



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

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

(11)



1-0023883

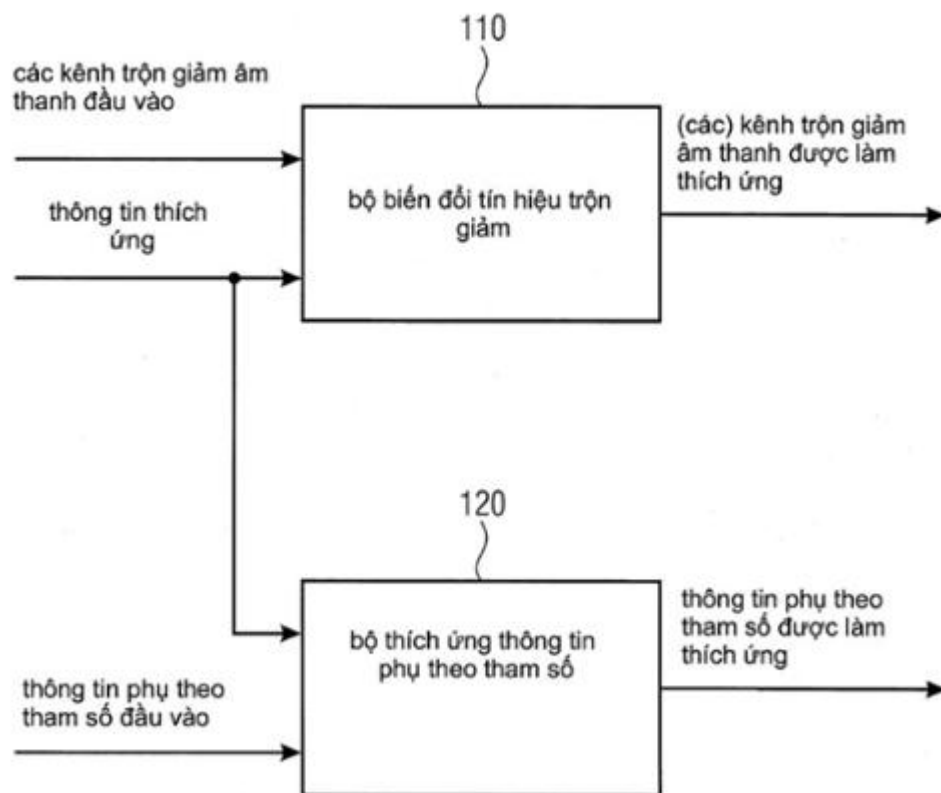
(51)<sup>7</sup> G10L 19/16

(13) B

- 
- (21) 1-2015-00739 (22) 28/06/2013  
(86) PCT/EP2013/063703 28/06/2013 (87) WO2014/023477 13/02/2014  
(30) 61/681,732 10/08/2012 US  
(45) 25/06/2020 387 (43) 25/05/2015 326A  
(73) Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. (DE)  
Hansastrasse 27c, 80686 Muenchen (DE)  
(72) KASTNER, Thorsten (DE); HERRE, Juergen (DE); TERENTIV, Leon (DE);  
HELLMUTH, Oliver (DE); PAULUS, Jouni (FI); RIDDERBUSCH, Falko (DE)  
(74) Công ty Luật TNHH AMBYS Hà Nội (AMBYS HANOI)
- 

(54) THIẾT BỊ VÀ PHƯƠNG PHÁP LÀM THÍCH ỨNG THÔNG TIN ÂM THANH ĐẦU VÀO, MÃ HÓA MỘT HOẶC NHIỀU ĐỐI TƯỢNG ÂM THANH ĐỂ THU ĐƯỢC THÔNG TIN ÂM THANH ĐƯỢC LÀM THÍCH ỨNG VÀ THIẾT BỊ TẠO RA MỘT HOẶC NHIỀU KÊNH ÂM THANH

(57) Sáng chế đề cập đến thiết bị và phương pháp làm thích ứng thông tin âm thanh đầu vào, mã hóa một hoặc nhiều đối tượng âm thanh để thu được thông tin âm thanh được làm thích ứng và thiết bị tạo ra một hoặc nhiều kênh âm thanh. Thông tin âm thanh đầu vào gồm có hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào và còn gồm có thông tin phụ theo tham số đầu vào. Thông tin âm thanh thích ứng gồm có hai hoặc nhiều hơn hai kênh trộn giảm âm thanh thích ứng và còn gồm có thông tin phụ theo tham số được làm thích ứng. Thiết bị gồm có bộ biến đổi tín hiệu trộn giảm (110) để làm thích ứng, phụ thuộc vào thông tin thích ứng, hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào để thu được một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng. Hơn nữa, thiết bị gồm có bộ thích ứng thông tin phụ theo tham số (120) để làm thích ứng, phụ thuộc vào thông tin thích ứng, thông tin phụ theo tham số đầu vào để thu được thông tin phụ theo tham số được làm thích ứng.



### **Lĩnh vực kỹ thuật được đề cập**

Sáng chế đề cập đến việc giải mã tín hiệu âm thanh và xử lý tín hiệu âm thanh, và cụ thể, đề cập đến bộ giải mã và các phương pháp làm thích ứng thông tin âm thanh trong mã hóa đối tượng âm thanh không gian (spatial-audio-object-coding (SAOC)).

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

Trong các hệ thống âm thanh số hiện đại, xu hướng chính cho phép các sửa đổi liên quan đến đối tượng âm thanh của nội dung được truyền trên phía máy nhận. Những sửa đổi này bao gồm các sửa đổi hệ số khuếch đại của các phần đã được chọn của tín hiệu âm thanh và/hoặc tái định vị trong không gian các đối tượng âm thanh đã được thiết kế trong trường hợp phát lại đa kênh qua các loa được phân bố trong không gian. Điều này có thể đạt được bằng cách phân phối riêng biệt các phần khác nhau của nội dung âm thanh đến các loa khác nhau.

Nói cách khác, trong kỹ thuật xử lý âm thanh, truyền âm thanh, và lưu trữ âm thanh, có sự tăng lên nhu cầu để cho phép một sự tương tác người dùng đối với sự phát lại nội dung âm thanh định hướng đối tượng cũng như yêu cầu sử dụng các khả năng mở rộng của sự phát lại đa kênh để biểu diễn riêng biệt các nội dung âm thanh hoặc các phần của các nội dung âm thanh để nâng cao được ấn tượng nghe. Bằng cách này, việc sử dụng nội dung âm thanh đa kênh mang đến những sự cải thiện đáng kể cho người dùng. Ví dụ, ấn tượng nghe ba chiều có thể thu được, nó đi cùng với tính cải thiện sự hài lòng người sử dụng về các ứng dụng giải trí. Tuy nhiên, nội dung âm thanh đa kênh cũng là hữu dụng trong các môi trường chuyên nghiệp, ví dụ, trong các ứng dụng hội thảo bằng điện thoại, bởi vì khả năng nghe rõ người nói có thể được cải thiện bằng cách sử dụng sự phát lại âm thanh đa kênh. Ứng dụng khác có thể là cho người nghe đoạn nhạc để điều chỉnh riêng biệt mức phát lại và/hoặc vị trí không gian của các phần khác nhau (còn được gọi là “đối tượng âm thanh”) hoặc các phần của

một đĩa hát, như phần xướng âm hoặc các nhạc cụ khác nhau. Người dùng có thể thực hiện việc điều chỉnh với các lý do sở thích cá nhân, để chuyển dịch dễ dàng hơn một hoặc nhiều phần từ một đoạn nhạc, cho các mục đích giáo dục, karaoke, buổi diễn tập, v.v.

Việc truyền rời rạc thẳng tất cả nội dung âm thanh đa đối tượng hoặc đa kênh dạng số, ví dụ, dưới dạng dữ liệu điều biến mã xung (pulse code modulation – PCM) hoặc các định dạng âm thanh được nén, yêu cầu dung lượng bit rất cao. Tuy nhiên, cũng có mong muốn truyền và lưu trữ dữ liệu âm thanh theo cách dung lượng bit hiệu quả. Do đó, người ta sẵn sàng chấp nhận một giao dịch hợp lý giữa chất lượng âm thanh và các yêu cầu dung lượng bit để tránh được việc tải nguồn quá mức gây ra bởi các ứng dụng đa đối tượng/đa kênh.

Gần đây, trong lĩnh vực mã hóa âm thanh, các kỹ thuật tham số cho sự truyền/lưu trữ hiệu quả dung lượng của các tín hiệu âm thanh đa đối tượng/đa kênh được đưa vào bởi, ví dụ, Nhóm các chuyên gia Fig.ảnh động (Moving Picture Experts Group (MPEG)) và khác nữa. Một ví dụ là MPEG Surround (MPS) là phương pháp định hướng kênh [MPS, BCC], hoặc mã hóa đối tượng âm thanh không gian MPEG (MPEG Spatial Audio Object Coding (SAOC)) như là phương pháp định hướng đối tượng [JSC, SAOC, SAOC1, SAOC2]. Phương pháp định hướng đối tượng khác được gọi là “phân chia nguồn được thông báo” [ISS1, ISS2, ISS3, ISS4, ISS5, ISS6]. Các kỹ thuật này nhằm tái tạo bối cảnh âm thanh đầu ra mong muốn hoặc đối tượng nguồn âm thanh mong muốn trên cơ sở của việc trộn các kênh/các đối tượng và thông tin phụ bổ sung mô tả bối cảnh âm thanh được lưu trữ/được truyền và/hoặc các đối tượng nguồn âm thanh trong đoạn âm thanh.

Việc ước tính và ứng dụng thông tin phụ liên quan đến kênh/đối tượng trong các hệ thống này được thực hiện theo cách thức có chọn lọc tần số-thời gian. Do đó, các hệ thống này sử dụng các phép biến đổi tần số-thời gian như phép biến đổi Fourier rời rạc (Discrete Fourier Transform (DFT)), phép biến đổi Fourier thời gian ngắn (Short Time Fourier Transform (STFT)) hoặc các giàn lọc như các giàn bộ lọc gương cầu phương (Quadrature Mirror Filter (QMF)), v.v. Nguyên lý cơ bản của các hệ thống này được mô tả trong Fig.3, ví dụ, sử dụng MPEG SAOC.

Trong trường hợp của STFT, kích thước theo thời gian được biểu diễn bởi số thời gian khối và kích thước phổ được thu giữ bởi số hệ số phổ (“ngăn”). Trong trường hợp của QMF, kích thước theo thời gian được biểu diễn bởi số khe thời gian và kích thước phổ được thu giữ bởi số dải phụ. Nếu độ phân giải phổ của QMF được cải thiện bởi ứng dụng tiếp theo của giai đoạn bộ lọc thứ hai, toàn bộ giàn lọc được gọi là QMF hỗn hợp và các dải phụ độ phân giải tốt được gọi là các dải phụ hỗn hợp.

Như đã chỉ ra ở trên, trong SAOC việc xử lý nhìn chung được tiến hành theo cách thức có chọn lọc tần số-thời gian và có thể được mô tả như dưới đây trong mỗi dải tần số, như được mô tả trên Fig.3:

- $N$  tín hiệu đối tượng âm thanh đầu vào  $s_1 \dots s_N$  được trộn giảm đến  $P$  kênh  $x_1 \dots x_P$  như là một phần việc xử lý của bộ mã hóa sử dụng ma trận trộn giảm gồm có các phần tử  $d_{1,1} \dots d_{N,P}$ . Ngoài ra, bộ mã hóa lấy thông tin phụ mô tả các đặc tính của các đối tượng âm thanh đầu vào (mô đun bộ ước lượng thông tin phụ (side-information-estimator - SEI)). Đối với MPEG SAOC, các mối tương quan hệ các năng lượng đối tượng với nhau là dạng cơ bản nhất của thông tin phụ này.

- Trộn giảm (các) tín hiệu và thông tin phụ được truyền/được lưu trữ. Để đạt được điều này, (các) tín hiệu âm thanh trộn giảm có thể được nén, ví dụ, sử dụng các bộ mã hóa âm thanh cảm quan đã biết như MPEG-1/2 lớp II hoặc III (aka .mp3), MPEG-2/4 Mã hóa âm thanh tiên tiến (Advanced Audio Coding - AAC) v.v.

- Ở đầu nhận, bộ giải mã về lý thuyết cố gắng phục hồi lại các tín hiệu đối tượng ban đầu (“sự chia tách đối tượng”) từ các tín hiệu trộn giảm (đã được giải mã) sử dụng thông tin phụ đã được truyền đi. Các tín hiệu đối tượng được lấy gần đúng này  $\hat{s}_1 \dots \hat{s}_N$  tiếp đó được trộn thành bối cảnh đích được biểu diễn bởi các kênh đầu ra âm thanh  $M$   $\hat{y}_1 \dots \hat{y}_M$  sử dụng ma trận biểu diễn được mô tả bởi các hệ số  $r_{1,1} \dots r_{N,M}$  trong Fig.3. Bối cảnh đích mong muốn, trong trường hợp đặc biệt, có thể là sự biểu diễn duy nhất một tín hiệu nguồn ra khỏi hỗn hợp (kịch bản chia tách nguồn), ngoài ra còn bối cảnh âm thanh tùy hứng bất kỳ khác gồm có các đối tượng được truyền. Ví dụ, đầu ra có thể là bối cảnh đích một kênh đơn lẻ, âm lập thể kênh 2 hoặc đa kênh 5.1.

Fig.6 mô tả dưới dạng giản đồ nguyên lý của sơ đồ mã hóa/giải mã âm thanh. Cụ thể, Fig.6 là sự mô tả nguyên tắc của chuỗi mã hóa/giải mã âm thanh.

Ở phía mã hóa, tín hiệu âm thanh được nén bởi sơ đồ mã hóa âm thanh (thường khai thác các hiệu ứng giác quan) và thông tin phụ theo tham số (Parametric Side Information-PSI) được tính toán (xem bộ mã hóa 601). Dòng bit thu được gồm có tín hiệu âm thanh đã được mã hóa và PSI được lưu trữ (hoặc được truyền) đến phía bộ giải mã, ở đó chúng có thể được giải mã bởi các thực thể bộ giải mã khác nhau 620, 621, 622, được đánh dấu là “A”, “B”, v.v trong Fig.6. Các thực thể bộ giải mã này có thể khác nhau (ví dụ, các mức phức tạp khác nhau trong thông số kỹ thuật tiêu chuẩn, ứng dụng hoặc các giới hạn thực hiện, v.v) [SAOC, SAOC1, SAOC2].

