



(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)



1-0025833

(51)⁷ G10L 19/008

(13) B

(21) 1-2015-04443

(22) 23/05/2014

(86) PCT/EP2014/060728 23/05/2014

(87) WO 2014/187986 A1 27/11/2014

(30) 61/827,246 24/05/2013 US

(45) 26/10/2020 391

(43) 25/02/2016 335A

(73) DOLBY INTERNATIONAL AB (SE)

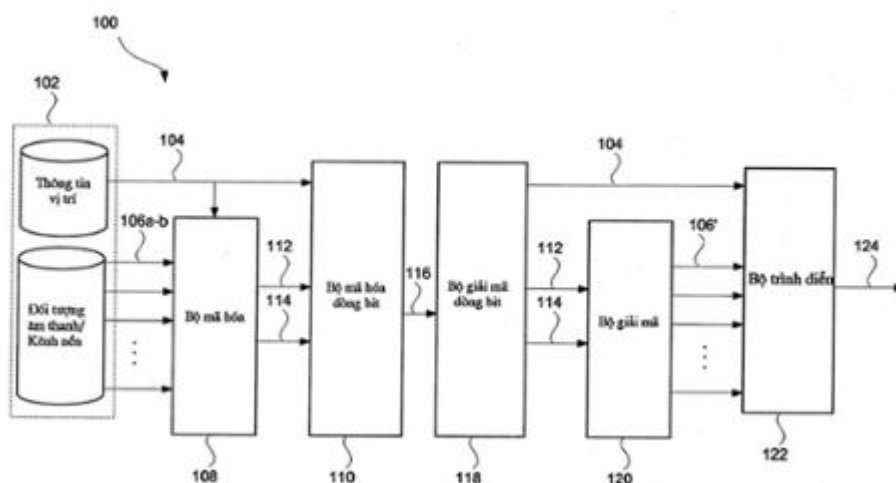
Apollo Building, 3E, Herikerbergweg 1-35, NL-1101 CN Amsterdam,
NETHERLANDS

(72) PURNHAGEN, Heiko (DE); VILLEMOES, Lars (DK); SAMUELSSON, Leif Jonas
(SE); HIRVONEN, Toni (FI).

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

(54) PHƯƠNG PHÁP MÃ HÓA VÀ GIẢI MÃ KHUNG LÁT THỜI GIAN-TÀN SỐ CỦA CẢNH ÂM THANH, BỘ MÃ HÓA VÀ BỘ GIẢI MÃ ĐỀ LẦN LƯỢT MÃ HÓA VÀ GIẢI MÃ KHUNG LÁT THỜI GIAN-TÀN SỐ CỦA CẢNH ÂM THANH, VÀ VẬT GHI ĐƯỢC BẰNG MÁY TÍNH BAO GỒM CÁC LỆNH ĐỌC ĐƯỢC BẰNG MÁY TÍNH ĐỂ THỰC HIỆN PHƯƠNG PHÁP NÀY

(57) Sáng chế theo các phương án làm ví dụ đề xuất phương pháp mã hóa và phương pháp giải mã, và bộ mã hóa và bộ giải mã kết hợp, để mã hóa và giải mã cảnh âm thanh mà ít nhất có một hoặc nhiều đối tượng âm thanh (106a). Bộ mã hóa (108, 110) tạo ra dòng bit (116) mà bao gồm các tín hiệu giảm mức (112) và thông tin phụ có các phần tử ma trận riêng lẻ (114) của ma trận tái cấu trúc cho phép tái cấu trúc một hoặc nhiều đối tượng âm thanh (106a) trong bộ giải mã (120). Sáng chế cũng đề xuất vật ghi đọc được bằng máy tính bao gồm các lệnh đọc được bằng máy tính để thực hiện phương pháp này.



Lĩnh vực kỹ thuật được đề cập

Nói chung, trong bản mô tả này, sáng chế đề cập đến lĩnh vực mã hóa và giải mã âm thanh. Cụ thể hơn, sáng chế đề cập tới việc mã hóa và giải mã cảnh âm thanh bao gồm các đối tượng âm thanh.

Tình trạng kỹ thuật của sáng chế

Hiện có các hệ thống mã hóa âm thanh để mã hóa âm thanh không gian theo tham số. Ví dụ, hệ thống mã hóa vòm MPEG mô tả hệ thống dùng để mã hóa đa kênh âm thanh không gian theo tham số. Hệ thống mã hóa đối tượng âm thanh không gian MPEG (SAOC-Spatial Audio Object Coding) mô tả hệ thống dùng để mã hóa các đối tượng âm thanh theo tham số.

Xét về phía bộ mã hóa, các hệ thống này thường giảm mức các kênh/các đối tượng thành tín hiệu giảm mức, mà thường là tín hiệu giảm mức mono (một kênh) hoặc tín hiệu giảm mức stereo (hai kênh), và trích thông tin phụ mô tả các đặc tính của các kênh/các đối tượng bằng các tham số như các chênh lệch mức độ và tương quan chéo. Sau đó, tín hiệu giảm mức và thông tin phụ được mã hóa và được gửi đến phía bộ giải mã. Ở phía bộ giải mã, các kênh/các đối tượng được tái cấu trúc, nghĩa là được tính gần đúng, từ tín hiệu giảm mức dưới sự kiểm tra các tham số của thông tin phụ.

Các hệ thống này có nhược điểm là việc tái cấu trúc thường phức tạp về mặt toán học và thường phải dựa trên các giả thuyết về các đặc tính của nội dung âm thanh mà không được mô tả một cách rõ ràng bởi các tham số được gửi dưới dạng thông tin phụ. Ví dụ, các giả thuyết này có thể cho rằng các kênh/các đối tượng được coi là không tương quan với nhau trừ khi tham số về tương quan chéo được gửi, hoặc tín hiệu giảm mức của các kênh/các đối tượng được tạo ra theo cách riêng. Hơn nữa, độ phức tạp về mặt toán học và nhu cầu về các giả định bổ sung gia tăng một cách đáng kể do số lượng các kênh tín hiệu giảm mức gia tăng.

Hơn thế nữa, các giả thuyết cần thiết được phản ánh về mặt bản chất trong các chi tiết thuật toán của bước xử lý được ứng dụng ở phía bộ giải mã. Điều đó cần được hiểu rằng rất nhiều tin tức được chứa đựng ở phía bộ giải mã. Điều này có nhược điểm là các thuật toán này khó có thể được nâng cấp hoặc cải biến một khi các bộ giải mã được sử dụng trong, ví dụ, các thiết bị tiêu dùng mà khó hoặc thậm chí không thể nâng cấp được.

Bản chất kỹ thuật của sáng chế

Để khắc phục các nhược điểm nêu trên, mục đích của sáng chế là đề xuất bộ mã hóa và bộ giải mã và các phương pháp kết hợp mà tạo ra việc tái cấu trúc linh hoạt hơn và ít phức tạp hơn cho các đối tượng âm thanh.

Tổng quan về bộ mã hóa

Theo khía cạnh thứ nhất, sáng chế theo các phương án làm ví dụ của nó đề xuất phương pháp mã hóa, bộ mã hóa, và sản phẩm là chương trình máy tính để mã hóa. Nói chung các phương pháp, bộ mã hóa và sản phẩm là chương trình máy tính do sáng chế đề xuất có thể nhìn chung là có cùng ưu điểm và các dấu hiệu.

Sáng chế theo các phương án làm ví dụ của nó đề xuất phương pháp mã hóa khung lát thời gian/tần số của cảnh âm thanh mà ít nhất là bao gồm N đối tượng âm thanh. Phương pháp này bao gồm các bước: tiếp nhận N đối tượng âm thanh; tạo ra M tín hiệu giảm mức trên cơ sở ít nhất N đối tượng âm thanh; tạo ra ma trận tái cấu trúc có các phần tử ma trận có khả năng tái cấu trúc ít nhất N đối tượng âm thanh từ M tín hiệu giảm mức; và tạo ra dòng bit bao gồm M tín hiệu giảm mức và ít nhất một số phần tử ma trận trong ma trận tái cấu trúc.

Số lượng N đối tượng âm thanh có thể bằng hoặc lớn hơn một. Số lượng M tín hiệu giảm mức có thể bằng hoặc lớn hơn một.

Do đó, nhờ phương pháp này dòng bit đã được tạo ra bao gồm M tín hiệu giảm mức và ít nhất một số phần tử ma trận trong ma trận tái cấu trúc dưới dạng thông tin phụ. Do gộp các phần tử ma trận riêng lẻ trong ma trận tái cấu trúc trong dòng bit, nên ít đòi hỏi sự thông minh ở phía bộ giải mã. Ví dụ, ở phía bộ giải mã không có nhu cầu phải tính toán phức tạp trong ma trận tái cấu trúc trên cơ sở các tham số đối tượng được truyền và các giả định bổ sung. Do đó, độ phức tạp toán học ở phía bộ giải mã được giảm một cách đáng kể. Hơn nữa, độ linh hoạt liên quan đến số lượng tín hiệu

giảm mức được tăng lên so với các phương pháp đã biết do độ phức tạp của phương pháp không phụ thuộc vào số lượng các tín hiệu giảm mức được sử dụng.

Nói chung, khi được dùng trong bản mô tả này, *cảnh âm thanh* chỉ môi trường âm thanh ba chiều mà bao gồm các phần tử âm thanh được kết hợp với các vị trí trong không gian ba chiều mà có thể tiến hành phát lại trên hệ thống âm thanh.

Khi được dùng trong bản mô tả này *đối tượng âm thanh* chỉ phần tử trong cảnh âm thanh. Thông thường, đối tượng âm thanh bao gồm tín hiệu âm thanh và thông tin bổ sung như vị trí của đối tượng trong không gian ba chiều. Thông thường, thông tin bổ sung được dùng để tiến hành một cách tối ưu đối tượng âm thanh trên hệ thống phát lại đã định.

Khi được dùng trong bản mô tả này *tín hiệu giảm mức* chỉ tín hiệu là tổ hợp của ít nhất N đối tượng âm thanh. Các tín hiệu khác trong cảnh âm thanh, như các kênh nền (sẽ được mô tả dưới đây), cũng có thể được tổ hợp thành dạng tín hiệu giảm mức. Ví dụ, M tín hiệu giảm mức có thể tương ứng với việc kết xuất cảnh âm thanh đối với cấu hình loa đã định, ví dụ, cấu hình 5.1 chuẩn. Thông thường, số lượng các tín hiệu giảm mức, dưới đây được tổ hợp bằng M, (nhưng không nhất thiết) thấp hơn tổng số lượng các đối tượng âm thanh và các kênh nền, giải thích tại sao M tín hiệu giảm mức được gọi là tín hiệu giảm mức.

Thông thường, các hệ thống mã hóa/giải mã âm thanh chia không gian tần số-thời gian thành các khung lát thời gian/tần số, ví dụ bằng cách áp dụng các dàn lọc thích hợp cho các tín hiệu âm thanh đầu vào. Nói chung, khung lát thời gian/tần số được hiểu là phần có không gian tần số-thời gian tương ứng với khoảng thời gian và dải phụ tần số. Thông thường, khoảng thời gian có thể tương ứng với độ dài của khung thời gian dùng trong hệ thống mã hóa/giải mã âm thanh. Thông thường, dải phụ tần số có thể tương ứng với một hoặc một số các dải phụ tần số lân cận xác định được bởi dàn lọc dùng trong hệ thống mã hóa/giải mã. Trong trường hợp dải phụ tần số tương ứng với một vài dải phụ tần số lân cận xác định được bởi dàn lọc, điều này cho phép có được các dải phụ tần số không đều trong quy trình giải mã tín hiệu âm thanh, ví dụ, các dải phụ tần số càng rộng dành cho các tần số càng cao của tín hiệu âm thanh. Trong trường hợp dải rộng, mà vì đó hệ thống mã hóa/giải mã âm thanh hoạt động trên toàn bộ vùng tần số, dải phụ tần số của khung lát thời gian/tần số có thể tương ứng với toàn bộ vùng tần số. Phương pháp nêu trên đề xuất các bước mã hóa để

mã hóa cảnh âm thanh trong một khung lát thời gian/tần số này. Tuy nhiên, cần phải hiểu rằng phương pháp này có thể được lặp lại đối với mỗi khung lát thời gian/tần số của hệ thống mã hóa/giải mã âm thanh. Cũng cần phải hiểu rằng một số khung lát thời gian/tần số có thể được mã hóa một cách đồng thời. Thông thường, các khung lát thời gian/tần số lân cận có thể chồng lấn bit theo thời gian và/hoặc tần số. Ví dụ, sự chồng lấn theo thời gian có thể là tương đương với phép nội suy tuyến tính của các phần tử trong ma trận tái cấu trúc theo thời gian, nghĩa là từ một khoảng thời gian này đến khoảng thời gian khác. Tuy nhiên, phần mô tả này hướng đến các phần khác của hệ thống mã hóa/giải mã và sự chồng lấn bất kỳ theo thời gian và/hoặc tần số giữa các khung lát thời gian/tần số lân cận được dành cho chuyên gia trong lĩnh vực kỹ thuật này thực hành.

Theo các phương án làm ví dụ, M tín hiệu giảm mức được sắp xếp trong trường thứ nhất của dòng bit đang sử dụng định dạng thứ nhất, và các phần tử ma trận được sắp xếp trong trường thứ hai của dòng bit đang sử dụng định dạng thứ hai, nhờ đó cho phép bộ giải mã chỉ hỗ trợ định dạng thứ nhất để giải mã và phát lại M tín hiệu giảm mức trong trường thứ nhất và bỏ qua các phần tử ma trận trong trường thứ hai. Điều này có lợi ở chỗ M tín hiệu giảm mức trong dòng bit là tương thích ngược với các bộ giải mã kế thừa mà không thực hiện tái cấu trúc đối tượng âm thanh. Nói cách khác, các bộ giải mã kế thừa có thể vẫn giải mã và phát lại M tín hiệu giảm mức của dòng bit, ví dụ, bằng cách ánh xạ mỗi tín hiệu giảm mức đến đầu ra kênh của bộ giải mã.

