



(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)



1-0023257

(51)⁷ G10L 19/008; H04S 3/00

(13) B

(21) 1-2016-00618

(22) 16/07/2014

(86) PCT/EP2014/065290 16/07/2014

(87) WO2015/010999A1 29/01/2015

(30) 13177357.4 22/07/2013 EP; 13177371.5 22/07/2013 EP; 13177378.0

22/07/2013 EP; 13189281.2 18/10/2013 EP

(45) 27/04/2020 385

(43) 25/04/2016 337A

(73) FRAUNHOFER-GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN FORSCHUNG E.V. (DE)

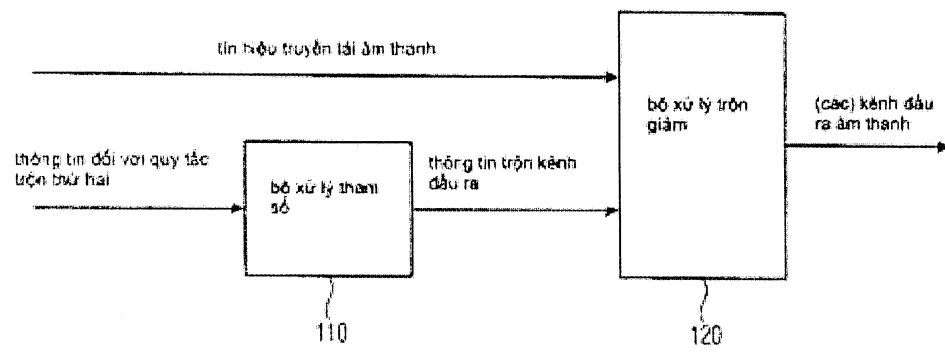
Hansastrasse 27c, 80686 Muenchen, Germany

(72) DISCH, Sascha (DE); FUCHS, Harald (DE); HELLMUTH, Oliver (DE); HERRE, Juergen (DE); MURTAZA, Adrian (RO); PAULUS, Jouni (FI); RIDDERBUSCH, Falko (DE); TERENTIV, Leon (DE)

(74) Công ty Luật TNHH 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, VÀ 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 để tạo ra một hoặc nhiều kênh đầu ra âm thanh, thiết bị và phương pháp tạo ra tín hiệu truyền tải âm thanh, và hệ thống các thiết bị này. Thiết bị này bao gồm bộ xử lý tham số (110) để tính toán thông tin trộn kênh đầu ra 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, trong đó hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được trộn vào tín hiệu truyền tải âm thanh, và trong đó số lượng của một hoặc nhiều kênh truyền tải âm thanh là nhỏ hơn số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh. Tín hiệu truyền tải âm thanh phụ thuộc vào quy tắc trộn thứ nhất và vào quy tắc trộn thứ hai. Quy tắc trộn thứ nhất biểu thị cách để trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh trộn trước. Hơn nữa, quy tắc trộn thứ hai biểu thị cách để trộn nhiều kênh trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh. Bộ xử lý tham số (110) được tạo cấu hình để nhận thông tin dựa trên quy tắc trộn thứ hai, trong đó thông tin dựa trên quy tắc trộn thứ hai biểu thị cách để trộn nhiều tín hiệu trộn trước sao cho thu được một hoặc nhiều kênh truyền tải âm thanh. Hơn nữa, bộ xử lý tham số (110) được tạo cấu hình để tính toán thông tin trộn kênh đầu ra phụ thuộc vào số lượng các đối tượng âm thanh biểu thị số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, phụ thuộc vào số lượng các kênh trộn trước biểu thị số lượng của nhiều kênh trộn trước, và phụ thuộc vào thông tin dựa trên quy tắc trộn thứ hai. 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 kênh đầu ra.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc mã hoá/giải mã âm thanh, cụ thể là, đến việc mã hóa âm thanh trong không gian và mã hóa đối tượng âm thanh trong không gian, và cụ thể hơn là đến thiết bị và phương pháp để thực hiện việc trộn giảm mã hóa đối tượng âm thanh trong không gian (Spatial Audio Object Coding – SAOC) của nội dung âm thanh ba chiều và đến thiết bị và phương pháp để giải mã một cách hiệu quả việc trộn giảm SAOC của nội dung âm thanh ba chiều.

Tình trạng kỹ thuật của sáng chế

Các công cụ mã hóa âm thanh trong không gian là đã được biết đến trong kỹ thuật hiện nay và là ví dụ, được chuẩn hoá trong âm thanh nổi theo tiêu chuẩn MPEG. Việc mã hóa âm thanh trong không gian bắt đầu từ các kênh đầu vào ban đầu như năm hoặc bảy kênh mà được nhận dạng bởi sự bố trí của chúng trong thiết lập tái tạo, nghĩa là kênh bên trái, kênh trung tâm, kênh bên phải, kênh âm thanh nổi bên trái, kênh âm thanh nổi bên phải và kênh tăng cường tần số thấp. Thông thường, bộ mã hoá âm thanh trong không gian suy ra một hoặc nhiều kênh trộn giảm từ các kênh ban đầu và, ngoài ra, suy ra dữ liệu tham số liên quan đến tín hiệu trong không gian như các chênh lệch mức liên kênh, các chênh lệch pha liên kênh, các chênh lệch thời gian liên kênh, v.v. Một hoặc nhiều kênh trộn giảm được truyền tải cùng với thông tin phụ tham số chỉ báo tín hiệu trong không gian đến bộ giải mã tín hiệu trong không gian mà giải mã kênh trộn giảm và dữ liệu tham số kết hợp để cuối cùng thu được kênh đầu ra mà là phiên bản gần đúng của các kênh đầu vào ban đầu. Việc bố trí các kênh trong thiết lập 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 phát thanh cụ thể ở vị trí đã cho. Việc tái tạo trung thực của loại định dạng này đòi hỏi thiết lập loa phát thanh trong đó người nói được đặt ở cùng vị trí như người nói mà

được sử dụng trong quá trình tạo ra tín hiệu âm thanh. Trong khi tăng số lượng loa phát thanh cải thiện việc tái tạo ngữ cảnh âm thanh ba chiều chìm đắm trung thực, ngày càng khó để đáp ứng yêu cầu này - đặc biệt là trong môi trường trong nhà như phòng khách.

Sự cần thiết của việc có thiết lập loa phát thanh cụ thể có thể được khắc phục bằng phương pháp dựa trên đối tượng trong đó tín hiệu loa phát thanh được kết xuất một cách cụ thể đối với thiết lập phát lại.

Ví dụ, công cụ mã hóa đối tượng âm thanh trong không gian là đã được biết đến rộng rãi trong kỹ thuật hiện nay và được chuẩn hoá trong tiêu chuẩn mã hóa đối tượng âm thanh trong không gian MPEG SAOC. Ngược với mã hóa âm thanh trong không gian bắt đầu từ các kênh ban đầu, việc mã hóa đối tượng âm thanh trong không gian bắt đầu từ các đối tượng âm thanh mà không được dành riêng một cách tự động đối với thiết lập tái tạo kết xuất nhất định. Thay vào đó, việc bố trí các đối tượng âm thanh trong ngữ cảnh tái tạo là có thể linh hoạt và có thể được xác định bởi người dùng bằng cách nhập thông tin kết xuất nhất định vào trong bộ giải mã mã hóa đối tượng âm thanh trong không gian. Theo cách khác hoặc ngoài ra, thông tin kết xuất, nghĩa là các thông tin mà vị trí trong thiết lập tái tạo, đối tượng âm thanh nhất định sẽ được đặt theo cách thông thường theo thời gian có thể được truyền như thông tin phụ bổ sung hoặc siêu dữ liệu. Để thu được việc nén dữ liệu nhất định, số lượng các đối tượng âm thanh được mã hoá bởi bộ mã hoá SAOC mà tính toán, từ các đối tượng đầu vào, một hoặc nhiều kênh truyền tải bằng cách trộn giảm các đối tượng theo các thông tin trộn giảm nhất định. Hơn nữa, bộ mã hoá SAOC tính toán các thông tin phụ tham số biểu diễn tín hiệu liên đối tượng như độ chênh lệch mức đối tượng (object level difference - OLD), trị số cố kết đối tượng, v.v. Các dữ liệu tham số liên đối tượng này được tính toán đối với các băng thời gian/tần số tham số, v.v, đối với khung nhất định của tín hiệu âm thanh bao gồm, ví dụ các mẫu 1024 hoặc 2048, 28, 20, 14 hoặc 10, v.v, các băng xử lý được xét đến sao cho, khi kết thúc, dữ liệu tham số tồn tại đối với mỗi khung và mỗi băng xử lý. Đối với ví dụ, khi mảnh âm thanh có 20 khung và khi mỗi khung được chia thành 28 băng xử lý, thì số lượng các ô thời gian/tần số bằng 560.

Trong phương pháp dựa trên đối tượng, trường âm thanh được mô tả bởi các

đối tượng âm thanh rời rạc. Việc này đòi hỏi siêu dữ liệu đối tượng mà mô tả trong số các yếu tố khác vị trí biến thể thời gian của mỗi nguồn âm thanh trong không gian ba chiều.

Khái niệm mã hóa siêu dữ liệu thứ nhất trong lĩnh vực kỹ thuật hiện nay là định dạng trao đổi mô tả âm thanh trong không gian (spatial sound description interchange format -SpatDIF), định dạng mô tả ngữ cảnh âm thanh mà vẫn đang phát triển [M1]. Sẽ được chỉ ra như định dạng trao đổi đối với ngữ cảnh âm thanh dựa trên đối tượng và không cung cấp phương pháp nén bất kỳ đối với quỹ đạo của đối tượng. SpatDIF sử dụng định dạng điều khiển âm thanh mở dựa trên ngữ cảnh (Open Sound Control - OSC) để cấu trúc siêu dữ liệu đối tượng [M2]. Tuy nhiên, phép biểu diễn dựa trên ngữ cảnh đơn giản không phải là sự lựa chọn để truyền nén quỹ đạo đối tượng.

Khái niệm siêu dữ liệu khác trong kỹ thuật hiện nay là định dạng mô tả ngữ cảnh âm thanh (Audio Scene Description Format - ASDF) [M3], giải pháp dựa trên cơ sở văn bản mà có cùng bất lợi. Dữ liệu được cấu trúc bằng việc mở rộng ngôn ngữ tích hợp đa phương tiện được đồng bộ hóa (Synchronized Multimedia Integration Language - SMIL) mà là tập hợp con của ngôn ngữ đánh dấu có thể mở rộng (Extensible Markup Language - XML) [M4], [M5].

Khái niệm siêu dữ liệu khác trong kỹ thuật hiện nay là định dạng nhị phân âm thanh đối với ngữ cảnh (AudioBIFS), định dạng nhị phân mà là một phần của đặc điểm kỹ thuật MPEG-4 [M6], [M7]. Nó liên quan một cách chặt chẽ với ngôn ngữ mô hình thực dựa trên XML (Virtual Reality Modeling Language - VRML) mà được phát triển để mô tả ngữ cảnh ba chiều nghe-nhìn và các áp dụng thực tương tác [M8]. Đặc điểm kỹ thuật AudioBIFS phức sử dụng biểu đồ ngữ cảnh để xác định rõ các đường di chuyển của đối tượng. Bất lợi chính của AudioBIFS là ở chỗ không được chỉ ra để vận hành thời gian thực trong đó hệ thống giới hạn trễ và ngẫu nhiên truy cập vào dòng dữ liệu là được đòi hỏi. Hơn nữa, việc mã hoá vị trí đối tượng không khai thác khả năng định vị hoá giới hạn của người nghe. Đối với vị trí người nghe cố định trong ngữ cảnh nghe-nhìn, dữ liệu đối tượng có thể được lượng tử hoá với số lượng bit thấp hơn nhiều [M9]. Do đó, việc mã hoá siêu dữ liệu đối tượng được áp dụng trong AudioBIFS là không hiệu quả khi xét đến 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, và 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 giải pháp được cải thiện để trộn giảm nội dung âm thanh. Mục đích của sáng chế được giải quyết bởi thiết bị và phương pháp nêu ở các điểm yêu cầu bảo hộ.

Theo các phương án, việc truyền tải hiệu quả là được thực hiện và có nghĩa là cách để giải mã trộn giảm đối với nội dung âm thanh ba chiều.

Thiết bị để tạo ra một hoặc nhiều kênh đầu ra âm thanh được đề xuất. Thiết bị bao gồm bộ xử lý tham số để tính toán thông tin trộn kênh đầu ra 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, trong đó hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được trộn vào tín hiệu truyền tải âm thanh, và trong đó số lượng của một hoặc nhiều kênh truyền tải âm thanh là nhỏ hơn số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh. Tín hiệu truyền tải âm thanh phụ thuộc vào quy tắc trộn thứ nhất và vào quy tắc trộn thứ hai. Quy tắc trộn thứ nhất biểu thị cách để trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh trộn trước. Hơn nữa, quy tắc trộn thứ hai biểu thị cách để trộn nhiều kênh trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh. Bộ xử lý tham số được cấu hình để nhận thông tin dựa trên quy tắc trộn thứ hai, trong đó thông tin dựa trên quy tắc trộn thứ hai biểu

thị cách để trộn nhiều tín hiệu trộn trước sao cho thu được một hoặc nhiều kênh truyền tải âm thanh. Hơn nữa, bộ xử lý tham số được tạo cấu hình để tính toán thông tin trộn kênh đầu ra phụ thuộc vào số lượng các đối tượng âm thanh biểu thị số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, phụ thuộc vào số lượng các kênh trộn trước biểu thị số lượng của nhiều kênh trộn trước, và phụ thuộc vào thông tin dựa trên quy tắc trộn thứ hai. 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 kênh đầu ra.

Hơn nữa, thiết bị để tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh được đề xuất. Thiết bị này bao gồm bộ trộn đối tượng để tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh từ hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, sao cho hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được trộn vào tín hiệu truyền tải âm thanh, và trong đó số lượng của một hoặc nhiều kênh truyền tải âm thanh là nhỏ hơn số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, và giao diện đầu ra để xuất ra tín hiệu truyền tải âm thanh. Bộ trộn đối tượng được cấu hình để tạo ra một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh phụ thuộc vào quy tắc trộn thứ nhất và phụ thuộc vào quy tắc trộn thứ hai, trong đó quy tắc trộn thứ nhất biểu thị cách để trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh trộn trước, và trong đó quy tắc trộn thứ hai biểu thị cách để trộn nhiều kênh trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh. Quy tắc trộn thứ nhất phụ thuộc vào số lượng các đối tượng âm thanh, biểu thị số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, và phụ thuộc vào số lượng các kênh trộn trước, biểu thị số lượng của nhiều kênh trộn trước, và trong đó quy tắc trộn thứ hai phụ thuộc vào số lượng các kênh trộn trước. Giao diện đầu ra được tạo cấu hình để xuất ra thông tin dựa trên quy tắc trộn thứ hai.

Hơn nữa, sáng chế cũng đề xuất hệ thống. Hệ thống này bao gồm thiết bị để tạo ra tín hiệu truyền tải âm thanh như được mô tả trên đây và thiết bị để tạo ra một hoặc nhiều kênh đầu ra âm thanh như được mô tả trên đây. Thiết bị để tạo ra một hoặc nhiều kênh đầu ra âm thanh được cấu hình để nhận tín hiệu truyền tải âm thanh và thông tin dựa trên quy tắc trộn thứ hai từ thiết bị để tạo ra tín hiệu truyền tải âm thanh. Hơn nữa,

thiết bị để tạo ra một hoặc nhiều kênh đầu ra âm thanh 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 dựa trên quy tắc trộn thứ hai.

