



(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)



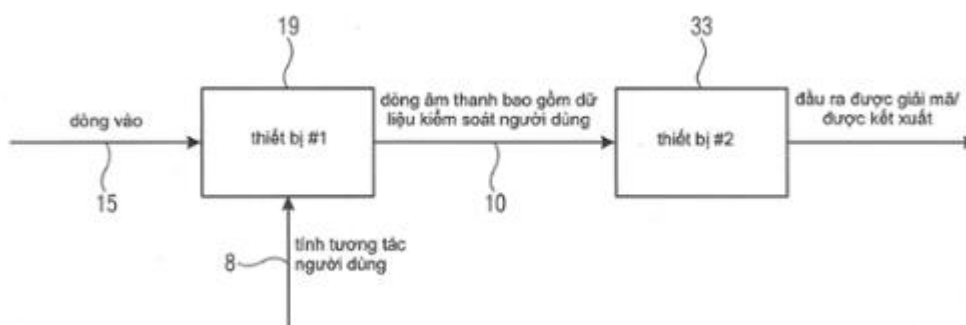
1-0029416

(51)⁷ H04N 21/81; H04N 21/435; H04N 21/4363; H04N 21/442; H04N 21/485; G10L 19/00; H04N 21/439 (13) B

- (21) 1-2016-05065 (22) 27/03/2015
(86) PCT/EP2015/056768 27/03/2015 (87) WO 2015/180866 A1 03/12/2015
(30) 14170416.3 28/05/2014 EP
(45) 25/09/2021 402 (43) 27/03/2017 348A
(73) FRAUNHOFER-GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN FORSCHUNG E.V. (DE)
Hansastraße 27c, 80686 Muenchen, Germany
(72) SCHREINER, Stephan (DE); FUEG, Simone (DE); FUCHS, Harald (DE); PLOGSTIES, Jan (DE); DOEHLA, Stefan (DE).
(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS HANOI)

(54) BỘ XỬ LÝ DỮ LIỆU ÂM THANH VÀ PHƯƠNG PHÁP XỬ LÝ DỮ LIỆU ÂM THANH

(57) Sáng chế đề cập đến bộ xử lý dữ liệu âm thanh và phương pháp xử lý dữ liệu âm thanh. Bộ xử lý dữ liệu âm thanh, bao gồm: giao diện bộ nhận để nhận dữ liệu âm thanh được mã hóa và siêu dữ liệu liên quan đến dữ liệu âm thanh được mã hóa; bộ phân tích siêu dữ liệu để phân tích siêu dữ liệu để xác định khả năng thao tác dữ liệu âm thanh; giao diện tương tác để nhận đầu vào tương tác và để tạo ra, từ đầu vào tương tác, dữ liệu điều khiển tương tác liên quan đến khả năng thao tác dữ liệu âm thanh; và bộ tạo dòng dữ liệu để thu được dữ liệu điều khiển tương tác và dữ liệu âm thanh được mã hóa và siêu dữ liệu và để tạo ra dòng dữ liệu đầu ra, dòng dữ liệu đầu ra bao gồm dữ liệu âm thanh được mã hóa, ít nhất một phần siêu dữ liệu, và dữ liệu điều khiển tương tác.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến bộ xử lý dữ liệu âm thanh, phương pháp xử lý dữ liệu âm thanh và chương trình máy tính để thực hiện phương pháp xử lý dữ liệu âm thanh.

Tình trạng kỹ thuật của sáng chế

Trong việc lắp đặt điện tử học dân dụng (Consumer Electronics - CE) trong gia đình, chức năng phổ biến trên các thiết bị được kết nối thông qua các giao diện được chuẩn hóa. Hơn nữa, dụng cụ (chất lượng cao) thường được xây dựng không chỉ trong thiết bị đơn, mà trong cả các thiết bị đơn phức tạp cũng có khả năng (xem xét các hộp chuyển đổi tín hiệu cáp tivi, thiết lập TV, bộ nhận AVR). Các thiết bị này giao tiếp thông qua các giao diện được chuẩn hóa (như HDMI).

Trong khi thiết bị thứ nhất trích các dòng mong muốn và cung cấp tất cả các giao diện cho người dùng thì thiết bị thứ hai thường thực hiện việc giải mã trong "chế độ thụ động" mà không có bất kỳ giao diện nào cho người dùng. Khi nó đi đến sự tương tác người dùng và sự kiểm soát của bộ giải mã, thì cần thiết truyền đạt thông tin người dùng này từ thiết bị #1 đến thiết bị #2 trong kịch bản này.

Ví dụ, như được thể hiện trên Fig.9, chương trình tivi thường được nhận bởi thiết bị thứ nhất như hộp chuyển đổi tín hiệu cáp tivi, mà lựa chọn các kênh truyền thích hợp và trích các dòng cơ bản liên quan chứa bản chất được tạo mã mong muốn. Các dòng được trích này có thể được nạp vào thiết bị thứ hai như bộ nhận âm thanh-video để tái sản xuất. Việc truyền giữa hai thiết bị này có thể được hoàn thành bằng cách hoặc là truyền sự biểu diễn được giải mã/giải nén (âm thanh PCM), hoặc trong sự biểu diễn được mã hóa, đặc biệt là nếu các giới hạn băng thông áp dụng trên đường liên kết được sử dụng.

Hơn nữa, vì việc lựa chọn các dòng mong muốn và/hoặc giao diện người dùng tùy chọn được hoàn thành trong thiết bị #1 (ví dụ hộp chuyển đổi tín hiệu cáp tivi), trong hầu hết các trường hợp, chỉ thiết bị này cung cấp giao diện điều khiển cho người dùng. Thiết bị thứ hai (ví dụ bộ nhận A/V) chỉ cung cấp giao diện cấu hình mà thường được truy cập chỉ một lần bởi người dùng khi thiết lập hệ thống và tác động theo "chế độ phụ thuộc" tại những lần hoạt động thông thường.

Các sơ đồ bộ mã hóa-giải mã âm thanh hiện đại không chỉ hỗ trợ việc mã hóa

của các tín hiệu âm thanh mà còn cung cấp các phương tiện cho tính tương tác người dùng để làm thích ứng việc phát lại âm thanh và để kết xuất thành sự ưu tiên của người nghe. Dòng dữ liệu âm thanh bao gồm số lượng các tín hiệu âm thanh được mã hóa, ví dụ các tín hiệu kênh hoặc các đối tượng âm thanh và thông tin siêu dữ liệu kèm theo mà mô tả cách các tín hiệu âm thanh này tạo thành cảnh âm thanh mà được kết xuất đến các loa phóng thanh.

Các ví dụ cho các đối tượng âm thanh là:

- cuộc hội thoại bằng các ngôn ngữ khác nhau,
- cuộc hội thoại bổ sung như mô tả âm thanh, hoặc
- âm nhạc và các hiệu ứng nền.

Các ví dụ cho thông tin siêu dữ liệu là:

- mức âm lượng mặc định của từng tín hiệu đối tượng (tức là nó phải được trộn thành tín hiệu hỗn hợp để biểu diễn loa phóng thanh với độ to như thế nào),
- vị trí không gian mặc định (tức là nó phải được kết xuất ở đâu),
- thông tin, nếu sự tương tác người dùng được cho phép cho đối tượng cụ thể, hoặc
- thông tin về cách người dùng được cho phép để tương tác, ví dụ các mức âm lượng tối thiểu/tối đa hoặc các giới hạn trên các vị trí người dùng có thể quét lại các đối tượng đến.
- sự phân loại và/hoặc mô tả của các đối tượng âm thanh

Để hoàn thành tính tương tác người dùng, các bộ giải mã/bộ kết xuất âm thanh (ví dụ thiết bị #2) cần cung cấp giao diện (đầu vào hoặc tương tác) bổ sung cho thông tin điều khiển đối với sự tương tác người dùng mong muốn.

Cách khác cũng có thể là đáng mong muốn để thực hiện sự điều khiển người dùng để lựa chọn và thao tác đối tượng âm thanh trong thiết bị #1 và nạp dữ liệu này vào thiết bị #2 khi việc giải mã và kết xuất được thực hiện trong thiết bị #2 và không được thực hiện trong thiết bị #1.

Tuy nhiên, việc truyền dữ liệu này bị giới hạn do thực tế là sự kết nối được chuẩn hóa hiện thời không hỗ trợ việc truyền dữ liệu điều khiển người dùng và/hoặc thông tin bộ kết xuất.

Theo cách khác, sự lựa chọn các dòng và sự tương tác người dùng như được mô

tả ở trên cho thiết bị #1, và việc giải mã như được mô tả ở trên cho thiết bị #2 có thể được xử lý bởi hai thành phần có chức năng riêng được chứa bên trong cùng thiết bị và có cùng các giới hạn trên việc truyền dữ liệu giữa cả hai thành phần, cụ thể là chỉ một giao diện cho dữ liệu được tạo mã và dữ liệu tương tác người dùng là có khả năng, tốt hơn là giao diện tương tác của thiết bị #1, trong khi giao diện thứ hai cho dữ liệu tương tác người dùng, tức là giao diện thường được cung cấp bởi thiết bị #2, có thể được bỏ qua. Thậm chí cả thiết bị #1 và thiết bị #2 được chứa hoặc được thực hiện trong cùng thiết bị (phần cứng), điều này dẫn đến cùng tình trạng như được mô tả đối với trường hợp của thiết bị được phân tách #1 và #2.

Để hoàn thành trường hợp sử dụng được mô tả và để khắc phục các hạn chế được mô tả ở trên, sáng chế được đề xuất để lòng dữ liệu thông tin điều khiển người dùng, hoặc dữ liệu tương tác nói chung, vào dòng dữ liệu âm thanh được mã hóa.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là cải thiện các bộ xử lý dữ liệu âm thanh hiện thời.

Nói chung, thiết bị thứ nhất có thể được tạo cấu hình như bộ xử lý dữ liệu âm thanh, bao gồm: giao diện bộ nhận để nhận dữ liệu âm thanh được mã hóa và siêu dữ liệu liên quan đến dữ liệu âm thanh được mã hóa; bộ phân tích siêu dữ liệu để phân tích siêu dữ liệu để xác định khả năng thao tác dữ liệu âm thanh; giao diện tương tác để nhận đầu vào tương tác và để tạo ra, từ đầu vào tương tác, dữ liệu điều khiển tương tác liên quan đến khả năng thao tác dữ liệu âm thanh; và bộ tạo dòng dữ liệu để thu được dữ liệu điều khiển tương tác và dữ liệu âm thanh được mã hóa và siêu dữ liệu và để tạo ra dòng dữ liệu đầu ra, dòng dữ liệu đầu ra bao gồm dữ liệu âm thanh được mã hóa, ít nhất một phần siêu dữ liệu, và dữ liệu điều khiển tương tác như được xác định theo điểm 1. Các phương án ưu tiên khác được xác định trong các điểm yêu cầu bảo hộ phụ thuộc và các điểm yêu cầu bảo hộ độc lập được kèm theo.

Dữ liệu âm thanh được mã hóa có thể bao gồm các đối tượng âm thanh được mã hóa riêng, trong đó ít nhất một phần siêu dữ liệu liên quan đến đối tượng âm thanh tương ứng, trong đó bộ phân tích siêu dữ liệu được tạo cấu hình để phân tích phần tương ứng cho các đối tượng âm thanh được mã hóa để xác định, đối với ít nhất là đối tượng âm thanh, khả năng thao tác đối tượng, trong đó giao diện tương tác được tạo cấu hình để tạo ra, đối với ít nhất một đối tượng âm thanh được mã hóa, dữ liệu điều khiển tương tác từ đầu vào tương tác liên quan đến ít nhất một đối tượng âm thanh

được mã hóa. Do đó, các đối tượng âm thanh có thể được thao tác dễ dàng và trực tiếp trong các khả năng thao tác đối tượng tương ứng của chúng được lưu trữ trong siêu dữ liệu bằng cách sử dụng dữ liệu điều khiển tương tác tương ứng.

Giao diện tương tác có thể được tạo cấu hình để biểu diễn, tới người dùng, khả năng thao tác dữ liệu âm thanh được suy ra từ siêu dữ liệu bằng bộ phân tích siêu dữ liệu, và để nhận, từ người dùng, đầu vào người dùng trên thao tác dữ liệu cụ thể của khả năng thao tác dữ liệu. Điều này có thể thực hiện cách thực tế để cung cấp giao diện người dùng tới người dùng để tương tác với thiết bị theo sáng chế, ví dụ, để thao tác các đối tượng âm thanh, tốt hơn là ở bên ngoài từ bộ giải mã.

Bộ tạo dòng dữ liệu có thể được tạo cấu hình để xử lý dòng dữ liệu bao gồm dữ liệu âm thanh được mã hóa và siêu dữ liệu được nhận bởi giao diện bộ nhận mà không giải mã dữ liệu âm thanh được mã hóa, hoặc để sao chép dữ liệu âm thanh được mã hóa và ít nhất một phần siêu dữ liệu mà không làm thay đổi dòng dữ liệu đầu ra, trong đó bộ tạo dòng dữ liệu được tạo cấu hình để bổ sung phần dữ liệu bổ sung chứa dữ liệu điều khiển tương tác cho dữ liệu âm thanh được mã hóa và/hoặc siêu dữ liệu trong dòng dữ liệu đầu ra. Điều này cung cấp lợi ích là độ phức tạp giảm vì bộ xử lý dữ liệu âm thanh không cần giải mã các tín hiệu âm thanh. Bộ xử lý dữ liệu âm thanh chỉ cần phân tích siêu dữ liệu và ghi nó trở lại với phần siêu dữ liệu của dòng dữ liệu âm thanh được mã hóa.

Bộ tạo dòng dữ liệu có thể được tạo cấu hình để tạo ra, trong dòng dữ liệu đầu ra, dữ liệu điều khiển tương tác trong cùng định dạng như siêu dữ liệu. Do đó, bất kỳ dữ liệu điều khiển tương tác nào cũng có thể được tích hợp một cách có lợi vào trong dòng dữ liệu đầu ra.

Bộ tạo dòng dữ liệu có thể được tạo cấu hình để kết hợp, với dữ liệu điều khiển tương tác, bộ nhận biết trong dòng dữ liệu đầu ra, bộ nhận biết khác với bộ nhận biết được kết hợp với siêu dữ liệu. Lợi ích của việc sử dụng bộ nhận biết khác đối với siêu dữ liệu được thao tác là ở chỗ bộ giải mã từ xa có thể được cho phép để nhận biết sự tương tác từ dòng dữ liệu được thao tác được nhận trong khi cũng nhận dữ liệu gốc.

Bộ tạo dòng dữ liệu có thể được tạo cấu hình để bổ sung, vào dữ liệu điều khiển sự tương tác, dữ liệu chữ ký, dữ liệu chữ ký biểu thị thông tin trên ứng dụng, thiết bị hoặc người dùng thực hiện sự tương tác, ví dụ sự thao tác dữ liệu âm thanh hoặc cung cấp đầu vào người dùng. Bằng cách vận chuyển dữ liệu gốc và dữ liệu được thao tác,

việc thiết lập lại siêu dữ liệu là có khả năng. Chữ ký trong siêu dữ liệu cho phép theo dõi nguồn gốc của thao tác.

Bộ phân tích siêu dữ liệu có thể được tạo cấu hình để nhận biết khả năng vô hiệu hóa đối với một hoặc nhiều đối tượng âm thanh được biểu diễn bởi dữ liệu âm thanh được mã hóa, trong đó giao diện tương tác được tạo cấu hình để nhận thông tin vô hiệu hóa đối với một hoặc nhiều đối tượng âm thanh, và trong đó bộ tạo dòng dữ liệu được tạo cấu hình để đánh dấu một hoặc nhiều đối tượng âm thanh như bị vô hiệu hóa trong dữ liệu điều khiển tương tác hoặc để loại bỏ một hoặc nhiều đối tượng âm thanh bị vô hiệu hóa ra khỏi dữ liệu âm thanh được mã hóa sao cho dòng dữ liệu đầu ra không bao gồm dữ liệu âm thanh được mã hóa cho một hoặc nhiều đối tượng âm thanh bị vô hiệu hóa. Do đó, dòng dữ liệu có thể được làm tích ứng với các đối tượng âm thanh đó mà có khả năng thực sự hoặc có khả năng hiện thời sao cho nội dung dữ liệu tổng của dòng bit hiện thời có thể bị giảm.

Bộ tạo dòng dữ liệu có thể được tạo cấu hình để tạo ra theo cách động học dòng dữ liệu đầu ra, trong đó, đáp ứng với đầu vào tương tác mới, dữ liệu điều khiển tương tác được cập nhật để phù hợp với đầu vào tương tác mới, và trong đó bộ tạo dòng dữ liệu được tạo cấu hình để bao gồm dữ liệu điều khiển tương tác được cập nhật trong dòng dữ liệu đầu ra. Do đó, dòng dữ liệu có thể được gửi với thông tin thời gian thực. Nói cách khác, đầu vào tương tác liên quan đến các giá trị cụ thể của đối tượng âm thanh bất kỳ có thể được cập nhật và được xử lý theo cách nhanh chóng, tốt hơn là trong thời gian thực.