Theo các phương án làm ví dụ, phương pháp này có thể còn bao gồm bước tiếp nhận dữ liệu vị trí tương ứng với mỗi đối tượng trong số N đối tượng âm thanh, trong đó M tín hiệu giảm mức được tạo ra trên cơ sở dữ liệu vị trí. Thông thường, dữ liệu vị trí kết hợp mỗi đối tượng âm thanh với vị trí trong không gian ba chiều. Vị trí của đối tượng âm thanh có thể thay đổi theo thời gian. Bằng cách sử dụng dữ liệu vị trí khi giảm mức các đối tượng âm thanh, các đối tượng âm thanh sẽ được trộn trong M tín hiệu giảm mức sao cho, ví dụ, nếu M tín hiệu giảm mức được nghe thấy trên hệ thống bằng M kênh đầu ra, các đối tượng âm thanh sẽ kêu vang như là chúng được đặt gần đúng ở các vị trí tương ứng của chúng. Ví dụ, điều này là có lợi nếu M tín hiệu giảm mức là tương thích ngược với bộ giải mã kế thừa.

Theo các phương án làm ví dụ, các phần tử ma trận trong ma trận tái cấu trúc là biến thời gian và tần số. Nói cách khác, các phần tử ma trận trong ma trận tái cấu trúc

có thể là khác nhau đối với các khung lát thời gian/tần số khác nhau. Độ linh hoạt lớn trong việc tái cấu trúc các đối tượng âm thanh đạt được theo cách này.

Theo các phương án làm ví dụ, cảnh âm thanh còn có các kênh nền. Ví dụ, điều này là thông thường trong các ứng dụng âm thanh trong rạp chiếu phim nơi mà nội dung âm thanh bao gồm các kênh nền bổ sung cho các đối tượng âm thanh. Trong các trường hợp này, M tín hiệu giảm mức có thể được tạo ra trên cơ sở ít nhất N đối tượng âm thanh và các kênh nền. Nói chung, kênh nền được hiểu là tín hiệu âm thanh tương ứng với vị trí cố định trong không gian ba chiều. Ví dụ, kênh nền có thể tương ứng với một trong số các kênh đầu ra của hệ thống mã hóa/giải mã âm thanh. Như vậy, kênh nền có thể được dịch nghĩa như đối tượng âm thanh có vị trí kết hợp trong không gian ba chiều như vị trí của một trong số các loa phát ra của hệ thống mã hóa/giải mã âm thanh. Do đó, kênh nền có thể được kết hợp với nhãn mà chỉ thể hiện vị trí của loa phát ra tương ứng.

Khi cảnh âm thanh bao gồm các kênh nền, ma trận tái cấu trúc có thể có các phần tử ma trận mà có khả năng tái cấu trúc các kênh nền từ M tín hiệu giảm mức.

Trong một số tình huống, cảnh âm thanh có thể bao gồm số lượng lớn các đối tượng. Để giảm độ phức tạp và lượng dữ liệu cần để thể hiện cảnh âm thanh, cảnh âm thanh có thể được đơn giản hóa bằng cách giảm số lượng các đối tượng âm thanh. Do đó, nếu cảnh âm thanh nguyên gốc có K đối tượng âm thanh, trong đó $K > N$, phương pháp này có thể còn bao gồm các bước tiếp nhận K đối tượng âm thanh, và giảm K đối tượng âm thanh thành N đối tượng âm thanh bằng cách gộp nhóm K đối tượng thành N nhóm và thể hiện mỗi nhóm này bằng một đối tượng âm thanh.

Để đơn giản hóa cảnh âm thanh, phương pháp này có thể còn bao gồm bước tiếp nhận dữ liệu vị trí tương ứng với mỗi đối tượng trong số K đối tượng âm thanh, trong đó việc gộp nhóm K đối tượng thành N nhóm được dựa trên cơ sở khoảng cách vị trí giữa K đối tượng như được tạo ra bởi dữ liệu vị trí của K đối tượng âm thanh. Ví dụ, các đối tượng âm thanh mà gần với nhau về mặt vị trí trong không gian ba chiều có thể được gộp nhóm với nhau.

Như nêu trên, phương pháp theo các phương án làm ví dụ là linh hoạt đối với nhiều tín hiệu giảm mức được sử dụng. Cụ thể là, phương pháp này có thể được sử dụng theo cách có lợi khi có nhiều hơn hai tín hiệu giảm mức, nghĩa là khi M lớn hơn hai. Ví dụ, năm hoặc bảy tín hiệu giảm mức tương ứng với các tổ hợp âm thanh 5.1

hoặc 7.1 thông thường có thể được sử dụng. Trái với các hệ thống đã biết, điều này là có lợi do độ phức tạp toán học của các nguyên lý mã hóa được đề xuất vẫn giữ nguyên mà không phụ thuộc vào số lượng các tín hiệu giảm mức được sử dụng.

Để tiếp tục cho phép tái cấu trúc nâng cao N đối tượng âm thanh, phương pháp này có thể còn bao gồm các bước: tạo ra L tín hiệu bổ trợ từ N đối tượng âm thanh; gộp các phần tử ma trận trong ma trận tái cấu trúc mà cho phép tái cấu trúc ít nhất N đối tượng âm thanh từ M tín hiệu giảm mức và L tín hiệu bổ trợ; và gộp L tín hiệu bổ trợ trong dòng bit. Do đó, các tín hiệu bổ trợ phục vụ như các tín hiệu hỗ trợ mà, ví dụ, có thể thu bắt các khía cạnh của các đối tượng âm thanh mà khó tái cấu trúc từ các tín hiệu giảm mức. Các tín hiệu bổ trợ còn có thể dựa trên cơ sở các kênh nền. Số lượng tín hiệu bổ trợ có thể bằng hoặc lớn hơn một.

Theo một phương án làm ví dụ, các tín hiệu bổ trợ có thể tương ứng với các đối tượng âm thanh đặc biệt quan trọng, như đối tượng âm thanh thể hiện đoạn hội thoại. Do đó, ít nhất một trong số L tín hiệu bổ trợ có thể như một trong số N đối tượng âm thanh. Điều này cho phép các đối tượng quan trọng được kết xuất với chất lượng cao hơn so với khi chúng được tái cấu trúc chỉ từ M kênh giảm mức. Trên thực tế, một số đối tượng âm thanh có thể được ưu tiên và/hoặc được gắn nhãn bởi người tạo lập nội dung âm thanh như các đối tượng âm thanh mà tốt hơn là được gộp một cách riêng biệt dưới dạng các đối tượng phụ trợ. Hơn thế nữa, điều này tạo ra sự cải biến/xử lý các đối tượng đó trước để cho ít giả tạo hơn. Do cân đối giữa tốc độ và chất lượng bit, cũng có thể gửi tổ hợp của hai hoặc nhiều đối tượng âm thanh dưới dạng tín hiệu bổ trợ. Nói cách khác, ít nhất một trong số L tín hiệu bổ trợ có thể được tạo ra dưới dạng tổ hợp của ít nhất hai trong số N đối tượng âm thanh.

Theo một phương án làm ví dụ, các tín hiệu bổ trợ thể hiện các kích thước tín hiệu của các đối tượng âm thanh mà đã bị mất trong quá trình tạo ra M tín hiệu giảm mức, ví dụ, do số lượng các đối tượng độc lập thường là cao hơn so với số lượng của các kênh giảm mức hoặc do hai đối tượng được kết hợp với các vị trí sao cho chúng được trộn trong cùng một tín hiệu giảm mức. Một ví dụ về trường hợp này là tình huống mà hai đối tượng chỉ nằm cách nhau theo chiều dọc song có chung vị trí khi được chiếu lên mặt phẳng nằm ngang, mà có nghĩa là chúng thường sẽ được kết xuất đối với cùng các kênh giảm mức của tổ hợp loa vòm 5.1 tiêu chuẩn, trong đó tất cả các loa được đặt trên cùng mặt phẳng nằm ngang. Cụ thể là, M tín hiệu giảm mức mở

rộng siêu mặt phẳng trong không gian tín hiệu. Bằng cách tạo ra các tổ hợp tuyến tính của M tín hiệu giảm mức, chỉ các tín hiệu âm thanh mà nằm trong siêu mặt phẳng có thể được tái cấu trúc. Để nâng cao việc tái cấu trúc, các tín hiệu bổ trợ có thể được gộp mà không nằm trong siêu mặt phẳng, nhờ đó cũng cho phép tái cấu trúc các tín hiệu mà không nằm trong siêu mặt phẳng. Nói cách khác, theo các phương án làm ví dụ, ít nhất một trong số các tín hiệu bổ trợ không nằm trong siêu mặt phẳng được mở rộng bởi M tín hiệu giảm mức. Ví dụ, ít nhất một trong số các tín hiệu bổ trợ có thể là vuông góc với siêu mặt phẳng được mở rộng bởi M tín hiệu giảm mức.

Sáng chế theo các phương án làm ví dụ đề xuất vật ghi đọc được bằng máy tính bao gồm các lệnh mã hóa bằng máy tính được làm thích ứng để thực hiện phương pháp bất kỳ theo khía cạnh thứ nhất khi thực hiện trên thiết bị có khả năng xử lý.

Sáng chế theo các phương án làm ví dụ đề xuất bộ mã hóa để mã hóa khung lát thời gian/tần số của cảnh âm thanh ít nhất có N đối tượng âm thanh, bao gồm: bộ phận tiếp nhận được tạo cấu hình để tiếp nhận N đối tượng âm thanh; bộ phận tạo ra tín hiệu giảm mức được tạo cấu hình để tiếp nhận N đối tượng âm thanh từ bộ phận tiếp nhận và để tạo ra M tín hiệu giảm mức trên cơ sở ít nhất N đối tượng âm thanh; bộ phận phân tích được tạo cấu hình để tạo ra ma trận tái cấu trúc có các phần tử ma trận có khả năng tái cấu trúc ít nhất N đối tượng âm thanh từ M tín hiệu giảm mức; và bộ phận tạo ra dòng bit được tạo cấu hình để tiếp nhận M tín hiệu giảm mức từ bộ phận tạo ra tín hiệu giảm mức và ma trận tái cấu trúc từ bộ phận phân tích và để tạo ra dòng bit có M tín hiệu giảm mức và ít nhất một số phần tử ma trận trong ma trận tái cấu trúc.

Tổng quan về bộ giải mã

Sáng chế theo các phương án làm ví dụ theo khía cạnh thứ hai, đề xuất phương pháp giải mã, thiết bị giải mã, và sản phẩm là chương trình máy tính để giải mã. Nói chung, phương pháp, thiết bị và sản phẩm là chương trình máy tính được đề xuất có thể có cùng các ưu điểm và các dấu hiệu.

Nói chung, các ưu điểm liên quan đến các dấu hiệu và các tổ hợp như được mô tả trong phần tổng quan của bộ mã hóa nêu trên có thể có hiệu lực đối với các dấu hiệu và các tổ hợp tương ứng đối với bộ giải mã.

Sáng chế theo các phương án làm ví dụ, đề xuất phương pháp giải mã khung lát thời gian-tần số của cảnh âm thanh mà ít nhất có N đối tượng âm thanh, phương pháp

này bao gồm các bước: tiếp nhận dòng bit có M tín hiệu giảm mức và ít nhất một số phần tử ma trận trong ma trận tái cấu trúc; tạo ra ma trận tái cấu trúc bằng cách sử dụng các phần tử ma trận; và tái cấu trúc N đối tượng âm thanh từ M tín hiệu giảm mức bằng cách sử dụng ma trận tái cấu trúc.

Theo các phương án làm ví dụ, M tín hiệu giảm mức được sắp xếp trong trường thứ nhất của dòng bit bằng cách sử dụng định dạng thứ nhất, và các phần tử ma trận được sắp xếp trong trường thứ hai của dòng bit bằng cách sử dụng định dạng thứ hai, nhờ đó cho phép bộ giải mã chỉ hỗ trợ định dạng thứ nhất để giải mã và phát lại M tín hiệu giảm mức trong trường thứ nhất và bỏ qua các phần tử ma trận trong trường thứ hai.

Theo các phương án làm ví dụ, các phần tử ma trận trong ma trận tái cấu trúc là biến thời gian và tần số.

Theo các phương án làm ví dụ, cảnh âm thanh còn bao gồm các kênh nền, phương pháp này còn bao gồm bước tái cấu trúc các kênh nền từ M tín hiệu giảm mức bằng cách sử dụng ma trận tái cấu trúc.

Theo các phương án làm ví dụ, số lượng M tín hiệu giảm mức lớn hơn hai.

Theo các phương án làm ví dụ, phương pháp này còn bao gồm các bước: tiếp nhận L tín hiệu bổ trợ được tạo ra từ N đối tượng âm thanh; tái cấu trúc N đối tượng âm thanh từ M tín hiệu giảm mức và L tín hiệu bổ trợ bằng cách sử dụng ma trận tái cấu trúc, trong đó ma trận tái cấu trúc bao gồm các phần tử ma trận mà cho phép tái cấu trúc ít nhất N đối tượng âm thanh từ M tín hiệu giảm mức và L tín hiệu bổ trợ.

