



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẢNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0026672

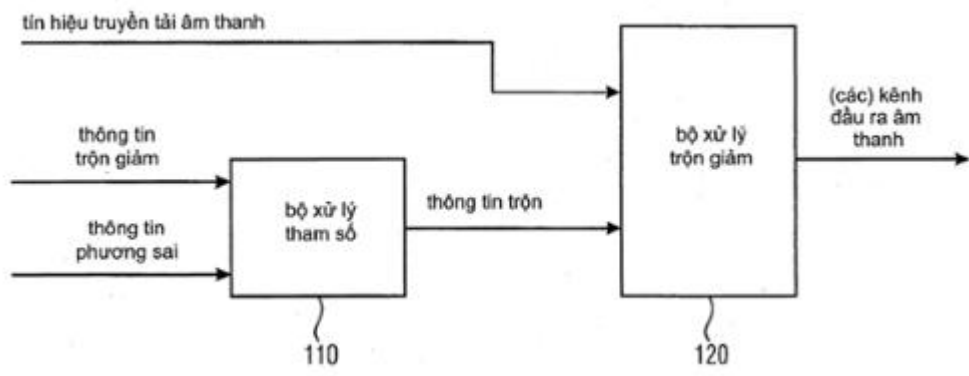
(51)⁷ G10L 19/008; H04S 3/00

(13) B

- (21) 1-2016-00386 (22) 17/07/2014
(86) PCT/EP2014/065427 17/07/2014 (87) WO2015/011024 A1 29/01/2015
(30) EP13177357 22/07/2013 EP; EP13177371 22/07/2013 EP; EP13177378 22/07/2013 EP; EP13189290 18/10/2013 EP
(45) 25/12/2020 393 (43) 25/05/2016 338A
(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)
Hansastrasse 27c, 80686 Muenchen, Germany
(72) HERRE, Juergen (DE); MURTAZA, Adrian (RO); PAULUS, Jouni (FI); DISCH, Sascha (DE); FUCHS, Harald (DE); HELLMUTH, Oliver (DE); RIDDERBUSCH, Falko (DE); TERENTIV, Leon (DE).
(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS HANOI)

(54) THIẾT BỊ VÀ PHƯƠNG PHÁP TẠO RA MỘT HOẶC NHIỀU KÊNH ĐẦU RA ÂM THANH, THIẾT BỊ VÀ PHƯƠNG PHÁP TẠO RA TÍN HIỆU TRUYỀN TẢI ÂM THANH, HỆ THỐNG TẠO RA TÍN HIỆU TRUYỀN TẢI ÂM THANH VÀ MỘT HOẶC NHIỀU KÊNH ĐẦU RA ÂM THANH

(57) Sáng chế đề cập đến thiết bị và phương pháp tạo ra một hoặc nhiều kênh đầu ra âm thanh, thiết bị và phương pháp tạo ra tín hiệu truyền tải âm thanh, hệ thống tạo ra tín hiệu truyền tải và một hoặc nhiều kênh đầu ra âm thanh. Thiết bị bao gồm bộ xử lý tham số (110) để tính toán thông tin trộn và bộ xử lý trộn giảm (120) để tạo ra một hoặc nhiều kênh đầu ra âm thanh. Bộ xử lý trộn giảm (120) được tạo cấu hình để nhận tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh. Một hoặc nhiều tín hiệu kênh âm thanh được trộn trong tín hiệu truyền tải âm thanh, và một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong tín hiệu truyền tải âm thanh, và trong đó số lượng của một hoặc nhiều kênh truyền tải âm thanh nhỏ hơn số lượng của một hoặc nhiều tín hiệu kênh âm thanh cùng với số lượng của một hoặc nhiều tín hiệu đối tượng âm thanh. Bộ xử lý tham số (110) được tạo cấu hình để nhận thông tin trộn giảm biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong một hoặc nhiều kênh truyền tải âm thanh, và trong đó bộ xử lý tham số (110) được tạo cấu hình để nhận thông tin phương sai. Hơn nữa, bộ xử lý tham số (110) được tạo cấu hình để tính toán thông tin trộn phụ thuộc vào thông tin trộn giảm và phụ thuộc vào thông tin phương sai. Bộ xử lý trộn giảm (120) được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu truyền tải âm thanh phụ thuộc vào thông tin trộn. Thông tin phương sai biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu kênh âm thanh và còn biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh. Tuy nhiên, thông tin phương sai không biểu thị thông tin tương quan đối với cặp bất kỳ của một trong số một hoặc nhiều tín hiệu kênh âm thanh và một trong số một hoặc nhiều tín hiệu đối tượng âm thanh.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc mã hóa/giải mã âm thanh, cụ thể là, đề cập đến mã hóa âm thanh trong không gian và mã hóa đối tượng âm thanh trong không gian, và, cụ thể hơn là, đề cập đến thiết bị và phương pháp mã hóa đối tượng âm thanh trong không gian được tăng cường.

Tình trạng kỹ thuật của sáng chế

Các công cụ mã hóa âm thanh trong không gian là công cụ đã biết trong lĩnh vực 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, được nhận biết bằng sự bố trí của chúng trong cài đặt tái tạo, 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 thu được một hoặc nhiều kênh trộn giảm từ các kênh ban đầu và, ngoài ra, thu được dữ liệu tham số liên quan đến các tín hiệu điều khiển trong không gian như các chênh lệch mức liên kênh ở các trị số 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 điều khiển 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. Sự bố trí của các kênh trong cài đặt đầu ra thường được cố định và là, ví dụ, định dạng 5.1, định dạng 7.1, v.v.

Các định dạng âm thanh dựa trên kênh này được sử dụng một cách rộng rãi để lưu trữ hoặc truyền nội dung âm thanh đa kênh, trong đó mỗi kênh liên quan đến loa phóng thanh cụ thể ở vị trí đã cho. Sự tái tạo chính xác của loại định dạng này yêu cầu việc cài đặt loa phóng thanh, trong đó các loa được đặt ở cùng vị trí như loa được sử dụng trong suốt quá trình tạo ra các tín hiệu âm thanh. Trong khi tăng số lượng các loa

phóng thanh cải thiện sự tái tạo các ngữ cảnh âm thanh 3D nhúng đích thực, trở nên khó khăn hơn nhiều để thực hiện yêu cầu đặc biệt trong môi trường trong nhà như phòng khách.

Sự cần thiết của việc có sự cài đặt loa phóng thanh cụ thể có thể được khắc phục bằng phương pháp dựa trên đối tượng, trong đó các tín hiệu loa phóng thanh được kết xuất một cách đặc biệt để cài đặt phát lại.

Ví dụ, các công cụ mã hóa đối tượng âm thanh trong không gian đã được biết trong lĩnh vực này và được chuẩn hóa theo tiêu chuẩn mã hóa đối tượng âm thanh trong không gian (spatial audio object coding - SAOC) MPEG. Trái lại với mã hóa âm thanh trong không gian bắt đầu từ các kênh ban đầu, mã hóa đối tượng âm thanh trong không gian bắt đầu từ các đối tượng âm thanh mà không tự động được dành riêng cho sự cài đặt tái tạo kết xuất nhất định. Thay vào đó, sự bố trí của các đối tượng âm thanh trong ngữ cảnh tái tạo là linh hoạt và có thể được xác định bởi người dùng bằng cách nhập một số thông tin kết xuất vào trong bộ giải mã mã hóa đối tượng âm thanh trong không gian. Ngoài ra hoặc theo cách khác, thông tin kết xuất, tức là, thông tin mà tại đó vị trí trong cài đặt tái tạo đối tượng âm thanh nhất định thường được đặt trên thời gian có thể được truyền như thông tin phụ hoặc siêu dữ liệu bổ sung. Để đạt được sự nén dữ liệu nhất định, số lượng 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 truyền tải bằng cách trộn giảm các đối tượng theo thông tin trộn giảm nhất định. Hơn nữa, bộ mã hóa SAOC tính toán thông tin phụ tham số biểu diễn các tín hiệu điều khiển liên đối tượng như các chênh lệch mức đối tượng (object level difference-OLD), các trị số 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 tham số, tức là, đối với khung nhất định của tín hiệu âm thanh bao gồm, ví dụ, 1024 hoặc 2048 mẫu, 28, 20, 14 hoặc 10, v.v., các băng xử lý được xem xét sao cho, cuối cùng, dữ liệu tham số tồn tại cho mỗi khung và mỗi băng xử lý. Như ví dụ, khi đoạn âm thanh có 20 khung và khi mỗi khung được chia nhỏ thành 28 băng xử lý, số lượng các ô thời gian/tần số tham số là 560.

Trong phương pháp dựa trên đối tượng, trường âm được mô tả bằng các đối tượng âm thanh rời rạc. Việc này yêu cầu siêu dữ liệu đối tượng mà mô tả trong số các yêu cầu khác vị trí biến đổi theo thời gian của mỗi nguồn âm trong không gian 3D.

Khái niệm mã hóa siêu dữ liệu thứ nhất trong tình trạng kỹ thuật là định dạng trao đổi mô tả âm trong không gian (spatial sound description interchange format - SpatDIF), định dạng mô tả ngữ cảnh âm thanh mà vẫn trong quá trình phát triển [M1]. Được thiết kế như định dạng trao đổi cho các ngữ cảnh âm dựa trên đối tượng và không cung cấp phương pháp nén bất kỳ cho các quỹ đạo của đối tượng. SpatDIF sử dụng định dạng điều khiển âm mở (Open Sound Control - OSC) dựa trên câu lệnh để cấu trúc siêu dữ liệu đối tượng [M2]. Tuy nhiên, việc biểu diễn dựa trên câu lệnh đơn giản không phải là lựa chọn sự truyền được nén của các quỹ đạo đối tượng.

Khái niệm siêu dữ liệu khác trong tình trạng kỹ thuật là định dạng mô tả ngữ cảnh âm thanh (Audio Scene Description Format -ASDF) [M3], giải pháp dựa trên câu lệnh mà có cùng nhược điểm. Dữ liệu được tạo cấu trúc bằng sự mở rộng của ngôn ngữ tích hợp đa phương tiện đồng bộ hóa (Synchronized Multimedia Integration Language - SMIL) mà là tập hợp con đánh dấu mở rộng được (Extensible Markup Language -XML) [M4], [M5].

Khái niệm siêu dữ liệu khác trong tình trạng kỹ thuật là định dạng nhị phân âm thanh đối với cảnh (AudioBIFS), định dạng nhị phân là phần của sự mô tả MPEG-4 [M6], [M7]. Khái niệm này gần tương quan với ngôn ngữ mô hình thực ảo (Virtual Reality Modeling Language - VRML) dựa trên XML mà đã được phát triển phần mô tả của các ngữ cảnh 3D thị giác-âm thanh và các ứng dụng thực ảo tương tác [M8]. Phần mô tả AudioBIFS phức sử dụng các biểu đồ ngữ cảnh để cụ thể các lộ trình dịch chuyển đối tượng. Nhược điểm chính của AudioBIFS là ở chỗ không được thiết kế cho hoạt động theo thời gian thực mà yêu cầu độ trễ hệ thống giới hạn và truy cập ngẫu nhiên vào dòng dữ liệu. Hơn nữa, việc mã hóa các vị trí đối tượng không khai thác sự thực hiện vị trí được giới hạn của người nghe. Đối với vị trí người nghe cố định trong ngữ cảnh thị giác-âm thanh, dữ liệu đối tượng có thể được lượng tử hóa với số lượng bit thấp hơn nhiều [M9]. Do đó, việc mã hóa siêu dữ liệu đối tượng được áp dụng ở AudioBIFS không hiệu quả đối với việc nén dữ liệu.

Tài liệu tham khảo

- [SAOC1] J. Herre, S. Disch, J. Hilpert, O. Hellmuth: "From SAC To SAOC - Recent Developments in Parametric Coding of Spatial Audio", 22nd Regional UK AES Conference, Cambridge, UK, April 2007.
- [SAOC2] J. Engdegård, B. Resch, C. Falch, O. Hellmuth, J. Hilpert, A. Hölzer, L. Terentiev, J. Breebaart, J. Koppens, E. Schuijers and W. Oomen: "Spatial Audio Object Coding (SAOC) – The Upcoming MPEG Standard on Parametric Object Based Audio Coding", 124th AES Convention, Amsterdam 2008.
- [SAOC] ISO/IEC, "MPEG audio technologies – Part 2: Spatial Audio Object Coding (SAOC)", ISO/IEC JTC1/SC29/WG11 (MPEG) International Standard 23003-2.
- [VBAP] Ville Pulkki, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning"; J. Audio Eng. Soc., Level 45, Issue 6, pp. 456-466, June 1997.
- [M1] Peters, N., Lossius, T. and Schacher J. C., "SpatDIF: Principles, Specification, and Examples", 9th Sound and Music Computing Conference, Copenhagen, Denmark, Jul. 2012.
- [M2] Wright, M., Freed, A., "Open Sound Control: A New Protocol for Communicating with Sound Synthesizers", International Computer Music Conference, Thessaloniki, Greece, 1997.
- [M3] Matthias Geier, Jens Ahrens, and Sascha Spors. (2010), "Object-based audio reproduction and the audio scene description format", Org. Sound, Vol. 15, No. 3, pp. 219-227, December 2010.
- [M4] W3C, "Synchronized Multimedia Integration Language (SMIL 3.0)", Dec. 2008.
- [M5] W3C, "Extensible Markup Language (XML) 1.0 (Fifth Edition)", Nov. 2008.
- [M6] MPEG, "ISO/IEC International Standard 14496-3 - Coding of audio-visual objects, Part 3 Audio", 2009.

- [M7] Schmidt, J.; Schroeder, E. F. (2004), "New and Advanced Features for Audio Presentation in the MPEG-4 Standard", 116th AES Convention, Berlin, Germany, May 2004.
- [M8] Web3D, "International Standard ISO/IEC 14772-1:1997 - The Virtual Reality Modeling Language (VRML), Part 1: Functional specification and UTF-8 encoding", 1997.
- [M9] Sporer, T. (2012), "Codierung räumlicher Audiosignale mit leichtgewichtigen Audio-Objekten", Proc. Annual Meeting of the German Audiological Society (DGA), Erlangen, Germany, Mar. 2012.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất các khái niệm được cải thiện để mã hóa đối tượng âm thanh trong không gian. Mục đích của sáng chế được giải quyết bởi thiết bị theo điểm 1 yêu cầu bảo hộ, thiết bị theo điểm 12 yêu cầu bảo hộ, hệ thống theo điểm 14 yêu cầu bảo hộ, phương pháp theo điểm 15 yêu cầu bảo hộ, phương pháp theo điểm 16 yêu cầu bảo hộ và chương trình máy tính theo điểm 17 và 18 yêu cầu bảo hộ.

Sáng chế đề xuất thiết bị để tạo ra một hoặc nhiều kênh đầu ra âm thanh. Thiết bị bao gồm bộ xử lý tham số tính toán thông tin trộn và bộ xử lý trộn giảm để tạo ra một hoặc nhiều kênh đầu ra âm thanh. Bộ xử lý trộn giảm được tạo cấu hình để nhận tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh. Một hoặc nhiều tín hiệu kênh âm thanh được trộn trong tín hiệu truyền tải âm thanh, và một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong tín hiệu truyền tải âm thanh, và trong đó số lượng của một hoặc nhiều kênh truyền tải âm thanh nhỏ hơn số lượng của một hoặc nhiều tín hiệu kênh âm thanh cộng với số lượng của một hoặc nhiều tín hiệu đối tượng âm thanh. Bộ xử lý tham số được tạo cấu hình để nhận thông tin trộn giảm biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong một hoặc nhiều kênh truyền tải âm thanh, và trong đó bộ xử lý tham số được tạo cấu hình để nhận thông tin phương sai. Hơn nữa, bộ xử lý tham số được tạo cấu hình để tính toán thông tin trộn phụ thuộc vào thông tin trộn giảm và phụ thuộc vào thông tin phương sai. Bộ xử lý trộn giảm được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu truyền tải âm

thanh phụ thuộc vào thông tin trộn. Thông tin phương sai biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu kênh âm thanh và còn biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh. Tuy nhiên, thông tin phương sai không biểu thị thông tin tương quan đối với cặp bất kỳ của một trong số một hoặc nhiều tín hiệu kênh âm thanh và một trong số một hoặc nhiều tín hiệu đối tượng âm thanh.

Hơn nữa, sáng chế đề xuất thiết bị để tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh. Thiết bị này bao gồm bộ trộn kênh/đối tượng để tạo ra một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh, và giao diện đầu ra. Bộ trộn kênh/đối tượng được tạo cấu hình để tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh bằng cách trộn một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh trong tín hiệu truyền tải âm thanh phụ thuộc vào thông tin trộn giảm biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh phải được trộn trong một hoặc nhiều kênh truyền tải âm thanh, trong đó số lượng của một hoặc nhiều kênh truyền tải âm thanh nhỏ hơn số lượng của một hoặc nhiều tín hiệu kênh âm thanh cộng với số lượng của một hoặc nhiều tín hiệu đối tượng âm thanh. Giao diện đầu ra được tạo cấu hình để xuất ra tín hiệu truyền tải âm thanh, thông tin trộn giảm và thông tin phương sai. Thông tin phương sai biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu kênh âm thanh và còn biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh. Tuy nhiên, thông tin phương sai không biểu thị thông tin tương quan đối với cặp bất kỳ của một trong số một hoặc nhiều tín hiệu kênh âm thanh và một trong số một hoặc nhiều tín hiệu đối tượng âm thanh.

Hơn nữa, sáng chế đề xuất hệ thống. Hệ thống này bao gồm thiết bị để tạo ra tín hiệu truyền tải âm thanh như được mô tả ở trên và thiết bị để tạo ra một hoặc nhiều kênh đầu ra âm thanh như được mô tả ở trên. Thiết bị để tạo ra một hoặc nhiều kênh đầu ra âm thanh được tạo cấu hình để nhận tín hiệu truyền tải âm thanh, thông tin trộn giảm và thông tin phương sai từ thiết bị để tạo ra tín hiệu truyền tải âm thanh. Hơn nữa, thiết bị để tạo ra các kênh đầu ra âm thanh được tạo cấu hình để tạo ra một hoặc

nhiều kênh đầu ra âm thanh tùy thuộc vào tín hiệu truyền tải âm thanh phụ thuộc vào thông tin trộn giảm và phụ thuộc vào thông tin phương sai.

Hơn nữa, sáng chế đề xuất phương pháp tạo ra một hoặc nhiều kênh đầu ra âm thanh. Phương pháp bao gồm:

- Nhận tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh, trong đó một hoặc nhiều tín hiệu kênh âm thanh được trộn trong tín hiệu truyền tải âm thanh, trong đó một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong tín hiệu truyền tải âm thanh, và trong đó số lượng của một hoặc nhiều kênh truyền tải âm thanh nhỏ hơn số lượng của một hoặc nhiều tín hiệu kênh âm thanh cộng với số lượng của một hoặc nhiều tín hiệu đối tượng âm thanh.
- Nhận thông tin trộn giảm biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong một hoặc nhiều kênh truyền tải âm thanh.
- Nhận thông tin phương sai.
- Tính toán thông tin trộn phụ thuộc vào thông tin trộn giảm và phụ thuộc vào thông tin phương sai. Và:
- Tạo ra một hoặc nhiều kênh đầu ra âm thanh.

Việc tạo ra một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu truyền tải âm thanh phụ thuộc vào thông tin trộn. Thông tin phương sai này biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu kênh âm thanh và còn biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh. Tuy nhiên, thông tin phương sai không biểu thị thông tin tương quan đối với cặp bất kỳ của một trong số một hoặc nhiều tín hiệu kênh âm thanh và một trong số một hoặc nhiều tín hiệu đối tượng âm thanh.

Hơn nữa, phương pháp tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh. Phương pháp bao gồm:

- Tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh bằng cách trộn một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh trong tín hiệu truyền tải âm thanh phụ thuộc vào

thông tin trộn giảm biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh phải được trộn trong một hoặc nhiều kênh truyền tải âm thanh, trong đó số lượng của một hoặc nhiều kênh truyền tải âm thanh nhỏ hơn số lượng của một hoặc nhiều tín hiệu kênh âm thanh cộng với số lượng của một hoặc nhiều tín hiệu đối tượng âm thanh.

Và:

- Xuất ra tín hiệu truyền tải âm thanh, thông tin trộn giảm và thông tin phương sai.

Thông tin phương sai biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu kênh âm thanh và còn biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh. Tuy nhiên, thông tin phương sai không biểu thị thông tin tương quan đối với cặp bất kỳ của một trong số một hoặc nhiều tín hiệu kênh âm thanh và một trong số một hoặc nhiều tín hiệu đối tượng âm thanh.

Hơn nữa, sáng chế đề xuất chương trình máy tính thực hiện phương pháp nêu trên khi được chạy trên máy tính hoặc bộ xử lý tín hiệu.