Sơ đồ mã hóa của tình trạng kỹ thuật không có khả năng thích ứng PSI với kịch bản hoặc nền ứng dụng đích riêng theo cách hiệu quả. Điều này có thể dẫn đến sự phức tạp tính toán (hơn cần thiết) cao hơn ở phía bộ giải mã hoặc có thể dẫn đến các vấn đề tương thích.

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

Mục đích của sáng chế là đề xuất các khía cạnh cải thiện đối với việc mã hóa đối tượng âm thanh. Mục đích của sáng chế được giải quyết bởi bộ giải mã theo điểm 1, bởi phương pháp mã hóa theo điểm 14 và bởi chương trình máy tính theo điểm 15.

Thiết bị làm thích ứng thông tin âm thanh đầu vào, mã hóa một hoặc nhiều đối tượng âm thanh, để thu được thông tin âm thanh thích ứng được đề xuất. Thông tin âm thanh đầu vào gồm có hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào và còn gồm có thông tin phụ theo tham số đầu vào. Thông tin âm thanh được làm thích ứng gồm có một hoặc nhiều kênh trộn giảm âm thanh thích ứng và còn gồm có thông tin phụ theo tham số được làm thích ứng.

Thiết bị gồm có bộ biến đổi tín hiệu trộn giảm để làm thích ứng, phụ thuộc vào thông tin thích ứng, hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào để thu được một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng.

Hơn nữa, thiết bị gồm có bộ thích ứng thông tin phụ theo tham số để làm thích ứng, phụ thuộc vào thông tin thích ứng, thông tin phụ theo tham số đầu vào để thu được thông tin phụ theo tham số thích ứng.

Theo một phương án, bộ biến đổi tín hiệu trộn giảm có thể được tạo cấu hình để làm thích ứng hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào phụ thuộc vào thông tin thích ứng, do đó số lượng của một hoặc nhiều kênh trộn giảm âm thanh thích ứng là nhỏ hơn số lượng của hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào.

Theo một phương án, thông tin thích ứng có thể phụ thuộc vào thực thể bộ giải mã. Bộ biến đổi tín hiệu trộn giảm có thể được tạo cấu hình để làm thích ứng hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào phụ thuộc vào thực thể bộ giải mã. Ở đây, các thuật ngữ “bộ giải mã”, và “thực thể bộ giải mã” có cùng một nghĩa.

Theo một phương án, thực thể bộ giải mã có thể có khả năng giải mã tối đa một lượng lớn nhất các kênh trộn giảm. Thông tin thích ứng có thể phụ thuộc vào số lượng lớn nhất các kênh trộn giảm đã nêu. Hơn nữa, bộ biến đổi tín hiệu trộn giảm có thể được tạo cấu hình để làm thích ứng với hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào phụ thuộc vào thông tin thích ứng để thu được một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng, do đó số lượng một hoặc nhiều kênh trộn giảm được làm thích ứng bằng với số lượng lớn nhất của các kênh trộn đã nêu.

Theo một phương án, thông tin thích ứng có thể gồm có ma trận thích ứng ( $\mathbf{D}_{dmx}^{DSM}$ ).

Theo một phương án, bộ biến đổi tín hiệu trộn giảm có thể được tạo cấu hình để làm thích ứng, phụ thuộc vào ma trận thích ứng ( $\mathbf{D}_{dmx}^{DSM}$ ), hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào ( $\mathbf{X}_{dmx}^{ENC}$ ) để thu được một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng ( $\mathbf{X}_{dmx}^{DSM}$ ),

Theo một phương án, bộ biến đổi tín hiệu trộn giảm có thể được tạo cấu hình để làm thích ứng, phụ thuộc vào ma trận thích ứng  $\mathbf{D}_{dmx}^{DSM}$ , hai hoặc nhiều hơn hai kênh

trộn giảm âm thanh đầu vào  $\mathbf{X}_{dmx}^{ENC}$  để thu được một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng  $\mathbf{X}_{dmx}^{DSM}$  bằng cách áp dụng công thức:

$$\mathbf{X}_{dmx}^{DSM} = \mathbf{D}_{dmx}^{DSM} \mathbf{X}_{dmx}^{ENC}.$$

Theo một phương án, bộ thích ứng thông tin phụ theo tham số có thể được tạo cấu hình để làm thích ứng, phụ thuộc vào ma trận thích ứng ( $\mathbf{D}_{dmx}^{DSM}$ ), thông tin phụ theo tham số đầu vào ( $\mathbf{D}_{dmx}^{ENC}$ ) để thu được thông tin phụ theo tham số được làm thích ứng ( $\mathbf{D}_{dmx}^{PSI}$ ).

Theo một phương án, bộ thích ứng thông tin phụ theo tham số có thể được tạo cấu hình để làm thích ứng, phụ thuộc vào ma trận thích ứng  $\mathbf{D}_{dmx}^{DSM}$ , thông tin phụ theo tham số đầu vào  $\mathbf{D}_{dmx}^{ENC}$  để thu được thông tin phụ theo tham số được làm thích ứng  $\mathbf{D}_{dmx}^{PSI}$  bằng cách áp dụng công thức:

$$\mathbf{D}_{dmx}^{PSI} = \mathbf{D}_{dmx}^{DSM} \mathbf{D}_{dmx}^{ENC}.$$

Theo một phương án, thông tin phụ theo tham số đầu vào ( $\mathbf{D}_{dmx}^{enc}$ ) có thể biểu thị ma trận trộn giảm ban đầu, do đó bằng cách áp dụng ma trận trộn giảm ban đầu ( $\mathbf{D}_{dmx}^{enc}$ ) trên một hoặc nhiều đối tượng âm thanh ( $\mathbf{S}$ ), hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào ( $\mathbf{X}_{dmx}^{enc}$ ) thu được. Bộ thích ứng thông tin phụ theo tham số có thể được tạo cấu hình để xác định ma trận trộn giảm thích ứng ( $\mathbf{D}_{dmx}^{PSI}$ ) là thông tin phụ theo tham số được làm thích ứng, do đó bằng cách áp dụng ma trận trộn giảm được làm thích ứng ( $\mathbf{D}_{dmx}^{PSI}$ ) trên một hoặc nhiều đối tượng âm thanh ( $\mathbf{S}$ ), một hoặc nhiều kênh trộn giảm âm thanh thích ứng ( $\mathbf{X}_{dmx}^{DSM}$ ) thu được.

Hơn nữa, theo một phương án, thiết bị tạo ra một hoặc nhiều kênh âm thanh từ thông tin âm thanh đầu vào mã hóa một hoặc nhiều đối tượng âm thanh được đề xuất.

Thiết bị tạo ra một hoặc nhiều kênh âm thanh gồm có thiết bị theo một trong số các phương án được mô tả ở trên để làm thích ứng thông tin âm thanh đầu vào để thu

được thông tin âm thanh được làm thích ứng, trong đó thông tin âm thanh đầu vào gồm có hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào và còn gồm có thông tin phụ theo tham số đầu vào, trong đó thông tin âm thanh được làm thích ứng gồm có một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng và còn gồm có thông tin phụ theo tham số được làm thích ứng.

Hơn nữa, thiết bị tạo ra một hoặc nhiều kênh âm thanh gồm có thực thể bộ giải mã, để giải mã, phụ thuộc vào thông tin phụ theo tham số được làm thích ứng, một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng để thu được một hoặc nhiều kênh âm thanh.

Theo một phương án, bộ thích ứng thông tin phụ theo tham số của thiết bị làm thích ứng thông tin âm thanh đầu vào có thể được tạo cấu hình để nhận dòng bit đầu vào gồm có thông tin phụ theo tham số đầu vào. Bộ thích ứng thông tin phụ theo tham số của thiết bị làm thích ứng thông tin âm thanh đầu vào có thể được tạo cấu hình để làm thích ứng thông tin phụ theo tham số đầu vào để thu được thông tin phụ theo tham số được làm thích ứng, và để đưa thông tin phụ theo tham số được làm thích ứng vào trong thực thể bộ giải mã. Thực thể bộ giải mã có thể được tạo cấu hình để giải mã một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng phụ thuộc vào thông tin phụ theo tham số được làm thích ứng.

Theo một phương án khác, bộ thích ứng thông tin phụ theo tham số của thiết bị làm thích ứng thông tin âm thanh đầu vào có thể được tạo cấu hình để nhận dòng bit đầu vào gồm có thông tin phụ theo tham số đầu vào. Bộ thích ứng thông tin phụ theo tham số của thiết bị làm thích ứng thông tin âm thanh đầu vào có thể được tạo cấu hình để thay thế thông tin phụ theo tham số đầu vào trong dòng bit đầu vào bởi thông tin phụ theo tham số được làm thích ứng để thu được dòng bit biến đổi. Bộ thích ứng thông tin phụ theo tham số của thiết bị làm thích ứng thông tin âm thanh đầu vào có thể được tạo cấu hình để đưa dòng bit biến đổi vào trong thực thể bộ giải mã. Hơn nữa, thực thể bộ giải mã có thể được tạo cấu hình để giải mã một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng phụ thuộc vào dòng bit biến đổi.

Hơn nữa, phương pháp làm thích ứng thông tin âm thanh đầu vào, mã hóa một hoặc nhiều đối tượng âm thanh, để thu được thông tin âm thanh được làm thích ứng. Thông tin âm thanh đầu vào gồm có hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào và còn gồm có thông tin phụ theo tham số đầu vào. Thông tin âm thanh được làm thích ứng gồm có một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng và còn gồm có thông tin phụ theo tham số được làm thích ứng. Phương pháp gồm có các bước:

- Làm thích ứng, phụ thuộc vào thông tin thích ứng, hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào để thu được một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng, và:

- Làm thích ứng, phụ thuộc vào thông tin thích ứng, thông tin phụ theo tham số đầu vào để thu được thông tin phụ theo tham số được làm thích ứng,

Hơn nữa, chương trình máy tính để thực hiện phương pháp được mô tả ở trên khi được thực hiện bởi máy tính hoặc bộ xử lý tín hiệu được đề xuất.

Các phương án ưu tiên sẽ được cung cấp trong các điểm yêu cầu bảo hộ phụ thuộc.

### **Mô tả vắn tắt các hình vẽ**

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

Fig.1 minh họa thiết bị làm thích ứng thông tin âm thanh đầu vào, mã hóa một hoặc nhiều đối tượng âm thanh, để thu được thông tin âm thanh được làm thích ứng theo một phương án.

Fig.2 minh họa thiết bị làm thích ứng thông tin âm thanh đầu vào, mã hóa một hoặc nhiều đối tượng âm thanh, để thu được thông tin âm thanh được làm thích ứng theo phương án khác.

Fig.3 thể hiện sơ đồ khối dưới dạng giản đồ tổng quan khái niệm của hệ thống SAOC.

Fig.4 thể hiện sơ đồ minh họa và dưới dạng giản đồ phép biểu diễn quang phổ theo thời gian của tín hiệu âm thanh đơn kênh,

Fig.5 thể hiện sơ đồ khối dưới dạng giản đồ của phép tính toán có chọn lọc tần số-thời gian của thông tin phụ trong bộ mã hóa SAOC.

Fig.6 mô tả giản lược nguyên lý của sơ đồ mã hóa/giải mã âm thanh.

Fig.7 minh họa thiết bị tạo ra một hoặc nhiều kênh âm thanh từ thông tin âm thanh đầu vào mã hóa một hoặc nhiều đối tượng âm thanh theo một phương án.

Fig.8 minh họa ứng dụng PSIA chung trong sơ đồ mã hóa/giải mã theo một phương án, và

Fig.9 minh họa ứng dụng PSIA rời trong sơ đồ mã hóa/giải mã theo một phương án.

### **Mô tả chi tiết sáng chế**

Trước khi mô tả các phương án của sáng chế, kiến thức về tình trạng kỹ thuật của các hệ thống SAOC được cung cấp.

Fig.3 thể hiện việc lắp ráp chung bộ mã hóa SAOC 10 và bộ giải mã SAOC 12. Bộ mã hóa SAOC 10 nhận  $N$  đối tượng đầu vào, và các tín hiệu âm thanh  $s_I$  đến  $s_N$ . Cụ thể, bộ mã hóa 10 gồm có bộ trộn giảm 16 mà nhận các tín hiệu âm thanh  $s_I$  đến  $s_N$  và trộn các tín hiệu âm thanh này thành tín hiệu trộn giảm 18. Hoặc, việc trộn giảm có thể được cung cấp theo bên ngoài ("sự trộn giảm nghệ thuật") và hệ thống ước lượng thông tin phụ bổ sung để làm cho sự trộn giảm được đề xuất khớp với sự trộn giảm được tính toán. Trong Fig.3, tín hiệu trộn giảm được thể hiện là tín hiệu kênh  $P$ . Do đó, cấu hình tín hiệu trộn giảm đơn âm ( $P=1$ ), âm lập thể ( $P=2$ ) hoặc đa kênh ( $P>2$ ) bất kỳ là có thể hiểu được.

Trong trường hợp trộn giảm âm lập thể, các kênh của tín hiệu trộn giảm 18 được chỉ ra là  $L0$  và  $R0$ , trong trường hợp trộn giảm đơn âm, các kênh của tín hiệu trộn giảm được chỉ ra đơn giản là  $L0$ . Để cho phép bộ giải mã SAOC 12 khôi phục các đối tượng riêng rẽ  $s_I$  đến  $s_N$ , bộ ước lượng thông tin phụ 17 cung cấp bộ giải mã SAOC 12 với thông tin phụ bao gồm các tham số SAOC. Ví dụ, trong trường hợp trộn giảm âm

lập thể, các tham số SAOC gồm có các chênh lệch mức đối tượng (object level differences-OLD), các tương quan giữa đối tượng (inter-object correlation-IOC) (các tham số tương quan chéo giữa đối tượng), các giá trị hệ số khuếch đại trộn giảm (downmix gain values-DMG) và các chênh lệch mức kênh trộn giảm (downmix channel level differences-DCLD). Thông tin phụ 20 bao gồm các tham số SAOC cùng với tín hiệu trộn giảm 18, tạo thành dòng dữ liệu đầu ra SAOC được nhận bởi bộ giải mã SAOC 12.