Hơn nữa, sáng chế cũng đề xuất phương pháp để 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 hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được trộn vào tín hiệu truyền tải âm thanh, và trong đó số lượng của một hoặc nhiều kênh truyền tải âm thanh là nhỏ hơn số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, trong đó tín hiệu truyền tải âm thanh phụ thuộc vào quy tắc trộn thứ nhất và vào quy tắc trộn thứ hai, trong đó quy tắc trộn thứ nhất biểu thị cách để trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh trộn trước, và trong đó quy tắc trộn thứ hai biểu thị cách để trộn nhiều kênh trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh.
- Nhận thông tin dựa trên quy tắc trộn thứ hai, trong đó thông tin dựa trên quy tắc trộn thứ hai biểu thị cách để trộn nhiều tín hiệu trộn trước sao cho thu được một hoặc nhiều kênh truyền tải âm thanh.
- Tính toán thông tin trộn kênh đầu ra phụ thuộc vào số lượng các đối tượng âm thanh biểu thị số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, phụ thuộc vào số lượng các kênh trộn trước biểu thị số lượng của nhiều kênh trộn trước, và phụ thuộc vào thông tin dựa trên quy tắc trộn thứ hai. Và:
- Tạo ra một hoặc nhiều kênh đầu ra âm thanh từ tín hiệu truyền tải âm thanh phụ thuộc vào thông tin trộn kênh đầu ra.

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

- Tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh từ hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.
- Xuất ra tín hiệu truyền tải âm thanh. Và:

- Xuất ra thông tin dựa trên quy tắc trộn thứ hai.

Tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh từ hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được thực hiện sao cho hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được trộn vào tín hiệu truyền tải âm thanh, trong đó số lượng của một hoặc nhiều kênh truyền tải âm thanh là nhỏ hơn số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh. Tạo ra một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh được thực hiện phụ thuộc vào quy tắc trộn thứ nhất và phụ thuộc vào quy tắc trộn thứ hai, trong đó quy tắc trộn thứ nhất biểu thị cách để trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh trộn trước, và trong đó quy tắc trộn thứ hai biểu thị cách để trộn nhiều kênh trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh. Quy tắc trộn thứ nhất phụ thuộc vào số lượng các đối tượng âm thanh, biểu thị số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, và phụ thuộc vào số lượng các kênh trộn trước, biểu thị số lượng của nhiều kênh trộn trước. Quy tắc trộn thứ hai phụ thuộc vào số lượng các kênh trộn trước.

Hơn nữa, chương trình máy tính để thực thi phương pháp được mô tả trên đây khi được chạy trên máy tính hoặc bộ xử lý tín hiệu được đề xuất

Mô tả vắn tắt các hình vẽ

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

- Fig.1 minh hoạ thiết bị để tạo ra một hoặc nhiều kênh đầu ra âm thanh theo phương án,
- Fig.2 minh hoạ 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 hoạ hệ thống theo phương án,
- Fig.4 minh hoạ phương án thứ nhất của bộ mã hoá âm thanh ba chiều,
- Fig.5 minh hoạ phương án thứ nhất của bộ giải mã âm thanh ba chiều,
- Fig.6 minh hoạ phương án thứ hai của bộ mã hoá âm thanh ba chiều,

- Fig.7 minh hoạ phương án thứ hai của bộ giải mã âm thanh ba chiều,
 Fig.8 minh hoạ phương án thứ ba của bộ mã hoá âm thanh ba chiều,
 Fig.9 minh hoạ phương án thứ ba của bộ giải mã âm thanh ba chiều,
 Fig.10 minh hoạ vị trí của đối tượng âm thanh trong không gian ba chiều từ nguồn gốc được biểu hiện theo góc phương vị, góc nâng và bán kính, và
 Fig.11 minh hoạ vị trí của đối tượng âm thanh và loa phát thanh thiết lập được giả định bởi bộ tạo kênh âm thanh.

Mô tả chi tiết sáng chế

Trước khi mô tả chi tiết phương án được ưu tiên của sáng chế, hệ thống mã hóa-giải mã âm thanh ba chiều sẽ được mô tả.

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

Hạn chế này được khắc phục bởi hệ thống mã hóa-giải mã âm thanh ba chiều mới.

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

Fig.4 minh hoạ bộ mã hoá âm thanh ba chiều theo phương án của sáng chế. Bộ mã hoá âm thanh ba chiều được tạo cấu hình để mã hoá dữ liệu đầu vào âm thanh 101 để thu được dữ liệu đầu ra âm thanh 501. Bộ mã hoá âm thanh ba chiều bao gồm giao diện đầu vào để nhận nhiều kênh âm thanh được biểu thị bởi CH và nhiều đối tượng âm thanh được biểu thị bởi OBJ. Hơn nữa, như được minh hoạ trên Fig.4, giao diện đầu vào 1100 còn nhận siêu dữ liệu liên quan đến một hoặc nhiều trong số nhiều đối tượng âm thanh OBJ. Hơn nữa, bộ mã hoá âm thanh ba chiều bao gồm bộ trộn 200 để trộn nhiều đối tượng và nhiều kênh để thu được nhiều kênh trộn trước, trong đó mỗi kênh 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 hoá âm thanh ba chiều bao gồm bộ mã hoá lõi 300 để mã hoá

lỗi tín hiệu đầu vào bộ mã hoá lỗi, bộ nén siêu dữ liệu 400 để nén siêu dữ liệu liên quan đến một hoặc nhiều trong số nhiều đối tượng âm thanh.

Hơn nữa, bộ mã hoá âm thanh ba chiều có thể bao gồm bộ điều khiển chế độ 600 để điều khiển bộ trộn, bộ mã hoá lỗi và/hoặc giao diện đầu ra 500 trong một trong số vài chế độ vận hành, trong đó trong chế độ thứ nhất, bộ mã hoá lỗi được cấu hình để mã hoá nhiều kênh âm thanh và nhiều đối tượng âm thanh được nhận bởi giao diện đầu vào 1100 mà không có tương tác bất kỳ bởi bộ trộn, nghĩa là, không có việc trộn bất kỳ bởi bộ trộn 200. Tuy nhiên, trong chế độ thứ hai, trong đó bộ trộn 200 hoạt động, bộ mã hoá lỗi mã hoá nhiều kênh trộn, nghĩa là, đầu ra được tạo ra bởi khối 200. Trong trường hợp sau này, được ưu tiên để không mã hoá dữ liệu đối tượng bất kỳ thêm nữa. Để thay thế, vị trí biểu thị siêu dữ liệu của đối tượng âm thanh được sử dụng bởi bộ trộn 200 để kết xuất các đối tượng 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 nhiều đối tượng âm thanh để kết xuất trước đối tượng âm thanh và sau đó đối tượng âm thanh được kết xuất trước được trộn với các kênh để thu được các kênh được trộn ở đầu ra của bộ trộn. Theo phương án này, đối tượng bất kỳ có thể không nhất thiết phải được truyền và việc này cũng áp dụng cho siêu dữ liệu nén như được xuất ra bởi khối 400. Tuy nhiên, nếu không phải tất cả các đối tượng được nhập vào giao diện 1100 được trộn mà chỉ có lượng đối tượng nhất định được trộn, thì chỉ có các đối tượng không trộn còn lại và siêu dữ liệu kết hợp tuy nhiên được truyền lần lượt vào bộ mã hoá lỗi 300 hoặc bộ nén siêu dữ liệu 400.

Fig.6 minh họa phương án khác về bộ mã hoá âm thanh ba chiều mà còn bao gồm bộ mã hoá SAOC 800. Bộ mã hoá 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 hoá đối tượng âm thanh trong không gian. Như được minh họa trên Fig.6, dữ liệu đầu vào bộ mã hoá đối tượng âm thanh trong không gian là các đối tượng mà không được xử lý bởi bộ kết xuất trước/bộ trộn. Theo cách khác, được đề xuất rằng bộ kết xuất trước/bộ trộn đã được chuyển qua như trong chế độ trong đó mã hóa kênh/đối tượng riêng rẽ là hoạt động, tất cả các đối tượng nhập vào giao diện đầu vào 1100 được mã hoá bởi bộ mã hoá SAOC 800.

Hơn nữa, như được minh hoạ trên Fig.6, bộ mã hoá lõi 300 tốt hơn là được thực thi như bộ mã hoá USAC, nghĩa là, như bộ mã hoá được xác định và tiêu chuẩn hoá theo tiêu chuẩn MPEG-USAC (USAC = Unified Speech and Audio Coding-mã hóa tiếng nói và âm thanh thống nhất). Đầu ra của toàn bộ bộ mã hoá âm thanh ba chiều được minh hoạ 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 tương tự bộ chứa đối với các loại dữ liệu riêng rẽ. 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 hoá OAM 400 để thu được dữ liệu OAM nén mà được nhập vào bộ mã hoá USAC 300 mà, như có thể được nhìn 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 hoá mà còn có dữ liệu OAM nén.

Fig.8 minh hoạ phương án khác về bộ mã hoá âm thanh ba chiều, trong đó ngược với Fig.6, bộ mã hoá SAOC có thể được tạo cấu hình để mã hoá, với thuật toán mã hoá SAOC, các kênh được cung cấp ở bộ kết xuất trước/bộ trộn 200 là không hoạt động trong chế độ này hoặc theo cách khác, để mã hoá SAOC các kênh kết xuất trước cộng đối tượng. Do đó, trên Fig.8, bộ mã hoá SAOC 800 có thể vận hành ba loại khác nhau của dữ liệu đầu vào, nghĩa là, các kênh mà không đối tượng kết xuất trước bất kỳ, các kênh và đối tượng kết xuất trước hoặc các đối tượng đơn lẻ. Hơn nữa, ưu tiên là cung cấp bộ giải mã OAM 420 bổ sung trên Fig.8 sao cho bộ mã hoá SAOC 800 sử dụng, để xử lý, cùng dữ liệu như ở phía bộ giải mã, nghĩa là, dữ liệu thu được do nén giảm bớt hơn là dữ liệu OAM ban đầu.

Fig.8 bộ mã hoá âm thanh ba chiều thể vận hành theo vài chế độ riêng rẽ.

Ngoài chế độ thứ nhất và thứ hai như được thảo luận trong ngữ cảnh trên Fig.4, bộ mã hoá âm thanh ba chiều trên Fig.8 còn có thể vận hành trong chế độ thứ ba trong đó bộ mã hoá lõi tạo ra một hoặc nhiều kênh truyền tải từ các đối tượng riêng rẽ khi bộ kết xuất trước/bộ trộn 200 là không hoạt động. Ngoài ra hoặc theo cách khác, trong chế độ thứ ba, bộ mã hoá SAOC 800 có thể tạo ra một hoặc nhiều kênh truyền tải thay thế hoặc bổ sung từ các kênh ban đầu, nghĩa là, một lần nữa khi bộ kết xuất trước/bộ trộn 200 tương ứng với bộ trộn 200 trên Fig.4 là không hoạt động.

Cuối cùng, bộ mã hoá SAOC 800 có thể mã hoá, khi bộ mã hoá âm thanh ba

chiều được cấu hình theo chế độ thứ tư, các kênh cộng với đối tượng kết xuất trước như được tạo ra bởi bộ kết xuất trước/bộ trộn. Do đó, trong chế độ thứ tư, việc áp dụng tốc độ bit thấp nhất sẽ tạo ra chất lượng tốt do thực tế rằng các kênh và đối tượng đã được biến đổi hoàn toàn thành các kênh truyền tải SAOC riêng rẽ và thông tin phụ kết hợp như được biểu thị trên các hình vẽ Fig.3 và Fig.5 như “SAOC-SI” và ngoài ra, siêu dữ liệu nén bất kỳ không phải được truyền trong chế độ thứ tư này.

Fig.5 minh hoạ bộ giải mã âm thanh ba chiều theo phương án của sáng chế. Bộ giải mã âm thanh ba chiều nhận, như tín hiệu vào, dữ liệu âm thanh được mã hoá, tức là, dữ liệu 501 trên 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.

Đặc biệt là, bộ giải mã âm thanh ba chiều được cấu hình để giải mã dữ liệu âm thanh được mã hoá và giao diện đầu vào được tạo cấu hình để nhận dữ liệu âm thanh được mã hoá, dữ liệu âm thanh được mã hoá bao gồm nhiều kênh được mã hoá và nhiều đối tượng được mã hoá và siêu dữ liệu nén liên quan đến nhiều đối tượng trong chế độ nhất định.

Hơn nữa, bộ giải mã lõi 1300 được tạo cấu hình để giải mã nhiều kênh được mã hoá và nhiều đối tượng được mã hoá 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 sử dụng siêu dữ liệu giải nén để thu được số lượng định trước của các kênh đầu ra bao gồm dữ liệu đối tượng và kênh được giải mã. Các kênh đầu ra này như được biểu thị ở 1205 sau đó được nhập vào trong 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 nhị phân hoặc định dạng đầu ra loa phát 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 mã hoá để phát hiện sự biểu thị 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 là tồn tại. Để thay thế, bộ giải mã âm thanh linh hoạt có thể được thiết lập trước bởi loại khác bất kỳ của dữ liệu đối chứng như đầu vào người dùng hoặc đối chứng 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 chuyển qua bộ xử lý đối tượng và để cấp các kênh giải mã vào trong bộ xử lý sau 1700. Đây là sự vận hành trong chế độ 2, nghĩa là, trong đó chỉ có các kênh kết xuất trước là được nhận, nghĩa là, khi chế độ 2 được áp dụng trong bộ mã hoá âm thanh ba chiều trên Fig.4. Theo cách khác, khi chế độ 1 được áp dụng trong bộ mã hoá âm thanh ba chiều, nghĩa là, khi bộ mã hoá âm thanh ba chiều thực hiện mã hóa kênh/đối tượng riêng rẽ, sau đó bộ xử lý đối tượng 1200 là không được chuyển qua, nhưng nhiều kênh được giải mã và nhiều đối tượng được giải mã được cấp vào trong bộ xử lý đối tượng 1200 cùng với siêu dữ liệu giải nén được tạo ra bởi bộ giải nén siêu dữ liệu 1400.

Tốt hơn là, chỉ báo liệu chế độ 1 hay chế độ 2 sẽ được áp dụng là được bao gồm trong dữ liệu âm thanh được mã hoá hay không và sau đó bộ điều khiển chế độ 1600 phân tích dữ liệu được mã hoá để phát hiện chỉ báo chế độ. Chế độ 1 được sử dụng khi chỉ báo chế độ chỉ báo rằng dữ liệu âm thanh được mã hoá bao gồm kênh được mã hoá và đối tượng được mã hoá và chế độ 2 được áp dụng khi chỉ báo chế độ chỉ báo rằng dữ liệu âm thanh được mã hoá không chứa đối tượng âm thanh bất kỳ, nghĩa là, chỉ chứa các kênh kết xuất trước thu được bởi chế độ 2 của bộ mã hoá âm thanh ba chiều trên Fig.4.

Fig.7 minh họa phương án được ưu tiên so với bộ giải mã âm thanh ba chiều trên Fig.5 và phương án trên Fig.7 tương ứng với bộ mã hoá âm thanh ba chiều trên Fig.6. Ngoài phương án về bộ giải mã âm thanh ba chiều trên 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 trên Fig.5 được thực thi 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 thi bởi bộ giải mã SAOC 1800.

Hơn nữa, bộ xử lý sau 1700 có thể được thực thi như bộ kết xuất lập thể 1710

hoặc bộ chuyển đổi định dạng 1720. Ngoài ra, đầu ra dữ liệu trực tiếp 1205 trên Fig.5 cũng có thể được thực thi như được minh hoạ bởi 1730. Do đó, ưu tiên là 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 sau đó để xử lý sau nếu định dạng nhỏ hơn là được đòi hỏi. Tuy nhiên, khi nó trở nên rõ ràng từ rất sớm mà chỉ có các định dạng khác với số lượng kênh nhỏ hơn như định dạng 5.1 là được đòi hỏi, thì được ưu tiên, như được biểu thị trên Fig.9 bởi đường tắt 1727, mà điều khiển nhất định đối với bộ giải mã SAOC và/hoặc bộ giải mã USAC có thể được áp dụng để ngăn ngừa các vận hành trộn tăng không cần thiết và các vận hành trộn giảm tiếp theo.