Giao diện bộ nhận có thể được tạo cấu hình để nhận dòng dữ liệu âm thanh chính bao gồm dữ liệu âm thanh được mã hóa và siêu dữ liệu liên quan đến dữ liệu âm thanh được mã hóa, và để nhận thêm dữ liệu âm thanh tùy chọn bao gồm đối tượng âm thanh tùy chọn, trong đó siêu dữ liệu liên quan đến đối tượng âm thanh tùy chọn đã nêu được chứa trong dòng dữ liệu âm thanh chính đã nêu. Với cấu hình này, bộ xử lý dữ liệu âm thanh có thể hợp nhất dữ liệu âm thanh được mã hóa của đối tượng âm thanh tùy chọn được lựa chọn vào trong dòng dữ liệu âm thanh chính dẫn đến dòng dữ liệu âm thanh đầu ra đầy đủ được tạo ra bởi bộ tạo dòng dữ liệu. Do đó, các đối tượng âm thanh tùy chọn có thể được cung cấp thêm cho người dùng sau đó hoặc theo yêu cầu.

Bộ phân tích siêu dữ liệu có thể được tạo cấu hình để xác định khả năng thao

tác âm thanh cho đối tượng âm thanh khuyết không được chứa trong dữ liệu âm thanh được mã hóa, trong đó giao diện tương tác được tạo cấu hình để nhận đầu vào tương tác cho đối tượng âm thanh khuyết, và trong đó giao diện bộ nhận được tạo cấu hình để yêu cầu dữ liệu âm thanh cho đối tượng âm thanh khuyết từ bộ cung cấp dữ liệu âm thanh hoặc để nhận dữ liệu âm thanh cho đối tượng âm thanh khuyết từ dòng con khác được chứa trong dòng phát rộng hoặc kết nối giao thức internet. Do đó, thiết bị hoặc người dùng có thể thao tác đối tượng âm thanh bổ sung có khả năng trước, tức là, trong khi nó khuyết thực sự. Đối tượng âm thanh bổ sung sau đó có thể được yêu cầu sau thông qua internet hoặc dòng phát rộng khác.

Bộ tạo dòng dữ liệu có thể được tạo cấu hình để gán, trong dòng dữ liệu đầu ra, loại gói nữa cho dữ liệu điều khiển tương tác, loại gói nữa này khác các loại gói cho dữ liệu âm thanh được mã hóa và siêu dữ liệu, hoặc trong đó bộ tạo dòng dữ liệu được tạo cấu hình để bổ sung, vào trong dòng dữ liệu đầu ra, dữ liệu điền đầy trong loại gói dữ liệu điền đầy, trong đó lượng dữ liệu điền đầy được xác định dựa trên yêu cầu tốc độ dữ liệu được xác định bởi giao diện đầu ra của bộ xử lý dữ liệu âm thanh. Do đó, chỉ một loại gói nữa cần được gán để lần lượt hoàn thành sự vận chuyển của siêu dữ liệu được thao tác hoặc dữ liệu điều khiển tương tác. Ngoài ra, bộ xử lý dữ liệu âm thanh có thể cần bổ sung dữ liệu điền đầy bổ sung vào dòng truyền dữ liệu sau để đáp ứng với cái đã cho, thường có sự yêu cầu tốc độ dữ liệu cao hơn cho đường liên kết này. Dữ liệu điền đầy này có thể không chứa thông tin và được kỳ vọng để được bỏ qua bởi bộ giải mã.

Bộ xử lý dữ liệu âm thanh có thể được thực hiện như thiết bị riêng biệt, trong đó giao diện bộ nhận có thể tạo thành đầu vào tới thiết bị riêng biệt thông qua kết nối có dây hoặc không dây, trong đó bộ xử lý dữ liệu âm thanh còn có thể bao gồm giao diện đầu ra được kết nối với bộ tạo dòng dữ liệu, giao diện đầu ra được tạo cấu hình để xuất ra dòng dữ liệu đầu ra, trong đó giao diện đầu ra thực hiện đầu ra của thiết bị và bao gồm giao diện không dây hoặc bộ kết nối có dây. Do đó, khả năng kết nối đơn giản, ví dụ trong mạng lưới, có thể được cung cấp.

Sáng chế còn có thể được thực hiện bằng phương pháp xử lý dữ liệu âm thanh, phương pháp bao gồm: nhận dữ liệu âm thanh được mã hóa và siêu dữ liệu liên quan đến dữ liệu âm thanh được mã hóa; phân tích siêu dữ liệu để xác định khả năng thao tác dữ liệu âm thanh; nhận đầu vào tương tác và tạo ra, từ đầu vào tương tác, dữ liệu

điều khiển tương tác liên quan đến khả năng thao tác dữ liệu âm thanh; và thu được dữ liệu điều khiển tương tác và dữ liệu âm thanh được mã hóa và siêu dữ liệu và tạo ra dòng dữ liệu đầu ra, dòng dữ liệu đầu ra bao gồm dữ liệu âm thanh được mã hóa, ít nhất một phần siêu dữ liệu, và dữ liệu điều khiển tương tác.

Sáng chế còn có thể được thực hiện bởi chương trình máy tính để thực hiện phương pháp xử lý dữ liệu âm thanh đã đề cập ở trên khi chạy trên máy tính hoặc bộ xử lý.

Sáng chế còn có thể được thực hiện bởi các phương án sau đây:

Khả năng thao tác dữ liệu âm thanh có thể được lựa chọn từ nhóm bao gồm ít nhất một trong số sự lựa chọn đối tượng, sự lựa chọn ngoài một vài ngôn ngữ, sự lựa chọn các đối tượng âm thanh bổ sung tùy chọn, thao tác đối tượng, sự thay đổi âm lượng của một hoặc nhiều đối tượng, sự thay đổi vị trí của các đối tượng như di chuyển lời thuyết minh bổ sung từ loa trung tâm đến loa bên phải hoặc vị trí tùy chọn ở giữa, sự lựa chọn các thiết lập trước, thay vì lựa chọn và thao tác từng đối tượng một cách riêng rẽ, trong đó sự thiết lập trước từ siêu dữ liệu được lựa chọn, trong đó sự thiết lập trước là sự lựa chọn trước các đối tượng được đề nghị bởi bộ tạo lập nội dung cho ứng dụng cụ thể hoặc cho kịch bản sử dụng cụ thể, trong đó sự thiết lập trước chứa tổ hợp của các đối tượng với ví dụ là các mức âm lượng, các vị trí, và độ to/dữ liệu nén dải động khác nhau được so sánh với sự biểu diễn mặc định.

Bộ tạo dòng dữ liệu có thể được tạo cấu hình để tạo ra dữ liệu điều khiển tương tác như thông tin độc lập hoặc như thông tin phụ thuộc, trong đó thông tin phụ thuộc là sự phụ thuộc vào siêu dữ liệu và các kết quả, nếu được áp dụng cho dữ liệu âm thanh được giải mã, cùng với siêu dữ liệu trong thao tác dữ liệu được xác định bởi đầu vào tương tác.

Dữ liệu âm thanh được mã hóa có thể bao gồm các đối tượng âm thanh tùy chọn và siêu dữ liệu có thể bao gồm siêu dữ liệu cho các đối tượng âm thanh tùy chọn, trong đó giao diện bộ nhận có thể được tạo cấu hình để nhận thêm dòng dữ liệu âm thanh chính có dữ liệu âm thanh chính, trong đó bộ tạo dòng dữ liệu có thể được tạo cấu hình để tạo ra dòng dữ liệu đầu ra để dòng dữ liệu đầu ra bao gồm thêm dữ liệu âm thanh chính.

Bộ tạo dòng dữ liệu có thể được tạo cấu hình để bổ sung dữ liệu bảo vệ lỗi cho dòng dữ liệu đầu ra và để gán loại gói nữa cho dữ liệu bảo vệ lỗi, trong đó bộ tạo dòng

dữ liệu được tạo cấu hình để suy ra dữ liệu bảo vệ lỗi từ dữ liệu âm thanh được mã hóa, siêu dữ liệu hoặc dữ liệu điều khiển tương tác.

Bộ tạo dòng dữ liệu có thể được tạo cấu hình để tạo ra dòng dữ liệu đầu ra như dòng dữ liệu để tạo dòng hoặc như tập tin chứa cơ sở trong định dạng tập tin như định dạng tập tin ISO MPEG-4.

Được đề xuất thêm rằng bộ xử lý dữ liệu âm thanh không có chức năng để giải mã dữ liệu âm thanh được mã hóa.

Bộ xử lý dữ liệu âm thanh có thể được thực hiện trong hộp chuyển đổi tín hiệu cáp tivi, thiết lập tivi hoặc bộ ghi-bộ nhận âm thanh/video.

Bộ xử lý dữ liệu âm thanh còn có thể bao gồm giao diện đầu ra để truyền dòng dữ liệu đầu ra đến thiết bị nữa thông qua sự kết nối HDMI.

Bộ xử lý dữ liệu âm thanh cũng có thể được cung cấp, tức là được tích hợp hoặc được thực hiện, cùng với bộ giải mã trong cùng thiết bị (phần cứng). Ví dụ, bộ xử lý dữ liệu âm thanh và bộ giải mã có thể cùng được cung cấp bên trong TV, hộp chuyển đổi tín hiệu cáp tivi, bộ nhận A/V, hoặc tương tự. Bộ xử lý dữ liệu âm thanh và bộ giải mã có thể kết nối thông qua các cấu trúc đường truyền chính dữ liệu bên trong. Cấu hình như vậy có thể được đặc biệt mong muốn trong các thiết bị TV bao gồm các phép giải hệ thống trên vi mạch (System-on-Chip - SoC).

Theo đó hoặc theo cách khác, bộ xử lý dữ liệu âm thanh có thể được thực hiện như thành phần chức năng độc lập và riêng rẽ trong cùng thiết bị giống với trường hợp được mô tả ở trên đối với trường hợp thiết bị riêng biệt, với chỉ sự khác biệt là giao diện đầu ra thực hiện đầu ra của bộ xử lý dữ liệu âm thanh trên sự kết nối bên trong thiết bị, ví dụ, sử dụng đường truyền chính dữ liệu bên trong.

Đối với các dấu hiệu được đề cập ở trên, bộ xử lý dữ liệu âm thanh theo sáng chế có khả năng cung cấp sự tương tác dễ dàng với thiết bị hoặc người dùng trong khi, tại cùng thời điểm, cung cấp thiết lập thiết bị đơn giản, tốt hơn là sử dụng các cài đặt hiện có.

Hơn nữa, bộ xử lý dữ liệu âm thanh theo sáng chế cung cấp giải pháp cho vấn đề được đề cập ở trên bằng cách lồng tương tác thiết bị hoặc tương tác người dùng như dữ liệu tương tác bổ sung bên trong dòng bit âm thanh. Bằng cách thực hiện các dấu hiệu được mô tả ở trên, việc thực hiện của bộ giải mã có thể chỉ yêu cầu một giao diện mà gồm cả dữ liệu biểu diễn được mã hóa và dữ liệu điều khiển tương tác. Các liên kết

hiện đã có có thể không cần thực hiện các kênh mới cho thông tin điều khiển, nhưng hiệu quả thực hiện tự được di chuyển vào trong bộ mã hóa-giải mã. Trong các thiết lập phức tạp, được đảm bảo thêm rằng thông tin điều khiển tương tác gắn chặt với bản chất được mã hóa và do đó không thể bị mất khi được nạp qua một vài giai đoạn xử lý.

Mô tả vắn tắt các hình vẽ

Các ví dụ của các phương án theo sáng chế được thể hiện trên các hình vẽ và sẽ được giải thích trong phần dưới đây, trong đó:

Fig.1 thể hiện bộ xử lý dữ liệu âm thanh theo sáng chế,

Fig.2 thể hiện phương pháp xử lý dữ liệu âm thanh theo sáng chế,

Fig.3 thể hiện ví dụ của dữ liệu âm thanh được mã hóa và siêu dữ liệu liên quan,

Fig.4 thể hiện ví dụ của việc xử lý dòng đầu vào và đầu ra,

Fig.5 thể hiện ví dụ nữa của việc xử lý dòng đầu vào và đầu ra,

Fig.6 thể hiện bộ xử lý dữ liệu âm thanh xử lý dữ liệu âm thanh tùy chọn,

Fig.7 thể hiện bộ xử lý dữ liệu âm thanh được thực hiện trong thiết bị riêng biệt,

Fig.8 thể hiện trường hợp sử dụng làm mẫu với thiết bị thứ nhất và thiết bị thứ hai, và

Fig.9 thể hiện kịch bản làm mẫu với hộp chuyển đổi tín hiệu cáp tivi và bộ nhận âm thanh-video.

Mô tả chi tiết sáng chế

Trong toàn bộ sáng chế này, và cụ thể là trong phần mô tả sau đây, thuật ngữ "sự tương tác" được sử dụng theo ý nghĩa là sự tương tác bởi người dùng hoặc sự tương tác bởi thiết bị, cũng như là sự tương tác nói chung, tức là sự tương tác theo ý nghĩa chung. Nói cách khác, "sự tương tác" có nghĩa là "sự tương tác người dùng" hoặc "sự tương tác thiết bị", hoặc sự tương tác nói chung. Trong các phần nhất định của bản mô tả, các thuật ngữ "người dùng" và "sự tương tác" được sử dụng cùng nghĩa. Ví dụ, giao diện người dùng có thể được sử dụng cùng nghĩa theo nghĩa giao diện người dùng và cách xung quanh khác.

Hơn nữa, "người dùng" có thể là người dùng là con người hoặc người dùng là máy, như thiết bị (phần cứng) hoặc thiết bị được thực hiện bởi phần mềm.

Hơn nữa, sự tương tác người dùng có thể có mặt như cấu hình thiết lập trước cụ thể cho thiết bị mà, một cách dành riêng hoặc ngoài đầu vào người dùng, có thể điều khiển sự thao tác dữ liệu.

Fig.1 thể hiện bộ xử lý dữ liệu âm thanh 1 theo sáng chế. Bộ xử lý dữ liệu âm thanh 1 bao gồm giao diện bộ nhận 2 để nhận dòng đầu vào được mã hóa 15 mà bao gồm dữ liệu âm thanh được mã hóa 3 và siêu dữ liệu 4. Siêu dữ liệu 4 liên quan đến dữ liệu âm thanh được mã hóa 3, mà sự liên quan được biểu thị bởi mũi tên 110. Ví dụ, dữ liệu âm thanh được mã hóa 3 có thể chứa các đối tượng âm thanh trong khi siêu dữ liệu 4 có thể chứa thêm thông tin về các khả năng thao tác của các đối tượng âm thanh đã nêu.

Bộ xử lý dữ liệu âm thanh 1 còn bao gồm bộ phân tích siêu dữ liệu 5 cho việc phân tích siêu dữ liệu 4 để xác định khả năng thao tác dữ liệu âm thanh. Ví dụ, mức âm lượng có thể điều chỉnh, vị trí không gian có thể điều chỉnh hoặc ngôn ngữ có thể lựa chọn có thể biểu diễn khả năng thao tác dữ liệu âm thanh của đối tượng âm thanh.

Hơn nữa, bộ xử lý dữ liệu âm thanh 1 bao gồm giao diện tương tác 6 cho việc nhận đầu vào tương tác 7. Giao diện tương tác 6 còn được tạo cấu hình để tạo ra dữ liệu điều khiển tương tác 8 dựa trên đầu vào tương tác 7. Dữ liệu điều khiển tương tác 8 đã nêu liên quan đến khả năng thao tác dữ liệu âm thanh được đề cập ở trên. Ví dụ, người dùng có thể tương tác với thiết bị bằng cách điều chỉnh mức âm lượng hoặc vị trí không gian của đối tượng âm thanh, hoặc bằng cách lựa chọn ngôn ngữ thông qua giao diện tương tác 6. Trong trường hợp này, giao diện tương tác 6 là giao diện người dùng 6 mà có thể tạo ra dữ liệu điều khiển người dùng tương ứng 8 mà liên quan đến sự lựa chọn của người dùng.

Ngoài ra hoặc theo cách khác, giao diện tương tác 6 có thể là giao diện tương tác thiết bị (cụ thể) 6. Trong trường hợp này, giao diện tương tác thiết bị 6 được tạo cấu hình để tạo ra dữ liệu điều khiển tương tác thiết bị 8 dựa trên đầu vào tương tác thiết bị 7. Ví dụ, thiết bị như tai nghe chụp đầu hoặc tương tự có thể được kết nối với giao diện tương tác 6. Sự kết nối giữa tai nghe chụp đầu và giao diện tương tác 6 có thể được phát hiện bởi bộ xử lý âm thanh và do đó được xem như đầu vào tương tác 7. Do đó, nhờ vào sự kết nối của tai nghe chụp đầu, giao diện tương tác 6 cung cấp dữ liệu điều khiển tương tác tai nghe chụp đầu cụ thể 8, như thao tác đối tượng âm thanh, ví dụ sự giảm âm lượng tự động, ngôn ngữ được lựa chọn trước hoặc sự điều chỉnh cấu hình phần cứng.