Bộ giải mã SAOC 12 gồm có bộ trộn tăng mà nhận tín hiệu trộn giảm 18 cũng như thông tin phụ 20 để khôi phục và biểu diễn các tín hiệu âm thanh  $s_I$  và  $s_N$  trên lựa chọn người dùng bất kỳ  $y_I$  đến  $y_M$ , với việc biểu diễn được mô tả sơ bộ bởi thông tin biểu diễn đầu vào 26 vào trong bộ giải mã SAOC 12.

Các tín hiệu âm thanh  $s_I$  đến  $s_N$  có thể được đưa vào trong bộ mã hóa 10 trong miền mã hóa bất kỳ, như, trong miền thời gian hoặc quang phổ. Trong trường hợp các tín hiệu âm thanh  $s_I$  đến  $s_N$  được đưa vào trong bộ mã hóa 10 trong miền thời gian, như PCM được mã hóa, bộ mã hóa 10 có thể sử dụng giàn lọc, như giàn QMF hỗn hợp, để truyền các tín hiệu vào trong miền quang phổ, trong đó các tín hiệu âm thanh được biểu diễn trong nhiều dải phụ được liên kết với các phần quang phổ khác nhau, với độ phân giải giàn lọc cụ thể. Nếu các tín hiệu âm thanh  $s_I$  đến  $s_N$  đã có trong phép biểu diễn được mong đợi bởi bộ mã hóa 10, các tín hiệu âm thanh không phải thực hiện sự phân giải quang phổ.

Fig.4 thể hiện tín hiệu âm thanh trong miền quang phổ vừa được nêu. Như có thể được thấy, tín hiệu âm thanh được thể hiện như là nhiều tín hiệu dải phụ. Mỗi tín hiệu dải phụ 30<sub>I</sub> đến 30<sub>K</sub> bao gồm một trình tự thời gian của các trị số dải phụ được chỉ ra bởi các ô nhỏ 32. Như có thể được thấy, các trị số dải phụ 32 của các tín hiệu dải phụ 30<sub>I</sub> đến 30<sub>K</sub> được đồng bộ hóa với nhau theo thời gian do đó, đối với mỗi khe thời gian của giàn lọc liên tiếp 34, mỗi dải phụ 30<sub>I</sub> đến 30<sub>K</sub> gồm có chính xác một trị số dải phụ 32. Như được minh họa bởi trục tần số 36, các tín hiệu dải phụ 30<sub>I</sub> đến 30<sub>K</sub> được liên kết với các vùng tần số khác nhau, và như được minh họa bởi trục thời gian 38, các khe thời gian của giàn lọc 34 được sắp đặt liên tiếp theo thời gian.

Như được chỉ ra ở trên, bộ trích xuất thông tin phụ 17 thuộc Fig.3 tính toán các tham số SAOC từ các tín hiệu âm thanh đầu vào  $s_1$  đến  $s_N$ . Theo tiêu chuẩn SAOC được thực hiện như hiện nay, bộ mã hóa 10 thực hiện việc tính toán này trong độ phân giải thời gian/tần số mà độ phân giải này có thể bị giảm đi tương đối so với độ phân giải thời gian/tần số ban đầu như được xác định bởi các khe thời gian của giàn lọc 34 và sự phân giải dải phụ, bởi một lượng nhất định, với lượng nhất định này được tín hiệu hóa đến phía bộ giải mã trong thông tin phụ 20. Các nhóm khe thời gian của giàn lọc liên tiếp 34 có thể tạo ra khung SAOC 41. Ngoài ra, số lượng dải tham số trong khung SAOC 41 được chuyển bên trong thông tin phụ 20. Do đó, miền thời gian/tần số được chia thành các ô thời gian/tần số được minh họa trong Fig.4 bởi các đường nét đứt 42. Trong Fig.4 các dải tham số được phân bố theo cùng cách thức trong các khung SAOC 41 được mô tả khác nhau do đó sự sắp xếp đều đặn các ô thời gian/tần số được thu. Tuy nhiên, nói chung, các dải tham số có thể thay đổi từ một khung SAOC 41 đến khung tiếp theo, phụ thuộc vào các nhu cầu khác nhau cho sự phân giải quang phổ trong các khung SAOC 41 tương ứng. Hơn nữa, chiều dài của các khung SAOC 41 cũng có thể thay đổi. Kết quả là, sự sắp đặt các ô thời gian/tần số có thể là không đều. Tuy nhiên, các ô thời gian/tần số trong khung SAOC 41 cụ thể thường có cùng thời gian và được sắp cùng hướng thời gian, cụ thể tất cả các ô t/f trong khung SAOC 41 đã nêu bắt đầu tại điểm bắt đầu khung SAOC 41 đã cho và kết thúc tại điểm kết thúc khung SAOC 41 đã nêu.

Bộ trích xuất thông tin phụ 17 được mô tả trong Fig.3 tính toán các tham số SAOC theo công thức dưới đây. Cụ thể, bộ trích xuất thông tin phụ 17 tính toán các chênh lệch mức đối tượng cho từng đối tượng  $i$  là

$$OLD_i^{l,m} = \frac{\sum_{n \in l} \sum_{k \in m} x_i^{n,k} x_i^{n,k*}}{\max_j \left( \sum_{n \in l} \sum_{k \in m} x_j^{n,k} x_j^{n,k*} \right)}$$

trong đó các tổng và chỉ số  $n$  và  $k$ , lần lượt, xuyên suốt tất cả các chỉ số theo thời gian 34, và tất cả các chỉ số quang phổ 30 thuộc về ô thời gian/tần số 42 nhất định, được tham chiếu bởi các chỉ số  $l$  cho khung SAOC (hoặc khe thời gian xử lý) và  $m$  cho dải tham số. Do đó, các năng lượng của tất cả các trị số dải phụ  $x_i$  của tín hiệu hoặc đối

tượng âm thanh  $i$  được tổng lại và chuẩn hóa thành giá trị số năng lượng cao nhất của ô đó trong số tất cả các đối tượng hoặc các tín hiệu âm thanh.  $x_i^{n,k*}$  để chỉ liên hợp phức của  $x_i^{n,k}$ .

Hơn nữa, bộ trích xuất thông tin phụ SAOC 17 có thể tính toán phép đo giống nhau của các ô thời gian/tần số tương ứng của cặp đối tượng đầu vào khác nhau  $s_I$  đến  $s_N$ . Mặc dù bộ trích xuất thông tin phụ SAOC 17 có thể tính toán phép đo giống nhau giữa tất cả các cặp đối tượng đầu vào  $s_I$  đến  $s_N$ , bộ trích xuất thông tin phụ 17 cũng có thể nén việc tạo tín hiệu các phép đo giống nhau hoặc giới hạn sự tính toán các phép đo giống nhau cho các đối tượng âm thanh  $s_I$  đến  $s_N$  mà các đối tượng này tạo thành các kênh bên trái hoặc bên phải của kênh âm lập thể thông thường. Trong trường hợp bất kỳ, phép đo giống nhau được gọi là tham số tương quan chéo liên đối tượng  $IOC_{i,j}^{l,m}$ . Việc tính toán là như sau

$$IOC_{i,j}^{l,m} = IOC_{j,i}^{l,m} = \text{Re} \left\{ \frac{\sum_{n \in l} \sum_{k \in m} x_i^{n,k} x_j^{n,k*}}{\sqrt{\sum_{n \in l} \sum_{k \in m} x_i^{n,k} x_i^{n,k*} \sum_{n \in l} \sum_{k \in m} x_j^{n,k} x_j^{n,k*}}} \right\}$$

với các chỉ số  $n$  và  $k$  xuyên suốt tất cả các trị số dải phụ thuộc ô thời gian/tần số nhất định 42,  $i$  và  $j$  biểu thị cặp đối tượng âm thanh nhất định  $s_I$  đến  $s_N$ , và  $\text{Re}\{ \}$  biểu thị hoạt động loại bỏ phần ảo của đối số phức.

Bộ trộn giảm 16 thuộc Fig.3 trộn các đối tượng  $s_I$  đến  $s_N$  bằng cách sử dụng các hệ số khuếch đại được áp dụng cho từng đối tượng  $s_I$  đến  $s_N$ . Tức là, hệ số khuếch đại  $d_i$  được áp dụng cho đối tượng  $i$  và sau đó tất cả các đối tượng được gán trọng số  $s_I$  đến  $s_N$  được cộng lại để thu được tín hiệu trộn giảm đơn âm, nó được minh họa trong Fig.3 nếu  $P=1$ . Trong một ví dụ khác, trường hợp của tín hiệu trộn giảm hai kênh, được mô tả trong Fig.3 nếu  $P=2$ , hệ số khuếch đại  $d_{l,i}$  được áp dụng cho đối tượng  $i$  và sau đó tất cả các đối tượng được khuếch đại nhờ hệ số khuếch đại này được cộng lại để thu được kênh trộn giảm bên trái  $L0$ , và các hệ số khuếch đại  $d_{r,i}$  được áp dụng cho đối tượng  $i$  và sau đó các đối tượng được khuếch đại nhờ hệ số khuếch đại được cộng lại

để thu được kênh trộn giảm bên phải  $R0$ . Xử lý tương tự với bên trên được áp dụng trong trường hợp trộn giảm đa kênh ( $P>2$ ).

Quy tắc trộn giảm này được tạo tín hiệu tới phía bộ giải mã bởi phương tiện là các hệ số khuếch đại trộn giảm  $DMG$  và, trong trường hợp của tín hiệu trộn giảm âm lập thể, là các chênh lệch mức kênh trộn giảm (downmix channel level differences)  $DCLD_i$

Các khuếch đại trộn giảm được tính toán theo công thức:

$$DMG_i = 20 \log_{10}(d_i + \varepsilon) \quad , \text{ (trộn giảm đơn âm),}$$

$$DMG_i = 10 \log_{10}(d_{1,i}^2 + d_{2,i}^2 + \varepsilon) \quad , \text{ (trộn giảm âm lập thể),}$$

trong đó  $\varepsilon$  là số nhỏ như  $10^{-9}$ .

Đối với DCLDs công thức sau đây áp dụng:

$$DCLD_i = 20 \log_{10} \left( \frac{d_{1,i}}{d_{2,i} + \varepsilon} \right).$$

Ở chế độ thông thường, bộ trộn giảm 16 tạo ra tín hiệu trộn giảm theo công thức:

$$(L0) = (d_i) \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix}$$

đối với trộn giảm đơn âm, hoặc

$$\begin{pmatrix} L0 \\ R0 \end{pmatrix} = \begin{pmatrix} d_{1,i} \\ d_{2,i} \end{pmatrix} \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix}$$

đối với trộn giảm âm lập thể, một cách lần lượt.

Do đó, theo các công thức được đề cập ở trên, các tham số  $OLD$  và  $IOC$  là hàm của các tín hiệu âm thanh và các tham số  $DMG$  và  $DCLD$  là hàm của  $d$ . Đồng thời, lưu ý rằng  $d$  có thể thay đổi theo thời gian và tần số.

Do đó, ở chế độ thông thường, bộ trộn giảm 16 trộn tất cả các đối tượng  $s_I$  đến  $s_N$  mà không có sự ưu tiên nào, tức là, xử lý tất cả các đối tượng  $s_I$  đến  $s_N$  như nhau.

Ở phía bộ giải mã, bộ trộn tăng thực hiện sự đảo ngược quy trình trộn giảm và việc thực hiện "thông tin kết xuất" 26 được biểu diễn bởi ma trận  $\mathbf{R}$  (trong tài liệu đôi khi còn được gọi là  $\mathbf{A}$ ) trong một bước tính toán, cụ thể, trong trường hợp trộn giảm hai kênh

$$\begin{pmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_M \end{pmatrix} = \mathbf{RED}^* (\mathbf{DED}^*)^{-1} \begin{pmatrix} L0 \\ R0 \end{pmatrix},$$

trong đó ma trận  $\mathbf{E}$  là hàm của các tham số OLD và IOC, và ma trận  $\mathbf{D}$  chứa các hệ số trộn giảm là

$$\mathbf{D} = \begin{pmatrix} d_{1,1} & \cdots & d_{1,N} \\ \vdots & \ddots & \vdots \\ d_{P,1} & \cdots & d_{P,N} \end{pmatrix}.$$

Ma trận  $\mathbf{E}$  là ma trận hiệp phương sai ước lượng của các đối tượng âm thanh  $s_I$  đến  $s_N$ . Trong cách thực hiện SAOC hiện thời, việc tính toán ma trận hiệp phương sai ước lượng  $\mathbf{E}$  thường được thực hiện trong độ phân giải theo quang phổ/thời gian của các tham số SAOC, cụ thể, đối với mỗi  $(l,m)$ , do đó ma trận hiệp phương sai ước lượng có thể được viết là  $\mathbf{E}^{l,m}$ . Ma trận hiệp phương sai ước lượng  $\mathbf{E}^{l,m}$  có kích thước  $N \times N$  với các hệ số của nó được định nghĩa là

$$e_{i,j}^{l,m} = \sqrt{OLD_i^{l,m} OLD_j^{l,m}} IOC_{i,j}^{l,m}.$$

Do đó, ma trận  $\mathbf{E}^{l,m}$  với

$$\mathbf{E}^{l,m} = \begin{pmatrix} e_{1,1}^{l,m} & \cdots & e_{1,N}^{l,m} \\ \vdots & \ddots & \vdots \\ e_{N,1}^{l,m} & \cdots & e_{N,N}^{l,m} \end{pmatrix}$$

có các chênh lệch mức đối tượng dọc theo đường chéo của ma trận, tức là,  $e_{i,j}^{l,m} = OLD_i^{l,m}$  đối với  $i=j$ , vì  $OLD_i^{l,m} = OLD_j^{l,m}$  và  $IOC_{i,j}^{l,m} = 1$  với  $i=j$ . Phía ngoài đường chéo của ma trận, ma trận hiệp phương sai ước lượng  $\mathbf{E}$  có các hệ số ma trận lần lượt

biểu diễn số trung bình nhân của các chênh lệch mức đối tượng của các đối tượng  $i$  và  $j$ , được gán trọng số với phép đo tương quan chéo liên đối tượng  $IOC_{i,j}^{l,m}$ .