Mô tả vắn tắt các hình vẽ

Sau đây, các phương án theo sáng chế được mô tả chi tiết hơn với sự tham chiếu đến các hình vẽ, trong đó:

Fig.1 minh họa thiết bị để tạo ra một hoặc nhiều kênh đầu ra âm thanh theo phương án,

Fig.2 minh họa thiết bị để tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh theo phương án,

Fig.3 minh họa hệ thống theo phương án,

Fig.4 minh họa phương án thứ nhất của bộ mã hóa âm thanh 3D,

Fig.5 minh họa phương án thứ nhất của bộ giải mã âm thanh 3D,

Fig.6 minh họa phương án thứ hai của bộ mã hóa âm thanh 3D,

Fig.7 minh họa phương án thứ hai của bộ giải mã âm thanh 3D,

Fig.8 minh họa phương án thứ ba của bộ mã hóa âm thanh 3D,

Fig.9 minh họa phương án thứ ba của bộ giải mã âm thanh 3D, và

Fig.10 minh họa cụm xử lý kết hợp theo phương án.

Mô tả chi tiết sáng chế

Trước khi mô tả chi tiết các phương án được ưu tiên theo sáng chế, hệ thống bộ mã hóa-giải mã âm thanh 3D được mô tả.

Trong tình trạng kỹ thuật, không có kỹ thuật linh hoạt tồn tại một mặt là mã hóa kênh kết 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.

Các hạn chế này được giải quyết bởi hệ thống bộ mã hóa-giải mã âm thanh ba chiều mới.

Trước khi mô tả chi tiết các phương án được ưu tiên, hệ thống bộ mã hóa-giải mã âm thanh ba chiều được mô tả.

Fig.4 minh họa bộ mã hóa âm thanh ba chiều theo phương án của sáng chế. Bộ mã hóa âm thanh ba chiều được tạo 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 âm thanh ba chiều bao gồm giao diện đầu vào để nhận nhiều kênh âm thanh được biểu thị bằng CH và nhiều đối tượng âm thanh được biểu thị bằng OBJ. Hơn nữa, như được minh họa trên Fig.4, giao diện đầu vào 1100 còn nhận siêu dữ liệu liên quan đến một hoặc nhiều đối tượng trong số nhiều đối tượng âm thanh OBJ. Hơn nữa, bộ mã hóa âm thanh ba chiều 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 âm thanh ba chiều 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 trong số nhiều đối tượng âm thanh.

Hơn nữa, bộ mã hóa âm thanh ba chiều 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 trong số vài chế độ hoạt động, trong đó ở chế độ thứ nhất, bộ mã hóa lõi được tạo cấu hình

để mã hóa nhiều kênh âm thanh và nhiều đối tượng âm thanh được nhận bằng giao diện đầu vào 1100 mà không cần sự tương tác bất kỳ bởi bộ trộn, tức là, mà không có việc trộn bất kỳ bởi bộ trộn 200. Tuy nhiên, ở 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. Trường hợp thứ hai này, tốt hơn là không mã hóa dữ liệu đối tượng bất kỳ nào nữa. Thay vào đó, siêu dữ liệu biểu thị các vị trí của các đối tượng âm thanh đã được sử dụng bởi bộ trộn 200 để kết xuất các đối tượng lên trên các kênh như được biểu thị bởi siêu dữ liệu. Nói cách khác, bộ trộn 200 sử dụng siêu dữ liệu liên quan đến các 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 của đầu ra của bộ trộn. Theo 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 được nhập vào giao diện 1100 được trộn nhưng chỉ lượng đối tượng nhất định được trộn, thì chỉ các đối tượng không được trộn còn lại và tuy nhiên siêu dữ liệu được kết hợp được truyền lần lượt đến bộ mã hóa lõi 300 hoặc bộ nén siêu dữ liệu 400.

Fig.6 minh họa phương án khác của bộ mã hóa âm thanh ba chiều mà còn bao gồm bộ mã hóa SAOC 800. Bộ mã hóa SAOC 800 được tạo cấu hình để tạo ra một hoặc nhiều kênh truyền tải 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.6, dữ liệu đầu vào bộ mã hóa đối tượng âm thanh trong không gian là các đối tượng mà đã không được xử lý bằng bộ kết xuất trước/bộ trộn. Theo cách khác, miễn là bộ kết xuất trước/bộ trộn đã bị bỏ qua như theo chế độ một mà việc mã hóa kênh/đối tượng riêng biệt hoạt động, tất cả các đối tượng nhập vào giao diện đầu vào 1100 được mã hóa bởi bộ mã hóa SAOC 800.

Hơn nữa, như được minh họa trên Fig.6, bộ mã hóa lõi 300 tốt hơn là được thực hiện như bộ mã hóa USAC, tức là, như bộ mã hóa như được xác định và tiêu chuẩn hóa trong tiêu chuẩn USAC-MPEG (Unified Speech and Audio Coding - USAC). Đầu ra của toàn bộ bộ mã hóa âm thanh ba chiều được minh họa trên Fig.6 là dòng dữ liệu MPEG 4, Dòng dữ liệu MPEG H hoặc dòng dữ liệu âm thanh ba chiều có các cấu trúc giống như vật chứa đối với các loại dữ liệu riêng biệt. Hơn nữa, siêu dữ liệu được biểu

thị như dữ liệu “OAM” và bộ nén siêu dữ liệu 400 trên Fig.4 tương ứng với bộ mã hóa OAM 400 để thu được dữ liệu OAM được nén mà được nhập vào bộ mã hóa USAC 300 mà, như có thể được thấy trên Fig.6, 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.8 minh họa phương án khác của bộ mã hóa âm thanh ba chiều, trong đó trái với Fig.6, bộ mã hóa SAOC có thể được tạo cấu hình hoặc để mã hóa, với thuật toán mã hóa SAOC, các kênh được cung cấp ở bộ kết xuất trước/bộ trộn 200 mà không hoạt động ở chế độ này hoặc, theo cách khác, để mã hóa SAOC các kênh được kết xuất trước cộng với các đối tượng. Do đó, trên Fig.8, bộ mã hóa SAOC 800 có thể hoạt động trên ba loại dữ liệu đầu vào khác nhau, tức là, các kênh mà không có đối tượng được kết xuất trước bất kỳ, các kênh và đối tượng được kết xuất trước hoặc riêng các đối tượng. Hơn nữa, tốt hơn là cung cấp bộ giải mã OAM bổ sung 420 trên Fig.8 sao cho bộ mã hóa SAOC 800 sử dụng, đối với việc xử lý của nó, cùng dữ liệu như trên phía bộ giải mã, tức là, dữ liệu thu được bởi sự nén tổn hao thay vì dữ liệu OAM ban đầu.

Bộ mã hóa âm thanh ba chiều của Fig.8 có thể hoạt động ở một vài chế độ riêng biệt.

Ngoài chế độ thứ nhất và thứ hai như được thảo luận trong nội dung của Fig.4, bộ mã hóa âm thanh ba chiều của Fig.8 có thể còn hoạt động theo chế độ thứ ba mà trong đó bộ mã hóa lõi tạo ra một hoặc nhiều kênh truyền tải từ các đối tượng riêng biệt khi bộ kết xuất trước/bộ trộn 200 không hoạt động. Theo cách khác hoặc ngoài ra, ở chế độ thứ ba, bộ mã hóa SAOC 800 có thể tạo ra một hoặc nhiều kênh truyền tải 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.4 không hoạt động.

Cuối cùng, bộ mã hóa SAOC 800 có thể mã hóa, khi bộ mã hóa âm thanh ba chiều được tạo cấu hình ở chế độ thứ tư, các kênh cộng với các đối tượng được kết xuất trước như được tạo ra bởi bộ kết xuất trước/bộ trộn. Do đó, ở 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 đối tượng đã được chuyển đổi hoàn toàn thành các kênh truyền tải SAOC riêng biệt và

thông tin phụ có được kết hợp như được biểu thị trên các Fig.3 và 5 như “SAOC-SI” và, ngoài ra, siêu dữ liệu được nén bất kỳ không phải được truyền ở chế độ thứ tư này.

Fig.5 minh họa bộ giải mã âm thanh ba chiều theo phương án của sáng chế. Bộ giải mã âm thanh ba chiều nhận, như đầu vào, dữ liệu âm thanh được mã hóa, tức là, dữ liệu 501 ở Fig.4.

Bộ giải mã âm thanh ba chiều 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 ba chiều được tạo cấu hình để giải mã dữ liệu âm thanh được mã hóa và giao diện đầu vào được tạo 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 và nhiều đối tượng được mã hóa và siêu dữ liệu được nén liên quan đến nhiều đối tượng theo chế độ nhất định.

Hơn nữa, bộ giải mã lõi 1300 được tạo 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 tạo cấu hình để giải nén siêu dữ liệu được nén.

Hơn nữa, bộ xử lý đối tượng 1200 được tạo 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 được xác định trước của các kênh đầu ra bao gồm dữ liệu đối tượng và các kênh được giải mã. Sau đó, các kênh đầu ra này như được biểu thị ở 1205 được nhập vào bộ xử lý sau 1700. Bộ xử lý sau 1700 được tạo cấu hình để chuyển đổi số lượng 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 âm thanh ba chiều bao gồm bộ điều khiển chế độ 1600 mà được tạo cấu hình để phân tích dữ liệu được mã hóa để phát hiện chỉ báo chế độ. Do đó, bộ điều khiển chế độ 1600 được kết nối với giao diện đầu vào 1100 trên Fig.5. Tuy nhiên, theo cách khác, bộ điều khiển chế độ không nhất thiết phải ở đó. Thay vào đó, bộ giải mã âm thanh linh hoạt có thể được thiết lập lại bởi loại dữ liệu điều khiển khác bất kỳ như đầu vào của người dùng hoặc điều khiển khác bất kỳ. Bộ giải mã âm

thanh ba chiều trên Fig.5 và, tốt hơn là được điều khiển bởi bộ điều khiển chế độ 1600, được tạo cấu hình để bỏ qua bộ xử lý đối tượng và để cấp nhiều kênh được giải mã vào bộ xử lý sau 1700. Đây là việc 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 được áp dụng trong bộ mã hóa âm thanh ba chiều ở Fig.4. Theo cách khác, khi chế độ 1 được áp dụng ở bộ mã hóa âm thanh ba chiều, tức là, khi bộ mã hóa âm thanh ba chiều đã thực hiện việc mã hóa kênh/đối tượng riêng biệt, sau đó bộ xử lý đối tượng 1200 bị bỏ qua, nhưng nhiều kênh được giải mã và nhiều đối tượng được giải mã được cấp 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 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 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 âm thanh ba chiều của Fig.4.

Fig.7 minh họa phương án ưu tiên so với bộ giải mã âm thanh ba chiều của Fig.5 và phương án của Fig.7 tương ứng với bộ mã hóa âm thanh ba chiều của Fig.6. Ngoài việc thực hiện bộ giải mã âm thanh ba chiều của Fig.5, bộ giải mã âm thanh ba chiều trên Fig.7 bao gồm bộ giải mã SAOC 1800. Hơn nữa, bộ xử lý đối tượng 1200 của Fig.5 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 bởi bộ giải mã SAOC 1800.

Hơn nữa, bộ xử lý sau 1700 có thể được thực hiện như 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.5 cũng có thể được thực hiện như được minh họa bởi 1730. Do đó, tốt hơn là để thực hiện việc xử lý trong bộ giải mã với số lượng kênh cao nhất như 22.2 hoặc 32 để có tính linh hoạt và để 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 mà chỉ định dạng nhỏ như định dạng 5.1 được yêu cầu, tốt hơn là, như được biểu thị bởi Fig.5 hoặc Fig.6 bởi phím tắt 1727,

rằng việc 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 theo sáng chế, bộ xử lý đối tượng 1200 bao gồm bộ giải mã SAOC 1800 và bộ giải mã SAOC được tạo cấu hình để giải mã một hoặc nhiều kênh truyền tải 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 các đối tượng âm thanh được kết xuất. Để đạt được điều này, đầu ra OAM được kết nối với hộp 1800.

Hơn nữa, bộ xử lý đối tượng 1200 được tạo 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 truyền tải SAOC nhưng được mã hóa riêng biệt trong các thành phần được tạo kênh đơn 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 truyền tải và thông tin phụ tham số được kết hợp biểu diễn các tín hiệu 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 tạo 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 sử dụng được 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 tạo cấu hình để tính toán các kênh âm thanh của định dạng đầu ra bằng cách sử dụng các kênh truyền tải được giải mã và thông tin phụ tham số được chuyển mã. Việc xử lý này được thực hiện bằng bộ xử lý sau có thể tương tự với việc xử lý bao quanh MPEG hoặc có thể là việc xử lý bất kỳ khác 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 tạo cấu hình để trộn tăng trực tiếp và kết xuất các tín hiệu cho định dạng đầu ra bằng cách sử dụng các kênh truyền tải được giải mã (bởi bộ giải mã lõi) và thông tin phụ tham số.

Hơn nữa, và quan trọng là, bộ xử lý đối tượng 1200 của Fig.5 còn bao gồm bộ trộn 1220 mà nhận, như đầu vào, dữ liệu được xuất ra bởi bộ giải mã USAC 1300 một cách trực tiếp khi các đối tượng được kết xuất trước được trộn với các kênh tồn tại, tức là, khi bộ trộn 200 của Fig.4 hoạt động. Ngoài ra, bộ trộn 1220 nhận dữ liệu từ bộ kết xuất đối tượng thực hiện đối tượng kết xuất mã 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 tạo 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 các đáp ứng xung phòng lập thể (binaural room impulse response - BRIR). Bộ chuyển đổi định dạng 1720 được tạo cấu hình để chuyển đổi các kênh đầu ra thành định dạng đầu ra có số lượng kênh thấp hơn các kênh đầu ra 1205 của bộ trộn và bộ chuyển đổi định dạng 1720 yêu cầu thông tin trên cách bố trí tái tạo như các loa 5.1 hoặc tương tự.

Bộ giải mã âm thanh ba chiều của Fig.9 khác với bộ giải mã âm thanh ba chiều của Fig.7 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 mà khi bộ mã hóa âm thanh ba chiều của Fig.8 đã được sử dụng và kết nối 900 giữa các kênh /các đối tượng được kết xuất trước và bộ mã hóa SAOC 800 giao diện đầu vào hoạt động.

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

Khối VBAP tốt hơn là nhận dữ liệu OAM được giải mã để thu được các ma trận kết xuất. Thông thường hơn, tốt hơn là yêu cầu thông tin hình học không chỉ của cách bố trí tái tạo mà còn của các vị trí mà các tín hiệu đầu vào nên được kết xuất trên cách bố trí tái tạo. Dữ liệu đầu vào hình học có thể là dữ liệu OAM cho các đối tượng hoặc thông tin vị trí kênh cho các kênh mà đã được truyền sử dụng SAOC.

Tuy nhiên, nếu chỉ giao diện đầu ra cụ thể được yêu cầu thì giai đoạn VBAP 1810 có thể đã cung cấp ma trận kết xuất yêu cầu đối với, ví dụ, đầu ra 5.1. Bộ giải mã SAOC 1800 sau đó thực hiện việc kết xuất trực tiếp từ các kênh truyền tải SAOC, dữ liệu tham số được kết hợp và siêu dữ liệu được giải nén, kết xuất trực tiếp thành định dạng đầu ra 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, trong đó một vài kênh được mã hóa SAOC nhưng không phải tất cả kênh được mã hóa SAOC hoặc trong đó một vài đối tượng được mã hóa SAOC nhưng không phải tất cả các đối tượng được mã hóa SAOC hoặc khi chỉ lượng đối tượng được kết xuất trước nhất định với các kênh được giải mã SAOC và các kênh còn lại không được xử lý SAOC thì bộ trộn sẽ đặt cùng 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.

Các ký hiệu toán học sau đây được sử dụng:

$N_{Objects}$	số tín hiệu đối tượng âm thanh đầu vào
$N_{Channels}$	số kênh đầu vào
N	số tín hiệu đầu vào; N có thể bằng $N_{Objects}$, $N_{Channels}$ hoặc $N_{Objects} + N_{Channels}$
N_{DmxCh}	số kênh trộn giảm (được xử lý)
$N_{Samples}$	số mẫu dữ liệu được xử lý
$N_{OutputChannels}$	số kênh đầu ra ở phía bộ giải mã
$\mathbf{D}$	ma trận trộn giảm, kích cỡ $N_{DmxCh} \times N$
$\mathbf{X}$	tín hiệu âm thanh đầu vào, kích cỡ $N \times N_{Samples}$
$\mathbf{E}_X$ $\mathbf{X} \mathbf{X}^H$	ma trận phương sai tín hiệu đầu vào, kích cỡ $N \times N$ được xác định là $\mathbf{E}_X = \mathbf{X} \mathbf{X}^H$
$\mathbf{Y}$ $= \mathbf{D} \mathbf{X}$	tín hiệu âm thanh trộn giảm, kích cỡ $N_{DmxCh} \times N_{Samples}$ được xác định là $\mathbf{Y} = \mathbf{D} \mathbf{X}$
$\mathbf{E}_Y$	ma trận phương sai của các tín hiệu trộn giảm, kích cỡ $N_{DmxCh} \times N_{DmxCh}$ được xác định là $\mathbf{E}_Y = \mathbf{Y} \mathbf{Y}^H$

G	ma trận ước tính nguồn tham số, kích cỡ $N \times N_{DmxCh}$ mà xấp xỉ $\mathbf{E}_X \mathbf{D}^H (\mathbf{D} \mathbf{E}_X \mathbf{D}^H)^{-1}$
$\hat{\mathbf{X}}$	các tín hiệu đầu vào được khôi phục theo tham số, kích cỡ $N_{Objects} \times N_{Samples}$ mà xấp xỉ $\mathbf{X}$ và được xác định là $\hat{\mathbf{X}} = \mathbf{G}\mathbf{Y}$
$(\cdot)^H$	toán tử tự liên hợp (Hermitian) biểu diễn hoán vị liên hợp của $(\cdot)$
R	ma trận kết xuất có kích cỡ $N_{OutputChannels} \times N$
S	ma trận tạo ra kênh đầu ra có kích cỡ $N_{OutputChannels} \times N_{DmxCh}$ được xác định là $\mathbf{S} = \mathbf{R}\mathbf{G}$
Z	các tín hiệu đầu ra, kích cỡ $N_{OutputChannels} \times N_{Samples}$, được tạo ở phía bộ giải mã từ các tín hiệu trộn giảm, $\mathbf{Z} = \mathbf{S}\mathbf{Y}$
$\hat{\mathbf{Z}}$	các kênh đầu ra mong muốn, kích cỡ $N_{OutputChannels} \times N_{Samples}$, $\hat{\mathbf{Z}} = \mathbf{R}\mathbf{X}$

Không làm mất nguyên tắc chung, để cải thiện tính dễ đọc của các công thức, đối với tất cả các biến được đưa vào mà thời gian biểu thị các chỉ số và sự phụ thuộc vào tần số được bỏ qua trong tài liệu này.

Trong ngữ cảnh âm thanh ba chiều, các kênh loa phóng thanh được phân bố ở một vài lớp cao, tạo ra các cặp kênh dọc và ngang. Việc mã hóa chung chỉ hai kênh như được xác định trong USAC không đủ để xem xét các mối liên hệ cảm giác và trong không gian giữa các kênh.

Để xem xét các mối quan hệ cảm giác và trong không gian giữa các kênh, trong ngữ cảnh âm thanh ba chiều, một là có thể sử dụng kỹ thuật tham số giống như SAOC để khôi phục các kênh đầu vào (các tín hiệu kênh âm thanh và các tín hiệu đối tượng âm thanh mà được mã hóa bởi bộ mã hóa SAOC) để thu được các kênh đầu vào được khôi phục $\hat{\mathbf{X}}$ ở phía bộ giải mã. Việc giải mã SAOC dựa trên thuật toán sai số bình phương trung bình tối thiểu (Minimum Mean Squared Error -MMSE):

$$\hat{\mathbf{X}} = \mathbf{G}\mathbf{Y} \text{ với } \mathbf{G} \approx \mathbf{E}_X \mathbf{D}^H (\mathbf{D} \mathbf{E}_X \mathbf{D}^H)^{-1}.$$

Thay vì tái cấu trúc các kênh đầu vào để thu được các kênh đầu vào được khôi phục $\hat{\mathbf{X}}$, các kênh đầu ra $\mathbf{Z}$ có thể trực tiếp được tạo ra ở phía bộ giải mã bằng cách tính đến ma trận kết xuất $\mathbf{R}$.

$$\mathbf{Z} = \mathbf{R} \hat{\mathbf{X}}$$

$$\mathbf{Z} = \mathbf{R} \mathbf{G} \mathbf{Y}$$

$$\mathbf{Z} = \mathbf{S} \mathbf{Y} ; \text{ với } \mathbf{S} = \mathbf{R} \mathbf{G}$$

Như có thể thấy được, thay vì khôi phục rõ ràng các đối tượng đầu vào và các kênh âm thanh đầu vào, các kênh đầu ra $\mathbf{Z}$ có thể được tạo ra trực tiếp bằng cách áp dụng ma trận tạo ra kênh đầu ra $\mathbf{S}$ ở tín hiệu âm thanh trộn giảm $\mathbf{Y}$.