Nói cách khác, thay vì sự tương tác người dùng bằng tay, giao diện tương tác 6 lựa chọn tự động các đối tượng hoặc điều chỉnh dựa trên sự phát hiện các thiết bị nhất

định. Giao diện tương tác 6 tạo ra dữ liệu điều khiển tương tác thiết bị cụ thể 8.

Bộ xử lý dữ liệu âm thanh 1 còn bao gồm bộ tạo dòng dữ liệu 9. Bộ tạo dòng dữ liệu 9 thu được dữ liệu điều khiển tương tác 8, dữ liệu âm thanh được mã hóa 3 và siêu dữ liệu 4. Bộ tạo dòng dữ liệu 9 được tạo cấu hình để tạo ra dòng dữ liệu đầu ra 10 mà bao gồm dữ liệu điều khiển tương tác 8 được đề cập ở trên, dữ liệu âm thanh được mã hóa 3 và siêu dữ liệu 4.

Fig.2 thể hiện phương pháp tương ứng để xử lý dữ liệu âm thanh theo sáng chế.

Trong bước 201, dữ liệu âm thanh được mã hóa 3 và siêu dữ liệu liên quan 4 được nhận.

Trong bước 202, siêu dữ liệu 4 được phân tích để xác định khả năng thao tác dữ liệu âm thanh.

Trong bước 203, đầu vào tương tác được nhận, trong đó dữ liệu điều khiển tương tác liên quan đến khả năng thao tác dữ liệu âm thanh được tạo ra từ đầu vào tương tác đã nêu trong bước 204.

Trong bước 205, dữ liệu điều khiển tương tác và dữ liệu âm thanh được mã hóa và siêu dữ liệu thu được, và dòng dữ liệu đầu ra được tạo ra, trong đó dòng dữ liệu đầu ra đã nêu bao gồm dữ liệu âm thanh được mã hóa, ít nhất một phần siêu dữ liệu và dữ liệu điều khiển tương tác.

Với sự tham chiếu đến Fig.3, dữ liệu âm thanh được mã hóa 3 bao gồm các đối tượng âm thanh được mã hóa riêng 11, 12. Hơn nữa, ít nhất một phần 13, 14 của siêu dữ liệu 4 liên quan (được biểu thị bởi các mũi tên 110, 120) đến đối tượng âm thanh tương ứng 11, 12. Ví dụ, phần 'MD1' 13 của siêu dữ liệu 4 liên quan đến đối tượng âm thanh được mã hóa tương ứng 'AO1' 11, trong khi phần 'MD2' 14 của siêu dữ liệu 4 liên quan đến đối tượng âm thanh được mã hóa tương ứng 'AO2' 12.

Bộ phân tích siêu dữ liệu 5 được tạo cấu hình để phân tích phần tương ứng 13, 14 cho các đối tượng âm thanh được mã hóa 11, 12 để xác định khả năng thao tác đối tượng của ít nhất một trong số các đối tượng âm thanh đã nêu 11, 12. Nói cách khác, bộ phân tích siêu dữ liệu 5 phân tích siêu dữ liệu 13, 14 cho các đối tượng âm thanh tương ứng 11, 12 để xác định khả năng thao tác đối tượng âm thanh cho từng đối tượng âm thanh 11, 12. Ví dụ, bộ phân tích siêu dữ liệu 5 xác định rằng đối tượng âm thanh 'AO1' 11 có thể bao gồm mức âm lượng có thể điều chỉnh được. Bộ phân tích siêu dữ liệu 5 có thể có tính điều chỉnh được mức âm lượng tiềm năng (khả năng thao

tác đối tượng âm thanh) cho người dùng thông qua giao diện người dùng 6.

Giao diện người dùng 6 được tạo cấu hình để tạo ra, đối với ít nhất một đối tượng âm thanh 11, 12, dữ liệu điều khiển người dùng 8 từ đầu vào người dùng 7 liên quan đến ít nhất một đối tượng âm thanh được mã hóa 11, 12. Ví dụ, người dùng có thể muốn điều chỉnh mức âm lượng của đối tượng âm thanh 'AO1' 11 và do đó cung cấp đầu vào tương ứng 7 thông qua giao diện người dùng 6. Giao diện người dùng 6 tạo ra dữ liệu điều khiển người dùng tương ứng 8 chứa thông tin mà, và tới mức độ nào, người dùng muốn điều chỉnh mức âm lượng của đối tượng âm thanh 'AO1' 11.

Theo đó, giao diện người dùng 6 được tạo cấu hình để biểu diễn cho người dùng khả năng thao tác đối tượng âm thanh của đối tượng âm thanh 11, 12 được suy ra từ siêu dữ liệu 4 bằng bộ phân tích siêu dữ liệu 5. Giao diện người dùng 6 còn được tạo cấu hình để nhận đầu vào người dùng 7 từ người dùng về thao tác dữ liệu cụ thể (ví dụ mức âm lượng hoặc ngôn ngữ cụ thể) của khả năng thao tác dữ liệu (ví dụ khoảng điều chỉnh mức âm lượng hoặc sự thiết lập các ngôn ngữ có thể dùng được).

Với sự tham chiếu đến Fig.1 và Fig.9, bộ tạo dòng dữ liệu 9 được tạo cấu hình để xử lý dòng dữ liệu 15 bao gồm dữ liệu âm thanh được mã hóa 3 và siêu dữ liệu 4 được nhận bởi giao diện bộ nhận 2 mà không giải mã dữ liệu âm thanh được mã hóa 3. Ví dụ, được giả định rằng bộ xử lý dữ liệu âm thanh 1 theo sáng chế được thực hiện trong hộp chuyển đổi tín hiệu cáp tivi 19, 26, nó có thể chuyển tiếp dòng dữ liệu đầu ra 10, 32 đến bộ nhận âm thanh-video bên ngoài 28, 33 mà bao gồm bộ giải mã. Trong trường hợp này, dòng dữ liệu đầu ra 10, 32 có thể vẫn được mã hóa vì việc giải mã sẽ không được thực hiện bởi hộp chuyển đổi tín hiệu cáp tivi 19, 26 mà bởi bộ nhận âm thanh-video 28, 33.

Theo cách khác, bộ tạo dòng dữ liệu 9 được tạo cấu hình để sao chép dữ liệu âm thanh được mã hóa 3 và siêu dữ liệu 4 mà không làm thay đổi trong dòng dữ liệu đầu ra 10.

Trong cả hai trường hợp, bộ tạo dòng dữ liệu 9 được tạo cấu hình để bổ sung phần dữ liệu bổ sung chứa dữ liệu điều khiển tương tác 8 cho dữ liệu âm thanh được mã hóa 3 và/hoặc siêu dữ liệu 4 trong dòng dữ liệu đầu ra 10, như có thể được thấy trên Fig.4.

Với sự tham chiếu đến Fig.5, bộ tạo dòng dữ liệu 9 còn được tạo cấu hình để hợp nhất hai dòng đầu vào 15a, 15b thành dòng đầu ra chung 10, trong đó phần dữ liệu

bổ sung chứa dữ liệu điều khiển tương tác 8 được bổ sung vào dữ liệu âm thanh được mã hóa 3 và/hoặc siêu dữ liệu 4 trong dòng dữ liệu đầu ra 10.

Tốt hơn là, bộ tạo dòng dữ liệu 9 được tạo cấu hình để tạo ra, trong dòng dữ liệu đầu ra 10, dữ liệu điều khiển tương tác 8 trong cùng định dạng như siêu dữ liệu 4. Do đó, dữ liệu điều khiển tương tác 8 có thể được tổ hợp dễ dàng với siêu dữ liệu có thể dùng được 4.

Như được đề cập ở trên, nếu bộ tạo dòng dữ liệu 9 sao chép siêu dữ liệu 4, thì siêu dữ liệu gốc 4 có thể còn lại trong dòng đầu ra 10 ngoài siêu dữ liệu được thao tác bất kỳ chứa dữ liệu điều khiển tương tác bổ sung 8 cho các đối tượng âm thanh tương ứng 11, 12. Cả siêu dữ liệu gốc và siêu dữ liệu được thao tác có thể được gửi đến bộ giải mã 28, 33 để cho phép bộ giải mã 28, 33 để hoặc là nhận biết các chênh lệch là kết quả của sự tương tác (người dùng) và để có được tất cả thông tin về các giá trị mặc định như được dự định bởi bộ tạo nội dung hoặc là để tính toán kết quả của sự tương tác (người dùng) từ siêu dữ liệu gốc 4 và siêu dữ liệu được thao tác 4' (hoặc dữ liệu điều khiển tương tác 8).

Hơn nữa, với sự tham chiếu đến Fig.1, bộ tạo dòng dữ liệu 9 được tạo cấu hình để tạo ra theo cách động học dòng dữ liệu đầu ra 10. Mỗi lần người dùng hoặc thiết bị cung cấp đầu vào tương tác mới 7 cho giao diện tương tác 6, dữ liệu điều khiển tương tác 8 được cập nhật theo để phù hợp với đầu vào tương tác mới 7 đã nêu. Bộ tạo dòng dữ liệu 9 bao gồm dữ liệu điều khiển tương tác được cập nhật 8 trong dòng dữ liệu đầu ra 10.

Fig.6 thể hiện bộ xử lý dữ liệu âm thanh 1 theo sáng chế, trong đó thông tin âm thanh tùy chọn được xử lý. Như có thể được thấy, dòng dữ liệu đầu vào 15 là dòng dữ liệu âm thanh chính bao gồm dữ liệu âm thanh được mã hóa 3 và siêu dữ liệu liên quan 4. Hơn nữa, giao diện bộ nhận 2 nhận thêm dữ liệu âm thanh tùy chọn 16 bao gồm đối tượng âm thanh tùy chọn 17.

Tuy nhiên, siêu dữ liệu liên quan đến đối tượng âm thanh tùy chọn bổ sung 'AOx' 17 đã nêu, tức là thông tin về các khả năng thao tác của đối tượng âm thanh tùy chọn 'AOx' 17 đã nêu, được chứa trong dòng dữ liệu âm thanh chính 15. Do đó, đối tượng âm thanh 17 được biết đến nhưng không có mặt và do đó tùy chọn.

Ví dụ, người dùng nghe dàn nhạc chứa các trống, các dây đàn và piano. Các nhạc khí thổi có thể được bao gồm một cách tùy chọn. Nếu người nghe muốn thêm

nhạc khí thổi, họ có thể làm như vậy bằng cách bổ sung nhạc khí thổi tùy chọn, ví dụ kèn, như đối tượng âm thanh tùy chọn 17. Vì các khả năng thao tác đối với các kèn đã nêu đã được chứa trong siêu dữ liệu 4 của dòng dữ liệu âm thanh chính 15, người dùng có khả năng thao tác kèn đã được bổ sung hiện thời theo mong muốn của họ.

Hơn nữa, với sự tham chiếu đến Fig.6, đối tượng âm thanh bổ sung đã nêu 'AOx' 17 có thể là đối tượng âm thanh khuyết mà không được chứa trong dữ liệu âm thanh được mã hóa 3 và/hoặc siêu dữ liệu 4. Do đó, đối tượng âm thanh 17 không được biết đến và do đó bị khuyết.

Trong trường hợp này, giao diện bộ nhận 2 được tạo cấu hình để yêu cầu dữ liệu âm thanh 16 thuộc về đối tượng âm thanh khuyết 17 đã nêu từ bộ cung cấp dữ liệu âm thanh 35. Giao diện bộ nhận 2 cũng được tạo cấu hình để nhận dữ liệu âm thanh đã nêu 16 từ dòng con khác được chứa trong dòng phát rộng 36. Giao diện bộ nhận 2 còn được tạo cấu hình để tìm kiếm dữ liệu âm thanh 16 đã nêu từ internet 37 thông qua kết nối giao thức internet.

Ví dụ, người dùng xem phim có thể chọn ngôn ngữ cụ thể từ thiết lập ngôn ngữ có sẵn chứa, ví dụ, tiếng Anh, tiếng Đức và tiếng Pháp. Ngôn ngữ thứ tư được biết đến nhưng không có mặt và do đó bị khuyết. Tuy nhiên, ngôn ngữ thứ tư có thể được cung cấp sau đó ví dụ thông qua internet.

Lại đề cập đến Fig.4 và Fig.5, dòng dữ liệu đầu vào 15 và dòng dữ liệu đầu ra 10 thường có thể có sẵn trong cấu trúc được gói hóa. Ví dụ, sự vận chuyển âm thanh MPEG-H trên chuỗi các giao diện được xác định bởi cú pháp vận chuyển MHAS (MPEG-H Audio Stream - MHAS) (xem đoạn 13 của N14459 (ISO/IEC 23008-3 Committee Draft Text) [1]). Cú pháp này được xác định theo cách được gói hóa.

Do đó, để hoàn thành sự vận chuyển của siêu dữ liệu được thao tác 4 hoặc dữ liệu điều khiển tương tác 8, chỉ một loại gói nữa cần được gán cho thông tin điều khiển mới.

Ngoài ra, thiết bị thứ nhất 'thiết bị#1' 19 bao gồm bộ xử lý dữ liệu âm thanh 1 có thể cần bổ sung dữ liệu điền đầy bổ sung 18 vào trong dòng truyền sau 10 để đáp ứng với cái đã cho, thường là yêu cầu tốc độ dữ liệu cao hơn nhiều đối với liên kết đó. Dữ liệu điền đầy 18 này có thể không chứa thông tin và được mong đợi để được bỏ qua bởi thiết bị thứ hai mà nhận dòng đầu ra được thao tác 10. Để hoàn thành điều này, loại gói dữ liệu nữa có thể được gán.

Hơn nữa, vì lớp vận chuyển của các dòng vào đến ‘thiết bị#1’ 19 có thể có sự bảo vệ lỗi của riêng chúng, nhưng liên kết ra không cung cấp lớp bảo vệ như vậy, nên thiết bị#1 có thể bổ sung các gói dữ liệu chứa dữ liệu kiểm tra tính chẵn lẻ. Chúng có thể được bổ sung vào các dòng MHAS như loại gói dữ liệu bổ sung.

Hơn nữa, vì lớp vận chuyển có thể truyền đạt dữ liệu bổ sung như thông tin phụ, các gói dữ liệu này cũng có thể được gói trong dòng âm thanh MHAS như loại gói khác. Ví dụ cho dữ liệu này là các bộ mô tả trong dòng vận chuyển và chương trình MPEG-2.

Ví dụ khác để lưu trữ dữ liệu âm thanh được mã hóa là định dạng tệp tin ISO mp4. Tương tự với định dạng dòng, cũng có khả năng đối với trường hợp định dạng tệp tin để đọc, thao tác và ghi ngược lại siêu dữ liệu đến tệp tin hoặc để lưu trữ dữ liệu điều khiển người dùng ngoài siêu dữ liệu gốc mà không thay đổi dữ liệu tín hiệu âm thanh được mã hóa.

Tài liệu tham khảo

- [1] ISO N14459 (ISO/IEC 23008-3 Committee Draft Text)
- [2] IEC 60958-3: “Digital audio interface – Part 3: Consumer applications”
- [3] IEC 61937-11, “Digital audio – Interface for non-linear PCM encoded audio bitstreams applying IEC 60958 – Part 11: MPEG-4 AAC and its extensions in LATM/LOAS”

Ngay sau đây đề cập đến Fig.7, bộ xử lý dữ liệu âm thanh 1 được thực hiện như thiết bị riêng biệt ‘thiết bị#1’ 19. Trong thiết bị riêng biệt 19 đã nêu, giao diện bộ nhận 2 tạo thành đầu vào 20 đến thiết bị riêng biệt 19 thông qua kết nối có dây 21 hoặc kết nối không dây 22.

Bộ xử lý dữ liệu âm thanh 1 còn bao gồm giao diện đầu ra 23 mà được kết nối với bộ tạo dòng dữ liệu 9 và cung cấp đầu ra cho thiết bị 19. Hơn nữa, giao diện đầu ra 23 được tạo cấu hình để xuất ra dòng dữ liệu đầu ra 10 thông qua giao diện không dây 24 hoặc bộ kết nối có dây 25.

Các cách khác của việc thực hiện sáng chế có thể được mô tả bằng cách ví dụ trong kịch bản mà trong đó hai thiết bị riêng biệt có sẵn. Thiết bị thứ nhất ‘thiết bị #1’ bao gồm bộ xử lý dữ liệu âm thanh theo sáng chế. Thiết bị thứ hai ‘thiết bị #2’ nhận dữ liệu âm thanh được xử lý, nhưng vẫn được mã hóa từ ‘thiết bị#1’ để giải mã dữ liệu âm thanh đã nêu.