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

Hơn nữa, bộ xử lý đối tượng 1200 được tạo cấu hình để kết xuất đối tượng được giải mã được xuất ra bởi bộ giải mã lõi mà không được mã hoá trong kênh truyền tải SAOC nhưng mà được mã hoá theo cách riêng rẽ trong các yếu tố kênh đơn lẻ thông thường 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 phát 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 kết hợp biểu diễn tín hiệu âm thanh được mã hoá hoặc kênh âm thanh được mã hoá, trong đó bộ giải mã mã hóa đối tượng âm thanh trong không gian được tạo cấu hình để chuyển mã thông tin tham số kết hợp và siêu dữ liệu giải nén thành thông tin phụ tham số được chuyển mã có thể sử dụng để kết xuất một cách trực tiếp định dạng đầu ra, như ví dụ được xác định trong phiên bản trước của SAOC. Bộ xử lý sau 1700 được tạo cấu hình để tính toán kênh âm thanh của định dạng đầu ra sử dụng kênh truyền tải giải mã và thông tin phụ tham số chuyển mã. Việc xử lý được thực hiện bởi bộ xử lý sau có thể là tương tự với xử lý âm thanh nổi theo tiêu chuẩn

MPEG hoặc có thể là xử lý khác bất kỳ như xử lý BCC hoặc v.v.

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 được tạo cấu hình để trộn tăng một cách trực tiếp và kết xuất tín hiệu kênh đối với định dạng đầu ra sử dụng kênh truyền tải được giải mã (bởi bộ giải mã lỗi) và thông tin phụ tham số.

Hơn nữa, và quan trọng là, bộ xử lý đối tượng 1200 trên Fig.5 còn bao gồm bộ trộn 1220 mà nhận, như đầu ra, dữ liệu được xuất ra bởi bộ giải mã USAC 1300 một cách trực tiếp khi đối tượng kết xuất trước được trộn với các kênh đang tồn tại, nghĩa là, khi bộ trộn 200 trên Fig.4 là hoạt động. Ngoài ra, bộ trộn 1220 nhận các dữ liệu từ bộ kết xuất đối tượng thực hiện việc kết xuất đối tượng mà không cần giải mã SAOC. Hơn nữa, bộ trộn nhận dữ liệu đầu ra bộ giải mã SAOC, nghĩa là, các đối tượng 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 sử dụng đầu liên quan đến hàm chuyển hoặc đáp ứng xung phòng lập thể (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 so với kênh đầu ra 1205 của bộ trộn và bộ chuyển đổi định dạng 1720 đòi hỏi thông tin đối với thiết kế tái tạo như loa 5.1 hoặc v.v.

Bộ giải mã âm thanh ba chiều trên Fig.9 khác với bộ giải mã âm thanh trên Fig.7 ở chỗ bộ giải mã SAOC có thể không chỉ tạo ra các đối tượng kết xuất mà còn cả các kênh kết xuất và đây là trường hợp khi bộ mã hoá âm thanh ba chiều trên Fig.8 được sử dụng và kết nối 900 giữa các kênh/đối tượng kết xuất trước và giao diện đầu vào bộ mã hoá SAOC 800 là hoạt động.

Hơn nữa, giai đoạn ghép nhấn biên độ cơ sở vector (vector base amplitude panning - VBAP) 1810 được tạo cấu hình tiếp nhận, từ bộ giải mã SAOC, thông tin dựa trên bố trí tái tạo và mã xuất ra ma trận kết xuất đến bộ giải mã SAOC sao cho bộ giải mã SAOC có thể, cuối cùng là, tạo ra các kênh được kết xuất mà không cần bất kỳ việc vận hành nào khác của bộ trộn ở định dạng kênh cao 1205, tức là 32 loa phát 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 thiết kế bố trí tái tạo mà còn là của các vị trí mà các tín hiệu đầu vào nên được kết xuất ở thiết kế bố trí tái tạo. Dữ liệu đầu vào hình học có thể là dữ liệu OAM đối với các đối tượng hoặc thông tin vị trí kênh đối với các kết mà được truyền sử dụng SAOC.

Tuy nhiên, nếu chỉ có giao diện đầu ra cụ thể được yêu cầu thì giai đoạn VBAP 1810 có thể đã tạo ra ma trận kết xuất yêu cầu đối với, ví dụ, đầu ra 5.1. Sau đó, bộ giải mã SAOC 1800 thực hiện việc kết xuất trực tiếp từ các kênh truyền tải SAOC, dữ liệu tham số có liên quan 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, nghĩa là, trong đó một vài kênh là SAOC được mã hóa 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 có lượng đối tượng được kết xuất trước nhất định với các kênh được giải mã SAOC và các kênh còn lại không phải là SAOC được xử lý thì bộ trộn sẽ đưa cùng với dữ liệu từ các phần đầu vào riêng biệt, nghĩa là trực tiếp từ bộ giải mã lõi 1300, từ bộ kết xuất đối tượng 1210 và từ bộ giải mã SAOC 1800.

Trong âm thanh ba chiều, góc phương vị, góc nâng và bán kính được sử dụng để xác định vị trí của đối tượng âm thanh. Hơn nữa, độ khuếch đại đối tượng âm thanh có thể được truyền.

Góc phương vị, góc nâng và bán kính xác định theo cách rõ ràng vị trí của đối tượng âm thanh trong không gian ba chiều từ ban đầu. Việc này được minh họa khi tham khảo đến Fig.10.

Fig.10 minh họa vị trí 410 của đối tượng âm thanh trong không gian ba chiều (3D) từ điểm gốc 400 được biểu hiện như góc phương vị, góc nâng và bán kính.

Góc phương vị làm rõ, ví dụ, góc trong mặt phẳng xy (mặt phẳng được xác định theo trục x và trục y). Góc nâng định rõ, ví dụ, góc trong mặt phẳng xz (mặt phẳng được xác định bởi trục x và trục z). Bằng cách xác định rõ góc phương vị và góc nâng, đường thẳng 415 đến điểm gốc 400 và vị trí 410 của đối tượng âm thanh có thể được xác định. Bằng cách xác định rõ hơn nữa bán kính, vị trí chính xác 410 của đối tượng

âm thanh có thể được xác định.

Theo phương án, góc phương vị được định rõ trong khoảng: $-180^\circ < \text{góc phương vị} \leq 180^\circ$, góc nâng được định rõ trong khoảng: $-90^\circ < \text{góc nâng} \leq 90^\circ$ và bán kính có thể, ví dụ, được xác định theo met [m] (bằng 0m hoặc lớn hơn). Hình cầu được mô tả bởi góc phương vị, góc nâng và góc có thể được chia thành hai bán cầu: bán cầu trái ($0^\circ < \text{góc phương vị} \leq 180^\circ$) và bán cầu phải ($-180^\circ < \text{góc phương vị} \leq 0^\circ$), bán cầu trên ($0^\circ < \text{góc nâng} \leq 90^\circ$) và bán cầu dưới ($-90^\circ < \text{góc nâng} \leq 0^\circ$)

Theo phương án khác, trong đó có thể, ví dụ, được giả định rằng tất cả các trị số x của vị trí đối tượng âm thanh trong hệ toạ độ xyz là lớn hơn hoặc bằng 0, góc phương vị có thể được định rõ trong khoảng: $-90^\circ \leq \text{góc phương vị} \leq 90^\circ$, góc nâng có thể được định rõ trong khoảng: $-90^\circ < \text{góc nâng} \leq 90^\circ$, và bán kính có thể, ví dụ, được xác định theo met [m].

Bộ xử lý trộn giảm 120 có thể, ví dụ được tạo cấu hình để tạo ra một hoặc nhiều kênh âm thanh phụ thuộc vào một hoặc nhiều tín hiệu đối tượng âm thanh phụ thuộc vào các trị thông tin siêu dữ liệu tái cấu trúc, trong đó trị số thông tin siêu dữ liệu tái cấu trúc có thể, ví dụ biểu thị vị trí của đối tượng âm thanh.

Theo phương án này, các trị số thông tin siêu dữ liệu có thể, ví dụ biểu thị góc phương vị được xác định đối với khoảng: $-180^\circ < \text{góc phương vị} \leq 180^\circ$, góc nâng được xác định đối với khoảng: $-90^\circ < \text{góc nâng} \leq 90^\circ$ và bán kính có thể, ví dụ được xác định theo met [m] (bằng 0m hoặc lớn hơn).

Fig.11 minh hoạ vị trí của đối tượng âm thanh và thiết lập loa phát thanh được giả định bởi bộ tạo kênh âm thanh. Ban đầu 500 của hệ toạ độ xyz được minh hoạ. Hơn nữa, vị trí 510 của đối tượng âm thanh thứ nhất và vị trí 520 của đối tượng âm thanh thứ hai là được minh hoạ. Hơn nữa, Fig.11 minh hoạ ngữ cảnh, trong đó bộ tạo kênh âm thanh 120 tạo ra bốn kênh âm thanh đối với bốn loa phát thanh. Bộ tạo kênh âm thanh 120 giả định rằng bốn loa phát thanh 511, 512, 513 và 514 được định vị ở các vị trí được thể hiện trên Fig.11.

Trên Fig.11, đối tượng âm thanh thứ nhất được đặt ở vị trí 510 gần với các vị trí giả định của loa phát thanh 511 và 512, và được đặt cách xa loa phát thanh 513 và 514.

Do đó, bộ tạo kênh âm thanh 120 có thể tạo ra bốn kênh âm thanh sao cho đối tượng âm thanh thứ nhất 510 được tái tạo bởi các loa phát thanh 511 và 512 nhưng không bởi các loa phát thanh 513 và 514.

Theo phương án khác, bộ tạo kênh âm thanh 120 có thể tạo ra bốn kênh âm thanh sao cho đối tượng âm thanh thứ nhất 510 được tái tạo với mức cao bởi các loa phát thanh 511 và 512 và với mức thấp bởi các loa phát thanh 513 và 514.

Hơn nữa, đối tượng âm thanh thứ hai được đặt ở vị trí 520 gần với các vị trí giả định của các loa phát thanh 513 và 514, và được đặt cách xa loa phát thanh 511 và 512. Do đó, bộ tạo kênh âm thanh 120 có thể tạo ra bốn kênh âm thanh sao cho đối tượng âm thanh thứ hai 520 được tái tạo bởi các loa phát thanh 513 và 514 nhưng không bởi các loa phát thanh 511 và 512.

Theo phương án khác, bộ xử lý trộn giảm 120 có thể tạo ra bốn kênh âm thanh sao cho đối tượng âm thanh thứ hai 520 được tái tạo với mức cao bởi các loa phát thanh 513 và 514 và với mức thấp bởi các loa phát thanh 511 và 512.

Theo phương án khác, chỉ có hai trị số thông tin siêu dữ liệu được sử dụng để định rõ vị trí của đối tượng âm thanh. Ví dụ, chỉ có góc phương vị và bán kính có thể được định rõ, ví dụ khi được giả định rằng tất cả các đối tượng âm thanh được đặt trong mặt phẳng đơn lẻ.

Theo phương án khác nữa, đối với mỗi đối tượng âm thanh, chỉ có trị số thông tin siêu dữ liệu đơn lẻ của tín hiệu siêu dữ liệu được mã hoá và được truyền như thông tin vị trí. Ví dụ, chỉ có góc phương vị có thể được xác định rõ như thông tin vị trí đối với đối tượng âm thanh (ví dụ, có thể được giả định rằng tất cả các đối tượng âm thanh được đặt trong cùng mặt phẳng có cùng khoảng cách với điểm trung tâm và do đó được giả định có cùng bán kính). Các thông tin góc phương vị có thể, ví dụ là đủ để xác định rằng đối tượng âm thanh được đặt gần với loa phát thanh bên trái và cách xa với loa phát thanh bên phải. Trong trường hợp này, bộ tạo kênh âm thanh 120 có thể, ví dụ, tạo ra một hoặc nhiều kênh âm thanh sao cho đối tượng âm thanh được tái tạo bởi loa phát thanh bên trái, mà không phải loa phát thanh bên phải.

Ví dụ, giai đoạn ghép nhẵn độ cơ sở vector có thể được dùng để xác định trọng

số của tín hiệu đối tượng âm thanh trong mỗi kênh đầu ra âm thanh trong số các kênh đầu ra âm thanh (xem, ví dụ, [VBAP]). Đối với VBAP, được giả định rằng tín hiệu đối tượng âm thanh được gán cho nguồn thực, và hơn nữa được giả định rằng kênh đầu ra âm thanh là kênh của loa phát thanh.

Theo các phương án, trị số thông tin siêu dữ liệu khác ví dụ, của tín hiệu siêu dữ liệu khác có thể xác định rõ âm lượng, ví dụ, độ khuếch đại (ví dụ, được biểu thị bằng decibel [dB]) đối với mỗi đối tượng âm thanh.

Ví dụ, trên Fig.11, trị số khuếch đại thứ nhất có thể được làm rõ bởi trị số thông tin siêu dữ liệu khác đối với đối tượng âm thanh thứ nhất được đặt ở vị trí 510 mà cao hơn so với trị số khuếch đại thứ hai là được xác định rõ bởi trị số thông tin siêu dữ liệu khác nữa đối với đối tượng âm thanh thứ hai được định vị ở vị trí 520. Trong trường hợp này, loa phát thanh 511 và 512 có thể tái tạo đối tượng âm thanh thứ nhất với mức cao hơn so với mức mà các loa phát thanh 513 và 514 tái tạo đối tượng âm thanh thứ hai.

Theo kỹ thuật SAOC, bộ mã hoá SAOC nhận nhiều tín hiệu đối tượng âm thanh X và trộn giảm chúng bằng cách dùng ma trận trộn giảm D để thu được tín hiệu truyền tải âm thanh Y bao gồm một hoặc nhiều kênh truyền tải âm thanh. Công thức

$$Y = DX$$

có thể được dùng. Bộ mã hoá SAOC truyền tín hiệu truyền tải âm thanh Y và các thông tin đối với ma trận trộn giảm D (ví dụ, các hệ số của ma trận trộn giảm D) đến bộ giải mã SAOC. Hơn nữa, bộ mã hoá SAOC truyền các thông tin đối với ma trận hiệp biến E (ví dụ, hệ số của ma trận hiệp biến E) đến bộ giải mã SAOC.

Ở phía bộ giải mã, tín hiệu đối tượng âm thanh X có thể được tái cấu trúc để thu được đối tượng âm thanh tái cấu trúc $\hat{X}$ bằng cách dùng công thức

$$\hat{X} = GY$$

trong đó G là ma trận ước lượng nguồn tham số với $G = E D^H (D E D^H)^{-1}$.

Sau đó, một hoặc nhiều kênh đầu ra âm thanh Z có thể được tạo ra bằng cách áp dụng ma trận kết xuất R đối với đối tượng âm thanh tái cấu trúc $\hat{X}$ theo công thức:

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

Tuy nhiên, việc tạo ra một hoặc nhiều kênh đầu ra âm thanh Z từ tín hiệu truyền tải âm thanh có thể cũng được thực hiện ở bước đơn lẻ bằng cách áp dụng ma trận U theo công thức:

$$Z = UY , \quad \text{với } U = RG .$$

Mỗi hàng của ma trận kết xuất R được kết hợp với một trong số các kênh đầu ra âm thanh mà sẽ được tạo ra. Mỗi hệ số trong một hàng của ma trận kết xuất R xác định trọng số của một trong số các tín hiệu đối tượng âm thanh tái cấu trúc trong kênh đầu ra âm thanh, mà dòng này của ma trận kết xuất R đề cập.

Ví dụ, ma trận kết xuất R có thể phụ thuộc vào thông tin vị trí đối với mỗi tín hiệu trong số tín hiệu đối tượng âm thanh được truyền đến bộ giải mã SAOC trong thông tin siêu dữ liệu. Ví dụ, tín hiệu đối tượng âm thanh có vị trí mà được đặt gần với vị trí loa phát thanh được giả định hoặc thực có thể, ví dụ, có trọng số cao hơn bên trong kênh đầu ra âm thanh của loa phát thanh này so với trọng số của tín hiệu đối tượng âm thanh, vị trí của nó được đặt cách xa loa phát thanh đã nêu (xem Fig.5). Ví dụ, giai đoạn ghép nhẵn độ cơ sở vector có thể được dùng để xác định trọng số của tín hiệu đối tượng âm thanh nằm trong mỗi trong số kênh đầu ra âm thanh (xem, ví dụ, [VBAP]). Đối với VBAP, được giả định rằng tín hiệu đối tượng âm thanh được gán cho nguồn thực và còn được giả định hơn nữa rằng kênh đầu ra âm thanh là kênh của loa phát thanh.