Để thu được ma trận tạo ra kênh đầu ra $\mathbf{S}$, ma trận kết xuất $\mathbf{R}$ có thể, ví dụ, được xác định hoặc có thể, ví dụ, đã có sẵn. Hơn nữa, ma trận ước tính nguồn tham số $\mathbf{G}$ có thể, ví dụ, được tính toán như được mô tả ở trên. Sau đó có thể thu được ma trận tạo ra kênh đầu ra $\mathbf{S}$ như sản phẩm của ma trận $\mathbf{S} = \mathbf{R} \mathbf{G}$ từ ma trận kết xuất $\mathbf{R}$ và ma trận ước tính nguồn tham số $\mathbf{G}$.

Hệ thống âm thanh ba chiều có thể yêu cầu chế độ kết hợp để mã hóa các kênh và các đối tượng.

Nói chung, đối với chế độ kết hợp này, việc mã hóa/ giải mã SAOC có thể được áp dụng theo hai cách khác nhau:

Một phương pháp có thể là để sử dụng một trường hợp về hệ thống tham số như SAOC, trong đó trường hợp này có khả năng xử lý các kênh và các đối tượng. Giải pháp này có nhược điểm ở chỗ đó là sự phức tạp tính toán, do số lượng tín hiệu đầu vào cao nên số lượng kênh truyền tải sẽ tăng để duy trì chất lượng khôi phục tương tự. Kết quả là kích cỡ của ma trận $\mathbf{D} \mathbf{E}_X \mathbf{D}^H$ sẽ tăng và độ phức tạp nghịch đảo sẽ tăng lên. Hơn nữa, giải pháp này có thể đưa nhiều hơn sự mất ổn định số học do kích cỡ của ma trận $\mathbf{D} \mathbf{E}_X \mathbf{D}^H$ tăng lên. Hơn nữa, nhược điểm khác là, sự nghịch đảo của ma trận $\mathbf{D} \mathbf{E}_X \mathbf{D}^H$ có thể dẫn đến nhiễu xuyên âm bổ sung giữa các kênh được khôi phục và các đối tượng được khôi phục. Việc này được gây ra do một số hệ số trong ma trận khôi phục $\mathbf{G}$ được cho là bằng 0 được thiết lập về trị số khác 0 do tính mất chính xác về số học.

Phương pháp khác có thể sử dụng hai trường hợp của hệ thống tham số như SAOC, một trường hợp cho việc xử lý dựa trên kênh và trường hợp khác cho việc xử lý dựa trên đối tượng. Phương pháp này sẽ có nhược điểm ở chỗ thông tin giống nhau được truyền hai lần để khởi tạo các giàn lọc và cấu hình bộ giải mã. Hơn nữa, không

thể trộn các kênh và đối tượng với nhau nếu việc này yêu cầu, và do đó không thể sử dụng các thuộc tính tương quan giữa các kênh và các đối tượng.

Để tránh các nhược điểm của phương pháp sử dụng các trường hợp khác nhau đối với các đối tượng âm thanh và các kênh âm thanh, các phương án sử dụng phương pháp thứ nhất và tạo ra hệ thống SAOC tăng cường có khả năng xử lý các kênh, đối tượng hoặc các đối tượng chỉ sử dụng một trường hợp hệ thống, theo cách hiệu quả. Mặc dù các kênh âm thanh và các đối tượng âm thanh lần lượt được xử lý bởi cùng trường hợp bộ mã hóa và bộ giải mã, các khái niệm hiệu quả được tạo ra, sao cho các nhược điểm của phương pháp có thể tránh được.

Fig.2 minh họa thiết bị để tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh theo phương án.

Phương pháp này bao gồm bộ trộn kênh/đối tượng 210 để tạo ra một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh, và giao diện đầu ra 220.

Bộ trộn kênh/đối tượng 210 được tạo cấu hình để tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh bằng cách trộn một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh trong tín hiệu truyền tải âm thanh phụ thuộc vào thông tin trộn giảm biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh phải được trộn trong một hoặc nhiều kênh truyền tải âm thanh.

Số lượng của một hoặc nhiều kênh truyền tải âm thanh nhỏ hơn số lượng của một hoặc nhiều tín hiệu kênh âm thanh cộng với số lượng của một hoặc nhiều tín hiệu đối tượng âm thanh. Do đó, bộ trộn kênh/đối tượng 210 có khả năng trộn giảm một hoặc nhiều tín hiệu kênh âm thanh với nhau và một hoặc nhiều tín hiệu đối tượng âm thanh, do bộ trộn kênh/đối tượng 210 được thích ứng để tạo ra tín hiệu truyền tải âm thanh mà có ít kênh hơn số lượng của một hoặc nhiều tín hiệu kênh âm thanh cùng với số lượng của một hoặc nhiều tín hiệu đối tượng âm thanh.

Giao diện đầu ra 220 được tạo cấu hình để xuất ra tín hiệu truyền tải âm thanh, thông tin trộn giảm và thông tin phương sai.

Ví dụ, bộ trộn kênh/đối tượng 210 có thể được tạo cấu hình để dẫn thông tin trộn giảm, mà được sử dụng để trộn giảm một hoặc nhiều tín hiệu kênh âm thanh và

một hoặc nhiều tín hiệu đối tượng âm thanh, vào giao diện đầu ra 220. Hơn nữa, ví dụ, giao diện đầu ra 220, có thể, ví dụ, được tạo cấu hình để nhận một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh và có thể được tạo cấu hình thêm nữa để xác định thông tin phương sai dựa trên một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh. Hoặc, giao diện đầu ra 220 có thể, ví dụ, được tạo cấu hình để nhận thông tin phương sai đã được xác định.

Thông tin phương sai biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu kênh âm thanh và còn biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh. Tuy nhiên, thông tin phương sai không biểu thị thông tin tương quan đối với cặp bất kỳ của một trong số một hoặc nhiều tín hiệu kênh âm thanh và một trong số một hoặc nhiều tín hiệu đối tượng âm thanh.

Fig.1 minh họa thiết bị để tạo ra một hoặc nhiều kênh đầu ra âm thanh theo phương án.

Phương pháp này bao gồm bộ xử lý tham số 110 tính toán thông tin trộn và bộ xử lý trộn giảm 120 để tạo ra một hoặc nhiều kênh đầu ra âm thanh.

Bộ xử lý trộn giảm 120 được tạo cấu hình để nhận tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh. Một hoặc nhiều tín hiệu kênh âm thanh được trộn trong tín hiệu truyền tải âm thanh. Hơn nữa, một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong tín hiệu truyền tải âm thanh. Số lượng của một hoặc nhiều kênh truyền tải âm thanh nhỏ hơn số lượng của một hoặc nhiều tín hiệu kênh âm thanh cùng với số lượng của một hoặc nhiều tín hiệu đối tượng âm thanh.

Bộ xử lý tham số 110 được tạo cấu hình để nhận thông tin trộn giảm biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong một hoặc nhiều kênh truyền tải âm thanh. Hơn nữa, bộ xử lý tham số 110 được tạo cấu hình để nhận thông tin phương sai. Bộ xử lý tham số 110 được tạo cấu hình để tính toán thông tin trộn phụ thuộc vào thông tin trộn giảm và phụ thuộc vào thông tin phương sai.

Bộ xử lý trộn giảm 120 được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu truyền tải âm thanh phụ thuộc vào thông tin trộn.

Thông tin phương sai biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu kênh âm thanh và còn biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh. Tuy nhiên, thông tin phương sai không biểu thị thông tin tương quan đối với cặp bất kỳ của một trong số một hoặc nhiều tín hiệu kênh âm thanh và một trong số một hoặc nhiều tín hiệu đối tượng âm thanh.

Theo phương án, thông tin phương sai có thể, ví dụ, biểu thị thông tin chênh lệch mức đối với mỗi trong số một hoặc nhiều tín hiệu kênh âm thanh và, có thể còn, ví dụ, biểu thị thông tin chênh lệch mức đối với mỗi trong số một hoặc nhiều tín hiệu đối tượng âm thanh.

Theo phương án, hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh có thể, ví dụ, được trộn trong tín hiệu truyền tải âm thanh và hai hoặc nhiều hơn hai tín hiệu kênh âm thanh có thể, ví dụ, được trộn trong tín hiệu truyền tải âm thanh. Thông tin phương sai có thể, ví dụ, biểu thị thông tin tương quan đối với một hoặc nhiều cặp của một tín hiệu kênh âm thanh thứ nhất trong số hai hoặc nhiều hơn hai tín hiệu kênh âm thanh và một tín hiệu kênh âm thanh thứ hai trong số hai hoặc nhiều hơn hai tín hiệu kênh âm thanh. Hoặc, thông tin phương sai có thể, ví dụ, biểu thị thông tin tương quan đối với một hoặc nhiều cặp của một tín hiệu đối tượng âm thanh thứ nhất trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh và một tín hiệu đối tượng âm thanh thứ hai trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh. Hoặc, thông tin phương sai có thể, ví dụ, biểu thị thông tin tương quan đối với một hoặc nhiều cặp của một tín hiệu kênh âm thanh thứ nhất trong số hai hoặc nhiều hơn hai tín hiệu kênh âm thanh và một tín hiệu kênh âm thanh thứ hai trong số hai hoặc nhiều hơn hai tín hiệu kênh âm thanh và biểu thị thông tin tương quan đối với một hoặc nhiều cặp của một tín hiệu đối tượng âm thanh thứ nhất trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh và một tín hiệu đối tượng âm thanh thứ hai trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

Thông tin chênh lệch mức đối với tín hiệu đối tượng âm thanh có thể, ví dụ, là chênh lệch mức đối tượng (object level difference - OLD). “Mức” có thể, ví dụ, liên quan đến mức năng lượng. “Chênh lệch” có thể, ví dụ, liên quan đến sự chênh lệch đối với mức tối đa trong số các tín hiệu đối tượng âm thanh.

Thông tin tương quan đối với cặp của một tín hiệu đối tượng âm thanh thứ nhất trong số các tín hiệu đối tượng âm thanh và một tín hiệu đối tượng âm thanh thứ hai trong số các tín hiệu đối tượng âm thanh có thể, ví dụ, là tương quan liên đối tượng (inter-object correlation - IOC).

Ví dụ, theo phương án, để đảm bảo sự biểu diễn tối ưu của SAOC 3D được khuyến cáo sử dụng các tín hiệu đối tượng âm thanh đầu vào có năng lượng tương thích. Sản phẩm của hai tín hiệu âm thanh đầu vào (được chuẩn hóa theo các dải thời gian/tần số tương ứng) được xác định là:

$$nrg_{i,j}^{l,m} = \frac{\sum_{n \in l} \sum_{k \in m} x_i^{n,k} (x_j^{n,k})^H}{\sum_{n \in l} \sum_{k \in m} 1} + \varepsilon .$$

Ở đây, i và j lần lượt là các chỉ số đối với các tín hiệu đối tượng âm thanh x_i và x_j , n biểu thị thời gian, k biểu thị tần số, l biểu thị tập hợp các chỉ số thời gian và m biểu thị tập hợp các chỉ số tần số. ε là hằng số bổ sung để tránh chia cho 0, ví dụ, $\varepsilon = 10^{-9}$.

Năng lượng đối tượng tuyệt đối (NRG) của đối tượng có năng lượng cao nhất có thể, ví dụ, được tính là:

$$NRG^{l,m} = \max_i (nrg_{i,i}^{l,m}) .$$

Tỷ số của các công suất của tín hiệu đối tượng đầu vào tương ứng (OLD) có thể, ví dụ, được cho bởi

$$OLD_i^{l,m} = \frac{nrg_{i,i}^{l,m}}{NRG^{l,m}} .$$

Số đo tương tự của các đối tượng đầu vào (IOC), có thể, ví dụ, được cho bởi độ tương quan chéo:

$$IOC_{i,j}^{l,m} = \text{Re} \left\{ \frac{nrg_{i,j}^{l,m}}{\sqrt{nrg_{i,i}^{l,m} nrg_{j,j}^{l,m}}} \right\} .$$

Ví dụ, trong phương án, các IOC có thể được truyền đối với tất cả các cặp tín hiệu âm thanh i và j , mà dòng bit có thể biến đổi $bsRelatedTo[i][j]$ được thiết lập là một.

Thông tin chênh lệch mức đối với tín hiệu kênh âm thanh có thể, ví dụ, là chênh lệch mức kênh (channel level difference - CLD). “Mức” có thể, ví dụ, liên quan đến mức năng lượng. “Chênh lệch” có thể, ví dụ, liên quan đến sự chênh lệch đối với mức tối đa trong số các tín hiệu kênh âm thanh.

Thông tin tương quan đối với cặp của một tín hiệu kênh âm thanh thứ nhất trong số các tín hiệu kênh âm thanh và một tín hiệu kênh âm thanh thứ hai trong số các tín hiệu kênh âm thanh có thể, ví dụ, là độ tương quan liên kênh (ICC).

Theo phương án, độ chênh lệch mức kênh (CLD) có thể được xác định theo cùng một cách là độ chênh lệch mức đối tượng (OLD) nêu trên, khi các tín hiệu đối tượng âm thanh trong công thức nêu trên được thay thế bằng các tín hiệu kênh âm thanh. Hơn nữa, độ tương quan liên kênh (ICC) có thể được xác định theo cùng cách như độ tương quan liên đối tượng (IOC) nêu trên, khi các tín hiệu đối tượng âm thanh trong công thức nêu trên được thay thế bởi các tín hiệu kênh âm thanh.

Trong SAOC, bộ mã hóa SAOC trộn giảm (theo thông tin trộn giảm, ví dụ, theo ma trận trộn giảm **D**) các tín hiệu đối tượng âm thanh để thu được (ví dụ, số lượng ít hơn) một hoặc nhiều kênh truyền tải âm thanh. Ở phía bộ giải mã, bộ giải mã SAOC giải mã một hoặc nhiều kênh truyền tải âm thanh sử dụng thông tin trộn giảm được nhận từ bộ mã hóa và sử dụng thông tin phương sai được nhận từ bộ mã hóa. Thông tin phương sai có thể, ví dụ, là các hệ số của ma trận phương sai **E**, mà biểu thị các chênh lệch đối tượng của các tín hiệu đối tượng âm thanh và các độ tương quan liên đối tượng giữa hai tín hiệu đối tượng âm thanh. Trong SAOC, ma trận trộn giảm được xác định **D** và ma trận phương sai được xác định **E** được sử dụng để giải mã các mẫu của một hoặc nhiều kênh truyền tải âm thanh (ví dụ, 2048 mẫu của một hoặc nhiều kênh truyền tải âm thanh). Bằng cách sử dụng khái niệm này, tốc độ bit được tiết kiệm so với việc truyền một hoặc nhiều tín hiệu đối tượng âm thanh mà không mã hóa.

Các phương án dựa trên phát hiện, rằng mặc dù các tín hiệu đối tượng âm thanh và các tín hiệu kênh âm thanh có các chênh lệch đáng kể, tín hiệu truyền tải âm thanh

có thể được tạo ra bởi bộ mã hóa SAOC được tăng cường, sao cho tín hiệu truyền tải âm thanh này, không chỉ các tín hiệu đối tượng âm thanh, mà còn các tín hiệu kênh âm thanh được trộn.

Các tín hiệu đối tượng âm thanh và các tín hiệu kênh âm thanh khác nhau đáng kể. Ví dụ, mỗi tín hiệu đối tượng âm thanh trong số nhiều tín hiệu đối tượng âm thanh có thể biểu diễn nguồn âm thanh của kịch bản âm thanh. Do đó, nói chung, hai đối tượng âm thanh có thể không tương quan nhiều. Ngược lại, các tín hiệu kênh âm thanh biểu diễn các kênh khác nhau của kịch bản âm, như thể được ghi bởi các tai nghe khác nhau. Nói chung, hai trong số các tín hiệu kênh âm thanh này được tương quan cao, cụ thể là, so với sự tương quan của hai tín hiệu đối tượng âm thanh, mà nhìn chung không có tương quan cao. Do đó, các phương án dựa trên phát hiện ra rằng các tín hiệu kênh âm thanh đặc biệt thu được từ việc truyền sự tương quan giữa cặp của hai tín hiệu kênh âm thanh và bằng cách sử dụng trị số tương quan được truyền này để giải mã.

Hơn nữa, các tín hiệu đối tượng âm thanh và các tín hiệu kênh âm thanh khác nhau ở chỗ, thông tin vị trí được chuyển thành các tín hiệu đối tượng âm thanh, ví dụ, biểu thị vị trí (được xem xét) của nguồn âm (ví dụ, đối tượng âm thanh) mà từ đó tín hiệu đối tượng âm thanh bắt đầu. Thông tin vị trí này (ví dụ, chứa trong thông tin siêu dữ liệu) có thể được sử dụng khi tạo ra các kênh đầu ra âm thanh từ tín hiệu truyền tải âm thanh ở phía bộ giải mã. Tuy nhiên, ngược lại, các tín hiệu kênh âm thanh không thể hiện vị trí, và không có thông tin vị trí nào được chuyển đến các tín hiệu kênh âm thanh. Tuy nhiên, các phương án dựa trên phát hiện ra rằng, tuy nhiên có hiệu quả để mã hóa SAOC các tín hiệu kênh âm thanh cùng với các tín hiệu đối tượng âm thanh, ví dụ như tạo ra các tín hiệu kênh âm thanh có thể được chia thành hai vấn đề nhỏ, cụ thể là, xác định thông tin giải mã (ví dụ, xác định ma trận **G** để không trộn, xem dưới đây), mà không có thông tin vị trí nào cần thiết, và xác định thông tin kết xuất (ví dụ, bằng cách xác định ma trận kết xuất **R**, xem dưới đây), mà thông tin vị trí ở các tín hiệu đối tượng âm thanh có thể được sử dụng để kết xuất các đối tượng âm thanh ở các kênh đầu ra âm thanh được tạo ra.

Hơn nữa, sáng chế dựa trên phát hiện ra rằng không có sự tương quan (hoặc ít nhất không đáng kể) tồn tại giữa cặp bất kỳ của một trong các tín hiệu đối tượng âm thanh và một trong số các tín hiệu kênh âm thanh. Do đó, khi bộ mã hóa không truyền

thông tin tương quan đối với cặp bất kỳ của một trong số một hoặc nhiều tín hiệu kênh âm thanh và một trong số một hoặc nhiều tín hiệu đối tượng âm thanh. Bằng cách này, băng thông truyền đáng kể được lưu và lượng thời gian đáng kể được lưu cho cả mã hóa và giải mã. Bộ giải mã được tạo cấu hình để không xử lý thông tin tương quan không đáng kể này lưu lượng thời gian tính toán khi xác định thông tin trộn (được sử dụng để tạo ra các kênh đầu ra âm thanh từ tín hiệu truyền tải âm thanh ở phía bộ giải mã).

Theo phương án, bộ xử lý tham số 110 có thể, ví dụ, được tạo cấu hình để nhận thông tin kết xuất biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong một hoặc nhiều kênh đầu ra âm thanh. Bộ xử lý tham số 110 có thể, ví dụ, được tạo cấu hình để tính toán thông tin trộn phụ thuộc vào thông tin trộn giảm, phụ thuộc vào thông tin phương sai và phụ thuộc vào thông tin kết xuất.

Ví dụ, bộ xử lý tham số 110 có thể, ví dụ, được tạo cấu hình để nhận các hệ số của ma trận kết xuất $\mathbf{R}$ như thông tin kết xuất, và có thể được tạo cấu hình để tính toán thông tin trộn phụ thuộc vào thông tin trộn giảm, phụ thuộc vào thông tin phương sai và phụ thuộc vào ma trận kết xuất $\mathbf{R}$. Ví dụ, bộ xử lý tham số có thể nhận các hệ số của ma trận kết xuất $\mathbf{R}$ từ phía bộ mã hóa, hoặc từ người dùng. Theo phương án khác, bộ xử lý tham số 110 có thể, ví dụ, được tạo cấu hình để nhận thông tin siêu dữ liệu, ví dụ, thông tin vị trí hoặc thông tin độ khuếch đại, và có thể, ví dụ, được tạo cấu hình để tính toán các hệ số của ma trận kết xuất $\mathbf{R}$ phụ thuộc vào thông tin siêu dữ liệu được nhận. Theo phương án khác, bộ xử lý tham số có thể được tạo cấu hình để nhận cả hai (thông tin kết xuất từ bộ mã hóa và từ người dùng) và tạo ra ma trận kết xuất dựa trên cả hai (về cơ bản tức là, tính tương tác được nhận ra).

