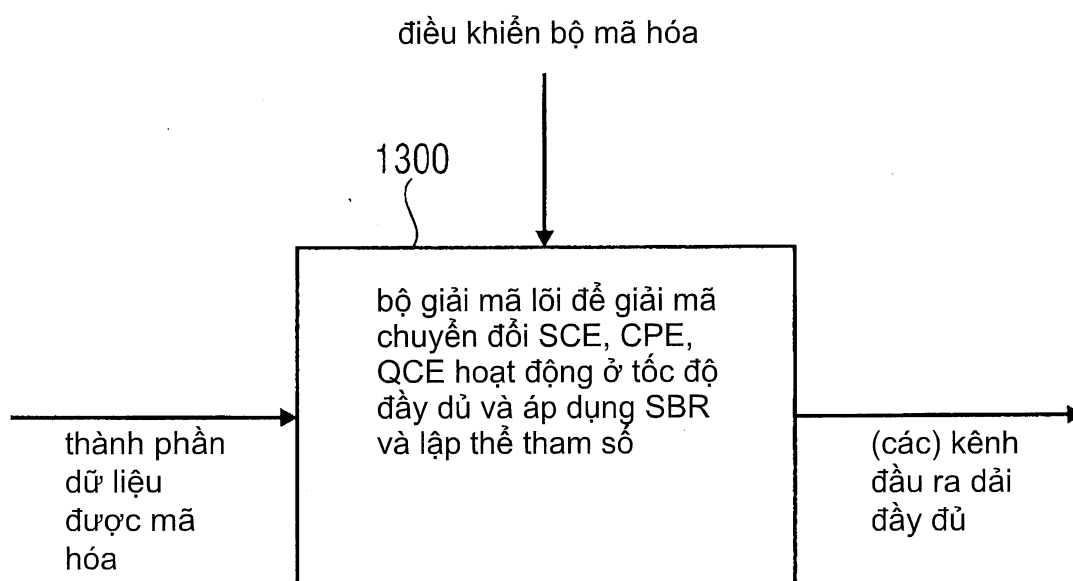


FIG 10

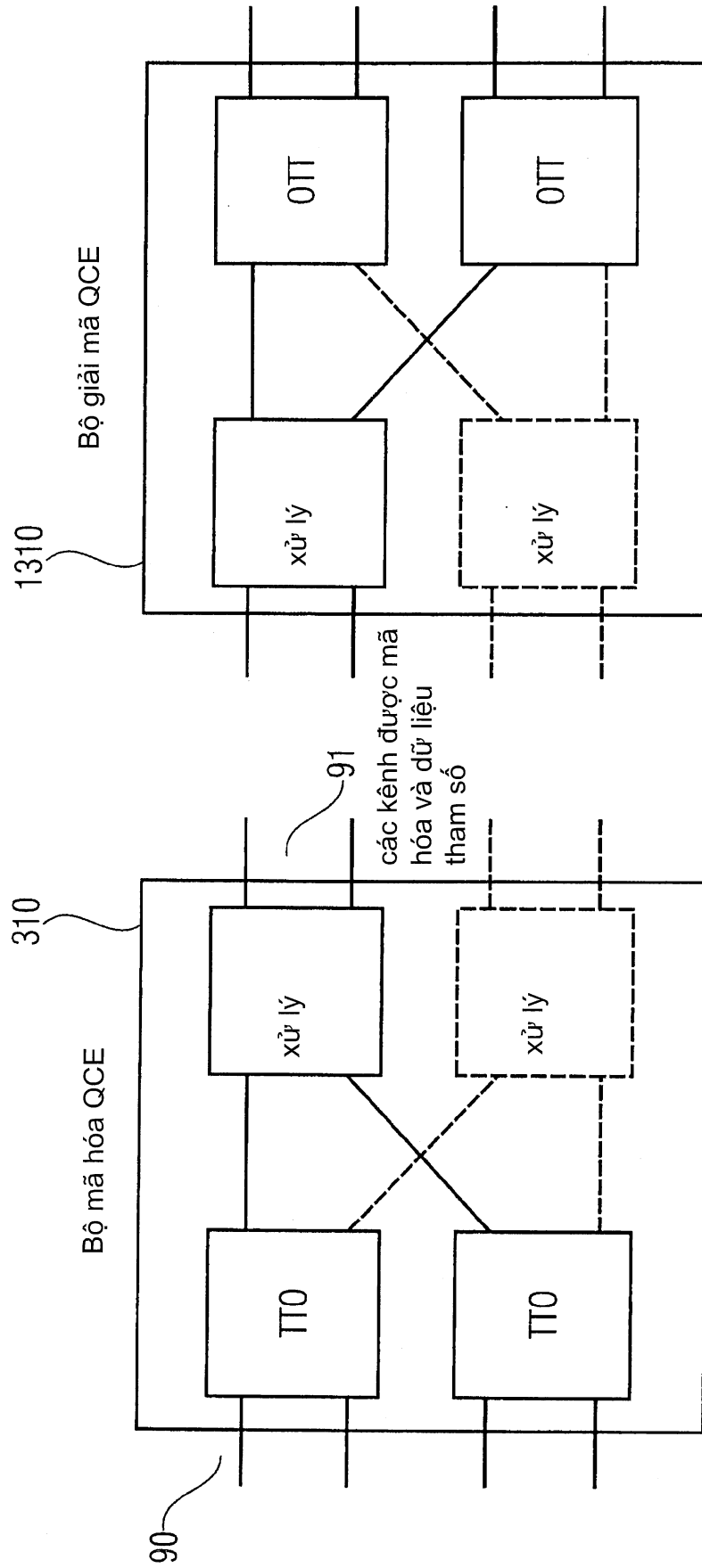


FIG 11