Trên các hình vẽ Fig.6 và Fig.8, bộ mã hoá SAOC 800 được mô tả. Bộ mã hoá SAOC 800 được sử dụng để mã hoá theo tham số số lượng đối tượng/kênh đầu vào bằng cách trộn giảm chúng đến số lượng thấp hơn của kênh truyền tải và chiết thông tin bổ trợ cần thiết mà được nhúng vào trong dòng bit âm thanh ba chiều.

Việc trộn giảm đến số lượng thấp hơn của kênh truyền tải được thực hiện bằng cách sử dụng các hệ số trộn giảm đối với mỗi tín hiệu đầu vào và kênh trộn giảm (ví dụ, bằng việc áp dụng ma trận trộn giảm).

Tình trạng kỹ thuật hiện nay để xử lý tín hiệu đối tượng âm thanh là hệ thống MPEG SAOC. Một đặc tính chính của hệ thống này là tín hiệu trộn giảm trung gian

(hoặc Các kênh Truyền tải SAOC theo các hình vẽ Fig.6 và Fig.8) có thể được nghe với thiết bị kế thừa không có khả năng giải mã thông tin SAOC. Việc này tạo giới hạn đối với các hệ số trộn giảm sẽ được sử dụng, mà thường được cung cấp bởi bộ tạo nội dung.

Hệ thống mã hóa-giải mã âm thanh ba chiều có mục đích sử dụng kỹ thuật SAOC để làm tăng hiệu quả mã hóa số lượng lớn đối tượng hoặc kênh. Việc trộn giảm số lượng lớn đối tượng vào số lượng nhỏ của kênh truyền tải tiết kiệm dòng bit.

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.

Thiết bị này bao gồm bộ trộn đối tượng 210 để tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh từ hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, sao cho hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được trộn vào tín hiệu truyền tải âm thanh, và trong đó số lượng của một hoặc nhiều kênh truyền tải âm thanh là nhỏ hơn số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

Hơn nữa, thiết bị này bao gồm giao diện đầu ra 220 để xuất ra tín hiệu truyền tải âm thanh.

Bộ trộn đối tượng 210 được tạo cấu hình để tạo ra một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh phụ thuộc vào quy tắc trộn thứ nhất và phụ thuộc vào quy tắc trộn thứ hai, trong đó quy tắc trộn thứ nhất biểu thị cách để trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh trộn trước, và trong đó quy tắc trộn thứ hai biểu thị cách để trộn nhiều kênh trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh. Quy tắc trộn thứ nhất phụ thuộc vào số lượng các đối tượng âm thanh, biểu thị số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, và phụ thuộc vào số lượng các kênh trộn trước, biểu thị số lượng của nhiều kênh trộn trước, và trong đó quy tắc trộn thứ hai phụ thuộc vào số lượng các kênh trộn trước. Giao diện đầu ra 220 được tạo cấu hình để xuất ra thông tin dựa trên quy tắc trộn thứ hai.

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

phương án.

Thiết bị này bao gồm bộ xử lý tham số 110 để tính toán thông tin trộn kênh đầu ra 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, trong đó hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được trộn vào tín hiệu truyền tải âm thanh, và trong đó số lượng của một hoặc nhiều kênh truyền tải âm thanh là nhỏ hơn số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh. Tín hiệu truyền tải âm thanh phụ thuộc vào quy tắc trộn thứ nhất và vào quy tắc trộn thứ hai. Quy tắc trộn thứ nhất biểu thị cách để trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh trộn trước. Hơn nữa, quy tắc trộn thứ hai biểu thị cách để trộn nhiều kênh trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh.

Bộ xử lý tham số 110 được tạo cấu hình để nhận thông tin dựa trên quy tắc trộn thứ hai, trong đó thông tin dựa trên quy tắc trộn thứ hai biểu thị cách để trộn nhiều tín hiệu trộn trước sao cho thu được một hoặc nhiều kênh truyền tải âm thanh. Bộ xử lý tham số 110 được tạo cấu hình để tính toán thông tin trộn kênh đầu ra phụ thuộc vào số lượng các đối tượng âm thanh biểu thị số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, phụ thuộc vào số lượng các kênh trộn trước biểu thị số lượng của nhiều kênh trộn trước, và phụ thuộc vào thông tin dựa trên quy tắc trộn thứ hai.

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 kênh đầu ra.

Theo phương án, thiết bị có thể, ví dụ được tạo cấu hình để nhận ít nhất một số lượng các đối tượng âm thanh và số lượng các kênh trộn trước.

Theo phương án khác, bộ xử lý tham số 110 có thể, ví dụ được tạo cấu hình để xác định, phụ thuộc vào số lượng các đối tượng âm thanh và phụ thuộc vào số lượng các kênh trộn trước, thông tin dựa trên quy tắc trộn thứ nhất, sao cho thông tin dựa trên quy tắc trộn thứ nhất biểu thị cách để trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh trộn trước. Theo phương án này, bộ xử lý tham số 110 có thể ví dụ được tạo cấu hình để tính toán thông tin trộn kênh đầu ra, phụ thuộc vào

thông tin dựa trên quy tắc trộn thứ nhất và phụ thuộc vào thông tin dựa trên quy tắc trộn thứ hai.

Theo phương án, bộ xử lý tham số 110 có thể, ví dụ được tạo cấu hình để xác định, phụ thuộc vào số lượng các đối tượng âm thanh và phụ thuộc vào số lượng các kênh trộn trước, nhiều hệ số của ma trận thứ nhất P làm thông tin dựa trên quy tắc trộn thứ nhất, trong đó ma trận thứ nhất P biểu thị cách để trộn nhiều kênh trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh. Theo phương án này, bộ xử lý tham số 110, có thể ví dụ được tạo cấu hình để nhận nhiều hệ số của ma trận thứ hai P làm thông tin dựa trên quy tắc trộn thứ hai, trong đó ma trận thứ hai Q biểu thị cách để trộn nhiều kênh trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh. Bộ xử lý tham số 110 của phương án này có thể, ví dụ được tạo cấu hình để tính toán thông tin trộn kênh đầu ra phụ thuộc vào ma trận thứ nhất P và phụ thuộc vào ma trận thứ hai Q.

Các phương án được dựa trên phát hiện ra rằng khi trộn giảm hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh X để thu được tín hiệu truyền tải âm thanh Y ở phía bộ mã hoá bằng cách dùng ma trận trộn giảm D theo công thức

$$Y = DX,$$

thì ma trận trộn giảm D có thể được chia thành hai ma trận nhỏ hơn P và Q theo công thức

$$D = QP.$$

Trong bản mô tả này, ma trận thứ nhất P thực hiện trộn từ tín hiệu đối tượng âm thanh X vào nhiều kênh trộn trước X_{pre} theo công thức:

$$X_{pre} = PX.$$

Ma trận thứ hai Q thực hiện trộn từ nhiều kênh trộn trước X_{pre} vào một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh Y theo công thức:

$$Y = Q X_{pre}.$$

Theo các phương án, thông tin dựa trên quy tắc trộn thứ hai, ví dụ trên hệ số của ma trận trộn thứ hai Q, được truyền đến bộ giải mã.

Các hệ số của ma trận trộn thứ nhất P không phải được truyền đến bộ giải mã. Để thay thế, bộ giải mã nhận thông tin đối với số lượng của tín hiệu đối tượng âm thanh và thông tin đối với số lượng các kênh trộn trước. Dựa vào các thông tin này, bộ giải mã có khả năng tái cấu trúc ma trận trộn thứ nhất P . Ví dụ, bộ mã hoá và bộ giải mã xác định ma trận trộn P theo cùng cách, khi trộn số lượng thứ nhất của tín hiệu đối tượng âm thanh N_{objects} vào số lượng thứ hai của kênh trộn trước N_{pre} .

Fig.3 minh hoạ hệ thống theo phương án. Hệ thống này bao gồm thiết bị 310 để tạo ra tín hiệu truyền tải âm thanh như được mô tả trên đây khi tham khảo đến Fig.2 và thiết bị 320 để tạo ra một hoặc nhiều kênh đầu ra âm thanh như được mô tả trên đây khi tham khảo đến Fig.1.

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 và thông tin dựa trên quy tắc trộn thứ hai từ thiết bị 310 để tạo ra tín hiệu truyền tải âm thanh. Hơn nữa, thiết bị 320 để tạo ra một hoặc nhiều kênh đầu ra âm thanh 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 dựa trên quy tắc trộn thứ hai.

Ví dụ, bộ xử lý tham số 110 có thể, ví dụ được tạo cấu hình để nhận thông tin siêu dữ liệu bao gồm thông tin vị trí đối với mỗi trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, và xác định thông tin dựa trên quy tắc trộn giảm thứ nhất phụ thuộc vào thông tin vị trí của mỗi trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, ví dụ bằng cách dùng giai đoạn ghép nhẵn độ cơ sở thẳng đứng. Ví dụ, bộ mã hoá cũng có truy cập đến thông tin vị trí của mỗi trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh và cũng có thể dùng giai đoạn ghép nhẵn độ cơ sở vector để xác định trọng số của tín hiệu đối tượng âm thanh trong các kênh trộn trước và bởi việc này xác định các hệ số của ma trận thứ nhất P theo cùng cách như được thực hiện sau đó bởi bộ giải mã (ví dụ, bộ mã hoá và bộ giải mã có thể giả định cùng định vị các loa phát thanh được giả định được gán cho các kênh trộn trước N_{pre}).

Bằng việc nhận các hệ số của ma trận thứ hai Q và bằng cách xác định ma trận thứ nhất P , bộ giải mã có thể xác định ma trận trộn giảm D theo $D = QP$.

Theo phương án, bộ xử lý tham số 110 có thể, ví dụ được tạo cấu hình để nhận

thông tin hiệp biến, ví dụ các hệ số của ma trận hiệp biến E (ví dụ, từ thiết bị để tạo ra tín hiệu truyền tải âm thanh), biểu thị độ chênh lệch mức đối tượng đối với mỗi trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, và có thể biểu thị một hoặc nhiều sự tương quan liên đối tượng giữa một trong số các tín hiệu đối tượng âm thanh và còn lại trong số tín hiệu đối tượng âm thanh.

Theo phương án này, bộ xử lý tham số 110 có thể được tạo cấu hình để tính toán thông tin trộn kênh đầu ra phụ thuộc vào số lượng các đối tượng âm thanh, phụ thuộc vào số lượng các kênh trộn trước, phụ thuộc vào thông tin dựa trên quy tắc trộn thứ hai, và phụ thuộc vào thông tin hiệp biến.

Ví dụ, bằng cách sử dụng ma trận hiệp biến E, tín hiệu đối tượng âm thanh X có thể được tái cấu trúc để thu được đối tượng âm thanh tái cấu trúc $\hat{X}$ bằng cách dùng công thức

$$\hat{X} = GY$$

trong đó G là ma trận ước lượng nguồn tham số với $G = E D^H (D E D^H)^{-1}$.

Sau đó, một hoặc nhiều kênh đầu ra âm thanh Z có thể được tạo ra bằng cách áp dụng ma trận kết xuất R đối với đối tượng âm thanh tái cấu trúc $\hat{X}$ theo công thức:

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

Tuy nhiên, việc tạo ra một hoặc nhiều kênh đầu ra âm thanh Z từ tín hiệu truyền tải âm thanh có thể, cũng được thực hiện ở bước đơn lẻ bằng cách dùng ma trận U theo công thức:

$$Z = UY, \quad \text{với } S = UG.$$

Ma trận S như vậy là ví dụ đối với thông tin trộn kênh đầu ra được xác định bởi bộ xử lý tham số 110.

Ví dụ, như được mô tả trên đây, mỗi hàng của ma trận kết xuất R có thể được kết hợp với một trong số các kênh đầu ra âm thanh mà sẽ được tạo ra. Mỗi hệ số trong một trong số các hàng của ma trận kết xuất R xác định trọng số của một trong số các tín hiệu đối tượng âm thanh tái cấu trúc trong kênh đầu ra âm thanh, mà dãy này của ma trận kết xuất R đề cập đến.

Theo phương án, trong đó bộ xử lý tham số 110 có thể, ví dụ được tạo cấu hình để nhận thông tin siêu dữ liệu bao gồm thông tin vị trí đối với mỗi trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, có thể ví dụ được tạo cấu hình để xác định thông tin kết xuất, ví dụ các hệ số của ma trận kết xuất R phụ thuộc vào thông tin vị trí của mỗi trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, và có thể, ví dụ được cấu hình để tính toán thông tin trộn kênh đầu ra (ví dụ ma trận S trên đây) phụ thuộc vào số lượng các đối tượng âm thanh, phụ thuộc vào số lượng các kênh trộn trước, phụ thuộc vào thông tin dựa trên quy tắc trộn thứ hai, và phụ thuộc vào thông tin kết xuất (ví dụ, ma trận kết xuất R).

Do đó, ma trận kết xuất R có thể, ví dụ phụ thuộc vào thông tin vị trí đối với mỗi tín hiệu đối tượng âm thanh trong số các tín hiệu đối tượng âm thanh được truyền vào bộ giải mã SAOC trong thông tin siêu dữ liệu. Ví dụ, tín hiệu đối tượng âm thanh có vị trí mà được định vị gần với vị trí loa phát thanh được giả định hoặc thực có thể, ví dụ có trọng số cao hơn trong kênh đầu ra âm thanh của loa phát thanh so với trọng số của tín hiệu đối tượng âm thanh, vị trí mà được định vị cách xa với loa phát thanh (xem Fig.5). Ví dụ, giai đoạn ghép nhẵn độ cơ sở vector có thể được dùng để xác định trọng số của tín hiệu đối tượng âm thanh trong mỗi trong số kênh đầu ra âm thanh (xem, ví dụ [VBAP]). Đối với VBAP, được giả định rằng tín hiệu đối tượng âm thanh được gán cho nguồn thực và được giả định hơn nữa rằng kênh đầu ra âm thanh là kênh của loa phát thanh. Hệ số tương ứng của ma trận kết xuất R (hệ số mà được gán cho kênh đầu ra âm thanh được xét đến và tín hiệu đối tượng âm thanh được xét đến) sau đó có thể được thiết lập đến trị số phụ thuộc vào trọng số. Ví dụ, tự bản thân trọng số là trị số của hệ số tương ứng trong ma trận kết xuất R .

Sau đây, các phương án thực hiện trộn giảm trong không gian đối với tín hiệu dựa trên đối tượng là được giải thích chi tiết.

Việc tham chiếu được thực hiện đến các ký hiệu và định nghĩa sau đây:

$N_{Objects}$	số lượng các tín hiệu đối tượng âm thanh đầu vào
$N_{Channels}$	số lượng các kênh đầu vào
N	số lượng các tín hiệu đầu vào;

	N có thể bằng với $N_{Objects}$, $N_{Channels}$ or $N_{Objects} + N_{Channels}$.
N_{DmxCh}	số lượng các kênh trộn giảm (được xử lý)
N_{pre}	số lượng các kênh trộn trước
$N_{Samples}$	số lượng các mẫu dữ liệu được xử lý
D	ma trận trộn giảm, kích thước $N_{DmxCh} \times N$
X	tín hiệu âm thanh đầu vào bao gồm hai hoặc nhiều hơn hai tín hiệu đầu vào âm thanh, kích thước $N \times N_{Samples}$
Y	tín hiệu âm thanh trộn giảm (tín hiệu truyền tải âm thanh), kích thước $N_{DmxCh} \times N_{Samples}$, được xác định là $Y = DX$
DMG	dữ liệu khuếch đại trộn giảm đối với mỗi tín hiệu đầu vào, kênh trộn giảm và tập hợp tham số
D_{DMG}	là ma trận ba chiều giữ dữ liệu DMG được lượng tử hoá và tạo bản đồ đối với mỗi tín hiệu đầu vào, kênh trộn giảm và tập hợp tham số

Không muốn bị mất theo cách thông thường, để cải thiện khả năng đọc của các phương trình, đối với tất cả các biến được đưa vào, chỉ số xác định sự phụ thuộc thời gian và tần số được bỏ qua.