Hoặc, bộ xử lý tham số có thể, ví dụ, nhận hai ma trận con kết xuất $\mathbf{R}_{ch}$, $\mathbf{R}_{obj}$, như thông tin kết xuất, trong đó $\mathbf{R}=(\mathbf{R}_{ch}, \mathbf{R}_{obj})$, trong đó $\mathbf{R}_{ch}$ ví dụ, biểu thị đối với cách mà để trộn các tín hiệu kênh âm thanh đến các kênh đầu ra âm thanh và trong đó $\mathbf{R}_{obj}$ có thể là ma trận kết xuất thu được từ thông tin OAM, trong đó $\mathbf{R}_{obj}$ có thể, ví dụ, được tạo ra bởi khối VBAP 1810 của Fig.9.

Theo phương án cụ thể, hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh có thể, ví dụ, được trộn trong tín hiệu truyền tải âm thanh, hai hoặc nhiều hơn hai tín hiệu

kênh âm thanh được trộn trong tín hiệu truyền tải âm thanh. Theo phương án này, thông tin phương sai có thể, ví dụ, biểu thị thông tin tương quan đối với một hoặc nhiều cặp của một tín hiệu kênh âm thanh thứ nhất trong số hai hoặc nhiều hơn hai tín hiệu kênh âm thanh và một tín hiệu kênh âm thanh thứ hai trong số hai hoặc nhiều hơn hai tín hiệu kênh âm thanh. Hơn nữa, theo phương án này, thông tin phương sai (cụ thể là ví dụ, được truyền từ phía bộ mã hóa đến phía bộ giải mã) không biểu thị thông tin tương quan đối với cặp bất kỳ của một tín hiệu đối tượng âm thanh thứ nhất trong số một hoặc nhiều tín hiệu đối tượng âm thanh và một tín hiệu đối tượng âm thanh thứ hai trong số một hoặc nhiều tín hiệu đối tượng âm thanh, do sự tương quan giữa các tín hiệu đối tượng âm thanh có thể quá nhỏ, nên nó có thể bị bỏ qua, và do đó, ví dụ, được truyền để lưu tốc độ bit và thời gian xử lý. Theo phương án này, bộ xử lý tham số 110 được tạo cấu hình để tính toán thông tin trộn phụ thuộc vào thông tin trộn giảm, phụ thuộc vào thông tin chênh lệch mức của mỗi trong số một hoặc nhiều tín hiệu kênh âm thanh, phụ thuộc vào thông tin chênh lệch mức thứ hai của mỗi trong số một hoặc nhiều tín hiệu đối tượng âm thanh, và phụ thuộc vào thông tin tương quan của một hoặc nhiều cặp của một tín hiệu kênh âm thanh thứ nhất trong số hai hoặc nhiều hơn hai tín hiệu kênh âm thanh và một tín hiệu kênh âm thanh thứ hai trong số hai hoặc nhiều hơn hai tín hiệu kênh âm thanh. Phương án này sử dụng việc phát hiện nêu trên rằng sự tương quan giữa các tín hiệu đối tượng âm thanh thường tương đối thấp và nên được bỏ qua, trong khi sự tương quan giữa hai tín hiệu kênh âm thanh thường tương đối cao và nên được xem xét. Bằng cách không xử lý thông tin tương quan không liên quan giữa các tín hiệu đối tượng âm thanh, thời gian xử lý có thể lưu lại được. Bằng cách xử lý sự tương quan liên quan giữa các tín hiệu kênh âm thanh, có thể tăng cường hiệu quả mã hóa.

Theo các phương án cụ thể, một hoặc nhiều tín hiệu kênh âm thanh được trộn trong nhóm thứ nhất của một hoặc nhiều kênh trong số các kênh truyền tải âm thanh, trong đó một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong nhóm thứ hai của một hoặc nhiều kênh trong số các kênh truyền tải âm thanh, trong đó mỗi kênh truyền tải âm thanh của nhóm thứ nhất không được bao gồm bởi nhóm thứ hai, và trong đó mỗi kênh truyền tải âm thanh của nhóm thứ hai không được bao gồm bởi nhóm thứ nhất. Theo các phương án này, thông tin trộn giảm bao gồm thông tin con trộn giảm

thứ nhất biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh được trộn trong nhóm thứ nhất của một hoặc nhiều kênh truyền tải âm thanh, và thông tin trộn giảm bao gồm thông tin con trộn giảm thứ hai biểu thị thông tin về cách mà một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong nhóm thứ hai của một hoặc nhiều kênh truyền tải âm thanh. Theo các phương án này, bộ xử lý tham số 110 được tạo cấu hình để tính toán thông tin trộn phụ thuộc vào thông tin con trộn giảm thứ nhất, phụ thuộc vào thông tin con trộn giảm thứ hai và phụ thuộc vào thông tin phương sai, và bộ xử lý trộn giảm 120 được tạo cấu hình để tạo ra một hoặc nhiều tín hiệu đầu ra âm thanh từ nhóm thứ nhất của một hoặc nhiều kênh truyền tải âm thanh và từ nhóm thứ hai của các kênh truyền tải âm thanh phụ thuộc vào thông tin trộn. Do hiệu quả mã hóa tiếp cận này tăng lên, như giữa các tín hiệu kênh âm thanh của kịch bản âm, sự tương quan cao tồn tại. Hơn nữa, các hệ số của ma trận trộn giảm biểu thị sự ảnh hưởng của các tín hiệu kênh âm thanh ở các kênh truyền tải âm thanh, mà mã hóa các tín hiệu đối tượng âm thanh, và ngược lại, không phải được tính toán bởi bộ mã hóa, không phải được truyền, và có thể được thiết lập về không bởi bộ giải mã mà không yêu cầu xử lý chúng. Việc này tiết kiệm băng thông truyền và thời gian tính toán cho bộ mã hóa và bộ giải mã.

Theo phương án, bộ xử lý trộn giảm 120 được tạo cấu hình để nhận tín hiệu truyền tải âm thanh trong dòng bit, bộ xử lý trộn giảm 120 được tạo cấu hình để nhận số đếm kênh thứ nhất biểu thị số lượng của các kênh truyền tải âm thanh chỉ mã hóa các tín hiệu kênh âm thanh, và bộ xử lý trộn giảm 120 được tạo cấu hình để nhận số đếm kênh thứ hai biểu thị số lượng của các kênh truyền tải âm thanh chỉ mã hóa các tín hiệu đối tượng âm thanh. Theo phương án này, bộ xử lý trộn giảm 120 được tạo cấu hình để nhận biết liệu kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh mã hóa các tín hiệu kênh âm thanh hoặc liệu kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh mã hóa các tín hiệu đối tượng âm thanh phụ thuộc vào số đếm kênh thứ nhất hoặc phụ thuộc vào số đếm kênh thứ hai, hoặc phụ thuộc vào số đếm kênh thứ nhất và số đếm kênh thứ hai. Ví dụ, trong dòng bit này, các kênh truyền tải âm thanh mã hóa các tín hiệu kênh âm thanh xuất hiện trước và các kênh truyền tải âm thanh mã hóa các tín hiệu đối tượng âm thanh xuất hiện sau đó. Sau đó, nếu số đếm kênh thứ nhất là, ví dụ, 3 và số đếm kênh thứ hai là, ví dụ, 2, bộ xử lý trộn giảm có kết luận rằng

ba kênh truyền tải âm thanh đầu tiên bao gồm tín hiệu kênh âm thanh được mã hóa và hai kênh truyền tải âm thanh sau đó bao gồm các tín hiệu đối tượng âm thanh được mã hóa.

Theo phương án, bộ xử lý tham số 110 được tạo cấu hình để nhận thông tin siêu dữ liệu bao gồm thông tin vị trí, trong đó thông tin vị trí biểu thị vị trí cho mỗi tín hiệu đối tượng âm thanh trong số một hoặc nhiều tín hiệu đối tượng âm thanh, và trong đó thông tin vị trí không biểu thị vị trí cho bất kỳ trong số một hoặc nhiều tín hiệu kênh âm thanh. Theo phương án này, bộ xử lý tham số 110 được tạo cấu hình để tính toán thông tin trộn phụ thuộc vào thông tin trộn giảm, phụ thuộc vào thông tin phương sai, và phụ thuộc vào thông tin vị trí. Ngoài ra hoặc theo cách khác, thông tin siêu dữ liệu còn bao gồm thông tin độ khuếch đại, trong đó thông tin độ khuếch đại biểu thị trị số khuếch đại cho mỗi tín hiệu đối tượng âm thanh trong số một hoặc nhiều tín hiệu đối tượng âm thanh, và trong đó thông tin độ khuếch đại không biểu thị trị số khuếch đại cho tín hiệu kênh âm thanh bất kỳ trong số một hoặc nhiều tín hiệu kênh âm thanh. Theo phương án này, bộ xử lý tham số 110 có thể được tạo cấu hình để tính toán thông tin trộn phụ thuộc vào thông tin trộn giảm, phụ thuộc vào thông tin phương sai, phụ thuộc vào thông tin vị trí, và phụ thuộc vào thông tin độ khuếch đại. Ví dụ, bộ xử lý tham số 110 có thể được tạo cấu hình để tính toán thông tin trộn hơn nữa phụ thuộc vào ma trận con $\mathbf{R}_{ch}$ nêu trên.

Theo phương án, bộ xử lý tham số 110 được tạo cấu hình để tính toán ma trận trộn $\mathbf{S}$ làm thông tin trộn, trong đó ma trận trộn $\mathbf{S}$ được xác định theo công thức $\mathbf{S} = \mathbf{R}\mathbf{G}$, trong đó $\mathbf{G}$ là ma trận giải mã phụ thuộc vào thông tin trộn giảm và phụ thuộc vào thông tin phương sai, trong đó $\mathbf{R}$ là ma trận kết xuất phụ thuộc vào thông tin siêu dữ liệu. Theo phương án này, bộ xử lý trộn giảm (120) có thể được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh bằng cách áp dụng công thức $\mathbf{Z} = \mathbf{S}\mathbf{Y}$, trong đó $\mathbf{Z}$ là tín hiệu đầu ra âm thanh, và trong đó $\mathbf{Y}$ là tín hiệu truyền tải âm thanh. Ví dụ, $\mathbf{R}$ có thể phụ thuộc vào các ma trận con $\mathbf{R}_{ch}$ và/hoặc $\mathbf{R}_{obj}$ (ví dụ, $\mathbf{R} = (\mathbf{R}_{ch}, \mathbf{R}_{obj})$) nêu trên.

Fig.3 minh họa hệ thống theo phương án. Hệ thống bao gồm thiết bị 310 để tạo ra tín hiệu truyền tải âm thanh như nêu trên và thiết bị 320 để tạo ra một hoặc nhiều kênh đầu ra âm thanh như nêu trên.

Thiết bị 320 để tạo ra một hoặc nhiều kênh đầu ra âm thanh được tạo cấu hình để nhận tín hiệu truyền tải âm thanh, thông tin trộn giảm và thông tin phương sai từ thiết bị 310 để tạo ra tín hiệu truyền tải âm thanh. Hơn nữa, thiết bị 320 để tạo ra các kênh đầu ra âm thanh được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh phụ thuộc từ tín hiệu truyền tải âm thanh phụ thuộc vào thông tin trộn giảm và phụ thuộc vào thông tin phương sai.

Theo các phương án, tính năng của hệ thống SAOC, là hệ thống hướng đối tượng mà nhận dạng mã hóa đối tượng, được mở rộng sao cho các đối tượng âm thanh (mã hóa đối tượng) và các kênh âm thanh (mã hóa kênh) hoặc cả hai kênh âm thanh và đối tượng âm thanh (mã hóa trộn) có thể được mã hóa.

Bộ mã hóa SAOC 800 ở Fig.6 và Fig.8 nêu trên được tăng cường, sao cho không chỉ có thể nhận các đối tượng âm thanh làm đầu vào, mà còn có thể nhận các kênh âm thanh làm đầu vào, và sao cho bộ mã hóa SAOC có thể tạo ra các kênh trộn giảm (ví dụ, các kênh truyền tải SAOC) trong đó các đối tượng âm thanh được nhận và các kênh âm thanh được nhận được mã hóa. Theo các phương án nêu trên, ví dụ, ở Fig.6 và Fig.8, bộ mã hóa SAOC 800 này nhận không chỉ các đối tượng âm thanh mà còn các kênh âm thanh là đầu vào và tạo ra các kênh trộn giảm (ví dụ, các kênh truyền tải SAOC) trong đó các đối tượng âm thanh được nhận và các kênh âm thanh được nhận được mã hóa. Ví dụ, bộ mã hóa SAOC ở Fig.6 và Fig.8 được thực hiện là thiết bị để tạo ra tín hiệu truyền tải âm thanh (bao gồm một hoặc nhiều kênh truyền tải âm thanh, ví dụ, một hoặc nhiều kênh truyền tải SAOC) như được mô tả dựa trên Fig.2, và các phương án ở Fig.6 và Fig.8 được biến đổi sao cho không chỉ các đối tượng mà còn một, một số hoặc tất cả các kênh được nạp vào bộ mã hóa SAOC 800.

Bộ giải mã SAOC 1800 ở Fig.7 và Fig.9 nêu trên được tăng cường, sao cho nó có thể nhận các kênh trộn giảm (ví dụ, các kênh truyền tải SAOC) trong đó các đối tượng âm thanh và các kênh âm thanh được mã hóa, và sao cho nó có thể tạo ra các kênh đầu ra (các tín hiệu kênh được kết xuất và các tín hiệu đối tượng được kết xuất) từ các kênh trộn giảm được nhận (ví dụ, các kênh truyền tải SAOC) trong đó các đối tượng âm thanh và các kênh âm thanh được mã hóa. Theo các phương án nêu trên, ví dụ, ở Fig.7 và Fig.9, bộ giải mã SAOC 1800 này nhận các kênh trộn giảm (ví dụ, các kênh truyền tải SAOC) trong đó không chỉ các đối tượng âm thanh mà còn các kênh

âm thanh được mã hóa và tạo ra các kênh đầu ra (các tín hiệu kênh được kết xuất và các tín hiệu đối tượng được kết xuất) từ các kênh trộn giảm được nhận (ví dụ, các kênh truyền tải SAOC) trong đó các đối tượng âm thanh và các kênh âm thanh được mã hóa. Ví dụ, bộ giải mã SAOC ở Fig.7 và Fig.9 được thực hiện là thiết bị để tạo ra một hoặc nhiều kênh đầu ra âm thanh như được mô tả dựa trên Fig.1, và các phương án của Fig.7 và Fig.9 được biến đổi sao cho một, một số hoặc tất cả các kênh được minh họa giữa bộ giải mã USAC 1300 và bộ trộn 1220 không được tạo ra (khôi phục) bởi bộ giải mã USAC 1300, nhưng thay vào đó được khôi phục bởi bộ giải mã SAOC 1800 từ các kênh truyền tải SAOC (các kênh truyền tải âm thanh).

Phụ thuộc vào ứng dụng, các ưu điểm khác nhau của hệ thống SAOC có thể được khai thác bằng cách sử dụng hệ thống SAOC được tăng cường.

Theo một số phương án, hệ thống SAOC được tăng cường này hỗ trợ số kênh trộn giảm duy nhất và kết xuất đến số kênh đầu ra duy nhất. Theo một số phương án, ví dụ, số lượng kênh trộn giảm (các kênh truyền tải SAOC) có thể bị giảm (ví dụ, ở thời điểm chạy), ví dụ, để giảm bớt tốc độ bit tổng đáng kể. Việc này dẫn đến tốc độ bit thấp.

Hơn nữa, theo một số phương án, bộ giải mã SAOC của hệ thống SAOC được tăng cường này có thể, ví dụ, có bộ kết xuất linh hoạt tích hợp mà có thể, ví dụ, cho phép sự tương tác người dùng. Bằng cách này, người dùng có thể thay đổi vị trí các đối tượng trong kịch bản âm thanh, làm giảm hoặc tăng mức đối tượng riêng biệt, hoàn toàn ức chế các đối tượng, v.v.. Ví dụ, xem xét các tín hiệu kênh là các đối tượng nền (BGO) và các tín hiệu đối tượng là các đối tượng cận cảnh (FGO), dấu hiệu tương tác của SAOC có thể được sử dụng cho các ứng dụng như tăng cường hội thoại. Bằng dấu hiệu tương tác này, người dùng có thể tự do để điều khiển, trong khoảng giới hạn, các BGO và FGO, để làm tăng tính đọc được hội thoại (ví dụ, hội thoại có thể được biểu diễn bởi các đối tượng cận cảnh) hoặc để thu được sự cân bằng giữa hội thoại (ví dụ, được biểu diễn bởi FGO) và nền xung quanh (ví dụ, được biểu diễn bởi các BGO).

Hơn nữa, theo các phương án, phụ thuộc vào độ phức tạp tính toán có sẵn ở phía bộ giải mã, bộ giải mã SAOC có thể tự động giảm bớt độ phức tạp tính toán bằng cách vận hành ở chế độ “độ phức tạp tính toán thấp”, ví dụ, bằng cách giảm số bộ giải tương quan, và/hoặc, ví dụ, bằng cách kết xuất trực tiếp đến thiết kế bố trí tái tạo và

bất hoạt bộ chuyển đổi định dạng sau 1720 mà đã được mô tả ở trên. Ví dụ, thông tin kết xuất có thể hướng đến đối với cách mà trộn giảm các kênh của hệ thống 22.2 đến các kênh của hệ thống 5.1.

Theo các phương án, bộ mã hóa SAOC tăng cường có thể xử lý số kênh đầu vào biến đổi ($N_{Channels}$) và các đối tượng đầu vào ($N_{Objects}$). Số kênh và đối tượng được truyền trong dòng bit để truyền tín hiệu đến phía bộ giải mã với sự có mặt của đường dẫn kênh. Các tín hiệu đầu vào đến bộ mã hóa SAOC luôn theo thứ tự sao cho các tín hiệu kênh là các tín hiệu đầu tiên và các tín hiệu đối tượng là các tín hiệu cuối cùng.

Theo các phương án khác, bộ trộn kênh/đối tượng 210 được tạo cấu hình để tạo ra tín hiệu truyền tải âm thanh sao cho số lượng của một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh phụ thuộc vào có tốc độ bit bao nhiêu là khả dụng để truyền tín hiệu truyền tải âm thanh.

Ví dụ, số kênh (vận chuyển) trộn giảm có thể, ví dụ, được tính toán là chức năng của tốc độ bit khả dụng và tổng số tín hiệu đầu vào:

$$N_{DmxCh} = f(\text{tốc độ bit}, N).$$

Các hệ số trộn giảm trong **D** xác định sự trộn các tín hiệu đầu vào (các kênh và đối tượng). Phụ thuộc vào ứng dụng, cấu trúc của ma trận **D** có thể được định rõ sao cho các kênh và đối tượng được trộn với nhau hoặc giữ riêng biệt.

Một số phương án, dựa trên phát hiện ra rằng có lợi khi không trộn các đối tượng cùng với các kênh. Để không trộn các đối tượng cùng với các kênh, ma trận trộn giảm có thể, ví dụ, được tạo cấu trúc là:

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_{ch} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{obj} \end{bmatrix}$$

Để truyền tín hiệu trộn riêng biệt vào dòng bit, các trị số của số kênh trộn giảm được chuyển vào đường dẫn kênh (N_{DmxCh}^{ch}) và số kênh trộn giảm được chuyển đến đường dẫn đối tượng (N_{DmxCh}^{obj}) có thể, ví dụ, được truyền.

Các ma trận trộn giảm theo khối $\mathbf{D}_{ch}$ và $\mathbf{D}_{obj}$ lần lượt có các kích cỡ: $N_{DmxCh}^{ch} \times N_{Channels}$ và $N_{DmxCh}^{obj} \times N_{Objects}$.

Tại bộ giải mã, các hệ số của ma trận ước tính nguồn tham số $\mathbf{G} \approx \mathbf{E}_X \mathbf{D}^H (\mathbf{D} \mathbf{E}_X \mathbf{D}^H)^{-1}$ được tính toán theo phương pháp khác. Sử dụng dạng ma trận, ma trận này có thể có biểu thức như sau:

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{ch} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_{obj} \end{bmatrix}$$

với:

- $\mathbf{G}_{ch} \approx \mathbf{E}_X^{ch} \mathbf{D}_{ch}^H (\mathbf{D}_{ch} \mathbf{E}_X^{ch} \mathbf{D}_{ch}^H)^{-1}$ có kích cỡ $N_{Channels} \times N_{DmxCh}^{ch}$
- $\mathbf{G}_{obj} \approx \mathbf{E}_X^{obj} \mathbf{D}_{obj}^H (\mathbf{D}_{obj} \mathbf{E}_X^{obj} \mathbf{D}_{obj}^H)^{-1}$ có kích cỡ $N_{Objects} \times N_{DmxCh}^{obj}$

Các trị số của phương sai tín hiệu của các kênh ($\mathbf{E}_X^{ch}$) và phương sai tín hiệu đối tượng ($\mathbf{E}_X^{obj}$) có thể, ví dụ, thu được từ ma trận phương sai tín hiệu đầu vào ($\mathbf{E}_X$) bằng cách chỉ lựa chọn các khối đường chéo tương ứng:

$$\mathbf{E}_X = \begin{bmatrix} \mathbf{E}_X^{ch} & \mathbf{E}_X^{ch,obj} \\ \mathbf{E}_X^{obj,ch} & \mathbf{E}_X^{obj} \end{bmatrix}$$

Là kết quả trực tiếp, tốc độ bit bị giảm bằng cách không gửi thông tin bổ sung (ví dụ, OLD, IOC) để khôi phục ma trận phương sai chéo giữa các kênh và đối tượng:

$$\mathbf{E}_X^{ch,obj} = (\mathbf{E}_X^{obj,ch})^H.$$

Theo một số phương án, $\mathbf{E}_X^{ch,obj} = (\mathbf{E}_X^{obj,ch})^H = 0$, và do đó:

$$\mathbf{E}_X = \begin{bmatrix} \mathbf{E}_X^{ch} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_X^{obj} \end{bmatrix}.$$

Theo phương án, bộ mã hóa SAOC tăng cường được tạo cấu hình để không truyền thông tin ở phương sai giữa một trong số đối tượng âm thanh bất kỳ đến bộ giải mã SAOC được tăng cường.