Như có thể được thấy trên Fig.8 và Fig.9, thiết bị thứ nhất 19, 26 nhận dòng đầu vào 15 bao gồm dữ liệu âm thanh được mã hóa 3 và siêu dữ liệu liên quan 4. Thiết bị thứ nhất 19, 26 đọc thông tin siêu dữ liệu 4 từ dòng dữ liệu âm thanh vào 15 hoặc dòng vận chuyển và dẫn dữ liệu tín hiệu âm thanh 3 chưa đề cập đến. Thiết bị thứ nhất 19, 26 phân tích siêu dữ liệu 4 và biểu diễn thông tin về các đối tượng tới ứng dụng ví dụ trong giao diện tương tác 6, bao gồm các giới hạn trên thao tác đối tượng mà là phần của siêu dữ liệu 4. Từ ứng dụng hoặc giao diện tương tác 6, người dùng có thể lựa chọn và thao tác các đối tượng để làm thích ứng sự biểu diễn âm thanh cho sở thích riêng của họ:

- Lựa chọn đối tượng: ví dụ lựa chọn một ngôn ngữ ngoài một vài ngôn ngữ, lựa chọn các đối tượng âm thanh bổ sung tùy chọn. v.v..
- Thao tác đối tượng: ví dụ thay đổi âm lượng của các đối tượng, thay đổi vị trí của các đối tượng như di chuyển lời thuyết minh bổ sung từ loa trung tâm đến loa bên phải hoặc vị trí tùy ý ở giữa.
- Lựa chọn các thiết lập trước: thay vì lựa chọn và thao tác từng đối tượng một cách riêng rẽ, người dùng cũng có thể lựa chọn thiết lập trước từ siêu dữ liệu. Thiết lập trước là lựa chọn trước của đối tượng được đề nghị bởi bộ tạo nội dung cho các ứng dụng cụ thể hoặc các kịch bản sử dụng. Thiết lập trước có thể chứa tổ hợp các đối tượng với, ví dụ các mức âm lượng khác nhau, các vị trí và độ to/dữ liệu nén dải động được so sánh với sự biểu diễn mặc định.

Trong bước tiếp theo, thiết bị thứ nhất 19, 26 lưu trữ thông tin về tính tương tác người dùng (dữ liệu điều khiển tương tác 8) vào trong dòng dữ liệu âm thanh được mã hóa 10, 32. Thiết bị thứ nhất 19, 26 có thể ghi các giá trị được thay đổi hoặc lượng thao tác, ví dụ, các giá trị độ lệch và thừa số nhân trở lại phần siêu dữ liệu 4 hoặc phần được dành riêng của dòng dữ liệu âm thanh được mã hóa sao cho đầu ra của thiết bị thứ nhất 19, 26 lại là dòng âm thanh được mã hóa có hiệu quả 10, 32.

Thiết bị thứ nhất 19, 26 có thể sử dụng bộ nhận biết, nhãn hoặc loại gói khác để lần lượt gói gọn siêu dữ liệu được thao tác hoặc dữ liệu điều khiển người dùng 8. Siêu dữ liệu gốc 4 có thể còn lại trong dòng đầu ra 10, 32 ngoài siêu dữ liệu được thao tác. Bộ nhận biết, nhãn hoặc loại gói khác lần lượt được sử dụng cho siêu dữ liệu được thao tác hoặc dữ liệu điều khiển tương tác 8 để cho phép thiết bị thứ hai 28, 33 nhận biết nếu siêu dữ liệu 4 đã được thao tác trước đó. Siêu dữ liệu gốc 4 còn lại trong dòng

10, 32 để cho phép thiết bị thứ hai 28, 33 để hoặc là nhận biết các chênh lệch là kết quả của sự tương tác người dùng hoặc thiết bị và để có được tất cả thông tin về các giá trị mặc định như được dự định bởi bộ tạo nội dung, hoặc là để tính toán kết quả của sự tương tác người dùng hoặc thiết bị từ siêu dữ liệu gốc 4 và siêu dữ liệu được thao tác (hoặc dữ liệu điều khiển người dùng 8).

Như phần của siêu dữ liệu được thao tác, dữ liệu chữ ký có thể được lồng trong siêu dữ liệu 4. Chữ ký có thể chứa thông tin về ứng dụng, thiết bị hoặc người dùng, mà đã thao tác siêu dữ liệu 4.

Các đối tượng âm thanh mà không được lựa chọn bởi người dùng có thể hoặc là được đánh dấu là bị vô hiệu hóa trong siêu dữ liệu 4 hoặc dữ liệu điều khiển tương tác 8, hoặc theo cách khác là phần âm thanh được mã hóa của các đối tượng âm thanh này có thể bị loại bỏ khỏi dòng âm thanh 10.

Quy trình của tính tương tác người dùng hoặc thiết bị có thể là động, tức là mọi lúc người dùng hoặc thiết bị thay đổi các thiết lập cho việc lựa chọn và thao tác, thiết bị thứ nhất 19, 26 ghi các giá trị được thay đổi này trở lại phần siêu dữ liệu 4 của dòng dữ liệu âm thanh được mã hóa 10.

Cũng có khả năng là thiết bị thứ hai 28, 33 thao tác thêm siêu dữ liệu 4 hoặc là do quy trình được tự động hóa (ví dụ để làm thích ứng cảnh âm thanh với trạng thái người nghe) hoặc là giao diện tương tác bổ sung. Trong trường hợp này, thiết bị thứ hai 28, 33 có thể ghi các giá trị được thao tác trở lại dòng dữ liệu âm thanh được mã hóa 10, ví dụ, ghi đè lên các giá trị được ghi bởi thiết bị thứ nhất 19, 26.

Tham chiếu lại Fig.6, cách nữa để thực hiện sáng chế được mô tả bằng cách ví dụ trong cái được gọi là Use case lai (Hybrid Use Case).

Dữ liệu âm thanh được mã hóa của các đối tượng âm thanh tùy chọn được lựa chọn 17 có thể không phải là phần của dòng dữ liệu âm thanh chính 15, nhưng có thể được phân phối sử dụng các kênh vận chuyển khác. Ví dụ, dòng dữ liệu âm thanh chính 15 được phân phối trên kênh phát rộng, trong khi dữ liệu âm thanh được mã hóa 3 của các đối tượng âm thanh tùy chọn 17 được phân phối theo yêu cầu trên sự kết nối IP 37.

Siêu dữ liệu đầy đủ 4 cho tất cả các đối tượng được chứa trong dòng dữ liệu âm thanh chính 15 sao cho tất cả thông tin về sự tương tác và sự lựa chọn đối tượng có sẵn trong thiết bị thứ nhất 19. Do đó quy trình tương tác và lưu trữ của dữ liệu điều khiển

tương tác 8 cho dòng 10 giống với trường hợp được mô tả ở trên.

Nếu người dùng lựa chọn đối tượng 17 và dữ liệu âm thanh được mã hóa 3 không phải là phần của dòng dữ liệu âm thanh chính 15, thiết bị thứ nhất 19 có thể nhận dữ liệu âm thanh được mã hóa 3 của đối tượng 17 này trên sự kết nối dữ liệu khác hơn là dòng dữ liệu âm thanh chính 15, ví dụ dòng con 36 khác trong dòng phát rộng hoặc kết nối IP 37.

Trong bước tiếp theo, thiết bị thứ nhất 19 hợp nhất dữ liệu âm thanh được mã hóa 3 của đối tượng được lựa chọn 17 thành dòng âm thanh chính 15 thu được dòng dữ liệu âm thanh đầy đủ 10 để phân phối tiếp đến thiết bị thứ hai 33.

Hơn nữa, khả năng được đề xuất để vận chuyển dữ liệu tương tác 8 được lồng trong dòng bit tương hợp MPEG-H 10.

Như có thể được thấy trên Fig.9, các hệ thống đa phương tiện tiên tiến thường không phải tích hợp tất cả chức năng mong muốn trong một thiết bị đơn, mà thực hiện các thành phần chức năng khác trong các thiết bị chuyên dụng ví dụ như các hộp chuyển đổi tín hiệu cáp tivi 26, thiết lập TV 27 hoặc bộ nhận AVR 28. Các thiết bị này giao tiếp thông qua các giao diện được chuẩn hóa như HDMI.

Tuy nhiên, có khả năng là ít nhất bộ xử lý dữ liệu âm thanh và bộ giải mã được tích hợp trong một thiết bị đơn. Bộ xử lý dữ liệu âm thanh có thể được cung cấp, tức là được tích hợp hoặc được thực hiện, cùng với bộ giải mã trong cùng thiết bị (phần cứng). Ví dụ, bộ xử lý dữ liệu âm thanh và bộ giải mã có thể cùng được cung cấp bên trong TV, hộp chuyển đổi tín hiệu cáp tivi, bộ nhận A/V, hoặc tương tự. Bộ xử lý dữ liệu âm thanh và bộ giải mã có thể kết nối thông qua các cấu trúc đường truyền chính dữ liệu bên trong. Cấu hình như vậy có thể được đặc biệt mong muốn trong các thiết bị TV bao gồm các phép giải hệ thống trên vi mạch (System-on-Chip - SoC).

Theo đó hoặc theo cách khác, bộ xử lý dữ liệu âm thanh có thể được thực hiện như thành phần chức năng độc lập và riêng biệt trong cùng thiết bị giống với trường hợp được mô tả ở trên đối với trường hợp hai thiết bị riêng biệt, với chỉ sự khác biệt là giao diện đầu ra thực hiện đầu ra của bộ xử lý dữ liệu âm thanh trên sự kết nối bên trong tới thiết bị, ví dụ, sử dụng đường truyền chính dữ liệu bên trong.

Một trường hợp sử dụng đối với sự tái tạo nội dung MPEG-H với nhiều thiết bị phức tạp là trường hợp, khi chương trình tivi được nhận bởi thiết bị thứ nhất 19 như hộp chuyển đổi dữ liệu tivi (set-top box - STB), mà lựa chọn kênh truyền thích hợp và

trích các dòng cơ sở liên quan chứa bản chất được tạo mã mong muốn. Sự tương tác, tức là sự điều khiển người dùng đối với sự lựa chọn phần tử âm thanh và sự tương tác/thao tác cũng thường được thực hiện ở đây.

Bộ giải mã MPEG-H 31 có thể không được định vị trong STB 26, mà thay vào đó sẽ được định vị trong bộ nhận âm thanh-video (Audio-Video-Receiver - AVR) 28. Trường hợp sử dụng này được mô tả trên Fig.9.

Trong trường hợp này, các dòng được trích cần được cấp vào AVR 28 để tái tạo; việc truyền giữa hai thiết bị 26, 28 này có thể được hoàn thành bằng cách hoặc là truyền sự biểu diễn được giải mã/giải nén (PCM với âm thanh), hoặc, đặc biệt là nếu các giới hạn băng thông áp dụng vào đường liên kết được sử dụng, trong sự biểu diễn được mã hóa.

AVR 28 sau đó chỉ cung cấp giao diện cấu hình mà thường được truy cập chỉ một lần bởi người dùng khi thiết lập hệ thống và tác động theo "chế độ phụ thuộc" tại những lần hoạt động thông thường.

Vì sự tương tác xảy ra trong STB (thiết bị #1) 26 và việc giải mã và kết xuất được thực hiện trong AVR (thiết bị #2) 28, điều đó là không thể tránh khỏi để có thể gửi thông tin tính tương tác người dùng từ STB 26 đến AVR 28.

Để hoàn thành trường hợp sử dụng được mô tả và khắc phục các hạn chế được mô tả, sáng chế được đề xuất để lồng dữ liệu thông tin tương tác 8 vào trong dòng dữ liệu âm thanh được mã hóa 10 như có thể được thấy trên Fig.8.

Thiết bị thứ nhất 19 đọc thông tin siêu dữ liệu 4 từ mpeg3daConfig() của dòng dữ liệu âm thanh vào 15 hoặc qua sự báo hiệu ngoài dải như các bộ mô tả MPEG-2 TS. Sau đó nó phân tích siêu dữ liệu phần tử âm thanh 4 và biểu diễn thông tin về các đối tượng trong giao diện tương tác 6, bao gồm các giới hạn trên thao tác đối tượng mà là phần của siêu dữ liệu 4. Từ giao diện tương tác 6, người dùng có thể lựa chọn và thao tác phần tử âm thanh để làm thích ứng sự biểu diễn âm thanh cho sở thích riêng của họ. "Dữ liệu tương tác người dùng" mô tả sự lựa chọn và sự thao tác đối tượng.

Trong bước tiếp theo, thiết bị thứ nhất 19 ghi dữ liệu này thành dòng dữ liệu âm thanh MHAS 32 sử dụng MHASPacketType mới. Đầu ra của thiết bị thứ nhất 19 lại là dòng âm thanh được mã hóa có hiệu quả 10, 32. Siêu dữ liệu gốc 4 trong mpeg3daConfig() và dữ liệu tín hiệu âm thanh được mã hóa 3 không được biến đổi.

Sự có mặt của gói MHASPacketType này cho phép thiết bị thứ hai 28, 33 nhận

biết rằng sự tương tác (người dùng) đã xảy ra. Siêu dữ liệu gốc 4 còn lại trong dòng 10, 32 để cho phép thiết bị thứ hai 10, 33 để có tất cả thông tin về các giá trị mặc định như được dự định bởi bộ tạo nội dung.

Các phần tử âm thanh mà không được lựa chọn bởi người dùng có thể hoặc là được đánh dấu là bị vô hiệu hóa trong siêu dữ liệu 4, hoặc theo cách khác phần âm thanh được mã hóa 3 của các đối tượng có thể được loại bỏ khỏi dòng âm thanh 10.

Quy trình tương tác, tức là của tính tương tác thiết bị hoặc người dùng có thể là động, tức là mọi lúc người dùng hoặc thiết bị thay đổi các thiết lập cho việc lựa chọn và thao tác, thiết bị thứ nhất 19, 26 ghi các giá trị được thay đổi này trở lại phần cụ thể của dòng dữ liệu âm thanh được mã hóa 10.

Mở rộng MHAS

Sự vận chuyển âm thanh MPEG-H trên chuỗi các giao diện được xác định bởi cú pháp vận chuyển MHAS (xem đoạn 13 của N14459 (ISO/IEC 23008-3 Committee Draft Text) [1]). Cú pháp này được xác định theo cách được gói hóa. Do đó, để hoàn thành việc vận chuyển của dữ liệu tương tác người dùng, chỉ một loại được gói nữa cần được gán cho thông tin điều khiển mới.

Bảng 1 — Cú pháp của MHASPacketPayload()

Cú pháp	Số lượng các bit	Gợi ý
<pre> MHASPacketPayload(MHASPacketType) { swich (MHASPacketType) { case PACTYP_SYNC: 0xA5; /* syncword*/ break; case PACTYP_MPEGH3DACFG: mpeg3daConfig(); break; case PACTYP_MPEGH3DAFRAME: mpeg3daFrame(); break; case PACTYP_SYNCGAP: syncSpacingLength = escapedValue(16,24,24); break; case PACTYP_MARKER: for (i=0; i< MHASPacketLength; i++) { marker_byte(i); } break; case PACTYP_USERINTERACTION: mpeg3daElementInteraction(); break; } ByteAlign(); } </pre>	8 16,40,64 8	uimbsf uimbsf

Bảng 2 — Giá trị của MHASPacketType

MHASPacketType	Giá trị
PACTYP_SYNC	0
PACTYP_MPEGH3DACFG	1
PACTYP_MPEGH3DAFRAME	2
/* reserved for ISO use */	3-6
PACTYP_SYNCGAP	7
PACTYPE_MARKER	8
PACTYP_USERINTERACTION	12
/* reserved for ISO use */	13-127
/* reserved for use outside of ISO scope */	128-261
/* reserved for ISO use */	262-389
/* reserved for use outside of ISO scope */	390-517
Lưu ý: Các giá trị MHASPacketType ứng dụng cụ thể được chỉ thị để nằm trong không gian được dự trữ cho việc sử dụng bên ngoài của phạm vi ISO. Những việc này được bỏ qua bởi bộ giải mã như cực tiểu của cấu trúc được yêu cầu bởi bộ giải mã để bỏ qua các mở rộng này.	

PACTYP_USERINTERACTION

MHASPacketType PACTYP_USERINTERACTION có thể được sử dụng để nạp dữ liệu tương tác phần tử vào bộ giải mã.

Đối với loại gói này, MHASPacketLabel có cùng giá trị như gói MHASPacketType PACTYP_MPEGH3DACFG, mà dữ liệu tương tác (người dùng) 8 đề cập đến.

Cú pháp

Phần tử cú pháp mpegH3daElementInteraction() (xem tài liệu [2]) từ giao diện để tương tác người dùng được sử dụng lại để mang dữ liệu tương tác (người dùng) 8 được lồng vào trong dòng bit 10.

Tài liệu tham khảo

[1] ISO/IEC JTC1/SC29/WG11 N14459, "Text of ISO/IEC 23008-3/CD, 3D audio"

[2] FhG, "Normative Interface for User Interaction", Input to the Meeting of AHG on 3D Audio, DRC and Audio Maintenance, June 2-3, 2014, Paris, France

Hơn nữa, cú pháp có khả năng được đề xuất cho giao diện tiêu chuẩn cho tính tương tác người dùng.