Theo các phương án làm ví dụ, ít nhất một trong số L tín hiệu bổ trợ là như một trong số N đối tượng âm thanh.

Theo các phương án làm ví dụ, ít nhất một trong số L tín hiệu bổ trợ là tổ hợp của N đối tượng âm thanh.

Theo các phương án làm ví dụ, M tín hiệu giảm mức mở rộng siêu mặt phẳng, và trong đó ít nhất một trong số các tín hiệu bổ trợ không nằm trong siêu mặt phẳng được mở rộng bởi M tín hiệu giảm mức.

Theo các phương án làm ví dụ, ít nhất một trong số các tín hiệu bổ trợ mà không nằm trong siêu mặt phẳng vuông góc với siêu mặt phẳng được mở rộng bởi M tín hiệu giảm mức.

Như nêu trên, các hệ thống mã hóa/giải mã âm thanh thường hoạt động trong miền tần số. Do đó, các hệ thống mã hóa/giải mã âm thanh thực hiện các chuyển đổi thời gian/tần số của các tín hiệu âm thanh bằng cách sử dụng các dàn lọc. Các loại khác của các chuyển đổi thời gian/tần số có thể được sử dụng. Ví dụ M tín hiệu giảm mức có thể được biểu diễn đối với miền tần số thứ nhất và ma trận tái cấu trúc có thể được biểu diễn đối với miền tần số thứ hai. Để giảm gánh nặng tính toán trong bộ giải mã, có lợi nếu chọn miền tần số thứ nhất và miền tần số thứ hai theo cách thông minh. Ví dụ, miền tần số thứ nhất và miền tần số thứ hai có thể được chọn như cùng miền tần số như miền biến đổi cosin rời rạc hiệu chỉnh (MDCT - Modified Discrete Cosine Transform). Theo cách này có thể tránh được sự biến đổi M tín hiệu giảm mức từ miền tần số thứ nhất đến miền thời gian được tiếp nối bởi sự biến đổi đến miền tần số thứ hai trong bộ giải mã. Theo cách khác, có thể chọn miền tần số thứ nhất và miền tần số thứ hai sao cho việc biến đổi từ miền tần số thứ nhất đến miền tần số thứ hai có thể được tiến hành cùng nhau sao cho không nhất thiết phải đi hết cả quá trình qua miền thời gian ở giữa.

Phương pháp này có thể còn bao gồm bước tiếp nhận dữ liệu vị trí tương ứng với N đối tượng âm thanh, và kết xuất N đối tượng âm thanh bằng cách sử dụng dữ liệu vị trí để tạo ra ít nhất một kênh âm thanh đầu ra. Theo cách này, việc tái cấu trúc N đối tượng âm thanh được ánh xạ trên các kênh đầu ra của hệ thống mã hóa/giải mã âm thanh trên cơ sở vị trí của chúng trong không gian ba chiều.

Tốt hơn, nếu việc kết xuất được thực hiện trong miền tần số. Để giảm gánh nặng tính toán trong bộ giải mã, tốt hơn nếu miền tần số của việc kết xuất được chọn theo cách thông minh đối với miền tần số trong đó các đối tượng âm thanh được tái cấu trúc. Ví dụ, nếu ma trận tái cấu trúc được biểu diễn đối với miền tần số thứ hai tương ứng với dàn lọc thứ hai, và việc kết xuất được thực hiện trong miền tần số thứ ba tương ứng với dàn lọc thứ ba, tốt hơn là dàn lọc thứ hai và dàn lọc thứ ba được chọn để ít nhất là một phần của cùng dàn lọc. Ví dụ, dàn lọc thứ hai và dàn lọc thứ ba có thể có miền bộ lọc gương cầu phương (Quadrature Mirror Filter - QMF). Theo cách khác, miền tần số thứ hai và miền tần số thứ ba có thể có dàn lọc MDCT. Theo phương án làm ví dụ, dàn lọc thứ ba có thể được làm bằng chuỗi dàn lọc, như dàn lọc QMF được tiếp nối bởi dàn lọc Nyquist. Nếu vậy, ít nhất một trong số các dàn lọc của

chuỗi này (dàn lọc thứ nhất trong chuỗi này) là như dàn lọc thứ hai. Theo cách này, dàn lọc thứ hai và dàn lọc thứ ba có thể coi là ít nhất một phần của cùng dàn lọc.

Theo các phương án làm ví dụ, sáng chế đề xuất vật ghi đọc được bằng máy tính có các lệnh mã hóa bằng máy tính được làm thích ứng để thực hiện phương pháp bất kỳ theo khía cạnh thứ hai khi thực hiện trên thiết bị có khả năng xử lý.

Theo các phương án làm ví dụ, sáng chế đề xuất bộ giải mã để giải mã khung lát thời gian-tần số trong cảnh âm thanh mà ít nhất có N đối tượng âm thanh, bao gồm: bộ phận tiếp nhận được tạo cấu hình để tiếp nhận dòng bit có M tín hiệu giảm mức và ít nhất một số phần tử ma trận trong ma trận tái cấu trúc; bộ phận tạo ra ma trận tái cấu trúc được tạo cấu hình để tiếp nhận các phần tử ma trận từ bộ phận tiếp nhận và dựa trên đó tạo ra ma trận tái cấu trúc; và bộ phận tái cấu trúc được tạo cấu hình để tiếp nhận ma trận tái cấu trúc từ bộ phận tạo ra ma trận tái cấu trúc và để tái cấu trúc N đối tượng âm thanh từ M tín hiệu giảm mức bằng cách sử dụng ma trận tái cấu trúc.

Mô tả vắn tắt các hình vẽ

Dưới đây, sáng chế theo các phương án làm ví dụ sẽ được mô tả chi tiết hơn và có dựa vào các hình vẽ kèm theo, trong đó:

Hình 1 là hình vẽ dạng sơ đồ thể hiện hệ thống mã hóa/giải mã âm thanh theo các phương án làm ví dụ;

Hình 2 là hình vẽ dạng sơ đồ thể hiện hệ thống mã hóa/giải mã âm thanh có bộ giải mã kế thừa theo các phương án làm ví dụ;

Hình 3 là hình vẽ dạng sơ đồ thể hiện phía mã hóa của hệ thống mã hóa/giải mã âm thanh theo các phương án làm ví dụ;

Hình 4 là sơ đồ của phương pháp mã hóa theo các phương án làm ví dụ;

Hình 5 là hình vẽ dạng sơ đồ thể hiện bộ mã hóa theo các phương án làm ví dụ;

Hình 6 là hình vẽ dạng sơ đồ thể hiện phía bộ giải mã của hệ thống mã hóa/giải mã âm thanh theo các phương án làm ví dụ;

Hình 7 là sơ đồ của phương pháp giải mã theo các phương án làm ví dụ;

Hình 8 là hình vẽ dạng sơ đồ thể hiện phía bộ giải mã của hệ thống mã hóa/giải mã âm thanh theo các phương án làm ví dụ; và

Hình 9 là hình vẽ dạng sơ đồ thể hiện các chuyển đổi thời gian/tần số được thực hiện ở phía bộ giải mã của hệ thống mã hóa/giải mã âm thanh theo các phương án làm ví dụ.

Tất cả các hình vẽ đều ở dạng sơ đồ và nói chung chỉ thể hiện các phần mà cần thiết để minh họa sáng chế, còn các phần khác có thể được bỏ qua hoặc chỉ được đề cập đến. Trừ khi có quy định khác, các số chỉ dẫn tương tự được dùng để chỉ các bộ phận tương tự trên các hình vẽ khác nhau.

Mô tả chi tiết sáng chế

Các phương án làm ví dụ

Hình 1 minh họa hệ thống mã hóa/giải mã 100 để mã hóa/giải mã cảnh âm thanh 102. Hệ thống mã hóa/giải mã 100 bao gồm bộ mã hóa 108, bộ phận tạo ra dòng bit 110, bộ phận giải mã dòng bit 118, bộ giải mã 120, và bộ kết xuất 122.

Cảnh âm thanh 102 được thể hiện bằng một hoặc nhiều đối tượng âm thanh 106a, nghĩa là các tín hiệu âm thanh, như N đối tượng âm thanh. Cảnh âm thanh 102 có thể còn có một hoặc nhiều kênh nền 106b, nghĩa là các tín hiệu mà tương ứng một cách trực tiếp với một kênh trong số các kênh đầu ra của bộ kết xuất 122. Cảnh âm thanh 102 còn được thể hiện bởi siêu dữ liệu có thông tin về vị trí 104. Ví dụ, thông tin về vị trí 104 được bộ kết xuất 122 sử dụng khi kết xuất cảnh âm thanh 102. Thông tin về vị trí 104 có thể kết hợp các đối tượng âm thanh 106a, và cũng có thể là các kênh nền 106b, với vị trí không gian trong không gian ba chiều dưới dạng hàm của thời gian. Siêu dữ liệu có thể còn có loại dữ liệu khác mà là hữu ích để kết xuất cảnh âm thanh 102.

Phần mã hóa của hệ thống 100 bao gồm bộ mã hóa 108 và bộ phận tạo ra dòng bit 110. Bộ mã hóa 108 tiếp nhận các đối tượng âm thanh 106a, các kênh nền 106b nếu thể hiện, và siêu dữ liệu có thông tin về vị trí 104. Dựa trên đó, bộ mã hóa 108 tạo ra Một hoặc nhiều tín hiệu giảm mức 112, như M tín hiệu giảm mức. Để làm ví dụ, các tín hiệu giảm mức 112 có thể tương ứng với các kênh $[L_f R_f C_f L_s R_s LFE]$ của hệ thống âm thanh 5.1. (“L” nghĩa là trái, “R” nghĩa là phải, “C” nghĩa là giữa, “f” nghĩa là trước, “s” nghĩa là vòm, và “LFE” nghĩa là các hiệu ứng tần số thấp).

Bộ mã hóa 108 còn tạo ra thông tin phụ. Thông tin phụ có ma trận tái cấu trúc. Ma trận tái cấu trúc có các phần tử ma trận 114 mà cho phép tái cấu trúc ít nhất các đối tượng âm thanh 106a từ các tín hiệu giảm mức 112. Ma trận tái cấu trúc có thể tiếp tục cho phép tái cấu trúc các kênh nền 106b.

Bộ mã hóa 108 truyền M tín hiệu giảm mức 112, và ít nhất một số phần tử ma trận 114 đến bộ phận tạo ra dòng bit 110. Bộ phận tạo ra dòng bit 110 tạo ra dòng bit

116 có M tín hiệu giảm mức 112 và ít nhất một số phần tử ma trận 114 bằng cách thực hiện việc lượng tử hóa và mã hóa. Bộ phận tạo ra dòng bit 110 còn tiếp nhận siêu dữ liệu bao gồm thông tin về vị trí 104 để có mặt trong dòng bit 116.

Phần giải mã của hệ thống bao gồm bộ phận giải mã dòng bit 118 và bộ giải mã 120. Bộ phận giải mã dòng bit 118 tiếp nhận dòng bit 116 và thực hiện việc giải mã và giải lượng tử hóa để trích M tín hiệu giảm mức 112 và thông tin phụ có ít nhất một số phần tử ma trận 114 trong ma trận tái cấu trúc. Sau đó, M tín hiệu giảm mức 112 và các phần tử ma trận 114 được nạp vào bộ giải mã 120 mà dựa trên đó tạo ra đối tượng tái cấu trúc 106' trong số N đối tượng âm thanh 106a và cũng có thể là các kênh nền 106b. Do đó, đối tượng tái cấu trúc 106' trong số N đối tượng âm thanh là gần đúng trong số N đối tượng âm thanh 106a và cũng có thể là của các kênh nền 106b.

Để làm ví dụ, nếu các tín hiệu giảm mức 112 tương ứng với các kênh $[Lf\ Rf\ Cf\ Ls\ Rs\ LFE]$ của cấu hình 5.1, bộ giải mã 120 có thể tái cấu trúc các đối tượng 106' chỉ bằng cách sử dụng các kênh toàn dải $[Lf\ Rf\ Cf\ Ls\ Rs\ LFE]$, do đó bỏ qua LFE. Điều này cũng được áp dụng cho các cấu hình kênh khác. Kênh LFE của tín hiệu giảm mức 112 có thể được gửi (về cơ bản không biến đổi) đến bộ kết xuất 122.

Các đối tượng âm thanh 106' đã được tái cấu trúc, cùng với thông tin về vị trí 104, sau đó được nạp vào bộ kết xuất 122. Trên cơ sở các đối tượng âm thanh 106' đã được tái cấu trúc và thông tin về vị trí 104, bộ kết xuất 122 kết xuất tín hiệu đầu ra 124 có định dạng mà là thích hợp để phát lại trên loa phát mong muốn hoặc cấu hình kiểu tai nghe. Định dạng đầu ra thông thường là tổ hợp âm thanh vòm 5.1 chuẩn (3 loa phát trước, 2 loa vòm, và 1 loa hiệu ứng tần số thấp, LFE) hoặc tổ hợp 7.1 + 4 (3 loa phát trước, 4 loa vòm, 1 loa LFE, và 4 loa tầm cao).