Hơn nữa, theo phương án, bộ giải mã SAOC được tăng cường được tạo cấu hình để không nhận thông tin ở phương sai giữa một trong số các đối tượng âm thanh bất kỳ và một trong số các kênh âm thanh bất kỳ.

Các thành phần theo khối đường chéo phụ của $\mathbf{G}$ không được tính toán, nhưng được thiết lập về không. Do đó, tránh được nhiễu xuyên âm có khả năng giữa các kênh và đối tượng được khôi phục. Hơn nữa, bằng cách này, việc làm giảm độ phức tạp tính toán đạt được do các hệ số nhỏ hơn $\mathbf{G}$ phải được tính toán.

Hơn nữa, theo các phương án, thay vì nghịch đảo ma trận lớn hơn:

$$\mathbf{D} \mathbf{E}_X \mathbf{D}^H \text{ có kích cỡ } [N_{DmxCh}^{ch} + N_{DmxCh}^{obj}] \times [N_{DmxCh}^{ch} + N_{DmxCh}^{obj}] ,$$

Hai ma trận nhỏ sau đây được lấy nghịch đảo:

$$\mathbf{D}_{ch} \mathbf{E}_X^{ch} \mathbf{D}_{ch}^H \text{ có kích cỡ } N_{DmxCh}^{ch} \times N_{DmxCh}^{ch}$$

$$\mathbf{D}_{obj} \mathbf{E}_X^{obj} \mathbf{D}_{obj}^H \text{ có kích cỡ } N_{DmxCh}^{obj} \times N_{DmxCh}^{obj}$$

Việc nghịch đảo các ma trận nhỏ hơn $\mathbf{D}_{ch} \mathbf{E}_X^{ch} \mathbf{D}_{ch}^H$ và $\mathbf{D}_{obj} \mathbf{E}_X^{obj} \mathbf{D}_{obj}^H$ rẻ hơn nhiều đối với độ phức tạp tính toán so với việc lấy nghịch đảo ma trận lớn hơn $\mathbf{D} \mathbf{E}_X \mathbf{D}^H$.

Hơn nữa, bằng cách lấy nghịch đảo các ma trận riêng biệt $\mathbf{D}_{ch} \mathbf{E}_X^{ch} \mathbf{D}_{ch}^H$ và $\mathbf{D}_{obj} \mathbf{E}_X^{obj} \mathbf{D}_{obj}^H$, tính mất ổn định số học có thể có bị giảm so với việc nghịch đảo ma trận lớn hơn $\mathbf{D} \mathbf{E}_X \mathbf{D}^H$. Ví dụ, trong cảnh trường hợp xấu hơn, khi các ma trận phương sai của các kênh truyền tải $\mathbf{D}_{ch} \mathbf{E}_X^{ch} \mathbf{D}_{ch}^H$ và $\mathbf{D}_{obj} \mathbf{E}_X^{obj} \mathbf{D}_{obj}^H$ có sự phụ thuộc tuyến tính do tính tương tự tín hiệu, ma trận đầy đủ $\mathbf{D} \mathbf{E}_X \mathbf{D}^H$ có thể là ma trận điều kiện xấu trong khi các ma trận riêng biệt nhỏ hơn tốt.

Sau khi

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{ch} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_{obj} \end{bmatrix}$$

được tính toán ở phía bộ giải mã, thì có thể, ví dụ, ước tính tham số các tín hiệu đầu vào để thu được các tín hiệu đầu vào được khôi phục $\hat{\mathbf{X}}$ (các tín hiệu kênh âm thanh đầu vào và các tín hiệu đối tượng âm thanh đầu vào), ví dụ, sử dụng:

$$\hat{\mathbf{X}} = \mathbf{G} \mathbf{Y} .$$

Hơn nữa, như nêu trên, việc kết xuất có thể được tiến hành ở phía bộ giải mã để thu được các kênh đầu ra $\mathbf{Z}$, ví dụ, bằng cách sử dụng ma trận kết xuất $\mathbf{R}$:

$$\mathbf{Z} = \mathbf{R}\hat{\mathbf{X}}$$

$$\mathbf{Z} = \mathbf{R}\mathbf{G}\mathbf{Y}$$

$$\mathbf{Z} = \mathbf{S}\mathbf{Y} ; \text{ với } \mathbf{S} = \mathbf{R}\mathbf{G}$$

Thay vì khôi phục rõ ràng các tín hiệu đầu vào (các tín hiệu kênh âm thanh đầu vào và các tín hiệu đối tượng âm thanh đầu vào) để thu được các kênh đầu vào được khôi phục $\hat{\mathbf{X}}$, các kênh đầu ra $\mathbf{Z}$ có thể được tạo ra trực tiếp ở phía bộ giải mã bằng cách áp dụng ma trận tạo ra kênh đầu ra $\mathbf{S}$ ở tín hiệu âm thanh trộn giảm $\mathbf{Y}$.

Như đã nêu trên, để thu được ma trận tạo ra kênh đầu ra $\mathbf{S}$, ma trận kết xuất $\mathbf{R}$ có thể, ví dụ, được xác định hoặc có thể, ví dụ, đã có sẵn. Hơn nữa, ma trận ước tính nguồn tham số $\mathbf{G}$ có thể, ví dụ, được tính toán như nêu trên. Sau đó, ma trận tạo ra kênh đầu ra $\mathbf{S}$ có thể thu được là sản phẩm ma trận $\mathbf{S} = \mathbf{R}\mathbf{G}$ từ ma trận kết xuất $\mathbf{R}$ và ma trận ước tính nguồn tham số $\mathbf{G}$.

Liên quan đến các tín hiệu đối tượng âm thanh được khôi phục, siêu dữ liệu được nén ở các đối tượng âm thanh được truyền từ bộ mã hóa đến bộ giải mã có thể được xem xét. Ví dụ, siêu dữ liệu ở các đối tượng âm thanh có thể biểu thị thông tin vị trí ở mỗi trong số các đối tượng âm thanh. Thông tin vị trí này có thể là, ví dụ, góc phương vị, góc nâng và bán kính. Thông tin vị trí này có thể biểu thị vị trí của đối tượng âm thanh trong không gian ba chiều. Ví dụ, khi đối tượng âm thanh được đặt gần với vị trí loa phóng thanh thực hoặc được xem xét, đối tượng âm thanh này có trọng số cao hơn ở kênh đầu ra đối với loa phóng thanh này so với trọng số của đối tượng âm thanh khác ở kênh đầu ra được đặt ra xa với loa phóng thanh này. Ví dụ, ghép nhẵn biên độ cơ sở vector (vector base amplitude panning - VBAP) có thể được sử dụng (xem, ví dụ, [VBAP]) để xác định các hệ số kết xuất của ma trận kết xuất $\mathbf{R}$ đối với các đối tượng âm thanh.

Hơn nữa, theo một số phương án, siêu dữ liệu nén có thể bao gồm trị số khuếch đại đối với mỗi trong số các đối tượng âm thanh. Ví dụ, đối với mỗi trong số tín hiệu đối tượng âm thanh, trị số khuếch đại có thể biểu thị hệ số khuếch đại đối với tín hiệu đối tượng âm thanh này.

Ngược với các đối tượng âm thanh, không có siêu dữ liệu thông tin vị trí được truyền từ bộ mã hóa đến bộ giải mã đối với các tín hiệu kênh âm thanh. Ma trận bỏ

sung (ví dụ, để chuyển đổi 22.2 thành 5.1) hoặc ma trận đơn vị (khi cấu hình đầu vào của các kênh bằng cấu hình đầu ra) có thể, ví dụ, được sử dụng để xác định các hệ số kết xuất của ma trận kết xuất **R** đối với kênh âm thanh.

Ma trận kết xuất **R** có thể có kích cỡ $N_{OutputChannels} \times N$. Ở đây, đối với mỗi trong số kênh đầu ra, một dòng tồn tại trong ma trận **R**. Hơn nữa, ở mỗi dòng của ma trận kết xuất **R**, N hệ số xác định trọng số của N tín hiệu đầu vào (các kênh âm thanh đầu vào và các đối tượng âm thanh đầu vào) trong kênh đầu ra tương ứng. Các đối tượng âm thanh này được đặt gần loa phóng thanh của kênh đầu ra này có hệ số lớn hơn so với hệ số của các đối tượng âm thanh được xa khỏi loa phóng thanh của kênh đầu ra tương ứng.

Ví dụ, việc ghép nhẵn biên độ cơ sở vector (VBAP) có thể được sử dụng (xem, ví dụ, [VBAP]) để xác định trọng số của tín hiệu đối tượng âm thanh trong mỗi kênh của các loa phóng thanh. Ví dụ, đối với VBAP, cần xem xét rằng đối tượng âm thanh liên quan đến nguồn ảo.

Ngược với các đối tượng âm thanh, do các kênh âm không có vị trí, các hệ số liên quan đến các kênh âm thanh trong ma trận kết xuất có thể, ví dụ, độc lập với thông tin vị trí.

Sau đó, cú pháp dòng bit, theo các phương án, được mô tả.

Trong ngữ cảnh của MPEG SAOC, việc báo hiệu của các chế độ hoạt động có khả năng (chế độ dựa trên kênh, dựa trên đối tượng, kết hợp) có thể đạt được bằng cách sử dụng, ví dụ, một trong hai khả năng sau đây (khả năng thứ nhất: sử dụng các cờ để truyền tín hiệu chế độ hoạt động; khả năng thứ hai: không sử dụng các cờ để truyền tín hiệu chế độ hoạt động):

Do đó, theo phương án thứ nhất, các cờ được sử dụng để truyền tín hiệu chế độ hoạt động.

Để sử dụng các cờ để truyền tín hiệu chế độ hoạt động, cú pháp của thành phần SAOCSpecifigConfig() hoặc thành phần SAOC3DSpecifigConfig() có thể, ví dụ, bao gồm:

bsSaocChannelFlag;

1 uimsbf

```

NumInputSignals = 0;

bsSaocCombinedModeFlag = 0;

if (bsSaocChannelFlag) {

    bsNumSaocChannels;                5    uimbsf

    bsNumSaocDmxChannels;             5    uimbsf

    NumInputSignals += bsNumSaocChannels + 1;

}

bsSaocObjectFlag;                    1    uimbsf

if (bsSaocObjectFlag) {

    bsNumSaocObjects;                 7    uimbsf

    bsNumSaocDmxObjects;              5    uimbsf

    bsSaocCombinedModeFlag;           1

    uimbsfNumInputSignals += bsNumSaocObjects + 1;

}

for ( i=0; i< bsNumSaocChannels+1; i++ ) {

    bsRelatedTo[i][i] = 1;

    for ( j=i+1; j< bsNumSaocChannels+1; j++ ) {

        bsRelatedTo[i][j];            1    uimbsf

        bsRelatedTo[j][i] = bsRelatedTo[i][j];

    }

}

for ( i= bsNumSaocChannels+1; i< bs NumInputSignals; i++ ) {

    cho( j=0; j< bsNumSaocChannels+1; j++ ) {

        bsRelatedTo[i][j] = 0

        bsRelatedTo[j][i] = 0

```

```

    }
}

for ( i= bsNumSaocChannels+1; i< bs NumInputSignals; i++ ) {

    bsRelatedTo[i][i] = 1;

    for ( j=i+1; j< NumInputSignals; j++ ) {

        bsRelatedTo[i][j];                1      uimbsf

        bsRelatedTo[j][i] = bsRelatedTo[i][j];

    }

}

```

Nếu biến dòng bit **bsSaocChannelFlag** được thiết lập về 1, các tín hiệu đầu vào **bsNumSaocChannels+1** thứ nhất được xử lý các tín hiệu dựa trên kênh. Nếu biến dòng bit **bsSaocObjectFlag** được thiết lập về 1, các tín hiệu đầu vào **bsNumSaocObjects+1** cuối cùng được xử lý giống như các tín hiệu đối tượng. Do đó, trong trường hợp cả hai biến dòng bit (**bsSaocChannelFlag**, **bsSaocObjectFlag**) khác không, sự có mặt của các kênh và đối tượng vào các kênh truyền tải âm thanh được báo hiệu.

Nếu biến dòng bit **bsSaocCombinedModeFlag** bằng 1 thì chế độ giải mã tổ hợp được truyền tín hiệu đến dòng bit và, bộ giải mã sẽ xử lý các kênh truyền tải **bsNumSaocDmxChannels** bằng cách sử dụng ma trận trộn giảm đầy đủ **D** (điều này tức là, các tín hiệu kênh và tín hiệu đối tượng được trộn với nhau).

Nếu biến dòng bit **bsSaocCombinedModeFlag** là 0 thì chế độ giải mã độc lập được truyền tín hiệu và bộ giải mã sẽ xử lý (**bsNumSaocDmxChannels+1**) + (**bsNumSaocDmxObjects+1**) các kênh truyền tải sử dụng ma trận trộn giảm theo khối như nêu trên.

Theo phương án thứ hai được ưu tiên, không cần chờ để truyền tín hiệu chế độ hoạt động.

Việc truyền tín hiệu chế độ hoạt động mà không sử dụng các cờ, có thể, ví dụ, được nhận ra bằng cách sử dụng ma trận sau đây

Báo hiệu:

Cú pháp của SAOC3DSpecificConfig():

bsNumSaocDmxChannels;	5	uimsbf
bsNumSaocDmxObjects;	5	uimsbf
NumInputSignals = 0;		
if (bsNumSaocDmxChannels > 0) {		
bsNumSaocChannels;	6	uimsbf
bsNumSaocLFES;	2	uimsbf
NumInputSignals += bsNumSaocChannels;		
}		
bsNumSaocObjects;	8	uimsbf
NumInputSignals += bsNumSaocObjects;		

Giới hạn sự tương quan chéo giữa các kênh và đối tượng về 0:

for (i=0; i<bsNumSaocChannels; i++) {		
bsRelatedTo[i][i] = 1;		
for (j=i+1; j< bsNumSaocChannels; j++) {		
bsRelatedTo[i][j];	1	uimsbf
bsRelatedTo[j][i] = bsRelatedTo[i][j];		
}		
}		
for (i=bsNumSaocChannels; i<NumInputSignals; i++) {		
for (j=0; j<bsNumSaocChannels; j++) {		
bsRelatedTo[i][j] = 0;		
bsRelatedTo[j][i] = 0;		
}		


```

}

for ( i=bsNumSaocChannels; i<NumInputSignals; i++ ) {

    bsRelatedTo[i][i] = 1;

    for ( j=i+1; j<NumInputSignals; j++ ) {

        bsRelatedTo[i][j];                                1           uimsbf

        bsRelatedTo[j][i] = bsRelatedTo[i][j];

    }

}

```

Độc độ khuếch đại trộn giảm khác nhau với các trường hợp khi các kênh âm thanh và đối tượng âm thanh được trộn ở các kênh truyền tải âm thanh khác nhau và khi chúng được trộn với nhau trong các kênh truyền tải âm thanh:

```

if (bsNumSaocDmxObjects==0) {

    for ( i=0; i< bsNumSaocDmxChannels; i++ ) {

        idxDMG[i] = EcDataSaoc(DMG, 0, NumInputSignals);

    }

} else {

    dmgIdx = 0;

    for ( i=0; i<bsNumSaocDmxChannels; i++ ) {

        idxDMG[i] = EcDataSaoc(DMG, 0, bsNumSaocChannels);

    }

    dmgIdx = bsNumSaocDmxChannels;

    if (bsSaocDmxMethod == 0) {

        for ( i=dmgIdx; i<dmgIdx + bsNumSaocDmxObjects; i++ )

{

```

```

        idxDMG[i] = EcDataSaoc(DMG, 0,
bsNumSaocObjects);

    }

}

if (bsSaocDmxMethod == 1) {

    for ( i= dmIdx; i<dmIdx + bsNumSaocDmxObjects; i++ )
{

    idxDMG[i] = EcDataSaoc(DMG, 0,
bsNumPremixedChannels);

    }

}

}

```

Nếu biến dòng bit **bsNumSaocChannels** khác không thì các tín hiệu đầu vào **bsNumSaocChannels** thứ nhất được xử lý giống như các tín hiệu dựa trên kênh. Nếu biến dòng bit **bsNumSaocObjects** khác không thì các tín hiệu đầu vào **bsNumSaocObjects** cuối cùng được xử lý giống như các tín hiệu đối tượng. Do đó, trong trường hợp cả hai biến dòng bit khác không, sự có mặt của các kênh và đối tượng vào kênh truyền tải âm thanh được báo hiệu.

Nếu biến dòng bit **bsNumSaocDmxObjects** bằng 0 thì chế độ giải mã tổ hợp được truyền tín hiệu vào dòng bit và, bộ giải mã sẽ xử lý các kênh truyền tải **bsNumSaocDmxChannels** sử dụng ma trận trộn giảm đầy đủ **D** (việc này tức là, các tín hiệu kênh và tín hiệu đối tượng được trộn với nhau).

Nếu biến dòng bit **bsNumSaocDmxObjects** khác 0 thì chế độ giải mã độc lập và bộ giải mã sẽ xử lý các kênh truyền tải **bsNumSaocDmxChannels** + **bsNumSaocDmxObjects** bằng cách sử dụng sự truyền tải ma trận trộn giảm theo khối như nêu trên.

Sau đây, các khía cạnh của việc xử lý trộn giảm, theo phương án, được mô tả:

Tín hiệu đầu ra của bộ xử lý trộn giảm (được biểu diễn trong miền QMF lai) được nạp vào giàn lọc tổng hợp tương ứng như được mô tả trong ISO/IEC 23003-1:2007 thu được đầu ra cuối cùng của bộ giải mã SAOC 3D.

Bộ xử lý tham số 110 ở Fig.1 và bộ xử lý trộn giảm 120 ở Fig.1 có thể được thực hiện làm cụm xử lý kết hợp. Cụm xử lý kết hợp này được minh họa ở Fig.1, trong đó các cụm U và R thực hiện bộ xử lý tham số 110 bằng cách cung cấp thông tin trộn.

Tín hiệu đầu ra $\hat{Y}$ được tính toán từ tín hiệu trộn giảm đa kênh X và tín hiệu đa kênh được giải tương quan X_d là:

$$\hat{Y} = P_{\text{dry}} RUX + P_{\text{wet}} M_{\text{post}} X_d.$$

trong đó U là ma trận không trộn tham số.

Ma trận trộn $P = \begin{pmatrix} P_{\text{dry}} & P_{\text{wet}} \end{pmatrix}$ là ma trận trộn.

Tín hiệu đa kênh được giải tương quan X_d được xác định là

$$X_d = \text{decorrFunc}(M_{\text{pre}} Y_{\text{dry}}).$$

Chế độ giải mã được điều khiển bởi thành phần dòng bit **bsNumSaocDmxObjects:**

bsNumSaocDmxObjects	Chế độ giải mã	Nghĩa
0	Tổ hợp	Các tín hiệu dựa trên kênh đầu ra và tín hiệu dựa trên đối tượng đầu vào được trộn với nhau vào N_{ch} kênh.
≥ 1	Độc lập	Các tín hiệu dựa trên kênh đầu vào được trộn giảm vào N_{ch} kênh. Các tín hiệu dựa trên đối tượng đầu vào được trộn giảm vào N_{obj} kênh.

Trong trường hợp chế độ giải mã được tổ hợp, ma trận không trộn tham số U được đưa ra bởi:

$$\mathbf{U} = \mathbf{E}\mathbf{D}^*\mathbf{J}.$$

Ma trận $\mathbf{J}$ có kích cỡ $N_{\text{dmx}} \times N_{\text{dmx}}$ được đưa ra bởi $\mathbf{J} \approx \mathbf{\Lambda}^{-1}$ với $\mathbf{\Lambda} = \mathbf{E}\mathbf{D}\mathbf{E}^*$.