Được đề xuất để trích giao diện cho việc điều khiển tương tác từ giao diện để điều khiển việc kết xuất. Phần tử cú pháp `mpegh3daAudioRendering()` sau đó xác định các tham số kết xuất và phần tử cú pháp `mpegh3daElementInteraction()` được xác định mới chứa thông tin mà được cần để tương tác (người dùng).

Định nghĩa phần tử cú pháp `mpegh3daElementInteraction()`

Phần tử cú pháp `mpegh3daElementInteraction()` cung cấp giao diện cho các tương tác (người dùng) có khả năng bất kỳ. Hai chế độ tương tác được định nghĩa.

Chế độ thứ nhất là chế độ tương tác tiên tiến, trong đó sự tương tác có thể được báo hiệu cho từng nhóm phần tử mà có mặt trong cảnh âm thanh. Chế độ này cho phép người dùng tự do chọn (trong các giới hạn của các định nghĩa nhóm chuyển đổi) mà nhóm để phát lại và để tương tác với tất cả chúng (trong các khoảng và các giới hạn đã cho).

Chế độ thứ hai là chế độ tương tác cơ bản, trong đó người dùng có thể chọn một trong số các `GroupPreset` được xác định (từ phần tử cú pháp `mae_AudioSceneInfo()`, xem 14.2 của [1]) như sự thiết lập trước. Trạng thái bật-tắt của các nhóm mà được tham chiếu trong các điều kiện của các `GroupPreset` được chọn sau đó được xác định và có thể không được thay đổi bởi người dùng. Người dùng có thể chỉ thay đổi trạng thái bật-tắt của các nhóm khác và vị trí và độ khuếch đại của tất cả các nhóm theo các thừa nhận hoặc các khoảng được xác định trước.

Bốn cải biến phần tử khác nhau được phân biệt.

- Tính tương tác bật/tắt: nhóm gồm các phần tử được chuyển đổi bật hoặc tắt (lưu ý về biên tập: trước đây được gọi là “thay đổi đối tượng” [4])
- Tính tương tác vị trí: các vị trí của nhóm các phần tử được thay đổi (góc phương vị, góc nâng và khoảng cách, lưu ý về biên tập: trước đây được gọi là “thay đổi vị trí” [4])
- Tính tương tác độ khuếch đại: mức độ/độ khuếch đại của nhóm các phần tử được thay đổi (lưu ý về biên tập: trước đây được gọi là “thay đổi độ khuếch đại” [4])
- Tính tương tác hữu tuyến: đầu ra hữu tuyến là đầu ra tùy chọn ngoài loa phóng thanh được kết nối. Nội dung âm thanh của các phần tử của nhóm được định

tuyến cho đầu ra hữu tuyến, ví dụ, nội dung đối với sự khiếm thính hoặc rãnh ngôn ngữ bổ sung.

Tất cả các cải biến được xác định trên nhóm của mức phân tử, vì các nhóm tập hợp các phân tử liên quan mà chỉ cần được thao tác cùng nhau.

Chữ ký được đưa vào để báo hiệu, ví dụ, trong thiết bị mà tính tương tác đã xảy ra.

Bảng 3a Cú pháp mpeg3daElementInteraction()

Cú pháp	Số lượng các bit	Gợi ý
mpeg3daElementInteraction() { ei_InteractionSignatureDataLength; if (ei_InteractionSignatureDataLength > 0) { ei_InteractionSignatureDataType; for (c = 0; c < bsInteractionSignatureDataLength + 1; c++) { ei_InteractionSignatureData[c]; } } ElementInteractionData(); }	8 8 8	uimbsf uimbsf uimbsf

ei_InteractionSignatureDataLength Trường này xác định độ dài của chữ ký tương tác sau đây theo byte.

ei_InteractionSignatureDataType Trường này xác định loại chữ ký. Các giá trị sau đây là có khả năng:

Phương án nữa của mpeg3daElementInteraction() được thể hiện trong bảng 3b:

Bảng 3b Cú pháp mpeg3daElementInteraction()

Cú pháp	Số lượng các bit	Gợi ý
mpeg3daElementInteraction() { ei_InteractionSignatureDataLength; if (ei_InteractionSignatureDataLength > 0) { ei_InteractionSignatureDataType; }	8 8	uimbsf uimbsf

```

    for ( c = 0; c < bsInteractionSignatureDataLength +
1; c++ ) {
        ei_InteractionSignatureData[c];
    }
}
ElementInteractionData();
hasLocalZoomAreaSize;
if (hasLocalZoomAreaSize) {
    LocalZoomAreaSize();
}
}

```

Bảng 4a Giá trị của ei_InteractionSignatureDataType

giá trị	nghĩa
0	Chuỗi chung
1-127	Dành riêng cho việc sử dụng ISO
128-255	Dành riêng cho việc sử dụng bên ngoài phạm vi ISO

ei_InteractionSignatureData Trường này chứa chữ ký xác định bộ khởi đầu của dữ liệu tương tác.

Phương án nữa của ei_InteractionSignatureDataType được thể hiện trong bảng 4b:

Bảng 4b Giá trị của ei_InteractionSignatureDataType

giá trị	nghĩa
0	Chuỗi chung trong UTF-8 theo ISO/IEC 10646
1-127	Dành riêng cho việc sử dụng ISO
128-255	Dành riêng cho việc sử dụng bên ngoài phạm vi ISO

ei_InteractionSignatureData Trường này chứa chữ ký xác định bộ khởi đầu của dữ liệu tương tác.

hasLocalZoomAreaSize Dấu hiệu mà xác định nếu thông tin về kích

thước vùng thu phóng cục bộ là có sẵn. Nếu dấu hiệu này được cho phép, việc ánh xạ lại đối tượng để thu phóng được áp dụng.

Bảng 5 - Cú pháp của ElementInteractionData()

Cú pháp	Số lượng các bit	Gợi ý
ElementInteractionData(); { ei_interactionMode; ei_numGroups; /* Channel, Object, HOA, SAOC */ if (ei_interactionMode == 0) { /* direct element group interaction = advanced */ ei_GroupInteractivityStatus(ei_numGroups); } else { /*GroupPreset interaction= basic */ ei_groupPresetID; ei_GroupInteractivityStatus(ei_numGroups); } }	1 7 5	bslbf uimsbf uimsbf

ei_interactionMode

Dấu hiệu mà xác định nếu loại tương tác tiên tiến hoặc chế độ tương tác cơ bản được chọn.

ei_numGroups

Trường này chứa số lượng các nhóm trong cảnh âm thanh.

ei_groupPresetID

Trường này chứa groupPresetID mà được xác định trong cảnh âm thanh. ID này phản xạ sự lựa chọn thiết lập trước của người dùng.

Bảng 6a Cú pháp của ei_GroupInteractivityStatus()

Cú pháp	Số lượng các bit	Gợi ý
ei_GroupInteractivityStatus (numGroups) { for (grp = 0; grp < numGroups; grp++) { ei_groupID[grp]; ei_onOff[grp]; ei_routeToWIRE[grp]; if (ei_routeToWIRE[grp] == 1) { routeToWireID[grp]; } } }	 7 1 1 4	 uimsbf bslbf bslbf uimsbf

if (ei_onOff [grp] == 1) {		
ei_changePosition[grp];/* position change */	1	bslbf
if (ei_changePosition[grp]) {		
ei_azOffset[grp];	8	uimsbf
ei_elOffset[grp];	6	uimsbf
ei_distFact[grp];	4	uimsbf
}		
ei_changeGain; /* gain change */	1	bslbf
if (ei_changeGain) {		
ei_gain;	7	uimsbf
}		
}		
}		

ei_groupID	GroupID đối với nhóm hiện thời mà sự tương tác được mô tả cho.
ei_routeToWIRE	Trường này được xác định nếu nội dung âm thanh của nhóm nên được định tuyến cho đầu ra hữu tuyến.
ei_routeToWireID	ID của đầu ra hữu tuyến mà nhóm nên được định tuyến cho.
ei_onOff	Xác định trạng thái bật-tắt của nhóm hiện thời. Trong trường hợp chế độ tương tác cơ bản (tương tác trên các GroupPreset) được chọn, giá trị này phải giống với trạng thái bật-tắt được xác định của nhóm với ei_groupID nếu nhóm này là phần của các điều kiện của GroupPreset được chọn với ei_groupPrersetId. Đối với chế độ tương tác cơ bản, không được cho phép để báo hiệu trạng thái bật-tắt khác nhau ở đây. Trạng thái bật-tắt của tất cả các nhóm mà không phải là phần của các điều kiện của GroupPreset được chọn, trạng thái bật-tắt có thể được báo hiệu một cách tùy ý.
ei_changePosition	Dấu hiệu này xác định nếu vị trí của các phân tử

	nhóm đã được thay đổi.
ei_azOffset	Sự thay đổi của góc phương vị được cho như độ lệch. Trường này có thể có các giá trị nằm trong khoảng $AzOffset = -180^\circ$ và $AzOffset = 180^\circ$: $AzOffset = 1.5 \cdot (ei_azOffset - 128)$ $AzOffset = \min(\max(AzOffset, -180), 180);$
ei_elOffset	Sự thay đổi của góc phương vị được cho như độ lệch. Trường này có thể có các giá trị nằm trong khoảng $ElOffset = -90^\circ$ và $ElOffset = 90^\circ$: $ElOffset = 3 \cdot (ei_elOffset - 32)$ $ElOffset = \min(\max(ElOffset, -90), 90);$
ei_distFact	Tính tương tác khoảng cách được cho như thừa số nhân. Trường có thể có các giá trị nằm trong khoảng 0 đến 15 dẫn đến DistFactor nằm trong khoảng 0,00025 và 8: $DistFactor = 2^{((ei_distFactor - 8) - 4)}$ $DistFactor = \min(\max(DistFactor, 0,00025), 8);$
ei_changeGain	Dấu hiệu này xác định nếu độ khuếch đại/mức độ của các phần tử nhóm đã được thay đổi.
ei_gain	Trường này xác định độ khuếch đại bổ sung của các phần tử của nhóm hiện thời. Trường có thể có các giá trị nằm trong khoảng 0 và 127 biểu diễn các giá trị độ khuếch đại nằm trong khoảng Độ khuếch đại = -63dB và Độ khuếch đại = 31dB trong các bước 1dB, với $Độ\ khuếch\ đại\ [dB] = ei_gain - 64$ $Độ\ khuếch\ đại\ [dB] = \min(\max(Gain, -63), 31);$ Nếu ei_gain được thiết lập đến 0, độ khuếch đại sẽ được thiết lập đến âm vô cực dB.

Phương án nữa của ei_GroupInteractivityStatus() được thể hiện trong Bảng 6b:

Bảng 6b Cú pháp của ei_GroupInteractivityStatus()

Cú pháp	Số lượng các bit	Gợi ý
ei_GroupInteractivityStatus (numGroups)		
{		
for (grp = 0; grp < numGroups; grp++) {		
ei_groupID[grp];	7	uimbsf
ei_onOff[grp];	1	bslbf
ei_routeToWIRE[grp];	1	bslbf
if (ei_routeToWIRE[grp] == 1) {		
routeToWireID[grp];	16	uimbsf
}		
if (ei_onOff [grp] == 1) {		
ei_changePosition[grp];/* position change */	1	bslbf
if (ei_changePosition[grp]) {		
ei_azOffset[grp];	8	uimbsf
ei_elOffset[grp];	6	uimbsf
ei_distFact[grp];	4	uimbsf
}		
ei_changeGain; /* gain change */	1	bslbf
if (ei_changeGain) {		
ei_gain;	7	uimbsf
}		
}		
}		
}		

Dữ liệu tương tác được xác định theo các trường phần tử âm thanh siêu dữ liệu (xem 14.2 trong [1]). Trong việc xác định phần tử âm thanh siêu dữ liệu, khoảng của tính tương tác đối với tính tương tác độ khuếch đại được cho theo dB, các khoảng của tính tương tác đối với góc phương vị và góc nâng được cho như các giá trị độ lệch cực tiểu và cực đại (với cùng độ phân giải như ở đây: 1,5° đối với góc phương vị và 3° đối với góc nâng) và các khoảng đối với khoảng cách được cho như các thừa số nhân. Do đó, giao diện được xác định theo cách tương tự.

Đối với sự tương tác người dùng, tất cả các thay đổi có khả năng (bật/tắt, độ khuếch đại, vị trí, dây) đối với tất cả các nhóm phần tử được báo hiệu trong một phần tử cú pháp ElementInteraction() thay vì sử dụng một phép lặp của các phần tử cú pháp

cho từng loại tương tác và từng nhóm hoặc phân tử.

Đầu ra hữu tuyến cũng xác định cho các nhóm mà bị tắt, vì chỉ sau đó, có khả năng định tuyến các rãnh ngôn ngữ khác cho đầu ra hữu tuyến. Logic nhóm chuyển đổi bị vi phạm có khả năng phải được xử lý trong bộ giải mã.

Loại nội dung (kênh, đối tượng, HOA) mà đã được đề xuất trước đây trong [4] được di chuyển đến đây. Đã hoàn toàn được biết đến bởi groupId, vì từng nhóm các phân tử có signalGroupType (các kênh, các đối tượng, SAOC, HOA) mà xác định loại nội dung tín hiệu và được báo hiệu trong phân tử cú pháp Signals3d()).

Sự tương tác cơ bản thông qua các GroupPreset

Nếu chế độ tương tác cơ bản được thiết lập, người dùng có thể chọn một trong số các GroupPreset được xác định (từ phân tử cú pháp mae_AudioSceneInfo() groupCollection, được đề xuất như công nghệ ứng viên trong w14464 [7]).

Các điều kiện từ thiết lập trước được chọn xác định trạng thái bật-tắt của các nhóm được tham chiếu. Người dùng không cần có khả năng cải biến trạng thái bật-tắt cho các nhóm này.

Trường tương ứng ei_onOff trong phân tử cú pháp ei_GroupInteractivityStatus [2] phải phản xạ điều kiện trong groupPreset được chọn cho tất cả các nhóm mà được tham khảo trong các điều kiện của groupPreset được chọn.

Trạng thái bật-tắt của tất cả các nhóm khác có thể được chọn bởi người dùng.

Kết luận

Đóng góp này đề xuất sự xác định cho giao diện đối với sự tương tác người dùng. Định nghĩa giao diện được cung cấp mà cho phép đối với hai chế độ tương tác khác nhau (cơ bản và tiên tiến). Được đề xuất để thừa nhận giao diện được đề xuất vào CD.

Tài liệu tham khảo

[1] ISO/IEC JTC1/SC29/WG11 N14459, "Text of ISO/IEC 23008-3/CD, 3D audio"

[2] FhG, "Rendering Interfaces to MPEG-H and Unification of Loudspeaker Signaling", Input to the Meeting of AHG on 3D Audio, DRC and Audio Maintenance, June 2-3, 2014, Paris, France

[3] ISO/IEC JTC1/SC29/WG11 N14463, "Normative Interface for Binaural

Data”, April 2014, Valencia, Spain

[4] ISO/IEC JTC1/SC29/WG11 M33134, “3D Audio Decoder Interfaces”, April 2014, Valencia, Spain (Input from Philips)

[5] ISO/IEC JTC1/SC29/WG11 M31427, “Thoughts on binaural parameterization of MPEG codecs”, October 2013, Geneva, Switzerland (Input from Orange)

[6] ISO/IEC JTC1/SC29/WG11 M30249, “BRIR interface format: update and implementation”, July 2013, Vienna, Austria (Input from Philips)

[7] ISO/IEC JTC1/SC29/WG11 N14464 “Candidate Technologies for 3D Audio”

Hơn nữa, đoạn 13 của tài liệu ISO/IEC 23008-3 CD [1] định nghĩa cú pháp MHAS được sử dụng để tóm lược các tài hữu ích của âm thanh MPEG-H 3D trong các dòng (thời gian thực). Với tài liệu này, các khả năng bổ sung để cho phép việc sử dụng đáng tin cậy của MHAS trong các hệ thống vận chuyển khác bên cạnh dòng vận chuyển MPEG-2 được đề xuất.

Dữ liệu điền đầy

Các kênh truyền nhất định có thể chỉ được vận hành tại tốc độ bit tức thời không đổi. Để hoàn thành việc này, loại gói dữ liệu điền đầy được đề xuất để có thể điền đầy dòng MHAS lên đến tốc độ bit đã cho nhất định.

CRC

Lớp của dòng vận chuyển MPEG-2 [2] lo việc bảo vệ dòng cơ sở MHAS được đóng gói. Do đó việc mất dữ liệu hoặc sai lệch dữ liệu có thể được nhận biết.

Mặt khác, các giao diện chuỗi phổ biến (ví dụ AES/EBU, S/PDIF, [3], [4], [5]) cung cấp sự bảo vệ lỗi không đủ. Loại gói CRC tùy chọn được đề xuất để cho phép phát hiện lỗi nếu MHAS được sử dụng trên các giao diện như vậy.