Theo một số phương án thực hiện, cảnh âm thanh nguyên gốc có thể có số lượng lớn các đối tượng âm thanh. Việc xử lý số lượng lớn các đối tượng âm thanh dẫn đến chi phí về độ phức tạp tính toán cao. Ngoài ra, lượng thông tin phụ (thông tin về vị trí 104 và việc tái cấu trúc các phần tử ma trận 114) cần được gắn vào trong dòng bit 116 tùy theo số lượng các đối tượng âm thanh. Thông thường, lượng thông tin phụ tăng tuyến tính với số lượng các đối tượng âm thanh. Do đó, để giảm độ phức tạp tính toán và/hoặc để giảm tốc độ bit cần thiết để mã hóa cảnh âm thanh, có thể có lợi nếu giảm số lượng các đối tượng âm thanh trước khi mã hóa. Đối với mục đích này, hệ thống mã hóa/giải mã âm thanh 100 có thể còn bao gồm mô đun đơn giản hóa

cảnh (không được thể hiện) nằm trước bộ mã hóa 108. Môđun đơn giản hóa cảnh lấy các đối tượng âm thanh nguyên gốc và cũng có thể là các kênh nền dưới dạng đầu vào và thực hiện việc xử lý đối với các đối tượng âm thanh đầu ra 106a. Môđun đơn giản hóa cảnh giảm số lượng các đối tượng âm thanh nguyên gốc, gọi là số lượng K thành số lượng N đối tượng âm thanh 106a khả thi hơn bằng cách thực hiện việc gộp nhóm. Chính xác hơn, môđun đơn giản hóa cảnh tổ chức K đối tượng âm thanh nguyên gốc và cũng có thể là các kênh nền thành N nhóm. Thông thường, các nhóm được xác định trên cơ sở sự lân cận về không gian trong cảnh âm thanh của K đối tượng âm thanh/kênh nền nguyên gốc. Để xác định sự lân cận về không gian, môđun đơn giản hóa cảnh có thể lấy thông tin về vị trí của các đối tượng âm thanh nguyên gốc/các kênh nền dưới dạng đầu vào. Khi môđun đơn giản hóa cảnh có N nhóm được tạo ra, nó xử lý để thể hiện mỗi nhóm bằng một đối tượng âm thanh. Ví dụ, đối tượng âm thanh thể hiện nhóm có thể được tạo ra dưới dạng tổng của các đối tượng âm thanh/các kênh nền tạo ra phần của nhóm. Cụ thể hơn, nội dung âm thanh của các đối tượng âm thanh/các kênh nền có thể được thêm vào để tạo ra nội dung âm thanh của đối tượng âm thanh minh họa. Hơn nữa, các vị trí của các đối tượng âm thanh/các kênh nền trong nhóm có thể được tính trung bình để tìm ra vị trí của đối tượng âm thanh minh họa. Môđun đơn giản hóa cảnh có các vị trí của các đối tượng âm thanh minh họa trong dữ liệu vị trí 104. Hơn nữa, môđun đơn giản hóa cảnh phát ra các đối tượng âm thanh minh họa mà cấu thành N đối tượng âm thanh 106a được thể hiện trên Hình 1.

M tín hiệu giảm mức 112 có thể được sắp xếp trong trường thứ nhất của dòng bit 116 bằng cách sử dụng định dạng thứ nhất. Các phần tử ma trận 114 có thể được sắp xếp trong trường thứ hai của dòng bit 116 bằng cách sử dụng định dạng thứ hai. Theo cách này, bộ giải mã chỉ hỗ trợ định dạng thứ nhất có khả năng giải mã và phát lại M tín hiệu giảm mức 112 trong trường thứ nhất và bỏ qua các phần tử ma trận 114 trong trường thứ hai.

Hệ thống mã hóa/giải mã âm thanh 100 được thể hiện trên Hình 1 hỗ trợ cả định dạng thứ nhất lẫn định dạng thứ hai. Chính xác hơn, bộ giải mã 120 được tạo cấu hình để giải nghĩa định dạng thứ nhất và định dạng thứ hai, có nghĩa rằng nó có khả năng tái cấu trúc các đối tượng 106' trên cơ sở M tín hiệu giảm mức 112 và các phần tử ma trận 114.

Hình 2 minh họa hệ thống mã hóa/giải mã âm thanh 200. Phần mã hóa 108, 110 của hệ thống 200 tương ứng với phần mã hóa được thể hiện trên Hình 1. Tuy nhiên, phần giải mã của hệ thống mã hóa/giải mã âm thanh 200 khác biệt so với phần giải mã của hệ thống mã hóa/giải mã âm thanh 100 được thể hiện trên Hình 1. Hệ thống mã hóa/giải mã âm thanh 200 bao gồm bộ giải mã kế thừa 230 mà hỗ trợ định dạng thứ nhất song không hỗ trợ định dạng thứ hai. Do đó, bộ giải mã kế thừa 230 của hệ thống mã hóa/giải mã âm thanh 200 không có khả năng tái cấu trúc các đối tượng âm thanh/các kênh nền 106a-b. Tuy nhiên, do bộ giải mã kế thừa 230 hỗ trợ định dạng thứ nhất, hệ thống này vẫn có thể giải mã M tín hiệu giảm mức 112 để tạo ra đầu ra 224 vốn là biểu diễn trên cơ sở kênh, như biểu diễn 5.1, thích hợp để phát lại trực tiếp trên tổ hợp loa đa kênh tương ứng. Đặc tính này của các tín hiệu giảm mức được gọi là sự tương thích sau mà có nghĩa rằng bộ giải mã kế thừa mà không hỗ trợ định dạng thứ hai, nghĩa là không có khả năng giải nghĩa thông tin phụ bao gồm các phần tử ma trận 114, cũng vẫn có thể giải mã và phát lại M tín hiệu giảm mức 112.

Hoạt động ở phía bộ mã hóa của hệ thống mã hóa/giải mã âm thanh 100 sẽ được mô tả một cách chi tiết hơn có dựa vào Hình 3 và sơ đồ được thể hiện trên Hình 4.

Hình 4 minh họa bộ mã hóa 108 và bộ phận tạo ra dòng bit 110 được thể hiện trên Hình 1 một cách chi tiết hơn. Bộ mã hóa 108 có bộ phận tiếp nhận (không được thể hiện), bộ phận tạo ra tín hiệu giảm mức 318 và bộ phận phân tích 328.

Ở bước E02, bộ phận tiếp nhận của bộ mã hóa 108 tiếp nhận N đối tượng âm thanh 106a và các kênh nền 106b nếu có. Bộ mã hóa 108 còn có thể tiếp nhận dữ liệu vị trí 104. Bằng cách sử dụng chỉ dẫn vector, N đối tượng âm thanh có thể được biểu thị bởi vector $S = [S1\ S2\ \dots\ SN]^T$, và các kênh nền bởi vector B . N đối tượng âm thanh và các kênh nền có thể cùng nhau được biểu diễn bởi vector $A = [B^T\ S^T]^T$.

Ở bước E04, bộ phận tạo ra tín hiệu giảm mức 318 tạo ra M tín hiệu giảm mức 112 từ N đối tượng âm thanh 106a và các kênh nền 106b nếu có mặt. Bằng cách sử dụng chỉ dẫn vector, M tín hiệu giảm mức có thể được biểu thị bởi vector $D = [D1\ D2\ \dots\ DM]^T$ có M tín hiệu giảm mức. Nói chung, tín hiệu giảm mức trong số các tín hiệu là tổ hợp của các tín hiệu, như tổ hợp tuyến tính của các tín hiệu. Để làm ví dụ, M tín hiệu giảm mức có thể tương ứng với cấu hình loa cụ thể, như cấu hình loa $[Lf\ Rf\ Cf\ Ls\ Rs\ LFE]$ theo cấu hình loa 5.1.

Bộ phận tạo ra tín hiệu giảm mức 318 có thể sử dụng thông tin về vị trí 104 khi tạo ra M tín hiệu giảm mức, sao cho các đối tượng có thể được tổ hợp thành các tín hiệu giảm mức khác nhau trên cơ sở vị trí của chúng trong không gian ba chiều. Điều này đặc biệt có liên quan khi M tín hiệu giảm mức này tương ứng với cấu hình loa cụ thể như trong ví dụ nêu trên. Ví dụ, bộ phận tạo ra tín hiệu giảm mức 318 có thể bắt nguồn từ ma trận thể hiện Pd (tương ứng với ma trận thể hiện được ứng dụng trong bộ kết xuất 122 được thể hiện trên Hình 1) trên cơ sở thông tin về vị trí và sử dụng ma trận này để tạo ra tín hiệu giảm mức theo $D = Pd * [B^T S^T]^T$.

N đối tượng âm thanh 106a và các kênh nền 106b nếu có mặt cũng được nạp vào bộ phận phân tích 328. Thông thường, bộ phận phân tích 328 hoạt động trên các khung lát thời gian/tần số riêng lẻ chứa các tín hiệu âm thanh đầu vào 106a-b. Đối với mục đích này, N đối tượng âm thanh 106a và các kênh nền 106b có thể được cấp qua dàn lọc 338, ví dụ, dàn QMF, mà thực hiện sự biến đổi tần số theo thời gian của các tín hiệu âm thanh đầu vào 106a-b. Cụ thể là, dàn lọc 338 được kết hợp với các dải phụ tần số. Độ phân giải tần số của khung lát thời gian/tần số tương ứng với một hoặc nhiều dải phụ tần số trong số các dải phụ tần số này. Độ phân giải tần số của các khung lát thời gian/tần số có thể là không đều, nghĩa là nó có thể biến thiên theo tần số. Ví dụ, độ phân giải tần số thấp hơn có thể được sử dụng cho các tần số cao, mà có nghĩa rằng khung lát thời gian/tần số trong vùng tần số cao có thể tương ứng với vài dải phụ tần số như được xác định bởi dàn lọc 338.

Ở bước E06, bộ phận phân tích 328 tạo ra ma trận tái cấu trúc, dưới đây được ký hiệu bằng $R1$. Ma trận tái cấu trúc được tạo ra được cấu thành bằng các phần tử ma trận. Ma trận tái cấu trúc $R1$ là ma trận cho phép tái cấu trúc (tính gần đúng) của các đối tượng âm thanh N 106a và cũng có thể là các kênh nền 106b từ M tín hiệu giảm mức 112 trong bộ giải mã.

Bộ phận phân tích 328 có thể có các cách khác nhau để tạo ra ma trận tái cấu trúc. Ví dụ, phương pháp dự báo sai số bình phương trung bình nhỏ nhất (Minimum Mean Squared Error - MMSE) có thể được sử dụng, phương pháp này lấy cả N đối tượng âm thanh/các kênh nền 106a-b dưới dạng đầu vào cũng như M tín hiệu giảm mức 112 dưới dạng đầu vào. Phương pháp này có thể được mô tả như phương pháp nhằm tìm ra ma trận tái cấu trúc mà giảm đến mức thấp nhất sai số bình phương trung bình của các đối tượng âm thanh/các kênh nền đã được tái cấu trúc. Cụ thể là, phương

pháp tái cấu trúc N đối tượng âm thanh/các kênh nền bằng cách sử dụng ma trận tái cấu trúc ứng viên và so sánh chúng với các đối tượng âm thanh/các kênh nền 106a-b được nạp về mặt sai số bình phương trung bình. Ma trận tái cấu trúc ứng viên mà giảm đến mức thấp nhất sai số bình phương trung bình được chọn làm ma trận tái cấu trúc và các phần tử ma trận 114 của nó là đầu ra của bộ phận phân tích 328.

Phương pháp MMSE cần đến các ước tính tương quan và các ma trận hiệp biến trong số N đối tượng âm thanh/các kênh nền 106a-b và M tín hiệu giảm mức 112. Theo phương pháp nêu trên, các mối tương quan và hiệp biến này được đo trên cơ sở N đối tượng âm thanh/các kênh nền 106a-b và M tín hiệu giảm mức 112. Theo cách khác, phương pháp dựa trên mô hình, bộ phận phân tích 328 lấy dữ liệu vị trí 104 dưới dạng đầu vào thay cho M tín hiệu giảm mức 112. Bằng cách lấy các giả thuyết nhất định, ví dụ, giả sử rằng N đối tượng âm thanh không có mối tương quan lẫn nhau, và bằng cách sử dụng giả thuyết này kết hợp với các nguyên tắc giảm mức được ứng dụng trong bộ phận tạo ra tín hiệu giảm mức 318, bộ phận phân tích 328 có thể tính toán các mối tương quan và hiệp biến theo yêu cầu cần thiết để thực hiện phương pháp MMSE được mô tả trên đây.

Sau đó, các phần tử trong ma trận tái cấu trúc 114 và M tín hiệu giảm mức 112 được nạp vào bộ phận tạo ra dòng bit 110. Ở bước E08, bộ phận tạo ra dòng bit 110 lượng tử hóa và mã hóa M tín hiệu giảm mức 112 và ít nhất một số phần tử ma trận 114 trong ma trận tái cấu trúc và sắp xếp chúng trong dòng bit 116. Cụ thể là, bộ phận tạo ra dòng bit 110 có thể sắp xếp M tín hiệu giảm mức 112 trong trường thứ nhất của dòng bit 116 bằng cách sử dụng định dạng thứ nhất. Hơn nữa, bộ phận tạo ra dòng bit 110 có thể sắp xếp các phần tử ma trận 114 trong trường thứ hai của dòng bit 116 bằng cách sử dụng định dạng thứ hai. Như được mô tả trên đây có dựa vào Hình 2, điều này cho phép bộ giải mã kế thừa chỉ hỗ trợ định dạng thứ nhất để giải mã và phát lại M tín hiệu giảm mức 112 và bỏ qua các phần tử ma trận 114 trong trường thứ hai.

Hình 5 minh họa bộ mã hóa 108 theo phương án khác. So với bộ mã hóa được thể hiện trên Hình 3, bộ mã hóa 508 được thể hiện trên Hình 5 còn cho phép gộp một hoặc nhiều tín hiệu bổ trợ trong dòng bit 116.