Trong trường hợp chế độ giải mã độc lập, ma trận không trộn $\mathbf{U}$ được đưa ra bởi:

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_{\text{ch}} & 0 \\ 0 & \mathbf{U}_{\text{obj}} \end{pmatrix},$$

$$\text{Trong đó, } \mathbf{U}_{\text{ch}} = \mathbf{E}_{\text{ch}}\mathbf{D}_{\text{ch}}^*\mathbf{J}_{\text{ch}} \text{ và } \mathbf{U}_{\text{obj}} = \mathbf{E}_{\text{obj}}\mathbf{D}_{\text{obj}}^*\mathbf{J}_{\text{obj}}.$$

Ma trận phương sai dựa trên kênh $\mathbf{E}_{\text{ch}}$ có kích cỡ $N_{\text{ch}} \times N_{\text{ch}}$ và ma trận phương sai dựa trên đối tượng $\mathbf{E}_{\text{obj}}$ có kích cỡ $N_{\text{obj}} \times N_{\text{obj}}$ thu được từ ma trận phương sai $\mathbf{E}$ bằng cách chỉ lựa chọn các khối đường chéo tương ứng:

$$\mathbf{E} = \begin{pmatrix} \mathbf{E}_{\text{ch}} & \mathbf{E}_{\text{ch,obj}} \\ \mathbf{E}_{\text{obj,ch}} & \mathbf{E}_{\text{obj}} \end{pmatrix},$$

trong đó, ma trận $\mathbf{E}_{\text{ch,obj}} = (\mathbf{E}_{\text{obj,ch}})^*$ là ma trận phương sai chéo giữa các kênh đầu vào và đối tượng đầu vào và không yêu cầu được tính toán.

Ma trận trộn giảm dựa trên kênh $\mathbf{D}_{\text{ch}}$ có kích cỡ $N_{\text{ch}}^{\text{dmx}} \times N_{\text{ch}}$ và ma trận trộn giảm dựa trên đối tượng $\mathbf{D}_{\text{obj}}$ có kích cỡ $N_{\text{obj}}^{\text{dmx}} \times N_{\text{obj}}$ thu được từ ma trận trộn giảm $\mathbf{D}$ bằng cách chỉ lựa chọn các khối đường chéo tương ứng:

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_{\text{ch}} & 0 \\ 0 & \mathbf{D}_{\text{obj}} \end{pmatrix}.$$

Ma trận $\mathbf{J}_{\text{ch}} \approx (\mathbf{D}_{\text{ch}}\mathbf{E}_{\text{ch}}\mathbf{D}_{\text{ch}}^*)^{-1}$ có kích cỡ $N_{\text{ch}}^{\text{dmx}} \times N_{\text{ch}}^{\text{dmx}}$ thu được từ việc xác định ma trận $\mathbf{J}$ đối với

$$\mathbf{\Lambda} = \mathbf{D}_{\text{ch}}\mathbf{E}_{\text{ch}}\mathbf{D}_{\text{ch}}^*.$$

Ma trận $\mathbf{J}_{\text{obj}} \approx (\mathbf{D}_{\text{obj}}\mathbf{E}_{\text{obj}}\mathbf{D}_{\text{obj}}^*)^{-1}$ có kích cỡ $N_{\text{obj}}^{\text{dmx}} \times N_{\text{obj}}^{\text{dmx}}$ thu được từ việc xác định ma trận $\mathbf{J}$ đối với

$$\mathbf{\Lambda} = \mathbf{D}_{\text{obj}}\mathbf{E}_{\text{obj}}\mathbf{D}_{\text{obj}}^*$$

Ma trận $\mathbf{J} \approx \mathbf{\Lambda}^{-1}$ được tính toán bằng cách sử dụng công thức sau đây:

$$\mathbf{J} = \mathbf{V} \mathbf{\Lambda}^{inv} \mathbf{V}^*.$$

Ở đây các vectơ đơn $\mathbf{V}$ của ma trận $\mathbf{\Lambda}$ thu được bằng cách sử dụng công thức sau đây

$$\mathbf{V} \mathbf{\Lambda} \mathbf{V}^* = \mathbf{\Lambda}.$$

Sự nghịch đảo theo quy tắc $\mathbf{\Lambda}^{inv}$ của ma trận trị số đơn đường chéo $\mathbf{\Lambda}$ được tính là

$$\lambda_{i,j}^{inv} = \begin{cases} \frac{1}{\lambda_{i,j}}, & \text{nếu } i = j \text{ và } \lambda_{i,j} \geq T_{reg}^{\Lambda} \\ 0, & \text{thì} \end{cases},$$

Vô hướng quy tắc tương đối T_{reg}^{Λ} được xác định bằng cách sử dụng ngưỡng tuyệt đối T_{reg} và trị số tối đa của $\mathbf{\Lambda}$ là

$$T_{reg}^{\Lambda} = \max(\lambda_{i,i}) T_{reg}, \quad T_{reg} = 10^{-2}.$$

Sau đây, ma trận kết xuất theo phương án được mô tả:

Ma trận kết xuất $\mathbf{R}$ được áp dụng với các tín hiệu âm thanh đầu vào $\mathbf{S}$ xác định đầu ra được kết xuất đích là $\mathbf{Y} = \mathbf{R} \mathbf{S}$. Ma trận kết xuất $\mathbf{R}$ có kích cỡ $N_{out} \times N$ được đưa ra bởi

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_{ch} & \mathbf{R}_{obj} \end{pmatrix},$$

trong đó, $\mathbf{R}_{ch}$ có kích cỡ $N_{out} \times N_{ch}$ biểu diễn ma trận kết xuất được kết hợp với các kênh đầu vào và $\mathbf{R}_{obj}$ có kích cỡ $N_{out} \times N_{obj}$ biểu diễn ma trận kết xuất được kết hợp với các đối tượng đầu vào.

Sau đây, tín hiệu đa kênh được giải tương quan $\mathbf{X}_d$ theo phương án được mô tả:

Các tín hiệu được giải tương quan $\mathbf{X}_d$, ví dụ, được tạo ra từ bộ giải tương quan được mô tả trong 6.6.2 của ISO/IEC 23003-1:2007, với $bsDecorConfig = 0$ và, ví dụ, chỉ số tương quan, X . Do đó, $decorrFunc()$ ví dụ, là quy trình giải tương quan:

$$\mathbf{X}_d = decorrFunc(\mathbf{M}_{pre} \mathbf{Y}_{dry}).$$

Mặc dù một số khía cạnh được mô tả ở nội dung 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.

Tín hiệu được phân giải có thể được lưu trữ trên vật ghi lưu trữ số hoặc có thể được truyền trên vật ghi truyền như vật ghi truyền không dây hoặc vật ghi truyền có dây như internet.

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 như vật ghi lưu trữ số, ví dụ, đĩa mềm, DVD, CD, ROM, PROM, và EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều khiển đọc được nhờ 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.

Một số phương án theo sáng chế bao gồm vật mang dữ liệu cố định có các tín hiệu điều khiển đọc được nhờ 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 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 các phương pháp được mô tả ở đây, được lưu trữ trên vật mang đọc được bằng máy.

Nói cách khác, phương án của các phương pháp sáng tạo là, do đó, chương trình máy tính có mã chương trình để thực hiện một trong 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 sáng tạo 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.

Do đó, phương án khác theo phương pháp sáng tạo là dòng dữ liệu hoặc trình tự 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 trình tự tín hiệu có thể, ví dụ, được tạo cấu hình để truyền qua kết nối giao tiếp dữ liệu, ví dụ, qua internet.

Phương án khác bao gồm các phương thức xử lý, ví dụ, máy tính hoặc thiết bị logic lập trình được, được tạo cấu hình, 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.

Theo một số phương án, thiết bị logic lập trình được (ví dụ, mảng cổng lập trình được dạng trường) có thể được sử dụng để thực hiện một vài hoặc tất cả các chức năng của các phương pháp được mô tả ở đây. Theo một số phương án, mảng cổng lập trình được dạng trường có thể kết hợp với bộ vi xử lý để thực hiện một trong các phương pháp được mô tả ở đây. Thông thường, các phương pháp này tốt hơn 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 của các sắp xếp và mô tả chi tiết ở đây sẽ rõ ràng với người có hiểu biết trung bình trong lĩnh vực kỹ thuật. Do đó, sáng chế chỉ bị giới hạn bởi phạm vi của các yêu cầu bảo hộ và không bị giới hạn bởi phần mô tả chi tiết được thể hiện bằng cách mô tả và giải thích các phương án ở đây.

YÊU CẦU BẢO HỘ

1. Thiết bị tạo ra một hoặc nhiều kênh đầu ra âm thanh, trong đó thiết bị bao gồm:

bộ xử lý tham số (110) để tính toán thông tin trộn, và

bộ xử lý trộn giảm (120) để tạo ra một hoặc nhiều kênh đầu ra âm thanh,

trong đó, bộ xử lý trộn giảm (120) được tạo cấu hình để nhận dòng dữ liệu bao gồm các kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh, trong đó một hoặc nhiều tín hiệu kênh âm thanh được trộn trong tín hiệu truyền tải âm thanh, trong đó một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong tín hiệu truyền tải âm thanh, và trong đó số lượng của các kênh truyền tải âm thanh nhỏ hơn số lượng của một hoặc nhiều tín hiệu kênh âm thanh cộng với số lượng của một hoặc nhiều tín hiệu đối tượng âm thanh,

trong đó bộ xử lý tham số (110) được tạo cấu hình để nhận thông tin trộn giảm biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong các kênh truyền tải âm thanh, và trong đó bộ xử lý tham số (110) được tạo cấu hình để nhận thông tin phương sai, và trong đó, bộ xử lý tham số (110) được tạo cấu hình để tính toán thông tin trộn phụ thuộc vào thông tin trộn giảm và phụ thuộc vào thông tin phương sai, và

trong đó, bộ xử lý trộn giảm (120) được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu truyền tải âm thanh phụ thuộc vào thông tin trộn,

trong đó, thông tin phương sai biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu kênh âm thanh và còn biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh, và trong đó thông tin phương sai không biểu thị thông tin tương quan đối với cặp bất kỳ của một trong số một hoặc nhiều tín hiệu kênh âm thanh và một trong số một hoặc nhiều tín hiệu đối tượng âm thanh,

trong đó một hoặc nhiều tín hiệu kênh âm thanh được trộn trong nhóm thứ nhất của một hoặc nhiều kênh truyền tải âm thanh trong số các kênh truyền tải âm thanh, trong đó một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong nhóm thứ hai của một hoặc nhiều kênh truyền tải âm thanh trong số các kênh truyền tải âm thanh, trong đó mỗi kênh truyền tải âm thanh của nhóm thứ nhất không được chứa bởi nhóm thứ

hai, và trong đó mỗi kênh truyền tải âm thanh của nhóm thứ hai không được chứa bởi nhóm thứ nhất, và

trong đó thông tin trộn giảm bao gồm thông tin con trộn giảm thứ nhất biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh được trộn trong nhóm thứ nhất của các kênh truyền tải âm thanh, và trong đó thông tin trộn giảm bao gồm thông tin con trộn giảm thứ hai, phụ thuộc vào biểu thị thông tin về cách mà một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong nhóm thứ hai của một hoặc nhiều kênh truyền tải âm thanh,

trong đó bộ xử lý tham số (110) được tạo cấu hình để tính toán thông tin trộn phụ thuộc vào thông tin con trộn giảm thứ nhất, phụ thuộc vào thông tin con trộn giảm thứ hai và phụ thuộc vào thông tin phương sai,

trong đó bộ xử lý trộn giảm (120) được tạo cấu hình để tạo ra một hoặc nhiều tín hiệu đầu ra âm thanh từ nhóm thứ nhất của các kênh truyền tải âm thanh và từ nhóm thứ hai của các kênh truyền tải âm thanh phụ thuộc vào thông tin trộn,

trong đó bộ xử lý trộn giảm (120) được tạo cấu hình để nhận số đếm kênh thứ nhất biểu thị số lượng của các kênh truyền tải âm thanh của nhóm thứ nhất của các kênh truyền tải âm thanh, và trong đó bộ xử lý trộn giảm (120) được tạo cấu hình để nhận số đếm kênh thứ hai biểu thị số lượng của các kênh truyền tải âm thanh của nhóm thứ hai của các kênh truyền tải âm thanh, và

trong đó bộ xử lý trộn giảm (120) được tạo cấu hình để nhận biết liệu kênh truyền tải âm thanh trong dòng dữ liệu thuộc về nhóm thứ nhất hoặc nhóm thứ hai phụ thuộc vào số đếm kênh thứ nhất hoặc phụ thuộc vào số đếm kênh thứ hai, hoặc phụ thuộc vào số đếm kênh thứ nhất và số đếm kênh thứ hai.

2. Thiết bị theo điểm 1, trong đó thông tin phương sai biểu thị thông tin chênh lệch mức đối với mỗi tín hiệu kênh âm thanh trong số một hoặc nhiều tín hiệu kênh âm thanh và còn biểu thị thông tin chênh lệch mức đối với mỗi tín hiệu đối tượng âm thanh trong số một hoặc nhiều tín hiệu đối tượng âm thanh.

3. Thiết bị theo điểm 1 hoặc điểm 2,

trong đó, hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được trộn trong tín hiệu truyền tải âm thanh, và trong đó hai hoặc nhiều hơn hai tín hiệu kênh âm thanh được trộn trong tín hiệu truyền tải âm thanh,

trong đó thông tin phương sai biểu thị thông tin tương quan đối với một hoặc nhiều cặp của một tín hiệu kênh âm thanh thứ nhất trong số hai hoặc nhiều hơn hai tín hiệu kênh âm thanh và một tín hiệu kênh âm thanh thứ hai trong số hai hoặc nhiều hơn hai tín hiệu kênh âm thanh, hoặc

trong đó, thông tin phương sai biểu thị thông tin tương quan đối với một hoặc nhiều cặp của một tín hiệu đối tượng âm thanh thứ nhất trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh và một tín hiệu đối tượng âm thanh thứ hai trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, hoặc

trong đó, thông tin phương sai biểu thị thông tin tương quan đối với một hoặc nhiều cặp của một tín hiệu kênh âm thanh thứ nhất trong số hai hoặc nhiều hơn hai tín hiệu kênh âm thanh và một tín hiệu kênh âm thanh thứ hai trong số hai hoặc nhiều hơn hai tín hiệu kênh âm thanh và biểu thị thông tin tương quan đối với một hoặc nhiều cặp của một tín hiệu đối tượng âm thanh thứ nhất trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh và một tín hiệu đối tượng âm thanh thứ hai trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

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

trong đó, thông tin phương sai bao gồm nhiều hệ số phương sai của ma trận phương sai $\mathbf{E}_X$ có kích cỡ $N \times N$, trong đó N biểu thị số lượng của một hoặc nhiều tín hiệu kênh âm thanh cộng với số lượng của một hoặc nhiều tín hiệu đối tượng âm thanh,

trong đó, ma trận phương sai $\mathbf{E}_X$ được xác định theo công thức

$$\mathbf{E}_X = \begin{bmatrix} \mathbf{E}_X^{ch} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_X^{obj} \end{bmatrix},$$

trong đó, $\mathbf{E}_X^{ch}$ biểu thị các hệ số của ma trận con phương sai thứ nhất có kích cỡ $N_{Channels} \times N_{Channels}$, trong đó $N_{Channels}$ biểu thị số lượng của một hoặc nhiều tín hiệu kênh âm thanh,

trong đó, E_X^{obj} biểu thị các hệ số của ma trận con phương sai thứ hai có kích cỡ $N_{Objects} \times N_{Objects}$, trong đó $N_{Objects}$ biểu thị số lượng của một hoặc nhiều tín hiệu đối tượng âm thanh,

trong đó, $\mathbf{0}$ biểu thị ma trận không,

trong đó, bộ xử lý tham số (110) được tạo cấu hình để nhận nhiều hệ số phương sai của ma trận phương sai E_X , và

trong đó, bộ xử lý tham số (110) được tạo cấu hình để thiết lập tất cả các hệ số của ma trận phương sai E_X về 0, mà không được nhận bởi bộ xử lý tham số (110).

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

trong đó, thông tin trộn giảm bao gồm nhiều hệ số trộn giảm của ma trận trộn giảm $\mathbf{D}$ có kích cỡ $N_{DmxCh} \times N$, trong đó N_{DmxCh} biểu thị số lượng của các kênh truyền tải âm thanh, và trong đó N biểu thị số lượng của một hoặc nhiều tín hiệu kênh âm thanh cộng với số lượng của một hoặc nhiều tín hiệu đối tượng âm thanh,

trong đó, ma trận trộn giảm $\mathbf{D}$ được xác định theo công thức

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_{ch} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{obj} \end{bmatrix},$$

trong đó, $\mathbf{D}_{ch}$ biểu thị các hệ số của ma trận con trộn giảm thứ nhất có kích cỡ $N_{DmxCh}^{ch} \times N_{Channels}$, trong đó N_{DmxCh}^{ch} biểu thị số lượng của các kênh truyền tải âm thanh của nhóm thứ nhất của các kênh truyền tải âm thanh, và trong đó $N_{Channels}$ biểu thị số lượng của một hoặc nhiều tín hiệu kênh âm thanh,

trong đó, $\mathbf{D}_{obj}$ biểu thị các hệ số của ma trận con trộn giảm thứ hai có kích cỡ $N_{DmxCh}^{obj} \times N_{Objects}$, trong đó N_{DmxCh}^{obj} biểu thị số lượng của một hoặc nhiều kênh truyền tải âm thanh của nhóm thứ hai của các kênh truyền tải âm thanh, và trong đó $N_{Objects}$ biểu thị số lượng của một hoặc nhiều tín hiệu kênh âm thanh,

trong đó, $\mathbf{0}$ biểu thị ma trận không,

trong đó, bộ xử lý tham số (110) được tạo cấu hình để nhận nhiều hệ số trộn giảm của ma trận trộn giảm $\mathbf{D}$, và

trong đó, bộ xử lý tham số (110) được tạo cấu hình để thiết lập tất cả các hệ số của ma trận trộn giảm $\mathbf{D}$ về 0, mà không được nhận bởi bộ xử lý tham số (110).

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

trong đó, bộ xử lý tham số (110) được tạo cấu hình để nhận thông tin kết xuất biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong một hoặc nhiều kênh đầu ra âm thanh,

trong đó, bộ xử lý tham số (110) được tạo cấu hình để tính toán thông tin trộn phụ thuộc vào thông tin trộn giảm, phụ thuộc vào thông tin phương sai và phụ thuộc vào thông tin kết xuất.

7. Thiết bị theo điểm 6,

trong đó, bộ xử lý tham số (110) được tạo cấu hình để nhận nhiều hệ số của ma trận kết xuất $\mathbf{R}$ làm thông tin kết xuất, và

trong đó, bộ xử lý tham số (110) được tạo cấu hình để tính toán thông tin trộn phụ thuộc vào thông tin trộn giảm, phụ thuộc vào thông tin phương sai và phụ thuộc vào ma trận kết xuất $\mathbf{R}$.

8. Thiết bị theo điểm 6,

trong đó, bộ xử lý tham số (110) được tạo cấu hình để nhận thông tin siêu dữ liệu như thông tin kết xuất, trong đó thông tin siêu dữ liệu bao gồm thông tin vị trí,

trong đó, thông tin vị trí biểu thị vị trí đối với mỗi tín hiệu đối tượng âm thanh trong số một hoặc nhiều tín hiệu đối tượng âm thanh,

trong đó, thông tin vị trí không biểu thị vị trí đối với tín hiệu kênh âm thanh bất kỳ trong số một hoặc nhiều tín hiệu kênh âm thanh,

trong đó, bộ xử lý tham số (110) được tạo cấu hình để tính toán thông tin trộn phụ thuộc vào thông tin trộn giảm, phụ thuộc vào thông tin phương sai, và phụ thuộc vào thông tin vị trí.

9. Thiết bị theo điểm 8,

trong đó, thông tin siêu dữ liệu còn bao gồm thông tin độ khuếch đại,

trong đó, thông tin độ khuếch đại biểu thị trị số khuếch đại đối với mỗi tín hiệu đối tượng âm thanh trong số một hoặc nhiều tín hiệu đối tượng âm thanh,

trong đó, thông tin độ khuếch đại không biểu thị trị số khuếch đại đối với tín hiệu kênh âm thanh bất kỳ trong số một hoặc nhiều tín hiệu kênh âm thanh,

trong đó, bộ xử lý tham số (110) được tạo cấu hình để tính toán thông tin trộn phụ thuộc vào thông tin trộn giảm, phụ thuộc vào thông tin phương sai, phụ thuộc vào thông tin vị trí, và phụ thuộc vào thông tin độ khuếch đại.

10. Thiết bị theo điểm 8 hoặc điểm 9,

trong đó, bộ xử lý tham số (110) được tạo cấu hình để tính toán ma trận trộn S như thông tin trộn, trong đó ma trận trộn S được xác định theo công thức:

$$S = RG,$$

trong đó, G là ma trận giải mã phụ thuộc vào thông tin trộn giảm và phụ thuộc vào thông tin phương sai,

trong đó, R là ma trận kết xuất phụ thuộc vào thông tin siêu dữ liệu,

trong đó, bộ xử lý trộn giảm (120) được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh của tín hiệu đầu ra âm thanh bằng cách áp dụng công thức

$$Z = SY,$$

trong đó, Z là tín hiệu đầu ra âm thanh, và trong đó Y là tín hiệu truyền tải âm thanh.