Đóng gói bộ mô tả

Các dòng MHAS được sử dụng để truyền đạt âm thanh được mã hóa đến hoặc từ các bộ mã hóa/bộ giải mã dòng vận chuyển MPEG-2. Được đề xuất để truyền đạt thông tin bộ mô tả liên quan trong loại gói MHAS bổ sung.

Bảng 7 — MHASPacketPayload()

Cú pháp	Số lượng các bit	Gợi ý
MHASPacketPayload(MHASPacketType)		
{		
switch (MHASPacketType) {		
case PACTYP_SYNC:		
0xA5; /* syncword*/	8	uimbsf
break;		
case PACTYP_MPEGH3DACFG:		
mpegh3daConfig();		
break;		
case PACTYP_MPEGH3DAFRAME:		
mpegh3daFrame();		
break;		
case PACTYP_FILLDATA:		
for (i=0; i< MHASPacketLength; i++) {		
mhas_fill_data_byte(i);	8	bslbf
}		
break;		
case PACTYP_SYNCGAP:		
syncSpacingLength = escapedValue(16,24,24);	16,40,64	uimbsf
break;		
case PACTYP_MARKER:		
for (i=0; i< MHASPacketLength; i++) {		
marker_byte(i);	8	
}		
break;		
case PACTYP_CRC16:		
mhasParity16Data;	16	bslbf
break;		
case PACTYP_CRC32:		
mhasParity32Data;	32	bslbf
break;		
case PACTYP_DESCRIPTOR:		
for (i=0; i< MHASPacketLength; i++) {		
mhas_descriptor_data_byte(i);	8	bslbf
}		
break;		
}		
ByteAlign();		
}		

Bảng 8a — Giá trị của MHASPacketType

MHASPacketType	Giá trị
PACTYP_SYNC	0
PACTYP_MPEGH3DACFG	1
PACTYP_MPEGH3DAFRAME	2
/* reserved for ISO use */	3-5
PACTYPE_FILLDATA	6
PACTYP_SYNCGAP	7
PACTYPE_MARKER	8
PACTYPE_CRC16	9
PACTYPE_CRC32	10
PACTYPE_DESCRIPTOR	11
/* reserved for ISO use */	13-127
/* reserved for use outside of ISO scope */	128-261
/* reserved for ISO use */	262-389
/* reserved for use outside of ISO scope */	390-517
Lưu ý: Các giá trị MHASPacketType ứng dụng cụ thể được chỉ thị để nằm trong không gian được dự trữ cho việc sử dụng bên ngoài của phạm vi ISO. Những việc này được bỏ qua bởi bộ giải mã như cực tiểu của cấu trúc được yêu cầu bởi bộ giải mã để bỏ qua các mở rộng này.	

Phương án nữa của MHASPacketType được thể hiện trong Bảng 8b:

Bảng 8b — Giá trị của MHASPacketType

MHASPacketType	Giá trị
PACTYP_SYNC	6
PACTYP_MPEGH3DACFG	1
PACTYP_MPEGH3DAFRAME	2
/* reserved for ISO use */	3-5
PACTYPE_FILLDATA	0
PACTYP_SYNCGAP	7
PACTYPE_MARKER	8
PACTYPE_CRC16	9
PACTYPE_CRC32	10
PACTYPE_DESCRIPTOR	11
/* reserved for ISO use */	13-127
/* reserved for use outside of ISO scope */	128-261
/* reserved for ISO use */	262-389
/* reserved for use outside of ISO scope */	390-517
Lưu ý: Các giá trị MHASPacketType ứng dụng cụ thể được chỉ thị để nằm trong không gian được dự trữ cho việc sử dụng bên ngoài của phạm vi ISO. Những việc này được bỏ qua bởi bộ giải mã như cực tiểu của cấu trúc được yêu cầu bởi bộ giải mã để bỏ qua các mở rộng này.	

mhasParity16Data trường 16-bit mà chứa giá trị CRC mà mang lại đầu ra bằng không của 16 bộ ghi trong bộ giải mã với đa thức:

$$x^{16} + x^{15} + x^5 + 1$$

và trạng thái ban đầu của bộ ghi dịch chuyển của 0xFFFF.

mhasParity32Data trường 32-bit mà chứa giá trị CRC mà mang lại đầu ra bằng không của 32 bộ ghi trong bộ giải mã với đa thức:

$$x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$$

và trạng thái ban đầu của bộ ghi dịch chuyển của

0xFFFFFFFF.

mhas_fill_data_byte các phân tử dữ liệu 8-bit, không có các giới hạn áp dụng
 Truy cập ngẫu nhiên/Bộ đánh dấu phát lại tức thời

Khi marker_byte thứ nhất của tải hữu ích gói là “0x02”, gói sau đây của loại PACTYP_MPEG3DAFRAME với MHASPacketLabel giống nhau được mã hóa theo các nguyên tắc được cho trong đoạn 5.5.5 “Audio Pre-Roll”.

Bộ đánh dấu biên chương trình

Khi marker_byte thứ nhất của tải hữu ích gói là “0x03”, biên chương trình xuất hiện tại thời điểm này và các gói sau thuộc về chương trình mới.

PSCTYP_CRC16 và PACTYP_CRC32

MHASPacketType PACTYP_CRC16 và PACTYP_CRC32 có thể được sử dụng để phát hiện các lỗi trong gói MHAS đứng trước với MHASPacketLabel được thiết lập đến cùng giá trị. Nó sẽ theo sau trực tiếp gói MHAS mà giá trị CRC của nó đề cập đến. Việc này có thể có lợi khi dòng MHAS được truyền đạt trên kênh dễ bị lỗi.

Phương pháp phát hiện lỗi sử dụng một trong số trạng thái đa thức sinh và trạng thái bộ ghi dịch chuyển kết hợp như được xác định lần lượt cho mhasParity16Data hoặc mhasParity32Data.

Các bit được chứa vào trong CRC-check là MHASPacketPayload() đầy đủ cho gói MHAS liên quan được theo sau lần lượt bởi các từ mhasParity16Data hoặc mhasParity32Data.

Trong trường hợp mà không có lỗi, từng đầu ra của bộ ghi dịch chuyển sẽ bằng không. Tại bộ mã hóa CRC, trường mhasParity16Data / mhasParity32Data được mã hóa với giá trị sao cho trường này được bảo đảm.

PACTYP_FILLDATA

MHASPacketType PACTYP_FILLDATA cung cấp khả năng để bổ sung dữ liệu điền đầy để điều chỉnh tốc độ bit tức thời. Điều này có thể là đáng mong muốn trong các ứng dụng thời gian thực nhất định sử dụng kênh truyền tốc độ không đổi.

Vì các gói của loại này không đề cập đến dữ liệu tải hữu ích nhất định, MHASPacketLabel được thiết lập đến 0.

Được mong đợi rằng bộ giải mã bỏ qua dữ liệu được truyền trong các gói typePACTYP_FILLDATA. Hơn nữa, các công cụ trung gian xử lý các dòng MHAS được cho phép để loại bỏ các gói này khỏi dòng.

Nó được cho phép để thiết lập MHASPacketLength đến 0. Việc này mang lại kích thước gói cực tiểu là bằng 2 byte.

PACTYP_DESCRIPTOR

PACTYP_DESCRIPTOR có thể được sử dụng để lồng các bộ mô tả MPEG-2 TS/PS trong các dòng MHAS. Dữ liệu đã truyền đạt như mhas_descriptor_data_byte có cùng cú pháp và các ngữ nghĩa hơn được xác định cho descriptor() trong ISO/IEC 13818-1.

Đối với loại gói này và đối với các bộ mô tả được truyền trong các chu trình bộ mô tả thứ nhất trong TS_program_map_section() (xem ISO/IEC 13818-1), MHASPacketLabel được thiết lập đến 0.

Đối với loại gói này và đối với các bộ mô tả được gán cho một dòng cơ bản (tức là chu trình bộ mô tả thứ hai trong TS_program_map_section()), MHASPacketLabel được thiết lập đến cùng giá trị như PACTYPE_CONFIG từ dòng cơ bản được kết hợp. Tài liệu tham khảo

[1] ISO/IEC JTC1/SC29/WG11 N14459, "Text of ISO/IEC 23008-3/CD, 3D audio"

[2] ISO/IEC 13818-1:2013, Information technology — Generic Coding of moving pictures and associated audio information: Systems

[3] IEC 60958-3: "Digital audio interface – Part 3: Consumer applications"

[4] IEC 61937-11, "Digital audio – Interface for non-linear PCM encoded audio bitstreams applying IEC 60958 – Part 11: MPEG-4 AAC and its extensions in LATM/LOAS"

[5] SMPTE 2041: Format for Non-PCM Audio and Data in AES-3 - MPEG-4 AAC and HE AAC Compressed Digital Audio in ADTS and LATM/LOAS Wrappers

Dù sáng chế đã được mô tả trong ngữ cảnh của các sơ đồ khối mà các khối biểu diễn các thành phần phần cứng thực sự hoặc theo logic, sáng chế cũng có thể được thực hiện bởi phương pháp được thực hiện trên máy tính. Trong trường hợp sau, các khối biểu diễn các bước phương pháp tương ứng mà các bước này đại diện các chức năng được thực hiện bởi các khối phần cứng theo logic hoặc vật lý.

Mặc dù một vài khía cạnh đã được mô tả trong ngữ cảnh của thiết bị, rõ ràng rằng các khía cạnh này cũng biểu diễn sự mô tả của phương pháp tương ứng, mà 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 phạm vi của bước phương pháp cũng biểu diễn sự mô tả của khối hoặc mục hoặc dấu hiệu tương ứng của thiết bị tương ứng. Một số hoặc tất cả các bước phương pháp có thể được thực hiện bởi (hoặc sử dụng) thiết bị phần cứng, như ví dụ, bộ vi xử lý, máy tính có thể lập trình hoặc mạch điện tử. Theo một số phương án, một số một hoặc nhiều bước phương pháp quan trọng nhất có thể được thực hiện bởi thiết bị như vậy.

Tín hiệu được truyền hoặc được mã hoá theo sáng chế có thể được lưu trữ trên vật ghi lưu trữ số hoặc có thể được truyền trên vật ghi truyền như vật ghi truyền không dây hoặc vật ghi truyền có dây như internet.

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. Việc thực hiện có thể được tiến hành sử dụng vật lưu trữ số, ví dụ, đĩa mềm, DVD, Blu-Ray, 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 có thể lập trình sao cho phương pháp tương ứng được thực hiện. Do đó, vật ghi lưu trữ số có thể đọc được bằng máy tính.

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

Thông thường, các phương án theo sáng chế có thể được thực hiện như sản phẩm 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. Ví dụ, mã chương trình có thể được lưu trữ trên vật mang có thể 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 có thể đọc được bằng máy.

Do đó, nói cách khác, 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.

Do đó, phương án nữa của phương pháp theo sáng chế là, vật mang dữ liệu

(hoặc vật ghi lưu trữ không chuyển tiếp như vật ghi lưu trữ số, hoặc vật ghi có thể đọc được bằng máy tính) bao gồm, đã được ghi lại trên đó, chương trình máy tính để thực hiện một trong số các phương pháp được mô tả ở đây. Vật mang dữ liệu, vật ghi lưu trữ số hoặc vật ghi được ghi lại thường là hữu hình và/hoặc không tạm thời.

Do đó, phương án nữa của phương pháp theo sáng chế là, dòng dữ liệu hoặc chuỗi các tín hiệu biểu diễn chương trình máy tính để thực hiện một trong các phương pháp được mô tả ở đây. Ví dụ, dòng dữ liệu hoặc chuỗi các tín hiệu được tạo cấu hình để được truyền tải thông qua sự kết nối truyền thông dữ liệu, ví dụ, thông qua internet.

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

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

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

Trong một số phương án, thiết bị logic có thể lập trình được (ví dụ, mảng công có thể lập trình được dạng trường) 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 công lập trình được dạng trường 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 tốt hơn là được thực hiện bởi thiết bị phần cứng bất kỳ.

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

YÊU CẦU BẢO HỘ

1. Bộ xử lý dữ liệu âm thanh (1) bao gồm:

giao diện bộ nhận (2) để nhận dữ liệu âm thanh được mã hóa (3) và siêu dữ liệu (4) liên quan đến dữ liệu âm thanh được mã hóa (3);

bộ phân tích siêu dữ liệu (5) để phân tích siêu dữ liệu (4) để xác định khả năng thao tác dữ liệu âm thanh;

giao diện tương tác (6) để nhận đầu vào tương tác (7) và để tạo ra, từ đầu vào tương tác (7), dữ liệu điều khiển tương tác (8) liên quan đến khả năng thao tác dữ liệu âm thanh, trong đó giao diện tương tác (6) được tạo cấu hình để biểu diễn, tới người dùng, khả năng thao tác dữ liệu âm thanh được suy ra từ siêu dữ liệu (4) bởi bộ phân tích siêu dữ liệu (5), và để nhận, từ người dùng, đầu vào tương tác (7) trên thao tác dữ liệu cụ thể của khả năng thao tác dữ liệu; và

bộ tạo dòng dữ liệu (9) để thu được dữ liệu điều khiển tương tác (8) và dữ liệu âm thanh được mã hóa (3) và siêu dữ liệu (4) và để tạo ra dòng dữ liệu đầu ra được mã hóa (10), dòng dữ liệu đầu ra được mã hóa (10) bao gồm dữ liệu âm thanh được mã hóa (3), siêu dữ liệu (4), và dữ liệu điều khiển tương tác (8),

trong đó bộ tạo dòng dữ liệu (9) được tạo cấu hình để xử lý dòng dữ liệu đầu vào (15) bao gồm dữ liệu âm thanh được mã hóa (3) và siêu dữ liệu (4) được nhận bởi giao diện bộ nhận (2) mà không giải mã dữ liệu âm thanh được mã hóa (3) để tạo ra dòng dữ liệu đầu ra được mã hóa (10), hoặc để sao chép dữ liệu âm thanh được mã hóa (3) và ít nhất một phần siêu dữ liệu (4) mà không làm thay đổi trong dòng dữ liệu đầu ra được mã hóa (10),

để lồng dữ liệu điều khiển tương tác (8) như phần dữ liệu bổ sung trong dòng dữ liệu đầu ra (10) và để chuyển tiếp dòng dữ liệu đầu ra được mã hóa (10) bao gồm dữ liệu âm thanh được mã hóa (3), siêu dữ liệu (4) và dữ liệu điều khiển tương tác (8) đến bộ giải mã ngoài (31) để giải mã dòng dữ liệu đầu ra được mã hóa (10).

2. Bộ xử lý dữ liệu âm thanh theo điểm 1, trong đó dữ liệu âm thanh được mã hóa (3) bao gồm các đối tượng âm thanh được mã hóa riêng biệt (11; 12), trong đó ít nhất một phần (13; 14) của siêu dữ liệu (4) liên quan đến đối tượng âm thanh (11; 12) tương ứng,

trong đó bộ phân tích siêu dữ liệu (5) được tạo cấu hình để phân tích phần (13; 14) tương ứng cho các đối tượng âm thanh được mã hóa (11; 12) để xác định, cho ít nhất một đối tượng âm thanh (11), khả năng thao tác đối tượng,

trong đó giao diện tương tác (6) được tạo cấu hình để tạo ra, đối với ít nhất một đối tượng âm thanh được mã hóa (11), dữ liệu điều khiển tương tác (8) từ đầu vào tương tác (7) liên quan đến ít nhất một đối tượng âm thanh được mã hóa (11).

3. Bộ xử lý dữ liệu âm thanh theo một trong số các điểm nêu trên, trong đó giao diện tương tác (6) được tạo cấu hình để tạo ra dữ liệu điều khiển tương tác (8) sao cho dữ liệu điều khiển tương tác (8) được biểu diễn bởi lượng thao tác so với siêu dữ liệu (4).

4. Bộ xử lý dữ liệu âm thanh theo một trong số các điểm nêu trên, trong đó, bộ tạo dòng dữ liệu (9) được tạo cấu hình để tạo ra, trong dòng dữ liệu đầu ra (10), dữ liệu điều khiển tương tác (8) theo cùng định dạng như siêu dữ liệu (4).

5. Bộ xử lý dữ liệu âm thanh theo một trong số các điểm nêu trên trong đó bộ tạo dòng dữ liệu (9) được tạo cấu hình để kết hợp, với dữ liệu điều khiển tương tác (8), bộ nhận biết trong dòng dữ liệu đầu ra (10), bộ nhận biết khác với bộ nhận biết được kết hợp với siêu dữ liệu (4).

6. Bộ xử lý dữ liệu âm thanh theo một trong số các điểm nêu trên, trong đó bộ tạo dòng dữ liệu (9) được tạo cấu hình để bổ sung, vào dữ liệu điều khiển tương tác (8), dữ liệu chữ ký, dữ liệu chữ kí biểu thị thông tin về ứng dụng, thiết bị hoặc người dùng thực hiện thao tác dữ liệu âm thanh hoặc cung cấp đầu vào tương tác.