Hình 5 thể hiện một nguyên lý có thể sử dụng để thực hiện một ví dụ của bộ ước lượng thông tin phụ (Side-information estimator-SIE) như là một phần của bộ mã hóa SAOC 10. Bộ mã hóa SAOC 10 gồm có bộ trộn 16 và bộ ước lượng thông tin phụ (SEI) 17. SEI về lý thuyết gồm có hai môđun: Một môđun 45 để tính toán phép biểu diễn t/f dựa trên ngắn hạn (ví dụ, STFT hoặc QMF) của mỗi tín hiệu. Phép biểu diễn t/f ngắn hạn được tính toán được đưa vào trong môđun thứ hai 46, môđun ước lượng thông tin phụ lựa chọn lọc t/f (t/f-SEI). Môđun t/f-SEI 46 tính toán thông tin phụ cho mỗi ô t/f. Trong các cách thực hiện SAOC hiện tại, sự biến đổi thời gian/tần số được cố định và giống nhau cho tất cả các đối tượng âm thanh  $s_1$  đến  $s_N$ . Hơn nữa, các tham số SAOC được xác định trên các khung SAOC mà các khung này là giống nhau cho tất cả các đối tượng âm thanh và có cùng độ phân giải thời gian/tần số cho tất cả các đối tượng âm thanh  $s_1$  đến  $s_N$ , do đó bỏ qua các yêu cầu đối tượng cụ thể cho độ phân giải theo thời gian tốt trong một số trường hợp hoặc độ phân giải quang phổ tốt trong các trường hợp khác.

Dưới đây, các phương án của sáng chế được mô tả.

Fig.1 minh họa thiết bị làm thích ứng thông tin âm thanh đầu vào, mã hóa một hoặc nhiều đối tượng âm thanh, để thu được thông tin âm thanh được làm thích ứng theo một phương án.

Thông tin âm thanh đầu vào gồm có hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào và còn gồm có thông tin phụ theo tham số đầu vào. Thông tin âm thanh được làm thích ứng gồm có một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng và còn gồm có thông tin phụ theo tham số được làm thích ứng.

Thiết bị gồm có bộ biến đổi tín hiệu trộn giảm (downmix signal modifier - DSM) 110 để làm thích ứng, phụ thuộc vào thông tin thích ứng, hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào để thu được một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng.

Hơn nữa, thiết bị gồm có bộ thích ứng thông tin phụ theo tham số (parametric side information adapter -PSIA) 120 để làm thích ứng, phụ thuộc vào thông tin thích ứng, thông tin phụ theo tham số đầu vào để thu được thông tin phụ theo tham số được làm thích ứng.

Fig.2 minh họa thiết bị làm thích ứng thông tin âm thanh đầu vào, mã hóa một hoặc nhiều đối tượng âm thanh, để thu được thông tin âm thanh được làm thích ứng theo phương án khác.

Theo một phương án, thông tin thích ứng có thể phụ thuộc vào thực thể bộ giải mã, và bộ biến đổi tín hiệu trộn giảm 110 có thể được tạo cấu hình để làm thích ứng với hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào phụ thuộc vào thực thể bộ giải mã.

Ví dụ, bộ biến đổi tín hiệu trộn giảm 110 thuộc Fig.2 làm thích ứng sự trộn giảm với các khả năng của thực thể bộ giải mã cụ thể.

Theo một phương án, bộ biến đổi tín hiệu trộn giảm 110 có thể được tạo cấu hình để làm thích ứng hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào phụ thuộc vào thông tin thích ứng, sao cho số lượng một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng nhỏ hơn số lượng của hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào.

Ví dụ, trong phương án thuộc Fig.2, bộ biến đổi tín hiệu trộn giảm 110 làm giảm số lượng kênh trộn giảm/vận chuyển.

Ví dụ, 22,2 kênh trộn giảm âm thanh đầu vào (= 24 kênh trộn giảm âm thanh đầu vào) có thể được giảm xuống 7.1 kênh trộn giảm âm thanh được làm thích ứng (= 8 kênh trộn giảm âm thanh được làm thích ứng).

Hoặc, ví dụ, 5.1 kênh trộn giảm âm thanh đầu vào (= 6 kênh trộn giảm âm thanh đầu vào) được giảm xuống 2.0 kênh trộn giảm âm thanh được làm thích ứng (= 2 kênh trộn giảm âm thanh được làm thích ứng).

Hoặc, ví dụ, 2 kênh trộn giảm âm thanh đầu vào được giảm xuống 1 kênh trộn giảm âm thanh được làm thích ứng.

Những sự kết hợp khác nhau của các kênh trộn giảm âm thanh đầu vào và các kênh trộn giảm âm thanh được làm thích ứng là có thể sử dụng.

Theo một phương án, thực thể bộ giải mã có khả năng giải mã nhiều nhất một lượng tối đa các kênh trộn giảm. Thông tin thích ứng có thể phụ thuộc vào số lượng tối đa các kênh trộn giảm đã nêu. Hơn nữa, bộ biến đổi tín hiệu trộn giảm 110 có thể được tạo cấu hình để làm thích ứng hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào phụ thuộc vào thông tin thích ứng để thu được một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng, sao cho số lượng một hoặc nhiều kênh trộn giảm được làm thích ứng bằng với số lượng tối đa các kênh trộn giảm đã nêu.

Ví dụ, bộ biến đổi tín hiệu trộn giảm 110 thuộc Fig.2 biến đổi sự trộn giảm thành tín hiệu âm thanh tương ứng với cấu hình kênh đầu ra được hỗ trợ tối đa của thực thể bộ giải mã cụ thể.

Theo một phương án, thông tin thích ứng có thể, ví dụ, gồm có ma trận thích ứng ( $\mathbf{D}_{dmx}^{DSM}$ ).

Bộ thích ứng thông tin phụ theo tham số 120 có thể, ví dụ, làm thích ứng PSI tương ứng với sự trộn giảm được biến đổi để làm giảm độ phức tạp tính toán cho bộ giải mã, và để làm giảm kích thước/dung lượng bit của dòng bit dữ liệu tương ứng mà không tạo ra ảnh hưởng tiêu cực đến chất lượng âm thanh đầu ra của bộ giải mã.

Ví dụ, PSIA 120 biến đổi dòng bit PSI tương ứng thay thế thông tin biểu diễn ma trận trộn giảm ban đầu bằng thông tin được cập nhật mô tả sự trộn giảm thu được (tính đến các biến đổi DSM) để tương ứng với thông số kỹ thuật cụ thể của bộ giải mã.

Ví dụ, bộ mã hóa SAOC cung cấp tín hiệu trộn giảm âm lập thể  $\mathbf{X}_{dmx}^{ENC}$  thu được từ việc ứng dụng ma trận trộn giảm của bộ mã hóa  $\mathbf{D}_{dmx}^{ENC}$  cho các tín hiệu đối tượng âm thanh đầu vào  $\mathbf{S}$ :

$$\mathbf{X}_{dmx}^{ENC} = \mathbf{D}_{dmx}^{ENC} \mathbf{S}.$$

Theo một phương án, bộ biến đổi tín hiệu trộn giảm 110 có thể được tạo cấu hình để làm thích ứng, phụ thuộc vào ma trận thích ứng  $\mathbf{D}_{dmx}^{DSM}$ , hai hoặc nhiều hơn hai

kênh trộn giảm âm thanh đầu vào  $\mathbf{X}_{dmx}^{ENC}$  để thu được một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng  $\mathbf{X}_{dmx}^{DSM}$ . Theo phương án, kết quả được rút ra, ví dụ, bằng cách áp dụng công thức  $\mathbf{X}_{dmx}^{DSM} = \mathbf{D}_{dmx}^{DSM} \mathbf{X}_{dmx}^{ENC}$ .

Ví dụ, theo một phương án, được giả định rằng thực thể bộ giải mã SAOC cụ thể hỗ trợ chỉ sự trộn giảm đơn âm (ví dụ SAOC biên dạng độ trễ thấp / Mức 1). Trong trường hợp này, DSM 110 biến đổi sự trộn giảm âm lập thể  $\mathbf{X}_{dmx}^{ENC}$  thành tín hiệu đơn âm  $\mathbf{X}_{dmx}^{DSM}$  sử dụng ma trận trộn giảm được định trước  $\mathbf{D}_{dmx}^{DSM}$  như sau:

$$\mathbf{X}_{dmx}^{DSM} = \mathbf{D}_{dmx}^{DSM} \mathbf{X}_{dmx}^{ENC}.$$

Theo một phương án, bộ thích ứng thông tin phụ theo tham số 120 có thể được tạo cấu hình để làm thích ứng, phụ thuộc vào ma trận thích ứng  $\mathbf{D}_{dmx}^{DSM}$ , thông tin phụ theo tham số đầu vào  $\mathbf{D}_{dmx}^{ENC}$  để thu được thông tin phụ theo tham số được làm thích ứng  $\mathbf{D}_{dmx}^{PSI}$ . Theo một phương án, kết quả này, ví dụ, được rút ra bằng cách áp dụng công thức:  $\mathbf{D}_{dmx}^{PSI} = \mathbf{D}_{dmx}^{DSM} \mathbf{D}_{dmx}^{ENC}$ .

Ví dụ, theo một phương án, PSIA 120 phân tích dòng bit PSI tương ứng; trích xuất thông tin mô tả ma trận trộn giảm  $\mathbf{D}_{dmx}^{ENC}$ ; thay thế những dữ liệu này bằng thông tin được cập nhật mà mô tả ma trận trộn giảm mới  $\mathbf{D}_{dmx}^{PSI}$ :

$$\mathbf{D}_{dmx}^{PSI} = \mathbf{D}_{dmx}^{DSM} \mathbf{D}_{dmx}^{ENC}.$$

Do đó, theo một phương án, thông tin phụ theo tham số đầu vào ( $\mathbf{D}_{dmx}^{enc}$ ) có thể biểu thị ma trận trộn giảm ban đầu, sao cho bằng cách áp dụng ma trận trộn giảm ban đầu ( $\mathbf{D}_{dmx}^{enc}$ ) cho một hoặc nhiều đối tượng âm thanh (S), hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào ( $\mathbf{X}_{dmx}^{enc}$ ) được thu. Bộ thích ứng thông tin phụ theo tham số có thể được tạo cấu hình để xác định ma trận trộn giảm được làm thích ứng ( $\mathbf{D}_{dmx}^{PSI}$ ) như là thông tin phụ theo tham số được làm thích ứng, sao cho bằng cách áp dụng ma trận trộn giảm được làm thích ứng ( $\mathbf{D}_{dmx}^{PSI}$ ) cho một hoặc nhiều đối tượng âm thanh (S), một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng ( $\mathbf{X}_{dmx}^{DSM}$ ) được thu.

Theo một phương án, PSIA định dạng dòng bit biến đổi mới hoặc chuyển trực tiếp các tham số này cho bộ giải mã.

Quy trình mã hóa và giải mã này được thực hiện bởi PSIA cũng có thể bao gồm sự biến đổi các định dạng thể hiện ma trận trộn giảm khác nhau (ví dụ hệ tọa độ cực đến hệ tọa độ Đề-các-tơ, v.v).

Chức năng được mô tả này của PSIA có thể giải quyết các vấn đề tương thích tiềm năng và làm giảm kích thước của dòng bit tương ứng.

Hình 7 minh họa thiết bị 700 để tạo ra một hoặc nhiều kênh âm thanh từ thông tin âm thanh đầu vào mã hóa một hoặc nhiều đối tượng âm thanh theo một phương án.

Thiết bị 700 để tạo ra một hoặc nhiều kênh âm thanh gồm có thiết bị 710 theo một trong số các phương án đã được mô tả ở trên để thích ứng thông tin âm thanh đầu vào để thu được thông tin âm thanh thích ứng. Thông tin âm thanh đầu vào gồm có hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào và còn gồm có thông tin phụ theo tham số đầu vào. Thông tin âm thanh được làm thích ứng gồm có một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng và còn gồm có thông tin phụ theo tham số được làm thích ứng.

Thiết bị 710 theo một trong số các phương án được mô tả ở trên để làm thích ứng thông tin âm thanh đầu vào gồm có bộ biến đổi tín hiệu trộn giảm 110 và bộ thích ứng thông tin phụ theo tham số 120.

Hơn nữa, thiết bị 700 để tạo ra một hoặc nhiều kênh âm thanh gồm có thực thể bộ giải mã 720, để giải mã, phụ thuộc vào thông tin phụ theo tham số thích ứng, một hoặc nhiều kênh âm thanh làm thích ứng để thu được một hoặc nhiều kênh âm thanh.

Theo một phương án, bộ thích ứng thông tin phụ theo tham số 120 của thiết bị 710 làm thích ứng thông tin âm thanh đầu vào có thể được tạo cấu hình để nhận dòng bit đầu vào gồm có thông tin phụ theo tham số đầu vào. Bộ thích ứng thông tin phụ theo tham số 120 của thiết bị 710 làm thích ứng thông tin âm thanh đầu vào có thể được tạo cấu hình để làm thích ứng thông tin phụ theo tham số đầu vào để thu được thông tin phụ theo tham số được làm thích ứng, và để đưa thông tin phụ theo tham số được làm thích ứng vào trong thực thể bộ giải mã 720. Thực thể bộ giải mã 720 có thể

được tạo cấu hình.để giải mã một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng phụ thuộc vào thông tin phụ theo tham số được làm thích ứng.