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

trong đó, hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được trộn trong tín hiệu truyền tải âm thanh, và trong đó hai hoặc nhiều hơn hai tín hiệu kênh âm thanh được trộn trong tín hiệu truyền tải âm thanh,

trong đó, thông tin phương sai biểu thị thông tin tương quan đối với một hoặc nhiều cặp của một tín hiệu kênh âm thanh thứ nhất trong số hai hoặc nhiều hơn hai tín hiệu kênh âm thanh và một tín hiệu kênh âm thanh thứ hai trong số hai hoặc nhiều hơn hai tín hiệu kênh âm thanh,

trong đó, thông tin phương sai không biểu thị thông tin tương quan đối với cặp bất kỳ của một tín hiệu đối tượng âm thanh thứ nhất trong số một hoặc nhiều tín hiệu đối tượng âm thanh và một tín hiệu đối tượng âm thanh thứ hai trong số một hoặc nhiều tín hiệu đối tượng âm thanh, và

trong đó, bộ xử lý tham số (110) được tạo cấu hình để tính toán thông tin trộn phụ thuộc vào thông tin trộn giảm, phụ thuộc vào thông tin chênh lệch mức của mỗi tín hiệu kênh âm thanh trong số một hoặc nhiều tín hiệu kênh âm thanh, phụ thuộc vào thông tin chênh lệch mức thứ hai của mỗi tín hiệu đối tượng âm thanh trong số một hoặc nhiều tín hiệu đối tượng âm thanh, và phụ thuộc vào thông tin tương quan của một hoặc nhiều cặp của một tín hiệu kênh âm thanh thứ nhất trong số hai hoặc nhiều hơn hai tín hiệu kênh âm thanh và một tín hiệu kênh âm thanh thứ hai trong số hai hoặc nhiều hơn hai tín hiệu kênh âm thanh.

12. Thiết bị tạo ra tín hiệu truyền tải âm thanh bao gồm các kênh truyền tải âm thanh, trong đó thiết bị bao gồm:

bộ trộn kênh/đối tượng (210) để tạo ra các kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh, và

giao diện đầu ra (220),

trong đó, bộ trộn kênh/đối tượng (210) được tạo cấu hình để tạo ra tín hiệu truyền tải âm thanh bao gồm các kênh truyền tải âm thanh bằng cách trộn một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh trong tín hiệu truyền tải âm thanh phụ thuộc vào thông tin trộn giảm biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh phải được trộn trong các kênh truyền tải âm thanh, trong đó số lượng của các kênh truyền tải âm thanh nhỏ hơn số lượng của một hoặc nhiều tín hiệu kênh âm thanh cộng với số lượng của một hoặc nhiều tín hiệu đối tượng âm thanh,

trong đó, giao diện đầu ra (220) được tạo cấu hình để xuất ra tín hiệu truyền tải âm thanh, thông tin trộn giảm và thông tin phương sai,

trong đó, thông tin phương sai biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu kênh âm thanh và còn biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh, và

trong đó, thông tin phương sai không biểu thị thông tin tương quan đối với cặp bất kỳ của một trong số một hoặc nhiều tín hiệu kênh âm thanh và một trong số một hoặc nhiều tín hiệu đối tượng âm thanh,

trong đó thiết bị được tạo cấu hình để trộn một hoặc nhiều tín hiệu kênh âm thanh trong nhóm thứ nhất của một hoặc nhiều kênh truyền tải âm thanh trong số các kênh truyền tải âm thanh, trong đó thiết bị được tạo cấu hình để trộn một hoặc nhiều tín hiệu đối tượng âm thanh trong nhóm thứ hai của một hoặc nhiều kênh truyền tải âm thanh, trong đó mỗi kênh truyền tải âm thanh của nhóm thứ nhất không được chứa bởi nhóm thứ hai, và trong đó mỗi kênh truyền tải âm thanh của nhóm thứ hai không được chứa bởi nhóm thứ nhất, và

trong đó thông tin trộn giảm bao gồm thông tin con trộn giảm thứ nhất biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh được trộn trong nhóm thứ nhất của các kênh truyền tải âm thanh, và trong đó thông tin trộn giảm bao gồm thông tin con trộn giảm thứ hai biểu thị thông tin về cách mà một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong nhóm thứ hai của các kênh truyền tải âm thanh,

trong đó thiết bị được tạo cấu hình để xuất ra số đếm kênh thứ nhất biểu thị số lượng của các kênh truyền tải âm thanh của nhóm thứ nhất của các kênh truyền tải âm thanh, và trong đó thiết bị được tạo cấu hình để xuất ra số đếm kênh thứ hai biểu thị số lượng của các kênh truyền tải âm thanh của nhóm thứ hai của các kênh truyền tải âm thanh.

13. Thiết bị theo điểm 12, trong đó bộ trộn kênh/đối tượng (210) được tạo cấu hình để tạo ra tín hiệu truyền tải âm thanh sao cho số lượng của các kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh phụ thuộc vào tốc độ bit bao nhiêu là khả dụng để truyền tín hiệu truyền tải âm thanh.

14. Hệ thống tạo ra tín hiệu truyền tải âm thanh và một hoặc nhiều kênh đầu ra âm thanh, hệ thống bao gồm:

thiết bị (310) theo điểm 12 hoặc 13 để tạo ra tín hiệu truyền tải âm thanh, và

thiết bị (320) theo điểm bất kỳ trong số các điểm từ 1 đến 11 để tạo ra một hoặc nhiều kênh đầu ra âm thanh,

trong đó, thiết bị (320) theo một trong số các điểm từ 1 đến 11 được tạo cấu hình để nhận tín hiệu truyền tải âm thanh, thông tin trộn giảm và thông tin phương sai từ thiết bị (310) theo điểm 12 hoặc 13, và

trong đó, thiết bị (320) theo điểm bất kỳ trong số các điểm từ 1 đến 11 được tạo cấu hình để tạo ra một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu truyền tải âm thanh phụ thuộc vào thông tin trộn giảm và phụ thuộc vào thông tin phương sai.

15. Phương pháp tạo ra một hoặc nhiều kênh đầu ra âm thanh, trong đó phương pháp bao gồm các bước:

nhận tín dòng dữ liệu bao gồm các kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh, trong đó một hoặc nhiều tín hiệu kênh âm thanh được trộn trong tín hiệu truyền tải âm thanh, trong đó một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong tín hiệu truyền tải âm thanh, và trong đó số lượng của các kênh truyền tải âm thanh nhỏ hơn số lượng của một hoặc nhiều tín hiệu kênh âm thanh cộng với số lượng của một hoặc nhiều tín hiệu đối tượng âm thanh,

nhận thông tin trộn giảm biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong các kênh truyền tải âm thanh,

nhận thông tin phương sai,

tính toán thông tin trộn phụ thuộc vào thông tin trộn giảm và phụ thuộc vào thông tin phương sai, và

tạo ra một hoặc nhiều kênh đầu ra âm thanh,

tạo ra một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu truyền tải âm thanh phụ thuộc vào thông tin trộn,

trong đó, thông tin phương sai biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu kênh âm thanh và còn biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh, và

trong đó, thông tin phương sai không biểu thị thông tin tương quan đối với cặp bất kỳ của một trong số một hoặc nhiều tín hiệu kênh âm thanh và một trong số một hoặc nhiều tín hiệu đối tượng âm thanh,

trong đó một hoặc nhiều tín hiệu kênh âm thanh được trộn trong nhóm thứ nhất của một hoặc nhiều kênh truyền tải âm thanh trong số các kênh truyền tải âm thanh, trong đó một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong nhóm thứ hai của một hoặc nhiều kênh truyền tải âm thanh, trong đó mỗi kênh truyền tải âm thanh của nhóm thứ nhất không được chứa bởi nhóm thứ hai, và trong đó mỗi kênh truyền tải âm thanh của nhóm thứ hai không được chứa bởi nhóm thứ nhất, và

trong đó thông tin trộn giảm bao gồm thông tin con trộn giảm thứ nhất biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh được trộn trong nhóm thứ nhất của các kênh truyền tải âm thanh, và trong đó thông tin trộn giảm bao gồm thông tin con trộn giảm thứ hai biểu thị thông tin về cách mà một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong nhóm thứ hai của các kênh truyền tải âm thanh,

trong đó thông tin trộn được tính toán phụ thuộc vào thông tin con trộn giảm thứ nhất, phụ thuộc vào thông tin con trộn giảm thứ hai và phụ thuộc vào thông tin phương sai,

trong đó một hoặc nhiều tín hiệu đầu ra âm thanh được tạo ra từ nhóm thứ nhất của các kênh truyền tải âm thanh và từ nhóm thứ hai của các kênh truyền tải âm thanh phụ thuộc vào thông tin trộn,

trong đó phương pháp còn bao gồm bước nhận số đếm kênh thứ nhất biểu thị số lượng của các kênh truyền tải âm thanh của nhóm thứ nhất của các kênh truyền tải âm thanh, và trong đó phương pháp còn bao gồm bước nhận số đếm kênh thứ hai biểu thị số lượng của các kênh truyền tải âm thanh của nhóm thứ hai của các kênh truyền tải âm thanh, và

trong đó phương pháp còn bao gồm bước nhận biết liệu kênh truyền tải âm thanh trong dòng dữ liệu thuộc về nhóm thứ nhất hoặc nhóm thứ hai phụ thuộc vào số đếm kênh thứ nhất hoặc phụ thuộc vào số đếm kênh thứ hai, hoặc phụ thuộc vào số đếm kênh thứ nhất và số đếm kênh thứ hai.

16. Phương pháp tạo ra tín hiệu truyền tải âm thanh bao gồm các kênh truyền tải âm thanh, trong đó phương pháp bao gồm các bước:

tạo ra tín hiệu truyền tải âm thanh bao gồm các kênh truyền tải âm thanh bằng cách trộn một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng

âm thanh trong tín hiệu truyền tải âm thanh phụ thuộc vào thông tin trộn giảm biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh và một hoặc nhiều tín hiệu đối tượng âm thanh phải được trộn trong các kênh truyền tải âm thanh, trong đó số lượng của các kênh truyền tải âm thanh nhỏ hơn số lượng của một hoặc nhiều tín hiệu kênh âm thanh cộng với số lượng của một hoặc nhiều tín hiệu đối tượng âm thanh, và

xuất ra tín hiệu truyền tải âm thanh, thông tin trộn giảm và thông tin phương sai,

trong đó, thông tin phương sai biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu kênh âm thanh và còn biểu thị thông tin chênh lệch mức đối với ít nhất một trong số một hoặc nhiều tín hiệu đối tượng âm thanh, và

trong đó thông tin phương sai không biểu thị thông tin tương quan đối với cặp bất kỳ của một trong số một hoặc nhiều tín hiệu kênh âm thanh và một trong số một hoặc nhiều tín hiệu đối tượng âm thanh,

trong đó một hoặc nhiều tín hiệu kênh âm thanh được trộn trong nhóm thứ nhất của một hoặc nhiều kênh truyền tải âm thanh trong số các kênh truyền tải âm thanh, trong đó một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong nhóm thứ hai của một hoặc nhiều kênh truyền tải âm thanh trong số các kênh truyền tải âm thanh, trong đó mỗi kênh truyền tải âm thanh của nhóm thứ nhất không được chứa bởi nhóm thứ hai, và trong đó mỗi kênh truyền tải âm thanh của nhóm thứ hai không được chứa bởi nhóm thứ nhất, và

trong đó thông tin trộn giảm bao gồm thông tin con trộn giảm thứ nhất biểu thị thông tin về cách mà một hoặc nhiều tín hiệu kênh âm thanh được trộn trong nhóm thứ nhất của các kênh truyền tải âm thanh, và trong đó thông tin trộn giảm bao gồm thông tin con trộn giảm thứ hai biểu thị thông tin về cách mà một hoặc nhiều tín hiệu đối tượng âm thanh được trộn trong nhóm thứ hai của các kênh truyền tải âm thanh, và

trong đó phương pháp còn bao gồm việc xuất ra số đếm kênh thứ nhất biểu thị số lượng của các kênh truyền tải âm thanh của nhóm thứ nhất của các kênh truyền tải âm thanh, và trong đó phương pháp còn bao gồm việc xuất ra số đếm kênh thứ hai

biểu thị số lượng của các kênh truyền tải âm thanh của nhóm thứ hai của các kênh truyền tải âm thanh.

17. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính để thực hiện phương pháp theo điểm 15 khi được chạy trên máy tính hoặc bộ xử lý tín hiệu.
18. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính để thực hiện phương pháp theo điểm 16 khi được chạy trên máy tính hoặc bộ xử lý tín hiệu.