Nếu không có các ràng buộc nào được xác định rõ xét đến các tín hiệu đầu vào (kênh hoặc đối tượng), các hệ số trộn giảm được tính toán theo cùng cách đối với các tín hiệu kênh đầu vào và tín hiệu đối tượng đầu vào. Ký hiệu đối với số lượng của tín hiệu đầu vào N được sử dụng.

Một vài phương án, có thể ví dụ được ký hiệu để trộn giảm các tín hiệu đối tượng theo cách khác với tín hiệu kênh, được dẫn bởi thông tin trong không gian sẵn có trong siêu dữ liệu đối tượng.

Trộn giảm có thể được tách thành hai bước:

- Ở bước thứ nhất, các đối tượng được kết xuất trước vào thiết kế tái tạo với số lượng loa phát thanh nhiều nhất N_{pre} (ví dụ, $N_{pre} = 22$ được đưa ra bởi cấu hình

22.2). Ví dụ, ma trận thứ nhất P có thể được dùng.

- Ở bước thứ hai, N_{pre} thu được biểu diễn tín hiệu được trộn giảm đến số lượng của kênh truyền tải sẵn có (N_{DmxCh}) (ví dụ, theo thuật toán phân bố trộn giảm trực giao). Ví dụ, ma trận thứ hai Q có thể được dùng.

Tuy nhiên, theo một vài phương án, trộn giảm được thực hiện ở bước đơn lẻ, ví dụ bằng cách dùng ma trận D được xác định theo công thức: $D = QP$, và bằng cách áp dụng $Y = DX$ với $D = QP$.

Không kể các yếu tố khác, lợi ích khác của các khái niệm được đề xuất là, ví dụ ở chỗ tín hiệu đối tượng đầu vào mà được cho là được kết xuất ở cùng vị trí trong không gian, trong ngữ cảnh âm thanh, được trộn giảm cùng với nhau trong cùng kênh truyền tải. Do đó, ở phía bộ giải mã, việc chia tách tốt hơn của tín hiệu kết xuất trước là thu được, ngăn ngừa việc chia tách của đối tượng âm thanh mà sẽ được trộn lại cùng với nhau trong ngữ cảnh tái tạo cuối cùng.

Theo phương án được ưu tiên đặc biệt, trộn giảm có thể được mô tả như phép nhân ma trận theo:

$$X_{pre} = PX \quad \text{và} \quad Y = QX_{pre}.$$

trong đó P có kích thước ($N_{pre} \times N_{Objects}$) và Q có kích thước ($N_{DmxCh} \times N_{pre}$) được tính toán như được giải thích sau đây.

Hệ số trộn trong P được cấu trúc từ siêu dữ liệu tín hiệu đối tượng (bán kính, khuếch đại, góc phương vị và góc nâng) sử dụng thuật toán ghép nhẵn (ví dụ, ghép nhẵn độ cơ sở vector). Thuật toán ghép nhẵn sẽ là giống với thuật toán được sử dụng ở phía bộ giải mã để cấu trúc các kênh đầu ra.

Các hệ số trộn trong Q được đưa ra ở phía bộ mã hoá đối với tín hiệu đầu vào N_{pre} và kênh truyền tải sẵn có N_{DmxCh} .

Để làm giảm độ phức tạp tính toán, trộn giảm hai bước có thể được đơn giản hoá thành một bước bằng cách tính toán độ khuếch đại trộn giảm cuối cùng như:

$$D = QP.$$

Sau đó, tín hiệu trộn giảm được đưa ra bởi:

$$Y = DX.$$

Các hệ số trộn trong P là không được truyền trong dòng bit. Để thay thế, chúng được tái cấu trúc ở phía bộ giải mã sử dụng cùng thuật toán đại. Do đó, tốc độ bit được giảm bằng cách chỉ gửi các hệ số trộn trong Q. Cụ thể là, như các hệ số trộn trong P là biến thể thời gian thông thường và P là không được truyền, có thể đạt được việc giảm tốc độ bit cao.

Sau đây, cú pháp dòng bit theo phương án được xét đến.

Để báo hiệu phương pháp trộn giảm được sử dụng và số lượng kênh N_{pre} để kết xuất trước đối tượng ở bước thứ nhất, cú pháp dòng bit MPEG SAOC được mở rộng với 4 bit:

bsSaocDmxMethod	Chế độ	Nghĩa
0	Chế độ trực tiếp	Ma trận trộn giảm được cấu trúc trực tiếp từ DMG giải lượng tử hoá (khuếch đại trộn giảm).
1,..., 15	Chế độ trộn trước	Ma trận trộn giảm được cấu trúc như sản phẩm của ma trận thu được từ DMG giải lượng tử hoá và ma trận trộn trước thu được từ các thông tin trong không gian của đối tượng âm thanh đầu vào.

bsNumPremixedChannels

bsSaocDmxMethod	bsNumPremixedChannels
0	0
1	22
2	11
3	10
4	8

5	7
6	5
7	2
8,..., 14	dự trữ
15	trị số biến mất

Trong nội dung của MPEG SAOC, việc này có thể được hoàn thành bởi cái biến sau đây:

bsSaocDmxMethod : biểu thị cách mà ma trận trộn giảm được cấu trúc

Syntax of SAOC3DSpecificConfig() - Signaling

bsSaocDmxMethod;	4	uimsbf
if (bsSaocDmxMethod == 15) {		
bsNumPremixedChannels;	5	uimsbf
}		