Theo một phương án khác, bộ thích ứng thông tin phụ theo tham số 120 của thiết bị 710 làm thích ứng thông tin âm thanh đầu vào có thể được tạo cấu hình.để nhận dòng bit đầu vào gồm có thông tin phụ theo tham số đầu vào. Bộ thích ứng thông tin phụ theo tham số 120 của thiết bị 710 làm thích ứng thông tin âm thanh đầu vào có thể được tạo cấu hình.để thay thế thông tin phụ theo tham số đầu vào trong dòng bit đầu vào bởi thông tin phụ theo tham số được làm thích ứng để thu được dòng bit biến đổi. Bộ thích ứng thông tin phụ theo tham số 120 của thiết bị 710 làm thích ứng thông tin âm thanh đầu vào có thể được tạo cấu hình.để đưa dòng bit được biến đổi vào trong thực thể bộ giải mã 720. Hơn nữa, thực thể bộ giải mã 720 có thể được tạo cấu hình.để giải mã một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng phụ thuộc vào dòng bit được biến đổi.

Các Fig.8 và 9 mô tả hai khả năng để kết hợp thiết bị thích ứng thông tin âm thanh đầu vào vào trong chuỗi xử lý giải mã.

Cụ thể, Fig.8 minh họa việc ứng dụng PSIA chung bên trong sơ đồ mã hóa/giải mã theo một phương án.

Fig.8 minh họa nhiều thiết bị 800, 801, 802 để tạo ra một hoặc nhiều kênh âm thanh từ thông tin âm thanh đầu vào mã hóa một hoặc nhiều đối tượng âm thanh, trong đó thiết bị 800 để tạo ra một hoặc nhiều kênh âm thanh gồm có thiết bị 810 làm thích ứng thông tin âm thanh đầu vào và thực thể bộ giải mã 820, trong đó thiết bị 801 để tạo ra một hoặc nhiều kênh âm thanh gồm có thiết bị 811 làm thích ứng thông tin âm thanh đầu vào và thực thể bộ giải mã 821, và trong đó thiết bị 802 để tạo ra một hoặc nhiều kênh âm thanh gồm có thiết bị 812 làm thích ứng thông tin âm thanh đầu vào và thực thể bộ giải mã 822. Cần lưu ý rằng, ví dụ, thiết bị 800 tạo ra một hoặc nhiều kênh âm thanh, gồm có thiết bị 810 làm thích ứng thông tin âm thanh đầu vào, và thực thể bộ giải mã 820, không được xem xét như là phần cứng đơn lẻ 800 nhưng thay vào đó có thể được nhận biết bởi hai bộ phận riêng biệt 810, 820 được kết nối có dây hoặc được kết nối không dây.

Việc thực hiện chung (được tích hợp) của thiết bị làm thích ứng thông tin âm thanh đầu vào có thể được thực hiện để làm giảm độ phức tạp tính toán đối với việc giải mã (xem Fig.8). Ngoài ra, việc này cho phép thực hiện giao diện không được lượng tử hóa (không được mã hóa) giữa thiết bị làm thích ứng thông tin âm thanh đầu vào và bộ giải mã. Việc này có thể là liên quan cụ thể cho các thiết bị ứng dụng di động để làm giảm sự tiêu tốn năng lượng.

Fig.9 minh họa ứng dụng PSIA rời rạc trong sơ đồ mã hóa/giải mã theo một phương án.

Cụ thể, Fig.9 minh họa nhiều thiết bị 900, 901, 902 để tạo ra một hoặc nhiều kênh âm thanh từ thông tin âm thanh đầu vào mã hóa một hoặc nhiều đối tượng âm thanh, trong đó thiết bị 900 tạo ra một hoặc nhiều kênh âm thanh gồm có thiết bị 910 làm thích ứng thông tin âm thanh đầu vào và thực thể bộ giải mã 920, trong đó thiết bị 901 tạo ra một hoặc nhiều kênh âm thanh gồm có thiết bị 911 làm thích ứng thông tin âm thanh đầu vào và thực thể bộ giải mã 921, và trong đó thiết bị 902 để tạo ra một hoặc nhiều kênh âm thanh gồm có thiết bị 912 làm thích ứng thông tin âm thanh đầu vào và thực thể bộ giải mã 922. Lưu ý rằng, ví dụ, thiết bị 900 để tạo ra một hoặc nhiều kênh âm thanh, gồm có thiết bị 910 làm thích ứng thông tin âm thanh đầu vào và thực thể bộ giải mã 920, không được xem xét là phần cứng đơn lẻ 900, nhưng thay vào đó có thể được thấy bởi hai bộ riêng biệt 910, 920 được kết nối bằng dây hoặc được kết nối không dây.

Việc thực hiện rời rạc (được chia tách) của thiết bị làm thích ứng thông tin âm thanh đầu vào có thể được thấy để làm giảm kích cỡ/tốc độ bit của bit của dòng bit dữ liệu tương ứng, xem Fig.9. Điều này có thể là có liên quan cụ thể đối với các thiết bị ứng dụng di động với sự lưu trữ và khả năng truyền bị hạn chế và các hệ thống bộ điều khiển đa điểm (Multi-point Control Unit-MCU) với các kênh truyền dữ liệu hẹp.

Mặc dù một vài khía cạnh đã được mô tả trong ngữ cảnh của thiết bị, rõ ràng là các khía cạnh này cũng thể hiện việc mô tả phương pháp tương ứng, trong đó khối hoặc thiết bị tương ứng với bước xử lý của phương pháp hoặc đặc điểm của bước xử lý của phương pháp. Tương tự, các khía cạnh được mô tả trong ngữ cảnh của bước xử lý

của phương pháp cũng thể hiện việc mô tả khối hoặc mục tương ứng hoặc đặc điểm của thiết bị tương ứng.

Tín hiệu đã được tách ra có thể được lưu trữ trên vật ghi lưu trữ dạng số hoặc có thể được truyền dẫn trên vật ghi truyền dẫn như vật ghi truyền dẫn không dây hoặc vật ghi truyền dẫn có dây như Internet.

Phụ thuộc vào các yêu cầu thực hiện nhất định, các phương án của sáng chế có thể được thực hiện trong phần cứng hoặc trong phần mềm. Phương án có thể được thực hiện sử dụng vật ghi lưu trữ dạng số, ví dụ, đĩa mềm, DVD, CD, ROM, PROM, EPROM, EEPROM hoặc bộ nhớ FLASH, có các tín hiệu điều khiển có thể đọc được bằng điện tử được lưu trữ trên đó, mà kết hợp (hoặc có thể 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 gồm có vật mang dữ liệu không tạm thời có các tín hiệu điều khiển có thể đọc được bằng điện tử, chúng có khả năng kết hợp với hệ thống máy tính lập trình được, do đó một trong số các phương pháp được mô tả ở đây được thực hiện.

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

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

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

Phương án khác của các phương pháp theo sáng chế là, vật mang dữ liệu (hoặc vật ghi lưu trữ dạng số, hoặc vật ghi đọc được bằng máy tính) gồm có, đã được ghi lại trên đó, chương trình máy tính để thực hiện một trong số các phương pháp được mô tả ở đây.

Phương án khác của các phương pháp theo sáng chế là, dòng dữ liệu hoặc chuỗi tín hiệu thể hiện chương trình máy tính để thực hiện một trong số các phương pháp đã được mô tả ở đây. Dòng dữ liệu hoặc chuỗi các tín hiệu có thể, ví dụ, được tạo cấu hình để được truyền dẫn nhờ sự kết nối truyền thông dữ liệu, ví dụ thông qua Internet.

Phương án khác gồm có phương tiện 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 gồm có máy tính có cài đặt 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) có thể được sử dụng để thực hiện một số 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 có thể kết hợp với bộ vi xử lý để thực hiện một trong số các phương pháp được mô tả ở đây. Thông thường, các phương pháp ưu tiên được thực hiện bởi thiết bị phần cứng bất kỳ.

Các phương án được mô tả ở trên chỉ mang tính minh họa cho các nguyên lý của sáng chế. Cần hiểu rằng các biến thể và biến đổi của cách sắp xếp và các chi tiết được mô tả ở đây sẽ là rõ ràng đối với người có hiểu biết trong lĩnh vực kỹ thuật tương ứng. Do đó, sáng chế bị giới hạn chỉ bởi phạm vi của các điểm yêu cầu bảo hộ và không bởi các chi tiết cụ thể được thể hiện bởi bản mô tả và sự giải thích các phương án ở đây.

Tài liệu tham khảo

[MPS] ISO/IEC 23003-1:2007, MPEG-D (MPEG audio technologies), Part 1: MPEG Surround, 2007

[BCC] C. Faller and F. Baumgarte, “Binaural Cue Coding - Part II: Schemes and applications,” IEEE Trans. on Speech and Audio Proc., vol. 11, no. 6, Nov. 2003

[JSC] C. Faller, “Parametric Joint-Coding of Audio Sources”, 120th AES Convention, Paris, 2006

[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.

[ISS1] M. Parvaix and L. Girin: “Informed Source Separation of underdetermined instantaneous Stereo Mixtures using Source Index Embedding”, IEEE ICASSP, 2010

[ISS2] M. Parvaix, L. Girin, J.-M. Brossier: “A watermarking-based method for informed source separation of audio signals with a single sensor”, IEEE Transactions on Audio, Speech and Language Processing, 2010

[ISS3] A. Liutkus and J. Pinel and R. Badeau and L. Girin and G. Richard: “Informed source separation through spectrogram coding and data embedding”, Signal Processing Journal, 2011

[ISS4] A. Ozerov, A. Liutkus, R. Badeau, G. Richard: “Informed source separation: source coding meets source separation”, IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2011

[ISS5] Shuhua Zhang and Laurent Girin: “An Informed Source Separation System for Speech Signals”, INTERSPEECH, 2011

[ISS6] L. Girin and J. Pintel: “Informed Audio Source Separation from Compressed Linear Stereo Mixtures”, AES 42nd International Conference: Semantic Audio, 2011

## YÊU CẦU BẢO HỘ

1. Thiết bị làm thích ứng thông tin âm thanh đầu vào, mã hóa một hoặc nhiều đối tượng âm thanh, để thu được thông tin âm thanh được làm thích ứng, trong đó thông tin âm thanh đầu vào gồm có hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào và còn gồm có thông tin phụ theo tham số đầu vào, trong đó thông tin âm thanh được làm thích ứng gồm có một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng và còn gồm có thông tin phụ theo tham số được làm thích ứng, trong đó thiết bị gồm có:

bộ biến đổi tín hiệu trộn giảm (110) làm thích ứng, phụ thuộc vào thông tin thích ứng, hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào để thu được một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng, và

bộ thích ứng thông tin phụ theo tham số (120) làm thích ứng, phụ thuộc vào thông tin thích ứng, thông tin phụ theo tham số đầu vào để thu được thông tin phụ theo tham số được làm thích ứng,

trong đó thông tin thích ứng gồm có ma trận thích ứng ( $\mathbf{D}_{dmx}^{DSM}$ ),

trong đó bộ biến đổi tín hiệu trộn giảm (110) được tạo cấu hình để làm thích ứng, phụ thuộc vào ma trận thích ứng ( $\mathbf{D}_{dmx}^{DSM}$ ), hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào ( $\mathbf{X}_{dmx}^{ENC}$ ) để thu được một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng ( $\mathbf{X}_{dmx}^{DSM}$ ),

trong đó bộ thích ứng thông tin phụ theo tham số (120) được tạo cấu hình để làm thích ứng, phụ thuộc vào ma trận thích ứng ( $\mathbf{D}_{dmx}^{DSM}$ ), thông tin phụ theo tham số đầu vào ( $\mathbf{D}_{dmx}^{ENC}$ ) để thu được thông tin phụ theo tham số được làm thích ứng ( $\mathbf{D}_{dmx}^{PSI}$ ).

2. Thiết bị theo điểm 1,

trong đó thông tin phụ theo tham số đầu vào ( $\mathbf{D}_{dmx}^{enc}$ ) biểu thị ma trận trộn ban đầu, sao cho bằng cách áp dụng ma trận trộn ban đầu ( $\mathbf{D}_{dmx}^{enc}$ ) trên một hoặc nhiều đối tượng âm thanh ( $\mathbf{S}$ ), hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào ( $\mathbf{X}_{dmx}^{enc}$ ) được thu, và

trong đó bộ thích ứng thông tin phụ theo tham số (120) được tạo cấu hình để xác định ma trận trộn giảm được làm thích ứng ( $\mathbf{D}_{dmx}^{PSI}$ ) như là thông tin phụ theo tham số được làm thích ứng, sao cho bằng cách áp dụng ma trận trộn giảm được làm thích ứng ( $\mathbf{D}_{dmx}^{PSI}$ ) trên một hoặc nhiều đối tượng âm thanh ( $\mathbf{S}$ ), một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng ( $\mathbf{X}_{dmx}^{DSM}$ ) được thu.

3. Thiết bị theo điểm 1 hoặc 2, trong đó bộ biến đổi tín hiệu trộn giảm (110) được tạo cấu hình để làm thích ứng hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào phụ thuộc vào thông tin thích ứng, sao cho số lượng của một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng là nhỏ hơn số lượng của hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào.

4. Thiết bị theo điểm bất kỳ trong số các điểm từ 1 đến 3, trong đó thông tin thích ứng phụ thuộc vào thực thể bộ giải mã, và trong đó bộ biến đổi tín hiệu trộn giảm (110) được tạo cấu hình để làm thích ứng hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào phụ thuộc vào thực thể bộ giải mã.

5. Thiết bị theo điểm 4,

trong đó thực thể bộ giải mã có khả năng giải mã tối đa một lượng lớn nhất kênh trộn giảm,

trong đó thông tin thích ứng phụ thuộc vào số lượng lớn nhất của kênh trộn giảm đã nêu, và

trong đó bộ biến đổi tín hiệu trộn giảm (110) được tạo cấu hình để làm thích ứng hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào phụ thuộc vào thông tin thích ứng để thu được một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng, sao cho số lượng của một hoặc nhiều kênh trộn giảm được làm thích ứng bằng với số lượng lớn nhất của các kênh trộn giảm đã nêu.