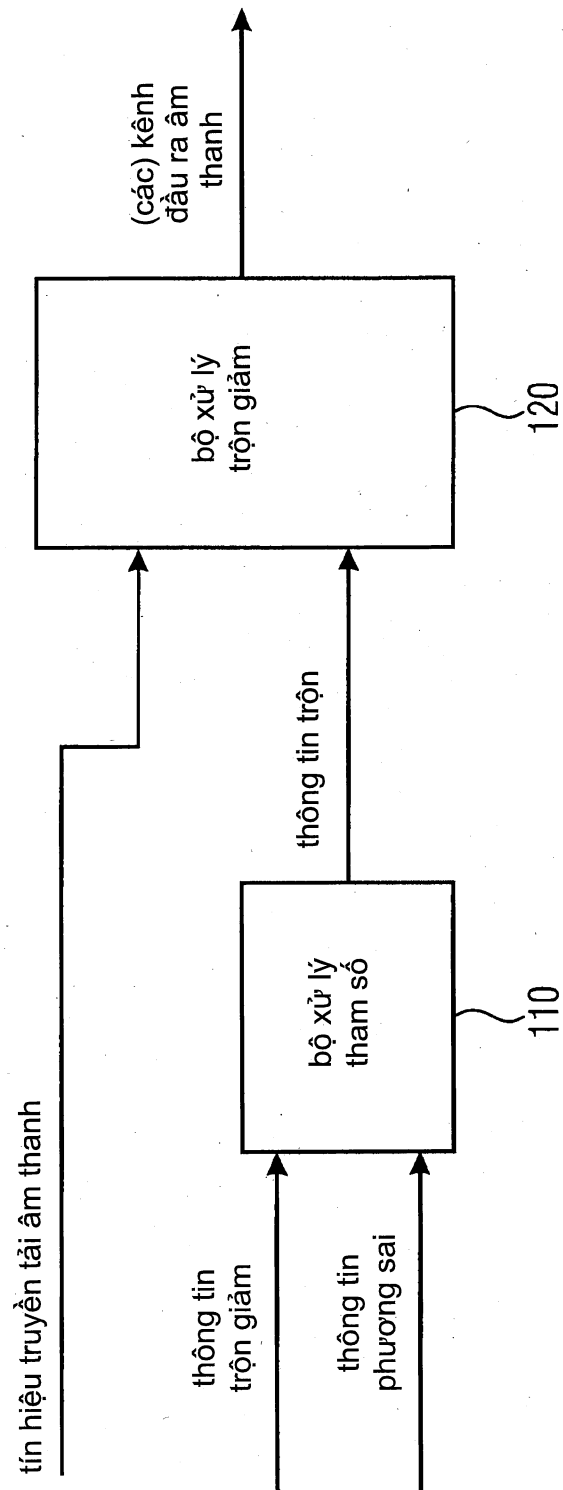


FIG.1

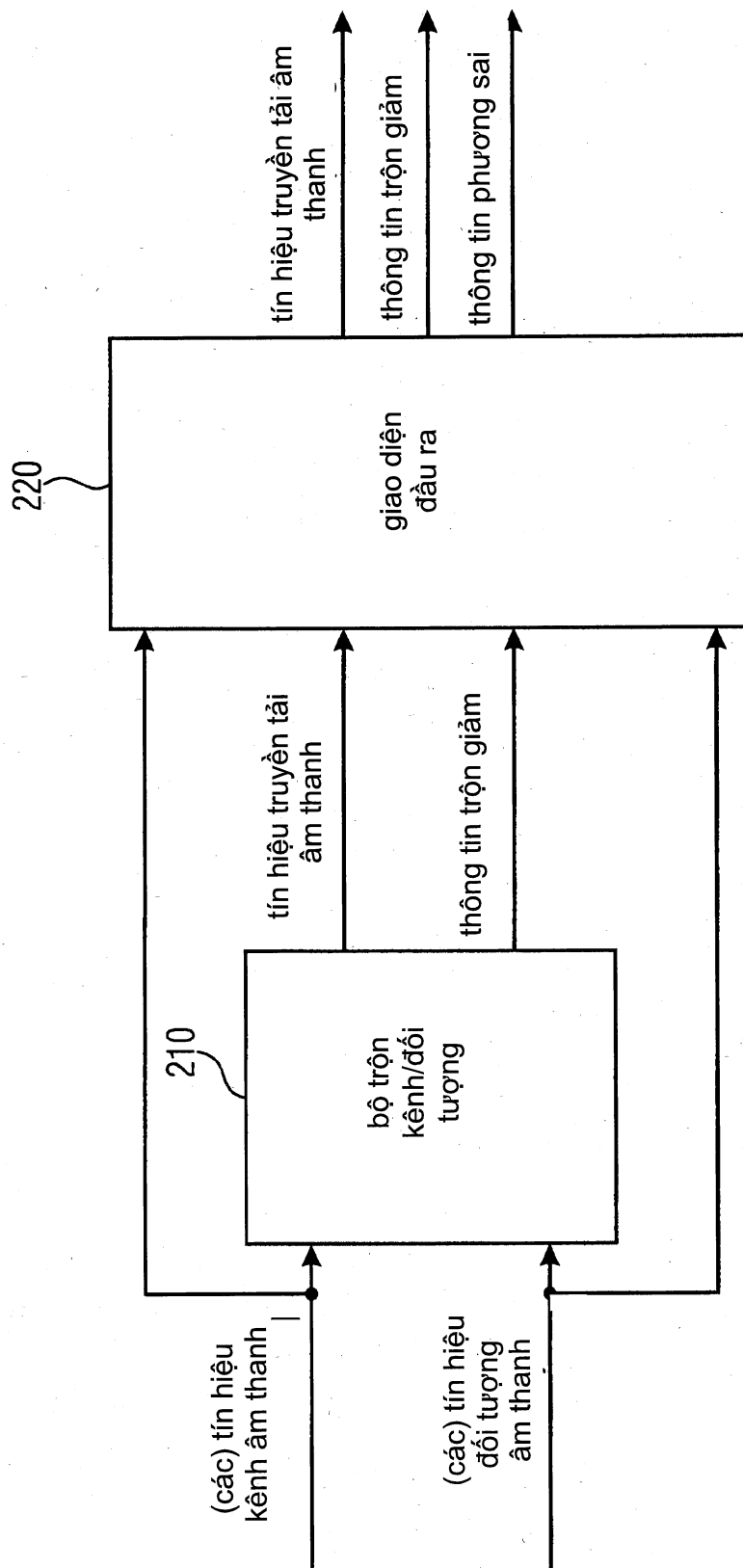


FIG.2

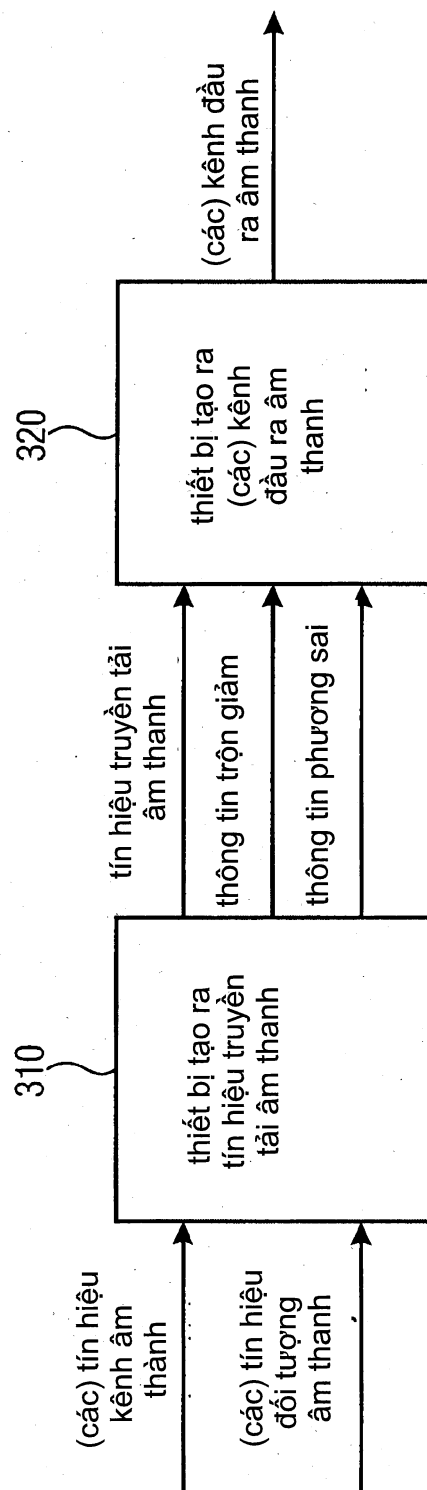
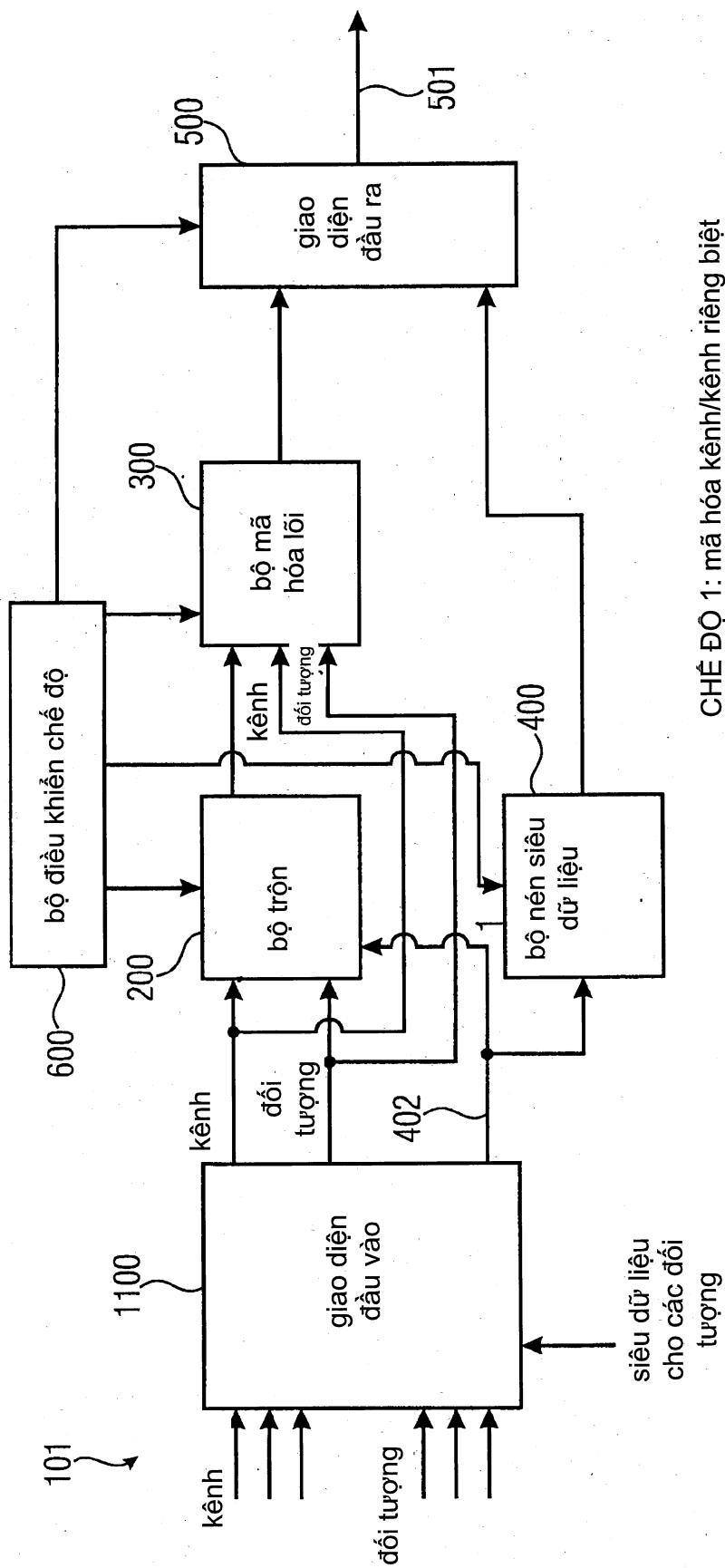
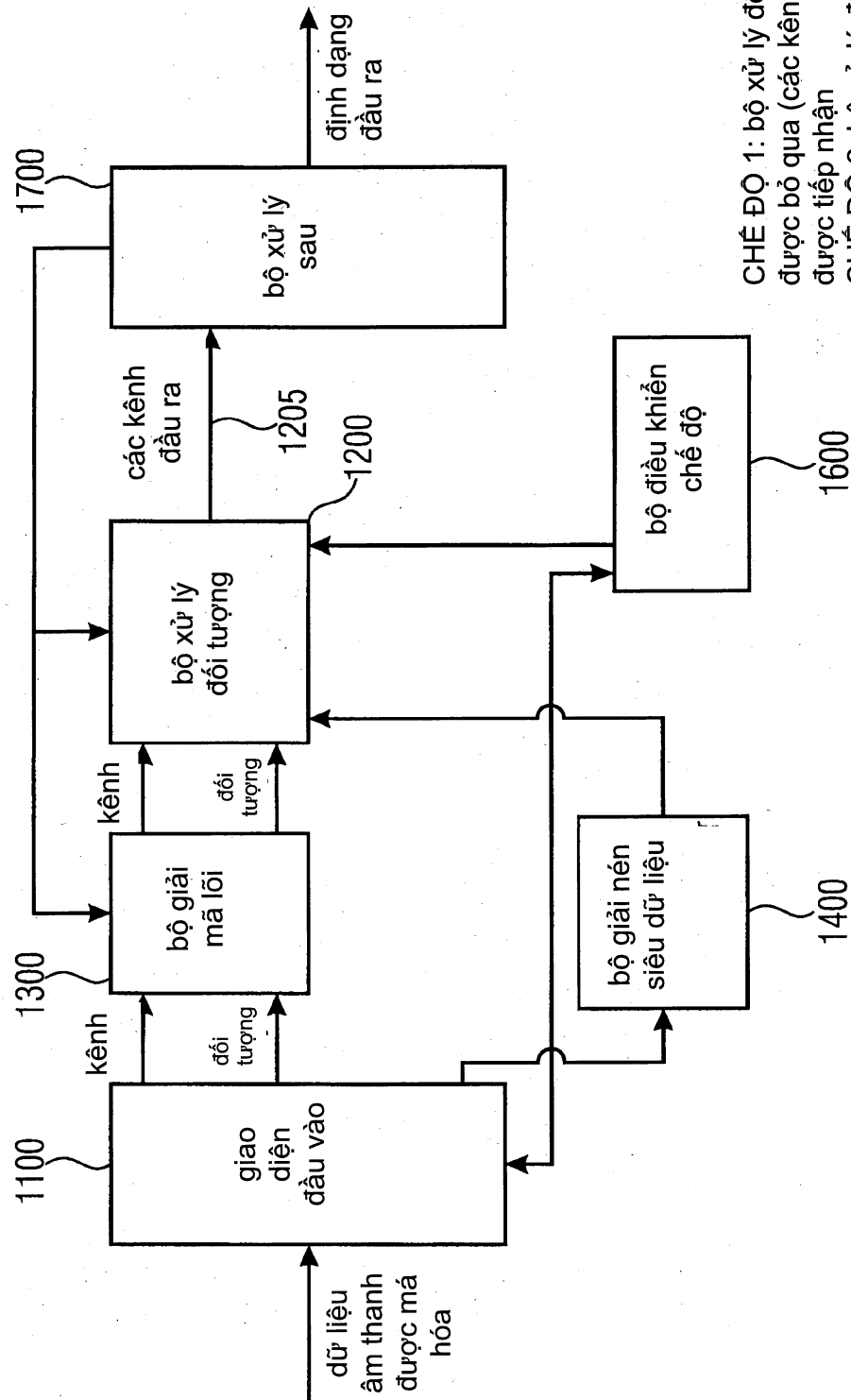


FIG.3



CHẾ ĐỘ 1: mã hóa kênh/kênh riêng biệt
CHẾ ĐỘ 2: trộn các kênh và các đổi tượng được kết xuất

FIG.4
(Bộ giải mã)



CHẾ ĐỘ 1: bộ xử lý đối tượng không được bỏ qua (các kênh và đối tượng được tiếp nhận)
 CHẾ ĐỘ 2: bộ xử lý đối tượng được bỏ qua (các kênh được tiếp nhận (chỉ được kết xuất trước))

FIG.5
(Bộ giải mã)

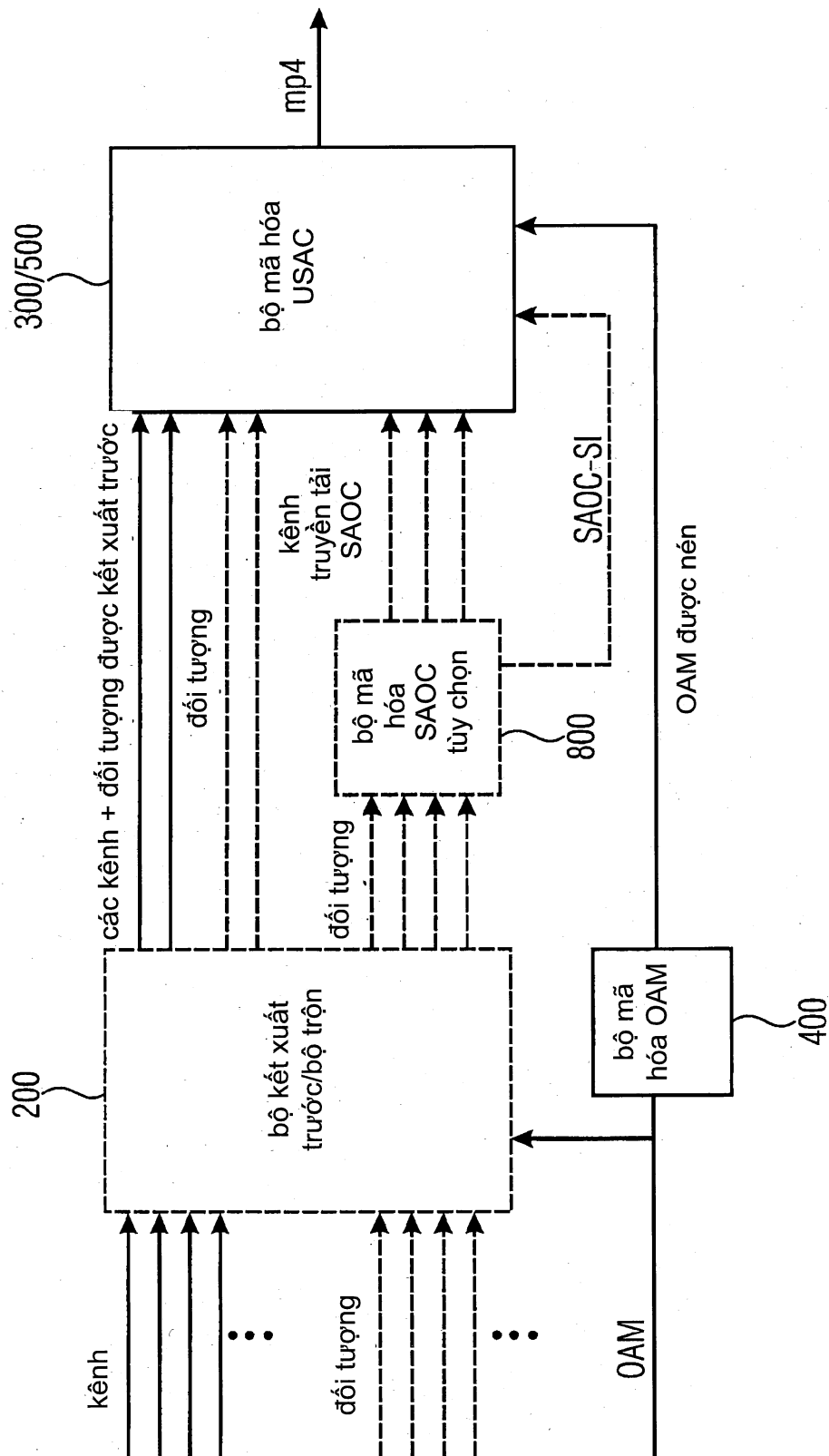


FIG. 6
(Bộ mã hóa)

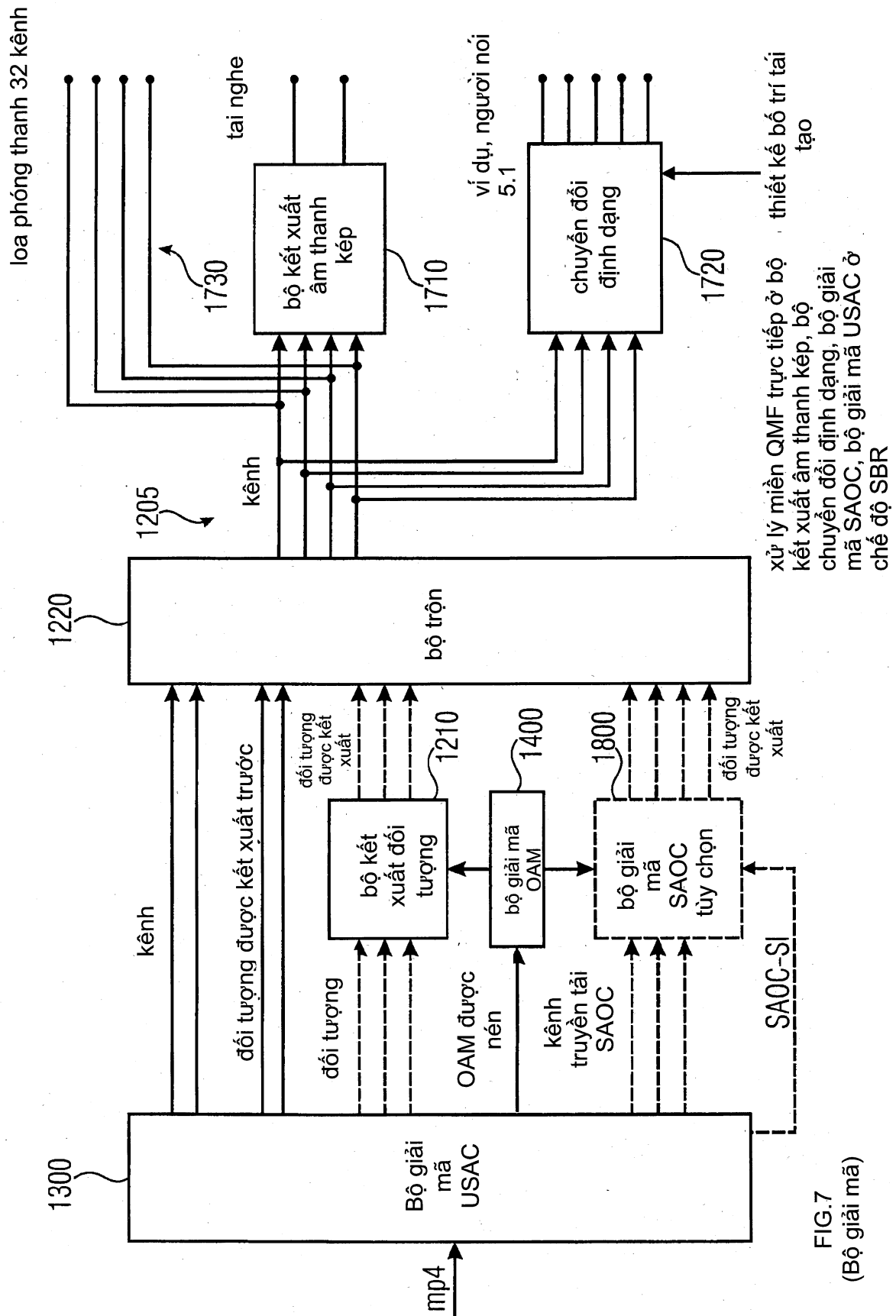


FIG. 7
(Bộ giải mã)

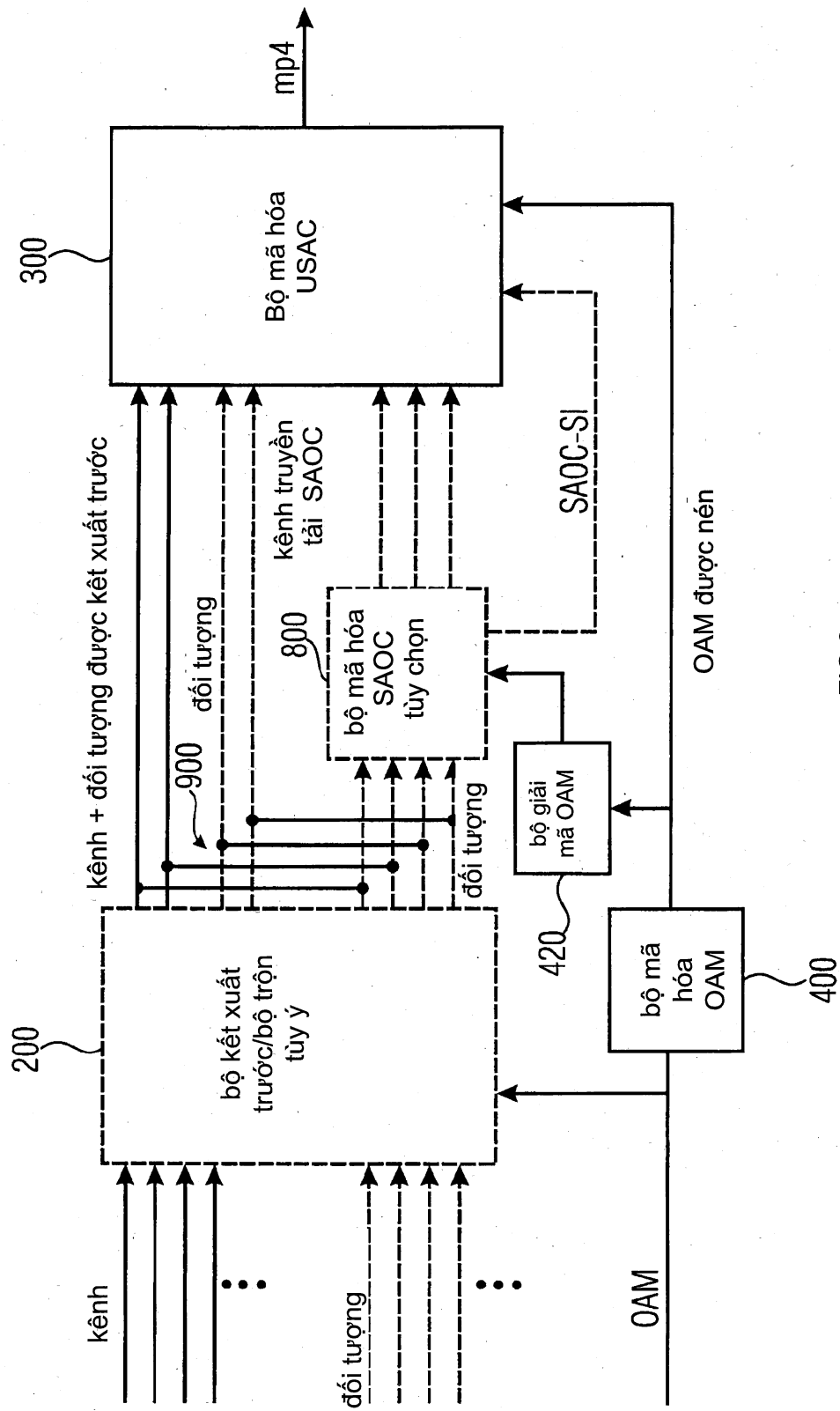


FIG.8
(Bộ mã hóa)

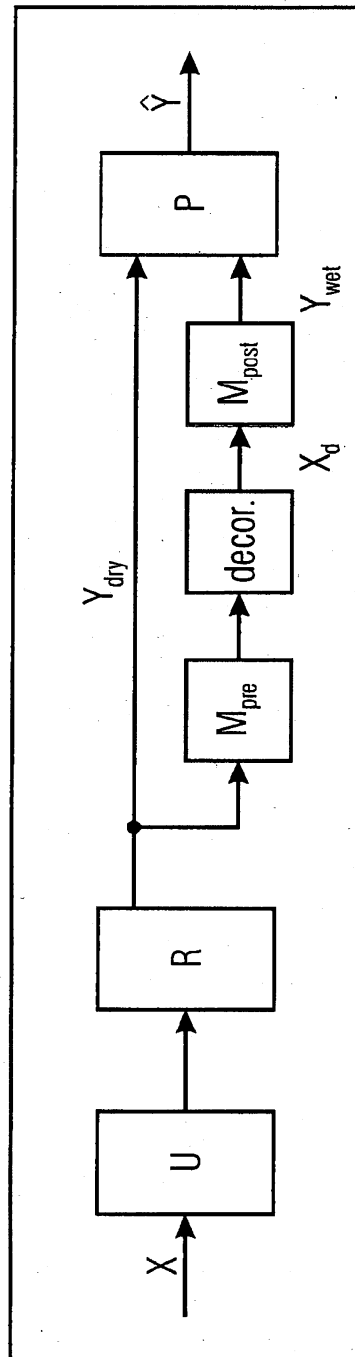


FIG.10