Syntax of Saoc3DFrame(): cách mà DMG được đọc đối với các chế độ phức

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);
        }
    } else {
        for( i= dmgIdx; i<dmgIdx + bsNumSaocDmxObjects; i++ )
        {
            idxDMG[i] = EcDataSaoc(DMG, 0,
bsNumPremixedChannels);
        }
    }
}

```

bsNumSaocDmxChannels

xác định số lượng kênh trộn giảm đối với nội dung dựa trên kênh. Nếu không có kênh nào có mặt trong trộn giảm bsNumSaocDmxChannels được thiết lập đến 0.

bsNumSaocChannels

xác định số lượng kênh đầu vào đối với các tham số SAOC 3D được truyền. Nếu bsNumSaocChannels = 0 không có kênh nào có mặt trong trộn giảm.

bsNumSaocDmxObjects

xác định số lượng kênh trộn giảm đối với nội dung dựa trên đối tượng. Nếu không có đối tượng nào có mặt trong trộn giảm bsNumSaocDmxObjects được thiết lập đến 0.

bsNumPremixedChannels

xác định số lượng các kênh trộn trước đối với đối tượng âm thanh đầu vào. Nếu bsSaocDmxMethod bằng 15 thì số lượng các kênh trộn trước thực được

báo hiệu trực tiếp bởi trị số `bsNumPremixedChannels`. Trong tất cả các trường hợp khác, `bsNumPremixedChannels` được thiết lập theo bản trên đây.

Theo phương án, ma trận trộn giảm $\mathbf{D}$ được áp dụng cho tín hiệu âm thanh đầu vào $\mathbf{S}$ xác định tín hiệu trộn giảm như

$$\mathbf{X} = \mathbf{D}\mathbf{S}.$$

Ma trận trộn giảm $\mathbf{D}$ có kích thước $N_{\text{dmx}} \times N$ thu được là:

$$\mathbf{D} = \mathbf{D}_{\text{dmx}} \mathbf{D}_{\text{premix}}.$$

Ma trận $\mathbf{D}_{\text{dmx}}$ và ma trận $\mathbf{D}_{\text{premix}}$ có kích thước khác nhau phụ thuộc vào chế độ xử lý.

Ma trận $\mathbf{D}_{\text{dmx}}$ thu được từ các tham số DMG là:

$$d_{i,j} = \begin{cases} 0 & \text{nếu không có dữ liệu DMG đối với cặp (i,j) là có mặt trong dòng bit} \\ 10^{0.05 \text{DMG}_{i,j}} & \text{thì} \end{cases},$$

Trong bản mô tả này, tham số trộn giảm giải lượng tử hoá thu được là:

$$\text{DMG}_{i,j} = \mathbf{D}_{\text{DMG}}(i, j, l).$$

Trong trường hợp chế độ trực tiếp, không có việc trộn trước nào được sử dụng. Ma trận $\mathbf{D}_{\text{premix}}$ có kích thước $N \times N$ và được đưa ra bởi: $\mathbf{D}_{\text{premix}} = \mathbf{I}$. Ma trận $\mathbf{D}_{\text{dmx}}$ có kích thước $N_{\text{dmx}} \times N$ và thu được từ các tham số DMG.

Trong trường hợp chế độ trộn trước, ma trận $\mathbf{D}_{\text{premix}}$ có kích thước $(N_{\text{ch}} + N_{\text{premix}}) \times N$ và được đưa ra bởi:

$$\mathbf{D}_{\text{premix}} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} \end{pmatrix},$$

trong đó ma trận trộn trước $\mathbf{A}$ có kích thước $N_{\text{premix}} \times N_{\text{obj}}$ được nhận như được nhập vào bộ giải mã ba chiều SAOC, từ bộ kết xuất đối tượng.

Ma trận D_{dmx} có kích thước $N_{dmx} \times (N_{ch} + N_{premix})$ và thu được từ các tham số DMG.

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 nội dung 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 liên mạng.

Tùy thuộc vào các yêu cầu thực thi nhất định, các phương án theo sáng chế có thể được thực hiện trong phần cứng hoặc phần mềm. Việc thực thi 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 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 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 làm 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 số các phương pháp được mô tả ở đây, được lưu trữ trên vật mang đọc được bằng máy.

Nói cách khác, phương án của 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 hàng 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. Hàng 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 liên mạng.

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 thích ứng, thực hiện một trong 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 các phương pháp được mô tả ở đây.

Theo một số phương án, thiết bị logic lập trình được (ví dụ, mảng công lập trình trường) có thể được sử dụng để thực hiện một vài hoặc tất cả các chức năng của các phương pháp được mô tả ở đây. Theo một số phương án, mảng công lập trình 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 phần 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 này bao gồm:

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

bộ xử lý trộn giảm (120) để tạo một hoặc nhiều kênh đầu ra âm thanh, trong đó 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, trong đó hai 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 hai hoặc nhiều tín hiệu đối tượng âm thanh,

trong đó tín hiệu truyền tải âm thanh phụ thuộc vào quy tắc trộn thứ nhất và quy tắc trộn thứ hai, trong đó quy tắc trộn thứ nhất biểu thị cách trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh được trộn trước, và trong đó quy tắc trộn thứ hai biểu thị cách trộn nhiều kênh được trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh,

trong đó bộ xử lý tham số (110) được tạo cấu hình để nhận thông tin về quy tắc trộn thứ hai, trong đó thông tin về quy tắc trộn thứ hai biểu thị cách trộn nhiều tín hiệu được trộn trước sao cho một hoặc nhiều kênh truyền tải âm thanh được thu nhận.

trong đó bộ xử lý tham số (110) được tạo cấu hình để tính toán thông tin trộn kênh đầu ra phụ thuộc vào số lượng đối tượng âm thanh biểu thị số lượng gồm hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, phụ thuộc vào số lượng kênh được trộn trước biểu thị số lượng của nhiều kênh được trộn trước, và phụ thuộc vào thông tin về quy tắc trộn thứ hai, 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 kênh đầu ra.

2. Thiết bị theo điểm 1, trong đó thiết bị này được tạo cấu hình để nhận ít nhất một đối tượng âm thanh trong số số lượng đối tượng âm thanh và số lượng kênh được trộn trước.

3. Thiết bị theo điểm 1,

trong đó bộ xử lý tham số (110) được tạo cấu hình để xác định, phụ thuộc vào

số lượng đối tượng âm thanh và phụ thuộc vào số lượng các kênh được trộn trước, thông tin về quy tắc trộn thứ nhất, ví dụ như thông tin về quy tắc trộn thứ nhất biểu thị cách trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh được trộn trước, và

trong đó bộ xử lý tham số (110) được tạo cấu hình để tính toán thông tin trộn kênh đầu ra, phụ thuộc vào thông tin về quy tắc trộn thứ nhất và phụ thuộc vào thông tin về quy tắc trộn thứ hai.

4. Thiết bị theo điểm 3,

trong đó bộ xử lý tham số (110) được tạo cấu hình để xác định, phụ thuộc vào số lượng đối tượng âm thanh và phụ thuộc vào số lượng kênh được trộn trước, nhiều hệ số của ma trận thứ nhất (P) như là thông tin về quy tắc trộn thứ nhất, trong đó ma trận thứ nhất (P) biểu thị cách trộn hai hoặc nhiều hơn hai kênh đối tượng âm thanh để thu được nhiều kênh trộn trước,

trong đó bộ xử lý tham số (110) được tạo cấu hình để nhận nhiều hệ số gồm ma trận thứ hai (Q) như là thông tin về quy tắc trộn thứ hai, trong đó ma trận thứ hai (Q) biểu thị cách trộn nhiều kênh được trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh, và

trong đó bộ xử lý tham số (110) được tạo cấu hình để tính toán thông tin trộn kênh đầu ra phụ thuộc vào ma trận thứ nhất (P) và phụ thuộc vào ma trận thứ hai (Q).