6. Thiết bị theo điểm bất kỳ trong số các điểm nêu trên, trong đó bộ biến đổi tín hiệu trộn giảm (110) được tạo cấu hình để làm thích ứng, phụ thuộc vào ma trận thích ứng  $\mathbf{D}_{dmx}^{DSM}$ , hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào  $\mathbf{X}_{dmx}^{ENC}$  để thu

được một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng  $\mathbf{X}_{dmx}^{DSM}$  bằng cách áp dụng công thức:

$$\mathbf{X}_{dmx}^{DSM} = \mathbf{D}_{dmx}^{DSM} \mathbf{X}_{dmx}^{ENC}$$

7. Thiết bị theo điểm bất kỳ trong số các điểm nêu trên, trong đó bộ thích ứng thông tin phụ theo tham số (120) được tạo cấu hình để làm thích ứng, phụ thuộc vào ma trận thích ứng  $\mathbf{D}_{dmx}^{DSM}$ , thông tin phụ theo tham số đầu vào  $\mathbf{D}_{dmx}^{ENC}$  để thu được thông tin phụ theo tham số được làm thích ứng  $\mathbf{D}_{dmx}^{PSI}$  bằng cách áp dụng công thức:

$$\mathbf{D}_{dmx}^{PSI} = \mathbf{D}_{dmx}^{DSM} \mathbf{D}_{dmx}^{ENC}$$

8. Thiết bị (700; 800, 801, 802; 900, 901, 902) tạo ra một hoặc nhiều kênh âm thanh từ thông tin âm thanh đầu vào mã hóa một hoặc nhiều đối tượng âm thanh, trong đó thiết bị gồm có:

thiết bị (710; 810, 811, 812; 910, 911, 912) theo điểm bất kỳ trong số các điểm từ 1 đến 6 làm thích ứng thông tin âm thanh đầu vào để thu được thông tin âm thanh được làm thích ứng, trong đó thông tin âm thanh đầu vào gồm có hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào và còn gồm có thông tin phụ theo tham số đầu vào, trong đó thông tin âm thanh được làm thích ứng gồm có một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng và còn gồm có thông tin phụ theo tham số được làm thích ứng, và

thực thể bộ giải mã (720; 820, 821, 822; 920, 921, 922) để giải mã, phụ thuộc vào thông tin phụ theo tham số được làm thích ứng, một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng để thu được một hoặc nhiều kênh âm thanh.

9. Thiết bị (700; 800, 801, 802) theo điểm 8,

trong đó bộ thích ứng thông tin phụ theo tham số (120) của thiết bị (710; 810, 811, 812) theo điểm bất kỳ trong số các điểm từ 1 đến 7 được tạo cấu hình để nhận dòng bit đầu vào gồm có thông tin phụ theo tham số đầu vào,

trong đó bộ thích ứng thông tin phụ theo tham số (120) của thiết bị (710; 810, 811, 812) theo điểm bất kỳ trong số các điểm từ 1 đến 7 được tạo cấu hình để làm

thích ứng thông tin phụ theo tham số đầu vào để thu được thông tin phụ theo tham số được làm thích ứng, và đưa thông tin phụ theo tham số được làm thích ứng vào trong thực thể bộ giải mã (720; 820, 821, 822), và

trong đó thực thể bộ giải mã (720; 820, 821, 822) được tạo cấu hình để giải mã một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng phụ thuộc vào thông tin phụ theo tham số được làm thích ứng.

10. Thiết bị (700; 900, 901, 902) theo điểm 8,

trong đó bộ thích ứng thông tin phụ theo tham số (120) của thiết bị (710; 910, 911, 912) theo điểm bất kỳ trong số các điểm từ 1 đến 7 được tạo cấu hình để nhận dòng bit đầu vào gồm có thông tin phụ theo tham số đầu vào,

trong đó bộ thích ứng thông tin phụ theo tham số (120) của thiết bị (710; 910, 911, 912) theo điểm bất kỳ trong số các điểm từ 1 đến 7 được tạo cấu hình để thay thế thông tin phụ theo tham số đầu vào trong dòng bit đầu vào bởi thông tin phụ theo tham số được làm thích ứng để thu được dòng bit biến đổi,

trong đó bộ thích ứng thông tin phụ theo tham số (120) của thiết bị (710; 910, 911, 912) theo điểm bất kỳ trong số các điểm từ 1 đến 7 được tạo cấu hình để đưa dòng bit biến đổi vào trong thực thể bộ giải mã (720; 920, 921, 922), và

trong đó thực thể bộ giải mã (720; 920, 921, 922) được tạo cấu hình để giải mã một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng phụ thuộc vào dòng bit biến đổi.

11. Phương pháp làm thích ứng thông tin âm thanh đầu vào, mã hóa một hoặc nhiều đối tượng âm thanh, để thu được thông tin âm thanh được làm thích ứng, trong đó thông tin âm thanh đầu vào gồm có hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào và còn gồm có thông tin phụ theo tham số đầu vào, trong đó thông tin âm thanh được làm thích ứng gồm có một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng và còn gồm có thông tin phụ theo tham số được làm thích ứng, trong đó phương pháp bao gồm các bước:

làm thích ứng, phụ thuộc vào thông tin thích ứng, hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào để thu được một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng, và

làm thích ứng, phụ thuộc vào thông tin thích ứng, thông tin phụ theo tham số đầu vào để thu được thông tin phụ theo tham số được làm thích ứng,

trong đó thông tin thích ứng gồm có ma trận thích ứng ( $\mathbf{D}_{dmx}^{DSM}$ ),

trong đó bước làm thích ứng hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào gồm có làm thích ứng, phụ thuộc vào ma trận thích ứng ( $\mathbf{D}_{dmx}^{DSM}$ ), hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào ( $\mathbf{X}_{dmx}^{ENC}$ ) để thu được một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng ( $\mathbf{X}_{dmx}^{DSM}$ ),

trong đó bước làm thích ứng thông tin phụ theo tham số đầu vào gồm có làm thích ứng, phụ thuộc vào ma trận thích ứng ( $\mathbf{D}_{dmx}^{DSM}$ ), thông tin phụ theo tham số đầu vào ( $\mathbf{D}_{dmx}^{ENC}$ ) để thu được thông tin phụ theo tham số được làm thích ứng ( $\mathbf{D}_{dmx}^{PSI}$ ).

12. Phương pháp theo điểm 11,

trong đó thông tin phụ theo tham số đầu vào ( $\mathbf{D}_{dmx}^{enc}$ ) biểu thị ma trận trộn giảm ban đầu, sao cho bằng cách áp dụng ma trận trộn giảm ban đầu ( $\mathbf{D}_{dmx}^{enc}$ ) trên một hoặc nhiều đối tượng âm thanh ( $\mathbf{S}$ ), hai hoặc nhiều hơn hai kênh trộn giảm âm thanh đầu vào ( $\mathbf{X}_{dmx}^{enc}$ ) được thu, và

trong đó bước làm thích ứng thông tin phụ theo tham số đầu vào gồm có việc xác định ma trận trộn giảm được làm thích ứng ( $\mathbf{D}_{dmx}^{PSI}$ ) như là thông tin phụ theo tham số được làm thích ứng, sao cho bằng cách áp dụng ma trận trộn giảm được làm thích ứng ( $\mathbf{D}_{dmx}^{PSI}$ ) trên một hoặc nhiều đối tượng âm thanh ( $\mathbf{S}$ ), một hoặc nhiều kênh trộn giảm âm thanh được làm thích ứng ( $\mathbf{X}_{dmx}^{DSM}$ ) được thu.

13. Vật ghi đọc được bằng máy tính bao gồm chương trình máy tính để thực hiện phương pháp theo điểm 11 hoặc 12 khi được chạy trên máy tính hoặc bộ xử lý tín hiệu.