Đối với mục đích này, bộ mã hóa 508 bao gồm bộ phận tạo ra các tín hiệu bổ trợ 548. Bộ phận tạo ra các tín hiệu bổ trợ 548 tiếp nhận các đối tượng âm thanh/các kênh nền 106a-b và dựa trên đó một hoặc nhiều tín hiệu bổ trợ 512 được tạo ra. Ví dụ,

bộ phận tạo ra các tín hiệu hỗ trợ 548 có thể tạo ra các tín hiệu hỗ trợ 512 dưới dạng tổ hợp của các đối tượng âm thanh/các kênh nền 106a-b. Việc biểu thị các tín hiệu hỗ trợ bằng vector $C = [C1 \ C2 \ ... \ CL]^T$, các tín hiệu hỗ trợ có thể được tạo ra dưới dạng, $C = Q * [B^T S^T]^T$, trong đó Q là ma trận mà có thể là biến thời gian và tần số. Việc thể hiện này bao gồm cả trường hợp mà các tín hiệu hỗ trợ như một hoặc nhiều các đối tượng âm thanh và trường hợp mà các tín hiệu hỗ trợ là các tổ hợp tuyến tính của các đối tượng âm thanh. Ví dụ, tín hiệu hỗ trợ có thể thể hiện đối tượng đặc biệt quan trọng, như đoạn hội thoại.

Vai trò của các tín hiệu hỗ trợ 512 là để cải thiện việc tái cấu trúc các đối tượng âm thanh/các kênh nền 106a-b trong bộ giải mã. Chính xác hơn, ở phía bộ giải mã, các đối tượng âm thanh/các kênh nền 106a-b có thể được tái cấu trúc trên cơ sở M tín hiệu giảm mức 112 cũng như L tín hiệu hỗ trợ 512. Do đó, ma trận tái cấu trúc sẽ có các phần tử ma trận 114 mà cho phép tái cấu trúc các đối tượng âm thanh/các kênh nền từ M tín hiệu giảm mức 112 cũng như L tín hiệu hỗ trợ.

Do đó, L tín hiệu hỗ trợ 512 có thể được nạp vào bộ phận phân tích 328 sao cho chúng sẽ được tính đến khi tạo ra ma trận tái cấu trúc. Bộ phận phân tích 328 cũng có thể gửi tín hiệu điều khiển đến bộ phận tạo ra các tín hiệu hỗ trợ 548. Ví dụ, bộ phận phân tích 328 có thể điều khiển các đối tượng âm thanh/các kênh nền mà được gộp trong các tín hiệu hỗ trợ mà các đối tượng này được gộp vào. Cụ thể là, bộ phận phân tích 328 có thể điều khiển sự lựa chọn ma trận Q . Ví dụ, việc điều khiển có thể dựa trên cơ sở phương pháp MMSE được mô tả trên đây sao cho các tín hiệu hỗ trợ được chọn sao cho các đối tượng âm thanh/các kênh nền đã được tái cấu trúc là gần nhất có thể với các đối tượng âm thanh/các kênh nền 106a-b.

Hoạt động ở phía bộ giải mã của hệ thống mã hóa/giải mã âm thanh 100 sẽ được mô tả một cách chi tiết hơn có dựa vào Hình 6 và sơ đồ được thể hiện trên Hình 7.

Hình 6 minh họa bộ phận giải mã dòng bit 118 và bộ giải mã 120 được thể hiện trên Hình 1 một cách chi tiết hơn. Bộ giải mã 120 bao gồm bộ phận tạo ra ma trận tái cấu trúc 622 và bộ phận tái cấu trúc 624.

Ở bước D02, bộ phận giải mã dòng bit 118 tiếp nhận dòng bit 116. Bộ phận giải mã dòng bit 118 giải mã và giải lượng tử hóa thông tin trong dòng bit 116 để trích

M tín hiệu giảm mức 112 và ít nhất một số phần tử ma trận 114 trong ma trận tái cấu trúc.

Bộ phận tạo ra ma trận tái cấu trúc 622 tiếp nhận các phần tử ma trận 114 và xử lý để tạo ra ma trận tái cấu trúc 614 ở bước D04. Bộ phận tạo ra ma trận tái cấu trúc 622 tạo ra ma trận tái cấu trúc 614 bằng cách sắp xếp các phần tử ma trận 114 ở các vị trí thích hợp trong ma trận. Nếu không phải toàn bộ các phần tử ma trận trong ma trận tái cấu trúc được tiếp nhận, bộ phận tạo ra ma trận tái cấu trúc 622 có thể, ví dụ, chèn các số không thay cho các phần tử thiếu.

Sau đó, ma trận tái cấu trúc 614 và M tín hiệu giảm mức được nạp vào bộ phận tái cấu trúc 624. Tiếp đó, ở bước D06, bộ phận tái cấu trúc 624 tái cấu trúc N đối tượng âm thanh và các kênh nền, nếu được áp dụng. Nói cách khác, bộ phận tái cấu trúc 624 tạo ra đối tượng gần đúng 106' trong số N đối tượng âm thanh/các kênh nền 106a-b.

Để làm ví dụ, M tín hiệu giảm mức có thể tương ứng với cấu hình loa cụ thể, như cấu hình loa $[Lf\ Rf\ Cf\ Ls\ Rs\ LFE]$ theo cấu hình loa 5.1. Nếu vậy, bộ phận tái cấu trúc 624 có thể dựa trên việc tái cấu trúc các đối tượng 106' chỉ trên các tín hiệu giảm mức tương ứng với các kênh toàn dải của cấu hình loa. Như nêu trên, tín hiệu hạn chế dải (tín hiệu LFE tần số thấp) có thể được gửi cơ bản không biến đổi đến bộ kết xuất.

Bộ phận tái cấu trúc 624 thường hoạt động trong miền tần số. Chính xác hơn, bộ phận tái cấu trúc 624 hoạt động trên các khung lát thời gian/tần số riêng lẻ của các tín hiệu đầu vào. Do đó, M tín hiệu giảm mức 112 thường phải qua biến đổi tần số theo thời gian 623 trước khi được nạp vào bộ phận tái cấu trúc 624. Biến đổi tần số theo thời gian 623 thường giống hoặc tương tự như biến đổi 338 được áp dụng ở phía bộ mã hóa. Ví dụ, biến đổi tần số theo thời gian 623 có thể là biến đổi QMF.

Để tái cấu trúc các đối tượng âm thanh/các kênh nền 106', bộ phận tái cấu trúc 624 ứng dụng cách hoạt động của ma trận. Cụ thể hơn, bằng cách sử dụng sự chỉ dẫn được nạp từ trước, bộ phận tái cấu trúc 624 có thể tạo ra đối tượng gần đúng A' trong số đối tượng âm thanh /các kênh nền dưới dạng $A' = R1 * D$. Ma trận tái cấu trúc $R1$ có thể biến thiên như hàm của thời gian và tần số. Do đó, ma trận tái cấu trúc có thể biến thiên giữa các khung lát thời gian/tần số khác nhau đã được xử lý bởi bộ phận tái cấu trúc 624.

Thông thường, các đối tượng âm thanh/các kênh nền 106' đã được tái cấu trúc được biến đổi ngược về miền thời gian 625 trước khi được xuất ra khỏi bộ giải mã 120.

Hình 8 minh họa tình huống khi dòng bit 116 còn bao gồm các tín hiệu bổ trợ. So với hệ thống theo phương án thực hiện được thể hiện trên Hình 7, trong hệ thống theo phương án thực hiện này bộ phận giải mã dòng bit 118 còn giải mã một hoặc nhiều tín hiệu bổ trợ 512 từ dòng bit 116. Các tín hiệu bổ trợ 512 được nạp vào bộ phận tái cấu trúc 624 nơi chúng được gộp vào trong việc tái cấu trúc các đối tượng âm thanh/các kênh nền. Cụ thể hơn, bộ phận tái cấu trúc 624 tạo ra các đối tượng âm thanh/các kênh nền bằng cách áp dụng hoạt động của ma trận $A' = R1 * [D^T C^T]^T$.

Hình 9 minh họa các biến đổi thời gian/tần số khác nhau được dùng ở phía bộ giải mã trong hệ thống mã hóa/giải mã âm thanh 100 được thể hiện trên Hình 1. Bộ phận giải mã dòng bit 118 tiếp nhận dòng bit 116. Bộ phận giải mã và giải lượng tử hóa 918 giải mã và giải lượng tử hóa dòng bit 116 để trích thông tin về vị trí 104, M tín hiệu giảm mức 112, và các phần tử ma trận 114 của ma trận tái cấu trúc.

Trong giai đoạn này, M tín hiệu giảm mức 112 thường được thể hiện trong miền tần số thứ nhất, tương ứng với bộ thứ nhất của các dàn lọc thời gian/tần số dưới đây được ký hiệu bằng T/F_C và F/T_C nhằm lần lượt biến đổi từ miền thời gian đến miền tần số thứ nhất và từ miền tần số thứ nhất đến miền thời gian. Thông thường, các dàn lọc tương ứng với miền tần số thứ nhất có thể thực hiện việc biến đổi cửa sổ chồng, như biến đổi MDCT và MDCT nghịch. Bộ phận giải mã dòng bit 118 có thể có bộ biến đổi 901 mà biến đổi M tín hiệu giảm mức 112 thành miền thời gian nhờ sử dụng dàn lọc F/T_C .

Thông thường, bộ giải mã 120, và cụ thể là bộ phận tái cấu trúc 624, xử lý các tín hiệu đối với miền tần số thứ hai. Miền tần số thứ hai tương ứng với bộ thứ hai của các dàn lọc thời gian/tần số dưới đây được ký hiệu bằng T/F_U và F/T_U nhằm lần lượt biến đổi từ miền thời gian đến miền tần số thứ hai và từ miền tần số thứ hai đến miền thời gian. Do đó, bộ giải mã 120 có thể có bộ biến đổi 903 mà biến đổi M tín hiệu giảm mức 112, mà được thể hiện trong miền thời gian, đến miền tần số thứ hai nhờ sử dụng dàn lọc T/F_U . Khi bộ phận tái cấu trúc 624 tái cấu trúc các đối tượng 106' trên cơ sở M tín hiệu giảm mức bằng cách tiến hành việc xử lý trong miền tần số thứ hai,

bộ biến đổi 905 có thể biến đổi các đối tượng 106' đã được tái cấu trúc ngược về miền thời gian nhờ sử dụng dàn lọc F/T_U .

Thông thường, bộ kết xuất 122 xử lý các tín hiệu đối với miền tần số thứ ba. miền tần số thứ ba tương ứng với bộ thứ ba của các dàn lọc thời gian/tần số dưới đây được ký hiệu bằng T/F_R và F/T_R nhằm lần lượt biến đổi từ miền thời gian đến miền tần số thứ ba và từ miền tần số thứ ba đến miền thời gian. Do đó, bộ kết xuất 122 có thể có bộ phận biến đổi 907 mà biến đổi các đối tượng âm thanh 106' đã được tái cấu trúc từ miền thời gian đến miền tần số thứ ba nhờ sử dụng dàn lọc T/F_R . Một khi bộ kết xuất 122, bằng bộ phận kết xuất 922, đã kết xuất các kênh đầu ra 124, các kênh đầu ra này có thể được biến đổi đến miền thời gian nhờ bộ biến đổi 909 bằng cách sử dụng dàn lọc F/T_R .

Như đã rõ từ phần mô tả nêu trên, phía bộ giải mã của hệ thống mã hóa/giải mã âm thanh có các bước biến đổi thời gian/tần số. Tuy nhiên, nếu miền tần số thứ nhất, thứ hai và thứ ba được chọn theo các cách nhất định, một số bước biến đổi thời gian/tần số trở nên thừa.

Ví dụ, một số miền trong số miền tần số thứ nhất, thứ hai và thứ ba có thể được chọn để giống hoặc có thể được tiến hành cùng nhau để đi một cách trực tiếp từ miền tần số này đến miền tần số khác mà không đi qua toàn bộ quá trình đến miền thời gian ở giữa. Ví dụ về việc tiến hành cùng nhau là trường hợp chỉ có sự khác biệt giữa miền tần số thứ hai và miền tần số thứ ba là ở chỗ việc bộ phận biến đổi 907 trong bộ kết xuất 122 sử dụng dàn lọc Nyquist dành cho độ phân giải tần số cao ở các tần số thấp để bổ sung cho dàn lọc QMF vốn là thông thường đối với cả bộ biến đổi 905 lẫn bộ biến đổi 907. Trong trường hợp này, các bộ biến đổi 905 và 907 có thể được tiến hành cùng nhau dưới dạng dàn lọc Nyquist, do đó giảm độ phức tạp tính toán.

Theo ví dụ khác, miền tần số thứ hai và miền tần số thứ ba là giống nhau. Ví dụ, miền tần số thứ hai và miền tần số thứ ba có thể đều là miền tần số QMF. Trong trường hợp này, bộ phận biến đổi 905 và 907 là thừa và có thể được loại bỏ, do đó giảm độ phức tạp tính toán.

Trong ví dụ khác, miền tần số thứ nhất và miền tần số thứ hai có thể giống nhau. Ví dụ miền tần số thứ nhất và miền tần số thứ hai có thể đều là miền MDCT.

Trong trường hợp này, bộ phận biến đổi thứ nhất 901 và bộ phận biến đổi thứ hai 903 có thể được loại bỏ, do đó giảm độ phức tạp tính toán.