7. Bộ xử lý dữ liệu âm thanh theo một trong số các điểm nêu trên, trong đó bộ phân tích siêu dữ liệu (5) được tạo cấu hình để nhận biết khả năng vô hiệu hóa đối với một hoặc nhiều đối tượng âm thanh (11, 12) được biểu diễn bởi dữ liệu âm thanh được mã hóa (3),

trong đó giao diện tương tác (6) được tạo cấu hình để nhận thông tin vô hiệu hóa đối với một hoặc nhiều đối tượng âm thanh (11, 12), và

trong đó bộ tạo dòng dữ liệu (9) được tạo cấu hình để đánh dấu một hoặc nhiều đối tượng âm thanh (11, 12) như được vô hiệu hóa trong dữ liệu điều khiển tương tác (8).

8. Bộ xử lý dữ liệu âm thanh theo một trong số các điểm nêu trên, trong đó bộ tạo dòng dữ liệu (9) được tạo cấu hình để tạo ra theo cách động học dòng dữ liệu đầu ra (10), trong đó đáp ứng với đầu vào tương tác mới (7), dữ liệu điều khiển tương tác (8) được cập nhật để khớp với đầu vào tương tác mới (7), và trong đó bộ tạo dòng dữ liệu (9) được tạo cấu hình để bao gồm dữ liệu điều khiển tương tác được cập nhật (8) trong dòng dữ liệu đầu ra (10).

9. Bộ xử lý dữ liệu âm thanh theo một trong số các điểm nêu trên, trong đó giao diện bộ nhận (2) được tạo cấu hình để nhận dòng dữ liệu âm thanh chính (15) bao gồm dữ liệu âm thanh được mã hóa (3) và siêu dữ liệu (4) liên quan đến dữ liệu âm thanh được mã hóa (3), và để nhận thêm dữ liệu âm thanh tùy chọn (16) bao gồm đối tượng âm thanh tùy chọn (17),

trong đó siêu dữ liệu (4) liên quan đến đối tượng âm thanh tùy chọn (17) đã nêu được chứa trong dòng dữ liệu âm thanh chính (15) đã nêu.

10. Bộ xử lý dữ liệu âm thanh theo một trong số các điểm nêu trên,

trong đó bộ phân tích siêu dữ liệu (5) được tạo cấu hình để xác định khả năng thao tác âm thanh đối với đối tượng âm thanh tùy chọn (17) không được chứa trong dữ liệu âm thanh được mã hóa (3),

trong đó giao diện tương tác (6) được tạo cấu hình để nhận đầu vào tương tác (7) cho đối tượng âm thanh tùy chọn (17), và

trong đó giao diện bộ nhận (2) được tạo cấu hình để yêu cầu dữ liệu âm thanh (16) cho đối tượng âm thanh tùy chọn (17) từ bộ cung cấp dữ liệu âm thanh (35) hoặc để nhận dữ liệu âm thanh (16) cho đối tượng âm thanh tùy chọn (17) từ dòng con khác được chứa trong dòng phát rộng (36) hoặc kết nối giao thức internet (37).

11. Bộ xử lý dữ liệu âm thanh theo một trong số các điểm nêu trên,

trong đó bộ tạo dòng dữ liệu (9) được tạo cấu hình để gán, trong dòng dữ liệu đầu ra (10), loại gói nữa cho dữ liệu điều khiển tương tác (8), loại gói nữa khác với các loại gói cho dữ liệu âm thanh được mã hóa (3) và siêu dữ liệu (4), hoặc

trong đó bộ tạo dòng dữ liệu (9) được tạo cấu hình để bổ sung, vào trong dòng dữ liệu đầu ra (10), dữ liệu điền đầy (18) trong loại gói dữ liệu điền đầy, trong đó

lượng dữ liệu điền đầy (18) được xác định dựa trên yêu cầu tốc độ dữ liệu được xác định bởi giao diện đầu ra của bộ xử lý dữ liệu âm thanh.

12. Bộ xử lý dữ liệu âm thanh theo một trong số các điểm nêu trên, trong đó bộ xử lý dữ liệu âm thanh được thực hiện như thiết bị thứ nhất (19) tách rời, mà tách rời khỏi thiết bị thứ hai (28) mà được tạo cấu hình để nhận dữ liệu âm thanh được xử lý, nhưng vẫn được mã hóa, từ thiết bị thứ nhất (19) để giải mã dữ liệu âm thanh đã nêu, trong đó giao diện bộ nhận (2) tạo thành đầu vào (20) tới thiết bị thứ nhất (19) tách rời thông qua kết nối có dây (21) hoặc không dây (22), trong đó bộ xử lý dữ liệu âm thanh (1) còn bao gồm giao diện đầu ra (23) được kết nối với bộ tạo dòng dữ liệu (9), giao diện đầu ra (23) được tạo cấu hình để xuất ra dòng dữ liệu đầu ra (10), trong đó giao diện đầu ra (23) thực hiện việc xuất ra của thiết bị thứ nhất (19) tách rời và bao gồm giao diện không dây (24) hoặc bộ kết nối có dây (25).

13. Phương pháp xử lý dữ liệu âm thanh, phương pháp bao gồm:

nhận dữ liệu âm thanh được mã hóa (3) và siêu dữ liệu (4) liên quan đến dữ liệu âm thanh được mã hóa (3);

phân tích siêu dữ liệu (4) để xác định khả năng thao tác dữ liệu âm thanh;

biểu diễn, tới người dùng, khả năng thao tác dữ liệu âm thanh được suy ra từ siêu dữ liệu (4) được phân tích, và nhận, từ người dùng, đầu vào tương tác (7) trên thao tác dữ liệu cụ thể của khả năng thao tác dữ liệu; và

tạo ra, từ đầu vào tương tác (7), dữ liệu điều khiển tương tác (8) liên quan đến khả năng thao tác dữ liệu âm thanh; và

thu được dữ liệu điều khiển tương tác (8) và dữ liệu âm thanh được mã hóa (3) và siêu dữ liệu (4) và tạo ra dòng dữ liệu đầu ra được mã hóa (10), dòng dữ liệu đầu ra được mã hóa (10) bao gồm dữ liệu âm thanh được mã hóa (3), siêu dữ liệu (4), và dữ liệu điều khiển tương tác (8),

xử lý dòng dữ liệu đầu vào (15) bao gồm dữ liệu âm thanh được mã hóa (3) và siêu dữ liệu (4) mà không giải mã dữ liệu âm thanh được mã hóa (3) để tạo ra dòng dữ liệu đầu ra được mã hóa (10), hoặc sao chép dữ liệu âm thanh được mã hóa (3) và ít nhất một phần siêu dữ liệu (4) mà không làm thay đổi trong dòng dữ liệu đầu ra được mã hóa (10),

lồng dữ liệu điều khiển tương tác (8) như phần dữ liệu bổ sung trong dòng dữ liệu đầu ra được mã hóa (10), và chuyển tiếp dòng dữ liệu đầu ra được mã hóa (10) bao gồm dữ liệu âm thanh được mã hóa (3), siêu dữ liệu (4) và dữ liệu điều khiển tương tác (8) đến bộ giải mã ngoài (31) để giải mã dòng dữ liệu đầu ra được mã hóa (10).

14. Vật ghi có thể đọc được bằng máy tính bao gồm chương trình máy tính để thực hiện phương pháp xử lý dữ liệu âm thanh theo điểm 13 khi chạy trên máy tính hoặc bộ xử lý.