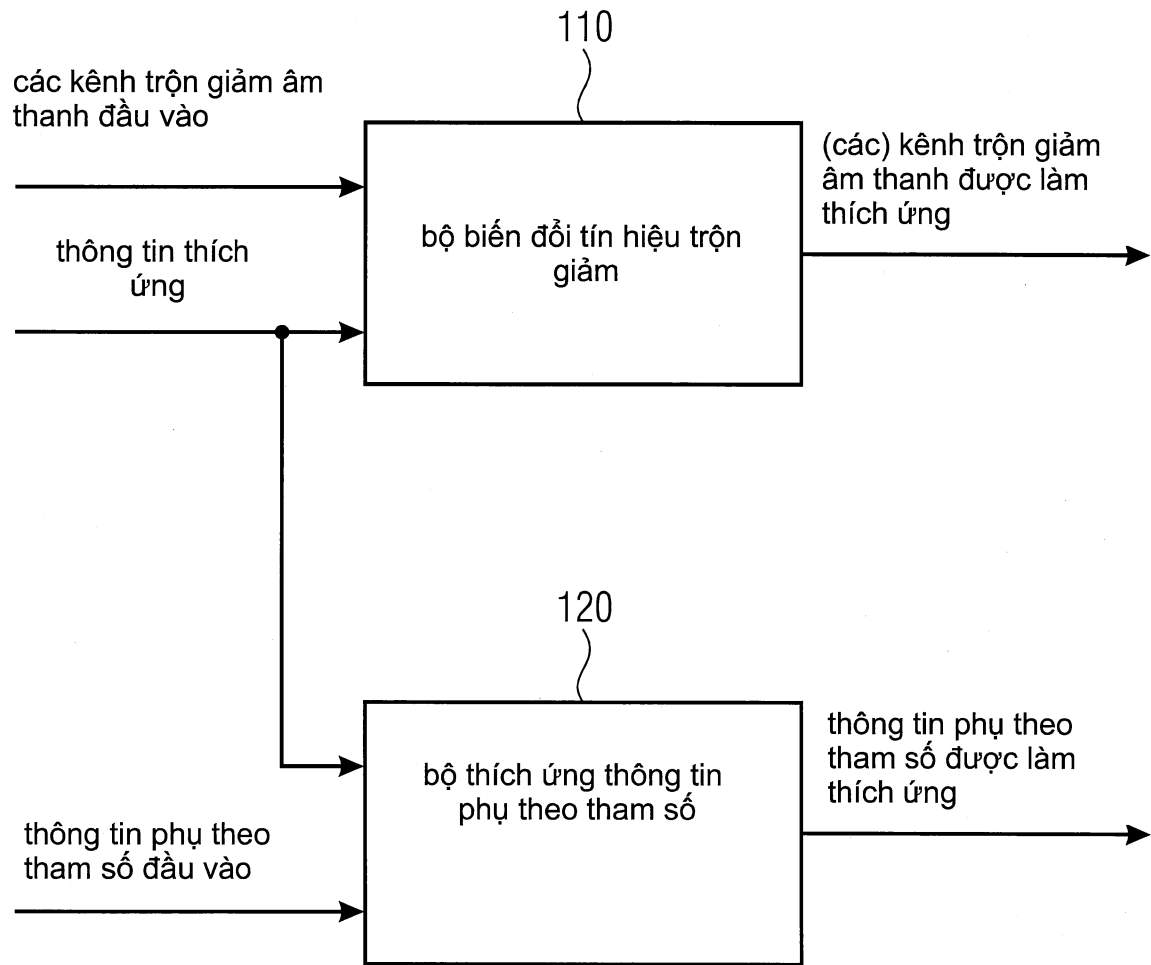


Fig.1

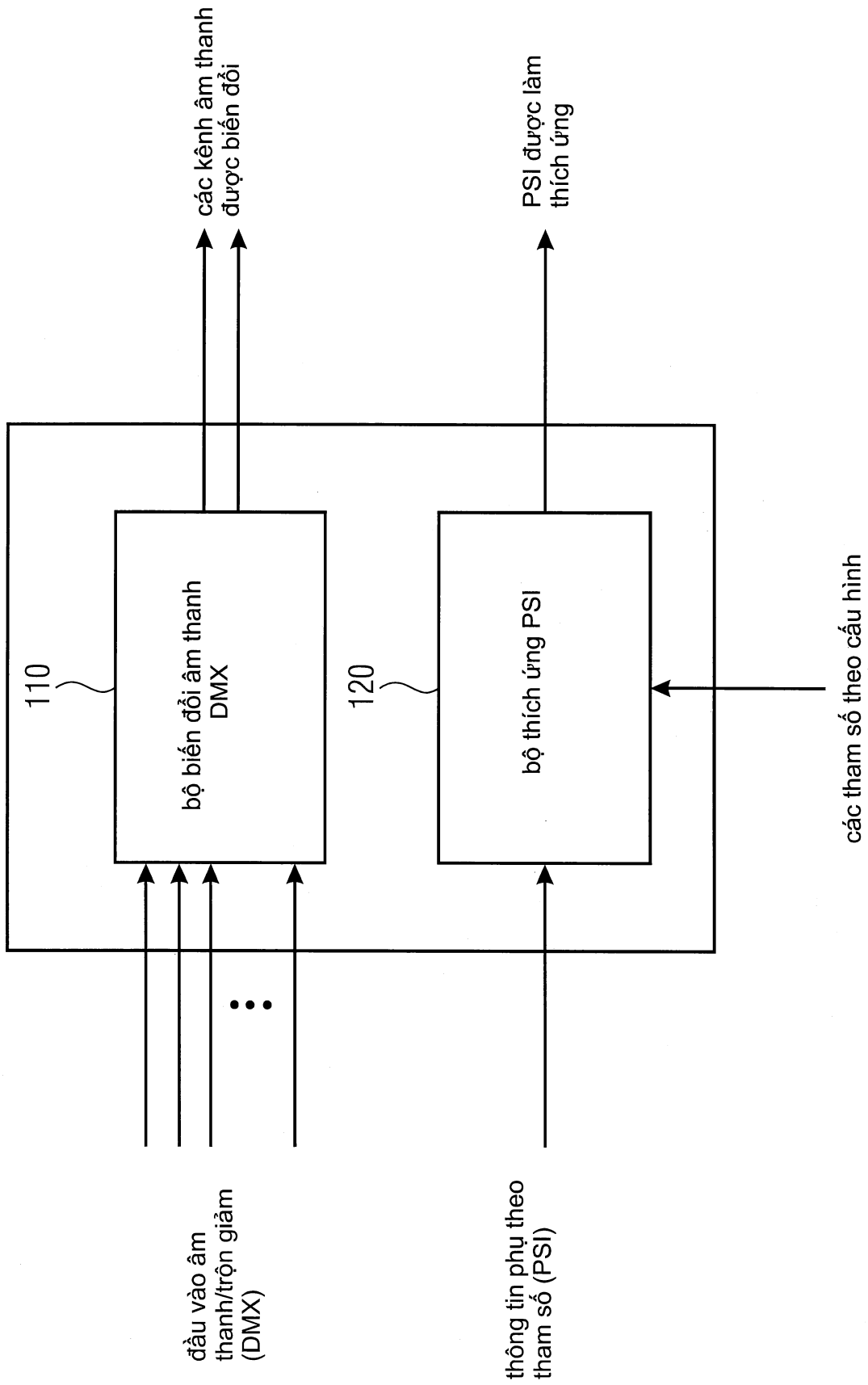


Fig.2

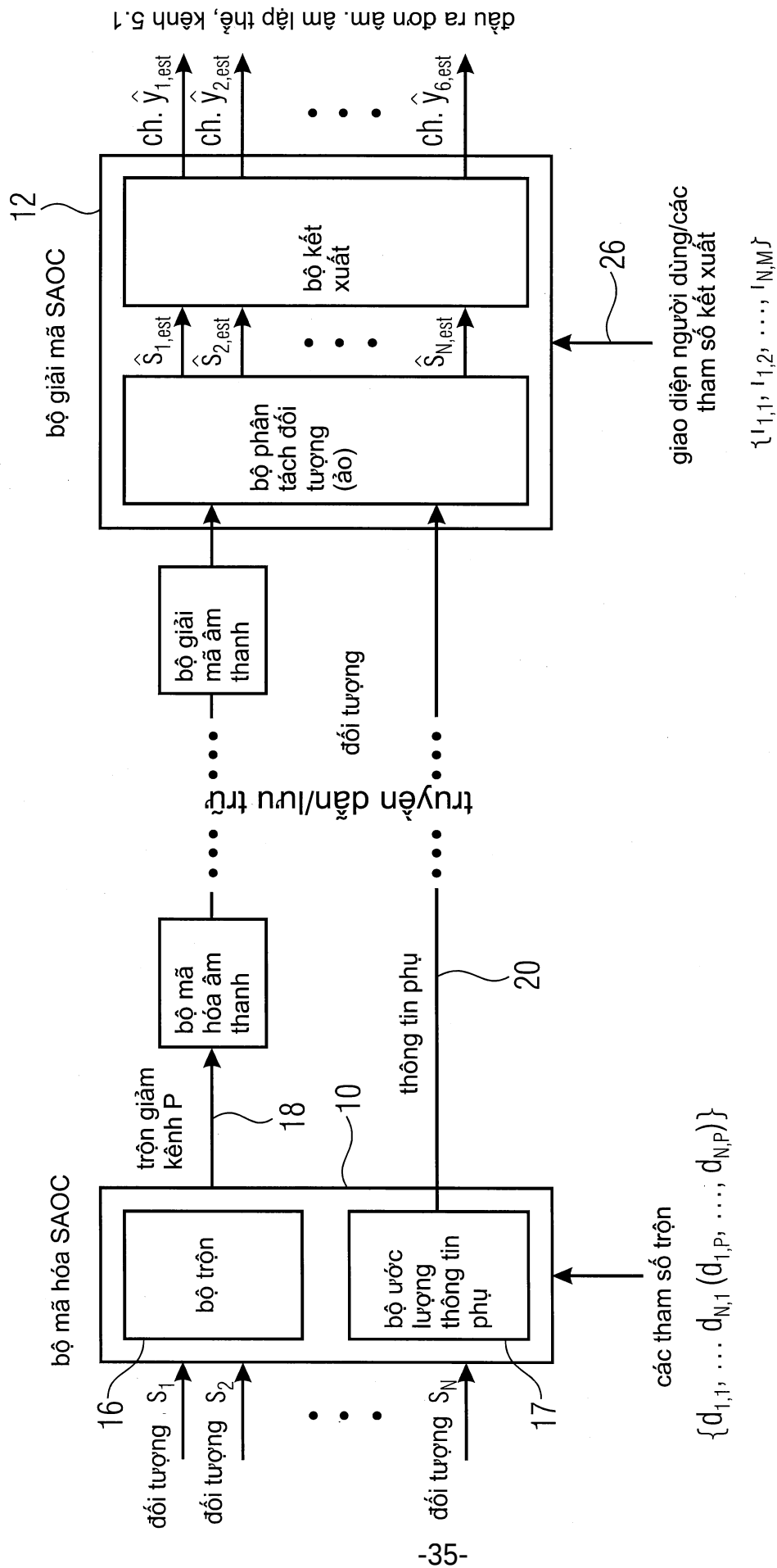


Fig.3

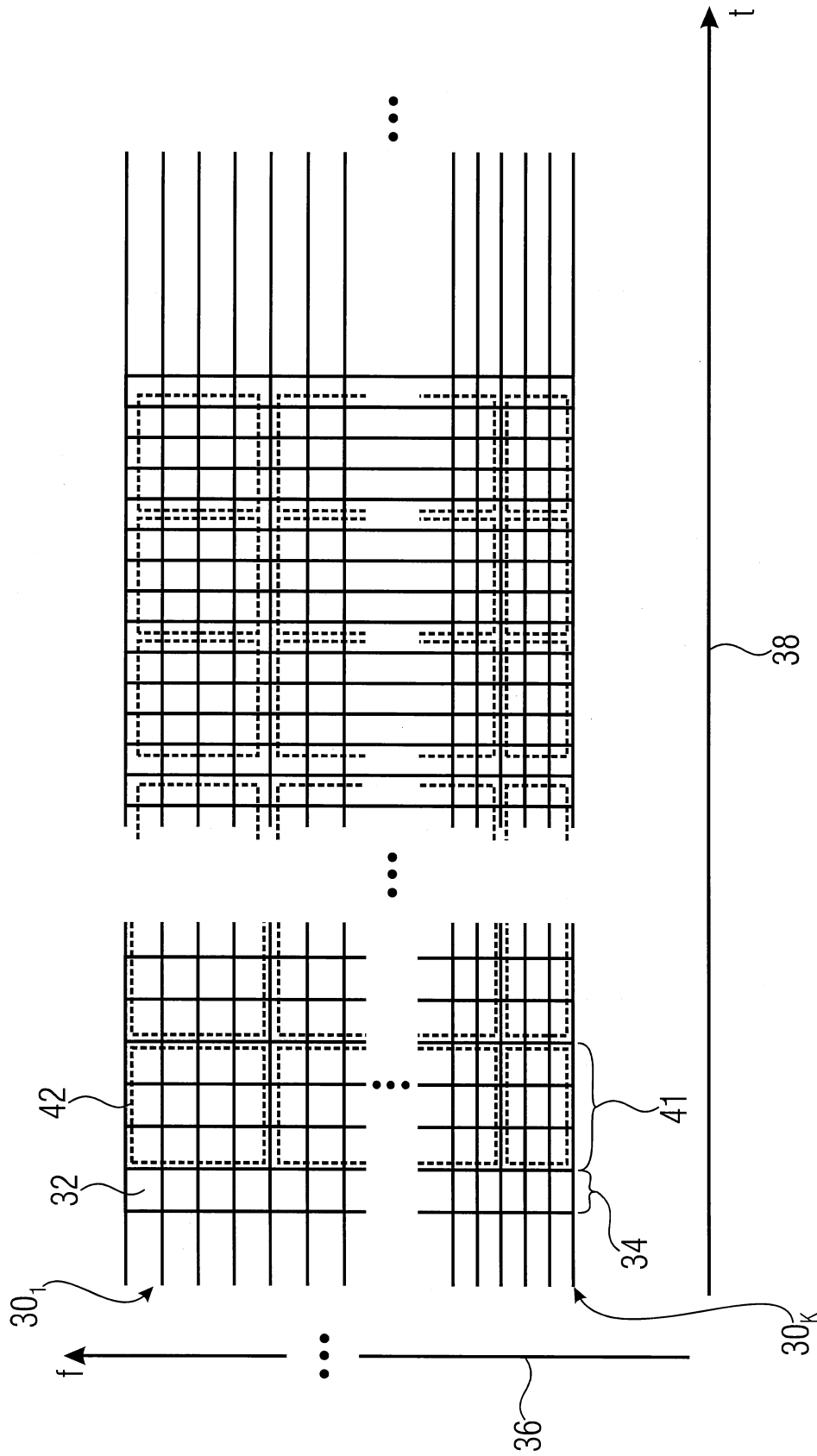


Fig.4

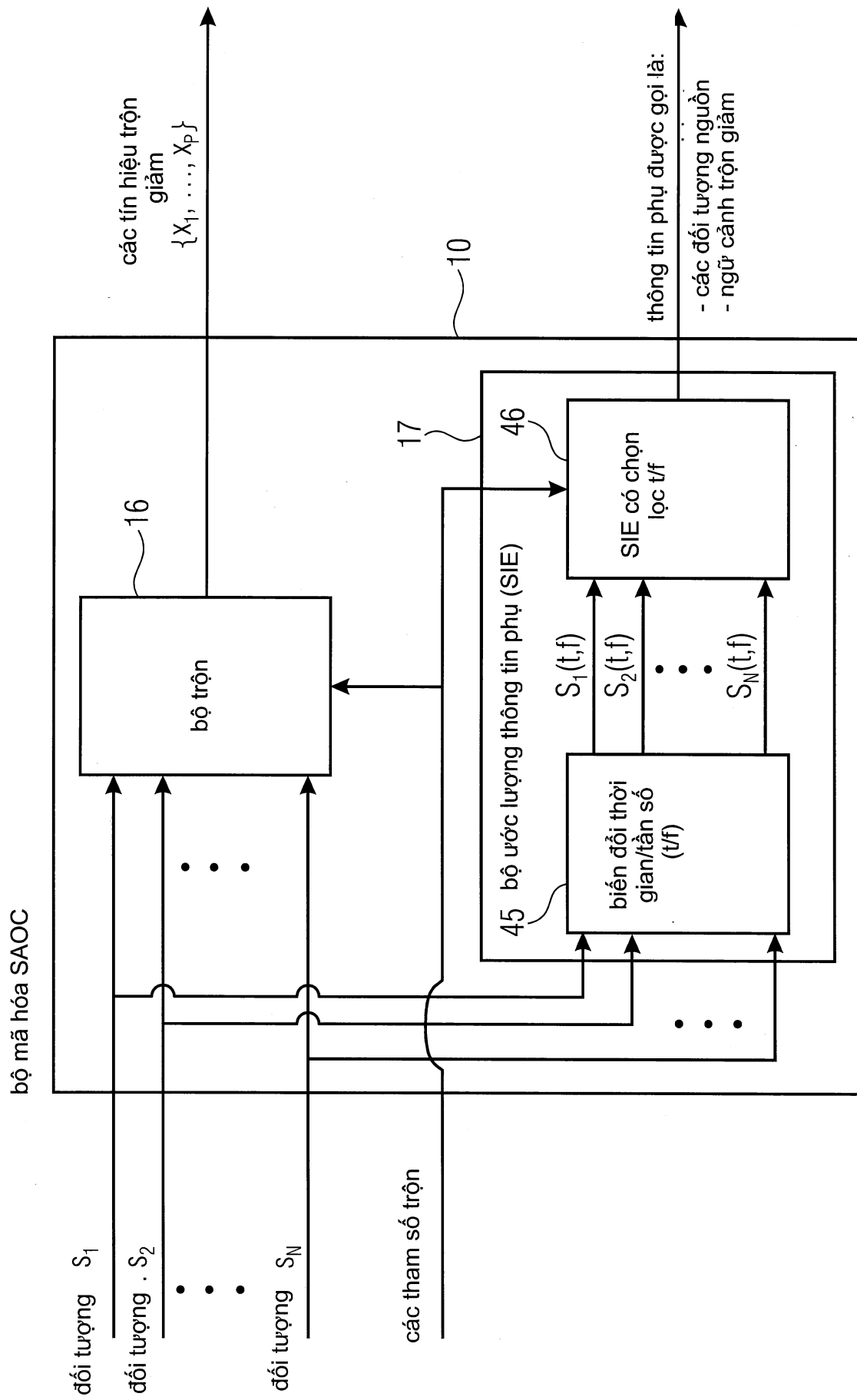