5. Thiết bị theo điểm 3,

trong đó bộ xử lý tham số (110) được tạo cấu hình để nhận thông tin siêu dữ liệu bao gồm thông tin vị trí cho mỗi tín hiệu đối tượng âm thanh trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh,

trong đó bộ xử lý tham số (110) được tạo cấu hình để xác định thông tin về quy tắc trộn thứ nhất phụ thuộc vào thông tin vị trí của mỗi tín hiệu đối tượng âm thanh trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

6. Thiết bị theo điểm 1,

trong đó bộ xử lý tham số (110) được tạo cấu hình để nhận thông tin siêu dữ liệu bao gồm thông tin vị trí cho mỗi tín hiệu đối tượng âm thanh trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh,

trong đó bộ xử lý tham số (110) được tạo cấu hình để xác định thông tin về quy

tắc trộn thứ nhất phụ thuộc vào thông tin vị trí của mỗi tín hiệu đối tượng âm thanh trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

7. Thiết bị theo điểm 5,

trong đó bộ xử lý tham số (110) được tạo cấu hình để xác định thông tin kết xuất phụ thuộc vào thông tin vị trí của mỗi tín hiệu đối tượng âm thanh trong số hai hoặc nhiều hơn hai 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 kênh đầu ra phụ thuộc vào số lượng đối tượng âm thanh, phụ thuộc vào số lượng kênh được trộn trước, phụ thuộc vào thông tin về quy tắc trộn thứ hai và phụ thuộc vào thông tin kết xuất.

8. Thiết bị theo điểm 1,

trong đó bộ xử lý tham số (110) được tạo cấu hình để nhận thông tin hiệp biến biểu thị chênh lệch mức đối tượng cho mỗi tín hiệu đối tượng âm thanh trong số hai hoặc nhiều hơn hai 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 kênh đầu ra phụ thuộc vào số lượng đối tượng âm thanh, phụ thuộc vào số lượng kênh được trộn trước, phụ thuộc vào thông tin về quy tắc trộn thứ hai và phụ thuộc vào thông tin hiệp biến.

9. Thiết bị theo điểm 8,

trong đó thông tin hiệp biến còn biểu thị ít nhất một tương quan liên đối tượng giữa một tín hiệu đối tượng âm thanh trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh và tín hiệu đối tượng âm thanh khác trong số hai hoặc nhiều hơn hai 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 kênh đầu ra phụ thuộc vào số lượng đối tượng âm thanh, phụ thuộc vào số lượng kênh được trộn trước, phụ thuộc vào thông tin về quy tắc trộn thứ hai, phụ thuộc vào chênh lệch mức đối tượng của mỗi tín hiệu đối tượng âm thanh trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh và phụ thuộc vào ít nhất một tương quan liên đối tượng giữa một tín hiệu đối tượng âm thanh trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh và tín hiệu đối tượng âm thanh khác trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

10. Thiết bị tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh, trong đó thiết bị này bao gồm:

bộ trộn đối tượng (210) để tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh từ hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, sao cho hai hoặc nhiều hơn hai 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 hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, và

giao diện đầu ra (220) để xuất ra tín hiệu truyền tải âm thanh, trong đó thiết bị được tạo cấu hình để truyền tín hiệu truyền tải âm thanh tới bộ giải mã,

trong đó bộ trộn đối tượng (210) được tạo cấu hình để tạo ra một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh phụ thuộc vào quy tắc trộn thứ nhất và quy tắc trộn thứ hai, trong đó quy tắc trộn thứ nhất biểu thị cách trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh được trộn trước, và trong đó quy tắc trộn thứ hai biểu thị cách trộn nhiều kênh được trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh,

trong đó quy tắc trộn thứ nhất phụ thuộc vào số lượng đối tượng âm thanh, biểu thị số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh và phụ thuộc vào số lượng kênh trộn trước, biểu thị số lượng của nhiều kênh được trộn trước, và trong đó quy tắc trộn thứ hai phụ thuộc vào số lượng kênh được trộn trước, và

trong đó bộ trộn đối tượng (210) được tạo cấu hình để tạo ra một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh phụ thuộc vào ma trận thứ nhất (P), trong đó ma trận thứ nhất (P) biểu thị cách trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh được trộn trước, và phụ thuộc vào ma trận thứ hai (Q), trong đó ma trận thứ hai (Q) biểu thị cách trộn nhiều kênh được trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh,

trong đó các hệ số thứ nhất của ma trận thứ nhất (P) biểu thị thông tin về quy tắc trộn thứ nhất và trong đó các hệ số thứ hai của ma trận thứ hai (Q) biểu thị thông tin về quy tắc trộn thứ hai,

trong đó thiết bị được tạo cấu hình để truyền các hệ số thứ hai của ma trận trộn thứ hai (Q) tới bộ giải mã, và trong đó thiết bị được tạo cấu hình để truyền các hệ số

thứ nhất của ma trận trộn thứ nhất (P) tới bộ giải mã.

11. Thiết bị theo điểm 10,

trong đó bộ trộn đối tượng (210) được tạo cấu hình để nhận thông tin vị trí cho mỗi tín hiệu đối tượng âm thanh của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, và

trong đó bộ trộn đối tượng (210) được tạo cấu hình để xác định quy tắc trộn thứ nhất phụ thuộc vào thông tin vị trí của mỗi tín hiệu đối tượng âm thanh trong số hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh.

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

thiết bị tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh, trong đó thiết bị bao gồm:

bộ trộn đối tượng để tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh từ hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, sao cho hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được trộn trong tín hiệu truyền tải âm thanh, và trong đó số lượng của một hoặc nhiều các kênh truyền tải âm thanh nhỏ hơn số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, và

giao diện đầu ra để xuất ra tín hiệu truyền tải âm thanh, trong đó thiết bị được tạo cấu hình để truyền tín hiệu truyền tải âm thanh tới bộ giải mã,

trong đó bộ trộn đối tượng được tạo cấu hình để tạo ra một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh phụ thuộc vào quy tắc trộn thứ nhất và quy tắc trộn thứ hai, trong đó quy tắc trộn thứ nhất biểu thị cách trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh được trộn trước, và trong đó quy tắc trộn thứ hai biểu thị cách trộn nhiều kênh được trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh,

trong đó quy tắc trộn thứ nhất phụ thuộc vào số lượng đối tượng âm thanh, biểu thị số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh và phụ thuộc vào số lượng kênh trộn trước, biểu thị số lượng của nhiều kênh được trộn trước, và trong đó quy tắc trộn thứ hai phụ thuộc vào số lượng kênh được trộn trước, và trong đó bộ trộn đối tượng được tạo cấu hình để tạo ra một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh phụ thuộc vào ma trận thứ nhất, trong đó ma trận thứ

nhất biểu thị cách trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh được trộn trước, và phụ thuộc vào ma trận thứ hai, trong đó ma trận thứ hai biểu thị cách trộn nhiều kênh được trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh,