Các biến thể tương đương, mở rộng, thay thế và các biến thể khác

Sáng chế theo các phương án thực hiện tiếp theo sẽ trở nên rõ ràng đối với chuyên gia trong lĩnh vực kỹ thuật này sau khi đọc phần mô tả nêu trên. Ngay cả trong bản mô tả và trên các hình vẽ thể hiện hệ thống theo các phương án thực hiện và các ví dụ, sáng chế không chỉ giới hạn ở các ví dụ cụ thể. Nhiều thay đổi và biến thể có thể được thực hiện mà không trệch khỏi phạm vi bảo hộ của sáng chế mà được xác định bởi các điểm yêu cầu bảo hộ kèm theo. Số chỉ dẫn bất kỳ nêu trong các điểm yêu cầu bảo hộ không được hiểu là giới hạn phạm vi bảo hộ của các điểm yêu cầu bảo hộ này.

Ngoài ra, các thay đổi đối với các phương án thực hiện đã được bộc lộ có thể được hiểu và được thực hiện bởi chuyên gia trong lĩnh vực kỹ thuật này khi thực hiện sáng chế, từ việc nghiên cứu các hình vẽ, bản mô tả, và các điểm yêu cầu bảo hộ kèm theo. Trong các điểm yêu cầu bảo hộ, từ "bao gồm" không được hiểu là loại trừ các chi tiết hoặc các bước khác, và từ chỉ số ít "một" hoặc "một cái" không loại trừ số nhiều của từ này. Trên thực tế, các số đo nhất định được chỉ ra trong các điểm yêu cầu bảo hộ phụ thuộc khác nhau không chỉ rằng tổ hợp của các số đo này không thể được sử dụng để có lợi.

Các hệ thống và phương pháp đã được bộc lộ trên đây có thể được thực hiện như phần mềm, phần sụn, phần cứng hoặc tổ hợp của chúng. Khi thực hiện phần cứng, sự phân chia nhiệm vụ giữa các cụm chức năng nêu trong phần mô tả trên đây không nhất thiết tương ứng với sự phân chia thành các cụm thực thể; trái lại, một bộ phận thực thể có thể có đa chức năng, và một nhiệm vụ có thể được thực hiện nhờ một số bộ phận thực thể kết hợp. Các bộ phận nhất định hoặc tất cả các bộ phận có thể được thực hiện như phần mềm được thực hiện bởi bộ xử lý hoặc bộ vi xử lý tín hiệu kỹ thuật số, hoặc được thực hiện dưới dạng phần cứng hoặc mạch tích hợp ứng dụng cụ thể. Phần mềm này có thể được phân phối trên vật ghi đọc được bằng máy tính, mà có thể bao gồm phương tiện lưu trên máy tính (hoặc phương tiện không chuyển tiếp) và phương tiện truyền thông (hoặc phương tiện chuyển tiếp). Các chuyên gia trong lĩnh vực kỹ thuật này đã biết rằng, thuật ngữ phương tiện lưu trên máy tính bao gồm cả phương tiện khả biến lẫn phương tiện không khả biến, phương tiện loại bỏ được và

không loại bỏ được được thực hiện theo phương pháp hoặc công nghệ bất kỳ hoặc để lưu thông tin như các lệnh đọc được bằng máy tính, các cấu trúc dữ liệu, các mô đun chương trình hoặc các dữ liệu khác. Phương tiện lưu trên máy tính bao gồm, nhưng không chỉ giới hạn ở, RAM, ROM, EEPROM, bộ nhớ tia chớp hoặc các công nghệ nhớ khác, CD-ROM, đĩa kỹ thuật số đa chức năng (Digital Versatile Disks-DVD) hoặc các phương tiện lưu đĩa quang khác, hộp băng từ, băng từ, phương tiện lưu đĩa từ hoặc các thiết bị lưu từ khác, hoặc môi trường khác bất kỳ mà có thể được sử dụng để lưu thông tin mong muốn và có thể truy cập được bằng máy tính. Hơn nữa, chuyên gia trong lĩnh vực kỹ thuật này đã biết phương tiện truyền thông thông thường thực hiện các lệnh đọc được bằng máy tính, các cấu trúc dữ liệu, các mô đun chương trình hoặc các dữ liệu khác trong tín hiệu dữ liệu đã được biến điệu như sóng mang hoặc các cơ cấu vận chuyển khác và bao gồm phương tiện phân phối thông tin bất kỳ.

YÊU CẦU BẢO HỘ

1. Phương pháp mã hóa khung lát thời gian/tần số của cảnh âm thanh mà ít nhất có N đối tượng âm thanh, phương pháp này bao gồm các bước:

tiếp nhận N đối tượng âm thanh;

tạo ra M tín hiệu giảm mức trên cơ sở ít nhất N đối tượng âm thanh;

tạo ra ma trận tái cấu trúc với các phần tử ma trận để tái cấu trúc ít nhất N đối tượng âm thanh từ M tín hiệu giảm mức, trong đó các đối tượng gần đúng trong số N đối tượng âm thanh thu được là các tổ hợp tuyến tính của ít nhất M tín hiệu giảm mức với các phần tử ma trận trong ma trận tái cấu trúc làm các hệ số trong các tổ hợp tuyến tính; và

tạo ra dòng bit bao gồm M tín hiệu giảm mức và ít nhất một số phần tử ma trận trong ma trận tái cấu trúc.

2. Phương pháp theo điểm 1, trong đó M tín hiệu giảm mức được sắp xếp trong trường thứ nhất của dòng bit bằng cách sử dụng định dạng thứ nhất, và các phần tử ma trận được sắp xếp trong trường thứ hai của dòng bit bằng cách sử dụng định dạng thứ hai, nhờ đó cho phép bộ giải mã chỉ hỗ trợ định dạng thứ nhất để giải mã và phát lại M tín hiệu giảm mức trong trường thứ nhất và bỏ qua các phần tử ma trận trong trường thứ hai.

3. Phương pháp theo điểm 2, trong đó phương pháp này còn bao gồm bước tiếp nhận dữ liệu vị trí tương ứng với mỗi đối tượng trong số N đối tượng âm thanh, trong đó M tín hiệu giảm mức được tạo ra trên cơ sở dữ liệu vị trí.

4. Phương pháp theo điểm 1, trong đó cảnh âm thanh còn có các kênh nền, trong đó M tín hiệu giảm mức được tạo ra trên cơ sở ít nhất N đối tượng âm thanh và các kênh nền.

5. Phương pháp theo điểm 4, trong đó ma trận tái cấu trúc có các phần tử ma trận để tái cấu trúc các kênh nền từ M tín hiệu giảm mức, trong đó các đối tượng gần đúng trong số N đối tượng âm thanh và các kênh nền thu được là các tổ hợp tuyến tính của

ít nhất M tín hiệu giảm mức với các phần tử ma trận trong ma trận tái cấu trúc làm các hệ số trong các tổ hợp tuyến tính.

6. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm bước:

tạo ra L tín hiệu bổ trợ từ N đối tượng âm thanh;

gộp các phần tử ma trận trong ma trận tái cấu trúc để tái cấu trúc ít nhất N đối tượng âm thanh từ M tín hiệu giảm mức và L tín hiệu bổ trợ, trong đó các đối tượng gần đúng trong số N đối tượng âm thanh thu được là các tổ hợp tuyến tính của M tín hiệu giảm mức và L tín hiệu bổ trợ với các phần tử ma trận trong ma trận tái cấu trúc làm các hệ số trong các tổ hợp tuyến tính; và

gộp L tín hiệu bổ trợ trong dòng bit.

7. Phương pháp theo điểm 6, trong đó M tín hiệu giảm mức mở rộng siêu mặt phẳng, và trong đó ít nhất một trong số các tín hiệu bổ trợ không nằm trong siêu mặt phẳng được mở rộng bởi M tín hiệu giảm mức.

8. Phương pháp theo điểm 7, trong đó ít nhất một trong số các tín hiệu bổ trợ là vuông góc với siêu mặt phẳng được mở rộng bởi M tín hiệu giảm mức.

9. Vật ghi bất biến đọc được bằng máy tính bao gồm các lệnh mã hóa bằng máy tính được làm thích ứng để thực hiện phương pháp theo điểm điểm 1 khi thực hiện trên thiết bị có khả năng xử lý.

10. Bộ mã hóa để mã hóa khung lát thời gian/tần số của cảnh âm thanh mà ít nhất có N đối tượng âm thanh, bao gồm ít nhất một trong số phần cứng và bộ xử lý kết hợp với bộ nhớ để khiến cho:

bộ tiếp nhận tiếp nhận N đối tượng âm thanh;

bộ tạo tín hiệu giảm mức để tiếp nhận N đối tượng âm thanh từ bộ tiếp nhận và tạo ra M tín hiệu giảm mức trên cơ sở ít nhất N đối tượng âm thanh;

bộ phân tích để tạo ra ma trận tái cấu trúc với các phần tử ma trận để tái cấu trúc ít nhất N đối tượng âm thanh từ M tín hiệu giảm mức, trong đó các đối tượng gần đúng trong số N đối tượng âm thanh thu được là các tổ hợp tuyến tính của ít nhất M

tín hiệu giảm mức với các phần tử ma trận trong ma trận tái cấu trúc làm các hệ số trong các tổ hợp tuyến tính; và

bộ tạo dòng bit để tiếp nhận M tín hiệu giảm mức từ bộ tạo tín hiệu giảm mức và ma trận tái cấu trúc từ bộ phân tích và tạo ra dòng bit bao gồm M tín hiệu giảm mức và ít nhất một số phần tử ma trận trong ma trận tái cấu trúc.

11. Phương pháp giải mã khung lát thời gian-tần số của cảnh âm thanh mà ít nhất có N đối tượng âm thanh, phương pháp này bao gồm các bước:

tiếp nhận dòng bit có M tín hiệu giảm mức và ít nhất một số phần tử ma trận trong ma trận tái cấu trúc;

tạo ra ma trận tái cấu trúc bằng cách sử dụng các phần tử ma trận; và

tái cấu trúc N đối tượng âm thanh từ M tín hiệu giảm mức bằng cách sử dụng ma trận tái cấu trúc, trong đó các đối tượng gần đúng trong số N đối tượng âm thanh được thu nhận dưới dạng các tổ hợp tuyến tính của ít nhất M tín hiệu giảm mức với các phần tử ma trận trong ma trận tái cấu trúc làm các hệ số trong các tổ hợp tuyến tính.

12. Phương pháp theo điểm 11, trong đó M tín hiệu giảm mức được sắp xếp trong trường thứ nhất của dòng bit bằng cách sử dụng định dạng thứ nhất, và các phần tử ma trận được sắp xếp trong trường thứ hai của dòng bit bằng cách sử dụng định dạng thứ hai, nhờ đó cho phép bộ giải mã chỉ hỗ trợ định dạng thứ nhất để giải mã và phát lại M tín hiệu giảm mức trong trường thứ nhất và bỏ qua các phần tử ma trận trong trường thứ hai.

13. Phương pháp theo điểm 1, trong đó cảnh âm thanh còn có các kênh nền, phương pháp này còn bao gồm bước tái cấu trúc các kênh nền từ M tín hiệu giảm mức bằng cách sử dụng ma trận tái cấu trúc, trong đó các đối tượng gần đúng trong số N đối tượng âm thanh và các kênh nền thu được là các tổ hợp tuyến tính của ít nhất M tín hiệu giảm mức với các phần tử ma trận trong ma trận tái cấu trúc làm các hệ số trong các tổ hợp tuyến tính.

14. Phương pháp theo điểm 11, trong đó phương pháp này còn bao gồm các bước:

tiếp nhận L tín hiệu bổ trợ được tạo ra từ N đối tượng âm thanh;

tái cấu trúc N đối tượng âm thanh từ M tín hiệu giảm mức và L tín hiệu bổ trợ bằng cách sử dụng ma trận tái cấu trúc, trong đó các đối tượng gần đúng trong số N đối tượng âm thanh được thu nhận dưới dạng các tổ hợp tuyến tính của M tín hiệu giảm mức và L tín hiệu bổ trợ với các phân tử ma trận trong ma trận tái cấu trúc làm các hệ số trong các tổ hợp tuyến tính.

15. Phương pháp theo điểm 14, trong đó M tín hiệu giảm mức mở rộng siêu mặt phẳng, và trong đó ít nhất một trong số các tín hiệu bổ trợ không nằm trong siêu mặt phẳng được mở rộng bởi M tín hiệu giảm mức.

16. Phương pháp theo điểm 15, trong đó ít nhất một trong số các tín hiệu bổ trợ mà không nằm trong siêu mặt phẳng vuông góc với siêu mặt phẳng được mở rộng bởi M tín hiệu giảm mức.

17. Phương pháp theo điểm 11, trong đó phương pháp này còn bao gồm các bước:

tiếp nhận dữ liệu vị trí tương ứng với N đối tượng âm thanh, và

kết xuất N đối tượng âm thanh bằng cách sử dụng dữ liệu vị trí để tạo ra ít nhất một kênh âm thanh đầu ra.

18. Phương pháp theo điểm 17, trong đó ma trận tái cấu trúc được biểu diễn đối với miền tần số thứ hai tương ứng với dàn lọc thứ hai, và việc kết xuất được thực hiện trong miền tần số thứ ba tương ứng với dàn lọc thứ ba, trong đó dàn lọc thứ hai và dàn lọc thứ ba ít nhất một phần là cùng dàn lọc.