Fig.5

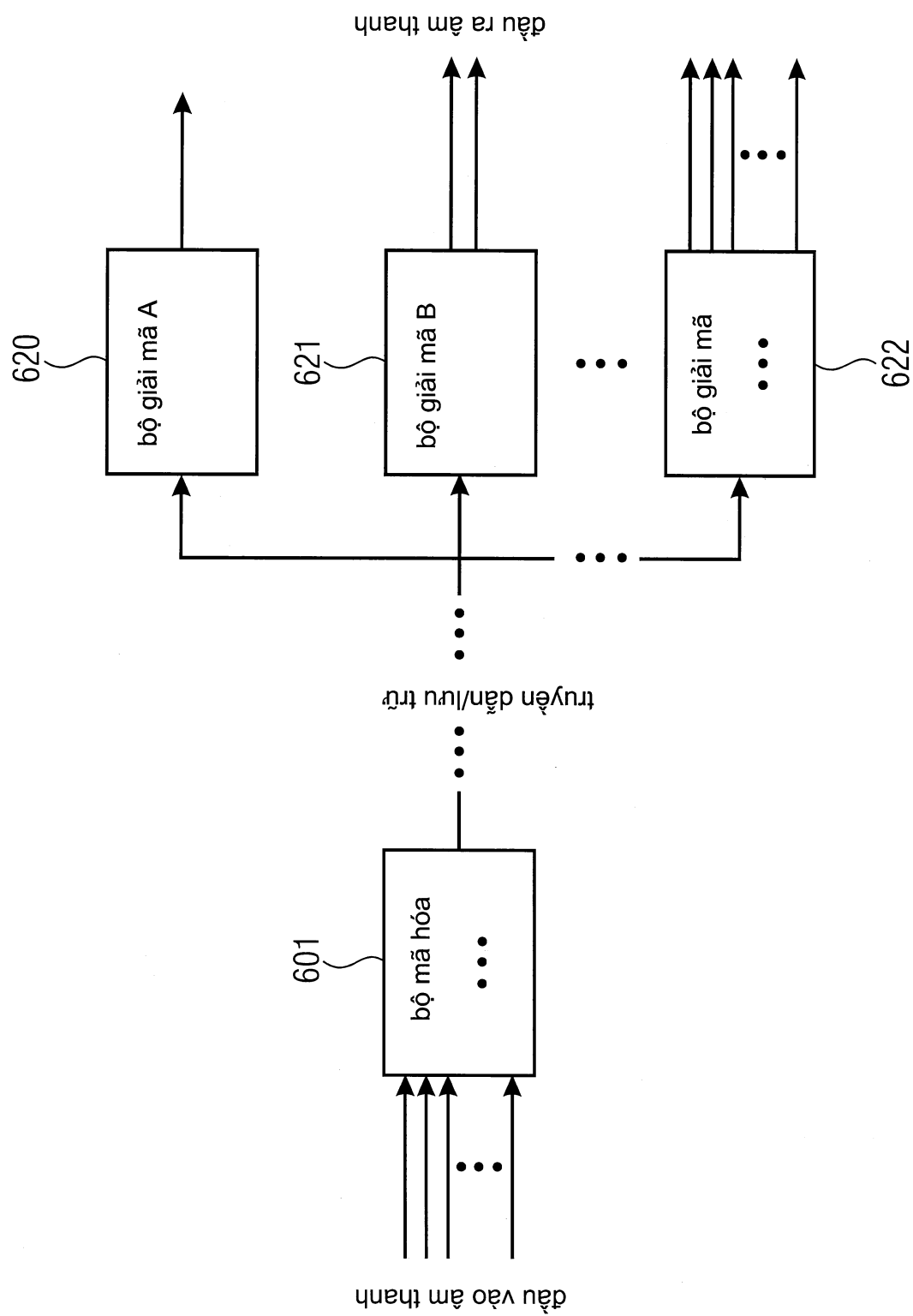


Fig.6

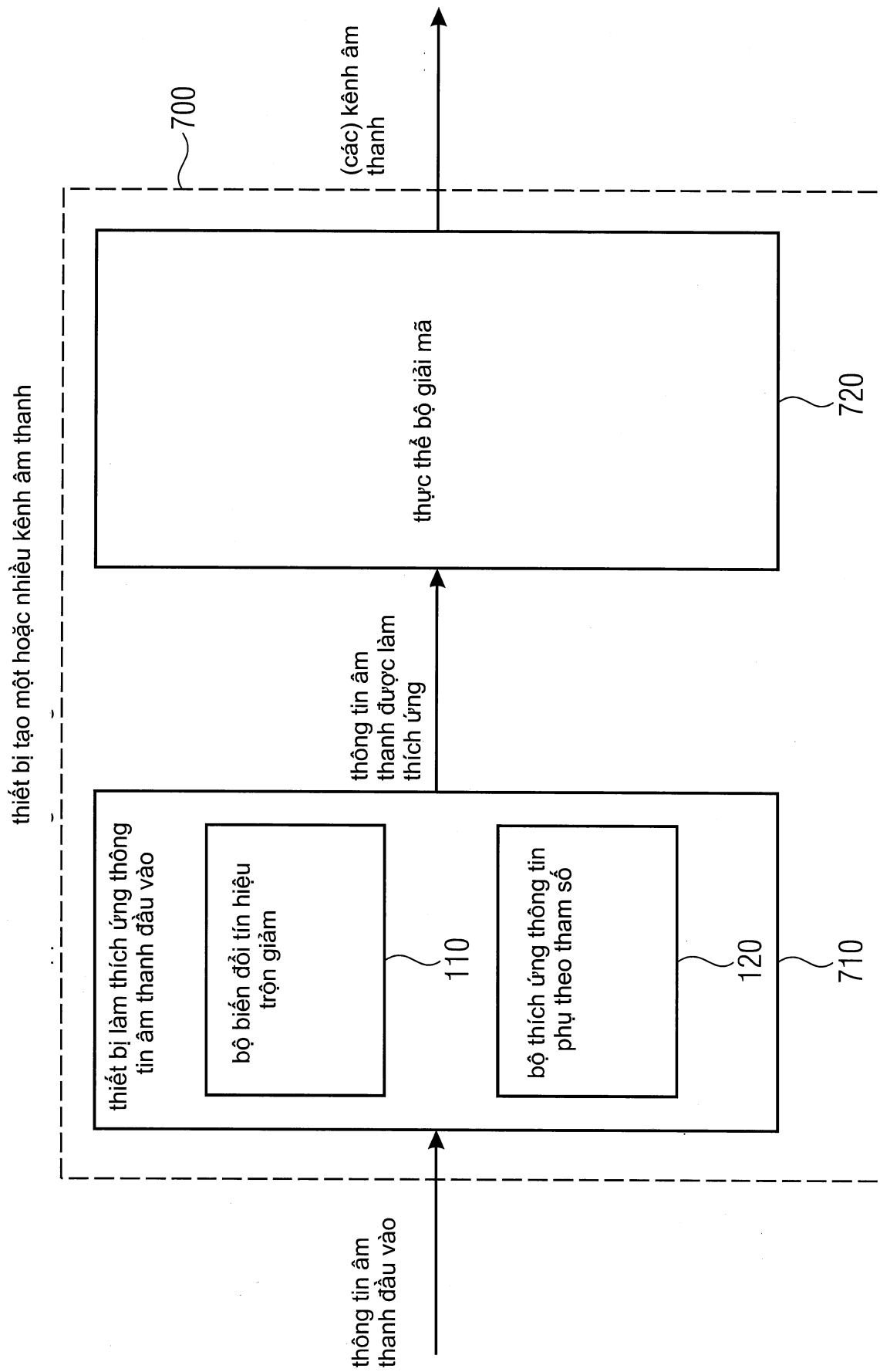


Fig.7

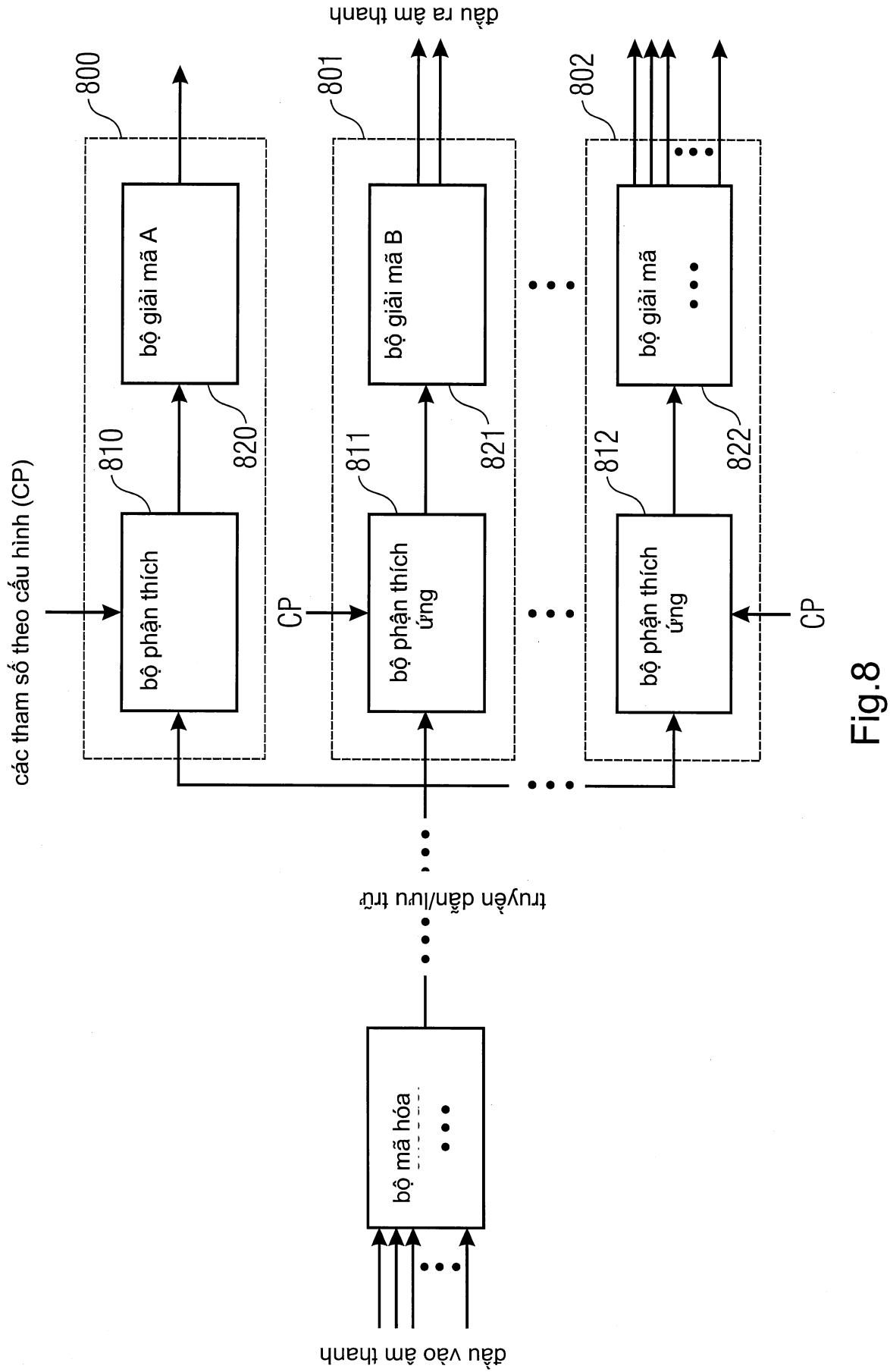


Fig.8

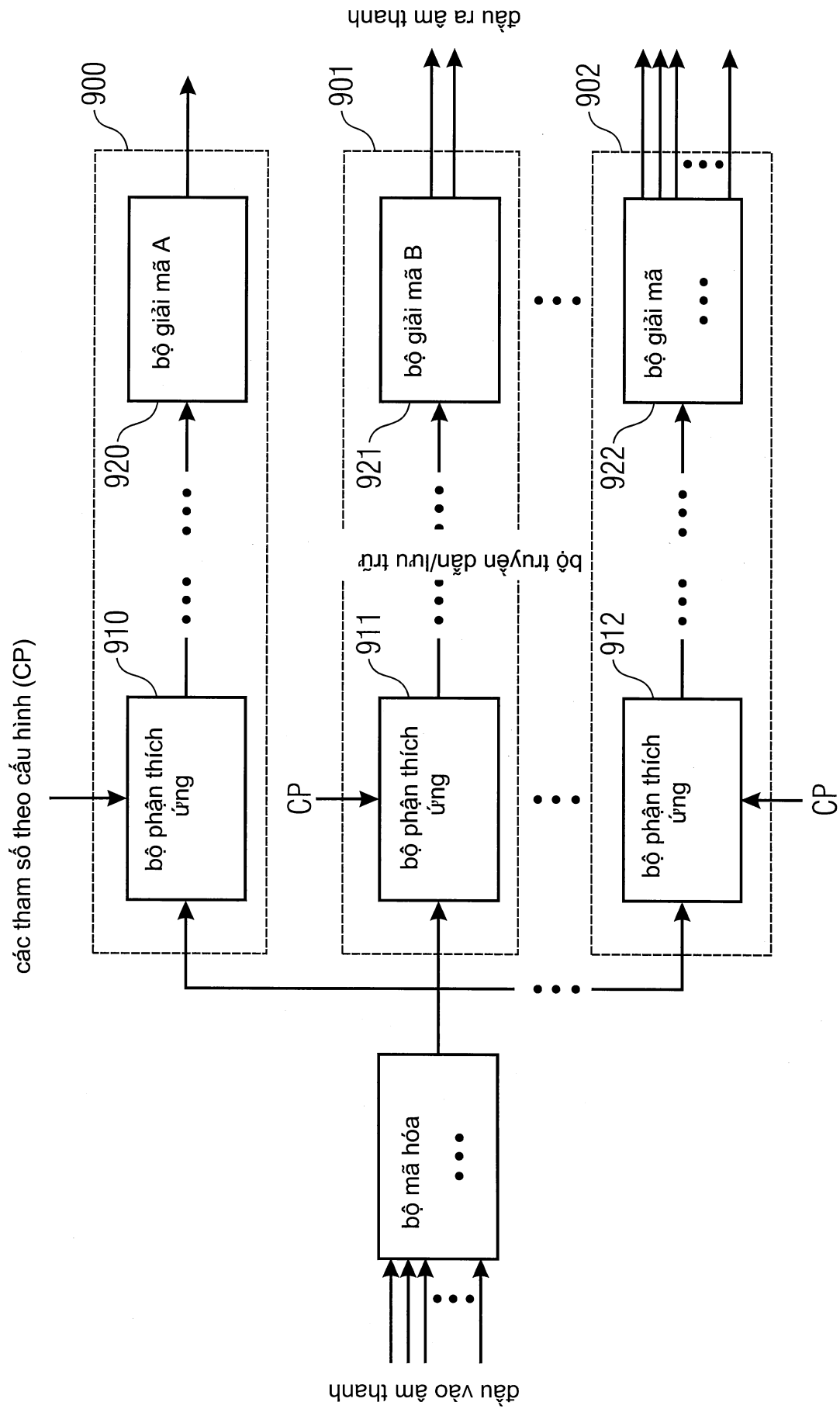


Fig.9