trong đó các hệ số thứ nhất của ma trận thứ nhất biểu thị thông tin về quy tắc trộn thứ nhất và trong đó các hệ số thứ hai của ma trận thứ hai biểu thị thông tin về quy tắc trộn thứ hai,

trong đó thiết bị được tạo cấu hình để truyền các hệ số thứ hai của ma trận trộn thứ hai tới bộ giải mã, và trong đó thiết bị được tạo cấu hình để không truyền các hệ số thứ nhất của ma trận trộn thứ nhất tới bộ giải mã, và

thiết bị tạo ra một hoặc nhiều kênh đầu ra, trong đó thiết bị bao gồm:

bộ xử lý tham số để tính toán thông tin trộn kênh đầu ra, và

bộ xử lý trộn giảm để tạo một hoặc nhiều kênh đầu ra âm thanh, trong đó 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, trong đó hai 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 hai hoặc nhiều tín hiệu đối tượng âm thanh,

trong đó tín hiệu truyền tải âm thanh phụ thuộc vào quy tắc trộn thứ nhất và quy tắc trộn thứ hai, trong đó quy tắc trộn thứ nhất biểu thị cách trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh được trộn trước, và trong đó quy tắc trộn thứ hai biểu thị cách trộn nhiều kênh được trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh,

trong đó bộ xử lý tham số được tạo cấu hình để nhận thông tin về quy tắc trộn thứ hai, trong đó thông tin về quy tắc trộn thứ hai biểu thị cách trộn nhiều tín hiệu được trộn trước sao cho một hoặc nhiều kênh truyền tải âm thanh được thu được.

trong đó bộ xử lý tham số được tạo cấu hình để tính toán thông tin trộn kênh đầu ra phụ thuộc vào số lượng đối tượng âm thanh biểu thị số lượng gồm hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, phụ thuộc vào số lượng kênh được trộn trước biểu thị số lượng của nhiều kênh được trộn trước, và phụ thuộc vào thông tin về quy tắc trộn thứ hai, và

trong đó 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 kênh đầu ra.

trong đó 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 và thông tin về quy tắc trộn thứ hai từ thiết bị để tạo ra tín hiệu truyền tải âm thanh, và

trong đó thiết bị để tạo ra một hoặc nhiều kênh đầu ra âm thanh 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 về quy tắc trộn thứ hai.

13. Phương pháp tạo ra một hoặc nhiều kênh đầu ra âm thanh, trong đó phương pháp này bao gồm:

nhận tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh, 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 số lượng của một hoặc nhiều kênh truyền tải âm thanh nhỏ hơn số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, trong đó tín hiệu truyền tải âm thanh phụ thuộc vào quy tắc trộn thứ nhất và quy tắc trộn thứ hai, trong đó quy tắc trộn thứ nhất biểu thị cách để trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh được trộn trước và trong đó quy tắc trộn thứ hai biểu thị cách để trộn nhiều kênh được trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh,

nhận thông tin về quy tắc trộn thứ hai, trong đó thông tin về quy tắc trộn thứ hai biểu thị cách trộn nhiều tín hiệu được trộn trước sao cho một hoặc nhiều kênh truyền tải âm thanh được thu được,

tính toán thông tin trộn kênh đầu ra phụ thuộc vào số lượng đối tượng âm thanh biểu thị số lượng gồm hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, phụ thuộc vào số lượng kênh được trộn trước biểu thị số lượng của nhiều kênh được trộn trước, và phụ thuộc vào thông tin về quy tắc trộn thứ hai, và

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

14. Phương pháp tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh trong đó phương pháp bao gồm:

tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh từ hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh,

xuất ra tín hiệu truyền tải âm thanh, và truyền tín hiệu truyền tải âm thanh tới bộ giải mã, và

truyền các hệ số thứ hai của ma trận trộn thứ hai (Q) tới bộ giải mã, và không truyền các hệ số thứ nhất của ma trận trộn thứ nhất (P) tới bộ giải mã,

trong đó việc tạo ra tín hiệu truyền tải âm thanh bao gồm một hoặc nhiều kênh truyền tải âm thanh từ hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được tiến hành sao cho hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh được trộn trong tín hiệu truyền tải âm thanh, và trong đó số lượng của một hoặc nhiều các kênh truyền tải âm thanh nhỏ hơn số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh, và

trong đó việc tạo ra một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh được tiến hành phụ thuộc vào quy tắc trộn thứ nhất và phụ thuộc vào quy tắc trộn thứ hai, trong đó quy tắc trộn thứ nhất biểu thị cách để trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu nhiều kênh được trộn trước và trong đó quy tắc trộn thứ hai biểu thị cách trộn nhiều kênh được trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh, trong đó quy tắc trộn thứ nhất phụ thuộc vào số lượng đối tượng âm thanh, biểu thị số lượng của hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh và phụ thuộc vào số lượng kênh được trộn trước, biểu thị số lượng của nhiều kênh được trộn trước, và trong đó quy tắc trộn thứ hai phụ thuộc vào số lượng kênh được trộn trước,

trong đó việc tạo ra một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh phụ thuộc vào ma trận thứ nhất (P), trong đó ma trận thứ nhất (P) biểu thị cách trộn hai hoặc nhiều hơn hai tín hiệu đối tượng âm thanh để thu được nhiều kênh được trộn trước, và phụ thuộc vào ma trận thứ hai (Q), trong đó ma trận thứ hai (Q) biểu thị cách trộn nhiều kênh được trộn trước để thu được một hoặc nhiều kênh truyền tải âm thanh của tín hiệu truyền tải âm thanh,

trong đó các hệ số thứ nhất của ma trận thứ nhất (P) biểu thị thông tin về quy tắc trộn thứ nhất và trong đó các hệ số thứ hai của ma trận thứ hai (Q) biểu thị thông tin về quy tắc trộn thứ hai,

15. Vật ghi lưu trữ dạng số bất khả biến có mã đọc được bằng máy tính được lưu

trữ trên đó để thực hiện phương pháp theo điểm 13 khi vật ghi lưu trữ này được thực hiện bởi máy tính hoặc bộ xử lý tín hiệu.

16. Vật ghi lưu trữ dạng số bất khả biến có mã đọc được bằng máy tính được lưu trữ trên đó để thực hiện phương pháp theo điểm 14 khi vật ghi lưu trữ này được thực hiện bởi máy tính hoặc bộ xử lý tín hiệu.

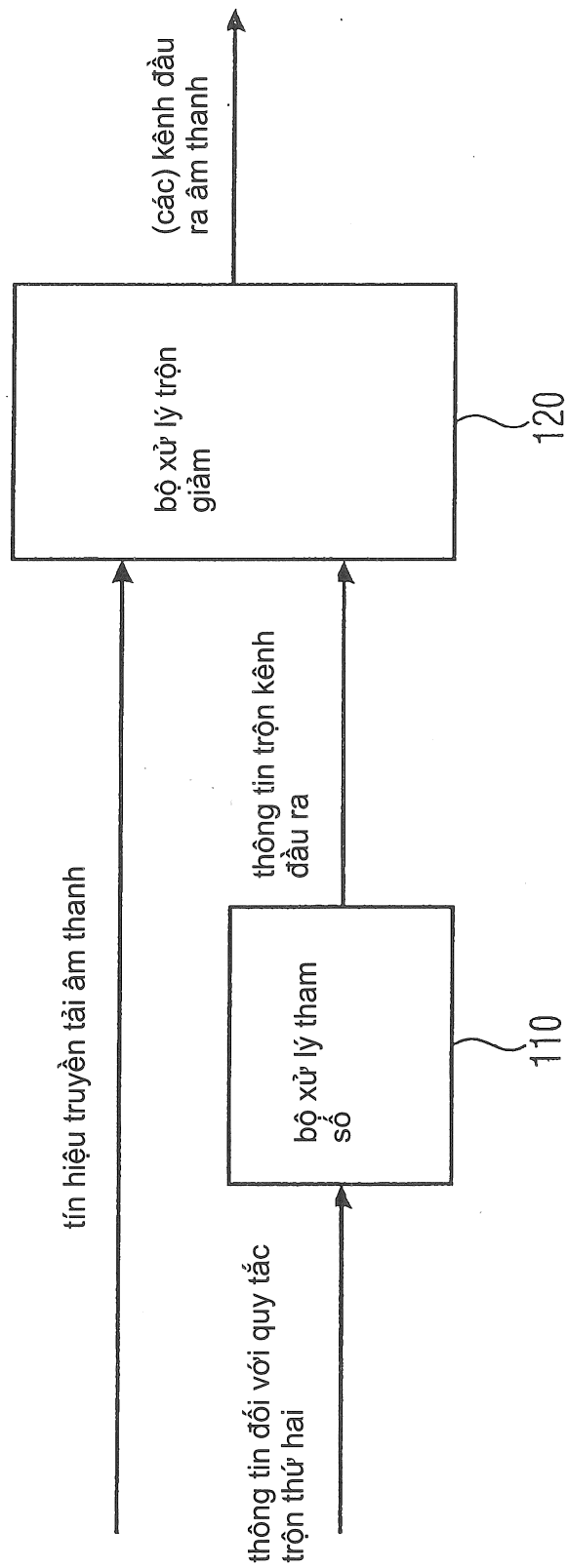


Fig.1

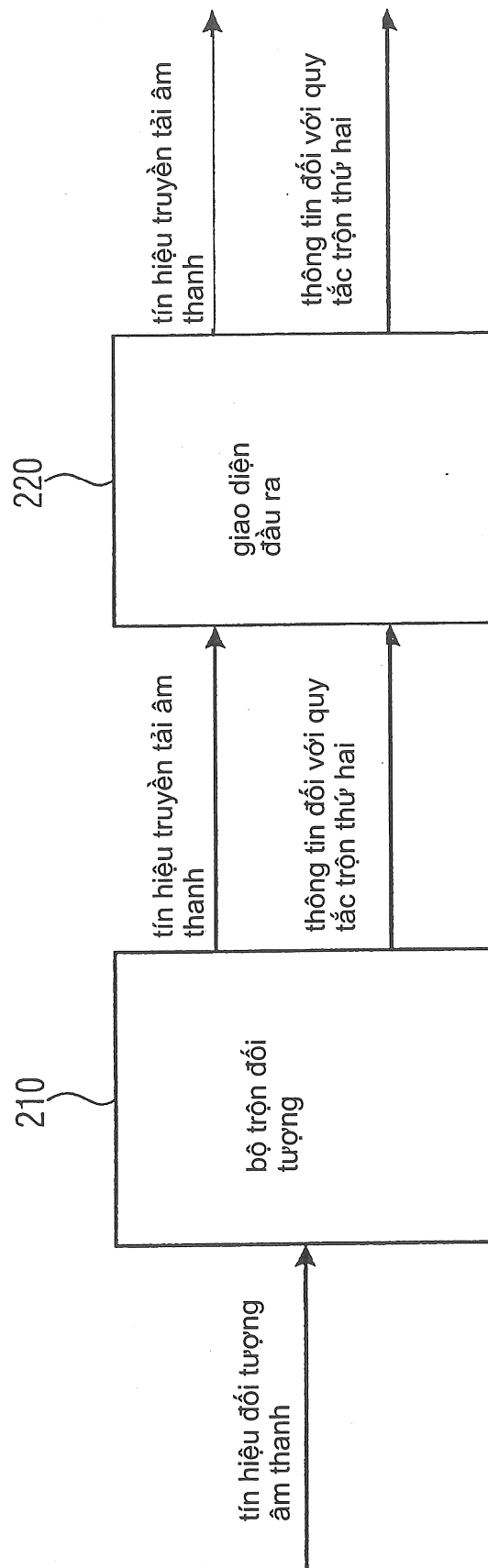


Fig.2

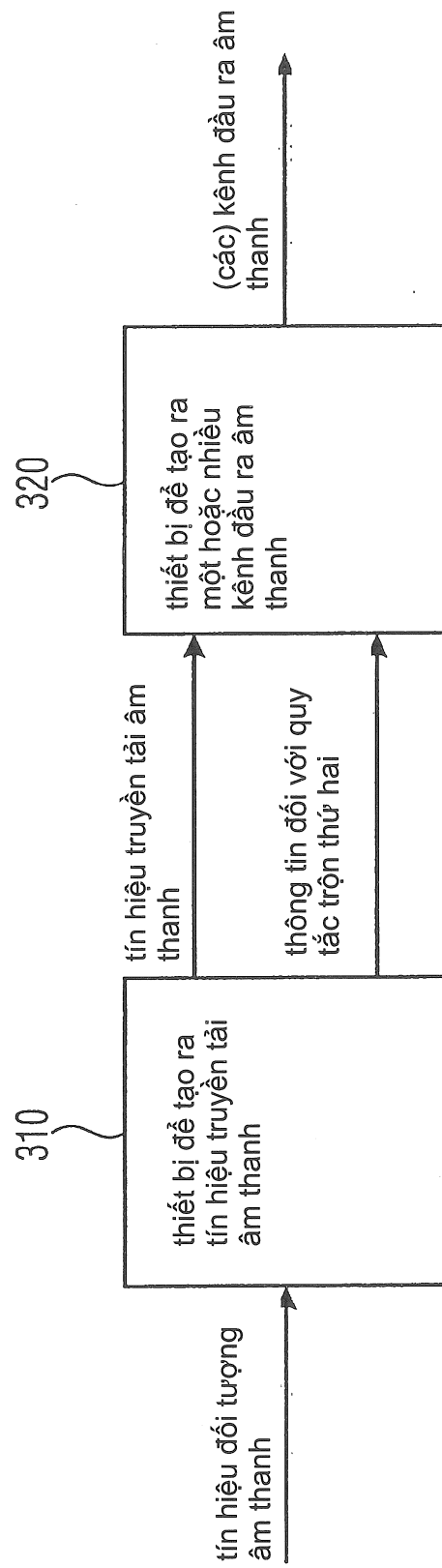
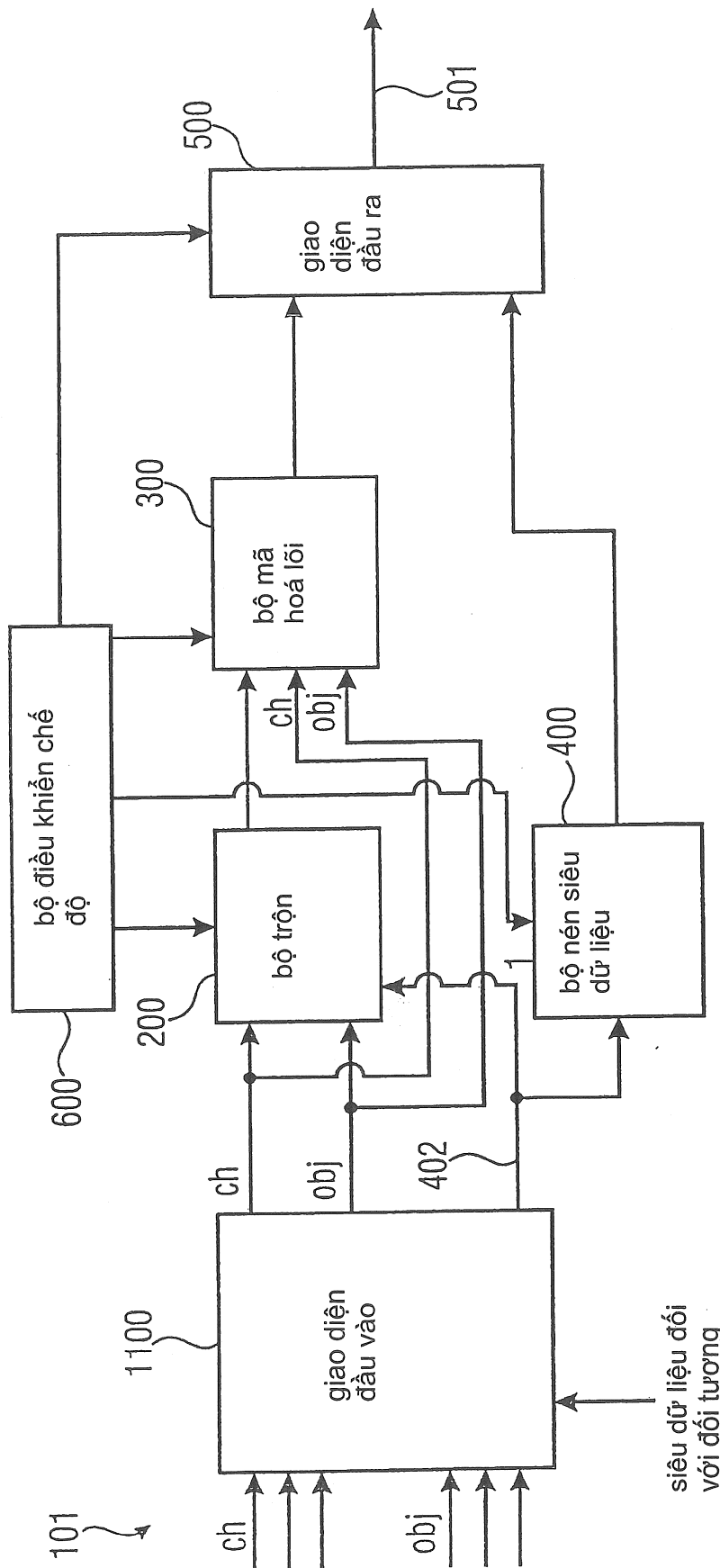
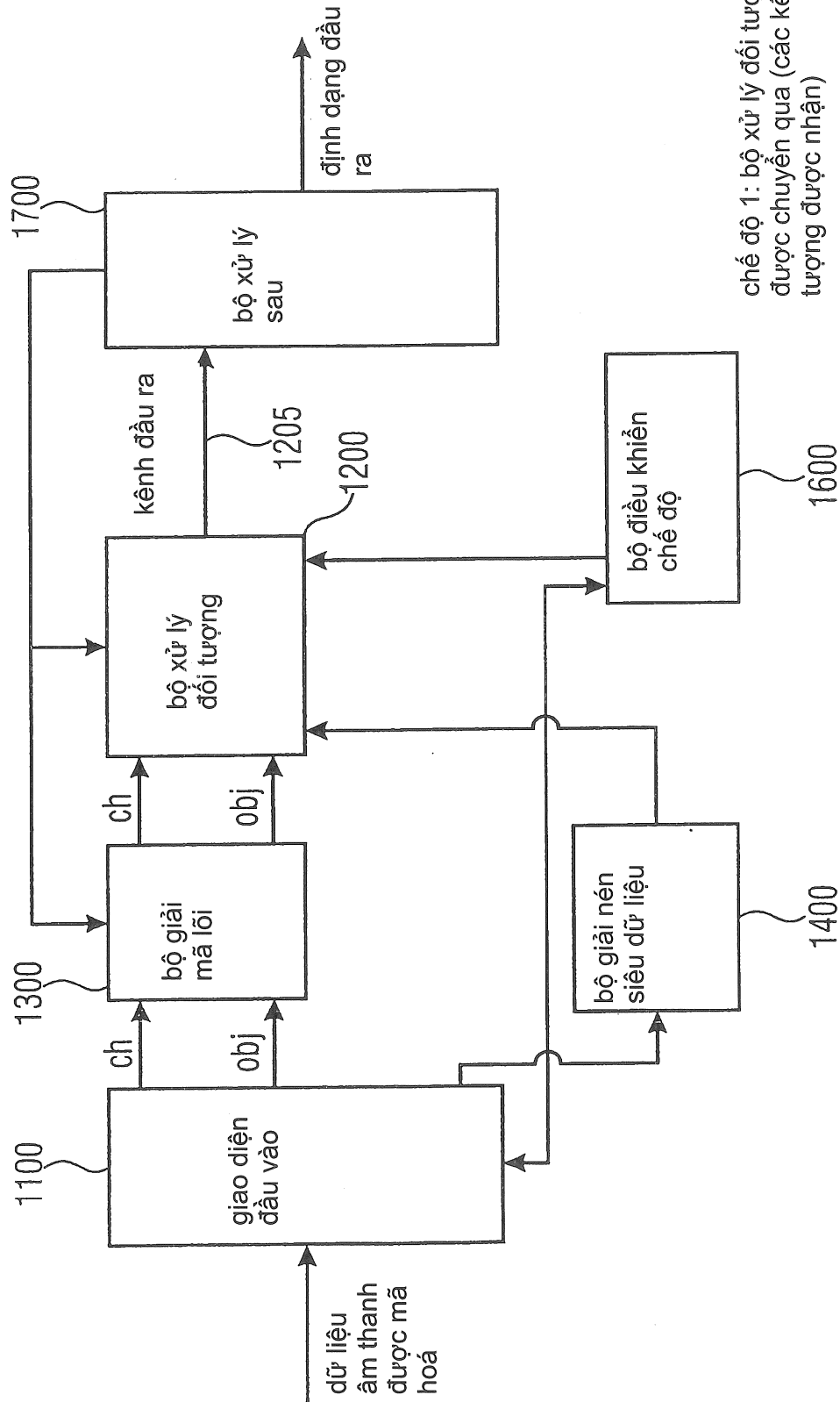


Fig.3



chế độ 1: lập mã kênh/đối tượng riêng rẽ
chế độ 2: trộn các kênh và các đối tượng kết xuất

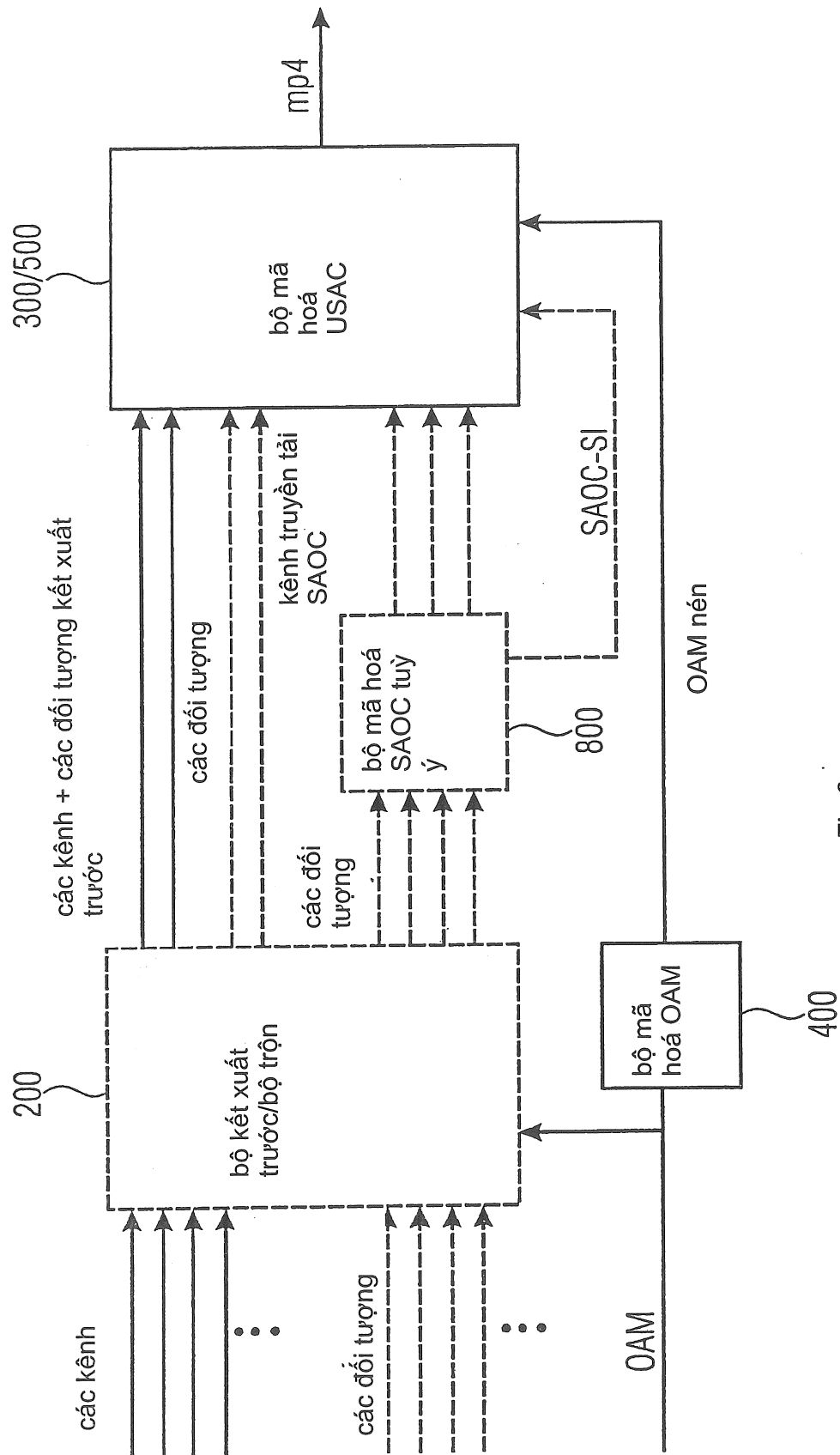
Fig.4
(Bộ mã hoá)



chế độ 1: bộ xử lý đối tượng là không được chuyển qua (các kênh và đối tượng được nhận)

chế độ 2: bộ xử lý đối tượng được chuyển qua (chỉ có các kênh được kết xuất trước được nhận)

Fig.5
(Bộ giải mã)

Fig.6
(Bộ mã hoá)

loa phát thanh 32ch

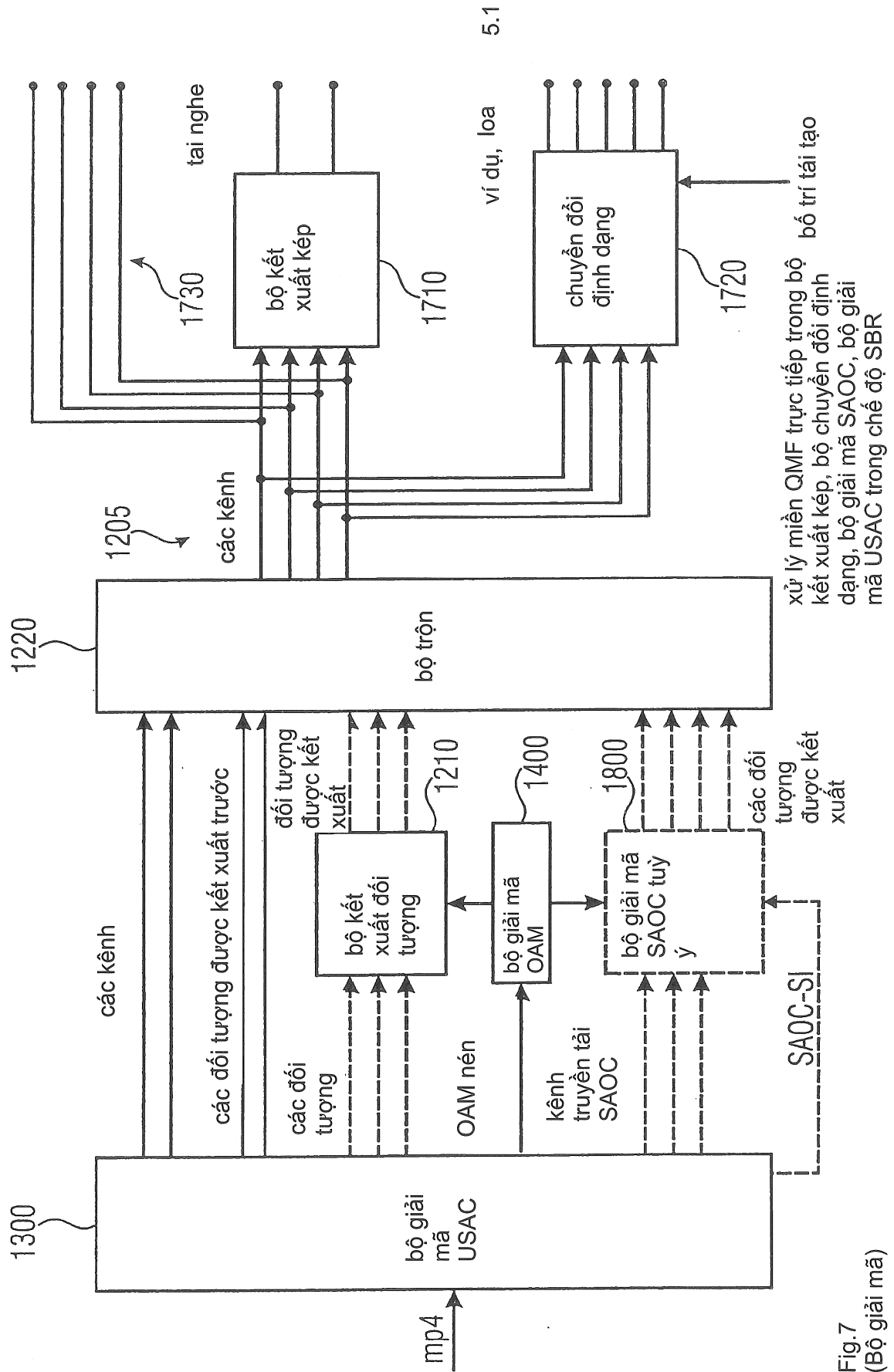
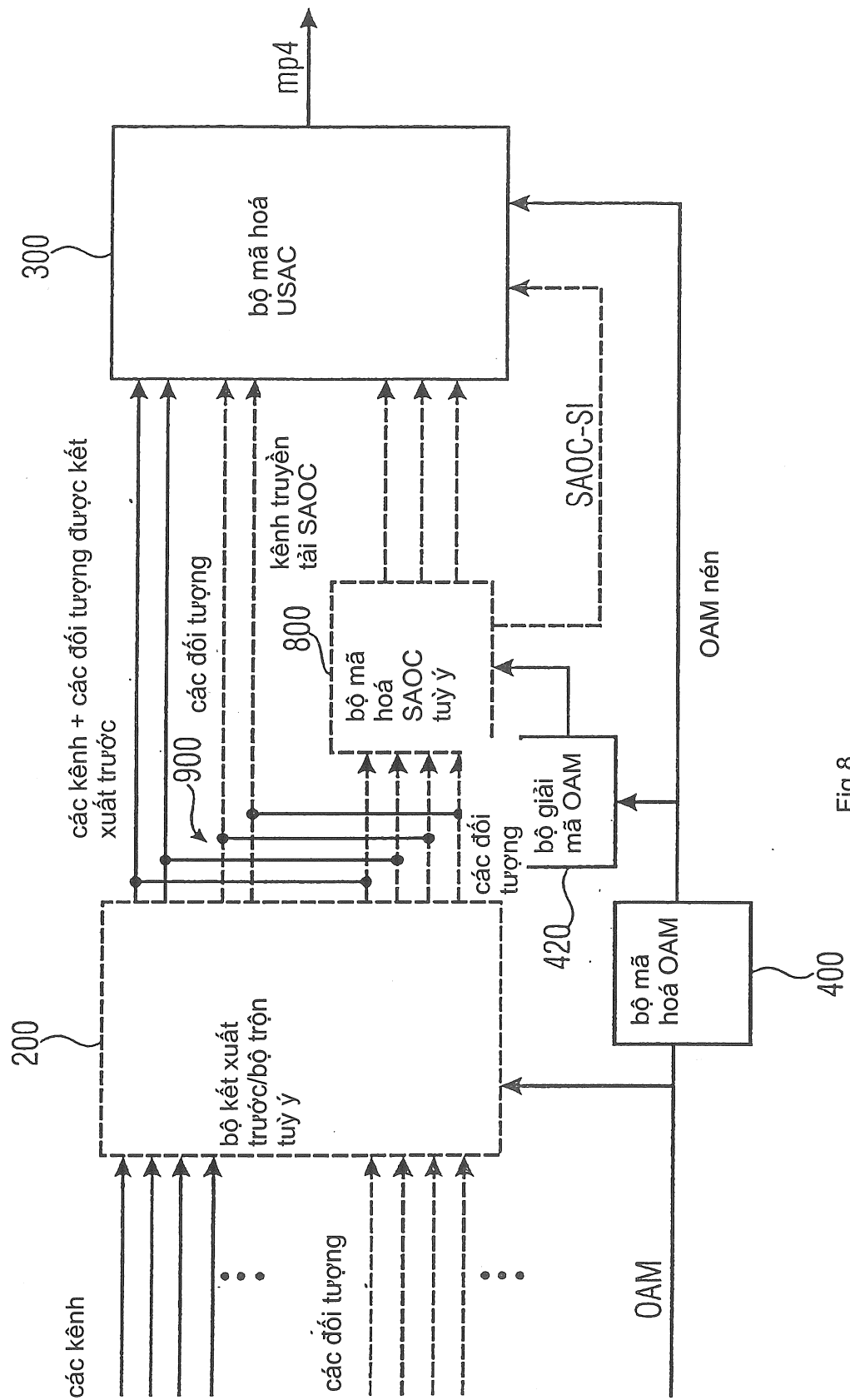
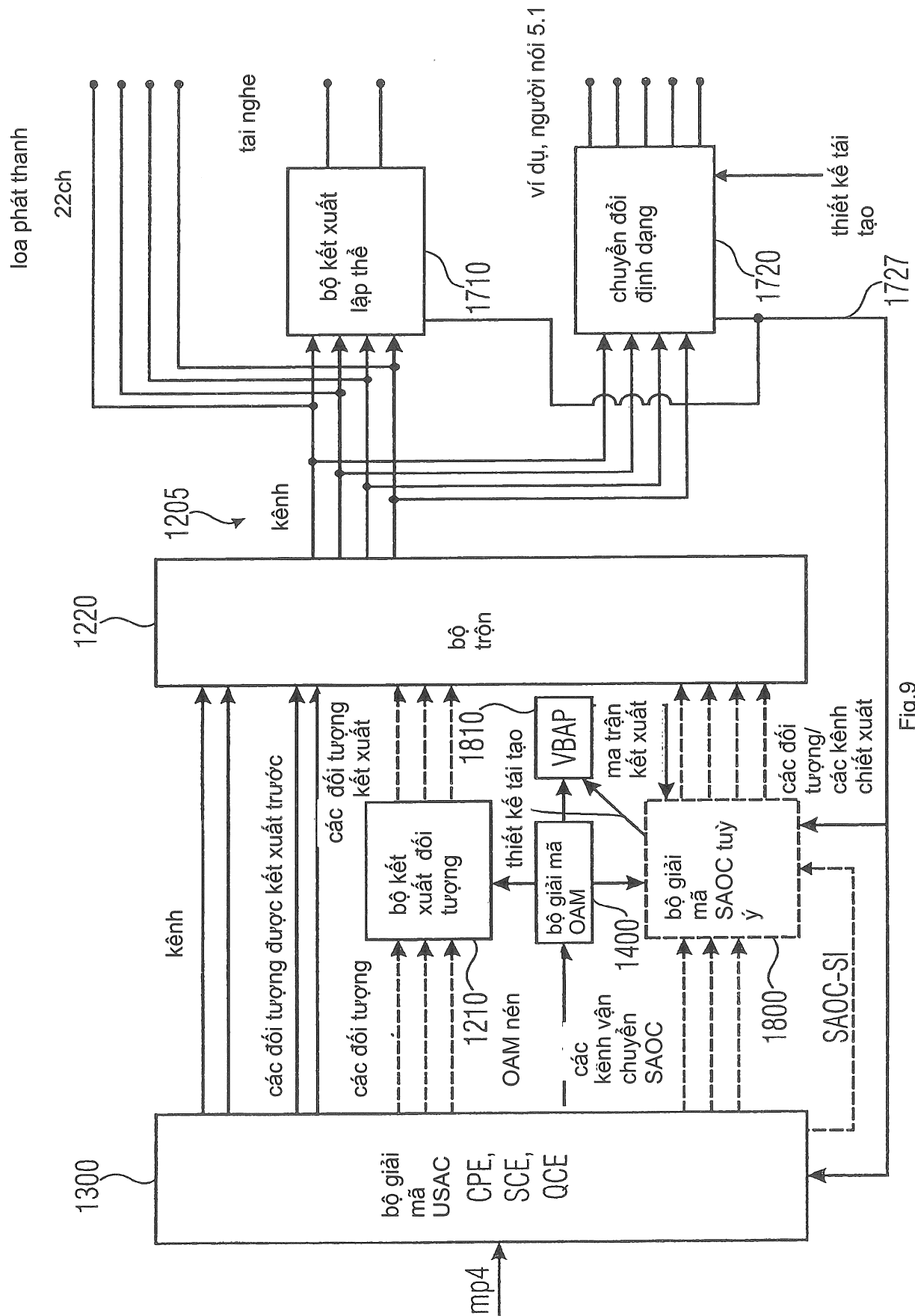


Fig.7
(Bộ giải mã)

Fig.8
(Bộ mã hoá)

Fig.9
(Bộ giải mã)

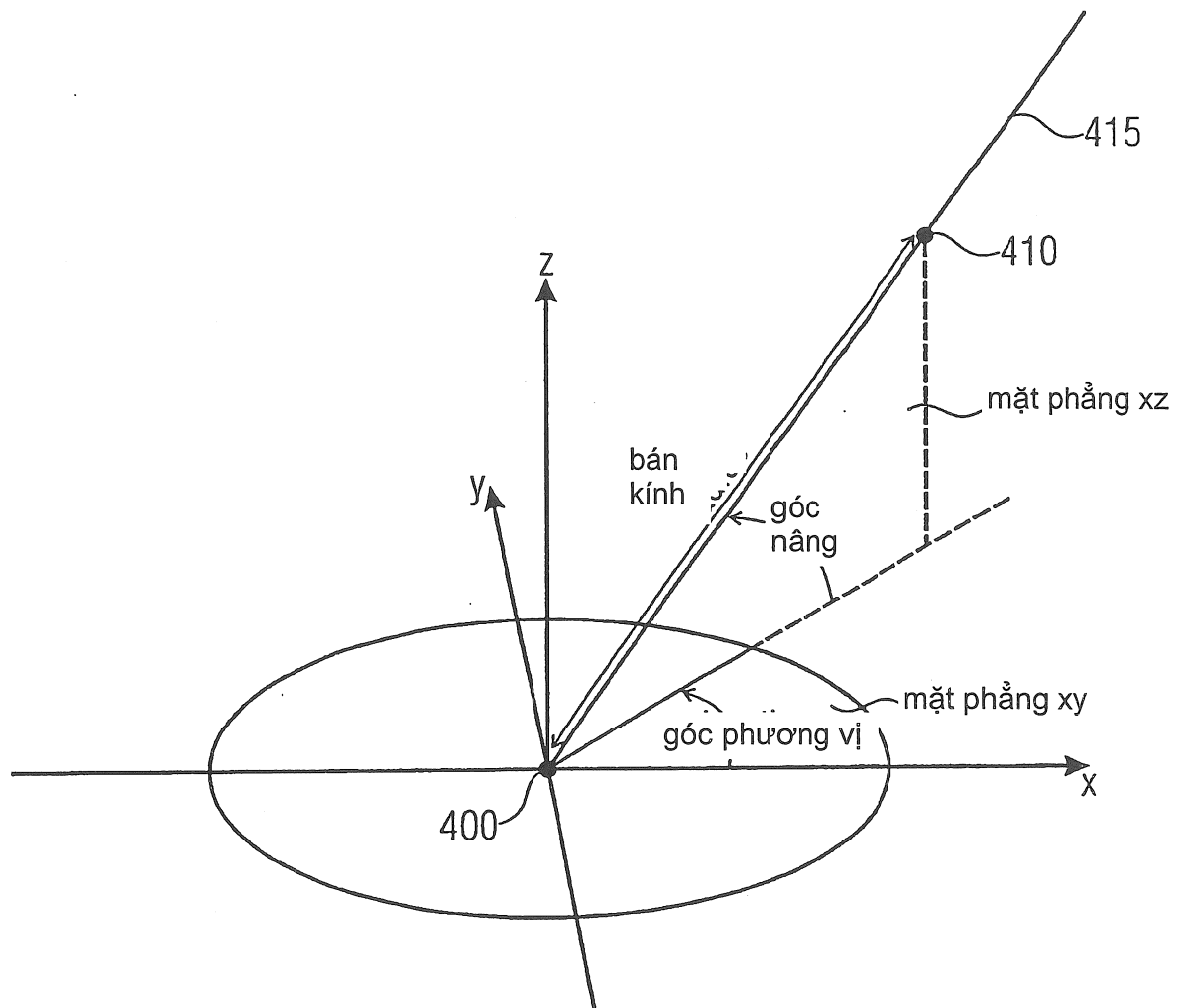


Fig.10

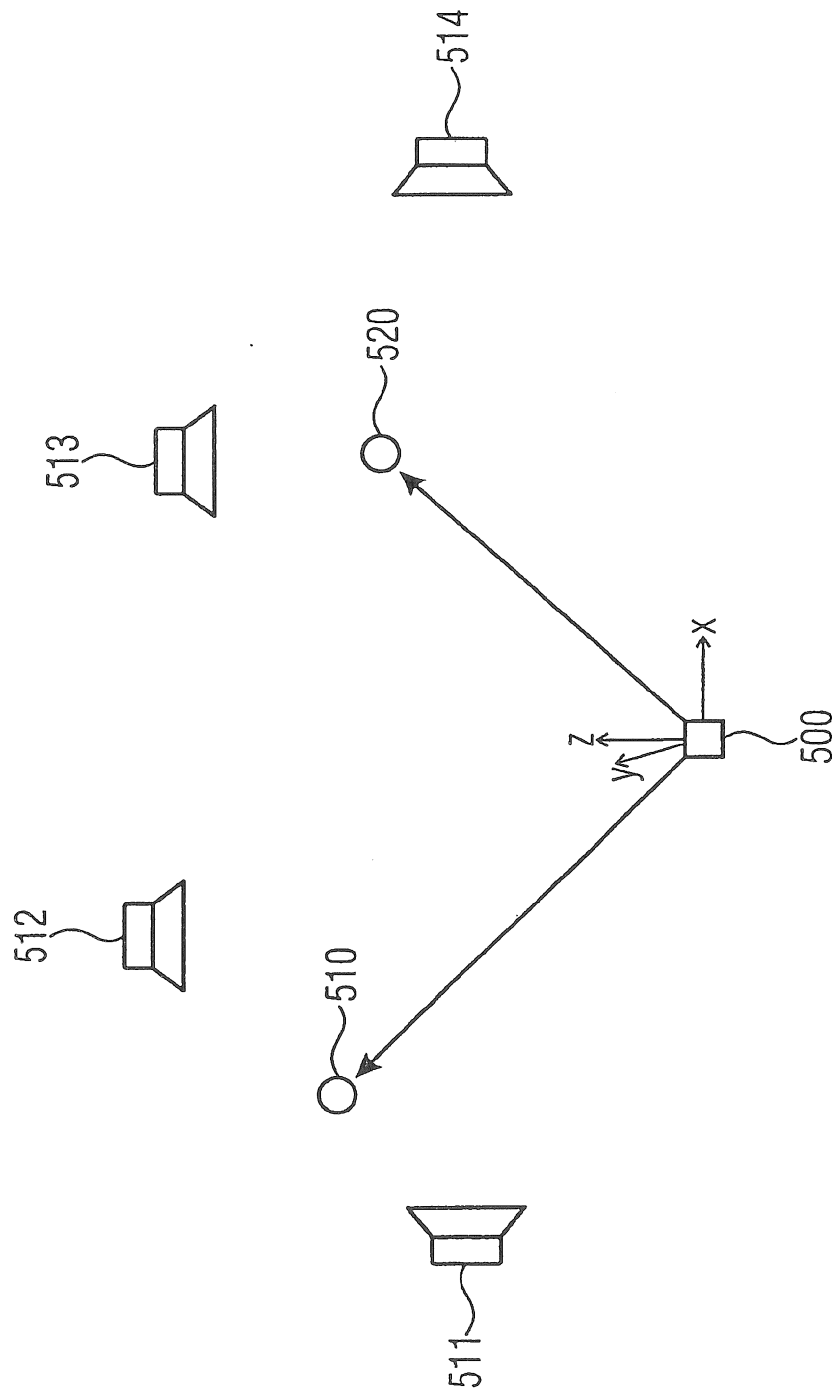


Fig.11