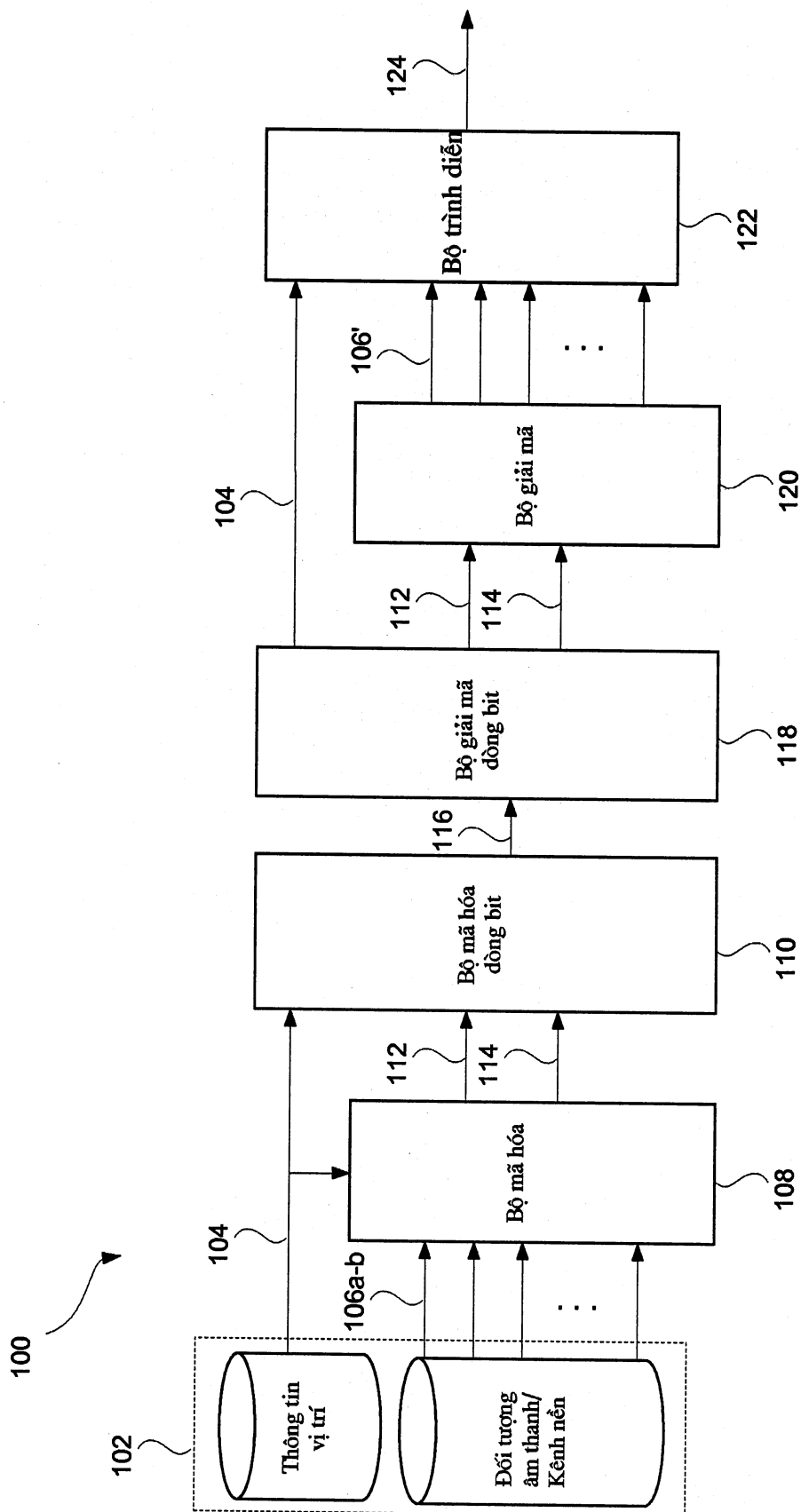
19. Vật ghi bất biến đọc được bằng máy tính bao gồm các lệnh mã hóa bằng máy tính được làm thích ứng để thực hiện phương pháp theo điểm 11 khi thực hiện trên thiết bị có khả năng xử lý.

20. Bộ giải mã để giải mã khung lát thời gian-tần số của cảnh âm thanh mà ít nhất có N đối tượng âm thanh, bao gồm ít nhất một trong số phần cứng và bộ xử lý kết hợp với bộ nhớ được tạo cấu hình để khiến cho:

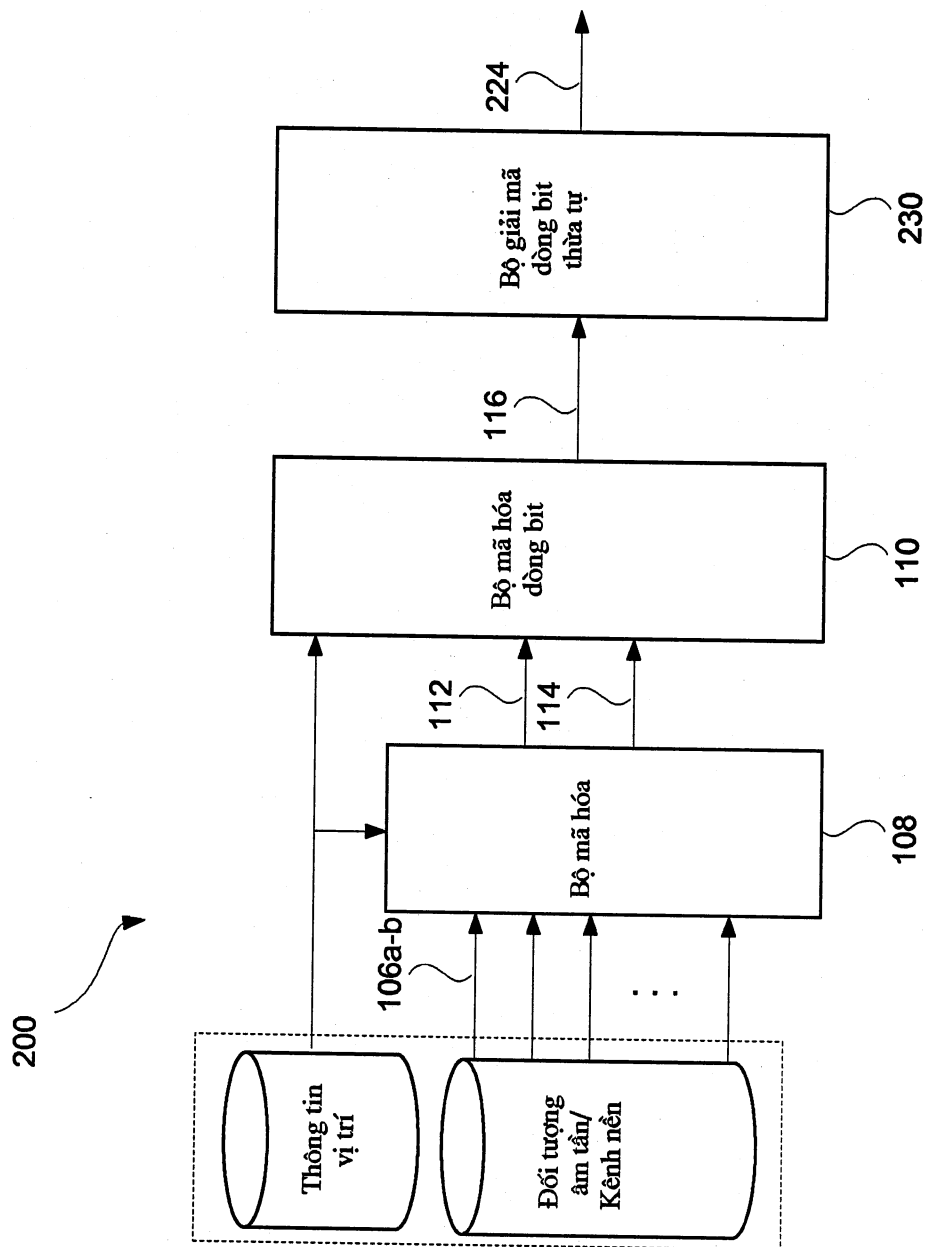
bộ tiếp nhận tiếp nhận dòng bit bao gồm M tín hiệu giảm mức và ít nhất một số phần tử ma trận trong ma trận tái cấu trúc;

bộ tạo ma trận tái cấu trúc để tiếp nhận các phần tử ma trận từ bộ tiếp nhận và dựa trên đó tạo ra ma trận tái cấu trúc; và

bộ tái cấu trúc để tiếp nhận ma trận tái cấu trúc từ bộ tạo ma trận tái cấu trúc và tái cấu trúc N đối tượng âm thanh từ M tín hiệu giảm mức bằng cách sử dụng ma trận tái cấu trúc, trong đó các đối tượng gần đúng trong số N đối tượng âm thanh được thu nhận dưới dạng các tổ hợp tuyến tính của ít nhất M tín hiệu giảm mức với các phần tử ma trận trong ma trận tái cấu trúc làm các hệ số trong các tổ hợp tuyến tính.

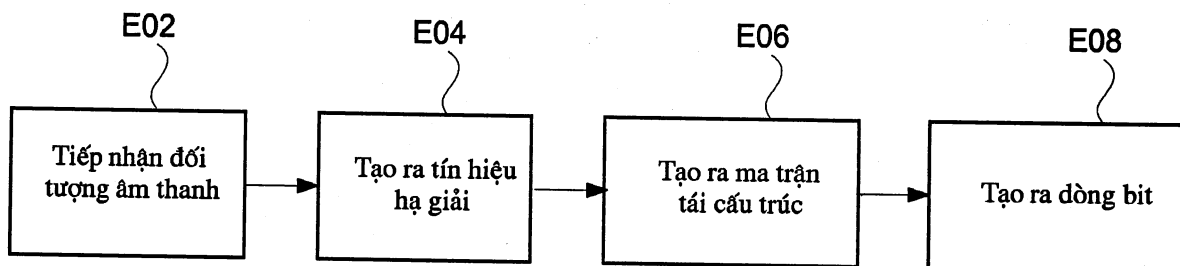
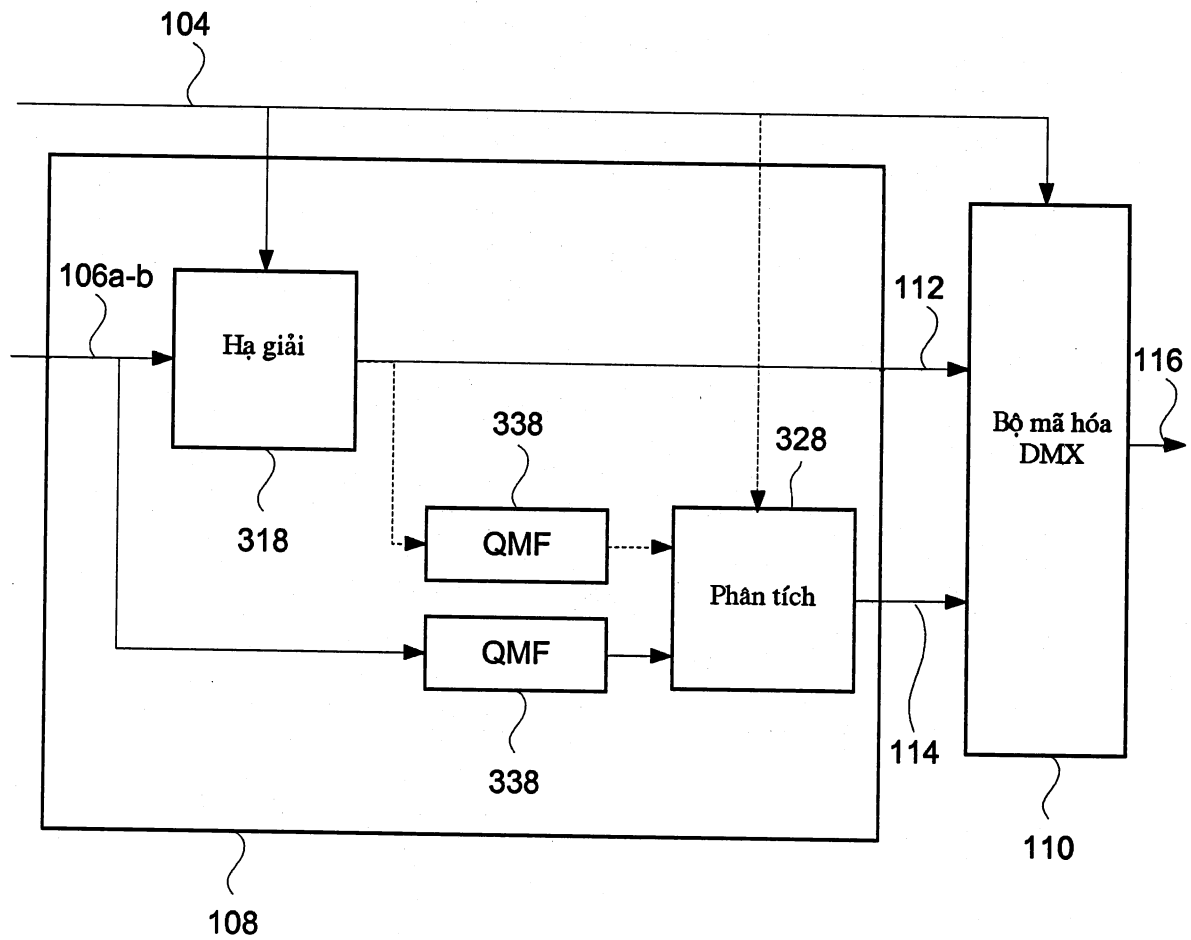