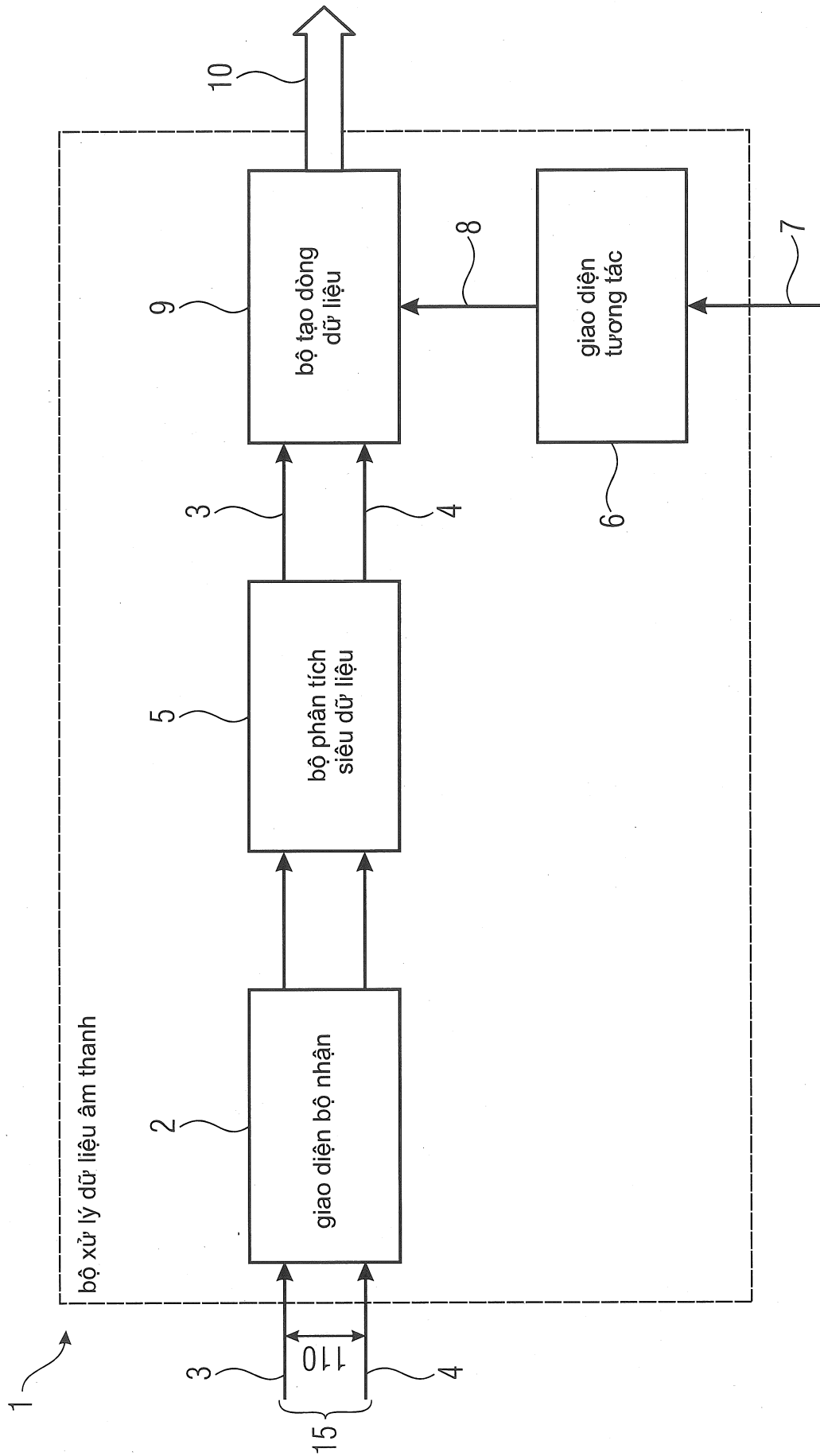


FIG 1

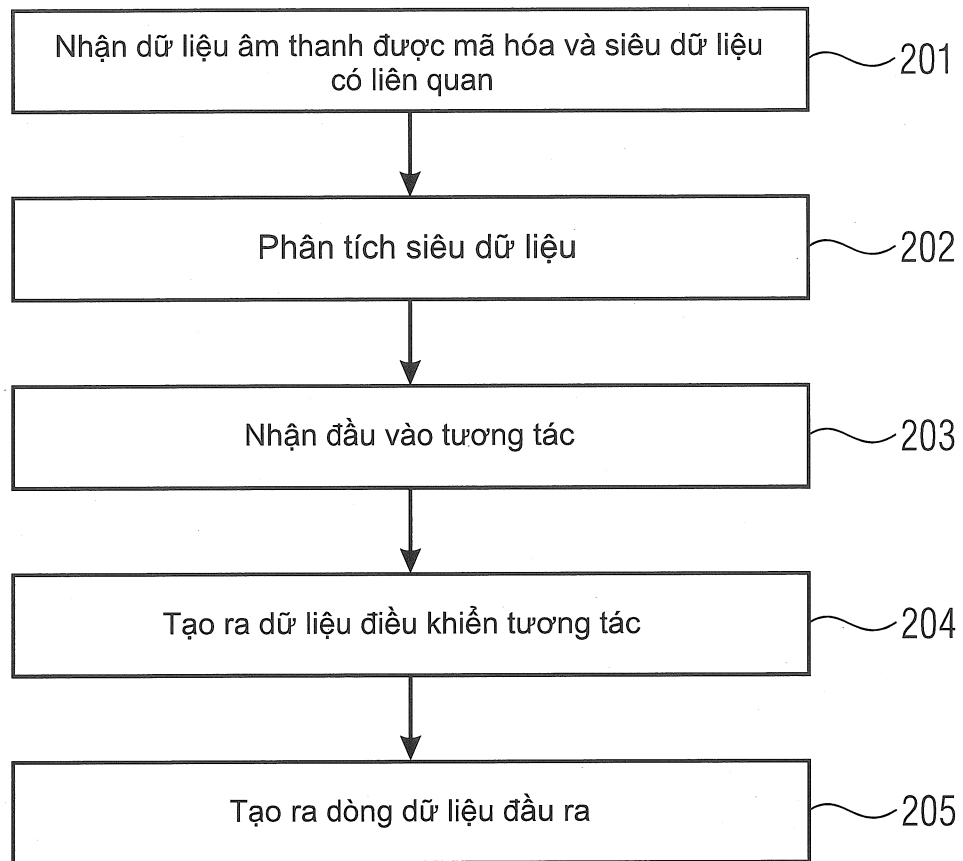


FIG 2

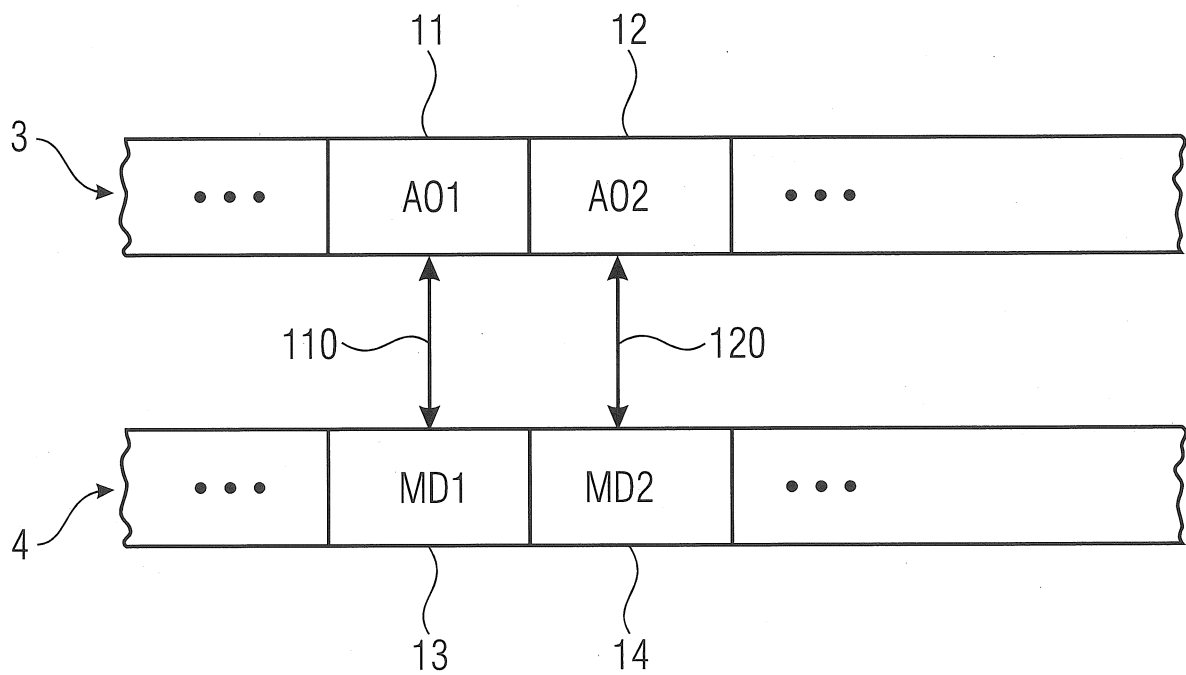


FIG 3

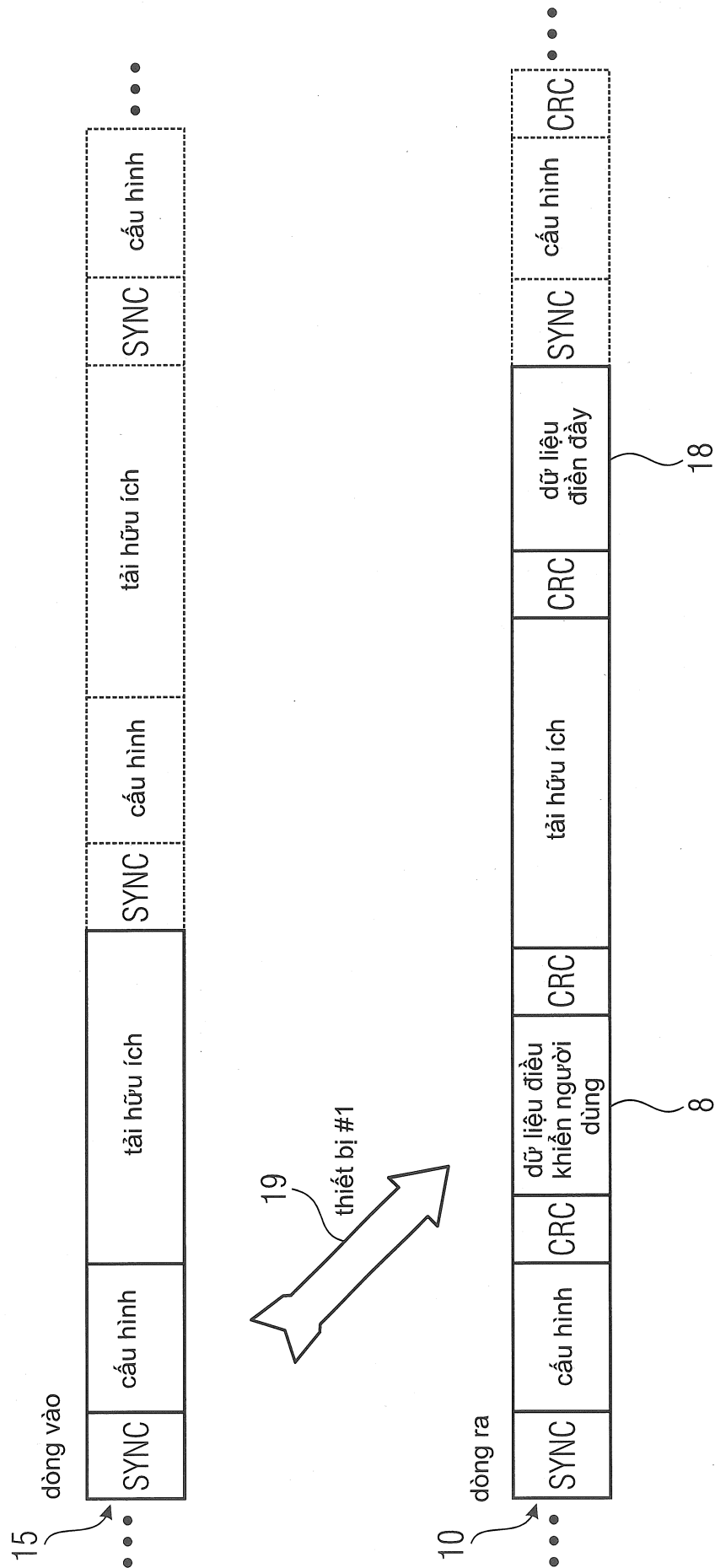


FIG 4

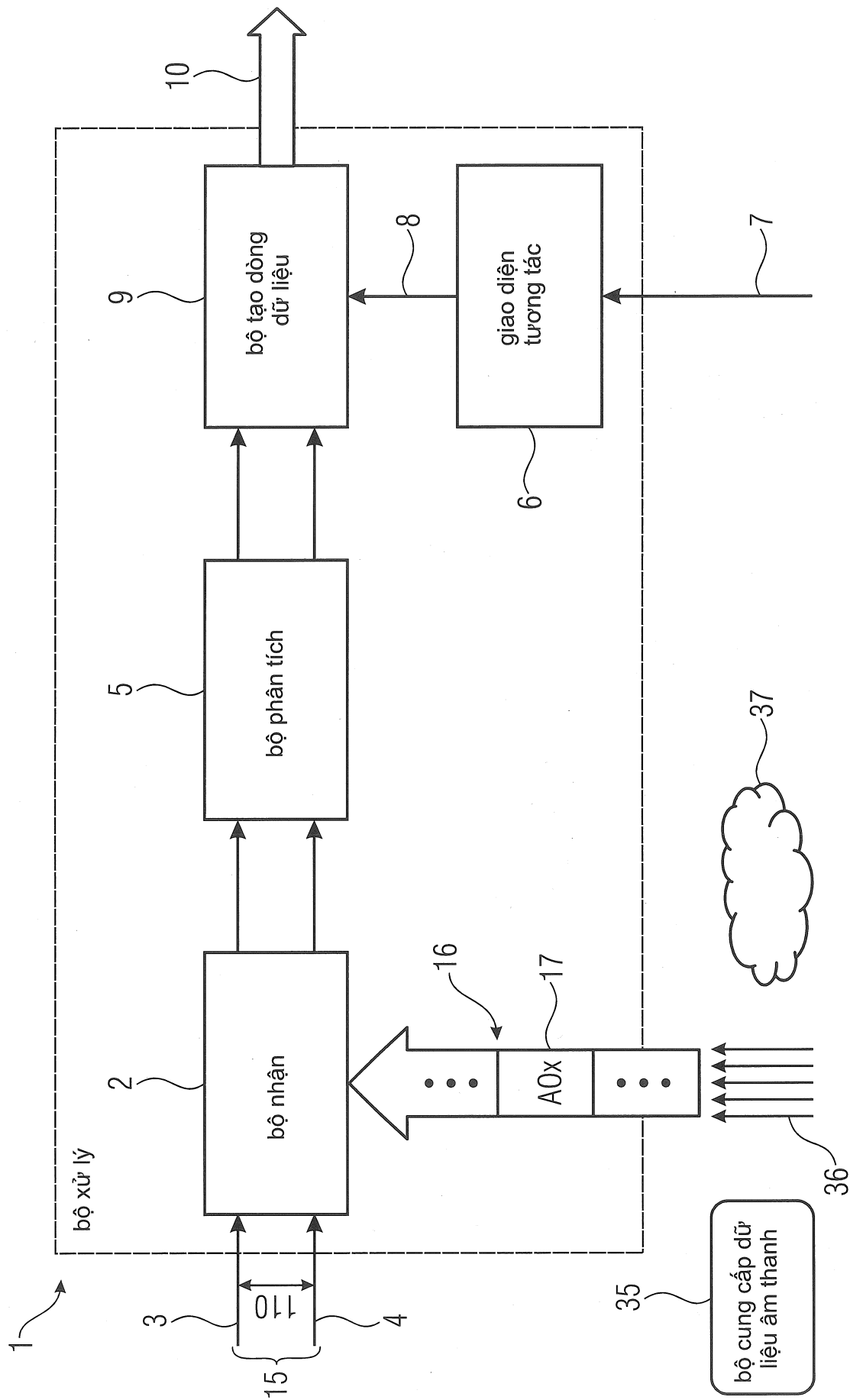


FIG 6

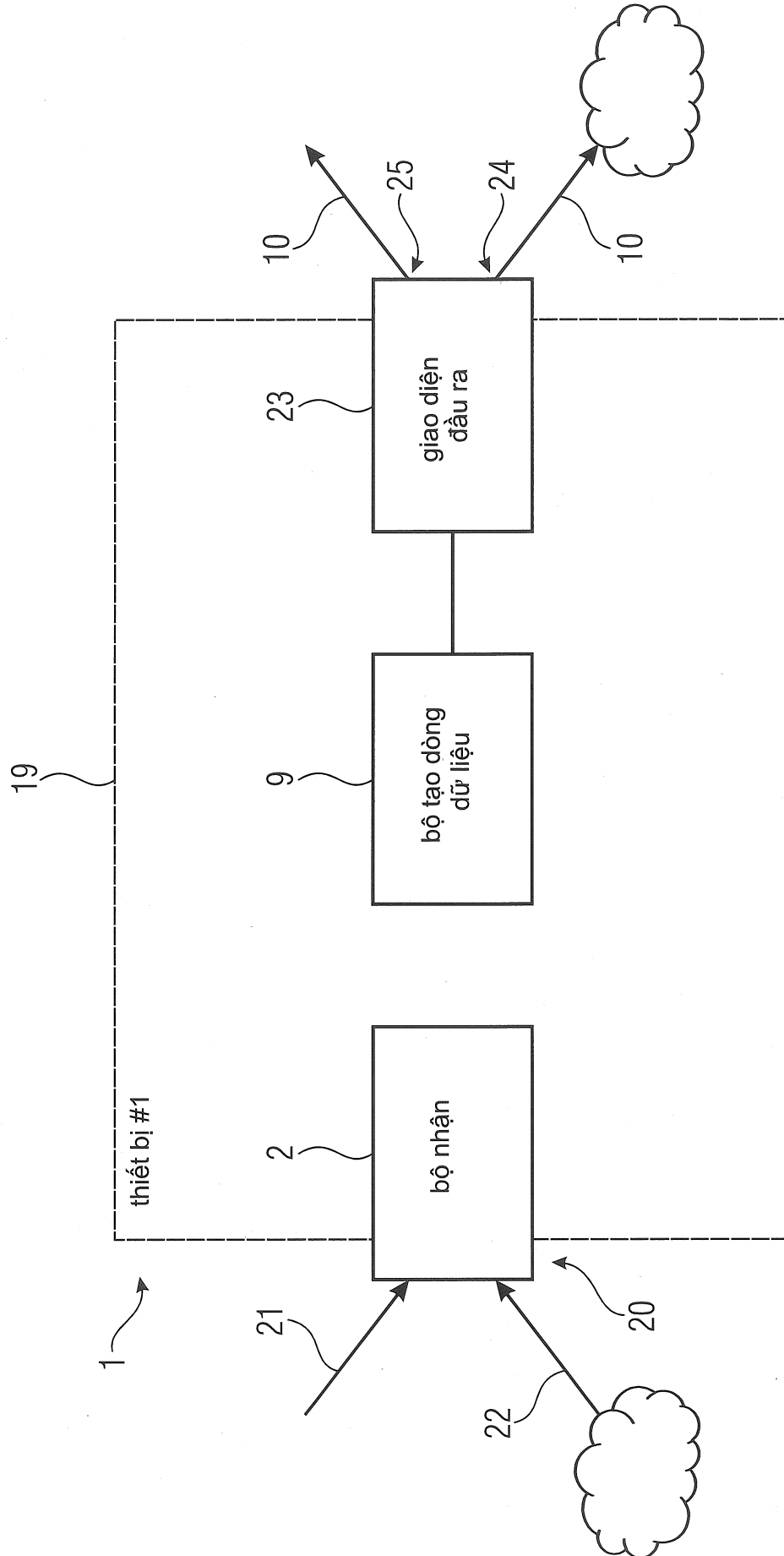


FIG 7

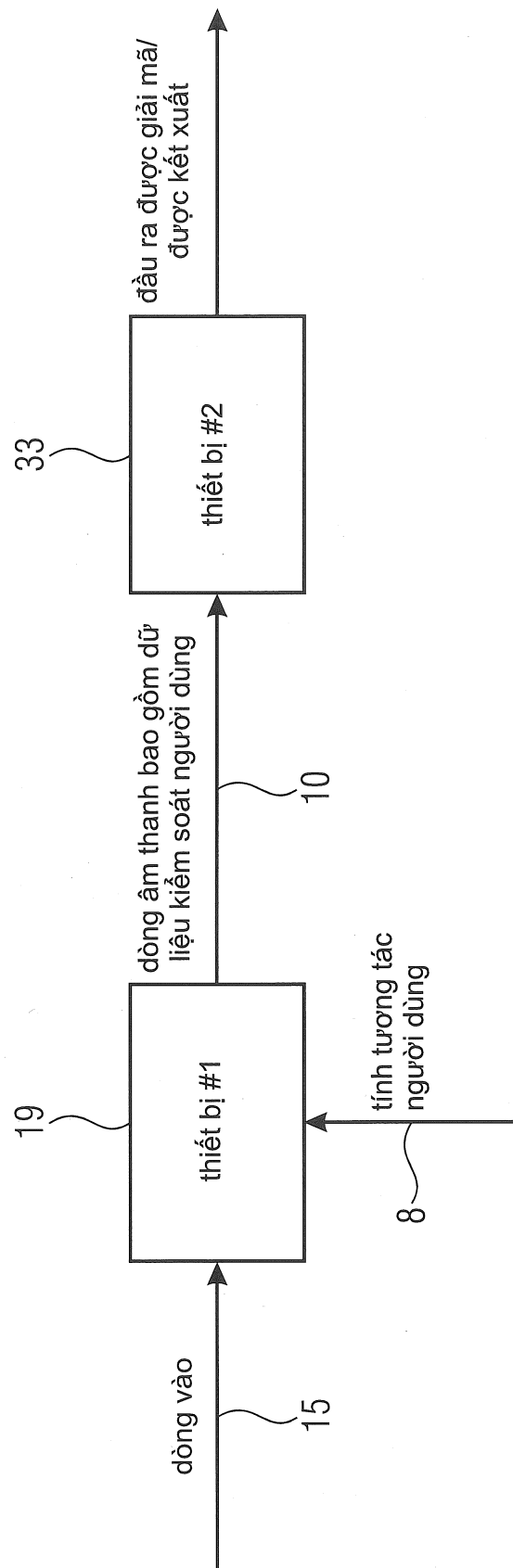


FIG 8

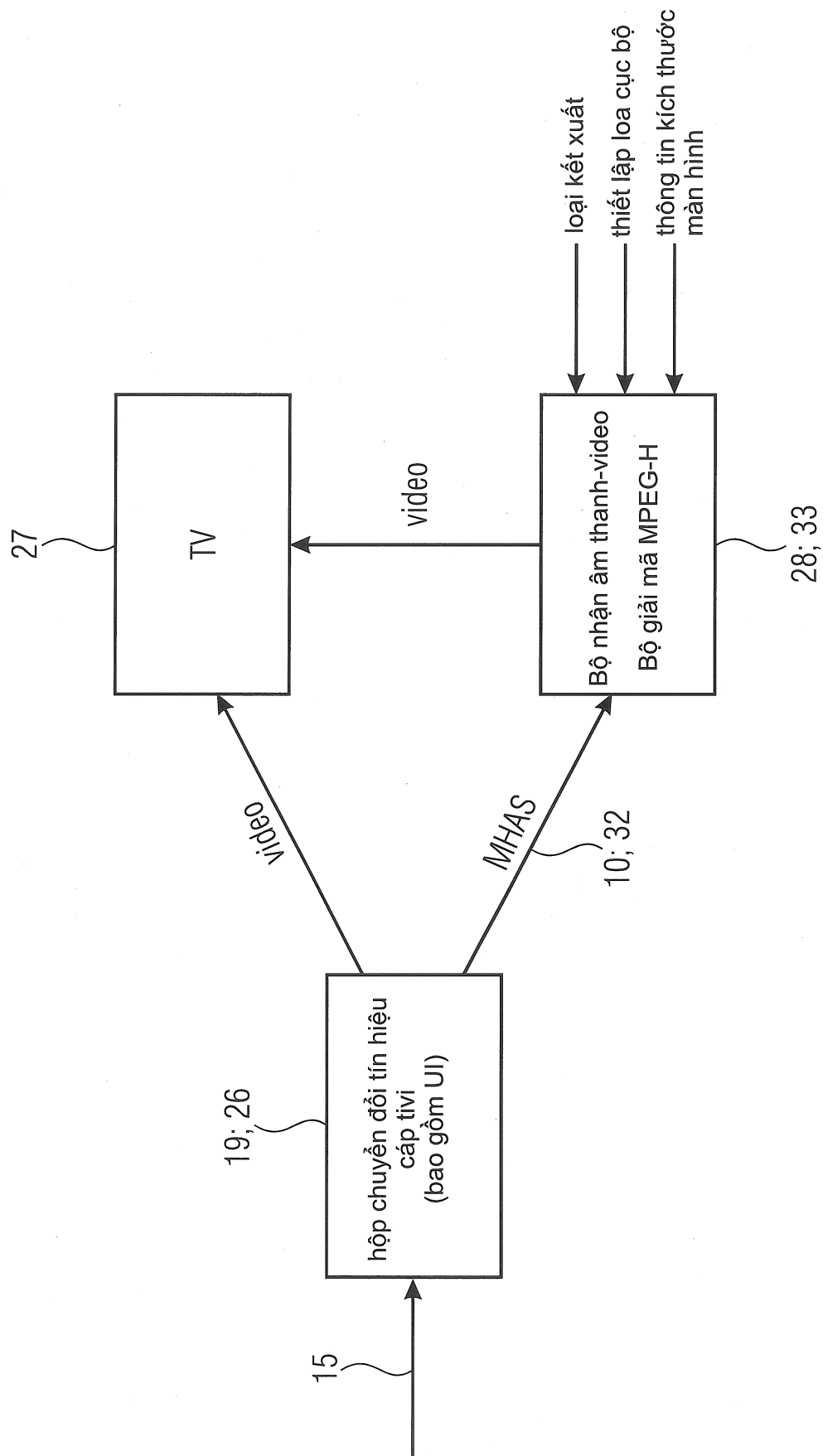


FIG 9