Hình 1



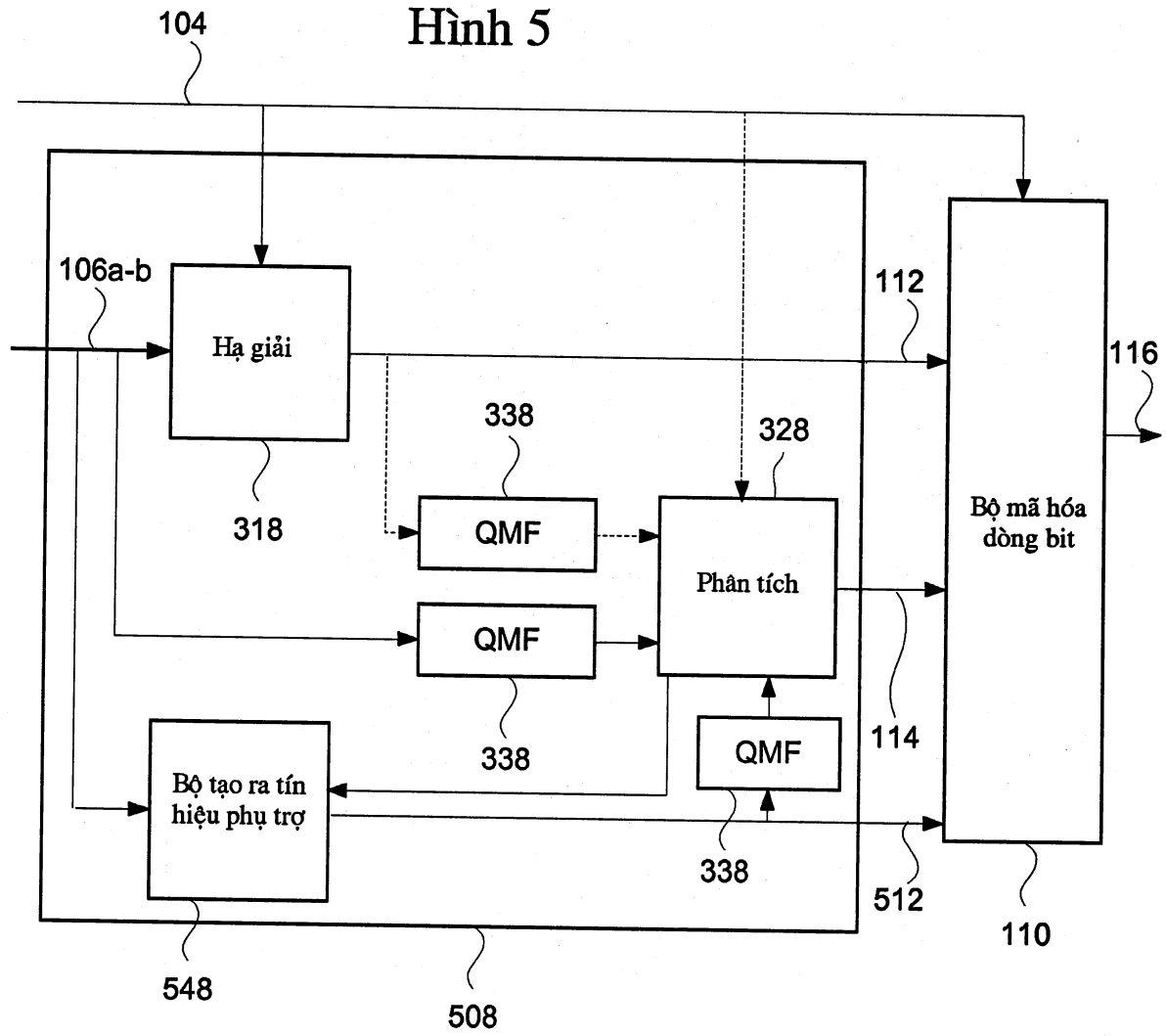
Hình 2

Hình 3

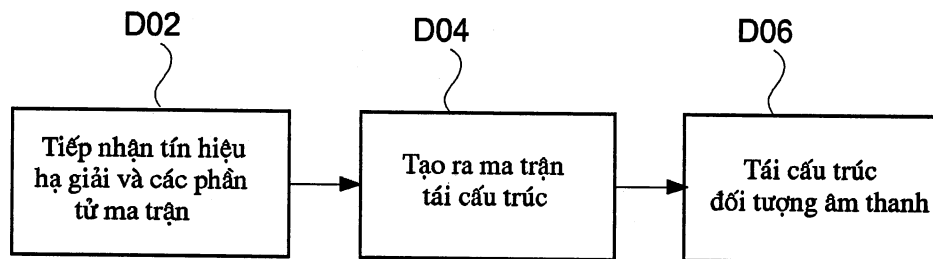
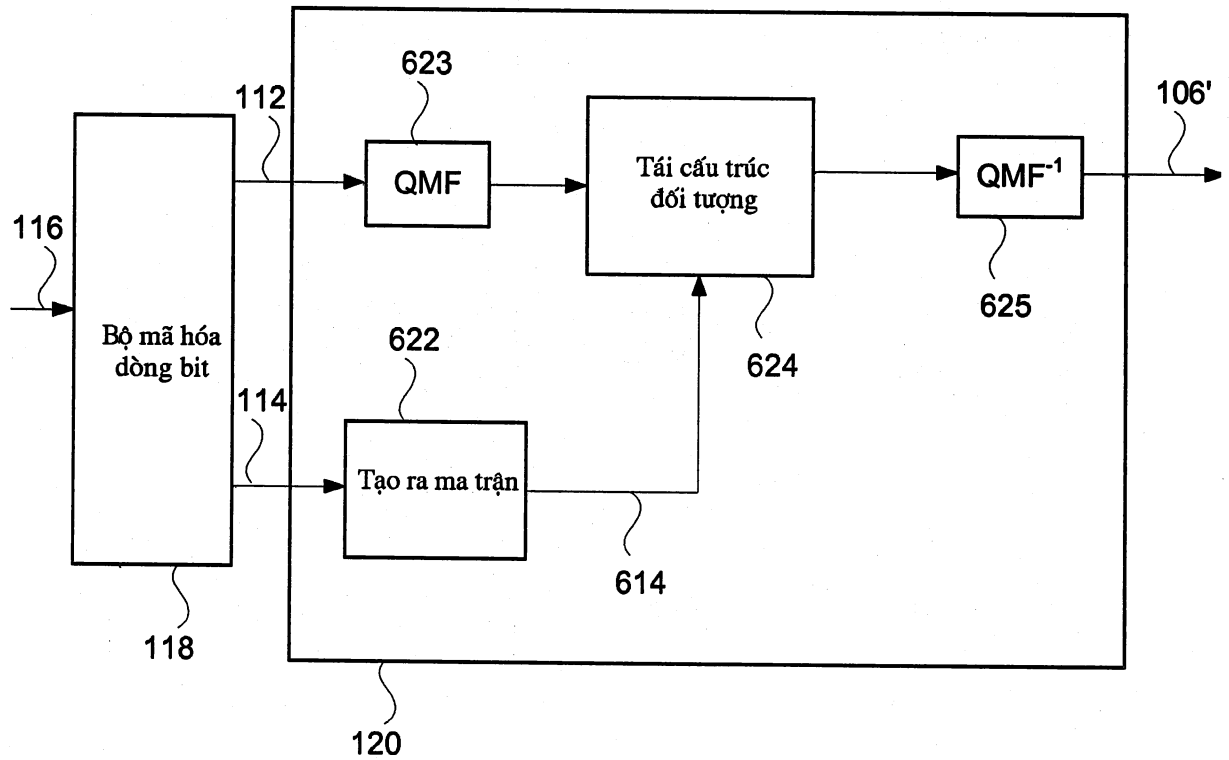


Hình 4

Hình 5

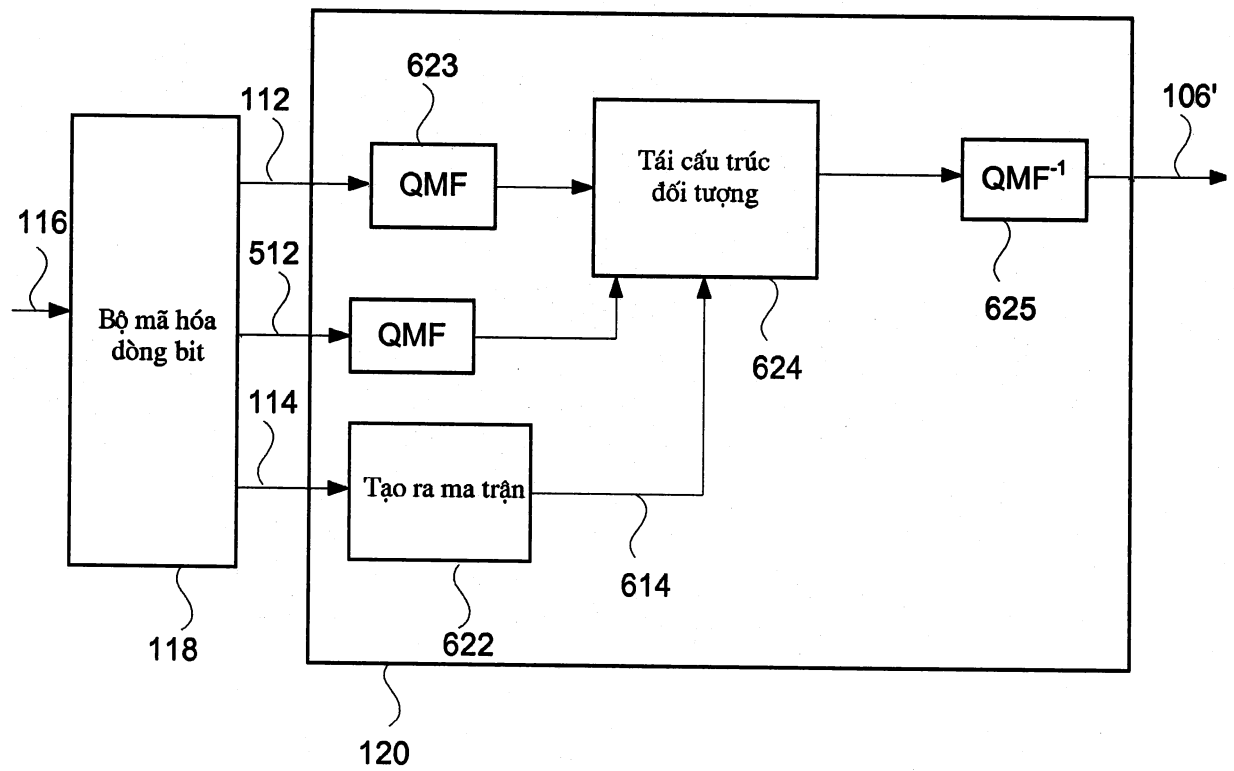


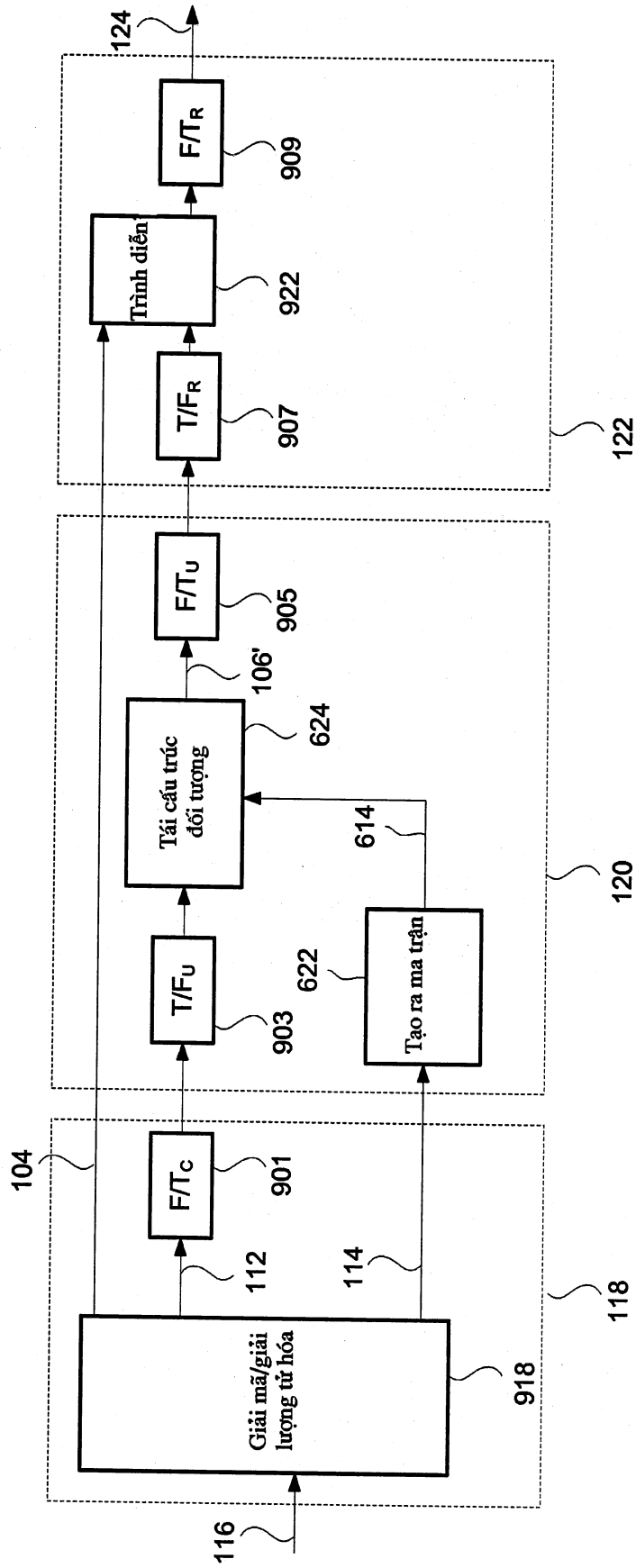
Hình 6



Hình 7

Hình 8





Hình 9